Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of an acylphosphatase from the hyperthermophilic archaeon *Pyrococcus horikoshii*

Yuk-Yin Cheung,[a] Mark D. Allen,[b] Mark Bycroft[b] and Kam-Bo Wong[a]*

[a]Department of Biochemistry, The Chinese University of Hong Kong, Hong Kong, and [b]MRC Centre for Protein Engineering, Hills Road, Cambridge CB2 2QH, England

Correspondence e-mail: kbwong@cuhk.edu.hk

Received 24 March 2004
Accepted 5 May 2004

Acylphosphatases catalyse the hydrolysis of the carboxyl phosphate bond in metabolites such as acetyl phosphate, 1,3-bisphosphoglycerate, succinoyl phosphate and carbamoyl phosphate. In this study, acylphosphatase (91 residues) from the hyperthermophilic archaeon *Pyrococcus horikoshii* has been cloned, overexpressed, purified and crystallized using the sitting-drop vapour-diffusion method using sodium formate as a precipitant at 289 K. The crystals belong to space group $P3_221$, with unit-cell parameters $a = b = 85.65$, $c = 75.51$ Å. The asymmetric unit contains two molecules of acylphosphatase, with a corresponding crystal volume per protein weight of 3.9 Å Da^{-1} and a solvent content of 68.6%. A data set diffracting to 1.6 Å resolution was collected from a single crystal at 100 K.

1. Introduction

Acylphosphatase (EC 3.6.1.7) is one of the smallest enzymes known (~10 kDa). It catalyses the hydrolysis of carboxyl phosphate bonds. Its substrates include 1,3-bisphosphoglycerate, succinoyl phosphate and carbamoyl phosphate, which are intermediates in glycolysis, the tricarboxylic acid cycle, pyrimidine and urea biosynthesis (Stefani *et al.*, 1997). Acylphosphatase has been isolated and characterized from a number of mammalian and avian species (*e.g.* horse, cow, human, pig, chicken, turkey; Stefani *et al.*, 1997), as well as from fish (Pazzagli *et al.*, 1998) and fruit fly (Degl'Innocenti *et al.*, 2003; Pieri *et al.*, 1998). Acylphosphatases from vertebrate sources are classified into two isoforms: muscle-type and common-type acylphosphatases. The structure of horse muscle-type acylphosphatase has been solved by NMR spectroscopy (Pastore *et al.*, 1992) and the structure of bovine common-type acylphosphatase has been solved by crystallography (Thunnissen *et al.*, 1997). The structure of a closely related protein, the N-terminal domain of hydrogenase maturation factor from *Escherichia coli*, has recently been solved (Rosano *et al.*, 2002). However, this domain does not possess any acylphosphatase activity, despite sharing both sequence and structural homology with acylphosphatase. There is no structure of any archaeal acylphosphatase available to date.

Mesophilic acylphosphatases have been studied extensively as a model for insoluble fibril formation in amyloid diseases (Chiti *et al.*, 1999, 2000, 2003; Chiti, Calamai *et al.*, 2002; Chiti, Taddei *et al.*, 2002; Taddei *et al.*, 2001). The protein has been induced to form *in vitro* amyloid fibrils that closely resemble those observed in clinical cases (Chiti *et al.*, 1999). Site-directed mutagenesis studies have shown that destabilizing mutations promote protein aggregation by populating a partially unfolded intermediate that serves as a precursor to insoluble aggregates (Chiti *et al.*, 2000). Recently, the aggregation rate was shown to correlate well with simple biophysical parameters such as hydrophobicity, secondary-structure propensity and net charges (Chiti *et al.*, 2003).

Acylphosphatase (91 residues) from *Pyrococcus horikoshii*, a hyperthermophilic archaeon that grows optimally at 371 K, is an excellent model for the study of how proteins remain stable and active at high temperatures. It is extremely thermostable and is resistant to denaturation at over 363 K (unpublished data), while the melting temperature of the mesophilic homologue is ~330 K (Taddei *et al.*, 1999). To our knowledge, no structures of thermophilic acylphosphatases have been solved. A structural comparison between thermophilic and mesophilic homologous acylphosphatase will provide important insights into the structural determinants of protein thermostability. As a first step toward structure determination, we report the crystallization, preliminary crystallographic analysis and molecular-replacement solution of acylphosphatase from *P. horikoshii*, for which diffraction data to 1.6 Å resolution were collected.

