The structure of an energy-coupling protein from bacteria, $IIB^{cellobiose}$, reveals similarity to eukaryotic protein tyrosine phosphatases

Rob LM van Montfort[1], Tjaard Pijning[1], Kor H Kalk[1], Jonathan Reizer[2], Milton H Saier Jr[2], Marjolein MGM Thunnissen[3], George T Robillard[1] and Bauke W Dijkstra[1]*

Background: The bacterial phosphoenolpyruvate-dependent phosphotransferase system (PTS) mediates the energy-driven uptake of carbohydrates and their concomitant phosphorylation. In addition, the PTS is intimately involved in the regulation of a variety of metabolic and transcriptional processes in the bacterium. The multiprotein PTS consists of a membrane channel and at least four cytoplasmic proteins or protein domains that sequentially transfer a phosphoryl group from phosphoenolpyruvate to the transported carbohydrate. Determination of the three-dimensional structure of the IIB enzymes within the multiprotein complex would provide insights into the mechanisms by which they promote efficient transport by the membrane channel IIC protein and phosphorylate the transported carbohydrate on the inside of the cell.

Results: The crystal structure of the IIB enzyme specific for cellobiose, $IIB^{cellobiose}$ (molecular weight 11.4 kDa), has been determined to a resolution of 1.8 Å and refined to an R factor of 18.7% (R_{free} of 24.1%). The enzyme consists of a single four-stranded parallel β sheet flanked by helices on both sides. The phosphorylation site (Cys10) is located at the C-terminal end of the first β strand. No positively charged residues, which could assist in phosphoryl-transfer, can be found in or near the active site. The fold of $IIB^{cellobiose}$ is remarkably similar to that of the mammalian low molecular weight protein tyrosine phosphatases.

Conclusions: A comparison between $IIB^{cellobiose}$ and the structurally similar low molecular weight protein tyrosine phosphatases provides insight into the mechanism of the phosphoryltransfer reactions in which $IIB^{cellobiose}$ is involved. The differences in tertiary structure and active-site composition between $IIB^{cellobiose}$ and the glucose-specific $IIB^{glucose}$ give a structural explanation why the carbohydrate-specific components of different families cannot complement each other.

Addresses: [1]Laboratory of Biophysical Chemistry, Biochemistry and BIOSON Research Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, [2]Department of Biology, University of California at San Diego La Jolla, California 92093-0116, USA and [3]Laboratory of Molecular Biology, Arrhenius Laboratories F3, Stockholm University, S-106 91 Stockholm, Sweden.

*Corresponding author.
E-mail: bauke@chem.rug.nl

Key words: carbohydrate transport, cysteine phosphorylation, IIB enzymes, PTS, X-ray structure

Received: 18 October 1996
Revisions requested: 13 November 1996
Revisions received: 25 November 1996
Accepted: 25 November 1996

Electronic identifier: 0969-2126-005-00217

Structure 15 February 1997, 5:217–225

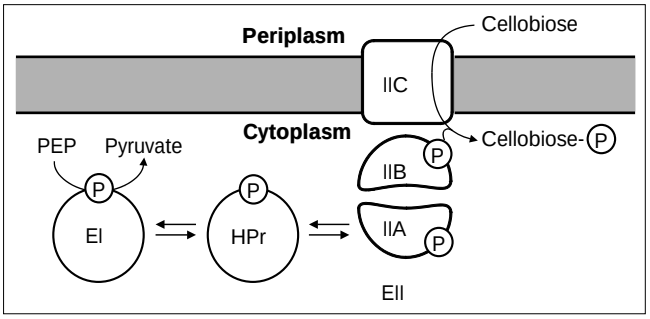
Introduction

The phosphoenolpyruvate-dependent phosphotransferase system (PTS) is a major system for the uptake and concomitant phosphorylation of carbohydrates in Gram-positive and Gram-negative bacteria [1,2]. In contrast to other transport systems, the PTS uses phosphoenolpyruvate (PEP) rather than ATP as an energy source to drive translocation, and it chemically modifies its substrate by phosphorylation. Apart from carbohydrate transport, the PTS has crucial roles in the global regulation of cellular metabolism [3] and in chemotaxis towards its substrates [4].

The PTS is composed of the general proteins enzyme I (E_I) and histidine containing phosphocarrier protein (HPr), and the carbohydrate-specific enzyme II complex (E_{II}). E_{II} usually consists of three functional domains, two cytoplasmic domains — IIA and IIB — and a transmembrane channel, IIC (Fig. 1). In an individual E_{II} complex, these domains may be covalently linked in various ways, or they may exist as distinct proteins. For example in the mannitol-specific E_{II} from *Escherichia coli*, the three functional domains are present on a single polypeptide chain [5], whereas in the cellobiose-specific E_{II} the three domains are present as three separate proteins [6,7].

Transfer of the phosphoryl group from PEP to the carbohydrate proceeds via a series of phosphoprotein intermediates (Fig. 1). In the first steps, the phosphoryl group is transferred from PEP to the IIA domain of the E_{II} complex, via E_I and HPr. IIA phosphorylates IIB, which in turn activates IIC-dependent carbohydrate transport across the cytoplasmic membrane with concurrent phosphorylation of

Figure 1



Schematic representation of the cellobiose-specific phosphotransferase system.

the carbohydrate [8]. The phosphorylated carbohydrate is released in the cytoplasm where it can be used as a first intermediate in bacterial metabolism.

To understand the mechanisms of carbohydrate transport and the regulatory processes in which the PTS is involved requires knowledge of the three-dimensional structures of its components. In recent years, detailed structural studies have been carried out on some PTS components, but for others only scarce or no structural information is available. For example, many attempts have been made to crystallize E_I, but the N-terminal domain of the *E. coli* E_I structure was only recently determined by X-ray crystallography [9]. In contrast, the phosphocarrier protein HPr has been studied in much greater detail, both by X-ray diffraction and by NMR. This has resulted in the determination of a number of structures of HPr proteins from different bacterial sources, and has provided an understanding of the molecular basis for the activities of HPr and its interactions with other proteins (for review, see [10]).

