Crystal structure of *Escherichia coli* pyruvate kinase type I: molecular basis of the allosteric transition

Andrea Mattevi¹*, Giovanna Valentini², Menico Rizzi¹, M Luisa Speranza², Martino Bolognesi¹,³ and Alessandro Coda¹

¹Department of Genetics and Microbiology, University of Pavia, via Abbiategrasso 207, 27100 Pavia, Italy, ²Department of Biochemistry, University of Pavia, via Taramelli 3b, 27100 Pavia, Italy and ³Department of Physics and IST Advanced Biotechnology Centre, University of Genova, viale Benedetto XV 10, 16132 Genova, Italy

**Background:** Pyruvate kinase (PK) plays a major role in the regulation of glycolysis. Its catalytic activity is controlled by the substrate phosphoenolpyruvate and by one or more allosteric effectors. The crystal structures of the non-allosteric PKs from cat and rabbit muscle are known. We have determined the three-dimensional structure of the allosteric type I PK from *Escherichia coli*, in order to study the mechanism of allosteric regulation.

**Results:** The 2.5 Å resolution crystal structure of the unligated type I PK in the inactive T-state shows that each subunit of the homotetrameric enzyme comprises a (β/α)₈-barrel domain, a flexible β-barrel domain and a C-terminal domain. The allosteric and active sites are located at the domain interfaces. Comparison of the T-state *E. coli* PK with the non-allosteric muscle enzyme, which is thought to adopt a conformation similar to the active R-state, reveals differences in the orientations of the β-barrel and C-terminal domains of each subunit, which are rotated by 17° and 15°, respectively. Moreover, the relative orientation of the four subunits differs by about 16° in the two enzymes. Highly conserved residues at the subunit interfaces couple these movements to conformational changes in the substrate and allosteric effector binding sites. The subunit rotations observed in the T-state PK induce a shift in loop 6 of the (β/α)₈-barrel domain, leading to a distortion of the phosphoenolpyruvate-binding site accounting for the low substrate affinity of the T-state enzyme.

**Conclusions:** Our results suggest that allosteric control of PK is accomplished through remarkable domain and subunit rotations. On transition from the T- to the R-state all 12 domains of the functional tetramer modify their relative orientations. These concerted motions are the molecular basis of the coupling between the active centre and the allosteric site.

**Structure** 15 July 1995, 3:729-741

Key words: active R-state, allosteric regulation, (β/α)₈-barrel, domain rotations, inactive T-state

**Introduction**

Pyruvate kinase (PK; EC 2.7.1.40) catalyzes the conversion of phosphoenolpyruvate (PEP) to pyruvate, coupled to the synthesis of one molecule of ATP:

\[
\text{PEP} + \text{Mg-ADP} + \text{H}^+ \rightarrow \text{Mg-ATP} + \text{pyruvate}
\]

The reaction is the last step in the glycolytic pathway and is irreversible under physiological conditions. The enzyme requires one equivalent of monovalent cations, normally K⁺ [1], and two equivalents of bivalent cations, usually Mg²⁺ or Mn²⁺ [2] for its activity. PK plays a central role in cellular metabolism. The product of the reaction, pyruvate, is involved in a number of metabolic pathways. Moreover, together with phosphofructokinase, PK is thought to be the major regulatory enzyme of glycolysis [3]. PK has been characterized from a number of prokaryotes and eukaryotes [4], and in nearly all organisms it shows allosteric properties in binding the substrate PEP. Furthermore, the enzymatic activity is heterotropically regulated by one or more allosteric effectors, whose precise chemical nature depends on the type of organism or tissue. The allosteric regulation is accomplished through the oligomeric organization of the enzyme, which is usually a tetramer of four identical subunits with an approximate molecular weight of 4×50000 Da [4].

Two forms of pyruvate kinase have been isolated from *Escherichia coli*. Both are homotropically activated by the substrate PEP, but differ in the nature of their allosteric effectors. The type I isoenzyme is heterotropically activated by fructose-1,6-bisphosphate (FBP) and inhibited by ATP [5], whereas the type II isoenzyme is activated by AMP and monophosphorylated sugars (i.e. ribose-phosphate). Most bacterial species have only one PK, kinetically similar to either the *E. coli* type I or type II isoenzymes [4]. In eukaryotes, PK normally exhibits cooperative binding of PEP and allosteric activation by FBP, and is therefore more similar to the type I than the type II isoenzyme of *E. coli*. An exception to this pattern is the mammalian muscle (M1) protein, which is the only characterized PK showing hyperbolic Michaelis–Menten kinetics and no cooperative properties [6]. The other three mammalian isoenzymes (M2, L and R) exhibit sigmoidal kinetics. The liver (L) isoenzyme, in particular, is regulated not only by the allosteric effectors but also by phosphorylation of a serine residue at the N terminus of the protein [3].

*Corresponding author.
The molecular basis of the allosteric regulation of PK is unknown. To date, only the crystal structures of the non-allosteric M1 PK from cat [6] and rabbit [7] muscle have been determined, at 2.6 A and 2.9 A resolution respectively. The M1 isoenzyme lacks sigmoidal kinetics and its structure most likely adopts an active R-type conformation [8]. This hypothesis is supported by kinetic and fluorescence data indicating that the M1 isoenzyme is in an active R-like state, whereas the M2 isoenzyme, which is identical to the M1 PK except for a stretch of 56 amino acids [9], is able to adopt alternate, active and inactive, conformations [10].

