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

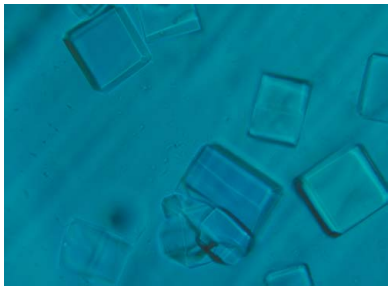
Overproduction, purification, crystallization and preliminary X-ray analysis of the peroxiredoxin domain of a larger natural hybrid protein from *Thermotoga maritima*

Carole Barbey,[a]* Nicolas Rouhier,[b] Ahmed Haouz,[c] Alda Navaza[a] and Jean-Pierre Jacquot[b]

[a]Laboratoire de Biophysique Moléculaire, Cellulaire et Tissulaire, UMR 7033, Université Paris 13, UFR SMBH, 74 Rue Marcel Cachin, 93017 Bobigny CEDEX, France, [b]Unité Mixte de Recherches 1136 INRA UHP (Interaction Arbres Microorganismes), IFR 110, Nancy Université BP 239, 54506 Vandoeuvre-lès-Nancy CEDEX, France, and [c]Plate-forme de Cristallogenèse et Diffraction des Rayons X, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris, France

Correspondence e-mail: carole.barbey@smbh.univ-paris13.fr

Received 21 September 2007
Accepted 29 November 2007



Thermotoga maritima contains a natural hybrid protein constituted of two moieties: a peroxiredoxin domain at the N-terminus and a nitroreductase domain at the C-terminus. The peroxiredoxin (Prx) domain has been overproduced and purified from *Escherichia coli* cells. The recombinant Prx domain, which is homologous to bacterial Prx BCP and plant Prx Q, folds properly into a stable protein that possesses biological activity. The recombinant protein was crystallized and synchrotron data were collected to 2.9 Å resolution. The crystals belonged to the tetragonal space group $I422$, with unit-cell parameters $a = b = 176.67$, $c = 141.20$ Å.

1. Introduction

Peroxiredoxins (Prxs) are non-haem-containing peroxidases that generally catalyze the reduction of peroxides to the corresponding alcohols through a catalytic cysteine that becomes oxidized to a sulfenic acid during the catalytic process (Chae *et al.*, 1994; Hofmann *et al.*, 2002; Wood *et al.*, 2003). Depending on their subunit composition and the mode of regeneration of the catalytic cysteine, peroxiredoxins have been classified into three types. The first group comprises the so-called typical 2-Cys peroxiredoxins, which are dimeric enzymes with a disulfide bridging two identical subunits. The second group, called 1-Cys peroxiredoxins, includes Prxs in which only one catalytic cysteine is conserved. The regenerating system used by the 1-Cys Prxs remains unknown, although ascorbate and glutathione have been shown to be efficient in some cases (Kang *et al.*, 1998; Monteiro *et al.*, 2007). The last group, called atypical 2-Cys peroxiredoxins, mostly contains Prxs that form an intramolecular disulfide bridge between two conserved cysteines. Nevertheless, in the case of plant type II Prxs, which were initially thought to belong to this group, the catalytic cysteine is conserved and essential but mutation of the noncatalytic cysteine does not totally abolish the activity (Rouhier *et al.*, 2002). This led to the introduction of another group of Prxs in plants, the Prx type II subgroup, which are Prxs that use only one cysteine (Rouhier & Jacquot, 2002). Usually, the thioredoxin system is needed for the reduction of the disulfide formed during catalysis and the regeneration of a reduced 'active' enzyme, but poplar type II Prx can use either thioredoxin or glutaredoxin as reductants (Rouhier *et al.*, 2001). The atypical 2-Cys Prx subgroup also contains the so-called glutathione peroxidases from some organisms such as plants, yeast and probably bacteria, which are in fact thioredoxin-dependent peroxidases and constitute another subgroup of the peroxiredoxin or thiol-peroxidase family (Rouhier & Jacquot, 2005; Herbette *et al.*, 2002; Navrot *et al.*, 2006; Tanaka *et al.*, 2005).

