

**Tomoko Adachi,^a Atsushi Izumi,^b
 Dean Rea,^a Sam-Yong Park,^b
 Jeremy R. H. Tame^b and
 David I. Roper^{a*}**

^aDepartment of Biological Sciences, University of Warwick, Coventry CV4 7AL, England, and

^bDivision of Protein Design, Yokohama City University, Suehiro 1-7-29, Yokohama, Kanagawa 230-0045, Japan

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

Received 19 June 2006
 Accepted 5 September 2006

Expression, purification and crystallization of 2-oxo-hept-4-ene-1,7-dioate hydratase (HpcG) from *Escherichia coli* C

The gene encoding 2-oxo-hept-3-ene-1,7-dioic acid (OHED) hydratase (HpcG) was cloned into the high-expression plasmid pET26b and overexpressed in *Escherichia coli* BL21(DE3). The enzyme was purified in three steps to greater than 95% purity prior to crystallization. Crystals were obtained by the hanging-drop vapour-diffusion method at 277 K in a number of screening conditions. Crystals measuring up to 1.5 mm in their longest dimension were grown from solutions containing polyethylene glycol 20 000. The crystals belonged to space group $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = 136$, $b = 136$, $c = 192$ Å. A complete data set was collected to 2.1 Å from a single cryocooled crystal at 100 K using synchrotron radiation.

1. Introduction

Degradation of 4-hydroxyphenylacetic acid (4-HPA) involves the *meta*-cleavage of the aromatic nucleus and proceeds *via* the longest known catabolic pathway (Cooper & Skinner, 1980). The 11 genes involved are clustered together in *Escherichia coli*, encoding eight enzymes as well as a transporter protein involved in the uptake of 4-HPA. These genes include a three-component hydroxylase required for destabilization of the aromatic nucleus (Sucharitakul *et al.*, 2005; Thotsaporn *et al.*, 2004), two regulatory genes, *hpaR* and *hpaA*, and a gene of uncharacterized function, *hpaX*, which encodes a protein related to the superfamily of transmembrane facilitators (Prieto *et al.*, 1996) and is co-transcribed with *hpaA*.

The genes required for 4-HPA degradation were first cloned from *E. coli* C (Jenkins & Cooper, 1988; Roper *et al.*, 1993) and *E. coli* W (Prieto *et al.*, 1996). Since the publication of these early studies, the availability of genomic DNA databases has revealed the presence of 4-HPA pathways in a wide range of bacteria, as well as homologues of certain gene products in the pathways in both plants and animals. Many of the enzymes have been characterized biochemically (Roper & Cooper, 1990*a,b*, 1993; Roper *et al.*, 1993, 1995) and crystal structures have been determined for several of the enzymes, including HpcD (Subramanya *et al.*, 1996), HpcE (Tame *et al.*, 2002) and HpcB (Vetting *et al.*, 2004). The final enzyme of the pathway, HpcH, has also been crystallized (Rea *et al.*, 2005). The enzymatic steps of the 4-HPA pathway include hydroxylation, dioxygenation, oxidation, decarboxylation, isomerization, hydration and a retro-aldol reaction. This pathway therefore provides an opportunity to study how these intermolecular rearrangements are coordinated on similar substrate molecules. A better understanding of bacterial aromatic degradative pathways could facilitate the design of future biodegradable chemicals or lead to the development of engineered organisms useful for bioremediation. In addition, aromatic degradative pathways utilize enzymatic activities which may be appropriate for use in synthetic chemistry for the stereoselective synthesis of organic compounds. The structure and function of the enzymes that precede and follow HpcG in the 4-HPA pathway (HpcE and HpcH, respectively) is the subject of ongoing research in our laboratories. HpcG links the chemistry of these two enzymes and its structure is being studied in order to understand the details of its mechanism and substrate specificity. The reaction catalyzed by HpcG is similar to that catalyzed by 2-hydroxypentadienoic acid hydratase (MhpD) from the *meta*-



