Crystallization and preliminary X-ray diffraction analysis of a putative carbon–carbon bond hydrolase from *Mycobacterium abscessus* 103

Zhang Zhang,^a Yong-Liang Jiang,^b Yi Wu^a and Yong-Xing He^{a*}

Received 7 January 2015
Accepted 24 January 2015

^aMOE Key Laboratory of Cell Activities and Stress Adaptations, School of Life Sciences, Lanzhou University, Lanzhou, Gansu 730000, People's Republic of China, and ^bHefei National Laboratory for Physical Sciences at the Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, People's Republic of China. *Correspondence e-mail: heyx@lzu.edu.cn

Keywords: PhlG; carbon–carbon bond hydrolase; *Mycobacterium abscessus* 103.

The PhlG protein from *Mycobacterium abscessus* 103 (mPhlG), which shares 30% sequence identity with phloretin hydrolase from *Eubacterium ramulus* and 38% sequence identity with 2,4-diacetylphloroglucinol hydrolase from *Pseudomonas fluorescens* Pf-5, is a putative carbon–carbon bond hydrolase. Here, the expression, purification and crystallization of mPhlG are reported. Crystals were obtained using a precipitant consisting of 100 mM citric acid pH 5.0, 1.0 M lithium chloride, 8% (w/v) polyethylene glycol 6000. The crystals diffracted to 1.87 Å resolution and belonged to space group $P2_1$, with unit-cell parameters $a = 71.0$, $b = 63.4$, $c = 74.7$ Å, $\alpha = 90.0$, $\beta = 103.2$, $\gamma = 90.0^\circ$. Assuming the presence of two mPhlG molecules in the asymmetric unit, V_M was calculated to be $2.5 \text{ Å}^3 \text{ Da}^{-1}$, which corresponds to a solvent content of 50%.

1. Introduction

Enzymes that cleave carbon–carbon bonds have great potential as industrial biocatalysts for the production of high-value chemical building blocks owing to their diverse range of substrates and their catalytic independence of cofactors (Siirola *et al.*, 2013). To date, 22 carbon–carbon bond hydrolases (EC 3.7.1.1–3.7.1.22) have been described in the KEGG enzyme database (Kanehisa & Goto, 2000). Among these, the phloretin hydrolase Phy (EC 3.7.1.4) from *Eubacterium ramulus* and the 2,4-diacetylphloroglucinol hydrolase PhlG from *Pseudomonas fluorescens* are so-called retro Friedel–Crafts hydrolases which cleave the carbon–carbon bond between an aryl moiety and a ketone moiety (Schoefer *et al.*, 2004; Bottiglieri & Keel, 2006). Phy and PhlG share moderate sequence homology, but display very distinct and strict substrate specificity. While PhlG hydrolyzes diacetylphloroglucinol (DAPG) by cleaving the acetyl group from the monoacetylphloroglucinol (MAPG) ring, Phy degrades phloretin by removing a much larger 3-(4-hydroxyphenyl) propionic acid moiety from the phloroglucinol ring (Schoefer *et al.*, 2004; Bottiglieri & Keel, 2006). The structure of PhlG from *P. fluorescens* Pf-5 has been solved, revealing a Bet v 1-like fold that is different from that of the classical α/β -fold hydrolases (He *et al.*, 2010). However, the structure determination of Phy was hampered by the poor diffraction quality of the Phy crystals and thus no structural explanation of its distinct substrate specificity was provided (Frank *et al.*, 2014).

Genomic context analysis has been shown to be a valuable tool to predict protein function in microorganisms since the

Table 1
Production of mPhlG.

Source organism	<i>M. abscessus</i> 103
5' sequence	NdeI site: CATATG
3' sequence	NotI site: GCGGCCGC
Cloning vector	pUC57
Expression vector	pET-28b-derived vector
Expression host	<i>E. coli</i> strain BL21 (DE3)
Complete amino-acid sequence of the construct produced	MGHHHHHMQAKTRRHAVHPITYYPVDTQRLVR-SNAERIRHKPYAHYFNPDAVPEEVFAALKAP-LEPEQVLGTSSSTELNRLLEPGYLEGETGYCGL-PDGAGYTSSLVRFPATPEMFRWWFWHSEFEP-ERYSLWHPWCHADIWRDPEPETETAPNLTDEQR-YVGSTHHINEYIGQDPLDIEITFIDPARWGFD-ADGFAAGIGAHACGSVLMKGSMLATMVHL-ARITDDGFELRSRYWIADRAEPRHDPVAGIAQ-LTTVPFGFSGERQAYEQLVHDQTEFNHLATFLP-DIYQEFQPR

