Ligation reaction specificities of an NAD⁺-dependent DNA ligase from the hyperthermophile *Aquifex aeolicus*

Jie Tong, Francis Barany and Weiguo Cao*

Department of Microbiology and Immunology, Hearst Microbiology Research Center and Strang Cancer Prevention Center, The Joan and Sanford I. Weill Medical College of Cornell University, 1300 York Avenue, Box 62, New York, NY 10021, USA

Received November 3, 1999; Revised and Accepted January 18, 2000

DDBJ/EMBL/GenBank accession no. AAC06838

ABSTRACT

An NAD⁺-dependent DNA ligase from the hyperthermophilic bacterium *Aquifex aeolicus* was cloned, expressed in *Escherichia coli* and purified to homogeneity. The enzyme is most active in slightly alkaline pH conditions with either Mg²⁺ or Mn²⁺ as the metal cofactor. Ca²⁺ and Ni²⁺ mainly support formation of DNA–adenylate intermediates. The catalytic cycle is characterized by a low k_{cat} value of 2 min⁻¹ with concomitant accumulation of the DNA–adenylate intermediate when Mg²⁺ is used as the metal cofactor. The ligation rates of matched substrates vary by up to 4-fold, but exhibit a general trend of T/A ≤ G/C < C/G < A/T on both the 3′- and 5′-side of the nick. Consistent with previous studies on *Thermus* ligases, this *Aquifex* ligase exhibits greater discrimination against a mismatched base pair on the 3′-side of the nick junction. The requirement of 3′ complementarity for a ligation reaction is reaffirmed by results from 1 nt insertions on either the 3′- or 5′-side of the nick. Furthermore, most of the unligatable 3′ mismatched base pairs prohibit formation of the DNA–adenylate intermediate, indicating that the substrate adenylation step is also a control point for ligation fidelity. Unlike previously studied ATP ligases, gapped substrates cannot be ligated and intermediate accumulation is minimal, suggesting that complete elimination of base pair complementarity on one side of the nick affects substrate adenylation on the 5′-side of the nick junction. Relationships among metal cofactors, ligation products and intermediate, and ligation fidelity are discussed.

INTRODUCTION

DNA ligases are universally found in bacteria, archaea and eukaryotes. Bacteria rely on NAD⁺-dependent DNA ligases to close nicks during DNA replication, repair and recombination. Early studies on *Escherichia coli* and T4 ligases have elucidated the basic enzymatic mechanism of DNA ligation and

demonstrated the different cofactor requirements. ATP-dependent ligases have been found in bacteriophages, archaea, eukaryotes and, recently, in bacteria. NAD⁺-dependent ligases have been found exclusively in bacteria. For both NAD⁺ and ATP ligases, the ligation reaction involves formation of two intermediates (first step, enzyme–adenylate; second step, DNA–adenylate) and a third nick closure step (1).

Bacterial ligases from thermophilic organisms have become model systems for structural and mechanistic studies of DNA ligation due to their high thermostability. The crystal structure of the N-terminal domain of an NAD⁺-dependent ligase from *Bacillus stearothermophilus* (Bst) has recently been solved at 2.8 Å resolution (2). The core of the NAD⁺-binding domain, which consists mainly of β-sheets, is similar to the core of the ATP-binding domain previously reported for the T7 ATP ligase (3). The N-terminal domain of the Bst ligase retains full activity in ligase–AMP formation with NAD⁺ but is poorly active in strand joining (4). NAD⁺-dependent ligases are highly conserved among *Thermus* spp. (5–7). Mutational analysis of *Thermus thermophilus* HB8 (Tth) NAD⁺-dependent ligase has identified a few amino acid residues important for different steps of the ligation pathway (8). Despite their high degree of homology, *Thermus* ligases differ in ligation fidelity. A NAD⁺-dependent ligase from *Thermus* sp. AK16D exhibits higher fidelity than Tth ligase with either Mg²⁺ or Mn²⁺ as the metal cofactor, emphasizing the subtle sequence changes required for substrate discrimination (7). The high ligation fidelity of Tth ligase has been exploited for the development of a ligase-based assay for detecting point mutations in genetic diseases and cancers (9–12). Studies of ligases from different sources provide the potential for the development of high sensitivity techniques for single base pair discrimination.

Aquifex aeolicus is a hyperthermophilic, hydrogen oxidizing, microaerophilic, obligate chemolithoautotroph (13), representing the deepest branch in the bacterial phylogeny (14,15). Its complete genome contains 1512 genes, two of which encode a putative NAD⁺-dependent ligase (AF633) and a putative ATP-dependent ligase (AF1394) (13). This work was undertaken to investigate the ligation activities of the putative NAD⁺-dependent DNA ligase from *Aquifex aeolicus* (Aae). We herein report biochemical characterization of the cloned Aae ligase.

