Nucleic Acids Research, 1997, Vol. 25, No. 4 727–734

Domain structure of vaccinia DNA ligase

JoAnn Sekiguchi and Stewart Shuman*

Molecular Biology Program, Sloan-Kettering Institute, New York, NY 10021, USA

Received November 18, 1996; Revised and Accepted December 31, 1996

ABSTRACT

The 552 amino acid vaccinia virus DNA ligase consists of three structural domains defined by partial proteolysis: (i) an amino-terminal 175 amino acid segment that is susceptible to digestion with chymotrypsin and trypsin; (ii) a protease-resistant central domain that contains the active site of nucleotidyl transfer (Lys-231); (iii) a protease-resistant carboxyl domain. The two protease-resistant domains are separated by a protease-sensitive interdomain bridge from positions 296 to 307. Adenylyltransferase and DNA ligation activities are preserved when the N-terminal 200 amino acids are deleted. However, the truncated form of vaccinia ligase has a reduced catalytic rate in strand joining and a lower affinity for DNA than does the full-sized enzyme. The 350 amino acid catalytic core of the vaccinia ligase is similar in size and protease-sensitivity to the full-length bacteriophage T7 DNA ligase.

INTRODUCTION

The ATP-dependent DNA ligases catalyze the joining of 5′ phosphate-terminated donor strands to 3′ hydroxyl-terminated acceptor strands via three sequential nucleotidyl transfer reactions (1,2). In the first step, attack on the α-phosphate of ATP by ligase results in liberation of pyrophosphate and formation of a covalent intermediate in which AMP is linked to the ε-amino group of a lysine. The nucleotide is then transferred to the 5′ end of the donor polynucleotide to form DNA-adenylate, an inverted (5′)–(5′) pyrophosphate bridge structure, AppN. Attack by the 3′ OH of the acceptor strand on the DNA-adenylate joins the two polynucleotides and liberates AMP.

Animal cells contain multiple ATP-dependent DNA ligase isozymes encoded by at least three genes (3–5). ATP-dependent DNA ligases are also encoded by fungi and plants, by eukaryotic DNA viruses, by the T-odd and T-even bacteriophages, and by archaea (6). The ATP-dependent DNA ligases belong to a superfamily of covalent nucleotidyl transferases that includes the GTP-dependent eukaryotic mRNA capping enzymes (7).

The ligase/capping enzyme superfamily is defined by a set of six short motifs arrayed in the same order and with similar spacing in nearly all family members (7). Conserved residues within these motifs are critical for covalent nucleotidyl transfer, as shown by mutational analysis (8–13). The recently reported crystal structure of T7 DNA ligase shows that the ATP binding site is made up of five of the six conserved motifs (14).

The *Chlorella* virus DNA ligase (a 298 amino acid polypeptide) is the smallest ATP-dependent ligase described to date (15). Cellular DNA ligases are much larger; for example, human ligases I, III, and IV are 919, 922, and 844 amino acid polypeptides, respectively (4). Vaccinia virus ligase (552 amino acids) is of intermediate size (16).

Sequence comparisons suggest that a catalytic domain common to all ATP-dependent ligases is embellished by additional isozyme-specific protein segments situated at their amino or carboxyl termini. Although the function of these segments is not known, there is reason to believe that they are not essential for catalysis. This is the case for human DNA ligase I, which retains activity after removal of 249 amino acids from the N-terminus and 16 amino acids from the C-terminus (12). Because further deletions of ligase I result in loss of function, it is surmised that a 654 amino acid segment of ligase I represents the catalytic core. In the case of the 755 amino acid *Saccharomyces cerevisiae* DNA ligase, a 650 amino acid C-terminal segment retains catalytic activity (17). These active domains of the cellular ligases are twice the size of the *Chlorella* virus DNA ligase. This raises the question of whether the catalytic core of the cellular DNA ligases is structurally more complex than the ‘minimal’ ligases exemplified by the *Chlorella* virus enzyme and the bacteriophage T7 and T3 enzymes (359 and 346 amino acid polypeptides, respectively).

We are exploring the structure and function of the eukaryotic DNA ligases using the vaccinia virus enzyme as a model. Vaccinia DNA ligase is strikingly similar at the amino acid sequence level to mammalian DNA ligases II and III (5,18). Because vaccinia ligase and mammalian ligases II and III are more similar to each other than to ligases I and IV, these three enzymes can be regarded as a distinct subgroup within the eukaryotic ligase family. The enzymatic properties of vaccinia ligase have been studied using recombinant enzyme produced in bacteria (13,19,20). Interestingly, the vaccinia enzyme requires high concentrations of ATP for strand joining. This is also the case for DNA ligase II, its cellular homologue. Mutational analysis identifies Lys-231 as the active site of adenylate transfer by vaccinia DNA ligase (13).

In this report we employ the classical approach of limited proteolysis to probe the domain structure of vaccinia DNA ligase. We find that the enzyme consists of three domains punctuated by protease-sensitive interdomain bridges. DNA ligase activity is preserved when the N-terminal domain is deleted. However, the truncated form of vaccinia ligase has a reduced catalytic rate and

*To whom correspondence should be addressed. Tel: +1 212 639 7143; Fax: +1 212 717 3623; Email: s-shuman@ski.mskcc.org

a lower affinity for DNA than does the full-sized enzyme. The 350 amino acid catalytic core of the vaccinia ligase is strikingly similar in size and domain organization to the minimal DNA ligases encoded by *Chlorella* virus PBCV-1 and the T-odd bacteriophages.

