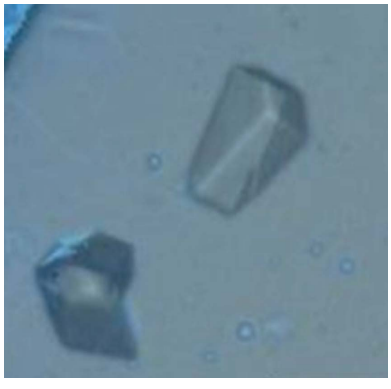
Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Katsumi Shimizu and Naoki Kunishima*

Advanced Protein Crystallography Research Group, RIKEN SPring-8 Center, Harima Institute, 1-1-1 Kouto, Sayo-cyo, Sayo-gun, Hyogo 679-5148, Japan

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

Received 14 December 2006
Accepted 2 March 2007



Purification, crystallization and preliminary X-ray diffraction study on pyrimidine nucleoside phosphorylase TTHA1771 from *Thermus thermophilus* HB8

Pyrimidine nucleoside phosphorylase (PYNP) catalyzes the reversible phosphorolysis of pyrimidines in the nucleotide-synthesis salvage pathway. In order to study the structure–thermostability relationship of this enzyme, PYNP from the extreme thermophile *Thermus thermophilus* HB8 (TTHA1771) has been cloned, overexpressed and purified. The TTHA1771 protein was crystallized at 291 K using the oil-microbatch method with PEG 4000 as a precipitant. A native data set was collected to 1.8 Å resolution using synchrotron radiation. The crystal belongs to the monoclinic space group $P2_1$, with unit-cell parameters a = 58.83, b = 76.23, c = 103.86 Å, β = 91.3°.

1. Introduction

Pyrimidine nucleoside phosphorylase (PYNP; EC 2.4.2.2) is an enzyme that catalyzes the reversible phosphorolysis of pyrimidine nucleosides, yielding pyrimidine and ribose-1-phosphate or deoxyribose-1-phosphate. PYNP can accept most pyrimidine nucleosides as substrates, except for 4-amino-substituted species such as deoxycytidine. PYNP plays an important role in the catalysis of glycosidic bond cleavage in pyrimidine nucleotides by the phosphorolytic mechanism, providing the reuse of free bases in the nucleotide-synthesis salvage pathway. Higher order organisms including mammals and some bacteria such as *Escherichia coli* have two distinct PYNPs: thymidine phosphorylase (TP; EC 2.4.2.4) and uridine phosphorylase (UP; EC 2.4.2.3) (Gallo *et al.*, 1967; Krenitsky *et al.*, 1965; Paege & Schlenk, 1952). However, in some lower order organisms such as *Bacillus stearothermophilus* (Saunders *et al.*, 1969) and *Haemophilus influenzae* (Scocca, 1971), only one PYNP is found and it does not discriminate between thymidine and uridine. Control of thymidine levels *in vivo* is thought to be critical for cell proliferation and hence inhibitors of TP have received considerable attention as potential cytotoxic antitumour agents (Schwartz & Milstone, 1988; Kirkwood *et al.*, 1980).

The structure of TP from *E. coli* was initially determined at 2.8 Å resolution (Walter *et al.*, 1990). The tetragonal crystals showed an S-shaped homodimer in which the subunits were related by a crystallographic twofold axis. The monomeric protomer was composed of two domains: an α-helical domain which was composed of six helices and an α/β-domain which included a six-stranded mixed β-sheet and a four-stranded antiparallel β-sheet. Similar S-shaped dimeric structures were observed in the homologous enzymes TP from *Homo sapiens* (Norman *et al.*, 2004) and PYNP from *B. stearothermophilus* (Pugmire & Ealick, 1998). For *E. coli* TP, two other crystal forms were observed: a monoclinic form and an orthorhombic form (Pugmire *et al.*, 1998). Comparison of the *E. coli* TP structures in the three different crystal forms revealed that domain movement occurs upon ligand binding. However, further studies using liganded enzymes are likely to contribute additional mechanistic insights. In addition, the structure–thermostability relationship of this enzyme remains to be elucidated, since most studies to date have focused on enzymes from mesophilic and mesothermophilic organisms. Here, we present the expression, purification and crystallization of a PYNP orthologue, the TTHA1771 protein from the extreme thermophile *Thermus thermophilus* HB8, which has an optimum growth temperature of 348 K. TTHA1771 protein shares 50% (220/433) sequence identity with

Table 1
Statistics of data collection.

Values in parentheses are for the highest resolution shell.