*To whom correspondence should be addressed. Tel: +1 212 746 6517; Fax: +1 212 746 8587; Email: wgc@mail.med.cornell.edu

MATERIALS AND METHODS

Reagents, media and strains

All routine chemical reagents were purchased from Sigma Chemical Co. (St Louis, MO) or Fisher Scientific (Fair Lawn, NJ). Restriction enzymes and the T4 DNA ligase were obtained from New England Biolabs (Beverly, MA). Oligonucleotide synthesis reagents, DNA sequencing kits and PCR kits were obtained from PE Biosystems (Foster City, CA). dNTPs, bovine serum albumin (BSA), IPTG and ATP were purchased from Boehringer-Mannheim (Indianapolis, IN). *Pfu* DNA polymerase was purchased from Stratagene (La Jolla, CA). *Escherichia coli* strain NovaBlue(DE3)pLysS and plasmid pET21a were purchased from Novagen Inc. (Madison, WI). The protein assay kit was from Bio-Rad (Hercules, CA). The HiTrap Blue affinity column was from Pharmacia (Piscataway, NJ). LB medium was prepared according to the standard formula (16). Sonication buffer consisted of 50 mM Tris-HCl, pH 8.0 and 1 mM EDTA. TE buffer consisted of 10 mM Tris-HCl, pH 8.0 and 1 mM EDTA.

Oligodeoxyribonucleotide synthesis

Oligodeoxyribonucleotides were synthesized using a 394 automated DNA synthesizer from PE Biosystems. PCR and sequencing primers were purified by ethanol precipitation and dissolved in TE buffer. PCR amplification primers for cloning the Aae DNA ligase gene into the pET21a vector were d(GCGCAT-ATGTTTACCCAGAAAGGGA) and d(GCGGTCGACT-CAAAATAGTCTGCCTATTT), where the *NdeI* and *SalI* sites are underlined and the initiation codon in the forward primer is shown in bold. Oligonucleotide substrates for ligation assay were purified on a denaturing sequencing gel (7 M urea-10% polyacrylamide) (16). 5'-Phosphorylation of oligonucleotides was achieved during synthesis using Chemical Phosphorylation Reagent (Glen Research, Sterling, VA). A fluorescent group was attached to the 3'-terminus using a Fluorescein CPG column (Glen Research).

DNA amplification, cloning and sequence analysis

PCR amplification (25 cycles) was carried out in a GeneAmp PCR System 9700 thermocycler (PE Biosystems) as described (17). The full-length Aae DNA ligase gene was amplified using PCR amplification primers as described above, digested with *NdeI* and *SalI*, ligated into the cloning vector pET21a treated with the same pair of restriction enzymes and transformed into *E. coli* strain NovaBlue(DE3)pLysS. Inserts in pET expression vectors were sequenced in both orientations to ensure that the plasmid constructs were free of PCR or ligation errors. Nucleic acid and protein sequence analyses were carried out using the DNASTAR package (Madison, WI).

Expression and purification of Aae DNA ligase

Escherichia coli NovaBlue(DE3)pLysS cells containing plasmid pET21a-AaeligA encoding the Aae DNA ligase gene

from a pET21a construct were propagated overnight at 37°C in LB medium containing 50 µg/ml ampicillin, 25 µg/ml chloramphenicol and 0.2% glucose. Overnight cultures were diluted 100-fold into the same medium, grown until the optical density of the culture reached 0.5 at 600 nm, then induced by the addition of IPTG to a final concentration of 1 mM and grown for an additional 4 h under the same conditions. Cells were collected by centrifugation, frozen/thawed at -20/23°C, disrupted by sonication and clarified by centrifugation as previously described (17). The resulting supernatants were heated at 70°C for 15 min to denature thermolabile *E. coli* proteins, placed on ice for 30 min to aggregate the denatured proteins and cleared of denatured proteins by microcentrifugation for 15 min at 4°C. The partially pure DNA ligase was further purified by chromatography using a 1 ml HiTrap Blue affinity column. Briefly, the column containing Aae DNA ligase was washed extensively with TE buffer (pH 7.8) containing 0.1 M NaOAc and the ligase eluted with TE buffer (pH 7.8) containing 2 M NaCl. After dialysis against TE buffer (pH 8.0) containing 0.2 M KCl and concentration using Centricon-30 (Amicon), protein concentration was assayed by the Bradford method with reagents supplied in the Bio-Rad Protein Assay Kit. The amount of protein was determined using BSA as the standard. The purity of the ligase was verified by 7.5% SDS-PAGE analysis followed by visualizing the overloaded gel with routine Coomassie brilliant blue R staining.

Substrates and ligation assay

The oligonucleotide perfect match substrate was formed by annealing two short oligonucleotides (33mer for LP3'C and 30mer for Com3F) with a 59mer complementary oligonucleotide (Glg). Oligonucleotides LP3'C and Glg were in 1.5-fold excess so that all the 3'-Fam-labeled Com3F represented nicked substrates. The T/G mismatch substrate was formed by annealing LP3'T, which introduced a single base pair mismatch at the 3'-end of the nick junction, along with Com3'F to the complementary strand (Glg). The nicked DNA duplex substrates were formed by denaturing DNA probes at 94°C for 2 min followed by re-annealing at 65°C for 2 min in ligation buffer. The sequences of the oligodeoxyribonucleotides are listed in Scheme 1 (p represents the 5'-phosphate group).

Ligation mixtures (20 µl) containing the indicated amount of DNA ligase and matched or mismatched substrates in ligation buffer (20 mM Tris-HCl, pH 7.6 at room temperature, 10 mM MgCl₂, 100 mM KCl, 10 mM DTT, 1 mM NAD⁺ and 20 µg/ml BSA) were incubated at 65°C for a predetermined time. Reactions were terminated by the addition of an equal volume of stop solution (50 mM EDTA, 80% formamide and 1% Blue Dextran). Samples (5 µl) were electrophoresed through an 8 M urea-10% polyacrylamide GeneScan gel according to the instruction manual (model 373 or 377; Perkin Elmer). The unreacted substrates were represented by the 30mer Com3F and products were represented by a ligated 63mer in the case of the matched substrate. Both the remaining substrates and

Scheme 1.

