Published in final edited form as:

Wiley Interdiscip Rev RNA. 2019 May ; 10(3): e1523. doi:10.1002/wrna.1523.

RNAs in the spliceosome: insight from cryoEM structures

Lingdi Zhang, Anne Vielle, Sara Espinosa, and Rui Zhao*

Department of Biochemistry and Molecular Genetics, University of Colorado Denver Anschutz Medical Campus

Graphical Abstract

Recent cryoEM structures offer new insight into the structure and function of spliceosomal RNAs

Overview

Pre-mRNA splicing is essential for gene expression in all eukaryotes. In mammals, an average of 95% of the nucleotides in the primary transcript (pre-mRNA) of a protein-encoding gene are introns¹. These introns need to be removed precisely by splicing before the mRNA can be transported out of the nucleus and translated. Alternative splicing greatly expands the gene coding capacity and is a fundamental component of eukaryotic gene regulation, influencing cell differentiation², development³, and many processes in the nervous system⁴. Over 90% of human genes are alternatively spliced⁵. Errors in splicing contribute to at least 30% of human genetic disorders⁶. Aberrant splicing also plays a significant role in the onset and development of many other diseases including cancer^{6,7}.

A typical intron contains a conserved 5' splice site (5' ss), a Branch Point Sequence (BPS) followed by a Polypyrimidine Tract (PYT), and a 3' splice site (3' ss)^{8,9}. Introns are removed through two transesterification reactions catalyzed by the spliceosome, a huge RNA-protein complex. The spliceosome contains five small nuclear RNAs (U1, U2, U4, U5, and U6 snRNAs) that form five snRNPs with their associated proteins, plus numerous other

*Corresponding author: rui.zhao@ucdenver.edu.

protein splicing factors (the total number of proteins in the spliceosome is well above 100) 10,11.

Spliceosome components assemble on the pre-mRNA in a step-wise manner (Fig. 1). The formation of the E complex involves the initial recognition of intron components by the spliceosome. In mammalian cells, the 5' ss is recognized by U1 snRNP, while the BPS, PYT, and 3' ss are recognized by SF1, U2AF65, and U2AF35, respectively 12–15. The *S. cerevisiae* (yeast) 5' ss, BPS, and the PYT are recognized by U1 snRNP, BBP, and Mud2, respectively 16–20, and there is no obvious spliceosomal component recognizing the 3' ss at this stage. How the splicing machinery accurately defines introns and exons in a large variety of pre-mRNAs is one of the most fundamental questions in pre-mRNA splicing. In lower eukaryotes whose pre-mRNAs typically contain small introns and large exons, the intron definition model (spliceosome assembly across an intron) seems to dominate 21. On the other hand, the exon definition model prevails in higher eukaryotes where small exons and large introns are prevalent 21. However, currently there is no direct structural evidence of either model and the molecular details governing these models are unclear.

Following initial intron recognition, the U2 snRNP replaces SF1 to form the A complex, assisted by RNA helicases Prp5 and Sub2 (UAP56 in higher eukaryotes) 22,23. The U4/U6.U5 tri-snRNP then joins the spliceosome to form the Pre-B complex 23. Helicase Prp28 disrupts the interaction between U1 snRNA and 5' ss, releasing the U1 snRNP and forming the B complex 24,25. Helicase Brr2 then unwinds the U4 and U6 snRNAs, releasing the U4 snRNP 26. The NineTeen Complex (NTD) and NTC-related complex (NTR) join the spliceosome at this stage, stabilizing the newly formed catalytic core and forming the B^{act} complex 27. Helicase Prp2 further rearranges the spliceosome to form the catalytically active B* complex that is ready for the first step reaction 28. Following the first transesterification reaction (also referred to as the branching reaction), the C complex emerges, encompassing the newly freed 5' exon and a lariat intermediate. The C complex further rearranges through the action of helicase Prp16 to form the C* complex, primed for the second transesterification reaction (also referred to as the ligation reaction) 29,30. In the ligation reaction, the two exons are joined together, forming the post-catalytic P complex that contains the ligated exons and the lariat. The exons are then released through the action of helicase Prp22 31, generating the intron-lariat spliceosome (ILS) complex, which only contains the lariat. ILS is further disassembled through the activity of the helicase Prp43 32,33.

The recent resolution revolution in cryoEM has brought about a slew of high-resolution structures of the spliceosome related complexes (including snRNPs), a previously unthinkable feat in the field (Table 1). The first high resolution cryoEM structure of the spliceosome was the *S. pombe* ILS complex determined by the Shi lab in 2015 34. In the following years, high resolution structures of the yeast U1 snRNP 35, tri-snRNP 36,37, A 38, pre-B 39, B 40, B^{act} 41, C 42,43, C* 44,45, P 46–48, and ILS 49 complexes, as well as the human B 50, B^{act} 51,52, C 53, and C* 54,55 complexes were determined. These structures provided unprecedented details on the organization of RNA and protein components in the spliceosome and their functions in the splicing reaction. Below we will highlight insight into

the structure and function of spliceosomal RNA components obtained from these cryoEM structures, with a focus on the yeast spliceosome.

