Nucleic Acids Research, 2002, Vol. 30, No. 4 **903–911**

Role of nucleotidyltransferase motifs I, III and IV in the catalysis of phosphodiester bond formation by *Chlorella* virus DNA ligase

Verl Sriskanda and Stewart Shuman*

Molecular Biology Program, Sloan-Kettering Institute, 1275 York Avenue, New York, NY 10021, USA

Received November 6, 2001; Revised and Accepted December 14, 2001

ABSTRACT

ATP-dependent DNA ligases catalyze the sealing of 5′-phosphate and 3′-hydroxyl termini at DNA nicks by means of a series of three nucleotidyl transfer steps. Here we have analyzed by site-directed mutagenesis the roles of conserved amino acids of *Chlorella* virus DNA ligase during the third step of the ligation pathway, which entails reaction of the 3′-OH of the nick with the DNA–adenylate intermediate to form a phosphodiester and release AMP. We found that Asp65 and Glu67 in nucleotidyltransferase motif III and Glu161 in motif IV enhance the rate of step 3 phosphodiester formation by factors of 20, 1000 and 60, respectively. Asp29 and Arg32 in nucleotidyltransferase motif I enhance the rate of step 3 by 60-fold. Gel shift analysis showed that mutations of Arg32 and Asp65 suppressed ligase binding to a pre-adenylated nick, whereas Asp29, Glu67 and Glu161 mutants bound stably to DNA–adenylate. We infer that Asp29, Glu67 and Glu161 are involved directly in the step 3 reaction. In several cases, the effects of alanine or conservative mutations on step 3 were modest compared to their effects on the composite ligation reaction and individual upstream steps. These results, in concert with available crystallographic data, suggest that the active site of DNA ligase is remodeled during the three steps of the pathway and that some of the catalytic side chains play distinct roles at different stages.

INTRODUCTION

DNA ligases are ubiquitous enzymes that repair DNA nicks containing 5′-phosphate and 3′-hydroxyl termini (1). The nick joining reaction depends on a high energy cofactor, which can be either ATP or NAD^+, depending on the source of the ligase. Ligation proceeds in a series of three partial reactions. In the first step, ligase reacts with the nucleotide cofactor to form a covalent intermediate (ligase–adenylate) in which AMP is linked via a phosphoamide bond to lysine. In the second step, the AMP is transferred to the 5′-phosphate at the nick to form DNA–adenylate (AppDNA). In the third step, ligase catalyzes attack by the 3′-OH of the nick on DNA–adenylate to join the two polynucleotides and release AMP. The ATP-dependent ligases are exemplified by the bacteriophage T7 and *Chlorella* virus enzymes, for which the atomic structures have been solved by X-ray crystallography (2,3). The *Chlorella* virus enzyme consists of a 188 amino acid N-terminal domain (domain 1) and a 110 amino acid C-terminal domain (domain 2). Within the N-terminal domain is an adenylate-binding pocket composed of five motifs (I, III, IIIa, IV and V) that define the polynucleotide ligase/mRNA capping enzyme superfamily of covalent nucleotidyltransferases (2–5) (Fig. 1). Motif I (^{27}KxDGxR) contains the lysine nucleophile to which AMP becomes covalently linked in the first step of the ligase reaction (3). Motifs III, IIIa, IV and V contain conserved side chains that contact AMP (2,3). The C-terminal domain comprises an oligomer-binding fold (OB-fold) consisting of a five-stranded antiparallel β-barrel and an α-helix. The OB-fold domain includes nucleotidyltransferase motif VI, which is uniquely required for step 1 of the ligase reaction (6).

The 298 amino acid *Chlorella* virus ligase is the smallest eukaryotic ATP-dependent ligase that has been characterized (7). It consists only of the catalytic core, without the large N- or C-terminal flanking domains found in eukaryotic cellular DNA ligases. The *Chlorella* virus ligase has an intrinsic nick-sensing function and it sustains mitotic growth and DNA repair in budding yeast when it is the only source of ligase in the cell (8–10). Our aim is to delineate structure–activity relationships for the *Chlorella* virus ligase that will contribute to an understanding of the general mechanism of covalent nucleotidyl transfer by the polynucleotide ligase/capping enzyme superfamily and provide insights into the basis for substrate specificity and nick recognition by the members of the ATP-dependent ligase subfamily. We are especially interested in the structural basis for step 3 of the ligation reaction, which distinguishes the polynucleotide ligases from the RNA capping enzymes. The crystal structure of the *Chlorella* virus ligase–adenylate intermediate, together with phylogenetic comparisons, provides a blueprint for mechanistic analysis by mutagenesis. Here we focus on the catalytic roles of individual amino acids in nucleotidyltransferase motifs I, III, IIIa and IV (Fig. 1A).

*To whom correspondence should be addressed. Tel: +1 212 639 7145; Fax: +1 212 717 3623; Email: s-shuman@ski.mskcc.org