MATERIALS AND METHODS

Proteolysis of vaccinia DNA ligase

Vaccinia DNA ligase was expressed in bacteria as an N-terminal His_6-tagged fusion protein and purified from bacterial lysates by Ni-agarose and phosphocellulose chromatography as described (13). Proteolysis reaction mixtures (20 µl) containing 40 mM Tris–HCl (pH 8.0), 0.4 M NaCl, 2 mM DTT, 0.1 mM EDTA, 0.1% Triton X-100, 8% glycerol, 6 µg of His_6-ligase, and trypsin or chymotrypsin were incubated at 22°C for 15 min. The samples were denatured in SDS and the proteolysis products were resolved by electrophoresis through a 12% polyacrylamide gel containing 0.1% SDS. The gel was soaked in 100 ml of electroblotting buffer {10 mM CAPS (3-[cyclohexylamino]-1-propanesulfonic acid), pH 11, 10% methanol} for 5 min at room temperature. Polypeptides were transferred to a PVDF membrane (Bio-Rad) in electroblotting buffer by electrophoresis at 500 mA for 2–4 h using a Hoefer Transphor apparatus (Model TE42). The membrane was rinsed with deionized water, then saturated with methanol. The membrane was stained with 0.1% Coomassie Blue in 40% ethanol/1% acetic acid for 1 min, followed by destaining with 50% methanol and rinsing with deionized water at 4°C for 16 h. The membrane was air-dried and slices containing individual proteolytic products were excised. Automated sequencing of the immobilized polypeptide was performed using a modified model 477A microsequencer (Applied Biosytems).

Expression and purification of His_{10}-ligase and N-terminal ligase deletion variants

An *Nde*I fragment containing the entire vaccinia ligase gene was excised from the plasmid pBS-ligase (13) and inserted into the *Nde*I site of the T7-based expression plasmid pET16b (Novagen) to generate pET-His_{10}-ligase. In this plasmid, a leader sequence encoding 10 tandem histidines was fused in-frame to the 5′ end of the ligase gene. (We found that the His_{10} leader peptide afforded a better affinity purification on Ni-agarose than did the His_6 tag used previously.) N-terminal deletion variants were generated by PCR-amplification using oligonucleotide primers designed to introduce *Nde*I restriction sites and to create methionine substitutions at internal positions within the ligase polypeptide. The PCR products were digested with *Nde*I and then inserted into the *Nde*I site of pET16b to yield plasmids pET-His10-ligase(176–552), pET-His10-ligase(191–552), and pET-His10-ligase(201–552). Wild type and deleted His_{10}-ligase expression plasmids were transformed into *Escherichia coli* BL21(DE3). Induction of ligase expression was performed as described previously (13). Cells from 200 ml bacterial cultures were harvested by centrifugation and the pellets were resuspended in 5 ml of lysis buffer [50 mM Tris–HCl (pH 8.0), 0.5 M NaCl, 10% sucrose]. The suspensions were adjusted to 0.2 mg/ml of lysozyme and 0.1% of Triton X-100. After incubation on ice for 30 min, the lysates were sonicated, then centrifuged for 15 min at 15 000 r.p.m. in a Sorvall SS34 rotor. The soluble supernatant fraction was adsorbed to Ni agarose; the resin was washed with buffers containing 5, 25, and 50 mM imidazole (13). His_{10}-ligase was recovered by step-elution with 100 mM imidazole. The Ni-agarose eluates were mixed with 4 vol of buffer A [50 mM Tris–HCl (pH 8.0), 1 mM EDTA, 2 mM dithiothreitol, 0.1% Triton X-100, 10% glycerol] and then applied to 1 ml columns of phosphocellulose that had been equilibrated with buffer A. The columns were eluted stepwise with buffer A containing 0.25, 0.5, and 1 M NaCl. The His_{10}-ligase proteins were recovered in the 0.5 M NaCl fraction. The protein concentration of the enzyme fractions was determined using the Biorad dye reagent with bovine serum albumin as a standard.

Ligase substrate

The standard substrate used in ligase assays was a 36 bp DNA duplex containing a centrally placed nick. This DNA was formed by annealing two 18mer oligonucleotides to a complementary 36mer strand (19). The 18mer constituting the donor strand was 5′ ^{32}P-labeled and gel-purified as described (13,19). The labeled donor was annealed to the complementary 36mer in the presence of a 3′ OH-terminated acceptor strand in 0.2 M NaCl by heating at 65°C for 2 min, followed by slow-cooling to room temperature. The molar ratio of the 18mer donor to 36mer complement to 18mer acceptor strands in the hybridization mixture was 1:4:4.

DNA ligation

Reaction mixtures (20 µl) containing 50 mM Tris–HCl (pH 8.0), 5 mM DTT, 10 mM $MgCl_2$, 1 mM ATP (where indicated), 500 fmol of 5′ ^{32}P-labeled DNA substrate, and enzyme were incubated at 22°C. Reactions were initiated by addition of enzyme and halted by the addition of 1 µl of 0.5 M EDTA and 5 µl of formamide. The samples were heated at 95°C for 5 min and then electrophoresed through a 17% polyacrylamide gel containing 7 M urea in TBE (90 mM Tris borate, 2.5 mM EDTA). The labeled 36mer ligation product was well-resolved from the 5′-labeled 18mer donor strand. The extent of ligation [36mer/(18mer + 36mer)] was determined by scanning the gel using a FUJIX BAS1000 phosphorimager.

RESULTS

Structure probing of vaccinia ligase by proteolysis

Recombinant vaccinia ligase containing an N-terminal His_6 tag was subjected to proteolysis with increasing amounts of trypsin. Initial scission of the 65 kDa His-ligase yielded two major products, a 29 kDa polypeptide (T29) and a polypeptide doublet of 35–36 kDa (Fig. 1). Sequencing of these cleavage products by automated Edman chemistry after transfer to a PVDF membrane revealed that T29 was a mixture of three species arising via trypsin cleavage at vicinal residues Lys-303, Lys-304, and Lys-305. These sites are denoted by arrows above the polypeptide sequence in Figure 2. (Note that amino acid residues are numbered according to their position in the native ligase polypeptide, not in the His-tagged version.) The T35 product arose via cleavage at an Arg residue within the His-tag (the leader peptide is underlined in Fig. 2). Thus, trypsin at limiting concentration cleaved the ligase into a 35 kDa N-terminal fragment and a 29 kDa carboxyl fragment.