Finally, the atypical 2-Cys Prxs also include a different type of Prx, known as Prx BCP (bacterioferritin comigratory proteins) in bacteria or Prx Q in plants. These have been described as monomeric enzymes and contain one or two conserved cysteines depending on the species. The first is the catalytic cysteine and the second, when present, is located five amino acids downstream (in *Escherichia coli* and in poplar). In the poplar enzyme the mutation of the first of these cysteines results in an inactive enzyme, but the enzyme lacking the second cysteine is still active provided the thioredoxin concentration

is increased (Rouhier *et al.*, 2004). Interestingly, the second cysteine mutant can also be reduced *via* glutaredoxin, while the wild-type protein is only regenerated *via* the thioredoxin system. Nevertheless, an intramolecular disulfide bridge can be formed between these two cysteines in plant enzymes and, as the presence of these two cysteines leads to a more active enzyme, the physiological regeneration mechanism used by the enzyme would preferably involve the formation of this particular disulfide bridge. In *E. coli*, it has been shown that a mutated protein possessing only the peroxidatic cysteine is still active and is regenerated by thioredoxin (Jeong *et al.*, 2000). Again, the classification is not adapted to this class of Prx as no intramolecular disulfide bridge can be formed in proteins that possess only one conserved cysteine. These intriguing biochemical data raise interest concerning the function of Prx BCP sequences in which the second cysteine is naturally absent. One of these sequences is present in *Thermotoga maritima* as a domain of a larger protein (321 amino acids; Genbank accession No. AAD35471) constituted of a natural hybrid between a Prx BCP module at the N-terminus (∼140 amino acids) and a nitroreductase module at the C-terminus (∼170 amino acids). Here, we present crystallization and diffraction data concerning the Prx BCP module of this larger protein.

2. Experimental procedures

2.1. Cloning, overproduction and purification

The Prx module of a hybrid protein from *T. maritima* constituted of a peroxiredoxin domain followed by a nitroreductase domain was cloned separately in pET3d using the primers 5′-CCCC**CCATGG**C-TAGGGTGAAGCACTTTGAA-3′ and 5′-CCCC**GGATCC**TTAGT-CTTCTTCTATCAGCG-3′ (*Nco*I and *Bam*HI restriction sites shown in bold). This module starts with the amino-acid sequence MARVKHF and ends with RLIEED. As many arginine codons unfavourable for expression in *E. coli* (AGG and AGA) are present in this sequence, the protein was overproduced in the *E. coli* BL21(DE3) strain in the presence of the helper plasmid pSBET (Schenk *et al.*, 1995). 2.6 l LB cultures were grown at 310 K and induced in the exponential phase by the addition of 100 µ*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG). After 4 h induction, the bacteria were pelleted by centrifugation for 15 min at 5000*g* and resuspended in buffer *A* (30 m*M* Tris–HCl pH 8.0, 1 m*M* EDTA, 200 m*M* NaCl). Bacteria were sonicated and centrifuged for 1 h at 16 000*g* to eliminate insoluble material. The soluble fraction was then precipitated with ammonium sulfate and the fraction precipitating between 40 and 80% saturation was collected. The ammonium sulfate precipitate was redissolved and loaded onto an ACA44 gel-filtration column (5 × 75 cm; BioSepra) equilibrated with buffer *A*. The fractions of interest were pooled, dialyzed against buffer *B* (buffer *A* without NaCl) and loaded onto a DEAE Sephacel column equilibrated with buffer *B*. The protein was eluted using a 0–0.4 *M* NaCl gradient, dialyzed again using buffer *B*, concentrated and finally stored at 243 K for crystallization trials.

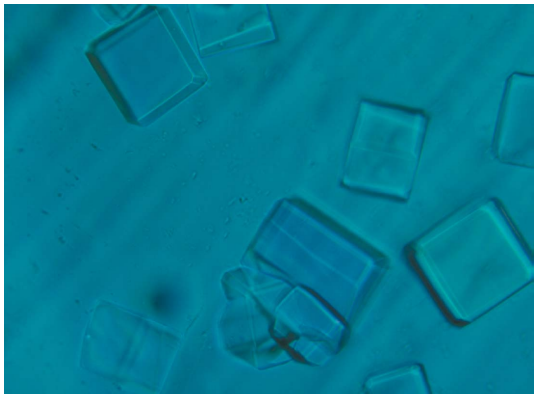


Figure 1
Crystals of the peroxiredoxin domain. The scale bar is 0.1 mm in length.

2.2. Crystallization strategy

Crystallization conditions for the purified peroxiredoxin domain were screened using the vapour-diffusion method with a Cartesian Technology workstation. Sitting drops composed of 200 nl protein solution at a concentration higher than 5 mg ml^{-1} and 200 nl mother liquor were equilibrated against 150 µl well solution on a Greiner plate (Santarsiero *et al.*, 2002). In this study, we used a Greiner microcrystallization plate with 96-reservoir wells and 288 sitting-drop circle shelves to perform the screening (Mueller *et al.*, 2001). Initial screens were performed using the commercially available sparse-matrix kits Structure Screens 1 and 2 from Molecular Dimensions Ltd, JBScreens 1–8 from Jena Biosciences and Crystal Screens I and II from Hampton Research. A Genesis Workstation 150 robot from Tecan was used to dispense the crystallization solutions (150 µl) in the well reservoirs.

