RNA (2002), *8*:412–425. Cambridge University Press. Printed in the USA.

DOI: 10.1017.S1355838202029321

A novel mechanism for protein-assisted group I intron splicing

AMANDA SOLEM,[1,*] PIYALI CHATTERJEE,[1,*] and MARK G. CAPRARA[1,2]

[1]Center for RNA Molecular Biology, Case Western Reserve University, School of Medicine, Cleveland, Ohio 44106-4960, USA
[2]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, USA

ABSTRACT

Previously it was shown that the *Aspergillus nidulans* (*A.n.*) mitochondrial COB intron maturase, I-*Ani*I, facilitates splicing of the COB intron in vitro. In this study, we apply kinetic analysis of binding and splicing along with RNA deletion analysis to gain insight into the mechanism of I-*Ani*I facilitated splicing. Our results are consistent with I-*Ani*I and *A.n.* COB pre-RNA forming a specific but labile encounter complex that is resolved into the native, splicing-competent complex. Significantly, kinetic analysis of splicing shows that the resolution step is rate limiting for splicing. RNA deletion studies show that I-*Ani*I requires most of the *A.n.* COB intron for binding suggesting that the integrity of the I-*Ani*I-binding site depends on overall RNA tertiary structure. These results, taken together with the observation that *A.n.* COB intron lacks significant stable tertiary structure in the absence of protein, support a model in which I-*Ani*I preassociates with an unfolded COB intron via a "labile" interaction that facilitates correct folding of the intron catalytic core, perhaps by resolving misfolded RNAs or narrowing the number of conformations sampled by the intron during its search for native structure. The active intron conformation is then "locked in" by specific binding of I-*Ani*I to its intron interaction site.

Keywords: catalytic RNA; ribonucleoprotein assembly; RNA–protein interactions; RNA splicing

INTRODUCTION

Group I introns are a class of large RNAs that catalyze their own excision from precursor RNA (pre-RNA) in a process called self-splicing. Through base pairing and multiple long-range tertiary interactions, group I introns fold into a conserved, compact structure required for catalysis. Studies in vitro have shown that group I introns face numerous kinetic and thermodynamic barriers that can slow overall folding to biologically irrelevant time scales (reviewed in Woodson, 2000; Thirumalai et al., 2001; Treiber & Williamson, 2001). In addition, many group I introns only self-splice under nonphysiological conditions such as elevated temperature and high divalent metal ion concentrations suggesting the requirement for cofactors that facilitate RNA folding in vivo (see Coetzee et al., 1994). Indeed, group I introns, like many other functional RNAs, are associated with specific proteins that function to promote active RNA structure (Lambowitz & Perlman, 1990; Lambowitz et al., 1999). However, the mechanisms by which proteins recognize "unfolded" RNA ligands and induce a defined three-dimensional RNA conformation are not well understood. The studies of protein-assisted group I intron splicing continue to provide a framework from which to address these challenging questions.

The key to group I introns' catalytic function lies in their tertiary structure, which, for the group I intron conserved core, is known at fairly high resolution. The 5′ and 3′ splice sites (SS) are defined by pairing to an intron structure called the internal guide sequence, forming the P1 and P10 helices, respectively (see Fig. 1B). The catalytic core is made up of two extended helices called domains; the P4–P6 domain is formed by the stacking of P5, P4, P6, and P6a helices whereas the P3–P9 domain is formed by the stacking of P8, P3, P7, and P9 (Michel & Westhof, 1990; Golden et al., 1998). Precise packing of these two domains via tertiary interactions forms the intron catalytic center that is capable of binding an exogenous guanosine (used as a nucleophile in the first splicing step) and the splice-site con-

Reprint requests to: Mark G. Caprara, Center for RNA Molecular Biology, Case Western Reserve University, School of Medicine, 10900 Euclid Avenue, Cleveland, Ohio 44106-4960, USA; e-mail: mgc3@po.cwru.edu.
*These authors contributed equally to this work.

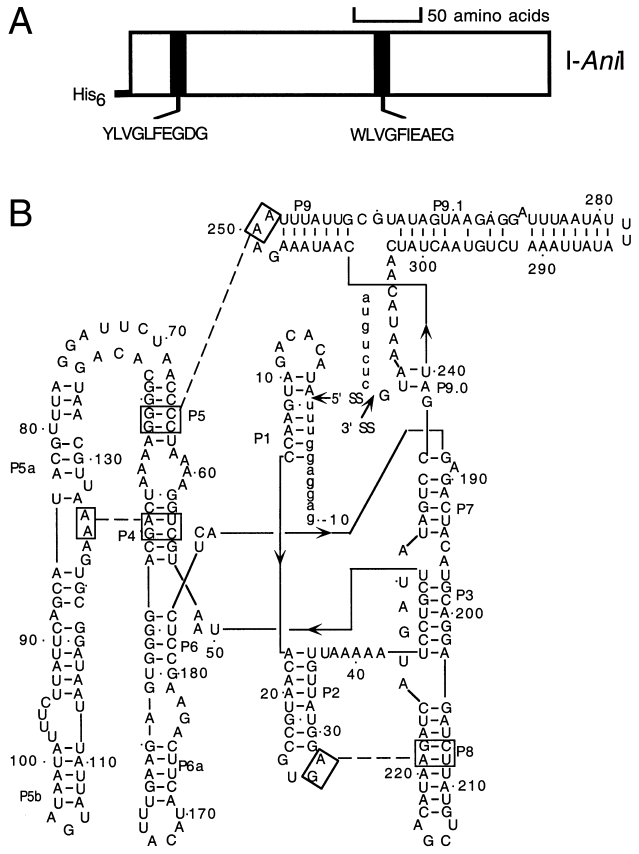

FIGURE 1. Schematic of I-*Ani*I and the secondary structure of *A.n.* COB pre-RNA. **A:** I-*Ani*I. The black boxes denote the conserved LAGLIDADG motifs. A vector tag of six histidines (His_6) at the N-terminus is used for purification. The protein schematic is drawn to scale; a marker with a length of 50 amino acids is shown. **B:** *A.n.* COB pre-RNA. Upper and lower case letters indicate intron and exon sequences, respectively. Arrows point to the splice sites. Predicted loop to helix tertiary interactions are shown by boxes connected by dashed lines (Michel & Westhof, 1990; Cate et al., 1996; Lehnert et al., 1996).

taining P1 and P10 helices (Michel & Westhof, 1990; Golden et al., 1998). Interconnecting peripheral structures surround the catalytic core and function to stabilize the active conformation, although their presence does not ensure self-splicing activity (Lehnert et al., 1996; Pan & Woodson, 1999; Engelhardt et al., 2000).

The identification of group I intron splicing cofactors has largely been through mitochondrial (mt) genetic analysis in fungi. These studies have shown that splicing cofactors include proteins either encoded by the introns themselves ("maturases") or various nuclear-encoded factors that often have an additional, RNA-related function in the cell (Lambowitz et al., 1999).

Two recombinant, nuclear-encoded proteins have been extensively studied in vitro; the yeast CBP2 protein that specifically promotes splicing of the fifth intron (bI5) in the mt cytochrome b (COB) gene (McGraw & Tzagoloff, 1983) and the *Neurospora* mt tyrosyl tRNA synthetase (CYT-18 protein) that promotes splicing of a variety of group I introns (Mannella et al., 1979; Wallweber et al., 1997). Importantly, these two unrelated proteins use fundamentally different mechanisms to facilitate splicing of similar group I introns.

CBP2 promotes bI5 splicing by a tertiary capture mechanism. Here, the secondary structure of the RNA is largely formed, but the RNA is otherwise unfolded. In the presence of low concentrations of Mg^{2+}, the catalytic core slowly folds, forming part of the CBP2 binding site in the P3–P9 domain and a peripheral structure, P7.1 (Shaw & Lewin, 1995; Weeks & Cech, 1995b). This is not a stable intron structure in low Mg^{2+}, but CBP2 binds strongly to "capture" the catalytically active intron (Weeks & Cech, 1995a, 1995b, 1996). A key biochemical signature in this mechanism is that the association rate of protein binding is limited by the equilibrium between the unfolded and folded states of the bI5 intron (Weeks & Cech, 1996).

In contrast, CYT-18-assisted splicing uses a scaffolding-like mechanism. Here the secondary structure of the P4–P6 domain is largely formed, but the rest of the intron structure is not stable. CYT-18 binds to the P4–P6 region, stabilizing the correct triple helical geometry at the junction of the P4 and P6 helices. This RNA/protein complex is thought to provide a scaffold for assembly of the P3–P9 domain that contains additional protein contact sites (Saldanha et al., 1995, 1996; Caprara et al., 1996a, 1996b, 2001). A key biochemical signature in this mechanism is stable protein binding to a single, folded intron domain that exists as an intermediate in the RNA/protein assembly process.

Compared to the CBP2 and CYT-18 examples, much less is known about the mechanism of maturase-assisted splicing. All identified maturases belong to a single family and promote splicing of the intron encoding them and, in one case, a very close relative (Lambowitz et al., 1999). For example, in the yeast mt cob gene, three group I introns encode maturases that are required for their own splicing (Lambowitz & Perlman, 1990). Introns that require maturases do not share obvious sequence or predicted structural similarities, suggesting that maturases recognize idiosyncratic features of their own unspliced RNAs.

All group I intron maturases known to date contain two copies of the amino acid motif LAGLIDADG. Other proteins that contain this motif are site-specific DNA endonucleases (Belfort & Roberts, 1997). The DNA endonucleases are located in group I and archeal introns, self-splicing protein introns ("inteins") and exist as freestanding nuclear encoded proteins (Dalgaard et al., 1997). Proteins in this class cleave long DNA target sites in a homologous allele that lacks the intron or intein and promote lateral transfer of their encoding intron or intein to these sites by a process called "homing" (Jurica & Stoddard, 1999; Chevalier & Stoddard, 2001). Some proteins possess maturase and DNA endonuclease activity, suggesting that the two activities are very closely associated and, furthermore, that ma-

turase function may be a secondary adaptation of these DNA endonucleases (see Lambowitz, 1989).