The spliceosome is a ribozyme

Biochemical and genetic studies have long fueled the speculation that the spliceosome, like the ribosome, is a ribozyme (reviewed in ^{56,57}). In 1982, Cech and colleagues described the ribozyme as a complex RNA secondary structure that recognizes specific RNA sequences and catalyzes the splicing reaction within their active center ⁵⁸. Among the different classes of ribozymes, the group II introns catalyze RNA self-splicing through two sequential transesterification reactions, where the introns excise themselves from surrounding exons and stitch the resulting pieces back together ⁵⁹. When Sharp and others described the mRNA precursor splicing process through an intermediate lariat structure, the speculation that the spliceosome and group II introns shared a common evolutionary mechanism arose ^{60–63}. Since then, numerous biochemical and genetic studies have shed light on the possible catalytic roles of snRNAs in the splicing reaction. For example, the U2/U6 helix 1 and the U5 snRNA loop 1 configure the 5' splice site for cleavage and play a role in exon ligation ^{64–66} whereas U6 snRNA intra-stemloop (ISL) and AGC triad seem to be the principal actors in catalysis ^{62,67}. Five non-bridging oxygens in U6 (one each in A59, G60, G78 and two in U80) are metal ligands and may be involved in either branching, or exon ligation by stabilizing the leaving groups, or both ^{68–70}. Although such evidence supports the spliceosome being a ribozyme, the definitive proof of whether the spliceosome uses a RNA-based catalytic mechanism requires the visualization of the catalytic site of the spliceosome.

The recent high-resolution cryoEM structures of the spliceosome provided just such visualization, especially with the structures of the B^{act}, C, C*, and P complexes which capture the spliceosome before the first step reaction, after the first step reaction, before the second step reaction, and after the second step reaction, respectively. In the structure of the B^{act} and C* complexes, the branch point and the 3' ss are not yet loaded and cannot be observed structurally. On the other hand, RNA components of the branching and ligation reaction have not yet left the active site in the C and P complexes after the reaction, providing the closest snapshots of various players in the two reactions.

The structure of the C complex reveals density likely corresponding to the two Mg²⁺ ions required for catalysis, M1 and M2, which are coordinated by nucleotides in U6 snRNA (M1 by G78 and U80, M2 by A59 and U80) ^{42,43} (Fig. 2a). M2 is about 6 Å away from the lariat junction in the C complex (which may help prevent the reversal of the branching reaction) and it is likely much closer to the branch point right before the branching reaction (for example, in the B* complex whose structure is not yet known). Before the branching reaction, M2 can foreseeably activate the 2' OH group of the branch point A for nucleophilic attack of the phosphodiester bond in the 5' ss, while the M1 ion stabilizes the leaving group (Fig. 2a). The structure of the P complex reveals that the M1 ion is in a position to activate the 3' OH group of 5' ss, which attacks the 3' ss, ligates the two exons and releases the intron ^{46–48} (Fig. 2b). The M2 ion is not consistently observed in all three P complex structures, possibly reflecting the post-catalytic nature of the complex.

Although the spliceosome is clearly a ribozyme just like the ribosome, the spliceosome is much more protein rich compared to the ribosome. Of these proteins, Prp8 and Snu114 from the U5 snRNP, Cef1 and Syf2 from NTC, as well as Bud31, Cwc2, Cwc15, Ecm2, Prp45, and Prp46 from NTR form a protein core surrounding the active site that is fairly stable from the B^{act} to ILS complex (Fig. 3). Prp8, in particular, is in the heart of the spliceosome and makes extensive contacts with the snRNA and pre-mRNA. Prp8 is one of the largest (Prp8 from all species are over 2,000 amino acids in length) and most conserved protein in the nucleus, with 62% sequence identity between its yeast and human homologs. Previous biochemical and genetic experiments demonstrate that Prp8 is crosslinked to all key elements of the pre-mRNA involved in the splicing reaction and many Prp8 mutants either exacerbate or suppress pre-mRNA splicing mutants (reviewed in ⁷¹). Prp8 is composed of the N-terminal domain, the large domain (which can be further divided into the helix bundle, reverse transcriptase-like (RT), thumb/X, linker, and endonuclease-like (EN) domains), the RNase H domain, and the MPN domain ^{34,72}. The N-terminal, linker, and RT domains of Prp8 form a cavity that cradles the RNA active site. Multiple structural elements, including the alpha finger (residues 1570–1615), the switch loop, and the RNase H domain containing the beta finger change conformation dramatically in different complexes, often help to stabilize and guide the precise conformation of RNA in different stages of the splicing cycle (reviewed in ⁷³).

