structural genomics communications

Acta Crystallographica Section F Structural Biology and Crystallization Communications

ISSN 1744-3091

Chao-Yu Yang,^a‡ Ko-Hsin Chin,^a‡ Chia-Cheng Chou,^{b,c} Andrew H.-J. Wang^{b,c} and Shan-Ho Chou^a*

^aInstitute of Biochemistry, National Chung-Hsing University, Taichung 40227, Taiwan, ^bInstitute of Biological Chemistry, Academia Sinica, Nankang, Taipei, Taiwan, and ^cCore Facility for Protein Crystallography, Academia Sinica, Nankang, Taipei, Taiwan

‡ These authors contributed equally.

Correspondence e-mail: shchou@nchu.edu.tw

Received 27 January 2006 Accepted 3 May 2006

PDB Reference: XC6422, 2fuk.



© 2006 International Union of Crystallography All rights reserved

Structure of XC6422 from *Xanthomonas campestris* at 1.6 Å resolution: a small serine α/β -hydrolase

XC6422 is a conserved hypothetical protein from Xanthomonas campestris pathovar campestris (Xcc), a Gram-negative yellow-pigmented pathogenic bacterium that causes black rot, one of the major worldwide diseases of cruciferous crops. The protein consists of 220 amino acids and its structure has been determined to 1.6 Å resolution using the multi-wavelength anomalous dispersion (MAD) method. Although it has very low sequence identity to protein sequences in the PDB (less than 20%), the determined structure nevertheless shows that it belongs to the superfamily of serine α/β -hydrolases, with an active site that is fully accessible to solvent owing to the absence of a lid domain. Modelling studies with the serine esterase inhibitor E600 indicate that XC6422 adopts a conserved Ser-His-Asp catalytic triad common to this superfamily and has a preformed oxyanion hole for catalytic activation. These structural features suggest that XC6422 is most likely to be a hydrolase active on a soluble ester or a small lipid. An extra strand preceding the first β -strand in the canonical α/β -hydrolase fold leads to extensive subunit interactions between XC6422 monomers, which may explain why XC6422 crystals of good diffraction quality can grow to dimensions of up to 1.5 mm in a few days.

1. Introduction

Structural genomics is a new and rapidly developing field that is under active investigation worldwide (Edwards et al., 2004; Zhang & Kim, 2004). It has great potential for the discovery of novel protein folds and the solution of representative protein structures to obtain a more thorough understanding of biology from a structural perspective. A structural genomics program for the local plant pathogen Xanthomonas campestris pathovar campestris strain 17 (Xcc) has recently been initiated to study its unique regulatory pathway toward pathogenicity. It does not use a cAMP-mediated signal transduction pathway, but instead uses Clp (cAMP-receptor like protein) as a global transcription factor, which has been found to regulate the protein expression of over 100 different cellular functions in Enterbacteriaceae and Bacillus (Ebright, 1993; Kolb et al., 1993). Clp is also possibly the major protein responsible for regulating a wide variety of genes necessary for the synthesis of exopolysaccharides, extracellular enzymes and components of apparatus for type II protein secretion, all of which are collectively required for Xcc pathogenicity (Chen & Tseng, 2006; de Crecy-Lagard et al., 1990).

