

Splicing of an intervening sequence by protein-free human snRNAs

Yasaman Jaladat, Bing Zhang, Afshin Mohammadi and Saba Valadkhan*

Center for RNA Molecular Biology; Case Western Reserve University; Cleveland, OH USA

Key words: U6, U2, snRNA, splicing, spliceosome, catalysis, ribozyme, intron

Significant structural and mechanistic similarities between the spliceosomal snRNAs and catalytically critical domains of self-splicing group II introns have led to the hypothesis that the spliceosomes and group II introns may be evolutionarily related. We have previously shown that *in vitro*-transcribed, protein-free U6 and U2 snRNAs can catalyze a two-step splicing reaction *in trans* on two short RNA oligonucleotides that is identical to the splicing reactions performed by many self-splicing group II introns. Here we show that the same two snRNAs can perform splicing *in cis* by removal of an intervening sequence from a model substrate. These results prove that the protein-free snRNAs are competent to perform splicing on pre-mRNAs and further strengthen the possibility of an evolutionary relationship to group II introns.

Almost all eukaryotic genes contain intervening sequences or introns which must be removed from primary genomic transcripts before they can be used by the cell. The vast majority of the intervening sequences, also called introns, are removed by the spliceosome, a gigantic ribonucleoprotein assembly found in all eukaryotes.¹⁻³ A small subset of introns, the self-splicing group I and group II introns, do not depend on the spliceosome for their removal. These introns, which have RNA or RNP-mediated enzymatic activity, catalyze their own removal from the primary transcripts in which they reside.^{4,5} Interestingly, the removal of introns, also called splicing, is performed by the spliceosome through a reaction which is identical in many respects to the one catalyzed by self-splicing group II introns.⁶⁻⁹ Both systems perform splicing through an identical reaction pathway that involves two consecutive transesterification reactions^{4,10} and use divalent cations in their catalytic strategy.¹¹⁻¹⁶ Perhaps most intriguingly, the RNA components of the spliceosome, called the small nuclear RNAs or snRNAs, have striking structural and functional similarities to the catalytically crucial domains of group II introns.^{6,8,9,17} Two of the five spliceosomal snRNAs, U2 and U6, form a functionally required basepaired complex in the spliceosomal catalytic core which closely resembles domain V of group II introns in both structure and function¹⁰ (Fig. 1A). These similarities, in aggregate, have led to the hypothesis that the snRNAs may be evolutionarily related to the group II introns,¹⁸⁻²⁰ which in turn raises the possibility that the snRNAs may play a catalytic role in the spliceosome.

As a first step toward defining the role of snRNAs in the spliceosomal catalytic core, we have analyzed the structure and function of the U6 and U2 snRNAs in isolation in the absence of all other spliceosomal factors. We have shown that upon incubation in the presence of divalent cations, the two snRNAs efficiently

form a basepaired complex closely similar to the one found in the activated spliceosomes.²¹ More recently, we showed that this *in vitro*-assembled, protein-free U6/U2 complex can perform a two-step *trans* splicing reaction on two short RNA oligonucleotides which play the role of small fragments of pre-mRNAs at the 5' and 3' splice sites (Fig. 1B).^{22,23} The first step of this *trans*-splicing reaction involves a U6/U2 catalyzed hydrolysis on the substrate carrying the 5' splice site equivalent, Exon1, resulting in the release of a fragment which has a free 3' OH group (Fig. 1C). In the second step of this reaction, similar to the second step of splicing in spliceosomes and group II introns, the newly released 3'OH of the Exon1 fragment undergoes a transesterification reaction with an internal phosphate in Exon2. This reaction leads to the formation of a linear RNA in which fragments of both substrates are joined together through a 3'-5' phosphodiester bond and is chemically identical to the splicing reaction catalyzed by group II self-splicing introns and the second step of spliceosomal splicing (Fig. 1C).^{22,23} Although current data indicate that the first step of spliceosomal splicing exclusively involves branching, in many self-splicing group II introns the first step of splicing occurs through hydrolysis both *in vivo* and *in vitro*.²⁴⁻²⁶ Thus, catalysis of the first step through hydrolysis is a physiological alternative to branching.

