

**Liping Gong,<sup>a,b</sup> Yi Xiao,<sup>c</sup> Qiang Liu,<sup>a</sup> Sumei Li,<sup>c</sup> Changsheng Zhang<sup>c\*</sup> and Jinsong Liu<sup>a,b\*</sup>**

<sup>a</sup>National Key Laboratory of Respiratory Disease, Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, 190 Kaiyuan Avenue, Guangzhou Science Park, Guangzhou 510530, People's Republic of China, <sup>b</sup>University of Science and Technology of China, Hefei, People's Republic of China, and <sup>c</sup>CAS Key Laboratory of Marine Bio-resources Sustainable Utilization, RNAM Center for Marine Microbiology, Guangdong Key Laboratory of Marine Materia Medica, South China Sea Institute of Oceanology, Chinese Academy of Sciences, 164 West Xingang Road, Guangzhou 510301, People's Republic of China

Correspondence e-mail: czhang@scsio.ac.cn, liu\_jinsong@gibh.ac.cn

Received 14 October 2010  
 Accepted 19 November 2010

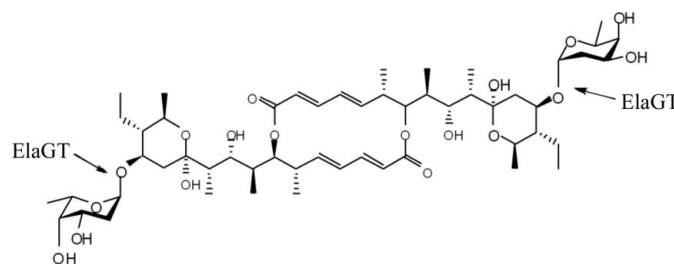
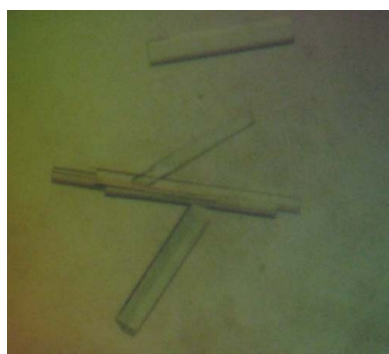
## Preliminary X-ray crystallographic analysis of the glycosyltransferase from a marine *Streptomyces* species

ElaGT is a glycosyltransferase from a marine *Streptomyces* species that is involved in the biosynthesis of elaiophylin. Here, the molecular cloning, protein expression and purification, preliminary crystallization and crystallographic characterization of ElaGT are reported. The rod-shaped crystals belonged to space group  $P2_122$ , with unit-cell parameters  $a = 66.7$ ,  $b = 131.7$ ,  $c = 224.6$  Å,  $\alpha = 90$ ,  $\beta = 90$ ,  $\gamma = 90^\circ$ . Data were collected to 2.9 Å resolution. A preliminary molecular-replacement solution implied the presence of two ElaGT molecules in the asymmetric unit.

### 1. Introduction

The polyketide elaiophylin contains an unusual 16-membered ring system comprising two identical 2-deoxy-L-fucoses with  $C_2$  symmetry (Fig. 1). This natural product has been isolated from many *Streptomyces* species (Fang *et al.*, 2000; Fiedler *et al.*, 1981; Haltli, 2006; Hammann *et al.*, 1990; Haydock *et al.*, 2004) and has also recently been found in the marine *Streptomyces* sp. SCSIO 01934 identified in sediment from the South China Sea (unpublished data). Elaiophylin exhibits broad antibacterial and antifungal activity (Fiedler *et al.*, 1981; Hammann *et al.*, 1990) and, as has often been found for other glycosidic antibiotics (Salas & Mendez, 2007; Thibodeaux *et al.*, 2008; Williams *et al.*, 2008), the deoxy-sugar moieties play an essential role in modulating its bioactivity (Evans & Fitch, 1997).