T29 was resistant to digestion by higher concentrations of trypsin; indeed, this species persisted even after all of the input

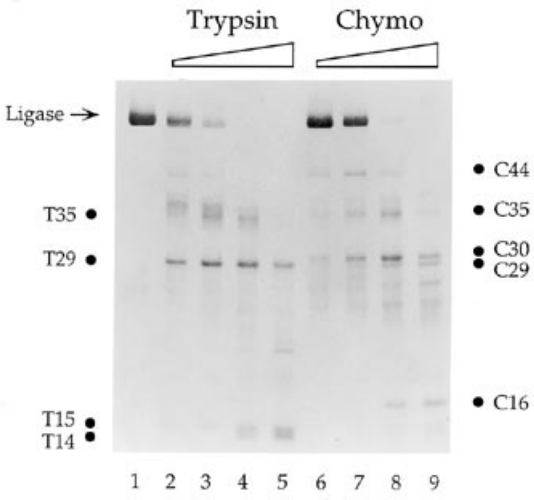

Figure 1. Digestion of vaccinia DNA ligase by trypsin and chymotrypsin. Proteolysis reaction mixtures contained 6 µg of purified His-Ligase (lane 1). The samples were digested with increasing amounts of trypsin (8, 20, 80, and 200 ng; lanes 2–5) or chymotrypsin (20, 50, 200, and 500 ng; lanes 6–9). The reaction products were resolved by SDS-PAGE, then transferred to a PVDF membrane and visualized by staining with Coomassie Blue dye. The intact ligase polypeptide is indicated at the left. The proteolytic fragments that were subjected to N-terminal sequencing are denoted by filled circles.

GSSHHHHHHSSGLVPRGSHMTSLREFRKLCCDIYHASGYKEKSKLIRDFITDRDDKYLIIK

LLLPGLDDRIYNMNDKQIIKLYSIIFKQSQEDMLQDLGYGYIGDTIRTFFKENTEIRPRDK

SILTLEDVDSFLTTLSSVTKESHQIKLLTDIASVCTCNDLKCVVMLIDKDLKIKAGPRYVL

NAISPNAYDVFRKSNNLKEIIENSSKQNLDSISISVMTPINPMLAESCDSVNKAFKKFPSG

MFAEVKYDGERVQVHKNNNEFAFFSRNMKPVLSHKVDYLKEYIPKAFKKATSIVLDSEIVL

VDEHNVPLPFGSLGIHKKKEYKNSNMCLFVFDCLYFDGFDMTDIPLYERRSFLKDVMVEIP

NRIVFSELTNISNESQLTDVLDDALTRKLEGLVLKDINGVYEPGKRRWLKIKRDYLNEGSM

ADSADLVVLGAYYGKGAKGGIMAVFLMGCYDDESGKWKTVTKCSGHDDNTLRVLQDQLTMI

KINKDPKKIPEWLVVNKIYIPDFVVEDPKQSQIWEISGAEFTSSKSHTANGISIRFPRFTR

IREDKIWKESTHLNDLVNLTKS

Figure 2. Sites of protease accessibility within native vaccinia DNA ligase. The amino acid sequence of the His-Ligase polypeptide is shown. The N-terminal His_6-tag is underlined. Conserved motifs I, III, IIIa, IV, V, and VI are highlighted in boxes. The sites of proteolysis by trypsin are denoted by arrows above the sequence; those of cleavage by chymotrypsin are indicated by arrows below the sequence.

His-ligase had been degraded (Fig. 1). In contrast, the N-terminal fragment T35 was converted at higher levels of trypsin to a doublet migrating at 14 kDa (denoted as T14 and T15 in Fig. 1). Both polypeptides of the doublet derived from trypsin cleavage between Lys-190 and Gln-191 (Fig. 2). The size of these species suggests that they extend from Gln-191 to the sites of initial trypsin cleavage at Lys-303 to Lys-305.

Treatment of His-ligase with limiting concentrations of chymotrypsin yielded polypeptides of 44 kDa (C44), 35 kDa (C35)

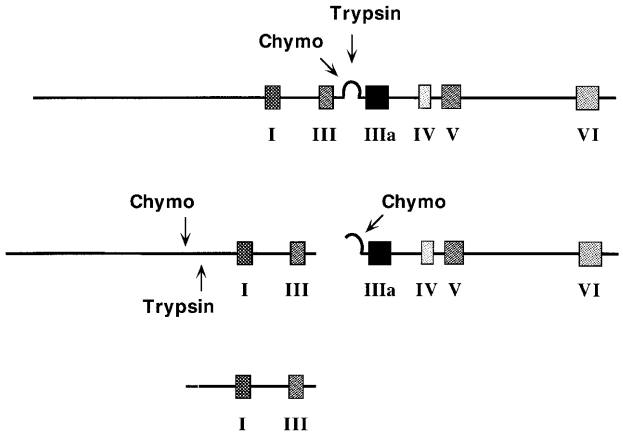

Figure 3. Domain structure of vaccinia DNA ligase. The ligase polypeptide is depicted as a straight line, with the positions of conserved motifs I, III, IIIa, IV, V, and VI denoted by filled boxes. Initial sites of cleavage by chymotrypsin and trypsin are closely clustered within an interdomain bridge, depicted as a loop between motifs III and IIIa. Scission within this loop generates the amino and carboxyl protein fragment shown below the intact ligase. Sites of secondary cleavages by proteases are indicated. The C-terminal domain is protease resistant, whereas the N-terminal polypeptide is converted into a protease resistant central domain that includes motifs I and III.

and 30 kDa (C30) (Fig. 1). C30 is a carboxyl fragment arising via cleavage between Phe-296 and Gly-297. These sites are denoted by arrows below the polypeptide sequence in Figure 2. Cleavage product C35 contains the original N-terminus of His-ligase. As the sum of the apparent molecular weights of C35 and C30 were equivalent to that of the full length protein, we presume that they were the products of a single proteolytic event. C44 derived from chymotryptic cleavage between Phe-175 and Arg-176 (Fig. 2).

