Structures of the Catalytically Activated Yeast Spliceosome Reveal the Mechanism of Branching

Ruixue Wan[1,4], Rui Bai[1,4], Chuangye Yan[1], Jianlin Lei[1,2], and Yigong Shi[1,3,5]

[1]Beijing Advanced Innovation Center for Structural Biology, Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences and School of Medicine, Tsinghua University, Beijing 100084, China

[2]Technology Center for Protein Sciences, Ministry of Education Key Laboratory of Protein Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China

[3]Institute of Biology, Westlake Institute for Advanced Study; School of Life Sciences, Westlake University, 18 Shilongshan Road, Xihu District, Hangzhou 310024, Zhejiang Province, China

[4]These authors contributed equally to this work.

[5]Lead contact

[*]To whom correspondence should be addressed. E-mail: shi-lab@tsinghua.edu.cn

1

Summary

Pre-mRNA splicing is executed by the spliceosome. Structural characterization of the catalytically activated complex (B*) is pivotal for mechanistic understanding of catalysis of the branching reaction by the spliceosome. In this study, we assembled the B* complex on two different pre-mRNAs from *Saccharomyces cerevisiae* and determined the cryo-EM structures of four distinct B* complexes at overall resolutions of 2.9-3.8 Å. The duplex between U2 snRNA and the branch point sequence (BPS) is located 13-20 Å away from the 5'-splice site (5'SS) in the B* complexes that are devoid of the step I splicing factors Yju2 and Cwc25. Recruitment of Yju2 into the active site brings the U2/BPS duplex into the vicinity of 5'SS, ready for branching. In the absence of Cwc25, the nucleophile from BPS is positioned about 4 Å away from, and remains to be activated by, the catalytic metal M2. This analysis reveals the functional mechanism of Yju2 and Cwc25 in branching. These four structures constitute compelling evidence for substrate-specific conformations of the spliceosome in a major functional state.

Introduction

Pre-mRNA splicing is executed by the spliceosome (Brody and Abelson, 1985; Frendewey and Keller, 1985; Grabowski et al., 1985), a supramolecular complex with exceptional dynamics in its composition and conformation (Shi, 2017a; Will and Luhrmann, 2011). The fully assembled spliceosome exists in at least eight major functional states: precursor of the pre-catalytic spliceosome (pre-B), pre-catalytic spliceosome (B), activated complex (B^{act}), catalytically activated complex (B^*), catalytic step I complex (C), step II activated complex (C^*), post-catalytic complex (P), and intron lariat spliceosome (ILS). Each splicing cycle, involving branching and exon ligation, results in the removal of the intervening RNA sequences between two target exons (Grabowski et al., 1984; Padgett et al., 1984; Ruskin et al., 1984). Branching and exon ligation proceed spontaneously in the B^* and C^* complexes, respectively (Shi, 2017b). The B^{act}-to-B^* and C-to-C^* transitions, however, are driven by the ATPase/helicases Prp2 and Prp16, respectively (Cordin et al., 2012; Jankowsky, 2011; Liu and Cheng, 2015).

The first near-atomic resolution structure of an assembled spliceosome, determined at 3.6 Å by single-particle cryo-EM analysis, was reported for the ILS complex from *S. pombe* in 2015 (Hang et al., 2015; Yan et al., 2015). Since then, 13 cryo-EM structures, mostly at resolutions between 3.3 and 5.8 Å, have been elucidated for seven distinct states of the assembled spliceosome from *S. cerevisiae* (Bai et al., 2018; Bai et al., 2017; Fica et al., 2017; Galej et al., 2016; Liu et al., 2017;

3

Plaschka et al., 2017; Rauhut et al., 2016a; Wan et al., 2016b; Wan et al., 2017; Wilkinson et al., 2017; Yan et al., 2016, 2017). Seven such structures have been reported for five distinct states of the human spliceosome (Bertram et al., 2017a; Bertram et al., 2017b; Haselbach et al., 2018; Zhan et al., 2018a, b; Zhang et al., 2017; Zhang et al., 2018). Among the eight known functional states of the spliceosome, only the B* complex remains structurally uncharacterized.

In this study, we report the cryo-EM structures of the B* complex from *S. cerevisiae*. The B* complex was prepared using two different pre-mRNAs, which give rise to an average resolution of 2.9 Å for the B* complex assembled on the *ACT1* pre-mRNA and 3.2 Å for that assembled on the *UBC4* pre-mRNA. For each of the two distinct B* complexes, two different conformational states were observed. Comparison of these conformational states reveals mechanistic insights into the branching reaction.

Results

Preparation and EM analysis of the B* complex

It is difficult to isolate the highly transient B* complex. Relying on direct purification of the spliceosomes from cells, we have been unable to purify the endogenous B* complex. We sought to reconstitute the B* complex using the *in vitro* assembly approach, in which the assembled B^{act} complex was remodeled into the B* complex upon incubation with recombinant Prp2 and Spp2 in the presence of ATP (Bao et al.,

4

2017; Ohrt et al., 2013; Roy et al., 1995; Warkocki et al., 2009; Warkocki et al., 2015) (Figure S1A-C). To maximize our chance of success, we used two different pre-mRNA: *ACT1* and *UBC4* (Figure S1C). The 3'-end nucleotides of the 5'-exon in *ACT1* are UCUG; in contrast, the corresponding nucleotides in *UBC4* are AAAG, which may make improved base-pairing interactions with the poly-U sequences of U5 loop I. In both cases, we were able to isolate the spliceosomes that appeared to be intact by negative staining EM analysis (Figure S1D,E).

For both samples, we prepared the cryo-EM grids and performed EM analysis (Figure S1F-H). 1.8 and 1.7 million particles were auto-picked for the spliceosomes assembled on the *ACT1* and *UBC4* pre-mRNA, respectively (Figures S2 & S3). Following two rounds of three-dimensional (3D) classification, 555,036 particles yielded a final reconstruction of the B* complex at an average resolution of 2.9 Å (Figures S2 & S4; Tables S1 & S2). Despite excellent EM density for much of the *ACT1* B* complex (Figures S5), the density in the 5'SS region is weak. Next, we processed the data on the spliceosomes assembled on *UBC4*. After two rounds of 3D classification, 132,125 particles yielded a final structure of the *UBC4* B* complex at an average resolution of 3.2 Å (Figures S3, S4, S6; Tables S1 & S3). Use of *UBC4* indeed led to improved EM density in the 5'SS region.

For each of the two B* complexes, the local resolution in the U2/BPS duplex region is relatively low. This is likely due to absence of the step I splicing factors

Cwc25 and Yju2, which may stabilize the conformation of the nucleophile as well as the acceptor (Galej et al., 2016; Wan et al., 2016b). To improve the local density, we applied a soft mask on the U2/BPS region and performed additional 3D classification. This effort led to two distinct conformational states of the *ACT1* B[*] complex at average resolutions of 3.6 and 3.2 Å, named B[*]-a1 and B[*]-a2, respectively (Figure 1A; Figures S2 & S7A,B). A similar analysis on the *UBC4* B[*] complex also resulted in two distinct conformations at 3.9 and 3.7 Å resolution, designated as B[*]-b1 and B[*]-b2, respectively (Figure 1B; Figures S3 & S7C,D).

Structure of the B[*] complex

Except for Yju2 in B[*]-b2, all four B[*] complexes share the same set of RNA and protein components (Figure 1). The overall structures of these components are very similar among the four B[*] complexes except for the step II splicing factor Prp17, the NTC component Isy1, and the local conformation of two RNA elements: the U2/BPS duplex and the 5'-exon-5'SS sequence. The final atomic model for each of the four B[*] complexes from *S. cerevisiae* contains a pre-mRNA, U6 small nuclear RNA (snRNA), U5 small nuclear ribonucleoprotein (snRNP), NineTeen complex (NTC), NTC-related (NTR), U2 snRNP core (U2 snRNA, Leal, Msl1, and the U2 Sm ring), and the splicing factors Cwc21, Cwc22 and Prp17 (Figure 1A). The shared components of these four B[*] complexes include 34 proteins and four RNAs (Tables S1-S3). The WD40 domain of Prp17 is invisible only in the B[*]-a1 complex but is clearly identified

6

in the other three B* complexes; it is positioned in slightly different locations in these complexes, presumably due to the relative movements of the U2/BPS duplex.

Replacement of *ACT1* by *UBC4* in the B* complex led to improvement of the EM density for the RNA elements in the splicing active site center (Figure S6B,C). Consequently, base-pairing interactions among the four pieces of RNA are unambiguously assigned in the *UBC4* B* complex (Figure S7E). The step I splicing factor Yju2, which is thought to be recruited into the B^{act} complex along with the NTC components (Chang et al., 2009; Liu et al., 2007; Warkocki et al., 2009), was not structurally identified in the B^{act} complex (Rauhut et al., 2016b; Yan et al., 2016). Notably, however, Yju2 is present only in the B*-b2 complex, but not in the other three B* complexes (Figure 1B). The N-terminus of Isy1 is positioned only in the active site of the B*-b2 complex but disordered in the other three complexes, likely due to the presence of Yju2 in B*-b2. A total of 548 nucleotides have been identified in the *UBC4* B* complex (Figure 2A). Among these, 491 are assigned to the three snRNAs and 57 to the *UBC4* pre-mRNA.

The RNA elements in the B* complex

A major difference among the four B* complexes is the conformation and the location of the key RNA elements in the catalytic center (Figures 1 & 2). These RNA elements include the U2/BPS, 5'SS, and U5/5'-exon duplexes. The conformation and precise positioning of the U2/BPS duplex in the active site require stabilization by the step I

factors Cwc25 and Yju2 (Galej et al., 2016; Schneider et al., 2015; Tseng et al., 2017; Wan et al., 2016b). In the absence of Cwc25 and Yju2, the U2/BPS duplex is not positioned in the close vicinity of the 5'SS and its exact conformation and location differ among the three B* complexes: a1, a2, and b1 (Figure 2B,C). The nucleophile from the BPS is located about 20 Å away from the phosphate of G_1 in the 5'SS in the B*-a1 complex (Figure 2B, left panel); this distance is shortened to approximately 13 Å in B*-a2 (Figure 2B, right panel) and 15 Å in B*-b1 (Figure 2C, left panel). In sharp contrast, the U2/BPS duplex is translocated to the vicinity of the 5'SS in the Yju2-containing B*-b2 complex, with the 2'-OH of BPS A_{70} poised for nucleophilic attack on the phosphate of G_1 in the 5'SS that is only 4.3 Å away (Figure 2C, right panel).

