Supplementary Materials for

Molecular mechanism of the catalysis for U12-type splicing by the human minor spliceosome

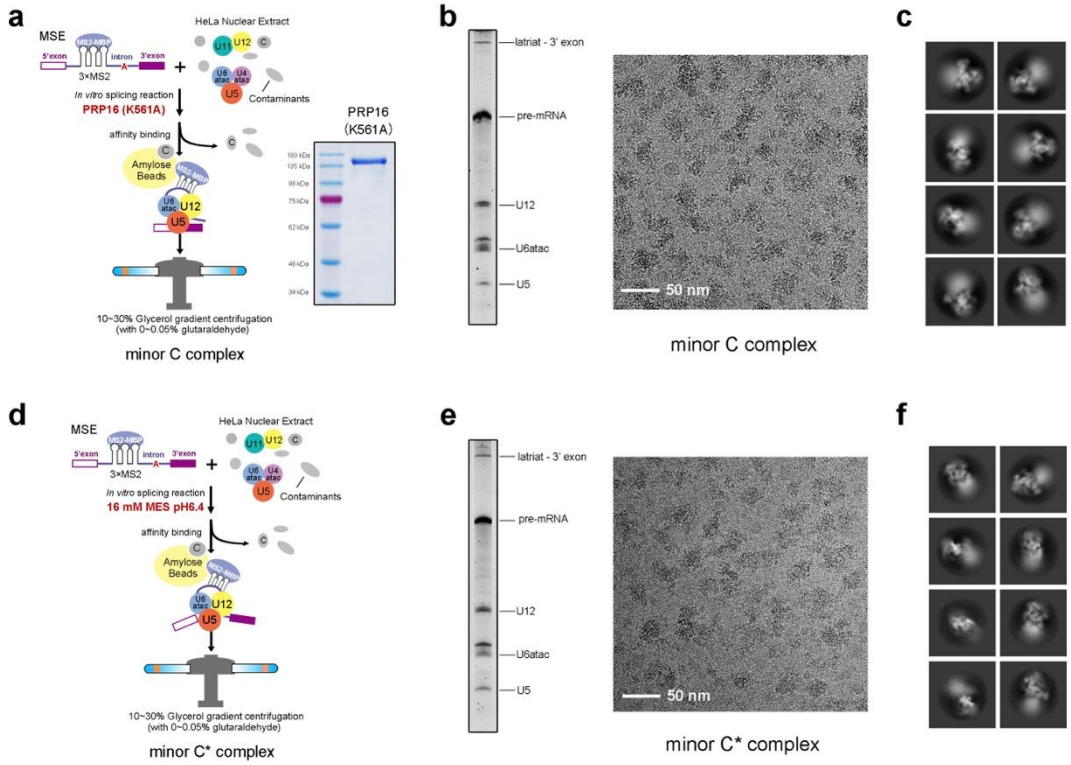
Rui Bai[1,2,3,4], Han Guo[1,2,3,4], Rui Sun[1,2,3], Yi Zhu[1,2,3], Yigong Shi[1,2,3] and Ruixue Wan[1,2,3*]

Corresponding author: wanruixue@westlake.edu.cn

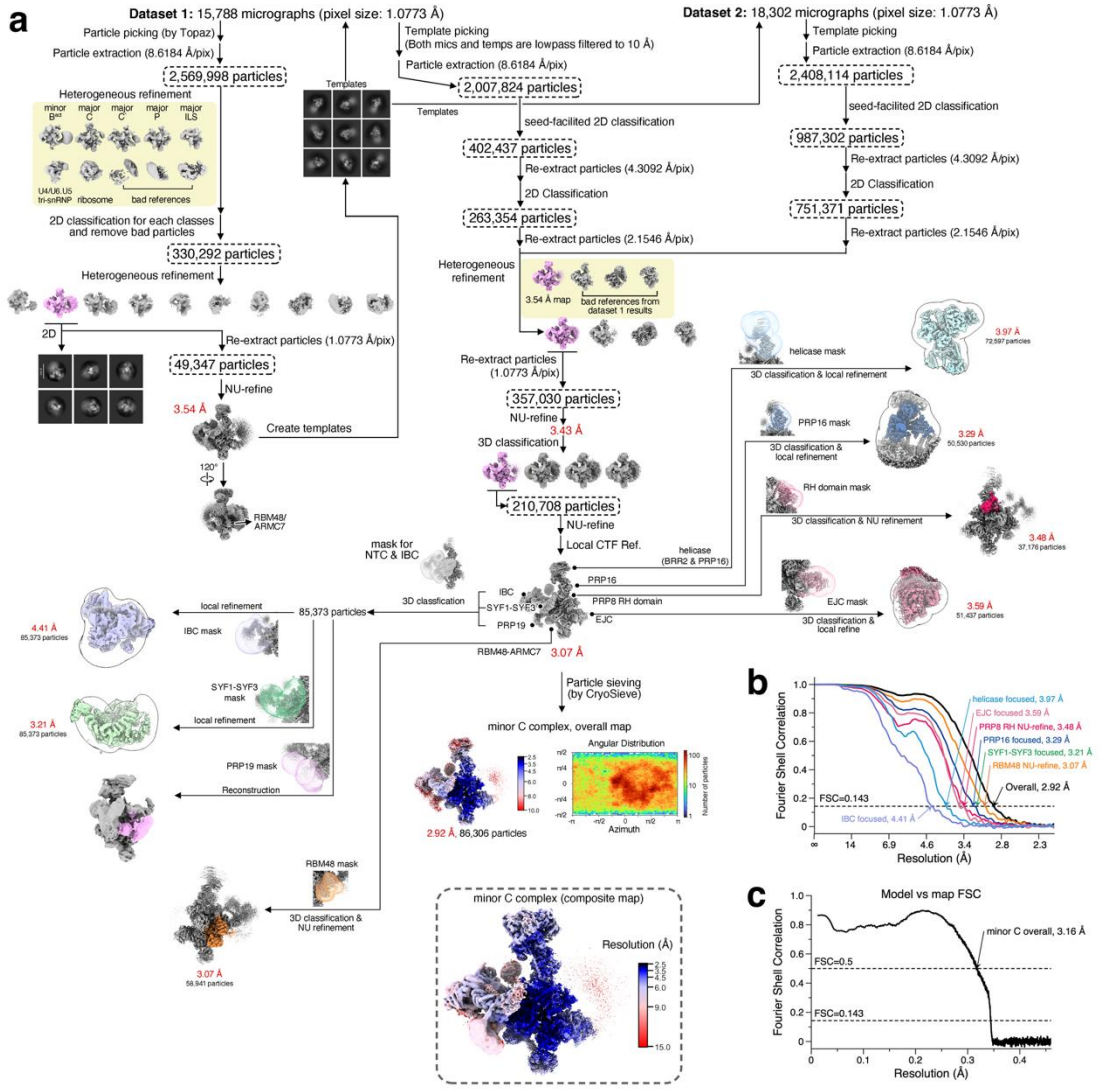
The PDF file includes:

Supplementary Figure S1 | Design and optimization of the U12-type pre-mRNA substrates. **a,** A schematic diagram of the U12-type pre-mRNA substrates. The MINX-U12 substrate[1] contains a 56-nt 5′ exon, a 228-nt intron and a 40-nt 3′ exon. The sequences of 5′SS, BPS and 3′SS are shown. Three tandem MS2-binding RNA aptamers are placed between 5′SS and BPS. MINXE-U12 (MINX with ENH1-U12) has an enhancer element ENH1 at the 3′ end of 3′ exon. MS-U12 (MINX and SCN4A-U12) contains a 57-nt 5′ exon, a 218-nt intron, and a 43-nt 3′ exon, with modified BPS, 3′SS, and 3′ exon. ENH1 was added to MS-U12 to generate MSE-U12 (MINX and SCN4A with ENH1-U12). **b,** Purified U12-type pre-mRNAs after *in vitro* transcription are shown on a urea PAGE gel. **c,** Splicing of MINX-U12 and MINXE-U12. Shown here is the RT-PCR result of an *in vitro* splicing assay. To inhibit the major or minor spliceosome, anti-sense DNA oligonucleotides that are complementary to U1/U2/U6 or U11/U12/U6atac are included in groups 2/5 or 3/6, respectively. Compared to the control (groups 1 & 4), MINX-U12 and MINXE-U12 are indifferent to elimination of U1, U2 and U6 snRNAs (groups 2 & 5), but sensitive to elimination of U11, U12 and U6atac snRNAs (groups 3 & 6). Compared to MINX-U12 (groups 1 & 2), enhancer element ENH1 can obviously improve the splicing

