Acta Crystallographica Section D
Biological
Crystallography
ISSN 0907-4449

Expression, purification, crystallization and preliminary X-ray studies of geranylgeranyl diphosphate synthase from *Thermus thermophilus* HB8

Kazuya Nishio,[a] Yuichi Nodake,[a] Kensaku Hamada,[a,b]† Kyoko Suto,[a]*‡ Noriko Nakagawa,[c] Seiki Kuramitsu[a,c] and Keiko Miura[d]

[a]RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, Japan, [b]Interdisciplinary Faculty of Science and Engineering, Shimane University, Matsue 690-8504, Japan, [c]Department of Biology, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan, and [d]Japan Synchrotron Radiation Institute (JASRI), 1-1-1 Kouto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, Japan

† Present address: X-ray Research Laboratory, Rigaku Corporation, Akisima, Tokyo 196-8666, Japan.
‡ Present address: Department of Biochemistry, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan.

Correspondence e-mail: ksuto@affrc.go.jp

Received 23 February 2003
Accepted 4 November 2003

Geranylgeranyl diphosphate (GGPP) synthase from *Thermus thermophilus* HB8 was expressed in *Escherichia coli*, purified to homogeneity and crystallized both as the recombinant native protein and its selenomethionine (SeMet) derivative. Well diffracting crystals of these proteins were obtained belonging to the tetragonal space group $P4_1$ or $P4_3$, with unit-cell parameters $a = b = 139.88$, $c = 73.37$ Å. There were two homodimers in the asymmetric unit. A native data set was collected to 1.55 Å resolution and a data set suitable for MAD phasing was collected to 2.40 Å resolution on beamline BL40B2 at SPring-8.

1. Introduction

Prenyltransferases (prenyl diphosphate synthases) catalyze the condensation of isopentenyl diphosphate with allylic diphosphate to give isoprenoids of defined chain length which are used as precursors in the synthesis of over 23 000 isoprenoid compounds and prenylated proteins. These enzymes are classified according to the chain length of the final product and the geometry formed by the double bonds in a condensation reaction (*E*- or *Z*-type). The short-chain prenyltransferase family consists of geranyl diphosphate synthase (GPP synthase, C_{10} product), farnesyl diphosphate synthase (FPP synthase, C_{15} product) and geranylgeranyl diphosphate synthase (GGPP synthase, C_{20} product). The structural genes for many kinds of *E*-type prenyltransferases have been cloned and characterized. Comparison of the amino-acid sequences has revealed two highly conserved aspartate-rich motifs, DD$X_{2–4}$D and DDXXD, which are essential for the condensation reaction (Chen *et al.*, 1994; Ogura & Koyama, 1998). The only known three-dimensional structures of *E*-type prenyltransferases are of farnesyl diphosphate synthase. The crystal structure of avian farnesyl diphosphate synthase determined by X-ray analysis consists of 12 mostly antiparallel α-helices that form a large cavity (Tarshis *et al.*, 1994). The two aspartate-rich motifs are located on the walls of the cavity forming the active site. The first aspartate-rich motif (FARM) interacts with the allylic substrate and the second aspartate-rich motif (SARM) interacts with isopentenyl diphosphate (Tarshis *et al.*, 1996). Using site-directed mutagenesis, it was shown that the region between the fifth amino-acid residue before the FARM motif and the end of the FARM sequence, called the chain-length determination region (CLD), fixes the chain length of the final product (Ohnuma *et al.*, 1996, 1997; Ohnuma & Wang, 1999).

GGPP synthase catalyzes the condensation of three molecules of isopentenyl diphosphate with dimethylallyl diphosphate to give a C_{20} compound which is used as a precursor in the synthesis of carotenoids, geranylgeranylated proteins, chlorophylls and archaeal ether-linked lipids for the cytoplasmic membrane. The GGPP synthase gene from the thermophilic bacterium *Thermus thermophilus* HB8 (ttGds) was cloned and overexpressed in *Escherichia coli* cells as a glutathione *S*-transferase fusion protein. The fusion protein formed a homodimer with high specific activity (Ohto *et al.*, 1999). GGPP synthases are classified into three types: type I (archaea) and type II (plants and eubacteria), according to differences in the CLD, and type III (yeast and mammals), which has a unique FARM. The ttGds differs from the typical type I by lacking inserted amino-acid residues in the FARM (Ohto *et al.*, 1999). Detailed structural information will allow us to understand the molecular basis of the chain-determination mechanism of isoprenyl diphosphate synthases and their origin as a result of molecular evolution. Here, we report on the purification, crystallization and collection of X-ray data for analysis of ttGds.