Recognition of the highly conserved 5'SS by U6 snRNA is nearly identical among the four B* complexes. Specifically, three contiguous nucleotides $U_4G_5U_6$ of the 5'SS form a duplex with $A_{47}C_{48}A_{49}$ of the ACAGA box of U6 snRNA through canonical base-pairing interactions (Figure 2D,E, upper panels). On the other hand, different sequences of *ACT1* and *UBC4* at the 3'-end of their 5'-exons lead to recognition of different strengths by U5 snRNA. The nucleotides $G_{-1}U_{-2}C_{-3}U_{-4}$ of the *ACT1* 5'-exon are bound loosely to the loop I of U5 snRNA, with U_{-2} of *ACT1* base-pairing to U_{96} of U5 snRNA (Figure 2D, lower panel). In contrast, the nucleotides $G_{-1}A_{-2}A_{-3}A_{-4}$ of the *UBC4* 5'-exon form a duplex with $U_{96}U_{97}U_{98}U_{99}$ of U5 snRNA through canonical base-pairing H-bonds (Figure 2E, lower panel).

The splicing active site and metal ions

In the B^*-b2 complex, the splicing active site comprises loop I of U5 snRNA, helix I of the U2/U6 duplex, intramolecular stem loop (ISL) of U6 snRNA and five associated metal ions (Figure 3A). The nucleophile and the acceptor of the branching reaction are already delivered into the active site. Four consecutive nucleotides at the 3'-end of the 5'-exon are anchored to U5 loop I. Helix Ib and three nucleotides of U6 snRNA form a characteristic catalytic triplex (Figures S6B & S7E). Remarkably, the nucleophile-containing A_{70} of the BPS interacts with G_1U_2 of the 5'SS, with the adenine base making canonical H-bonds to the uracil while stacking against the guanine (Figure 3B). The nucleophile 2'-OH of A_{70} is located about 4.3 Å away from the phosphorous atom of G_1, ready to be activated by a catalytic magnesium ion (M2) for the branching reaction.

Among the five metal ions, three are coordinated exclusively by the ISL and may play a structural role by neutralizing the negative charges of the RNA (Figure 3A). These three metals are identically positioned in all four B^* complexes. The other two metals, presumably Mg^{2+}, have been identified as M1, which stabilizes the leaving group during the branching reaction, and M2, which activates the nucleophile (Fica et al., 2014; Fica et al., 2013; Keating et al., 2010; Steitz and Steitz, 1993; Yean et al., 2000). In the B^*-a1 and -a2 complexes, M1 is coordinated by the phosphate oxygens of G_{78} and U_{80} from U6 snRNA and M2 is bound by a phosphate oxygen from U_{80} of U6 snRNA (Figure 3C,D). M1 is identically coordinated in the B^*-b1

complex (Figure 3E) but is already bound to the phosphate oxygen of G_1 of the 5'SS in the B^*-b2 complex (Figure 3F). Compared to B^*-a1/a2, M2 gains an additional ligand (a phosphate oxygen from G_{60} of U6 snRNA) in the B^*-b1 complex (Figure 3E) and two additional ligands (two phosphate oxygens from A_{59} and G_{60} of U6 snRNA) in the B^*-b2 complex (Figure 3F). But the 2'-OH of A_{70} is yet to be activated by M2, with a distance of approximately 4.3 Å. The observation that the phosphate of U_{80} of U6 snRNA coordinates both M1 and M2 in all four B^* complexes is consistent with the finding that sulfur substitution of one of the non-bridging phosphoryl oxygens in U_{80} abolishes the branching reaction (Fica et al., 2014; Yean et al., 2000).

Among the four B^* complexes, B^*-b2 appears to be nearly ready for the branching reaction. In response to Yju2 binding, the 5'-exon-5'SS junction undergoes a major reorganization in the B^*-b2 complex compared to the other three (Figure 3C-F). In particular, the phosphate group of G_1 of the 5'SS is flipped inside out to coordinate the M1 metal. Consistent with published observations (Bao et al., 2017), such a dynamic change is afforded by the dissociation of Cwc24 and Prp11, which protect G_1 in the B^{act} complex (Rauhut et al., 2016b; Yan et al., 2016), and the recruitment of Yju2, which stabilizes the local conformation in the B^*-b2 complex. The recruitment of Cwc25 may allow slight rearrangement of the active site RNA elements and consequent activation of the nucleophile by M2 (Chiu et al., 2009; Galej et al., 2016; Wan et al., 2016b; Warkocki et al., 2009).

Structural changes of the Bact-to-B* transition

The Bact-to-B* transition requires the ATPase/helicase Prp2 and involves a major reorganization of the spliceosome (Bao et al., 2017; Kim and Lin, 1996; Lardelli et al., 2010; Ohrt et al., 2012; Warkocki et al., 2009; Warkocki et al., 2015) (Figure 4A). During this transition, at least 16 proteins are dissociated and several proteins undergo conformational changes. In the Bact complex, the SF3b complex and the RES complex recognize the U2/BPS duplex and the downstream RNA sequences of the intron, respectively; Prp11 from the SF3a complex and the splicing factor Cwc24 are placed in between the U2/BPS duplex and the 5'SS (Yan et al., 2016) (Figure 4A). Upon the action of Prp2, SF3a, SF3b, RES and the splicing factors Cwc24 and Cwc27 are dissociated to release the U2/BPS duplex and the 5'SS (Lardelli et al., 2010; Ohrt et al., 2012; Warkocki et al., 2009). U5 snRNP, NTR, and the splicing factors Cwc21 and Cwc22 remain static. In contrast, the U2 snRNP core and the superhelical proteins Syf1 and Clf1 of the NTC complex undergo marked translocation. These changes are similar to those described for the Bact-to-C transition (Wan et al., 2016b).

Concurrent with the flux of the protein components, some of the RNA elements also undergo drastic structural changes. Compared to the Bact complex, the U2 snRNA sequences downstream of nucleotide 30 and the pre-mRNA downstream of the 5'SS are drastically translocated in the B*-a1 complex (Figure 4B). The interaction between stem IIb of U2 snRNA and the C-terminal domain of Ecm2 is clearly present in the four B* complexes (Figure S5F). In addition, stem IIc of U2

11

snRNA, which is thought to promote the 5'SS cleavage (Hilliker et al., 2007; Perriman and Ares, 2007), is formed during the transition from the B^{act} to B^*-a1 complex (Figure S7E). The U2/BPS duplex, which was protected by the SF3b complex and located ~50 Å away from the active site in the B^{act} complex (Yan et al., 2016), undergoes a movement of 30-40 Å into the vicinity of the 5'SS in the B^*-a1 complex. In contrast, U5 and U6 snRNAs and the first 30 nucleotides of U2 snRNA remain unchanged during the transition (Figure 4B). Consequently, the RNA elements in the active site remain nearly identical between the B^{act} and B^* complexes (Figure 4C). The only notable difference is that, compared to the B^*-a1 complex, M1 is yet to be loaded in the B^{act} complex (Yan et al., 2016) (Figure 4D; Figure S6C, left panel).

In the B^{act} complex (Yan et al., 2016), the RNA elements at the active site (including pre-mRNA and snRNAs) are coated by protein components and stabilized by Prp11, Cwc24 and the 1585-loop (also known as α-finger) of Prp8, which allow little movement of the RNA elements. Consequently, the M1 metal is not loaded, because the phosphate oxygens from G78 and U80 of U6 snRNA are not in the right positions to coordinate M1. In the B^* complex, most of the coated proteins have been dissociated through the action of Prp2. The RNA elements at the active site are less constrained; the ISL of U6 snRNA undergoes slight conformational adjustment to allow coordination of M1 by the phosphate oxygens from G78 and U80.

Structural changes of the B^*-to-C transition

Branching occurs in the B^* complex, resulting in the C complex. Unlike the B^{act}-to-B^* transition, the B^*-to-C transition has been predicted to involve few structural changes (Shi, 2017b; Wan et al., 2016b). Indeed, the snRNA elements and much of the pre-mRNA are nearly identical between the B^*-b2 and C complexes (Wan et al., 2016b) (Figure 5A). The RNA elements in the active site of the B^*-b2 complex can be superimposed to those of the C complex with a root-mean-squared deviation of 0.46 Å for 80 nucleotides (Figure 5B). The only apparent difference is that the 3'-5' covalent linkage between the 5'-exon and 5'SS in the B^*-b2 complex is broken and replaced by the 2'-5' phosphodiester bond between A_{70} of the BPS and G_1 of the 5'SS in the C complex (Wan et al., 2016b) (Figure 5C). As a consequence of branching, coordination of the catalytic metals in the C complex is different from that in the B^*-b2 complex (Figure 5D; Figure S6C, middle and right panels). In the C complex, M1 is coordinated in a planar fashion by four ligands: 3'-OH of G_{-1} of the 5'-exon and three phosphate oxygens from G_{78} and U_{80} of U6 snRNA and G_1 of 5'SS (Wan et al., 2016b). In the B^*-b2 complex, M1 is only coordinated by three ligands but does not interact with 3'-OH of G_{-1} of the 5'-exon. Notably, both M2 and the phosphates of G_{-1} and G_1 undergo apparent positional shifts in the B^*-to-C transition (Figure 5D). In the C complex, M2 is bound by two phosphate oxygens from A_{59} and U_{80} of U6 snRNA (Wan et al., 2016b); in the B^*-b2 complex, M2 is coordinated by three phosphates from A_{59}, G_{60} and U_{80} of U6 snRNA.

In the B*-b2 complex, the step I splicing factor Yju2 is already loaded into the active site, with its N-terminal loop and β-sheet domain binding to 5'SS and U2/BPS duplex, respectively (Figure 5E). Consistent with previous biochemical analysis (Liu et al., 2007; Ohrt et al., 2012; Warkocki et al., 2009), the presence of Yju2 in the structure suggests its recruitment prior to or during the formation of the B^{act} complex because no recombinant Yju2 is required in our preparation of the B* complex. However, the other step I factor Cwc25, which is thought to be required after Yju2 to promote branching (Chiu et al., 2009; Warkocki et al., 2009), is still absent in the B*-b2 complex. A close-up comparison of the region occupied by the step I factors between the B*-b2 and C complexes reveals apparent positional shifts for the β-sheet domain of Yju2 and the U2/BPS duplex (Figure 5F). These differences are accompanied by conformational changes of the U2/BPS duplex.

Notably, Cwc25 is the last protein factor recruited to the spliceosome before branching and the recruitment of Cwc25 is thought to require Yju2 (Chiu et al., 2009). Consistent with this conclusion, the active site conformation in the B*-b2 complex, which is stabilized by Yju2, is conducive for Cwc25 recruitment (Figure 5E). The recruitment of Cwc25, in turn, likely leads to a positional shift of the Yju2 β-sheet domain and movement of U2/BPS towards the active site center (Figure 5F).

