Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and X-ray analysis of a bacterial non-haem iron-containing phenylalanine hydroxylase from the Gram-negative opportunistic pathogen *Pseudomonas aeruginosa*

Fredrik Ekström,[a] Gunter Stier,[b] Julian T. Eaton[c] and Uwe H. Sauer[a]*

[a]Umeå Centre for Molecular Pathogenesis, UCMP, Umeå University, SE-901 87, Umeå, Sweden, [b]EMBL, Meyerhofstrasse 1, D-69017 Heidelberg, Germany, and [c]Department of Biochemistry and Molecular Biology, University College London, Gower Street, London WC1E 6BT, England

Correspondence e-mail: uwe@ucmp.umu.se

Received 10 February 2003
Accepted 8 May 2003

Monooxygenases are frequently involved in the pathways that mediate the pivotal role of microorganisms in recycling carbon from the environment. A structural study of a monooxygenase from *Pseudomonas aeruginosa* that was identified as a phenylalanine hydroxylase has been initiated. The single-domain monomeric protein harbours a non-haem iron at the active site. The sequence identity to the catalytic domains of tyrosine and tryptophan hydroxylases suggests that the enzyme is not restricted to the substrate phenylalanine alone. Here, the cloning, purification and crystallization of native and SeMet-labelled *P. aeruginosa* phenylalanine hydroxylase are reported. Crystals grew in space group $P6_1$, with unit-cell parameters $a = b = 210.5$, $c = 100.7$ Å, and diffracted to a d spacing of 2.0 Å. Crystals of SeMet-labelled protein were used to collect a three-wavelength multiple anomalous dispersion (MAD) data set around the Se K edge.

1. Introduction

The non-haem iron-containing enzyme phenylalanine hydroxylase (phenylalanine 4-monooxygenase; EC 1.14.16.1; PheOH, also abbreviated as PAH) belongs to a family of (6*R*)-L-*erythro*-5,6,7,8-tetrahydrobiopterin (BH_4) dependent aromatic amino-acid monooxygenases and is homologous to tyrosine hydroxylase (TyrOH, tyrosine 3-monooxygenase; EC 1.14.16.2) and tryptophan hydroxylase (TrpOH, tryptophan 5-monooxygenase; EC 1.14.16.4). Mammalian PheOH initiates the rate-limiting step in the catabolism of L-phenylalanine, a process requiring the cofactor BH_4 and dioxygen (O_2) to hydroxylate L-phenylalanine to L-tyrosine (Kaufman, 1993). Mutations in the human phenylalanine hydroxylase gene result in deficient enzyme activity and cause hyperphenylalaninaemia, *e.g.* the autosomal recessive disease phenylketonuria (PKU; Hoang *et al.*, 1996) that affects about 1 in 10 000 newborns (DiLella *et al.*, 1986). The tetrameric mammalian PheOH is composed of an N-terminal regulatory domain (residues 1–142), a central catalytic domain (residues 143–410) and a C-terminal tetramerization domain (residues 411–452). Recently, crystal structures of truncated variants of human PheOH have become available (Erlandsen *et al.*, 1997; Fusetti *et al.*, 1998; Kobe *et al.*, 1999).

The best-characterized prokaryotic phenylalanine hydroxylases are the protein from *Chromobacterium violaceum* (CvPheOH; Nakata *et al.*, 1979; Chen & Frey, 1998), the crystal structure of which was recently determined (Erlandsen *et al.*, 2002), and the protein from *Pseudomonas aeruginosa* (PaPheOH, also termed PhhA; Zhao *et al.*, 1994). Although the *P. aeruginosa* enzyme is homologous to the catalytic domain of human PheOH (34% sequence identity over 215 aligned residues), the substrate specificity of PaPheOH might not be restricted to phenylalanine alone. Indirect support comes from the observation that the residues of the secondary-structure elements surrounding the active-site cleft of PaPheOH display about 40% sequence identity over 96 aligned amino acids to human phenylalanine, tyrosine and tryptophan hydroxylases. Experimental evidence that substrate specificity can be extended to include tyrosine was provided by Song *et al.* (1999). We intend to use PaPheOH as a model enzyme to further elucidate the underlying mechanisms of the enzymatic hydroxylation reaction and have initiated the structure determination of this archetypal monooxygenase. Here, we describe the expression, purification and crystallization of PheOH from the opportunistic pathogen *P. aeruginosa*. The effects of different additives on crystal morphology and diffraction quality are discussed. Data collection and preliminary analysis of native data as well as MAD data collected around the Fe K and Se K edges are reported.