Figure 1. Motif/domain organization of a putative NAD⁺-dependent DNA ligase from *Aquifex aeolicus* (ligA, GenBank accession no. AAC06838). The putative adenylation site is marked (inverted triangle). HhH, helix–hairpin–helix motif; BRCT, BRCT domain previously identified in BRCA1 and other proteins involved in cell cycle checkpoints and DNA repair (39).

ligation products were quantified using GeneScan analysis software 672 (v.2.0; Perkin Elmer).

Steady-state kinetics

Steady-state kinetic constants were determined by measuring initial rates of the ligation reaction at a given substrate concentration (nicked DNA duplex substrate concentration ranging from 25 to 400 nM) and a given ligase concentration (0.25 nM) in a 100 μ l reaction containing 20 mM Tris–HCl (pH 8.0), 150 mM KCl, 10 mM DTT, 1 mM NAD⁺, 20 μ g/ml BSA and 5 mM Mg²⁺ at 65°C. A 5 μ l aliquot was removed at 0, 2, 4, 6 and 8 min and mixed with 5 μ l of stop solution. The remaining substrate was separated from ligation product by GeneScan gel as described above. Initial rates of the ligation reactions were calculated from the generation of ligation product over time. The K_m and k_{cat} values were determined using the computer software Ultrafit (Biosoft, Ferguson, MO).

RESULTS

Cloning, expression and purification of DNA ligase from *Aquifex aeolicus*

The putative Aae NAD⁺-dependent ligase (ligA, GenBank accession no. AAC06838) consists of 720 amino acids, which is within the range of bacterial ligases (~600–800 amino acids). The Aae NAD⁺-dependent ligase shares 43% amino acid sequence identity with the Tth ligase and 37% with the *E. coli* ligase. This gene possesses all the typical motifs found in other bacterial NAD⁺-dependent ligases (Fig. 1): a KxxG adenylation motif in the N-terminal region, a zinc binding motif (C-X₂-C-X₁₂₋₁₅-C-X₄₋₅-C), four helix–hairpin–helix (HhH) motifs and a BRCT domain in the C-terminal region. To biochemically characterize this ligase from the hyperthermophile, we PCR amplified the full-length Aae ligase gene from *A. aeolicus* genomic DNA (a kind gift from Dr Ron Swanson) and cloned it in *E. coli*. The Aae NAD⁺-dependent ligase protein, which was expressed to approximately 10% of total cellular proteins, was separated from most *E. coli* host proteins by thermal denaturation (70°C for 15 min) and further purified by Cibacron blue-based affinity chromatography (Pharmacia). The resulting preparation was homogeneous as judged by Coomassie brilliant blue staining (data not shown).

Matched substrate ligation kinetics

A synthetic nicked substrate consisting of three oligonucleotides was employed for the Aae NAD⁺-dependent ligase activity assay (see Materials and Methods for sequence). A successful ligation reaction generates a fluorescently labeled product of 63 nt.

Figure 2. Steady-state kinetics of the Aae ligase. (A) Time course of ligation reactions containing 0.25 nM Aae ligase and the following substrate concentrations: 25 (closed square), 50 (open circle), 100 (triangle), 200 (diamond), 300 (open square) and 400 nM (closed circle). (B) The velocity as a function of substrate concentration follows Michaelis–Menten kinetics. (Inset) Double reciprocal plot of $1/V$ as a function of $1/[S]$. The sequence of the nicked substrate is shown in Materials and Methods. Kinetic constants were calculated as the average of multiple independent measurements and the experimental error is estimated to be 20%.

Plots of accumulation of the ligation product versus time showed excellent linearity, allowing us to accurately determine the kinetic constants (Fig. 2A). The enzyme exhibited typical Michaelis–Menton steady-state kinetics with a K_m of 76 nM and a k_{cat} of 2 min^{−1} for this nicked substrate (Fig. 2). The K_m is similar to that of *E. coli* (18) and *Thermus* ligases (7) and ~10-fold higher than that of the *Vaccinia* virus ATP-dependent ligase

(19). The k_{cat} value is 14-fold lower than that of the *E. coli* ligase (18) and ~25-fold lower than that of *Thermus* ligases (7).

Effects of external factors on the ligation reaction

The optimal pH for the Aae ligase was 8.0, with greater than 70% activity observed between pH 7.8 and 8.8 (Table 1). This value is similar to the pH optimum we previously determined for *Thermus* ligases (7). However, unlike *Thermus* ligases, Aae ligase was active over a relatively broad salt concentration range. The optimum KCl concentration was 150 mM, with greater than 70% activity observed between 50 and 200 mM. The enzyme required 0.5 mM NAD^+ for optimal activity (Table 1).

