Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Initiating a structural study of 2-keto-3-deoxy-6-phosphogluconate aldolase from *Escherichia coli*

Louise V. Buchanan,[a] Nupur Mehta,[b] Luka Pocivavsek,[b] S. Niranjanakumari,[b] Eric J. Toone[b] and James H. Naismith[a]*

[a]Centre for Biomedical Sciences, North Haugh, The University, St Andrews, Fife KY16 9ST, Scotland, and [b]Department of Chemistry, Duke University, Durham, North Carolina, USA

Correspondence e-mail: naismith@st-and.ac.uk

Received 7 July 1999
Accepted 25 August 1999

2-Keto-3-deoxy-6-phosphogluconate aldolase (KDPG aldolase, E.C. 4.1.2.14) is a member of the pyruvate/phospho*enol*pyruvate aldolase family. It is also a synthetically useful enzyme, capable of catalyzing the stereoselective aldol addition of pyruvate to a range of unnatural electrophilic substrates. The recombinant protein was purified by a two-step HPLC protocol involving anion-exchange and hydrophobic chromatography. Dynamic light-scattering experiments indicated the protein to be monodisperse. Crystals were obtained using the sitting-drop vapour-diffusion method, with PEG 6K as precipitant. Diffraction data were collected on a frozen crystal to a resolution of 2.26 Å on station PX9.6 at the Daresbury synchrotron. The crystal belongs to space group $P2_12_12_1$, with unit-cell parameters a = 53.2, b = 77.9, c = 146.8 Å.

1. Introduction

Naturally occurring enzymes that catalyze aldol addition reactions are firmly established as powerful tools for stereocontrolled carbon–carbon bond formation in organic synthesis (Fessner & Walter, 1997; Wong *et al.*, 1995). Aldolases have been used in numerous efficient syntheses of stereochemically complex molecules, most notably compounds of importance in the pharmaceutical industry and in the field of carbohydrate synthesis (Gijsen *et al.*, 1996).

Aldolases can be categorized according to nucleophile type and include the dihydroxyacetone phosphate, acetaldehyde and pyruvate/phospho*enol*pyruvate aldolases (Fessner, 1998; Henderson & Toone, 1999). 2-Keto-3-deoxy-6-phosphogluconate aldolase (KDPG aldolase, E.C. 4.1.2.14), a key component of the Entner–Doudoroff glycolysis pathway, is a member of the pyruvate/phospho*enol*pyruvate family of aldolases. Unlike the dihydroxyacetone phosphate aldolases, KDPG aldolase operates exclusively under kinetic control to provide stereochemically pure products (Shelton *et al.*, 1996; Fig. 1). KDPG aldolase catalyzes the stereospecific aldol addition of pyruvate to a range of unnatural electrophilic substrates at rates practical for preparative-scale organic synthesis. It is interesting to note that KDPG aldolase is a constitutive enzyme, while the related 2-keto-3-deoxy-6-phosphogalactonate (KDPGal) aldolase is inducible and is expressed only during growth on galactose or galactonate. A separate classification for aldolases is by group according to catalytic mechanism: class I (lysine Schiff base) and class II (metal-dependent). The mechanism of each class of enzyme is under active investigation (Hall *et al.*, 1999; Blom & Sygush, 1997).

The X-ray structure of KDPG aldolase from *Pseudomonas putida* was solved to 3.5 Å in 1976 (Mavridis & Tulinsky, 1976) and subsequently refined to 2.8 Å in 1982; these coordinates were not deposited (Mavridis *et al.*, 1982). The structure of this enzyme is the classical TIM barrel arrangement (Banner *et al.*, 1975) of eight α-helices around a barrel of eight β-strands. The enzyme is a class I aldolase and functions *via* a Schiff base at the active-site lysine residue. This residue is found inside the barrel. The KDPG aldolase from *P. putida* is a trimer in the crystal and to our knowledge is the only trimeric aldolase. Aldolases from other sources are dimeric or tetrameric (for references, see Fessner, 1998). The enzymes from *P. putida* and *E. coli* are closely related and share 45% sequence identity. To assist in our efforts to enhance the synthetic utility of KDPG aldolase by both rational and random modification, we require detailed structural information regarding the active site of the enzyme. In particular, the details of the substrate-recognition residues are not known experimentally for KDPG aldolase. Here, we describe the purification, crystallization and preliminary structural studies of *E. coli* KDPG aldolase.

