

# Experimental Tests of Two Proofreading Mechanisms for 5'-Splice Site Selection

Yangming Wang and Scott K. Silverman\*

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, Illinois 61801

**ABSTRACT** Self-splicing group II intron RNAs catalyze a two-step process in which the intron is excised as a lariat by two successive phosphodiester exchange reactions. The reversibility of the first step has been hypothesized to act as a proofreading mechanism for improper 5'-splice site selection. However, without synthetic access to mis-spliced RNAs, this hypothesis could not be tested. Here, we used a deoxyribozyme to synthesize several branched RNAs that are derived from the ai5 $\gamma$  group II intron and mis-spliced at the 5'-splice site. Unlike the correctly spliced ai5 $\gamma$  RNAs, the mis-spliced RNAs are observed not to undergo the reverse of the first step. This is well-controlled negative evidence against the hypothesis that first-step reversibility is a proofreading mechanism for 5'-splice site selection. In a reaction equivalent either to the hydrolytic first step of splicing or to the hydrolytic reverse of the second step of splicing, a mis-spliced 5'-exon can be "trimmed" to its proper length by the corresponding mis-spliced intron, and in one case, the trimmed 5'-exon was observed to proceed correctly through the second step of splicing. These findings are the first direct evidence that this second proofreading mechanism can occur with a group II intron RNA that is mis-spliced at the 5'-splice site. On the basis of the likely structural and evolutionary relationship between group II introns and the spliceosome, we suggest that this second proofreading mechanism may be operative in the spliceosome.

The RNA self-splicing pathway catalyzed by group II intron RNAs (1–5) has two reaction steps (Figure 1, panel a). In the first step, an internal adenosine 2'-hydroxyl group from domain 6 attacks the 5'-splice site phosphodiester linkage, forming a lariat–3'-exon intermediate with departure of the 5'-exon. In the second step, the 5'-exon attacks the 3'-splice site, forming the ligated exons and excising the intron as a lariat. Although the first step of splicing is reversible (6), the underlying explanation for this reversibility is unclear. The lack of natural 5'-splice site mis-splicing (7–9) suggests that a proofreading mechanism may exist. One specific hypothesis is that first-step reversibility is itself a proofreading mechanism: any intron that mis-splices at the first step by choosing an improper 5'-splice site can return to the initial unspliced state by the reverse of the first step (6). This would provide the RNA another opportunity to splice correctly, rather than waste the RNA molecule in a dead-end route or lead to improperly spliced exons after subsequently proceeding through the second step of splicing. This first proofreading mechanism requires that mis-spliced introns which have been formed by use of the improper 5'-splice site during the first step will be competent to proceed through the reverse of the first step. However, this mechanism has never been tested experimentally, because the required mis-spliced RNAs could not be synthesized using previously available methods.

Because a mis-spliced 5'-exon that proceeds through the second step would lead to improperly ligated exons, a natural mechanism to repair a mis-spliced 5'-exon would be valuable. A second potential proofreading mechanism for RNAs that are mis-spliced at the 5'-splice site is for the incorrect 5'-exon to be converted to the correct length 5'-exon and subsequently proceed through the second step. When the incorrect 5'-exon has one or more extra nucleotides due to mis-splicing

\*To whom correspondence should be addressed.

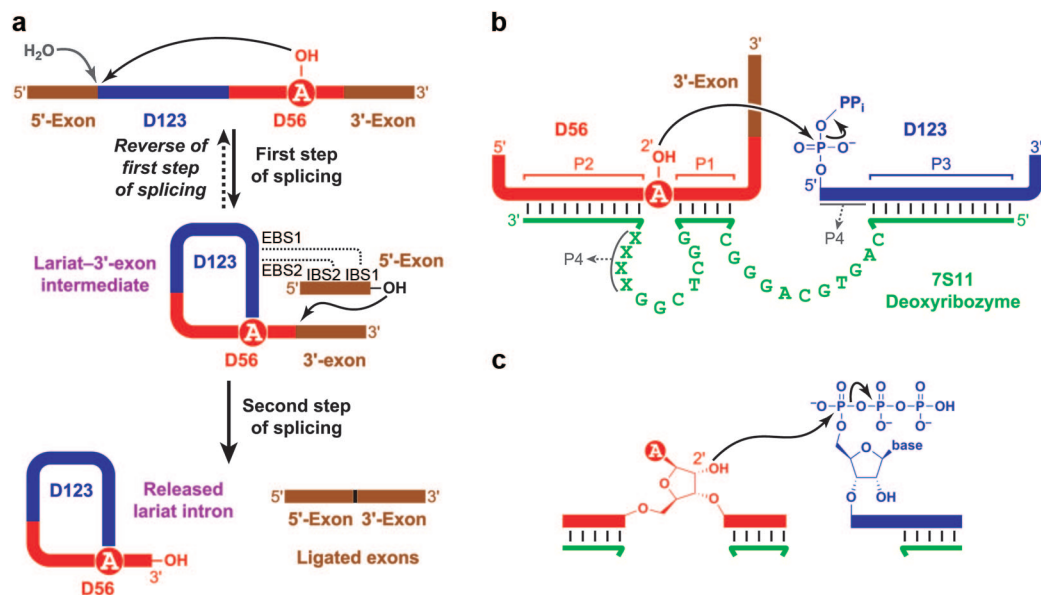
E-mail: scott@scs.uiuc.edu.

Received for review April 6, 2006  
and accepted May 6, 2006.

