Biol. Chem., Vol. 388, pp. 693–697, July 2007 • Copyright © by Walter de Gruyter • Berlin • New York. DOI 10.1515/BC.2007.080

Minireview

The spliceosome: a ribozyme at heart?

Saba Valadkhan

Center for RNA Molecular Biology, Case Western Reserve University, Cleveland, OH 44106, USA

e-mail: saba.valadkhan@case.edu

Abstract

The spliceosome, the multi-megadalton molecular machine that performs splicing, consists of over 200 different proteins and five small nuclear RNAs (snRNAs). Extensive mechanistic and structural similarities to self-splicing group II introns, large ribozymes found in prokaryotes and lower eukaryotes that catalyze an identical reaction, strongly suggest that the spliceosomal RNAs are in fact the catalytic components of the spliceosome. Of the five spliceosomal RNAs, U2 and U6 are the only ones that are absolutely required for both steps of splicing. These two snRNAs form an elaborate base-paired complex that might in fact constitute the active site of the spliceosome.

Keywords: active site; catalysis; ribozyme; splicing mechanism; U6 snRNA.

Introduction

Splicing of pre-messenger RNAs to mature transcripts is a crucial and elaborate step in the expression of most eukaryotic genes. Almost all human pre-messenger RNAs undergo multiple splicing events, and alternative splicing is not only one of the most important means of regulation of gene expression, but is also largely responsible for generating proteomic diversity in eukaryotes. Commensurate with its crucial role, the spliceosome, the multi-megadalton ribonucleoprotein assembly that catalyses the splicing reaction, is the largest and one of the most complicated cellular machines known (reviewed in Nilsen, 1998; Will and Luhrmann, 2006). The spliceosome consists of five small nuclear RNAs, or snRNAs, named U1, U2, U4, U5 and U6, and over 200 different proteins. Early on, structural and mechanistic similarities between the spliceosome and self-splicing group II introns led many to develop the hypothesis that the spliceosome is also a ribozyme and a descendent of group II-like introns. Consistent with this hypothesis, it has been shown that snRNAs play key roles in several aspects of the splicing reaction, including splice-site recognition and substrate positioning in the active site throughout the spliceosomal cycle.

At early stages in spliceosome assembly, U1 snRNA forms a base-pairing interaction with the 5′-splice site, thus marking the upstream boundary of the intron. U2 snRNA similarly forms a base-pairing interaction with the branch site, a region close to the 3′-splice site that contains a highly conserved adenosine residue. Importantly, this base-pairing interaction positions the conserved adenosine in a bulged, extrahelical conformation, which is important for its catalytic function (Chu et al., 1998, Figure 1). Next, U4, U5 and U6 snRNAs and their associated proteins join the spliceosome. U4 snRNA, which is initially base-paired with U6 snRNA through two base-paired helices, dissociates from U6 and leaves the spliceosome. The departure of U4 allows U6 snRNA to replace U1 at the 5′-splice site, and at the same time form extensive base-pairing interactions with U2 snRNA. The U6-U2 base-pairing interactions help to bring the 5′-splice site and the branch site close to each other (Figure 1). During the first step of splicing, the 2′-OH of the bulged branch-site adenosine is activated through an unknown mechanism to make a nucleophilic attack at the 5′-splice site, releasing the 5′-exon and forming a 2′-5′-linkage with the phosphate at the 5′-splice site through a transesterification reaction. U5 snRNA forms non-canonical base-pairing interactions with the two exons, keeping the released 5′-exon in alignment with the 3′-exon for the second step of splicing. During the second step, the 3′-OH of the first exon, which was the leaving group in the first step of splicing, becomes the second-step nucleophile in a transesterification reaction that results in release of the lariat intron and ligation of the two exons (Brow, 2002; Figure 1).

A catalytic role for snRNAs in splicing

As detailed above, the ability of RNA molecules to form strong, specific interactions via base-pairing with another RNA is extensively utilized in the spliceosome. Base-pairing interactions contribute to substrate recognition (U1 and U2), positioning the branch site in a strained, catalytic conformation (U2), regulation of the activity of another snRNA (U4), and juxtaposition of reactive substrates (U2, U5 and U6). Similar RNA-RNA interactions play identical or closely related roles in group II introns, and it is conceivable that snRNAs have retained their roles throughout the evolution of the spliceosome from group II-like introns.

