Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of PEP carboxykinase from the succinate-producing *Actinobacillus succinogenes*: a new conserved active-site motif

Yvonne A. Leduc,[a] Lata Prasad,[a] Maris Laivenieks,[b] J. Gregory Zeikus[b] and Louis T. J. Delbaere[a]*

[a]Department of Biochemistry, University of Saskatchewan, Saskatoon, Saskatchewan S7N 5E5, Canada, and [b]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, USA

Correspondence e-mail: louis.delbaere@usask.ca

Received 6 February 2005
Accepted 17 March 2005

PDB References: PEP carboxykinase, 1ygg, r1yggsf; complex with manganese and pyruvate, 1ylh, r1ylhsf.

Actinobacillus succinogenes can produce, *via* fermentation, high concentrations of succinate, an important industrial commodity. A key enzyme in this pathway is phospho*enol*-pyruvate carboxykinase (PCK), which catalyzes the production of oxaloacetate from phospho*enol*pyruvate and carbon dioxide, with the concomitant conversion of adenosine 5′-diphosphate to adenosine 5′-triphosphate. 1.85 and 1.70 Å resolution structures of the native and a pyruvate/Mn^{2+}/phosphate complex have been solved, respectively. The structure of the complex contains sulfhydryl reducing agents covalently bound to three cysteine residues *via* disulfide bonds. One of these cysteine residues (Cys285) is located in the active-site cleft and may be analogous to the putative reactive cysteine of PCK from *Trypanosoma cruzi*. Cys285 is also part of a previously unreported conserved motif comprising residues 280–287 and containing the pattern N*X*E*X*G*X*Y(/F)A(/G); this new motif appears to have a structural role in stabilizing and positioning side chains that bind substrates and metal ions. The first few residues of this motif connect the two domains of the enzyme and a fulcrum point appears to be located near Asn280. In addition, an active-site Asp residue forms two coordinate bonds with the Mn^{2+} ion present in the structure of the complex in a symmetrical bidentate manner, unlike in other PCK structures that contain a manganese ion.

1. Introduction

1.1. *Actinobacillus succinogenes*

Succinate, a four-carbon dicarboxylic acid, is an important industrial and specialty chemical currently produced from petrochemicals. Because of shrinking petroleum resources and the desire for 'greener' technology and lower costs, fermentation is being investigated as an alternative way to produce succinate and other industrial chemicals (Lee *et al.*, 2004; Zeikus *et al.*, 1999).

A. succinogenes, an anaerobic bacterium isolated from the rumen of cattle (Guettler *et al.*, 1999), produces one of the highest succinate concentrations ever reported *via* the fermentation of simple carbon substrates. Furthermore, *A. succinogenes* can tolerate the high sugar, salt and organic acid concentrations of the fermentation environment (Guettler *et al.*, 1999; Zeikus *et al.*, 1999).

1.2. Phospho*enol*pyruvate carboxykinase

The first step in the formation of succinate in *A. succinogenes* is catalyzed by phospho*enol*pyruvate carboxykinase (PCK), a 538-residue enzyme ($M_r \simeq 59$ kDa). PCK [ATP-

oxaloacetate carboxylyase (transphosphorylating); EC 4.1.1.49] catalyzes the formation of oxaloacetate (OAA) from the reversible carboxylation of phospho*enol*pyruvate (PEP) and the concomitant phosphorylation of nucleoside diphosphate (NDP) to nucleoside triphosphate (NTP; Utter & Kolenbrander, 1972). A divalent metal cation such as Mg^{2+} or Mn^{2+} is required for catalysis,

$$\mathrm{PEP} + \mathrm{NDP} + \mathrm{CO_2} \longleftrightarrow \mathrm{OAA} + \mathrm{NTP}.$$

This reaction fixes carbon dioxide, a known 'greenhouse' gas, and is one reason that succinate fermentation is considered a 'greener' technology (Zeikus *et al.*, 1999). In most organisms, PCK primarily catalyzes the reverse direction of this reaction as the first committed step of gluconeogenesis. In *A. succinogenes*, however, PCK catalyzes the forward direction of this reaction for the eventual formation of succinate.

PCKs are classified according to nucleotide specificity, with ATP-dependent enzymes present in most bacteria (including *A. succinogenes*), yeast, trypanomastids and plants, and the GTP-dependent PCKs [GTP-oxaloacetate carboxylyase (transphosphorylating), EC 4.1.1.32] present in a variety of other eukaryotes (including mammals) and in some bacteria (Matte *et al.*, 1996).

Enzymes within their respective ATP- or GTP-dependent class share high amino-acid sequence identity, but homology between the two classes is not statistically significant (Table 1). Nonetheless, the key residues for substrate binding and metal coordination are known to be conserved between classes (Fig. 1). The nucleotide-binding residues are conserved within each class. Conservation was confirmed at the structural level by an examination of the X-ray crystallographic structures of PCK published to date from *Escherichia coli* (EcPCK; Matte *et al.*, 1996; Tari *et al.*, 1996, 1997), *Trypanosoma cruzi* (TcPCK; Trapani *et al.*, 2001) and human (hPCK; Dunten *et al.*, 2002).

A. succinogenes PCK (AbsPCK) has been cloned and overexpressed in *E. coli* and purified. The purified protein has been crystallized and structures of the native form and a complex with pyruvate, Mn^{2+} and phosphate have been solved to 1.85 and 1.70 Å resolution, respectively.

