Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

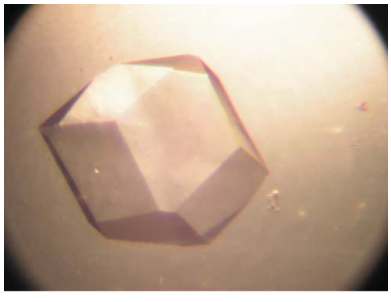
Expression, purification and preliminary crystallographic analysis of 2,4-dihydroxy-hepta-2-ene-1,7-dioate aldolase (HpcH) from *Escherichia coli* C

Dean Rea,[a] Vilmos Fülöp,[a] Timothy D. H. Bugg[b] and David I. Roper[a]*

[a]Department of Biological Sciences, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, England, and [b]Department of Chemistry, University of Warwick, Gibbet Hill Road, Coventry CV4 7AL, England

Correspondence e-mail: david.roper@warwick.ac.uk

Received 23 June 2005
Accepted 19 July 2005
Online 31 August 2005



The gene encoding 2,4-dihydroxy-hepta-2-ene-1,7-dioate (HHED) aldolase (HpcH; EC 4.1.2) from *Escherichia coli* C was cloned into the high-expression plasmid pProEx-HTa and overexpressed in *E. coli* BL21 (DE3). The 28 kDa enzyme was purified using immobilized metal-affinity and size-exclusion chromatography prior to crystallization. Crystals were obtained by the hanging-drop vapour-diffusion method at 277 K from a number of screening conditions. Type I crystals grown in a solution containing 0.4 *M* ammonium dihydrogen phosphate belong to space group *R*32, with unit-cell parameters $a = b = 128.92$, $c = 175.30$ Å. Type II crystals grown in a solution containing 0.5 *M* sodium chloride, 0.1 *M* sodium citrate pH 5.5 belong to space group *I*222, with unit-cell parameters $a = 133.39$, $b = 155.39$, $c = 168.80$ Å. Complete data sets were collected to 1.6 and 2.0 Å from type I and type II crystals, respectively, using synchrotron radiation.

1. Introduction

The benzene ring is one of the most prevalent chemical entities in the biosphere (Diaz, 2004). Recycling of the aromatic nucleus is mainly achieved by the specific catabolic activities of bacteria and fungi. In addition to the recycling of naturally occurring aromatic compounds, these organisms can evolve the ability to or be engineered to degrade xenobiotic compounds such as herbicides, pesticides and industrial effluents, many of which contain a benzene nucleus (Diaz *et al.*, 2001; Diaz, 2004; Bielefeldt & Cort, 2005). A better understanding of bacterial aromatic degradative pathways may facilitate the design of biodegradable chemicals or lead to the development of engineered organisms useful for bioremediation. Also, aromatic degradative pathways utilize a diverse array of enzymatic activities, which may be appropriate for use in synthetic chemistry for the stereoselective synthesis of organic compounds.

In bacteria, degradation of the aromatic nucleus is initiated by dihydroxylation of the aromatic ring, affording a more reactive species that is amenable to ring-opening (Diaz, 2004). This essentially irreversible step occurs either between the two hydroxyl groups (*ortho*-cleavage) or adjacent to one of the hydroxyl groups (*meta*-cleavage). Products of the ring cleavage are subsequently incorporated into the organism's central metabolic pathways.

One of the longest and most well characterized catabolic pathways involves the degradation of 4-hydroxyphenylacetic acid (4-HPA) to pyruvate and succinic semialdehyde (Cooper & Skinner, 1980). The catalytic pathway includes 11 genes encoding proteins for eight enzyme-catalysed steps, as well as a transporter protein involved in the uptake of 4-HPA and a regulatory protein. The genes required for 4-HPA degradation have been cloned from *Escherichia coli* C (Jenkins & Cooper, 1988; Roper *et al.*, 1993) and *E. coli* W (Prieto *et al.*, 1996). In the first step, 4-HPA is hydroxylated to give 3,4-dihydroxyphenylacetic acid (homoprotocatechuate; HPC). Hence, for historical reasons the pathway is known as the HPC pathway in *E. coli* C and the 4-HPA pathway in *E. coli* W.

