Expression, purification and crystallization of phosphoribosyl transferase from a mycobacteriophage

Lei Li,^a Eungyu Lee^b and Neil Shaw^{b,c,*}

Received 30 December 2017

Accepted 11 February 2018

Edited by K. K. Kim, Sungkyunkwan University School of Medicine, Republic of Korea

Keywords: mycobacteriophages; tuberculosis; X-ray diffraction; phosphoribosyl transferases; phage BTCU-1.

^aState Key Laboratory of Biotherapy and Cancer Center/National Collaborative Innovation Center for Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, People's Republic of China, ^bJOY EL International Academy, 8 Laiguangying East Road, Chaoyang District, Beijing 100 102, People's Republic of China, and ^cNational Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing 100101, People's Republic of China.

*Correspondence e-mail: neilshaw@moon.ibp.ac.cn

Tuberculosis (TB) continues to remain a leading cause of death globally. Of particular concern is the emergence and rise in incidence of multidrug-resistant and extremely drug-resistant cases of TB. To counter this threat, it is important to explore alternative therapies, including phage therapy. Phage BTCU-1 specifically infects *Mycobacterium* spp. and kills the majority of them. Intriguingly, many proteins from the phage do not share high amino-acid sequence identity with proteins from species other than phages. Here, the expression, purification and crystallization of one such protein, a putative phosphoribosyl transferase from phage BTCU-1, is reported. The crystals belonged to space group $C222_1$, with unit-cell parameters $a = 59.71$, $b = 64.42$, $c = 65.32$ Å, $\alpha = \beta = \gamma = 90^\circ$. The crystals diffracted X-rays to 2.2 Å resolution.

1. Introduction

In 2016, more than 1.5 million people died from tuberculosis (TB), a disease associated with infection by *Mycobacterium tuberculosis* (World Health Organization, 2017). The high number of deaths associated with this disease makes it one of the major causes of death globally. Of particular concern are the rise in cases where the bacterium has gained resistance to the first-line drugs for the treatment of TB: isoniazid and rifampicin. TB caused by such bacteria is referred to as multidrug-resistant TB (MDR-TB) and almost half a million cases of MDR-TB were reported in 2016. Although second-line drugs are available for the treatment of MDR-TB, these therapies take much longer and require the use of drugs that have toxic side effects. More recently, cases of bacterial resistance to the second-line drugs, referred to as extremely drug-resistant TB (XDR-TB), have been reported. Another area of alarm related to the epidemiology of TB is the increase in cases of co-infection by *M. tuberculosis* in patients suffering from acquired immunodeficiency syndrome (AIDS) caused by *Human immunodeficiency virus* (HIV). More than 1.4 million people infected with HIV are reported to have developed TB owing to co-infection with the tubercule bacillus. It is hard to treat such patients with currently available therapies.

An alternative strategy to overpower the tubercule bacillus would be to use phages that cause the death of the bacterium. There are more than 4200 phages that specifically infect *Mycobacterium* spp. (Hatfull, 2010). These phages, which are referred to as mycobacteriophages, can kill the bacteria during

© 2018 International Union of Crystallography

Table 1
Macromolecule-production information.

Source organism	Phage BTCU-1
DNA source	Synthesized by GenScript
Forward primer†	CATGCCATGGGCAGCAGCCATCATCATCAT CATCACATGGATTCCGTGAGACGAGTTCTG
Reverse primer‡	CGGGATCCTCAGTCTTCTCCTTCAGGTATCC
Cloning vector	pET-28a
Expression vector	pET-28a
Expression host	<i>E. coli</i> BL21 (DE3)
Complete amino-acid sequence of the construct produced§	MGSSHHHHHMDSVRRVLVTGSRDWKDRIT VWAALRQELEQFGSLVVHGAARGADDI ADRAWGMAQAGYQVHVEAHPADWDGLG KAAGVIRNQQMVDLGADICHAFFLPGSI GTWDCIGRAEAAGIRVVNHGYLKEKT

