Structure-Function Analysis of T4 RNA Ligase 2*

Shenmin Yin, C. Kiong Ho, and Stewart Shuman‡

From the Molecular Biology Program, Sloan-Kettering Institute, New York, New York 10021

Received for publication, January 24, 2003, and in revised form, February 27, 2003
Published, JBC Papers in Press, February 27, 2003, DOI 10.1074/jbc.M300817200

Bacteriophage T4 RNA ligase 2 (Rnl2) exemplifies a polynucleotide ligase family that includes the trypanosome RNA-editing ligases and putative RNA ligases encoded by eukaryotic viruses and archaea. Here we analyzed 12 individual amino acids of Rnl2 that were identified by alanine scanning as essential for strand joining. We determined structure-activity relationships via conservative substitutions and examined mutational effects on the isolated steps of ligase adenylylation and phosphodiester bond formation. The essential residues of Rnl2 are located within conserved motifs that define a superfamily of nucleotidyl transferases that act via enzyme-(lysyl-N)-NMP intermediates. Our mutagenesis results underscore a shared active site architecture in Rnl2-like ligases, DNA ligases, and mRNA capping enzymes. They also highlight two essential signature residues, Glu34 and Asn46, that flank the active site lysine nucleophile (Lys35) and are unique to the Rnl2-like ligase family.

RNA ligases join 3’-OH and 5’-PO4 RNA termini through a series of three nucleotidyl transfer steps similar to the pathway used by DNA ligases (1–5). Step 1 is the reaction of ligase with ATP to form a covalent ligase-(lysyl-N)-AMP intermediate and pyrophosphate. In step 2, the AMP is transferred from ligase adenylate to the 5’-PO4 RNA end to form an RNA-adenylate intermediate (AppRNA). In step 3, attack by an ATP 3’-OH on the RNA-adenylate seals the two ends via a phosphodiester bond and releases AMP. Bacteriophage T4 RNA ligase 1 (Rnl1) is the founding member of the RNA ligase family (1). The active site lysine of Rnl1 is located within a conserved sequence element, KX/D/N/G (motif 1), that defines a superfamily of covalent nucleotidyl transferases, which includes DNA ligases and mRNA capping enzymes (6–8). DNA ligases and capping enzymes share a common tertiary structure composed of five conserved motifs of Rnl2 (13). We showed that Lys35 in motif I, Glu204 in motif IV, and Lys225 and Lys227 in motif V are required for overall strand joining and particularly for the enzyme adenylation reaction (step 1) of the ligation pathway. His37 in motif I is not required for ligase adenylation but plays a critical role downstream of step 1. These initial results suggested that the partial reactions of the RNA ligation pathway may be catalyzed by distinct constellations of active site residues.

To further delineate which of the conserved side chains of the nucleotidyl transferase motifs are functionally relevant, we have extended the mutational analysis of T4 Rnl2, focusing on residues in motifs I, III, IIIa, IV, and V (indicated by ● in Fig. 1). We identified individual amino acids that are required for strand joining, determined structure-activity relationships via conservative substitutions, and then stratified the mutational effects on the isolated steps of ligase adenylylation (step 1) and phosphodiester formation (step 3). We find that all five motifs are essential for covalent nucleotidyl transfer by Rnl2, as they are for DNA ligases and capping enzymes, and we locate essential structural “signatures” that are unique to Rnl2-like ligases.

EXPERIMENTAL PROCEDURES

Rnl2 Mutants—Amino acid substitution mutations were introduced into the rnl2 gene by PCR using the two-stage overlap extension method as described previously (13). The PCR products were digested with NdeI and BamHI and inserted into pET16b (Novagen). The inserts of the mutant pET-RNL2 plasmids were sequenced completely to exclude the acquisition of unwanted changes during amplification and cloning. pET-RNL2 plasmids were transformed into Escherichia coli BL21(DE3). Induction of Rnl2 production with isopropyl-1-thio-D-galactopyranoside and purification of Rnl2 from soluble bacterial extracts by nickel-agarose chromatography were performed as described previously (13). The wild-type and mutant Rnl2 preparations were stored at −80 °C.

