Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray analysis of PH1010 from *Pyrococcus horikoshii* OT3, a member of the archaeal DUF54 family of proteins

Michio Shirokane, Yoriko Sawano, Ken-ichi Miyazono, Koji Nagata and Masaru Tanokura*

Department 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

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

Received 26 February 2007
Accepted 18 May 2007

PH1010 from *Pyrococcus horikoshii* OT3, a member of the archaeal DUF54 family of proteins, was expressed, purified and crystallized. Crystallization was performed by the sitting-drop vapour-diffusion method using PEG 3350 as the precipitant. The crystal diffracted X-rays to 1.90 Å resolution using a synchrotron-radiation source. The space group of the crystal was determined to be $P2_12_12_1$, with unit-cell parameters $a = 46.9$, $b = 49.5$, $c = 132.7$ Å. The crystal contained two PH1010 molecules in the asymmetric unit (V_M = 2.4 Å^3 Da^{-1}) and had a solvent content of 48%.

1. Introduction

Pyrococcus horikoshii OT3 is an obligate anaerobic hyperthermophilic archaeon that has an optimal growth temperature of 371 K in the presence of sulfur (Kawarabayasi *et al.*, 1998). Proteins from hyperthermophiles such as *P. horikoshii* are frequently selected as structural proteomics targets as they are highly stable and can easily be purified and crystallized. Their high thermostability is suitable for applications in the chemical, food and medicinal industries. According to the DOGAN database (http://www.bio.nite.go.jp/dogan/Top), PH1010 (16 056 Da, 138 residues) is a hypothetical protein of unknown function. PH1010 shares over 80% sequence identity with archaeal DUF54-family proteins from *P. abyssi* GE5 (Cohen *et al.*, 2003), *P. furiosus* DSM 3638 (Maeder *et al.*, 1999) and *Thermococcus kodakarensis* KOD1 (Fukui *et al.*, 2005). DUF54-family proteins are found in archaea but not in bacteria or eukaryotes. To date, no tertiary structure or biological function has been reported for the DUF54-family proteins. Thus, we aimed to determine the crystal structure of PH1010 as an approach to revealing the function of DUF54-family proteins. Here, we describe the expression, purification, crystallization and preliminary X-ray analysis of PH1010.

2. Methods and results

2.1. Expression and purification

The PH1010 gene was amplified by the polymerase chain reaction (PCR) with KOD-plus DNA polymerase (Toyobo) from the genomic DNA of *P. horikoshii* OT3. The amplified gene fragment was digested with the restriction enzymes *Nde*I and *Bam*HI. The digested product was ligated into the pET-26b(+) vector (Novagen) digested with the same restriction enzymes. The plasmid possessing the PH1010 gene was transformed into *Escherichia coli* Rosetta(DE3) for protein expression. The transformants were cultivated at 298 K in LB medium. Overexpression was induced by adding 1.0 m*M* isopropyl β-D-thiogalactopyranoside (IPTG) when the optical density at 600 nm reached 0.8 and the culture was continued overnight. The cells were suspended in 20 m*M* Tris–HCl pH 8.0 and 300 m*M* NaCl and were disrupted by sonication. The supernatant was incubated at 363 K for 30 min and then gently stirred in the presence of 0.1%(*w*/*v*) polyethyleneimine at 277 K for 10 min. Denatured proteins and precipitated nucleic acids were removed by centrifugation. After ammonium sulfate precipitation with 70% saturated (final concen-

Table 1
Data-collection and phasing statistics.

Values in parentheses are for the highest resolution shell.

	Native	SeMet
Beamline	Photon Factory BL-5A	SPring-8 BL41XU
Space group	$P2_12_12_1$	$P2_12_12_1$
Unit-cell parameters (Å)	a = 46.9, b = 49.5, c = 132.7	a = 45.5, b = 47.9, c = 130.1
Wavelength (Å)	1.00000	0.97908
Resolution (Å)	30.0–1.90 (1.97–1.90)	20.0–2.80 (2.90–2.80)
No. of observations	172095	98578
No. of unique reflections	25193	7495
Data completeness (%)	99.7 (99.8)	99.7 (97.9)
Redundancy	6.9 (6.6)	13.3 (11.9)
R_{sym}†	0.061 (0.290)	0.076 (0.179)
$\langle I\rangle/\langle\sigma(I)\rangle$	16.5 (6.5)	11.1 (8.5)
R_{cullis}		0.471
Phasing power		3.06
FOM		0.510

† $R_{sym} = \sum_{hkl}[(\sum_i |I_i - \langle I\rangle|)/\sum_i |I_i|]$, where I_i is the ith intensity measurement of reflection hkl, including symmetry-related reflections, and $\langle I\rangle$ is its average.

tration) ammonium sulfate, the supernatant was loaded onto a Resource PHE column (GE Healthcare) equilibrated with 50 m*M* sodium phosphate pH 7.0 containing 1.5 *M* ammonium sulfate and eluted with a linear 1.5–0 *M* ammonium sulfate gradient. PH1010 was further purified by gel filtration on a HiLoad 26/60 Superdex 75 prep-grade column (GE Healthcare) with 20 m*M* Tris–HCl pH 8.0 and 300 m*M* NaCl. Dimerization of PH1010 was observed on gel filtration using a Superdex75 HR 10/30 column (GE Healthcare; the elution volume of PH1010 is approximately 10.6 ml). The purified protein was dialyzed against 10 m*M* Tris–HCl pH 8.0 containing 200 m*M* NaCl and concentrated to 15 mg ml^{-1} by ultrafiltration using Centriprep and Microcon (3000 Da molecular-weight cutoff; Millipore).

