Mutational Analysis of Bacteriophage T4 RNA Ligase 1

DIFFERENT FUNCTIONAL GROUPS ARE REQUIRED FOR THE NUCLEOTIDYL TRANSFER AND PHOSPHODIESTER BOND FORMATION STEPS OF THE LIGATION REACTION

Received for publication, April 24, 2003, and in revised form, May 20, 2003
Published, JBC Papers in Press, May 24, 2003, DOI 10.1074/jbc.M304320200

Li Kai Wang, C. Kiong Ho, Yi Pei, and Stewart Shuman†

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

T4 RNA ligase 1 (Rnl1) exemplifies an ATP-dependent RNA ligase family that includes fungal tRNA ligase (Trl1) and a putative baculovirus RNA ligase. Rnl1 acts via a covalent enzyme-AMP intermediate generated by attack of Lys-99 N\text{ε} on the \(\alpha\) phosphorus of ATP. Mutation of Lys-99 abolishes ligase activity. Here we tested the effects of alanine mutations at 19 conserved positions in Rnl1 and thereby identified 9 new residues essential for ligase activity: Arg-54, Lys-119, Glu-227, Gly-228, Lys-240, and Lys-242. Seven of the essential residues are located within counterparts of conserved nucleotidyltransferase motifs I (\(\text{KEDG}^{102}\)), Ia (\(\text{SK}^{118}\)), IV (\(\text{EGYA}^{231}\)), and V (\(\text{HFKIK}^{242}\)) that comprise the active sites of DNA ligases, RNA capping enzymes, and T4 RNA ligase 2. Three other essential residues, Arg-54, Lys-75 and Phe-77, are located upstream of the AMP attachment site within a conserved domain unique to the Rnl1-like ligase family. We infer a shared evolutionary history and active site architecture in Rnl1 (a tRNA repair enzyme) and Trl1 (a tRNA splicing enzyme). We determined structure-activity relationships via conservative substitutions and examined mutational effects on the isolated steps of Rnl1 adenyllylation (step 1) and phosphodiester bond formation (step 3). Lys-75, Lys-240, and Lys-242 were found to be essential for step 1 and overall ligation of 5′-phosphorylated RNA but not for phosphodiester bond formation. These results suggest that the composition of the Rnl1 active site is different during steps 1 and 3. Mutations at Arg-54 and Lys-119 abolished the overall RNA ligation reaction without affecting steps 1 and 3. Arg-54 and Lys-119 are thereby implicated as specific catalysts of the RNA adenylation reaction (step 2) of the ligation pathway.

RNA ligases join 3′ OH and 5′ PO₄ RNA termini via a series of three nucleotidyl transfer steps similar to those of DNA ligases: (i) RNA ligase reacts with ATP to form a covalent ligase-(lysyl-AMP)-AMP intermediate plus pyrophosphate; (ii) AMP is transferred from ligase-adenylate to the 5′ PO₄ RNA end to form an RNA-adenylate intermediate (AppRNA); and (iii) ligase catalyzes attack by an RNA 3′ OH on the RNA-adenylate to seal the two ends via a phosphodiester bond and release AMP (1–5). Bacteriophage T4 RNA ligase 1 (Rnl1) is the founding member of the RNA ligase family (1). The function of Rnl1 in vivo is to repair a break in the anticodon loop of *Escherichia coli* tRNA\(^{5\text{-}\text{ps}}\) triggered by phage activation of a host-encoded anticodon nuclease (6).

T4 Rnl1 is a 374-amino acid polypeptide (7). Gait and coworkers (8, 9) mapped the site of covalent adenylation to Lys-99 and demonstrated the essential role of Lys-99 in catalysis by site-directed mutagenesis. The active site lysine of Rnl1 is located within a conserved sequence element (KXE/D/N/G; motif I in Fig. 1) that defines a superfamily of covalent nucleotidyltransferases, which includes DNA ligases and mRNA capping enzymes (10, 11). DNA ligases and capping enzymes have a common tertiary structure composed of five conserved motifs (I, III, IIIa, IV, and V) that contain amino acid functional groups responsible for nucleotide binding and catalysis (12–16). It has been suggested that DNA ligases and capping enzymes evolved from a common ancestral nucleotidyltransferase (11), possibly from an ancient RNA strand-joining enzyme like Rnl1, in which case the fold and active site of Rnl1 should resemble that of DNA ligases and RNA capping enzymes. This model remains speculative because there is no atomic structure available for Rnl1, and no mutational analysis has been conducted outside of the KXE/D/N/G motif.

We recently identified and characterized a second T4 RNA ligase (Rnl2) encoded by bacteriophage T4 gene 24.1 (17). Rnl2 exemplifies a novel family of RNA ligases that contain all five motifs found in DNA ligases and capping enzymes, plus a set of unique and essential structural “signatures” (17, 18). The Rnl2-like ligase family includes the RNA-editing ligases of *Trypanosoma* and *Leishmania* and a group of putative RNA ligases encoded by eukaryotic viruses and many species of Archaea. Thus, the Rnl2-like ligases are present in all three phylogenetic domains.

In contrast, T4 Rnl1 exemplifies a distinct subfamily of RNA ligases with a narrow phylogenetic distribution. The Rnl1-like proteins include a putative RNA ligase/polyribonucleotide kinase of *Autographa californica* nuclear polyhedrosis virus (AcNPV,\(^1\) a baculovirus) and the tRNA ligases of fungi (19–22). The primary structures of the Rnl1-like and Rnl2-like ligases are quite dissimilar (no cross-recognition by multiround PSI-BLAST searches), with the only shared features being the presence of three of the nucleotidyltransferase motifs (I, IV, and V) that comprise the NMP-binding site of DNA ligases and RNA capping enzymes (Fig. 1). Alignment of the primary structures of the Rnl1-like ligases highlights three notable features as follows: (i) an absence of obvious counterparts of nucleotidyltransferase motifs III and IIIa shared among DNA ligases, capping enzymes, and Rnl2-like ligases; (ii) the presence of clusters of conserved residues

\(^*\) This work was supported by National Institutes of Health Grant GM42498. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 just to indicate this fact.

