Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction studies of a fungal hydrolase from *Ophiostoma novo-ulmi*

Michail N. Isupov,[a] Amanda A. Brindley,[a] Edward J. Hollingsworth,[a] Garib N. Murshudov,[b] Alexei A. Vagin[b] and Jennifer A. Littlechild[a]*

[a]Henry Wellcome Building for Biocatalysis, School of Biological and Chemical Sciences, University of Exeter, Exeter EX4 4QD, England, and [b]Department of Chemistry, University of York, Heslington, York YO1 5DD, England

Correspondence e-mail: j.a.littlechild@exeter.ac.uk

Received 15 July 2004
Accepted 22 July 2004

Dutch elm disease fungus *Ophiostoma novo-ulmi* contains a hydrolase activity which catalyses the resolution of racemic ethyl naproxen to the corresponding acid. The recombinant enzyme has been crystallized by the vapour-diffusion method in two crystal forms. The crystals of the first form belong to space group $P2_12_12$, with unit-cell parameters a = 115.9, b = 174.4, c = 62.1 Å. The enzyme also crystallizes in space group $P2_12_12$, with unit-cell parameters a = 72.9, b = 212.7, c = 61.7 Å. Synchrotron data have been collected for both crystal forms to 2.6 and 2.3 Å, respectively. A molecular-replacement solution has been found using a remote starting model of a bacterial esterase (23% sequence identity) for both crystal forms. Multicrystal averaging has resulted in interpretable electron-density maps.

1. Introduction

Hydrolases are a large family of enzymes that are responsible for the degradation of amide- and ester-type bonds in a wide range of substrate molecules. These enzymes can utilize a variety of natural and unnatural substrates and this has led to their use as stereoselective catalysts in the synthesis of optically pure molecules for the pharmaceutical and agrochemical industries (Faber, 1997). Esterases are part of the hydrolase family and are ubiquitous enzymes, having been identified in eukaryotes, bacteria and archaeal species.

A novel hydrolase was isolated from the fungus *Ophiostoma novo-ulmi*, the causative agent of Dutch elm disease. The 435-amino-acid enzyme has a molecular mass of 47 601, a predicted pI of 5.44 and is capable of catalysing the resolution of racemic ethyl naproxen. Naproxen, like ketoprofen and ibruprofen, is a 2-arylpropionoic acid that acts as a non-steroidal anti-inflammatory drug in the treatment of arthritis and other related connective-tissue diseases. *S*-Naproxen has been shown to be the active isomer. Optically pure naproxen can be produced by chemical methods that are generally complicated and costly (Fadel, 1992). The so-called 'naproxen esterase' described in this paper is capable of utilizing *S*-ethyl naproxen as a substrate with 98% enantiomeric excess.

Characterization of this esterase (Brindley, 1998) shows that this enzyme is not inhibited by phenylmethylsulfonyl fluoride (PMSF), an irreversible inhibitor of serine hydrolase enzymes. The enzyme belongs to the esterase family VIII (Arpigny & Jaeger, 1999) and has a sequence similarity to penicillin-binding proteins. These latter enzymes bind or react with penicillin and other β-lactam antibiotics and do not possess a classical catalytic triad of Ser-His-Asp-(Glu), but have a Ser-*X*-*X*-Lys motif where serine is the catalytic residue (Kelly *et al.*, 1985).

A structure has been elucidated recently of another representative of the esterase VIII family: EstB from *Burkholderia gladioli* (Wagner *et al.*, 2002). This enzyme shares 23% sequence identity with *O. novo-ulmi* esterase and has a tertiary structure similar to that of penicillin-binding proteins. Although EstB shares some active-site residues with penicillin-binding proteins, it shows no catalytic activity towards the β-lactam ring (Petersen *et al.*, 2001). This has been attributed to the difference in shape of the active-site tunnel (Wagner *et al.*, 2002).

Here, we report the purification, crystallization and preliminary X-ray analysis of the naproxen esterase.

2. Materials and methods

2.1. Protein purification

A λ-phage expression library of the *O. novo-ulmi* cDNA was made in *Escherichia coli* and initially screened for ability to use *S*-ethyl ketoprofen as a substrate. The resultant gene was then subjected to mutagenesis with hydroxylamine and screened for the ability to use *S*-ethyl naproxen as a substrate. This was monitored by the appearance of clearing zones of the emulsified substrate overlaid on an agar plate. The positive colonies were picked from the plate and the resultant naproxen esterase gene was overexpressed in *E. coli* using two different expression systems: pKK233-2 (Amersham Biosciences) and a variant of

Table 1
Processing statistics for the native data sets from the two different crystal forms.