Although this snRNA-catalyzed *trans*-splicing reaction has significant similarities to the splicing reactions observed in nature both in terms of its chemistry and the required snRNA sequences,^{10,22} an important remaining question is whether the U6/U2 complex can catalyze a canonical splicing reaction *in cis*. In other words, if U6/U2 can perform the removal of an intervening sequence and exon ligation on an intact pre-mRNA-like substrate, which contains the exonic and intronic sequences in a spatially constrained arrangement on a single, continuous strand

*Correspondence to: Saba Valadkhan; Email: saba.valadkhan@case.edu

Submitted: 12/02/10; Revised: 02/18/11; Accepted: 02/22/11

DOI: 10.4161/rna.8.3.15386

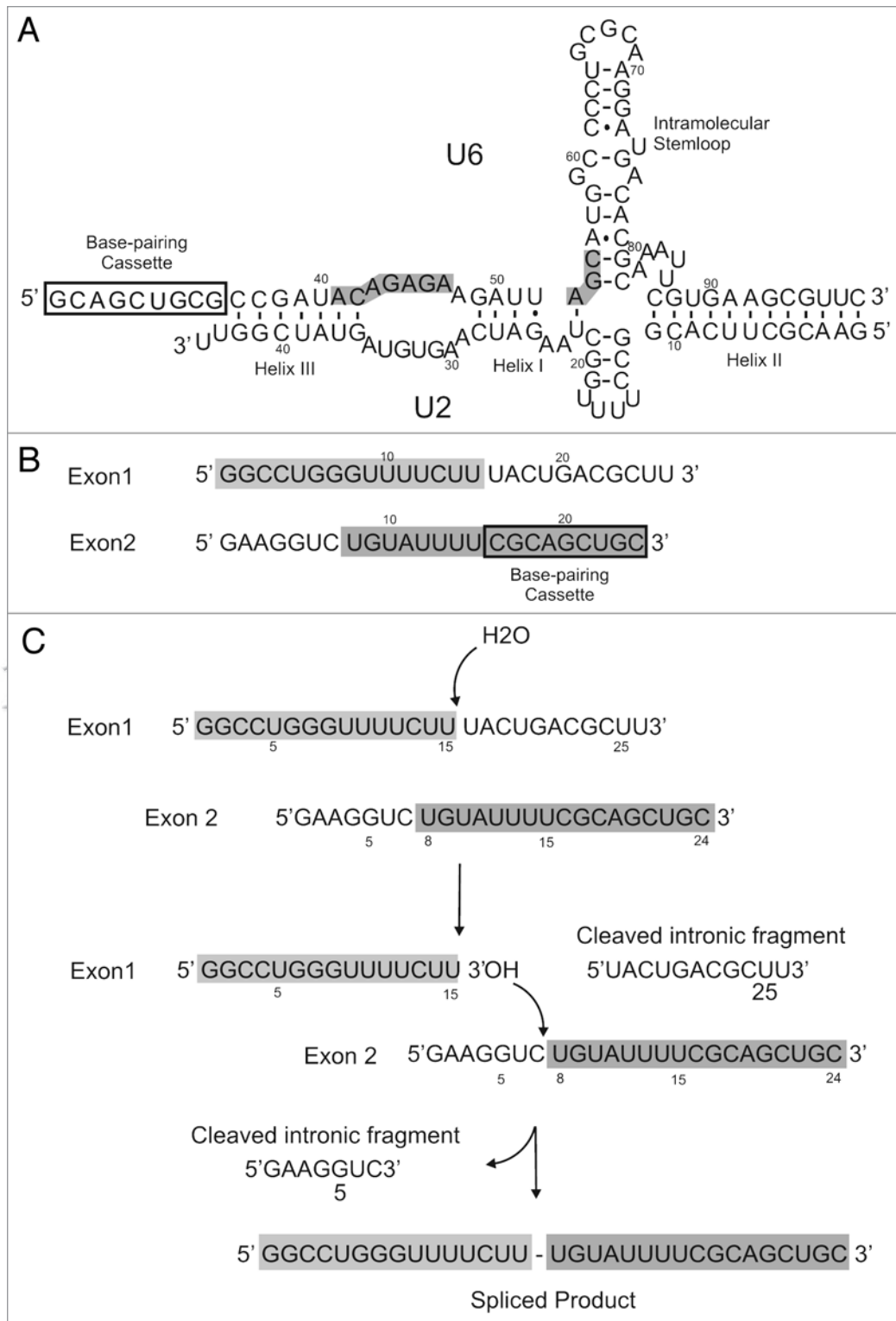


Figure 1. The snRNA-catalyzed trans splicing reaction. (A) The U6/U2 basepaired complex. The basepaired helices I, II and III and the intramolecular stemloop of U6 are shown. The highlighted sequences are evolutionarily invariant U6 sequences. The boxed nucleotides basepair to a short RNA sequence at the 3' end of Exon2 substrate. The numbers close to U6 and U2 nucleotides reflect the human numbering system for these snRNAs. (B) The sequence of the trans splicing substrates. The boxed nucleotides basepair to a short sequence at the 5' end of U6 snRNA. The numbers indicate position from the 5' end. The highlighted regions are the exonic sequences and are incorporated into the product. The rest of the substrates is removed as intronic sequences during the trans-splicing reaction. (C) The trans splicing reaction. The substrates and fragments resulting from each step of the reaction are shown. Numbers reflect position from the 5' end. See the text for details.

of RNA. Since the snRNAs may have evolved from a self-splicing ribozyme, it is important to determine the extent to which their primordial catalytic activity has been preserved during the evolution of the eukaryotic spliceosomes.

