Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 400–404, January 1991
Biochemistry

Location of the active site for enzyme-adenylate formation in DNA ligases

(DNA replication/lysine-adenylate/pyridoxal 5'-phosphate/evolutionarily conserved enzyme)

ALAN E. TOMKINSON*, NICHOLAS F. TOTTY†, MICHELLE GINSBURG*, AND TOMAS LINDAHL*

*Imperial Cancer Research Fund, Clare Hall Laboratories, South Mimms, Hertfordshire EN6 3LD, United Kingdom; and †Ludwig Institute for Cancer Research, Middlesex Hospital, University College Branch, London W1P 8BT, United Kingdom

Communicated by Charles C. Richardson, October 11, 1990 (received for review September 20, 1990)

ABSTRACT **The enzyme-AMP reaction intermediate of the 102-kDa bovine DNA ligase I was digested with trypsin, and the adenylylated peptide was isolated by chromatography under conditions that maintain the acid-labile phosphoramidate bond. Microsequencing of the peptide showed that it contains an internal trypsin-resistant lysine residue, as expected for the site of adenylylation. Inhibition of DNA ligase I activity by pyridoxal 5'-phosphate also indicated the presence of a reactive lysine residue in the catalytic domain of the enzyme. Comparison of the known primary structures of several other DNA ligases with the adenylylated region of mammalian DNA ligase I allows their active sites to be tentatively assigned by sequence homology. The ATP-dependent DNA ligases of mammalian cells, fission yeast, budding yeast, vaccinia virus, and bacteriophages T3, T4, and T7 contain the active site motif Lys-Tyr/Ala-Asp-Gly-(Xaa)-Arg, with the reactive lysine residue flanked by hydrophobic amino acids. The distance between the postulated adenylylation site and the carboxyl terminus of the polypeptide is very similar in these ATP-dependent DNA ligases, whereas the size of the amino-terminal region is highly variable.**

DNA ligases catalyze the formation of phosphodiester bonds at single-strand breaks with adjacent 3' hydroxyl and 5' phosphate termini in double-helical DNA. In the first step of the joining reaction, DNA ligase interacts with a nucleotide derivative, either ATP for virus-encoded and eukaryotic enzymes or NAD^+ for bacterial enzymes, to form a covalent enzyme-adenylate intermediate with the concomitant release of pyrophosphate or NMN (1). A lysine-AMP residue with a phosphoramidate bond between the nucleotide and the ε-amino group of lysine has been isolated after proteolytic degradation of the *Escherichia coli* and phage T4 DNA ligase-adenylate intermediates, demonstrating that a lysine residue in the protein is adenylylated (2). The chemical stability and sensitivity to acidic hydroxylamine of the complex also indicate the presence of a phosphoramidate bond (2). Similar results have been obtained with the enzyme-adenylate form of mammalian DNA ligase I (3). The latter enzyme is the major DNA ligase activity in proliferating mammalian cells (4–6).

DNA sequences have been determined for the genes encoding the DNA ligases of bacteriophages T4 (7), T7 (8), and T3 (9); *E. coli* (10); *Saccharomyces cerevisiae* (11); *Schizosaccharomyces pombe* (12); and vaccinia virus (13) as well as for human DNA ligase I cDNA (14). With regard to the T4 and *E. coli* DNA ligases, overexpression of the cloned genes has facilitated purification of the proteins (15, 16). Nevertheless, the position of the reactive lysine residue has not been determined in any of these DNA ligase sequences. The location of the adenylylation site in the related enzyme T4 RNA ligase is known, but this protein only shows extremely weak sequence homology with T4 DNA ligase, not allowing a prediction of the active site of the latter enzyme (17, 18). In the present work, we have isolated and sequenced an adenylylated tryptic peptide from a bovine DNA ligase I-AMP complex, determined its corresponding position in the cDNA encoding human DNA ligase I, and compared it with related regions in other DNA ligases.

