Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray crystallographic studies of chorismate synthase from *Helicobacter pylori*

Hyung Jun Ahn, Jin Kuk Yang, Byung Il Lee, Hye Jin Yoon, Hyung Wook Kim and Se Won Suh*

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

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

Received 11 October 2002
Accepted 2 January 2003

Chorismate synthase (EC 4.6.1.4) catalyzes the transformation of 5-enolpyruvylshikimate 3-phosphate to chorismate in the last step of the shikimate pathway. Chorismate synthase from *Helicobacter pylori* fused with an eight-residue C-terminal tag was overexpressed in soluble form in *Escherichia coli*. It was crystallized at 296 K using polyethylene glycol 400 as a precipitant. A set of X-ray diffraction data was collected to 2.5 Å resolution using synchrotron radiation. The crystals belong to the tetragonal space group *I*4, with unit-cell parameters $a = b = 145.79$, $c = 130.98$ Å. The asymmetric unit contains a tetramer, giving a crystal volume per protein mass (V_M) of 2.13 $Å^3$ Da^{-1} and a solvent content of 42.3%.

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

1. Introduction

The seven-step shikimate pathway links the metabolism of carbohydrates to the biosynthesis of aromatic amino acids and many aromatic secondary metabolites, including tetrahydrofolate and ubiquinone (Pittard, 1996; Meganathan, 2001). It provides excellent potential targets for antimicrobial agents, antiparasitic agents and herbicides (Roberts *et al.*, 2002; Schonbrunn *et al.*, 2001) because it is absent in animals but is apparently essential in bacteria, fungi, parasites and plants. As the last enzyme in the shikimate pathway, chorismate synthase catalyzes an unusual reaction, the conversion of 5-enolpyruvylshikimate 3-phosphate (EPSP) to chorismate by *anti*-1,4-elimination of the 3-phosphate group and the C(6pro-*R*) hydrogen from EPSP (Hill & Newkome, 1969; Onderka & Floss, 1969; Macheroux *et al.*, 1999). Although the reaction does not involve an overall change in redox states, it has an absolute requirement for reduced FMN as a cofactor (Morell *et al.*, 1967; Welch *et al.*, 1974). Evidence for a radical mechanism has been obtained for *Escherichia coli* chorismate synthase (Osborne *et al.*, 2000). Spectroscopic data also suggest that a major structural change in *E. coli* chorismate synthase is induced by flavin and substrate binding (Macheroux *et al.*, 1998).

Structural information is necessary for structure-based inhibitor discovery and for a better understanding of the catalytic mechanism. However, no three-dimensional structure of chorismate synthase has yet been reported. In this work, we have initiated structure determination of chorismate synthase from *Helicobacter pylori*, a 365-residue protein encoded by the *aroC* gene. Its amino-acid sequence shows 46% identity to the well characterized *E. coli* chorismate synthase. As the first step, we have overexpressed *H. pylori* chorismate synthase in *E. coli* and crystallized it. We report here preliminary X-ray crystallographic data as well as crystallization conditions.

2. Experimental

2.1. Protein expression and purification

The *aroC* gene (HP0663) encoding chorismate synthase was amplified by the polymerase chain reaction using the genomic DNA of *H. pylori* strain 26695 as a template. The forward and reverse oligonucleotide primers designed using the published genome sequence (Tomb *et al.*, 1997) were AGG GAA TTC **CAT ATG** AAC ACT TTG GGG CGT TTT and GGG CCG **CTC GAG** ATT CTC ATT ATA AAT CGT TTT, respectively. The bases shown in bold represent the *Nde*I and *Xho*I restriction-enzyme cleavage sites, respectively. The amplified DNA was digested with the enzymes *Nde*I and *Xho*I and was then inserted into the *Nde*I/*Xho*I-digested expression vector pET-21a(+) (Novagen). This vector construction adds an eight-residue tag (LEHHHHHH) to the C-terminus of the recombinant protein to facilitate protein purification. The enzyme was overexpressed in *E. coli* C41(DE3) cells (Miroux & Walker, 1996). Cells were grown in Luria–Bertani medium to an OD_{600} of 0.6 at 291 K and expression of the recombinant enzyme was induced by 1.0 m*M* isopropyl β-D-thiogalactopyranoside (IPTG) at 291 K. Cell growth continued at 291 K for 12 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 mM Tris–HCl pH 7.9, 0.50 M sodium chloride, 50 mM imidazole, 1 mM phenylmethylsulfonyl fluoride) and was then homogenized by sonication. The crude lysate was centrifuged at 36 000g (18 000 rev min^{-1}; Hanil Supra 21K rotor) for 30 min at 277 K.

