Articles

Effects of Maturase Binding and Mg^{2+} Concentration on Group II Intron RNA Folding Investigated by UV Cross-Linking†

James W. Noah and Alan M. Lambowitz*

Institute for Cellular and Molecular Biology, Department of Chemistry and Biochemistry, and Section of Molecular Genetics and Microbiology, School of Biological Sciences, University of Texas at Austin, Austin, Texas 78712

Received July 28, 2003; Revised Manuscript Received September 3, 2003

ABSTRACT: The *Lactococcus lactis* Ll.LtrB group II intron encodes a reverse transcriptase/maturase (LtrA protein) that promotes RNA splicing by stabilizing the catalytically active RNA structure. Here, we mapped 17 UV cross-links induced in both wild-type Ll.LtrB RNA and Ll.LtrB-Δ2486 RNA, which has a branch-point deletion that prevents splicing, and we used these cross-links to follow tertiary structure formation under different conditions in the presence or absence of the LtrA protein. Twelve of the cross-links are long-range, with six near known tertiary interaction sites in the active RNA structure. In a reaction medium containing 0.5 M NH_4Cl, eight of the 17 cross-links were detected in the absence of Mg^{2+} or the presence of EDTA, and all were detected at 5 mM Mg^{2+}, where efficient splicing requires the LtrA protein. The frequencies of all but four cross-links increased with increasing Mg^{2+} concentrations, becoming maximal between 4 and 50 mM Mg^{2+}, where the intron is self-splicing. These findings suggest that a high Mg^{2+} concentration induces self-splicing by globally stabilizing tertiary structure, including key tertiary interactions that are required for catalytic activity. Significantly, the binding of the maturase under protein-dependent splicing conditions (0.5 M NH_4Cl and 5 mM Mg^{2+}) increased the frequency of only nine cross-links, seven of which are long-range, suggesting that, in contrast to a high Mg^{2+} concentration, LtrA promotes splicing by stabilizing critical tertiary structure interactions, while leaving other regions of the intron relatively flexible. This difference may contribute to the high rate of protein-dependent splicing, relative to the rate of self-splicing. The propensity of the intron RNA to form tertiary structure even at relatively low Mg^{2+} concentrations raises the possibility that the maturase functions at least in part by tertiary structure capture. Finally, an abundant central wheel cross-link, present in >50% of the molecules at 5 mM Mg^{2+}, suggests models in which group II intron domains I and II are either coaxially stacked or aligned in parallel, bringing the 5′-splice site together with the 3′-splice site and catalytic core elements at JII/III. This and other cross-links provide new constraints for three-dimensional structural modeling of the group II intron catalytic core.

Group II introns are catalytic RNAs that splice via two sequential transesterification reactions involving the formation of an intron lariat, and they are widely believed to be ancestors of nuclear pre-mRNA introns (*1*). Some group II introns self-splice *in vitro*, but this self-splicing is generally inefficient and requires nonphysiological conditions, such as high salt, Mg^{2+}, or temperature, implying that proteins are required *in vivo* to help the intron RNA fold into the catalytically active structure (*2*). Although host-encoded proteins play a role in splicing some group II introns, a key protein required for splicing mobile group II introns is the intron-encoded reverse transcriptase (RT),[1] which also functions in intron mobility (*3*−*5*). After splicing, the RT remains associated with the excised intron lariat RNA to promote mobility by a novel target DNA-primed reverse transcription mechanism in which the intron RNA reverse splices directly into a DNA target site and is then reverse transcribed by the intron-encoded protein (IEP) (*6*−*9*). If mobile group II introns were the evolutionary ancestors of nuclear pre-mRNA introns, further studies of their splicing and mobility mechanisms may provide insight into the evolution of introns and splicing mechanisms in higher organisms.

† This work was supported by National Institutes of Health Grant GM37951 to A.M.L. J.W.N. is the recipient of National Institutes of Health Postdoctoral Fellowship GM20780.

* To whom correspondence should be addressed. Telephone: (512) 232-3418. Fax: (512) 232-3420. E-mail: lambowitz@mail.texas.edu.

[1] Abbreviations: AMV, avian myeloblastosis virus; BTBE, 30 mM bis(2-hydroxyethyl)iminotris(hydroxymethyl)methane, 30 mM boric acid, and 2.5 mM EDTA (pH 6.8); DI−DVI, domains I−VI, respectively; E1, 5′-exon; E2, 3′-exon; EBS, exon-binding site; EDTA, ethylenediaminetetraacetic acid; IBS, intron-binding site; IEP, intron-encoded protein; JII/III, junction region between domains II and III; ORF, open reading frame; phenol-CIA, phenol, chloroform, and isoamyl alcohol (25:24:1); RNP, ribonucleoprotein; RT, reverse transcriptase; SD, Shine-Dalgarno; TBE, 90 mM Tris, 90 mM boric acid, and 2 mM EDTA (pH 8.2).

10.1021/bi035339n CCC: $25.00
Published on Web 10/11/2003

The catalytic activity of group II intron RNAs is dependent upon the formation of a conserved three-dimensional structure that forms an active site containing catalytically essential Mg^{2+} ions (*10*−*15*). This active site binds the 5′- and 3′-splice sites and aligns and activates the appropriate phosphodiester bonds for catalysis. The conserved group II intron secondary structure consists of six double-helical domains, denoted domains I−VI (DI−DVI, respectively; see Figure 3 below). DI contains three short sequence elements [exon-binding sites 1 and 2 (EBS1 and EBS2, respectively) and δ in subgroup IIA introns], which base pair with flanking 5′- and 3′-exon sequences [intron-binding sites 1 and 2 (IBS1 and IBS2, respectively) and δ'] for exon definition during RNA splicing and for DNA target site recognition during intron mobility. DV, the most highly conserved domain, interacts with DI to form the minimal catalytic core, while DVI contains the branch-point nucleotide residue, generally a bulged A residue. DII and DIII contribute to RNA folding and formation of the catalytic core, with DIII functioning as a catalytic effector (*16*). DIV, which encodes the intron open reading frame (ORF), also contributes to RNA folding, but is not required for ribozyme activity (*17*, *18*). A number of conserved tertiary interactions between different group II intron domains have been identified (denoted with Greek letters) (*1*, *10*, *19*−*22*); an X-ray crystal structure has been determined for an RNA containing DV and DVI (*23*), and three-dimensional models of the catalytic core elements have been constructed (*11*, *13*).

To study group II intron RNA splicing and mobility reactions in detail, we have developed an experimental system based on the *Lactococcus lactis* Ll.LtrB intron, a group IIA intron closely related to yeast mtDNA introns aI1 and aI2 (*24*, *25*). A key experimental advantage of the Ll.LtrB intron is the availability of an efficient *Escherichia coli* expression system, which makes it possible to obtain large amounts of purified, active IEP (denoted LtrA protein) for biochemical and structural studies (*4*, *26*). By using this system, we showed that the purified LtrA protein is both necessary and sufficient for efficient splicing of the Ll.LtrB intron at physiologically relevant Mg^{2+} concentrations (5 mM Mg^{2+}) and that it is an intron specific splicing factor, which binds to and splices the Ll.LtrB intron, but not other group II introns (*4*). We also showed that the purified LtrA protein could be reconstituted with the excised intron lariat RNA to form RNPs that carry out the reverse splicing and target DNA-primed reverse transcription reactions for intron mobility (*4*). Thus, the interaction of the LtrA protein with the intron RNA is central to both RNA splicing and intron mobility.

We previously investigated the interaction of the LtrA protein with the Ll.LtrB intron by using intron RNA deletion mutants (*18*), chemical RNA structure mapping and footprinting (*27*), and *in vitro* selection and mutagenesis (*28*). These studies showed that under protein-dependent *in vitro* splicing conditions (0.5 M NH_4Cl and 5 mM Mg^{2+}), the Ll.LtrB RNA forms most of the conserved group II intron secondary structure and at least some tertiary structure, and that LtrA binds to this partially folded intron to stabilize tertiary interactions required for catalytic activity. LtrA has a primary high-affinity binding site in intron subdomain DIVa, an idiosyncratic stem−loop structure that is located at the beginning of DIV and contains the ribosome-binding site and initiation codon of the LtrA ORF (*18*, *27*, *28*). The binding of LtrA to this region downregulates its own translation and appears to serve as an anchor for additional contacts with conserved regions of the intron that stabilize the catalytically active RNA structure. The RNA structure mapping and footprinting experiments identified potential secondary binding sites in DI, DII, and DVI and showed directly that the binding of the maturase promotes the formation of conserved long-range tertiary interactions (e.g., $\delta-\delta'$, $\epsilon-\epsilon'$, $\gamma-\gamma'$, $\kappa-\kappa'$, $\lambda-\lambda'$, and $\zeta-\zeta'$), which do not form stably in the free RNA at low Mg^{2+} concentrations (*27*).

Here, we used UV cross-linking methods (*29*, *30*) to study conformational changes in the Ll.LtrB intron RNA induced by LtrA binding or by high Mg^{2+} concentrations that support self-splicing. Our results indicate that LtrA binds to the partially folded intron to stabilize critical tertiary structure interactions that are required for RNA catalytic activity. In contrast to a high Mg^{2+} concentration, which globally stabilizes tertiary structure and may also stabilize non-native structures, the binding of LtrA at low Mg^{2+} concentrations stabilizes critical interactions, while leaving other regions of the intron RNA relatively flexible. We suggest that this difference accounts at least in part for the finding that the rate of maturase-promoted splicing at 5 mM Mg^{2+} is ≥30-fold faster than the maximum rate of self-splicing. Finally, the cross-links identified here suggest models in which group II introns DI and DII are either coaxially stacked or aligned in parallel, bringing the 5′- and 3′-splice sites together at JII/III, a key region of the catalytic core, and they provide constraints for further three-dimensional modeling of the group II intron RNA structure.

MATERIALS AND METHODS

Preparation of the LtrA Protein. The LtrA protein was synthesized in *E. coli* BL21(DE3) using the expression plasmid pImp-1P (*4*). This plasmid contains the LtrA ORF cloned downstream of a tac promoter in the intein-based expression vector pCYB2 (Impact, New England Biolabs, Beverly, MA), with the C-terminus of the ORF fused in-frame to a cassette consisting of the *Saccharomyces cerevisiae* VMA1 intein and a *Bacillus circulans* chitin-binding domain. The expressed LtrA protein was purified via chromatography on a chitin affinity column (*4*). Protein concentrations were determined by the Bradford assay (*31*), using the LtrA protein whose concentration had been determined by A_{280} as a standard. The protein preparations were >98% pure, as judged by Coomassie blue-stained SDS−polyacrylamide gels.

In Vitro Transcription. Ll.LtrB RNA contains a 902-nucleotide ΔORF derivative of the *L. lactis* Ll.LtrB intron flanked by 5′- and 3′-exons of 32 and 37 nucleotides, respectively, and Ll.LtrB-ΔA2486 RNA is the corresponding RNA deleted for branch-point nucleotide residue A2486. These RNAs were transcribed from plasmids pJNΔORF and pJNΔORF-ΔA2486, respectively, which were constructed by recloning the Ll.LtrB transcription units of pGMΔORF (*4*) or pGMΔORF-ΔA2486 (*27*) amplified with primers 5′-CGCGCGGATCCGGGCCCACGTCGATCGTGAACAC-ATCC and 5′-CGCGCGGATCCGTATAAAGATTCGTA-GAATTAAAATG into the *Bam*HI site of pUC19. For *in vitro* transcription, the plasmids were linearized with *Bam*HI

and transcribed with phage T7 RNA polymerase (Stratagene, La Jolla, CA) (*4*). For UV cross-linking, gel-purified transcripts were 5′-labeled with [γ-^{32}P]ATP using phage T4 polynucleotide kinase (*32*). Transcripts used for RNA splicing assays were internally labeled by *in vitro* transcription in the presence of [α-^{32}P]UTP (*18*). Prior to use, RNAs were renatured by being heated to 50 °C for 1 min in the reaction medium, and then slowly cooled over the course of ~20 min to the desired temperature.

