ELSEVIER

snRNAs as the catalysts of pre-mRNA splicing

Saba Valadkhan

The spliceosome, the gigantic molecular machine that performs pre-mRNA splicing in eukaryotes, contains over 200 different proteins and five RNA molecules. The central role played by the spliceosomal RNAs in splicing has led to the hypothesis that, like the ribosome, the spliceosome is an RNA-centric enzyme and a relic from the RNA world. Recent structural studies have provided the first glimpses of the structural features of spliceosomal RNAs, and mutational analyses *in vivo* and *in vitro* have uncovered new functional roles for a catalytically essential domain. An emerging model for the active site of group II introns, a closely related class of natural ribozymes, is likely to provide a wealth of insights on structure and function of the active site of the spliceosome.

Addresses
Center for RNA Molecular Biology, Case Western Reserve University School of Medicine, Cleveland, Ohio, 44106, USA

Corresponding author: Valadkhan, Saba (saba.valadkhan@case.edu)

Current Opinion in Chemical Biology 2005, **9**:603–608

This review comes from a themed issue on
Biopolymers
Edited by Scott A Strobel and Tom W Muir

Available online 20th October 2005

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2005.10.008

Introduction

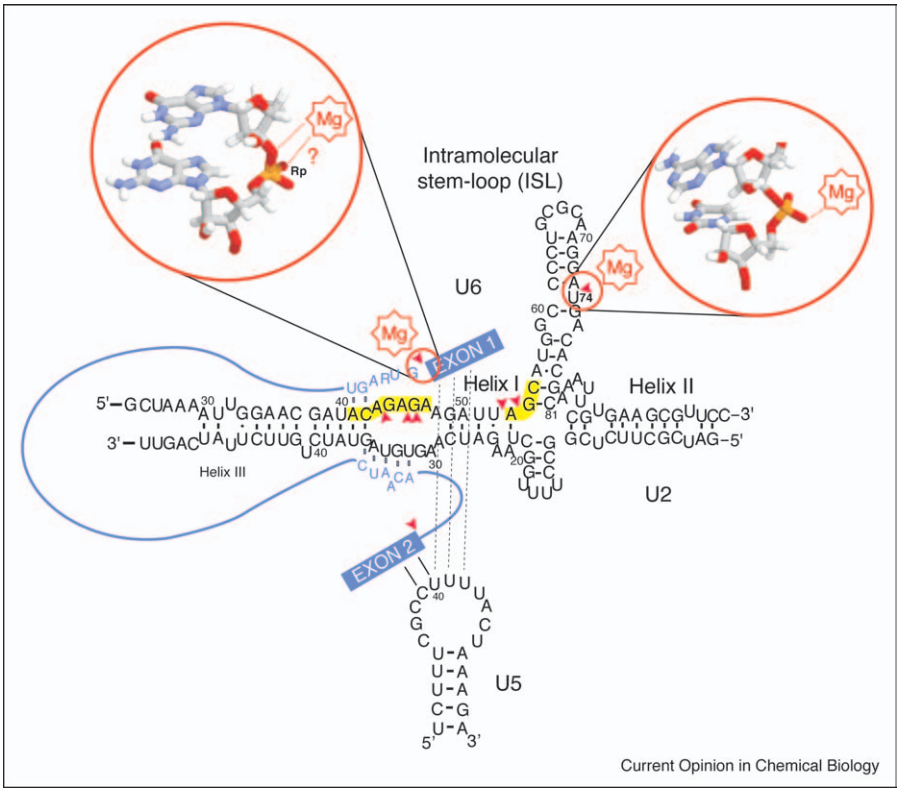
Removal of the intervening sequences (introns) from pre-mRNAs is a crucial and ubiquitous step in eukaryotic gene expression. Almost all primary transcripts undergo multiple splicing events, and alternative splicing and its regulation play a central role in cellular homeostasis. Commensurate with its critical role, the spliceosome, the gigantic ribonucleoprotein assembly that performs splicing, is the largest and most complicated cellular machine known. The spliceosome is composed of over 200 different proteins and five RNA components (U1, U2, U4, U5 and U6 small nuclear RNAs, or snRNAs) that form a dynamic and elaborate network of RNA–RNA, RNA–protein and protein–protein interactions [1]. The snRNAs are highly conserved and play a central role in all aspects of the splicing reaction. The 5′ splice site and branch site are recognized partially via base-pairing interactions with U1 and U2 snRNAs, respectively. Subsequently, U6 snRNA replaces U1 at the 5′ splice site, and forms a base-pairing interaction with U2 snRNA that juxtaposes the 5′ splice site and branch site, the reactants of the first of the two transesterification reactions of splicing. Finally, the two exons are bound and kept in alignment partially via interactions with a conserved loop in U5 snRNA [2,3] (Figure 1).

