Crystal structure of the *Thermoplasma acidophilum* protein Ta1207

Ganesh Ramnath Pathare,^{a,b,*} István Nagy,^a Ágnes Hubert,^{a,c} Dennis R. Thomas^{a,d} and Andreas Bracher^{e,*}

^aDepartment of Molecular Structural Biology, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany, ^bDepartment of Quantitative Biology, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland, ^cDepartment of Medical Biochemistry, Semmelweis University, 37-47 Tuzolto Street, Budapest 1094, Hungary, ^dCenter for Integrative Proteomics Research, Rutgers University, 174 Frelinghuysen Road, Piscataway, NJ 08854-8076, USA, and ^eDepartment of Cellular Biochemistry, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany. *Correspondence e-mail: ganesh.pathare@fmi.ch, bracher@biochem.mpg.de

Received 16 February 2017

Accepted 12 May 2017

Edited by M. S. Weiss, Helmholtz-Zentrum Berlin für Materialien und Energie, Germany

Keywords: pentamers; archaeon; thermo-stability; nanotechnology; *Thermoplasma acidophilum*; Ta1207.

PDB reference: Ta1207, 5m86

Supporting information: this article has supporting information at journals.iucr.org/f

The crystal structure of the Ta1207 protein from *Thermoplasma acidophilum* is reported. Ta1207 was identified in a screen for high-molecular-weight protein complexes in *T. acidophilum*. In solution, Ta1207 forms homopentamers of 188 kDa. The crystal structure of recombinant Ta1207 solved by Se-MAD at 2.4 Å resolution revealed a complex with fivefold symmetry. In the crystal lattice, calcium ions induce the formation of a nanocage from two pentamers. The Ta1207 protomers comprise two domains with the same novel α/β topology. A deep pocket with a binding site for a negatively charged group suggests that Ta1207 functions as an intracellular receptor for an unknown ligand. Homologues of Ta1207 occur only in Thermoplasmatales and its function might be related to the extreme lifestyle of these archaea. The thermostable Ta1207 complex might provide a useful fivefold-symmetric scaffold for future nanotechnological applications.

1. Introduction

Cryo-electron tomography of single cells allows the structural study of the molecular landscape of macromolecular complexes in their natural environment under physiological conditions. The technique is gaining increasing attention with recent advances in direct electron-detector, focused-ion-beam (FIB) milling and phase-plate technologies (Mahamid *et al.*, 2016; Schaffer *et al.*, 2016). However, many structures, even if they are detected, cannot be identified owing to the lack of high-resolution structures of these complexes. This problem occurs more often in less well characterized organisms. Most microbial genomes present a substantial fraction of unique proteins with unknown functions, protein associations and folds. For this reason, a template library of macromolecular complexes with high-resolution structures would be helpful. The structural proteomics initiatives have already made it possible to obtain one or more representative three-dimensional structures for almost all of the protein-domain families present in nature (Yee *et al.*, 2003). The structural information obtained from these approaches provides a valuable resource for the structural and functional characterization of proteins.

T. acidophilum is a thermoacidophilic archaeon which thrives in hot and acidic environments (Darland *et al.*, 1970; Segerer & Stetter, 1992). It grows optimally at 59°C and pH 2.

The cells are pleiomorphic because of their lack of a cell wall, and spread out on EM grids. This is highly advantageous for cryo-electron tomography, which requires sample thicknesses of less than 200 nm. With 1.5 Mbases, the genome of *T. acidophilum* was among the smallest at the time of sequencing (Ruepp *et al.*, 2000). Around 35 macromolecular complexes with sizes of greater than 300 kDa were identified in the cytosol of *T. acidophilum* (Sun *et al.*, 2009). These large complexes are prime candidates for single-particle EM analysis in order to create template libraries for a comprehensive cellular atlas of cellular protein complexes (Nickell *et al.*, 2006). In the past, the structures of ancestral proteins and protein complexes from thermophilic organisms have provided a basis for the molecular understanding of many eukaryotic protein homologues for which structures were elusive (Jap *et al.*, 1993; Hegerl *et al.*, 1991). However, there are many proteins and protein complexes in thermophiles which have yet to be explored and characterized. Their structure might provide valuable insights into the physiology and lifestyle of these fascinating organisms.