XC6422 was selected as a target protein for this structural genomics project because no sequence identity higher than 20% could be found in the PDB. In this manuscript, we describe the crystal structure determination of XC6422 using the MAD method. The crystals diffracted to a good resolution of 1.6 Å and the solved structure reveals that the protein belongs to the small serine α/β -hydrolase superfamily (Heikinheimo *et al.*, 1999; Nardini & Dijkstra, 1999) that lacks a lid, leaving the active site fully accessible to solvent. It adopts the conserved Ser-His-Asp catalytic triad (Ser119-His195-Asp166) common to this superfamily, with Ser119 positioned at the very sharp turn between strand β 5 and helix α C, the so-called nucleophile elbow (Heikinheimo et al., 1999; Nardini & Dijkstra, 1999). A structural homology search conducted using the DALI method (Holm & Sander, 1995) revealed more than 20 protein homologues with Z scores larger than 15, including a variety of enzymes with esterase (Arndt et al., 2005; Janda et al., 2004; Kim et al., 1997; Murayama et al., 2005), lipase (Martinez et al., 1994; Nardini et al., 2000; van Pouderoven et al., 2001; Wei et al., 1998) and peptidase (Bartlam et al., 2004; Fülöp et al., 1998; Goettig et al., 2002; Medrano et al., 1998) activities. This is consistent with the notion that the α/β -hydrolase fold family is one of the most versatile and widespread observed to date (Heikinheimo et al., 1999; Nardini & Dijkstra, 1999). Modelling studies with a serine esterase inhibitor E600 indicate that XC6422 adopts a preformed oxyanion-hole conformation and is poised for catalysis without the requirement of further structural rearrangement (Martinez et al., 1994). From these detailed structural studies, we suggest that XC6422 is most likely to be an esterase active on a soluble ester (Arndt et al., 2005; Kim et al., 1997; Martinez et al., 1994; Murayama et al., 2005) or a lipase requiring no interfacial activation to act on a lipid or triacylglycerol substrate (Hjorth et al., 1993; van Pouderoyen et al., 2001; Wei et al., 1998).

2. Materials and methods

2.1. Cloning, expression and purification

Data on the gene cloning, protein expression, purification, crystallization and diffraction of native XC6422 have been reported previously (Yang *et al.*, 2005). SeMet-labelled XC6422 was produced using a non-auxotrophic *Escherichia coli* strain BL21(DE3) as host in the absence of methionine but with ample amounts of SeMet (100 mg l⁻¹). Induction was conducted at 310 K for 4 h by the addition of 0.5 mM IPTG in M9 medium consisting of 1 g NH₄Cl, 3 g KH₂PO₄ and 6 g Na₂HPO₄ supplemented with 20%(w/v) glucose, 0.3%(w/v) MgSO₄ and 10 mg FeSO₄. Incorporation of SeMet was almost complete and mass data showed a major molecular-weight peak of 24 439 Da (data not shown). Purification and crystallization of the SeMet-labelled XC6422 were performed using the protocols established for the native protein (Yang *et al.*, 2005).

2.2. Crystallization

For crystallization, the protein was concentrated to 35.5 mg ml⁻¹ in 20 mM Tris pH 8.0 and 70 mM NaCl using an Amicon Ultra-10 (Millipore). Crystallization screening was performed using the sittingdrop vapour-diffusion method in 96-well plates (Hampton Research Crystal Screen and Crystal Screen 2 kits) at 295 K by mixing 0.5 µl protein solution with 0.5 µl reagent solution. Initial screens were performed using a Gilson C240 crystallization workstation. Parallelepiped-shaped and prism-shaped crystals appeared in 1 d from a reservoir solution comprising 0.1 M HEPES buffer pH 7.5, 2.0 M (NH₄)₂SO₄ and 2% PEG 400. This initial condition was then optimized by varying the concentrations of ammonium sulfate. Crystals suitable for diffraction experiments were grown by mixing 1.5 µl protein solution with 1.5 µl reagent solution and reached maximum dimensions of $2.0 \times 1.5 \times 0.4$ mm after one week. Crystals of SeMet-labelled XC6422 were stabilized in mother liquor containing 22% PEG 2K MME (polyethylene glycol monomethyl ether) plus 20% glycerol before freezing in liquid nitrogen.



Figure 1

Crystal structure of XC6422. All figures were produced with the program *PDBViewer* (Guex & Peitsch, 1997). (*a*) Ribbon diagram of XC6422 colour coded from the N-terminus (blue) to the C-terminus (red). α -Helices A–E and β -strands (β 1– β 8) are annotated. (*b*) Diagram showing the secondary-structure elements of XC6422 superimposed on its primary sequence. The α -helices (green rectangles) and β -strands (red arrows) are indicated. The disordered regions are depicted by a dotted line. The consensus elbow sequence is boxed in red, with the catalytic triad Ser119-His195-Asp166 coloured red and the oxyanion-binding residues Leu47 and Phe119 blue.