The first in vitro maturase system was developed only recently. Purified recombinant maturase protein encoded by the *Aspergillus nidulans* (*A.n.*) mt COB group I intron (I-*Ani*I) promoted splicing of the *A.n.* COB intron as well as exhibited site-specific DNA endonuclease activity (Ho et al., 1997). Kinetic measurements of splicing suggested that I-*Ani*I binds the COB intron tightly with an equilibrium constant (K_d) in the picomolar range (Ho et al., 1997; Ho & Waring, 1999). RNase T1 footprinting experiments provided evidence that I-*Ani*I facilitates folding of the *A.n.* COB intron, which has limited, if any, tertiary structure in the absence of protein (Ho & Waring, 1999). Intriguingly, *A.n.* COB peripheral domains that function to stabilize intron tertiary structure could not be deleted or truncated without significant decreases in protein-assisted splicing activity, suggesting that extensive RNA folding is required for protein recognition (Geese & Waring, 2001).

These observations raise the question of how I-*Ani*I recognizes the "unfolded" or "misfolded" COB intron and how binding is translated into conformational rearrangements that lead to a final native RNA structure. In this study, we apply kinetic analysis of binding and splicing along with RNA deletion analysis to gain insight into the mechanism of I-*Ani*I-facilitated splicing. Our results are consistent with I-*Ani*I and *A.n.* COB pre-RNA forming an initial, specific but labile encounter complex that is resolved into the native, splicing-competent complex. Kinetic analysis of splicing shows that the resolution step is rate limiting for splicing at neutral pH. RNA deletion studies show that I-*Ani*I binding requires most of the *A.n.* COB pre-RNA, suggesting that integrity of the I-*Ani*I binding site depends on the overall RNA tertiary structure. These results, taken together with the observation that the tertiary structure of *A.n.* COB intron is incomplete in the absence of protein, are incorporated in a model for I-*Ani*I function. First, I-*Ani*I preassociates with an unfolded COB intron producing a "labile" interaction that likely facilitates correct folding of the intron catalytic core, perhaps by resolving misfolded intermediates or narrowing the number of conformations sampled in the search for native RNA structure. The active intron conformation is then "locked in" by specific interaction of I-*Ani*I to its intron binding site.

RESULTS

Improved method for expression and purification of I-*Ani*I and construction of a suitable pre-RNA for structural analysis

Methods originally described for expression and purification of the recombinant N-terminal His-tagged I-*Ani*I from *Escherichia coli* yields a modest ~1 mg/L culture resulting in very dilute and unstable protein preparations (Ho et al., 1997 and M.G. Caprara, unpubl. observations). Because larger amounts of protein would expand the range of approaches that could be used to address fundamental questions in this system, we optimized the production of I-*Ani*I resulting in a final yield of ~10 mg protein/L culture (see Materials and Methods). The purified protein is active in *A.n.* COB pre-RNA splicing (see below) and site-specific DNA cleavage (P. Chatterjee, A. Solem, & M.G. Caprara, in prep.).

The original *A.n.* COB precursor RNA contains 112 nt of 5′ exon, 311 nt of intron, and 209 nt of 3′ exon. The large exon sizes create a cumbersome substrate for insightful in vitro structural analysis such as chemical structure and interference mapping. With this in mind, we constructed and characterized a pre-RNA with 5′ and 3′ exon sequences shortened to 22 and 19 nt, respectively. As shown below, the splicing properties of the mini-exon derivative (denoted *A.n.* COBme pre-RNA) are similar to those of the original construct (*A.n.* COB RNA; Ho & Waring, 1999; Geese & Waring, 2001).

I-*Ani*I-mediated group I intron splicing

All group I introns require divalent metal ions and a guanosine cofactor for splicing. The *A.n.* COB intron is able to self-splice providing that the Mg^{2+} concentration is at least 15 mM Mg^{2+} whereas the protein stimulates splicing at Mg^{2+} concentrations as low as 2 mM (Hur et al., 1997; Ho & Waring, 1999). We have found that 150 mM Mg^{2+} is optimal for self-splicing of *A.n.* COBme pre-RNA whereas protein-dependent splicing is relatively insensitive to Mg^{2+} between 5 and 25 mM Mg^{2+} (less than a twofold difference; data not shown). The self-splicing reaction at 150 mM Mg^{2+} for *A.n.* COBme pre-RNA shows that ~75% of the RNA undergoes the first step in splicing at a rate of 0.07 ($\pm$0.01) min^{-1} whereas the remainder is inactive (Fig. 2). In the presence of saturating concentrations of I-*Ani*I, at 5 mM Mg^{2+}, ~90% of the *A.n.* COBme pre-RNA splices at a rate of 2.1 ($\pm$0.48) min^{-1} and ~5% splices at ~0.003 min^{-1} (Fig. 2). The slow reacting species may represent a population of RNA that undergoes a slow folding step required for protein binding or RNA catalysis. These data show that I-*Ani*I binding causes a 30-fold rate enhancement of splicing relative to optimal self-splicing conditions.

The nature of the rate-limiting step for *A.n.* COBme pre-RNA splicing was investigated by pH titration experiments. As shown for a number of group I introns, pH dependence of the splicing reaction has revealed apparent pK_as close to neutrality (e.g. Herschlag & Khosla, 1994; Weeks & Cech, 1995a). Above pH 7–8, the reactions become independent of pH, consistent with the proposal that RNA conformational changes rather than the chemical steps of phosphodiester bond cleavage are rate limiting for splicing. To determine the pH

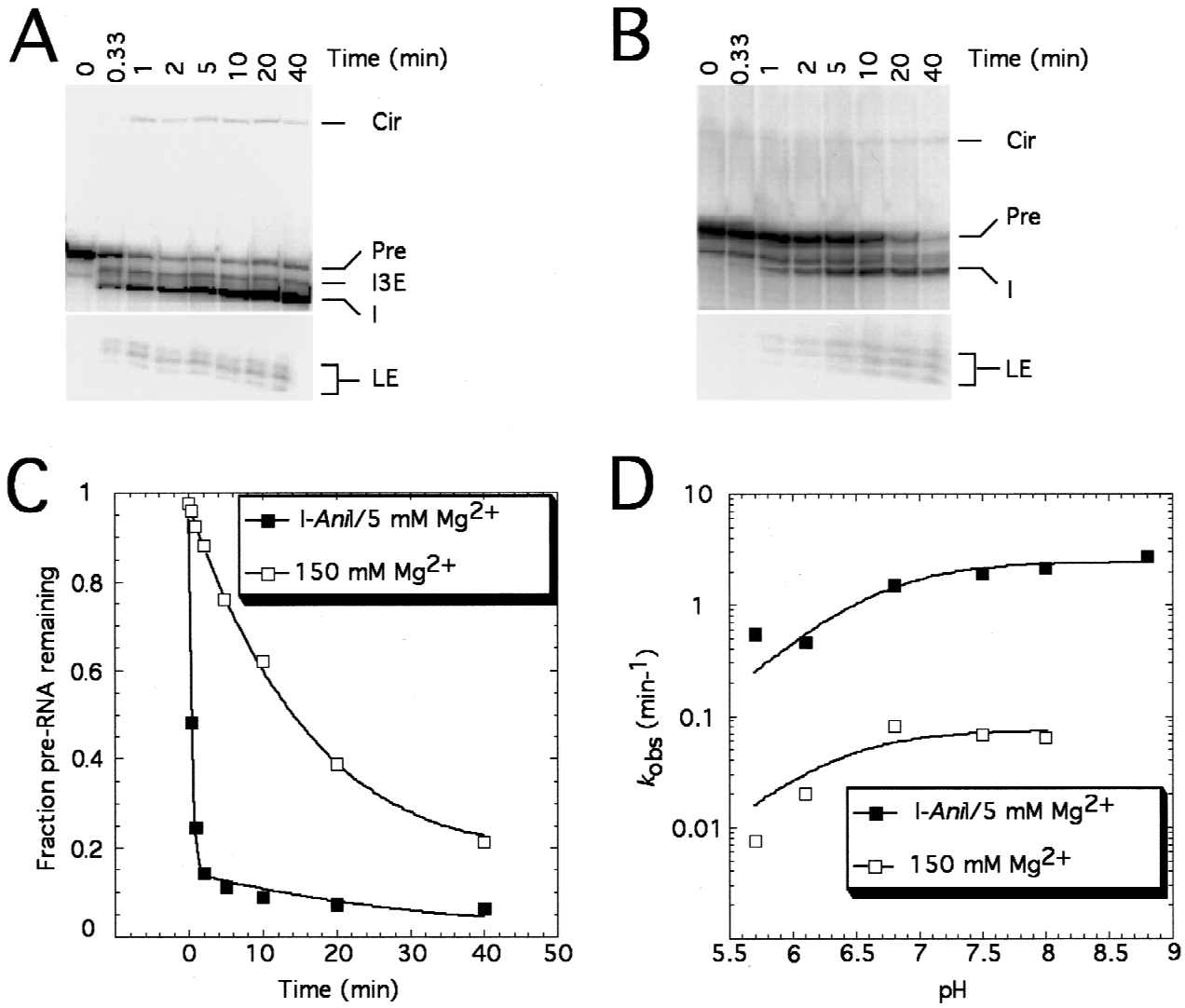

FIGURE 2. Protein-dependent and self-splicing of *A.n.* COBme pre-RNA. **A:** Time course of protein-assisted splicing. Internally ^{32}P-labeled *A.n.* COBme pre-RNA (7 nM) was preincubated in splicing buffer (50 mM Tris-HCl, pH 7.5, 100 mM NaCl; TN) supplemented with 5 mM $MgCl_2$ and 1 mM guanosine at 37 °C for 20 min and the reaction initiated by the addition of I-*Ani*I (14 nM final concentration). Aliquots were withdrawn at the indicated times and the products separated by denaturing gel electrophoresis. Control experiments showed that increasing I-*Ani*I concentration by twofold did not result in a rate increase, confirming that the protein is saturating under these conditions. Cir: circular form of the excised intron produced by attack of the intron's terminal guanosine to an intron residue near the 5′ end (the identity of this nucleotide has not yet been determined); Pre: precursor; I3E: splicing intermediate after 5′ SS cleavage with the 3′ exon still linked to the intron; I: free intron; LE: ligated exon. Note that the ligated exons appear as multiple bands due to 3′ end heterogeneity caused by in vitro transcription. **B:** Time course of *A.n.* COBme pre-RNA self-splicing. *A.n.* COBme pre-RNA (7 nM) was preincubated in 2× TN buffer supplemented with 300 mM $MgCl_2$ and at 37 °C for 20 min and the reaction initiated by the addition of an equal volume of guanosine (1 mM final concentration) yielding a final reaction mix of 1× TN buffer and 150 mM $MgCl_2$. **C:** Plot of the fraction of *A.n.* COBme pre-RNA spliced versus time. The data from **A** and **B** were quantified with a phosphorimager. The data for the protein-dependent reaction were fit to a double exponential with the fast phase describing 90% of the reactive RNA. The self-splicing reactions were fit to a single exponential. The average splicing rates for at least four independent determinations are 2.1 (±0.48) and 0.07 (±0.01) min^{-1} for I-*Ani*I-dependent and self-splicing, respectively. **D:** pH dependence of protein-facilitated and self-splicing reactions. Splicing reactions were carried out as for **A** and **B** and quantified as in **C**. The rate of self-splicing at pH 8.8 could not be determined due to extensive RNA hydrolysis at 150 mM Mg^{2+}. The data represent averages of at least two independent determinations. The error was less than 20% in all but one case (protein-dependent, pH 6.1, error was 30%). The data were fit to the equation, $k_{obs} = k_{max}/1 + 10^{(pKa-pH)}$ (Campbell et al., 2002).