2. Expression and purification

The AcP coding sequence was obtained by polymerase chain reaction from genomic DNA of *P. horikoshii* and was subcloned into

Table 1
Data-processing statistics for *P. horikoshii* acylphosphatase.

Values in parentheses are for the highest resolution shell (1.69–1.60 Å)

Resolution range (Å)	26.44–1.60
No. measurements	878617
No. unique reflections	40890
Redundancy	21.5 (20.2)
Completeness (%)	96.3 (91.1)
$R_{\rm merge}$† (%)	7.5 (22.0)
Mean $I/\sigma(I)$	37.8 (10.9)

† $R_{\rm merge} = \sum_{hkl} |I - \langle I \rangle| / \sum_{hkl} I$.

pET507a, an in-house modified vector in which a multiple cloning site was inserted between the *Nco*I and *Bam*HI sites of the pET3d vector (Novagen). The expression plasmid was transformed into *E. coli* C41 strain (Miroux & Walker, 1996). The bacterial culture was grown in M9ZB medium at 310 K until the absorbance at 600 nm reached 0.4–0.8, when 0.2 m*M* isopropyl-β-thiogalactopyranose was added in order to induce protein expression. Cells were harvested after 16 h and then lysed by sonication in buffer *A* (20 m*M* Tris pH 8.5) with 0.5 m*M* phenylmethylsulfonyl fluoride. The lysed bacteria were centrifuged at 13 000*g* for 30 min. The supernatant was then heated at 333 K for 15 min to denature heat-labile *E. coli* proteins, which were removed by centrifugation at 13 000*g* for 30 min. The supernatant was loaded onto a HiTrap SP HP column (Amersham Biosciences) pre-equilibrated with buffer *A*. The protein was eluted using a linear gradient of 0–1.0 *M* NaCl in buffer *A* over a volume of 300 ml. The eluted protein was then loaded onto a HiLoad 26/60 Superdex 75 column (Amersham Biosciences) pre-equilibrated with 20 m*M* Tris, 0.2 *M* NaCl pH 8.5. The purified protein was dialysed against buffer *A* and concentrated to 24 mg ml^{-1}.

3. Crystallization and data processing

Screening for crystallization conditions was performed by the sitting-drop vapour-diffusion method at 290 K using kits (Crystal Screen I and II, Index) from Hampton Research. Each drop contained 1 µl 24 mg ml^{-1} protein sample and 1 µl reservoir solution. The initial conditions obtained were 2.0 *M* sodium formate, 0.1 *M* sodium acetate pH 4.6 (solution 34 of Crystal Screen I from Hampton Research). Crystallization conditions were optimized by a grid screen with 0.5–4.0 *M* sodium formate pH 4–8. Crystals for diffraction data collection were grown using the sitting-drop vapour-diffusion method at 290 K by mixing 2 µl 24 mg ml^{-1} protein sample with 2 µl reservoir solution containing 1.2 *M* sodium formate, 0.1 *M* MES pH 6.0 on a Cryschem Plate (Hampton Research) with 500 µl reservoir solution. Hexagonal crystals grew to maximum dimensions of 0.3 × 0.3 × 1 mm in 48 h. The crystals were cryoprotected by soaking crystals in mother-liquor solution containing 20%(*v*/*v*) glycerol for 1 min at room temperature. The crystals were then loop-mounted and flash-cooled in liquid nitrogen before transfer into the cryostream at 100 K. X-ray diffraction data were collected at a wavelength of 1.488 Å at beamline 14.1 of Daresbury Synchrotron Radiation Source using an ADSC Quantum-4R CCD detector. A total of 360 images were collected from a single crystal in 1° oscillation steps with a crystal-to-detector distance of 66.9 mm and an exposure time of 60 s per image (Fig. 1).

