Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Glutamate-5-kinase from *Escherichia coli*: gene cloning, overexpression, purification and crystallization of the recombinant enzyme and preliminary X-ray studies

Isabel Pérez-Arellano,[a] Fernando Gil-Ortiz,[b] Javier Cervera[a] and Vicente Rubio[b]*

[a]Fundación Valenciana de Investigaciones Biomédicas (FVIB), Spain, and [b]Instituto de Biomedicina de Valencia (IBV-CSIC), Spain

Correspondence e-mail: rubio@ibv.csic.es

Received 31 July 2004
Accepted 24 September 2004

Glutamate-5-kinase (G5K) catalyzes the first step of proline (and, in mammals, ornithine) biosynthesis. It is a key regulatory point of these routes, since it is the subject of feedback allosteric inhibition by proline or ornithine. The *Escherichia coli* gene (*pro*B) for G5K was cloned in pET22, overexpressed in *E. coli*, purified in a few steps in high yield to 95% homogeneity in the highly active proline-inhibitable form and was shown by cross-linking to be a tetramer. It was crystallized by the hanging-drop vapour-diffusion method at 294 K in the presence of ADP, $MgCl_2$ and L-glutamate using 1.6 *M* $MgSO_4$, 0.1 *M* KCl in 0.1 *M* MES pH 6.5 as the crystallization solution. The tetragonal bipyramid-shaped crystals diffracted to 2.5 Å resolution using synchrotron radiation. The crystals belong to space group $P4_{1(3)}2_12$, with unit-cell parameters $a = b = 101.1$, $c = 178.6$ Å, and contain two monomers in the asymmetric unit, with 58% solvent content.

1. Introduction

Microorganisms, plants and animals synthesize proline from glutamate in three steps catalyzed by glutamate 5-kinase (G5K), glutamyl 5-phosphate reductase (G5PR) and pyrroline 5-carboxylate reductase (Aral & Kamoun, 1997). Mammals also use G5K and G5PR to synthesize ornithine (Alonso & Rubio, 1989). Proline synthesis is an important process that has various implications, since proline, in addition to being a protein building block, plays an important role as an osmoprotectant in microorganisms (including pathogens) and plants; its intracellular levels increase in response to osmotic stress (Schwan *et al.*, 1998; Hong *et al.*, 2000). In addition, ornithine biosynthesis in mammals is crucial for proper ammonia detoxification, as one G5K mutation has been shown to cause human hyperammonaemia (Baumgartner *et al.*, 2000). G5K is an initial key catalyst and feedback-controlling point in these processes (Aral & Kamoun, 1997), but there are still limited data on this enzyme (Fujita *et al.*, 2003). It transfers the terminal phosphoryl group of ATP to the γ-carboxyl group of glutamate, activating this carboxylate by formation of the mixed carboxylic–phosphoric anhydride. Sequence comparisons have identified G5K as a member of the amino-acid kinase enzyme family (PF00696; http://www.sanger.ac.uk/Software/Pfam), which also includes acetylglutamate kinase, carbamate kinase, aspartokinase and UMP kinase. These enzymes, similarly to G5K, catalyze the formation of phosphoric anhydrides, generally with a carboxylate, and use ATP as the source of the phosphoryl group (Ramón-Maiques *et al.*, 2002). Although the three-dimensional structures of two enzymes of this family, carbamate kinase (Marina *et al.*, 1999; Ramón-Maiques *et al.*, 2000) and *Escherichia coli* acetylglutamate kinase (Ramón-Maiques *et al.*, 2002), have been determined, the structure of G5K would not only add valuable information concerning formation of the acyl-phosphate, but would also shed light on allosteric control, multi-enzyme architecture and gene expression. In contrast to the two enzymes of the amino-acid kinase family for which the three-dimensional structure is known, G5K is the subject of allosteric control either by proline (the case of the G5K studied here; Smith *et al.*, 1984) or by ornithine (Hu *et al.*, 1999). Secondly, G5K, together with G5PR, is part of a bifunctional polypeptide assembly that occurs both in plants and animals (Fujita *et al.*, 2003; Hu *et al.*, 1999). It has been suggested that owing to the instability of glutamyl 5-phosphate (Katchalsky & Paecht, 1954), even the microbial monofunctional G5Ks (*e.g. E. coli* G5K; Smith *et al.*, 1984) associate *in vivo* with G5PR to facilitate the channelling of glutamyl 5-phosphate (Seddon *et al.*, 1989). *E. coli* G5K has also been implicated in the control of gene expression (Ogura & Tanaka, 1996); this function has been attributed (Aravind & Koonin, 1999) to the presence of a 79-residue putative PUA-type domain at the enzyme C-terminus that is believed to be involved in RNA binding and for which only few examples have been characterized structurally so far (Ishitani *et al.*, 2002; Pan *et al.*, 2003). As a first