EXPERIMENTAL PROCEDURES

Adenylylation of Bovine DNA Ligase I. The enzyme was purified to homogeneity from calf thymus as described (6). Protein concentrations were determined by the method of Bradford (19) with bovine serum albumin used as a standard and a molecular mass of 102 kDa for DNA ligase I (14). Prior to use, DNA ligase I (fraction VII) was dialyzed extensively at 4°C against 25 mM Bicine-NaOH, pH 8.0/1 mM EDTA/0.5 mM dithiothreitol (DTT)/1 M NaCl and then against the same buffer containing only 0.1 M NaCl. The adenylylation reaction mixture (150 μl) contained 60 mM Tris·HCl (pH 8.0), 10 mM $MgCl_2$, 0.5 mM DTT, 30 mM NaCl, 50 μg of bovine serum albumin per ml, 5 μM ATP, 20 μCi of [α-^{32}P]ATP (3000 Ci/mmol; 1 Ci = 37 GBq; Amersham), and 8 pmol of DNA ligase I. After incubation at 25°C, aliquots (10 μl) were removed at various time intervals from 10 sec to 20 min and the reactions were stopped by the addition of 2 μl of 0.6 M Tris·HCl, pH 6.8/165 mM DTT/10% SDS. The samples were heated at 90°C for 10 min and then electrophoresed through a 7.5% SDS/polyacrylamide gel (20). After fixing in 10% acetic acid, the gel was dried and polypeptide-adenylate complexes were detected by autoradiography. The amount of DNA ligase I-adenylate was quantitated by scanning densitometry (LKB Ultrascan XL).

Inhibition of Ligase Adenylylation by Pyridoxal Phosphate. Pyridoxal 5'-phosphate (PLP; Sigma) was dissolved in 25 mM Bicine-NaOH (pH 8.0) to 10 mM immediately prior to use. Bovine DNA ligase I (5 pmol) was incubated in a reaction mixture (50 μl) containing 25 mM Bicine-NaOH (pH 8.0), 0.2 mM DTT, 40 mM NaCl, 10 mM $MgCl_2$, and 0, 50, 100, or 200 μM PLP at 25°C. At various time intervals, 8-μl samples were removed and added to 2 μl of 25 μM ATP containing 1 μCi of [α-^{32}P]ATP (3000 Ci/mmol). After incubation for 15 sec, reactions were stopped and enzyme-adenylate formation was quantitated as described above. Because of the short incubation period for the adenylylation reaction, it was not necessary to stabilize the metastable enzyme-PLP complexes by reduction with $NaBH_4$ (21) prior to assay.

Isolation of Adenylylated Peptide. DNA ligase I (fraction VII; 1.6 nmol) was incubated at 20°C for 30 min in 3 ml of 60 mM Tris·HCl, pH 8.0/10 mM $MgCl_2$/0.5 mM DTT/5 μM ATP/100 μCi of [2,5,8-^{3}H]ATP (70 Ci/mmol; Amersham). The reaction was terminated by the addition of 150 μl of 0.2

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: DTT, dithiothreitol; PLP, pyridoxal 5'-phosphate.

M EDTA (pH 8.0) and excess ATP was removed by extensive dialysis at 4°C against 100 mM Tris·HCl, pH 8.2/20 mM NaCl/0.5 mM DTT. Adenylylated DNA ligase I was incubated with 3 μg of trypsin (sequencing grade; Boehringer Mannheim) overnight at 37°C. The resulting peptides were fractionated by HPLC on a C_8/RP300 column (220 × 2.1 mm; Applied Biosystems) using a linear gradient from 30 mM sodium acetate, pH 5.5/1% acetonitrile to 35 mM sodium acetate, pH 5.5/80% acetonitrile at a flow rate of 0.2 ml/min. The eluate was monitored for absorbance at 214, 260, and 280 nm. Fractions (100 μl) were collected and aliquots were removed to determine radioactivity. The fractions containing the main peak of radioactivity were pooled and an aliquot was taken for microscale sequencing by automated Edman degradation (22). After removal of acetonitrile by evaporation, peptides in the pooled fraction were refractionated by HPLC on a C_{18} column (100 × 2.1 mm; Applied Biosystems) using a linear gradient of 1–80% acetonitrile in 30 mM sodium acetate (pH 5.5) at a flow rate of 0.2 ml/min. Fractions (50 μl) were collected and aliquots were removed to determine radioactivity. Peptides were then subjected to microscale sequencing.

Data Base Searches and Hydropathy Profiles. The OWL 9 protein data base was searched using the PROSRCH program (23, 24) based on the Smith and Waterman algorithm (25). The cDNA sequence of human DNA ligase I corresponding to the adenylylated peptide from bovine DNA ligase I was used as a probe; 60, 100, and 250 PAM (percent accepted mutation) tables (26) were used to score conservative substitutions and stringent penalties were applied to minimize gaps in any similarities reported. Hydropathy profiles were generated by using the program GeneWorks (IntelliGenetics) applying the Kyte and Doolittle algorithm with amino- and carboxyl-terminal spans of five residues. The predicted amino acid sequences were compiled from the Protein Identification Resource 24.0, Swiss-Prot 14.0, and EMBL 23.0 data bases.