2. Expression of *T. thermophilus* HB8 geranylgeranyl diphosphate synthase in *E. coli* cells

To obtain the cDNA fragment encoding ttGds, polymerase chain reaction (PCR) amplification was performed using *T. thermophilus* HB8 chromosomal DNA as a template with the forward primer 5′-ATATCATATGGTACCCG-

© 2004 International Union of Crystallography
Printed in Denmark – all rights reserved

CGCCCGAGGCCATCCGGCAGG-3′ and the reverse primer 5′-ATATAGATCTTTA-TTATGCCCTGCGCTCCACCAAAGCG-GCG-3′. The PCR products were purified on a gel and then ligated into the pT7Blue vector. After confirmation of the DNA sequence, the DNA fragment encoding ttGds was excised by digestion with *Nde*I and *Bgl*II and ligated into the expression vector pET-3a previously digested with the same enzymes. The *gds* gene encoded a 330-residue polypeptide with a molecular weight of 36 503 Da (deposited in GenBank with ID BAC77651), compared with 344 amino-acid residues (37 946 Da) as determined by Ohto *et al.* (1999). The resultant plasmid, pET-ttGds, was introduced into *E. coli* BL21(DE3)pLysE and the recombinant protein was induced with 1 m*M* isopropyl 1-thio-β-D-galactoside (IPTG). Over-production occurred and the ttGds protein was analyzed by SDS–PAGE using 12% polyacrylamide gels (Laemmli, 1970).

3. Purification of the recombinant ttGds

E. coli BL21(DE3)pLysE transformants harbouring the expression vector pET-ttGds were cultured at 310 K in 1 l of Luria–Bertani broth containing 50 µg ml^{-1} ampicillin. After cultivation, the cells were harvested by centrifugation at 14 000*g* for 30 min and the pellet was washed twice in 20 m*M* Tris–HCl buffer pH 8.0 containing 5 m*M* 2-mercaptoethanol (buffer *A*) and resuspended in the same buffer. Cells were sonicated at ice-bath temperature. The crude lysate was incubated at 343 K for 10 min to denature *E. coli* proteins and the resultant solution was centrifuged at 18 000*g* for 30 min.

After dialysis of the supernatant fraction against buffer *A*, the solution was introduced into a SuperQ-Toyopearl 650M column (Tosoh, 2.0 × 10 cm) equilibrated with the same buffer and eluted with a linear gradient of 0–0.5 *M* NaCl in buffer *A*. The protein fractions were collected and further purified using a Mono Q HR 10/10 column (Pharmacia) equilibrated with the same buffer and eluted with a linear gradient of 0–0.5 *M* NaCl in the same buffer. Each protein fraction was subjected to SDS–PAGE. The fractions in 0.25 *M* NaCl showed the same molecular weight as ttGds and were pooled. In the final step, the protein was purified by gel filtration using a 16/60 Hiload Superdex 200pg column (Pharmacia) which had been equilibrated with 10 m*M* Tris–HCl buffer pH 8.0 containing 150 m*M* NaCl and 10 m*M* dithiothreitol (DTT). The selenomethionylated ttGds protein, which was obtained by culture of a methionine auxotroph in LeMaster medium (Le Master & Richards, 1985) containing selenomethionine, was purified by the same procedure used to purify wild-type ttGds protein from *E. coli* cells.

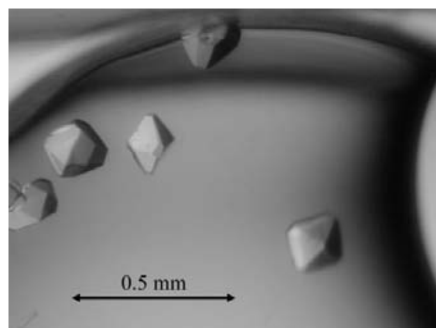

Figure 1
Bipyramidal crystals of native ttGds.

Table 1
Statistics of X-ray data sets of native and SeMet crystals.

Values in parentheses refer to the highest resolution shell.

Data set	Native	SeMet remote 1	SeMet edge	SeMet peak	SeMet remote 2
Wavelength (Å)	1.00000	0.97880	0.97777	0.97760	0.96860
Resolution (Å)	30.00–1.55 (1.61–1.55)	30.00–2.40 (2.49–2.40)	30.00–2.40 (2.49–2.40)	30.00–2.40 (2.49–2.40)	30.00–2.40 (2.49–2.40)
Total No. of reflections	617224	459219	574730	456169	373117
No. unique reflections	182752	55088	55156	55006	54970
Average $I/\sigma(I)$	42.2 (10.9)	43.8 (20.8)	41.4 (23.1)	47.4 (23.8)	40.0 (18.7)
Completeness (%)	89.2 (81.8)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
$R_{\rm merge}$† (%)	3.9 (11.5)	6.1 (15.6)	7.4 (16.2)	6.9 (15.0)	6.9 (16.0)

† $R_{\rm merge} = \sum_h \sum_i |I_{h,i} - \langle I_h \rangle| / \sum_h \sum_i |I_{h,i}|$, where h refers to unique reflection indices and i indicates symmetry-equivalent indices.