2. Materials and methods

2.1. Cloning expression and purification

The *phhA* gene coding for the full-length PaPheOH protein (residues 1–262) was amplified simultaneously with its downstream neighbour, the *phhB* gene, from a crude

© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

Table 1
Data-collection and processing statistics.

Values for the highest resolution shell are given in parentheses.

	Native	SeMet peak	SeMet inflection	SeMet remote	Fe peak
Space group	$P6_1$	$P6_1$	$P6_1$	$P6_1$	$P6_1$
Unit-cell parameters (Å)	$a = b = 210.51$, $c = 100.70$	$a = b = 210.26$, $c = 102.35$	$a = b = 210.26$, $c = 102.35$	$a = b = 210.26$, $c = 102.35$	$a = b = 210.11$, $c = 102.21$
Wavelength (Å)	0.9340	0.978619	0.97880	0.90492	1.7387
Resolution range (Å)	20–2.0 (2.05–2.0)	20–2.6 (2.69–2.60)	20–2.6 (2.69–2.60)	20–2.6 (2.69–2.60)	20–3.4 (3.65–3.39)
Total reflections	1755022	672736	601145	595450	224371
Unique reflections	170881 (113759)	77581 (7268)	70933 (3704)	70911 (3714)	67960 (13501)
$I/\sigma(I)$	24.7 (3.8)	31.2 (9.7)	28.1 (7.5)	31.4 (7.2)	20.2 (11.3)
Completeness (%)	99.9 (100.0)	98.2 (92.3)	89.7 (47.0)	89.6 (47.1)	99.0 (98.3)
$R_{\rm merge}$†	7.7 (62.6)	5.9 (22.0)	5.2 (17.2)	5.2 (23.7)	6.1 (10.4)

† $R_{\rm merge}(I) = [\sum_h \sum_{i=1}^{N} |I_i(h) - \langle I(h)\rangle| / \sum_h \sum_{i=1}^{N} I_i(h)] \times 100$.

P. aeruginosa lysate by PCR using the 5′ primer CCC*TC ATG AAA A*CG ACG CAG TAC GTG GCC and the 3′ primer TTT *GGT ACC* TCA TTT GCG CCC CTC GGC GGT TTT C. The PCR product was cut with *Bsp*HI and *Kpn*I and subsequently cloned into the compatible *Nco*I and *Kpn*I sites of a modified pET9d expression vector (Novagen) that contained coding sequences for an N-terminal six-His tag and a thrombin protease recognition site preceding the *Nco*I site. The plasmid was transformed into *Escherichia coli* strain BL21(DE3) (Novagen). Cells were cultured at 310 K in Luria–Bertani medium to an OD_{600} of 1.0. After induction with 0.2 m*M* isopropyl thiogalactopyranoside (IPTG) for 16 h at 293 K, the cells were harvested by centrifugation and immediately stored at 193 K. Prior to purification, the bacterial pellet was thawed and resuspended in lysis buffer (20 m*M* Tris pH 8.0, 200 m*M* NaCl, 10 m*M* imidazole). The cells were lysed using a combination of lysozyme and sonication in the presence of DNaseI and were subsequently centrifuged at 30 000*g* for 50 min. The supernatant was batch-incubated with Ni–NTA–agarose (Qiagen) for 15 min and applied to an Econo chromatography column (BioRad). Unspecific bound proteins were removed by washing with five column volumes of high-salt buffer (20 m*M* Tris pH 8.0, 1.0 *M* NaCl) followed by equilibration with buffer containing 200 m*M* NaCl. The protein was stepwise eluted in buffer containing 20 m*M* Tris pH 8.0, 200 m*M* NaCl and 500 m*M* imidazole. The six-His tag was cleaved off with thrombin protease (Pharmacia) during dialysis against 20 m*M* Tris pH 8.0, 50 m*M* NaCl. Uncleaved protein was removed with Ni–NTA–agarose. PaPheOH was further purified using a MonoQ 10/10 ion-exchange column and a Superdex75 16/60 size-exclusion column (Amersham Pharmacia Biotech). The protein was concentrated in Centriprep-10 concentrators (Amicon), the purity was analyzed by SDS–PAGE and the concentration was determined with Coomassie Protein Plus solution (Pierce). A standard protocol (Hendrickson *et al.*, 1990) was followed to label the protein with SeMet for MAD experiments using the methionine-auxotropic *E. coli* strain B834(DE3) (Novagen).