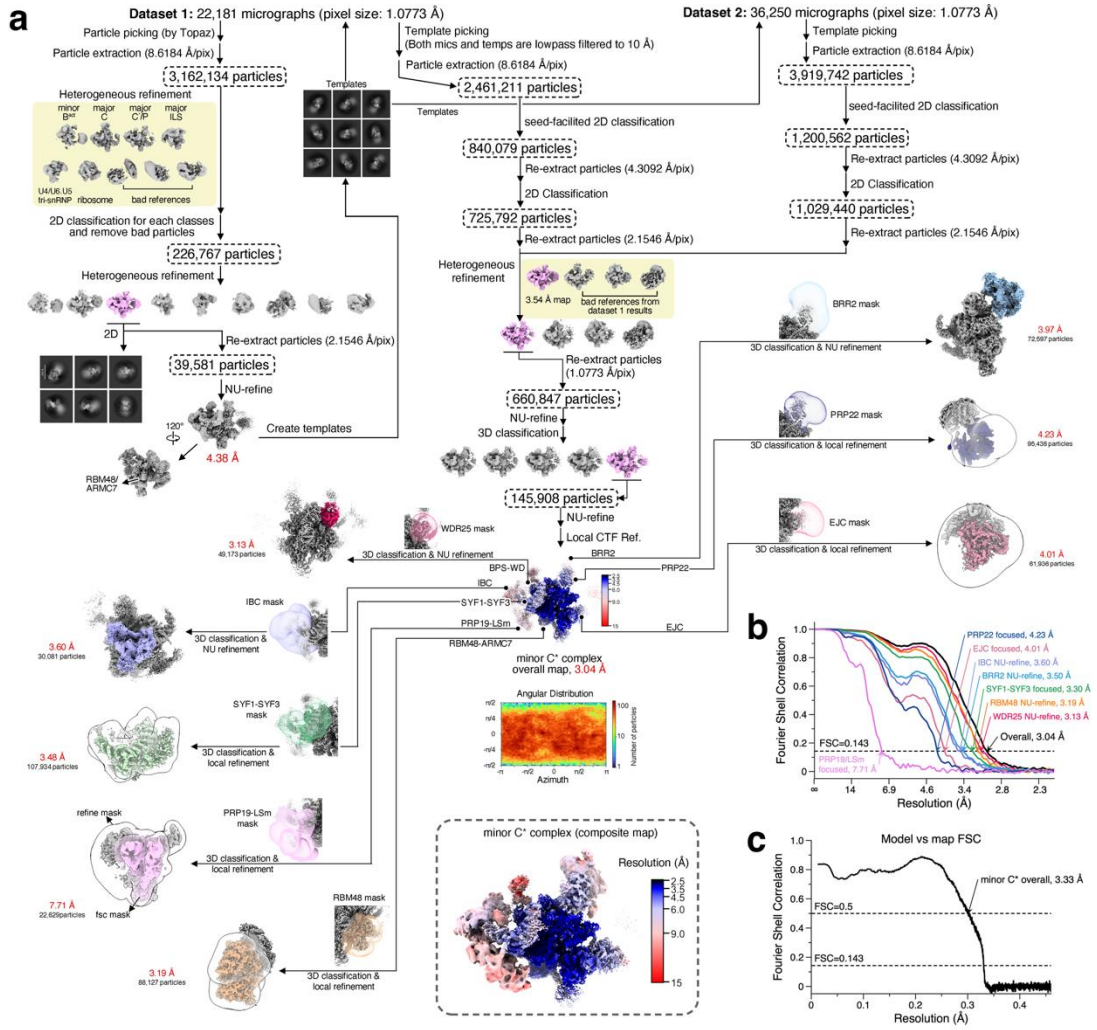
activity (groups 4 & 5). **d,** Splicing of MS-U12 and MSE-U12. MS-U12 and MSE-U12 are only sensitive to elimination of U11, U12 and U6atac snRNAs (groups 7-13). Compared to MINX-U12 (groups 1 & 2), MS-U12 (groups 7 & 8) displays improved splicing efficiency. The splicing efficiency is further improved for MSE-U12 (groups 10 & 11).



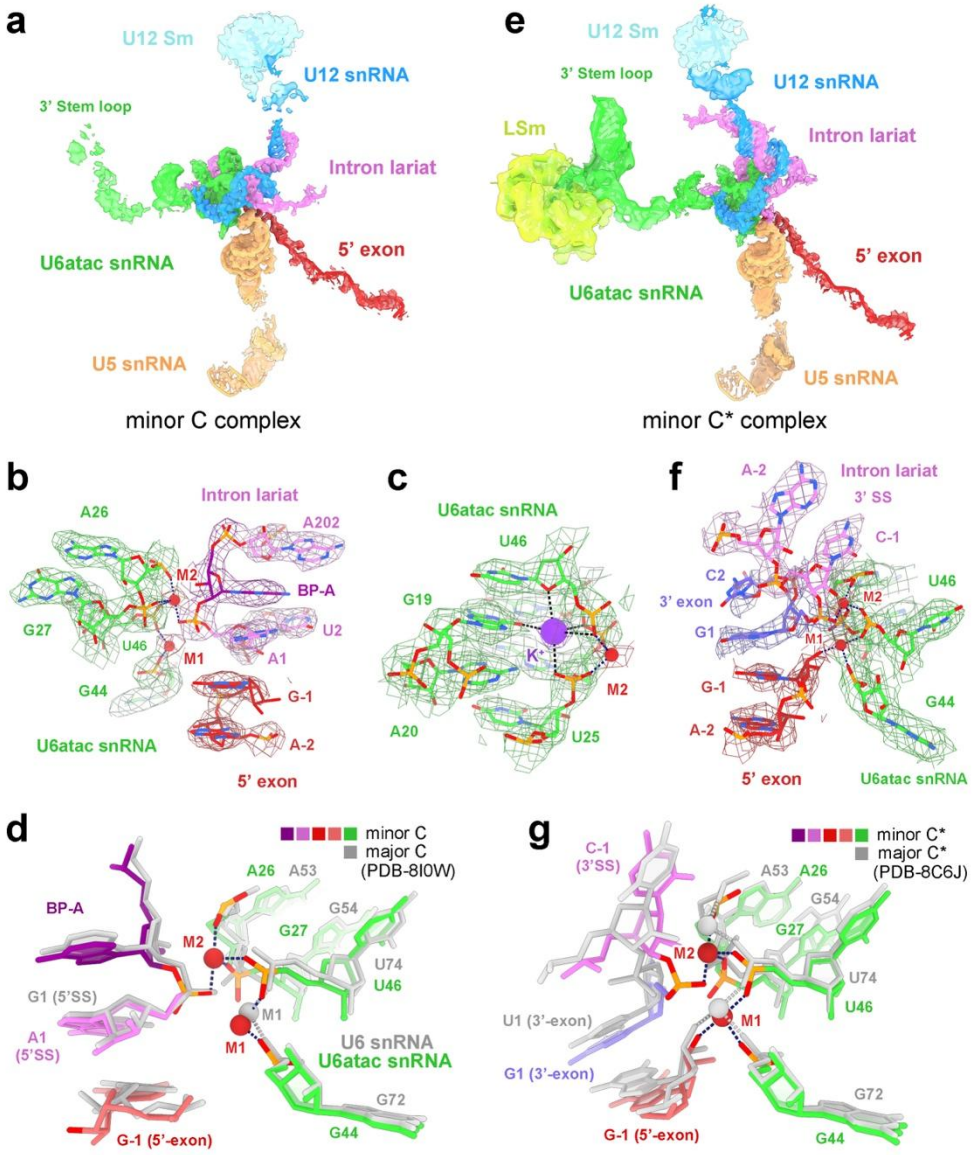
Supplementary Figure S2 | Purification and EM analysis of the human minor C and C* complex. **a,** A schematic diagram of purification for the human minor C complex. The helicase-defective PRP16 mutant (PRP16 K561A) was applied in the splicing reaction, aiming to stall the assembly of the minor spliceosome at the C complex stage. Coomassie-stained SDS-PAGE gel of the purified recombinant protein PRP16 (K561A) is shown on the right. **b,** The purified human minor C complex was visualized on a urea PAGE gel (left) and imaged under cryogenic condition on a Titan Krios microscope with a K3 detector (right). Scale bar, 50 nm. The sample contains U12, U6atac and U5 snRNAs, intron lariat-3′ exon intermediate (left). Unspliced pre-mRNA are also present. **c,** 2D class averages of the human minor C complex particles. **d,** A schematic diagram of purification for the human minor C* complex. A mildly acidic condition (16 mM MES, pH 6.4) was used. **e,** The purified human minor C* complex was visualized on a urea PAGE gel (left) and imaged under cryo-EM (right). Scale bar, 50 nm. **f,** 2D class averages of the human minor C* complex particles.



Supplementary Figure S3 | Cryo-EM data processing and analysis of the human minor C complex. **a,** A flow chart of cryo-EM data processing for the human minor C complex. Please refer to Methods and Supplementary Table S2 for details. **b,** The Fourier shell correlations (FSCs) of the overall map, seven local maps of the human minor C complex. The resolution was estimated on the basis of the FSC criterion of 0.143[2]. **c,** Model versus map FSC curves for the human minor C complex. The FSC curves were calculated using PHENIX[3]. The entire model of the human minor C complex was refined against the 2.92-Å overall map. All images of the EM maps were created using UCSF ChimeraX[4].

