STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Crystal structure determination of the halogenase CtcP from *Streptomyces aureofaciens*

Lijuan Yin*

Department of Pathology, Changhai Hospital, The Second Military Medical University, 168 Changhai Road, Shanghai 200433, People's Republic of China. *Correspondence e-mail: yinlijuan2@163.com

Received 20 April 2022
Accepted 24 June 2022

Edited by M. J. van Raaij, Centro Nacional de Biotecnología – CSIC, Spain

Keywords: crystal structure; CtcP; halogenases; chlortetracycline; *Streptomyces aureofaciens*; antibiotics.

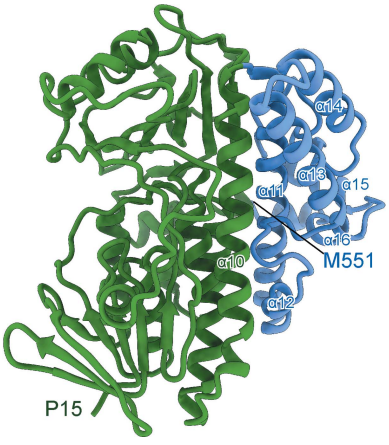
PDB reference: CtcP, 7xgb

Supporting information: this article has supporting information at journals.iucr.org/f

Chlortetracycline (CTC), a derivative of tetracycline (TC), is a broadly used antibiotic that inhibits the synthesis of bacterial proteins by competing with the A-site tRNA on ribosomes. A recent study showed that during the biosynthesis of CTC in *Streptomyces aureofaciens*, the halogenase CtcP catalyzes the final chlorination reaction and transforms TC into CTC. However, the structure of this fundamental enzyme is still lacking. Here, selenomethionine-derivatized CtcP from *S. aureofaciens* was overexpressed and purified and its structure was determined at 2.7 Å resolution. The structure of CtcP reveals the conserved monooxygenase domain shared by all flavin-dependent halogenases and a unique C-terminal domain. Although FAD was not observed in the structure, the monooxygenase domain has a conserved FAD-binding pocket and active center. The C-terminal domain displays an α-helical bundle fold, which could contribute to substrate specificity. This work provides a molecular basis for enzyme engineering to improve the industrial production of CTC.

1. Introduction

The tetracycline (TC) family contains some of the most commonly used broad-spectrum antibiotics against numerous pathogens, including Gram-positive and Gram-negative bacteria, as well as atypical organisms (Bahrami *et al.*, 2012; Nelson & Levy, 2011; Grossman, 2016). TC inhibits the translation of bacterial proteins by binding to 70S ribosomes and competing with A-site tRNA (Wilson, 2009; Brodersen *et al.*, 2000; Pioletti *et al.*, 2001; Jenner *et al.*, 2013). To date, multiple derivatives of TC have been discovered or semi-synthesized, including chlortetracycline (CTC). As a derivative, CTC has a chlorine group at the C7 position of its tetracycline backbone. It was the first member of the TC family to be isolated from *Streptomyces aureofaciens* more than 70 years ago (Duggar, 1948).



Intriguingly, *S. aureofaciens* produces both CTC and TC (Zhu *et al.*, 2013). Under normal culture conditions this bacterium mainly synthesizes CTC, while TC is only a side product. However, CTC production can be inhibited by adding aminopterin to switch to demeclocycline production, or by adding mercaptothiazole or potassium bromide to switch to TC production. Extensive studies have already uncovered the biosynthetic pathway of TC and CTC in this bacterium (Dairi *et al.*, 1995; Nakano *et al.*, 2000, 2004). Moreover, a recent study found that the rate-limiting chlorination step in CTC biosynthesis is catalyzed by two enzymes: the flavin adenine dinucleotide (FAD)-dependent halogenase CtcP and the flavin reductase CtcQ (Zhu *et al.*, 2013).