Spatial organizations of the snRNAs and Pre-mRNA through the splicing cycle

From the initial recognition of pre-mRNA by the U1 snRNA at the 5' ss (E complex), several dramatic changes involving the snRNA occur to achieve spliceosomal activation. These changes include the recognition of the BPS through basepairing with U2 snRNA in the A complex, the joining of the tri-snRNP to form the pre-B complex, the removal of U1 snRNP to form the B complex, and the removal of U4 snRNP to form the B^{act} complex. Once the spliceosome is activated (starting from the B^{act} complex), although the 3' end of U2 snRNA changes positions to a certain extent, the snRNA core surrounding the active site (U2/U6 helix 1 and 2, the U2 snRNA region that basepairs with the BPS, the U6 intra-stemloop (ISL) and ACAGAGA box, and U5 snRNP loop 1) remains largely unchanged through the rest of the splicing cycle (reviewed in ^{73–77}) (Fig. 4).

On the other hand, the pre-mRNA undergoes dramatic movements through the two transesterification reactions (Fig. 5). In the B^{act} complex, the branch point A is about 50 Å away from the 5' ss. The branch point A presumably moves close to the 5' ss in the B* complex and carries out the branching reaction. The intronic RNA downstream of the BPS remains in the active site in the C complex, which is moved away in the C* complex to make room for the incoming 3' ss. The 3' ss is then loaded into the active site for the ligation reaction, which remains in the active site after exon ligation in the post-catalytic P complex. Ligated exons are then released, forming the ILS.

The spliceosomal helicases are largely responsible for the movement of the pre-mRNA, either directly or indirectly. The pre-mRNA splicing process is remarkable in the large

number of DExD/H-box RNA helicases required in the splicing cycle (Fig. 1). Four of these RNA helicases (Prp5, Sub2, Prp28 and Brr2) are involved in spliceosomal assembly and activation in early stages of the splicing cycle, up to the B^{act} complex. Interestingly, the four helicases (Prp2, Prp16, Prp22, and Prp43) in the later stages of the splicing cycle all belong to the DEAH-box sub-family with their helicase motif II taking the form of DEAH and the four proteins share similar domain organizations^{33,78–80}. Prp2 is responsible for removing SF3b and exposing the BPS in preparation for the first catalytic reaction in the transition from the B^{act} to the B* complex²⁸. Prp16 vacates the space near the active site in the C complex so that the 3' ss can enter²⁹. Prp22 acts in the P complex to release ligated exons³¹, and Prp43 helps disassemble the ILS^{32,33}. These helicases are recruited to specific stages of the spliceosome and discarded after their dedicated actions. For example, Prp2 is recruited to B^{act}, and released from B^{act} after ATP hydrolysis, allowing Prp16 to bind. The action of Prp16 induces the dissociation of Prp16 and a number of step one factors, allowing Prp22 and step two factors to bind and form C*. Prp22 remains bound in the P complex, whose action releases ligated exons and Prp22 is released with the exon, allowing Prp43 to bind. These observations suggest that the helicases have to be specifically recruited to the right place at the right time and their activities tightly regulated. In addition, it is clear from the P complex structure that Prp22 binds to the single stranded 3' end of the 3' intron, and pulls it from the 3' to 5' direction. This observation is consistent with previous biochemical observations that Prp22 uses a winching mode to fulfill its function in kinetic proofreading⁸¹, and suggests that translocation on the RNA instead of unwinding by these helicases is important for their functions in splicing. It is interesting to note that three of the four helicases that are resolved in the cryoEM structures all bind to the periphery of the spliceosome and close to the 3' end of their potential target RNAs (Prp2 is close to the 3' end of intron⁴¹, Prp22 binds to the 3' end of 3' intron^{46–48}, and Prp43 is close to intron/U6 snRNA⁴⁹). It is possible that these helicases fulfill their function through a similar mechanism of action.

RNA secondary structures in pre-mRNA

One intriguing observation in the P complex structure is that the region between the BPS and 3' ss forms a stem-like secondary structure⁴⁶ (Fig. 6). This secondary structure could potentially serve to bring the 3' ss close to the 5' ss and BPS (which are in close proximity of each other after the branching reaction) to facilitate the recognition of the 3' ss and subsequent ligation reaction. Almost all the yeast introns with long BPS to 3' ss distances are predicted to form some secondary structures⁴⁶. The distance between the 5' ss and BPS is typically much longer. It would not be surprising if secondary structures form in this region as well to bring the BPS and 5' ss close together and facilitate the branching reaction.

The existence of secondary structures in introns of isolated genes that are important for splice site selection have previously been observed through biochemical and genetic experiments (reviewed in⁸²). For example, intramolecular secondary structures in specific yeast introns are essential for spliceosome assembly and 5' splice site selection^{83–85}. Studies performed on *UBC13* and *RPS20B* with long branch point (BP) to 3' ss distances demonstrated impaired splicing when the proposed secondary structures are disrupted^{86,87}. The structure of the P complex provided the first physical evidence of such secondary

structures and points to the potentially general role of these secondary structures in facilitating splice site selection. Multiple methods have been developed in recent years to probe the genome-wide secondary structures of RNAs⁸⁸, and these approaches may also be used to provide insight into the extent of secondary structures in all yeast introns and their functions. In addition to facilitating splicing by bringing important intron elements close together, secondary structures can also inhibit splicing by masking splice sites or blocking the binding of splicing factors⁸⁹. Not surprisingly, secondary structures have been found to regulate alternative splicing in higher eukaryotes⁹⁰ and can potentially be a point of modulation by small molecules in the future⁸⁹.

