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*]

[1]State Key Laboratory of Gene Expression, Zhejiang Key Laboratory of Structural Biology, Research Center for Industries of the Future and School of Life Sciences, Westlake University, Hangzhou, Zhejiang, China

[2]Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China

[3]Institute of Biology, Westlake Institute for Advanced Study, Hangzhou, Zhejiang, China.

[4]These authors contributed equally to this work.

[*]To whom correspondence should be addressed. E-mail: wanruixue@westlake.edu.cn

Abstract

Precursor messenger RNA (pre-mRNA) splicing is a central process in eukaryotic gene expression. Among the introns removed, the rare but essential U12-type introns are excised by the minor spliceosome, which in most multicellular eukaryotes constitutes a rate-limiting step with distinct regulatory potential. Despite its critical role, the catalytic mechanism of this distinctive spliceosome remains largely unknown. Here we report the isolation and structure determination of two catalytic states of the human minor spliceosome: the branching-completed C complex and the exon-ligation-ready C* complex. These two cryo-EM structures, at average resolutions of 2.9-3.0 Å, reveal the configuration of key RNA elements and the positions of catalytic metals at the splicing active site. The U12-type catalysis is safeguarded by splicing factors that are unique to the minor spliceosome: MMTAG2 facilitates the stabilization of the branch point sequence (BPS)/U12 duplex to promote branching reaction (step-I); WDR25 and FAM204A orient the BPS/U12 duplex to dock the 3′ splice site (3′SS) for exon-ligation (step-II); RBM41 assists spliceosome binding of the helicase PRP22 to prepare for the mRNA release. These findings reveal catalytic principles of the U12-type splicing and establish new molecular links to development and diseases.

Splicing stands at the core of eukaryotic gene expression, responsible for the removal of introns from precursor messenger RNAs (pre-mRNAs) to generate mature mRNAs[1-4]. This process is catalyzed by the spliceosome, a dynamic ribonucleoprotein complex that must cycle through precise assembly, activation, catalysis, and disassembly[5]. Most multicellular eukaryotes employ two parallel splicing systems[6]: the U2-type major spliceosome, which processes the vast majority of introns, and the U12-type minor spliceosome, which excises a rare but evolutionarily conserved class of introns with characteristic 5′ splice site (5′SS) and BPS, but without a defined polypyrimidine tract upstream of the 3′SS[7,8]. The evolutionary maintenance of two distinct splicing machineries must serve a non-redundant biological purpose. This rationale is evident in the specialized regulatory role of minor splicing: U12-type introns are highly enriched in essential genes and function as solitary, slow-processing units within transcripts[6,9]. Their delayed excision establishes a conserved rate-limiting checkpoint[10], potentially adding a critical layer of regulation to the expression of vital genes.

The minor spliceosome comprises five small nuclear ribonucleoprotein particles (snRNPs), four of which (U11, U12, U4atac, and U6atac) are unique to this pathway[11-13]. Its assembly initiates with recognition of 5′SS and BPS by the U11/U12 di-snRNP, followed by recruitment of the pre-assembled U4atac/U6atac • U5 tri-snRNP[14]. Despite this established framework and its physiological importance, the molecular mechanism of the minor spliceosome has remained poorly understood. Studies have long depended on genetic and bioinformatic analyses of its substrate introns, which have successfully mapped related genes and disease links[15-17]. However,

these indirect approaches provide limited information about the inner workings of the spliceosomal machine itself. A major obstacle to direct biochemical dissection is its extreme scarcity[8,14]: minor snRNPs are present at merely 1% of the level of its major counterparts, with U6atac being even rarer[18]. This severe low abundance has historically precluded biochemical isolation and functional reconstitution, leaving fundamental questions about its assembly, catalytic mechanism, and regulation unresolved. Therefore, direct insight into its operational dynamics is essential to not only to elucidate its mechanism but also to define the functional niche that ensured its evolutionary persistence alongside the major spliceosome.

Early efforts resulted in the isolation of the U11/U12 di-snRNP[19], which likely represents the most stable and abundant subcomplex. More recently, the optimization of U12-type *in vitro* splicing systems enabled the visualization of the assembly intermediates, including the minor pre-B and B^{act} complexes[20,21], thereby clarifying early steps in the minor splicing pathway coupled with other genetic and structural advances[22-26]. However, these advances stop at the threshold of catalysis. The complexes that execute the two-step transesterification reaction, the chemical core of splicing, have remained entirely uncharacterized. This fundamental gap has precluded any mechanistic insight into critical features including catalytic metal coordination, active-site configuration, substrate delivery, and step-specific control by protein factors, representing the central unresolved problem to completely understand minor splicing.

In this study, we employed a markedly improved U12-type *in vitro* splicing system to capture the human minor spliceosome in both catalytic states, the

post-branching C complex and the pre-exon-ligation C^* complex, and determine their cryo-EM structures at 2.9 Å and 3.0 Å resolution, respectively. These high-resolution views enable us to directly visualize the catalytic core of the minor spliceosome, offering unprecedented biochemical insight into its unique principles when catalyzing two-step transesterification reaction. Our findings elucidate the molecular logic underlying U12-type splicing and reveal how conserved splicing principles are uniquely adapted in this parallel eukaryotic splicing machinery, thereby resolving a central and long-standing question in RNA splicing.

Capturing the minor spliceosome during catalysis

We previously assembled the human minor pre-B and B^{act} complexes using the MINX-U12 pre-mRNA[20,21]. However, all attempts to assemble sufficient catalytic-state minor spliceosomes using this pre-mRNA failed, presumably due to its low splicing efficiency *in vitro*. To overcome this obstacle, we optimized the substrate through parallel screening of pre-mRNA scaffolds and splicing regulatory elements (Fig. 1a; Supplementary Fig. S1a,b). Replacing the 3′ exon with that of SCN4A and incorporating the ENH1 enhancer were found to boost the U12-type splicing efficiency (Supplementary Fig. S1c,d). Subsequent refinement of the U12-type intron sequences yielded the high-performance substrate MSE-U12 (Fig. 1a,b; Supplementary Fig. S1d).

Using the helicase-deficient PRP16 mutant (K561A), which is known to arrest the major spliceosome after first step of catalysis[27], we sought to capture the

human minor C complex on MSE-U12 pre-mRNA (Supplementary Fig. S2a). The purified sample contained U12, U6atac, U5 snRNAs, and the lariat-3′ exon intermediate; the particles appeared intact on the cryo-EM micrograph (Supplementary Fig. S2b,c). In parallel, we attempted to trap the human minor C* complex by lowering the pH of the splicing reaction[28] (Supplementary Fig. S2d-f). Cryo-EM analysis of both samples yielded high-resolution reconstructions (2.9 Å for C complex; 3.0 Å for C* complex), affording an atomic-level view of most regions especially for the catalytic core (Supplementary Fig. S3-8). Furthermore, the application of local refinement strategies improved maps in the most distal areas with resolvable secondary structure features (Fig. 1c,d; Supplementary Figs. S3 and S4). Aided by mass spectrometry (MS) analysis, these data enabled the construction of high-fidelity atomic models, thereby providing a comprehensive structural framework for elucidating the catalysis of the minor spliceosome (Supplementary Tables S1-S5).

Structures of the human minor C and C* complexes

The structure of the human minor C complex contains 49 proteins, three snRNAs (U12, U6atac, and U5), and two MSE-U12 fragments (5′ exon and lariat-3′ exon intermediate) (Fig. 1c). The protein components include those from U5 and U12 snRNPs, NineTeen Complex (NTC), NTC related (NTR), the intron binding complex (IBC), the exon junction complex (EJC), and 13 splicing factors. Notably, the previously uncharacterized protein MMTAG2, originally linked to multiple myeloma[29], is now identified as a step-I factor in minor splicing. The DEAH-box

ATPase/helicase PRP16 is unambiguously resolved and is located close to the intron sequences downstream of the BPS. Unexpectedly, the step-II factor SLU7 is already recruited in this state. In contrast to the major C complex[27], CIR1, which tethers U2 snRNP to PRP8 RNaseH, is absent here.

The structure of the minor C* complex comprises 59 proteins, three snRNAs, a free 5′ exon, and a lariat-3′ exon intermediate (Fig. 1d). Compared to the minor C complex, six step-I factors dissociate and nine proteins are recruited, including the ATPase/helicase PRP22. Among the newly incorporated components, we identify three novel factors, WDR25, FAM204A, and RBM41, each without a defined role in minor splicing. Their recruitment coincides with large-scale, PRP16-driven rearrangements of peripheral modules that collectively primes the complex for exon ligation. Surprisingly, we observe the stable anchoring of the heptameric LSm2-8 complex between the PRP19 core and Aquarius, where it protects the U6atac 3′ end. This specific interaction is absent during catalysis in the major spliceosome[30,31]. Despite these changes, core modules including U5 snRNP, NTC, NTR, EJC, and four splicing factors remain structurally invariant. In contrast to the major C* complex, the minor C* complex lacks seven step-II factors essential for the major spliceosome catalysis: PRP17, SDE2, PRKRIP1, FAM50A, CXORF56, NOSIP, and TLS1[32].

Together, the distinct composition and architecture of the minor spliceosome reflect unique strategies for catalytic center organization and substrate handling, which establishes a specialized mechanistic framework for its splicing pathway.

Catalytic mechanism of U12-type splicing

The atomic models of the minor C and C* complexes provide a near-complete visualization of the RNA conformational dynamics and reveal the unique features that govern both steps of U12-type splicing. The catalytic step-I conformation is assembled by U12, U6atac, and U5 snRNAs, together with the 5′ exon and the intron of MSE-U12 (Fig. 2a; Supplementary Fig. S5a; Supplementary Table S4). The 5′SS is recognized via duplex formation between the intron sequences downstream of the AU dinucleotide and a specific region of U6atac snRNA, the $A_{12}AGGA_{16}$ box and flanking nucleotides, termed 5′SS-paring site (5′SS-PS) (Fig. 2a,b). Concurrently, the BPS form an extended duplex with U12 snRNA, with the invariant branch point adenosine (BP-A) flipped out. The 5′ exon is anchored by loop I of U5 snRNA. The covalent linkage between the 2′-oxygen of the BP-A and the phosphorus of nucleotide A1 at the 5′SS confirms the completion of lariat formation (Fig. 2b).

Within this overall scaffold, the active site harbors a fully assembled heteronuclear metal-ion core organized by the U6atac intramolecular stem loop (ISL) by helix I of U12/U6atac duplex (Fig. 2b; Supplementary Fig. S5b,c). A defining feature is the stable presence of the catalytic metal M2, which activates the BP-A 2′-OH nucleophile, in contrast to its absence in the major C complex[27,33] (Supplementary Fig. S5d). M2 is coordinated by four phosphate groups from U6atac snRNA and A1 of 5′SS. Notably, following the completed branching reaction, the nucleophile no longer coordinates M2, as indicated by their separation of ~2.5 Å (Fig. 2c). The 3′-terminal guanine (G-1) of 5′ exon undergoes a configurational flip, displacing its

newly liberated 3′-OH group (i.e. the leaving group) ~10 Å from the A1 phosphate and away from the metal M1, which stabilizes the leaving group during branching (Fig. 2c). A putative potassium ion, coordinated by U6atac snRNA in a position analogous to the K1 site in group II introns[34], likely stabilizes the two catalytic metals (Supplementary Fig. S5c).

Transition to the minor C* complex involves significant structural resolving, including the previously unassigned regions downstream of the U6atac central stem loop (nucleotides 66–86, 91–101, and 104–125) (Fig. 2d; Supplementary Fig. S5e; Supplementary Table S5). This is accompanied by an ~80° rotation of the entire BPS/U12 duplex, reconfiguring the active site for the second step catalysis (Fig. 2e). The 2′-5′ linkage is displaced outward, creating space for the 3′SS and the 3′ exon, the first two nucleotides of which become visible (Fig. 2f; Supplementary Fig. S5f).