\(‡\) To whom correspondence should be addressed. E-mail: s-shuman@ski.mskcc.org.

\(^1\) The abbreviation used is: AcNPV, *A. californica* nuclear polyhedrosis virus.
upstream of the motif I lysine nucleophile that are not found in Rnl2-like ligases or in DNA ligases and capping enzymes; and (iii) sequence similarity between Rnl1 and the AcNPV ligase in the segment separating motifs I and IV, which is not found in the tRNA ligases of the Rnl1-like group (Fig. 1). In order to delineate which of the conserved side chains are functionally relevant, we have performed a mutational analysis of 20 positions of T4 Rnl1 (indicated by ● in Fig. 1), focusing on motifs I, IV, and V, the segment proximal to motif I, and the region separating motifs I and IV. We find that motifs IV and V are essential for covalent nucleotidyl transfer by Rnl1. We also identify essential residues outside the nucleotidyltransferase motifs, suggesting that Rnl1-like enzymes have distinct signature components of their active sites.

EXPERIMENTAL PROCEDURES

Recombinant T4 Rnl1—A DNA fragment containing the g63 open reading frame (hereafter RNL1) was amplified by PCR from T4 genomic DNA (a gift of Ken Kreuzer, Duke University) with oligonucleotide primers designed to introduce an NdeI restriction site at the translation start codon and a BamHI site 3′ of the stop codon. The PCR product was digested with NdeI and BamHI and inserted into the pet16b (Novagen) vector to express plasmid pET-RNL1. Amino acid substitution mutations were introduced into the RNL1 gene by PCR using the previously (18)), and Rnl1 as specified were incubated for 30 min at 37°C. The reactions were quenched by adding 5 μl of 95% formamide, 20 mM EDTA. The samples were analyzed by electrophoresis through a 18% polyacrylamide gel containing 7 M urea in 0.5× TBE (45 mM Tris borate, 1 mM EDTA). The ligation products were visualized by autoradiography of the gel and quantitated with a PhosphorImager.

Sealing of Preformed RNA-Adenylate—Reaction mixtures (10 μl) containing 50 mM Tris-HCl (pH 8.0), 2 mM dithiothreitol, 10 mM MgCl2, 20 μM [α-32P]ATP. The labeled 18-mer was purified by electrophoresis through a 18% polyacrylamide gel containing 7 M urea in 0.5× TBE (45 mM Tris borate, 1 mM EDTA). The ligation products were visualized by autoradiography of the gel and quantitated with a PhosphorImager.

RESULTS

Recombinant T4 Rnl1—Wild-type Rnl1 was produced in E. coli as a His_{6}^-tagged fusion protein, and the 45-kDa recombinant protein was purified from the crude soluble bacterial extract by adsorption to Ni-agarose and elution with buffer containing imidazole (Fig. 2, W7). The adenyllyltransferase activity of recombinant Rnl1 (step 1 of the ligation pathway) was evinced by label transfer from 20 μM [α-32P]ATP to the Rnl1 polypeptide to form a covalent enzyme-adenylate adduct (not shown). The reaction was magnesium-dependent (optimal at 1–5 mM MgCl2), and the yield of Rnl1-[32P]AMP was optimal at pH 8.0 (not shown). We calculated that ~70% of the input Rnl1 protein was adenylated with [32P]AMP. The remaining ~30% of the Rnl1 preparation may consist of pre-formed Rnl1-AMP intermediate.

Reaction of recombinant Rnl1 with a 5′-[32P] labeled 18-mer RNA oligonucleotide (rRNA) and magnesium in the presence of

---

**Fig. 1. Rnl1-like family of RNA ligases.** The amino acid sequence of T4 Rnl1 from residues 51–250 is aligned to the sequence of the putative RNA ligase of the AcNPV baculovirus and two homologous segments of the tRNA ligase (Tyr) enzymes of Saccharomyces cerevisiae (Sc) and Schizosaccharomyces pombe (Sp). Nucleotidyltransferase motifs I, IV, and V are underlined. Positions of Rnl1 that were subjected to mutational analysis are indicated by ●. Residues found to be essential for RNA ligase activity are highlighted in shaded boxes.
20 μM ATP resulted in the formation of two new radiolabeled RNAs (Fig. 3, lane WT). The major product, migrating −1.5 nucleotide steps faster than the input 18-mer pRNA strand, corresponds to a covalently closed 18-mer circle formed by intramolecular ligation of the 5′-PO₄ and 3′-OH termini of the substrate strand (1, 17, 18). A minor labeled product, migrating −1 nucleotide step slower than the input 18-mer, corresponded to the RNA-adenylate (AppRNA) generated by AMP transfer from Rnl1-AMP to the 5′ end of the input 18-mer RNA (3, 17, 18).

Mutational Analysis of Rnl1 Motifs I, IV, and V Identifies Residues Essential for Catalysis—The presence of putative counterparts of nucleotidyltransferase motifs I, IV, and V in Rnl1 raises the question of whether and how they contribute to RNA ligase function. Heaphy et al. (9) showed previously that mutating Lys-99 of motif I to asparagine abolished overall ligation and the ability to form the Rnl1-AMP intermediate, as expected from the assignment of Lys-99 as the site of covalent adenylylation (8). Changing Glu-100 to glutamine or threonine was well tolerated, which is consistent with the lack of side chain conservation at this position (Fig. 1). Several different mutations of Asp-101 (to asparagine, serine, or glutamate) abolished the composite ligation reaction but had no effect on formation of the Rnl1-AMP intermediate (9). The lack of a requirement of the aspartate residue in motif I for the first nucleotidyl transfer step is a property shared with DNA ligase (14, 23–26), mRNA capping enzyme (27, 28), and Rnl2 (17, 18). Yet the finding that Rnl1 cannot tolerate asparagine at this position for subsequent reaction steps is notable given that asparagine is naturally present at the equivalent position in yeast tRNA ligases (Fig. 1), and the equivalent residue is a histidine in Rnl2 (17). The glycine position of motif I is invariant in all RNA ligases, DNA ligases, and cellular capping enzymes, but the contribution of this residue in RNA ligation has not been addressed. Thus, we substituted Gly-102 with alanine. To query the roles of the defining residues in motifs IV (consensus (D/E)Gφφφφ, where φ is a hydrophobic side chain) and V (consensus φφKφφK), we introduced alanine substitutions at Glu-227 and Gly-228 in motif IV and Lys-240 and Lys-242 in motif V. We also mutated Asp-244 of motif V to alanine, and an acidic side chain is present at the equivalent positions in Ac-NPV ligase and S. cerevisiae tRNA ligase and in several DNA ligases and RNA capping enzymes. As a control we also changed Lys-99 to Ala.