Many of the enzymes have been biochemically characterized (Roper & Cooper, 1993; Roper *et al.*, 1995) and crystal structures have been determined for two of the enzymes (Subramanya *et al.*, 1996; Tame *et al.*, 2002). The final step of the pathway involves the conversion of 2,4-dihydroxy-hepta-2-ene-1,7-dioate (HHED) into

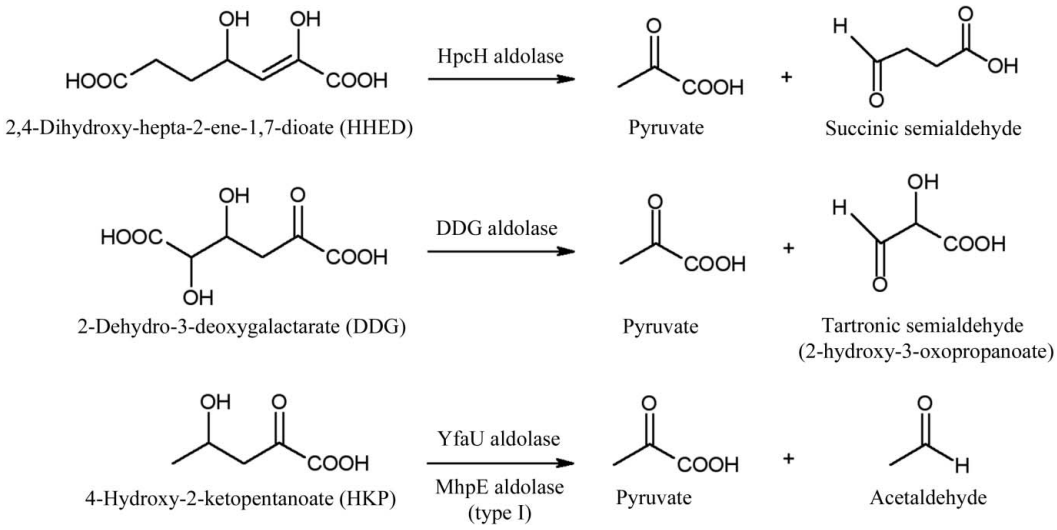


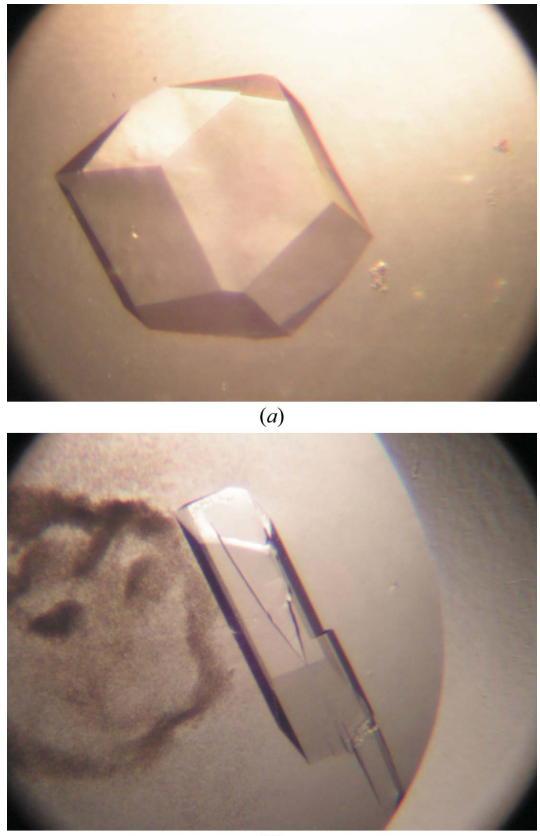
Figure 1
Comparison of the substrates and products of the HpcH, DDG and YfaU aldolases. The substrate for YfaU is unknown, but we have shown that this protein is able to catalyse the cleavage of 4-hydroxy-2-keto-pentanoate (HKP) to pyruvate and acetate. HKP is the substrate of MhpE aldolase from the phenylproprioate pathway of *E. coli* (Diaz *et al.*, 1988).