In addition to binding and coordinating the splice sites and the branch site, the snRNAs are implicated in catalysis of the splicing reactions [4,5]. Mutagenesis data have defined two catalytically crucial domains in U6 — the evolutionarily invariant ACAGAGA box and the AGC triad (Figure 1) — where mutations lead to a block to the first or second step of splicing. In addition, phosphorothioate backbone substitutions in several positions in U6 are incompatible with splicing, and at least in one case (U6 position 74, Figure 1), the substituted phosphate was shown to be involved in a direct interaction with a functionally important divalent cation [4,6]. Interestingly, an NMR study on the intramolecular stemloop (ISL) of U6 provided evidence for the intrinsic ability of U6 ISL to bind a divalent metal ion at this position [7]. Taken together, the above data clearly indicated that protein-free snRNAs were capable of binding functionally important metals that could be potentially used in catalysis of splicing reactions, thus, strengthening the case for RNA catalysis in the spliceosome. This review focuses on progress over the past two years in understanding the structure and function of spliceosomal snRNAs and their possible catalytic role in the spliceosome.

Role of metal ions in splicing catalysis

As mentioned above, it has been shown that a conserved stemloop in U6 snRNA binds a functionally required divalent cation. Interestingly, there is strong evidence for the involvement of metal ions in splicing catalysis from phosphorothioate substitution studies in pre-mRNA, where thio substitution of either the Rp or the 3′-linked oxygen at the 5′ splice site resulted in a block to splicing [8,9]. Intriguingly, splicing of the 3′-linked thio-substituted pre-mRNA could be rescued by the addition of thiophilic metal ions, proving direct coordination of the metal to this oxygen, which is the leaving group in the first step of splicing [9] (Figure 1). The coordinated metal can assist the departure of the leaving group by stabilizing the negative charge developing on the leaving oxyanion, consistent with the proposed catalytic strategy for ribozyme-catalyzed phosphoryl transfer reactions [10–12]. The other functional groups involved in the positioning of this metal ion are not known, but crosslinking studies have provided evidence for the proximity of the ACAGAGA catalytic domain to the 5′ splice site. Interestingly, thio substitutions at several positions in ACAGAGA result

Figure 1



RNA–RNA interactions in the spliceosome catalytic core. The base-pairing interactions between U6, U2 and U5 snRNAs and the pre-mRNA are shown. The location of U6–U2 helices I, II and III and the U6 intramolecular stemloop (ISL) are indicated. The invariant ACAGAGA and AGC catalytic domains in U6 are highlighted in yellow. Red arrowheads mark the sites of backbone phosphorothioate interference. Insets show the known metal binding sites in U6 and pre-mRNA in greater detail.

in a block to splicing (Figure 1). One can speculate that in the folded structure of U6, the metal binding domain of the ISL might also be positioned close to the splice sites and contribute to formation of the active site (see below).