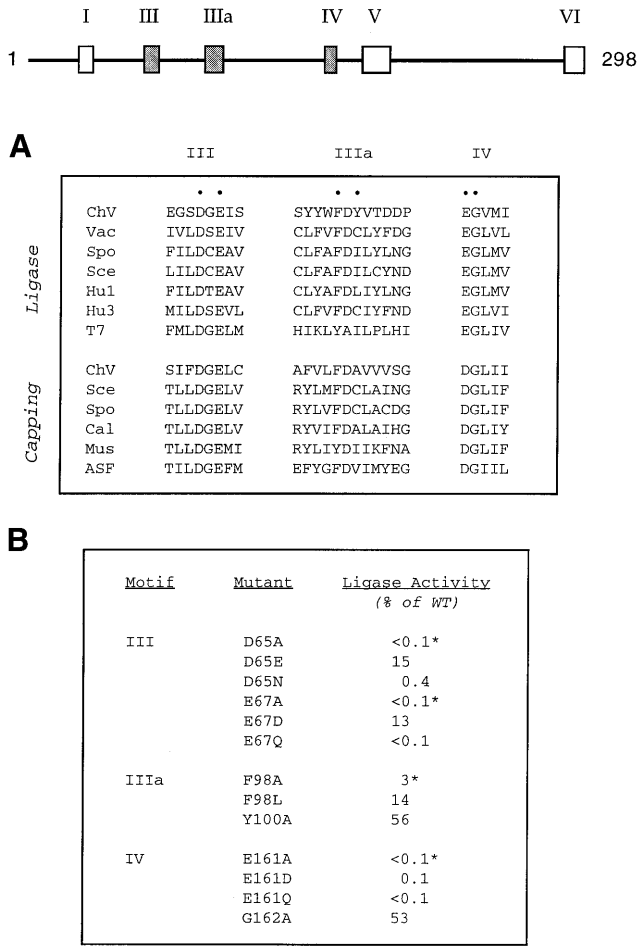

Figure 1. Mutational analysis of nucleotidyltransferase motifs III, IIIa and IV. The 298 amino acid *Chlorella* virus DNA ligase polypeptide is depicted as a straight line, with the positions of conserved motifs I, III, IIIa, IV, V and VI denoted by boxes. (**A**) The sequences of motifs III, IIIa and IV of *Chlorella* virus ligase (ChV) are aligned with the corresponding sequences of the following ATP-dependent DNA ligases: vaccinia virus (Vac), *Schizosaccharomyces pombe* (Spo), *S.cerevisiae* (Sce), human ligase I (Hu1), human ligase III (Hu3) and bacteriophage T7 (T7). Also included in the alignment are motifs III, IIIa and IV from the capping enzymes of *Chlorella* virus, *S.cerevisiae*, *S.pombe*, *Candida albicans* (Cal) mouse (Mus) and African swine fever virus (ASF). Amino acid residues of ChV ligase that were targeted for mutagenesis in the present study are denoted by dots. (**B**) The specific activities of wild-type (wild-type) and mutant ligases were determined by enzyme titration and the values were then normalized to the wild-type specific activity (defined as 100%). The specific activities of four of the alanine mutants were reported previously (3) and are denoted by *. All other values are derived from the present study.

MATERIALS AND METHODS

Ligase mutants

Missense mutations were introduced into the pET-ChVLig or pET-ChVLigCΔ5 expression plasmid as described (6,8). The entire ligase gene was sequenced in every case to confirm the desired mutation and exclude the acquisition of unwanted changes during PCR amplification and cloning. The expression plasmids were transformed into *Escherichia coli* BL21(DE3). Mutant and wild-type ligases were purified from the soluble lysates of IPTG-induced BL21(DE3) cells by Ni–agarose and phosphocellulose chromatography as described (6,8). The protein concentrations of the phosphocellulose enzyme preparations were determined using the Bio-Rad dye reagent with bovine serum albumin as the standard. SDS–PAGE analysis of the recombinant protein preparations is shown in Figure 2.

Assay of nick joining

Reaction mixtures (20 µl) containing 50 mM Tris–HCl, pH 7.5, 5 mM DTT, 10 mM $MgCl_2$, 1 mM ATP, 1 pmol 5′-^{32}P-labeled nicked duplex DNA substrate (6) and aliquots of serial 2-fold dilutions of wild-type or mutant ligases were incubated at 22°C for 10 min. The products were resolved by denaturing PAGE and the extents of ligation were determined by scanning the gel using a FUJIX BAS2000 phosphorimager. The specific activities of the wild-type and mutant ligases were determined from the slopes of the titration curves in the linear range of enzyme dependence.

Ligation at a pre-adenylated nick

The nicked DNA–adenylate substrate is shown in Figure 5. The 5′-adenylated ^{32}P-labeled 18mer strand was synthesized and purified by gel electrophoresis as described (11). The DNA–adenylate ligation reaction mixtures (20 µl) contained 50 mM Tris–HCl, pH 7.5, 5 mM DTT, 5 mM $MgCl_2$, 200 fmol nicked DNA–adenylate substrate and wild-type or mutant CΔ5 protein as specified. The mixtures were incubated for 30 min at 22°C. The products were resolved by denaturing PAGE and the extents of ligation were determined by scanning the gel with a phosphorimager. For kinetic analysis, reaction mixtures (20 µl) containing 200 fmol nicked DNA–adenylate substrate and other components as specified above were incubated at 22°C. The sealing reactions were initiated by adding enzyme. Aliquots (20 µl) were withdrawn at the times specified in the figures and quenched immediately with EDTA and formamide (6).

Binding of CΔ5 ligase to DNA–adenylate

Reaction mixtures (20 µl) containing 50 mM Tris–HCl, pH 7.5, 5 mM DTT, 200 fmol nicked DNA–adenylate and 4 pmol wild-type or CΔ5 mutants as specified were incubated for 10 min at 22°C. Glycerol was added to 5% and the samples were analyzed by electrophoresis (for 2 h at 70 V) through a native 6% polyacrylamide gel containing 90 mM Tris–borate, 2.5 mM EDTA (8). Free AppDNA and ligase–DNA complexes of retarded mobility were visualized by autoradiography of the dried gel.

RESULTS

Structure-based mutational strategy

Motifs III, IIIa and IV comprise β-strands that line the adenylate-binding pocket of domain 1 of *Chlorella* virus DNA ligase (3). The motifs are conserved in order and spacing in the primary structures of other ATP-dependent DNA ligases and in the GTP-dependent mRNA capping enzymes (Fig. 1A). Indeed, motifs III, IIIa and IV occupy similar positions in the tertiary structures of the *Chlorella* virus and T7 DNA ligases and in the *Chlorella* virus mRNA capping enzyme (2,3,5). Motif III contains two acidic side chains (Asp65 and Glu67)

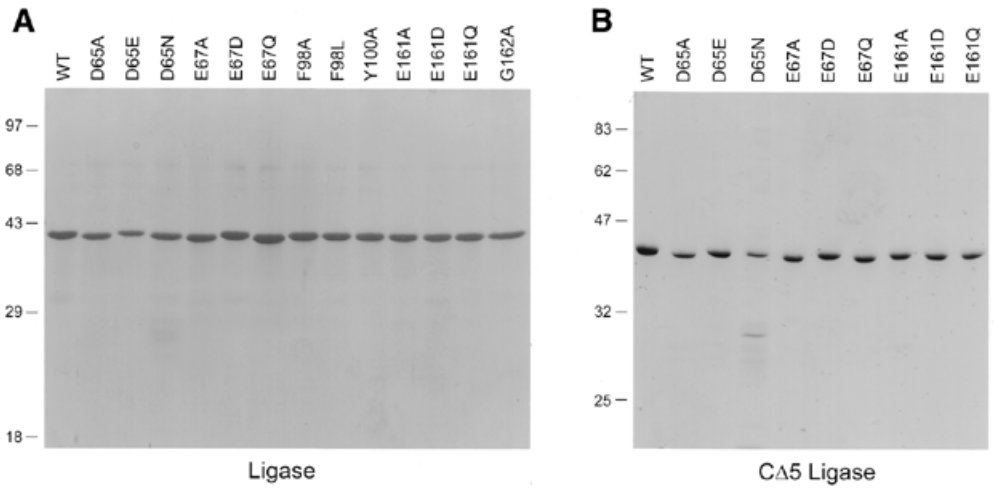

Figure 2. Protein purification. (**A**) Aliquots (1.5 μg) of the phosphocellulose fractions of full-length wild-type ligase and the indicated mutant proteins were analyzed by SDS–PAGE. (**B**) Aliquots (2 μg) of the phosphocellulose fractions of wild-type CΔ5 and the indicated mutant CΔ5 proteins were analyzed by SDS–PAGE. Polypeptides were visualized by staining the gel with Coomassie Brilliant Blue dye. The positions and sizes (in kDa) of marker proteins are indicated on the left.

