Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Cloning, expression and crystallization of pyrimidine nucleoside phosphorylase from *Bacillus stearothermophilus*

Min Zhou,[a]† Matthew J. Pugmire,[b] Bao Q. Vuong[a]‡ and Steven E. Ealick[a]*

[a]Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853, USA, and [b]Department of Chemistry, Cornell University, Ithaca, NY 14853, USA

† Present address: Alexander and Parker Corporation, 2665 Long Lake Road, Suite 330, Roseville, MN 55113, USA.
‡ Present address: Department of Microbiology, Columbia University, New York, NY 10032, USA.

Correspondence e-mail: see3@cornell.edu

Received 13 April 1998
Accepted 19 June 1998

Pyrimidine nucleoside phosphorylase (PYNP) from *B. stearothermophilus* has been cloned and purified for crystallization. Crystals of a potential protein–inhibitor complex have been prepared by co-crystallization techniques using the substrate analog pseudouridine. These crystals provide good-quality diffraction images to 2.7 Å and belong to space group $P2_1$. The asymmetric unit contains the dimer structure of PYNP with unit-cell parameters a = 53.9, b = 71.9, c = 123.3 Å and β = 96.9°.

1. Introduction

Pyrimidine nucleoside phosphorylase catalyzes the reversible phosphorolysis of pyrimidine nucleosides, which is an important step in the salvage pathway of nucleotide synthesis (E.C. 2.4.2.4). Mammalian and other higher organisms have two separate PYNP enzymes, one that is specific for thymidine (thymidine phosphorylase) and one that is specific for uridine (uridine phosphorylase). However, many lower organisms, such as *B. stearothermophilus* (Saunders *et al.*, 1969) and *Haemophilus influenzae* (Scocca, 1971), have a single PYNP that does not discriminate between uridine and thymidine. Such PYNP sequences are homologous to thymidine phosphorylase (TP) (~40% sequence identity) whereas there appears to be no relationship with the uridine phosphorylase sequences (<20% sequence identity). TP has traditionally been of interest due to its potential as a drug target in several chemotherapeutic strategies (Kirkwood *et al.*, 1980; Schwartz & Milstone, 1988; Schwartz *et al.*, 1995). Additional interest has surfaced with the recent findings that the sequence of human TP is identical to that of both platelet-derived endothelial cell-growth factor (Miyadera *et al.*, 1995) and gliostatin (Asai *et al.*, 1992). The fact that a single peptide (TP) functions as both an enzyme in an important metabolic pathway and as a growth factor for both endothelial cells and neurons makes this an exciting protein to study.

TP functions as a dimer which is made up of two identical subunits related to each other by a twofold axis of symmetry. The dimer molecular mass ranges from 90 kDa in *Escherichia coli* to 110 kDa in mammals (Degranges *et al.*, 1981; Schwartz, 1978). Structural details were provided by the 2.8 Å crystal structure of *E. coli* TP (Walter *et al.*, 1990) which showed each subunit to consist of a small α-helical domain separated by a cleft from a larger α/β-domain. It was also hypothesized that rigid-body domain movement, which would close the cleft, is necessary for catalysis to proceed. The crystal structure of TP in a fully closed conformation (presumably triggered by binding substrate), would prove the domain-closure hypothesis and provide insights into a possible closing mechanism. In an attempt to show this closed conformation, we report here the cloning and purification of PYNP from the thermophilic bacteria *B. stearothermophilus*, as well as the co-crystallization with the substrate-analog inhibitor pseudouridine.

2. Materials, methods and results

2.1. Cloning

B. stearothermophilus ATCC 12980 (American Type Culture Collection) genomic DNA was partially digested with *Sau*3A. The 2–3 kbp fragments were ligated into the pBluescriptSK+ (Stratagene) vector (which had been previously digested with *Bam*HI), and the resulting constructs were then transformed into *E. coli* DH5α cells. About 5000 white colonies were selected for on Xgal/ampicillin plates, and were pooled as a genomic library. These plasmid DNAs were then transformed into PNP-deficient *E. coli* cells SΦ312 (a gift from Dr Per Nygaard, Institute of Biological Chemistry, Copenhagen, Denmark) and positive clones were selected for on deoxyinosine/ampicillin/casamino acid M63 minimal plates. 27 colonies were examined for PNP activity using inosine or adenosine as substrates in a previously described assay (Hoffee *et al.*, 1978). The colonies were sequenced and two major clones identified: clone *A* and clone *B*. Both clones contained the PNP (purine nucleoside phosphorylase) and PYNP gene; however, the PNP-coding sequence of clone *B* lacked the first ten N-terminal amino acids. Digestion of clone *A* with *Pst*I removed the PNP-coding sequence to obtain the *E. coli* SΦ312 PYNP clone. Fig. 1