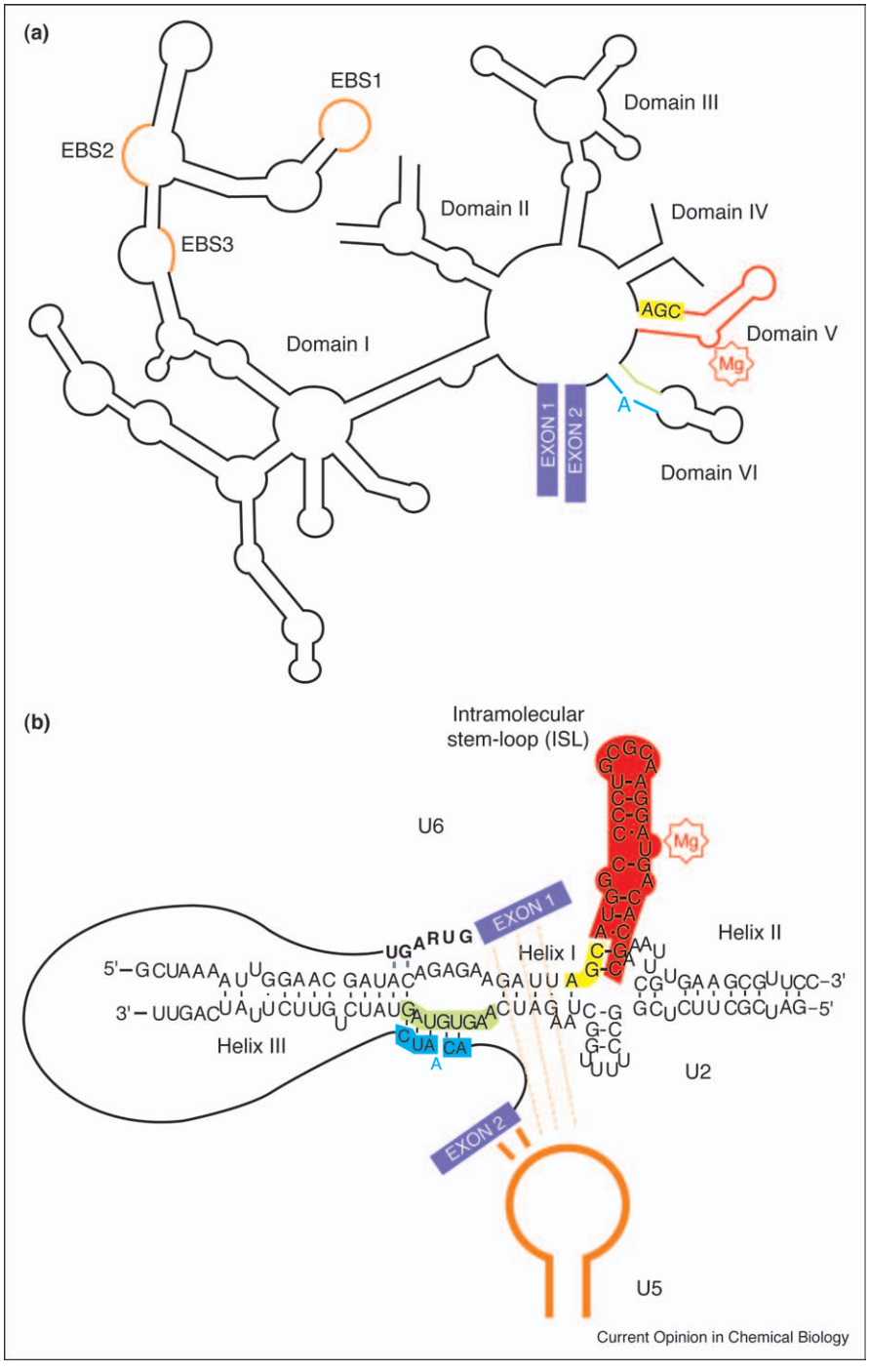
The case for RNA-based catalysis in the spliceosome is further strengthened by the parallels between the spliceosome and group II introns, self-splicing ribozymes that catalyze a reaction mechanistically identical to that catalyzed by the spliceosome [13] (Figure 2). Intriguing structural and functional similarities between the U6 ISL and the catalytically crucial domain V of group II introns makes it likely that the active site and the catalytic strategies used in the two systems are closely related. It is thought that group II-like introns were the evolutionary ancestors of the modern spliceosome with some of their RNA elements replaced by proteins during the course of evolution.

