Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression, purification, crystallization and initial crystallographic analysis of the lysine-biosynthesis LysX protein from *Thermus thermophilus* HB8

Marina N. Vassylyeva,[a,b]† Hiroaki Sakai,[a]† Takanori Matsuura,[a] Shun-ichi Sekine,[b,c] Makoto Nishiyama,[d] Takaho Terada,[a,b] Mikako Shirouzu,[a,b] Seiki Kuramitsu,[a,b,e] Dmitry G. Vassylyev[b,c]* and Shigeyuki Yokoyama[a,b,c,f]*

[a]Genomic Sciences Center, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan, [b]Structurome Project, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [c]Cellular Signaling Laboratory, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [d]Biotechnology Research Center, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan, [e]Department of Biology, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, and [f]Department of Biophysics and Biochemistry, Graduate School of Science, University of Tokyo, Bunkyo-Ku, Tokyo 113-0033, Japan

† These authors contributed equally to this work.

Correspondence e-mail: dmitry@yumiyoshi.harima.riken.go.jp, yokoyama@biochem.s.u-tokyo.ac.jp

Received 28 February 2003
Accepted 1 July 2003

The gene encoding LysX, an essential component of the lysine-biosynthesis pathway in *Thermus thermophilus* (molecular weight $\simeq$ 31 000 Da), was cloned and expressed and the purified protein was crystallized by the hanging-drop vapour-diffusion technique in two different space groups, *C*2 (unit-cell parameters a = 124.7, b = 51.4, c = 103.6 Å, β = 122.8°) and *R*3 ($a = b$ = 122.6, c = 97.6 Å). Crystals improved by macroseeding diffracted to beyond 2.3 and 3 Å resolution for the *C*2 and *R*3 crystal forms, respectively. Complete diffraction data sets were collected for the *C*2 and *R*3 crystal forms at 2.5 and 3.1 Å resolution, respectively. Crystals of selenomethionine-containing LysX protein were obtained by cross-microseeding, using the native microcrystals as a seed. Structure determination is now in progress.

1. Introduction

There are two completely independent and unrelated pathways for lysine biosynthesis in prokaryotes and eukaryotes: the diaminopimelate and α-aminoadipate pathways (Broquist, 1971; Eggeling, 1994; Kosuge & Hoshino, 1998; Kobashi *et al.*, 1999). Most bacteria utilize the diaminopimelate pathway, whereas the α-aminoadipate pathway has only been found in yeast (Broquist, 1971) and fungi (Vogel, 1964; Irvin & Bhattacharjee, 1998). The α-aminoadipate pathway begins with the synthesis of homocitrate from 2-oxoglutarate and acetyl-CoA by homocitrate synthase and proceeds *via* α-aminoadipate and saccharopine to lysine. Recently, a thermophilic bacterium, *Thermus thermophilus*, was shown to synthesize lysine through a modified α-aminoadipate pathway which is similar to that of fungi and yeast (Kobashi *et al.*, 1999; Nishida *et al.*, 1999; Kosuge & Hoshino, 1998). The lysine-biosynthesis genes found in *T. thermophilus* are also present in the hyperthermophilic and anaerobic archaeon *Pyrococcus horikoshii*, in which they form a very similar gene cluster to that of *T. thermophilus* (Nishida *et al.*, 1999).

T. thermophilus contains the *lysX* gene within the gene cluster for lysine biosynthesis. The *lysX* gene encodes a protein with a molecular weight of 31 kDa. The LysX protein is likely to be an essential component of the lysine-biosynthesis pathway, as it has been proposed to catalyze the modification of α-aminoadipate to N^2-acetyl-α-aminoadipate (Nishida *et al.*, 1999). However, the precise catalytic mechanism has not yet been elucidated. On the other hand, the *T. thermophilus* LysX protein shares homology with the *Escherichia coli* ribosomal protein S6 modification protein RimK, with a sequence identity of 28%. According to PSI-BLAST sequence analysis, the *E. coli* RimK protein belongs to the ATP-grasp fold family of proteins, which catalyze the ligation of a carboxylate-containing molecule to an amino or thiol group-containing molecule in an ATP-dependent manner (Galperin & Koonin, 1997). However, in *E. coli* a lysine-biosynthesis α-aminoadipate pathway similar to that of *T. thermophilus* has not been found.