The C30 species was largely resistant to digestion by concentrations of chymotrypsin sufficient to cleave all the input ligase (Fig. 1). Some breakdown of the C30 to a 29 kDa polypeptide was evident at higher levels of protease (Fig. 1). The sequence of the C29 species indicated that the secondary cleavage event by chymotrypsin occurred between residues Tyr-307 and Lys-308 (Fig. 2). The N-terminal C35 species and the C44 polypeptide were degraded at higher levels of chymotrypsin; this occurred as a novel 16 kDa product was formed. C16 originated via cleavage between Tyr-307 and Lys-308. The size of the C16 species suggested that it extends from Lys-308 to the site of initial chymotrypsin cleavage at Phe-175.

These results suggests a tripartite domain structure for vaccinia ligase, as illustrated in Figure 3. An interdomain bridge located between motifs III and IIIa separates a protease-resistant C-terminal domain from a proximal structural domain that is composed of a protease-sensitive N-terminal region and a protease-resistant central domain. The latter includes the active site of adenylate transfer (motif I).

The N-terminal domain is dispensable for nucleotidyl transfer

Three deletion mutants of vaccinia DNA ligase were constructed. Ligase(176–552) and ligase(191–552) were deleted from the N-terminus to the upstream margins of the central domain, as defined by sites of cleavage by chymotrypsin (at Phe-175) and trypsin (at Lys-190). Ligase(201–552) was truncated up to Met-201, which is located 30 residues upstream of the active site residue, Lys-231 (Fig. 2). The ligase genes were inserted into an

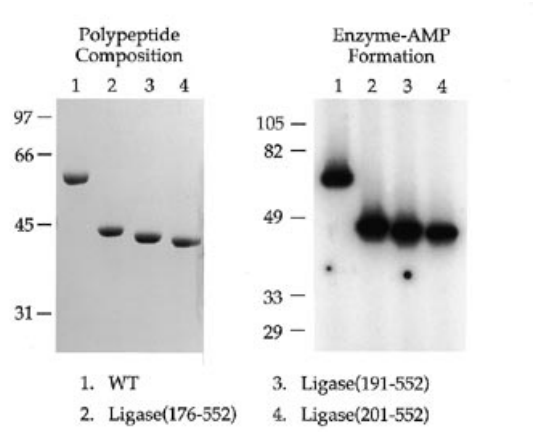

Figure 4. Purification and adenylyltransferase activity of N-terminal deletion mutants of vaccinia DNA ligase. (Left) Aliquots (5 µg) of the phosphocellulose preparations of recombinant ligase protein were analyzed by SDS-PAGE. A Coomassie blue-stained gel is shown. The positions and sizes (kDa) of marker proteins are shown at the left of the gel. (Right) Enzyme-adenylate formation. Reaction mixtures (20 µl) containing 50 mM Tris–HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT, 10 µM [α-^{32}P]ATP, and 100 ng of the indicated preparations of recombinant ligase were incubated at 22°C for 10 min. Reactions were stopped by adding SDS to 1%. The samples were heated at 95°C for 5 min, then analyzed by SDS-PAGE. An autoradiogram of the gel is shown. The positions and sizes (kDa) of pre-stained marker proteins are shown at the left.

inducible T7 RNA polymerase-based pET vector such that a histidine-rich N-terminal leader (His-10) was fused to each ligase polypeptide. The pET expression plasmids were introduced into *E.coli* BL21(DE3), a strain that contains the T7 RNA polymerase gene under the control of a *lac*UV5 promoter. Polypeptides corresponding to wild type ligase, ligase(176–552), ligase(191–552), and ligase(201–552) were recovered in soluble extracts of IPTG-induced bacteria (not shown). The His-tag allowed for rapid enrichment of the ligase proteins by adsorption to Ni-agarose and elution with 100 mM imidazole. Each protein was purified further by adsorption to phosphocellulose and step-elution with 0.5 M NaCl. SDS-PAGE analysis of the phosphocellulose preparations is shown in Figure 4 (left panel). Note that the electrophoretic mobilities of the recombinant proteins increased with serial N-terminal deletions, as expected.

The initial step in DNA ligation involves formation of a covalent enzyme-adenylate intermediate, EpA. The formation of EpA by vaccinia ligase can be detected by label transfer from [α^{32}P]ATP to the enzyme. Incubation of purified recombinant wild type ligase in the presence of [α-^{32}P]ATP and a divalent cation resulted in the formation of a nucleotidyl-protein adduct that migrated as a single 60 kDa species during SDS-PAGE (Fig. 4 right panel). Ligase(176–552), ligase(191–552) and ligase(201–552) also formed enzyme-adenylate complexes. We conclude that the N-terminal 200 amino acids of vaccinia DNA ligase are not required for the first step of covalent nucleotidyl transfer.

Effects of N-terminal deletions on DNA ligation

We assayed the ability of the recombinant ligases to seal a 36mer synthetic duplex DNA substrate containing a single nick (19). The structure of the substrate is shown in Figure 5. Ligase activity was evinced by conversion of the 5′ ^{32}P-labeled 18mer donor strand into an internally-labeled 36mer product (19). The standard ligase reaction mixtures contained 500 fmol of labeled DNA substrate. The extent of strand joining by wild type ligase in the presence of 1 mM ATP increased linearly with protein up to 10 fmol (Fig. 5, +ATP). The reaction saturated at ≥20 fmol of enzyme with 85% of the labeled donor strand converted to 36mer in 10 min. This upper limit of ligation probably reflected incomplete annealing of all three component strands to form the nicked substrate. In the linear range of enzyme-dependence in this experiment, the recombinant wild type ligase joined ~41 fmol of DNA ends per fmol of enzyme. For the purpose of estimating the ratio of product to enzyme, the enzyme molarity was calculated based on total protein concentration, assuming enzyme homogeneity. It was also assumed that all enzyme molecules in the preparation were catalytically active.