Discussion

Spliceosome remodeling from the B^{act} to the B^* complex is driven by Prp2 and its cofactor Spp2 (Kim and Lin, 1993; King and Beggs, 1990; Roy et al., 1995). SF3a, SF3b, RES and the splicing factors Cwc24 and Cwc27 are dissociated from the B^{act} complex (Lardelli et al., 2010; Ohrt et al., 2012), and the step I splicing factors Yju2 and Cwc25 are required for branching after the action of Prp2 (Chiu et al., 2009; Krishnan et al., 2013; Liu et al., 2007). Yju2 by itself is insufficient to yield the intron lariat-3'-exon intermediate (Liu et al., 2007; Warkocki et al., 2009). Branching is dramatically stimulated by the addition of Cwc25, suggesting its indispensable role in promoting efficient branching (Krishnan et al., 2013; Warkocki et al., 2009). Despite these biochemical studies, how Yju2 and Cwc25 differentially facilitate the branching reaction remains unclear. Our structural analysis provides a plausible answer to this key question.

Structure determination of the B^* complex allows proposition of the structural changes at the active site during the B^{act}-to-B^*-to-C transition in *S. cerevisiae* (Figure 6A). This transition is triggered by Prp2, which may pull the 3'-end of the intron (Liu and Cheng, 2012; Yan et al., 2016), dissociating proteins that are associated with the BPS and 5'SS and allowing translocation of the U2/BPS duplex into the active site (Lardelli et al., 2010; Ohrt et al., 2012; Warkocki et al., 2009; Warkocki et al., 2015). The nucleophile-containing A_{70} of the BPS is stabilized through base-pairing interaction with U_2 and base-stacking interaction with G_1 at the 5'-end of the 5'SS. This conformation is supported by Isy1 and Yju2. Upon

recruitment of Yju2 and Cwc25 into the active site, the nucleophile from U2/BPS duplex is moved into the close proximity of the acceptor in the 5'SS, forming the B^* complex. Branching occurs instantaneously, resulting in a covalent linkage between the 2'-OH of BPS A_{70} and the 5'-phosphate of 5'SS G_1 in the C complex (Galej et al., 2016; Wan et al., 2016b).

On the basis of previous biochemical evidence and our structural observations, we further propose that the B^{act}-to-B^* transition may occur in several distinct steps (Figure 6B). First, the ATPase/helicase Prp2 and its cofactor Spp2 mediate the dissociation of the SF3a and SF3b complexes and the splicing factors Cwc24 and Cwc27, freeing the 5'SS and the U2/BPS duplex (Lardelli et al., 2010; Ohrt et al., 2012; Roy et al., 1995; Warkocki et al., 2009). In the resulting intermediary spliceosome (Figure 6B, bottom left), a number of the components and subcomplexes, which include Brr2, U2 snRNP core and two NTC proteins Syf1 and Clf1 of NTC, may exhibit dynamic conformation. Next, recruitment of Isy1 and the WD40 domain of Prp17 in the active site may lead to reorganization of the surrounding components, allowing translocation of the U2/BPS duplex into the approximately 20-Å range of the 5'SS and forming a partially catalytically activated complex (Figure 6B, bottom middle). This spliceosome may exhibit a range of dynamic conformations, exemplified by the a1 and a2 conformations of the *ACT1* B^* complex and the b1 conformation of the *UBC4* B^* complex prior to Yju2 loading. Then, the recruitment of Yju2 further reorganizes the active site, pushing the U2/BPS

duplex into the close proximity of the 5'SS as observed in the B^*-b2 complex (Figure 6B, bottom right). Finally, Cwc25 is recruited, allowing fine adjustment of the active site elements and consequent execution of the branching reaction (Wan et al., 2016b). The resulting spliceosome is the C complex.

Structure of the Yju2-containing B^* complex reveals, for the first time, the interaction between the highly conserved dinucleotide G_1U_2 at the 5'-end of the 5'SS and the invariant A_{70} of the BPS just before branching (Figure 3B). In particular, the nucleophile-containing A_{70} is recognized and partially stabilized by U_2 through base-pairing interactions. Replacement of U_2 by any other nucleotide may weaken or disrupt the interaction thus reducing the splicing level (Aebi et al., 1986; Aebi et al., 1987; Lamond et al., 1987). This analysis suggests that formation of these interactions may represent a potential check-point for proofreading the first transesterification. A similar check-point may exist just prior to exon ligation, whereby the invariant G_1 of the 5'SS and A_{70} of the BPS recognize the dinucleotide AG of the 3'SS through base-paring and stacking interactions as observed in the P complex (Bai et al., 2017; Liu et al., 2017; Wilkinson et al., 2017) (Figure 7A). Utilization of such conserved nucleotides towards specific interactions may constitute an additional layer of regulation to safeguard the splicing specificity and efficiency.

Previous studies reported the proofreading function of Prp16 and Prp22 before branching and exon ligation, respectively (Burgess and Guthrie, 1993;

17

Koodathingal et al., 2010; Mayas et al., 2006; Query and Konarska, 2004; Semlow et al., 2016; Villa and Guthrie, 2005). In the structure of the four B* complexes, the binding site for Prp16 is exposed. But the stable association of Prp16 requires the presence of both Cwc25 and Yju2, as observed in the structure of the C complex (Galej et al., 2016; Wan et al., 2016b). Consistently, in our structures of the B* complex, the EM density for Prp16 is insufficient for its conclusive assignment.

The overall structures of the *ACT1* B* complex and the *UBC4* B* complex are unsurprisingly similar, because the two pre-mRNA molecules share exactly the same sequences in the three signature elements: 5'SS, BPS, and 3'SS. Consequently, the 5'SS is recognized identically by the ACAGA box of U6 snRNA in either *ACT1* or *UBC4* B* complex (Figure 2D,E, upper panels). The BPS also forms the same duplex with the conserved sequences of U2 snRNA in either B* complex (Figure 2B,C). Such structural similarities even extend to a small molecule, which has been identified as inositol hexaphosphate (IP6). This small molecule is identically bound to the N-domain of Prp8 not just in all four structures of the B* complex (Figure 7B) but perhaps more importantly in all yeast spliceosomes between the functional states of B^{act} and ILS. The function of this small molecule in splicing remains to be investigated.

The *ACT1* B* complex and the *UBC4* B* complex differ in the 3'-end nucleotides of their 5'-exons. These differences determine the strengths of interactions

with loop I of U5 snRNA. Only a single nucleotide from the 5'-exon of *ACT1* base-pairs with loop I, whereas four nucleotides from the 5'-exon of *UBC4* form a duplex with loop I (Figure 2D,E, lower panels). The sequence variation in *ACT1* versus *UBC4* also results in differences in overall structure of the spliceosome (Figure 1), conformation of key RNA elements (Figure 2), and coordination of catalytic metal ions (Figure 3). These structurally documented differences constitute compelling evidence for substrate-specific conformations of the spliceosome in a major functional state – in this case the B* complex. These structural observations are consistent with the biochemical finding that depletion or mutation of spliceosomal core components exhibits differential effects on splicing substrates (Campion et al., 2010; Clark et al., 2002; Kawashima et al., 2009; Park et al., 2004; Pleiss et al., 2007; Saltzman et al., 2011). We further speculate that substrate-specific conformation may be a general phenomenon for all major functional states of the spliceosome.

The B* complex, if simultaneously containing Cwc25 and Yju2, would instantaneously proceed to branching and is thus extremely transient. To obtain the B* complex, we took the *in vitro* assembly approach using recombinant Prp2 and Spp2 (Bao et al., 2017; Warkocki et al., 2009). This approach allowed us to observe four distinct conformations of the B* complex assembled on two different pre-mRNA. Yju2 is only present in one of the *UBC4* B* conformers. These four structures might represent the B* complex just preceding the spliceosomal state that catalyzes branching. Three structures show that, in the absence of Cwc25 and Yju2, the U2/BPS

duplex cannot be placed into the correct location for the branching reaction,

explaining the essential function of the step I splicing factors (Chiu et al., 2009;

Krishnan et al., 2013; Liu et al., 2007; Warkocki et al., 2009). The recruitment of Yju2

into the active site allows the U2/BPS duplex to move into the close proximity of the

5'SS; but in the absence of Cwc25, the nucleophile remains to be activated by the M2

metal and the overall active site conformation is subtly unsuitable for the branching

reaction.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the

following:

- KEY RESOURCES TABLE

- CONTACT FOR REAGENT AND RESOURCE SHARDING

- EXPERIMENTAL MODEL AND SUBJECT DETAILS

    o Cell lines

- METHOD DETAILS

    o CBP tagging of in *S. cerevisiae*

    o Preparation of the yeast whole cell extract

    o Expression and purification of Prp2 and Spp2

    o Preparation of the pre-mRNA

    o Assembly and purification of the yeast B^{act} ($\triangle$Prp2) complex

- - Remodeling of the B^{act} ($\triangle$Prp2) complex by Prp2 and Spp2

  - EM data acquisition

  - Data processing

  - Model building and refinement

- QUANTIFICATION AND STATISTICAL ANALYSIS

- DATA AND SOFTWARE AVAILABILITY

  - Data Resources

SUPPLEMENTAL INFORMATION

Supplemental Information includes three tables and seven figures and can be found with this article online at XXX.