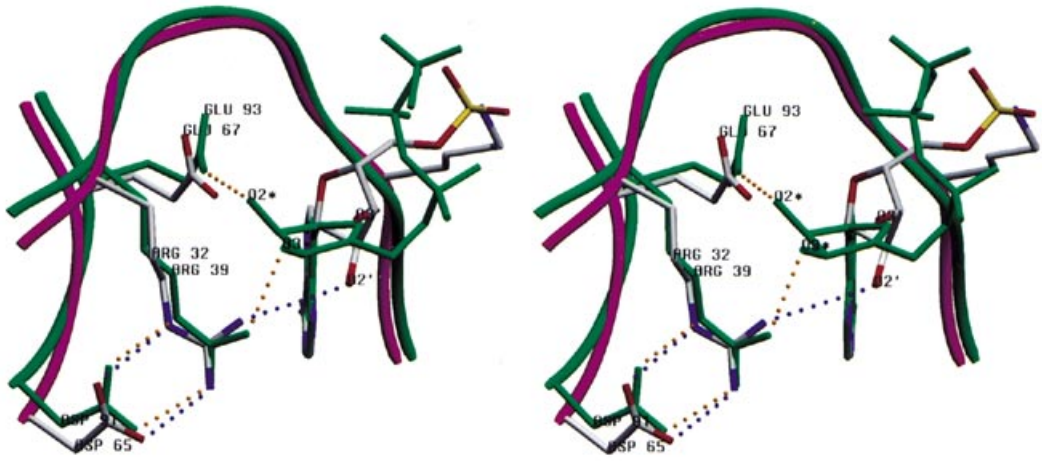

Figure 3. Stereo view of the adenylate-binding pocket. The figure shows interactions of the motif I and III side chains with the ribose sugar of the ChV lysyl-adenylate adduct. The ChV ligase polypeptide backbone is colored purple and the side chains are in CPK. The equivalent structural elements of the T7 ligase–ATP complex (colored in green) are aligned with the ChV structure based on superimposition of the adenine base. Side chain–side chain and side chain–sugar contacts are denoted by dashed lines, blue for ChV ligase and yellow for T7 ligase.

that are widely conserved in other members of the nucleotidyltransferase superfamily. Asp65 of motif III forms a salt bridge to Arg32 in motif I (^{27}KxDGxR) (depicted in Fig. 3 in CPK coloring); the identical ion pair is observed in the structure of T7 ligase (depicted in green in Fig. 3). Arg32 is essential for nick joining by *Chlorella* virus ligase (8). Glu67 of motif III does not contact the adenylate in the crystal structure of the *Chlorella* virus ligase–AMP intermediate, but it does contact the ribose 2′-O in the structures of T7 ligase with ATP (Fig. 2) and the capping enzyme–GMP intermediate (5). Motif IIIa contains a conserved aromatic residue (Phe98 in *Chlorella* virus ligase) that stacks on the purine base of the nucleotide. Motif IV is conserved among the covalent nucleotidyltransferases and consists of a (Glu/Asp)Gly dipeptide followed by a triplet of hydrophobic residues. The motif IV carboxylate (Glu161 in *Chlorella* virus ligase) may coordinate a divalent cation bound to the phosphate of AMP, as surmised by the position of a lutetium atom soaked into the ligase–adenylate crystal (3).

An initial test of the function of the conserved side chains of motifs III, IIIa and IV was conducted previously by replacing Asp65, Glu67, Phe98 and Glu161 individually by alanine and then scoring the ability of the mutated enzymes to seal a nicked duplex DNA substrate *in vitro* (3). The specific activities of the D65A, E67A and E161A proteins were <0.1% of the wild-type ligase activity and the F98A protein was 3% as active as wild-type ligase (Fig. 1B). Here we have extended the alanine scan to two additional residues, the invariant glycine of motif IV (Gly162) and Tyr100 in motif III, a position that is occupied by a hydrophobic residue in other DNA ligases (Fig. 1A). We

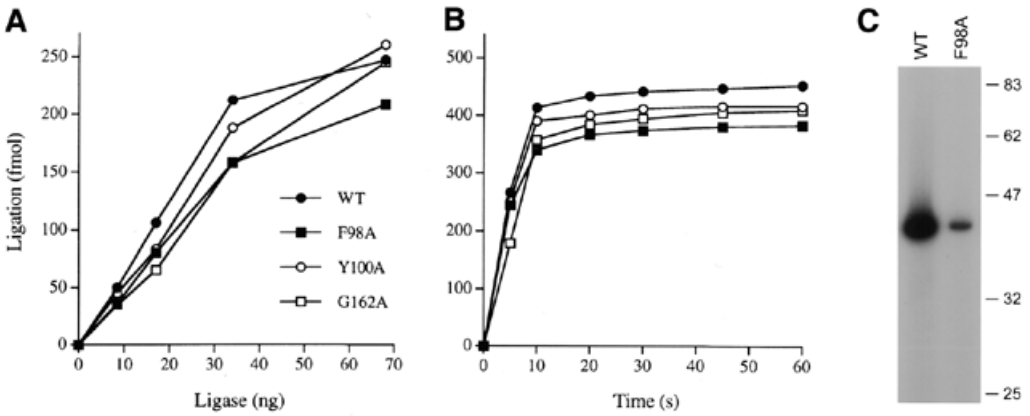

Figure 4. Single turnover nick ligation. (**A**) Ligase titration. Reaction mixtures (20 µl) containing 50 mM Tris–HCl, pH 7.5, 5 mM DTT, 10 mM $MgCl_2$, 500 fmol 5′-^{32}P-labeled nicked duplex DNA substrate and wild-type, F98A, Y100A or G162A ligase as specified were incubated at 22°C for 10 min. The extent of strand joining is plotted as a function of input ligase. (**B**) Kinetics. Reaction mixtures (20 µl) containing 50 mM Tris–HCl, pH 7.5, 5 mM DTT, 10 mM $MgCl_2$, 600 fmol 5′-^{32}P-labeled nicked duplex DNA substrate and 136 ng of wild-type, F98A, Y100A or G162A ligase were incubated at 22°C. The reactions were initiated by addition of ligase. Aliquots (20 µl) were withdrawn at the times specified and quenched immediately with EDTA and formamide. The extent of ligation is plotted as a function of time. (**C**) Ligase adenylation. Reaction mixtures (20 µl) containing 50 mM Tris–HCl, pH 8.0, 5 mM DTT, 5 mM $MgCl_2$, 5 µM [α-^{32}P]ATP and 4 pmol ligase preparation as indicated were incubated for 5 min at 22°C. Reactions were quenched by adding SDS to 1%. The reaction products were resolved by SDS–PAGE. An autoradiogram of the dried gel is shown.

found that the purified recombinant G162A and Y100A proteins (Fig. 2A) were 53 and 56% as active in nick joining as the wild-type ligase under steady-state conditions (Fig. 1B). We have imposed a 5-fold activity decrement as the criterion of significance for the effects of alanine substitution. Residues are deemed 'important' when alanine substitution reduces specific activity to 6–20% of the wild-type value. Our operational definition of an 'essential' residue is one at which alanine substitution reduces specific activity to ≤5% of the wild-type activity. By these criteria, Asp65, Glu67, Phe98 and Glu161 are essential for the overall ligation reaction, whereas Tyr100 and Gly162 do not contribute significantly to ligation.

