Biochemical characterization of an ATP-dependent DNA ligase from the hyperthermophilic crenarchaeon *Sulfolobus shibatae*

**Abstract** A gene encoding a putative ATP-dependent DNA ligase was identified in the genome of the hyperthermophilic archaeon *Sulfolobus shibatae* and expressed in *Escherichia coli*. The 601 amino acid recombinant polypeptide was a monomeric protein capable of strand joining on a singly nicked DNA substrate in the presence of ATP ($K_m = 34 \mu M$) and a divalent cation (Mg$^{2+}$, Mn$^{2+}$, or Ca$^{2+}$). dATP was partially active in supporting ligation catalyzed by the protein, but GTP, CTP, UTP, dGTP, dCTP, dTTP, and NAD$^+$ were inactive. The cloned *Ssh* ligase showed an unusual metal cofactor requirement; it was significantly more active in the presence of Mn$^{2+}$ than in the presence of Mg$^{2+}$ or Ca$^{2+}$. Unexpectedly, the native *Ssh* ligase preferred Mg$^{2+}$ and Ca$^{2+}$ rather than Mn$^{2+}$. Both native and recombinant enzymes displayed optimal nick-joining activity at 60–80°C. *Ssh* ligase discriminated against substrates containing mismatches on the 3′-side of nick junction and was more tolerant of mismatches at the penultimate 5′-end. The enzyme showed little activity on a 1-nucleotide gapped substrate. This is the first biochemical study of a DNA ligase from the crenarchaeotal branch of the archaea domain.

**Key words** DNA ligase · Crenarchaeota · *Sulfolobus shibatae* · ATP · Metal cofactor · Thermostability

**Introduction**

DNA ligases catalyze the formation of a phosphodiester bond between adjacent 3′-hydroxyl and 5′-phosphoryl groups at a single-stranded break in double-stranded DNA (Lehman 1974). These enzymes are an essential component in DNA replication, recombination, and repair systems and are therefore indispensable in all organisms (Lindahl and Barnes 1992). Based on energy cofactor requirement, DNA ligases fall into two groups: those requiring ATP for activity and those requiring NAD$^+$ for activity. ATP- and NAD$^+$-dependent DNA ligases employ a common three-step mechanism in ligation reaction (Engler and Richardson 1982). In the first step the enzyme is activated through the covalent addition of the adenylate group from ATP or NAD$^+$ to the conserved active site lysine. The second step involves the transfer of AMP from the enzyme to the 5′-phosphoryl group at a nick site on the DNA strand. In the final step the nick is sealed with the release of AMP from the adenylated DNA intermediate.

NAD$^+$-dependent ligases are found only in bacteria, whereas ATP-dependent ligases are encoded in eukarya, archaea, bacteriophages as well as bacteria. The two classes of ligases share limited sequence homology but have a common core architecture (Doherty and Suh 2000). These enzymes are characterized structurally by a set of six short motifs (I, III, IIIa, IV, V, and VI) conserved in sequence, order and spacing (Aravind and Koonin 1999). The conserved motifs play critical roles in nucleotide binding, nick recognition, and nucleotidyl transfer, as revealed in structural and mutational studies (Doherty and Daftorn 2000; Doherty and Wigley 1999; Lee et al. 2000; Odell et al. 2000; Singleton et al. 1999; Sriskanda and Shuman 1998; Subramanya et al. 1996).

Biochemical and genetic studies have been performed on ligases from eukarya, bacteria, viruses, and recently archaea. As the third domain of life, archaea have been shown to be similar to eukarya in many aspects of DNA metabolism despite their morphological and structural resemblance to bacteria (Edgell and Doolittle 1997; Olsen and Woese 1997). Culturable archaea are divided into two
kingdoms: the Crenarchaeota and the Euryarchaeota. Kletzin (1992) identified the first archaeal ligase gene from a hyperthermophilic crenarchaeon, Desulfovirulobus ambivalens, and showed that the gene encoded a protein similar at the amino acid sequence level to eukaryal ATP-dependent DNA ligases. In recent years a number of genes encoding putative ATP-dependent ligases have been identified from sequenced archaeal genomes. Phylogenetic analysis indicates that all archaeal DNA ligases are closely related to DNA ligase I from eukarya (Nakatani et al. 2000). The availability of these putative genes has offered good opportunities for studying DNA ligases in archaea. However, only two archaeal DNA ligases have been biochemically characterized so far. An ATP-dependent DNA ligase from the thermophilic euryarchaeon Methanobacterium thermoautotrophicum was the first DNA ligase to be characterized in archaea (Sriskanda et al. 2000). This was followed by the biochemical study of a DNA ligase from the hyperthermophilic euryarchaeon Thermococcus kodakaraensis (Nakatani et al. 2000). Both enzymes are capable of closing nicks in an ATP-dependent manner at elevated temperatures characteristic of those required for the optimal growth of the host organisms. Interestingly, the T. kodakaraensis enzyme appears to have a low but significant activity when using NAD+ as the cofactor. No biochemical studies have been conducted on DNA ligases from crenarchaeta, which have been shown to differ from euryarchaeta in DNA replication and cell cycle process (She et al. 2001).

