Ken-ichi Miyazono,^a Kanae
Tsutsumi,^a Yoshizumi Ishino^b
and Masaru Tanokura^{a*}

^aDepartment of Applied Biological Chemistry,
Graduate School of Agricultural and Life
Sciences, The University of Tokyo, 1-1-1 Yayoi,
Bunkyo-ku, Tokyo 113-8657, Japan, and

^bDepartment of Bioscience and Biotechnology,
Graduate School of Bioresource and
Bioenvironmental Sciences, and Faculty of
Agriculture, Kyushu University, Fukuoka, Japan

Correspondence e-mail:
amtanok@mail.ecc.u-tokyo.ac.jp

Received 13 February 2014

Accepted 1 June 2014

Expression, high-pressure refolding, purification, crystallization and preliminary X-ray analysis of a novel single-strand-specific 3'–5' exonuclease PhoExo I from *Pyrococcus horikoshii* OT3

PhoExo I is a single-strand-specific 3'–5' exonuclease from *Pyrococcus horikoshii* OT3 and is thought to be involved in a Thermococcales-specific DNA-repair pathway. The recombinant PhoExo I protein was produced as inclusion bodies in *Escherichia coli* cells. Solubilization of the inclusion bodies was performed by the high-pressure refolding method and highly purified protein was subjected to crystallization by the sitting-drop vapour-diffusion method at 20°C. A crystal of PhoExo I was obtained in a reservoir solution consisting of 0.1 M Tris–HCl pH 8.9, 27% PEG 6000 and diffracted X-rays to 1.52 Å resolution. The crystal of PhoExo I belonged to space group *H*32, with unit-cell parameters $a = b = 112.07$, $c = 202.28$ Å. The crystal contained two PhoExo I molecules in the asymmetric unit.

1. Introduction

Nucleases are enzymes that catalyze the hydrolysis of phosphodiester bonds in DNA and RNA. Nucleases are indispensable for living cells because these enzymes participate in all DNA transactions, including replication, repair and recombination (Yang, 2011). In addition, some nucleases (restriction endonucleases) function as a host-defence system against exogenous DNA. In a previous study, we identified a novel single-strand-specific 3'–5' exonuclease, PfuExo I, from the hyperthermophilic archaeon *Pyrococcus furiosus* (Tori *et al.*, 2013). Although more than 140 archaeal genomes have been completely sequenced, homologous enzymes of PfuExo I are found only in the Thermococcales. Because PfuExo I showed no sequence similarity to any other protein with a known function, it was impossible to predict the function of PfuExo I by amino-acid sequence comparison.

Since thermophiles are always threatened by high temperatures, which accelerate spontaneous mutation in DNA, and other DNA-damaging factors such as ionizing radiation and chemical agents (Watrin & Prieur, 1996; Jacobs & Grogan, 1998), it has been suggested that thermophiles must have extremely efficient and specialized DNA-repair systems to overcome such an inhospitable environment. The chromosome fragmentation of *P. furiosus* cells caused by exposure to ionizing radiation was quickly repaired by incubation of the cells at 95°C (DiRuggiero *et al.*, 1997). However, details of the components and mechanisms of Thermococcales-specific DNA-repair systems remain unclear. The fact that the amount of mRNA for PfuExo I increased after ionizing irradiation (Williams *et al.*, 2007) may indicate that PfuExo I is one of the indispensable enzymes for the DNA-repair system of the Thermococcales.

PfuExo I forms a homotrimer and excises a single-stranded DNA (ssDNA), but not a double-stranded DNA, at every two nucleotides from the 3' to the 5' direction. PfuExo I requires more than seven nucleotides to excise ssDNA and does not recognize 3'-phosphorylated ssDNA. In addition, PfuExo I prefers dT as a substrate (Tori *et al.*, 2013). PhoExo I (DDBJ accession No. AB935327) is a homologous enzyme from *P. horikoshii* OT3 and shares 76% amino-acid sequence identity with PfuExo I. PhoExo I consists of 229 amino-acid residues with a molecular weight of 25 634 Da. To elucidate the structural basis for the catalytic mechanism and characteristic features of this family of exonucleases, PhoExo I was overexpressed

as inclusion bodies in *Escherichia coli*, solubilized by the high-pressure refolding method, purified and crystallized. The PhoExo I crystal diffracted X-rays to 1.52 Å resolution.

