Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure of a circularly permuted phosphoglycerate kinase

Pierre Tougard,[a] Thierry Bizebard,[b] Monica Ritco-Vonsovici,[a]† Philippe Minard[a] and Michel Desmadril[a]*

[a]Laboratoire de Modélisation et d'Ingénierie des Protéines, UMR 8619, Université de Paris-Sud, Bâtiment 430, F-91405 Orsay CEDEX, France, and [b]Laboratoire d'Enzymologie et Biologie Structurale, UPR 901, Avenue de la Terrasse, F-91128 Gif-sur-Yvette, France

† Present address: Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, England.

Correspondence e-mail: michel.desmadril@mip.u-psud.fr

Received 11 March 2002
Accepted 28 August 2002

PDB Reference: circularly permuted phosphoglycerate kinase, 1fw8, r1fw8sf.

The crystallographic structure of a circularly permuted form of yeast PGK, 72p yPGK, has been determined to a resolution of 2.3 Å by molecular replacement. In this engineered protein, the C- and N-terminal residues of the wild-type protein are directly connected by a peptide bond and new N- and C-terminal residues are located within the N-terminal domain. The overall fold of the protein is very similar to that of the wild-type protein, directly demonstrating that the continuity of a folding unit is not relevant to the folding process of the whole protein. Only limited structural changes were observed: these were in the regions associated with the new connection, in a long flexible loop in the permuted domain and in the vicinity of Arg38, a functionally important residue. The relative positions of the two domains suggested that this permuted protein adopts one of the most open/twisted conformations seen amongst PGKs of known structure. The effect of the mutation on the functional properties is more easily accounted for by a restriction of hinge-bending motion than by structural changes in the protein.

1. Introduction

The genetically engineered circular permutation of a polypeptide sequence was first described in 1989 by Kirschner and coworkers for the TIM-barrel enzyme phosphorybosyl anthranilate isomerase (Luger *et al.*, 1989). This elegant strategy has since proved incisive for investigating fundamental aspects of protein folding. As protein folding probably occurs *in vivo* as a co-translational process, it has long been thought that the translation order and folding speed of the successive sequential elements could be critical features directing the folding process to produce the correct structure. In circularly permuted proteins, however, structures that are structurally adjacent in the wild-type protein are moved apart at the sequence level. New C- and N-terminal residues are created and the natural termini of the wild-type polypeptide chain are linked to each other in the permuted protein. This artificial separation of individual residues or secondary-structure elements that usually come into close contact early in the folding process may lead to perturbations in folding kinetics and possibly to the incorrect folding of the polypeptide chain.

However, since the pioneering work of Kirschner's group (Luger *et al.*, 1989), this strategy has been successfully applied to many different proteins and collectively these studies provide strong evidence that the order of translation of structural elements is not relevant to the folding efficiency of natural proteins (Buchwalder *et al.*, 1992; Horlick *et al.*, 1992; Yang & Schachman, 1993; Zhang *et al.*, 1993; Hahn *et al.*, 1994; Protasova *et al.*, 1994; Jung & Lee, 2001).

Circular permutation can also be used to challenge received wisdom concerning the folding process of multidomain

Table 1
Crystallographic data and refinement statistics.

Crystallographic data	
Space group	$P2_12_12$
Unit-cell parameters (Å)	a = 85.74, b = 104.77, c = 44.34
Resolution limit (Å)	2.3
No. of observations	64243
No. of unique reflections	18306
Redundancy	3.5 (3.3)†
Completeness (%)	99.8 (99.4)†
R_{sym}‡	0.079 (0.322)†
$I/\sigma(I)$	15.0 (3.5)†
Refinement statistics	
Total No. of selected reflections	17625
Working set	16817
Resolution range (Å)	7.0–2.3
No. of non-H atoms	3353
Mean B factor for non-H atoms	26.5 (14.1)§
No. of water molecules	181
Mean B factor for the water molecules	35.4 (11.9)§
R.m.s.d. from ideal values	
Bond lengths (Å)	0.011
Angle values (°)	1.7
Dihedral angles (°)	24.1
Ramachandran plot	
Most favoured regions (%)	87.7
Allowed regions (%)	12.0
Disallowed (%)	0.3
Refinement R‡ value	0.211
Free R value¶	0.324