2. Overexpression and purification

BL21(DE3) pLysS cells were transformed with a plasmid pTC190 (Egan *et al.*, 1992) containing the gene for the *E. coli* aldolase. In the presence of ampicillin the cells constitutively express the protein. Cells were grown at

Table 1
Data collection on PX 96 at the CLRC.

Resolution range was 30–2.26 Å. Values in parentheses refer to the highest resolution shell, 2.37–2.26 Å. Diffraction was observed to 2.10 Å; however, this is in the corners of the square CCD and data from 2.26–2.10 Å is approximately 50% complete. This adds a further 3533 unique reflections.

Space group	$P2_12_12_1$
Unit-cell parameters	a = 53.2, b = 77.79, c = 146.83 Å; $\alpha = \beta = \gamma = 90^\circ$
Solvent content (%)	62 (dimer) or 44 (trimer)
Unique reflections	27601 (2318)
$I/\sigma(I)$	18 (4)
Average redundancy	5.1 (2.5)
Data completeness (%)	94 (80)
$R_{\rm merge}$ (%)	5.7 (14.2)

310 K in Terrific Broth (Maniatis *et al.*, 1982) containing 100 µg ml^{-1} ampicillin and 50 µg ml^{-1} chloramphenicol to an OD_{600} of 0.5. The cells were harvested by centrifugation (15 min, 8000*g*, 277 K) and frozen. They were then resuspended in 20 m*M* Tris–HCl pH 8, 1 m*M* EDTA, 1 m*M* DTT. After incubation at room temperature for 30 min, the viscosity of the mixture was reduced by addition of 20 µg ml^{-1} DNAse I. To ensure complete lysis, the cells were sonicated on ice (six cycles of 30 s sonication followed by 30 s rest). Cell debris was removed by centrifugation (20 min, 18000*g*, 277 K). The supernatant was brought to 20% ammonium sulfate saturation and centrifuged (20 min, 18000*g*, 277 K). The supernatant was dialysed against three changes of buffer (20 m*M* Tris–HCl pH 8.0) and filtered through a 0.2 µm membrane. The resulting solution was passed over a POROS-HQ HPLC column (BioCad Sprint) and protein was eluted with a 0–1000 m*M* NaCl gradient. A protein with a molecular weight equivalent to KDPG aldolase eluted at 125 m*M* NaCl. Fractions corresponding to this peak were pooled and concentrated under nitrogen pressure in an Amicon ultrafiltration cell. Ammonium sulfate was added gradually to 30% saturation. The filtered sample was loaded on a POROS high-density phenyl HPLC column (BioCad Sprint) equilibrated in 30% $(NH_4)_2SO_4$, 20 m*M* Na_2HPO_4 pH 8.0. Elution was achieved with a decreasing gradient of $(NH_4)_2SO_4$. The target protein eluted at 10% $(NH_4)_2SO_4$.

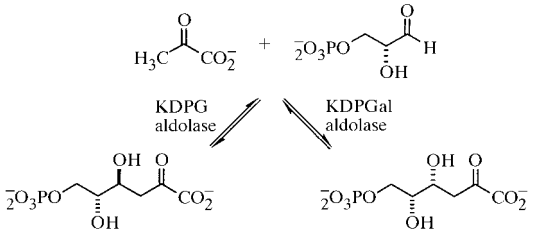

Figure 1
Reaction catalyzed by KDPG aldolase.

3. Protein analysis

Following HPLC purification, the protein appeared as a single band on both SDS–PAGE and IEF gels (silver stain). Dynamic light-scattering results indicated the protein to be monodisperse and to have an apparent molecular mass of 52 kDa, which is indicative of a dimer. N-terminal sequencing confirmed the crystallized protein to be fully intact KDPG aldolase. The yield was 8 mg of aldolase per litre of Terrific Broth (Maniatis *et al.*, 1982) by the Bradford method (Bradford, 1976).