Strand joining by wild type ligase could be detected in the absence of added ATP, but only at high levels of input enzyme (Fig. 5, –ATP). ATP-independent ligation is attributable to

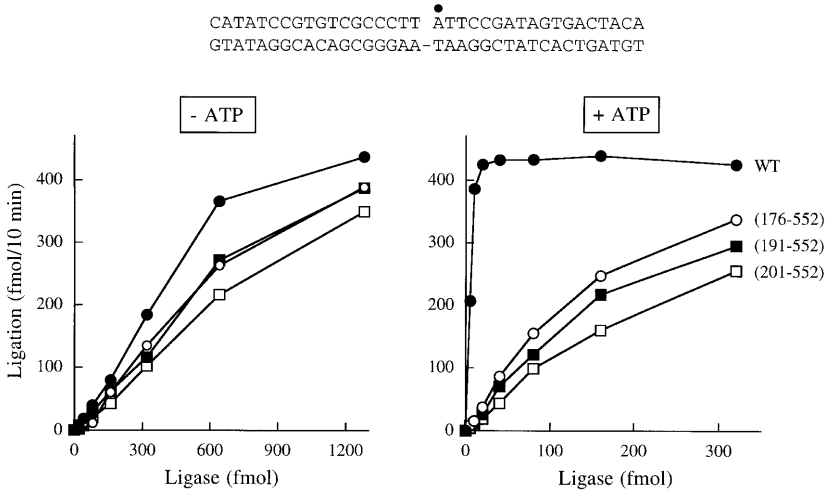

Figure 5. DNA ligation. The structure of the nicked duplex DNA substrate is shown with the 5′ ^{32}P-labeled nucleotide on the donor strand indicated by a closed circle. Ligation reaction mixtures contained 500 fmol nicked DNA, the indicated amounts of wild type or truncated ligases and either 1 mM ATP (+ATP; left) or no added ATP (–ATP; right). The reactions were halted after incubation for 10 min at 22°C. The yield of 36mer ligation product is plotted as a function of the molar amount of input ligase.

ligase-adenylate in the enzyme preparation. The linear dependence of ATP-independent strand joining on enzyme indicated that 0.58 fmol of ends were sealed per fmol of ligase, i.e. 58% of the enzyme molecules had AMP bound at the active site.

The activity of deletion variants ligase(176–552), ligase(191–552), and ligase(201–552) in ATP-independent strand joining as a function of the molar amount of input protein was similar to that of wild type ligase (Fig. 5, –ATP). We calculated the percent of enzyme-adenylate in the preparations as follows: 42% for ligase(176–552), 42% for ligase(191–552), and 39% for ligase(201–552). Clearly, these deleted enzymes retained the ability to catalyze strand joining. However, when activity was assayed in the presence of ATP, the truncated enzymes were less active on a molar basis than the wild type protein (Fig. 5, +ATP). Whereas specific activity in wild type ligation was stimulated 40-fold by inclusion of ATP, the truncated proteins were stimulated only 4- to 5-fold. This implied that the mutant proteins, while catalytically competent under single-turnover conditions (in the absence of ATP), were impaired relative to wild type under multiple turnover conditions (in the presence of ATP and excess DNA substrate). The basis for this effect was revealed by kinetic analyses described below.

Kinetics of ATP-independent ligation

Strand joining by pre-adenylylated ligase in the absence of ATP is a two step reaction entailing: (i) AMP transfer to the donor strand to form DNA-adenylate and (ii) attack on DNA-adenylate by the acceptor strand with displacement of AMP. Because reversal of the ligation step is both slow and dependent on high (≥5 mM) concentrations of exogenous AMP (Sekiguchi and Shuman, unpublished), and because the forward reaction proceeds to completion at stoichiometric levels of EpA, we regard ATP-independent ligation as an essentially irreversible reaction under the conditions employed in our experiments. We analyzed the kinetics of ATP-independent ligation by wild type ligase and by each of the N-terminal deletion mutants. The concentration of nicked DNA substrate in the ligation mixtures was 12.5 nM; enzyme-adenylate was included at several concentrations. Results are shown in Figure 6 for the wild type enzyme and for ligase(176–552).

Table 1. Kinetics of ATP-independent ligation

Enzyme	k_{lig} (s^{-1})	K_m (nM)
Wild Type	0.5	9
Ligase(176–552)	0.04	33
Ligase(191–552)	0.04	55
Ligase(201–552)	0.04	59

The rate of single-turnover ligation by ligase(176–552) varied with the concentration of input enzyme-adenylate in the range of 11–210 nM (Fig. 6A). Observed rate constants (k_{obs}) were calculated by fitting each time-course to a single exponential. A double reciprocal plot of $1/k_{obs}$ versus 1/[EpA] fit well to a straight line (Fig. 6B). A rate constant for single-turnover ligation by EpA (k_{lig}), extrapolated to infinite EpA concentration, was determined from the *y*-intercept of the double-reciprocal plot (*y*-intercept = $1/k_{lig}$). The value for k_{lig} was 0.04 s^{-1} (Table 1). The Michaelis constant of the substrate for EpA was calculated from

