

# Protein-free small nuclear RNAs catalyze a two-step splicing reaction

Saba Valadkhan<sup>1,2</sup>, Afshin Mohammadi<sup>1</sup>, Yasaman Jaladat, and Sarah Geisler

Center for RNA Molecular Biology, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106

Edited by Joan A. Steitz, Yale University, New Haven, CT, and approved May 13, 2009 (received for review March 3, 2009)

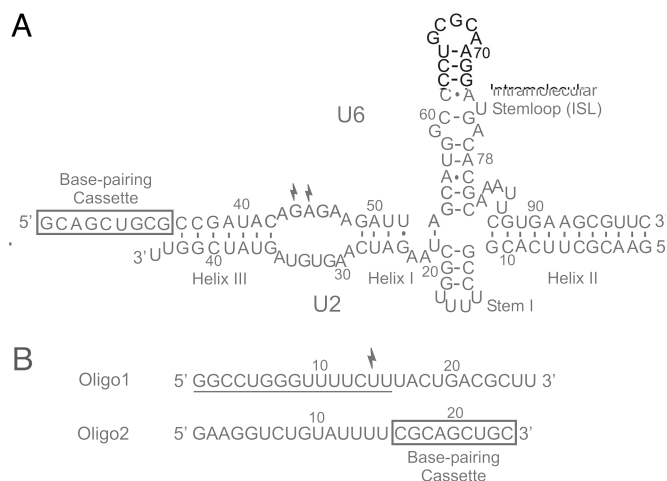
Pre-mRNA splicing is a crucial step in eukaryotic gene expression and is carried out by a highly complex ribonucleoprotein assembly, the spliceosome. Many fundamental aspects of spliceosomal function, including the identity of catalytic domains, remain unknown. We show that a base-paired complex of U6 and U2 small nuclear RNAs, in the absence of the  $\approx 200$  other spliceosomal components, performs a two-step reaction with two short RNA oligonucleotides as substrates that results in the formation of a linear RNA product containing portions of both oligonucleotides. This reaction, which is chemically identical to splicing, is dependent on and occurs in proximity of sequences known to be critical for splicing *in vivo*. These results prove that the complex formed by U6 and U2 RNAs is a ribozyme and can potentially carry out RNA-based catalysis in the spliceosome.

catalysis | ribozymes | snRNAs | spliceosome | U6

Extensive mechanistic and structural similarities between spliceosomal small nuclear RNAs (snRNAs) and self-splicing group II introns (1, 2) have led to the hypothesis that the snRNAs are evolutionary descendents of group II-like introns and thus could similarly have a catalytic role in the spliceosome (3–5). However, because of the extreme complexity of the spliceosome (6–9), a dynamic cellular machine composed of more than 200 different proteins in addition to the snRNAs (U1, U2, U4, U5, and U6), it has not been possible to determine whether the snRNAs can indeed catalyze the splicing reaction without the help of spliceosomal proteins.

In activated spliceosomes, U6 and U2, which are the only snRNAs required for both steps of splicing, form an extensively base-paired complex (7, 10, 11) (Fig. 1A). Considerable data support the functional importance of both the secondary and tertiary structure of the U6/U2 complex (10–13). Furthermore, an evolutionarily invariant region in U6, the ACAGAGA box (Fig. 1A), is in close proximity to the splice sites during splicing catalysis (14–18), and mutagenesis studies have shown that this domain plays a crucial role in catalysis of the splicing reaction (7, 10, 11). Two other conserved regions, the AGC triad and an asymmetric bulge in the intramolecular stemloop of U6 (ISL), are also thought to play important roles in spliceosomal catalysis (11). Recent structural studies indicate that sequences equivalent to these three regions form the active site of group II introns (19). In both systems, the ACAGAGA sequence and its equivalent in group II introns seem to be positioned in close proximity to a conserved asymmetric loop in the middle of the ISL or domain V, the corresponding stemloop in group II introns (19, 20).

Previously we provided evidence that *in vitro*-transcribed, protein-free RNAs containing the conserved central domains of human U6 and U2 could form a base-paired complex *in vitro* that was similar to the complex formed in activated spliceosomes (13) (Fig. 1A). In addition, this *in vitro*-assembled, protein-free U6/U2 complex could catalyze a number of branching-like reactions on a short RNA substrate containing the consensus sequence of the branch site of introns (Oligo1, Fig. 1B), during which an internal functional group in Oligo1 formed a covalent linkage with U6 or a substrate containing the 5' splice site sequence. Although these reactions resembled the first step of splicing in terms of their sequence and ionic requirements, the chemistry of the reactions was either distinct from splicing or the reaction was so inefficient that



**Fig. 1.** The U6/U2 complex and splicing substrates. (A) The U6/U2 construct used in this study, which contains the central domains of the human U6 and U2 snRNAs. The invariant domains of U6 are shaded in gray. Numbers reflect the human numbering system. Thunderbolts point to nucleotides crosslinked to Oligo1. The position of U6/U2 helices I, II, and III is shown. The base-pairing cassette, which provides a binding sequence for Oligo2, is marked by a rectangle. (B) The substrates. Thunderbolt indicates the location of crosslinking probe. The numbers reflect position from the 5' end. The minimal required nucleotides in Oligo1 are underlined. Nucleotides 16–22 of Oligo1 comprise the branch site consensus sequence. Oligo2 nucleotides marked by a rectangle can base pair to a complementary sequence at the 5' end of the U6 construct.

it could not be chemically characterized (21–23). Thus, the role of the spliceosomal RNAs in catalysis of the chemistry of the splicing reaction remains uncertain. The recent discovery of an RNase H motif in Prp8 (24), a constituent of the spliceosomal active site, further underscores the importance of a thorough understanding of the role of the snRNAs in splicing and the development of a simple model system for biochemical analysis of the role of RNA and protein components of the spliceosomal active site.

Here we show that, by altering the sequence of the RNA oligonucleotide substrates, we have enabled the base-paired complex of U6 and U2 snRNAs to catalyze a two-step reaction on two short RNA substrates that leads to the cleavage and subsequent ligation of fragments of the two RNA substrates into a linear product, similar to fully spliced mRNAs. This reaction, which does not require any of the  $\approx 200$  other spliceosomal factors, resembles the splicing reaction catalyzed by group II introns and the spliceosome, not only in its chemistry and ionic requirements but also

Author contributions: S.V. designed research; S.V., A.M., Y.J., and S.G. performed research; S.V. analyzed data; and S.V. wrote the paper.