With the aim of understanding the structural basis for the mechanisms of allosteric transition and cooperative substrate binding in PK we have undertaken the X-ray analysis of the type I enzyme from E. coli. This protein is well characterized in its kinetic and regulatory properties [5,11,12]. The enzyme is a tetramer of four identical subunits, each consisting of 470 residues with an amino acid sequence identity of 49% with respect to both cat and rabbit M1 isoenzymes [13]. Type I PK exhibits typical sigmoidal kinetics with a Hill coefficient of 3.15. The unligated protein is in the inactive T-state and is activated upon binding of either the substrate, PEP, or the allosteric effector, FBP. The analysis of the enzymatic reaction has shown that the allosteric regulation can be described according to the sequential model of Monod et al. [14]. Here we report the structure of the unligated protein in the inactive T-state determined by the method of multiple isomorphous replacement (MIR). The structure is compared with that of the M1 R-like isoenzyme, revealing a striking variation in the conformation of the two proteins that allows us to propose a mechanism for the allosteric control of PK.

Results and discussion

The structure of the unligated type I PK from E. coli was solved by MIR, followed by phase improvement using solvent flattening and density averaging. Attempts to solve the structure by molecular replacement using the cat M1 PK model [6] failed. The crystallographic refinement was carried out at 2.5 A resolution using data collected at 100 K from a crystal belonging to space group P2_12_1 with an entire tetramer in the asymmetric unit. During refinement it became evident that the published amino acid sequence [13] did not agree with the observed electron density. In particular, the electron density indicated the presence of a C-terminal extension of eight amino acids, corresponding to a polypeptide chain of 470 rather than 462 residues.

In the current model all amino acids are in well defined electron density (Fig. 1), except for residues 346–351 in all four subunits and residues 75–105 in subunit 3. A total of 434 ordered water molecules have been placed in the asymmetric unit; they are all engaged in at least one hydrogen bond with a protein or solvent atom. The crystallographic R-factor is 20.5% for 72,305 unique reflections in the 6.0–2.5 A resolution range (99.9% complete, no sigma cutoff); the free R-factor (calculated using 4% of the reflections that have not been included in the refinement) is 30.7%. The model has good geometry with a root mean square (rms) deviation from ideality of 0.016 A for bond distances and 1.6° for bond angles. The four crystallographically independent subunits are very similar except that one domain is differently oriented in two of the four chains (see below). Subunit 4 has the lowest average B-factor and will be used as a reference subunit for the analysis of the structure. The nomenclature proposed by Muirhead et al. [6] for identification of the secondary-structure elements has been adopted to describe E. coli type I PK.

Overall conformation

A single subunit of E. coli type I PK is shown schematically in Figure 2. The subunit can be divided into three domains, called A, B and C. Domain A (residues 1–70 and 171–345) is the largest of the three and has a typical (B/4)-barrel topology. The barrel is characterized by three additional alpha-helical segments located on loops 6, 7 and 8 on the C-terminal side of the eight-stranded parallel beta-sheet. These helices are named Aalpha6’, Aalpha7’, Aalpha8’ (Fig. 3) and, as described below, have a central role in catalysis and allosteric regulation. Domain B (residues 71–170), adjacent to the C-terminal side of domain A and inserted on loop 3, forms a lid covering the (B/4)-barrel. Domain B consists of a mixed beta-barrel with only one short alpha-helix. It interacts very loosely with...
The allosteric regulation of the enzymatic activity [15].

The PK domains play a key role in the overall conformation of the individual B domains, as indicated by rms deviations of 0.39 Å and 0.34 Å between the Cα atoms of the B domains in subunits 1 and 4 and 3 and 4, respectively. As described below, relative motions between the PK domains play a key role in the allosteric regulation of the enzymatic activity [15].

Superposition of the crystallographically independent PK subunits indicates that domains A and C have very similar conformational variations associated with protein regions having a direct functional role. Herzberg and Moult [16] that strain is frequently associated with residues belonging to helices A6, A7 and A8 (Fig. 6). There are 11 hydrogen bonds, 2 of which involve backbone atoms, whereas the remaining 9 involve only side chains. The active centre is built up by residues belonging to the three longest loops (loops 6, 7 and 8) of the (β/α)8-barrel (Fig. 2). Interestingly, Ser312, the only amino acid residue that has dihedral angles outside the allowed regions of the Ramachandran plot in all four crystallographically independent subunits, is positioned on loop 8 and is part of the active site (Fig. 1). This observation is in agreement with the proposal of Herzberg and Moult [16] that strain is frequently associated with protein regions having a direct functional role.

The tetramer

The four subunits are associated in a tetramer with 222 symmetry (Fig. 5a). The three twofold axes are labelled p, q and r. The 222 symmetry is exact only for the A and C domains, whereas the B domains of subunits 1 and 3 are not related to each other by exact twofold axes because of their different orientations. The inter-subunit contacts involve residues of the A and C domains, whereas the B domains do not take part in any inter-subunit interaction. Two types of inter-subunit interfaces can be distinguished; the first type involves residues of the A domains related by the r molecular axis. The second type only involves amino acids of the C domains related by the q axis. The two contact areas are of approximately equal size. Upon formation of the tetramer, 1062 Å² (6%) and 904 Å² (5%) of the solvent accessible surface area of the monomer are buried at the r and q interfaces, respectively.