The diffraction data were indexed, integrated, scaled and merged using the programs *MOSFLM*, *SCALA* and *TRUNCATE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994).

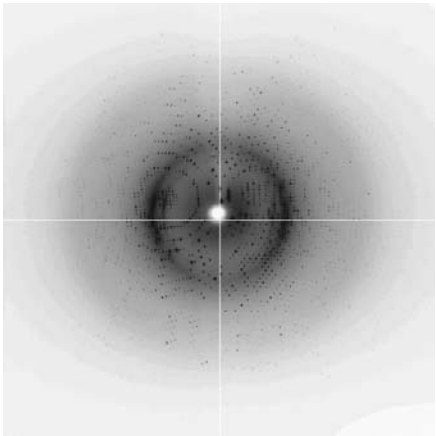

Figure 1
A diffraction image of *P. horikoshii* acylphosphatase collected at beamline 14.1, Daresbury SRS, UK. The edge of the frame corresponds to 1.6 Å resolution.

Table 2
The alignment-improved model yielded a clearer solution for molecular replacement.

Only the results for the translation search in the $P3_221$ space group are shown. Two monomers per asymmetric unit were found and their Eulerian angles and fractional coordinates are shown. The values in parentheses are those found for the next peaks.

	Eulerian angles			Fractional coordinates				
Search model	α (°)	β (°)	χ (°)	*x*	*y*	*z*	Correlation coefficient	*R* factor
Polyserine	13.40	15.63	107.62	0.020	0.560	0.075		
	142.66	167.02	−142.63	−0.220	0.439	0.176	0.188 (0.185)	0.573 (0.574)
Polyalanine	14.31	15.36	107.66	0.020	0.560	0.075		
	138.34	166.87	−144.74	−0.225	0.443	0.179	0.167 (0.163)	0.570 (0.582)
All-atom	15.35	156.55	207.57	0.660	0.218	0.500		
	80.11	−19.39	−77.31	0.534	−0.017	0.407	0.249 (0.244)	0.547 (0.549)
Alignment-improved	20.84	14.78	100.02	0.030	0.559	0.094		
	136.71	159.51	−152.89	−0.230	0.443	0.153	0.316 (0.285)	0.527 (0.541)

Pseudo-precession images generated with the program *HKLVIEW* (Collaborative Computational Project, Number 4, 1994) revealed $\bar{3}m$ Laue symmetry. The reflection condition $l = 3n$ was observed, indicating the presence of a threefold screw axis parallel to the *l* axis. The space group belongs to one of the enantiomorphic pair $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 85.65$, $c = 75.51$ Å. Assuming the asymmetric unit to contain two, three or four molecules of protein, the solvent content of the crystal is 68.6, 52.9 or 37.2% and the Matthews constant is 3.9, 2.6 or 2.0 Å^3 Da^{-1}, respectively. The statistics for data processing are summarized in Table 1.

4. A molecular-replacement solution was obtained only when an improved search model based on sequence alignment was used

Molecular replacement was performed with the program *MOLREP* (Vagin & Teplyakov, 1997) using all data in the resolution range 26.4–3.5 Å, with translation searches performed in both $P3_121$ and $P3_221$ space groups. The crystal structure of bovine acylphosphatase (PDB code 2acy), with a sequence identity of 31%, was used as the search model. No obvious solution was obtained when polyserine, polyalanine and all-atom models were used: the correlation coefficients and the *R* factor for the best peaks were similar to those of the next peaks (Table 2).

A clearer solution was obtained when an improved search model based on sequence alignment was used (Table 2). To create this alignment-improved model, the sequences of the bovine and *P. horikoshii* acylphosphatases were first aligned using the program *CLUSTALW* (Thompson *et al.*, 1994). For the 30 residues (highlighted in Fig. 2) that were identical in both sequences, all atoms were retained in the search model.

Figure 2
Sequence of the alignment-improved model. Sequences of the *P. horikoshii* and bovine common-type acylphosphatases were aligned using the program *CLUSTALW* (Thompson *et al.*, 1994). Conserved residues were highlighted, while residues that were found only in the bovine sequences are underlined.