Wavelength (Å)	1.0
Unit-cell parameters (Å, °)	a = 58.83, b = 76.23, c = 103.86, β = 91.3
Space group	$P2_1$
Resolution range (Å)	40.0–1.8 (1.86–1.8)
No. of unique reflections	84627 (8406)
Completeness (%)	99.5 (99.4)
R_{merge} (%)	8.2 (28.2)
Redundancy	3.6 (3.5)
$\langle I/\sigma(I)\rangle$	8.9 (2.0)

B. stearothermophilus PYNP (PDB code 1brw), 44% (185/420) with *H. sapiens* TP (PDB code 1uou) and 41% (173/420) with *E. coli* TP (PDB code 1tpt). The structural determination of the TTHA1771 protein and structural comparison with homologous enzymes may provide insight into the reaction and thermostabilization mechanism of this family of enzymes.

2. Experimental

2.1. Protein expression and purification

The gene encoding TTHA1771 protein was cloned from the *T. thermophilus* HB8 genome (Yokoyama *et al.*, 2000). The TTHA1771 protein is composed of 423 amino acids (residues 1–423) with a molecular weight of 45.4 kDa. Protein expression and purification were performed by the Structurome Research Group at the RIKEN SPring-8 Center as follows using a routine procedure. *E. coli* BL21 (DE3) cells were transformed with the recombinant expression plasmid pET-11a carrying the TTHA1771 gene 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 20 000*g* for 4 min and suspended in 20 m*M* Tris–HCl pH 8.0 (buffer *A*) containing 500 m*M* NaCl, 5 m*M* 2-mercaptoethanol and 1 m*M* phenylmethylsulfonyl fluoride. The cells obtained were disrupted by sonication and heated at 343 K for 13 min. After the heat treatment, cell debris and denatured proteins were removed by centrifugation (21 600*g*, 30 min) and the supernatant was used as the crude extract for purification. The crude extract was desalted using a HiPrep 26/10 desalting column (Amersham Biosciences) and applied onto a Super Q Toyopearl 650M column (Tosoh) equilibrated with buffer *A*. Proteins were eluted with a linear gradient of 0–0.3 *M* NaCl. After buffer replacement with buffer *A*, the fraction containing TTHA1771 was subjected to a Resource Q column (Amersham Biosciences) equilibrated with buffer *A* and eluted with a linear gradient of 0–0.3 *M* NaCl in buffer *A*. After buffer replacement with 10 m*M* phosphate–NaOH pH 7.0 (buffer *B*), the fraction containing TTHA1771 was applied onto a Bio-Scale CHT-20-I column (Bio-Rad) equilibrated with buffer *B* and eluted with a linear gradient of 10–100 m*M* NaCl in buffer *B*. The fraction containing TTHA1771 was concentrated by ultrafiltration (Vivaspin, 10 kDa cutoff, Vivascience) 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 and N-terminal sequence analysis. Finally, the purified TTHA1771 was concentrated to 21.7 mg ml^{-1} by ultrafiltration and kept at 203 K.

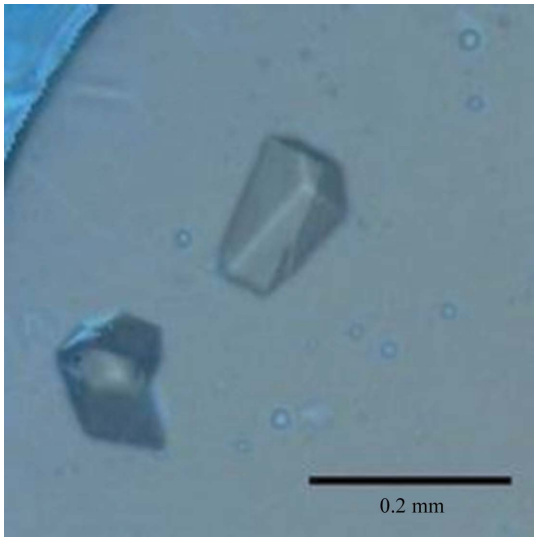


Figure 1
Optical micrograph of TTHA1771 crystals obtained by the oil-microbatch method. The scale bar represents 0.2 mm.

2.2. Dynamic light scattering

The oligomeric state of the purified TTHA1771 was examined by a dynamic light-scattering (DLS) experiment using a DynaPro MS/X instrument (Protein Solutions). The concentration of the protein solution was 20.0 mg ml^{-1} in 20 m*M* Tris–HCl buffer pH 7.6 with 200 m*M* NaCl. Several measurements were taken at 291 K and analyzed using the *DYNAMICS* software v.3.30 (Protein Solutions). A particle-size distribution with a polydispersity of 18.7% was observed and the molecular weight was estimated to be 86.3 kDa, which is consistent with a dimeric state of this protein in solution.