Values in parentheses are for the highest resolution shell.

	Crystal *A*	Crystal *B*
Space group	$P2_12_12$	$P2_12_12$
Unit-cell parameters (Å)	a = 115.9, b = 174.4, c = 62.1	a = 72.9, b = 212.7, c = 61.7
Resolution range (Å)	25–2.6 (2.64–2.6)	20–2.32 (2.36–2.32)
No. measured reflections	97261	136392
No. unique reflections	36757	40786
Completeness	92.6 (85.6)	95.7 (94.7)
$I > 2\sigma(I)$ (%)	74.7	80.4
$\langle I\rangle/\sigma\langle I\rangle$	6.8	16.7
$R_{\rm sym}$† (%)	14.1 (48.9)	6.5 (25.1)

† $R_{\rm sym} = \sum|I - \langle I\rangle|/\sum I$, where I is the intensity of the reflection.

pUC19. This work was carried out at Chiroscience, Cambridge Science Park, UK in collaboration with Dr Richard Wisdom.

The *E. coli* cells harbouring the naproxen esterase gene were lysed with lysozyme [0.1%(*w*/*v*) in lysis buffer: 0.1 *M* sodium bicarbonate, 10 m*M* EDTA, 1% Triton X-100 pH 9] by shaking for 4 h at room temperature. Polyethylenimine (0.25% final concentration) was added to remove the nucleic acids and the lysed cells were shaken for a further 30 min. After centrifugation to remove the cell debris, the resultant supernatant was dialysed against 10 m*M* Bicine, 0.2 m*M* EDTA pH 8.5 (buffer *A*).

A 30% ammonium sulfate (enzyme-grade) precipitation step was used to remove unwanted protein and the supernatant was dialysed against buffer *A*. This was then adjusted to 20% $(NH_4)_2SO_4$ before application onto a benzamidine Sepharose affinity column (Sigma) equilibrated in the same buffer. The naproxen esterase was eluted with a negative linear gradient of 20–0% $(NH_4)_2SO_4$ in buffer *A* over one column volume.

Esterase activity was detected using the *p*-nitrophenyl acetate assay (Sobek & Gorisch, 1988) and the purity of the protein was assessed by SDS–PAGE (Laemmli, 1970). Active fractions were precipitated with 80% $(NH_4)_2SO_4$. The resulting pellet was resuspended in a minimum of 10 m*M* HEPES, 0.2 m*M* EDTA pH 7.2 (buffer *B*) and dialysed against the same buffer. The dialysed sample was loaded onto a pre-packed Pharmacia MonoQ anion-exchange column 5/5 equilibrated with buffer *B*. Protein was eluted with a linear gradient of 0–0.5 *M* NaCl over 20 column volumes. The esterase activity was eluted at 0.3 *M* NaCl and these active fractions were pooled and concentrated by the addition of 80% $(NH_4)_2SO_4$. The resultant pellet was resuspended in a minimal amount of buffer *A* with 5% $(NH_4)_2SO_4$. After centrifugation, the supernatant was applied onto a pre-packed Pharmacia Superose 12 column (3.5 × 30 cm) equilibrated in buffer *A* with 5% $(NH_4)_2SO_4$. Active fractions from this final step were pooled and concentrated using Amicon Centricons (10 kDa cutoff).

The purified enzyme was subjected to sedimentation-equilibrium and velocity analysis (NCMH, Universities of Nottingham and Leicester) to determine the oligomeric state of the protein and laser-desorption mass spectrometry (University of Aberdeen, Protein Sequencing Facility) to assess the protein homogeneity and subunit mass.

2.2. Crystallization

Crystallization was carried out by the vapour-diffusion method at 290 K. A wide screen of precipitants was used as a preliminary step to determine conditions from which crystals would grow. These included ammonium sulfate with 10 m*M* Bicine pH 8.5, 20 m*M* Tris–HCl pH 8.0 and 10 m*M* HEPES pH 7.5 with protein concentrations of between 10 and 15 mg ml^{-1}, polyethylene glycols 400 and 600 in 10 m*M* Bicine pH 8.5 and 10 m*M* HEPES pH 7.5 with protein concentrations of 15–20 mg ml^{-1} and Crystal Screen I and II (Hampton Research, USA) with protein at 10 mg ml^{-1}. Crystals were set up initially as 5 µl hanging drops and later as sitting drops.