† Values in parentheses are for last resolution shell (2.38–2.30 Å). ‡ $R_{sym} = \sum_{hkl}\sum_i |\langle I_i\rangle - I_i|/\sum_{hkl}\sum_i I_i$, $R = \sum_{hkl} ||F_o| - |F_c||/\sum_{hkl}|F_o|$. § B-factor mean standard deviations are given in parentheses. ¶ The free R value (Brünger, 1992) was calculated with a test set consisting of 5% of the structure factors, randomly selected and not used for structure refinement.

proteins. It is often stated that in multidomain proteins the individual domains fold first and then interact with each other. Circular permutation of a protein consisting of several continuous domains can be used to disrupt the sequence continuity of one of the domains presumed to be a folding unit. The following questions may then be addressed. Is sequence continuity within a domain required for that domain to be a folding unit? Does circular permutation modify the overall folding process of the protein or only that of the permuted domain? Does perturbation of the temporal sequence of folding events prevent the correct folding of the protein? We have constructed circularly permuted variants of yeast phosphoglycerate kinase (Ritco-Vonsovici *et al.*, 1995). This protein is a monomeric enzyme of 415 amino acids that folds into two continuous globular domains of approximately the same size (Watson *et al.*, 1971). The N- and C-terminal residues are very close to each other (about 5.3 Å apart) and can therefore be directly connected. In a previous study, we described the construction of two circularly permuted variants and their functional properties (Ritco-Vonsovici *et al.*, 1995). We found that the natural continuity of the domains is not required for this protein to achieve a native-like active structure. These mutated proteins are less stable than the wild-type protein, but their ability to fold into a functional protein is not fundamentally affected. This suggests that the circular permutations did not disrupt structural units essential for folding and that the native network of short-range interactions can be acquired despite the discontinuity introduced within each domain.

We have investigated the structural consequences of permutation by determining the X-ray crystal structure of a yeast phosphoglycerate kinase mutant in which the new N-terminus was located at residue 72 (72p yPGK). This position offers the advantage of being located on an external loop exposed to solvent and not directly involved in the active site. We describe here the structure of this protein and present a detailed comparison with other known PGK structures.

2. Materials and methods

2.1. Crystallization

The 72p circularly permuted mutant of yPGK was constructed and purified as previously described (Ritco-Vonsovici *et al.*, 1995). The protein was crystallized by the hanging-drop technique using the same buffer as described for wild-type yeast PGK (Watson *et al.*, 1971) or Arg203Pro yeast PGK (Tougard *et al.*, 1996). The precipitant used was ammonium sulfate at a concentration of 2.0–2.6 *M* and a pH of 6.1. The optimum protein concentration was 7–10 mg ml^{-1} and the crystallization temperature was about 293 K. Crystals formed and grew over a period of three to six months, giving long prism-shaped needles 0.05–0.1 mm thick and 0.3–1 mm long.

2.2. Data acquisition

In order to avoid radiation damage to the crystal, data were collected under cryogenic conditions. We soaked the crystals in a series of buffers containing 2.8 *M* ammonium sulfate with increasing concentrations of glycerol [0–25%(*v*/*v*)] for cryoprotection.

Crystals of about 200 × 60 × 30 µm in size were mounted on a 200 µm nylon loop, quickly frozen in a stream of nitrogen at about 120 K (Oxford Cryogenic System) and exposed to 1.375 Å synchrotron radiation from beamline D41 (Fourme *et al.*, 1992) of LURE-DCI Orsay. Diffraction data were collected at a resolution of 2.3 Å on an 18 cm diameter MAR Research image plate using a single crystal and were processed with *DENZO* and *SCALEPACK* (Otwinowski, 1993).

The crystals were found to be orthorhombic, with space group $P2_12_12$; the unit-cell parameters were a = 85.74, b = 104.77, c = 44.34 Å.

Using the standard mean value of 2.5 Å^3 Da^{-1} for crystal volume per unit protein molecular mass (Matthews, 1968) and considering the molecular mass of 44 500 indicated there to be one molecule per asymmetric unit. A total of 18 306 unique reflections were collected. The experimental data are given in Table 1.

