Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray analysis of shikimate dehydrogenase from *Escherichia coli*

John Maclean,[a] Samantha A. Campbell,[b] Kevin Pollock,[b] Sureka Chackrewarthy,[b] John R. Coggins[b] and Adrian J. Lapthorn[a]*

[a]Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, and [b]Division of Biochemistry and Molecular Biology, Institute of Biological and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland

Correspondence e-mail: a.lapthorn@chem.gla.ac.uk

Received 14 October 1999
Accepted 9 February 2000

Shikimate dehydrogenase from *Escherichia coli* has been crystallized by the vapour-diffusion method using ammonium sulfate as a precipitant. Mass spectrometry confirmed the purity of the enzyme and dynamic light scattering was used to find the appropriate additives to yield a monodisperse enzyme solution. The crystals are monoclinic, space group $C2$, with unit-cell parameters a = 110.0, b = 139.8, c = 102.6 Å, β = 122.2° (at 100 K). Native crystals diffract to 2.3 Å in-house on a rotating-anode X-ray source. The asymmetric unit is likely to contain four molecules, related by 222 symmetry, corresponding to a packing density of 2.86 Å^3 Da^{-1}.

1. Introduction

The shikimate pathway is essential to plants and microorganisms for the biosynthesis of aromatic compounds, but is absent from animals. The enzymes of this pathway are therefore potential targets for the development of non-toxic antimicrobial agents (Davies *et al.*, 1994) and herbicides (Coggins, 1989).

Shikimate dehydrogenase (SDH; E.C. 1.1.1.25) catalyses the fourth reaction in the pathway, the NADPH-dependent reduction of 3-dehydroshikimate to shikimate. The protein from *E. coli* K12 has been purified to homogeneity (Chaudhuri & Coggins, 1985) and shown to be a monomer of molecular weight 30 kDa. The gene (*aroE*) from *E. coli* for shikimate dehydrogenase has been cloned, sequenced and overexpressed at moderate levels (Anton & Coggins, 1988).

SDHs represent a distinct group of dehydrogenases based on their amino-acid sequences. The only other obviously related dehydrogenases at the sequence level are the inducible quinate (shikimate) dehydrogenases from fungi and a gene product ydiB (of unknown function) found in the bacteria *E. coli* and *Haemophilus influenzae*. In most bacteria, SDH is present as a monomer, which is quite uncommon for dehydrogenases (Chaudhuri & Coggins, 1985). In plants, SDH is linked with a type I dehydroquinase, forming a bifunctional enzyme (Deka *et al.*, 1994), while in fungi such as *Neurospora crassa*, SDH forms the fourth domain of the pentafunctional AROM polypeptide (Lumsden & Coggins, 1977), which catalyses five of the seven steps of the shikimate pathway.

The structure analysis of shikimate dehydrogenase will provide information about substrate recognition by the enzyme and will suggest how inhibitors could be designed. The relatively small size of SDH and the fact that it is a monomer makes the enzyme of interest to those wishing to investigate NADP-linked dehydrogenases.

2. Overexpression of *aroE*

The gene *aroE* has previously been expressed in a construct called pIA321 (Anton & Coggins, 1988). This was produced by inserting a 1.27 kbp fragment encoding the *aroE* gene into the plasmid pKK223 and was expressed in an *E. coli aroE* auxotrophic mutant AB2834. Expression in this system results in moderate expression of shikimate dehydrogenase with a 300 times higher specific activity for the enzyme in the crude extract in comparison with native *E. coli* K12.

PCR primers were designed to incorporate restriction sites for the enzymes *Nde*I and *Bgl*II. PCR was performed with DNA from pIA321 and Vent polymerase, resulting in a single fragment of 0.82 kbp. This fragment was cloned into pTB361, a T7 RNA polymerase expression plasmid (Studier & Moffatt, 1986). Shikimate dehydrogenase expressed from this construct accounted for 30% of the total cellular protein.

3. Enzyme purification and mass-spectrometric analysis

Cells (14 g wet weight) were suspended in 20 ml 10 m*M* Tris–HCl pH 7.5 containing 0.4 m*M* DTT and 1.3 m*M* EDTA (buffer *A*) and were broken by two passes through a French pressure cell operated at 6.55 MPa. This material was then diluted with 80 ml buffer *A*; 0.5 mg of DNAase was added and the mixture was left stirring at 277 K; it was then centrifuged at 17 000 rev min^{-1} for 1 h. Shikimate dehy-

Table 1
Dynamic light-scattering measurements of *E. coli* shikimate dehydrogenase.