cleavage phenylpropionate degradation pathway (Pollard & Bugg, 1998). The crystal structure of MhpD has recently been deposited in the PDB (code 1sv6). Interestingly, this protein belongs to the fumarylacetoacetate hydrolase family of enzymes, which includes HpcE, the enzyme preceding HpcG in the 4-HPA pathway. HpcE is a bifunctional isomerase and decarboxylase containing two very similar domains with subtly different active-site topology (Tame *et al.*, 2002). Whilst there is very little sequence similarity between the domains of HpcE and HpcG, it seems likely that they may have evolved from a common ancestral precursor, particularly given their similar substrate structures. The kinetics and stereochemistry of HpcG have been investigated using an elegant series of experiments by Burks *et al.* (1998). The enzyme processes both 2-oxo-hept-4-ene-1,7-dioate and 2-hydroxy-2,4-heptadiene-1,7-dioate. Further kinetic analysis demonstrated that the mechanism proceeds *via* isomerization of 2-oxo-hept-4-ene-1,7-dioate to its $\alpha\beta$ -unsaturated ketone form followed by Michael addition of water, although the participation of one or two bases in the mechanism could not be defined (Burks *et al.*, 1998). HpcG utilizes a coordinated divalent metal ion, reported to be magnesium (Roper *et al.*, 1995; Burks *et al.*, 1998). This is consistent with other hydratase enzymes from *meta*-cleavage pathways such as BphH from *Burkholderia xenoverans* LB400 (Wang & Seah, 2005) and dmpE from *Pseudomonas putida* (Shingler *et al.*, 1992). A comparison of hydratase activities is shown in Fig. 1. Thus, solving the crystal structure of HpcG will provide further insight into the evolution of these proteins. This has implications for the development of biochemical pathways in general and may provide further understanding of the chemistry involved in this type of enzyme-catalyzed reaction.

2. Materials and methods

2.1. Cloning, overexpression and purification

The gene encoding *E. coli* C HpcG was amplified by PCR from *E. coli* C chromosomal DNA using primers incorporating *Nde*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 pET26b expression vector (Novagen) prepared with complementary ends. Positive recombinants were selected by restriction digest and the resulting pET-HpcG plasmid containing the C-terminally hexahistidine-tagged HpcG gene was used to transform *E. coli* BL21 (DE3), which was also cotransformed with the chaperone coexpression vector pTF16 (Takara Inc). Bacteria were grown at 310 K in 2 l LB medium supplemented with reagents required for automatic

induction of the culture (Studier, 2005) as well as 50 mg ml⁻¹ kanamycin, 35 mg ml⁻¹ chloroamphenicol and 0.02% L-arabinose. The cells were cultured overnight at 293 K prior to harvesting by centrifugation at 6000g for 20 min. The cells were resuspended in 50 mM HEPES pH 8.0 prior to sonication and clarification of the extract by centrifugation at 50 000g for 45 min. Ammonium sulfate was added to 40% saturation and the extract was clarified by centrifugation at 50 000g for 30 min. The ammonium sulfate pellet was resuspended in 50 mM HEPES, 1 mM DTT pH 8.0 and dialysed overnight into the same buffer. The clarified extract was applied to 3 × 5 ml HisTrap HP (GE Biosciences) columns connected in series equilibrated in 50 mM HEPES pH 8.0, 0.3 M NaCl, 2.5 mM imidazole (buffer A). The column was washed extensively with buffer A and the enzyme was eluted with a linear increasing gradient of imidazole from 2.5 to 250 mM. Under these conditions, it was found that HpcG bound with weak affinity to the column. Fractions containing HpcG were combined, dialysed against buffer A and reappplied onto the HisTrap HP columns, which were developed in the same way as described above. This reapplication procedure yielded higher purity protein samples for application to a size-exclusion chromatography column. Protein samples were concentrated in a centrifugal concentrator and applied onto a Superdex 75 gel-filtration column equilibrated in 20 mM HEPES pH 8.0, 500 mM NaCl and 3 mM DTT. Enzyme eluting from this column appeared to be greater than 99% pure by SDS-PAGE and was dialysed into 10 mM HEPES pH 8.0 and concentrated to 25 mg ml⁻¹ prior to crystallization. Selenomethionine-substituted protein was prepared using a similar protocol after initial growth in M9-based minimal media supplemented with 20 mg ml⁻¹ L-selenomethionine (Studier, 2005) and growth in the methionine-auxotrophic B834(DE3) strain.

