Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic analysis of thioesterase I from *Escherichia coli*

Yu-Chih Lo,[a,c] Ya-Lin Lee,[b] Jei-Fu Shaw[b] and Yen-Chywan Liaw[a]*

[a]Institute of Molecular Biology, Academia Sinica, Taipei 11529, Taiwan, [b]Institute of Botany, Academia Sinica, Taipei 11529, Taiwan, and [c]Graduate Institute of Life Sciences, National Defense Medical Center, Taipei 202, Taiwan

Correspondence e-mail: mbycliaw@ccvax.sinica.edu.tw

Received 29 November 1999
Accepted 16 March 2000

The *Escherichia coli* thioesterase I specifically catalyzes the deacylation of fatty acyl–CoA thioesters, especially those with long acyl groups (C_{12}–C_{18}). Single crystals of thioesterase I (E.C. 3.1.2.2) from *E. coli* have been obtained using methoxypolyethylene glycol 5000 (PEG–MME 5K) as a precipitant at room temperature in 21 d. The crystals belong to the tetragonal space group $P4_12_12$ or its enantiomorph $P4_32_12$, with unit-cell parameters $a = b = 50.85$ (7), $c = 171.5$ (1) Å. The crystals diffract to beyond 2.4 Å resolution. There is one molecule of molecular weight 20.5 kDa in the asymmetric unit, with a solvent content of 55%.

1. Introduction

The *E. coli* thioesterase I (E.C. 3.1.2.2) belongs to a discrete group of widely distributed esterases (ester hydrolases; E.C. 3.1), which hydrolyze ester bonds (Bronner & Bloch, 1972; Bjorkling *et al.*, 1991). It specifically catalyzes the hydrolytic cleavage of fatty acyl–coenzyme A (CoA) thioesters and also cleaves fatty acyl–acyl carrier protein (ACP) thioesters at 10^3 to 10^4 lower rates than acyl–CoA esters of the same length (Spencer *et al.*, 1978). Thioesterase I, unlike thioesterase II [which is a tetramer of 127.4 kDa, has a broader substrate specificity (C_6–C_{18}) and is insensitive to serine-esterase inhibitors (Barnes & Walkil, 1968; Barnes, 1975; Naggert *et al.*, 1991)], specifically catalyzes the deacylation of long acyl groups (C_{12}–C_{18}, especially the palmitoyl group; Bronner & Bloch, 1972; Cho & Cronan, 1993). In addition, this enzyme has been demonstrated to be a multifunctional enzyme possessing protease, lipase and arylesterase functions (Ichihara *et al.*, 1993; Cho & Cronan, 1994; Lee *et al.*, 1997). It possesses stereoselectivity for protease substrates. The catalytic efficiency (k_{cat}/K_m) of the enzyme for an amino-acid derived substrate L-NBPNPE was 23 times higher than that for the enantiomer D-NBPNPE, which was considered to be useful for kinetic resolution of racemic mixtures of chemicals in industry. It can also hydrolyze several short acyl-chain aromatic esters and short acyl-chain triacylglycerols (Lee *et al.*, 1997). Thioesterase I is therefore expected to have industrial potential in processes such as stereospecific synthesis and hydrolysis of esters (Aragozzini *et al.*, 1992; Cambillau *et al.*, 1996).

Thioesterase I encoded by the *tesA* gene consists of 183 amino-acid residues and has a molecular mass of 20.5 kDa (Cho & Cronan, 1993). According to the sequence alignment and biochemical assay, thioesterase I has been shown to contain a novel Gly-Asp-Ser-Leu-Ser motif, which is located close to the N-terminal region and differs from the Gly-Xaa-Ser-Xaa-Gly motif found in the middle regions of other typical lipases and esterases (Winkler *et al.*, 1990; Brzozowski *et al.*, 1991; Schrag *et al.*, 1991; Sussman *et al.*, 1991; Martinez *et al.*, 1992), suggesting a unique three-dimensional folding of thioesterase. With this Gly-Asp-Ser-Leu-Ser motif, thioesterase I and seven other proteins from different species form a new subfamily of lipolytic enzymes (Upton & Buckley, 1995). The amino-acid sequences of these proteins share the same five consensus-sequence blocks arranged in the same order. At present, no X-ray crystal structure is available for any member of this family of proteins. Therefore, crystallographic studies of this subfamily of lipolytic enzymes have become an interesting structural biology target. In this paper, we report the crystallization and preliminary studies of this novel *E. coli* thioesterase I.

