Structure of the activated human minor spliceosome

Rui Bai[1,2,3]*, Ruixue Wan[1,2,3]*†, Lin Wang[4], Kui Xu[4], Qiangfeng Zhang[4], Jianlin Lei[4,5], Yigong Shi[1,2,3,4]†

[1]Key Laboratory of Structural Biology of Zhejiang Province, School of Life Sciences, Westlake University, Xihu District, Hangzhou 310024, Zhejiang Province, China.
[2]Westlake Laboratory of Life Sciences and Biomedicine, Xihu District, Hangzhou 310024, Zhejiang Province, China. [3]Institute of Biology, Westlake Institute for Advanced Study, Xihu District, Hangzhou 310024, Zhejiang Province, China. [4]Beijing Advanced Innovation Center for Structural Biology and Advanced Research Center for Biological Structure, Tsinghua-Peking Joint Center for Life Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China. [5]Technology Center for Protein Sciences, Ministry of Education Key Laboratory of Protein Sciences, School of Life Sciences, Tsinghua University, Beijing 100084, China.

*These authors contributed equally to this work.

†Corresponding author. Email: wanruixue@westlake.edu.cn (R.W.); syg@westlake.edu.cn (Y.S.)

The minor spliceosome mediates splicing of the rare but essential U12-type pre-mRNA. Here we report the atomic features of the activated human minor spliceosome determined by cryo-electron microscopy at 2.9-Å resolution. The 5′-splice site and branch point sequence of the U12-type intron are recognized by U6atac and U12 small nuclear RNA (snRNA), respectively. Five newly identified proteins stabilize the conformation of the catalytic center. The zinc finger protein SCNM1 functionally mimics the SF3a complex of the major spliceosome. The RBM48/ARMC7 complex binds the γ-monomethyl phosphate cap at the 5′-end of U6atac snRNA. The U-box protein PPIL2 coordinates loop I of U5 snRNA and stabilizes U5 snRNP. CRIPT stabilizes U12 snRNP. Our study provides a framework for mechanistic understanding of the function of the minor spliceosome.

Pre-mRNA splicing, involving removal of non-coding introns and ligation of the coding exons, is achieved by the spliceosome (*1–3*). An intron is defined by three sequence elements: 5′-splice site (5′SS), branch point sequence (BPS), and 3′-splice site (3′SS) (*4*). Most introns conform to the U2-type and are removed by the major spliceosome that contains the U2 small nuclear RNA (snRNA) (*4–7*). A very small percentage of introns belong to the U12-type (*8, 9*), which were initially identified to have the dinucleotides AT at the 5′-end of their 5′SS and AC as their 3′SS (*10, 11*). Compared to the U2-type, the U12-type introns are characterized by highly conserved 5′SS and BPS (*12–16*). The U12-type genes play an essential role in development and are present in all major eukaryotic taxa including fungi, plants and animals (*17–21*). Pre-mRNA splicing involving U12-type introns is mediated by the minor spliceosome (*22, 23*), which contain five snRNAs: U11, U12, U4atac, U5, and U6atac (*24–27*). Of these snRNAs, only U5 is shared between the major and minor spliceosomes.

Due to its scarcity in cells (*8, 9, 28*), the minor spliceosome represents a challenge for biochemical studies. To date, there is no published protocol on the purification of the minor spliceosome. Our knowledge on the protein components of the minor spliceosome has been mostly derived from mass spectrometry analysis of the purified U11/U12 di-snRNP (*29*) and immunoprecipitation studies on the tri-snRNP (*30*). There is no structural information on the minor spliceosome. It is even unclear how many such U12-specific components are present in each major functional state of the minor spliceosome. Our lack of understanding on the minor

spliceosome contrasts sharply with our comprehensive knowledge of the major spliceosome, for which the major functional states from *S. cerevisiae* and *H. sapiens* have been structurally characterized by cryo-electron microscopy (cryo-EM) since 2015 (*31–34*).

In this study, we report in vitro assembly and purification of the activated human minor spliceosome (minor B[act] complex). In the major B[act] complex, the splicing active site has just been formed but productive docking of the BPS is yet to happen (*35, 36*). We determined the cryo-EM structure of the minor B[act] complex at an average resolution of 2.89 Å. This structure reveals a wealth of information.

Reconstitution and purification of the human minor B[act] complex

The in vitro splicing assay using HeLa nuclear extract was established specifically to study the minor spliceosome 25 years ago (*27*). The U2-type 5′SS, BPS and 3′SS of the intron in the MINX pre-mRNA were replaced by those of the U12-type, resulting in the MINX-U12 pre-mRNA. The specificity of MINX-U12 was confirmed in our in vitro splicing assay modified from the established protocol (*27*) (fig. S1A). Relying on endogenous RNaseH activity in the nuclear extract, two sets of anti-sense DNA oligonucleotides were used to digest either U1/U2/U6 snRNAs (groups 2 and 5) or U11/U12/U6atac snRNAs (groups 3 and 6). With application of the anti-U1/U2/U6 oligos, splicing of MINX was severely diminished (fig. S1A, groups 1 and 2); splicing of MINX-U12 was unaffected (groups 4 and 5). Conversely, with application of anti-

U11/U12/U6atac oligos, splicing of MINX was unaffected (fig. S1A, groups 1 and 3); splicing of MINX-U12 became undetectable (groups 4 and 6).

Next, we truncated MINX-U12 18 nucleotides downstream of the BPS to generate the MINX-U12Δ pre-mRNA, which lacks the binding site for the ATPase/helicase Prp2. As previously shown for the major B^{act} complex (35–37), such a pre-mRNA disables Prp2-mediated spliceosome remodeling, thus favoring accumulation of the minor B^{act} complex. To facilitate purification, three tandem MS2 binding sites were inserted between the 5′SS and BPS of MINX-U12Δ. Following incubation of the HeLa nuclear extract with MINX-U12Δ, the minor B^{act} complex was purified through affinity chromatography and glycerol gradient centrifugation (fig. S1B). Chemical crosslinking was used to stabilize the minor spliceosome. Relying on these strategies, we were able to obtain the minor B^{act} complex as confirmed by the presence of U12, U5, and U6atac snRNAs (fig. S1C). The spliceosomal particles appear intact on the cryo-EM micrograph (fig. S1D). To facilitate protein identification in subsequent cryo-EM analysis, the purified sample was analyzed by mass spectrometry (MS) (table S1).

Overall structure of the human minor B^{act} complex
20,390 micrographs were recorded using a Gatan K3 detector mounted on a Titan Krios microscope, yielding 3.5 million auto-picked particles and good two-dimensional averages (fig. S1E). Through multiple cycles of multi-reference three-dimensional (3D) classifications, 101,443 particles were selected to yield a reconstruction of the minor B^{act} complex at an average resolution of 2.89 Å (Fig. 1A; fig. S2; fig. S3, A to D; and table S2). The core region of the spliceosome reaches 2.6 Å (fig. S3B). Improved reconstruction for the SF3b complex and the 5′-end region of U6atac snRNA was obtained through focused 3D classification and refinement (fig. S2). The quality of the EM maps allows atomic modeling of the human minor B^{act} complex and identification of many previously unrecognized features (figs. S4 to S8).

The final atomic model contains 45 proteins, three snRNAs (U12, U5, U6atac) and the MINX-U12Δ pre-mRNA, amounting to a molecular mass of 1.7 mega-Dalton (Fig. 1A). All 45 proteins are detected by MS (table S1). The minor B^{act} complex comprises U12 snRNP, U5 snRNP, U6atac snRNA, MINX-U12Δ pre-mRNA, two components of the NTC (CDC5L and SYF3/CRNKL1), three components of the NTR (CWC15/AD002, SKIP, PLRG1/PRL1), the RES complex (SNIP1, RBMX2, BUD13), two candidate prolyl peptidyl isomerases (PPIs) (PPIL2 and CWC27/NY-CO-10), and nine splicing factors (ARMC7, CRIPT, CWC22, GPKOW, PRP2, RBM48, RNF113A, SRm160, SRm300) (Fig. 1A). Among the splicing factors, GPKOW (Spp2 in *S. cerevisiae*) is the co-activator of the DEAH-box ATPase/helicase PRP2. U12 snRNP consists of the core (U12 snRNA and U12-Sm ring), the SF3b complex and a newly identified protein SCNM1.

Compared to the major B^{act} complex (35, 36, 38, 39), the cryo-EM structure of the minor B^{act} complex contains four new proteins: SCNM1 of U12 snRNP and three splicing factors ARMC7, CRIPT, and RBM48 (Fig. 1B). All four proteins are clearly defined by EM density maps (figs. S6 to S8). U11/U12-65K is a known components of U12 snRNP (29, 40); but the EM density around the U11/U12-65K region is poor and inadequate for judgment (fig. S6A). Nonetheless, U11/U12-65K shows up in the MS analysis (table S1). Although PPIL2 is also a component of the major B^{act} complex (35, 41), it is atomically modeled only in our current structure. As will be detailed later, the location and putative function of these unique proteins in the minor spliceosome provide an explanation to the departure of a number of proteins that are unique to the major spliceosome.

RNA elements in the human minor B^{act} complex
We modeled 63 nucleotides in U12 snRNA (1-51 and 73-84), 90 nucleotides in U5 (8-71, 85-103 and 110-116), 61 nucleotides in U6atac (1-56 and 61-65), 12 nucleotides in 5′-exon (-12 through -1), and nucleotides 1-19 and 192-223 in the intron of MINX-U12Δ (Fig. 2, A and B; figs. S4 and S5; and table S3). The active site center comprises the intramolecular stem loop (ISL) of U6atac snRNA, helix I of the U12/U6atac duplex, loop I of U5 snRNA, and six metal ions. Helix Ib of the U12/U6atac duplex forms a catalytic triplex with G19, A20 and U46 of U6atac (fig. S4, B to D). Notably, the 3′-end sequences of U6atac snRNA form an internal stem loop (named 3′-stem loop) that is placed in approximately the same location as that of U2/U6 helix II in the major B^{act} complex (35). The γ-monomethyl phosphate cap at the 5′-end of U6atac snRNA is unambiguously defined by the EM map (Fig. 2A and fig. S4C).

Overall organization of the RNA elements in the minor B^{act} complex resembles that in the major B^{act} complex (35), with important local differences (Fig. 3A). Compared to U6 snRNA, U6atac lacks the 5′-stem loop but has a unique 3′-stem loop. Unlike that in the major spliceosome, the U12/U6atac duplex lacks helix II because there are no nucleotides 5′ to helix Ib in U12 snRNA (Fig. 2B). Recognition of the highly conserved U12-type 5′SS (5′-AUAUCCUUU-3′) by U6atac involves six canonical Watson-Crick base pairs and one Hoogsteen base pair, which together offer more stringent sequence specificity than that in the major B^{act} complex (35) (Fig. 3B and fig. S5A). Similarly, the highly conserved U12-type BPS (5′-UCCUUAACUC-3′) is recognized by U12 snRNA through nine Watson-Crick base pairs, leaving only the nucleophile-containing adenine base unpaired (Fig. 3C and fig. S5B). In contrast, the BPS in the major B^{act} complex is recognized by five Watson-Crick base pairs (35).

RNA elements in the active site adopt a nearly identical conformation as those in the major spliceosome (Fig. 3D). Unlike the major B^{act} complex (*35*, *36*), both catalytic metal ions M1 and M2 are loaded in the active site of the minor B^{act} complex (Fig. 3E and fig. S5, C and D). M1, which stabilizes the leaving group during the branching reaction, is coordinated by two phosphate groups from G44 and U46 of U6atac. M2, which activates the nucleophile, is coordinated by three phosphate groups from A26, G27 and U46. M1 and M2 are located 4.0 Å and 3.9 Å away from the phosphate group of the first nucleotide (A1) of 5′SS, respectively (Fig. 3E). The distance between M1 and the 3′-oxygen atom from the scissile phosphodiester bond is 3.3 Å. Therefore, the catalytic metals are poised for catalysis of the branching reaction in the minor B^{act} complex. Coordination of M2 in the minor B^{act} complex is identical to that in the major spliceosome (Fig. 3F), in which M2 is fixed by the phosphates from A53, G54 and U74 of U6 snRNA.

