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

Experimental Procedures

DNA oligonucleotides and RNA transcription templates. For preparing the branched RNAs using the 7S11 deoxyribozyme, the RNA substrates were transcribed using T7 RNA polymerase and an appropriate DNA template. The D123 plasmid templates were constructed as follows. First, a DNA fragment was amplified by PCR using Taq polymerase and the pJD20 plasmid. The forward primer was 5′-GGTGAAAACCTCTGACACAT**GAATTC**AATTAATACGACTCACTATA**X**CGGTCTGAAAGTTATCAT-3′ (**X** = GGAG for –1 RNA; GAG for +1 RNA; AG for +2 RNA; and G for +3 RNA; T7 promoter site is underlined). The reverse primer was 5′-TCGCGCGTTTCGGTGATGAC**AAGCTT**CTCTTCACCTATAGTATAAGTTAGCAGATTTTCATCT-3′. The boldface EcoRI and HindIII sites were introduced to enable subcloning, and the underlined EarI site was introduced to enable linearization. The PCR product was ligated into pCR2.1-TOPO (Invitrogen), and the EcoRI-HindIII fragment was subcloned into pUC19.

The mutated pJD20 plasmid required for preparing the –1 RNA (with a C:G base-pair flip) was prepared using the QuikChange Multi Site-Directed Mutagenesis Kit (Stratagene). The two primers for introducing mutations at positions –1 (the last nt of IBS1) and +329 (the first nt of EBS1) were 5′-GTGGGACATTTT**G**GAGCGGTCTGAAAGTTATC-3′ (IBS1 underlined) and 5′-ATGTTATGTATT**C**GAAATGAGCATA-CGATAAATCATATAACC-3′ (EBS1 underlined). The two boldface nucleotides introduced the C→**G** mutation at the –1 position and the G→**C** mutation at the +329 position.

The double-stranded D56 transcription templates were prepared by PCR using pJD20 as the template. The primers were as follows. Primer 1, 5′-ACGCACGCTGTAATACGACTCACTATAGCGTGAGCCGTATG-CGAT-3′ (forward primer, with T7 promoter sequence underlined); primer 2, 5′-ACCATGCCGGCC**ATCCCGA-TAGGTAGACCTTTACA**-3′ (for introducing the HDV ribozyme sequence; overlap sequence with the HDV ribozyme is underlined and overlap sequence with D56 is boldface); primer 3, 5′-GGTCCCATTCGCCATTC-CGAAGAATGTTGCCCAGCCGGCGCCAGCGAGGAGGCTGGGACCATGCCGGCC-3′ (complementary to the HDV ribozyme sequence; underlined sequence overlaps with primer 2); primer 4, 5′-GGTCCCATTCGCCATTC-3′ (reverse primer). The PCR reaction used 100 pmol each of primers 1 and 4 plus 10 pmol each of primers 2 and 3 in 100 μl reaction volume with Taq polymerase. Following PCR, the DNA template was ethanol-precipitated and used without further purification. The DNA template for transcribing D56 with a 6-nt 3′-exon was prepared by PCR in a similar manner, with primer 2 changed to 5′-ACCATGCCGGCC*CATAGT***ATCCCGATAGGTAGACCTT**-3′ (the six italicized nucleotides are the complement of the 3′-exon sequence). When the 3′-exon was not present at the 3′-terminus of D56, the 2′,3′-cyclic phosphate was removed by treatment with T4 polynucleotide kinase (PNK; Fermentas) as follows. A sample of D56 (600 pmol) was incubated in 50 mM Tris (pH 7.6), 10 mM $MgCl_2$, 5 mM DTT, 0.1 mM spermidine, and 0.1 mM EDTA in 120 μl total volume with 120 U of T4 PNK at 37 ºC for 1 h, then extracted with phenol-chloroform and ethanol-precipitated.

