Cell, Vol. 109, 149–152, April 19, 2002, Copyright ©2002 by Cell Press

Spliceosomal snRNAs: Mg^{2+}-Dependent Chemistry at the Catalytic Core?

Minireview

Tommaso Villa, Jeffrey A. Pleiss, and Christine Guthrie[1]
Department of Biochemistry and Biophysics
University of California, San Francisco
San Francisco, California 94143

Since the discovery of self-splicing RNAs, it has been suspected that the snRNAs are the catalytic components of the spliceosome. Recent evidence supports both the catalytic potential of the spliceosomal snRNAs and their resemblance to elements of group II introns.

Thanks to the stunning realization that alternative mRNA splicing is a major source of genetic diversity (Black, 2000), a new and richly diversified contingent of biologists are being driven to understand the mechanism and regulation of this amazing class of RNA processing reactions. While our knowledge base has grown exponentially, the most basic question remains unanswered: what catalyzes splicing? The spliceosome, like its sister RNA-protein (RNP) machine the ribosome, is large (~40S) and complex (5 snRNAs and more than 70 proteins; Stevens et al., 2002). With the recent crystallographic analyses bolstering the conclusion that the ribosome is a ribozyme (the nearest protein is 18 Å from the active site), the odds seem good that the spliceosome's RNA components will also lie at the heart of the catalytic core (although at least one protein component, Prp8p/p220, is in close proximity to the active site; Collins and Guthrie, 2000). Here, we review recent evidence supporting this idea.

It was first shown in 1984 that mRNA splicing proceeds through a novel intermediate termed the "lariat," so named for the 2′-5′ branch formed via the nucleophilic attack by the 2′OH of an internal adenosine ("branchpoint A" or BP) on the phosphate at the 5′ exon/intron junction (5′ splice site or 5′SS; Figure 1; Burge et al., 1998). In the second chemical step, the 3′OH of the 5′ exon attacks the phosphate at the intron/3′ exon boundary (3′ splice site or 3′SS). The mechanistic basis of this bizarre chemical reaction remained mysterious until it was shown that group II self-splicing introns (predominantly found in organellar genes of bacteria, plants, and fungi) also employ a 2-step transesterification reaction with a lariat intermediate. The implication was immediate and profound: the spliceosomal mechanism must be fundamentally RNA-based. Several scenarios were proposed. In one extreme view, the *trans*-acting snRNAs are considered as "escaped domains" from the highly conserved secondary structure of group II introns. At the other extreme, commonalities between group II introns and the spliceosomal machinery merely reflect the constraints of RNA chemistry, i.e., convergent evolution (Weiner, 1993). The resolution of this debate will require a detailed description of the active sites and catalytic strategies of both systems.

Model for the Spliceosomal RNA Catalytic Core

A combination of phylogenetics, the genetics of compensatory base-pair mutations, and photochemical crosslinking in both mammals and budding yeast led to a unifying secondary structure model of the spliceosomal catalytic core dominated by interactions between the U2 and U6 snRNAs (Figure 2). Early in vitro experiments suggested that snRNPs assemble on the pre-mRNA in an ordered fashion (Burge et al., 1998). Interestingly, both U1 and U4 snRNPs are released or destabilized from the spliceosome prior to catalysis (Figure 1), and could thus be eliminated as candidates for catalytic core components. Furthermore, although base pairs can also form between the invariant loop of U5 snRNA and divergent sequences in the exons, the specific U5 loop sequences are themselves dispensable in vitro (thus, the essential role of the RNA may be to present Prp8/p220 to the catalytic core; Collins and Guthrie, 2000). In the U2/U6-based model of the catalytic core, the phylogenetically conserved base-pairing network comprises one intramolecular helix (U6:U6) and intermolecular helices between U6 and the 5′ splice site (U6:5′SS), U2 and the branchpoint region (U2:BP), and U2 and U6. For simplicity, in Figure 2 we depict only a single U2:U6 helix (Helix I; Madhani and Guthrie, 1992), though additional (or alternative) U2:U6 interactions have been demonstrated in the mammalian spliceosome (Burge et al., 1998; Sun and Manley, 1995).

