Annual Review of Biochemistry

RNA Splicing by the Spliceosome

Max E. Wilkinson,* Clément Charenton,* and Kiyoshi Nagai*

MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom; email: mwilkin@mrc-lmb.cam.ac.uk, ccharent@mrc-lmb.cam.ac.uk

Annu. Rev. Biochem. 2020. 89:359–88

First published as a Review in Advance on December 3, 2019

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

https://doi.org/10.1146/annurev-biochem-091719-064225

*All authors contributed equally to this article

M.E. Wilkinson and C. Charenton would like to dedicate this review to the memory of Kiyoshi Nagai (1949–2019)

Keywords

spliceosome, RNA catalysis, pre-mRNA splicing, helicases, cryo-EM, cryo–electron microscopy, crystallography

Abstract

The spliceosome removes introns from messenger RNA precursors (pre-mRNA). Decades of biochemistry and genetics combined with recent structural studies of the spliceosome have produced a detailed view of the mechanism of splicing. In this review, we aim to make this mechanism understandable and provide several videos of the spliceosome in action to illustrate the intricate choreography of splicing. The U1 and U2 small nuclear ribonucleoproteins (snRNPs) mark an intron and recruit the U4/U6.U5 tri-snRNP. Transfer of the 5′ splice site (5′SS) from U1 to U6 snRNA triggers unwinding of U6 snRNA from U4 snRNA. U6 folds with U2 snRNA into an RNA-based active site that positions the 5′SS at two catalytic metal ions. The branch point (BP) adenosine attacks the 5′SS, producing a free 5′ exon. Removal of the BP adenosine from the active site allows the 3′SS to bind, so that the 5′ exon attacks the 3′SS to produce mature mRNA and an excised lariat intron.

Contents

Protein-coding sequences of eukaryotic genes are often interrupted by noncoding introns. Hence, introns must be removed from messenger RNA precursors (pre-mRNA), and protein-coding segments known as exons must be spliced together to form mature messenger RNAs (mRNAs). This essential process in eukaryotic gene expression is known as pre-mRNA splicing. In the yeast *Saccharomyces cerevisiae*, there are approximately 5,000 protein-coding genes, of which approximately 400 genes contain single introns (1, 2). In contrast, in humans there are approximately 20,000 protein-coding genes, which on average contain 8 introns each with a median length of approximately 1 kb (3). The cell therefore faces a considerable challenge in correctly identifying the exons within a sea of intron sequences. Adding to this challenge, approximately 95% of human genes are alternatively spliced (4, 5). Hence, a single gene can produce multiple protein isoforms either by including or skipping an exon or by choosing an alternative exon. This enormously expands the proteome that can be derived from a limited number of genes, contributing to the vast complexity of higher organisms (6).

CHEMISTRY OF THE SPLICING REACTION

Introns are defined by three important sites, the 5′ splice site (5′SS), branch point (BP) adenosine, and 3′ splice site (3′SS), all of which are defined by short conserved sequences (**Figure 1*a***;

Supplemental Material >

Supplemental Videos 1 and **2**). In yeast, the 5′SS is followed by a highly conserved sequence, GUAUGU, and the 3′SS is preceded by the YAG trinucleotide, in which Y is a pyrimidine. The BP adenosine is located 18–40 nucleotides upstream of the 3′SS in a highly conserved sequence, UACUA**A**C, in which the bold A denotes the BP adenosine (1, 2). In humans, the nucleotide sequences surrounding the 5′SS and BP adenosine are less stringently conserved, and the 3′SS YAG trinucleotide is preceded by a poly-pyrimidine tract (7) (**Figure 1*a***). Biochemical characterization of splicing intermediates in 1984 established a two-step phosphoryl transfer mechanism of splicing (8–11) similar to that of group II intron self-splicing. In the first reaction (termed branching), the 2′ hydroxyl group of the BP adenosine attacks the phosphodiester group at the 5′SS, producing a cleaved 5′ exon and a lariat-intron–3′ exon intermediate in which the 5′ phosphate of the first intron nucleotide (G) is linked to the 2′ oxygen of the BP adenosine (**Figure 1*b***). In the second step (exon ligation), the newly exposed 3′ hydroxyl group of the 5′ exon attacks the phosphodiester group at the 3′SS, ligating the 5′ and 3′ exons to form mRNA and releasing the lariat intron. These two reactions are catalyzed by a molecular machine termed the spliceosome (12). Steitz & Steitz (13) proposed a two-metal-ion mechanism for splicing catalysis, in which two metal ions in the active site of the spliceosome stabilize the pentacovalent transition states of the splicing transesterification reactions (**Figure 1*c***). For the first phosphoryl transfer reaction, one of the two metal ions (M1) stabilizes the leaving group, the 3′ hydroxyl of the last 5′ exon nucleotide, and the second metal ion (M2) activates the attacking nucleophile, the 2′ hydroxyl group of the BP adenosine. One of the reactants of the first reaction, the BP adenosine, must leave the active site to allow the binding of the 3′SS to the active site. In the second phosphoryl transfer reaction (exon ligation), M1 activates the 3′ hydroxyl group of the 5′ exon, and M2 stabilizes the leaving group, the 3′ hydroxyl group of the last intron nucleotide. M1 and M2 are both coordinated by RNA (14); therefore, the spliceosome is a ribozyme.

THE SPLICEOSOME IS A COMPLEX MOLECULAR MACHINE

The spliceosome is not a preassembled enzyme; instead, it is formed anew on its substrate from 5 small nuclear RNAs (snRNAs) and approximately 100 proteins (15) (**Figure 1*d***; **Supplemental Video 1**). U1, U2, U4, and U5 snRNAs are transcribed by RNA polymerase II and acquire a tri-methyl-guanosine cap, whereas U6 snRNA is transcribed by RNA polymerase III and has a γ-monomethyl guanosine cap (16). Seven homologous Sm proteins assemble into a ring around the U-rich sequence known as the Sm site located toward the 3′ end of U1, U2, U4, and U5 snRNAs (17–19), whereas a U-rich sequence at the 3′ end of U6 snRNA threads through a preassembled ring of seven paralogous LSm proteins (LSm2–8) (20–22). Each of these snRNAs binds a specific set of additional proteins and forms a small nuclear ribonucleoprotein (snRNP) particle, pronounced "snurp" for short (17, 18). Several non-snRNP–associated proteins and protein complexes, including splicing factors and eight ATP-dependent helicases, are also involved in splicing. Within the spliceosome, the snRNAs perform the essential roles of catalysis and substrate recognition.

The U1 and U2 snRNPs recognize the 5′SS and the BP sequence, respectively, and form the prespliceosome or A complex. The prespliceosome then associates with the preassembled U4/U6.U5 tri-snRNP to form the fully assembled spliceosome (**Figure 1*d***; **Supplemental Video 1**). U6 snRNA, which ultimately folds to form the active site of the spliceosome, is extensively base-paired with U4 snRNA within the tri-snRNP (15, 23, 24). The DEAD-box helicase Prp28 releases the 5′SS from U1 snRNP and transfers it to the ACAGAGA box within U6 snRNA (25). The RNA helicase Brr2 then separates U4 snRNA from U6 snRNA (26, 27) and allows the U6 snRNA sequence adjacent to the 5′SS-bound ACAGAGA box to fold and associate

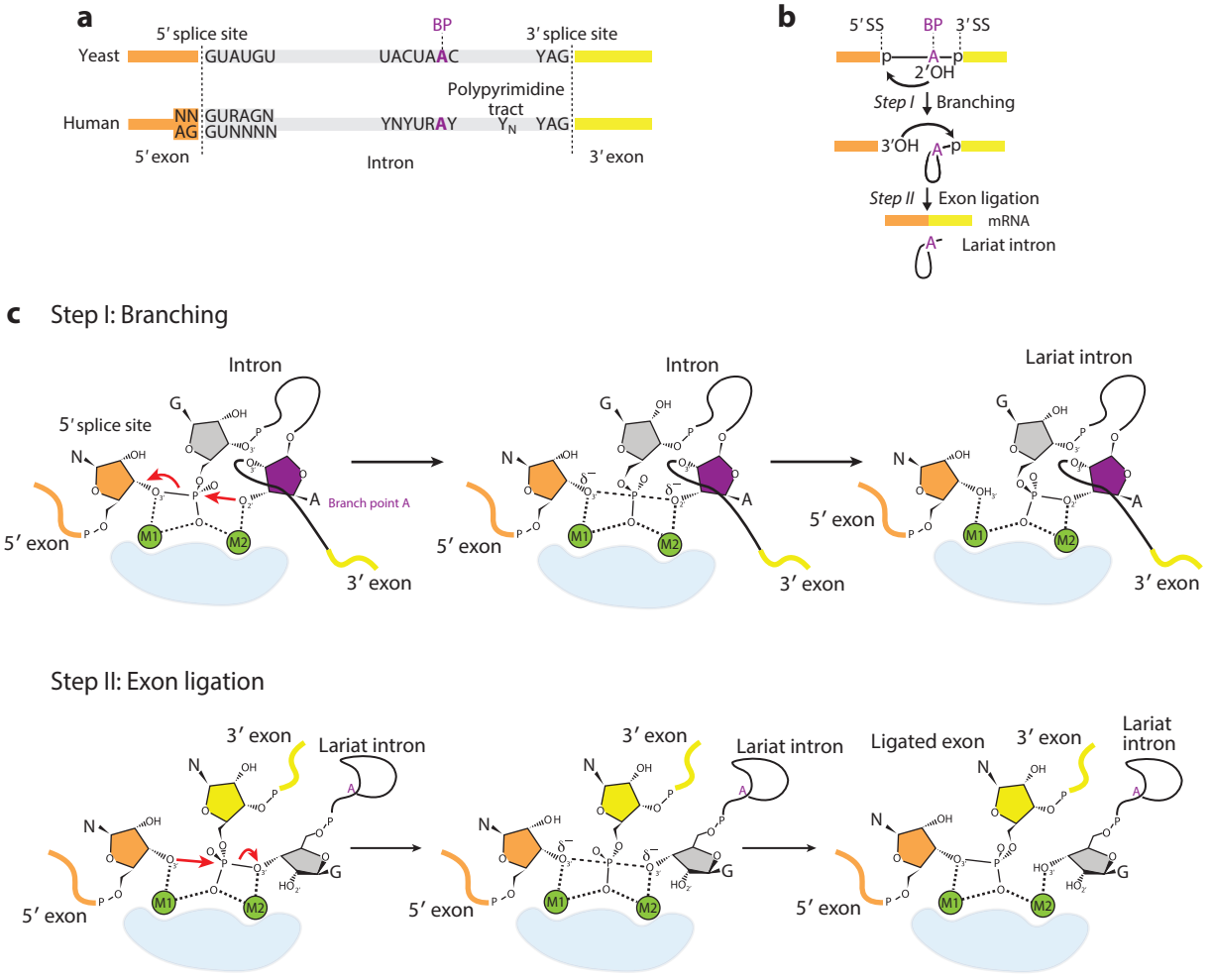


Figure 1

Schematic representations of yeast and human pre-mRNA substrates, the splicing reaction, and the splicing cycle. (*a*) Introns are characterized by three short conserved sequences, the 5′SS, BP sequence, and 3′SS. Purple indicates the BP adenosine. The sequences surrounding these sites are stringently conserved in yeast but more degenerate in humans. These sequences are recognized multiple times during the splicing cycle to maintain the fidelity of the splicing reactions. (*b*) Introns are removed by two transesterification reactions, branching and exon ligation, that are catalyzed at a single active site. (*c*) The two-metal-ion mechanism, originally proposed by Steitz & Steitz (13), proceeds via a pentacovalent transition state. For the branching reaction, the 5′SS is first positioned at the active site and the BP adenosine nucleophile is docked into the active site to attack the phosphorus of the 5′SS, producing the free 5′ exon and lariat–3′ exon intermediate. In the resulting lariat–3′ exon intermediate, the phosphorus atom of the first intron nucleotide is linked to the 2′O of the BP adenosine. The 5′ exon remains in the active site, but for the exon ligation reaction the BP adenosine moves away to allow the 3′SS to dock into the active site. The 5′ and 3′ exons are ligated by the nucleophilic attack of the 5′ exon 3′OH group at the phosphorus atom of the 3′SS. (*d*) The spliceosome is assembled in a highly ordered manner, activated to form the active site, and remodeled extensively to perform the branching and exon ligation reactions, release mRNA (ligated exons), and disassemble the spliceosome. ATPases in the DEAD-box, DEAH-box, and Ski-2 families (*red*) play crucial roles in remodeling processes. Abbreviations: BP, branch point; ILS, intron-lariat spliceosome; M1 and M2, catalytic metal ions one and two; mRNA, messenger RNA; NTC, Prp19-associated complex; NTR, Prp19-related complex; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein; SS, splice site.

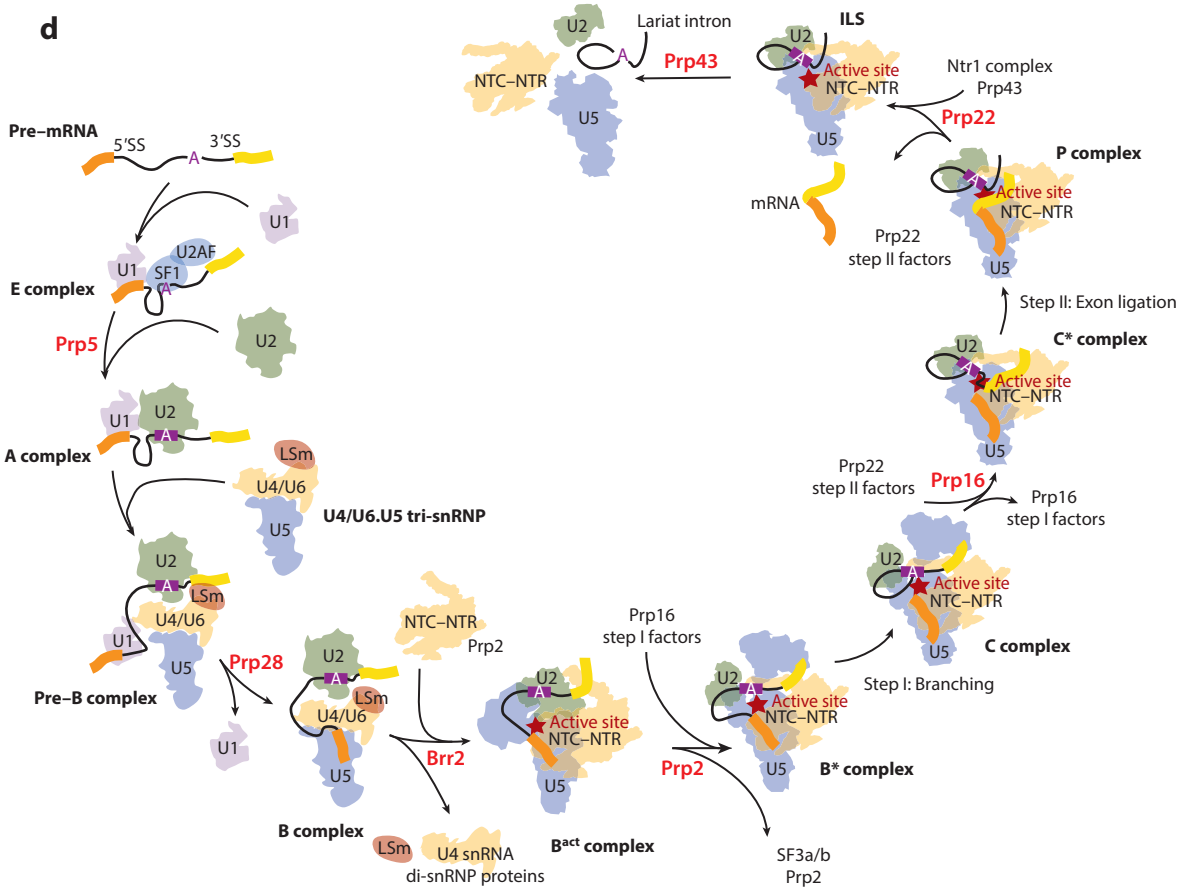


Figure 1

(*Continued*)

with part of U2 snRNA to yield the active site harboring two catalytic metal ions (13, 14, 28). The 5′SS is positioned at the M1 metal ion (**Figure 1*c***). When the BP adenosine is docked into the active site, the branching reaction produces the cleaved 5′ exon and the lariat-intron intermediate (29, 30). The 5′ exon remains in the active site, but the BP adenosine must vacate the active site for the incoming 3′SS site for the exon–ligation reaction. Finally, the 5′ and 3′ exons are ligated, and the resulting mRNA (ligated exons) is released from the active site. The spliceosome choreographs the intricate movements of these substrates in and out of the active site (**Supplemental Video 1**).

