Xiaolu Duan,^{a‡} Liang Zhang,^{a‡}
Daming Zhou,^a Kaihua Ji,^b
Ting Ma,^b Wenqing Shui,^a
Guoqiang Li^{b*} and Xin Li^{a*}

^aState Key Laboratory of Medicinal Chemical Biology, College of Life Sciences, Nankai University, Tianjin 300071, People's Republic of China, and ^bKey Laboratory of Molecular Microbiology and Technology, Ministry of Education, College of Life Sciences, Nankai University, Tianjin 300071, People's Republic of China

‡ These authors contributed equally to this work.

Correspondence e-mail: gqli@nankai.edu.cn, lix1980@nankai.edu.cn

Received 2 March 2013
Accepted 24 April 2013

PDB Reference: DszC, 4jek

Crystallization and preliminary structural analysis of dibenzothiophene monooxygenase (DszC) from *Rhodococcus erythropolis*

Dibenzothiophene (DBT) and its derivatives are typical sulfur compounds found in fossil fuels. These compounds show resistance to the hydrodesulfurization treatment that is commonly used in industry. Dibenzothiophene monooxygenase (DszC) is responsible for the oxidation of DBT, which is the first and the rate-limiting step in the DBT enzymatic desulfurization 4S pathway. In this study, the crystal structure of DszC from *Rhodococcus erythropolis* DS-3 is reported. The crystal of native DszC belonged to space group *P*1, with unit-cell parameters $a = 96.16$, $b = 96.27$, $c = 98.56$ Å, $\alpha = 81.03$, $\beta = 67.57$, $\gamma = 85.84^\circ$. To determine the phase, SAD X-ray diffraction data were collected from a SeMet-derivative DszC crystal, which also belonged to space group *P*1, with unit-cell parameters $a = 95.379$, $b = 95.167$, $c = 94.891$ Å, $\alpha = 87.046$, $\beta = 70.536$, $\gamma = 79.738^\circ$. Further structural analysis of DszC is in progress.

1. Introduction

With the ever-growing demand for energy, worldwide crude-oil consumption is continuously rising. The increasing use of high-sulfur crude oil is known to produce petroleum products with a high sulfur content (Torktaz *et al.*, 2012). Combustion of these sulfur-containing fossil fuels has contributed to severe environmental pollution (Li *et al.*, 2008). Hydrodesulfurization is a widely used chemical process to lower the sulfur content of fossil fuels, but certain sulfur compounds in fossil fuels such as dibenzothiophene (DBT) and its derivatives are resistant to hydrodesulfurization (Kabe *et al.*, 1992; Monticello & Finnerty, 1985). Therefore, an alternative enzymatic method has drawn increasing attention for the desulfurization of DBT and its derivatives. The typical DBT catalytic desulfurization 4S pathway involves two classes of enzymes: flavin-dependent monooxygenases, such as DszA and DszC, and the desulfinase DszB. In this pathway, DBT is first transformed into dibenzothiophene sulfoxide (DBTO) and dibenzothiophene sulfone (DBTO₂) catalyzed by DszC; DszA then further oxidizes DBTO₂ to 2-hydroxybiphenyl-2-sulfinic acid (HBPS), and finally DszB hydrolyzes HBPS to 2-hydroxybiphenyl (HBP) and sulfite (Fig. 1; Calzada *et al.*, 2009; Olmo *et al.*, 2005; Denome *et al.*, 1994; Gray *et al.*, 1996; Gupta *et al.*, 2005; Kilbane, 2006; Piddington *et al.*, 1995). Notably, the initial transformation catalyzed by DszC is regarded to be the major bottleneck that determines the desulfurization efficiency of the entire 4S pathway.

To date, the structures of several dehydrogenases and hydroxylases with sequence similarity to the dibenzothiophene monooxygenase (DszC) from *Rhodococcus erythropolis* DS-3 have been solved. However, none of these enzymes have been reported to catalyze the DBT oxidation reaction catalyzed by DszC. In this study, we crystallized DszC from *R. erythropolis* DS-3 and solved its structure using the single-wavelength anomalous diffraction phasing method.