An essential role of SCNM1 in the minor spliceosome

U12 snRNP differs from U2 snRNP in both composition and overall organization (Fig. 4A). U2-A′ and U2-B′′ in the U2 core are known to be undetected in the U12 snRNP (*29*) and probably replaced by U11/U12-65K given its homology to U2-B′′ (*40*). In fact, the C-terminal portion of U11/U12-65K exhibits structural homology with the N-terminal portion of U2-B′′ (*40*) (fig. S6A). Consistent with previous studies (*29*), the SF3a complex, which bridges SF3b and U2 core in the major spliceosome, is absent in the minor spliceosome. The N- and C-domains of the newly identified component SCNM1 are located on two opposing sides of U12 snRNP, spanning a distance of about 80 Å (Fig. 4A). A newly identified splicing factor CRIPT, which contains two zinc finger (ZnF) motifs, stabilizes U12 snRNP through simultaneous interactions with its three components (SF3b14b, SF3b145, SF3b155) (Fig. 4B). CRIPT also directly binds to the splicing factor RNF113A, an ortholog of the yeast splicing factor Cwc24.

SCNM1, a modifier of the severity of sodium channelopathy in mice (*42*, *43*), functionally mimics the SF3a complex of the major spliceosome. SCNM1 comprises an N-domain (residues 1-75), a C-domain (residues 183-226), and an intervening flexible linker (residues 76-182) (Fig. 4C and fig. S6C). Both N- and C-domains of SCNM1 are highly conserved in vertebrates (fig. S9). The N-domain contains a C2H2-type ZnF and shares sequence homology with that of SF3a66 from the SF3a complex (Fig. 4D). Similar to SF3a66, the SCNM1 ZnF interacts with the BPS/U12 duplex and the proteins SF3b155, SF3b145 and CDC5L (Fig. 4E). The functional mimicry is underscored by structural similarity between the ZnF of SCNM1 and that of SF3a66, which occupy the same corresponding location in the minor and major B^{act} complexes (Fig. 4F).

Just like SF3a66 (*35*, *36*, *38*), the N terminus of SCNM1 inserts into the active site to interact with the 5′SS, U6atac snRNA, and RNF113A (Fig. 4G). Similar to that in the major spliceosome (*35*, *36*, *38*), the adenine nucleotide at the 5′-end of 5′SS (A1) is sandwiched by the aromatic side chains of Phe213 and Lys218 of RNF113A. The aromatic side chain of Phe3 from SCNM1 is located next to the backbone of U2 and between the nucleobases of A1 and A3 from 5′SS. The positively charged side chains of Lys4 and Arg5 from SCNM1 donate H-bonds to the backbone phosphates of U24 and A26 from U6atac snRNA. These interactions stabilize the recognition of 5′SS by RNF113A and help maintain the conformation of U6atac. Notably, all three N-terminal residues of SCNM1 are invariable in other vertebrates (fig. S9).

The C-domain of SCNM1 interacts with SF3b130 and SF3b145 (Fig. 4H). The spatial location occupied by the C-domain of SCNM1 overlaps with that occupied by SF3a60, a component of the SF3a complex. In the major B^{act} complex, SF3a60 stabilizes the SF3b complex by interacting with SF3b130 and SF3b145 (*35*, *36*). Therefore, the C-domain of SCNM1 can be regarded as a functional mimicry of SF3a60. Among the three components of the SF3a complex, SF3a60 connects SF3a66 to SF3a120, which interacts with U2-A′ and the Sm ring in the major B^{act} complex (*35*, *36*). SF3a120 is absent in the minor spliceosome. Together, the above analysis explains the absence of the SF3a complex in the minor spliceosome.

PPIL2 stabilizes loop I and U5 snRNP

U5 snRNP is the only shared snRNP between the minor and major spliceosomes (*16*, *22*, *23*). In the minor B^{act} complex, U5 snRNP is stabilized by the newly identified component PPIL2 (Fig. 5A). PPIL2 is one of eight nuclear cyclophilins in human cells (*44*) and is conserved from *S. pombe* to human (fig. S10A). PPIL2 is thought to contain a U-box E3 ubiquitin ligase domain at its N-terminal half and a prolyl peptidyl isomerase (PPIase) domain at its C-terminal half. Our EM reconstruction allows identification of four discrete fragments in the N-terminal half of PPIL2 (Fig. 5, A and B, and fig. S7A). The quality of the EM density for the C-terminal half is insufficient for sequence assignment; but its appearance allows docking of the crystal structure of the PPIL2 PPIase domain [PDB code 1ZKC (*45*)].

The structure of PPIL2 comprises fragment I (residues 2-26), fragment II (residues 33-163), fragment III (residues 191-225), fragment IV (residues 239-263), and a PPIase domain (residues 271-457) (Fig. 5B). Each fragment appears to play a uniquely important role. Fragment I reaches into the catalytic center and interacts with loop I of U5 snRNA, N-domain of PRP8, N terminus of SNU114, and the WD40 domain of the NTR component PLRG1 (Fig. 5C). The N terminus of PPIL2 inserts into the groove on the back side of the duplex between

loop I and 5′-exon. Six contiguous polar or charged residues of PPIL2 (Lys3-Arg4-Gln5-His6-Gln7-Lys8) make specific interactions to loop I (Fig. 5D and fig. S7B). Lys3, Arg4, Gln5 and Gln7 may make direct H-bonds to the bases of A44, U33/U34, C45/C45 and U35 of loop I, respectively. His6, Lys8 and Tyr12 may mediate H-bonds to the phosphate groups of U42, G37 and U34. These interactions likely stabilize the conformation of loop I, which may help anchor the 5′-exon. Notably, the N-terminal residues of PPIL2 are nearly invariable in other organisms except *S. pombe* (fig. S10A).

Unexpectedly, fragment II of PPIL2 contains two U-box E3 ubiquitin-ligase domains: one as predicted (*46*) (U-box I, residues 41 to 91) and the other previously unknown (U-box II, residues 102 to 158) (Fig. 5B and fig. S7C). Structures of the two U-box domains can be superimposed with a pair-wise root mean squared deviation (RMSD) of 1.86 Å over 50 aligned residues (fig. S7D). Close examination reveals conservation of key residues between U-boxes I and II (fig. S10B). U-box I contacts Prp8 and U-box II interacts with SNU114 (Fig. 5E). Fragments III and IV of PPIL2 adopt an extend conformation and interact with PRP8, SNU114, the NTC component CDC5L and the NTR component SKIP, likely stabilizing association of U5 snRNP with NTC and NTR (Fig. 5F). Because NTC and NTR are required for formation of the splicing active site, the multi-faceted interactions by PPIL2 may contribute to not just assembly of the minor spliceosome but also the splicing reaction itself.

The RBM48-ARMC7 complex binds the 5′-end of U6atac snRNA

The RNA binding motif (RRM) containing protein RBM48 is essential for splicing of U12-type introns in both plants and animals (*47, 48*). It interacts with the protein armadillo repeat containing 7 (ARMC7) (*47–49*). RBM48-deficient human cell lines and maize endosperm display aberrant splicing of the U12-type introns, which may lead to developmental defects (*47, 48*). However, the role of RBM48 in U12-dependent splicing remains unknown. In our structure, RBM48 and ARMC7 form a complex, which associates with the U6atac/5′SS duplex and binds the characteristic γ-monomethyl phosphate cap of U6atac (*26*) (Fig. 6A and fig. S8). In addition, RBM48 interacts with the N-terminal α-helix in the N-domain of PRP8, which further binds to the top face of U5-40K.

The γ-monomethyl phosphate cap of the guanine nucleotide (G1) at the 5′-end of U6atac snRNA is recognized by RBM48 (Fig. 6B). The guanine base stacks against the aromatic side chain of Tyr39 and are specifically recognized by two H-bonds from the side chains of Tyr121 and Glu44. The β-phosphate of the cap is recognized by two H-bonds from the side chains of Tyr39 and Asp108. In contrast, the α- and γ-phosphates are each coordinated by a single H-bond from

the main chain amide groups of Ala35 and His15, respectively (Fig. 6B). All key residues involved in recognition of the 5′-cap are invariable in all organisms examined (fig. S11). In addition, the side chain of Asn39 from ARMC7 contacts the phosphate group of U2 from U6atac.

The nucleotides at the 5′-end of U6atac are also stabilized by additional interactions. The aromatic side chain of Tyr16 from ARMC7 stacks against the nucleobases of G3 and U4 from U6atac (Fig. 6C). The guanidinium group of Arg31 from RBM48 donates two H-bonds to the phosphate groups of U4 and U5. As a whole, coordination of the 5′-end sequences of U6atac differs from that in the major spliceosome (Fig. 6D). In contrast to the minor spliceosome, the 5′-end sequences of U6 snRNA form a stem loop in the major B^{act} complex (*35, 36*). RBM22 and BUD31 from the NTR and the splicing factor PRP17 stabilize the 5′-stem loop and downstream sequences of U6 snRNA. RBM22 also directly contacts the extended intron sequences from the U6/5′SS duplex. Compared to the minor B^{act} complex, U5-40K along with the N-terminal helix of PRP8 in the major B^{act} complex are rotated by approximately 45 degrees to facilitate association of RBM22, BUD31 and PRP17.

Discussion

In this study, we report the in vitro assembly and purification of the human minor B^{act} complex. The key here is to obtain active minor spliceosome. To achieve this goal, we first reconstituted an in vitro splicing assay using HeLa cell extract and demonstrated that the U12-type intron can be removed independent of the major snRNAs (fig. S1A). To ensure enrichment of the minor B^{act} complex, we truncated MINX-U12 18-nt downstream of the BPS to eliminate the Prp2-binding site. Notably, this design disables Prp2-mediated spliceosome remodeling but has no impact on Prp2 binding to the spliceosome. These strategies proved effective, resulting in purification of the minor B^{act} complex and subsequent determination of its cryo-EM structure. Gratifyingly, the newly identified protein components in the cryo-EM structure are all detected by MS (table S1). This general approach may be adapted for isolation of other major functional states of the minor spliceosome.

The U12-type introns were initially thought to have the dinucleotide AT the 5′-end of the 5′SS and AC as the 3′SS (*10, 11*), hence the name U4atac and U6atac (*26*). Subsequent studies found GT and AG to be more prevalent in 5′SS and 3′SS, respectively (*14, 15*). Nonetheless, because the dinucleotide AT is more complementary to sequences in U11 snRNA, we chose to have <u>AT</u>ATCCTTT as the 5′SS to favor the minor splicing pathway. Importantly, however, as was first observed in the yeast B^{act} complex (*38, 39*), the first three nucleotides at the 5′-end of the 5′SS remain unpaired because sufficient latitude must be given to the 5′-end nucleotide for the

branching reaction. This structural feature, together with the fact that the U12-type introns have lengthy consensus sequences for 5′SS and BPS, strongly favor the recognition of U12-type introns by the minor spliceosome (*12, 15–17*).

13 proteins of the major B^{act} complex, including five NTC and NTR components, are absent in the structure of the minor B^{act} complex (Fig. 1, A and B). Three proteins of the SF3a complex are substituted by SCNM1, and U2-A′ and U2-B′′ could be replaced by U2-65K as previously known. For the NTR components RBM22 and BUD31, their function has been partially replaced by RBM48-ARMC7 complex, which in addition to binding the 5′-cap of U6atac may also facilitate U12-type splicing by stabilizing the intron sequences (Fig. 7). The explanation for absence of the other proteins remain to be investigated. The absence of PRPF19 and SYF1 in the minor B^{act} complex can be explained by dynamic conformation or assembly.