2.3. Crystallization and X-ray diffraction study

An initial screening of crystallization conditions was carried out using the TERA automatic crystallization system (Sugahara & Miyano, 2002), which is based on the oil-microbatch method (Chayen *et al.*, 1990), using Nunc HLA plates (Nalge Nunc International). After the initial screening, the crystallization conditions for TTHA1771 protein were optimized. The optimized precipitant solution consisted of 27.5%(*w*/*v*) PEG 4000 as precipitant, 0.1 *M* calcium chloride as additive and 0.1 *M* HEPES–NaOH pH 7.5 as buffer. A 0.5 µl aliquot of the optimized precipitant solution was mixed with 0.5 µl protein solution (21.7 mg ml^{-1} protein, 0.2 *M* NaCl, 20 m*M* Tris–HCl pH 8.0) in a well of the HLA plate. Subsequently, the 1 µl crystallization drop was covered with 15 µl of a 7:3(*v*/*v*) paraffin oil:silicon oil mixture, allowing slow evaporation of water in the drop, and stored at 291 K. Crystals with suitable dimensions for X-ray diffraction study were reproducibly obtained using the TERA system. The crystals were flash-cooled under a nitrogen-gas stream at 100 K using an oil-based cryoprotectant comprising 90% Paratone-N and 10% glycerol by weight (Kwong & Liu, 1999). X-ray diffraction data were collected using a Rigaku Jupiter 210 CCD detector at BL26B1 of SPring-8, Japan (Ueno *et al.*, 2006). Diffraction images were integrated and scaled using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997).

3. Results

We have established the expression, purification and crystallization of the TTHA1771 protein. After one week at 291 K, monoclinic crystals grew with typical dimensions of 0.13 × 0.13 × 0.07 mm (Fig. 1). The crystals diffracted X-rays to 1.8 Å resolution at 100 K using

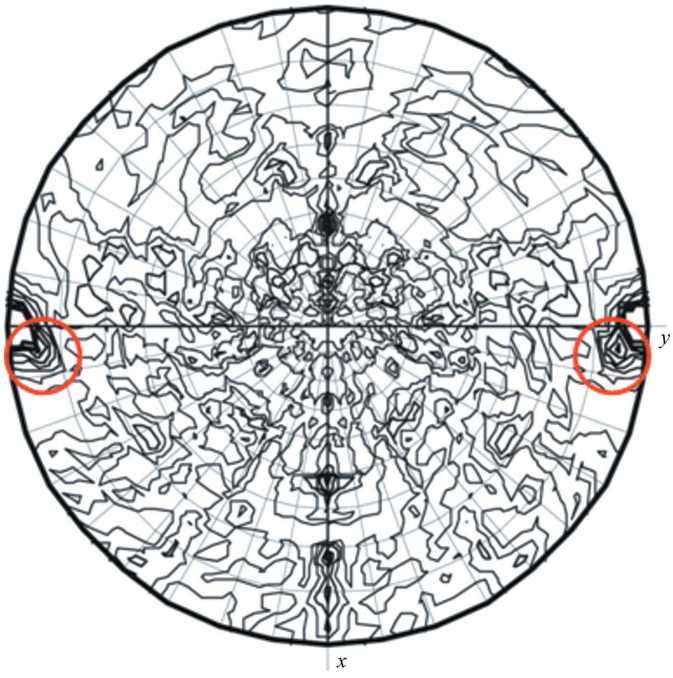


Figure 2
Projection map of the self-rotation function at the $\chi = 180^\circ$ section. The radius of integration was 20 Å and the higher resolution limit was 3 Å. On the map, contour lines are drawn at 0.5σ from a lowest level of $+0.5\sigma$. A strong signal arising from twofold noncrystallographic symmetry is observed at $\theta = 84.4$, $\varphi = 84.9^\circ$ and is indicated as red circles. The plot was generated with the help of *MOLREP*.