Adenylyltransferase Assay—Reaction mixtures (20 µl) containing 50
and 20 mM EDTA, and the samples were analyzed by electrophoresis beled using T4 polynucleotide kinase and 
borate, 1 mM EDTA. The products were visualized by autoradiography of the wet gel and then eluted from an excised gel slice 
cated by 
of T4 Rnl2 from residues 1 to 227 is aligned to the sequences of the 5.5, 6.0, 6.5, or 7.0, or Tris-HCl, pH 7.0, 7.5, 8.0, or 8.5, 5 mM DTT, 2 mM 
labeling 18-mer single strand (data not shown). The broken loop of the 18-mer single strand substrate occurred with efficiency comparable with circu-
larization of an 18-mer single strand (Fig. 2, D). Thus, the physiological 
substrate for Rnl1 is a folded two-piece RNA molecule with the reactive ends held in proximity by the secondary structure of the RNA. It was of interest to determine whether Rnl2 was capable of joining such a substrate or whether its action was confined to single-stranded RNAs. We constructed a two-piece substrate designed to mimic the anticodon stem and broken anticodon loop of tRNA^lys (17) via hybridization of a 5'-32P-labeled 18-mer RNA to a complementary 15-mer RNA with 5'-OH and 3'-OH termini (Fig. 2). The broken loop of the synthetic two-piece substrate consists of a two-nucleotide single-
stranded tail on one strand of the stem duplex that provides the ligatable 3'-OH and a 5'-nucleotide single strand tail on the other strand that provides the ligatable 5'-PO_4 end (17). Reaction of this substrate with Rnl2 resulted in joining of the 5'-32P-labeled 18-mer to the unlabeled 15-mer to form a 33-mer ligation product (Fig. 2, RNA:pRNA). Thus, Rnl2 is capable of joining ends that are brought together by secondary structure. Enzyme titration experiments showed that ligation of the stem-
loop substrate occurred with efficiency comparable with circularization of an 18-mer single strand (data not shown).

The yield of the two-piece ligation product was optimal between pH 6.0 and 7.0. Reducing the pH to 5.5 or 5.0 suppressed strand joining while stimulating the formation of the RNA-adenylate intermediate (Fig. 2). Similar pH effects on AppRNA accumulation were noted previously for Rnl2-mediated circularization of a single-stranded RNA substrate (13). We conclude that the step of phosphodiester bond formation becomes rate-limiting at mildly acidic pH, independent of the type of RNA substrate used. Raising the pH to 7.5 suppressed the

![Fig. 2. Ligation of a two-piece stem-loop substrate.](image)

**RESULTS**

Two-piece Strand Joining by Rnl2—The strand joining activity of T4 Rnl2 was demonstrated initially with a 5'-32P-labeled single-stranded 18-mer RNA substrate. Cyclization of the 18-mer RNA was the predominant outcome of the ligation reaction rather than formation of linear multimers (13). The same preference for cyclization is displayed by T4 Rnl1 with substrates of similar size (3) and is construed to reflect proximity of the intramolecular 3'-OH terminus to the active site. The function of Rnl1 in vivo is to repair a break in the anticodon loop of E. coli tRNA^{lys} triggered by phage activation of a host-encoded anticodon nuclease (17, 18). Thus, the physiological substrate for Rnl1 is a folded two-piece RNA molecule with the reactive ends held in proximity by the secondary structure of the RNA. It was of interest to determine whether Rnl2 was capable of joining such a substrate or whether its action was confined to single-stranded RNAs. We constructed a two-piece substrate designed to mimic the anticodon stem and broken anticodon loop of tRNA^{lys} (17) via hybridization of a 5'-32P-labeled 18-mer RNA to a complementary 15-mer RNA with 5'-OH and 3'-OH termini (Fig. 2). The broken loop of the synthetic two-piece substrate consists of a two-nucleotide single-stranded tail on one strand of the stem duplex that provides the ligatable 3'-OH and a 5'-nucleotide single strand tail on the other strand that provides the ligatable 5'-PO_4 end (17). Reaction of this substrate with Rnl2 resulted in joining of the 5'-32P-labeled 18-mer to the unlabeled 15-mer to form a 33-mer ligation product (Fig. 2, RNA:pRNA). Thus, Rnl2 is capable of joining ends that are brought together by secondary structure. Enzyme titration experiments showed that ligation of the stem-loop substrate occurred with efficiency comparable with circularization of an 18-mer single strand (data not shown).

The yield of the two-piece ligation product was optimal between pH 6.0 and 7.0. Reducing the pH to 5.5 or 5.0 suppressed strand joining while stimulating the formation of the RNA-adenylate intermediate (Fig. 2). Similar pH effects on AppRNA accumulation were noted previously for Rnl2-mediated circularization of a single-stranded RNA substrate (13). We conclude that the step of phosphodiester bond formation becomes rate-limiting at mildly acidic pH, independent of the type of RNA substrate used. Raising the pH to 7.5 suppressed the
two-piece ligation reaction without trapping the RNA-adenylate intermediate (Fig. 2).

We reported previously that the inclusion of 1 mM ATP in the strand joining reaction with a single-stranded RNA substrate promoted accumulation of AppRNA and suppressed formation of ligated circles (13). The explanation offered for the ATP effect is that Rnl2, like Rnl1 (3), is prone to dissociate from the newly formed RNA-adenylate product of step 2 and that an immediate reaction of Rnl2 with ATP to form ligase-adenylate precludes it from rebinding to the RNA-adenylate for subsequent catalysis of strand joining. The experiment in Fig. 3 shows a smooth transition from circular product to AppRNA as the ATP concentration was increased, with a midpoint at ~10 μM ATP. The ATP concentration dependence of the trapping of RNA-adenylate roughly parallels the ATP concentration dependence of ligase-adenylate formation, which is saturated at 20 μM ATP (13). We considered the possibility that Rnl2 might be less prone to dissociate from the two-piece stem-loop substrate in which the 3’-OH end is tethered in the vicinity of the 5’ end, but found that inclusion of ATP in the two-piece ligation reaction had the same effect of trapping RNA-adenylate as it did with the single-stranded RNA substrate (not shown). Thus, we surmise that Rnl2 is generally liable to dissociate from the step 2 product under the reaction conditions employed.