Five new proteins (ARMC7, CRIPT, RBM48, SCNM1, and PPIL2) have been structurally identified in the minor B^{act} complex (Fig. 7). Among them, PPIL2 is also present in the major pre-B^{act} complex (*50*). It is possible that CRIPT may also be a component of the major spliceosome because of its unblocked location. SCNM1, RBM48, and ARMC7 appear to be unique components of the minor spliceosome (Figs. 1 and 7). These newly identified proteins remain poorly understood. Our study defines these five proteins as hallmarks of the minor B^{act} complex. Given its predominant interactions with the SF3b complex, SCNM1 likely dissociates during the B^{act}-to-B* transition. The other three proteins may stay around to other major functional states.

SCNM1 was first identified as a key modifier of the severity of sodium channelopathy through modulating the proportion of correctly spliced transcripts of the *Scn8a* (Nav1.6) gene in mice (*42, 43*). Due to a 5′SS mutation in *Scn8a*, two neighboring exons are skipped in 90 percent of the mRNA, leading to channelopathy in mice (*51*). Notably, the intron between these two exons belongs to the U12-type. A C-terminal truncation or an internal deletion in SCNM1 increases exon skipping to 95 percent of the mRNA, resulting in lethality (*51, 52*). The channelopathy phenotype could be rescued by transgenic expression of WT SCNM1 (*43*). In our structure, SCNM1 is identified as a key component of the minor spliceosome, with its N- and C-domains each playing a vital role. The pathogenic C-terminal truncation of SCNM1 results in removal of the residues 187-230, thus eliminating the C-domain. The internal truncation (Δ133-196) compromises the integrity of the C-domain and no longer allows a functionally required distance between the N- and C-domains.

RBM48 is thought to be an essential splicing factor for U12-type introns, as its mutations cause abnormality in genome-wide U12-type splicing (*47, 48*). Together with ARMC7, RBM48 binds the 5′-end of U6atac snRNA, which lacks the

5′-stem loop compared to U6 snRNA in the major spliceosome. This interaction also stabilizes six unpaired nucleotides at the 5′-end of U6atac and orients the extended sequences of the intron downstream the U6atac/5′SS duplex in the minor B^{act} complex. In this sense, the RBM48-ARMC7 complex appears to serve a similar role as that by the splicing factor Prp17 and the NTR components RBM22 and BUD31 in the major B^{act} complex (*35, 36*).

PPIL2 is thought to be enriched only in the B^{act} complex (*41*). This observation is consistent with the structural finding that, during the B-to-B^{act} transition, the N-domain of PRP8 and SNU114 as a single bloc are rotated 30 degrees relative to the core domain of Prp8 (*32*), which creates a new surface cleft for recruitment of PPIL2. Unexpectedly, PPIL2 actually contains two U-boxes and both are bound in the surface cleft between SNU114 and the core domain of PRP8. The interface between U-box II and SNU114 in our structure is somewhat similar to that reported recently between the predicted U-box (U-box I) and SNU114 in the major pre-B^{act} complex (*50*). Docking of the PPIase domain of PPIL2 in our minor B^{act} complex resembles that in the major pre-B^{act} complex (*50*). During the B^{act}-to-C transition of the major spliceosome, the binding site for the PPIase domain of PPIL2 is occupied by the NTC component SYF2 (*53*), again consistent with the reported enrichment of PPIL2 only in the B^{act} complex (*41*).

The U12-type introns are present in genes that play an essential role in cells. These genes control DNA replication and repair, transcription, RNA processing and translation, cytoskeletal organization, vesicular transport and voltage-gated ion channel activity (*54*). Usually only one U12-type intron is present in each gene but is likely the rate-limiting step in splicing due to the scarcity of the minor spliceosome. Curiously, however, although U12-type introns are present in all major eukaryotic taxa, they are absent in the popular model organisms *Saccharomyces cerevisiae* and *Caenorhabditis elegans* (*17, 55*). In this regard, investigation of regulators that are unique to minor spliceosome, such as SCNM1, RBM48 and ARMC7, may hold key to understanding the evolution, function, and disease relevance of the U12-type genes.

Although the minor spliceosome has been known for decades, relatively little is known about its composition, functional states, catalysis and regulation, especially when compared to the major spliceosome. Similarly, little is known about its remodeling and the associated ATPase/helicases. The fact that Prp2 and Spp2 are bound to the minor B^{act} complex may constitute unequivocal evidence that the Prp2/Spp2 complex in the minor spliceosome functions similarly as in the major spliceosome. Identical to that observed in the major B^{act} complex (*56*), Prp2 directly interacts with Rse1 and Hsh155 of the minor spliceosome, and this interaction is strengthened by Spp2. Scrutiny of the structural features of the human minor B^{act} complex – the first of its kind – may

reveal additional findings about the minor spliceosome. This study provides a framework for mechanistic understanding of the function of the minor spliceosome.

Materials and Methods

Preparation of the pre-mRNA
The U12-type pre-mRNA for the in vitro splicing assay was modified from the MINX gene, in which the 5′SS, BPS and 3′SS are replaced by consensus sequences of the U12-type intron. The 5′SS, BPS and 3′SS have sequences 5′-AUAUCCUUU-3′, 5′-UCCUUAACUC-3′ and 5′-CAC-3′, respectively. The altered MINX pre-mRNA (referred to as MINX-U12) comprises a 57-nucleotide (nt) 5′-exon, a 228-nt intron, a 51-nt 3′-exon, and three tandem MS2-binding RNA aptamers at the 3′-end of the 3′-exon. To enrich the minor B[act] complex in our in vitro assembly assay (*37*), we further modified MINX-U12 by deleting all sequences beyond nucleotide 18 counting from the 3′-end of the BPS. We then inserted three tandem MS2-binding RNA aptamers between the 5′SS and the BPS. The resulting pre-mRNA, referred to as MINX-U12Δ, was used for spliceosome assembly and purification. The DNA templates for in vitro transcription were generated using PCR, and the RNA substrates were synthesized using the method of T7 runoff transcription.

In vitro splicing assay and RT-PCR
Nuclear extract from HeLa S3 cell lines was prepared for in vitro splicing as described (*57*). The in vitro splicing reaction, assembled in a 20-μl volume, was performed in the presence of 1 nM pre-mRNA substrate and 50% nuclear extract, in the buffer containing 20 mM HEPES-KOH, pH 7.9, 65 mM KCl, 2 mM ATP, 20 mM creatine phosphate, and 3 mM MgCl$_2$. To examine the specificity and activity of MINX-U12, we depleted U1, U2 and U6 snRNAs (so as to examine the impact on U2-type splicing), or U11, U12 and U6atac snRNAs (so as to examine the impact on U12-type splicing) in the nuclear extract using endogenous RNaseH at 30°C for 30 min prior to the splicing reaction (fig. S1A). This was accomplished by incubating the reaction with 1 μM antisense DNA oligonucleotides for each of the U1, U2 and U6 snRNAs (anti-U1 DNA oligo: 5′-CAGGTAAGTAT-3′; anti-U2 DNA oligo: 5′-GAACAGATACTACACTTGA-3′; anti-U6 DNA oligo: 5′-CTTCTCTGTATCGTTCCAATTTTAGTAT-3′), or for each of the U11, U12 and U6atac snRNAs (anti-U11 DNA oligo: 5′-CACGACAGAAGCCCTTTT-3′, 5′-TTCCGCACGCAGAGCAATCG-3′, 5′-GCTTCCGAAATCTCTTGATG-3′, and 5′-GGGCGCCGGGACCAACGATC-3′; anti-U12 DNA oligo: 5′-CCTTACTCATAAGTTTAAGGCA-3′, 5′-GTGAGGATTCGGGCGTCACCCC-3′, 5′-AGGCATCCCGCAAAGTAGGCGGG-3′, and 5′-

CCTTGAGGGCGACCTTTACCCGC-3′; anti-U6 DNA oligo: 5′-CTAACCTTCTCTCCTTTCATACAAC-3′, 5′-CGATGGTTAGATGCCACG AAG-3′, 5′-GAGGGCCTCTTCCATCCTTGTC-3′, and 5′-CAATGCCTTAACCG TATGCGTG-3′). The splicing reaction mixture was incubated at 30°C for varying time points of 0 min, 30 min, 60 min and 120 min, followed by proteinase K digestion. RNA from the in vitro splicing assay was extracted using phenol:chloroform:isopentanol at a volume ratio of 25:24:1 (Coolaber Science & Technology). Reverse transcription was performed using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and random hexamers. The RT-PCR products were resolved on 2% (w/v) agarose gel and stained by GoldView (Beijing SBS Genetech Co Ltd.) (fig. S1A).

Assembly and purification of the human minor B[act] complex
The protocol for assembly and purification of the minor B[act] complex was modified from that for the major B[act] complex (*35*, *36*) (fig. S1B). Briefly, the splicing reaction was performed in a volume of 40 mL or multiples thereof, containing 20 mM HEPES-KOH, pH 7.9, 65 mM KCl, 2 mM ATP, 20 mM creatine phosphate, and 3 mM MgCl$_2$, in the presence of 10 nM pre-mRNA, 450 nM MS2-MBP and 50% splicing extract. The pre-mRNA was pre-bound to MS2-MBP for 30 min on ice. The reaction mixture was incubated for 4 hours at 30°C, and then centrifuged at 3,000 g for 15 min to remove aggregates. The supernatant was loaded onto amylose resin (NEB), allowed to flow through the resin 3 times, and washed using the G50K buffer (10 mM HEPES-KOH, pH 7.9, 50 mM KCl, 1.5 mM MgCl$_2$, 0.01% NP40 and 5% (v/v) glycerol). The spliceosomal complexes were eluted using 30 mM maltose. The eluent was loaded onto a 10~30% glycerol gradient with 0%–0.01% glutaraldehyde (Sigma), and centrifuged at 23,000 rpm for 10 hours at 4°C in a SW41Ti rotor (fig. S1B). Fractions of 1-mL aliquot each were quenched by 50 mM Tris (pH 7.6), and then examined on denaturing RNA gels (fig. S1C). The fractions that contained the minor B[act] complex were pooled and dialyzed against Buffer D (20 mM HEPES-KOH, pH 7.9, 50 mM KCl, 1.5 mM MgCl$_2$, 0.01% NP40) to remove glycerol prior to sample preparation for electron microscopy (EM). The dialyzed sample was concentrated for cryo-EM studies.

Northern blot
Samples of the purified minor B[act] complex were heated at 95°C for 10 min, separated on 10% polyacrylamide gels containing 8M urea. RNAs were transferred onto Hybond-N+ membranes (Amersham) in 1xTBE (Coolaber Science & Technology) for 1 hour at 400 mA. Membranes were fixed by UV irradiation and hybridized to ^{32}P-end-labeled DNA oligonucleotide probes. The U12 snRNA specific probes are derived

from a published study (*24*) and include: 5'-CCTTACTCATAAGTTTAAGGCA-3', 5'-GTGAGGATTCGGGCGTCACCCC-3', AGGCATCCCGCAAAGTAGGCGGG-3', and 5'-CCTTGAGGGCGACCTTTACCCGC-3'. The specific probes of U6atac snRNA include: 5'-CTAACCTTCTCTCCTTTCATACAAC-3', CGATGGTTAGATGCCACGAAG-3', 5'-GAGGGCCTCTTCCATCCTTGTC-3', and 5'-CAATGCCTTAACCGTATGCGTG-3'. The specific probes of U5 snRNA include: 5'-GCGATCTGAAGAGAAACCAGAG-3', 5'- CTTGCCAAAGCAAGGCCTC-3', 5'-GGGTTAAGACTCAGAGTTGTTC-3', and 5'-CTCCACGGAAATCTTTAGTAAAAGGC-3'. Membranes were exposed to phosphorimager screens (Amersham Biosciences) (fig. S1C). The exposure time is approximately 48 hours for the DNA probes targeting U12, U6atac and U5 snRNAs. The phosphorimager screen was scanned by Personal Molecular Imager (BioRad).