At the secondary structure level, the U2:U6 helix would allow the 2′OH of the branchpoint A to be juxtaposed to the 5′ splice site G for the first phosphoryl transfer reaction. Notably, the regions of U2 and U6 in Figure 2 are highly conserved (especially the ACAGAG and AGC sequences of U6). While such evolutionary conservation is consistent with an essential function, the rules of phylogenetic sequence comparisons require sequence covariations as "proof" of a helix. The surprising discovery of the "minor" spliceosome several years ago was the example that proved the rule: novel U2 (U12) and U6 (U6atac) snRNAs diverge in primary sequence, yet maintain each of the 4 helical elements shown in Figure 2. Finally, the proposed secondary structure could be folded into a more compact form by a putative long-range interaction (identified by genetic selection in budding yeast) between the terminal G of the ACAGAG box and a bulged residue in U2 (Figure 2; cited in Collins and Guthrie, 2000). We will come back to the significance of this interaction in the context of the mammalian catalytic core.

Metal Ion Catalysis in the Spliceosome

A key issue in evaluating the spliceosome as a potential ribozyme is the role of metal ions. While it is well known that divalent cations are crucial for the correct folding of RNA (e.g., by neutralizing negative charge in closely apposed RNA helices), it now appears that many ribozymes are metallo-enzymes, as are many protein phosphoryl transferases (Narlikar and Herschlag, 1997). A powerful technique for identifying potentially catalytic

[1]Correspondence: guthrie@cgl.ucsf.edu

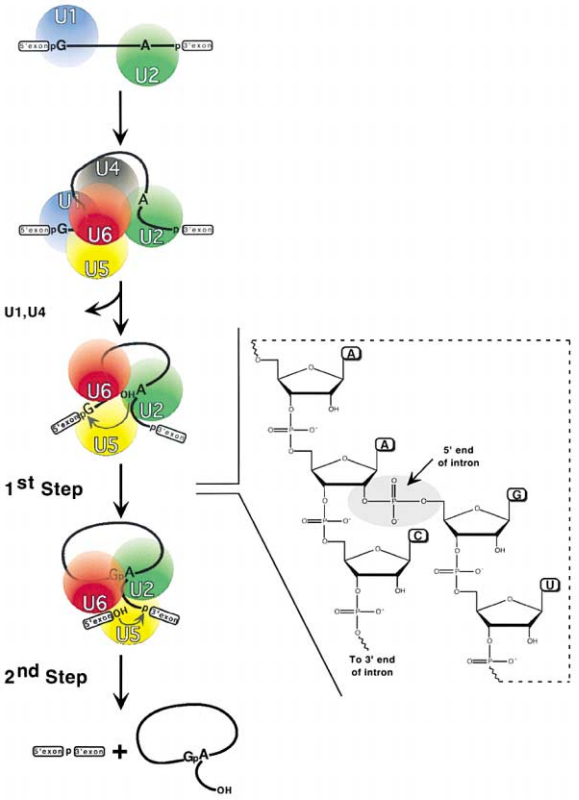

Figure 1. The pre-mRNA Splicing Pathway

Highlighted is the chemical structure of the 2′-5′ branched lariat intermediate.

metal ion binding sites are so-called "metal specificity switch" experiments in which one or more phosphate oxygens in the RNA are substituted with sulfur (termed a phosphorothioate when the substitution is for a non-bridging oxygen, and a phosphorothiolate when for a bridging oxygen). Because such sulfur-substituted phosphates are less efficient at coordinating Mg^{2+}, activity is inhibited upon substitution at critical metal binding sites. Addition of Mn^{2+} or other "soft" metals which coordinate to sulfur better than to oxygen may then rescue catalytic activity. While failure to rescue cannot be reliably interpreted, a successful specificity switch is taken as strong evidence of a catalytic role for the metal ion at that particular site. Although a critical structural role cannot be formally excluded, the binding sites for structural ions are generally thought to be less sensitive to substitution than are those for catalytic ions.