2.3. Data collection and structure refinement

Diffraction data sets were collected at 100 K from a single SeMetlabelled crystal using SPring-8 beamline SP12B2 with an ADSC Quantum 4 CCD detector. The diffraction data were indexed and processed using *HKL*2000 (Otwinowski & Minor, 1997). The structure of XC6422 was determined to 1.6 Å resolution using the MAD method. The refinement of Se-atom positions, phase calculation, density modification and building of the initial model were performed using the program *SOLVE/RESOLVE* (Terwilliger & Berendzen, 1999). The model was manually built using the *XtalView/Xfit* package (McRee, 1999). *CNS* (Brünger *et al.*, 1998) was then used for model refinement to a final R_{cryst} of 22.8% and R_{free} of 24.7%. The unit-cell parameters of SeMet-labelled XC6422 crystals are identical to those of native protein and are summarized in Table 1.

3. Results and discussion

3.1. The quality of the model

The refined model of XC6422 contains one single polypeptide chain per asymmetric unit that includes 218 amino acids (residues 3–220) and 207 water molecules. Of all non-glycine residues, 89.1 and 10.4% are in the most favoured and allowed regions of the Ramachandran plot, respectively, as verified with *PROCHECK* (Laskowski *et al.*, 1993). However, the catalytic residue Ser119 adopts an ε -conformation ($\varphi = 54.5$, $\psi = -112.6^{\circ}$) and lies outside the allowed range. It is located at a sharp turn between strand $\beta 5$ and helix αC and its strained conformation is a standard feature of all α/β hydrolases (Heikinheimo *et al.*, 1999; Nardini & Dijkstra, 1999). The other residue that also has a strained conformation is Glu49 ($\varphi = -102.9$, $\psi = -132.9^{\circ}$), which is located at the $\beta 3-\alpha A$ L(46)-S-T-E-G-G(51) loop turn and stabilized by two hydrogen bonds between its side-chain carboxylate and the main-chain carbonyl and side-chain hydroxyl groups of Thr48.

3.2. Overview of the structure

The refined XC6422 model reveals a canonical α/β -hydrolase fold, albeit with one additional strand adjacent to the first β -strand (Fig. 1). However, this additional strand does not lie parallel to nor form any

Table 1

Structural refinement statistics for selenomethionine-substituted XC6422 crystals.

Values in parentheses are for the highest resolution shell.

Resolution range (Å)	24.8-1.60 (1.66-1.60)
Data cutoff $[\sigma(F)]$	0.0
Completeness of reflections used (%)	99.0 (91.3)
No. of reflections used	18909 (2047)
$R_{\rm free}$ test-set size (%)	5.0
$R/R_{\rm free}$ (%)	22.8/24.7 (23.5/25.4)
No. of non-H atoms	
Protein	1680
Solvent	207
R.m.s. deviations	
Bond lengths (Å)	0.005
Bond angles (°)	1.3

specific hydrogen bonds with the first β -strand (Fig. 4). The molecule thus consists of a mostly parallel eight-stranded β -sheet adopting a 12435678 topology and is surrounded by two α -helices on one side and three on the other side. Ser119 is positioned on the nucleophile elbow, as identified by the consensus sequence Sm-X-Nu-X-Sm (where Sm is a small residue, X is any residue and Nu is a nucleophile), corresponding to Gly117-Phe118-Ser119-Phe120-Gly121 in XC6422.

The Ser119-His195-Asp166 catalytic triad is located at the bottom of a pocket that is surrounded by four small loops connecting the $\beta 3-\alpha A$, $\beta 6-\beta 7$, $\beta 7-\alpha D$ and $\beta 8-\alpha E$ secondary-structural elements (Fig. 1). There is therefore no lid domain to hinder the substrate entry from above the pocket.

3.3. Comparison with other α/β -hydrolases

XC6422 is 220 residues long and is relatively small compared with other known α/β -hydrolase fold enzymes that possess a lid domain. A structural homology search conducted with the *DALI* server (Holm & Sander, 1995) returned 20 structures ranging in size from 179 to 726 residues with high Z scores (>15). Despite the low sequence identity, the central β -sheet and the catalytic triad superimpose very well, with r.m.s.d. values ranging from 1.3 to 1.6 Å. Of these structural homologues, many of them are peptidases such as acylamino-acid releasing enzyme (PDB code 1ve6), dipeptidyl peptidase (1n1m), proline iminopeptidases (1mt3 and 1azw) and prolyl oligopeptidases (1qfm



