



## Supplementary Materials for **Structure of a transcribing RNA polymerase II–U1 snRNP complex**

Suyang Zhang, Shintaro Aibara, Seychelle M. Vos, Dmitry E. Agafonov, Reinhard Lührmann,  
Patrick Cramer\*

\*Corresponding author. Email: [patrick.cramer@mpibpc.mpg.de](mailto:patrick.cramer@mpibpc.mpg.de)

Published 15 January 2021, *Science* **371**, 305 (2021)  
DOI: 10.1126/science.abf1870

**This PDF file includes:**

Materials and Methods  
Figs. S1 to S9  
Table S1  
Captions for Movies S1 and S2  
References

**Other Supplementary Materials for this manuscript include the following:**  
(available at [science.sciencemag.org/content/371/6526/305/suppl/DC1](http://science.sciencemag.org/content/371/6526/305/suppl/DC1))

MDAR Reproducibility Checklist (.pdf)  
Movies S1 and S2 (.mp4)

## Materials and Methods

### Expression and protein purifications

The porcine Pol II was purified from *Sus scrofa domesticus* thymus as described (15). In brief, the thymus was flash frozen in liquid nitrogen and homogenized in a blender. After removing the fat tissue, the supernatant was precipitated with 0.04% polyethylenimine (PEI). The PEI pellet was dissolved in 0.4 M ammonium sulfate buffer and applied onto a column packed with Macro-Prep High Q resin (Biorad). The eluted peak was then precipitated with 4 M ammonium sulfate and the re-dissolved pellet applied onto an 8WG16 antibody column. The eluate was further purified on a Uno Q1 column (Biorad) with a final size exclusion purification step using the Sephacryl S-300 16/60 column (GE Healthcare) in 10 mM HEPES pH 7.25, 150 mM NaCl, 10  $\mu$ M ZnCl<sub>2</sub> and 10 mM DTT. Human U1 snRNP was purified from HeLa cell nuclear extract by immunoaffinity purification and a 10-30% glycerol gradient as described (16). The fractions containing U1 snRNP were diluted to 100 mM NaCl and applied onto a heparin sepharose column (GE Healthcare). U1 snRNP was eluted with 20 mM Tris-HCl pH 7.9, 450 mM NaCl, 1.5 mM MgCl<sub>2</sub> and 1 mM DTT. Human P-TEFb, PAF, DSIF and SPT6 were purified as described (15).

### RNA preparation

The following RNA constructs were generated for this study:

15-nt RNA: 5'-*GAG AGG GAA CCC ACU*-3'

145-nt RNA: 5'-3xMS<sub>2</sub>-CUU GGA UCG GAA ACC CGU CGG CCU CCG **ACA GGU AAG UAU** UAA CCG *GAG AGG GAA CCC ACU*-3'

155-nt RNA: 5'-3xMS<sub>2</sub>-CUU GGA UCG GAA ACC CGU CGG CCU CCG **ACA GGU AAG UAU** CUA GCA UGU AUA ACC *GGA GAG GGA ACC CAC U*-3'

189-nt RNA: 5'-3xMS<sub>2</sub>-CUU GGA UCG GAA ACC CGU CGG CCU CCG **ACA GGU AAG UAU** CUA GCA UGU AGA ACU GGU UAC CUG CAG CCC AAG CUU GCU GCA CGU AAC CGG *AGA GGG AAC CCA CU*-3'

3xMS<sub>2</sub>: 5'-CGU ACA CCA UCA GGG UAC GAG CUA GCC CAU GGC GUA CAC CAU CAG GGU ACG ACU AGU AGA UCU CGU ACA CCA UCA GGG UAC G-3'

The sequence that is complementary to U1 snRNA is in bold and the sequence that is covered within Pol II is in italic. The 15-nt RNA was synthesized by IDT. All other constructs contained modified MINX pre-mRNA sequence (31) and were cloned into the pUC18 vector with