The structural characterization of E_{II} complexes has proven difficult as a result of their complexity and the pronounced differences between them. On the basis of sequence comparisons, E_{II} complexes have been divided into four families: the glucose–sucrose family, the mannose family, the mannitol–fructose family and the lactose–cellobiose family [1,2]. Sequence identities within one family are usually greater than 25% and a component of an E_{II} complex can often be complemented by a homologous component from the same family without severe loss of activity. Between components of different families, however, the sequence similarities are mostly limited to the phosphorylation sites, and complementation by components from different families usually abolishes carbohydrate uptake and phosphorylation. This suggests that major structural and/or functional differences exists between members of the different E_{II} families.

Structure determinations of the IIC domains have been hampered by the difficulties generally encountered in the structural analysis of membrane proteins. Most structural information is available for the cytoplasmic IIA domains. From the glucose–sucrose family X-ray and NMR structures are available of the glucose-specific IIA^{glc} from *E. coli* [11,12] and from *Bacillus subtilis* [13,14]. Also, the X-ray structure of *E. coli* IIA^{glc} complexed with glycerol kinase has been elucidated [15]. Recently the X-ray structure of the mannose-specific IIA domain was reported [16], and an NMR study has revealed the secondary structure of the mannitol-specific IIA domain (IIA^{mtl}) from *E. coli* [17]. Progress has also been made for IIA domains in the fourth class, of which the lactose-specific IIA (IIA^{lac}) has been crystallized [18]. Indeed, the structural information now available for three of the four E_{II} families shows that the structures of the IIA domains differ between families.

Structural detail on the IIB enzymes is rather limited, despite their essential role in the activation of carbohydrate transport by the IIC domains and subsequent phosphorylation of the transported carbohydrate. The only known three-dimensional structure of a IIB enzyme is the very recently solved NMR structure of the IIB^{glc} [19]. It is composed of a single α/β domain with a topology that has so far been observed only for the small domain of the arginine repressor ArgR [20].

In this article, we describe the first X-ray structure of a IIB enzyme from the lactose–cellobiose family, the cellobiose-specific enzyme IIB (IIB^{cel}). IIB^{cel} is a protein of 106 amino acids (molecular weight 11.4 kDa). Like most other IIB enzymes, IIB^{cel} is phosphorylated on a cysteine residue, but it is not covalently linked to the membrane-bound IIC channel. As is the case for the IIA domains of different families, the amino acid sequences of IIB^{glc} and IIB^{cel} are entirely different, as are their three-dimensional structures and the organization of their active-site residues.

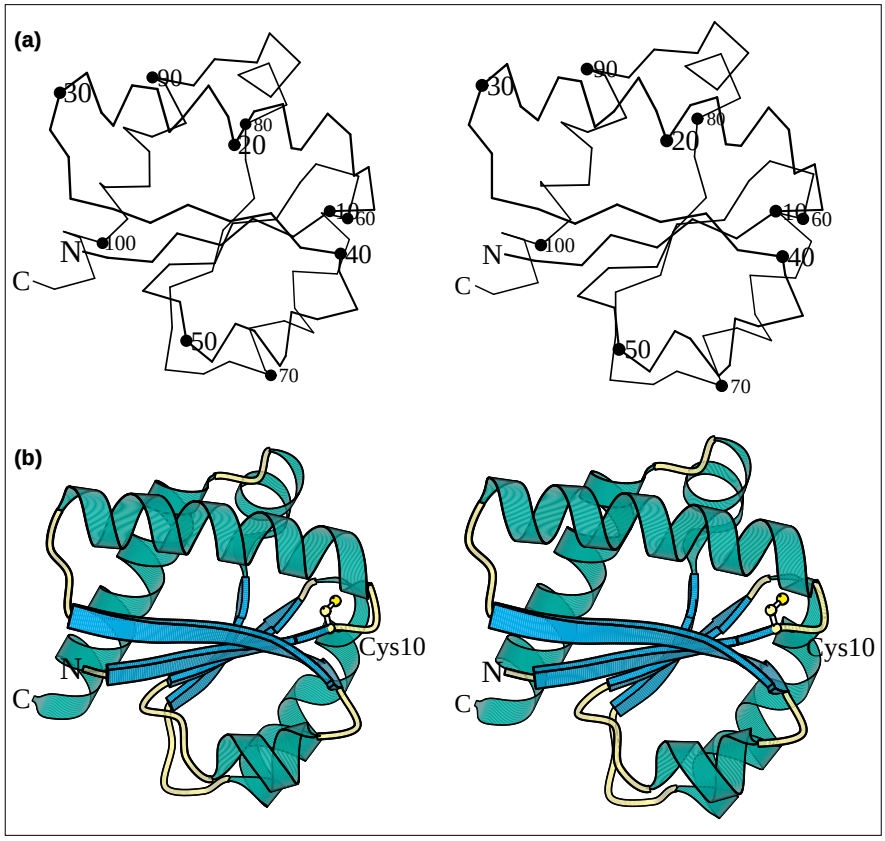
Results and discussion

Protein structure

The IIB^{cel} molecule comprises a single domain with approximate dimensions of 20×27×38 Å. Its structure is composed of a central four-stranded parallel open twisted β sheet, which is flanked by three α helices on the concave side and two on the convex side of the β sheet (Fig. 2). The β sheet is composed mainly of hydrophobic residues and its strand order is 2, 1, 3, 4. The structure contains two right-handed βαβ motifs. The first is formed by β1, α1 and β2, and the second of β3, α3 and β4. The phosphorylation site Cys10 [6] is located at the C-terminal end of the first β strand. Two strictly conserved residues that are probably important during catalysis are Tyr84, the sidechain of which points towards the phosphorylation site, and Gln59, which hovers

Figure 2

The structure of IIB^{cel}. **(a)** Stereoview of the Cα trace of IIB^{cel}, with N and C termini and every tenth residue labeled. **(b)** Ribbon stereoview of the structure of IIB^{cel} generated using the program MOLSCRIPT [46]. Strands are shown in blue, helices in turquoise and loops in light yellow. The catalytic Cys10 is shown in a ball-and-stick representation with the sulphur atom shown in yellow. The N and C termini are indicated.