The intermolecular contacts along the r interface are mostly contributed by residues belonging to helices A6, A7 and A8 (Fig. 6). There are 11 hydrogen bonds, 2 of which involve backbone atoms, whereas the remaining 9 involve only side chains. At the q interface, a predominance of main chain–main chain interactions is observed: out of 10 intermolecular hydrogen bonds, 8 involve backbone atoms. The most relevant structural elements building this interface are strands Cβ5 of the two opposite (q-related) subunits which, by running antiparallel to each other, extend the central β-sheet of the C domains to generate a 10-stranded intermolecular β-sheet (Fig. 5a).

The active centre

By analogy with the M1 structures [6,7], the active site in each subunit can be located in the crevice between the A and B domains (Fig. 2). The catalytic centre is built up by residues belonging to the three longest loops (loops 6, 7 and 8) of the (β/α)8-barrel (Fig. 2). Interestingly, Ser312, the only amino acid residue that has dihedral angles outside the allowed regions of the Ramachandran plot in all four crystallographically independent subunits, is positioned on loop 8 and is part of the active site (Fig. 1). This observation is in agreement with the proposal of Herzberg and Moult [16] that strain is frequently associated with protein regions having a direct functional role.

PK requires both divalent and monovalent cations for its activity [1,2]. Despite the fact that the crystals of the unligated type I PK were grown in the presence of 10 mM Mg2+ and 10 mM K+, no electron density could be attributed to either of these two cations in the active-site region. Moreover, no significant difference electron density peaks could be observed after soaking crystals in Mn2+- or Tl+-containing solutions (data not shown). On the other hand, the most powerful heavy-atom derivatives used for the structure determination were obtained by soaking the crystals in Gd3+ and Lu3+, which both bind in the catalytic centre and are coordinated by the side-chain carboxylates of Asp246 and Glu222. This binding site is identical to that for Mn2+ observed in the rabbit M1 PK–pyruvate complex [7]. Thus, in the T-state, type I
Fig. 3. Structure-based alignment of the *E. coli* type I and the rabbit M1 (RabM1) sequences. Residues belonging to α-helices and β-strands are indicated by the letters 'h' and 'e', respectively. The secondary structures of domains A and C are named according to (6). Helix Cal is present only in the M1 structure, whereas the corresponding residues of the *E. coli* enzyme are disordered and not visible in the electron-density map. Vertical bars identify the domain boundaries. Amino acid identities are shown by shading.

Fig. 4. Stereoview of superposed Ca tracings of the four crystallographically independent *E. coli* PK subunits. Subunit 1 is shown in red, subunit 2 in blue, subunit 3 in green and subunit 4 in yellow. The picture was obtained by superposing the Ca atoms of the A and C domains (residues 1–70 and 171–470) of the four monomers. The different orientation of the B domains in subunits 1 and 3 is evident (Table 2). The orientation is the same as in Figure 2.
Fig. 5. (a) The tetrameric *E. coli* type I PK in the T-state viewed along the p axis away from the centre of the molecule. Domains A, B and C of subunit 4 are shown in green, blue and red, respectively. (b) Structure of the rabbit M1 PK [7] viewed along the p axis. In subunit 4, the N-terminal additional domain (see text) is shown in yellow. (c) Schematic comparison of the T-state (left) and R-like (right) tetramers outlining the set of motions affecting the domain and subunit orientations (see text). The subunits are labelled 1, 2, 3 and 4. In the active M1 enzyme (right) the B domain adopts the ‘closed’ conformation narrowing the active-site cleft. Due to the ‘open’ conformation, in the T-state the B domains interact very loosely with their own subunit and do not take part in the inter-subunit interactions. The effector binding site at the interface between domains A and C is wider in the R-like structure as a consequence of the domain rotation. (Parts (a) and (b) drawn with MOLSCRIPT [45].)
PK seems to be incapable of binding divalent cations but is, however, able to bind trivalent cations. In this context, it should be noted that the enzyme retains its catalytic activity at the pH of the crystallization medium (pH 6.2).

Comparison with the M1 R-like pyruvate kinase

As mentioned above, the three-dimensional structures of the M1 PK from cat and rabbit muscle have both been determined [6,7]. It has been shown that the M1 isoenzyme is stabilized in an R-like conformation which confers activity but abolishes the sigmoidal kinetics typical of the allosterically regulated proteins [8,10]. Comparison of the structures of E. coli type I PK (in the T-state) with the M1 isoenzyme (R-like) can therefore offer the first clues about the mechanism of allosteric regulation in this enzyme. The rabbit M1 PK structure was used for this analysis. However, the major conclusions drawn from the comparison would be the same if the cat M1 PK structure were considered. Residues of M1 PK mentioned in the following discussion have been numbered according to the amino acid sequence of E. coli type I PK.

Type I PK from E. coli shares 49% sequence identity with the rabbit M1 protein (Fig. 3). The identity is higher for domain A (58%) than for domains B (34%) and C (35%) (Table 1). Moreover, the M1 protein differs from the E. coli enzyme in having a 42-residue N-terminal extension which forms a small additional helical region [6,7]. The absence of this extension in the structure of E. coli PK generates the hole visible along the p axis of the tetramer (compare Fig. 5a,b). The N-terminal extension is lacking in all bacterial PK enzymes, whereas it is characteristic not only of the M1 enzyme but also of all the eukaryotic allosteric PKs [4]. Therefore, the presence of this N-terminal domain does not seem to relate to the lack of allostery in the M1 enzyme.