In *S. aureofaciens*, CtcP depletion can completely block the production of CTC without potassium bromide supplementation. Moreover, overexpression of CtcP in this bacterium remarkably increases the yield of CTC. In addition, CtcP

preferentially catalyzes the chlorination reaction on authentic CTC (4*S*) rather than 4-*epi*-CTC (4*R*) *in vitro* (Zhu *et al.*, 2013). Thus, CtcP is an active halogenase responsible for the final chlorination reaction during CTC biosynthesis. CtcP, formerly known as Cts4, contains 555 amino-acid residues and shares 72% sequence identity with DacE (Wang *et al.*, 2012), 47% with PrnA (Dong *et al.*, 2005) and 41% with PltM (Zhang & Parry, 2007; Mori *et al.*, 2019). It harbors the classic monooxygenase domain shared by all flavin-dependent halogenases, which contains all of the FAD/FMN-binding motifs that are necessary to coordinate $FADH_2$ during the reaction (Dym & Eisenberg, 2001; Phintha *et al.*, 2020; Blasiak & Drennan, 2009). Similar to other halogenases, CtcP harbors a C-terminal domain that may contribute to its substrate recognition (Phintha *et al.*, 2020; Blasiak & Drennan, 2009). However, no structural information on CtcP is yet available.

Here, selenomethionine-derivatized CtcP was over-expressed, purified and crystallized. Using single-wavelength anomalous diffraction (SAD), its structure was solved at 2.7 Å resolution in an apo form. Structure analysis revealed that the N-terminus of CtcP contains the classic flavin-dependent halogenase fold and the C-terminus forms an α-helical bundle, which together with the N-terminal domain forms the substrate-binding pocket. In summary, the structure provides a molecular basis for enzyme engineering to improve the industrial production of CTC.

2. Materials and methods

2.1. CtcP protein purification

The CtcP gene was amplified from *S. aureofaciens* genomic DNA using the following primers: forward primer, GATCGC GGATCCATGACCGACACAACCGCGGATCAG; reverse primer, GCACCGGAATTCCTACGGCTTGACCCGCATC GCCG. The DNA fragment corresponding to the correct CtcP gene was purified and then digested using BamHI and EcoRI restriction enzymes overnight at 37°C. Digested fragments were inserted into a modified pET-28a vector. The final CtcP construct (pET-28a-CtcP) contains a 6×His tag followed by a SUMO protein tag at its N-terminus. Macromolecule-production information is summarized in Table 1.

The plasmids were transformed into *Escherichia coli* BL21 (DE3) cells by heat-shock transformation and the cells were cultured in 10 ml LB medium overnight. They were then transferred to 2 l 2×YT medium supplemented with kanamycin (50 µg ml^{-1}) and culture continued at 37°C. When the OD_{600} of the culture reached 1.0, expression of the CtcP protein was induced by adding 0.1 m*M* isopropyl β-D-1-thio-galactopyranoside (IPTG) at 15°C overnight. The cells were harvested by centrifugation, resuspended in lysis buffer (20 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl) and lysed using a French press. The cell lysate was then centrifuged at 15 000 rev min^{-1} for 15 min and the supernatant was loaded onto Ni–NTA resin (GE Healthcare) pre-equilibrated with lysis buffer. The bound protein was washed three times with wash buffer (20 m*M* Tris–HCl pH 8.0, 300 m*M* NaCl, 20 m*M* imidazole) and digested on-column by adding Ulp1 protease overnight to elute the CtcP protein. The cleaved protein was further purified by ion-exchange chromatography and size-exclusion chromatography, concentrated using an Amicon Ultra concentrator and stored in buffer consisting of 10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 5 m*M* DTT.

Table 1
Macromolecule-production information.

