Purification and crystallographic analysis of a FAD-dependent halogenase from *Streptomyces* sp. JCM9888

Yanqun Zhao,^{a,b,†} Baohua Yan,^{a,b,†} Ting Yang,^{a,b} Jian Jiang,^{a,b} Heng Wei^{a,b} and Xiaofeng Zhu^{a,b,*}

Received 28 April 2015

Accepted 22 May 2015

Edited by R. A. Pauptit, Macclesfield, England

† These authors contributed equally to this work.

Keywords: FAD-dependent halogenase; regioselective halogenation; adenine.

^aCollege of Life Science, Key Laboratory of Bio-Resources and Eco-Environment of the Ministry of Education, Center for Growth, Metabolism and Aging, Sichuan University, Chengdu 610064, People's Republic of China, ^bState Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, People's Republic of China, and ^cCollaborative Innovation Centre for Biotherapy, People's Republic of China. *Correspondence e-mail: zhuxiaofeng@scu.edu.cn

A new FAD (flavin adenine dinucleotide)-dependent halogenase HalY from *Streptomyces* sp. JCM9888 was reported to be involved in the regioselective halogenation of adenine. HalY is a variant B FAD-dependent halogenase that is most similar to the halogenase PltA involved in pyoluteorin biosynthesis. This study reports the overexpression and purification of HalY with an N-terminal hexahistidine tag, followed by crystallization experiments and X-ray crystallographic analysis. HalY was purified as a monomer in solution and crystallized to give X-ray diffraction to a resolution of 1.7 Å. The crystal belonged to the monoclinic space group $P2_1$, with unit-cell parameters $a = 41.4$, $b = 113.4$, $c = 47.6$ Å, $\alpha = \gamma = 90$, $\beta = 107.4^\circ$, and contained one monomer of HalY in the asymmetric unit, with a calculated Matthews coefficient of $2.3 \text{ Å}^3 \text{ Da}^{-1}$ and a solvent content of 46%. The structure of the halogenase CndH was used as a search model in molecular replacement to obtain the initial model of HalY. Manual model building and structure refinement of HalY are in progress.

1. Introduction

More than 4500 halogenated natural products have been isolated to date and the majority of these contain either chlorine or bromine (Smith *et al.*, 2013). Many of these halogenated compounds are biologically active, ranging from vancomycin, which depends on chlorine for its bactericidal activity (Williams & Bardsley, 1999), to the potent anticancer compounds cryptophycin A (Sailler & Glombitza, 1999) and rebeccamycin (Pereira *et al.*, 1996). While stereoselective or regioselective halogen incorporation is a challenge in organic chemistry, the corresponding biosynthetic pathways remain elusive and it is difficult to identify the correct regioselective halogenase and the substrate for halometabolite formation (van Pée, 2012).

Owing to the thriving process of genome sequencing, increasing numbers of halogenases with regioselectivity have been identified and characterized. The halogenases can be classified into three groups based on the cofactor of the reaction (Smith *et al.*, 2013): (i) FAD-dependent halogenases, which catalyze the electrophilic halide attack of electron-rich substrates, (ii) nonhaem iron and α -ketoglutarate-dependent halogenases, which use a substrate radical mechanism to allow the halogenation of an aliphatic carbon centre, and (iii) SAM (S-adenosylmethionine)-dependent halogenases, including the natural fluorinase first isolated from *Streptomyces cattleya*, which catalyze the formation of 5'-fluorodeoxyadenosine through nucleophilic fluorination.

The FAD-dependent halogenases comprise the largest and the most extensively characterized group of halogenases. The three-dimensional structures of FAD-dependent halogenases reveal two discrete reaction modules (Dong *et al.*, 2005; Zhu *et al.*, 2009). The conserved FAD-binding module uses FADH₂ to form a hypohalous acid, which is then delivered through a 10 Å channel to the substrate-binding module; the substrate-

binding module is diverse, determines the substrate specificity of various FAD-dependent halogenases and is where the electrophilic halogenation occurs. Based on the form of the substrate, FAD-dependent halogenases can be divided into variant A and variant B subgroups (Buedenbender *et al.*, 2009): variant A halogenases accept the free form of the substrate, such as the amino acid tryptophan, while variant B

Figure 1

Multiple sequence alignment of the FAD-dependent halogenases from *Streptomyces* sp. JCM9888 against the homologous halogenases from the biosynthetic pathways of pyoluteorin (PltA, AAY92059.1), chondrochloren (CndH, Q7S6N4) and chloramphenicol (CmlS, Q9AL91).

halogenases require a carrier protein to link the substrate covalently, which makes it difficult to establish an *in vitro* enzyme assay for variant B halogenases.

