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

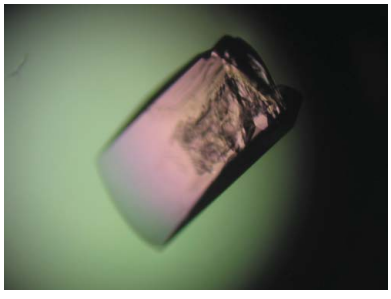
Overexpression, purification and crystallization of an archaeal DNA ligase from *Pyrococcus furiosus*

Hirokazu Nishida,[a] Daisuke Tsuchiya,[a] Yoshizumi Ishino[b] and Kosuke Morikawa[a]*

[a]Department of Structural Biology, Biomolecular Engineering Research Institute (BERI), 6-2-3 Furuedai, Suita, Osaka 565-0874, Japan, and [b]Department of Genetic Resources Technology, Faculty of Agriculture, Kyushu University, 6-10-1 Hakozaki, Fukuoka-shi, Fukuoka 812-8581, Japan

Correspondence e-mail: morikako@protein.osaka-u.ac.jp

Received 27 October 2005
Accepted 22 November 2005
Online 29 November 2005



DNA ligases seal single-strand breaks in double-stranded DNA and their function is essential to maintain the integrity of the genome during various aspects of DNA metabolism, such as replication, excision repair and recombination. DNA-strand breaks are frequently generated as reaction intermediates in these events and the sealing of these breaks depends solely on the proper function of DNA ligase. Crystals of the archaeal DNA ligase from *Pyrococcus furiosus* were obtained using 6.6%(*v*/*v*) ethanol as a precipitant and diffracted X-rays to 1.7 Å resolution. They belong to the monoclinic space group $P2_1$, with unit-cell parameters a = 61.1, b = 88.3, c = 63.4 Å, β = 108.9°. The asymmetric unit contains one ligase molecule.

1. Introduction

DNA ligases are universally found in all organisms (bacteria, eukarya and archaea), including viruses and bacteriophages (reviewed in Timson *et al.*, 2000; Tomkinson & Mackey, 1998; Wilkinson *et al.*, 2001). DNA ligases catalyze phosphodiester-bond formation between adjacent 5′-phosphoryl and 3′-hydroxyl groups at a single-strand break in double-stranded DNA. These enzymes are essential for maintaining genome information during DNA replication and repair. The replication and repair pathways converge at a common final step in which the continuity of the replicated lagging strands or the repaired DNA strands is restored by DNA ligases. Unsealed nicks are potential lesions that may cause double-strand breaks, which are lethal if they are not repaired through homologous recombination or ligase-mediated non-homologous end joining. Therefore, DNA ligases are indispensable in all organisms.

DNA ligases are categorized into two groups on the basis of the required cofactor for activity: one group requires ATP and the other group requires NAD^+ (Lehman, 1974). ATP-dependent ligases are present in viruses, bacteriophages, eukarya, archaea and bacteria, whereas NAD^+-dependent ligases have been found exclusively in bacteria. ATP-dependent and NAD^+-dependent DNA ligases catalyze reactions using a common mechanism. The reaction proceeds *via* three sequential nucleotidyl-transfer steps. In the first step, the conserved active-site lysine of the protein is activated by the covalent addition of AMP *via* a phosphoamide bond to the N^{ζ} atom of the lysine, accompanied by the release of PP_i or nicotinamide mononucleotide from the cofactor (ATP or NAD^+). In the second step, AMP is transferred from the protein to the 5′-phosphoryl group of the nick on the DNA to form the DNA–adenylate intermediate. In the last step, the phosphodiester bond is formed with the concomitant release of AMP from the DNA–adenylate intermediate.

Biochemical and genetic studies of the DNA ligases from various organisms have been performed. Viruses and bacteriophages each appear to have a single type of DNA ligase, whereas eukaryotic organisms possess multiple enzymes. The ATP-dependent enzymes show a wide range of molecular weights, which is mainly a consequence of diversity in the N-terminal region of the DNA ligases. DNA ligases from archaea and eukarya have a large extra extension at their N-termini and their molecular weights range from 23 kDa (euryarchaeota) to 64 kDa (mammalian DNA ligase I), with the catalytic core domains towards the C-terminus. The N-terminal extra region of human DNA ligase I is essential for interactions with

proliferating cell nuclear antigen (PCNA) and some types of DNA polymerase (Levin *et al.*, 1997; Prasad *et al.*, 1996; Mossi *et al.*, 1998). These results indicate that the N-terminal extra regions of the eukaryotic and archaeal ATP-dependent DNA ligases are utilized for interaction with other protein factors. On the other hand, the NAD^+-dependent DNA ligases have a long extension at their C-terminus. The C-terminal half of the protein forms a circular structure in collaboration with the N-terminal catalytic core domains and the resultant central hole in the molecule is capable of holding double-stranded DNA (Lee *et al.*, 2000). This structure implies that the C-terminal extra region of the NAD^+-dependent DNA ligases from bacteria is required for binding DNA solely by the ligase molecule, in contrast to the N-terminal extra region of the ATP-dependent DNA ligases from higher organisms.

