JOURNAL OF BACTERIOLOGY, Aug. 1991, p. 5047–5053
0021-9193/91/165047-07$02.00/0

Vol. 173, No. 16

Cloning, Nucleotide Sequence, and Engineered Expression of *Thermus thermophilus* DNA Ligase, a Homolog of *Escherichia coli* DNA Ligase

GAIL LAUER,[1] EDWIN A. RUDD,[1]† DIANE L. McKAY,[1] ABDUL ALLY,[2] DELPHINE ALLY,[2] AND KEITH C. BACKMAN[1,2]*

*BioTechnica Diagnostics, Inc., Cambridge, Massachusetts 02138,[1] and BioTechnica International, Inc., 85 Bolton Street, Cambridge, Massachusetts 02140[2]**

Received 15 January 1991/Accepted 3 June 1991

We have cloned and sequenced the gene for DNA ligase from *Thermus thermophilus*. A comparison of this sequence and those of other ligases reveals significant homology only with that of *Escherichia coli*. The overall amino acid composition of the thermophilic ligase and the pattern of amino acid substitutions between the two proteins are consistent with compositional biases in other thermophilic enzymes. We have engineered the expression of the *T. thermophilus* gene in *Escherichia coli*, and we show that *E. coli* proteins may be substantially removed from the thermostable ligase by a simple heat precipitation step.

DNA ligases (EC 6.5.1.1 or 6.5.1.2) are essential enzymes in the replication, recombination, and repair of DNA. As such, we might expect that an archetypical ligase arose at an early point in evolution and then was disseminated throughout the biota. Several points stand in contrast to such a view. The sequences of ligases from eukaryotes (5–7), prokaryotes (22), bacteriophages (1, 17), and viruses (39) exhibit significant dissimilarity. Ligases use one of two different high-energy adenylates to drive their reactions. In an unusual reaction, ligases determined by prokaryotic cells use the diphosphate linkage of NAD like a phosphoanhydride linkage of ATP (28, 42). All other ligases use ATP itself (5, 6, 20a, 39, 46). Various ligases also exhibit dissimilar abilities to join noncanonical substrates, such as blunt-ended duplexes (37, 41, 50) or RNA-DNA hybrids (2, 19).

Despite the differences in the primary sequence, substrate, and cofactor aspects of ligases, all ligases appear to act via an identical sequence of reaction steps (18, 40). In the first step, the AMP moiety of a high-energy adenylate (NAD or ATP) is covalently linked to the enzyme, releasing NMN or pyrophosphate. Next, this AMP moiety is transferred to the phosphorylated 5′ end of a DNA strand, creating a new diphosphate linkage. Finally, the 3′ hydroxyl of an adjacent DNA strand attacks the new diphosphate linkage, which releases AMP, creates a phosphodiester linkage, and thereby completes the covalent joining of the DNA strands. Either the various ligases are truly homologous in a way which is not readily apparent at the primary sequence level (but exists at a secondary or tertiary structure level), or they have convergently evolved identically detailed reaction mechanisms.

Thermus thermophilus determines a DNA ligase that is active at temperatures above 75°C (42). This enzyme is present in low abundance, and its purification (42) is complicated by the presence of host proteases. Although it is known to utilize NAD, its relatedness to *Escherichia coli* DNA ligase was unexplored. We wished to obtain greater quantities of this enzyme free from thermostable contaminants, and we wished to understand its phylogenetic relationship to other ligases, as well as aspects of its composition which contribute to its thermostability. Accordingly, we cloned and sequenced its structural gene. We significantly increased expression of the gene by attaching highly functional *E. coli* transcription and translation signals. We found that the differential thermostability of the enzyme in its new host can be used in enzyme purification starting from our overproducing strain. In addition, we have compared the deduced amino acid sequence of the *T. thermophilus* ligase with the sequences of other ligases and find that it is similar only to that from *E. coli*, with which it also shares common biochemical properties. Comparison of the sequences of these two enzymes reveals changes in composition which may be related to thermostability.

MATERIALS AND METHODS

Chemicals and reagents. Cibacron blue beaded agarose and Fractogel HW-50S were obtained from Pierce Chemical. Centricon-10 and Centriprep-10 with 10,000-molecular-weight cutoff membranes were obtained from Amicon. Protein molecular weight standards were from Bio-Rad. Reagents for DNA manipulation were of the highest quality commercially available. DNA ligase from *T. thermophilus* cells was purified by a modification of the procedure of Takahashi et al. (42, 43). Radiochemicals were from NEN-DuPont; ^{14}C-labelled protein standards were from Amersham.

Strains and strain growth. The bacterial strains and plasmids used are described in Table 1. Luria broth (LB) and minimal medium (M9) recipes are from Miller (27). *T. thermophilus* was grown on ATCC medium 697 (20). Plasmid selection was generally performed on LB supplemented with ampicillin (100 μg/ml). The clone bank was selected on LB supplemented with tetracycline (5 μg/ml). KB649/pGL628 cells for ligase production were grown at 33°C in 10 liters of minimal medium in a Chemap CF 2000 fermentor to an optical density of 5 at 600 nm and induced by shifting the temperature to 42°C. Cells were harvested 2 h after temperature shift by centrifugation and stored frozen.

DNA manipulation and clone bank. Standard methods of DNA manipulation in vitro were used (9). Plasmid DNA

* Corresponding author.
† Present address: Syncap, Cambridge, MA 02138.

