The use of simple model systems to study spliceosomal catalysis

SABA VALADKHAN[1] and JAMES L. MANLEY[2]

[1]Center for RNA Molecular Biology, Case Western Reserve University, Cleveland, Ohio 44106, USA
[2]Department of Biological Sciences, Columbia University, New York, New York 10027, USA

ABSTRACT

Since direct analysis of many aspects of spliceosomal function is greatly hindered by the daunting complexity of the spliceosome, the development of functionally validated simple model systems can be of great value. The critical role played by a base-paired complex of U6 and U2 snRNAs in splicing in vivo suggests that this complex could be a suitable starting point for the development of such a simple model system. However, several criteria must be satisfied before such a snRNA-based in vitro system can be considered a valid model for the spliceosomal catalytic core, including similarities at the level of reaction chemistry and cationic and sequence requirements. Previous functional analyses of in vitro assembled base-paired complexes of human U2 and U6 snRNAs have been promising, providing insight into catalysis. Furthermore, they strongly suggest that with further optimization, these RNAs might indeed be able to recapitulate the function of the spliceosomal catalytic core, thus opening the door to several lines of study not previously possible.

Keywords: minimal splicing systems; U6; U2; snRNA; ribozyme; catalysis

Due to the intractable complexity of biological systems, the vast majority of scientific research is performed on simplified study models. Similarly, the enormous complexity of the human spliceosome, one of the largest and most critical cellular machines known (Nilsen 2003), has necessitated the use of in vitro systems based on cellular extracts or synthetic molecules as well as reliance on yeast genetics as a more tractable in vivo model. The contribution of these model systems to our understanding of spliceosomal function has been invaluable.

Mechanistic and structural similarities between group II self-splicing introns and the spliceosome together with cross-linking and mutagenesis data suggest a critical role for spliceosomal snRNAs in splicing catalysis (Valadkhan 2007; Dayie and Padgett 2008). In order to investigate the catalytic potential of the snRNAs, we have analyzed the ability of in vitro synthesized, protein-free snRNAs to perform catalysis on short oligonucleotide substrates. Such an approach has been widely used in other catalytic systems consisting of RNA and protein, such as RNase P and group

I and II introns, and the results have made significant inroads into our understanding of the function of the RNAs in their in vivo context (Lilley 2005; Strobel and Cochrane 2007).

Our work has revealed two related catalytic activities from protein-free snRNAs that share several similarities with each other and with spliceosomal catalysis (Valadkhan and Manley 2001, 2003; Valadkhan et al. 2007). In one reaction, incubation of an RNA oligonucleotide containing the intronic branch site consensus sequence with a previously characterized in vitro-assembled base-paired complex of U6 and U2 snRNAs (Valadkhan and Manley 2000) led to the formation of a covalent linkage between the branch site adenosine in the branch site oligonucleotide and the conserved AGC triad of U6 (Valadkhan and Manley 2001). The reaction was dependent on divalent cations, the presence of a hydroxyl group at the 2′ position of the branch site adenosine, and the ability of the branch site oligonucleotide to base-pair with U2 snRNA. Further, mutations in the conserved ACAGAGA and AGC sequences of U6, which block splicing in vivo (Madhani et al. 1990; Madhani and Guthrie 1992; Datta and Weiner 1993; McPheeters 1996), also blocked this reaction. Taken together, the reaction suggested that the branch site adenosine was juxtaposed to the AGC triad in the folded structure of U6/U2/branch site oligonucleotide complex

and was induced to react with the AGC triad in a Mg^{++} and ACAGAGA-dependent manner, most likely through its 2′OH. In a follow up work, we could induce the branch site oligonucleotide to instead react with a short RNA sequence that carried the 5′ splice site consensus sequence when both were incubated with the U6/U2 complex (Valadkhan et al. 2007). This reaction similarly required divalent cations and base-pairing interaction between the branch site oligonucleotide and U2, and is sensitive to mutations in the ACAGAGA, AGC, and the U6 intra-molecular stemloop (ISL), another element suspected to be important for splicing catalysis in vivo (McPheeters 1996; Butcher and Brow 2005). However, despite these similarities with authentic splicing, the reactions are very inefficient, which has hindered conclusive and complete characterization of the products. More recently, though, by redesigning the substrate oligonucleotides, we observed an ACAGAGA-, AGC-, and ISL-dependent activity from the U6/U2 base-paired complex that involves removal of an intervening sequence from a pre-mRNA-like construct in a two-step reaction that is chemically identical to a splicing reaction (A. Mohammadi, Y. Jaladat, S. Geisler, and S. Valadkhan, unpubl.).