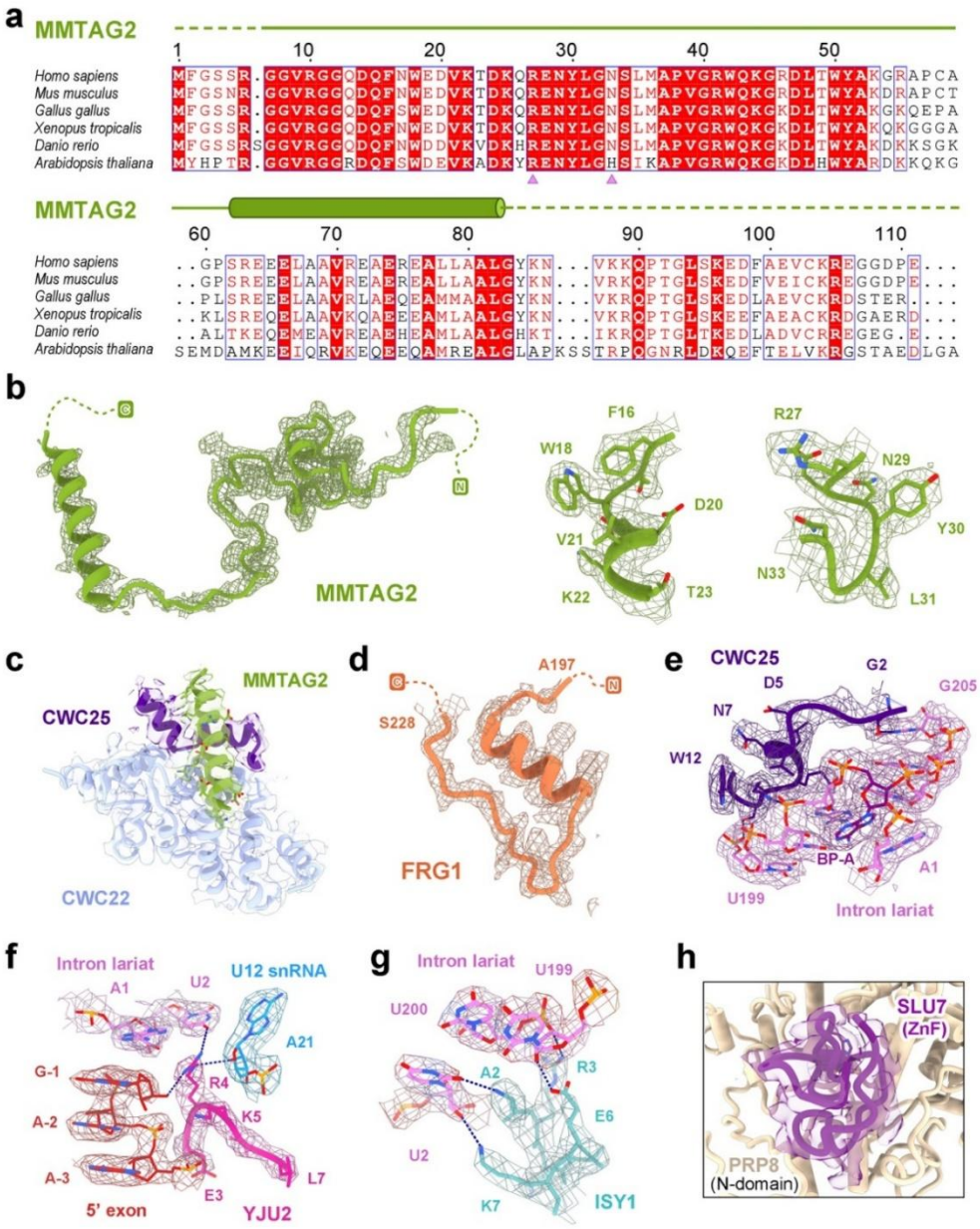


Supplementary Figure S4 | Cryo-EM data processing and analysis of the human minor C* complex. **a,** A flow chart of cryo-EM data processing for the human minor C* complex. Please refer to Methods and Supplementary Table S3 for details. **b,** The Fourier shell correlations (FSCs) of the overall map, eight local maps of the human minor C* complex. The resolution was estimated on the basis of the FSC criterion of 0.143[2]. **c,** Model versus map FSC curves for the human minor C* complex. The FSC curves were calculated using PHENIX[3]. The entire model of the human minor C* complex was refined against the 3.04-Å overall map.

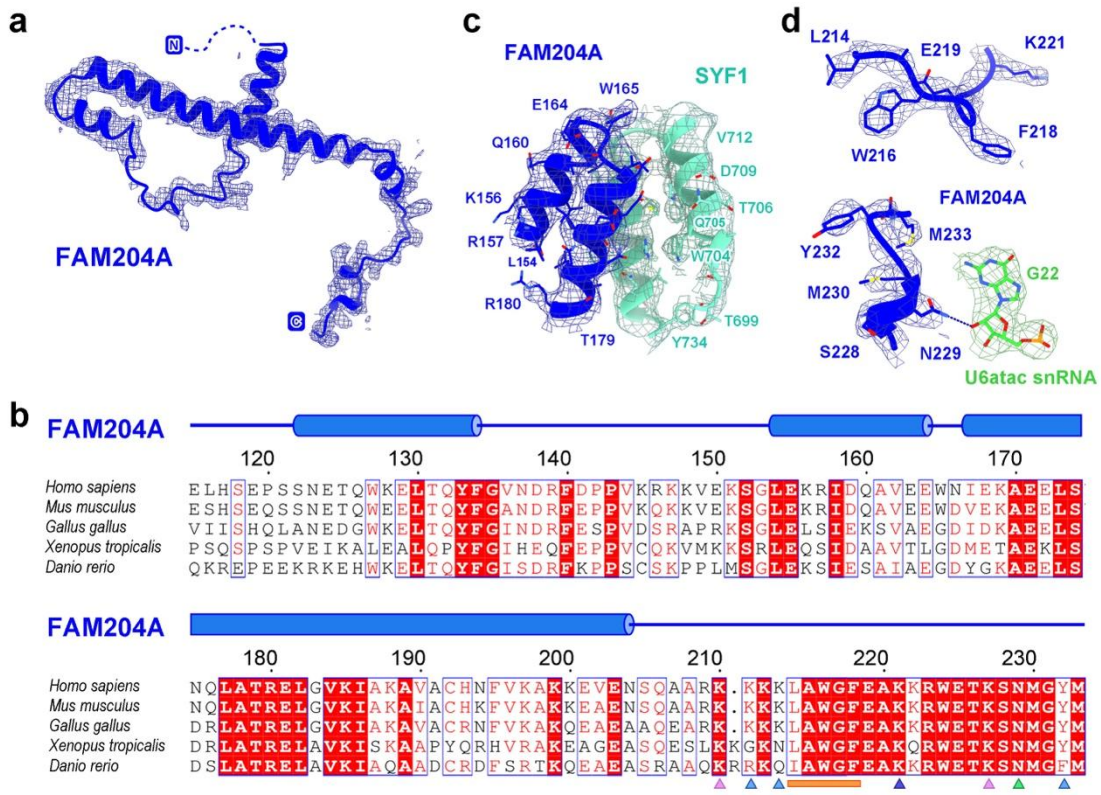


Supplementary Figure S5 | The EM density maps for the RNA elements in the human minor C and C* complexes. **a,** The EM density map of all the RNA elements in the human minor C complex. U12, U5, and U6atac snRNAs are colored marine, light orange, and green, respectively. The 5′ exon and intron of MSE-U12 are colored red and violet, respectively. The EM density map for U12 Sm ring is displayed and colored light blue. **b,** The EM density map for the RNA elements within the splicing active site in the minor C complex. Following the branching reaction, the 3′-5′ phosphodiester bond at the exon-intron junction is cleaved, and the 2′-oxygen of the branch point adenosine forms a covalent bond with the phosphorus atom of the guanine nucleotide at the 5′ end of the 5′SS. Two catalytic metal ions (M1 and M2) are colored red. **c,** The EM density map for the RNA elements of U6atac snRNA that coordinate the potassium ion and the catalytic metal ion M2 in the minor C complex. The potassium and magnesium coordination bonds are indicated as black and blue dashed lines, respectively. **d,** Comparison of the metal coordination between the human minor C complex and the major C complex[5]. M2 is not resolved in the major C complex. **e,** The EM density map of all the RNA elements in the human minor C* complex. The RNA elements are colored the same as in panel a. The

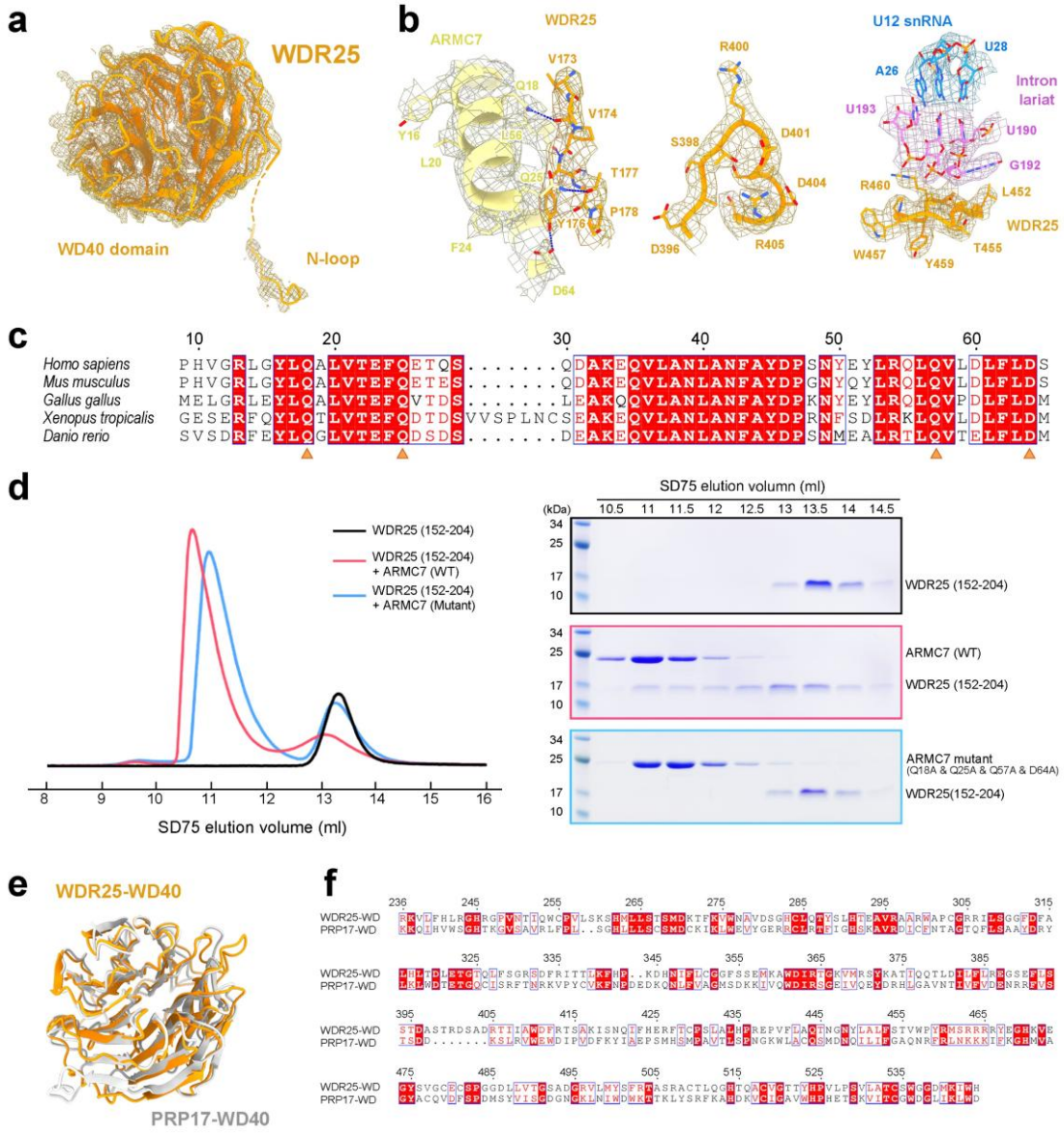
EM density maps for U12 Sm ring and U6atac LSm ring are colored light blue and yellow green, respectively. In the minor C* complex, the EM densities of the stem loop in the middle portion of U12 snRNA, the 3′ stem loop of U6atac snRNA, and the LSm complex bound to the 3′ end of U6atac snRNA are clearly distinguishable. **f,** The EM density map for the RNA elements within the splicing active site in the minor C* complex. The 3′SS-3′ exon is positioned in the active site of the minor C* complex, with the free 3′-OH of the 5′ exon poised for exon ligation. The 3′ exon sequences are colored slate. **g,** Comparison of metal coordination between the human minor C* complex and the major C* complex[6].