RNA Splicing. Self-splicing was carried out with 20 nM ^{32}P-labeled precursor RNA in 1 mL of reaction medium containing 0.5 or 1.5 M NH_4Cl, 50 mM $MgCl_2$, and 40 mM Tris-HCl (pH 7.5) at 30 °C. After renaturation in 0.5 M NH_4Cl, 5 mM Mg^{2+}, and 40 mM Tris-HCl (pH 7.5) as described above, self-splicing was initiated by adding NH_4Cl and/or $MgCl_2$ to the final concentrations. Aliquots (50 μL) were withdrawn at different times, and the reaction was terminated by extraction with phenol, chloroform, and isoamyl alcohol (phenol-CIA in a 25:24:1 ratio). LtrA-promoted splicing reactions were carried out with 20 nM ^{32}P-labeled precursor RNA and 200 nM LtrA protein in 1 mL of reaction medium containing 0.5 M NH_4Cl, 5 mM $MgCl_2$, and 40 mM Tris-HCl (pH 7.5) at 30 °C. After renaturation in the reaction medium, splicing was initiated by adding LtrA protein, aliquots were withdrawn at different times, and the reaction was terminated, as described above. The products were then ethanol precipitated and analyzed in a denaturing 4% polyacrylamide gel containing TBE [90 mM Tris, 90 mM boric acid, and 2 mM EDTA (pH 8.2)], which was dried and quantitated with a PhosphorImager (Amersham, Piscataway, NJ). The spliced fraction was calculated as the ratio of splicing products to total RNA in each lane (*27*), and the data were best fit to equations with one or two exponentials using KaleidaGraph (Synergy Software, Reading, PA) to obtain k_1 and k_2 (*4*).

To test splicing of cross-linked RNAs, 2 μM ^{32}P-labeled wild-type Ll.LtrB RNA was irradiated, and the individual cross-linked RNA species were gel-purified (see below); ~2 pmol (20 nM) of each cross-linked RNA (concentration determined by the specific activity) was renatured by heating to 50 °C in 100 μL of 0.5 M NH_4Cl, 5 mM $MgCl_2$, and 40 mM Tris-HCl (pH 7.5) and slowly cooling to 30 °C. The samples were then split and incubated in the presence or absence of 200 nM LtrA protein for 30 min at 30 °C. The reaction was terminated by phenol-CIA extraction and ethanol precipitation, and the products were analyzed in a 4% polyacrylamide gel (19:1 acrylamide:bisacrylamide ratio) containing 8.3 M urea and BTBE buffer [30 mM bis(2-hydroxyethyl)iminotris(hydroxymethyl)methane, 30 mM boric acid, and 2.5 mM EDTA (pH 6.8)] (*29*).

UV Cross-Linking. Analytical UV cross-linking was carried out with 20 nM RNA in the presence or absence of 50 nM LtrA protein in 100 μL of reaction medium containing 0.5 M NH_4Cl, 40 mM Tris-HCl (pH 7.5), and different $MgCl_2$ concentrations. For preparative purposes, scaled-up cross-linking reactions used 2 μM RNA by itself or 200 nM RNA with 500 nM LtrA. For samples containing LtrA, the protein was incubated with the RNA for 15 min at 30 °C to form RNP complexes prior to cross-linking. Samples were then cross-linked for 5 min in a stirred quartz cuvette (4 mm × 4 mm × 30 mm, QS 0.500, Helma, Mulheim, Germany) with a hand-held UV lamp (two 8 W bulbs, $\lambda_{max} = 312$ nm; Spectroline, Westbury, NY) at a distance of 1 cm. Immediately after irradiation, $^1/_{10}$ of the final sample volume each of 0.5 M EDTA, 10% SDS, and proteinase K (10 mg/mL) were added, and the samples were incubated at 37 °C for 30 min, then extracted with phenol-CIA, back-extracted, and ethanol precipitated. Cross-linked RNAs were analyzed in a 4% polyacrylamide gel (19:1 acrylamide:bisacrylamide ratio) containing 8.3 M urea and BTBE (see above).

Primer Extension Mapping of Intramolecular RNA Cross-Links. To map cross-linking sites, labeled cross-linked Ll.LtrB RNAs were isolated from preparative gel slices by ultracentrifugation through cushions containing 2 M CsCl and 0.2 M EDTA (pH 7.4) at 40 000 rpm for 12 h (*33*), then dissolved in 400 μL of distilled water, extracted with phenol-CIA, and ethanol precipitated. The cross-linked nucleotide residues were mapped by primer extension with avian myeloblastosis virus (AMV) RT (Life Sciences, St. Petersburg, FL), using nine different 20-mer DNA primers complementary to different regions of the Ll.LtrB intron and 3′-exon. Cross-links were identified as strong reverse transcription stops one base prior to the cross-linked nucleotide residue compared to control lanes containing unirradiated RNA or the irradiated but un-cross-linked parent (P) RNA band. In cases where band diffusion or RNA degradation during preparative experiments resulted in reverse transcription stops in more than one fraction, the cross-link was assigned to the fraction containing the highest-intensity stop. In one case where a double stop was seen, the cross-linked nucleotide residues were assigned according to the position of the first reverse transcription stop.

Table 1: Kinetic Analysis of *in Vitro* Splicing of Ll.LtrB RNA under Different Conditions[a]

conditions	LtrA	k_1 (min^{-1})	k_2 (min^{-1})
0.5 M NH_4Cl and 5 mM Mg^{2+}	−	<0.001	
0.5 M NH_4Cl and 5 mM Mg^{2+}	+	2.9 (75%)	0.18 (11%)
0.5 M NH_4Cl and 50 mM Mg^{2+}	−	0.10 (70%)	
0.5 M NH_4Cl and 50 mM Mg^{2+}	+	0.38 (49%)	0.04 (19%)
1.5 M NH_4Cl and 50 mM Mg^{2+}	−	0.09 (72%)	

[a] Splicing time courses were determined with 20 nM ^{32}P-labeled Ll.LtrB RNA in the presence or absence of 200 nM LtrA in the indicated reaction medium at 30 °C. k_1 and k_2 were derived from curves best fit to either one- or two-exponential equations.

RESULTS

Experimental Strategy and Kinetic Analysis of Protein-Dependent and Self-Splicing under UV Cross-Linking Conditions. The experimental strategy was to UV cross-link Ll.LtrB RNA under protein-dependent and self-splicing conditions in the presence or absence of LtrA protein to identify cross-links that might be correlated with the acquisition of splicing activity. The Ll.LtrB RNA used in this work contains a 902-nucleotide Ll.LtrB-ΔORF intron, with the flanking 5′- and 3′-exons shortened to 32 and 37 nucleotides, respectively, to minimize extraneous cross-links. Table 1 summarizes kinetic analysis of the splicing of this Ll.LtrB RNA with a 10-fold molar excess of LtrA protein. The reactions were carried out at 30 °C under previously established protein-dependent (0.5 M NH_4Cl and 5 mM Mg^{2+}) and self-splicing conditions (0.5 or 1.5 M NH_4Cl and 50 mM Mg^{2+}) (*4*, *27*). As discussed previously, 0.5 M NH_4Cl is required in the reaction medium for optimal activity of

the LtrA protein, which aggregates and binds nonspecifically under lower-salt conditions (*4*).

First, under protein-dependent conditions (0.5 mM NH_4Cl and 5 mM Mg^{2+}), self-splicing of the wild-type Ll.LtrB RNA was just barely detectable (see Figure 5 below), but the intron spliced efficiently in the presence of the LtrA protein. As reported previously, the splicing reaction in the presence of excess LtrA was biphasic (*4*), with 75% of the RNA reacting at a rate of 2.9 min^{-1} (k_1), 11% at a rate of 0.18 min^{-1} (k_2), and the remainder inactive, presumably reflecting the presence of different RNA conformers, which fold at different rates. In different experiments, the proportion of intron RNA splicing at the fast rate ranged from 75 to 86%, consistent with a homogeneous active structure for this proportion of the RNA. We note that the behavior of the Ll.LtrB with shorter exons differed from that with longer exons, which was not detectably self-splicing under these conditions (confirmed in parallel splicing assays, not shown). It was found previously that exon length could significantly affect the self-splicing rate of group II introns (*34*).

At higher Mg^{2+} concentrations (0.5 M NH_4Cl and 50 mM Mg^{2+}), the Ll.LtrB RNA self-spliced at an appreciable but still relatively slow rate ($k_1 = 0.10$ min^{-1}), which could be increased by adding LtrA protein ($k_1 = 0.38$ min^{-1}). Significantly, the rate of LtrA-promoted splicing at 50 mM Mg^{2+} was ~7-fold slower than at 5 mM Mg^{2+}, and the proportion of spliceable RNA was lower, possibly reflecting the decreased conformational flexibility of the RNA and/or stabilization of non-native structures at the higher Mg^{2+} concentration (see Discussion). In reaction medium containing 1.5 M NH_4Cl and 50 mM Mg^{2+}, previously determined to be optimal for self-splicing (*27*), the reaction still occurred with a k_1 of ~0.1 min^{-1}, ~30-fold slower than the rate of protein-dependent splicing at 5 mM Mg^{2+}.

For all conditions that were tested, the protein-promoted splicing rates of the Ll.LtrB RNA with shorter 5′- and 3′-exons were within 2-fold of those for the Ll.LtrB RNA with longer exons, but the self-splicing rates were generally 4−5-fold faster (*27*). As in the previous RNA footprinting study, we used a branch-point deletion mutant Ll.LtrB-ΔA2486 when necessary to prevent RNA splicing during the incubations (*27*). This branch-point deletion mutation, which presumably traps the RNA structure just prior to the first step of splicing, does not affect the binding affinity for the LtrA protein and is the only one of the DV or DVI mutations that were tested that completely blocks splicing (*27*, *35*).

UV Cross-Linking of Wild-Type Ll.LtrB and Ll.LtrB-ΔA2486 RNAs in Reaction Medium Containing 5 mM Mg^{2+}. We first compared the UV cross-linking patterns of the wild-type Ll.LtrB and Ll.LtrB-ΔA2486 RNAs at room temperature (RT), 30 °C, and 37 °C under protein-dependent conditions (0.5 M NH_4Cl and 5 mM Mg^{2+}) in the absence of LtrA protein so that the wild-type RNA could not splice during the incubations. The RNAs were renatured and UV irradiated at relatively low intensity (~1.5 W/cm^2) in a stirred cuvette for 5 min, and the products were analyzed in a denaturing 4% polyacrylamide gel, which separates the molecules on the basis of the size and conformation of the covalently closed cross-linked loops. As shown in Figure 1, the pattern of cross-linked bands was essentially the same for the wild-type Ll.LtrB and Ll.LtrB-ΔA2486 RNAs and did not change markedly for either RNA at different

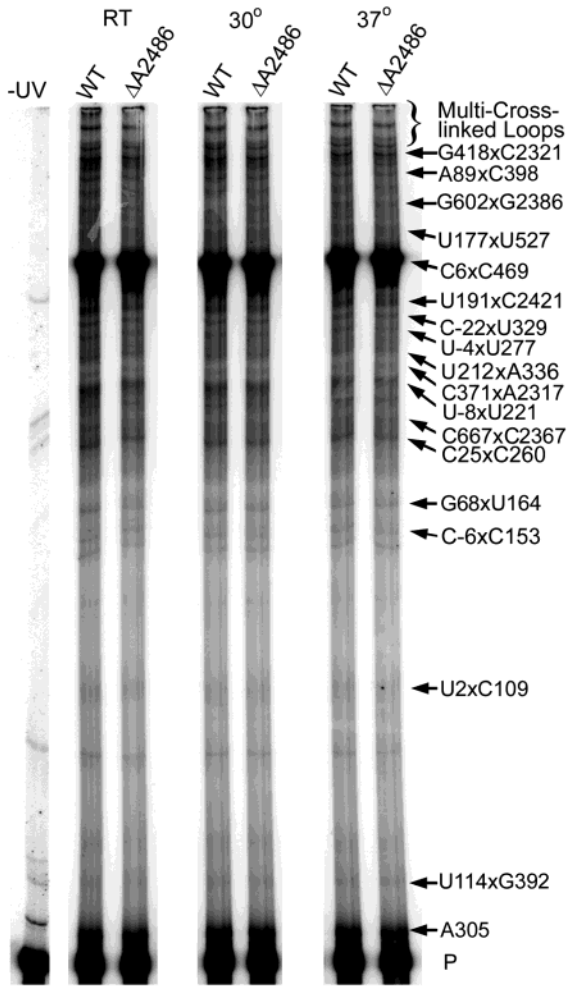

FIGURE 1: UV cross-linking patterns of wild-type (WT) and Ll.LtrB-ΔA2386 RNAs. The 5′-labeled Ll.LtrB RNAs were UV irradiated in reaction medium containing 0.5 NH_4Cl and 5 mM Mg^{2+} at room temperature (RT), 30 °C, and 37 °C and analyzed in a denaturing 4% polyacrylamide gel, which was dried and scanned with a phosphorimager. Mapped cross-links are identified to the right of the 37 °C gel. Unirradiated (−UV) RNA is shown in the leftmost lane. P denotes the irradiated but un-cross-linked parent RNA band at the bottom of the gel.

temperatures. In separate experiments, we confirmed that all the cross-linked products comigrated with unirradiated monomer RNA in a 2% agarose gel, indicating that their decreased electrophoretic mobility in polyacrylamide gels results from intramolecular and not intermolecular cross-links (not shown). As discussed below, the same cross-linked bands were also observed at higher Mg^{2+} concentrations and in the presence of the LtrA protein, but their relative frequencies differed, reflecting tertiary structure stabilization under these conditions.