RNA interactions in splice site recognition

The 5' ss of pre-mRNA is largely recognized through Watson-Crick basepairing. It is initially recognized through basepairing with the 5' end of the U1 snRNA, as can be seen in the structure of the A complex³⁸ (Fig. 7a). During the spliceosomal activation process, U1 snRNP is removed, and the 5' ss forms Watson-Crick basepairs with U6 snRNA instead in the B^{act} complex⁴¹ (Fig. 7b). The BPS is initially recognized by the Branch Bridging Protein (BBP)¹³ and is then handed over to U2 snRNA which recognizes the BPS through Watson-Crick basepairing³⁸ (Fig. 7c).

The CryoEM structure of the post-catalytic P complex reveals that the 3' ss is mainly recognized through non-Watson-Crick basepairing between the 3' ss and the 5' ss as well as the BPS (Fig. 7d). The last nucleotide of the intron (G in the 3' ss) forms a non-Watson Crick basepair with the first nucleotide of the intron (G in the 5' ss), and the penultimate nucleotide of the intron (A upstream of the 3' ss) forms another non-Watson Crick basepairing with the branch point A. These nucleotides are surrounded mainly by the U6 snRNA, Prp8, and Prp18.

The above structural observation helps explain the strong preference for the AG dinucleotide at the end of introns. In yeast, the last two nucleotides of all introns are strictly AG⁹¹ and in humans, they are predominantly AG with some low frequency non-canonical 3' ss⁹². In the P complex structure, substituting A or G in the dinucleotide to the other three bases loses all or partial hydrogen bonds with the 5' ss G (barring potential conformational changes around these nucleotides) (Fig. 7e, f), potentially explaining the conservation of the AG dinucleotide at the 3' ss.

In summary, recent high resolution cryoEM structures revealed unambiguously that the spliceosome is a ribozyme. The U2, U5, and U6 snRNAs form a fairly constant active site core stabilized by proteins such as Prp8 once the spliceosome is activated. The pre-mRNA moves significantly to bring key elements to the active site for catalysis, and most of these movements are facilitated by RNA helicases. Both RNA structures and RNA-RNA interactions likely play important roles in splice site recognition. Although the core of the spliceosome generally has high resolutions in the cryoEM structures determined to date, resolutions of regions that are more flexible, like those in the periphery of the spliceosome, are typically much lower. For example, the helicases such as Prp2, Prp16, Prp22, and Prp43 that play important roles in RNA remodeling and spliceosome assembly/activation are

located in the periphery of the spliceosome. Prp16 was not observed in the C complex structure, and the other three helicases also have much lower resolution than the core of the spliceosome, hindering a detailed mechanistic understanding of their action. As sample preparation and calculation methods continue to improve, we will likely to gain more high-resolution information on the RNAs in the spliceosome and the proteins that interact with them.

Acknowledgements

We would like to acknowledge support from the NIH R01GM114178 and R01GM126157 (R. Z.). We thank Drs. Quentin Vicens, Brian Wimberly, and members of the Zhao lab for helpful discussions.

References

1. Mattick JS & Gagen MJ The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* 18, 1611–30 (2001). [PubMed: 11504843]
2. Venables JP Alternative splicing in the testes. *Curr Opin Genet Dev* 12, 615–19 (2002). [PubMed: 12240623]
3. Lopez AJ Alternative splicing of pre-mRNA: developmental consequences and mechanisms of regulation. *Annu Rev Genet* 32, 279–305 (1998). [PubMed: 9928482]
4. Grabowski PJ & Black DL Alternative RNA splicing in the nervous system. *Prog Neurobiol* 65, 289–308 (2001). [PubMed: 11473790]
5. Wang ET et al. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–6 (2008). [PubMed: 18978772]
6. Faustino NA & Cooper TA Pre-mRNA splicing and human disease. *Genes Dev* 17, 419–37 (2003). [PubMed: 12600935]
7. Philips AV & Cooper TA RNA processing and human disease. *Cell Mol Life Sci* 57, 235–49 (2000). [PubMed: 10766020]
8. Moore MJ, Query CC & Sharp PA Splicing of precursors to mRNA by the spliceosome in *The RNA World*. (eds. Gesteland R & Atkins J) 303–357 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1993).
9. Hastings ML & Krainer AR Pre-mRNA splicing in the new millennium. *Curr Opin Cell Biol* 13, 302–9 (2001). [PubMed: 11343900]
10. Jurica MS & Moore MJ Pre-mRNA splicing: awash in a sea of proteins. *Mol Cell* 12, 5–14 (2003). [PubMed: 12887888]
11. Nilsen TW The spliceosome: the most complex macromolecular machine in the cell? *Bioessays* 25, 1147–9 (2003). [PubMed: 14635248]
12. Black DL, Chabot B & Steitz JA U2 as well as U1 small nuclear ribonucleoproteins are involved in premessenger RNA splicing. *Cell* 42, 737–50 (1985). [PubMed: 2996775]
13. Berglund JA, Chua K, Abovich N, Reed R & Rosbash M The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* 89, 781–7 (1997). [PubMed: 9182766]
14. Ruskin B, Zamore PD & Green MR A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* 52, 207–19 (1988). [PubMed: 2963698]
15. Wu S, Romfo CM, Nilsen TW & Green MR Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* 402, 832–5 (1999). [PubMed: 10617206]
16. Ruby SW & Abelson J An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science* 242, 1028–35 (1988). [PubMed: 2973660]
17. Legrain P, Seraphin B & Rosbash M Early commitment of yeast pre-mRNA to the spliceosome pathway. *Mol Cell Biol* 8, 3755–60 (1988). [PubMed: 3065622]