TABLE 1. Bacteria and plasmids

Strain or plasmid	Relevant feature(s)	Reference or source
T. thermophilus		
HB8		29; ATCC 27634
E. coli strains		
HB101	Host for cloning	11
JM109	Host for subcloning	48
KB649	*E. coli* YMC9 (3) containing phage lambda repressor gene (*c*I857)	This laboratory
Plasmids		
pGL516	Expression vector; phage lambda p_L promoter cloned into pBR322 (10)	This laboratory
pGL600	Source of DNA fragments for sequencing ligase gene	This work
pGL628	Production plasmid for ligase	This work
pTR264	Positive selection vector; similar to pTR262 (33) but carries *bla* gene	This laboratory
pUC18	Cloning vector	48
pUC19	Cloning vector	48
M13mp18	Sequencing vector	48
M13mp19	Sequencing vector	48

isolation (13), plasmid miniprep procedures (8, 15, 21), agarose gel electrophoresis (23), DNA fragment preparation (44), and cell transformation (14) have been described elsewhere. A clone bank was constructed in strain HB101 by cloning 7- to 30-kb DNA fragments generated by partial *Sau*3AI digestion of *T. thermophilus* chromosomal DNA into pTR264 digested with *Bcl*I and treated with calf intestinal phosphatase. Subclones were prepared in strain JM109 by using vectors pUC18 and pUC19.

Screening of the clone bank. Pools of colonies (30 pools of 70 colonies each in the first round, 10 pools of 7 colonies each in the second round, and 7 single colonies in the third round) were grown in 100 ml of LB to near saturation, and a 20-ml aliquot was harvested and stored frozen. Cell pellets were disrupted by sonication in 10 ml of disruption buffer (50 mM Tris-HCl, 10 mM $MgCl_2$, 50 mM KCl, 2 mM $MnCl_2$, 1 mM dithiothreitol, 10 mM NH_4Cl, pH 7.7), and the debris was removed by centrifugation (20 min, 30,000 × *g*, 4°C). The supernatant was incubated at 80°C for 30 min, and the coagulated protein was removed by centrifugation (20 min, 5,000 × *g*, 4°C). The supernatant was concentrated in two steps (using Centriprep and Centricon concentrators), first to about 0.5 ml and finally to about 50 μl. The sample was again heated to 80°C for 20 min in a 0.5-ml microcentrifuge tube, and the precipitate was removed by centrifugation (4 min, top speed). The final supernatant was labelled at 80°C with [^{32}P]NAD (1.0 μl of a 1:1 dilution of the NEN stock with 0.2 M Tris-HCl, pH 7.6). After 2 h, the reaction was stopped with 25 μl of sodium dodecyl sulfate (SDS) sample buffer, analyzed by SDS-polyacrylamide gel electrophoresis according to Laemmli (25), and subjected to autoradiography.

DNA chemistry. Synthetic DNA was prepared on an Applied Biosystems model 380B DNA synthesis machine, using protocols supplied by the manufacturer.

DNA sequencing was performed by the method of Sanger et al. (35) on a set of overlapping fragments derived from pGL600 and cloned into M13mp18 or M13mp19. About 2.4 kb of the insert was completely sequenced on both strands. Sequence information was aligned, and a consensus sequence was generated on IBM or fully compatible computers, using the DNAStar sequence analysis program, which was also used for other sequence analyses.

Ligase partial purification and assay. Frozen KB649/pGL628 induced cells (15 g) were suspended in 60 ml of disruption buffer (as described above) and disrupted at 14,000 lb in^2 in a French pressure cell chilled to 4°C. Cell debris was removed by centrifugation (20 min, 4°C, 30,000 × *g*). The supernatant (fraction I) was incubated at 80°C for 30 min, and the resulting precipitate was removed by centrifugation (20 min, 4°C, 30,000 × *g*). Streptomycin sulfate (10% stock solution, 2% final concentration) was added to the supernatant (fraction II), and the solution was stirred for 20 min on ice. The resulting precipitate was removed by centrifugation (20 min, 4°C, 16,000 × *g*), yielding supernatant fraction III. This was loaded onto a Cibacron blue agarose column (2.2 by 24 cm) equilibrated with buffer A (20 mM Tris-HCl, 50 mM KCl, pH 7.6). The column was washed with buffer A until the A_{280} of the eluent returned to baseline. Ligase was eluted with a 300-ml linear gradient of KCl (50 mM to 1 M) in 20 mM Tris-HCl, pH 7.6. The flow rate was 2 ml/min, and 5-ml fractions were collected. Active fractions were pooled and concentrated in a Centriprep to a volume of 2.5 ml (fraction IV). In some preparations, fraction IV had a noticeable pink tinge, which was not removed by subsequent steps. Fraction IV was applied to a Fractogel 50S column (2.2 by 32 cm) equilibrated with 10 mM KPO_4–100 mM KCl–1 mM $MgCl_2$, pH 7.6, and eluted with the same buffer at a flow rate of 0.5 ml/min. Fractions of 2.5 ml were collected, and active fractions were pooled (fraction V). For storage, fraction V was concentrated in a Centriprep to 0.5 ml, diluted with 5 ml of 20 mM Tris-HCl–100 mM KCl, pH 7.6, and reconcentrated in a Centriprep to 0.8 ml. To this