Source organism	*Streptomyces aureofaciens*
DNA source	*Streptomyces aureofaciens* genomic DNA
Forward primer	GATCGCGGATCCATGACCGACACAACCGCG GATCAG
Reverse primer	GCACCGGAATTCCTACGGCTTGACCCGCAT CGCCG
Expression vector	pET-28a
Expression host	*E. coli* BL21 (DE3)
Complete amino-acid sequence of construct produced	SMTDTTADQTRHGDRPYDVVIIGSGLSGTM LGSILAKHGFRIMLLDGAHHPRFAVGES TIGQTLVVLRLISDRYGVPEIANLASFQ DVLANVSSSHGQKSNFGFMFHRDGEEPD PNETSQFRIPSIVGNAAHFFRQDTDSYM FHAAVRYGCDARQYYRVENIEFDDGGVT VSGADGSTVRARYLVDASGFRSPLARQL GLREEPSRLKHHARSIFTHMVGVDAIDD HVDTPAELRPPVPWNDGTMHHIFERGWM WIIPFNNHPGATNPLCSVGIQLDERRYP ARPDLTPEEEFWSHVDRFPAVQRQLKGA RSVREWVRTDRMQYSSSRTVGERWCLMS HAAGFIDPLFSRGLSNTCEIINALSWRL MAALREDDFAVERFAYVEELEQGLLDWN DKLVNNSFISFSHYPLWNSVFRIWASAS VIGGKRILNALTRTKETGDDSHCQALDD NPYPGLWCPLDFYKEAFDELTELCEAVD AGRTTAEEAARVLEQRVRESDWMLPALG FNDPDTHHINPTADKMIRIAEWATGHHR PEIRELLAASAEEVRAAMRVKP

To purify selenomethionine-derivatized protein, after the overnight preculture the transformants were transferred into M9 medium and culture continued at 37°C. When the OD_{600} of the culture reached 0.6–0.8, the cells were switched into a shaker at 15°C. The target CtcP protein was induced by adding 0.1 m*M* IPTG and 50 mg l^{-1} selenomethionine to the medium overnight.

2.2. Crystallization and data collection

Initial crystallization screening was carried out using commercially available kits from Hampton Research at 4 and 18°C. The final crystals of selenomethionine-derivatized CtcP protein were grown at 18°C using the hanging-drop vapor-diffusion method by mixing equal volumes of protein solution and crystallization buffer consisting of 0.1 *M* bis-Tris pH 6.0, 0.2 *M* magnesium sulfate, 20%(*w*/*v*) polyethylene glycol 3350. Crystals were slowly equilibrated with a cryoprotectant buffer containing reservoir buffer plus 5%(*v*/*v*) glycerol and were flash-cooled in liquid nitrogen. All data sets were collected on beamline BL17U1 at the Shanghai Synchrotron Radiation Facility (SSRF) using an ADSC Quantum 315r detector. The exposure time was 0.5 s per frame and the oscillation range was 1° per frame. The data were processed using *HKL*-2000 (Otwinowski & Minor, 1997). In order to obtain crystal structures of CtcP with FAD or TC, crystal soaking was carried out. Different concentrations of FAD and TC were

Table 2
Crystallization.

Method	Hanging-drop vapor diffusion
Plate type	24-well plate
Temperature (K)	291
Protein concentration (mg ml^{-1})	10
Buffer composition of protein solution	10 m*M* Tris–HCl pH 8.0, 150 m*M* NaCl, 5 m*M* DTT
Composition of reservoir solution	0.1 *M* bis-Tris pH 6.0, 0.2 *M* magnesium sulfate heptahydrate, 20%(*w*/*v*) polyethylene glycol 3350
Volume and ratio of drop	2 µl, 1:1 ratio
Volume of reservoir (µl)	200

incubated with the CtcP crystals for different times. However, none of the conditions that were tried preserved the diffraction of the crystals. The crystals even dissolved when incubated with higher concentrations of the molecules for longer times. Thus, in the end, no complex structures could be obtained. Crystallization information is provided in Table 2.

2.3. Structure determination

The structure was determined by the single-wavelength anomalous dispersion (SAD) method with *Phaser* (McCoy *et al.*, 2007) from the *Phenix* suite (Liebschner *et al.*, 2019). A total of 49 selenomethionine residues could be found in one asymmetric unit. The atomic model was built manually using *Coot* (Emsley *et al.*, 2010) and was refined with *phenix.refine* (Adams *et al.*, 2010). The structure was validated using *MolProbity* (Chen *et al.*, 2010). All figures were generated by *ChimeraX* (Goddard *et al.*, 2018). The data-collection and final refinement statistics are summarized in Tables 3 and 4.