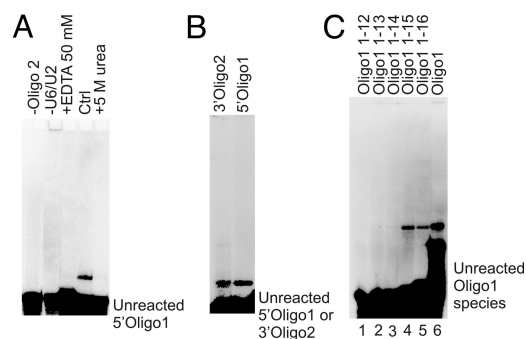
The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

S.V. and A.M. contributed equally to this work.

To whom correspondence should be addressed. E-mail: saba.valadkhan@case.edu.

This article contains supporting information online at [www.pnas.org/cgi/content/full/090202106/DCSupplemental](http://www.pnas.org/cgi/content/full/090202106/DCSupplemental).



**Fig. 2.** Formation of the product. (A) Requirements for product formation. Lane labeled Ctrl contains a typical reaction. Lanes marked –Oligo2 and –U6/U2 lack Oligo2 or U6/U2, respectively. Lanes marked + 5 M urea or +EDTA 50 mM contain typical reactions to which the indicated chemicals have been added. (B) Product formation with end-labeled substrates. The identity of the labeled substrate in each reaction is shown above each lane. (C) The effect of Oligo1 truncation mutations. The Oligo1 fragment used is shown above each lane.

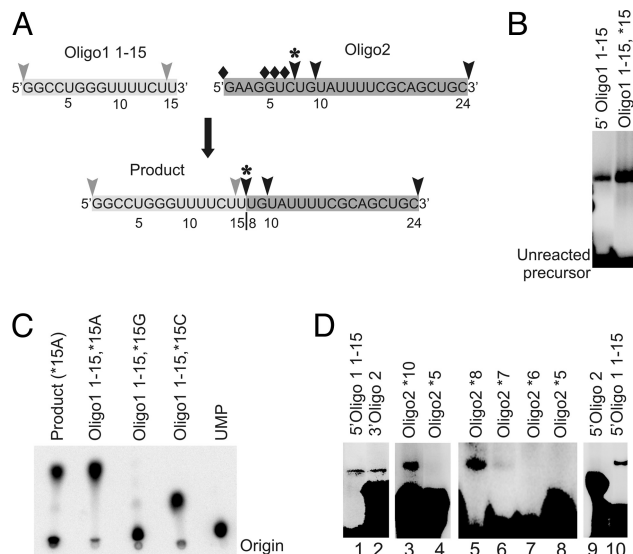
because mutational and complementation analyses indicate that U6 sequences critical for splicing in the spliceosome and group II introns are also required for this reaction. Furthermore, the reaction occurs in the vicinity of nucleotides known to be close to the 5' splice site at the time of catalysis of the first step of splicing in the spliceosome. Thus, it recapitulates several aspects of spliceosomal catalysis.

## Results

**Formation of a New U6/U2 Catalyzed Product.** To investigate the ability of spliceosomal snRNAs to catalyze the chemistry of the splicing reaction, we added a short stretch of nucleotides to the 3' end of one of our RNA oligonucleotides (Oligo2, Fig. 1B). The added nucleotides could form base-pairing interactions with a complementary stretch of nucleotides added to the 5' end of U6 (Fig. 1A) in a way that would orient the rest of Oligo2 toward the catalytically crucial ACAGAGA box of U6. We reasoned that, by manipulating the sequence of this RNA, we might enable it to interact productively with the U6/U2 complex, leading to a U6/U2 catalyzed reaction between this oligonucleotide and Oligo1 that more closely resembled the splicing reaction. We designed a number of Oligo2 substrates with slight variations in their degree of complementarity to ACAGAGA and also ensured that the substrates did not contain stable intramolecular structures (see *SI Text*, Supporting note 1). We next incubated the *in vitro*-assembled U6/U2 complex with 5'-radiolabeled Oligo1 and various unlabeled Oligo2 species under several cationic and pH conditions and looked for the formation of new RNA species that could result from a splicing-like reaction.

Interestingly, we observed a new RNA species after 6 hours of incubation of one set of our substrates with the U6/U2 complex that formed at an efficiency of  $\approx 2\%$  compared to the radiolabeled Oligo1 input (Fig. 2A). Omitting either the U6/U2 complex or Oligo2 resulted in complete inhibition of the formation of this RNA product (Fig. 2A). Similar to reactions with 5'-labeled Oligo1, reactions containing 3'-labeled Oligo2 resulted in incorporation of the radiolabel into the product, indicating the presence of the 5' end of Oligo1 and 3' end of Oligo2 in the product (Fig. 2B). Thus, the product resulted from a reaction between the two oligonucleotides. Product formation was dependent on the presence of  $Mg^{2+}$  in the buffer, and cobalt hexammine or molar concentrations of monovalent salts could not replace  $Mg^{2+}$  (Fig. 2A and Fig. S1A) (1, 25, 26).

Eliminating the base-pairing potential between U6 and Oligo2 by mutating either the U6 or Oligo2 side of the base-pairing cassette (Fig. 1A and B) blocked product formation, whereas restoring the base-pairing by complementary mutations resulted in complete restoration of activity (Fig. S1B), indicating that this interaction is



**Fig. 3.** Defining the constituents of the product. Numbers after an asterisk indicate the position of the site-specific radiolabel. (A) Summary of site-specific labeling experiments. The sites of radiolabels that were (arrows) or were not (diamonds) incorporated into the product are shown. Numbers indicate position from the 5' end of each substrate. The 5'-most site-specific label in Oligo2 that is incorporated into the product is marked by an asterisk. (B) The product contains the entire Oligo1 1–15. The identity of site-specifically labeled Oligo1 used in each reaction is shown above each lane. (C) Thin layer chromatography (TLC) analysis of product made with Oligo1 species containing a radiolabel 5' to nt 15, and an U to A mutation at this position, Product (\*15A). The purified product is subjected to complete Nuclease P1 digestion, followed by loading onto the TLC. The Oligo1 precursors containing a radiolabel at position 15 and an A, C, or G mutation at this position were subjected to Nuclease P1 digestion along with the product to mark the location of each NMP. (D) Product formation using Oligo2 site-specifically labeled at internal positions or the 5' or 3' ends. Reactions with 5' Oligo1 1–15 are included as controls.

required for product formation. To our surprise, a 3' truncation mutant of Oligo1 that lacked the branch site consensus sequence was completely active in product formation (Fig. 2C, lane 5). This indicated that the interaction of the branch site consensus sequence with U2 was not required for the reaction, and that the product likely did not result from a branching reaction. Further deletion mutagenesis indicated that the first 15 nucleotides of Oligo1 constituted the minimal sequences required for the reaction (Fig. 2C, also see Fig. 1B).

