Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray crystallographic analysis of γ-carboxymucolactone decarboxylase from *Sulfolobus solfataricus*

Ho Yeon Lee and Jin Kuk Yang*

Department of Chemistry, College of Natural Sciences, Soongsil University, Seoul 156-743, Republic of Korea

Correspondence e-mail: jinkukyang@ssu.ac.kr

Received 26 August 2009
Accepted 16 October 2009

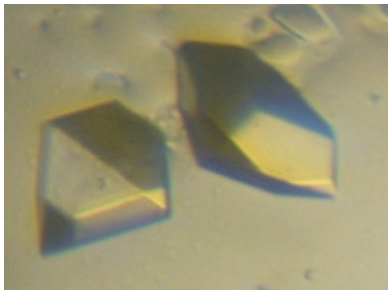
γ-Carboxymucolactone decarboxylase (γ-CMD; EC 4.1.1.44) catalyzes the conversion of γ-carboxymucolactone to β-ketoadipate enol-lactone in the β-ketoadipate pathway, which is a key part of the degradation process of aromatic compounds in bacteria and in some eukaryotes such as fungi and yeast. γ-CMD from the thermophilic archaeon *Sulfolobus solfataricus* (Ss γ-CMD) is encoded by the *pcaC* gene and is composed of 139 amino-acid residues with a molecular mass of 15 945 Da. Ss γ-CMD was crystallized and X-ray data were collected to 2.40 Å resolution. The crystal belonged to space group $P4_32_12$, with unit-cell parameters $a = b = 66.66$, $c = 184.82$ Å. The Matthews coefficient and solvent content were estimated to be 2.14 Å^3 Da^{-1} and 42.6%, respectively, assuming that the asymmetric unit contained three recombinant protein molecules.

1. Introduction

The catabolism of aromatic compounds has tremendous significance in two main aspects. Firstly, natural aromatic compounds are abundant in the biosphere and their recycling is an important part of the earth's carbon cycle. Notably, the complex aromatic polymer lignin comprises about 25% of the land-based biomass on earth (Harwood & Parales, 1996). Secondly, environmental contamination arising from industrial activities results in the accumulation of unnatural aromatic compounds, which raises a serious environmental issue. In an effort to resolve this problem, microbiologists have found that some soil bacteria can degrade those aromatic contaminants (Habe *et al.*, 2005).

Various aromatic compounds are degraded to two kinds of common diphenolic intermediates: catechol and protocatechuate. These two intermediates are converted to tricarboxylic acid cycle intermediates by the reactions of the β-ketoadipate pathway, which is widely distributed in soil bacteria and fungi. γ-Carboxymucolactone decarboxylase (γ-CMD; EC 4.1.1.44) catalyzes the conversion of γ-carboxymucolactone to β-ketoadipate enol-lactone in the protocatechuate branch of the β-ketoadipate pathway (Fig. 1).

γ-CMD from *Sulfolobus solfataricus* (Ss γ-CMD) is encoded by the *pcaC* gene. It consists of 139 amino-acid residues with a molecular weight of 15 945 Da and shares 32% sequence identity over a range of 118 residues with γ-CMD from *Methanobacterium thermoautotrophicum* (Mt γ-CMD; PDB code 2af7; S. M. Vorobiev, A. Kuzin, T. Skarina, A. Savchenko, A. Semesi, C. Arrowsmith, A. Edwards, G. T. Montelione & L. Tong, unpublished work). Although γ-CMD was purified for the first time in 1973 from the Gram-negative soil bacterium *Pseudomonas putida* (Parke *et al.*, 1973), detailed studies of its biochemical and structural features have not yet been reported. Here, we report the crystallization of Ss γ-CMD as an initial step in its structural characterization.



