Extremophiles (2001) 5:161–168
Digital Object Identifier (DOI) 10.1007/s007920100187

ORIGINAL PAPER

Jae-Hwan Lim · Juhyun Choi · Soo-Jin Han · Sung Hou Kim
Hye-Zin Hwang · Dong-Kyu Jin · Byung-Yoon Ahn
Ye Sun Han

Molecular cloning and characterization of thermostable DNA ligase from *Aquifex pyrophilus*, a hyperthermophilic bacterium

Received: June 20, 2000 / Accepted: February 16, 2001 / Published online: May 22, 2001

Abstract A DNA ligase gene from the hyperthermophilic bacterium *Aquifex pyrophilus* (*Ap*) was cloned and sequenced. An open reading frame of 2,157 bp that codes for a 82-kDa protein showed 40%–60% homology with a series of NAD^+-dependent DNA ligases from different organisms. The recombinant enzyme *Ap* DNA ligase expressed in *Escherichia coli* was purified to homogeneity and characterized. The activity of *Ap* DNA ligase gradually increased in proportion to the concentration of monovalent salt up to 200 mM NaCl, 150 mM KCl, 200 mM NH_4Cl, and 350 mM potassium glutamate. The optimum temperature and pH of *Ap* DNA ligase were greater than 65°C and 8.0–8.6, respectively, for nick-closing activity. More than 75% of the ligation activity was retained after incubation at 95°C for 60 min, whereas the half-lives of *Thermus aquaticus* and *Escherichia coli* DNA ligases at 95°C were ≤15 min and 5 min, respectively. Thermostable *Ap* DNA ligase was applied to repeat expansion detection (RED) and could be a useful enzyme in DNA diagnostics.

Key words DNA ligase · *Aquifex pyrophilus* · Hyperthermophile · Thermostability · Repeat expansion detection

Communicated by J. Wiegel

J.-H. Lim · J. Choi · S.-J. Han · S.H. Kim · Y.S. Han (✉)
Structural Biology Research Center, Korea Institute of Science and Technology, 39-1 Hawallkok-dong, Sungbuk-ku, Seoul, Korea
Tel. +82-2958-5933; Fax +82-2958-5939
e-mail: yshan2@kist.re.kr

B.-Y. Ahn
Graduate School of Biotechnology, Korea University, Seoul, Korea

S.H. Kim
Department of Chemistry and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA, USA

H.-Z. Hwang
Department of Biology, Ewha Woman's University, Seoul, Korea

D.-K. Jin
Department of Pediatrics, Samsung Medical Center, Sungkyunkwan University, Seoul, Korea

Introduction

DNA ligase is an essential enzyme for a number of important cellular processes including repair, replication, and recombination of DNA (Lehman 1974). It catalyzes the formation of phosphodiester bonds at single-strand breaks in double-stranded DNA. DNA ligases from bacteriophages, Archaea, eukaryotes, and viruses use ATP as a cofactor (Tomkinson et al. 1991; Kletzin 1992), whereas eubacterial DNA ligases require NAD^+ (Shark and Conway. 1992). Interestingly, an ATP-dependent DNA ligase encoded by the genome of the bacterium *Haemophilus influenza* has been reported (Cheng and Shuman 1997). As the entire genome sequences of more species are determined, the difference between DNA ligases from prokaryotes and eukaryotes becomes ambiguous. For example, both types of DNA ligases were shown to be encoded in the genome of *Aquifex aeolicus* (Deckert et al. 1998).

Molecular masses of DNA ligases range from 103 kDa for the human DNA ligase I to 41 kDa for the bacteriophage T7 enzyme (Dunn and Studier 1981; Barnes et al. 1990). In the mammalian enzyme, the large molecular weight results from the presence of a domain in the protein that contains the signal for localization of the protein in the cell or the site for specific protein–protein interaction (Robins and Lindahl 1996). In the case of NAD^+-dependent DNA ligases, however, the enzymes are about 70–90 kDa in size irrespective of their origin.

Likewise, very little sequence similarity exists between the ATP-dependent DNA ligases and the bacterial DNA ligases, except for a short stretch that bears the AMP-binding site (Thorbjarnardottir et al. 1995), in contrast to the high level of sequence homology conserved within each class of DNA ligases. These data imply that the two classes of DNA ligases are not closely related phylogenetically in spite of their functional similarity.

Ongoing research in DNA diagnostics is providing automated, rapid, and inexpensive analysis for DNA sequences associated with genetic, malignant, and infectious diseases. Thermostable DNA ligases are used in assays that amplify

DNA and distinguish single-base substitution. The important applications of this technology include the detection of single-base mutations or specific nucleotide sequences by ligase chain reaction (LCR) and the use of repeat expansion detection (RED) for detection of genetic disease caused by trinucleotide repeats (Barany 1991; Shalling et al. 1993).

The sequences of four thermophilic DNA ligases, from *Thermus thermophilus* (Barany and Gelfand 1991; Lauer et al. 1991), *Rhodothermus marinus* (Thorbjarnardottir et al. 1995), *Thermus scotoductus* (Jonsson et al. 1994), and *Bacillus stearothermophilus* (Brannigan et al. 1999), have been published. Recently, the crystal structure of the NAD^+-dependent DNA ligase from *Thermus filiformis*, a 667-residue multidomain protein, has been determined (Lee et al. 2000). A unique circular arrangement of its four distinct domains, i.e., domain 1 (adenylation), domain 2 (oligonucleotide/oligosaccharide binding-fold, OB-fold), domain 3 (zinc finger and HhH motif), and domain 4 (BRCA [breast cancer]-1 C-terminus, BRCT), leads to a hole large enough to hold a double-stranded DNA. In addition, two hyperthermostable DNA ligases, named *Pfu* ligase (Stratagene, La Jolla, CA, USA) and Ampligase (Epicentre Technology, Madison, WI, USA) are commercially available.

In this study, we cloned and overexpressed the *Aquifex pyrophilus* DNA ligase (hereafter named *Ap* DNA ligase) in *E. coli*. *A. pyrophilus* (Huber et al. 1992) is a hyperthermophilic marine bacterium that grows at temperatures between 67° and 95°C, with an optimum growing temperature of 85°C. *A. pyrophilus* was discovered from hot marine sediments (depth, 106 m) and represents the deepest branching with the Bacteria. We report here the effects of temperature on *Ap* DNA ligase activity and thermostability of the recombinant enzyme.