RESULTS

Formation of Enzyme-Adenylate Complex. The reaction of bovine DNA ligase I with ATP to generate a covalent enzyme-adenylate complex proceeded rapidly at 25°C and was essentially complete after 10–20 sec. The reaction intermediate was then stable for at least 20 min, but the complex was readily dissociated by the addition of pyrophosphate or nicked DNA as described (3). The K_m for ATP in the adenylylation of DNA ligase I was ≈0.6 μM (data not shown).

Inhibition of DNA Ligase Adenylylation by PLP. Several enzymes that interact with nucleotide substrates contain an unusually reactive lysine residue in their active site, which can form a Schiff base with PLP (21, 27, 28). Thus, PLP can act as a group-specific reagent, and its inhibition of enzyme activity is indicative of the presence of a reactive and essential lysine moiety. Formation of the DNA ligase I-adenylate complex was 85% inhibited by a 5-min preincubation with 50 μM PLP and it was >95% inhibited with 200 μM PLP (Fig. 1). The results are similar to those obtained with DNA polymerase α (21) and adenylosuccinate synthetase (27).

Isolation of an Adenylylated Tryptic Peptide from DNA Ligase I. Homogeneous bovine DNA ligase I (1.6 nmol) was adenylylated with [^{3}H]ATP and then digested with trypsin. The predicted amino acid sequence from the human DNA ligase I cDNA (14) indicates that ≈103 tryptic peptides would be generated. Chromatographic isolation of a single peptide containing a radioactive adenylyl group was attempted. Since the amino acid-AMP bond is acid labile (2, 3), reverse-phase HPLC fractionation was carried out at pH 5.5, rather than in the more commonly used trifluoroacetic acid-based solvent system. After initial separation on a C_8 column, a single major

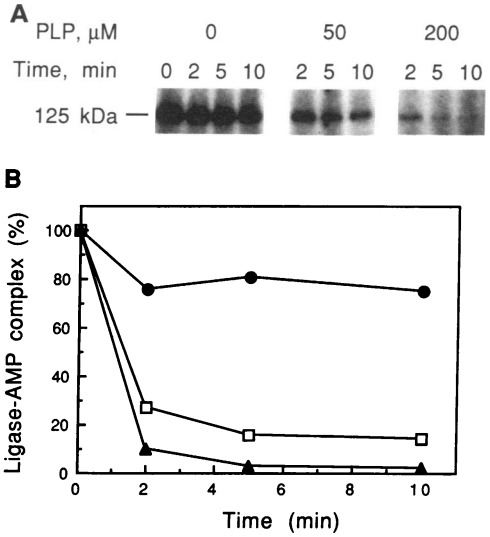

FIG. 1. Inhibition of DNA ligase I-adenylate complex formation by PLP. Bovine DNA ligase I was incubated with PLP for the times indicated prior to adenylylation. The ligase-adenylate complex [125 kDa as measured by SDS/PAGE (6)] was detected by autoradiography for 3–15 hr. (*A*) The relevant portion of the autoradiogram (15-hr exposure) is shown. Each aliquot contained 0.8 pmol of DNA ligase I. (*B*) The inhibition of adenylylation by PLP was quantitated by scanning densitometry. The data were obtained with no PLP (●), 50 μM PLP (□), or 200 μM PLP (▲).

peak of radioactive material was observed, which cochromatographed with a minor peak of A_{214}-absorbing peptides (Fig. 2*A*). Several minor peaks of radioactive material (Fig. 2*A*) were ascribed to partial digestion products. The fractions containing the major peak of radioactivity were pooled. The ultraviolet absorption spectrum of this material was consistent with the presence of adenylyl residues. Amino acid sequence analysis of this pooled fraction indicated that it contained three different peptides.

On further chromatographic purification of this fraction, using a C_{18} column, the radioactive material coeluted with a split peak of A_{214}-absorbing material, peptides IIa and IIb (Fig. 2*B*). Two additional, well separated peaks of A_{214}-absorbing material, peptides I and III, were also observed (Fig. 2*B*). The ultraviolet absorption spectra of peptides IIa and IIb were identical and exhibited an A_{260}/A_{280} ratio of 2.0, indicative of the presence of adenylylated material. The nonradioactive peptide I had an A_{260}/A_{280} ratio of 0.56, close to the predicted value for a tyrosine-containing peptide. Treatment of peptides IIa/IIb with acidic hydroxylamine (2, 3) released >90% of the radioactive material, indicative of the presence of an amino acid-AMP residue with a phosphoramidate bond.

