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

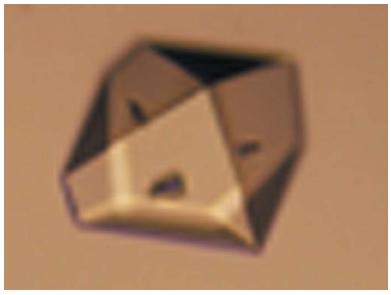
Expression, purification, crystallization and preliminary crystallographic analysis of a thermostable DNA ligase from the archaeon *Thermococcus sibiricus*

T. E. Petrova,[a,b]* E. Y. Bezsudnova,[a] B. D. Dorokhov,[c] E. S. Slutskaya,[a] K. M. Polyakov,[d,e] P. V. Dorovatovskiy,[d,f] N. V. Ravin,[c] K. G. Skryabin,[c,d] M. V. Kovalchuk[d,f] and V. O. Popov[a,d]

[a]Bach Institute of Biochemistry, RAS, Leninsky pr. 33, Moscow 119071, Russian Federation, [b]Institute of Mathematical Problems of Biology, RAS, Institutskaja str. 4, Pushchino 142290, Russian Federation, [c]Centre 'Bioengineering', RAS, Pr. 60-let Oktyabrya, Bldg 7-1, Moscow 117312, Russian Federation, [d]RSC 'Kurchatov Institute', Acad. Kurchatov sq. 1, Moscow 123182, Russian Federation, [e]Engelgardt Institute of Molecular Biology, RAS, Vavilov str. 32, Moscow 119991, Russian Federation, and [f]Shubnikov Institute of Crystallography, RAS, Leninsky pr. 59, Moscow 119333, Russian Federation

Correspondence e-mail: petrova@impb.psn.ru

Received 28 October 2011
Accepted 26 November 2011



DNA ligases join single-strand breaks in double-stranded DNA by catalyzing the formation of a phosphodiester bond between adjacent 5′-phosphate and 3′-hydroxyl termini. Their function is essential to maintain the integrity of the genome in DNA replication, recombination and repair. A recombinant ATP-dependent DNA ligase from the hyperthermophilic anaerobic archaeon *Thermococcus sibiricus* was expressed in *Escherichia coli* and purified. Crystals were grown by vapour diffusion using the hanging-drop method with 17%(*w*/*v*) PEG 4000 and 8.5%(*v*/*v*) 2-propanol as precipitants. A diffraction experiment was performed with a single crystal, which diffracted X-rays to 3.0 Å resolution. The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 58.590$, $b = 87.540$, $c = 126.300$ Å.

1. Introduction

DNA ligases play an essential role in joining the breaks in double-stranded DNA by catalyzing the formation of a phosphodiester bond between adjacent 5′-phosphate and 3′-hydroxyl termini. Their function is essential in DNA replication, recombination and repair (Lehman, 1974). DNA ligases are universally present in all living organisms and may be classified into two families according to their high-energy cofactor requirements. ATP-dependent ligases (EC 6.5.1.1) are found in eukaryotes, archaea, viruses and some bacteria. The second group consists of NAD^+-dependent DNA ligases (EC 6.5.1.2), which are present in bacteria and in some eukaryotic viruses. Apart from the difference in their cofactor requirement, both ATP-dependent and NAD^+-dependent DNA ligases catalyze reactions *via* three common sequential nucleotidyl-transfer reactions: covalent addition of the AMP group of ATP or NAD^+ (step 1), AMP transfer from the enzyme to the 5′-phosphoryl group of the nick on the DNA (step 2) and phosphodiester-bond formation with AMP release (step 3) (Pascal, 2008).