In an accompanying Perspective, Smith and Konarska (2009a) raise a number of very important points about the role and validity of such simple model systems in the study of the splicing reaction and biological phenomena in general that merit further discussion. While by nature minimal model systems fall short of capturing the full complexity of biological phenomena, they have been invaluable in helping to solve otherwise intractable questions. The study of in vitro-synthesized ribozymes (and also protein enzymes) has provided significant insights into their biological function and, in studied cases, the presence of additional factors including proteins in vivo does not seem to affect the function of the RNAs in a fundamental way (Hsieh et al. 2004). Indeed, the spliceosome's closest relatives, group II introns, require proteins for splicing in vivo; however, a large share of our knowledge on these introns comes from the study of protein-free self-splicing systems. Thus, three decades of experience in the ribozyme field suggests (but certainly does not guarantee) that performing an in vitro analysis on protein-free snRNAs similar to those done on known natural ribozymes will provide valuable insights into their potential catalytic activity that could be highly relevant in vivo.

We are in full agreement with Smith and Konarska (2009a) that several important points should be taken into consideration when analyzing the catalytic activity of a potential natural ribozyme. First, the activity should originate from the sequences present in the biologically relevant complex, in this case, the U2 and U6 snRNAs found to be present in mature spliceosomes and necessary for their function. It should be ascertained that the activity is not in any way related to any added structural modules or altered sequences. Also, the observed activity should involve sequence elements known to be critical for the activity of the RNA in vivo due to their high conservation or in vivo mutagenesis data. A crucial role for sequences such as ACAGAGA and AGC domains and the ISL in spliceosomal U6 (Nilsen 1998) or the conserved domain V in group II introns (Lehmann and Schmidt 2003; Pyle and Lambowitz 2006) in the catalytic activity observed in vitro suggest that the reaction could be related to the biological function of the RNA. While a dependence on such conserved sequence elements is necessary for biological relevance, we agree with Smith and Konarska (2009a) that it is not by itself sufficient to guarantee that the activity observed in the in vitro system recapitulates the one occurring in vivo.

The nature of the observed chemistry is also important. However, it should be taken into consideration that both ribozymes and protein enzymes can catalyze a number of related reactions from the same active site, depending on the exact complement of substrates they encounter (Que 2000; Lehmann and Schmidt 2003). The versatile catalytic activity of group I and group II introns is well known. For example, group II introns can perform hydrolysis rather than branching during the first step of splicing (Chu et al. 1998), and it has been recently demonstrated that the spliceosomal active site is also capable of hydrolytic cleavage of the 3′ splice site (Tseng and Cheng 2008). Thus, while catalysis of a reaction identical to the one performed in vivo is certainly an optimal outcome, closely related but distinct reactions can also result from an active site fully competent to catalyze the in vivo-observed reaction. Careful and thorough analysis of the product(s) of any reaction will help define the relationship of the observed chemistry to that performed in vivo. However, limitations of the currently available techniques might prohibit a full characterization. For example, both our group and the Konarska group have not been able to detect or characterize small RNA fragments that were released from substrates during RNA-catalyzed reactions and migrated among a significant excess of degradation-induced fragments (Valadkhan et al. 2007; Smith and Konarska 2009b). However, even in such cases, often the reaction can be reasonably well characterized by employing less direct analysis methods such as careful characterization of the other RNAs formed during the reaction and ruling out alternative possibilities.

In the case of the spliceosome, further investigation of the level of similarity between a promising in vitro reaction and the reaction occurring in vivo is hindered by our limited understanding of spliceosomal catalysis. Nonetheless, several aspects of spliceosomal catalysis are clear. The reaction involves two transesterification reactions, leading to the formation of a 2′–5′ bond in the lariat intron and a 3′–5′ linkage in the spliced product (Nilsen 1998). Metal ion coordination by the 3′-linked oxygen of the bridging phosphate at the 5′ splice site suggests the involvement of

at least one inner sphere coordinated divalent cation (Sontheimer et al. 1997). It has also been shown that at least the first step of the splicing reaction occurs in the vicinity of a conserved sequence in U6, the above-mentioned ACAGAGA box (Sontheimer and Steitz 1993; Konarska et al. 2006). As mentioned above, two other regions of U6, the AGC triad and a conserved asymmetric bulge in the middle of the U6 ISL, have also shown a high sensitivity to nucleobase and backbone mutations (Valadkhan 2007). While the observed mutational sensitivity does not necessarily indicate catalytic relevance, a recent high resolution structure has indicated that corresponding sequences in the group II introns form the active site, along with the ACAGAGA equivalent in that system (Toor et al. 2008). This finding raises the likelihood that these three regions are similarly involved in forming the active site in the spliceosome. Further, it has been shown that, in fully assembled spliceosomes, the asymmetric bulge area of ISL and ACAGAGA sequence are in close proximity (Rhode et al. 2006). The equivalents of these two sequences are in close proximity in the high resolution Group II intron structure (Toor et al. 2008).