Within the remodeled active site, metal M2, which stabilizes the leaving group during exon ligation, adopts a distinctive coordination involving the phosphate of the first 3′-exon nucleotide and three phosphates from U6atac, contrasting with the simpler coordination seen in the major C* complex[32,35] (Fig. 2g; Supplementary Fig. S5g). Simultaneously, the nucleophile of exon ligation (i.e. 3′-OH of the 5′ exon) is repositioned to coordinate M1 and resides ~3.3 Å from the scissile phosphate, priming the complex for the second transesterification.

MMTAG2 stabilizes the branching conformation

During branching, we identify the conserved nuclear protein MMTAG2[29], previously

unrecognized in splicing, as an essential structural component that stabilizes the active site conformation (Fig. 3a; Supplementary Fig. S6a). Notably, MMTAG2 was originally identified as a candidate oncogene in multiple myeloma[29,36,37], suggesting a potential molecular link between the regulation of minor splicing and tumorigenesis. In the structure of the minor C complex, residues 7–83 of MMTAG2 are unambiguously resolved in an extended conformation that integrates deeply into the step-I assembly, where it directly binds to PRP8 RNaseH, CWC25, CWC22, and the intron sequences immediately downstream of the BPS (Fig. 3a,b; Supplementary Fig. S6b,c).

Structurally, the N-terminal segment (residues 9–13) of MMTAG2 forms a β-strand that pairs up with two β-strands of PRP8 RNaseH to create a hybrid β-sheet, a mechanism that contributes to locking the BPS/U12 duplex into its branching conformation by stabilizing PRP8 RNaseH together with FRG1 (Fig. 3b,c; Supplementary Fig. S6d). Downstream of this β-strand, MMTAG2 and PRP8 RNaseH jointly form a positively charged surface that lines the intron exit channel (Fig. 3d). MMTAG2 specifically recognizes the intron through hydrogen bonds (H-bonds) from Asn33 and Arg27 to the nucleobase of U210, while the guanidinium group of Arg27 forms a cation–π interaction with U211, collectively maintaining the local intron conformation (Fig. 3e). Further stability is provided by a C-terminal α-helix (residues 63–82) of MMTAG2, which packs against the central helix of CWC25, with this helical unit anchored by the MA3 domain of CWC22 (Supplementary Fig. S6c).

Through these multifaceted interactions, MMTAG2 is positioned to stabilize

both the BPS/U12 duplex and the surrounding protein network of the active site during branching. While conserved step-I factors such as YJU2, CWC25, and ISY1 stabilize core RNA elements in a manner analogous to the major spliceosome[27,38](Fig. 3f; Supplementary Fig. S6e-g), MMTAG2 appears to provide a distinct, pathway-specific stabilization critical for U12-type branching. Intriguingly, the step-II factor SLU7 is already recruited in this complex via its zinc-finger domain bound to the PRP8 N-domain (Fig. 3a; Supplementary Fig. S6h). Thus, MMTAG2 emerges as a dedicated architectural factor for the minor spliceosome catalysis and a compelling molecular candidate linking U12-type splicing to oncogenic processes.

WDR25 and FAM204A stabilize the exon ligation conformation

For exon ligation, the minor spliceosome is stabilized by a distinct set of proteins, which includes three newly characterized U12-specific factors, WDR25, FAM204A, and RBM41 in the structure of the minor C* complex (Fig. 4a). Notably, database analyses indicate both WDR25 and FAM204A as prognostic markers in multiple carcinomas[39], underscoring their potential clinical relevance and implicating a broader involvement of minor splicing in oncogenic pathways. WDR25 and FAM204A spatially replace and functionally mimic the major step-II factors PRP17/CDC40 and SDE2, respectively[32]. Occupying a substantial region within the minor C* complex, they contact PRP8, Cactin, CDC5L, SYF1, and SYF3, cooperatively locking the BPS/U12 duplex in the exon ligation conformation (Fig. 4b).

FAM204A adopts an extended conformation in its C-terminal half (residues

120–233), which is clearly resolved in our structure (Supplementary Fig. S7a,b). A coiled-coil segment and flanking helices interact with SYF1 C-terminus (Supplementary Fig. S7c), while the C-terminal tail projects deep into the splicing active site (Fig. 4b; Supplementary Fig. S7d). WDR25 contains a C-terminal WD40 domain and an N-terminal fragment (residues 172–182, referred as N-loop), which is anchored to ARMC7 through conserved, specific interactions (Fig. 4a; Supplementary Fig. S8a-d). Despite limited sequence identity, the WD40 domains of WDR25 exhibit high structural similarity to that of PRP17 and occupies equivalent positions in the C* complexes of both spliceosomes[32,35] (Fig. 4b; Supplementary Fig. S8e,f).

WDR25 contacts the tip of the BPS/U12 duplex and, together with the C-terminal region of FAM204A, forms a positively charged surface that stabilizes the duplex (Fig. 4c). Four conserved arginine residues in WDR25 (Arg460/Arg463/Arg464/Arg465) mediate H-bonds to the backbones of both U12 snRNA and the intron, while Arg503 interacts directly with the intronic nucleobases of G188 and A189 (Fig. 4d; Supplementary Fig. S9). A lysine-rich stretch in FAM204A further donates H-bonds to backbone phosphates of the intron and U12 snRNA (Fig. 4d). Thus, the WDR25-FAM204A pair collectively stabilizes the BPS/U12 duplex, facilitating the 3′SS docking during U12-type exon ligation.

Beyond the duplex stabilization, WDR25 and FAM204A also directly maintain the splicing active site during exon ligation (Fig. 4e). The two proteins sandwich and stabilize the β-finger of PRP8, which together with the 1585-loop (also known as α-finger[40]) clamps the lariat junction (Fig. 4e,f). In addition, the C-terminus

of FAM204A interacts with the 1585-loop and intercalates between helix Ia of the U12/U6atac duplex and the BPS/U12 duplex (Fig. 4e,g).

In summary, WDR25 and FAM204A represent key structural determinants that ensure the fidelity of U12-specific exon ligation. Their established role as cancer prognostic markers, coupled with their essential function in stabilizing the catalytic step-II conformation of the minor spliceosome, position them as promising molecular links between splicing regulation and disease mechanisms, including tumorigenesis.

Conserved mechanism of association and activation for PRP16 and PRP22

Our structures of the minor C and C* complexes enable the unambiguous assignment of the DEAH-box helicases PRP16 and PRP22 to their respective catalytic states (Fig. 5; Supplementary Fig. S10a-e). Both helicases are peripherally positioned along the pre-mRNA 3′ exit channel and resembles the association mode of PRP2 in the minor B^{act} complex[21] (Fig. 5a,b; Supplementary Fig. S10f).

In analogy to PRP2 that engages the spliceosomal RNP core via its CTD (binding SF3B1/SF3B3) and RecA2 (contacting BUD13), PRP16 employs its CTD to interact with the PRP8 Core domain and YJU2, while its RecA2 to bind CWC25 in the minor C complex (Fig. 5a,c; Supplementary Fig. S10c). In the minor C* complex, PRP22 similarly uses its CTD to bind the PRP8 Core, but its RecA2 to contact a previously uncharacterized protein RBM41, which in turn bridges PRP8 RNaseH and the step-II factor SLU7 (Fig. 5b,d). Moreover, the C-terminus of PRP22 extends along the lateral surface of PRP8 Core, placing it adjacent to the 1585-loop that directly

anchors the 3′SS in the active site (Fig. 5b,d; Supplementary Fig. S10e).

A key mechanistic divergence lies in helicase activation. Unlike PRP2, which strictly requires the G-patch coactivator GPKOW (Spp2 in yeast) that simultaneously tethers the RNP core and PRP2[41] (Supplementary Fig. S10f,g), PRP16 and PRP22 function without one. The structural basis for this coactivator independence has remained unclear, largely due to the poorly resolved N-terminal extensions in previous studies[27,32,33,42-44]. Our structures now reveal that PRP16 and PRP22 operate via a built-in, coactivator-independent mechanism.

This intrinsic mechanism is mediated by an N-terminal extension harboring two conserved motifs: N-fragment I and N-fragment II. N-fragment I, comprising an α-helix followed by a flexible loop, anchors to the RNP core; N-fragment II, containing two α-helices and an extended intervening loop, simultaneously engages the helicase's own CTD and RecA2 domain (Fig. 5a-d; Supplementary Fig. S10b,e,g). This binding mode mirrors the tethering function of GPKOW/Spp2 but is encoded entirely within the helicase itself. The sequences of N-fragment II are conserved between PRP16 and PRP22, with conserved residues contacting both the CTD and RecA2 domains (Fig. 5e). Most strikingly, the binding site for N-fragment II on PRP16 or PRP22 precisely overlaps with the site occupied by GPKOW/Spp2 on PRP2, explaining why no external G-patch protein is required.

Taken together, our structural and mechanistic analysis reveals that PRP16 and PRP22 have evolved an integrated regulatory architecture that bypasses the need for a separate G-patch coactivator. Concurrently, the recruitment of PRP22 is

specifically assisted by RBM41, which physically links the helicase to the SLU7-dependent step-II machinery, highlighting a unique mechanism for coordinating helicase action with catalytic progression in the minor spliceosome.

Discussion

Specified proteins orchestrate a conserved catalytic core

Our structures resolve the molecular basis of both catalytic steps in U12-type splicing. revealing that the reactions are driven by a conserved catalytic core: a two-metal-ion center organized by the U6atac ISL and the U12/U6atac helix I (Fig. 2), which together adopt a near-identical conformation to the corresponding U6 ISL and U2/U6 helix I in the major spliceosome. Meanwhile, despite the RNA-based catalytic core, the newly identified pathway-specific proteins add distinct layers of stabilization: MMTAG2 adopts an extended conformation to support the active site during branching (Fig. 3), whereas WDR25 and FAM204A directly position and stabilize the BPS/U12 duplex during exon ligation (Fig. 4).

Extending this principle of specific regulation, our structures uncover how the previously uncharacterized protein RBM41 mediates the recruitment of the helicase PRP22. RBM41 utilizes its N-terminal domain, composed of seven α-helices, to directly interacts with PRP22, consistent with a previous study[45], while simultaneously contacting BRR2, Cactin, SLU7, and PRP8 RNaseH (Supplementary Fig. S11). Through an extended structural module formed by helices α1/α2/α6/α7 and associated β-strands (Supplementary Fig. S11c-e), RBM41 effectively bridges PRP22

with the step-II machinery, thereby integrating helicase activity into the exon-ligation network (Supplementary Fig. S11c-e).

Together with the minor B^{act} structure[21], our current study also outlines a coherent remodeling pathway of protein displacement and recruitment that drives the U12-type catalytic progression (Supplementary Fig. S12). During the B^{act}-to-B^{*} transition, PRP2 and its cofactor GPKOW mediate the dissociation of the SF3b complex, SCNM1, and a set of splicing factors, thereby freeing the 5'SS and the BPS/U12 duplex. Subsequent rearrangements, along with the recruitment of six shared splicing factors and the unique component MMTAG2, translocate the BPS/U12 duplex into the vicinity of the splicing active site, enabling branching. Following PRP16 action, the step-I factors are dissociated, making way for nine step-II factors. Among these, RBM41, WDR25, and FAM204A are specific to the minor spliceosome, whereas several step-II factors essential in the major spliceosome (including PRP17, SDE2, PRKRIP1, FAM50A, CXORF56, NOSIP, and TLS1) are absent, highlights the compositional specialization of the minor spliceosome.

U6atac snRNA as a central organizer for U12-type catalysis

Structural analysis reveals that the unique protein composition of the minor spliceosome is likely governed by the catalytic component U6atac snRNA. Unlike U6 snRNA, U6atac lacks a 5′ stem loop and does not form helix II, two hallmarks of the major spliceosome catalytic center[32,46] (Fig. 6a). Instead, the sequences downstream of ISL fold into a central stem loop, a single-stranded segment that binds along the positively charged surface of SYF3, and a 3′ stem loop anchored by SYF1

(Supplementary Fig. S13a,b). Notably, the RNA segment emerging from the 3′ stem loop enters the heptameric LSm2-8 complex (Fig. 6b,c; Supplementary Fig. S5e). In contrast to the major spliceosome where the LSm2-8 complex dissociates during activation[30], it remains stably bound in the minor spliceosome through interactions with PRP19 core, SYF1, and Aquarius (Supplementary Fig. S13c).