**Table 1. Structural comparison between the individual domains of rabbit M1 and E. coli T-state pyruvate kinases.**

<table>
<thead>
<tr>
<th>Structural unit</th>
<th>No. of Ca atom pairs*</th>
<th>Rms deviation (Å)</th>
<th>Sequence identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain A</td>
<td>220</td>
<td>1.1</td>
<td>58</td>
</tr>
<tr>
<td>Domain B</td>
<td>92</td>
<td>0.9</td>
<td>34</td>
</tr>
<tr>
<td>Domain C⁺</td>
<td>110 (95)</td>
<td>1.8 (1.2)</td>
<td>35</td>
</tr>
<tr>
<td>Monomer</td>
<td>422</td>
<td>3.6</td>
<td>49</td>
</tr>
<tr>
<td>q dimer</td>
<td>844</td>
<td>5.8</td>
<td>49</td>
</tr>
<tr>
<td>r dimer</td>
<td>844</td>
<td>4.8</td>
<td>49</td>
</tr>
<tr>
<td>Tetramer</td>
<td>1688</td>
<td>8.7</td>
<td>49</td>
</tr>
</tbody>
</table>

*All residues except for those belonging to loop regions (see Fig. 3) characterized by insertions or deletions were used in the superpositions. The rms deviation obtained omitting helix Ca5 from the superposition is given in parentheses (see text).

**Domain rotations**

Comparison of the individual domains of E. coli and M1 PK reveals a strong structural conservation. After optimal superposition of the A domains the rms deviation for 220 Ca atom pairs is 1.1 Å. A high degree of structural similarity is also shown by the B domains (rms deviation of 0.9 Å for 92 Ca atom pairs of the two enzymes) despite the lower sequence identity (Table 1). The superposition of the C domains, on the other hand, yields an rms deviation of 1.8 Å for 110 Ca pairs. This relatively large difference is mostly accounted for by a shift of the Ca5 helix, which is located after a six-residue insertion in the sequence of the M1 protein (Fig. 3). When the residues belonging to Ca5 are omitted from the superposition the rms deviation decreases to 1.2 Å for 95 Ca pairs.

The analysis of the tertiary structure organization of the E. coli and M1 subunits reveals large variations in the orientation of their domains. Taking domain A as a reference, domain B appears to be rotated by 17.2° in the M1 PK relative to the E. coli enzyme (Table 2; Fig. 7). This hinge-bending movement has the pivot point on the domain linkers (Gly70 and Lue169) and causes a narrowing of the active-site cleft of the M1 protein. There is only one direct hydrogen bond between domains A and B in the ‘open’ T-state of E. coli PK and the B domains do not interact with any other subunit. By contrast, in the ‘closed’ R-like conformation of the M1 enzyme the B domain interacts with the A domains from its own subunit as well as from the r-related monomer (Fig. 8b). The ‘closed’ conformation of the R-state is probably required for proper binding of the ADP substrate [6]. Interestingly, this movement is reminiscent of the ‘open’ and ‘closed’ conformations observed in protein kinases [17] which, although completely unrelated to PK, also catalyze a phosphotransfer reaction.

After superposition of the A domains, domain C of the M1 structure is rotated by 14.9° relative to the T-state subunit (Table 2). The rotation has the effect of narrowing the binding pocket for the allosteric activator, FBP, in the T-state enzyme (Fig. 7).
have shown that Lys382 on domain C of E. coli PK is involved in the binding of this allosteric effector [11,12]. The closest residue to Lys382 across the cleft on domain A is Arg271, in both T- and R-like states, and the Ca-Ca distance between them widens from 7.3 Å to 10.9 Å in the more open R-like conformation. In the absence of structural information for the complex between PK and the allosteric effector, it is not possible to predict the effects of FBP binding on the enzyme conformation. The crystals of the unligated type I PK crack immediately when soaked in a solution containing FBP. However, it seems plausible that binding of the activator could affect the relative orientation of the domains, widening the binding pocket and stabilizing a conformation similar to that observed in the active M1 subunit.

Changes at the subunit interfaces
The comparison between the structures of the T-state E. coli PK and the R-like M1 enzyme reveals large and complex changes in the quaternary structure of the two proteins (Fig. 5; Table 2). The magnitude of these variations is indicated by the fact that, after optimal superposition of the two entire tetramers, the rms deviation is 8.7 Å for 422 Ca atom pairs. For a detailed analysis of the reorientations of domains A and C at the q and r interfaces, the M1 and E. coli PK tetramers were compared as follows: firstly, the centres of mass of the two tetramers were translated to make them coincide with the coordinate origin; secondly, the molecular p, q and r axes were aligned with the Cartesian x, y and z axes, respectively. Structural unit II is the A domain of the r-related subunit. Structural unit II is the C domain of the q-related subunit.

The rotation of domain C occurs about an axis located close to Phe345, on the segment linking domains C and A. The linker is partly disordered in E. coli PK, whereas it adopts a helical conformation in the muscle enzyme (helix Cα1; Fig. 3). It is noteworthy that these residues are part of the only region of the polypeptide chain exhibiting sequence differences between the non-allosteric M1 PK and the allosteric M2 protein [9]. Thus, these residues may confer a reduced flexibility on the M1 enzyme while allowing relative domain movements in the allosteric PKs.