genes involved in the same metabolic pathways tend to cluster together in the genome (Huynen *et al.*, 2000). The gene encoding PhlG is located within a genomic context that is involved in the biosynthesis of DAPG, an important antibiotic produced by certain types of *P. fluorescens*. PhlG is supposed to prevent the self-toxicity of DAPG by degrading it to the less toxic MAPG. Interestingly, through a search for PhlG homologues in currently sequenced bacterial genomes, we found that PhlG homologues exist across diverse bacterial species, a large number of which do not possess DAPG-biosynthetic gene clusters. It is likely that different substrate specificities may have evolved in different branches of PhlG homologues with very different genomic contexts (such as the phloretin hydrolase from *E. ramulus*). To gain further insight into the structural basis of the substrate specificity of the PhlG protein family, a structural investigation was initiated on a PhlG homologue from *Mycobacterium abscessus* 103 (mPhlG), which displays 30 and 38% sequence identity to phloretin hydrolase from *E. ramulus* and PhlG from *P. fluorescens* Pf-5, respectively. Here, we report the overexpression, purification and crystallization, as well as the preliminary X-ray diffraction analysis, of mPhlG.

2. Materials and methods

2.1. Expression and purification of His₆-PhlG

The codon-optimized mPhlG gene for expression in *Escherichia coli* was synthesized (Genewiz) and cloned into a pET-28b-derived vector with an N-terminal 6×His tag. The sequence was confirmed by automated DNA sequencing. The recombinant plasmid was transformed into *E. coli* strain BL21 (DE3) (Novagen). The cells were grown to an OD_{600 nm} of 0.6–0.8 at 37°C and expression of the recombinant protein was induced by 0.2 mM isopropyl β-D-1-thiogalactopyranoside for 20 h at 16°C. Cells were collected by centrifugation and were resuspended in 20 mM Tris–HCl pH 8.0, 100 mM NaCl, 10 mM imidazole. After 30 min of sonication and centrifugation at 12 000g for 25 min, the supernatant containing the soluble target protein was collected and loaded onto Ni²⁺–

nitrilotriacetate affinity resin (Ni-NTA, Qiagen) equilibrated with 20 mM Tris–HCl pH 8.0, 100 mM NaCl. The target protein was eluted with 20 mM Tris–HCl pH 8.0, 100 mM NaCl, 300 mM imidazole and further purified by gel filtration (HiLoad 16/60 Superdex 75, GE Healthcare) equilibrated with 20 mM Tris–HCl pH 8.0, 100 mM NaCl. The peak fractions were collected and concentrated to 20 mg ml^{−1} for crystallization. The purity of the protein was evaluated by SDS–PAGE and the protein sample was stored at −80°C. The details of mPhlG production are summarized in Table 1.

2.2. Crystallization

The mPhlG crystals were grown by the hanging-drop vapour-diffusion method at 14°C, with the initial condition consisting of mixing 1 µl 10 mg ml^{−1} protein sample with an equal volume of reservoir solution consisting of 100 mM citric acid pH 5.0, 1.0 M lithium chloride, 8% (w/v) polyethylene glycol 6000. Crystals appeared in 1 d and grew to full size within 2 d.

Figure 1
(a) Gel filtration of mPhlG using a HiLoad 16/60 Superdex 75 column. The flow rate was 1 ml min^{−1}. (b) 15% SDS–PAGE analysis of mPhlG. Lane 1, purified recombinant mPhlG; Lane M, low-molecular-weight marker (labelled in kDa)

2.3. Data collection and processing

The crystals were soaked in cryoprotectant (reservoir solution supplemented with 25% glycerol) and flash-cooled in liquid nitrogen. X-ray data were collected at 100 K in a liquid-nitrogen gas stream on beamline BL17U1 at Shanghai Synchrotron Radiation Facility (SSRF) using a Q315r CCD detector (ADSC). All diffraction data were indexed, integrated and scaled with *HKL-2000* (Otwinowski & Minor, 1997).

Figure 2
A diffraction-quality crystal of mPhlG obtained using the hanging-drop vapour-diffusion method with dimensions of $0.35 \times 0.05 \times 0.01$ mm.