Our efforts to characterize large protein complexes of *T. acidophilum* for cryo-electron tomography of whole cells was continued by purifying the endogenous Ta1207 complex. Ta1207 was detected as a complex of approximately 200 kDa in a proteomics experiment that aimed to identify protein complexes in *T. acidophilum* (Sun *et al.*, 2007). Sequence homologues to Ta1207 appear to be present only in Thermoplasmatales. The putative orthologue in *T. volcanium* shares 70.8% sequence identity with Ta1207. In this study, we present data demonstrating that Ta1207 forms homopentamers in solution. We furthermore report the crystal structure of Ta1207, revealing a novel α/β -fold architecture. Based on this structure and its sequence conservation in Thermoplasmatales, we propose that Ta1207 may function as a receptor for an unknown ligand with a negatively charged motif.

2. Materials and methods

2.1. Macromolecule production

2.1.1. Preparation of *Thermoplasma* extract and purification of endogenous Ta1207. *T. acidophilum* was cultured aerobically using standard protocols (Robb & Place, 1995) with minor modifications as described previously (Sun *et al.*, 2009). Cells from a 500 ml culture were harvested by centrifugation at 4000g and 4°C for 10 min after 48 h of incubation. The cells were resuspended after washing them in 5 ml Milli-Q water containing protease inhibitors, 5 mM MgCl₂, 1 mM ATP and 1 mM dithiothreitol. The pH of the cell suspension was titrated with 1 M unbuffered Tris base to pH 8, whereupon immediate cell lysis occurred. The released nucleic acids were degraded by adding 0.5 mg Benzonase and the lysate was centrifuged at 34 800g and 4°C for 30 min to remove cell debris.

To purify endogenous Ta1207, a three-step purification method was used. Firstly, the Ta0328 (FixA)-specific mouse

Table 1

Macromolecule-production information.

Source organism	<i>Thermoplasma acidophilum</i>
DNA source	<i>T. acidophilum</i> genomic DNA
Forward primer	<u>CGAATTC</u> CCATGGGCATAATCATGGTTCCTCAGG- ATATTG
Reverse primer	CGAATTCCTCGAGCATGGGTGTTATCTCATCGAT- TTTCATG
Cloning vector	pET-28a
Expression vector	pET-28a
Expression host	<i>E. coli</i> BL21 (DE3)
Native expression source	<i>T. acidophilum</i>
Complete amino-acid sequence of the construct produced	MGIIIMVLQDIDGIFDSQAILAGRFHNDLTIVINEK- YDLFYLMPTINEKKIVSFYIFLQDDQIPERH- REEIESVLNKFNPVIKNGIWKIYLDTESFKLS- EPFSTFFGIDSIVFDMGSMKGGEMLLPVRFIS- KDKDALVNSIIDSAGYGENIYLRIGQNKGF- YSFIAIKLLDQVYKLTLSIDNPHVMHGIFAET- KKNIAWRRESKAPHKDNTEYIYALDDTHTIP- DILIDTAYTGEKGTVYIGKHSNYDIYRAFFGD- ALTNHMSVMISENVYLLRRWSKYEDGKFLY- FYTTVDLRLIPAILDSTRKNFPKVMKIDEI- TPMLEHHHHH

antibody was employed for affinity chromatography. This antibody was selected from a phage-display library raised against high-molecular-weight protein complexes of *T. acidophilum* (Hubert *et al.*, 2014). In the second step, the proteins eluted from the first column were subjected to size-exclusion chromatography (SEC), and the protein-containing fractions were analyzed by one-dimensional polyacrylamide gel electrophoresis (1D-PAGE) and LC-MS/MS. The fractions of the FixABCX complex contained two dominant proteins, Ta1207 and Ta0784, with a size of approximately 200 kDa. These proteins appeared in the molecular-weight range of around 34 kDa on SDS-PAGE. Ta1207 was further separated from the other proteins using gel-blotting technology (Knispel *et al.*, 2012) and was analyzed by cryo-EM.

