Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Expression, purification and preliminary X-ray diffraction analysis of a ketoreductase from a type II polyketide synthase

Watchrra Teartasin,[a,b] Claire Limpkin,[a,c] Frank Glod,[a,b] James Spencer,[a,d] Russell J. Cox,[a,b] Thomas J. Simpson,[a,b] John Crosby,[a,b] Matthew P. Crump[a,b] and Andrea T. Hadfield[a,c]*

[a]Molecular Recognition Centre, University of Bristol, School of Medical Sciences, University Walk, Bristol BS8 1TD, England, [b]Organic and Biological Chemistry, School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, England, [c]Department of Biochemistry, University of Bristol, School of Medical Sciences, University Walk, Bristol BS8 1TD, England, and [d]Department of Pathology and Microbiology, University of Bristol, School of Medical Sciences, University Walk, Bristol BS8 1TD, England

Correspondence e-mail: a.t.hadfield@bris.ac.uk

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

Received 10 March 2004
Accepted 31 March 2004

Polyketide metabolites produced by bacteria and other organisms include antibiotics, anticancer and antifungal compounds. In type II polyketide synthesis, three enzymes are sufficient to form a polyketide product of the requisite chain length, although the fidelity of the first cyclization is variable. Addition of ketoreductase (KR) to this system results in the formation of a product with correct cyclization and reduction. This paper reports the cloning of the *Streptomyces coelicolor actIII* ORF5 gene that codes for the ketoreductase. The 261-amino-acid protein has been overexpressed with a 20-residue His tag, purified by affinity chromatography and crystallized in space group $P3_221$, with unit-cell parameters $a = b = 103.9$, $c = 123.1$ Å. The crystals diffract to 2.5 Å resolution. A complete data set has been collected and structure solution and refinement is under way.

1. Introduction

Polyketides are a chemically diverse group of secondary metabolites produced by bacteria, fungi, plants and marine organisms (Hopwood, 1990, 1997; Staunton & Weissman, 2001). These natural products are of enormous pharmaceutical interest and include oxytetracycline (an antibiotic), doxorubicin (anticancer) and resveratrol (antifungal). The polyketide synthase (PKS) biosynthetic complex is organized as either a covalent (type I) or non-covalent (type II) assembly of catalytic domains analogous to fatty-acid synthases (FASs). Both polyketide and fatty-acid biosyntheses proceed *via* the iterative decarboxylative condensation of malonyl extender units to a starter unit, which is usually acetate.

In type II PKSs, the minimal number of enzymes required to form a polyketide of the requisite chain length are the ketosynthase (KSα), chain-length factor (KSβ) and the acyl-carrier protein (ACP) (McDaniel *et al.*, 1994, 1995). The KR is the first enzyme to modify the nascent polyketide chain, its primary function being to reduce the carbonyl group at C-9. KR may also enhance the efficiency of the aldol condensation between C-7 and C-12 which is necessary to form the first aromatic ring correctly, since addition of the *act* KR to the minimal PKS eliminates incorrectly cyclized products (Fu *et al.*, 1994*a*,*b*; Fu, Ebert-Khosla *et al.*, 1994).

To date, only four type II aromatic PKS components have been solved. These are three solution NMR structures of the *act* (Crump *et al.*, 1997), *otc* (Findlow *et al.*, 2003) and *fren* (Li *et al.*, 2003) ACPs and a crystal structure of the unusual priming β-ketosynthase from the R1128 polyketide-biosynthetic pathway (Pan *et al.*, 2002). There is currently no three-dimensional structure of a polyketide type II ketoreductase or indeed any other downstream enzyme from these pathways. The *act* KR (29 kDa, 261 amino acids) is a member of the short-chain dehydrogenase/reductase evolutionary family of NADP(H) dependent enzymes that contains over 60 genes in the human genome alone (Jornvall *et al.*, 1995; Oppermann *et al.*, 2001, 2003; Kallberg *et al.*, 2002).