The protein was deposited first into all wells, followed by line dispensing of the crystallizing agent into the 96 wells with an eight-tip Cartesian robot. The microplate was then manually sealed with a transparent self-adhesive foil (CrystalClear) purchased from Hampton Research and incubated at 291 K.

Images of the drops were taken 1, 3, 7 and 30 d after mixing the protein with the crystallizing agents. We used a Nikon microscope (Eclipse E600) equipped with a DXM 1200 video camera and an *xy* computer driver plate holder. The *Lucia* version G software (Nikon) was used to obtain and analyze the images. We obtained crystallization hits in many conditions using ammonium sulfate as a crystallizing agent, especially with JBScreen 6 conditions A2 and C4.

Crystals obtained from drops set up using the Cartesian robot were small or in needle-cluster form. In order to increase their quality and size, they were reproduced manually using the hanging-drop vapour-diffusion method in Linbro 24-well plates. We also tested the effect of varying the crystallizing agent or protein concentration, pH and drop volume (McPherson, 1999).

The best crystals were obtained by mixing 1.5 µl 24.8 mg ml^{-1} protein solution in 30 m*M* Tris–HCl pH 8.0, 1 m*M* EDTA with 1.5 µl reservoir solution containing 1.8 *M* ammonium sulfate, 0.1 *M* Tris–HCl pH 8.5, 7.5%(*v*/*v*) ethylene glycol. The crystals grew within a week at 291 K.

2.3. Collection and processing of diffraction data

Crystals were prepared for cryocrystallography by transferring them to reservoir solution containing 20%(*v*/*v*) glycerol. A single crystal was picked up with a cryo-loop (Hampton Research) and flash-cooled to 100 K in a nitrogen-gas stream (Rodgers, 1994).

Diffraction data were collected on beamline ID14-1 at the European Synchrotron Radiation Facility. 1° oscillation photographs were collected. X-ray data were indexed and integrated using *MOSFLM* v.6.2.6 for image plates and CCDs (Leslie, 1992). Intensities were further converted to amplitudes and scaled using programs from the *CCP*4 suite v.6.0 (Collaborative Computational Project, Number 4, 1994; Evans, 1997).

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Beamline	ID14-1, ESRF, France
Space group	*I*422
Unit-cell parameters (Å)	$a = b = 176.67$, $c = 141.20$
Resolution range (Å)	30–2.90 (3.06–2.90)
Measured reflections	174631
Unique reflections	25000
Completeness (%)	99.9 (100.0)
Redundancy (multiplicity)	7.0 (6.8)
$R_{\rm merge}$ (%)	11.3 (42.0)
Average $I/\sigma(I)$	16.1 (3.9)

3. Results and discussion

Crystals grew as small cubes as shown in Fig. 1. The dimensions of the crystal used for the final data collection were 200 × 200 × 300 µm. From the symmetry of the diffraction data set and from the absence of systematic extinctions along the fourfold axis, the space group was determined to be *I*422, with unit-cell parameters $a = b = 176.67$, $c = 141.20$ Å, $\alpha = \beta = \gamma = 90^\circ$. The crystal showed ordered diffraction to 2.9 Å resolution, with a mosaicity of around 0.5°. Details of the data-collection statistics are reported in Table 1. The symmetry of the crystal implies that at least six and possibly as many as eight molecules are present in the asymmetric unit; these would correspond to 51 or 35% solvent content, respectively. Preliminary calculations deduced from the self-rotation function (*POLARRFN*; Collaborative Computational Project, Number 4, 1994) calculated using standard parameters emphasize significant noncrystallographic axes: twofold and threefold axes (peaks that reach a maximum of 37% of the normalized value of the crystallographic fourfold axis appear at coordinates $\omega = 45$, $\varphi = 90$, $\kappa = 180^\circ$ and $\omega = 55$, $\varphi = 45$, $\kappa = 120^\circ$) are found in agreement with a trimer of homodimers, but the presence of a fourfold axis ($\omega = 90$, $\varphi = 0$, $\kappa = 90^\circ$) corresponding to an octamer solution is also indicated. Molecular-replacement trials using the program *AMoRe* (Navaza, 1994) are currently in progress and are being performed using various Prx structures available in the PDB as search models. To date, a large number of peroxiredoxin structures (43) have been deposited in the PDB. Most of them are of typical 2-Cys peroxiredoxins, 1-Cys peroxiredoxins and type II peroxiredoxins. Yeast nuclear thiol peroxidase (PDB code 2a4v), *Plasmodium falciparum* mitochondrial 2-Cys peroxiredoxin (PDB code 2c0d) and *Aeropyrum pernix* K1 peroxiredoxin (PDB codes 2cx3 and 2cx4) exhibit the largest amino-acid sequence identity to the *T. maritima* peroxiredoxin module described here (30–31%; Altschul *et al.*, 1997), but only the yeast and *A. pernix* proteins belong to the Prx Q/Prx BCP family. In addition, these two proteins possess the two highly conserved cysteines that are present in approximately half of the Prx BCP family members, whereas the Prx domain of the *T. maritima* hybrid protein displays only the peroxidatic cysteine and no putative resolving cysteines. The regeneration mechanism used by this Prx is thus presently unknown and as there is no identified thioredoxin (the usual reductant for Prx Q and BCP) in the genome of *T. maritima*, it is expected that this structural study should provide data that will lead to a better understanding of the function and regeneration of these enzymes. Last but not least, this Prx module is associated with a nitroreductase module and whether and how the two modules interact in catalysis is still uncertain, especially as the structure of the full-length protein is not available. The low identity between the Prx domain of *T. maritima* and these models, together with the large number of independent molecules to place, have so far prevented an unambiguous identification of the molecular-replacement solution. Different modifications of the coordinates (*i.e.* the replacement of the non-aligned residues or of all the residues by alanines or the use of a poly-C_α model) and the use of oligomeric models are currently being performed in order to determine the phases by molecular replacement.