For expression of the selenomethionine (SeMet) derivative of PH1010, the cells were grown in LB medium with 20 µg ml^{-1} kanamycin and harvested by centrifugation at an OD_{600} of 0.3. The cells were gently resuspended in M9 medium containing 0.1 mg l^{-1} thiamine, 0.4% glucose, 1 m*M* $MgSO_4$, 4.2 mg l^{-1} $FeSO_4$ and 50 mg l^{-1} of the amino acids Ile, Leu, Val and L-SeMet and 100 mg l^{-1} Lys, Phe and Thr. Overexpression was induced by adding 1.0 m*M* IPTG at an OD_{600} of 0.8 at 298 K. The purification method for the SeMet-derivative protein was identical to that for the native protein, except for the addition of 10 m*M* DTT to the purification buffers and of 2 m*M* DTT to the crystallization medium to prevent oxidation of the SeMet derivative. The purified SeMet-derivative protein was concentrated to 20 mg ml^{-1}. The purified native and SeMet-derivative proteins were assessed to be over 95% pure by SDS–PAGE (Fig. 1).

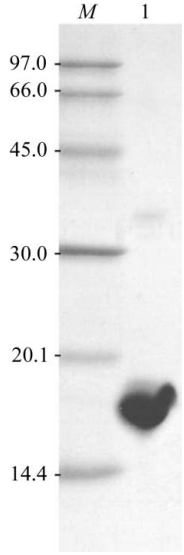
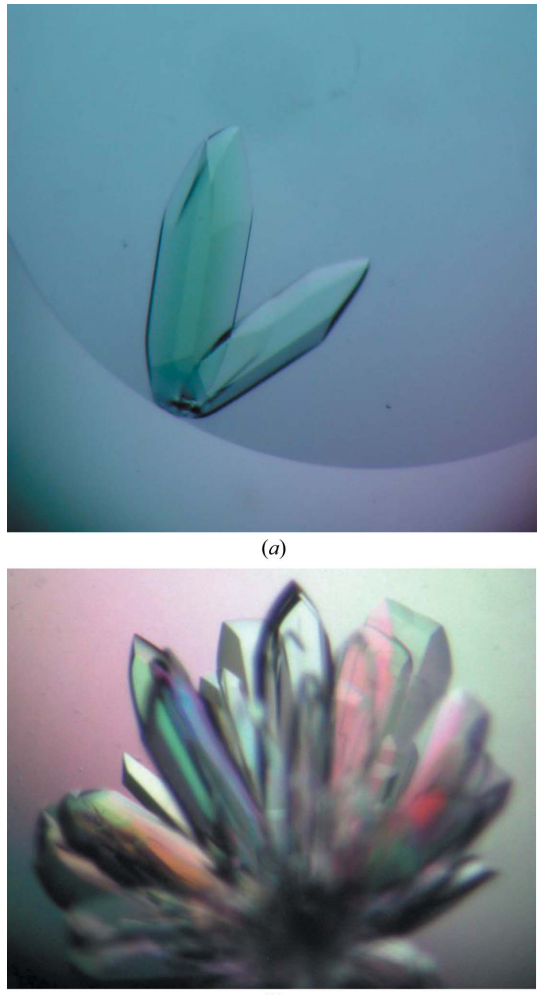


Figure 1
SDS–PAGE of the purified PH1010. Lane *M*, molecular-weight markers (kDa). Lane 1, PH1010.

2.2. Crystallization

All crystallization experiments were performed at 293 K using the sitting-drop vapour-diffusion method. Initial screening was performed using the Index HT crystallization screening kit (Hampton Research). After optimization of the crystallization conditions, the best crystals of the native and SeMet-derivative proteins were obtained in 18 and 25%(*w*/*v*) PEG 3350, respectively, in 0.2 *M* triammonium citrate pH 7.0. The drops were formed by mixing 1.5 µl protein solution with 1.5 µl reservoir solution and were equilibrated against 500 µl reservoir solution. Figs. 2(*a*) and 2(*b*) show images of typical crystals of native protein and SeMet-derivative protein,



(*a*)

(*b*)

Figure 2
Images of typical PH1010 crystals. (*a*) Crystals of the native protein of approximate dimensions 0.25 × 0.1 × 0.05 mm. (*b*) Crystals of the SeMet derivative of approximate dimensions 0.25 × 0.1 × 0.05 mm.