Supplementary Figure S6 | Representative EM densities for select splicing factors in the human minor C complex. **a,** Sequence alignment of MMTAG2 from human, mouse, chicken, frog, zebrafish and plant. Secondary structural elements are indicated above the alignment. Identical residues are shaded red. Conserved residues are colored red in blue boxes. **b,** The overall EM density map for the newly identified protein MMTAG2 and densities for two fragments at its N-terminus. **c,** The EM density map for the middle helix of MMTAG2 and its interacting motifs: the central helix of CWC25 and the MA3 domain of CWC22. **d,** The EM density map for the FRG1 fragment (residues 197-228) that inserts between the Linker and Endo domains of PRP8. **e,** The EM density map for the N-terminus of CWC25 and its interacting RNA elements. **f,** The EM density map for the N-terminus of YJU2 and its interacting RNA elements. **g,** The EM density map for the N-terminus of ISY1 and its interacting RNA elements. **h,** The EM density map for the zinc-finger domain (ZnF) of SLU7 in the minor C complex.

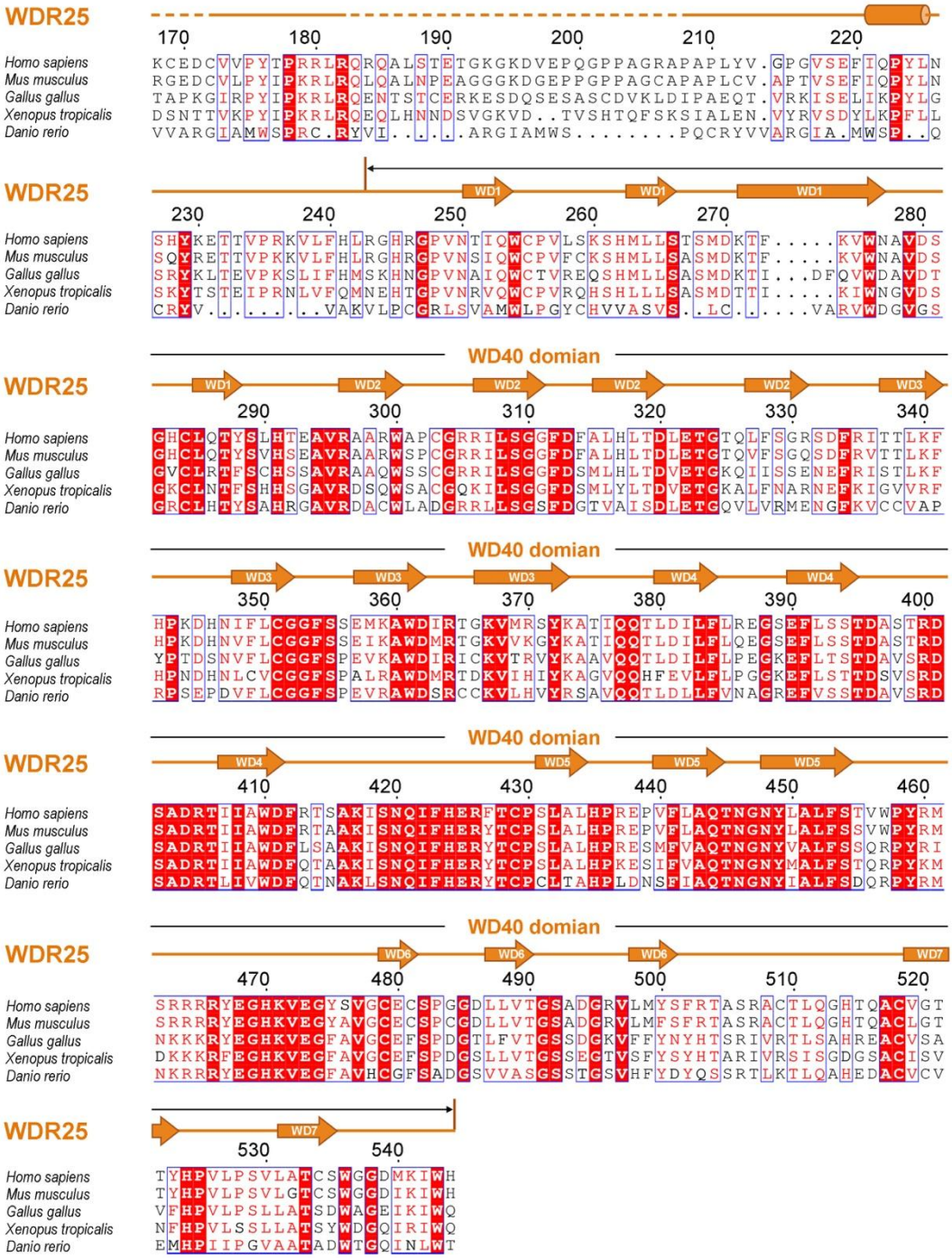


Supplementary Figure S7 | The U12-type step-II factor FAM204A in the human minor C* complex. **a,** The overall EM density map for FAM204A. **b,** Sequence alignment of FAM204A from human, mouse, chicken, frog and zebrafish. Secondary structural elements are indicated above the alignment. Identical residues are shaded red. Conserved residues are colored red in blue boxes. Residues that are involved in the interaction with the intron, U12 snRNA, PRP8 β-finger, and U6atac snRNA are marked by triangles colored violet, blue, purple, and green, respectively. **c,** The EM density map for the coiled-coil fragment of FAM204A and its interacting fragment from SYF1. **d,** The EM density maps for the C-terminus of FAM204A.

