Assembly of an Active Group II Intron−Maturase Complex by Protein Dimerization†

Robert P. Rambo and Jennifer A. Doudna*,‡

Department of Molecular Biophysics and Biochemistry and Howard Hughes Medical Institute, Yale University, New Haven, Connecticut 06520

Received January 9, 2004; Revised Manuscript Received April 1, 2004

ABSTRACT: Group II intron-encoded proteins promote both splicing and mobility of the intron RNA through formation of a specific RNA−protein complex. The *Lactococcus lactis* L1.LtrB intron encodes a maturase, LtrA, with reverse transcriptase homology and specific binding affinity for two domains of the intron RNA. The catalytically active ribonucleoprotein (RNP) has splicing, endonuclease, and reverse transcriptase activity, enabling efficient insertion of the intron sequence by a retro-homing mechanism. To determine the composition and assembly mechanism of the RNP complex, purified LtrA protein was analyzed for its ability to recognize a series of intron-derived RNAs. Equilibrium dissociation measurements show that LtrA recognizes two intronic domains, DI and DIV. However, distinct electrostatic requirements for binding imply different modes of molecular recognition in each case. Stoichiometric binding experiments show that the functional RNP complex consists of a dimeric protein species bound to a single intron RNA. Significant differences between the measured equilibrium dissociation constants and kinetically derived values suggest that conformational changes accompany assembly of the intron−maturase complex, and results of limited proteolysis and fluorescence spectroscopy experiments suggest that significant RNA-dependent structural changes within the maturase occur upon association with the intron. These results support a mutually induced fit model in which RNA-dependent conformational changes within LtrA enable stable association of the protein dimer with two independent intron domains to form a functional RNP.

Protein-assisted RNA splicing by group I and group II introns and by the spliceosome involves folding of catalytically important RNA structures through the ordered assembly of RNA−protein domains. In some group I and II introns, an open reading frame within the intron itself encodes a maturase protein that binds the intron RNA to form a catalytically active ribonucleoprotein (RNP)[1] particle. Maturases from the *Lactococcus lactis* L1.ltrB intron and the *Saccharomyces cerevisiae* mtDNA aI1 and aI2 introns are essential for efficient *in vivo* intron splicing, and each forms a complex with its cognate intron that enables the RNA to be a mobile genetic element (*1*−*4*).

Detailed studies of the aI1, aI2, and L1.ltrB group II introns revealed that in each case the intron RNP complex, whose formation regulates translation of the intron-encoded maturase, comprises the catalytically competent structure required for splicing (*5*). Following excision of an intron from a precursor transcript, the intron RNP complex functions as a mobile retro-homing particle capable of site specific DNA cleavage, target DNA-primed reverse transcription, and integration of the intron copy by recombination or repair mechanisms (*6*). The variety of activities associated with mobile group II intron-encoded proteins suggests that specific functional protein domains must exist for recognition of the different substructures of the intron required for translational regulation, splicing, and retro-homing (*7*−*11*).

On the basis of the sequence alignment, group II intron-encoded maturases are organized into three conserved domains: an N-terminal reverse transcriptase (RT) domain, a unique maturase (X) domain, and a C-terminal DNA endonuclease (ENDO) domain (*12*). The RT domain belongs to the family of reverse transcriptases found in non-long terminal repeat (non-LTR) retrotransposable elements, which includes the R2 element of arthropods and the LINE-1 (L1) retrotransposons found throughout the human genome (*13*−*15*). The RT domain is the largest conserved domain of group II intron-encoded ORFs, containing the seven conserved regions typical of reverse transcriptases as well as an N-terminal Z motif common in non-LTR retroelements (*16*). The X domain is the putative splicing factor, or maturase domain, and is uniquely found in related group II intron-encoded ORFs (*12*). The domain is highly basic, consistent with nucleic acid binding proteins, and is the most commonly retained sequence motif in ORFs lacking identifiable RT and ENDO domains (*12*). *In vivo* studies have demonstrated that mutations in domain X are responsible for deficiencies in splicing, supporting the possibility of its involvement in mediating contacts with the intronic RNA to promote protein-

† This work was supported by Grant GM22778 from the National Institutes of Health.

* To whom correspondence should be addressed. Phone: (510) 643-0225. Fax: (510) 643-0080. E-mail: doudna@uclink.berkeley.edu.

‡ Present address: Department of Molecular and Cell Biology and Department of Chemistry, Howard Hughes Medical Institute, University of California at Berkeley, Berkeley, CA 94720.

[1] Abbreviations: SRP, signal recognition particle; MES, 2-(4-morpholino)ethanesulfonic acid; MPD, (±)-2-methyl-2,4-pentanediol; HEPES, 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid; DTT, 1,4-dithio-DL-threitol; TCEP, tris(2-carboxyethyl)phosphine; βME, β-mercaptoethanol; RNP, ribonucleoprotein.

10.1021/bi049912u CCC: $27.50
Published on Web 05/06/2004

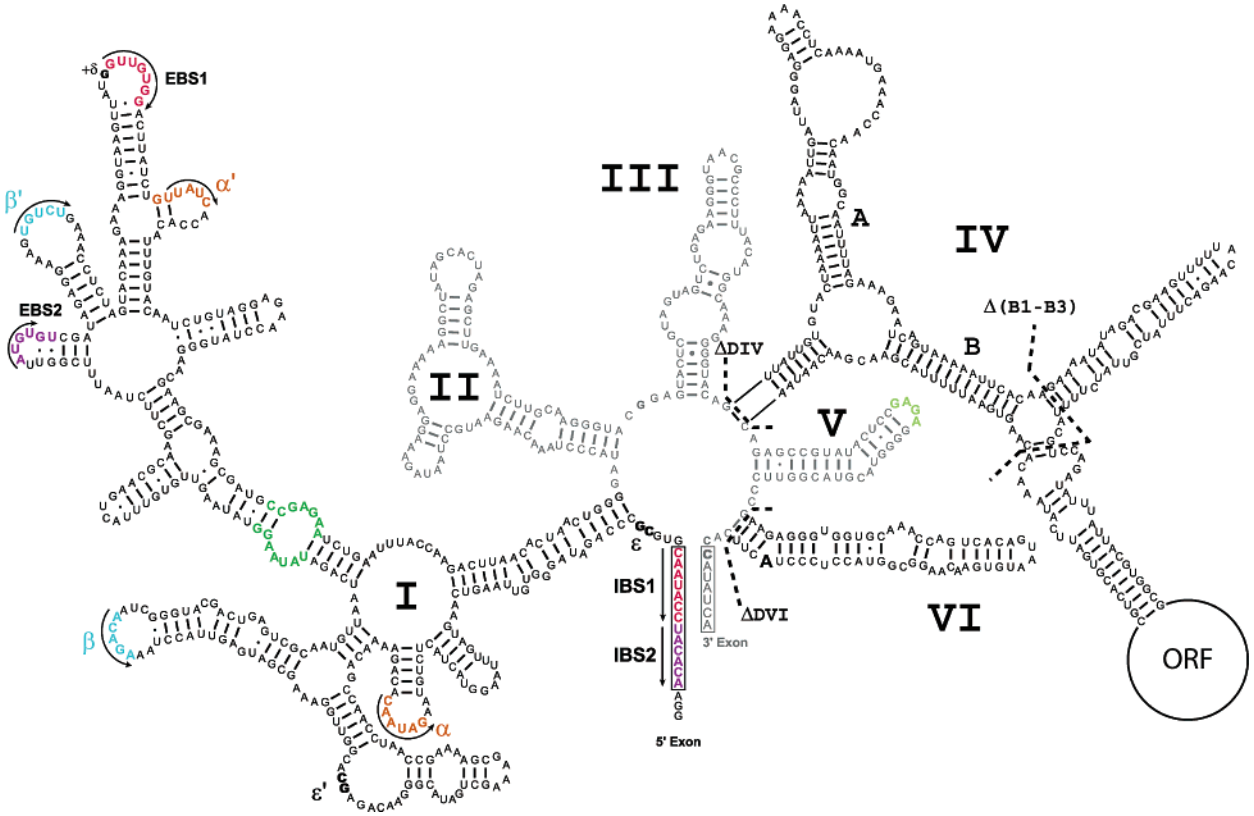

FIGURE 1: Predicted secondary structure of the *L. lactis* L1.LtrB intron. Large black Roman numerals denote the six individual helical domains of group II introns; intron domains represented by bold nucleotide residues denote the individual domains cloned from the L1.ltrB intron used in this study. Dashed lines demarcate internal deletions of individual domains or subdomains, labeled accordingly with a preceding Δ character. Within domain IV, regions A and B refer to the major subdomain divisions DIVA and DIVB, respectively, and SD refers to the Shine−Dalgarno sequence for prokaryotic ribosomal assembly. ORF refers to the intron-encoded open reading frame, truncated to contain only an additional 66 nucleotide residues not represented in the above secondary structure. Sequences in color and denoted with Greek letters highlight tertiary interactions present within group II introns; EBS1, EBS2, IBS1, and IBS2 refer to intron−exon interactions. The 5′ exon used in this study is shown, and the 3′ exon contains additional nucleotides of single-stranded RNA not represented in the above secondary structure.

assisted splicing (*2*, *17*). The ENDO domain is the DNA endonuclease domain consisting of both a DNA binding region and conserved sequence motifs characteristic of H−N−H DNA endonucleases. The ENDO domain binds a single Mg^{2+} ion, which appears to be required for the prescribed endonuclease activity (*18*). Deletion of the ENDO domain severely attenuates RT activity but still maintains its ability to facilitate protein-dependent splicing (*18*). The domain arrangement of mobile group II intron-encoded proteins provides for a multifunctional protein and, within the context of the group II intron RNA, enables intron mobility or retro-homing.