4. Crystallization

After the final purification step, KDPG aldolase was dialysed against three changes of buffer (20 m*M* Tris–HCl pH 8.0) and concentrated to 3 mg ml^{-1}. Prior to crystallization, DTT was added to a final concentration of 4 m*M*. Crystals (Fig. 2) were grown using the sitting-drop vapour-diffusion method (Ducruix & Giegé, 1992), with 2.5 µl protein and 2.5 µl mother liquor (20% PEG 6K, 0.075 *M* citric acid pH 4.0) used in the drops.

5. Data collection

Data were collected from a single crystal on station PX9.6 using the ADSC Quantum-4 CCD detector at the CCLRC synchrotron at Daresbury. The crystal was soaked for 120 s in the crystallization solution containing 15% glycerol prior to data collection. Data were recorded at 100 K as 180 90 s 1° non-overlapping images. The crystal-to-detector distance was 230 mm and the wavelength was set to 0.87 Å. Data were indexed and integrated with *DENZO* version 1.12d (Otwinowski & Minor, personal communication) and merged with *SCALEPACK* (Otwinowski, 1993). Data could be processed and observed to 2.1 Å resolution; however, completeness at this resolution is affected by the square geometry of the detector. We, therefore, report statistics to 2.26 Å, where completeness is still high. The space group was assigned as $P2_12_12_1$ following observation of a $2n$ condition along all three axis (at least 11 systematically absent reflections were measured for each axis). Table 1 summarizes the quality of the data.

6. Molecular replacement

The unit cell of the crystal gives a Matthews number (Matthews, 1968) of 3.3 Å^3 Da^{-1} for a dimer in the asymmetric unit and 2.2 Å^3 Da^{-1} for a trimer, both of which are feasible. From self-rotation analysis it is unclear whether a dimer or a trimer is present. As the space group has no twofold axis, to have one and a half dimers in the asymmetric unit is not possible. Given the enzymes have 45% sequence identity (Fig. 3), it is inconceivable that our structure does not share the same fold and is not structurally very similar to that observed for *P. putida* aldolase. Professor Tulinsky kindly provided us with 2.2 Å refined, but unpublished, coordinates of the enzyme from *P. putida.* Attempts to find a molecular-replacement solution using the trimer observed for *P. putida* have been unsuccessful. We excluded variable loops and removed non-conserved side chains, with all calculations being performed with the *AMoRe* program (Navaza, 1994) as implemented in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Despite extensive testing of resolution ranges, no solution has emerged using a trimeric search model. Similar failure has characterized our attempts with a monomeric search model. The reasons for these failures will only become apparent once our structure has been determined. We were, of course, unable to construct a dimer based on the trimeric *P. putida* structure.

7. Selenomethionine variant

The cells containing constitutive expression vector were unable to grow satisfactorily on

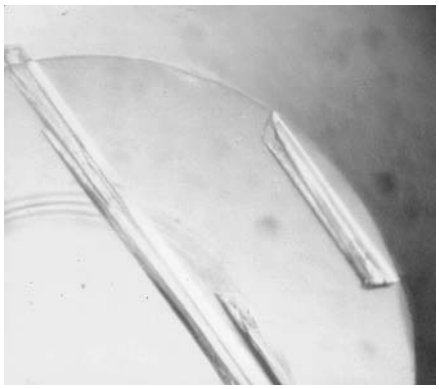

Figure 2
KPDG aldolase crystals; maximum dimensions were 0.6 × 0.3 × 0.2 mm.

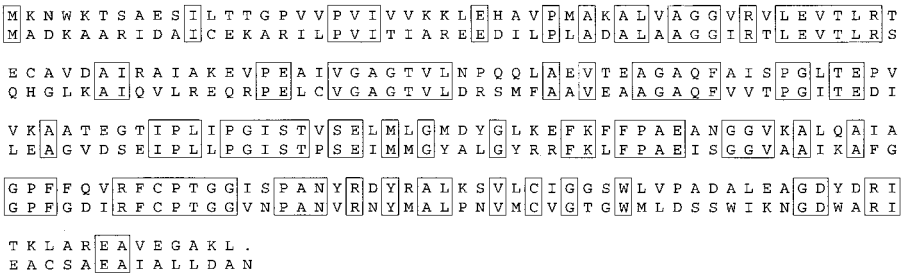

Figure 3
Alignment displaying the sequence homology between KDPG aldolase from *E. coli* (top) and *P. putida* (bottom). The conserved residues are boxed.