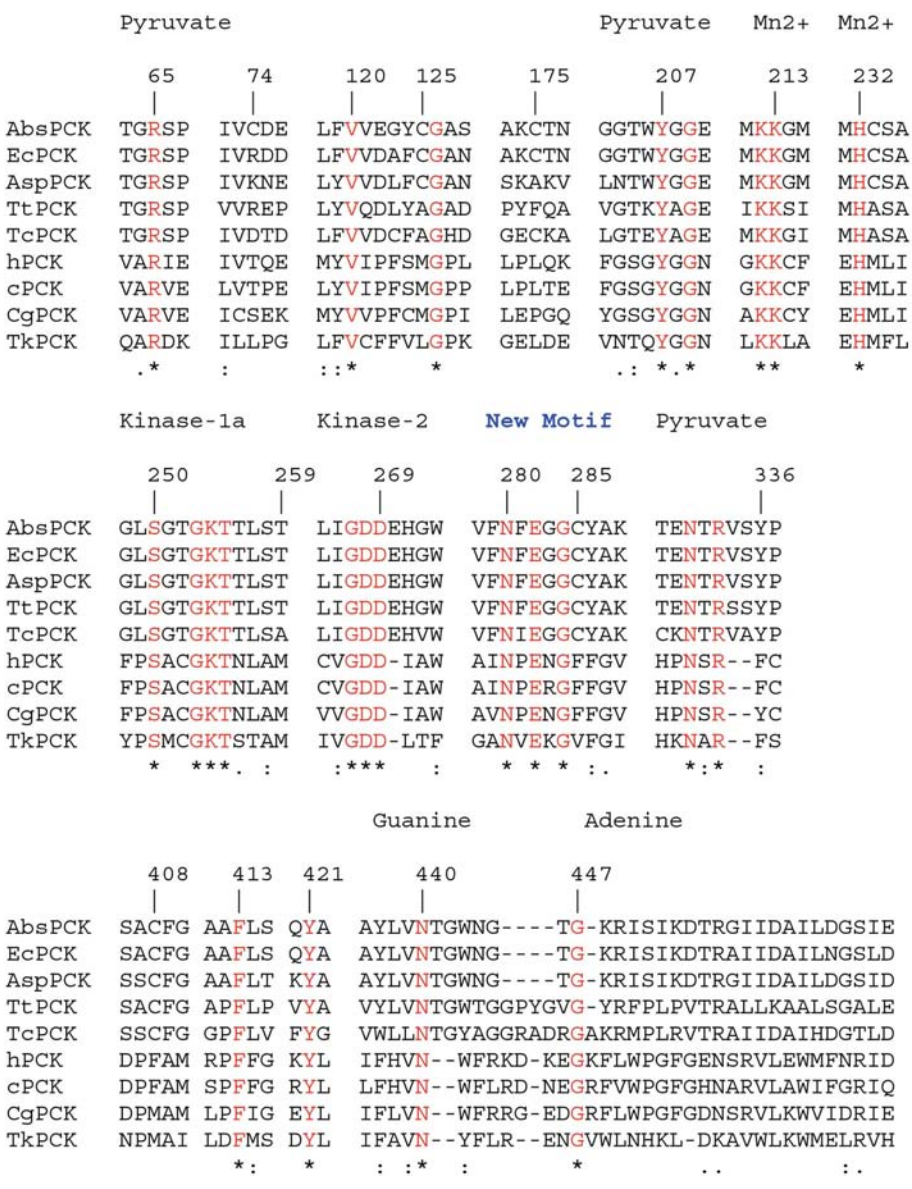

Figure 1
Multiple sequence alignment of selected regions of representative ATP- and GTP-dependent PCKs. The selected regions are indicated above the sequence (new motif in blue) and are regions that contain conserved residues that interact with ligands and that contain the six cysteine residues found in AbsPCK. Sequences were aligned using *CLUSTALW* v.1.82 (Chenna *et al.*, 2003). Residue numbering is that of *E. coli*. Residues in red and marked '*' are identical in all sequences in the alignment. Conserved substitutions of residues in a column are marked ':'. Semi-conserved substitutions of residues in a column are marked '.'. Primary accession numbers: AbsPCK, Q6W6X5 (*Actinobacillus succinogenes*); EcPCK, P22259 (*Escherichia coli*); AspPCK, O09460 (*Anaerobiospirillum succiniciproducens*); TtPCK, Q7SIC6 (*Thermus thermophilus*); TcPCK, P51058 (*Trypanosoma cruzi*); hPCK, P35558 (human cytosolic); cPCK, P21642 (chicken mitochondrial); CgPCK, Q9AEM1 (*Corynebacterium glutamicum*); TkPCK, Q6F494 (*Thermococcus kodakaraensis*).

2. Materials and methods

2.1. Cloning, protein expression and purification

The coding region for the cloned *pckA* gene (Kim *et al.*, 2004) of *A. succinogenes* 130Z was PCR-amplified and subcloned into the *Ehe*I–*Sal*I sites of vector pProEx-1 (Invitrogen, Carlsbad, CA, USA; plasmid pAsPCK). AbsPCK was overexpressed as a recombinant six-His-tagged protein in

Table 1
Sequence identity of various ATP- and GTP-dependent PCKs with AbsPCK.

Organism	Identity† with AbsPCK (%)	Nucleotide specificity
Escherichia coli	74	ATP
Anaerobiospirillum succiniciproducens	65	ATP
Thermus thermophilus	45	ATP
Trypanosoma cruzi	41	ATP
Human (cytosolic)	<10	GTP
Chicken (mitochondrial)	<10	GTP
Corynebacterium glutamicum	<10	GTP
Thermococcus kodakaraensis	<10	GTP

† Identity calculated with *bl2seq* (Tatusova & Madden, 1999).

E. coli PB25 and purified as previously described for *Anaerobiospirillum succiniciproducens* PCK (AspPCK; Laivenieks *et al.*, 1997). The purified protein was concentrated with an Amicon stirred ultrafiltration cell with a 30 kDa molecular-weight cutoff YM30 membrane and the elution buffer [20 m*M* Tris–HCl pH 8.5, 100 m*M* KCl, 100 m*M* imidazole, 10 m*M* 2-mercaptoethanol, 10%(*v*/*v*) glycerol] was exchanged for storage buffer [50 m*M* MES–NaOH pH 6.5, 2 m*M* dithiothreitol (DTT), 50 m*M* trehalose and 1.0% PEG 8000] at 277 K. The protein migrated through a 12% SDS–polyacrylamide gel with a mobility suggesting a size of ∼59 kDa.

The protein at a concentration of 4 mg ml^{-1} in storage buffer was then flash-frozen in liquid nitrogen, lyophilized and stored at 253 K. Prior to crystallization, a weighed quantity of dehydrated protein and storage buffer was redissolved in distilled water. Using a Millipore Ultrafree (30 kDa molecular-weight cutoff) centrifugal filter device, the protein was concentrated and the storage buffer-exchanged with 25 m*M* MES–NaOH pH 6.5, 2 m*M* DTT and 1 m*M* EDTA at 277 K.

2.2. Crystallization

The native protein was crystallized by vapour diffusion in hanging drops by mixing 1 µl of 8 mg ml^{-1} protein in 25 m*M* MES–NaOH pH 6.5, 2 m*M* DTT and 1 m*M* EDTA with 1 µl reservoir solution containing 30% saturated ammonium sulfate, 0.2 *M* potassium sodium tartrate and 0.1 *M* sodium citrate pH 5.5 and incubating at 285 K. Rod-shaped crystals approximately 0.2 mm in length appeared after about four weeks. Crystals were transferred briefly to a cryoprotectant consisting of 35% glycerol, 30% saturated ammonium sulfate, 0.15 *M* potassium sodium tartrate and 0.075 *M* sodium citrate pH 5.5 and flash-cooled in liquid nitrogen.

The protein with substrates was crystallized by vapour diffusion in hanging drops by mixing 1 µl of 8 mg ml^{-1} protein in 25 m*M* MES–NaOH pH 6.5, 2 m*M* DTT, 1 m*M* EDTA, 5 m*M* PEP, 5 m*M* $MgCl_2$ and 5 m*M* $MnCl_2$ with 1 µl reservoir solution containing 20% PEG 4000, 0.2 *M* sodium formate and 0.1 *M* Tris–HCl pH 8.0 and incubating at room temperature. Rod-shaped crystals approximately 0.1 mm in length appeared in a cluster after about five months. A piece of this cluster was transferred briefly to a cryoprotectant consisting of 40% ethylene glycol, 20% PEG 4000, 0.15 *M* sodium formate and 0.075 *M* Tris–HCl pH 8.0 and flash-cooled in liquid nitrogen.