pyruvate and succinic semialdehyde. This reaction is catalyzed by 2,4-dihydroxy-hepta-2-ene-1,7-dioate aldolase (HpcH). HpcH (262 amino acids, 28 kDa) belongs to a subclass of type II divalent metal ion-dependent aldolases for which there is relatively little structural and mechanistic information. Indeed, the only crystal structure of an enzyme of this type is that of 2-dehydro-3-deoxygalactarate (DDG) aldolase (Izard & Blackwell, 2000). This enzyme is proposed to catalyse the breakage of the carbon–carbon bond in its substrate *via* a novel mechanism involving the participation of a phosphate bound in the active site that acts as a base, abstracting a proton from the substrate. The crystallization of YfaU from *E. coli* K12, a homologue of HpcH and DDG, has also recently been reported (Wright *et al.*, 2002). Interestingly, *E. coli* K12 does not contain HPC-pathway genes and cannot survive on 4-HPA or HPC as sole carbon sources. Analysis of sequence data from *E. coli* C (Roper *et al.*, 1995) and *E. coli* B (Prieto *et al.*, 1996) indicates that the HPC gene cluster has been inserted in a region between the *tsr* gene for methyl-accepting chemotaxis protein and a putative carbon-starvation protein, YjiY, relative to the *E. coli* K12 genome. Comparison of the crystal structures of DDG, HpcH, YfaU and related enzymes should enhance our understanding of the substrate specificity and catalytic mechanism of this class of magnesium-dependent class II aldolases (Fig. 1), particularly concerning the participation of phosphate in the reaction mechanism.

2. Materials and methods

2.1. Cloning, overexpression and purification

The gene encoding *E. coli* C HpcH was amplifed by PCR from *E. coli* C chromosomal DNA using primers incorporating *Nco*I and *Xho*I sites at the 5′ and 3′ ends of the gene, respectively. The PCR fragment was digested with these enzymes and ligated into the pProEx-HTa expression vector, which was also cut with these enzymes. Positive recombinants were selected by restriction digest and the resulting pProEx-HTa-HpcH plasmid containing the N-terminally hexahistidine-tagged HpcH gene was used to transform *E. coli* BL21 (DE3). Bacteria were grown at 310 K in 2 l LB medium containing 100 mg ml^{-1} ampicillin. Protein expression was induced by the addition of IPTG to 1 m*M* and growth continued at 293 K for 18 h. The cells were harvested by centrifugation at 6000*g* for 15 min and resuspended in 50 m*M* HEPES pH 8.0, 0.5 *M* NaCl, 25 m*M* imidazole (buffer *A*) prior to sonication and clarification of the extract by centrifugation at 20 000*g* for 30 min. The cell extract was applied onto a nickel Sepharose column equilibrated in buffer *A*. The column was washed with buffer *A* and buffer *B* (50 m*M* HEPES pH



(*a*)

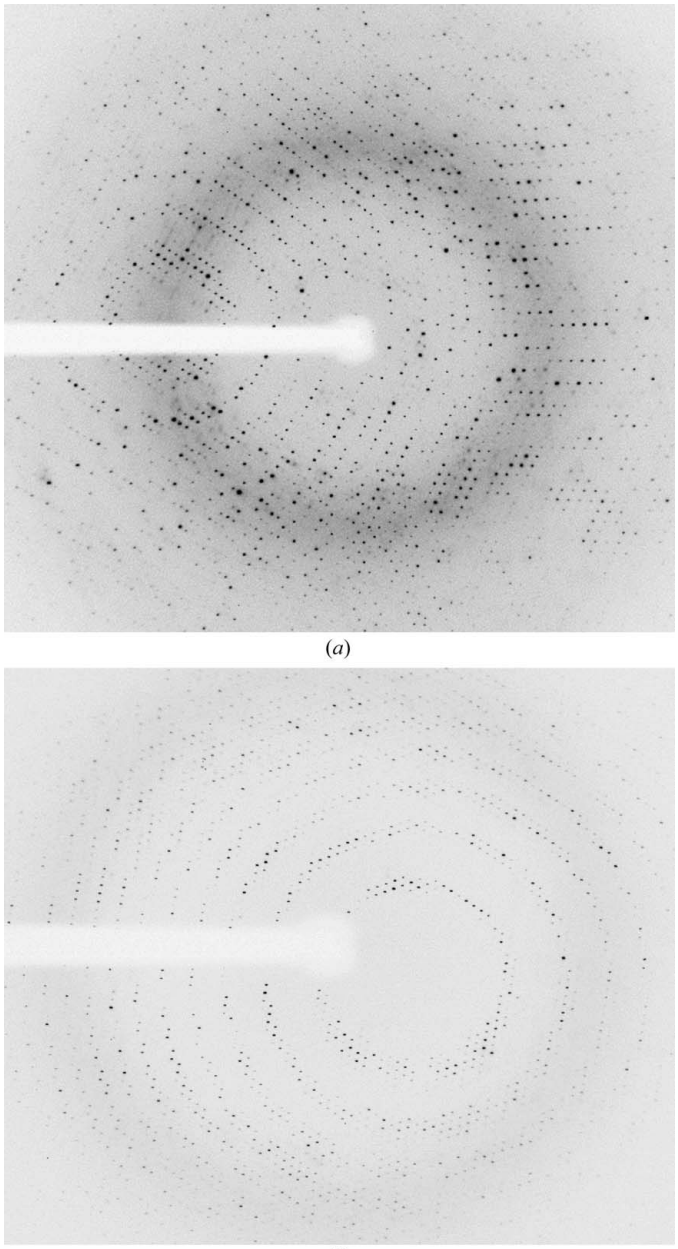
(*b*)