The experiments were carried out on the same day using 1 mg ml^{-1} of enzyme at 293 K. The values quoted are the mean values of at least 15 measurements. The hydrodynamic radius represents the median particle size present in the sample cell. The polydispersity indicates the standard deviation of the spread of particle sizes about the reported average radius. An increase in polydispersity relative to the average radius represents greater spread in the size distribution. The estimated M_w is calculated from the hydrodynamic radius (R_H) using an empirically derived relationship between the R_H and M_w values for a number of well characterized globular proteins in buffered solution.

Additives	Hydrodynamic radius (nm)	Polydispersity (nm)	Estimated M_w	Count (s^{-1})
None	2.8	1.2	35000	49300
NADPH (5 m*M*)	2.2	—	20000	31000
Shikimate (5 m*M*)	4.6	1.8	120000	57200
NADPH (5 m*M*) + shikimate (5 m*M*)	7.1	4.0	340000	49300

drogenase was purified from the resulting cell-free extract. All subsequent steps were performed at 277 K.

Solid $(NH_4)_2SO_4$ was added to the supernatant to 30% saturation. The mixture was stirred for 15 min and the precipitate was then removed by centrifugation at 17 000 rev min^{-1} for 15 min. The supernatant was adjusted to 55% saturation with solid $(NH_4)_2SO_4$ (162 g l^{-1}) and stirred for 15 min, followed by centrifugation as before. The precipitated protein was collected and dissolved in buffer *A* and the sample was dialysed overnight against 50 m*M* Tris–HCl pH 7.5, 0.4 m*M* DTT and 50 m*M* KCl (buffer *B*).

The protein was loaded on to a DEAE-Sephacel anion-exchange column (21 × 14 cm, 36 ml h^{-1} flow rate) which had been equilibrated with buffer *B*. The column was then washed with the same buffer until the A_{280} of the eluate was less than 0.3. SDH was eluted with a linear gradient of 50–350 m*M* KCl in 700 ml buffer *B*; 6 ml fractions were collected and assayed as described by Chaudhuri & Coggins (1985). Fractions containing SDH activity were pooled and dialysed overnight against 25 m*M* Tris–HCl pH 7.5, 0.4 m*M* DTT (buffer *C*).

The enzyme solution was concentrated by ultra-filtration against 50 m*M* Tris–HCl pH 7.5 containing 500 m*M* KCl, 0.4 m*M* DTT and 20%(*v*/*v*) glycerol (buffer *D*). The concentrated protein was applied to a Sephacryl S200 (superfine grade) column (8.5 × 2.1 cm, 6 ml h^{-1} flow rate) which had been equilibrated with buffer *D*. The enzyme was eluted with buffer *D*. Fractions containing SDH activity were pooled and dialysed overnight against buffer *C*.

The dialysed enzyme was loaded on to an ADP-Sepharose column (5 ml bed volume) which had been equilibrated with buffer *C*. The column was eluted with buffer *C* containing 0.1 m*M* NADP. Fractions containing SDH activity were dialysed against buffer *C* containing 50%(*v*/*v*) glycerol before long-term storage at 253 K.

Mass spectrometry was performed on a VG Platform quadrupole mass spectrometer fitted with a pneumatically assisted electrospray source (VG Biotech Ltd, Altrincham, Cheshire, England). For experimental details, see Krell *et al.* (1995). The purified protein was shown to be of high purity: a single species of molecular mass 29 414 Da was observed in the electrospray mass-spectrometry spectra, compared with a molecular weight of 29 413.64 Da calculated from the DNA sequence.