Table 1. Effects of external factors on ligation activity of Aae ligase

	pH ^a	KCl (mM) ^b	NAD^+ (mM) ^c	Mg^{2+} (mM) ^d	Mn^{2+} (mM) ^d
Optimum	8.0	150	0.5	5.0	5.0
Range (70% activity)	7.8–8.8	50–200	0.2–1.5	2.0–8.0	1.0–10

^apH effect. Reactions were performed in a 20 μl mixture containing 200 nM nicked duplex substrate, 62.5 pM Aae ligase, 20 mM Tris–HCl (pH values were determined at room temperature), 10 mM MgCl_2 , 100 mM KCl, 10 mM DTT, 1 mM NAD^+ and 20 $\mu\text{g/ml}$ BSA at 65°C for 10 min.

^bSalt effect. Reactions were performed in a similar 20 μl mixture with 20 mM Tris–HCl (pH 8.0) and the indicated amount of KCl at 65°C for 10 min.

^c NAD^+ effect. Aae ligation reactions were performed in a similar 20 μl mixture with 20 mM Tris–HCl (pH 8.0), 5 mM MgCl_2 , 150 mM KCl and the indicated concentration of NAD^+ at 65°C for 10 min.

^dReaction mixtures (20 μl) containing 20 nM nicked duplex substrate, 5 nM Aae ligase, 20 mM Tris–HCl (pH 8.0), 150 mM KCl, 10 mM DTT, 1 mM NAD^+ , 20 $\mu\text{g/ml}$ BSA and various concentrations of Mg^{2+} or Mn^{2+} were incubated at 65°C for 10 min.

The enzyme was active with either Mg^{2+} or Mn^{2+} as the metal cofactor (Fig. 3). Ca^{2+} supported formation of the DNA–adenylate intermediate (AppDNA), but very little of this was converted into ligation product. In the presence of Ni^{2+} , a limited amount of substrate was converted into AppDNA, but this in turn could not be ligated. The enzyme could not use Co^{2+} , Cu^{2+} or Zn^{2+} as the metal cofactor for formation of either AppDNA or the ligation product. In the single time point experiment, the enzyme appeared to be more active with Mn^{2+} than with Mg^{2+} (Fig. 3). The higher yield using Mn^{2+} as the metal cofactor than with Mg^{2+} is probably due to a higher nick closure rate, resulting in less accumulation of AppDNA (Fig. 3). To verify this observation, we performed a time course study (Fig. 4A and B). The Aae NAD^+ -dependent ligase was consistently more active with Mn^{2+} throughout the 20 min period. The enzyme also accumulated more AppDNA with Mg^{2+} as the metal cofactor. Furthermore, this enzyme was more active with Mn^{2+} at all metal concentrations tested (Fig. 4C and D), reaffirming that the Aae NAD^+ -dependent ligase is indeed more active when Mn^{2+} is used as the metal cofactor.

Ligation of matched substrates

Previous studies have provided an estimation of reaction rates of matched substrate ligations for bacterial NAD^+ -dependent ligases (20), but the actual ligation rates of related matched substrates have not been determined. In this study we measured the individual ligation rates by varying matched base

Figure 3. Divalent cation dependence of Aae ligase activity. (A) Ligation reactions with different divalent ions as the metal cofactor. Reaction mixtures (20 μl) containing 10 nM nicked duplex substrate, 10 nM Aae ligase, 20 mM Tris–HCl (pH 8.0), 150 mM KCl, 10 mM DTT, 1 mM NAD^+ , 20 $\mu\text{g/ml}$ BSA and 5 mM indicated divalent cation were incubated at 65°C for 20 min with either Mg^{2+} or Mn^{2+} as the metal cofactor. The incubation time with other metal ions was extended to 90 min due to slow ligation rates. (B) Chromatogram of a representative GeneScan gel illustrating formation of the ligation product and AppDNA. Co^{2+} may have caused precipitation of the DNA substrate, resulting in its loss.

pairs on both the 3'- and 5'-sides. The alterations of base pairs occurred within the same flanking sequence (Fig. 5A and B). In the first set of experiments, shown in Figure 5C, the base pair on the 5'-side of the nick was kept as A/T while that on the 3'-side of the nick was varied. The ligation rates of the matched substrates ranged from 12 fmol/min for T/A, to 19 fmol/min for G/C, to 37 fmol/min for C/G, to 55 fmol/min for A/T, resulting in a more than 4-fold difference. On the other hand, the rates obtained when the base pair on the 5'-side of the nick was varied were within 2-fold of one another (Fig. 5D). In addition, all reactions generated AppDNA, suggesting that the nick closure step is rate limiting for these matched substrates. The ligation rates in general followed a similar order of $\text{T/A} \leq \text{G/C} < \text{C/G} < \text{A/T}$ for the matched substrates on both the 3'- and 5'-side of the nick.

