Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic investigations of an unusual type III polyketide synthase PKS18 from *Mycobacterium tuberculosis*

Raju Rukmini,[a] Vellaiah M. Shanmugam,[a] Priti Saxena,[b] Rajesh S. Gokhale[b] and Rajan Sankaranarayanan[a]*

[a]Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India, and [b]National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

Correspondence e-mail: sankar@ccmb.res.in

Received 19 December 2003
Accepted 29 January 2004

The biosynthetic machinery of polyketide synthases involves various sequential enzymatic reactions, such as initiation, elongation and cyclization, to produce polyketides. PKS18 protein from *Mycobacterium tuberculosis* belongs to the type III polyketide synthase family and displays an unusual starter-unit specificity to catalyze the formation of α-pyrones. This enzyme uses malonyl-CoA to iteratively extend long-chain aliphatic coenzyme A (C_{12} to C_{20}) molecules, producing triketide and tetraketide pyrone products. In order to aid in understanding the structural basis of this long-chain specificity and to further characterize the enzymatic mechanism of PKS18, the protein has been crystallized. The crystal belongs to the triclinic space group $P1$, with unit-cell parameters a = 59.9, b = 80.7, c = 99.6 Å, α = 108.2, β = 93.0, γ = 103.7°.

1. Introduction

Polyketide synthases (PKSs) are a class of enzymes that are involved in the biosynthesis of secondary metabolites such as rapamycin, erythromycin, tetracycline, lovastatin and resveratrol. PKSs are widely distributed in nature and are found in various bacteria, fungi and plants (O'Hagan, 1991). These enzymes catalyze various sequential enzymatic reactions, carrying out the biosynthesis of polyketide products. Polyketides have a broad spectrum of biological activity, exhibiting potential for pharmaceutical and industrial applications (O'Hagan, 1993). Despite the structural diversity of these natural products, PKSs utilize a common strategy for polyketide synthesis. The starter molecule primes the initiation reaction step, which is followed by a repetitive decarboxylative condensation of coenzyme A (CoA) analogues of simple dicarboxylic acids. The growing polyketide chain is elongated either by repetitive use of the single active site to perform multiple condensation reactions or by utilization of a modular assembly-line mechanism (Katz & Donadio, 1993; Hopwood, 1997; Khosla *et al.*, 1999).

PKSs have been classified into three distinct families on the basis of protein architecture and the mechanism of polyketide biosynthesis. Type I PKSs are large assemblies of multifunctional polypeptides which use both iterative and modular mechanisms of biosynthesis. Type II PKSs are large multienzyme complexes of discrete proteins and resemble the type II fatty-acid synthases found in bacteria and plants. Type III PKSs have recently been discovered in bacteria and have been affiliated to the superfamily of plant chalcone synthases (CHSs; Moore & Hopke, 2001). These homodimeric proteins are structurally and mechanistically quite distinct from type I and II PKSs (Gokhale & Tuteja, 2001). CHSs are ubiquitously present in higher plants and catalyze the biosynthesis of starter molecules for many flavonoids (Schroder, 2000). Three-dimensional crystal structure analysis of plant type III PKSs has shown a conserved $\alpha\beta\alpha\beta\alpha$-fold architecture that is characteristic of the superfamily of condensing enzymes (Ferrer *et al.*, 1999). There are no reports of crystallographic studies of bacterial type III PKSs, although some critical residues in the catalytic cavity of the plant CHS family are also conserved in bacterial type III polyketide synthases.

The genome sequence of *Mycobacterium tuberculosis* has been shown to contain a remarkable array of genes that are homologous to PKSs (Cole *et al.*, 1998; Funa *et al.*, 1999). To date, no type III PKS product has been characterized from *M. tuberculosis*, although there are indications that some of its PKS genes are implicated in the biosynthesis of complex lipids (Kolattukudy *et al.*, 1997). Recently, a new family of type III polyketide synthases has been characterized in *M. tuberculosis* (Saxena *et al.*, 2003). One of its members, PKS18, a 393-amino-acid protein with a molecular weight of 45.0 kDa, exhibits a very high specificity towards long-chain aliphatic acyl-coenzyme A (C_{12} to C_{20}) substrates. PKS18 catalyzes the synthesis of triketide and tetraketide α-pyrones by repetitively extending these long-chain starter substrates with malonyl-CoA. The high specificity of PKS18 for long-chain starter substrates is unprecedented in the chalcone synthase

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

family of condensing enzymes (Saxena *et al.*, 2003). In order to elucidate the structural basis of this unusual starter-molecule specificity and to understand the enzymatic mechanism, we have undertaken a structural study of PKS18.

