Santosh Kumar Chaudhary,^a
Jeyaraman Jeyakanthan^b and
Kanagaraj Sekar^{c*}

^aDepartment of Physics, Indian Institute of
Science, Bangalore 560 012, India,

^bDepartment of Bioinformatics, Alagappa
University, Karaikudi 630 003, India, and

^cSupercomputer Education and Research Centre,
Indian Institute of Science, Bangalore 560 012,
India

Correspondence e-mail:
sekar@physics.iisc.ernet.in,
sekar@serc.iisc.ernet.in

Received 12 November 2012

Accepted 8 December 2012

Cloning, expression, purification, crystallization and preliminary X-ray crystallographic study of thymidylate kinase (TTHA1607) from *Thermus thermophilus* HB8

Nucleotide biosynthesis plays a key role in cell survival and cell proliferation. Thymidylate kinase is an enzyme that catalyses the conversion of dTMP to dTDP using ATP-Mg²⁺ as a phosphoryl-donor group. This enzyme is present at the junction of the *de novo* and salvage pathways; thus, any inhibitor designed against it will result in cell death. This highlights the importance of this enzyme as a drug target. Thymidylate kinase from the extremely thermophilic organism *Thermus thermophilus* HB8 has been expressed, purified and crystallized using the microbatch method. The crystals diffracted to a resolution of 1.83 Å and belonged to the orthorhombic space group *P*2₁2₁2₁, with unit-cell parameters *a* = 39.50, *b* = 80.29, *c* = 122.55 Å. Preliminary studies revealed the presence of a dimer in the asymmetric unit with a Matthews coefficient (*V*_M) of 2.18 Å³ Da⁻¹.

1. Introduction

Thermus thermophilus HB8 is a thermophilic, aerobic Gram-negative eubacterium with an optimum growth temperature ranging from 338 to 345 K. It was isolated from thermal vents in the hot springs of Izu, Japan. The proteins present in *T. thermophilus* are thermally stable compared with their mesophilic homologues (Oshima & Imahori, 1974; Yoshida & Oshima, 1971). This makes them industrially important and ideal tools for structural biologists. The present work is based on an enzyme, thymidylate kinase, from *T. thermophilus* HB8. The enzyme is involved in the nucleotide-biosynthesis pathway which is essential for cell survival and proliferation. Nucleotides (purine and pyrimidine) are biosynthesized by either the *de novo* pathway or the salvage pathway. The enzyme thymidylate kinase catalyses the reversible conversion of deoxythymidine monophosphate (dTMP) to deoxythymidine diphosphate (dTDP) in the presence of ATP-Mg²⁺ which acts as a phosphoryl donor in the reaction. In the *de novo* pathway of dTMP synthesis, either the phosphorolysis of dUTP or deamination of dCMP yields dUMP, which is further converted to dTMP by thymidylate synthase (Reichard, 1988). In the salvage pathway of dTMP synthesis, thymidine kinase synthesizes dTMP using thymidine as a substrate (Arnér & Eriksson, 1995). In later steps thymidylate kinase converts dTMP to dTDP and subsequently dTDP is phosphorylated to dTTP by dNDP kinase for DNA synthesis (Reichard, 1988). Therefore, thymidylate kinase occurs at the junction of the *de novo* and salvage pathway and plays a vital role in dTTP synthesis; thus, it is an important drug target for cancer, bacterial and viral diseases (Anderson, 1973; Neuhaud & Nygaard, 1987).

The three-dimensional crystal structures of thymidylate kinase from various sources have been solved using X-ray crystallography and include structures from *Staphylococcus aureus* (PDB entry 4f4i; Midwest Center for Structural Genomics, unpublished work), *Burkholderia thailandensis* (PDB entry 3v9p; Seattle Structural Genomics Center for Infectious Disease, unpublished work) and *Ehrlichia chaffeensis* (PDB entry 3ld9; Leibly *et al.*, 2011). The crystal structures of the two native thymidylate kinases from extremely thermophilic organisms, *Aquifex aeolicus* (PDB entry 2pbr; RIKEN Structural Genomics/Proteomics Initiative, unpublished work) and

Sulfolobus tokodaii (PDB entry 2plr; RIKEN Structural Genomics/Proteomics Initiative, unpublished work) have been submitted to the PDB by our group. The crystal structures of thymidylate kinase in complex with various native ligands and inhibitors such as dTDP, dTMP, ADP, AZTMP, TP5A, AppNHp, FLTMP, APPNHp, NH2TMP and (*E*)-5-(2-bromovinyl)-2'-deoxyuridine-5'-monophosphate have also been investigated using X-ray crystallography. Most of the enzymes studied have an almost similar active-site topology but display different catalytic activity. The sequence alignment (Fig. 1) of thymidylate kinase from *T. thermophilus* (198 amino acids, ~22 kDa) shows an identity of 23.6% to the enzyme from *Homo sapiens*, 29% to that from *Mycobacterium tuberculosis*, 28% to that from *Saccharomyces cerevisiae*, 20.8% to that from *Vaccinia virus* and 44.6% to that from *A. aeolicus*.