Materials and methods

Cloning of the DNA ligase gene

Construction of a plasmid library and a genomic λ library of *A. pyrophilus* was reported previously (Lim et al. 1997). One of the recombinant plasmids contained sequences highly homologous to DNA ligases, which were identified by the BLAST (Basic Local Alignment Search Tool) programs. The recombinant plasmid, which is 0.7 kDa in size, was used to prepare the probe for cloning the DNA ligase gene. Polymerase chain reaction (PCR) amplification was performed to make the probe for the DNA ligase gene using two universal primers (T7 and T3 promoter primers). A genomic library was plated and the plaque hybridization screening was done using the enhanced chemiluminescence labeling- (ECL) directed system (Amersham, Piscataway, NJ, USA) with the probe mentioned earlier. Positive plaques were selected and confirmed by secondary screening. Phage DNA of the isolated single plaque was prepared using the Quiagen Lambda Midi Kit (Hilden, Germany).

The full (2,157-bp) sequence coding for the DNA ligase was determined by sequential sequencing with synthetic primers and phage DNA containing the fragment of the DNA ligase gene as a template using an ABI 373 DNA automated sequencer. All the oligonucleotide primers were synthesized by BioSynthesis (Lewisville, TX, USA). For the expression of a recombinant *Ap* DNA ligase, the DNA ligase gene was amplified by direct PCR of the *A. pyrophilus* genomic DNA with two synthetic oligonucleotides as primers. The upstream primer, LIG-F, contained an *Nde*I restriction site, which has the translational initiation site, ATG, and the downstream primer, LIG-R, contained a *Bam*HI restriction site (LIG-F: 5′-GCCTCACGTTCACATATGTTCACCC CCGAAAGGGAAAGG-3′, LIG-R: 5′-GCTAGGCATGTCGGATCCTTAAAATAGCC TTCCCATCTTAACCTC-3′). The nucleotide sequence of the PCR-amplified *Ap* DNA ligase gene was confirmed to be the same as that of the DNA ligase gene cloned from the genomic library.

The PCR product digested with *Nde*I and *Bam*HI was isolated and ligated into *Nde*I/*Bam*HI-digested pET3a vector (Novagen, Madison, WI, USA). The ligation mixture was transformed into *E. coli* BL21 (DE3) containing plasmid pSJS1240 (a gift from Dr. R. Kim, University of California at Berkeley, CA, USA), which carries the tRNA genes for arginine (AGA) and isoleucine (ATA). Plasmid pSJS1240 was spectinomycin resistant.

Overexpression and purification of recombinant protein

The *E. coli* cells harboring the plasmid containing the DNA ligase gene were grown at 37°C in Luria broth containing 100 μg/μl ampicillin and 50 μg/μl spectinomycin. The *Ap* DNA ligase protein was induced by the addition of 1 mM isopropyl β-D-thiogalactopyranoside to the culture at $OD_{600} \geq 0.8$, and incubation was continued overnight at 30°C. The cells were harvested by centrifugation for 10 min at 3,000 *g* and were resuspended in lysis buffer A (50 mM Tris-HCl, pH 7.5, 1 mM dithiothreitol [DTT], 10 mM $MgCl_2$, and 0.1 mM phenylmethylsulfonyl fluoride PMSF]). The resuspended cells were lysed using a French pressure cell (SLM Instruments, Rochester, NY, USA) at 12,000 psi. DNase I was added to the lysate at a final concentration of 20 μg/μl. The lysate was incubated for 30 min in an ice bath and then centrifuged for 40 min at 25,000 *g*. The soluble extract was pooled and heated at 80°C for 40 min.

After centrifugation at 25,000 *g* for 30 min, the supernatant was applied to a S-Sepharose Fast Flow (40 μl; Pharmacia, Uppsala, Sweden) column, previously equilibrated with buffer A without PMSF. The column was washed extensively and eluted with a linear gradient from 0 to 1 M NaCl in buffer A. DNA ligase was eluted using 0.4 M NaCl. The fraction was applied to a heparin Sepharose column (15 μl; Pharmacia), which was preequilibrated with buffer A. The protein was eluted with a linear gradient of 0–2 M NaCl in buffer A. DNA ligase eluted at 1 M NaCl was diluted with buffer A to reduce the salt concentration to 100 mM NaCl. The proteins were applied to a HiTrap-Blue (5 μl, Pharmacia) column equilibrated with buffer A containing 100 mM NaCl. The protein was eluted with a linear gradient of NaCl in buffer A. Collected fractions were concentrated to a final protein concentration of 10 μg/μl using an Amicon concentrator. Subsequently, the concentrated protein solution was

loaded on a Superdex-S200 gel filtration (Pharmacia) column equilibrated with buffer A containing 100 mM NaCl. The DNA ligase fractions were pooled and the purity of the protein sample was checked by electrophoresis using a 10% sodium dodecyl sulfate- (SDS) polyacrylamide gel. Protein concentration of the enzyme was determined using the Bio-Rad dye reagent with bovine serum albumin (BSA) as a standard. The zinc ion concentration of the enzyme was determined by atomic absorption spectrophotometry (Varian SpectrAA800; Varian, Palo Alto, CA, USA).

Preparation of a nicked DNA substrate

To check the nick-closing activity of the DNA ligase, three oligonucleotides were synthesized by BioSynthesis. The substrate used in the ligase assay was a DNA duplex containing a centrally placed nick. The sequence of the substrate was as follows: LA 1:5′-GGTAAAGCAATGGGCAAACAGGGAAGCTATG-3′, LA 2: 5′-GACATAAGAGGTCTCGGTGATGACCCAGTAAAGCT-3′, LA 3: 5′-GGAGCTTTACTGGGTCATCACCGAGACCTCTTATGTCCATAGCTTCCCTGTTTGCCCATTGCTTTACCCTC-3′.

First, 50 pmol of the gel-purified oligonucleotide LA-2 was radiolabeled by 100 μmCi [γ-^{32}P]ATP (3,000 Ci/mmol; Pharmacia-Amersham Biotech, Buckinghamshire, UK) and 50 U T4 polynucleotide kinase (Promega, Madison, WI, USA) for 60 min at 37°C. After incubation of the enzyme for 10 min at 70°C, the unincorporated radiolabeled ATP was removed by centrifugation through a Quick Spin Column (Boehringer Mannheim, Germany). The labeled oligonucleotide, LA-2, and the other oligonucleotide, LA-1, were annealed to the complementary 71-mer oligonucleotide LA-3 in annealing buffer (50 mM Tris-HCl, pH 7.5, 200 mM NaCl) by heating at 90°C for 2 min, followed by slowly cooling to room temperature. The molar ratio of the LA-1, LA-2, and LA-3 fractions in the hybridization mixture was 1:1.2:1.2, respectively. The melting temperature (T_m) of the annealed oligonucleotide duplex was analyzed by circular dichroism (CD) and differential scanning calorimeter (DSC). The oligonucleotide concentration was 2 μM for CD and 0.67 μM for DSC.