Three chromatographic steps were used to purify the recombinant protein in the supernatant fraction. The first step utilized the C-terminal hexahistidine tag by metal-chelate chromatography on Ni–NTA resin (Qiagen). Next, gel filtration was performed on a HiLoad XK16 Superdex 200 prep-grade column (Amersham Biosciences), which was previously equilibrated with buffer A (20 mM potassium phosphate pH 6.4) containing 100 mM sodium chloride and 1 mM β-mercaptoethanol. Further purification was achieved by ion-exchange chromatography on a Mono S HR5/5 column (Amersham Biosciences), which was previously equilibrated with buffer A. Before loading the protein sample onto this column, the salt concentration was lowered to 50 mM by diluting it with buffer A. The protein was eluted with a linear gradient of 0–1.0 M sodium chloride in buffer A. The homogeneity of the purified protein was judged by polyacrylamide gel electrophoresis in the presence of 0.1%(w/v) sodium dodecyl sulfate (Laemmli, 1970). After dialysis of the protein solution against buffer A containing 100 mM sodium chloride and 1 mM dithiothreitol, it was concentrated using an YM10 membrane (Amicon) to about 30 mg ml^{-1}. The protein concentration was estimated by measuring the absorbance at 280 nm employing the calculated molar extinction coefficient of 11 520 M^{-1} cm^{-1} (SWISS-PROT, http://www.expasy.ch/).

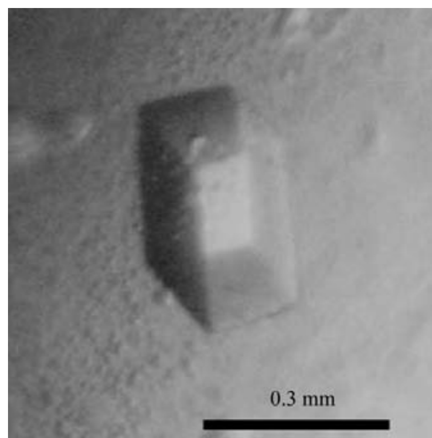

Figure 1
A tetragonal crystal of chorismate synthase from *H. pylori*. Its approximate dimensions are 0.4 × 0.2 × 0.15 mm.

2.2. Dynamic light-scattering studies

A dynamic light-scattering experiment was performed on a Model DynaPro-801 instrument from Protein Solutions (Lakewood, New Jersey, USA). The data were measured at 297 K with 2.5 mg ml^{-1} of protein in buffer A containing 100 mM sodium chloride.

2.3. Crystallization and X-ray data collection

Crystallization was performed by the hanging-drop vapour-diffusion method using 24-well tissue-culture plates (Hampton Research) at 296 K. Each hanging drop, prepared by mixing 2 µl each of the protein solution and the reservoir solution, was placed over a 1.0 ml reservoir solution. Initial crystallization conditions were screened using Crystal Screen I and II kits (Hampton Research).

For X-ray data collection, a crystal was transferred to a solution consisting of 34%(w/v) PEG 400, 100 mM sodium chloride, 200 mM calcium chloride, 100 mM HEPES pH 7.5 within 1 min in one step before being flash-frozen. X-ray diffraction data were collected at 100 K on an imaging-plate detector system (MacScience DIP2030b) at experimental station 6B, Pohang Light Source, Korea. The crystal was rotated through a total of 120°, with a 1.0° oscillation range per frame. Data were processed and scaled using *DENZO* and *SCALEPACK* from the *HKL* program suite (Otwinowski & Minor, 1997).

3. Results

Chorismate synthase from *H. pylori* fused with an eight-residue tag (LEHHHHHH) was highly overexpressed in *E. coli* C41(DE3) cells upon induction by 1.0 mM IPTG at 291 K. About 90% of the expressed protein was in the soluble fraction. The yield was ~50 mg of purified protein per litre of culture. The molecular mass of the recombinant enzyme in the native state was estimated to be 198 kDa by dynamic light-scattering analysis, with a polydispersity of 27%. This is consistent with the protein being a tetramer. Well diffracting crystals were obtained with the optimized reservoir solution comprising 28%(w/v) PEG 400, 100 mM sodium chloride, 200 mM calcium chloride, 100 mM HEPES pH 7.5. Crystals grew reproducibly to approximate dimensions of 0.4 × 0.2 × 0.15 mm within a week (Fig. 1). A set of diffraction data was collected to 2.5 Å resolution at 100 K using synchrotron X-rays. A total of 466 523 measured reflections were merged into 40 795 unique reflections with an $R_{\rm merge}$ (on intensity) of 4.6%. The crystals belong to the tetragonal space group $I4$, with unit-cell parameters a = b = 145.79 (6), c = 130.98 (19) Å, where the estimated standard deviations are given in parentheses. The presence of four monomers of chorismate synthase in the asymmetric unit gives a crystal volume per protein mass ($V_{\rm M}$) of 2.13 Å^3 Da^{-1} (Matthews, 1968), with a corresponding solvent content of 42.3%. Table 1 summarizes the data-collection statistics.

Table 1
Data-collection statistics.

