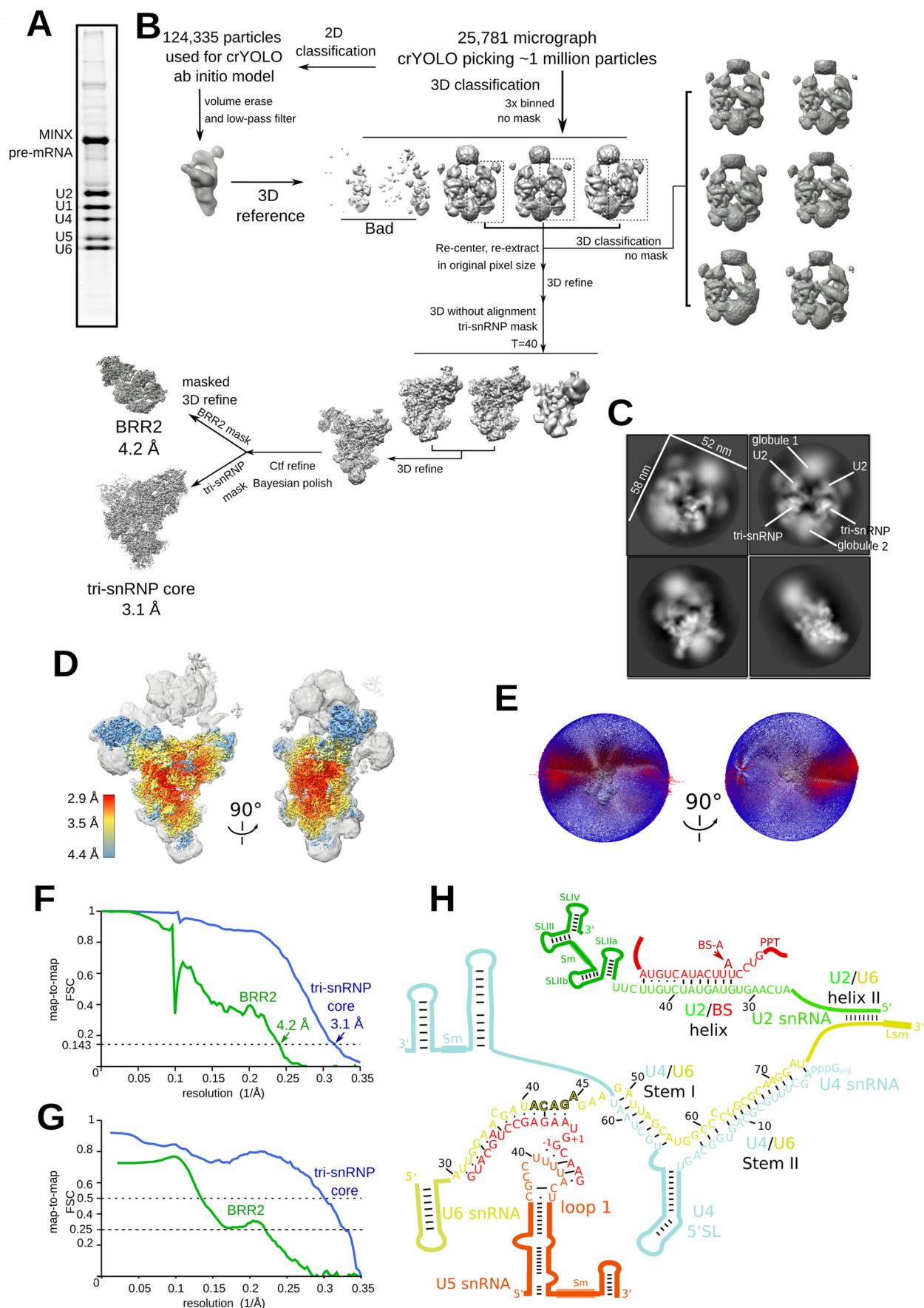


Expanded View Figures

Figure EV1. Cryo-EM and image-processing of the hB complex.