Whether or not any of the protein factors in the spliceosome participate in catalysis remains unknown. The most likely candidate, Prp8/p220, is the most conserved protein in the spliceosome and crosslinking and mutagenesis data have shown that it comes in contact with both 5′ and 3′ splice sites and the branch site [14]. Current data point to a dual fitter/coordinator role for Prp8/p220 in the spliceosome. On the one hand, evidence suggests that Prp8 provides a structural scaffold for exact positioning of various elements of the active site. On the other hand, through its extensive interactions with the other spliceosomal proteins, Prp8/p220 probably plays an important role in coordinating the progression of the catalytic cycle of the spliceosome. However, an additional catalytic role for Prp8/p220 has not yet been ruled out.

Solution structure of a fragment of yeast U6–U2 snRNA complex

Butcher and colleagues [15••,16–18] have determined the solution structure of a fragment of yeast U6–U2 complex that contains the U6 ISL and surrounding regions. Interestingly, they found that this region of the U6–U2 complex forms a four-way junction that bears similarities to the proposed structure for human U6–U2 complex [19] and domain V of group II introns (Figures 2 and 3). In this structure, the U6 ISL contains additional base-pairs at the base of the stemloop that incorporate the catalytic AGC triad into the ISL, in a position identical to the AGC triad in domain V. Additionally, the longer U6 ISL shows