Figure 2

Stereoview of the superimposition of XC6422 (220 residues, purple), 1vkh (253 residues, cyan), 1uxo (184 residues, blue), 1ufo (184 residues, green), 1auo (218 residues, red) and 1i6w (179 residues, black) performed with the program *PDBViewer* (Guex & Peitsch, 1997). β -Strands 1–8 and α -helices A–E are labelled red and green, respectively. Only the catalytic triad residues of XC6422 are shown in ball-and-stick representation.

and 1yr2). These peptidases contain a large insertion domain after strand β 6, which typically serves as a tunnel to prevent large peptides and proteins from accessing the central cavity and undergoing accidental hydrolysis (Bartlam *et al.*, 2004; Fülöp *et al.*, 1998; Goettig *et al.*, 2002). These large enzymes are therefore excluded for comparison since XC6422 has no lid domain and possibly does not belong to this family.

We have therefore restricted our comparison to other small α/β hydrolases of similar size, such as TT1662 (PDB code 1ufo; 234 residues) from *Thermus thermophilus* (Murayama *et al.*, 2005), YDR428C (1vkh; 253 residues) from *Saccharomyces cerevisiae* (Arndt *et al.*, 2005), YdeN (1uxo; 184 residues) from *Bacillus subtilis* (Janda *et al.*, 2004), a *B. subtilis* lipase (1i6w; 179 residues; van Pouderoyen *et al.*, 2001), a *Pseudomonas fluorescens* carboxylesterase (1auo, 218 residues; Kim *et al.*, 1997), a *Fusarium solani* cutinase (1oxm, 213 residues; Martinez *et al.*, 1992) and a *Streptomyces exfoliates* lipase (1jfr; 260 residues; wei *et al.*, 1998). Superimposition of several selected small α/β -hydrolases is shown in Fig. 2. From this comparison, it is clear that the central β -sheet and the surrounding α -helices C, D and E overlap well. Importantly, the loops connecting the $\beta_{3-\alpha}A$, $\beta_{7-\alpha}D$ and $\beta_{8-\alpha}E$ structural elements are also well superimposed which, together with the $\beta_{5-\alpha}C$ loop, collectively accommodate the Ser-His-Asp catalytic triad and the oxyanion pocket (Heikinheimo *et al.*, 1999; Nardini & Dijkstra, 1999). However, considerable variations in the position and length are observed for the loops connecting the $\beta_{4-\alpha}B$ and $\beta_{6-\beta}7$ elements, which may participate in the formation of the substrate-binding site and define the enzyme's specificity. Fig. 2 also reveals that no lid covering the active sites is present in these small α/β -hydrolases. Their active sites are consequently fully accessible to solvent.

3.4. The oxyanion hole and binding pocket

To obtain insight into substrate binding, we have modelled a covalently bound inhibitor in the active site of XC6422. Diethyl-p-nitrophenyl phosphate (E600) is a phosphorylating inhibitor that has been used in studies on serine protease mechanisms (Kraut, 1977). The covalent adduct with the active serine is believed to mimic



Figure 3

Stereoview of the structural details of the active-site pocket in the XC6422–E600 complex. (a) The E600 forming a covalent adduct with the catalytic Ser119 residue is shown in the centre region. The hydrogen-bonding network in the active site is connected by dotted green lines. (b) The E600 inhibitor is shown as a CPK model with the residues that possibly participate in substrate binding shown in ball-and-stick representation. The C atoms are coloured grey, O atoms red, N atoms blue and P atoms orange.