Supplemental Material >

Since 2015, cryo–electron microscopy (cryo-EM) has captured the structures of both yeast and human spliceosomes in several key states during assembly, activation, catalysis, and disassembly (reviewed in 15, 23, 24). These structures, together with the results of more than 30 years of extensive genetic and biochemical experiments, have provided mechanistic insights into the molecular mechanism of pre-mRNA splicing. The aim of this review is to describe a complete picture of the inner workings of the spliceosome. We also provide seven videos that show transitions between every major state of the spliceosome (**Supplemental Videos 1–7**).

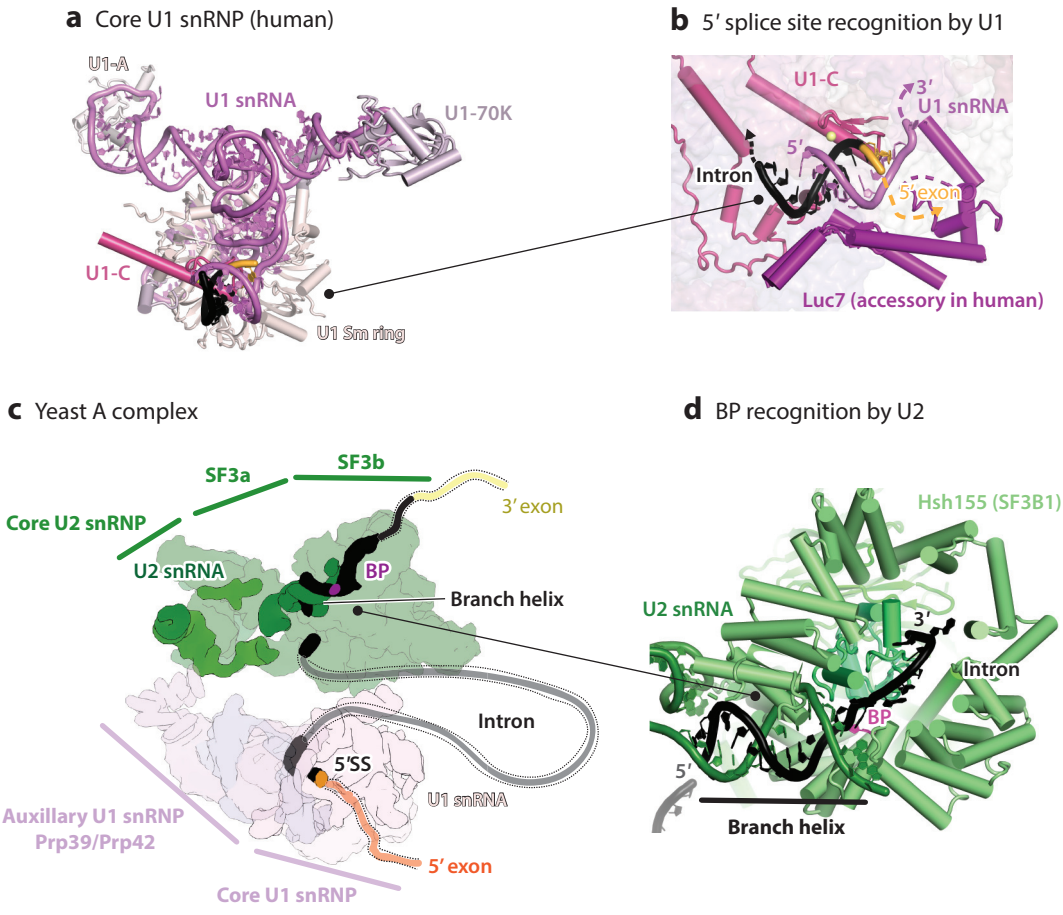


Figure 2

The A complex (prespliceosome) is formed by the U1 and U2 snRNPs bound to the 5′SS and the BP sequence. These two sites can be 1,000 nucleotides apart in yeast. The two reactants of the branching reaction are brought into a single complex by A-complex formation. (*a*) The crystal structure of human U1 snRNP bound to the 5′SS (34, 35). (*b*) Binding of the 5′SS to yeast U1 snRNP. The yeast U1 snRNP is considerably larger than the human counterpart, but its functional core is highly similar to human U1 snRNP. In free yeast U1 snRNP (37), the 5′ end of U1 snRNA and Luc7 are disordered, but upon binding of the 5′SS the U1–5′SS duplex is formed and sandwiched between the zinc-finger domains of U1-C and Luc7 (31). (*c*) Schematic of the cryo-EM structure of the yeast A complex (31). U1 and U2 snRNPs interact loosely side by side. The intron between the 5′SS and the BP sequence loops out from the A complex and is indicated by a gray dotted line. (*d*) The BP sequence (UACUAAC in yeast) pairs with a conserved sequence in U2 snRNA to form the branch helix within SF3b of the U2 snRNP. The branch helix persists until the spliceosome is disassembled. The BP adenosine is flipped out from the branch helix and cradled by Hsh155 (SF3B1 in humans). Abbreviations: BP, branch point; cryo-EM, cryo–electron microscopy; snRNP, small nuclear ribonucleoprotein; SS, splice site.

EARLY ASSEMBLY PROCESS AND THE FORMATION OF THE PRESPLICEOSOME

Supplemental Material >

The early steps of spliceosome assembly entail recognition and marking of an intron by the U1 and U2 snRNPs within the prespliceosome (31) (**Figures 1*d*** and **2*c***; **Supplemental Video 2**). In higher eukaryotes, this stage is subjected to extensive regulation, by both *cis*-acting sequence elements and *trans*-acting splicing factors, and is believed to be the main determinant for splice-site selection during alternative splicing, as it generally commits an intron to being removed (6).

E-Complex Formation

Early spliceosome assembly (E-complex formation) is initiated by binding of U1 snRNP to the 5′SS through base-pairing between the 5′SS and the 5′ end of U1 snRNA (1, 32, 33) (**Figure 2*a***). Human U1 snRNP comprises the Sm ring and three U1-specific proteins (U1-70k, U1A, and U1C) (18), and the zinc-finger domain of U1C directly contacts this RNA duplex to stabilize the 5′SS/U1 snRNA interaction (34, 35) (**Figure 2*a***). Yeast U1 snRNP contains seven additional stably associated proteins (31, 36–38), four of which have human counterparts that facilitate the binding U1 snRNP to weak 5′SSs, thereby functioning as alternative splicing factors. For instance, yeast Luc7 contacts the backbone of the U1/5′SS helix (31), and both yeast and human Luc7 (LUC7L in humans) affect selection of 5′SSs (39) (**Figure 2*b***).

In the metazoan E complex, the pre-mRNA BP sequence is bound by SF1/mBBP, while the U2AF65–U2AF35 heterodimer cooperatively binds the downstream polypyrimidine tract and 3′SS (40, 41) (**Figure 1*d***). In yeast, the Msl5–Mud2 dimer fulfils a similar function in recognizing the BP sequence but does not recognize the 3′SS (42).

A-Complex Formation

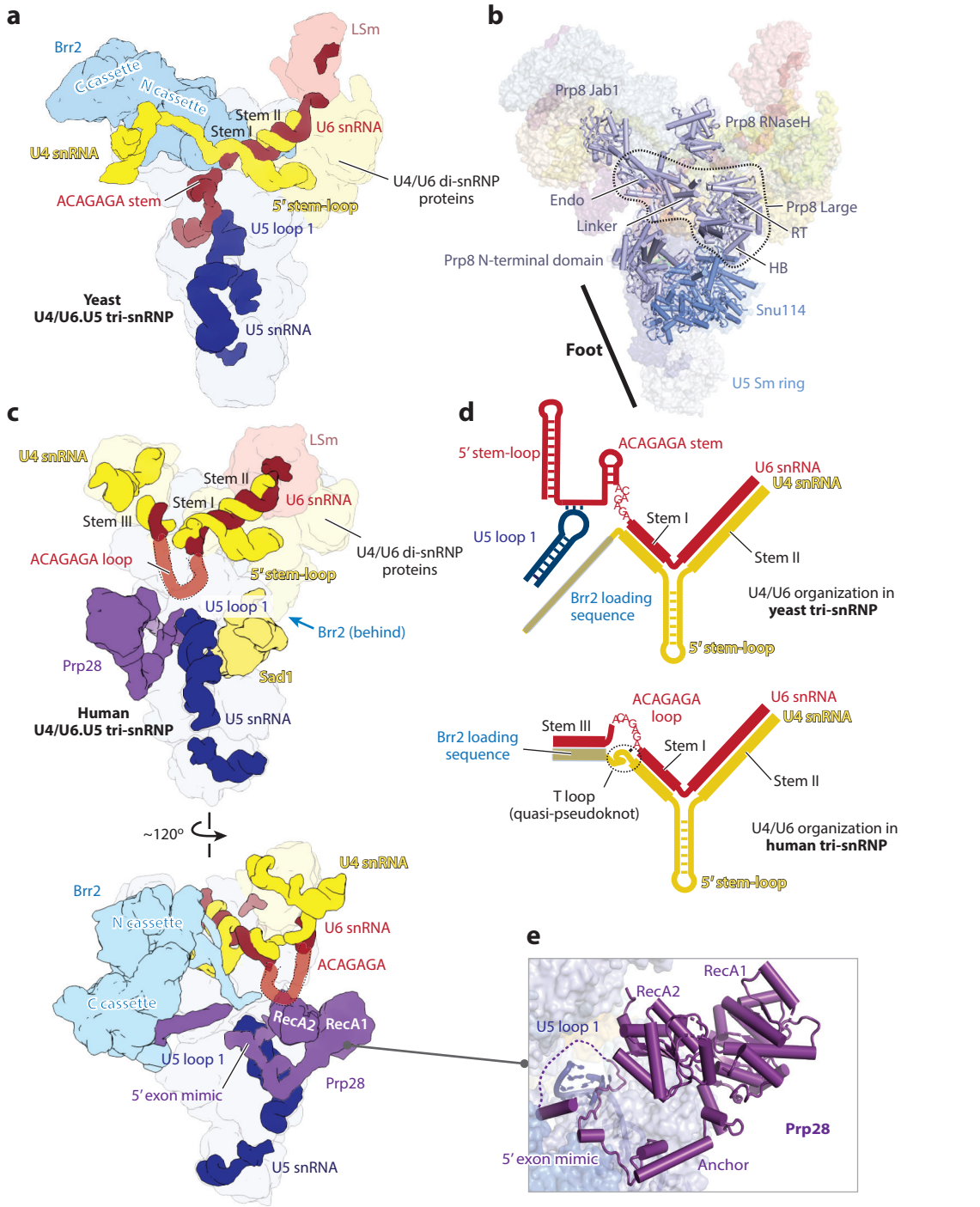
The DEAD-box helicases Prp5 and Sub2 displace SF1 and U2AF and recruit the U2 snRNP to the BP sequence, yielding the prespliceosome or the A complex (43–45) (**Figure 1*d***; **Supplemental Video 2**). U2 snRNP has a bilobed architecture organized along the U2 snRNA. The SF3b subcomplex (comprising six to seven proteins) binds the 5′ half of U2 snRNA near the BP-binding sequence (46) (**Figure 2*c***). The Sm proteins assemble around the Sm site near the 3′ end of U2 snRNA to form the U2 core domain together with the U2B″–U2A′ (Msl1–Lea1 in yeast) (31, 47). The trimeric SF3a complex bridges the SF3b and core domains of U2 snRNP. Upon stable integration of the U2 snRNP into the prespliceosome, the U2 snRNA pairs with the BP region of the pre-mRNA to form the branch helix (48–50) encapsulated within the SF3b protein SF3B1 (Hsh155 in yeast). The BP adenosine base is flipped out from the branch helix and interacts with SF3B1 (**Figure 2*d***). As a result, its 2′OH group, which acts as a nucleophile for the first reaction, is not yet easily accessible (51). Displacement of SF3b from the BP adenosine is a prerequisite for the branching reaction.

In the yeast prespliceosome, U1 snRNP and U2 snRNP bind side by side, but their interface is not very extensive (**Figure 2*c***; **Supplemental Video 2**). The U1 components Prp39–Prp42 form a heterodimer bridging the U1 core domain to the 5′ region of the U2 snRNP (31). In metazoans, the PRPF39 homodimer seems to mirror this function but as an accessory component of the U1 snRNP.