Amino Acid Sequence Analysis. The sequences of peptides I, IIa, and IIb were determined by automated Edman degradation. Peptide I had the sequence Val-Pro-Tyr-Leu-Ala-Val-Ala-Arg. This sequence occurs in the amino-terminal half of mammalian DNA ligase I, in a region that shows little or no homology with other DNA ligases (14). Peptides IIa and IIb yielded the same sequence, Phe-Glu-Glu-Ala-Ala-Phe-Thr-(Xaa)-Glu-Tyr-Lys-Tyr-Asp-Gly-Gln-Arg. The absence of a signal at position 8 may have reflected the occurrence of a cysteine residue, since the phenylthiohydantoin-derived cysteine is unstable. Moreover, partial oxidation of a cysteine residue could account for the split appearance of peak II. The amount of radioactive material in the peak fraction corresponded to 25 pmol of [^{3}H]AMP. The yield of phenylthiohydantoin-derived amino acids during Edman degradation of the same fraction indicated that 25–35 pmol of peptide II was present. Moreover, the A_{260}/A_{280} ratio of 2.0 observed for

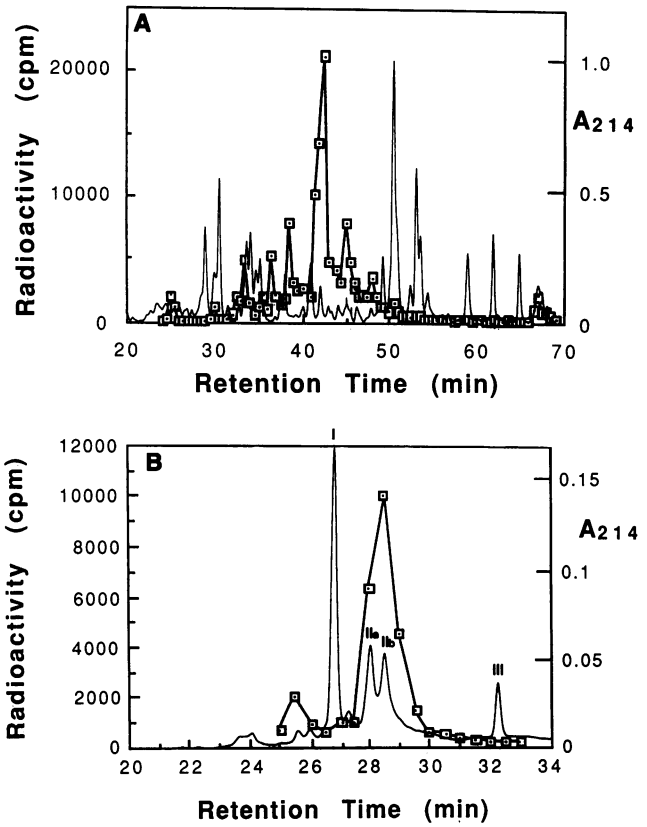

FIG. 2. Purification of adenylylated peptide from trypsin-digested DNA ligase I. Homogeneous bovine DNA ligase I was adenylylated with [^{3}H]ATP and then digested with trypsin. (*A*) Tryptic peptides separated by reverse-phase HPLC on a C_8 column. (*B*) Further separation of fractions containing the adenylylated peptide by reverse-phase HPLC on a C_{18} column. Peptides were detected by absorbance at 214 nm (——) and adenylylated moieties were detected by liquid scintillation counting (⊡).

peptide II was close to the theoretical value of 2.2 calculated for a peptide containing two tyrosine residues and one AMP residue. We conclude that peptide II contained a single adenylyl group.

The acidic conditions used in Edman degradation cleave a phosphoramidate bond, so the AMP residue was removed from the peptide during the first degradation cycle. However, the presence of an internal, trypsin-resistant single lysine residue in peptide II provides good evidence that this residue was the site of adenylylation. Trypsin resistance of a lysine-AMP residue has also been observed for the adenylylated reaction intermediate of the yeast tRNA splicing enzyme tRNA ligase (29).

Active Site Homology Between DNA Ligases. The sequence of the adenylylated tryptic peptide of bovine DNA ligase I corresponds to residues 558–573 of the predicted amino acid sequence of human DNA ligase I, Phe-Glu-Glu-Ala-Ala-Phe-Thr-Cys-Glu-Tyr-Lys-Tyr-Asp-Gly-Gln-Arg, which has a cysteine residue at position 8 (14). Moreover, the cDNA sequence predicts that residue 557 is a trypsin-sensitive arginine residue, so the bovine and human enzymes appear identical in this region.