4. Crystallization

Recombinant native ttGds was concentrated to 20 mg ml^{-1} in 5 m*M* Na HEPES buffer pH 7.5 using a Centricon-30 concentrator. Crystallization trials were carried out at 296 K by the vapour-diffusion method in sitting drops using a sparse-matrix screen (Jancarik & Kim, 1991; Crystal Screen I, Hampton Research). Each drop was prepared by mixing 2 µl of protein solution with 2 µl of reservoir solution and the mixture was allowed to equilibrate against 100 µl of reservoir solution. Two kinds of crystals, with bipyramidal and thin needle-like shapes, were obtained within one week. The bipyramidal crystals were obtained with reagent No. 37 of Crystal Screen I, which contained 8%(*w*/*v*) PEG 4000 in 0.1 *M* sodium acetate buffer pH 4.6. The crystals grew to maximum dimensions of 0.8 × 0.4 × 0.4 mm in a week (Fig. 1). Thin needle-like crystals were obtained when reagents Nos. 18, 22, 40 and 41 containing 20–30%(*w*/*v*) PEG 8000 or PEG 4000 were used as precipitants. Crystals of the SeMet derivative were obtained by microseeding with native bipyramidal crystals using the same crystallization conditions as employed for native bipyramidal crystals. They grew to dimensions of 0.15 × 0.15 × 0.15 mm in 2 d.

5. X-ray data collection and crystallographic study

The recombinant native and SeMet-derivative crystals were flash-cooled in a nitrogen stream after soaking them in a cryoprotectant reagent containing 30%(*w*/*v*) trehalose or 25%(*v*/*v*) glycerol, respectively, for a few seconds. All data sets were collected on beamline BL40B2 at SPring-8, Japan and were processed using the *HKL*2000 program suite (Otwinowski & Minor, 1997).

The bipyramidal crystals diffracted to better than 1.55 Å resolution and were more suitable as single crystals for X-ray experiments than the thin needle crystals. Therefore, the bipyramidal crystals were used in the following experiments. A native data set was collected at a 200 mm crystal-to-detector distance and 1.0° oscillation angle using a Rigaku R-AXIS IV imaging-plate detector. The total oscillation range covered was 95°. The crystals belonged to the tetragonal space group $P4_1$ or $P4_3$, with unit-cell parameters $a = b = 139.88$ (6), $c = 73.37$ (3) Å. Assuming the presence of two homodimers in an asymmetric unit, the crystal-packing parameter $V_{\rm M}$ (Matthews, 1968) was 2.46 Å^3 Da^{-1}, corresponding to a solvent content of 50%.

Prior to data collection employing the multiple-wavelength anomalous dispersion (MAD) method, an X-ray fluorescence spectrum of a SeMet crystal was measured in order to optimize the wavelengths. The MAD data sets were collected at four different wavelengths using an ADSC

Quantum 4R CCD detector under the following conditions: a crystal-to-detector distance of 170 mm, an oscillation angle of 1.0° and an total oscillation range of 200°. Crystals of the SeMet derivative belonged to the same space group $P4_1$ or $P4_3$ as the bipyramidal crystals of the recombinant native protein; the unit-cell parameters were $a = b = 139.44$ (7), $c = 72.99$ (2) Å.

Intensity data for the recombinant native and SeMet-derivative crystals were processed to 1.55 Å resolution with 89.2% completeness and to 2.40 Å resolution with 100.0% completeness, respectively. The statistics for each data set are tabulated in Table 1.

Phase-determination experiments are now in progress.

This work was supported by a grant from the Collaborative Project on Advanced SR Research at SPring-8 between JASRI and the RIKEN Harima Institute. We gratefully acknowledge the advice and assistance of Dr Masahide Kawamoto in the collection of data.

References

Chen, A., Kroon, P. & Poulter, C. D. (1994). *Protein Sci.* **3**, 600–607.

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

Laemmli, U. K. (1970). *Nature (London)*, **227**, 680–685.

Le Master, D. M & Richards, F. M. (1985). *Biochemistry*, **24**, 7263–7268.

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

Ogura, K. & Koyama, T. (1998). *Chem. Rev.* **98**, 1263–1276.

Ohnuma, S., Hirooka, K., Hemmi, H., Ishida, C., Ohto, C. & Nishino, T. (1996). *J. Biol. Chem.* **271**, 18831–18837.

Ohnuma, S., Hirooka, K., Ohto, C. & Nishino, T. (1997). *J. Biol. Chem.* **272**, 5192–5198.

Ohnuma, S. & Wang, K. (1999). *Trends Biochem Sci.* **24**, 445–451.

Ohto, C., Ishida, C., Koike-Takeshita, A., Yokoyama, K., Muramatsu, M., Nishino, T. & Obata, S. (1999). *Biosci. Biotechnol. Biochem.* **63**, 261–270.

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

Tarshis, L. C., Proteau, P. J., Kellogg, B. A., Sacchettini, J. C. & Poulter, C. D. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 15018–15023.

Tarshis, L. C., Yan, M., Poulter, C. D. & Sacchettini, J. C. (1994). *Biochemistry*, **33**, 10871–10877.