dependence for *A.n.* COBme pre-RNA, splicing rates were measured under both protein-assisted and self-splicing conditions (Fig. 2D). Under both conditions, the apparent pK_a of the reaction is ~6.5. Importantly, both protein-assisted and self-splicing reactions are pH independent above pH 6.8, as was found for the original *A.n.* COB construct at 25 mM Mg^{2+} and under protein-assisted conditions (Ho & Waring, 1999). These

data suggest that both protein-assisted and self-splicing of the *A.n.* COBme pre-RNA at neutral pH are limited by changes in RNA conformation and thus the different splicing rates between the two conditions report variations in RNA folding.

Stoichiometry and stability of the I-*Ani*I/*A.n.* COBme pre-RNA complex

The mechanism of I-*Ani*I-facilitated *A.n.* COBme pre-RNA splicing is directly related to how the protein and RNA assemble into an active ribonucleoprotein (RNP) complex. As a first step to understand the assembly process, the stoichiometry of the complex as well as the kinetic constants for dissociation (k_{off}) and association (k_{on}) were determined by nitrocellulose filter binding assays. To determine binding stoichiometry, ^{32}P-labeled *A.n.* COBme pre-RNA (30 nM) was incubated with increasing concentrations of I-*Ani*I (2.5–75 nM) and filtered. A plot of the complex formed versus protein concentration was linear between 2.5 and 30 nM protein, yielding a slope of ~1, consistent with I-*Ani*I binding as a monomer in a 1:1 stoichiometry with the *A.n.* COBme pre-RNA (Fig. 3A).

To measure k_{off}, I-*Ani*I/*A.n.* COBme pre-RNA complexes were diluted in buffer containing an 800-fold excess of unlabeled (relative to labeled) *A.n.* COBme pre-RNA that functioned to bind unassociated protein. Control experiments showed that dilution and addition of unlabeled RNA simultaneously with protein was sufficient to reduce binding to undetectable levels throughout the experiment. The rate of dissociation best fit a double exponential with ~20% of the complexes dissociating quickly (~0.4 min^{-1}) whereas the vast majority dissociated with a slow rate of 0.012 (±0.003) min^{-1} (Table 1; Fig. 3B). The slow rate is within twofold of the I-*Ani*I turnover rate (~0.02 min^{-1}) measured in multiple turnover splicing reactions consistent with dissociation being quite slow (data not shown; see also Ho et al., 1997). The RNA species in the initial fast rate presumably represents a fraction of misfolded RNA that does not strongly associate with I-*Ani*I.

Evidence for a multistep mechanism for I-*Ani*I/*A.n.* COBme pre-RNA complex formation

To measure the association rate, trace amounts of ^{32}P-labeled RNA (50 pM) were mixed with several different concentrations of I-*Ani*I (2–32 nM) and complex formation was quenched at different times by addition of a 1,000-fold excess of unlabeled (relative to labeled) *A.n.* COBme pre-RNA followed by filtration (Riggs et al., 1970). At all protein concentrations, the data best fit a single exponential, demonstrating that binding showed good pseudo-first-order behavior (Fig. 4A). For a simple bimolecular association, a plot of the measured association rate versus I-*Ani*I concentration is expected to be linear with the slope related to k_{on}. However, the data clearly show a hyperbolic relationship (Fig. 4B).

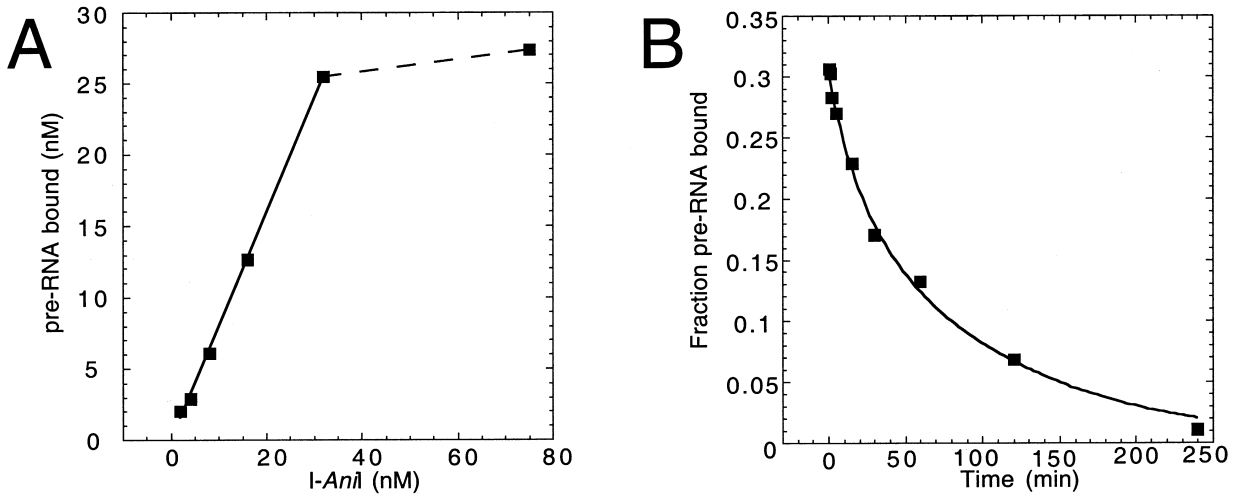

FIGURE 3. I-*Ani*I binding to *A.n.* COBme pre-RNA. **A:** Binding stoichiometry. Stoichiometry of the complex was determined by incubating increasing concentrations of I-*Ani*I (2.5–75 nM) with 32 nM ^{32}P-labeled *A.n.* COBme pre-RNA for 5 min prior to filtration. The data were corrected for the efficiency of nitrocellulose detection (30%, see Materials and Methods and Yarus & Berg, 1970). The slope of the solid line is 0.8. The average slope in three independent experiments is 0.9 (±0.1) indicating that I-*Ani*I binds *A.n.* COBme pre-RNA in a 1:1 stoichiometry. **B:** k_{off} determination. Dissociation of the *A.n.* COBme pre-RNA/I-*Ani*I complex was initiated by dilution in RNA-binding buffer containing an 800-fold molar excess of unlabeled to labeled *A.n.* COBme pre-RNAs followed by filtration at the times indicated. The data were fit to a double exponential with 20% of the complexes dissociating quickly and the remainder much more slowly. The fast rate presumably represents an RNA population that does not bind I-*Ani*I tightly (see text for details). The k_{off}, determined from the slow phase of four independent experiments, is 0.012 (±0.003) min^{-1}.

TABLE 1. Dissociation kinetics of *A.n.* COBme and mutant pre-RNAs

RNA	Fraction	Rate (min^{-1})
A.n. COBme	0.8 (±0.1)	0.012 (±0.003)
Δ5E	0.8 (±0.3)	0.023 (±0.016)
Δ P1	0.3 (±0.2)	0.090 (±0.012)
Δ P2	0.5 (±0.3)	0.23 (±0.15)
Δ 3E	0.8 (±0.1)	0.040 (±0.028)

Dissociation of the protein/RNA complexes in all cases showed two phases. The fraction in the slow (specific) phase and the corresponding dissociation rates (k_{off}) are shown. The values represent averages from at least three independent experiments and the standard deviations are shown in parentheses.

The behavior of the association kinetics is consistent with binding occurring in at least two steps. In one model, I-*Ani*I, free RNA, and an initial ("encounter") complex come to a rapid equilibrium that is followed by a relatively slow conformational change:

$$\text{I-}Ani\text{I} + \text{tRNA} \overset{K_1}{\rightleftharpoons} \text{I-}Ani\text{I/RNA} \xrightarrow{k_2} \text{I-}Ani\text{I/RNA*}.$$

In this model, the second (or native) complex, I-*Ani*I/RNA*, represents the species detected by nitrocellulose filter binding. The "encounter" complex, I-*Ani*I/RNA, is a labile complex that is either retained at a low efficiency on the filter or is not detected by filter binding due to its rapid dissociation and binding of free protein by the unlabeled RNA in the reaction. The concentration dependence of the reaction to form the second complex reflects saturation of the initial complex; at the upper limit the rate approaches an isomerization constant k_2 (Johnson, 1992; Fierke & Hammes, 1995; Sullivan et al., 1997). The data were an excellent fit (R value 0.99) to the equation that describes this model: $k_{obs} = K_1 k_2[\text{I-}Ani\text{I}]/(K_1[\text{I-}Ani\text{I}] + 1)$, where K_1 equals 7.7×10^7 M^{-1}, k_2 is equal to 2.7 min^{-1}, and k_{on}, which equals $K_1 k_2$, is 2.2×10^8 M^{-1} min^{-1}. In contrast, the data were a relatively poor fit (R value 0.89) to an equation that describes an alternative model in which a conformational change in the RNA is required prior to protein binding (data not shown; Fersht & Requena, 1971). The K_d for the native complex determined by dividing k_{off}/k_{on} is ~50 pM.

