ISSN 0907-4449

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

^aSection of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853, USA, and ^bDepartment 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

© 1999 International Union of Crystallography Printed in Great Britain – all rights reserved Received 13 April 1998 Accepted 19 June 1998

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

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 $\beta = 96.9^{\circ}$.

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 domainclosure 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 substrateanalog inhibitor pseudouridine.

2. Materials, methods and results

2.1. Cloning

B. stearothermophilus ATCC 12980 (American Type Culture Collection) genomic DNA was partially digested with Sau3A. The 2-3 kbp fragments were ligated into the pBluescriptSK+ (Stratagene) vector (which had been previously digested with BamHI), and the resulting constructs were then transformed into E. coli DH5a 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 SØ312 (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 PstI removed the PNP-coding sequence to obtain the E. coli SΦ312 PYNP clone. Fig. 1

crystallization papers

Myge_Tp Mypn_Tp Bsub_Pynp Ecoli_Tp Human_Tp Meti_Tp Bst_Pynp	MLF	LKV	RVI	 	b L	ĒŇ	: : : : : : : : : :	: : : :	 i N	S E	b i	к	 	a ì	, F	ΡQ	D F		v v	E	F K	G	κĒ		i c	 	CPUN 0011 20.420-011 44 40
Myge_Tp Mypn_Tp Mypi_Tp Bsub_Pynp Ecoli_Tp Human_Tp Metj_Tp Bst_Pynp	 H S S	 M A T T L	 A L M I N F	 МТР З G E	G T I G	 	 		 Р G V K	 D F E L 	S C G V		 	 D I	a L	PD TI	 Р S К H		 L E E K	MMMMFKPM1	NNNRLQKR				NSELRRRA		
Myge_Tp Mypn_Tp Mypi_Tp Bsub_Pynp Ecoll_Tp Human_Tp Metj_Tp Bst_Pynp	KKKKIGZK Rrzzodog R	AAKEARKA	L A T E Q T E E A T C E E A T C E E A T C E E A T C E E A T C E E A T C E E A T C E E A T C E A T C E A T C E A T C E A T C E A T E A T A T E A T A T A T A T A T A T A T A T A T A T		WFFFFFI		AVLY		KKKGZGGG	T L L S S S T S K D	ALPPSQTP	YYY GALY		TSSSAGSS	FFLLLMFL		AAAIASA	WWWFFRY	KONOHENE	00000000		PSDDMLME			FFYAVSEA		
Myge_Tp Mypn_Tp Mypi_Tp Bsub_Pynp Ecoli_Tp Human_Tp Metj_Tp Bst_Pynp	К А А М К К А А М М А А А М О А В О О О О О О О О О О О О О О О О О	V V I V R A A V	GGGGGGGGGG	ILK ILK ILDEN IL IL I I I I I I I I I I I I I I I I	FFFLSKP.S	P I S L S I	HHHY ELL WR. R.	GOK .NOH .	KKKGGQGG	SSLKI .IK			SSSSSSSS *	000000000000000000000000000000000000000		OCOROCOC	DZDDIDDD 9	>>>Z+>>+	SSASSALL	AAASMVLV		PPPPPPP		T T V A A A A A A A A A A A A A A A A A	SASAAASSO	G D L V G G C L V	
Myge_Tp Mypn_Tp Mypl_Tp Bsub_Pynp Ecoli_Tp Human_Tp Meti_Tp Bst_Pynp	88022 88022 8002 800 800 800 800 800 800	KKKKMMKK 1 1 1 1 1 1 1 1 1 1 1	GRCGGRCGGRCGGRCGGRCGGRCGGRCGGRCGGRCGGRC		YYFHHHSH	GGGGGGGAG	T T T T T T T T T T					GGNGGG .G	VVVFFF.F130	TTTVIVVV			Q T K D D P H D Q H X H Z H Z H D	AAS .RQL .	QQKIRQRI 14	ANKKEVVRO			NNADVANN	DDXKGGGG		I I I V V I V I V V I V V V V V V V V V	
Myge_Tp Mypi_Tp Bsub_Pynp Ecoli_Tp Human_Tp Meti_Tp Bst_Pynp	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	KENGSELG	A P P A A A A P P P A A A A P P P A A A A P P A A A A A P P A A A A T A V A T A V A T A V		VVLKRIIK	GAAAAAAA YYYYYYYY	LLRRRR.R *			GAGGAALA	>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>>	SSSSSSPS				000000000 *		SSSSASAS	* ×××××××× * ×××××××××		AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	MNEGGGGG	EDDDDSND	YYYAAL	FFFVVLV		
Myge_Tp Mypn_Tp Bsub_Pynp Ecoll_Tp Human_Tp Metj_Tp Bst_Pynp	L KYY I KYY V KYT V KKF I P K 200	GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	AAAAAAAA	KKHDTTNSK2	KKIEYQILO	IMTDLQEE	NKKAERSR		KONKEKRR	L M M A I T L K M 220			KKKNNADK	SQKNGSRR	KKKGGGRG		L SYYMTAEM	VI FAAACA		SSSSTTTS	D M M M M M M M M M M M M M M M M M M M	NNNSNDGS240			KKNESRRY	A V T I V A A V A V	The second secon

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 = M-thanococcus 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).