To this end, we designed pre-mRNA-like substrates which contained the sequence of Exon1 linked to the 5' end of Exon2 by a RNA linker sequence (Fig. 2A). We performed an in-silico screening to select a linker sequence that did not induce the formation of stable intramolecular structures in order to increase the likelihood that the sequences in the designed substrate (E1-int-E2, Fig. 2A) will be available for forming interactions with the U6/U2 complex. Next, we incubated the in vitro-transcribed E1-int-E2 substrate with the U6/U2 complex under a variety of conditions and analyzed the reaction mixture for product formation.

If U6/U2 could indeed perform a splicing reaction in cis on the E1-int-E2 substrate, the product should contain fragments from the 5' and 3' ends of the substrate ligated to each other, with the intervening nucleotides removed as an intron. It was possible that the Exon1- and Exon2-like sequences present in each E1-int-E2 molecule could interact with U6/U2 in the same way that the Exon1 and Exon2 oligonucleotides did in the trans splicing reaction,^{22,23} leading to the formation of a product similar to the one formed in the trans-splicing reactions. On the other hand, they could form alternative interactions, leading to the formation of multiple products or a product with a different length compared to the one formed in the trans splicing reaction (Fig. 1C). Since in these reactions, the E1-int-E2 precursor will be larger than the product(s) formed, the degradation products resulting from the high pH and magnesium concentrations would prohibit the direct visualization of the product using radiolabeling techniques. To circumvent this problem, we took a PCR-based approach to detect any products which may have formed as a result of incubation of the E1-int-E2 substrate with the U6/U2 complex. After the incubation period was over, we used gel electrophoresis on a denaturing PAGE to fractionate the RNAs in the reaction mixture based on size. Next, we used short PCR primers targeting the nucleotides at the beginning and end of the E1-int-E2 (Fig. 2A) to probe the size-fractionated RNAs for the presence of any RNA species that may correspond to a spliced product. We set up control reactions which lacked the U6/U2 complex, along with parallel trans-splicing reactions which contained Exon1 and Exon2 substrates as positive controls and size markers.