Some substrate analogues of thymidylate kinase bind competitively to the enzyme and are converted to the corresponding diphosphate form and subsequently incorporated into DNA. This incorporation results in chain termination of the newly synthesized DNA strand (Furman *et al.*, 1986). It has been found that the conversion of the monophosphate substrate analogue to its diphosphate form is the rate-limiting step which results in the accumulation of the monophosphate form in the cell (Frick *et al.*, 1988; Fridland *et al.*, 1990). Thus, there is a need to design better substrate analogues than those currently available. Thymidylate kinase from *T. thermophilus* may be a suitable system for developing better DNA base analogues for targeting cancer cells. Insights into the reaction mechanism or strategies for drug design based on this enzyme may be applicable to the enzyme from pathogenic microorganisms.

Figure 1
Multiple sequence alignment of thymidylate kinase from *T. thermophilus*, *A. aeolicus*, *S. cerevisiae*, *M. tuberculosis*, *H. sapiens* and *Vaccinia virus*.

2. Materials and methods

2.1. Cloning, expression and purification

The thymidylate kinase gene (TTHA1607) was amplified from *T. thermophilus* HB8 genome using the primers 5'-ATATCATA-TGCCGGGGCTCTTCCTCACCTCGA-3' and 5'-ATATGGATCC-TTATTATGGCAGAAGGGGCCGGAGGTG-3' by the polymerase chain reaction. The amplified fragment was cloned into the *Escherichia coli* expression vector pET11a (Novagen, Madison, Wisconsin, USA). The plasmid was transformed into *E. coli* Rosetta(DE3) (Novagen) cells for protein expression. The transformants were cultured overnight at 310 K in 2 l LB medium containing 100 µg ml⁻¹ ampicillin. The cells were harvested by centrifugation and resuspended in a buffer consisting of 20 mM Tris-HCl pH 8 and 50 mM NaCl, and lysed by sonication. The cell lysate was heated in a dry bath at 343 K for 15 min and the denatured proteins and other cell debris (from the cell lysate) were then separated by centrifugation at 15 000g for 45 min at 277 K. The supernatant was desalted using a Sephadex G-25 (GE Healthcare Biosciences) column pre-equilibrated with 20 mM Tris-HCl pH 8. The desalted fraction was passed through a Sepharose S cation-exchange column (GE Healthcare Biosciences) pre-equilibrated with 20 mM Tris-HCl buffer pH 8 and the protein was eluted using a linear gradient of 0–1.0 M NaCl in 20 mM Tris-HCl pH 8. The fraction containing the protein was dialysed against a buffer consisting of 20 mM Tris-HCl, 50 mM NaCl pH 7.4. The dialysed protein sample was further concentrated using a Centricon device (10 kDa molecular weight cutoff). The protein concentration was estimated by measuring the absorbance at 280 nm. The yield of the purified protein was 10 mg per litre of culture.

2.2. Crystallization experiments

The concentration of the protein solution used for crystallization of the enzyme was 5 mg ml⁻¹ in 20 mM Tris-HCl pH 7.4, 50 mM NaCl. A preliminary screening of the crystallization conditions was performed using Hampton Research Crystal Screen and Crystal Screen 2 by the microbatch-under-oil (Greiner plate) technique. The drops contained 2 µl protein solution and 2 µl crystallization condition. 7 ml of a 1:1 mixture of silicone oil and paraffin oil (1:1) were used in the plate. Diffraction-quality crystals (Fig. 2) were obtained after 2 weeks in the Crystal Screen condition No. 6 consisting of 0.2 M magnesium chloride hexahydrate, 0.1 M Tris-HCl pH 8.5, 30% (w/v) PEG 4000 at 293 K.

Figure 2
Crystal used for X-ray data collection.

Table 1

X-ray data-collection statistics.

Values in parentheses are for the last resolution bin.

Crystal habit	Primitive orthorhombic
Wavelength (Å)	1.83
Temperature (K)	100
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	$a = 39.50, b = 80.29, c = 122.55$
Matthews coefficient (Å ³ Da ⁻¹)	2.18
Solvent content (%)	43.6
No. of molecules in asymmetric unit	2
Resolution range (Å)	40.83–1.83 (1.92–1.83)
No. of observed reflections	251551
No. of unique reflections	33877 (4078)
Completeness (%)	95.5 (80.2)
$R_{\text{merge}}^{\dagger}$ (%)	7.2 (57.2)
Multiplicity	7.4 (7.4)
Average $\langle I/\sigma(I) \rangle$	17 (3.3)
Wilson B factor (Å ²)	23

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the observed intensity and $\langle I(hkl) \rangle$ is the average intensity for multiple measurements.