Author Contributions

R.B. prepared the whole cell extract, pre-mRNA substrates, and recombinant proteins Prp2 and Spp2. R.W. and R.B. purified the yeast spliceosomes and prepared the cryo-EM samples. R.W., R.B., and J.L. collected and processed the EM data. R.W. calculated the EM map and built the atomic model. C.Y. gave advices on data processing and model building. All authors contributed to structure analysis. R.W. and R.B. contributed to manuscript preparation. Y.S. designed and guided the project and wrote the manuscript.
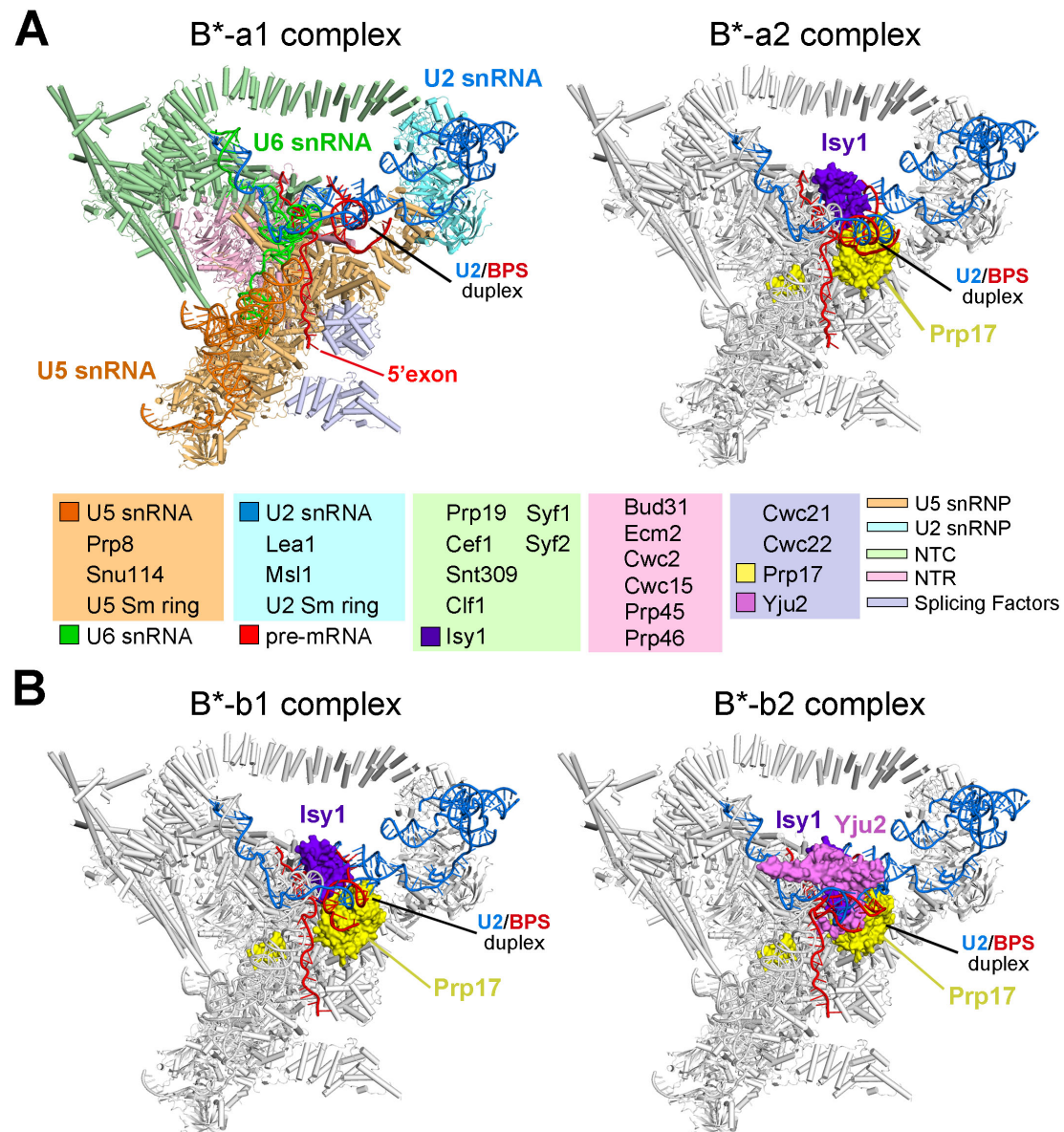
ACKNOWLEDGMENTS

Figures

Figure 1   Cryo-EM structures of the catalytically activated spliceosomes (B* **complexes) from *Saccharomyces cerevisiae* (*S. cerevisiae*).**   (A) Structures of the B* complex assembled on the *ACT1* pre-mRNA. Two distinct conformational states were captured and named B*-a1 (left panel) and B*-a2 (right panel). The final atomic model includes 34 proteins, three snRNAs, and 60 nucleotides of the *ACT1* pre-mRNA. The snRNPs and major components of the B*-a1 complex are color-coded.

The pre-mRNA, U2, U5 and U6 snRNAs are colored red, marine, orange and green, respectively. The splicing factors are colored light purple. Compared to B*-a1, only those components that undergo changes are colored in the B*-a2 complex, with the rest shown in grey. The protein and RNA components are tabulated below the structures.　 (B) Structures of the B* complex assembled on the *UBC4* pre-mRNA. Two distinct conformational states were captured and named B*-b1 (left panel) and B*-b2 (right panel). Compared to B*-a1, only those components that undergo changes are colored in these complexes. Compared to the other three B* complexes, B*-b2 contains an extra protein – the step I splicing factor Yju2 (violet). All structural images were created using PyMol (DeLano, 2002).
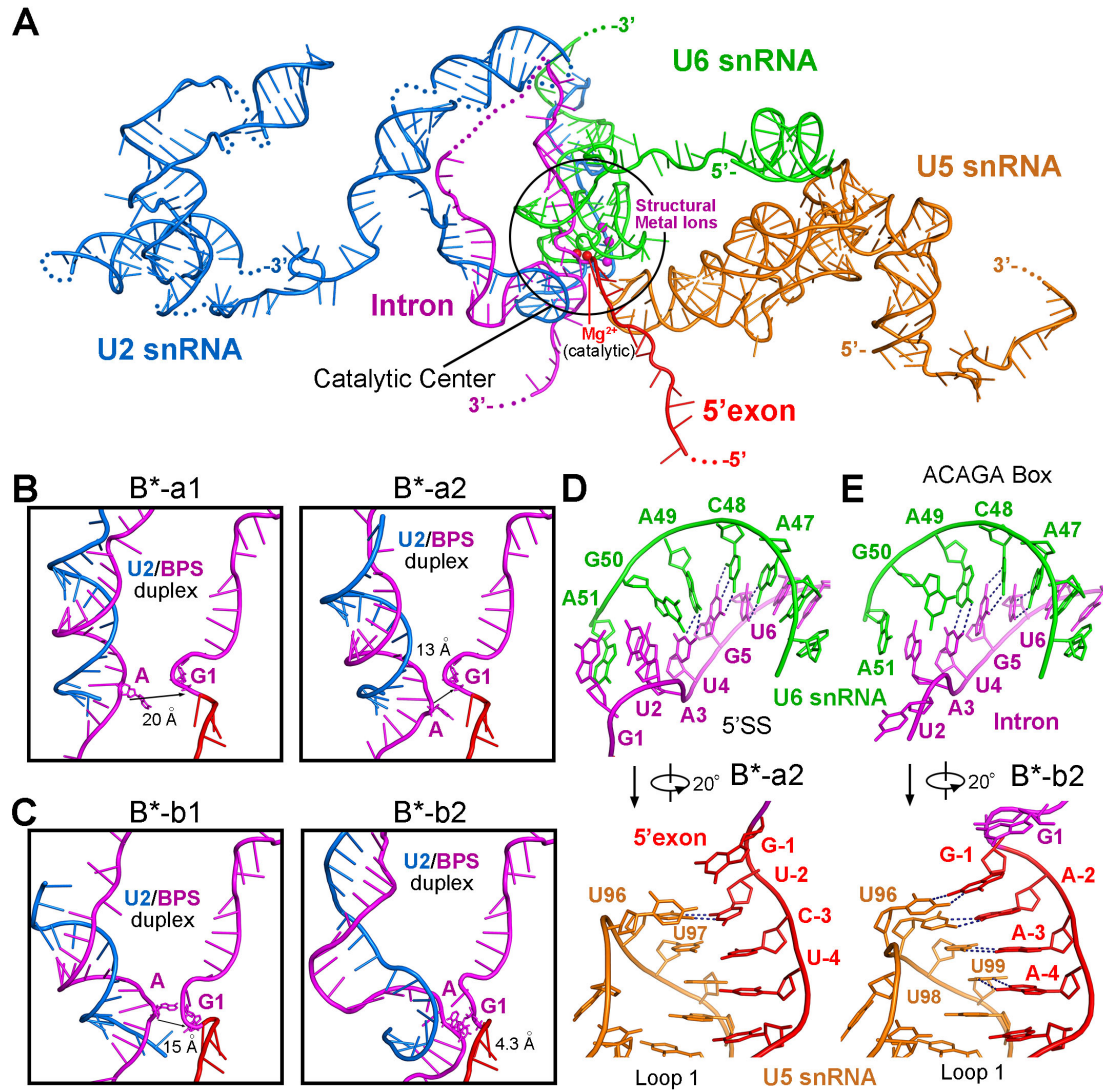
Figure 2   The RNA elements in the *S. cerevisiae* B* complexes.   (A) Structure of the RNA elements in the B*-b2 complex. The *UBC4* pre-mRNA and U2, U5, and U6 snRNAs are colored red, marine, orange, and green, respectively. The catalytic and structural metal ions are shown as red and magenta spheres, respectively.   (B) Comparison of the U2/BPS duplexes from the B*-a1 and B*-a2 complexes. The nucleophile is located about 20 and 13 Å away from the acceptor in the B*-a1 and B*-a2 complexes, respectively.   (C) Comparison of the U2/BPS duplexes from the B*-b1 and B*-b2 complexes. In the B*-b1 complex (left panel), the nucleophile is about 15 Å away from the acceptor. In the B*-b2 complex (right panel), the nucleophile is

positioned 4 Å away and almost ready for the branching reaction.   (D) A close-up

view on the 5'SS/U6 duplex (upper panel) and the 5'-exon/U5 duplex (bottom panel)

in the B*-a2 complex. The 5'SS is recognized by the ACAGA box of U6 snRNA

through base-paring interactions between $U_4G_5U_6$ of pre-mRNA and $A_{47}C_{48}A_{49}$ of U6

snRNA. The 5'-exon is recognized by loop I of U5 snRNA, with U_{-2} of 5'-exon base-

pairing with U_{97} of U5 snRNA.   (E) A close-up view on the structure of the 5'SS/U6

duplex (upper panel) and the 5'-exon/U5 duplex (bottom panel) in the B*-b2 complex.