Table 2

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL17U1, SSRF
Wavelength (Å)	0.97915
Temperature (K)	100
Detector	ADSC Q315r CCD
Crystal-to-detector distance (mm)	250
Rotation range per image (°)	1
Total rotation range (°)	360
Exposure time per image (s)	1
Space group	$P2_1$
Unit-cell parameters (Å, °)	$a = 71.0, b = 63.4, c = 74.7$ $\alpha = 90.0, \beta = 103.2, \gamma = 90.0$
Mosaicity (°)	0.3
Resolution range (Å)	50.00–1.87 (1.94–1.87)
Total No. of reflections	161357 (15744)
No. of unique reflections	52356 (5248)
Completeness (%)	97.8 (99.0)
Multiplicity	3.1 (3.0)
Mean $I/\sigma(I)$	13.8 (2.6)
$R_{\text{r.i.m.}}^\dagger$	0.046 (0.327)
Overall B factor from Wilson plot (Å ²)	24.6

[†] The redundancy-independent merging R factor $R_{\text{r.i.m.}}$ was estimated by multiplying the linear R_{merge} value by the factor $[N/(N-1)]^{1/2}$, where N is the data multiplicity

3. Results and discussion

The mPhlG protein was expressed in *E. coli* strain BL21 (DE3) with a N-terminal 6×His tag to facilitate affinity purification of the protein. The gel-filtration method was used as the final purification step and the purity of the target protein was further checked by SDS-PAGE (Fig. 1). Crystals of mPhlG were obtained using a precipitant consisting of 100 mM citric acid pH 5.0, 1.0 M lithium chloride, 8%(w/v) polyethylene glycol 6000 (Fig. 2).

Figure 3
X-ray diffraction pattern of the mPhlG crystals. Black circles indicate resolution rings at 2.00 and 1.80 Å resolution (inner and outer, respectively). The inset represents a magnified view of the section indicated by the corresponding square.

A complete diffraction data set was collected to 1.87 Å resolution and the data-collection statistics are listed in Table 2. A typical diffraction image for mPhlG is shown in Fig. 3. By using *POINTLESS* (Evans, 2006) in the *CCP4* suite (Winn *et al.*, 2011), the space group was determined to be $P2_1$, with unit-cell parameters $a = 71.0$, $b = 63.4$, $c = 74.7$ Å, $\alpha = 90.0$, $\beta = 103.2$, $\gamma = 90.0^\circ$. Determination of the Matthews coefficient (Matthews, 1968) suggested 50% solvent content ($V_M = 2.5$ Å³ Da⁻¹), which corresponds to the presence of two molecules per asymmetric unit. The structure solution was found by the molecular-replacement method with *MOLREP* (Vagin & Teplyakov, 2010), using the *P. fluorescens* PhlG structure (PDB entry 3hwp, sequence identity 38%; He *et al.*, 2010) as the search model. Structure refinement is in progress.

Acknowledgements

We thank the staff of beamline BL17U1 at the Shanghai Synchrotron Radiation Facility for assistance during the data-collection process. This work was supported by grants from the National Natural Science Foundation of China (grant No. 31300616) and Lanzhou University (grant Nos. lzujbky-2014-

85 and lzujbky-2013-bt05). The funders had no role in the study design, data collection and analysis, the decision to publish or the preparation of the manuscript.

References

- Bottiglieri, M. & Keel, C. (2006). *Appl. Environ. Microbiol.* **72**, 418–427.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Frank, A., Siirola, E., Kroutil, W. & Grogan, G. (2014). *Top. Catal.* **57**, 376–384.
- He, Y.-X., Huang, L., Xue, Y., Fei, X., Teng, Y.-B., Rubin-Pitel, S. B., Zhao, H. & Zhou, C.-Z. (2010). *J. Biol. Chem.* **285**, 4603–4611.
- Huynen, M., Snel, B., Lathe, W. & Bork, P. (2000). *Genome Res.* **10**, 1204–1210.
- Kanehisa, M. & Goto, S. (2000). *Nucleic Acids Res.* **28**, 27–30.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Schoefer, L., Braune, A. & Blaut, M. (2004). *Appl. Environ. Microbiol.* **70**, 6131–6137.
- Siirola, E., Frank, A., Grogan, G. & Kroutil, W. (2013). *Adv. Synth. Catal.* **355**, 1677–1691.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.