(A) RNA composition of purified B complex dimers. Human (h) B complexes were affinity-purified and RNA from the peak gradient fractions was isolated, separated on a NuPAGE gel, and visualized by staining with SyBr gold. U1 snRNA, that is no longer base paired to the 5' splice site of the MINX pre-mRNA is still present, together with U2, U4, U5 and U6 snRNA, and the MINX pre-mRNA. It is likely that, under our low salt purification conditions, U1 snRNP remains bound via protein-protein interactions (i.e., in the poorly resolved, globular EM densities). (B) Cryo-EM computation sorting scheme for the hB complexes. All major image-processing steps are depicted. (C) Representative cryo-EM 2D class averages of the human B complex dimers. (D) Local resolution estimation of the tri-snRNP core region of the B complex. (E) Orientation distribution plot for the particles contributing to the reconstruction of the tri-snRNP core region. (F) Fourier shell correlation (FSC) values for the listed parts of the B complex, indicate a resolution of 3.1 Å for the tri-snRNP core and 4.2 Å for the BRR2 region. (G) Map versus model FSC curves generated for the tri-snRNP core and BRR2 regions of hB using PHENIX mtriage. (H) Schematic of the RNA-RNA network in the hB complex. Red nucleotides are from the MINX pre-mRNA.



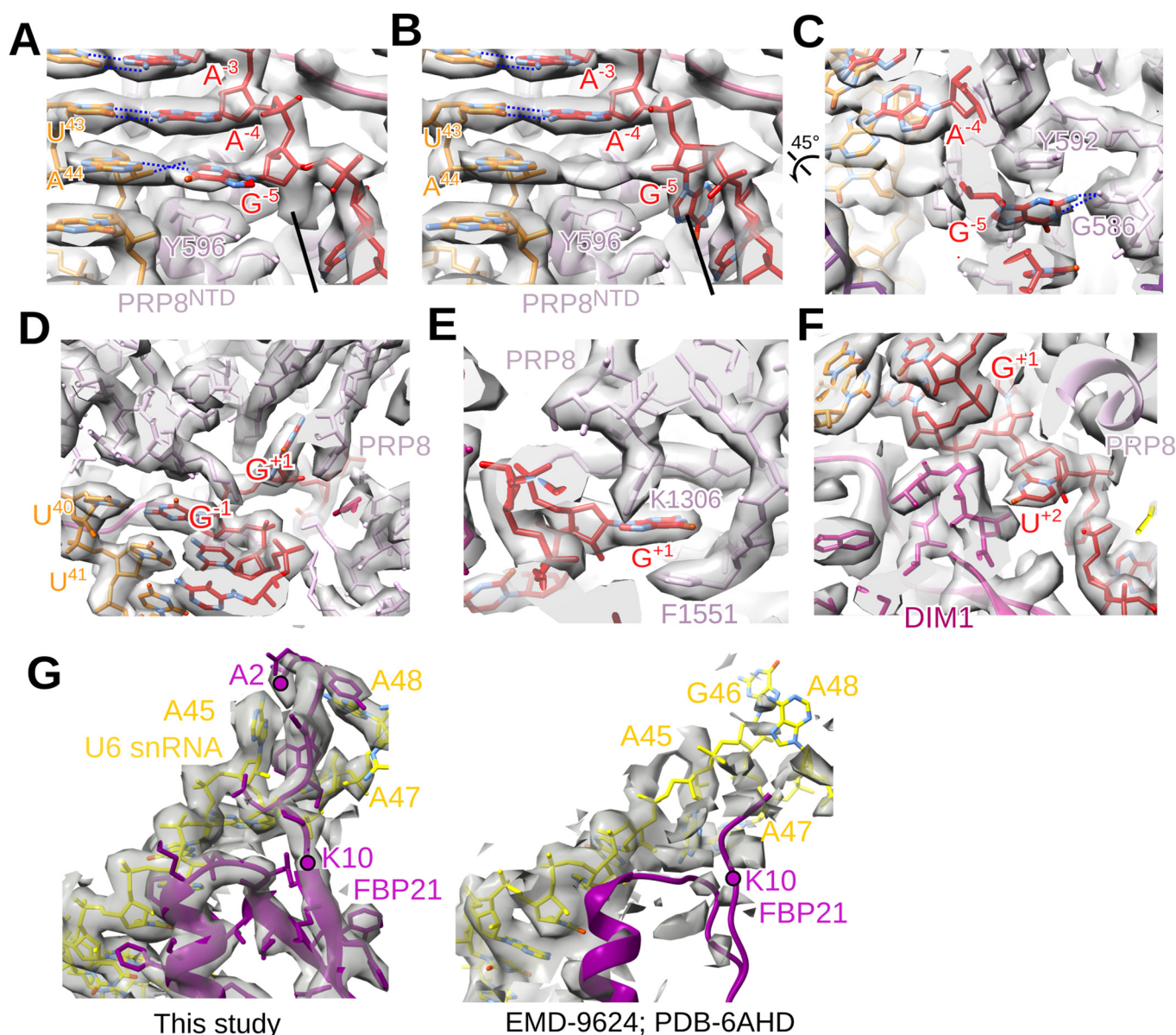


Figure EV2. Molecular interactions with the 5'ss nucleotides in hB.

(A–C) Two different conformations (panels A and B) of the base of G⁻⁵ of the 5' exon and their stabilization by neighboring residues of the PRP8 N-terminal domain (PRP8^{NTD}). The fit of U5 snRNA (orange) and 5' exon nucleotides (red), where numbers