2.3. Data collection

A protein crystal was soaked briefly (10 s) in a cryoprotectant solution consisting of mother liquor supplemented with 20% (v/v) glycerol, picked up using a nylon loop and flash-cooled at 100 K in an N₂ cold stream (McFerrin & Snell, 2002). The diffraction data were collected at 100 K using a MAR 345 imaging-plate detector mounted on a Rigaku UltraX-18 rotating-anode X-ray generator (operated at 40 kV and 80 mA) using the home source available at the Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India. The crystal-to-detector distance was 145 mm and the crystal diffracted to a resolution of 1.83 Å. The diffraction data were processed using *iMOSFLM* (Battye *et al.*, 2011) and scaled using *SCALA* (Evans, 2006) in the *CCP4* program suite (Winn *et al.*, 2011). Details of data collection are summarized in Table 1.

3. Results and discussion

The thymidylate kinase from *T. thermophilus* consists of 198 amino acids and has an approximate molecular mass of 22 kDa. A data set was collected from a native protein crystal to 1.83 Å resolution at 100 K. A total of 251 551 reflections are merged into 33 877 unique reflections with an overall R_{merge} of 7.2% and a completeness of 95.5%. The crystal belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters $a = 39.50, b = 80.29, c = 122.55$ Å. The Matthews coefficient V_M (Matthews, 1968) was found to be 2.18 Å³ Da⁻¹, which suggests the presence of two molecules in the asymmetric unit with a solvent content of 43.6%. The crystal structure of the *T. thermophilus* thymidylate kinase was solved by molecular replacement using the program *MOLREP* (Vagin & Teplyakov, 2010) available in the *CCP4* program suite. The three-dimensional atomic coordinates of the thymidylate kinase (PDB entry 2pbr, RIKEN Structural Genomics/Proteomics Initiative, unpublished work) from *A. aeolicus* were used as a search model for the molecular-replacement calculations (using default parameters). The resultant output model was further subjected to restrained refinement using *REFMAC5* (Murshudov *et al.*, 2011). A total of 5% of the reflections were used for R_{free} calculations (Brünger, 1992). The R_{work} and R_{free} values of the partially refined model were 43.5 and 46.9%, respectively. Further refinement of the model and the map fitting using *Coot* (Emsley & Cowtan, 2004) are in progress.

The authors (SKC and KS) thank the Interactive Graphics Based Molecular Modeling Facility and the Supercomputer Education and

Research Centre. The authors thank the RIKEN Structural Genomics/Proteomics Initiative (RSGI), the National Project on Protein Structural and Functional Analyses, Ministry of Education, Culture, Sports, Science and Technology of Japan for providing the clone.

References

- Anderson, E. (1973). *The Enzymes*, 3rd ed., edited by P. D. Boyer, Vol. 9, pp. 49–96. New York: Academic Press.
- Arnér, E. S. & Eriksson, S. (1995). *Pharmacol. Ther.* **67**, 155–186.
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
- Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Frick, L. W., Nelson, D. J., St Clair, M. H., Furman, P. A. & Krenitsky, T. A. (1988). *Biochem. Biophys. Res. Commun.* **154**, 124–129.
- Fridland, A., Connelly, M. C. & Ashmun, R. (1990). *Mol. Pharmacol.* **37**, 665–670.
- Furman, P. A., Fyfe, J. A., St Clair, M. H., Weinhold, K., Rideout, J. L., Freeman, G. A., Lehrman, S. N., Bolognesi, D. P., Broder, S., Mitsuya, H. & Barry, D. W. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 8333–8337.
- Leibly, D. J., Abendroth, J., Bryan, C. M., Sankaran, B., Kelley, A., Barrett, L. K., Stewart, L. & Van Voorhis, W. C. (2011). *Acta Cryst.* **F67**, 1090–1094.
- McFerrin, M. B. & Snell, E. H. (2002). *J. Appl. Cryst.* **35**, 538–545.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Neuhard, J. & Nygaard, P. (1987). *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, edited by F. C. Neidhardt, J. L. Ingraham, B. K. Low, B. Magasanik, M. Schaechter & H. E. Umbarger, pp. 445–473. Washington DC: American Society for Microbiology.
- Oshima, T. & Imahori, K. (1974). *Int. J. Syst. Bacteriol.* **24**, 102–112.
- Reichard, P. (1988). *Annu. Rev. Biochem.* **57**, 349–374.
- Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* **D66**, 22–25.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yoshida, M. & Oshima, T. (1971). *Biochem. Biophys. Res. Commun.* **45**, 495–500.