Fig. 7. Comparison between the PK subunit in the T-state (black) and M1 protein (red). The picture has been obtained by superposing the Ca atoms of the A domain of subunit 4 of the M1 PK onto subunit 4 of the E. coli enzyme. The B and C domains are differently oriented by 17.2° and 14.9°, respectively. The orientation is the same as in Figure 2.
Fig. 8. (a) Close-up view of the r interface in the E. coli T-state enzyme. The orientation is approximately the same as in Figure 6. Secondary-structure elements and residues belonging to the r-related subunit (subunit 2) are labelled (r). For the sake of clarity, the side chain of Met279 is not shown. (b) Close-up view of the r interface of the M1 protein. The orientation is the same as in (a). (c) Superposition of the PEP-binding sites of the T-state and R-like structures. The shaded bonds connect the T-state atoms while the white bonds are used for the M1 structure. Loop 6 of the T-state model is shown in green. The orientation is the same as in (a) and (b). The picture has been obtained by superposing the A domain of the M1 subunit onto the same domain of the E. coli subunit 4. (d) Schematic drawing showing the concerted movements at the subunit interface and active site occurring in the transition from the T-state to the R-like state. The helices belonging to the opposite subunits are shown by differential shading. In the T-state, Asp297 makes a salt bridge with Arg292(r) located on the A domain of the r-related subunit. Upon subunit rotation, helix Aα7(r) is moved towards the opposite chain establishing interaction with Asp297 via its N-terminal backbone nitrogen. Meanwhile, Arg292(r) changes its conformation and interacts with Asp127. The latter is located on the B domain, which, in the R-like structure, adopts the ‘closed’ conformation. The restructuring of the subunit interface forces helix Aα6’ on loop 6 to move towards the inner part of the active site in order to take part in the binding of phosphoenolpyruvate.

The r interface is composed of residues belonging to the A domains of two opposite subunits. When compared with the T-state structure, the two r-related A domains in the M1 enzyme differ in their relative orientation by 18.9° (Table 2). This rotation occurs about an axis which is perpendicular to the r axis and affects the subunit interface in the region close to loop 6. Here, a major role is played by Arg292, located at the N terminus of helix Aα7. In the T-state, this side chain is directed towards the inner part of the interface forming a salt bridge with Asp297 of an r-related monomer (Fig. 8a,d). In the M1 structure, the conformation of Arg292 is completely different (Fig. 8b); it points towards the surface of the molecule to make a salt bridge with Asp127, which belongs to domain B of the r-related subunit. The conformational change of Arg292 is the result of a rotation about the χ1 and χ2 angles, which vary by 95° and 101°, respectively.

A remarkable restructuring of the r interface is observed upon transition from the T-state to the R-like state.
Due to the subunit rotation, in the R-like state the N terminus of helix Aα7 is moved towards the r-related subunit, close to the site occupied by Arg292 in the T-state (Fig. 8a,b,d). In this way, the positive charge of the arginine is replaced by the N terminus of Aα7, which bears the partial positive charge associated with the helix dipole [18]. Furthermore, no hydrogen-bond donors or acceptors are left unsatisfied. Whereas in the T-state Asp297 and the backbone oxygen of Met279 (on loop 7) receive an inter-subunit hydrogen bond from Arg292 (Fig. 8a), in the R-like structure they interact with the main-chain nitrogen of Ala293, located at the N terminus of Aα7 on the r-related subunit (Fig. 8b).

The reorganization of the r interface is coupled in a concerted manner to the rotation of domain B occurring within the individual subunits. In the R-state, Arg292 swings round and anchors domain B of the r-related monomer in the 'closed' conformation by means of the inter-subunit salt bridge with Asp127 (Fig. 8b). Conformational changes of arginine residues have also been observed in the allosteric and effector sites of glycogen phosphorylase [19] and lactate dehydrogenase [20]. PK exploits the flexibility of the arginine side chain to generate different types of inter-subunit interactions that stabilize the quaternary structure of the active and inactive forms.

The residues taking part in the intermolecular contacts at the r interface are highly conserved among PK sequences [4]. Very high sequence identity (76%) exists between the E. coli and the M1 proteins in the polypeptide segment comprising strand Aβ6 to helix Aα8 (residues 241-330; Fig. 3). Furthermore, nine of the twelve residues that constitute the Aα7 helix (including Arg292) are strictly conserved in all 16 known PK sequences [4]. Strict conservation is also observed for Asp127, underlining the functional role played by this residue in the allosteric transition.

The C domains of two opposite subunits interact with each other along the g-axis interface. Comparison of the M1 and E. coli tetrameric PK structures shows that the C domains of the two g-related subunits differ in their relative orientation by 29.4° (Table 2). Such domain reorientation occurs about an axis which lies in between, and parallel to, the two Cβ5 strands forming the intermolecular β-sheet located at the heart of the g interface (Fig. 5a,b). In this way, the interactions between the two strands are preserved in the R-like and T-state tetramers despite the large subunit rotation.

The PEP-binding site

The three-dimensional structure of the rabbit M1 pyruvate kinase in complex with the reaction product, pyruvate [7], identifies the residues involved in substrate binding and the phosphotransfer reaction. These amino acids are all located on domain A, at the C-terminal side of the parallel β-sheet of the barrel. Comparison between the M1 enzyme and the T-state PK structure shows that all the catalytically relevant amino acids display a highly conserved conformation, with the exception of loop 6. Compared with the T-state, in the R-like PK structure the short helix, Aα6' (residues 245–249), located in loop 6, performs a sliding movement towards the inner part of the active-site cleft. The shift, of some 1.6 Å, occurs almost exactly along the helix axis (Fig. 8c). The complex of M1 PK with pyruvate and Mn2+ [7] shows that Aα6' plays a key role in binding the cooperative PEP substrate. The pyruvate carboxylate oxygen atoms make two hydrogen bonds with the main-chain nitrogen atoms of Gly245 and Asp246, located at the N terminus of this helix. Moreover, the side chain of Asp246 plays a crucial role by coordinating the catalytically essential divalent cation (Fig. 8c).