step towards structural insight into these processes and into proline synthesis, the cloning of the *E. coli* gene (*proB*) for G5K and the overexpression, purification, crystallization and initial X-ray diffraction of the recombinant enzyme, a 367-residue polypeptide (sequence-deduced weight 39.06 kDa), are reported here.

2. Materials and methods

2.1. Protein expression and purification

The *proB* gene was PCR-amplified from *E. coli* (strain DH5α, from Clontech) genomic DNA, using Expand High Fidelity PCR System (from Roche) with the primers 5′-CAGAGACATATGAGTGACAGCC-3′ and 5′-TTGTTCCAGGATCCGCCTGC-TCC-3′, which introduce *Nde*I and *Bam*HI sites at the initiator ATG and downstream of the stop codon. The *Nde*I- and *Bam*HI-digested amplified fragment was ligated using T4 ligase into the *Nde*I and *Bam*HI sites of plasmid pET22b (Novagen) and *E. coli* DH5α cells (from Clontech) were transformed. Plasmid pGKE was isolated and was shown by restriction analysis and automated DNA sequencing to harbour the full *proB* gene. *E. coli* BL21 (DE3) cells (from Novagen) were transformed with pGKE and grown to an A_{600} of 0.6 at 310 K in 0.5 l LB broth containing 75 µg ml^{-1} ampicillin. They were then induced with 0.5 m*M* isopropyl-β-D-thiogalactoside for 3 h and harvested by centrifugation. Subsequent steps were carried out at 277 K. The cells were suspended in 10 ml buffer *A* (50 m*M* Tris–HCl, 1 m*M* dithiothreithiol pH 7.2) and then sonicated and centrifuged (30 min, 35 000*g*). The supernatant was precipitated with $(NH_4)_2SO_4$ at 20% saturation, the pellet was resuspended in 20 ml buffer *A* and the solution was subjected, in two batches of 10 ml, to ion-exchange chromatography using an FPLC (Amersham Pharmacia) system fitted with a MonoQ 5/5 column equilibrated and run with buffer *A*. After washing with 15 ml buffer *A*, a 30 ml linear gradient of 0–0.5 *M* NaCl in buffer *A* was applied. Enzyme-rich fractions eluted at approximately 0.3 *M* NaCl. They were pooled and precipitated with $(NH_4)_2SO_4$ at 35% saturation and the pellet was kept at 193 K.