Primer Extension Mapping of Cross-Linked Nucleotide Residues. To map the cross-links, RNA bands were isolated from preparative gel fractions, and the cross-linked nucleotide residues were identified as strong stops to AMV RT-catalyzed primer extension, using a series of 20-mer DNA primers complementary to different regions of the intron and 3′-exon. The presence of a cross-link generally arrests primer extension one base prior to the cross-linked nucleotide residue. The primer extension mapping identified 17 different cross-links, of which mapping data for 14 are shown in Figure 2. The identified cross-links are indicated to the right

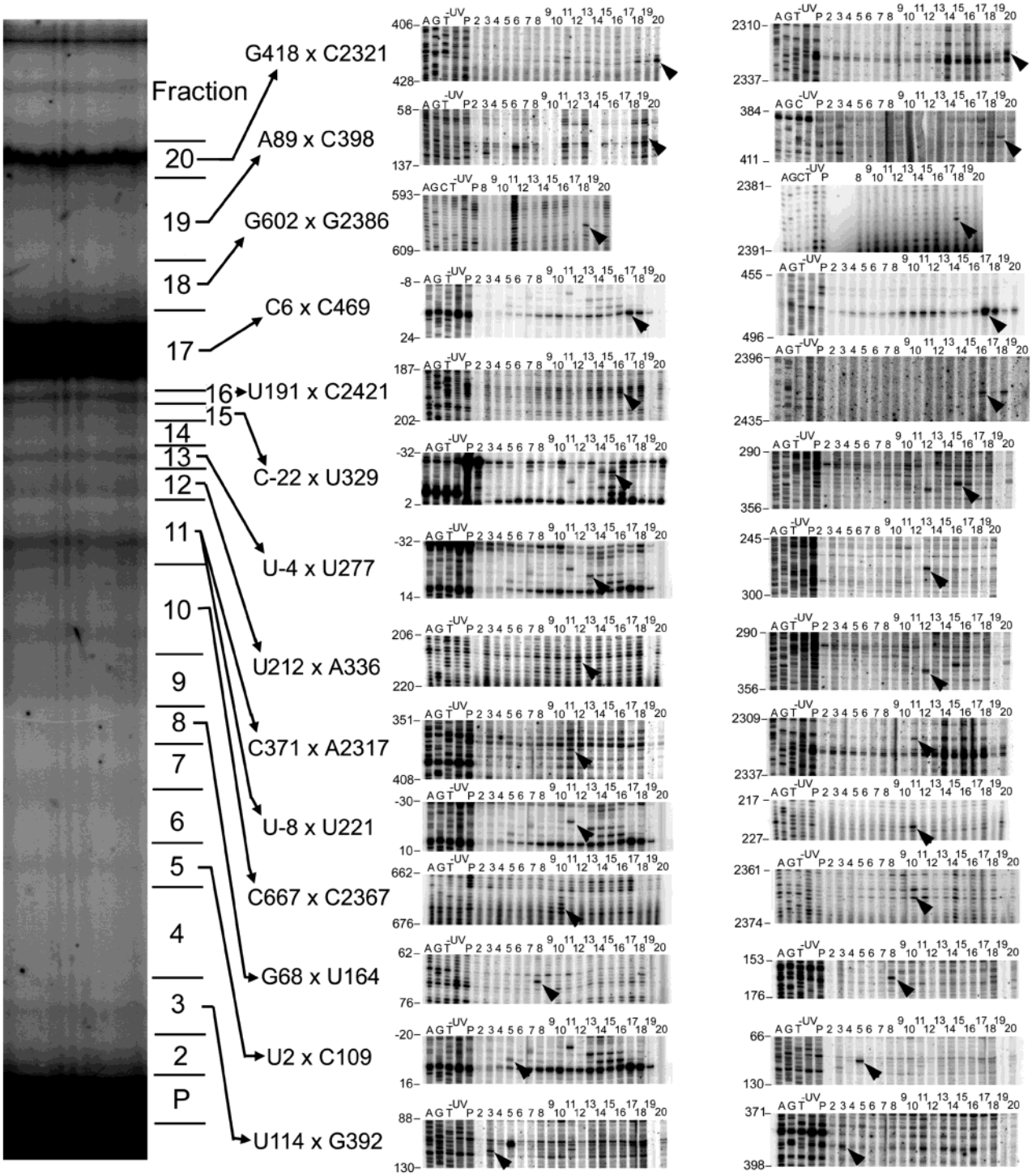

FIGURE 2: Primer extension mapping of cross-linked nucleotide residues. Scaled-up UV cross-linking reactions were carried out with Ll.LtrB-ΔA2486 RNA in reaction medium containing 0.5 M NH_4Cl and 5 mM Mg^{2+} at 30 °C, and bands corresponding to cross-linked RNAs were isolated from a preparative gel, a segment of which is shown to the left along with the fraction boundaries. The panels to the right show primer extension mapping of the cross-linked nucleotide residues with AMV RT, using a series of 20-mer DNA primers complementary to different regions of the RNA. Cross-linked nucleotide residues were identified as strong reverse transcription stops (black arrowheads) with relative intensities substantially higher than that seen in the unirradiated (−UV) control or the irradiated but un-cross-linked parent (P) RNA from the bottom of the gel. In most cases, the cross-links could be readily assigned to two prominent primer extension stops that appear in the same gel fraction. Fraction 11 contains two different cross-links, which were assigned on the basis of the relative intensities of the paired primer extension stops. In cases where the same stop was seen in multiple fractions due to band diffusion or RNA degradation during preparative experiments (C-22, U-4, C6, G418, C469, and C2421), the cross-link was assigned to the fraction containing the highest-intensity stop. In one case (C667xC2367), a double stop was seen for both sides of the cross-link, and the cross-linked nucleotide residues were assigned according to the position of the first reverse transcription stop. In some cases, the same primer extension stops reappear in higher-molecular weight fractions due to the presence of multiply cross-linked molecules. The identification of each cross-link was confirmed in five independent experiments similar to the one shown here. Lanes A, G, C, and T show dideoxy sequencing ladders obtained with unirradiated RNA. Nucleotide positions of the cross-links in the Ll.LtrB intron sequence (*24*) are indicated to the right of the preparative gel panel.

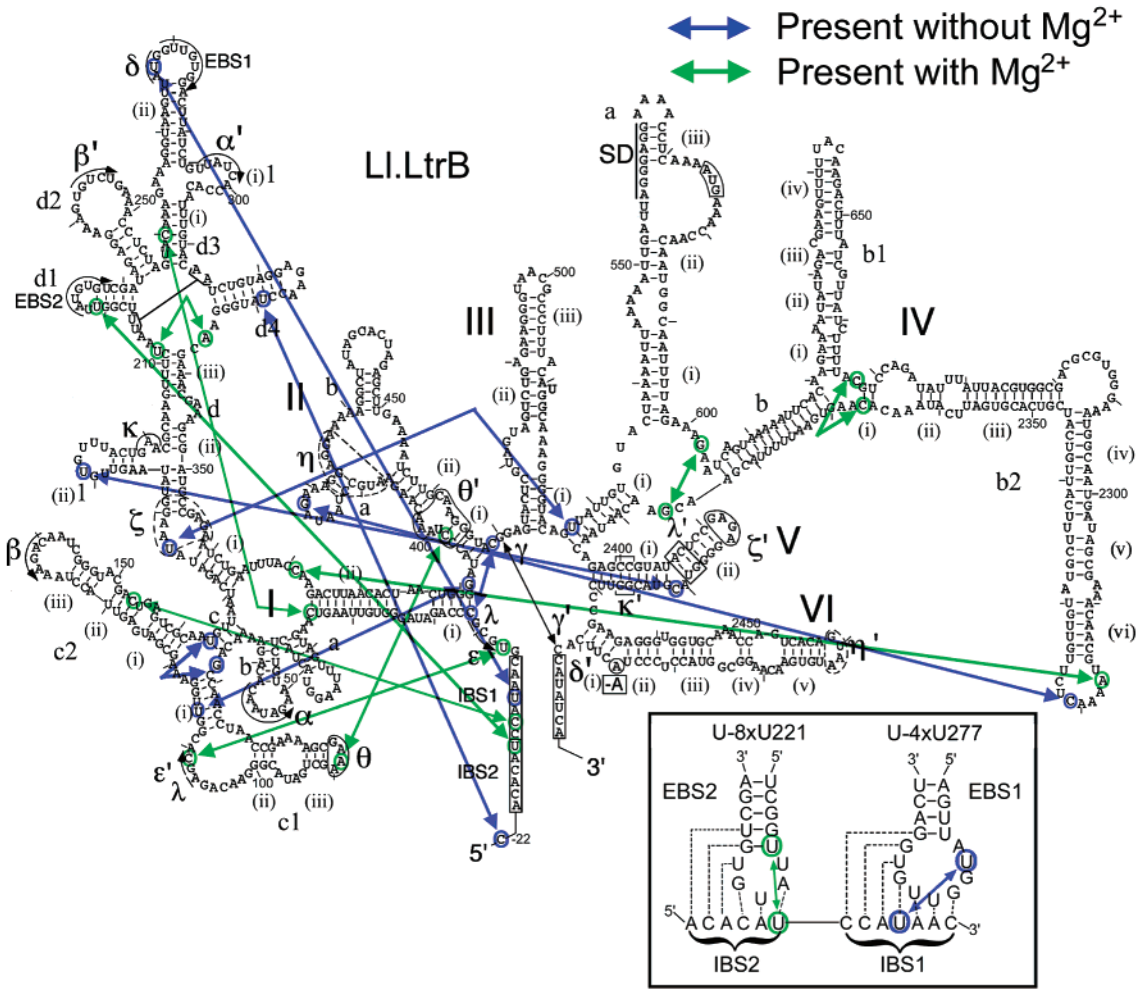

FIGURE 3: Secondary structure model of the Ll.LtrB-ΔORF intron and flanking exons showing cross-links mapped by primer extension. Blue arrows represent cross-links detected in the absence of Mg^{2+} or presence of EDTA, and green arrows represent cross-links detected only in the presence of Mg^{2+} (see Figure 5). SD is the Shine-Dalgarno sequence for the LtrA ORF, and the boxed AUG is the initiation codon (*28*). Greek letters indicate predicted tertiary structure interaction sites, with the dashed lines delineating *η*, *η*′, and *ζ* indicating previously discussed uncertainties in those assignments (*27*). The branch-point nucleotide residue, A2386, is circled, and its deletion in Ll.LtrB-ΔA2486 is denoted with a boxed −A. The inset at the bottom shows the IBS1−EBS1 and IBS2−EBS2 interactions between the intron and flanking 5′-exon sequences, with the U-4xU277 and U-8xU221 cross-links. Potential base pairing interactions between the intron and 5′-exon are represented by dashed lines.

of the gel in Figure 1 and depicted schematically on the secondary structure model of Ll.LtrB RNA in Figure 3. The characteristics of all the cross-links analyzed in this study are summarized in Table 2.

The most abundant cross-link was the C6xC469 cross-link between central wheel positions at the 5′-base of DI and just 3′ of DII. This cross-link was present in >50% of the irradiated RNA at 5 mM Mg^{2+}, presumably reflecting a highly favored RNA conformation. As discussed below, this cross-link suggests models in which the DI and DII helices are either coaxially stacked or aligned in parallel in arrangements that would bring both the 5′- and 3′-splice sites together with catalytic core elements in the proximity of JII/III (see Discussion). There were four additional short-range cross-links at three-way junctions [G68xU164 in DIc, U212xA336 in DId(iii), G602xG2386 in DIV, and C667xC2367 in DIVb] and 12 long-range cross-links, six of which involve nucleotide residues at or near known tertiary interaction sites. The latter are U-8xU221 (IBS2−EBS2), U-4xU277 (IBS1−EBS1), U2xC109 (*ϵ*−*ϵ*′), U114xG392 (*λ*−*λ*′), A89xC398 (*θ*−*θ*′), and U191xC2421 (*κ*−*κ*′). The IBS1−EBS1 and IBS2−EBS2 cross-links are similar in that a U residue in the base-paired 5′-exon region is cross-linked to a distal U residue in the intron stem−loop structure, suggesting a compact structure that brings the cross-linked U residues into proximity (Figure 3, inset).