The attachment of sugar moieties to natural products is generally catalyzed by glycosyltransferases, which have proven to be potential enzymatic tools for enhancing the diversity and activity of natural products (Thibodeaux *et al.*, 2008; Williams *et al.*, 2008; Zhang, Griffith *et al.*, 2006). Elucidation of the elaiophylin-biosynthetic gene cluster identified a single glycosyltransferase ElaGT (Haydock *et al.*, 2004; Haltli, 2006) which is putatively responsible for the sequential attachment of the two deoxy sugars. Iteratively acting glycosyltransferases have been demonstrated to be involved in the biosynthesis of landomycin and avermectin (Luzhetskyy *et al.*, 2005; Zhang, Albermann *et al.*, 2006). However, the catalytic mechanism of this class of glycosyltransferases remains unresolved. Structural elucidation of glycosyltransferases has facilitated understanding of their mechanisms of action and substrate specificities (Roychoudhury & Pohl, 2010) and is also important for the rational engineering of glycosyltransferases for glycodiversification (Williams *et al.*, 2008).



**Figure 1**  
 Chemical structure of elaiophylin. Two deoxy sugars are putatively biosynthesized by a single glycosyltransferase ElaGT.

In order to understand the catalytic mechanism of ElaGT and lay a foundation for further rational protein engineering, we have carried out a structural analysis of ElaGT. Here, we report the crystallization and preliminary X-ray diffraction data of ElaGT.

## 2. Materials and methods

### 2.1. Molecular cloning

A 1.7 kb DNA fragment containing the elaiophylin glycosyltransferase gene *elaGT* was amplified by PCR using a pair of degenerate primers TEIIa (5'-TCCCGGCGGBCACTTCTWCS-3') and EPIs (5'-AGCGCGACGAABSCGTGBCC-3') designed to match conserved regions based on published sequence information (Haltli, 2006; Haydock *et al.*, 2004). The primers were synthesized by Sangong Inc. (Shanghai, China) and the DNA sequences were determined by Invitrogen Technologies (Guangzhou, China). The PCR product was cloned, sequenced and deposited in GenBank with Accession No. HQ318716.

The 1254 bp *elaGT* gene was PCR-amplified from genomic DNA of *Streptomyces* sp. SCSIO 01934 with Pyrobest high-fidelity DNA polymerase (TaKaRa) using the primers *ela13s* (forward, 5'-AACCGCATATGCGTATCCTTTTCG-3'; *NdeI*) and *ela13a* (reverse, 5'-GTAGGATCCGGGCACCGCCAGC-3'; *BamHI*). The PCR product was cloned into pET28a to form pCSG100 and the insert was confirmed by sequencing.

### 2.2. Protein expression and purification

*Escherichia coli* BL21 (DE3) cells containing the plasmid pCSG100 were grown at 310 K in Luria–Bertani medium containing 50 µg ml<sup>-1</sup> kanamycin until the OD<sub>600nm</sub> reached 0.6–0.8. The cultures were induced with 0.2 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown for a further 16 h at 293 K. The cells were pelleted and resuspended in buffer A (50 mM Tris pH 8.0, 150 mM NaCl, 50 mM imidazole). The cells were lysed by sonication and then centrifuged at 12 000 rev min<sup>-1</sup> for 30 min. The soluble fraction was loaded onto an Ni–NTA column (Qiagen). After washing with lysis buffer (~10 column volumes), the target protein was eluted with buffer B (50 mM Tris pH 8.0, 150 mM NaCl, 500 mM imidazole). The eluate was loaded onto a Source-Q anion-exchange column pre-equilibrated with buffer C (50 mM Tris–HCl pH 8.0, 180 mM NaCl). The protein was eluted by application of a linear NaCl gradient (100–1000 mM). Fractions containing the protein were pooled and, after concentration, were loaded onto a Superdex 200 16/60 column (GE Healthcare) pre-equilibrated with 10 mM Tris–HCl pH 7.3, 60 mM NaCl. The elution volume for the target protein was 81.6 ml, corresponding to a molecular weight of around 51.3 kDa. This was close to the theoretical molecular weight of ElaGT (48.4 kDa) and indicated that ElaGT eluted as a monomer. The peak fraction containing the target protein was concentrated to 6 mg ml<sup>-1</sup> and incubated with the co-factor UDP at a final concentration of 2 mM overnight at 295 K for crystallization trials.