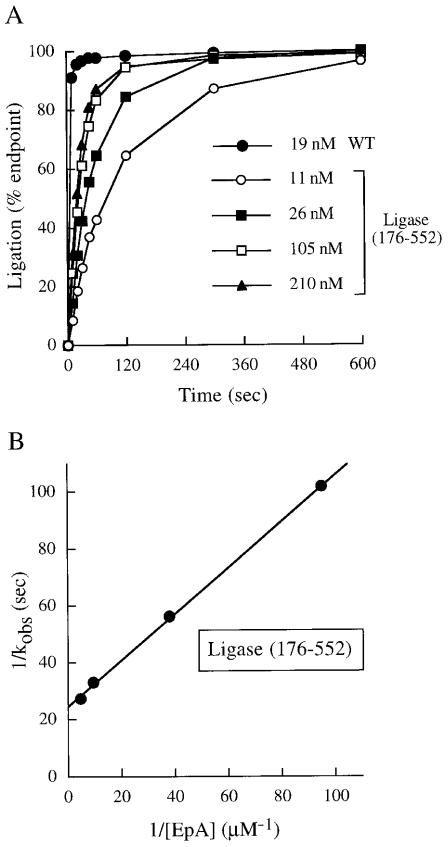

Figure 6. Kinetic analysis of ATP-independent ligation. (**A**) reaction mixtures containing (per 20 µl) 50 mM Tris–HCl (pH 8.0), 10 mM $MgCl_2$, 5 mM DTT, 12.5 nM nicked DNA substrate, and either wild type (WT) ligase or ligase(176–552) at the concentrations indicated were incubated at 22°C. Enzyme concentrations are expressed as the molar concentration of ligase-adenylate included in the reaction mixtures. The reactions were initiated by the addition of enzyme. Aliquots (20 µl) were withdrawn at the times indicated and quenched immediately. The 'time 0' sample was taken prior to the addition of enzyme. The extent of ligation at each time point was normalized to the observed end-point value for strand joining at each ligase concentration. The normalized data (% of endpoint) are plotted as a function of time. A plot of the percent of input DNA cleaved versus time established endpoint values for cleavage. The observed rate constants for ATP-independent ligation (k_{obs}) at each concentration of ligase(176–552) were determined by fitting the data to the equation (100 – % of endpoint) = $100e^{-kt}$. A double reciprocal plot of $1/k_{obs}$ versus 1/[EpA] is shown in (**B**).

the *x*-intercept of the double-reciprocal plot (*x*-intercept = $-1/K_m$). The K_m value was 33 nM (Table 1).

Single-turnover ATP-independent ligation by 19 nM wild type EpA was essentially complete within 10 s (Fig. 6A). Additional experiments showed that the extent of ligation by 10 nM wild type EpA at 5 s was 80% of the endpoint value (±1.9%; average of four experiments); ligation by 70 nM wild type EpA at 5 s was 90% of the endpoint value (±1.1%; average of four experiments). These values were used to calculate minumum values for k_{obs}. From a double reciprocal plot of $1/k_{obs}$ versus 1/[EpA], we determined the following kinetic parameters for the wild type enzyme: k_{lig} was 0.5 s^{-1}; K_m value was 9 nM (Table 1). Comparing these values to those of ligase(176–552), we surmise that deletion of the N-terminal 175 amino acids slowed the rate of ATP-independent ligation by an order of magnitude and reduced the affinity of EpA for the nicked DNA substrate.

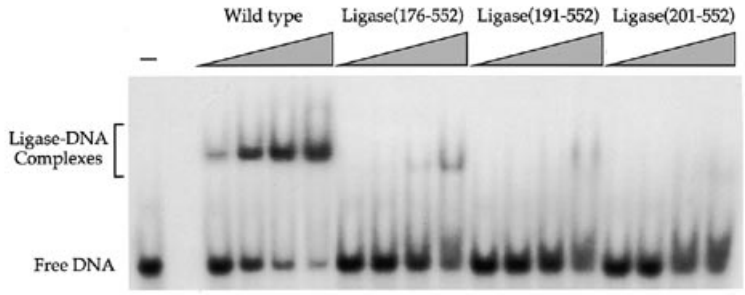

Figure 7. Mutational effects on ligase binding to nicked DNA. Reaction mixtures (20µl) contained 50 mM Tris–HCl (pH 8.0), 5 mM DTT, 500 fmol of ^{32}P-labeled nicked DNA (lane –), and 0.25, 0.5, 1, or 2.5 pmol (proceeding from left to right within each titration series) of the indicated preparation of vaccinia ligase. (Ligase concentrations are the molar amount of ligase protein). After incubation for 10 min at 22°C, glycerol was added to 5%, and the samples were electrophoresed for 1.5 h at 100 V through a native 6% polyacrylamide gel in 0.25× TBE (22.5 mM Tris-borate, 0.6 mM EDTA). An autoradiogram of the dried gel is shown. Labeled species corresponding to free DNA and the ligase–DNA complexes are indicated at the left.

Kinetic parameters in ATP-independent ligation were also determined for ligase(191–552) and ligase(201–552) (Table 1). The k_{lig} values (0.04 s^{-1}) were identical to that of ligase (176–552), whereas the K_m values were higher: 55 and 59 nM, respectively. Thus, incremental deletion of the 25 residues from 177 to 200 did not exacerbate the catalytic defect, but did enhance the effect on DNA binding affinity.

Effects of N-terminal deletions on the binding of vaccinia ligase to nicked duplex DNA

A native gel mobility shift assay was employed to examine the binding of purified recombinant ligase to the ^{32}P-labeled nicked duplex DNA (19). Binding reactions were performed in the absence of magnesium so as to preclude conversion of substrate to product during the incubation (19). Mixing the wild type ligase with nicked substrate resulted in the formation of a discrete protein–DNA complex that migrated more slowly than the free DNA during electrophoresis through a 6% native polyacrylamide gel (Fig. 7). The abundance of this complex increased in proportion to the amount of input ligase. In order to estimate binding affinity, the gel was scanned using a phosphorimager. The dependence of protein–DNA complex formation on input ligase is shown in Figure 8. The apparent dissociation constant, estimated according to Riggs *et al.* (21), was ~15 nM.

A more rapidly migrating protein–DNA complex was detected when ligase(176–552) was incubated with the nicked DNA (Fig. 7). The altered mobility was in keeping with the smaller size of the deleted protein. This mutant ligase formed lower amounts of ligase–DNA complex than did the wild type enzyme at comparable levels of input protein, and there was a smear of radiolabeled material trailing behind the free DNA (Fig. 7). The protein-dependence of DNA binding by ligase(176–552) is shown in Figure 8. By comparing the extents of binding at equivalent concentration of ligase, we estimated that the N-terminal deletion resulted in a ~10-fold decrement in binding affinity relative to wild type ligase. Formation of a discrete protein–DNA complex was even lower with deletion mutants ligase(191–552) and ligase(201–552) (Fig. 7). Thus, the mutational effects on substrate affinity determined from kinetic analysis of single-turnover ligation were confirmed by direct assay of DNA binding via electrophoretic mobility shift.