In this study we identified a gene encoding a putative ATP-dependent DNA ligase in the genome of the hyperthermophilic crenarchaeon Sulfolobus shibatae. The putative protein shares sequence homology with known ATP-dependent DNA ligases and consists of the six conserved sequence motifs common to all ATP-dependent DNA ligases. Sequence comparison shows that putative ligases encoded by S. shibatae and other crenarchaeotal species are more closely related at the amino acid sequence level to eukaryal ATP-dependent ligases than their euryarchaeotal counterparts. In order to investigate the enzymatic properties of the S. shibatae protein and compare them with those of other ligases we overexpressed the gene coding for the protein in Escherichia coli and purified the recombinant protein to apparent homogeneity. We show that the protein displays optimal nick-joining activity at temperatures between 60° and 80°C in the presence of ATP and therefore is indeed an ATP-dependent DNA ligase. We also demonstrate that the identified ligase gene is expressed in S. shibatae. A biochemical characterization of Ssh ligase is presented.

Materials and methods

Growth of Sulfolobus shibatae

S. shibatae 51178 was purchased from the American Type Culture Collection (Rockville, Md., USA). The organism was grown at 75°C in Brock’s medium (Brock et al. 1972) supplemented with 0.2% tryptone and 0.1% yeast extract in a 5-l Bioflo fermentor (New Brunswick Scientific, Edison, N.J., USA).

Computer analysis of protein sequences

Protein sequences were retrieved from public sequence databases. The ClustalW program (Thompson et al. 1994) was employed to produce a multiple sequence alignment and to identify conserved regions.

Isolation of the putative Ssh ligase gene

Based on motif I and VI sequences in selected known archaeal and eukaryal ATP-dependent DNA ligases, the following pair of degenerate oligonucleotide primers were synthesized: 5′-AARTARGATGWRHRMGWKGKWCA RRTWCA/5′-GGRAATCKYADNGMDWANCC (R: A/ G, Y: C/T, M: A/C, K: G/T, W: A/T, H: A/C/T, D: A/G/T, N: A/C/G/T). A PCR reaction was carried out with the S. shibatae genomic DNA as the template, using the above primers, in the presence of 40 μM digoxigenin-11-dUTP (DIG; Boehringer-Mannheim, Mannheim, Germany). A HindIII digest of the S. shibatae genomic DNA was resolved by electrophoresis in an agarose gel. DNA fragments identified by hybridization with the DIG-labeled probe were recovered and cloned into the HindIII site of pUC18. After transformation into E. coli strain DH5α, recombinant plasmids were screened by colony hybridization with the DIG-labeled probe. Inserts in plasmids isolated from positive clones were sequenced in both directions.

Expression of the Ssh ligase gene

The putative Ssh ligase gene was amplified by PCR using a pair of primers (5′-CGGGGATTCTATATGGATGTTTAAAA GTTAT/5′-CCCCGGATCCGGATATTTTAAACGCTC) derived from the 5′- and 3′-ends of the gene. The primers were designed to introduce NdeI and BamHI sites at the 5′- and 3′-ends, respectively, of the Ssh ligase gene. S. shibatae genomic DNA was used as the template for PCR. The PCR product was cleaved with NdeI and BamHI, and cloned into the NdeI and BamHI sites of the T7-based expression plasmid pET-30a(+) (Novagen, Madison, Wis., USA). The sequence of the insert was verified by sequencing in both directions. The recombinant plasmid was transformed into E. coli BL21(DE3) (Stratagene, La Jolla, Calif., USA). The transformant was grown in Luria-Bertani medium containing 30 μg kanamycin/ml at 37°C until the optical density at 600 nm reached approx. 0.5. isopropyl-d-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Incubation was continued for 3 h.

Purification of recombinant Ssh ligase

Fully induced cells of the constructed Ssh ligase overproducer were harvested by centrifugation (5,000 g, 10 min,