It has been found that *Streptomyces* sp. JCM9888 produces a halogenated adenosine analogue which has antibiotic activity (Isono *et al.*, 1984). Searching the draft genome sequence revealed only two open reading frames which are homologous to regioselective halogenases (Zhao *et al.*, 2014), referred to here as HalX and HalY. They have high amino-acid sequence identity to each other (61% sequence identity) and the highest similarity to the FAD-dependent variant B halogenase PltA (51 and 68% sequence identity and similarity, respectively; Fig. 1). PltA catalyzes the dichlorination of a pyrrolyl-S-carrier protein during pyoluteorin biosynthesis, with an established *in vitro* enzyme assay (Dorrestein *et al.*, 2005). However, the lack of structural data has delayed our understanding of the reaction details of PltA. The available structural homologues of HalX or HalY are the chondrochloren halogenase CndH (PDB entry 3e1t; 27% sequence identity; Buedenbender *et al.*, 2009) and the chloramphenicol halogenase CmlS (PDB entry 3i3l; 21% sequence identity; Podzelinska *et al.*, 2010) (Fig. 1). The enzyme activities of both CndH and CmlS have yet to be reconstituted *in vitro*. As the

Table 1

Macromolecule-production information.

Source organism and DNA source	<i>Streptomyces</i> sp. JCM9888
Forward primer†	GGCAGCCATATGGTGACCGATTTCGATATCG
Reverse primer‡	GCTCGAATTCTCAGAAAAGCGGTGCG
Clone and expression vector	pET-28a
Expression host	<i>E. coli</i> BL21 (DE3) PlysS
Complete amino-acid sequence of the construct produced§	MGSSHHHHHSSGLVPRGSHMVTDFDIGIIGGGP-GGTAAASYLARAGLSVAVFEGELFPREHVGES-LVPATTPVLREIGALAKLDEAGFPRKYGAAWT-SPSGARASDTFGMADLAFEDVLGGQSYTY-HVDRGRFDQLLEHSSTLGAKVFEGARVRRVD-FDGPPIVHASGGSESRTGVRMVIDASGRGT-LLGRQLGLKVTDPFTDQYAVHGWFTGLDREAI-AGRDALADYLFVHFLPAEDSWWWQIPITDTIT-SVGVSRSRLRQAGGDREAFFWGTLEHRPEL-AAALRACTPVRPLKVDGFSYAMRQITGDGWA-MVGDAARFVDPIFSSGVGIAMNGARLLSQDVL-AAAEAGDFARQFAEYKQRRGLTNWYNFIA-VYYRLNILFTVFVQDPRYRLDVLRLQGNVYE-TDHPVLTMMQGVVETVESNPDHLWPHLGLSL-RAPAAAPLF

† The NdeI site is underlined. ‡ The EcoRI site is underlined. § The residues from the pET-28a vector are underlined.

role of HalX or HalY in adenine halogenation has not yet been determined, the ongoing structural determination and enzyme assays will help to elucidate the possible mechanism of adenine halogenation, which could further the understanding of the reaction of PltA. Here, we report the overexpression, purification and crystallization of the halogenase HalY. Preliminary X-ray diffraction analysis has also been performed.

2. Materials and methods

2.1. Macromolecule production

2.1.1. Cloning and overexpression of HalY. The gene encoding the FAD-dependent halogenase HalY was amplified from the genomic DNA of *Streptomyces* sp. JCM9888 by polymerase chain reaction (PCR) and cloned into the expression vector pET-28a between the NdeI and EcoRI sites with an N-terminal histidine tag. The constructed plasmid pET-28a-HalY was transformed into *Escherichia coli* DH5 α cells for plasmid preparation and the encoding region of HalY was verified by DNA sequencing. The pET-28a-HalY plasmid was transformed into *E. coli* BL21 (DE3) pLysS cells for expression (Table 1). A single colony of the transformants from LB-agar solid medium supplemented with 50 $\mu\text{g ml}^{-1}$ kanamycin and 34 $\mu\text{g ml}^{-1}$ chloramphenicol was picked and cultured overnight at 310 K in 100 ml LB liquid medium containing identical antibiotics as in the solid medium. The 100 ml overnight culture was used to inoculate 1 l LB liquid medium in a ratio of 1:100 with the essential antibiotics. When the cell density reached an OD_{600 nm} of 0.5, the culture temperature was decreased from 310 to 289 K and 0.1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) was added to induce the overexpression of HalY. After additional growth at 289 K for 20 h, the cells were harvested by centrifugation and were stored at 193 K until use.