2. Methods and materials

2.1. Cloning and transformation

The gene encoding dibenzothiophene monooxygenase (DszC) was PCR-amplified from the *R. erythropolis* DS-3 genome. Sequences for

the *EcoRI* and *HindIII* restriction sites were added to the sense (5'-CCGGAATTCGCGATGACACTGTACCTG-3') and antisense (5'-CCCAAGCTTCTCAGGAGGTGAAGCCG-3') primers, respectively. PCR fragments were purified by agarose-gel electrophoresis and digested with the two restriction enzymes prior to insertion into the pET28a(+) (Novagen) vector. The correctly constructed recombinant plasmid was transformed into *Escherichia coli* BL21 (DE3) competent cells (Novagen) and the auxotrophic *E. coli* B834 (DE3) strain (Novagen) for expression.

2.2. Expression

The *E. coli* BL21 (DE3) cells containing the recombinant plasmid were cultured in Luria-Bertani (LB) medium supplemented with 100 µg ml⁻¹ kanamycin. The bacterial cells were grown at 310 K and 200 rev min⁻¹ until the OD₆₀₀ reached 0.8. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the culture to a final concentration of 500 µM for induction. N-terminally His-tagged native DszC protein was expressed by continuing the culture for an additional 20 h at 289 K and 200 rev min⁻¹.

To obtain SeMet-derivatized DszC (Se-DszC), *E. coli* B834 (DE3) bacterial cells containing the recombinant plasmid that had been cultured overnight in adapted LB medium (20% LB and 80% PBS buffer) were collected by centrifugation (Eppendorf Centrifuge 5804R). The pellet was resuspended in selenomethionine medium [3% (w/v) sucrose, 6.5% (w/v) yeast nitrogen base, 47 mM Na₂HPO₄, 22 mM KH₂PO₄, 18.7 mM NH₄Cl, 8 mM NaCl] supplemented with 30 µg ml⁻¹ L-selenomethionine (Sigma-Aldrich) and 100 µg ml⁻¹ kanamycin. The culture was continued at 310 K and 200 rev min⁻¹ until the OD₆₀₀ reached 0.8. Another 30 mg of L-selenomethionine was added per litre of culture and IPTG was also added to a final concentration of 300 µM for induction. The Se-DszC protein was expressed at 289 K and 200 rev min⁻¹ for an additional 24 h.

2.3. Purification

The *E. coli* BL21 (DE3) cells were harvested by centrifugation (Beckman/Coulter Avanti J-26XP) at 5000 rev min⁻¹ and 277 K for 30 min. The cell pellet was resuspended in cold lysis buffer consisting of 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% (v/v) glycerol and lysed by sonication on ice. The supernatant was obtained by centri-

fuging the cell lysate at 18 000 rev min⁻¹ and 277 K (Beckman/Coulter Avanti J-26XP) for 1 h. Standard nickel-affinity chromatography (GE Healthcare) was then performed for preliminary purification of the native His-tagged DszC protein from the supernatant. The Ni-NTA resin was pre-equilibrated with lysis buffer and the supernatant was then applied. After extensively washing the resin with lysis buffer supplemented with 20 mM imidazole, the bound protein was eluted using elution buffer [20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% (v/v) glycerol, 330 mM imidazole]. The eluted DszC protein was dialyzed against 20 mM Tris-HCl pH 8.0, 10 mM NaCl overnight. The protein solution was then loaded onto an anion-exchange chromatography column [1 ml HiTrap Q HP column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 8.0 at 283 K] for further purification. The DszC protein eluted from the HiTrap Q HP column was finally purified by size-exclusion chromatography using a Superdex 200 10/300 column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl. The final yield of purified DszC was about 10 mg per litre of LB medium. The purity of the DszC protein was evaluated by SDS-PAGE and it was concentrated to 20 mg ml⁻¹ for crystallization (Fig. 2). The Se-DszC protein was purified using the same method as used for the native DszC protein described above; the yield was about 5 mg of purified Se-DszC per litre of medium.