18. Seraphin B, Kretzner L & Rosbash M A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO J* 7, 2533–8 (1988). [PubMed: 3056718]
19. Seraphin B & Rosbash M Mutational analysis of the interactions between U1 small nuclear RNA and pre-mRNA of yeast. *Gene* 82, 145–51 (1989). [PubMed: 2684770]
20. Abovich N & Rosbash M Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* 89, 403–12 (1997). [PubMed: 9150140]
21. De Conti L, Baralle M & Buratti E Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip Rev RNA* 4, 49–60 (2013). [PubMed: 23044818]
22. Parker R, Siliciano PG & Guthrie C Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* 49, 229–39 (1987). [PubMed: 3552247]
23. Konarska MM & Sharp PA Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* 49, 763–74 (1987). [PubMed: 2953438]
24. Staley JP & Guthrie C An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol Cell* 3, 55–64 (1999). [PubMed: 10024879]
25. Boesler C et al. A spliceosome intermediate with loosely associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. *Nat Commun* 7, 11997 (2016). [PubMed: 27377154]
26. Raghunathan PL & Guthrie C RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr Biol* 8, 847–55 (1998). [PubMed: 9705931]
27. Chan SP, Kao DI, Tsai WY & Cheng SC The Prp19p-associated complex in spliceosome activation. *Science* 302, 279–82 (2003). [PubMed: 12970570]
28. Lardelli RM, Thompson JX, Yates JR, 3rd & Stevens SW Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *Rna* 16, 516–28 (2010). [PubMed: 20089683]
29. Tseng CK, Liu HL & Cheng SC DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *RNA* 17, 145–54 (2011). [PubMed: 21098140]
30. Schwer B & Guthrie C A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *Embo J* 11, 5033–9 (1992). [PubMed: 1464325]
31. Schwer B & Gross CH Prp22, a DEXH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *Embo J* 17, 2086–94 (1998). [PubMed: 9524130]
32. Martin A, Schneider S & Schwer B Prp43 is an essential RNA-dependent ATPase required for release of lariat-intron from the spliceosome. *J Biol Chem* 277, 17743–50 (2002). [PubMed: 11886864]
33. Arenas JE & Abelson JN Prp43: An RNA helicase-like factor involved in spliceosome disassembly. *Proc Natl Acad Sci U S A* 94, 11798–802 (1997). [PubMed: 9342317]
34. Yan C et al. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* 349, 1182–91 (2015). [PubMed: 26292707]
35. Li X et al. CryoEM structure of *Saccharomyces cerevisiae* U1 snRNP offers insight into alternative splicing. *Nat Commun* 8, 1035 (2017). [PubMed: 29051543]
36. Wan R et al. The 3.8 Å structure of the U4/U6.U5 tri-snRNP: Insights into spliceosome assembly and catalysis. *Science* 351, 466–75 (2016). [PubMed: 26743623]
37. Nguyen TH et al. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* 530, 298–302 (2016). [PubMed: 26829225]
38. Plaschka C, Lin PC, Charenton C & Nagai K Prespliceosome structure provides insights into spliceosome assembly and regulation. *Nature* 559, 419–422 (2018). [PubMed: 29995849]
39. Bai R, Wan R, Yan C, Lei J & Shi Y Structures of the fully assembled *Saccharomyces cerevisiae* spliceosome before activation. *Science* (2018).
40. Plaschka C, Lin PC & Nagai K Structure of a pre-catalytic spliceosome. *Nature* 546, 617–621 (2017). [PubMed: 28530653]
41. Yan C, Wan R, Bai R, Huang G & Shi Y Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* 353, 904–11 (2016). [PubMed: 27445306]
42. Wan R, Yan C, Bai R, Huang G & Shi Y Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science* 353, 895–904 (2016). [PubMed: 27445308]