**Phosphodiester Formation at a Preadenylated RNA 5’ End**—The production of high levels of RNA-adenylate by Rnl2 in the presence of ATP allowed us to synthesize and gel purify a preadenylated RNA substrate (AppRNA) for analysis of step 3 of the ligation pathway in isolation. Formation of a phosphodiester at the activated 5’ end was manifest by the appearance of a sealed circular RNA product, the yield of which was proportional to the amount of input Rnl2 (Fig. 4A). More than 90% of the substrate was converted to circular RNA at saturating Rnl2 concentrations. Circularization of RNA-adenylate required a divalent cation cofactor (data not shown). A small fraction of the input AppRNA was apparently deadenylated during the reaction to yield pRNA, which migrated between RNA-adenylate and the ligated circle (Fig. 4A). Deadenylation is the reverse of step 2 of the ligation pathway. Circularization of AppRNA was optimal at pH 6.0 to 7.0 and fell off sharply at a more acidic pH; activity declined gradually at alkaline pH such that the yield of circular product at pH 8.5 was ~25% of the value at pH 7.0 (data not shown).

We also prepared a preadenylated version of the two-piece stem-loop substrate. Rnl2 catalyzed efficient joining of the activated 5’ end of the labeled 18-mer to the 3’-OH of the unlabeled 5’ end of the labeled 18-mer to the 3’-OH of the unlabeled 15-mer RNA to form a circular RNA product, the yield of which was proportional to the amount of input Rnl2 (Fig. 4B). More than 90% of the substrate was converted to circular RNA at saturating Rnl2 concentrations. Circularization of RNA-adenylate required a divalent cation cofactor (data not shown). A small fraction of the input AppRNA was apparently deadenylated during the reaction to yield pRNA, which migrated between RNA-adenylate and the ligated circle (Fig. 4A). Deadenylation is the reverse of step 2 of the ligation pathway. Circularization of AppRNA was optimal at pH 6.0 to 7.0 and fell off sharply at a more acidic pH; activity declined gradually at alkaline pH such that the yield of circular product at pH 8.5 was ~25% of the value at pH 7.0 (data not shown).

We also prepared a preadenylated version of the two-piece stem-loop substrate. Rnl2 catalyzed efficient joining of the activated 5’ end of the labeled 18-mer to the 3’-OH of the unlabeled 15-mer RNA to form a circular RNA product, the yield of which was proportional to the amount of input Rnl2 (Fig. 4A). More than 90% of the substrate was converted to circular RNA at saturating Rnl2 concentrations. Circularization of RNA-adenylate required a divalent cation cofactor (data not shown). A small fraction of the input AppRNA was apparently deadenylated during the reaction to yield pRNA, which migrated between RNA-adenylate and the ligated circle (Fig. 4A). Deadenylation is the reverse of step 2 of the ligation pathway. Circularization of AppRNA was optimal at pH 6.0 to 7.0 and fell off sharply at a more acidic pH; activity declined gradually at alkaline pH such that the yield of circular product at pH 8.5 was ~25% of the value at pH 7.0 (data not shown).

We also prepared a preadenylated version of the two-piece stem-loop substrate. Rnl2 catalyzed efficient joining of the activated 5’ end of the labeled 18-mer to the 3’-OH of the unlabeled 15-mer RNA to form a circular RNA product, the yield of which was proportional to the amount of input Rnl2 (Fig. 4A). More than 90% of the substrate was converted to circular RNA at saturating Rnl2 concentrations. Circularization of RNA-adenylate required a divalent cation cofactor (data not shown). A small fraction of the input AppRNA was apparently deadenylated during the reaction to yield pRNA, which migrated between RNA-adenylate and the ligated circle (Fig. 4A). Deadenylation is the reverse of step 2 of the ligation pathway. Circularization of AppRNA was optimal at pH 6.0 to 7.0 and fell off sharply at a more acidic pH; activity declined gradually at alkaline pH such that the yield of circular product at pH 8.5 was ~25% of the value at pH 7.0 (data not shown).
beled 15-mer to yield a 33-mer product (Fig. 4B). Again, a small fraction of the substrate was deadenylated to produce pRNA. Also, a trace amount of circular product was formed, most likely reflecting the presence of a small fraction of unhybridized 18-mer AppRNA strand in the substrate preparation.