above the active site with its sidechain in close proximity to the sulphur atom of the catalytic cysteine (Fig. 3). A large part of the active site is formed by the residues of the loop connecting the first β strand to the first α helix. The α helix points with its N-terminal end towards the active site. This suggests that the phosphocysteine intermediate can be stabilized by the macrodipole of the helix, as has been observed in many proteins that bind phosphate or sulphate moieties [21]. Although no significant sequence homology is apparent, the structure of IIB^{cel} appears remarkably

Figure 3

Stereoview of the active site of IIB^{cel} showing the putative P loop in red (color representation is the same as in Fig. 2). Cys10 is shown in ball-and-stick representation. Also shown are the conserved residues Pro58, Gln59 and Tyr84, with carbon atoms shown in light yellow, oxygen in red, nitrogen in blue and sulphur in yellow.

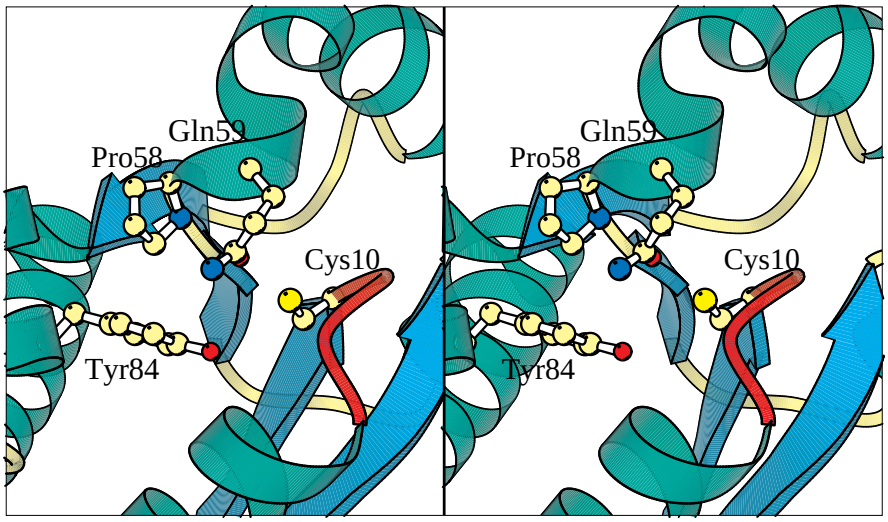
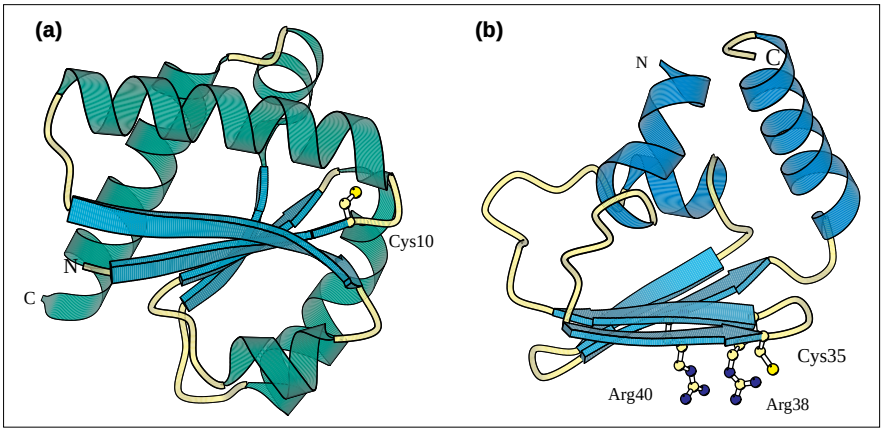


Figure 4



Comparison of the structures of (a) IIB^{cel} (colors are the same as in Fig. 2) and (b) IIB^{glc} (PDB entry code 1IBA). The catalytic cysteines are shown in ball-and-stick representation. In IIB^{glc}, Arg38 and Arg40 are also shown.

similar to the folds of the chemotaxis protein CheY [22], the recently solved IIA domain of the mannose-specific E_{II} [16] and the low molecular weight protein tyrosine phosphatases (LMW PTPases) [23].

Comparison with glucose-specific enzyme IIB, IIB^{glc}

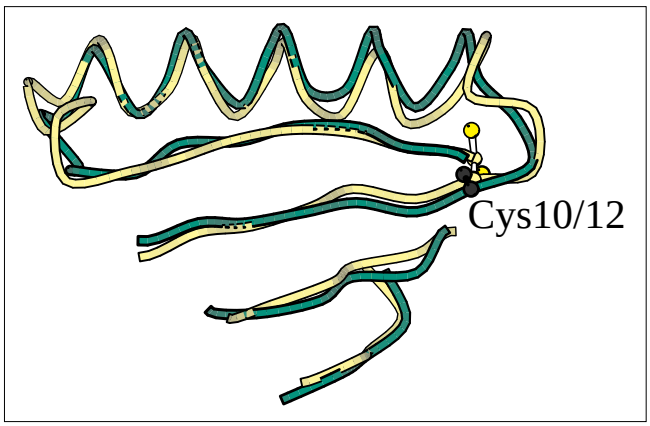
Structural information is now available on two different IIB enzymes, IIB^{glc} and IIB^{cel} (Fig. 4). IIB^{glc} belongs to the glucose–sucrose family, whereas IIB^{cel} is a member of the lactose–cellobiose family. IIB^{glc} is a protein domain of 74 well defined residues [19], which is connected to the membrane-embedded IIC via a flexible linker region [24]. IIB^{cel} is somewhat bigger (106 amino acids) and is a separate protein. IIB^{glc} is composed of an antiparallel four-stranded β sheet packed with three α helices on one side of the sheet. IIB^{cel} contains a parallel four-stranded β sheet flanked with helices on both sides. The catalytic cysteine (Cys35) in IIB^{glc} is located at the C-terminal end of the first β strand and is highly exposed to the solvent. Two arginine residues (Arg38 and Arg40) have been suggested to be involved in the stabilization of the thiolate form of the cysteine and in the stabilization of the negatively charged phosphocysteine intermediate. Moreover, they are both essential in the phosphoryltransfer from IIB^{glc} to the carbohydrate [19]. In contrast, the phosphorylation site of IIB^{cel} (Cys10) is more buried and no positive charges are found in or near the active site. This suggests that activation and stabilization of the catalytic cysteine is different in IIB^{cel} than in IIB^{glc}, and supports the observation that the E_{II} components of different families cannot complement each other.