2.1.2. Cloning of the Ta1207 gene. The Ta1207 ORF was amplified by PCR from *T. acidophilum* genomic DNA using *Pfu* polymerase (Thermo Scientific) and the primer pair specified in Table 1. The PCR product was digested with NcoI and XhoI restriction endonucleases (the corresponding DNA sequences are underlined in Table 1) and cloned between the NcoI and XhoI sites of the pET-28a plasmid using standard cloning methods. Using this method, we fused a His₆ affinity tag encoded by the plasmid to the C-terminus of the protein. Thus, the additional sequence LEHHHHHH was introduced into the expressed protein as a cloning artifact. The correctness of the plasmid insert was confirmed by DNA sequencing.

2.1.3. Expression and purification. The C-terminally His₆-tagged Ta1207 was overexpressed in *Escherichia coli* BL21 (DE3) cells overnight at 37°C in 5 l LB medium. The cells were sedimented at 4000g and washed with sterile water. The cell pellet (50 g) was washed and resuspended in lysis buffer [50 mM Tris pH 8.0, 10 mM imidazole, 150 mM NaCl and one Complete protease-inhibitor cocktail tablet (Roche)]. Because the obtained cell lysate was very viscous, Benzonase (Novagen) was added to degrade nucleic acids and the sample was ultrasonicated on ice. The cell debris was sedimented by centrifugation at 21 000g, and the supernatant was applied

Table 2
Crystallization.

Method	Sitting-drop vapour diffusion
Plate type	96-well 2-drop MRC Crystallization Plates (Swissci) UVB
Temperature (K)	291
Protein concentration (mg ml ⁻¹)	7
Buffer composition of protein solution	10 mM Tris pH 7.5, 100 mM NaCl, 50 mM KCl, 2 mM DTT
Composition of reservoir solution	50 mM MES pH 6.0, 200 mM calcium acetate, 20%(w/v) PEG 3350
Volume and ratio of drop	2 µl (1:1)
Volume of reservoir (µl)	50

Table 3
Data collection and processing.

Diffraction source	ESRF beamline ID29
Wavelength (Å)	0.9767
Temperature (K)	100
Detector	PILATUS 6M
Crystal-to-detector distance (mm)	566.96
Rotation range per image (°)	0.05
Total rotation range (°)	180
Exposure time per image (s)	0.037
Space group	C2
<i>a</i> , <i>b</i> , <i>c</i> (Å)	228.64, 112.82, 221.33
α , β , γ (°)	90, 118.70, 90
Mosaicity (°)	0.083
Resolution range (Å)	49.17–2.39 (2.43–2.39)
Total No. of reflections	663454 (27342)
No. of unique reflections	191535 (8492)
Completeness (%)	98.7 (88.4)
Multiplicity	3.5 (3.2)
$\langle I/\sigma(I) \rangle$	13.6 (2.1)
CC _{1/2}	0.997 (0.621)
<i>R</i> _{r.i.m.}	0.090 (0.839)
<i>R</i> _{p.i.m.}	0.048 (0.452)
Overall <i>B</i> factor from Wilson plot (Å ²)	46.17

onto Ni-NTA beads (Qiagen) and processed according to the supplier's recommendations. The fractions containing Ta1207 were pooled, concentrated and further purified on a Superdex 200 (GE Healthcare) column equilibrated with 10 mM Tris pH 7.5, 100 mM NaCl, 50 mM KCl, 2 mM DTT.