Relatively little is known about DNA ligases from the Archaea, the third domain of life. Thermostable DNA ligases from several hyperthermophilic archaea have been isolated and functionally characterized (*e.g.* Sriskanda *et al.*, 2000; Lai *et al.*, 2002; Keppetipola & Shuman, 2005; Smagin *et al.*, 2008). Only a few crystal structures of archaeal DNA ligases have been reported: those from *Pyrococcus furiosus* (Nishida *et al.*, 2006), *Archaeoglobus fulgidus* (Kim *et al.*, 2009), *Thermococcus* sp. 1519 (Bezsudnova *et al.*, 2009) from the Euryarchaeota and from *Sulfolobus solfataricus* (Pascal *et al.*, 2006) and *Sulfophobococcus zilligii* (Supangat *et al.*, 2010) from the Crenarchaeota. These data confirmed that archaeal DNA ligases consist of three domains: a DNA-binding domain, an adenylation domain and an OB-fold domain. The DNA ligases from *P. furiosus* and *A. fulgidus* adopted a closed conformation (Nishida *et al.*, 2006; Kim *et al.*, 2009), whereas *S. solfataricus* DNA ligase had an open extended conformation (Pascal *et al.*, 2006).

Functional studies of archaeal DNA ligases showed that all of them utilized ATP as a cofactor and it was presumed that archaeal ligases would exclusively utilize ATP as their nucleotide substrate. However,

this idea was challenged by studies of DNA ligases from euryarchaea of the order *Thermococcales*. The ATP-dependent DNA ligases from *Thermococcus kodakaraensis*, *T. fumicolans*, *T. onnurineus* and *Pyrococcus abyssi* were found to have the ability to also use NAD^+ as a cofactor (Nakatani *et al.*, 2000; Rolland *et al.*, 2004; Kim *et al.*, 2006). Thus, it was suggested that the dual specificity of DNA ligases from *Thermococcales* illustrates an intermediate phase of the evolution of an ancestral enzyme towards 'NAD^+-only' bacterial DNA ligases (Sun *et al.*, 2008). However, dual cofactor specificity is not a general feature of DNA ligases from *Thermococcales*, as exemplified by DNA ligase from *P. horikoshii*, which can only use ATP and not NAD^+ (Keppetipola & Shuman, 2005).

Previously, we determined the complete genome sequence of the hyperthermophilic euryarchaeon *T. sibiricus* MM 739 (Mardanov *et al.*, 2009) and identified the DNA ligase gene. The *T. sibiricus* DNA ligase (LigTsib) shares a high level of amino-sequence identity (77–79%) with the DNA ligases of *T. kodakaraensis*, *T. fumicolans*, *T. onnurineus* and *P. abyssi*, suggesting that they may have similar cofactor requirements. However, our functional analysis of LigTsib shows that this enzyme can utilize only ATP and not NAD^+ (to be published elsewhere). Further comparative analysis of structures of LigTsib and DNA ligases from other *Thermococcales* would reveal the molecular features that determine the cofactor specificity of these DNA ligases and contribute to an understanding of their evolution. Here, we report the overexpression, purification and preliminary crystallographic studies of LigTsib.

2. Expression and purification

The gene encoding the ATP-dependent LigTsib (Tsib_0885) was cloned into the expression vector pQE30 (Qiagen) by adding an MRGSHHHHHHGS tag to the N-terminus of the recombinant enzyme. The recombinant vector pQELigTsib was introduced into *Escherichia coli* DLT1270/pRARE-2 strain. The transformants were cultivated in LB medium containing 100 µg ml^{-1} ampicillin and 20 µg ml^{-1} chloramphenicol at 310 K until the optical density at 600 nm reached 0.5. Isopropyl β-D-1-thiogalactopyranoside was added to a final concentration of 1 m*M* to induce gene expression, which continued for 15 h.

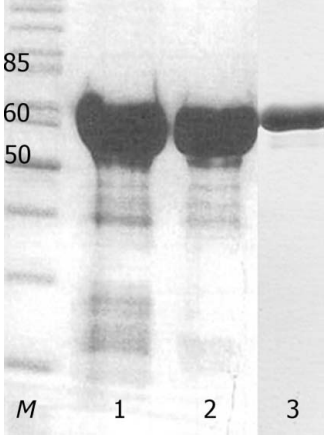


Figure 1
SDS–PAGE analysis of LigTsib during purification. Proteins were analysed on 12%(*w*/*v*) SDS–PAGE and stained with Coomassie Brilliant Blue. Lane 1, LigTsib after Ni-Sepharose FF column chromatography; lane 2, purified LigTsib after cation-exchange chromatography; lane 3, purified LigTsib after hydrophobic chromatography; lane *M*, molecular-weight markers (labelled in kDa).