**Characterization of the Reaction.** We characterized the observed reaction using the minimally-required Oligo1 sequences. Product formation with Oligo1 1–15 was identical to the reactions with full-length Oligo1 in its requirement for U6/U2 and Oligo2 and its cationic requirements (Fig. S2A and B). The product formed with the same efficiency when the reaction was incubated in the dark, indicating that it was not a photo-adduct (Fig. S2A). The product did not form upon incubation at 4 °C (Fig. S2C). Replacing either of the substrates with DNA oligonucleotides of the same sequence blocked product formation (Fig. S2D). Purified product was resistant to denaturation, indicating that it did not result from a hyperstable noncovalent interaction between the substrates (Fig. S2E). The product accumulated to  $\approx 2\%$  of the input Oligo1 1–15 after 6–8 hours of incubation in the presence of 60 mM  $MgCl_2$  at 45 °C (Fig. S2F). Product formation was dependent on the concentration of the U6/U2 complex and the two substrates, and showed a plateau at higher Oligo2 and U6/U2 concentrations, consistent with a catalyzed reaction (Fig. S2G).

To determine the composition of the product, we used site-specifically labeled substrates in the reaction. The results of these experiments are summarized in Fig. 3A. Radiolabels placed at the 5' end or the 3'-most phosphate in Oligo1 appeared in the product

(Fig. 3B). We formed the product using Oligo1 1–15 species that contained a single radioactive phosphate 5' to the nucleotide (nt) at position 15, and a C, A, or G nucleobase at this position, followed by RNase P1 digestion of the purified product. The results indicated that in addition to the phosphate located 5' to the last nucleotide in Oligo1, the sugar and nucleobase of this residue were also present in the product (Fig. 3C, Fig. S3A; see also *SI Text*, Supporting Note 2). Taken together with sequencing data (see below), the above results indicate that Oligo1 1–15 is incorporated into the product in its entirety.

Interestingly, whereas labels located 5' to nt 8 and downstream of this residue in Oligo2 were incorporated into the product, radioactive labels at the 5' end, or on phosphates 5' to nts 5, 6 and 7 did not appear in the product (Fig. 3D; compare lanes 2, 3, and 5 with lanes 6, 7, 8, and 9; see also *SI Text*, Supporting Note 3). The use of Oligo2 species that had 17 or 40 nts added to their 5' or 3' ends, respectively, further confirmed that nt 7 and nucleotides upstream of it were removed during product formation, whereas nts added to the 3' end of Oligo2 were incorporated into the product (Fig. S3 B and C).

**Formation of a Linear Product via Transesterification.** Taken together, the above results suggested that the product contained full-length Oligo1 1–15 and nts 8–24 of Oligo2 (Fig. 3A). To determine how these two RNA fragments were joined together in the product, we formed the product using 5'-labeled Oligo1 1–15 that had a U to G mutation at nt 15, and performed limited RNase T1, RNase T2, or alkaline hydrolysis reactions on the purified product. The RNase T1 fingerprint and uniform ladders resulting from limited RNase T2 and alkaline hydrolysis (Fig. 4 A and B) clearly indicated that the product was a linear RNA molecule in which nt 15 at the 3' end of Oligo1 1–15 was joined to U8 of an Oligo2 fragment containing nts 8–24 (Fig. 3A). Reverse transcription polymerase chain reaction (RT-PCR) reactions on purified product further confirmed this conclusion (Fig. S3D). The absence of a gap in the RNase T2 digestion ladder also suggested that the linkage between nt 15 of Oligo1 1–15 and nt 8 of Oligo2 is a 3'-5' phosphodiester linkage, as RNase T2 is highly specific for 3'-5' RNA linkages (27). To confirm this result, we performed the reaction using an Oligo2 that was site-specifically labeled at the phosphate 5' to nt 8, followed by RNase P1 digestion. Because RNase P1 is also highly specific for 3'-5' phosphodiester linkages (28), the absence of a nuclease-resistant species on TLC plates indicated that the linkage between nt 15 of Oligo1 1–15 and nt 8 of Oligo2 is a canonical 3'-5' phosphodiester linkage (Fig. 4C and *SI Text*, Supporting Note 4). Importantly, we were able to demonstrate the transfer of a radiolabeled phosphate from 5' end of nt 8 in Oligo2 to the 3' end of nt 15 of Oligo1 by RNase T1 digestion assays (Fig. 4D and Fig. S4), further indicating that product formation involves a transesterification reaction.

Taken together, the above data suggest that during product formation, the 3' OH of the last nucleotide of Oligo1 1–15 forms a 3'-5' linkage with the phosphate 5' to nt 8 in Oligo2, and that the first seven nucleotides of Oligo2 are removed during the reaction. The most plausible reaction mechanism for the formation of the observed product entails a nucleophilic attack by 3' OH of U15 of Oligo1 1–15 on the phosphate 5' to U8 in Oligo2, similar to the mechanism of the second step of splicing in the spliceosome and group II introns (1, 2, 7, 11). Results of 2' deoxy, 3' deoxy or dideoxy substitution of U15 in Oligo1 1–15 suggested that product formation requires a 3' OH at nt 15 of Oligo1 1–15, whereas the loss of 2' OH alone had no observable effect (Fig. 4E, Fig. S5). Taken together, the above results suggest that human U6 and U2 snRNAs can catalyze a transesterification reaction between two short RNA oligonucleotides leading to the formation of a 3'-5' linkage.