These distinct structural features create a binding landscape that selectively recruits minor-specific factors. For example, in the major C* complex, PRP17 anchors its N-terminus to the 5′ stem loop of U6 snRNA and the NTR components RBM22 and BUD31, whereas its functional partner SDE2 is partly stabilized by SYF2, which binds to U2/U6 helix II (Fig. 6d,e). In the minor C* complex, the 5′ end sequences of U6atac snRNA are encapsulated by the RBM48-ARMC7 complex (Fig. 6b,c), which occludes the binding site for the N-domain of PRP17 but provides a binding interface for the N-loop of WDR25. Overall, of the nine unique U12-type splicing factors (excluding previously characterized U11/U12 di-snRNP proteins), five are directly linked to the specific conformation of U6atac snRNA, accounting for over half of these specialized components. Therefore, U6atac snRNA serves as a central organizer for the minor spliceosome.

Architectural insights define a regulatory hub in development

Moving beyond catalysis, our structural elucidation reveals that the minor spliceosome emerges not only as a parallel catalytic machine, but also as a critical regulatory node in the gene expression of multicellular eukaryotes. The identification of multiple disease-associated proteins within the machine further underscores its

physiological importance. Its precise regulation is essential for normal development: defects in its components are linked to developmental disorders[47],while its dysregulation may enable cells to bypass key checkpoints[29,36,37], thereby contributing to tumorigenesis. Several of the factors identified here are established cancer prognostic markers. Given that U12-type splicing is often rate-limiting for its target transcripts, the minor spliceosome act as a tunable gateway, making it particularly susceptible to both loss-of-function and gain-of-function disease mechanisms. Our work thus provides new molecular handles for understanding how splicing fidelity is maintained in health and disrupted in disease.

Acknowledgements

We thank Kui Xu for help during A^2-Net facilitated model building, and the staff at the cryo-EM facility, the high-performance computing center of Westlake University for technical assistance and support. This work was supported by funds from the National Natural Science Foundation of China (32522048 to R.B., 32571396 and 32171214 to R.W.), the Zhejiang Provincial Natural Science Foundation of China (LR25C050001 to R.B. and LDQ23C050001 to R.W.) and Start-up funds from Westlake University (to R.W.).

Author contributions: R.B. and R.W. designed the experiments. R.B. and H.G. purified the spliceosome and prepared cryo-EM samples. R.B., R.W., and H.G. collected the EM data. R.W. processed the EM data, calculated the EM map, and built

the atomic model. R.S. and Y.Z. performed MS analysis of the spliceosome samples. All authors contributed to data analysis. R.W., R.B. and Y.S. wrote the manuscript. R.W. supervised the project.

References:

1 Berget, S. M., Moore, C. & Sharp, P. A. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* **74**, 3171-3175 (1977). https://doi.org:10.1073/pnas.74.8.3171

2 Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **12**, 1-8 (1977). https://doi.org:10.1016/0092-8674(77)90180-5

3 Wright, C. J., Smith, C. W. J. & Jiggins, C. D. Alternative splicing as a source of phenotypic diversity. *Nat Rev Genet* **23**, 697-710 (2022). https://doi.org:10.1038/s41576-022-00514-4

4 Rogalska, M. E., Vivori, C. & Valcarcel, J. Regulation of pre-mRNA splicing: roles in physiology and disease, and therapeutic prospects. *Nat Rev Genet* **24**, 251-269 (2023). https://doi.org:10.1038/s41576-022-00556-8

5 Yan, C., Wan, R. & Shi, Y. Molecular Mechanisms of pre-mRNA Splicing through Structural Biology of the Spliceosome. *Cold Spring Harb Perspect Biol* **11** (2019). https://doi.org:10.1101/cshperspect.a032409

6 Patel, A. A. & Steitz, J. A. Splicing double: insights from the second spliceosome. *Nat Rev Mol Cell Biol* **4**, 960-970 (2003). https://doi.org:10.1038/nrm1259

7 Sharp, P. A. & Burge, C. B. Classification of introns: U2-type or U12-type. *Cell* **91**, 875-879 (1997). https://doi.org:10.1016/s0092-8674(00)80479-1

8 Will, C. L. & Luhrmann, R. Splicing of a rare class of introns by the U12-dependent spliceosome. *Biol Chem* **386**, 713-724 (2005). https://doi.org:10.1515/BC.2005.084

9 Turunen, J. J., Niemela, E. H., Verma, B. & Frilander, M. J. The significant other: splicing by the minor spliceosome. *Wiley Interdiscip Rev RNA* **4**, 61-76 (2013). https://doi.org:10.1002/wrna.1141

10 Patel, A. A., McCarthy, M. & Steitz, J. A. The splicing of U12-type introns can be a rate-limiting step in gene expression. *EMBO J* **21**, 3804-3815 (2002). https://doi.org:10.1093/emboj/cdf297

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

12 Tarn, W. Y. & Steitz, J. A. A novel spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vitro. *Cell* **84**, 801-811 (1996). https://doi.org:10.1016/s0092-8674(00)81057-0

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

14 Tarn, W. Y. & Steitz, J. A. Pre-mRNA splicing: the discovery of a new spliceosome doubles the

challenge. *Trends Biochem Sci* **22**, 132-137 (1997). https://doi.org:10.1016/s0968-0004(97)01018-9

15 Alioto, T. S. U12DB: a database of orthologous U12-type spliceosomal introns. *Nucleic Acids Res* **35**, D110-115 (2007). https://doi.org:10.1093/nar/gkl796

16 Olthof, A. M., Hyatt, K. C. & Kanadia, R. N. Minor intron splicing revisited: identification of new minor intron-containing genes and tissue-dependent retention and alternative splicing of minor introns. *BMC Genomics* **20**, 686 (2019). https://doi.org:10.1186/s12864-019-6046-x

17 Olthof, A. M. *et al.* Taxonomy of introns and the evolution of minor introns. *Nucleic Acids Res* **52**, 9247-9266 (2024). https://doi.org:10.1093/nar/gkae550

18 Yu, Y.-T., Scharl, E.C., Smith, C.M., and Steitz, J.A. *The growing world of small nuclear ribonucleoproteins*. 487–524 (Cold Spring Harbor Laboratory Press, 1999).

19 Will, C. L. *et al.* The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA* **10**, 929-941 (2004). https://doi.org:10.1261/rna.7320604

20 Bai, R. *et al.* Structural basis of U12-type intron engagement by the fully assembled human minor spliceosome. *Science* **383**, 1245-1252 (2024). https://doi.org:10.1126/science.adn7272

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

22 Bai, F. *et al.* RNA Binding Motif Protein 48 Is Required for U12 Splicing and Maize Endosperm Differentiation. *Plant Cell* **31**, 715-733 (2019). https://doi.org:10.1105/tpc.18.00754

23 Zuo, Y., Feng, F., Qi, W. & Song, R. Dek42 encodes an RNA-binding protein that affects alternative pre-mRNA splicing and maize kernel development. *J Integr Plant Biol* **61**, 728-748 (2019). https://doi.org:10.1111/jipb.12798

24 de Wolf, B. *et al.* Chromosomal instability by mutations in the novel minor spliceosome component CENATAC. *EMBO J* **40**, e106536 (2021). https://doi.org:10.15252/embj.2020106536

25 Suzuki, T. *et al.* The DROL1 subunit of U5 snRNP in the spliceosome is specifically required to splice AT-AC-type introns in Arabidopsis. *Plant J* **109**, 633-648 (2022). https://doi.org:10.1111/tpj.15582

26 Zhao, J., Peter, D., Brandina, I., Liu, X. & Galej, W. P. Structural basis of 5' splice site recognition by the minor spliceosome. *Mol Cell* **85**, 652-664 e654 (2025). https://doi.org:10.1016/j.molcel.2024.12.017

27 Bertram, K. *et al.* Structural Insights into the Roles of Metazoan-Specific Splicing Factors in the Human Step 1 Spliceosome. *Mol Cell* **80**, 127-139 e126 (2020). https://doi.org:10.1016/j.molcel.2020.09.012

28 Bertram, K. *et al.* Cryo-EM structure of a human spliceosome activated for step 2 of splicing. *Nature* **542**, 318-323 (2017). https://doi.org:10.1038/nature21079

29 Tian, J. Y. *et al.* Cloning and sequence analysis of tumor-associated gene hMMTAG2 from human multiple myeloma cell line ARH-77. *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)* **35**, 143-148 (2003).

30 Townsend, C. *et al.* Mechanism of protein-guided folding of the active site U2/U6 RNA during spliceosome activation. *Science* **370** (2020). https://doi.org:10.1126/science.abc3753

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

111024

32 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

33 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

34 Marcia, M. & Pyle, A. M. Visualizing group II intron catalysis through the stages of splicing. *Cell* **151**, 497-507 (2012). https://doi.org:10.1016/j.cell.2012.09.033

35 Zhang, X. *et al.* An Atomic Structure of the Human Spliceosome. *Cell* **169**, 918-929 e914 (2017). https://doi.org:10.1016/j.cell.2017.04.033

36 Luo, S. Q. *et al.* C1orf35 contributes to tumorigenesis by activating c-MYC transcription in multiple myeloma. *Oncogene* **39**, 3354-3366 (2020). https://doi.org:10.1038/s41388-020-1222-7

37 Meier, T., Timm, M., Montani, M. & Wilkens, L. Gene networks and transcriptional regulators associated with liver cancer development and progression. *BMC Med Genomics* **14**, 41 (2021). https://doi.org:10.1186/s12920-021-00883-5

38 Zhan, X., Yan, C., Zhang, X., Lei, J. & Shi, Y. Structure of a human catalytic step I spliceosome. *Science* **359**, 537-545 (2018). https://doi.org:10.1126/science.aar6401

39 Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science* **357** (2017). https://doi.org:10.1126/science.aan2507

40 Nguyen, T. H. D. *et al.* Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 A resolution. *Nature* **530**, 298-302 (2016). https://doi.org:10.1038/nature16940

41 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

42 Garbers, T. B., Enders, M., Neumann, P. & Ficner, R. Crystal structure of Prp16 in complex with ADP. *Acta Crystallogr F Struct Biol Commun* **79**, 200-207 (2023). https://doi.org:10.1107/S2053230X23005721

43 Felisberto-Rodrigues, C. *et al.* Structural and functional characterisation of human RNA helicase DHX8 provides insights into the mechanism of RNA-stimulated ADP release. *Biochem J* **476**, 2521-2543 (2019). https://doi.org:10.1042/BCJ20190383

44 Hamann, F., Enders, M. & Ficner, R. Structural basis for RNA translocation by DEAH-box ATPases. *Nucleic Acids Res* **47**, 4349-4362 (2019). https://doi.org:10.1093/nar/gkz150

45 Norppa, A. J. *et al.* Distinct functions for the paralogous RBM41 and U11/U12-65K proteins in the minor spliceosome. *Nucleic Acids Res* **52**, 4037-4052 (2024). https://doi.org:10.1093/nar/gkae070

46 Hang, J., Wan, R., Yan, C. & Shi, Y. Structural basis of pre-mRNA splicing. *Science* **349**, 1191-1198 (2015). https://doi.org:10.1126/science.aac8159

47 Norppa, A. J., Shcherbii, M. V. & Frilander, M. J. Connecting genotype and phenotype in minor spliceosome diseases. *RNA* **31**, 284-299 (2025). https://doi.org:10.1261/rna.080337.124

48 Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci* **27**, 14-25 (2018). https://doi.org:10.1002/pro.3235

49 Wu, Q. & Krainer, A. R. Purine-rich enhancers function in the AT-AC pre-mRNA splicing pathway and do so independently of intact U1 snRNP. *Rna* **4**, 1664-1673 (1998). https://doi.org:Doi 10.1017/S1355838298981432