2. Materials and methods

2.1. Cloning and expression

The *Streptomyces coelicolor actIII* ORF5 gene was cloned from genomic DNA isolated from the organism by PCR amplification using primers that added an *Nde*I restriction site at the N-terminus and a *Hin*dIII site at the C-terminus. The restriction sites were digested and the gene was first ligated into the commercial vector pET26b (Novagen). Following successful sequencing, the gene was excised from this non-tagged vector by a restriction digest using *Nde*I and *Xho*I (inherent to pET26b). The same restriction sites were used to insert the gene into the plasmid pET15b to incorporate an N-terminal His_6 tag at the N-terminus, which results in a total of 20 additional residues (MGSSHH-HHHHSSGLVPRGSH). This plasmid was transformed into *Escherichia coli* strain BL21(DE3) for IPTG-induced expression. Cultures were produced by inoculation of single colonies from LB agar plates (1 ml). These precultures were grown overnight at

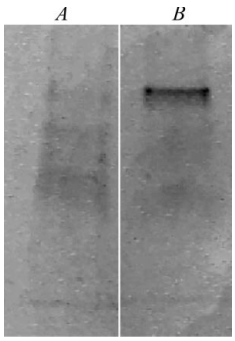

Figure 1
Native PAGE analysis of ketoreductase. Lane *A* shows 17 µl ketoreductase in Tris buffer pH 9. Lane *B* shows 17 µl ketoreductase in the same buffer plus 20 m*M* NADPH.

Table 1
Summary of diffraction data analysis.

Space group	Hexagonal, $P3_221$
Unit-cell parameters (Å, °)	a = 103.9, b = 103.9, c = 123.1, α = 90, β = 90, γ = 120
Diffracting volume (mm^3)	2.5×10^{-4}
Matthews coefficient (Å^3 Da^{-1})	3.3
Solvent content (%)	62.5
No. molecules in the AU	2
Total No. unique reflections collected	174794
No. observed reflections (average multiplicity)	26930 (6.5)
Wavelength used (Å)	0.97
Resolution range (Å)	20.0–2.4
Highest resolution shell observed (Å)	2.53–2.4
Highest resolution shell used (Å)	2.69–2.53
Overall completeness to 2.53 Å (%)	100
Completeness in highest resolution shell used (%)	100
$R_{\rm merge}$	0.073
$R_{\rm merge}$ for highest resolution shell used	0.448
$\langle I/\sigma(I)\rangle$	20.0

310 K. Overnight cultures were then used to typically inoculate 0.5–1 l of LB liquid media. These were grown at 310 K to an A_{595} of between 0.6 and 0.8. Cultures were then induced by adding IPTG to a final concentration of 1 m*M*. The cells were disrupted by sonication in 3 × 30 s bursts. The cell debris was removed by centrifugation at 6000*g* for 30 min. The fully active enzyme was purified from the cell extract in a single step by affinity chromatography on an Ni^{2+} column (Novagen). Fractions containing KR were desalted into buffer *A* (100 m*M* phosphate, 10% glycerol pH 7.2) by FPLC and stored at 193 K.

The activity assay was performed by adding ketoreductase, NADPH (cofactor) and malonyl CoA (substrate) to the PKS minimal system. Typically, the assay contained 50 µ*M* ACP, 1 µ*M* KS/CLF, 10 µ*M* KR, 0.1 m*M* DTT, 2 m*M* EDTA, 1 m*M* malonyl CoA, 2 m*M* NADPH and 10% glycerol, buffered with 0.1 m*M* KH_2PO_4 pH 7.3 in a final volume of 100 µl. Polyketide metabolites were extracted with ethyl acetate, freeze-dried and dissolved in methanol for HPLC analysis. The presence of mutactin was used to confirm the presence of active KR.