2.2. Crystallization and cryoprotection

Crystals of Ta1207 were grown as specified in Table 2. For mounting, crystals were transferred stepwise into reservoir solution containing an additional 15%(v/v) glycerol and were subsequently cryocooled by immersion in liquid nitrogen.

2.3. Data collection and processing

Diffraction data were collected at the European Synchrotron Radiation Facility (ESRF), Grenoble, France (Table 3). Data were processed with *XDS* (Kabsch, 2010) and converted to CCP4 format using *POINTLESS* (Evans, 2006), *AIMLESS* (Evans & Murshudov, 2013) and *CTRUNCATE* (French & Wilson, 1978) as implemented in the *CCP4* interface (Winn *et al.*, 2011).

2.4. Structure solution and refinement

The crystal structure was solved by three-wavelength Se-MAD using *SHELXC/D/E* as implemented in *HKL2MAP*

Table 4
Structure refinement.

Resolution range (Å)	30–2.4 (2.43–2.40)
Completeness (%)	99.0
σ Cutoff	$F > 1.920\sigma(F)$
No. of reflections, working set	181077 (5885)
No. of reflections, test set	9569 (333)
Final <i>R</i> _{cryst}	0.177 (0.282)
Final <i>R</i> _{free}	0.203 (0.301)
No. of non-H atoms	
Total	27183
Protein	26350
Ligand	80
Solvent	753
Total No. of atoms	53263
R.m.s. deviations	
Bonds (Å)	0.006
Angles (°)	0.9
Average <i>B</i> factors, non-H atoms (Å ²)	
Protein	56.8
Ligand	51.1
Solvent	53.9
R.m.s. ΔB (Å ²)	
Main chains	1.8
Side chains	2.9
Ramachandran plot	
Most favoured (%)	97.3
Allowed (%)	2.6

(Pape & Schneider, 2004; Schneider & Sheldrick, 2002). *SHARP* was used to refine the 91 Se sites found and to calculate experimental phases (de La Fortelle & Bricogne, 1997). Density modification was performed with *RESOLVE*, resulting in a partial model and a readily interpretable electron-density map (Terwilliger, 2000). The model of one subunit was completed manually using *Coot* and placed onto other partially built subunits, resulting in a complete model. Subsequently, the model was improved by iterative cycles of manual model building in *Coot* (Emsley & Cowtan, 2004) and refinement using *REFMAC5* with local noncrystallographic symmetry (NCS) restraints and a translation–libration–screw (TLS) description of *B* factors (Murshudov *et al.*, 2011). In the final stages of refinement, *phenix.refine* was used (Afonine *et al.*, 2012; Table 4). The stereochemical quality of the model was assessed with *MolProbity* (Chen *et al.*, 2010). The model is complete apart from residues 1–4 and 328–333 of each chain, for which no electron density was detected. Several residues with disordered side chains were truncated after their C^β atom. The model has good stereochemistry, with 97% of the peptide bonds in the favoured region of the Ramachandran plot and one outlier. Each subunit contains one ordered glycerol molecule, a putative chloride ion trapped inside a cavity and a calcium ion interconnecting two subunits. Coordinates and structure factors have been deposited in the wwPDB with accession code 5m86.

3. Results and discussion

3.1. Ta1207

The open reading frame Ta1207 of *T. acidophilum* encodes a hypothetical protein of 324 amino acids with a calculated molecular weight of 37.7 kDa (Ruepp *et al.*, 2000). A close

sequence homologue is known from the related organism *T. volcanium* (TVN1256). Two more distantly related sequences were discovered in a mine-drainage metagenome containing two unknown Thermoplasmatales species (Yelton *et al.*, 2013; Supplementary Fig. S1). In the genome of *T. acidophilum*, Ta1207 is located between a nitrilase-related gene (Ta1208) and another gene coding for a hypothetical protein (Ta1206). Quantitative proteomic analysis of *T. acidophilum* grown aerobically and anaerobically suggests that the Ta1207 protein is constitutively expressed irrespective of the O₂ content of the culture medium (Sun *et al.*, 2010).