In order to improve the quality of the crystals grown from ammonium sulfate the protein concentration and pH were varied and additives were tested by adding 10% of different Crystal Screen I reagents to a reservoir of 35% ammonium sulfate in Bicine buffer. The best crystals grown from 35% ammonium sulfate and 0.5% PEG 4000 as an additive were further optimized by varying the protein concentration. Crystal size was improved by employing an oil barrier (Chayen, 1997) spread onto the reservoir prior to positioning the cover slip. The oil barrier used was 35% paraffin oil/65% silicone fluid in 500 µl.

The second crystal form of the enzyme was grown from 38% saturated sodium malonate in 100 m*M* Bicine buffer pH 9.0. The crystals were improved by screening the addition of organic additives to the reservoir. The best crystals were grown by the addition of 3% ethanol or ethylene glycol.

2.3. Crystallographic data collection

For X-ray data collection, the crystals of first form were harvested into a cryoprotectant of 45% ammonium sulfate, 25% glycerol, 10 m*M* Bicine pH 8.5. Native data were collected at 100 K at beamline 7.2 of the Daresbury Synchrotron source from a single crystal at a wavelength of 1.488 Å. Putative derivative data were collected at EMBL station BW7B, DESY Synchrotron Hamburg on two crystals soaked in cryoprotectant containing either 5 m*M* uranyl acetate for 1 h or 10 m*M* mercury acetate for 4 h.

Crystals of the second crystal form were harvested into cryoprotectant containing 50% saturated sodium malonate in 100 m*M* Bicine buffer pH 9.0. Native data were collected at 100 K from a single crystal at Max-Planck Institute station BW6, DESY Synchrotron Hamburg at a wavelength of 1.05 Å. The data were collected on EMBL station BW7A from a crystal soaked for 1 h in cryoprotectant containing 10 m*M* mercury acetate. All data were processed using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

The naproxen esterase eluted as a single peak from the gel-filtration column and appeared to run as a single band on SDS–PAGE with an apparent molecular weight of 44 kDa.

Crystallization trials carried out with enzyme expressed from the pUC19 variant plasmid were unsuccessful. A sample of this protein, pure with respect to SDS–PAGE, exhibited two protein species, one of 48 288 kDa and one of 47 456 kDa, present in a 7:3 ratio when examined by mass spectrometry. N-terminal sequencing determined that part of the pUC19 plasmid had been transcribed upstream of the naproxen esterase gene. As a consequence, the sample contained two protein species that were essentially identical, only differing by nine amino acids at the N-terminus. This phenomenon may have resulted from either

two separate translation reactions or from post-translational proteolysis of the extra nine amino acids.

Naproxen esterase expressed from the pKK233-2 expression system was shown to be a single species by N-terminal sequencing and sedimentation-equilibrium and velocity analysis. The protein was determined to be a monomer of 47 179 ± 340 Da. Initial crystallization trials produced clusters of very fine needles from 34–36% $(NH_4)_2SO_4$,10 m*M* Bicine pH 8.5. Larger crystals were grown using 0.5% PEG 4000 as an additive. Oil barriers spread over the reservoir reduced the amount of clustering, which allowed the growth of single crystals. Rod-like crystals achieved dimensions of 0.7 × 0.2 × 0.1 mm.

The heterogeneity of protein samples expressed from the pUC19 variant plasmid is believed to have prohibited crystallization. A naproxen esterase sample that was shown to be homogeneous crystallized readily. This is in agreement with observations by Abergel *et al.* (1991) that nucleation of turkey egg-white lysozyme is inhibited by controlled contamination with hen and quail lysozyme.

The native data for the first crystal form were collected to 2.6 Å (Table 1). The crystals belong to the orthorhombic space group $P2_12_12$, with unit-cell parameters a = 115.9, b = 174.4, c = 62.1 Å. The solvent content of the crystals, which contain two molecules in the asymmetric unit, has been estimated at 62% (V_M = 3.3 Å Da^{-1}; Matthews, 1968).