a T7 promoter and three MS<sub>2</sub>-tags at the 5' end and an HDV (hepatitis delta virus) ribozyme at the 3' end. The plasmids were prepared using the Maxiprep kit (Qiagen) and digested with XbaI overnight. The digested plasmids were purified by phenol chloroform extraction and ethanol precipitation. RNA was prepared by *in vitro* transcription using T7 RNA polymerase (5 U/μl final, Thermo Fisher Scientific) at 37 °C for 4 hours in the presence of 100 ng/μl digested plasmid, 40 mM Tris-HCl pH 7.9, 30 mM MgCl<sub>2</sub>, 10 mM DTT, 10 mM NaCl, 2 mM spermidine, 4 mM NTP and 0.001% Triton X-100. The transcription reaction was incubated at 60 °C for 20 min to facilitate HDV ribozyme cleavage, and then precipitated with ethanol, dissolved in 1x RNA loading dye (47.5% formamide, 0.01% bromophenol blue and 0.5 mM EDTA) and loaded onto 6% denaturing gels (6% acrylamide/bis-acrylamide, 7M urea and 1x TBE). The RNA bands were excised from the gels using UV shadowing and incubated in 0.3 M sodium acetate pH 5.2 overnight at -80 °C for extraction. The RNA was subsequently precipitated with isopropanol at -80 °C for 1 hour and dissolved in water.

RNAs were 5' capped using the Vaccinia capping system (NEB). For each reaction of 20 μl, the RNA (10 μg) and nuclease-free water were combined to a final volume of 15 μl, heated at 65 °C for 5 min, put on ice for 5 min and the following components were added in the specified order: 1x capping buffer (NEB), 0.5 mM GTP, 0.1 mM SAM and 1 μl Vaccinia capping enzyme (10 units/μl, NEB). The reaction was incubated at 37 °C for 1 h and the capped RNA was purified by phenol chloroform extraction and ethanol precipitation at -20 °C overnight. The precipitated RNA was dissolved in water and flash frozen in small aliquots in liquid nitrogen.

#### Sample preparation for cryo-EM

The transcribing Pol II-U1 snRNP complex was formed on a DNA scaffold with the following sequences: template DNA 5'-/Desthiobiotin/-GCT CCC AGC TCC CTG CTG GCT CCG AGT GGG TTC TGC CGC TCT CAA TGG-3', non-template DNA 5'-CCA TTG AGA GCG GCC CTT GTG TTC AGG AGC CAG CAG GGA GCT GGG AGC-3'. The DNA scaffold contains 14 nucleotides of upstream DNA, 23 nucleotides of downstream DNA and a mismatch bubble of 11 nucleotides, of which 9 nucleotides of the template DNA base-pairs with the RNA. The DNA oligos were synthesized by IDT and dissolved in water.

RNA (15-nt, 145-nt or 155-nt) and template DNA were mixed in equimolar ratio (30 μM) in 20 mM HEPES pH 7.4, 100 mM NaCl and 3 mM MgCl<sub>2</sub>, and annealed by heating up at 60 °C

for 4 min followed by decreasing the temperature by 1 °C min<sup>-1</sup> steps to a final temperature of 30 °C in a thermocycler. *Sus scrofa domesticus* Pol II (50 pmol) and the RNA–template hybrid (100 pmol) were incubated for 10 min at 30 °C, followed by addition of the non-template DNA (200 pmol) and incubation at 30 °C for another 10 min. Phosphorylation of Pol II was performed with 0.4 μM P-TEFb and 1 mM ATP in 20 mM HEPES pH 7.4, 100 mM NaCl, 3 mM MgCl<sub>2</sub> and 1 mM DTT in a total volume of 25 μl for 30 min at 30 °C. P-TEFb phosphorylation was performed as previous studies showed that phosphorylated Pol II CTD can serve as a platform to recruit splicing factors (32-36). The assembled transcribing Pol II complex was then applied onto a Superdex 200 Increase 3.2/300 column (GE Healthcare) equilibrated with the SEC150 buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.5 mM tris(2-carboxyethyl)phosphine (TCEP)). Purified U1 snRNP from the heparin column was further purified on the Superose 6 Increase 3.2/300 column (GE Healthcare) in SEC150 buffer. The peak fraction of the Pol II complex was mixed with the peak fraction of U1 snRNP at a molar ratio of 1:1.5, incubated on ice for 30 min and used for freezing grids. Similar results were obtained when combining the Pol II complex and U1 snRNP before the size exclusion purification step.

Samples were diluted to a concentration of 60-65 nM for the Pol II-U1 snRNP complex and 2 μl was applied to each side of the R2/2 UltrAuFoil grids (Quantifoil) that had been glow-discharged for 100 s. After incubation of 10 s and blotting for 4 s, the grid was vitrified by plunging it into liquid ethane with a Vitrobot Mark IV (FEI Company) operated at 4 °C and 100% humidity.