2. Materials and methods

2.1. Overexpression, high-pressure refolding and purification

The gene fragment of PhoExo I (Tori *et al.*, 2013) was amplified by PCR from genomic DNA of *P. horikoshii* OT3 and was cloned into the *NdeI/BamHI* site of the pET-26b plasmid (Novagen). The expression construct was designed to overexpress PhoExo I in the cytoplasm of *E. coli* without any tag. PhoExo I was overexpressed in *E. coli* Rosetta(DE3) (Novagen) cells harbouring the constructed plasmid. The *E. coli* transformants were cultivated at 37°C in LB medium until the optical density at 600 nm reached 0.6. PhoExo I expression was induced by the addition of 0.1 mM (final concentration) isopropyl β -D-1-thiogalactopyranoside (IPTG). The cells were cultured for an additional 18 h at 18°C. After cultivation, the cells were harvested by centrifugation at 5000g for 10 min and disrupted by sonication in 50 mM Tris-HCl pH 8.0. After centrifugation at 40 000g for 30 min, the pellet was collected as an inclusion-body sample.

PhoExo I that was expressed as inclusion bodies was solubilized by the high-pressure refolding method as described previously (Okai *et al.*, 2012). Before the high-pressure refolding, the collected pellet was resuspended in 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM TCEP, 0.5 mM EDTA, 125 mM nondetergent sulfobetaine 201 (NDSB201;

Calbiochem), 5% glycerol and centrifuged at 8000g for 15 min to wash the inclusion bodies. After the centrifugation, the pellet was resuspended in 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM TCEP, 0.5 mM EDTA, 5% glycerol and centrifuged at 8000g for 15 min to remove NDSB201 from the sample. This step (removal of NDSB201) was repeated twice. The pellet was resuspended in the refolding buffer consisting of 50 mM Tris-HCl pH 8.0, 500 mM arginine to a concentration of 0.5 mg ml⁻¹ protein and was packed into a 39 ml Quick-Seal polypropylene tube (Beckman). Refolding of PhoExo I was performed in a high-pressure vessel (Barofold). The resuspended PhoExo I solution was incubated under a high-pressure condition (200 MPa) for 16 h at room temperature and was depressurized by 25 MPa every 5 min until the pressure of the solution reached 0.1 MPa. The solubilized solution was centrifuged at 40 000g for 30 min and the supernatant was dialyzed against 20 mM Tris-HCl pH 8.0. The solution was purified using an anion-exchange chromatography column (Resource Q, 6 ml; GE Healthcare) pre-equilibrated with 20 mM Tris-HCl pH 8.0 and was then eluted with a linear gradient of 0–1 M NaCl. PhoExo I was eluted at an NaCl concentration of 0.22 M. The eluted PhoExo I was dialyzed against 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 200 mM MgCl₂ and concentrated to 6 mg ml⁻¹ for crystallization. Because the activity of PhoExo I is dependent on a divalent cation (Tori *et al.*, 2013), a high concentration of MgCl₂ was added to the protein solution to obtain the PhoExo I–Mg²⁺ complex crystal. For exonuclease activity assay, the eluted PhoExo I was dialyzed against 10 mM Tris-HCl pH 8.0, 5 mM EDTA.

2.2. Exonuclease assay

Exonuclease activity of the purified PhoExo I was detected using the 5'-fluorescein-labelled 25 nt polythymine (5'-fluorescein-TTTTTTTTTTTTTTTTTTTTTT-3') as a substrate. 0.1 μ M of the fluorescein-labelled DNA and 0.15 μ M of the refolded PhoExo I were mixed in 20 μ l 20 mM Tris-HCl pH 8.0, 50 mM NaCl with or without 5 mM MgCl₂ and incubated at 65°C for 20 min. The solutions were separated on a denaturing 18% polyacrylamide gel in 0.5×TBE and 7 M urea, and fluorescence was measured using an LAS4000 Mini system (Fujifilm, Tokyo, Japan).