The K99A, G102A, E227A, G228A, K240A, K242A, and D244A mutants were expressed in bacteria and purified from soluble lysates by Ni-agarose chromatography (Fig. 2). Muta-
tional effects on RNA ligation are shown in Fig. 3. The proteins were assayed under conditions of enzyme excess in order to illuminate the most severe catalytic defects. We found that mutants K99A, G102A, E227A, G228A, K240A, and K242A, and D244A mutants were expressed in bacteria and purified from soluble lysates by Ni-agarose chromatography (Fig. 2). Mutational effects on RNA ligation are shown in Fig. 3. The proteins were assayed under conditions of enzyme excess in order to illuminate the most severe catalytic defects. We found that mutants K99A, G102A, E227A, G228A, K240A, and K242A, and D244A mutants were expressed in bacteria and purified from soluble lysates by Ni-agarose chromatography (Fig. 2). Mutational effects on RNA ligation are shown in Fig. 3. The proteins were assayed under conditions of enzyme excess in order to illuminate the most severe catalytic defects. We found that mutants K99A, G102A, E227A, G228A, K240A, and K242A, and D244A mutants were expressed in bacteria and purified from soluble lysates by Ni-agarose chromatography (Fig. 2). Mutational effects on RNA ligation are shown in Fig. 3. The proteins were assayed under conditions of enzyme excess in order to illuminate the most severe catalytic defects. We found that mutants K99A, G102A, E227A, G228A, K240A, and K242A, and D244A mutants were expressed in bacteria and purified from soluble lysates by Ni-agarose chromatography (Fig. 2).
The Rnl1-Ala preparations were assayed for adenylyltransferase activity as described under “Experimental Procedures.” The specific activities were normalized to the activity of the wild-type (WT) protein (defined as 100%). The 10% threshold for defining an amino acid as essential for step 1 activity is indicated by a horizontal line.

Mutational Analysis of Other Conserved Amino Acids—The fact that Rnl1 and its baculovirus and fungal homologs contain only three of the five covalent nucleotidyltransferase motifs raises the prospect that either (i) the active site of Rnl1-like ligases is simpler than that of DNA ligases, capping enzymes, and Rnl2-like ligases or (ii) Rnl1 incorporates into its active site additional functional groups that are unique to this branch of the nucleotidyltransferase superfamily. To explore the latter scenario, we introduced single alanine changes at 13 additional amino acids of Rnl1, 9 of which (Arg-54, Gly-55, Arg-71, Lys-75, Phe-76, Phe-77, Asn-78, Glu-81, and Lys-119) are conserved in the putative baculovirus RNA ligase and the fungal tRNA ligases and 4 of which (Arg-166, Arg-182, Asn-184, and Glu-185) are conserved only in the baculovirus protein (Fig. 1). We focused in particular on basic residues, which we viewed as potential ligands for the ATP or RNA substrates, acidic residues as potential ligands for a divalent cation cofactor, and aromatic residues as candidates to engage in π stacking interactions with the purine base of the NTP (12–16).

The R54A, G55A, R71A, K75A, F76A, F77A, N78A, E81A, K119A, R166A, R182A, N184A, and E185A mutants were expressed in bacteria and purified from soluble bacterial lysates by Ni-agarose chromatography (Fig. 2). Their adenylyltransferase-specific activities were assayed by protein titration; the adenylyltransferase-specific activities were normalized to that of wild-type Rnl1 (Fig. 4). As expected, the K99A mutant was inert in the ligase adenylation reaction. The instructive findings were that the G102A, E227A, G228A, K240A, and K242A mutants were either inert or severely defective (≤2% of wild-type activity) in ligase adenylation, thereby explaining their inability to perform the composite pRNA ligation reaction. In contrast, the D244A mutant retained near wild-type adenylyltransferase activity, consistent with its near wild-type activity in pRNA ligation. We conclude that the conserved motif I glycine and the defining residues of motifs IV and V are essential for Rnl1 function.

Mutational Effects on Phosphodiester Formation at a Pre-adenylated RNA—We synthesized and purified a pre-adenylated 18-mer AppRNA and 800 ng (~20 pmol) of wild-type (WT) or mutant Rnl1 as specified. Rnl1 was omitted from a control reaction (lane –). The step 3 (phosphodiester bond formation) and reverse step 2 (RNA deadenylation) reactions are illustrated schematically at bottom.

Consistent with their severe step 1 defects, the K75A and F77A mutants were either unreactive or severely impaired in RNA ligation (Fig. 3). Mutants G55A, R71A, F76A, N78A, N184A, and E185A that were active in step 1 were also active in the composite ligation reaction. The striking finding was that K119A was inert in RNA ligation (Fig. 3), despite being fully active at the ligase adenylation step. The R54A mutant was also unreactive in strand joining, although it retained 30% of wild-type activity in ligase adenylation. We infer that Arg-54 and Lys-119 are essential for step(s) of the ligation pathway subsequent to formation of the Rnl1-AMP intermediate.

Mutants E81A and R182A were active in overall RNA ligation (Fig. 3), even though they displayed reduced activity in ligase adenylation. (Note that the use of excess enzyme in the ligation reactions may mask modest mutational effects on strand joining.) R166A was fully active in step 1 but displayed a partial defect in strand joining, characterized by reduced yield of RNA circles and enrichment of the AppRNA intermediate (Fig. 3).