We identified Ta1207 as an interactor of the FixA mouse antibody selected from a phage-display library raised against high-molecular-weight protein complexes of *T. acidophilum* (Hubert *et al.*, 2014). This antibody was also selective for the formate-dehydrogenase-related protein Ta0425 and its putative interaction partner Ta0424. The interaction between the proteins was weak, if there was any, as the Ta0425–Ta0424 complex eluted at 500 kDa, while Ta1207 eluted at 200 kDa, in SEC experiments. Ta1207 was further separated from contaminating proteins by native PAGE, transferred to EM grids using gel-blotting technology (Knispel *et al.*, 2012) and analyzed by cryo-EM. Numerous, fairly uniform particles of ~110 Å diameter were readily detectable in the sample (Fig. 1*a*). Two-dimensional classification suggested a particle with fivefold symmetry (Fig. 1*a*). Three-dimensional single-particle reconstruction revealed a C₅-symmetric homopentamer structure for Ta1207 (Fig. 1*a*).

3.2. Characterization of recombinant Ta1207

For further structural characterization of the Ta1207 complex, the gene was cloned and expressed in *E. coli*. Recombinant Ta1207 was expressed in large amounts as a soluble protein, and it could be purified in high yield almost to homogeneity using a two-step chromatographic procedure. Ta1207 appeared as a sharp, symmetrical peak in size-

exclusion chromatography (SEC), suggesting uniform particle size. Cryo-EM analysis of these fractions revealed homopentameric particles, similar to the endogenous complex from *T. acidophilum* (Figs. 1*a* and 1*b*). Therefore, the recombinant Ta1207 was deemed to be suitable for structural analysis by X-ray crystallography.

The molecular weight determined by multi-angle laser light scattering (MALS) was $187\,676 \pm 310$ Da, which is close to the expected molecular weight of a Ta1207 pentamer (189 000 Da) (Supplementary Fig. S2*a*). Guinier plot analysis of small-angle X-ray scattering (SAXS) intensity curves of Ta1207 at four different concentrations indicated an average radius of gyration of 4.64 ± 0.18 nm (Supplementary Fig. S2*b*). Ta1207 readily crystallized, forming monoclinic and tetragonal crystals in the presence of Ca²⁺. The latter crystals only diffracted to 3.4 Å resolution and were not investigated further.

3.3. Crystal structure of the Ta1207 homo-oligomer

We solved the crystal structure of selenomethionine-labelled Ta1207 protein at a resolution of 2.4 Å by Se-MAD using direct methods (Tables 3 and 4). The experimental electron-density map was of excellent quality and was readily interpretable (Fig. 2*a*). The asymmetric unit contains a Ta1207 homodecamer with approximate D₅ symmetry (Figs. 2*b* and 2*c*). The particle dimensions are roughly 115 × 115 × 125 Å. The subunits form a hollow sphere with large pores on the symmetry axes. The contacts between the pentamers are mediated by pairs of Ca²⁺ ions from the precipitant (Fig. 2*b*). The association of pentamers *via* bound Ca²⁺ ions presumably represents a crystallization artifact, in line with the MALS and SAXS data (Supplementary Figs. S2*a* and S2*b*). The SAXS scattering curve fits better to the Ta1207 pentamer than to the decamer (Supplementary Fig. S2*b*). The theoretical R_g for the Ta1207 pentamer calculated by CRY SOL was 4.3 nm. Of note, the Ta1207 pentamer also fits well to the cryo-EM envelope (Supplementary Fig. S3).