indicate the nucleotide position relative to the 5'ss GU dinucleotide at the 5' end of the intron, to the EM density is shown. Dashed lines, hydrogen bonds. Panel (C) shows a different view of the conformation shown in panel (B). In one conformation, G⁻⁵ base pairs with U5-A44 via two hydrogen bonds, one with the base of U5-A44, and the other with the preceding phosphate group. Moreover, the position of the G⁻⁵ base is further stabilized by stacking interactions with Y592 of PRP8^{NTD} (panel A). In the alternate conformation, the G⁻⁵ base is flipped back by about 0.6 nm (panel B) and is stabilized by stacking interactions with Y592 and by hydrogen bonds with the protein main chain at G586 of PRP8^{NTD} (panel C). The functional relevance of these alternative conformations, is not clear. However, after step 1 (i.e., in the human spliceosomal C complex), the base-paired conformation appears to be favored (Bertram et al, 2020), which would help to tether the cleaved 5' exon to the spliceosome prior to exon ligation. (D–F) Fit of 5'ss nucleotides and neighboring RNAs and proteins to the hB EM density. (G) Fit of the N-terminus of FBP21, together with the U6 nts with which it interacts, into the hB EM density (left panel). Comparison with a previous hB complex model (PDB-6AHD) (Zhan et al, 2018), in which U6 snRNA nts 46–47 were incorrectly placed in the density occupied by the N-terminus of FBP21 (right panel). In the B complex from the yeast *S. cerevisiae*, the 5'ss and U6 ACAGA box region, as well as the 5' exon and 5' stem-loop of the U6 snRNA, appear to be organized somewhat differently compared to the hB complex (Plaschka et al, 2017; Bai et al, 2018).