2.2. Crystallization

Initial crystallization conditions for the native protein were established using 24-well Linbro plates with the hanging-drop vapour-diffusion technique. Purified HpcG was initially crystallized using the sitting-drop vapour-diffusion method under paraffin oil at 291 K using 1 μ l protein solution and 1 μ l mother liquor. Needle-shaped crystals were grown in 100 mM Tris-HCl pH 8.0, 0.2 M KSCN, 8% polyethylene glycol 20 000, 8% polyethylene glycol 550 monomethyl ether, derived from the Clear Strategy screen (Brzozowski & Walton, 2001). These crystals proved difficult to cryoprotect and handle, but further screening with Additive Screens 1–3 (Hampton Research) yielded superior crystals that were suitable for data

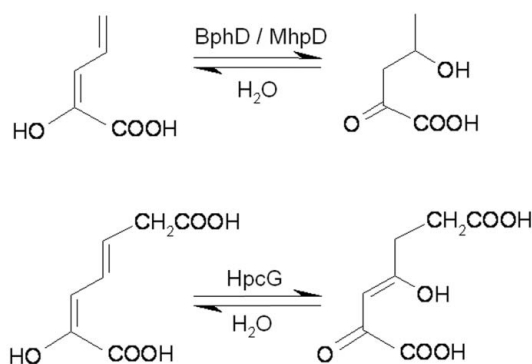


Figure 1
Comparison of the substrate selectivity of HpcG and analogous BphH and MhpD hydratases.

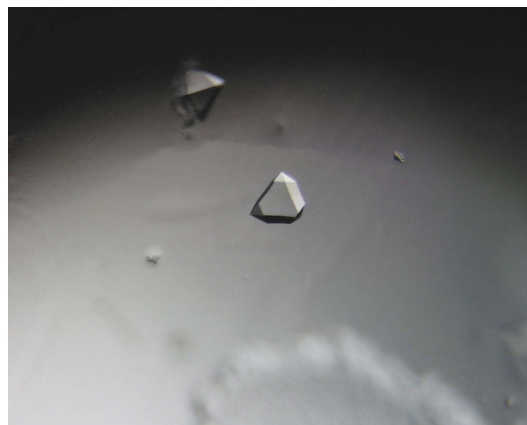


Figure 2
Crystals of HpcG grown as detailed in the text with dimensions of 0.1–0.2 mm.

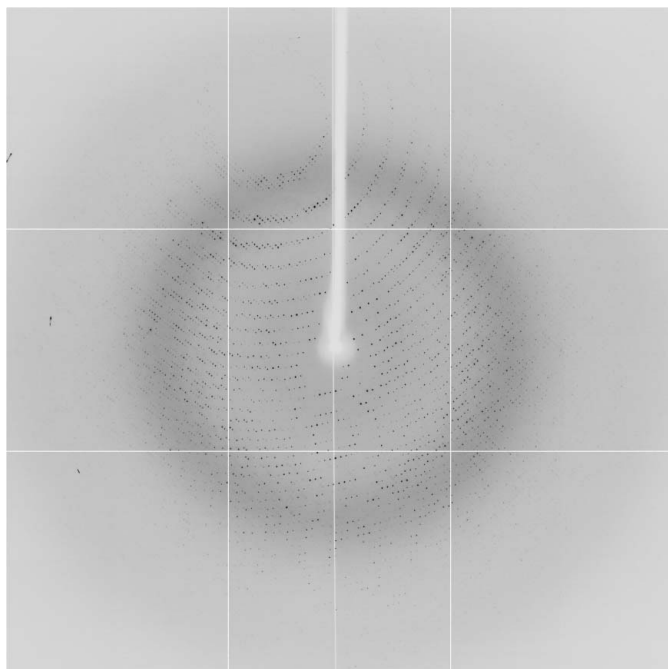


Figure 3
A typical diffraction image of an HpcG crystal collected on beamline BL5A at the Photon Factory using an ADSC 315 detector. Resolution at the edge is 2.0 Å.

collection. The final HpcG crystals (Fig. 2) used for data collection were grown in 4 d using 6 µl drops containing 2.5 µl protein, 1.4 µl H₂O, 1.5 µl mother liquor (6.4% polyethylene glycol 20 000, 5% polyethylene glycol 550 monomethyl ether, 0.7 M KSCN, 50 mM MES pH 5.5) and 0.6 µl 0.1 M DTT in sitting-drop experiments under paraffin oil.