Ligation of mismatched substrates

A mismatched base pair is structurally more distinct from a matched base pair. Therefore, the ligation rate may be more strongly affected by a mismatch at the nick. We evaluated the ability of the Aae NAD^+ -dependent ligase to discriminate all possible mismatches on both the 3'- and 5'-sides of the nick junction. Initial experiments showed that some of the ligation rates of mismatched substrates were too slow to be measured

Figure 4. Ligation profiles of a C/G matched substrate in the presence of Mg^{2+} or Mn^{2+} . Square, ligation product; circle, AppDNA. (A) Time course of the ligation reaction in the presence of Mg^{2+} . (B) Time course of the ligation reaction in the presence of Mn^{2+} . Reaction mixtures (100 μ l) containing 20 nM nicked duplex substrate, 5 nM Aae ligase, 20 mM Tris-HCl (pH 8.0), 150 mM KCl, 10 mM DTT, 1 mM NAD⁺, 20 μ g/ml BSA and 5 mM Mg^{2+} or Mn^{2+} were incubated at 65°C. Aliquots (5 μ l) were removed at the indicated times and reactions stopped by adding an equal volume of stop solution. (C) Ligase activity in the presence of various concentrations of $MgCl_2$. (D) Ligation reaction in the presence of various concentrations of $MnCl_2$. Reactions were performed with the indicated concentrations of Mg^{2+} or Mn^{2+} at 65°C for 2 min.

accurately. We therefore extended the incubation time to 3 h to record the product yields from the low level ligation reactions. As shown in Figure 5E, even after extended incubation the enzyme was not able to ligate most of the 3' mismatched substrates. The three observable mismatch ligations on the 3'-side of the nick were with C/A, T/G, G/T. Significant accumulation of AppDNA was observed only with the ligatable T/G and G/T mismatches on the 3'-side of the nick. Other mismatches either did not generate or generated minimal amounts of the intermediate (Fig. 5E).

The enzyme is more tolerant of 5' mismatched substrates (Fig. 5F). Strand ligation was detected in all mismatched substrates except for T/C. The most significant ligations occurred with G/G, C/A, T/T and T/G mismatches. In comparison, C/A, T/T, T/G, A/C, A/A, G/A and G/T mismatches are ligated by the Tth NAD⁺-dependent ligase (20). Since identical mismatched substrates were used for both the previous Tth ligase study and the current Aae ligase study, the differences in mismatch ligation profile on the 5'-side of the nick junction could not be due to flanking sequence variations. Rather, they hint at subtle differences in protein structures for recognizing different mismatched base pairs.

Ligation of gapped or inserted DNA duplex substrates

Gapped substrates eliminate base pair complementarity on one side of the nick. To determine the influence of the loss of base pair complementarity on ligation rates, 1 or 2 nt from the 3'-OH end of oligonucleotide LP3'C (see Materials and Methods

for sequence) were deleted to form a 'one-gap' or 'two-gap' substrate. In both cases, the enzyme was incapable of catalyzing the formation of ligation product and accumulation of AppDNA was minimal (data not shown). Similar results were obtained in previous studies on the *Thermus* ligase (7).

A 3' insertion also inhibited the ligation reaction, further emphasizing the importance of 3' complementarity in binding and/or catalysis (Fig. 6A). The small amount of ligation of the 3' A insertion was likely due to base pairing of this inserted A base with T in the template (Fig. 6A). This resulted in a 5' A insertion structure which can be ligated by the ligase, as demonstrated in Figure 6B. In addition, other 5' base insertions (G, T or C) were also ligated, suggesting that activation of the catalytic center is less dependent on the base pair complementarity on the 5'-side of the nick junction.

Ligation with Mn^{2+} as the metal cofactor

The high ligation activity observed with Mn^{2+} as the metal cofactor has prompted us to compare the ligation rates with Mn^{2+} to those with Mg^{2+} . The enzyme consistently exhibited higher ligation rates with Mn^{2+} for all three substrates tested: 1.2-fold for C/G matched, 9.3-fold for T/G mismatched, 62-fold for T/G mismatched on the penultimate 3'-side (Table 2). Apparently, the enzyme was more capable of ligating mismatched substrates when Mn^{2+} was used as the metal cofactor. This is in keeping with the notion that Mn^{2+} relaxes the specificity of NAD⁺-dependent DNA ligases (7). The low ligation rate of a T/G mismatch on the penultimate 3'-side with Mg^{2+} suggests

Figure 5. Ligation of matched and mismatched substrates by the Aae ligase. (A) Schematic illustration of the substrates with different nucleotides on the 3'-side of the nick. (B) Schematic illustration of the substrates with different nucleotides on the 5'-side of the nick. (C) Ligation rates of 3' matched substrates. (D) Ligation rates of 5' matched substrates. Reaction mixtures (50 μ l) containing 12.5 nM nicked duplex substrate, 12.5 nM Aae ligase, 20 mM Tris-HCl (pH 8.0), 150 mM KCl, 10 mM DTT, 1 mM NAD⁺, 20 μ g/ml BSA and 5 mM Mg²⁺ were incubated at 65°C. Five microliter aliquots from a 50 μ l reaction mixture were removed at 0, 10, 20, 30, 40 and 50 s for reactions containing matched substrates and at 3 h for reactions containing mismatched substrates. Samples (5 μ l) were electrophoresed through an 8 M urea–10% polyacrylamide gel as described. Fluorescently labeled ligation products were analyzed and quantified using GeneScan 672 v.2.0 software. The results were plotted using DeltaGraph Pro3 software (DeltaPoint Inc., Monterey, CA). The initial rates were determined as the slopes of the linear curves with the *x*-axis as the time and the *y*-axis as the amount of ligation product generated. (E) Ligation of a 3' mismatch. —, no intermediate observed; ±, minimal intermediate observed; +, significant intermediate observed. (F) Ligation of a 5' mismatch.

that the enzyme is highly sensitive to a mismatched base pair even when it is located 1 bp away from the 3'-side of the nick junction.