Published online June 16, 2006

10.1021/cb6001569 CCC: \$33.50

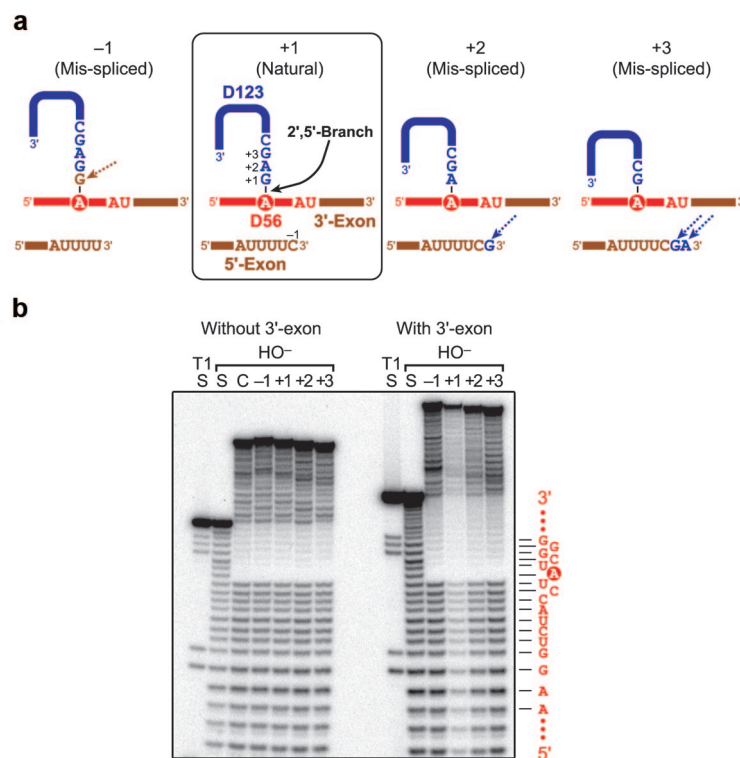
© 2006 by American Chemical Society



**Figure 1.** RNA splicing and branched RNA. **a)** The two steps of group II intron splicing. Interactions between the intron binding sites (IBS) and exon binding sites (EBS) are also present during the first step (not depicted). For the ai5 $\gamma$  intron, D123 is 675 nt and D56 is 77 nt; D4 is not depicted because it is dispensable for catalytic activity (52). First-step reversibility has been observed for correctly spliced RNA (6); whether or not this reversibility applies for mis-spliced RNAs is one focus of this manuscript. Hydrolytic cleavage at the 5'-splice site (4, 10–12) is also depicted (gray water molecule); this leads to linear instead of branched RNA. **b)** Synthesis of 2',5'-branched ai5 $\gamma$  RNA catalyzed by the 7S11 deoxyribozyme (20, 21). For natural branched RNAs, the branch-site nucleotide is almost always adenosine (circled). For applying 7S11 to synthesize branched RNAs that correspond to selection of an improper 5'-splice site, the DNA sequence of paired region P4 (denoted here with Xs) is chosen to maintain base pairing with the RNA sequence near the 5'-splice site. **c)** The key components of the branch-formation reaction, shown in greater atomic detail.

downstream of the correct 5'-splice site, this requires “trimming” of the incorrect 5'-exon by removal of the extra nucleotide(s), followed by the second step. Such trimming of the 5'-exon could correspond to the well-known hydrolytic first step of splicing, in which water instead of the branch-site adenosine acts as the nucleophile that attacks the 5'-splice site (4, 10–12). Alternatively, the trimming reaction could instead correspond to the hydrolytic reverse of the second step of splicing (11, 13), analogous to spliced exons reopening (SER) (14–16). Hydrolysis has been observed for short RNA oligonucleotides that are analogues of the 5'-exon (7, 8, 13). However, catalysis of oligonucleotide hydrolysis and the subsequent second step have never been observed by mis-spliced intron RNAs, because the required mis-spliced RNAs could not be synthesized. Therefore, a direct test of this second proofreading mechanism for 5'-splice site selection has not been achievable.

Our laboratory has identified many artificial deoxyribozymes (DNA enzymes) (17) for RNA ligation (18, 19). Several of these deoxyribozymes such as 7S11 (20, 21) create 2',5'-branched RNA (Figure 1, panel b) (20–24), where a 2',5'-branch is the key structural element of the lariat RNAs that are the intermediates in biological RNA splicing (25). In reactions catalyzed by group II introns, branched RNA (which lacks the closed loop of a lariat) is as functionally competent as lariat RNA (4, 26). Here, we have used synthetic branched RNAs prepared by the 7S11 deoxyribozyme to allow explicit experimental tests of the two proofreading mechanisms for 5'-splice site selection. Mis-spliced RNAs are observed not to undergo the reverse of the first step of splicing, which provides evidence against the first proofreading mechanism. In contrast, direct evidence is obtained for the operation of the second proofreading mechanism.



**Figure 2.** Correctly spliced and mis-spliced ai5 $\gamma$  branched RNAs, and verification of their structures. **a)** The ai5 $\gamma$  bipartite D123/D56 system, showing correctly spliced (box) and mis-spliced branch–3'-exon intermediates after the first step of splicing. Before the first step, the 5'-exon is joined at its 3'-end to the 5'-end of D123. The 5'-splice site locations are denoted +1 for the natural splice site (see boxed structure) and either –1, +2, or +3 for an improper splice site. The branch-site adenosine nucleotide is circled. Nucleotides found at the incorrect position within each mis-spliced RNA are marked with dashed arrows. For the –1 mis-spliced RNA, the final nucleotide of the original 5'-exon sequence, which is connected directly to the branch-site adenosine after the first step, has been changed from C to G (see text for details). **b)** Partial alkaline hydrolysis to verify the structures of the ai5 $\gamma$  branched RNAs synthesized using the 7S11 deoxyribozyme (20% PAGE; T1 denotes G-specific RNase T1 digestion and S = standard for ladder calibration). “C” denotes a branched RNA synthesized using the conventional ai5 $\gamma$  forward splicing reaction. Prior to alkaline hydrolysis, the core of each branched RNA was excised using two 10–23 deoxyribozymes (38, 39).