In the T-state, the sliding of helix Aα6' displaces Gly245 and Asp246 by 1.9 Å, distorting the PEP-binding site. The distance between Nε of Lys220, which is thought to interact with the phosphate moiety of PEP [6,21], and the peptide nitrogen of Gly245, which is hydrogen bonded to the carboxylate of the substrate, increases from 8.3 Å in the M1 active site to 9.8 Å in the T-state catalytic centre. This indicates that in the transition to the T-state, two of the most important elements participating in the binding of the cooperative substrate move away from each other by 1.5 Å.

The present analysis suggests that the reorganization of loop 6 underlies the low PEP affinity of the T-state PK. The movement of the loop is necessary to accommodate the subunit rotations occurring during the allosteric transition. In the R-like state, loop 6 is forced to change its conformation to avoid collision with the backbone atoms of helix Aα7 from the r-related subunit. Loop 6 is therefore a crucial element which couples the subunit rotation to the restructuring of the cooperative substrate binding site. The role played in the allosteric transition by this loop is emphasized by the strict conservation of its amino acid sequence (residues 243–249) in all known eukaryotic and prokaryotic enzymes [4]. Intriguingly, the function of loop 6 as a flexible element involved in catalysis and substrate binding is not unprecedented in (B/α)8-barrel enzymes. For instance, in triosephosphate isomerase a 'closed' and 'open' conformations of this loop have been observed depending on the presence of the substrate [22]. In ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), the equivalent loop moves 12 Å upon activation of the protein [23]. In PK, loop 6 has a crucial function. It is instrumental to the allosteric regulation by adapting its own conformation to the quaternary structural transitions in the protein.

Comparison with other allosteric systems characterized at atomic resolution

The comparison between the structures of the T-state type I PK with the M1 enzyme [6,7] reveals a complicated set of motions which probably reflect the structural changes occurring in the transition from the T to the R
form. This notion is supported by the observation that the individual domains of the two proteins have remarkably similar conformations despite the large differences in their relative orientations. Allosteric transitions in tetrameric enzymes have been studied in phosphofructokinase [24] and fructose-1,6-bisphosphatase [25]. The two proteins behave like dimers of dimers, which change their relative orientations during the activation. Relative movements of each of the four subunits occur in the allosteric transition of the lactate dehydrogenase tetramer where, however, no domain rotations are observed [20]. 

E. coli type I PK displays a complex mechanism for the allosteric regulation, which is accomplished through concerted rotations of both the domains within the subunits and the subunits within the tetramer (Fig. 5c). The modular nature of PK allows the three domains of each subunit to change their articulation so that one domain closes to form a catalytic site while the other domain rotates to widen the allosteric effector binding pocket.

### Biological implications

The two major regulatory enzymes in glycolysis are phosphofructokinase and pyruvate kinase (PK). Each catalyzes an irreversible step of this metabolic pathway and their activity is allosterically controlled. In particular, PK is the major regulatory enzyme of the section of glycolysis and controls the flux from fructose-1,6-bisphosphate to pyruvate. The enzyme is allosterically activated by phosphoenolpyruvate and fructose-1,6-bisphosphate. No structural model for the allosteric transition of PK is available. Here we report the structure of PK from Escherichia coli in the inactive T-state; this is the first time that an atomic model for an allosteric PK has been described.

E. coli PK is a homotetramer. Each subunit can be divided into three domains, a large (β/α)₈-barrel domain (A), a small β-barrel domain (B), and a C-terminal domain (C) with an α/β open sheet topology. The allosteric effector binding site is located at the interface between domains A and C, whereas the active site is in a cleft between domains A and B. Both functional sites are entirely built by residues belonging to one subunit, none of them being located at the interface between subunits of the tetramer. The domain linkers appear to be flexible. The polypeptide segment connecting domains A and C (residues 346–351) is disordered and not visible in the electron-density map. Domain B has a slightly different orientation in two of the four polypeptide chains of the tetramer, underlining the modular nature of the PK subunits.

The structure of the T-state E. coli PK can be compared with the known structure of the muscle M1 isoenzyme. The latter is the only known PK that is not allosteric and is thought to adopt an active conformation similar to the R-state. The comparison reveals striking differences between the two structures. In the M1 model, the orientations of domains B and C differ by 17.2° and 14.9° respectively, relative to the T-state PK. These rotations are functionally relevant in that they affect the shape of the catalytic and allosteric sites located at the domain interfaces within each subunit. Furthermore, the M1 and the T-state structures display relative reorientations of the four subunits within the tetramers, with rotations larger than 16°. Highly conserved residues at the subunit interfaces couple these quaternary structure variations to the domain rotations occurring within the monomers. The differences between the T-state and M1 structures probably reflect the structural changes occurring during the R- to T-state transition. Most importantly, they illustrate how the modular nature of PK is instrumental to the allosteric regulation of the enzymatic activity.