C/G ligation with Mn²⁺ generates little AppDNA when the reaction is performed under optimal conditions (Figs 3 and 4). However, when a T/G mismatch was introduced at the 3' nick junction or on the penultimate 3'-side, accumulation of the intermediate was significant (data not shown), suggesting that the rate of nick closure was greatly reduced as compared with the rate of substrate adenylation.

Table 2. Aae ligase ligation rates with Mn²⁺^a

Metal	<i>V</i> ₀ (C/G) (fmol/min)	<i>V</i> ₀ (T/G) (fmol/min)	<i>V</i> ₀ (T/G)* ^b (fmol/min)
Mn ²⁺	1.0 × 10 ²	1.4 × 10 ⁻¹	1.6 × 10 ¹
Mg ²⁺	8.1 × 10 ¹	1.5 × 10 ⁻²	2.6 × 10 ⁻¹

^aThe initial rates were determined as the slopes of the linear curves with the *x*-axis as the time and the *y*-axis as the amount of ligation product generated. Results were calculated as the average of at least two experiments. A schematic illustration of matched and mismatched substrates is as follows:

C/G match at 3'-end	T/G mismatch at 3'-end	T/G mismatch at penultimate 3'-end
—GTC p—F	—GTT p—F	—GTC p—F
—CAG—	—CAG—	—CGG—

^b*V*₀ (T/G)* indicates the initial rate of ligation when the substrate contains a T/G mismatch at the penultimate 3'-end.

DISCUSSION

Our previous studies of NAD⁺-dependent ligases have centered on thermostable ligases from *Thermus* spp. (7,8). The availability of whole genome sequence data provides unprecedented opportunities to perform comparative studies of NAD⁺-dependent ligases from other sources. At least two DNA ligases have been identified from the *A. aeolicus* genome, however, their biochemical activities are unknown. Given the high rate of gene transfer among thermophilic archaea and bacteria (21) and based on our sequence analysis, one of the ligase genes (ATP-dependent ligase) is likely from the archaea. The NAD⁺-dependent ligase is therefore probably the endogenous ligase for this organism. Sequence comparison suggests that this putative ligase has all the structural elements found in other bacterial ligases. This work describes some of the biochemical properties of the Aae NAD⁺-dependent ligase.

Metal cofactors and ligation activity

Ligases are metal-dependent enzymes. As with our previously characterized *Thermus* ligases (7), the Aae ligase can use either Mg²⁺ or Mn²⁺ as the metal cofactor (Fig. 3). However, a few differences are noted. First, although both ligases are capable of coordinating hard metals such as Mg²⁺ or Mn²⁺ throughout the catalytic cycle, the metal-binding pocket of the Aae ligase appears to be more adapted to utilizing Mn²⁺ as the metal

Figure 6. Ligation of 1 nt insert substrates by Aae ligase. (A) Insertion on the 3'-side of the nick. (B) Insertion on the 5'-side of the nick. Reactions were performed in 20 μ l mixtures containing 12.5 nM nicked duplex substrate and 12.5 nM Aae ligase in the reaction buffer as described in Figure 5 at 65°C for 4 h.

cofactor, as demonstrated by the consistently higher ligation rates and the minimal accumulation of AppDNA species (Fig. 4 and Table 2). Since the experiments were performed under multiple turnover conditions, it is not clear whether Mn^{2+} accelerated catalysis of nick closure or product dissociation. It is also unclear whether the high activity with Mn^{2+} is unique to the Aae NAD⁺-dependent ligase or is shared by NAD⁺-dependent ligases from other hyperthermophiles. Second, Aae ligase is less active in converting the intermediate into ligation product than *Thermus* ligases in the presence of Ca^{2+} (7). Third, nick closure or subsequent product dissociation is the rate limiting step for the Aae ligase in the presence of Mg^{2+} , which may account for the large k_{cat} difference between the two enzymes (Figs 2 and 3).

The accumulation of AppDNA species with Mg^{2+} is analogous to an observation made on *Chlorella* virus ATP-dependent ligase using a 1 nt gapped substrate (22), in which the ATP-dependent ligase failed to complete conversion of AppDNA to

ligation product. It is suggested that the inability of the *Chlorella* virus ligase to remain bound to the gapped AppDNA accounts for accumulation of AppDNA. A similar mechanism may explain the constant yield of AppDNA generated by the Aae NAD⁺-dependent ligase (Fig. 4A). However, the virtually complete conversion of AppDNA to ligation product by this NAD⁺-dependent ligase in the presence of Mn^{2+} may suggest that the ligase–AppDNA– Mn^{2+} tertiary complex is more stable than the ligase–AppDNA– Mg^{2+} complex.

It is noteworthy that at low Mg^{2+} or Mn^{2+} concentrations more AppDNA accumulated (Fig. 4C and D, open circles). These results suggest that at the substrate adenylation step (step 2 of the ligation reaction), the metal-binding site may have a higher affinity for Mg^{2+} or Mn^{2+} . Once the AMP moiety is transferred to the substrate, the enzyme may experience a conformational change which lowers the affinity of the metal-binding site for Mg^{2+} or Mn^{2+} . As such, the metal-binding site may not be filled at all times, resulting in a low nick closure rate at low Mg^{2+} or Mn^{2+} concentrations. Similar reasoning may help explain formation of only AppDNA with Ca^{2+} and Ni^{2+} . At the substrate adenylation step, the metal-binding site is able to accommodate Ca^{2+} or Ni^{2+} , thus the AppDNA species is observed. However, the subsequent conformational change may alter the metal-binding site geometry such that it becomes difficult to accommodate Ca^{2+} or Ni^{2+} and causes dissociation of the ligase–AppDNA– M^{2+} complex, thereby limiting conversion of the intermediate into ligation products.