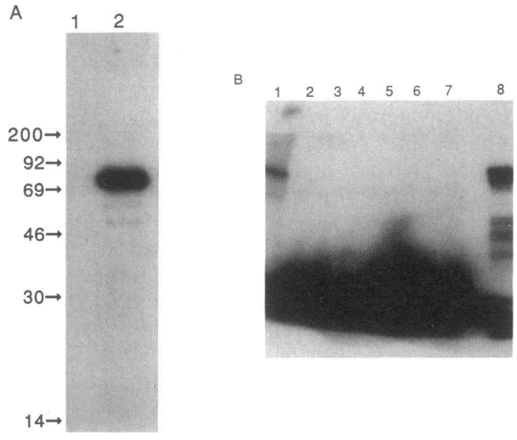

FIG. 1. Labelling of ligase and screening of clones. A pool of seven *E. coli* isolates harboring plasmids with *T. thermophilus* DNA inserts had been assayed as a pool and shown to express thermostable DNA ligase. Ligase purified from *T. thermophilus* was labelled with [^{32}P]NAD. Extracts made from each member of the pool were separately labelled with [^{32}P]NAD. Labelled samples were analyzed by gel electrophoresis as described in Materials and Methods. (A) Lanes: 1, ^{14}C-labelled protein standards (M_r in thousands indicated at the left; the bands are faint in this reproduction); 2, ligase purified from *T. thermophilus*. (B) Lanes: 1 to 7, individual *E. coli* isolates from the pool; 8, ligase purified from *T. thermophilus*.

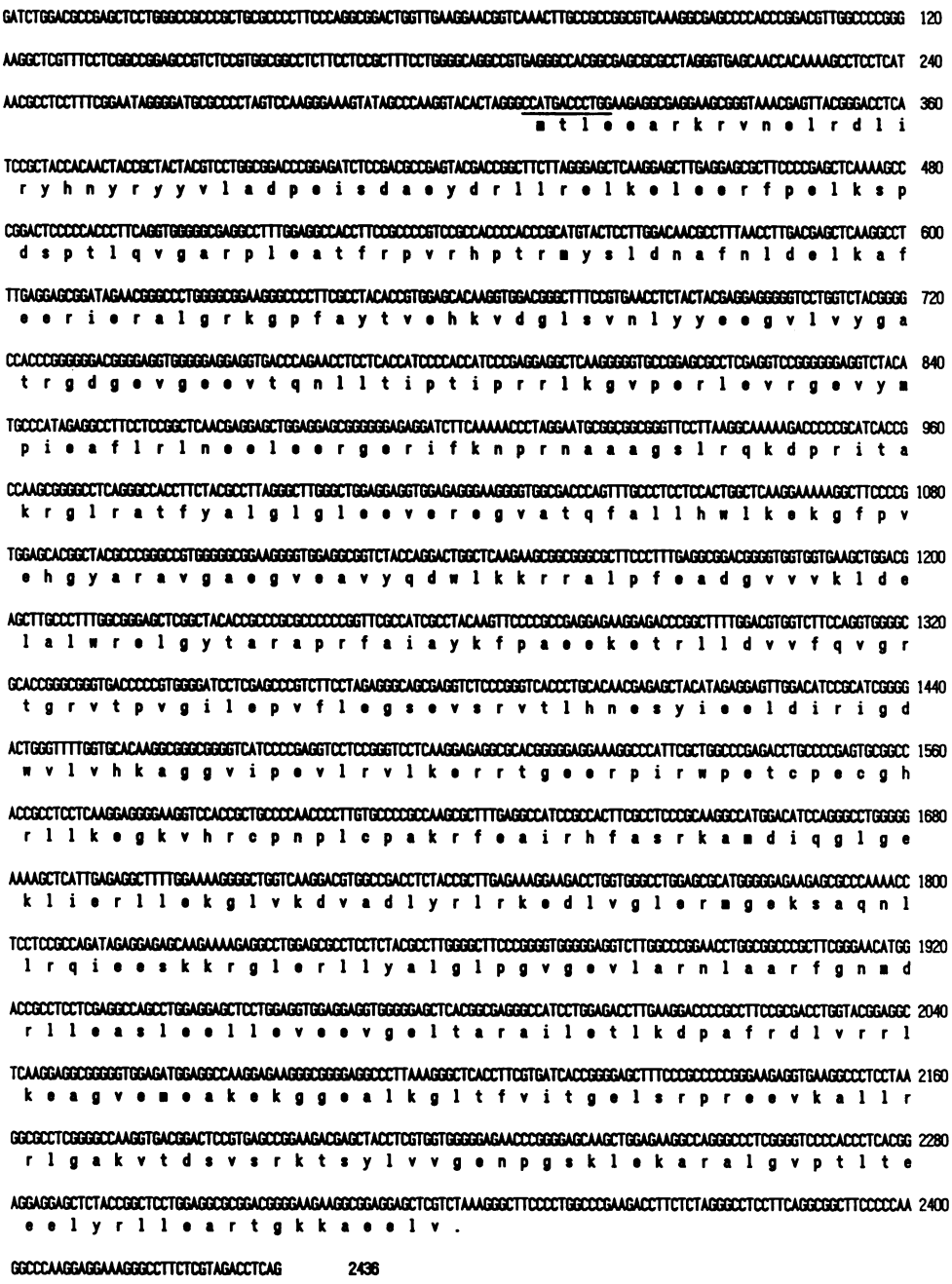

FIG. 2. Sequence of the *T. thermophilus* ligase gene. The nucleotide sequence of the ligase gene is given in capital letters, and the deduced amino acid sequence of the protein is given in lowercase letters aligned with corresponding codons. Underlined is the unique cleavage site for *Bst*XI which was used in constructing the production vector pGL628.

was added an equal volume of glycerol, and the ligase was stored at −20°C.