Figure 1 Electron-microscopy single-particle analysis. (*a*, *b*) Cryo-EM micrographs of Ta1207 purified from *T. acidophilum* cell extract (*a*) and recombinantly expressed Ta1207 purified from *E. coli* lysate (*b*). The insert shows averages from two-dimensional particle classification on the same scale; the white scale bar indicates a length of 100 Å. On the right, the three-dimensional reconstructions of the particle at 14 Å resolution are shown, revealing fivefold rotational symmetry. Details of cryo-EM data collection, image processing and three-dimensional refinement can be found in the Supporting Information.

The ten Ta1207 monomers in the crystal lattice are virtually identical, with residual mean-square deviations (r.m.s.d.s on

C^α atoms) of 0.1–1.3 Å. The Ta1207 monomer consists of two globular domains (domain 1, residues 13–155; domain 2,

Figure 2

Crystal structure and domain architecture of Ta1207. (a) Experimental electron-density map at 2.4 Å resolution. Electron density at 2.0σ is shown in mesh representation. The backbone trace of the final model in rainbow colouring is superposed. (b) View of the asymmetric unit in the crystal lattice, where the two pentamers of Ta1207 form a homodecamer. The Ta1207 homodecamer has roughly D_5 symmetry. The protein chains are shown in ribbon representation. One of the five subunits is coloured according to its domain structure, with the N- and C-terminal domains indicated in blue and gold, respectively. The yellow spheres at the interface of the pentameric rings are calcium ions interconnecting two subunits. The green spheres designate bound chloride ions. (c) View of the Ta1207 pentamer along the fivefold axis. (d) Ribbon representation of the subunit structure. Front and back views are shown. The representation and colouring are the same as in (b). N- and C-termini are indicated. (e) Cartoon representation of the secondary-structure topology of Ta1207. N- and C-termini are indicated. (f) Superposition of domains 1 and 2.

residues 173–324) with the same architecture (Figs. 2*d*, 2*e* and 2*f*). Each domain contains a twisted eight-stranded β -sheet with topology $\beta I-\beta A-\beta H-\beta G-\beta B-\beta C-\beta E-\beta D$ flanked by α -helices (Fig. 2*e*). $\beta B-\beta C$, $\beta D-\beta E$ and $\beta G-\beta H$ form β -hairpin structures interspersed with the flanking helices $\alpha 1$, $\alpha 2$, $\alpha 3$ and $\alpha 4$ (Fig. 2*e*). The interspersed helices are all on the same face of the β -sheet. In domain 2, the second β -hairpin contains an insertion that forms β -contacts to βB in domain 1. Helix 5, which is located in the interdomain linker, contacts both domains (Figs. 2*d* and 2*e*). The two domains, which share only 4.7% sequence identity, differ by an r.m.s.d. of 2.1 Å across 74 matching C^α positions (Fig. 2*f*). The fold of Ta1207 appears to be previously unknown. A DALI search for related protein

structures only yielded hits for the fragment containing the final four β -strands of Ta1207 (Holm & Rosenström, 2010).

The contacts between the subunits in the pentamer are formed by $\alpha 4$ and βI in domain 1 and $\alpha 8$ and βR in domain 2. The average buried surface area at these interfaces is 728.6 Å² as calculated with PISA (Krissinel & Henrick, 2005). The interface comprises the exposed conserved residues Asp133, Val136, Ile140, Gly144, Tyr153, Gly155, Pro301, Ile318 and Ile321. In contrast, the buried surface area between subunits at the equator of the homodecamer is only 159.8 Å². The sequence conservation of the residues contacting the Ca²⁺ ions, Asp219, Asp220, Thr221 and Asn247, is poor, in line with our interpretation of this as a crystallization artifact. Thus, we

Figure 3

The putative ligand-binding pocket of Ta1207. (a) Overview of the Ta1207 pentamer. Bound chloride ions are shown as green spheres. The boxed area represents one putative binding site. (b) Detailed view of the putative ligand-binding pocket of Ta1207. The molecular surface of Ta1207 is represented as a semi-transparent envelope. In the top right corner, the outer surface and the entrance to the pocket are visible. The dark grey tube in the centre represents the inner part of the binding pocket. In the crystal structure, a presumed chloride ion is clamped by two arginines at the bottom of a narrow channel, which winds down from the perimeter of the pentamer. Arginine–chloride interactions are represented as dashed lines. The experimental density for chloride at 2.0 σ is shown in mesh representation. (c) Excerpt from a sequence alignment of Ta1207 homologues (Supplementary Fig. S1). The conserved arginines clamping the chloride ions in the Ta1207 structure are indicated by green arrows.

predict that all known Ta1207 homologues form homopentamers in solution.