2. Materials and methods

His-tagged thioesterase I was purified by the protocol described previously (Lee *et al.*, 2000). The purified protein in 10 m*M* sodium phosphate buffer (pH 7.0) was concentrated to 76.8 mg ml^{-1}. Crystallization experiments were performed at room temperature using a modification of the hanging-drop vapour-diffusion technique (Wlodawer & Hodgson, 1975). The initial crystallization conditions were screened using the sparse-matrix sampling method (Jancarik & Kim, 1991), mixing 2 µl of the protein solution with an equal volume of reservoir solution. Small diamond-shaped crystalsz were obtained under conditions containing 0.2 *M* ammonium sulfate and 30%(*w*/*w*) PEG–MME 5K in 0.1 *M* 2-[*N*-morpholino]ethanesulfonic acid (MES)

Table 1
Crystallographic data and data-collection statistics.

Standard deviations are shown in parentheses.

Temperature (K)	113
Space group	$P4_32_12$ or $P4_12_12$
Unit-cell parameters (Å)	
$a = b$	50.85 (0.07)
c	171.45 (0.12)
Mosaicity (°)	0.52
Total observations	97547
Unique reflections	9601
Multiplicity	10.2
Complete resolution range (Å)	40.0–2.4
Outermost shell resolution (Å)	2.44–2.4
Overall completeness (%)	98.7
Outermost shell completeness (%)	100.0
Overall $I/\sigma(I)$	15.4
Outermost shell $I/\sigma(I)$	4.5
Overall $I/\sigma(I) > 3$ (%)	78.5
Outermost shell $I/\sigma(I) > 3$ (%)	33.6
Overall $R_{\rm merge}$ (%)	5.2
Outermost shell $R_{\rm merge}$ (%)	36.1

buffer pH 6.5. To optimize the crystallization conditions, the concentration of PEG–MME 5K was decreased to 27%(*w*/*w*); the crystal appeared after 7 d and grew slowly to dimensions of 0.9 × 0.5 × 0.35 mm after 21 d (see Fig. 1).

Diffraction data were collected on an R-AXIS IV image plate using Cu $K\alpha$ radiation generated by a Rigaku RU-300 rotating-anode generator operating at a voltage of 50 kV and a current of 100 mA. The thioesterase I crystal was soaked in mother liquor containing 15% glycerol as a cryo-protectant, was retrieved with a 0.7–1.0 mm nylon loop and was flash-frozen in the nitrogen stream from an X-stream cooler operated at a temperature of 113 K. The crystal-to-detector distance was set at 145 mm and diffraction images were recorded with 2° oscillation per image and an exposure time of 25 min per frame. All data were processed using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

A homology sequence search with Smith–Waterman algorithm (Smith & Waterman, 1981) was performed using *GeneWeb* with the Bio-XLP accelerator (Compugen Inc.) to locate similar sequences over the whole sequence database. The sequence similarity was calculated by using the sequence-alignment analysis program of the *GCG* package.

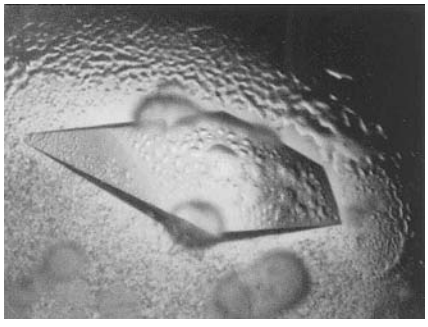

Figure 1
The crystal of thioesterase I from *E. coli* has a bipyramidal shape. The dimensions of the largest crystal are 0.9 × 0.5 × 0.35 mm.

3. Results

According to the results deduced from *DENZO* indexing program, the crystals of thioesterase I belong to the primitive tetragonal lattice. Examination of systematic absences allowed the identification of the space group $P4_12_12$ or its enantiomorph $P4_32_12$ and the unit-cell parameters $a = b = 50.85$ (7), $c = 171.5$ (1) Å. A solvent content of 55% and an acceptable crystal packing density V_m of 2.73 Å^3 Da^{-1} were calculated using the method of Matthews (1968), assuming the thioesterase I crystals to contain one molecule per asymmetric unit. Although this crystal diffracted to better than 2.2 Å resolution, it decayed severely in the last few frames. A total of 97 547 reflections were collected, 9601 reflections being unique, which represents a completeness of 98.7% at 2.4 Å resolution. The $R_{\rm merge}$ for the reduced data set was 5.2% in this resolution range. Data statistics are given in Table 1.