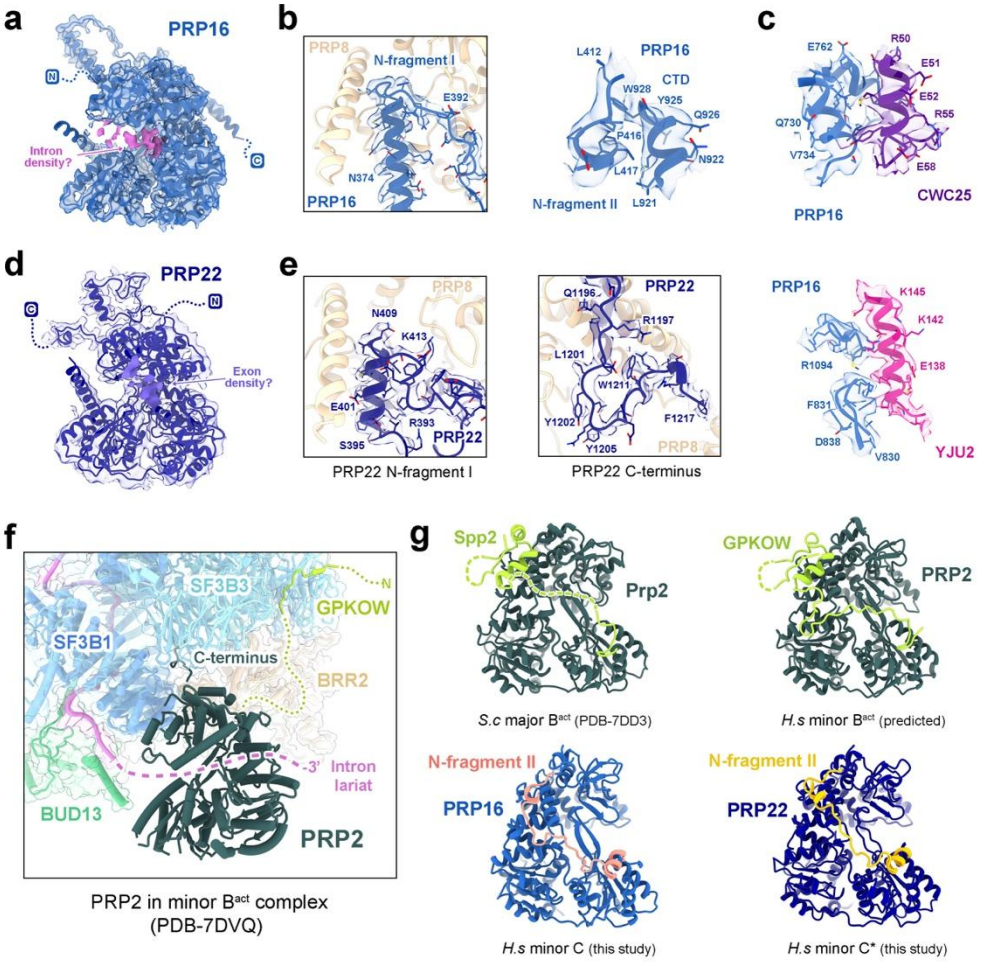


Supplementary Figure S8 | The U12-type step 2 factor WDR25 in the human minor C* complex. **a,** The overall EM density map for WDR25. **b,** Representative EM maps for three WDR25 fragments: the N-terminal region that interacts with ARMC7 (left), a segment that is inserted between PRP8 β-finger and intron/U6atac AAGGA duplex (middle), and a fragment that contacts the BPS/U12 duplex (right). **c,** Sequence alignment for the ARMC7 fragment that interacts with the N-loop of WDR25. The ARMC7 orthologs from human, mouse, chicken, frog and zebrafish are shown. Invariant residues are shaded red. Conserved residues are colored red in blue boxes. Residues that are involved in the interaction with WDR25 are marked by orange triangles. **d,** WDR25 N-fragment directly binds to ARMC7. The interactions were examined using gel filtration. Shown on the left panel are the gel filtration chromatograms of the WDR25 fragment (residues 152-204) alone, WDR25 in complex with WT ARMC7, and WDR25 in complex with mutant ARMC7. Fractions from gel filtration were analyzed on SDS-PAGE gels stained by Coomassie blue (right panels). In the presence of WT ARMC7, the peak for WDR25 (residues 152-204) is shifted to earlier fractions, indicating direct interactions. The

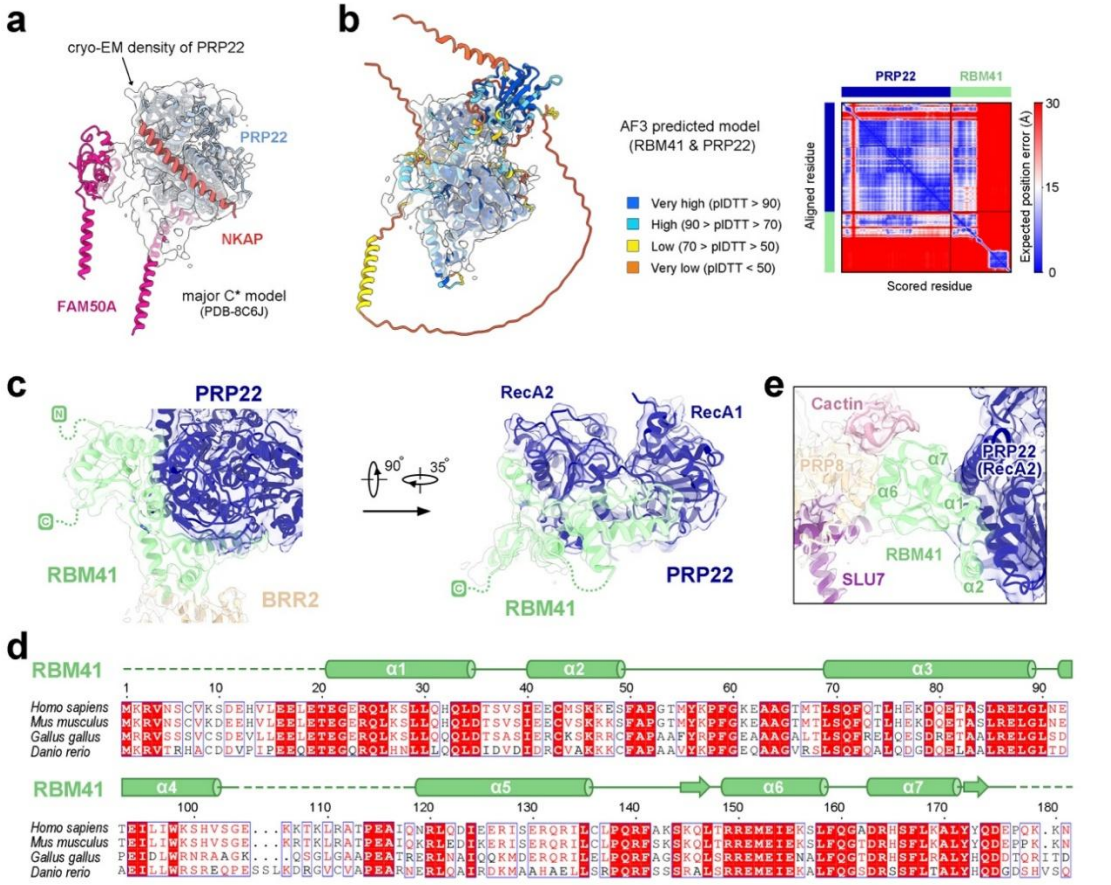
ARMC7 mutations Q18A/Q25A/Q57A/D62A that target the interface with WDR25 abolish the interactions. **e,** Structural comparison between the WD40 domains of WDR25 and PRP17. WDR25 is colored orange and PRP17 is colored gray. The WD40 domains of the two proteins can be superimposed with a root mean squared deviation (RMSD) of 1.34 Å over 250 aligned Cα atoms. **f,** Sequence alignment between the WD40 domains of WDR25 and PRP17. Conserved sequences are boxed. Identical residues are shaded red.



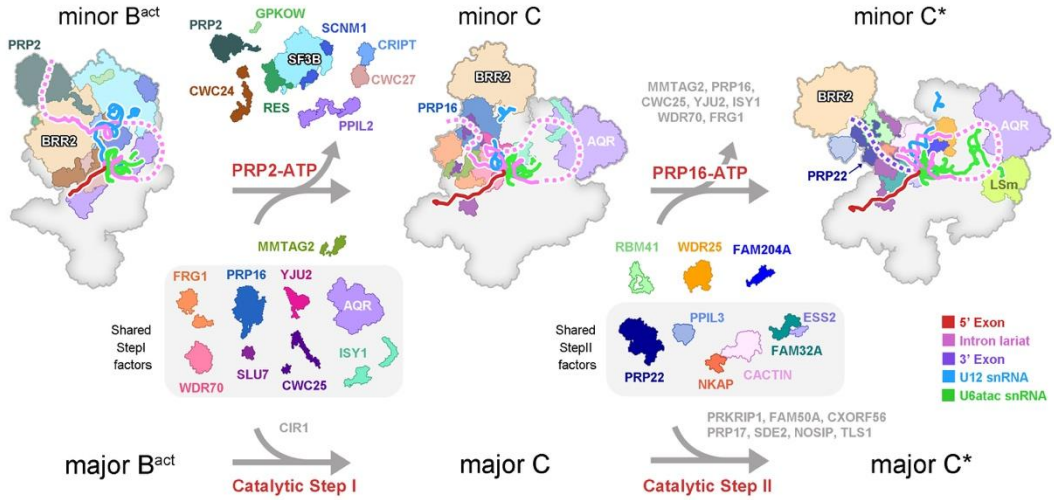
Supplementary Figure S9 | Sequence alignment of WDR25. Sequence alignment of WDR25 from human, mouse, chicken, frog and zebrafish. Secondary structural elements are indicated above the alignment. Identical residues are shaded red. Conserved residues are colored red in blue boxes.



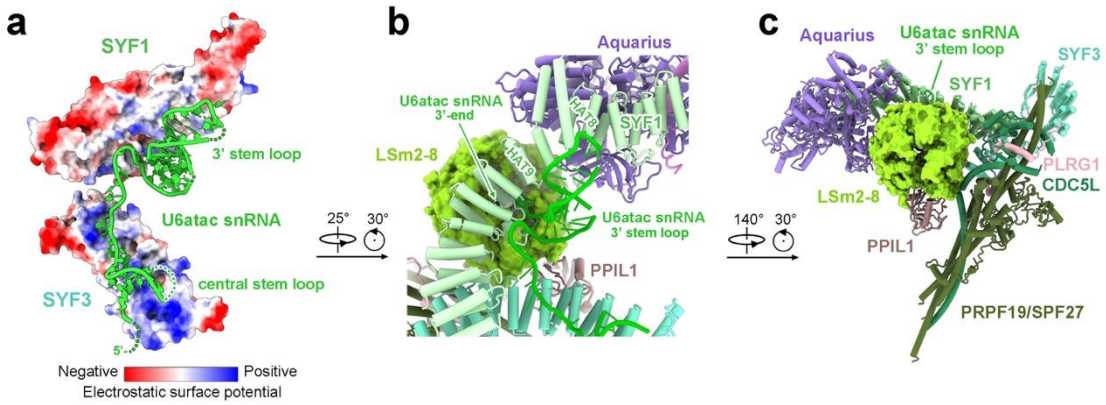
Supplementary Figure S10 | The ATPases PRP16 and PRP22 in the human minor spliceosomes. **a,** The EM density map for PRP16 in the minor C complex. The density map is shown as a transparent surface. The additional density, colored violet, between RecA and CTD of PRP16 is potentially derived from the 3′ end of the pre-mRNA. **b,** The EM density maps for the N-fragment I (left) and II (right) of PRP16. **c,** The EM density maps for the local regions of PRP16 that interact with CWC25 (upper) and YJU2 (bottom). **d,** The EM density map for PRP22 in the minor C* complex. The additional density, colored slate blue, between RecA and CTD of PRP22 is potentially derived from the 3′ end of the 3′ exon. **e,** The EM density maps for the N-fragment I (left) and C-terminus (right) of PRP22. **f,** Structure of PRP2 in the minor B^{act} complex (PDB: 7DVQ[1]) and its interactions with surrounding spliceosomal components. PRP2 is anchored to the minor B^{act} complex through interactions with SF3B1 (light blue), SF3B3 (light cyan), BRR2 (wheat), BUD13 (spring green), and its coactivator GPKOW (yellow green). Violet dashed lines indicate the possible path of the sequences downstream of the BPS. **g,** Comparison between the structures of Prp2-Spp2 (PDB code: 7DD3[7]), PRP2-GPKOW (predicted using AlphaFold3[8]), PRP16 in the human minor C complex, and PRP22 in the human minor C* complex. The binding modes of PRP16 N-fragment II (residues 412-447) and PRP22 N-fragment II (residues 444-479) to the helicase core (RecA and CTD) resemble the interaction between Spp2 G-patch and Prp2.