Group II intron-encoded proteins appear to be intron specific splicing factors. Detailed kinetic and *in vivo* studies on the L1.ltrB-encoded protein, LtrA, have demonstrated that LtrA binds to a primary high-affinity binding site located within domain IV (DIV) of the intron (Figure 1) (*19*, *20*). This binding region consists of the ribosome-binding site and initiation codon for the encoded ORF. As with other group II introns, DIV is not critical for *in vitro* self-splicing reactions or for mediating maturase-dependent splicing (unpublished results) (*19*, *21*). Therefore, the functional relevance of this interaction remained unclear until recently, when it was shown to be involved in translation regulation of the intron-encoded ORF (*22*). A subregion of DIV, DIVa, is required for maximally efficient maturase-dependent splicing and for intron homing (*10*, *19*). These observations imply that additional secondary interactions within the intron that are required for assembly of the RNP complex into a maturase-dependent splicing competent structure must exist. Consistent with this hypothesis, deletion analysis demonstrated that domain I (DI), including interactions with the 5′ exon, are important for enhanced interactions with LtrA. Fragments of the intron containing either the 5′ exon and DI or domains V and VI with the 3′ exon bound significantly better than nonspecific RNA (*19*, *20*).

The group II intron is characterized by a highly conserved secondary structure divided into six distinct domains (DI−DVI) (*23*). Domains I and V play fundamental roles in the catalytic properties of group II introns comprising the catalytic core required for splicing (*24*−*26*). Domain IV, which varies in size and sequence, contains the ORF for encoded proteins. In the yeast aI1 and aI2 introns, the ORFs are translated in-frame with the 5′ exon and proteolytic processing produces the mature protein. In contrast, the ORF encoded by the L1.ltrB intron is translated independently of the 5′ exon such that it contains the necessary Shine−Dalgarno sequence required for prokaryotic translation (*19*, *27*).

To determine how the LtrA protein forms a functional complex with its cognate LtrB intron, we developed an improved method for purification of LtrA and tested its binding activity with a series of substrate RNAs derived from the intron. Different electrostatic requirements for the

interaction between LtrA and each of its primary intron binding sites, DI and DIV, imply distinct modes of molecular recognition in each case. Furthermore, measured equilibrium dissociation constants differ significantly from the kinetically reported values, providing evidence for a mutually induced fit model for assembly of the group II intron RNP complex (*28*−*30*). The stoichiometry of formation of the LtrA−intron complex suggests that the functional RNP comprises a dimeric protein species on the RNA. Finally, limited proteolysis experiments with LtrA in the absence and presence of intron-derived RNAs show that the protein structure is dramatically stabilized upon association with the RNA. Together, these data support a model in which significant RNA-dependent conformational changes within LtrA enable the stable association of the protein dimer with two independent regions of the intron structure in forming the functional RNP.

MATERIALS AND METHODS

Protein Expression and Purification. LtrA was cloned from the pImp-1p expression vector kindly provided by A. M. Lambowitz (University of Texas, Austin, TX) into the pThioHis-B vector from the His-Patch ThioFusion Expression System (Invitrogen). Using the 5′ primer (5′-AGCGCGGGGGATCCGGTGAAAACCTGTACTTCCAGGGGATGAAACCAACAATGGCAATTTTA-3′) which adds both a *Bam*HI and TEV protease cleavage site and the 3′ primer (5′-CCGGGCCCGAATTCTCACTACTTGTGTTTATGAATCACGTG-3′) which adds an *Eco*RI site, LtrA was PCR amplified from pImp-1p and cloned into pThioHis-B digested with *Bam*HI and *Eco*RI restriction enzymes. Resulting recombinant plasmid pThB-LtrA contains the TEV protease recognition sequence and LtrA gene fused in-frame to the recombinant 11.7 kDa thioredoxin gene. The plasmid was sequenced to verify the correct insertion and sequence.

Expression of recombinant LtrA was performed using *Escherichia coli* BL21(DE3)-RIL CodonPlus competent cells (Stratagene). Several colonies from a freshly transformed plate were used to make a starter culture, inoculating 100 mL of LB supplemented with 100 μL of 100 mg/mL ampicillin and 20 μg/mL chloramphenicol. The starter culture was allowed to grow for approximately 10 h at 37 °C, at which point 8 mL was used to inoculate 750 mL of LB in a 2 L shaker flask supplemented with 100 μg/mL ampicillin and 20 μg/mL chloramphenicol. The cultures were incubated at 25 °C and induced for expression at an OD_{600} of 0.6 with 750 μL of 1 M isopropyl β-D-thiogalactoside (IPTG). Expression was allowed to proceed for an additional 3.5 h at 25 °C. Cells were harvested by centrifugation, and the resulting cell pellet was resuspended in 25 mL of buffer containing 25 mM Tris-HCl (pH 7.8), 50 mM PO_4^{2-} (pH 7.8), 500 mM NaCl, 0.1% NP-40, 5 mM imidazole, 5 mM βME, 0.5 mg/mL lysozyme, 0.2 mg/mL DNase, 57 mM PMSF, 800 μg/mL leupeptin, 4 mM $MgCl_2$, and 20 μM $CaCl_2$. Resuspended cells were lysed via three freeze−thaw cycles, thawing at 25 °C. The bacterial lysate was clarified by centrifugation at 4 °C (20 min at 20000*g*).

The clarified lysate was passed over an equilibrated Ni^{2+}−NTA Superflo Column (Qiagen). The column was prepared with 5 column volumes of the lysis buffer (excluding lysozyme, DNase, $MgCl_2$, $CaCl_2$, and protease inhibitors). The supernatant was loaded at 1 mL/min and washed with an additional 1 column volume of lysis buffer. To remove unbound proteins, the column was washed with 2.5 column volumes of pH 6.0 buffer containing 50 mM MES (pH 6.0), 20 mM PO_4^{2-}, 500 mM NaCl, 0.1% NP-40, 5 mM imidazole, and 5 mM βME. The recombinant protein was eluted with a 200 mM imidazole buffer containing 20 mM Tris-HCl (pH 8.5), 20 mM PO_4^{2-}, 500 mM NaCl, 10% glycerol, and 5 mM βME.

Contaminating nucleic acids were removed via a secondary purification using a 5 mL heparin HiTrap column (Pharmacia) equilibrated in buffer A containing 20 mM Tris-HCl (pH 8.1), 10% glycerol, 10 mM DTT, and 0.1 mM EDTA at 4 °C. The undialyzed protein was loaded at 1 mL/min and washed with buffer A. The recombinant protein was eluted with a linear 100 mL buffer gradient from 100% buffer A to 100% buffer B, where buffer B contains 1 M NaCl in addition to the contents of buffer A. Fractions were collected at 5 mL intervals. A UV absorbance scan of the purified protein was used to verify removal of bound nucleic acids. The Ni^{2+} affinity tag was removed by incubating the recombinant protein overnight with TEV protease at 4 °C. The protein was concentrated to 25−30 mL and loaded onto a Superdex 200 HR 26/60 gel filtration column preequilibrated with 2.5 column volumes of buffer containing 1 M NaCl, 20 mM Tris-HCl (pH 8.1), 5 mM βME, and 10% glycerol. Purification was performed twice as only 12−15 mL can be loaded during each gel filtration run; purified fractions containing LtrA were pooled. The concentration of the purified protein was determined using the Bradford assay with BSA (66 kDa) as a standard. This final purification produces >98% pure protein as judged by denaturing SDS−PAGE gels visualized by Coomassie staining. The purified LtrA can be stored at 1−3 mg/mL for up to 6 months at 4 °C in 1 M NaCl buffer containing 20 mM Tris-HCl (pH 8.1), 5 mM βME, and 10% glycerol.

RNA Cloning. The pL1.ltrB plasmid used for synthesis of full-length L1.ltrB intronic RNA was cloned from plasmid pGMΔORF kindly provided by A. M. Lambowitz. Standard PCR methods were used to incorporate a T7 RNA polymerase promoter, where the resulting PCR fragment was inserted into pUC19 using *Bam*HI and *Eco*RI restriction sites. Internal domain deletions were generated using the two-step PCR fusion technique; plasmids pL1.ltrBΔDIV and pL1.ltrBΔDVI were cloned from pL1.ltrB. Exact domain deletions are depicted in Figure 1. PCR fragments corresponding to the correct nucleotide size were digested with *Bam*HI and *Eco*RI and ligated into pUC19. Plasmids pDI, pDIV, and pDVI corresponding to domains I, IV, and VI, respectively, were cloned from pL1.ltrB with hammerhead and hepatitis delta virus ribozymes on the 5′ and 3′ ends, respectively. Domains I and IV were cloned directly from pL1.ltrB using standard PCR techniques, while DVI was cloned from two primers covering the entire cloned sequence where a PCR was used to extend the ends of the primers prior to restriction digestion and cloning into *Bam*HI- and *Eco*RI-digested pUC19. Precise domain boundaries of the individual intronic domains are depicted in Figure 1. Plasmids containing the intron and intron derivatives were linearized with *Sap*I, and plasmids containing the individual domains were linearized with *Bam*HI for T7-based *in vitro* transcription reactions.