An amino acid sequence identical at 13 of 16 positions with the adenylylated peptide derived from bovine DNA ligase I occurs at residues 406–421 in DNA ligase from *Sc. pombe* (12). The overall amino acid identity between these two DNA ligases is 44% (14), and the strong local sequence conservation (81%) observed here implies that Lys-416 is the adenylylated lysine residue in the fission yeast enzyme. Partial homology was also observed between the region surrounding the reactive lysine residue in the mammalian enzyme and sequences present in other DNA ligases. Thus, a putative assignment of the location of the active site lysine can be made for these enzymes (Fig. 3). The active site motif Lys-(Tyr/Ala)-Asp-Gly-(Xaa)-Arg is present consistently in the ATP-dependent DNA ligases. A shorter version, Lys-(Leu)-Asp-Gly, is also present in the NAD^+-dependent DNA ligase of *E. coli*. The T4 RNA ligase shows little or no overall sequence homology with DNA ligases (17, 18), although it acts by a very similar reaction mechanism, but the known adenylylation site of the RNA ligase (18) also contains the Lys-(Xaa)-Asp-Gly motif (Fig. 3).

The molecular mass of DNA ligases varies between 40 kDa for the T3 ligase (9) and 102 kDa for the mammalian ligase I (14). However, when the sequences were aligned with regard to the active site lysine, this amino acid residue was observed to be present consistently at a distance of 332 ± 20 residues from the carboxyl terminus of the seven ATP-dependent DNA ligases (Fig. 4). Consequently, the great size variation between these enzymes is accounted for by very different lengths of the amino-terminal regions. Armstrong *et al.* (7) noted previously that on alignment of homologous sequences in the T4 and T7 DNA ligases the carboxyl termini coincided almost exactly, indicating the presence of a larger domain in the amino-terminal part of the T4 DNA ligase. A comparison of the amino acid sequences between the putative active site and the carboxyl termini of the ATP-dependent DNA ligases (7, 11–14) showed several regions of sequence homology. Moreover, the similarity of the hydrophobicity plots of the same amino acid sequences (Fig. 4) indicated that even in regions of little sequence homology, there is conservation of structural features in the carboxyl-terminal domains of these enzymes.

		⬇	
H. sapiens	543	MLAHPTRGISEVLKRFEEAAFTCEY**KYDGQRA**QIHALEGGEVKIFSRNQEDNTGKYPDI	601
Sch. pombe	391	MLAKPTKQISEVLNTFDQAAFTCEY**KYDGERA**QVHFTEDGKFYVFSRNSENMSVRYPDI	449
S. cerevisiae	394	MLAKPTKAINEVLDRFQGETFTSEY**KYDGERA**QVHLLNDGTMRIYSRNGENMTERYPEI	452
Vaccinia	207	MLA ESCDSVNKAFKKFPSGMFAEV**KYDGER**VQVH KNNNEFAFFSRNMKPVLSHKVDY	263
T3	9	FKAVSFVESAVKKALETSGYLIADC**KYDGVR**GNIVVDNVAEAAWLSRVSKFIPALEHLN	67
T7	9	FKAVSFVESAIKKALDNAGYLIAEI**KYDGVR**GNICVDNTANSYWLSRVSKTIPALEHLN	67
T4	136	MLASSYDEKGINKNIKFPAF AQL**KADGAR**CFAEVRGDELDDVRLLSRAGNEYLGLDL	192
E.coli	91	FLAFNKRVQDRLKNNEKVTWC CEL**KLDGLA**VSILYENGVLVSAATRGDGTTGEDITSN	148
T4 RNA	74	EKFFNLNENPFTMNIDLNDVDYILT**KEDGSL**VSTYLDGDEILFKSKGSIKSEQALMANG	132

FIG. 3. Alignment of sequences from DNA and RNA ligases showing partial homology with the active site region of mammalian DNA ligase I. The alignment with the apparently related sequences in the DNA ligase genes of *Sc. pombe* (12); *S. cerevisiae* (11); vaccinia virus (13); bacteriophages T3 (8), T7 (9), and T4 (7); and *E. coli* (10), and the RNA ligase gene of bacteriophage T4 (17) are shown. Gaps have been inserted to maximize homology and the region containing the most highly conserved residues is indicated in boldface. Arrow indicates the postulated or known site of adenylylation.