Close inspection of the electron density revealed a strong peak between the side chains of Arg201 and Arg277 in every subunit (Figs. 3*a* and 3*b*). These residues are located in strands β K or β P of domain 2. This peak is likely to correspond to a chloride ion clamped between the guanidine groups of the arginine residues, as judged by the difference in *B* factor from the contacting arginine-residue atoms and the chemical environment. No corresponding structure was found in domain 1, arguing against a generic function of the bound ion in stabilizing the domain fold. The chloride ion is located at the bottom of a deep pocket which can be accessed by solvent from the perimeter of the pentamer (Fig. 3*b*). The walls lining the binding pocket, comprising residues Val110, Arg201, Glu202, Ser203, Pro206, Arg276 and Arg277, are conspicuously conserved among the sequences of Ta1207 homologues (Fig. 3*c*). Thus, we propose that the Ta1207 homologues in Thermoplasmatales function as receptors for an unknown ligand with a negative charge. The absence of Ta1207 homologues outside of Thermoplasmatales suggests a relationship of this interaction to the extreme thermophilic and acidophilic lifestyle of these organisms.

3.4. Potential application of Ta1207 as a pentameric scaffold in nanotechnology

Pentameric rotational symmetry is frequently found in the capsid structures of viruses and carboxysomes. In comparison to dimers, trimers, tetramers and hexamers, few soluble homomeric protein complexes with fivefold symmetry are found in nature, as judged by the numbers of data sets in the PDB with the respective numbers of chains in the biological assembly. Nanotechnology employs such molecular symmetries as templates for the self-assembly of functional structures. Thus, the Ta1207 pentamer may serve as an interesting thermostable pentameric scaffold in nanotechnological applications, such as the controlled immobilization of biomolecules on solid substrates. The N- and C-termini are facing the pore in the Ta1207 pentamer structure. Moreover, the Ta1207 pentamers form a nanocage with pseudo-*D*₅ symmetry in the presence of calcium, which connects the two pentameric rings. The inner diameter of this nanocage at the equator is ~ 53 Å and the height of the cavity is ~ 38 Å, resulting in a volume of 56×10^3 Å³, which is sufficient to enclose a 46 kDa protein. The cage is thus similar in size to the *cis*-ring of the GroEL–GroES complex, and may be used as a promising delivery vehicle for macromolecules.

4. Related literature

The following references are cited in the Supporting Information for this article: Gouet *et al.* (1999), Konarev *et al.* (2003) and Vachette *et al.* (2003).

Acknowledgements

We thank the JSBG group at the ESRF Grenoble and the staff at the MPIB Crystallization Facility and the MPIB Core Facility for their excellent support.