2.2. Crystallization and data collection

The sparse-matrix-sampling vapour-diffusion method (Jancarik & Kim, 1991) was used for crystallization tests, which were carried out at 294 K in drops hanging in multiwell plates utilizing commercial kits (Crystal Screens I and II from Hampton Research). The drops contained 1.5 µl reservoir solution and 1.5 µl of a 10 mg ml^{-1} G5K solution prepared by repeated centrifugal ultrafiltration (Microsep 10K, Pall Filtron) of the enzyme in 50 m*M* Tris–HCl pH 7.2, 1 m*M* dithioerythritol, 160 m*M* L-glutamate, 30 m*M* $MgCl_2$, 20 m*M* KCl and 6 m*M* ADP. The best crystals, which reached about 0.3 mm in the largest dimension, were grown at 294 K in about four to five months using 1.6 *M* magnesium sulfate, 0.1 *M* potassium chloride in 0.1 *M* MES pH 6.5 as reservoir solution. The same solution supplemented with 10% glycerol was used for crystal harvesting prior to flash-cooling in the gas flow of an Oxford cryosystem. X-ray data collection was carried out at 100 K using synchrotron radiation (ESRF, Grenoble; beamline BM16) and a MAR CCD detector. The data set was processed and scaled with *MOSFLM* (Leslie, 1990) and *SCALA* (Evans, 1997).

2.3. Enzyme activity and other assays

G5K activity was assayed at 310 K as glutamate-dependent ADP release, monitoring the $A_{340\,nm}$ in a solution containing 0.15 *M* sodium L-glutamate, 20 m*M* ATP, 80 m*M* $MgCl_2$, 1 m*M* dithiothreitol, 50 m*M* Tris–HCl pH 7.2, 5 m*M* phosphoenol-pyruvate, 0.25 m*M* NADH, 0.08 mg ml^{-1} pyruvate kinase and 0.03 mg ml^{-1} lactate dehydrogenase. Protein was determined with the Bradford assay (Bradford, 1976) using bovine serum albumin as standard and SDS–PAGE was performed according to Laemmli (1970). Cross-linking assays with dimethylsuberimidate (from Pierce) were carried out and analyzed by SDS–PAGE as originally described by Davies & Stark (1970).

3. Results and discussion

E. coli BL21 (DE3) cells transformed with plasmid pGKE, but not those transformed with the empty pET22b plasmid (Fig. 1*a*, lane 1), expressed G5K in large amounts in soluble form, as shown by the prominent 40 kDa band (expected weight 39 kDa) observed by SDS–PAGE (Fig. 1*a*, lane 2) and by the large G5K activity (data not shown) found in the centrifuged extracts of the pGKE-transformed cells. The over-expressed protein was isolated in few steps (see §2) and high yield (35 mg per litre of culture) to >95% purity (SDS–PAGE estimate; Fig. 1*a*, lane 3) and exhibited high G5K activity (106 ± 10 µmol min^{-1} mg^{-1}), higher than the highest activities (approximately 55 µmol min^{-1} mg^{-1}) reported previously for the naturally produced *E. coli* enzyme (Seddon *et al.*, 1989). As expected (Smith *et al.*, 1984), the enzyme activity was completely inhibited by the addition of 1 m*M* L-proline. Automated Edman sequencing yielded the expected N-terminal sequence (SDSQTLVV) except for the lack of the initial methionine, which is removed post-translationally. MALDI–TOF mass spectrometry of the purified protein yielded a molecular weight for the polypeptide of 38 927 Da, which is in excellent agreement