2.3. Data collection and processing

X-ray diffraction data sets from both crystals were collected at 100 K on BioCARS beamline 14-ID-B at the Advanced Photon Source, Argonne National Laboratories, Argonne, IL, USA using a MAR 165 CCD detector, a wavelength of 0.9984 Å and 1.0° rotation images with a crystal-to-detector distance of 125 mm for the native and 110 mm for the complex. Diffraction data were processed using *HKL* v.1.97.3, *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

2.4. Molecular replacement and crystallographic refinement

The native structure was solved by molecular replacement using *AMoRe* (Navaza, 1994) from the *CCP*4 package of programs (Collaborative Computational Project, Number 4, 1994) and *E. coli* PCK (PDB code 1oen) as the search model. The structure of the complex was similarly solved by molecular replacement using *AMoRe*, with the two domains of the native enzyme as search models. The correlation coefficients and starting *R* factors were 0.61 and 0.39, respectively, for the native data set and 0.74 and 0.35, respectively, for the complex data set.

The models were subsequently refined with *CNS* (Brünger *et al.*, 1998) and the model visualized and manually adjusted with *TURBO-FRODO* v.5.5 (Roussel & Cambillau, 1991). The final $R_{\rm work}$ was 0.199 and 0.190 for the native and complex, respectively.

Each structure was validated during and at the end of the refinement process with the validation suite of programs located at http://biotech.ebi.ac.uk:8400/ with *PROCHECK* v.3.5 (Laskowski *et al.*, 1993) and *WHAT IF* v.4.99 (Vriend, 1990).

2.5. Sequence and structural analysis and comparison

Sequences of PCKs were compared for conserved residues using *CLUSTALW* v.1.82 (Chenna *et al.*, 2003) from the European Bioinformatics Institute server and the program *Blast 2 Sequences* (*bl2seq*; Tatusova & Madden, 1999) located at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov). The topologies of the AbsPCK, EcPCK, TcPCK and hPCK structures were aligned using *LSQKAB* (Kabsch, 1976) and *TOPP* (Lu, 1996) from the *CCP*4 suite of programs (Collaborative Computational Project, Number 4, 1994; Potterton *et al.*, 2002, 2003; Winn, 2003) and visualized with *TURBO-FRODO* to confirm the conserved residues and compare and analyze their environments by overlaying the structures. A comparison of domain alignment and movement amongst the structures of AbsPCK, EcPCK and TcPCK was also performed using the program *DynDom* (Hayward & Berendsen, 1998) located at http://www.sys.uea.ac.uk/dyndom and the program *LSQKAB*.

The sequence numbering of EcPCK is used in this article. The equivalent AbsPCK residue is one number less than

Table 2
Data-collection and refinement statistics for AbsPCK structures.

Values in parentheses are for the highest resolution shell.

	Native	Complex
Data collection		
Space group	$P4_3$	$P2_1$
Unit-cell parameters		
a (Å)	102.09	56.74
b (Å)	102.09	55.09
c (Å)	72.12	90.08
β (°)	90	106.17
No. of molecules in AU	1	1
Resolution range (Å)	50.0–1.85 (1.92–1.85)	50.0–1.70 (1.76–1.70)
No. of reflections measured	459360	222834
No. of unique reflections	62780 (6231)	58868 (5857)
R_{sym}†	0.069 (0.433)	0.043 (0.269)
Completeness (%)	99.8 (99.3)	99.9 (99.7)
Redundancy	7.3	3.8
Mean $I/\sigma(I)$	25.6 (3.2)	29.0 (5.5)
Refinement statistics		
Resolution limits (Å)	26.36–1.85 (1.97–1.85)	26.26–1.70 (1.81–1.70)
R_{free}‡	0.207 (0.265)	0.215 (0.255)
R_{work}‡	0.199 (0.262)	0.190 (0.219)
R_{cryst}‡	0.202 (0.286)	0.195 (0.248)
Total No. of reflections	59533 (8244)	57207 (8365)
No. of non-H protein atoms	4056	4033
No. of heteroatoms	11	45
No. of water molecules	157	197
Mean B factors		
Protein atoms (Å^2)	32.73	19.37
Water molecules (Å^2)	31.34	22.77
R.m.s. deviation from ideal geometry		
Bond distances (Å)	0.010	0.010
Bond angles (°)	1.5	1.6
Dihedral angles (°)	23.9	23.9
Improper angles (°)	0.91	0.93
Ramachandran statistics		
Most favored (%)	90.9	90.2
Favored (%)	8.6	9.3
Additional allowed (%)	0.2	0.2
Disallowed (%)	0.2	0.2

† $R_{sym} = \sum|\langle I_{hkl}\rangle - I_{hkl}|/|I_{hkl}|$, where $\langle I_{hkl}\rangle$ is the average intensity over symmetry-related reflections and I_{hkl} is the observed intensity. ‡ R value $= \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are the observed and calculated structure factors. For R_{free}, the sum is performed on the test-set reflections (~5% of total reflections), for R_{work} on the remaining reflections and for R_{cryst} on all reflections included in the resolution range.

EcPCK for the first 88 residues and two numbers less for the remainder of the protein.

The Macromolecular Structure Database located at http://www.ebi.ac.uk/msd-srv/msdsite/index.jsp with 27 927 structures from the Protein Data Bank (PDB) was searched for instances of DTT and 2-mercaptoethanol covalently bonded to the terminal S atom of a cysteine residue and for the known geometries of sodium-ion binding.

Figs. 2–7 were produced with *PyMOL* v.0.98beta (DeLano, 2002).

3. Results and discussion

3.1. Overall structure

The structure of AbsPCK in the native form was solved by molecular replacement using EcPCK (PDB code 1oen; Matte *et al.*, 1996) as the search model. The structure of the complex was subsequently solved by molecular replacement using the native AbsPCK structure (PDB code 1ygg) as the model. Statistics of data collection and the final refined structures are summarized in Table 2.

The native crystal belonged to space group $P4_3$, with a solvent content of 59% and a Matthews coefficient (Matthews, 1968) of 3.1 Å^3 Da^{-1} (assuming a molecular weight of 59.9 kDa and a molecular density of 1.30 g cm^{-3}). The crystal of the complex belonged to space group $P2_1$, with a solvent content of 43% and a Matthews coefficient of 2.3 Å^3 Da^{-1} calculated with the same assumptions.

In both structures, the N-terminal His tag, the first two residues of the molecule and the last residue are not visible in the electron density. Most residues comprising a β-hairpin region in both structures (approximately residues 390–400) were also not visible. The corresponding region in the structures of EcPCK, hPCK and TcPCK also has missing residues. It has been hypothesized that this flexible surface β-hairpin may function as an active-site lid in both EcPCK and TcPCK (Trapani *et al.*, 2001). Poorly defined residues with missing atoms in both AbsPCK structures were located on surface regions of the protein where flexible conformations are expected.

