Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Purification, crystallization and preliminary crystallographic analysis of the glycine-cleavage system component T-protein from *Pyrococcus horikoshii* OT3

Neratur K. Lokanath, Chizu Kuroishi, Nobuo Okazaki and Naoki Kunishima*

Highthroughput Factory, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo-gun, Hyogo 679-5148, Japan

Correspondence e-mail: kunisima@spring8.or.jp

Received 5 April 2004
Accepted 27 May 2004

The glycine-cleavage system component T-protein is a folate-dependent enzyme that catalyzes the formation of ammonia and 5,10-CH_2-tetrahydrofolate from the aminomethyl intermediate bound to the lipoate cofactor of H-protein. T-protein from *Pyrococcus horikoshii* OT3 has been cloned, overexpressed in *Escherichia coli*, purified and crystallized by the microbatch method using PEG 4000 as a precipitant at 296 K. X-ray diffraction data have been collected to 1.50 Å resolution at 100 K using synchrotron radiation. The crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 78.980$, $b = 95.708$, $c = 118.331$ Å. Assuming one homodimer per asymmetric unit gives a V_M value of 2.4 Å^3 Da^{-1} and a solvent content of 49.0%.

1. Introduction

The glycine-cleavage system (GCS; EC 2.1.2.10) is a multicomponent enzyme system that catalyzes the cleavage of glycine to yield carbon dioxide, ammonia, 5,10-CH_2-tetrahydrofolate (Motokawa & Kikuchi, 1974) and NADH + H^+. This system is composed of four proteins: a pyridoxal phosphate enzyme (P-protein), a carrier protein containing covalently bound lipoic acid (H-protein), a folate-dependent enzyme (T-protein) and a protein exhibiting lipoamide dehydrogenase activity (L-protein). The system catalyzes the oxidative cleavage of glycine, a major catabolic pathway of glycine in various organisms. Firstly, the P-protein catalyzes the decarboxylation of the glycine molecule and subsequent transfer of the residual methylamine to the oxidized H-protein (H_{ox}), generating the methylamine-loaded H-protein (H_{met}). Secondly, the T-protein catalyzes the transfer of a methylene C atom from H_{met} to tetrahydrofolate, resulting in the release of ammonia and the generation of reduced H-protein (H_{red}). Lastly, the hydrolipoyl group of H_{red} is oxidized by the L-protein and H_{ox} is generated, thereby completing the catalytic cycle (Motokawa *et al.*, 1995). The GCSs distributed in bacteria, mitochondria of plants and mammals have been studied extensively (Perham, 2000; Douce *et al.*, 2001). In humans, a genetic disease caused by GCS absence, termed non-ketoic hyperglycinaemia, has been reported to cause a dramatic accumulation of glycine in the blood, leading to a neurological disorder (Tada & Hayasaka, 1987; Toone *et al.*, 2001).

Although the H-protein (Pares *et al.*, 1994, 1995; Cohen-Addad *et al.*, 1995; Faure *et al.*, 2000; Nakai, Ishijima *et al.*, 2003) and the L-protein (Mattevi *et al.*, 1991, 1992, 1993; Mande *et al.*, 1996; Li de la Sierra *et al.*, 1997; Toyoda, Kobayashi *et al.*, 1998; Toyoda, Suzuki *et al.*, 1998; Faure *et al.*, 2000) from several sources have been well studied by X-ray crystallography, the three-dimensional structures of the P-protein and the T-protein have not yet been reported. Recently, the crystallization of P-protein from *Thermus thermophilus* has been published (Nakai, Nakagawa *et al.*, 2003). T-protein is a folate-dependent enzyme that catalyzes the formation of ammonia and 5,10-CH_2-tetrahydrofolate from the methylamine-loaded H-protein (Motokawa *et al.*, 1995). Therefore, the T-protein should have interactions or binding site(s) with the H-protein, the folate cofactor and possibly the P-protein, which directly participates in the formation of the methylamine-loaded H-protein. Kinetic studies on the forward (Fujiwara *et al.*, 1984) and reverse (Okamura-Ikeda *et al.*, 1987) reactions catalyzed by the T-protein have revealed the importance of specific amino-acid residues participating in binding or catalysis. In this paper, we report the expression, purification and preliminary crystallographic analysis of the T-protein from *Pyrococcus horikoshii* OT3. Knowledge of the three-dimensional structure of the T-protein could help in understanding the molecular mechanism of the GCS multienzyme complex.

2. Experimental

2.1. Protein expression and purification

The T-protein from *P. horikoshii* OT3 used in this study has a molecular weight of 46.17 kDa and consists of 401 amino-acid residues. The plasmid encoding the T-protein

was digested with *Nde*I and *Bgl*II and the fragment was inserted into the expression vector pET-11a (Novagen) linearized with *Nde*I and *Bam*HI. *Escherichia coli* BL21 Codon Plus (DE3)-RIL cells were transformed with the recombinant plasmid and grown at 310 K in Luria–Bertani medium containing 50 µg ml^{-1} ampicillin for 20 h. The cells were harvested by centrifugation at 6500 rev min^{-1} for 5 min at 277 K, suspended in 20 m*M* Tris–HCl pH 8.0 containing 0.5 *M* NaCl and 5 m*M* 2-mercaptoethanol and finally disrupted by sonication and heated at 363 K for 10 min. The cell debris and denaturated protein were removed by centrifugation (14 000 rev min^{-1}, 30 min). The supernatant solution was used as the crude extract for purification.