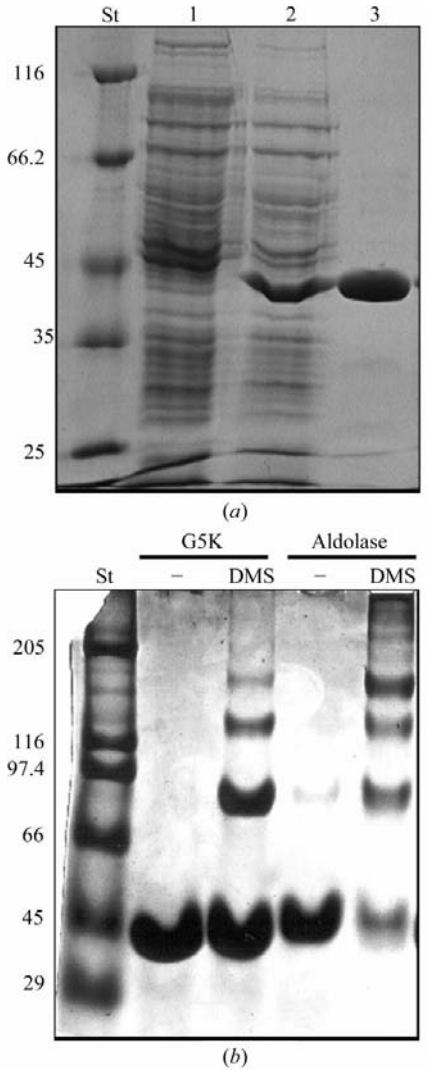

Figure 1
Expression, purification and cross-linking of G5K. (*a*) SDS–PAGE of cell extracts of *E. coli* BL21(DE3) transformed with plasmids pET22 (lane 1) or pGKE (lane 2). Lane 3, purified enzyme. (*b*) Cross-linking of G5K or of aldolase (used as a positive control for a tetramer; subunit weight 40 kDa), as reported by Davies & Stark (1970), at a concentration of the indicated proteins of 0.5 mg ml^{-1}. The minus sign and DMS indicate the absence or the presence in the mixture of the cross-linking agent dimethylsuberimidate. In both (*a*) and (*b*) the left lane corresponds to protein standards (St) of the weights (in kDa) indicated at the side.

with the sequence-deduced molecular weight (38 926 Da without Met1). New bands consistent with dimers, trimers and tetramers were observed by SDS–PAGE after cross-linking with dimethylsuberimidate (DMS; Fig. 1*b*, labelled DMS under G5K). This indicates that *E. coli* G5K forms tetramers under the conditions of the cross-linking assays.

Bipyramidal-shaped crystals of approximately 0.3 mm maximal dimension were obtained in the presence of glutamate and MgADP at 294 K (Fig. 2*a*), although the crystal growth was very slow (>4 months). X-ray data were collected to 2.5 Å resolution with the use of synchrotron radiation (Table 1 and Figs. 2*b* and 2*c*). The space group was identified as tetragonal $P4_{1(3)}2_12$. Packing-density calculations (Matthews, 1968) for the observed cell size ($a = b = 101.1$, $c = 178.6$ Å) and for the monomeric weight of 38 926 Da agree with the presence of two enzyme monomers in the asymmetric unit ($V_M = 2.92$ Å^3 Da^{-1}, giving a solvent content of 58%). If the enzyme is tetrameric, as suggested by the cross-linking studies, and since carbamate kinase and *E. coli* acetylglutamate kinase are homodimers (Marina *et al.*, 1999; Ramón-Maiques *et al.*, 2000, 2002), the G5K tetramer might be a dimer of dimers. The $\chi = 90$ and 180° sections of the self-rotation function (Rossmann & Blow, 1962) calculated with *MOLREP* (Vagin & Teplyakov, 1997) clearly show strong peaks corresponding to the fourfold and twofold axes of space group $P4_{1(3)}2_12$ (Fig. 3). In addition, two weaker peaks are also present in the $\chi = 180°$ section at angles of 23 and 67° from c^*. These peaks could correspond to a non-crystallographic molecular dyad axis and would be in agreement with a tetrameric organization of the enzyme molecule, with two monomers in the asymmetric unit and a molecular dyad coincident with the crystal twofold along the diagonal in the *ab* (or *c*) face. Attempts at finding an initial solution by molecular replacement [using the programs *AMoRe* (Navaza, 1994) and *MOLREP* (Vagin & Teplyakov, 1997)] with the polyalanine model encompassing residues 1–258 of acetylglutamate kinase from *E. coli* (Ramón-Maiques *et al.*, 2002) as a search model are under way, but they have not yet given a conclusive solution. Furthermore, selenomethionine-substituted protein is presently being prepared.

Table 1
X-ray data-collection statistics.

Values in parentheses are for the highest resolution shell (2.64–2.50 Å).