_Tp _Tp Fynp Tp n_Tp n_Tp Tp	G N V I E V N E A V N F L K Q D L D Q V G Q D F I D L MQT I VI NI L L E T K Q A K T K Q K A G N A L E V L E V V H F L K R N WT E V G A D F I QL ME Q I V T E I L I E T K R A P N K R A A G N A I E V K E A I D F L K N N. S D V G T Y F K K L MF D L V T L I L L K T K K C K T K K E A G N A L E V K E A I D T L K G E G P E . D. L H E L V L T L G S Q M V V L A K K A D T L D E A G N A V E V R E A V Q F L T G E Y R N . P R L F D V T MA L C V E ML I S G K L A K D D A E A G H A L E V E E A L L C M D G A G P P . D. L R D L V T T L G G A L W L S G H A G T Q A Q G G P A L E A K E A L L A L . E D Y T Q A P T S L V E K S I S L A G I L L E M G G V A P T . G E G G N A L E V K E A I E T L K G N G P H . D. L T E L C L T L G S H M V Y L A E K A P S L D E A 250 260 260 270 280 290 280 290
ΤΡ ΤΡ Γρ Ργηρ Τρ τ_Τρ τ_Τρ γηρ	V ALY HATLEGEK PWQRFLKFI ELGGS WERFLDL BGFFN VALY HATLEGEK PWQRFLKFI ELGGS WERFLDL DTFFK KEKI NY VLENKI AFNNFCNWI ELONGNI AKI KN DTFFK RAKLEEVMKNGKALEKFKDFLKNQGGDSSI VDDPS KLPG RAKLQAVLDNGKAAEVFGRMVAAQKGPTDFVENYA
_Tp _Tp Tp Pynp Tp n_Tp Tp ynp	P K Y K A S I K A E K S G I L H F T D P I D L A K I G I N L G A G R M K K T D Q I D P M A G L F P K Y K A P V L A S Q S G I L S Y T S P V D L A M V S I S L G A G R M V K T D L I D P M A G I K P K Y WT N I A A W K S G K I S Y K S I I E L A E I G V D L G S G R R K K E D K I D F Q A G I Y A A Y Q I D V P A K E A G V V S E I V A D E I G V A A M L L G A G R A T K E D E I D L A V G I N A M L T K A V Y A D T E G F V S E M D T R A L G M A V V M G G G R R Q A S D T I D Y S V G F T A R E Q E E L L A P A D G T V E L V R A L P L A L V L H E L G A G R S R A G E P L R L G V G A E G K Y K A D I H S P I D G Y V T R I S N A G I T K I A K E A G A P N D K K
Тр Тр Рупр Тр п_Тр гр упр	L MKKDNESVAVGDTVLNLYSSSPISNEYISAA.QKTIIINK. LVKQANEVVKAGDTVLELYSSKPITPAHIEAA.QHTIIINK. LHAKSNEKIKIKDKILTLYSSKPIKQDLIDKA.KKIIKIS LRKKVGDKVEKGEPLVTLYANRENVDEVIAK.VYDNIRIAAE.AKAP DMARLGDQVDGQRPLAVIHAKDENNWQEAAKAVKAAIKLADKAPESTP LLVDVGQRLRRGTPWLRVHRDGPALSGPQSRALQEALVLSDRAPFAAP LNVKVGNKVEKGDVLYTIYSDSEERLKSAIKLARILYPIKVEGMLLQK HKKIGDRVQKGEALATIHSNRPDVLDVKEK.IEAAIRLSPQPVARPP 380 400 410 420
_Tp _Tp _Pynp .Pynp .Tp n_Tp Tp ynp	K L I H T L I T E T V Y R R I S E L P F A E L V L P P Q I S R F L I Y E T I V

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 S Φ 312 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 17500g 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 17500g for 10 min at 277 K to pellet the precipitated protein. The pellet

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 17500g for 10 min at 277 K. The resulting pellet was resuspended in 10 ml (per 40 g of cell paste prepared) of 20 mM Tris pH 7.0 and dialyzed overnight against 21 of the same buffer. The dialyzed sample was centrifuged at 17500g 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 \times 50 cm) equilibrated with 20 mM Tris pH 7.0. The column was washed with 1 column volume of 20 mM Tris pH 7.0 followed by 2 column volumes of 20 mM Tris pH 7.0, 100 mM NaCl. A linear gradient of 1 column volume of 100-500 mM NaCl in 20 mM Tris pH 7.0 was passed over the

was discarded and the supernatant was

column. The PYNP eluted at approximately 200 mM NaCl. Active fractions were pooled and buffer-exchanged with a tenfold dilution (three times) into 20 mM Tris pH 7.0 in an Amicon pressure concentrator at 3.1 \times 10⁶ 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 \times 50 ml) and washed with 100 ml of 20 mM Tris pH 7.0. The PYNP was eluted from the column with 100 mM NaCl, 20 mM Tris pH 7.0. The final product was buffer-exchanged with a tenfold dilution (three times) into 10 mM 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.							

39.3 (2.79-2.71)
61540
21835
83.8 (52.3)
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 mM citrate buffer at pH 4.8 and 4 mM 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-substrateanalog complex structure, the protein sample was next prepared for co-crystallization experiments by adding to it 0.175 ml of 10 mM 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 mM 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 Xray data were collected on a San Diego Multiwire area-detector system using Cu Ka radiation from a Rigaku RU 200 rotatinganode 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 = 2nobserved, indicating space group $P2_1$. Assuming a dimer in the asymmetric unit, a Matthews number (Matthews, 1968) of 2.57 \AA^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 CCP4 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 PYNPcomplex 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. D50, 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.