The rate of conversion from intermediate to stable complex, k_2, is equal to ~3 min^{-1}, which is very near the rate measured for the protein-assisted splicing reactions under saturating protein concentrations (2.1 min^{-1} (±0.48); see above). Because the pH profile of the protein-assisted splicing reactions suggested that the observed splicing rate represents a conformational change rather than a chemical step in splicing, we investigated whether these two measured values represent the same step in RNA/protein assembly (i.e., does $k_2 = k_{splice}$?). If splicing reports the same step, then a plot of reaction rate as a function of I-*Ani*I concentration should be similar to the plot of association rates. Again, using trace amounts of ^{32}P-labeled RNA (50 pM), splicing reactions were initiated by addition of I-*Ani*I at the same concentrations used in the binding experiments (2–32 nM; Fig. 4C,D). As for the association kinetics, under these conditions the splicing rate showed good pseudo-first-order behavior with respect to I-*Ani*I concentration. As expected from our earlier analysis, the data from individual reactions fit a double exponential with ~90% of the active RNA reacting in a fast phase. A plot of these data versus I-*Ani*I concentration again showed a clear hyperbolic relationship. The splicing data were also an excellent fit to the equation $k_{obs} = K_1 k_2[\text{I-}Ani\text{I}]/(K_1[\text{I-}Ani\text{I}] + 1)$ and revealed a K_1 of 6.0 (± 1.1) $\times 10^7$ M^{-1}, a k_2 of 2.4 (±0.14) min^{-1} and k_{on} of 1.4 (± 0.16) $\times 10^8$ M^{-1} min^{-1}, the same values, within error, for those determined for binding (see Table 2 for comparison).[1] Taken together, the kinetic data suggest that the splicing reaction reports the rate of conversion for the I-*Ani*I/*A.n.*COBme pre-RNA encounter complex to a native, splicing-competent complex, I-*Ani*I/*A.n.*COBme pre-RNA*.

The existence of an intermediate ribonucleoprotein in the assembly of I-*Ani*I and *A.n.*COBme pre-RNA raises the question of whether this complex consists entirely of nonspecific interactions. The K_d of the encounter complex is quite low ($K_d = 1/K_1 = 13$ nM) relative to I-*Ani*I's nonspecific RNA affinity as measured for other non-self-splicing group I introns and an RNA hairpin (K_ds in the micromolar range; data not shown). The more tightly bound RNA in the encounter complex may be the result of I-*Ani*I making a subset of specific interactions with *A.n.* COBme pre-RNA that are present in the native complex.

I-*Ani*I recognition requires the catalytic core and peripheral structures for binding *A.n.* COBme pre-RNA

The specificity and mechanism with which I-*Ani*I binds *A.n.* COBme pre-RNA suggests that I-*Ani*I may recognize idiosyncratic features of the *A.n.* COBme pre-RNA. To determine the RNA requirements for binding, a series of 5′ and 3′ truncations of the *A.n.* COBme pre-RNA were constructed and binding affinities measured kinetically. The location of each truncation is shown schematically in Figure 5A. The effect on binding for each construct is discussed first in terms of their

[1] Earlier estimates of k_{on} using splicing as an assay were ~20-fold faster and the hyperbolic nature of the curve was not considered in the analysis (Ho & Waring, 1999; R.B. Waring, pers. comm.). The faster k_{on} estimate relative to the one reported here is most likely due, in part, to the imprecision in measuring fast reactions rates, the analysis of a small range of protein concentrations that excluded the hyperbolic nature of the splicing behavior, and the possibility that, under the previous conditions used, the isomerization step may be ~10-fold faster (R.B. Waring, pers. comm.).

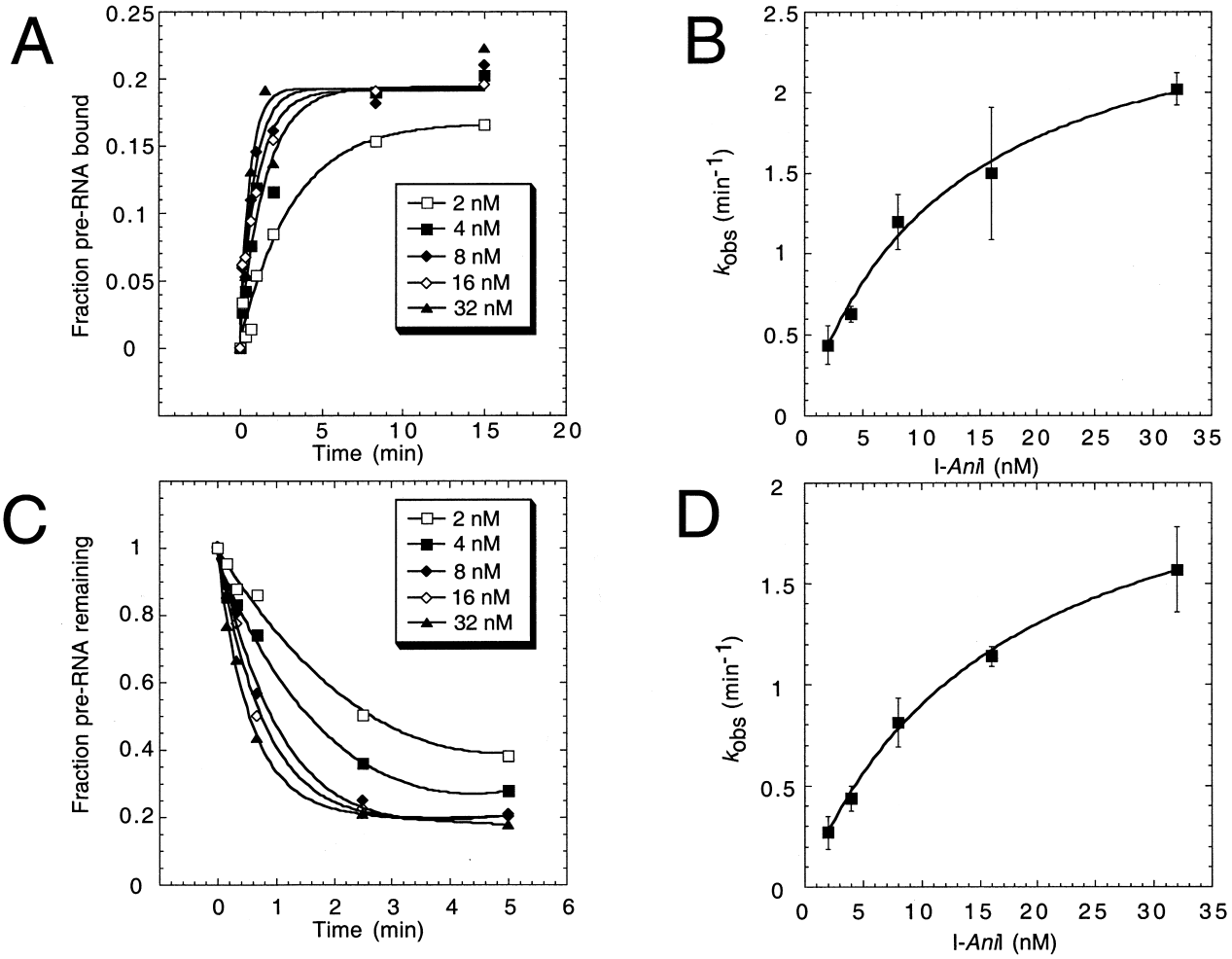

FIGURE 4. Monitoring the assembly of I-*Ani*I with *A.n.* COBme pre-RNA by association and splicing kinetics. **A:** Association kinetics. Trace amounts of ^{32}P-labeled RNA (50 pM) were incubated with a range of I-*Ani*I concentrations and, at the times indicated, the reactions were stopped by the addition of a ~1000-fold excess of unlabeled to labeled *A.n.* COBme pre-RNA and immediately filtered. The data were fit to a single exponential to obtain the pseudo-first-order rate constant, k_{obs}. **B:** Plot of the observed association rates as a function of I-*Ani*I concentration. The data are averages from three independent experiments and the error bars represent the standard deviation from the mean. The data are consistent with a model in which binding occurs in two steps: first an equilibrium (K_1) between the free RNA, protein, and an "encounter complex" is rapidly established followed by a slower unimolecular isomerization (k_2) of the complex to the native complex. The data were fit to the following equation that describes this model: $k_{obs} = K_1 k_2[\text{I-}Ani\text{I}]/(K_1[\text{I-}Ani\text{I}] + 1)$ where K_1 equals 7.7 ($\pm$2.6) $\times$ 10^7 M^{-1}, k_2 is equal to 2.7 ($\pm$0.23) min^{-1} and k_{on}, which equals $K_1 k_2$, is 2.2 ($\pm$0.62) $\times$ 10^8 M^{-1} min^{-1}. **C:** *A.n.* COBme pre-RNA splicing kinetics at various I-*Ani*I concentrations. Trace amounts of ^{32}P-labeled pre-RNA (50 pM) were incubated with 1 mM guanosine cofactor and splicing initiated by addition of I-*Ani*I (concentrations ranged from 2–32 nM). **D:** Plot of *A.n.* COBme pre-RNA splicing rate as a function of I-*Ani*I concentration. The data are averages from three independent experiments and the error bars represent the standard deviation from the mean. The curve is essentially the same as that observed in the binding experiments in **B** and the data were fit to the same equation. For the splicing experiments, K_1 equals 6.0 ($\pm$ 1.1) $\times$ 10^7 M^{-1}, k_2' is equal to 2.4 min^{-1} ($\pm$ 0.12), k_{on}, which equals $K_1 k_2$, is 1.4 ($\pm$ 0.17) $\times$ 10^8 M^{-1} min^{-1}. For **B** and **D**, errors for the constants are based on variations of fits to the data excluding any single point (see McConnell & Cech, 1995).

dissociation rates and the fraction of specifically bound RNA (Table 1; Fig. 5B). Starting from the 5′ end, deletion of the 5′ exon (Δ5E) increased k_{off} by 2-fold, whereas removal of the remainder of the P1 helix (ΔP1) decreased the fraction of specifically bound RNA to 30% and increased k_{off} by ~8-fold. A further truncation that deleted the adjacent P2 helix (ΔP1/P2) also decreased the fraction of RNA that was specifically bound and dramatically increased k_{off} by ~20-fold. Further deletion of the 5′ strand of P3 (ΔP1–P3) eliminated detectable binding. From the 3′ end, deleting the 3′ exon (Δ3E) increased k_{off} by ~3-fold, whereas elimination of the 3′ strand of the peripheral P9.1 helix abolished binding. Taken together, these data suggest that I-*Ani*I requires the entire catalytic core (i.e., P3–P9) and some peripheral structures for specific intron binding.