The six remaining cross-links are between positions not previously known to be in proximity in the three-dimensional structure: C-22xU329, between the 5′-exon and DId4; C-6xC153, between IBS1 and DIc2; C25xC260, between DI(ii) and DId3; C371xA2317, between DI and DIVb2; G418xC2321, between DIIa and DIVb2; and U177xU527, between DI and the base of DIV. The bands near the top of the gel contain multiply cross-linked molecules, typically one of the other identified cross-links, in combination with the prominent C6xC469 central wheel cross-link, which is present in >50% of the molecules. We also identified one orphan primer extension stop (A305), just above the parent RNA band, for which no partner could be found. The rapid migration of this RNA suggests a very small cross-linked loop (<50 nucleotides), making the cross-linked partner difficult to identify by bracketed primer extension. Because the formation of a UV cross-link requires a favorable geometry and distance between suitably reactive moieties,

Table 2: Intramolecular RNA Cross-Links Identified in Ll.LtrB

cross-link	location	interaction	Mg^{2+} [a]	LtrA[b]	frequency (%)[c]
C-22xU329	5′-exon−DId4		↑	↑	0.4
U-8xU221	IBS2−DId1	IBS2−EBS2	↑	↑	0.4
C-6xC153	IBS1−DIc2		↑	−	<0.1
U-4xU277	IBS1−DId3	IBS1−EBS1	↑	↑	0.5
U2xC109	DI−DIc1	$\epsilon-\epsilon'$	↑	−	0.1
C6xC469	DI		↑	−	58
C25xC260	DI(ii)−DId3		↑	↓	<0.1
G68xU164	DIc		−	−	0.1
A89xC398	DIc1−DII	$\theta-\theta'$	↑	−	0.3
U114xG392	DIc1	$\lambda-\lambda'$	↓	−	0.1
U177xU527	DI−DIV		↓	↑	0.6
U191xC2421	DId(ii)1−DV	$\kappa-\kappa'$	↑	↑	2.0
U212xA336	DId(iii)		↑	↑	0.1
C371xA2317	DI−DIVb2		↑	↑	0.1
G418xC2321	DIIa−DIVb2		−	↑	1.0
G602xG2386	DIV		↑	↑	0.5
C667xC2367	DIVb(i)		↑	−	0.1

[a] ↓, ↑, and − indicate the frequency increased to maximum between 4 and 50 mM Mg^{2+}, decreased, and did not change with an increase in the Mg^{2+} concentration, respectively. [b] ↓, ↑, and − indicate the frequency increased, decreased, and did not change after LtrA binding at 5 mM Mg^{2+}, respectively. [c] Frequencies of cross-links in reaction medium containing 5 mM Mg^{2+}, calculated as the percentage of total RNA in the gel lane (Figures 5 and 6).

the failure to detect a cross-link between two structures does not necessarily imply the absence of an interaction. In addition, the procedures used here would not detect cross-links resulting in a small loop size, e.g., a cross-link involving an E motif pointed out by an anonymous reviewer in DIII (*36*, *37*).

Splicing Activity of the Cross-Linked RNA Molecules. To test whether the cross-links capture active RNA conformations, internally labeled wild-type Ll.LtrB RNA was UV cross-linked in 0.5 M NH_4Cl and 5 mM Mg^{2+} in the absence of LtrA protein so that minimal splicing would occur during cross-linking. The cross-linked RNA bands were then excised from a preparative gel and incubated in the same reaction medium at 30 °C with excess LtrA protein to promote splicing. As shown in Figure 4, each of 16 tested RNA bands gave rise to splicing products when incubated with LtrA protein (only one band, C-6xC153, was not recovered in sufficient yield to be tested). Although the splicing efficiency of the excised bands was generally low, in each case we detected a product corresponding to the ligated exons, indicative of both steps of splicing (bottom gel), as well as a prominent product that migrated either faster or slower than un-cross-linked lariat RNA and presumably corresponds to lariat RNA containing the cross-link (top gel). (The faster migrating lariat products include those containing the abundant C6xC469 cross-link.) We confirmed separately that UV irradiation of purified lariat RNA yields products that migrate with similar mobilities, either faster or slower than the un-cross-linked lariat RNA (not shown). In addition, all three cross-linked species found by primer extension to contain a cross-link between the intron and 5′-exon gave an additional larger splicing product, which could correspond to the spliced intron lariat RNA with the ligated exons still attached via the cross-link (albeit with different effects on electrophoretic mobility; lanes 8, 11, and 12). Considered together, these results suggest that most if not all of the identified cross-links reflect at least a partially active RNA conformation. The relatively inefficient splicing of the cross-linked RNA species could reflect the fact that the cross-links decrease the conformational flexibility required for catalysis and/or impede RNA refolding after purification in a denaturing gel.

Effect of Mg^{2+} on the Pattern and Frequency of Cross-Links. To follow progressive changes in RNA structure between protein-dependent and self-splicing conditions, we carried out UV cross-linking of Ll.LtrB-ΔA2486 RNA in reaction medium containing 0.5 M NH_4Cl and increasing concentrations of Mg^{2+}. An additional sample was cross-linked under the optimal self-splicing conditions (1.5 M NH_4Cl and 50 mM Mg^{2+}) to assess the effect of higher monovalent cation concentrations. Figure 5 shows a representative gel for samples incubated at 30 °C, and Figure 6 plots the relative frequency of each cross-link based on quantitation of that gel. The degree of splicing of wild-type Ll.LtrB RNA in the same reaction medium is shown below the gel. In each case, we confirmed by primer extension mapping of isolated bands that the cross-linking sites at 50 mM Mg^{2+} were the same as those at 5 mM Mg^{2+} (not shown). We also obtained similar results at room temperature and 37 °C (not shown).

Surprisingly, eight of the cross-links were present even in the absence of Mg^{2+} or in reaction medium containing 1 mM EDTA in place of Mg^{2+} (blue arrows in Figures 3 and 5). These include the prominent C6xC469 central wheel cross-link and cross-links U114xG392 ($\lambda-\lambda'$), G68xU164 (DIc the three-way junction), U177xU527 (DI−DIV), G418xC2321 (DIIa−DIVb2), U191xC2421 ($\kappa-\kappa'$), C-22xU329 (5′-exon−DId4), and U-4xU277 (IBS1−EBS1). In five of eight cases, the sites of cross-linking in the presence of EDTA were confirmed by primer extension mapping of isolated bands, while in the remaining three cases (U191xC2421, C-22xU329, and U-4xU277), recovery from the gel was insufficient and the cross-links were assigned on the basis of comigration with bands in other gel lanes (see Figure 5). The detection of almost half of the cross-links in the absence of Mg^{2+} suggests that significant tertiary structure forms at least transiently under these conditions.

The remaining nine cross-links were detected only in the presence of Mg^{2+} (green arrows in Figures 3 and 5). Seven of these became faintly visible between 1 and 5 mM Mg^{2+} and relatively prominent at 5 mM Mg^{2+}, while the remaining two cross-links (C-6xC153 and C25xC260) became faintly visible only at 5 mM Mg^{2+}. The frequencies of all but four cross-links increased with increasing Mg^{2+} concentrations, becoming maximal between 4 and 50 mM Mg^{2+}, as expected for tertiary structure stabilization. The four exceptions were the G68xU164 (DIc three-way junction) and G418xC2321 (DIIa−DIVb2) cross-links, which did not appear to be strongly affected by increasing Mg^{2+} concentrations, and the U114xG392 (DIc1−central wheel) and U177xU527 (ζ−DIV) cross-links, the frequencies of which decreased with increasing Mg^{2+} concentrations. The frequency of an additional short-range cross-link C667xC2367 increased up to 4 mM Mg^{2+}, but decreased between 5 and 20 mM Mg^{2+}. Surprisingly, many of the cross-links were already near maximal at 5 mM Mg^{2+}, with relatively little further increase at 50 mM Mg^{2+}.

The C-6xC153 (IBS1−DIc2) and C25xC260 [DI(ii)−DId3] cross-links, which were barely detectable at 5 mM

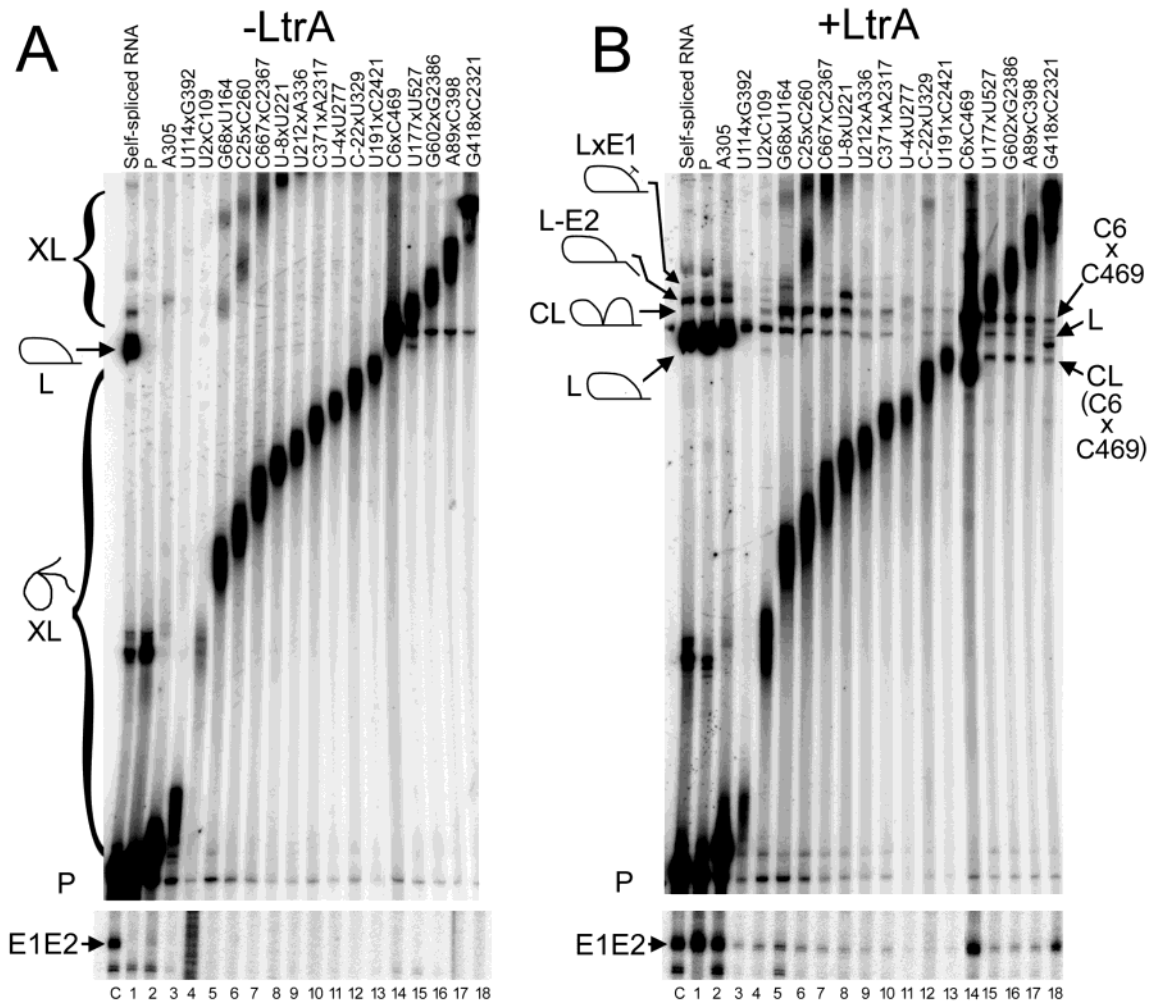

FIGURE 4: Splicing of isolated cross-linked RNA molecules in the presence of LtrA protein. Individual cross-linked bands obtained from wild-type Ll.ltrB RNA were isolated from preparative gels, renatured, and split into two portions, which were incubated in 0.5 M NH_4Cl, 5 mM $MgCl_2$, and 40 mM Tris-HCl (pH 7.5) for 30 min at 30 °C in the absence (A) or presence (B) of a 10-fold molar excess of LtrA protein. After phenol-CIA extraction and ethanol precipitation, the RNAs were analyzed in denaturing 4% (top) and 8% (bottom) polyacrylamide gels to detect formation of lariat RNA and ligated exons, respectively. Lane C contained unirradiated Ll.LtrB RNA self-spliced in 1.5 M NH_4Cl, 50 mM $MgCl_2$, and 40 mM Tris-HCl (pH 7.5). Lanes 1−18 contained products obtained after incubation of the irradiated parent (P) band and individual cross-linked bands as indicated at the top of the gel. Bands present in both the −LtrA and +LtrA panels were not considered spliced products. Arrows to the left and right of gel indicate bands corresponding to cross-linked Ll.LtrB RNA (XL), intron lariat RNA (L), cross-linked intron lariat RNA (CL), the lariat−3′-exon intermediate (L−E2), lariat with cross-linked 5′-exon (LxE1), and ligated exons (E1E2). The structures are depicted schematically to the left of each gel.