Lys35 is the site of covalent attachment of AMP to Rnl2; its replacement by alanine abolishes ligase-AMP formation and, perforce, the formation of the RNA-adenylate intermediate (13). An instructive finding was that the K35A mutant of Rnl2 was capable of sealing a preadenylated two-piece stem-loop RNA substrate, albeit less efficiently than did wild-type Rnl2 (Fig. 4B). As expected, the K35A mutant was unable to deadenylate the AppRNA substrate to form the pRNA species seen in the wild-type Rnl2 reaction. This result underscores that the lysine nucleophile is not strictly essential for the chemical step of phosphodiester bond formation. Similar results concerning the ability of motif I lysine mutants to catalyze phosphodiester bond formation have been reported for ATP-dependent DNA ligases (19-21).

New Alanine-scanning Mutagenesis—The five putative nucleotidyl transferase motifs of Rnl2 and related ligases are highlighted in Fig. 1. Motif I, which contains the active site lysine, adheres to the consensus sequence EKXH/DGXXN, which differs from the XKKDGXR sequence characteristic of ATP-dependent DNA ligases and most mRNA capping enzymes (22). To evaluate the functional relevance of the signature Glu24 and Asn40 side chains, we substituted each with alanine. Alanine changes were also introduced at five other conserved positions: Arg55, located between motifs I and III; Glu99 in motif III; Phe119 and Asp120 in motif IIIa; and Lys209 flanking motif IV. A single nonconserved residue, Lys189, was also changed to alanine. The E34A, N40A, R55A, E99A, F119A, D120A, K189A, and K209A mutants were produced in E. coli as His6-tagged fusions and purified from soluble bacterial extracts by nickel-agarose chromatography (Fig. 5A). The 42-kDa Rnl2 polypeptide was the predominant species detected by SDS-PAGE, and the extents of purification were comparable to those of the protein preparation likely comprises preadenylated Rnl2, apparently caused by the loss of the acidic carboxylate moiety (see below).

The adenylyltransferase activity of recombinant Rnl2 was assessed by label transfer from [γ-32P]ATP to the Rnl2 polypeptide to form the covalent enzyme-adenylate intermediate. The extent of ligase-adenylate formation by wild-type Rnl2 was proportional to input protein (Fig. 5B). We estimated from the slope of the titration curve that 65% of the input enzyme molecules became labeled with [32P]AMP. The residual fraction of the protein preparation likely comprises preadenylated Rnl2 (see below). The E34A, N40A, F119A, D120A, and E204A mutants were effectively inert over the same range of input enzyme, i.e. their specific activities were ≈1% of the wild-type value. Other mutants displaying significant defects in step 1 adenylate formation were R55A (2.3%), K209A (2.7%), K227A (4.1%), and N40A (15%). Only the K198A mutant displayed weak activity in the electrophoretic mobility of Rnl2, apparently caused by the loss of the acidic carboxylate moiety (see below).

The adenylyltransferase activity of recombinant Rnl2 was assayed by label transfer from [γ-32P]ATP to the Rnl2 polypeptide to form the covalent enzyme-adenylate intermediate. The extent of ligase-adenylate formation by wild-type Rnl2 was proportional to input protein (Fig. 5B). We estimated from the slope of the titration curve that 65% of the input enzyme molecules became labeled with [32P]AMP. The residual fraction of the protein preparation likely comprises preadenylated Rnl2 (see below). The E34A, N40A, F119A, D120A, and E204A mutants were effectively inert over the same range of input enzyme, i.e. their specific activities were ≈1% of the wild-type value. Other mutants displaying significant defects in step 1 adenylate formation were R55A (2.3%), K209A (2.7%), K227A (4.1%), and N40A (15%). Only the K198A mutant displayed near wild-type adenylyltransferase activity (Fig. 5B and Table 1).

Mutational effects on ligation of a 5'-32P-labeled single-stranded RNA substrate in the presence and absence of ATP under conditions of enzyme excess are shown in Fig. 6. Wild-type Rnl2 efficiently circularized the RNA in the absence of ATP (reflecting the catalysis of steps 2 and 3 by preformed Rnl2-AMP in the enzyme preparation) but generated predominantly AppRNA in the presence of ATP. The ligation activity of the K198A mutant was indistinguishable from that of wild-type Rnl2. Mutants E34A, E99A, and D120A failed to form any ligated RNA circles or RNA-adenylate intermediate. R55A and

![Fig. 5. Purification and adenylyltransferase activity of Rnl2-Ala mutants. A, aliquots (3 μg) of the nickel-agarose preparations (0.1 M imidazole eluates) of wild-type (WT) Rnl2 and the indicated mutants were analyzed by SDS-PAGE. The Coomassie Blue-stained gel is shown. B, reaction mixtures contained 50 mM Tris-acetate (pH 6.5), 5 mM DTT, 1 mM MgCl2, 20 μM [γ-32P]ATP, and wild-type or mutant Rnl2 as specified. Ligate-adenylate formation is plotted as a function of input Rnl2 protein.](image-url)
ligases and most mRNA capping enzymes. Twenty new Rnl2 mutants were produced in *E. coli* and purified from soluble bacterial extracts by nickel-agarose chromatography. The Rnl2 polypeptide was the predominant species detected by SDS-PAGE, and the extents of purification were comparable for mutant and wild-type Rnl2 (Fig. 7). The D120N mutation caused the same increase in the electrophoretic mobility seen with D120A, which was rectified when Asp120 was replaced by Glu (Fig. 7, middle panel). The K227R mutation also caused an increase in electrophoretic mobility, which was not seen with K227Q (Fig. 7, bottom panel) or K227A (13).