Supplementary Figure S11 | The U12-type step-II factor RBM41 in the human minor C* complex. **a,** PRP22 and its interacting partners from the human major C* complex (PDB code: 8C6J[6]) do not align well with the local EM map of PRP22 in the human minor C* complex. **b,** The AlphaFold3-predicted model[8] for the helicase core of PRP22 (residues 439-1183) in complex with full-length RBM41 fits well with the local EM map around PRP22 in the minor C* complex. The model is colored by AlphaFold per-residue model confidence score (plDDT, upper right). RBM41 consists of an N-terminal domain (NTD) and a C-terminal RRM. AlphaFold prediction suggests binding of the RBM41 NTD to the PRP22 RecA domain. The predicted aligned error (PAE) plot for the model of PRP22 (439-1183)-RBM41 complex is shown on the right. Cross-peaks between PRP22 RecA domain and RBM41 NTD show lower PAE, indicating high confidence for relative orientation of these regions. **c,** The EM density map for RBM41 in the human minor C* complex. Two views are shown. The NTD of RBM41 interacts with the RecA domain of PRP22 and the C-terminal cassette of BRR2. RBM41, PRP22, and BRR2 are colored pale green, navy, and bisque, respectively. **d,** Sequence alignment of RBM41 NTD from human, mouse, chicken, and zebrafish. Secondary structural elements are indicated above the alignment. Conserved residues are colored red in blue boxes. **e,** The EM density map for the local region of RBM41 that is responsible for recruiting PRP22 to the spliceosome.



Supplementary Figure S12 | A schematic diagram of the remodeling of the minor spliceosome for the branching and exon ligation reactions. PRP2-mediated remodeling of the minor B^{act} complex results in dissociation of the SF3b complex, RES, and seven splicing factors, and recruitment of nine factors. This remodeling converts the minor spliceosome into the Step-I conformation. During this transition, eight of the nine recruited factors are shared between the minor and major C complexes, including five active site proteins (CWC25, YJU2, ISY1, FRG1, and WDR70), the helicases PRP16 and Aquarius, and SLU7. They are recruited through conserved binding interfaces. During the PRP16-driven C-to-C^* transition, the minor spliceosome recruits three minor-specific factors: the WDR25–FAM204A pair, which stabilizes the BPS/U12 duplex, and RBM41, which bridges PRP22, through unique structural features. In addition, the C-to-C^* transition allows recruitment of six proteins that are shared with the major C^* complex: four proteins surrounding the active site, the helicase PRP22, and ESS2.



Supplementary Figure S13 | Structural features of the 3′-region of U6atac snRNA in the minor C* complex. **a,** Sequences downstream of the central stem loop of U6atac snRNA bind along a positively-charged surface formed by two superhelical proteins SYF3 and SYF1. **b,** The 3′ stem loop of U6atac snRNA is anchored by the helices between the HAT8 and HAT9 repeats of SYF1. **c,** The U6atac LSm ring is surrounded by the PRP19 core, SYF1, and Aquarius.

Supplementary Table S1 Summary of mass spectrometry analysis of the human minor C and C* complexes.

Spliceosomal subcomplex	Protein	Uniprot ACC. No.	No. of Peptides		pFind Score		Protein found in EM map?	
			minor C	minor C*	minor C	minor C*	minor C	minor C*
U5 snRNP	PRP8	Q6P2Q9	131	118	5.519	6.068	yes	yes
	SNU114	Q15029	64	55	6.931	7.09	yes	yes
	BRR2	O75643	126	101	5.673	6.042	yes	yes
	U5-40K	Q96DI7	25	21	5.475	7.057	yes	yes
	SmB	P14678	16	13	4.604	4.724	yes	yes
	SmD1	P62314	8	6	5.816	5.537	yes	yes
	SmD2	P62316	12	9	5.833	5.773	yes	yes
	SmD3	P62318	11	9	4.146	5.961	yes	yes
	SmF	P62306	12	10	4.532	5.562	yes	yes
	SmE	P62304	12	8	5.896	6.605	yes	yes
	SmG	P62308	10	2	3.641	2.399	yes	yes
U6atac snRNP	LSm2	Q9Y333	8	7	4.874	6.264	no	yes
	LSm3	P62310	2	2	2.076	2.284	no	yes
	LSm4	Q9Y4Z0	7	8	4.195	4.503	no	yes
	LSm5	Q9Y4Y9	—	—	—	—	no	yes
	LSm6	P62312	3	3	2.118	2.33	no	yes
	LSm7	Q9UK45	3	2	1.671	1.536	no	yes
	LSm8	O95777	4	4	4.109	4.889	no	yes
U12 snRNP	SmB	P14678	16	13	4.604	4.724	yes	yes
	SmD1	P62314	8	6	5.816	5.537	yes	yes
	SmD2	P62316	12	9	5.833	5.773	yes	yes
	SmD3	P62318	11	9	4.146	5.961	yes	yes
	SmF	P62306	12	10	4.532	5.562	yes	yes
	SmE	P62304	12	8	5.896	6.605	yes	yes
	SmG	P62308	10	2	3.641	2.399	yes	yes
NTC	PRPF19	Q9UMS4	19	20	6.691	6.8	yes	yes
	SPF27	O75934	18	12	4.863	5.407	yes	yes
	CDC5L	Q99459	28	25	4.319	4.977	yes	yes
	SYF1	Q9HCS7	38	27	5.004	4.725	yes	yes
	SYF3/CRNL1	Q9BZJ0	32	29	3.519	4.693	yes	yes
	ISY1	Q9ULR0	10	6	3.623	3.796	yes	yes
NTR	SKIP	Q13573	18	16	4.001	6.348	yes	yes
	PLRG1	O43660	19	19	4.53	4.873	yes	yes
	CWC15	Q9P013	6	2	3.843	2.082	yes	yes
	PPIL1	Q9Y3C6	8	6	4.329	3.55	yes	yes

Spliceosomal subcomplex	Protein	Uniprot ACC. No.	No. of Peptides		pFind Score		Protein found in EM map?	
			minor C	minor C*	minor C	minor C*	minor C	minor C*
IBC	AQR	O60306	71	60	6.398	6.29	yes	yes
	PPIE	Q9UNP9	13	9	3.821	3.554	yes	yes
EJC	eIF4AIII	P38919	25	22	5.273	5.695	yes	yes
	MAGOH	P61326	6	8	4.488	4.04	yes	yes
	Y14/RBM8A	Q9Y5S9	6	6	4.736	5.399	yes	yes
Splicing Factors (Shared)	SLU7	O95391	9	5	3.32	2.965	yes	yes
	SRm300	Q9UQ35	55	39	5.87	5.71	yes	yes
	CWC22	Q9HCG8	28	13	4.222	3.231	yes	yes
	PPWD1	Q96BP3	36	22	4.709	4.956	yes	yes
	CCDC12	Q8WUD4	5	4	4.856	4.923	yes	yes
	RBM48	Q5RL73	5	5	4.205	4.831	yes	yes
	ARMC7	Q9H6L4	8	3	4.352	3.268	yes	yes
minor C specific (Step I Factors)	PRP16	Q92620	42	24	4.565	4.511	yes	no
	FRG1	Q14331	6	3	3.065	3.034	yes	no
	WDR70	Q9NW82	14	6	4.384	2.443	yes	no
	CWC25	Q9NXE8	9	5	2.823	2.356	yes	no
	YJU2	Q9BW85	18	2	4.248	2.708	yes	no
	MMTAG2	Q9BU76	7	4	3.404	2.6	yes	no
minor C* specific (Step II Factors)	FAM32A	Q9Y421	1	1	2.159	2.778	no	yes
	CACTIN	Q8WUQ7	11	13	4.804	5.141	no	yes
	PPIL3	Q9H2H8	9	6	3.752	4.279	no	yes
	ESS2	Q96DF8	3	6	3.547	3.895	no	yes
	NKAP	Q8N5F7	1	1	1.572	1.549	no	yes
	PRP22	Q14562	49	38	6.022	5.303	no	yes
	WDR25	Q64LD2	12	10	4.972	3.531	no	yes
	FAM204A	Q9H8W3	7	6	3.332	3.315	no	yes
	RBM41	Q96IZ5	19	10	3.643	3.756	no	yes

The two samples that yielded the cryo-EM reconstruction of the human minor C and C* complexes were analyzed by mass spectrometry. Shown here is a summary of all components of the human minor C and C* complexes that were identified by this analysis. The pFind score represents the confidence and ranking of the target protein, with higher score represents higher confidence and higher ranking.

Supplementary Table S2 Cryo-EM data collection, refinement, and validation (the minor C complex).