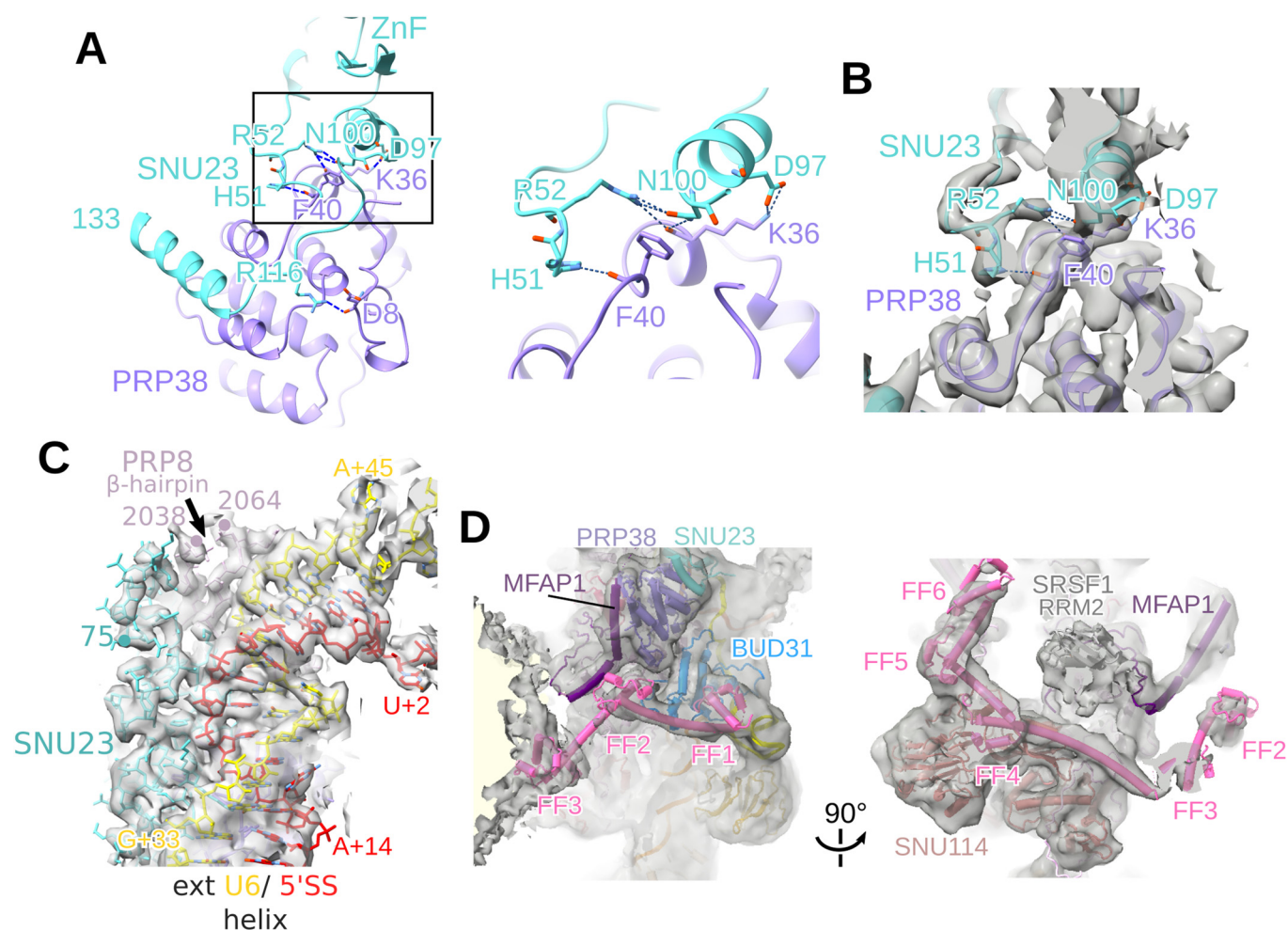


Figure EV3. Interface of the SNU23 zinc finger and PRP38 N-terminal region, and localization of TCERG1 and SRSF1.

(A) Interface between the SNU23 zinc finger (ZnF) and PRP38 N-terminal region. The boxed region is expanded in the right panel. (B) Fit of the SNU23 and PRP38 residues shown at the right in panel (A) into the hB EM density. (C) Fit of SNU23, the PRP8 RNase H-Jab1 linker and the U6/5'ss helix to the hB EM density. (D) Fit of the TCERG1 FF1-FF6 domains and SRSF1 RRM2 into the EM density of the hB monomer. Although the α -helix bridging FF2 and FF3 cannot be visualized in hB, fitting of the FF1-FF3 structure from pre-B^{act-1} into our B complex indicates that FF3 is located in a small globular density that directly contacts the lower globule 2.