Ligase was assayed quantitatively by a modification of the method of Barker et al. (4). A unit of ligase catalyzes the formation of 1 μmol of product per min. Protein concentration was determined by the method of Bradford (12), using bovine serum albumin as a standard. Protein analysis by polyacrylamide gel electrophoresis in the presence of SDS was performed as described by Laemmli (25).

Nucleotide sequence accession number. The sequence of the *T. thermophilus* ligase gene has been submitted to GenBank and assigned accession number M36417.

RESULTS

Proceeding from the observation that other ligases can be covalently adenylylated (24, 26, 49), we determined that the DNA ligase from *T. thermophilus* was stably labelled by [^{32}P]NAD at elevated temperatures (Fig. 1A). We then showed that the ability of *E. coli* ligase (and other *E. coli*

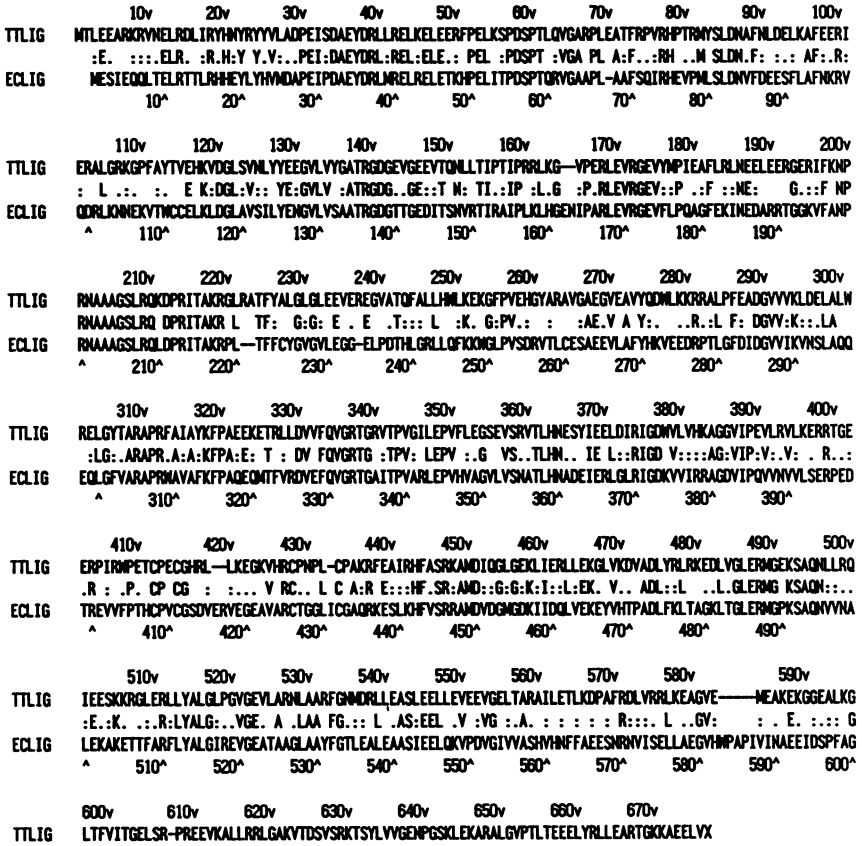

FIG. 3. Comparison of DNA ligase amino acid sequences. The deduced amino acid sequences of DNA ligases from *T. thermophilus* (TTLIG) and *E. coli* (ECLIG) are aligned. Gaps introduced to improve alignment are indicated by dashes. Between the two sequences are printed identically conserved residues. Conservative replacements are indicated by colons or dots. We present the *E. coli* ligase sequence determined in our laboratory, which differs from the published sequence (22) at one position. We assign codon 69 as A, whereas Ishino et al. (22) assigns it as R.

proteins) to be stably labelled with this reagent was destroyed by treatment at those elevated temperatures (Fig. 1B, lanes 2 to 7). This provided a screening assay for our clone bank. Reconstruction experiments (mixing *T. thermophilus* and *E. coli* cells) indicated that we should be able to find 1 clone in a pool of about 70. We screened our clone bank accordingly, found one weakly active pool, and identified the active clone in that pool (Fig. 1B, lane 1). A DNA joining assay confirmed that the labelled protein was in fact a thermostable DNA ligase. The plasmid in the clone (pGL600) contained an insert of approximately 8.3 kb, which was subcloned into pUC18 and pUC19. The subcloning procedure localized the gene to a 3.3-kb region and determined its orientation, since the *lac* promoter on the vector affected the amount of ligase produced. We sequenced about 2.4 kb of that region (Fig. 2) and identified an open reading frame corresponding to ligase. We identified the ligase gene by the correspondence of the size of its expected product (676 amino acids; 76,913 Da) with that observed for the native protein (79,000 Da; 39) and by the significant homology it exhibits with the ligase from *E. coli* (Fig. 3). Our engineering experiments (below) confirm our assignment of the initiation codon.

