Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic studies of 3-deoxy-*manno*-octulosonate cytidylyltransferase from *Haemophilus influenzae*

Min-Je Ku,† Hye-Jin Yoon,† Hyung Jun Ahn, Hyung-Wook Kim, Seung-Hun Baek and Se Won Suh*

Structural Proteomics Laboratory, School of Chemistry and Molecular Engineering, Seoul National University, Seoul 151-742, South Korea

† These authors contributed equally to this work.

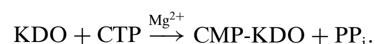
Correspondence e-mail: sewonsuh@snu.ac.kr

Received 4 July 2002
Accepted 28 October 2002

The enzyme 3-deoxy-*manno*-octulosonate cytidylyltransferase (CMP-KDO synthetase; CKS) catalyzes the activation of 3-deoxy-*manno*-octulosonate (KDO) by forming CMP-KDO. It is essential for the biosynthesis of lipopolysaccharides in Gram-negative bacteria and is a potential target for the discovery of antibacterial agents. L-CKS from *Haemophilus influenzae* was overexpressed with a C-terminal hexahistidine tag in *Escherichia coli* and crystallized in the presence of the substrate KDO at 297 K using PEG 4000 as a precipitant and ethylene glycol as an additive. The diffraction limit and spot shape of the native crystal could be improved significantly by dehydration/annealing. X-ray diffraction data were collected to 2.5 Å resolution from a native crystal. The crystals are orthorhombic, belonging to the space group $P2_12_12_1$, with unit-cell parameters a = 48.6, b = 83.1, c = 117.3 Å. The presence of two monomers of recombinant L-CKS in the crystallographic asymmetric unit gives a reasonable V_M of 2.05 Å^3 Da^{-1}, with a solvent content of 40.0%.

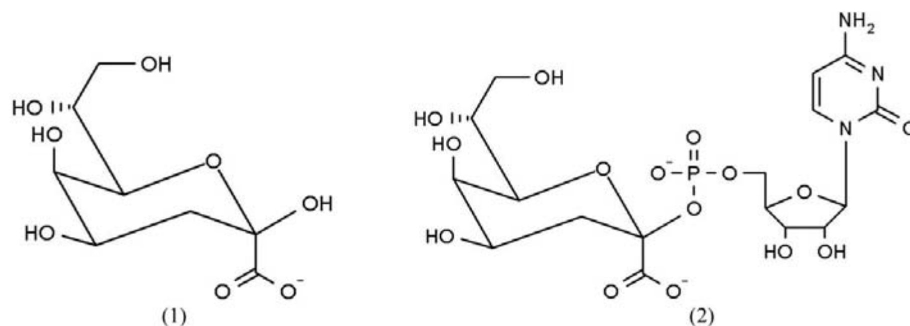
1. Introduction

The sugar 3-deoxy-D-*manno*-octulosonate [or 2-keto-3-deoxy-*manno*-octonic acid, KDO (1)] is an essential component of lipopolysaccharides (LPS) in Gram-negative bacteria. It serves as the link between lipid A and the core oligosaccharides of LPS (Raetz, 1996). Incorporation of KDO into lipopolysaccharides is catalyzed by the membrane-bound KDO transferase(s) that require(s) a nucleotide derivative of the sugar, CMP-KDO (2) (Royo *et al.*, 2000). Synthesis of CMP-KDO is catalyzed by 3-deoxy-*manno*-octulosonate cytidylyltransferase (CMP-KDO synthetase; CKS; E.C. 2.7.7.38) according to the reaction



$$\mathrm{KDO} + \mathrm{CTP} \xrightarrow{\mathrm{Mg}^{2+}} \mathrm{CMP\text{-}KDO} + \mathrm{PP_i}.$$

This reaction is unique to Gram-negative bacteria and therefore absent from mammalian cells (Jelakovic & Schulz, 2001). Therefore, CKS is an attractive target for the development of antibacterial agents (Goldman *et al.*, 1987). Two functionally distinct isozymes, L-CKS and K-CKS, have been found in *Escherichia coli*. L-CKS is encoded by the *kdsB* gene and is involved in the biosynthesis of LPS, while K-CKS is encoded by the *kpsU* gene and participates in the capsule expression (Rosenow *et al.*, 1995). These two isozymes show an amino-acid sequence identity of 44% and catalyze the same reaction, but with different kinetic properties and under different conditions in the bacterial cell (Rosenow *et al.*, 1995). In order to elucidate the structural basis of their functional differences, it is necessary to compare their three-dimensional structures. However, structural information is presently available for only K-CKS from *E. coli* (Jelakovic *et al.*, 1996; Jelakovic & Schulz, 2001, 2002). Therefore, we have initiated the structure determination of L-CKS from *H. influenzae* (Fleischmann *et al.*, 1995). The protein is composed of 254 residues with



