A. A. Trofimov,^{a,b,*}
E. A. Slutskaya,^b
K. M. Polyakov,^{a,b,c}
P. V. Dorovatovskii,^c
V. M. Gumerov^d and
V. O. Popov^{b,c}

^aEngelhardt Institute of Molecular Biology, Russian Academy of Sciences, ul. Vavilova 32, Moscow 119991, Russian Federation, ^bBach Institute of Biochemistry, Russian Academy of Sciences, Leninsky pr. 33, Moscow 119071, Russian Federation, ^cRussian Research Centre 'Kurchatov Institute', Academic Kurchatov sq. 1, Moscow 123182, Russian Federation, and ^dCentre 'Bioengineering', Russian Academy of Sciences, Prosp. 60-let Oktyabrya 7-1, Moscow 117312, Russian Federation

Correspondence e-mail: burayatina@gmail.com

Received 18 July 2012

Accepted 31 August 2012

PDB Reference: Tspro, 4fkc

Influence of intermolecular contacts on the structure of recombinant prolidase from *Thermococcus sibiricus*

Prolidases are peptidases that are specific for dipeptides with proline as the second residue. The structure of recombinant prolidase from the hyperthermophilic archaeon *Thermococcus sibiricus* (Tsprol) was determined at 2.6 Å resolution. The homodimer of Tsprol is characterized by a complete lack of interactions between the N- and C-terminal domains of the two subunits and hence can be considered to be the most open structure when compared with previously structurally studied prolidases. This structure exists owing to intermolecular coordination bonds between cadmium ions derived from the crystallization solution and histidine residues of a His tag and aspartate and glutamate residues, which link the dimers to each other. This linking leads to the formation of a crystal with a loose packing of protein molecules and low resistance to mechanical influence and temperature increase.

1. Introduction

Prolidases are metallopeptidases that specifically hydrolyze dipeptides with proline at the C-terminus and a nonpolar amino acid at the N-terminus (Lowther & Matthews, 2002). The majority of prolidases that have been structurally characterized to date are homodimers with a subunit made up of two domains: an N-terminal domain, which may vary significantly in different enzymes, and a C-terminal catalytic domain, which is characterized by a highly conserved pitta-bread structure (Maher *et al.*, 2004; Jeyakanthan *et al.*, 2009). The active site is located in the catalytic domain and contains two metal ions coordinated by conserved residues: two Asp residues, two Glu residues and a His residue. Under aerobic conditions, prolidase exhibits maximal activity if it has two Co²⁺ ions or two Mn²⁺ ions bound in the active site (Lowther & Matthews, 2002). Under anaerobic conditions, maximal activity is achieved with two Fe²⁺ ions in the active site (Du *et al.*, 2005).

In this work, the crystal structure of recombinant prolidase (Tsprol) from the hyperthermophilic anaerobic organotrophic archaeon *Thermococcus sibiricus* is characterized and the joint effect of a His tag and cadmium ions from the crystallization solution on the structure is discussed.

2. Materials and methods

2.1. Cloning and expression

The gene Tsib_0821 encoding a putative prolidase was identified in the complete genome sequence of *T. sibiricus* (Mardanov *et al.*, 2009). Tsib_0821 (GenBank accession No. ACS89882.1; 365 amino-acid residues) was PCR-amplified from *T. sibiricus* genomic DNA using the primers F_BamHI (5'-TTGGATCCATGGATTACAAGAGA-AG-3') and R_HindIII (5'-TTAAGCTTCACCCCAAACATC-3'). The PCR fragment was digested with BamHI and HindIII and inserted into the expression vector pQE80L (Qiagen), which enables the production of a recombinant protein containing a six-histidine tag at the N-terminus (MRGSHHHHHGS). The recombinant vector,

Figure 1
Crystals of *Tsprol*.

pQE80L_TSIB0821, was introduced into *Escherichia coli* strain Rosetta-gami (DE3) (Novagen). The transformants were grown at 310 K in Luria–Bertani medium to an optical density of 0.5 at 600 nm. Isopropyl β -D-1-thiogalactopyranoside was added to a final concentration of 1 mM for induction of gene expression; cultivation was then continued at 310 K for an additional 20 h.