† The underlined sequence denotes the *Nco*I restriction site. ‡ The underlined sequence denotes the *Bam*HI restriction site. § The hexahistidine tag is underlined in the amino-acid sequence of the protein.

the course of their life cycle. Recently, one such phage named BTCU-1 that specifically infects *Mycobacterium* was isolated from a soil sample in Taiwan (Lai *et al.*, 2015). Images of the phage captured using TEM reveal that the phage resembles members of the *Siphoviridae* family. It has an icosahedral head with a very long tail. The icosahedral head encapsulates a 46 kb stretch of linear double-stranded DNA that makes up the genome of the phage.

The genome of BTCU-1 encodes a number of proteins that have been annotated with putative functions. A remarkable feature of these proteins is that many of them share no sequence similarities with any structures deposited in the Protein Data Bank (PDB; Berman *et al.*, 2000). One such example is a 14.05 kDa protein annotated as a putative phosphoribosyl transferase (PRT). The primary sequence of this protein retrieves absolutely no matches when the PDB is searched for homologues. We decided to conduct structural characterization of this protein to shed light on its architecture and gain clues about its function. Here, we report the results of the cloning, expression, purification, crystallization and preliminary X-ray diffraction studies of the PRT from phage BTCU-1.

2. Materials and methods

2.1. Macromolecule production

The gene for the PRT (gene *CKC61*; UniProt ID R9R4E0) was synthesized (GenScript, USA). The full-length gene was subcloned into a pET-28a vector (Novagen, USA). For this, the gene was first PCR-amplified using the primers listed in Table 1. The amplified PCR products were gel-purified using a QIAquick Gel Extraction Kit (Qiagen, People's Republic of China) and digested with *Nco*I and *Bam*HI. The restriction-enzyme-treated products were ligated into a pET-28a vector. The resulting plasmid, pET-28-PRT, was sequenced to verify the accuracy of the clone. The protein was produced in *Escherichia coli* by first transforming the pET-28-PRT plasmid into *E. coli* BL21 (DE3) cells. The cells were grown at 310 K in Luria–Bertani (LB) medium in the presence of 50 µg ml⁻¹

Table 2
Crystallization.

Method	Hanging-drop vapour diffusion
Plate type	Linbro 24-well plate
Temperature (K)	277
Protein concentration (mg ml ⁻¹)	20
Buffer composition of protein solution	20 mM Tris–HCl pH 8.0, 250 mM NaCl, 5% glycerol
Composition of reservoir solution	100 mM Tris–HCl buffer pH 8.5, 25% PEG 3350
Volume and ratio of drop	2 µl; 1:1
Volume of reservoir (µl)	100

kanamycin. When the cell density (OD_{660 nm}) reached 0.8, the cultures were first cooled at 277 K for 2 h and then induced with 0.2 mM IPTG at 289 K for 20 h. At the end of the incubation time, the cells were harvested by centrifugation at 4570g for 15 min (Beckman Coulter, USA). The harvested cells were resuspended in lysis buffer consisting of 100 mM Tris–HCl pH 8.0, 250 mM NaCl, 5% glycerol. The cells were lysed using a high-pressure cell disruptor. Unbroken cells and cell debris were removed by centrifugation at 20 000g for 30 min at 277 K. The clarified supernatant containing the recombinant PRT protein was purified using nickel-affinity chromatography with Ni–NTA agarose beads (GE Healthcare, USA). The protein eluted in 20 mM Tris–HCl pH 8.0, 250 mM NaCl, 5% glycerol, 500 mM imidazole was concentrated to 1 ml using 10 kDa cutoff centrifugal concentrators (Millipore, USA) and loaded onto a Superdex G75 gel-filtration column (GE Healthcare, USA) equilibrated with 20 mM Tris–HCl pH 8.0, 250 mM NaCl, 5% glycerol. The protein eluted as a single peak. Fractions corresponding to the protein were pooled, concentrated to approximately 20 mg ml⁻¹ and stored at 277 K.