Figure 2
Photograph of (*a*) type I HpcH crystals grown in a solution containing 0.4 *M* ammonium dihydrogen phosphate and (*b*) type II HpcH crystals grown in a solution containing 0.5 *M* sodium chloride, 0.1 *M* sodium citrate pH 5.5.

8.0, 0.5 *M* NaCl, 100 m*M* imidazole) and the enzyme was eluted with buffer *C* (50 m*M* HEPES pH 8.0, 0.5 *M* NaCl, 500 m*M* imidazole). Fractions containing HpcH were combined, concentrated in a centrifugal concentrator and applied onto a Superdex 75 gel-filtration column equilibrated in 50 m*M* Tris–HCl pH 8.0, 200 m*M* NaCl. HpcH eluting from this column appeared to be >99% pure by SDS–PAGE and was dialysed into 10 m*M* Tris–HCl pH 8.0 and concentrated to 12 mg ml^{-1} prior to crystallization.

2.2. Crystallization

Initial crystallization trials were carried out using the sitting-drop vapour-diffusion method with 96-well plates (Greiner Bio-One Ltd, UK) containing 0.05 ml reservoir solutions taken from the Hampton Research Crystal Screen HT kit (Jancarik & Kim, 1991). Drops consisting of 2 µl protein solution and 1 µl reservoir solution were used throughout. Crystals appeared after 16 h using a number of conditions and conditions A3 (0.4 *M* ammonium dihydrogen phosphate) and F4 (0.5 *M* NaCl, 2% ethyleneimine polymer, 0.1 *M* trisodium citrate pH 5.6) were optimized using 24-well Linbro plates containing 0.2 ml reservoir solution. Type I (0.4 *M* ammonium dihydrogen phosphate, 30% glycerol) and type II (0.5 *M* sodium chloride, 0.1 *M* sodium citrate pH 5.5, 20% glycerol) crystals appeared to be suitable for diffraction studies, with dimensions in excess of 0.1 × 0.1 × 0.2 mm (Fig. 2).



(*a*)

(*b*)

Figure 3
A typical diffraction image of (*a*) type I crystals and (*b*) type II crystals collected at beamline BM14 at the ESRF using a MAR CCD detector. The oscillation range was 1° and the resolution at the edge is 1.6 and 2.0 Å, respectively.

Table 1
Data-collection and processing statistics.

Values in parentheses refer to the highest resolution shell.

	Type 1	Type 2
Wavelength (Å)	0.979	0.979
Space group	*R*32	*I*222
Unit-cell parameters		
a (Å)	128.92	133.35
b (Å)	128.92	155.39
c (Å)	175.30	168.84
Molecules per AU	2	6
Matthews coefficient (Å^3 Da^{-1})	2.25	2.33
Solvent content (%)	45	47
Resolution range (Å)	41–1.6 (1.66–1.6)	50–2.0 (2.07–2.00)
Total observations	803467	719529
Unique reflections	73674	117205
Average $I/\sigma(I)$	33.6 (2.5)	18.2 (1.9)
$R_{\rm merge}$† (%)	0.071 (0.713)	0.082 (0.655)
Completeness (%)	100.0 (100.0)	99.3 (99.1)

† $R_{\rm merge} = \sum_j \sum_h |I_{h,j} - \langle I_h \rangle| / \sum_j \sum_h \langle I_h \rangle$, where $I_{h,j}$ is the jth observation of reflection h and $\langle I_h \rangle$ is the mean intensity of that reflection.