Mutational Effects on Phosphodiester Formation at a Pre-adenylated RNA 5’ End—We synthesized and purified a pre-adenylated 18-mer RNA substrate (AppRNA) for analysis of step 3 of the ligation pathway in isolation. Rnl1 was reacted with AppRNA in enzyme excess in the absence of ATP. The ability of wild-type Rnl1 to form a phosphodiester at the activated 5’ end was manifest by the appearance of a sealed circular RNA product (Fig. 5). 90% of the AppRNA substrate was converted to circular RNA, and essentially no 5’-PO₄ RNA strand (pRNA) was formed by wild-type Rnl1; thus the reaction was biased toward ligation (step 3) and away from deadenylation (the reverse of step 2 of the ligation pathway). Circularization of RNA-adenylate required a divalent cation cofactor (data not shown).

Mutant proteins G55A, R71A, F76A, N78A, E81A, N184A, R166A, E185A, and D244A all retained step 3 activity (Fig. 5),
consistent with their ability to perform the composite pRNA ligation reaction (Fig. 3). The K99A mutant was unreactive with AppRNA, indicating that the motif I lysine is essential for phosphodiester formation even though the step 3 reaction does not entail a covalent intermediate. This result is consistent with the earlier report that mutation of Lys-99 to asparagine abrogated the step 3 activity of T4 Rnl1 (9). Mutants G102A and G228A were unreactive in sealing AppRNA (Fig. 5), just as they were unreactive in the ligase adenylation step (Fig. 4) and overall pRNA ligation (Fig. 3). The F77A mutant displayed feeble step 3 activity, consistent with its impairment at the enzyme adenylation step and in overall pRNA ligation.

Mechanistically instructive findings emerged from the apparent lack of effect of the K119A, K240A, and K242A changes on the isolated step 3 reaction (Fig. 5), which contrasts sharply with their abrogation of the overall pRNA ligation reaction (Fig. 3). Note that K119A retained step 1 activity, whereas K240A and K242A were defective. These results provide evidence that different constellations of side chain functional groups are responsible for catalysis of the three separate steps of the ligation pathway. Differential sparing of step 3 function was also observed for mutants K75A and E227A, which were capable of sealing the preadenylated substrate (albeit less effectively than wild-type Rnl1; Fig. 5) despite being unreactive in ligation of pRNA (Fig. 3). The R54A mutant was poorly active in circularizing AppRNA and was distinguished from the other mutants in that it generated deadenylated pRNA in roughly equal proportion to sealed RNA circles (Fig. 5).

A surprising finding was that the R182A enzyme was defective in the isolated step 3 reaction, while retaining apparently wild-type activity in ligase adenylation and overall pRNA ligation. This result suggested that the sealing of exogenous preformed AppRNA may differ in some respects from the sealing of the AppRNA intermediate formed in situ on the enzyme (see below).

Structure-Activity Relationships at Essential Residues of Rnl1—The present results highlight 9 individual amino acids in addition to the motif I lysine nucleophile (Lys-99) that are essential for Rnl1 activity: Arg-54, Lys-75, Phe-77, Gly-102, Lys-119, Glu-227, Gly-288, Lys-240, and Lys-242. To better evaluate the contributions of these residues to the RNA ligation reaction, we tested the effects of conservative substitutions at 7 of the essential positions (all except the glycines, for which alanine is already construed to represent a conservative change). Arginine was replaced by lysine and glutamine, glutamate by glutamine and aspartate, lysine by arginine and glutamine, and phenylalanine by leucine. Also, Arg-182 was replaced with lysine and glutamine in order to clarify its role in the isolated step 3 reaction. Fifteen new Rnl1 mutants were produced in E. coli and purified from soluble bacterial extracts by Ni-agarose chromatography (Fig. 6).

Each mutant was assayed in pRNA ligation under conditions of enzyme excess (Fig. 7). Conservative replacement of Glu-227 with Asp or Gln elicited a severe catalytic defect comparable with that seen with the E227A mutant. These data establish the requirement for a carboxylate residue at positions 227 (in motif IV) 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 this position in motif IV of tRNA ligases and the baculovirus Rnl1 homolog (Fig. 1). Although replacement of Lys-119 by Arg revived ligase activity to wild-type level, the K119Q protein remained catalytically inert, like the alanine mutant (Fig. 7). Replacement of Arg-54 by Lys resulted in a significant gain of ligase function compared with R54A (which was unreactive), but the R54Q change had no salutory effect (Fig. 7). Similarly, the introduction of Arg in lieu of Lys-75 partially restored ligation activity, whereas the K75Q mutant was unreactive (Fig. 7). We surmise that the positive charges at positions 54, 75, and 119 are critical for pRNA strand joining. (Note that the sites corresponding to Rnl1 Lys-75 and Lys-119 are naturally occupied by arginine in the putative baculovirus Rnl1 homolog; see Fig. 1.) In contrast, Lys-240 and Lys-242 in motif V of Rnl1 were strictly essential and could not be functionally replaced by either Arg or Gln (Fig. 7). The F77L mutant was impaired in pRNA ligation and only marginally more active than F77A. This finding attests to the importance of the aromatic side chain at position 77. The R182K and R182Q proteins were active in pRNA ligation, as expected given that R182A was also active.

The adenylyltransferase specific activities of the conservative mutants were assayed by protein titration and normalized to that of wild-type Rnl1 (Fig. 8). Mutational effects on step 1 were concordant with the effect on overall ligation at only some
of the residues. At Lys-75, a position deemed essential for step 1 in the alanine scan, we found that the arginine change restored adenylyltransferase activity to 40% of wild-type (consistent with the gain of function in pRNA ligation), whereas the glutamine mutation was just as deleterious as alanine. The E227D and E227Q mutants were both defective in ligase adenylation (5 and 7% as active as wild-type, respectively), thereby accounting for their profound defects in RNA ligation. F77L was also impaired in step 1 (15% of wild-type activity), thereby explaining its weak overall ligation activity.

