Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes

Berthold Kastner,[1] Cindy L. Will,[1] Holger Stark,[2] and Reinhard Lührmann[1]

[1]Department of Cellular Biochemistry, Max Planck Institute for Biophysical Chemistry, D-37077 Göttingen, Germany

[2]Department of Structural Dynamics, Max Planck Institute for Biophysical Chemistry, D-37077 Göttingen, Germany

Correspondence: reinhard.luehrmann@mpibpc.mpg.de

SUMMARY

The spliceosome is a highly complex, dynamic ribonucleoprotein molecular machine that undergoes numerous structural and compositional rearrangements that lead to the formation of its active site. Recent advances in cyroelectron microscopy (cryo-EM) have provided a plethora of near-atomic structural information about the inner workings of the spliceosome. Aided by previous biochemical, structural, and functional studies, cryo-EM has confirmed or provided a structural basis for most of the prevailing models of spliceosome function, but at the same time allowed novel insights into splicing catalysis and the intriguing dynamics of the spliceosome. The mechanism of pre-mRNA splicing is highly conserved between humans and yeast, but the compositional dynamics and ribonucleoprotein (RNP) remodeling of the human spliceosome are more complex. Here, we summarize recent advances in our understanding of the molecular architecture of the human spliceosome, highlighting differences between the human and yeast splicing machineries.

Outline

Cold Spring Harbor Perspectives in Biology

B. Kastner et al.

1 INTRODUCTION

Pre-mRNA introns are defined by three short consensus sequences, the 5′ splice site (SS), the 3′SS, and the branch site (BS,) which is often followed by a pyrimidine-rich region (the PPT) (Fig. 1A) (Will and Lührmann 2011). Pre-mRNA splicing involves two sequential SN2-type transesterification reactions (Fig. 1B). First, the 2′-OH group of the branch site adenosine (BS-A) carries out a nucleophilic attack at the phosphodiester bond of the 5′SS, producing the cleaved 5′ exon and the intron lariat-3′ exon intermediates. During step 2, the 3′-OH group of the 5′ exon attacks the 3′SS, leading to exon ligation and intron excision. In higher eukaryotes, most pre-mRNAs undergo alternative splicing (Pan et al. 2008), during which different sets of exons from a pre-mRNA are spliced together, leading to different mRNA isoforms and vastly amplifying the palette of translated proteins. Furthermore, errors in splicing or its regulation are either the basis for or modulate the severity of many human diseases (Daguenet et al. 2015; Lee and Rio 2015).

Nuclear pre-mRNA introns contain relatively little conserved secondary or tertiary structure information and thus depend on the spliceosome to align their reactive sites (i.e., the 5′SS, BS, and 3′SS) for splicing catalysis. The major spliceosomal building blocks are the U1, U2, U4/U6, and U5 small nuclear ribonucleoproteins (snRNPs). Each consists of a small nuclear RNA (snRNA) (or two, in the case of U4/U6), seven Sm proteins that form a ring (or seven Lsm proteins in the case of U6), and a variable number of snRNP-specific proteins (Fig. 1F) (Will and Lührmann 2011). Under splicing conditions, the U4/U6 and U5 snRNPs form a 25S U4/U6.U5 tri-snRNP, in which the U4 and U6 snRNAs are base-paired, forming a three-way junction. The func-

tionally important regions of the snRNAs are highly conserved, as is—for the most part—the protein composition of the snRNPs, although some proteins that are snRNP-associated in yeast are instead recruited to the human spliceosome as non-snRNP associated factors, and vice versa.

Spliceosome assembly proceeds stepwise and the generation of its active site requires extensive structural and compositional rearrangements (Fig. 1C) (Will and Lührmann 2011). These are driven mainly by eight conserved DEXD/H-box ATPases or RNA helicases that trigger RNA–RNA and RNA–protein rearrangements (Cordin and Beggs 2013). Initially, U1 snRNP interacts with the pre-mRNA, via base-pairing between the 5′ end of U1 snRNA and the 5′SS, and is stabilized by U1-associated proteins. Subsequently, U2 snRNP stably interacts with the BS, forming the spliceosomal A complex. In the latter, U2 snRNA base pairs with the BS, whereby the BS-A bulges out from the U2/BS helix. The U2–BS interaction is reinforced by the U2AF65/35 heterodimer, which binds the PPT, and by the U2-associated, heteromeric protein complexes SF3a and SF3b that contact the intron 6 to 26 nucleotides upstream of the BS-A (the so-called anchoring site) (Gozani et al. 1996). Next, the preformed U4/U6.U5 tri-snRNP joins the A complex, generating the pre-B complex in which the tri-snRNP is not yet stably bound (Boesler et al. 2016). Initial tri-snRNP docking is facilitated by the formation of a short duplex (U2/U6 helix II) between the 5′ and 3′ ends of the U2 and U6 snRNA, respectively (Fig. 1D). In humans, stable tri-snRNP integration requires the action of the helicase PRP28 (Boesler et al. 2016), which disrupts the U1/5′SS helix, allowing the latter to base-pair with the conserved U6 ACAGA box (Staley and Guthrie 1999). This destabilizes the U1 snRNP and yields the precatalytic B complex. In the latter, loop 1 of U5 snRNA interacts with

Figure 1. (*See figure on following page.*) Assembly pathway, RNA network, and compositional dynamics of the spliceosome. (*A*) Conserved sequences at the 5′ splice site (SS), branch site (BS), and 3′SS of U2-type (major) pre-mRNA introns in metazoans and the yeast *Saccharomyces cerevisiae*. (*B*) Schematic of the two-step splicing reaction. p, phosphates at the 5′ and 3′ splice sites, which are conserved in the splicing products. (*C*) Assembly, catalytic activation, and disassembly pathway of the spliceosome. For simplicity, the ordered interactions of the U1, U2, U4/U6, and U5 small nuclear ribonucleoproteins (snRNPs), but not non-snRNP spliceosomal proteins, are shown. Helicases required for splicing in both yeast and humans are indicated and include the Ski2-like helicase BBR2, the DEAD-box helicases UAP56, PRP5, and PRP28, and the DEAH-box helicases PRP2, PRP16, PRP22, and PRP43. (*D*) Dynamic network of RNA–RNA interactions in the spliceosomal B and B^act complexes. (*E*) 3D structure of the catalytic RNA network (in the human C* complex), showing the coordination of the catalytic magnesium ions M1 and M2. (*F*) Dynamic exchange of spliceosomal proteins during splicing. Proteins present in pre-B, B, B^act, C, or C* human spliceosomal complexes are indicated by a square, in which blue denotes highly abundant and gray moderately abundant proteins. Serine-arginine (SR) dipeptide-rich proteins and heterogeneous nuclear ribonucleoproteins (hnRNPs), as well as those present in very low amounts, are not shown. The table summarizes the proteomes of various, recently purified human spliceosomal complexes (Agafonov et al. 2011; De et al. 2015; Boesler et al. 2016; Bertram et al. 2017b), as well as our unpublished work. *Prp38 and Snu23 are associated with the tri-snRNP in *S. cerevisiae*. (*G*) Differential pre-organization of yeast nineteen complex (NTC) proteins in humans. (*H*) Alternative secondary structures of the U2 small nuclear RNA (snRNA).
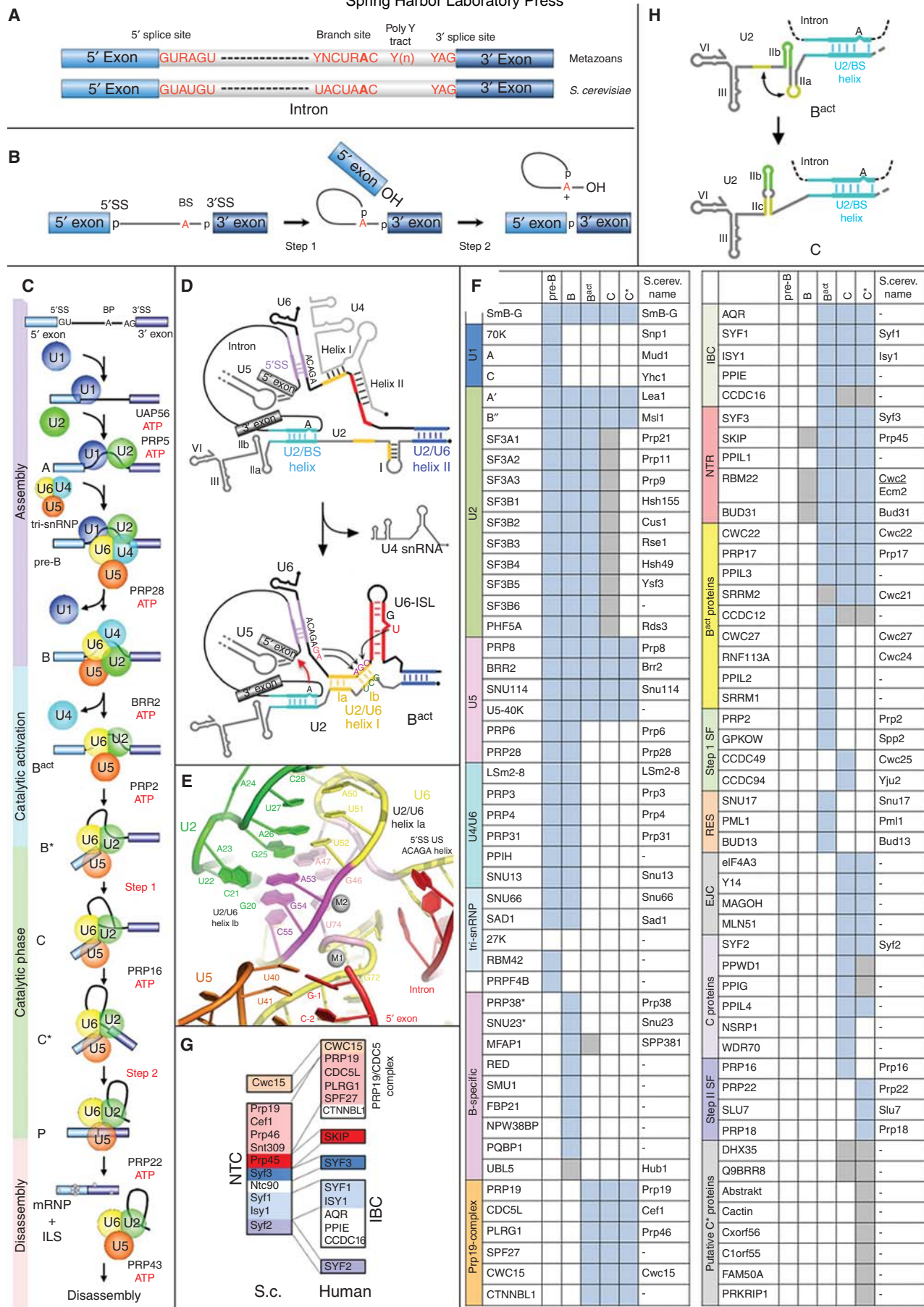
Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a032417

A — 5′ splice site / Branch site / Poly Y tract / 3′ splice site

5′ Exon	GURAGU – – – – – YNCURAC Y(n) YAG	3′ Exon	Metazoans
5′ Exon	GUAUGU – – – – – UACUAAC YAG	3′ Exon	S. cerevisiae

Intron

B — Step 1 / Step 2 splicing chemistry (5′SS, BS, 3′SS; 5′ exon, 3′ exon)

C — Assembly / Catalytic activation / Catalytic phase / Disassembly pathway (U1, U2, U4, U5, U6, tri-snRNP, pre-B, B, Bact, B*, C, C*, P, mRNP + ILS, Disassembly; UAP56 ATP, PRP5 ATP, PRP28 ATP, BRR2 ATP, PRP2 ATP, PRP16 ATP, PRP22 ATP, PRP43 ATP)

D — U6/U4/U5/U2/intron snRNA rearrangements; U2/BS helix, U2/U6 helix I/II, U6-ISL, Bact

E — U2/U6 helix structural model (U2, U5, U6, intron, 5′ exon)

G — NTC / IBC protein correspondence (S.c. / Human): Cwc15; Prp19, Cef1, Prp46, Snt309, Prp45, Syf3, Ntc90, Syf1, Isy1, Syf2 → CWC15, PRP19, CDC5L, PLRG1, SPF27, CTNNBL1 (PRP19/CDC5 complex); SKIP; SYF3; SYF1, ISY1, AQR, PPIE, CCDC16 (IBC); SYF2

H — U2/BS helix, Bact → C transition (U2, IIa, IIb, IIc, III, VI, Intron, A)