Mass spectrometry analysis

About 40 μl of the minor spliceosome sample was mixed with 10 μl of 5×SDS sample loading buffer (GenScript Biotech, China) supplemented with 150 mM dithiothreitol. The sample was incubated at 95°C for 5 min, and resolved using a 4%–12% gradient SDS-PAGE gel. The proteins were subjected to in-gel proteolytic digestion as described (*58*). Peptides were purified using Pierce C18 Spin Tips (Thermo Fisher, USA) prior to LC-MS/MS analysis using Ultimate 3000 nanoLC system coupled with Q Exactive HF-X Hybrid Quadrupole-Orbitrap (Thermo Fisher Scientific, San Jose, USA). About 500 ng peptides were separated over 90 min using a linear LC gradient of 3-28% (buffer A: 2% acetonitrile, 0.1% formic acid; buffer B: 98% acetonitrile, 0.1% formic acid) at a flow rate of 300 nL/min. The top 20 peptides were subjected to MS2 analysis. MS2 spectra were acquired at the resolution of 30,000 (at m/z 200) in the orbitrap using an AGC target of 1e5, and max IT of 80 ms. Dynamic exclusion was applied with a repeat count of 1 and an exclusion time of 25 s. The resultant mass spectrometric data were analyzed using pFind (*59*) (Version 3.1.5) against the *Homo sapiens* FASTA database downloaded from UniProtKB (version on 27-Apr-2020), which contains 20,365 reviewed protein sequences. Cysteine carbamidomethyl was set as fixed modification and methionine oxidation was set as variable modification. A summary of mass spectrometric analysis for the human minor B^{act} complex is listed in table S1.

EM data acquisition and processing of the minor B^{act} complex

Grids for cryo-EM data collection were carried out essentially as described (*60*). Briefly, the Quantifoil R1.2/1.3 grids coated with homemade continuous carbon film of ~4 nm thickness were used for cryo-EM specimen preparation. Cryo-EM grids were prepared using Vitrobot Mark IV (FEI Company) at 8°C and 100% humidity. 4-μl aliquots of the sample at a concentration of ~0.2 mg/mL were applied to glow-discharged grids, blotted for 1.5 s using the standard vitrobot filter paper (Ø55/20mm, Ted Pella), and plunged into liquid ethane cooled by liquid nitrogen. The glow-discharged grids were prepared for 30 s using the "Low" setting of the Plasma Cleaner (Harrick, Plasma Cleaner PDC-32G).

The grids were loaded onto a FEI Titan Krios electron microscope equipped with a GIF Quantum energy filter (slit width 20 eV) and operating at 300 kV with a nominal magnification of 81,000x. Images were recorded using a Gatan K3 detector (Gatan Company) in the super-resolution mode, with a pixel size of 0.5371 Å (fig. S1D). Each image was dose-fractionated to 32 frames with a dose rate of 22.54 counts/sec per physical pixel (~19.531 e-/sec per Å^2) and a total exposure time of 2.56 s. Total electron dose for each image is about 50 e$^-$/Å^2. AutoEMation2 (written by Jianlin Lei) was used for all data collection (*61*). All 32 frames in each stack were aligned and summed using the whole-image motion correction program MotionCor2 (*62*) and binned to a pixel size of 1.0742 Å. The defocus value of each image was set from 1.3 to 2.5 μm and was determined by Gctf (*63*).

EM data processing for the human minor B^{act} complex

3,500,973 particles were automatically picked by DeepPicker (*64*) from 20,390 micrographs (fig. S2). These particles were extracted using a pixel size of 4.2968 Å and subjected to three parallel guided multi-reference global search 3D classification using RELION3.0 (*65*). The volumes representing the human B^{act} complex [EMDB: EMD-6891 (*36*)], 17S U2 snRNP [EMDB: EMD-10689 (*66*)], tri-snRNP [EMDB:EMD-6581 (*67*)], B complex [EMDB: EMD-9624 (*68*)], C complex [EMDB: EMD-6864 (*53*)], and four bad classes were low-pass filtered to 35 Å and used as the initial references (fig. S2). The particles from the best classes were merged and the duplicated particles were removed as described (*69*). Another subsequent three parallel local search 3D classification (30 iterations, T=4, local angular search range=15°) was performed to further identify good particles, using the same references above (Round 1). The resulting 1,217,469 particles were re-centered and re-extracted with a pixel size of 2.1484 Å. After one round of refinement, the resulting map is low-pass filtered to 15 Å, 25 Å, 35 Å, 60 Å. Together with the refinement map, we obtained five references with a resolution gradient. Three parallel guided multi-reference local 3D classifications (30 iterations, T=4, local angular search range=15°) using the five references above were performed, yielding 101,443 good particles. These particles were further re-centered and re-extracted using a pixel size of 1.0742 Å,

yielded a reconstruction at an average resolution of 3.2 Å for the entire human minor B^{act} complex. After CTF refinement by RELION3.0 and Bayesian polishing by RELION3.1, the resolution was further improved to 2.89 Å (figs. S2 and S3A). To improve the map quality for the SF3B region, focused refinement with a soft mask on the SF3b complex was performed (fig. S2). Another round of focused 3D refinement with the mask for one of the WD40 domains from SF3b130 further improved the map quality for this region. For the 5′-end of U6atac snRNA (RBM48/ARMC7 region), focused 3D classification (30 iterations, T=4, local angular search range=15°) was performed with a soft mask around this region. 66,297 particles were selected and subjected to focused 3D refinement using the same mask as 3D classification.

In the final EM map of the human minor B^{act} complex, the local resolution reaches 2.6 Å in the core region (fig. S3B). The angular distribution of the particles used for the final reconstruction of the human minor B^{act} complex is reasonable (fig. S3C). The refinement of the atomic coordinates did not suffer from severe over-fitting (fig. S3D). The EM maps show fine features for individual component of the human minor B^{act} complex (figs. S4 to S8) and allow atomic modeling of five previously unidentified proteins (PPIL2, SCNM1, CRIPT, RBM48 and ARMC7) (figs. S6 to S8).

Reported resolutions were calculated on the basis of the FSC 0.143 criterion, and the FSC curves were corrected using high-resolution noise substitution methods (70). Prior to visualization, all density maps were corrected for the modulation transfer function (MTF) of the detector and sharpened by applying a negative B-factor that was estimated using automated procedures (71). Local resolution variations were estimated using RELION.

Model building and refinement of the human minor B^{act} complex

We combined homology modeling, rigid docking of known structures, de novo modeling and manual adjustment with an AI-guided deep natural network method (table S3). The majority of individual protein components were identified using the atomic coordinates of the human B^{act} complex [PDB code: 5Z56, 5Z58 (36), 6FF7 (35)]. Structures of most components (which include U5 snRNP, U6atac snRNA, U12 snRNA and U12 Sm ring of U12 snRNP, SF3b complex of U12 snRNP, RES complex, two components of NTC, three components of NTR and CWC27, PRP2, RNF113A, SRm160, CWC22, SRm300) were docked into the density map and manually adjusted using COOT (72). Position of the C terminus of U11/U12-65K, which shares homology with U2-B′′, was predicted according to the location of U2-B′′ using the crystal structure 3EGN (73). However, the EM density for this region is inadequate for identification of U11/U12-65K. The location of the N terminus of GPKOW (Spp2 in *S. cerevisiae*) was

identified according to the 2.5 Å cryo-EM structure of the *S. cerevisiae* B^{act} complex (PDB code 7DCO) (56).

After these steps, a number of EM density patches still remained unaccounted for, including the regions around loop I of U5 snRNA, SF3b155, SF3b130, and the 5′-end of U6atac. The N-terminal zinc finger region of SF3a66 (homolog of *S. cerevisiae* Prp11) clearly did not fit into the well-resolved density map. To identify the unknown protein, we used an improved version of the A2-Net deep neural network method (74) to recognize the side chains of amino acids from EM density maps and to connect adjacent amino acids to generate protein fragments. Together with mass spectrometry data, this practice revealed protein candidates. The sequences of the selected protein candidates and the density map were then fed into the deep neural network to generate the atomic model. Models of PPIL2, SCNM1, CRIPT and RBM48 were generated using this approach. Human RBM48 was reported to interact with the armadillo repeat-containing protein 7 (ARMC7) (47, 49). Next to RBM48, there was indeed an EM density lobe that closely resembles the overall shape of the armadillo repeat helices. An initial model for ARMC7 was generated using the protein structure prediction web server trRosetta (75). The predicted structure of ARMC7 was docked into the density map and manually adjusted using COOT (72).

The entire model of the human minor Bact complex was refined against the 2.89-Å map using PHENIX (76) in real space with secondary structure and geometry restraints. Model overfitting was monitored by refining the model in one of the two independent maps and testing the refined model against the other map (77) (fig. S3D). Model quality was assessed using the Molprobity scores and the Ramachandran plots (table S2). Molprobity scores were calculated as described (78). All EM map images were created using UCSF Chimera (79).

REFERENCES AND NOTES

1. E. Brody, J. Abelson, The "spliceosome": Yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **228**, 963–967 (1985). doi:10.1126/science.3890181 Medline

2. P. J. Grabowski, S. R. Seiler, P. A. Sharp, A multicomponent complex is involved in the splicing of messenger RNA precursors. *Cell* **42**, 345–353 (1985). doi:10.1016/S0092-8674(85)80130-6 Medline

3. D. Frendewey, W. Keller, Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell* **42**, 355–367 (1985). doi:10.1016/S0092-8674(85)80131-8 Medline

4. C. B. Burge, T. Tuschl, P. A. Sharp, in *The RNA World, Second Edition*, R. F. Gesteland, T. R. Cech, J. F. Atkins, Eds. (Cold Spring Harbor Laboratory Press, 1999), pp. 525–560.

5. M. J. Moore, C. C. Query, P. A. Sharp, in *The RNA World*, R. F. Gesteland, J. F. Atkins, Eds. (Cold Spring Harbor Laboratory Press, 1993), pp. 303–357.

6. J. P. Staley, C. Guthrie, Mechanical devices of the spliceosome: Motors, clocks, springs, and things. *Cell* **92**, 315–326 (1998). doi:10.1016/S0092-8674(00)80925-3 Medline

7. M. C. Wahl, C. L. Will, R. Lührmann, The spliceosome: Design principles of a dynamic RNP machine. *Cell* **136**, 701–718 (2009). doi:10.1016/j.cell.2009.02.009 Medline

8. A. Levine, R. Durbin, A computational scan for U12-dependent introns in the human genome sequence. *Nucleic Acids Res.* **29**, 4006–4013 (2001). doi:10.1093/nar/29.19.4006 Medline

9. N. Sheth, X. Roca, M. L. Hastings, T. Roeder, A. R. Krainer, R. Sachidanandam, Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res.* **34**, 3955–3967 (2006). doi:10.1093/nar/gkl556 Medline

10. I. J. Jackson, A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res.* **19**, 3795–3798 (1991). doi:10.1093/nar/19.14.3795 Medline

11. S. L. Hall, R. A. Padgett, Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J. Mol. Biol.* **239**, 357–365 (1994). doi:10.1006/jmbi.1994.1377 Medline

12. P. A. Sharp, C. B. Burge, Classification of introns: U2-type or U12-type. *Cell* **91**, 875–879 (1997). doi:10.1016/S0092-8674(00)80479-1 Medline

13. W. Y. Tarn, J. A. Steitz, Pre-mRNA splicing: The discovery of a new spliceosome doubles the challenge. *Trends Biochem. Sci.* **22**, 132–137 (1997). doi:10.1016/S0968-0004(97)01018-9 Medline

14. Q. Wu, A. R. Krainer, Splicing of a divergent subclass of AT-AC introns requires the major spliceosomal snRNAs. *RNA* **3**, 586–601 (1997). Medline

15. R. C. Dietrich, R. Incorvaia, R. A. Padgett, Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. *Mol. Cell* **1**, 151–160 (1997). doi:10.1016/S1097-2765(00)80016-7 Medline

16. A. A. Patel, J. A. Steitz, Splicing double: Insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* **4**, 960–970 (2003). doi:10.1038/nrm1259 Medline