2. Experimental procedures

2.1. Protein expression and purification

The gene *pcaC* encoding Ss γ-CMD was amplified by polymerase chain reaction from the genomic DNA of *S. solfataricus*. The expression vector was constructed by ligation of the amplified gene into pET26b (Novagen) and the nucleotide sequence of the insert

was confirmed by sequencing. The inserted gene has a stop codon immediately after the coding sequence; the expressed protein is therefore expected to be the full-length protein without any extra sequence such as a His tag. *Escherichia coli* strain Rosetta2 (DE3) was transformed with the expression vector and cultured at 310 K in 2 l Luria–Bertani medium containing 30 µg ml^{-1} kanamycin. When the OD_{600} of the culture reached 0.6, the culture was cooled to 298 K and expression of Ss γ-CMD was induced by addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 m*M*. The cells were further incubated at 298 K with shaking for 16 h after IPTG induction and were then harvested by centrifugation at 4410*g* for 30 min (Hanil Supra 22K with rotor 11).

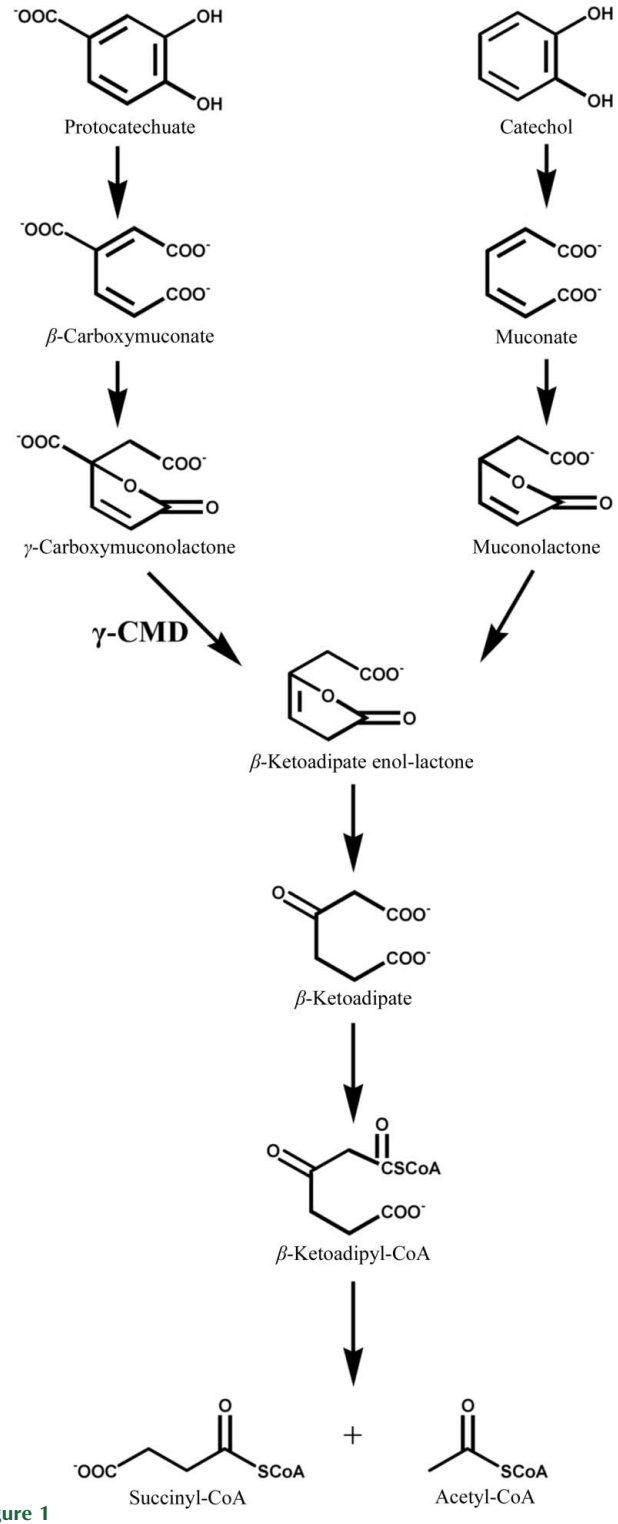


Figure 1
The β-ketoadipate pathway, which includes γ-carboxymucolactone decarboxylase (γ-CMD).