For the non-conserved residues, side chains were truncated to alanine or glycine if the corresponding residues in the *P. horikoshii* sequence were alanine or glycine, respectively. The side chains of other larger residues were truncated to serine. The atoms of those residues that are absent in the *P. horikoshii* sequence (underlined residues in Fig. 2) were omitted from the search model. The above algorithm was implemented in an in-house-written Perl script, which is available from the author's home page (http://smart.bch.cuhk.edu.hk/kbwong/index.htm). The sequence of the alignment-improved model is shown in Fig. 2.

Using the alignment-improved model, a solution with two protein molecules per asymmetric unit was found in space group $P3_221$. The correlation coefficient and R factor of the solution were 0.316 and 0.527, compared with values of 0.285 and 0.541 for the next best solution (Table 2). Searching for additional protein molecules did not improve the correlation coefficient and R factor, confirming the presence of two protein molecules in the asymmetric unit. No unfavourable molecular contacts were observed in the crystal packing. After rigid-body refinement and energy minimization using the program *CNS* (Brünger *et al.*, 1998), the $R_{\rm cryst}$ and $R_{\rm free}$ factors fell to 44 and 47%, respectively. Model building and structure refinement are now under way.

In order to understand the structural basis of the thermostability of *P. horikoshii* AcP, we are preparing its mutants to measure their thermodynamic parameters (such as melting temperature and free energy of unfolding). The crystal structure of *P. horikoshii* AcP will be reported together with these thermodynamic results.

We thank Professor Alan Fersht for his generous support during KBW's sabbatical visit to Cambridge. This work was partially supported by a Strategic Grant (SRP0102) from the Chinese University of Hong Kong.

References

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.
Chiti, F., Calamai, M., Taddei, N., Stefani, M., Ramponi, G. & Dobson, C. M. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 16419–16426.
Chiti, F., Stefani, M., Taddei, N., Ramponi, G. & Dobson, C. M. (2003). *Nature (London)*, **424**, 805–808.
Chiti, F., Taddei, N., Baroni, F., Capanni, C., Stefani, M., Ramponi, G. & Dobson, C. M. (2002). *Nature Struct. Biol.* **9**, 137–143.
Chiti, F., Taddei, N., Bucciantini, M., White, P., Ramponi, G. & Dobson, C. M. (2000). *EMBO J.* **19**, 1441–1449.
Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G. & Dobson, C. M. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 3590–3594.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Degl'Innocenti, D., Ramazzotti, M., Marzocchini, R., Chiti, F., Raugei, G. & Ramponi, G. (2003). *FEBS Lett.* **535**, 171–174.
Miroux, B. & Walker, J. E. (1996). *J. Mol. Biol.* **260**, 289–298.
Pastore, A., Saudek, V., Ramponi, G. & Williams, R. J. (1992). *J. Mol. Biol.* **224**, 427–440.
Pazzagli, L., Manao, G., Cappugi, G., Caselli, A., Camici, G., Moneti, G. & Ramponi, G. (1998). *Biochim. Biophys. Acta*, **1387**, 264–274.
Pieri, A., Magherini, F., Liguri, G., Raugei, G., Taddei, N., Bozzetti, M. P., Cecchi, C. & Ramponi, G. (1998). *FEBS Lett.* **433**, 205–210.
Rosano, C., Zuccotti, S., Bucciantini, M., Stefani, M., Ramponi, G. & Bolognesi, M. (2002). *J. Mol. Biol.* **321**, 785–796.
Stefani, M., Taddei, N. & Ramponi, G. (1997). *Cell Mol. Life Sci.* **53**, 141–151.
Taddei, N., Capanni, C., Chiti, F., Stefani, M., Dobson, C. M. & Ramponi, G. (2001). *J. Biol. Chem.* **276**, 37149–37154.
Taddei, N., Chiti, F., Paoli, P., Fiaschi, T., Bucciantini, M., Stefani, M., Dobson, C. M. & Ramponi, G. (1999). *Biochemistry*, **38**, 2135–2142.
Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Thunnissen, M. M., Taddei, N., Liguri, G., Ramponi, G. & Nordlund, P. (1997). *Structure*, **5**, 69–79.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.