To determine the three-dimensional structure of LysX protein, which would shed light on its enzymatic mechanism, its substrate specificity and its evolutional relationship to homologous proteins utilizing different substrates, we have cloned and expressed the *lysX* gene from *T. thermophilus* HB8. The purified LysX protein has been crystallized.

2. Experimental procedures and results

2.1. Cloning, expression and purification

The *lysX* gene from *T. thermophilus* HB8 was amplified by PCR from *T. thermophilus* HB8 genomic DNA. The PCR product was cloned into the pET26b(+) expression vector (Novagen).

For overexpression of the native protein, the plasmid was transformed into *E. coli* strain BL21(DE3). Cells were harvested and sonicated in 20 m*M* Tris–HCl buffer pH 8.0 containing 300 m*M* NaCl. The cell lysate was heated at 343 K for 20 min to denature the *E. coli* proteins. The heat-treated solution was centrifuged at 15 000*g* for 20 min and the supernatant was dialyzed against 20 m*M* Tris–HCl buffer pH 8.0 containing 1 m*M* DTT. Q-Sepharose Fast Flow, Resource Q, Mono Q

and Superdex 75 columns (Amersham Bioscience) were subsequently employed to purify LysX from *T. thermophilus* HB8 to homogeneity. The purified protein was concentrated to approximately 12 mg ml^{-1} using a Centricon YM-10 filter (Millipore). The estimated yield was 13.5 mg of purified protein per litre of culture. The selenomethionine-substituted (SeMet) LysX protein was synthesized using the *E. coli* cell-free protein synthesis system (Kigawa *et al.*, 1999) and was purified as described above.

2.2. Crystallization and data collection

Hampton Research Crystal Screens (Jancarik & Kim, 1991) were used to determine the initial crystallization conditions for the LysX protein. Crystallization was carried out by the sitting-drop vapour-diffusion technique at 293 K; the Crystal Screen precipitant solutions were diluted sixfold with water in the initial screening. 2 µl aliquots of these diluted precipitant solutions were added to 2 µl drops containing the protein solution at a protein concentration of 10 mg ml^{-1}. Occasionally, after two weeks of equilibration, a single large plate-like crystal (dimensions 0.15 × 0.15 × 0.5 mm) was found to have grown in Crystal Screen I solution No. 10. We were not able to reproduce this crystal type; therefore, the only single mono-crystal was split into small parts and these tiny crystals were subjected to macroseeding using drops prepared under the same conditions. After a few days of equilibration, the small crystals grew to typical dimensions of 0.25 × 0.25 × 0.4 mm (Fig. 1*a*) and diffracted to beyond 2.3 Å resolution at an in-house X-ray generator. As revealed by indexing the diffraction data and by the $R_{\rm merge}$ statistics provided by the *HKL*2000 program package (Otwinowski & Minor, 1997) (Table 1), the crystals belong to space group *C*2, with unit-cell parameters a = 124.7, b = 51.4, c = 103.6 Å, β = 122.8°. The final crystallization solution was composed of 5% PEG 4000, 17 m*M* sodium acetate pH 4.6 and 0.35 m*M* ammonium acetate. At the same time, small hexagonal crystals (dimensions of 0.1 × 0.1 × 0.1 mm) were grown from Crystal Screen I solution No. 3. The crystallization conditions were then adjusted, using additives to the initial solution and variation of the concentration of solution No. 3. The small crystals were subjected to macroseeding using the optimized solution and could be enlarged in size

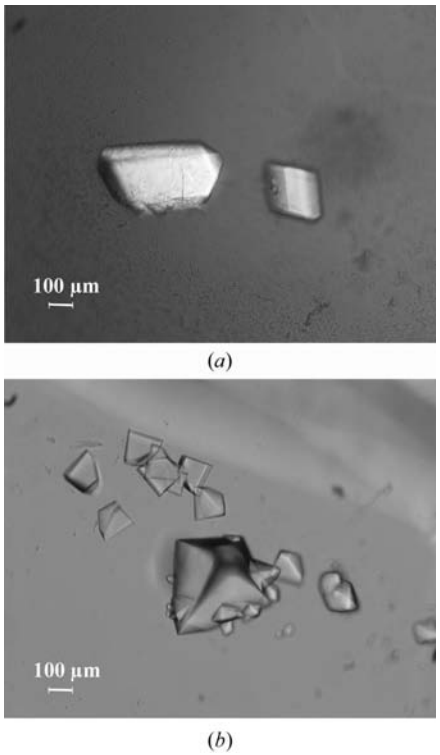

Figure 1
Crystals of the LysX protein. (*a*) *C*2 crystal form. (*b*) *R*3 crystal form.