#### Analytical size exclusion assay

Analytical size exclusion assay was performed using the same DNA scaffold as for the cryo-EM studies. EC\* had to be used instead of Pol II to allow separation of the U1 snRNP bound complex from free U1 snRNP. The same procedure was carried out for three different RNA constructs (15-nt, 155-nt and 189-nt RNAs). RNA and template DNA were mixed in equimolar ratio (25 μM) in 20 mM HEPES pH 7.4, 100 mM NaCl and 3 mM MgCl<sub>2</sub>, and annealed by heating up at 60 °C for 4 min followed by decreasing the temperature by 1 °C min<sup>-1</sup> steps to a final temperature of 30 °C in a thermocycler. *Sus scrofa domesticus* Pol II (25 pmol) and the RNA–template hybrid (30 pmol) were incubated for 10 min at 30 °C, followed by addition of the non-template DNA (50 pmol) and incubation at 30 °C for another 10 min. After addition of the elongation factors DSIF, PAF and SPT6 (75 pmol each), phosphorylation of EC\* was performed



with 0.4  $\mu$ M P-TEFb and 1 mM ATP in 20 mM HEPES pH 7.4, 100 mM NaCl, 3 mM MgCl<sub>2</sub> and 1 mM DTT for 30 min at 30 °C. The assembled EC\* was kept on ice for 5 min and mixed with 3x molar excess of U1 snRNP (75 pmol) in a total volume of 37  $\mu$ l. Following incubation on ice for 15 min, the sample was applied onto a Superose 6 Increase 3.2/300 column (GE Healthcare) equilibrated with the SEC100 buffer (20 mM HEPES pH 7.4, 100 mM NaCl, 3 mM MgCl<sub>2</sub>, 0.5 mM TCEP). Control runs of U1 snRNP alone and U1 snRNP with phosphorylated elongation factors were performed.

The fractions 1-12 were loaded individually onto 4-12% NuPAGE Bis-Tris gels (the amount loaded was calculated to have the same amount of Pol II in fraction #2) and run in 1x MOPS at 180 V for 50 min. Gels were stained with Instant Blue. To visualize U1 snRNA, fraction #2 of all samples as well as a positive control of U1 snRNP were incubated with 1  $\mu$ l of proteinase K (NEB, 800 units/ml) for 30 min at 37 °C and loaded onto a denaturing gel (8% acrylamide/bis-acrylamide, 7M urea and 1x TBE). The denaturing gel was stained with SybrGold (Thermo Fisher Scientific) and visualized using a Typhoon 9500 FLA Imager (GE Healthcare). The U1-70k bands were confirmed by Western blotting using an antibody against U1-70k (Santa Cruz Biotechnology, sc-390899) and a secondary anti-mouse HRP antibody (Abcam, ab5870). The loading control was confirmed with an antibody against RPB3 (BETHYL, A303-771A) and a secondary anti-rabbit HRP antibody (GE Healthcare, NA934).

#### RNA extension assay

The RNA extension assay was performed with a complementary DNA scaffold that corresponds to modified *JUNB* sequence with the 5'SS sequence inserted (in bold). Template DNA: 5'-TCT GGC GCG ATA GCT TTC CTG GCG TCG AAA CAC CCA GCA CCC AGC ACC CAG CAG GCA CCG **ATA CTT ACC TGG** TCC GCT CTC AAT GG-3'. Non-template DNA: 5'-CCA TTG AGA GCG GAC **CAG GTA AGT ATC** GGT GCC TGC TGG GTG CTG GGT GCT GGG TGT TTC GAC GCC AGG AAA GCT ATC GCG CCA GA-3'. RNA: 5'-/6-FAM/-UUU UUU CCA GGU AAG-3'. The scaffold has 14 nucleotides of upstream DNA, 63 nucleotides of downstream DNA, a 9 base-paired RNA·DNA hybrid and 6 nucleotides of exiting RNA. The oligos were synthesized by IDT.

RNA and template DNA were mixed in equimolar ratio (10  $\mu$ M) in 20 mM HEPES pH 7.4, 100 mM NaCl and 3 mM MgCl<sub>2</sub>, and annealed by heating up at 60 °C for 4 min followed by