	Minor C complex, overall EMD-YYYY PDB: XXXX	Minor C with more stable RBM48 EMD-YYYY	SYF region, focused EMD-YYYY	PRP16 region, focused EMD-YYYY	Minor C with more stable PRP8-RH EMD-YYYY	EJC region, focused EMD-YYYY	Helicase region, focused EMD-YYYY	IBC region, focused EMD-YYYY
Data collection and processing								
Microscope	Titan Krios							
Voltage (keV)	300							
Magnification	81,000							
Camera	K3							
Pixel size (Å/pixel)	1.0773							
Total electron exposure (e–/Å^2)	50							
Number of frames	32							
Defocus range (μm)	1.3~2.3							
Automation software	EPU							
Energy filter slit width	20							
Micrographs used (no.)	34,090							
Total extracted particles (no.)	4,415,938							
For each reconstruction								
Refined particles (no.)	86,306	58,941	85,373	50,530	37,176	51,437	72,597	85,373
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1	C1
Resolution (global, Å)								
FSC 0.143	2.92	3.07	3.21	3.29	3.48	3.59	3.97	4.41
Resolution range (local, Å)	0~20							
Map sharpening B factor (Å^2)	-52.8	-48.8	-71.7	-79.5	-42.2	-91.1	-108.0	-142.4
Model composition								
Non-hydrogen atoms	126,724							
Protein residues	15,260							
Ligands	19							
RNA/DNA nucleotides	282							
Model refinement								
Refinement package	Phenix.real_space_refine							
Resolution cutoff	3.0							
Model-Map CC	0.83							
FSC (Model-map) = 0.5	3.16							
B factors (Å^2)								
Protein	207.51							
Ligand	135.54							
RNA/DNA nucleotides	114.54							
R.m.s. deviations								
Bond lengths (Å)	0.003							
Bond angles (°)	0.767							
Validation								
MolProbity score	1.71							
CaBLAM outliers	1.75							
Clashscore	5.63							
Poor rotamers (%)	1.62							
C-beta deviations	0							
EMRinger score	1.80							
Ramachandran plot								
Favored (%)	96.32							
Allowed (%)	3.56							
Outliers (%)	0.12							

Supplementary Table S3 Cryo-EM data collection, refinement, and validation (the minor C* complex).

	Minor C*, overall EMD-YYYY PDB: XXXX	Minor C* with stable WDR25 EMD-YYYY	RBM48 region, focused EMD-YYYY	SYF region, focused EMD-YYYY	Minor C* with stable BRR2 EMD-YYYY	Minor C* with stable IBC EMD-YYYY	EJC region, focused EMD-YYYY	PRP22 region, focused EMD-YYYY	PRP19/LSm, focused EMD-YYYY
Data collection and processing									
Microscope	Titan Krios								
Voltage (kV)	300								
Magnification	81,000								
Camera	K3								
Pixel size (Å/pixel)	1.0773								
Total electron exposure (e–/Å^2)	50								
Number of frames	32								
Defocus range (μm)	1.3~2.3								
Automation software	EPU								
Energy filter slit width	20								
Micrographs used (no.)	58,431								
Total extracted particles (no.)	6,380,953								
For each reconstruction									
Refined particles (no.)	145,908	49,173	88,127	107,934	30,195	30,081	61,936	95,438	22,629
Symmetry imposed	C1								
Resolution (global, Å)									
FSC 0.143	3.04	3.13	3.19	3.30	3.50	3.60	4.01	4.23	7.71
Resolution range (local, Å)	0~20								
Map sharpening B factor (Å^2)	-62.5	-46.2	-74.0	-83.7	unsharpened	unsharpened	-99.1	-142.2	unsharpened
Model composition									
Non-hydrogen atoms	131,880								
Protein residues	16,227								
Ligands	18								
RNA/DNA nucleotides	345								
Model refinement									
Refinement package	Phenix.real_space_refine								
Resolution cutoff	3.1								
Model-Map CC	0.73								
FSC (Model-map) = 0.5	3.33								
B factors (Å^2)									
Protein	246.61								
Ligand	388.58								
RNA/DNA nucleotides	266.64								
R.m.s. deviations									
Bond lengths (Å)	0.003								
Bond angles (°)	0.780								
Validation									
MolProbity score	1.66								
CaBLAM outliers	1.99								
Clashscore	5.70								
Poor rotamers (%)	1.29								
C-beta deviations	0								
EMRinger score	1.63								
Ramachandran plot									
Favored (%)	96.14								
Allowed (%)	3.75								
Outliers (%)	0.11								

Supplementary Table S4 Summary of model building for the human minor C complex.

	Components	Chain ID	Uniprot No.	Total length	Modeled	Modeling template	Modeling method	Resolution (Å)
mRNA	MSE-U12	0	-	334	5'-exon: -25~-1	7dvq, 8i0w	Docked & rebuilt	2.5~4.0
		1			Intron: 1~20,189~214			
U5 snRNP	U5 snRNA	5	-	117	9~70	7dvq	Docked & rebuilt	2.5~3.5
					3~8, 71~78		Rigid docking	3.5~5.5
					Sm binding: 85~117			
	PRP8	A	Q6P2Q9	2335	N-domain: 25~284,292~800	7dvq	Docked & rebuilt	2.5~4.0
					Core: 800~1521,1526~1754			
					RH-like: 1755~2019	8i0w		
					Jab1/MPN: 2067-2335	8i0w	Rigid docking	3.0~4.5
					Ligand: IP6	7dco	Docked & adjusted	~3.0
	SNU114	B	Q15029	972	56~955	7dvq	Docked & rebuilt	2.5~3.5
					Ligand: GTP, Mg^{2+}			
	BRR2	C	O75643	2136	100~184, 251~366, 404~2125	8i0w	Docked & rebuilt	3.5~6.0
	U5-40K	D	Q96DI7	357	WD40: 58-356	7dvq	Docked & rebuilt	3.5~6.0
	SmB	a	P14678	240	6~51, 62~88	7dvq	Docked & rebuilt or Rigid docking according to local resolution	3.5~6.0
	SmD1	b	P62314	119	1~82			
	SmD2	c	P62316	118	10~77, 87~116			
	SmD3	d	P62318	126	1~85			
	SmE	e	P62304	86	14~90			
	SmF	f	P62306	92	3~75			
	SmG	g	P62308	76	3~76			
U6atac snRNP	U6atac snRNA	6	-	125	1~56, 61~65	7dvq	Docked & rebuilt	2.5~4.0
U12 snRNP	U12 snRNA	2	-	130	1-30	7dvq	Docked & rebuilt	2.6~3.5
					Sm binding: 73-84		Rigid docking	>8.0
	SmB	h	P14678	240	6~50, 61~88	7dvq	Rigid docking	>8.0
	SmD1	i	P62314	119	2~82			
	SmD2	j	P62316	118	10~77, 87~116			
	SmD3	k	P62318	126	4~84			
	SmE	l	P62304	86	14~90			
	SmF	m	P62306	92	3~75			
	SmG	n	P62308	76	3~76			
NTC	PRP19	q-t	Q9UMS4	504	q: 1~137; r: 1~135; s: 1~138; t:1~136	C* (this study)	Rigid docking	>8.0
	SPF27	u	O75934	225	1~225	C* (this study)	Rigid docking	>8.0
	CDC5L	E	Q99459	802	4~111,116~233	7dvq	Docked & rebuilt	2.5~3.5
					461~487	AF3 predicted	Rigid docking	6.0~8.0
					511~802	C* (this study)	Rigid docking	>8.0
	SYF1	F	Q9HCS7	855	19~313,322~384,395~736	8i0w	Docked & rebuilt	>3.5
	CRNKL1/SYF3	G	Q9BZJ0	848	174~755	7dvq	Docked & rebuilt	2.5~6.0
	ISY1	I	Q9ULR0	285	2~101, 116~137, 148~195	8i0w	Docked & rebuilt	>2.5
NTC Related proteins (NTR)	SKIP/hPrp45	J	Q13573	536	86~161, 203~216, 222~262, 278~319	7dvq	Docked & rebuilt	2.5~4.0
					1~13, 171~178, 196~202	C* (this study)	Docked & adjusted	3.0~6.0
	PLRG1/hPrp46	K	O43660	514	1-59	C* (this study)	Rigid docking	>8.0
					93~118	C* (this study)	Docked & adjusted	3.0~4.0
					WD40: 184~504	7dvq	Docked & rebuilt	2.5~3.5
	CWC15	N	Q9P013	229	25~78,182~229	7dvq, 8i0w	Docked & rebuilt	3.0~3.5
	PPIL1	O	Q9Y3C6	166	8~116	8i0w	Rigid docking	>8.0
IBC	AQR	Q	O60306	1485	21,497, 502~526, 529~752, 773~1373	8i0w	Rigid docking	>6.0
	PPIE	P	Q9UNP9	334	8~166			
EJC	eIF4AIII	v	P38919	411	22~411	8i0w	Docked & rebuilt	3.0~6.0
	MAGOH	w	P61326	146	5~148			
	RBM8A/Y14	x	Q9Y5S9	174	64~154			
Splicing factors	CWC25	R	Q9NXE8	425	2~77	8i0w	Docked & rebuilt	2.5~4.0
					122~148	AF3 predicted	Docked & rebuilt	3.5~5.0
	YJU2	Y	Q9BW85	323	2~183	8i0w	Docked & rebuilt	2.5~4.0
	PRP16/DHX38	T	Q92620	1227	361~469, 523~1192	AF	Docked & rebuilt	3.0~4.0
	SRm300/SRRM2	U	Q9UQ35	2752	1~28	7dvq	Docked & rebuilt	2.5~3.0
					36~106	8i0w	Docked & rebuilt	3.5~4.0
					122~161	AF3 predicted	Docked & rebuilt	3.5~5.0
	CWC22	V	Q9HCG8	908	149~406, 447~652	7dvq	Docked & rebuilt	3.0~5.0
	SLU7	S	O95391	586	ZnF: 112~152	C* (this study)	Docked & adjusted	3.0~4.0
	FRG1	X	Q14331	258	38~188, 197~258	AF	Docked & rebuilt	3.0~5.0
	hMMTAG2	z	Q9BU76	263	7~57, 59~83	A^2-Net & AF3	Docked & rebuilt	3.0~5.0
	WDR70	Z	Q9NW82	654	WD40: 163~503	8i0w	Docked & rebuilt	3.0~5.0
	RBM48	L	Q5RL73	367	2~18, 23~146	7dvq	Docked & rebuilt	3.0~4.0
	ARMC7	M	Q9H6L4	198	11~183	7dvq	Docked & rebuilt	3.0~4.0
	PPWD1	4	Q96BP3	646	62~411, 430~486	8i0w	Rigid docking	3.5~6.0
	CCDC12	H	Q8WUD4	848	107~150	AF3 predicted	Docked & adjusted	4.0~6.0

Under the column labeled "Modeling template", AF: models from AlphaFold DB; AF3: models predicted using AlphaFold 3; A^2-Net: models generated by A2-Net deep neural network method; HM: homology modeling.

