Crystallization and preliminary X-ray diffraction analysis of the diterpene cyclooctatin synthase (CYC) from *Streptomyces* sp. LZ35

Xiulei Zhang,^a Guijun Shang,^b
Lichuan Gu^{b*} and Yuemao
Shen^{a*}

^aKey Laboratory of Chemical Biology (Ministry of Education), School of Pharmaceutical Sciences, Shandong University, Jinan, Shandong 250012, People's Republic of China, and ^bState Key Laboratory of Microbial Technology, School of Life Sciences, Shandong University, Jinan, Shandong 250100, People's Republic of China

Correspondence e-mail: lcgu@sdu.edu.cn,
yshen@sdu.edu.cn

Received 9 December 2013

Accepted 11 February 2014

Terpenoids are a large and highly diverse group of natural products, with the most chemically diverse pool of structures. Terpene synthase is the key enzyme in the process of terpenoid synthesis. In this paper, the first diterpene synthase (CYC) of bacterial origin was successfully crystallized. Native and SeMet-derivative crystals diffracted to 1.75 and 2.6 Å resolution, respectively. The native crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 59.10$, $b = 101.73$, $c = 108.93$ Å, and contained two molecules per asymmetric unit. The SeMet-derivative crystal belonged to space group $P2_1$, with unit-cell parameters $a = 58.64$, $b = 109.47$, $c = 58.73$ Å, $\beta = 119.35^\circ$, and had two molecules per asymmetric unit.

1. Introduction

With more than 60 000 members known to date according to the Dictionary of Natural Products (<http://dnp.chemnetbase.com/>), terpenoids comprise the largest, and structurally most diverse, family of natural products. As a consequence of their structural and stereochemical diversity, terpenoids exhibit various biological functions, including plant growth promotion, chemical communication and defence in living organisms. Furthermore, some terpenoids have great industrial value as pharmaceuticals, fragrances, fuels, flavourings and insecticides (Tholl, 2006; Gershenzon & Dudareva, 2007; Bohlmann & Keeling, 2008).

The tremendous structural diversity of terpenoids is a result of the cyclization of linear polyprenyl diphosphate substrates, such as geranyl diphosphate (GPP, C_{10}), farnesyl diphosphate (FPP, C_{15}) and geranylgeranyl diphosphate (GGPP, C_{20}). In the biosynthetic processes of terpenoids, terpene synthases play a key role in catalyzing linear isoprenoid precursors to form the structurally and stereochemically diverse ring skeletons of (poly)cyclic terpenoids. The terpene-cyclization reaction cascade is initiated by the formation of a reactive carbocation intermediate. The formation of this carbocation is through either ionization of polyprenyl diphosphate substrate by diphosphate abstraction or protonation of the terminal olefin (Christianson, 2006; Smanski *et al.*, 2012; Meguro *et al.*, 2013).

While terpene biosynthesis has been extensively studied in plants and fungi, unique terpenoids have now been identified in bacteria and they are likely to harbour a significant reservoir of new terpene synthases (Dairi, 2005). X-ray crystal structures of a number of terpenoid cyclases from fungi and plants are available (Lesburg *et al.*, 1997; Wendt *et al.*, 1999; Rynkiewicz *et al.*, 2001; Caruthers *et al.*, 2000; Shishova *et al.*, 2007; Christianson, 2008; Aaron *et al.*, 2010; Gao *et al.*, 2012). The diterpenoid taxol shows remarkable efficacy in cancer chemotherapy (Arbuck & Blaylock, 1995). In 2011, the crystal structure of a taxadiene synthase of plant origin was reported, which played an important role in the biosynthesis of taxol (Köksal *et al.*, 2010). However, no structures of bacterial diterpene synthases are known to date. In this paper, we report the expression, purification, crystallization and preliminary X-ray diffraction analysis of cyclooctatin synthase (CYC) from *Streptomyces* sp. LZ35.