The sequence does not contain transcription or translation initiation sites that would be expected to operate efficiently in *E. coli*. Our strategy to attach such elements to the structural gene hinged upon a unique *Bst*XI site near the beginning of the coding sequence (Fig. 2). We synthesized a DNA fragment containing a highly functional Shine-Dalgarno sequence (38) for efficient initiation of translation. The synthetic DNA had sticky ends that directed its insertion between an *Nhe*I cleavage site on the expression vector (pGL516) and the *Bst*XI site within the ligase gene. This results in a strong promoter and ribosome binding site directly upstream of the ligase gene. We tried several different Shine-Dalgarno sequences. We found that a designed sequenced (i.e., not modelled on a known natural ribosome binding site) (Fig. 4), which has worked well in expressing other foreign genes in *E. coli* (30), directed the synthesis of the highest observed levels of ligase. The resulting plasmid was named pGL628. Upon induction, strain KB649/pGL628 produces about 3 to 5% of its total protein as the thermostable ligase (Fig. 5).

In our screening assay, we established that *T. thermophilus* ligase remains soluble and active after heat denaturation and precipitation of *E. coli* proteins. This fact can be exploited in purifying thermostable ligase from extracts of induced KB649/pGL628 cells. Following lysis, the large

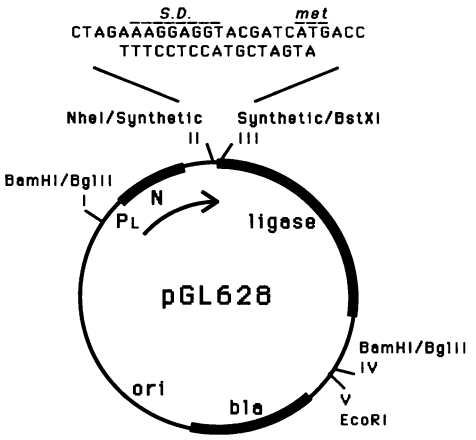

FIG. 4. Production plasmid pGL628. A map (approximately to scale) of plasmid pGL628 is represented as a circle, with thickened regions indicating genes. Genetic elements (genes, the p_L promoter, and the origin of replication) are labelled inside the circle, and an arrow indicating the direction of transcription from p_L is shown. Junctions of component DNA fragments are indicated by hash marks outside the circle which are labelled with roman numerals. The vector is based on pBR322, which is present as an *Eco*RI-to-*Bam*HI fragment (3,986 bp) between junctions V and I. A *Bgl*II-to-*Nhe*I fragment (1,032 bp) carrying the p_L promoter and gene *N* lies between junctions I and II. A synthetic DNA fragment containing the ribosome binding site lies between junctions II and III; its sequence is shown at the top, and the Shine-Dalgarno (S.D.) sequence and initiator methionine codon are labelled. A *Bst*XI-to-*Bgl*II DNA fragment (about 2,900 bp) from *T. thermophilus*, which includes the gene for ligase, lies between junctions III and IV. In constructing pGL628, a small portion of polylinker (*Bam*HI to *Eco*RI; 21 bp) from pUC19 was carried along and lies between junctions IV and V.

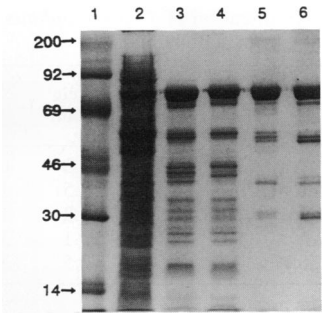

FIG. 5. Analysis of ligase fractions by SDS-polyacrylamide gel electrophoresis. Electrophoresis was carried out on a 10% acrylamide gel (15 by 20 by 0.15 cm) as described in Materials and Methods. The gel was stained with Coomassie brilliant blue. Lanes: 1, protein standards (M_r in thousands indicated at the left); 2, 123 μg of fraction I; 3, 15 μg of fraction II; 4, 12 μg of fraction III; 5, 6 μg of fraction IV; 6, 6 μg of fraction V.

majority of *E. coli* proteins are removed by a thermal denaturation step. This results in significant purification in a single nonchromatographic step (Fig. 5). Nucleic acids, the major remaining physical contaminant at this point, are removed by streptomycin sulfate precipitation. At this stage, the ligase preparation exhibits undesirable nuclease contamination. This can be removed by chromatographic steps, such as affinity chromatography on Cibacron blue agarose and size fractionation on Fractogel 50S. Typical results are presented in Fig. 5 and Table 2.

DISCUSSION

Initial attempts to clone the *T. thermophilus* ligase gene by complementation of *E. coli* mutants (not shown) were unsuccessful, complicated by reversion of the mutant strains, low expression of the *T. thermophilus* gene, and low activity of the enzyme at mesophilic temperatures. A screen based on labelling of the ligase with [^{32}P]NAD was successful for four reasons: first, background is inherently low because few proteins are capable of being stably labelled by this reagent; second, the *E. coli* proteins capable of being labelled could be killed by heat treatment which was not harmful to the thermostable ligase; third, the enzyme can be labelled to high specific activity because of its low K_m for NAD (20 nM; 42); and fourth, tiny amounts of ligase could be separated from vast excesses of host proteins by virtue of differential thermal stability, thus permitting resolution of the samples by gel electrophoresis. The combination of these unusual circumstances resulted in a sensitive screening assay which identified our primary clone.