DNA ligase assay

DNA ligase activity of the recombinant protein was alternatively assayed with λDNA digested with *Hin*dIII. The ligation reaction was performed in a 10-μl reaction mixture containing 20 mM Tris-HCl, pH 7.5 or 8.0, 10 mM $MgCl_2$, 5 mM DTT, 5 mM NAD^+, 0.5 μg *Hin*dIII-digested λDNA, and 1.2 pmol of the recombinant enzyme. The reaction mixtures were incubated at various temperatures for 2 h and stopped by adding 5-μl stop/loading buffer (3×: 30% sucrose, 150 mM ethylenediaminetetraacetic acid [EDTA], 0.15% SDS, and 0.03% bromophenol blue), followed by heating at 90°C for 2 min. Each sample was subjected to electrophoresis in a 0.8% agarose gel at 60 V and stained with ethidium bromide.

The assay for nick-closing activity of DNA ligase from *A. pyrophilus* was performed according to previously published methods (Doherty et al. 1996; Shuman and Ru 1995). The annealed DNA substrate (500 fmol) was incubated in a 10-μl reaction mixture (50 mM Tris-HCl, pH 8, 10 mM $MgCl_2$, 10 mM DTT, 5 mM NAD^+, 100 mM KCl, 0.1% Triton X-100, and 1.2 pmol DNA ligase) at 50°C for 1 h unless otherwise stated. After incubation, the reactions were terminated by adding stop buffer (95% formamide, 1× TBE, 0.05% bromophenol blue, 0.05% xylene cyanol, and 0.2% SDS) followed by heating at 95°C for 5 min. The samples were chilled on ice and analyzed by electrophoresis on a 10% polyacrylamide gel containing 7 M urea in TBE buffer (90 mM Tris-borate, 2.5 mM EDTA) at constant power of 20 W. Gels were dried under vacuum, and ligation products were visualized by autoradiography. For quantitation of the product, the dried gel was scanned using a FUJIX BAS 2000 phosphoimager. To determine the heat resistance of *Ap* DNA ligase, the protein was incubated at 95°C in 50 mM Tris-HCl buffer, pH 7.5, 10 mM $MgCl_2$, and 10 mM DTT. At various intervals during heating, aliquots were taken, chilled for 10 min, and the remaining activity of the enzyme was subsequently measured.

Repeat expansion detection

The repeat expansion detection (RED) assay was performed by the standard method with some modifications (Lindblad et al. 1995). A reaction mixture (10 μl) consisting of *Ap* DNA ligase of various concentrations from 1.2 to 6.1 pmol, 1 μl 10× buffer (50 mM Tris-HCl, pH 8.4, 10 mM $MgCl_2$, 1 mM NAD^+, 50 mM KCl, 0.1% Triton X-100), 5 pmol ^{32}P-labeled oligonucleotide $(CTG)_{10}$, and 600 ng human genomic DNA, which bears a few hundred repeats of the CTG sequence, was prepared. All reactions were performed on a Geneamp PCR system 9600 (Perkin-Elmer Cetus, Norwalk, CT, USA) with 500 cycles consisting of 30 s at 76°C, followed by 94°C for 10 s. The ligated product was loaded onto 6% acrylamide gel and analyzed by electrophoresis at 15 W for 1.5 h. The dried gel was exposed to X-ray film and the ligated products were analyzed by their mobility.

Results and discussion

Sequence analysis of the *Ap* DNA ligase

Three positive phage recombinants were analyzed for identification of the *Ap* DNA ligase gene. The full DNA sequence was determined by sequential sequencing with synthetic primers and phage DNA containing the fragment of the DNA ligase gene as a template (GeneBank accession number AF 152998).

The amino acid sequence of *Ap* DNA ligase was compared, using the SEQSEE program (Wishart et al. 1994), with *Aquifex aeolicus* and *Thermus aquaticus* ligases, which have identities of 84% and 44% with that of *Aquifex pyrophilus* (Fig. 1). The *Ap* DNA ligase contains conserved

motifs of NAD+-dependent DNA ligase as reported previously (Thorbjarnardottir et al. 1995). The Lys-142 was suggested to be the AMP-binding residue that is a part of the active site of nucleotidyl transfer enzymes such as DNA, RNA, and tRNA ligases and mRNA capping enzymes (Shuman and Schwer 1995). Several pieces of evidence obtained by site-directed mutagenesis support this notion (Timson and Wigley 1999).

The sequence of the N-terminal 160 residues of *Ap* DNA ligase matched well with the critical residues of *E. coli* adenylate kinase for AMP/ATP binding (Berry et al. 1994) (see Fig. 1). The structure of *E. coli* adenylate kinase with bound AMP and AMPPNP revealed unambiguously the location of sites of AMP and ATP binding. The nucleoside monophosphate (NMP) kinase fold usually had been subdivided into the three domains: CORE, NMP bind, and Lid (Schulz et al. 1990). In our laboratory, we constructed an *Ap* DNA ligase mutant in which the N-terminal 90 residues of *Ap* DNA ligase were deleted. The mutant was found to have a significantly reduced amount of adenylation (unpublished