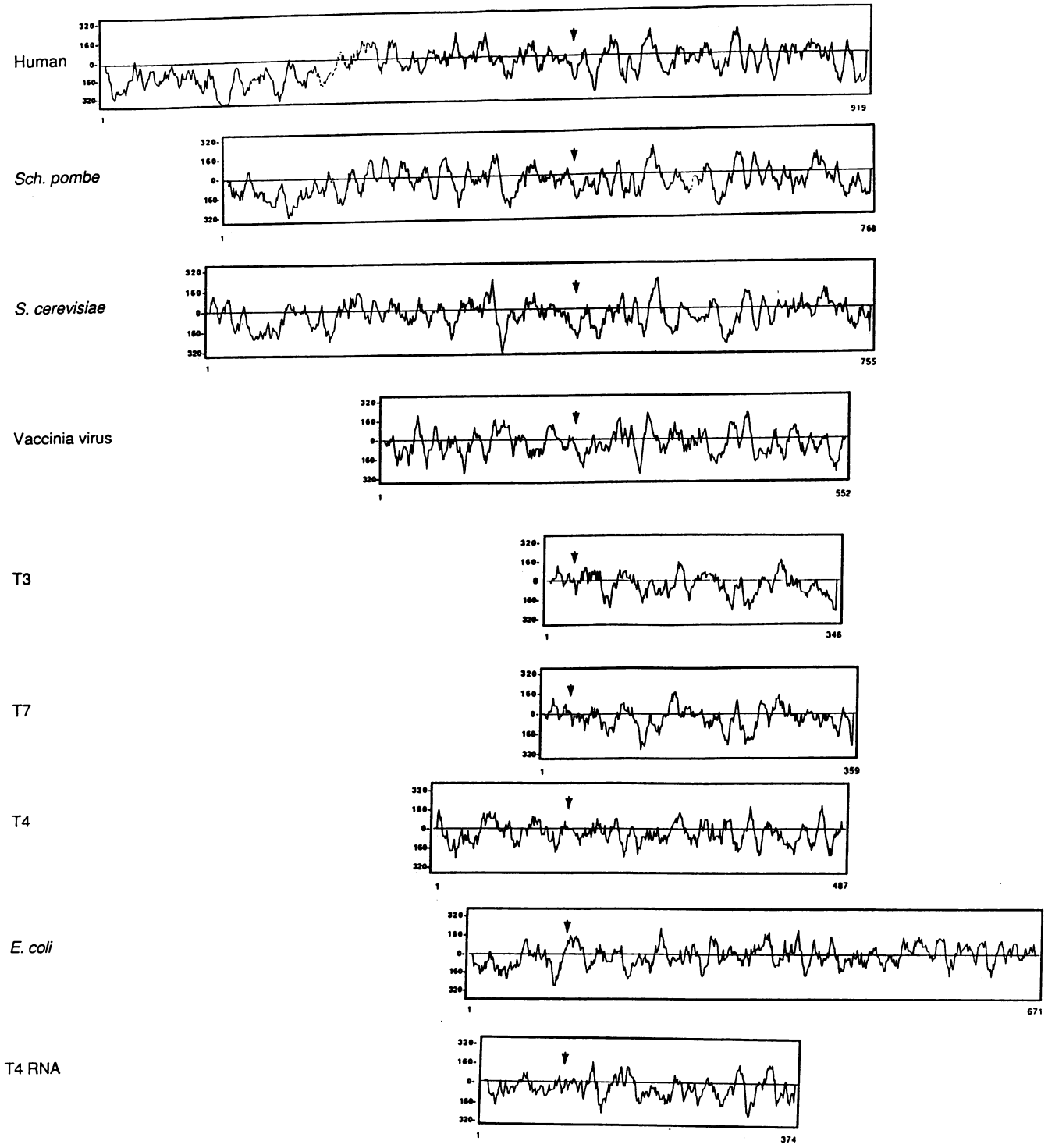

FIG. 4. Hydropathy profiles of DNA and RNA ligases. The profiles were generated, using the Kyte and Doolittle algorithm, for the complete sequences of the ligases described in Fig. 3. Hydrophilic regions are below and hydrophobic regions are above the line. The position of the presumed active site lysine is marked with an arrow and the plots are aligned with reference to this residue.

E. coli DNA ligase uses NAD^+ instead of ATP as a cofactor. A data base search with the sequence of the carboxyl-terminal one-third of this enzyme (Fig. 4) indicated weak homology with several NAD^+-dependent enzymes, suggesting that the carboxyl-terminal part of *E. coli* DNA ligase contains an NAD^+-binding site.