2.3. Structural determination

As 72p yPGK has 14% of the enzymatic activity of the wild-type enzyme (Ritco-Vonsovici *et al.*, 1995), we expected the structures of the 72p yPGK and wild-type enzymes to be similar. We therefore used the molecular-replacement method with the *AMoRe* program (Navaza, 1994) to determine the structure of the mutant. However, to take into account

possible rotation of N- and C-terminal domains around the 72p yPGK hinge region, calculations were made using the individual domains. The molecular structure of R65Q yPGK (McPhillips *et al.*, 1996), the most similar available yPGK structure resolved to high resolution and available in the Protein Data Bank (Bernstein *et al.*, 1977), was chosen as a molecular model and split into two blocks, with amino acids 6–67 and 74–190 in one block and amino acids 203–403 in the other (wild-type yPGK numbering).

The diffraction data used for molecular-replacement studies extended over the resolution range 10–3.5 Å. A clear and contrasted solution was obtained giving, after rigid-body refinement, a correlation factor CC of 0.55 between the observed and calculated structure factors. The second solution of the molecular replacement had a CC of 0.20. The molecular model was then subjected to 100 cycles of *X-PLOR* conjugate-gradient refinement (Brünger *et al.*, 1987), resulting in a reliability index R of 0.26 at a resolution of 2.8 Å. At this stage, a σ_A-weighted (Read, 1986) $2|F_o| - |F_c|$ difference Fourier map was calculated and used to model the 415–1 peptide link and the new terminal polypeptide residues (residues 71 and 72). Five cycles of alternate rebuilding from electron-density maps and *X-PLOR* conjugate-gradient refinement gave an R index of 0.23 at a resolution of 2.7 Å.

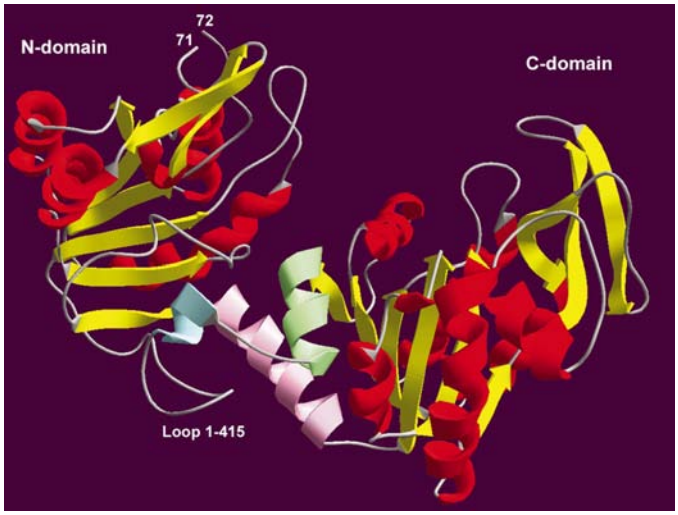

Figure 1
Schematic structure of 72p yPGK. The new link 1–415 and the new extremities are indicated. Helix V is shown in light pink, helix XIII in light green and helix XIV in light blue.

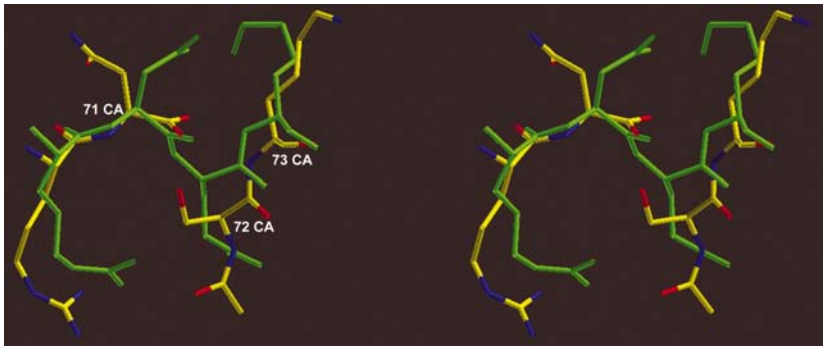

Figure 2
Stereoview of the refined model of the 70–73 region of 72p yPGK together with the backbone of WT yPGK (green) of the same region.

Model rebuilding from examination of a weighted $2|F_o| - |F_c|$ map (Read, 1986) plotted at one standard deviation above the mean level of the map was performed with *O* (version 5.9; Jones *et al.*, 1991) on a Silicon Graphics workstation. Each cycle of *X-PLOR* refinement consisted of 100 cycles of atomic position refinement followed by 20 cycles of temperature-factor refinement. The introduction of 17 625 unique reflections at 2.3 Å (consisting of a working set of 16 817 reflections and a test set of 808 reflections) in an *X-PLOR* simulated-annealing refinement gave an R index of 0.26.