2.4. Crystallization and data collection

Crystal screening was performed at 293 K by the sitting-drop vapour-diffusion method. 1 µl protein solution (10 mg ml⁻¹) was mixed with 1 µl reservoir solution and equilibrated against 100 µl reservoir solution. Commercial crystallization kits from Hampton Research and Emerald BioSystems were used for screening. Initial crystals of native DszC were observed in Index (Hampton Research) condition Nos. 55 and 67 and Crystal Screen (Hampton Research) condition No. 26 after one week. All three crystallization conditions were optimized using the hanging-drop vapour-diffusion method: 2 µl protein solution (20 mg ml⁻¹) was mixed with 2 µl reservoir solution and equilibrated against 300 µl reservoir solution. Finally, an optimized crystallization reservoir solution for native DszC was determined to consist of 17.5% PEG 3350, 200 mM bis-tris pH 7.5, 200 mM ammonium sulfate.

Figure 1

A brief description of the 4S pathway for the biodesulfurization of DBT by *R. erythropolis* DS-3. DszC oxidizes DBT first to DBTO and further to DBTO₂. In the third step, DszA catalyzes the transformation of DBTO₂ to HPBS. In the last step, DszB hydrolyzes HPBS to HPB and sulfite.

Crystallization of the Se-DszC protein began with the optimized reservoir solution used for the crystallization of native DszC. After changing the buffer pH and adjusting the precipitant concentration, diffraction-quality Se-DszC crystals were obtained using a reservoir solution consisting of 25% PEG 3350, 100 mM PIPES pH 7.0, 200 mM ammonium sulfate.

Diffraction data for native DszC were collected on beamline BL17U at Shanghai Synchrotron Radiation Facility (SSRF), Shanghai, People's Republic of China. Selenomethionine anomalous diffraction data for Se-DszC were collected on the BL5A beamline at the Photon Factory (PF), Tsukuba, Japan.

2.5. Structure determination and refinement

The data sets were processed with the *HKL-2000* (Otwinowski & Minor, 1997) and *CCP4* (Winn *et al.*, 2011) program suites. A 2.9 Å resolution structure of Se-DszC was solved by the single-wavelength anomalous diffraction phasing method at the selenium absorption edge using the program *PHENIX* (Adams *et al.*, 2010). The 2.9 Å resolution phases were extended to 2.4 Å using the native data set

and the program *Phaser* (McCoy *et al.*, 2007). Further refinement was performed using *REFMAC5* (Murshudov *et al.*, 2011), *PHENIX* and *Coot* (Emsley *et al.*, 2010). The Ramachandran plot for the final model was checked with *MolProbity* (Chen *et al.*, 2010). Structure figures were produced using *PyMOL* (DeLano, 2002).

3. Results

The purity of the native DszC and Se-DszC proteins was evaluated by SDS-PAGE and was adequate for crystallization (Fig. 2). Diffraction-quality plate-shaped crystals of native DszC were obtained from hanging drops equilibrated for one week against the optimal crystallization reservoir solution (Fig. 3). A 2.4 Å resolution diffraction data set was collected at the SSRF synchrotron; the crystal belonged to space group *P1*, with unit-cell parameters $a = 96.16$, $b = 96.27$, $c = 98.56$ Å, $\alpha = 81.03$, $\beta = 67.57$, $\gamma = 85.84^\circ$ (Fig. 4). In order to solve the structure, we initially tried the molecular-replacement method. The five different protein structures with the highest protein-sequence similarity to DszC [PDB entries 2vig (25% identity; A. C. W. Pike, N. Pantic, E. Parizotto, O. Gileadi, E. Ugochukwu, F. Von Delft, J. Weigelt, C. H. Arrowsmith, A. Edwards & U. Oppermann, unpublished work), 3pfd (24% identity; Abendroth *et al.*, 2011), 1jqj (24% identity; Battaille *et al.*, 2002), 3nfd (24% identity; Seattle

Figure 2
Coomassie Blue-stained SDS-PAGE images showing recombinant native DszC and Se-DszC before crystallization. Lane *M* contains Protein Ruler I (TransGen Biotech; labelled in kDa), lane NDC contains native DszC protein with a His tag and lane SDC contains Se-DszC protein with a His tag.