F — Spliceosomal protein composition across complexes

	Protein	pre-B	B	Bact	C	C*	S.cerev. name
	SmB-G						SmB-G
U1	70K						Snp1
U1	A						Mud1
U1	C						Yhc1
U2	A′						Lea1
U2	B″						Msl1
U2	SF3A1						Prp21
U2	SF3A2						Prp11
U2	SF3A3						Prp9
U2	SF3B1						Hsh155
U2	SF3B2						Cus1
U2	SF3B3						Rse1
U2	SF3B4						Hsh49
U2	SF3B5						Ysf3
U2	SF3B6						-
U2	PHF5A						Rds3
U5	PRP8						Prp8
U5	BRR2						Brr2
U5	SNU114						Snu114
U5	U5-40K						-
U5	PRP6						Prp6
U5	PRP28						Prp28
U4/U6	LSm2-8						LSm2-8
U4/U6	PRP3						Prp3
U4/U6	PRP4						Prp4
U4/U6	PRP31						Prp31
U4/U6	PPIH						-
U4/U6	SNU13						Snu13
tri-snRNP	SNU66						Snu66
tri-snRNP	SAD1						Sad1
tri-snRNP	27K						-
tri-snRNP	RBM42						-
tri-snRNP	PRPF4B						-
B-specific	PRP38*						Prp38
B-specific	SNU23*						Snu23
B-specific	MFAP1						SPP381
B-specific	RED						-
B-specific	SMU1						-
B-specific	FBP21						-
B-specific	NPW38BP						-
B-specific	PQBP1						-
B-specific	UBL5						Hub1
Prp19-complex	PRP19						Prp19
Prp19-complex	CDC5L						Cef1
Prp19-complex	PLRG1						Prp46
Prp19-complex	SPF27						-
Prp19-complex	CWC15						Cwc15
Prp19-complex	CTNNBL1						-

	Protein	pre-B	B	Bact	C	C*	S.cerev. name
IBC	AQR						-
IBC	SYF1						Syf1
IBC	ISY1						Isy1
IBC	PPIE						-
IBC	CCDC16						-
NTR	SYF3						Syf3
NTR	SKIP						Prp45
NTR	PPIL1						-
NTR	RBM22						Cwc2 / Ecm2
NTR	BUD31						Bud31
Bact proteins	CWC22						Cwc22
Bact proteins	PRP17						Prp17
Bact proteins	PPIL3						-
Bact proteins	SRRM2						Cwc21
Bact proteins	CCDC12						-
Bact proteins	CWC27						Cwc27
Bact proteins	RNF113A						Cwc24
Bact proteins	PPIL2						-
Bact proteins	SRRM1						-
Step 1 SF	PRP2						Prp2
Step 1 SF	GPKOW						Spp2
Step 1 SF	CCDC49						Cwc25
Step 1 SF	CCDC94						Yju2
RES	SNU17						Snu17
RES	PML1						Pml1
RES	BUD13						Bud13
EJC	eIF4A3						-
EJC	Y14						-
EJC	MAGOH						-
EJC	MLN51						-
C proteins	SYF2						Syf2
C proteins	PPWD1						-
C proteins	PPIG						-
C proteins	PPIL4						-
C proteins	NSRP1						-
C proteins	WDR70						-
Step II SF	PRP16						Prp16
Step II SF	PRP22						Prp22
Step II SF	SLU7						Slu7
Step II SF	PRP18						Prp18
Putative C* proteins	DHX35						-
Putative C* proteins	Q9BRR8						-
Putative C* proteins	Abstrakt						-
Putative C* proteins	Cactin						-
Putative C* proteins	Cxorf56						-
Putative C* proteins	C1orf55						-
Putative C* proteins	FAM50A						-
Putative C* proteins	PRKRIP1						-

Figure 1. (*See legend on previous page.*)

Cold Spring Harbor Perspectives in Biology

the 3′-terminal nucleotides of the 5′ exon, and later aids in positioning both exons for step 2 of splicing (Newman and Norman 1992; Sontheimer and Steitz 1993).

Catalytic activation occurs stepwise, with RNA helicase BRR2 initially unwinding the U4/U6 snRNA duplex (Cordin and Beggs 2013). After displacement of U4 snRNP and integration of numerous proteins, the B^{act} complex is generated. Subsequent action of RNA helicase PRP2 transforms B^{act} into a catalytically competent B* complex, which catalyzes step 1 of splicing (branching), yielding the C complex. The latter contains the cleaved 5′ exon and the intron lariat-3′ exon. Remodeling of the C complex by the helicase PRP16 generates the C* complex, which performs the second catalytic step (exon ligation), yielding the P (postsplicing) complex (Horowitz 2012). The release of the mRNA, which generates the intron lariat spliceosome (ILS), is catalyzed by the helicase PRP22. Finally, the ILS is dismantled by PRP43 helicase, and the U2, U5, and U6 snRNPs that are released are recycled for additional rounds of splicing.

The spliceosomal RNA network is restructured into a catalytically active conformation during the transformation of B into B* (Fig. 1) (Will and Lührmann 2011). After disruption of U4/U6 stems I and II, U6 snRNA is extensively restructured. The U6 internal stem-loop (U6 ISL) is generated, and U6 forms additional base-pairs with U2, yielding the U2/U6 helices Ia and Ib. A catalytic U6 RNA triplex is also formed, whereby two nucleotides of the U6 ACAGA-GAG sequence and U6-U74 form Hoogsteen base-pair interactions with the U6-catalytic triad AGC sequence of helix Ib (Fig. 1D,E) (Fica et al. 2014; Hang et al. 2015). Phosphorothioate substitution studies showed that U6 catalyzes both steps of pre-mRNA splicing by positioning two Mg^{2+} ions that stabilize the leaving groups during each reaction (Fica et al. 2013). The active site RNA conformations of the spliceosome and group II introns are highly similar (Zhao and Pyle 2017), but the former requires proteins for its formation and stabilization as described below.

Human and *Saccharomyces cerevisiae* (subsequently referred to as yeast) spliceosomes share a common set of about 85 core spliceosomal proteins that are evolutionarily conserved (Fabrizio et al. 2009). In higher eukaryotes, about 80 additional proteins, with no counterparts in yeast, copurify with the splicing machinery throughout its assembly and catalytic stages (Jurica and Moore 2003; Agafonov et al. 2011). These include, among others, (1) numerous SR and hnRNP proteins, which play a major part in regulating alternative splicing (Zhou and Fu 2013), and (2) proteins with potential enzymic functions that may act as molecular switches or conformational driving forces—that is, kinases, several RNA helicases including Aquarius (AQR, also termed IBP160), and seven peptidyl-prolyl isomerases (PPIases). Human spliceosomal proteins are also subject to substantially more posttranslational modification than yeast proteins. Many spliceosomal proteins, in addition to the snRNPs, are pre-organized into large functional ensembles, which simplifies the assembly landscape and enhances the kinetics of spliceosome assembly. These include the retention and splicing (RES) complex (Dziembowski et al. 2004), the splicing essential PRP19/CDC5L complex (Makarova et al. 2004), and the pentameric intron-binding complex (IBC) (De et al. 2015) (Fig. 1F). The yeast counterparts of some IBC and PRP19/CDC5L complex proteins are pre-organized in a different manner (Fig. 1G). That is, they are found in the large Prp19 complex (NTC, nineteen complex) together with additional spliceosomal proteins, some of which in human are recruited independently (i.e., SKIP, SYF2, SYF3) and/or at a different stage (SYF2) of splicing. Thus, not only is the assembly of the human spliceosome more complex, with likely additional assembly intermediates, but also more highly regulated.

The protein composition of the spliceosome changes from one spliceosomal complex to the next (Fig. 1F) (Will and Lührmann 2011). For example, during the pre-B to B complex transition, nine B-specific proteins are stably integrated. The greatest exchange of proteins occurs during the B to B^{act} transition, in which in humans about 20 proteins are released/destabilized, and more than 25 proteins are newly integrated. During this transition, U4/U6- and B-specific proteins are destabilized/released, whereas the PRP19/CDC5L and RES complexes, the IBC, and other proteins, including those operationally designated PRP19/CDC5L-related (nineteen-related, NTR) are stably integrated. Unlike PRP19/CDC5L and IBC proteins, PRP19-related proteins are not recruited as part of a large preformed complex. How this dramatic exchange of proteins takes place is presently unclear, but likely proceeds through several intermediates, in which the loss of certain proteins provides novel binding sites required to recruit others.

2   STRUCTURE OF HUMAN SPLICEOSOMAL snRNPs

Whereas the molecular architecture of the human U1 snRNP was revealed by X-ray crystallography (Pomeranz Krummel et al. 2009; Weber et al. 2010), a high-resolution structure of the isolated 17S U2 snRNP is currently lacking. Cryo-EM was successfully used to elucidate the molecular organization of the human (h) and yeast (y) U4/U6.U5 tri-snRNP (Nguyen et al. 2015, 2016; Agafonov et al. 2016; Wan et al. 2016a), the largest spliceosomal subunit. The U4/U6 snRNP is located in the upper part of the (h)tri-snRNP (Fig. 2A). Stems I and II of the U4/U6 three-way junction are stacked co-axially and stabilized by U4/U6 proteins that also extensively interact with each other. Con-
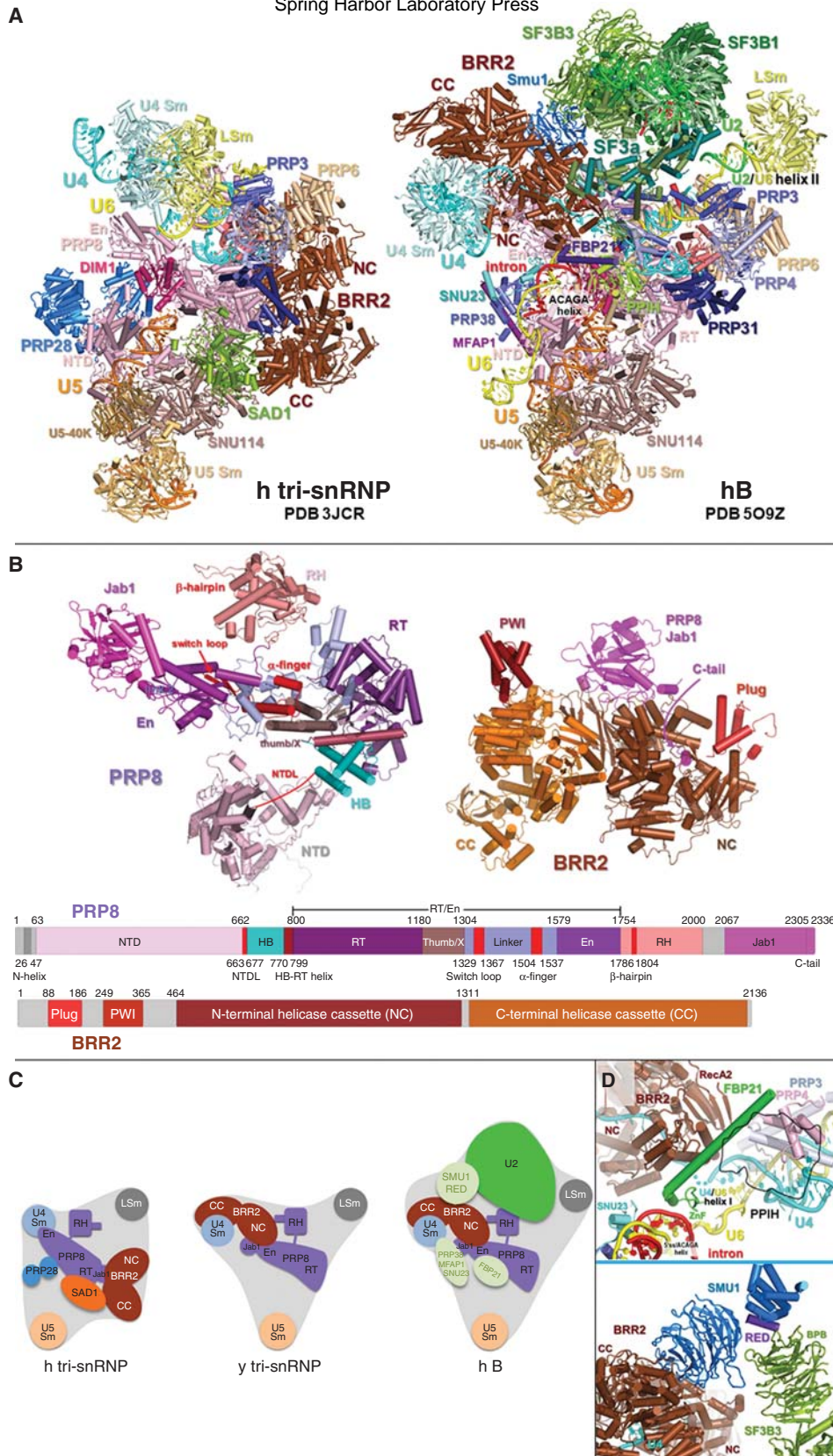
Figure 2. Structure of the human tri-snRNP and spliceosomal B complex. (*A*) Overview of the structure of the human U4/U6.U5 tri-snRNP and B complex. (*B*) Schematic of the domain organization of human PRP8 and BRR2. PRP8 contains an amino-terminal domain (NTD), helical bundle (HB), reverse transcriptase-like (RT) domain, and thumb/X domain, followed by a Linker region and an endonuclease-like (En) domain, plus carboxy (C)-terminal RNase-H-like (RH) and Jab1/MPN domains. (*Legend continues on following page.*)

sistent with previous studies (Schaffert et al. 2004), PRP6 plays a key bridging function in the tri-snRNP, contacting PRP4 and other U4/U6 proteins, and U5 PRP8.