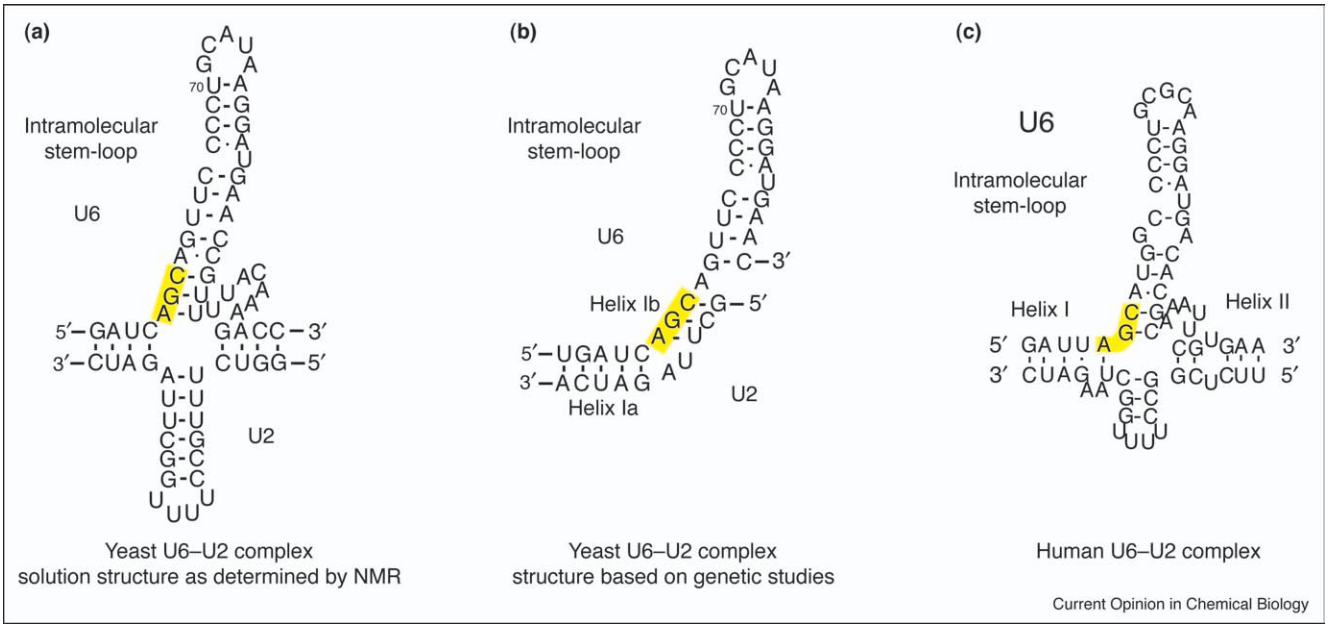
Figure 2



Structural parallels between **(a)** group II introns and **(b)** spliceosomal snRNAs. Homologous structures in the two systems are highlighted in the same color. Location of the AGC triad, branch site adenosine and Mg^{2+} binding site are indicated in both systems. Domain V and U6 ISL are in red, and the exon-binding sequences (EBS1, 2 and 3) and U5 are in orange. Exons are shown in magenta and branch site regions in the two systems are in blue and green.

strong similarities in size and the overall structure to the solution structure of domain V [20], however, there are important differences in the asymmetric bulge region between the two structures. Formation of this elongated ISL is mutually exclusive with the structure proposed on the basis of genetic studies in yeast [2], in which the AGC triad forms a short helix (helix Ib) with U2 (Figure 3). It is likely that in yeast spliceosome helix Ib is stabilized by

Figure 3



Comparison of the proposed structures for the U6 intramolecular stemloop (ISL). **(a)** Solution structure of a fragment of yeast U6–U2 complex as determined by NMR. **(b,c)** Structure of the central domain of U6–U2 complex in (b) yeast and (c) human as determined by genetic complementation studies *in vivo*.

proteins, because the autonomous folding of the RNA is incompatible with its formation. Butcher *et al.* proposed a model in which formation of helix Ib is driven by proteins during the structural rearrangements between the two steps of splicing. This model is appealing because it also reconciles the difference between the human and yeast U6–U2 structures deduced by genetic interactions. Intriguingly, computer modeling on a larger fragment of U6–U2 shows that the four-way junction folding has the potential to bring the metal-binding region of the U6 ISL (nucleotide U74) to the vicinity of 5′ splice site and the catalytic ACAGAGA domain. Thus, similar to the hairpin ribozyme, a four way junction in the U6–U2 complex might contribute to the formation of the active site of the splicing reaction [21,22].

The AGC catalytic triad revisited

A mutagenesis study *in vivo* by Staley and colleagues [23••] has provided further evidence for the catalytically crucial role of the AGC triad and the base-pairing interactions of this domain in the yeast spliceosome. Using a well-characterized genetic assay in yeast, the authors incorporated compensatory mutations that would be compatible with formation of either helix Ib or incorporation of the AGC into the base of an extended U6 ISL (Figure 3). Mutation of the A residue to C or G could be efficiently suppressed by a complementary mutation in U2 that would restore the formation of helix Ib, whereas complementary mutations that would allow the formation of an extended U6 ISL were not viable. Likewise, mutations in the G were not suppressed by compensatory mutations that restored the formation of an extended U6 ISL, but a G to A mutation in the AGC triad could be suppressed with a C to U mutation that restored helix Ib. Other point mutations in G could not be suppressed by compensatory mutations. This result is in agreement with previous mutagenesis data, which showed that the G to A mutations in the AGC triad generally have milder effects than other mutations at this position [24,25]. Mutations of the C of the AGC triad were mostly well tolerated. Whether suppression of the lethal AGC mutant phenotypes resulted from an enhancing effect on the first or the second step of splicing remains to be determined. However, the extreme sensitivity of the G in the AGC triad to mutations is similar to the observed effect in group II introns [26], consistent with a catalytic role for this nucleotide.