Mg^{2+} and the frequencies of which increased substantially at 50 mM Mg^{2+}, have the characteristics expected for interactions correlated with the acquisition of self-splicing activity. However, neither of these cross-links was enhanced by the binding of the LtrA protein at 5 mM Mg^{2+}, and it is possible that they reflect intermediate or aberrant structures that are stabilized by high Mg^{2+} concentrations. Several other unmapped cross-links (asterisks), which were present in the absence of Mg^{2+}, disappeared at higher Mg^{2+} concentrations, and thus may also reflect intermediate or aberrant tertiary structures that differ from the final catalytically active structure.

Together, the results described above indicate that most of the productive RNA structure monitored by UV cross-linking is formed at least transiently at 5 mM Mg^{2+} and that acquisition of self-splicing activity at higher Mg^{2+} concentrations reflects either a global increase in the stability of the RNA tertiary structure or additional interactions that are not detected readily by UV cross-linking.

Effect of the LtrA Protein on Cross-Link Frequency. Finally, to investigate how the binding of the LtrA protein affects RNA structure, we carried out UV cross-linking of the Ll.LtrB-ΔA2486 RNA in reaction medium containing 5 or 50 mM Mg^{2+} in the presence or absence of the protein (Figures 7 and 8). In addition to the gel that is shown, which is for samples cross-linked at 30 °C, similar results were obtained in repeated experiments at 4 and 25 °C (not shown). In each case, the identity of the cross-linked band was confirmed by primer extension mapping.

The results show that LtrA binding at 5 mM Mg^{2+} reproducibly increased the frequency of nine cross-links. The strongest increases were found for the G602xC2386 (DIV), U177xU527 (DI−DIV), and G418xC2321 (DIIa−DIVb) cross-links, all of which involve DIV, the location of LtrA's high-affinity binding site. None of these cross-links was comparably enhanced by high Mg^{2+} concentrations. LtrA binding also resulted in substantial increases for the U-4xU277 (EBS1−IBS1), U-8xU221 (EBS2−IBS2), and C-22xU329

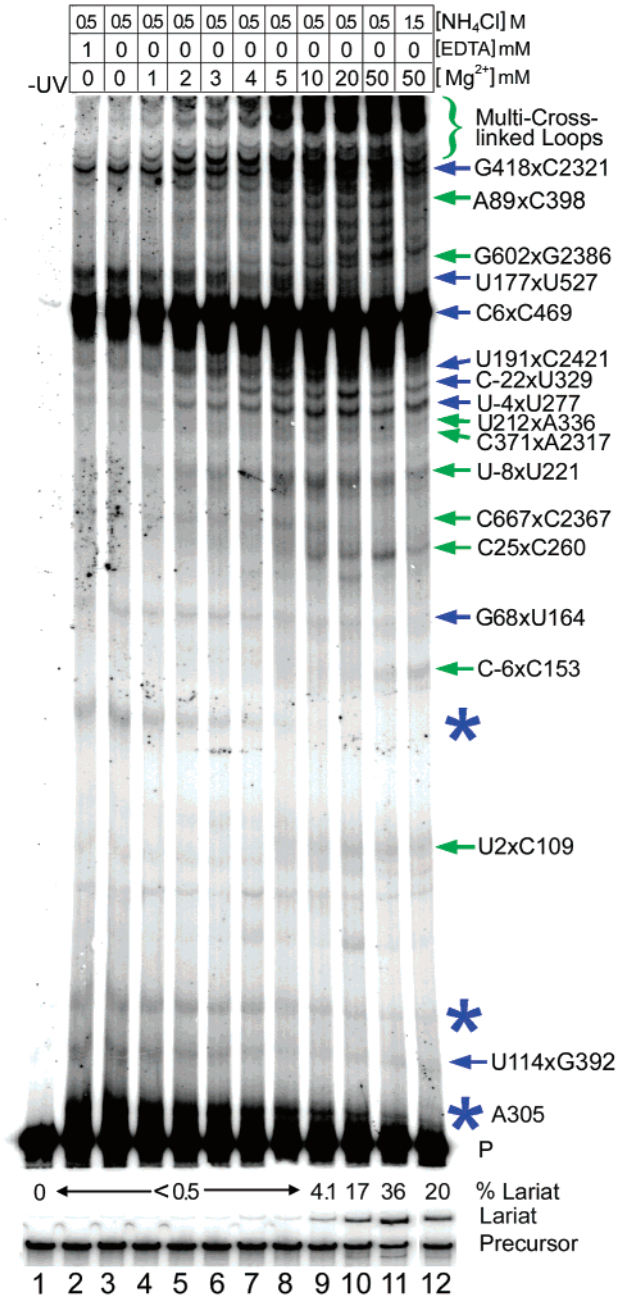

FIGURE 5: Effects of Mg^{2+} concentration on the pattern of intramolecular RNA−RNA cross-links. The 5′-labeled Ll.LtrB-ΔA2386 RNA was UV irradiated at 30 °C in the indicated reaction medium, and the cross-linked RNAs were analyzed in a denaturing 4% polyacrylamide gel, which was dried and scanned with a phosphorimager. Identified cross-links that are present in EDTA or in the absence of Mg^{2+} are denoted with blue arrows, and those detected only in the presence of Mg^{2+} are denoted with green arrows. Blue asterisks denote unidentified cross-links, whose frequencies decreased with an increase in the Mg^{2+} concentration. The lanes at the bottom show the extent of self-splicing of unirradiated wild-type Ll.LtrB RNA incubated in the same reaction medium for 30 min at 30 °C and then analyzed in a denaturing 4% polyacrylamide gel. The percentage of lariat RNA is indicated above each lane. The experiment was repeated three times with essentially the same results. Abbreviations are the same as those used in Figure 1.

(5′-exon−DId4) cross-links, which involve contacts between the 5′-exon and DI, and a smaller but highly reproducible increase for the U191xC2421 ($\kappa-\kappa'$) cross-link in three independent experiments. Frequency changes in other cross-links were not reproducible. The samples cross-linked in the presence of LtrA also gave a single orphan primer extension stop in fraction 11, which mapped to G357, a nucleotide residue located in tertiary structure element ζ (not shown). This cross-link could be related to the formation of the $\zeta-\zeta'$ interaction, with the failure to detect its partner being caused by a cross-link geometry that does not constitute a strong block to primer extension (cf. ref *29*). This orphan primer extension stop is unlikely to correspond to an RNA−protein cross-link, since the cross-linked RNAs were digested with proteinase K and thus are expected to contain only very short peptides (two or three amino acids) (*38*), which should not affect migration relative to the parent RNA band.

Notably, LtrA binding did not appreciably affect the frequency of a number of cross-links, including C6xC469 (the prominent central wheel cross-link), U2xC109 ($\epsilon-\epsilon'$), A89xC398 ($\theta-\theta'$), U114xG392 (DIc1−JI/II), C-6xC153 (IBS1−DId1), C68xU164 (DIc), and C667xC2367 [DIVb-(i) loop]. These cross-links presumably reflect parts of the RNA structure that are already sufficiently stable for splicing at 5 mM Mg^{2+}, but include three (C-6xC153, U2xC109, and A89xC398) that were enhanced significantly in the free RNA at higher Mg^{2+} concentrations (Figures 5 and 6 and Table 1). Moreover, a fourth cross-link {C25xC260 [DI(ii)−DId3]}, which was also strongly enhanced in the free RNA at high Mg^{2+} concentrations, was decreased strongly by LtrA binding at either low or high Mg^{2+} concentrations (Figures 7 and 8). This cross-link could reflect a competing, non-native structure of DI that is not present after LtrA binding, or it could be occluded directly by LtrA binding, a possibility consistent with its proximity to the potential LtrA interaction site DId4 (*27*). Together, these findings suggest that LtrA promotes splicing by stabilizing RNA tertiary structure interactions in some regions, while leaving other regions of the intron relatively flexible, potentially facilitating conformational changes required between the different steps of splicing.

DISCUSSION

In the work presented here, we identified 17 UV cross-links that are present in both the wild-type Ll.LtrB intron and the branch-point deletion mutant and used them to follow tertiary structure formation under different conditions. Surprisingly, eight of the cross-links were detected in absence of Mg^{2+}, and an additional seven became detectable between 1 and 5 mM Mg^{2+}, with only two (C-6xC153 and C25xC260) first becoming detectable at 5 mM Mg^{2+}. These findings suggest that the Ll.LtrB intron forms much of the tertiary structure required for catalytic activity at least transiently at 5 mM Mg^{2+}, even though the intron RNA does not self-splice efficiently under these conditions. LtrA binding and a higher Mg^{2+} concentration then act in somewhat different ways to provide additional structural stabilization required for catalytic activity. All but four of the cross-links increased in frequency with an increase in the Mg^{2+} concentration, reaching a maximum between 4 and 50 mM Mg^{2+}. Thus, a high Mg^{2+} concentration appears to globally stabilize the intron's tertiary structure, including key interactions required for self-splicing. In contrast, LtrA binding at 5 mM Mg^{2+} enhanced only a subset of cross-links, including U-4xU277 (IBS1−EBS1), U-8xU221 (IBS2−EBS2), and U191xC2421 ($\kappa-\kappa'$), along with several long-range cross-links involving

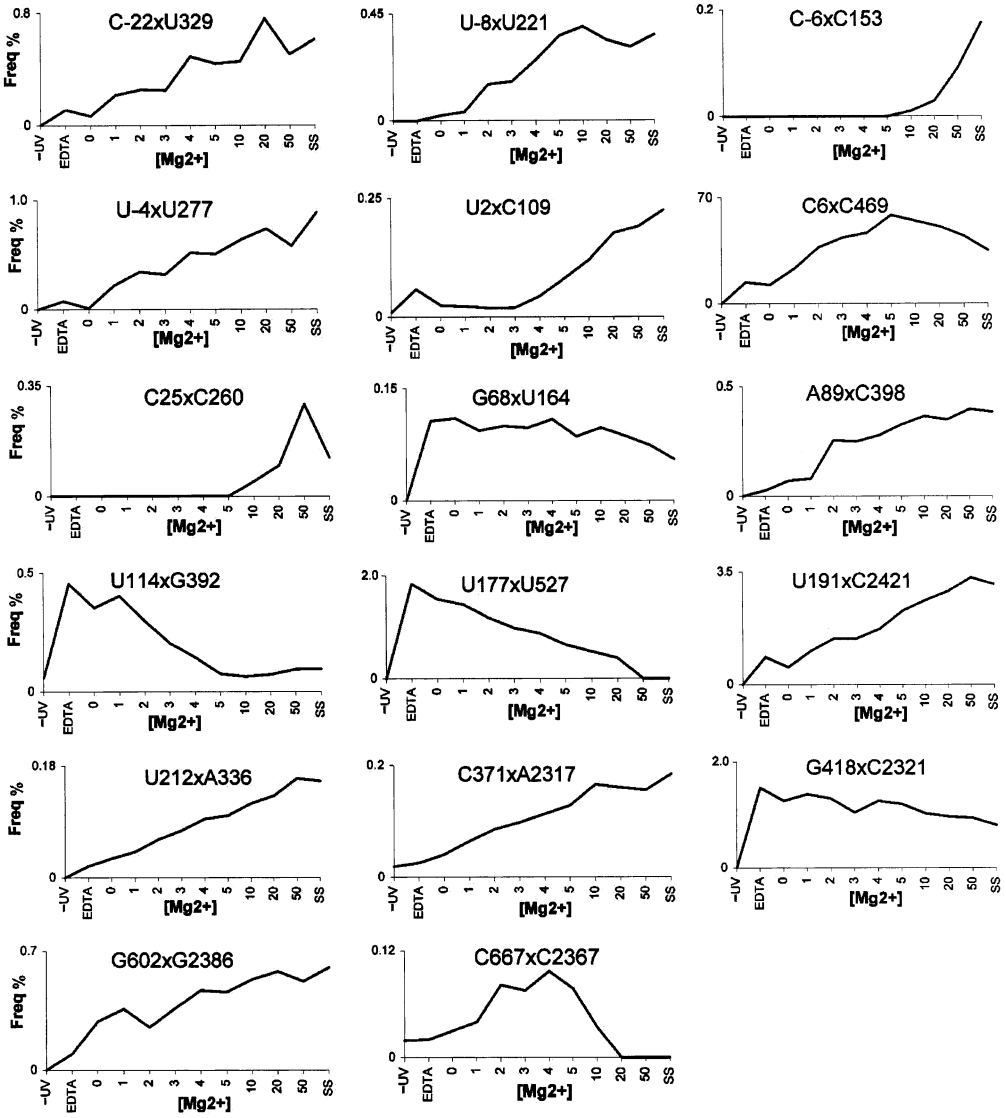

FIGURE 6: Plots showing the frequency of individual cross-links as a function of Mg^{2+} concentration. Frequency (percent of total RNA) was calculated by dividing the radioactivity in each cross-linked RNA band by the sum of that in all RNA bands in the gel of Figure 5, after subtraction of background radioactivity in adjacent, equivalent-sized fractions. −UV indicates unirradiated RNA, and EDTA indicates RNA irradiated in reaction medium containing 1 mM EDTA.