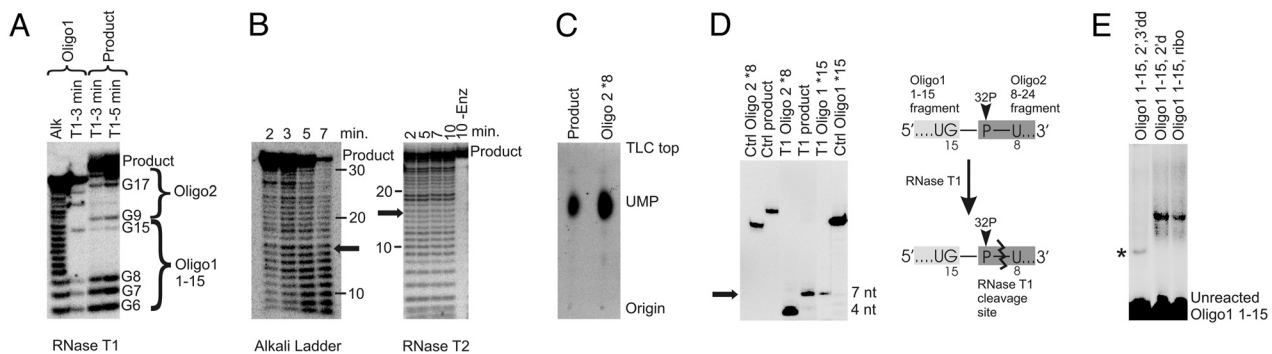
**Catalysis of Two Consecutive Splicing-Like Reactions.** As mentioned above, the original Oligo1 species, which contained 11 additional nucleotides at its 3' end compared to the minimal Oligo1 1–15

substrate, could form a product with mobility and ionic and sequence requirements identical to the product formed with Oligo1 1–15. The use of Oligo1 species containing 2 or 5 extra nucleotides at the 3' end compared with Oligo1 1–15 indicated that the number of nucleotides downstream of nt 15 of Oligo1 did not affect the formation or mobility of the product (Fig. S6A). RT-PCR sequencing of the product formed with full-length Oligo1 proved that it indeed had the same sequence as the product formed with Oligo1 1–15 (Fig. 5A). Thus, the extra 11 nucleotides at the 3' end of Oligo1 must necessarily be removed during product formation. Indeed, site specific labeling experiments on full-length Oligo1 confirmed the removal of nucleotides 3' to nt 15 (Fig. 5B and *SI Text*, Supporting Note 5) and the formation of a 3'-5' phosphodiester linkage between the 3' OH of nt 15 of Oligo1 and nt 8 of Oligo2 (Fig. 5C and *SI Text*, Supporting Note 4).

A one-step reaction mechanism that would lead to the formation of a linkage between the 3' OH of nt 15 of Oligo1 and the phosphate 5' to nt 8 of Oligo2 would not be plausible, as both are engaged in stable phosphodiester linkages. Thus, cleavage of Oligo1 and Oligo2 and ligation of the two fragments should involve at least two steps. Either Oligo1 or Oligo2 can be cleaved first, generating 3' OH and 5'-phosphate ends, respectively, followed by a reaction that couples the cleavage of the second substrate to ligation of the two fragments. Alternatively two cleavage events result in generation of the two fragments, followed by a third ligation step. Incubation of Oligo2 8–24 fragment containing a monophosphate at the 5' end with full-length Oligo1 or Oligo1 1–15 did not result in product formation (Fig. S6B) as expected. Because both these reactions require expenditure of energy to activate the 5' phosphate, they are unlikely to occur under the reaction conditions. The only remaining possibility would be cleavage of Oligo1 nucleotides 3' to nt 15 as the first event, generating an Oligo1 1–15 with a free 3' OH, which would then proceed to carry out the transesterification reaction described above. Consistent with this mechanism, reactions containing full-length Oligo1 and Oligo1 1–15 not only form an identical product, but also have similar reaction kinetics and sequence and ionic requirements. Thus it is possible that the reaction with Oligo1 1–15 is the rate-limiting step for the reaction involving full-length Oligo1.

To gain further insight into the mechanism of removal of the residues 3' to nt 15 in Oligo1, we reasoned that if the removal of the nucleotides at the 3' end of Oligo1 was the result of a nucleophilic attack by another functional group in U6/U2 or elsewhere in the substrates, for example in the intronic part of Oligo2, the released fragment of Oligo1 should be attached to that RNA, resulting in a slower-mobility RNA species. We could not detect such species in experiments containing an Oligo1 that was labeled on the released fragment (Fig. S6C), which indicated that either the removal of the 3' fragment of Oligo1 did not involve an attack by another RNA, or it involved a short-lived covalent intermediate. Also, an Oligo1 species that contained a 2'-deoxy substitution at nt 15 (Oligo1 15d) was fully active in product formation, indicating that the mechanism of cleavage of the 3' fragment of Oligo1 cannot be an attack by the adjacent 2' OH (Fig. 5D). Furthermore, a cleavage mechanism involving the adjacent 2' OH and cyclic end formation would result in incorporation of the site-specific radiolabel at position 16 into the product, which was not observed (Fig. 5B and *SI Text*, Supporting Note 6). Thus, removal of the 3' fragment of Oligo1 is likely to occur by a hydrolytic mechanism (Fig. 5E).

To determine whether the hydrolytic cleavage at nt 15 is a spontaneous or catalyzed reaction, we reasoned that uncatalyzed hydrolytic cleavage at nt 15 of Oligo1 15d, which cannot form 2' OH-mediated cleavage fragments at position 15, would lead to accumulation of fragments resembling Oligo1 1–15 2'd, which should react with the same efficiency as observed in reactions containing Oligo1 1–15 (Fig. 4E). Thus, there should be a  $\approx 50$ -fold excess of cleaved, unreacted Oligo1 1–15-like fragments compared with the formed product, which was not observed (Fig. S7A), ruling