The effect on k_{on} for each mutant is shown in Figure 5C with values reported in Table 2. No complex formation for mutant ΔP1/P2 was detected under these

TABLE 2. Association kinetics of *A.n.* COBme and mutant pre-RNAs

RNA	K_1 ($\times 10^7$ M^{-1})	k_2' (min^{-1})	k_{on} ($\times 10^8$ M^{-1} min^{-1})
A.n. COBme	7.7 (±2.6)	2.7 (±0.23)	2.2 (±0.62)
Δ5E	8.4 (±2.3)	3.0 (±0.29)	2.5 (±0.89)
Δ P1	11 (±0.70)	2.6 (±0.23)	2.9 (±0.23)
Δ P2	n.d.	n.d.	n.d.
Δ 3E	8.1 (±3.3)	3.3 (±0.20)	2.7 (±0.84)

Plots of association rate versus protein concentration were fit to an equation that describes a two-step binding process. K_1 represents a rapid equilibrium between free protein, RNA, and the first complex. k_2 is the first-order rate constant describing the conformational change from the first to the second complex. The association rate, k_{on}, is equal to $K_1 k_2$. Errors are based on variations of the fit to the data excluding any single point. n.d.: not determined due to complex instability.

conditions except at the highest protein concentration and after long incubation times consistent with the results of k_{off} experiments that showed I-*Ani*I does not bind this RNA tightly. The remaining mutants showed the same hyperbolic concentration dependence for binding as the wild-type *A.n.* COBme pre-RNA (Fig. 5C). The values for K_1 and k_2' and k_{on} are shown in Table 2 and are essentially the same, within error, as the wild-type.

DISCUSSION

The goal of the current work was to develop a versatile experimental model system to study how group I intron-encoded maturases recognize their cognate intron and facilitate splicing. Our work began with the in vitro system first developed by Waring and colleagues who showed that the *Aspergillus* mt COB intron encodes an RNA maturase (I-*Ani*I) that is required for its splicing at physiological concentrations of Mg^{2+}. We have adapted the in vitro system to be more accessible to structural studies that will be implemented in the future. In work presented here, we have shown that the protein binds to the COB pre-RNA in a 1:1 stoichiometry, that the complex is extremely stable, and that the protein readily discriminates against noncognate RNAs. Association kinetics of protein/RNA complex formation and splicing revealed that I-*Ani*I binding is a multistep process that includes an intermediate "encounter" complex that is relatively slowly resolved into a native, splicing-competent complex. Deletion analysis shows that protein does not bind to an independently folded domain(s) of the *A.n.* COB intron but instead requires most of the *A.n.* COB catalytic core and peripheral domains for binding.

I-*Ani*I maturase activity

We have optimized the overexpression and purification of recombinant I-*Ani*I to facilitate in vitro structural studies. In particular, the yield of pure protein is now sufficient for such widely used and informative spectral approaches such as CD and fluorescence spectroscopy as well as structural approaches such as X-ray crystallography. We have also generated a shortened form of the COB pre-RNA that should facilitate these studies as well as the use of chemical cleavage as a probe for RNA structure.

The splicing activity of *A.n.* COBme pre-RNA resembles that of the original construct. Self-splicing is detected only at Mg^{2+} concentrations of 15 mM or greater and the rate of splicing at 150 mM Mg^{2+} is ~6-fold slower than that reported for the original construct (Ho & Waring, 1999). The protein-dependent splicing rate at 5 mM Mg^{2+} is ~10-fold slower than that estimated for the original construct using 1 mM guanosine (~20 min^{-1}; Geese & Waring, 2001). Importantly, pH titration experiments show that both constructs in protein-dependent and self-splicing reactions are pH independent above pH 7, presumably reflecting that conformational changes, not chemistry, are rate limiting for activity under these conditions.

The differences in rate between the two constructs may be due to a stimulatory effect of longer exon sequences on intron splicing. Precedent for this possibility has been found for the self-splicing *Tetrahymena* LSU intron in which structures within the natural 5′ exon have been found to positively influence the folding rate of the catalytic core (Pan & Woodson, 1998). Although this may be the case, the splicing rate of the I-*Ani*I/*A.n.* COBme pre-RNA complex is 2- and 14-fold faster than the CBP2 and CYT-18-assisted intron splicing reactions, respectively, providing evidence that the maturase remains a very efficient cofactor for group I intron splicing (Saldanha et al., 1995; Weeks & Cech, 1995a).

Properties of the I-*Ani*I/*A.n.* COBme pre-RNA complex

Stoichiometry binding experiments show that I-*Ani*I binds in a 1:1 molar ratio with the *A.n.* COBme pre-RNA, demonstrating that the protein functions as a monomer. This result is consistent with active-site titration splicing experiments that showed the same stoichiometry (data not shown; Ho et al., 1997). The functional quaternary structure of RNA maturases is therefore similar to that of the related LAGLIDADG homing endonucleases. The DNA endonucleases that contain two copies of the LAGLIDADG motif are monomers whereas proteins with a single copy of the motif are dimers (Jurica & Stoddard, 1999; Chevalier & Stoddard, 2001).[2] I-*Ani*I, which has two copies of the motif,

[2]One exception is I-*Sce*II, encoded by yeast mt aI4α intron that contains two copies of the motif but exists as a dimer. It is not known if the dimer cleaves one or two substrates (Wernette et al., 1990).

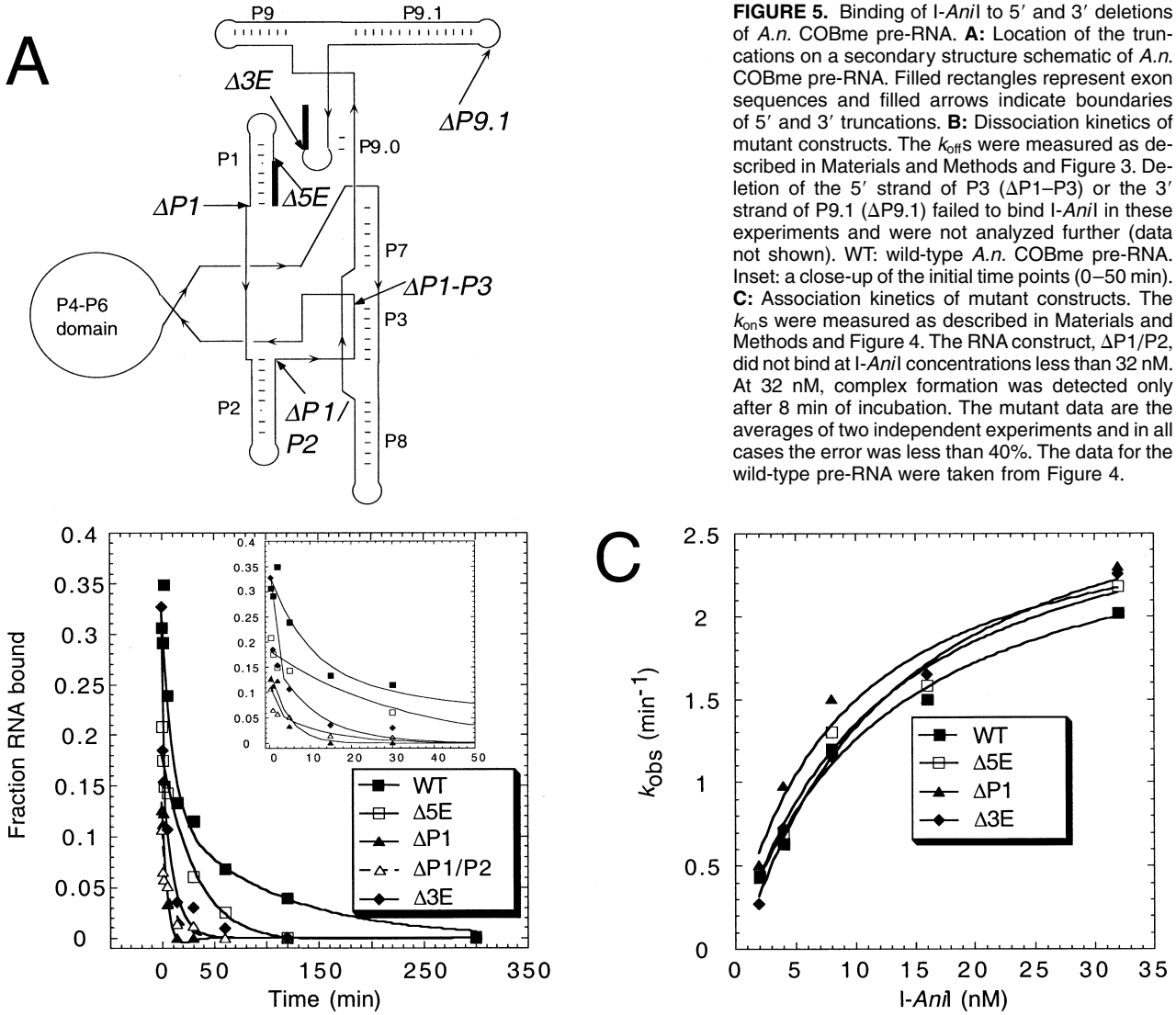

FIGURE 5. Binding of I-*Ani*I to 5′ and 3′ deletions of *A.n.* COBme pre-RNA. **A:** Location of the truncations on a secondary structure schematic of *A.n.* COBme pre-RNA. Filled rectangles represent exon sequences and filled arrows indicate boundaries of 5′ and 3′ truncations. **B:** Dissociation kinetics of mutant constructs. The k_{off}s were measured as described in Materials and Methods and Figure 3. Deletion of the 5′ strand of P3 (ΔP1–P3) or the 3′ strand of P9.1 (ΔP9.1) failed to bind I-*Ani*I in these experiments and were not analyzed further (data not shown). WT: wild-type *A.n.* COBme pre-RNA. Inset: a close-up of the initial time points (0–50 min). **C:** Association kinetics of mutant constructs. The k_{on}s were measured as described in Materials and Methods and Figure 4. The RNA construct, ΔP1/P2, did not bind at I-*Ani*I concentrations less than 32 nM. At 32 nM, complex formation was detected only after 8 min of incubation. The mutant data are the averages of two independent experiments and in all cases the error was less than 40%. The data for the wild-type pre-RNA were taken from Figure 4.

fits this trend well as it functions as a monomer when binding to or cleaving its DNA substrate (P. Chatterjee, A. Solem, & M.G. Caprara, in prep.). Taken together, the RNA and DNA substrate data provide evidence that the functional quaternary structure of I-*Ani*I is not changed when carrying out either of its unrelated activities.