Phe98 is not required for single turnover nick ligation

Nick joining by the wild-type ligase, the weakly active F98A protein, and the minimally affected Y100A and G162A proteins was assayed under single turnover conditions, i.e. in the absence of added ATP. The linear dependence of the extent of nick joining on the amount of input wild-type ligase suggested that 15% of the enzyme molecules in the preparation contained covalently bound AMP (Fig. 4A). The titration profiles of the F98A, Y100A and G162A proteins were similar to that of wild-type ligase, indicating that ligase–adenylate comprised 10–13% of the mutant enzyme preparations. We then performed a kinetic analysis of single turnover strand joining by wild-type and mutant ligases added in 6-fold molar excess over the nicked DNA substrate. The reactions proceeded to similar end-points, with ~55–65% of the input substrate being sealed in 10–20 s (Fig. 4B). The rates of approach to the end-point were virtually identical for the wild-type and mutant enzymes. The instructive finding was that F98A was just as active as the wild-type ligase in nick joining catalyzed by pre-formed ligase–adenylate (reflecting steps 2 and 3 of the ligase pathway), but was only 3% as active under steady-state conditions. Additional enzyme titration experiments showed that whereas wild-type ligase was stimulated 50-fold by 1 mM ATP, the activity of F98A was enhanced by only a factor of 2 (not shown). We surmise that F98A is specifically defective in step 1 of the reaction of ligase with free ATP to form ligase–adenylate. (The steady-state activity of F98A was too low to reliably determine a K_m for ATP.) A comparison of the reaction of wild-type ligase and F98A with [α-^{32}P]ATP *in vitro* to form a covalent ligase–[^{32}P]AMP adduct verified that F98A was impaired in the ligase adenylation step (Fig. 4C).

These findings and inferences are in keeping with the crystallographic evidence that Phe98 stacks on the adenine base of ATP (3). To further explore the role of the aromatic ring, we substituted Phe98 conservatively by leucine. (We eschewed testing the effects of a tyrosine in light of the fact that other ligases and capping enzymes, including T7 ligase, naturally have a tyrosine at the equivalent position in motif IIIa.) We found that the F98L change partially restored ligase function to 14% of the wild-type specific activity (Fig. 1B). We conclude that: (i) a hydrophobic interaction of the motif IIIa residue with the adenine base of ATP is critical for *Chlorella* virus ligase activity; (ii) a π stack of adenine on an aromatic side chain is superior to interactions with an aliphatic side chain.

Effects of conservative motif III and IV mutations on nick joining

Motif III contains two essential acidic residues and motif IV contains a single essential carboxylate (Fig. 1B). To evaluate the roles of charge, hydrogen bonding potential and steric constraints in the functions of these amino acids, we replaced Glu67 and Glu161 by glutamine and aspartate and substituted Asp65 by asparagine and glutamate. The recombinant mutant ligases were purified from soluble bacterial extracts by Ni–agarose and phosphocellulose column chromatography (Fig. 2A). The specific activity of each mutant was determined under steady-state conditions by protein titration and normalized to the specific activity of wild-type ligase; the results are summarized in Figure 1B. The salient findings were that the D65N, E67Q

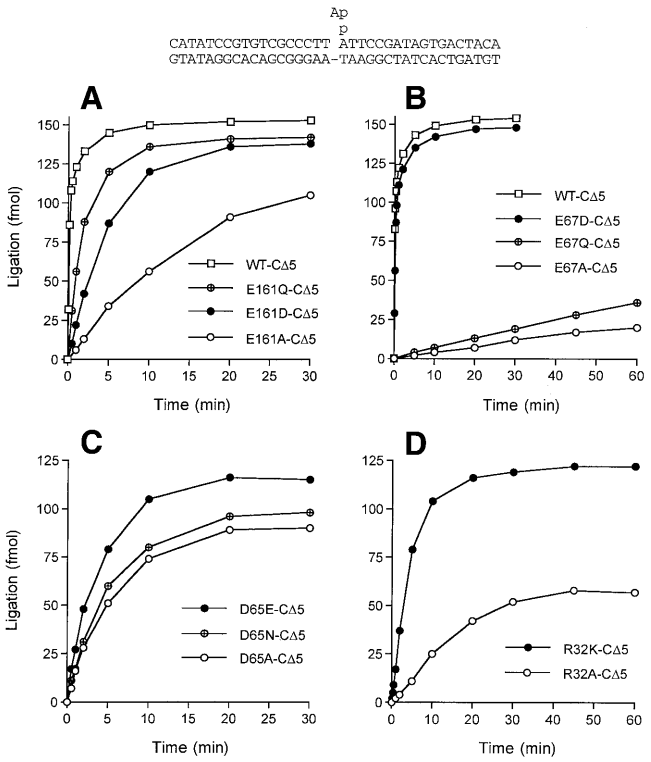

Figure 5. Effects of Glu161, Glu67, Asp65 and Arg32 mutations on strand joining at a pre-adenylated nick. The structure of the nicked DNA–adenylate substrate is shown. (**A**) Reaction mixtures (20 µl) contained 200 fmol nicked DNA–adenylate substrate and 1 pmol wild-type, E161A, E161D or E161Q CΔ5. (**B**) Reaction mixtures (20 µl) contained 200 fmol nicked DNA–adenylate substrate and 2 pmol wild-type, E67A, E67D or E67Q CΔ5. (**C**) Reaction mixtures (20 µl) contained 200 fmol nicked DNA–adenylate substrate and 2 pmol wild-type, D65A, D65E or D65N CΔ5. (**D**) Reaction mixtures (20 µl) contained 200 fmol nicked DNA–adenylate substrate and 2 pmol R32A or R32K CΔ5. The extent of ligation is plotted as a function of reaction time.

and E161Q mutations abolished ligase activity, indicating that a carboxylate group is critical at all three positions. Introduction of a glutamate in place of Asp65 or an aspartate in place of Glu67 in motif III partially restored ligase activity to 15 and 13% of the wild-type level, respectively. This represents at least a two order of magnitude gain of function compared to the catalytically inert D65A and E67A mutants. We surmise that lengthening or shortening the distance from the main chain to the motif III carboxylates can be tolerated with a significant, but not catastrophic, loss of activity compared to wild-type ligase. In contrast, the E161D change in motif IV failed to revive the nick joining activity. Contraction of the main chain–carboxylate distance of the motif IV glutamate is apparently not tolerated.