Nick recognition and ligation activity

This work provides complete rate measurements for related matched substrate ligation on both the 3'- and 5'-side of the nick junction for an NAD⁺-dependent ligase (Fig. 5). As expected, the matched substrate ligation rates vary with different base pairs. However, we observed a similar trend of ligation activity for the four different matched base pairs on both the 3'- and 5'-sides. The structural basis underlying this observation is not known. This regularity of matched substrate ligation on the 3'- and 5'-sides may suggest that similar structural elements are used for DNA recognition on both sides of the nick in some steps of the ligation cycle. The C-terminal fragment of Bst ligase is able to bind a nicked DNA substrate (4). Two kinds of motifs at the C-terminus, the four HhH motifs and the zinc-binding motif (8,23–30), are likely candidates responsible for DNA binding and nick recognition.

The DNA–adenylate intermediate and ligation fidelity

We have examined accumulation of the AppDNA species and ligation activities for matched and mismatched substrates as well as substrates with a 1-bp gap or insertion. Matched substrates are ligated far more efficiently than other substrates, such as mismatches (Table 2). However, match ligation rates can differ as much as 4-fold (Fig. 5). These results suggest that the enzyme is not only sensitive to the structural differences in mismatched substrates on the 3'-side of the nick, but is also sensitive to differences in the matched base pairs.

Consistent with previous studies on *Thermus* ligases and ATP-dependent ligases (7,31–33), the Aae NAD⁺-dependent ligase is more sensitive to mismatched base pairs on the 3'-side of the nick (Fig. 5). This global similarity in mismatch ligation by NAD⁺- and ATP-dependent ligases is in keeping with the conserved nature of the catalytic elements in the N-terminal

part of the ligases. Yet individual mismatch ligation rates could vary substantially. For example, a purine–purine mismatch such as G/G on the 5′-side is a good substrate for the Aae ligase (Fig. 5F) and the *Vaccinia* virus ATP-dependent ligase (33), but not for the Tth ligase (20).

The accumulation of AppDNA species reveals a relationship between the second step of the ligation cycle (substrate adenylation) and ligation fidelity. Our observations suggest that when some mismatches are placed on the 3′-side, Aae NAD⁺-dependent ligase could not even transfer the AMP moiety from the adenylated enzyme to the 5′-phosphate of the nick (Fig. 5E), indicating that the substrate adenylation step plays an important role in controlling ligation fidelity. The contribution of the third step (nick closure) to ligation fidelity cannot be easily deduced, because this ligase accumulates AppDNA even during matched substrate ligation. However, the importance of the third step in controlling ligation fidelity has been demonstrated for the *Thermus* ligases (7). A 3′ G/T mismatch ligation reaction results in accumulation of the ligation product as well as AppDNA, whereas a match ligation reaction does not result in accumulation of AppDNA for *Thermus* ligase (7). Taken together, it is likely that for NAD⁺-dependent ligases both the second (substrate adenylation) and third (nick closure) steps play an important role in determining ligation specificity. This conclusion is consistent with studies on *Chlorella* ligase (34), human ligases I and II (35) and *E. coli* NAD⁺-dependent ligase using 3′-dideoxy substrates (36).

The ability of DNA ligases to ligate gapped or inserted substrates has been assessed (7,33,37,38; this work). ATP ligases from *Hemophilus influenzae*, *Chlorella* virus and *Vaccinia* virus are either able to form ligation product or to accumulate significant amounts of AppDNA (33,37,38). On the other hand, NAD⁺-dependent ligases failed to support either nick closure or substrate adenylation. These results suggest that for bacterial NAD⁺-dependent ligases effective transfer of the AMP moiety from the adenylated ligase to the 5′-phosphate of the substrate is not possible with a 1 bp gap. The complete loss of 3′ or 5′ complementarity, such as by insertion of a 1 bp gap, is apparently sufficient to abolish substrate adenylation. This observation, combined with the lack of AMP transfer to the 5′-phosphate with mismatches on the 3′-side, as described above, emphasizes the importance of 3′ recognition at the second step of the ligation cycle for NAD⁺-dependent ligases.

In summary, this work has revealed several unique properties of the NAD⁺-dependent DNA ligase from the hyperthermophilic bacterium *A. aeolicus*. The low catalytic efficiency is likely a reflection of a slow rate of nick closure, which results in accumulation of AppDNA with Mg²⁺ as the metal cofactor. The thorough investigation of ligation activities of matched, mismatched and other substrates has provided new insights into the steps that control nick recognition.

ACKNOWLEDGEMENTS

We thank Drs Neal Lue and Jianmin Huang for discussions and Jing Lu for technical assistance. We are grateful to Dr Ron Swanson of Diversa Corp. for providing *Aquifex aeolicus* genomic DNA and discussions. This work was supported by

grants from the National Institutes of Health (GM-41337-07 and PO1-CA65930-02-04).