43. Galej WP et al. Cryo-EM structure of the spliceosome immediately after branching. *Nature* 537, 197–201 (2016). [PubMed: 27459055]
44. Yan C, Wan R, Bai R, Huang G & Shi Y Structure of a yeast step II catalytically activated spliceosome. *Science* 355, 149–155 (2017). [PubMed: 27980089]
45. Fica SM et al. Structure of a spliceosome remodelled for exon ligation. *Nature* (2017).
46. Liu S et al. Structure of the yeast spliceosomal postcatalytic P complex. *Science* (2017).
47. Bai R, Yan C, Wan R, Lei J & Shi Y Structure of the Post-catalytic Spliceosome from *Saccharomyces cerevisiae*. *Cell* 171, 1589–1598 e8 (2017). [PubMed: 29153833]
48. Wilkinson ME et al. Postcatalytic spliceosome structure reveals mechanism of 3'-splice site selection. *Science* 358, 1283–1288 (2017). [PubMed: 29146871]
49. Wan R, Yan C, Bai R, Lei J & Shi Y Structure of an Intron Lariat Spliceosome from *Saccharomyces cerevisiae*. *Cell* 171, 120–132 e12 (2017). [PubMed: 28919079]
50. Bertram K et al. Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. *Cell* 170, 701–713 e11 (2017). [PubMed: 28781166]
51. Zhang X et al. Structure of the human activated spliceosome in three conformational states. *Cell Res* (2018).
52. Haselbach D et al. Structure and Conformational Dynamics of the Human Spliceosomal B(act) Complex. *Cell* 172, 454–464 e11 (2018). [PubMed: 29361316]
53. Zhan X, Yan C, Zhang X, Lei J & Shi Y Structure of a human catalytic step I spliceosome. *Science* (2018).
54. Zhang X et al. An Atomic Structure of the Human Spliceosome. *Cell* 169, 918–929 e14 (2017). [PubMed: 28502770]
55. Bertram K et al. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* 542, 318–323 (2017). [PubMed: 28076346]
56. Collins CA & Guthrie C The question remains: is the spliceosome a ribozyme? *Nat Struct Biol* 7, 850–4 (2000). [PubMed: 11017191]
57. Valadkhan S The spliceosome: a ribozyme at heart? *Biol Chem* 388, 693–7 (2007). [PubMed: 17570821]
58. Kruger K et al. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. *Cell* 31, 147–57 (1982). [PubMed: 6297745]
59. Peebles CL et al. A self-splicing RNA excises an intron lariat. *Cell* 44, 213–23 (1986). [PubMed: 3510741]
60. Padgett RA, Konarska MM, Grabowski PJ, Hardy SF & Sharp PA Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* 225, 898–903 (1984). [PubMed: 6206566]
61. Cech TR The generality of self-splicing RNA: relationship to nuclear mRNA splicing. *Cell* 44, 207–10 (1986). [PubMed: 2417724]
62. Valadkhan S, Mohammadi A, Jaladat Y & Geisler S Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci U S A* 106, 11901–6 (2009). [PubMed: 19549866]
63. Butcher SE The spliceosome as ribozyme hypothesis takes a second step. *Proc Natl Acad Sci U S A* 106, 12211–2 (2009). [PubMed: 19622742]
64. Madhani HD & Guthrie C A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* 71, 803–17 (1992). [PubMed: 1423631]
65. O'Keefe RT & Newman AJ Functional analysis of the U5 snRNA loop 1 in the second catalytic step of yeast pre-mRNA splicing. *EMBO J* 17, 565–74 (1998). [PubMed: 9430647]
66. Mefford MA & Staley JP Evidence that U2/U6 helix I promotes both catalytic steps of pre-mRNA splicing and rearranges in between these steps. *RNA* 15, 1386–97 (2009). [PubMed: 19458033]
67. Fabrizio P & Abelson J Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science* 250, 404–9 (1990). [PubMed: 2145630]
68. Fabrizio P & Abelson J Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res* 20, 3659–64 (1992). [PubMed: 1641331]