The 5'SS is recognized identically as that in the B*-a2 complex; the 5'-exon, however,

is now anchored to loop I of U5 snRNA through base-pairing interactions between $A_{-4}A_{-3}A_{-2}G_{-1}$ of the 5'-exon and $U_{96}U_{97}U_{98}U_{99}$ of U5 snRNA.
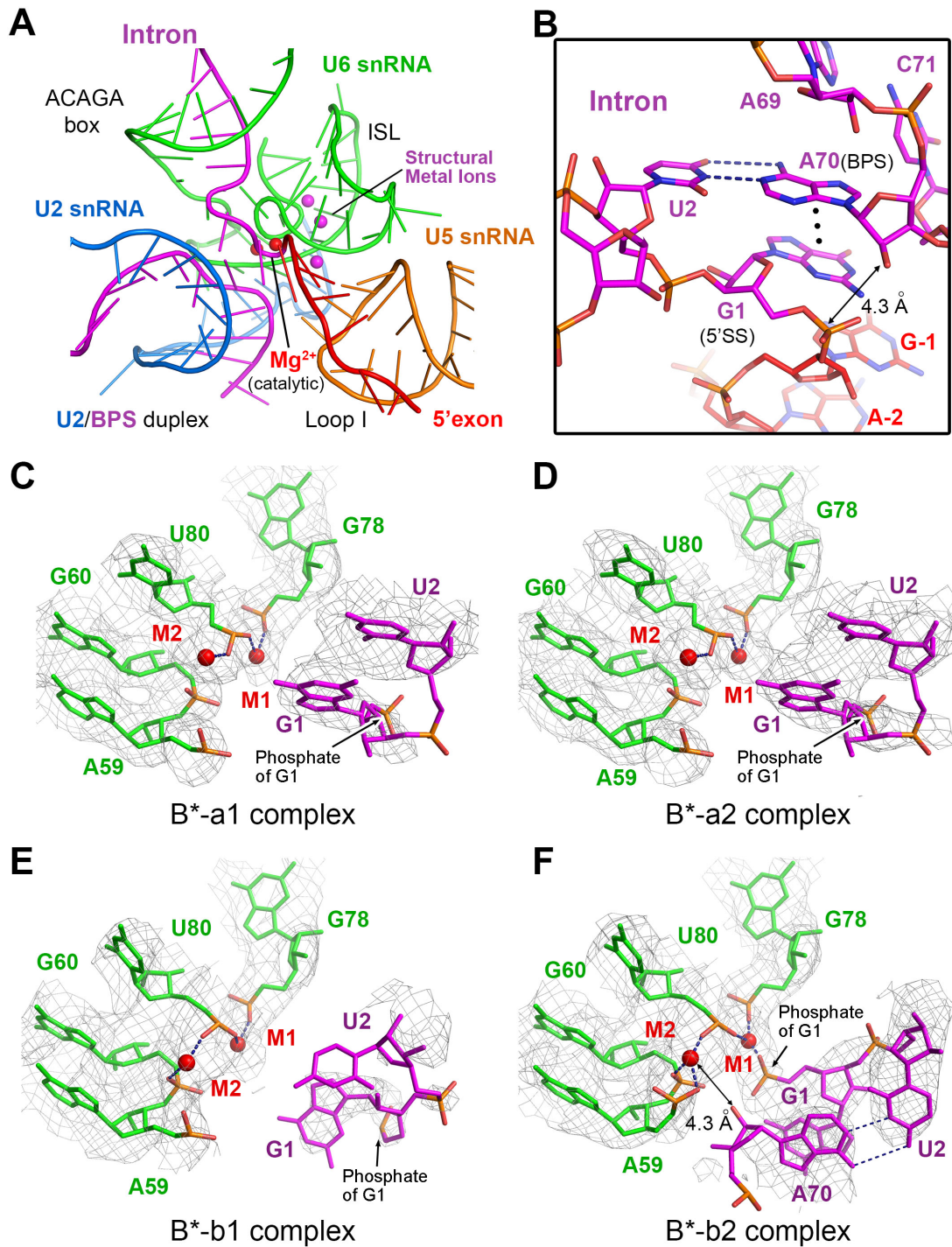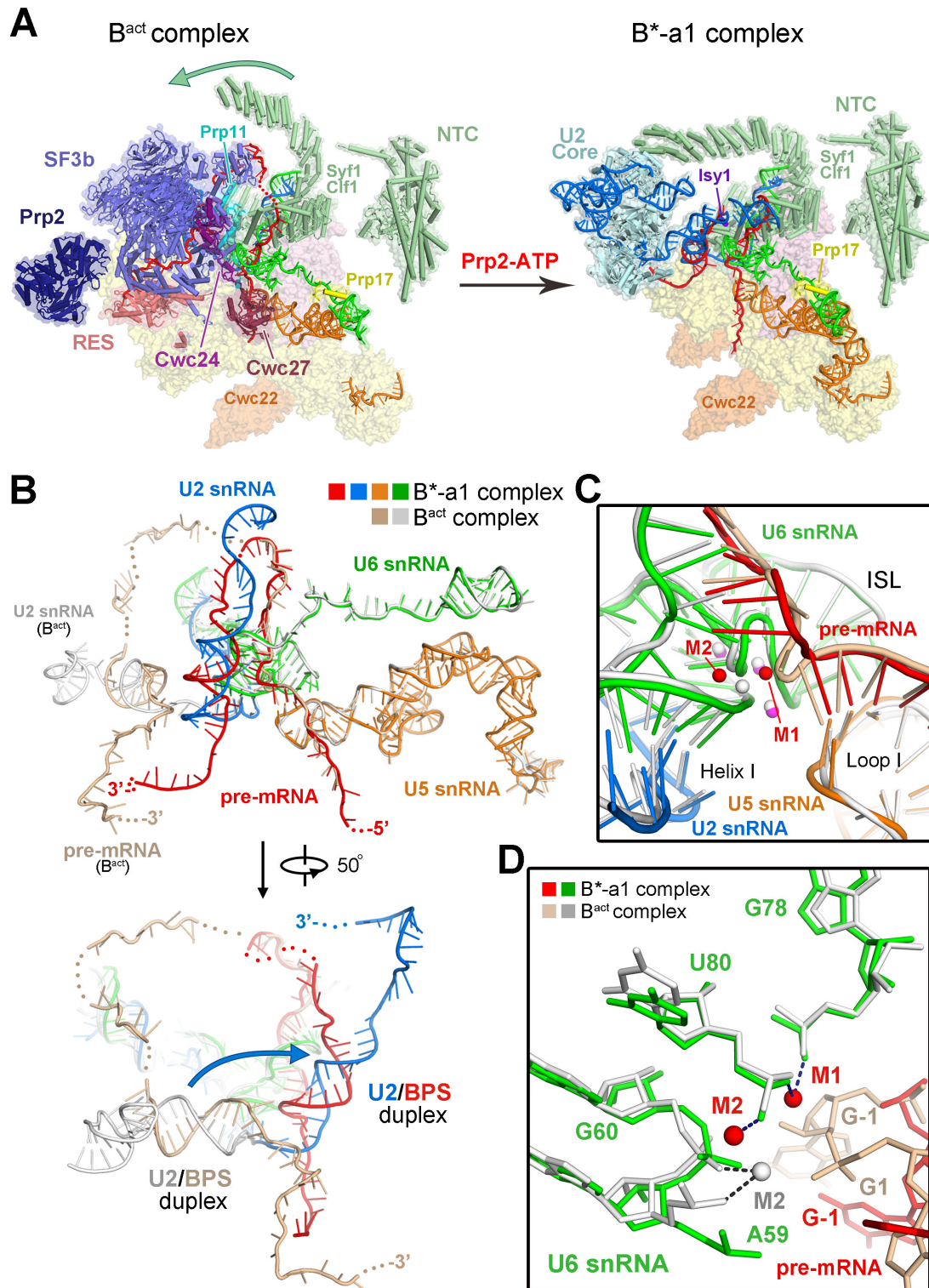
Figure 3   The splicing active site in the B*-b2 complex and the coordination of catalytic metal ions in the B* complexes.   (A) Structure of the splicing active site in the B*-b2 complex. The catalytic center comprises the intramolecular stem loop (ISL) of U6 snRNA and the associated Mg^{2+} ions, helix I of the U2/U6 duplex, and loop I of U5 snRNA. The catalytic and structural metal ions are shown as red and magenta

spheres, respectively. The U2/BPS duplex is positioned in close proximity to the 5'SS in the active site. (B) A close-up view on the recognition of the invariant adenine nucleotide A_{70} of the BPS at the active site center of the B^*-b2 complex. A_{70} of the BPS is recognized by the conserved dinucleotide G_1U_2 at the 5'-end of the 5'SS, through A_{70}-U_2 base-pairing and A_{70}-G_1 base stacking. The nucleophile 2'-OH of A_{70} is located only ~4.3 Å away from the phosphate of G_1, nearly ready for nucleophilic attack. (C) A close-up view on the coordination of the catalytic metals M1 and M2 in the B^*-a1 complex. M2 is coordinated by a phosphate from U_{80} of U6 snRNA. M1 is bound by the phosphate oxygens of G_{78} and U_{80} from U6 snRNA. (D) A close-up view on the coordination of M1 and M2 in the B^*-a2 complex, which is identical to that in the B^*-a1 complex. (E) A close-up view on the coordination of M1 and M2 in the B^*-b1 complex. Compared to that in B^*-a1 or -a2, M2 is coordinated by an additional phosphate from G_{60} of U6 snRNA. The coordination of M1 is identical to that in the B^*-a1 and -a2 complexes. (F) A close-up view on the coordination of M1 and M2 in the B^*-b2 complex. M2 is coordinated by three phosphates from A_{59}, G_{60} and U_{80} of U6 snRNA. But the 2'-OH of A_{70} is yet to be activated by M2 in this complex, with a distance of approximately 4.3 Å. M1 stabilizes the leaving group, the phosphate of G_1 in the 5'SS, through direct interaction. M1 is also coordinated by the phosphate oxygens of G_{78} and U_{80} from U6 snRNA.

Figure 4   Structural changes of the B^{act}-to-B* transition in *S. cerevisiae*.   (A)

Structural changes of the protein components during the transition from the B^{act} to B*-

a1 complex. In the B^{act} complex (Yan et al., 2016), the SF3b complex and the RES