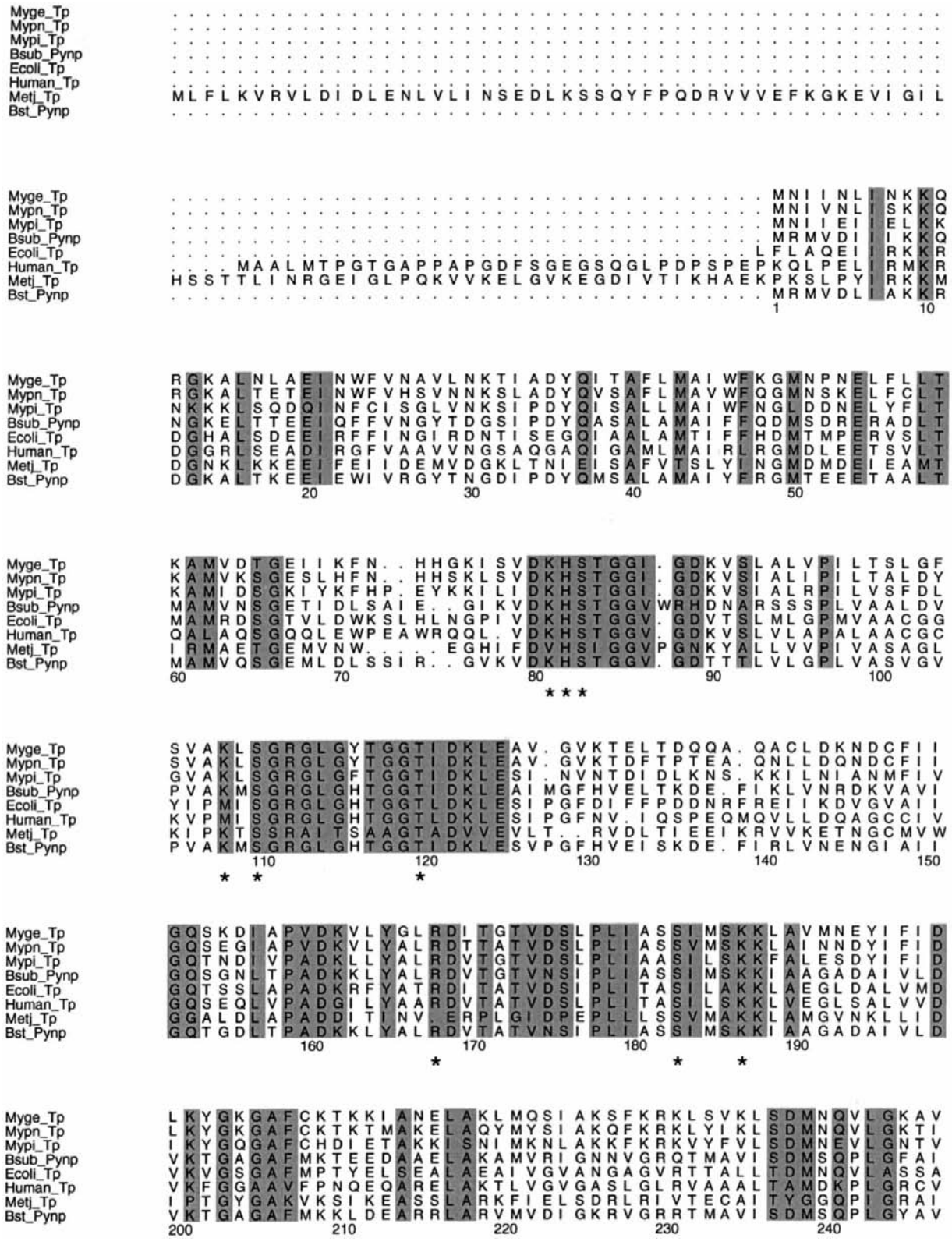

Figure 1
Multiple sequence alignment of eight thymidine phosphorylase sequences. Myge_Tp = *Mycoplasma genitalium* thymidine phosphorylase; Mypn_Tp = *M. pneumoniae* thymidine phosphorylase; Mypi_Tp = *M. pirum* thymidine phosphorylase; Bsub_Pynp = *Bacillus subtilis* pyrimidine nucleoside phosphorylase; E. coli_Tp = *Escherichia coli* thymidine phosphorylase; Human_Tp = human thymidine phosphorylase/platelet-derived endothelial cell-growth