a selenomethionine-containing minimal media, although the cells did grow on methionine-containing minimal media. We re-cloned the enzyme cDNA into a pET-28b(+) vector. This vector produces His-tagged protein after induction with IPTG. The protein was purified in an analogous manner to that already described. The fraction obtained after anion exchange was divided in two. One half was immediately purified by hydrophobic chromatography, while the other half was digested with thrombin overnight, thus removing the His-tag. The cleaved protein was then purified by hydrophobic chromatography. The integrity of both fractions was confirmed by N-terminal sequencing. Only the protein lacking the His-tag crystallized. The methionine auxotrophic *E. coli* strain B834 was transformed with the expression vector and grown on rich selenomethionine-containing media (Budisa *et al.*, 1997). Se-Met protein was purified by the same method as the native protein, incorporating thrombin treatment. The presence of selenium was confirmed by mass spectroscopy. Se-Met crystals have been obtained and are not isomorphous with the native data. Se atoms, and thus methionines, will require location by anomalous diffraction methods. Two methionines are conserved between the *P. putida* and *E. coli* enzymes (Fig. 3). From simple modelling, we can predict the locations of the other two methionines in the secondary structure. This should allow us to position the *P. putida* structure in our unit cell and begin refinement. If this fails, we will determine the structure by multiwavelength anomalous dispersion methods.

We are grateful to Professor Tulinsky for allowing us access to the 2.2 Å refined data of the enzyme from *P. putida*. We also acknowledge Professor Tyrell Conway for the generous gift of pTC191 containing the *E. coli* aldolase gene and Daresbury Laboratory for providing synchrotron-radiation facilities. We thank W. Minor and Z. Otwinowski for providing us with *HKL* version 1.12. We thank the editor and referees for their helpful suggestions in revising this manuscript.

References

Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. I., Milman, J. D., Offord, R. E., Priddle, J. D. & Waley, S. G. (1975). *Nature (London)*, **255**, 609–614.
Blom, N. & Sygush, J. (1997). *Acta Cryst.* D**54**, 461–466.
Bradford, M. M. (1976). *Anal. Biochem.* **72**, 248–254.
Budisa, N., Karnbrock, W., Steinbacher, S., Humm, A., Prade, L., Neuefeind, T., Moroder, L. & Huber, R. (1997). *J. Mol. Biol.* **270**, 616–623.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Ducruix, A. & Giegé, R. (1992). *Crystallization of Nucleic Acids and Proteins.* Oxford: IRL Press.
Egan, S. E., Fliege, R. A, Tong, S., Shibata, A., Wolf, R. E. Jr & Conway, T. (1992). *J. Bacteriol.* **174**, 4638–4646.
Fessner, W. D. (1998). *Curr. Opin. Chem. Biol.* **2**, 85–97.
Fessner, W. D. & Walter, C. (1997). *Top. Curr. Chem.* **184**, 97–194.
Gijsen, H. J. M., Qiao, L., Fitz, W. & Wong, C. H. (1996). *Chem. Rev.* **96**, 443–473.
Hall, D. R., Leonard, G. A., Reed, C. D., Watt, C. I., Berry, A. & Hunter, W. N. (1999). *J. Mol. Biol.* **287**, 383–394.
Henderson, D. P. & Toone, E. J. (1999). *Comprehensive Natural Products Chemistry*, Vol. 3, *Carbohydratres and their Derivatives Including Tannins, Cellulose and Related Lignins*, edited by B. M. Pinto, pp. 367–441. New York: Elsevier.
Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mavridis, I. M., Hatada, M. H., Tulinsky, A. & Lebioda, L. (1982). *J. Mol. Biol.* **162**, 419–444.
Mavridis, I. M. & Tulinsky, A. (1976). *Biochemistry*, **15**, 4410–4417.
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.
Shelton, M. C., Cotterill, I. C., Novak, S. T. A., Poonawala, R. M., Sudarshan, S. & Toone, E. J. (1996). *J. Am. Chem. Soc.* **118**, 2117–2125.
Wong, C. H., Halcomb, R. L., Ichikawa, Y. & Kajimoto, T. (1995). *Angew. Chem. Int. Ed. Engl.* **34**, 521–546.