The adenylyltransferase activities of the mutants were determined by enzyme titration and normalized to the wild-type value (Table I). Conservative replacement of Glu34, Glu99, or Glu204 with Asp or Gln elicited severe catalytic defects comparable with those seen with the respective Ala mutants. These data establish the requirement for a carboxylate residue at positions 34, 99, and 204 and a minimum distance from the main chain to the carboxylate that is met by glutamate but not aspartate. Note that glutamates are strictly conserved at these three positions in the trypanosome, baculovirus, and entomopoxvirus Rnl2 homologs (Fig. 1). Changing Asp120 to Asn reduced ligase-AMP formation to 0.2% of the wild-type level.

### Table I

**Mutational effects on Rnl2 adenylyltransferase activity**

The recombinant Rnl2 preparations were titrated for adenylyltransferase activity as described in the legend to Fig. 5B. The specific activities were calculated from the slopes of the titration curves and normalized to the specific activity of wild-type Rnl2 (defined as 100%). The right-hand column lists the side chain present at the equivalent position of DNA ligases and the relevant side chain contacts seen in the available crystal structures of DNA ligases and RNA capping enzymes. Contacts to the 5'-PO₄ are inferred from the crystal structure of *Chlorella* virus DNA ligase-adenylate, which contains a sulfate bound to the enzyme surface adjacent to the AMP phosphate. This sulfate is proposed to mimic the 5'-PO₄ of the DNA nick.

| Motif | Mutation | Specific activity | DNA ligase amino acids/contacts | % of wild type |
|-------|----------|------------------|--------------------------------|---------------|
| I     | E34A     | 0.4              |                                |               |
|       | E34D     | 0.2              |                                |               |
|       | E34Q     | 2.1              |                                |               |
|       | N40A     | 15               | Arg                            |               |
|       | N40D     | 99               | Ribose O                       |               |
|       | N40Q     | 21               |                                |               |
|       | N40R     | 0.2              |                                |               |
|       | R55A     | 2.3              | Arg                            |               |
|       | R55K     | 30               | γ-PO₄ and 5'-PO₄               |               |
|       | R55Q     | 1.3              |                                |               |
| III   | E99A     | 0.8              | Glu                            |               |
|       | E99D     | 1.3              | Ribose O                       |               |
|       | E99Q     | 0.2              |                                |               |
| IIIa  | F119A    | <0.1             | Phe                            |               |
|       | F119L    | 3.9              | Adenine                        |               |
|       | D120A    | <0.1             | Asp                            |               |
|       | D120E    | 12               | Salt bridge                    |               |
|       | D120N    | 0.2              |                                |               |
|       | K189A    | 7.0              |                                |               |
| IV    | E204A    | 0.2              | Glu                            |               |
|       | E204D    | 2.2              | Metal                          |               |
|       | E204Q    | <0.1             |                                |               |
|       | K209A    | 2.7              | Arg/Lys                        |               |
|       | K209R    | 25               | Salt bridge                    |               |
|       | K209Q    | 7.3              |                                |               |
| V     | K225A    | 0.2              | Lys                            |               |
|       | K225R    | 2.2              | α-PO₄ and 5'-PO₄               |               |
|       | K225Q    | <0.1             |                                |               |
|       | K227A    | 4.1              | Lys                            |               |
|       | K227R    | <0.1             | α-PO₄                          |               |
|       | K227Q    | 3.9              |                                |               |

**Fig. 6.** Effects of alanine mutations on RNA strand joining. Reaction mixtures contained 50 mM Tris acetate (pH 6.5), 5 mM DTT, 1 mM MgCl₂, 1 pmol of pRNA substrate, 160 ng (3.8 pmol) of wild-type (WT) Rnl2 or mutant Rnl2, and either 1 mM ATP (+) or no ATP (-). The products were resolved by PAGE and visualized by autoradiography. A control reaction lacking Rnl2 is shown in lane C. The step 2 and 3 reactions of the ligation pathway are illustrated schematically at the bottom left.

**Fig. 7.** Purification of conservative mutants of Rnl2. Aliquots (3 µg) of the nickel-agarose preparations (0.1 M imidazole eluates) of wild-type (WT) Rnl2 and the indicated mutants were analyzed by SDS-PAGE. The Coomassie Blue-stained gels are shown. The positions and sizes (kDa) of marker polypeptides are indicated on the left.
similar to the D120A mutant, but activity was partially restored by the glutamate substitution (to 12% of wild type). Thus, the carboxylate functional group is critical at position 120, but there is some flexibility in accommodating the longer Glu side chain instead of Asp. This is notable in light of the fact that Glu is present at the corresponding motif IIIa position of the baculovirus and entomopoxvirus Rnl2 homologs (Fig. 1).