U5 PRP8 and SNU114 are located in the central region of the tri-snRNP, whereas the U5 Sm core is located at the bottom (Fig. 2A). Prp8 is a large, highly conserved scaffolding protein that contains several functionally important domains (Fig. 2B). Cross-linking studies showed that Prp8 interacts with all of the chemically reactive groups of the pre-mRNA (5′SS, BS, and 3′SS) (Grainger and Beggs 2005), and these interactions were subsequently confirmed by cryo-EM of various spliceosomal complexes (Scheres and Nagai 2017; Shi 2017). In the isolated tri-snRNP (and after its integration into the spliceosome), the Prp8 amino-terminal domain (Prp8^{NTD}) tightly grasps U5 stem I, holding it in a rigid conformation and at a later stage positioning U5 loop 1 close to the active site. PRP8 also interacts extensively with the GTP-binding protein SNU114 mainly via its amino-terminal domain (NTD). SNU114 shares structural homology with the ribosome elongation factor EF-2/EF-G. In all currently available high-resolution cryo-EM structures of the spliceosome, Snu114 shows the same compact structure and appears to bind GTP, suggesting that it does not hydrolyse GTP to facilitate a conformational change in the spliceosome (for detailed discussion, see Plaschka et al. 2019).

The U5 helicase BRR2 contains two helicase cassettes (Fig. 2B), but only the amino-terminal cassette (NC) actively unwinds U4/U6 (Santos et al. 2012). BBR2^{NC} is tightly bound by the PRP8 carboxy-terminal Jab1 domain (PRP8^{Jab1}) throughout the splicing cycle, and as a consequence BRR2 is the only spliceosomal helicase that does not dissociate from the spliceosome after its action. Several mechanisms exist to prevent premature BRR2 activity including (1) auto-inhibition by its large amino-terminal region, which contains a helical PLUG and PWI-like domain (Fig. 2B) that folds back onto BRR2's helicase cassettes, preventing ATP hydrolysis and RNA binding (Absmeier et al. 2015) and (2) insertion of the carboxy-terminal tail of PRP8^{Jab1} into the RNA binding channel of BRR2^{NC}, inhibiting accommodation of U4 snRNA (Mozaffari-Jovin et al. 2013). In (h)tri-snRNPs, not only is BRR2 repressed by its amino-terminal region, but its NC is located about 10 nm from its U4/U6 substrate (Agafonov et al. 2016).

Surprisingly, although the region comprising the PRP8^{NTD}, SNU114, and U5 Sm core is structurally very similar, the remaining parts of yeast and human tri-snRNPs differ in their conformation or position to varying degrees (Fig. 2C). Most strikingly, in (y)tri-snRNPs, BRR2 is located ~20 nm away from its position in (h)tri-snRNPs (Fig. 2). Moreover, as yeast Brr2 is loaded onto U4 snRNA and is not auto-inhibited (Nguyen et al. 2016; Wan et al. 2016a), it appears to be ready to unwind the U4/U6 duplex. Indeed, unlike purified (h)tri-snRNPs, purified (y)tri-snRNPs dissociate in the presence of ATP in a BRR2-dependent manner (Nguyen et al. 2015; Agafonov et al. 2016). The different molecular organization of BRR2 could (1) reflect differences in the early stages of the spliceosome activation process in yeast versus humans and/or (2) be the result of the absence of Sad1 in (y)tri-snRNPs, which plays a key role in stabilizing the tri-snRNP in yeast extract (Huang et al. 2014). In (h)tri-snRNPs, SAD1 binds SNU114, BRR2's PWI domain and U4/U6-PRP31, thereby stabilizing the interaction between U5 and U4/U6, and also tethering BRR2 at a distance from U4/U6 (Agafonov et al. 2016).

3 STRUCTURE OF THE PRECATALYTIC B COMPLEX

Cryo-EM of human (Bertram et al. 2017a) and yeast (Plaschka et al. 2017) spliceosomal B complexes showed that their globular head domain is comprised of the U2 snRNP, and their triangular body contains mainly the U4/U6.U5 tri-snRNP (Fig. 2A). In human B (hB), the 5′ domain of U2 containing the heptameric SF3b protein complex is spatially separated from its 3′ domain comprised of the Sm RNP core and U2 A′ and B′ proteins, with proteins of the SF3a complex bridging both domains. SF3b sequesters the U2/BS helix in a similar manner to that in the B^{act} complex (see below). U2 snRNP is connected to the tri-snRNP by the U2/U6 helix II, which is stabilized by PRP3, and by bridges between BRR2 and SF3B3 involving SMU1 (Fig. 2A,D), and between the U2 SF3b domain and PRP8^{NTD}, which likely contains the intron and SF3A1.

In hB, intron nucleotides +4 to +6 base-pair with the U6 ACAGA-box and, accordingly, U1 snRNP and PRP28 are absent. This U6/5′SS interaction triggers a major structural change that converts human pre-B into a stable B complex (Boesler et al. 2016). U6 nucleotides directly upstream of the

Figure 2. (*Continued*) The RT, En, RH, and Jab1/MPN domains are deconstructed enzymes that do not show enzymatic activities. Within the Linker, several distinct structural elements play important roles within the spliceosome including the switch-loop and the α-finger. The 3D structure of PRP8 shown is from the hB complex, and that of BRR2 (except for its PWI domain) is from hB^{act}. (*C*) Cartoon of the organization of selected proteins in the human and yeast tri-snRNPs, and human B complex. (*D*) *Upper* panel: FBP21's zinc finger contacts the U6/intron helix, whereas its α-helix keeps U4/U6 stem I, via PPIH (shown as a black outline), at a distance from Brr2^{NC}. *Lower* panel: SMU1 and RED bridge BRR2 and SF3B3.

Advanced Online Article. Cite this article as *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a032417

ACAGA-box also form base pairs (including noncanonical ones) with additional intron nucleotides downstream from the 5′SS. This extended U6/intron helix, which is also found in human B^{act}, C, and C* complexes, appears to be a structural feature of spliceosomes solely in higher eukaryotes and may help to stabilize the short U6/5′SS helix. In hB, the extended U6/intron helix and, unlike in yeast also the 5′ region of U6, are accommodated by $PRP8^{NTD}$ (Fig. 2A).

Most human tri-snRNP proteins undergo substantial structural rearrangements on tri-snRNP integration into the spliceosome (Fig. 2C) (Bertram et al. 2017a). The most dramatic is observed for BRR2 and its binding partner $PRP8^{Jab1}$, which are located in hB near the PRP8 En domain ($Prp8^{En}$) ~20 nm from their position in the isolated tri-snRNP. Moreover, $BRR2^{NC}$ is loaded onto U4 snRNA, and is poised to unwind the U4/U6 duplex. SAD1, which tethers BRR2 in its pre-activation position, is largely displaced from the B complex, which may play a role in triggering tri-snRNP rearrangements. Finally, PRP8 adopts a partially closed conformation, in which several loops emanating from Prp8's Linker are juxtaposed with its NTD and the U5 DIM1 protein. Together, they form a pocket that binds the 5′SS GU nucleotides in an extended conformation, revealing an unexpected role for DIM1 in 5′SS recognition (Bertram et al. 2017a).

The structural organization of U5 and U4/U6 proteins, and PRP38 and SNU23, is highly similar in human and yeast B complexes, and, surprisingly, also in the purified yeast tri-snRNP (Fig. 2C) (Nguyen et al. 2016; Wan et al. 2016a). This suggests that, in contrast to the situation in humans, the yeast tri-snRNP is already rearranged before docking with the A complex. Although BRR2 binds its U4/U6 substrate and is no longer auto-inhibited in both B complexes, U4 and U6 snRNA dissociate solely from purified yeast B complexes in the presence of ATP (Bao et al. 2017), suggesting that BRR2 is still negatively regulated in the purified human, but not yeast B complex. This implies that tri-snRNP rearrangements leading to BRR2 substrate loading and the triggering of BRR2 activity are both more complex in higher eukaryotes.

The hB cyro-EM structure revealed the likely functions of several B-specific proteins, which include those lacking a yeast counterpart, such as SMU1, RED, and FBP21, and conserved proteins like PRP38, SNU23, and MFAP1. Because of their mutually exclusive binding sites, incorporation of the latter proteins in hB requires prior displacement of PRP28. PRP38 and SNU23 contact the U6/5′SS helix, and may facilitate its repositioning during activation, enabling the juxtapositioning of the 5′SS and U2/U6 catalytic center. The heterodimer SMU1/RED bridges U2 and the tri-snRNP (Fig. 2D), and via its interaction with $BRR2^{NC}$, it may potentially modulate BRR2 activity. Intriguingly, an α-

helical element of FBP21, which was implicated in regulating BRR2 activity (Henning et al. 2017), is positioned between $BRR2^{NC}$ and U4/U6 proteins that are bound to U4/U6 stem I (Fig. 2D). This architecture suggests that FBP21 may act to keep U4/U6 stem I at a distance from BRR2's NC, and that its release from the spliceosome, which occurs early during activation, may play a part in triggering BRR2 activity. Thus, the initial steps toward the formation of an activated spliceosome appear to differ between yeast and humans.

4 STRUCTURE OF THE ACTIVATED (B^{act}) SPLICEOSOME

The cryo-EM structure of the B^{act} complex (Fig. 3A) (Rauhut et al. 2016; Yan et al. 2016; Haselbach et al. 2018; Zhang et al. 2018) provides important insights into the mechanism of spliceosome activation, and elucidates the function of proteins integrated at this stage. As first indicated by biochemical studies, the catalytic U2/U6 RNA network is already in place in B^{act}—that is, the U6 ISL, U2/U6 helices Ia and Ib and catalytic RNA triple helix have formed. However, in contrast to catalytically active C or C* complexes, where catalytic metal ions M1 and M2 are positioned ~4.2 Å apart (Fig. 1E), in B^{act} density for M1 is either missing or displaced both in humans and yeast. Thus, B^{act} lacks a fully functional catalytic center.

The rearranged catalytic U2/U6 RNA core is stabilized to a large extent by interactions with PRP8 (Fig. 3B). During the B to B^{act} transition, PRP8 undergoes a further conformational change that leads to a closed conformation of its cavity that accommodates the U2/U6 RNA network. The U6 ILS fits snuggly into a basic pocket of $PRP8^{NTD}$, close to U5 loop 1, and is stabilized by $PRP8^{NTDL}$. The nearby PRP8 HB domain accommodates the U2 strand of U2/U6 helix Ib. The α-finger of the PRP8 Linker (also called the 1585-loop in yeast) further stabilizes the catalytic center by interacting with U2 nucleotides between U2/U6 helix I and the U2/BS helix. This leaves a gap between the U2/U6 catalytic center and the PRP8 Linker region in which part of RNF113A (Cwc24 in yeast) and SF3A2 are accommodated. During catalytic activation, RNF113A and SF3A2 are replaced in this gap by the U2/BS helix (Fig. 3B), thereby juxtapositioning the first-step reactants. $PRP8^{NTD}$ plays an important part in orienting the scissile bond of the 5′SS toward the catalytic center. However, the bases of G+1 and U+2 of the 5′SS are contacted by the carboxy-terminal zinc finger of RNF113A and amino terminus of SF3A2, thus protecting the 5′SS from premature nucleophilic attack (Fig. 3C).