Ramachandran statistics for both structures were good, with greater than 90% of residues being in the most favoured region (Table 2). The residues outside the favoured regions were confirmed against the electron density and were well matched. Residue Ile477 in the disallowed region is equivalent to Met477 in EcPCK, which is also in the disallowed region (Matte *et al.*, 1996). In TcPCK, the equivalent residue is Gly456 and therefore has fewer restrictions on its geometry. In hPCK, the equivalent residue is Ile555 which is within the favoured region of the Ramachandran plot. In all structures, this residue is part of a γ-turn between two antiparallel β-strands adjacent to and interacting *via* hydrogen bonds with the C-terminus.

3.2. Ligands and substrates

The native structure contained one sodium ion and two sulfate ions. The structure of the complex contained three formates, two phosphate ions, one Mn^{2+} ion, one pyruvate, one covalently bound molecule of sulfonated DTT and two covalently bound molecules of 2-mercaptoethanol.

The manganese ion in the structure of the AbsPCK complex (Fig. 2) is octahedrally coordinated as in other PCK structures, with the conserved residues Lys213, His232 and Asp269 and three water molecules. However, in the AbsPCK structure, Asp269 forms two coordination bonds to the metal ion in the less common symmetrical bidentate manner, unlike the equivalent Asp residues in EcPCK (PDB code 1aq2) and hPCK (PDB codes 1khb, 1khe, 1khf, 1khg), which form only one bond to the manganese ion (Figs. 2 and 3) in an asymmetrical *syn*-monodentate manner. The bonds formed and their distances are given in Table 3. Manganese coordination numbers, including seven, and their coordination patterns with carboxylates in both small molecules (Harding, 1999) and

proteins (Harding, 2004) have been discussed previously. In the PDB, bonds between a manganese ion and the side chain of an Asp or Glu residue are very common, with asymmetrical *syn*-monodentate being the predominant arrangement.

Table 3
Bond distances (Å) in the Mn^{2+} coordination geometry of some PCKs.

Bond	AbsPCK†	EcPCK‡	hPCK§
Mn—His232	2.23	2.35	2.22
Mn—Lys213	2.29	2.46	2.32
Mn—Asp269 $O^{\gamma 1}$	2.57	3.08¶	3.20¶
Mn—Asp269 $O^{\gamma 2}$	2.35	2.30	2.18
Mn—H_2O	2.20	2.30	2.20
Mn—H_2O	2.39	2.40	2.24
Mn—H_2O	2.40	—	2.25
Mn—$O^{\gamma 1}$ of ATP	—	2.1	—
No. of bonded atoms	7	6	6

† AbsPCK complex with pyruvate and Mn^{2+} (PDB code 1ylh). ‡ EcPCK complex with ATP, Mg^{2+}, Mn^{2+} and pyruvate (PDB code 1aq2). § hPCK complex with Mn^{2+} and PEP (PDB code 1khf). ¶ Distance too great for metal coordination bond.

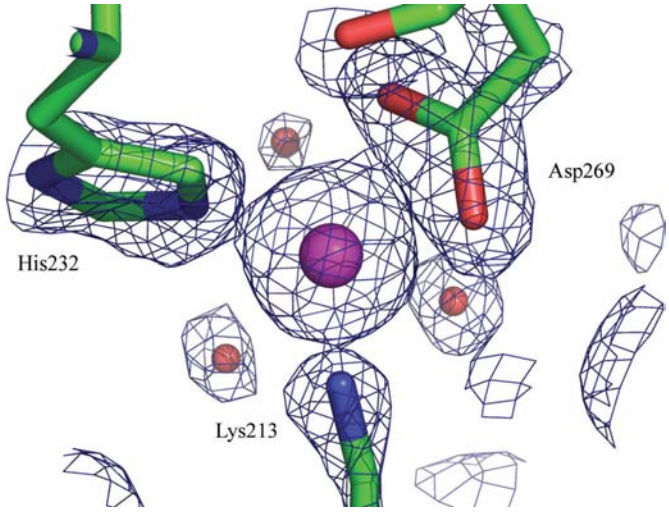

Figure 2
Manganese coordination site of the structure of the AbsPCK complex, showing Lys213, His232, and Asp269. Small spheres are water and the large magenta sphere is Mn^{2+}; electron density was contoured at 2.0σ.

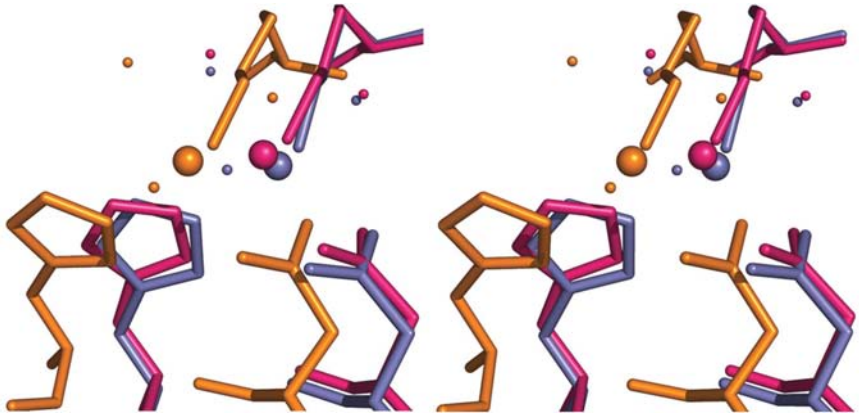

Figure 3
Stereoview of an overlay of the manganese coordination sites of hPCK (PDB code 1khf; orange), EcPCK (PDB code 1aq2; pink) and AbsPCK–pyruvate–Mn^{2+} (PDB code 1ylh; blue). Small spheres are water, large spheres are Mn^{2+}.

Additionally, manganese bonded to seven different atoms is frequently found in the PDB. Troponin-C in complex with manganese (PDB code 1ncy) is one example of an octahedrally coordinated manganese ion with seven bonded atoms including two from a Glu residue in a symmetrical bidentate arrangement. The bond distances from the Glu residue to the ion were comparable to the bond distances from the Asp residue to the ion in the AbsPCK structure, with one bond from the residue being slightly longer than the second bond from the same residue.

In the structure of the complex, the PEP present during crystallization appears to have been hydrolyzed to pyruvate and phosphate. It is known that PEP binds tightly to the enzyme in the presence of Mn^{2+} and that such a reaction is possible because Mn^{2+} is believed to interact directly with the transferred phosphate group (Hebda & Nowak, 1982; Miller *et al.*, 1968). A similar reaction occurred when EcPCK was crystallized in the presence of ADP, PEP, Mg^{2+} and Mn^{2+}: ATP and pyruvate were found in the structure (Tari *et al.*, 1997).

The pyruvate in the complex is in the same postively charged pocket near the manganese-binding site as the substrates (pyruvate, oxalate, acetate and PEP) found in other PCK structures. In the structure of the complex of AbsPCK, pyruvate is hydrogen bonded to the terminal hydroxyl group of Tyr207, N^{ω} of Arg333, the N^{δ} atom of Arg65 and three water molecules, two of which are coordinating the Mn^{2+} ion.