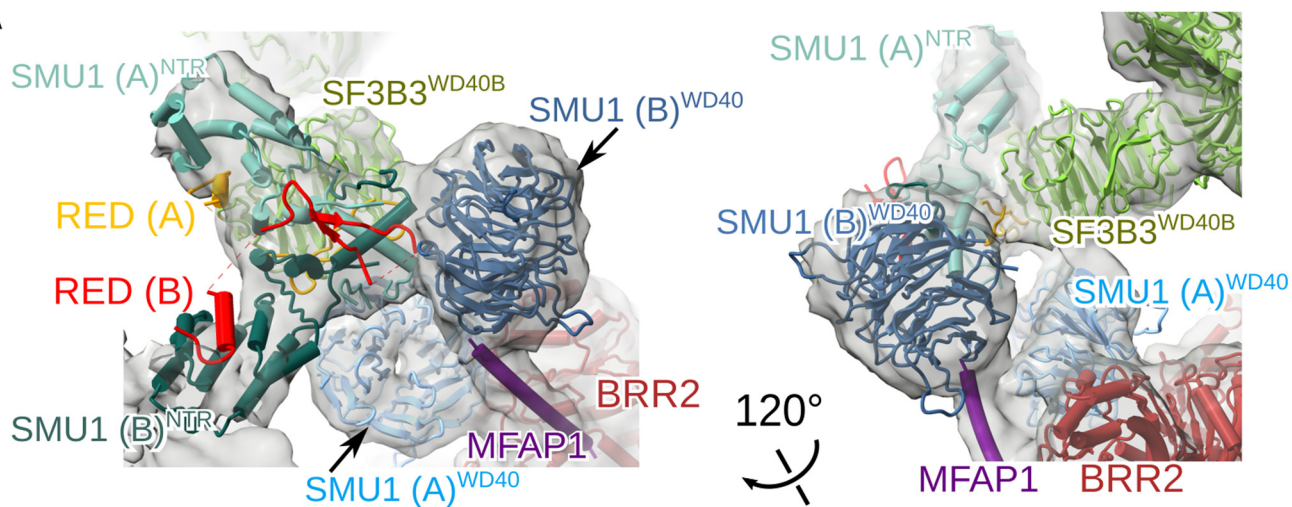
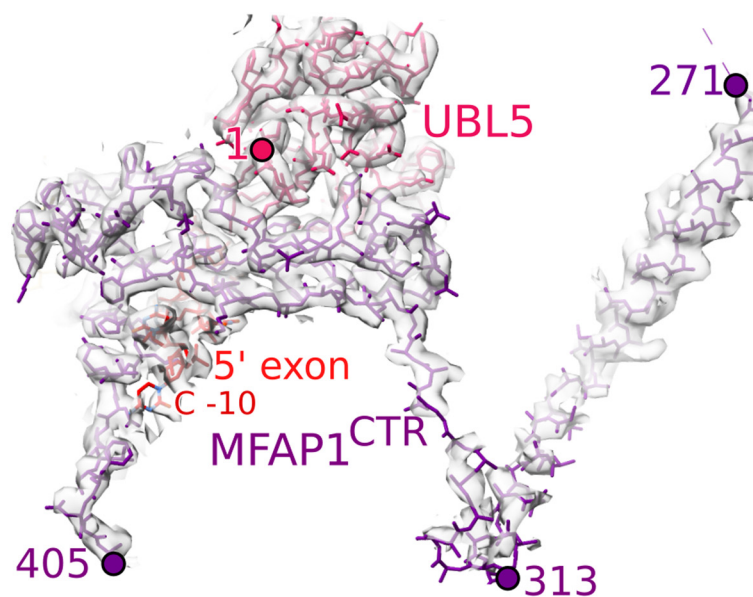