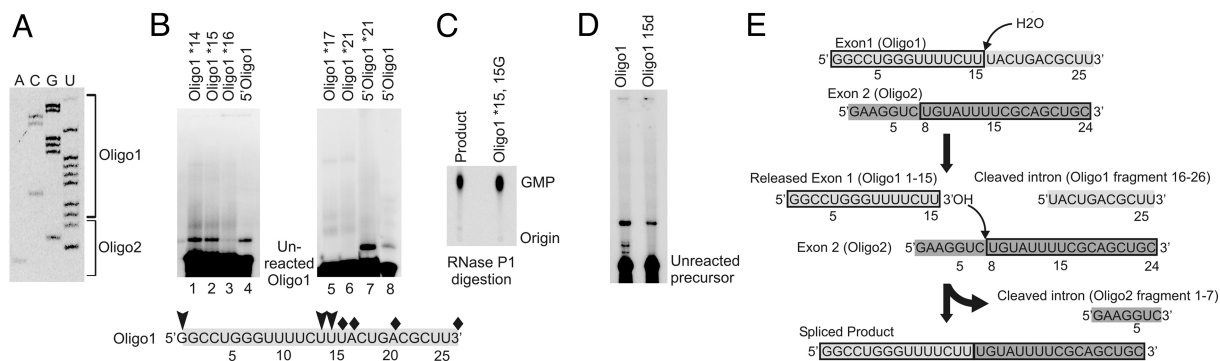


**Fig. 4.** Characterization of the linkage between Oligo1 1–15 and Oligo2 fragment in the product. Min, minutes of RNase T1, T2, or alkaline hydrolysis reaction. (A) Limited RNase T1 digestion. Alk, limited alkaline hydrolysis ladder. The duration of RNase T1 treatment is shown above each lane. Position of the undigested product, and nucleotides in the product that originate from Oligo1 1–15 or Oligo2 are shown to the right. Full-length Oligo1 is used in T1 and alkaline hydrolysis reactions as control. (B) Limited alkaline hydrolysis and RNase T2 digestion of the product. Numbers to the side of panels indicate fragment size in nucleotides. –Enz, no enzyme added. Numbers above each lane indicate the duration of alkaline or RNase T2 treatment. The position of undigested product is shown to the right. Arrows point to the junction between Oligo1 1–15 and Oligo2-derived sequences in the product. (C) TLC analysis of complete RNase P1 digestion of product formed with Oligo2 labeled 5' to nt 8 (Oligo2\*8). The precursor Oligo2\*8 is similarly digested to mark the position of UMP on the TLC plate. (D) Nucleotide 15 of Oligo1 1–15 is linked to nt 8 of Oligo2 through a phosphodiester bond. Product formed with Oligo2\*8 is subjected to complete RNase T1 digestion along with Oligo2\*8 and Oligo1\*15. Arrow points to the site of the labeled fragment released from the product after RNase T1 digestion, which contains nts 9–15 of Oligo1 (Fig. S4). Length of the released fragments is shown to the right. Ctrl, undigested product or precursor. Numbers following asterisk indicate the position of the radiolabel. Diagram shows the junction between Oligo1- (light gray) and Oligo2- (dark gray) derived sequences in the product. Serrated line marks the site of RNase T1 cleavage, which results in transfer of the label from nt 8 of Oligo2 to nt 15 of Oligo1. The position of the radiolabel is shown by an arrowhead. (E) Product formation with 2' deoxy- (2'd) and 2', 3' dideoxy-substituted (2',3'dd) Oligo1 1–15. The identity of the Oligo1 species used in each lane is shown on top. The band marked by an asterisk is a preparation artifact.

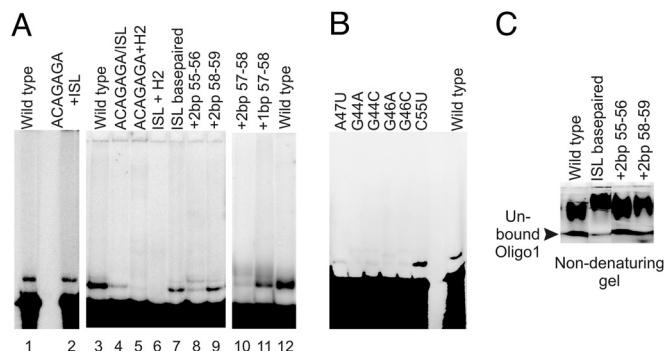
out the possibility of an uncatalyzed hydrolysis reaction at nt 15 of Oligo1. Furthermore, preincubation of reactions that were missing one or the other substrate, which would result in accumulation of randomly cleaved fragments, did not result in faster reaction kinetics once the missing ingredient was added (Fig. S7B and *SI Text*, Supporting Note 7). Taken together, these results suggest that the mechanism of cleavage of Oligo1 was a catalyzed attack by a water/hydroxyl group from the solvent.

**Relevance of U6/U2-Mediated Protein Free Splicing to Spliceosomal Catalysis.** Our data suggest that product formation is essentially a splicing reaction: an “intronic” fragment is removed from Oligo1 (Exon1), generating a free 3'OH, which then attacks the “3' splice site” on Oligo2 (Exon2), resulting in removal of an intronic fragment and a transesterification reaction joining the exonic fragments of Exon1 and Exon2 (Fig. 5E). As mentioned above, our analyses indicate that the formation of the product is a catalyzed

reaction, and is also U6/U2-dependent. We defined the sequences in U6 and U2 that were essential for product formation via mutagenesis (Fig. S8A). Even complete removal of the U2 stem I and the sequences in U6 and U2 that form helix II did not affect the reaction (Fig. 6A, lane 2). In contrast, removal of the intramolecular stem loop (ISL) or the invariant ACAGAGA sequence and its adjoining helices resulted in complete loss of activity (Fig. 6A, lanes 5 and 6). Although a construct that only contained the ACAGAGA-containing stem and the ISL linked together through a flexible junction (construct labeled ACAGAGA+ISL) was fully active, increasing the rigidity of the junction to induce coaxial stacking of these two elements (construct labeled ACAGAGA/ISL) substantially reduced activity (Fig. 6A; compare lanes 1, 2, 3 and 4; Fig. S8A). Several point mutations in the ACAGAGA sequence almost completely blocked product formation (Fig. 6B, Fig. S8B). Point mutations in the G of the AGC triad were similarly incompatible with product formation, while mutation of the C