complex recognize the U2/BPS duplex and the downstream RNA sequences of the

intron, respectively. Upon Prp2-mediated remodeling, SF3a, SF3b, RES and the splicing factors Cwc24 and Cwc27 are dissociated to release the U2/BPS duplex and the 5'SS. The U2 snRNP core and the NTC proteins Syf1 and Clf1 undergo marked translocation. Isy1 associates with the RNA elements at the active site center. Consequently, the U2/BPS duplex is delivered close to the active site. (B) Structural comparison of the overall RNA elements between the B^{act} complex (Yan et al., 2016) and the B^*-a1 complex. Two views are shown. The RNA elements in the B^*-a1 complex are colored identically as those in Figure 2A; in the B^{act} complex, pre-mRNA is colored wheat and all other RNA elements are shown in gray. During the B^{act}-to-B^* transition, the U2/BPS duplex undergoes a drastic translocation to the vicinity of the active site. U5 and U6 snRNAs and the first 30 nucleotides of U2 snRNA remain unchanged. (C) Structural comparison of the active site between the B^{act} (Yan et al., 2016) and B^*-a1 complexes. Compared to the B^*-a1 complex, M1 is yet to be loaded in the B^{act} complex, M2 undergoes a positional shift. (D) Comparison on the coordination of the catalytic metals between the B^{act} and B^*-a1 complexes. In the B^{act} complex, M1 is yet to be loaded, M2 is coordinated by two phosphates from A_{59} and G_{60} of U6 snRNA. In the B^*-a1 complex, M1 is already loaded into the active site and coordinated by two phosphates from G_{78} and U_{80} of U6 snRNA. M2 is coordinated by a phosphate from U_{80} of U6 snRNA.
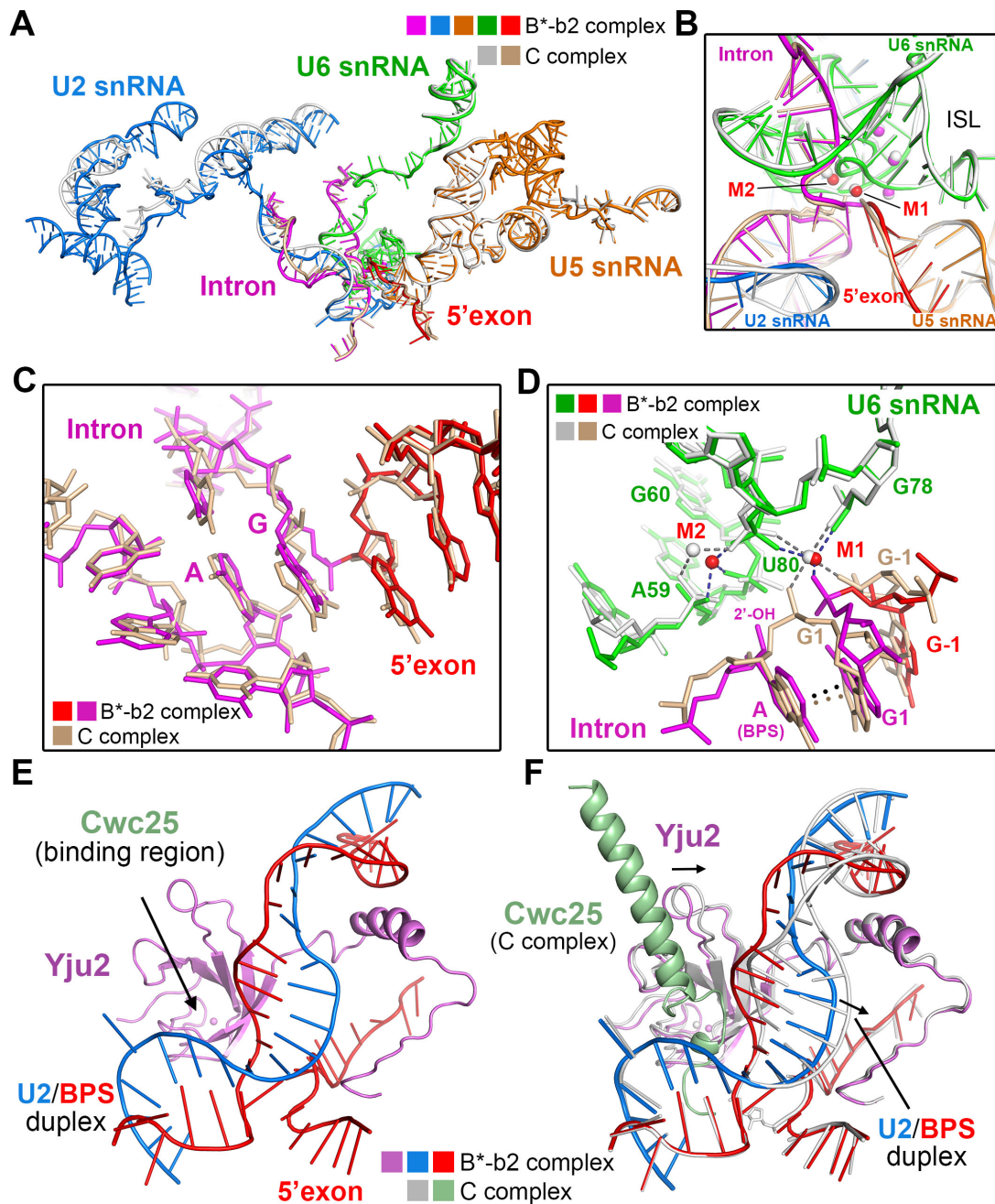
Figure 5   Structural comparison between the B*-b2 complex and the C complex from *S. cerevisiae*.   (A) Overall structural comparison of the RNA elements between the B*-b2 complex and the C complex (Wan et al., 2016b). The RNA elements in the B*-b2 complex are colored identically as those in Figure 2A; in the C complex, pre-mRNA is colored wheat and all other RNA elements are shown in gray. Except for minor differences in the 3'-end region of U2 snRNA and the BPS region, structures of

all other RNA elements are nearly identical.   (B) A close-up view on the structural comparison of the active site between the B^*-b2 and C complexes.   (C) A close-up view on the structural comparison of the region surrounding A_{70} of the BPS and G_1 of the 5'SS between the B^*-b2 and C complexes. The nucleophile and the acceptor are positioned close to each other in the B^*-b2 complex but are covalently linked together in the C complex.   (D) A close-up view on the structural comparison of catalytic metal coordination between the B^*-b2 and C complexes. In the C complex, M1 is coordinated in a planar fashion by four ligands: 3'-OH of G_{-1} of the 5'-exon and three phosphate oxygens from G_{78} and U_{80} of U6 snRNA and G_1 of 5'SS. In the B^*-b2 complex, M1 is only coordinated by three ligands and no longer interacts with 3'-OH of G_{-1} of the 5'-exon. In the C complex, M2 is bound by two phosphate oxygens from A_{59} and U_{80} of U6 snRNA; in the B^*-b2 complex, M2 is coordinated by three phosphates from A_{59}, G_{60} and U_{80} of U6 snRNA.   (E) A close-up view on the region occupied by the step I splicing factors in the B^*-b2 complex. The N-terminal loop and downstream β-sheet domain of Yju2 bind to the 5'SS and the U2/BPS duplex, respectively. Cwc25 remains to be recruited into the active site.   (F) A close-up view on the region occupied by the step I splicing factors between the B^*-b2 complex and the C complex. The β-sheet domain of Yju2 and the U2/BPS duplex exhibit apparent positional shifts between the B^*-b2 and C complexes.
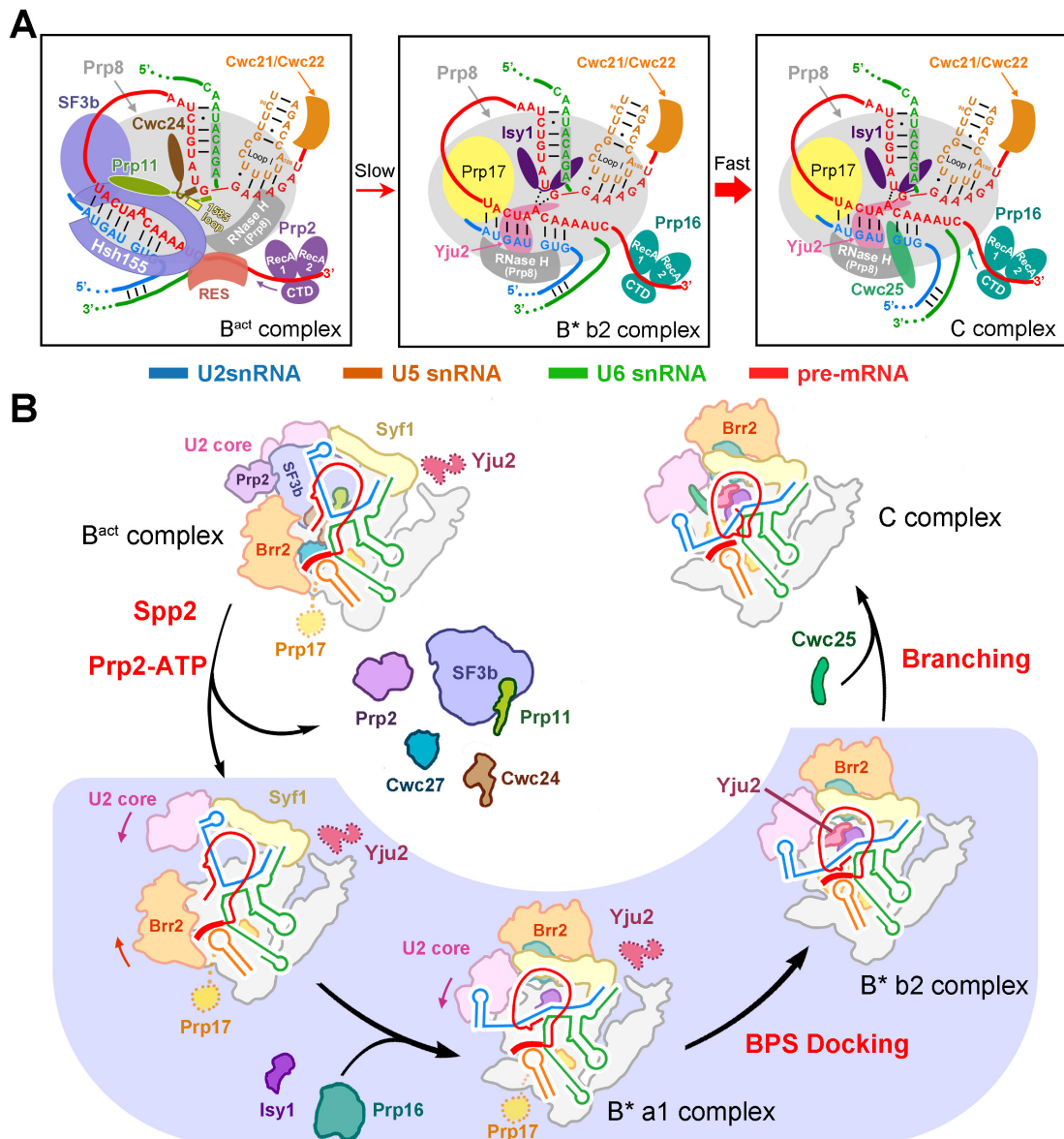
Figure 6   A mechanistic model of branching catalyzed by the spliceosome in *S. cerevisiae*. (A) A cartoon diagram of the structural changes at the active site during the B^{act}-to-B*-to-C transition. In this model, Prp2 may pull the 3'-end of the intron, dissociating proteins that are associated with the BPS and 5'SS and allowing the translocation of the U2/BPS duplex into the active site. The nucleophile-containing adenine nucleobase in BPS is stabilized through base-pairing with U$_2$ and base-stacking with G$_1$ at the 5'-end of the 5'SS. Upon binding of the step I splicing factor Cwc25, the U2/BPS duplex is pushed closer to the 5'SS, allowing the branching to