A Minimal U6/U2 splicing system

Direct probing of the catalytic potential of protein-free spliceosomal snRNAs has shown that snRNAs do indeed possess intrinsic splicing-related catalytic activity. An *in vitro*-assembled, protein-free human U6–U2 complex was incubated with a short RNA oligonucleotide that contained the branch site consensus sequence [27,28•]. The 5′ splice site substrate was absent in this minimal system. Intriguingly, the U6–U2 complex could catalyze a splicing-related reaction that resulted in a covalent linkage between the branch site adenosine, the reacting group of the first step of splicing, and the AGC triad. The reaction

was dependent on the presence of Mg^{2+} in the buffer, the branch site consensus sequence and the ACAGAGA and AGC catalytic domains. Although the chemistry of the reaction was different from the first step of splicing, the parallels between the two reactions were intriguing. Despite their small size, protein-free U6 and U2 could recognize and bind the branch site substrate, one of the two substrates of the first step of splicing. Additionally, they could form part of the active site of the first step of splicing by juxtaposing the catalytic AGC triad and the branch site adenosine and catalyzing a reaction between the two without the help of proteins. Thus, spliceosomal RNAs have intrinsic splicing-related catalytic activity, consistent with an RNA active site in the spliceosome. However, the minimal size of the snRNAs suggests that spliceosomal proteins play a significant role in modulating the kinetics of the reaction. In the most likely scenario, the snRNAs form the catalytic core of the spliceosome, with a myriad of protein factors fine-tuning the folding of the active site, the positioning of the substrates, and the progress of the splicing reaction from the first to the second step.

Supporting the above hypothesis, Query and Konarka [29•] have shown that the ability of Prp8 to suppress multiple mutations in both 5′ and 3′ splice sites and the branch site probably stems from its role as the gatekeeper at a fidelity checkpoint between the first and second step of splicing. In a genetic screen in yeast for suppressors of a branch site mutation, the authors isolated Prp8 alleles that not only could suppress branch site mutants, but could also suppress mutations at the 5′ and 3′ splice sites. A re-examination of previously isolated Prp8 suppressors of splice site mutations showed that those Prp8 alleles also shared the ability to suppress mutations in branch site and both splice sites via improving the efficiency of the second splicing step. These data suggest that the Prp8 suppressor alleles probably act by affecting the transition of the catalytic core between the two steps of splicing and not via direct participation in catalysis.

An active site model for group II introns

In a recent study, Pyle and co-workers [30] provided evidence for novel interactions at the active site of the group II introns by studying the effect of deletion of small fragments of domains II and III of a group II intron (Figure 2) on catalysis. The authors demonstrated a significant rate enhancement effect for domain III, and proved a novel role for domain II in the positioning of the 5′ splice site, domain III, and other active site elements, which is suggestive of a stepwise, hierarchical assembly pathway for the group II intron active site. Such ordered, intricate RNA–RNA interaction networks are reminiscent of similar interaction networks in the spliceosome. Interestingly, another interaction involving domain II is thought to play an important role in a conformational rearrangement between the two steps of splicing in group II introns, indicating yet another exciting new parallel between the spliceosome and group II introns.

Drawing on their recently published structural data [31••], de Lencastre *et al*. proposed a model for the active site of group II introns that also incorporates all previously characterized structural constraints [13]. Using a UV crosslinking approach, the authors demonstrated novel long-range interactions at the active site of a group II intron that juxtaposed the splice sites with the known catalytic domains. In addition, several other intronic elements were shown to be required for the architecture and function of the active site. Taken together, the wealth of available data on group II intron catalysis and the detailed model developed by de Lencastre *et al*. provide promising tools for guiding future mechanistic and structural studies on group II intron catalysis and elucidation of the functional relationship between the group II introns and the eukaryotic splicing machines.

Conclusions and perspective

The past two years have seen the spliceosome catalysis field revisiting old, unsolved problems and tackling challenging questions. Significant progress on determination of the structure of catalytic core of the spliceosome and group II introns has been made, and evidence for catalytic potential of spliceosomal RNAs is accumulating. Structural studies on folding of larger fragments of U6–U2 complex in human and yeast and further characterization of the catalytic potential of protein-free snRNAs promise to answer the question of the identity of spliceosomal active site. The role of the snRNAs in catalysis of the second step of splicing remains to be determined.