The harvested cells were resuspended in lysis buffer [20 m*M* Tris–HCl, 5%(*v*/*v*) glycerol, 0.1 m*M* TCEP, 1 m*M* phenylmethylsulfonyl fluoride pH 8.0] and lysed by sonication with a VCX500 ultrasonic processor (Sonics). The cell lysate was centrifugated at 24 650*g* for 30 min (Hanil Supra 22K with rotor 7) and the supernatant was subjected to heat treatment at 348 K for 20 min. The sample was chilled on ice for 10 min and subjected to centrifugation at 24 650*g* for 30 min (Hanil Supra 22K with rotor 7). The supernatant was applied onto a Q-Sepharose column (GE Healthcare Bioscience) employing a linear gradient of 0–1.0 *M* sodium chloride in working buffer [20 m*M* Tris–HCl, 5%(*v*/*v*) glycerol, 0.1 m*M* TCEP pH 8.0]. The fractions containing Ss γ-CMD were pooled and concentrated by centrifugal ultrafiltration (Amicon Ultra, Millipore). Gel filtration was then performed on a Superdex 200 column (GE Healthcare Bioscience) which was pre-equilibrated with working buffer containing 100 m*M* sodium chloride. SDS–PAGE was used to determine the fractions that contained Ss γ-CMD and to judge the purity of the sample at each step of the purification procedure. The purified sample was concentrated to 10 mg ml^{-1} and stored at 203 K. The concentration of the protein was determined by measuring the absorbance at 280 nm, employing the calculated extinction coefficient of 0.98 mg^{-1} ml cm^{-1} (http://www.expasy.org).

2.2. Crystallization and X-ray data collection

Crystallization conditions were initially screened using commercial screening kits from Hampton Research, Emerald BioSystems and Qiagen. Crystals were grown by the hanging-drop vapour-diffusion method. 2 µl protein solution was mixed with an equal volume of reservoir solution and the mixture was equilibrated against 0.5 ml reservoir solution at 295 K. The final refined composition of the reservoir solution was 22%(*w*/*v*) polyethylene glycol 3350, 0.2 *M* sodium chloride, 0.1 *M* Na HEPES pH 7.0 and 10%(*v*/*v*) glycerol. Immediately after the crystal was taken out of the drop, it was soaked in a 10 µl droplet of well solution for 3–5 s for cryoprotection and flash-cooled in liquid nitrogen.

X-ray diffraction data were collected at a wavelength of 1.0000 Å from an Ss γ-CMD crystal at 100 K on synchrotron beamline 4A, Pohang Light Source, Republic of Korea. Intensity data were processed, merged and scaled with *MOSFLM* and *SCALA* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994).

3. Results and discussion

γ-CMD from *S. solfataricus* was overexpressed in *E. coli* and purified, with a yield of 71 mg from a 2 l culture of Luria–Bertani medium. The purity of the protein sample was >95% as judged by SDS–PAGE. In the initial crystallization trials using commercial screening kits, small crystals of less than 20 µm in size were grown using condition Nos. 71 and 72 of the Qiagen Classics II Suite. Subsequent refinement of the initial conditions led to the growth of larger well diffracting crystals of 100–200 µm in size within one week using a reservoir solution with the following composition: 22%(*w*/*v*) polyethylene glycol 3350, 0.2 *M* sodium chloride, 0.1 *M* Na HEPES pH 7.0 and 10%(*v*/*v*) glycerol (Fig. 2).

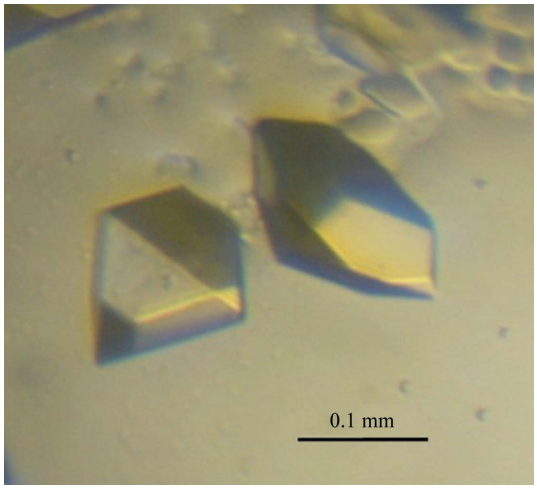


Figure 2
Crystals of γ-carboxymucolactone decarboxylase from *S. solfataricus*.