© 2014 International Union of Crystallography
All rights reserved

2. Materials and methods

2.1. Molecular cloning

The *cyc* gene encoding diterpene synthase was PCR-amplified from *Streptomyces* sp. LZ35 genomic DNA using FastPfu DNA polymerase. The forward and reverse primers, with flanking *Eco*RI and *Hind*III sites, were 5'-TGTGAGGAATTCATGACGACCGG-ACCT-3' and 5'-TCCCCAAGCTTTCACCGGATGCGGGAG-3', respectively. The amplified segment was inserted into the pET-28a(+) vector (Novagen, Merck). The resulting plasmid pET-28a-CYC containing an N- and C-terminally His₆-tagged sequence facilitates the purification procedure. For the expression of selenomethionine (SeMet)-derivative protein, the *cyc* gene was alternatively constructed into the pET-22b(+) vector at *Nde*I and *Xho*I sites using the forward primer 5'-TGTGAGCATATGACGACCGGACCTT-CCACG-3' and the reverse primer 5'-TATATACTCGAGCCGG-ATGCGGGAGTTGAC-3', resulting in a C-terminally His₆-tagged protein. The correctness of the constructs was confirmed by sequencing. The recombinant plasmids were transformed into *Escherichia coli* strain BL21(DE3) cells for protein expression.

2.2. Protein expression and purification

E. coli BL21(DE3) cells transformed with the recombinant DNA were grown at 310 K in Luria broth (LB) medium supplemented with 100 µg ml⁻¹ kanamycin. The cells were grown to an OD₆₀₀ of 0.6 before induction with IPTG at a final concentration of 0.1 mM and were then grown at 310 K for a further 3 h. The cells were harvested by centrifugation (12 000g at 277 K for 30 min). The cells were resuspended in buffer (25 mM Tris-HCl, 200 mM NaCl) and lysed by sonication. The cell debris was removed following centrifugation (30 000g at 277 K for 45 min). The His-tagged protein was purified using affinity chromatography by binding the protein to an Ni-chelating Sepharose affinity column (GE Healthcare) pre-equilibrated with buffer A (25 mM Tris-HCl pH 8.0, 200 mM NaCl). After 1 h incubation, the column was washed with buffer B (25 mM Tris-HCl pH 8.0, 200 mM NaCl, 40 mM imidazole). The His-tagged protein was eluted with buffer C (25 mM Tris-HCl pH 8.0, 100 mM NaCl, 300 mM imidazole) and concentrated to a volume of 2.0 ml using an Amicon Ultra 3 kDa (Millipore) cutoff concentrator. Finally, the target protein was purified using a Superdex 200 size-exclusion column (GE Healthcare) with buffer D (10 mM Tris-HCl, 300 mM NaCl). Samples at each purification step were analysed by SDS-PAGE. Fractions containing the CYC protein were pooled according to the protein purity. SeMet-derivative CYC protein was

overexpressed by inhibiting endogenous methionine biosynthesis in M9 minimal medium supplemented with SeMet (Doublé, 1997) and was purified as described above. All protein purifications were performed at 277 K. For crystallization, the purified proteins were concentrated to approximately 10 mg ml⁻¹ in buffer E (10 mM Tris pH 8.0, 100 mM NaCl).

2.3. Crystallization and optimization

Initial sitting-drop vapour-diffusion crystallization trials were performed using commercial crystallization screening kits. For the initial screen, a 1:1 ratio of protein solution and commercial screen solutions (1 µl each) was used. Prism-like single crystals of CYC protein grew within a few days from a condition from the SaltRx 2 screen consisting of 1.0 M NaH₂PO₄, K₂HPO₄ pH 8.2. Attempts to optimize the crystallization condition by altering the pH range and salt concentration to obtain a large single crystal were successful. The best crystals grew within one month using a reservoir solution consisting of 0.8 M NaH₂PO₄/K₂HPO₄ pH 7.8 (Fig. 1). The SeMet-derivative protein was obtained by streak-seeding (Zhu *et al.*, 2005). The native protein crystals that were used as seeds were smashed into crystalline particles and used to make a stock in a stabilizing solution. A cat's whisker was used as a seeding wand and was dipped into the seeds. The whisker was drawn through pre-equilibrated SeMet-derivative protein solution drops, depositing the seeds in a streak line.