50 Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. Accurate transcription initiation by RNA

polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res* **11**, 1475-1489 (1983). https://doi.org:10.1093/nar/11.5.1475

51 Bai, R. *et al.* Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. *Science* (2020). https://doi.org:10.1126/science.abe8863

52 Bai, R., Yan, C., Wan, R., Lei, J. & Shi, Y. Structure of the Post-catalytic Spliceosome from Saccharomyces cerevisiae. *Cell* **171**, 1589-1598 e1588 (2017). https://doi.org:10.1016/j.cell.2017.10.038

53 Zhu, T. *et al.* DPHL: A DIA Pan-human Protein Mass Spectrometry Library for Robust Biomarker Discovery. *Genomics Proteomics Bioinformatics* **18**, 104-119 (2020). https://doi.org:10.1016/j.gpb.2019.11.008

54 Wang, L. H. *et al.* pFind 2.0: a software package for peptide and protein identification via tandem mass spectrometry. *Rapid Commun Mass Spectrom* **21**, 2985-2991 (2007). https://doi.org:10.1002/rcm.3173

55 Bai, R., Wan, R., Yan, C., Lei, J. & Shi, Y. Structures of the fully assembled Saccharomyces cerevisiae spliceosome before activation. *Science* **360**, 1423-1429 (2018). https://doi.org:10.1126/science.aau0325

56 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332 (2017). https://doi.org:10.1038/nmeth.4193

57 Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153-1160 (2019). https://doi.org:10.1038/s41592-019-0575-8

58 Zhang, X. *et al.* Structures of the human spliceosomes before and after release of the ligated exon. *Cell Res* **29**, 274-285 (2019). https://doi.org:10.1038/s41422-019-0143-x

59 Agafonov, D. E. *et al.* Molecular architecture of the human U4/U6.U5 tri-snRNP. *Science* **351**, 1416-1420 (2016). https://doi.org:10.1126/science.aad2085

60 Zhu, J. *et al.* A minority of final stacks yields superior amplitude in single-particle cryo-EM. *Nat Commun* **14**, 7822 (2023). https://doi.org:10.1038/s41467-023-43555-x

61 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

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

63 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004). https://doi.org:10.1107/S0907444904019158

64 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612 (2004). https://doi.org:10.1002/jcc.20084

65 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589 (2021). https://doi.org:10.1038/s41586-021-03819-2

66 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

67 Xu, K., Wang, Z., Shi, J. P., Li, H. S. & Zhang, Q. C. A(2)-Net: Molecular Structure Estimation from Cryo-EM Density Volumes. *Thirty-Third Aaai Conference on Artificial Intelligence / Thirty-First*

Innovative Applications of Artificial Intelligence Conference / Ninth Aaai Symposium on Educational Advances in Artificial Intelligence, 1230-1237 (2019).

68 Zhou, L. *et al.* Crystal structures of the Lsm complex bound to the 3' end sequence of U6 small nuclear RNA. *Nature* **506**, 116-120 (2014). https://doi.org:10.1038/nature12803

69 Wang, W. *et al.* trRosettaRNA: automated prediction of RNA 3D structure with transformer network. *Nat Commun* **14**, 7266 (2023). https://doi.org:10.1038/s41467-023-42528-4

70 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

71 Davis, I. W. *et al.* MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* **35**, W375-383 (2007). https://doi.org:10.1093/nar/gkm216

Figure Legends

Fig.1 Cryo-EM structures of the human minor C and C* complexes. **a,** A schematic diagram for the optimization of the U12-type pre-mRNA substrate. The final pre-mRNA substrate MSE-U12 differs from the original substrate MINX-U12 in four elements: BPS, sequences preceding the 3′SS, 3′ exon, and a splicing enhancer. **b,** Comparison of the splicing efficiency between MINX-U12 (left panel) and MSE-U12 (right penal). Shown here are the RT-PCR results of *in vitro* splicing reactions. **c,** Structure of the human minor C complex. Shown here is a composite EM map, where the spliceosomal components are color-coded and tabulated below. **d,** Structure of the human minor C* complex. Shown here is a color-coded, composite EM map. The state-specific splicing factors are color-coded. Proteins specific to the structure of the major C/C* complexes are indicated by dashed gray boxes. All structural images were created using ChimeraX[48].

Fig.2 The RNA elements in the human minor C and C* complexes. **a,** Overall structure of the RNA elements in the minor C complex. U12, U5, and U6atac snRNAs are colored marine, orange, and green, respectively. The 5′ exon and intron of MSE-U12 are colored red and violet, respectively. **b,** Key RNA elements in the catalytic center of the minor C complex. The BP-A (purple) and the guanine residue at the 5′ end of the 5′SS are shown in stick representation. The 2′-5′ phosphodiester linkage is already formed. **c,** Coordination of the catalytic metals in the minor C complex. Both catalytic metals M1 and M2 are coordinated by phosphate groups of U6atac

snRNA. M2, which activates the nucleophile during branching reaction, is also coordinated by the phosphate group of the first adenine residue (A1) at 5′SS. The terminal guanine residue (G-1) at the 3′ end of the 5′ exon is flipped and no longer coordinates M1 that stabilizes the leaving group. **d,** Overall structure of the RNA elements in the minor C* complex. The RNA elements are colored the same as those in panel a, except for the 3′ exon, which is colored slate blue. In the minor C* complex, the 3′SS is docked into the active site for exon ligation. The 3′-half of U6atac snRNA can be resolved. **e,** Structural comparison of U12 snRNA and pre-mRNA between the minor C and C* complexes. The RNA elements in the minor C complex are colored the same as in panel a, whereas those in the C* complex are colored in light shades. During the transition from C to C* complex, the U12/BPS duplex undergoes an ~80° rotational rearrangement. **f,** Key RNA elements in the catalytic center of the minor C* complex. The orientation is the same as that in panel b. The 2′-5′ linkage is repositioned outside the active site to accommodate 3′SS docking. **g,** Coordination of the catalytic metals in the minor C* complex. M2, which stabilizes the leaving group during exon ligation, is coordinated by the terminal cytosine residue (C-1) at 3′SS and U6atac snRNA. The nucleophile (the 3′-OH of 5′ exon G-1) coordinates M1, is positioned 3.3 Å from the scissile phosphate and primed for the exon ligation.

Fig.3 MMTAG2 may stabilize the step-I conformation in the human minor C complex. **a,** An overview of key proteins that are recruited to and functionally

important for the minor C complex. The newly identified protein MMTAG2 is colored green. The overall structure of the minor C complex is displayed in the background with faded colors. **b,** MMTAG2, FRG1, and CWC25 associate with PRP8 RNaseH domain to facilitate the stabilization of the BPS/U12 duplex. The N-terminus of MMTAG2 projects toward the U12/BPS duplex. **c,** A close-up view on the N-terminus of MMTAG2, which inserts into the RNaseH domain of PRP8. Residues 9 to 13 of MMTAG2 constitute a β-strand that pairs up with a β-strand from PRP8 RNaseH. **d,** The positively charged surface formed by MMTAG2 and PRP8 RNaseH may help orient the intron sequences downstream the BPS. **e,** A close-up view on the interactions between MMTAG2 and the intron. **f,** An overlay of the active site RNA elements and surrounding proteins between the human minor and major C (PDB-8I0W[33]) complexes. In both complexes, the active site RNA elements are stabilized by CWC25, YJU2, and ISY1.

Fig.4 WDR25 and FAM204A stabilize the step-II conformation for catalysis in the human minor C* complexes. **a,** An overview of key proteins that are recruited to and functionally important for the minor C* complex. Three newly identified U12-type step 2 factors WDR25, FAM204A, and RBM41 are colored orange, blue, and light green, respectively. The minor C* complex is displayed in the background. **b,** FAM204A and the C-terminal WD40 domain of WDR25 bind within the cavity formed by PRP8, Cactin, CDC5L, SYF1, and SYF3, thereby stabilizing the BPS/U12 duplex in the minor C* complex. **c,** The BPS/U12 duplex is bound by a positively

charged surface that is formed by the C-terminus of FAM204A and the WD40 domain of WDR25. **d,** A close-up view on the recognition of the BPS/U12 duplex and the region upstream of BPS by WDR25 and FAM204A. **e,** WDR25 and FAM204A contact the β-finger and 1585-loop of PRP8 and stabilize the active site conformation for exon ligation in the minor C* complex. The β-finger and 1585-loop are colored purple and hot pink, respectively. **f,** A close-up view on the interactions between the β-finger and its surrounding proteins. The β-finger of PRP8 is sandwiched by WDR25 and FAM204A. **g,** A close-up view on the interactions between the 1585-loop and its surrounding proteins.

Fig.5 Conserved mechanism of spliceosome association and helicase activation for PRP16 and PRP22. **a,** Structure of PRP16 in the minor C complex and its interactions with surrounding components. Two views are shown. PRP16 is anchored to the minor C complex through interactions with PRP8 (wheat), BRR2 (wheat), CDC5L (aquamarine), and two step I factors YJU2 (medium violet) and CWC25 (purple). The possible path of the 3′ region of the pre-mRNA is indicated as violet dashed lines. The N-fragment I (residues 365-403) and N-fragment II (residues 412-447) of PRP16 are colored light sky blue and coral, respectively. **b,** Structure of PRP22 in the minor C* complex and its interactions with surrounding components. PRP22 is anchored to the minor C* complex through interactions with PRP8 (wheat) and RBM41 (pale green). The possible path of the 3′ exon sequences is indicated as dashed slate lines. The N-fragment I (residues 396-435) and N-fragment II (residues

444-479) of PRP22 are colored deep sky blue and yellow, respectively. **c,** A schematic diagram of the sequence features of PRP16 in the minor C complex. Regions included in the atomic model are indicated by solid lines above the sequence. The interacting proteins are tabulated below the sequences. **d,** The sequence features and the interacting proteins of PRP22 in the minor C* complex. **e,** Sequence alignment between the N-fragment II of PRP16 and that of PRP22. Conserved sequences are boxed. Identical residues are shaded red. Residues involved in the interactions with the respective CTD and RecA2 are conserved between PRP16 and PRP22.

Fig.6 U6atac snRNA plays an essential role in recruiting the U12-type specific step II factors. **a,** A structural overlay between U6atac snRNA (green) of the human minor C* complex and U6 snRNA (light gray) of the human major C* complex (PDB-8C6J[32]). These two complexes are aligned using their respective ISLs. **b,** A number of key proteins stabilize U6atac snRNA and help recruit WDR25 and FAM204A in the minor C* complex. **c,** A schematic diagram of U6atac snRNA and surrounding proteins in the minor C* complex. **d,** Five proteins bind to U6 snRNA and help recruit PRP17 and SDE2 in the major C* complex. The orientation and scale are the same as the structure in panel b. **e,** A schematic diagram of U6 snRNA and surrounding proteins in the major C* complex.

Materials and Methods

Preparation of the pre-mRNA

The U12-type pre-mRNAs for the *in vitro* splicing assay were modified from the MINX and SCN4A genes. MINX-U12, used as a substrate for purified the human minor B^{act} complex[21], was added an exonic splicing enhancer (ESE) sequence at the 3′ end of the 3′ exon (Supplementary Fig. S1a, upper panel).The ESE sequence is 5′-AGGAUCCGGAAGAAUU-3′, which reported as ENH1[49]. MS-U12, retained the major sequences of MINX-U12, was modified the BPS and 3′SS sequences, and replaced the 3′ exon by a 43-nucleotide (nt) sequence of the SCN4A exon. MSE-U12 placed an ENH1 enhancer element at the 3′ end of MS-U12 (Supplementary Fig. S1a, bottom panel). The three tandem MS2-binding RNA aptamers were inserted between the 5′SS and the BPS. The DNA templates for *in vitro* transcription were generated using PCR, and the RNA substrates were synthesized using the method of T7 runoff transcription and analyzed on a urea-PAGE gel (Supplementary Fig. S1b).