3. Results

3.1. The overall structure of CtcP

To obtain the crystal structure, the full-length CtcP protein with one extra serine residue at the N-terminus left by Ulp1 protease cleavage was used. Wild-type and selenomethionine-derivatized CtcP proteins were purified to homogeneity, which showed that CtcP is a monomer in solution. Both native and selenomethionine-derivatized CtcP crystals could be obtained, but only the latter diffracted well. Using a data set collected from a selenomethionine-derivatized CtcP crystal, the structure was determined by the single-wavelength anomalous dispersion (SAD) method at 2.7 Å resolution. Four molecules were observed in one crystallographic asymmetric unit packed in space group $P22_12_1$. Although the full-length protein was used for crystallization, only residues ranging 7–551 could be built in the final model. Unfortunately, despite the purified protein displaying a yellowish color indicative of the presence of FAD, no reliable density for FAD molecules could be built. Soaking crystals in solutions with high concentrations of FAD or TC also failed to lead to the solution of structures of their complexes with CtcP.

Like other FAD-dependent halogenases, the overall structure of CtcP can be clearly divided into N-terminal and C-terminal domains (Phintha *et al.*, 2020; Dym & Eisenberg, 2001; Blasiak & Drennan, 2009). The N-terminal domain displays a Rossmann-like fold and harbors all of the FAD-binding motifs, which are supposed to bind to $FADH_2$ during the reaction (Phintha *et al.*, 2020; Dym & Eisenberg, 2001; Blasiak & Drennan, 2009). The C-terminal domain displays an α-helical bundle fold that contains seven protein α-helices. It binds closely to the N-terminal domain and contributes to the formation of the substrate-binding pocket (Fig. 1).

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL17U, SSRF
Wavelength (Å)	0.97923
Temperature (K)	100
Detector	ADSC Q315r
Crystal-to-detector distance (mm)	400
Rotation range per image (°)	1
Total rotation range (°)	360
Exposure time per image (s)	0.5
Space group	$P22_12_1$
a, *b*, *c* (Å)	116.393, 121.281, 179.194
α, β, γ (°)	90, 90, 90
Mosaicity (°)	0.651
Resolution range (Å)	50–2.7 (2.8–2.7)
Total No. of reflections	1038277
No. of unique reflections	70271
Completeness (%)	100
Multiplicity	14.8 (14.9)
$\langle I/\sigma(I)\rangle$	15.5 (4.4)
$R_{\rm r.i.m.}$ (%)	0.177 (0.812)
Overall *B* factor from Wilson plot (Å2)	37.030

Table 4
Structure refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	46.11–2.70 (2.797–2.700)
Completeness (%)	99.77 (98.15)
No. of reflections, working set	68153 (6598)
No. of reflections, test set	2010 (206)
Final $R_{\rm cryst}$ (%)	0.1766 (0.2324)
Final $R_{\rm free}$ (%)	0.2240 (0.3227)
No. of non-H atoms	
Total	17336
Solvent	288
Protein	17048
R.m.s. deviations	
Bond lengths (Å)	0.010
Angles (°)	1.16
Average *B* factors (Å^2)	
Overall	38.35
Macromolecules	38.44
Solvent	33.08
Ramachandran plot	
Favored (%)	97.00
Additionally allowed (%)	2.82
Outliers (%)	0.18

3.2. Structure comparison of CtcP with other halogenases

The *S. aureofaciens* CtcP structure solved here was then compared with available structures of all types of halogenases from various species. The results revealed that the *S. aureofaciens* CtcP structure had the highest similarity to that of the phenolic halogenase PltM from *Pseudomonas protegens* (Mori *et al.*, 2019). A low r.m.s.d. value (1.026 Å) was obtained by a

direct structural alignment of 2465 atoms between CtcP and PltM (PDB entry 6bza), confirming their high conformational similarity (Mori *et al.*, 2019). Although both the N-terminal and the C-terminal domains are structurally comparable in