Comparison with bovine liver LMW PTPase

A superimposition of IIB^{cel} on the structure of bovine liver LMW PTPase [23] shows that, in particular, the central β sheets and the first α helices superimpose very closely (Fig. 5). Both enzymes have reaction mechanisms that feature a phosphocysteine intermediate [25], and their catalytic cysteines are located in equivalent positions.

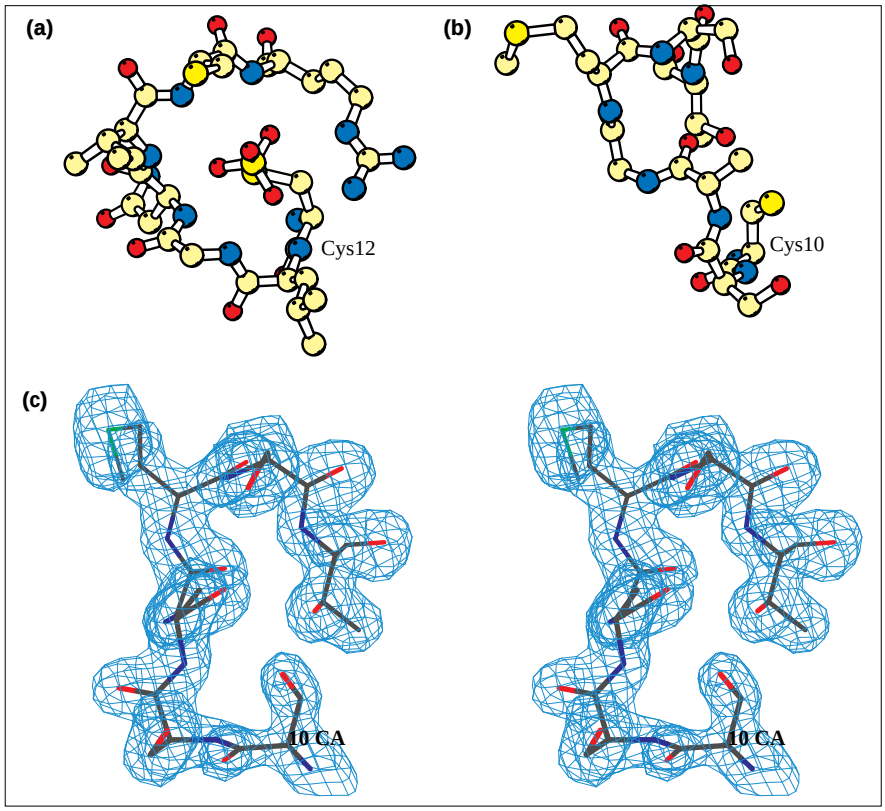
Nevertheless, there are significant differences in the catalytic properties of IIB^{cel} and the PTPase. The PTPase catalytic cysteine residue is part of the PTPase signature sequence Cys-X-X-X-X-X-Arg [25], which connects the first β strand to the first α helix. This sequence forms the rigid phosphate-binding loop (P loop; Fig. 6a) that is stabilized by extensive interactions with surrounding residues. The cradle-like conformation of the loop is structurally conserved among the PTPase family [26]. The mainchain amide groups point towards the centre of the loop, which together with the conserved arginine of the signature sequence and the N-terminal end of the connected α helix provide an excellent environment for binding the phosphate group of the substrate. Moreover, the positively charged P loop probably promotes stabilization of the negatively charged thiolate.

Figure 5



Backbone comparison of the central β sheets and the first α helices of IIB^{cel} (green) and the LMW PTPase from bovine liver (yellow). For 65 Cα atoms, the rms fit is 1.7 Å, as calculated with the lsq options in the program O. Also indicated are the positions of the respective catalytic cysteines.

Figure 6



A comparison of the P loops of bovine liver LMW PTPase and IIB^{cel}. (a) Ball-and-stick representation of the bovine liver PTPase P loop with a bound sulphate ion in the centre of the loop. (b) Ball-and-stick representation of this loop in IIB^{cel}. In (a) and (b) sulphur atoms are shown in yellow and carbon atoms are shown in light yellow. (c) Stereoview of the $2F_o–F_c$ electron-density map for the corresponding loop in IIB^{cel}. The map was contoured at 1σ. The atoms shown are from the final model. Sulphur atoms are shown in green and carbon atoms are shown in grey. Oxygen atoms are shown in red and nitrogen atoms in blue.