2.2. Crystallization and data collection

The protein was concentrated to 10 mg ml^{-1} in 50 m*M* phosphate buffer pH 8 with 10% glycerol. Comparison with other SDR structures suggests that the ketoreductase oligomeric state is likely to be tetrameric. This was investigated by native PAGE (Fig. 1). In the absence of NADPH, the enzyme appears to exist in a mixture of oligomeric states (tetramer, dimer and monomer). Addition of NADPH to the enzyme under otherwise identical conditions results in a tight band representing the tetramer. $NADP^+$ was therefore added to the concentrated protein, which reduced the phosphate concentration and also ensured that the enzyme existed in the holo form and in a homogenous oligomeric state. The mother liquor used for crystallization contained 5 mg ml^{-1} protein and 5 m*M* $NADP^+$ in 25 m*M* phosphate buffer pH 8 with 10%(*v*/*v*) glycerol. Crystallization conditions were identified using a sparse-matrix screen (MDL Screen I and II) using hanging-drop vapour diffusion at 291 K. Data were collected from crystals obtained from wells containing 4.0–4.4 *M* sodium formate and 10%(*v*/*v*) glycerol after combination of equal volumes of reservoir and mother liquor to form 3 µl hanging drops. The crystals belong to space group $P3_221$, with unit-cell parameters $a = b = 103.9$, $c = 123.1$ Å, $\alpha = \beta = 90.0$, $\gamma = 120.0°$. The small (0.1 × 0.05 × 0.05 mm) gem-like crystals were transferred briefly into a cryoprotectant made up of 25%(*v*/*v*) glycerol and 75%(*v*/*v*) well solution prior to flash-cooling directly in the nitrogen-gas stream. Diffraction data that were 100% complete to 2.5 Å resolution were collected at 100 K using synchrotron radiation at station 14.2 at the Daresbury Laboratory SRS (25 369 reflections in the resolution range 20–2.5 Å) and reduced using *HKL*2000 (HKL Research Inc) (Table 1). The solvent content of the crystals suggests that each asymmetric unit contains a dimer. To be consistent with a tetrameric oligomerization state, this requires a twofold symmetry axis of the tetramer to be coincident with a twofold crystallographic axis and the remaining twofold symmetry axes to be non-crystallographic.

The authors would like to thank David Hopwood for the gift of *S. coelicolor* genomic DNA.

References

Crump, M. P., Crosby, J., Dempsey, C. E., Parkinson, J. A., Murray, M., Hopwood, D. A. & Simpson, T. J. (1997). *Biochemistry*, **36**, 6000–6008.
Findlow, S. C., Winsor, C., Simpson, T. J., Crosby, J. & Crump, M. P. (2003). *Biochemistry*, **42**, 8423–8433.
Fu, H., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1994). *J. Am. Chem. Soc.* **116**, 4166–4170.
Fu, H., McDaniel, R., Hopwood, D. A. & Khosla, C. (1994*a*). *Biochemistry*, **33**, 9321–9326.
Fu, H., McDaniel, R., Hopwood, D. A. & Khosla, C. (1994*b*). *Chem. Biol.* **1**, 205–210.
Hopwood, D. A. (1990). *Annu. Rev. Genet.* **24**, 37–66.
Hopwood, D. A. (1997). *Chem. Rev.* **97**, 2465–2497.
Jornvall, H., Persson, B., Krook, M., Atrian, S., Gonzalezduarte, R., Jeffery, J. & Ghosh, D. (1995). *Biochemistry*, **34**, 6003–6013.
Kallberg, Y., Oppermann, U., Jornvall, H. & Persson, B. (2002). *Protein Sci.* **11**, 636–641.
Li, Q., Khosla, C., Puglisi, J. D. & Liu, C. W. (2003). *Biochemistry*, **42**, 4648–4657.
McDaniel, R., Ebert-Khosla, S., Fu, H., Hopwood, D. A. & Khosla, C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 11542–11546.
McDaniel, R., Ebert-Khosla, S., Hopwood, D. A. & Khosla, C. (1995). *Nature (London)*, **375**, 549–554.
Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X. Q., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. & Jornvall, H. (2003). *Chem. Biol. Interact.* **143**, 247–253.
Oppermann, U. C. T., Filling, C. & Jornvall, H. (2001). *Chem. Biol. Interact.* **130**, 699–705.
Pan, H., Tsai, S. C., Meadows, E. S., Miercke, L. J. W., Keatinge-Clay, A. T., O'Connell, J., Khosla, C. & Stroud, R. M. (2002). *Structure*, **10**, 1559–1568.
Staunton, J. & Weissman, K. J. (2001). *Nat. Prod. Rep.* **18**, 380–416.