In vitro splicing assay and RT-PCR

Nuclear extract from HeLa S3 cell lines was prepared for *in vitro* splicing as described[50]. 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 the

U12-type pre-mRNAs, 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 ℃ for 30 minutes prior to the splicing reaction (Supplementary Fig. S1c,d). This was accomplished by incubating the reaction with 1.5 μM antisense DNA oligonucleotides for each of six snRNAs. The sequences of the antisense DNA oligonucleotides (Sangon Biotech) were as described[21]. The splicing reaction mixture was incubated at 30 ℃ for varying time points of 0 minute, 30 minutes, 60 minutes and 120 minutes, 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 3% (w/v) agarose gel and stained by GoldView (Beijing SBS Genetech Co Ltd.) (Supplementary Fig. S1c,d).

Expression and purification of human PPR16

The purification protocol of spliceosomal ATPase/helicase is as described[20,51]. The optimized coding sequences of human full-length PRP16 (DHX38) was synthesized by GenScript®. The dominant negative mutant PRP16- K561A, which is unable to function as an ATPase/helicase, can block the splicing reaction after catalytic step I, resulting in the spliceosome stalling at the C complex (Supplementary Fig. S2a). The DNA fragment encoding PRP16 with a K561A mutation was cloned into the pESC-

TRP vector with a C-terminal Flag tag. The construct was transformed into the a wild-type (WT) *Saccharomyces cerevisiae (S. cerevisiae)* strain JDY52 *(trp-)*[52] by the lithium acetate method. Correct transformants were selected on minimal medium lacking tryptophan (Coolaber Science & Technology). Cells were grown to an OD_{600} of 1.8~2.0 at 30 ℃ in the medium supplemented with 2% (w/v) raffinose, and then pelleted and resuspended in fresh medium supplemented with 2% (w/v) galactose to induce overexpression of PRP16 mutant. After 14~16 hours in the galactose medium, the yeast cells were collected and resuspended in lysis buffer that contains 25 mM HEPES-KOH, pH 7.4, 500 mM NaCl, 1.5 mM $MgCl_2$ and 20% glycerol, and then disrupted by the SPEX 6870 Freezer Mill. The recombinant proteins were purified using the anti-Flag M2 affinity gel (Millipore). The eluted proteins were fractionated through a heparin column (GE Healthcare) to remove the non-specifically bound nucleic acids. Finally, the proteins were applied to gel filtration (Superdex-200 10/300 GL, GE Healthcare) in the buffer containing 20 mM HEPES-KOH, pH 7.9, 500 mM NaCl and 5% (v/v) glycerol. The peak fractions were analyzed on an SDS-PAGE gel (GenScript®) (Supplementary Fig. S2a) and stored at -80 ℃.

Assembly and purification of the human minor C complex

The protocol for assembly of the minor C complex was modified from that for the major C complex[27] and the minor pre-B complex[20] (Supplementary Fig. S2a). 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, 3 mM $MgCl_2$ and 0.5 μM PRP16- K561A, 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 minutes on ice. To reduce contamination by the major spliceosome, we depleted U1, U2 and U6 snRNAs in the nuclear extract using endogenous RNaseH at 30 ℃ for 30 minutes prior to the splicing reaction. The reaction mixture was incubated for 2 hours at 30 ℃, and then centrifuged at 3,000g for 15 minutes to remove aggregates. The supernatant was loaded onto amylose resin (NEB), and washed using the G100K buffer (10 mM HEPES-KOH, pH 7.9, 100 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%–0.05% glutaraldehyde (Sigma), and centrifuged at 23,000 rpm for 10 hours at 4 ºC in a SW32Ti rotor. The fractions that contained the minor C complex were quenched by 25 mM Tris (pH 7.6), 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 analyzed on a 10% urea PAGE gel for detection of the RNA elements and concentrated for cryo-EM studies (Supplementary Fig. S2b,c).

Assembly and purification of the human minor C* complex

The protocol for assembly and purification of the minor C* complex was modified

from that for the major C^* complex[28] and the minor B^{act} complex[21] (Supplementary Fig. S2d). Briefly, the splicing reaction was performed in a volume of 40 mL or multiples thereof, containing 4 mM HEPES-KOH, pH 7.9, 16 mM MES-NaOH, pH 6.4, 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 purification method is similar with that of the minor C complex as above. The final sample was analyzed on denaturing RNA gels and concentrated for cryo-EM studies (Supplementary Fig. S2e,f).

Mass spectrometry analysis

About 40 µl of the minor spliceosome samples were 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 minutes and resolved using a 4%–12% gradient SDS-PAGE gel. The proteins were subjected to in-gel proteolytic digestion as described[53]. 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 minutes 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 milliseconds. Dynamic exclusion was applied with a repeat count of 1 and an exclusion time of 25 seconds. The resultant mass spectrometric data were analyzed using pFind[54] (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 C and C* complexes are listed in Supplementary Table S1.

EM data acquisition and pre-processing

Cryo-EM grids for data collection were prepared using Vitrobot Mark IV (FEI Company) at 8 ºC and 100% humidity, largely as described[55]. Briefly, 4-µl aliquots of the sample at a concentration of ~0.2 mg/mL were applied to Quantifoil R2/1 grids coated with homemade continuous carbon film of ~2 nm thickness, which are glow-discharged for 30 seconds using the "Low" setting of the Plasma Cleaner (Harrick, Plasma Cleaner PDC-32G). After blotted for 1.5 seconds using the standard Vitrobot filter paper (Ø55/20mm, Ted Pella), the grids were plunged into liquid ethane cooled by liquid nitrogen.

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 0.53865 Å

(Supplementary Fig. S2b,e). Each image was dose-fractionated to 32 frames with a dose rate of ~22.49 counts/sec per physical pixel (~19.38 e-/sec per $Å^2$) and a total exposure time of 2.58 seconds. Total electron dose for each image is about 50 $e^-/Å^2$. All the data were collected using EPU (ThermoFisher Scientific) with a preset defocus range from -1.3 to -1.9 μm. Using the sample of human minor C complexes, we collected a total of 34,090 micrographs, comprising an earlier dataset (dataset 1) of 15,788 micrographs and a later dataset (dataset 2) of 18,302 micrographs. For human minor C^* complexes, the total micrographs collected were 58,431, consisting of an earlier dataset (dataset 1) of 22,181 micrographs and a later dataset (dataset 2) of 36,250 micrographs.

The image stacks were motion-corrected using MotionCor2[56] and binned to a pixel size of 1.0773 Å. The dose-weighted micrographs were processed for Patch CTF estimation in CryoSPARC (v4.3.0). Micrographs were discarded if showing obvious ice contamination, excessive drift, or damage.

Cryo-EM data processing

The data processing pipeline is summarized in Supplementary Figs. S3 and S4. For human minor C complex, we used dataset 1 for initial data analysis (Supplementary Fig. S3a). To avoid missing the spliceosomal particles, we used a reduced threshold for Topaz[57] auto-picking, yielding 2,569,998 particle coordinates. Particles were initially extracted using a pixel size of 8.6184 Å and a box size of 80 pixels. To

identify spliceosomal particles, we performed a heterogeneous refinement using reference volumes representing the human minor B^{act} complex (EMDB code: EMD-30875[21]), the major C complex (EMDB code: EMD-6864[38]), the major C^{*} complex (EMDB code: EMD-6721[35]), the major P complex (EMDB code: EMD-9645[58]), the major ILS complex (EMDB code: EMD-9647[58]), the major tri-snRNP (EMDB code: EMD-6581[59]), the ribosome and three bad classes from previous results, which were were low-pass filtered to 20 Å. We further performed 2D classifications on each class from the heterogeneous refinement to enrich particles with reasonable size. 330,292 particles were selected and subjected to an additional round of heterogeneous refinement with the initial references. Only class derived from C reference shows fine features. The resulting 49,347 particles were re-extracted with a 1x binning factor and processed for non-uniform (NU) refinement, yielded a reconstruction at an average resolution of 3.54 Å, representing the human minor C complex, featured by RBM48-ARMC7 bound at 5'-end of U6atac snRNA.

To enhance the accuracy of particle picking, we performed template picking on dataset 1 and 2 using templates created from the 3.54 Å reconstruction, resulting in 2,007,824 and 2,408,114 particles, respectively. Subsequently, multiple rounds of seed-facilitated 2D classification were performed to maximize the utilization of the collected datasets and achieve rapid preliminary particle alignment. We initially generated 500,000 simulated high-quality particles using the 3.54 Å reconstruction as our seed dataset. The template-picked particles from dataset 1 and 2 were divided into four and five subgroups respectively, with each subgroup containing a similar number

of particles as the seed dataset. Each subgroup was then mixed with the simulated seed dataset and subjected to 2D classification to select particles with fine features. After removing simulated particles, 402,437 particles from dataset 1 and 987,302 particles from dataset 2 were re-extracted with a 4x binning factor and processed through similar seed-facilitated 2D classifications, yielding a total of 1,014,725 particles (263,354 from dataset 1 and 751,371 from dataset 2).

These particles were re-extracted with 2x binning factor and subjected to three parallel runs of heterogeneous refinement to obtain more particles. The 3.54 Å map was rescaled to a pixel size of 2.1546 Å and used as a "good" reference, while three other maps generated during processing of dataset 1 served as bad references. Particles from good classes were merged and duplicated particles were removed. The resulting 357,030 particles were re-extracted at a pixel size of 1.0773 Å and processed for NU refinement, yielding a reconstruction at an average resolution of 3.43 Å. After one round of global 3D classification without image alignment, 210,708 particles were selected, processed for NU refinement and CTF refinement, improving the resolution to 3.07 Å for the entire human minor C complex. We further selected 86,306 particles through an iterative particle sorting algorithm call CryoSieve[60], resulting in a reconstruction at an overall resolution of 2.92 Å (Supplementary Fig. S3a,b). The local resolution reaches 2.5 Å in the core region.

In contrast to the core region, the peripheral regions of the human minor C complex exhibit flexibility. To improve map quality, we performed localized 3D classification and refinement on eight regions using region-specific soft masks. This

approach yielded high-resolution densities for seven regions: the RBM48-ARMC7 region (3.07 Å; 58,941 particles), the PRP8 RNaseH-like domain (3.48 Å; 37,176 particles), the EJC (3.59 Å; 51,437 particles), PRP16 (3.29 Å; 50,530 particles), and the region encompassing BRR2 and PRP16 (3.97 Å; 72,597 particles), the IBC region (4.41 Å; 85,373 particles), the SYF1-SYF3 region (3.21 Å; 85,373 particles). However, flexibility limited the resolution of the PRP19-associated region to a level that revealed only its global architecture. Overall, these targeted refinements significantly enhanced the local resolution of peripheral features compared to the global 2.92-Å map.

For the human minor C* complex, a similar data processing procedure was carried out with slight modification (Supplementary Fig. S4a). After multiple rounds of heterogeneous refinement and 2D classification of 3,162,134 Topaz-picked particles, 39,581 particles, yielding a reconstruction at 4.38-Å resolution. The same procedure of template picking, seed-facilitated 2D classification, hetero-refinement, and 3D classification was performed as described above. A total of 145,908 particles were selected, re-extracted with a pixel size of 1.0773 Å, and subjected to NU and CTF refinement, yielding a 3.04-Å reconstruction. Local resolutions were improved via focused classification and refinement for several regions: RBM48-ARMC7 (3.19 Å; 88,127 particles), SYF1/3 (3.48 Å; 107,934 particles), EJC (4.01 Å; 61,936 particles), PRP22 (4.23 Å; 95,438 particles). For BRR2, the BPS-proximal WD40 protein, and IBC, we performed NU refinements using the region-specific particle subsets to improve local map quality. Focused classification of the PRP19-associated

region (22,629 particles) revealed its overall fold and the Sm-like heptameric complex, with a local resolution of 7.71 Å after refinement. (Supplementary Fig. S4a,b).