69. Yean SL, Wuenschell G, Termini J & Lin RJ Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. *Nature* 408, 881–4 (2000). [PubMed: 11130730]
70. Fica SM et al. RNA catalyses nuclear pre-mRNA splicing. *Nature* 503, 229–34 (2013). [PubMed: 24196718]
71. Grainger RJ & Beggs JD Prp8 protein: at the heart of the spliceosome. *Rna* 11, 533–57 (2005). [PubMed: 15840809]
72. Galej WP, Oubridge C, Newman AJ & Nagai K Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature* 493, 638–43 (2013). [PubMed: 23354046]
73. Galej WP, Toor N, Newman AJ & Nagai K Molecular Mechanism and Evolution of Nuclear Pre-mRNA and Group II Intron Splicing: Insights from Cryo-Electron Microscopy Structures. *Chem Rev* 118, 4156–4176 (2018). [PubMed: 29377672]
74. Shi Y The Spliceosome: A Protein-Directed Metalloribozyme. *J Mol Biol* 429, 2640–2653 (2017). [PubMed: 28733144]
75. Shi Y Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat Rev Mol Cell Biol* (2017).
76. Scheres SH & Nagai K CryoEM structures of spliceosomal complexes reveal the molecular mechanism of pre-mRNA splicing. *Curr Opin Struct Biol* 46, 130–139 (2017). [PubMed: 28888105]
77. Fica SM & Nagai K Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. *Nat Struct Mol Biol* 24, 791–799 (2017). [PubMed: 28981077]
78. Chen JH & Lin RJ The yeast PRP2 protein, a putative RNA-dependent ATPase, shares extensive sequence homology with two other pre-mRNA splicing factors. *Nucleic Acids Res* 18, 6447 (1990). [PubMed: 2147058]
79. Burgess S, Couto JR & Guthrie C A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. *Cell* 60, 705–17 (1990). [PubMed: 2138057]
80. Company M, Arenas J & Abelson J Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* 349, 487–93 (1991). [PubMed: 1992352]
81. Semlow DR, Blanco MR, Walter NG & Staley JP Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. *Cell* 164, 985–98 (2016). [PubMed: 26919433]
82. Buratti E & Baralle FE Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* 24, 10505–14 (2004). [PubMed: 15572659]
83. Goguel V & Rosbash M Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. *Cell* 72, 893–901 (1993). [PubMed: 8458083]
84. Goguel V, Wang Y & Rosbash M Short artificial hairpins sequester splicing signals and inhibit yeast pre-mRNA splicing. *Mol Cell Biol* 13, 6841–8 (1993). [PubMed: 8413277]
85. Charpentier B & Rosbash M Intramolecular structure in yeast introns aids the early steps of in vitro spliceosome assembly. *RNA* 2, 509–22 (1996). [PubMed: 8718681]
86. Meyer M, Plass M, Perez-Valle J, Eyraas E & Vilardell J Deciphering 3' splice site selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing. *Mol Cell* 43, 1033–9 (2011). [PubMed: 21925391]
87. Gahura O, Hammann C, Valentova A, Puta F & Folk P Secondary structure is required for 3' splice site recognition in yeast. *Nucleic Acids Res* 39, 9759–67 (2011). [PubMed: 21893588]
88. Bevilacqua PC, Ritchey LE, Su Z & Assmann SM Genome-Wide Analysis of RNA Secondary Structure. *Annu Rev Genet* 50, 235–266 (2016). [PubMed: 27648642]
89. Warf MB & Berglund JA Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem Sci* 35, 169–78 (2010). [PubMed: 19959365]
90. Solnick D Alternative splicing caused by RNA secondary structure. *Cell* 43, 667–76 (1985). [PubMed: 4075405]
91. Spingola M, Grate L, Haussler D & Ares M, Jr. Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *Rna* 5, 221–34 (1999). [PubMed: 10024174]
92. Parada GE, Munita R, Cerda CA & Gysling K A comprehensive survey of non-canonical splice sites in the human transcriptome. *Nucleic Acids Res* 42, 10564–78 (2014). [PubMed: 25123659]

Fig. 1.
A schematic representation of the *S. cerevisiae* splicing cycle.

Fig. 2.

CryoEM structures demonstrate that the splicing reaction is catalyzed by RNA with the help of two Mg^{2+} ions. (a) The left panel is a model of the spliceosome (presumably the B* complex) immediately before the branching reaction, generated based on the structure of the C complex. The model illustrates that the M2 ion can potentially activate the 2' OH of BP A, which will attack the 5' ss, free the 5' exon and generate a lariat intermediate. Bases at the -1 and +1 position from the BP on the 5' intron are not shown to avoid over-crowding of the figure. The right panel shows structures of RNAs around the active site in the C complex (PDB ID 5GMK). (b) The left panel is a model of the spliceosome immediately before the ligation reaction, generated based on the structure of the P complex. The model illustrates that the M1 ion can potentially activate the 3' OH of 5' exon, which will attack the 3' ss and ligate the two exons. The right panel shows structures of RNAs around the active site in the P complex (PDB ID 6BK8).

Fig. 3.

Multiple proteins including Prp8 and Snu114 from U5 snRNP and proteins of the NTC and NTR complexes form a protein core around the active site indicated by the red circle. This protein core is fairly stable from the B^{act} to the ILS complex (the structure of the P complex with PDB ID 6BK8 is shown here as an example).

Fig. 4.

The snRNA core (in the black circle) around the active site (U2/U6 helix 1 and 2, the U2 snRNA region that basepairs with the BPS, U6 ISL and ACAGAGA box, and U5 snRNP loop 1) remain largely invariant from the B^{act} to the ILS complex. PDB IDs for the B^{act}, C, C*, P, and ILS complex structures used to generate this figure are 5GM6, 5GMK, 5MQ0, 6BK8, and 5Y88, respectively.

Fig. 5. Pre-mRNA changes its conformation dramatically through the splicing cycle. Different regions of the pre-mRNA are shown in the color code depicted in the schematic representation of the pre-mRNA. snRNAs are shown in grey. The same PDBs as in Fig. 4 were used to generate this figure.