Preparation of Group II Intron RNA Samples. All RNA samples were prepared using T7 RNA polymerase runoff transcription reactions in buffer containing 20 μg/mL linearized plasmid DNA, ATP, GTP, CTP, and UTP (32 mM each), 30 mM Tris-HCl (pH 8.1), 0.01% Triton X-100, 2 mM spermidine, and 34 mM $MgCl_2$. Transcription reactions were optimized for 4 h incubations at 37 °C. Reactions were extracted with phenol and chloroform followed by ethanol precipitation and 70% ethanol wash prior to electrophoretic separation on preparative 8 M denaturing polyacrylamide gels. Gels were visualized by UV shadowing, and bands corresponding to the RNA of interest were excised and eluted overnight into a buffer containing 300 mM sodium acetate (pH 5.6) and 0.1 mM EDTA. The eluant was ethanol precipitated, washed, and resuspended in buffer containing 10 mM Tris-HCl (pH 7.0) and 0.1 mM EDTA.

Body-labeled RNA samples were prepared using the previously described conditions for T7 RNA polymerase-based transcriptions with the exception that transcription reactions were optimized for maximal incorporation of [α-^{32}P]ATP using the following conditions: 30 mM $MgCl_2$, GTP, UTP, and CTP (8 mM each), 0.5 mM ATP, and 25 μL of 3000 Ci/mmol (10 mCi/mL) [α-^{32}P]ATP. Radiolabeled RNA samples were purified as described above except that samples were visualized by autoradiography and quantitated by standardized scintillation counting.

To prepare (5′-^{32}P) end-labeled RNA samples, 10 μg of RNA was incubated at 37 °C for 0.5 h in a 20 μL reaction mixture containing 3 μL of T4 polynucleotide kinase (PNK), 1× PNK buffer (New England Biolabs), and 1 μL with 5 mCi of [α-^{32}P]ATP. RNA was gel-purified as described above and visualized by autoradiography.

Equilibrium Binding Studies. Apparent dissociation constants for RNA−protein interactions were determined using trace quantities of α-^{32}P-labeled RNA (1−25 pM) based on a modification of the nitrocellulose filter binding method of Riggs et al. (1968). Unless otherwise specified, intronic RNA samples were annealed by heating to 65 °C for 5 min and slow cooled in an aluminum block to room temperature in a buffer containing 100 mM NaCl, 5 mM $MgCl_2$, and 10 mM Tris (pH 7.5). Annealed samples were mixed with varying amounts of protein to a final reaction volume of 200 μL containing the following at their final concentrations: 250 or 450 mM NaCl, 5 mM $MgCl_2$, 10 mM Tris-HCl (pH 7.5), and 5 mM βME. Binding reaction mixtures were equilibrated for 1−1.5 h at room temperature prior to partitioning through a 96-well dot blot vacuum apparatus containing presoaked membranes, a nitrocellulose membrane layered on top of a positively charged Hybond-N+ membrane (Amersham). Binding reaction mixtures incubated for 1 or 12 h were identical, indicating a 1 h incubation is sufficient for equilibrium. Partitioned samples were washed once with 200 μL of the appropriate binding buffer. Filters were dried, and radioactivity corresponding to free and bound states was quantitated by phosphorimaging (Molecular Dynamics). The apparent dissociation constant was determined from data prepared using Microsoft Excel and, as described, fit to either eq 1 or 2 using nonlinear least-squares methods as implemented in Kalidegraph.

$$\theta = \frac{[P]}{K_d + [P]} \quad (1)$$

$$\theta = \frac{[P]^n}{K_d^n + [P]^n} \quad (2)$$

where θ is the fraction of RNA bound, [P] is the concentration of LtrA, K_d is the apparent dissociation constant, and n is the Hill coefficient.

Stoichiometric Filter Binding Assay. Stoichiometric filter binding assays were performed with unlabeled RNA samples (100−800 nM) supplemented with trace quantities of α-^{32}P-labeled RNA added for detection (<1 nM). RNA samples were annealed and mixed with protein samples as described above. For binding reactions with 900 mM NaCl, annealed RNA samples were diluted in a buffer containing 900 mM NaCl and mixed with an equal volume of protein at 900 mM NaCl. The extent of binding was determined as described above. Data were analyzed and fit to eq 3, which describes complex formation

$$\theta = \frac{r + K_d + n - \sqrt{(r + K_d + n)^2 - 4rn}}{2n} \quad (3)$$

where θ is the fraction of RNA bound, r is the molar ratio of [RNA]:[LtrA] where specified, n is the stoichiometric equivalence point, and K_d is the apparent dissociation constant.

Limited Proteolysis Experiments. Reaction mixtures contained 2 μM LtrA, 450 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris-HCl (pH 7.5), and 5 mM BME. Where specified, experiments included 4.4 μM RNA. The RNA samples were annealed as described for the equilibrium binding assays. After 45 min at room temperature, proteolytic cleavage reactions were initiated by the addition of trypsin to a final concentration of 12.5 ng/μL in a total volume of 200 μL. Aliquots were quenched at designated times in an equal volume of 2× SDS−PAGE loading buffer and boiled immediately at 95 °C for 10 min. Aliquots used for mass spectrometry were added to an equal volume of 8 M guanidinium-HCl in 0.1% TFA. Reaction products were electrophoretically separated using 4 to 20% gradient denaturing SDS−PAGE gels visualized with Coomassie brilliant blue P250 staining.

Samples submitted for N-terminal sequencing were transferred to Immobolin-PSQ membranes by Western blotting techniques. Transferred protein samples were visualized by Coomassie staining and visualized bands excised and submitted for sequencing to the Keck Biotech Facility at Yale University.

Sequence Alignments. Preliminary sequence alignments were generated using a protein−protein BLAST search with the wild-type LtrA protein sequence (GenBank entry Q57005). The LtrA alignment was further refined into the BLAST search results using a secondary search at the Pfam server at http://pfam.wustl.edu/, version 8.0, and manually refined with Seaview.

Mass Spectrometry. Mass spectra were recorded using time-of-flight matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. Sinapinic acid (10%) dissolved in 30% acetonitrile and 0.07% trifluoroacetic acid served as

the matrix. Prior to matrix formation, protein samples were desalted using C_4 ZipTip hydrophobic P-10 (Millipore) pipet tips as described by the manufacturer.

Fluorescence Spectroscopy. Fluorescence spectra were recorded at room temperature, measured using a Photon Technology International fluorescence spectrometer. Measurements were taken using 30 nM LtrA in a buffer containing 450 mM NaCl, 5 mM $MgCl_2$, 20 mM Tris (pH 7.5), and 1 mM TCEP in a total volume of 500 μL. Due to inner filter effects of the RNA ligand between 260 and 300 nm, a correction factor must be incorporated into the observed intensities. Using *N*-acetyltryptophanamide (NAWA) as a fluorescent standard, DIV RNA absorption correction parameters were determined at excitation wavelengths of 282 and 290 nm. Equation 4 was fit to data collected at 282 nm, and eq 5 was fit to data collected at 290 nm. The observed fluorescence intensity at 10 μM NAWA was monitored and fit to the following equation.

$$C_f^{282} = \frac{e^{-aLd} - e^{-aL}}{aL(1-d)} \quad (4)$$

$$C_f^{290} = \frac{1 - 10^{-eL}}{2.303eL} \quad (5)$$

where L is the total ligand concentration and collectively a, d, and e are parameters numerically determined through nonlinear regression of the above equations (Igor Pro). Observations at 282 nm produced an a of 7.20×10^5 M^{-1} cm^{-1} and a d of 1.15, and observations at 290 nm produced an A of 5.35×10^5 M^{-1} cm^{-1}. For each RNA concentration in the titration, corrections to the observed intensities were incorporated using the following equation:

$$F_{corr} = F_{obs}\frac{1}{C_f} \quad (6)$$

where F_{corr} and F_{obs} are the corrected and observed fluorescence intensities, respectively, and C_f is the correction factor as described above.

Gel Filtration Chromatography. Multiangle laser light scattering experiments were performed at room temperature with an HR-10/30 Superdex-200 size exclusion column coupled to instrumentation monitoring changes in light scattering (690 nm), absorbance (280 nm), and the refractive index (Wyatt Technology Corp.). Sample volumes between 100 and 150 μL were injected at flow rates of 0.5 mL/min. Data analysis was performed using the Astra software package using dn/dC values of 0.185 and 0.17 for protein and RNA, respectively. Reported values are for samples at 1 mg/mL. Prior to injection, protein samples were exchanged into the chromatography running buffer consisting of 1 M NaCl, 1 mM TCEP, 5 mM $MgCl_2$, and 25 mM Tris-HCl (pH 7.5). The amount of protein recovered in the monomer peak represented essentially all of the protein loaded onto the column. RNA samples were annealed as previously described and exchanged into a running buffer containing 450 mM NaCl, 5 mM $MgCl_2$, 1 mM TCEP, and 25 mM Tris-HCl (pH 7.5). Analytical grade bovine serum albumin was used as a calibration standard prior to the experimental runs.