RNAs can perform the above-mentioned tasks simply and even more effectively than proteins. However, when it comes to catalysis, proteins seem to have an advantage over RNA, at least in the case of natural ribozymes (Emilsson et al., 2003). Thus, the remaining question is whether the catalytic role that RNA plays in group II introns has been similarly preserved during the evolution

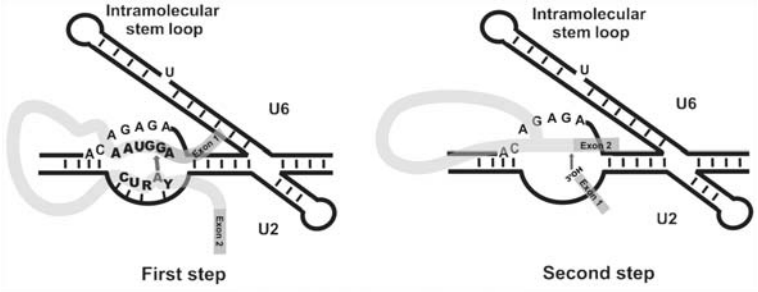

Figure 1 The base-paired complex of human U6 and U2 and the pre-mRNA during the two splicing steps.
The snRNAs are shown by black lines and letters. The introns are shown in light grey and the exons in dark grey. The U6 intramolecular stem loop is shown. Arrows mark the direction of nucleophilic attacks during the first and second steps of splicing. The bulged branch-site adenosine is shown in the left panel.

of the spliceosome, in other words, if the snRNAs can perform splicing catalysis without the help of proteins. Alternatively, the spliceosomal proteins might have replaced some of the RNA components of the active site, contributing to productive positioning of the reactive groups, activation of the nucleophile for the first or second step of splicing, stabilization of the transition states, or stabilizing the developing charge on the leaving groups, thus assisting their departure. On the other hand, while it is very likely that spliceosomal proteins make important contributions to substrate positioning in the active site, and might even assist in one or more aspects of the catalysis mechanism, these contributions might be redundant, and thus do not necessarily rule out RNA catalysis in the spliceosome. If the spliceosome is indeed a ribozyme, while proteins may contribute to efficiency and accuracy of splicing, the RNA components should be able to perform splicing in the absence of proteins, albeit with lower efficiency, as has been observed for other natural ribozymes (for a review see Hsieh et al., 2004).

Evidence for RNA catalysis in the spliceosome

At the time of splicing catalysis, U1 and U4 snRNAs have already left the spliceosome, and the conserved loop 1 of U5 snRNA, which helps to align the exons, is not required for the first step of splicing, at least *in vitro* (O'Keefe et al., 1996; Segault et al., 1999). This leaves U6 and U2 snRNAs as the only spliceosomal RNAs that are absolutely needed for both splicing steps. Thus, if snRNAs are the catalytic components of the spliceosome, U6 and U2 snRNAs should be able to catalyze the splicing reaction. Consistent with having a critical role in splicing, these two snRNAs are extremely conserved. This is particularly true for U6 snRNA, which contains two evolutionarily invariant regions, the ACAGAGA box and the AGC triad (Figure 2). Interestingly, an intramolecular stem loop in U6 has significant structural similarity to the catalytically crucial domain V of group II introns (Figure 2). Both stem loops have an AGC sequence in similar positions, an asymmetric internal loop, and a GNRA-type loop capping the helix. Numerous functional similarities between the two stem loops suggest that the two loops play similar catalytic roles in the two systems (Figure 2 and see below).

If U6 and U2 snRNAs are indeed the catalytic components of the spliceosome, then certain residues in them should form the active site of the splicing reactions. These residues are expected to be highly sensitive to mutations, and it should also be possible to prove their spatial proximity to the reacting groups of the splicing reaction. Interestingly, the invariant ACAGAGA box of U6 meets both these criteria. Several mutations in this conserved region are incompatible with splicing, or result in blocking of splicing after the first step (McPheeters,