2.2. Crystallization

The protein was centrifuged at 8000g for 10 min at 277 K prior to screening for crystallization. Crystallization screening was performed at 277 K using the hanging-drop vapour-diffusion method. A total of 300 conditions from commercially available sparse-matrix screens (Hampton Research, USA) were set up during the first round of screening. 1 µl protein solution was mixed with 1 µl mother liquor and the drop was equilibrated over 100 µl reservoir solution. Crystals appeared in several conditions after a week. Crystals grown in 0.1 M Tris–HCl buffer pH 8.5, 25% PEG 3350 (condition No. 45 of the Index screen) were of sufficient size for their ability to diffract X-rays to be tested. Details of the crystallization are shown in Table 2. Crystals were soaked in mother liquor consisting of 20 mM Tris–HCl buffer pH 8.5, 25% PEG 3350, 1 mM HgCl₂, 20% glycerol for 1 h prior to data collection.

2.3. Data collection and processing

Initial tests on the crystals revealed that although they diffracted X-rays well, they needed cryoprotection. Soaking crystals in mother liquor containing 20% glycerol improved the diffraction pattern. Cryoprotected crystals were flash-cooled in liquid nitrogen prior to data collection. Data were

collected on beamline BL18U1 of the Shanghai Synchrotron Radiation Facility (SSRF). The crystals were kept cool in a stream of nitrogen gas at 100 K during data collection. The distance between the crystal and the detector was set to

350 mm. The crystal was oscillated by 0.5° after every 0.1 s exposure. A total of 720 images were collected during a 360° scan of the crystal. The data were processed using *HKL-3000* (Table 3; Minor *et al.*, 2006).

Figure 1

(a) Sequence alignment of homologues of the PRT from phage BTCU-1. The top ten matches in the *BLASTP* search for homologues of the PRT from phage BTCU-1 include proteins from viruses and phages such as the Heldan strain (95% sequence identity) to the Anubis strain (70% sequence identity) that infect *Mycobacterium*. (b) Sequence alignment of homologues of the PRT from phage BTCU-1 from *Mycobacterium* spp. Sequence alignment of PRT with the top four matches of homologues from *Mycobacterium* spp. is shown. The top match from *M. abscessus* shares only 56% sequence identity with the PRT from phage BTCU-1. Only the name of the species of *Mycobacterium* is shown in the alignment. The phage BTCU-1 sequence is named 'phage' in the alignment. Sequences were aligned using *Clustal Omega* (Sievers *et al.*, 2011).

Table 3
Data collection and processing.

Values in parentheses are for the outer shell.

Diffraction source	BL18U1, SSRF
Wavelength (Å)	0.97776
Temperature (K)	100
Detector	PILATUS 6M
Crystal-to-detector distance (mm)	350
Rotation range per image (°)	0.5
Total rotation range (°)	360
Exposure time per image (s)	0.1
Space group	C222 ₁
a, b, c (Å)	59.71, 64.42, 65.32
α, β, γ (°)	90, 90, 90
Resolution range (Å)	50.00–2.20 (2.24–2.20)
Total No. of reflections	52235
No. of unique reflections	6741
Completeness (%)	99.7 (99.7)
Multiplicity	7.7 (5.9)
<i>I</i> / <i>σ</i> (<i>I</i>)	26.79 (3.75)
<i>R</i> _{int}	0.061 (0.246)
<i>R</i> _{merge}	0.160 (0.587)