## RESULTS AND DISCUSSION

**Choice of the Splicing System.** We used the ai5 $\gamma$  group II intron RNA (27, 28) as the self-splicing RNA to test the first proofreading mechanism for 5'-splice site selection. The ai5 $\gamma$  RNA is a common model system for understanding group II intron structure and function, and it has been studied using several biochemical approaches (4). One of the most useful approaches to assemble ai5 $\gamma$  is with the bipartite D123/D56 system

(Figure 2, panel a, shown schematically after the first step of splicing; see boxed structure for correctly spliced RNA). In this bipartite system, the branch-site adenosine nucleotide of domain 6 is located within the 77-nt D56 RNA that comprises domains 5 and 6, plus the 3'-exon (if any is included; the first splicing step does not require the 3'-exon). The 5'-exon is joined to the 5'-end of domains 1–3 (D123), which encompasses 675 nt. A key advantage of studying ai5 $\gamma$  splicing is that a lariat topology of D123/D56 is not necessary, because both splicing steps proceed well in the simpler 2',5'-branched RNA (4, 26).

### Synthesis of Mis-Spliced ai5 $\gamma$ Variants by the 7S11 Deoxyribozyme.

Experimentally testing the proofreading mechanisms required synthesis of 2',5'-branched RNAs that correspond to the mis-spliced intermediates after the first step of splicing. The 7S11 deoxyribozyme synthesizes branched RNA from two RNA substrates with few restrictions on their nucleotide sequences (20, 21). To prepare ai5 $\gamma$  variants using 7S11, the RNA substrate providing the 2'-hydroxyl nucleophile was a T7 polymerase transcript corre-

sponding to D56, and the RNA substrate providing the 5'-portion of the 2',5'-linkage was a D123 transcript (Figure 1, panel b).

Using the 7S11 deoxyribozyme, the ai5 $\gamma$  branched intermediate with the natural (correct) spliced sequence was prepared along with three sequence variants. Each variant represents a mis-spliced RNA that would be formed if an improper 5'-splice site phosphodiester linkage near the 5'-end of D123 were attacked by the

branch-site 2'-hydroxyl group from D6 (Figure 2, panel a, three unboxed structures). The RNA sequence near the natural 5'-splice site is 5'...CpGAGC...3', where the G is the first nucleotide of the intron (D123 domain) and p denotes the natural 5'-splice site phosphodiester linkage. This natural 5'-splice site is conventionally designated as position +1, with the +1 value pertaining to the nucleotide whose 5'-hydroxyl is attached to the splice site phosphorus atom (*i.e.*, G). The nucleotides one or two positions to the 3'-side of the natural 5'-splice site (here, A and G) are denoted +2 and +3, respectively. The +1, +2, and +3 ai5 $\gamma$  intermediates were prepared using the 7S11 deoxyribozyme with appropriate RNA substrate sequences that have entirely wild-type nucleotides. The intermediate from splicing at a position one nucleotide to the 5'-side of the natural splice site (*i.e.*, at position -1) was also prepared using 7S11. To accommodate the modest substrate sequence requirements of 7S11, which requires a purine as the branch-site nucleotide (20, 21), the -1 nucleotide was changed from C to G. To compensate, the corresponding base-paired nucleotide within the exon binding site (EBS1) of D123 was changed from G to C. Consistent with expectations (29, 30), this C=G base-pair flip did not substantially affect either the yield or the rate of the forward first step of splicing (Supplementary Figure 1).

We focused on this small window of phosphodiester linkages near the natural 5'-splice site (-1, +1, +2, and +3) because the -1 nucleotide is already within IBS1 that is required for 5'-splice site fidelity (1, 7, 8, 31) and because the +3 nucleotide is involved directly in the  $\varepsilon$ - $\varepsilon'$  tertiary interaction (32). Therefore, it seems unlikely that studying improper 5'-splice site selection outside of this window will be biologically meaningful, because the mis-spliced intron will be missing crucial key components. It should be noted that all six nucleotides at the start of domain 1 (nucleotides +1 through +6) are highly conserved and may contribute to catalytic activity (33, 34).

For all four splice site variants (-1, +1, +2, and +3), the branched RNA intermediate was synthesized either omitting or including a short (6-nt) 3'-exon sequence. Therefore,  $4 \times 2 = 8$  branched RNAs were prepared. Only when the 3'-exon is omitted can the first step of splicing be examined in either direction without intervention of the second step. As a positive control (denoted "C"), the natural (+1) branched intermediate without 3'-exon was also synthesized using the

conventional forward splicing reaction of the ai5 $\gamma$  RNA without using the 7S11 deoxyribozyme. Of course, the control branch that includes the 3'-exon could not be synthesized in more than trace amounts using ai5 $\gamma$  catalysis, because the second step of splicing is rapid (1, 35-37).

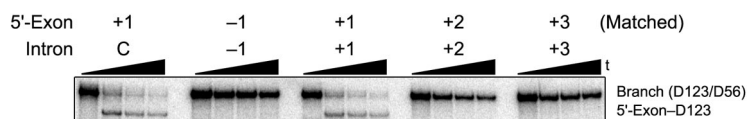
This study is the first in which we report direct biochemical applications of branched RNAs that are synthesized by deoxyribozymes. Therefore, we empirically established the nucleotide connectivities of all branched ai5 $\gamma$  intermediates using partial alkaline hydrolysis, after initially excising the branched core using two 10-23 deoxyribozymes (38, 39). For all eight synthetic branches as well as for the "C" control, the anticipated branch sites were confirmed unambiguously (Figure 2, panel b).