Supplemental Material >

STRUCTURE OF TRI-snRNP

The U4/U6.U5 tri-snRNP is the largest preassembled spliceosomal complex (52–56) (**Figure 3**). It contains protein and RNA building blocks that become part of the active site after activation, including U6 snRNA, which pairs with the 5′SS and folds to form the active site. In tri-snRNP, U6 snRNA is paired with the U4 snRNA, which acts as a chaperone to keep U6 in a precatalytic conformation (57–59). U5 snRNA loop 1 is important for tethering the 5′ exon during branching and aligning the 5′ and 3′ exons during exon ligation (60, 61). All these RNA elements are organized by an extensive protein exoskeleton. Some of these proteins remain bound to the catalytic complexes, but others dissociate during activation (15). The complex organization of tri-snRNP ensures the spliceosome active site is not formed until substrate pre-mRNA integration. Cryo-EM structures of yeast and human tri-snRNPs have revealed unexpected differences in organization (52–56) (**Figure 3**). Therefore, we first describe features in common between yeast



(Caption appears on following page)

Figure 3 (*Figure appears on preceding page*)

Cryo-EM structures of the yeast and human U4/U6.U5 tri-snRNPs. (*a*) Schematic overview of the yeast U4/U6.U5 tri-snRNP in which RNA components and Brr2 are highlighted (52–54). The single-stranded region of U4 snRNA is already loaded in the active site of Brr2 helicase (N-terminal cassette). (*b*) Four domains of Prp8 are connected by flexible linkers so these domains can move with respect to each other. These domains organize RNA elements of the spliceosome, and the Large domain accommodates the RNA-based active site after spliceosome activation. Subdomains of the Large domain (indicated by the *dotted line*) are labeled as helical bundle (HB), reverse-transcriptase (RT), and endonuclease (endo). Snu114 and the Prp8 N-terminal domain form the foot domain. (*c*) Two different views of the human U4/U6.U5 tri-snRNP (55, 56). It is distinctively different from its yeast counterpart in that Brr2 is held on the opposite face of the tri-snRNP by its interaction with Sad1. The DEAD-box helicase Prp28 is stably bound. (*d*) The organization of the U4, U6, and U5 snRNAs in the yeast (*top*) and human (*bottom*) tri-snRNP. The U4 and U6 snRNAs are extensively base-paired to form stems I and II and the 5′ stem-loop. In yeast, the 5′ end of U6 snRNA forms the ACAGAGA stem and interacts with loop 1 of U5 snRNA. These interactions prevent formation of stem III and keep the Brr2 binding sequence available for Brr2 binding. The 5′ end of U6 snRNA was attributed to pre-mRNA by Wan et al. (54), but their tri-snRNP structure within the pre–B complex (38) reveals the same RNA arrangement as in the tri-snRNP by Nguyen et al. (53). In human U4 and U6, snRNAs form stem III, and stem I is capped with a quasi-pseudoknot (T loop). These interactions make the ACAGAGA sequence protrude as a flexible loop toward Prp28, ready to pair with the 5′SS released from U1 snRNP by Prp28. (*e*) In humans, the two RecA domains of Prp28 are preceded by the anchor sequence that interacts with the Prp8 N-terminal domain, Snu114, and the C-terminal helicase cassette of Brr2. The anchor sequence is not conserved in yeast Prp28, but the 5′ exon mimic sequence that interacts with U5 snRNA loop 1 is conserved from yeast and human (56). Abbreviations: mRNA, messenger RNA; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein.

and human and then discuss important differences that give rise to different activation pathways in these two systems.

Common Features Between the Yeast and Human Tri-snRNP

The largest and most conserved splicing factor, Prp8, occupies a central position within the tri-snRNP and assembled spliceosomes (23) (**Figure 3*b***). It comprises four domains—N-terminal, Large, RNaseH, and Jab1—flexibly connected by long linker peptides (62) that each act as assembly nodes for other splicing factors. Prp8 organizes RNA elements of the spliceosome and, most importantly, accommodates the RNA-based active site within its Large domain after spliceosome activation (62, 63), a function inherited from its ancestral group II intron-encoded maturase protein (64).

The Prp8 N-terminal domain forms the stable spliceosome foot domain along with the U5 snRNA, its associated Sm ring, and Snu114, a GTPase closely related to the ribosomal translocase EF-G/EF2 (65) (**Figure 3*b***). U5 snRNA forms a long helix capped by an invariant uridine-rich loop 1, accommodated in a groove in the Prp8 N-terminal domain (52–56) (**Figure 3*a*,*c***). The Large domain of Prp8, comprising the closely linked helix bundle (HB), reverse transcriptase (RT), linker domain, and endonuclease (endo) domain (53, 62), is located at the center of the tri-snRNP assembly (**Figure 3*b***). The Prp8 RNaseH domain, C-terminal to the Large domain, moves dramatically between the distinct spliceosome intermediates. It forms a protein–protein interaction hub and changes protein partners depending on its rotation state (15, 23, 24, 66). Finally, the C-terminal Jab1 domain of Prp8 stably binds Brr2 (67, 68) (**Figure 3*a*,*b***), the Ski2-like helicase responsible for spliceosome activation, tracking its substantial movements throughout the splicing cycle and ensuring it remains tethered to the main body of Prp8. Brr2 itself consists of two helicase cassettes, of which only the N-terminal cassette is catalytically active (**Figure 3*a*,*c***).

U6 snRNA eventually forms the binding site for catalytic metal ions in the active site, but within the tri-snRNP, it is maintained in an inactive conformation through extensive pairing with the U4 snRNA (**Figure 3*d***). The U4 and U6 snRNA form an extended duplex divided into two coaxially stacked stems (I and II) separated by the 5′ stem-loop of U4 snRNA that protrudes from the duplex, forming a three-way junction (**Figure 3*d***). Stem II and the U4 5′ stem-loop are bound by U4/U6 di-snRNP proteins (Snu13, Prp31, Prp3, and Prp4), while the LSm ring binds the

U-rich sequence at the end of U6 snRNA (21) (**Figure 3*a*,*c***). The U4/U6 di-snRNP and the U5 snRNP are bridged by the Prp8 Large and RNaseH domains to form the U4/U6.U5 tri-snRNP (**Figure 3*a*,*b***). The ACAGAGA box of U6 snRNA is found upstream of U4/U6 stem I and is functionally important, as it receives the 5′SS from U1 snRNA and positions it at the active site metal M1 during spliceosome activation.

Structural Differences Between the Yeast and Human Tri-snRNPs

In addition to the common protein components discussed above, human tri-snRNP contains additional proteins Sad1 (69), Prp28, SNRNP-27k, and RBM42 (56), whereas yeast tri-snRNP contains the B-complex-specific proteins Prp38, Snu23, and Spp381 (51). This difference in protein composition and the subtle differences in U4 and U6 snRNA sequences give rise to large architectural differences, most notably the position of Brr2 and the availability of the U6 ACAGAGA box (**Figure 3*a–d***).

In human tri-snRNP, Brr2 is held more than 100 Å away from its U4 snRNA substrate (55) (**Figure 3*c***). This position is stabilized by interactions between the N-terminal PWI domain of Brr2 (70) and the Sad1 protein bound across the Prp8 N-terminal domain and Snu114 in the tri-snRNP foot (55) and by contacts between the extended N terminus of Brr2 and the Prp8 Large and RNaseH domains (56). Furthermore, the high-resolution structure of human tri-snRNP (56) showed that the Brr2 loading sequence in U4 snRNA pairs with U6 snRNA to form stem III, a third helix not found in yeast, which prevents Brr2 loading on U4 snRNA (**Figure 3*d***). As a result, the U6 ACAGAGA box, found between U4/U6 stems I and III, protrudes as a flexible loop ready to pair with the 5′SS when it is released from U1 snRNP by Prp28 (**Figure 3*c*,*d***). This configuration is associated with a T loop (originally named quasi-pseudoknot) adjacent to U4/U6 stem I, stabilized by RBM42.

In contrast, in yeast tri-snRNP Sad1 is not bound and the Brr2 helicase is already loaded onto the single-stranded region of U4 snRNA adjacent to stem I, ready to translocate on U4 snRNA and unwind the U4/U6 snRNA duplex (52–54) (**Figure 3*a***). The U6 ACAGAGA box in U6 snRNA folds into a hairpin, and the preceding single-stranded U6 sequence pairs with U5 snRNA loop 1 (53) (**Figure 3*a*,*d***). These interactions prevent formation of the U4/U6 stem III found in the human counterpart, thereby keeping the Brr2 loading sequence of U4 snRNA free and available for Brr2 binding (**Figure 3*d***).

In Humans, Prp28 Is Positioned Proximal to the snRNA Elements That Recognize the Pre-mRNA

Prp28 is the initiator of spliceosome activation. It disrupts the U1/5′SS duplex, thereby facilitating the transfer of the 5′SS to the U5 and U6 snRNAs (25). In humans, Prp28 stably binds to the tri-snRNP foot through its N-terminal anchor sequence, positioning its catalytic RecA domains adjacent to the U6 snRNA ACAGAGA box loop and U5 snRNA loop 1 (**Figure 3*c*,*e***). U5 loop 1 is the element that pairs with the last few nucleotides of the 5′ exon during splicing catalysis (60, 61). Therefore, Prp28 is perfectly positioned in human tri-snRNP to deliver the 5′SS (**Figure 3*c*,*e***). Interestingly, because the end of the 5′ exon does not have a strong consensus sequence (apart for a slight preference for AG at the very end of human exons) (1, 33) (**Figure 1*a***), U5 loop 1 is apparently capable of pairing promiscuously, likely due to its high uridine content. Therefore, loop 1 is blocked by the 5′ exon mimic peptide in the N-terminal domain of Prp28, which covers loop 1 and likely prevents its premature pairing with nonspecific RNA (**Figure 3*e***). Yeast Prp28 lacks the anchor sequence and does not stably bind to tri-snRNP (71, 72). Interestingly, yeast Prp28 retains the 5′ exon mimic sequence (73), whose binding to U5 loop 1 would be mutually exclusive with

that of the U6 ACAGAGA region. Hence, the mechanism of Prp28 recruitment to yeast spliceosome is unclear, but its binding to tri-snRNP is likely to result in a different conformation of the U6 ACAGAGA box to that observed in the available structures.

FORMATION OF THE FULLY ASSEMBLED PRE–B COMPLEX

The pre-mRNA is integrated into the spliceosome by joining of the prespliceosome to the tri-snRNP to form the fully assembled pre–B complex (**Figures 1*d*** and **4**; **Supplemental Video 3**). In yeast and humans, the tri-snRNP structure remains virtually unchanged when joining the prespliceosome to form the pre–B complex (38, 56). Hence, the differences between the human and yeast tri-snRNPs described above persist in their respective pre–B complexes. During prespliceosome formation, binding of the BP sequence to U2 snRNA frees the 5′ end of U2 snRNA (45), enabling it to pair with an exposed sequence at the 3′ end of U6 snRNA in tri-snRNP, forming U6/U2 snRNA stem II within the pre–B complex (74, 75) (**Figure 4*a***). In humans, an additional contact between tri-snRNP and the prespliceosome is observed through a C-terminal extension to the U2 component SF3A1 that loosely binds to the U4/U6 region in tri-snRNP (56). Pre–B complex formation brings into one complex the substrates for the first catalytic reaction (branching), namely the 5′SS and the BP sequence, as well as the RNA elements in the U2, U5, and U6 snRNAs required to form the spliceosome active site. However, U6 snRNA is still chaperoned in its inactive conformation by U4 snRNA, and the 5′SS is still bound to U1 snRNP.

Supplemental Material >

5′SS TRANSFER AND B-COMPLEX FORMATION IN HUMANS

A recent structure of human pre–B complex led to a detailed mechanistic understanding of pre-mRNA integration and 5′SS transfer (56) (**Supplemental Video 3**). In humans, pairing between the 5′SS and the ACAGAGA box (76–78) and between the 5′ exon and U5 loop 1 (60, 61) is completed during the pre–B complex to B complex transition (**Figure 4*a*,*b***). At the pre-B stage, U1 snRNP docks the 5′SS/U1 duplex between the RecA domains of Prp28, and ATP-dependent closure of these RecA domains around the U1 snRNA strand frees the 5′SS from U1 snRNP, allowing the 5′SS to anneal to the flexible U6 ACAGAGA loop (56). Formation of the 5′SS/U6 duplex induces extensive remodeling of the spliceosome, whereby the Prp8 RNaseH domain rotates by 180° and Sad1 dissociates from its binding site on Snu114. During activation, the Brr2 loading site on U4 snRNA, which was paired with U6 snRNA to form stem III, becomes single stranded and exposed as the U6 half of the stem anneals with the 5′SS. This allows Brr2 to load onto the U4 snRNA, ready to unwind the U4/U6 duplex for spliceosome active site formation (**Figure 4*a*,*b***). Transfer of the 5′SS is further accompanied by the release of Prp28 and U1 snRNP and by subsequent binding of B-complex proteins, which stabilizes the nascent 5′SS/U6 duplex (51, 56, 79). In summary, in humans 5′SS release from the U1 snRNP by Prp28 induces a chain of events that results in a dramatic relocation of the Brr2 helicase poised to activate the spliceosome (56) (**Supplemental Video 3**). This functional coupling explains how active site formation is coordinated to integration of a bona fide pre-mRNA substrate.

In yeast, the mechanism of 5′SS transfer from U1 to U6 snRNA is not yet clear, as Prp28 has yet to be visualized on the yeast tri-snRNP or pre–B complex (see the sidebar titled Spliceosome Activation in Yeast).

FORMATION OF THE ACTIVE SITE DURING THE TRANSITION FROM B COMPLEX TO B^{act} COMPLEX

When U4 snRNA is unwound from U6 snRNA by Brr2 (26, 27), U2 and U6 snRNAs, which already pair on one end to form helix II, can fold together to form the active site of the