2.2. Purification of the recombinant prolidase *Tsprol*

The induced cells were harvested by centrifugation and resuspended in 50 mM Tris–HCl buffer pH 7.5 containing 200 mM NaCl, 20 mM imidazole, 0.1% (v/v) Triton X-100, 10% (v/v) glycerol. The cells were disrupted using a French press and centrifuged at 10 000g for 30 min. The supernatant was heated at 343 K for 10 min and the lysate was clarified by centrifugation at 10 000g for 30 min. The supernatant was loaded onto a HisTrap HP column (Amersham Biosciences) equilibrated with 50 mM Tris–HCl pH 7.5 containing 500 mM NaCl, 20 mM imidazole, 0.1% (v/v) Triton X-100. Recombinant *Tsprol* was eluted with a linear gradient of imidazole (20–500 mM) in the same buffer without Triton X-100. *Tsprol* activity was detected in one peak. Protein purity was analyzed by SDS–PAGE; the molecular weight was consistent with that calculated from the amino-acid sequence of recombinant *Tsprol* (41.1 kDa). The protein was transferred into 50 mM Tris–HCl buffer pH 7.5 containing 150 mM NaCl using a Sephadex G-25 Superfine column (HiTrap Desalting column, Amersham Biosciences) and concentrated using a 30 kDa cutoff centrifugal filter device (Millipore). Immediately after this, the protein was used for crystallization.

2.3. Crystallization

A search for crystallization conditions was performed by the hanging-drop vapour-diffusion method at 277 K using kits from Hampton Research. 3 μ l drops were prepared by combining equal volumes of protein solution and reservoir solution. The protein

Table 1

Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Space group	$P4_12_12$
Unit-cell parameters ($\text{\AA}$, $^\circ$)	$a = b = 89.7$, $c = 165.2$, $\alpha = \beta = \gamma = 90$
X-ray source	Kurchatov SNC, Moscow
Wavelength ($\text{\AA}$)	0.9887
Temperature (K)	100
Oscillation range ($^\circ$)	1
No. of frames	153
Mosaicity ($^\circ$)	0.27
Resolution ($\text{\AA}$)	19.9–2.6 (2.8–2.6)
Total No. of reflections	253240 (47886)
No. of unique reflections	21360 (4182)
Completeness (%)	99.2 (99.8)
$R_{\text{meas}}^\dagger$	0.089 (0.728)
Mean $I/\sigma(I)$	24.5 (4.2)
Matthews coefficient ($\text{\AA}^3 \text{Da}^{-1}$)	4.14
B factor from Wilson plot ($\text{\AA}^2$)	69.1
R_{work}	0.231
R_{free}	0.269
Root-mean-square deviations	
Bond lengths ($\text{\AA}$)	0.012
Bond angles ($^\circ$)	1.384
No. of protein residues	370
No. of cadmium ions	12
No. of water molecules	53
Ramachandran plot statistics, residues in (%)	
Most favoured regions	89.3
Additionally allowed regions	9.5
Generously allowed regions	0.6
Disallowed regions	0.6

$^\dagger R_{\text{meas}} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where N is the multiplicity of the observed reflection hkl .

solution consisted of 10.8 mg ml $^{-1}$ *Tsprol* and 150 mM NaCl in 50 mM Tris buffer pH 7.6. *Tsprol* crystals grew in several conditions, but a resolution better than 3 $\text{\AA}$ was only achieved using crystals obtained from condition No. 12 of Crystal Screen 2, which consisted of 100 mM CdCl $_2$ and 30% (v/v) PEG 400 in 100 mM sodium acetate buffer pH 4.6. These crystals grew within 3 d and had maximal dimensions of 0.3 mm (Fig. 1).