To obtain an improved cryoprotectant, the protein was crystallized from sodium malonate, which produced the second crystal form, which grew to dimensions of 0.8 × 0.15 × 0.1 mm. These crystals belong to space group $P2_12_12$, with unit-cell parameters a = 73.0, b = 212.7, c = 61.7 Å. The native data from this crystal form have been collected to 2.3 Å (Table 1). The solvent content of the crystals, which contain two molecules in the asymmetric unit, has been estimated at 51% (V_M = 2.5 Å Da^{-1}).

The self-rotation function calculated for both crystal forms did not show any significant features (data not shown). The similarity of the unit-cell parameter c between the two lattices suggests that the two unit cells might be related. To check this hypothesis, the cross-rotation function was calculated between the two native data sets using the *BLANC* complex of crystallographic programs (Vagin *et al.*, 1998). The cross-rotation function showed strong features at section $\beta = 0, \alpha - \gamma = 47^\circ$ (Fig. 1). This indicates that molecules in the two crystal lattices are related by a rotation of 47° around the crystallographic z axis.

Crystal non-isomorphism in both crystal forms prevented the use of the MIR approach for structure solution. Initial attempts to use molecular replacement (MR) with either the D-alanyl-D-alanine carboxypeptidase (Kelly *et al.*, 1985; sequence identity 14%) or the *E. coli* Ampc β-lactamase (Patera *et al.*, 2000; sequence identity 16%) as a model failed to find a solution which could be refined further.

When a monomeric model of EstB from *B. gladioli* (Wagner *et al.*, 2002), which shares 23% sequence identity with the naproxen esterase, became available the program *MOLREP* (Vagin & Teplyakov, 1997) was used for MR. The correct MR solutions for two naproxen esterase monomers in the asymmetric unit of the first crystal form were found only when a new FILE_SEQUENCE option of *MOLREP* was used. When this option is activated *MOLREP* reads the sequence of the target structure and truncates the search model according to the sequence alignment. Default values were used for other *MOLREP* options. The correct rotation-function peaks were the third and twelfth in the rotation-function list. Subsequent MR studies have shown that the sequence-based truncation of the model only marginally improved the contrast for the translation search. However, without this truncation the correct solution for one of the two molecules was not amongst the highest 200 peaks of the rotation function.

The EstB model positioned in the first crystal form was subjected to 100 cycles of positional and B-factor refinement in *REFMAC* (Murshudov *et al.*, 1997). Although R_{free} did not improve during the course of this refinement, the output model was successfully used for MR in the second crystal form. The MR solution could not be found for this crystal form with the unrefined EstB model. Multicrystal averaging using the program *DMMULTI* (Cowtan & Main, 1998; Collaborative Computational Project, Number 4, 1994) produced an interpretable electron-density map. The atomic models have been refined by *REFMAC* in both crystal forms to an R_{free} of below 30%.

The asymmetric unit of each crystal form contains a naproxen esterase dimer. The r.m.s. deviation of the C^α-atom positions between the dimers in the two different crystal forms is only 1.13 Å. However, the two monomers in the naproxen esterase dimer are not related by a dyad, but by a rotation of 139°. As was originally suggested from the cross-rotation function, the dimers in two crystal forms are related by a rotation of 47° around the crystallographic z axis. One subunit of the dimer in the second crystal form is significantly less ordered than the other subunit. This could explain the earlier failure to find the translation-function solution in this crystal form using