#### Assaying Mis-Spliced RNAs for the Reverse of the First Step of Splicing (First Proofreading Mechanism).

When the newly prepared synthetic mis-spliced RNAs were used, it was straightforward to test the first proofreading mechanism, which requires that mis-spliced RNAs undergo the reverse of the first step of splicing. Each of the four internally  $^{32}\text{P}$ -radiolabeled synthetic branched RNA variants without the 3'-exon (and the "C" control) was separately incubated with an excess of the matched unradiolabeled 5'-exon oligonucleotide (the 5'-exon is "matched" because reverse splicing would restore the original unspliced RNA sequence). If reverse splicing occurred at all, it was readily detected by a gel shift (Figure 3). To prevent the second step of splicing from occurring, the 3'-exon was not present on D56 in these experiments. When the branched intermediate has the natural 5'-splice site, reverse splicing is observed, with equivalent efficiency for both the control ("C") and synthetic (+1) branched RNA. In contrast, reverse splicing fails to occur (<2%) when the branched intermediate has the incorrect (-1, +2, or +3) 5'-splice site and the matched 5'-exon is provided. In this direct and well-controlled test, the results do not support the first proofreading mechanism.

#### Assaying Mis-Spliced RNAs for the Second Step of Splicing (Second Proofreading Mechanism).

In the absence of the 3'-exon sequence, the only possible reaction of the branched RNA is the reverse of the first step, as described above. However, when the 3'-exon is included with D56, the second step of splicing can also occur by attack of the 5'-exon at the 3'-splice site of the branched RNA. The availability of synthetic mis-spliced



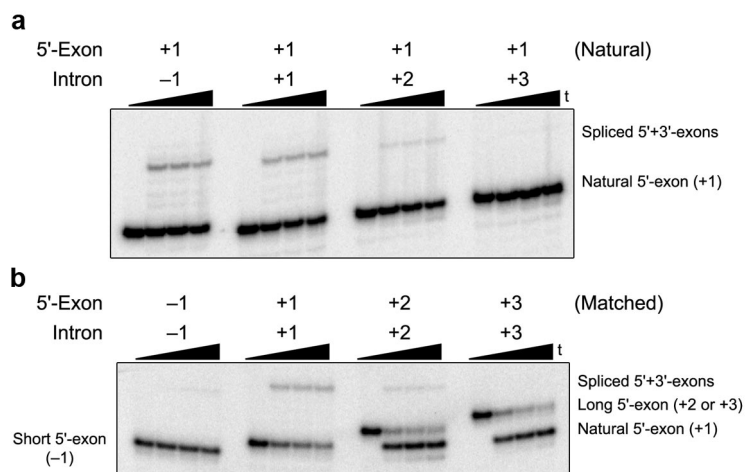
**Figure 3.** Assays for the reverse of the first step of splicing show no reaction for the mis-spliced RNAs, which is evidence against the first proofreading mechanism for 5'-splice site selection. The internally  $^{32}\text{P}$ -radiolabeled branched RNA without 3'-exon was the limiting reagent (unradiolabeled 5'-exon in 200-fold excess;  $t = 0, 0.5, 1,$  and  $2\text{ h}$ ; 5% PAGE). The control "C" used the +1 splice site and was prepared by forward splicing using the catalytic activity of ai5 $\gamma$  itself. The other four branched RNAs were prepared using the 7S11 deoxyribozyme (Figure 1, panel b). Consistent results were observed with analogous assays that used either an excess of radiolabeled 5'-exon or an excess of branched RNA (Supplementary Figures 2 and 3). The overall loss of signal intensity with time has been observed numerous times by others (15, 16, 57–60).

branched RNAs allowed us to examine whether the mis-spliced branched RNAs are competent for the second step of splicing, independent of their ability to undergo the reverse of the first step.

We first examined the competence of the mis-spliced branch–3'-exon RNAs to proceed through the second step using the natural (+1) 5'-exon in all cases (Figure 4, panel a). The primary purpose of this assay was to provide baseline information on the catalytic competence of the mis-spliced branches, before testing them with their matched 5'-exons. The –1, +1, and +2 branches were found to catalyze the second step reasonably well, whereas the +3 branch does not. The 10–20% efficiency observed for the natural +1 reaction is similar to that observed for related systems (10, 29).

We then tested each of the four synthetic branch–3'-exon RNAs (–1, +1, +2, and +3) by incubating each RNA with a limiting amount of the matched 5'-exon (Figure 4, panel b). In the case of the natural (+1) branched RNA and matched natural 5'-exon, normal second-step splicing activity was found (as also shown in

Figure 4, panel a). For the +2 mis-spliced RNA, the matched (+2) 5'-exon, which is longer than the natural (+1) 5'-exon at its 3'-end by one nucleotide, was rapidly and efficiently "trimmed" with loss of a single nucleotide to the correct length. This trimmed 5'-exon, now equivalent in length to the +1 5'-exon, then proceeded through the second step of splicing with 2–3% final splicing yield, versus 10–20% yield for the +1 RNA. This indicates that the efficiency of this second proofreading mechanism is on the order of 10–30% relative to formation of correctly spliced exons by the properly branched RNA intermediate. Therefore, for the +2 mis-spliced RNA, the correctly ligated exons can indeed be formed via the second proofreading mechanism. The sequence of the correctly ligated exons was verified by RT-PCR and sequencing of the splice junction (see Supporting Infor-

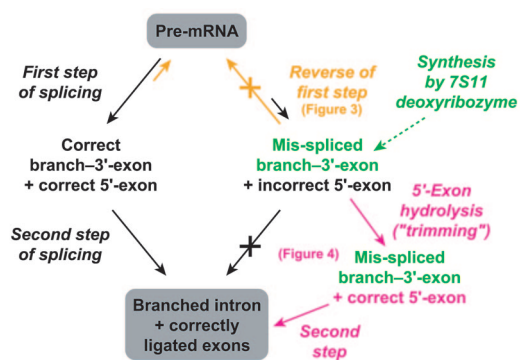


**Figure 4.** Assays for the second step of splicing directly demonstrate the operation of the second proofreading mechanism for 5'-splice site selection using RNAs mis-spliced at the 5'-splice site. a) Using branch–3'-exon RNA and natural-length (+1) 5'-exon. b) Using branch–3'-exon RNA and matched 5'-exon. In both sets of experiments, the 5'-exon was  $^{32}\text{P}$ -radiolabeled, and a 10-fold excess of unradiolabeled branched RNA was added ( $t = 0, 0.5, 1,$  and  $2\text{ h}$ ; 20% PAGE). When expected, the bands at the top of each gel near the wells showed substantial reverse first-step splicing in competition with the second step (Supplementary Figure 4). The overall loss of signal intensity with time has been observed by others (15, 16, 57–60).