The harvested cells were suspended in buffer *A* [50 m*M* Tris–HCl pH 7.5, 250 m*M* NaCl, 1 m*M* phenylmethanesulfonyl fluoride (PMSF), 5 m*M* β-mercaptoethanol (β-ME), 0.1%(*v*/*v*) Triton X-100, 10%(*v*/*v*) glycerol] for 30 min and disrupted using a French press. Cell debris was removed by centrifugation at 12 000*g* for 30 min and the supernatant was collected. The soluble fraction of the cell-free extract was heat-treated at 343 K for 15 min and the precipitate was removed by centrifugation (12 000*g* for 30 min). The collected supernatant was stirred with 0.15%(*v*/*v*) polyethyleneimine for 15 min at 277 K and nucleic acids were removed by centrifugation (12 000*g* for 30 min). The resulting nucleotide-free solution was applied onto an affinity column (Ni-Sepharose FF, 5 ml; Amersham Pharmacia Biotech) and equilibrated with buffer *B* (50 m*M* Tris–HCl pH 7.5, 500 m*M* NaCl, 1 m*M* PMSF, 20 m*M* imidazole). After washing with buffer *B*, the enzyme was eluted with a linear imidazole gradient from 100 to 500 m*M* in buffer *B*. The peak fractions eluted at 150 m*M* imidazole and were dialyzed overnight against buffer *C* [25 m*M* Tris–HCl pH 6.5, 50 m*M* KCl, 5 m*M* EDTA, 1 m*M* PMSF, 5%(*v*/*v*) glycerol] with 5%(*v*/*v*) glycerol and 5 m*M* EDTA added to the collected fractions. The solution was then applied onto a cation-exchange column equilibrated with buffer *C* (Mono S, 10 ml; Amersham Pharmacia Biotech) and the fractions were eluted with a linear KCl gradient from 0 to 500 m*M*. The homogeneity and purity of the fractions were determined by 12%(*w*/*v*) SDS–PAGE with Coomassie Brilliant Blue staining (Fig. 1). The purest fractions after cation-exchange chromatography were diluted twice with 3 *M* ammonium sulfate and loaded onto a hydrophobic column (Source 15Phe, Amersham Pharmacia Biotech) equilibrated with buffer *D* [100 m*M* sodium phosphate pH 6.7, 1.5 *M* ammonium sulfate, 0.2 *M* sucrose, 5%(*v*/*v*) glycerol, 1 m*M* PMSF]. The enzyme was eluted with buffer *D* without ammonium sulfate, transferred into crystallization buffer (50 m*M* bis-tris pH 6.0, 100 m*M* NaCl) and concentrated to 13.3 mg ml^{-1}. 10% glycerol was added to the protein solution immediately after concentration. The purity of the prepared sample was judged using 12%(*w*/*v*) SDS–PAGE (Fig. 1, lane 3). The protein concentration was determined by the Bradford method using BSA as a standard.

3. Crystallization

All crystallization experiments were performed at 291 K using the hanging-drop vapour-diffusion method. 1 µl protein solution was mixed with the same volume of precipitant solution. Initial crystal-

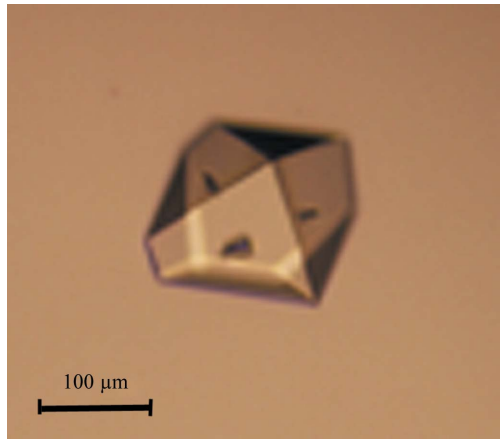


Figure 2
Crystal of LigTsib.