2. Experimental

2.1. Expression and purification of PKS18

The *pks18* gene was cloned into expression vector pET21c with a hexahistidine tag at the C-terminus of the pks18 gene product, as previously reported (Saxena *et al.*, 2003). The recombinant PKS18 protein with the C-terminal hexahistidine tag was transformed and overexpressed in *Escherichia coli* BL21(DE3) strain cells. Cells were grown in Luria–Bertani medium containing 100 µg ml^{-1} ampicillin at 303 K until OD_{600} reached ∼1.5 absorbance units and then induced at 291 K with 0.5 m*M* isopropyl-D-thiogalactopyranoside. Cells were harvested by centrifugation at 4000*g* for 10 min at 277 K. The pellet was resuspended in ice-cold lysis buffer containing 50 m*M* Tris–HCl pH 8.0 and lysed by sonication in ice. The crude cell extract was centrifuged at 40 000*g* for 45 min at 277 K. Purification of the supernatant was performed by two chromatographic steps. The first step involved purification of the hexahistidine-tagged protein by passing the supernatant through an Ni^{2+}–NTA agarose column. PKS18 was eluted by washing the column with 50 m*M* and 100 m*M* imidazole in 50 m*M* Tris–HCl pH 8.0 buffer. Both of the eluted protein fractions were pooled and further purified to homogeneity by gel filtration with a Superdex-75 chromatography column. The purity of the protein was assayed by 10% SDS–PAGE and was found to be more than 95% pure. Fractions containing purified protein in 50 m*M* Tris–HCl pH 8.0 buffer with 150 m*M* NaCl, 1 m*M* DTT and 1 m*M* EDTA were pooled and concentrated to 3 mg ml^{-1} by ultrafiltration using Centriprep and Centricon YM-10 devices. Protein concentration was determined by Bradford reagent assay.

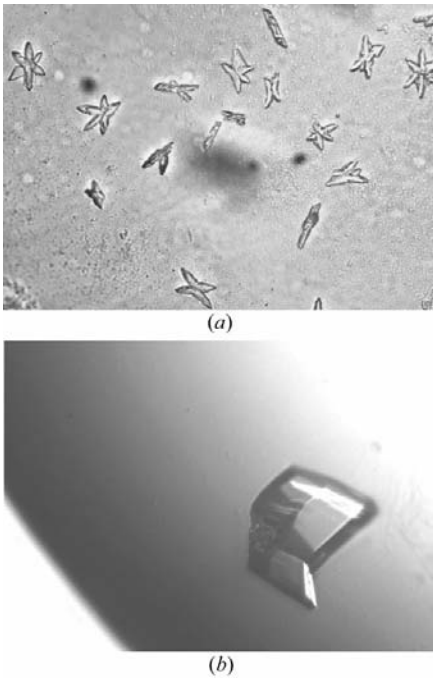

Figure 1
Crystals of PKS18 from *M. tuberculosis*: (*a*) star-like clusters of crystals grown at room temperature and (*b*) butterfly-shaped crystals obtained by refining growth conditions at 277 K.

2.2. Crystallization and preliminary X-ray data

Crystallization experiments were set up by the hanging-drop vapour-diffusion method using 24-well plates. Initial crystallization attempts were carried out at room temperature using Crystal Screens from Hampton Research. A hanging drop was prepared on a siliconized cover slip by mixing equal volumes (2 µl each) of protein solution containing 3 mg ml^{-1} protein and the reservoir solution.

Preliminary diffraction data for PKS18 were collected on an in-house MAR Research MAR-345dtb image-plate detector with Cu $K\alpha$ X-rays generated by a Rigaku RU-H3R rotating-anode generator equipped with an Osmic mirror system and operated at 50 kV and 100 mA. Prior to flash-cooling in a liquid-nitrogen stream at 100 K for data collection, crystals were soaked in a cryoprotectant solution containing 35% ethylene glycol in addition to the mother liquor for 10–15 s. X-ray data were processed using *DENZO* (Otwinowski & Minor, 1997) and the subsequent scaling and merging of intensities was carried out using *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results and discussion

Recombinant PKS18 with a C-terminal hexahistidine tag was expressed in soluble form in *E. coli*. 1 l of culture yielded up to 2–3 mg pure protein after affinity and gel-filtration chromatography. Star-like clusters of PKS18 crystals appeared in 10%(*w*/*v*) PEG 8000 and 0.1 *M* HEPES pH 7.5 with 8%(*v*/*v*) ethylene glycol at room temperature (Fig. 1*a*). Well diffracting butterfly-shaped crystals were obtained after optimization in 10%(*w*/*v*) PEG 8000, 0.1 *M* HEPES pH 7.7 and 10%(*v*/*v*) ethylene glycol and were grown at 277 K (Fig. 1*b*) after altering the original crystallization conditions. Crystals of smaller size (∼50–70 µm) diffracted X-rays to a resolution limit of 2.25 Å, whereas larger crystals had a very high mosaicity and were not used for data collection.

Table 1
Crystallographic data.

Values in parentheses are for the highest resolution shell.