mation). As revealed by data points taken at shorter times (not shown), the second-step splicing product was formed without any detectable induction period relative to formation of the trimmed 5'-exon. This suggests that for the +2 mis-spliced intron RNA, the second step is fast (*i.e.*, not rate-determining), which is also true for the correctly spliced intron RNA (1, 35–37). For the +3 mis-spliced branch, 5'-exon trimming was observed. However, the subsequent second-step efficiency was too low to be detected, as expected on the basis of the analogous experiment with the +1 branch (in Figure 4, panel a; this may be due to the participation of the +3 nucleotide in the  $\epsilon$ - $\epsilon'$  tertiary interaction (32), which in the +3 mis-spliced RNA may be positioned suboptimally). For the -1 mis-spliced 5'-exon, no trimming is possible because the 5'-exon is already too short by one nucleotide, and almost no second-step product is observed.

**Implications of the Experimental Data for the Two Proofreading Mechanisms.** Using synthetic branched RNAs prepared by the 7S11 deoxyribozyme (Figure 2), we have tested the long-standing hypothesis (5) that reversibility of the first step of group II intron splicing acts as a proofreading mechanism for 5'-splice site selection (6). Because first-step reversibility was not observed for any of the three mis-spliced RNAs that were tested (Figure 3), our data are evidence against this first proofreading mechanism. In contrast, direct evidence was obtained for operation of the alternative, second proofreading mechanism catalyzed by mis-spliced RNA, in which 5'-exon hydrolysis is followed by the second step of splicing (Figure 4). A summary of the experimental observations highlighting both proofreading mechanisms and their relationship is provided (Figure 5).

Our evidence against first-step reversibility is inherently negative evidence; that is, reversibility was not observed (Figure 3). Complicating factors could contribute to this lack of reversibility. For example, a synthetic mis-spliced RNA prepared using a deoxyribozyme might adopt a conformation that does not permit the reverse of the first step, in contrast to a mis-spliced RNA that was formed through the forward operation of the first step using the improper 5'-splice site. It is also possible that protein cofactors which are not required *in vitro* for the forward first step of splicing might nonetheless be required for the reverse of the first step by synthetic mis-spliced RNAs. Such possibilities



**Figure 5. Summary of the experimental observations, highlighting the first and second proofreading mechanisms (orange and magenta, respectively) and synthesis of the key mis-spliced branch-3'-exon RNAs by the 7S11 deoxyribozyme (green).**

are intrinsically challenging (if not impossible) to disprove. Instead, we emphasize that our experiments do *not* have the simplest outcome that would be expected if first-step reversibility were to act as a proofreading mechanism for 5'-splice site selection.

In sharp contrast, we have provided positive evidence for the second proofreading mechanism, which is "trimming" of the 5'-exon by hydrolysis (7, 8, 13) followed by the second step. In particular, for the +2 mis-spliced RNA (where the nucleotide one position to the 3'-side of the natural 5'-splice site is the site of branching), both of these reaction steps were directly observed (Figure 4). Therefore, we have provided unambiguous evidence for the *in vitro* operation of this second proofreading mechanism. All previous studies, which could not use mis-spliced RNAs, merely suggested that such a mechanism might be applicable for RNAs that are mis-spliced at the 5'-splice site (7, 8, 13). The efficiency of the second proofreading mechanism is relatively modest (on the order of 10–30%), but this level of proofreading may plausibly be biologically relevant. In addition, protein cofactors could potentially be required *in vivo* for optimal proofreading efficiency.

**Natural Roles of First-Step Reversibility and of the Second Proofreading Mechanism.** Improper selection of the 5'-splice site in group II introns is rare because the IBS1–EBS1 interaction strictly enforces proper splice-site selection (1, 7, 8, 31). This suggests that an efficient first-step proofreading mechanism may be unnecessary in nature; indeed, our results directly show that first-step reversibility does not contribute to the fidelity of

5'-splice site selection (6). We agree with the previous suggestion (6) that first-step reversibility likely evolved to enable retrotransposition (40, 41), and not for any proofreading function. In contrast to the irreversibility of the first step for the mis-spliced RNAs, our data directly demonstrate the second proofreading mechanism for improper 5'-splice site selection (Figure 4). This second proofreading mechanism may operate in those rare cases for which improper 5'-splice site selection has indeed occurred, at least in cases where mis-splicing has occurred to the 3'-side of the natural 5'-splice site. We do not observe the second step of splicing using a mis-spliced 5'-exon unless 5'-exon hydrolysis (trimming) occurs first. This could be important biologically because the second step of splicing using an untrimmed mis-spliced 5'-exon would give improperly ligated exons and therefore an incorrect transcript. Together, the combination of trimming and the second step constitutes the second proofreading mechanism,

which we have now shown directly to occur for RNAs that are mis-spliced at the 5'-splice site.

#### Implications for Spliceosomal RNA Processing.

Spliceosomal pre-mRNA processing (25, 42) appears to be mechanistically related to group II intron splicing (2, 16, 43–45). Although reversibility of the first step has not been explicitly demonstrated in the spliceosome, such reversibility is likely on the basis of the relationship of the spliceosome to group II introns. The data presented here therefore suggest that reversibility of the first step is not a proofreading mechanism in pre-mRNA splicing. Proper splice site selection during pre-mRNA splicing requires many protein and RNA factors (46–49), which suggests that proofreading mechanisms would be valuable. In the absence of first-step reversibility, we suggest that the second proofreading mechanism for 5'-splice site selection (involving 5'-exon hydrolysis followed by the second step of splicing) may be operative within the spliceosome.