Although inorganic phosphate was not present during crystallization, two phosphate ions are visible in the structure of the complex, presumably produced from the hydrolysis of PEP as noted above. One is present on the protein's surface and does not appear to play an important role. The other is hydrogen bonded to the main chain and to the O^{γ} of both Thr252 and Thr255. This phosphate ion is in the same position as the putative sulfate (or phosphate) ion in TcPCK, which is bound to the main chain and also to O^{γ} of the analogous residue to Thr255 (Trapani *et al.*, 2001). Both ions are in a position corresponding to that of the α-phosphate of bound ATP in EcPCK, in the kinase-1a pocket created by residues 250–255.

In the native structure of AbsPCK there is a sulfate ion present in an analogous position to the phosphate ion in the pyruvate/Mn^{2+} structure. It is hydrogen bonded to the main chain and to O^{γ} of Thr252 and Thr255.

There is also a water molecule present in the native structure that is not present in the structure of the complex in this phosphate/sulfate ion-binding region, forming one hydrogen bond with the sulfate ion.

A putative sodium ion is present in the native structure, bonded to the main-chain O atom of Tyr29, $O^{\delta 1}$ of Asp307 and a water molecule. An identical interaction was not found in any structure in the PDB. However, numerous entries for sodium ions are present in the PDB with a great variety of

interactions, many of them with the $O^{\delta1}$ or $O^{\delta2}$ of an Asp residue.

Even though Mg^{2+} was present during crystallization of the complex, it was not visible in the structure of the complex. This is not unexpected since Mg^{2+} interacts with the nucleotide and none was present.

3.3. Sequence, structural homology and new conserved active-site motif

AbsPCK shares high sequence identity with other ATP-dependent PCKs and low identity with GTP-dependent PCKs (Table 1). The sequence alignment of some representative PCKs with AbsPCK identified the known conserved residues and motifs (Fig. 1). These were confirmed by topology alignments and visualization with modelling software for the PCKs of known structures. The sequence alignment also identified a previously unreported conserved motif comprising residues 280–287 and containing the pattern N*X*E*X*G*X*Y(/F)A(/G), where *X* is any amino-acid residue. This motif is completely conserved in all sequenced PCKs except for two species (*Aeropyrum pernix* from the archaea and *Lactobacillus plantarum*, a bacterium) in which this motif is poorly conserved. In ATP-dependent PCKs this sequence is N*X*EGGCYA. In GTP-dependent PCKs the sequence is N*X*E*X*G*X*FG, except for one species where F is instead Y and some species where the final G is replaced by A.

The motif forms one side of the active-site cleft (Fig. 4), indicating its importance and the reason for its conservation. Its function appears to be structural; it helps to stabilize and position the residues that bind substrates and metal cofactors. The nucleotide-binding region is on the other side of the cleft.

In AbsPCK, Asn280–Gly283 form a random coil connecting two consecutive β-strands, one in the N-terminal domain (residues 1–216 and 283–341) and the other in the C-terminal domain (residues 217–282 and 342–540), thus connecting the two domains. $N^{\delta2}$ of Asn280 hydrogen bonds to a water molecule and to the carbonyl O atom of Ile266. This bond helps position a turn near Ile266 containing the conserved residues Asp268 and Asp269. $O^{\delta1}$ of Asn280 hydrogen bonds to another water and to the backbone N atom of Glu282. These bonds help create and stabilize the backbone turn of this motif. Glu282 $O^{\varepsilon1}$ forms a salt bridge with Lys212 N^{ε}, another conserved residue adjacent to Lys213. Additionally, Glu282 $O^{\varepsilon1}$ hydrogen bonds with the same water molecule as Asn280 $O^{\delta1}$.

Gly284–Ala287 are the beginning of the next β-strand and form hydrogen bonds with the adjacent antiparallel β-strand comprising Arg333–Pro337. Arg333 is a conserved residue that hydrogen bonds with the γ-phosphate group of the nucleotide and with the substrate (OAA, PEP). Gly284 and Ala287 appear to be conserved for steric reasons. There is no space for a larger side chain in these regions of the enzyme. Cys285 is not present in GTP-dependent PCKs. Its role in ATP-dependent PCKs is discussed in §3.4. Tyr286 also appears to be conserved for steric reasons. A planar hydrophobic side chain fits best in the available space and may help position the conserved residue Tyr207. In AbsPCK, the terminal hydroxyl group of Tyr207 hydrogen bonds with pyruvate.

3.4. Reduced cysteines

AbsPCK contains six cysteine residues. Three of these (Cys125, Cys233 and Cys408) are relatively inaccessible to solvent. The other three, located either on the protein surface (Cys74 and Cys175) or near the open active-site cleft (Cys285), were covalently bound to a sulfhydryl reducing agent through a disulfide bond in the structure of the complex. Cys175 was covalently bound to a DTT molecule. The unbonded sulfur on the opposite end of the DTT molecule had been oxidized to a sulfonate, (2*S*,3*S*)-2,3-dihydroxy-4-sulfanylbutane-1-sulfonate. Cys74 and Cys285 (Fig. 5) were covalently bound to 2-mercaptoethanol.

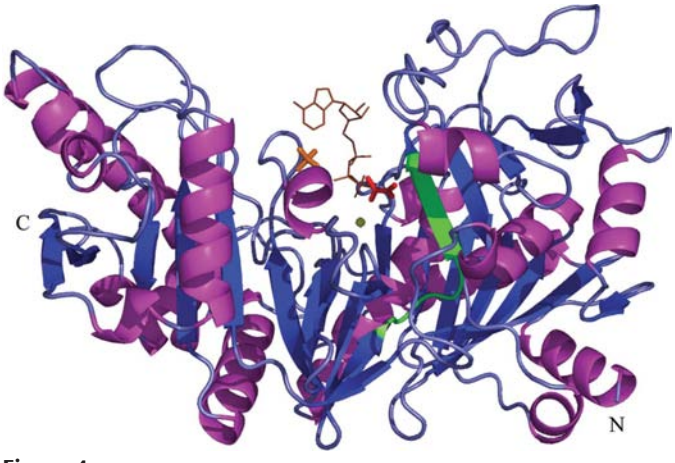

Figure 4
Tertiary structure of AbsPCK showing secondary-structural elements and location of new motif (highlighted in green) in the open active site. ATP (brown) has been modelled into active site to show where it binds in relation to the new motif. Manganese is shown as a yellow sphere, pyruvate is red and phosphate is orange.