© 2003 International Union of Crystallography
Printed in Denmark – all rights reserved

headxl;M_r = 28 124 (HI0058; http://www.tigr.org/tigr-scripts/CMR2/CMRHomePage.spl) and shows a high level (63%) of amino-acid sequence identity to L-CKS from *E. coli* (Goldman *et al.*, 1986). In comparison, it shows a lower level (43%) of sequence identity to K-CKS from *E. coli*. In this study, we overexpressed L-CKS from *H. influenzae* with a C-terminal hexahistidine tag in *E. coli* and crystallized it in the presence of the substrate KDO. Here, we report its crystallization conditions and preliminary X-ray crystallographic data.

2. Experimental

2.1. Protein expression and purification

The *kdsB* gene encoding L-CKS was amplified by the polymerase chain reaction using the *H. influenzae* genomic DNA as a template. The sequences of the forward and reverse oligonucleotide primers designed from the published genome sequence (Fleischmann *et al.*, 1995) were 5′-G GAA TTC **CAT ATG** TCA TTT ACC GTG ATT ATC CCC-3′ and 5′-CCG CCG **CTC GAG** ATT CGC CGC TAA AAT TGC CCG C-3′, respectively. The bases in bold represent the *Nde*I and *Xho*I digestion sites, respectively. The amplified DNA was inserted into the *Nde*I/*Xho*I-digested expression vector pET-21a (Novagen). This vector construction has six histidine residues to the C-terminus of the *kdsB* gene product to facilitate protein purification (see below). The protein was overexpressed in *E. coli* B834(DE3) cells. Cells were grown at 310 K to an OD_{600} of 0.5 in Luria–Bertani medium containing 50 µg ml^{-1} ampicillin and protein expression was induced by 0.5 m*M* isopropyl-β-D-thiogalactopyranoside (IPTG). Cell growth continued at 303 K for 4 h after IPTG induction and cells were harvested by centrifugation at 4200*g* (6000 rev min^{-1}; Sorvall GSA rotor) for 10 min at 277 K. The cell pellet was resuspended in ice-cold lysis buffer (20 m*M* Tris–HCl pH 7.9, 500 m*M* NaCl, 5 m*M* imidazole and 1 m*M* phenylmethylsulfonyl fluoride) and was homogenized with an ultrasonic processor. The crude cell extract was centrifuged at 36 000*g* (18 000 rev min^{-1}; Hanil Supra 21K rotor) for 30 min at 277 K and the recombinant protein in the supernatant fraction was purified by two chromatographic steps. The first step utilized the C-terminal hexahistidine tag by metal-chelate chromatography on Ni-NTA resin (Qiagen). Gel filtration was then performed on a HiLoad XK 16 Superdex 200 prep-grade column (Amersham-Pharmacia), which was previously equilibrated with buffer *A* (50 m*M* Tris–HCl pH 8.0, 1 m*M* EDTA) containing 100 m*M* sodium chloride. Homogeneity of the purified protein was assessed by SDS–PAGE (Laemmli, 1970). The protein solution was concentrated using a YM10 ultrafiltration membrane (Amicon). The protein concentration was estimated by measuring the absorbance at 280 nm, employing the calculated extinction coefficient of 15 930 M^{-1} cm^{-1} (SWISS-PROT; http://www.expasy.ch/).

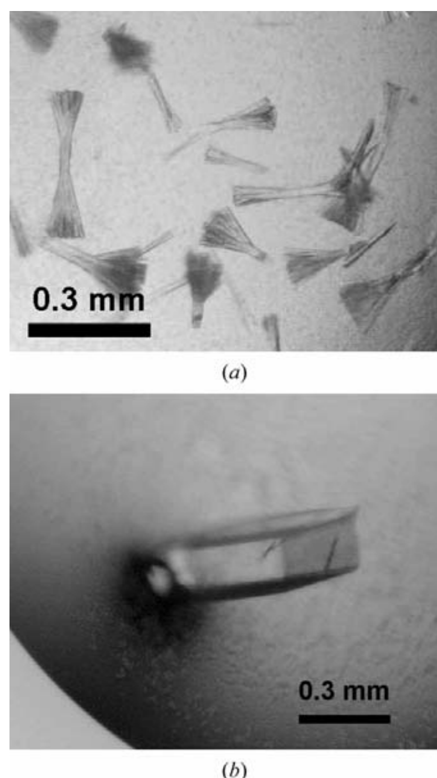


Figure 1
Crystals of CMP-KDO synthetase from *H. influenzae* grown in the presence of its substrate KDO. (*a*) Broom-shaped crystals grown in the absence of ethylene glycol. (*b*) A rectangular crystal grown in the presence of 3.0%(*v*/*v*) ethylene glycol as an additive. Its approximate dimensions are 0.4 × 0.3 × 0.15 mm.