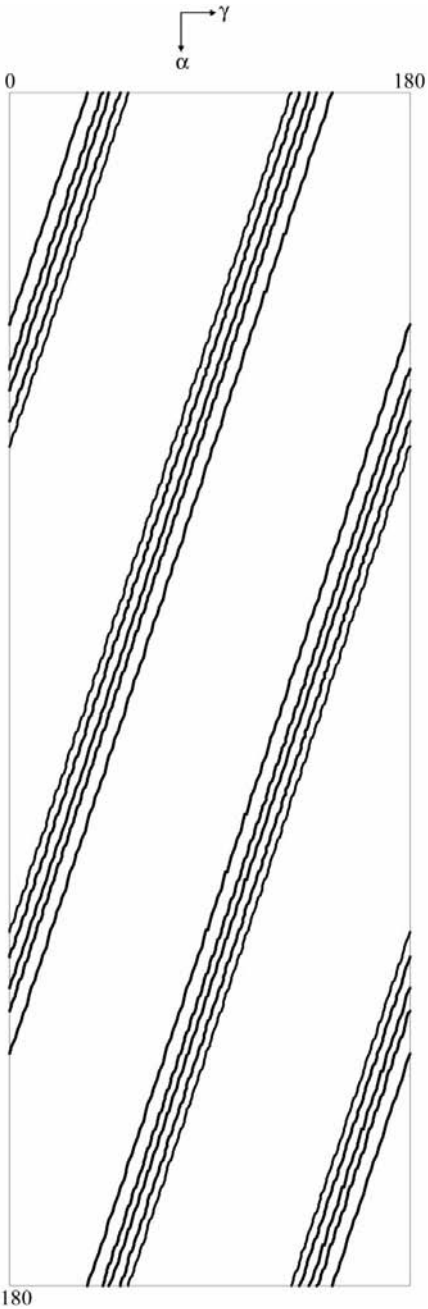

Figure 1
β = 0° section of the cross-rotation function calculated between the data from the two crystal forms of naproxen esterase. The calculations were perfomed with an integration radius of 30 Å at 14–3 Å resolution using the program suite *BLANC* (Vagin *et al.*, 1998). The strong feature at $\alpha - \gamma = 47^\circ$ indicates that molecules in the two crystal lattices are related by a rotation around the crystallographic z axis.

the relationship between the two crystal forms.

Although macromolecules often crystallize in several crystal forms, a cross-rotation function between the data is not routinely calculated. It would appear that use of the cross-rotation function calculated between data sets in different lattices could benefit difficult phasing problems when one of the lattices is triclinic and contains a single biological unit. The cross-rotation function calculation could also be useful for higher symmetry crystal lattices in instances when there is a putative relationship between their unit-cell parameters. Naproxen esterase is an example of a case where the cross-rotation function calculated for data in the two orthorhombic lattices has provided useful information about the relative arrangement of the molecules between the two crystal forms prior to structure solution.

We are currently refining the structure of the naproxen esterase in order to understand the relationship of this enzyme to β-lactamases. Experiments to soak or co-crystallize the naproxen esterase with substrates/inhibitors are under way in order to define the enzyme's active site and its substrate specificity.

Thanks are given to Gleb Bourenkov at the Max Planck Institute, Hamburg for help with data collection. This work was supported by the Biotechnology and Biological Sciences Research Council as a CASE studentship to AAB (with Chirotech Technology Ltd, Cambridge, UK) and postdoctoral fellowships to MNI and AAV. GNM thanks the Wellcome Trust. We acknowledge support under the European Community Access to Research Infrastructure Action of the Improving Human Potential Programme to EMBL Hamburg Outstation, contract No. HPRI-CT-1999-00017.

References

Abergel, C., Nesa, M. P. & Fontecilla-Camps, J. C. (1991). *J. Cryst. Growth*, **110**, 11–19.
Arpigny, J. L. & Jaeger, K. E. (1999). *Biochem. J.* **343**, 177–183.
Brindley, A. A. (1998). PhD thesis, University of Exeter, England.
Chayen, N. E. (1997). *J. Appl. Cryst.* **30**, 198–202.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Cowtan, K. & Main, P. (1998). *Acta Cryst.* D**54**, 487–493.
Faber, K. (1997). *Pure Appl. Chem.* **69**, 1613–1632.
Fadel, A. (1992). *Synlett*, **1**, 48–50.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Kelly, J. A., Knox, J. R., Moews, P. C., Hite, G. J., Bartolone, J. B., Zhao, H., Joris, B., Frere, J. M. & Ghuysen, J. M. (1985). *J. Biol. Chem.* **260**, 6449–6458.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **267**, 21839–21843.
Patera, A., Blaszczak, L. C. & Shoichet, B. K. (2000). *J. Am. Chem. Soc.* **122**, 10504–10512.
Petersen, E. I., Valinger, G., Solkner, B., Stubenrauch, G. & Schwab, H. (2001). *J. Biotechnol.* **89**, 11–25.
Sobek, H. & Gorisch, H. (1988). *J. Biochem* **250**, 453–458.
Vagin, A. A., Murshudov, G. N. & Strokopytov, B. V. (1998). *J. Appl. Cryst.* **31**, 98–102.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Wagner, U. G., Petersen, E. I., Schwab, H. & Kratky, C. (2002). *Protein Sci.* **11**, 467–478.