synchrotron radiation. Statistics of the native diffraction data are shown in Table 1. The crystals belonged to space group $P2_1$. Assuming the presence of two chains of TTHA1771 in the asymmetric unit gives a reasonable value of 2.6 $Å^3$ Da^{-1} for the Matthews coefficient V_M (Matthews, 1968), corresponding to a solvent content of 51.7%, which is consistent with the result of the DLS experiment, which showed a dimeric state of this protein in solution (see §2). To examine the local symmetry in the asymmetric unit, self-rotation functions were calculated using *MOLREP* (Collaborative Computational Project, Number 4, 1994; Vagin & Teplyakov, 1997; Fig. 2). The search was carried out for χ angles of 180, 120, 90 and 60° to check for the presence of twofold, threefold, fourfold and sixfold axes, respectively. A local peak with a signal-to-noise ratio of 4.5σ was found at $\chi = 180^\circ$, demonstrating the existence of twofold noncrystallographic symmetry at $\theta = 84.4$, $\varphi = 84.9^\circ$. We have obtained a clear molecular-replacement solution using the coordinates of the *B. stearothermophilus* PYNP (PDB code 1brw), which shows the highest sequence similarity to TTHA1771 in the PDB. The molecular-replacement result confirmed that the asymmetric unit contains a dimer with local pseudo-twofold symmetry. The pseudo-twofold axis of the dimer could be calculated by the superposition of one subunit onto the other using *LSQKAB* (Kabsch, 1976) as having spherical polar angles $\omega = 84.6$, $\varphi = 84.4$, $\chi = 175.7^\circ$, which was consistent with the result of the self-rotation function. A detailed discussion of the crystal structure will be published elsewhere.

The authors thank the staff of the RIKEN Structurome Research Group for providing the plasmid, the technical staff of the RIKEN SPring-8 Center for large-scale protein production and dynamic light-scattering experiments and the beamline staff at BL26B1 of SPring-8 for assistance during data collection. This work (TTHA1771/HTPF00532) was supported by the 'National Project on Protein Structural and Functional Analyses' funded by the MEXT of Japan.

References

Chayen, N. E., Shaw Stewart, P. D., Maeder, D. L. & Blow, D. M. (1990). *J. Appl. Cryst.* **23**, 297–302.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Gallo, R. C., Perry, S. & Breitman, T. R. (1967). *J. Biol. Chem.* **242**, 5059–5068.
Kabsch, W. (1976). *Acta Cryst.* A**32**, 922–923.
Kirkwood, J. M., Ensminger, W., Rosowsky, A., Papathanasopoulos, N. & Frei, E. (1980). *Cancer Res.* **40**, 107–113.
Krenitsky, T. A., Mellors, J. W. & Barclay, R. K. (1965). *J. Biol. Chem.* **240**, 1281–1286.
Kwong, P. D. & Liu, Y. (1999). *J. Appl. Cryst.* **32**, 102–105.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Norman, R. A., Barry, S. T., Bate, M., Breed, J., Colls, J. G., Ernill, R. J., Luke, R. W. A., Minshull, C. A., McAlister, M. S. B., McCall, E. J., McMiken, H. H. J., Paterson, D. S., Timms, D., Tucker, J. A. & Pauptit, R. A. (2004). *Structure*, **12**, 75–84.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Paege, L. M. & Schlenk, F. (1952). *Arch. Biochem. Biophys.* **40**, 42–49.
Pugmire, M. J., Cook, W. J., Jasanoff, A., Walter, M. R. & Ealick, S. E. (1998). *J. Mol. Biol.* **281**, 285–299.
Pugmire, M. J. & Ealick, S. E. (1998). *Structure*, **6**, 1467–1479.
Saunders, P. P., Wilson, B. A. & Saunders, G. F. (1969). *J. Biol. Chem.* **244**, 3691–3697.
Schwartz, P. M. & Milstone, L. M. (1988). *Biochem. Pharmacol.* **37**, 353–355.
Scocca, J. J. (1971). *J. Biol. Chem.* **246**, 6606–6610.
Sugahara, M. & Miyano, M. (2002). *Tanpakushitsu Kakusan Koso*, **47**, 1026–1032.
Ueno, G., Kanda, K., Hirose, R., Ida, K., Kumasaka, T. & Yamamoto, Y. (2006). *J. Struct. Funct. Genomics*, **7**, 15–22.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Walter, M. R., Cook, W. J., Cole, L. B., Short, S. A., Koszalka, G. W., Krenitsky, T. A. & Ealick, S. E. (1990). *J. Biol. Chem.* **265**, 14016–14022.
Yokoyama, S., Hirota, H., Kigawa, T., Yabuki, T., Shirouzu, M., Terada, T., Ito, Y., Matsuo, Y., Kuroda, Y., Nishimura, Y., Kyogoku, Y., Miki, K., Masui, R. & Kuramitsu, S. (2000). *Nature Struct. Biol.* **7**, 943–945.