For preparing the control (‘C’) branched RNA by ai5γ catalysis, the 5′-exon–D123 RNA was transcribed using an EarI-linearized plasmid derived from plasmid pJD20 as the template. The plasmid was constructed by PCR using pJD20 as the template. The forward primer was 5′-ACGCACGCTGTAATACGACTCACTATA-3′ (T7 promoter site is underlined); the reverse primer was 5′-TCGC-GCGTTTCGGTGATGAC**AAGCTT**CTCTTCACCTATAGTATAAGTTAGCAGATTTTCATCT-3′. In the reverse primer, the

boldface HindIII site was introduced to enable subcloning, and the underlined EarI site was introduced to enable linearization. The PCR product was ligated into pCR2.1-TOPO (Invitrogen), and the EcoRI-HindIII fragment was subcloned into pUC19. The D56–3′-exon RNA was transcribed using a double-stranded PCR product derived from plasmid pJD20 as the template. The forward primer was 5′-ACGCACGCTGTAATACGACTCACTATAGCGTGAGCCGTATGCGAT-3′ (T7 promoter site is underlined); the reverse primer was 5′-CTAGGCATACCATTAATACCTAAGAAATGCAT-3′.

Synthesis of branched RNA. The 7S11 deoxyribozymes used to synthesize the branched RNAs were as follows (enzyme region loops are underlined, and the P4 sequence is boldface):

–1: 5′-TGATAACTTTCAGACCGCAGTGCAGGGCCCCGAGGCTCGG**CTCC**AGGTAGACCTTTAC-3′

+1: 5′-ATGATAACTTTCAGACCCAGTGCAGGGCCCCGAGGCTCGG**GCTC**AGGTAGACCTTTAC-3′

+2: 5′-TATGATAACTTTCAGACCAGTGCAGGGCCCCGAGGCTCGG**CGCT**AGGTAGACCTTTAC-3′

+3: 5′-TTATGATAACTTTCAGACAGTGCAGGGCCCCGAGGCTCGG**CCGC**AGGTAGACCTTTAC-3′

Disruptor DNA oligonucleotides were designed for inclusion during ai5γ branched RNA formation by the 7S11 deoxyribozyme. The deoxyribozyme must bind with intron RNA sequences near the branch site, but secondary structure in the intron itself competes with the deoxyribozyme binding. Therefore, disruptors were used to compete with the secondary structure and allow the deoxyribozyme to bind. The disruptor sequences were as follows:

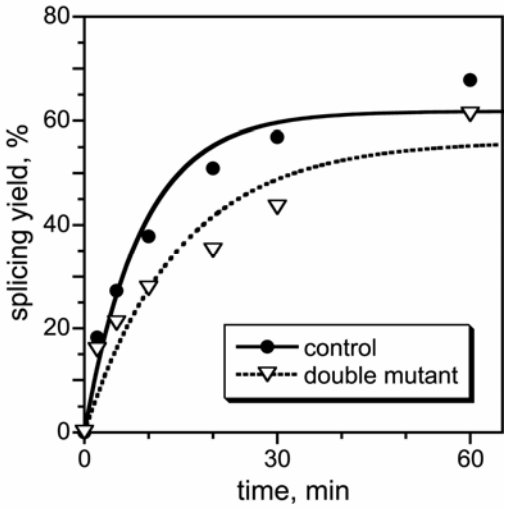
Disruptor for D123: 5′-AATTTATAGGGTTTATT**A**TGTTTT**GCTG**TAAATACGTAAATATCTAACTTAGC-3′