factor; Metj_Tp = *Methanococcus jannashii* thymidine phosphorylase; Bst_Pynp = *B. stearothermophilus* pyrimidine nucleoside phosphorylase. The numbering refers to the Bst_Pynp sequence. Highly conserved regions have been shaded, and '*' indicates putative active-site residues. This figure was prepared using *ALSCRIPT* (Barton, 1993).

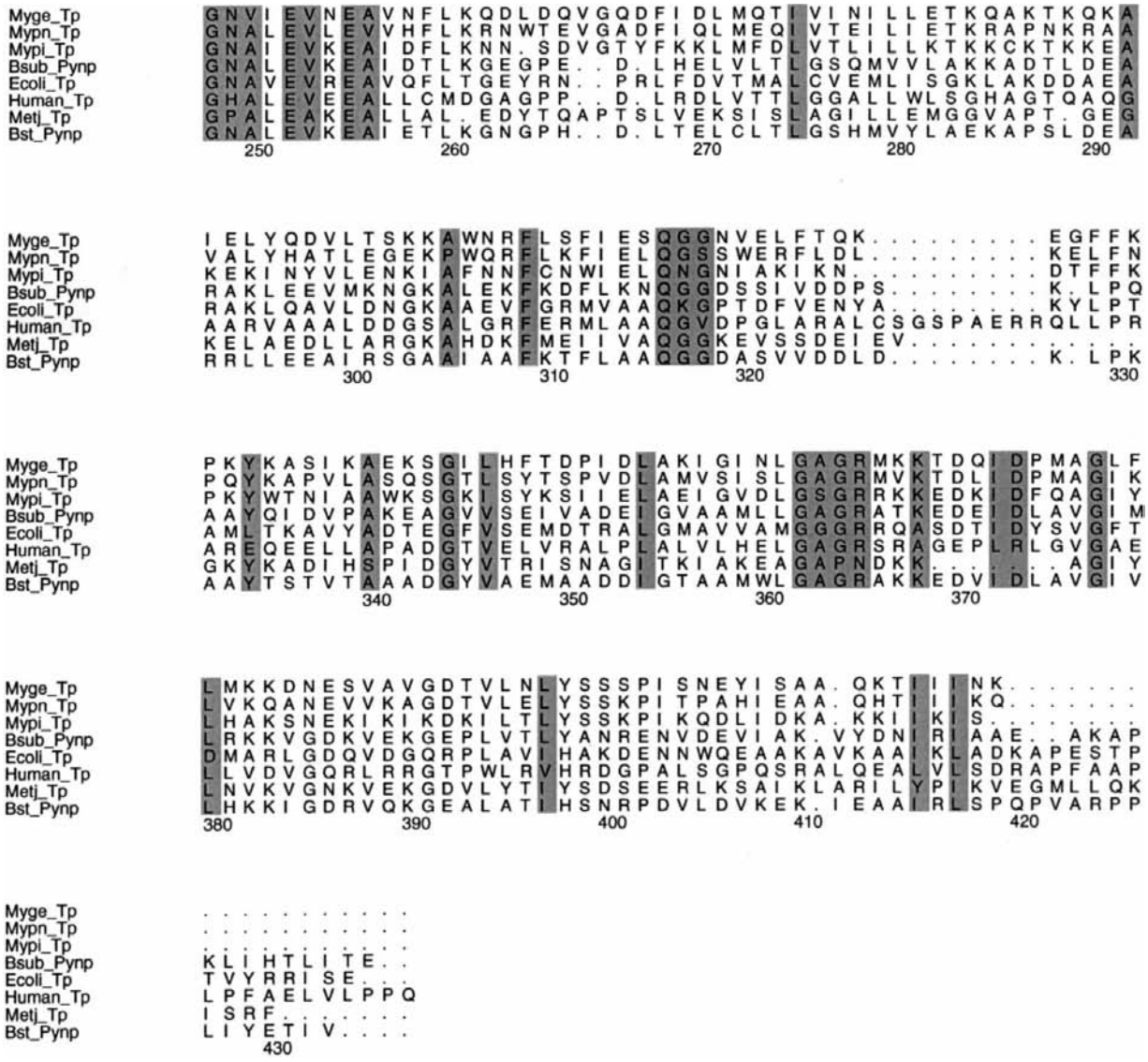

Figure 1 (continued)

shows the amino-acid sequence of this clone aligned with several homologous TP sequences.