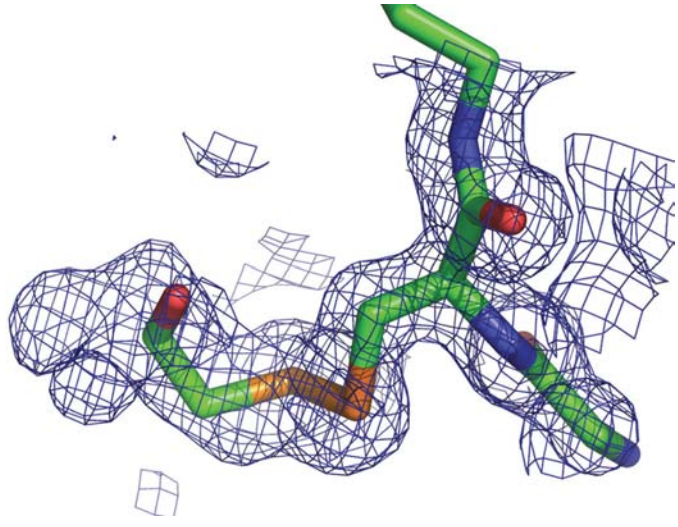

Figure 5
Electron density at 1.0σ around Cys285 and bound 2-mercaptoethanol. The disulfide bond is coloured orange.

Although 2-mercaptoethanol and DTT were present during purification and crystallization, respectively, for both crystals, these molecules were not present in the native structure. In TcPCK, it has been hypothesized (Jurado *et al.*, 1996) that the metal-ion cofactor affects the reactivity of an unidentified cysteine to sulfhydryl modifying reagents. The presence of covalently bound sulfhydryl reducing agents in the structure of AbsPCK with manganese and the absence of them in the structure of the native (no manganese) seems to support this conclusion.

The presence of a reactive cysteine and the effect of sulfhydryl reducing agents on enzyme activity have been studied in several PCKs with little consensus (Bazaes *et al.*, 1993; Jurado *et al.*, 1996; Lewis *et al.*, 1989; Makinen & Nowak, 1989; Rojas *et al.*, 1993). Part of the difficulty is determining which cysteine is responsible for any observed change in activity. The other factor is the variety of sulfhydryl modifying agents used in the studies.

Cysteine residues are not generally conserved between ATP- and GTP-dependent PCKs. Within the GTP class, there is considerable conservation of cysteine residues amongst the sequences, with the cysteine residue located in the kinase-1a loop being the most well conserved, but absent in ATP-dependent PCKs. In the ATP class, Cys285 and Cys408 are almost completely conserved in the known sequences of ATP-dependent PCKs. Both are absent in GTP-dependent PCKs except for the slime mold *Dictyostelium discoideum*.

Cys285 is in the active-site cleft and part of the conserved motif NFEGGCYA (§3.3 and Fig. 4). This motif helps to position residues involved in the binding of the metal-ion cofactors, nucleotide and substrate(s). Cys285 is analogous to Cys258 in TcPCK and may be the putative cysteine residue in TcPCK that is reduced by some sulfhydryl reducing agents, resulting in loss of enzyme activity (Jurado *et al.*, 1996; Trapani *et al.*, 2001). The degree of inactivation depended on the type of agent and metal-ion cofactor and was slowed or prevented in some cases by the presence of ADP. On the other hand, EcPCK was not inactivated by the presence of various sulfhydryl reducing agents (Bazaes *et al.*, 1993). However, the agents tested did not include 2-mercaptoethanol or DTT.

There is limited space near Cys285 or its analogous residue for a sulfhydryl reducing agent to bind, even when the enzyme is in the open conformation (Fig. 6). This may explain why bulkier agents have less of an inhibitory effect on activation of the enzyme. It is possible to model 2-mercaptoethanol into the structures of the closed conformation of EcPCK (PDB code 1aq2) and of TcPCK (PDB code 1ii2) at this residue with distances between atoms greater than 2.5 Å if two water molecules are displaced. Larger molecules such as DTT would not fit in either structure, even in TcPCK where the equivalent Thr259 is replaced with an alanine (Fig. 6*d*).

A comparative analysis of AbsPCK and other PCK structures shows that enzyme inhibition through modification of Cys285 may be a consequence of a change in the hydrogen-bonding network between the two domains of the protein at the active site. This would lead to hindrance of domain closure. The presence of the 2-mercaptoethanol in AbsPCK (Fig. 6*a*) displaces two water molecules that are present in the native AbsPCK structure (Fig. 6*b*), as well as in the structures of TcPCK (Fig. 6*d*) and nucleotide-bound EcPCK (Fig. 6*c*). This may have an impact on domain closure (§3.5).

Table 4
Alignment of the two domains of native AbsPCK with the two domains of other ATP-dependent PCKs.

sc means that the two structures have the same conformation; *i.e.* no dynamic domains were found according to the criteria of the program *DynDom* (Hayward & Berendsen, 1998).

	EcPCK	EcPCK	EcPCK	TcPCK
PDB code	1oen	1aq2	1ayl	1ii2
ATP present	No	Yes	Yes	No
Reported conformation	Open	Closed	Closed	Closed
AbsPCK native (no ATP)				
Fixed N-terminal domain				
Best fit r.m.s.d. main chain (Å)	sc	0.90	1.10	2.31
Movement of the unfixed C-terminal domain necessary to achieve overlap				
Rotation (°)	sc	18.7	18.7	20.4
Translation (Å)	sc	0.2	0.2	0.5

In the structure of EcPCK with bound nucleotide (PDB code 1aq2), Tyr336, a residue which is almost completely conserved in the ATP class of PCKs and which is part of the N-terminal domain, forms a hydrogen bond with one of these water molecules, which in turn forms a hydrogen-bond network with the other water molecule, Thr259 $O^{\gamma 1}$ and the hydroxyl group of Ser258 on the other side of the active-site cleft and part of the C-terminal domain. Thr259 $O^{\gamma 1}$ hydrogen bonds directly or through another water molecule to atoms which form the kinase-1a loop. The effect of all these hydrogen bonds appears to be to bring the two sides of the active-site cleft together (and thus the two domains) at this region and ultimately affects the positioning of substrates. A similar bonding pattern is apparent in the structure of TcPCK, although nucleotide is not present because the enzyme is in a conformation with the domains closed. However, because the equivalent Thr259 is replaced with Ala259, this residue does not contribute as much to the hydrogen-bonding network. Instead, the equivalent Ser258 is present and binds through the two water molecules to the equivalent Tyr336 on the opposite side of the cleft.

In hPCK, Tyr336 is replaced by a phenylalanine residue which has no hydroxyl group. Furthermore, the residue analogous to Cys285 is also a phenylalanine. Therefore, the mechanism by which sulfhydryl reducing agents inhibit activation of GTP-dependent PCKs is expected to be different.

In the structure of the AbsPCK complex, there is no water molecule to bind to Ser258 and the side chain points away from Tyr336 (Fig. 6*a*). The terminal hydroxyl group of 2-mercaptoethanol forms hydrogen bonds with that of Tyr336 and a water molecule. This water molecule is not present in the native structure of either EcPCK or AbsPCK, nor in the structure of TcPCK and nucleotide-bound EcPCK. The water molecule also forms a hydrogen bond with Asp268 $O^{\delta 2}$. It is possible that the terminal hydroxyl group of 2-mercaptoethanol also hydrogen bonds to Asp268 $O^{\delta 2}$. The electron density around the terminal hydroxyl group of 2-mercapto-

ethanol in AbsPCK allows more than one orientation of the terminal group (Fig. 5), which would easily bring it into hydrogen-bonding distance from Asp268 $O^{\delta2}$. Furthermore, unlike in the native structure (Fig. 6*b*), the electron density around the carboxyl group of Asp268 in the structure of the complex is not well defined (Fig. 6*a*). In contrast, in the structures of both EcPCK and TcPCK, $O^{\delta1}$ and $O^{\delta2}$ of Asp268 are available to form hydrogen bonds directly or through water molecules to help position the kinase-1a loop and bring together the two domains of the protein. However, in the structure of the AbsPCK complex, these potential bonds are disrupted by the bound 2-mercaptoethanol which can bind to Asp268 and which also displaces critical water molecules. This contributes to prevent domain closure and substrate binding.