3. X-ray diffraction analysis

Crystals were picked up from the crystallization drop using a nylon loop, transferred directly into a cryostream, cooled to 100 K and stored in liquid nitrogen until needed for data collection. Complete data sets were collected from type I and type II crystals to resolutions of 1.6 and 2.0 Å, respectively, on BM14 at the European Synchrotron Radiation Facility using a MAR CCD detector with an oscillation angle of 1°. Typical diffraction images are shown in Fig. 3. All data were indexed, integrated and scaled using the *HKL* suite of programs (Otwinowski & Minor, 1997). Type I crystals belong to space group *R*32, with unit-cell parameters $a = b = 128.92$, $c = 175.30$ Å. The Matthews probability calculation suggests the presence of two molecules in the asymmetric unit, with a $V_{\rm M}$ value of 2.25 Å^3 Da^{-1} and a solvent content of 45%. Type II crystals grown in a solution containing 0.5 *M* sodium chloride, 0.1 *M* sodium citrate pH 5.5 belong to space group *I*222, with unit-cell parameters $a = 133.35$, $b = 155.39$, $c = 168.4$ Å. The Matthews probability calculation suggests the presence of six molecules in the asymmetric unit, with a $V_{\rm M}$ value of 2.33 Å^3 Da^{-1} and a solvent content of 47%. Data-collection and processing statistics are given in Table 1. We are currently in the process of attempting a molecular-replacement solution using the coordinates for 2-dehydro-3-deoxygalactarate aldolase (Izard & Blackwell, 2000), which shares 46% sequence identity with HpcH.

We are grateful for access to and user support at the synchrotron-radiation facility of the ESRF, Grenoble. This work is funded by the BBSRC (grant reference BBC0044501) and supported by equipment

grant funding from the Wellcome Trust (grant references 068598, 075861 and 071998).

References

Bielefeldt, A. R. & Cort, T. (2005). *Biotechnol. Bioeng.* **89**, 680–689.
Cooper, R. A. & Skinner, M. (1980). *J. Bacteriol.* **143**, 302–306.
Diaz, E. (2004). *Int. Microbiol.* **7**, 173–180.
Diaz, E., Ferrandez, A. & Garcia, J. L. (1988). *J. Bacteriol.* **180**, 2915–2923.
Diaz, E., Ferrandez, A., Prieto, M. A. & Garcia, J. L. (2001). *Microbiol. Mol. Biol. Rev.* **65**, 523–569.
Izard, T. & Blackwell, N. C. (2000). *EMBO J.* **19**, 3849–3856.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Jenkins, J. R. & Cooper, R. A. (1988). *J. Bacteriol.* **170**, 5317–5324.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Prieto, M. A., Diaz, E. & Garcia, J. L. (1996). *J. Bacteriol.* **178**, 111–120.
Roper, D. I. & Cooper, R. A. (1993). *Eur. J. Biochem.* **217**, 575–580.
Roper, D. I., Fawcett, T. & Cooper, R. A. (1993). *Mol. Gen. Genet.* **237**, 241–250.
Roper, D. I., Stringfellow, J. M. & Cooper, R. A. (1995). *Gene*, **156**, 47–51.
Subramanya, H. S., Roper, D. I., Dauter, Z., Dodson, E. J., Davies, G. J., Wilson, K. S. & Wigley, D. B. (1996). *Biochemistry*, **35**, 792–802.
Tame, J. R. H., Namba, K., Dodson, E. J. & Roper, D. I. (2002). *Biochemistry*, **41**, 2982–2989.
Wright, A., Blewett, A., Fülöp, V., Cooper, R., Burrows, S., Jones, C. & Roper, D. (2002). *Acta Cryst.* D**58**, 2191–2193.