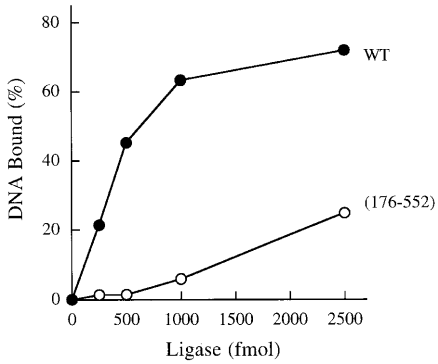

Figure 8. Protein-dependence of ligase–DNA complex formation. The extents of DNA binding by wild type ligase and ligase(176–552) at each level of input protein were quantitated by scanning the gel with a FUJIX BAS1000 Bio-Imaging analyzer. A plot of DNA binding, [bound/(bound + free)] × 100, versus input ligase is shown.

DISCUSSION

We have employed partial proteolysis and deletion mutagenesis to identify structural and functional domains of vaccinia virus DNA ligase. Three significant findings emerge from the experiments presented: (i) the native ligase consists of (at least) three structural domains; (ii) the N-terminal 200 amino acids are not essential for ligase activity; (iii) N-terminal deletions reduce affinity of the ligase for nicked DNA and decrease the rate of strand joining at a step subsequent to enzyme-adenylate formation.

Proteolytic domain mapping

Vaccinia ligase consists of three structural domains. These are: (i) an amino-terminal domain that is susceptible to digestion with chymotrypsin and trypsin; (ii) a protease-resistant central domain, and (iii) a protease-resistant carboxyl domain. The two protease-resistant domains are separated by a protease-sensitive interdomain bridge. The bridge extends from Phe-296 to Tyr-307 (sites accessible to chymotrypsin), and includes a cluster of trypsin-sensitive lysines at positions 303, 304, and 305. The interdomain bridge of the vaccinia ligase is located between conserved sequence motifs III and IIIa of the ligase/capping enzyme superfamily. Mutational and structural studies of family

members indicate that motifs III and IIIa form part of the nucleotide binding pocket (9,13,14).

The domains defined by proteolysis of the vaccinia ligase are similar to those demarcated for the T7 enzyme (22), except that the vaccinia protein contains the N-terminal domain that is absent from the T7 protein. Limited digestion of T7 ligase with endoproteinase Arg-C yielded an N-terminal 16 kDa fragment and a 26 kDa carboxyl fragment. As with the vaccinia enzyme, the sites of protease accessibility in the T7 ligase were located between motifs III and IIIa (22). In the crystal structure of T7 DNA ligase, this polypeptide segment is a disordered loop that extends from the surface of an otherwise tightly folded protein (14). This would account for the observed protease susceptibility. The observation that the same segment of the vaccinia DNA ligase is protease-sensitive suggests that it is also exposed on the surface. Thus, both the T7 and the vaccinia ligases have two protease-resistant domains, one containing motifs I and III and another including motifs IIIa, IV, V, and VI.

The size of the protease-sensitive polypeptide segment separating motifs III and IIIa varies considerably within the covalent nucleotidyl transferase superfamily (7,15). Among the DNA ligases, the bacteriophage enzymes have the longest interval (58, 55, and 49 amino acids for the T4, T3, and T7 enzymes, respectively), whereas vaccinia ligase and *Chlorella* virus ligase have the shortest interval (27 and 24 amino acids, respectively). In the capping enzymes, motifs III and IIIa are separated by only 3–12 amino acids.

Deletion mutations

The N-terminal domain of vaccinia ligase was less stable to proteolysis than were the central and carboxyl domains. This protein segment, extending to the site of trypsin cleavage at position 190, has no counterpart in the smaller DNA ligases encoded by the T-odd bacteriophages and by the *Chlorella* virus PBCV-1. Our analysis of the effects of N-terminal deletions on vaccinia ligase activity indicated that the N-terminal domain is not essential for adenylyl transfer or strand joining, insofar as the deleted proteins retained catalytic function. Yet, removal of the N-terminus did have consequences with respect to catalytic rate and DNA substrate affinity.

Deletion of the N-terminal 176–200 amino acids resulted in a 10-fold decrement in the rate constant for single-turnover ligation by enzyme-adenylate. This could arise from mutational effects either on AMP transfer from EpA to nicked DNA to form DNA-adenylate (step 2 of the overall ligation reaction) or on enzyme-mediated attack of the acceptor strand on DNA-adenylate to form a phosphodiester bond (step 3 of the complete reaction). Previous studies had indicated that DNA-adenylate formation (step 2) is the rate-limiting step during DNA ligation catalyzed by the wild type vaccinia ligase. DNA-adenylate is detectable only under reaction conditions that strongly suppress the strand closure step (19). Because DNA-adenylate did not accumulate during the course of single-turnover ligation by the deletion mutants of vaccinia ligase (Sekiguchi and Shuman, unpublished), we surmise that the observed effects of N-terminal truncations on k_{lig} were caused primarily by a slowing of step 2. We cannot exclude the possibility that the N-terminal deletions affected the rate of step 3; however, potential rate effects on step 3 are apparently not significant (i.e. not rate-determining) in the context of the single-turnover ligation reaction.