Values in parentheses refer to the highest resolution shell (2.59–2.50 Å).

X-ray wavelength (Å)	0.9794
Temperature (K)	100
Resolution range (Å)	50.0–2.50
Total/unique reflections	466523/40795
Space group	$I4$
Unit-cell parameters (Å)	a = 145.79, b = 145.79, c = 130.98
Data completeness (%)	87.0 (80.5)
Average $I/\sigma(I)$	32.9 (9.1)
$R_{\rm merge}$† (%)	4.6 (16.7)

† $R_{\rm merge} = \sum_h \sum_i |I(h)_i - \langle I(h)\rangle| / \sum_h \sum_i I(h)_i$, 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.

The self-rotation function R (ω, φ, κ) was calculated using the program *POLARRFN* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). It showed two strong peaks (74.9% of the origin peak) at ω = 90, φ = 45° and at ω = 90, φ = 135° in the κ = 180° section, while the κ = 90° section showed no significant peaks. This indicates that twofold non-crystallographic symmetry axes relate the four monomers in the asymmetric unit and that the four monomers are likely to form a tetramer of 222 symmetry. We checked the possibility of twinning (Yeates, 1997) and found that the crystal is partially merohedrally twinned. A twin fraction α of 28.0% in the range 20–2.5 Å was obtained using the program *CNS* (Brünger *et al.*, 1998). However, partial twinning did not hamper solution of the phase problem using multi-wavelength anomalous diffraction data collected at 3.2 Å resolution from a crystal of the selenomethionine-substituted protein, which was partially twinned to a similar extent. The model has been refined against the native data without detwinning, giving $R_{\rm cryst}$ and $R_{\rm free}$ values of 21.0 and 27.9% to 2.5 Å, respectively. Refinement of the model against the detwinned data gave lower $R_{\rm cryst}$ and $R_{\rm free}$ values of 13.6 and 20.3%, respectively. These twinned R values probably

underestimate the true crystallographic R values to some degree (Redinbo & Yeates, 1993).

We thank Dr H. S. Lee and his staff for assistance during data collection at beamline BL-6B, Pohang Light Source, Korea. We also thank Professor N. Sakabe and his staff for assistance during data collection at beamline BL-18B, Photon Factory, Japan (Proposal No. 01G147). This work was supported by the Center for Functional Analysis of Human Genome under the 21C Frontier R&D Program of the Korea Ministry of Science and Technology. HJA, JKY, BIL, HJY and HWK are recipients of the BK21 Fellowship.

References

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Hill, R. K. & Newkome, G. R. (1969). *J. Am. Chem. Soc.* **91**, 5893–5894.

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

Macheroux, P., Schmid, J., Amrhein, N. & Schaller, A. (1999). *Planta*, **207**, 325–334.

Macheroux, P., Schonbrunn, E., Svergun, D. I., Volkov, V. V., Koch, M. H., Bornemann, S. & Thorneley, R. N. (1998). *Biochem. J.* **335**, 319–327.

Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–493.

Meganathan, R. (2001). *FEMS Microbiol. Lett.* **203**, 131–139.

Miroux, B. & Walker, J. E. (1996). *J. Mol. Biol.* **260**, 289–298.

Morell, H., Clark, M. J., Knowles, P. F. & Sprinson, D. B. (1967). *J. Biol. Chem.* **242**, 82–90.

Onderka, D. K. & Floss, H. G. (1969). *J. Am. Chem. Soc.* **91**, 5894–5896.

Osborne, A., Throneley, R. N. F., Abell, C. & Bornemann, S. (2000). *J. Biol. Chem.* **275**, 35825–35830.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–325.

Pittard, A. J. (1996). *Escherichia coli and Salmonella*, edited by F. C. Neidhardt, pp. 458–484. Washington, DC: ASM Press.

Redinbo, M. R. & Yeates, T. O. (1993). *Acta Cryst.* D**49**, 375–380.

Roberts, C. W., Roberts, F., Lyons, R. E., Kirisits, M., Mui, E. J., Finnerty, J., Johnson, J. J., Ferguson, D. J. P., Coggins, J. R., Krell, T., Coombs, G. H., Milhous, W. K., Kyle, D. E., Tzipori, S., Barnwell, J., Dame, J. B., Carlton, J. & McLeod, R. (2002). *J. Infect. Dis.* **185**, S25–S36.

Schonbrunn, E., Eschenburg, S., Shuttleworth, W. A., Schloss, J. V., Amrhein, N., Evans, J. N. & Kabsch, W. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 1376–1380.

Tomb, J.-F. *et al.* (1997). *Nature (London)*, **388**, 539–547.

Welch, G. R., Cole, K. W. & Gaertner, F. H. (1974). *Arch. Biochem. Biophys.* **165**, 505–518.

Yeates, T. O. (1997). *Methods Enzymol.* **276**, 344–358.