An amino-acid sequence analysis indicates that thioesterase I has the Gly-Asp-Ser-Leu-Ser sequence homologous to the consensus Gly-Xaa-Ser-Xaa-Gly motif of the serine proteases which function *via* a Ser-Asp-His catalytic triad. A sequence comparison of the *E. coli* thioesterase I with other available serine-protease crystal structures shows very low sequence homology, which does not justify molecular-replacement trials. Further sequence-similarity searches of the SWISS-PROT sequence database and EMBL database revealed 51.7% homology between the *E. coli* thioesterase I and the *Vibrio mimicus* arylesterase (Shaw *et al.*, 1994). Both enzymes belong to the new subfamily of lipolytic enzymes possessing the Gly-Asp-Ser-Leu-Ser motif and the first serine residue of this motif was shown to be the one belonging to the catalytic triad (Cho & Cronan, 1993). At present, no X-ray crystal structure has been determined for this family of enzymes. Crystallographic three-dimensional structure studies of thioesterase I are important to elucidate the structure differences between our thioesterase and other related esterases and help us to understand the substrate specificity and stereoselectivity of these enzymes. Structural information from these proteins would be of great benefit for protein-engineering applications such as the development of engineered enzymes for stereospecific synthesis and the hydrolysis of some specific esters. A search for heavy-atom derivatives for a structure solution with multiple isomorphous replacement is currently under way.

This work was supported by a grant from Academia Sinica to Y.-C. Liaw and a grant (NSC 85-2321-B001-017-A18) from the National Science Council, Republic of China to J-FS. Y-CL is a National Defense Medical Center graduate student fellow.

References

Aragozzini, F., Valenti, M., Santaniello, E., Ferraboschi, P. & Grisenti, P. (1992). *Biocatalysis*, **5**, 325–332.

Barnes, E. M. (1975). *Methods Enzymol.* **35**, 102–109.

Barnes, E. M. & Walkil, S. J. (1968). *J. Biol. Chem.* **243**, 2955–2962.

Bjorkling, F., Godtfredsen, S. E. & Kirk, O. (1991). *Trends Biotechnol.* **9**, 360–363.

Bronner, W. M. & Bloch, K. (1972). *J. Biol. Chem.* **247**, 3123–3133.

Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Huge-Jensen, B., Patkar, S. A. & Thim, L. (1991). *Nature (London)*, **351**, 491–494.

Cambillau, C., Longhi, S., Nicolas, A. & Matinez, C. (1996). *Curr. Opin. Struct. Biol.* **6**, 449–455.

Cho, H. & Cronan, J. E. Jr (1993). *J. Biol. Chem.* **268**, 9238–9245.

Cho, H. & Cronan, J. E. Jr (1994). *J. Bacteriol.* **176**, 1793–1795.

Ichihara, S., Matsubara, Y., Kato, C., Akasaka, K. & Mizushima, S. (1993). *J. Bacteriol.* **175**, 1032–1037.

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.

Lee, Y.-L., Chen, J.-C. & Shaw, J.-F. (1997). *Biochem. Biophys. Res. Commun.* **231**, 452–456.

Lee, Y.-L., Su, M.-S., Huang, T.-H. & Shaw, J.-F. (2000). In the press.

Martinez, C., De Geus, P., Lauwereys, M., Matthyssens, G. & Cambillau, C. (1992). *Nature (London)*, **356**, 615–618.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

Naggert, J., Narashimhan, M. L., Deveaux, L., Cho, H., Randhawa, Z. I., Cronan, J. E. Jr, Green, B. N. & Smith, S. (1991). *J. Biol. Chem.* **266**, 11044–11050.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Schrag, J. D., Li, Y., Wu, S. & Cygler, M. (1991). *Nature (London)*, **351**, 761–764.

Shaw, J.-F., Chang, R.-C., Yen, Y.-T., Wang, F.-F. & Wang, Y.-J. (1994). *Biochem. J.* **298**, 675–680.

Smith, T. F. & Waterman, M. S. (1981). *Adv. Appl. Math.* **2**, 482–489.

Spencer, A. K., Greenapsn, A. D. & Cronan, J. E. Jr (1978). *J. Biol. Chem.* **253**, 5922–5926.

Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L. & Silman, I. (1991). *Science*, **253**, 872–879.

Upton, C. & Buckley, J. T. (1995). *Trends Biochem. Sci.* **20**, 178–179.

Winkler, F. K., D'Arcy, A. & Hunziker, W. (1990). *Nature (London)*, **343**, 771–774.

Wlodawer, A. & Hodgson, K. O. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 398–399.