RT-PCR reactions indicated that all except one of the fractions did not contain any product-like RNA species. In fractions corresponding to RNA sizes close to or identical to that of the precursor E1-int-E2, we obtained PCR products which resulted from its amplification, as expected (data not shown). The only fraction which contained a product-like, RT-PCR-amplified species was the one which contained RNAs ranging from 25 to 35 nucleotides in size. RT-PCR reactions on this fraction indicated the presence of a single RNA species which was identical in size to the 32-nucleotide-long product formed in the trans-splicing reactions (Fig. 2B). The formation of this product was dependent on the presence of U6 and U2 snRNAs in the reaction and formed with the same efficiency as the trans-splicing reaction, as shown by sensitive radiolabeled PCR assays (Fig. 2B and data not

shown). This product was not observed in any other fractions or in reactions in which the fractionated RNAs or the PCR reagents were omitted (Fig. 2C). Further, replacing one of the PCR primers with a primer containing its complementary sequence did not result in product formation (Fig. 2C). The formation of such a product requires the removal of 35 nucleotides from the sequences located in the middle of E1-int-E2 and ligation of the sequences at its 5' and 3' ends together into a linear RNA species which can be amplified by RT-PCR. The fact that only one such RNA was formed in these reactions and that it is identical in size to the product formed in trans-splicing reactions suggests that this product likely forms through a reaction pathway similar to the one observed in trans-splicing, except that it occurs in cis.

To ensure that the reaction indeed occurred in cis, it was important to ascertain that the E1-int-E2 substrate was not cleaved into two Exon1 and Exon2-like fragments through a random cleavage event as a prerequisite prior to the formation of this product. In order to investigate this possibility, and also to gain further insight into the requirements for this reaction, we designed a substrate that had a linker sequence shortened by seven nucleotides (E1-int-E2-7, Fig. 2A). We reasoned that if the reaction occurs in cis, a shorter linker region should result in spatial constraints for positioning of the reaction sites in the U6/U2 complex and thus, should affect product formation. On the other hand, if a random cleavage reaction was a prerequisite for product formation, shortening of the linker sequence by 7 nucleotides should not have a major effect. Interestingly, the use of E1-int-E2-7 substrate in catalytic assays completely blocked product formation (Fig. 2D). While these results strongly suggested that product formation does not result from trans splicing of randomly degraded products of the E1-int-E2 substrate, it was important to ensure that the pattern of random cleavage of the E1-int-E2 and E1-int-E2-7 substrates were similar under the reaction conditions. If a significantly higher percentage of the E1-int-E2 substrate were randomly cleaved during the reaction compared to E1-int-E2-7, the product formed from the longer substrate could still be the result of a reaction in trans. We analyzed the pattern of degradation of the two substrates under the reaction conditions, which indicated a very similar pattern and extent of random cleavage events (Fig. 2E).

Finally, the lack of reactivity of the substrate with the shorter linker could be due to stable intramolecular interactions which prohibited its binding to the U6/U2 complex or blocked one of the catalytic steps of the reaction. While our in silico studies and analysis of the magnesium dependent degradation patterns described above suggested that the shorter substrate did not significantly differ from E1-int-E2 in terms of forming intramolecular structures, they don't completely rule out this possibility. To partially address this shortcoming, we tested the ability of the two substrates for interacting with the U6/U2 complex in a previously characterized functional assay.²³ We added an excess amount of E1-int-E2 or E1-int-E2-7 to trans-splicing reactions to determine if these substrates could compete with the Exon1 and Exon2 for binding to U6/U2. We reasoned that if they compete for binding to the same binding site, the addition of the cis-splicing substrates should block the trans-splicing reaction, which was