Whereas replacement of Arg209 by Lys resulted in a significant gain of function (to 30% of wild type) compared with R55A (2.3%), the R55Q change had no salutary effect. Similarly, the gain of function (to 30% of wild type) compared with R55A and R55K and K227Q, respectively (Table I). WT, E119L, D120E, and E204A were inert in the isolated step of ligation adenylation (Fig. 8). Thus, although F119L was defective in ligase adenylation in vitro, it obviously did react with ATP to form ligase-adenylate in vivo in E. coli. Adenylation in vivo of D120E, K209R, and K209Q is in keeping with their partial adenylyltransferase activities in vitro (between 7 and 25% of wild-type Rnl2). The accumulation of AppRNA seen in Fig. 8 suggested that F119L, D120E, K209R, or K209Q might be defective in step 3 of the ligation pathway.

Mutational Effects on Phosphodiester Formation at a Preadenylated RNA 5’ End—Although the effects on step 1 sufficed to explain the loss of overall ligation activity of many of our Rnl2 mutants, we were interested in determining whether the enzyme functional groups implicated for step 1 are also required for phosphodiester formation (step 3). We showed above (Fig. 4) that step 3 can be studied in isolation by bypassing the requirement for steps 1 and 2 and gauging the ability of Rnl2 to seal a preadenylated RNA substrate. Thus, the 13 Rnl2-Ala mutants in our collection were reacted with a single-stranded 18-mer labeled RNA, compared with 90% circles in the wild-type Rnl2 (Fig. 7). Although F119L had a slight gain of function at position 3, it was completely inert in the isolated step of phosphodiester formation (Fig. 8). E204D, K209R, and K209Q were inert in the isolated step of phosphodiester formation, just as they were virtually inert in the isolated step 1 ligase adenylation reaction. R55A and K209A, which were 2–3% as active as wild-type Rnl2 in step 1, formed only trace amounts of circular product in the isolated

**Fig. 8.** Effects of conservative mutations on RNA strand joining. Reaction mixtures contained 50 mM Tris acetate (pH 6.5), 5 mM DTT, 1 mM MgCl₂, 1 pmol of pRNA substrate, 160 ng (3.8 pmol) of wild-type (WT) or mutant Rnl2, and either 1 mM ATP (+) or no ATP (−). The products were resolved by PAGE and visualized by autoradiography.
step 3 reaction (Fig. 9). A simple interpretation of these results is that Glu34, Arg26, Glu99, Phe119, Asp120, Glu204, and Lys209 play equivalent roles in the first and third steps of the ligation reaction. The obvious common feature of the step 1 and 3 reactions is that they entail recognition of the adenylate moiety of the ATP and AppRNA substrates, respectively.

Nonetheless, as discussed above for the K35A mutation, it is not a forgone conclusion that any change that abrogates or severely affects step 1 will exert the same effect on step 3. For example, K225A is just as defective in step 1 as E34A or E204A (Table I), but it clearly can catalyze ligation of AppRNA, albeit less efficiently than wild-type Rnl2 (Fig. 9). K227A is more active in step 3 than K209A (Fig. 9), although they have comparable activity in step 1 (3–4% of wild type).

Are any of the putative active site constituents of Rnl2 specific for the step of phosphodiester bond formation? We showed previously that the motif I mutant H37A was fully active in step 1 ligase adenyllylation and capable of transferring the adenylate to the 5'-PO4 of RNA to form RNA-adenylate but seemingly unable to form RNA circles during the pRNA ligation reaction (13). Here we found that H37A was severely and selectively impaired at the isolated step of phosphodiester bond formation (Fig. 9). Thus, His37 is specifically implicated in catalysis of step 3. Replacing His37 by Asp rectified the pRNA strand joining defect of the H37A mutant (13). Here we see that the H37D change restored wild-type activity in the isolated step 3 reaction (Fig. 9). Note that Asp is normally present at the equivalent motif I position of the entomopoxvirus and baculovirus Rnl2 homologs (Fig. 1) and in motif I of the majority of known DNA ligase enzymes. The signature Asn residue of the Rnl2 family may also play a key role in step 3, insofar as the N40A mutant appears to be more defective in step 3 (Fig. 9) than it does in step 1 (15% of wild-type adenyltransferase activity).

Additional insight into the requirements for phosphodiester bond formation were gleaned from an analysis of the effects of conservative mutations on the isolated step 3 reaction (Fig. 10). E34D, E34Q, E99D, E99Q, E204D, and E204Q were inert or severely defective in sealing AppRNA, just like the respective Glu-to-Ala mutants. The N40D and N40Q changes restored step 3 function compared with the defective N40A mutant, whereas N40R was inactive in step 3. The R55K change revived phosphodiester bond formation relative to R55A, whereas the R55Q mutant remained defective. The K209R and K209Q mutations both resulted in gains of step 3 activity relative to the severely defective K209A protein.