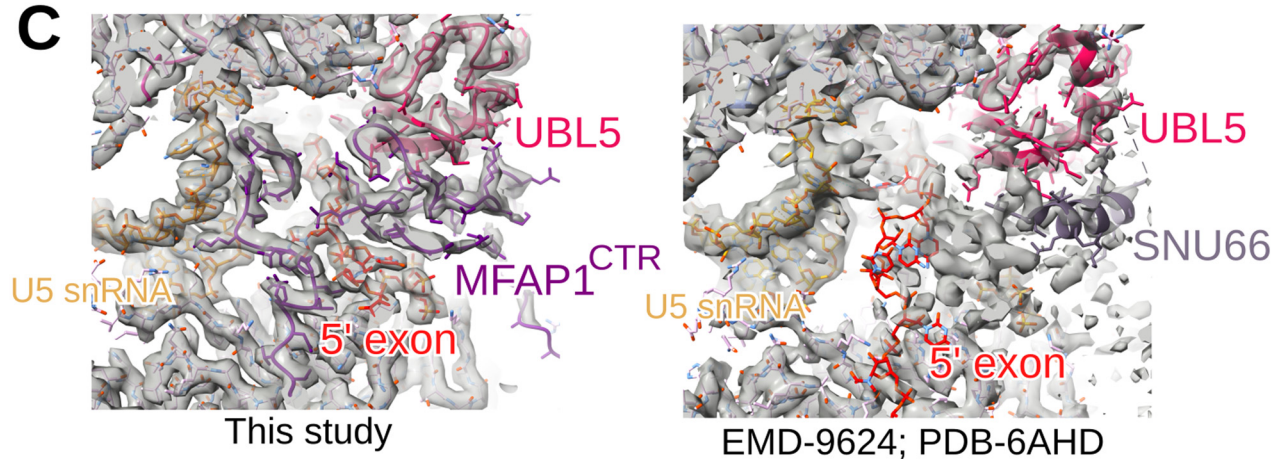
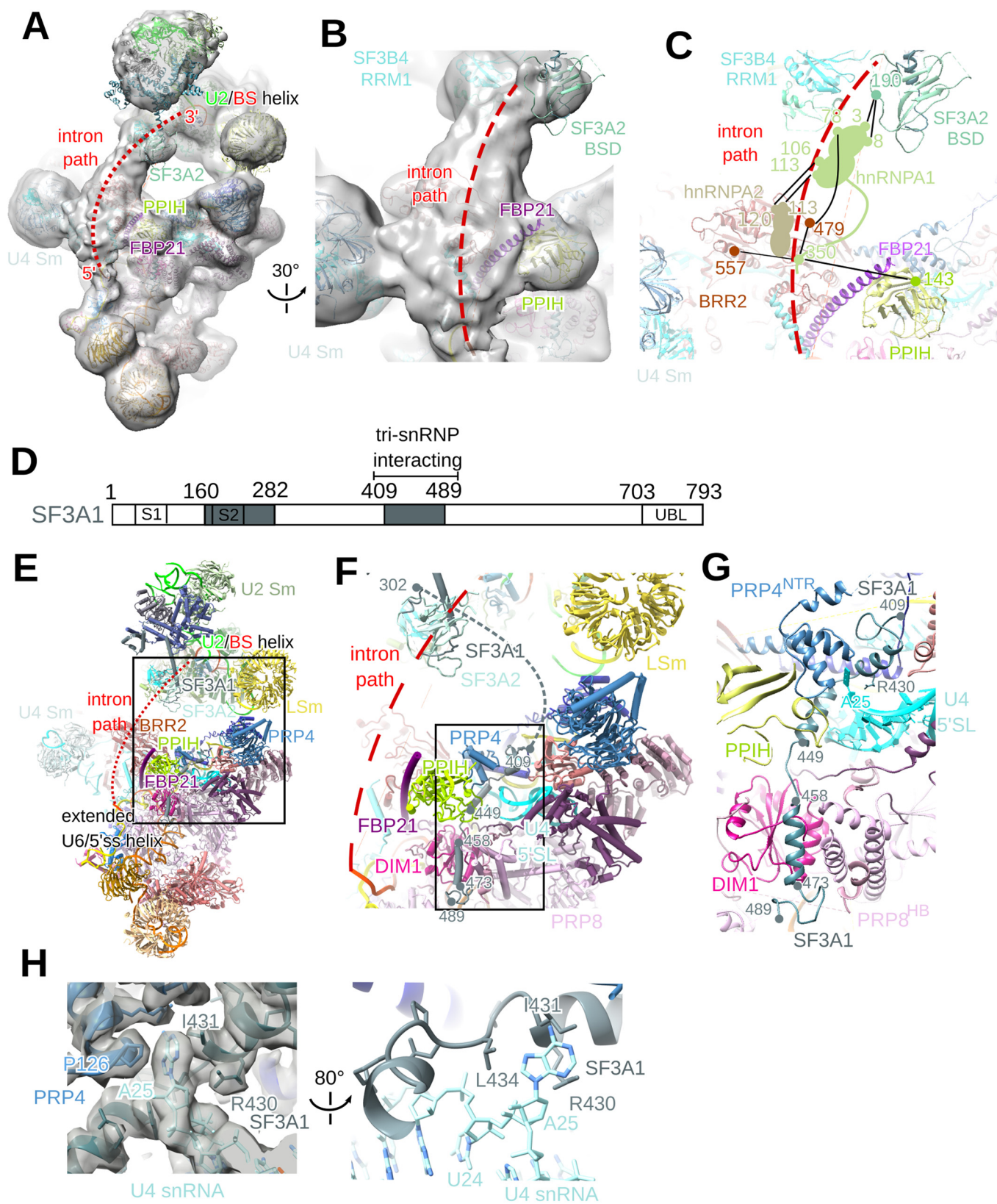
A**B****C**