decreasing the temperature by 1 °C min<sup>-1</sup> steps to a final temperature of 30 °C in a thermocycler. All concentrations refer to the final concentration in the assays. *Sus scrofa domesticus* Pol II (150 nM) and the RNA–template hybrid (100 nM) were incubated for 10 min at 30 °C, followed by addition of the non-template DNA (200 nM) and incubation at 30 °C for another 10 min. The phosphorylation reaction was then performed with 100 nM P-TEFb and 1 mM ATP in 20 mM HEPES pH 7.4, 100 mM NaCl, 3 mM MgCl<sub>2</sub>, 4% glycerol and 1 mM DTT for 15 min at 30 °C for either Pol II alone or in the presence of elongation factors (DSIF, PAF, SPT6 and RTF1, 150 nM each). The reactions were then cooled down on ice and a titration of U1 snRNP from 0 to 1.5 μM was added. Transcription was initiated upon addition of 50 μM NTP (Pol II alone) or 10 μM NTP (EC\*) at 20 °C. The reactions were stopped at various timepoints by mixing 5 μl of transcription reaction with 5 μl of 2x stop buffer (6.4 M urea, 50 mM EDTA pH 8.0, 1x TBE). Samples were treated with 2 μl of proteinase K (NEB, 800 units/ml) for 20 min at 37 °C and 5 μl of each sample was applied onto 20% denaturing gels (20% acrylamide/bis-acrylamide, 7 M urea and 1xTBE, run in 0.5x TBE at 300 V for 100 min). The gels were visualized by the 6-FAM label on the RNA primer using the Typhoon 9500 FLA Imager (GE Healthcare).

#### Cryo-EM data collection and processing

Cryo-EM data were collected on the 300 kV FEI Titan Krios with a K3 summit direct detector (Gatan) and a GIF quantum energy filter (Gatan) operated with a slit width of 20 eV. Automated data collection was performed with SerialEM (37) with a tilt angle of 40° at a nominal magnification of 81,000x, corresponding to a pixel size of 1.05 Å/pixel. Image stacks of 40 movie frames were collected with a defocus range of -0.5 to -2.0 μm in electron counting mode and a dose rate of 1.01 e<sup>-</sup>/Å<sup>2</sup>/frame. A total of 17,321 image stacks were collected across four datasets for the Pol II-U1 snRNP complex with 145-nt RNA. A total of 5,660 image stacks were collected for the Pol II-U1 snRNP complex with 155-nt RNA.

All movie frames were aligned and the contrast transfer function (CTF) parameters were calculated in Warp (38). Particles in 400 pixels x 400 pixels were selected by automatic particle picking in Warp. The following steps were performed in RELION 3.0 (39) to exclude bad particles from the dataset: 1) Two-dimensional (2D) classification was performed and particles in bad classes with poorly recognizable features were excluded. 2) In the second round of 2D classification, free U1 snRNP particles were excluded, leaving only Pol II containing particles. 3)

The remaining particles were refined using three-dimensional (3D) refinement, followed by multiple rounds of CTF refinement to correct the CTF parameters. Beam-induced particle motion was corrected using Bayesian polishing in RELION 3.0 (40, 41). The particles were then refined with a soft mask on Pol II and divided into six classes using 3D classification in RELION with local fine-angle search (0.9 degree) (**fig. S2, A and B**). All 3D classes with bad particles or the class with missing clamp were discarded.

For the dataset of the Pol II-U1 snRNP complex with 145-nt RNA, all good particles with Pol II were combined and 3D refined using a soft mask on Pol II, resulting in a reconstruction of Pol II at 2.6 Å resolution. To separate Pol II alone particles from the Pol II-U1 snRNP particles, signal subtraction of Pol II followed by focused 3D classification of the subtracted particles without alignment was performed using a large spherical mask near the RNA exit site (where the extra density is) (**fig. S2, C and D**). The classes with an extra density corresponding to U1 snRNP were combined (17.5%), reverted to original particles and 3D refined with a solvent mask. This resulted in a Pol II-U1 snRNP reconstruction with 220,922 particles at an overall resolution of 3.0 Å (**fig. S2E**). A second round of signal subtraction and focused 3D classification without alignment using a soft mask on U1-70k/stem loop I (the interface density) was performed to separate the movements of U1 snRNP (**fig. S2G**). All four classes have good densities corresponding to U1-70k/stem loop I, yet a clear movement relative to Pol II was observed when they were superimposed. The class with the best density for U1-70k/stem loop I was selected, reverted to original particles and 3D refined with a solvent mask. This resulted in the final Pol II-U1 snRNP reconstruction with 61,596 particles at an overall resolution of 3.6 Å (**fig. S2H**).

To improve local resolutions for model building, the particles were also refined with a soft mask on either Pol II or the Pol II-U1 snRNP interface region (3.3 Å and 3.2 Å, respectively, **fig. S2I**). Following 3D refinement with a soft mask on Pol II, signal subtraction of Pol II was performed, leaving only the signal of U1 snRNP. The subtracted particles were then refined with a soft mask on U1 snRNP, resulting in a U1 snRNP reconstruction of 5.9 Å (**fig. S2I**). Due to the extended structure of U1 snRNP, the Sm-ring and RNA stem-loops (SL) 2-4 that are further away from the Pol II-U1 snRNP interface showed large movements relative to Pol II and their fit relied on the focused refined map of U1 snRNP (**fig. S3 and fig. S4**).