Figure 2

(a) Purification profile of HalY from size-exclusion chromatography. The solid curve shows the UV absorption at 280 nm and the elution volume for the major peak of HalY is labelled. (b) SDS-PAGE of purified HalY fractions from size-exclusion chromatography; the middle three fractions were used for crystallization. Lane 1 shows molecular-weight markers labelled in kDa.

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	24-well plate
Temperature (K)	293
Protein concentration (mg ml ⁻¹)	15
Buffer composition of protein solution	20 mM Tris-HCl pH 7.8, 0.15 M NaCl
Composition of reservoir solution	0.1 M bis-tris propane pH 6.3, 16%(w/v) PEG 3350, 0.2 M NaCl
Volume and ratio of drop	2 µl (1:1 ratio of protein solution and reservoir solution)
Volume of reservoir (µl)	400

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	3W1A, BSRF
Wavelength (Å)	0.97924
Temperature (K)	100
Detector	MAR 165 CCD
Crystal-to-detector distance (mm)	120
Rotation range per image (°)	0.5
Total rotation range (°)	180
Exposure time per image (s)	10
Space group	<i>P</i> 2 ₁
Unit-cell parameters (Å, °)	<i>a</i> = 41.35, <i>b</i> = 113.09, <i>c</i> = 47.56, <i>α</i> = <i>γ</i> = 90, <i>β</i> = 107.38
Mosaicity (°)	0.102
Resolution range (Å)	21.18–1.70
Total No. of reflections	202230 (10791)
No. of unique reflections	45022 (3008)
Completeness (%)	99 (90.2)
Multiplicity	4.5 (3.6)
<i>I</i> / <i>σ</i> (<i>I</i>)	18 (6.4)
<i>R</i> _{meas}	0.065 (0.208)
Overall <i>B</i> factor from Wilson plot (Å ²)	21.329

2.1.2. Purification of HalY. The cell pellets containing the halogenase HalY from 2 l of culture were resuspended in 100 ml lysis buffer (20 mM Tris-HCl pH 7.8, 500 mM NaCl, 10 mM imidazole) and disrupted using a low-temperature high-pressure cell disruptor (JNBIO, People's Republic of China). The lysate was clarified by centrifugation at 35 000*g* for 30 min at 277 K and the supernatant was loaded onto a 5 ml HisTrap HP column (GE Healthcare, Sweden). During sample loading, a yellow protein layer accumulated at the top of the column, indicating the presence of bound FAD in HalY. After a six-column-volume wash with lysis buffer, the target protein HalY (yellow fraction) was eluted at an early stage of a linear gradient elution with elution buffer (20 mM Tris-HCl pH 7.8, 500 mM NaCl, 500 mM imidazole). The purity of HalY was further improved by running the sample through a HiLoad Superdex 200 16/60 size-exclusion chromatography column (GE Healthcare, Sweden) in sample buffer (20 mM Tris-HCl pH 7.8, 150 mM NaCl; Fig. 2*a*). The major peak fractions remained yellow and were analyzed by 12% SDS-PAGE (Fig. 2*b*).

2.2. Crystallization

For crystallization experiments, HalY was concentrated to 15 mg ml⁻¹ in sample buffer (20 mM Tris-HCl pH 7.8,

150 mM NaCl). Crystallization trials were set up in 96-well sitting-drop plates using the JCSG+ and PACT crystallization screening kits (Qiagen, Germany). Each drop consisted of 0.6 µl protein solution plus 0.6 µl precipitant and was equilibrated by vapour diffusion against 70 µl precipitant. Micro-needle crystals appeared after 3 d at 293 K. The pH range and precipitant concentration were optimized based on the initial crystallization hits. Yellow crystals with the best appearance were obtained by the hanging-drop vapour-diffusion method in 24-well plates using an optimized crystallization condition [0.1 M bis-tris propane pH 6.3, 16%(w/v) PEG 3350, 0.2 M NaCl] without supplementary FAD (Table 2). Several cryoprotectants were screened and 20% ethylene glycol was

(a)

(b)

Figure 3

(*a*) Crystals of HalY grown in 0.1 M bis-tris propane pH 6.3, 16%(w/v) PEG 3350, 0.2 M NaCl, which were used in X-ray diffraction data collection. (*b*) Diffraction image of HalY on BSRF beamline 3W1A. The circle indicates 1.70 Å resolution.

chosen as the cryoprotectant to supplement the optimized crystallization condition. The crystals were mounted in a nylon loop and flash-cooled in liquid nitrogen before diffraction.

2.3. Data collection and processing

X-ray diffraction data were collected from a single HalY crystal at 100 K on beamline 3W1A at Beijing Synchrotron Radiation Facility (BSRF) using a MAR 165 CCD detector (MAR Research, Germany). A total of 360 images were collected in a single pass of 180° with 0.5° oscillation. The diffraction images were indexed, integrated and merged in *XDS* (Kabsch, 2010) and scaled with *SCALA* (Evans, 2006) as implemented in the *xia2* package (Winter, 2010). Data-collection and processing statistics are summarized in Table 3.