Finally, after ten cycles of alternate model rebuilding, conjugate-gradient refinement, introduction of 181 water molecules and four glycerol molecules, and taking into account the replacement of Glu72 by an *N*-acetylserine group (Ritco-Vonsovici *et al.*, 1995), we obtained the final structure with an R index of 0.211 (free R value of 0.324) at a resolution of 2.3 Å.

Water molecules were selected as peaks above 1.5σ on weighted $2|F_o| - |F_c|$ maps that also fulfilled stereochemical criteria: the peak was retained if it was at an acceptable distance from a hydrogen-bond donor or acceptor compatible with hydrogen-bond formation. Finally, water-molecule temperature factors over 70 Å^2 were not considered.

3. Results

Comparison of 72p yPGK with wild-type PGK showed no major differences in the overall conformation of the polypeptide chain (Fig. 1). The root-mean-square deviation between all C$^\alpha$ equivalent positions in the two structures was about 0.5 Å for the N-terminal domain and 0.9 Å for the C-terminal domain. For the sequence stretches 65–71 and 72–77, corresponding to the new extremities, the r.m.s.d. for C$^\alpha$ atoms was 1.2 Å, whereas that for the new link 415–1 was 1.3 Å.

The main structural features used for comparison with other known PGK structures were the C$^\alpha$ positions for each domain, the rotations of the N- and C-terminal domains and conformations of the 'nose' region (a long loop in the N-terminal domain, amino acids 128–142), the newly formed 415–1 link region, the new termini regions and the region around Asp23 and Arg38, both of which are residues involved in catalytic activity.

3.1. Comparison of domains

The molecular structures of the 72p yPGK mutant domains are very similar to those of WT yPGK and other PGKs. Table 2 gives the r.m.s.d.s between C$^\alpha$ atoms of various regions of 72p yPGK and the C$^\alpha$ atoms of the same regions in other PGKs (Watson *et al.*, 1971; Harlos *et al.*, 1992; Davies *et al.*, 1994; McPhillips *et al.*, 1996; Tougard *et al.*, 1996; Auerbach *et al.*, 1997; Bernstein & Hol, 1998; Bernstein *et al.*, 1998),

excluding the 'nose' region (*i.e.* amino acids 128–142) because this region is absent [in 1php (Davies *et al.*, 1994) and 1vpe PGKs (Auerbach *et al.*, 1997)] or very different (in 13pk PGK; Bernstein & Hol, 1998; Bernstein *et al.*, 1998).

The high variability of PGK N- and C-terminal domain relative angular rotations around the hinge region (Table 2), even if partially induced by crystal packing, reflects the internal flexibility necessary for PKG activity (Banks *et al.*, 1979; Pickover *et al.*, 1979; Wilson *et al.*, 1988; Desmadril *et al.*, 1991; Guilbert *et al.*, 1996). Analysis of rotation-axis equation coefficients showed that rotation axes could be classified into at least three groups according to the various orientations in the coordinate system of 72p yPGK (Table 2). The first mean direction of rotation between C-terminal domains, after superimposition of N-terminal domains, lies almost parallel to helix V and is defined by C^{α} atoms 186 and 194. The second direction of rotation lies roughly parallel to a line from C^{α} atom 21 to C^{α} atom 186. The third rotation-axis direction is parallel to a line passing through C^{α} atoms 187 and 384. These results indicate that the motion between the two domains is probably not a simple hinge motion, but a more complicated convolution of several rotations.

Finally, looking along the various rotation axes from the 72p yPGK N-terminal domain side, shows that a clockwise rotation is necessary in each case to bring the C terminal domain of other PGKs into the same position as the C-terminal domain of 72p yPGK.

3.2. New N- and C-terminal extremities region

The r.m.s.d.s between C^{α} positions 65–77 of 72p yPGK and those of other yeast PGKs are 0.7, 1.2 and 1.1 Å for R65Q yPGK, WT yPGK and R203P yPGK, respectively. The creation of two new extremities increases the degrees of freedom for new hydrogen bonding, but the low r.m.s.d. with the same region of other yeast PGKs suggests only minor changes in C^{α}-residue positions (Fig. 2).