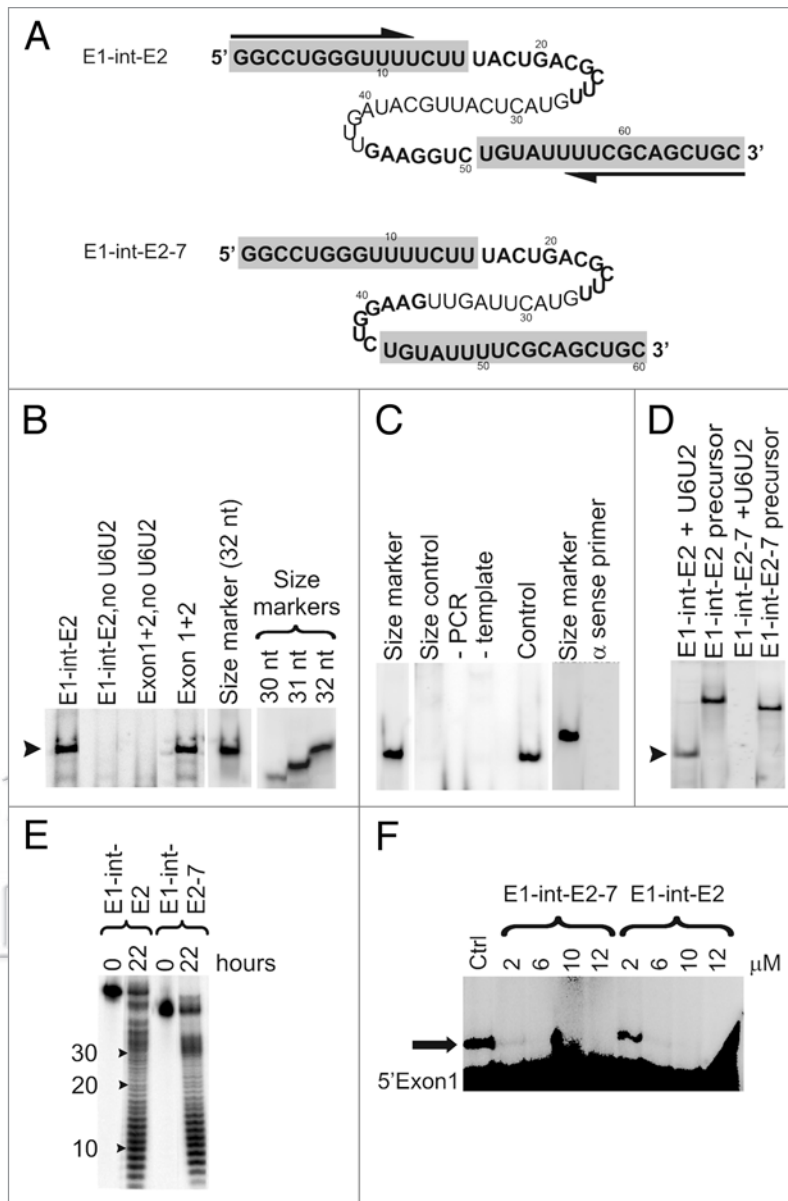


Figure 2. The protein-free snRNAs can catalyze the removal of intervening sequences in cis. (A) The cis splicing substrates. The exonic sequences are highlighted in gray. Numbers indicate position from the 5' end. The sequences derived from the Exon1 and Exon2 trans-splicing substrates are in bold face font. The arrows on E1-int-E2 mark the sequences that bind the RT-PCR primers. (B) The U6/U2 complex can catalyze the removal of intervening sequences in a cis splicing reaction. Arrowhead points to the DNA species resulting from RT-PCR on the 25–35 nucleotide long fraction of the reactions containing E1-int-E2, as detailed in the text. Lanes marked Exon1 + 2 contain the result of RT-PCR on trans-splicing reactions used as controls. Lanes marked no U6U2 contain the result of RT-PCR on control reactions which lack U6/U2. A 5' labeled double stranded DNA containing the sequence of the trans-splicing product is used as size marker (lane labeled size marker, 32 nt). Two additional size markers, missing one or two basepairs (lanes marked 30 nt and 31 nt) are shown. (C) The requirements for the formation of the RT-PCR-generated species corresponding to the cis-splicing product. The lane marked size control contains the fraction with RNAs 35–50 nucleotides in length. The PCR mixture is omitted in the reaction loaded into the -PCR lane. In the lane marked -template the RT-PCR reaction is performed without the addition of the fractionated RNAs from the cis-splicing reactions. The lane marked “ α sense primer” contains a PCR reaction in which one of the originally used primers was replaced by one containing its complementary sequence. (D) Shortening of the linker sequence blocks the cis splicing reaction. The lanes marked precursor contain the result of RT-PCR reactions performed on unreacted E1-int-E2 or E1-int-E2-7 substrates. The other two lanes contain the 25–35 nucleotide long fraction of the cis-splicing reactions on substrates indicated above each lane. Arrowhead points to the cis-splicing product. (E) The cis splicing substrates have a similar extent and pattern of random cleavage events under the reaction conditions. The duration of incubation under the reaction conditions is shown above each lane. The identity of each substrate is shown. Numbers to the left indicate the length of each fragment. (F) The cis-splicing substrate can compete with trans-splicing substrates for binding to the U6/U2 complex. The identity and amount of the cis-splicing substrate added is shown on top. No competing cis-splicing substrates has been added to the lane marked Ctrl. Arrow points to the trans-splicing product. The location of the unreacted 5' Exon1 is shown.