How similar then is the relationship between the snRNA-catalyzed reactions and splicing in vivo? The RNA sequences used in our protein-free catalytic assays are rather small, 75 and 45 nucleotides long, but contain all the minimal sequence elements known to be essential for splicing in the authentic spliceosome. All reactions we have observed from the protein-free snRNAs required divalent cations and involved the same three regions of U6 snRNA, the equivalents of which form the active site of group II introns. Would it be possible for these three domains of U6 snRNA to play a role in catalysis of splicing in the spliceosome and in the in vitro, protein-free reactions in completely unrelated ways? While the many similarities between the snRNA-catalyzed in vitro reactions and the in vivo splicing reaction argue against this possibility, in order to completely rule out such a scenario it should be demonstrated that the protein-free U6/U2 complex is indeed capable of catalyzing the chemistry of splicing in an ACAGAGA-, AGC-, and ISL-dependent manner. While it is possible for similar sequences to form different active sites, and different sequences may form very similar active sites that catalyze the same reaction (Shih and Been 2002), for such an snRNA-catalyzed splicing reaction to be unrelated to in vivo splicing, the same short sequences should contribute to the formation of two completely unrelated active sites that catalyze the same reaction. To our knowledge, this has never been observed. In biological systems, ribozymes capable of catalyzing splicing are large and complex RNA molecules, suggesting the paucity of such activities in sequence space. The fact that, despite the efforts of several groups in the past, no catalytic activity beyond the ones described by us has been reported from unaltered sequences of U6 and U2 snRNAs

further indicates the nonpromiscuous nature of these RNAs.

While defining the catalytic ability of a potential natural ribozyme is of interest by itself, it also raises interesting questions about the role of the RNAs in their in vivo RNP context. All natural eukaryotic ribozymes require protein cofactors for function in vivo, but as long as there is no evidence for direct involvement of proteins in catalysis, the RNA can be considered a ribozyme in vivo. In studied natural ribozymes proteins play a structural supportive role, stabilizing the active site and alleviating the need for high concentrations of divalent cations (Hsieh et al. 2004). In the specific case of the spliceosome, a highly conserved U5 snRNP-associated protein, Prp8, is the most likely candidate protein for having a catalytic role. However, the existing data are most consistent with a structural and not a catalytic role for Prp8, as currently characterized mutations in Prp8 seem to mainly affect the rearrangements between different active site conformations at different stages of splicing reaction (Grainger and Beggs 2005; Liu et al. 2007). Recently two groups (Pena et al. 2008; Yang et al. 2008) have indicated the presence of an RNase H-like motif in a region of Prp8 previously shown to be close to the 5′ splice site in the spliceosomal B complex (Reyes et al. 1999). The catalytic core of this motif, however, is degenerate in Prp8 and cannot bind either RNA or divalent cations by itself. While mutation of a number of residues forming the RNase H-like motif leads to a lethal phenotype, at least in the case of one of these residues, the one most closely corresponding to a metal binding residue in the actual RNase H active site, this phenotype seems to reflect misfolding of the protein rather than disruption of a catalytic activity (Yang et al. 2008). Analysis of the integrity of U5 snRNP particle and spliceosome formation in these mutants will help determine if this is indeed the case. However, conclusive determination of the role of Prp8 in the spliceosomal active site will require detailed biochemical analysis, and a simple, well-characterized snRNA-based system that partially recapitulates the spliceosomal active site will provide an ideal tool for addressing such questions.

We are at the very early stages of understanding the catalytic potential of snRNAs and crucial questions remain unanswered: Is the recently observed protein-free splicing-related activity of snRNAs in vitro a vestigial remnant from their group II-like ancestors, or is spliceosomal catalysis indeed completely protein-free? Are the same exact functional groups involved? How does the catalytic strategy employed by snRNAs compare to that of known natural ribozymes? How do the spliceosomal proteins modulate the structure and catalytic activity of the snRNAs? As also noted by (Smith and Konarska 2009a), providing definitive answers to these questions and the many others that will follow requires time, parallel experiments in vitro and in vivo, and technical advances in RNA biology. However, it

should be kept in mind that scientific advances occur in steps, and no story is ever complete.