Initial examination of a $|F_o| - |F_c|$ difference map showed no major electron-density features in this area, except peaks accounting for the presence of the *N*-acetyl modification of Ser72. This preliminary examination suggested two possible rotamers of the *N*-acetylserine group around the 72 C^{α}—72 C bond. The two different structures were refined and comparison of temperature factors unambiguously selected one of the orientations of the *N*-acetylserine group with the main-chain NH of Ser72 forming a hydrogen bond to the terminal carboxyl group of Asn71. This position of the *N*-acetylserine does not significantly displace the C^{α} positions, with only a few changes observed in the hydrogen-bond network. Moreover, buried water molecules were present at equivalent positions in both the 72p yPGK and the R65Q mutants, with a similar

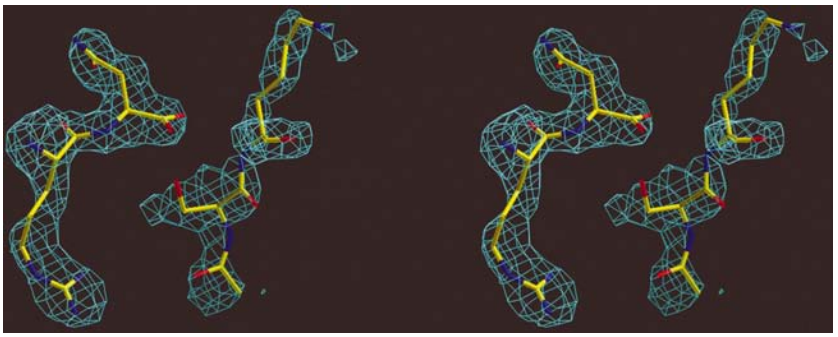

Figure 3
Stereoview of an omit $2|F_o| - |F_c|$ map contoured at one standard deviation above the mean level of the map as computed with *O* (version 7.0; Jones *et al.*, 1991). The refined model of the 70–73 region of 72p yPGK is superimposed.

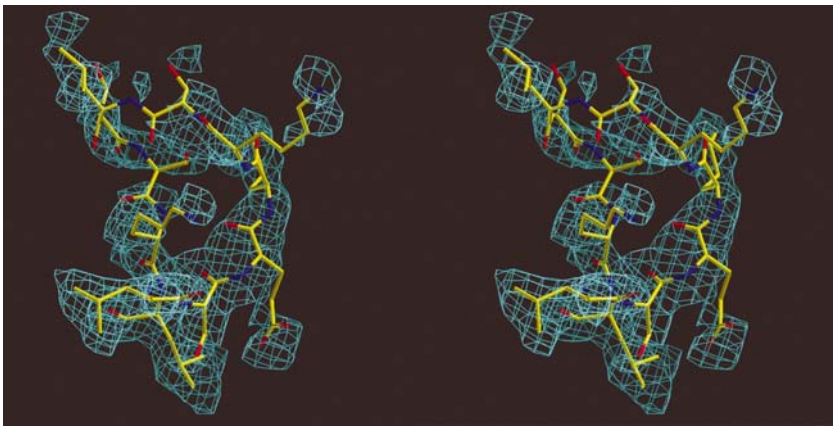

Figure 4
Stereoview of an omit $2|F_o| - |F_c|$ map contoured at 0.7 standard deviations above the mean level of the map, computed with *O* (version 7.0; Jones *et al.*, 1991). The refined model of the 6–411 region of 72p yPGK is superimposed.

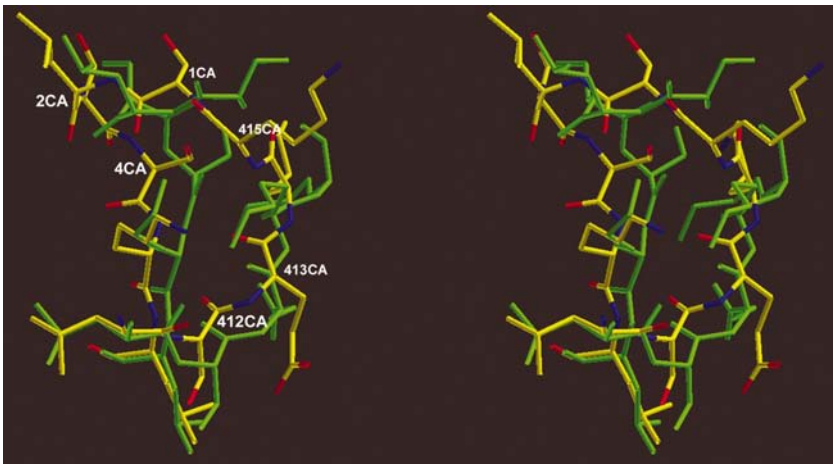

Figure 5
Stereoview of the refined model of the 6–411 region of 72p yPGK together with the backbone of WT yPGK (green) of the same region.

pattern of hydrogen bonding observed in both proteins. Fig. 3 shows the omit $2F_o - F_c$ electron-density map drawn for the region 70–72 together with the modelling of the polypeptide-chain conformation through the electron density. As can be observed, the newly formed extremities are clearly defined and separated.