In B^{act}, the catalytic RNA network is also contacted by PRP19/CDC5L complex and PRP19-related proteins (i.e., CDC5L, CWC15, PLRG1, SKIP, and SYF3). These proteins
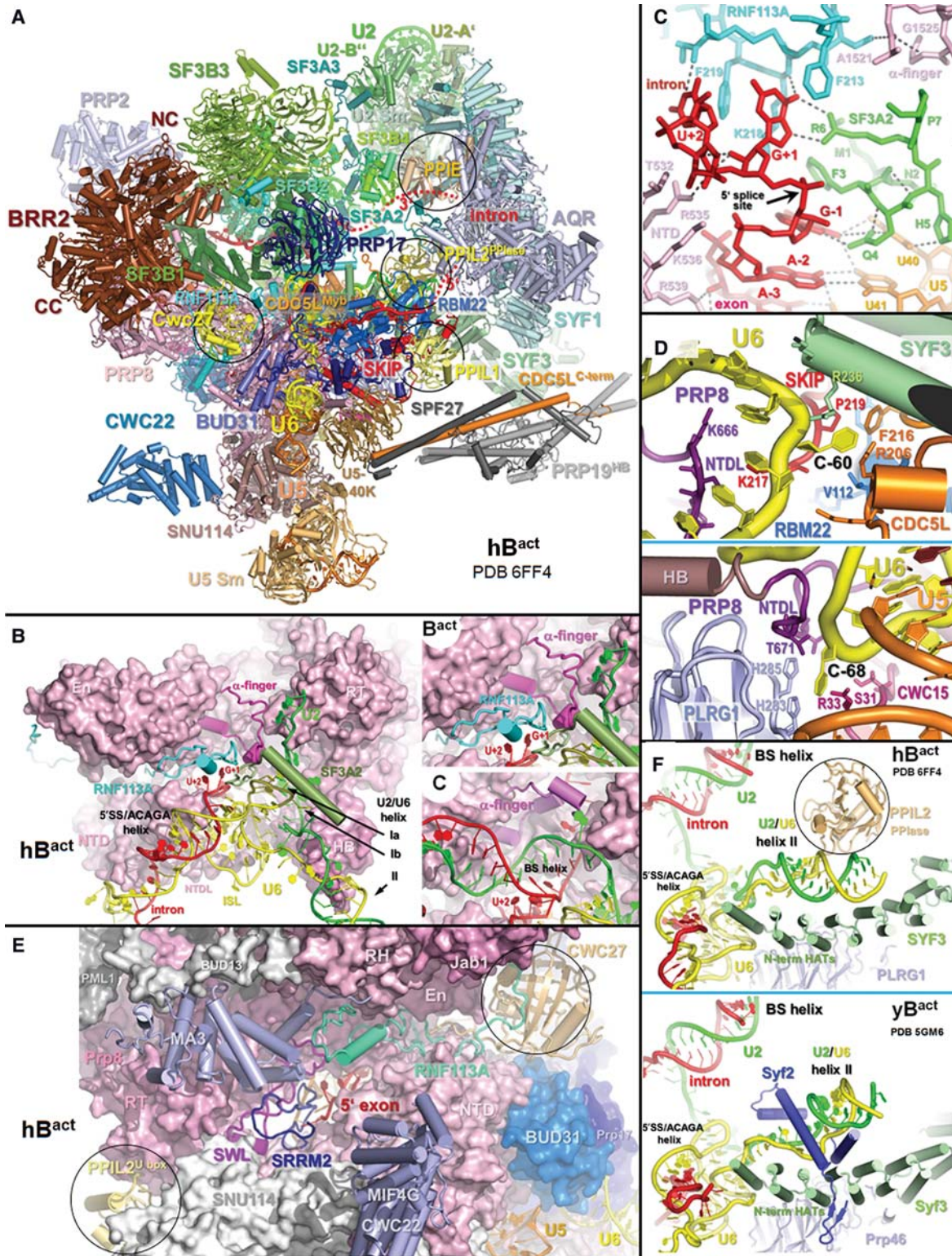
Figure 3. Molecular organization of RNA−protein interactions involving the catalytic RNA network, 5′SS, and 5′ exon in the human B^{act} complex. (*A*) Overview of the structural organization of the human B^{act} complex. The positions of PPIases are circled in black. (*B*) Accommodation of the catalytic RNA network that is first formed in B^{act} by the PRP8 amino-terminal domain (NTD) and helical bundle (HB) domains and α-finger of its Linker domain. (*Legend continues on following page.*)

make only limited direct contacts, interacting predominantly with flipped-out nucleotides of U6 snRNA such as U68 and C60 (Fig. 3D), consistent with studies showing that the yeast NTC plays a role in stabilizing the association of U6 with the spliceosome (Chan et al. 2003). The resulting protein–RNA network serves to both stabilize and fine-tune the conformation of the catalytic RNA network, and may transmit structural changes in the catalytic center to the periphery of the spliceosome and vice versa. As CDC5L, SKIP, CWC15, and PLRG1 interact with PRP8, they likely also indirectly help to maintain the RNA in a catalytically active conformation by stabilizing the closed conformation of PRP8. The extensive protein–RNA network that is established in the catalytic RNP core, which is conserved between yeast and humans (Rauhut et al. 2016; Yan et al. 2016), underscores the essential role of proteins in establishing a catalytically active RNA structure in the spliceosome. In fact, the spliceosome may be considered an RNP enzyme, in which the actual catalytic steps are facilitated by a catalytic metal–RNA center.

When PRP8 adopts its closed conformation in B^{act}, a narrow 5′ exon binding channel (EBC) is created between PRP8's RT/Linker and NTD (Fig. 3E). Exon nucleotides −5 to −7, upstream of where U5 loop 1 interacts with the 5′ exon, are contacted by loops of PRP8^{Linker}. The 5′ exon is then contacted by the amino-terminal region of SRRM2 (Cwc21 in yeast) and the Prp8 switch loop (PRP8^{SWL}), which has moved ∼7 nm from the PRP8^{En} in B, to PRP8^{RT} in B^{act}. The EBC is straddled by CWC22 whose MIF4G and MA3 domains bind to SNU114 and PRP8^{RT}, respectively (Fig. 3E). In B^{act}, PRP8^{En} is in an upright position relative to PRP8^{RT}, and does not interact with PRP8^{NTD}. Both domains are kept at a distance by the presence of amino-terminal regions of both RNF113A and CWC27 in a cleft between PRP8^{En} and PRP8^{NTD} (Fig. 3B,E). RNF113A tightly interacts with the PPIase domain of CWC27—the only human spliceosomal PPIase with a yeast counterpart. Release of these proteins during catalytic activation likely triggers the conformational change in the PRP8 RT/En domain

relative to PRP8^{NTD} that is observed during the B^{act} to C complex transition (see also Fig. 7B). This leads to tightening of PRP8's grasp on the 3′ end of the 5′ exon, thereby ensuring that it is stably bound after 5′SS cleavage.

In hB^{act}, the extended U6/intron helix binds to a positively charged edge of PRP8^{NTD} and is further stabilized by contact with RBM22. Intron nucleotides downstream from this helix are enclosed in a positively charged channel formed by RBM22 and further downstream they are likely handed over to the RNA helicase AQR (Fig. 3A and see below). Thus, RBM22 plays an important part in chaperoning the 5′ region of the intron. In yB^{act}, similar tasks are performed by Ecm2 and Cwc2, which contain one or more domains homologous to those found in RBM22, indicating that the latter is a fused version of Ecm2 and Cwc2 (Rasche et al. 2012).

The conformation of the U2/U6 RNA and RNP network is largely similar in yeast and human B^{act}, and the structure of the spliceosome's core domain does not change substantially in the subsequently formed catalytically active spliceosomal complexes. However, U2/U6 helix II adopts a unique conformation in hB^{act} (Fig. 3F). In yeast B^{act} and in subsequently formed human and yeast complexes, U2/U6 helix II is bound by Syf2. In contrast, in hB^{act}, helix II is clamped between SYF3 and the PPIase domain of PPIL2, preventing SYF2 from binding. This explains, at least in part, why in humans SYF2 binds at a later stage and is not preorganized into a large complex like its yeast counterpart. SYF3 interacts with SYF1, which also shares a large interface with the RNA helicase AQR that is absent in yeast. Rearrangement of U2/U6 helix II and the incorporation of hSYF2 during the B^{act} to C transition in higher eukaryotes may thus involve remodeling by AQR, whose ATPase activity is required for efficient catalytic activation of human spliceosomes (De et al. 2015).

The bipartite U2 snRNP is located at the top of the hB^{act} structure (Fig. 3A). As in hB, the 5′ and 3′ domains of U2 are bridged by SF3a proteins that contact the U2 Sm ring and SF3b proteins in the 5′ domain. SF3B3, which contains

Figure 3. (*Continued*) Part of RNF113A (Cwc24) and SF3A2 (Prp11) are inserted into the gap between the RT/En domain and the catalytic RNA network and these proteins are replaced by the U2/BS helix in the C complex (see insets). (*C*) Close-up of interactions between the 5′SS nucleotides G+1 and U+2 and residues of RNF113A, and between SF3A2 and G+1 and G−1. (*D*) PRP19/CDC5L complex proteins and related proteins directly contact primarily flipped-out U6 nucleotides in the catalytic RNA core. For example, the base of U6-C60 in the lower stem of the ISL lies in a pocket formed by amino acids of SYF3, CDC5L, RBM22, and SKIP (*upper* panel) and flipped-out U6-U68 of the U6 ISL loop is bound cooperatively by amino acids of PLRG1 and CWC15 (*lower* panel). (*E*) Spatial organization of the 5′ exon binding channel in hB^{act}. Black circles indicate the position of PPIL2's U box domains and CWC27's PPIase domain. (*F*) U2/U6 helix II adopts a unique position in human (*upper* panel) compared with yeast (*lower* panel) B^{act}. In yB^{act}, helix II lies in a more upright position and interacts with Syf2, whereas in hB^{act} it lies more perpendicular (relative to the U6 ACAGA/5′SS helix) and contacts the PPIase domain of the PPIL2 protein, which is not present in the *S. cerevisiae* spliceosome. Black circle indicates the position of PPIL2's PPIase domain.
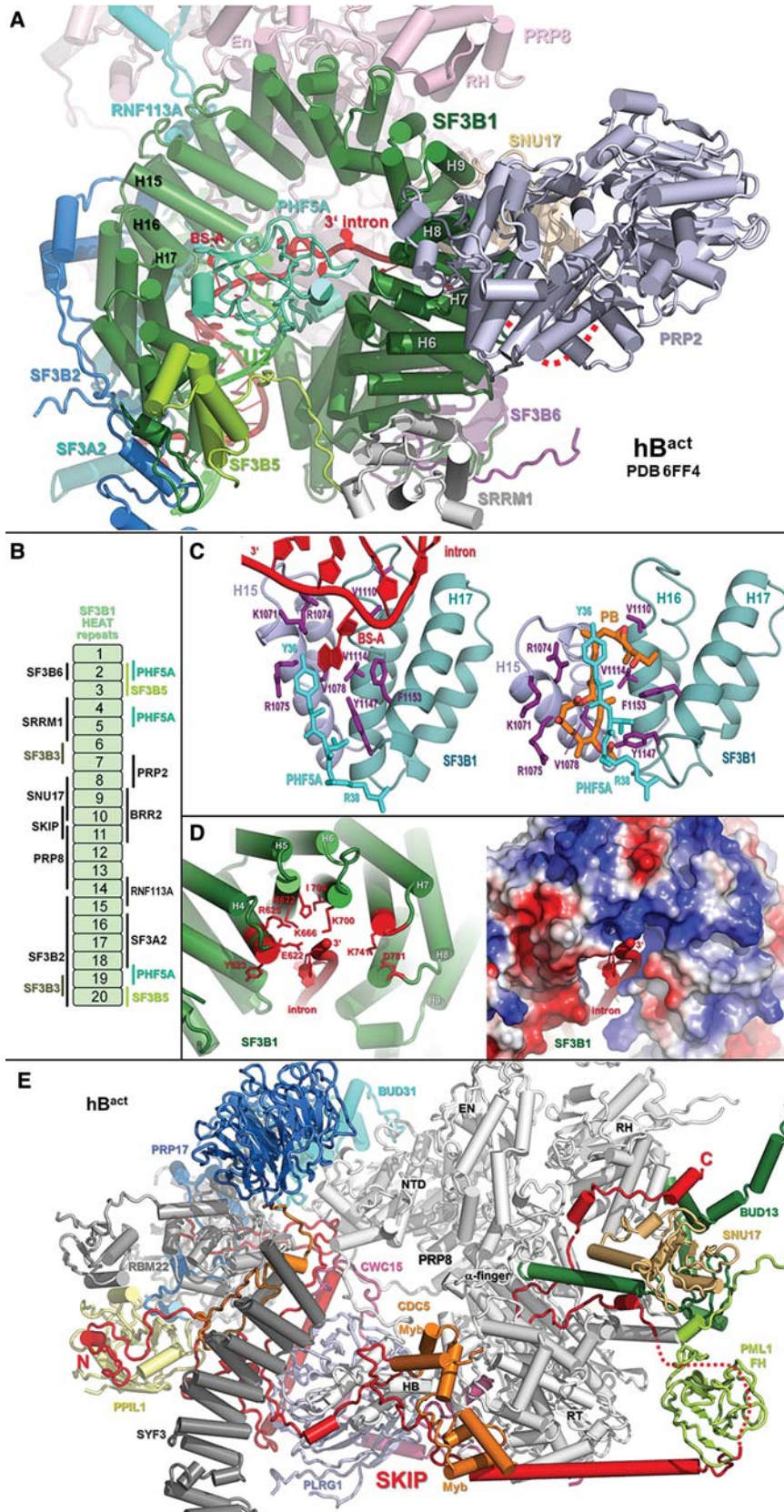
Figure 4. (*See legend on following page.*)

three β propeller domains (BPD) A-C, SF3B5, and PHF5A form an intricate protein network with the carboxy-terminal HEAT domain of SF3B1 (SF3B1^{HEAT}), similar to that observed in the crystal structure of the human SF3b core complex (Cretu et al. 2016). The SF3B1^{HEAT} domain forms a superhelical structure that clamps the U2/BS helix between its terminal HEAT repeats (Fig. 4A), consistent with cross-linking studies showing SF3B1 contacts the pre-mRNA on both sides of the BS (Gozani et al. 1998). The base of the BS-A nucleotide resides in a pocket comprised of SF3B1 HEAT repeats H15–H17 and a Tyr side-chain of PHF5A (Fig. 4C). Thus, the first catalytic step nucleophile is sequestered by protein and also located ∼5 nm from the 5′SS, preventing premature 5′SS cleavage.