2.4. X-ray data collection and processing

In order to prevent radiation damage, the native and SeMet-derivative crystals were flash-cooled in a nitrogen stream at 100 K in the presence of reservoir buffer containing 15% (v/v) glycerol. X-ray diffraction data were collected using a Q315r CCD detector on beamline BL17U1 at the Shanghai Synchrotron Radiation Facility (SSRF), People's Republic of China. The native and SeMet-derivative crystals were rotated through 180° and 360°, respectively,

Figure 1

Crystal of the CYC protein. The crystal was grown in a reservoir solution consisting of 0.8 M NaH₂PO₄/K₂HPO₄ pH 7.8 using the hanging-drop vapour-diffusion method at 293 K.

Figure 2

Diffraction pattern of native CYC protein collected using a Q315r CCD detector on beamline BL17U1 at the SSRF. The data collected were processed to 1.75 Å resolution.

with 1° oscillation and 0.2 s exposure per frame. The data sets were processed using *HKL-2000* (Otwinowski & Minor, 1997). Data-collection and processing statistics are summarized in Table 1.

3. Results and discussion

The overexpression of the *cyc* gene was carried out in the *Streptomyces* sp. LZ35 Δ heng strain by conjugal transfer to afford the LZ35 Δ heng-*cyc* strain. Fermentation of LZ35 Δ heng-*cyc* led to the isolation of cyclooctatin (Kim *et al.*, 2009; Aoyama *et al.*, 1992). Moreover, the sequence of the *cyc* gene showed 97% identity to *cotB2* (Kim *et al.*, 2009). These results indicated that the CYC protein is a diterpene synthase which may have a similar function to CotB2.

The native crystal of CYC protein diffracted to 1.75 Å resolution (Fig. 2) and belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 59.10$, $b = 101.73$, $c = 108.93$ Å. The solvent content was about 47.99%, with a corresponding Matthews coefficient of $2.36 \text{ Å}^3 \text{ Da}^{-1}$ (Matthews, 1968) and two molecules per asymmetric unit. Searching the PDB with the CYC sequence retrieved no obvious homologous targets. The top hit was human apolipoprotein A-I (PDB entry 1av1; Borhani *et al.*, 1997) with 25% identity for 73 aligned residues. Moreover, the sequence identity to plant taxadiene synthase (PDB entry 3p5p; Köksal *et al.*, 2010) was also very low (17.5% with 307 aligned residues). The CYC protein thus could not be solved by the molecular-replacement (MR) method. We attempted to solve the structure of CYC by preparing selenomethionine-substituted protein and proceeded with crystallization. However, when we expressed the pET-28a-CYC construct using the SeMet-substitution protocol, the protein formed inclusion bodies owing to the extra methionines derived from the vector. An alternative construct was obtained by inserting *cyc* into the pET-22b(+) vector at *Nde*I and *Xho*I restriction sites, which resulted in the successful expression of SeMet-CYC protein with a C-terminal His₆-tag (LEHHHHHH). The SeMet-CYC protein was crystallized at 293 K by the hanging-drop vapour-diffusion method, using the same reservoir solution as used for the native crystals. However, the initial crystals of the SeMet-CYC protein appeared as a cluster of crystals that were not suitable for X-ray diffraction. For optimization, the obtained native crystals were used as seeds to optimize the crystals by a streak-seeding method

Figure 3
Initial electron-density map of the CYC model contoured at 1.6σ generated by *RESOLVE*.

Table 1

X-ray diffraction data-collection and processing statistics.

Values in parentheses are for the outermost shell.

	Native	SeMet derivative
Wavelength (Å)	0.9790	0.9789
Space group	$P2_12_12_1$	$P2_1$
Molecules in asymmetric unit (<i>N</i>)	2	2
Unit-cell parameters (Å, °)	$a = 59.10$, $b = 101.73$, $c = 108.93$	$a = 58.64$, $b = 109.47$, $c = 58.73$, $\beta = 119.35^\circ$
Resolution (Å)	50–1.75 (1.81–1.75)	50–2.60 (2.69–2.60)
Total No. of reflections	67227 (6608)	19811 (1931)
No. of unique reflections	9337 (957)	2677 (263)
Completeness (%)	99.9 (99.5)	100 (100)
Multiplicity	7.2 (6.9)	7.4 (7.5)
$\langle I/\sigma(I) \rangle$	33.6 (5.2)	35.6 (9.1)
$R_{\text{merge}}^\dagger$ (%)	7.4 (42)	19.1 (47)

$^\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ over *i* observations.