3.3. New 415–1 peptide-link region

The newly built 1–415 loop has been modelled in an electron-density map which is poorly defined in this region. Examination of the temperature factors of this region shows higher values than in the rest of the molecule (up to about three times the mean value of the molecule for some residues). Fig. 4 shows the omit $2F_o - F_c$ map drawn for the region 6–411 (excluding these residues from the computation) together with the modelling of the polypeptide chain. The electron density is absent or poor for chain segments between the C^α atoms of residues 1 and 2 and between the C^α atoms of residues 414 and 415. Owing to the lack of definition in this region, it must be noticed that if residues 413, 414 and 415 are found in generously allowed regions of a φ–ψ Ramachandran plot, residue 5 is in a disallowed region. Furthermore, some bond-angle discrepancies between standard values for the polypeptide chain and those for residues 412, 413 and 415 are observed. Finally, the higher values of the atomic *B* factors found in this region suggest that if this interpretation is a plausible one, there may be more than one unique conformation in this region of the 72p yPGK crystal structure. The r.m.s.d.s between C^α positions 409–413 and 1–7 of 72p yPGK and the equivalent C^α atoms in other yeast PGK proteins are 1.41, 1.40 and 1.30 Å for R65Q yPGK, WT yPGK and R203P yPGK, respectively. Therefore, the introduction of a new covalent bond between the natural extremities induced, as expected, very limited local structural changes (Fig. 5). However, these limited local changes may be associated with functionally relevant constraints in rotation between the two domains.

3.4. Region around Arg38 and Asp23

The region around Arg38 and Asp23 presents a minor rearrangement of the hydrogen-bond network compared with that observed in R65Q or wild-type yPGK. In 72p yPGK, as in R65Q, Arg38 is linked to Asp23 by a hydrogen bond. However, the hydrogen bond between Arg38 and the carbonyl group of Asn25, present in R65Q is absent in 72p yPGK as in wild-type yPGK, because of the rotation of the guanidinium group. In this region, 72p yPGK seems to retain one glycerol and some water molecules in positions equivalent to those of water molecules in R65Q yPGK.

Table 2
General comparison of the N- and C-terminal domains of various PGK crystals studied, excluding the nose region, and comparison of the rotation angles of C-terminal domains with respect to the C-terminal domain of 72p yPGK, without changing the position of the N-terminal domain.

The C^α segments chosen for comparisons were 12–27, 33–54, 57–64, 73–94, 113–124 and 146–172 for the N-terminal domain and 201–210, 228–244, 251–267, 275–288, 308–357 and 366–389 for the C-terminal domain; for pig and horse PGK, amino-acid C^α atoms at positions 369 and beyond were not introduced.

	Domain	C^α	R.m.s.d. (Å)	Rotation angle† (°)	Direction cosines of the average rotation axis‡
WT yPGK	N	107	0.5	12	−0.849, −0.219, −0.479
	C	132	0.9		
R203P yPGK	N	107	0.5	11.7	
(Tougard *et al.*, 1996)	C	132	0.8		
R65Q yPGK	N	107	0.4	19.4	
(McPhillips *et al.*, 1996)	C	132	0.6		
1php PGK	N	107	0.6	11.4	
(Davies *et al.*, 1994)	C	132	0.8		
Horse	N	107	0.6	13.6	
(Banks *et al.*, 1979)	C	112	0.5		
16pk PGK	N	107	0.5	16.7	−0.543, −0.670, 0.446
(Bernstein *et al.*, 1998)	C	132	1.1		
Pig	N	107	0.7	11.1	
(Harlos *et al.*, 1992)	C	112	0.4		
1vpe PGK	N	107	0.7	29.3	−0.532, 0.018, 0.847
(Auerbach *et al.*, 1997)	C	132	0.8		
13pk PGK	N	107	0.6	31.5§	
(Bernstein & Hol, 1998)	C	132	0.9		

† Rotation angles were calculated with the *SUPERPK* program (P. Alzari, personal communication), superimposing the N-terminal domain of p72 yPGK on that of another yeast PGK and then superimposing the C-terminal domain of the second PGK on that of p72 yPGK. ‡ Direction cosines are given in an orthogonal coordinates system built on the main axis of p72 yPGK. § Mean value for the four molecules found in the asymmetric unit listed in the PDB.