Loop 6 of the (β/α)₈-barrel takes part in the binding of phosphoenolpyruvate. In the inactive state, this loop undergoes a conformational change that distorts the active site. This alteration of the catalytic centre explains the low affinity of the T-state for the cooperative substrate and is directly coupled to the subunit rotations. A subunit having loop 6 in the T-state conformation would not be compatible with the R-state quaternary structure. The adaptability of loop 6 therefore plays a crucial role by linking domain and subunit rotations to modifications of the active-site geometry. The role of loop 6 as a flexible element participating in substrate binding and catalysis is a feature common to several other (β/α)₈-barrel-containing enzymes.

In the two-state model of allosteric control proposed by Monod et al. [14] regulation of enzyme activity is thought to arise from the alteration of the equilibrium between two or more activity states of the enzyme. Their model predicted that allosteric enzymes would probably be symmetric oligomers in which variations of the quaternary structure would be involved in the allosteric transition. PK provides a particularly sophisticated example of the two-state paradigm wherein complex and concerted rotations, involving both the domains and the subunits, occur upon transition from the inactive to the active state of the enzyme.

### Materials and methods

#### Crystallization

Crystals of the unligated type I pyruvate kinase from E. coli were grown from 16% (w/v) PEG 8000, 10 mM MgSO₄, 10 mM KCl, 100 mM MES/NaOH (pH 6.2) using the
hanging drop method, as described in [26]. They belong to space group C2221 with cells dimensions a=76.8 Å, b=247.5 Å, c=132.6 Å and one dimer in the asymmetric unit. Soaking with the allosteric activators, phosphoenolpyruvate and FBP, leads to the immediate cracking of the crystals. This observation indicates that the crystallized protein is in the inactive T-state in agreement with the kinetic data [5], which show that the unligated enzyme is the inactive form.

**Data collection and MIR phasing**

Attempts to solve the structure by molecular replacement using the cat M1 atomic coordinates [6] as a search model failed. No attempts were made using the rabbit M1 model [7] as its coordinates became available only in the later stages of the structure determination. Therefore, the MIR method was used for solving the structure. The application of this technique, however, was severely hampered by the lack of isomorphism between the native crystals. It was only possible to obtain an interpretable difference Patterson map after collecting a number of native and derivative data sets. For data collection and heavy-atom screening the crystals were transferred in a stabilizing solution containing 25% (w/v) PEG 8000, 10 mM MgSO4, 10 mM KCl, 100 mM MES/NaOH pH 6.2. All data sets used for the MIR phasing were collected at room temperature on a Rigaku RAXIS IIC imaging plate system, using CuKα radiation. The images were evaluated using a modified version of MOSFLM (A Leslie) while the CCP4 suite [27] was used in the data reduction. Table 3 gives a summary of the statistics for the data sets which were used in the structure determination. The Gd3+ and Lu3+ isomorphous difference Patterson maps were solved using SHELX-90 [28]. The iodide-binding sites were located by difference Fourier techniques. Although two of the four Gd3+-binding sites coincide with the two sites of the Lu3+ derivative, both derivatives were used for phasing. The heavy-atom parameters were refined using MLPHARE [29]. The overall figure of merit was 0.43 for data between 20.0 Å and 3.6 Å resolution.

**Density averaging and map interpretation**

If there is one dimer in the asymmetric unit, the solvent content of the crystals is 60% [30]. Self-rotation function calculations carried out using the program GLRF [31] gave a strong peak at κ=180°, ϕ=55°, ψ=90° [32] indicating that one of the molecular twofold axes of the 222 symmetric tetramer was coincident with the crystallographic b axis. The positions of the Lu3+-binding sites allowed the non-crystallographic twofold axis to be located in the unit cell. The position of the twofold axis was confirmed by the 'real space translation function' implemented in the program GLRF [31]. An envelope around the protein could be calculated from a local correlation map [33] enclosing 47% of the asymmetric unit. The MIR phases were improved by twofold averaging and solvent flattening using the program DEMON [34] with gradual phase extension from 5.5 Å to 3.5 Å. The resulting map was interpretable for domains A and C but not for the region corresponding to domain B (residues 71-170). A partial model for domains A and C was built using the program O [35]. The cat M1 PK structure [6] made interpretation of the electron-density map easier. The initial model (77% of the residues) was refined by energy minimization with non-crystallographic symmetry constraints using the program X-PLOR [36]. The native data, up to 3.2 Å resolution, were employed for refinement. The phases from the refined model were combined with the phases obtained after twofold averaging using SIGMAA [37]. The resulting map was significantly improved and domain B could be built. This provided the initial model for the crystallographic refinement at 2.5 Å resolution.

**High-resolution data collection and crystallographic refinement**

Data at higher resolution were obtained using synchrotron radiation at the X11 beam line of the EMBL outstation at DESY (Hamburg, Germany). The data were collected at 100 K by means of cryocrystallographic techniques [38] using 25% (w/v) PEG 400 as cryoprotectant. A total of 140 images with a width of 0.5° were recorded on a Mar Research imaging plate at a wavelength of 0.92 Å. The intensities were evaluated and internally scaled using the programs DENZO and SCALEPACK [39]. The data set is 99% complete up to 2.5 Å resolution with an overall Rmerge of 8.8% (Table 3). The data processing revealed that the crystal used belonged to space group P212121 (a=73.9 Å, b=129.6 Å, c=241.4 Å) rather than to the usual C2221, space group. The P212121 symmetry has been observed occasionally in cryocooling experiments conducted in-house but never for PK crystals for which data have been collected at room temperature. This suggests that