Application of the metal specificity switch approach to the pre-mRNA substrate has provided evidence that metal ions stabilize the leaving groups during both steps of splicing (Gordon et al., 2000, and references therein). However, application of this approach to the spliceosome per se has been hampered by the requirement that sulfur-substituted snRNAs must then be reconstituted into functional snRNPs in order for their activity to be evaluated. In a significant advance, Lin and coworkers (Yean et al., 2000) have now employed chemical synthesis to introduce sulfur subsitutions at either of the two non-bridging oxygens at specific positions in the U6 snRNA, followed by in vitro reconstitution of the U6 snRNP. Substitution of either phosphorothioate 5′ of the highly conserved U_{80} residue in the U6 intramolecular helix (Figure 2) impaired splicing in the presence of Mg^{2+}. Activity was rescued by addition of the more thiophilic cations Cd^{2+} or Mn^{2+} (although with only one of the two U6/sU_{80} phosphorothioates, validating the vesatility of chemical versus enzymatic RNA synthesis). The role of the coordinated Mg^{2+} is very likely catalytic, since spliceosome assembly is not affected by sulfur substitution. In sum, these results lend strong support to an RNA-based catalytic mechanism.

Group II versus snRNA Structure: Ancestry or Convergent Evolution?

If group II introns and spliceosomal snRNAs share a common origin, it is predicted that this relationship will be reflected in the conservation of functionally important structural features. In both the spliceosome and group II introns, the branchpoint A is bulged from an RNA helix (Burge et al., 1998). In the former case, the helix (U2:BP, Figure 2) is intermolecular. In group II, the BP resides in an intramolecular helix, domain VI. Interestingly, recent crystallographic data indicate that not only the BP but also an adjacent nucleotide are bulged from domain VI (Zhang and Doudna, 2002; Figure 2). It remains to be seen whether this 2-base bulge is unique to self-splicing

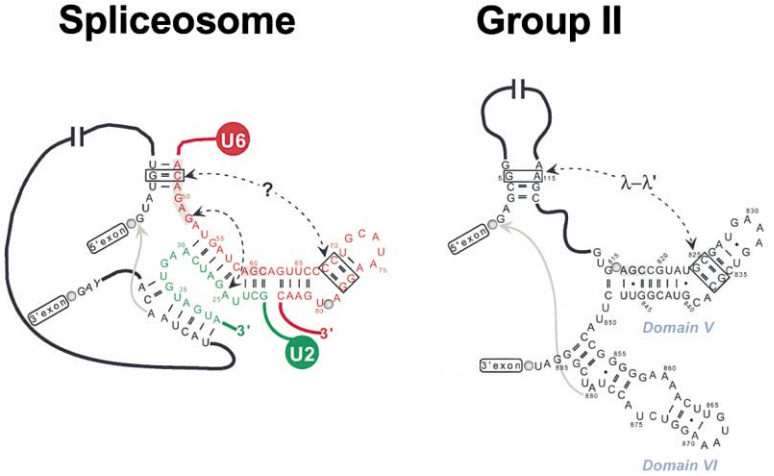

Figure 2. Proposed Secondary Structure Model of the Spliceosomal Catalytic Core Compared to Domains V and VI of Self-Splicing Group II Introns

Phylogenetically conserved sequence elements are highlighted. Dashed arrows indicate proven or proposed (Boudvillain et al., 2000) long-range tertiary interactions. Gray spheres represent critical Mg^{2+} ion binding sites (see text and Gordon and Piccirilli, 2001). Introns are represented by thick black lines. Y: pyrimidine. Numbering for the spliceosome and group II introns is from *S. cerevisiae* snRNAs and the mitochondrial ai5γ intron, respectively. Alternative interactions can be drawn for the AGC triad in the mammalian spliceosome (Sun and Manley, 1995); the conformation depicted is from *S. cerevisiae*.