Beamline	ESRF BM16
Wavelength (Å)	0.979
Space group	$P4_12_12$ or $P4_32_12$
Unit-cell parameters (Å)	$a = b = 101.1$, $c = 178.6$
Resolution range (Å)	48.8–2.5
Total No. reflections	367219
Unique reflections	61566
Data completeness (%)	99.5 (99.5)
R_{sym}† (%)	9.1 (41.6)
$I/\sigma(I)$	6.8 (1.7)

† $R_{sym} = \sum|I - \langle I\rangle|/\sum I$, where I is the observed intensity and $\langle I\rangle$ is the average intensity of multiple observations of symmetry-related reflections.

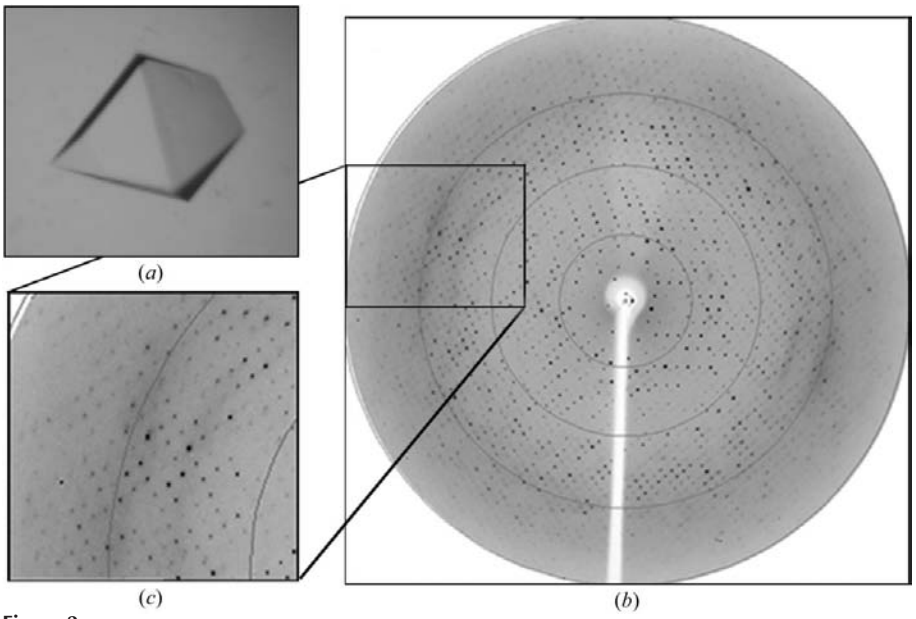

Figure 2
Typical crystal (*a*) and diffraction pattern (*b*, *c*) of *E. coli* G5K. The arcs indicate 2.4, 3.1, 4.7 and 9.5 Å resolution, respectively.

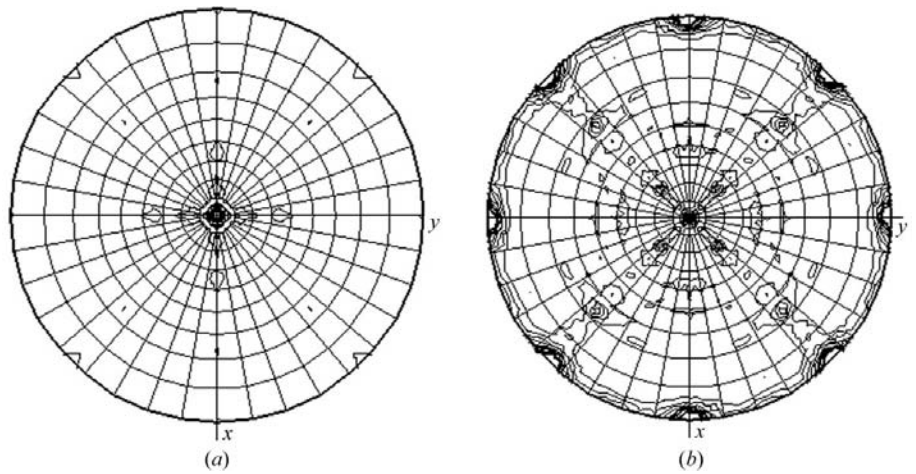

Figure 3
Representation of the (*a*) $\chi = 90°$ and (*b*) $\chi = 180°$ sections of the self-rotation function from *E. coli* G5K crystals. The start level is 1σ and the step size is 0.75σ. The orthogonalization code follows the PDB convention. See text for interpretation.