<table>
<thead>
<tr>
<th>Table 3. Data collection and MIR statistics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data set</td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Native*</td>
</tr>
<tr>
<td>Gd3+ 2 mM*</td>
</tr>
<tr>
<td>Lu3+ 0.2 mM*</td>
</tr>
<tr>
<td>IrCl6 40 mM*</td>
</tr>
<tr>
<td>Native 100 K</td>
</tr>
</tbody>
</table>

*All the data sets used for the MIR phasing were collected at room temperature. The soaking time was 12 h for all three heavy-atom derivatives. The native high resolution data were collected at 100 K using cryocrystallographic techniques (see text). Rmerge=\[\sum_{j=1}^{N} |F_{ij}|-|F_{ij}^n|/\sum_{j=1}^{N} |F_{ij}|,\] where |F_{ij}| is the intensity of an observation of reflection j and |F_{ij}^n| is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j. The Rmerge is the average intensity for reflection j.

---

T-state pyruvate kinase Mattevi et al. 739
the change of the space group is probably induced by the 'shock' freezing. Because of the different space group, molecular replacement was carried out using as a search model the structure built in the map obtained after density averaging and partial model refinement of the C2221 crystal form. The program AMORE [40] was employed. The molecular replacement solution was immediately evident, and showed that the molecular packing in the P212121 crystals is virtually the same as in the C2221 space group crystals. The difference between the two crystals forms is that in P212121, all three molecular twofold axes of the tetramer are non-crystallographic whereas one of them is coincident with the crystallographic b axis in the C2221 crystals. Thus, in the P212121 crystal the asymmetric unit contains one full tetramer.

The crystallographic refinement to 2.5 Å was carried out using the program PROLSQ/PROFFT [41] with tight non-crystallographic symmetry restraints and Engh and Huber stereochemical parameters [42]. The program O was used for manual rebuilding of the model [35]. A random sample containing 4% of the total data (2890 reflections) was excluded from the refinement and used for the calculation of the free R-factor [43]. When the R-factor dropped below 25%, water molecules were added at positions with density >1σ in the 2F o-F c map and >3σ in the F o-F c map. A second criterion for incorporating waters was that they had to be engaged in at least one hydrogen bond with a protein or other solvent atom.

The electron-density maps calculated after density averaging and during the crystallographic refinement agreed well with the published amino acid sequence [13] with the exception of the residues located at the C terminus of the polypeptide chain. In particular, the electron density unambiguously indicated the presence of eight additional C-terminal amino acids. On the basis of the crystallographic information, the sequence of the gene encoding type I PK was analyzed revealing that a frameshift error, resulting from the spurious insertion of a T between residues 450-470 (Fig. 3) and all other known PK enzymes [4]. Moreover, all backbone and side-chain atoms belonging to these residues are located in extremely well defined electron density.

The current model contains 1760 residues and 434 water molecules (see Table 4 for model refinement statistics). Residues 346-351 are poorly ordered and not visible in the electron density of all four independent subunits. Moreover, residues 75-105 of subunit 3 display poor density and are not included in the final model. These residues are, however, well defined in the other three monomers. In the Ramachandran plot (Fig. 9), 90% of the residues are in the most favoured regions [44]. Ser312 is the only residue which in all four subunits has an energetically unfavourable (φ,ψ) pair. This residue has well defined electron density in all four subunits (Fig. 1).

All the calculations related to the analysis of the model and the comparison with the M1 structure were carried using the CCP4 package [27] and the program O [35]. The coordinates of the T-state E. coli PK have been deposited with the Brookhaven Protein Data Bank (entry 1PKY).

Acknowledgements: We wish to thank Hugo Monaco (Pavia), Silvia Onesti (Pavia) and Massimo Malcovati (Milano) for their continuous interest in the project and for many helpful discussions. The supervision of Thomas Schneider and Alex Teplyakov (DESY, Hamburg) during synchrotron data collection is most gratefully acknowledged. We thank Ivan Rayment (Madison) and Hilary Muirhead (Bristol) for providing the M1 PK coordinates.

### Table 4. Model refinement statistics.

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of non-hydrogen protein atoms</td>
<td>13637</td>
</tr>
<tr>
<td>No. of solvent molecules</td>
<td>434</td>
</tr>
<tr>
<td>Resolution limits</td>
<td>6.0–2.5 Å</td>
</tr>
<tr>
<td>R-factor</td>
<td>20.5%</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>72305</td>
</tr>
<tr>
<td>Free R-factor</td>
<td>30.7%</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>2890</td>
</tr>
<tr>
<td>Rms deviation in bond lengths</td>
<td>0.016 Å</td>
</tr>
<tr>
<td>Rms deviation in bond angles</td>
<td>1.6°</td>
</tr>
<tr>
<td>Rms deviation B bonded atoms</td>
<td>2.8 Å2</td>
</tr>
<tr>
<td>Average B-factor subunit 1</td>
<td>41.4 Å2</td>
</tr>
<tr>
<td>Average B-factor subunit 2</td>
<td>37.0 Å2</td>
</tr>
<tr>
<td>Average B-factor subunit 3</td>
<td>44.2 Å2</td>
</tr>
<tr>
<td>Average B-factor subunit 4</td>
<td>34.8 Å2</td>
</tr>
<tr>
<td>Average B-factor water molecules</td>
<td>43.0 Å2</td>
</tr>
</tbody>
</table>

References


Received: 21 Apr 1995; revisions requested: 16 May 1995; revisions received: 2 Jun 1995. Accepted: 2 Jun 1995.