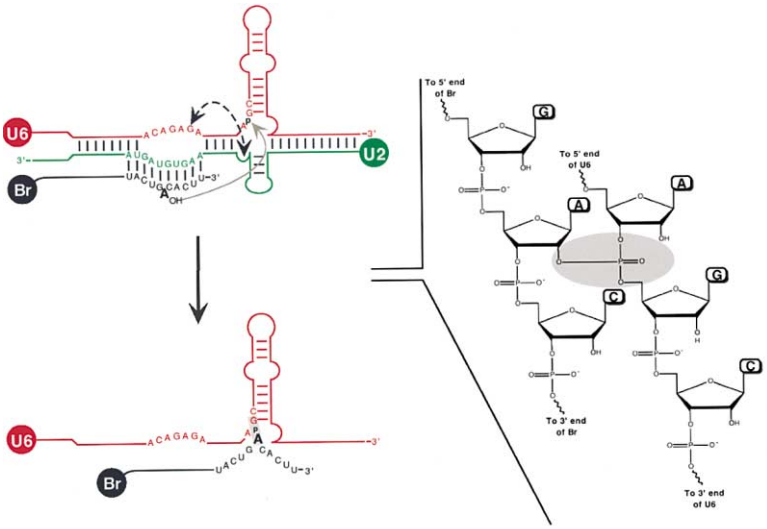

Figure 3. Proposed Chemistry of RNA X

Schematic representation of the human U2/U6 complex and of its interaction with the Br oligonucleotide. The dashed arrow indicates a tertiary interaction captured by UV-crosslinking (see text). Attack of the 2′OH of the branched adenosine on the phosphodiester bond between the A and G residues in the U6 AGC triad is shown by a gray arrow. The proposed chemical structure of the phosphotriester bond in RNA X is highlighted (Valadkhan and Manley, 2001).

introns. Available biochemical and phylogenetic data for the spliceosome do not suggest such a feature, but comparable structural information is conspicuously lacking.

A more extensive analogy can be drawn between domain V of group II introns and the region of the catalytic core comprising the U2:U6 and intramolecular U6 helices. Domain V is the single most highly conserved secondary structure element in group II phylogeny; it is an irregular helix, due to an internal bulge, and it contains an invariant AGC trinucleotide near the bottom of the helix. Curiously, the requirement for the AGC sequence is greater than for its pairing partner. U6 also has an invariant AGC, which participates in the U2:U6 helix, and genetic analysis in yeast revealed a similar asymmetry: the complementary sequence in the U2 pairing partner is markedly more tolerant to mutation than is the U6 trinucleotide (Madhani and Guthrie, 1992). Moreover, the U6 intramolecular helix, which likely stacks on this U2:U6 helix, itself has a highly conserved structure, including a bulge on the 3′ side of the helix.

So what is the specific function of domain V in group II introns? Notably, in isolation it can activate the first chemical step in *trans* with domain I, which in turn forms hydrogen bonds with the 5′ slice site. Importantly, sulfur substitution of one of the non-bridging oxygens 5′ of the adenosine of AGC is inhibitory, but can be rescued by Cd^{2+} (Gordon and Piccirilli, 2001). This provides direct evidence for a coordinated metal ion in a group II intron. Moreover, the internal bulge in domain V is sensitive to both phosphorothioate substitutions, although, unlike the case described above for the U6 bulge, no metal rescue has been observed at this position. However, the bulge was independently identified as a metal ion binding site by monitoring Tb^{3+}-induced cleavage (Sigel et al., 2000). Thus, domain V can be thought of as a "metal ion binding platform" which can be delivered to domain I in *trans* to activate catalysis.