17. C. B. Burge, R. A. Padgett, P. A. Sharp, Evolutionary fates and origins of U12-type introns. *Mol. Cell* **2**, 773–785 (1998). doi:10.1016/S1097-2765(00)80292-0 Medline

18. M. Lynch, A. O. Richardson, The evolution of spliceosomal introns. *Curr. Opin. Genet. Dev.* **12**, 701–710 (2002). doi:10.1016/S0959-437X(02)00360-X Medline

19. A. G. Russell, J. M. Charette, D. F. Spencer, M. W. Gray, An early evolutionary origin for the minor spliceosome. *Nature* **443**, 863–866 (2006). doi:10.1038/nature05228 Medline

20. M. Dávila López, M. A. Rosenblad, T. Samuelsson, Computational screen for spliceosomal RNA genes aids in defining the phylogenetic distribution of major and minor spliceosomal components. *Nucleic Acids Res.* **36**, 3001–3010 (2008). doi:10.1093/nar/gkn142 Medline

21. S. Bartschat, T. Samuelsson, U12 type introns were lost at multiple occasions during evolution. *BMC Genomics* **11**, 106 (2010). doi:10.1186/1471-2164-11-106 Medline

22. C. L. Will, R. Lührmann, Splicing of a rare class of introns by the U12-dependent spliceosome. *Biol. Chem.* **386**, 713–724 (2005). doi:10.1515/BC.2005.084 Medline

23. J. J. Turunen, E. H. Niemelä, B. Verma, M. J. Frilander, The significant other: Splicing by the minor spliceosome. *WIREs RNA* **4**, 61–76 (2013). doi:10.1002/wrna.1141 Medline

24. K. A. Montzka, J. A. Steitz, Additional low-abundance human small nuclear ribonucleoproteins: U11, U12, etc. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8885–8889 (1988). doi:10.1073/pnas.85.23.8885 Medline

25. S. L. Hall, R. A. Padgett, Requirement of U12 snRNA for in vivo splicing of a minor class of eukaryotic nuclear pre-mRNA introns. *Science* **271**, 1716–1718 (1996). doi:10.1126/science.271.5256.1716 Medline

26. W. Y. Tarn, J. A. Steitz, Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. *Science* **273**, 1824–1832 (1996). doi:10.1126/science.273.5283.1824 Medline

27. W. Y. Tarn, J. A. Steitz, A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* **84**, 801–811 (1996). doi:10.1016/S0092-8674(00)81057-0 Medline

28. Y.-T. Yu, E. C. Scharl, C. M. Smith, J. A. Steitz, in *The RNA World, Second Edition*, R. F. Gesteland, T. R. Cech, J. F. Atkins, Eds. (Cold Spring Harbor Laboratory Press, 1999), pp. 487–524.

29. C. L. Will, C. Schneider, M. Hossbach, H. Urlaub, R. Rauhut, S. Elbashir, T. Tuschl, R. Lührmann, The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA* **10**, 929–941 (2004). doi:10.1261/rna.7320604 Medline

30. C. Schneider, C. L. Will, O. V. Makarova, E. M. Makarov, R. Lührmann, Human U4/U6.U5 and U4atac/U6atac.U5 tri-snRNPs exhibit similar protein compositions. *Mol. Cell. Biol.* **22**, 3219–3229 (2002). doi:10.1128/MCB.22.10.3219-3229.2002 Medline

31. R. Wan, R. Bai, Y. Shi, Molecular choreography of pre-mRNA splicing by the spliceosome. *Curr. Opin. Struct. Biol.* **59**, 124–133 (2019). doi:10.1016/j.sbi.2019.07.010 Medline

32. R. Wan, R. Bai, X. Zhan, Y. Shi, How Is Precursor Messenger RNA Spliced by the Spliceosome? *Annu. Rev. Biochem.* **89**, 333–358 (2020). doi:10.1146/annurev-biochem-013118-111024 Medline

33. C. Plaschka, A. J. Newman, K. Nagai, Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast. *Cold Spring Harb. Perspect. Biol.* **11**, a032391 (2019). doi:10.1101/cshperspect.a032391 Medline

34. B. Kastner, C. L. Will, H. Stark, R. Lührmann, Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **11**, a032417 (2019). doi:10.1101/cshperspect.a032417 Medline

35. D. Haselbach, I. Komarov, D. E. Agafonov, K. Hartmuth, B. Graf, O. Dybkov, H. Urlaub, B. Kastner, R. Lührmann, H. Stark, Structure and Conformational Dynamics of the Human Spliceosomal B^{act} Complex. *Cell* **172**, 454–464.e11 (2018). doi:10.1016/j.cell.2018.01.010 Medline

36. X. Zhang, C. Yan, X. Zhan, L. Li, J. Lei, Y. Shi, Structure of the human activated spliceosome in three conformational states. *Cell Res.* **28**, 307–322 (2018). doi:10.1038/cr.2018.14 Medline

37. S. Bessonov, M. Anokhina, A. Krasauskas, M. M. Golas, B. Sander, C. L. Will, H. Urlaub, H. Stark, R. Lührmann, Characterization of purified human B^{act} spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA* **16**, 2384–2403 (2010). doi:10.1261/rna.2456210 Medline

38. C. Yan, R. Wan, R. Bai, G. Huang, Y. Shi, Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* **353**, 904–911 (2016). doi:10.1126/science.aag0291 Medline

39. R. Rauhut, P. Fabrizio, O. Dybkov, K. Hartmuth, V. Pena, A. Chari, V. Kumar, C.-T. Lee, H. Urlaub, B. Kastner, H. Stark, R. Lührmann, Molecular architecture of the *Saccharomyces cerevisiae* activated spliceosome. *Science* **353**, 1399–1405 (2016). doi:10.1126/science.aag1906 Medline

40. H. Benecke, R. Lührmann, C. L. Will, The U11/U12 snRNP 65K protein acts as a molecular bridge, binding the U12 snRNA and U11-59K protein. *EMBO J.* **24**, 3057–3069 (2005). doi:10.1038/sj.emboj.7600765 Medline

41. D. E. Agafonov, J. Deckert, E. Wolf, P. Odenwälder, S. Bessonov, C. L. Will, H. Urlaub, R. Lührmann, Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. *Mol. Cell. Biol.* **31**, 2667–2682 (2011). doi:10.1128/MCB.05266-11 Medline

42. L. K. Sprunger, A. Escayg, S. Tallaksen-Greene, R. L. Albin, M. H. Meisler, Dystonia associated with mutation of the neuronal sodium channel Scn8a and identification of the modifier locus Scnm1 on mouse chromosome 3. *Hum. Mol. Genet.* **8**, 471–479 (1999). doi:10.1093/hmg/8.3.471 Medline

43. D. A. Buchner, M. Trudeau, M. H. Meisler, SCNM1, a putative RNA splicing factor that modifies disease severity in mice. *Science* **301**, 967–969 (2003). doi:10.1126/science.1086187 Medline

44. C. Rajiv, T. L. Davis, Structural and Functional Insights into Human Nuclear Cyclophilins. *Biomolecules* **8**, 161 (2018). doi:10.3390/biom8040161 Medline

45. T. L. Davis, J. R. Walker, V. Campagna-Slater, P. J. Finerty, R. Paramanathan, G. Bernstein, F. MacKenzie, W. Tempel, H. Ouyang, W. H. Lee, E. Z. Eisenmesser, S. Dhe-Paganon, Structural and biochemical characterization of the human cyclophilin family of peptidyl-prolyl isomerases. *PLOS Biol.* **8**, e1000439 (2010). doi:10.1371/journal.pbio.1000439 Medline

46. S. Hatakeyama, M. Yada, M. Matsumoto, N. Ishida, K. I. Nakayama, U box proteins as a new family of ubiquitin-protein ligases. *J. Biol. Chem.* **276**, 33111–33120 (2001). doi:10.1074/jbc.M102755200 Medline

47. F. Bai, J. Corll, D. N. Shodja, R. Davenport, G. Feng, J. Mudunkothge, C. J. Brigolin, F. Martin, G. Spielbauer, C.-W. Tseung, A. E. Siebert, W. B. Barbazuk, S. Lal, A. M. Settles, RNA Binding Motif Protein 48 Is Required for U12 Splicing and Maize Endosperm Differentiation. *Plant Cell* **31**, 715–733 (2019). doi:10.1105/tpc.18.00754 Medline

48. A. E. Siebert, J. Corll, J. P. Gronevelt, L. Levine, L. M. Hobbs, C. Kenney, R. Davenport, A. M. Settles, W. B. Barbazuk, R. J. Westrick, G. J. Madlambayan, S. Lal, Evolutionary Conservation and Essential Function of Human and Maize RNA

Binding Motif Protein 48 (RBM48) in U12-Type Intron Splicing. bioRxiv 2020.07.18.209528 [Preprint]. 19 July 2020. https://doi.org/10.1101/2020.07.18.209528.

49. T. Hart, M. Chandrashekhar, M. Aregger, Z. Steinhart, K. R. Brown, G. MacLeod, M. Mis, M. Zimmermann, A. Fradet-Turcotte, S. Sun, P. Mero, P. Dirks, S. Sidhu, F. P. Roth, O. S. Rissland, D. Durocher, S. Angers, J. Moffat, High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. Cell 163, 1515–1526 (2015). doi:10.1016/j.cell.2015.11.015 Medline

50. C. Townsend, M. N. Leelaram, D. E. Agafonov, O. Dybkov, C. L. Will, K. Bertram, H. Urlaub, B. Kastner, H. Stark, R. Lührmann, Mechanism of protein-guided folding of the active site U2/U6 RNA during spliceosome activation. Science 370, eabc3753 (2020). doi:10.1126/science.abc3753 Medline

51. J. A. Kearney, D. A. Buchner, G. De Haan, M. Adamska, S. I. Levin, A. R. Furay, R. L. Albin, J. M. Jones, M. Montal, M. J. Stevens, L. K. Sprunger, M. H. Meisler, Molecular and pathological effects of a modifier gene on deficiency of the sodium channel Scn8a (Na$_v$1.6). Hum. Mol. Genet. 11, 2765–2775 (2002). doi:10.1093/hmg/11.22.2765 Medline

52. D. A. Buchner, M. Trudeau, A. L. George Jr., L. K. Sprunger, M. H. Meisler, High-resolution mapping of the sodium channel modifier Scnm1 on mouse chromosome 3 and identification of a 1.3-kb recombination hot spot. Genomics 82, 452–459 (2003). doi:10.1016/S0888-7543(03)00152-6 Medline

53. X. Zhan, C. Yan, X. Zhang, J. Lei, Y. Shi, Structure of a human catalytic step I spliceosome. Science 359, 537–545 (2018). doi:10.1126/science.aar6401 Medline

54. B. Verma, M. V. Akinyi, A. J. Norppa, M. J. Frilander, Minor spliceosome and disease. Semin. Cell Dev. Biol. 79, 103–112 (2018). doi:10.1016/j.semcdb.2017.09.036 Medline

55. H. W. Mewes, K. Albermann, M. Bähr, D. Frishman, A. Gleissner, J. Hani, K. Heumann, K. Kleine, A. Maierl, S. G. Oliver, F. Pfeiffer, A. Zollner, Overview of the yeast genome. Nature 387, 7–8 (1997). doi:10.1038/387s007 Medline

56. R. Bai, R. Wan, C. Yan, Q. Jia, J. Lei, Y. Shi, Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. Science 371, eabe8863 (2021). doi:10.1126/science.abe8863 Medline

57. J. D. Dignam, R. M. Lebovitz, R. G. Roeder, Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. 11, 1475–1489 (1983). doi:10.1093/nar/11.5.1475 Medline