3.5. 'Nose' region

We calculated the r.m.s.d. between nose region C^α positions and the C^α positions of other PGK structures for the same region, *i.e.* residues 128–142. We obtained values of 0.91, 4.60, 4.27 and 1.93 Å for R65Q yPGK, WT yPGK, R203P yPGK and pig PGK, respectively. These figures, which are significantly higher than those of any other region of the protein, are probably accounted for by the intrinsic flexibility of this region.

4. Discussion

As already suggested by the method used to solve the 72p yPGK structure (*i.e.* molecular replacement), the structural similarities between the N-terminal and C-terminal domains of circularly permuted yeast 72p yPGK and other domains of PGKs from various sources are not that surprising. Clearly, the circularly permuted yeast 72p yPGK adopted typical PGK-like folding despite the new topology induced by circular permutation of the sequence.

Although the wild-type structure is strongly conserved, the permuted mutant is less stable than WT yPGK. Moreover, its specific activity is significantly lower, representing 14% of that of WT yPGK. Various features may account for this decrease in specific activity. It has been suggested that residue Arg38 plays a significant role in the expression of enzymatic activity (Sherman *et al.*, 1992). The minor rearrangement in the

hydrogen-bond network observed in the region around residues Asp23–Arg38 may be involved in this decrease in specific activity, although differences in crystallization conditions between R65Q, 72p and wild-type yPGK could also account for such rearrangement.

However, the main structural cause of the observed change in specific activity may be related to the internal dynamics of the mutant protein. PGK is the archetypal protein in which enzyme activity is directly related to the hinge-bending motion of its structural domains. This motion was first proposed by Banks *et al.* (1979) and has been confirmed by various experimental (Pickover *et al.*, 1979; Wilson *et al.*, 1988; Desmadril *et al.*, 1991; McHarg *et al.*, 1999) and theoretical studies (Guilbert *et al.*, 1996; Chandra *et al.*, 1998). Crystal structures have since shown that PGK can crystallize in various forms from a closed form to an open form (Bernstein *et al.*, 1977; Banks *et al.*, 1979; Harlos *et al.*, 1992; Davies *et al.*, 1994; Auerbach *et al.*, 1997; Bernstein *et al.*, 1998; Szilagyi *et al.*, 2001) according to the species and/or the substrate present.

Comparison of the various PGK structures available supports the results of previous normal-mode analysis suggesting shearing and propeller-twist motions for low-frequency motions, involving, in particular, helices V and XIII (Guilbert *et al.*, 1996). However, more specifically, the clockwise rotations of the C-terminal domains of various PGKs with respect to the C-terminal domains of 72p yPGK indicates that the conformation of 72p yPGK is close to being one of the most open/twisted conformations. This may be a consequence of the presence of the newly formed 1–415 bond, which may hamper the slight displacements of helix V required for rotation around the axis of this helix, or may separate the N-terminal domain from the C-terminal domain by maintaining helix XIV close to loop 161–171. This conformation should not be considered to be rigid, but merely to have been stabilized as a primary effect of the new 415–1 link. Finally, all these observations can lead us to speculate that the new peptide link in the permuted protein can subtly interfere with the hinge-binding motion of the protein, which in turn affects its enzymatic activities (Ritco-Vonsovici *et al.*, 1995). Such an indirect effect on the PGK hinge-bending motion was previously suggested in a study of the R203P PGK mutant (Tougard *et al.*, 1996), in which the replacement of a proline by an arginine led to a decrease in enzyme activity, even though the mutation was located far from the active site and there was no major conformational change.

This work was supported by the Centre National de la Recherche Scientifique and the Ministère de l'Enseignement Supérieur et de la Recherche. We thank Dr J. Janin and Dr M. Knossow for providing access to the facilities of their laboratory. We thank the LURE technical staff for running the synchrotron and the computing facilities.

References

Auerbach, G., Huber, R., Grattinger, M., Zaiss, K., Schurig, H., Jaenicke, R. & Jacob, U. (1997). *Structure*, **5**, 1475–1483.