The crude extract was desalted with a HiPrep 26/10 desalting column (Amersham Biosciences) and applied onto a Super Q Toyopearl 650M (Tosoh) column equilibrated with 20 m*M* Tris–HCl pH 8.0 (buffer *A*). After elution with a linear gradient of 0–0.5 *M* NaCl, the fraction containing the T-protein was desalted with a HiPrep 26/10 desalting column (Amersham Biosciences) with buffer *A*. The sample was loaded onto a Resource Q column (Amersham Biosciences) equilibrated with buffer *A*. After elution with a linear gradient of 0–0.3 *M* NaCl, the fraction containing the T-protein was desalted with a HiPrep 26/10 desalting column with 10 m*M* sodium phosphate pH 7.0. The sample was then applied onto a Bio-Scale CHT-20-I column (Bio-Rad) equilibrated with 10 m*M* sodium phosphate pH 7.0 and eluted with a linear gradient of 10–500 m*M* sodium phosphate. The sample was concentrated by ultrafiltration (Vivaspin) and loaded onto a HiLoad 16/60 Superdex 200 prep-grade column (Amersham Biosciences) equilibrated with buffer *A* containing 0.2 *M* NaCl. The homogeneity and identity of the purified sample were assessed by SDS–PAGE (Laemmli, 1970) and N-terminal sequence analysis. Finally, the purified T-protein was concentrated to 30.3 mg ml^{-1} using ultrafiltration and was stored at 203 K.

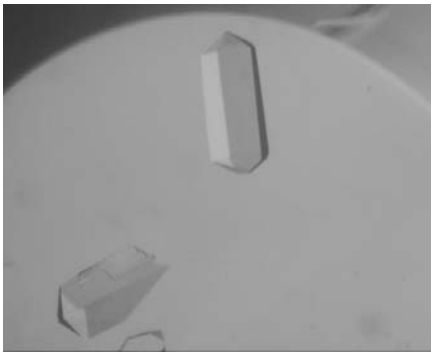

Figure 1
Crystals of T-protein from *P. horikoshii* OT3. The crystals have approximate dimensions of 0.2 × 0.15 × 0.15 mm.

2.2. Dynamic light-scattering study

The oligomerization state of the purified T-protein was examined by a dynamic light-scattering experiment using a DynaPro MS/X instrument (Protein Solutions), which was performed at a protein concentration of 20 mg ml^{-1} in 20 m*M* Tris–HCl pH 7.6 with 200 m*M* NaCl. Several measurements were taken at 291 K and analyzed by the *DYNAMICS* software, v.3.30 (Protein Solutions). A bimodal analysis resulted in a molecular weight of 75 kDa, which is consistent with the formation of a homodimer.

2.3. Crystallization and X-ray data collection

Crystals of T-protein were obtained by the microbatch method using Nunc HLA plates (Nalge Nunc International). Each crystallization drop was prepared by mixing 1.0 µl precipitant solution (12.5% PEG 4000, 0.1 *M* MES pH 5.9) and 1.0 µl protein solution at 25.43 mg ml^{-1}. The crystallization drop was overlaid with a 1:1 mixture of silicon/paraffin oils, allowing slow evaporation of water in the drop, and stored at 296 K. Initial crystallization conditions were established using the TERA (automatic crystallization) system (Sugahara & Miyano, 2002) from 144 independent conditions.

A crystal of the T-protein from *P. horikoshii* was flash-frozen using a cryoprotectant solution comprising 12.5% PEG 4000, 0.1 *M* MES and 30%(*v*/*v*) glycerol. X-ray diffraction data were collected from a flash-cooled crystal at 100 K using synchrotron radiation on a Rigaku R-AXIS V image-plate detector on beamline BL26B1 at SPring-8, Japan. Each frame was exposed for 45 s with a 1° oscillation at a crystal-to-detector distance of 250 mm. Data were processed and scaled using *HKL*2000 (Otwinowski & Minor, 1999).

3. Results

Well diffracting crystals of the T-protein were obtained using PEG 4000 as a precipitant. Crystals appeared about 15–18 d after setup and grew to approximate dimensions of 0.2 × 0.15 × 0.15 mm after four weeks (Fig. 1). A set of X-ray diffraction data was collected to 1.50 Å resolution at 100 K using synchrotron radiation. Data-collection statistics are summarized in Table 1. A total of 875 037 measured reflections in the resolution range 40–1.50 Å were merged into 142 355 unique reflections with an $R_{\rm merge}$ of 6.2%. The completeness of the merged data set at 1.50 Å resolution is 99%. The crystal belongs to the orthorhombic space group $P2_12_12_1$. The asymmetric unit probably contains two molecules of the T-protein, giving a crystal volume per protein weight ($V_{\rm M}$) of 2.4 Å^3 Da^{-1} and a solvent content of 49.0% (Matthews, 1968). A dynamic light-scattering experiment shows a consistent result of a dimeric state of this protein in solution (see §2). Combining these observations, the asymmetric unit is most likely to contain a homodimer of T-protein, although a self-rotation search did not clearly provide a local twofold symmetry. The high quality of the T-protein crystals will probably allow us to apply the multiple anomalous dispersion (MAD) method (Hendrickson *et al.*, 1990) using selenomethionine or other heavy-atom derivative crystals, which should be precise enough to phase the crystal structure of *P. horikoshii* T-protein at an atomic resolution.