Six DNA-ligase motifs are commonly found in the catalytic core domains of DNA ligases (Shuman *et al.*, 1994). The roles of these motifs were clarified in the structural analyses of bacteriophage T7 DNA ligase (Subramanya *et al.*, 1996) and *Chlorella* virus DNA ligase (Odell *et al.*, 2000). When the sequences of the ATP-dependent DNA ligases from bacteriophages, viruses, archaea and eukarya were compared with each other, the six motifs aligned well without long gaps or insertions. However, extensions ranging from 20 to 35 residues were commonly found after the sixth motif at the C-termini of the sequences from archaea and eukarya, whereas the bacteriophage and virus sequences terminated shortly after the sixth motif. The role of this C-terminal extension is a subject of interest.

One crystal structure of a eukaryotic DNA ligase has been reported to date (Pascal *et al.*, 2004), but this molecule harbours the substrate DNA in the open conformation. We have crystallized the substrate-free archaeal DNA ligase, in which a large conformational change is expected. Here, we report the overexpression, purification and preliminary crystallographic studies of the ATP-dependent DNA ligase from *Pyrococcus furiosus*.

2. Expression and purification

The *P. furiosus* DNA ligase (*Pfu*Lig) gene was cloned from *P. furiosus* genomic DNA and inserted into the expression vector pET-21d (Novagen) for *Escherichia coli* cells (Takayama & Ishino, unpublished). The resulting plasmid, pET-lig, was introduced into *E. coli* BL21-Codon Plus(DE3)-RIL (Stratagene). The transformants were cultivated in Luria–Bertani medium containing 100 µg ml^{-1} ampicillin and 20 µg ml^{-1} chloramphenicol at 310 K until the optical density at 660 nm reached 0.6. Isopropyl-β-D-thiogalactopyranoside was added to a final concentration of 1 m*M* to induce gene expression, which continued for 12 h.

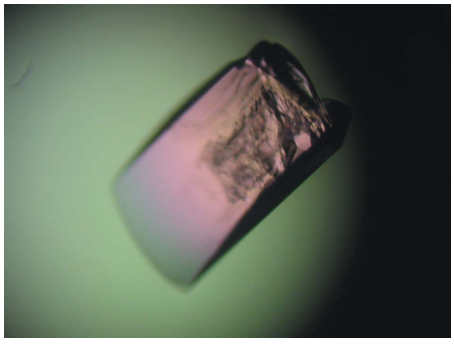


Figure 1
Crystal of *Pfu* DNA ligase grown in ethanol solution.

The harvested cells were suspended in buffer *A* (50 m*M* Tris–HCl pH 8.0, 250 m*M* NaCl, 0.5 m*M* DTT) and disrupted by sonication. The cell debris was removed by centrifugation at 30 000*g* for 15 min and the supernatant was collected. The soluble fraction of the cell-free extract was heat-treated at 353 K for 20 min and the precipitate was removed by centrifugation (30 000*g*, 15 min). The collected supernatant was stirred with 0.15%(*v*/*v*) polyethyleneimine for 15 min at 277 K and the nucleic acids were then removed by centrifugation (30 000*g*, 15 min).

The resulting nucleotide-free solution was applied onto an affinity column (HiTrap Heparin, 5 ml; Amersham Pharmacia Biotech), equilibrated with buffer *B* (50 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 0.5 m*M* DTT). After washing with buffer *B*, the enzyme was eluted with a linear gradient of 0.1–1.0 *M* NaCl in buffer *B*. The peak fractions eluted between 0.40 and 0.50 *M* NaCl and were collected and dialyzed against buffer *B*. The solution was then applied onto an anion-exchange column equilibrated with buffer *B* (HiTrap Q, 5 ml; Amersham Pharmacia Biotech) and the flowthrough fractions were collected. The enzyme solution was brought to 80% saturation by adding solid ammonium sulfate with stirring. The precipitate was recovered by centrifugation for 20 min at 30 000*g*. The pellet was resuspended in buffer *A* and dialyzed against the same buffer for 12 h. The solution was concentrated and loaded onto a gel-filtration column (Superdex 200 HiLoad 26/60, Amersham Pharmacia Biotech) equilibrated with buffer *B* and eluted with the same buffer. The fractions in a single peak were collected and used as purified *Pfu*Lig in the following experiments. The protein concentration was determined with the Bio-Rad protein assay system, with BSA as the standard.