3. X-ray diffraction analysis

Crystals were picked up from the crystallization drops, soaked in mother liquor containing 30% glycerol for 30 s, transferred to a cryostream and cooled to 100 K. The quality of the diffraction was checked using a home-source X-ray generator and an R-AXIS IV⁺⁺ detector (Rigaku). A complete data set was collected on beamline BL5A at the Photon Factory using an ADSC 315 detector with an oscillation angle of 1° (Fig. 3). Data were indexed using *MOSFLM* (Leslie, 1992) and scaled and merged using *SCALA* (Collaborative Computational Project, Number 4, 1994). Data-collection and processing statistics are shown in Table 1. HpcG crystals belong to the tetragonal space group *P*₄₁₂ or *P*₄₃₂. The Matthews coefficient suggests that there may be five or six molecules in the asymmetric unit. In the case of five molecules, V_M is 3.1 Å³ Da⁻¹ and the solvent content is around 60%. For six molecules, V_M is 2.5 Å³ Da⁻¹ and the solvent content is around 51%. We are attempting molecular replacement using the *Escherichia coli* 2-hydroxypentadienoic acid hydratase coordinates (PDB code 1sv6),

Table 1
Data-collection and processing statistics.

Values in parentheses refer to the highest resolution shell.

Wavelength (Å)	1.00
Space group	<i>P</i> ₄ ₁ ₂ or <i>P</i> ₄ ₃ ₂
Unit-cell parameters (Å)	<i>a</i> = 136.25, <i>b</i> = 136.25, <i>c</i> = 192.42
Molecules per ASU	6
Matthews coefficient (Å ³ Da ⁻¹)	2.53
Solvent content (%)	51
Resolution range (Å)	39.3–2.1
Total observations	1016569
Unique reflections	105928
Average <i>I</i> /σ(<i>I</i>)	5.6 (1.6)
<i>R</i> _{sym} † (%)	10.5 (47.0)
Completeness (%)	100 (100)

† $R_{\text{sym}} = \sum_{\mathbf{h}} \sum_l |I_{\mathbf{h}l} - \langle I_{\mathbf{h}} \rangle| / \sum_{\mathbf{h}} \sum_l \langle I_{\mathbf{h}} \rangle$, where $I_{\mathbf{h}l}$ is the *l*th observation of reflection **h** and $\langle I_{\mathbf{h}} \rangle$ is the weighted average intensity for all observations *l* of reflection **h**.

which has 32% sequence identity to HpcG. Selenomethionine-substituted protein has also been prepared for structure determination using MAD.

This work is funded in part by the BBSRC (grant ref. BBC0044501) and supported by equipment grant funding from the Wellcome Trust (grant refs. 068598, 075861 and 071998).

References

- Brzozowski, A. M. & Walton, J. (2001). *J. Appl. Cryst.* **34**, 97–101.
- Burks, E. A., Johnson, W. H. Jr & Whitman, C. P. (1998). *J. Am. Chem. Soc.* **120**, 7665–7675.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Cooper, R. A. & Skinner, M. A. (1980). *J. Bacteriol.* **143**, 302–306.
- Jenkins, J. R. & Cooper, R. A. (1988). *J. Bacteriol.* **170**, 5317–5324.
- Leslie, A. G. W. (1992). *Int CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- Pollard, J. R. & Bugg, T. D. (1998). *Eur. J. Biochem.* **251**, 98–106.
- Prieto, M. A., Diaz, E. & Garcia, J. L. (1996). *J. Bacteriol.* **178**, 111–120.
- Rea, D., Füllöp, V., Bugg, T. D. H. & Roper, D. I. (2005). *Acta Cryst.* **F61**, 821–824.
- Roper, D. I. & Cooper, R. A. (1990a). *FEBS Lett.* **266**, 63–66.
- Roper, D. I. & Cooper, R. A. (1990b). *FEBS Lett.* **275**, 53–57.
- 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.
- Shingler, V., Powlowski, J. & Marklund, U. (1992). *J. Bacteriol.* **174**, 711–724.
- Studier, F. W. (2005). *Protein Expr. Purif.* **41**, 207–234.
- 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.
- Sucharitakul, J., Chaiyen, P., Entsch, B. & Ballou, D. P. (2005). *Biochemistry*, **44**, 10434–10442.
- Tame, J. R., Namba, K., Dodson, E. J. & Roper, D. I. (2002). *Biochemistry*, **41**, 2982–2989.
- Thotsaporn, K., Sucharitakul, J., Wongratana, J., Suadee, C. & Chaiyen, P. (2004). *Biochim. Biophys. Acta*, **1680**, 60–66.
- Vetting, M. W., Wackett, L. P., Que, L. Jr, Lipscomb, J. D. & Ohlendorf, D. H. (2004). *J. Bacteriol.* **186**, 1945–1958.
- Wang, P. & Seah, S. Y. (2005). *FEBS J.* **272**, 966–974.