2.2. Expression and protein purification

E. coli SΦ312 PYNP cells were grown in LB media, harvested and stored at 193 K. The cell pellet was resuspended in 3 ml 20 m*M* Tris buffer pH 7.0 per gram of wet cells on ice. The resuspended cells were passed through a French press twice. Cellular debris was removed by centrifugation at 17500*g* for 30 min at 277 K. The supernatant was removed and ammonium sulfate was slowly added to a final concentration of 35% at 277 K. The solution was allowed to stir for 30 min at 277 K and then centrifuged at 17500*g* for 10 min at 277 K to pellet the precipitated protein. The pellet was discarded and the supernatant was brought to a final ammonium sulfate concentration of 65% at 277 K and allowed to stir for 30 min. The precipitated protein was centrifuged at 17500*g* for 10 min at 277 K. The resulting pellet was resuspended in 10 ml (per 40 g of cell paste prepared) of 20 m*M* Tris pH 7.0 and dialyzed overnight against 2 l of the same buffer. The dialyzed sample was centrifuged at 17500*g* for 10 min at 277 K to remove precipitated protein that did not resolubilize. The soluble dialyzed sample was placed over a Q-Sepharose column (2.5 × 50 cm) equilibrated with 20 m*M* Tris pH 7.0. The column was washed with 1 column volume of 20 m*M* Tris pH 7.0 followed by 2 column volumes of 20 m*M* Tris pH 7.0, 100 m*M* NaCl. A linear gradient of 1 column volume of 100–500 m*M* NaCl in 20 m*M* Tris pH 7.0 was passed over the column. The PYNP eluted at approximately 200 m*M* NaCl. Active fractions were pooled and buffer-exchanged with a tenfold dilution (three times) into 20 m*M* Tris pH 7.0 in an Amicon pressure concentrator at 3.1 × 10^6 Pa, 277 K on a YM-10 membrane with gentle stirring. The desalted sample was placed over a Blue Sepharose 6 Fast Flow column (2.5 cm × 50 ml) and washed with 100 ml of 20 m*M* Tris pH 7.0. The PYNP was eluted from the column with 100 m*M* NaCl, 20 m*M* Tris pH 7.0. The final product was buffer-exchanged with a tenfold dilution (three times) into 10 m*M* potassium phosphate pH 7.0 in an Amicon pressure concentrator at 3.1 × 10^6 Pa, 277 K on a YM-10 membrane with gentle stirring, concentrated to 10 mg ml^{-1} and run on a 12% SDS–PAGE gel to examine protein purity.

Table 1
Data-collection statistics.

Values in parentheses refer to the highest resolution shell.

Resolution (Å)	39.3 (2.79–2.71)
Total number of reflections	61540
Number of unique reflections	21835
Completeness (%)	83.8 (52.3)
R_{sym} (%)	7.2 (23.9)

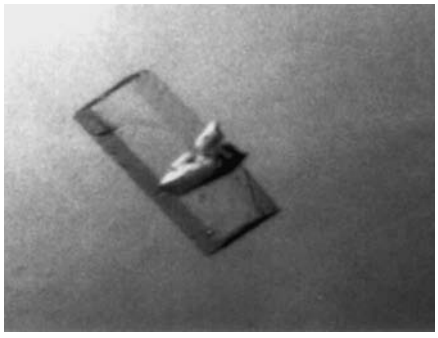

Figure 2
Photograph of PYNP co-crystal with pseudouridine grown by vapor diffusion.