2.4. Data collection, structure solution and refinement

The *Tsprol* crystals were not stable upon transfer to room temperature and were easily disrupted by touching them with a plastic loop. Crystals fished out from the drop and flash-cooled in a nitrogen stream gave diffraction circles consistent with water ice formation. Therefore, *Tsprol* crystals needed to be immersed in their reservoir solution for a few seconds before flash-cooling. A full data set was collected from one crystal to 2.6 $\text{\AA}$ resolution on beamline K4.4 of the Kurchatov Synchrotron Nano Centre (SNC), Moscow using a MAR CCD detector. Data were processed using *XDS* (Kabsch, 2010) and *SCALEPACK* (Otwinowski & Minor, 1997).

Figure 2
The dimer of *Tsprol*. One *Tsprol* subunit is represented as a ribbon model and the other is shown as a worm model. The His tag is coloured cyan; the N-terminal domain, the linker and the C-terminal domain are coloured green/yellow, grey and red/pink, respectively. Cadmium ions in the active site are shown as cyan spheres.

The structure of *Ts*prol was solved using *BALBES* (Long *et al.*, 2008). Subunit A of prolidase from *Pyrococcus horikoshii* OT3 (2.25 Å resolution; PDB entry 1wn1; RIKEN Structural Genomics/Proteomics Initiative, unpublished work) was used as a starting model for solution of the *Ts*prol structure using the molecular-replacement method. There is one *Ts*prol subunit per asymmetric unit. The structure was refined using *REFMAC* (Murshudov *et al.*, 2011) and manual corrections to the model in accordance with the electron-density maps were made using *Coot* (Emsley & Cowtan, 2004). The data-collection and refinement statistics are summarized in Table 1. Structural data are available in the RCSB Protein Data Bank as entry 4fk. Structural figures were produced using *CCP4mg* (Winn *et al.*, 2011).

3. Results and discussion

The structure of *Ts*prol was refined to an R_{work} of 0.231 ($R_{\text{free}} = 0.269$) at 2.6 Å resolution. The model contained 365 amino-acid residues of the native protein and five residues of the His tag (HHHGS) and fitted to the electron-density map without gaps.

In the crystal, *Ts*prol reveals a homodimeric structure (Fig. 2). The subunit of *Ts*prol is composed of an N-terminal domain (residues 6–132), an α -helical linker (residues 133–142) and a C-terminal catalytic domain (residues 143–370), like previously structurally studied prolidases. The active site of *Ts*prol contains a characteristic dinuclear metal cluster coordinated by the side chains of Asp240, Asp229, Glu347, Glu333 and His304 (Fig. 3). Cadmium ions were placed in the metal positions of the cluster based on their presence in the crystallization solution at high concentration and the fact that substitution of the metal ions in the active site of prolidase from *P. furiosus* has previously been detected upon crystallization (Maher *et al.*, 2004). In the structure of *Ts*prol, the relative arrangement of the metal ions of the cluster and the residues coordinating them is almost the same as in the structures of other prolidases (PDB entries 1wy2 and 1pv9; RIKEN Structural Genomics/Proteomics Initiative, unpublished work; Maher *et al.*, 2004). The distance between the metal ions in these structures is about 3.1 Å.

The hyperthermostable prolidase from the archaeon *P. horikoshii* OT3 is the closest enzyme to *Ts*prol (52% identity) with known structure (PDB entry 2how; Jeyakanthan *et al.*, 2009). The subunit

structures of *Ts*prol and 2how can be superimposed with an r.m.s. deviation of 1.1 Å using equivalent C^α atoms of 355 residues. If their C-terminal domains are superimposed in a similar manner (r.m.s.d. of 0.7 Å), the deviation of the majority of equivalent C^α atoms in the N-terminal domains is 4–7 Å (Fig. 4). According to *PISA* calculations (Krissinel & Henrick, 2007), the dimer interface area of *Ts*prol in the crystal is 989 Å² (the solvation free-energy gain is $-23.9 \text{ kJ mol}^{-1}$), which is similar to the value for 2how (1170 Å²; solvation free-energy gain of $-28.5 \text{ kJ mol}^{-1}$). The dimers of *Ts*prol, 2how and other prolidases (PDB entries 1wy2 and 1pv9) are organized in a similar manner: the subunits are related to each other by the twofold axis of the dimer (crystallographic in the case of *Ts*prol), the N-terminal domain interacts with the N-terminal domain and the C-terminal domain interacts with the C-terminal domain (Fig. 2). However, the dimer structures 2how, 1wy2 and 1pv9 reveal additional relatively weak interactions between the N- and C-terminal domains which are absent from the structure of *Ts*prol. Hence, the dimer of *Ts*prol can be considered to be the most open dimer when compared with previously structurally studied prolidases.