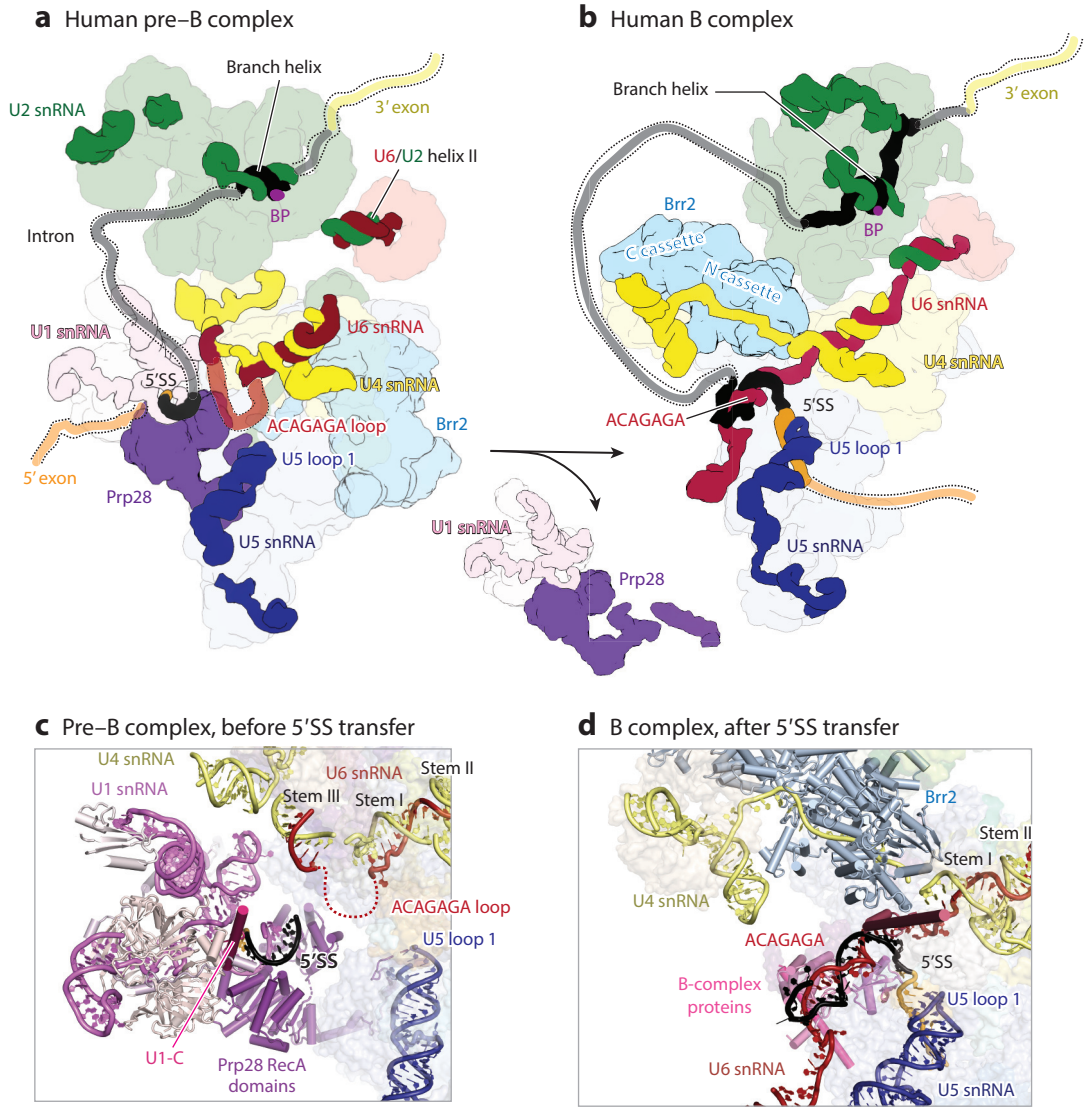


Figure 4

Activation of the human spliceosome, triggered by 5′SS transfer from the U1 snRNP to the U6 ACAGAGA box. (*a*) Schematic representation of the human pre–B complex (56). During pre–B complex formation, the structure of tri-snRNP remains essentially unchanged. The U2 snRNP carrying the branch helix is tethered to the tri-snRNP through formation of U2/U6 helix II and sits on the U4/U6 di-snRNP domain. The U1 snRNP carrying the 5′SS inserts the U1/5′SS helix between the two RecA domains of Prp28 located near the mobile U6 ACAGAGA loop. Upon ATP binding, the RecA domains close and the 5′SS is released and handed over to the U6 ACAGAGA loop. The binding of the 5′SS induces rearrangement of U4 and U6 snRNAs, relocation of Brr2, and dissociation of the U1 snRNP. (*b*) Schematic of the human B complex (79, 150). The single-stranded region of U4 snRNA is loaded in the active site of the relocated Brr2. Upon formation of the 5′SS–U6 ACAGAGA helix, the 5′ exon is positioned to pair with U5 snRNA loop 1 in the B complex. (*c*) Detailed view of the docking of the U1 snRNP to the tri-snRNP in humans. The mobile ACAGAGA loop is formed between stems I and III. The U1/5′SS helix stabilized by U1-C is inserted between the two RecA domain of Prp28. (*d*) Detailed view of RNA interactions in the human B complex. The duplex between the 5′SS and U6 ACAGAGA box is stabilized by the B-complex proteins. Abbreviations: BP, branch point; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein; SS, splice site.

SPLICEOSOME ACTIVATION IN YEAST

In yeast, the timing of Prp28 recruitment and the mechanism of 5′SS transfer are not yet clear, as Prp28 is not stably associated with the yeast tri-snRNP and is not visible in the cryo-EM structure of the yeast pre–B complex (38). The yeast B-complex structure (51) and cross-linking data (143) indicate that, even at the B-complex stage, the 5′SS is not fully paired with U6 and is instead tethered by a few base pairs to nucleotides just upstream of the ACAGAGA sequence, at the tip of the ACAGAGA stem-loop (51, 53). This finding suggests that U4/U6 unwinding and possibly the recruitment of additional factors are necessary for stable pairing of the 5′SS with U6 and U5 snRNAs in yeast (92, 143). Furthermore, 5′SS transfer in yeast does not trigger major remodeling of the tri-snRNP, because Brr2 is already relocated and loaded on U4 snRNA at the tri-snRNP stage. Indeed, in vitro the yeast tri-snRNP disintegrates upon addition of ATP (52), suggesting that Brr2 in the yeast tri-snRNP is active, whereas the human tri-snRNP is stable in the presence of ATP, as Brr2 is held away from U4 snRNA. Hence, how 5′SS integration and active site formation are coupled in yeast remains elusive.

spliceosome in the B^{act} complex (80–82), the first of the catalytic states of the spliceosomes (**Figures 1*d*** and **5*a***; **Supplemental Video 4**). U2 and U6 snRNAs form two short helices, Ia and Ib, which are separated by a bulge and are adjacent to an internal stem-loop (ISL) in U6 snRNA (83) (**Figure 5*b***). A base triple formed between a bulged U (U80 in yeast U6) in the U6 ISL and helix Ib brings two phosphates from the ISL close to two phosphates from the U6 strand of helix Ib. Among these four U6 phosphates, five nonbridging oxygens are used to coordinate the two catalytic metal ions that activate the nucleophiles and stabilize the leaving groups during the two splicing transesterification reactions (13, 14) (**Figures 1*c*** and **5*b–d***). This structure of the active site RNA was predicted on the basis of the group II intron active site structure (84), with conservation of the base triples in the spliceosome supported by base substitution experiments (80). The phosphate oxygen ligands for the catalytic metal ions were identified by sulfur-substitution and metal-rescue experiments (14). The architecture of the catalytic core of the spliceosome has now been visualized in multiple structures of the catalytic states of the spliceosomes (29, 30, 81, 82, 85–91) (**Figure 5*a***).

Supplemental Material >

Extensions of the U2 and U6 snRNAs next to this active site juxtapose the splice sites with the catalytic metal ions (**Figure 5*d–e***). The U6 ACAGAGA box is found immediately upstream of helix Ia and positions the 5′SS at the catalytic metal ions. In yeast, the ACA sequence within the U6 snRNA **ACA**GAGA box pairs with the $U_{+4}G_{+5}U_{+6}$ of the intron 5′SS (77, 78), while the final GA of this sequence forms two base triples with helix Ib—a tertiary interaction that docks the 5′SS into the active site (29, 30, 80, 82) (**Figure 5*c,d***). This so-called catalytic triplex is completed by the above-mentioned base triple between the U6 snRNA ISL (U80) and helix Ib (**Figure 5*c***). Downstream of helix Ib in the U2 snRNA strand (i.e., equivalent to upstream in the U6 snRNA strand), U2 pairs with the BP sequence, with the BP adenosine bulged out of the branch helix.

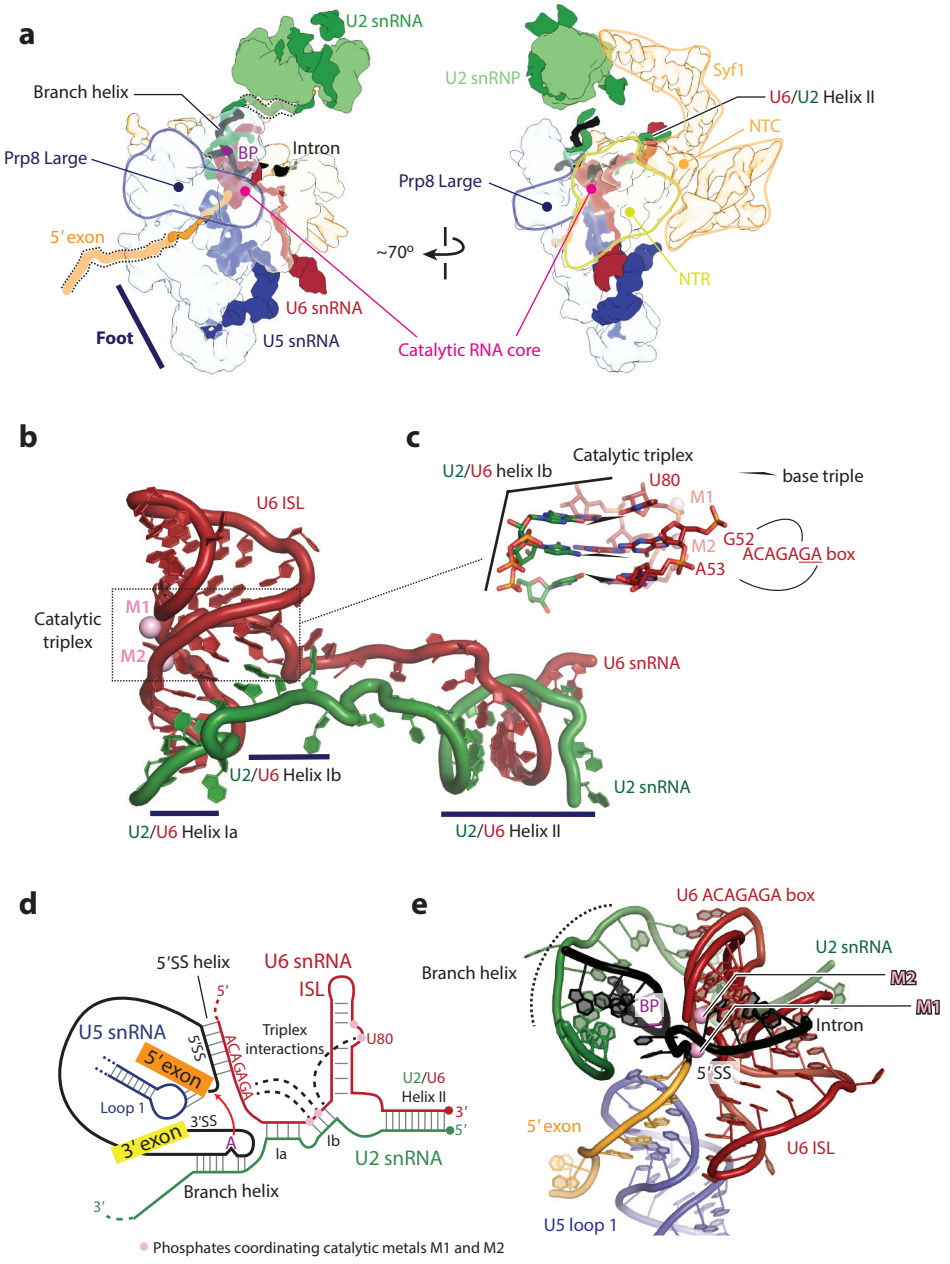
U5 snRNA loop 1 protrudes into the active site and pairs with the 5′ exon sequence immediately upstream of the 5′SS (60, 61) (**Figure 5*d,e***). This contact forms prior to branching and persists through exon ligation, tethering the 5′ exon in place. The active site structure formed by the snRNAs remains remarkably static from the B^{act} complex to the intron-lariat spliceosome (ILS) complex, as the RNA structure is stabilized by numerous proteins, with the exception of dramatic movements of the branch helix between these states (66).

The Active Site RNA Is Cradled by NTC and NTR Proteins

The RNA core of the spliceosomes in the catalytic phase (B^{act} to ILS complex) is cradled by a common protein shell comprising the U5 and U2 snRNP proteins, the Prp19-associated

Supplemental Material >

complex (NTC) and the Prp19-related complex (NTR) (92–94) (**Figure 5*a***; **Supplemental Video 4**). The concave surface of the Large domain of Prp8 comprising the linker and HB domains accommodates the catalytic RNA core, while the NTC and NTR proteins constrain the RNA core onto Prp8, stabilizing the active site RNA (**Figure 5*a***). The NTR proteins (Cwc2,



Spliceosome active site organization (B* stage)

(*Caption appears on following page*)

Figure 5 (*Figure appears on preceding page*)

The structure of the catalytic core of the spliceosome. (*a*) The catalytic RNA core is formed between the Prp8 Large domain and the NTC and NTR proteins. The catalytic core is formed in the B^{act} complex and remains unchanged during the catalytic phase of the splicing cycle ($B^{act} \rightarrow B^* \rightarrow C \rightarrow C^* \rightarrow P \rightarrow ILS$). This example structure shows the yeast P complex (87). (*b*) The catalytic RNA core of the spliceosome. U6 snRNA forms the ISL and helices Ia and Ib with U2 snRNA. The U6 snRNA ACAGAGA box preceding helix Ia folds back and forms two triplex pairs to helix Ib. U80 (yeast numbering), a bulged nucleotide in the ISL, forms a third triplex pair to helix Ia. These three stacked triple base pairs create binding sites for two catalytic metal ions (see also **Figure 1*c***). (*c*) Detail of the catalytic triplex. (*d*) Secondary structures of the snRNAs and substrate around the catalytic core during branching, the first reaction. Tertiary interactions forming the catalytic triplex are indicated (*dashed lines*); these interactions bring together U6 phosphates (*pink circles*) to coordinate the catalytic metal ions. (*e*) Structure of the catalytic core during branching (29, 30, 91). Abbreviations: BP, branch point; ILS, intron-lariat spliceosome; ISL, internal stem-loop; NTC, Prp19-associated complex; NTR, Prp19-related complex; snRNA, small nuclear RNA; snRNP, small nuclear ribonucleoprotein; SS, splice site.

Ecm2, Bud31, Cwc15, Prp45, Prp46; in humans, Cwc2 and Ecm2 are fused into RBM22) form both a unit that binds the 5′ end of U6 snRNA as it exits the active site and a series of loops that wind around the RNAs of the active site. In contrast, the NTC proteins (Prp19, Cef1, Clf1, Syf1, Syf2, Ntc20, Snt309) are all almost entirely alpha-helical and form the opposite side of the active site cavity to the Prp8 Large domain, closely binding the U2/U6 snRNA structures of the active site and U2/U6 helix II (**Figure 5*a***). The NTC proteins Syf1 and Clf1 form a remarkable series of helical arches that are a striking peripheral feature of the catalytic spliceosomes (**Figure 5*a***). The U2 snRNP core domain is connected to the spliceosome via the N terminus of the Syf1 helical arch and via the branch helix (**Figure 5*a***). The position of U2 snRNP is variable, tracking the substantial movements of the branch helix during splicing and in turn transmitting these movements to Syf1 and the NTC (**Supplemental Videos 1–7**).