3.5. Domain movement

Independent superposition of all atoms of the N- and C-terminal domains of the native and the pyruvate/Mn^{2+} structures of AbsPCK gave a positional r.m.s. displacement of 0.75 Å for both domains. This exercise also revealed a 2.1° rotation difference between the two structures, with the complex counterintuitively slightly more open than the native.

Both structures of AbsPCK were observed to be in an open conformation when compared with the open and closed conformations of EcPCK and with the closed conformation of TcPCK (Table 4). Analogously to what was seen for EcPCK (Tari *et al.*, 1996), a rotation of ∼20° of one of the AbsPCK domains relative to the other is necessary to align the AbsPCK structures with the closed conformation of TcPCK and EcPCK.

An open conformation would be expected to occur with the native structure since no substrates were present and it is thought that binding of nucleotide in PCK closes the structure (Matte *et al.*, 1997, 1998; Tari *et al.*, 1996). A comparison of the native AbsPCK structure with the structure of nucleotide-free TcPCK (Table 4) suggests that TcPCK, paradoxically, is in the closed conformation. This was also the conclusion regarding

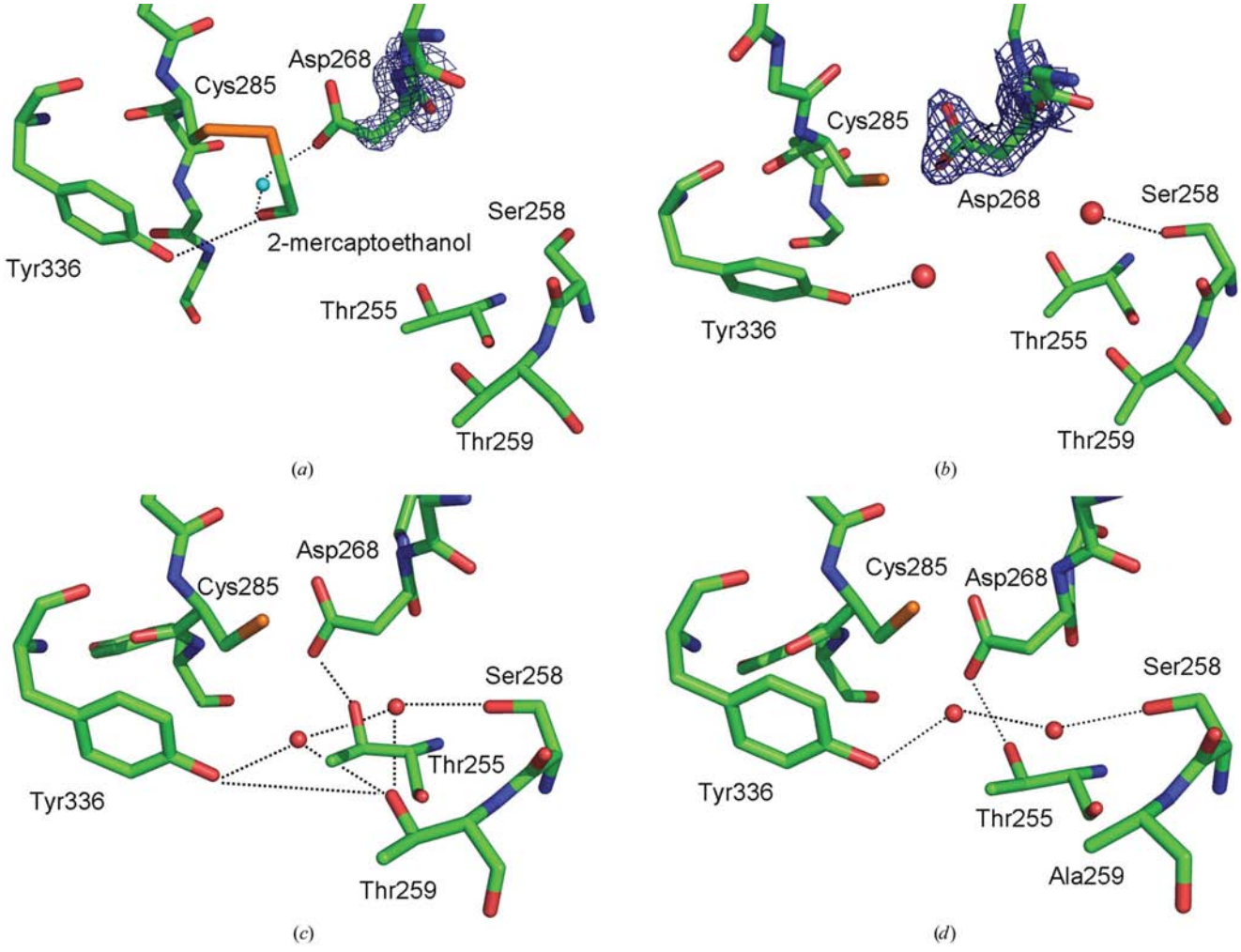

Figure 6
Hydrogen-bonding network of a portion of the active-site cleft in the structures of (*a*) AbsPCK complex (PDB code 1ylh), (*b*) native AbsPCK (PDB code 1ygg), (*c*) EcPCK (PDB code 1aq2) and (*d*) TcPCK (PDB code 1ii2). Numbering is that of EcPCK. Electron density around Asp268 for the AbsPCK structures is contoured at 2σ. Hydrogen bonds are represented by dashed lines. Water molecules are small spheres.

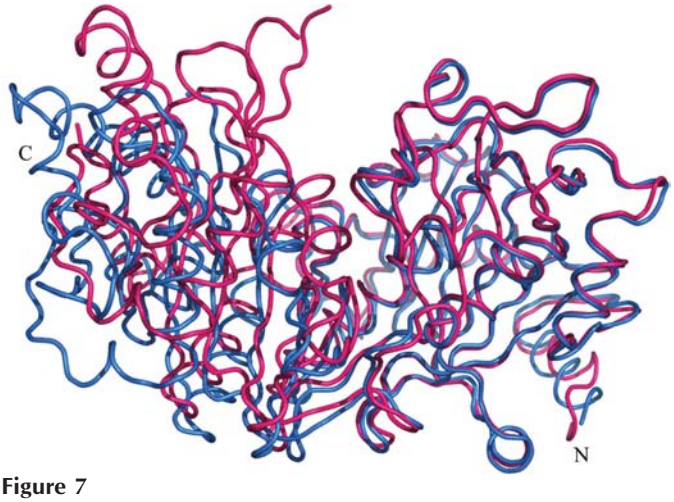

Figure 7
Overlay of the structure of the AbsPCK complex (PDB code 1ylh; blue) with the closed conformation of the structure of EcPCK (PDB code 1aq2; pink), indicating the former's open conformation. The N-terminal domains were aligned with the use of *LSQKAB* (Kabsch, 1976).