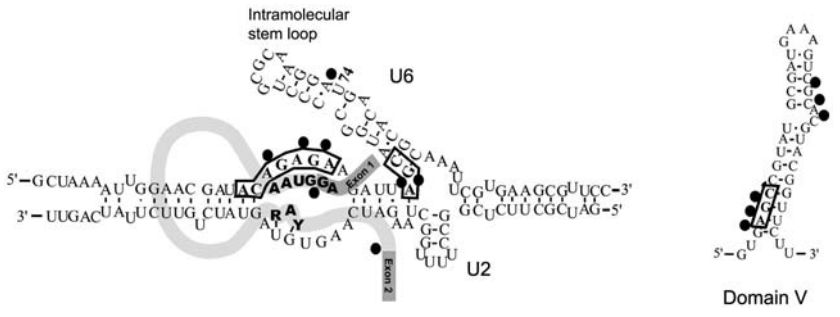

Figure 2 The U6/U2 base-paired complex and domain V of group II introns.
The human U6 and U2 snRNAs and domain V are shown in black, and the pre-mRNA is shown in light grey (introns) or dark grey (exons). The bulged branch-site adenosine is shown. The ACAGAGA and AGC conserved regions are boxed. The site of the intramolecular stem loop of U6 is indicated, and U74 residue in U6 is shown. Bullets mark the site of phosphorothioate interference.

1996, and references therein). In addition, cross-linking and complementation studies have proved the proximity of the 5′-splice site to the last nucleotides of the ACAG**AG**A box (Sontheimer and Steitz, 1993; Kim and Abelson, 1996; Luukkonen and Seraphin, 1998; Figure 2). Furthermore, phosphorothioate substitution analysis of the backbone phosphate groups in U6 snRNA has identified a number of phosphate oxygens that, when substituted with sulfur, result in blocking of the first or second step of splicing. Intriguingly, most of the phosphorothioate interference sites fall within the ACAGAGA and AGC sequences (Valadkhan and Manley, 2002; Figure 2). The close correlation between the phosphorothioate interference sites in U6 and those in domain V of group II introns adds further to the parallels between the two systems (Figure 2). Taken together, these results provide strong evidence supporting a catalytic role for the ACAGAGA and AGC conserved domains of U6 in the active site of the spliceosome. While U2 snRNA is also absolutely required for the splicing reaction, current data suggest that this is likely due to its role in positioning of the first-step nucleophile rather than direct participation in catalysis. Consistent with this possibility, deletion of the U2-like domain of group II introns, domain VI, does not compromise the catalytic activity of these introns *in vitro*. Rather, group II introns lacking domain VI efficiently catalyze splicing by performing hydrolysis as an alternative pathway to branching during the first step (Lehmann and Schmidt, 2003).

Further evidence of the catalytic competence of U6 snRNA is provided by direct catalytic assays, in which purified, protein-free fragments of human U6 and U2 snRNAs were incubated with short RNA oligonucleotides that resembled splicing substrates. The protein-free fragments of U6 and U2 were not only able to form the base-paired complex that is observed in the activated spliceosomes, but could also bind the splicing substrates in a sequence-specific manner (Valadkhan and Manley, 2000, 2001, 2003). Furthermore, this *in vitro*-assembled U6/U2 complex could perform a number of splicing-related catalytic reactions (Valadkhan and Manley, 2001, 2003). Although these RNA-catalyzed reactions are slow and inefficient, they demonstrate clear similarities to the splicing reaction in terms of the chemistry of the reaction and the requirement for the ACAGAGA box of U6 and the presence of Mg^{2+} in the reaction buffer. Thus, these data provide direct evidence for the catalytic activity of spliceosomal snRNAs and their competence to act as the catalytic components of the spliceosome. Considering the very minimal length of spliceosomal snRNAs, the low efficiency of their catalytic activity likely reflects their dependence on other spliceosomal factors for folding and stabilization of their secondary and tertiary structure, roles commonly played by proteins in catalytic ribonucleoprotein complexes (Hsieh et al., 2004).

The base-paired U6/U2 complex as the spliceosome active site

As mentioned above, it is highly likely that the base-paired complex between U6 and U2 snRNAs contains all the elements required for catalysis of the splicing reaction, and that the ACAGAGA box of U6 might directly participate in catalysis. Another RNA element that likely forms part of the spliceosomal active site is the intramolecular stem loop of U6 (Figure 2; see also above). While the structure of the base of this stem loop is slightly different between human and yeast spliceosomes (Valadkhan, 2005), the global structure of the stem loop is well supported by complementation studies in both systems (Nilsen, 1998; Will and Luhrmann, 2006). A recent hydroxyl-radical cleavage analysis performed in catalytically activated spliceosomes indicated that the U6 stem loop was in proximity to the ACAGAGA box both before and after the first step of splicing (Rhode et al., 2006). This result is highly appealing, since it provides direct evidence of the juxtaposition of the functionally critical elements in the U6 stem loop and the ACAGAGA box in catalytically active spliceosomes (Figure 2). Genetic complementation analyses in human spliceosome and NMR studies on U6/U2 fragments indicate the presence of a four-way junction in the U2/U6 complex, which can provide a structural basis for the proximity of the U6 stem loop and the ACAGAGA box (Sun and Manley, 1995; Sashital et al., 2004).