Data collection was carried out at ESRF (EMBL Grenoble) under the auspices of the EU. This work was supported by grants BMC2001-2182 and BFU2004-04472 of the Spanish Ministry of Education and Science and GV04B/587 of the Generalitat Valenciana. IP-A and FG-O were post-doctoral fellows of the FVIB-Bancaja and Fundación Ferrer, respectively. We thank J. J. Calvete (Instituto de Biomedicina de Valencia) for N-terminal sequencing and mass spectrometry and I. Fita (IBMB-CSIC) for critical reading of the manuscript and for helpful suggestions.

References

Alonso, E. & Rubio, V. (1989). *Biochem. J.* **259**, 131–138.
Aral, B. & Kamoun, P. (1997). *Amino Acids*, **13**, 189–217.
Aravind, L. & Koonin, E. V. (1999). *J. Mol. Evol.* **48**, 291–302.

Baumgartner, M. R., Hu, C. A., Almashanu, S., Steel, G., Obie, C., Aral, B., Rabier, D., Kamoun, P., Saudubray, J. M. & Valle, D. (2000). *Hum. Mol. Genet.* **22**, 2853–2858.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Davies, G. E. & Stark, G. R. (1970). *Proc. Natl Acad. Sci. USA*, **66**, 651–656.
Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.
Fujita, T., Maggio, A., Garcia-Rios, M., Stauffacher, C., Bressan, R. A. & Csonka, L. N. (2003). *J. Biol. Chem.* **278**, 14203–14210.
Hong, Z., Lakkineni, K., Zhang, Z. & Verma, D. P. S. (2000). *Plant Physiol.* **122**, 1129–1136.
Hu, C. A., Lin, W. W., Obie, C. & Valle, D. (1999). *J. Biol. Chem.* **274**, 6754–6762.
Ishitani, R., Nureki, O., Fukai, S., Kijimoto, T., Nameki, N., Watanabe, M., Kondo, H., Sekine, M., Okada, N., Nishimura, S. & Yokohama, S. (2002). *J. Mol. Biol.* **318**, 665–677.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Katchalsky, A. & Paecht, M. (1954). *J. Am. Chem. Soc.* **76**, 6042–6044.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Leslie, A. G. W. (1990). *Crystallographic Computing.* Oxford University Press.
Marina, A., Alzari, P. M., Bravo, J., Uriarte, M., Barcelona, B., Fita, I. & Rubio, V. (1999). *Protein Sci.* **8**, 934–940.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Ogura, M. & Tanaka, T. (1996). *J. Bacteriol.* **178**, 216–222.
Pan, H., Agarwalla, S., Moustakas, D. T., Finer-Moore, J. & Stroud, R. M. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 12648–12653.
Ramón-Maiques, S., Marina, A., Gil, F., Fita, I. & Rubio, V. (2002). *Structure*, **10**, 329–340.
Ramón-Maiques, S., Marina, A., Uriarte, M., Fita, I. & Rubio, V. (2000). *J. Mol. Biol.* **299**, 463–476.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Schwan, W. R., Coulter, S. N., Ng, E. Y., Langhorne, M. H., Ritchie, H. D., Brody, L. L., Westbrock-Wadman, S., Bayer, A. S., Folger, K. R. & Stover, C. K. (1998). *Infect. Immun.* **66**, 567–572.
Seddon, A. P., Zhao, K. Y. & Meister, A. (1989). *J. Biol. Chem.* **264**, 11326–11335.
Smith, C. J., Deutch, A. H. & Rushlow, K. E. (1984). *J. Bacteriol.* **157**, 545–551.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022-1025.