REFERENCES

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.

Datta, B. and Weiner, A.M. 1993. The phylogenetically invariant ACAGAGA and AGC sequences of U6 small nuclear RNA are more tolerant of mutation in human cells than in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **13:** 5377–5382.

Dayie, K.T. and Padgett, R.A. 2008. A glimpse into the active site of a group II intron and maybe the spliceosome, too. *RNA* **14:** 1697–1703.

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

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.

Konarska, M.M., Vilardell, J., and Query, C.C. 2006. Repositioning of the reaction intermediate within the catalytic center of the spliceosome. *Mol. Cell* **21:** 543–553.

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.

Lilley, D.M. 2005. Structure, folding and mechanisms of ribozymes. *Curr. Opin. Struct. Biol.* **15:** 313–323.

Liu, L., Query, C.C., and Konarska, M.M. 2007. Opposing classes of prp8 alleles modulate the transition between the catalytic steps of pre-mRNA splicing. *Nat. Struct. Mol. Biol.* **14:** 519–526.

Madhani, H.D. and Guthrie, C. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **71:** 803–817.

Madhani, H.D., Bordonne, R., and Guthrie, C. 1990. Multiple roles for U6 snRNA in the splicing pathway. *Genes & Dev.* **4:** 2264–2277.

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* (eds. R.W. Simons and M. Grunberg-Manago), pp. 279–307. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Nilsen, T.W. 2003. The spliceosome: The most complex macromolecular machine in the cell? *Bioessays* **25:** 1147–1149.

Pena, V., Rozov, A., Fabrizio, P., Lührmann, R., and Wahl, M.C. 2008. Structure and function of an RNase H domain at the heart of the spliceosome. *EMBO J.* **27:** 2929–2940.

Pyle, A.M. and Lambowitz, A.M. 2006. Group II introns: Ribozymes that splice RNA and invade DNA. In *The RNA world* (eds. R.F. Gesteland et al.), pp. 469–506. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Que Jr., L. 2000. One motif—Many different reactions. *Nat. Struct. Biol.* **7:** 182–184.

Reyes, J.L., Gustafson, E.H., Luo, H.R., Moore, M.J., and Konarska, M.M. 1999. The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5′ splice site. *RNA* **5:** 167–179.

Rhode, B.M., Hartmuth, K., Westhof, E., and Lührmann, R. 2006. Proximity of conserved U6 and U2 snRNA elements to the 5′ splice site region in activated spliceosomes. *EMBO J.* **25:** 2475–2486.

Shih, I.H. and Been, M.D. 2002. Catalytic strategies of the hepatitis delta virus ribozymes. *Annu. Rev. Biochem.* **71:** 887–917.

Smith, D.J. and Konarska, M.M. 2009a. A critical assessment of the utility of protein-free splicing systems. *RNA* **15:** (this issue). doi: 10.1261/rna.1322709.

Smith, D.J. and Konarska, M.M. 2009b. Identification and characterization of a short 2′–3′ bond-forming ribozyme. *RNA* **15:** (this issue). doi: 10.1261/rna.1321909.

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.

Strobel, S.A. and Cochrane, J.C. 2007. RNA catalysis: Ribozymes, ribosomes, and riboswitches. *Curr. Opin. Chem. Biol.* **11:** 636–643.

Toor, N., Keating, K.S., Taylor, S.D., and Pyle, A.M. 2008. Crystal structure of a self-spliced group II intron. *Science* **320:** 77–82.

Tseng, C.K. and Cheng, S.C. 2008. Both catalytic steps of nuclear pre-mRNA splicing are reversible. *Science* **320:** 1782–1784.

Valadkhan, S. 2007. The spliceosome: A ribozyme at heart? *Biol. Chem.* **388:** 693–697.

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. 2003. Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA* **9:** 892–904.

Valadkhan, S., Mohammadi, A., Wachtel, C., and Manley, J.L. 2007. Protein-free spliceosomal snRNAs catalyze a reaction that resembles the first step of splicing. *RNA* **13:** 2300–2311.

Yang, K., Zhang, L., Xu, T., Heroux, A., and Zhao, R. 2008. Crystal structure of the beta-finger domain of Prp8 reveals analogy to ribosomal proteins. *Proc. Natl. Acad. Sci.* **105:** 13817–13822.