We acknowledge the European Synchrotron Radiation Facility for providing access to synchrotron radiation and would like to thank Emanuela Fioravanti for assistance in using beamline ID14-1, Philippe Benas and Thierry Prangé for their help during data collection and processing, and Jorge Navaza for access to novel features of the *AMoRe* program.

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.
Chae, H. Z., Chung, S. J. & Rhee, S. G. (1994). *J. Biol. Chem.* **269**, 27670–27678.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.
Herbette, S., Lenne, C., Leblanc, N., Julien, J. L., Drevet, J. R. & Roeckel-Drevet, P. (2002). *Eur. J. Biochem.* **269**, 2414–2420.
Hofmann, B., Hecht, H.-J. & Flohé, L. (2002). *Biol Chem.* **383**, 347–364.
Jeong, W., Cha, M. K. & Kim, I. H. (2000). *J. Biol. Chem.* **275**, 2924–2930.
Kang, S. W., Baines, I. C. & Rhee, S. G. (1998). *J. Biol. Chem.* **273**, 6303–6311.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
McPherson, A. (1999). *Crystallization of Biological Macromolecules.* New York: Cold Spring Harbor Laboratory Press.
Monteiro, G., Horta, B. B., Pimenta, D. C., Augusto, O. & Netto, L. E. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 4886–4891.
Mueller, U., Nyarsik, L., Horn, M., Rauth, H., Przewieslik, T., Saenger, W., Lehrach, H. & Eickhoff, H. (2001). *J. Biotechnol.* **85**, 7–14.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Navrot, N., Collin, V., Gualberto, J., Gelhaye, E., Hirasawa, M., Rey, P., Knaff, D. B., Issakidis, E., Jacquot, J. P. & Rouhier, N. (2006). *Plant Physiol.* **142**, 1364–1379.
Rodgers, D. W. (1994). *Structure*, **2**, 1135–1140.
Rouhier, N., Gelhaye, E., Gualberto, J. M., Jordy, M. N., de Fay, E., Hirasawa, M., Duplessis, S., Lemaire, S. D., Frey, P., Martin, F., Manieri, W., Knaff, D. B. & Jacquot, J. P. (2004). *Plant Physiol.* **134**,1027–1038.
Rouhier, N., Gelhaye, E. & Jacquot, J. P. (2002). *J. Biol. Chem.* **277**, 13609–13614.
Rouhier, N., Gelhaye, E., Sautiere, P. E., Brun, A., Laurent, P., Tagu, D., de Fay, E., Meyer, Y. & Jacquot, J. P. (2001). *Plant Physiol.* **127**, 1299–1309.
Rouhier, N. & Jacquot, J. P. (2002). *Photosynth. Res.* **74**, 259–268.
Rouhier, N. & Jacquot, J. P. (2005). *Free Radic. Biol. Med.* **38**, 1413–1421.
Santarsiero, B. D., Yegian, D. T., Lee, C. C., Spraggon, G., Gu, J., Scheibe, D., Uber, D. C., Cornell, E. W., Nordmeyer, R. A., Kolbe, W. F., Jin, J., Jones, A. L., Jaklevic, J. M., Schultz, P. G. & Stevens, R. C. (2002). *J. Appl. Cryst.* **35**, 278–281.
Schenk, P. M., Baumann, S., Mattes, R. & Steinbiss, H. H. (1995). *Biotechniques*, **19**, 196–200.
Tanaka, T., Izawa, S. & Inoue, Y. (2005). *J. Biol. Chem.* **280**, 42078–42087.
Wood, Z. A., Schröder, E., Harris, J. R. & Poole, L. B. (2003). *Trends Biochem. Sci.* **28**, 32–40.