4. Dynamic light scattering

Experiments were carried out using a DYNA-PRO 801 dynamic light-scattering/molecular-sizing instrument (Protein Solutions, Buckinghamshire, England). Protein solutions (1 mg ml^{-1}) were in 50 m*M* Tris–HCl pH 7.6. Dynamic light-scattering experiments were carried out in order to assess which combination of product and co-factor were appropriate to obtain a monodisperse enzyme solution. Monodispersity is not an infallible indicator that diffraction-quality crystals can be obtained from a protein; instead, it reflects the observation that monodisperse protein solutions have a much higher probability of producing crystals than those which are not (Ferré-D'Amaré & Burley, 1997). Although several conditions were tried (Table 1), it was shown that monodispersity could only be achieved in the presence of NADPH. Inclusion of the product shikimate with or without co-factor led to significant aggregation. In the presence of NADPH, the estimated molecular weight obtained,

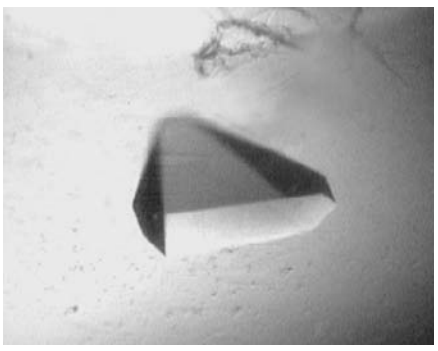

Figure 1
A crystal of *E. coli* shikimate dehydrogenase. The crystal is approximately 0.3 × 0.3 × 0.15 mm in size.

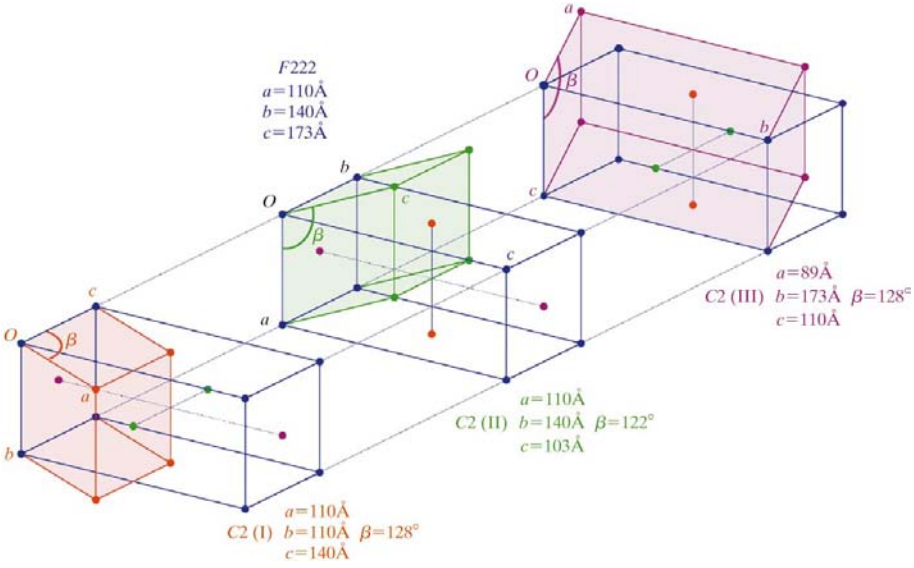

Figure 2
The three possible monoclinic cells (I–III) are shown in red, green and purple. They are superimposed on the original *F*222 cell, which is shown in blue. The *F*222 lattice points corresponding to the face centering are highlighted according to their use in the monoclinic lattices. The cell marked *C*2 (II) and shown in green is the *C*2 cell chosen as the true unit cell.

Table 2
Statistics for the X-ray data.

(*a*) Processed in *F*222.

d_{min} (Å)	R_{sym} (*I*)	(I/σ)	No. of unique reflections	Completeness (%)	Multiplicity (%)
4.91	0.068	26.5	3144	98.4	3.4
3.90	0.072	26.2	3030	98.0	3.2
3.41	0.095	20.8	3019	98.4	3.2
3.09	0.146	14.5	3029	99.1	3.5
2.87	0.230	9.2	3032	99.5	3.7
2.70	0.290	6.7	3006	99.5	3.8
2.57	0.356	5.5	3035	99.8	3.8
2.46	0.411	4.3	3036	99.9	3.8
2.36	0.466	3.5	3050	99.9	3.7
2.28	0.521	2.9	2238	74.9	3.5
Total	0.141	14.1	29619	98.3	3.6

(*b*) Processed in *C*2.