## METHODS

**DNA Oligonucleotides and RNA Transcripts.** All nucleic acid samples were purified by PAGE as described previously (18, 50). The natural and mis-spliced 5'-exons were prepared by *in vitro* transcription using T7 RNA polymerase with an appropriate double-stranded DNA template (51). The natural 5'-exon was 5'-GCGUGGUGGGACAUUUUC-3', where IBS2 is bold and IBS1 is underlined. Each D123 RNA was transcribed using an Ear-linearized plasmid derived from plasmid pJD20 (52) as the template. Each D56 RNA (with or without 3'-exon) was transcribed using a double-stranded PCR product derived from plasmid pJD20 as the template. A homogeneous 3'-terminus for D56 (with or without the 3'-exon) was provided by intramolecular hepatitis delta virus ribozyme cleavage (53, 54). When the 3'-exon was not included at the 3'-terminus of D56, the 2',3'-cyclic phosphate was removed with T4 polynucleotide kinase in the absence of ATP (55). 5'-<sup>32</sup>P-Radiolabeling was performed with  $\gamma$ -[<sup>32</sup>P]-ATP and T4 polynucleotide kinase. Internal <sup>32</sup>P-radiolabeling was performed by including in the transcription solution a trace amount of  $\alpha$ -[<sup>32</sup>P]-CTP along with all four unradiolabeled NTPs, each at 4 mM.

**Synthesis of Branched RNA.** The branched RNAs were synthesized using the 7S11 deoxyribozyme, with the enzyme region sequence as described previously and the binding arms complementary to appropriate portions of the RNA substrates (20, 21). Disruptor oligonucleotides (56) were required for both D123 and D56 to enable proper binding of 7S11 to the RNA substrates (see Supporting Information for details). The D123 substrate (20 pmol), deoxyribozyme (30 pmol), D56 substrate (with or without 3'-exon; 60 pmol), and disruptor oligonucleotides (200 pmol of D123 disruptor and 600 pmol of D56 disruptor) were combined in 11.8  $\mu$ L of 5 mM HEPES (pH 7.5), 15 mM NaCl, and 0.1 mM EDTA. The sample was annealed by heating at 95 °C for 3 min and cooling on ice for 5 min. The sample volume was raised to 20  $\mu$ L with final concentrations of 50 mM EPPS (pH 8.5), 150 mM NaCl, 2 mM KCl, and 40 mM MgCl<sub>2</sub>. The

sample was incubated at 37 °C for 1.5 h, then quenched onto 25  $\mu$ L of stop solution (80% formamide, 1  $\times$  TB [89 mM each Tris and boric acid, pH 8.3], 50 mM EDTA, and 0.025% each bromophenol blue and xylene cyanol). Samples were purified by 5% denaturing PAGE.

To prepare the control ("C") branched RNA by ai5 $\gamma$  catalysis to study the reverse of the first step of splicing, 5'-exon-D123 (293 + 675 = 968 nt; 20 pmol) and D56-3'-exon (77 + 180 = 257 nt; 400 pmol) were combined in 64  $\mu$ L of 5 mM MOPS (pH 7.0) and 1 mM EDTA and annealed by heating at 95 °C for 1 min and cooling at room temperature for 1 min. The sample volume was then raised to 80  $\mu$ L with final concentrations of 40 mM MOPS (pH 7.0), 100 mM MgCl<sub>2</sub>, and 500 mM NH<sub>4</sub>Cl. The sample was incubated at 45 °C for 1 h, quenched onto 80  $\mu$ L of stop solution, and purified by 5% denaturing PAGE.

**Verification of Branched RNA Structures.** The core of each branched RNA (45 nt for +1 RNA) was prepared by two sequential 10–23 deoxyribozyme cleavage reactions of the large branched RNA (752 nt for +1 RNA). Then, the location of each branch site was determined by partial alkaline hydrolysis of the small branched RNA as previously described (22, 39).

**Splicing Assays.** For the assay of Figure 3, internally <sup>32</sup>P-radiolabeled branched RNA (0.1 pmol) and unradiolabeled matched 5'-exon (20 pmol) were combined in 16  $\mu$ L of 5 mM MOPS (pH 7.0) and 0.1 mM EDTA and annealed by heating at 95 °C for 1 min and cooling at room temperature for 1 min. The sample volume was raised to 20  $\mu$ L with final concentrations of 40 mM MOPS (pH 7.0), 100 mM MgCl<sub>2</sub>, and 500 mM NH<sub>4</sub>Cl. The sample was incubated at 45 °C. Aliquots were withdrawn at desired time points, quenched onto stop solution, and separated by 5% denaturing PAGE. Gel images were acquired using a PhosphorImager. Similar results were observed using an alternative high-salt assay buffer (26) of 40 mM HEPES (pH 7.5), 100 mM MgSO<sub>4</sub>, and 500 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> (data not shown). For the assays of Figure 4, internally <sup>32</sup>P-radiolabeled branched RNA (0.2 pmol) and 5'-<sup>32</sup>P-radiolabeled matched 5'-exon

(0.005 pmol) were assayed as described above (20% denaturing PAGE).

**Acknowledgment:** This research was supported by the Burroughs Wellcome Fund (New Investigator Award in the Basic Pharmacological Sciences), the March of Dimes Birth Defects Foundation (Research Grant No. 5-FY02-271), the National Institutes of Health (GM-65966), the American Chemical Society Petroleum Research Fund (38803-G4), and the UIUC Department of Chemistry (all to S.K.S.). S.K.S. is the recipient of a fellowship from The David and Lucile Packard Foundation. We thank P. Perlman for providing us with the pJD20 (ai5 $\gamma$ ) plasmid. We also thank members of the Silverman lab for discussions.