Finally, using the elegant chemico-genetic technique of nucleotide analog interference mapping (NAIM) and suppression (NAIS) first developed by S. Strobel, A.M. Pyle and coworkers (Boudvillain et al., 2000, and references therein) recently identified a conserved tertiary interaction between the upper stem of domain V and the domain I:5′SS helix (λ-λ′; Figure 2). The authors point out a potential parallel in the spliceosome; namely, the fifth nucleotide of the 5′SS is an invariant G and forms a Watson-Crick base pair with the ACAGAG box of U6. A strong prediction of this analogy is the existence of a tertiary interaction between the C:G pairs in the upper stem of U6 and the U6:5′SS helix. Such a demonstration would provide powerful ammunition for the controversial argument that the spliceosome is an evolutionary descendent of group II introns.

Chemistry at the "Catalytic Core" of the Spliceosome

Ancestral relationships notwithstanding, the operative question remains: what mediates catalysis by the contemporary spliceosome?

Early in the game, a number of devout "RNA World" believers tried the obvious experiment of mixing all five, protein-free snRNAs together in a test tube and looking for splicing of a synthetic transcript. In retrospect, this approach was doomed if for no other reason than that the spliceosome is highly dynamic and we now know that the snRNAs undergo profound conformational rearrangements between their so-called ground state and catalytically active forms (see, e.g., Madhani and Guthrie, 1992). Now, Manley and coworkers have devised a logical extension of the RNA-only approach by exclusively using selected regions of human U2 and U6 snRNAs (including those depicted in Figure 2 as well as flanking sequences which form U2:U6 helices II and III; Burge et al., 1998). His group (Valadkhan and Manley, 2001, and references therein) first demonstrated that these pieces of RNA could form a stable complex in the presence of Mg^{2+}, and secondary structure mapping confirmed the expected inter- and intramolecular helices. Notably, the U2 and U6 pieces derived from human sequences could also be efficiently UV crosslinked at the corresponding sites of tertiary interaction proposed previously based on genetic selection in yeast (see Figures 2 and 3).

Emboldened by the implication that these short syn-

thetic RNAs were forming a structurally interesting and perhaps biologically significant tertiary fold, in subsequent experiments Valadhkan and Manley included short oligonucleotides containing 5′SS or BP sequences. Remarkably, incubation with the BP-containing oligo (termed Br in their system; Figure 3) for up to 24 hr at room temperature resulted in the Mg^{2+}-dependent formation of a novel RNA product with slow mobility on high percentage polyacrylamide gels, a hallmark of spliceosomal lariat-containing intermediates and products. However, in contrast to the "Y" structure of lariats, the 5′ and 3′ ends of both U2 and U6 snRNAs are free, suggesting that this novel RNA may be X shaped. Using a combination of alkaline and enzymatic (RNases T1, T2, and P1) digestion, and iodoethanol cleavage of randomly phosphorothioate-substituted RNAs, the authors conclude that the structure of RNA X is a phospho*triester*, with the linkage joining the 2′ oxygen of the branchpoint A in the Br oligo to the phosphate between A_{53} and G_{54} in the AGC triad of U6 (Figure 3). Acknowledging that formation of such an adduct without a leaving group would be entropically unfavorable, the authors propose that a small leaving group such as H_2O is consistent with their proposal. Finally, they demonstrate that the efficiency of the Mg^{2+}-dependent reaction is influenced by parameters similar to those seen in the true splicing reaction. For example, the attacking nucleophile shows a preference for a bulged A and an intact ACAGAG box is required. The authors conclude that their experiments "provide direct evidence for RNA-based catalysis in the spliceosome."

Taken at face value, the paper clearly demonstrates a novel chemical reaction dependent on regions of U2 and U6 which are themselves critical for catalysis in the spliceosome. Notably, the canonical lariat mechanism was in fact precluded because no 5′SS oligo was present (this oligo did not promote RNA X formation). Nevertheless, as the authors point out, attacks on the backbone of U6 as a result of abnormal splicing reactions have been documented biochemically (Yu et al., 1993), potentially explaining the presence of an intron in a fungal U6 gene immediately adjacent to the AGC triad of U6 (Tani and Oshima, 1991, and references therein). Together with the requirement of the reaction for an extrahelical A and the ACAGAG sequence, as well as the specific dependence on Mg^{2+} as the divalent cation, these characteristics make a plausible case that this reaction is using the same catalytic core elements as the spliceosome.