58. T. Zhu, Y. Zhu, Y. Xuan, H. Gao, X. Cai, S. R. Piersma, T. V. Pham, T. Schelfhorst, R. R. G. D. Haas, I. V. Bijnsdorp, R. Sun, L. Yue, G. Ruan, Q. Zhang, M. Hu, Y. Zhou, W. J. Van Houdt, T. Y. S. Le Large, J. Cloos, A. Wojtuszkiewicz, D. Koppers-Lalic, F. Böttger, C. Scheepbouwer, R. H. Brakenhoff, G. J. L. H. van Leenders, J. N. M. Ijzermans, J. W. M. Martens, R. D. M. Steenbergen, N. C. Grieken, S. Selvarajan, S. Mantoo, S. S. Lee, S. J. Y. Yeow, S. M. F. Alkaff, N. Xiang, Y. Sun, X. Yi, S. Dai, W. Liu, T. Lu, Z. Wu, X. Liang, M. Wang, Y. Shao, X. Zheng, K. Xu, Q. Yang, Y. Meng, C. Lu, J. Zhu, J. Zheng, B. Wang, S. Lou, Y. Dai, C. Xu, C. Yu, H. Ying, T. K. Lim, J. Wu, X. Gao, Z. Luan, X. Teng, P. Wu, S. Huang, Z. Tao, N. G. Iyer, S. Zhou, W. Shao, H. Lam, D. Ma, J. Ji, O. L. Kon, S. Zheng, R. Aebersold, C. R. Jimenez, T. Guo, DPHL: A DIA Pan-human Protein Mass Spectrometry Library for Robust Biomarker Discovery. Genomics Proteomics Bioinformatics 18, 104–119 (2020). doi:10.1016/j.gpb.2019.11.008 Medline

59. L. H. Wang, D.-Q. Li, Y. Fu, H.-P. Wang, J.-F. Zhang, Z.-F. Yuan, R.-X. Sun, R. Zeng, S.-M. He, W. Gao, pFind 2.0: A software package for peptide and protein identification via tandem mass spectrometry. Rapid Commun. Mass Spectrom. 21, 2985–2991 (2007). doi:10.1002/rcm.3173 Medline

60. R. Bai, R. Wan, C. Yan, J. Lei, Y. Shi, Structures of the fully assembled Saccharomyces cerevisiae spliceosome before activation. Science 360, 1423–1429 (2018). doi:10.1126/science.aau0325 Medline

61. J. Lei, J. Frank, Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. J. Struct. Biol. 150, 69–80 (2005). doi:10.1016/j.jsb.2005.01.002 Medline

62. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: Anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat. Methods 14, 331–332 (2017). doi:10.1038/nmeth.4193 Medline

63. K. Zhang, Gctf: Real-time CTF determination and correction. J. Struct. Biol. 193, 1–12 (2016). doi:10.1016/j.jsb.2015.11.003 Medline

64. F. Wang, H. Gong, G. Liu, M. Li, C. Yan, T. Xia, X. Li, J. Zeng, DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. J. Struct. Biol. 195, 325–336 (2016). doi:10.1016/j.jsb.2016.07.006 Medline

65. J. Zivanov, T. Nakane, B. O. Forsberg, D. Kimanius, W. J. H. Hagen, E. Lindahl, S. H. W. Scheres, New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife 7, e42166 (2018). doi:10.7554/eLife.42166 Medline

66. Z. Zhang, C. L. Will, K. Bertram, O. Dybkov, K. Hartmuth, D. E. Agafonov, R. Hofele, H. Urlaub, B. Kastner, R. Lührmann, H. Stark, Molecular architecture of the human 17S U2 snRNP. Nature 583, 310–313 (2020). doi:10.1038/s41586-020-2344-3 Medline

67. D. E. Agafonov, B. Kastner, O. Dybkov, R. V. Hofele, W.-T. Liu, H. Urlaub, R. Lührmann, H. Stark, Molecular architecture of the human U4/U6.U5 tri-snRNP. Science 351, 1416–1420 (2016). doi:10.1126/science.aad2085 Medline

68. X. Zhan, C. Yan, X. Zhang, J. Lei, Y. Shi, Structures of the human pre-catalytic spliceosome and its precursor spliceosome. Cell Res. 28, 1129–1140 (2018). doi:10.1038/s41422-018-0094-7 Medline

69. R. Wan, R. Bai, C. Yan, J. Lei, Y. Shi, Structures of the Catalytically Activated Yeast Spliceosome Reveal the Mechanism of Branching. Cell 177, 339–351.e13 (2019). doi:10.1016/j.cell.2019.02.006 Medline

70. S. Chen, G. McMullan, A. R. Faruqi, G. N. Murshudov, J. M. Short, S. H. W. Scheres, R. Henderson, High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. Ultramicroscopy 135, 24–35 (2013). doi:10.1016/j.ultramic.2013.06.004 Medline

71. P. B. Rosenthal, R. Henderson, Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. J. Mol. Biol. 333, 721–745 (2003). doi:10.1016/j.jmb.2003.07.013 Medline

72. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. Acta Cryst. D60, 2126–2132 (2004). doi:10.1107/S0907444904019158 Medline

73. C. Netter, G. Weber, H. Benecke, M. C. Wahl, Functional stabilization of an RNA recognition motif by a noncanonical N-terminal expansion. RNA 15, 1305–1313 (2009). doi:10.1261/rna.1359909 Medline

74. K. Xu, Z. Wang, J. Shi, H. Li, Q. C. Zhang, A2-Net: Molecular Structure Estimation from Cryo-EM Density Volumes. Proc. AAAI Conf. Artif. Intell. 33, 1230–1237 (2019); https://doi.org/10.1609/aaai.v33i01.33011230.

75. J. Yang, I. Anishchenko, H. Park, Z. Peng, S. Ovchinnikov, D. Baker, Improved protein structure prediction using predicted interresidue orientations. Proc. Natl. Acad. Sci. U.S.A. 117, 1496–1503 (2020). doi:10.1073/pnas.1914677117 Medline

76. P. V. Afonine, B. K. Poon, R. J. Read, O. V. Sobolev, T. C. Terwilliger, A. Urzhumtsev, P. D. Adams, Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Cryst. D74, 531–544 (2018). doi:10.1107/S2059798318006551 Medline

77. A. Amunts, A. Brown, X. C. Bai, J. L. Llácer, T. Hussain, P. Emsley, F. Long, G. Murshudov, S. H. W. Scheres, V. Ramakrishnan, Structure of the yeast mitochondrial large ribosomal subunit. Science 343, 1485–1489 (2014). doi:10.1126/science.1249410 Medline

78. I. W. Davis, A. Leaver-Fay, V. B. Chen, J. N. Block, G. J. Kapral, X. Wang, L. W. Murray, W. B. Arendall 3rd, J. Snoeyink, J. S. Richardson, D. C. Richardson, MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 35, W375–W383 (2007). doi:10.1093/nar/gkm216 Medline

79. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612 (2004). doi:10.1002/jcc.20084 Medline

80. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 27, 14–25 (2018). doi:10.1002/pro.3235 Medline

81. W. L. DeLano, The PyMOL Molecular Graphics System (2002); https://pymol.org/2/.

82. V. M. Howell, J. M. Jones, S. K. Bergren, L. Li, A. C. Billi, M. R. Avenarius, M. H. Meisler, Evidence for a direct role of the disease modifier SCNM1 in splicing. Hum. Mol. Genet. 16, 2506–2516 (2007). doi:10.1093/hmg/ddm206 Medline

SUPPLEMENTARY MATERIALS

science.sciencemag.org/cgi/content/full/science.abg0879/DC1
Figs. S1 to S11
Tables S1 to S3
MDAR Reproducibility Checklist

9 December 2020; accepted 18 January 2021
Published online 28 January 2021
10.1126/science.abg0879

A

U5 snRNP
- U5 snRNA
- U5 Sm ring
- PRP8
- BRR2
- SNU114
- U5-40K

U6atac snRNA

U12 snRNP		
Core	SF3B	
U12 snRNA	SF3b155	SCNM1
U12 Sm ring	SF3b145	
	SF3b130	SF3A
U2-A'	SF3b49	SF3a120
U2-B''	SF3b14a	SF3a66
	SF3b14b	SF3a60
	SF3b10	

RES
- SNIP1
- RBMX2
- BUD13

PPIs
- PPIL2
- CWC27
- PPIL1
- CypE

NTC
- CDC5L
- SYF3
- PRPF19
- SPF27
- SYF1

NTR
- CWC15
- SKIP
- PLRG1
- BUD31
- RBM22

pre-mRNA

Splicing Factors
- PRP2
- RNF113A
- SRm160
- CRIPT
- ARMC7
- GPKOW
- CWC22
- SRm300
- RBM48
- PRP17

B

minor B^{act} complex

major B^{act} complex

Fig. 1. Cryo-EM structure of the activated human minor spliceosome (B^act complex). (A) Overall structure of the human minor B^act complex. Two surface views are shown. Spliceosomal components are tabulated below the images. Proteins that are found only in the structure of the major, but not minor, B^act complex are indicated by dashed gray boxes. Compared to the major B^act complex (*35, 36*), the minor spliceosome contains four new proteins ARMC7, CRIPT, RBM48, and SCNM1. Although PPIL2 (colored deep purple) is also a component of the major B^act complex (*35*), its atomic model is reported here for the first time. (B) Overall structural comparison between the minor and major B^act complexes. The orientation and scale are the same for the two complexes. Identical protein components are shown in the background with faded colors. RNA elements and protein components that differ in the two complexes are color-coded. Structural images in this figure were prepared using ChimeraX (*80*).
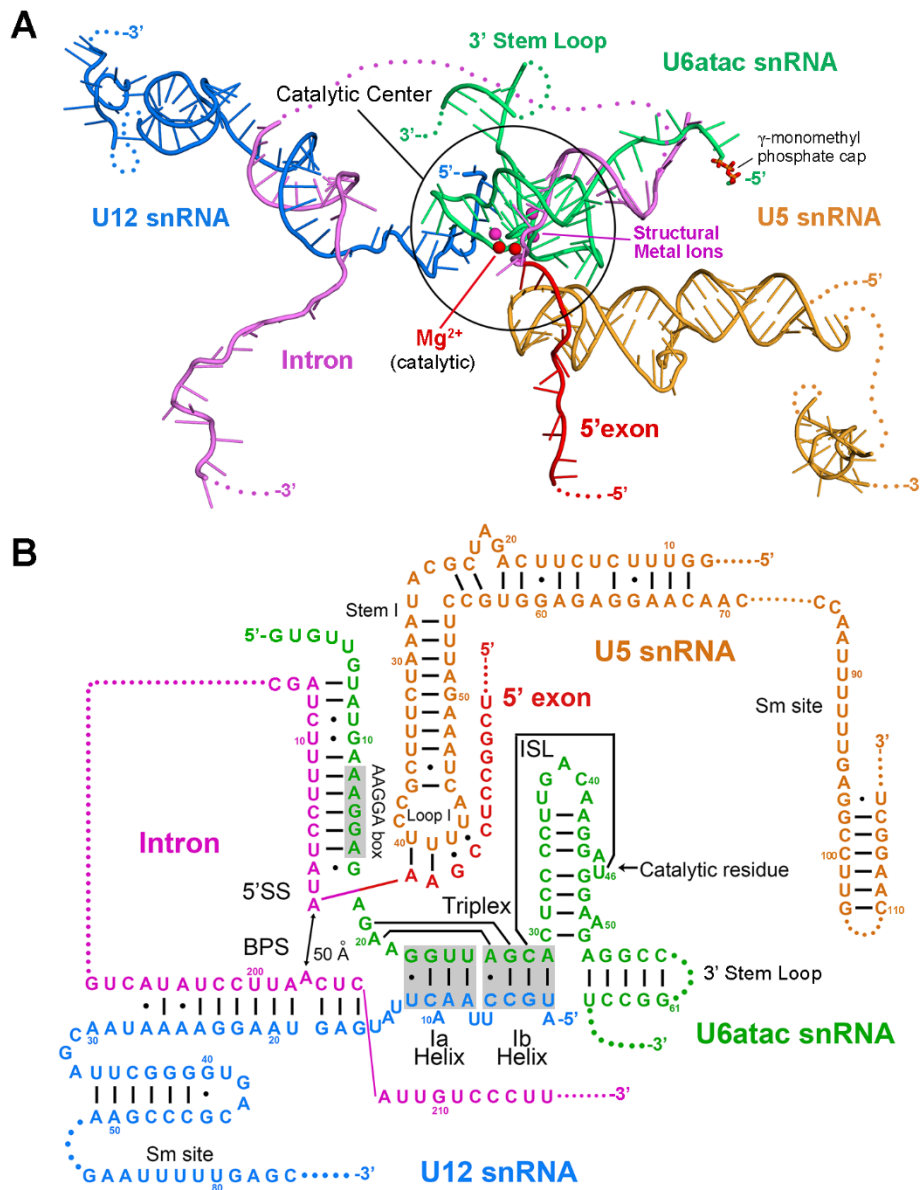
Fig. 2. Organization and structure of the RNA elements in the human minor B^act complex. (A) Overall structure of the RNA elements. U12, U5 and U6atac snRNAs are colored marine, orange and green, respectively. The 5′-exon and intron of MINX-U12Δ are colored red and violet, respectively. The γ-monomethyl phosphate cap of U6atac snRNA is shown in stick. The catalytic center comprises the ISL of U6atac snRNA, helix I of the U12/U6atac duplex and loop I of U5 snRNA. (B) Summary of the base-pairing interactions among the RNA elements. Helix Ib of the U12/U6atac duplex forms a catalytic triplex with three nucleotides of U6atac. Canonical Watson-Crick and noncanonical base-pairing interactions are identified by solid lines and dots, respectively. Structural images in this and following figures were prepared using PyMol (81).