Fig. 6.

A stem-like secondary structure is observed in the intronic region between the BPS and 3' ss in the P complex structure (PDB ID 6BK8), which may bring the 3' ss close to the 5' ss and BPS to facilitate the recognition of the 3' ss before the ligation reaction.

Fig. 7.

RNA interactions in splice site recognition. (a) The 5' ss region is first recognized through basepairing with the 5' end of U1 snRNA as observed in the A complex structure (PDB ID 6G90). (b) The 5' ss region is subsequently basepaired with U6 snRNA after spliceosomal activation in the B^{act} complex (PDB ID 5GM6). (c) The BPS is recognized through basepairing with U2 snRNA as observed in the A complex (PDB ID 6G90). (d) The 3' ss is recognized mainly through non-Watson Crick basepairing between the last two nucleotides of the intron (nucleotides AG) with the 5' ss (nucleotide G) and BP (nucleotide A) as observed in the P complex (PDB ID 6BK8). Black dashed lines indicate hydrogen bonds. (e) Base substitution at the last nucleotide G in the intron with A, C, and U result in partial or complete loss of hydrogen bonds with the 5' ss. (f) Base substitution of the second to last nucleotide A in the intron with C, G, and U result in the loss of hydrogen bonds with the branch point A.

Table 1.

Recently determined high resolution cryoEM structures of spliceosome-related complexes.

Complex	Species	PDB ID	Resolution(Å)	Lab	Reference
ILS	<i>S. pombe</i>	3JB9	3.6	Yigong Shi	Yan <i>et al.</i> 2015 ³⁴
U1 snRNP	<i>S. cerevisiae</i>	5UZ5	3.6	Rui Zhao Hong Zhou	Li <i>et al.</i> 2017 ³⁵
Tri-snRNP	<i>S. cerevisiae</i>	5GAN	3.7	Kiyoshi Nagai	Nguyen <i>et al.</i> 2016 ³⁷
Tri-snRNP	<i>S. cerevisiae</i>	3JCM	3.8	Yigong Shi	Wan <i>et al.</i> 2016 ³⁶
A	<i>S. cerevisiae</i>	6G90	4.0	Kiyoshi Nagai	Plaschka <i>et al.</i> 2018 ³⁸
pre-B	<i>S. cerevisiae</i>	5ZWM	3.4	Yigong Shi	Bai <i>et al.</i> 2018 ³⁹
B	<i>S. cerevisiae</i>	5ZWO	3.9	Yigong Shi	Bai <i>et al.</i> 2018 ³⁹
B	<i>S. cerevisiae</i>	5NRL	7.2	Kiyoshi Nagai	Plaschka <i>et al.</i> 2017 ⁴⁰
B ^{act}	<i>S. cerevisiae</i>	5GM6	3.5	Yigong Shi	Yan <i>et al.</i> 2016 ⁴¹
C	<i>S. cerevisiae</i>	5GMK	3.4	Yigong Shi	Wan <i>et al.</i> 2016 ¹⁹
C	<i>S. cerevisiae</i>	5LJ3 5LJ5	3.8 10.0	Kiyoshi Nagai	Galej <i>et al.</i> 2016 ²⁰
C*	<i>S. cerevisiae</i>	5WSG	4.0	Yigong Shi	Yan <i>et al.</i> 2017 ²¹
C*	<i>S. cerevisiae</i>	5MPS 5MQ0	3.85 4.17	Kiyoshi Nagai	Fica <i>et al.</i> 2017 ²²
P	<i>S. cerevisiae</i>	6BK8	3.3	Rui Zhao Hong Zhou	Liu <i>et al.</i> 2017 ²³
P	<i>S. cerevisiae</i>	5YLZ	3.6	Yigong Shi	Bai <i>et al.</i> 2017 ⁴⁷
P	<i>S. cerevisiae</i>	6EXN	3.7	Kiyoshi Nagai	Wilkinson <i>et al.</i> 2017 ²⁵
ILS	<i>S. cerevisiae</i>	5Y88	3.5	Yigong Shi	Wan <i>et al.</i> 2017 ⁴⁹
B	human	5O9Z	4.5	Reinhard Luhrmann Holger Stark	Bertram <i>et al.</i> 2017 ⁵⁰
B ^{act}	human	5Z58 5Z56 5Z57	4.9 5.1 6.5	Yigong Shi	Zhang <i>et al.</i> 2018 ⁵¹
B ^{act}	human	6FF4	3.4	Reinhard Luhrmann Holger Stark	Haselbach <i>et al.</i> 2018 ⁵²
C	human	5YZG	4.1	Yigong Shi	Zhan <i>et al.</i> 2018 ⁵³
C*	human	5XJC	3.8	Yigong Shi	Zhang <i>et al.</i> 2017 ³⁰
C*	human	5MQF	5.9	Reinhard Luhrmann Holger Stark	Bertram <i>et al.</i> 2017 ³¹