Effects of Asp65, Glu67 and Glu161 mutations on ligase–AMP formation

As noted above, incubation of wild-type ligase in the presence of [α-^{32}P]ATP and a divalent cation resulted in the formation of a nucleotide–protein adduct. The D65A, E67A and E161A mutants were apparently inert in enzyme–adenylate formation (not shown). Thus, their loss of function in the nick joining reaction can be attributed to a failure to form the ligase–adenylate intermediate, either *in vitro* or during expression of the mutant proteins in bacteria. Conservative replacements of the motif III and IV carboxylates with amide groups (D65N, E67Q and E161Q) also abrogated ligase–adenylate formation (data not shown), just as they did the composite nick ligation reaction. The D65E mutant displayed reduced activity in ligase adenylation, consistent with its reduced level of nick joining. The E67D mutant formed trace levels of the AMP adduct. We also observed trace levels of adenylation of the E161D mutant, which was severely defective in the composite ligation reaction (not shown).

Effects of Asp65, Glu67 and Glu161 mutations on phosphodiester formation at a pre-adenylated nick

Although effects on step 1 sufficed to explain the loss of activity upon removal of the motif III and IV carboxylates, we were interested in determining whether the step 1 catalytic residues are also required for downstream steps of the ligase reaction pathway. We focused on the roles of individual side chains in the third step of the ligation reaction (phosphodiester formation), because this chemical step is unique to the polynucleotide ligases and not part of the repertoire of the mRNA capping enzymes.

Step 3 of the *Chlorella* virus ligase reaction can be studied in isolation by assaying the ability of wild-type and mutant enzymes to seal a pre-adenylated nicked duplex DNA (8) (Fig. 5). Ligase reacts with this substrate in the absence of ATP to catalyze phosphodiester bond formation. In bypassing the requirement for steps 1 and 2, we can assess the capacity of step 1-defective or step 2-defective mutants to recognize the nicked DNA–adenylate and catalyze strand closure. Defects in step 3 catalysis can be gauged only qualitatively by assays of full-length wild-type and mutant ligases. A quantitative analysis of mutational effects on step 3 requires that the mutations of interest be transferred into the CΔ5 deletion variant of *Chlorella* virus ligase, which lacks the C-terminal five amino acids that comprise part of nucleotidyltransferase motif VI. The rate of ligation of a pre-adenylated nick by CΔ5 is 16-fold faster than the rate of full-length ligase (6). We propose that the loss of motif VI overcomes a rate limiting conformational step that is unique to the reaction of ligase with an exogenous nicked DNA–adenylate substrate (6). The relevant issue for the structure–function analysis is that mutational effects on step 3 catalysis can be detected and quantitated with at least 10-fold higher sensitivity in the CΔ5 background than they can in the context of the full-sized ligase.

We introduced into the CΔ5 protein a collection of alanine and conservative substitutions at essential motif III residues Asp65 (D65A, D65E and D65N) and Glu67 (E67A, E67D and E67Q) and motif IV residue Glu161 (E161A, E161D and E161Q). The 'wild-type' CΔ5 and the nine mutants were purified from soluble bacterial extracts by Ni–agarose and phosphocellulose chromatography (Fig. 2B).

The extent of sealing of the nicked DNA–adenylate (200 fmol input substrate) during a 30 min reaction was proportional to the amount of input wild-type CΔ5; 75% of the input substrate was ligated at saturating levels of enzyme (≥0.5 pmol wild-type CΔ5; data not shown). A kinetic analysis of the step 3 reaction in enzyme excess is shown in Figure 5A. The wild-type CΔ5 reaction proceeded rapidly, attaining half of the end-point

value in ~10 s (k_{obs} = 4.0 min^{-1}). The Glu161 mutants reacted slowly (Fig. 5A). The apparent rate constants were 0.06 min^{-1} for E161A CΔ5, 0.18 min^{-1} for E161D CΔ5 and 0.48 min^{-1} for E161Q CΔ5. Comparing E161A to wild-type CΔ5, we conclude that the Glu161 side chain accelerates the rate of the step 3 reaction by a factor of 60. Conservative substitutions with aspartate and glutamine restored activity by 3- and 8-fold, respectively, above that of E161A. The partial losses of function elicited by Glu161 mutations in the sealing of DNA–adenylate contrast sharply with their total abrogation of the full nick joining reaction (Fig. 1B). Whereas the carboxylate functional group is required absolutely for overall nick joining and for ligase adenylation, we find that the isolated reaction of ligase with DNA–adenylate is reduced by only a factor of 8 when a glutamine is present at position 161. We conclude that the invariant motif IV carboxylate plays a dominant essential role in step 1 catalysis and only a supportive 'important' role in step 3, with differences in structure–activity relationships during the two steps.

The E67A and E67Q changes in the CΔ5 ligase elicited marked decreases in phosphodiester formation. Kinetic analysis of step 3 in enzyme excess showed that the E67A and E67Q CΔ5 proteins displayed linear kinetics for 60 min (longer times were not tested) (Fig. 5B). Although we could not determine rate constants for these mutants without an end-point, the initial rates were ≤0.1% of the initial rate of wild-type CΔ5. Thus, the Glu67 carboxylate enhances the rate of phosphodiester formation by about three orders of magnitude. An instructive finding was that the E67D mutation revived step 3 activity to near wild-type levels. This was apparent in the recovery of rapid reaction kinetics (k_{obs} = 2.5 min^{-1}) (Fig. 5B). Thus, the enzyme tolerates contraction of the linker from the main chain to the carboxylate.

The effects of Asp65 mutations on step 3 are shown in Figure 5C. The D65A and D65N CΔ5 mutants attained saturation at a slightly lower extent of strand joining than wild-type CΔ5. The apparent rate constants were 0.24 min^{-1} for D65E CΔ5, 0.18 min^{-1} for D65N CΔ5 and 0.18 min^{-1} for D65A CΔ5. Thus, the Asp65 carboxylate accelerates the rate of the step 3 reaction by a factor of 20.