Supplemental Material >

B^{act} Complex: Activated to Perform Splicing but Inhibited

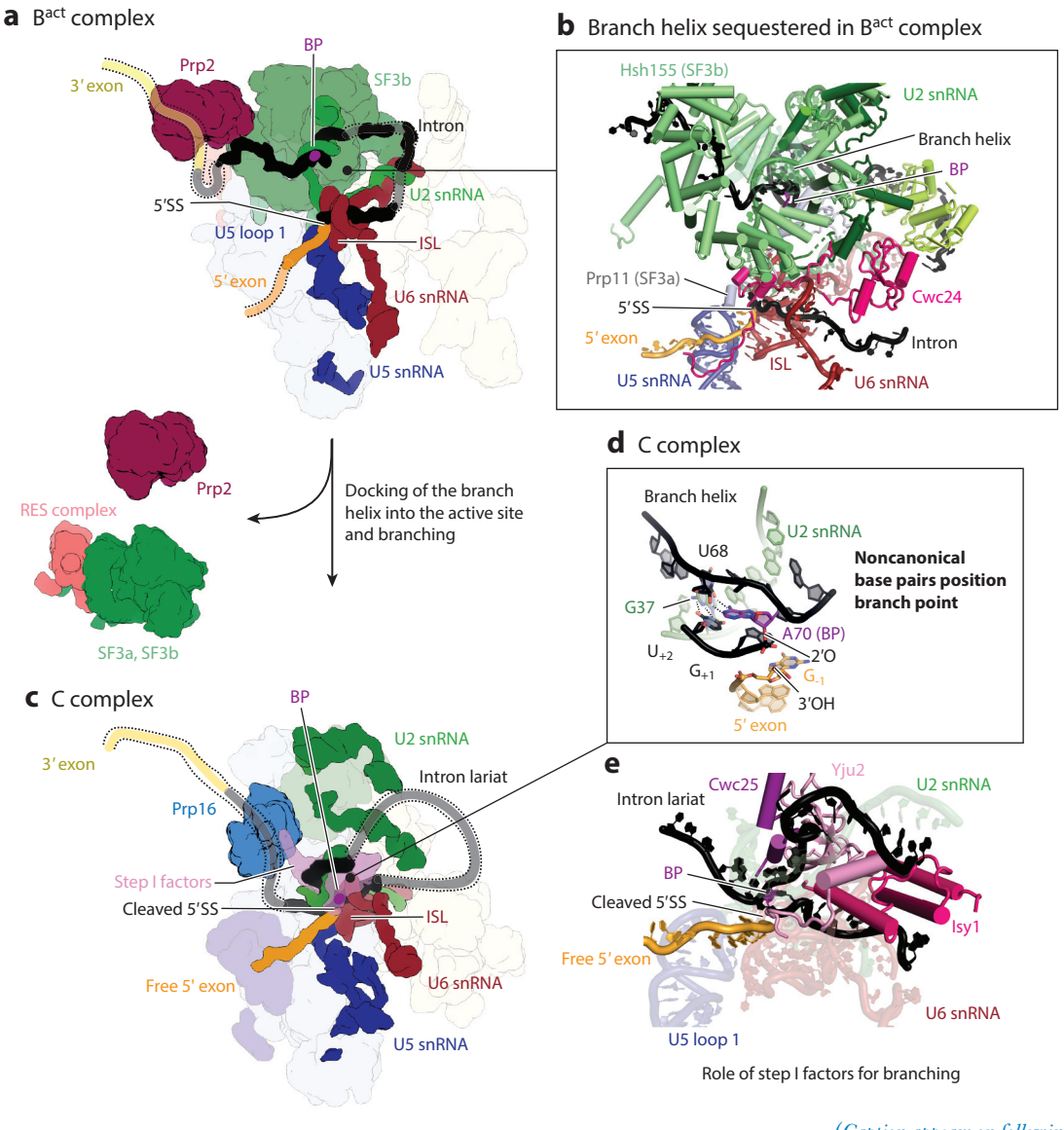
As described above, the B^{act} complex is the first spliceosome intermediate with a fully formed RNA catalytic core (**Figures 1*d*** and **6*a***; **Supplemental Video 4**). The 5′SS is properly loaded into the active site, but the 2′OH group of the BP adenosine, which acts as the nucleophile for the branching reaction, is 50 Å away from the active site as the branch helix is encapsulated by Hsh155 (SF3B1 in humans) within the SF3b complex (81, 82, 95, 96) (**Figure 6*a*,*b***; **Supplemental Video 4**). The 5′SS is also protected by the SF3a protein Prp11 (SF3A2 in humans) and the B^{act} complex–specific factor Cwc24 (RNF113A in humans), which shields 5′SS $G_{+1}U_{+2}$ with a compact zinc finger (97) (**Figure 6*b***; **Supplemental Video 4**). Hence, the two reactants for the branching reaction cannot reach each other at the active site. This inhibited conformation is stabilized by Brr2, which binds SF3b via its N-terminal helicase cassette and thus retains SF3b on the spliceosome. SF3b is also stabilized by the trimeric pre-mRNA REtention and Splicing (RES) complex (Bud13, Snu17, and Pml1) that bridges SF3b to the Prp8 Large domain (98).

REMODELING THE B^{act} COMPLEX TO THE BRANCHING B* AND C COMPLEXES

To release this inhibition and enable branching, the contacts described above must be disrupted, thus allowing docking of the BP adenosine into the active site (**Supplemental Video 5**). This remodeling is achieved by Prp2, the first of four DEAH-box ATPases to act on the spliceosome, the others being Prp16, Prp22, and Prp43 (99). Members of this family of ATPases contain two RecA-like domains that form a helicase core, which can bind RNA and hydrolyze ATP (or other

NTPs), and a C-terminal domain that also contributes to RNA binding. These ATPases bind to single-stranded 3′ overhangs and can translocate in the 3′ to 5′ direction in vitro to disrupt RNA duplexes or protein–RNA interactions (100). Prp2, Prp16, Prp22, and Prp43 all localize to the periphery in solved spliceosome structures and tend to be rather mobile and thus poorly resolved in the cryo-EM maps (66). It is not yet clear precisely how these ATPases effect spliceosome remodeling. A plausible unifying mechanism is that they all act from a distance by pulling on RNA, rather than translocating through RNA–protein contacts (101).

Prp2 is located downstream of the BP adenosine and probably relieves branch helix sequestration by pulling on the substrate intron, disrupting interactions of the RNA with SF3b (81, 82)



(*Caption appears on following page*)

Figure 6 (*Figure appears on preceding page*)

Docking of the branch helix for the branching reaction. (*a*) Schematic overview of the yeast B^{act} complex (81, 82). The catalytic RNA core is already formed, and the 5′SS is accommodated in the active site. The branch helix is sequestered by SF3b within the U2 snRNP, keeping the 2′OH nucleophile of the BP adenosine 50 Å away from the active site. (*b*) Detailed view of the B^{act} complex active site. In the B^{act} complex, the 5′SS is protected by Cwc24 and Prp11, and Hsh155 (SF3B1) sequesters the branch helix within SF3b. (*c*) Upon Prp2-induced dissociation of the RES complex, SF3a, and SF3b, the branch helix docks into the active site with the help of step I (branching) factors Yju2, Cwc25, and Isy1. The nucleophile 2′OH group of the BP adenosine attacks the 5′SS, producing the cleaved 5′ exon and the lariat–3′ exon intermediate within the C complex (29, 30). (*d*) The BP adenosine (A70) base-pairs with U68 to allow the 2′O of A70 to attack the 5′SS. As a result, the phosphate of the first intron nucleotide G_{+1} is linked to the 2′OH group of the BP adenosine (A70) (29). In the recent B* complex structure (91), the branch helix was not properly docked into the active site in the absence of Cwc25. The BP adenosine was proposed to form a Watson-Crick base pair with the invariant U_{+2} of the 5′SS, but the cryo-EM density does not exclude the possibility of the BP adenosine pairing with U68 as in the C complex. (*e*) The branching factors Yju2, Cwc25, and Isy1 wrap around the branch helix, fully docking it into the active site to allow branching chemistry. Abbreviations: BP, branch point; cro-EM, cryo–electron microscopy; ISL, internal stem-loop; RES complex, REtention and Splicing complex; snRNA, small nuclear RNA; SS, splice site.

(**Figure 6*a***). Prp2 activity results in destabilization of the U2-associated complexes, SF3a and SF3b, leaving behind only the U2 snRNP core domain (102). Prp2 action probably also directly leads to dissociation of the RES complex, which also binds the intron downstream of the BP. Prp2 itself also dissociates after stimulating remodeling. With the release of SF3b, the branch helix is free to dock next to the 5′SS, where it would displace Cwc24 (**Supplemental Video 5**).

Supplemental Material >

THE FIRST CATALYTIC STEP WITHIN B* COMPLEX TO FORM C COMPLEX

The result of B^{act} complex remodeling by Prp2 is the B* complex, in which the branch helix is docked and the active site is fully competent to catalyze the branching reaction, producing the C complex (**Figure 5*e***). Despite their different names, the B* and C complexes are essentially equivalent and representative of the branching conformation of the spliceosome, differing only in whether the phosphate group of the first intron nucleotide (G_{+1}) is linked to the 3′OH group of the 5′ exon (B* complex) or to the 2′OH of the BP adenosine (C complex) (**Figure 1*c***). In the cryo-EM structure of the yeast C complex (29, 30), the 3′OH group of the cleaved 5′ exon and the 5′ phosphate of the first intron nucleotide remain close to the catalytic metal ions (**Figure 6*d***), suggesting that the pre-branching structure (B* complex) can be restored by a small structural rearrangement (**Figure 1*c***). Chemistry in the B* complex is promoted by the branching factors Yju2, Cwc25, and the NTC component Isy1. In the absence of these branching factors, the branch helix in the B* complex is mobile and does not stably dock the BP adenosine into the active site (91). When present in the B* or C complex, Yju2, with the aid of Isy1, clamps around the branch helix and distorts it from the canonical A form, aiding its docking into the narrow active site and explaining why Yju2 is required for catalysis (103, 104) (**Figure 6*c*,*e***; **Supplemental Video 5**). However, the structure of the B* complex (91), which lacks Cwc25, shows that branch helix docking alone is not sufficient for chemistry, as the 2′OH of the BP adenosine is still ∼4 Å from the scissile phosphate at the 5′SS. Comparison with structures of the C complex (29, 30) suggests that Cwc25 is further required to rigidify the branch helix and push the intron sequences around the BP such that the BP adenosine 2′OH can attack the 5′SS (102, 105) (**Supplemental Video 5**).

In B* and C complexes, the position of the nucleophile 2′OH is reinforced by a non–Watson-Crick base pair between the BP adenosine and the U68 two nucleotides upstream (UAC**U**AA**C**) (29, 30) (**Figure 6*c***). This positioning creates a backbone conformation such that the 2′OH can project toward the 5′SS site, explaining why adenosine is preferred as the BP nucleotide for the

branching reaction (106, 107). The invariant U of the second intron nucleotide (U_{+2}) hydrogen bonds with U2 snRNA G37 and plays a crucial role in positioning the 5′SS at the catalytic metals (**Figure 6*c***).

REMODELING FROM THE BRANCHING CONFORMATION TO EXON-LIGATION CONFORMATION

For the second step of splicing, exon ligation, the same active site is used (**Figure 1*c***). This requires displacement of the branch helix after the first step during remodeling to a distinct exon-ligation conformation of the spliceosome (29, 30, 85, 86) (**Figure 7*a*,*b***; **Supplemental Video 6**). The DEAH-box ATPase Prp16 promotes this remodeling (108). Prp16 binds the intron downstream of the BP (29, 109), and its activity is associated with translocation toward the branch helix that results in displacement of the branching factors Yju2 and Cwc25 and dissociation of Prp16 from the spliceosome (101, 106) (**Figure 7*a*,*b***). The structure of the B* complex suggests that the branch helix is not docked into the active site in the absence of Yju2 and Cwc25. However, in the postreaction state, the branch helix is still tethered via a 2′–5′ phosphodiester bond from the BP adenosine to the 5′SS. Therefore, the structural consequence of Prp16 activity is rotation of the branch helix by 75° to vacate a space in the active site for docking of the 3′SS (85, 86, 110) (**Figure 7*a*,*b***; **Supplemental Video 6**). To stabilize the inherently mobile branch helix in its new position, the β-propeller of Prp17, an exon-ligation factor pre-bound by an N-terminal anchor region to the U6 snRNA 5′ stem-loop, moves inward to buttress the branch helix against the NTC, and the Prp8 RNaseH domain rotates to clasp the branch helix using a long β-hairpin (**Supplemental Video 6**). The branch helix reverts to a canonical A-form helix when undocked. Branch helix movement is associated with a shift of the linked U2 snRNP core and a concerted movement of the helical arches of the NTC, which contact U2 snRNP (**Supplemental Video 6**). During Prp16-mediated remodeling, subtle restructuring of the pairing between the 5′SS and U6 snRNA ACAGAGA box results in formation of a noncanonical pair between 5′SS U_{+2} and A51 (ACAG**A**GA) (85, 86), which creates a landing platform for the 3′SS G_{-1} to stack on A51 (61, 87–89) (**Figure 7*c***).

EXON LIGATION WITHIN C* COMPLEX TO FORM THE P COMPLEX

The resultant spliceosome after Prp16-mediated remodeling adopts the exon-ligation conformation (**Figure 7*b***). Before exon ligation, it is referred to as the C*complex; after exon ligation, it becomes the postcatalytic spliceosome (P complex) (**Figure 7*c***). The active site in the C* complex has a vacancy for the 3′SS (85, 86), which can dock with the aid of the exon-ligation factors Prp18 and Slu7 (101, 111) (**Figure 7*d***; **Supplemental Video 6**). Slu7 binds the exon-ligation spliceosome, aided by Prp18, and promotes 3′SS docking, while Prp18 may independently stabilize pairing of the exons to U5 snRNA loop 1 (112, 113).

Virtually all introns from eukaryotes end with an AG dinucleotide, almost always (>95% of the time) preceded at the −3 position by a pyrimidine (1, 7). Mutations at any of these three positions can block splice-site usage (113). This short YAG motif has limited information content and is usually selected by being the first such motif more than 10 nt downstream of the BP (113).