The above reported resolutions were calculated according to the FSC 0.143 criterion with a high-resolution noise substitution method[61]. Prior to visualization, all density maps were sharpened by applying a negative B-factor that was estimated using automated procedures[62] in cryoSPARC. Local resolution variations were estimated using cryoSPARC (Supplementary Tables S2 & S3).

Model building and refinement

Model building of minor C and C* complex was carried out using COOT[63] and UCSF Chimera[64]. We combined appropriate modeling methods including docking and manual adjustment, *de novo* modeling, homology modeling, AI-facilitated modeling, and rigid-body docking according to resolution levels of the various local maps to generate the atomic models (Supplementary Tables S4 & S5).

For the minor C complex, components were identified using the atomic coordinates of the human minor B^{act} complex (PDB code: 7DVQ[21]) and the human major C complex (PDB code: 8I0W[33]). For initial model building, we first aligned the locally refined maps with the consensus map to integrate high-resolution information. This composite map was then used to guide the docking of relevant structures from the human minor B^{act} complex (PDB code: 7DVQ[21]), the major C complex (PDB

code: 8I0W[33]), and the AlphaFold-predicted models[65] of PRP16 and FRG1 into the maps using UCSF Chimera[64]. The docked models were subsequently truncated as needed, manually adjusted and extended (if necessary) using COOT[63]. Structural models for the BRR2 and IBC regions were generated by rigid-body docking of coordinates from the major C complex, while the U12 Sm ring and associated RNA were docked from the minor B^{act} complex. A model of the complex comprising the PRP19-tetramer, SPF27, the C-terminal fragment of CDC5L, and the N-terminus of PRP46 was generated by docking a refined Alphafold3-predicted complex model[66], which was manually truncated and adjusted against the local map for this region in the minor C^* complex.

After these steps, several unassigned EM density patches remained, including regions adjacent to the PRP8 RNaseH-like domain and the C-half of CWC22. The A2-Net method[67] was applied to its density to recognize amino acid side chains, which, combined with MS data, identified the protein as MMTAG2. The AlphaFold model of the N-terminal fragment of MMTAG2 fitted well into the density and was manually adjusted, revealing that its downstream sequence extends into the CWC22-proximal density. An AlphaFold3-based interaction screen of the CWC22-MMTAG2 interface identified a fragment of CWC25 as the interacting partner occupying this density. Such method was used to model other missing fragments, including CDC5L (residues 461-487), SRm300 (residues 122~161), and CCDC12 (residues 107~150), into remaining densities.

Model building of the human minor C^* complex followed a procedure

analogous to that described above. Based on the EM maps of the human minor C* complex, we selected appropriate segments of the atomic coordinates from the human minor B^{act} complex (PDB code: 7DVQ[21]) and the major C* complex (PDB code: 8C6J[32]), and the AlphaFold2 models of SYF1, SYF3, PRP22, Aquarius, and PPIE, docked them into the EM densities and performed manual adjustments, extensions, or rigid-body docking as required (Supplementary Table S5). To determine the identity of the unassigned density in the vicinity of the BPS/U12 duplex, A2-Net-facilitated modeling[67] was employed. This approach generated the models for FAM204A and WDR25. Additionally, the A2-Net method was used to model interacting fragments of SKIP and PLRG1 with the RBM48-ARMC7 complex.

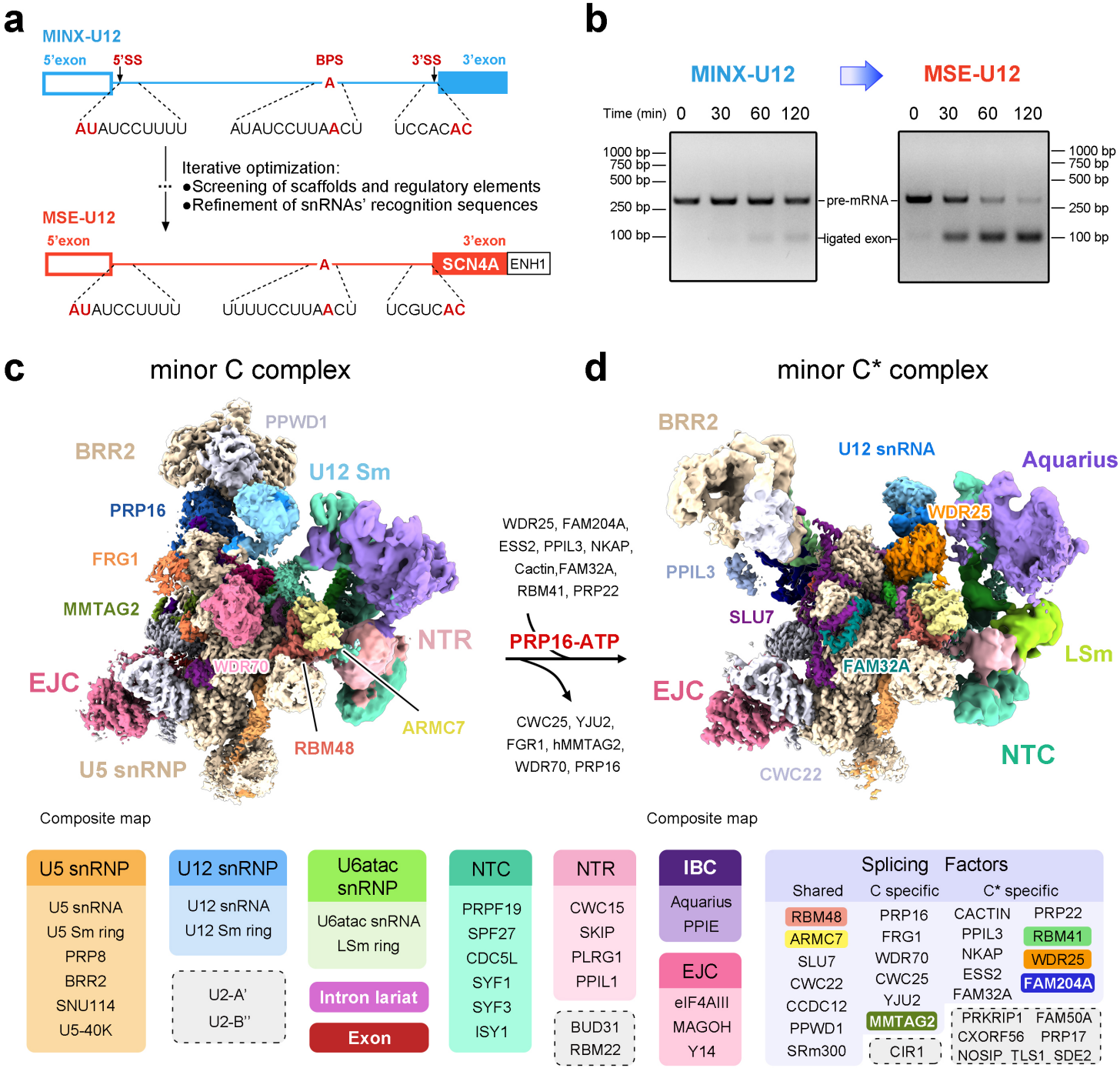
To identify the binding partner of PRP22 RecA2 domain in the minor C* complex, we employed AlphaFold3[66] to screen against PRP22. This analysis predicted a complex between PRP22 and the N-terminal region of RBM41, which was validated by its exceptional fit to the adjacent density adjacent to PRP22 and consistent with proximity-labeling data[45]. Additional fragments, including CDC5L (residues 289–327, 359–387, 461–487), SRm300 (residues 122–161), and CCDC12 (residues 107–150), were incorporated based on AlphaFold3 predictions. For the PRP19-associated region, a complex model was predicted using AlphaFold3, docked as a rigid body, and manually refined according to local map. The model for 3′ region of U6atac snRNA and associated LSm2-8 heptameric complex were generated by homology modeling based on the *S. cerevisiae* Lsm2-8-U6 snRNA complex structure (PDB code: 4M7A[68]) and prediction of structure for 3′ stem loop of U6atac snRNA by

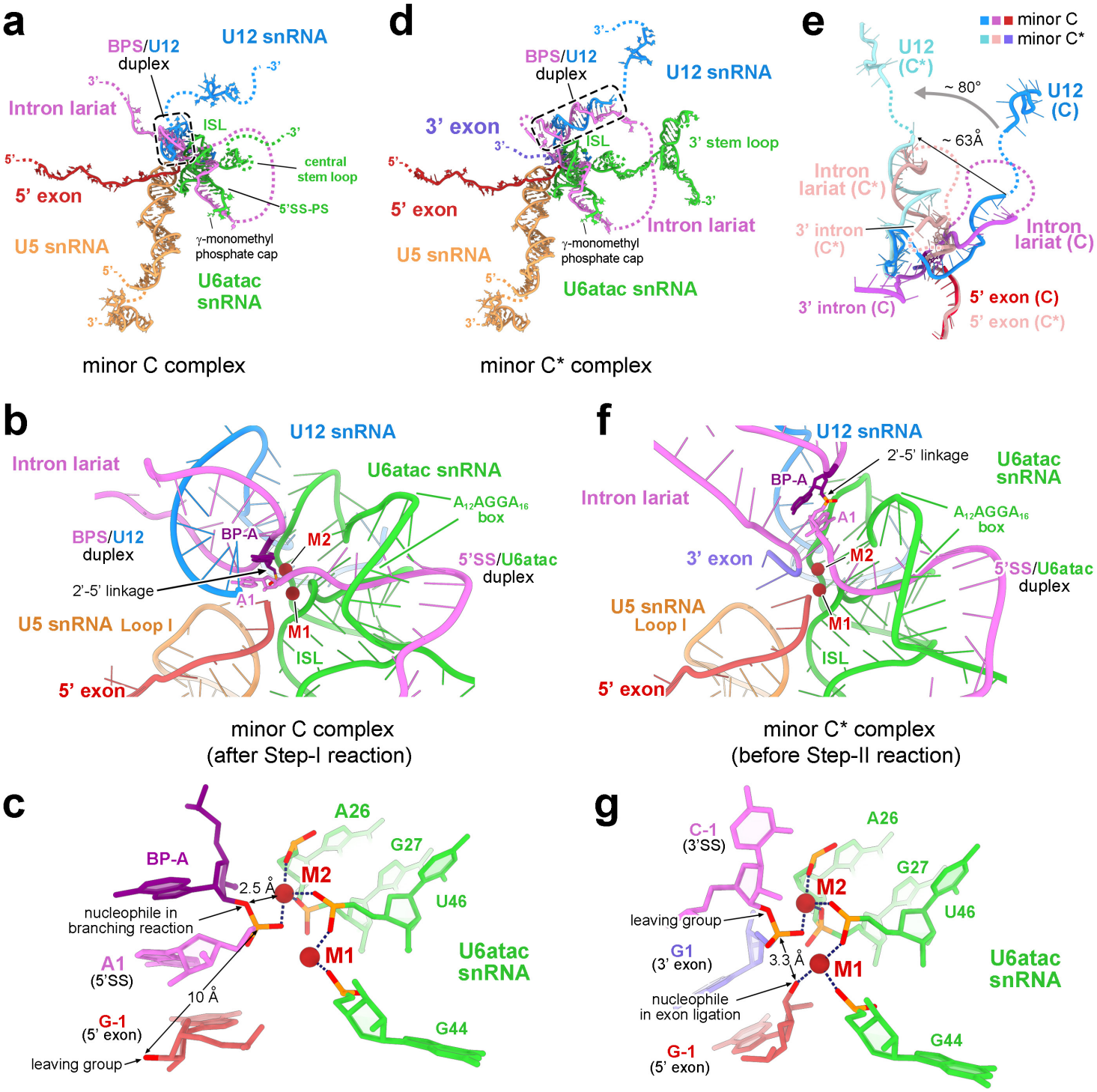
trRosettaRNA[69]. After docking the two models, we manually built extensions for U6atac snRNA (residues 66-79, 117-120).