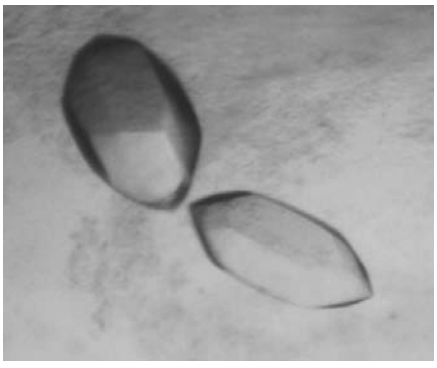

Figure 1
Crystals of PaPheOH grew at 291 K in the presence of 100 m*M* 1,6-hexanediol to dimensions of approximately 0.5 × 0.5 × 1.0 mm. They belong to space group $P6_1$ and diffract to a *d* spacing of 2.0 Å.

2.2. Crystallization and data collection

Initial crystallization conditions were identified using the hanging-drop vapour-diffusion method and a protein concentration of 10 mg ml^{-1}. The crystallization conditions were refined at 291 K for the native protein; selenomethionine-labelled protein crystals grew under conditions similar to those of the native protein. To ensure a high occupancy of the active-site iron, 1 m*M* $FeSO_4$ was added to the crystallization buffer.

The crystals were mounted free-standing in a nylon loop (Sauer & Ceska, 1997) and flash-cooled to 100 K in a cold nitrogen-gas stream (Oxford Cryosystems cryohead). The oscillation angle for all data sets was 0.5° per exposure and the total oscillation range covered at least 150° per data set in order to obtain high redundancy. Native data were collected at ID14-1 at ESRF, France on a MAR Research CCD detector (Table 1). The iron MAD data around the Fe *K* edge were collected at a wavelength of 1.7387 Å (7149.0 keV) (Table 1), the inflection-point data at 1.7400 Å (7143.7 keV) and the remote data at 0.9130 Å (13614.5 keV) using a MAR Research CCD detector at beamline BW7A at the EMBL Outstation at DESY in Hamburg, Germany. The SeMet MAD data were collected at beamline BM14 at ESRF, France on a fast-readout low-noise ('FReLoN') CCD camera coupled to an X-ray image intensifier (XRII) system developed at the ESRF (Hammersley *et al.*, 1997). Appropriate corrections for both spatial distortions and non-uniform response were carried out with the program *FIT2D* (Hammersley, 1998). All intensities were indexed, integrated and scaled with *XDS*/*XSCALE* (Kabsch, 1993) or *DENZO* (Otwinowski & Minor, 1997) (Table 1).

3. Results and discussion

Protein purification yielded about 10 mg of highly pure and homogeneous PaPheOH per litre of bacterial culture. The dicistronic expression unit in which the *phhB* gene is expressed as a second open reading frame downstream of the *phhA* gene was crucial for high-level expression of PaPheOH. As predicted from the sequence and verified by size-exclusion chromatography, the PaPheOH protein elutes as a monomer of MW = 30.3 kDa. The protein crystallized within 3 d at 291 K from mother liquor containing 1.4–1.7 *M* $LiSO_4$ and 0.1 *M* HEPES adjusted to pH 7.5. The elliptically elongated crystals, similar in shape to American footballs, grew to dimensions of 0.3 × 0.3 × 0.6 mm and diffracted to a *d*