Figure EV4. A heterotetrameric SMU1-RED complex is present in the human B complex.

(A) Fit of the SMU1 N-terminal region (NTR) and the WD40 domain of each SMU1 subunit (denoted **A** or **B**) of the SMU1 dimer, plus 2 copies of RED, into the hB EM density. The orientation of the individual NTR domains (**A** or **B**) cannot be determined unambiguously. That the B-specific proteins SMU1 and RED interact with each other was initially shown by co-precipitation experiments (Chung et al, 2009) and luciferase complementation assays (Fournier et al, 2014). The crystal structure of a minimal SMU1-RED complex from *C. elegans* and human, comprised of the SMU1 N-terminal region and a middle region of RED, showed that SMU1 dimerizes via its LiSH motif and a C-terminal α -helix, and that the dimerization module binds two copies of RED (Ashraf et al, 2019; Ulrich et al, 2016a). The human structure suggests that the SMU1 N-terminal region must first form a dimer before it can interact with two copies of the middle domain of RED (Ashraf et al, 2019). (B) Fit of MFAP1 helix 217-313 and its C-terminal region (CTR), as well as UBL5 and the 3' end of the 5' exon, to the hB EM density. (C) Fit of amino acids in the MFAP1 C-terminal region to the hB EM density. Comparison with a previous hB complex model (PDB-6AHD) (Zhan et al, 2018), in which an α -helix of SNU66 was incorrectly placed in the density occupied by the MFAP1 C-terminal region.



◀ **Figure EV5. Location of intron nucleotides upstream of the branch site and a C-terminal region of SF3A1 in hB.**

(A) Proposed path (dashed line) of the intron between the U2/BS helix and U6/5'ss helix. (B, C) Intron nucleotides directly upstream of the BS are sandwiched between SF3B4 RRM1 and the β -sandwich domain (BSD) of SF3A2, and nucleotides further upstream are likely contacted by hnRNP proteins. Panel (B) shows the EM density that appears to accommodate the intron, as well as adjacent protein domains. Panel (C) shows the proposed position of hnRNPA1 and A2, based on protein crosslinks with neighboring B complex proteins. Intermolecular crosslinks are indicated by a line, with the position of the crosslinked residues indicated. (D) Domain organization of SF3A1. The modeled regions are indicated as colored boxes. S1 and S2, SURP domains 1 and 2; UBL, ubiquitin-like domain. (E) Overview of the hB complex molecular architecture. The boxed region is expanded in panel (F). (F, G). SF3A1 aa 409–449 interact with U4 SL1, PRP4 and PPIH, while SF3A1 aa 454–489, which form an α -helix, are located between DIM1 and the PRP8 helical bundle (PRP8^{HB}). The region boxed in panel (F) is expanded in panel (G). (H) Fit of indicated U4 nucleotides, and amino acids of PRP4 and SF3A1 into the hB EM density. The base of A25 interacts with several hydrophobic protein side chains (i.e., I431, L434 of SF3A1 and P126 of PRP4), while R430 of SF3A1 contacts the backbone of U4-A25.