RESULTS AND DISCUSSION

Distinct Modes of LtrA Recognition of Different Intron Domains. Domain IV of the L1.ltrB intron is a major binding site for the LtrA protein, and deletion analysis showed that the binding site is localized to the DIVA subdomain (Figure 1) (*20*). However, intron fragments, including the 5′ exon and DI, or domains V and VI and the 3′ exon, bind significantly tighter than nonspecific RNA (*19*), implying additional sites of LtrA recognition. To compare the specificity and affinity of LtrA for DI, DIV, and DVI RNAs (Figure 1), equilibrium binding measurements were performed at varying ionic strengths using a nitrocellulose filter binding assay. While no detectable binding of DVI occurred at 450 mM NaCl, the previously established condition required for optimal protein-dependent splicing (*31*), both DIV and DI were bound with high affinity (Figure 2A,B). The LtrA−DIV RNP complex has an apparent K_d of 8.3 ± 0.1 nM with a binding isotherm weakly described by the Hill equation with a Hill coefficient, α_H, of 1.2 ± 0.1, consistent with a simple two-state binding reaction. However, the measured K_d differs by approximately 10−1000-fold from the reported kinetically determined K_d (9.3−480 pM) calculated as the ratio of k_{off}/k_{on} rates (*20*, *31*). While this observed difference could result from a weaker affinity of the LtrA protein dimer than of the LtrA−intron complex, the bimolecular behavior of the complex at even very low (20−120 pM) LtrA concentrations (*20*, *28*) argues against this possibility. An alternative explanation is that formation of the LtrA−intron complex involves a multistep induced fit mechanism in which a kinetically stable intermediate assembles prior to the native equilibrium state. Similar assembly mechanisms have been demonstrated for other RNP complexes, including the eukaryotic SRP19−7S RNA and CYT18−group I intron complexes (*28*−*30*).

The electrostatic contributions to binding were evaluated within the LtrA−DIV RNP complex by monitoring the equilibrium dissociation constant at varying NaCl concentrations (*32*−*36*). There is an exponential increase in the apparent K_d above 250 mM NaCl (Figure 2B). Analytically, these differential effects of NaCl on the apparent affinity can be described by considering that the binding of a protein (P) to an RNA (R) surface displaces Na^+ counterions condensed on the RNA surface. This relationship is of the form

$$P + R \leftrightarrow RP + n\psi Na^+ \quad (7)$$

where n is the number of ion pairs formed and ψ is the effective fraction of phosphate groups neutralized by the condensed Na^+ counterions. Plotted as the ln K_{obs} versus ln[NaCl] (Figure 2C), the observed log-linear dependence has a slope that is proportional to the number of counterions displaced from the RNA surface upon assembly. If a value of 0.89 is assumed for ψ, standard thermodynamic definitions describe the linear relationship by (*33*)

$$\ln K_a = -n\psi \ln[Na^+] + \ln K^o \quad (8)$$

Therefore, the linear relationship between 150 and 450 mM NaCl indicates there are approximately three to four phosphate groups associated with binding of LtrA to DIV

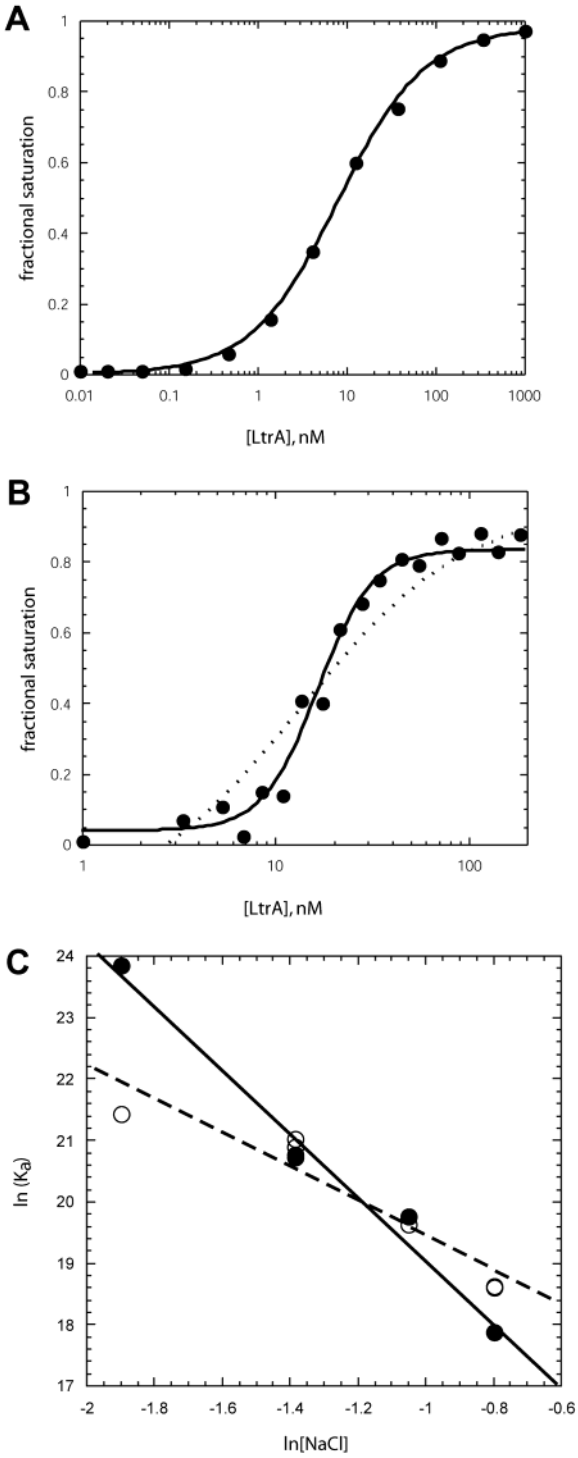

FIGURE 2: Equilibrium K_d measurements for LtrA−domain RNP interactions performed in 450 mM NaCl, 5 mM $MgCl_2$, 5 mM BME, and 20 mM Tris-HCl (pH 7.5) incubated at room temperature for 1 h. (A) Nitrocellulose filter binding results for the LtrA−DIV RNP complex. Data were fit to the semiempirical Hill equation as described in Materials and Methods, yielding a K_d of 8.3 ± 0.1 nM with an associated Hill coefficient of 1.2 ± 0.1. (B) Nitrocellulose filter binding results for the LtrA−DI RNP complex. The solid line represents a nonlinear fit to the Hill equation as described in Materials and Methods, yielding a K_d of 16.6 ± 0.4 nM with an associated Hill coefficient of 3.0 ± 0.2. The dashed line is the best nonlinear fit to the rectangular hyperbolic Langmuir binding isotherm. (C) Effects of varying NaCl concentration on measured K_d values for DIV (○) and DI (●) plotted as natural log K_d vs natural log of [NaCl]. The linear region describes the ionic dependency of the apparent association constant.

Table 1: Equilibrium Binding Measurements

	K_d (nM)	α_H, Hill coefficient	[NaCl] (mM)
intronic domain			
DI[a]	0.95 ± 0.02	1.99 ± 0.07	250
DI[a]	16.6 ± 0.4	3.0 ± 0.2	450
DIV[a]	0.98 ± 0.04	1.67 ± 0.09	250
DIV[a]	8.3 ± 0.1	1.2 ± 0.1	450
DIVΔ(b1-b3)[a]	3.0 ± 0.2	1.4 ± 0.1	250
DVI[a]	nb[e]	nb[e]	450
domain deletions[b]			
L1.ltrB[c]	1.05 ± 0.07	0.95 ± 0.01	450
L1.ltrBΔDVI[d]	1.20 ± 0.04	0.98 ± 0.03	450
L1.ltrBΔDIV[a]	1.99 ± 0.06	1.68 ± 0.07	450
L1.ltrBΔDIIDIV[a]	29.7 ± 0.8	2.01 ± 0.09	450

[a] Reported values are a result of at least three independent experiments of separate preparations of LtrA and RNA. [b] Domain deletions refer to internal deletions of individual domains within the context of the intron. [c] Represents eight independent experiments with three separate preparations of LtrA and four separate preparations of RNA. Measurements were within error between 1.5 and 24 h of incubation. [d] Represents six independent experiments with two separate preparations of LtrA and three separate preparations of RNA. [e] No detectable binding.

RNA. The 10-fold increase in binding affinity in 250 mM NaCl, a less stringent condition, indicates that the strength of the interactions is largely electrostatic (Table 1). Deletion of more than half of DIV [construct DIVΔ(b1−b3), Figure 1] has little effect on LtrA binding affinity, consistent with previous results showing that the DIVA subdomain contains the necessary elements for specific binding by LtrA (*20*, *22*).

The apparent K_d for the LtrA−DI RNP complex is 16.6 ± 0.4 nM in 450 mM NaCl, similar to the value measured for DIV (Figure 2B). Here, however, the binding isotherm is described strongly by the Hill equation with an α_H value of 3.0 ± 0.2, suggesting a multimerization of LtrA on the RNA. Variation in the salt concentration from 150 to 450 mM NaCl results in an exponential change in the apparent dissociation constant, and the relationship from 150 to 450 mM NaCl (Figure 2C) is consistent with approximately six ion pairs formed upon association of the LtrA−DI RNP complex. Comparatively, the interactions within the LtrA−DI RNP complex have a stronger electrostatic component than those involving DIV. These observations support the hypothesis that LtrA uses distinct modes of recognition of the DI versus DIV regions of the LtrB intron.