X-ray crystallography (Cretu et al. 2018) and cryo-EM (Finci et al. 2018) of the recombinant SF3b core bound by the splicing inhibitor/modulator Pladienolide B (PB) or its derivative E7107, respectively, recently revealed that these antitumor compounds bind to SF3B1 and PHF5A residues that ultimately form the BS-A binding pocket, and that, when mutated, lead to compound resistance (Fig. 4C). Thus, inhibitor binding and accommodation of the BS-A appear to be mutually exclusive, suggesting that these SF3B inhibitors compete with the BS for SF3b binding. SF3B1^{HEAT} shows a more open conformation in the apo SF3b core structure (Cretu et al. 2016), compared with its conformation in the B and B^{act} complex (Rauhut et al. 2016; Yan et al. 2016; Bertram et al. 2017a; Plaschka et al. 2017; Haselbach et al. 2018; Zhang et al. 2018). By binding to HEAT repeats H15–H17, PB appears to act as a wedge, "locking" SF3B1^{HEAT} in the open form (Fig. 4C) (Cretu et al. 2018). In the absence of PB, interaction of the U2/BS helix with the open conformation likely induces the closed SF3B1 conformation, with H15–H17 acting as a conformational switch.

U2 and intron nucleotides upstream of the BS form an RNA helical structure comprised of many noncanonical base pairs. This extended U2/BS helix contains part of the U2 anchoring site and is contacted by SF3A2 and SF3A3, consistent with cross-linking studies (Gozani et al. 1996; Schneider et al. 2015). As SF3A2's amino terminus also contacts the 5′SS GU nucleotides, there is a direct link between the extended U2/BS helix and the 5′SS. The structure of SF3b and its interactions with the extended U2/BS helix are very similar in human and yeast B^{act}. However, in higher eukaryotes SF3b contains an additional subunit, the RRM protein SF3B6 (p14). In hB^{act}, SF3B6 is located close to HEAT repeats 1–2 on the convex (outer) side of SF3B1^{HEAT} and is spatially separated by ∼5 nm from the BS-A (Fig. 4A) (Haselbach et al. 2018; Zhang et al. 2018). As SF3B6 is cross-linked to the BS-A in early spliceosomal complexes (Will et al. 2001), it appears to be substantially repositioned before the formation of B^{act}.

SF3B1^{HEAT} not only sequesters the U2/BS helix, but also acts as a scaffold, interacting not only with SF3a/b proteins, but also several other spliceosomal proteins (Fig. 4A,B). SNU17's RES complex binding partner BUD13 interacts with PRP8^{RT} and thus the RES proteins bridge the major structural modules of the U2 (the SF3B complex) and U5 snRNP (PRP8) (Fig. 4E), consistent with recent studies indicating that they play an important structural role during catalytic activation (Bao et al. 2017).

5 STRUCTURAL BASIS OF PRP2-MEDIATED CATALYTIC ACTIVATION

PRP2 binds HEAT repeats H7–H8, ∼6 nm from its actual target, the U2/BS helix (Fig. 4A). PRP2 requires ∼25–30 intron nucleotides downstream from the BS to facilitate catalytic activation (Liu and Cheng 2012). Human B^{act} complexes analyzed by cryo-EM (Haselbach et al. 2018; Zhang et al. 2018) were formed on truncated pre-mRNAs with 10 or 19 nucleotides downstream from the BS. These downstream intron nucleotides contact PHF5A and exit the SF3B1 HEAT ring at repeats H4–H6, before they are bound

Figure 4. Sequestration of the U2/BS helix by SF3B1 and location of SF3B1 hotspot mutations and the Pladienolide B binding site. (A) The superhelical SF3B1 HEAT domain sequesters the U2/BS helix and acts as a scaffold that binds numerous proteins. Intron nucleotides downstream from the branch site (BS) exit the HEAT domain at repeats H4–H6 and are subsequently bound by the retention and splicing (RES) protein SNU17, followed by the helicase PRP2. SF3B3, which interacts with HEAT repeats H5 and H6 and H19 and H20, has been omitted. (B) Schematic of the SF3B1 HEAT repeats and their interaction partners. Proteins contacting both amino and carboxy-terminal HEAT repeats are colored as in A. (C) Overlapping binding site of Pladienolide B (PB) and the branch site adenosine (BS-A). (Left) Residues of SF3B1 HEAT repeats H15–H17 and PHF5A that form the BS-A binding pocket in the human B^{act} complex, in which the HEAT domain shows a closed conformation. (Right) The H15–H17 hinge region of SF3B1 bound by PB (orange sticks), which locks the SF3B1 HEAT domain in an open conformation. (D) SF3B1 residues frequently mutated in various human cancers (red lettering) cluster near the 3′ end of the intron where it exits the HEAT superhelix (left). (Right) Space filling model with surface potential (red, negatively charged amino acids; blue, positively charged) of the SF3B1 region shown in the left panel. Many hotspot mutations are nonconserved changes that alter the amino acid's charge. (E) Meandering path of the SKIP protein (red) in hB^{act} and its numerous interaction partners.

Cold Spring Harbor Perspectives in Biology

by SNU17, but do not extend to the adjacent PRP2 binding site (as observed in yB^{act}). Thus, PRP2's RecA domains cannot access the intron, providing a structural basis for why PRP2 cannot catalytically activate B^{act} complexes formed on these truncated pre-mRNAs. During B* formation, the U2/BS helix must be freed from SF3B1^{HEAT}, and also from other SF3b and SF3a proteins, to juxtapose the BS-A and the scissile bond of the 5′SS. At the same time, RNF113A and the amino terminus of SF3A2 must be repositioned to free the 5′SS and allow docking of the U2/BS helix into the catalytic center. The scaffolding function of SF3B1^{HEAT} likely plays an important role in the coordinated release of multiple proteins during PRP2-mediated formation of B*. That is, PRP2 has been proposed to induce, in an intron-binding and ATP-dependent manner, a conformational change in the HEAT domain curvature (Rauhut et al. 2016) that would disrupt its interaction with the U2/BS helix, setting free the BS-A, and destabilize the interaction of proteins which bind SF3B1, including SF3A2 and RNF113A, as well as its binding partner CWC27. Indeed, previous biochemical studies showed that PRP2 action leads to the destabilization/release of these proteins (Warkocki et al. 2009; Lardelli et al. 2010; Ohrt et al. 2012). Thus, PRP2 facilitates remodeling of the U2/BS from a distance.

6 STRUCTURAL CONSEQUENCES OF SF3B1 CANCER-RELATED MUTATIONS

SF3B1 mutations that occur very frequently in patients with myelodysplastic syndromes and various types of cancer (i.e., "hot spot" mutations) are located predominantly within the intrarepeat loops of HEAT repeats H4–H7 (Fig. 4) (reviewed in Jenkins and Kielkopf 2017). In the B and B^{act} complexes, these cancer-related mutations are located close to where the 3′ region of the intron exits the HEAT superhelix (Fig. 4A,D). This region of SF3B1 forms a basic channel (Fig. 4D), and thus hot spot mutations such as K666E or K700E, will change the surface potential and thus potentially alter intron binding. In vivo, SF3B1 hot spot mutations cause alternative BS selection, leading to cryptic 3′SS usage and the formation of aberrantly spliced mRNAs (Darman et al. 2015; Alsafadi et al. 2016). The mechanism whereby these mutations affect AS is unclear, but could also be caused by changes (1) in the conformation of SF3B1^{HEAT} that destabilize/alter its interaction with the U2/BS helix and/or the 3′ end of the intron (Jenkins and Kielkopf 2017) or (2) that alter its interaction with one or more protein, as recently shown for the RNA helicase Prp5 (Tang et al. 2016). As BS selection occurs predominantly during A complex formation, more detailed insights into the structural consequences of SF3B1 cancer-related muta-

tions await the determination of the cryo-EM structure of the human A and earlier complexes.

7 INSIGHTS INTO THE FUNCTION OF PPIases AND IBC PROTEINS DURING SPLICING

Numerous human-specific spliceosomal proteins, including several PPIases, are located in peripheral domains of hB^{act} and their molecular organization provides insights into their functions (Fig. 3A). One function of the spliceosomal PPIases appears to be to bridge various protein modules. For example, PPIL1 helps to anchor the PRP19 helical bundle (PRP19^{HB}) to the B^{act} main body (Fig. 3A and below). The RRM domain of PPIE (PPIE^{RRM}) binds SYF1, and its PPIase domain interacts with SF3B4^{RRM1} and the β-sandwich of SF3A2. PPIL2^{PPIase} also interacts with SF3A2, while its amino-terminal U box domains bind to SNU114 and a nearby region of PRP8^{RT} (Fig. 3A,E). The PPIase domains of PPIL2 and PPIE also interact with RNA, contacting U2/U6 helix II or the intron upstream of the extended U2/BS helix, respectively (Fig. 3A,F). Whether these PPIases act as molecular switches during splicing, isomerizing proline-peptide bonds, and thereby catalyzing protein/RNP rearrangements within the spliceosome, is not clear (Thapar 2015).

The helicase AQR shares a large interface with SYF1's amino-terminal HAT (half a TPR) repeats 2–8, which extend close to the U2 3′ domain (Figs. 3A, 5A). SYF1 and SYF3 form long, curved α-helical solenoids that cross one another close to U2/U6 helix II (Figs. 3A and 5B). Via SYF1 and its interaction partners, AQR is thus indirectly connected to the catalytic U2/U6 RNP core and U2/U6 helix II. Within spliceosomes, AQR cross-links to intron nucleotides 30–40 upstream of the BS-A (Hirose et al. 2006). Consistent with this, intron nucleotides upstream of the extended U2/BS helix are contacted by SF3B4^{RRM1} and PPIE, before apparently being transferred to AQR in hB^{act} and hC (Figs. 3A and 5B) (Haselbach et al. 2018; Zhan et al. 2018; Zhang et al. 2018). AQR's β-barrel also directly contacts RBM22^{RRM}, which binds intron nucleotides downstream from the U6/intron helix. Thus, AQR connects and potentially fixes, directly or indirectly, intron nucleotides that lie upstream of the extended U2/BS helix and downstream from the extended U6/intron helix (Fig. 3A). This would allow the intron located between these regions, which potentially contains intronic small nucleolar RNAs (snoRNAs) that are processed during spliceosome assembly (Hirose et al. 2006), to loop out, and thereby prevent it from interfering with the spliceosome's core machinery.

Cryo-EM has revealed many examples of proteins that are intrinsically unfolded in the isolated state, but adopt ordered structures in the spliceosome, apparently in a coor-
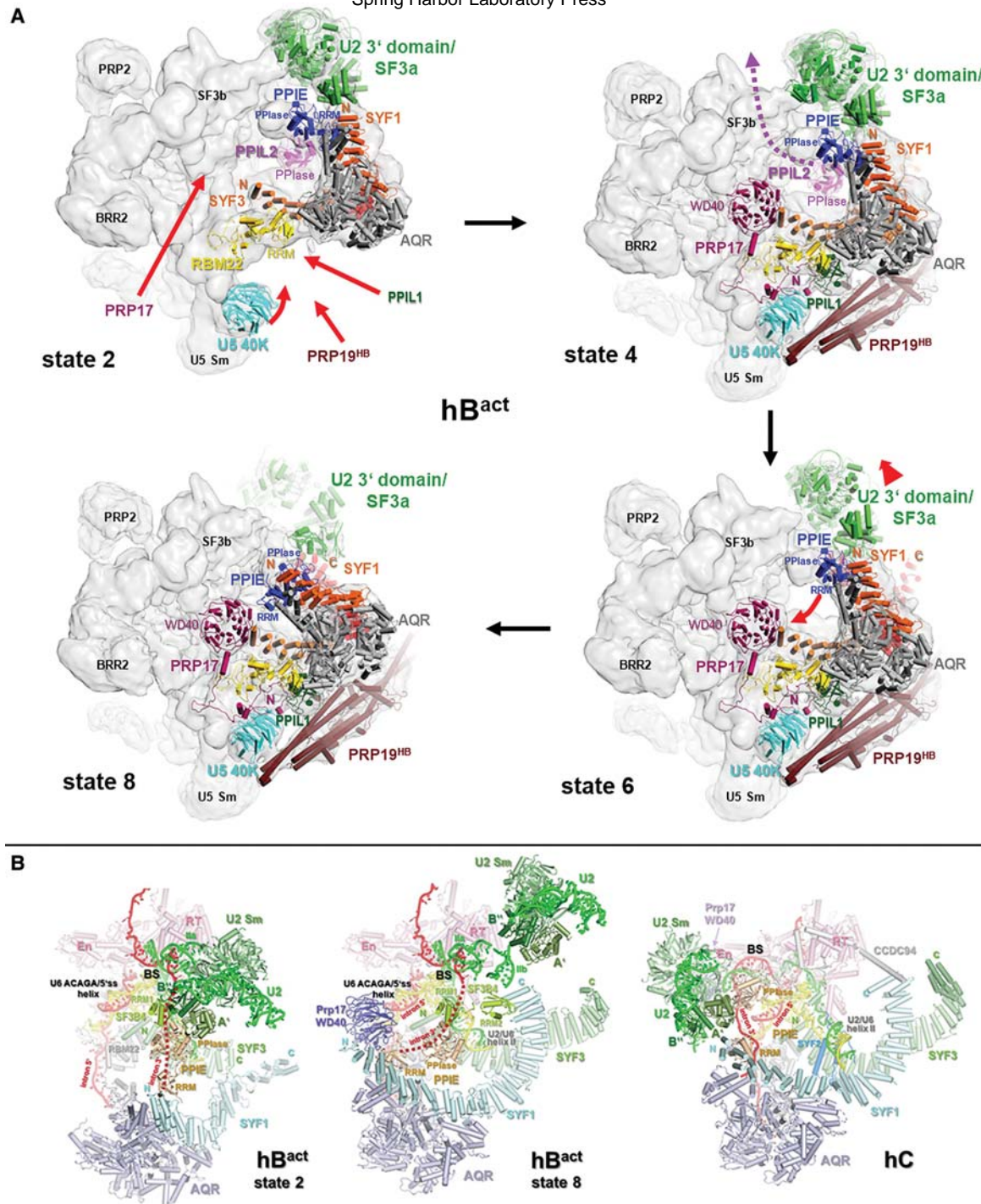
Figure 5. Conformational dynamics of the human B^{act} complex. (*A*) Eight major conformational states of the human B^{act} complex were obtained by computational sorting of images. Only states 2, 4, 6, and 8 are shown. The positions of the most mobile proteins/protein domains are indicated in the electron microscopic (EM) density of B^{act}. State 4 corresponds to the high-resolution hB^{act} structure from Haselbach et al. (2018) that is shown in Figures 3 and 4. Density for the PRP19^{HB}, PPIL1, and most of PRP17 cannot be recognized in states 1 and 2, even though there is evidence that these proteins, as well as nearly all other hB^{act} proteins, are already present in the B^{act} complex represented by stage 1. This means that they bind in a highly flexible manner and are simply not discernable by EM, as opposed to being first recruited to hB^{act} at a subsequent step. In state 4, U5-40K^{WD40} has shifted upward by 2 nm, generating a docking site for one end of the PRP19^{HB}, stabilizing its conformation and leading to its appearance in state 4. Similarly, density for PRP17 and PPIL1 is also first detected in state 4, in which PPIL1 now interacts with PRP19^{HB}. In state 6, PPIL2 is no longer detected, and in state 8, PPIE moves toward PRP17. The U2 3′ domain also moves by several nm from states 2 to 8. (*B*) Comparison of the spatial organization of the U2 3′ domain, SYF1, SYF3, AQR, PPIE, and the PRP17 WD40 domain in states 2 and 8 of hB^{act} and in the hC complex.