Other conservative mutations had differential effects on step 1 versus overall ligation. For example, K119R and K119Q were both quite active in step 1 (as was K119A), yet only K119R was competent in pRNA ligation. The R54K and R54Q mutants were also active in step 1 (50–60% of wild-type; slightly higher than R54A), yet only R54K had RNA ligase activity. These results underscore the critical role played by these two positively charged residues at a step subsequent to ligase adenylation.

Conservative changes in the motif V lysines had distinctive effects on step 1 function. The K240R mutation restored adenylyltransferase activity to 30% of wild type, whereas the K240Q mutant was as defective as K240A, suggesting that a positive charge at position 240 was sufficient for ligase adenylation (albeit not for pRNA ligation; Fig. 7). The K242R and K242Q mutant had modestly improved step 1 activity (14 and 11% of wild-type) compared with the alanine mutant (3%), but neither was able to ligate pRNA efficiently. Thus lysine is specifically required at this position for activity in step 1 and overall ligation.

Effects of Conservative Mutations on Sealing of RNA-Adenylate—Distinct Requirements for the Three Steps of the Ligation Pathway—Conservative mutants K119R and K119Q were both active in the isolated step 3 reaction (Fig. 9). This stands in contrast to the overall pRNA ligation reaction in which the K119R protein was active but the K119Q mutant was inert (Fig. 6). The glutamine change at Lys-119 elicited the same set of biochemical defects as the alanine mutation, i.e. preservation of steps 1 and 3, but loss of overall ligase activity (summarized in Table I). These findings implicate Lys-119 as a specific catalyst of the RNA adenylation step (step 2) of the ligase reaction.

Distinctive structure-function relationships in step 3 were evident at the two essential lysines of motif V. Conservative mutations K242R and K242Q at the downstream lysine did not affect the isolated step 3 reaction (Fig. 9), consistent with the findings for the K242A mutant (Fig. 5). Conservative mutations K240R and K240Q at the upstream lysine resulted in impaired sealing of pre-formed AppRNA (Fig. 9), unlike the K240A mutant (Fig. 5). We surmise that the essentiality of Lys-240 and Lys-242 in overall pRNA ligation reflects their catalytic roles during the ligase adenylation reaction (step 1; see Table I). It is possible that Lys-242 also participates in AMP transfer to RNA, insofar as the K242Q mutant displayed partial step 1 activity but remained unreactive in pRNA ligation.

Replacing Arg-54 with lysine or glutamine resulted in a gain of step 3 function (Fig. 9) compared with the step 3-defective R54A mutant. (Note that the R54K and R54Q still generated a small amount of deadenylated pRNA, as seen with R54A.) Thus, the arginine side chain is not strictly essential for phosphodiester formation, a polar functional group suffices. The instructive finding was that R54Q is defective in overall pRNA ligation despite being active in steps 1 and 3 (Table I). This, plus the fact that lysine at position 54 restored pRNA ligation activity, implies a critical role for a positively charged side chain at position 54 in the RNA-adenylation reaction (step 2) of the ligase pathway.

Introducing arginine or glutamine in lieu of Lys-75 restored step 3 activity to wild-type levels (Fig. 9). The K75R and K75Q enzymes were more active in step 3 than K75A (Fig. 5). The properties of the K75Q mutant in particular (i.e. defective in step 1, active in step 3, and defective in overall pRNA ligation) highlight the specific contributions of Lys-75 during the adenylyltransferase reaction of the ligase pathway (Table I). Partial restoration adenylyltransferase of overall ligase activity for the K75R enzyme indicated that positive charge is the key property at position 75.

The F77L enzyme was partially active in ligating the pre-adenylated RNA (Fig. 9), implying that the aromatic side chain was important for full activity in step 3, as it was in step 1 and the overall pRNA ligation reaction (Table I). The biochemical defects of the Phe-77 mutants could reflect a role for Phe-77 in a process common to steps 1 and 3 (e.g. an interaction with the adenosine nucleotide of ATP or AppRNA); alternatively,
steps 1 and 3 of the ligation pathway, thereby accounting for E227A change. We infer that Glu-227 plays a major role in Asp or Gln caused severe step 3 defects, even more so than the sealing of the AppRNA intermediate formed AppRNA requires a positive charge at position 119, whereas a lysine at this position fully revived step 3 function, but the glutamine mutant remained unreactive in step 3 (Fig. 9), de- pite its activity in ligase adenylation and overall pRNA ligation. Although there has been no targeted mutational analysis of fungal tRNA ligases, we would speculate that the conserved essential positions demonstrated here for Rnl1 are likely to be important for Trl1 activity as well. Indeed, given that the physiological role of T4 Rnl1 in thwarting tRNA restriction is quite similar to the splicing step catalyzed by Trl1 (6, 37, 38), we propose that T4 Rnl1 and Trl1 have a shared evolutionary history whereby they descend from an ancestral ligase devoted to repairing broken tRNAs. We suspect that the baculovirus Rnl1 homolog may play a similar role in RNA repair, insofar as its primary structure is made up of separate domains resembling T4 RNA ligase, T4 polynucleotide kinase, and T4 polynucleotide 3′ phosphatase, respectively (19, 39).

**Table I**

| Rnl1 mutant | Overall ligation | Step 1 | Step 3 |
|-------------|-----------------|--------|--------|
| R54A        | –               | +      | +      |
| R54K        | ++              | +      | +      |
| R54Q        | ++              | +      | +      |
| K75A        | –               | –      | –      |
| K75R        | +               | +      | +      |
| K75Q        | –               | –      | –      |
| F77A        | –               | –      | –      |
| F77L        | +               | +      | +      |
| K99A        | –               | –      | –      |
| G102A       | –               | –      | –      |
| K119A       | –               | +      | +      |
| K119R       | ++              | +      | +      |
| K119Q       | –               | –      | –      |
| R182A       | –               | –      | –      |
| R182K       | –               | –      | –      |
| R182Q       | ++              | +      | +      |
| E227A       | –               | –      | –      |
| E227D       | –               | –      | –      |
| E227Q       | –               | –      | –      |
| K240A       | –               | –      | –      |
| K240R       | –               | –      | –      |
| K240Q       | –               | –      | –      |
| K242A       | –               | –      | –      |
| K242R       | –               | –      | –      |
| K242Q       | –               | –      | –      |