The I-*Ani*I/AnCOB pre-RNA complex is very stable, characterized by a half-life of approximately 1 h. Slow dissociation of group I intron cofactors is a common theme with both CBP2 and CYT-18 having half-lives of ~75 min and 100 h, respectively (Saldanha et al., 1995; Weeks & Cech, 1996). Such tight binding is apparently required to stabilize active RNA conformations over competing, presumably near isoenergetic structures (Weeks, 1997). Despite this similarity, CBP2, CYT-18, and, as we show here, I-*Ani*I use different mechanisms to facilitate intron splicing.

Mechanism of I-*Ani*I association with *A.n.* COBme pre-RNA

The kinetic analysis of binding shows that I-*Ani*I intron recognition proceeds through a fundamentally different mechanism than either CBP2 or CYT-18. CYT-18 binds strongly to the P4–P6 domain structure at a diffusion-controlled rate (Saldanha et al., 1995, 1996; Caprara et al., 2001). Once bound to this structure, the P3–P9 domain folds and associates with other protein and RNA interaction sites. On the other hand, the association rate of CBP2 is limited by the formation of a transient bI5 intron tertiary structure that must occur prior to protein binding (Weeks & Cech, 1996; Webb & Weeks, 2001). I-*Ani*I, by contrast, appears to bind through an initial, labile, encounter complex that is resolved into a stable "native" complex. The finding that the K_d for the encounter complex is much lower than

the K_d of I-*Ani*I for nonspecific RNAs suggests that this complex may be structurally specific. A structurally specific, but unstable intermediate has also been detected in the assembly of the human signal recognition particle by time-resolved RNA footprinting techniques (Rose & Weeks, 2001). The finding that protein-dependent splicing of *A.n.* COBme pre-RNA is limited by the rate of transition from the encounter to native complex shows that this process is a key step in the protein activation of splicing.

The mechanism of I-*Ani*I binding raises a question about the nature of the conformational changes that occur in the RNP. In principle, these could reflect changes in the protein, RNA, or both. RNA-structure-mapping experiments with RNase T1 showed that secondary structure in the substrate P1 helix, the core elements P3, P7, P9, and peripheral helices P5a and P9.1 are not stable in 5 mM Mg^{2+}, but become protected in the presence of 25 mM Mg^{2+} or by binding of I-*Ani*I (Ho & Waring, 1999). Structural instability of these critical folding elements at 5 mM Mg^{2+} may indicate that little tertiary structure exists in the COB pre-RNA in the absence of protein. Consistent with this possibility, the use of hydroxyl radical cleavage as a probe for RNA tertiary structure has failed to reveal higher order folding of *A.n.* COBme pre-RNA at 5 mM Mg^{2+} (A. Solem & M.G. Caprara, in prep.). The finding that binding of I-*Ani*I induces structural rearrangements in COB pre-RNA suggests that RNA folding is an integral component of the conformational changes detected in the binding experiments.

The truncation experiments indicate that I-*Ani*I requires most of the intact COB intron for binding. Truncation of the intron that deletes the 5′ strand of P3 (ΔP1-P3) or within the P9.1 (ΔP9.1) peripheral helix abolished detectable binding. In contrast, deletion of all P1 or the 3′ exon resulted in a slightly higher rate of dissociation relative to the wild type but had no effect on RNA/protein assembly. These data suggest that these elements make only a modest contribution to binding affinity. A further truncation of sequence that deleted the P2 helix (ΔP1/P2) reduced k_{on} and k_{off} considerably, which may indicate that P2 forms part of the I-*Ani*I binding site. However, the presence of P2 does not ensure binding as the ΔP9.1 mutation, which contains P2, failed to bind I-*Ani*I. Previous analysis of the effect of internal deletions on protein-dependent splicing also showed that the RNA was extremely sensitive to mutation, leading to the suggestion that the protein requires an intact intron for recognition (Geese & Waring, 2001). Taken together, these observations suggest that I-*Ani*I binds to regions of the COB pre-RNA that are far apart in primary sequence but brought together in close proximity in the folded core. Thus, unlike CYT-18 that binds specifically to the intron's P4–P6 domain, I-*Ani*I does not bind to an independently folded domain in the catalytic core. Additionally, the finding that deletion of PI either alone or in combination with P2 decreases the fraction of RNA that can bind protein suggest that these elements may play a key role in intron folding events required for I-*Ani*I recognition. Both P1 and P2 are expected to associate with the intron core during 5′SS docking (Costa & Michel, 1995), and it is tempting to speculate that this step facilitates RNA folding events required for I-*Ani*I binding.

A "preassociation" model for I-*Ani*I facilitation

The current observations can be integrated in a model for I-*Ani*I function. The RNA deletion analysis shows that protein binding is extremely sensitive to RNA alterations, suggesting that the integrity of the I-*Ani*I-binding site depends on overall intron tertiary structure. Because structure-mapping experiments have shown that the core is not completely folded in the absence of protein, an important question is raised: how is the I-*Ani*I complex binding site formed? Association and splicing kinetics were inconsistent with a unimolecular conformational change occurring in the RNA prior to binding, as is the case for CBP2. One hypothesis for RNA folding takes into account the encounter complex that is in rapid equilibrium between the free protein and RNA prior to consolidation to the native complex (see Fig. 6). In this view, I-*Ani*I may initially bind the RNA via nonspecific or a subset of specific interactions to form this labile collision complex. A consequence of this interaction may be to facilitate correct folding of COB pre-RNA, perhaps by destabilizing misfolded conformations of the RNA or by restricting the number of conformations sampled by the intron during folding. Thus, the RNA in the complex undergoes conformational rearrangements that lead to its native structure (and formation of the I-*Ani*I binding site) that is then "locked in" by specific interactions with the protein. This "preassociation" binding mechanism was originally proposed by Herschlag (1995) as a mechanism by which specific RNA-binding proteins could stabilize RNA folding by both facilitating correct folding and stabilizing tertiary structure. Future work will be focused on establishing the RNA structure in both the encounter and native complexes.

MATERIALS and METHODS

Plasmids and oligonucleotides

The plasmid pEC was used to express the I-*Ani*I protein and contains the 880-nt open reading frame cloned into the *Bam*HI/*Hin*dIII sites of the expression vector pET28b (Novagen, Madison, Wisconsin) under the control of a T7 RNA polymerase promoter (Ho et al., 1997). This construct produces a protein with a 6× histidine tag at the N-terminus that is used for

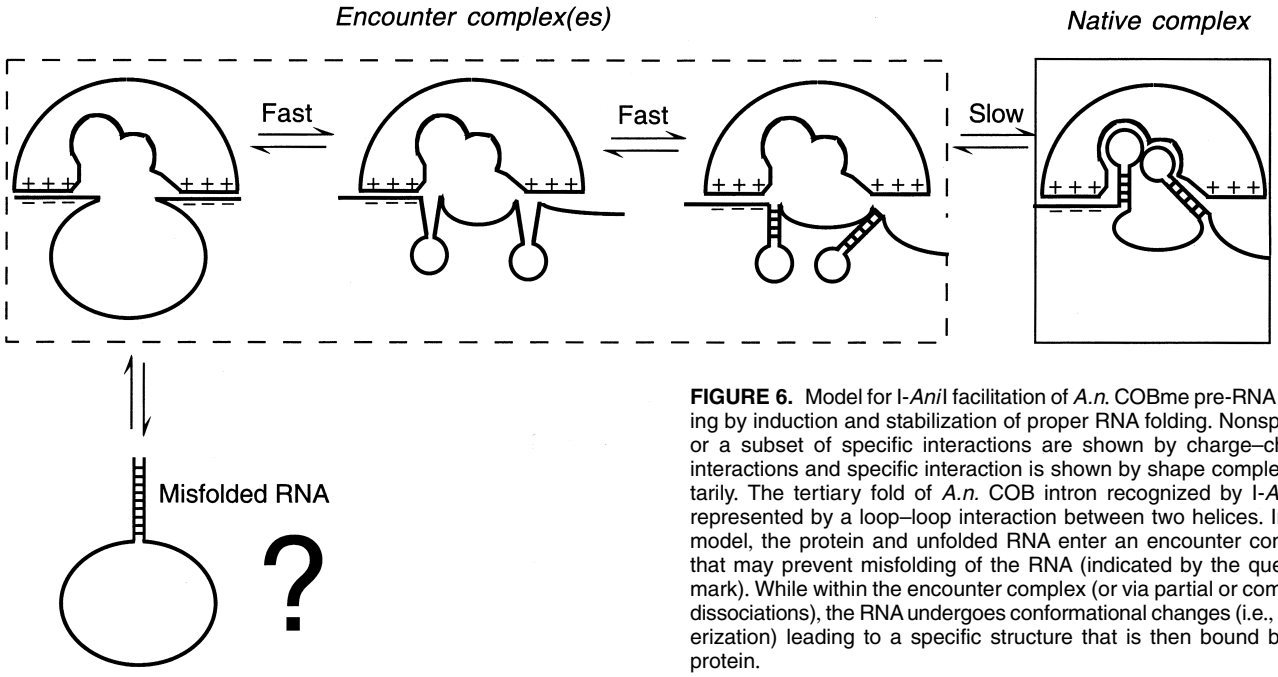

FIGURE 6. Model for I-*Ani*I facilitation of *A.n.* COBme pre-RNA splicing by induction and stabilization of proper RNA folding. Nonspecific or a subset of specific interactions are shown by charge–charge interactions and specific interaction is shown by shape complementarily. The tertiary fold of *A.n.* COB intron recognized by I-*Ani*I is represented by a loop–loop interaction between two helices. In this model, the protein and unfolded RNA enter an encounter complex that may prevent misfolding of the RNA (indicated by the question mark). While within the encounter complex (or via partial or complete dissociations), the RNA undergoes conformational changes (i.e., isomerization) leading to a specific structure that is then bound by the protein.

affinity chromatography. Plasmid pT7AnCOBme was used to generate *A.n.* mt ΔORF COBme intron-containing precursor RNA. This construct contains a T7 RNA polymerase promoter, 22 nt of 5′ exon, 311 nt of intron, and 15 nt of 3′ exon cloned into the *Eco*RI and *Bam*HI sites of pUC19 (New England Biolabs, Beverly, Massachusetts). This construct was made by amplification of the plasmid pCOBMEsal (Hur et al., 1997), which contains the ΔORF COB intron and flanking exon using the polymerase chain reactions (PCR) and oligonucleotides COB5E and COB3E (see below). The final clone was sequenced completely.