spacing of 2.6 Å. Addition of 100 m*M* 1,6-hexanediol to the crystallization buffer led to hexagonal rod-shaped crystals with well defined edges and improved size (0.5 × 0.5 × 1.0 mm; Fig. 1) without affecting the resolution. Finally, the best-diffracting crystals grew from a buffer containing 0.1 *M* HEPES pH 7.5, 1.4–1.7 *M* $LiSO_4$, 100 m*M* 1,6-hexanediol and 1 m*M* $FeSO_4$. They were used to collect a 2.0 Å resolution native data set as well as the Fe MAD data sets (Table 1). The space group was determined to be $P6_1$, with unit-cell parameters $a = b = 210.5$, $c = 100.7$ Å. Considering the volume of the unit cell and the molecular weight of PaPheOH, we obtain packing parameter values V_M (Matthews, 1968) of between 3.6 and 1.8 $Å^3\ Da^{-1}$, suggesting a minimum of six or a maximum of 12 molecules in the asymmetric unit. Molecular-replacement trials with hPheOH and hTyrOH as search models using *CNS* (Brünger *et al.*, 1998) did not result in usable solutions. Exploiting the active-site iron, we collected a three-wavelength MAD data set to 3.4 Å resolution around the iron *K* edge ($\lambda = 1.743$ Å) (Table 1). However, the anomalous signal of the iron MAD data was not sufficient for phasing trials using *SnB* (Weeks & Miller, 1999) and *SOLVE* (Terwilliger & Berendzen, 1999).

Useful phase information could be extracted from a MAD data set collected from SeMet-labelled protein crystals. The labelled protein crystallized under native conditions and diffracted to a resolution of 2.6 Å (Table 1). The electron density could be interpreted and structure refinement is in progress.

The authors thank Dr Dietrich Suck, EMBL, Heidelberg, Germany and Professor Hans Eklund, Swedish University of Agricultural Sciences (SLU) for support during part of the project. We thank Drs Andy Thomson, Sean McSweeney and Gordon Leonard from the EMBL Outstation in Grenoble, France for expert help with SeMAD data collection. The authors acknowledge the help of Ehmke Pohl, Cristofer Enroth and Alexander Popov from the EMBL Outstation at DESY, Hamburg, Germany with Fe MAD data collection. We gratefully appreciate the critical comments of Elisabeth Sauer-Eriksson on the manuscript. This work was supported by a grant from the Swedish Natural Science Research Council.

References

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Chen, D. & Frey, P. A. (1998). *J. Biol. Chem.* **273**, 25594–25601.

DiLella, A. G., Kwok, S. C., Ledley, F. D., Marvit, J. & Woo, S. L. (1986). *Biochemistry*, **25**, 743–749.

Erlandsen, H., Fusetti, F., Martinez, A., Hough, E., Flatmark, T. & Stevens, R. C. (1997). *Nature Struct. Biol.* **4**, 995–1000.

Erlandsen, H., Kim, J. Y., Patch, M. G., Han, A., Volner, A., Abu-Omar, M. M. & Stevens, R. C. (2002). *J. Mol. Biol.* **320**, 645–661.

Fusetti, F., Erlandsen, H., Flatmark, T. & Stevens, R. C. (1998). *J. Biol. Chem.* **273**, 16962–16967.

Hammersley, A. P. (1998). *FIT2D v.*9.129 *Reference Manual v.*3.1.

Hammersley, A. P., Brown, K., Burmeister, W., Claustre, L., Gonzalez, A., McSweeney, S., Mitchell, E., Moy, J.-P., Svensson, S. O. & Thompson, A. (1997). *J. Synchrotron Rad.* **4**, 67–77.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. (1990). *EMBO J.* **9**, 1665–1672.

Hoang, L., Byck, S., Prevost, L. & Scriver, C. R. (1996). *Nucleic Acids Res.* **24**, 127–131.

Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.

Kaufman, S. (1993). *Adv. Enzymol. Relat. Areas Mol. Biol.* **67**, 77–264.

Kobe, B., Jennings, I. G., House, C. M., Michell, B. J., Goodwill, K. E., Santarsiero, B. D., Stevens, R. C., Cotton, R. G. & Kemp, B. E. (1999). *Nature Struct. Biol.* **6**, 442–448.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Nakata, H., Yamauchi, T. & Fujisawa, H. (1979). *J. Biol. Chem.* **254**, 1829–1833.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Sauer, U. H. & Ceska, T. H. (1997). *J. Appl. Cryst.* **30**, 71–72.

Song, J., Xia, T. & Jensen, R. A. (1999). *J. Bacteriol.* **181**, 2789–2796.

Terwilliger, T. C. & Berendzen, J. (1999). *Acta Cryst.* D**55**, 849–861.

Weeks, C. M. & Miller, R. (1999). *J. Appl. Cryst.* **32**, 120–124.

Zhao, G., Xia, T., Song, J. & Jensen, R. A. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 1366–1370.