33

occur and resulting in the C complex.   (B) A schematic diagram of spliceosome remodeling by Prp2 and the branching reaction catalyzed by the spliceosome in *S. cerevisiae*. During the B^{act}-to-B^* transition, the ATPase/helicase Prp2 and its cofactor Spp2 mediate the dissociation of the SF3a and SF3b complexes and the splicing factors Cwc24 and Cwc27, freeing the 5'SS and the U2/BPS duplex. The reorganization of the protein components may allow translocation of the U2/BPS duplex into the vicinity of the splicing active site, forming a partially catalytically activated complex (bottom middle), exemplified by the B^*-a1 complex. Next, Yju2 is loaded to the spliceosome catalytic center, pushing the U2/BPS duplex closer to the 5'SS (bottom right), forming the B^*-b2 complex. In the final step, Cwc25 is recruited, allowing fine adjustment of the active RNA and protein elements and consequent branching reaction. The resulting spliceosome is the C complex.
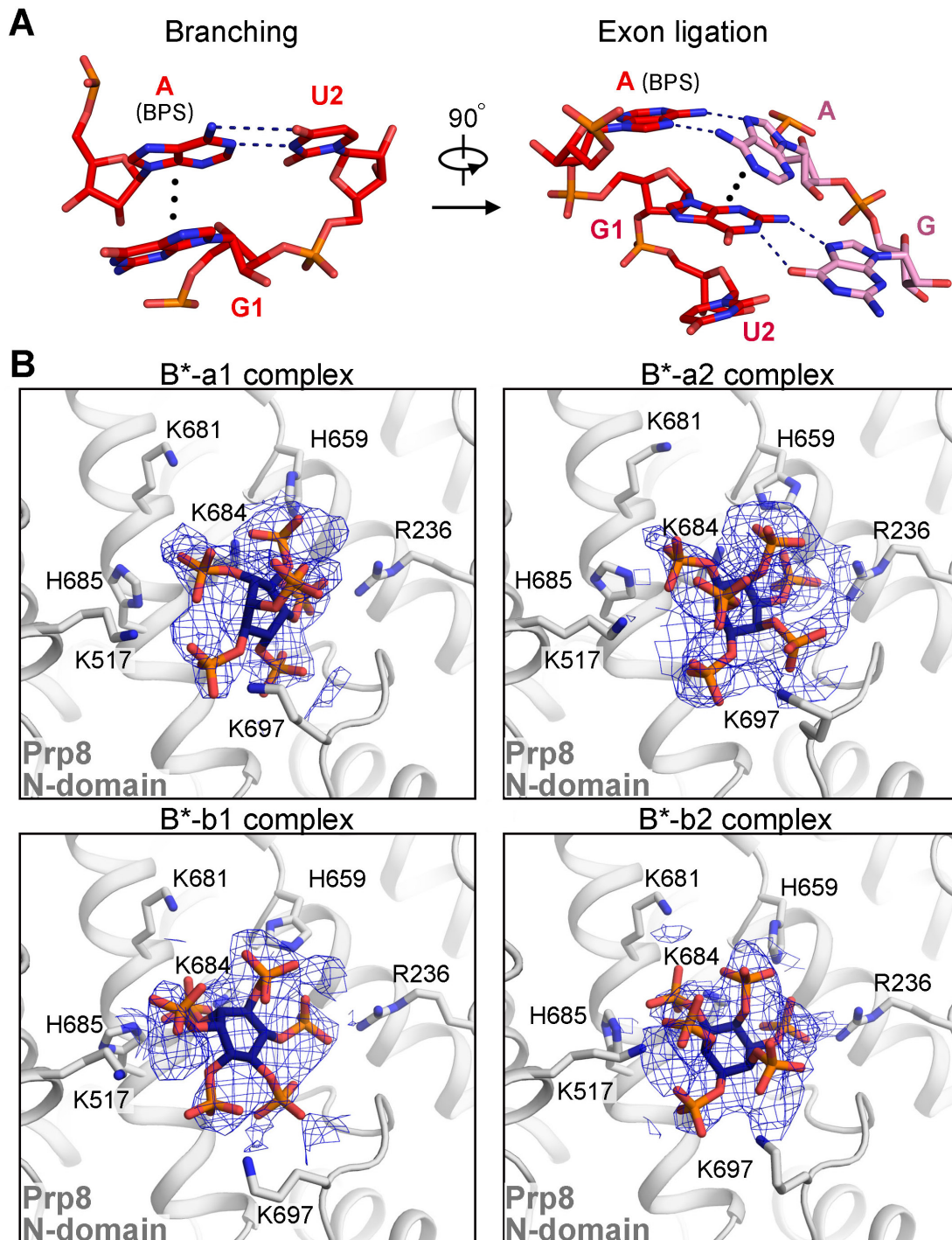
Figure 7   Conserved structural features in the B* complex and other functional states of the assembled spliceosome.   (A) Interactions among conserved intron sequences in branching and exon ligation. In the B* complex preceding the branching reaction (left panel), the nucleophile-containing A_{70} of the BPS is recognized through base-pairing with U_2 and base-stacking against G_1 of the 5'SS. In the P complex

following exon ligation (right panel), the dinucleotides AG at the 3'-end of the 3'-splice site (3'SS) are stabilized by the invariant A_{70} of the BPS and G_1 of the 5'SS through base-pairing and base-stacking interactions, respectively. (B) A small molecule interacts with Prp8 in all four B^* complexes. This small molecule, identified as inositol hexaphosphate (IP6), is identically present in the yeast B^{act} through ILS complexes. The EM density maps of IP6 and its interactions with Prp8 in the four distinct conformations of the B^* complex are shown here. IP6 is located in a positively charged cavity formed in the N-domain of Prp8.

References:

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., *et al*. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography *66*, 213-221.

Aebi, M., Hornig, H., Padgett, R.A., Reiser, J., and Weissmann, C. (1986). Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. Cell *47*, 555-565.

Aebi, M., Hornig, H., and Weissmann, C. (1987). 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. Cell *50*, 237-246.

Amunts, A., Brown, A., Bai, X.C., Llacer, J.L., Hussain, T., Emsley, P., Long, F., Murshudov, G., Scheres, S.H., and Ramakrishnan, V. (2014). Structure of the yeast mitochondrial large ribosomal subunit. Science *343*, 1485-1489.

Bai, R., Wan, R., Yan, C., Lei, J., and Shi, Y. (2018). Structures of the fully assembled Saccharomyces cerevisiae spliceosome before activation. Science *360*, 1423-1429.

Bai, R., Yan, C., Wan, R., Lei, J., and Shi, Y. (2017). Structure of the Post-catalytic Spliceosome from Saccharomyces cerevisiae. Cell *171*, 1589-1598 e1588.

Bao, P., Hobartner, C., Hartmuth, K., and Luhrmann, R. (2017). Yeast Prp2 liberates the 5' splice site and the branch site adenosine for catalysis of pre-mRNA splicing. Rna *23*, 1770-1779.

Bertram, K., Agafonov, D.E., Dybkov, O., Haselbach, D., Leelaram, M.N., Will, C.L., Urlaub, H., Kastner, B., Luhrmann, R., and Stark, H. (2017a). Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. Cell *170*, 701-713 e711.

Bertram, K., Agafonov, D.E., Liu, W.T., Dybkov, O., Will, C.L., Hartmuth, K., Urlaub, H., Kastner, B., Stark, H., and Luhrmann, R. (2017b). Cryo-EM structure of a human

spliceosome activated for step 2 of splicing. Nature *542*, 318-323.

Brody, E., and Abelson, J. (1985). The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. Science *228*, 963-967.

Burgess, S.M., and 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*, 1377-1391.

Campion, Y., Neel, H., Gostan, T., Soret, J., and Bordonne, R. (2010). Specific splicing defects in S. pombe carrying a degron allele of the Survival of Motor Neuron gene. The EMBO journal *29*, 1817-1829.

Chang, K.J., Chen, H.C., and Cheng, S.C. (2009). Ntc90 is required for recruiting first step factor Yju2 but not for spliceosome activation. Rna *15*, 1729-1739.

Chen, S., McMullan, G., Faruqi, A.R., Murshudov, G.N., Short, J.M., Scheres, S.H., and Henderson, R. (2013). High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy *135*, 24-35.

Chiu, Y.F., Liu, Y.C., Chiang, T.W., Yeh, T.C., Tseng, C.K., Wu, N.Y., and Cheng, S.C. (2009). Cwc25 is a novel splicing factor required after Prp2 and Yju2 to facilitate the first catalytic reaction. Molecular and cellular biology *29*, 5671-5678.

Clark, T.A., Sugnet, C.W., and Ares, M., Jr. (2002). Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. Science *296*, 907-910.

Cordin, O., Hahn, D., and Beggs, J.D. (2012). Structure, function and regulation of spliceosomal RNA helicases. Current opinion in cell biology *24*, 431-438.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S.*, et al*. (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic acids research *35*, W375-383.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System. on World Wide Web http://wwwpymolorg/.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D *60*, 2126-2132.

Fica, S.M., Mefford, M.A., Piccirilli, J.A., and Staley, J.P. (2014). Evidence for a group II intron-like catalytic triplex in the spliceosome. Nat Struct Mol Biol *21*, 464-471.

Fica, S.M., Oubridge, C., Galej, W.P., Wilkinson, M.E., Bai, X.C., Newman, A.J., and Nagai, K. (2017). Structure of a spliceosome remodelled for exon ligation. Nature *542*, 377-380.

Fica, S.M., Tuttle, N., Novak, T., Li, N.S., Lu, J., Koodathingal, P., Dai, Q., Staley, J.P., and Piccirilli, J.A. (2013). RNA catalyses nuclear pre-mRNA splicing. Nature *503*, 229-234.

Frendewey, D., and Keller, W. (1985). Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell *42*, 355-367.

Galej, W.P., Wilkinson, M.E., Fica, S.M., Oubridge, C., Newman, A.J., and Nagai, K. (2016). Cryo-EM structure of the spliceosome immediately after branching. Nature *537*, 197-201.

Gietz, R.D., and Schiestl, R.H. (2007). Quick and easy yeast transformation using the

LiAc/SS carrier DNA/PEG method. Nat Protoc *2*, 35-37.

Grabowski, P.J., Padgett, R.A., and Sharp, P.A. (1984). Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. Cell *37*, 415-427.

Grabowski, P.J., Seiler, S.R., and Sharp, P.A. (1985). A multicomponent complex is involved in the splicing of messenger RNA precursors. Cell *42*, 345-353.

Hang, J., Wan, R., Yan, C., and Shi, Y. (2015). Structural basis of pre-mRNA splicing. Science *349*, 1191-1198.

Haselbach, D., Komarov, I., Agafonov, D.E., Hartmuth, K., Graf, B., Dybkov, O., Urlaub, H., Kastner, B., Luhrmann, R., and Stark, H. (2018). Structure and Conformational Dynamics of the Human Spliceosomal B(act) Complex. Cell *172*, 454-464 e411.

Hilliker, A.K., Mefford, M.A., and Staley, J.P. (2007). U2 toggles iteratively between the stem IIa and stem IIc conformations to promote pre-mRNA splicing. Genes & development *21*, 821-834.

Jankowsky, E. (2011). RNA helicases at work: binding and rearranging. Trends in biochemical sciences *36*, 19-29.

Kawashima, T., Pellegrini, M., and Chanfreau, G.F. (2009). Nonsense-mediated mRNA decay mutes the splicing defects of spliceosome component mutations. Rna *15*, 2236-2247.