indeed the case for both E1-int-E2 and E1-int-E2-7 (Fig. 2F). These results indicate that both the E1-int-E2 and E1-int-E2-7 compete with at least one of the two trans-splicing substrates for binding to U6/U2 and thus, a complete inability to interact with the U6/U2 complex is not the basis for lack of reactivity of the E1-int-E2-7 substrate. Taken together, the above results strongly suggest that the removal of the intervening sequences from the E1-int-E2 substrate during product formation is performed in cis. Thus, the lack of reactivity of the E1-int-E2-7 in product formation likely reflects an incorrect positioning on the U6/U2 complex caused by the spatial constraints resulting from the short linker sequence. Importantly, these experiments also rule out the possibility of a reaction in trans involving E1-int-E2 or E1-int-E2-7, since in such a scenario the radiolabeled Exon1 should have been able to form a product with the cis substrates across the range of concentrations tested, which was not observed.

Existing data indicate that both group II introns and the spliceosome can efficiently perform trans-splicing reactions on split introns in vitro.^{4,27,28} However, in vivo the vast majority, if not all, of splicing reactions catalyzed by the major spliceosome are performed in cis on substrates in which the 5' and 3' splice sites are on the same RNA strand. While the cis and trans splicing reactions follow an identical catalytic pathway, there are important differences between them in both the spliceosome and group II self-splicing introns, including the presence of a rate-limiting conformational change before the second catalytic step in cis splicing in both systems.^{13,16} Since a branch site sequence is not present in our minimal cis splicing substrates, the first step of the cis splicing reaction described above by necessity occurs through hydrolysis, as previously described in self-splicing group II introns.²⁴ Thus, our data indicate that similar to the spliceosomes and group II introns, the protein-free snRNAs can perform splicing both in cis and in trans, further strengthening the possible evolutionary link between the two systems. Since the snRNAs are likely descendants of primordial RNA enzymes, these data indicate that despite their highly minimal nature, the vestigial catalytic activity retained in the snRNAs is sufficient to enable them to perform splicing reactions similar to those catalyzed by the extant splicing systems. The extent to which the observed inherent catalytic activity of the snRNAs contributes to spliceosomal catalysis remains an intriguing question. Based on mutational studies and characterization of the catalytic function of protein-free snRNAs performed in this study and previous reports, spliceosomal catalysis and snRNA-catalyzed splicing occur through closely similar reaction pathways and in addition require almost identical snRNA sequences.^{6,10,17,22,23} While these observations strengthen the possibility of RNA catalysis in the spliceosome, unequivocally defining the share of snRNAs in spliceosomal catalysis and dissecting the role of spliceosomal proteins in this process awaits further detailed studies.