(Bergfors, 2003; Zhu *et al.*, 2005). Finally, prism-like crystals of SeMet-CYC protein were obtained that diffracted to 2.6 Å resolution and belonged to space group $P2_1$, with unit-cell parameters $a = 58.64$, $b = 109.47$, $c = 58.73$ Å, $\beta = 119.35^\circ$. The solvent content was about 48.49%, with a corresponding Matthews coefficient of $2.39 \text{ Å}^3 \text{ Da}^{-1}$ and two molecules per asymmetric unit. The structure of the CYC protein has been solved using the single-wavelength anomalous diffraction (SAD) method (Terwilliger, 2003). 17 heavy atoms (Se) were found using *SOLVE* and the initial electron-density map generated by *RESOLVE* clearly shows the secondary structure (α -helix) of the protein (Fig. 3). The solved structure will be published elsewhere.

We are grateful to the staff of beamline BL17U1 at the Shanghai Synchrotron Radiation Facility for support during data collection. This project was supported by the 973 Programs (2010CB833802 and 2012CB721005).

References

- Aaron, J. A., Lin, X., Cane, D. E. & Christianson, D. W. (2010). *Biochemistry*, **49**, 1787–1797.
- Aoyama, T., Naganawa, H., Muraoka, Y., Aoyagi, T. & Takeuchi, T. (1992). *J. Antibiot.* **45**, 1703–1704.
- Arbuck, S. & Blaylock, B. (1995). *Taxol: Science and Applications*, edited by M. Suffness, pp. 379–416. Boca Raton: CRC Press.
- Bergfors, T. (2003). *J. Struct. Biol.* **142**, 66–76.
- Bohlmann, J. & Keeling, C. I. (2008). *Plant J.* **54**, 656–669.
- Borhani, D. W., Rogers, D. P., Engler, J. A. & Brouillette, C. G. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 12291–12296.
- Caruthers, J. M., Kang, I., Rynkiewicz, M. J., Cane, D. E. & Christianson, D. W. (2000). *J. Biol. Chem.* **275**, 25533–25539.
- Christianson, D. W. (2006). *Chem. Rev.* **106**, 3412–3442.
- Christianson, D. W. (2008). *Curr. Opin. Chem. Biol.* **12**, 141–150.
- Dairi, T. (2005). *J. Antibiot.* **58**, 227–243.
- Doublé, S. (1997). *Methods Enzymol.* **276**, 523–530.
- Gao, Y., Honzatko, R. B. & Peters, R. J. (2012). *Nat. Prod. Rep.* **29**, 1153–1175.
- Gershenzon, J. & Dudareva, N. (2007). *Nature Chem. Biol.* **3**, 408–414.
- Kim, S.-Y., Zhao, P., Igarashi, M., Sawa, R., Tomita, T., Nishiyama, M. & Kuzuyama, T. (2009). *Chem. Biol.* **16**, 736–743.
- Köksal, M., Jin, Y., Coates, R. M., Croteau, R. & Christianson, D. W. (2010). *Nature (London)*, **469**, 116–120.
- Lesburg, C. A., Zhai, G., Cane, D. E. & Christianson, D. W. (1997). *Science*, **277**, 1820–1824.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Meguro, A., Tomita, T., Nishiyama, M. & Kuzuyama, T. (2013). *ChemBiochem*, **14**, 316–321.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Rynkiewicz, M. J., Cane, D. E. & Christianson, D. W. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 13543–13548.

- Shishova, E. Y., Di Costanzo, L., Cane, D. E. & Christianson, D. W. (2007). *Biochemistry*, **46**, 1941–1951.
- Smanski, M. J., Peterson, R. M., Huang, S.-X. & Shen, B. (2012). *Curr. Opin. Chem. Biol.* **16**, 132–141.
- Terwilliger, T. C. (2003). *Methods Enzymol.* **374**, 22–37.
- Tholl, D. (2006). *Curr. Opin. Plant Biol.* **9**, 297–304.
- Wendt, K. U., Lenhart, A. & Schulz, G. E. (1999). *J. Mol. Biol.* **286**, 175–187.
- Zhu, D.-Y., Zhu, Y.-Q., Xiang, Y. & Wang, D.-C. (2005). *Acta Cryst.* **D61**, 772–775.