F119L displayed feeble step 3 activity, which was nonetheless an improvement compared with F119A. The D120E change partially restored step 3 function relative to the D120A mutant, but the D120N mutant remained defective. The isolated step 3 defects of the F119L and D120E proteins may account for their accumulation of RNA-adenylate during the composite ATP-independent ligation reaction (Fig. 8).

In motif V, the K225R protein was partially active in step 3 (like K225A), whereas K225Q was severely defective. At position 227, however, the arginine substitution abolished step 3 activity, whereas the glutamine mutant displayed nearly wild-type step 3 activity. The disparate effects of conservative substitutions indicate that the two motif V lysines play distinct roles at the step of phosphodiester bond formation, i.e. the function of Lys225 depends on its positive charge, whereas Lys227 function likely depends on its hydrogen bonding capacity.

**DISCUSSION**

We have now identified 12 essential residues of bacteriophage T4 Rnl2 and defined structure-activity relationships via conservative substitutions. We find that all five nucleotidyl transferase motifs are essential in Rnl2, just as they are in DNA ligases and RNA capping enzymes (23–30). These results underscore a shared structural basis for catalysis among Rnl2-like ligases, DNA ligases, and capping enzymes. They also highlight two essential signature residues in motif I (Glu34 and Asn40) that are unique to Rnl2-like ligases.

We report mutational effects on overall RNA ligation and the isolated steps of ligase adenyllylation and phosphodiester formation. Our inferences about which features of the individual side chains are required for activity have been discussed in detail above. Thus, we focus here on how the findings may be interpreted in light of the crystal structures available for other members of the covalent nucleotidyl transferase superfamily. In the absence of an atomic structure for any RNA ligase, we assume based on concordance of mutational data that the fold of the N-terminal nucleotidyl transferase domain of Rnl2 (which includes the five motifs) resembles that of Chlorella virus DNA ligase (the minimal eukaryotic ATP-dependent DNA ligase) and that there is a direct correspondence between the essential amino acids of the Rnl2-like RNA ligase family and the amino acids found at "equivalent" positions of DNA ligases, which are listed in Table I along with the atomic contacts made by these side chains, as revealed by the struc-
tures of the ligase-AMP intermediate of Chlorella virus DNA ligase and the ATP-bound bacteriophage T7 DNA ligase (9, 11). Note that eight of the ten essential Rnl2 positions listed in Table I are occupied by an identical or closely related amino acid in ATP-dependent DNA ligases. The two exceptions are the signature motif I residues of Rnl2, Glu<sup>40</sup> and Asn<sup>40</sup>.

Rnl2 residues Arg<sup>206</sup>, Glu<sup>209</sup>, Phe<sup>119</sup>, and Glu<sup>204</sup> are critical for the first and third steps of the RNA ligation reaction, which, by analogy to DNA ligase, will entail docking of the adenylate moiety of the ATP or the AppRNA substrate into an AMP-binding pocket of Rnl2. The counterparts of Rnl2 residues Arg<sup>206</sup>, Glu<sup>209</sup>, Phe<sup>119</sup>, and Glu<sup>204</sup> in the DNA ligases make direct contacts with constituents of the adenosine nucleotide.

The motif III glutamate contacts the ribose sugar; the motif IIIa aromatic group engages in a π stack on the adenine base; the motif IV glutamate is implicated in coordinating a divalent cation; the two motif V lysines coordinate the α-phosphate; and the conserved arginine located between motifs I and III is proposed to contact the γ-phosphate of ATP in step 1 and the 5′-PO<sub>4</sub> of the nucleic acid substrate during subsequent steps. We impute similar functions to the respective essential side chains of Rnl2.

The counterparts of the essential Rnl2 residues Asp<sup>120</sup> (motif IIIa) and Lys<sup>209</sup> (motif IV) do not contact the nucleotide in any of the available DNA ligase (or capping enzyme) structures. Rather they form ion pairs with oppositely charged side chains of the respective enzymes. The motif IIIa Asp is located next to the essential aromatic amino acid that forms a hydrophobic pocket for the purine base. This position is occupied by aspartate in the majority of ATP-dependent DNA ligases and RNA capping enzymes. The participation of the aspartate in a salt bridge may be crucial for the active site architecture of Rnl2, insofar as the neutral asparagine substitution elicited defects in steps 1 and 3 of the Rnl2 reaction. It is remarkable that the positively charged partner in the ion pair formed by the conserved motif IIIa Asp is located at very different places in the amino acid sequences of the several ligases and capping enzymes for which atomic structures have been solved.

In the Chlorella virus DNA ligase structure, the motif IIIa Asp forms an ion pair with the arginine-flanking motif IV, corresponding to the essential Lys<sup>209</sup> residue of Rnl2. The significant gain of function seen for the K209R mutant of Rnl2 is in keeping with a putative ionic interaction for Lys<sup>209</sup> be it with Asp<sup>120</sup> (analogous to the ion pair of Chlorella virus DNA ligase) or with some other acidic residue in Rnl2. Asn<sup>40</sup> is one of the signature amino acids of the Rnl2-like family. An essential arginine side is present at the equivalent position in motif I of ATP-dependent DNA ligases and capping enzymes; this motif I arginine interacts with the ribose sugar of the adenosine or guanosine nucleoside. Asn<sup>40</sup> of Rnl2 may also be concerned with recognition of the adenosine sugar of ATP or AppRNA, albeit functioning as a hydrogen bond acceptor. Alternatively, it may play different role in substrate binding or catalysis. In any event, it is clear that Rnl2 has evolved a unique structural requirement for this side chain, which cannot be fulfilled (but is instead antagonized) by the arginine normally present in other covalent nucleotidyl transferases.