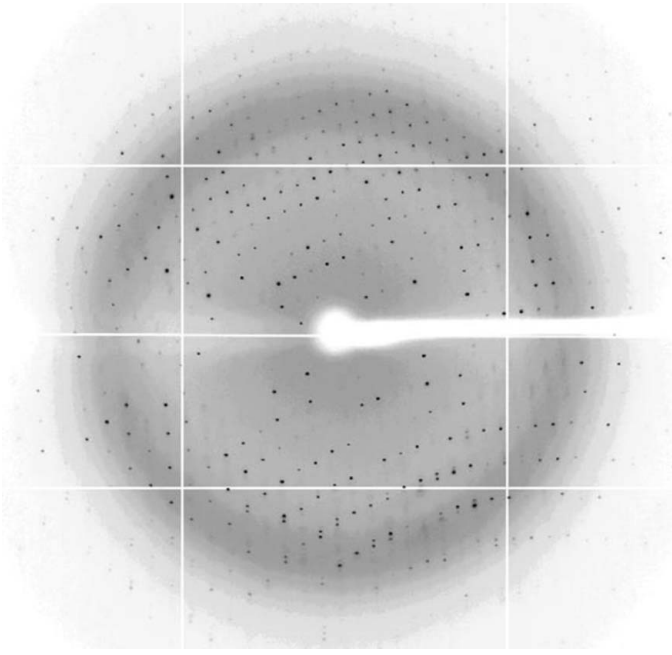


Figure 3
An X-ray diffraction image from a typical crystal of native protein. Diffraction data are detectable to 1.90 Å. The edge of the detector corresponds to a resolution of 1.90 Å.

respectively. The approximate dimensions of both the native and the SeMet-derivative crystals are 0.25 × 0.1 × 0.05 mm.

2.3. X-ray data collection and processing

Crystals of PH1010 were soaked in reservoir solution containing 20%(*v*/*v*) glycerol as a cryoprotectant and mounted for flash-cooling at 100 K. X-ray diffraction data were collected at 100 K on BL-5A at Photon Factory, Tsukuba, Japan and on BL41XU at SPring-8, Harima, Japan. The wavelength used to collect the anomalous data from the SeMet-derivative crystal was determined by fluorescent scanning. Fig. 3 shows an X-ray diffraction image from a crystal of the native protein. The data were indexed, integrated and scaled using the program *HKL*-2000 (Otwinowski & Minor, 1997). The diffraction data from the native protein were used to 1.90 Å resolution in order to avoid a decreased $R_{\rm sym}$. The diffraction data from the SeMet-derivative protein were used to 2.80 Å resolution to determine its anomalous data accurately. The native and SeMet-derivative crystals belonged to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 46.9, b = 49.5, c = 132.7 Å and a = 45.5, b = 47.9, c = 130.1 Å, respectively. The unit-cell parameters differ significantly between the native and SeMet-derivative proteins owing to the differing crystallization conditions of the two crystals. The $R_{\rm sym}$ of the native data set was 6.1%. The asymmetric unit contains two protein molecules, corresponding to a solvent content of 48% (the MW of PH1010 is 16 056 Da; $V_{\rm M}$ = 2.4 Å^3 Da^{-1}; Matthews, 1968). Details of the data-collection statistics are given in Table 1. The structure of PH1010 has been solved by the single-wavelength anomalous dispersion (SAD) method using the anomalous signal from the Se atoms of the SeMet-derivative protein. The Se sites were determined using the program *SHELXD* (Schneider & Sheldrick, 2002). Refinement of those sites and phasing were performed using the program *SHARP* (Bricogne *et al.*, 2003). Phasing statistics are summarized in Table 1. Model building is currently in progress. The detailed structure will be reported elsewhere.

We are grateful to Mr Tomohiro Ohishi for his kind help. The synchrotron-radiation experiments were performed on BL-5A at the Photon Factory, Tsukuba, Japan (Proposal No. 2003S2-002) and on BL41XU at SPring-8, Harima, Japan with the approval of the Japan Synchrotron Radiation Research Institute (Proposal Nos. 03S2-002 and 2005A0820-NL1-np, respectively). This research was supported by the National Project on Protein Structural and Functional Analyses of the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan and by Grants-in-Aid for Scientific Research from the MEXT of Japan.

References

Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. (2003). *Acta Cryst.* D**59**, 2023–2030.
Cohen, G. N., Barbe, V., Flament, D., Galperin, M., Heilig, R., Lecompte, O., Poch, O., Prieur, D., Querellou, J., Ripp, R., Thierry, J. C., Van der Oost, J., Weissenbach, J., Zivanovic, Y. & Forterre, P. (2003). *Mol. Microbiol.* **47**, 1495–1512.
Fukui, T., Atomi, H., Kanai, T., Matsumi, R., Fujiwara, S. & Imanaka, T. (2005). *Genome Res.* **15**, 352–363.
Kawarabayasi, Y. *et al.* (1998). *DNA Res.* **5**, 55–76.
Maeder, D. L., Weiss, R. B., Dunn, D. M., Cherry, J. L., Gonzalez, J. M., DiRuggiero, J. & Robb, F. T. (1999). *Genetics*, **152**, 1299–1305.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* D**58**, 1772–1779.