structural genomics communications

the tetrahedral intermediate of the reaction. We have superimposed the XC6422 structure on a cutinase crystal structure containing the E600 adduct (Martinez et al., 1994) and merged the E600 coordinates with the XC6422 coordinates to obtain the initial XC6422-E600 complex. This initial complex structure was further energy-minimized using the DS-Modeler program (Accelrys Inc.) to obtain the final complex structure, which does not change significantly from the initial complex structure (with an overall r.m.s.d. value of only 0.2 Å). This indicates that XC6422 requires no structural rearrangement for catalysis. The active-site pocket with the E600 adduct of the final XC6422 is shown Fig. 3(a). It is located at the bottom of a small cleft between four loops consisting of residues 46-49, 143-147, 166-169 and 194-197. The catalytic triad is locked in position by a hydrogenbonding network: Asp166 $O^{\delta 2}$ forms a hydrogen bond with His195 N^{ϵ_1} and Asp166 O^{δ_1} form two hydrogen bonds with the mainchain amides of Ile168 and Val169. The Ser119 hydroxyl group does not form a hydrogen bond with His195 $N^{\epsilon 2}$ in the adduct complex; it is now shifted away from His195 N^{ϵ_2} and forms a weak hydrogen bond with the O2 atom of E600 (3.26 Å). The main-chain N atoms of Leu46 and of Phe120, the residue adjacent to the catalytic serine, are within hydrogen-bonding distance of the E600 O3 atom (2.72 and 2.52 Å, respectively). The inhibited XC6422 structure thus made it possible to locate the oxyanion hole.

We also found many hydrophobic residues (Ile168, Val169, Pro143, Trp147, Phe196 and Leu46) above the active-site crevice; these residues possibly form the binding site for an ester or a lipid substrate. Interestingly, although one of the ethyl groups of the diethylphosphate inhibitor was found to fit into the crevice very well (Fig. 3*b*), there seems to be room for a longer carbon chain in the upper part of the crevice. The correct substrate for XC6422 remains to be established in a further study.

3.5. Crystal contacts

XC6422 formed high-quality crystals that grew to 1.5 mm in a few days and diffracted to a good resolution of 1.6 Å. Detailed subunitinteraction studies between XC6422 monomers revealed that the extra strand preceding the first β -strand of the canonical α/β hydrolase fold made a considerable contribution to the crystal packing. As shown in Fig. 4, the XC6422 crystal packs along a 2 axis and is mediated by extensive intermolecular hydrogen-bonding interactions, including Ser2 H^N...Leu187 C=O, Leu5 H^N...Gln184 C=O, Leu5 C=O···Gln184 side-chain side-chain amide, Ala12 C= $O \cdots Gln184 H^N$, Ala12 H^N····Glu182 C==O and Thr8 H^N...Glu179 C=O. Altogether, excluding possible hydrophobic contributions, a single Gln184 residue contributes half (three)



Figure 4

XC6422 crystal structure. (a) Packing of XC6422 along a 2 axis mediated by the extra strand prior to the β 1 strand in the canonical α/β -hydrolase fold. (b) Close-up of the intermolecular hydrogen-bonding network that stabilizes the crystal packing.

In conclusion, XC6422 was chosen as a target protein in our Xcc structural genomics project since no sequence identity larger than 20% could be found in the PDB. After its structure had been determined and a structural homology search performed, it was found that it belongs to a subfamily of the α/β -hydrolase fold that does not possess a lid. The active-site residues and the oxyanion hole were further identified with the aid of a XC6422–inhibitor modelling study and the hydrophobic residues around the active site were also located. Since the active-site pocket is open and fully accessible to solvent with no lid domain to block the entry of substrate, we tentatively conclude that XC6422 is a small esterase active on a soluble ester without the need for structural rearrangement or a lipase active on a lipid or triacylglycerol substrate requiring no interfacial activation.

This work is supported by the Academic Excellence Pursuit grant from the Ministry of Education and by the National Science Council, Taiwan to S-HC. We thank the Core Facilities for Protein X-ray Crystallography in Academia Sinica, Taiwan, in the National Synchrotron Radiation Research Center, Taiwan and in the SPring-8 Synchrotron facility, Japan for assistance during X-ray data collection. The National Synchrotron Radiation Research Center is a user facility supported by the National Science Council, Taiwan and the Protein Crystallography Facility is supported by the National Research Program for Genomic Medicine, Taiwan.

References

- Arndt, J. W., Schwarzenbacher, R., Page, R., Abdubek, P., Ambing, E., Biorac, T., Canaves, J. M., Chiu, H.-J., Dai, X. & Wilson, I. A. (2005). *Proteins*, 58, 755–758.
- Bartlam, M., Wang, G., Yang, H., Gao, S., Feng, Y. & Rao, Z. (2004). *Structure*, **12**, 1481–1488.
- Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Acta Cryst. D54, 905–921.

Chen, M.-C. & Tseng, Y.-H. (2006). Submitted.