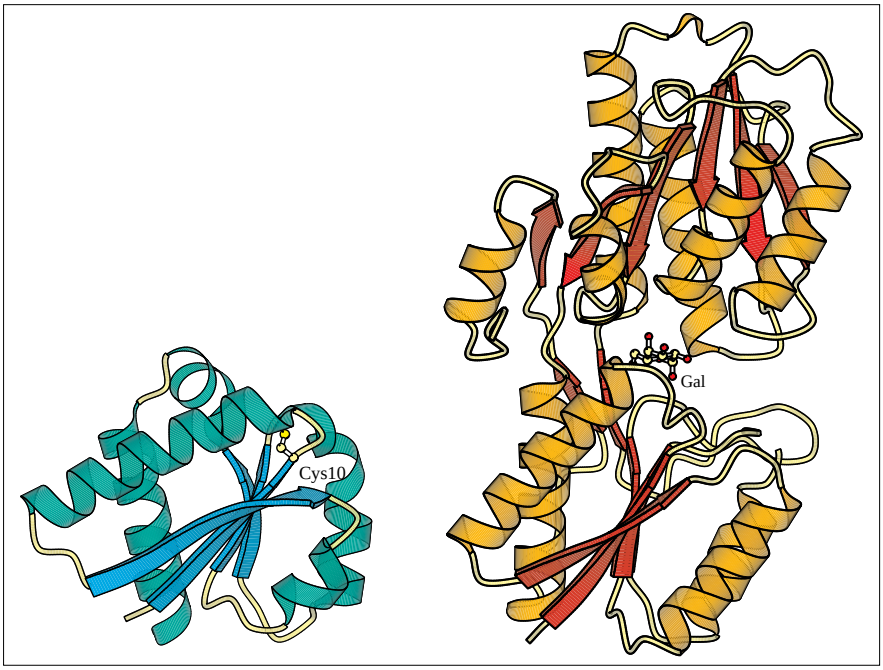
IIB^{cel} contains an active-site loop similar to that of the PTPases, but it lacks the arginine of the signature sequence. In contrast to the PTPase loop, the IIB^{cel} loop is completely exposed to the solvent and does not have the characteristic cradle conformation (Fig. 6b). Stabilization of the catalytic thiolate can only be provided by the backbone amide groups of Ser11 and Ala12 and the sidechain amide group of Gln59, but not by the macrodipole of the following helix as in the PTPases. In the crystal structure, the IIB^{cel} loop is well ordered (Fig. 6c), and crystal contacts prevent the loop forming a conformation that allows positioning of the phosphocysteinyl intermediate in line with the axis of the α helix. One can argue, however, that the loop of the non-phosphorylated form of the protein in solution is able to adopt different conformations and that, upon binding to phosphocysteinyl-IIA^{cel}, it becomes fixed in a cradle-like loop conformation that allows an energetically favourable positioning of the phosphoryl group near the N terminus of the α helix. Indeed, the recently solved NMR structure of IIB^{cel} (C10S mutant) shows an appreciable conformational flexibility in the loop, suggesting that it is disordered in solution [27].

Phosphotransfer reactions

In phosphotyrosine hydrolysis as well as in the PTS phosphotransfer reactions, the transfer of the phosphoryl group proceeds via an associative mechanism with a pentavalent phosphorous intermediate [28,29]. Usually, positively charged residues, often arginines, are important for the stabilization of this negatively charged reaction intermediate. For example, in the PTS, Arg17 of HPr is thought to stabilize the reaction intermediates in the phosphoryltransfer from E_I to HPr [9], and from HPr to IIA^{glc} [30]. Also, the phosphoryltransfer from IIB^{glc} to glucose seems to involve arginine residues (Arg38 and Arg40) of the IIB domain [19]. Moreover, stabilization of the intermediate in the PTPases is provided by the arginine in the signature sequence [31]. IIB^{cel} lacks not only this residue but also any other positively charged residue that could fulfil this role in or near the active site. This suggests that in the phosphotransfer reactions between IIA^{cel} and IIB^{cel}, and between IIB^{cel} and the translocated cellobiose, stabilization of the respective transition states is carried out by residues located on IIA^{cel} and the membrane-bound IIC^{cel}, respectively. In this respect, it is interesting that it has been suggested that at least one histidine residue of IIA^{glc} is required for transfer of the phosphoryl group from IIA^{glc} to IIB^{glc} [13].

In the first step of the reaction catalysed by the LMW PTPases, a strictly conserved aspartic acid protonates the leaving tyrosine. This aspartic acid residue, in its charged form, activates a water molecule required for the hydrolysis

Figure 7



Comparison of the structures of IIB^{cel} and the arabinose-binding protein (ABP; PDB entry code 8ABP). Strands of IIB^{cel} are shown in blue and helices in turquoise. Strands of ABP are shown in red and helices in gold. The galactose (Gal) in ABP is bound at the C-terminal end of the β strands. Superimposition of the N-terminal domain of ABP on IIB^{cel} would position the galactose close to the active site of IIB^{cel}, near Tyr84.

of the phosphocysteinyl intermediate in the subsequent step [23,32]. Because phosphotransfer from IIB^{cel} to cellobiose occurs only after translocation of the carbohydrate, the absence of an aspartic acid or equivalent residue in IIB^{cel} may be a clever way to prevent premature hydrolysis of the phosphocysteinyl intermediate by an activated water molecule. Activation of the cellobiose C6-hydroxyl group is more likely to be mediated by the IIC domain, most probably by residues located in the loop between the putative transmembrane helices 4 and 5, which are known to be somehow involved in substrate binding and phosphorylation [33,34].

Similarity to periplasmic sugar-binding proteins

A structural alignment search using the program DALI [35] revealed not only the homology with CheY and the LMW PTPases but also a strong structural homology with the periplasmic sugar-binding proteins. These proteins consist of two similar domains separated by a deep substrate-binding cleft [36]. Superimposition of IIB^{cel} on the N-terminal domain of the arabinose-binding protein (ABP) shows that the conserved Tyr84 in IIB^{cel} is located just beneath the monosaccharide substrate complexed in ABP (Fig. 7). Since tyrosine residues are known to be involved in carbohydrate binding [37], and Tyr84 is within reach of the phosphoryl-binding site in IIB^{cel}, it is enticing to assume that this residue is part of the recognition site for the IIC^{cel}–cellobiose complex. Unfortunately, soaking studies with the product analogue glucose-6-phosphate destroyed the crystals. Experiments to try to overcome this problem are in progress.

Biological implications

The phosphoenolpyruvate-dependent phosphotransferase system (PTS) is a major system for the uptake of carbohydrates in bacteria; it is also intimately involved in metabolic and transcriptional regulation and in chemotaxis towards PTS substrates. The different functions of the PTS require a flexible system that is responsive to various signals, and this requirement is reflected in the complexity of the multiprotein PTS. To obtain a complete understanding of the mechanism of carbohydrate transport as well as the regulatory properties of the PTS, knowledge of the three-dimensional structures of its components is essential.