Effects of motif I Arg32 mutations on phosphodiester formation at a pre-adenylated nick

The Asp in motif III makes no direct contact with ATP or AMP in the DNA ligase crystal structures (Fig. 3), yet we find that Asp65 is essential for ligase–adenylate formation and phosphodiester bond formation. The crystal structure shows that Asp65 makes a bidentate salt bridge to the conserved Arg32 in motif I. Because this arginine coordinates the ribose sugar of the covalently bound AMP (and of ATP in the T7 ligase crystal), we suspected that the deleterious effects of Asp65 mutations might reflect the key role of the carboxylate in positioning the Arg32 side chain. We had shown previously that a R32A mutation abrogates the overall nick joining reaction and ligase–adenylate formation (8). Thus, it was of interest to determine the contribution of Arg32 to step 3 and compare it to the results obtained for Asp65. To do so, we engineered the R32A mutation and the conservative change R32K into the CΔ5 protein. The purities of recombinant R32A and R32K CΔ5 were equivalent to that of the wild-type CΔ5 in Figure 2B (data not shown). The apparent rate constants for step 3 ligation in enzyme excess were 0.06 min^{-1} for R32A CΔ5 and 0.18 min^{-1} for R32K CΔ5 (Fig. 5D). Thus, the Arg32 guanidinium group accelerates the rate of the step 3 reaction by ~60-fold and a lysine does not suffice for performance of this step. The results imply a requirement for polyvalent interactions of the arginine with the ribose sugar and the motif III aspartate and they are consistent with the mutational effects at Asp65 being consequent to indirect effects on Arg32.

Effects of motif I Asp29 mutations on phosphodiester formation at a pre-adenylated nick

The conserved Asp29 side chain in motif I is essential for the composite nick joining reaction, but is not required for ligase adenylation step 1 (8). Indeed, the crystal structure of the *Chlorella* virus ligase–AMP intermediate was initially determined with the D29A mutant (3). The D29A–adenylate intermediate is defective in the transfer of AMP to the 5′-phosphate of the nick (step 2). Kinetic analysis of the reaction of D29A–adenylate with nicked DNA under single turnover conditions showed that Asp29 contributes an ~6000-fold enhancement of the reaction rate (9). As there was no accumulation of DNA–adenylate during this reaction, it was inferred that DNA–adenylate formation (step 2) was rate limiting for D29A. Does Asp29 contribute at all to the step 3 reaction? To explore this issue, we introduced D29A, D29E and D29N changes into the CΔ5 background (Fig. 6A). The titration profiles for D29E and D29N CΔ5 in the sealing of nicked DNA–adenylate were nearly identical to that of wild-type CΔ5 (Fig. 6C), whereas D29A CΔ5 displayed a lower extent of ligation during the standard 30 min reaction. The rates of step 3 ligation by D29E and D29N CΔ5 in enzyme excess were identical to the rate of wild-type CΔ5 (not shown). In contrast, the D29A CΔ5 enzyme reacted slowly (k_{obs} = 0.06 min^{-1}) (Fig. 6B). Thus, the Asp29 residue accelerates the rate of the step 3 reaction by ~60-fold, but the essential function of this side chain during step 3 can be fulfilled by either glutamate or asparagine. We infer that hydrogen bonding contacts, rather than electrostatic interactions, underlie Asp29 function during step 3. This is not the case for the composite ligation reaction, wherein the D29N mutant is severely impaired (8).

Mutational effects on binding of CΔ5 ligase to DNA–adenylate

Mixture of 200 nM wild-type CΔ5 with 10 nM nicked DNA–adenylate in the absence of a divalent cation (omission of which precludes phosphodiester formation) resulted in the formation of a discrete ligase–AppDNA complex that migrated slower than free DNA–adenylate during native gel electrophoresis (Figs 6B and 7). The majority of the input DNA–adenylate was bound to the ligase under these conditions. We exploited this assay to gauge mutational effects on the DNA–adenylate binding step of the isolated step 3 reaction. We found that CΔ5 mutants E67A, E67D, E67Q, E161A, E161D and E161Q bound to nicked DNA–adenylate (Fig. 7), as did mutants D29A, D29E and D29N (Fig. 6B). In contrast, we detected no stable binding of CΔ5 mutants D65A and D65N to DNA–adenylate and only trace levels of ligase–AppDNA complex with mutants R32A, R32K and D65E (Fig. 7). These results highlight two classes of functional groups that participate in the isolated step 3 reaction: those that are essential for stable binding to DNA–adenylate and for strand sealing (Arg32 and

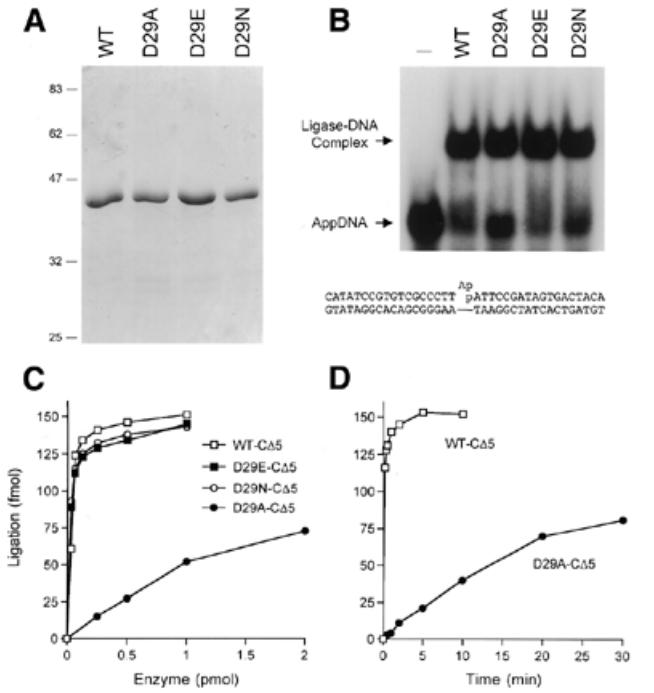

Figure 6. Effects of Asp29 mutations on step 3. (**A**) Protein purification. Aliquots (5 μg) of the phosphocellulose fractions of wild-type and the indicated Asp29 mutant CΔ5 proteins were analyzed by SDS–PAGE. Polypeptides were visualized by staining the gel with Coomassie Brilliant Blue dye. (**B**) Binding of the indicated CΔ5 ligases to the nicked DNA–adenylate substrate was assayed as described in Materials and Methods. An autoradiogram of the native gel is shown. (**C**) Reaction mixtures containing 200 fmol nicked DNA–adenylate and CΔ5 ligase as specified were incubated for 30 min at 22°C. The extent of ligation is plotted as a function of input enzyme. (**D**) Reaction mixtures (20 μl) contained 200 fmol nicked DNA–adenylate substrate and 2 pmol wild-type or D29A CΔ5. Ligation is plotted as a function of reaction time.

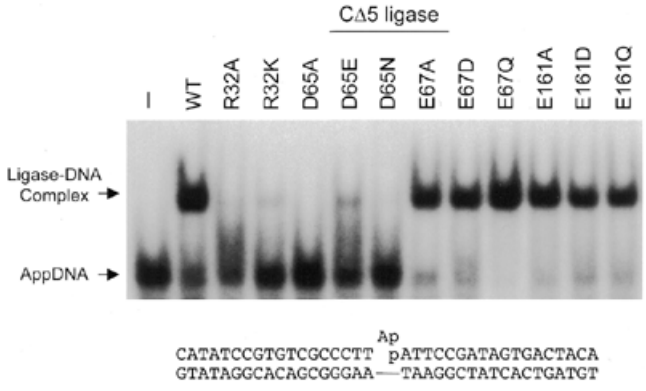

Figure 7. Mutational effects on binding of CΔ5 ligase to DNA–adenylate. Binding of the indicated CΔ5 ligases to the nicked DNA–adenylate substrate was assayed as described in Materials and Methods. An autoradiogram of the native gel is shown. The positions of free AppDNA and the ligase–AppDNA complex are indicated by arrows.