Deletion of the N-terminal domain had a definite impact on the affinity of the vaccinia ligase for nicked duplex DNA. Kinetic analysis indicates that the wild type K_m of nicked DNA for enzyme is 9 nM. The N-terminal deletions increased the K_m value by a factor of 4–6. The kinetic experiments were supported by native gel-shift assays of the binding of ligase to nicked DNA. Detection of a discrete ligase–DNA complex in the mobility shift assay requires that the protein remain bound stably to the DNA ligand during the electrophoresis procedure. Formation of lower levels of discrete shifted complexes by the truncated ligases, and the presence of a diffuse smear of DNA trailing behind the unbound ligand, presumably reflects both a lower binding affinity for nicked DNA in solution (as revealed initially by the kinetic experiments) and reduced stability of the mutant protein–DNA complex during the electrophoresis. The latter may account for the observation that the mutational effects on DNA binding in the gel-shift assay were slightly greater than those determined kinetically.

In summary, we have defined at low resolution a domain structure for vaccinia ligase and shown that a 352 amino acid derivative, ligase(201–552), is active in enzyme-adenylate formation and strand ligation. The size of the vaccinia catalytic domain is comparable to that of the full-length 359 amino acid T7 enzyme; indeed the vaccinia and T7 enzymes have similar domain boundaries. Vaccinia ligase(201–552) contains only 30 amino acids upstream of the active site lysine residue in motif I. The *Chlorella* virus and T7 DNA ligases contain 26 and 33 amino acids upstream of their respective active sites (14,15).

Although the N-terminal domain of vaccinia ligase is not essential, we suspect that it does play a role in ligase function. This is based on the observations that catalytic rate and DNA affinity are impacted by loss of the N-terminus, plus the fact that the amino acid sequence of the entire N-terminal domain of vaccinia ligase is conserved in mammalian DNA ligase III and in mammalian DNA ligase I (3–5). The proximal region of similarity to the vaccinia N-terminal domain is recessed from the amino terminus in ligase III and ligase I because these cellular ligases contain sizable protein segments at their amino termini (175 and 292 amino acids respectively) that have no counterpart in the vaccinia virus protein (5). This N-terminal domain of mammalian DNA ligase I contains two elements involved in protein targeting: a nuclear localization signal and a signal that directs ligase to sites of DNA replication within the nucleus (23). Vaccinia virus, which replicates in the cytoplasm, would have no need for such a domain. It is tempting to speculate that the shared N-terminal domain of the mammalian and vaccinia ligases might, in addition to facilitating catalysis and DNA binding as described herein, also participate in protein–protein interactions with other components of the DNA repair or replication machinery. In this respect, it is noteworthy that vaccinia DNA ligase localizes to the cytoplasmic sites of DNA replication in virus-infected cells (24); the deletion mutants described in this report may prove useful in delineating the requirement for targeting of vaccinia ligase *in vivo*.

REFERENCES

1 Lehman, I. R. (1974) *Science* **186**, 790–797.
2 Lindahl, T., and Barnes, D. E. (1992) *Annu. Rev. Biochem.* **61**, 251–281.
3 Barnes, D. E., Johnston, L. H., Kodama, K., Tomkinson, A. E., Lasko, D. D., and Lindahl, T. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6679–6683.

4 Wei, Y., Robins, P., Carter, K., Caldecott, K., Pappin, D., Yu, G., Wang, R., Shell, B., Nash, R., Schar, P., Barnes, D., Haseltine, W., and Lindahl, T. (1995) *Mol. Cell. Biol.* **15**, 3206–3216.
5 Chen, J., Tomkinson, A. E., Ramos, W., Mackey, Z. B., Danehower, S., Walter, C. A., Schultz, R. A., Besterman, J. M., and Husain, I. (1995) *Mol. Cell. Biol.* **15**, 5412–5422.
6 Kletzin, A. (1992) *Nucleic Acids Res.* **20**, 5389–5396.
7 Shuman, S., and Schwer, B. (1995) *Mol. Microbiol.* **17**, 405–410.
8 Cong, P., and Shuman, S. (1993) *J. Biol. Chem.* **268**, 7256–7260.
9 Cong. P., and Shuman, S. (1995) *Mol. Cell. Biol.* **15**, 6222–6231.
10 Schwer, B., and Shuman, S. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4328–4332.
11 Shuman, S., Liu, Y., and Schwer, B. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 12046–12050.
12 Kodama, K., Barnes, D. E., and Lindahl, T**.** (1991) *Nucleic Acids Res.* **19**, 6093–6099.
13 Shuman, S., and Ru. X. (1995) *Virology* **211**, 73–83.
14 Subramanya, H. S., Doherty, A. J., Ashford, S. R., and Wigley, D. B. (1996) *Cell* **85**, 607–615.
15 Ho, C. K., Van Etten, J. L., and Shuman, S. (1997) *J. Virol.* **71**, (in press).
16 Smith, G. L., Chan, Y. S., and Kerr, S. M. (1989) *Nucleic Acids. Res.* **17**, 9051–9062.
17 Tomkinson, A. E., Tappe, N. J., and Friedberg, E. C. (1992) *Biochemistry* **31**, 11762–11771.
18 Husain, I., Tomkinson, A. E., Burkhart, W. A., Moyer, M. B., Ramos, W., Mackey, Z. B., Besterman, J. M., and Chen, J. (1995) *J. Biol. Chem.* **270**, 9683–9690.
19 Shuman, S. (1995) *Biochemistry* **34**, 16138–16147.
20 Odell, M., Kerr, S. M., and Smith G. L. (1996) *Virology* **221**, 120–129.
21 Riggs, A. D., Suzuki, H., and Bourgeois, S. (1970) *J. Mol. Biol.* **48**, 67–83.
22 Doherty, A. J., Ashford, S. R., and Wigley, D. B. (1996) *Nucleic Acids. Res.* **24**, 2281–2287.
23 Montecucco, A., Savini, E., Weighardt, F., Rossi, R., Ciarrocchi, G., Villa, A., and Biamonti, G. (1995) *EMBO J.* **14**, 5379–5386.
24 Kerr, S. M., Johnston, L. H., Odell, M., Duncan, S. A., Law, K. M., and Smith, G. L. (1991) *EMBO J.* **10**, 4343–4350.