Crystal structure of pyrimidine-nucleoside phosphorylase from *Bacillus subtilis* in complex with imidazole and sulfate

V. V. Balaev,^{a*} I. I. Prokofev,^a A. G. Gabdoulkhakov,^a C. Betzel^b and A. A. Lashkov^a

Received 3 October 2017

Accepted 19 February 2018

Edited by N. Sträter, University of Leipzig, Germany

Keywords: X-ray analysis; protein crystallography; nucleoside phosphorylases; nucleosides; imidazole; pyrimidine-nucleoside phosphorylase; *Bacillus subtilis*.

PDB reference: pyrimidine-nucleoside phosphorylase, complex with imidazole and sulfate, 5oln

Supporting information: this article has supporting information at journals.iucr.org/f

^aA. V. Shubnikov Institute of Crystallography, Leninsky Prospect 59, Moscow 119333, Russian Federation, and

^bLaboratory for Structural Biology of Infection and Inflammation, University of Hamburg, Institute of Biochemistry and Molecular Biology, c/o DESY, Building 22a, Notkestrasse 83, Hamburg, Germany. *Correspondence e-mail: vlad_balaev@mail.ru

Pyrimidine-nucleoside phosphorylase catalyzes the phosphorolytic cleavage of thymidine and uridine with equal activity. Investigation of this protein is essential for anticancer drug design. Here, the structure of this protein from *Bacillus subtilis* in complex with imidazole and sulfate is reported at 1.9 Å resolution, which is an improvement on the previously reported structure at 2.6 Å resolution. The localization and position of imidazole in the nucleoside-binding site reflects the possible binding of ligands that possess an imidazole ring.

1. Introduction

Pyrimidine-nucleoside phosphorylases catalyze the phosphorolysis of pyrimidine nucleosides to the corresponding nucleoside bases. These enzymes are widely used as biocatalysts for the biotechnological production of nucleosides (Utagawa, 1999). In eukaryotic cells and in some types of prokaryotes (for example, *Escherichia coli* and *Salmonella typhimurium*), pyrimidine-nucleoside phosphorylases are represented by narrow-specificity thymidine phosphorylases (TPs; EC 2.4.2.4) and uridine phosphorylases (EC 2.4.2.3). In some lower organisms (for example *Bacillus subtilis*, *Geobacillus stearothermophilus* and *Staphylococcus aureus*), a nonspecific pyrimidine-nucleoside phosphorylase (PyNP; EC 2.4.2.2) with the same catalytic activity cleaves both uridine and thymidine (Pugmire & Ealick, 2002). PyNPs are similar to thymidine phosphorylases, sharing 40% of their residues (Pugmire & Ealick, 1998). Both types of enzyme act as a homodimer with a two-domain subunit structure and with the active site located at the interface between the two domains.

The ability of some types of tumours to accumulate pyrimidine-nucleoside phosphorylases allows these enzymes to be used to activate antimetabolites based on pyrimidine nucleosides (Bronckaers *et al.*, 2009). For example, 5-(*E*)-(2-bromovinyl)-2'-deoxyuridine, 5-trifluorothymidine, 5-fluorouracil and 5-fluoro-5'-deoxyuridine (an intermediate metabolite of capecitabine) are used in clinical practice against breast and colon cancer metastases (Bronckaers *et al.*, 2009). However, some organisms that infect human cells (for instance, *Mycoplasma hyorhinitis*) hinder the use of nucleoside anticancer agents, which highlights another direction of research: the design of inhibitors of pyrimidine-nucleoside phosphorylase that are not able to bind to the active site of TP (Vande Voorde *et al.*, 2012).

Table 1
Macromolecule-production information.