**Fig. 5.** Characterization of product formation with full-length Oligo1. (A) Dideoxy sequencing reactions on purified product. The nucleotides originating from Oligo1 and Oligo2 are shown. The identity of nucleotides labeled in each sequencing lane is shown on top. (B) Product formation with site-specifically labeled full-length Oligo1. Numbers following an asterisk indicate the position of the site-specific label. 5'-Labeled Oligo1 is used in reactions as control. 5'Oligo1\*21 carries two labels. Diagram shows the site of labels that do (arrowheads) or do not incorporate (diamonds) into the product. Numbers refer to position from the 5' end. (C) TLC analysis of complete RNase P1 digestion of product made with Oligo1 U15G mutant labeled 5' to nt 15. The precursor (Oligo1\*15, 15G) is similarly digested to mark the location of GMP. (D) Product formation with Oligo1 containing a deoxy substitution at position 15 (Oligo1 15d). The lane marked Oligo1 contains a typical reaction. (E) Product formation. Substrates, product, and fragments cleaved after each step of the reaction are shown. Numbers indicate position from the 5' end of each substrate. Rectangles mark the nucleotides in each substrate that are incorporated into the product. The site of hydrolytic cleavage of Oligo1 is shown.

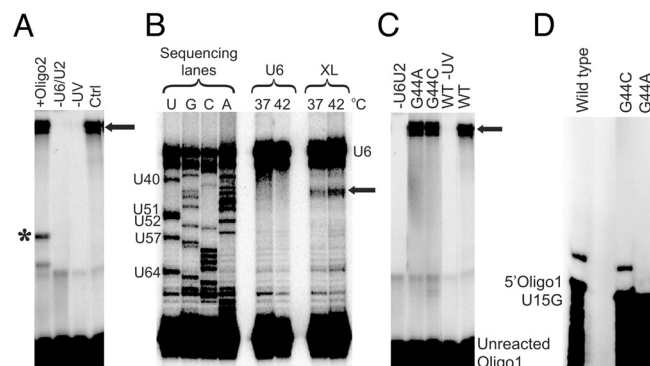


**Fig. 6.** Role of the U6/U2 complex in product formation. (A) The effect of deletion and insertion mutations on product formation. The designations +1 and +2 bp reflect the number of inserted base pairs, followed by the U6 nucleotides between which the added base pairs are inserted. For the description of each mutation, see text and Fig. S8A. (B) Effect of ACAGAGA point mutations. (C) ISL insertion mutants can bind Oligo1. The location of unbound Oligo1 is shown. The identity of the U6/U2 used is shown above each lane.

residue of this triad had no effect (Fig. 6B, Fig. S8B). Converting the ISL to a fully base-paired helix (construct labeled ISL base-paired) considerably reduced activity, whereas substrate binding was not decreased (Figs. S8A, S6A, lane 7, and Fig. S6C), indicating a functional role for this region of ISL in product formation. To determine whether the spatial orientation of the ISL toward the rest of U6 is important for product formation, we made a series of 1- or 2-bp insertions in the lower stem of the ISL (Fig. S8A). All four insertion mutations substantially reduced product formation without affecting substrate binding (Fig. 6A, lanes 8, 9, 10, 11; Fig. 6C). Taken together, these results indicate that the ISL and the ACAGAGA-containing stem contain a number of critical sequence elements that are necessary and sufficient for the observed reaction to occur. The importance of spatial orientation and flexibility between the ISL and the ACAGAGA-containing stem suggests that the reaction requires the formation of a certain three-dimensional arrangement between these elements.

To determine the sequences in the U6/U2 complex that are located close to the site of the reaction and thus might play a direct role in it, we performed site-specific UV crosslinking with an Exon1 containing a 4-thio-U substitution at nt 14, which is one nucleotide upstream of the reaction site. Remarkably, primer extension reactions on the obtained crosslinked species indicated that nt 14 of Exon1 was positioned in close proximity to U6 nts 44 and 45 in the ACAGAGA box (Fig. 1, Fig. 7A, B). Furthermore, a faster migrating crosslinked species that was dependent on the presence of Exon2 and its ability to base pair to U6/U2 was also observed (Fig. 7A), indicating juxtaposition of Exon1 nt 14 to Exon2. Because the base-pairing interaction between Exon2 and U6 orients Exon2 toward the ACAGAGA sequence (Fig. 1), these results suggest that at least in a fraction of the U6/U2 complexes, both Exon1 and Exon2 are positioned in the vicinity of the catalytically critical ACAGAGA domain of U6.

Mutations of G44 in ACAGAGA to A and C, both of which result in loss of product formation (Fig. 6B), did not alter the binding of Exon1, as the formation of the crosslink between Exon1 nt 14 and U6/U2 was not affected (Fig. 7C). Because the binding of Exon2 through its base-pairing cassette to U6 is not likely to be affected by ACAGAGA mutations either, this result suggests that the detrimental effect of at least some of the point mutations in the ACAGAGA box is not caused by a gross defect in substrate positioning. Importantly, the catalytic defect of the G44C mutant could be suppressed by a U to G mutation at nt 15 of Exon1 (Fig. 7D), whereas the activity of the G44A mutant is not rescued, reflecting allele specificity. These results suggest that, rather than affecting ground state substrate binding, the lack of activity of G44C



**Fig. 7.** Defining the U6/U2 sequences that are close to the site of the reaction. (A) Formation of a U6/U2-dependent, UV-dependent crosslink between U6 and nt 14 of Oligo1. Arrow points to the crosslinked product. Plus and minus signs indicate the added or omitted ingredient compared to a typical crosslinking reaction (Ctrl). Asterisk marks an Oligo2-dependent crosslinked species. (B) Primer extension reactions map the position of the crosslink in U6 to G44 and A45 in the ACAGAGA box. Arrow points to primer extension stops in the lanes containing the crosslinked species (lanes marked XL). Control reactions on non-crosslinked U6 are included (lanes marked U6 on the top). Location of the DNA species resulting from primer extension on intact U6 is shown to the right. The primer extension reaction temperature is indicated above each lane. The identity of each band in the left-most sequencing lane is shown to the left. (C) Point mutations in the ACAGAGA domain of U6 do not prevent the formation of a crosslink between Oligo1 and U6. Arrow points to the site of crosslink. WT: wild type. Lane marked WT-UV has not received UV irradiation. Lane marked -U6/U2 lacks the U6/U2 complex. Oligo1 used in these reactions contained a 4-thio-U substitution at nt 14. (D) Suppression of a mutation in the ACAGAGA box of U6 by a point mutation in nt 15 of Oligo1. Product formation is restored in U6 G44 to C mutant when nt 15 of Oligo1 is also mutated to a G. The identity of U6/U2 species used is shown above each lane.

mutant results from the loss of a functionally critical interaction with the nucleotide immediately upstream of the cleavage site. Taken together, the above data suggest that the reaction occurs in the proximity of the ACAGAGA sequence of U6, and that some residues in this sequence likely play a critical role in sculpting the active site of the observed reaction.