While these intriguing data are certainly consistent with the notion that the sequences in the U2/U6 core can promote catalysis in the spliceosome, many questions remain to be answered in order to prove a direct link between the observed chemistry and catalytic activity relevant to the spliceosome. Particularly because the proposed chemical mechanism is unprecedented in biology, an authentic standard containing an RNA phosphotriester should be chemically synthesized and characterized for its chemical properties. While such analyses have been undertaken for certain organic phosphotriesters, the relevance of these data to the stability of a phosphotriester like that proposed in RNA X remains uncertain. Additionally, RNA X should be purified and subjected to higher-resolution analyses, preferably by NMR but at least using mass spectrometry, in order to confirm its chemical structure. Because of the low level of product formation (about 0.1%), these analyses would also be instrumental in formally ruling out the unlikely possibility that RNA X is a biologically irrelevant artifact resulting from modification or contamination of the active subpopulation. These chemical details notwithstanding, perhaps the most persuasive argument for RNA-based catalysis at the core of the spliceosome would come from the development of a 5′SS oligo-dependent reaction which supports legitimate branch formation.

Conclusions and Perspectives

Previous arguments for mechanistic parallels between self-splicing group II introns and the spliceosome—including lariat formation via transesterification and similar stereospecificity at the splice sites—have been considerably strengthened by recent evidence for catalytic metal ions within the most highly conserved domains of RNA. While it thus seems likely that both machines operate as metalloenzymes, proposed relationships between the two types of catalytic cores, while intriguing, await critical tests. The publication of a 3 Å crystal structure containing two catalytically important domains of a group II intron is thus an exciting and important advance. Moreover, recent progress in cryo-EM imaging of the spliceosome bodes well for the ultimate crystallographic determination of RNA structure in the context of the RNP.

Selected Reading

Black, D.L. (2000). Cell *108*, 367–370.

Boudvillain, M., de Lencastre, A., and Pyle, A.M. (2000). Nature *406*, 315–318.

Burge, C.B., Tuschl, T.H., and Sharp, P.A. (1998). In The RNA World II, R.F. Gesteland, T.R. Cech, and J.F. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp 525–560.

Collins, C.A., and Guthrie, C. (2000). Nat. Struct. Biol. *7*, 850–854.

Gordon, P.M., and Piccirilli, J.A. (2001). Nat. Struct. Biol. *8*, 893–898.

Gordon, P.M., Sontheimer, E.J., and Piccirilli, J.A. (2000). RNA *6*, 199–205.

Madhani, H.D., and Guthrie, C. (1992). Cell *71*, 803–817.

Narlikar, G.J., and Herschlag, D. (1997). Annu. Rev. Biochem. *66*, 19–59.

Sigel, R.K.O., Vaidya, A., and Pyle, A.M. (2000). Nat. Struct. Biol. *7*, 1111–1116.

Stevens, S.W., Ryan, D.E., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D., and Abelson, J. (2002). Mol. Cell *9*, 31–44.

Sun, J.S., and Manley, J.L. (1995). Genes Dev. *9*, 843–854.

Tani, T., and Oshima, Y. (1991). Genes Dev. *5*, 1022–1031.

Valadkhan, S., and Manley, J.L. (2001). Nature *413*, 701–707.

Weiner, A.M. (1993). Cell *72*, 161–164.

Yean, S.-L., Wuenschell, G., Termini, J., and Lin, R.-J. (2000). Nature *408*, 881–884.

Yu, Y.-T., Maroney, P.A., and Nilsen, T.W. (1993). Cell *75*, 1049–1059.

Zhang, L., and Doudna, J.A. (2002). Science *295*, 2084–2088.