Oligonucleotides used to generate the *A.n.* COBme pre-RNA clone (pT7AnCOBme) and deletion mutants from this plasmid were as follows. In all cases, a copy of the bacteriophage T7 RNA polymerase promoter was included in the design of the 5′-oligonucleotides and an extra two guanosines were added to the 5′ end of the RNA to facilitate transcription. In addition, the 3′-oligonucleotides were designed to contain a *Bam*HI restriction enzyme site that, upon digestion and RNA transcription, added four non-*A.n.* COBme sequences to the 3′ end of each construct. The wild-type substrate was generated with COB5E (5′-GCGGAATTCG CTGCAGTAATACGACTCACTATAGGGAGTTTATTTGAGG AG) and COB3E (5′-GCGTAAGGATTCGCATTATTTACAG AGCATTTATG) that creates a full-length ΔORF *A.n.* COBme intron. For truncations in the 5′ end of the precursor, COBΔ5E (5′-GCGGAATTCGCTGCAGTAATACGACTCACTATAGGGC ACAGATGAACCACA) combined with COB3E generates a product that deletes the 5′ exon. COBΔP1 (5′-GCGGAATT CGCTGCAGTAATACGACTCACTATAGGGACAATGCGGTG ACGT) combined with COB3E generates a product that deletes all of the P1 helix. COBΔP2 (5′-GCGGAATTCGCT GCAGTAATACGACTCACTATAGGGTAAAAATCCTGCTTA) combined with COB3E generates a product that deletes through the P2 helix. COBΔP3 (5′-GCGGAATTCGCTGCAG TAATACGACTCACTATAGGGTTAATGCTGGAAAAT) combined with COB3E generates a product that deletes the 5′ strand of P3.

The first truncation in the 3′ end of the precursor used primer COBΔ3E (5′-GCGTAAGGATCCCATTTTATGTTGAT AGTTACAGA) combined with COB5E to generate a product that deletes the 3′ exon. To delete the 3′ strand of P9.1, this product was digested with *Ssp*I, which truncates at position 280 within the P9.1 loop (see Fig. 1).

Expression and purification of I-*Ani*I

The I-*Ani*I protein was expressed from the plasmid pEC that was grown in either BL21(DE3) or HMS174(DE3) cells (Novagen Inc., Madison, Wisconsin) in 4 mL of SOB media (Sambrook et al., 1989) containing 40 μg/mL kanamycin overnight at 37 °C. One milliliter of this culture was added to into 500 mL SOB media supplemented with 40 μg/mL kanamycin and grown at 37 °C until an $O.D._{600}$ of ~0.7–0.9, at which point 5 mL of 100 mM IPTG was added to induce expression and the culture was grown at 30 °C for 18–20 h.

The cells were harvested at 4 °C and resuspended in 10 mL of 20 mM Tris-HCl, pH 7.9, 500 mM NaCl (TN) supplemented with 5 mM imidazole to which 10 μL of 10 mg/mL lysozyme were added. The cells were lysed by three freeze/thaw cycles at −70 and 25 °C followed by sonication on ice three times. The insoluble material was removed by spinning the lysate at 16,000 × *g* for 15 min at 4 °C. The I-*Ani*I protein in the cleared lysate was chelated with 2.5 mL Ni-NTA agarose resin (Qiagen, Valencia, California) per 500 mL culture and mixed at 4 °C for 1 h.

The mixture was loaded onto a 2 × 10 cm column and the resin washed with 30 mL of TN buffer supplemented with 5 mM imidazole. The column was then washed with 20 mL of TN supplemented with 15 mM imidazole. Finally, the protein was eluted by three successive 3 mL washes of TN supplemented with 60, 100, and 1,000 mM imidazole. Fractions

containing the pure protein were combined and dialyzed against 2 L of TN buffer containing 10% glycerol, for 8–10 h at 4 °C and concentrated by a second dialysis, for 12–18 h, against 2 L of TN buffer with 50% glycerol. The protein was stored at −20 °C. Protein concentration was determined by absorption at 279 nm in 6 M guanidine-HCL, 20 mM Na-phosphate buffer, pH 6.5, using the calculated extinction coefficient of 35,530 $M^{-1}cm^{-1}$. The amount of active protein was measured by stoichiometric RNA splicing and DNA endonuclease assays. In general, the protein preparations were 90–95% active.

Synthesis of in vitro transcripts

In vitro transcription reactions were in 50 or 100 μL of reaction medium containing 2–5 μg of plasmid or PCR generated DNA (for the mutant RNAs), 5 U/μL phage T7 RNA polymerase, 25 mM NaCl, 8 mM $MgCl_2$, 40 mM Tris-HCl, pH 8.0, 2 mM spermidine, 10 mM dithiothreitol (DTT), 0.5–1 mM NTPs, and 1 U/μL RNasin (Promega, Madison, Wisconsin) for 60–120 min at 37 °C. ^{32}P-labeled transcripts used in splicing and binding stoichiometry measurements were synthesized by adding 1 μCi/μL [α-^{32}P]UTP (3,000 Ci/mmol; ICN Biomedicals, Irvine, California) to a transcription mix containing 0.5 mM UTP and 1 mM of each remaining NTP. For higher specific activity transcripts used in k_{off}, k_{on}, and K_d measurements, the 25- or 50-μL transcription mix contained 3 μCi/μL [α-^{32}P]UTP with 0.5 mM ATP, CTP, and GTP, and 10 μM UTP. After transcription, the DNA template was digested with DNase I (0.5 U/μL; FPLC-purified; Pharmacia Biotech Inc, Piscataway, New Jersey) for 20 min at 37 °C. Transcripts were extracted with phenol-chloroform-isoamyl alcohol (phenol-CIA; 25:24:1 v/v/v), centrifuged through a Sephadex G-50 (Sigma Chemical Co., St. Louis, Missouri) spun column and purified by denaturing gel electrophoresis. Prior to use, precursor RNA was heated to 90 °C in H_2O for 20 s and placed on ice. This was found to be important for consistent behavior between different RNA preparations.

RNA splicing reactions

For protein-facilitated splicing reactions, precursor RNA was preincubated at 37 °C in 50 mM Tris-HCl, pH 7.5, and 100 mM NaCl (TN) supplemented with 5 mM $MgCl_2$ in the presence of 1 mM guanosine for 20 min, after which I-*Ani*I was added to initiate splicing. Self-splicing reactions were preincubated in 2× TN and 300 mM $MgCl_2$, at 37 °C for 20 min in the absence of guanosine. The splicing reaction was initiated by the addition of an equal volume of guanosine (1 mM, final concentration) to make a final reaction mixture of 1× TN and 150 mM $MgCl_2$. The final concentrations of RNA and protein are indicated in each experiment shown (see legends for Figs. 2 and 4). The following buffers, at a concentration of 50 mM, were used in the pH titration experiments: MES, pH 5.7, PIPES, pH 6.18, Tris-HCl for pH 6.8, 7.5, 8.0, and 8.8. All splicing reactions were stopped by the addition of EDTA to a final concentration of 100 mM and extraction with phenol-CIA. The reaction products were separated on a denaturing 6% polyacrylamide/7 M urea gel and the dried gels were quantified with a phosphorimager (Molecular Dynamics, Sunnyvale, California). The self-splicing data were fit to a first-order equation with a single exponential: Fraction pre-RNA spliced = $A(e^{-kt})$ where A is the amplitude of RNA and k represents the first-order rate constant. The protein-dependent data were fit to a first-order equation with a double exponential: Fraction RNA spliced = $A(e^{-kt}) + B(e^{-kt})$ where A and B are the amplitude of RNA in each phase and k represents the first order rate constants for each phase. The pH dependence of the splicing reactions were fit to the equation: $k_{obs} = k_{max}/1 + 10^{(pKa-pH)}$ where k_{obs} is the observed rate constant at a particular pH, k_{max} is the activity of the deprotonated species and it is assumed that protonated species are inactive (Campbell et al., 2002).

RNA binding assays

For k_{off} measurements, 1 nM ^{32}P-labeled RNA was preincubated in 50 μL of TN supplemented with 5 mM $MgCl_2$, 10% glycerol, 0.1 mg/mL BSA, and 10 mM DTT (TNMGBD) at 37 °C for 20 min. Complexes were formed by addition of I-*Ani*I to 30 nM and further incubation for 5 min. The reaction mixture was then mixed with 1 mL of TNMGBD containing 100 nM unlabeled *A.n.* COBme pre-RNA to bind unassociated I-*Ani*I protein. At different times after mixing, 90 μL were filtered through nitrocellulose, washed with 3 mL of TN, dried, and counted using a scintillation counter. The data were fit to a first-order equation with a double exponential: Fraction RNA bound = $A(e^{-kt}) + B(e^{-kt})$ where A and B are the amplitude of RNA in each phase and k represents the first-order rate constants for each phase.

To measure k_{on}, 50 pM ^{32}P-labeled RNA was preincubated in TNMGBD at 37 °C for 20 min. Reactions were performed by mixing 100 μL of the RNA mixture with 2 μL of a 50× concentration of I-*Ani*I prewarmed at 37 °C. The reaction was incubated for 10–900 s at 37 °C. Complex formation was stopped by the addition of unlabeled *A.n.* COBme pre-RNA to a concentration of 60 nM and subsequently filtered through nitrocellulose and processed as above. The data were fit to a first-order equation: Fraction RNA bound = $A(1 - e^{-kt})$ where A is the amplitude of RNA and k represents the pseudo-first-order rate constant, k_{obs}.

The stoichiometry of the complex was measured by incubating 32 nM ^{32}P-RNA with different concentrations of I-*Ani*I (2.5–75 nM) in 50 μL of TNMGBD at 37 °C for 5 min. The complexes were filtered and processed as above. The data were corrected for the efficiency of nitrocellulose detection by first measuring the amount of RNA bound with a large excess of I-*Ani*I and then correcting the raw data by this factor (30% efficiency; see Yarus & Berg, 1970).