Disruptor for D56: 5′-AAGTTTTCCCCC**A**GTAAGAACCGTACGTGCGACTTTCAT-3′

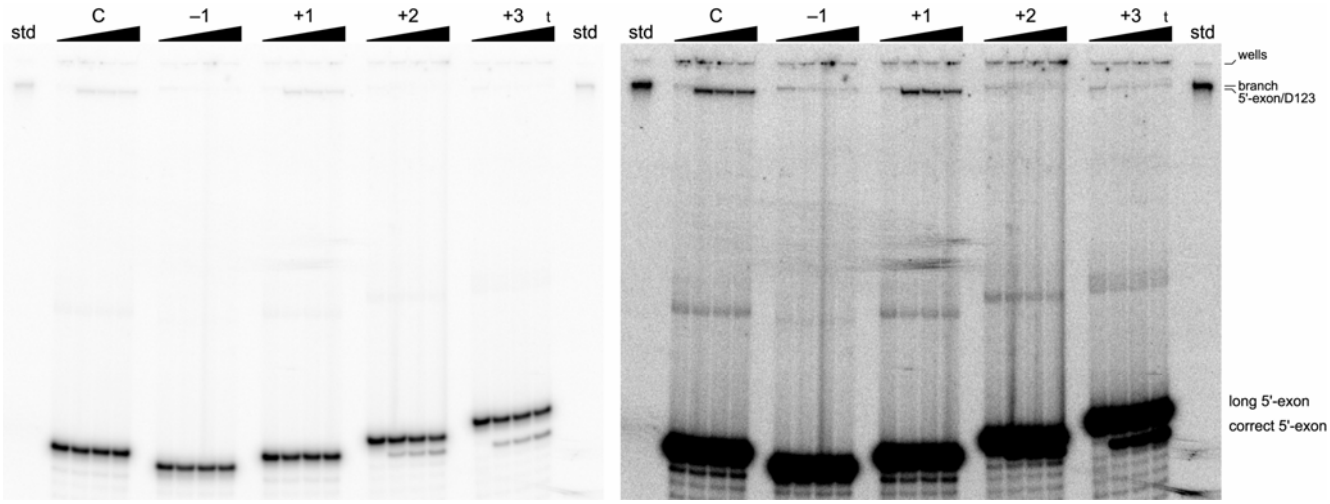
The disruptor for D123 is complementary to nucleotides G379–U431 of the ai5γ sequence. The disruptor for D56 is complementary to nucleotides A827–U865 of the ai5γ sequence. Underlined nucleotides are those with potential base-pairs between the disruptor and the 7S11 deoxyribozyme binding arms. Double-underlined nucleotides are those with Watson-Crick matches between the disruptor and the 7S11 deoxyribozyme. Boldface nucleotides indicate those positions where mismatches or G-U wobble pairs were intentionally created between the disruptor and the 7S11 deoxyribozyme. The intentional mismatches and wobble pairs were created whenever a stretch of three or more nucleotides would have been Watson-Crick matches between the disruptor and the 7S11 deoxyribozyme.

Verification of branched RNA structures. First, each large branched RNA was cleaved by the 10–23 deoxyribozyme 5′-GTAGACCTTTACAAGGGCTAGCTACAACGATTTCCCCCGG-3′ (which cleaves between A862 and C863 in D56; the binding arms are underlined). The cleavage product (700 nt for the +1 RNA) was 5′-^{32}P-radiolabeled and purified by 5% denaturing PAGE. Second, the RNA was cleaved by the 10–23 deoxyribozyme 5′-TTATATGGTAAATATTATTTAGGCTAGCTACAACGAGATAACTTTCAGACC-3′ (which cleaves between A20 and U21 in D123; the binding arms are underlined). The cleavage product (45 nt for the +1 RNA) was purified by 20% denaturing PAGE. To prepare the standard RNAs, D56 and D56–3′-exon were cleaved by the first 10-23 deoxyribozyme, and the 3′-fragments of the cleavage products were 5′-^{32}P-radiolabeled.