Source organism	<i>B. subtilis</i> strain 168
EC No.	2.4.2.2
GenBank ID	BAA08339.1
Gene	<i>pdp</i>
Forward primer	5'-CGCCATATGAGAATGGTAGATATCATC ATC-3'
Reverse primer	5'-CGCCTCGAGTTATTCTGTAATTAACGT ATGAATC-3'
Cloning vector	pET-15b
Expression vector	pET-15b
Expression host	<i>E. coli</i>
Complete amino-acid sequence of the construct produced	AMRMVDIIKKQNGKELTTEEIQFFVNGYT DGSIPDYQASALAMAIFFQDMSDRERAD LTMAMVNSGETIDLSAIEGKVDKHSSTG GVGDTTTLVLAPLVAALDVPVAKMSGRG LGHTGGTIDKLEAIMGFHVELTKDEFIK LVNRDKVAVIGQSGNLTPADKKLYALRD VTGTVNSIPLIASSIMSKKIAAGADAIV LDVKTGAGAFMKTEEDAELAKAMVRIG NNVGRQTMAVISDMSQPLGFAIGNALEV KEAIDTLKGEGPEDLHELVLTLGSQMVV LAKKADTLDEARAKLEEVMMKNGKALEKF KDFLNQGGDSSIVDDPSKLPQAAYQID VPAKEAGVSEIVADEIGVAAMLLGAGR ATKEDEIDLAVGIMLRKKVGDKEKGEF LVTLYANRENVDEVIKVDNIRIAAEA KAPKLIHTLITE

In a previous paper, we compared the three-dimensional structure of *B. subtilis* PyNP (*BsPyNP*) complexed with sulfate with that of thymidine phosphorylase from *S. typhimurium* (*StTP*) (Balaev, Lashkov, Prokofev *et al.*, 2016). We reported differences in the domain movements of these enzymes and in their phosphate-binding sites. In this paper, we describe the three-dimensional structure of *BsPyNP* in complex with imidazole and sulfate. Several authors have reported compounds containing imidazole rings that were able to inhibit thymidine phosphorylase (Reigan *et al.*, 2005; Cole *et al.*, 2003; Rajabi *et al.*, 2011). The similarity of the nucleoside-binding sites in PyNP and TP implies that such drugs could also inhibit PyNP. X-ray analysis of *BsPyNP* complexed with imidazole will provide a good basis for further drug design. We

Figure 1
Crystals of the pyrimidine-nucleoside phosphorylase from *B. subtilis* complexed with imidazole and sulfate.

Table 2
Crystallization.

Method	Vapour diffusion, sitting drop
Plate type	96-well plate
Temperature (K)	288
Protein concentration (mg ml ⁻¹)	16
Buffer composition of protein solution	20 mM Tris-HCl, 500 mM KCl
Composition of reservoir solution	0.2 M ammonium acetate, 0.1 M bis-tris pH 5.5
Volume and ratio of drop	2 µl protein solution + 2 µl reservoir solution
Volume of reservoir (µl)	400

also describe some features of the new structure at 1.9 Å resolution compared with a previous structure at 2.6 Å resolution.

2. Materials and methods

2.1. Macromolecule production

The cloning and expression of *BsPyNP* has been extensively described in Balaev, Lashkov, Prokofev *et al.* (2016).

2.1.1. Biomass preparation and isolation and purification of the protein. The cell pellet was suspended in a binding buffer consisting of 20 mM Tris-HCl pH 7.5, 500 mM KCl, 30 mM imidazole. The cells were lysed by sonication for three cycles of 20 s. Unlysed cells were precipitated by centrifugation at 9000 rev min⁻¹ for 20 min at 4°C. Chromatographic purification was performed on an ÄKTA Start system (GE Healthcare). The resulting total protein was loaded onto a 2 ml Ni-NTA column (Qiagen) equilibrated with binding buffer and washed until the optical density decreased to the original value.

Elution of the protein was performed with a step gradient of imidazole at concentrations of 50, 150 and 300 mM in 20 mM Tris-HCl buffer pH 7.5, 500 mM KCl.

2.2. Crystallization

Screening of crystallization conditions was performed using The PACT Suite and The JCSG+ Suite crystallization kits (Qiagen, Germany). Macromolecule-production information is given in Table 1.

Crystals were grown by the sitting-drop vapour-diffusion method at 15°C. The crystallization and protein solution compositions are listed in Table 2. Crystals grew in a month (Fig. 1) and were used for X-ray data collection. The approximate dimensions of the crystals were 400 × 50 × 50 µm (Fig. 1).