3. Crystallization

Crystallization was initiated at 293 K using the hanging-drop vapour-diffusion method with Crystal Screen and Crystal Screen 2 from Hampton Research (Jancarik & Kim, 1991). In the initial screening, 10, 20 and 40 mg ml^{-1} concentrations of *Pfu*Lig were used, with 1 µl of protein solution mixed with half the amount of precipitant solution. Large single crystals (0.3 × 0.5 × 1 mm) were obtained after 2–3 d from a mixture of protein solution [40 mg ml^{-1} protein, 10%(*v*/*v*) glycerol, 0.5 m*M* dithiothreitol, 100 m*M* NaCl, 15 m*M* $MgCl_2$, 100 m*M* Tris–HCl pH 8.5] and precipitant solution [20%(*v*/*v*) ethanol, 100 m*M* Tris–HCl pH 8.5; Crystal Screen 2 No. 44, Hampton Research] (Fig. 1).

4. X-ray analysis

Crystals were cryoprotected by soaking briefly in mother liquor supplemented with 20%(*v*/*v*) 2-methyl-2,4-pentanediol before flash-freezing in a boiled-off N_2 stream. The crystals belong to space group $P2_1$, with unit-cell parameters $a = 60.8$, $b = 88.5$, $c = 63.1$ Å, $\beta = 108.7^\circ$. Assuming the presence of one molecule in the asymmetric unit, the solvent content of the crystals is 51%, corresponding to a Matthews coefficient (Matthews, 1968) V_M of 2.5 $Å^3$ Da^{-1}. The native data set was collected at beamline BL38 (SPring-8, Japan) to a resolution of 1.75 Å. Diffraction data were processed and scaled with *HKL*2000 (Otwinowski & Minor, 1997) and the data-collection statistics are summarized in Table 1.

The crystal described in this report showed diffraction limits of better than 1.8 Å. We expect that structural information from this substrate-free archaeal DNA ligase at high resolution will help to further clarify the catalytic mechanism in detail and will explain the

Table 1
Data-collection and processing statistics.

Values in parentheses are for the last shell.

Space group	$P2_1$
Radiation	BL38, Spring-8
Unit-cell parameters (Å, °)	a = 61.1, b = 88.3, c = 63.4, β = 108.9
Temperature (K)	100
Wavelength (Å)	1.0
Oscillation range (°)	1.0
No. of frames	260
Resolution limit (Å)	50.0–1.75 (1.82–1.75)
Total reflections	295859
Independent reflections	58797 (5904)
Average redundancy	5.0 (4.9)
Average $I/\sigma(I)$	15.5 (6.5)
Completeness (%)	92.0 (83.9)
$R_{\rm merge}$†	0.065 (0.481)

† $R_{\rm merge} = \sum_{hkl}\sum_j |I(hkl)_j - \langle I(hkl)\rangle| / \sum_{hkl}\sum_j I(hkl)_j$, where $I(hkl)_j$ is the jth measurement of the intensity of reflection hkl and $\langle I(hkl)\rangle$ is the mean intensity of reflection hkl.

functional significance of the N- and C-terminal extensions in the DNA ligases from higher organisms.

We are grateful to Dr T. Oyama for helpful discussions. This work was partly supported by a research grant endorsed by the New Energy and Industrial Technology Development Organization (NEDO) and YI was supported by a grant from the Human Frontier Science Program (HFSP) and a Grant-in-aid from the Ministry of Education, Science and Sports of Japan.

References

Jancarik, J. & Kim, S.-H. (1991). *J. Appl. Cryst.* **24**, 409–411.
Lee, J. Y., Chang, C., Song, H. K., Moon, J., Yang, J. K., Kim, H.-K., Kwon, S.-T. & Suh, S. W. (2000). *EMBO J.* **19**, 1119–1129.
Lehman, I. (1974). *Science*, **186**,790–797.
Levin, D. S., Bai, W., Yao, N., O'Donnell, M. & Tomkinson, A. E. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 12863–12868.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Mossi, R., Ferrari, E. & Hübscher, U. (1998). *J. Biol. Chem.* **273**, 14322–14330.
Odell, M., Sriskanda, V., Shuman, S. & Nikolov, D. B. (2000). *Mol. Cell*, **6**, 1183–1193.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pascal, J. M., O'Brien, P. J., Tomkinson, A. E. & Ellenberger, T. (2004). *Nature (London)*, **432**, 473–478.
Prasad, R., Singhal, R. K., Srivastava, D. K., Molina, J. T., Tomkinson, A. E. & Wilson, S. H. (1996). *J. Biol. Chem.* **271**, 16000–16007.
Shuman, S., Liu, Y. & Schwer, B. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 12046–12050.
Subramanya, H. S., Doherty, A. J., Ashford, S. R. & Wigley, D. B. (1996). *Cell*, **85**, 607–615.
Timson, D. J., Singleton, M. R. & Wigley, D. B. (2000). *Mutat. Res.* **460**, 301–318.
Tomkinson, A. E. & Mackey, Z. B. (1998). *Mutat. Res.* **407**, 1–9.
Wilkinson, A., Day, J. & Bowater, R. (2001). *Mol. Microbiol.* **40**, 1241–1248.