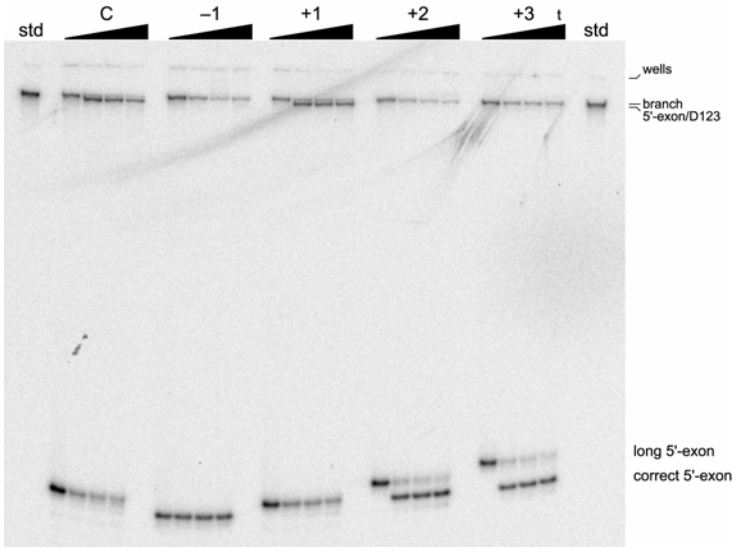
Sequencing of ligated exons formed from +2 mis-spliced RNA. To confirm the sequences of the ligated exons from the +2 mis-spliced RNA, the D56–3′-exon RNA was transcribed with inclusion of a specific and arbitrary 62-nucleotide sequence after the 6-nt 3′-exon. The last 22 nucleotides of this sequence were used as a primer binding site for RT-PCR, using SuperScript II RT (Invitrogen) and Taq polymerase (NEB). The 108-base pair PCR product was purified by 2% agarose gel electrophoresis and then cloned (TOPO-TA; Invitrogen). Automated sequencing over the exon ligation junction was performed. The sequences of eight clones revealed the correctly ligated exons; see representative data in Figure S5.