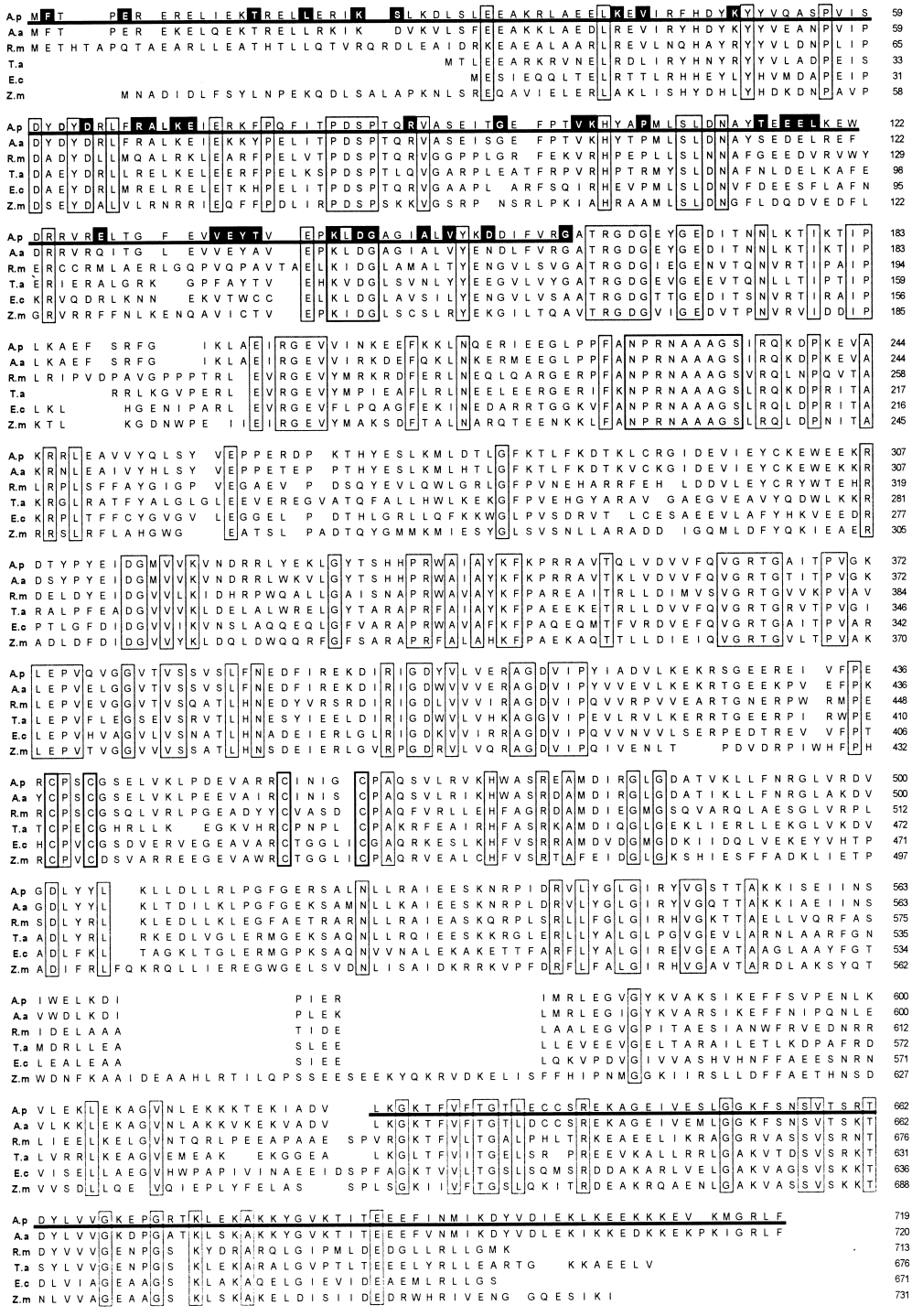

Fig. 1. Amino acid sequence alignment of NAD+-dependent DNA ligases was accomplished with the SEQSEE program. Residues common to all the ligases are *boxed*. The proposed AMP-binding site, KXDG, is *boxed in bold*. The most conserved motif in the bacterial ligase and the zinc-finger-like residues also are *boxed in bold*. Sequences at both termini homologous to *Escherichia coli* adenylate kinase (EcADK) and *Drosophila* replication factor C (RF-C) are *underlined*. The matched residues between Ap DNA ligase and EcADK are indicated by the *black box*. *A.p, Aquifex pyrophilus; A.a, Aquifex aeolicus; R.m, Rhodothermus marinus; T.a, Thermus aquaticus; E.c, Escherichia coli; Z.m, Zymomonas mobilisk*

results). This result suggested that the N-terminal 90 residues, located forward of the AMP-binding site, may also contribute to the adenylation of *Ap* DNA ligase.

Several motifs of NAD^+-dependent DNA ligase, including NPRNAAAGS and CPXC-(aa_{12-14})-C-(aa_{4-5})-CXA, were conserved as reported previously (Thorbjarnardottir et al. 1995). Jonsson et al. (1994) reported that the motif NPRNAAAGS did not align with any other protein sequence to allow deduction of its function. However, the motif CPXC-(aa_{12-14})-C-(aa_{4-5})-CXA aligned with the eukaryotic zinc finger motif. The presence of zinc ions in the *Ap* DNA ligase was analyzed because of the presence of the motif CPXC-(aa_{12-14})-C-(aa_{4-5})-CXA, a putative zinc finger. When purified *Ap* DNA ligase was analyzed by atomic absorption spectrometry, zinc ions at a molar ratio of 0.38 were detected.

The C-terminal region of DNA ligase is responsible for the binding to DNA containing a nick (Timson and Wigley 1999). The isoeletric point (pI) of *Ap* DNA ligase was inferred to be 8.5 on the basis of the amino acid sequence. The pI value is higher than those of ligases from other species. For *Thermus aquaticus* and *E. coli*, the pI values of the ligases are 6.2 and 5.2, respectively. This difference is partly due to the C-terminus (558–719) of *Ap* DNA ligase, which has a pI of 9.8, much higher than that of the other species [4.6 in *E. coli* (532–671) and 6.9 in *T. aquaticus* (533–676)]. The high pI value of *Ap* DNA ligase may suggest a high binding affinity to DNA.

The C-terminal region of the *Ap* ligase has similarity to the domain of the eukaryotic replication factor C (RF-C) (Burbelo et al. 1993; Lu et al. 1993). Burbelo et al. cloned the gene encoding a DNA-binding protein (A1, activator 1, also called replication factor C) by Southwestern screening of a murine cDNA library and found that the middle portion of the 1,131-amino-acid protein has a region homologous to bacterial DNA ligases (Burbelo et al. 1993). The discovery that the region required for A1 binding to DNA has amino acid sequence similarity to a DNA ligase domain suggests that this region may be utilized by both proteins in recognizing DNA.

Overexpression and purification of the *Ap* DNA ligase

The identified gene, which codes for DNA ligase, was cloned into the pET3a vector for overexpression of the target protein. The pET3a vector containing the *Ap* DNA ligase gene was introduced into *E. coli* BL21(DE3), BL21(DE3)-[pLysS], and BL21(DE3)-[pSJS1240]. Efficient production of the ligase protein was achieved only in *E. coli* BL21(DE3)-[pSJS1240]. This result can be explained as follows: AGA and ATA are rarely used codons in *E. coli*. Only 4% of all arginines, which can be encoded by a total of six codons (AGA, AGG, CGA, CGT, CGG, and CGC), are encoded by the codon AGA in *E. coli*, whereas the codon ATA accounts for 7% of the three codons (ATA, ATC, ATT) for isoleucine. In contrast to *E. coli*, *A. pyrophilus* utilizes the codons AGA and ATA at a frequency of 28% and 69%, respectively (Choi et al. 1997). The use of the *E. coli* BL21(DE3)-[pSJS1240] strain would have contributed to the successful overexpression of the recombinant protein.