Update

Recent work by Nielsen and colleagues [32••] has shown that a group I-like ribozyme in the slime mold *Didymium Iridis* catalyzes the formation of a 2′-5′ linkage between the first and the third nucleotide of its mRNA. This reaction, which is coupled to the cleavage of this mRNA from the 3′ end of the ribozyme, resembles the first step of splicing in eukaryotic spliceosome and group II introns. Discovery of 2′-5′ branching activity in an evolutionarily unrelated natural ribozyme suggests that branching may be a common RNA-catalyzed reaction.

Acknowledgements

I thank Jim Manley and Tim Nilsen for discussions and comments on the manuscript, and Kinship foundation for a Searle Scholar award and the National Cancer Institute for funding.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Nilsen TW: **The spliceosome: the most complex macromolecular machine in the cell?** *BioEssays* 2003, **25**:1147-1149.

2. Nilsen TW: **RNA-RNA interactions in nuclear pre-mRNA splicing**. In *RNA Structure and Function.* Edited by Simons R, Grunberg-Manago M. Cold Spring Harbor Laboratory Press; 1998:279-307.

3. Burge CB, Tuschl T, Sharp PA: **Splicing of precursors to mRNAs by the spliceosomes**. In *The RNA World.* Edited by Gesteland RF, Cech TR, Atkins JF. Cold Spring Harbor Laboratory Press; 1999:525-560.

4. Valadkhan S, Manley JL: **Intrinsic metal binding by a spliceosomal RNA**. *Nat Struct Biol* 2002, **9**:498-499.

5. Collins CA, Guthrie C: **The question remains: is the spliceosome a ribozyme?** *Nat Struct Biol* 2000, **7**:850-854.

6. Yean SL, Wuenschell G, Termini J, Lin RJ: **Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome**. *Nature* 2000, **408**:881-884.

7. Huppler A, Nikstad LJ, Allmann AM, Brow DA, Butcher SE: **Metal binding and base ionization in the U6 RNA intramolecular stem-loop structure**. *Nat Struct Biol* 2002, **9**:431-435.

8. Maschhoff KL, Padgett RA: **Phosphorothioate substitution identifies phosphate groups important for pre-mRNA splicing**. *Nucleic Acids Res* 1992, **20**:1949-1957.

9. Sontheimer EJ, Sun S, Piccirilli JA: **Metal ion catalysis during splicing of premessenger RNA**. *Nature* 1997, **388**:801-805.

10. Steitz TA, Steitz JA: **A general two-metal-ion mechanism for catalytic RNA**. *Proc Natl Acad Sci USA* 1993, **90**:6498-6502.

11. Pyle AM: **Metal ions in the structure and function of RNA**. *J Biol Inorg Chem* 2002, **7**:679-690.

12. Lilley DM: **The origins of RNA catalysis in ribozymes**. *Trends Biochem Sci* 2003, **28**:495-501.

13. Lehmann K, Schmidt U: **Group II introns: structure and catalytic versatility of large natural ribozymes**. *Crit Rev Biochem Mol Biol* 2003, **38**:249-303.

14. Grainger RJ, Beggs JD: **Prp8 protein: at the heart of the spliceosome**. *RNA* 2005, **11**:533-557.

15. •• Sashital DG, Cornilescu G, McManus CJ, Brow DA, Butcher SE: **U2-U6 RNA folding reveals a group II intron-like domain and a four-helix junction**. *Nat Struct Mol Biol* 2004, **11**:1237-1242.

The solution structure of a fragment of yeast U6–U2 snRNAs was determined via NMR spectroscopy. Interestingly, the U6–U2 fragment folds into a four-way junction that has the potential to juxtapose catalytically crucial regions of the snRNAs.