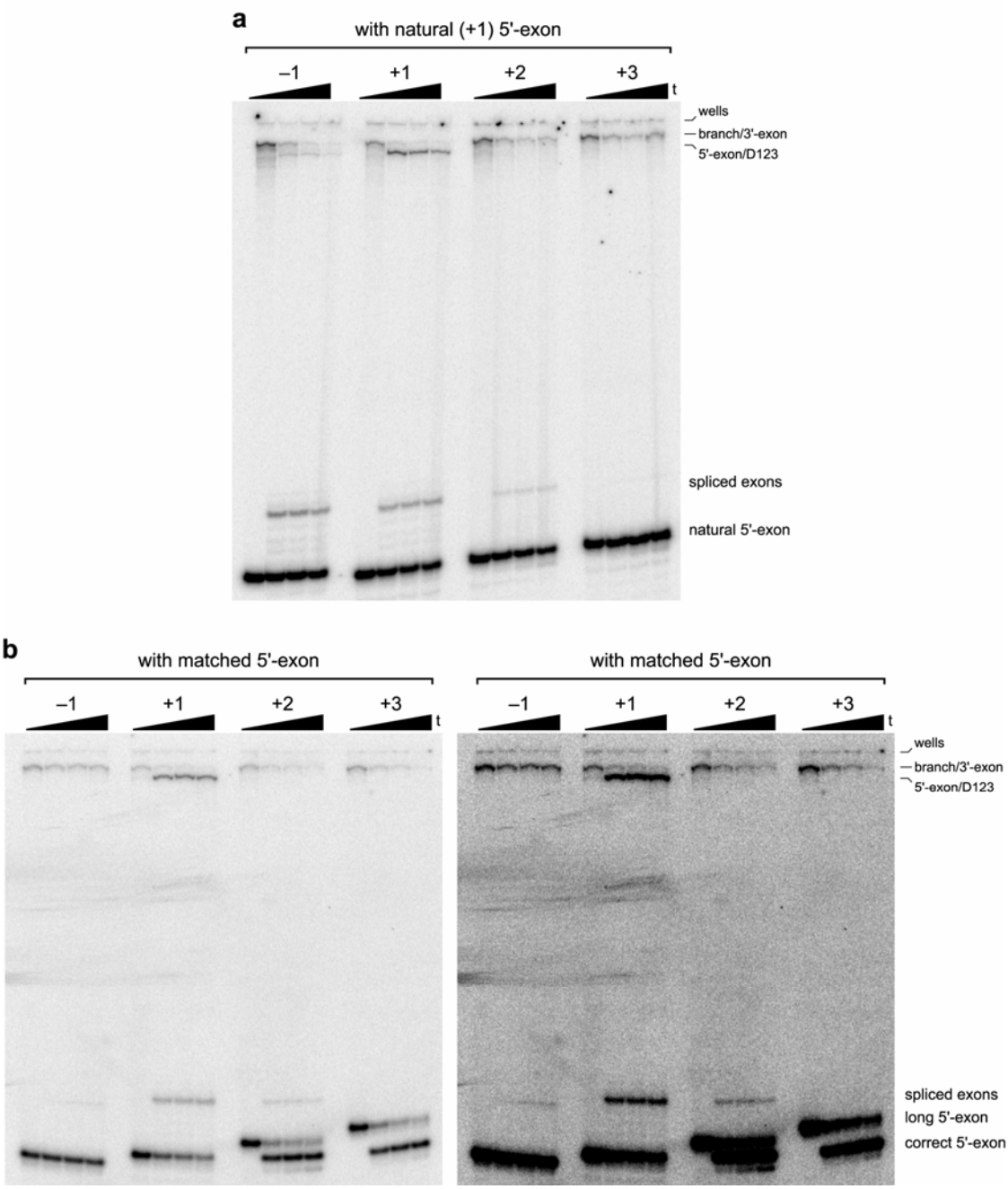
Supplementary Figure 1. Changing the –1 nucleotide of the 5′-exon along with the corresponding EBS1 nucleotide of D123 from C:G to G:C does not substantially affect the forward first step of splicing (the first step of the two-step pathway is rate-limiting). The fit values of k_{obs} were 0.11 min^{-1} and 0.07 min^{-1} (control and double mutant, respectively). Procedure: 5 pmol of pre-mRNA in 8 µl of 5 mM MOPS (pH 7.0) and 1 mM EDTA were annealed by heating at 95 °C for 1 min and cooling at room temperature for 1 min. The sample volume was raised to 10 µl with final concentrations of 40 mM MOPS (pH 7.0), 100 mM $MgCl_2$, and 500 mM NH_4Cl. The sample was incubated at 45 °C. Aliquots were withdrawn at desired timepoints and quenched onto stop solution (80% formamide, 1× TB, 50 mM EDTA, 0.025% each bromophenol blue and xylene cyanol), followed by 5% denaturing PAGE and exposure to a PhosphorImager screen. The pre-mRNA was transcribed from a HindIII-linearized pJD20 plasmid (wild-type control or double mutant, prepared as described in the Experimental Procedures of this Supporting Information). The transcripts were 1501 nt in length (293-nt 5′-exon + 887-nt intron + 321-nt 3′-exon).