Glu<sup>40</sup> occupies the position immediately preceding the active site lysine nucleophile. Whereas this residue is invariant in Rnl2-like proteins, the equivalent side chain in DNA ligases is typically hydrophobic. It is worth noting that T4 Rnl1 has no counterpart of Glu<sup>40</sup> in motif I: rather there is a threonine immediately preceding the motif I lysine. Thus, Glu<sup>40</sup> stands out as an essential and defining feature of the Rnl2 family (required for steps 1 and 3 of the Rnl2 ligation pathway). Delineation of the atomic contacts made by this Rnl2-specific side chain will obviously hinge on crystallization of Rnl2 or one of its orthologs.

REFERENCES

1. Silber, R., Malathi, V. G., and Hurwitz, J. (1972) Proc. Natl. Acad. Sci. U. S. A. 69, 3099–3103.
2. Cranston, J. W., Silber, R., Malathi, V. G., and Hurwitz, J. (1974) J. Biol. Chem. 249, 7447–7456.
3. Sugino, A., Noppek, T. J., and Cozzarelli, N. R. (1978) J. Biol. Chem. 252, 1732–1738.
4. Uhlenbeck, O. C., and Gumport, R. I. (1982) Enzymes 15, 51–58.
5. Engler, M. J., and Richardson, C. C. (1982) Enzymes 15, 3–29.
6. Thogerson, H. C., Morris, H. R., Rand, K. N., and Gait, M. J. (1985) Eur. J. Biochem. 147, 325–339.
7. Hephby, S., Singh, M., and Gait, M. J. (1987) Biochemistry 26, 1688–1696.
8. Shuman, S., and Schwer, B. (1995) Mol. Microbiol. 17, 405–410.
9. Subramanya, H. S., Doherty, A. J., Ashford, S. R., and Wigley, D. B. (1996) Cell 85, 607–615.
10. Håkansson, K., Doherty, A. J., Shuman, S., and Wigley, D. B. (1997) Cell 89, 545–553.
11. Odell, M., Sriskanda, V., Shuman, S., and Nikolov, D. (2000) Mol. Cell 6, 1183–1193.
12. Lee, J. Y., Chang, C., Song, H. K., Moon, J., Yang, J., Kim, H. K., Kwon, S. T., and Suh, S. W. (2000) EMBO J. 19, 1119–1129.
13. Ho, C. K., and Shuman, S. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 12709–12714.
14. Schnauffer, A., Panigrahi, A. K., Panici, B., Igo, R. P., and Suh, S. W. (2000) EMBO J. 19, 1119–1129.
15. McManus, M. T., Shimamura, M., Graus, J., and Hajduk, S. L. (2001) RNA 7, 167–175.
16. Rusche, L. N., Huang, C. E., Piller, K. J., Hemann, M., Wirtz, E., and Solomon, A. (2001) Mol. Cell. Biol. 21, 979–989.
17. Amitsur, M., Levitz, R., and Kaufman, G. (1987) EMBO J. 6, 2499–2503.
18. Kaufman, G. (2000) Trends Biochem. Sci. 25, 70–74.
19. Sekiguchi, J., and Shuman, S. (1997) J. Virol. 71, 9679–9684.
20. Sriskanda, V., and Shuman, S. (1998) Nucleic Acids Res. 26, 4618–4625.
21. Sriskanda, V., Kelman, Z., Hurwitz, J., and Shuman, S. (2000) Nucleic Acids Res. 28, 2221–2228.
22. Shuman, S. (2001) Cold Spring Harbor Symp. Quant. Biol. 66, 301–312.
23. Shuman, S. (2000) Proc. Natl. Acad. Sci. Mol. Biol. 66, 1–40.
24. Shuman, S., and Ru, X. (1995) Virology 211, 73–83.
25. Sriskanda, V., and Shuman, S. (1998) Nucleic Acids Res. 26, 525–531.
26. Doherty, A. J., and Dafforn, T. R. (2000) J. Mol. Biol. 296, 43–56.
27. Sriskanda, V., and Shuman, S. (2002) Nucleic Acids Res. 30, 903–911.
28. Sriskanda, V., and Shuman, S. (2002) J. Biol. Chem. 277, 9661–9667.
29. Shuman, S., Liu, Y., and Schwer, B. (1994)Proc. Natl. Acad. Sci. U. S. A. 91, 12046–12050.
30. Wang, S. P., Deng, L., Ho, C. K., and Shuman, S. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 8573–8578.