### 2.3. Crystallization

Initial crystallization experiments were set up using the vapour-diffusion method in CrystalQuick 96-well sitting-drop plates (Greiner) using Index, Crystal Screen, Crystal Screen 2 and PEG/Ion kits from Hampton Research. The sitting drops were made up of 1 µl protein solution and 1 µl well solution and the mixture was equilibrated over 60 µl well solution. All crystallization trials were performed at 295 K. Small needle-shaped crystals appeared in a few

**Table 1**

Crystal parameters and data-collection statistics for the crystal of ElaGT.

Values in parentheses are for the highest resolution shell.

Space group	P2 <sub>1</sub> 22
Unit-cell parameters (Å, °)	<i>a</i> = 66.7, <i>b</i> = 131.7, <i>c</i> = 224.6, α = β = γ = 90
Resolution (Å)	46.87–2.90 (3.06–2.90)
Unique reflections	44849 (6416)
Multiplicity	4.1 (4.1)
Average <i>I</i> /σ( <i>I</i> )	9.7 (2.7)
Completeness (%)	100 (100)
<i>R</i> <sub>merge</sub> † (%)	11.2 (48.7)
<i>R</i> <sub>r.i.m.</sub> ‡ (%)	13.0 (55.4)
<i>R</i> <sub>p.i.m.</sub> § (%)	6.4 (27.0)

†  $R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}$ , where  $I_i(hkl)$  is the *i*th observation of reflection *hkl* and  $\langle I(hkl) \rangle$  is the weighted average intensity of all observations *i* of reflection *hkl*. ‡ *R*<sub>r.i.m.</sub> is the redundancy-independent (multiplicity-weighted) *R*<sub>merge</sub>. § *R*<sub>p.i.m.</sub> is the precision-indicating (multiplicity-weighted) *R*<sub>merge</sub>.

days. Optimization of the initial condition was carried out by changing the buffer pH and the precipitant concentration. Additive screening was carried out with additive added to both the drop and the well solution and the effectiveness of the additive was assessed from the crystal size and diffraction image. After several rounds of refinement, well diffracting single crystals were obtained. The final condition was 22% PEG 3350, 0.15 M DL-malic acid pH 7.2, 24% propylene glycol.

### 2.4. Data collection and processing

Before data collection, crystals were soaked in reservoir solution with an additional 6% propylene glycol to give a final propylene glycol concentration of 30% as the cryoprotectant. The crystal was then flash-cooled in an N<sub>2</sub> stream and used for X-ray diffraction data collection at 100 K using synchrotron radiation on beamline BL17U1 at the Shanghai Synchrotron Radiation Facility.

A complete diffraction data set consisting of 200 images was collected from one crystal with an oscillation angle of 0.5° per image and a crystal-to-detector distance of 240 mm. The data were indexed and integrated using *MOSFLM* (Leslie, 1992) and scaled using the program *SCALA* from the *CCP4* package (Collaborative Computational Project, Number 4, 1994). Data-collection and processing statistics are listed in Table 1.