Interestingly, phosphorothioate substitution studies have shown that the phosphate 5′ to a highly conserved U residue (U74 in human numbering, Figure 2) is involved in a functionally required metal coordination (Yean et al., 2000; Huppler et al., 2002). Similar phosphorothioate substitution analyses at the 5′-splice site have indicated that the spliceosome takes advantage of metal ions, likely magnesium, to assist in stabilization of the developing charge on the leaving group during the first step of splicing (Sontheimer et al., 1997). Whether the U6 stem loop assists in coordination of this catalytic metal ion or another active-site cation remains to be established; however, recent data suggest that U74 might be facing away from the active site and thus might have a structural rather than a catalytic role (Rhode et al., 2006). Thus, based on current data, the ACAGAGA box and the intramolecular stem loop of U6 are the primary candidates for forming the spliceosomal active site (Butcher and Brow, 2005).

The other invariant region of U6, the AGC triad, is also highly sensitive to mutations in nucleobases and backbone elements (Valadkhan and Manley, 2002; Hilliker and Staley, 2004), and data for group II introns suggest a direct catalytic role for the G residue in that system (Konforti et al., 1998; Gordon and Piccirilli, 2001). *In vivo* complementation and cross-linking analyses in protein-free human U6/U2 complex suggest that in the tertiary structure of the U6/U2 complex, the AGC triad is in spatial proximity to the ACAGAGA box (Madhani and Guthrie, 1994; Valadkhan and Manley, 2000). Thus, similar to the group II introns, the AGC sequence might directly participate in formation of the spliceosomal active site.

Recently, a working model for the active site of the group II introns has been developed (de Lencastre et al., 2005). Based on this model, several RNA elements of group II introns, including the U6-like domain V, participate in formation of the active site. While most of these components have direct counterparts in the spliceo-

some, there are no equivalents for a number of RNAs that have a role in providing structural support in the group II intron active site. It is quite likely that during the course of evolution, these supporting elements have been replaced by spliceosomal proteins. Indeed, a highly conserved spliceosomal protein, Prp8, is known to contact both splice sites and the branch site in assembled spliceosomes, and current data suggest that it likely has a role in providing structural support for other active-site elements (Grainger and Beggs, 2005). While direct participation of Prp8 in catalysis has not been formally ruled out, catalytic activity of protein-free snRNAs is consistent with a structural support role for this spliceosomal protein (see above).

Conclusion

Existing data suggest that the RNA elements of the spliceosome are almost certainly the catalytic components of this large cellular machine. However, spliceosomal proteins most likely play significant roles in buttressing the active-site elements and remodeling the active site between the two steps of splicing. Owing to a lack of high-resolution structural data and the technical limitations imposed by the spliceosome assembly process, our current knowledge of the organization of the spliceosomal active site and the role of individual RNA and protein elements remains very incomplete. Recent advances in purification of activated spliceosomes stalled before the first or second step of splicing, combined with the development of minimalistic splicing model systems, promise to provide significant new insights into the structure and function of the spliceosomal active site in the near future. These studies would also shed light on the evolutionary origin of the spliceosome and its relationship to the self-splicing group II introns. In addition to its physiological and medical significance, the spliceosome, a likely descendent of group II-like introns, constitutes an ideal model system for studying the evolution of the RNA world to the modern, protein-dominated one.

Acknowledgments

I thank Tim Nilsen for critical review of the manuscript, the Kinship Foundation for a Searle Scholar Award, and the National Institute of Health for Grant no. GM078572.

References

Brow, D.A. (2002). Allosteric cascade of spliceosome activation. Annu. Rev. Genet. *36*, 333–360.

Butcher, S.E. and Brow, D.A. (2005). Towards understanding the catalytic core structure of the spliceosome. Biochem. Soc. Trans. *33*, 447–449.

Chu, V.T., Liu, Q., Podar, M., Perlman, P.S., and Pyle, A.M. (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.

de Lencastre, A., Hamill, S., and Pyle, A.M. (2005). A single active-site region for a group II intron. Nat. Struct. Mol. Biol. *12*, 626–627.