DISCUSSION

The presence of an amino acid-adenylate residue containing a phosphoramidate bond, and having the chromatographic properties of lysine-AMP, in the enzyme-adenylate reaction intermediates of *E. coli* and T4 DNA ligases was first demonstrated by Gumport and Lehman (2). Here, we have extended these observations to a mammalian DNA ligase by showing that (*i*) adenylylation of bovine DNA ligase I is inhibited by the lysine-specific reagent PLP; (*ii*) the amino acid sequence of the adenylylated peptide generated by tryptic digestion of bovine DNA ligase I contains an internal trypsin-resistant lysine residue; (*iii*) the absorption spectrum of the adenylylated peptide and separate quantitation of the AMP and peptide entities yield a 1:1 molar ratio; (*iv*) the

sensitivity to acidic hydroxylamine of the bovine peptide-adenylate complex indicates the presence of a phosphoramidate bond between the internal lysine and the adenylyl group; (*v*) the adenylylated peptide contains a single lysine but no histidine residue.

The formation of an enzyme-adenylate reaction intermediate by mammalian DNA ligase I was prevented by PLP. Most enzymes are not susceptible to PLP, and such inhibition is indicative of the presence of a reactive lysine residue in the vicinity of the active site. Thus, PLP modification of one susceptible residue of the 22 lysines in *E. coli* adenylosuccinate synthetase causes inhibition of enzymatic activity (27), lamb kidney Na^+,K^+-ATPase is inhibited by the reaction of PLP with a single essential lysine residue in the putative ATP-binding site (28), and *Drosophila melanogaster* DNA polymerase α exhibits two binding sites for PLP located in the nucleotide binding sites of the polymerase and primase subunits, respectively (21, 30). PLP also can react with the amino-terminal residue of proteins. In mammalian DNA ligase I, the amino-terminal region is not required for catalytic activity (6), so the enzyme inhibition by PLP may be ascribed to the occurrence of a reactive lysine residue in the catalytic domain of the enzyme. These data are in agreement with the ε-amino group of a lysine residue being the nucleophilic catalyst in the formation of the ligase-AMP intermediate and do not support a less likely, alternative model in which the AMP residue would bind to the imidazole N^3 position of a histidine residue, followed by migration of the adenylyl group during isolation of the enzyme-AMP complex (2, 31).

The adenylylated amino acid residue in mammalian DNA ligase I has been identified as Lys-568. In spite of the extensive studies on DNA ligases over the past 25 years, the active site for the formation of the first reaction intermediate, enzyme-adenylate, has not been mapped previously in a DNA ligase. Barker *et al.* (12) observed that a stretch of 24 amino acid residues close to the carboxyl terminus was almost identical between the DNA ligases of *S. cerevisiae* (residues 568–591) and *Sc. pombe* (residues 565–588). Since this was the most highly conserved region between the two proteins, they suggested that a lysine residue in this region might be the site of adenylylation. However, the colinear sequence is not well conserved in the partly homologous vaccinia virus DNA ligase (13) and mammalian DNA ligase I (14), and no direct evidence has been obtained that this particular sequence contains the reactive lysine moiety. Instead, we propose (Fig. 3) that the active residue in *Sc. pombe* DNA ligase is Lys-416, and in *S. cerevisiae* it is Lys-419.

The ATP-dependent DNA ligases of eukaryotes and viruses contain the sequence Lys-(Tyr/Ala)-Asp-Gly-(Xaa)-Arg, with the lysine residue flanked by hydrophobic residues. By homology with mammalian DNA ligase I, this short region may be assigned as the site of adenylylation. The known active site for enzyme-adenylate formation in T4 RNA ligase (18) contains the same Lys-(Xaa)-Asp-Gly motif. Another distantly related protein, the yeast splicing enzyme tRNA ligase, has a similar sequence at its site of adenylylation, Ile-Lys-Ala-Asn-Gly (29), and this further supports the present assignment of the active site location in DNA ligases. Heaphy *et al.* (32) demonstrated by site-specific mutagenesis at the active site of T4 RNA ligase that replacement of Lys-99 with a histidine or asparagine residue caused total loss of activity. Furthermore, conservation of the adjacent Asp-101 residue in the Lys^{99}-(Xaa)-Asp^{101}-Gly^{102} sequence was essential for the joining function of the enzyme but not for enzyme-adenylate formation, whereas a number of amino acid substitutions could be introduced at position 100 without functional impairment. Similar site-specific mutagenesis to alter different amino acid residues at the active site of a DNA ligase can now be performed.