**Supporting Information Available:** This material is available free of charge via the Internet.

## REFERENCES

1. Michel, F., and Ferat, J. L. (1995) Structure and activities of group II introns, *Annu. Rev. Biochem.* **64**, 435–461.
2. Jacquier, A. (1996) Group II introns: elaborate ribozymes, *Biochimie* **78**, 474–487.
3. Martinez-Abarca, F., and Toro, N. (2000) Group II introns in the bacterial world, *Mol. Microbiol.* **38**, 917–926.
4. Fedorova, O., Su, L. J., and Pyle, A. M. (2002) Group II introns: highly specific endonucleases with modular structures and diverse catalytic functions, *Methods* **28**, 323–335.
5. Lehmann, K., and Schmidt, U. (2003) Group II introns: structure and catalytic versatility of large natural ribozymes, *Crit. Rev. Biochem. Mol. Biol.* **38**, 249–303.
6. Chin, K., and Pyle, A. M. (1995) Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5'-splice site selection, *RNA* **1**, 391–406.
7. Xiang, Q., Qin, P. Z., Michels, W. J., Freeland, K., and Pyle, A. M. (1998) Sequence specificity of a group II intron ribozyme: multiple mechanisms for promoting unusually high discrimination against mismatched targets, *Biochemistry* **37**, 3839–3849.
8. Su, L. J., Qin, P. Z., Michels, W. J., and Pyle, A. M. (2001) Guiding ribozyme cleavage through motif recognition: the mechanism of cleavage site selection by a group II intron ribozyme, *J. Mol. Biol.* **306**, 655–668.
9. Soller, M. (2006) Pre-messenger RNA processing and its regulation: a genomic perspective, *Cell. Mol. Life Sci.* **63**, 796–819.
10. Jacquier, A., and Rosbash, M. (1986) Efficient trans-splicing of a yeast mitochondrial RNA group II intron implicates a strong 5' exon-intron interaction, *Science* **234**, 1099–1104.
11. Podar, M., Perlman, P. S., and Padgett, R. A. (1995) Stereochemical selectivity of group II intron splicing, reverse splicing, and hydrolysis reactions, *Mol. Cell. Biol.* **15**, 4466–4478.
12. Vogel, J., and Borner, T. (2002) Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing, *EMBO J.* **21**, 3794–3803.
13. Michels, W. J., Jr., and Pyle, A. M. (1995) Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships, *Biochemistry* **34**, 2965–2977.
14. Jarrell, K. A., Peebles, C. L., Dietrich, R. C., Romiti, S. L., and Perlman, P. S. (1988) Group II intron self-splicing. Alternative reaction conditions yield novel products, *J. Biol. Chem.* **263**, 3432–3439.
15. Koch, J. L., Boulanger, S. C., Dib-Hajji, 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.
16. Sontheimer, E. J., Gordon, P. M., and Piccirilli, J. A. (1999) Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome, *Genes Dev.* **13**, 1729–1741.
17. Silverman, S. K. (2004) Deoxyribozymes: DNA catalysts for bioorganic chemistry, *Org. Biomol. Chem.* **2**, 2701–2706.
18. Flynn-Charlebois, A., Wang, Y., Prior, T. K., Rashid, I., Hoadley, K. A., Coppins, R. L., Wolf, A. C., and Silverman, S. K. (2003) Deoxyribozymes with 2'-5' RNA ligase activity, *J. Am. Chem. Soc.* **125**, 2444–2454.
19. Purtha, W. E., Coppins, R. L., Smalley, M. K., and Silverman, S. K. (2005) General deoxyribozyme-catalyzed synthesis of native 3'-5' RNA linkages, *J. Am. Chem. Soc.* **127**, 13124–13125.
20. Coppins, R. L., and Silverman, S. K. (2004) A DNA enzyme that mimics the first step of RNA splicing, *Nat. Struct. Mol. Biol.* **11**, 270–274.
21. Coppins, R. L., and Silverman, S. K. (2005) A deoxyribozyme that forms a three-helix-junction complex with its RNA substrates and has general RNA branch-forming activity, *J. Am. Chem. Soc.* **127**, 2900–2907.
22. Wang, Y., and Silverman, S. K. (2003) Deoxyribozymes that synthesize branched and lariat RNA, *J. Am. Chem. Soc.* **125**, 6880–6881.
23. Pratico, E. D., Wang, Y., and Silverman, S. K. (2005) A deoxyribozyme that synthesizes 2',5'-branched RNA with any branch-site nucleotide, *Nucleic Acids Res.* **33**, 3503–3512.
24. Wang, Y., and Silverman, S. K. (2005) Efficient one-step synthesis of biologically related lariat RNAs by a deoxyribozyme, *Angew. Chem., Int. Ed.* **44**, 5863–5866.
25. Burge, C. B., Tuschl, T., and Sharp, P. A. (1999) In *The RNA World*, 2nd ed. (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 525–560, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26. Perlman, P. S., and Podar, M. (1996) Reactions catalyzed by group II introns in vitro, *Methods Enzymol.* **264**, 66–86.
27. Peebles, C. L., Perlman, P. S., Mecklenburg, K. L., Petrillo, M. L., Tabor, J. H., Jarrell, K. A., and Cheng, H. L. (1986) A self-splicing RNA excises an intron lariat, *Cell* **44**, 213–223.
28. Bonitz, S. G., Coruzzi, G., Thalenfeld, B. E., Tzagoloff, A., and Macino, G. (1980) Assembly of the mitochondrial membrane system. Structure and nucleotide sequence of the gene coding for subunit 1 of yeast cytochrome oxidase, *J. Biol. Chem.* **255**, 11927–11941.
29. Jacquier, A., and Michel, F. (1987) Multiple exon-binding sites in class II self-splicing introns, *Cell* **50**, 17–29.
30. Michel, F., and Jacquier, A. (1987) Long-range intron-exon and intron-intron pairings involved in self-splicing of class II catalytic introns, *Cold Spring Harbor Symp. Quant. Biol.* **52**, 201–212.
31. Jacquier, A., and Jacquesson-Breuleux, N. (1991) Splice site selection and role of the lariat in a group II intron, *J. Mol. Biol.* **219**, 415–428.
32. Jacquier, A., and Michel, F. (1990) Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns, *J. Mol. Biol.* **213**, 437–447.
33. Peebles, C. L., Belcher, S. M., Zhang, M., Dietrich, R. C., and Perlman, P. S. (1993) Mutation of the conserved first nucleotide of a group II intron from yeast mitochondrial DNA reduces the rate but allows accurate splicing, *J. Biol. Chem.* **268**, 11929–11938.
34. Boudvillain, M., and Pyle, A. M. (1998) Defining functional groups, core structural features and inter-domain tertiary contacts essential for group II intron self-splicing: a NAIM analysis, *EMBO J.* **17**, 7091–7104.
35. Mikheeva, S., Murray, H. L., Zhou, H., Turczyk, B. M., and Jarrell, K. A. (2000) Deletion of a conserved dinucleotide inhibits the second step of group II intron splicing, *RNA* **6**, 1509–1515.
36. Bar-Shalom, A., and Moore, M. J. (2000) Tri-partite assay for studying exon ligation by the ai5 $\gamma$  group II intron, *Biochemistry* **39**, 10207–10218.
37. Gordon, P. M., Sontheimer, E. J., and Piccirilli, J. A. (2000) Kinetic characterization of the second step of group II intron splicing: role of metal ions and the cleavage site 2'-OH in catalysis, *Biochemistry* **39**, 12939–12952.
38. Santoro, S. W., and Joyce, G. F. (1997) A general purpose RNA-cleaving DNA enzyme, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4262–4266.