Table 1
Statistics of X-ray data measurement for T-protein crystals.

Values in parentheses correspond to the highest resolution shell.

Space group	$P2_12_12_1$
Unit-cell parameters (Å)	a = 78.980, b = 95.708, c = 118.331
Wavelength (Å)	1.0
Resolution range (Å)	40–1.5 (1.62–1.55)
Total observations	875037
Unique reflections	142355
Redundancy	6.2 (5.7)
Completeness (%)	99.1 (97.2)
Mean $I/(I)$	19.2 (3.2)
$R_{\rm merge}$† (%)	4.9 (28.1)

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

The authors thank the staff of RIKEN Genomic Science Centre, Yokohama for providing the plasmid and thank the technical staff of the Highthroughput Factory, Y. Uchida for large-scale protein production, T. Hanamura for automatic crystallization and M. Yamamoto for manual crystallization. We also thank M. Yamamoto and his staff for assistance during data collection at beamline BL26B1 of SPring-8. This work (PH1146/HTPF11087) was supported by the 'National Project of

Protein Structural and Functional Analysis' funded by the MEXT of Japan.

References

Cohen-Addad, C., Pares, S., Sieker, L., Neuburger, M. & Douce, R. (1995). *Nature Struct. Biol.* **2**, 63–68.

Douce, R., Bourguignon, J., Neuburger, M. & Rebeille, F. (2001). *Trends Plant Sci.* **6**, 167–176.

Faure, M., Bourguignon, J., Neuburger, M., Macherel, D., Sieker, L., Ober, R., Kahn, R., Cohen-Addad, C. & Douce, R. (2000). *Eur. J. Biochem.* **267**, 2890–2898.

Fujiwara, K., Okamura-Ikeda, K. & Motokawa, Y. (1984). *J. Biol. Chem.* **259**, 10664–10668.

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.

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

Li de la Sierra, I., Pernot, L., Prange, T., Saludjian, P., Schiltz, M., Fourme, R. & Padron, G. (1997). *J. Mol. Biol.* **269**, 129–141.

Mande, S. S., Sarfaty, S., Allen, M. D., Perham, R. N. & Hol, W. G. J. (1996). *Structure*, **4**, 277–286.

Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J. & Hol, W. G. (1993). *J. Mol. Biol.* **230**, 1200–1215.

Mattevi, A., Obmolova, G., Sokatch, J. R., Betzel, C. & Hol, W. G. (1992). *Proteins*, **13**, 336–351.

Mattevi, A., Schierbeck, A. J. & Hol, W. G. (1991). *J. Mol. Biol.* **220**, 975–994.

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

Motokawa, Y., Fujiwara, K. & Okamura-Ikeda, K. (1995). *Bithiols in Health and Disease*, edited by L. Packer & E. Cadenas, pp. 389–407. New York: Marcel Dekker.

Motokawa, Y. & Kikuchi, G. (1974). *Arch. Biochem. Biophys.* **164**, 624–633.

Nakai, T., Ishijima, J., Masui, R., Kuramitsu, S. & Kamiya, N. (2003). *Acta Cryst.* D**59**, 1610–1618.

Nakai, T., Nakagawa, N., Maoka, N., Masui, R., Kuramitsu, S. & Kamiya, N. (2003). *Acta Cryst.* D**59**, 554–557.

Okamura-Ikeda, K., Fujiwara, K. & Motokawa, Y. (1987). *J. Biol. Chem.* **262**, 6746–6749.

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

Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M. & Douce, R. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 4850–4853.

Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M. & Douce, R. (1995). *Acta Cryst.* D**51**, 1041–1051.

Perham, R. N. (2000). *Annu. Rev. Biochem.* **69**, 961–1004.

Sugahara, M. & Miyano, M. (2002). *Tanpakushitsu Kakusan Koso*, **47**, 1026–1032.

Tada, K. & Hayasaka, K. (1987). *Eur. J. Pediatr.* **146**, 221–227.

Toone, J. R., Applegarth, D. A., Coulter-Mackie, M. B. & James, R. E. (2001). *Mol. Genet. Metab.* **72**, 322–325.

Toyoda, T., Kobayashi, R., Sekiguchi, T., Koike, K., Koike, M. & Takenaka, A. (1998). *Acta Cryst.* D**54**, 982–985.

Toyoda, T., Suzuki, K., Sekiguchi, T., Reed, L. J. & Takenaka, A. (1998). *J. Biochem.* **123**, 668–674.