dinated manner with other spliceosomal proteins. Such proteins, which include RNF113A and SKIP, often form complex protein–protein networks. Indeed, SKIP meanders throughout B^{act}, contacting at least 10 other proteins, and is therefore thought to play an important part in the conformational maturation and stabilization of B^{act} (Fig. 4E). Proteins like SKIP are also well-suited for transferring information from the periphery of the spliceosome to the catalytic RNP core and thus, potentially, for influencing the kinetics of the splicing reaction. A particularly interesting case of long-range communication is observed with proteins of the PRP19/CDC5L complex. Whereas, the two amino-terminal Myb domains of CDC5L bind $PRP8^{RT}$ and help to stabilize the catalytic U2/U6 RNA structure, CDC5L's carboxy-terminal region forms, together with SPF27 and the coiled-coil regions of four copies of PRP19, an elongated helical bundle ($PRP19^{HB}$) (Figs. 3A and 5A). In B^{act}, the latter is docked to its main body, via PPIL1, SYF1 and the U5-40K WD40 domain ($U5\text{-}40K^{WD40}$), which communicate via PRP17 and SKIP with other regions of B^{act}, including its catalytic core. Via these multiple contacts, the $PRP19^{HB}$ potentially plays a role in transferring information from other molecular machines to the spliceosome's catalytic core.

8 ANALYSIS OF THE CONFORMATIONAL DYNAMICS OF hB^{act}

Whereas the spliceosome's core is generally stable and resolved at 3.0–4.5 Å in the various complexes analyzed by cryo-EM, peripheral regions are less well-resolved and the most highly mobile elements or protein domains are not visible in many of the EM density maps. Recently, the conformational dynamics of hB^{act} were investigated by a quantitative method (Haselbach et al. 2017) that combines iterative 3D classification with a newly developed approach based on 3D principal-component analysis (PCA) (Haselbach et al. 2018). Eight major conformational states of B^{act} present in the same cryo-EM sample were obtained (Haselbach et al. 2018). Transitions between these different conformational states can be achieved by the available thermal energy. Comparison of their cryo-EM structures (Fig. 5) allows insight into the conformational variability of the more mobile parts/proteins of B^{act}, which include the PRP19^{HB}, IBC proteins, the U2 snRNP 3′ domain, as well as SKIP, PRP17, PPIL1, and PPIE. During its procession from state 1 to 8, certain conformations of hB^{act} are first generated after the coordinated translocation or stable integration of other proteins (Fig. 5). Many of the newly created protein–protein and protein–RNA interactions, and protein conformations observed in state 8, are most similar to those in human C or C* complexes (Bertram et al. 2017b;

Zhang et al. 2017; Zhan et al. 2018). For example, the position and conformation of the SYF1 and SYF3 HAT repeats change from state 2 to 8, but in state 8 they are most similar to those observed in C complexes (Fig. 5B). This supports the idea that the conformational maturation of B^{act} proceeds from state 1 to 8, with the latter likely representing the latest maturation state of hB^{act} before it is remodeled by PRP2 into a catalytically active complex. Thus, this novel method can potentially be used to both identify and temporally order the many distinct conformational states of the spliceosome during its assembly/work cycle.

9 STRUCTURE OF HUMAN C AND C* COMPLEXES

The cryo-EM structures of the human C and C* complexes (Bertram et al. 2017b; Zhang et al. 2017; Zhan et al. 2018) revealed the spliceosome's molecular architecture during the catalytic steps of splicing, and also RNP rearrangements likely mediated by the RNA helicase PRP16 during C* formation (Fig. 6). The catalytic RNA network is highly similar in both hC and yC complexes (Galej et al. 2016; Wan et al. 2016b; Zhan et al. 2018). Furthermore, in both complexes U2 snRNA has rearranged, now forming stem IIc instead of stem IIa (Fig. 1H), which in earlier complexes interacts with SF3b proteins. This U2 restructuring may thus be triggered by the release/destabilization of SF3b during the B^{act} to C transition. Stem IIc persists in the C* complex and, at least in yeast, in the postsplicing P complex (Bai et al. 2017; Liu et al. 2017; Wilkinson et al. 2017). Release of the U2/BS helix from SF3B1 by PRP2 allows it to dock into the catalytic center of the C complex, in which it is tethered by the step 1 splicing factors CCDC49 (Cwc25) and CCDC94 (Yju2), which also interact with each other. Residues in the amino terminus of CCDC94 contact ISY1 and PRP8^{NTD}, as well as the 5′ exon and U2/BS helix (Fig. 6B) (Zhan et al. 2018). CCDC49 is sandwiched between the U2/BS helix and U2/U6 helix Ia, and contacts the Prp8 α-finger. The close contact of CCDC49 and CCDC94 with the U2/BS helix is consistent with their important role in promoting step 1 of splicing (Liu et al. 2007; Chiu et al. 2009; Warkocki et al. 2009), where they likely help to position the BS-A for step 1 catalysis. Both proteins interact with BRR2 and PRP16, which are located at the periphery of hC, and thus link these helicases to the catalytic core.

Before step 2, the branched intron structure (BIS) generated during step 1 of splicing must be displaced from the catalytic center to allow juxtapositioning of the step 2 reactants, the 3′-OH of the 5′ exon and the 3′SS. PRP16 action leads to its own release and destabilization/reorganization of CCDC49 and CCDC94 (Horowitz 2012), and reorganization of ISY1, which remains bound to hC* (Bertram et al. 2017b). This in turn enables the repositioning of the BIS
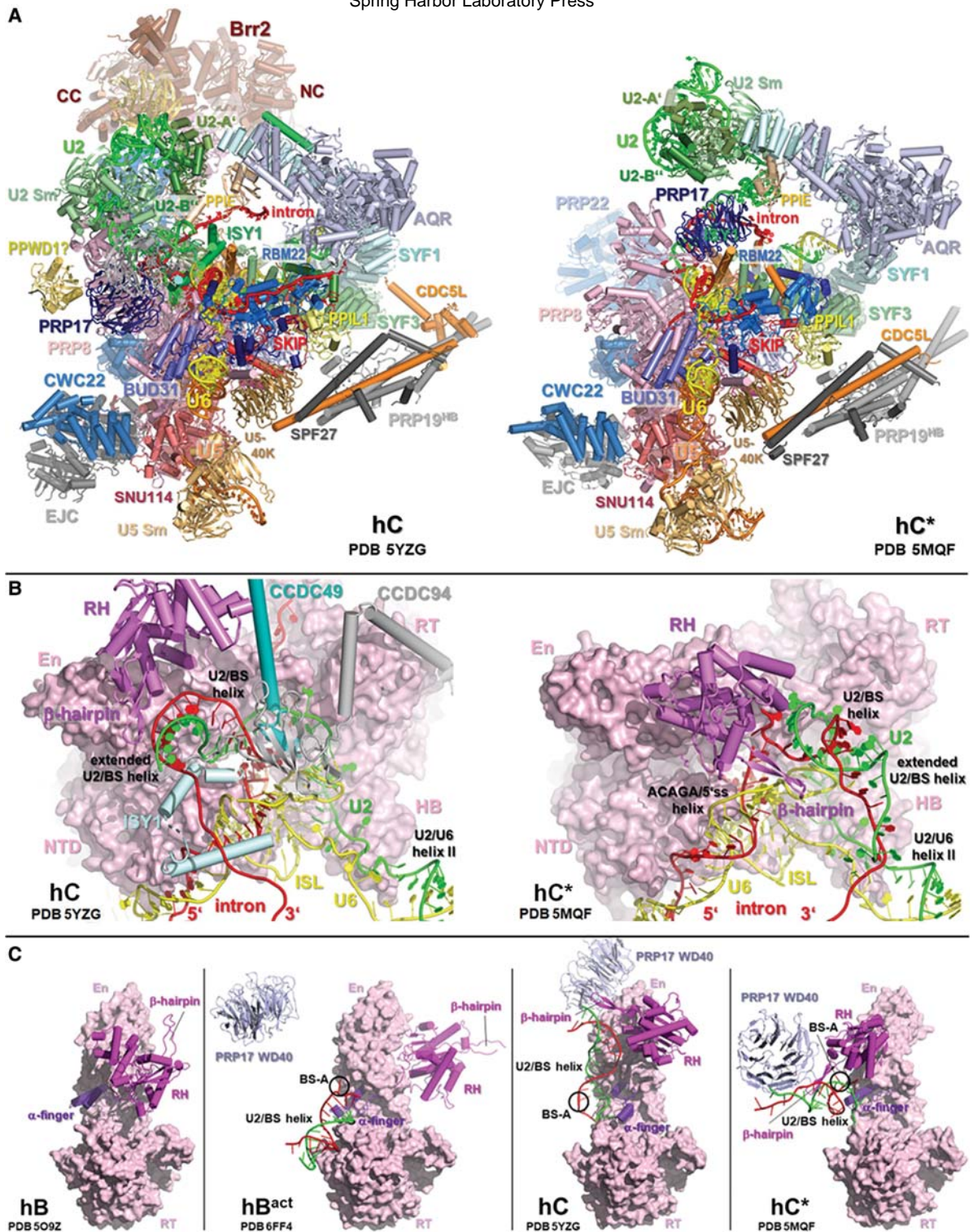
Figure 6. Molecular organization of the human C and C* complexes. (*A*) Overview of the structural organization of the human C and C* complexes. In higher eukaryotes, the exon junction complex (EJC) is deposited 20–24 nucleotides upstream of the exon–exon junction of spliced mRNAs (Le Hir et al. 2000). In hC*, eIF4AIII binds 6 nts of the 5′ exon at a position close to where the EJC should be deposited, and at the same time interacts with MAGOH and MLN51, and U5-SNU114. (*B*) Interaction of PRP8^{RH} with the U2/BS helix in the C (*left*) and C* complex (*right*), as well as interactions of CCDC49 (CWC25), CCDC94 (YJU2) and ISY1 within the RNP core of the C complex (*left*). (*C*) Conformational dynamics of the PRP8 RH domain and movements of the U2/BS helix. The position of the branched intron structure formed after step 1 of splicing is stabilized by contacts with PRP8^{RH} and Prp17. The branch site adenosine (BS-A) is indicated by a black circle.