DIV, the location of LtrA's high-affinity binding site, without enhancing other cross-links that were enhanced by a high Mg^{2+} concentration. Thus, LtrA binding appears to stabilize critical tertiary structure interactions, while leaving other regions of the intron comformationally flexible.

The prominent C6xC469 central wheel cross-link, which was present in >50% of the UV-irradiated molecules at 5 mM Mg^{2+}, links a position at the 5′-base of DI to one just 3′ of DII. This cross-link suggests structural models in which DI and DII are either coaxially stacked or aligned in parallel, as shown in panels A and B of Figure 10, respectively. The stacking of DI and DII is expected to be energetically favorable (*39*) and would account for the very high frequency of the cross-link, even in the absence of Mg^{2+}. Both the stacked and parallel helices models lead to the confluence of several key tertiary interactions, including $\epsilon-\epsilon'$, $\theta-\theta'$, $\gamma-\gamma'$, and $\lambda-\lambda'$, and bring the 5′- and 3′-splice sites together with DV and DVI in the proximity of JII/III. Both models also accommodate the additional constraint imposed by the U114xG392 cross-link, which connects JII/III to a region of DI involved in the $\epsilon-\epsilon'$ and $\lambda-\lambda'$ interactions.

Analysis of a database consisting of 176 aligned group IIA introns showed significant covariation between central wheel nucleotide residues G392 and C469 (Table 3; nucleotide residues highlighted in yellow in Figure 10). In the stacked helices model, these nucleotide residues form a base pair that stacks between the DI−DII junction. This GC combination is conserved in 112 of 176 group IIA intron sequences (*40*) and is replaced in most of the remaining introns by other base pairs or by base mismatch combinations found to be favored at stacked helical junctions in known RNA secondary structures [CA, AA, GA, GG, and UU, albeit with GA underrepresented, possibly reflecting additional constraints (Table 3) (*39*)]. In addition, the adjoining unpaired residue A393 is conserved in 84.8% of group IIA introns and is U in an additional 14%, implying that this position is also highly constrained. We did not find similar conservation or covariation in group IIB introns, nor was an analogous central wheel cross-link observed in a previous UV cross-linking study of yeast group IIB intron aI5γ (*41*). However, the failure to observe this cross-link in aI5γ could reflect the possibility that the corresponding positions in group IIB

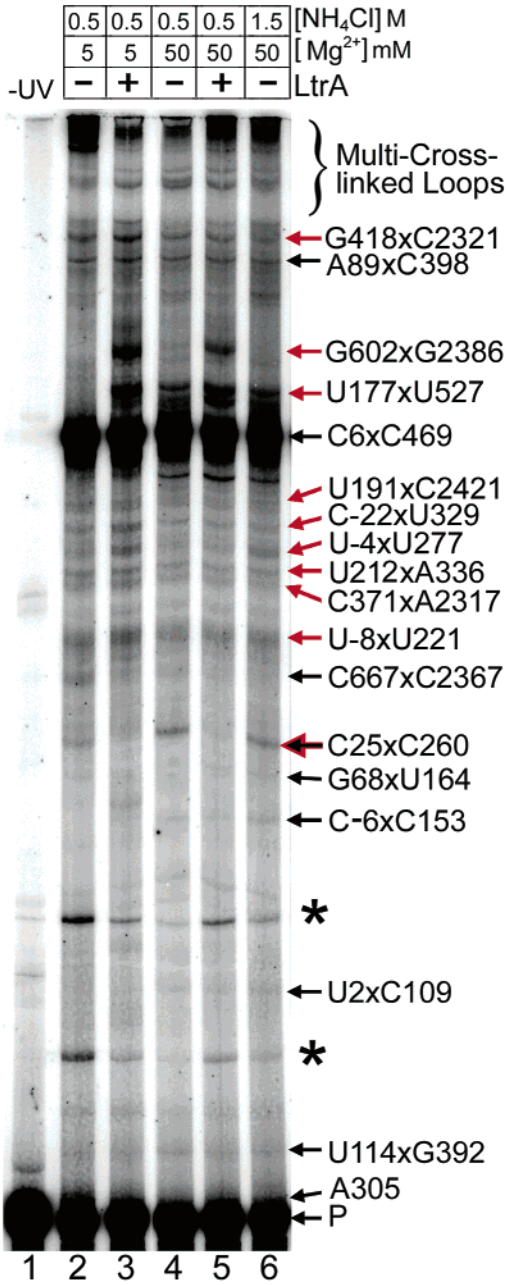

FIGURE 7: Effect of the LtrA protein on the pattern of intramolecular RNA−RNA cross-links. The 5′-labeled Ll.LtrB-ΔA2386 RNA was UV irradiated at 30 °C in the indicated reaction medium in the presence or absence of LtrA protein, and the cross-linked RNAs were analyzed in a denaturing 4% polyacrylamide gel, which was dried and scanned with a phosphorimager. Cross-links whose frequency increased reproducibly in the presence of LtrA protein in three independent experiments are denoted with red arrows; those whose frequency did not increase reproducibly in all three experiments are denoted with black arrows, and one that strongly decreased in frequency in the presence of LtrA protein is denoted with a two-tone black and red arrow. For each band that is indicated, increased frequency in the presence of LtrA was confirmed by mapping of primer extension stops corresponding to the cross-link (not shown). Asterisks denote bands that could not be reproduced in other experiments. Abbreviations are the same as those used in Figure 1.

introns are occupied by combinations of purine residues, which may not be amenable to efficient UV cross-linking.

Six of the cross-links were found within regions of the intron RNA known to be involved in tertiary structure interactions. These are the U-4xU277 (IBS1−EBS1), U-8xU221 (IBS2−EBS2), U2xC109 ($\epsilon-\epsilon'$), U114xG392 ($\lambda-\lambda'$), A89xC398 ($\theta-\theta'$), and U191xC2421 ($\kappa-\kappa'$) cross-links. The IBS1−EBS1 and IBS2−EBS2 base pairing interactions between two regions of DI and the 5′-exon are required for exon definition and positioning of the 5′-splice site at the intron's catalytic core. Each contains one cross-link between a base-paired U residue in the 5′-exon and a distal U residue in the intron stem−loop structure (Figure 3, inset), implying that both stem−loop structures can adopt a compact configuration that brings the cross-linked nucleotide residues into proximity. The IBS1−EBS1 and IBS2−EBS2 cross-links were enhanced both by high Mg^{2+} concentrations and by the binding of LtrA protein, presumably reflecting stabilization of the catalytically active structure (*42*). An additional cross-link (C-6xC153) was detected between IBS2 and DIc2(ii) and could reflect an interaction that stabilizes the EBS−IBS interaction.

Three cross-links were between nucleotide residues in DIc1 and regions near the base of DI and DII. Two of these confirm previously proposed interactions. The U2xC109 cross-link reflects the formation of the $\epsilon-\epsilon'$ interaction, a base-paired pseudoknot interaction, which places the cross-linked nucleotide residues diagonally opposite each other (Figures 3 and 10), and the A89xC398 cross-link involves nucleotide residues in the $\theta-\theta'$ interaction. The Ll.LtrB $\theta-\theta'$ interaction is a GAAA tetraloop−receptor interaction, analogous to one whose structure was determined as part of the *Tetrahymena thermophila* LSU intron P4−P6 domain (*43*). Using that structure as a model, the shortest distance between the adenine C8 and cytidine C5−C6 residues corresponding to the cross-linked nucleotide residues is 6.67 Å (Figure 11A), within the range of 12 UV-induced cross-links in *E. coli* 16S rRNA (average distance of 5.46 ± 1.42 Å) (*30*, *44*). The remaining cross-link (U114xG392) links DIc1 to JII/III and provided an additional constraint in the structure modeling of Figure 10. All three cross-links involving DIc1

Table 3: Covariation Analysis of G392 and C469[a]

		469 (%)				
		A	C	G	U	−
392 (%)	A	6	0	1	1	0
	C	10	0	0	0	0
	G	4	64	2	0	0
	U	6	0	3	2	1
	−	0	0	0	0	2

G392−C469	actual no. of sequences	predicted no. of sequences
GC	112	79
CA*	17	4.6
AA*	11	3.2
UA*	11	5.0
GA	7	32
UG*	5	1.2
GG	4	7.4
UU*	3	0.39
−	3	0.07

[a] The top section shows frequencies of different combinations of nucleotide residues at positions 392 and 469 in 176 group IIA introns (*40*). The bottom section shows the actual and predicted number of different combinations; only those combinations that occurred in three or more introns are shown. Asterisks denote significant covariation of a nucleotide pair. A "−" indicates the absence of a nucleotide at the specified position.

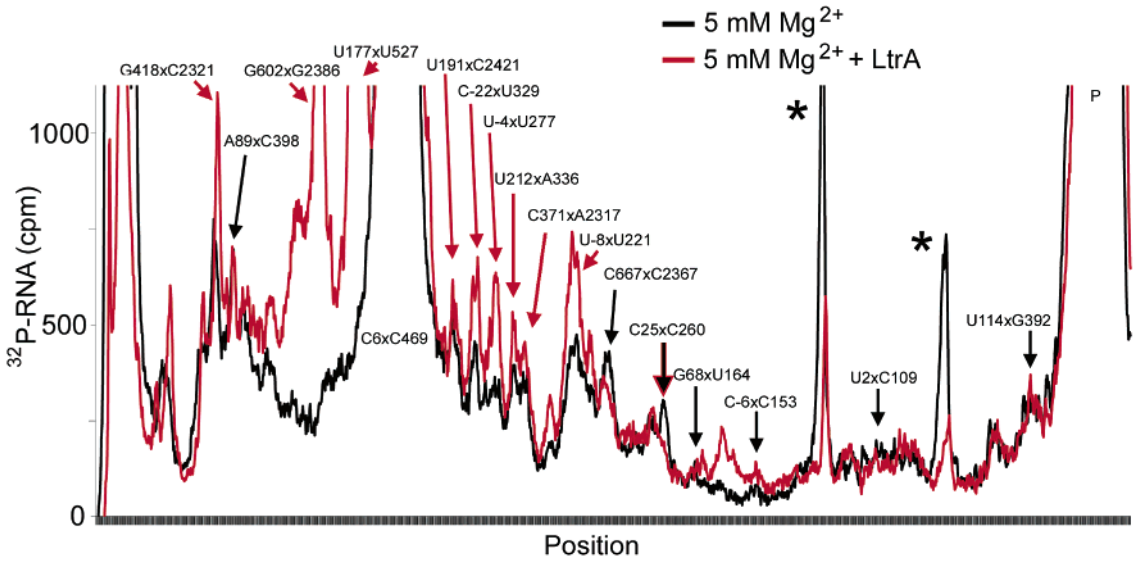

FIGURE 8: Phosphorimager scans showing the effect of LtrA binding on cross-link frequency. Scans showing the distribution of cross-linked ^{32}P-labeled RNAs in lane 2 of Figure 7 (5 mM Mg^{2+}, black) and lane 3 of Figure 7 (5 mM Mg^{2+} and LtrA, red) were superimposed. Peaks corresponding to cross-links whose frequency increased reproducibly upon LtrA binding in three independent experiments are denoted with red arrows; those whose frequency did not reproducibly increase in all three experiments are denoted with black arrows, and one that decreased in frequency upon LtrA binding is denoted with a two-tone black and red arrow. Cross-link frequency was determined by the area under the peak after subtraction of background. Asterisks indicate bands that could not be reproduced in other experiments. The experiment was repeated three times with similar results.