3. Results and discussion

The PRT gene of phage BTCU-1 encodes a 130-amino-acid protein. A search for homologues using the *BLASTP* suite at NCBI (Altschul *et al.*, 1997) retrieved sequences with two obvious trends. Firstly, the closest matches (sequence identities of between 95 and 60%) included about 60 proteins of almost similar length (125–145 amino acids). Intriguingly, all of these proteins are either from phages (the majority) or viruses (a few) that infect *Mycobacterium* (Fig. 1*a*). None of them are from bacteria or any other species, suggesting that PRT and PRT-like proteins are specifically found in phages

and viruses that infect *Mycobacterium*. Secondly, proteins that exhibited less than 56% sequence identity include those that predominantly originate from *Mycobacterium* spp. (Fig. 1*b*). This suggests that either part of the gene was acquired by the mycobacteriophages from *Mycobacterium* or some part of the gene from mycobacteriophages mimics the function of the protein from *Mycobacterium*. Sequencing of the genomes of mycobacteriophages and comparison of their genomes indicates that horizontal gene transfer is common between the phages and their mycobacterial hosts (Pedulla *et al.*, 2003). It is quite likely that the PRT gene or parts of it were acquired by the mycobacteriophages from their host and the gene subsequently diverged further. However, given the fact that the PRT gene shares only about 55% (or less) sequence identity with proteins from *Mycobacterium*, this gene could very well be of mycobacteriophage origin and might be assisting the phage in adapting to its niche.

The PRT protein was cloned in a pET-28a vector such that the protein was produced with a hexahistidine tag at the N-terminus. The expression of the protein in *E. coli* strain BL21 (DE3) was high. Approximately 120 mg protein could be purified from a 2 l culture. After the size-exclusion chromatography step, the purity of the protein was >95% as judged by the Commassie-stained protein band on an SDS–PAGE gel (Fig. 2*a*). The retention time of the protein during size-exclusion chromatography suggests a molecular weight of around 14 kDa for the protein. Thus, the protein seems to exist as a monomer in solution.

The crystals were tested for their ability to diffract X-rays (Fig. 2*b*), and diffracted X-rays to 2.5 Å resolution when tested using an in-house X-ray source (Fig. 3). A higher resolution data set was collected at SSRF. The crystals

Figure 2
(*a*) Purification of PRT from phage BTCU-1. The elution profile of PRT during gel filtration on a Superdex G75 column is shown. The protein eluted as a single peak. The protein was concentrated and two different aliquots, 1 and 2 µl, were loaded onto an SDS–PAGE gel before setting up crystallization. Lane *M* contains molecular-mass marker, with the position of the 20 kDa standard marked. (*b*) Crystals of PRT from phage BTCU-1. Crystals grown in the condition 0.1 M Tris–HCl buffer pH 8.5, 25% PEG 3350 were used for data collection.

Figure 3
X-ray diffraction pattern of a crystal of PRT from page BTCU-1.

belonged to space group $C222_1$, with unit-cell parameters $a = 59.71$, $b = 64.42$, $c = 65.32$ Å, $\alpha = \beta = \gamma = 90^\circ$. The asymmetric unit contains one molecule of the protein with a solvent content of 44% and a Matthews coefficient of 2.23 Å³ Da⁻¹ (Matthews, 1968). Phasing and model building are currently under way.

Acknowledgements

We thank Dr Lijie Wu at beamline BL18U1 of Shanghai Synchrotron Radiation Facility for his generous help with collection of diffraction data.

Funding information

Funding for this research was provided by: The National Natural Science Foundation of China (grant Nos. 31000332, 31100208, 31100877, 31170782 and 31110103915).

References

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). *Nucleic Acids Res.* **25**, 3389–3402.
- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.
- Hatfull, G. F. (2010). *Annu. Rev. Microbiol.* **64**, 331–356.
- Lai, M.-J., Liu, C.-C., Jiang, S.-J., Soo, P.-C., Tu, M.-H., Lee, J.-J., Chen, Y.-H. & Chang, K.-C. (2015). *Molecules*, **20**, 19277–19290.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. (2006). *Acta Cryst. D* **62**, 859–866.
- Pedulla, M. L. *et al.* (2003). *Cell*, **113**, 171–182.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D. & Higgins, D. G. (2011). *Mol. Syst. Biol.* **7**, 539.
- World Health Organization (2017). *Global Tuberculosis Report 2017*. Geneva: World Health Organization. http://www.who.int/tb/publications/global_report/en/.