2.3. Data collection and processing

Data sets for the *BsPyNP* complex were collected at 100 K on beamline 14.1 at BESSY II, Helmholtz-Zentrum, Berlin, Germany. 30% PEG was used as a cryoprotectant.

All data were indexed, merged and processed using the XDS package (Kabsch, 2010) and SCALA (Evans, 2011) within the CCP4 suite (Winn *et al.*, 2011). The data statistics are summarized in Table 3.

Table 3

Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	Beamline 14.1, BESSY II
Wavelength (Å)	0.92
Temperature (K)	100
Detector	PILATUS 6M
Crystal-to-detector distance (mm)	396
Rotation range per image (°)	0.1
Total rotation range (°)	180
Exposure time per image (s)	1
Space group	$P2_12_12_1$
a, b, c (Å)	64.37, 93.01, 139.85
α, β, γ (°)	90, 90, 90
Mosaicity (°)	0.17
Resolution range (Å)	77.44–1.88 (1.98–1.88)
Total No. of reflections	755376 (112933)
No. of unique reflections	71954 (11543)
Completeness (%)	98.4 (97.7)
Multiplicity	11.1 (11.6)
$\langle I/\sigma(I) \rangle$	8.6 (1.95)
$R_{\text{r.i.m.}}$	0.207 (1.103)
$R_{\text{p.i.m.}}$	0.062
Overall B factor from Wilson plot (Å ²)	26.8

Table 4

Structure solution and refinement.

Values in parentheses are for the outer shell.

Resolution range (Å)	46.50–1.88 (2.29–1.88)
Completeness (%)	98.4 (97.2)
σ Cutoff	$F > 1.970\sigma(F)$
No. of reflections, working set	67757 (27980)
No. of reflections, test set	3363 (1502)
Final R_{cryst}	0.200 (0.244)
Final R_{free}	0.239
Cruickshank DPI (Å)	0.14
No. of non-H atoms	
Protein	6473
Ion	5
Ligand	19
Water	450
Total	6951
R.m.s. deviations	
Bonds (Å)	0.007
Angles (°)	0.86
Average B factors (Å ²)	
Protein	31.1
Ion	31.1
Ligand	38.6
Water	29.8
Ramachandran plot	
Most favoured (%)	98.0
Allowed (%)	2.0

2.4. Structure solution and refinement

The structure of *BsPyNP* complexed with imidazole was solved by the molecular-replacement method using *Phaser* (McCoy, 2007) from the *CCP4* suite. The previous structure of *BsPyNP* (PDB entry 5ep8; A. A. Lashkov, V. V. Balaev, A. G. Gabdoulkhakov, C. Betzel & A. M. Mikhailov, unpublished work) at 2.66 Å resolution was used as a search model.

The *BsPyNP* structure was refined to an R_{work} of 20.0% and an R_{free} of 23.9% using *PHENIX* (Adams *et al.*, 2010) and *REFMAC5* (Murshudov *et al.*, 2011). Monitoring of structure quality and manual correction were performed using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010). Localization of

enzyme and ligand atoms was carried out using σ_A -weighted maps calculated with $mF_o - DF_c$ and $2mF_o - DF_c$ coefficients. In the final stage, refinement was performed using an anisotropic approximation for thermal factors for all atoms of the structure. The quality of the refinement results was estimated using *Coot* (Emsley & Cowtan, 2004; Emsley *et al.*, 2010) and *MolProbity* (Chen *et al.*, 2010). Refinement statistics for the *BsPyNP* model are shown in Table 4.

3. Results and discussion

Pyrimidine-nucleoside phosphorylase from *B. subtilis* is a member of the NP-II family (Pugmire & Ealick, 2002), implying that it has a two-domain structure and functions as a dimer. Its presence in some microorganisms infecting human cells hinders nucleoside-based anticancer therapy. Thus, the need arises to be able to inhibit this enzyme while not affecting human thymidine phosphorylase. We made an attempt to design such an inhibitor (Balaev, Lashkov, Prokofev *et al.*, 2016) based on a previous structure (PDB entry 5ep8) at only 2.66 Å resolution. We now present structural data for the same enzyme with imidazole bound (PDB entry 5oln). The most improved characteristics are the estimated coordinate errors: DPI (0.14 Å compared with 0.41 Å), maximal estimated error (0.09 Å compared with 0.36 Å) and B factor (31 Å² compared with 75 Å²). The quality of the structure with PDB code 5oln is also better in terms of R factors: an R_{work} of 19.9% and an R_{free} of 24.1% compared with an R_{work} of 21.5% and an R_{free} of 29.3% for PDB entry 5ep8. Some improvements are found in

Figure 2

The structure of the *BsPyNP* homodimer. The small α -domain is marked in orange and the α/β -domain is in blue. This figure was produced using *PyMOL* (v.1.3r1; Schrödinger).