The distance between the presumed active site lysine and the carboxyl terminus of the polypeptide is remarkably similar between all ATP-dependent DNA ligases. The T3 and T7 DNA ligases, which have the putative active lysine close to their amino terminus, may define a minimum size for a catalytically active ATP-dependent DNA ligase. The highly variable amino-terminal regions of the larger DNA ligases presumably are required in other roles, such as protein–protein interactions (12, 14). In agreement with this notion, a catalytically active domain of mammalian DNA ligase I, obtained by partial proteolytic digestion, represents the carboxyl-terminal region of the enzyme (6).

We thank Dr. M. D. Waterfield for discussions and G. Daly for technical assistance.

1. Engler, M. J. & Richardson, C. C. (1982) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 15B, pp. 3–29.
2. Gumport, R. I. & Lehman, I. R. (1971) *Proc. Natl. Acad. Sci. USA* **68**, 2559–2563.
3. Soderhall, S. & Lindahl, T. (1973) *J. Biol. Chem.* **248**, 672–675.
4. Soderhall, S. & Lindahl, T. (1975) *J. Biol. Chem.* **250**, 8438–8444.
5. Teraoka, H. & Tsukada, K. (1982) *J. Biol. Chem.* **257**, 4758–4763.
6. Tomkinson, A. E., Lasko, D. D., Daly, G. & Lindahl, T. (1990) *J. Biol. Chem.* **265**, 12611–12617.
7. Armstrong, J., Brown, R. S. & Sugita, A. (1983) *Nucleic Acids Res.* **11**, 7145–7156.
8. Dunn, J. J. & Studier, F. W. (1981) *J. Mol. Biol.* **148**, 303–330.
9. Schmitt, M. P., Beck, P. J., Kearney, C. A., Spence, J. L., DiGiovanni, D., Condreay, P. & Molineux, I. J. (1987) *J. Mol. Biol.* **193**, 479–495.
10. Ishino, Y., Shinagawa, H., Makino, K., Tsunasawa, S., Sakiyama, F. & Nakata, A. (1986) *Mol. Gen. Genet.* **204**, 1–7.
11. Barker, D. G., White, J. H. M. & Johnston, L. H. (1985) *Nucleic Acids Res.* **13**, 8323–8337.
12. Barker, D. G., White, J. H. M. & Johnston, L. H. (1987) *Eur. J. Biochem.* **162**, 659–667.
13. Smith, G. L., Chan, Y. S. & Kerr, S. M. (1989) *Nucleic Acids Res.* **17**, 9051–9062.
14. Barnes, D. E., Johnston, L. H., Kodama, K., Tomkinson, A. E., Lasko, D. D. & Lindahl, T. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6679–6683.
15. Murray, N. E., Bruce, S. A. & Murray, K. (1979) *J. Mol. Biol.* **132**, 493–505.
16. Panasenko, S. M., Alzard, R. J. & Lehman, I. R. (1978) *J. Biol. Chem.* **253**, 4590–4592.
17. Rand, K. N. & Gait, M. J. (1984) *EMBO J.* **3**, 397–402.
18. Thogerson, H. C., Morris, H. R., Rand, K. N. & Gait, M. J. (1985) *Eur. J. Biochem.* **147**, 325–329.
19. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254.
20. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
21. Diffley, J. F. X. (1988) *J. Biol. Chem.* **263**, 14669–14677.
22. Downward, J., Yarden, Y., Mayes, E., Scrace, G., Totty, N., Stockwell, P., Ullrich, A., Schlessinger, J. & Waterfield, M. D. (1984) *Nature (London)* **307**, 521–527.
23. Collins, J. F., Coulson, A. F. W. & Lyall, A. (1988) *Comput. Appl. Biosci.* **4**, 67–71.
24. Coulson, A. F. W., Collins, J. F. & Lyall, A. (1987) *Comput. J.* **30**, 420–424.
25. Smith, T. F. & Waterman, M. S. (1981) *J. Mol. Biol.* **147**, 195–197.
26. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. M. (1978) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. (Natl. Biomed. Res. Found., Silver Spring, MD), Vol. 5, Suppl. 3, pp. 345–352.
27. Dong, Q. & Fromm, H. J. (1990) *J. Biol. Chem.* **265**, 6235–6240.
28. Hinz, H. R. & Kirley, T. L. (1990) *J. Biol. Chem.* **265**, 10260–10265.
29. Xu, Q., Teplow, D., Lee, T. D. & Abelson, J. (1990) *Biochemistry* **29**, 6132–6138.
30. Diffley, J. F. X. (1988) *J. Biol. Chem.* **263**, 19126–19131.
31. Yang, S. L. & Frey, P. A. (1979) *Biochemistry* **18**, 2980–2984.
32. Heaphy, S., Singh, M. & Gait, M. J. (1987) *Biochemistry* **26**, 1688–1696.