T. thermophilus ligase is a homolog (46% conserved residues) of *E. coli* DNA ligase; this is consistent with the biochemical similarities of the two ligases. The important features of a comparison, indeed, may be in the nonconserved residues, where information concerning differences in thermostability between the respective enzymes may reside. A full understanding of the significance of differences may require a determination of the structures of the two ligases. However, we note several trends. The strongest trend is a significant increase in the proportion of charged residues (glutamic acid, arginine, and to a lesser extent lysine) in thermophilic ligase. These residues may be involved in protein-stabilizing salt bridges or in ionic interactions with substrates (NAD and DNA), either of which might be important for activity at elevated temperatures. The net increase in these charged residues is offset to a significant extent by a decrease in the abundance of alanine, glutamine, and serine. It is interesting that the abundance of aspartic acid, itself a charged residue, actually decreases. This may reflect its higher chemical reactivity compared with glutamic acid. Other reactive or thermally less stable residues are either slightly (methionine, cysteine, and asparagine) or more severely (glutamine) decreased in relative abundance. Among hydrophobic residues, valine, isoleucine, and to a lesser extent phenylalanine show decreased abundance, offset by large increases in the abundance of leucine and tyrosine. Indeed, the most common replacement is leucine (in *T. thermophilus*) for valine (in *E. coli*). This may reflect a trend toward larger side chains which more completely fill the internal volume of the protein. Similar, but not identical, compositional trends have been observed in a xylose isomerase isolated from *T. thermophilus* (16); differences in trends between these examples may in part result from the fact that the highly thermostable xylose isomerase from *T. thermophilus* was compared with moderately thermostable xylose isomerases from various actinomycetes rather than from *E. coli*.

Since *T. thermophilus* is not closely related to *E. coli*, it seems reasonable to speculate that all prokaryotic NAD-utilizing DNA ligases are related to the enzymes discussed here, although this is by no means an established fact. Mammalian, yeast, and viral ligases, which operate in eukaryotic cells, have some conserved sequences (7, 39) and

TABLE 2. Purification of *T. thermophilus* ligase from KB649/pGL628

Fraction	Step	Vol (ml)	Activity (U/ml)	Protein (mg/ml)	Sp act (U/mg of protein)	Recovery (%)
I	Extract	61.5	2.2	12.3	0.2	100
II	Heat treatment	54.0	2.1	1.5	1.4	84
III	Streptomycin	62.0	2.3	1.2	1.9	105
IV	Blue agarose	2.0	3.7	12.1	3.1	55
V	Fractogel 50S	7.5	9.3	2.5	3.7	52

share an immunological epitope (45). This leaves the ligases determined by bacteriophages as the idiosyncratic examples of one member classes. Despite the lack of similarity across ligases broadly, there are features (first noticed by Barker et al. [5] and Tomkinson et al. [45a]) which seem to be conserved among all ligases (including RNA ligases). These conserved motifs (Table 3) lie in the vicinity of the (in most cases presumed) adenylylated lysine residue. The motif identified by Tomkinson et al. (45a) has been shown to contain the adenylylated lysine residue in one case. If this is generally true, then the function of the motif identified by Barker et al. (5) remains unknown. The two motifs overlap at the adenylylated lysine in the RNA ligases, complicating the analysis in those cases.

We achieved high-level expression of *T. thermophilus* ligase in *E. coli* by providing a strong promoter (p_L) to ensure transcription of the gene and a good Shine-Dalgarno sequence to ensure efficient translation of the message. The high G+C content of the gene (66%) presented two potential problems for expression: the possibility for mRNA secondary structure, and unusual codon usage. To ensure transcriptional elongation across this gene, we provided the antitermination gene *N* (32) and its site of utilization, *nutL* (34). Although we took no specific steps to address codon usage, we obtained satisfactory expression of the gene. In expressing the enzyme, we observed a strong influence of the sequence and placement of the Shine-Dalgarno sequence with respect to the initiation codon. Best results were obtained with a concocted sequence previously shown to work well in expression of foreign genes (30). This Shine-Dalgarno sequence has eight bases of complementarity to the 3′ end of 16S rRNA and results in a nine-base separation of the conserved AGGA motif from the initiation codon. This is consistent with the observations of others concerning translation initiation in *E. coli* (36).

We present a partial purification which is adequate to prepare nuclease-free ligase, although the material is not homogeneous. In the enzyme purification, a thermal denaturation step very efficiently removes the bulk of the host proteins, yielding significant purification without column chromatography. This phenomenon has also been found to occur with *Thermus* xylose isomerase cloned in *E. coli* (16). Residual nuclease activity can be removed from the ligase at this stage by using columns of modest capacity, since the material going onto the column is already highly enriched for ligase. We are developing protocols which will reproducibly yield homogeneous ligase. The availability of a plentiful source of thermostable ligase will facilitate future studies.