~20 Å away from the catalytic center, creating space for the 3′ exon to dock near the 5′ exon (Fig. 6B) and the recruitment of the step 2 factors SLU7 and PRP18 and the RNA helicase PRP22. PRP8^{RH} is a highly mobile domain whose position changes throughout the splicing cycle (Fig. 6C). In hC, it interacts with the extended region of the U2/BS helix, CCDC49 and the U2 3′ domain. During the C to C* transition, PRP8^{RH} is rotated by 70 degrees and its β-hairpin is inserted between the U2/BS helix and the U6/intron helix; it thus appears to play an important part in stabilizing the conformation/position of the BIS in C* (Bertram et al. 2017b; Zhang et al. 2017). Whereas the position of the amino-terminal domain of PRP17—which has an extended conformation and contacts multiple proteins (Fig. 4E)—does not change, its WD40 domain (PRP17^{WD40}) moves ~70 Å, contacting the PRP8^{RH} and the extended U2/BS helix and adjacent U2 nucleotides in hC* (Fig. 6). The step 2 factor SLU7 also interacts tightly with PRP8^{RH}, stabilizing its rotated conformation. Thus, both PRP17 and SLU7 likely aid in stabilizing the step 2 conformation of the spliceosome, consistent with them acting as step 2 factors (Horowitz 2012). Although the step 2 factor PRP18 was not detected in hC* complexes, in yC* it interacts with Slu7 and PRP8^{RH}, and thus also contributes to the formation of a step 2 competent spliceosome (Fica et al. 2017; Yan et al. 2017).

In both hC and hC*, the 5′ exon is tightly tethered to the spliceosome via base-pairing with U5 loop 1 nucleotides U40–U42 and via multiple contacts with the exon binding channel proteins (see above). Consistent with previous studies showing that CWC22 facilitates exon junction complex (EJC) formation (Barbosa et al. 2012), in C*, the amino-terminal MIF4G domain of CWC22 contacts the RNA helicase eIF4AIII of the EJC that additionally contains the proteins Y14, MAGOH, and MLN51 (Fig. 6A).

10 INSIGHTS INTO THE MECHANISM OF ACTION OF SPLICEOSOMAL HELICASES

CWC22 binds across the EBC and, together with its interaction partners SLU7 and PRP18, was shown in yeast to be released after step 2 by the action of PRP22, leading to release of the spliced mRNA (Fourmann et al. 2013). PRP22 is located at the periphery of the human C* complex and interacts via its carboxy-terminal OB domain with PRP8, and a long α-helix of SKIP (Fig. 6A). As PRP8 interacts with CWC22, PRP22 action may indirectly lead to release of CWC22. Cryo-EM revealed that PRP2, PRP16 and PRP22 bind sequentially to the periphery of the spliceosome, and all appear to be located far from their RNA targets (i.e., the U2/BS in the case of PRP2 or at/near the spliceosome's catalytic center in the case of PRP16/PRP22) (Fig. 7). This supports the idea that these helicases act indirectly, at a distance from their RNA targets (Rauhut et al. 2016; Semlow et al. 2016). The cryo-EM structures also revealed that each of these helicases interacts primarily with those proteins that are destabilized/released up on their action, and that their stable association with the spliceosome does not require RNA binding. Protein release or RNP remodeling by these helicases, which is dependent on their interaction with intronic RNA downstream from the BS, generates new conformational space that allows the subsequent integration of new spliceosomal proteins or the stabilization/rearrangement of protein domains or proteins already present in the spliceosome.

11 LARGE-SCALE MOVEMENTS OF SPLICEOSOMAL COMPONENTS

Cryo-EM has revealed that many spliceosomal proteins and large protein ensembles undergo large-scale movements during splicing. The major structural rearrangements that occur during the transitions from one spliceosomal complex to the next are summarized in Figure 7. These rearrangements require a dynamic network of molecular interactions among numerous spliceosomal components. The ~300-kDa U2 snRNP 3′ domain is a prime example, which unexpectedly undergoes complex translocations of 10–16 nm between complexes B, B^{act}, C, and C* (Fig. 7). SYF1, which contacts the U2 3′ domain via the U2-A′ protein in the latter three complexes, and SYF3 undergo movements in a concerted manner with the U2 3′ domain, suggesting they aid in its repositioning (Fig. 7A). The movement of the U2 3′ domain first toward the catalytic center during B^{act} formation, and then away from it after catalytic step 1, may be driven by the repositioning of the U2/BS helix, whose rotational movements closely mirror those of the U2 3′ domain (Figs. 6C and 7A). Four nucleotides of the human U2 snRNA (A29 to U32), located between the U2/BS helix and U2 nucleotides that form U2/U6 helix Ia, appear to act as a hinge and enable these complex U2 movements. Brr2, together with PRP8^{Jab1} (which tethers BRR2 to the spliceosome) also undergoes large scale movements during most transitions from one spliceosomal complex to the next (Fig. 7A). As BRR2 helicase activity is negatively regulated in B^{act}, C, and C* complexes, its restructuring may reflect its potential role in stabilizing the position/conformation of its interaction partners at each of these stages.

Multiple domains of PRP8 undergo numerous rearrangements during spliceosome assembly and catalytic activation. As described above, its RT, Linker, and En domains undergo substantial changes relative to its NTD during the integration of the tri-snRNP and during the B to B^{act} and B^{act} to C transitions (Fig. 7B); these major PRP8 domains
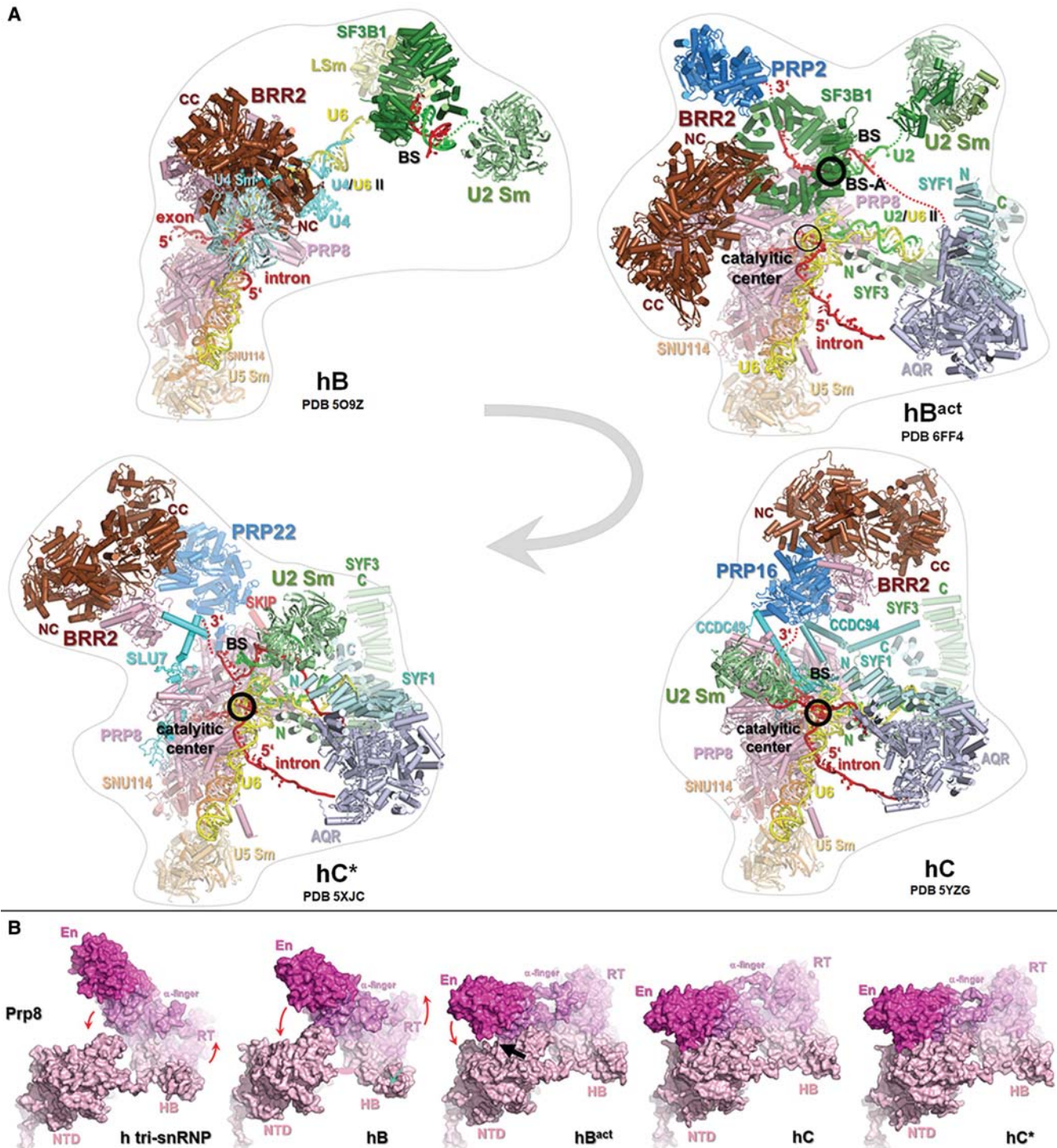
Figure 7. Structural dynamics of the spliceosome. (*A*) Cartoon depicting the major structural rearrangements that occur during the transitions from the human B to B^{act}, B^{act} to C, and C to C* complex. (*B*) Rearrangements of the Prp8 reverse transcriptase-like (RT) domain, Linker (α-finger), endonuclease-like (En) domain, helical bundle (HB), and amino-terminal (NTD) domain during B, B^{act}, and C complex formation (indicated by arrows in the preceding complex). Green arrow indicates rotation of HB relative to the RT domain. Black arrow indicates the cleft between the Prp8 NTD and En domains in the B^{act} complex that are occupied initially by regions of RNF113A and CWC27.

remain largely unchanged after formation of the C complex. PRP8[RH] is repositioned at each spliceosomal stage analyzed to date and both the PRP8 α-finger and switch loop (SWL) are repositioned and/or rearranged during the spliceosome's assembly and catalytic phases (Fig. 6C and see above). The SWL comprises part of the 5′ exon binding channel in hB[act], hC, and C* and toggles between two different conformations. In C* and the yeast post step 2 P complex, the α-finger contacts the 3′SS and nucleotides near the 5′ end of the 3′ exon (Bai et al. 2017; Liu et al. 2017; Wilkinson et al. 2017). Thus, together with U5 loop 1, the PRP8 α-finger aids in stabilizing/positioning the 3′SS and 3′ exon, both before and after exon ligation. Whether this and other functionally important interactions after exon ligation are conserved in humans and yeast is presently not clear, as high-resolution, cryo-EM structures of the human P complex and ILS remain to be determined.

12 FUTURE DIRECTIONS

Currently available cryo-EM structures of the human spliceosome and tri-snRNP provide important insights into the potential structural and mechanistic effects of disease-related mutations in spliceosomal proteins. These mutations may also have kinetic effects, altering the rate at which one spliceosomal complex is converted into the next, which may influence which proteins (including regulatory ones) ultimately bind. As many alternative splicing choices are made during the earliest stages of spliceosome assembly, future cryo-EM analyses of human spliceosomal complexes formed before complex B, will undoubtedly provide key insights into not only the structural basis for splice site recognition by early splicing factors, but also how mutations in these factors may alter splicing outcomes. Structural analyses of spliceosomes formed on alternatively spliced pre-mRNAs may also reveal the complex mechanisms whereby splice site choices are regulated. The structural information provided by cryo-EM can also be exploited to understand how splicing inhibitors/modulators act, and to design structure-based derivatives that enhance their therapeutic efficacy. In the past three years, cryo-EM has revolutionized the structural analysis of the spliceosome, and it will certainly provide a plethora of intriguing new insights into early spliceosome assembly events and the regulation of alternative splicing in higher eukaryotes in the future.

ACKNOWLEDGMENTS

REFERENCES

Absmeier E, Wollenhaupt J, Mozaffari-Jovin S, Becke C, Lee CT, Preussner M, Heyd F, Urlaub H, Lührmann R, Santos KF, et al. 2015. The large N-terminal region of the Brr2 RNA helicase guides productive spliceosome activation. *Genes Dev* **29:** 2576–2587.

Agafonov DE, Deckert J, Wolf E, Odenwälder P, Bessonov S, Will CL, Urlaub H, Lührmann R. 2011. Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol Cell Biol* **31:** 2667–2682.

Agafonov DE, Kastner B, Dybkov O, Hofele RV, Liu WT, Urlaub H, Lührmann R, Stark H. 2016. Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* **351:** 1416–1420.

Alsafadi S, Houy A, Battistella A, Popova T, Wassef M, Henry E, Tirode F, Constantinou A, Piperno-Neumann S, Roman-Roman S, et al. 2016. Cancer-associated *SF3B1* mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat Commun* **7:** 10615.

Bai R, Yan C, Wan R, Lei J, Shi Y. 2017. Structure of the post-catalytic spliceosome from *Saccharomyces cerevisiae. Cell* **171:** 1589–1598.

Bao P, Will CL, Urlaub H, Boon KL, 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:** 2416–2429.

Barbosa I, Haque N, Fiorini F, Barrandon C, Tomasetto C, Blanchette M, Le Hir H. 2012. Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex formation. *Nat Struct Mol Biol* **19:** 983–990.