39. Pyle, A. M., Chu, V. T., Jankowsky, E., and Boudvillain, M. (2000) Using DNAzymes to cut, process, and map RNA molecules for structural studies or modification, *Methods Enzymol.* **317**, 140–146.
40. Lambowitz, A. M., and Belfort, M. (1993) Introns as mobile genetic elements, *Annu. Rev. Biochem.* **62**, 587–622.
41. Lambowitz, A. M., and Zimmerly, S. (2004) Mobile group II introns, *Annu. Rev. Genet.* **38**, 1–35.
42. Collins, C. A., and Guthrie, C. (2000) The question remains: is the spliceosome a ribozyme? *Nat. Struct. Biol.* **7**, 850–854.
43. Weiner, A. M. (1993) mRNA splicing and autocatalytic introns: distant cousins or the products of chemical determinism? *Cell* **72**, 161–164.
44. Gaur, R. K., McLaughlin, L. W., and Green, M. R. (1997) Functional group substitutions of the branchpoint adenosine in a nuclear pre-mRNA and a group II intron, *RNA* **3**, 861–869.
45. Newman, A. (1997) RNA splicing: out of the loop, *Curr. Biol.* **7**, R418–420.
46. Krämer, A. (1996) The structure and function of proteins involved in mammalian pre-mRNA splicing, *Annu. Rev. Biochem.* **65**, 367–409.
47. Reed, R. (2000) Mechanisms of fidelity in pre-mRNA splicing, *Curr. Opin. Cell Biol.* **12**, 340–345.
48. Brow, D. A. (2002) Allosteric cascade of spliceosome activation, *Annu. Rev. Genet.* **36**, 333–360.
49. Jurica, M. S., and Moore, M. J. (2003) Pre-mRNA splicing: awash in a sea of proteins, *Mol. Cell* **12**, 5–14.
50. Wang, Y., and Silverman, S. K. (2003) Characterization of deoxyribozymes that synthesize branched RNA, *Biochemistry* **42**, 15252–15263.
51. Milligan, J. F., Groebe, D. R., Witherell, G. W., and Uhlenbeck, O. C. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates, *Nucleic Acids Res.* **15**, 8783–8798.
52. Jarrell, K. A., Dietrich, R. C., and Perlman, P. S. (1988) Group II intron domain 5 facilitates a trans-splicing reaction, *Mol. Cell. Biol.* **8**, 2361–2366.
53. Grosshans, C. A., and Cech, T. R. (1991) A hammerhead ribozyme allows synthesis of a new form of the *Tetrahymena* ribozyme homogeneous in length with a 3' end blocked for transesterification, *Nucleic Acids Res.* **19**, 3875–3880.
54. Ferré-D'Amaré, A. R., and Doudna, J. A. (1996) Use of *cis*- and *trans*-ribozymes to remove 5' and 3' heterogeneities from milligrams of in vitro transcribed RNA, *Nucleic Acids Res.* **24**, 977–978.
55. Miduturu, C. V., and Silverman, S. K. (2005) DNA constraints allow rational control of macromolecular conformation, *J. Am. Chem. Soc.* **127**, 10144–10145.
56. Wang, Y., and Silverman, S. K. (2006) A general two-step strategy to synthesize lariat RNAs, *RNA* **12**, 313–321.
57. Kwakman, J. H., Konings, D., Pel, H. J., and Grivell, L. A. (1989) Structure-function relationships in a self-splicing group II intron: a large part of domain II of the mitochondrial intron *al5* is not essential for self-splicing, *Nucleic Acids Res.* **17**, 4205–4216.
58. Schmidt, U., Riederer, B., Morl, M., Schmelzer, C., and Stahl, U. (1990) Self-splicing of the mobile group II intron of the filamentous fungus *Podospora anserina* (COI 1) in vitro, *EMBO J.* **9**, 2289–2298.
59. Hebbar, S. K., Belcher, S. M., and Perlman, P. S. (1992) A maturase-encoding group IIA intron of yeast mitochondria self-splices in vitro, *Nucleic Acids Res.* **20**, 1747–1754.
60. Fedorova, O., Mitros, T., and Pyle, A. M. (2003) Domains 2 and 3 interact to form critical elements of the group II intron active site, *J. Mol. Biol.* **330**, 197–209.