Key enzymes in the PTS are the IIB enzymes, which in their phosphorylated state activate the translocation of the carbohydrate by the membrane channel IIC protein. Moreover, the IIB enzymes phosphorylate the sugar moiety once it has crossed the cytoplasmic membrane. The crystal structure of IIB^{cellobiose}, the IIB enzyme of the cellobiose-specific PTS, is composed of a central four-stranded parallel **β** sheet flanked by **α** helices on both sides. The structure of IIB^{cel} has a striking resemblance to the fold of the low molecular weight protein tyrosine phosphatases (LMW PTPases), mammalian proteins that are involved in signal transduction pathways. Like IIB^{cel}, these proteins feature a phosphocysteine intermediate in their reaction mechanism. Another class of proteins that shows a remarkable structural similarity to IIB^{cel}, is the class of the periplasmic sugar-binding proteins (PBPs). The structural similarity of IIB^{cel} to LMW PTPases and

PBPs is intriguing given that IIB^{cel} couples the phosphorylation cascade and sugar transport.

IIB^{cel} is the first member of the lactose–cellobiose PTS family for which the three-dimensional structure has been determined. It may serve as a model for other IIB enzymes of this family and their interactions with other PTS components. In contrast to most of the other IIB enzymes, IIB^{cel} is not covalently linked to the membrane-bound IIC^{cel}, so IIB^{cel} may prove to have additional, separate regulatory functions. Additionally, the absence of positively charged residues in or near to the active site of IIB^{cel}, which for IIB^{glucose} assist in the phosphoryltransfer, suggests that EII enzymes of different families may differ in their detailed mode of action.

Materials and methods

Crystallization and data collection

The IIB^{cel} C10S-mutant was crystallized as described previously [38]. The crystals belong to the space group $P2_1$, with cell dimensions a = 53.6 Å, b = 31.7 Å, c = 60.0 Å and β = 101.7°. The asymmetric unit contains two IIB^{cel} molecules. A native dataset to 2.6 Å (native 1) and all derivative data (see Table 1) were collected on a FAST area detector (Enraf Nonius, Delft, The Netherlands) connected to an Elliot GX 21 rotating anode generator, and reduced with the programs MADNES [39], XDS [40] and the software of the BIOMOL package (protein crystallography group, University of Groningen). A 1.8 Å resolution dataset (native 2) was collected on the MAR Research image plate system at the X-31 beamline of the EMBL Outstation at DESY, Hamburg. This dataset was processed using XDS and the programs of the CCP4 suite [41] and was subsequently scaled to and merged with native 1 with the BIOMOL software (native 3).

Selenomethionine incorporation in IIB^{cel}

The methionine auxotrophic *E. coli* strain LE392 (PROMEGA, Madison, USA) was used for production and overexpression of [SeMet]IIB^{cel}. Transformants of this strain with the plasmid PJL503 were grown on M9 medium supplied with the 19 other amino acids and 50 mg ml^{-1} L-selenomethionine. Isolation and purification of [SeMet]IIB^{cel} was carried out using a procedure similar to that for IIB^{cel}. Crystals could be grown using a strategy similar to that for the C10S mutant, although the optimal precipitant concentrations were considerably lower.

Multiple isomorphous replacement (MIR) analysis

The other heavy atom compounds were prepared by soaking the crystals in solutions of a standard mother liquor, containing 30% (w/v) PEG4000, 5% (v/v) 2-propanol, 25 mM NaCl, 50 mM BES/NaOH buffer (pH 6.5), 0.25% (v/v) glycerol, 0.5 mM NaN_3, and saturated chloro-2,2′:6′,2″-terpyridine platinum(II)chloride (Pt), or 2,6 dichloro-mercury-4-nitrophenol (Hg1), or mercury-phenoxy-glyoxal (Hg2), or Hg2 and Hg1, or 3 mM $K_3UO_2F_5$ (U). Heavy-atom parameters were refined with MLPHARE [41]. The final figure of merit was 0.82 for data to 3.5 Å. The MIR maps were further improved by solvent flattening and histogram matching, using DM [42], followed by 11 cycles of twofold NCS averaging at 3.5 Å with the programs of the DEMON suite [43]. The initial model was fitted to the density using O [44].

Refinement and quality of the model

The model was subjected to simulated annealing refinement with X-PLOR [45]. This was followed by several rounds of positional refinement alternated with individual B factor refinement and manual rebuilding steps. The current model contains the two molecules in the asymmetric unit and 128 water molecules. Each monomer consists of residues 3 to 105. The model has been refined at 1.8 Å to an R factor of 18.7%, R_{free} = 24.1%, with 94.3% of the residues in the most favoured regions and 5.7% in the additional allowed regions of a Ramachandran plot (Fig. 8). The root mean square (rms) deviations from ideality in bond lengths, angles and B factors are 0.009 Å, 1.68 Å and 1.79 Å^2, respectively. Recently, data have been collected for native IIB^{cel}, which revealed no difference other than the Oγ/Sγ substitution in the sidechain of residue 10.

Accession numbers

The coordinates of IIB^{cel} have been deposited in the Brookhaven Protein Data Bank with the entry code 1IIB.

Table 1

Data collection and MIR analysis.

Crystal/derivative*	Native 1	Native 2	Native 3	Pt	Hg1	Hg2	Pt/Hg2	U	Se
Data									
Resolution	2.6	1.8	1.8	2.8	3.5	3.5	2.7	3.5	2.4
No. unique reflections	5461	16 165	17 061	4770	2562	2763	4195	2418	6201
Completeness	80.2	82.8	87.3	91.0	79.7	84.5	94.6	77.8	91.2
R_{merge}†	2.5	6.1	10.5	6.9	9.3	5.3	8.0	8.8	9.2
R_{deriv}‡				10.0	24.9	16.7	20.9	24.3	14.5
MIR analysis									
Resolution used				3.0	3.5	3.5	3.5	3.5	3.5
No. of sites				2	1	4	2	3	5
Phasing power§ centric				1.8	0.6	0.9	1.3	1.0	1.2
Phasing power§ acentric				2.3	0.8	1.3	1.8	1.4	1.5