Bertram K, Agafonov DE, Dybkov O, Haselbach D, Leelaram MN, Will CL, Urlaub H, Kastner B, Lührmann R, Stark H. 2017a. Cryo-EM structure of a pre-catalytic human spliceosome primed for activation. *Cell* **170:** 701–713.

Bertram K, Agafonov DE, Liu WT, Dybkov O, Will CL, Hartmuth K, Urlaub H, Kastner B, Stark H, Lührmann R. 2017b. Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **542:** 318–323.

Boesler C, Rigo N, Anokhina MM, Tauchert MJ, Agafonov DE, Kastner B, Urlaub H, Ficner R, Will CL, Lührmann R. 2016. A spliceosome intermediate with loosely associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. *Nat Commun* **7:** 11997.

Chan SP, Kao DI, Tsai WY, Cheng SC. 2003. The Prp19-associated complex in spliceosome activation. *Science* **302:** 279–282.

Chiu YF, Liu YC, Chiang TW, Yeh TC, Tseng CK, Wu NY, Cheng SC. 2009. Cwc25 is a novel splicing factor required after Prp2 and Yju2 to facilitate the first catalytic reaction. *Mol Cell Biol* **29:** 5671–5678.

Cordin O, Beggs JD. 2013. RNA helicases in splicing. *RNA Biol* **10:** 83–95.

Cretu C, Schmitzová J, Ponce-Salvatierra A, Dybkov O, De Laurentiis EI, Sharma K, Will CL, Urlaub H, Lührmann R, Pena V. 2016. Molecular architecture of SF3b and structural consequences of its cancer-related mutations. *Mol Cell* **64:** 307–319.

Cretu C, Agrawal AA, Cook A, Will CL, Fekkes P, Smith PG, Lührmann R, Larsen N, Buonamici S, Pena V. 2018. Structural basis of splicing modulation by antitumor macrolide compounds. *Mol Cell* **70:** 265–273.

Daguenet E, Dujardin G, Valcárcel J. 2015. The pathogenicity of splicing defects: Mechanistic insights into pre-mRNA processing inform novel therapeutic approaches. *EMBO Rep* **16:** 1640–1655.

Darman RB, Seiler M, Agrawal AA, Lim KH, Peng S, Aird D, Bailey SL, Bhavsar EB, Chan B, Colla S, et al. 2015. Cancer-associated *SF3B1* hotspot mutations induce cryptic 3′ splice site selection through use of a different branch point. *Cell Rep* **13:** 1033–1045.

De I, Bessonov S, Hofele R, dos Santos K, Will CL, Urlaub H, Lührmann R, Pena V. 2015. The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. *Nat Struct Mol Biol* **22:** 138–44.

Dziembowski A, Ventura AP, Rutz B, Caspary F, Faux C, Halgand F, Laprévote O, Séraphin B. 2004. Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. *EMBO J* **23:** 4847–4856.

Fabrizio P, Dannenberg J, Dube P, Kastner B, Stark H, Urlaub H, Lührmann R. 2009. The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol Cell* **36:** 593–608.

Fica SM, Tuttle N, Novak T, Li NS, Lu J, Koodathingal P, Dai Q, Staley JP, Piccirilli JA. 2013. RNA catalyses nuclear pre-mRNA splicing. *Nature* **503:** 229–234.

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:** 464–471.

Fica SM, Oubridge C, Galej WP, Wilkinson ME, Bai XC, Newman AJ, Nagai K. 2017. Structure of a spliceosome remodelled for exon ligation. *Nature* **542:** 377–380.

Finci LI, Zhang X, Huang X, Zhou Q, Tsai J, Teng T, Agrawal A, Chan B, Irwin S, Karr C, et al. 2018. The cryo-EM structure of the SF3b spliceosome complex bound to a splicing modulator reveals a pre-mRNA substrate competitive mechanism of action. *Genes Dev* **32:** 309–320.

Fourmann JB, Schmitzová J, Christian H, Urlaub H, Ficner R, Boon KL, Fabrizio P, Lührmann R. 2013. Dissection of the factor requirements for spliceosome disassembly and the elucidation of its dissociation products using a purified splicing system. *Genes Dev* **27:** 413–428.

Galej WP, Wilkinson ME, Fica SM, Oubridge C, Newman AJ, Nagai K. 2016. Cryo-EM structure of the spliceosome immediately after branching. *Nature* **537:** 197–201.

Gozani O, Feld R, Reed R. 1996. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes Dev* **10:** 233–243.

Gozani O, Potashkin J, Reed R. 1998. A potential role for U2AF–SAP155 interactions in recruiting U2 snRNP to the branch site. *Mol Cell Biol* **18:** 4752–4760.

Grainger RJ, Beggs JD. 2005. Prp8 protein: At the heart of the spliceosome. *RNA* **11:** 533–557.

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

Haselbach D, Schrader J, Lambrecht F, Henneberg F, Chari A, Stark H. 2017. Long-range allosteric regulation of the human 26S proteasome by 20S proteasome-targeting cancer drugs. *Nat Commun* **8:** 15578.

Haselbach D, Komarov I, Agafonov DE, Hartmuth K, Graf B, Dybkov O, Urlaub H, Kastner B, Lührmann R, Stark H. 2018. Structure and conformational dynamics of the human spliceosomal B[act] complex. *Cell* **172:** 454–464.

Henning LM, Santos KF, Sticht J, Jehle S, Lee CT, Wittwer M, Urlaub H, Stelzl U, Wahl MC, Freund C. 2017. A new role for FBP21 as regulator of Brr2 helicase activity. *Nucleic Acids Res* **45:** 7922–7937.

Hirose T, Ideue T, Nagai M, Hagiwara M, Shu MD, Steitz JA. 2006. A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol Cell* **23:** 673–684.

Horowitz DS. 2012. The mechanism of the second step of pre-mRNA splicing. *Wiley Interdiscip Rev RNA* **3:** 331–350.

Huang YH, Chung CS, Kao DI, Kao TC, Cheng SC. 2014. Sad1 counteracts Brr2-mediated dissociation of U4/U6.U5 in tri-snRNP homeostasis. *Mol Cell Biol* **34:** 210–220.

Jenkins JL, Kielkopf CL. 2017. Splicing factor mutations in myelodysplasias: Insights from spliceosome structures. *Trends Genet* **33:** 336–348.

Jurica MS, Moore MJ. 2003. Pre-mRNA splicing: Awash in a sea of proteins. *Mol Cell* **12:** 5–14.

Lardelli RM, Thompson JX, Yates JR III, Stevens SW. 2010. Release of SF3 from the intron branchpoint activates the first step of pre-mRNA splicing. *RNA* **16:** 516–528.

Lee Y, Rio DC. 2015. Mechanisms and regulation of alternative splicing. *Annu Rev Biochem* **84:** 291–323.

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:** 6860–6869.

Liu HL, Cheng SC. 2012. The interaction of Prp2 with a defined region of the intron is required for the first splicing reaction. *Mol Cell Biol* **32:** 5056–5066.

Liu YC, Chen HC, Wu NY, Cheng SC. 2007. A novel splicing factor, Yju2, is associated with NTC and acts after Prp2 in promoting the first catalytic reaction of pre-mRNA. *Mol Cell Biol* **27:** 5403–5413.

Liu S, Li X, Zhang L, Jiang J, Hill RC, Cui Y, Hansen KC, Zhou ZH, Zhao R. 2017. Structure of the yeast spliceosomal postcatalytic P complex. *Science* **358:** 1278–1283.

Makarova OV, Makarov EM, Urlaub H, Will CL, Gentzel M, Wilm M, Lührmann R. 2004. A subset of human 35S proteins, including Prp19, function prior to catalytic step 1 of splicing. *EMBO J* **23:** 2381–2391.

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:** 80–84.

Newman AJ, Norman C. 1992. U5 snRNA interacts with exon sequences at 5′ and 3′ splice sites. *Cell* **68:** 743–754.

Nguyen THD, Galej WP, Bai XC, Savva CG, Newman AJ, Scheres SH, Nagai K. 2015. The architecture of the spliceosomal U4/U6.U5 tri-snRNP. *Nature* **523:** 47–52.

Nguyen THD, Galej WP, Bai XC, Oubridge C, Newman AJ, Scheres SHW, Nagai K. 2016. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. *Nature* **530:** 298–302.

Ohrt T, Prior M, Dannenberg J, Odenwälder P, Dybkov O, Rasche N, Schmitzová 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.

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:** 1413–1415.

Plaschka C, Lin PC, Nagai K. 2017. Structure of a pre-catalytic spliceosome. *Nature* **546:** 617–621.

Plaschka C, Newman AJ, Nagai K. 2019. Structural basis of nuclear pre-mRNA splicing: Lessons from yeast. *Cold Spring Harb Perspect Biol* doi: 10.1101/cshperspect.a032391.

Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K. 2009. Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* **458:** 475–480.

Rasche N, Dybkov O, Schmitzová J, Akyildiz B, Fabrizio P, Lührmann R. 2012. Cwc2 and its human homologue RBM22 promote an active conformation of the spliceosome catalytic centre. *EMBO J* **31:** 1591–1604.

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

Cold Spring Harbor Perspectives in Biology

Santos KF, Jovin SM, Weber G, Pena V, Lührmann R, Wahl MC. 2012. Structural basis for functional cooperation between tandem helicase cassettes in Brr2-mediated remodeling of the spliceosome. *Proc Natl Acad Sci* **109:** 17418–17423.

Schaffert N, Hossbach M, Heintzmann R, Achsel T, Lührmann R. 2004. RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. *EMBO J* **23:** 3000–3009.

Scheres SH, Nagai K. 2017. CryoEM structures of spliceosomal complexes reveal the molecular mechanism of pre-mRNA splicing. *Curr Opin Struct Biol* **46:** 130–139.

Schneider C, Agafonov DE, Schmitzová J, Hartmuth K, Fabrizio P, Lührmann 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 step1 catalysis in yeast spliceosomes. *PLoS Genet* **11:** e1005539.

Semlow DR, Blanco MR, Walter NG, Staley JP. 2016. Spliceosomal DEAH-box ATPases remodel pre-mRNA to activate alternative splice sites. *Cell* **164:** 985–998.

Shi Y. 2017. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. *Nat Rev Mol Cell Biol* **18:** 655–670.

Sontheimer EJ, Steitz JA. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* **262:** 1989–1996.

Staley JP, Guthrie C. 1999. An RNA switch at the 5′ splice site requires ATP and the DEAD box protein Prp28p. *Mol Cell* **3:** 55–64.

Tang Q, Rodriguez-Santiago S, Wang J, Pu J, Yuste A, Gupta V, Moldón A, Xu YZ, Query CC. 2016. SF3B1/Hsh155 HEAT motif mutations affect interaction with the spliceosomal ATPase Prp5, resulting in altered branch site selectivity in pre-mRNA splicing. *Genes Dev* **30:** 2710–2723.

Thapar R. 2015. Roles of prolyl isomerases in RNA-mediated gene expression. *Biomolecules* **5:** 974–999.

Wan R, Yan C, Bai R, Wang L, Huang M, Wong CC, Shi Y. 2016a. The 3.8 Å structure of the U4/U6.U5 tri-snRNP: Insights into spliceosome assembly and catalysis. *Science* **351:** 466–475.

Wan R, Yan C, Bai R, Huang G, Shi Y. 2016b. Structure of a yeast catalytic step I spliceosome at 3.4Å resolution. *Science* **353:** 895–904.

Warkocki Z, Odenwälder P, Schmitzová J, Platzmann F, Stark H, Urlaub H, Ficner R, Fabrizio P, Lührmann R. 2009. Reconstitution of both steps of *Saccharomyces cerevisiae* splicing with purified spliceosomal components. *Nat Struct Mol Biol* **16:** 1237–1243.

Weber G, Trowitzsch S, Kastner B, Lührmann R, Wahl MC. 2010. Functional organization of the Sm core in the crystal structure of human U1 snRNP. *EMBO J* **29:** 4172–4184.

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:** 1283–1288.

Will CL, Lührmann R. 2011. Spliceosome structure and function. *Cold Spring Harb Perspect Biol* **3:** a003707.

Will CL, Schneider C, MacMillan AM, Katopodis NF, Neubauer G, Wilm M, Lührmann R, Query CC. 2001. A novel U2 and U11/U12 snRNP protein that associates with the pre-mRNA branch site. *EMBO J* **20:** 4536–4546.

Yan C, Wan R, Bai R, Huang G, Shi Y. 2016. Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* **353:** 904–911.

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

Zhao C, Pyle AM. 2017. Structural insights into the mechanism of group II intron splicing. *Trends Biochem Sci* **42:** 470–482.

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

Zhang X, Yan C, Hang J, Finci L, Lei J, Shi Y. 2017. An atomic structure of the human spliceosome. *Cell* **169:** 1–12.

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:** 307–322.

Zhou Z, Fu XD. 2013. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma* **122:** 191–207.

Cold Spring Harbor Perspectives in Biology

Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes

Berthold Kastner, Cindy L. Will, Holger Stark and Reinhard Lührmann

Cold Spring Harb Perspect Biol published online February 14, 2019

Subject Collection   RNA Worlds