d_{min} (Å)	R_{sym} (*I*)	(I/σ)	No. of unique reflections	Completeness (%)	Multiplicity (%)
4.91	0.040	22.5	5741	95.0	2.2
3.90	0.046	21.8	5854	97.8	2.1
3.41	0.059	16.9	5878	98.7	2.1
3.09	0.083	11.4	5923	99.4	2.1
2.87	0.122	7.0	5940	99.7	2.0
2.70	0.164	5.1	5886	99.8	2.0
2.57	0.197	4.2	5942	99.7	2.0
2.46	0.244	3.3	5925	99.4	2.0
2.36	0.285	2.7	5903	99.2	2.0
2.28	0.368	2.3	5546	94.3	1.9
Total	0.081	10.7	58538	98.3	2.0

20 000 Da, was a third lower than the molecular weight of the enzyme, but was consistent with the expected monomeric state of the enzyme. The apoenzyme gave a higher molecular-weight estimate, suggesting that addition of the cofactor results in significant closure of the structure, as would be expected. In the light of these results, crystallization trials were carried out in the presence of NADPH.

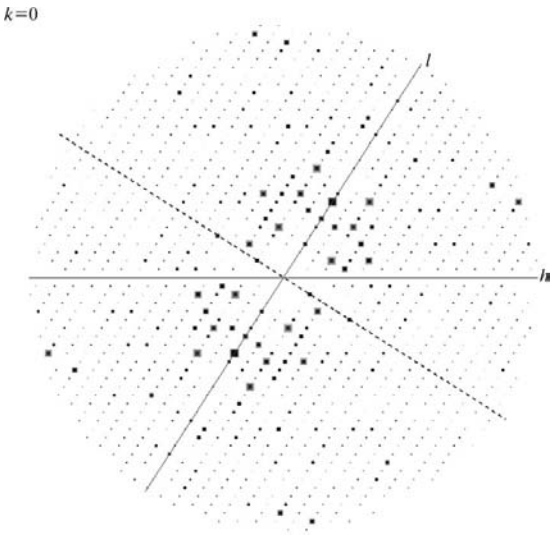

Figure 3
A pseudo-precession plot of the *h*0*l* zone using data to 3.4 Å calculated using the program *HKLVIEW* (Collaborative Computational Project, Number 4, 1994). In addition to *h*0*l* = $\bar{h}0\bar{l}$, an NCS twofold axis (marked with a dashed line) perpendicular to *l* makes *h*0*l* ≃ (*h*0$\bar{l}$ = $\bar{h}$0*l*).

5. Crystallization

The enzyme was stored in 50%(*v*/*v*) glycerol, which was removed during concentration by Centricon-10 centrifugal concentrators (Amicon, Stonehouse, Gloucestershire, England). The enzyme was concentrated to 10 mg ml^{-1} in 10 m*M* Tris–HCl pH 7.6, 0.4 m*M* DTT and 2.5 m*M* NADPH. Crystallization experiments were performed at 293 K using the sitting-drop vapour-diffusion technique. An exhaustive set of conditions, comprising commercially available and local sparse-matrix screens (Jancarik & Kim, 1991; Cudney *et al.*, 1994) and linear screens were tried without success. Attempts to optimize conditions that had yielded promising crystalline precipitate were also unsuccessful.

It was observed that the protein did not precipitate significantly in ammonium sulfate between pH 6.0 and 7.0. The protein concentration was therefore increased to 20 mg ml^{-1} and a linear screen was constructed varying the pH between 5.8 and 7.0 using 100 m*M* cacodylate, PIPES, MOPS and BTP as buffers and a range of ammonium sulfate concentrations between 1.0 and 2.2 *M*. Two conditions, 1.7 and 1.9 *M* $(NH_4)_2SO_4$ in cacodylate, yielded crystals. These initial conditions were optimized to 1.65*M* ammonium sulfate, 100 m*M* cacodylate buffer pH 5.8. Typically, 1.5 µl of protein (20 mg ml^{-1}) was mixed with an equal volume of reservoir solution to form the drop. Crystals appeared after 10–12 d and continued to grow as bisphenoids to maximum dimensions of 0.3 × 0.3 × 0.15 mm (Fig. 1).