up to 0.25 × 0.25 × 0.2 mm (Fig. 1*b*). As revealed by indexing the diffraction data and by the $R_{\rm merge}$ statistics provided by the *HKL*2000 program package (Otwinowski & Minor, 1997) (Table 1), the crystals belong to space group *R*3, with unit-cell parameters $a = b$ = 122.6, c = 97.6 Å. The final reservoir solution used for growing the *R*3 form crystals contained 33 m*M* ammonium phosphate, 2.5% PEG 4000, 8 m*M* sodium acetate pH 4.6, 16 m*M* ammonium acetate and 5% glycerol. However, the diffraction quality of these crystals was substantially poorer than that of the *C*2 crystal form and they diffracted to only 3 Å resolution. Complete diffraction data sets for the *C*2 and *R*3 crystal forms were collected at 2.5 and 3.1 Å resolution, respectively, at the SPring-8 BL45 beamline using a Rigaku R-AXIS V imaging-plate detector or at the BL44XU beamline using an Oxford CCD detector. The data were collected at 100 K using the mother-liquor solution with 35% glycerol as a cryoprotectant. The data were processed using the *HKL*2000 program (Otwinowski & Minor, 1997) (Table 1). To phase the diffraction data, the LysX protein containing selenomethionine (SeMet) was prepared for crystallization. As for the native protein, we were unable to reproduce the *C*2-form crystals. Therefore, the SeMet protein was subjected to cross-microseeding, in which the native microcrystals of the *C*2 crystal form were seeded in drops containing the SeMet protein. The drops were prepared under the same conditions (for the *C*2 crystal form) as for the native protein. This technique produced SeMet-containing crystals of the *C*2 crystal form with diffraction quality almost identical to that of the native crystals. The high quality of the SeMet LysX crystals will probably allow us to collect multiple anomalous dispersion data at the absorption edge for the Se atoms, which should be precise enough to phase the diffraction data at an atomic resolution. The structure determination is now in progress.

Table 1
Data-collection statistics for the LysX protein.

Values in parentheses indicate data in the highest resolution shell.

Data collection	Form 1	Form 2
Space group	*C*2	*R*3
Wavelength used (Å)	1.02	0.95
Resolution range (Å)	40.0–2.5 (2.59–2.5)	200–3.1 (3.21–3.1)
Total observations	80369	29990
Unique reflections	19195	9136
Redundancy	4.2 (4.2)	3.3 (3.1)
Completeness (%)	99.3 (100)	97.5 (98.4)
Mean $I/\sigma(I)$	15.6 (15.5)	15.6 (6.4)
$R_{\rm merge}$† (%)	4.0 (10.5)	5.6 (18.1)

† $R_{\rm merge} = \sum_{hkl}\sum_j |I_j(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_j \langle I(hkl)\rangle$, where $I_j(hkl)$ and $\langle I(hkl)\rangle$ are the intensity of measurement j and the mean intensity for the reflection with indices hkl, respectively.

References

Broquist, H. P. (1971). *Methods Enzymol.* **17**, 112–129.
Eggeling, L. (1994). *Amino Acids*, **6**, 261–272.
Galperin, M. Y. & Koonin, E. V. (1997). *Protein Sci.* **6**, 2639–2643.
Irvin, S. D. & Bhattacharjee, J. K. (1998). *J. Mol. Evol.* **46**, 401–408.
Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Kigawa, T., Yabuki, T., Yoshida, Y., Tsutsui, M., Ito, Y., Shibata, T. & Yokoyama, S. (1999). *FEBS Lett.* **442**, 15–19.
Kobashi, N., Nishiyama, M. & Tanokura, M. (1999). *J. Bacteriol.* **181**, 1713–1718.
Kosuge, T. & Hoshino, T. (1998). *FEMS Microbiol. Lett.* **169**, 361–367.
Nishida, H., Nishiyama, M., Kobashi, N., Kosuge, T., Hoshino, T. & Yamane, H. (1999). *Genome Res.* **9**, 1175–1183.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Vogel, H. J. (1964). *Am. Nat.* **98**, 435–446.