Asp65) and those required for the step 3 reaction chemistry, but not for stable binding to a pre-adenylated nick (Asp29, Glu67 and Glu161).

Mutations D65A, E67A and E161A abolish ligase activity *in vivo*

Budding yeast provides a surrogate genetic assay for scoring mutational effects on *Chlorella* virus DNA ligase function *in vivo* (10). Viability of a *Saccharomyces cerevisiae cdc9*Δ strain deleted at the chromosomal locus encoding the essential Cdc9 DNA ligase is contingent on maintenance of an extrachromosomal *CDC9* gene on a *CEN URA3* plasmid. Hence, *cdc9*Δ cells cannot grow on medium containing 5-FOA, a drug that selects against the *URA3 CDC9* plasmid. *cdc9*Δ cells can grow on 5-FOA if they have been transformed with a second *CEN TRP1* plasmid containing the wild-type *Chlorella* virus ligase gene driven by a constitutive yeast promoter (10). Here we found that the *D65A*, *E67A* and *E161A* ligase alleles were unable to complement *cdc9*Δ (data not shown). Thus, the *in vivo* lethality of motif III and IV mutations correlated with their profound catalytic defects in nick joining *in vitro*. Motif I mutants *D29A* and *R32A* were also lethal *in vivo* (10).

DISCUSSION

ATP-dependent DNA ligases and GTP-dependent mRNA capping enzymes represent distinct branches of a covalent nucleotidyltransferase superfamily defined by a common protein fold (2–6). Ligases and capping enzymes catalyze two chemically similar partial reactions entailing the transformation of a phosphoanhydride bond (in ATP or GTP) into a phosphoamide (lysyl-AMP or lysyl-GMP) and then back into a phosphoanhydride (AppDNA or GpppRNA). Whereas the guanylyltransferase reaction is complete after formation of GpppRNA, the DNA ligases catalyze a third step in which a phosphoanhydride is transformed into a phosphodiester. The peptide motifs that comprise the nucleotide-binding pocket are conserved between DNA ligases and capping enzymes, suggesting that specific functional groups might be put to similar uses in ligation and capping (12). But, what about the third step? Are the same residues also exploited by DNA ligases in catalysis of strand sealing and, if so, do these residues function similarly at each step in the pathway?

The present study provides an analysis of the roles of specific functional groups in step 3 of the DNA ligation reaction. We focused first on conserved carboxylates of motifs III (Asp65 and Glu57) and IV (Glu161) that are essential for nick joining by *Chlorella* virus DNA ligase and, as we show here, are required for the reaction of ligase with ATP to form ligase–adenylate. We found that although all three acidic residues also participate in the conversion of DNA–adenylate into a phosphodiester, their quantitative contributions to step 3 and the effects of conservative structural changes on step 3 did not inevitably correlate with catalytic defects in the composite ligation reaction. The same is true of Asp29 in motif I. As discussed in detail below, our findings suggest that the active site of DNA ligase is remodeled during the three steps of the pathway and that at least some of the catalytic side chains may enhance activity by a different mechanism at different steps.

Comparison of the available crystal structures indicates that the nucleotide binding interactions of DNA ligase are altered after the ligase–adenylate intermediate is formed. The adenosine nucleoside of the *Chlorella* virus ligase–AMP adduct is in the *anti* conformation (Fig. 3). The adenosine is also in the *anti*

conformation in the *Thermus filiformis* ligase–AMP crystal (13). This is in contrast to the *syn* conformation of adenosine in the crystal of T7 ligase bound to ATP (Fig. 3). In addition to a shift of the nucleoside from *syn* to *anti*, the structure of *Chlorella* virus ligase–AMP reveals a remodeling of the contacts between the enzyme and the ribose oxygens (Fig. 3). The 2′-O of ligase–AMP is engaged by the side chain of Arg32 in motif I, which is itself held in place by the salt bridge to Asp65 in motif III. The T7 ligase has the same network of interactions between the motif I Arg and motif III Asp, except that the motif I Arg makes a hydrogen bond to the ribose 3′-O, rather than the 2′-O seen in the ligase–AMP structure (Fig. 3). The contact between the ribose 2′-O and the conserved glutamate side chain of motif III characteristic of the ground states in the T7 ligase–ATP and capping enzyme–GTP crystals is absent in the ligase–adenylate intermediate. The present finding that elimination of the Glu67 carboxylate abolishes ligase adenylation provides evidence that the ground state of the *Chlorella* virus ligase with respect to the sugar contacts resembles that of the T7 ligase–ATP complex, not that of ligase–adenylate. Because Glu67 is not located near the site of step 1 chemistry (i.e. it interacts neither with the α-phosphate nor the motif I lysine nucleophile), we infer that the essential function of Glu67 is to hold the adenosine nucleotide in the proper conformation within the active site pocket via a hydrogen bond to the 2′-O. This model explains why dATP is an extremely poor substrate for ligase adenylation (7), i.e. elimination of the 2′-OH group phenocopies the effect of removing the Glu67 side chain that coordinates the 2′-OH.

A novel insight of the present study is that Glu67 is essential for step 3, where it provides a rate enhancement of about three orders of magnitude. Our interpretation of this finding in light of the crystal structures is that the nucleoside conformation and the specific sugar contacts seen in the ligase–AMP intermediate (which we propose reflects the state of the ligase–adenylate immediately prior to step 2 catalysis) are not the ones that prevail during step 3. Indeed, the functional data suggest that the active site is remodeled again after completion of step 2, such that the nucleoside reverts to the *anti* conformation and the contact of Glu67 to the ribose 2′-O is restored. Glu67 is apparently not strictly required for the initial binding of ligase to exogenous DNA–adenylate (Fig. 7); rather, we speculate that its essential role is to fix the orientation of the adenosine and position the AMP phosphate (the step 3 leaving group) apical to the attacking 3′-OH nucleophile. Aspartate can functionally replace Glu67 in phosphodiester formation, whereas glutamine cannot, implying that the carboxylate is a hydrogen bond acceptor in its proposed contact with the ribose.