## Discussion

We have shown that protein-free RNAs containing the unaltered sequence of the central domains of human U6 and U2 snRNAs, which correspond to the minimal domains required for spliceosomal function, can catalyze a splicing reaction on two short RNA substrates. This reaction is chemically identical to the second step of the splicing reaction catalyzed by Group II introns and the spliceosome. The first step of splicing in this reaction apparently proceeds through hydrolysis rather than branching; however, catalysis of the first step of splicing through hydrolysis is widely observed in group II introns as a physiological alternative to branching both *in vitro* and *in vivo* (29, 30). Indeed, under the conditions used in our catalytic assays, hydrolysis is the dominant or even the sole reaction pathway for a number of well-studied introns that can efficiently perform branching in the presence of higher salt concentrations (1, 2). Recently, it has been shown that the spliceosomal active site can also catalyze hydrolytic reactions at splice sites by cleavage of the 2nd exon from lariat intermediates or spliced mRNAs (31).

Our data suggest that the function of the snRNAs in the observed reaction closely parallels their function in the spliceosomal active site. First, mutagenesis data show that the regions in U6 that are required for spliceosomal catalysis, the ACAGAGA and AGC sequences and the ISL, are also necessary and sufficient to promote the protein-free splicing activity of U6/U2 (7, 10, 11, 14–16, 20, 32–34). Equivalent sequences form the active site of self-splicing group II introns (19). Second, both the snRNA-catalyzed reac-

tion and spliceosomal splicing occur in close proximity of the ACAGAGA sequence (17, 18), and in both systems this sequence seems to help sculpt the active site by interacting with nucleotides adjacent to the site of catalysis (11). Third, footprinting experiments indicate that the ISL is in the vicinity of the ACAGAGA sequence in activated spliceosomes (20), and similarities to group II introns suggest that the active site of the spliceosome is at least partly formed through juxtaposition of these two elements. Sensitivity of the snRNA-catalyzed reaction to changes in spatial orientation of the ISL relative to ACAGAGA and AGC and the detrimental effect of mutations that convert ISL to a fully base-paired helix likely reflect a similar three-dimensional organization.

The observed snRNA-catalyzed splicing reaction was specific; no other product was formed under the various ionic and pH conditions and RNA concentrations or with any of the different RNA species tested, except a previously described, related product forming with much lower efficiency at lower pH (21). The efficiency of the observed splicing reaction approaches the self-splicing efficiency of several less efficient group I and group II introns (35–37). Further, similar to the spliceosomal catalysis (25, 26), the reaction showed a clear requirement for divalent cations. The clear relevance of the chemistry of the reaction to the reaction occurring in the authentic spliceosome, the arrangement and identity of the snRNA sequences involved in the reaction, and the requirement for divalent cations strongly suggest that the observed reaction might originate from a protein-free version of the spliceosomal active site.

Demonstration of the ability of unmodified spliceosomal RNAs to catalyze a full splicing reaction suggests the presence of another naturally occurring RNA enzyme in eukaryotes, and further strengthens the possibility of a common origin between the eukaryotic spliceosome and group II introns. Interestingly, the other natural ribozymes that catalyze complex, multistep reactions such as splicing are often much larger in size (e.g.,  $\approx 400$  nts for a minimal group II intron that can still perform a two-step splicing reaction vs.  $\approx 120$  nts for U6/U2) (1), making the human U6/U2 complex the smallest known natural ribozyme capable of performing splicing.

At least one spliceosomal protein, Prp8, is present in the spliceosomal active site along with snRNAs, playing an important role in substrate recognition and positioning and stabilization of snRNA structures (38). Many catalytic RNAs exist as ribonucleoprotein complexes *in vivo*; however, in studied cases, the presence of proteins does not seem to alter the reaction mechanism, which remains RNA-catalyzed even in the presence of the associated proteins, albeit with much higher efficiency (39). Although the data presented here indicate that snRNAs can perform splicing without the help of proteins, it is likely that in the spliceosomal active site Prp8 or another protein assists the snRNAs in assuming the catalytically active conformation, positioning of the substrates, coordinating catalytic metal ions or even performing chemistry (24). In addition to demonstrating the ability of snRNAs to catalyze splicing, the reaction described above provides the basis for development of an *in vitro* system that would allow the function of RNA and protein elements found in the active site of the spliceosome, including Prp8, to be biochemically analyzed in detail.

## Materials and Methods

*In vitro* transcribed central domains of human U6 and U2 snRNAs were annealed together in 40 mM Tris pH 7.2 and 20 mM MgCl<sub>2</sub> by heating to 75 °C followed by gradual cooling to room temperature. After annealing, the Oligo1 and Oligo2 substrates were added at concentrations ranging from 50 nM to 10  $\mu$ M and MgCl<sub>2</sub> concentration and pH was adjusted to 60 mM and 8.0–8.4, respectively. The final concentration of the U6/U2 complex in a typical reaction was 2  $\mu$ M. The reaction mixtures were incubated at 45 °C for 6 hours, 37 °C for 15 hours, or 32 °C for 40 hours, followed by analysis on 12–20% denaturing polyacrylamide gel electrophoresis (PAGE). The gels were exposed to PhosphorImager screens. Further experimental details are provided in the online [SI Text](#).

**ACKNOWLEDGMENTS.** The authors thank John Burke, Rick Padgett, Tim Nilsen, Mike Harris, Pieter deHaseth, Mark Caprara, Jim Bruzik, Jonatha Gott, and Jo Ann Wise for suggestions, insights and comments on the manuscript, and Catherine Tran for technical assistance. This work was supported by a Searle Scholar award from Kinship Foundation, by an institutional research grant (IRG-91-022) from the American Cancer Society, and by a National Institutes of Health grant (GM078572) (to S.V.).