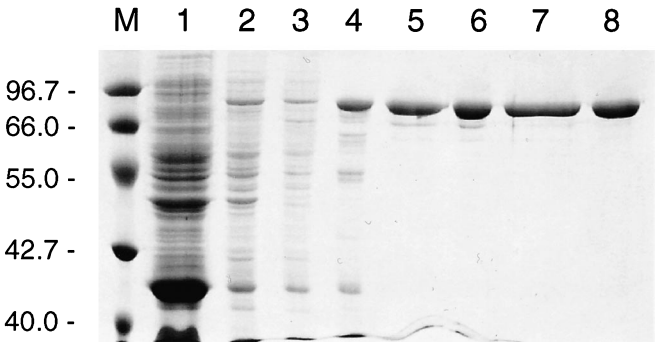

Fig. 2. Identification of recombinant *Ap* DNA ligase. Purification of *Ap* DNA ligase of aliquots at each step of purification were subjected to electrophoresis on a 12.5% polyacrylamide gel. *Lane 1*, crude extract of uninduced cells; *lane 2*, crude extract of induced cells; *lane 3*, soluble fraction after centrifugation; *lane 4*, soluble fraction after heat treatment; *lanes 5–8*, fractions containing DNA ligase from S-sepharose, heparin-sepharose, HiTrap-Blue, and Superdex-S200 chromatography, respectively. The amount of total protein in each sample from lane 2 to lane 8 was adjusted to 10 μg. The sizes of the standard (kDa) are given on the *left*. *Lane M*, marker. The gel was stained with Coomassie blue dye

The overproduced ligase was purified through sequential treatment with S-sepharose, heparin sepharose, HiTrap-Blue, and Superdex-S200 columns. Fractions were taken from each purification step and subjected to electrophoresis on a 12.5% SDS-polyacrylamide gel (Fig. 2). A protein with an apparent molecular mass of 82 kDa was purified to near homogeneity.

Effect of pH and temperature on enzyme activity

An optimum pH of nick-closing activity for *Ap* DNA ligase was in the range pH 8.0–8.6 in Tris-HCl and pH 7.7–8.8 for cohesive end ligation activity, at 65°C (Fig. 3a). More than 90% of the catalytic activity of DNA ligase was retained within this pH range.

Ligation of *Hin*dIII-cleaved λDNA fragments by *Ap* DNA ligase was examined at various temperatures. The *Ap* DNA ligase activity increased up to 65°C (Fig. 3b). The concatemeric λDNA which is too large to exit the well, was trapped when the reaction mixture was incubated at 65°C. The detergents SDS, NP-40, and Triton X-100, plus Proteinase K, were added to the product mixture but could not dissociate the concatemeric λDNA.

To investigate the temperature dependency of *Ap* DNA ligase activity and thermostability, ^{32}P-labeled oligo-duplex substrate (71 mer) was used for the reaction. The T_m of oligonucleotide duplex as *Ap* DNA ligase substrate was 95°C, which was analyzed with DSC and CD. The optimum temperature of *Ap* DNA ligase was about 65°C (Fig. 3c). The optimum temperature of DNA ligase was reported to depend on the length and sequence of the substrate DNA (Takahashi et al. 1984). The activity of the *Ap* DNA ligase was compared with those of *E. coli* DNA ligase (*Ec* ligase; New England Biolabs, Beverly, MA, USA) and *T. aquaticus* DNA ligase (*Taq* ligase; New England Biolabs) using the same substrate and conditions (Fig. 3c). The ligation activi-