Binding experiments performed using the full-length intron and derivatives deleted for various domains showed that the presence or absence of the DIV region governs the mode of LtrA binding observed for the intron. At 450 mM NaCl and 5 mM $MgCl_2$, binding isotherms determined for RNP complexes involving both the L1.ltrB and L1.ltrBΔDVI intronic RNAs are described by a simple Langmuir binding isotherm where the fit is indistinguishable from the Hill equation with an α_H of 1 (Figure 3). Thus, the apparent cooperative association determined for isolated DI RNA was not observed (Table 1). However, binding assays using the L1.ltrBΔDIV intron, an internal deletion of DIV within the context of the full-length intron, recapitulated the apparent cooperative association (Figure 4).

In all cases, proper RNA folding was essential for specific and high-affinity LtrA binding as well as splicing activity, as demonstrated by the effects of different annealing condi-

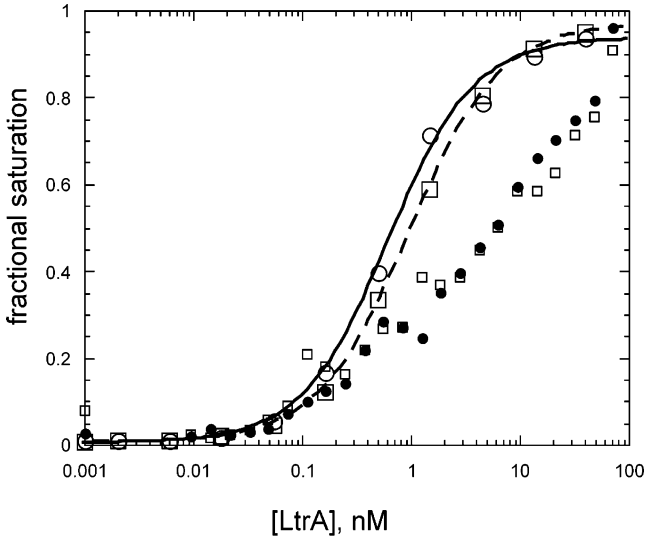

FIGURE 3: Equilibrium nitrocellulose filter binding measurements of LtrA complexed with L1.ltrB or L1.ltrBΔDVI intron constructs. Binding reaction mixtures were incubated at room temperature in a buffer containing 450 mM NaCl, 5 mM $MgCl_2$, 5 mM βME, and 20 mM Tris-HCl (pH 7.5). Intron RNAs L1.ltrB (●) and L1.ltrBΔDVI (small □) annealed by heating to 65 °C for 5 min followed by slow cooling in the presence of 5 mM $MgCl_2$ produced RNAs incapable of specific recognition by LtrA. In contrast, intron RNAs L1.ltrB (○) and L1.ltrBΔDVI (large □) annealed in the presence of 100 mM NaCl produced RNAs specifically bound by LtrA. The binding data were described by the simple Langmuir binding isotherm, yielding K_d values of 1.05 ± 0.07 and 1.20 ± 0.04 nM for L1.ltrB and L1.ltrBΔDVI introns, respectively.

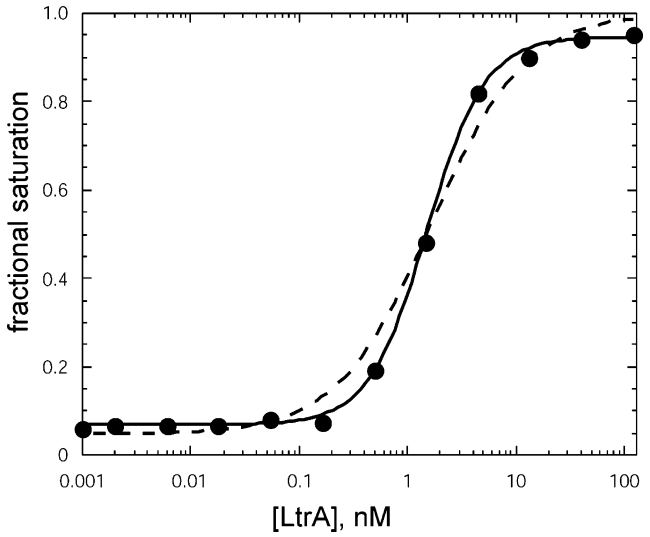

FIGURE 4: Equilibrium nitrocellulose filter binding measurements of LtrA complexed with L1.ltrBΔDIV intron RNA. Data were plotted as fractional saturation vs the log of the total LtrA concentration. The solid line represents the nonlinear fit to the semiempirical Hill equation with a K_d of 1.99 ± 0.06 and an associated Hill coefficient of 1.68 ± 0.07. The dashed line illustrates the best fit to the rectangular hyperbolic Langmuir binding isotherm.

tions. Intronic RNAs slowly annealed in the presence of either 10 mM $MgCl_2$ or 2.5 mM KCl produced RNAs incapable of being specifically recognized by LtrA (Figure 3), and <10% of the L1.ltrB RNA was capable of splicing (data not shown). In contrast, annealing splicing competent constructs slowly in the presence of 100 mM NaCl produced folded RNAs capable of specific binding to LtrA (Figure 3), and this annealing condition also supports the full splicing activity of the complex.

Dimerization of LtrA within Intron RNA Complexes. The apparent cooperative binding of LtrA to DI RNA suggested that LtrA might form a higher-order complex within the assembled RNP. Previous studies on the full-length intron have provided evidence for LtrA dimerization (*31*), though whether dimerization occurs through protein−protein interactions or by distinct interaction of two protein molecules on a single RNA species has been unclear (*31*). Mass spectrometry and multiangle laser light scattering of the free protein showed the molecular mass of the free protein to be 70 077.3 ± 45 and 78 650 ± 90 g/mol, respectively, where the theoretical molecular mass is expected to be 70 219 g/mol. To determine the multimerization state of LtrA in various RNP complexes, a series of stoichiometric filter binding experiments were performed using RNA constructs derived from the L1.ltrB intron. These experiments were performed using several RNA concentrations ranging from 100 to 750 nM, and at varying ionic strengths. While there is a salt-dependent change in the stoichiometric equivalence point, presumably due to nonspecific protein binding at the lower salt concentrations, binding of LtrA to DI RNA at 450 mM NaCl shows a protein:RNA ratio of 2:1 (Figure 5A,B). Similar results were obtained with DIV RNA (Table 2), while nonspecific binding of the small DVI hairpin RNA has a stoichiometric ratio close to 1:1 (Table 2). Stoichiometric binding determinations using the full-length and deletion constructs of the intron at 450 mM NaCl likewise showed protein:RNA ratios of 2:1. Interestingly, the ratio remains unchanged at 900 mM NaCl (Figure 6), providing evidence for a strongly associated complex that includes both protein−protein and protein−RNA interactions. These data suggest that the functional intronic RNP complex consists of a dimeric complex of two LtrA proteins per intron RNA, and that a stable protein dimer forms only upon association with the cognate RNA.

Stabilization of LtrA Domain Structure in the RNP Complex. To assess structural changes in the LtrA protein that occur upon RNA binding, limited proteolysis was conducted in the presence or absence of bound RNAs. Complete trypsin digestion of the 70 kDa, highly basic LtrA protein would produce a 3.6 kDa peptide fragment as the largest possible proteolytic product, indicating that tryptic cleavage sites are dispersed well throughout the protein sequence. Limited proteolysis experiments were performed using a saturating concentration of LtrA in the absence and presence of various RNAs derived from the LtrB intron. Proteolytic digestion of LtrA alone or in the presence of excess nonspecific RNA (poly-A) produces no stable proteolytic fragments of >30 kDa throughout the time course of the reaction. Comparison with the predicted structure of LtrA (Figure 7A,B) suggests the protein domains are not stably folded under these conditions, consistent with previous results (*31*). In contrast, partial tryptic digestion of LtrA in the presence of intronic RNAs produced four major proteolytic fragments (B1−B4) (Figure 7C). Fragment B1 has an apparent molecular mass of just less than 60 kDa, and N-terminal sequencing showed that this fragment consists of the residues at the beginning of the intact protein. Presumably, this fragment is produced by removal of the 12.2 kDa C-terminal ENDO domain from the intact protein, where there is a single lysine residue in a small region

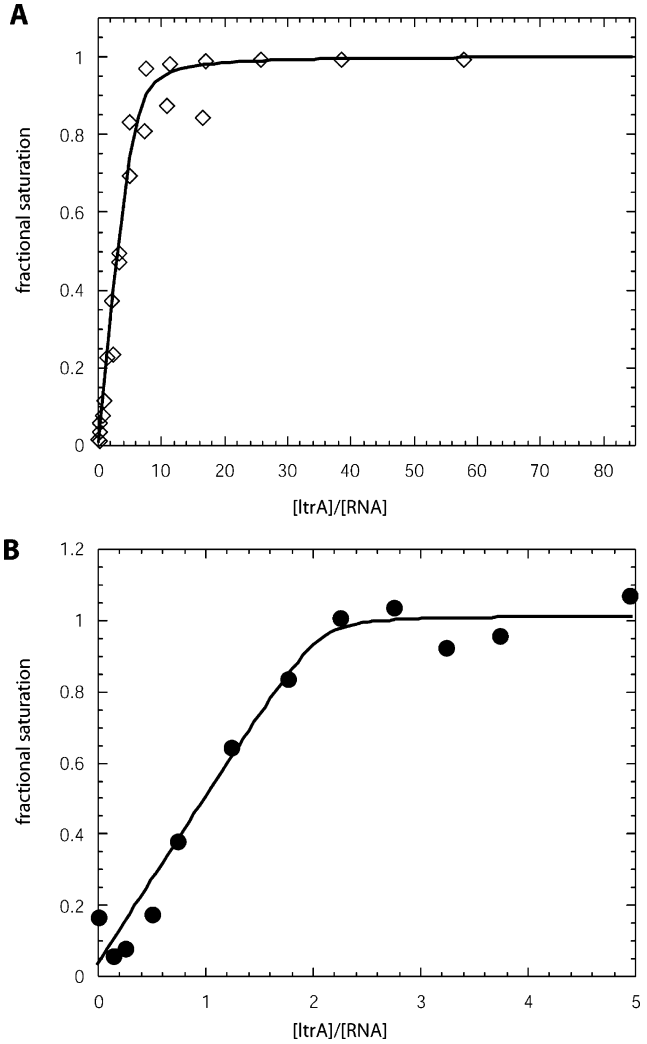

FIGURE 5: Stoichiometric filter binding experiments with LtrA−DI RNP complexes at varying salt concentrations. All reaction mixtures were incubated at room temperature for 1 h in a buffer containing 5 mM $MgCl_2$, 5 mM DTT, and 20 mM Tris-HCl (pH 7.5). Data were fit to a quadratic equation described in Materials and Methods. The titration is plotted as the ratio of protein to RNA vs fraction bound. (A) Binding experiments performed at 250 mM NaCl determine a stoichiometric ratio of 6.0 ± 0.2, indicating six proteins are bound per RNA. (B) At 450 mM NaCl, the stoichiometric ratio shifts to 2.1 ± 0.2.