References

- Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). *Acta Cryst.* **D68**, 352–367.
- Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- Darland, G., Brock, T. D., Samsonoff, W. & Conti, S. F. (1970). *Science*, **170**, 1416–1418.
- Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.
- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Evans, P. R. & Murshudov, G. N. (2013). *Acta Cryst.* **D69**, 1204–1214.
- French, S. & Wilson, K. (1978). *Acta Cryst.* **A34**, 517–525.
- Gouet, P., Courcelle, E., Stuart, D. I. & Métoz, F. (1999). *Bioinformatics*, **15**, 305–308.
- Hegerl, R., Pfeifer, G., Pühler, G., Dahlmann, B. & Baumeister, W. (1991). *FEBS Lett.* **283**, 117–121.
- Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
- Hubert, Á., Mitani, Y., Tamura, T., Boicu, M. & Nagy, I. (2014). *J. Microbiol. Methods*, **98**, 15–22.
- Jap, B., Pühler, G., Lücke, H., Typke, D., Löwe, J., Stock, D., Huber, R. & Baumeister, W. (1993). *J. Mol. Biol.* **234**, 881–884.
- Kabsch, W. (2010). *Acta Cryst.* **D66**, 125–132.
- Knispel, R. W., Kofler, C., Boicu, M., Baumeister, W. & Nickell, S. (2012). *Nature Methods*, **9**, 182–184.
- Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.
- Krissinel, E. & Henrick, K. (2005). *CompLife 2005*, edited by edited by M. R. Berthold, R. Glen, K. Diederichs, O. Kohlbacher & I. Fischer, pp. 163–174. Berlin, Heidelberg: Springer-Verlag.
- La Fortelle, E. de & Bricogne, G. (1997). *Methods Enzymol.* **276**, 472–494.
- Mahamid, J., Pfeffer, S., Schaffer, M., Villa, E., Danev, R., Kuhn Cuellar, L., Forster, F., Hyman, A. A., Plitzko, J. M. & Baumeister, W. (2016). *Science*, **351**, 969–972.
- 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.
- Nickell, S., Kofler, C., Leis, A. P. & Baumeister, W. (2006). *Nature Rev. Mol. Cell Biol.* **7**, 225–230.
- Pape, T. & Schneider, T. R. (2004). *J. Appl. Cryst.* **37**, 843–844.
- Robb, F. T. & Place, A. R. (1995). *Archaea: A Laboratory Manual*, pp. 60–61. New York: Cold Spring Harbor Laboratory Press.
- Ruepp, A., Graml, W., Santos-Martinez, M. L., Koretke, K. K., Volker, C., Mewes, H. W., Frishman, D., Stocker, S., Lupas, A. N. & Baumeister, W. (2000). *Nature (London)*, **407**, 508–513.
- Schaffer, M., Mahamid, J., Engel, B. D., Laugs, T., Baumeister, W. & Plitzko, J. M. (2016). *J. Struct. Biol.* **197**, 73–82.
- Schneider, T. R. & Sheldrick, G. M. (2002). *Acta Cryst.* **D58**, 1772–1779.
- Segerer, A. & Stetter, K. O. (1992). *The Prokaryotes*, edited by A. Balows, H. G. Trüper, M. Dworkin, W. Harder & K.-H. Schleifer, pp. 712–718. New York: Springer.
- Sun, N., Beck, F., Knispel, R. W., Siedler, F., Scheffer, B., Nickell, S., Baumeister, W. & Nagy, I. (2007). *Mol. Cell. Proteomics*, **6**, 492–502.
- Sun, N., Pan, C., Nickell, S., Mann, M., Baumeister, W. & Nagy, I. (2010). *J. Proteome Res.* **9**, 4839–4850.
- Sun, N., Tamura, N., Tamura, T., Knispel, R. W., Hrabe, T., Kofler, C., Nickell, S. & Nagy, I. (2009). *Proteomics*, **9**, 3783–3786.

- Terwilliger, T. C. (2000). *Acta Cryst.* **D56**, 965–972.
- Vachette, P., Koch, M. H. & Svergun, D. I. (2003). *Methods Enzymol.* **374**, 584–615.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 355–367.
- Yee, A., Pardee, K., Christendat, D., Savchenko, A., Edwards, A. M. & Arrowsmith, C. H. (2003). *Acc. Chem. Res.* **36**, 183–189.
- Yelton, A. P., Comolli, L. R., Justice, N. B., Castelle, C., Deneff, V. J., Thomas, B. C. & Banfield, J. F. (2013). *BMC Genomics*, **14**, 485.