TcPCK, which showed no domain movement when compared with the nucleotide-bound structure of EcPCK (Trapani *et al.*, 2001). It was hypothesized that either crystal packing or the presence of the putative sulfate (or phosphate) ion at the ATP α-phosphate position in the TcPCK structure caused domain closure. The authors noted that TcPCK in a sulfate-free solution seemed to be in the open conformation as measured by small-angle X-ray scattering. However, in the native structure of AbsPCK there is a sulfate ion in the equivalent position, yet it is in an open conformation relative to nucleotide-bound EcPCK and the TcPCK structure.

The N- and C-terminal domains of the native AbsPCK structure and the closed conformation of the EcPCK structure (PDB code 1aq2) were independently superposed, excluding the residues of the interdomain connection 280–283 and several residues to either side from the superposition. The resulting superposed models were visually inspected in three dimensions in the region of the new conserved motif, residues 280–287. It appeared that a fulcrum point about which the two domains bend is located near Asn280.

As with the structure of the native, the structure of the complex was also observed to be in an open conformation relative to the open and closed conformations of EcPCK (Fig. 7) and in an open conformation relative to TcPCK, with similar amounts of rotation and translation necessary to align the unfixed domain. These observations were confirmed with *LSQKAB*. The structure of the complex also included a phosphate ion in the ATP α-phosphate position. If the presence of a phosphate or sulfate ion was essential for domain closure as mentioned above, one would also expect this structure to have a closed conformation. However, it can also be hypothesized that closure has been prevented by the presence of 2-mercaptoethanol bound to Cys285.

LTJD is a Canada Research Chair in Structural Biochemistry. This research was funded by operating grant ROP-10162 of the Canadian Institutes of Health Research–Saskatchewan Regional Partnership Program to LTJD. We thank Guy Macha for data collection on the BioCARS beamline at the Advanced Photon Source, Argonne, Illinois. Use of the Advanced Photon Source was supported by the US Department of Energy, Basic Energy Sciences, Office of Science under contract No. W-31-109-Eng-38. Use of the BioCARS Sector 14 was supported by the National Institutes of Health, National Center for Research Resources under grant No. RR07707.

References

Bazaes, S., Silva, R., Goldie, H., Cardemil, E. & Jabalquinto, A. M. (1993). *J. Protein Chem.* **12**, 571–577.
Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. & Thompson, J. D. (2003). *Nucleic Acids Res.* **31**, 3497–3500.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *PyMOL.* DeLano Scientific, San Carlos, CA, USA.
Dunten, P., Belunis, C., Crowther, R., Hollfelder, K., Kammlott, U., Levin, W., Michel, H., Ramsey, G. B., Swain, A., Weber, D. & Wertheimer, S. J. (2002). *J. Mol. Biol.* **316**, 257–264.
Guettler, M. V., Rumler, D. & Jain, M. K. (1999). *Int. J. Syst. Bacteriol.* **49**, 207–216.
Harding, M. (1999). *Acta Cryst.* D**55**, 1432–1443.
Harding, M. (2004). *Acta Cryst.* D**60**, 849–859.
Hayward, S. & Berendsen, H. J. (1998). *Proteins*, **30**, 144–154.
Hebda, C. A. & Nowak, T. (1982). *J. Biol. Chem.* **257**, 5515–5522.
Jurado, L. A., Machin, I. & Urbina, J. A. (1996). *Biochim. Biophys. Acta*, **1292**, 188–196.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Kim, P., Laivenieks, M., McKinlay, J., Vieille, C. & Zeikus, J. G. (2004). *Plasmid*, **51**, 108–115.
Laivenieks, M., Vieille, C. & Zeikus, J. G. (1997). *Appl. Environ. Microbiol.* **63**, 2273–2280.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Lee, S. Y., Hong, S. H., Lee, S. H. & Park, S. J. (2004). *Macromol. Biosci.* **4**, 157–164.
Lewis, C. T., Seyer, J. M. & Carlson, G. M. (1989). *J. Biol. Chem.* **264**, 27–33.
Lu, G. (1996). *Protein Data Bank Quart. Newsl.* **78**, 10–11.
Makinen, A. L. & Nowak, T. (1989). *J. Biol. Chem.* **264**, 12148–12157.
Matte, A., Goldie, H., Sweet, R. M. & Delbaere, L. T. (1996). *J. Mol. Biol.* **256**, 126–143.
Matte, A., Tari, L. W. & Delbaere, L. T. (1998). *Structure*, **6**, 413–419.
Matte, A., Tari, L. W., Goldie, H. & Delbaere, L. T. (1997). *J. Biol. Chem.* **272**, 8105–8108.
Matthews, B. (1968). *J. Mol. Biol.* **33**, 491–497.
Miller, R. S., Mildvan, A. S., Chang, H. C., Easterday, R. L., Maruyama, H. & Lane, M. D. (1968). *J. Biol. Chem.* **243**, 6030–6040.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). *Acta Cryst.* D**59**, 1131–1137.
Potterton, E., McNicholas, S., Krissinel, E., Cowtan, K. & Noble, M. (2002). *Acta Cryst.* D**58**, 1955–1957.

Rojas, M. C., Encinas, M. V., Kemp, R. G., Latshaw, S. P. & Cardemil, E. (1993). *Biochim. Biophys. Acta*, **1164**, 143–151.
Roussel, A. & Cambillau, C. (1991). *Silicon Graphics Geometry Partners Directory*, p. 81. Silicon Graphics, Mountain View, CA, USA.
Tari, L. W., Matte, A., Goldie, H. & Delbaere, L. T. (1997). *Nature Struct. Biol.* **4**, 990–994.
Tari, L. W., Matte, A., Pugazhenthi, U., Goldie, H. & Delbaere, L. T. (1996). *Nature Struct. Biol.* **3**, 355–363.
Tatusova, T. A. & Madden, T. L. (1999). *FEMS Microbiol. Lett.* **174**, 247–250.
Trapani, S., Linss, J., Goldenberg, S., Fischer, H., Craievich, A. F. & Oliva, G. (2001). *J. Mol. Biol.* **313**, 1059–1072.
Utter, M. F. & Kolenbrander, H. M. (1972). *The Enzymes* Vol. 6, 3rd ed., edited by P. D. Boyer, pp. 117–168. New York: Academic Press.
Vriend, G. (1990). *J. Mol. Graph.* **8**, 52–56.
Winn, M. D. (2003). *J. Synchrotron Rad.* **10**, 23–25.
Zeikus, J. G., Jain, M. K. & Elankovan, P. (1999). *Appl. Microbiol. Biotechnol.* **51**, 545–552.