3. Results and discussion

A new FAD-dependent halogenase HalY from *Streptomyces* sp. JCM9888, which has been reported to play a role in adenine halogenation (Zhao *et al.*, 2014), was successfully cloned and expressed in *E. coli* and purified to homogeneity. The molecular weight of HalY with an N-terminal histidine tag was confirmed to be 50 kDa by SDS-PAGE (Fig. 2*b*). The elution volume of HalY in size-exclusion chromatography indicated a monomeric form in solution. The protein sample appeared to be yellow throughout the purification and crystallization procedures, suggesting tight binding of FAD to HalY (Fig. 3*a*).

The crystal of HalY diffracted to a highest resolution of 1.7 Å (Fig. 3*b*). The data were processed in the monoclinic space group *P*₂₁ and are summarized in Table 3. The overall *R*_{merge} and *R*_{p.i.m} were 5.7 and 3.0%, respectively. Cell-content analysis showed one HalY monomer per asymmetric unit, with a solvent content of 46% and a Matthews coefficient of 2.3 Å³ Da⁻¹. The available structures of CndH (Buedenbender *et al.*, 2009) and CmlS (Podzelinska *et al.*, 2010) were both used as search models in molecular replacement with *Phaser* using the automatic resolution-selection strategy (McCoy *et al.*, 2007). However, only a search model of CndH prepared using *CHAINS*AW (Stein, 2008) gave an initial model of HalY, with RFZ = 8.9, TFZ = 5.6 and LLG = 63. Although the low TFZ value indicated an ambiguous solution, refinement using *phenix.refine* (Adams *et al.*, 2010) led to an *R*_{free} of 48.9%. The resulting electron-density map was visually

inspected and matched the FAD-binding module of HalY well. The initial model was further automatically built using *AutoBuild* from the *PHENIX* package (Adams *et al.*, 2010), which decreased the *R* and *R*_{free} factors to 22.8 and 26.1%, respectively. The electron-density map was unambiguous for the bound FAD cofactor and clear for the majority of the protein molecule. Manual model building and refinement are currently in progress.

Acknowledgements

The staff of beamline 3W1A at BSRF are appreciated for their assistance in data collection. This work was supported by the Start-up Fund of Sichuan University and the 1000 Plan of Sichuan Province.

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Buedenbender, S., Rachid, S., Müller, R. & Schulz, G. E. (2009). *J. Mol. Biol.* **385**, 520–530.
- Dong, C., Flecks, S., Unversucht, S., Haupt, C., van Pée, K.-H. & Naismith, J. H. (2005). *Science*, **309**, 2216–2219.
- Dorrestein, P. C., Yeh, E., Garneau-Tsodikova, S., Kelleher, N. L. & Walsh, C. T. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 13843–13848.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Isono, K., Uramoto, M., Kusakabe, H., Miyata, N., Koyama, T., Ubukata, M., Sethi, S. K. & McCloskey, J. A. (1984). *J. Antibiot.* **37**, 670–672.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
- Pée, K.-H. van (2012). *Methods Enzymol.* **516**, 237–257.
- Podzelinska, K., Latimer, R., Bhattacharya, A., Vining, L. C., Zechel, D. L. & Jia, Z. (2010). *J. Mol. Biol.* **397**, 316–331.
- Rodrigues Pereira, E., Belin, L., Sancelme, M., Prudhomme, M., Ollier, M., Rapp, M., Severe, D., Riou, J. F., Fabbro, D. & Meyer, T. (1996). *J. Med. Chem.* **39**, 4471–4477.
- Sailler, B. & Glombitza, K. W. (1999). *Nat. Toxins*, **7**, 57–62.
- Smith, D. R., Grischow, S. & Goss, R. J. (2013). *Curr. Opin. Chem. Biol.* **17**, 276–283.
- Stein, N. (2008). *J. Appl. Cryst.* **41**, 641–643.
- Williams, D. H. & Bardsley, B. (1999). *Angew Chem Int. Ed.* **38**, 1173–1193.
- Winter, G. (2010). *J. Appl. Cryst.* **43**, 186–190.
- Zhao, C., Qi, J., Tao, W., He, L., Xu, W., Chan, J. & Deng, Z. (2014). *PLoS One*, **9**, e114722.
- Zhu, X., De Laurentis, W., Leang, K., Herrmann, J., Ihlefeld, K., van Pée, K.-H. & Naismith, J. H. (2009). *J. Mol. Biol.* **391**, 74–85.