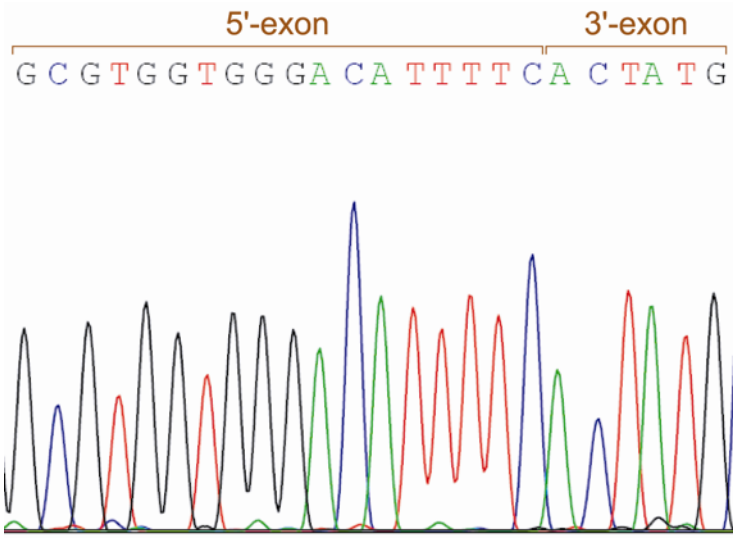
Supplementary Figure 2. Additional assays for the reverse of the first step of splicing, using an excess of radiolabeled 5′-exon. The image on the right is at 100× exposure relative to the image on the left. The 5′-exon was 5′-^{32}P-radiolabeled and present in 20-fold excess relative to internally ^{32}P-radiolabeled branched RNA (t = 0, 0.5, 1, and 2 h; 20% PAGE; std = +1 branched RNA). Note the hydrolytic cleavage activity evident for the +2 and +3 5′-exons at the bottom of the gel. Timepoints taken after 7 h showed 40–50% hydrolytic cleavage. At the top of the gel, no reverse splicing product (5′-exon–D123) was detected at 7 h for the mis-spliced –1, +2, or +3 introns. At the top of the gel, the product (5′-exon–D123) bands are more intense than the reactant (branch–3′-exon) bands because the 5′-^{32}P-radiolabeled 5′-exon is of higher specific activity than the internally ^{32}P-radiolabeled branched RNA. The procedure was as described in the text, using internally ^{32}P-radiolabeled branched RNA (0.025 pmol) and 5′-^{32}P-radiolabeled matched 5′-exon (0.5 pmol).



Supplementary Figure 3. Additional assays for the reverse of the first step of splicing, using an excess of the branched RNA. The internally ^{32}P-radiolabeled branched RNA was present in 10-fold excess relative to 5′-^{32}P-radiolabeled 5′-exon (t = 0, 0.5, 1, and 2 h; 20% PAGE; std = +1 branched RNA). Note that hydrolytic cleavage proceeds essentially to completion within 0.5 h for the +2 and +3 5′-exons. At the top of the gel, no reverse splicing product (5′-exon–D123) was detected at 7 h for the mis-spliced –1, +2, or +3 introns. The procedure was as described in the text, using internally ^{32}P-radiolabeled branched RNA (0.1 pmol) and 5′-^{32}P-radiolabeled matched 5′-exon (0.01 pmol).



Supplementary Figure 4. Expanded view of gel images from Figure 4, which depicts assays for the second step of splicing. a) Image from Figure 4a. As expected, the bands at the top of the gel provide clear evidence for the reverse of the first step of splicing for the +1 branch–3′-exon RNA. In addition, the –1 mis-spliced branch–3′-exon RNA intermediate is observed to participate in the reverse of the first step of splicing, but in nature the –1 intermediate would never have the opportunity to react with the +1 5′-exon. b) Image from Figure 4b. The image on the right is at 10× exposure relative to the image on the left, which is at the same exposure as the image in Figure 4b itself. Again note that the upper bands provide clear evidence for the reverse of the first step of splicing for the +1 branch–3′-exon RNA. Note also that the –1 branch–3′-exon RNA does not undergo the reverse of the first step of splicing with the matched (–1) 5′-exon, which is the 5′-exon that would be present naturally. At the top of the gel, the product (5′-exon–D123) bands are more intense than the reactant (branch–3′-exon) bands because the 5′-^{32}P-radiolabeled 5′-exon is of higher specific activity than the internally ^{32}P-radiolabeled branched RNA.



Supplementary Figure 5. Representative automated sequencing run that illustrates the expected sequence of the correctly ligated exons obtained from the +2 mis-spliced RNA after RT-PCR. Eight such clones had the correct sequence at the splice junction.