Materials and Methods

The E1-int-E2 and E1-int-E2-7 substrates were designed based on the previously-described trans-splicing substrates.^{22,23} Several

substrates were designed which differed in the sequence of the linker region and were screened for lack of stable secondary structures using the RNAstructure algorithm.²⁹ The candidates with the least stable intramolecular structures were selected and PCR-generated templates containing their sequence were transcribed into RNA by in vitro transcription using T7 RNA polymerase as previously described in references 21, 30 and 31. The in vitro-transcribed RNAs were gel purified and a small fraction was end-labeled and run on a 16% denaturing PAGE for quality control. RNase T1 reactions were performed as described in references 21 and 30 to ensure the accuracy of the sequence of the transcribed RNAs. In the E1-int-E2 construct, the first and last exons are 15 and 17 nucleotides long, respectively, with a 35 nucleotide long intervening sequence in between. The snRNA-mediated splicing reactions were performed as described in references 22 and 23, except that instead of the trans-splicing substrates, E1-int-E2 or E1-int-E2-7 were used at concentrations ranging from 2–10 μ M. The reactions were incubated at 35°C for 10–15 hours followed by loading onto denaturing PAGE. Gel fragments containing RNAs approximately 15–25, 25–35, 35–50 and 50–80 were cut from the PAGE and the RNAs in each gel piece were eluted and subjected to RT-PCR reactions using radiolabeled PCR primers as described in reference 22. The result of the PCR reactions were loaded onto a 16% non-denaturing PAGE. A double stranded DNA containing the deduced sequence of the trans-splicing product was labeled and loaded on the PAGE as size marker. Additional size markers that had an identical sequence with one or two nucleotides deleted were loaded on PAGE to determine the resolution of the PAGE. Similar PCR reactions were performed on the E1-int-E2 and E1-int-E2-7 substrates to mark the location of the precursors on the gel.

To determine if under the reaction conditions the E1-int-E2 and E1-int-E2-7 substrates had a different pattern of degradation, end-labeled substrates were incubated under the reaction conditions for 22 hours followed by loading onto a 20% denaturing PAGE. To determine if the E1-int-E2 and E1-int-E2-7 substrates could compete with the trans-splicing substrates for binding and interacting with the U6/U2 complex, we set up typical trans-splicing reactions using end-labeled Exon1 substrates as described in references 22 and 23, followed by the addition of 2–12 μ M unlabeled E1-int-E2 or E1-int-E2-7 substrates. The results of the reactions were analyzed on 16% denaturing PAGE.

Acknowledgements

We thank Tim Nilsen for suggestions and comments on the manuscript. We also thank Aaron Kim, Amanda Black, Young-Min Park, Avanti Golikeri, Justin Pruttivarasin and Miriam Israelowitz for technical assistance. A.M. designed the substrates and performed the initial optimization assays. Y.J. and B.Z. performed the experiments. S.V. designed the experiments, analyzed the data and prepared the manuscript. This work was supported by a Searle Scholar award from Kinship Foundation, an institutional research grant (IRG-91-022) from the American Cancer Society and NIH grant no. GM078572 to S.V.