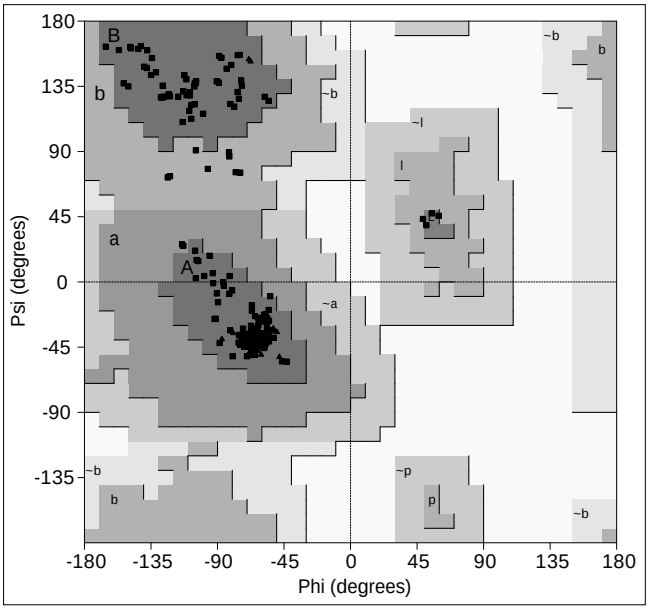
*Heavy-atom derivatives: Pt: chloro-2,2′:6′,2″-terpyridine platinum(II)chloride, Hg1: 2,6 dichloro-mercury-4-nitrophenol, Hg2: mercury-phenoxy-glyoxal, U: $K_3UO_2F_5$, Se: [SeMet]IIB^{cel}.

†$R_{merge} = \Sigma_h\Sigma_i|I_{(h,i)} - <I_{(h)}>|\Sigma_h\Sigma_i<I>_{(h)}$; $I_{(h,i)}$ is the scaled intensity of the ith observation of reflection h and $<I>_{(h)}$ is the mean value.

‡$R_{deriv} = \Sigma_h||F_{PH}| - |F_P||/\Sigma_h|F_P|$; F_{PH} and F_P are the structure factors for native and derivative data.

§Phasing power = $[\Sigma_h|F_{H,calc}|^2/\Sigma_h(|F_{PH,obs}| - |F_{PH,calc}|)^2]^{1/2}$.

Figure 8



Ramachandran plot of the refined model of IIB^{cel}. Glycines are represented as triangles and squares represent all other residues. Disallowed, generously allowed, favorable and most favorable regions are indicated by progressively darker grey shading. The plot was calculated with the program PROCHECK [47].

Acknowledgements

We thank the staff of the EMBL outstation at DESY, Hamburg, for their help during data collection, and the European Union for support of the work at the EMBL Hamburg through the HCMP Access to Large Installations Project, Contract Number CHGE-CT93-0040. The investigations were, in part, supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for Scientific Research (NWO).