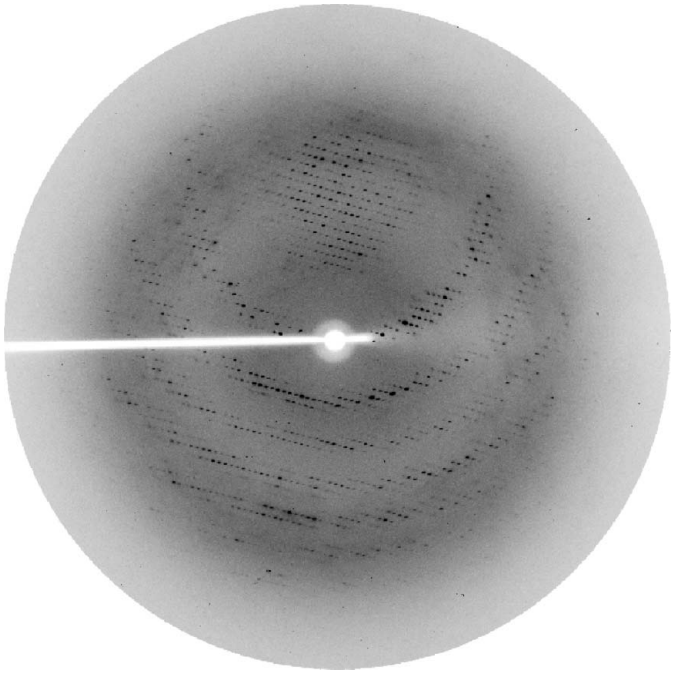


Figure 3
A typical 0.45° oscillation image obtained during data collection from the LigTsib crystal. The edge of the oscillation image corresponds to 3.0 Å resolution.

lization screening of LigTsib was carried out with the Crystal Screen, Index and Crystal Screen Cryo (Hampton Research) kits. A single crystal of truncated square-bipyramidal habit with dimensions of about 150–200 µm was obtained after 10–14 d in Crystal Screen Cryo (Hampton Research) condition No. 41 [0.085 *M* HEPES pH 7.5, 8.5%(*v*/*v*) 2-propanol, 17%(*w*/*v*) PEG 4000, 15%(*v*/*v*) glycerol; Fig. 2]. Further refinement of the crystallization conditions did not result in crystals of better quality.

4. X-ray analysis

An X-ray data set was collected at 100 K on the 'Belok' synchrotron beamline at the National Research Center 'Kurchatov Institute' using a Rayonix SX165 detector (Rayonix LLC, USA; Fig. 3). The crystals were directly flash-cooled in a stream of cold nitrogen gas at 100 K using a Cryostream Plus cooling device (Oxford Cryosystems Ltd, England). There was no need for prior crystal transfer to cryoprotectant solution. Data collection was performed using the *DNA* software package (http://www.dna.ac.uk). Experimental details are summarized in Table 1.

X-ray diffraction images were indexed, integrated and scaled using *XDS* (Kabsch, 2010). The *CCP*4 package (Winn *et al.*, 2011) was used for data reduction. Data-collection statistics are given in Table 1.

The structure of LigTsib was solved using the program *BALBES* (Long *et al.*, 2008). The best solution found by *BALBES* was obtained using the DNA ligase from *P. furiosus* (PDB entry 2cfm; Nishida *et al.*, 2006) as a starting model for molecular replacement. The best solution had one molecule in the asymmetric unit of the unit cell, showed good crystal packing and gave an *R* factor of 0.32 for data in the resolution range 29.3–3.0 Å. The model consisted of three domains and showed a closed domain arrangement, as previously observed for DNA ligases from *P. furiosus* and *A. fulgidus*. Refinement of the structure is currently in progress and will be published elsewhere.

Table 1
Experimental setup and statistics of data collection and processing.