**Discussion**

**Defining the Rnl1 Ligase Family**—To understand the structural requirements for RNA ligation, we initiated an alanine-scanning mutational analysis of selected residues of T4 Rnl1 that are conserved in fungal tRNA ligases and/or the putative baculovirus Rnl1 homolog. We report that conserved residues in nucleotidyltransferase motifs I (Lys-99 and Gly-102), IV (Glu-227 and Gly-228), and V (Lys-240 and Lys-242) are essential for the ligase activity of Rnl1 (see Table I). The corresponding side chains were shown previously to be essential for the activities of DNA ligases (23–26, 30–33) and mRNA capping enzymes (27, 28, 34–36) and also for the activity of T4 RNA ligase 2 (17, 18), which belongs to a different RNA ligase family than Rnl1. These results hint that the structural basis for nucleotidyl transfer is at least partially conserved among RNA ligases, DNA ligases, and mRNA capping enzymes. Such conservation is consistent with the speculation that RNA-joining enzymes that evolved during a primordial RNA/protein world are the ancestors of present day DNA ligases and mRNA capping enzymes (11).

At the same time, our results highlight several signature residues (Arg-54, Lys-75, and Phe-77), located upstream of the motif I lysine nucleophile (Lys-99), that are found in all of the Rnl1-like ligases. They are essential for Rnl1 ligase activity and have no apparent counterparts in ATP-dependent DNA ligases, RNA capping enzymes, and Rnl2-like RNA ligases. Although there has been no targeted mutational analysis of fungal tRNA ligases, we would speculate that the conserved essential positions demonstrated here for Rnl1 are likely to be important for Trl1 activity as well. Indeed, given that the physiological role of T4 Rnl1 in thwarting tRNA restriction is quite similar to the splicing step catalyzed by Trl1 (6, 37, 38), we propose that T4 Rnl1 and Trl1 have a shared evolutionary history whereby they descend from an ancestral ligase devoted to repairing broken tRNAs. We suspect that the baculovirus Rnl1 homolog may play a similar role in RNA repair, insofar as its primary structure is made up of separate domains resembling T4 RNA ligase, T4 polynucleotide kinase, and T4 polynucleotide 3′ phosphatase, respectively (19, 39).

**Insights from Mutagenesis to Rnl1 Mechanism**—We assessed alanine and conservative mutational effects on overall RNA ligation and the isolated steps of ligase adenylylation and phosphodiester formation. Nine residues in addition to the Lys-99 nucleophile were deemed essential for overall ligation, and one (Arg-182) was paradoxically critical for the isolated step 3 reaction but not for overall ligation (Table I). As reported previously (9), the motif I lysine nucleophile was essential for both steps 1 and 3. Other positions found to be important for all phases of the Rnl1 ligation reaction include Gly-102 (in motif I), Glu-227, and Gly-228 (in motif IV), and Phe-77. Three residues, Lys-75, Lys-240, and Lys-242, were identified as essential for phosphodiester bond formation. Thus, we surmise that the composition of the active site is different during steps 1 and 3. Mutations at Arg-54 and Lys-119 abolished the overall pRNA ligation reaction without affecting steps 1 and 3. Arg-54 and Lys-119 are thereby implicated as specific catalysts of the RNA adenylation reaction (step 2) of the ligation pathway.

Our inferences about which features of the individual side chains are required for activity have been discussed above. Thus, we focus here on the mechanistic implications and how the findings may be interpreted in light of the crystal structures available for other members of the covalent nucleotidyltransferase superfamily. In the absence of an atomic structure for any RNA ligase, we assume a correspondence between the essential amino acids in motifs I, IV, and V of Rnl1 and the amino acids found at “equivalent” positions of motifs I, IV, and V of DNA ligases and capping enzymes (12–16).

The Rnl1 motif IV glutamate (Glu-227) was critical for Rnl1 ligation activity and for the individual steps of ligase adenylation and phosphodiester formation. Conservative changes were not tolerated. Similar mutational effects on step 1, step 3, and overall ligation activity were observed for T4 Rnl2 (17, 18). Alanine and conservative changes at the motif IV glutamate of *Chlorella* virus DNA ligase abrogated the overall DNA nick joining reaction and the ligase adenylation reaction (step 1) and reduced the efficiency of phosphodiester bond formation on a pre-formed nicked DNA-adenylate substrate (32). Thus, the acidic side chain in motif IV appears to be universally essential for the polynucleotide ligation pathway, independent of whether it involves a single-stranded RNA or duplex DNA.
substrate. We infer a direct role for the glutamate side chain at the active site of Rnl1, perhaps in coordinating the divalent cation cofactor, as suggested for the motif IV glutamate of DNA ligase (14).

The two Rnl1 motif V lysines (Lys-240 and Lys-242) were essential for overall ligation activity and specifically for step 1 of the pathway but not for step 3. The motif V lysines of Rnl2 (Lys-225 and Lys-227) are also required for ligase adenylation and overall pRNA ligation, but some mutations of the Rnl2 lysine strongly affect step 3 of the pathway (18). For example, the sealing of AppRNA by Rnl2 is abrogated by arginine substitution at the distal motif V lysine, but this change did not affect Rnl1.

The available crystal structures of capping enzymes and DNA ligases reveal that these two lysines contact the \( \alpha \) or \( \gamma \) phosphates of the NTP substrate, the phosphate of the covalent lysyl-NMP intermediate, or a sulfate/phosphate ion on the enzyme surface (12–16). The sulfate ion is proposed to occupy the position of 5'-phosphate of the polynucleotide undergoing ligation, i.e. the strand that become 5'-adenylated during the second step of the ligation pathway (14). In light of these structures of related enzymes, we hypothesize a direct role for the Rnl1 motif V lysines in catalysis of step 1, possibly in stabilization of the transition state of the \( \alpha \) phosphorus of ATP or in directing the PP, leaving group to an apical position with respect to the Lys-99 nucleophile (13).