REFERENCES

- Lehman, I.R. (1974) *Science*, **186**, 790–797.
- Singleton, M.R., Hakansson, K., Timson, D.J. and Wigley, D.B. (1999) *Structure*, **7**, 35–42.
- Subramanya, H.S., Doherty, A.J., Ashford, S.R. and Wigley, D.B. (1996) *Cell*, **85**, 607–615.
- Timson, D.J. and Wigley, D.B. (1999) *J. Mol. Biol.*, **285**, 73–83.
- Barany, F. and Gelfand, D. (1991) *Gene*, **109**, 1–11.
- Jonsson, Z.O., Thorbjarnardottir, S.H., Eggertsson, G. and Palsdottir, A. (1994) *Gene*, **151**, 177–180.
- Tong, J., Cao, W. and Barany, F. (1999) *Nucleic Acids Res.*, **27**, 788–794.
- Luo, J. and Barany, F. (1996) *Nucleic Acids Res.*, **24**, 3079–3085.
- Barany, F. (1991) *PCR Methods Appl.*, **1**, 5–16.
- Barany, F. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 189–193.
- Day, D.J., Speiser, P.W., White, P.C. and Barany, F. (1995) *Genomics*, **29**, 152–162.
- Khanna, M., Park, P., Zirvi, M., Cao, W., Picon, A., Day, J., Paty, P. and Barany, F. (1999) *Oncogene*, **18**, 27–38.
- Deckert, G., Warren, P.V., Gaasterland, T., Young, W.G., Lenox, A.L., Graham, D.E., Overbeek, R., Snead, M.A., Keller, M., Aujay, M., Huber, R., Feldman, R.A., Short, J.M., Olsen, G.J. and Swanson, R.V. (1998) *Nature*, **392**, 353–358.
- Pitulle, C., Yang, Y., Marchiani, M., Moore, E.R., Siefert, J.L., Aragno, M., Jurtshuk, P., Jr and Fox, G.E. (1994) *Int. J. Syst. Bacteriol.*, **44**, 620–626.
- Burggraf, S., Olsen, G.J., Stetter, K.O. and Woese, C.R. (1992) *Syst. Appl. Microbiol.*, **15**, 353–356.
- Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Cao, W., Lu, J. and Barany, F. (1997) *Gene*, **197**, 205–214.
- Modrich, P., Anraku, Y. and Lehman, I.R. (1973) *J. Biol. Chem.*, **248**, 7495–7501.
- Sekiguchi, J. and Shuman, S. (1997) *Nucleic Acids Res.*, **25**, 727–734.
- Luo, J., Bergstrom, D.E. and Barany, F. (1996) *Nucleic Acids Res.*, **24**, 3071–3078.
- Feng, D.F., Cho, G. and Doolittle, R.F. (1997) *Proc. Natl Acad. Sci. USA*, **94**, 13028–13033.
- Skisanda, V. and Shuman, S. (1998) *Nucleic Acids Res.*, **26**, 4618–4625.
- Tchou, J., Michaels, M.L., Miller, J.H. and Grollman, A.P. (1993) *J. Biol. Chem.*, **268**, 26738–26744.
- O'Connor, T.R., Graves, R.J., de Murcia, G., Castaing, B. and Laval, J. (1993) *J. Biol. Chem.*, **268**, 9063–9070.
- Gradwohl, G., Menissier de Murcia, J.M., Molinete, M., Simonin, F., Koken, M., Hoeijmakers, J.H. and de Murcia, G. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 2990–2994.
- Navaratnam, S., Myles, G.M., Strange, R.W. and Sancar, A. (1989) *J. Biol. Chem.*, **264**, 16067–16071.
- Skisanda, V., Schwer, B., Ho, C.K. and Shuman, S. (1999) *Nucleic Acids Res.*, **27**, 3953–3963.
- Doherty, A.J., Serpell, L.C. and Ponting, C.P. (1996) *Nucleic Acids Res.*, **24**, 2488–2497.
- Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H. and Kraut, J. (1994) *Science*, **264**, 1891–1903.
- Sawaya, M.R., Pelletier, H., Kumar, A., Wilson, S.H. and Kraut, J. (1994) *Science*, **264**, 1930–1935.
- Housby, J.N. and Southern, E.M. (1998) *Nucleic Acids Res.*, **26**, 4259–4266.
- Skisanda, V. and Shuman, S. (1998) *Nucleic Acids Res.*, **26**, 3536–3541.
- Shuman, S. (1995) *Biochemistry*, **34**, 16138–16147.
- Odell, M. and Shuman, S. (1999) *J. Biol. Chem.*, **274**, 14032–14039.
- Yang, S.W. and Chan, J.Y. (1992) *J. Biol. Chem.*, **267**, 8117–8122.
- Modrich, P. and Lehman, I.R. (1973) *J. Biol. Chem.*, **248**, 7502–7511.
- Ho, C.K., Van Etten, J.L. and Shuman, S. (1997) *J. Virol.*, **71**, 1931–1937.
- Cheng, C. and Shuman, S. (1997) *Nucleic Acids Res.*, **25**, 1369–1374.
- Bork, P., Hofmann, K., Bucher, P., Neuwald, A.F., Altschul, S.F. and Koonin, E.V. (1997) *FASEB J.*, **11**, 68–76.