Figure 3
Plate-shaped crystals of native DszC.

Figure 4
X-ray diffraction pattern of the native DszC crystal. The circles from the edge to the centre indicate 2.4, 3, 4, 6 and 10 Å resolution, respectively.

Structural Genomics Center for Infectious Disease, unpublished work) and livh (22% identity; Tiffany *et al.*, 1997)] were used as search models. Unfortunately, none of the models led to correct phases; therefore, we purified and crystallized the Se-DszC protein. It is noteworthy that the purified selenomethionine-labelled protein seemed to contain a minute amount of molecular chaperone, probably because it was expressed in a harsher medium than the LB medium. This contaminant may have affected the crystallization to a certain degree, as the resulting crystals were much thinner, smaller and more fragile compared with the native DszC crystals (Fig. 5). Fortunately, we were able to acquire high-quality diffraction data from an Se-DszC crystal at the Photon Factory, Tsukuba, Japan. An anomalous scattering data set was collected to a maximal resolution of 2.9 Å. The space group was also *P*1, the same as that of the native DszC crystal, with unit-cell parameters *a* = 95.379, *b* = 95.167, *c* = 94.891 Å, α = 87.046, β = 70.536, γ = 79.738°. Detailed statistics for

the data sets are summarized in Table 1. The coordinates and structure factors have been deposited in the Protein Data Bank (Bernstein *et al.*, 1977; Berman *et al.*, 2000, 2003) as entry 4jek.

The structure of DszC was solved and refined. The asymmetric unit contained eight DszC subunits assembled as two tetramers (Fig. 6*a*). The structure of each DszC subunit revealed that the DszC structure was composed of two parts: one was formed by α -helices (residues 19–104 and 257–417), while the other was made up by β -strands

Table 1
Data-collection and refinement statistics.

	Se-DszC (peak)	Native DszC
Data collection		
Space group	<i>P</i> 1	<i>P</i> 1
Unit-cell parameters (Å, °)	<i>a</i> = 95.379, <i>b</i> = 95.167, <i>c</i> = 94.891, α = 87.046, β = 70.536, γ = 79.738	<i>a</i> = 96.16, <i>b</i> = 96.27, <i>c</i> = 98.56, α = 81.03, β = 67.57, γ = 85.84
Wavelength (Å)	0.9790	1.0000
Resolution (Å)	50–2.9 (2.95–2.90)	50–2.4 (2.49–2.40)
Average <i>I</i> / σ (<i>I</i>)	30.12 (5.61)	15.01 (2.25)
Total reflections	308566	465998
Unique reflections	66880	122302
Multiplicity	4.6 (4.8)	3.8 (3.7)
Completeness (%)	93.1 (98.3)	97.4 (97.1)
<i>R</i> _{merge} [†] (%)	13.3 (45.0)	7.7 (48.2)
Refinement		
Resolution (Å)	50–2.9	50–2.4
Average <i>B</i> factor (Å ²)	NA	37.55
<i>R</i> _{work} / <i>R</i> _{free} [‡] (%)	NA	20.69/24.61
R.m.s.d., bond lengths§ (Å)	NA	0.014
R.m.s.d., bond angles§ (°)	NA	1.249
Ramachandran plot, residues in (%)		
Favoured regions	NA	98.6
Allowed regions	NA	1.4
Generally allowed regions	NA	0
Disallowed regions	NA	0

[†] $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is an individual intensity measurement and $\langle I(hkl) \rangle$ is the average intensity for all reflections. [‡] $R_{\text{work}} / R_{\text{free}} = \sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure factors, respectively. [§] R.m.s. deviations relate to the Engh and Huber parameters (Engh & Huber, 1991).

Figure 5
Crystals of Se-DszC.