2.2. Crystallization

Crystallization was performed at 297 K by the hanging-drop vapour-diffusion method using 24-well tissue-culture plates. Before the crystallization setup, the substrate KDO was added to the protein solution in 15-fold excess over the subunit concentration and the protein solution was incubated at 277 K for 1 d. Each hanging drop was prepared on a siliconized cover slip by mixing equal volumes (2 µl each) of the protein solution (at 15 mg ml^{-1} protein concentration) and the reservoir solution. When a number of additives were examined to improve the crystal quality, the hanging drop was prepared by mixing 2 µl of the protein solution, 2 µl of the reservoir solution and 0.4 µl of the additive solution, but the additive was not added to the reservoir solution. Each hanging drop was placed over a 1.0 ml reservoir solution. Initial crystallization conditions were established using screening kits from Hampton Research (Crystal Screens I, II and MembFac) and from Emerald BioStructures Inc. (Wizard I and II).

2.3. X-ray diffraction experiment

A crystal of the native protein was flash-frozen with a cryoprotectant solution, which consisted of 100 m*M* sodium citrate pH 5.6, 50 m*M* ammonium acetate, 30%(*w*/*v*) PEG 4000 and 20%(*v*/*v*) PEG 400. Crystals were soaked in 5 µl of the cryoprotectant solution for 10 s before being flash-frozen in liquid nitrogen. Since the crystal gave elongated spot shapes upon exposure to X-rays, it was removed from the cold nitrogen-gas stream and placed in 100 µl of the cryoprotectant solution in a spot plate (Hampton Research) without sealing and the cryoprotectant solution was dried out slowly for about 30 min. The dehydrated and annealed crystal was again flash-cooled. X-ray diffraction data were collected at 100 K with an ADSC Quantum 4R CCD detector at the BL-18B experimental station of the Photon Factory, Japan. The crystal was oscillated by 1.0° per frame over a total range of 108°. Intensity data were processed and scaled using the programs *MOSFLM* (Leslie, 1994) and *SCALA* (Collaborative Computational Project, Number 4, 1994). X-ray data from a crystal of selenomethionine-substituted protein were similarly collected on a MacScience imaging-plate detector at the BL-6B experimental station of Pohang Light Source, Korea, except that the crystal dehydration/annealing step was omitted.

3. Results

The recombinant CKS from *H. influenzae* with a C-terminal hexahistidine tag was highly overexpressed in soluble form in *E. coli* cells, with a yield of ~75 mg of the purified enzyme per litre of culture. Initially, broom-shaped crystals (Fig. 1) were obtained using a reservoir solution at pH 5.6 containing 100 m*M* sodium citrate, 30%(*w*/*v*) PEG 4000 and 50 m*M* ammonium acetate. However, their size and quality were not suitable for X-ray data collection. Further optimization of the crystallization conditions was achieved by screening addi-

Table 1
Data-collection statistics.

Values in parentheses refer to the highest resolution shells (2.64–2.50 Å for native data and 2.28–2.20 Å for SeMet data).

Data	Native	SeMet
X-ray wavelength (Å)	0.9500 (PF 18B)	0.9792 (PLS 6B)
Temperature (K)	100	100
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 48.64, b = 83.13, c = 117.33	a = 48.42, b = 82.61, c = 115.71
Resolution range (Å)	30.0–2.50	30.0–2.20
Total/unique reflections	150 357/15 313	396 736/24 031
$R_{\rm merge}$† (%)	7.7 (33.2)	5.1 (13.3)
Data completeness (%)	88.8 (88.8)	93.9 (86.6)
Average $I/\sigma(I)$	8.2 (2.3)	44.7 (11.0)