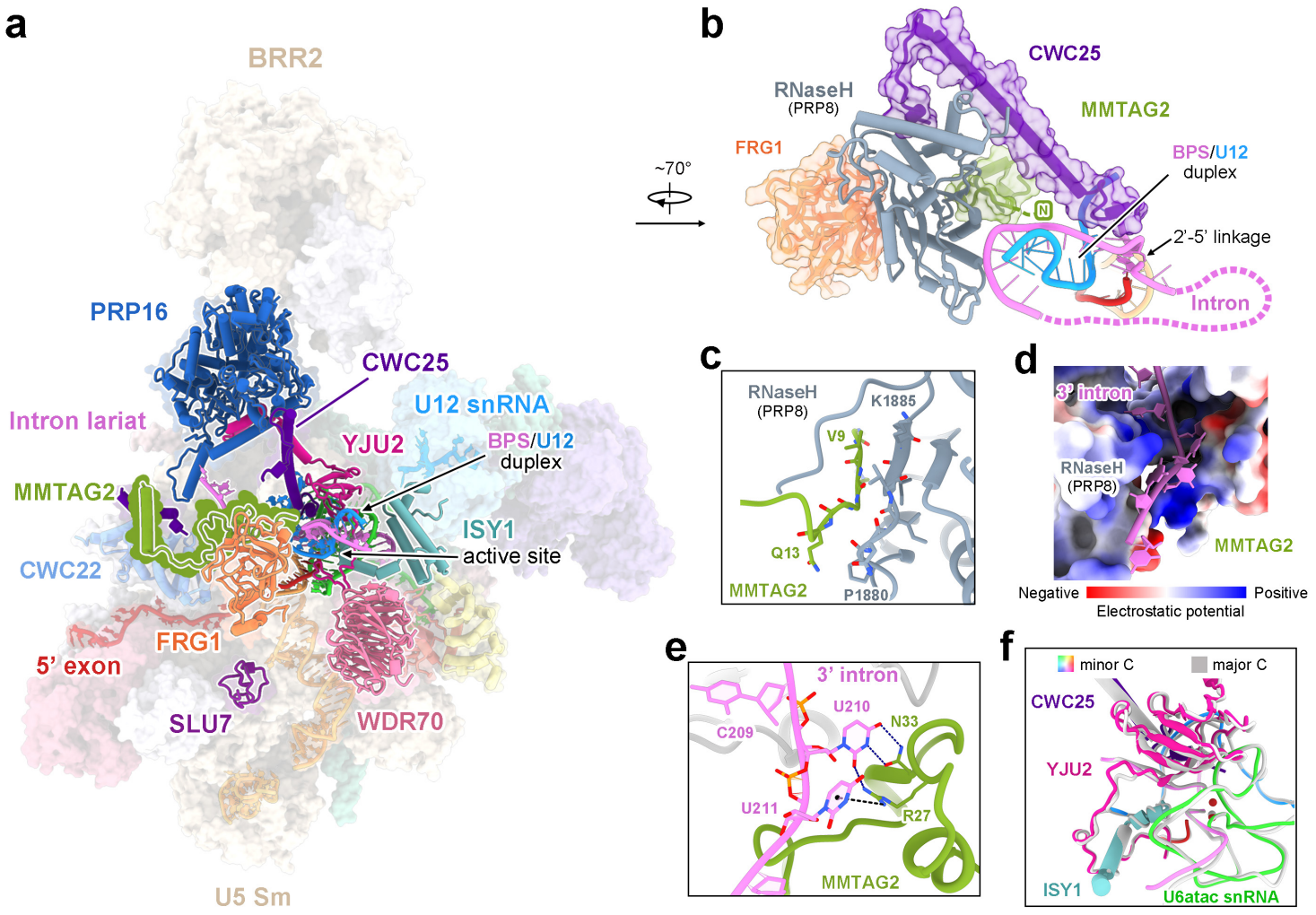
The entire models of the human minor C and C* complexes were respectively refined against the 2.92-Å map and 3.04-Å map using PHENIX[70] in real space with secondary structure and geometry restraints. Model overfitting was monitored by model-to-map FSC curves (Supplementary Figs. S3c and S4c). Model quality was assessed using the Molprobity scores and the Ramachandran plots (Supplementary Tables S2 & S3). Molprobity scores were calculated as described[71]. Components included in the final coordinates are summarized in Supplementary Tables S4 & S5.

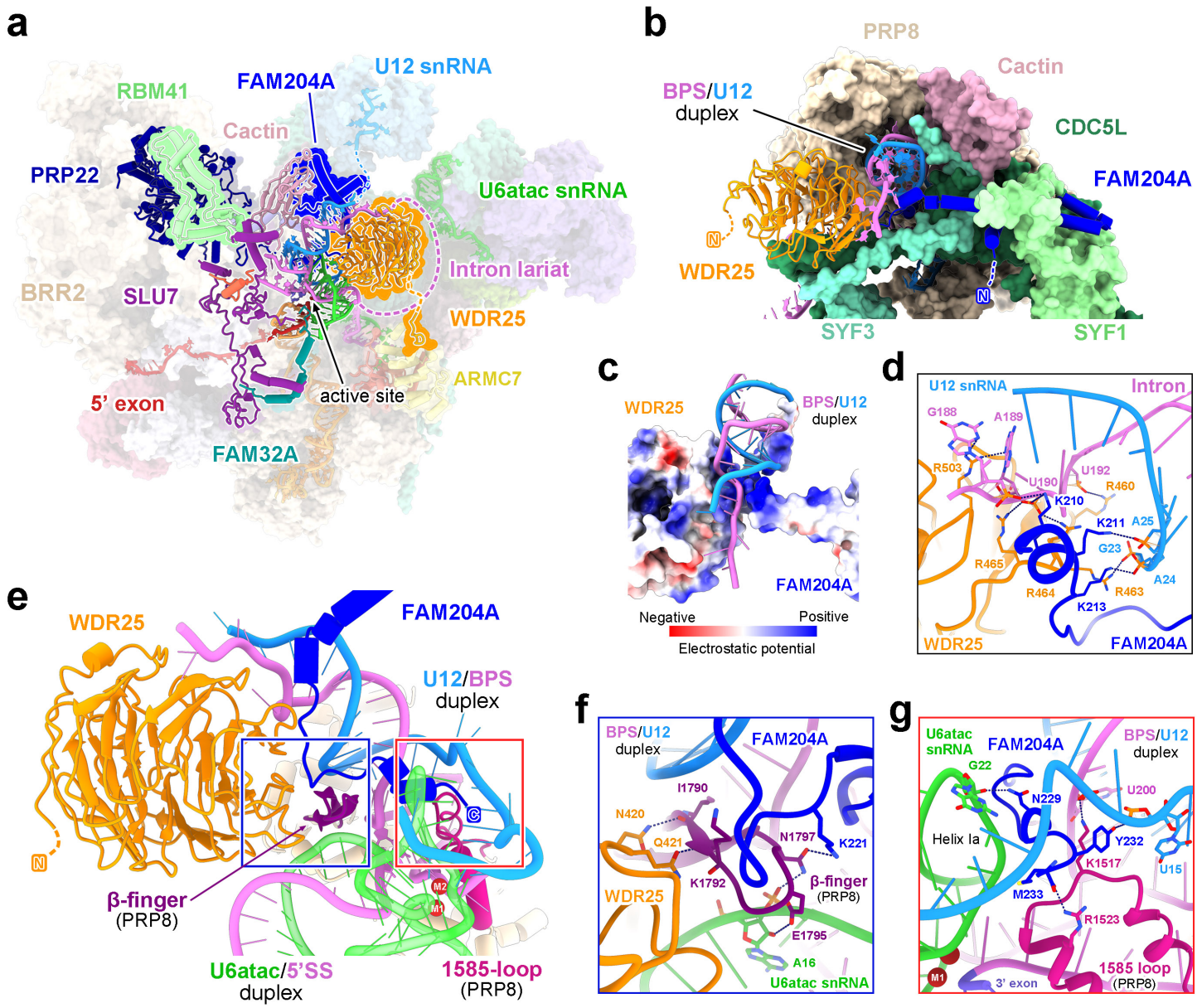
Competing interests: The authors declare no competing financial interests.

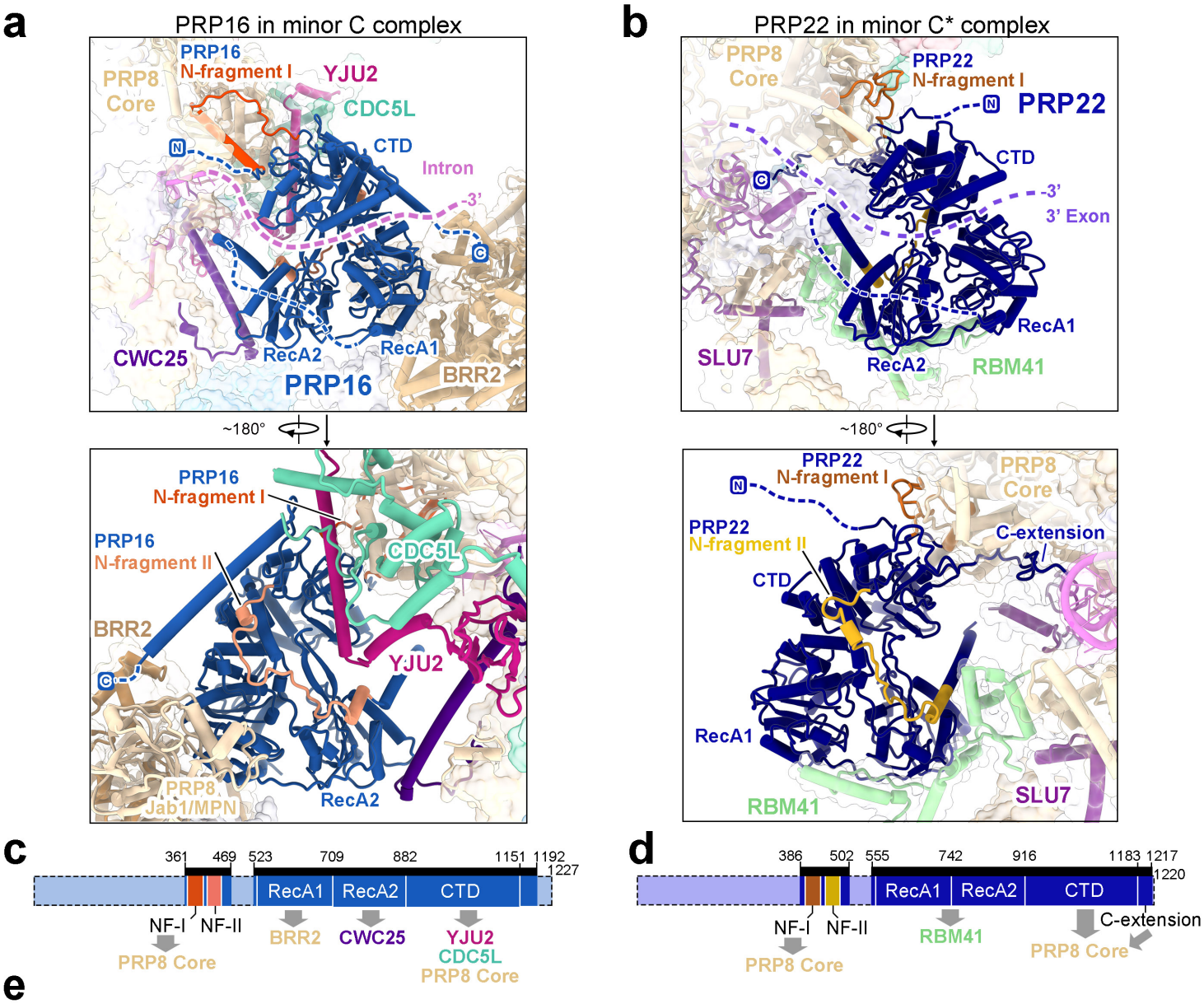
Data and materials availability: The atomic coordinates have been deposited in the Protein Data Bank. The EM maps have been deposited in the EMDB. All data and materials are available from the corresponding author (wanruixue@westlake.edu.cn) upon reasonable request.