References

1. Lengeler, J.W., Jahreis, K. & Wehmeier, U.F. (1994). Enzymes II of the phosphoenolpyruvate-dependent phosphotransferase systems: their structure and function in carbohydrate transport. *Biochim. Biophys. Acta* 1188, 1–28.
2. Postma, P.W., Lengeler, J.W. & Jacobson, G.R. (1993). Phospho-enolpyruvate-carbohydrate phosphotransferase systems of bacteria. *Microbiol. Rev.* 57, 543–594.
3. Saier Jr., M.H. & Reizer, J. (1994). The bacterial phosphotransferase system: new frontiers 30 years later. *Mol. Microbiol.* 13, 755–764.
4. Lux, R., Jahreis, K., Bettenbrock, K., Parkinson, J.S. & Lengeler, J.W. (1995). Coupling the phosphotransferase system and the methyl-accepting chemotaxis protein-dependent chemotaxis signaling pathways of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* 92, 11583–11587.
5. Lee, C.A. & Saier Jr., M.H. (1983). Mannitol-specific enzyme II of the bacterial phosphotransferase system III. The nucleotide sequence of the permease gene. *J. Biol. Chem.* 258, 10761–10767.
6. Reizer, J., Reizer, A. & Saier Jr., M.H. (1990). The cellobiose permease of *Escherichia coli* consists of three proteins and is homologous to the lactose permease of *Staphylococcus aureus*. *Res. Microbiol.* 141, 1061–1067.
7. Lai, X. & Ingram, L.O. (1993). Cloning and sequencing of a cellobiose phosphotransferase system operon from *Bacillus stearothermophilus* XL-65-6 and functional expression in *E. coli*. *J. Bacteriol.* 175, 6441–6450.
8. Lolkema, J.S., Ten Hoeve-Duurkens, R.H., Swaving-Dijkstra, D. & Robillard, G.T. (1991). Mechanistic coupling of transport and phosphorylation activity by enzyme II^{mtl} of the *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system. *Biochemistry* 30, 6716–6721.
9. 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 *E. coli* PEP: sugar phosphotransferase system and a model of the phosphotransfer complex with HPr. *Structure* 4, 861–872.
10. Herzberg, O. & Klevit, R. (1994). Unraveling a bacterial hexose transport pathway. *Curr. Opin. Struct. Biol.* 4, 814–822.
11. 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.
12. Pelton, J.G., Torchia, D.A., Meadow, N.D., Wong, C.Y. & Roseman, S. (1991). ^{1}H, ^{15}N, and ^{13}C NMR signal assignments of III^{glc}, a signal-transducing protein of *Escherichia coli*, using three-dimensional triple-resonance techniques. *Biochemistry* 30, 10043–10057.
13. Liao, D.-I., Kapadia, G., Reddy, P., Saier Jr., M.H., 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.
14. Fairbrother, W.J., Gippert, G.P., Reizer, J., Saier, M.H.J. & Wright, P.E. (1992). Low resolution solution structure of the *Bacillus subtilis* glucose permease IIA domain derived from heteronuclear three-dimensional NMR spectroscopy. *FEBS Lett.* 296, 148–152.
15. Hurley, J.H., *et al.*, & Remington, S.J. (1993). Structure of the regulatory complex of *Escherichia coli* III^{glc} with glycerol kinase. *Science* 259, 673–677.
16. Nunn, R.S., *et al.*, & Erni, B. (1996). Structure of the mannose transporter from *Escherichia coli* at 1.7 Å resolution. *J. Mol. Biol.* 259, 502–511.
17. Kroon, G.J.A., Grötzinger, 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 heteronuclear three dimensional spectroscopy. *Protein Sci.* 2, 1331–1341.
18. Celikel, R., *et al.*, & Reizer, J. (1991). Crystallization an 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.
19. Eberstadt, M., Gradnolik, 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.
20. van Duyne, G.D., Ghosh, G., Maas, W.K. & Sigler, P.B. (1996). Structure of the oligomerization and L-arginine binding domain of the arginine repressor of *Escherichia coli*. *J. Mol. Biol.* 256, 377–391.
21. Hol, W.G.J., van Duijnen, P.T. & Berendsen, H.J.C. (1978). The α helix dipole and the properties of proteins. *Nature* 273, 443–446.
22. Voltz, K. & Matsumura, P. (1991). Crystal structure of *Escherichia coli* CheY refined at 1.7 Å resolution. *J. Biol. Chem.* 266, 15511–15519.
23. Su, X.D., Taddei, N., Stefani, M., Ramponi, G. & Nordlund, P. (1994). The crystal structure of a low-molecular-weight phosphotyrosine protein phosphatase. *Nature* 370, 575–578.
24. Hummel, U., Nuoffer, C., Zanolari, B. & Erni, B. (1992). A functional protein hybrid between the glucose transporter and the *N*-acetylglucosamine transporter of *Escherichia coli*. *Protein Sci.* 1, 356–362.
25. Cirri, P., *et al.*, & Ramponi, G. (1993). The role of Cys12, Cys17 and Arg18 in the catalytic mechanism of low-Mr cytosolic phosphotyrosine protein phosphatase. *Eur. J. Biochem.* 214, 647–657.
26. Barford, D. (1995). Protein phosphatases. *Curr. Opin. Struct. Biol.* 5, 728–734.
27. AB, E., *et al.*, & Robillard, G.T. (1996). The NMR sidechain assignments and solution structure of enzyme IIB^{cellobiose} of the phosphoenolpyruvate-dependent phosphotransferase system of *Escherichia coli*. *Protein Sci.*, in press.
28. Su, X.D. (1994). Structural studies on key enzymes involved in cell proliferation. PhD thesis, Karolinska Institutet, Stockholm.
29. Mueller, E.G., Khandekar, S.S., Knowles, J.R. & Jacobson, G.R. (1990). Stereochemical course of the reactions catalyzed by the bacterial phosphoenolpyruvate:mannitol phosphotransferase system. *Biochemistry* 29, 6892–6896.
30. Herzberg, O. (1992). An atomic model for protein-protein phosphoryl group transfer. *J. Biol. Chem.* 267, 24819–24823.
31. Zhang, Z.-Y., *et al.*, & Dixon, J.E. (1994). The Cys(X)$_5$Arg catalytic motif in phosphoester hydrolysis. *Biochemistry* 33, 15266–15270.
32. Taddei, N., *et al.*, & Ramponi, G. (1994). Aspartic-129 is an essential residue in the catalytic mechanism of the low M(r) phosphotyrosine protein phosphatase. *FEBS Lett.* 350, 328–332.

33. Ruyter, G.J.G., van Meurs, G., Verwey, M.A., Postma, P.W. & van Dam, K. (1992). Analysis of mutations that uncouple transport from phosphorylation in enzyme II^{glc} of the *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system. *J. Bacteriol.* 174, 2843–2850.
34. Weng, Q.-P., Elder, J. & Jacobson, G.R. (1992). Site-specific mutagenesis of residues in the *Escherichia coli* mannitol permease that have been suggested to be important for its phosphorylation and chemoreception functions. *J. Biol. Chem.* 267, 19529–19535.
35. Holm, L. & Sander, C. (1994). The FSSP database of structurally aligned protein fold families. *Nucl. Acids Res.* 22, 3600–3609.
36. Quiocho, F.A. (1991). Atomic structures and function of periplasmic receptors for active transport and chemotaxis. *Curr. Opin. Struct. Biol.* 1, 922–933.
37. Lawson, C.L., *et al.*, & Dijkstra, B.W. (1994). Nucleotide sequence and X-ray structure of cyclodextrin glycosyltransferase from *Bacillus circulans* strain 251 in a malto-dependent crystal form. *J. Mol. Biol.* 236, 590–600.
38. van Montfort, R.L.M., *et al.*, & Dijkstra, B.W. (1994). Crystallization of enzyme IIB of the cellobiose-specific phosphotransferase system of *Escherichia coli*. *J. Mol. Biol.* 239, 588–590.
39. Messerschmidt, A. & Pflugrath, J.W. (1987). Crystal orientation and X–ray pattern prediction routines for area-detector diffractometer systems in macromolecular crystallography. *J. Appl. Cryst.* 20, 306–315.
40. Kabsch, W. (1988). Evaluation of single crystal X-ray diffraction data from a position-sensitive detector. *J. Appl. Cryst.* 21, 916–924.
41. Collaborative Computational Project No. 4 (1994). The CCP4 Suite: programs for protein crystallography. *Acta Cryst. D* 50, 760–763.
42. Cowtan, K.D. & Main, P. (1996). Phase combination and cross validation in iterated density modifcation calculations. *Acta. Cryst. D* 52, 756–764.
43. Vellieux, F.M.D.A.P. (1995). DEMON/ANGEL: a suite of programs to carry out density modification. *J. Appl. Cryst.* 28, 347–351.
44. Jones, T.A., Zou, J.-Y., Cowan, S.W. & Kjelgaard, M. (1991). Improved methods for building protein models in electron-density maps and the location of errors in these models. *Acta Cryst. A* 47, 110–119.
45. Brünger, A.T., Kuriyan, J. & Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. *Science* 235, 458–460.
46. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946–950.
47. 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.