References

1. Wahl MC, Will CL, Lührmann R. The spliceosome: design principles of a dynamic RNP machine. *Cell* 2009; 136:701-18.
2. Smith DJ, Query CC, Konarska MM. "Nought may endure but mutability": spliceosome dynamics and the regulation of splicing. *Mol Cell* 2008; 30:657-66.
3. Will CL, Lührmann R. Spliceosome structure and function. In: Gesteland R, Cech T, Atkins J, Eds. *The RNA World*. Third edition. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press 2006; 369-400.
4. Pyle AM, Lambowitz AM. Group II Introns: Ribozymes That Splice RNA and Invade DNA. In: Gesteland R, Cech T, Atkins J, Eds. *The RNA World*. Third edition. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press 2006; 469-506.
5. Houglund JL, Piccirilli JA, Forconi M, Lee J, Herschlag D. How the Group I Intron Works: A Case Study of RNA Structure and Function. In: Gesteland R, Cech T, Atkins J, Eds. *The RNA World*. Third edition. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press 2006; 133-205.
6. Valadkhan S. Role of the snRNAs in spliceosomal active site. *RNA Biol* 2010; 7:345-53.
7. Keating KS, Toor N, Perlman PS, Pyle AM. A structural analysis of the group II intron active site and implications for the spliceosome. *RNA* 2010; 16:1-9.
8. Michel F, Costa M, Westhof E. The ribozyme core of group II introns: a structure in want of partners. *Trends Biochem Sci* 2009; 34:189-99.
9. Dayie KT, Padgett RA. A glimpse into the active site of a group II intron and maybe the spliceosome, too. *RNA* 2008; 14:1697-703.
10. Nilsen TW. RNA-RNA interactions in nuclear pre-mRNA splicing. In: Simons R, Grunberg-Manago M, Editors. *RNA Structure and Function*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press 1998; 279-307.
11. Gordon PM, Fong R, Piccirilli JA. A second divalent metal ion in the group II intron reaction center. *Chem Biol* 2007; 14:607-12.
12. Gordon PM, Piccirilli JA. Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. *Nat Struct Biol* 2001; 8:893-8.
13. Sontheimer EJ, Gordon PM, Piccirilli JA. Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. *Genes Dev* 1999; 13:1729-41.
14. Sontheimer EJ, Sun S, Piccirilli JA. Metal ion catalysis during splicing of premessenger RNA. *Nature* 1997; 388:801-5.
15. Gordon PM, Sontheimer EJ, Piccirilli JA. Kinetic characterization of the second step of group II intron splicing: role of metal ions and the cleavage site 2'-OH in catalysis. *Biochemistry* 2000; 39:12939-52.
16. Gordon PM, Sontheimer EJ, Piccirilli JA. Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. *RNA* 2000; 6:199-205.
17. Valadkhan S. The spliceosome: a ribozyme at heart? *Biol Chem* 2007; 388:693-7.
18. Cech TR. The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* 1986; 44:207-10.
19. Sharp PA. Five easy pieces. *Science* 1991; 254:663.
20. Sharp PA. On the origin of RNA splicing and introns. *Cell* 1985; 42:397-400.
21. Valadkhan S, Manley JL. A tertiary interaction detected in a human U2-U6 snRNA complex assembled in vitro resembles a genetically proven interaction in yeast. *RNA* 2000; 6:206-19.
22. Valadkhan S, Mohammadi A, Jaladat Y, Geisler S. Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci USA* 2009; 106:11901-6.
23. Lee C, Jaladat Y, Mohammadi A, Sharifi A, Geisler S, Valadkhan S. Metal binding and substrate positioning by evolutionarily invariant U6 sequences in catalytically active protein-free snRNAs. *RNA* 2010; 16:2226-38.
24. Chu VT, Liu Q, Podar M, Perlman PS, Pyle AM. More than one way to splice an RNA: branching without a bulge and splicing without branching in group II introns. *RNA* 1998; 4:1186-202.
25. Podar M, Chu VT, Pyle AM, Perlman PS. Group II intron splicing in vivo by first-step hydrolysis. *Nature* 1998; 391:915-8.
26. Vogel J, Börner T. Lariat formation and a hydrolytic pathway in plant chloroplast group II intron splicing. *EMBO J* 2002; 21:3794-803.
27. Fedorova O, Zingler N. Group II introns: structure, folding and splicing mechanism. *Biol Chem* 2007; 388:665-78.
28. Anderson K, Moore MJ. Bimolecular exon ligation by the human spliceosome. *Science* 1997; 276:1712-6.
29. Reuter JS, Mathews DH. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* 2010; 11:129.
30. Valadkhan S, Manley JL. Splicing-related catalysis by protein-free snRNAs. *Nature* 2001; 413:701-7.
31. Valadkhan S, Manley JL. Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA* 2003; 9:892-904.

© 2011 Landes Bioscience.
Do not distribute.