TABLE 3. Conserved amino acid motifs among ligases[a]

Source of ligase	Motif I	Motif II	Reference
Thermus thermophilus	118 KVDGLS	289 DGVVVK	This work
Escherichia coli	115 KLDGLA	285 DGVVIK	22
Homo sapiens	568 KYDGQR	720 EGLMVK	7
Saccharomyces cerevisiae	419 KYDGER	570 EGLMVK	5
Schizosaccharomyces pombe	416 KYDGER	567 EGLMVK	6
Vaccinia virus	231 KYDGER	377 EGLVIK	39
Phage T4	159 KADGAR	344 EGIILK	1
Phage T7	34 KYDGVR	216 EGLIVK	17
S. cerevisiae (tRNA)	114 KANGCI	110 D-VTIK	47
Phage T4 (RNA)	99 KEDGSL	94 DYILTK	31

[a] Several DNA ligases and two RNA ligases are listed along with the locations and sequences of conserved motif I (45a) and motif II (5). Lysine residues known to be adenylylated are residue 568 in the human ligase (45a) and residue 99 in phate T4 RNA ligase (31). A single amino acid gap (indicated by a dash) has been added to the *S. cerevisiae* tRNA ligase motif II to improve overall alignment.

ACKNOWLEDGMENTS

We thank M. Radjai and J. Matheny of BioTechnica International for growth of KB649/pGL628 in the fermentor.

REFERENCES

1. **Armstrong, J., R. S. Brown, and A. Tsugita.** 1983. Primary structure and genetic organization of phage T4 DNA ligase. Nucleic Acids Res. **11**:7145–7156.
2. **Arrand, J. E., A. E. Willis, I. Goldsmith, and T. Lindahl.** 1986. Different substrate specificities to the two DNA ligases of mammalian cells. J. Biol. Chem. **261**:9079–9082.
3. **Backman, K., Y.-M. Chen, and B. Magasanik.** 1981. Physical and genetic characterization of the *glnA-glnG* region of the *Escherichia coli* chromosome. Proc. Natl. Acad. Sci. USA **78**:3743–3747.
4. **Barker, D. G., A. L. Johnson, and L. H. Johnson.** 1985. An improved assay for DNA ligase reveals temperature sensitive activity in *cdc9* mutants of *Saccharomyces cerevisiae*. Mol. Gen. Genet. **200**:458–462.
5. **Barker, D. G., J. H. M. White, and L. H. Johnson.** 1985. The nucleotide sequence of the DNA ligase gene (*CDC9*) from *Saccharomyces cerevisiae*: a gene which is cell cycle regulated and induced in response to DNA damage. Nucleic Acids Res. **13**:8323–8337.
6. **Barker, D. G., J. H. M. White, and L. H. Johnson.** 1987. Molecular characterization of the DNA ligase gene CDC17 from the fission yeast *Schizosaccharomyces pombe*. Eur. J. Biochem. **162**:659–667.
7. **Barnes, D. E., L. H. Johnston, K. Kodama, A. E. Tomkinson, D. D. Lasko, and T. Lindahl.** 1990. Human DNA ligase I cDNA: cloning and functional expression in *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA **87**:6679–6683.
8. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
9. **Bolivar, F., and K. Backman.** 1980. Plasmids of *Escherichia coli* as cloning vectors. Methods Enzymol. **65**:245–267.
10. **Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
11. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–472.
12. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248–254.
13. **Clewell, D. B., and D. R. Helinski.** 1970. Properties of a supercoiled deoxyribonucleic acid protein relaxation complex and strand specificity of the relaxation event. Biochemistry **9**:4428–4440.
14. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Non-

chromosomal antibiotic resistance in bacteria: genetic transformation by R-factor DNA. Proc. Natl. Acad. Sci. USA **69**:2110–2114.

15. **Davis, R. W., M. Thomas, J. Cameron, T. P. St. John, S. Scherer, and R. A. Padgett.** 1980. Rapid DNA isolation for enzymatic and hybridization analysis. Methods Enzymol. **65**: 404–411.
16. **Dekker, K., H. Yamagata, K. Sakaguchi, and S. Udaka.** 1991. Xylose (glucose) isomerase gene from the thermophile *Thermus thermophilus*: cloning, sequencing, and comparison with other thermostable xylose isomerases. J. Bacteriol. **173**:3078–3083.
17. **Dunn, J. J., and F. W. Studier.** 1981. The complete nucleotide sequence of bacteriophage T7 DNA, and the locations of T7 genetic elements. J. Mol. Biol. **148**:303–330.
18. **Engler, M. J., and C. C. Richardson.** 1982. DNA ligases, p. 3–29. *In* P. D. Boyer (ed.), The Enzymes, vol. 15. Academic Press, New York.
19. **Fareed, G. C., E. M. Wilt, and C. C. Richardson.** 1971. Enzymatic breakage and joining of DNA. Part 8. Hybrids of ribonucleotide and deoxyribonucleotide homopoymers as substrates for polynucleotide ligase of bacteriophage T4. J. Biol. Chem. **246**:925–932.
20. **Gerna, R., P. Pienta, and R. Cote (ed.).** 1989. ATCC catalog of bacteria and bacteriophages, 17th ed., p. 317. American Type Culture Collection, Rockville, Md.

20a. **Hinkle, D. C., and C. C. Richardson.** 1975. Bacteriophage T7 replication *in vitro*: purification and properties of the gene 4 protein of bacteriophage T7. J. Biol. Chem. **250**:5523–5529.