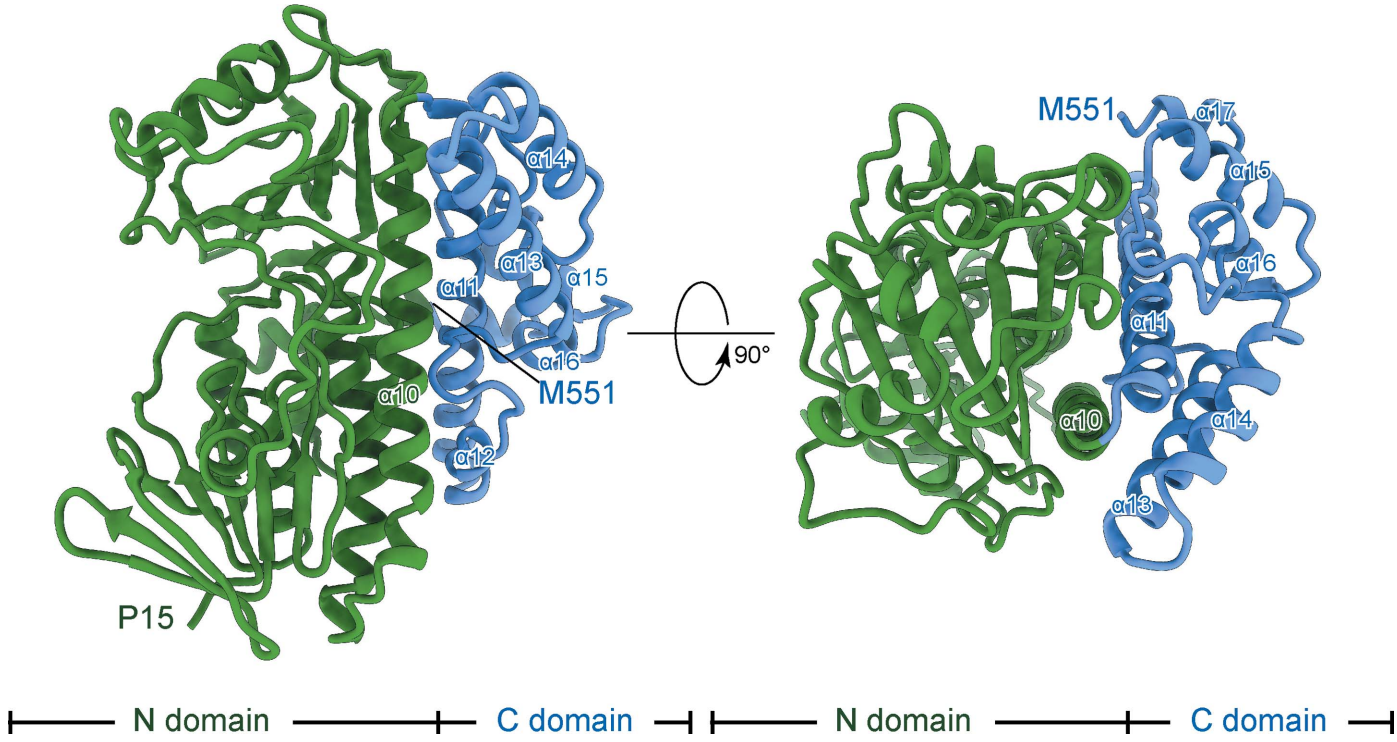


Figure 1
The crystal structure of CtcP. The crystal structure of CtcP from *S. aureofaciens* is shown in ribbon representation in two different views. The N-terminal (green) and C-terminal (blue) domains of CtcP are color-coded. α-Helix 10 from the N-terminal domain and α-helices 11–17 from the C-terminal domain are labeled. Due to the slight difference among the four molecules in the asymmetric unit, only residues 15–551 were built in this chosen copy.