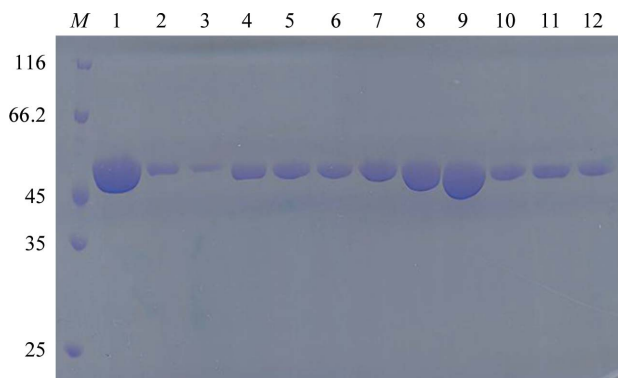
## 3. Results and discussion

The elaiophylin-biosynthetic gene cluster has been identified in *Streptomyces* sp. DSM4137 (Haydock *et al.*, 2004) and *Streptomyces* sp. NRRL 30748 (Haltli, 2006). These two independently characterized gene clusters from two different strains showed identical organization patterns, in which the putative glycosyltransferase-encoding gene was flanked by two genes encoding a type II thioesterase and a dTDP-4-keto-6-deoxyglucose 3,5-epimerase. Since sequence information was unavailable for the elaiophylin-biosynthetic locus of *Streptomyces* sp. SCSIO 01934, we designed a pair of degenerate primers TEIIa (targeting a conserved region of type II thioesterases) and EPIs (targeting a conserved region of dTDP-4-keto-6-deoxyglucose 3,5-epimerases) to successfully amplify a 1.7 kb DNA fragment by PCR from the genomic DNA of *Streptomyces* sp. SCSIO 01934. Cloning and sequencing of the 1.7 kb DNA fragment (GenBank accession No. HQ318716) revealed a glycosyltransferase-encoding gene designated *elaGT*, the nucleotide sequence of which displayed 95% identity to the previously reported nucleotide sequence from *Streptomyces* sp. NRRL 30748 (Haltli, 2006). Subse-

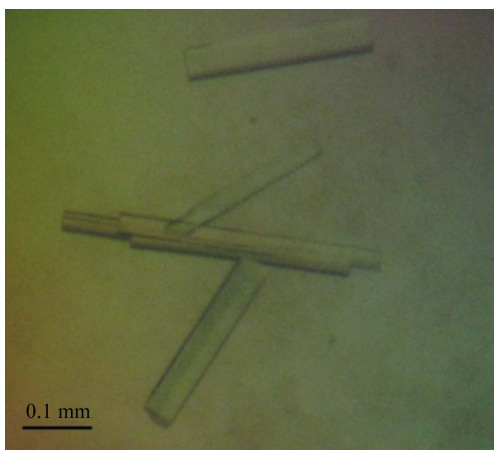
quently, the 1254 bp *elaGT* was PCR-amplified from the genomic DNA of *Streptomyces* sp. SCSIO 01934 and cloned into pET28, resulting in the expression plasmid pCSG100.

ElaGT was successfully overexpressed in *E. coli* BL21 (DE3) containing pCSG100 in a soluble form and purified to homogeneity. Gel-filtration analysis showed that protein was a monomer in solution. The purity was checked by SDS-PAGE (Fig. 2).

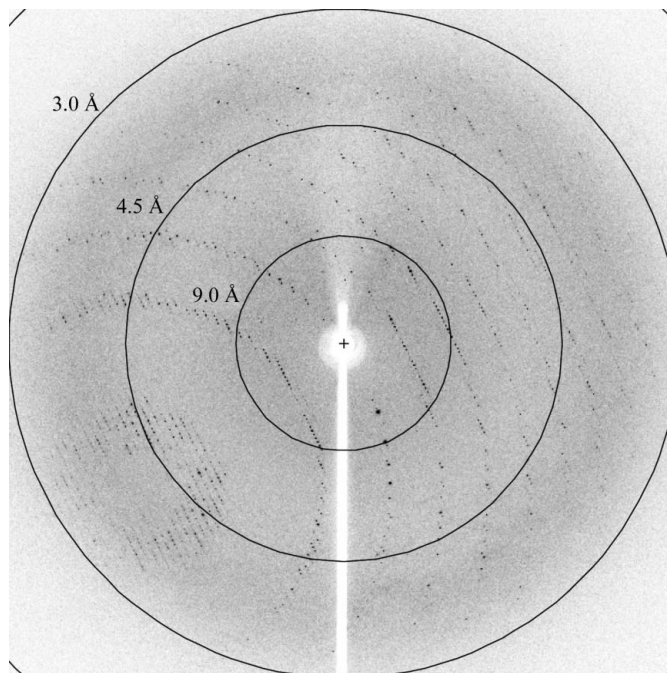
A shower of small needle-shaped crystals appeared after 2–6 d in a crystallization condition consisting of 20% PEG 3350, 0.15 M DL-malic acid pH 7.0 (Index condition No. 91). This initial condition was further optimized at the same temperature using the same type of plates. Approximately 1000 droplets were set up and propylene glycol was identified as the best additive. The final condition for diffraction-quality crystals (shown in Fig. 3) was 22% PEG 3350, 0.15 M DL-malic acid pH 7.2, 24% propylene glycol. The crystal diffracted to 2.9 Å resolution on beamline BL17U1 at Shanghai Synchrotron Radiation Facility. A diffraction image is shown in Fig. 4 and the statistics of data collection and processing are summarized in Table 1. The crystal belonged to the orthorhombic space group  $P2_122$ . The systematic absences were consistent with the presence of one twofold screw axis. No indication of twinning was observed. A cell-content analysis suggested that there were possibly two or four molecules in the asymmetric unit, with a corresponding calculated Matthews coefficient ( $V_M$ ) of 5.34 or 2.67 Å<sup>3</sup> Da<sup>-1</sup> and a solvent content of 77 or 54% (Matthews, 1968).