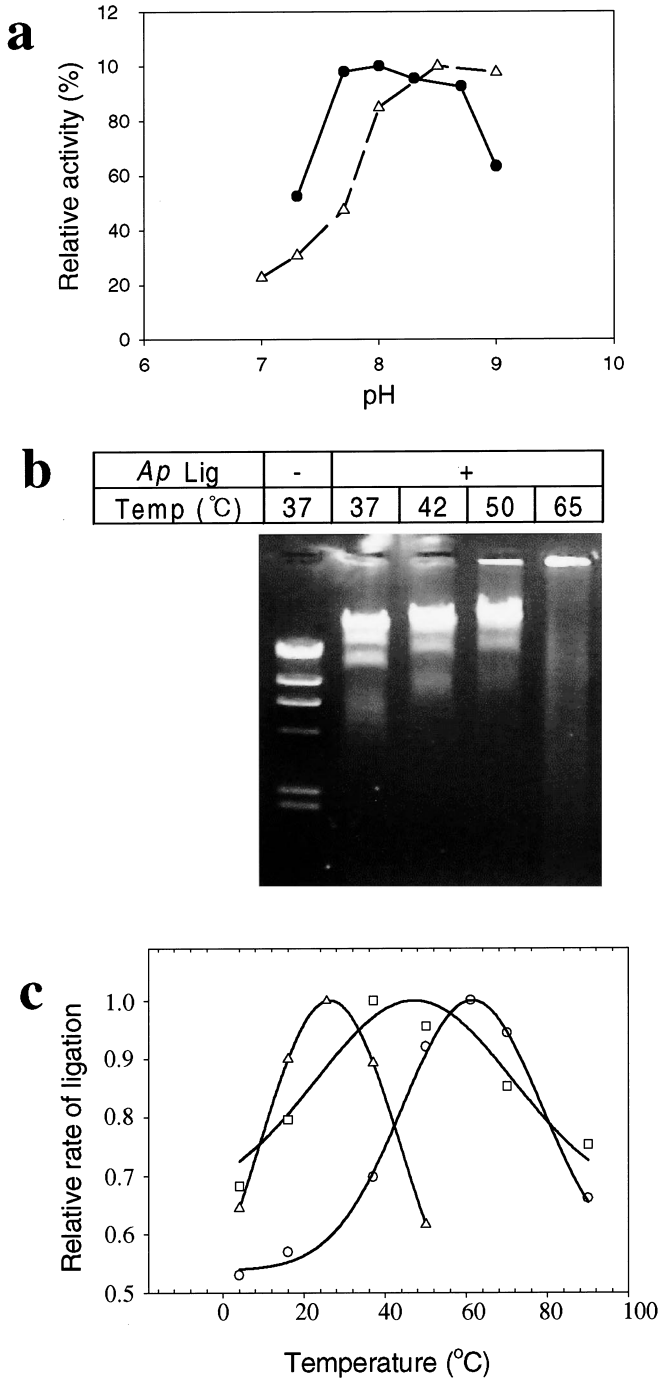

Fig. 3a–c. Relative activity of *Ap* DNA ligase at various pH levels and temperatures. **a** pH dependency. Ligations of cohesive end (*closed circles*) or nick-closing activity (*open triangles*) were determined at various pH levels as described in Materials and methods. All reaction buffers were 50 mM Tris-HCl, adjusted as indicated for pH. **b** Temperature dependence on cohesive end ligation. *Hin*dIII-digested λDNA was incubated with *Ap* DNA ligase (1 μg) at the indicated temperatures, and the reaction was analyzed by electrophoresis on a 1% agarose gel. **c** Temperature dependence of the DNA ligases *E. coli* DNA ligase (*triangles*), *Thermus thermophilus* DNA ligase (*squares*), and *Ap* DNA ligase (*circles*) on nick-closing activity. The annealed DNA substrate (500 fmol) was incubated in a 10-μl reaction mixture (1.2 pmol *Ap* DNA ligase, 50 mM Tris-HCl, pH 8, 10 mM $MgCl_2$, 10 mM DTT, 5 mM NAD^+, 100 mM KCl, and 0.1% Triton X-100) for 1 h at the indicated temperatures. The ligation products were analyzed by a FujiX 2000 BAS Phosphoimage analyzer after electrophoresis on a 10% polyacrylamide gel. The ratio was compared with the highest activity of each ligase in the supplied buffer

ties of *Ec* ligase and *Taq* ligase reached maximum levels at 28° and 48°C, respectively. In contrast, the activity of *Ap* DNA ligase reached maximum at 65°C and decreased above 65°C.

Because *Aquifex pyrophilus* grows between 67° and 95°C, the ligation reaction was performed at high temperature, up to 100°C, using nicked plasmids, pBluescript (2.9 kb), and pMet vector (8 kb). However, the nicked plasmids were hydrolyzed in the buffer condition during the minimum ligation reaction time of 1 h at 80°C in the presence of 100 mM NaCl or potassium glutamate (data not shown). The *Ap* DNA ligase reactions were performed with potassium glutamate concentrations from 150 mM to 1 M at 80°C for 2 h. Some *Ap* DNA ligation products were shown, but most of the template was hydrolyzed (data not shown).

Monovalent and divalent ion dependence of enzyme activity

When each of the monovalent salts NaCl, KCl, and NH_4Cl was added to the reaction mixture, the activity of *Ap* DNA ligase gradually increased in proportion to the concentration of salt, up to 600 mM NaCl, KCl, NH_4Cl, and potassium glutamate (Fig. 4a). This property was different from that of the *T. thermophilus* ligase (*Tt* ligase). In the case of *Tt* ligase, the maximum catalytic activity of the enzyme was observed with low concentrations of the monovalent cations K^+ and NH_4^+ (≤10 mM), whereas ligation activity was inhibited by the presence of the Na^+ cation (Tong et al. 1999). However, K^+, NH_4^+, and Na^+ stimulated the activity of *Ap* DNA ligase, whereas the ligation activity was decreased at a concentration above 150 mM KCl. Because the Cl anion could inhibit the activity of *Ap* DNA ligase, salts were substituted by potassium glutamate. As expected, the enzyme activity was increased up to 350 mM potassium glutamate (Fig. 4a).

Divalent ions are necessary for a ligation reaction. Most DNA-binding enzymes, including ATP- and NAD^+-dependent DNA ligase, prefer the presence of Mg^{2+} ions. To determine the dependency of ligation activity on the divalent ion concentration, we prepared metal-free apoprotein by dialyzing *Ap* DNA ligase against the denaturation and reconstitution conditions as previously described (Lim et al. 1997). After dialysis for demetallization, *Ap* DNA ligase still contained 10% molar ratio of Mg^{2+} ions and retained about 15% of ligation activity. Thus, after substraction of residual activity, the data shown in Fig. 4b were obtained. Divalent metals in the reaction mixture were substituted for $MgCl_2$ to identify divalent cation specificity (Fig. 4b). When 5 mM Mn^{2+} and Ca^{2+} were substituted for $MgCl_2$, 50% and 40% relative activity were observed, respectively (Fig. 4b). *Tt* ligase was reported to use Mn^{2+} and Ca^{2+} as a cofactor for ligation activity (Tong et al. 1999), and ATP-dependent DNA ligase from *Chlorella virus* PBCV-1 uses Mn^{2+} and Co^{2+} (Ho et al. 1997). When Cu^{2+} was used as a metal cofactor for *Ap* DNA ligase, a small amount of ligation activity was observed. The crystal structure of NAD^+-dependent *Thermus filiformis* DNA ligase suggested that the AMP-binding pocket may include some of the highly conserved residues to participate in magnesium ion binding (Lee et al. 2000).

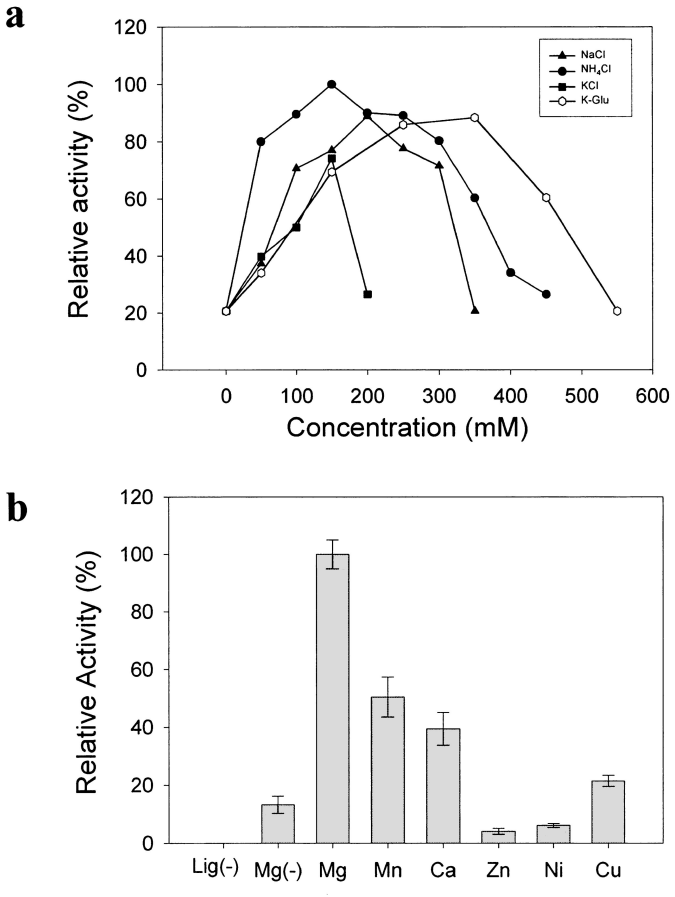

Fig. 4a,b. Effect of salt concentration and divalent cation specificity in the ligation reaction. **a** The *Ap* DNA ligase was assayed at various salt concentrations with oligonucleotide substrate under the conditions described in Materials and methods. Ammonium chloride (*circles*), sodium chloride (*triangles*), potassium chloride (*squares*), and potassium glutamate (*hexagons*) were added to each reaction buffer. **b** The assay of ligation was performed with demetallized enzyme, which has about 10% of the maximum activity. The indicated divalent metals were substituted for $MgCl_2$ in this reaction mixture. All metals were added as the chloride salt, and their concentrations in the reaction mixture were adjusted to 5 mM. The *negative* (–) symbol denotes the component in the reaction mixture that was removed