21. **Ish-Horowicz, D., and J. F. Burke.** 1981. Rapid and efficient cosmid cloning. Nucleic Acids Res. **9**:2989–2998.
22. **Ishino, Y., H. Shinagawa, K. Makino, S. Tsunasawa, F. Sakiyama, and K. Nakata.** 1986. Nucleotide sequence of the *lig* gene and primary structure of DNA ligase of *Escherichia coli*. Mol. Gen. Genet. **204**:1–7.
23. **Keller, W., and I. Wendell.** 1975. Stepwise relaxation of supercoiled SV40 DNA. Cold Spring Harbor Symp. Quant. Biol. **39**:199–208.
24. **Knopf, K.-W.** 1977. Simple isolation method and assay for T4 DNA ligase and characterization of the purified enzyme. Eur. J. Biochem. **73**:33–38.
25. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
26. **Little, J. W., S. B. Zimmerman, C. K. Oshinsky, and M. Gellert.** 1967. Enzymatic joining of DNA strands. II. An enzyme adenylate intermediate in the diphosphopyridine nucleotide dependent enzyme DNA ligase reaction in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **58**:2004–2011.
27. **Miller, J. H.** 1972. Experiments in molecular genetics, p. 431–435. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
28. **Olivera, B. M., and I. R. Lehman.** 1967. Diphosphopyridine nucleotide: a cofactor for the polynucleotide joining enzyme from *Escherichia coli*. Proc. Natl. Acad. Sci. USA **57**:1700–1704.
29. **Oshima, T., and K. Imahori.** 1974. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa. Int. J. Syst. Bacteriol. **24**:102–112.
30. **Pallas, D. C., C. Schley, M. Mahoney, E. Harlow, B. S. Schaffhausen, and T. M. Roberts.** 1986. Polyomavirus small t antigen: overproduction in bacteria, purification, and utilization for monoclonal and polyclonal antibody production. J. Virol. **60**: 1075–1084.
31. **Rand, K. N., and M. J. Gait.** 1984. Sequence and cloning of bacteriophage T4 gene 63 encoding RNA ligase and tail fiber attachment activities. EMBO J. **3**:397–402.
32. **Roberts, J. W.** 1969. Termination factor for RNA synthesis. Nature (London) **224**:1168–1174.
33. **Roberts, T. M., S. L. Swanberg, A. Poteete, G. Reidel, and K. Backman.** 1980. A plasmid cloning vector allowing a positive selection for inserted fragments. Gene **12**:123–127.
34. **Salstrom, J. S., and W. Szybalski.** 1978. Coliphage lambda *nutL*$^-$: a unique class of mutants defective in the site of gene N product utilization for antitermination of leftward transcription. J. Mol. Biol. **124**:195–221.
35. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with DNA chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
36. **Schneider, T. D., G. D. Stormo, L. Gold, and A. Ehrenfeucht.** 1986. Information content of binding sites on nucleotide sequences. J. Mol. Biol. **188**:415–431.
37. **Sgaramella, V., and H. G. Khorana.** 1972. Studies on polynucleotides. CXVI. A further study of the T4 ligase catalyzed joining of DNA at base paired ends. J. Mol. Biol. **72**:493–502.
38. **Shine, J., and L. Dalgarno.** 1974. The 3′ terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
39. **Smith, G. L., Y. S. Chan, and S. M. Kerr.** 1989. Transcriptional mapping and nucleotide sequence of a vaccinia virus gene encoding a polypeptide with extensive homology to DNA ligases. Nucleic Acids Res. **17**:9051–9062.
40. **Soderhall, S., and T. Lindahl.** 1976. DNA ligases of eukaryotes. FEBS Lett. **67**:1–8.
41. **Takahashi, M., and T. Uchida.** 1986. Thermophilic HB-8 DNA ligase: effects of polyethylene glycols and polyamines on blunt end ligation of DNA. J. Biochem. **100**:123–131.
42. **Takahashi, M., E. Yamaguchi, and T. Uchida.** 1984. Thermophilic DNA ligase. J. Biol. Chem. **259**:10041–10047.
43. **Takahashi, M., E. Yamaguchi, and T. Uchida.** 1986. Purification of HB-8 DNA ligase by red sepharose chromatography. Agric. Biol. Chem. **50**:1333–1334.
44. **Thuring, R. W. J., J. P. M. Sanders, and P. Borst.** 1975. A freeze squeeze method for recovering long DNA from agarose gels. Anal. Biochem. **66**:213–220.
45. **Tomkinson, A. E., D. D. Lasko, G. Daly, and T. Lindahl.** 1990. Mammalian DNA ligases. J. Biol. Chem. **265**:12611–12617.

45a. **Tomkinson, A. E., N. F. Totty, M. Ginsburg, and T. Lindahl.** 1991. Location of the active site for enzyme-adenylate formation in DNA ligases. Proc. Natl. Acad. Sci. USA **88**:400–404.

46. **Weiss, B., and C. C. Richardson.** 1967. Enzymatic breakage and joining of deoxyribonucleic acid. I. Repair of single stranded breaks in DNA by an enzyme system from *Escherichia coli* infected with T4 bacteriophage. Proc. Natl. Acad. Sci. USA **57**:1021–1028.
47. **Xu, Q., D. Teplow, T. D. Lee, and J. Abelson.** 1990. Domain structure in yeast tRNA ligase. Biochemistry **29**:6132–6138.
48. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
49. **Zimmerman, S. B., and C. K. Oshinsky.** 1969. Enzymatic joining of deoxyribonucleic acid strands. J. Biol. Chem. **244**:4689–4695.
50. **Zimmerman, S. B., and B. H. Pheiffer.** 1983. Macromolecular crowding allows blunt-end ligation by DNA ligases from rat liver or *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80**:5852–5856.