2.3. Crystallization

Initial crystallization experiments consisted of hanging-drop vapor-diffusion methods at room temperature with 3 µl drops (2 µl protein plus 1 µl well solution), using a well solution consisting of 15% ammonium sulfate in 50 m*M* citrate buffer at pH 4.8 and 4 m*M* DTT. This resulted in a large number of small block-shaped crystals that were 50–100 µm on each edge but were unsuitable for diffraction experiments. In an attempt to obtain a protein–substrate-analog complex structure, the protein sample was next prepared for co-crystallization experiments by adding to it 0.175 ml of 10 m*M* pseudouridine (a non-cleavable substrate analog from Sigma Chemical Co.) for every 1 ml of protein sample. This ratio was such that the concentration of pseudouridine was ten times that of the protein. This new protein sample was allowed to sit for 24 h, before washing three times with 10 m*M* potassium phosphate buffer by spinning the sample in a centrifuge using Amicon Centricon concentrator tubes, each time bringing the protein concentration back to 10 mg ml^{-1}. New crystals were obtained by again using the hanging-drop vapor-diffusion method, with 4 µl drops (2 µl protein plus 2 µl well solution) method at room temperature, using a well solution consisting of 25% PEG 6000 in 0.1 *M* MES buffer at pH 6.0. This resulted in rectangular crystals that grew out of a precipitate in three to four weeks and were 250 µm in the longest dimension (see Fig. 2).

2.4. Crystallographic analysis

The crystals were mounted at room temperature in glass capillary tubes, and X-ray data were collected on a San Diego Multiwire area-detector system using Cu $K\alpha$ radiation from a Rigaku RU 200 rotating-anode generator operating at 40 kV and 100 mA. The crystals diffracted to a maximum resolution of 2.7 Å and were determined to be monoclinic with cell parameters $a = 53.9$, $b = 71.9$, $c = 123.3$ Å and $\beta = 96.9^\circ$. Precession photographs revealed systematic absences along $0k0$, with $k = 2n$ observed, indicating space group $P2_1$. Assuming a dimer in the asymmetric unit, a Matthews number (Matthews, 1968) of 2.57 Å^3 Da^{-1} and a solvent content of 52% were calculated. A complete data set was collected at room temperature (data-collection statistics are shown in Table 1). The data were processed with software provided by the San Diego Multiwire system (Xuong *et al.*, 1985) and scaled using *SCALA* from the *CCP*4 package (Collaborative Computational Project, Number 4, 1994). Although there is no direct evidence that the crystals contain a substrate-analog complex, the observation that crystal size and diffraction quality improved only after the co-crystallization with pseudouridine lead us to be hopeful of having obtained a PYNP-complex structure. A detailed structural analysis is currently under way.

The authors would like to acknowledge the generous support provided by the W. M. Keck Laboratory for Molecular Structure, the Lucille P. Markey Charitable Trust and an NIH Training Grant (T2GM08500A).

References

Asai, K., Hirano, T., Kaneko, S., Moriyama, A., Nakanishi, K., Isobe, I., Eksioglu, Y. Z. & Kato, T. (1992). *J. Neurochem.* **59**, 307–317.
Barton, G. J. (1993). *Protein Eng.* **6**, 37–40.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Degranges, C., Razaka, G., Rabaud, M. & Bricaud, H. (1981). *Biochim. Biophys. Acta*, **654**, 211–218.
Hoffee, P. A., May, R. & Robertson, B. C. (1978). *Methods Enzymol.* **51**, 517–524.
Kirkwood, J. M., Ensminger, W., Rosowsky, A., Papathanasopoulos, N. & Frei, E. (1980). *Cancer Res.* **40**, 107–113.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Miyadera, K., Dohmai, N., Takio, K., Sumizawa, T., Haraguchi, M., Furukawa, T., Yamada, Y. & Akiyama, S. (1995). *Biochem. Biophys. Res. Commun.* **212**, 1040–1045.
Saunders, P. P., Wilson, B. A. & Saunders, G. F. (1969). *J. Biol. Chem.* **244**, 3691–3697.
Schwartz, E. L., Baptiste, N., Wadler, S. & Makower, D. (1995). *J. Biol. Chem.* **270**, 19073–19077.
Schwartz, M. (1978). *Methods Enzymol.* **51**, 442–445.
Schwartz, P. M. & Milstone, L. M. (1988). *Biochem. Pharmacol.* **37**, 353–355.
Scocca, J. J. (1971). *J. Biol. Chem.* **246**, 6606–6610.
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.
Xuong, N. H., Nielsen, C., Hamlin, R. & Anderson, D. (1985). *J. Appl. Cryst.* **18**, 342–350.