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

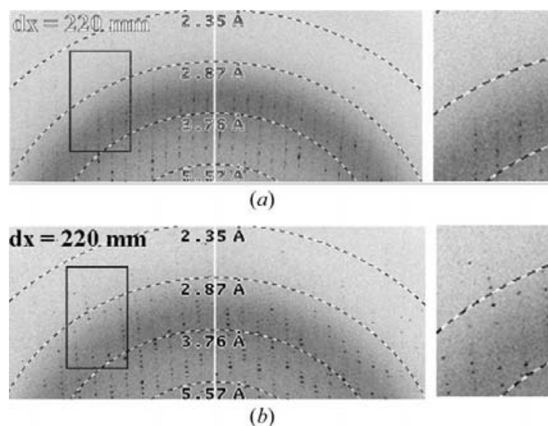


Figure 2
A comparison of X-ray diffraction images (*a*) before and (*b*) after dehydration/annealing of a native crystal. This technique resulted in a significant improvement of the diffraction limit and the mosaicity. The crystal-to-detector distance, dx, was 220 mm. A magnified view of a boxed region is shown on the right.

tives. Better crystals were obtained when ethylene glycol was added to the hanging drop, its final concentration being adjusted to 3.0%(v/v). Rectangular crystals grew to dimensions of 0.4 × 0.3 × 0.15 mm within a week (Fig. 1). They were reasonably stable for mechanical treatment and X-ray exposure.

Since the flash-frozen crystals diffracted to ~3.0 Å and their spot shapes were elongated, we attempted crystal dehydration/annealing, which significantly improved the diffraction quality. The diffraction pattern extended to ~2.5 Å and the mosaicity became smaller (Fig. 2). It is conceivable that the protein molecules in the crystal lattice slightly rearranged themselves during crystal dehydration/annealing and the crystal became better ordered. A total of 15 313 unique reflections were measured with a redundancy of 9.8, with a completeness of 88.8% and an $R_{\rm merge}$ of 7.7%. The space group was determined to be $P2_12_12_1$ from the systematic absences of reflections; the unit-cell parameters were a = 48.6, b = 83.1, c = 117.3 Å, $\alpha = \beta = \gamma = 90°$. The unit-cell parameters were fixed for processing the data. Table 1 summarizes the statistics of data collection. The presence of two monomers of the recombinant CKS in the crystallographic asymmetric unit gives a crystal volume per protein mass ($V_{\rm M}$) of 2.05 Å^3 Da^{-1}, with a solvent content of 40.0% (Matthews, 1968).

Employing the model of *E. coli* K-CKS (PDB code 1he7; Jelakovic & Schulz, 2001) as a probe, molecular-replacement trials gave a solution which could not readily be refined to a reasonably low R value. Therefore, we prepared a SeMet-substituted protein and crystallized it under conditions identical to those for the native crystal. We found that the SeMet crystals diffracted to a higher resolution even without crystal dehydration/annealing. Owing to a shortage of synchrotron beam time, however, we only collected X-ray data to 2.2 Å at a single wavelength (Table 1). The unit-cell parameters are a = 48.42 (1), b = 82.61 (12), c = 115.71 (17) Å, where the estimated standard deviations are given in parentheses. Single-wavelength data were not sufficient for phasing. In order to solve the structure by the multi-wavelength anomalous diffraction (MAD) method, we plan to collect MAD data at three different wavelengths when sufficient synchrotron beam time becomes available.

We thank Professor N. Sakabe and his staff for assistance during data collection at beamline BL-18B of the Photon Factory, Japan. We also thank the staff at beamline BL-6B of Pohang Light Source. This work was supported by a grant from the Korea Ministry of Science and Technology (NRL-2001, grant No. M1-0104-00-0132). MJK, HJY, HJA and HWK are supported by BK21 Fellowships.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Fleischmann, R. D. *et al.* (1995). *Science*, **269**, 496–512.
Goldman, R. C., Bolling, T. J., Kohlbrenner, W. E., Kim, Y. & Fox, J. L. (1986). *J. Biol. Chem.* **261**, 15831–15835.
Goldman, R., Kohlbrenner, W., Lartey, P. & Pernet, A. (1987). *Nature (London)*, **329**, 162–164.
Jelakovic, S., Jann, K. & Schulz, E. S. (1996). *FEBS Lett.* **391**, 157–161.
Jelakovic, S. & Schulz, E. S. (2001). *J. Mol. Biol.* **312**, 143–155.
Jelakovic, S. & Schulz, E. S. (2002). *Biochemistry*, **41**, 1174–1181.
Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.
Leslie, A. G. W. (1994). *MOSFLM User Guide.* MRC–LMB, Cambridge, England.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–493.
Raetz, C. R. H. (1996). *Escherichia Coli and Salmonella*, 2nd ed., edited by F. C. Neidhardt, pp. 1035–1063. Washington, DC: ASM Press.
Rosenow, C., Roberts, I. S. & Jann, K. (1995). *FEMS Microbiol. Lett.* **125**, 159–164.
Royo, J., Gomez, E. & Hueros, G. (2000). *J. Biol. Chem.* **275**, 24993–24999.