Table 2: Stoichiometry Binding Experiments

RNA	[LtrA]/[RNA]	[NaCl] (mM)
DI[a]	6.0 ± 0.2	250
DIV[b]	2.87 ± 0.05	250
DVI[a]	0.82 ± 0.02	250
DI[b]	2.0 ± 0.02	450
DIV[b]	2.5 ± 0.4	450
DIV[a,c]	0.4 ± 0.1	450
L1.ltrB[d]	2.3 ± 0.1	450
L1.ltrBΔDIV[a]	2.06 ± 0.06	450
L1.ltrB[b]	2.25 ± 0.04	900

[a] Reported values are a result of at least three independent experiments with separate preparations of LtrA and RNA. [b] Represents eight independent experiments with three separate preparations of LtrA and four separate preparations of RNA. [c] Determined from the fluorescence titration of the protein with RNA. The ratio is written as the inverse where the equivalence point is [RNA]/[LtrA]. [d] Represents six independent experiments with two separate preparations of LtrA and three separate preparations of RNA.

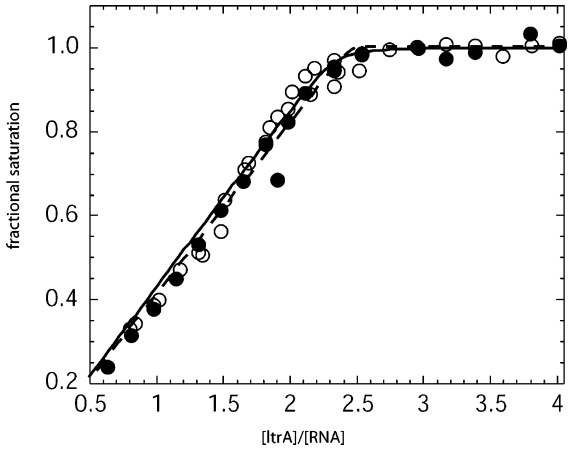

FIGURE 6: Stoichiometric filter binding experiments with LtrA−L1.ltrB RNP complexes at 450 (○) and 900 mM NaCl (●). Binding was performed with varying concentrations of unlabeled intron RNA between 100 and 750 nM supplemented with trace amounts of ^{32}P body-labeled RNA for detection. Reaction mixtures at the specified salt concentration were incubated in a buffer containing 5 mM βME, 5 mM $MgCl_2$, and 10 mM Tris-HCl (pH 7.5). Stoichiometric ratios of 2.3 ± 0.1 and 2.25 ± 0.04 were determined at 450 and 900 mM NaCl, respectively.

connecting the X and ENDO domains, leaving the RT and X domains (Figure 7A,B). The ~60 kDa fragment is produced in the presence of DI, DIV, or DVI RNA, but only persists in the presence of DI or DIV, showing that specific RNA binding stabilizes this substructure of LtrA.

Fragments B2 and B4, with molecular masses determined by mass spectrometry of 42 915 and 26 119 Da, respectively, were the most stable proteolysis products generated in the presence of DI or DIV RNA (Figure 7C). N-Terminal sequencing identified fragment B2 as the N-terminal region of the intact protein, suggesting this fragment results from cleavages in the β variable region of the alignment to yield the RT domain. Fragment B4 was identified as a C-terminal cleavage product, where the sequenced residues map to the β variable region of the alignment (Figure 7A). This fragment consists of both the X and ENDO domains and therefore represents the complementary proteolytic product to fragment B2. The 32 kDa fragment, B3, is a proteolytic fragment of the RT domain that includes the N-terminal residues of the intact protein. This fragment is uniquely present in digestions with DVI RNA and accumulates significantly after 10 min. The appearance of B3 corresponds with depletion of B2, suggesting that B3 results from a secondary proteolytic cleavage of fragment B2 specifically within the α variable region of the alignment. This proteolytic susceptibility subdivides the RT domain into the RTα and RTβ domains as proposed by the sequence alignment analysis.

For each of the three RNAs that were tested, the distinctive proteolytic cleavage patterns observed can be explained by the unique interactions each makes with the LtrA protein. DVI, which does not bind detectably to the LtrA protein under the conditions that were used, provided the least amount of proteolytic protection with the exception of transient accumulation of fragment B3. In contrast, the proteolytic susceptibility of the α variable region is greatly attenuated in the presence of DI, DIV, or the intact intron, suggesting these RNAs either occlude trypsin from the protein or promote a proteolytically resistant structure within LtrA. Specifically, the reactions involving DIV RNA result in a distinctive proteolytic cleavage pattern in which ad-