6. X-ray analysis

X-ray diffraction data were collected in-house on a Nonius FR591 rotating-anode generator using a MacScience DIP2000 imaging-plate system. The crystals proved to be radiation-sensitive and therefore cryo-cooling was essential. Crystals were loop-mounted in a cryoprotectant containing 20%(*v*/*v*) glycerol and were cryo-cooled to 100 K using an Oxford Cryosystems Cryostream. A native data set extending to 2.3 Å resolution was collected using 1° oscillation frames. Data were processed with *DENZO* and scaled with *SCALEPACK* (Otwinowski, 1993). From autoindexing in *DENZO*, the crystals were found to index in a face-centred orthorhombic lattice corresponding to space group *F*222, with unit-cell parameters *a* = 110.0, *b* = 139.8, *c* = 173.5 Å. An assumption of two molecules per asymmetric unit lead to an acceptable packing density, V_m, of 3.03 Å^3 Da^{-1}, corresponding to a solvent content of 60% (Matthews, 1968). However, merging the data in this space group gave poor statistics which were especially poor at high resolution (Table 2*a*) and involved a large number of rejections. The crystals were therefore scaled and merged in the monoclinic space group *C*2, which is a subgroup of *F*222, with unit-cell parameters *a* = 110.0, *b* = 139.8, *c* = 102.5 Å, β = 122.2° (Table 2*b*). There were three possible *C*2 cells (Fig. 2) as a consequence of the *F*222 pseudo-lattice, so the monoclinic cell chosen was that which provided the best merging statistics. In *C*2, an acceptable packing density is obtained for three to six molecules per asymmetric unit. A pseudo-precession plot of the X-ray data clearly shows an NCS twofold perpendicular to the crystallographic twofold (Fig. 3). Analysis of the self-rotation function showed two perpendicular twofold non-crystallographic symmetry (NCS) axes with correlation coefficients of 82%; this pseudo-222 symmetry indicated that the asymmetric unit contained four molecules, giving a V_m of 2.86 Å^3 Da^{-1} and a solvent content of 57%. There are many examples of dehydrogenases whose quaternary structures are tetrameric and show 222 symmetry. This arrangement may be reflected in the NCS present in shikimate dehydrogenase, although gel filtration (Chaudhuri & Coggins, 1985) and dynamic light-scattering studies (reported here) have shown the protein to be monomeric under physiological conditions. We have attempted to solve the structure by molecular replacement with *AMoRe* (Navaza, 1990) using various dehydrogenases including lactate dehydrogenase (PDB code 1ldn), NADP-dependent alcohol dehydrogenase (1kev), and porcine mitochondrial acyl-COA dehydrogenase (0acd, now obsolete) as search models. These attempts have been without success and the structure will therefore be

solved by multiple isomorphous replacement.

This research was supported by BBSRC and the University of Glasgow. We would like to especially thank Mr John Greene and Dr Tino Krell for technical assistance in protein purification and mass spectrometry, and Dr Aleksander Roszak for useful discussions.

References

Anton, I. A. & Coggins, J. R. (1988). *Biochem. J.* **249**, 319–326.
Chaudhuri, S. & Coggins, J. R. (1985). *Biochem. J.* **226**, 217–223.
Coggins, J. R. (1989). In *Herbicides and Plant Metabolism*, edited by A. Dodge. Cambridge University Press.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. (1994). *Acta Cryst.* D**50**, 414–423.
Davies, G. M., Barrett-Bee, K. J., Jude, D. A., Lehan, M., Nichols, W. W., Pinder, P. E., Thain, J. L., Watkins, W. J. & Wilson, R. G. (1994). *Antimicrob. Agents Chemother.* **38**, 403–406.
Deka, R. K., Anton, I. A., Dunbar, B. & Coggins, J. R. (1994). *FEBS Lett.* **349**, 397–402.
Ferré-D'Amaré, A. R. & Burley, S. K. (1997). *Methods Enzymol.* **276**, 157–166.
Ferrige, A. G., Seddon, M. J., Green, B. N., Jarvis, S. A. & Skilling, J. (1992). *Rapid Commun. Mass Spectrom.* **6**, 707–711.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Krell, T., Pitt, A. R. & Coggins, J. R. (1995). *FEBS Lett.* **360**, 93–96.
Lumsden, J. & Coggins, J. R. (1977). *Biochem. J.* **161**, 599–607.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (1990). *Acta Cryst.* A**46**, 619–620.
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.
Studier, F. W. & Moffatt, B. A. (1986). *J. Mol. Biol.* **189**, 113–130.