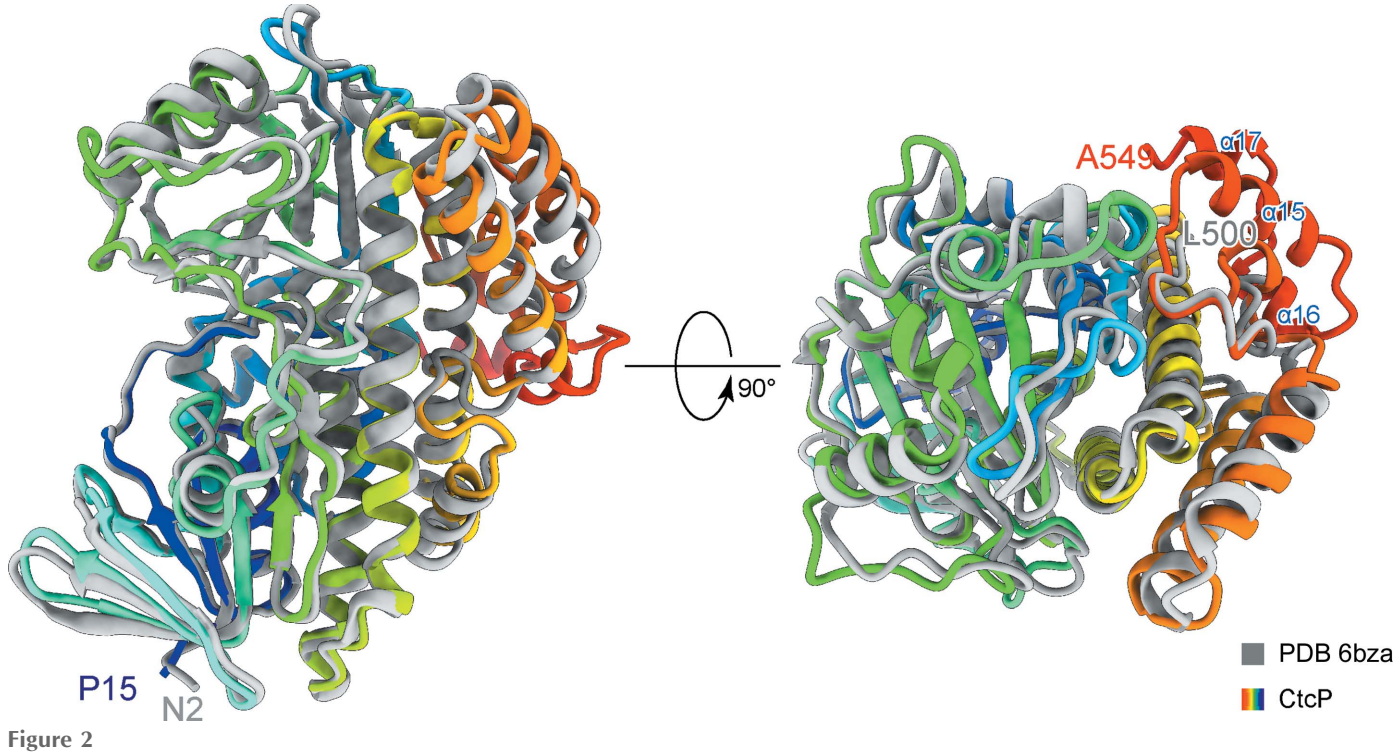


Figure 2
Structural comparison of CtcP with PltM. Two different views of a structural comparison between the *P. protegens* halogenase PltM in complex with phloroglucinol (PDB entry 6bza) and CtcP (this study). The CtcP protein is rainbow-colored, while PltM is colored gray.