Emilsson, G.M., Nakamura, S., Roth, A., and Breaker, R.R. (2003). Ribozyme speed limits. RNA *9*, 907–918.

Gordon, P.M. and Piccirilli, J.A. (2001). Metal ion coordination by the AGC triad in domain 5 contributes to group II intron catalysis. Nat. Struct. Biol. *8*, 893–898.

Grainger, R.J. and Beggs, J.D. (2005). Prp8 protein: at the heart of the spliceosome. RNA *11*, 533–557.

Hilliker, A.K. and Staley, J.P. (2004). Multiple functions for the invariant AGC triad of U6 snRNA. RNA *10*, 921–928.

Hsieh, J., Andrews, A.J., and Fierke, C.A. (2004). Roles of protein subunits in RNA-protein complexes: lessons from ribonuclease P. Biopolymers *73*, 79–89.

Huppler, A., Nikstad, L.J., Allmann, A.M., Brow, D.A., and Butcher, S.E. (2002). Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure. Nat. Struct. Biol. *9*, 431–435.

Kim, C.H. and Abelson, J. (1996). Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5′ splice site. RNA *2*, 995–1010.

Konforti, B.B., Abramovitz, D.L., Duarte, C.M., Karpeisky, A., Beigelman, L., and Pyle, A.M. (1998). Ribozyme catalysis from the major groove of group II intron domain 5. Mol. Cell *1*, 433–441.

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.

Luukkonen, B.G. and Seraphin, B. (1998). Genetic interaction between U6 snRNA and the first intron nucleotide in *Saccharomyces cerevisiae*. RNA *4*, 167–180.

Madhani, H.D. and Guthrie, C. (1994). Randomization-selection analysis of snRNAs *in vivo*: evidence for a tertiary interaction in the spliceosome. Genes Dev. *8*, 1071–1086.

McPheeters, D.S. (1996). Interactions of the yeast U6 RNA with the pre-mRNA branch site. RNA *2*, 1110–1123.

Nilsen, T.W. (1998). RNA-RNA interactions in nuclear pre-mRNA splicing. In: RNA Structure and Function, R.W. Simons and M. Grunberg-Manago, eds. (Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press), pp. 279–307.

O'Keefe, R.T., Norman, C., and Newman, A.J. (1996). The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. Cell *86*, 679–689.

Rhode, B.M., Hartmuth, K., Westhof, E., and 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.

Sashital, D.G., Cornilescu, G., McManus, C.J., Brow, D.A., and Butcher, S.E. (2004). U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction. Nat. Struct. Mol. Biol. *11*, 1237–1242.

Segault, V., Will, C.L., Polycarpou-Schwarz, M., Mattaj, I.W., Branlant, C., and Luhrmann, R. (1999). Conserved loop I of U5 small nuclear RNA is dispensable for both catalytic steps of pre-mRNA splicing in HeLa nuclear extracts. Mol. Cell. Biol. *19*, 2782–2790.

Sontheimer, E.J. and Steitz, J.A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. Science *262*, 1989–1996.

Sontheimer, E.J., Sun, S., and Piccirilli, J.A. (1997). Metal ion catalysis during splicing of premessenger RNA. Nature *388*, 801–805.

Sun, J.S. and Manley, J.L. (1995). A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. Genes Dev. *9*, 843–854.

Valadkhan, S. (2005). snRNAs as the catalysts of pre-mRNA splicing. Curr. Opin. Chem. Biol. *9*, 603–608.

Valadkhan, S. and Manley, J.L. (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.

Valadkhan, S. and Manley, J.L. (2001). Splicing-related catalysis by protein-free snRNAs. Nature *413*, 701–707.

Valadkhan, S. and Manley, J.L. (2002). Intrinsic metal binding by a spliceosomal RNA. Nat. Struct. Biol. *9*, 498–499.

Valadkhan, S. and Manley, J.L. (2003). Characterization of the catalytic activity of U2 and U6 snRNAs. RNA *9*, 892–904.

Will, C.L. and Luhrmann, R. (2006). Spliceosome structure and function. In: The RNA World, R.F. Gesteland, T.R. Cech and J.F. Atkins, eds. (Cold Spring Harbor, NY, USA: Cold Spring Harbor Laboratory Press), pp. 369–400.

Yean, S.L., Wuenschell, G., Termini, J., and Lin, R.J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. Nature *408*, 881–884.