**A New Motif Conserved in the Polynucleotide Ligase/Capping Enzyme Superfamily**—The essential Lys-119 residue of Rnl1, which is conserved as a basic side chain in all of the Rnl1-like ligases, is located 20–21 amino acids downstream of the motif I lysine nucleophile, where it is flanked by conserved hydroxymino acid (Ser or Thr) immediately upstream (Fig. 1). A Ser-Arg dipeptide motif is also found 20–23 amino acids downstream of the motif I lysine in numerous ATP-dependent DNA ligases, including exemplary enzymes from Archaea, bacteriophages, eukaryotic viruses, fungi, and mammals (40). The crystal structure of the Chlorella virus DNA ligase-AMP intermediate reveals that the Ser-Arg motif is located on the enzyme surface, where the arginine (Arg-42) coordinates a sulfate ion on the protein surface ~5 Å from the \( \alpha \)-phosphate of AMP (14). The sulfate contacts are proposed to mimic the interactions of DNA ligase with the reactive 5'-phosphate of the polynucleotide (and with the \( \gamma \)-phosphate of ATP). Indeed, mutation of this arginine in the Chlorella virus DNA ligase reduces nick joining activity by a factor of 50 and abrogates the binding of DNA ligase to nicked duplex DNA (14). In this light, it is instructive that mutations of the homologous Lys-119 side chain in Rnl1 appeared to specifically interfere with step 2 of the ligation pathway, which entails recognition of the 5'-phosphate of the pRNA substrate.

Cellular and viral mRNA capping enzymes also contain a conserved arginine located 21–26 amino acids downstream of the motif I lysine nucleophile (36). The residue immediately upstream of the arginine is typically aspartate or threonine in the capping enzymes. Mutational analysis of mammalian capping enzyme shows that this arginine is essential for activity (36). The arginine is located at a position in crystal structures of Chlorella virus and Candida albicans capping enzymes equivalent to that of the Ser-Arg motif in ATP-dependent DNA ligases (13, 16). In the Chlorella virus guanylyltransferase-GTP complex, the Arg-106 side chain coordinates the \( \gamma \)-phosphate of GTP. In the covalent guanylyltransferase-GMP intermediate, the arginine coordinates a sulfate ion on the protein surface (perhaps mimicking contacts to the 5’ terminus of the RNA substrate to be capped with GMP) (13). The equivalent arginine of the C. albicans guanylyltransferase-GMP intermediate coordinates a phosphate on the enzyme surface located close to the lysyl-GMP phosphate; this phosphate is proposed to mark the positions of the 5’-diphosphate RNA terminus prior to the second GMP transfer step of the capping pathway (16).

T4 DNA ligase 2 and its orthologs also have a conserved arginine situated 20–24 amino acids downstream of the motif I lysine nucleophile (17). Recent mutational analysis of Rnl2 shows that the arginine is essential for pRNA ligation (i.e. alanine and glutamine substitutions abolish Rnl2 activity) and that conservative substitution with lysine restores ligation activity (18).

Based on these functional and structural data, we conclude that the conserved Arg/Lys side chain downstream of motif I is an essential component of the active sites of ATP-dependent RNA and DNA ligases and GTP-dependent mRNA capping enzymes. Thus we designate the (S/T/D)-R(K) peptidase as “motif Ia” of the covalent nucleotidyltransferase superfamily (Fig. 1).

**Signature Features of the Rnl1-like Enzymes—Essential residues** Arg-54, Lys-75, and Phe-77 are conserved in all of the Rnl1-like ligases. ATP-dependent DNA ligases, RNA capping enzymes, and Rnl2-like RNA ligases have no obvious equivalents of these residues. Indeed, there is no recognizable homology between the entire N-terminal segment preceding motif I of Rnl1-like ligases and the N-terminal segments of ATP-dependent DNA ligases, RNA capping enzymes, and Rnl2-like RNA ligases. In fact, no catalytic residues have been mapped at a distance upstream of motif I in any of the latter group of enzymes. The occurrence of a unique N-terminal functional domain in Rnl1-like ligases may have to do with the seeming absence in Rnl1 of counterparts of motifs III and IIIa that are essential for the function of ATP-dependent DNA ligases, RNA guanylyltransferases, and Rnl2. Motif III (consensus sequence \( \text{GxxxGxxxGxxx} \)) is typically located 40–50 amino acids downstream of motif I, and it includes an essential glutamate that coordinates the ribose sugar of the NTP substrate (12, 13). Motif IIIa is located 10–30 amino acids downstream of motif III. Motif IIIa forms part of the nucleotide binding pocket, to which it contributes a conserved aromatic residue (either Tyr or Phe) that forms a \( \pi \) stack on the purine base of the NTP substrate and the covalent NMP intermediate (12–16).

Although the fold of the N-terminal domain of Rnl1 and the contacts of Arg-54, Lys-75, and Phe-77 are still unknown, we can speculate based on the present biochemical findings what their function might be. Arg-54 is specifically implicated in step 2 of the Rnl1 ligation pathway. Therefore, Arg-54 is a candidate to coordinate the 5'-phosphate of the pRNA substrate (together with Lys-119, as discussed above). The fact that a lysine can function in lieu of Arg-54 is consistent with an electrostatic interaction with phosphate. Lys-75 is required for the Rnl1 adenylation step but not for phosphodiester formation. This functional profile is consistent with an interaction between Lys-75 and the \( \gamma \)-phosphate of ATP in step 1. (Note that there is no equivalent of a \( \gamma \)-phosphate in the AppRNA substrate in step 3.) Phe-77 is implicated in steps 1 and 3. As noted above, this aromatic group is a candidate to form part of the adenine binding pocket (for ATP and AppRNA) via \( \pi \) stacking interactions with the adenine base, such as those made by the aromatic side chains in motif IIIa of other covalent nucleotidyltransferases.