- Lehmann K, Schmidt U (2003) Group II introns: Structure and catalytic versatility of large natural ribozymes. *Crit Rev Biochem Mol Biol* 38:249–303.
- Pyle AM, Lambowitz AM (2006) Group II introns: Ribozymes that splice RNA and invade DNA. *The RNA World*, eds Gesteland RF, Cech TR, Atkins JF (Cold Spring Harbor Laboratory Press, Cold Spring Harbor), pp 469–506.
- Sharp PA (1985) On the origin of RNA splicing and introns. *Cell* 42:397–400.
- Cech TR (1986) The generality of self-splicing RNA: Relationship to nuclear mRNA splicing. *Cell* 44:207–210.
- Collins CA, Guthrie C (2000) The question remains: Is the spliceosome a ribozyme? *Nat Struct Biol* 7:850–854.
- Nilsen TW (2003) The spliceosome: The most complex macromolecular machine in the cell? *Bioessays* 25:1147–1149.
- Will CL, Luhrmann R (2006) Spliceosome structure and function. *The RNA World*, eds Gesteland RF, Cech TR, Atkins JF (Cold Spring Harbor Laboratory Press, Cold Spring Harbor), pp 369–400.
- Staley JP, Guthrie C (1998) Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell* 92:315–326.
- Valadkhan S (2007) The spliceosome: Caught in a web of shifting interactions. *Curr Opin Struct Biol* 17:310–315.
- Valadkhan S (2007) The spliceosome: A ribozyme at heart? *Biol Chem* 388:693–697.
- Nilsen TW (1998) RNA–RNA interactions in nuclear pre-mRNA splicing. *RNA Structure and Function*, eds Simons RW, Grunberg-Manago M (Cold Spring Harbor Laboratory Press, Cold Spring Harbor), pp 279–307.
- Madhani HD, Guthrie C (1994) Randomization-selection analysis of snRNAs *in vivo*: Evidence for a tertiary interaction in the spliceosome. *Genes Dev* 8:1071–1086.
- Valadkhan S, Manley JL (2000) A tertiary interaction detected in a human U2–U6 snRNA complex assembled *in vitro* resembles a genetically proven interaction in yeast. *RNA* 6:206–219.
- Kandels-Lewis S, Seraphin B (1993) Involvement of U6 snRNA in 5' splice site selection. *Science* 262:2035–2039.
- Sawa H, Abelson J (1992) Evidence for a base-pairing interaction between U6 small nuclear RNA and 5' splice site during the splicing reaction in yeast. *Proc Natl Acad Sci USA* 89:11269–11273.
- Sawa H, Shimura Y (1992) Association of U6 snRNA with the 5'-splice site region of pre-mRNA in the spliceosome. *Genes Dev* 6:244–254.
- Konarska MM, Vilardell J, Query CC (2006) Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol Cell* 21:543–553.
- Sontheimer EJ, Steitz JA (1993) The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* 262:1989–1996.
- Toor N, Keating KS, Taylor SD, Pyle AM (2008) Crystal structure of a self-spliced group II intron. *Science* 320:77–82.
- Rhode BM, Hartmuth K, Westhof E, Luhrmann R (2006) Proximity of conserved U6 and U2 snRNA elements to the 5' splice site region in activated spliceosomes. *EMBO J* 25:2475–2486.
- Valadkhan S, Manley JL (2001) Splicing-related catalysis by protein-free snRNAs. *Nature* 413:701–707.
- Valadkhan S, Manley JL (2003) Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA* 9:892–904.
- Valadkhan S, Mohammadi A, Wachtel C, Manley JL (2007) Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing. *RNA* 13:2300–2311.
- Abelson J (2008) Is the spliceosome a ribonucleoprotein enzyme? *Nat Struct Mol Biol* 15:1235–1237.
- Sontheimer EJ, Sun S, Piccirilli JA (1997) Metal ion catalysis during splicing of pre-messenger RNA. *Nature* 388:801–805.
- Gordon PM, Sontheimer EJ, Piccirilli JA (2000) Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: Extending the parallels between the spliceosome and group II introns. *RNA* 6:199–205.
- Nilsen TW, Maroney PA, Robertson HD, Baglioni C (1982) Heterogeneous nuclear RNA promotes synthesis of (2',5')oligoadenylate and is cleaved by the (2',5')oligoadenylate-activated endoribonuclease. *Mol Cell Biol* 2:154–160.
- Fujimoto M, Kuninaka A, Yoshino H (1974) Nuclease from *Penicillium citrinum*. III. Substrate specificity of nuclease P1. *Agricultural Biol Chem* 38:1555–1561.
- Chu VT, Liu Q, Podar M, Perlman PS, Pyle AM (1998) More than one way to splice an RNA: Branching without a bulge and splicing without branching in group II introns. *RNA* 4:1186–1202.
- Podar M, Chu VT, Pyl AM, Perlman PS (1998) Group II intron splicing *in vivo* by first-step hydrolysis. *Nature* 391:915–918.
- Tseng CK, Cheng SC (2008) Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* 320:1782–1784.
- Wassarman DA, Steitz JA (1992) Interactions of small nuclear RNA's with precursor messenger RNA during *in vitro* splicing. *Science* 257:1918–1925.
- Sashital DG, Cornilescu G, McManus CJ, Brow DA, Butcher SE (2004) U2–U6 RNA folding reveals a group II intron-like domain and a four-helix junction. *Nat Struct Mol Biol* 11:1237–1242.
- Yean SL, Wuenschell G, Termini J, Lin RJ (2000) Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* 408:881–884.
- Vicens Q, Paukstelis PJ, Westhof E, Lambowitz AM, Cech TR (2008) Toward predicting self-splicing and protein-facilitated splicing of group I introns. *RNA* 14:2013–2029.
- Jacquier A, 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.
- Jarrell KA, Peebles CL, Dietrich RC, Romiti SL, Perlman PS (1988) Group II intron self-splicing. Alternative reaction conditions yield novel products. *J Biol Chem* 263:3432–3439.
- Grainger RJ, Beggs JD (2005) Prp8 protein: At the heart of the spliceosome. *RNA* 11:533–557.
- Hsieh J, Andrews AJ, Fierke CA (2004) Roles of protein subunits in RNA-protein complexes: Lessons from ribonuclease P. *Biopolymers* 73:79–89.