2.3. Oligomeric state analysis by gel-filtration chromatography

The PhoExo I purified using an anion-exchange chromatography column was loaded onto a Superdex 200 HR 10/30 (GE Healthcare) column and eluted with buffer consisting of 10 mM Tris-HCl pH 8.0, 200 mM NaCl. To estimate the multimerization state of PhoExo I, the following standard proteins were used: thyroglobulin (M_r = 669 000), ferritin (M_r = 440 000), conalbumin (M_r = 75 000), ovalbumin (M_r = 44 000), chymotrypsinogen A (M_r = 25 000) and ribonuclease A (M_r = 13 700). The experiments were performed at 20°C.

2.4. Crystallization and data collection

All crystallization experiments were performed at 20°C using the sitting-drop vapour-diffusion method. Initial crystallization screens were performed using the commercial crystal screening kits Crystal Screen HT and Index HT from Hampton Research as well as The JCSG Core I suite and The JCSG Core II suite from Qiagen. Each crystallization drop was prepared by mixing 0.75 μ l protein solution and 0.75 μ l reservoir solution. The reservoir-solution conditions that yielded protein crystals were optimized to produce crystals suitable for X-ray diffraction analysis. In the optimization step, the crystallization drops were prepared by mixing 2.0 μ l protein solution and 1.0 μ l reservoir solution to obtain larger crystals.

Figure 1
Purification of PhoExo I. (a) SDS-PAGE of PhoExo I. Lane M, molecular-weight marker (labelled in kDa). Lane 1, resuspended inclusion bodies. Lane 2, solubilized fraction from lane 1 by the high-pressure refolding method. Lane 3, purified and concentrated PhoExo I by anion-exchange chromatography. (b) Mg²⁺-dependent exonuclease activity of PhoExo I. (c) Oligomeric state analysis of PhoExo I by gel-filtration chromatography. The peak positions of the marker proteins in kDa are indicated by the black triangles at the top of the chromatogram.

Table 1

Summary of data-collection statistics.

Values in parentheses are for the highest resolution shell.

Beamline	AR-NE3A, Photon Factory
Wavelength (Å)	1.0000
Space group	<i>H</i> 32
Unit-cell parameters	
$a = b$ (Å)	112.07
c (Å)	202.28
Resolution range (Å)	20–1.52 (1.60–1.52)
Multiplicity	6.6 (3.1)
Completeness (%)	98.6 (91.1)
R_{merge} (%)	3.3 (53.6)
$\langle I/\sigma(I) \rangle$	30.2 (2.1)

The crystals were mounted on cryoloops and flash-cooled at -178°C in a nitrogen stream for data collection. For cryoprotection, the crystals of PhoExo I were soaked in reservoir solution supplemented with 20% (v/v) ethylene glycol for a few seconds. The X-ray

Figure 2

A crystal of PhoExo I. The scale bar is 200 μm in length.

Figure 3

X-ray diffraction image of PhoExo I. The circle indicates a resolution of 1.52 Å.

diffraction data set was collected on the AR-NE3A beamline at the Photon Factory, Tsukuba, Japan. The diffraction data were indexed and integrated with *XDS* (Kabsch, 2010) and scaled with *SCALA* in the *CCP4* suite (Winn *et al.*, 2011).

3. Results and discussion

The recombinant PhoExo I was produced with high efficiency in the *E. coli* cells. However, most of the overexpressed PhoExo I formed inclusion bodies. Therefore, we tried to solubilize it by the high-pressure refolding method. After high-pressure refolding, approximately 20% of the protein was solubilized (Fig. 1a). Since the inclusion bodies contained a large amount of PhoExo I, the solubilized PhoExo I was purified to homogeneity using only a single column-chromatography step (Fig. 1a). The apparent purity of PhoExo I estimated by SDS-PAGE analysis was more than 95%. We obtained 3 mg of PhoExo I per litre of cell culture. The exonuclease assay showed that the refolded PhoExo I exhibits Mg^{2+} -dependent 3′–5′ exonuclease activity similar to PfuExo I (Fig. 1b; Tori *et al.*, 2013). The molecular weight of the refolded PhoExo I was estimated to be 75 000 Da by gel-filtration analysis (Fig. 1c). These results indicate that the refolded PhoExo I forms an active homotrimer as observed for PfuExo I prepared from the normal soluble protein fraction (Tori *et al.*, 2013). Recently, the refolding technique using high hydrostatic pressure has been reported by several groups (Chura-Chambi *et al.*, 2008; Arana *et al.*, 2010; Fradkin *et al.*, 2010; Fraga *et al.*, 2010; Okai *et al.*, 2012). A high hydrostatic pressure of 100–300 MPa can physically dissociate aggregates and oligomers without the addition of denaturing agent because the high hydrostatic pressure conditions disrupt both the ionic and hydrophobic interactions without disruption of hydrogen bonds (Silva *et al.*, 2006; Qoronfleh *et al.*, 2007; Seefeldt *et al.*, 2009). Therefore, high hydrostatic pressure conditions may be useful to solubilize a protein from inclusion bodies under relatively mild and simple conditions. The high-pressure refolding method is a good technique for protein

preparation for crystallography because inclusion bodies contain a large amount of heterologously expressed protein with high purity.