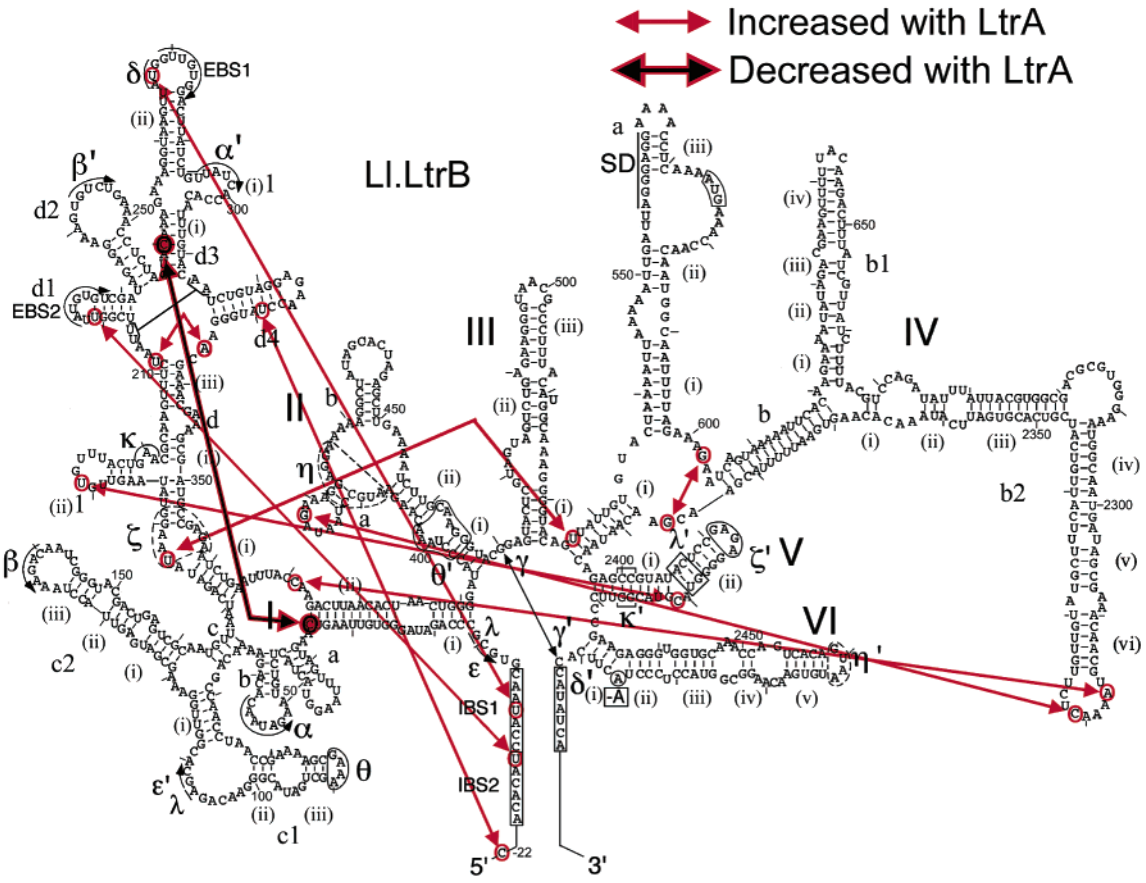

FIGURE 9: Secondary structure model of the Ll.LtrB-ΔORF intron and flanking exons showing UV cross-links, whose frequency increased significantly in the presence of LtrA protein. This figure is based on data in Figures 7 and 8. Cross-links that increased in frequency after binding the LtrA protein are denoted with red arrows, and the C25xC260 cross-link, whose frequency decreased in the presence of LtrA, is denoted with a two-tone black and red arrow.

were detected at 5 mM Mg^{2+}, with the U2xC109 and A89xC398 cross-links becoming further enhanced at 50 mM Mg^{2+}, consistent with previous chemical modification data showing an increased frequency of protection for nucleotide residues involved in the $\epsilon-\epsilon'$, $\lambda-\lambda'$, and $\theta-\theta'$ interactions at high Mg^{2+} concentrations (*27*). However, none of these cross-links was affected significantly by LtrA binding, suggesting that the corresponding structural interactions are

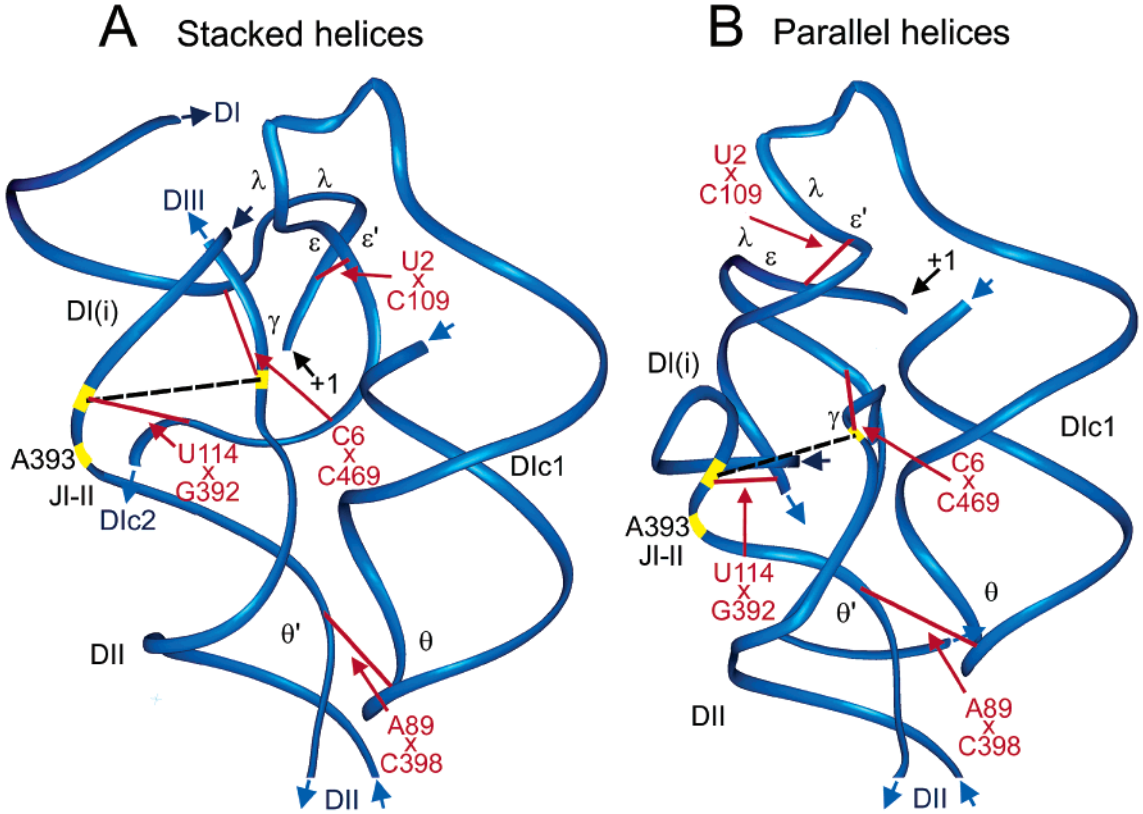

FIGURE 10: Models of stacked (A) and parallel (B) arrangements of Ll.LtrB DI(i) and DII. Helices were constructed using MANIP (*49*, *50*) and docked using Insight II (Accelrys, San Diego, CA). Cross-links and tertiary interactions ($\epsilon-\epsilon'$, $\theta-\theta'$, and $\lambda-\lambda'$) were used as constraints. The $\theta-\theta'$ tetraloop−receptor interaction was modeled on the basis of the X-ray crystal structure of a GAAA tetraloop−receptor interaction in the *T. thermophila* group I intron P4−P6 domain (*43*) (see Figure 11). Cross-links are shown in red. The conserved A393 and covarying nucleotide residues G392 and C469 are shown in yellow, with a dashed line denoting the covariation.

already sufficiently stable for RNA splicing at 5 mM Mg^{2+} and do not require further stabilization by the protein.

The U191xC2421 cross-link involves nucleotide residues near those involved in the $\kappa-\kappa'$ interaction between DI and DV, with C2421 corresponding to an unpaired nucleotide residue in the elbow region of DV (Figure 11B). The X-ray crystal structure of a 70-nucleotide RNA comprising DV and DVI of the aI5γ group II intron (*23*) shows that the corresponding C residue in the elbow region protrudes into the major groove of the DV helix, presumably making it accessible for cross-linking. A model based on known tertiary interactions of the yeast aI5γ group II intron catalytic core depicts the DV elbow region as being buried within the RNA's catalytic core, a helical diameter from the $\kappa-\kappa'$ interaction, and in the proximity of the $\lambda-\lambda'$ interaction between DI and DV (*13*). Our results suggest a similar orientation in which the longer Ll.LtrB κ stem−loop structure extends farther into the RNA core to dock within the DV major groove, placing U191 in proximity to cross-link to C2421. Chemical modification experiments indicate that the Ll.LtrB DV backbone is only moderately protected even at high Mg^{2+} concentrations (*27*), suggesting a more accessible Ll.LtrB core region for accommodating the docking of the κ stem−loop structure in the DV major groove. Docking of an internal loop analogous to κ into the distorted major groove of an RNA helix was observed in the crystal structure of the L.11 binding site in the *E. coli* large rRNA (*44*).

Three cross-links (U177xU527, C371xA2317, and G418xC2321) occurred between conserved regions of the catalytic core and different regions of DIV, which contains the high-affinity binding site for LtrA. All three of these cross-links were enhanced by LtrA binding, possibly reflecting bridging interactions between LtrA bound to DIV and the catalytic core. The G418xC2321 cross-link links DIVb2 to DIIa, which is itself a potential secondary binding site for LtrA (*27*). It was present in the absence of Mg^{2+}, and the frequency diminished with an increase in the Mg^{2+} concentration; the C371xA2317 cross-link, which links DI(ii) and DIVb2(iv), was enhanced by both an increase in the Mg^{2+} concentration and LtrA binding. Surprisingly, the U177xU527 cross-link, which was stimulated by LtrA binding, places the ζ element close to the base of DIV rather than its interacting partner ζ' in DV. This cross-link was seen without Mg^{2+}, and its frequency decreased at higher Mg^{2+} concentrations. It could reflect a transient interaction or non-native RNA structure, or that LtrA stabilizes a structure in which the base of DIV is brought close to the $\zeta-\zeta'$ interaction.

The four short-range cross-links that occurred at multihelix junctions [G68xU164 in DIc, U212xA336 in DId(iii), G602xG2386 in DIV, and C667xC2367 in DIVb] presumably reflect specific internal structures at these junctions. Two of these cross-links (U212xA336 and G602xG2386) were enhanced by both LtrA binding and high Mg^{2+} concentrations. Another cross-link (C-22xU329) links DId4, previously identified as a potential secondary binding site for LtrA (*27*), to a region of the 5′-exon that corresponds to part of the DNA target site recognized by LtrA for intron mobility (*46*). If these exon sequences can also be recognized in RNA, the enhancement of this cross-link by LtrA binding could reflect