Wavelength λ (Å)	1.5418
Space group	*P*1
Unit-cell parameters	
a (Å)	59.9
b (Å)	80.7
c (Å)	99.6
α (°)	108.2
β (°)	93.0
γ (°)	103.7
Resolution (Å)	25.0–2.25 (2.33–2.25)
Observations	134485 (12524)
Unique reflections	76938 (7360)
Completeness (%)	95.2 (91.2)
Multiplicity	1.74 (1.70)
$R_{\rm merge}$† (%)	7.0 (30.7)
$I/\sigma(I)$	9.5 (2.0)
Solvent content (%)	50.4
$V_{\rm M}$ (Å^3 Da^{-1})	2.5
Monomers per AU	4

† $R_{\rm merge} = \sum|I(h) - \langle I(h)\rangle|/\sum I(h)$, where $I(h)$ is the observed intensity and $\langle I(h)\rangle$ is the mean intensity of reflection h over all measurements of $I(h)$.

Analysis of the diffraction pattern indicated that the crystal belongs to the triclinic space group *P*1, with unit-cell parameters a = 59.9, b = 80.7, c = 99.6 Å, α = 108.2, β = 93.0, γ = 103.7°. Unit-cell volume calculations suggested four monomers of PKS18 per asymmetric unit. A complete data set was collected by rotating the crystal by 0.5° oscillations through a total of 160.5°, with a crystal-to-detector distance of 170 mm. The mosaicity of the crystal was found to be 0.74. The crystal data-collection statistics are shown in Table 1. The overall completeness and $R_{\rm merge}$ value are 95.2 and 7.0%, respectively. A self-rotation function map calculated using data in the 10.0–5.5 Å resolution range showed three peaks corresponding to a rotation of 180° with heights of 75, 38 and 35% of the origin peak, thus indicating the presence of two independent dimeric molecules in the asymmetric unit. In addition, a single peak was observed corresponding to a rotation of 48.3° with a height of 35% of the origin peak. Initial structure solution by the molecular-replacement method using the program *MOLREP-AUTO MR* (Vagin & Teplyakov, 1997), as implemented in the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994), was attempted with alfalfa chalcone synthase (PDB code 1cgz) as a model (Ferrer *et al.*, 1999). PKS18 exhibits a

sequence homology of 42% with alfalfa CHS (Saxena *et al.*, 2003). A poly-alanine monomer search model enabled the location of four molecules in the asymmetric unit, as expected from the cell-content analysis. The resulting solution was found to be consistent with the peaks obtained in the self-rotation function analysis. The two functional dimeric molecules *AB* and *CD* are related by the second and third largest peaks obtained with 180° rotation, whereas the dimers are related by the highest peak in the self-rotation map. In addition, the monomers *AD* and *BC* are related by a 48° rotation. Further refinement and analysis of the structure are in progress.

The authors thank Ms Sirisha Tatavarti for technical help. RS and RSG are Wellcome Trust International Senior Research Fellows in India.

References

Cole, S. T. *et al.* (1998). *Nature (London)*, **393**, 537–544.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Ferrer, J. L., Jez, J. M., Bowman, M. E., Dixon, R. A. & Noel, J. P. (1999). *Nature Struct. Biol.* **6**, 775–784.
Funa, N., Ohnishi, Y., Fujii, I. & Shibuya, M. (1999). *Nature (London)*, **400**, 897–899.
Gokhale, R. S. & Tuteja, D. (2001). In *Biotechnology*, edited by H.-J. Rehm, Vol. 10. Weinheim: Wiley–VCH.
Hopwood, D. A. (1997). *Chem. Rev.* **97**, 2465–2498.
Katz, L. & Donadio, S. (1993). *Annu. Rev. Microbiol.* **47**, 875–912.
Khosla, C., Gokhale, R. S., Jacobsen, J. R. & Cane, D. E. (1999). *Annu. Rev. Biochem.* **68**, 219–253.
Kolattukudy, P. E., Fernandes, N. D., Azad, A. K., Fitzmaurice, A. M. & Sirakova, T. D. (1997). *Mol. Microbiol.* **24**, 263–270.
Moore, B. S. & Hopke, J. N. (2001). *Chembiochem*, **2**, 35–38.
O'Hagan, D. (1991). In *The Polyketide Metabolites.* Chichester: Ellis Horwood.
O'Hagan, D. (1993). *Nat. Prod. Rep.* **10**, 593–624.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Saxena, P., Yadav, G., Mohanty, D. & Gokhale, R. S. (2003). *J. Biol. Chem.* **278**, 44780–44790.
Schroder, J. (2000). In *Evolution of Metabolic Pathways*, edited by J. T. Romeo, R. K. Ibrahim, L. Varin & V. De Luca, Vol. 34, pp. 55–89. Amsterdam: Pergamon Press.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022-1025.