In the crystal, each *Ts*prol subunit has two identical areas of intermolecular contacts (the average interface area is 743.7 Å²; the solvation free-energy gain is 14.2 kJ mol^{-1}), which are formed by the surface regions of the N-terminal domain and the C-terminal domain of another subunit (Fig. 5). Coordination bonds formed by cadmium ions were detected in these regions. One of the cadmium ions (CdI) is coordinated by the side chains of the His-tag residue His2 and Asp147 from the same subunit and takes part in structuring the N-terminal region. Another ion (CdII) is coordinated by the side chains of His3 and Glu259 from different subunits. The third cadmium ion (CdIII) is bound to the side chain of Glu366; its coordination sphere interacts with the main chain of the His-tag residue Gly4 from another subunit. The fourth cadmium ion (CdIV) is coordinated by the side chains of Glu99 and Glu157 from different subunits. Therefore, during crystallization the dimers of *Ts*prol are linked to each other owing to the formation of coordination bonds to cadmium ions from the crystallization solution; it is apparently these strong interactions that lead both to the distortion of the subunit of *Ts*prol and to the 'opening' of its dimer.

The crystal of *Ts*prol is characterized by a rather loose packing of protein molecules (the Matthews coefficient is $4.14 \text{ Å}^3 \text{ Da}^{-1}$) and a low resistance to mechanical influence and temperature increase, which can also be attributed to the anomalously strong

Figure 3

The active site of *Ts*prol. Cadmium ions are shown as brown spheres. Coordination bonds are indicated by dashed lines; the bond distances are labelled in Å.

Figure 4

Superposition of *Ts*prol and 2how subunits using equivalent C^α atoms of the C-terminal domains. *Ts*prol is coloured from red at the N-terminus to purple at the C-terminus. 2how is coloured grey.

Figure 5

Intermolecular contacts in the crystal of *TspI*. One *TspI* subunit is represented as a ribbon model and the other two are shown as worm models. The His tag is coloured cyan; the N-terminal domain, the linker and the C-terminal domain are coloured green, yellow and pink, respectively. The intermolecular cadmium ions are shown as brown spheres. The residues coordinating them are represented as ball-and-stick models. The active-site cadmium ions are shown as cyan spheres.

intermolecular interactions. The contribution of the cadmium ions and the histidine residues of the His tag to the formation of the *TspI* crystal is likely to be pivotal.

This work was supported by the Ministry of Education and Science of the Russian Federation (State Contract 16.512.11.2175).

References

- Du, X., Tove, S., Kast-Hutcherson, K. & Grunden, A. M. (2005). *FEBS Lett.* **579**, 6140–6146.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Jeyakanthan, J., Takada, K., Sawano, M., Ogasahara, K., Mizutani, H., Kunishima, N., Yokoyama, S. & Yutani, K. (2009). *J. Biophys.* **2009**, 434038.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
- Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* **D64**, 125–132.
- Lowther, W. T. & Matthews, B. W. (2002). *Chem. Rev.* **102**, 4581–4608.
- Maher, M. J., Ghosh, M., Grunden, A. M., Menon, A. L., Adams, M. W., Freeman, H. C. & Guss, J. M. (2004). *Biochemistry*, **43**, 2771–2783.
- Mardanov, A. V., Ravin, N. V., Svetlitchnyi, V. A., Beletsky, A. V., Miroshnichenko, M. L., Bonch-Osmolovskaya, E. A. & Skryabin, K. G. (2009). *Appl. Environ. Microbiol.* **75**, 4580–4588.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.