Figure 3

The three-dimensional structures of the nucleoside-binding sites in subunits *A* (a) and *B* (b) of *BsPyNP* complexed with imidazole (IMD). An OMIT map is shown for the IMD molecule.

the Ramachandran plot statistics: 93.5% compared with 89.7% of residues are in preferred regions. The new structure is more precise and is more suitable for further drug design, in particular of imidazole-based drugs.

The monomer of the pyrimidine-nucleoside phosphorylase from the bacterium *B. subtilis* consists of 434 residues. It forms a homodimer, with the subunits interfacing *via* the α -domains (Fig. 2). The *BsPyNP* subunit, like that of TP (Walter *et al.*, 1990), consists of two domains. The first domain includes residues 1–65 and 159–191 and consists of α -helices (the α -domain), while the other domain includes both α -helices and β -strands (α/β -domain) (Fig. 3). The authors of the first structural investigations of TP assumed that movement of the domains towards each other is essential for the reaction to proceed (Walter *et al.*, 1990). PyNP might not act in this way. We have previously discussed this in terms of the substrate specificity of these enzymes (Balaev, Lashkov, Prokofev *et al.*, 2016). We made the assumption that nucleoside cleavage in PyNPs is performed without subunit closure and thus is possible both for thymidine and for uridine.

An imidazole (IMD) molecule was localized in the active site of both subunits (Fig. 3) of the *BsPyNP* homodimer (PDB entry 5oln). IMD interacts with the enzyme residues *via* hydrogen bonds and a stacking interaction. Superposition of the active sites of molecules *A* and *B* gives a distance of 0.7 Å between the IMD centres. Superposition of the nucleoside-binding sites of the complexes of *BsPyNP* with IMD and of *SrTP* with thymidine (PDB entry 4yek, chain *B*; Balaev, Lashkov, Gabdulkhakov *et al.*, 2016) gives distances of 1.5 and 1.4 Å between the centres of the thymidine pyrimidine (in *SrTP*) and IMD (*BsPyNP*) rings (a superposition of the subunit *B* nucleoside-binding sites is shown in Fig. 4). Both rings interact with Tyr165 (Tyr168 in *SrTP*), Arg168 (Arg171 in *SrTP*) and Ser183 (Ser186 in *SrTP*). This similarity indicates that imidazole-based inhibitors (Cole *et al.*, 2003; Reigan *et al.*, 2005; Rajabi *et al.*, 2011) might interact with TP in the

nucleoside-binding sites *via* the imidazole base. Peptide drugs containing imidazole rings are promising as TP inhibitors, being less toxic than those currently used.

The mean distance between the two domains was calculated for both *BsPyNP* subunits (PDB entry 5oln), each with imidazole bound in the active site, for both subunits of PyNP from *G. stearothermophilus* (PDB entry 1brw; Pugmire & Ealick, 1998), which are known to be in the closed subunit conformation, and for the half-closed subunit *B* of *SrTP* complexed with thymidine (PDB entry 4yek). The values are 31.4 and 31.7 Å for *BsPyNP*, 26.3 and 29.1 Å for *GsPyNP* and 30.1 Å for *SrTP*. These values indicate that imidazole does not result in the closure of *BsPyNP* subunits, even to the half-closed conformation as found in *SrTP* bound to thymidine.

Figure 4

Superposition of the nucleoside-binding sites in subunits *B* of *SrTP* (in green) complexed with thymidine (THM) and in *BsPyNP* (in blue) complexed with IMD (violet). The corresponding PDB entries are 4yek and 5oln, respectively.