Values in parentheses are for the highest resolution shell.

Space group	$P2_12_12_1$
Radiation source	National Research Center 'Kurchatov Institute'
Unit-cell parameters (Å)	a = 58.590, b = 87.540, c = 126.300
Temperature (K)	100
Wavelength (Å)	0.985
Crystal-to-detector distance (mm)	190
Oscillation range (°)	1.0
Mosaicity (°)	0.37
No. of frames	103
Resolution limit (Å)	29.30–3.00 (3.30–3.00)
Total reflections	54735 (13682)
Molecules per asymmetric unit	1
Solvent content (%)	52
Matthews coefficient (Å^3 Da^{-1})	2.52
Independent reflections	13400 (3285)
Average $I/\sigma(I)$	12.2 (3.20)
Completeness (%)	98.9 (99.7)
$R_{\rm merge}$†	0.11 (0.48)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of the ith measurement of reflection hkl and $\langle I(hkl)\rangle$ is the average intensity of the reflection.

This work was performed using the scientific equipment of the Core Research Facility of the 'Bioengineering' Center of RAS and was supported by the Ministry of Education and Sciences of Russia (contract 02.740.11.0765).

References

Bezsudnova, E. Y., Kovalchuk, M. V., Mardanov, A. V., Poliakov, K. M., Popov, V. O., Ravin, N. V., Skryabin, K. G., Smagin, V. A., Stekhanova, T. N. & Tikhonova, T. V. (2009). *Acta Cryst.* F**65**, 368–371.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Keppetipola, N. & Shuman, S. (2005). *J. Bacteriol.* **187**, 6902–6908.
Kim, D. J., Kim, O., Kim, H.-W., Kim, H. S., Lee, S. J. & Suh, S. W. (2009). *Acta Cryst.* F**65**, 544–550.
Kim, Y. J., Lee, H. S., Bae, S. S., Jeon, J. H., Yang, S. H., Lim, J. K., Kang, S. G., Kwon, S.-T. & Lee, J.-H. (2006). *Biotechnol. Lett.* **28**, 401–407.
Lai, X., Shao, H., Hao, F. & Huang, L. (2002). *Extremophiles*, **6**, 469–477.
Lehman, I. R. (1974). *Science*, **186**, 790–797.
Long, F., Vagin, A. A., Young, P. & Murshudov, G. N. (2008). *Acta Cryst.* D**64**, 125–132.
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.
Nakatani, M., Ezaki, S., Atomi, H. & Imanaka, T. (2000). *J. Bacteriol.* **182**, 6424–6433.
Nishida, H., Kiyonari, S., Ishino, Y. & Morikawa, K. (2006). *J. Mol. Biol.* **360**, 956–967.
Pascal, J. M. (2008). *Curr. Opin. Struct. Biol.* **18**, 96–105.
Pascal, J. M., Tsodikov, O. V., Hura, G. L., Song, W., Cotner, E. A., Classen, S., Tomkinson, A. E., Tainer, J. A. & Ellenberger, T. (2006). *Mol. Cell*, **24**, 279–291.
Rolland, J. L., Gueguen, Y., Persillon, C., Masson, J. M. & Dietrich, J. (2004). *FEMS Microbiol. Lett.* **236**, 267–273.
Smagin, V. A., Mardanov, A. V., Bonch-Osmolovskaia, E. A. & Ravin, N. V. (2008). *Prikl. Biokhim. Mikrobiol.* **44**, 523–528.
Sriskanda, V., Kelman, Z., Hurwitz, J. & Shuman, S. (2000). *Nucleic Acids Res.* **28**, 2221–2228.
Sun, Y., Seo, M. S., Kim, J. H., Kim, Y. J., Kim, G. A., Lee, J. I., Lee, J.-H. & Kwon, S.-T. (2008). *Environ. Microbiol.* **10**, 3212–3224.
Supangat, S., An, Y. J., Sun, Y., Kwon, S.-T. & Cha, S.-S. (2010). *Acta Cryst.* F**66**, 1583–1585.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.