ACKNOWLEDGMENTS

We thank Dr. Richard B. Waring (Temple University) for the generous gift of plasmid stocks and Dr. Alan M. Lambowitz (University of Texas–Austin) in whose laboratory these studies were initiated. We also thank Drs. R.B. Waring, A.M. Lambowitz, Pieter DeHaseth, Timothy W. Nilsen, and David Setzer (Case Western Reserve University School of Medicine) for comments on the manuscript. This work was supported by American Cancer Society Grant IRG9102206 to M.G.C. and National Institutes of Health Grant GM37951 to A.M.L.

Received November 29, 2001; accepted without revision January 15, 2002

REFERENCES

Belfort M, Roberts RJ. 1997. Homing endonucleases: Keeping the house in order. *Nucleic Acids Res 25*:3379–3388.

Campbell FE Jr, Cassano AG, Anderson VE, Harris ME. 2002. Pre-steady-state and stopped-flow fluorescence analysis of *Escherichia coli* ribonuclease III: Insights into mechanism and conformational changes associated with binding and catalysis. *J Mol Biol 317*:21–40.

Caprara MG, Lehnert V, Lambowitz AM, Westhof E. 1996b. A tyrosyl-tRNA synthetase recognizes a conserved tRNA-like structural motif in the group I intron catalytic core. *Cell 87*:1135–1145.

Caprara MG, Mohr G, Lambowitz AM. 1996a. A tyrosyl-tRNA synthetase protein induces tertiary folding of the group I intron catalytic core. *J Mol Biol 257*:512–531.

Caprara MG, Myers CA, Lambowitz AM. 2001. Interaction of the *Neurospora crassa* mitochondrial tyrosyl-tRNA synthetase (CYT-18 protein) with the group I intron P4–P6 domain. Thermodynamic analysis and the role of metal ions. *J Mol Biol 308*:165–190.

Cate JH, Gooding AR, Podell E, Zhou K, Golden BL, Kundrot CE, Cech TR, Doudna JA. 1996. Crystal structure of a group I ribozyme domain: Principles of RNA packing. *Science 273*:1678–1685.

Chevalier BS, Stoddard BL. 2001. Homing endonucleases: Structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res 29*:3757–3774.

Coetzee T, Herschlag D, Belfort M. 1994. *Escherichia coli* proteins, including ribosomal protein S12, facilitate in vitro splicing of phage T4 introns by acting as RNA chaperones. *Genes & Dev 8*:1575–1588.

Costa M, Michel F. 1995. Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J 14*:1276–1285.

Dalgaard JZ, Klar AJ, Moser MJ, Holley WR, Chatterjee A, Mian IS. 1997. Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. *Nucleic Acids Res 25*:4626–4638.

Engelhardt MA, Doherty EA, Knitt DS, Doudna JA, Herschlag D. 2000. The P5abc peripheral element facilitates preorganization of the *Tetrahymena* group I ribozyme for catalysis. *Biochemistry 39*:2639–2651.

Fersht AR, Requena Y. 1971. Equilibrium and rate constants for the interconversion of two conformations of a-chymotrypsin. *J Mol Biol 60*:279–290.

Fierke CA, Hammes GG. 1995. Transient kinetic approaches to enzyme mechanisms. *Methods Enzymol 249*:3–37.

Geese WJ, Waring RB. 2001. A comprehensive characterization of a group IB intron and its encoded maturase reveals that protein-assisted splicing requires an almost intact intron RNA. *J Mol Biol 308*:609–622.

Golden BL, Gooding AR, Podell ER, Cech TR. 1998. A preorganized active site in the crystal structure of the *Tetrahymena* ribozyme. *Science 282*:259–264.

Herschlag D. 1995. RNA chaperones and the RNA folding problem. *J Biol Chem 270*:20871–20874.

Herschlag D, Khosla M. 1994. Comparison of pH dependencies of the Tetrahymena ribozyme reactions with RNA 2′-substituted and phosphorothioate substrates reveals a rate-limiting conformational step. *Biochemistry 33*:5291–5297.

Ho Y, Kim SJ, Waring RB. 1997. A protein encoded by a group I intron *in Aspergillus nidulans* directly assists RNA splicing and is a DNA endonuclease. *Proc Natl Acad Sci USA 94*:8994–8999.

Ho Y, Waring RB. 1999. The maturase encoded by a group I intron from *Aspergillus nidulans* stabilizes RNA tertiary structure and promotes rapid splicing. *J Mol Biol 292*:987–1001.

Hur M, Geese WJ, Waring RB. 1997. Self-splicing activity of the mitochondrial group-I introns from *Aspergillus nidulans* and related introns from other species. *Curr Genet 32*:399–407.

Johnson KA. 1992. Transient-state kinetic analysis of enzyme reaction pathways. In: Boyer PD, ed. *Enzymes*, 4th ed., Vol. 20. New York: Academic Press. pp 1–60.

Jurica MS, Stoddard BL. 1999. Homing endonucleases: Structure, function and evolution. *Cell Mol Life Sci 55*:1304–1326.

Lambowitz AM. 1989. Infectious introns. *Cell 56*:323–326.

Lambowitz AM, Caprara MG, Zimmerly S, Perlman PS. 1999. Group I and group II ribozymes as RNPs: Clues to the past and guides to the future. In: Gesteland RF, Atkins JF, Cech TR, eds. *The RNA world II*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press. pp 451–485.

Lambowitz AM, Perlman PS. 1990. Involvement of aminoacyl-tRNA synthetases and other proteins in group I and group II intron splicing. *Trends Biochem Sci 15*:440–444.

Lehnert V, Jaeger L, Michel F, Westhof E. 1996. New loop-loop tertiary interactions in self-splicing introns of subgroup IC and ID: A complete 3D model of the *Tetrahymena thermophila* ribozyme. *Chem Biol 12*:993–1009.

Mannella CA, Collins RA, Green MR, Lambowitz AM. 1979. Defective splicing of mitochondrial rRNA in cytochrome-deficient nuclear mutants of *Neurospora crassa*. *Proc Natl Acad Sci USA 76*:2635–2639.

McConnell TS, Cech TR. 1995. A positive entropy change for guanosine binding and for the chemical step in the Tetrahymena ribozyme reaction. *Biochemistry 34*:4056–4067.

McGraw P, Tzagoloff A. 1983. Assembly of the mitochondrial membrane system. Characterization of a yeast nuclear gene involved in the processing of the cytochrome b pre-mRNA. *J Biol Chem 258*:9459–9468.

Michel F, Westhof E. 1990. Modelling of the three-dimensional architecture of group I catalytic introns based on comparative sequence analysis. *J Mol Biol 216*:585–610.

Pan J, Woodson SA. 1998. Folding intermediates of a self-splicing RNA: Mispairing of the catalytic core. *J Mol Biol 280*:597–609.

Pan J, Woodson SA. 1999. The effect of long-range loop-loop interactions on folding of the *Tetrahymena* self-splicing RNA. *J Mol Biol 294*:955–965.

Riggs AD, Bourgeois S, Cohn M. 1970. The lac repressor-operator interaction. 3. Kinetic studies. *J Mol Biol 53*:401–417.

Rose MA, Weeks KM. 2001. Visualizing induced fit in early assembly of the human signal recognition particle. *Nature Struct Biol 8*: 515–520.

Saldanha R, Ellington A, Lambowitz AM. 1996. Analysis of the CYT-18 protein binding site at the junction of stacked helices in a group I intron RNA by quantitative binding assays and in vitro selection. *J Mol Biol 261*:23–42.

Saldanha RJ, Patel SS, Surendran R, Lee JC, Lambowitz AM. 1995. Involvement of *Neurospora* mitochondrial tyrosyl-tRNA synthetase in RNA splicing. A new method for purifying the protein and characterization of physical and enzymatic properties pertinent to splicing. *Biochemistry 34*:1275–1287.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: A laboratory manual*, 2nd ed. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Shaw LC, Lewin AS. 1995. Protein-induced folding of a group I intron in cytochrome b pre-mRNA. *J Biol Chem 270*:21552–21562.

Sullivan JJ, Bjornson KP, Sowers LC, deHaseth PL. 1997. Spectroscopic determination of open complex formation at promoters for *Escherichia coli* RNA polymerase. *Biochemistry 36*:8005–8012.

Thirumalai D, Lee N, Woodson SA, Klimov D. 2001. Early events in RNA folding. *Annu Rev Phys Chem 52*:751–762.

Treiber DK, Williamson JR. 2001. Beyond kinetic traps in RNA folding. *Curr Opin Struct Biol 11*:309–314.

Wallweber GJ, Mohr S, Rennard R, Caprara MG, Lambowitz AM. 1997. Characterization of *Neurospora* mitochondrial group I introns reveals different CYT-18 dependent and independent splicing strategies and an alternative 3′ splice site for an intron ORF. *RNA 2*:114–131.

Webb AE, Weeks KM. 2001. A collapsed state functions to self-chaperone RNA folding into a native ribonucleoprotein complex. *Nature Struct Biol 8*:135–140.

Weeks KM. 1997. Protein-facilitated RNA folding. *Curr Opin Struct Biol 7*:336–342.

Weeks KM, Cech TR. 1995a. Efficient protein-facilitated splicing of the yeast mitochondrial bI5 intron. *Biochemistry 34*:7728–7738.

Weeks KM, Cech TR. 1995b. Protein facilitation of group I intron splicing by assembly of the catalytic core and the 5′ splice site domain. *Cell 82*:221–230.
Weeks KM, Cech TR. 1996. Assembly of a ribonucleoprotein catalyst by tertiary structure capture. *Science 271*:345–348.
Wernette CM, Saldanha R, Perlman PS, Butow RA. 1990. Purification of a site-specific endonuclease, I-Sce II, encoded by intron 4 alpha of the mitochondrial coxI gene of *Saccharomyces cerevisiae*. *J Biol Chem 265*:18976–18982.
Woodson SA. 2000. Recent insights on RNA folding mechanisms from catalytic RNA. *Cell Mol Life Sci 57*:796–808.
Yarus M, Berg P. 1970. On the properties and utility of a membrane filter assay in the study of isoleucyl-tRNA synthetase. *Anal Biochem 35*:450–465.