Figure 5
The displacement of the loop region (residues 110–115) owing to sulfate binding in *BsPyNP*.

Sulfate was located in subunit *B* of *BsPyNP*. Comparison with the previous structure (PDB entry 5ep8) gives a distance of 1.3 Å between sulfate centres (in the *B* subunits). Binding of sulfate is likely to result in a shift of the loop region containing residues 110–115 towards it (Fig. 5). The magnitude of this shift varies from 1 to 2 Å. The effect of phosphate/sulfate on the stabilization of this loop in TP was originally hypothesized by Pugmire *et al.* (1998). However, further studies (Norman *et al.*, 2004; El Omari *et al.*, 2006; Mitsiki *et al.*, 2009; Balaev, Lashkov, Gabdulkhakov *et al.*, 2016) did not confirm this hypothesis.

In conclusion, the reported structure provides a description of a procedure for crystal growth that results in crystals that diffract to a resolution higher than 2 Å, a high-quality three-dimensional structure for investigation of the structure–function relationship of PyNPs *via* computational chemistry and a basis for the rational drug design of imidazole-based compounds.

Funding information

This study was funded by the Federal Agency of Scientific Organizations (Agreement No. 007-GZ/CH3363/26) for the protein structure investigation and by a grant from the President of the Russian Federation (No. MK-9246.2016.3).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
- Balaev, V. V., Lashkov, A. A., Gabdulkhakov, A. G., Dontsova, M. V., Seregina, T. A., Mironov, A. S., Betzel, C. & Mikhailov, A. M. (2016). *Acta Cryst.* **F72**, 224–233.
- Balaev, V. V., Lashkov, A. A., Prokofev, I. I., Gabdulkhakov, A. G., Seregina, T. A., Mironov, A. S., Betzel, C. & Mikhailov, A. M. (2016). *Crystallogr. Rep.* **61**, 830–841.
- Bronckaers, A., Gago, F., Balzarini, J. & Liekens, S. (2009). *Med. Res. Rev.* **29**, 903–953.
- Cole, C., Reigan, P., Gbaj, A., Edwards, P. N., Douglas, K. T., Stratford, I. J., Freeman, S. & Jaffar, M. (2003). *J. Med. Chem.* **46**, 207–209.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- El Omari, K., Bronckaers, A., Liekens, S., Pérez-Pérez, M. J., Balzarini, J. & Stammers, D. K. (2006). *Biochem. J.* **399**, 199–204.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486–501.
- Evans, P. R. (2011). *Acta Cryst.* **D67**, 282–292.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- McCoy, A. J. (2007). *Acta Cryst.* **D63**, 32–41.
- Mitsiki, E., Papageorgiou, A. C., Iyer, S., Thiagarajan, N., Prior, S. H., Sleep, D., Finnis, C. & Acharya, K. R. (2009). *Biochem. Biophys. Res. Commun.* **386**, 666–670.
- 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.
- Norman, R. A., Barry, S. T., Bate, M., Breed, J., Colls, J. G., Ernill, R. J., Luke, R. W., Minshull, C. A., McAlister, M. S., McCall, E. J., McMiken, H. H., Paterson, D. S., Timms, D., Tucker, J. A. & Pauptit, R. A. (2004). *Structure*, **12**, 75–84.
- 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.
- Pugmire, M. J. & Ealick, S. E. (2002). *Biochem. J.* **361**, 1–25.
- Rajabi, M., Mansell, D., Freeman, S. & Bryce, R. A. (2011). *Eur. J. Med. Chem.* **46**, 1165–1171.
- Reigan, P., Edwards, P. N., Gbaj, A., Cole, C., Barry, S. T., Page, K. M., Ashton, S. E., Luke, R. W., Douglas, K. T., Stratford, I. J., Jaffar, M., Bryce, R. A. & Freeman, S. (2005). *J. Med. Chem.* **48**, 392–402.
- Utagawa, T. (1999). *J. Mol. Catal. B Enzym.* **6**, 215–222.
- Vande Voorde, J., Gago, F., Vrancken, K., Liekens, S. & Balzarini, J. (2012). *Biochem. J.* **445**, 113–123.
- 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.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **67**, 235–242.