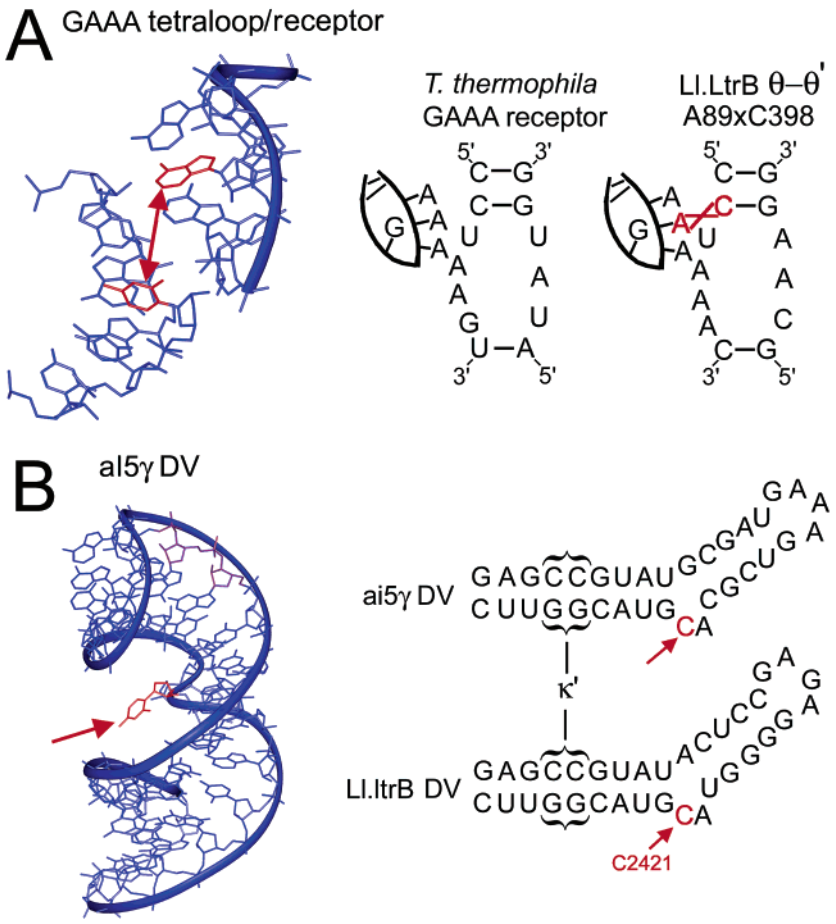

FIGURE 11: Correlation of Ll.LtrB UV cross-linking sites with X-ray crystal structures of analogous interactions or structures in other RNAs. (A) GAAA tetraloop−receptor interaction A89xC398 ($\theta-\theta'$). The left panel shows the X-ray crystal structure of the GAAA tetraloop and partial receptor in the *T. thermophila* group I intron P4−P6 domain (*43*), with the nucleotide residues corresponding to the A89xC398 cross-link highlighted in red and linked by a red arrow. (B) Bulged nucleotide residue in the elbow of DV involved in the U191xC2421 cross-link near the $\kappa-\kappa'$ interaction. The left panel shows the X-ray crystal structure of DV from an RNA containing DV and DVI of the yeast aI5γ intron (*23*), with the nucleotide residue corresponding to C2421 in Ll.LtrB RNA highlighted in red. The secondary structures of the aI5γ and Ll.LtrB DVs are compared at the right.

direct interaction of both sides of the cross-link with the protein.

The previous RNA structure mapping experiments suggested that the Ll.LtrB intron could form most of the conserved secondary structure and some tertiary structures in DI and DII in reaction medium containing 0.5 M NH_4Cl and 5 mM Mg^{2+} in the absence of the LtrA protein (*27*). However, the cross-links identified here go beyond DI and DII and include a number of long-range interactions. Because of the limited half-life of the excited state of a nucleotide (several hundred nanoseconds to a millisecond) (*47*) and the short time scales necessary for formation of a covalent cross-link, UV cross-linking can capture transient states that would not be seen by chemical footprinting. The detection of all 17 cross-links in the reaction medium containing 5 mM Mg^{2+} is consistent with the detectable albeit very slow self-splicing of the Ll.LtrB RNA with short exons under these conditions, implying that the active structure can form at least transiently.

Importantly, our results show significant differences in the active RNA structures induced by the LtrA protein and high Mg^{2+} concentrations. Thus, of the 17 cross-links identified here, only seven (C-22xU329, U-8xU221, U-4xU277, U191xC2421, U212xA336, C371xA2317, and G602xG2386) were enhanced by both the LtrA protein and a high Mg^{2+} concentration. Further, none of the three cross-links that was most strongly enhanced between 5 and 50 mM Mg^{2+} (C-6xC153, U2xC109, and C25xC260) was enhanced by LtrA binding, and the three cross-links that were most strongly enhanced by LtrA binding (U177xU527, G418xC2321, and G602xC2386) either decreased or were not enhanced comparably by high Mg^{2+} concentrations. Another cross-link {C25xC260 [DI(ii)−DId3]}, which was enhanced by high Mg^{2+} concentrations, disappeared almost completely after LtrA binding at either low or high Mg^{2+} concentrations. This cross-link could reflect a non-native structure of DI that is stabilized by high Mg^{2+} concentrations, or its disappearance could reflect direct occlusion of nucleotide residues by LtrA. The previous RNA footprinting study also indicated that the active structures induced by the LtrA protein or high Mg^{2+} concentrations differ, with the $\kappa-\kappa'$, ζ, and DIII terminal loop regions showing strong protections at 50 mM Mg^{2+}, which were not detected in the presence of the LtrA protein at 5 mM Mg^{2+} (*27*). These RNA structural differences can be correlated with the finding that the rate of maturase-promoted splicing at 5 mM Mg^{2+} is 7-fold faster than the rate of maturase-promoted splicing at 50 mM Mg^{2+} and at least 30-fold faster than the maximal rate of self-splicing (Table 1) (*27*). We suggest that these rate differences reflect at least in part the fact that a high Mg^{2+} concentration stabilizes non-native structures that constitute kinetic traps,

and/or decreases conformational flexibility in key regions of the intron, thereby impeding structural transitions required between the different steps of splicing.

Finally, the finding that substantial parts of the active intron RNA tertiary structure are formed at least transiently at 5 mM Mg^{2+} suggests that after binding DIVa, interaction of LtrA with the catalytic core could occur at least in part by tertiary structure capture (*48*) rather than by tertiary structure induction. In this view, the initial tight binding to DIVa would serve to increase the local concentration of LtrA and possibly to orient the protein for interaction with other regions of the intron's catalytic core. Further insight into this possibility may be obtained by kinetic analysis of trans-cleavage reactions, which would enable comparison of reactions rates for ribozymes with and without prebound maturase.

ACKNOWLEDGMENT

We thank Manabu Matsuura, Georg Mohr, Paul Wollenzien (North Carolina State University, Raleigh, NC), and Steven Zimmerly (University of Calgary, Calgary, AB) for comments on the manuscript, Robin Gutell for help with the covariation analysis, and Eric Westhof (Strasbourg, France) for helpful discussions and for suggesting the parallel arrangement of DI and DII.

REFERENCES

1. Michel, F., and Ferat, J. L. (1995) *Annu. Rev. Biochem. 64*, 435−461.
2. Lambowitz, A. M., and Perlman, P. S. (1990) *Trends Biochem. Sci. 15*, 440−444.
3. Carignani, G., Groudinsky, O., Frezza, D., Schiavon, E., Bergantino, E., and Slonimski, P. P. (1983) *Cell 35*, 733−742.
4. Saldanha, R., Chen, B., Wank, H., Matsuura, M., Edwards, J., and Lambowitz, A. M. (1999) *Biochemistry 38*, 9069−9083.
5. Lambowitz, A. M., Caprara, M. G., Zimmerly, S., and Perlman, P. S. (1999) in *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkin, J. F., Eds.) 2nd ed., pp 451−485, Cold Spring Harbor Laboratory Press, Plainview, NY.
6. Zimmerly, S., Guo, H., Perlman, P. S., and Lambowitz, A. M. (1995) *Cell 82*, 545−554.
7. Zimmerly, S., Guo, H., Eskes, R., Yang, J., Perlman, P. S., and Lambowitz, A. M. (1995) *Cell 84*, 529−538.
8. Yang, J., Zimmerly, S., Perlman, P. S., and Lambowitz, A. M. (1996) *Nature 381*, 332−335.
9. Belfort, M., Derbyshire, V., Parker, M. M., Cousineau, B., and Lambowitz, A. M. (2002) in *Mobile DNA II* (Craig, N. L., Craigie, R., Gellert, M., and Lambowitz, A. M., Eds.) pp 761−783, ASM Press, Herndon, VA.
10. Qin, P. Z., and Pyle, A. M. (1998) *Curr. Opin. Struct. Biol. 8*, 301−308.
11. Costa, M., Michel, F., and Westhof, E. (2000) *EMBO J. 19*, 5007−5018.
12. Gordon, P. M., Sontheimer, E. J., and Piccirilli, J. A. (2000) *Biochemistry 39*, 12939−12952.
13. Swisher, J., Duarte, C. M., Su, L. J., and Pyle, A. M. (2001) *EMBO J. 20*, 2051−2061.
14. Gordon, P. M., and Piccirilli, J. A. (2001) *Nat. Struct. Biol. 8*, 893−898.
15. Pyle, A. M. (2002) *J. Biol. Inorg. Chem. 7*, 679−690.
16. Fedorova, O., Mitros, T., and Pyle, A. M. (2003) *J. Mol. Biol. 330*, 197−209.
17. Koch, J. L., Boulanger, S. C., Dib-Hajj, S. D., Hebbar, S. K., and Perlman, P. S. (1992) *Mol. Cell. Biol. 12*, 1950−1958.
18. Wank, H., San Filippo, J., Singh, R. N., Matsuura, M., and Lambowitz, A. M. (1999) *Mol. Cell 4*, 239−250.
19. Costa, M., and Michel, F. (1995) *EMBO J. 14*, 1276−1285.
20. Chanfreau, G., and Jacquier, A. (1996) *EMBO J. 15*, 3466−3476.
21. Boudvillain, M., and Pyle, A. M. (1998) *EMBO J. 17*, 7091−7104.
22. Boudvillain, M., de Lencastre, A., and Pyle, A. M. (2000) *Nature 406*, 315−318.
23. Zhang, L., and Doudna, J. A. (2002) *Science 295*, 2084−2088.
24. Mills, D. A., McKay, L. L., and Dunny, G. M. (1996) *J. Bacteriol. 178*, 3531−3538.
25. Shearman, C., Godon, J. J., and Gasson, M. (1996) *Mol. Microbiol. 21*, 45−53.
26. Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J., Mohr, G., Cavanagh, S., Dunny, G. M., Belfort, M., and Lambowitz, A. M. (1997) *Genes Dev. 11*, 2910−2924.
27. Matsuura, M., Noah, J. W., and Lambowitz, A. M. (2001) *EMBO J. 20*, 7259−7270.
28. Singh, R. N., Saldanha, R. J., D'Souza, L. M., and Lambowitz, A. M. (2002) *J. Mol. Biol. 318*, 287−303.
29. Wilms, C., Noah, J. W., Zhong, D., and Wollenzien, P. (1997) *RNA 3*, 602−612.
30. Noah, J. W., and Wollenzien, P. (1998) *Biochemistry 37*, 15442−15448.
31. Bradford, M. (1976) *Anal. Biochem. 72*, 248−254.
32. Cannistraro, V. J., Wice, B. M., and Kennell, D. E. (1985) *J. Biochem. Biophys. Methods 11*, 163−175.
33. Wilms, C., and Wollenzien, P. (1994) *Anal. Biochem. 221*, 204−205.
34. Nolte, A., Chanfreau, G., and Jacquier, A. (1998) *RNA 4*, 694−708.
35. D'Souza, L. M., and Zhong, J. (2002) *BMC Mol. Biol. 3*, 17.
36. Leontis, N. B., and Westhof, E. (1998) *J. Mol. Biol. 283*, 571−583.
37. Branch, A. D., Benenfeld, B. J., and Robertson, H. D. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 6590−6594.
38. Bernabeu, C., Conde, P., Vasquez, D., and Ballesta, J. (1979) *Eur. J. Biochem. 93*, 527−533.
39. Kim, J., Walter, A. E., and Turner, D. H. (1996) *Biochemistry 35*, 13753−13761.
40. Cannone, J. J., Subramanian, S., Schnare, M. N., Collett, J. R., D'Souza, L. M., Du, Y., Feng, B., Lin, N., Madabusi, L. V., Muller, K. M., Pande, N., Shang, Z., Yu, N., and Gutell, R. R. (2002) *BMC Bioinf. 3*, 2.
41. Podar, M., Dib-Hajj, S., and Perlman, P. S. (1995) *RNA 1*, 828−840.
42. Costa, M., and Michel, F. (1999) *EMBO J. 18*, 1025−1037.
43. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1685.
44. Noah, J. W., Shapkina, T. S., Nanda, K., Huggins, W., and Wollenzien, P. (2003) *Biochemistry* (in press).
45. Conn, G. L., Draper, D. E., Lattman, E. E., and Gittis, A. G. (1999) *Science 284*, 1171−1174.
46. Singh, N. N., and Lambowitz, A. M. (2001) *J. Mol. Biol. 309*, 361−386.
47. Salet, C., and Bensasson, R. (1975) *Photochem. Photobiol. 22*, 231−235.
48. Weeks, K. M., and Cech, T. R. (1996) *Science 271*, 345−348.
49. Massire, C., and Westhof, E. (1998) *J. Mol. Graphics Modell. 16*, 197−205.
50. Massire, C., and Westhof, E. (1998) *J. Mol. Graphics Modell. 16*, 255−257.

BI035339N