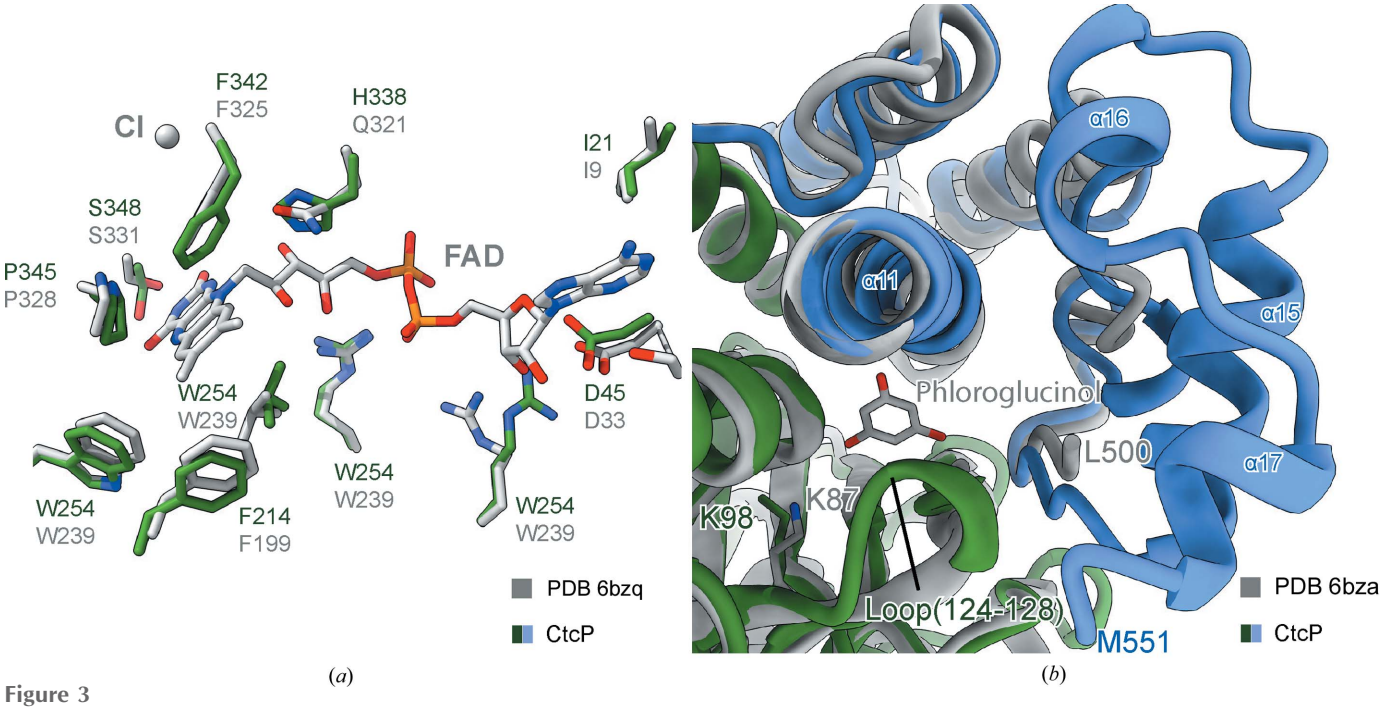


Figure 3
Structural comparison of the FAD- and substrate-binding sites. (*a*) All of the essential residues which coordinate FAD in both CtcP and PltM are shown in stick representation. CtcP is colored green, while PltM is colored gray. The FAD molecule and the Cl^- ion are from the model of PltM. (*b*) Comparison of the substrate-binding site in CtcP and PltM. The enzymatic active-center residues (Lys87 in PltM and Lys98 in CtcP) and the phloroglucinol from the PltM structure are shown in stick representation. CtcP is colored green and blue, while PltM is colored gray.

CtcP and PltM, the C-terminus of CtcP contains three extra α-helices that closely fold together with the rest of the C-terminal domain and form a larger helical bundle fold (Fig. 2).

The bacterial halogenase PltM is capable of using diverse halides for the halogenation of various types of phenolic compounds, including phloroglucinol (Mori *et al.*, 2019). In a previous study, its structure was solved with a FAD ligand and a phloroglucinol substrate. Comparison with the structure of PltM in complex with FAD shows that all of the essential residues that recognize FAD are highly conserved, indicating that CtcP is likely to be capable of binding FAD. Moreover, the conformation of the residue that transfers the halogen atom to the substrate (Lys87 in PltM and Lys98 in CtcP) in the active center is also conserved (Fig. 3).

Interestingly, the structure of the substrate-binding pocket of PltM with phloroglucinol differs from that of *S. aureofaciens* CtcP. Due to the extra three α-helices in the C-terminal domain and a longer internal loop region (124–128) in the N-terminal domain, the substrate-binding pocket of CtcP is wider and thus allows tetracycline to be accommodated, which is larger than phloroglucinol (Fig. 3*b*).