Figure 3 illustrates that a shift in the motif I Arg contact from the ribose 3′-O in the ground state to the 2′-O in the *Chlorella* virus ligase–adenylate leaves the 3′-O without any direct contacts to the protein. Instead, the 3′-O coordinates a sulfate on the surface of the protein, which is believed to mimic the 5′-phosphate at the DNA nick (3). In keeping with the above proposal for remodeling of the adenylate-binding pocket after step 2, we envision that the requirement for Arg32 in step 3 (where it contributes a 60-fold rate enhancement) entails the restitution of contacts to the ribose 3′-O. Arg32 (unlike Glu67) is implicated in the initial binding of ligase to exogenous DNA–adenylate (Fig. 7) and the conservative R32K mutant remains defective in DNA–adenylate binding and sealing. The requirement for the polyvalent guanidinium group at position 32 reflects the critical role of the salt bridge to Asp65. Indeed, disruption of the salt bridge by replacement of Asp65 with either Ala or Asn phenocopies the Arg32 mutations with respect to the resulting defects in binding and sealing of the DNA–adenylate substrate.

Each of the three steps in the ligation pathway of *Chlorella* virus DNA ligase is strictly dependent on a divalent cation cofactor. Glu161 in motif IV is implicated in metal binding to the ligase–adenylate intermediate. A lutetium atom soaked into the crystal was coordinated by the non-bridging phosphate oxygen of the covalently bound adenylate and the carboxylate side chain of Glu161 (3). The structure suggested a catalytic mechanism whereby the metal ion activates the α-phosphate for attack by the motif I lysine (in step 1) or the DNA 5′-PO_4 (in step 2) and stabilizes a pentacoordinate transition state. Structure–activity relationships for Glu161 are consistent with a metal binding function in the overall nick joining pathway and step 1 in particular, i.e. changing glutamate to glutamine abolished both of these reactions. Remarkably, the Glu161 carboxylate plays a quantitatively less important role in step 3, where substitution by glutamine restored activity to 12% of the wild-type level. This result hints that the metal-binding site(s) may change during the different steps of the nick joining reaction. Indeed, whereas it is sensible that coordination of a metal to the AMP phosphate is critical for steps 1 and 2 (because the AMP phosphate is where the chemistry is occurring), step 3 entails chemistry at the DNA 5′-phosphate of DNA–adenylate and AMP is the leaving group. Therefore, in a new metal-binding site formed during step 3, Glu161 may not make a direct contact to the divalent cation.

Asp29 in motif I is essential for overall nick joining, but not for the formation of ligase–adenylate. Asp29 plays a critical role in step 2, the transfer of AMP from ligase–adenylate to the DNA 5′-phosphate to form DNA–adenylate. Here we have shown that Asp29 also plays a role in step 3, enhancing the rate of closure of nicked DNA–adenylate by a factor of 60. The contribution of Asp29 to step 3 is less than the rate enhancement of 6000 that Asp29 lends to the step 2 plus step 3 reaction (single turnover nick ligation) (9). The mechanistically instructive finding was that Asn substitution for Asp29 reduced the rate constant for single turnover nick ligation to 2% of the wild-type rate (8), but, as noted here, had no discernable impact on the rate of step 3. Thus, the carboxylate moiety is critical for catalysis of step 2, but an amide suffices during step 3. We infer that the Asp29 side chain functions in different ways during the second and third steps of the reaction pathway. The crystal structure of *Chlorella* virus ligase suggested that Asp29 would be in a position to interact with the metal ion in a lutetium soak of the D29A–AMP crystal (3). If Asp29 is a metal-binding residue, then it appears to function as such only during the catalysis of step 2. This contrasts with Glu161, which coordinates to lutetium and (unlike Asp29) is essential for step 1. It is conceivable that the metal-binding site, like the nucleoside-binding pocket, is remodeled as the reaction proceeds, such that Glu161 is the critical metal-coordinating residue for step 1 and Asp29 is critical during step 2, either *per se* or in concert with Glu161. (We were unable to assess the role of Glu161 in step 2, because the E161A mutant cannot be adenylated and hence cannot perform step 2.) Given that D29N appears to be fully active in sealing DNA–adenylate, we suspect that Asp29

may not be directly coordinating a metal during step 3 of the pathway.

ACKNOWLEDGEMENT

Supported by NIH grant GM63611.

REFERENCES

1. Lehman,I.R. (1974) DNA ligase: structure, mechanism and function. *Science*, **186**, 790–797.
2. Subramanya,H.S., Doherty,A.J., Ashford,S.R. and Wigley,D.B. (1996) Crystal structure of an ATP-dependent DNA ligase from bacteriophage T7. *Cell*, **85**, 607–615.
3. Odell,M., Sriskanda,V., Shuman,S. and Nikolov,D.B. (2000) Crystal structure of eukaryotic DNA ligase–adenylate illuminates the mechanism of nick sensing and strand joining. *Mol. Cell*, **6**, 1183–1193.
4. Shuman,S. and Schwer,B. (1995) RNA capping enzyme and DNA ligase—a superfamily of covalent nucleotidyl transferases. *Mol. Microbiol.*, **17**, 405–410.
5. Håkansson,K., Doherty,A.J., Shuman,S. and Wigley,D.B. (1997) X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell*, **89**, 545–553.
6. Sriskanda,V. and Shuman,S. (1998) Mutational analysis of *Chlorella* virus DNA ligase: catalytic roles of domain I and motif VI. *Nucleic Acids Res.*, **26**, 4618–4625.
7. Ho,C.K., Van Etten,J.L. and Shuman,S. (1997) Characterization of an ATP-dependent DNA ligase encoded by *Chlorella* virus PBCV-1. *J. Virol.*, **71**, 1931–1937.
8. Sriskanda,V. and Shuman,S. (1998) *Chlorella* virus DNA ligase: nick recognition and mutational analysis. *Nucleic Acids Res.*, **26**, 525–531.
9. Odell,M. and Shuman,S. (1999) Footprinting of *Chlorella* virus DNA ligase bound at a nick in duplex DNA. *J. Biol. Chem.*, **274**, 14032–14039.
10. Sriskanda,V., Schwer,B., Ho,C.K. and Shuman,S. (1999) Mutational analysis of *E. coli* DNA ligase identifies amino acids required for nick-ligation *in vitro* and for *in vivo* complementation of the growth of yeast cells deleted for *CDC9* and *LIG4*. *Nucleic Acids Res.*, **27**, 3953–3963.
11. Sekiguchi,J. and Shuman,S. (1997) Nick sensing by DNA ligase requires a 5′ phosphate at the nick and occupancy of the adenylate binding site on the enzyme. *J. Virol.*, **71**, 9679–9684.
12. Wang,S.P., Deng,L., Ho,C.K. and Shuman,S. (1997) Phylogeny of mRNA capping enzymes. *Proc. Natl Acad. Sci. USA*, **94**, 9573–9578.
13. Lee,J.Y., Chang,C., Song,H.K., Moon,J., Yang,J., Kim,H.K., Kwon,S.T. and Suh,S.W. (2000) Crystal structure of NAD^+-dependent DNA ligase: modular architecture and functional implications. *EMBO J.*, **19**, 1119–1129.