Recognition of the 3′SS AG dinucleotide is a major exception to the general rule that intron sequences are recognized by base-pairing to snRNAs. Instead, when the 3′SS docks into the core of the C* complex, it is recognized by non–Watson-Crick base-pairing to intron nucleotides in the 5′SS and BP adenosine (87, 88). The 3′SS G_{-1} pairs with G_{+1} of the 5′SS (114), and 3′SS A_{-2} pairs with the BP adenosine (**Figure 7*c***; **Supplemental Video 6**). As 5′SS G_{+1} and the BP

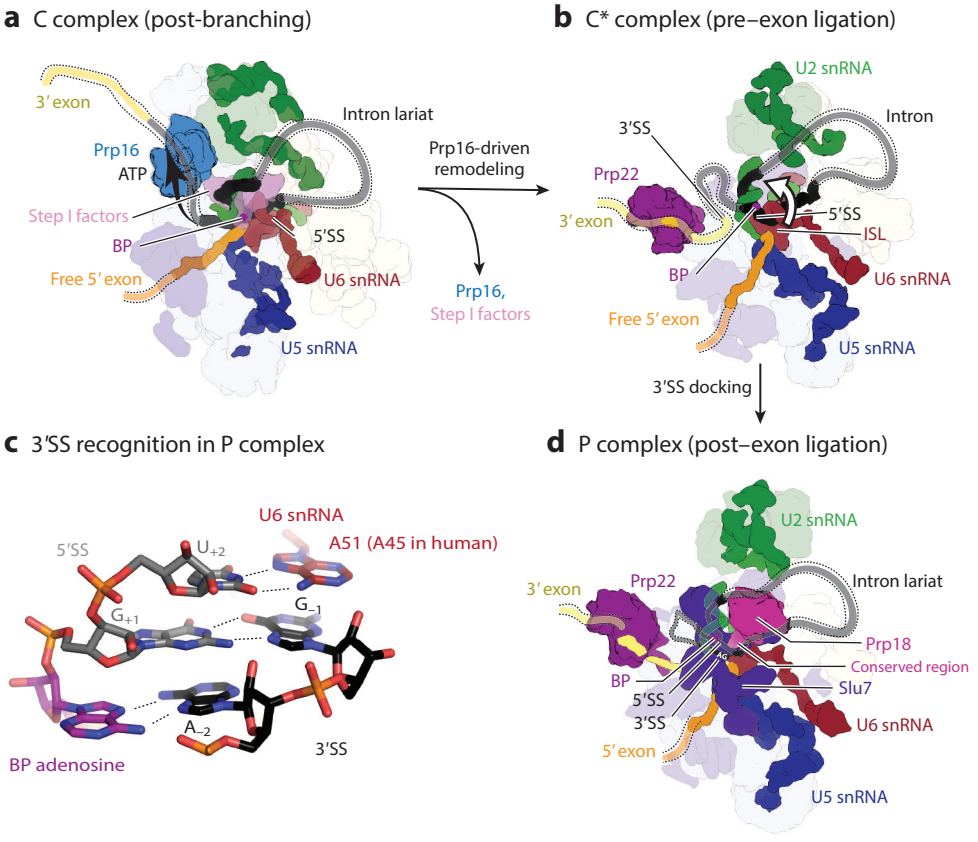


Figure 7

Remodeling to the exon-ligation conformation. (*a*) Schematic of yeast C complex (29, 30) as in **Figure 6*c***. (*b*) ATP-driven remodeling by Prp16 helicase to the C* complex results in loss of step I factors and Prp16, and as a consequence the branch helix rotates out of the active site (85, 86, 110). Prp22 replaces Prp16 downstream of the BP. In the absence of step II factors, the 3′SS does not dock into the active site (87, 101). (*c*) When docked, the 3′SS stacks on the noncanonical pair between U6 snRNA A51 and 5′SS U_{+2}. The 3′SS is recognized by pairing to the 5′SS G_{+1} and BP adenosine, as shown by two structures of the yeast P-complex spliceosome (87, 88). In a third P-complex structure (89), the interaction between 3′SS A_{-2} and the BP adenosine is different from that observed in the other two structures. (*d*) Schematic of the P-complex spliceosome (87) showing binding of the step II factors Prp18 and Slu7 to the periphery of the spliceosome, from where Prp18 can insert a conserved loop to contact the 3′SS (87, 88). In contrast, the third P-complex structure (89) lacks the Prp18 conserved loop, which may explain why the BP adenosine does not form a Hoogsteen base pair with 3′SS A_{-2} in this structure. Abbreviations: BP, branch point; ISL, internal stem-loop; snRNA, small nuclear RNA; SS, splice site.

adenosine are covalently linked by the branching reaction, this means branching creates a structure in the intron—the branched lariat—utilized by the spliceosome during exon ligation for splice-site recognition. This mechanism explains the absolute conservation of the GU and AG dinucleotides at the 5′ and 3′ splice sites and the identity of the BP nucleotide, as all of these positions form non–Watson-Crick base pairs during exon ligation. The only exception is the AT-AC class of introns spliced by the minor spliceosome, in which the 5′SS G_{+1} to 3′SS G_{-1} is replaced by an almost isosteric 5′SS A_{+1} to 3′SS C_{-1} pair (87, 115). Conservation of the GU dinucleotide at the 5′ end

of the intron and the BP adenosine is also explained by their interactions during the branching stage, as explained above.

Compared with the branching factors, which influence reactant docking very directly, the exon-ligation factors promote chemistry in a more subtle way (**Figure 7*d***). The most conserved exon-ligation factors are Prp17, Prp18, and Slu7, of which Slu7 is the only essential gene in yeast (113). Prp17 most likely stabilizes the exon-ligation conformation of the spliceosome over the branching conformation, driving the equilibrium of Prp16-mediated remodeling forward (116). Prp18 and Slu7 bind each other and are visible on the periphery of the C*- and P-complex structures (85–89). Prp18 binds the Prp8 RNaseH domain and inserts a conserved loop through a channel between the Prp8 Large and RNaseH domains to buttress the docked 3′SS against the active site (87, 88) (**Figure 7*d***; **Supplemental Video 6**). This loop is proposed to be especially important for exon ligation in pre-mRNAs with weak pairing to U5 snRNA (113). Slu7, by contrast, is sprawling and largely consists of loops binding exposed surfaces of Prp8, anchored by its zinc knuckle domain near Cwc22. Many Slu7 loops line the spliceosomal cavity through which the 3′SS would dock after Prp16-mediated remodeling, consistent with a postulated role in promoting 3′SS docking (101, 117, 118), although the precise mechanism for this is unclear. Prp17, Prp18, and Slu7 are the only identified exon-ligation factors in yeast, but a number of additional factors have recently been identified in humans (**Figure 8**; see the sidebar titled Exon Ligation in the Human Spliceosome).

Supplemental Material >

RELEASE OF THE LIGATED EXONS

The DEAH-box ATPase Prp22 catalyzes the release of the ligated exons (119, 120). Prp22 interacts with nucleotides upstream of the 3′SS before exon ligation (109) and downstream of the 3′SS within the 3′ exon after exon ligation (87, 120) (**Figure 7*a*,*b***; **Supplemental Video 7**). Upon ATP hydrolysis, Prp22 translocates along the mRNA with 3′ to 5′ directionality, leading to release of the mRNA and loss of exon-ligation factors Prp18 and Slu7 from the spliceosome (112, 120) (**Figure 9**; **Supplemental Video 7**). In metazoans, mRNA released from the spliceosome contains the exon-junction complex (EJC) upstream of the exon junction. The EJC is recruited to the 5′ exon by the integral spliceosome component Cwc22, which forms two helical bundles on either side of a channel in Prp8 for the 5′ exon (**Figure 8*b*,*c***). The N-terminal bundle, an MIF4G domain, binds Snu114 and the EJC component eIF4AIII (29, 121–124) (**Figure 8*c***). The position of Cwc22 acts as a molecular ruler for deposition of the EJC 20–24 nucleotides upstream of exon junctions (29), a positioning that is important for the functional roles of the EJC in nonsense-mediated mRNA decay (125). The EJC is bound to Cwc22 and possibly the 5′ exon in structures of human C, C*, and P complexes (110, 124, 126–128). During Prp22-mediated mRNA release, the EJC is released and bound to the 5′ exon, along with Cwc22. Because Prp22 does not translocate through the mRNA-U5 snRNA loop 1 duplex to release the mRNA (101) and the ligated mRNA in P complex is tightly bound by many contacts to the U5 snRNP and the EJC, the precise mechanism of mRNA release is currently not clear.

DISASSEMBLY OF THE SPLICEOSOME

The spliceosome state after Prp22-mediated release of the ligated exons (mRNA), the ILS, must be disassembled to allow release and decay of the intron lariat and recycling of the snRNAs and associated factors for subsequent rounds of splicing (**Figure 9**). Disassembly is carried out by the versatile helicase Prp43, which functions both in splicing (129) and ribosome biogenesis (130, 131) (**Supplemental Video 7**). The specific utilization of Prp43 in splicing cycle termination is mediated by the Ntr1 complex (not to be confused with the NTR complex);

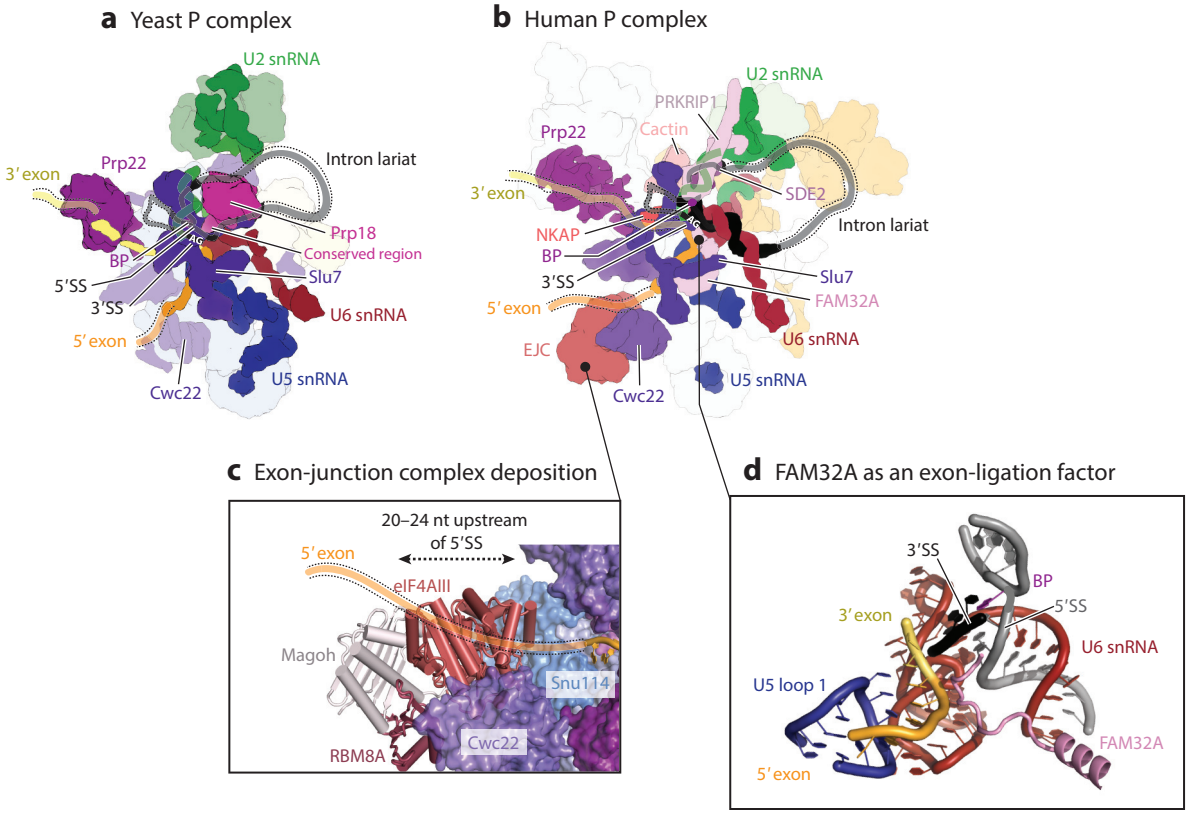


Figure 8

The human spliceosome adopts a more complex solution for exon ligation. (*a*) Schematic of the yeast P complex (87) as in **Figure 7*d***. (*b*) Similar view of the human P complex (127) showing several factors not found in yeast that bind near the 3′SS and branch helix. (*c*) The human spliceosome recruits the EJC to the 5′ exon via Cwc22, which binds the exon channel and measures a distance of 20–24 nt upstream of the exon junction. (*d*) FAM32A is an exon-ligation factor that inserts a C-terminal peptide between the 3′ exon and 3′SS to directly promote exon ligation. Abbreviations: BP, branch point; EJC, exon-junction complex; snRNA, small nuclear RNA; SS, splice site.

the Ntr1 complex consists of Ntr1/Spp382, Ntr2, and Cwc23 and can stably bind Prp43 (132, 133). The Ntr1 N-terminal G-patch domain binds Prp43 and stimulates its helicase activity, resulting in disassembly of the spliceosome into the U6 snRNA, U2 snRNP core, U5 snRNP, and individual NTC proteins (134) (**Supplemental Video 7**). Recruitment and activation of the disassembly machinery are finely regulated to restrict its irreversible action to terminal stages of the spliceosome cycle—either postcatalytic spliceosomes or spliceosomes harboring suboptimal substrates that are rejected by proofreading (135).

Supplemental Material >

The exact structure targeted by Prp43 to trigger spliceosome disassembly has been unclear. Recent biochemical evidence points to two potentially non-mutually exclusive RNA targets: disruption of the branch helix by translocating on the intron (136) and/or disruption of U2/U6 snRNA interactions by translocating from the 3′ end of U6 snRNA (137, 138). Structures of both *S. cerevisiae* and human ILS complex contain Prp43 (90, 128), and Prp43 binds the extended arch formed by the NTC protein Syf1, proximal to U6/U2 helix II (**Figure 9*b***; **Supplemental Video 7**). Prp43 is therefore ideally positioned to bind the free 3′ end of U6 snRNA to initiate disassembly, consistent with Prp43 disrupting U2/U6 snRNA interactions. However, it is not

EXON LIGATION IN THE HUMAN SPLICEOSOME

Human spliceosomes contain many more protein components in addition to the orthologs of yeast proteins (**Figure 8*a*,*b***), but the roles of these additional factors have mostly remained unclear (144, 145). A recent high-resolution structure of the human P complex showed that the 3′SS recognition mechanism is conserved in humans but identified four new proteins important for exon ligation: Cactin, SDE2, FAM32A, and NKAP (127) (**Figure 8*b***). However, Prp18, essential for exon ligation in yeast, was not bound in the active site. This finding is highly surprising, as the exon–ligation reaction of several pre-mRNAs including globin pre-mRNA is inhibited by depletion of Prp18 from HeLa cell nuclear extract (146). FAM32A was found to interact with Slu7 and insert its conserved C terminus into the active site to bridge the 5′ and 3′ exons and the 3′SS (**Figure 8*d***), and its depletion was shown to inhibit exon ligation, thus identifying it as a novel exon-ligation factor (127). FAM32A is also known as OTAG-12 (ovarian tumor associated gene-12), which promotes splicing of proapoptotic genes, thereby acting as a tumor suppressor (147). This finding suggests that splicing of different pre-mRNAs is regulated by different sets of exon-ligation factors in humans, providing the opportunity for tissue-specific regulation of alternative splicing. Cactin, SDE2, and NKAP interact with the branch helix and stabilize the exon-ligation conformation of the spliceosome (**Figure 8*b***). The orthologs of these proteins are known to regulate alternative splicing in *Schizosaccharomyces pombe*, so they may operate similarly in humans (148, 149).

possible to trace these RNAs to Prp43 in the cryo-EM maps; hence, these structures cannot exclude either possibility (90, 128). In either case, pulling on these RNA elements would trigger an unknown cascade of events that would lead to spliceosome disassembly.