Figure 6
Ribbon diagrams of the DszC structure. (a) The protein content in the asymmetric unit. Each DszC subunit is coloured differently. (b) Ribbon diagram of a DszC subunit. Residues 19–104 (α domain I), 257–417 (α domain II) and 125–246 (β domain) are coloured red, grey and blue, respectively.

(residues 125–246), as shown in Fig. 6(b). The R factor and R_{free} for the structural model are 0.2069 and 0.2461, respectively. DszC contains 417 residues, but the first 18 residues are missing owing to disorder. Of the 399 residues observable in the model, 98.6 and 1.4% are localized in the favoured and the allowed regions of the Ramachandran plot, respectively. Detailed refinement statistics are summarized in Table 1.

Further structural and functional studies of DszC are in progress.

We thank beamline BL17U at Shanghai Synchrotron Radiation Facility (SSRF, People's Republic of China) and beamline BL5A at the Photon Factory, Japan for data collection. This study was supported by the Young Scientists Fund of the National Natural Science Foundation of China (Grant Nos. 31000345 and 31000056), the National Basic Research Program of China (973 Program; Grant No. 2010CB833600) and the Natural Science Foundation of Tianjin, China (Grant No. 09JCZDJC18000).

References

- Abendroth, J., Gardberg, A. S., Robinson, J. I., Christensen, J. S., Staker, B. L., Myler, P. J., Stewart, L. J. & Edwards, T. E. (2011). *J. Struct. Funct. Genomics*, **12**, 83–95.
- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Battaille, K. P., Molin-Case, J., Paschke, R., Wang, M., Bennett, D., Vockley, J. & Kim, J.-J. P. (2002). *J. Biol. Chem.* **277**, 12200–12207.
- Berman, H., Henrick, K. & Nakamura, H. (2003). *Nature Struct. Biol.* **10**, 980.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *Eur. J. Biochem.* **80**, 319–324.
- Calzada, J., Zamarro, M. T., Alcón, A., Santos, V. E., Díaz, E., García, J. L. & Garcia-Ochoa, F. (2009). *Appl. Environ. Microbiol.* **75**, 875–877.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
- Denome, S. A., Oldfield, C., Nash, L. J. & Young, K. D. (1994). *J. Bacteriol.* **176**, 6707–6716.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Engh, R. A. & Huber, R. (1991). *Acta Cryst.* **A47**, 392–400.
- Gray, K. A., Pogrebinsky, O. S., Mrachko, G. T., Xi, L., Monticello, D. J. & Squires, C. H. (1996). *Nature Biotechnol.* **14**, 1705–1709.
- Gupta, N., Roychoudhury, P. K. & Deb, J. K. (2005). *Appl. Microbiol. Biotechnol.* **66**, 356–366.
- Kabe, T., Ishihara, A. & Tajima, H. (1992). *Ind. Eng. Chem. Res.* **31**, 1577–1580.
- Kilbane, J. J. II (2006). *Curr. Opin. Biotechnol.* **17**, 305–314.
- Li, G.-Q., Li, S.-S., Zhang, M.-L., Wang, J., Zhu, L., Liang, F.-L., Liu, R.-L. & Ma, T. (2008). *Appl. Environ. Microbiol.* **74**, 971–976.
- 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.
- Monticello, D. J. & Finnerty, W. R. (1985). *Annu. Rev. Microbiol.* **39**, 371–389.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Olmo, C. H. del, Santos, V. E., Alcon, A. & Garcia-Ochoa, F. (2005). *Biochem. Eng. J.* **22**, 229–237.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Piddington, C. S., Kovacevich, B. R. & Rambosek, J. (1995). *Appl. Environ. Microbiol.* **61**, 468–475.
- Tiffany, K. A., Roberts, D. L., Wang, M., Paschke, R., Mohsen, A. W., Vockley, J. & Kim, J.-J. P. (1997). *Biochemistry*, **36**, 8455–8464.
- Torktaz, I., Etemadifar, Z. & Derikvand, P. (2012). *Bioinformation*, **8**, 246–250.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.