A

U2/U6 Helix II

minor Bᵃᶜᵗ
major Bᵃᶜᵗ

U12 snRNA

3' Stem Loop

U6atac snRNA

U6 snRNA 5' Stem Loop

Intron

5' exon

U5 snRNA

B

A18 G17 A16 G15
            G14
         A3   A13
      U4   A12
   C5   U9
U9   U10
   A8   C11
U6atac
snRNA

Intron

5'-A₈ U G A G A A A G G A G A
     · · | | | | | | | | · ·
3'- U C U U U U C C U A U A₁
              ⌐──── 5'SS ────⌐

A45 G44 A43 C42
A3   A41
U4   U40
G5   A39
C37   C9
A35   A10
U6 snRNA   A11

5'-A₃₅ A C G A U A C A G A
      · | | | | | · · | | ·
3'- G A A C U A U G U A U G₁
              ⌐──── 5'SS ────⌐

C

A24 G23
         A21
U   A
   C   A
G18   C
A17   U
G16   C
U12 snRNA

BPS
         A
5'- U C C U U A C U C
    | | | | | | · | |
3'- A G G A A U G A G₁₆

U39 A38
         G36
U   G
   U   A
G33   C
U32   C
G31   U
U2 snRNA

BPS
         A
5'- U U A C U G C C U
    · · | | | · | | |
3'- U A U G A U G U G₃₁

D

U6atac snRNA

ISL

AAGGA box

Structural Metal Ions

M2
M1

U12 snRNA

U5 snRNA

5' exon

Loop I

E   minor Bᵃᶜᵗ

A45

U46   G44

                              A1
G27   M1
         M2   3.3 Å
A26   3.9 Å
         4 Å   A-1

F   major Bᵃᶜᵗ

A73

U74   G72

G54   G1
         M2
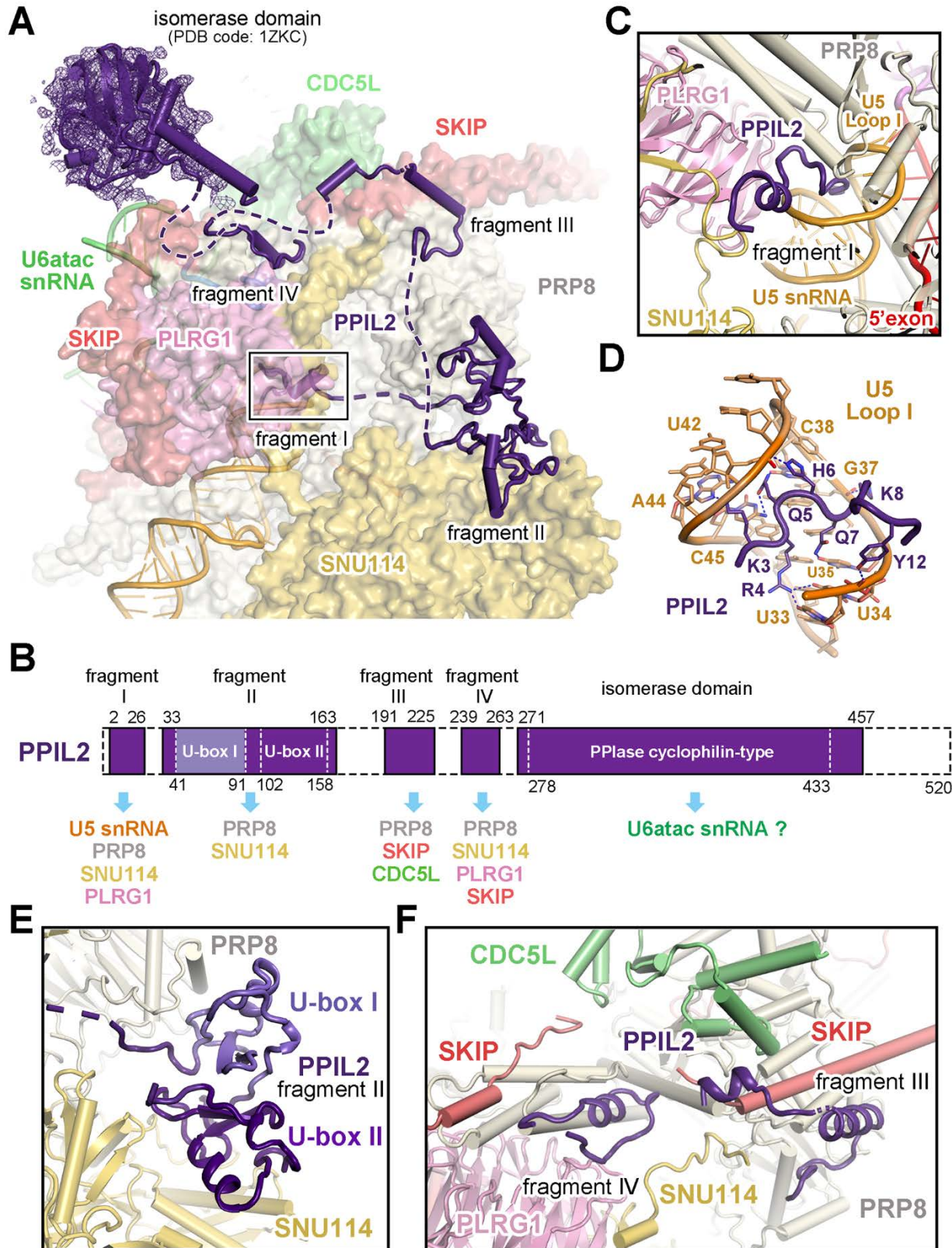A53   3.9 Å
                 G-1

Fig. 3. Conservation and variation of the RNA structures between the minor and major B^{act} complexes.
(**A**) Overall structural comparison of the RNA elements between the minor and major B^{act} complexes. The RNA elements from the minor spliceosome are color-coded, whereas those from the major spliceosome are shown in gray. Compared to U6 snRNA (*35*), U6atac lacks the 5′-stem loop, cannot form helix II with U12, but forms a unique 3′-stem loop. (**B**) Recognition of 5′SS by U6atac (left panel) involves more base-pairing interactions than that by U6 (*35*) (right panel). (**C**) Recognition of BPS by U12 (left panel) involves more base-pairing interactions than that by U2 (*35*) (right panel). (**D**) The active site conformation in the minor B^{act} complex is nearly identical to that in the major B^{act} complex (*35*). Shown here is an overlay of the active site RNA elements from both complexes. (**E**) Coordination of the catalytic metals in the minor B^{act} complex. Unlike the major B^{act} complex (*35*), both catalytic metals M1 and M2 are loaded in the minor B^{act} complex. (**F**) Coordination of the catalytic metals in the major B^{act} complex [PDB code 6FF7 (*35*)]. The orientation is the same as that in panel (E). M1 is yet to be loaded.
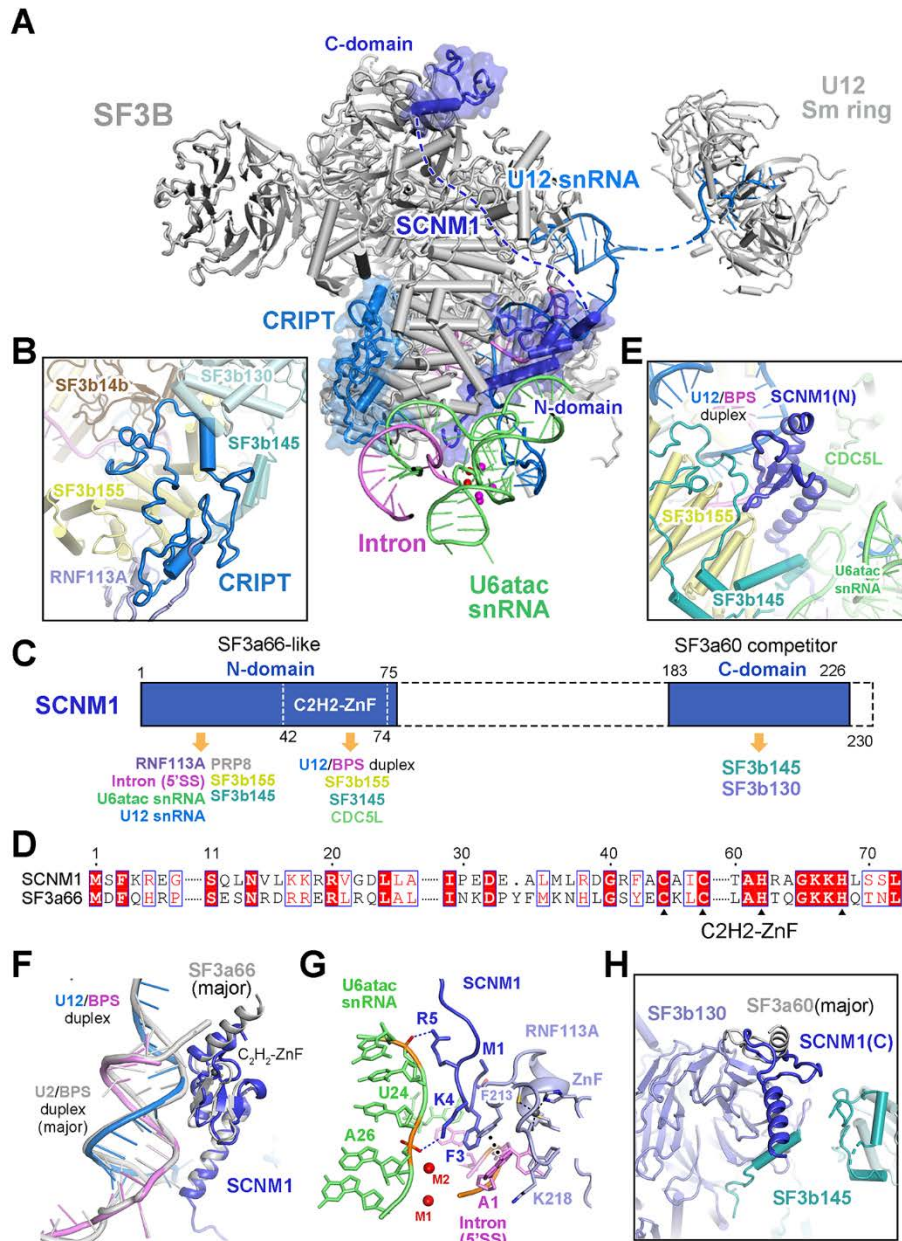
Fig. 4. Structure of U12 snRNP in the human minor B^{act} complex. (**A**) Overall structure of U12 snRNP. U12 snRNP consists of the core (U12 snRNA and Sm ring), the SF3b complex, and the newly identified component SCNM1 (*82*). A newly identified splicing factor CRIPT is also shown. (**B**) A close-up view on the cysteine-rich PDZ-binding protein CRIPT and its surrounding proteins. CRIPT interacts with SF3b155, SF3b145, SF3b14b and the splicing factor RNF113A. (**C**) A schematic diagram of the sequence features of SCNM1. The interacting proteins and RNA elements are tabulated below the sequences. (**D**) Sequence alignment between the N-domains of SCNM1 and SF3a66. Conserved sequences are boxed and identical residues are shaded red. (**E**) The N-domain of SCNM1 interacts with the BPS/U12 duplex and the proteins SF3b145, SF3b155 and CDC5L. (**F**) The C2H2-type zinc finger of SCNM1 closely resembles that of SF3a66. Shown here is a structural overlay. The RNA and protein components are color-coded in the minor spliceosome and gray in the major spliceosome (*36*). (**G**) A close-up view on the N terminus of SCNM1 and its interactions with surrounding proteins and RNA elements. The N terminus of SCNM1 is inserted between RNF113A and U6atac. (**H**) The C-domain of SCNM1 interacts with SF3b130 and SF3b145 in the minor B^{act} complex. The spatial location of the SCNM1 C-domain overlaps with that of SF3a60 in the major B^{act} complex (*36*).