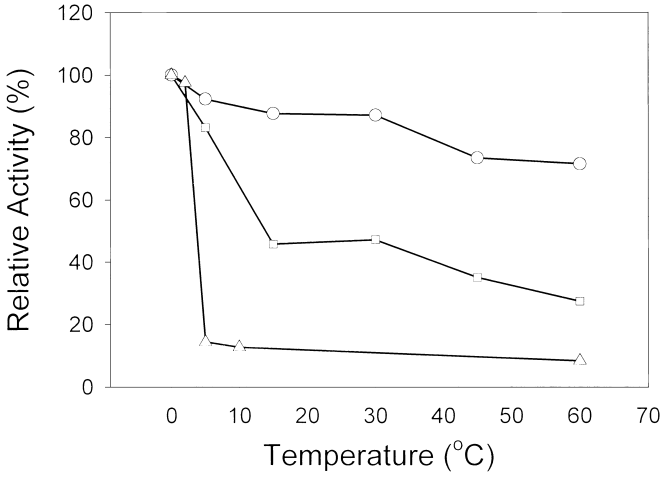

Fig. 5. Thermal inactivation of *E. coli* DNA ligase (*triangles*), *Thermus thermophilus* DNA ligase (*squares*), and *Ap* DNA ligase (*circles*). Each enzyme (25 µg/µl in 50 mM Tris-HCl, pH 8.0, 10 mM DTT, 10 mM $MgCl_2$) was incubated at 95°C, and an aliquot taken at regular intervals was immediately chilled on ice to quench the activity. The residual activity of each aliquot was determined according to the standard assay method (nick-closing assay) in the reaction buffers optimized for each enzyme

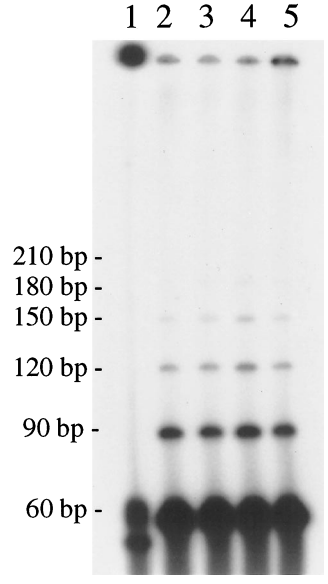

Fig. 6. Repeat expansion detection. An autoradiogram of the integers on the *left* of the figure depicts the expanded products, which are multiples of the designated 30-bp oligonucleotide. The amount of *Ap* DNA ligase added to the ligation reaction was as follows: 1.2 pmol (*lane 1*), 2.4 pmol (*lane 2*), 3.6 pmol (*lane 3*), 4.8 pmol (*lane 4*), 6.1 pmol (*lane 5*). The reaction and detection methods are described in Materials and methods

Thermostability of *Ap* DNA ligase

To measure protein thermostability, *Ap* DNA ligase (0.1 µg/µl in 50 mM Tris-HCl, pH 8.0, 10 mM DTT, 10 mM $MgCl_2$) was incubated at 95°C. Aliquots were taken at regular intervals and immediately chilled on ice to quench the activity. The residual activity of each aliquot was analyzed using a standard nick-closing assay method. The estimated half-lives of *T. aquaticus* and *E. coli* ligases were ≤15 min and ≤5 min, respectively, under the conditions mentioned here. In the case of *Ap* DNA ligase, more than 75% of the ligation activity remained after heating at 95°C for 60 min (Fig. 5). According to previously published results, the half-lives of ligases from *Thermus thermophilus* HB8, *Thermus scotoductus*, and *Rhodomarinus marinus* were 26, 26, and 7 min, respectively, at 91°C (Thorbjarnardottir et al. 1995; Jonsson et al. 1994; Takahashi et al. 1984). The half-life of *Bacillus stearothermophilus* was 20 min at 65°C (Brannigan et al. 1999).

RED provides a method to determine if repeat expansions are associated with human genetic diseases. Long trinucleotide repeats are able to serve as templates to catalyze ligation of oligonucleotides into larger multimers, providing information on the size of the largest trinucleotide repeats in the genome. To apply RED and LCR to DNA ligase, the enzyme should be resistant to heat because the reaction is performed at high temperatures. The *Ap* DNA ligase was

found to be thermostable, and it retains substantial catalytic activity after exposure to 95°C. Thus far, two thermostable DNA ligases, *Pyrococcus furiosus* DNA ligase (Mathur et al. 1996) and Ampligase, have been purified and applied to the diagnostic method. The thermostable *Ap* DNA ligase was subjected to RED (Fig. 6), and we observed 150 bp of $(CTG)_{10}$ repeats from the human genomic DNA. The optimum concentration of *Ap* DNA ligase was 4.8 pmol in RED analysis.

Several recent studies performed using RED have described a significant association between longer CAG/ATG repeats and both bipolar affective disorder (Lindblad et al. 1995) and schizophrenia (Morris et al. 1995). Thermostable NAD^+-dependent *Ap* DNA ligase has many advantages for the use of RED and LCR.

Acknowledgments We gratefully acknowledge financial support from the Korea Institute of Science and Technology and a grant from the Ministry of Science and Technology, Korea. We thank Je-An Park (Advanced Analysis Center, KIST) for assistance with atomic absorption measurements.