In conclusion, the crystal structure of the halogenase CtcP from *S. aureofaciens*, which is responsible for the final crucial step in CTC biosynthesis, was solved. In future studies, molecular-dynamics simulation or small-molecular docking can be applied to reveal its substrate specificity. Since CTC is a widely used antibiotic both in human medicine and animal feed, the structure can provide structural information to guide enzyme engineering to improve the production of CTC or TC.

Acknowledgements

I would like to thank the staff members of beamline 17U at the SSRF for their assistance with X-ray data collection.

References

Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. & Zwart, P. H. (2010). *Acta Cryst.* D**66**, 213–221.
Bahrami, F., Morris, D. L. & Pourgholami, M. H. (2012). *Mini Rev. Med. Chem.* **12**, 44–52.
Blasiak, L. C. & Drennan, C. L. (2009). *Acc. Chem. Res.* **42**, 147–155.
Brodersen, D. E., Clemons, W. M. Jr, Carter, A. P., Morgan-Warren, R. J., Wimberly, B. T. & Ramakrishnan, V. (2000). *Cell*, **103**, 1143–1154.
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.* D**66**, 12–21.
Dairi, T., Nakano, T., Aisaka, K., Katsumata, R. & Hasegawa, M. (1995). *Biosci. Biotechnol. Biochem.* **59**, 1099–1106.
Dong, C., Flecks, S., Unversucht, S., Haupt, C., van Pée, K.-H. & Naismith, J. H. (2005). *Science*, **309**, 2216–2219.
Duggar, B. M. (1948). *Ann. N. Y. Acad. Sci.* **51**, 177–181.
Dym, O. & Eisenberg, D. (2001). *Protein Sci.* **10**, 1712–1728.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H. & Ferrin, T. E. (2018). *Protein Sci.* **27**, 14–25.
Grossman, T. H. (2016). *Cold Spring Harb. Perspect. Med.* **6**, a025387.
Jenner, L., Starosta, A. L., Terry, D. S., Mikolajka, A., Filonava, L., Yusupov, M., Blanchard, S. C., Wilson, D. N. & Yusupova, G. (2013). *Proc. Natl Acad. Sci. USA*, **110**, 3812–3816.

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., Stockwell, D. H., Terwilliger, T. C., Urzhumtsev, A. G., Videau, L. L., Williams, C. J. & Adams, P. D. (2019). *Acta Cryst.* D**75**, 861–877.
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.
Mori, S., Pang, A. H., Thamban Chandrika, N., Garneau-Tsodikova, S. & Tsodikov, O. V. (2019). *Nat. Commun.* **10**, 1255.
Nakano, T., Miyake, K., Endo, H., Dairi, T., Mizukami, T. & Katsumata, R. (2004). *Biosci. Biotechnol. Biochem.* **68**, 1345–1352.
Nakano, T., Miyake, K., Ikeda, M., Mizukami, T. & Katsumata, R. (2000). *Appl. Environ. Microbiol.* **66**, 1400–1404.
Nelson, M. L. & Levy, S. B. (2011). *Ann. N. Y. Acad. Sci.* **1241**, 17–32.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Phintha, A., Prakinee, K. & Chaiyen, P. (2020). *Enzymes*, **47**, 327–364.
Pioletti, M., Schlünzen, F., Harms, J., Zarivach, R., Glühmann, M., Avila, H., Bashan, A., Bartels, H., Auerbach, T., Jacobi, C., Hartsch, T., Yonath, A. & Franceschi, F. (2001). *EMBO J.* **20**, 1829–1839.
Wang, P., Kim, W., Pickens, L. B., Gao, X. & Tang, Y. (2012). *Angew. Chem. Int. Ed.* **51**, 11136–11140.
Wilson, D. N. (2009). *Crit. Rev. Biochem. Mol. Biol.* **44**, 393–433.
Zhang, X. & Parry, R. J. (2007). *Antimicrob. Agents Chemother.* **51**, 946–957.
Zhu, T., Cheng, X., Liu, Y., Deng, Z. & You, D. (2013). *Metab. Eng.* **19**, 69–78.