The interface of Prp43 on Syf1 is available in all spliceosomes containing the NTC, as is its potential substrate, the 3′ end of U6 snRNA. However, the Ntr1 complex has numerous binding interfaces on the catalytic spliceosome that are variously occluded by SF3b in the B^{act} complex,

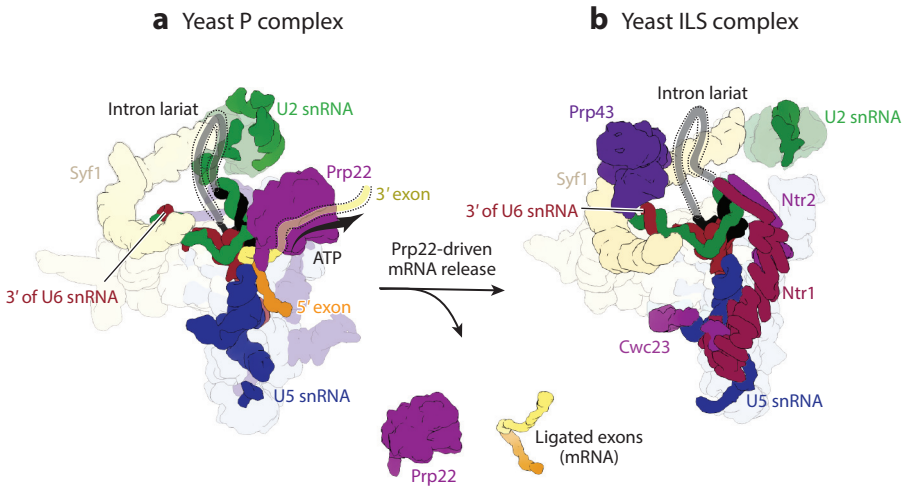


Figure 9

Product release and spliceosome disassembly. (*a*) Schematic of yeast P complex (87). Prp22 helicase pulls on the 3′ exon, triggering release of the ligated exons as mRNA. (*b*) Schematic of the resultant ILS complex (90). The Ntr1 complex (Ntr1, Ntr2, Cwc23) binds to stimulate Prp43 activity. Prp43 is near the 3′ end of U6 snRNA, which is its likely substrate for triggering spliceosome disassembly (138). Abbreviations: ILS, intron-lariat spliceosome; mRNA, messenger RNA; snRNA, small nuclear RNA.

Prp16 in the B* and C complexes, and Prp22 in the C* and P complexes and only become fully available once Prp22 leaves following completion of catalysis. Therefore, this coactivator complex serves in part to prevent premature activity of Prp43 on spliceosomes still in the process of catalysis.

A notable exception is the kinetic proofreading that can occur on the spliceosome. Both Prp16 and Prp22, in addition to their canonical roles in remodeling the spliceosome, can promote an ATP-driven cycle of sampling alternative BPs or 3′SSs within spliceosomes that fail to perform catalysis within a kinetic window (101, 139). This cycle is broken either by successful catalysis or by a discard pathway driven by Prp43 (135, 140). In this context, Prp43-mediated disassembly of earlier spliceosome intermediates could be licensed by the Ntr1 complex competing with Prp16 or Prp22 for binding to the spliceosome and activating disassembly. Dissociation of Prp16 or Prp22 could thus be a signal of a stalled spliceosome that requires disassembly (141).

FUTURE ISSUES

1. Splicing is cotranscriptional (142). The speed of RNA polymerase can directly influence splicing efficiency and splice-site choice, and the polymerase C-terminal domain (CTD) probably plays an active role in cotranscriptional spliceosome assembly. How direct the interactions are between RNA polymerase and the spliceosome requires further investigation.
2. Since its initial discovery in the yeast C* complex (85), several cryo-EM structures of the spliceosome have shown the presence of a small-molecule ligand, inositol hexaphosphate, bound to Prp8. In yeast, this ligand joins during formation of the B^{act} complex, whereas in humans, it is already bound in tri-snRNP. During exon ligation, it additionally contacts Slu7. The functional relevance of this ligand has yet to be tested.
3. Cryo-EM structures of the spliceosome provide vivid snapshots of splicing, but the transitions between these states are not well understood. The videos accompanying this review show possible ways of interpolating between these states, but many routes are available. Further experiments (e.g., by time-resolved cryo-EM or single-molecule methods) (94) are required to dissect the order of events during spliceosome remodeling.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Sebastian Fica, Andy Newman, Charlotte Capitanchik, Morgan Jones, and Martin Jinek for helpful discussion and comments on the manuscript. We thank Christine Norman for narrating the accompanying videos. K.N. thanks the past and present members of the Nagai group who have worked on the spliceosome project since 1989. We thank Reinhard Lührmann and his colleagues for their inspirational and rigorous pioneering work. This work was supported by the Medical Research Council (MC_U105184330) and a European Research Council (ERC) Advanced Grant (AdG-693087-SPLICE3D). M.E.W. was supported by a Cambridge–Rutherford Memorial PhD scholarship, and C.C. was supported by an ERC grant, a European Molecular Biology Organization (EMBO) long-term fellowship, and a Marie Skłodowska-Curie individual fellowship.

LITERATURE CITED

1. Spingola M, Grate L, Haussler D, Ares M. 1999. Genome-wide bioinformatic and molecular analysis of introns in *Saccharomyces cerevisiae*. *RNA* 5(2):221–34
2. Qin D, Huang L, Wlodaver A, Andrade J, Staley JP. 2016. Sequencing of lariat termini in *S. cerevisiae* reveals 5′ splice sites, branch points, and novel splicing events. *RNA* 22(2):237–53
3. Hong X, Scofield DG, Lynch M. 2006. Intron size, abundance, and distribution within untranslated regions of genes. *Mol. Biol. Evol.* 23(12):2392–404
4. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40(12):1413–15
5. Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, et al. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456(7221):470–76
6. Ule J, Blencowe BJ. 2019. Alternative splicing regulatory networks: functions, mechanisms and evolution. *Mol. Cell* 76:329–45
7. Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R. 2006. Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res.* 34(14):3955–67
8. Padgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA. 1984. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* 225(4665):898–903
9. Ruskin B, Krainer AR, Maniatis T, Green MR. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* 38(1):317–31
10. Rodriguez JR, Pikielny CW, Rosbash M. 1984. In vivo characterization of yeast mRNA processing intermediates. *Cell* 39(3, Part 2):603–10
11. Domdey H, Apostol B, Lin RJ, Newman AJ, Brody E, Abelson J. 1984. Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. *Cell* 39(3, Part 2):611–21
12. Brody E, Abelson J. 1985. The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* 228(4702):963–67
13. Steitz TA, Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. *PNAS* 90(14):6498–502
14. Fica SM, Tuttle N, Novak T, Li N-S, Lu J, et al. 2013. RNA catalyses nuclear pre-mRNA splicing. *Nature* 503(7475):229–34
15. Kastner B, Will CL, Stark H, Lührmann R. 2019. Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring Harb. Perspect. Biol.* 11(11):a032417
16. Reddy R, Busch H. 1988. Small nuclear RNAs: RNA sequences, structure, and modifications. In *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, ed. ML Birnstiel, pp. 1–37. Berlin: Springer
17. Lerner MR, Steitz JA. 1979. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *PNAS* 76(11):5495–99
18. Bringmann P, Lührmann R. 1986. Purification of the individual snRNPs U1, U2, U5 and U4/U6 from HeLa cells and characterization of their protein constituents. *EMBO J.* 5(13):3509–16
19. Leung AKW, Nagai K, Li J. 2011. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* 473(7348):536–39
20. Séraphin B. 1995. Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *EMBO J.* 14(9):2089–98
21. Achsel T, Brahms H, Kastner B, Bachi A, Wilm M, Lührmann R. 1999. A doughnut-shaped heteromer of human Sm-like proteins binds to the 3′-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. *EMBO J.* 18(20):5789–802
22. Zhou L, Hang J, Zhou Y, Wan R, Lu G, et al. 2013. Crystal structures of the Lsm complex bound to the 3′ end sequence of U6 small nuclear RNA. *Nature* 506(7486):116–20
23. Plaschka C, Newman AJ, Nagai K. 2019. Structural basis of nuclear pre-mRNA splicing: lessons from yeast. *Cold Spring Harb. Perspect. Biol.* 11(5):a032391
24. Yan C, Wan R, Shi Y. 2019. Molecular mechanisms of pre-mRNA splicing through structural biology of the spliceosome. *Cold Spring Harb. Perspect. Biol.* 11(1):a032409

25. Staley JP, Guthrie C. 1999. An RNA switch at the 5′ splice site requires ATP and the DEAD box protein Prp28p. *Mol. Cell* 3(1):55–64
26. Laggerbauer B, Achsel T, Lührmann R. 1998. The human U5–200kD DEXH-box protein unwinds U4/U6 RNA duplices in vitro. *PNAS* 95(8):4188–92
27. Raghunathan PL, Guthrie C. 1998. RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. *Curr. Biol.* 8(15):847–55
28. Hang J, Wan R, Yan C, Shi Y. 2015. Structural basis of pre-mRNA splicing. *Science* 349(6253):1191–98
29. Galej WP, Wilkinson ME, Fica SM, Oubridge C, Newman AJ, Nagai K. 2016. Cryo-EM structure of the spliceosome immediately after branching. *Nature* 537(7619):197–201
30. Wan R, Yan C, Bai R, Huang G, Shi Y. 2016. Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science* 353(6302):895–904
31. Plaschka C, Lin P-C, Charenton C, Nagai K. 2018. Prespliceosome structure provides insights into spliceosome assembly and regulation. *Nature* 559(7714):419–22
32. Zhuang Y, Weiner AM. 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell* 46(6):827–35
33. Wong MS, Kinney JB, Krainer AR. 2018. Quantitative activity profile and context dependence of all human 5′ splice sites. *Mol. Cell* 71(6):1012–13
34. Pomeranz Krummel DA, Oubridge C, Leung AKW, Li J, Nagai K. 2009. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* 458(7237):475–80
35. Kondo Y, Oubridge C, van Roon A-MM, Nagai K. 2015. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5′ splice site recognition. *eLife* 4:360
36. Neubauer G, Gottschalk A, Fabrizio P, Séraphin B, Lührmann R, Mann M. 1997. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *PNAS* 94(2):385–90
37. Li X, Liu S, Jiang J, Zhang L, Espinosa S, et al. 2017. CryoEM structure of *Saccharomyces cerevisiae* U1 snRNP offers insight into alternative splicing. *Nat. Commun.* 8(1):1035
38. Bai R, Wan R, Yan C, Lei J, Shi Y. 2018. Structures of the fully assembled *Saccharomyces cerevisiae* spliceosome before activation. *Science* 360(6396):1423–29
39. Puig O, Bragado-Nilsson E, Koski T, Séraphin B. 2007. The U1 snRNP-associated factor Luc7p affects 5′ splice site selection in yeast and human. *Nucleic Acids Res.* 35(17):5874–85
40. Zamore PD, Patton JG, Green MR. 1992. Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* 355(6361):609–14
41. Berglund JA, Abovich N, Rosbash M. 1998. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev.* 12(6):858–67
42. Wang Q, Zhang L, Lynn B, Rymond BC. 2008. A BBP-Mud2p heterodimer mediates branchpoint recognition and influences splicing substrate abundance in budding yeast. *Nucleic Acids Res.* 36(8):2787–98
43. Kistler AL, Guthrie C. 2001. Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase. *Genes Dev.* 15(1):42–49
44. Liang W-W, Cheng S-C. 2015. A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. *Genes Dev.* 29(1):81–93
45. Perriman R, Ares M. 2010. Invariant U2 snRNA nucleotides form a stem loop to recognize the intron early in splicing. *Mol. Cell* 38(3):416–27
46. Krämer A, Grüter P, Gröning K, Kastner B. 1999. Combined biochemical and electron microscopic analyses reveal the architecture of the mammalian U2 snRNP. *J. Cell Biol.* 145(7):1355–68
47. Price SR, Evans PR, Nagai K. 1998. Crystal structure of the spliceosomal U2B″U2A′ protein complex bound to a fragment of U2 small nuclear RNA. *Nature* 394(6694):645–50
48. Parker R, Siliciano PG, Guthrie C. 1987. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* 49(2):229–39
49. Zhuang YA, Goldstein AM, Weiner AM. 1989. UACUAAC is the preferred branch site for mammalian mRNA splicing. *PNAS* 86(8):2752–56

50. Wu J, Manley JL. 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev*. 3(10):1553–61
51. Plaschka C, Lin P-C, Nagai K. 2017. Structure of a pre-catalytic spliceosome. *Nature* 546(7660):617–21
52. Nguyen THD, Galej WP, Bai X-C, Savva CG, Newman AJ, et al. 2015. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* 523(7558):47–52
53. Nguyen THD, Galej WP, Bai X-C, Oubridge C, Newman AJ, et al. 2016. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* 530(7590):298–302
54. Wan R, Yan C, Bai R, Wang L, Huang M. 2016. The 3.8 Å structure of the U4/U6.U5 tri-snRNP: insights into spliceosome assembly and catalysis. *Science* 351(6272):466–75
55. Agafonov DE, Kastner B, Dybkov O, Hofele RV, Liu W-T, et al. 2016. Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* 351(6280):1416–20
56. Charenton C, Wilkinson ME, Nagai K. 2019. Mechanism of 5′ splice site transfer for human spliceosome activation. *Science* 364(6438):362–67
57. Bringmann P, Appel B, Rinke J, Reuter R, Theissen H, Lührmann R. 1984. Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing. *EMBO J.* 3(6):1357–63
58. Hashimoto C, Steitz JA. 1984. U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. *Nucleic Acids Res*. 12(7):3283–93
59. Brow DA, Guthrie C. 1988. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature* 334(6179):213–18
60. Newman AJ, Norman CM. 1992. U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. *Cell* 68(4):743–54
61. Sontheimer EJ, Steitz JA. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* 262(5142):1989–96
62. Galej WP, Oubridge C, Newman AJ, Nagai K. 2013. Crystal structure of Prp8 reveals active site cavity of the spliceosome. *Nature* 493(7434):638–43
63. Turner IA, Norman CM, Churcher MJ, Newman AJ. 2006. Dissection of Prp8 protein defines multiple interactions with crucial RNA sequences in the catalytic core of the spliceosome. *RNA* 12(3):375–86
64. Galej WP, Toor N, Newman AJ, Nagai K. 2018. Molecular mechanism and evolution of nuclear pre-mRNA and group II intron splicing: insights from cryo-electron microscopy structures. *Chem. Rev.* 118(8):4156–76
65. Fabrizio P, Laggerbauer B, Lauber J, Lane WS, Lührmann R. 1997. An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. *EMBO J.* 16(13):4092–106
66. Fica SM, Nagai K. 2017. Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. *Nat. Struct. Mol. Biol.* 24(10):791–99
67. Nguyen THD, Li J, Galej WP, Oshikane H, Newman AJ, Nagai K. 2013. Structural basis of Brr2-Prp8 interactions and implications for U5 snRNP biogenesis and the spliceosome active site. *Struct. Fold. Des.* 21(6):910–19
68. Mozaffari-Jovin S, Wandersleben T, Santos KF, Will CL, Lührmann R, Wahl MC. 2013. Inhibition of RNA helicase Brr2 by the C-terminal tail of the spliceosomal protein Prp8. *Science* 341(6141):80–84
69. Huang Y-H, Chung C-S, Kao D-I, Kao T-C, Cheng S-C. 2014. Sad1 counteracts Brr2-mediated dissociation of U4/U6.U5 in tri-snRNP homeostasis. *Mol. Cell. Biol.* 34(2):210–20
70. Absmeier E, Wollenhaupt J, Mozaffari-Jovin S, Becke C, Lee C-T, et al. 2015. The large N-terminal region of the Brr2 RNA helicase guides productive spliceosome activation. *Genes Dev*. 29(24):2576–87
71. Gottschalk A, Neubauer G, Banroques J, Mann M, Lührmann R, Fabrizio P. 1999. Identification by mass spectrometry and functional analysis of novel proteins of the yeast [U4/U6.U5] tri-snRNP. *EMBO J.* 18(16):4535–48
72. Stevens SW, Abelson J. 1999. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. *PNAS* 96(13):7226–31
73. Jacewicz A, Schwer B, Smith P, Shuman S. 2014. Crystal structure, mutational analysis and RNA-dependent ATPase activity of the yeast DEAD-box pre-mRNA splicing factor Prp28. *Nucleic Acids Res*. 42(20):12885–98