**Figure 2**  
SDS-PAGE analysis of purified ElaGT. Lane M, markers (labelled in kDa); lanes 1–12, eluted fractions from a HiLoad 16/60 Superdex 200 size-exclusion column containing ElaGT.



**Figure 3**  
Crystals of ElaGT.



**Figure 4**  
An X-ray diffraction frame collected from a crystal of ElaGT.

The glycosyltransferase UrdGT2 from *S. fradiae* (PDB code 2p6p; Mittler *et al.*, 2007), which has 26% sequence identity to ElaGT, was used to obtain a preliminary molecular-replacement solution using *BALBES* (Long *et al.*, 2008). The *R* factor and free *R* factor after one round of refinement using the molecular-replacement solution were 46.6% and 50.9%, respectively. This initial molecular-replacement solution suggested the presence of two ElaGT molecules per asymmetric unit. Further model improvement and refinement is ongoing. The crystal structure of ElaGT will ideally provide structural insight into the catalytic mechanism of ElaGT and be helpful for future protein-engineering work with this enzyme.

This work was supported by grants from the One Hundred Person Project of the Chinese Academy of Sciences and the Knowledge Innovation Program of the Chinese Academy of Sciences (KSCX2-YW-R-084, KSCX2-YW-G-065 and KZCX2-YW-216). We thank the staff members at beamline BL17U1 at the Shanghai Synchrotron Radiation Facility (SSRF) of China for technical support in diffraction data collection.

## References

- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Evans, D. A. & Fitch, D. M. (1997). *J. Org. Chem.* **62**, 454–455.
- Fang, A., Wong, G. K. & Demain, A. L. (2000). *J. Antibiot.* **53**, 158–162.
- Fiedler, H. P., Worner, W., Zahner, H., Kaiser, H. P., Keller-Schierlein, W. & Muller, A. (1981). *J. Antibiot.* **34**, 1107–1118.
- Haltli, B. (2006). US Patent 20060141583.
- Hammann, P., Kretschmar, G. & Seibert, G. (1990). *J. Antibiot.* **43**, 1431–1440.
- Haydock, S. F., Mironenko, T., Ghoorahoo, H. I. & Leadlay, P. F. (2004). *J. Biotechnol.* **113**, 55–68.
- Leslie, A. G. W. (1992). *Jnt CCP4/ESF-EACBM Newsl. Protein Crystallogr.* **26**.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* **D64**, 125–132.
- Luzhetskyy, A., Fedoryshyn, M., Durr, C., Taguchi, T., Novikov, V. & Bechthold, A. (2005). *Chem. Biol.* **12**, 725–729.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

- Mittler, M., Bechthold, A. & Schulz, G. E. (2007). *J. Mol. Biol.* **372**, 67–76.
- Roychoudhury, R. & Pohl, N. L. (2010). *Curr. Opin. Chem. Biol.* **14**, 168–173.
- Salas, J. A. & Mendez, C. (2007). *Trends Microbiol.* **15**, 219–232.
- Thibodeaux, C. J., Melançon, C. E. III & Liu, H.-W. (2008). *Angew. Chem. Int. Ed. Engl.* **47**, 9814–9859.
- Williams, G. J., Gantt, R. W. & Thorson, J. S. (2008). *Curr. Opin. Chem. Biol.* **12**, 556–564.
- Zhang, C., Albermann, C., Fu, X. & Thorson, J. S. (2006). *J. Am. Chem. Soc.* **128**, 16420–16421.
- Zhang, C., Griffith, B. R., Fu, Q., Albermann, C., Fu, X., Lee, I.-K., Li, L. & Thorson, J. S. (2006). *Science*, **313**, 1291–1294.