- Crecy-Lagard, V. de, Glaser, P., Lejeune, O., Sismeiro, C. E., Barber, C. E., Daniels, M. J. & Danchin, A. (1990). *J. Bacteriol.* **172**, 5877–5883.
- Ebright, R. H. (1993) Mol. Microbiol. 8, 797-802.
- Edwards, A. F., Yee, A., Savchenko, A., Edwards, A. M. & Arrowsmith, C. H. (2004). *Curr. Opin. Chem. Biol.* **8**, 42–48.
- Fülöp, V., Bocskei, Z. & Polgar, L. (1998). Cell, 94, 161-170.
- Goettig, P., Groll, M., Kim, J.-S., Huber, R. & Brandstetter, H. (2002). EMBO J. 21, 5343–5352.
- Guex, N. & Peitsch, M. C. (1997). Electrophoresis, 18, 2714-2723.
- Heikinheimo, P., Goldman, A., Jeffries, C. & Ollis, D. L. (1999). Structure, 7, R141–R146.
- Hjorth, A., Carriere, F., Cudrey, C., Woldike, H., Boel, E., Lawson, D. M., Ferrato, F., Cambillau, C., Dodson, G. G., Thim, L. & Verger, R. (1993). *Biochemistry*, **32**, 4702–4707.
- Holm, L. & Sander, C. (1995). Trends Biochem. Sci. 20, 478-480.
- Janda, I., Devedjiev, Y., Cooper, D., Chruszcz, M., Derewenda, U., Gabrys, A., Minor, W., Joachimiak, A. & Derewenda, Z. S. (2004). Acta Cryst. D60, 1101–1107.
- Kim, K. K., Song, H. K., Shin, D. H., Hwang, K. Y., Choe, S., Yoo, O. J. & Suh, S. W. (1997). *Structure*, 5, 1571–1584.
- Kolb, A., Busby, S., Buc, H., Garges, S. & Adhya, S. (1993). Annu. Rev. Biochem. 62, 749–795.
- Kraut, J. (1977). Annu. Rev. Biochem. 46, 331-358.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). J. Appl. Cryst. 26, 283–291.
- McRee, D. E. (1999). J. Struct. Biol. 125, 156-165.
- Martinez, C., De Geus, P., Lauwereys, M., Matthyssens, G. & Cambillau, C. (1992). Nature (London), 356, 615–618.
- Martinez, C., Nicolas, A., van Tilbeurgh, H., Egloff, M.-P., Cudrey, C., Verger, R. & Cambillau, C. (1994). *Biochemistry*, **33**, 83–89.
- Medrano, F. J., Alonso, J., Garcia, J. L., Romero, A., Bode, W. & Gomis-Ruth, F. X. (1998). *EMBO J.* **17**, 1–9.
- Murayama, K., Shirouzu, M., Terada, T., Kuramitsu, S. & Yokoyama, S. (2005). Proteins, 58, 982–984.
- Nardini, M. & Dijkstra, B. W. (1999). Curr. Opin. Struct. Biol. 9, 732-737.
- Nardini, M., Lang, D. A., Liebeton, K., Jaeger, K. L. & Dijkstra, B. W. (2000). J. Biol. Chem. 275, 31219–31225.
- Otwinowski, Z. & Minor, W. (1997). Methods Enzymol. 276, 307-326.
- Pouderoyen, G. van, Eggert, T., Jaeger, K.-E. & Dijkstra, B. W. (2001). J. Mol. Biol. 309, 215–226.
- Terwilliger, T. C. & Berendzen, J. (1999). Acta Cryst. D55, 849-861.
- Wei, Y., Swenson, L., Castro, C., Derewenda, U., Minor, W., Arai, H., Aoki, J., Inoue, K., Servin-Gonzalez, L. & Derewenda, Z. S. (1998). *Structure*, 6, 511–519.
- Yang, C.-Y., Chin, K.-H., Chou, C.-C., Shr, H.-L., Gao, F. P., Lyu, P.-C., Wang, A. H.-J. & Chou, S.-H. (2005). Acta Cryst. F61, 703–705.
- Zhang, C. & Kim, S.-H. (2004). Curr. Opin. Chem. Biol. 7, 28-32.