Banks, R. D., Blake, C. C., Evans, P. R., Haser, R., Rice, D. W., Hardy, G. W., Merrett, M. & Phillips, A. W. (1979). *Nature (London)*, **279**, 773–777.

Bernstein, B. E. & Hol, W. G. (1998). *Biochemistry*, **37**, 4429–4436.

Bernstein, B. E., Williams, D. M., Bressi, J. C., Kuhn, P., Gelb, M. H., Blackburn, G. M & Hol, W. G. (1998). *J. Mol. Biol.* **279**, 1137–1148.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. E., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Brünger, A. T. (1992). *Nature (London)*, **355**, 472–474.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Buchwalder, A., Szadkowski, H. & Kirschner, K. (1992). *Biochemistry*, **31**, 1621–1630.

Chandra, N. R., Muirhead, H., Holbrook, J. J., Bernstein, B. E., Hol, W. G. & Sessions, R. B. (1998). *Proteins*, **30**, 372–380.

Davies, G. J., Gamblin, S. J., Littlechild, J. A., Dauter, Z., Wilson, K. S. & Watson, H. C. (1994). *Acta Cryst.* D**50**, 202–209.

Desmadril, M., Minard, P., Ballery, N., Gaillard-Miran, S., Hall, L. & Yon, J. M. (1991). *Proteins*, **10**, 315–324.

Fourme, R., Dhez, P., Benoît, J. P., Kahn, R., Dubuisson, J. M., Besson, P. & Frouin, J. (1992). *Rev. Sci. Instrum. B*, **63**, 982–987.

Guilbert, C., Pecorari, F., Perahia, D. & Mouawad, L. (1996). *Chem. Phys.* **204**, 327–336.

Hahn, M., Piotukh, K., Borriss, R. & Heinemann, U. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 10417–10421.

Harlos, K., Vas, M. & Blake, C. F. (1992). *Proteins*, **12**, 133–144.

Horlick, R. A., George, H. J., Cooke, G. M., Tritch, R. J., Newton, R. C., Dwivedi, A., Lischwe, M., Salemme, F. R., Weber, P. C. & Horuk, R. (1992). *Protein Eng.* **5**, 427–431.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Jung, J. & Lee, B. (2001). *Protein Sci.* **10**, 1881–1886.

Luger, K., Hommel, U., Herold, M., Hofsteenge, J. & Kirschner, K. (1989). *Science*, **243**, 206–210.

McHarg, J., Kelly, S. M., Price, N. C., Cooper, A. & Littlechild, J. A. (1999). *Eur. J. Biochem.* **259**, 939–945.

McPhillips, T. M., Hsu, B. T., Sherman, M. A., Mas, M. T. & Rees, D. C. (1996). *Biochemistry*, **35**, 4118–27.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 499–501.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Pickover, C. A., McKay, D. B., Engelman, D. M. & Steitz, T. A. (1979). *J. Biol. Chem.* **254**, 11323–11329.

Protasova, N. Y., Kireeva, M. L., Murzina, N. V., Murzin, A. G., Uversky, V. N., Gryaznova, O. I. & Gudkov, A. T. (1994). *Protein Eng.* **7**, 1373–1377.

Read, R. J. (1986). *Acta Cryst.* A**42**, 140–149.

Ritco-Vonsovici, M., Minard, P., Desmadril, M. & Yon, J. M. (1995). *Biochemistry*, **34**, 16543–16551.

Sherman, M. A., Fairbrother, W. J. & Mas, M. T. (1992). *Protein Sci.* **1**, 752–760.

Szilagyi, A. N., Ghosh, M., Garman, E. & Vas, M. (2001). *J. Mol. Biol.* **306**, 499–511.

Tougard, P., Le, T. H., Minard, P., Desmadril, M., Yon, J. M., Bizebard, T., Lebras, G. & Dumas, C. (1996). *Protein Eng.* **9**, 181–187.

Watson, H. C., Wendell, P. L. & Scopes, R. K. (1971). *J. Mol. Biol.* **57**, 623–625.

Wilson, H. R., Williams, R. J., Littlechild, J. A. & Watson, H. C. (1988). *Eur. J. Biochem.* **170**, 529–538.

Yang, Y. R. & Schachman, H. K. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 11980–11984.

Zhang, T., Bertelsen, E., Benvegnu, D. & Alber, T. (1993). *Biochemistry*, **32**, 12311–12318.