A isomerase domain (PDB code: 1ZKC)

CDC5L

SKIP

fragment III

U6atac snRNA

SKIP

PLRG1

fragment IV

PPIL2

PRP8

fragment I

fragment II

SNU114

C

PRP8

PLRG1

PPIL2

U5 Loop I

fragment I

SNU114

U5 snRNA

5'exon

D U5 Loop I

U42

C38

G37

H6

K8

A44

Q5

Q7

C45

K3

U35

Y12

R4

PPIL2

U33

U34

B

fragment I   fragment II   fragment III   fragment IV   isomerase domain

PPIL2

2 26   33   163   191 225   239 263   271   457

| U-box I | U-box II |

41   91 102   158

278   433   520

PPIase cyclophilin-type

U5 snRNA   PRP8   PRP8   PRP8   U6atac snRNA ?
PRP8   SNU114 SKIP   SNU114
SNU114   CDC5L   PLRG1
PLRG1   SKIP

E

PRP8

U-box I

PPIL2 fragment II

U-box II

SNU114

F

CDC5L

SKIP

PPIL2

SKIP

fragment III

fragment IV

SNU114

PLRG1

PRP8

Fig. 5. Multi-faceted roles of PPIL2 in the human minor B^{act} complex. (A) PPIL2 forms at least five distinct interfaces with the minor B^{act} complex and spans a distance of over 100 Å. The quality of the EM density allows unambiguous identification of four discrete fragments I through IV in the N-terminal half of PPIL2, but not for the C-terminal half. (B) A schematic diagram of the sequence features of PPIL2. PPIL2 comprises four fragments and a PPIase domain. The interacting proteins and RNA elements are tabulated below the sequences. (C) A close-up view on fragment I of PPIL2 and its surrounding environment. This fragment interacts with loop I of U5 snRNA, the N-domain of PRP8, the WD40 domain of PLRG1 and the N terminus of SNU114. (D) A close-up view on the interface between the N terminus of fragment I and loop I of U5 snRNA. The N terminus of PPIL2 is inserted into the groove on the back side of the duplex between loop I and 5′-exon. (E) A close-up view on fragment II of PPIL2 and its surrounding environment. Unexpectedly, this fragment contains two U-box E3 ubiquitin-ligase domains: one as predicted (46) (U-box I, light purple), and the other previously unknown (U-box II, deep purple). U-box I contacts Prp8 and U-box II interacts with SNU114. (F) A close-up view on fragments III and IV of PPIL2. These two fragments adopt an extend conformation and interact with SKIP, PRP8, CDC5L and SNU114.
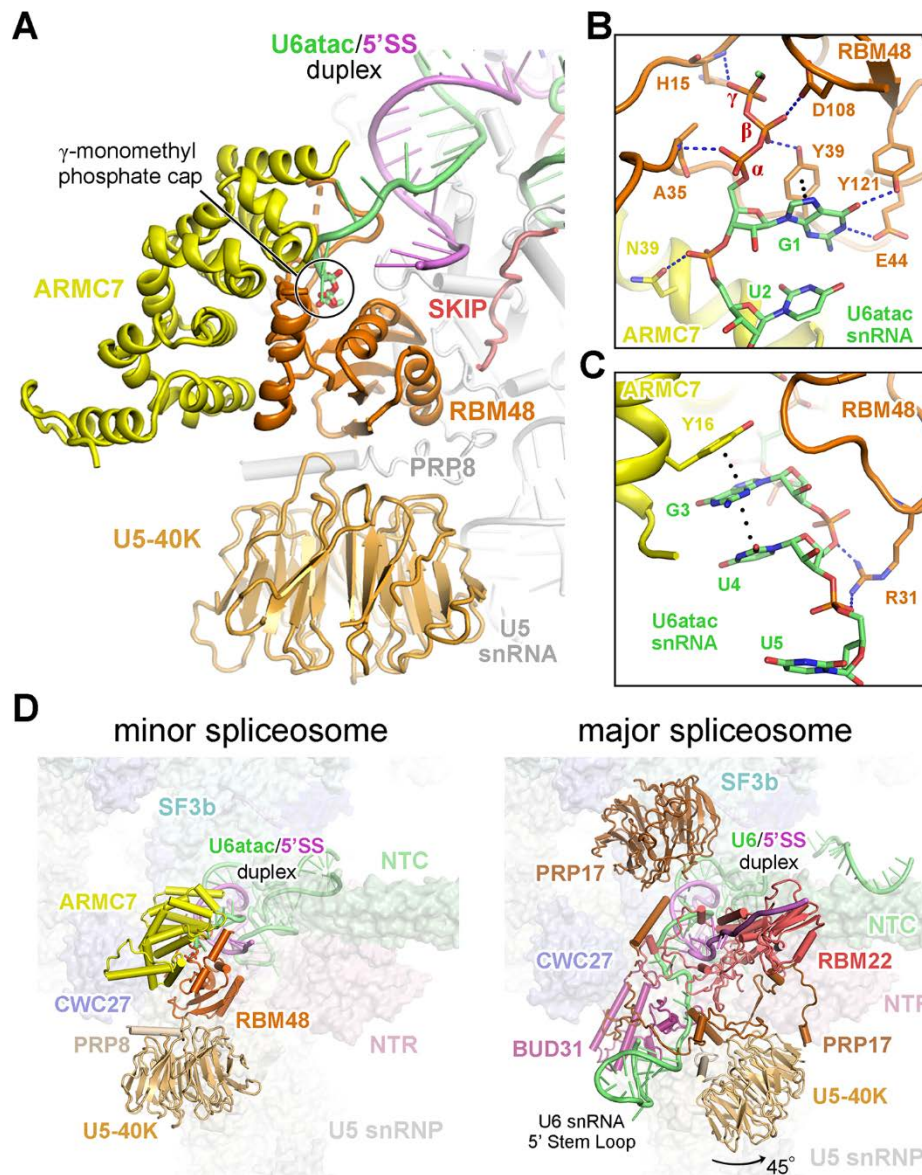
Fig. 6. The RBM48-ARMC7 complex binds the 5′-end of U6atac snRNA in the minor B^{act} complex. (**A**) The RBM48-ARMC7 complex binds the 5′-end of U6atac snRNA. The γ-monomethyl phosphate cap at the 5′-end of U6atac is shown in stick. (**B**) A close-up view on the recognition of the guanine nucleotide (G1) at the 5′-end of U6atac. G1 is recognized by RBM48. (**C**) A close-up view on the stacking interactions among the aromatic side chain of Tyr17 from ARMC7 and the bases of G3 and U4 from U6atac. (**D**) Variations in the coordination of the 5′-end sequences from U6 and U6atac. In the major B^{act} complex [PDB code 5Z56 (*36*)] (right panel), the 5′-end sequences of U6 form a stem loop. RBM22, BUD13 and PRP17 recognize the 5′-stem loop and its downstream sequences. U6atac lacks the 5′-stem loop and its 5′-end is bound by the RBM48-ARMC7 complex (left panel).
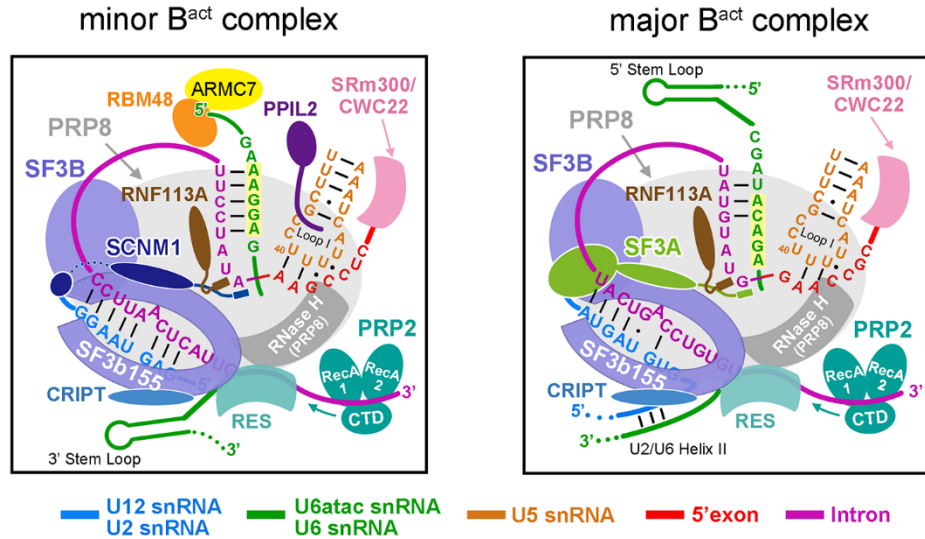
Fig. 7. Structural differences between the human minor and major B^act complexes. Shown here are cartoon diagrams of the structures around the active site of minor and major B^act complexes. In the minor spliceosome (left panel), the last five nucleotides UCCUU of 5'SS are recognized by the AAGGA box of U6atac snRNA. The first adenine nucleobase in 5'SS is sandwiched by Phe213 and Lys218 from RNF113A. The N terminus of SCNM1 inserts into the active site, stabilizing 5'SS recognition by RNF113A. In the major B^act complex (right panel), 5'SS recognition by U6 snRNA involves much fewer base pairs, and recognition of the first guanine nucleotide by RNF113A is stabilized by SF3a66. In addition, SF3a60 in the major spliceosome is replaced by the C terminus of SCNM1 in the minor B^act complex. In these regards, SCNM1 in the minor spliceosome mimics the SF3a complex in the major spliceosome. This is likely true not only in the minor B^act complex but also during spliceosome assembly. In the minor spliceosome, the RBM48-ARMC7 complex binds the 5'-end of U6atac snRNA, which lacks the 5'-stem loop compared to U6 snRNA in the major spliceosome. The N terminus of PPIL2 binds loop I of U5 snRNA, interacts with several components from U5 snRNP, NTC and NTR in the minor B^act complex. In contrast, the role of PPIL2 in the major B^act complex is yet to be defined. Structural features of the minor spliceosome that are similar to those of the major spliceosome are not discussed here. For example, in both cases, the HEAT repeat protein SF3b155 serves as the central scaffold of U2 snRNP and wraps around the U2/BPS duplex.

Science

Structure of the activated human minor spliceosome

Rui Bai, Ruixue Wan, Lin Wang, Kui Xu, Qiangfeng Zhang, Jianlin Lei and Yigong Shi

published online January 28, 2021