Supplementary Table S5 Summary of model building for the human minor C* complex.

	Components	Chain ID	Uniprot No.	Total length	Modeled	Modeling template	Modeling method	Resolution (Å)
mRNA	MSE-U12	0	-	334	5'-Exon: -25~-1	7dvq, 8c6j	Docked & rebuilt	2.5~4.0
		1			Intron: 1~22,185~208, 212-218			
					3'-Exon: 219-220			
U5 snRNP	U5 snRNA	5	-	117	9~70	7dvq	Docked & rebuilt	2.5~3.5
					3~8, 71~78		Rigid docking	3.5~5.5
					Sm binding: 85~117			
	PRP8	A	Q6P2Q9	2335	N-domain: 25~284,292~800	7dvq	Docked & rebuilt	2.5~4.0
					Core: 801~1754			2.6~4.0
					RH-like: 1755~2036	8c6j	Docked & rebuilt	~5.0
					Jab1/MPN: 2067-2335	8c6j	Rigid docking	3.0~4.5
					Ligand: IP6	7dco	Docked & adjusted	~3.0
	SNU114	B	Q15029	972	56~954	7dvq	Docked & rebuilt	2.6~3.5
					Ligand: GTP, Mg^{2+}			
	BRR2	C	O75643	2136	100~184, 251~366, 404~2125	8c6j	Rigid docking	>5.0
	U5-40K	D	Q96DI7	357	WD40: 50-357	7dvq	Docked & rebuilt	3.5~5.0
	SmB	a	P14678	240	6~51, 62~88	7dvq	Rigid docking	3.0~6.0
	SmD1	b	P62314	119	1~82			
	SmD2	c	P62316	118	10~77, 87~116			
	SmD3	d	P62318	126	1~85			
	SmE	e	P62304	86	14~90			
	SmF	f	P62306	92	3~75			
	SmG	g	P62308	76	3~76			
U6atac snRNP	U6atac snRNA	6	-	125	1~56, 61~65	7dvq	Docked & rebuilt	2.5~4.0
					66~79, 117~120	-	*de novo*	4.0~8.0
					3' stem loop: 80~101, 104~116	trRosettaRNA	Docked & adjusted	7.0~8.0
					121-125	HM (4m7a)	Rigid docking	7.0~8.0
	LSm3	L3	P62310	102	16~100	HM (4m7a)	Rigid docking	7.0~8.0
	LSm2	L2	Q9Y333	95	1~95			
	LSm8	L8	O96777	96	1~96			
	LSm4	L4	Q9Y4Z0	139	1~80			
	LSm7	L7	Q9UK45	103	11~99			
	LSm5	L5	Q9Y4Y9	91	10~91			
	LSm6	L6	P62312	80	4~78			
U12 snRNP	U12 snRNA	2	-	130	1-30	7dvq	Docked & rebuilt	2.5~3.5
					Sm binding: 73-84		Rigid docking	>8.0
	SmB	h	P14678	240	6~50, 61~88			
	SmD1	i	P62314	119	2~82			
	SmD2	j	P62316	118	10~77, 87~116			
	SmD3	k	P62318	126	4~84			
	SmE	l	P62304	86	14~90			
	SmF	m	P62306	92	3~75			
	SmG	n	P62308	76	3~76			
NTC	PRP19	q-t	Q9UMS4	504	q: 1~137; r: 1~135; s: 1~138; t:1~136	AF3 predicted	Rigid docking	7.0~8.0
	SPF27	u	O75934	225	1~225			
	CDC5L	E	Q99459	802	4~111, 121~233,	7dvq	Docked & rebuilt	2.5~3.5
					289~327, 359~387, 461~487,	AF3 predicted	Docked & rebuilt	3.5~6.0
					511~802	AF3 predicted	Rigid docking	>6.0
	SYF1	F	Q9HCS7	855	19~313, 322~384, 395~749, 762~789	AF	Docked & rebuilt	3.5~8.0
	CRNKL1/SYF3	G	Q9BZJ0	848	174~812	AF	Docked & rebuilt	2.5~6.0
	ISY1	I	Q9ULR0	285	148~195	8c6j	Rigid docking	>7.0
NTC Related proteins (NTR)	SKIP/hPrp45	J	Q13573	536	1~13, 16~30, 51~78,	8c6j	Rigid docking	5.0~7.0
					87~161, 203~216, 222~262, 276~319	7dvq	Docked & rebuilt	3.0~4.0
					171~178, 196~202	A^2-Net	Docked & rebuilt	3.0~4.0
	PLRG1/hPrp46	K	O43660	514	1-59	AF3 predicted	Rigid docking	>6.0
					93~118	A^2-Net	Docked & rebuilt	3.0~4.0
					WD40: 184~504	7dvq	Docked & rebuilt	2.5~3.5
	CWC15	N	Q9P013	229	25~78,186~229	7dvq, 8c6j	Docked & rebuilt	3.0~4.0
	PPIL1	O	Q9Y3C6	166	8~166	8c6j	Rigid docking	>6.0
IBC	AQR	Q	O60306	1485	21,497, 502~526, 529~752, 773~1373	AF	Rigid docking	>6.0
	PPIE	P	Q9UNP9	334	5~83	AF		
EJC	eIF4AIII	v	P38919	411	22~411	8c6j	Docked & rebuilt	3.0~6.0
	MAGOH	w	P61326	146	5~148			
	RBM8A/Y14	x	Q9Y5S9	174	64~154			
Splicing factors	SLU7	S	O95391	586	31~196, 265~370	8c6j	Docked & rebuilt	3.0~4.5
	DHX8/PRP22	T	Q14562	1220	386~502, 555~1217	AF	Docked & rebuilt	3.5~5.0
	SRm300/SRRM2	U	Q9UQ35	2752	1~28	7dvq	Docked & rebuilt	2.5~3.0
					36~106	8c6j	Docked & rebuilt	3.5~4.0
					122~161	AF3 predicted	Docked & rebuilt	3.5~5.0
	CWC22	V	Q9HCG8	908	149~406, 447~652	7dvq	Docked & rebuilt	3.0~5.0
	CACTIN	X	Q8WUQ7	772	617~758	8c6j	Docked & rebuilt	3.0~5.0
	FAM32A	Y	Q14331	112	53~111	8c6j	Docked & rebuilt	3.0~4.0
	NKAP	Z	Q8N5F7	415	328~358	8c6j	Docked & rebuilt	3.5~4.5
	RBM48	L	Q5RL73	367	1~18, 12~146	7dvq	Docked & rebuilt	3.0~4.0
	ARMC7	M	Q9H6L4	198	11~183	7dvq	Docked & rebuilt	3.0~4.5
	WDR25	W	Q64LD2	544	172~182, 208~544	A^2-Net	Docked & rebuilt	2.8~3.5

	FAM204A	z	Q9H8W3	233	120~233	A^2-Net	Docked & rebuilt	2.5~4.0
	RBM41	y	Q96IZ5	413	18~104, 118~177	AF3 predicted	Docked & rebuilt	4.0~5.0
	ESS2	p	Q96DF8	476	37~65	8c6j	Docked & rebuilt	3.0~5.0
	PPIL3	R	Q9H2H8	161	1~160	8c6j	Docked & rebuilt	3.0~4.0
	PPWD1	4	Q96BP3	646	62~411, 430~486	8c6j	Rigid docking	>7.0
	CCDC12	H	Q8WUD4	848	107~150	AF3 predicted	Docked & adjusted	3.5~6.0

Under the column labeled "Modeling template", AF: models from AlphaFold DB; AF3: models predicted using AlphaFold 3; A^2-Net: models generated by A2-Net deep neural network method; HM: homology modeling.

References:

1 Bai, R. *et al.* Structure of the activated human minor spliceosome. *Science* **371**, eabg0879 (2021). https://doi.org:10.1126/science.abg0879

2 Chen, S. *et al.* High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy* **135**, 24-35 (2013). https://doi.org:10.1016/j.ultramic.2013.06.004

3 Afonine, P. V. *et al.* Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* **74**, 531-544 (2018). https://doi.org:10.1107/S2059798318006551

4 Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**, 70-82 (2021). https://doi.org:10.1002/pro.3943

5 Zhan, X., Lu, Y. & Shi, Y. Molecular basis for the activation of human spliceosome. *Nat Commun* **15**, 6348 (2024). https://doi.org:10.1038/s41467-024-50785-0

6 Dybkov, O. *et al.* Regulation of 3' splice site selection after step 1 of splicing by spliceosomal C* proteins. *Sci Adv* **9**, eadf1785 (2023). https://doi.org:10.1126/sciadv.adf1785

7 Bai, R. *et al.* Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. *Science* **371** (2021). https://doi.org:10.1126/science.abe8863

8 Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493-500 (2024). https://doi.org:10.1038/s41586-024-07487-w