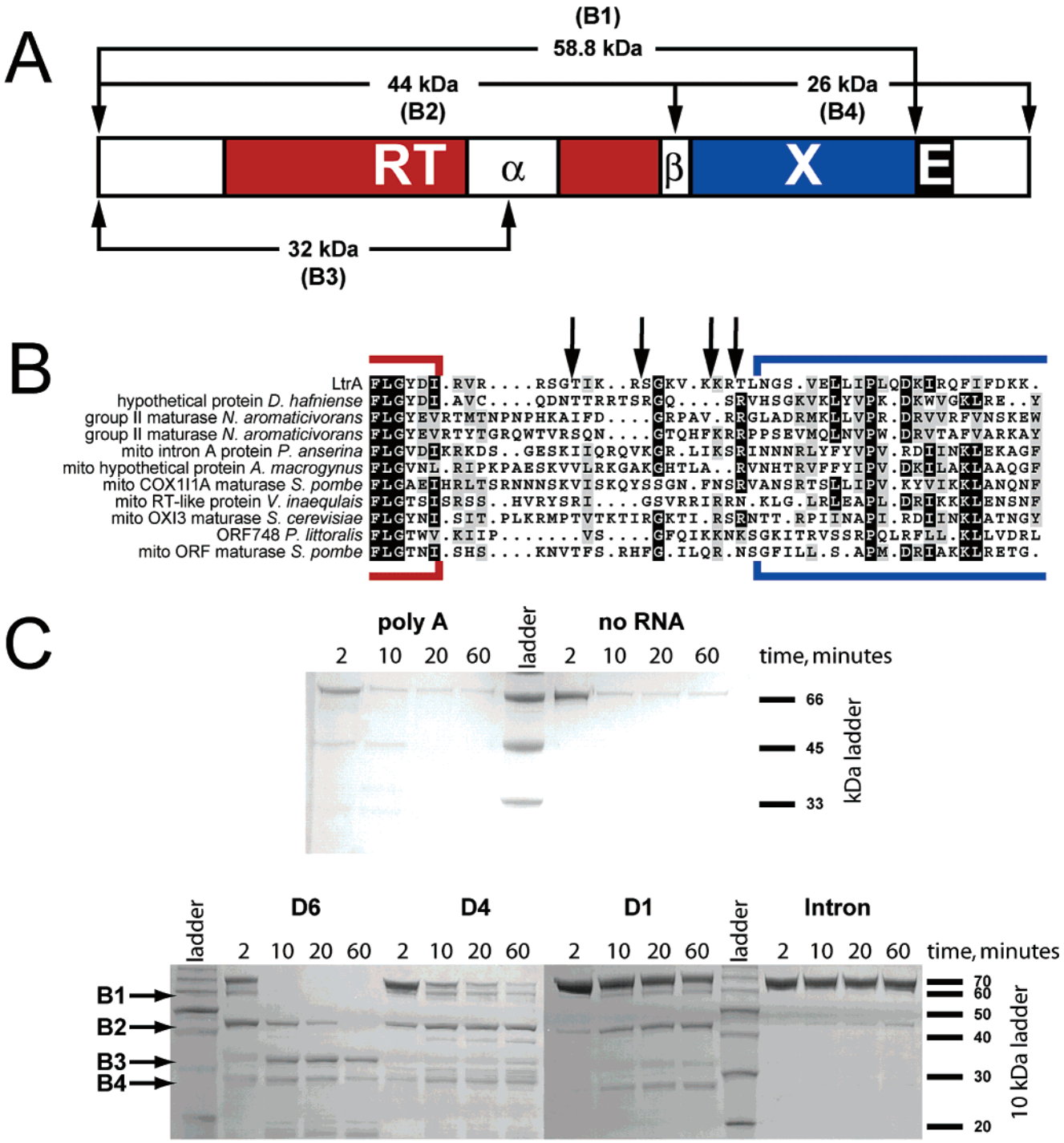

FIGURE 7: Limited proteolysis of LtrA and LtrA−RNA complexes. (A) Schematic representation of the proposed domain organization of LtrA. Colored boxes represent regions of conservation: red for RT, the reverse transcriptase domain; blue for X, the putative maturase domain; and black for E, the endonuclease domain. Boxes α and β correspond to regions of variable sequence in the alignment. Assigned molecular masses describe regions bracketed by arrows and correspond to B1−B4. (B) Expanded region of the β variable region between the RT and X domains of a sequence alignment. Boxed regions are the putative protein domains as described above. Arrows indicate sites of proteolytic cleavage by trypsin as determined by N-terminal sequencing. The sequence for LtrA is explicitly labeled, while all other sequences are labeled according to their respective ORF and biological source. (C) Denaturing 4 to 20% gradient SDS−PAGE analysis of the limited proteolysis reaction of LtrA with trypsin as visualized by Coomassie staining in the presence of poly-A and four different intronic RNAs. Numbers above each lane denote the time points taken (in minutes) for analysis during the cleavage reaction. Poly-A, DI, DIV, DVI, and intron refer to the RNAs in each time course reaction. Arrows labeled B1−B4 identify major proteolytic products. The intact protein is at 70 219 Da, and the numbers on the far right indicate the migration positions of molecular mass protein standards in the ladder. In each experiment, LtrA was at 2 μM with each RNA tested at a 2.5-fold molar excess over the protein.

ditional fragments are observed between 25 and 40 kDa (Figure 8B). Although not positively identified, these fragments could arise from a primary cleavage event within the β variable region, producing the RT domain (44 kDa) and the X−ENDO domain (26 kDa), followed by secondary cleavages within the RT domain. This observation of hypersensitive cleavage sites suggests there may be DIV specific structural changes that occur within the RT domain.

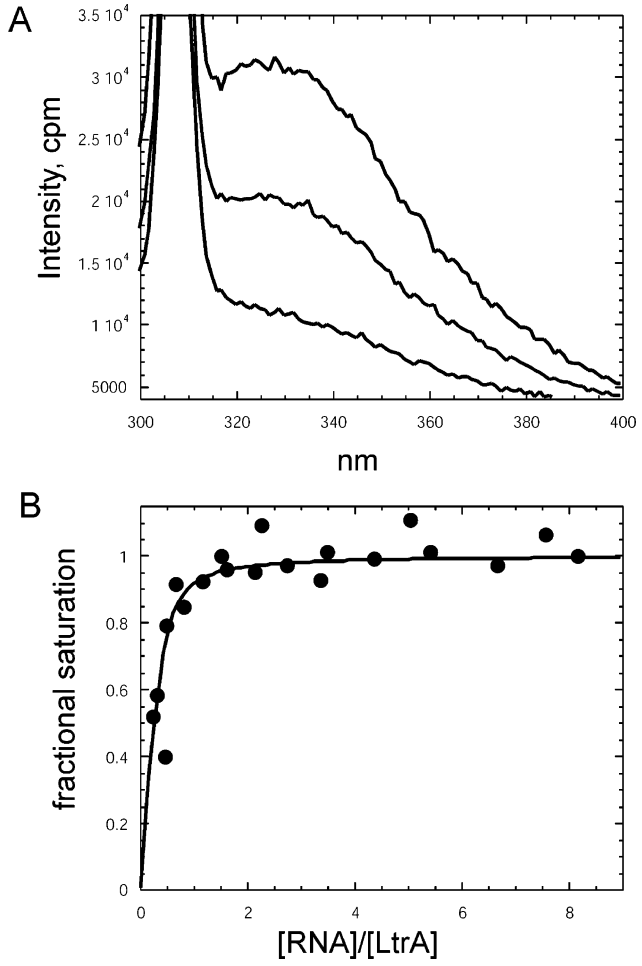

FIGURE 8: Effects of DIV RNA on the intrinsic fluorescence properties of LtrA. (A) Fluorescence emission spectra of LtrA in the presence and absence of DIV RNA: free protein (top), LtrA complexed with 6.6 nM DIV RNA (middle), and LtrA complexed with 244.4 nM DIV RNA (bottom). (B) Corrected fluorescence emission intensities at 331 nm plotted as fractional saturation vs the [RNA]/[LtrA] ratio.

Proteolysis in the presence of DI RNA produces slow stoichiometric conversion of the intact protein into the RT and X−ENDO domains, an observation remarkably similar to the result obtained with the HIV-RT protein. The X-ray crystal structure of the HIV-RT heterodimer revealed that the p66 subunit is divided into two major structural domains, termed the finger/palm and thumb/connection domains (*37*). Partial proteolysis experiments show that the linker region connecting the two domains is susceptible to proteolytic cleavage, thus stably separating the two structural domains (*38*, *39*). Though LtrA and HIV-RT are highly divergent reverse transcriptases, these data suggest that the structural core of the RT domain, consisting of seven blocks of identical sequence, is largely preserved in both.

Intrinsic Fluorescence of LtrA. As an independent assay for conformational changes induced in LtrA upon RNA binding, the intrinsic protein fluorescence was measured before and after complexation with DIV RNA. LtrA contains six tryptophan residues that might be expected to undergo changes in environment, and hence fluorescence properties, upon association of LtrA with cognate RNA. Fluorescence was measured for the free LtrA protein and for varying ratios of RNA to protein at concentrations above the measured dissociation constant for the complex. Titratable RNA-dependent quenching of the intrinsic fluorescence of LtrA is observed, with saturation of the signal at a stoichiometric equivalence point of a 0.4 ± 0.1 RNA:protein ratio (Figure 8). These data are consistent with an RNP complex consisting of at least one RNA molecule per two protein molecules. Together, the limited proteolysis and fluorescence results are consistent with the idea that structural changes in LtrA accompany RNA binding.

CONCLUSIONS

Assembly of the L1.LtrB group II intron of *L. lactis* with its encoded LtrA protein produces an active ribonucleoprotein that catalyzes both RNA splicing and intron retrotransposition. To investigate the molecular interactions involved in complex formation, we analyzed the affinity and specificity of binding, as well as the structural changes that accompany RNP assembly. Equilibrium binding measurements of LtrA in complex with various intron-derived RNAs reveal a significant discrepancy between kinetically determined versus experimentally observed K_d values, consistent with a multistep, mutually induced fit mechanism for RNP assembly. LtrA is a highly positively charged protein in which ~1/6 of the residues are the basic amino acids lysine and arginine. Its rapid association with intron RNA is likely a result of initial strong but nonspecific electrostatic contacts, followed by a slower equilibration to the active structure of the complex. Equilibrium binding data for complexes that include intron domain DIV, the highest-affinity LtrA binding site within the intron, are described by a binding isotherm with an α_H of 1, while those missing DIV but containing DI are instead described by a Hill equation with an α_H of >1.5. This latter result implies a distinct mechanism for assembly of multiple LtrA proteins on the RNA when DI is the dominant site of LtrA interaction in the complex. Consistent with the multimeric state of LtrA, stoichiometric determinations for LtrA bound to DI, DIV, or the intact intron revealed a 2:1 ratio of protein to RNA at 450 mM NaCl, the previously determined condition for optimal maturase activity. The integrity of this complex was unchanged at 900 mM NaCl, implying that electrostatic interactions are not predominant in the complex.

The results of limited proteolysis experiments on LtrA indicate that the group II intron-encoded protein exists in at least three major conformational states. As a free protein or in the presence of noncognate RNA, LtrA adopts a flexible conformation that is completely susceptible to proteolytic digestion. In contrast, the RTα, RTβ, X, and ENDO domains of the protein undergo conformational rearrangement in the presence of DIV RNA, characterized by the complete stabilization of the RT domain with the concomitant appearance of hypersensitive sites. This specific stabilization and hypersensitivity of the protein are unique to DIV RNA, consistent with published work showing that DIV RNA contains the primary high-affinity binding site and is important for correct positioning of the RT domain during reverse transcription of the 3′ exon in the retro-homing pathway (*20*). In the presence of DI RNA or the intact intron, the protein is almost entirely protected from proteolysis; the limited cleavage observed occurs at the boundary between the RT and X domains, confirming that they are uniquely separate domains within the group II intron-encoded protein as has been observed for HIV-RT (*39*). The structural

conservation of the reverse transcriptase catalytic core is consistent with phylogenetic studies where it has been determined that the RT domain is the only component common in all non-LTR elements (*13*, *40*).

The data presented here as well as previous studies are consistent with a mutual induced fit model for binding of LtrA to the L1.ltrB group II intron in which initial binding of LtrA is achieved through formation of an encounter complex followed by specific recognition of properly formed RNA elements within DI and DIV. A dimerization model of LtrA provides a plausible mechanism for the recognition of two distinct intronic regions. Structural changes in LtrA evidenced by limited proteolysis and fluorescence quenching suggest that a significant contribution to the induced fit comes from an accompanying conformational change within the protein. Structural rearrangement of the intron may coincide with changes in LtrA structure, providing a plausible mechanism for folding the intron into a catalytically competent structure.