The crystal of Ss γ-CMD diffracted to 2.40 Å resolution on beamline 4A, Pohang Light Source, Republic of Korea. The crystal belonged to the primitive tetragonal system, with unit-cell parameters $a = b = 66.66$, $c = 184.82$ Å. Systematic absences of reflections indicated that the space group was $P4_12_12$ or $P4_32_12$. If three molecules of Ss γ-CMD were assumed to be present in the asymmetric unit, the Matthews coefficient was 2.14 Å^3 Da^{-1} and the corresponding solvent content was 42.6% (Matthews, 1968). The diffraction data set containing 17 153 unique reflections was 100.0% complete with a redundancy of 12.3 and an $R_{\rm merge}$ of 6.9%. Table 1 summarizes the statistics for the data collection.

Structure solution was attempted by molecular replacement using *Phaser* from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994; McCoy *et al.*, 2007) with the Mt γ-CMD structure (PDB code 2af7) as a search model. Three molecules of Mt γ-CMD were successfully located in the asymmetric unit of the Ss γ-CMD crystal and the correct space group was determined to be $P4_32_12$ from the two enantiomeric space groups. Model building and structure refinement are under way.

γ-CMD belongs to the CMD family, which also includes alkylhydroperoxidase (AhpD), in the Pfam protein-family database (Finn *et al.*, 2008) and a distinct conserved protein has recently been recognized as another member of the CMD family (Ito *et al.*, 2006). Even though γ-CMD was purified and crystallized for the first time in 1973 from the Gram-negative soil bacterium *P. putida* (Parke *et al.*, 1973), its crystal structure has not yet been reported. In the last 5 years, four γ-CMD structures from other bacteria and archaea (PDB codes 2af7, 3d7i, 2qeu and 1vke; Joint Center for Structural Genomics, unpublished work) have been deposited in the Protein Data Bank by structural genomics consortia, but their structural characters have not been analyzed or discussed in any published papers. The crystal structure of Ss γ-CMD from the current structural analysis will enhance understanding of the γ-CMD structure and its functional mechanism.

Table 1
Summary of data-collection statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P4_32_12$
Unit-cell parameters (Å)	$a = b = 66.66$, $c = 184.82$
Resolution (Å)	50.00–2.40 (2.53–2.40)
Wavelength (Å)	1.0000
No. of measured reflections	210651
No. of unique reflections	17153
$R_{\rm merge}$† (%)	6.9 (30.9)
Completeness (%)	100.0 (99.9)
$\langle I/\sigma(I)\rangle$	7.6 (2.1)
Redundancy	12.3 (11.7)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I(hkl)$ is the intensity of reflection hkl, $\sum_{hkl}$ is the sum over all reflections and $\sum_i$ is the sum over i measurements of reflection hkl.

This work was supported by the Soongsil University Research Fund. The authors thank the staff of Pohang Light Source beamline 4A.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, S. J., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L. & Bateman, A. (2008). *Nucleic Acids Res.* **36**, D281–D288.
Habe, H., Chung, J. S., Ishida, A., Kasuga, K., Ide, K., Takemura, T., Nojiri, H., Yamane, H. & Omori, T. (2005). *Microbiology*, **151**, 3713–3722.
Harwood, C. S. & Parales, R. E. (1996). *Annu. Rev. Microbiol.* **50**, 553–590.
Ito, K., Arai, R., Fusatomi, E., Kamo-Uchikubo, T., Kawaguchi, S., Akasaka, R., Terada, T., Kuramitsu, S., Shirouzu, M. & Yokoyama, S. (2006). *Protein Sci.* **15**, 1187–1192.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
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.
Parke, D., Meagher, R. B. & Ornston, L. N. (1973). *Biochemistry*, **12**, 3537–3542.