Keating, K.S., Toor, N., Perlman, P.S., and Pyle, A.M. (2010). A structural analysis of the group II intron active site and implications for the spliceosome. Rna *16*, 1-9.

Kim, S.H., and Lin, R.J. (1993). Pre-mRNA splicing within an assembled yeast spliceosome requires an RNA-dependent ATPase and ATP hydrolysis. Proc Natl Acad Sci U S A *90*, 888-892.

Kim, S.H., and Lin, R.J. (1996). Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. Molecular and cellular biology *16*, 6810-6819.

Kimanius, D., Forsberg, B.O., Scheres, S.H., and Lindahl, E. (2016). Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. eLife *5*.

King, D.S., and Beggs, J.D. (1990). Interactions of PRP2 protein with pre-mRNA splicing complexes in Saccharomyces cerevisiae. Nucleic acids research *18*, 6559-6564.

Koodathingal, P., Novak, T., Piccirilli, J.A., and Staley, J.P. (2010). The DEAH box ATPases Prp16 and Prp43 cooperate to proofread 5' splice site cleavage during pre-mRNA splicing. Molecular cell *39*, 385-395.

Krishnan, R., Blanco, M.R., Kahlscheuer, M.L., Abelson, J., Guthrie, C., and Walter, N.G. (2013). Biased Brownian ratcheting leads to pre-mRNA remodeling and capture prior to first-step splicing. Nat Struct Mol Biol *20*, 1450-1457.

Lamond, A.I., Konarska, M.M., and Sharp, P.A. (1987). A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. Genes & development *1*, 532-543.

Lardelli, R.M., Thompson, J.X., Yates, J.R., 3rd, and Stevens, S.W. (2010). Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. Rna *16*, 516-528.

Lei, J., and Frank, J. (2005). Automated acquisition of cryo-electron micrographs for

single particle reconstruction on an FEI Tecnai electron microscope. Journal of structural biology *150*, 69-80.

Liu, H.L., and Cheng, S.C. (2012). The interaction of Prp2 with a defined region of the intron is required for the first splicing reaction. Molecular and cellular biology *32*, 5056-5066.

Liu, S., Li, X., Zhang, L., Jiang, J., Hill, R.C., Cui, Y., Hansen, K.C., Zhou, Z.H., and Zhao, R. (2017). Structure of the yeast spliceosomal postcatalytic P complex. Science *358*, 1278-1283.

Liu, Y.C., Chen, H.C., Wu, N.Y., and 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. Molecular and cellular biology *27*, 5403-5413.

Liu, Y.C., and Cheng, S.C. (2015). Functional roles of DExD/H-box RNA helicases in Pre-mRNA splicing. J Biomed Sci *22*, 54.

Mayas, R.M., Maita, H., and Staley, J.P. (2006). Exon ligation is proofread by the DExD/H-box ATPase Prp22p. Nat Struct Mol Biol *13*, 482-490.

Nicholls, R.A., Fischer, M., McNicholas, S., and Murshudov, G.N. (2014). Conformation-independent structural comparison of macromolecules with ProSMART. Acta crystallographica Section D, Biological crystallography *70*, 2487-2499.

Ohrt, T., Odenwalder, P., Dannenberg, J., Prior, M., Warkocki, Z., Schmitzova, J., Karaduman, R., Gregor, I., Enderlein, J., Fabrizio, P., *et al.* (2013). Molecular dissection of step 2 catalysis of yeast pre-mRNA splicing investigated in a purified system. Rna *19*, 902-915.

Ohrt, T., Prior, M., Dannenberg, J., Odenwalder, P., Dybkov, O., Rasche, N., Schmitzova, J., Gregor, I., Fabrizio, P., Enderlein, J., *et al.* (2012). Prp2-mediated protein rearrangements at the catalytic core of the spliceosome as revealed by dcFCCS. Rna *18*, 1244-1256.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984). Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. Science *225*, 898-903.

Park, J.W., Parisky, K., Celotto, A.M., Reenan, R.A., and Graveley, B.R. (2004). Identification of alternative splicing regulators by RNA interference in Drosophila. Proc Natl Acad Sci U S A *101*, 15974-15979.

Perriman, R.J., and Ares, M., Jr. (2007). Rearrangement of competing U2 RNA helices within the spliceosome promotes multiple steps in splicing. Genes & development *21*, 811-820.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. Journal of computational chemistry *25*, 1605-1612.

Plaschka, C., Lin, P.C., and Nagai, K. (2017). Structure of a pre-catalytic spliceosome. Nature *546*, 617-621.

Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. PLoS biology *5*, e90.

Query, C.C., and Konarska, M.M. (2004). Suppression of multiple substrate mutations

by spliceosomal prp8 alleles suggests functional correlations with ribosomal ambiguity mutants. Molecular cell *14*, 343-354.

Rauhut, R., Fabrizio, P., Dybkov, O., Hartmuth, K., Pena, V., Chari, A., Kumar, V., Lee, C.T., Urlaub, H., Kastner, B., *et al*. (2016a). Molecular architecture of the Saccharomyces cerevisiae activated spliceosome. Science *353*, 1399-1405.

Rauhut, R., Fabrizio, P., Dybkov, O., Hartmuth, K., Pena, V., Chari, A., Kumar, V., Lee, C.T., Urlaub, H., Kastner, B., *et al*. (2016b). Molecular architecture of the Saccharomyces cerevisiae activated spliceosome. Science.

Rosenthal, P.B., and Henderson, R. (2003). Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. Journal of molecular biology *333*, 721-745.

Roy, J., Kim, K., Maddock, J.R., Anthony, J.G., and Woolford, J.L., Jr. (1995). The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. RNA *1*, 375-390.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell *38*, 317-331.

Saltzman, A.L., Pan, Q., and Blencowe, B.J. (2011). Regulation of alternative splicing by the core spliceosomal machinery. Genes & development *25*, 373-384.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. Journal of structural biology *180*, 519-530.

Schneider, C., Agafonov, D.E., Schmitzova, J., Hartmuth, K., Fabrizio, P., and Luhrmann, R. (2015). Dynamic Contacts of U2, RES, Cwc25, Prp8 and Prp45 Proteins with the Pre-mRNA Branch-Site and 3' Splice Site during Catalytic Activation and Step 1 Catalysis in Yeast Spliceosomes. PLoS genetics *11*, e1005539.

Semlow, D.R., Blanco, M.R., Walter, N.G., and Staley, J.P. (2016). Spliceosomal DEAH-Box ATPases Remodel Pre-mRNA to Activate Alternative Splice Sites. Cell *164*, 985-998.

Shi, Y. (2017a). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. Nature reviews Molecular cell biology *18*, 655-670.

Shi, Y. (2017b). The Spliceosome: A Protein-Directed Metalloribozyme. Journal of molecular biology *429*, 2640-2653.

Steitz, T.A., and Steitz, J.A. (1993). A general two-metal-ion mechanism for catalytic RNA. Proc Natl Acad Sci U S A *90*, 6498-6502.

Tseng, C.K., Chung, C.S., Chen, H.C., and Cheng, S.C. (2017). A central role of Cwc25 in spliceosome dynamics during the catalytic phase of pre-mRNA splicing. Rna *23*, 546-556.

Villa, T., and Guthrie, C. (2005). The Isy1p component of the NineTeen complex interacts with the ATPase Prp16p to regulate the fidelity of pre-mRNA splicing. Genes & development *19*, 1894-1904.

Wan, R., Yan, C., Bai, R., Huang, G., and Shi, Y. (2016a). Structure of a yeast catalytic step I spliceosome at 3.4 A resolution. Science.

Wan, R., Yan, C., Bai, R., Huang, G., and Shi, Y. (2016b). Structure of a yeast catalytic step I spliceosome at 3.4 A resolution. Science *353*, 895-904.

Wan, R., Yan, C., Bai, R., Lei, J., and Shi, Y. (2017). Structure of an Intron Lariat

Spliceosome from Saccharomyces cerevisiae. Cell *171*, 120-132 e112.

Wang, F., Gong, H., Liu, G., Li, M., Yan, C., Xia, T., Li, X., and Zeng, J. (2016). DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. Journal of structural biology *195*, 325-336.

Warkocki, Z., Odenwalder, P., Schmitzova, J., Platzmann, F., Stark, H., Urlaub, H., Ficner, R., Fabrizio, P., and Luhrmann, R. (2009). Reconstitution of both steps of Saccharomyces cerevisiae splicing with purified spliceosomal components. Nat Struct Mol Biol *16*, 1237-1243.

Warkocki, Z., Schneider, C., Mozaffari-Jovin, S., Schmitzova, J., Hobartner, C., Fabrizio, P., and Luhrmann, R. (2015). The G-patch protein Spp2 couples the spliceosome-stimulated ATPase activity of the DEAH-box protein Prp2 to catalytic activation of the spliceosome. Genes & development *29*, 94-107.

Wilkinson, M.E., Fica, S.M., Galej, W.P., Norman, C.M., Newman, A.J., and Nagai, K. (2017). Postcatalytic spliceosome structure reveals mechanism of 3'-splice site selection. Science *358*, 1283-1288.

Will, C.L., and Luhrmann, R. (2011). Spliceosome structure and function. Cold Spring Harb Perspect Biol *3*.

Yan, C., Hang, J., Wan, R., Huang, M., Wong, C.C., and Shi, Y. (2015). Structure of a yeast spliceosome at 3.6-angstrom resolution. Science *349*, 1182-1191.

Yan, C., Wan, R., Bai, R., Huang, G., and Shi, Y. (2016). Structure of a yeast activated spliceosome at 3.5 A resolution. Science *353*, 904-911.

Yan, C., Wan, R., Bai, R., Huang, G., and Shi, Y. (2017). Structure of a yeast step II catalytically activated spliceosome. Science *355*, 149-155.

Yean, S.L., Wuenschell, G., Termini, J., and Lin, R.J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. Nature *408*, 881-884.

Zhan, X., Yan, C., Zhang, X., Lei, J., and Shi, Y. (2018a). Structure of a human catalytic step I spliceosome. Science *359*, 537-545.

Zhan, X., Yan, C., Zhang, X., Lei, J., and Shi, Y. (2018b). Structures of the human pre-catalytic spliceosome and its precursor spliceosome. Cell Res.

Zhang, K. (2016). Gctf: Real-time CTF determination and correction. Journal of structural biology *193*, 1-12.

Zhang, X., Yan, C., Hang, J., Finci, L.I., Lei, J., and Shi, Y. (2017). An Atomic Structure of the Human Spliceosome. Cell *169*, 918-929 e914.

Zhang, X., Yan, C., Zhan, X., Li, L., Lei, J., and Shi, Y. (2018). Structure of the human activated spliceosome in three conformational states. Cell research *28*, 307-322.

Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nature methods *14*, 331-332.