16. Reiter NJ, Nikstad LJ, Allmann AM, Johnson RJ, Butcher SE: **Structure of the U6 RNA intramolecular stem-loop harboring an S(P)-phosphorothioate modification**. *RNA* 2003, **9**:533-542.

17. Sashital DG, Allmann AM, Van Doren SR, Butcher SE: **Structural basis for a lethal mutation in U6 RNA**. *Biochemistry* 2003, **42**:1470-1477.

18. Reiter NJ, Blad H, Abildgaard F, Butcher SE: **Dynamics in the U6 RNA intramolecular stem-loop: a base flipping conformational change**. *Biochemistry* 2004, **43**:13739-13747.

19. Sun JS, Manley JL: **A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing**. *Genes Dev* 1995, **9**:843-854.

20. Sigel RK, Sashital DG, Abramovitz DL, Palmer AG, Butcher SE, Pyle AM: **Solution structure of domain 5 of a group II intron ribozyme reveals a new RNA motif**. *Nat Struct Mol Biol* 2004, **11**:187-192.

21. Ferre-D'amare AR: **Rupert PB: The hairpin ribozyme: from crystal structure to function**. *Biochem Soc Trans* 2002, **30**:1105-1109.

22. Butcher SE, Brow DA: **Towards understanding the catalytic core structure of the spliceosome**. *Biochem Soc Trans* 2005, **33**:447-449.

23. •• Hilliker AK, Staley JP: **Multiple functions for the invariant AGC triad of U6 snRNA**. *RNA* 2004, **10**:921-928.

Using an *in vivo* mutagenesis-complementation approach in yeast, the paper provides new evidence for the catalytically crucial role of the G nucleotide in the catalytic AGC triad.

24. Datta B, Weiner AM: **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* 1993, **13**:5377-5382.

25. Fabrizio P, Abelson J: **Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing**. *Science* 1990, **250**:404-409.

26. Konforti BB, Abramovitz DL, Duarte CM, Karpeisky A, Beigelman L, Pyle AM: **Ribozyme catalysis from the major groove of group II intron domain 5**. *Mol Cell* 1998, **1**:433-441.

27. Valadkhan S, Manley JL: **Splicing-related catalysis by protein-free snRNAs**. *Nature* 2001, **413**:701-707.

28. • Valadkhan S, Manley JL: **Characterization of the catalytic activity of U2 and U6 snRNAs**. *RNA* 2003, **9**:892-904.

Using a minimal U6–U2 spliceosome, the paper provides evidence for the importance of the AGC sequence in a splicing-related catalytic activity observed in spliceosomal snRNAs.

29. • Query CC, Konarska MM: **Suppression of multiple substrate mutations by spliceosomal prp8 alleles suggests functional correlations with ribosomal ambiguity mutants**. *Mol Cell* 2004, **14**:343-354.

Genetic analysis of Prp8 mutant alleles in yeast provides novel explanation for the ability of these alleles to suppress a large number of mutations at the splice junctions and branch site.

30. Fedorova O, Mitros T, Pyle AM: **Domains 2 and 3 interact to form critical elements of the group II intron active site**. *J Mol Biol* 2003, **330**:197-209.

31. •• de Lencastre A, Hamill S, Pyle AM: **A single active-site region for a group II intron**. *Nat Struct Mol Biol* 2005, **12**:626-627.

Using site-specific UV crosslinking, the authors show evidence for new interactions in the catalytic core of the group II intron aI5γ, and provide a model for the organization of the active site of group II introns.

32. •• Nielsen H, Westhof E, Johansen S: **An mRNA is capped by a 2′, 5′ lariat catalyzed by a group I-like ribozyme**. *Science* 2005, **309**:1584-1587.

The authors show that a group I-like ribozyme removes an mRNA from its 3′ end by a branching reaction. The very small lariat formed probably serves as a cap substitute, protecting the 5′ end of this mRNA against nucleases.