74. Hausner TP, Giglio LM, Weiner AM. 1990. Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes Dev.* 4(12A):2146–56
75. Boesler C, Rigo N, Anokhina MM, Tauchert MJ, Agafonov DE, et al. 2016. A spliceosome intermediate with loosely associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. *Nat. Commun.* 7(1):11997
76. Wassarman DA, Steitz JA. 1992. Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. *Science* 257(5078):1918–25
77. Lesser CF, Guthrie C. 1993. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science* 262(5142):1982–88
78. Kandels-Lewis S, Séraphin B. 1993. Involvement of U6 snRNA in 5′ splice site selection. *Science* 262(5142):2035–39
79. Zhan X, Yan C, Zhang X, Lei J, Shi Y. 2018. Structures of the human pre-catalytic spliceosome and its precursor spliceosome. *Cell Res.* 28(12):1129–40
80. Fica SM, Mefford MA, Piccirilli JA, Staley JP. 2014. Evidence for a group II intron-like catalytic triplex in the spliceosome. *Nat. Struct. Mol. Biol.* 21(5):464–71
81. Rauhut R, Fabrizio P, Dybkov O, Hartmuth K, Pena V, et al. 2016. Molecular architecture of the *Saccharomyces cerevisiae* activated spliceosome. *Science* 353(6306):1399–405
82. Yan C, Wan R, Bai R, Huang G, Shi Y. 2016. Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* 353(6302):904–11
83. Madhani HD, 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(5):803–17
84. Toor N, Keating KS, Taylor SD, Pyle AM. 2008. Crystal structure of a self-spliced group II intron. *Science* 320(5872):77–82
85. Fica SM, Oubridge C, Galej WP, Wilkinson ME, Bai X-C, et al. 2017. Structure of a spliceosome remodelled for exon ligation. *Nature* 542(7641):377–80
86. Yan C, Wan R, Bai R, Huang G, Shi Y. 2016. Structure of a yeast step II catalytically activated spliceosome. *Science* 355(6321):149–55
87. Wilkinson ME, Fica SM, Galej WP, Norman CM, Newman AJ, Nagai K. 2017. Postcatalytic spliceosome structure reveals mechanism of 3′-splice site selection. *Science* 358(6368):1283–88
88. Liu S, Li X, Zhang L, Jiang J, Hill RC, et al. 2017. Structure of the yeast spliceosomal postcatalytic P complex. *Science* 358(6368):1278–83
89. Bai R, Yan C, Wan R, Lei J, Shi Y. 2017. Structure of the post-catalytic spliceosome from *Saccharomyces cerevisiae*. *Cell* 171(7):1589–98.e8
90. Wan R, Yan C, Bai R, Lei J, Shi Y. 2017. Structure of an intron lariat spliceosome from *Saccharomyces cerevisiae*. *Cell* 171(1):120–32.e12
91. Wan R, Bai R, Yan C, Lei J, Shi Y. 2019. Structures of the catalytically activated yeast spliceosome reveal the mechanism of branching. *Cell* 177(2):339–51.e13
92. Chan S-P, Kao D-I, Tsai W-Y, Cheng S-C. 2003. The Prp19p-associated complex in spliceosome activation. *Science* 302(5643):279–82
93. Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, et al. 2009. The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol. Cell* 36(4):593–608
94. Hoskins AA, Rodgers ML, Friedman LJ, Gelles J, Moore MJ. 2016. Single molecule analysis reveals reversible and irreversible steps during spliceosome activation. *eLife* 5:504
95. Zhang X, Yan C, Zhan X, Li L, Lei J, Shi Y. 2018. Structure of the human activated spliceosome in three conformational states. *Cell Res.* 28(3):307–22
96. Haselbach D, Komarov I, Agafonov DE, Hartmuth K, Graf B, et al. 2018. Structure and conformational dynamics of the human spliceosomal B^{act} complex. *Cell* 172(3):454–64.e11
97. Wu N-Y, Chung C-S, Cheng S-C. 2016. The role of Cwc24 in the first catalytic step of splicing and fidelity of 5′ splice site selection. *Mol. Cell. Biol.* 37:e00580-16
98. Bao P, Will CL, Urlaub H, Boon K-L, Lührmann R. 2017. The RES complex is required for efficient transformation of the precatalytic B spliceosome into an activated B^{act} complex. *Genes Dev.* 31(23–24):2416–29

99. Cordin O, Hahn D, Beggs JD. 2012. Structure, function and regulation of spliceosomal RNA helicases. *Curr. Opin. Cell Biol.* 24(3):431–38
100. Pyle AM. 2008. Translocation and unwinding mechanisms of RNA and DNA helicases. *Annu. Rev. Biophys.* 37:317–36
101. Semlow DR, Blanco MR, Walter NG, Staley JP. 2016. Spliceosomal DEAH-Box ATPases remodel pre-mRNA to activate alternative splice sites. *Cell* 164(5):985–98
102. Warkocki Z, Odenwälder P, Schmitzová J, Platzmann F, Stark H, et al. 2009. Reconstitution of both steps of *Saccharomyces cerevisiae* splicing with purified spliceosomal components. *Nat. Struct. Mol. Biol.* 16(12):1237–43
103. Liu Y-C, Chen H-C, Wu N-Y, Cheng S-C. 2007. A novel splicing factor, Yju2, is associated with NTC and acts after Prp2 in promoting the first catalytic reaction of pre-mRNA splicing. *Mol. Cell. Biol.* 27(15):5403–13
104. Villa T, Guthrie C. 2005. The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. *Genes Dev.* 19(16):1894–904
105. Chiu Y-F, Liu Y-C, Chiang T-W, Yeh T-C, Tseng C-K, et al. 2009. Cwc25 is a novel splicing factor required after Prp2 and Yju2 to facilitate the first catalytic reaction. *Mol. Cell. Biol.* 29(21):5671–78
106. Tseng C-K, Liu H-L, Cheng S-C. 2011. DEAH-box ATPase Prp16 has dual roles in remodeling of the spliceosome in catalytic steps. *RNA* 17(1):145–54
107. Query CC, Strobel SA, Sharp PA. 1996. Three recognition events at the branch-site adenine. *EMBO J.* 15(6):1392–402
108. Schwer B, Guthrie C. 1992. A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *EMBO J.* 11(13):5033–39
109. McPheeters DS, Muhlenkamp P. 2003. Spatial organization of protein-RNA interactions in the branch site-3′ splice site region during pre-mRNA splicing in yeast. *Mol. Cell. Biol.* 23(12):4174–86
110. Bertram K, Agafonov DE, Liu W-T, Dybkov O, Will CL, et al. 2017. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* 542(7641):318–23
111. Ohrt T, Odenwälder P, Dannenberg J, Prior M, Warkocki Z, et al. 2013. Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. *RNA* 19(7):902–15
112. James S-A, Turner W, Schwer B. 2002. How Slu7 and Prp18 cooperate in the second step of yeast pre-mRNA splicing. *RNA* 8(8):1068–77
113. Horowitz DS. 2012. The mechanism of the second step of pre-mRNA splicing. *Wiley Interdiscip. Rev. RNA* 3(3):331–50
114. Parker R, Siliciano PG. 1993. Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. *Nature* 361(6413):660–62
115. Patel AA, Steitz JA. 2003. Splicing double: insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* 4(12):960–70
116. Jones MH, Frank DN, Guthrie C. 1995. Characterization and functional ordering of Slu7p and Prp17p during the second step of pre-mRNA splicing in yeast. *PNAS* 92(21):9687–91
117. Brys A, Schwer B. 1996. Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3′ splice site. *RNA* 2(7):707–17
118. Chua K, Reed R. 1999. The RNA splicing factor hSlu7 is required for correct 3′ splice-site choice. *Nature* 402(6758):207–10
119. Company M, Arenas J, Abelson J. 1991. Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* 349(6309):487–93
120. Schwer B. 2008. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol. Cell* 30(6):743–54
121. Alexandrov A, Colognori D, Shu M-D, Steitz JA. 2012. Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay. *PNAS* 109(52):21313–18
122. Barbosa I, Haque N, Fiorini F, Barrandon C, Tomasetto C, et al. 2012. Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly. *Nat. Struct. Mol. Biol.* 19(10):983–90

123. Steckelberg A-L, Boehm V, Gromadzka AM, Gehring NH. 2012. CWC22 connects pre-mRNA splicing and exon junction complex assembly. *Cell Rep*. 2(3):454–61
124. Zhang X, Yan C, Hang J, Finci LI, Lei J, Shi Y. 2017. An atomic structure of the human spliceosome. *Cell* 169(5):918–29.e14
125. Le Hir H, Izaurralde E, Maquat LE, Moore MJ. 2000. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon–exon junctions. *EMBO J.* 19(24):6860–69
126. Zhan X, Yan C, Zhang X, Lei J, Shi Y. 2018. Structure of a human catalytic step I spliceosome. *Science* 359(6375):537–45
127. Fica SM, Oubridge C, Wilkinson ME, Newman AJ, Nagai K. 2019. A human postcatalytic spliceosome structure reveals essential roles of metazoan factors for exon ligation. *Science* 363(6428):710–14
128. Zhang X, Zhan X, Yan C, Zhang W, Liu D, et al. 2019. Structures of the human spliceosomes before and after release of the ligated exon. *Cell Res*. 29:274–85
129. Arenas JE, Abelson J. 1997. Prp43: an RNA helicase-like factor involved in spliceosome disassembly. *PNAS* 94(22):11798–802
130. Leeds NB, Small EC, Hiley SL, Hughes TR, Staley JP. 2006. The splicing factor Prp43p, a DEAH box ATPase, functions in ribosome biogenesis. *Mol. Cell. Biol.* 26(2):513–22
131. Combs DJ, Nagel RJ, Ares M, Stevens SW. 2006. Prp43p is a DEAH-box spliceosome disassembly factor essential for ribosome biogenesis. *Mol. Cell. Biol.* 26(2):523–34
132. Tsai R-T, Fu R-H, Yeh F-L, Tseng C-K, Lin Y-C, et al. 2005. Spliceosome disassembly catalyzed by Prp43 and its associated components Ntr1 and Ntr2. *Genes Dev*. 19(24):2991–3003
133. Tanaka N, Aronova A, Schwer B. 2007. Ntr1 activates the Prp43 helicase to trigger release of lariat-intron from the spliceosome. *Genes Dev*. 21(18):2312–25
134. Fourmann JB, Schmitzová J, Christian H, Urlaub H, Ficner R, et al. 2013. Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev*. 27(4):413–28
135. Koodathingal P, Novak T, Piccirilli JA, Staley JP. 2010. The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5′ splice site cleavage during pre-mRNA splicing. *Mol. Cell* 39(3):385–95
136. Fourmann JB, Dybkov O, Agafonov DE, Tauchert MJ, Urlaub H, et al. 2016. The target of the DEAH-box NTP triphosphatase Prp43 in *Saccharomyces cerevisiae* spliceosomes is the U2 snRNP-intron interaction. *eLife* 5:11798
137. Bohnsack MT, Martin R, Granneman S, Ruprecht M, Schleiff E, Tollervey D. 2009. Prp43 bound at different sites on the pre-rRNA performs distinct functions in ribosome synthesis. *Mol. Cell* 36(4):583–92
138. Toroney R, Nielsen KH, Staley JP. 2019. Termination of pre-mRNA splicing requires that the ATPase and RNA unwindase Prp43 acts on the catalytic snRNA U6. *Genes Dev*. 33:1555–74
139. Burgess SM, Guthrie C. 1993. A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. *Cell* 73(7):1377–91
140. Mayas RM, Maita H, Semlow DR, Staley JP. 2010. Spliceosome discards intermediates via the DEAH box ATPase Prp43p. *PNAS* 107(22):10020–25
141. Chen H-C, Tseng C-K, Tsai R-T, Chung C-S, Cheng S-C. 2013. Link of NTR-mediated spliceosome disassembly with DEAH-box ATPases Prp2, Prp16, and Prp22. *Mol. Cell. Biol.* 33(3):514–25
142. Herzel L, Ottoz DSM, Alpert T, Neugebauer KM. 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat. Rev. Mol. Cell Biol.* 18(10):637–50
143. Chan S-P, Cheng S-C. 2005. The Prp19-associated complex is required for specifying interactions of U5 and U6 with pre-mRNA during spliceosome activation. *J. Biol. Chem.* 280(35):31190–99
144. Agafonov DE, Deckert J, Wolf E, Odenwälder P, Bessonov S, et al. 2011. Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol. Cell. Biol.* 31(13):2667–82
145. Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ. 2002. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* 8(4):426–39
146. Horowitz DS, Krainer AR. 1997. A human protein required for the second step of pre-mRNA splicing is functionally related to a yeast splicing factor. *Genes Dev*. 11(1):139–51

147. Chen X, Zhang H, Aravindakshan JP, Gotlieb WH, Sairam MR. 2011. Anti-proliferative and pro-apoptotic actions of a novel human and mouse ovarian tumor-associated gene OTAG-12: downregulation, alternative splicing and drug sensitization. *Oncogene* 30(25):2874–87
148. Lorenzi LE, Bah A, Wischnewski H, Shchepachev V, Soneson C, et al. 2015. Fission yeast Cactin restricts telomere transcription and elongation by controlling Rap1 levels. *EMBO J.* 34(1):115–29
149. Thakran P, Pandit PA, Datta S, Kolathur KK, Pleiss JA, Mishra SK. 2018. Sde2 is an intron-specific pre-mRNA splicing regulator activated by ubiquitin-like processing. *EMBO J.* 37(1):89–101
150. Bertram K, Agafonov DE, Dybkov O, Haselbach D, Leelaram MN, et al. 2017. Cryo-EM structure of a pre-catalytic human spliceosome primed for activation. *Cell* 170(4):701–11

RELATED RESOURCES LIST

The Nagai laboratory website hosts a number of PyMOL sessions set up for various spliceosome structures, which make exploring these structures easier. The site also includes a guide for using PyMOL to explore the spliceosome. **https://www2.mrc-lmb.cam.ac.uk/groups/nagai/resources/**

The Spliceosome Database hosted by the Jurica laboratory collates information about the different spliceosome intermediates, including composition and the details of solved structures. **http://spliceosomedb.ucsc.edu/**

Annual Review of Biochemistry

Volume 89, 2020

Contents

Errata

An online log of corrections to *Annual Review of Biochemistry* articles may be found at http://www.annualreviews.org/errata/biochem