References

Barany F (1991) Genetic disease detection and DNA amplification using cloned thermostable ligase. Proc Natl Acad Sci USA 88:189–193

Barany F, Gelfand DH (1991) Cloning, overexpression and nucleotide sequence of a thermostable DNA ligase-encoding gene. Gene (Amst)109:1–11

Barnes DE, Johnston LH, Kodama K, Tomkinson AE, Lasko DD, Lindahl T (1990) Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA 87:6679–6683

Berry MB, Meador B, Bilderback T, Liang P, Glaser M, Phillips GN Jr (1994) The closed conformation of a highly flexible protein: the structure of *E. coli* adenylate kinase with bound AMP and AMPPNP. Proteins 19:183–198

Brannigan JA, Ashford SR, Doherty AJ, Timson DJ, Wigley DB (1999) Nucleotide sequence, heterologous expression and novel purification of DNA ligase from *Bacillus stearothermophilus* (1). Biochim Biophys Acta 1432:413–418

Burbelo PD, Utani A, Pan ZQ, Yamada Y (1993) Cloning of the large subunit of activator 1 (replication factor C) reveals homology with bacterial DNA ligases. Proc Natl Acad Sci USA 90:11543–11547

Cheng C, Shuman S (1997) Characterization of an ATP-dependent DNA ligase encoded by *Haemophilus influenzae*. Nucleic Acids Res 25:1369–1374

Choi IG, Kim SS, Ryu JR, Han YS, Bang WG, Kim SH, Yu YG (1997) Random sequence analysis of genomic DNA of a hyperthermophile: *Aquifex pyrophilus*. Extremophiles 1:125–134

Deckert G, Warren PV, Gaasterland T, Young WG, Lenox AL, Graham DE, Overbeek R, Snead MA, Keller M, Aujay M, Huber R, Feldman RA, Short JM, Olsen GJ, Swanson RV (1998) The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*. Nature (Lond) 392:353–358

Doherty AJ, Ashford SR, Subramanya HS, Wigley DB (1996) Characterization of proteolytic fragments of bacteriophage T7 DNA ligase. J Biol Chem 271:11083–11089

Dunn JJ, Studier FW (1981) Nucleotide sequence from the genetic left end of bacteriophage T7 DNA to the beginning of gene 4. J Mol Biol 148:303–330

Ho CK, Van Etten JL, Shuman S (1997) An essential surface motif (WAQKW) of yeast RNA triphosphatase mediates formation of the mRNA capping enzyme complex with RNA guanylyltransferase. J Virol 71:1931–1937

Huber R, Wilharm T, Huber D, Tricone A, Burggraf S, Koning H, Rachel R, Rockinger I, Fricke H, Stetter KO (1992) *Aquifex pyrophilus*, gen. nov. sp. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. Syst Appl Microbiol 15:349–351

Jonsson ZO, Thorbjarnardottir SH, Eggertsson G, Palsdottir A (1994) Sequence of the DNA ligase-encoding gene from *Thermus scotoductus* and conserved motifs in DNA ligases. Gene (Amst) 151:177–180

Kletzin A (1992) Molecular characterisation of a DNA ligase gene of the extremely thermophilic archaeon *Desulfurolobus ambivalens* shows close phylogenetic relationship to eukaryotic ligases. Nucleic Acids Res 20:5389–5396

Lauer G, Rudd EA, McKay DL, Ally A, Ally D, Backman KC (1991) Cloning, nucleotide sequence, and engineered expression of *Thermus thermophilus* DNA ligase, a homolog of *Escherichia* DNA ligase. J Bacteriol 173:5047–5053

Lee JY, Chang C, Song HK, Moon J, Yang JK, Kim H-K, Kwon S-T, Suh SW (2000) Crystal structure of NAD^+-dependent DNA ligase: modular architecture and functional implications. EMBO J 19:1119–1129

Lehman IR (1974) DNA ligase: structure, mechanism, and function. Science 186:790–797

Lim JH, Yu YG, Choi IG, Ryu JR, Ahn BY, Kim SH, Han YS (1997) Cloning and expression of superoxide dismutase from *Aquifex pyrophilus*, a hyperthermophilic bacterium. FEBS Lett 406:142–146

Lindblad K, Nylander PO, De bruyn A, Sourey D, Zander C, Engstrom C, Holmgren G, Hudson T, Chotai J, Mendlewicz J, Van Broeckhoven C, Schalling M, Adolfsson R (1995) Detection of expanded CAG repeats in bipolar affective disorder using the repeat expansion detection (RED) method. Neurobiol Dis 2:55–62

Lu Y, Zeft AS, Reigel AT (1993) Cloning and expression of a novel human DNA binding protein, PO-GA. Biochem Biophys Res Commun 193:779–786

Mathur EJ, Marsh EJ, Schoettlin WE (1996) Purified thermostable *Pyrococcus furiosus* DNA ligase. U.S. Patent No. 5,506,137, April 9, 1996

Morris AG, Gaitonde E, McKenna PJ, Mollon JD, Hunt DM (1995) CAG repeat expansions and schizophrenia: association with disease in females and with early age-at-onset. Hum Mol Genet 4:1957–1961

Robins P, Lindahl T (1996) DNA ligase IV from HeLa cell nuclei. J Biol Chem 271:24257–24261

Schalling M, Hudson TJ, Buetow KH, Housman DE (1993) Direct detection of novel expanded trinucleotide repeats in the human genome. Nat Genet 4:135–139

Schulz GE, Muller CW, Diederichs K (1990) Induced-fit movements in adenylate kinases. J Mol Biol 213:627–630

Shark KB, Conway T (1992) Cloning and molecular characterization of the DNA ligase gene (lig) from *Zymomonas mobilis*. FEMS Microbiol Lett 75:19–26

Shuman S, Ru XM (1995) Mutational analysis of vaccinia DNA ligase defines residues essential for covalent catalysis. Virology 211:73–83

Shuman S, Schwer B (1995) RNA capping enzyme and DNA ligase: a superfamily of covalent nucleotide transferases. Mol Microbiol 17(3):405-410

Takahashi M, Yamaguchi E, Uchida T (1984) Thermophilic DNA ligase. J Biol Chem 259:10041–10047

Thorbjarnardottir SH, Jonsson ZO, Andresson OS, Kristjansson JK, Eggertsson G, Palsdottir A (1995) Cloning and sequence analysis of the DNA ligase-encoding gene of *Rhodothermus marinus*, and overproduction, purification and characterization of two thermophilic DNA ligases. Gene (Amst) 161:1–6

Timson DJ, Wigley DB (1999) Functional domains of an NAD^+-dependent DNA ligase. J Mol Biol 285:73–83

Tomkinson AE, Roberts E, Daly G, Totty NF, Lindahl T (1991) Location of the active site for enzyme-adenylate formation in DNA ligase. J Biol Chem 266:21728–21735

Tong J, Cao W, Barany F (1999) Biochemical properties of a high fidelity DNA ligase from *Thermus* species AK16D. Nucleic Acids Res 27:788–794

Wishart DS, Boyko RF, Willard L, Richards FM, Sykes BD (1994) SEQSEE: a comprehensive program suite for protein sequence analysis. Comput Appl Biosci 10:121–132