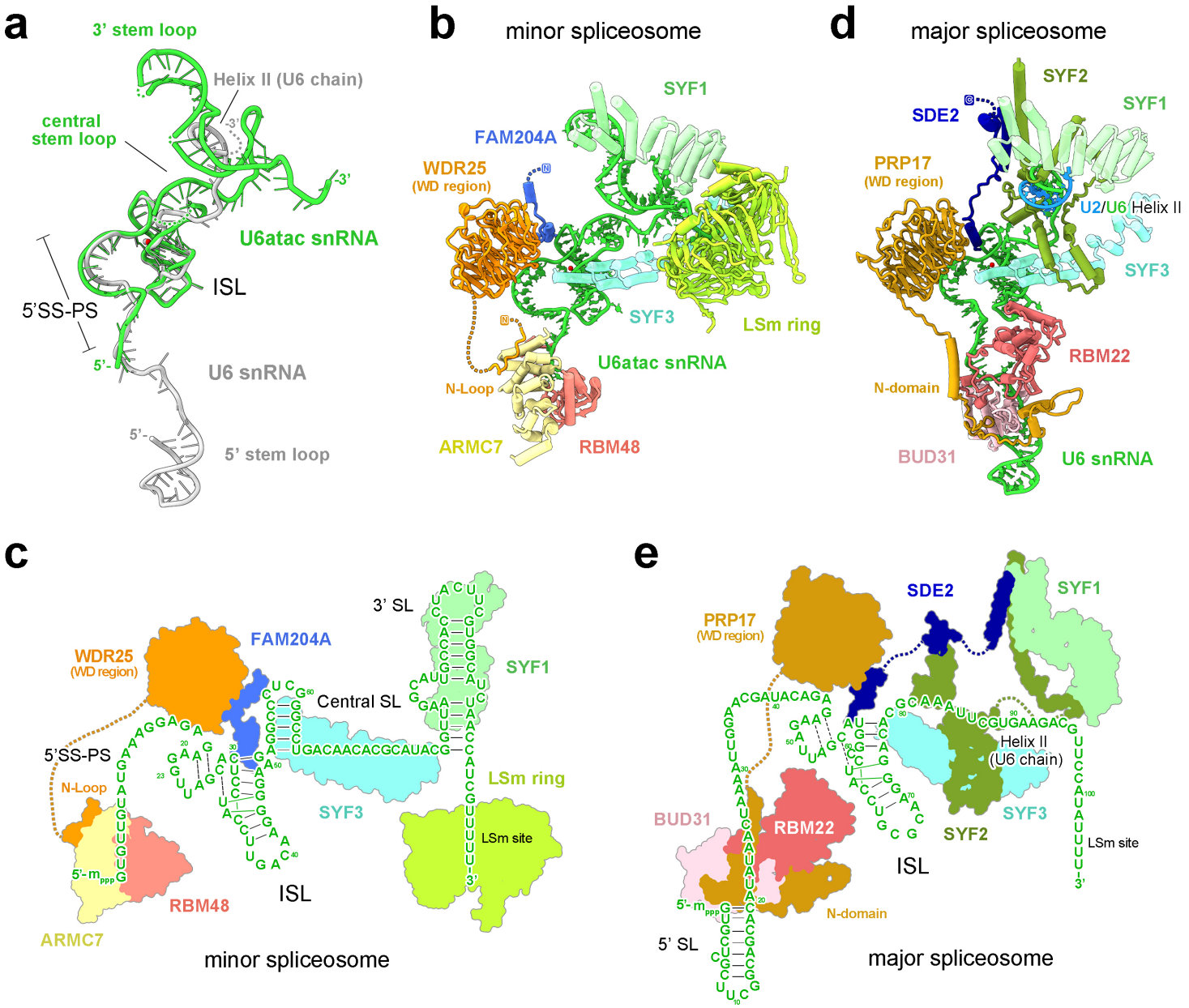

a
MINX-U12
5'exon
5'SS
BPS
3'SS
3'exon
A
AUAUCCUUUU
AUAUCCUUAACU
UCCACAC
Iterative optimization:
●Screening of scaffolds and regulatory elements
●Refinement of snRNAs' recognition sequences
MSE-U12
5'exon
3'exon
SCN4A
ENH1
A
AUAUCCUUUU
UUUUCCUUAACU
UCGUCAC
b
MINX-U12
MSE-U12
Time (min) 0 30 60 120
0 30 60 120
1000 bp
750 bp
500 bp
250 bp
100 bp
pre-mRNA
ligated exon
c
minor C complex
PPWD1
BRR2
U12 Sm
PRP16
FRG1
MMTAG2
NTR
EJC
WDR70
ARMC7
RBM48
U5 snRNP
Composite map
WDR25, FAM204A,
ESS2, PPIL3, NKAP,
Cactin,FAM32A,
RBM41, PRP22
PRP16-ATP
CWC25, YJU2,
FGR1, hMMTAG2,
WDR70, PRP16
d
minor C* complex
BRR2
U12 snRNA
Aquarius
WDR25
PPIL3
SLU7
LSm
FAM32A
EJC
NTC
CWC22
Composite map
U5 snRNP
U5 snRNA
U5 Sm ring
PRP8
BRR2
SNU114
U5-40K
U12 snRNP
U12 snRNA
U12 Sm ring
U2-A'
U2-B''
U6atac snRNP
U6atac snRNA
LSm ring
Intron lariat
Exon
NTC
PRPF19
SPF27
CDC5L
SYF1
SYF3
ISY1
NTR
CWC15
SKIP
PLRG1
PPIL1
BUD31
RBM22
IBC
Aquarius
PPIE
EJC
eIF4AIII
MAGOH
Y14
Splicing Factors
Shared
RBM48
ARMC7
SLU7
CWC22
CCDC12
PPWD1
SRm300
C specific
PRP16
FRG1
WDR70
CWC25
YJU2
MMTAG2
CIR1
C* specific
CACTIN
PPIL3
NKAP
ESS2
FAM32A
PRP22
RBM41
WDR25
FAM204A
PRKRIP1
FAM50A
CXORF56
PRP17
NOSIP
TLS1
SDE2


a
BPS/U12 duplex
U12 snRNA
Intron lariat
ISL
central stem loop
5'SS-PS
5' exon
U5 snRNA
γ-monomethyl phosphate cap
U6atac snRNA
minor C complex
d
BPS/U12 duplex
U12 snRNA
3' exon
ISL
3' stem loop
5' exon
Intron lariat
U5 snRNA
γ-monomethyl phosphate cap
U6atac snRNA
minor C* complex
e
minor C
minor C*
U12 (C*)
~ 80°
U12 (C)
~ 63Å
Intron lariat (C*)
3' intron (C*)
Intron lariat (C)
3' intron (C)
5' exon (C)
5' exon (C*)
b
U12 snRNA
Intron lariat
U6atac snRNA
A12AGGA16 box
BPS/U12 duplex
BP-A
M2
5'SS/U6atac duplex
2'-5' linkage
A1
U5 snRNA Loop I
M1
ISL
5' exon
minor C complex
(after Step-I reaction)
f
U12 snRNA
2'-5' linkage
U6atac snRNA
BP-A
Intron lariat
A1
A12AGGA16 box
3' exon
M2
5'SS/U6atac duplex
U5 snRNA Loop I
M1
ISL
5' exon
minor C* complex
(before Step-II reaction)
c
A26
BP-A
G27
2.5 Å
M2
U46
nucleophile in branching reaction
M1
U6atac snRNA
A1
(5'SS)
10 Å
G-1
(5' exon)
G44
leaving group
g
C-1
(3'SS)
A26
G27
M2
leaving group
U46
G1
(3' exon)
3.3 Å
M1
U6atac snRNA
nucleophile in exon ligation
G-1
(5' exon)
G44


a
BRR2
PRP16
CWC25
U12 snRNA
Intron lariat
YJU2
BPS/U12
duplex
MMTAG2
ISY1
active site
CWC22
FRG1
5' exon
SLU7
WDR70
U5 Sm
b
CWC25
RNaseH
(PRP8)
MMTAG2
FRG1
~70°
BPS/U12
duplex
N
2'-5' linkage
Intron
c
RNaseH
(PRP8)
K1885
V9
Q13
MMTAG2
P1880
d
3' intron
RNaseH
(PRP8)
MMTAG2
Negative
Positive
Electrostatic potential
e
3' intron
U210
C209
N33
U211
R27
MMTAG2
f
minor C
major C
CWC25
YJU2
ISY1
U6atac snRNA


a
RBM41
FAM204A
U12 snRNA
Cactin
PRP22
U6atac snRNA
Intron lariat
BRR2
SLU7
WDR25
ARMC7
5' exon
active site
FAM32A
b
PRP8
Cactin
BPS/U12
duplex
CDC5L
FAM204A
WDR25
SYF3
SYF1
c
WDR25
BPS/U12
duplex
FAM204A
Negative
Positive
Electrostatic potential
d
U12 snRNA
Intron
G188
A189
R503
U192
U190
R460
K210
K211
A25
G23
R465
A24
R464
R463
K213
WDR25
FAM204A
e
WDR25
FAM204A
U12/BPS
duplex
β-finger
(PRP8)
U6atac/5'SS
duplex
1585-loop
(PRP8)
f
BPS/U12
duplex
FAM204A
I1790
N420
Q421
N1797
K221
K1792
β-finger
(PRP8)
WDR25
E1795
A16
U6atac snRNA
g
U6atac
snRNA
FAM204A
BPS/U12
duplex
G22
N229
U200
Helix Ia
Y232
K1517
U15
M233
R1523
1585 loop
(PRP8)
3' exon


a
PRP16 in minor C complex
PRP16
N-fragment I
YJU2
PRP8
Core
CDC5L
CTD
Intron
-3'
CWC25
RecA2
RecA1
PRP16
BRR2
~180°
PRP16
N-fragment I
PRP16
N-fragment II
CDC5L
BRR2
YJU2
PRP8
Jab1/MPN
RecA2
b
PRP22 in minor C* complex
PRP8
Core
PRP22
N-fragment I
PRP22
CTD
-3'
3' Exon
RecA1
SLU7
RecA2
RBM41
~180°
PRP22
N-fragment I
PRP8
Core
PRP22
N-fragment II
C-extension
CTD
RecA1
RBM41
SLU7
c
361
469
523
709
882
1151
1192
1227
RecA1
RecA2
CTD
NF-I
NF-II
BRR2
CWC25
YJU2
CDC5L
PRP8 Core
PRP8 Core
d
386
502
555
742
916
1183
1217
1220
RecA1
RecA2
CTD
NF-I
NF-II
RBM41
C-extension
PRP8 Core
PRP8 Core
e
PRP16-N-fragment II (412-447) LVPPFLDGRIVFTKQPEPVIPVKDATSDLAIIARKG
PRP22-N-fragment II (444-479) EEPPFLRGHTKQSMDMSPIKIVKNPDGSLSQAAMMQ


a
3' stem loop
Helix II (U6 chain)
central
stem loop
-3'
-3'
U6atac snRNA
ISL
5'SS-PS
5'-
U6 snRNA
5'-
5' stem loop
b
minor spliceosome
SYF1
FAM204A
WDR25
(WD region)
SYF3
LSm ring
U6atac snRNA
N-Loop
ARMC7
RBM48
d
major spliceosome
SYF2
SYF1
SDE2
PRP17
(WD region)
U2/U6 Helix II
SYF3
RBM22
N-domain
BUD31
U6 snRNA
c
3' SL
FAM204A
WDR25
(WD region)
SYF1
Central SL
5'SS-PS
LSm ring
N-Loop
SYF3
LSm site
ISL
RBM48
ARMC7
minor spliceosome
e
SDE2
SYF1
PRP17
(WD region)
Helix II
(U6 chain)
BUD31
RBM22
SYF3
SYF2
ISL
LSm site
N-domain
5' SL
major spliceosome