ACKNOWLEDGMENT

We thank Nigel Grindley and Enrique DeLa Cruz for help during the final stages of this project and Marlene Belfort and Alan Lambowitz for sharing unpublished data.

REFERENCES

1. Lazowska, J., Meunier, B., and Macadre, C. (1994) Homing of a group II intron in yeast mitochondrial DNA is accompanied by unidirectional co-conversion of upstream-located markers, *EMBO J. 13*, 4963−4972.
2. Moran, J. V., Mecklenburg, K. L., Sass, P., Belcher, S. M., Mahnke, D., Lewin, A., and Perlman, P. (1994) Splicing defective mutants of the COXI gene of yeast mitochondrial DNA: initial definition of the maturase domain of the group II intron aI2, *Nucleic Acids Res. 22*, 2057−2064.
3. Moran, J. V., Zimmerly, S., Eskes, R., Kennell, J. C., Lambowitz, A. M., Butow, R. A., and Perlman, P. S. (1995) Mobile group II introns of yeast mitochondrial DNA are novel site-specific retroelements, *Mol. Cell. Biol. 15*, 2828−2838.
4. Mills, D. A., Manias, D. A., McKay, L. L., and Dunny, G. M. (1997) Homing of a group II intron from *Lactococcus lactis* subsp. lactis ML3, *J. Bacteriol. 179*, 6107−6111.
5. Lambowitz, A. M., Caprara, M. G., Zimmerly, S., and Perlman, P. S. (1999) Group I and group II ribozymes as RNPs: clues to the past and guides to the future, in *The RNA world: the nature of modern RNA suggests a prebiotic RNA* (Atkins, J. F., Ed.) 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
6. Belfort, M., Derbyshire, V., Parker, M. M., Cousineau, B., and Lambowitz, A. (2002) Mobile Introns: Pathways and Proteins, in *Mobile DNA II* (Craig, N., Ed.) 2nd ed., ASM Press, Washington, DC.
7. Zhong, J., and Lambowitz, A. M. (2003) Group II intron mobility using nascent strands at DNA replication forks to prime reverse transcription, *EMBO J. 22*, 4555−4565.
8. Schafer, B., Gan, L., and Perlman, P. S. (2003) Reverse transcriptase and reverse splicing activities encoded by the mobile group II intron cobI1 of fission yeast mitochondrial DNA, *J. Mol. Biol. 329*, 191−206.
9. Noah, J. W., and Lambowitz, A. M. (2003) Effects of maturase binding and Mg^{2+} concentration on group II intron RNA folding investigated by UV cross-linking, *Biochemistry 42*, 12466−12480.
10. Huang, H. R., Chao, M. Y., Armstrong, B., Wang, Y., Lambowitz, A. M., and Perlman, P. S. (2003) The DIVa maturase binding site in the yeast group II intron aI2 is essential for intron homing but not for in vivo splicing, *Mol. Cell. Biol. 23*, 8809−8819.
11. Aizawa, Y., Xiang, Q., Lambowitz, A. M., and Pyle, A. M. (2003) The pathway for DNA recognition and RNA integration by a group II intron retrotransposon, *Mol. Cell 11*, 795−805.
12. Mohr, G., Perlman, P. S., and Lambowitz, A. M. (1993) Evolutionary relationships among group II intron-encoded proteins and identification of a conserved domain that may be related to maturase function, *Nucleic Acids Res. 21*, 4991−4997.
13. Malik, H. S., Burke, W. D., and Eickbush, T. H. (1999) The age and evolution of non-LTR retrotransposable elements, *Mol. Biol. Evol. 16*, 793−805.
14. Michel, F., and Lang, B. F. (1985) Mitochondrial class II introns encode proteins related to the reverse transcriptases of retroviruses, *Nature 316*, 641−643.
15. Xiong, Y., and Eickbush, T. H. (1990) Origin and evolution of retroelements based upon their reverse transcriptase sequences, *EMBO J. 9*, 3353−3362.
16. Xiong, Y., and Eickbush, T. H. (1988) Similarity of reverse transcriptase-like sequences of viruses, transposable elements, and mitochondrial introns, *Mol. Biol. Evol. 5*, 675−690.
17. Kennell, J. C., Moran, J. V., Perlman, P. S., Butow, R. A., and Lambowitz, A. M. (1993) Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria, *Cell 73*, 133−146.
18. San Filippo, J., and Lambowitz, A. M. (2002) Characterization of the C-terminal DNA-binding/DNA endonuclease region of a group II intron-encoded protein, *J. Mol. Biol. 324*, 933−951.
19. Matsuura, M., Noah, J. W., and Lambowitz, A. M. (2001) Mechanism of maturase-promoted group II intron splicing, *EMBO J. 20*, 7259−7270.
20. Wank, H., SanFilippo, J., Singh, R. N., Matsuura, M., and Lambowitz, A. M. (1999) A reverse transcriptase/maturase promotes splicing by binding at its own coding segment in a group II intron RNA, *Mol. Cell 4*, 239−250.
21. Koch, J. L., Boulanger, S. C., Dib-Hajj, S. D., Hebbar, S. K., and Perlman, P. S. (1992) Group II introns deleted for multiple substructures retain self-splicing activity, *Mol. Cell. Biol. 12*, 1950−1958.
22. Singh, R. N., Saldanha, R. J., D'Souza, L. M., and Lambowitz, A. M. (2002) Binding of a group II intron-encoded reverse transcriptase/maturase to its high affinity intron RNA binding site involves sequence-specific recognition and autoregulates translation, *J. Mol. Biol. 318*, 287−303.
23. Michel, F., and Ferat, J. L. (1995) Structure and activities of group II introns, *Annu. Rev. Biochem. 64*, 435−461.
24. Qin, P. Z., and Pyle, A. M. (1998) The architectural organization and mechanistic function of group II intron structural elements, *Curr. Opin. Struct. Biol. 8*, 301−308.
25. Peebles, C. L., Zhang, M., Perlman, P. S., and Franzen, J. S. (1995) Catalytically critical nucleotide in domain 5 of a group II intron, *Proc. Natl. Acad. Sci. U.S.A. 92*, 4422−4426.
26. Boudvillain, M., de Lencastre, A., and Pyle, A. M. (2000) A tertiary interaction that links active-site domains to the 5′ splice site of a group II intron, *Nature 406*, 315−318.
27. Matsuura, M., Saldanha, R., Ma, H., Wank, H., Yang, J., Mohr, G., Cavanagh, S., Dunny, G. M., Belfort, M., and Lambowitz, A. M. (1997) A bacterial group II intron encoding reverse transcriptase, maturase, and DNA endonuclease activities: biochemical demonstration of maturase activity and insertion of new genetic information within the intron, *Genes Dev. 11*, 2910−2924.
28. Bassi, G. S., and Weeks, K. M. (2003) Kinetic and thermodynamic framework for assembly of the six-component bI3 group I intron ribonucleoprotein catalyst, *Biochemistry 42*, 9980−9988.
29. Rose, M. A., and Weeks, K. M. (2001) Visualizing induced fit in early assembly of the human signal recognition particle, *Nat. Struct. Biol. 8*, 515−520.
30. Webb, A. E., Rose, M. A., Westhof, E., and Weeks, K. M. (2001) Protein-dependent transition states for ribonucleoprotein assembly, *J. Mol. Biol. 309*, 1087−1100.
31. Saldanha, R., Chen, B., Wank, H., Matsuura, M., Edwards, J., and Lambowitz, A. M. (1999) RNA and protein catalysis in group II intron splicing and mobility reactions using purified components, *Biochemistry 38*, 9069−9083.
32. Record, M. T., Jr., Lohman, M. L., and De Haseth, P. (1976) Ion effects on ligand-nucleic acid interactions, *J. Mol. Biol. 107*, 145−158.
33. Lohman, T. M., deHaseth, P. L., and Record, M. T., Jr. (1980) Pentalysine-deoxyribonucleic acid interactions: a model for the general effects of ion concentrations on the interactions of proteins with nucleic acids, *Biochemistry 19*, 3522−3530.

34. Katsamba, P. S., Myszka, D. G., and Laird-Offringa, I. A. (2001) Two functionally distinct steps mediate high affinity binding of U1A protein to U1 hairpin II RNA, *J. Biol. Chem*. *276*, 21476−21481.
35. Hall, K. B., and Stump, W. T. (1992) Interaction of N-terminal domain of U1A protein with an RNA stem/loop, *Nucleic Acids Res*. *20*, 4283−4290.
36. Weeks, K. M., and Crothers, D. M. (1992) RNA binding assays for Tat-derived peptides: implications for specificity, *Biochemistry 31*, 10281−10287.
37. Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., and Steitz, T. A. (1992) Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor, *Science 256*, 1783−1790.
38. Kumar, A., Kim, H. R., Sobol, R. W., Becerra, S. P., Lee, B. J., Hatfield, D. L., Suhadolnik, R. J., and Wilson, S. H. (1993) Mapping of nucleic acid binding in proteolytic domains of HIV-1 reverse transcriptase, *Biochemistry 32*, 7466−7474.
39. Lowe, D. M., Aitken, A., Bradley, C., Darby, G. K., Larder, B. A., Powell, K. L., Purifoy, D. J., Tisdale, M., and Stammers, D. K. (1988) HIV-1 reverse transcriptase: crystallization and analysis of domain structure by limited proteolysis, *Biochemistry 27*, 8884−8889.
40. Eickbush, T. H. (1997) Telomerase and retrotransposons: which came first? *Science 277*, 911−912.

BI049912U