Finally, we noted paradoxical effects of Lys-182 mutations that abrogated the sealing of pre-formed AppRNA by Rnl1 but spared the overall pRNA ligation reaction, which perchance requires catalysis of step 3. This result echoes the finding that the rate of sealing of an exogenous pre-adenylated nicked DNA substrate (AppDNA) by Chlorella virus DNA ligase was much slower than the rate of the composite pDNA ligation reaction (41). Similar findings were reported by Modrich and Lehman.
for E. coli DNA ligase. The explanation offered for this anomaly was that sealing of an exogenous AppDNA substrate entails a slow step in the binding of free AppDNA that does not apply when the adenylated polynucleotide intermediate is formed in situ at the ligase-active site (41). We invoke a similar explanation for the effects of the Lys-182 mutations on the isolated step 3 reaction of RnI, i.e. that Arg-182 facilitates the binding of exogenous AppRNA but is dispensable when the RNA-adenylate is formed in situ. Arg-182 is located within a peptide sequence 182RENERTGEY189 that is conserved in the putative baculovirus Rnl1 homolog, but there is no obvious equivalent of this motif in the fungal tRNA ligases. The lack of conservation of Arg-182 in all of RnI-like enzymes is consistent with our inference that Arg-182 is not a true component of the active site.

REFERENCES
1. Silber, R., Malathi, V. G., and Hurwitz, J. (1972) Proc. Natl. Acad. Sci. U. S. A. 69, 3009–3013
2. Cranston, J. W., Silber, R., Malathi, V. G., and Hurwitz, J. (1974) J. Biol. Chem. 249, 7447–7456
3. Sugino, A., Snopk, T. J., and Cozzarelli, N. R. (1978) J. Biol. Chem. 252, 1732–1738
4. Uhlenbeck, O. C., and Gunport, R. I. (1982) Enzymes 15, 31–58
5. Engler, M. J., and Richardson, C. C. (1982) Enzymes 15, 3–29
6. Amtur, M., Levitz, R., and Kaufman, G. (1987) EMBO J. 6, 2499–2503
7. Rand, K. N., and Gait, M. J. (1984) EMBO J. 3, 397–402
8. Thogerson, H. C., Morris, H. R., Rand, K. N., and Gait, M. J. (1985) Eur. J. Biochem. 147, 325–329
9. Hepch, S., Singh, M., and Gait, M. J. (1987) Biochemistry 26, 1688–1696
10. Shuman, S., and Schwer, B. (1995) Mol. Microbiol. 17, 405–410
11. Shuman, S. (2000) Prog. Nucleic Acids Res. Mol. Biol. 66, 1–40
12. Subramanya, H. S., Dobert, A. J., Ashford, S. R., and Wigley, D. B. (1996) Cell 85, 607–615
13. Håkansson, K., Dobert, A. J., Shuman, S., and Wigley, D. B. (1997) Cell 89, 545–553
14. Odell, M., Sriskanda, V., Shuman, S., and Nikolov, D. (2000) Mol. Cell 6, 1183–1193
15. 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
16. Fabrega, C., Shen, V., Shuman, S., and Lima, C. D. (2003) Mol. Cell. 11, 1549–1561
17. Ho, C. K., and Shuman, S. (2002) Proc. Natl. Acad. Sci. U. S. A. 99, 12709–12714
18. Yin, S., Ho, C. K., and Shuman, S. (2003) J. Biol. Chem. 278, 17601–17608
19. Duranton, D., Crozier, L., Ayres, M. D., Crozier, G., Possee, R. D., and Lopez-Ferber, M. (1999) J. Gen. Virol. 79, 629–637
20. Westaway, S. K., Phizicky, E. M., and Abelson, J. (1988) J. Biol. Chem. 262, 3171–3176
21. Xu, Q., Teplow, D., Lee, T. D., and Abelson, J. (1990) Biochemistry 29, 6132–6138
22. Baymiller, J., Jennings, S., Kienitz, B. K., Gorman, J. A., Kelly, R., and McCullough, J. E. (1994) Gene (Amst.) 142, 129–134
23. Kodama, K., Barnes, D. E., and Lindahl, T. (1991) Nucleic Acids Res. 19, 6093–6099
24. Luo, J., and Barnay, F. (1996) Nucleic Acids Res. 12046–12049
25. Sriskanda, V., and Shuman, S. (1998) Nucleic Acids Res. 26, 525–531
26. Sriskanda, V., Kelman, Z., Hurwitz, J., and Shuman, S (2000) Nucleic Acids Res. 28, 2221–2228
27. Cong, P., and Shuman, S. (1993) J. Biol. Chem. 268, 7256–7260
28. Schwer, B., and Shuman, S. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 4328–4332
29. Ho, S. N., Hunt, H. D., Horton, R. M., Pulles, J. K., and Pease, L. R. (1989) Gene (Amst.) 77, 51–59
30. Shuman, S., and Xu, X. (1995) Virology 211, 73–83
31. Sriskanda, V., and Shuman, S. (2002) J. Biol. Chem. 277, 9661–9667
32. Sriskanda, V., and Shuman, S. (2002) Nucleic Acids Res. 30, 903–911
33. Doherty, A. J., and Dafforn, T. R. (2000) J. Mol. Biol. 296, 43–56
34. Shuman, S., Liu, Y., and Schwer, B. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 12045–12050
35. Wang, S. P., Deng, L., Ho, C. K., and Shuman, S. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 9573–9578
36. Sawaya, R., and Shuman, S. (2003) Biochemistry, in press
37. Kaufman, G. (2000) Trends Biochem. Sci. 25, 70–74
38. Abelson, J., Tennet, R. C., and Li, H. (1998) J. Biol. Chem. 273, 12685–12688
39. Wang, L. K., and Shuman, S. (2002) Nucleic Acids Res. 30, 1073–1080
40. Kletzin, A. (1992) Nucleic Acids Res. 20, 5389–5396
41. Sriskanda, V., and Shuman, S. (1998) Nucleic Acids Res. 26, 4618–4625
42. Modrich, P., and Lehman, I. R. (1973) J. Biol. Chem. 257, 8117–8122