The crystals of PhoExo I were obtained under a reservoir-solution condition consisting of 0.1 M Tris-HCl pH 8.9, 27% PEG 6000. The best crystal grew to dimensions of $0.15 \times 0.10 \times 0.10$ mm (Fig. 2) and diffracted X-rays to 1.52 Å resolution (Fig. 3). The preliminary crystallographic analysis indicated that the crystal belonged to space group *H32*, with unit-cell parameters $a = b = 112.07$, $c = 202.28$ Å. Statistics of the data collection are summarized in Table 1. According to the Matthews coefficient calculation (Matthews, 1968), the crystal contained two PhoExo I molecules in the asymmetric unit with a solvent content of 48.4%. Since PfuExo I shows no sequence similarity to proteins of known structure, it is difficult to determine the structure of PhoExo I by the molecular-replacement method. Structure determination by either single-wavelength or multiple-wavelength anomalous dispersion methods using selenomethionine-derivative PhoExo I crystals is currently under way.

The synchrotron-radiation experiments were performed on beamline AR-NE3A at the Photon Factory (Proposal Nos. 2008G137 and 2008S2-001). This work was supported by the National Project on Protein Structural and Functional Analyses, the Targeted Proteins Research Program (TPRP) and the Platform for Drug Discovery, Informatics and Structural Life Science from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

- Arana, M. E., Powell, G. K., Edwards, L. L., Kunkel, T. A. & Petrovich, R. M. (2010). *Protein Expr. Purif.* **70**, 163–171.
- Chura-Chambi, R. M., Genova, L. A., Affonso, R. & Morganti, L. (2008). *Anal. Biochem.* **379**, 32–39.
- DiRuggiero, J., Santangelo, N., Nackerdien, Z., Ravel, J. & Robb, F. T. (1997). *J. Bacteriol.* **179**, 4643–4645.
- Fradkin, A. H., Boand, C. S., Eisenberg, S. P., Rosendahl, M. S. & Randolph, T. W. (2010). *Biotechnol. Prog.* **26**, 743–749.
- Fraga, T. R., Chura-Chambi, R. M., Gonçalves, A. P., Morais, Z. M., Vasconcellos, S. A., Morganti, L. & Martins, E. A. (2010). *J. Biotechnol.* **148**, 156–162.
- Jacobs, K. L. & Grogan, D. W. (1998). *Arch. Microbiol.* **169**, 81–83.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
- Okai, M., Ohtsuka, J., Asano, A., Guo, L., Miyakawa, T., Miyazono, K., Nakamura, A., Okada, A., Zheng, H., Kimura, K., Nagata, K. & Tanokura, M. (2012). *Protein Expr. Purif.* **84**, 214–218.
- Qoronfleh, M. W., Hesterberg, L. K. & Seefeldt, M. B. (2007). *Protein Expr. Purif.* **55**, 209–224.
- Seefeldt, M. B., Rosendahl, M. S., Cleland, J. L. & Hesterberg, L. K. (2009). *Curr. Pharm. Biotechnol.* **10**, 447–455.
- Silva, J. L., Cordeiro, Y. & Foguel, D. (2006). *Biochim. Biophys. Acta*, **1764**, 443–451.
- Tori, K., Ishino, S., Kiyonari, S., Tahara, S. & Ishino, Y. (2013). *PLoS One*, **8**, e58497.
- Watrin, L. & Prieur, D. (1996). *Curr. Microbiol.* **33**, 377–382.
- Williams, E., Lowe, T. M., Savas, J. & DiRuggiero, J. (2007). *Extremophiles*, **11**, 19–29.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yang, W. (2011). *Q. Rev. Biophys.* **44**, 1–93.