The dataset for the Pol II-U1 snRNP complex with 155-nt RNA was processed in the same way, resulting in a final reconstruction with 14,497 particles at 3.9 Å resolution (**fig. S8**). The same

protein interface between Pol II and U1 snRNP was observed, whereas the linking RNA showed higher mobility and more diffuse densities.

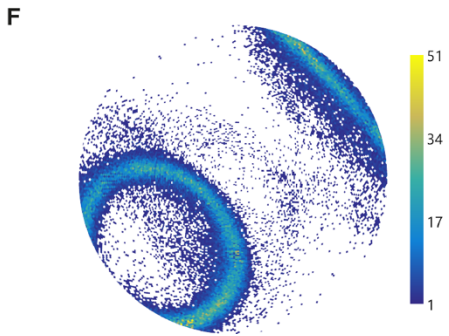
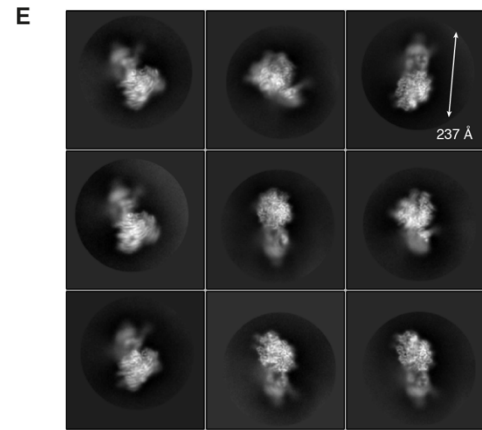
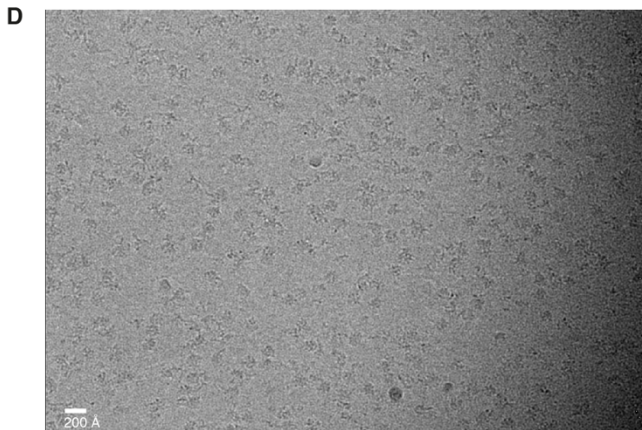
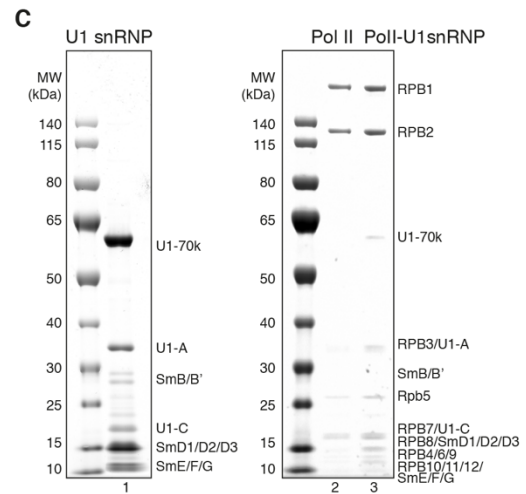
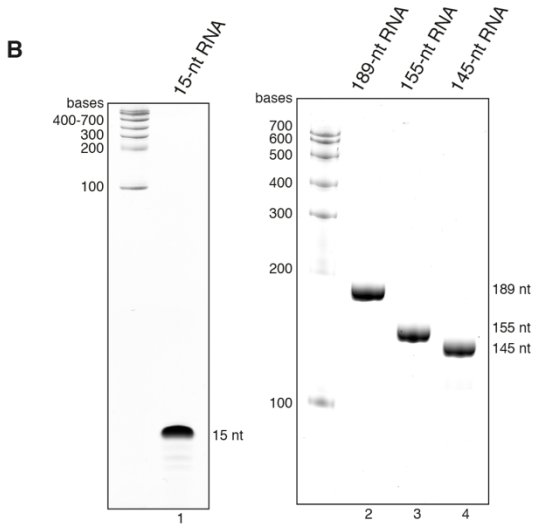
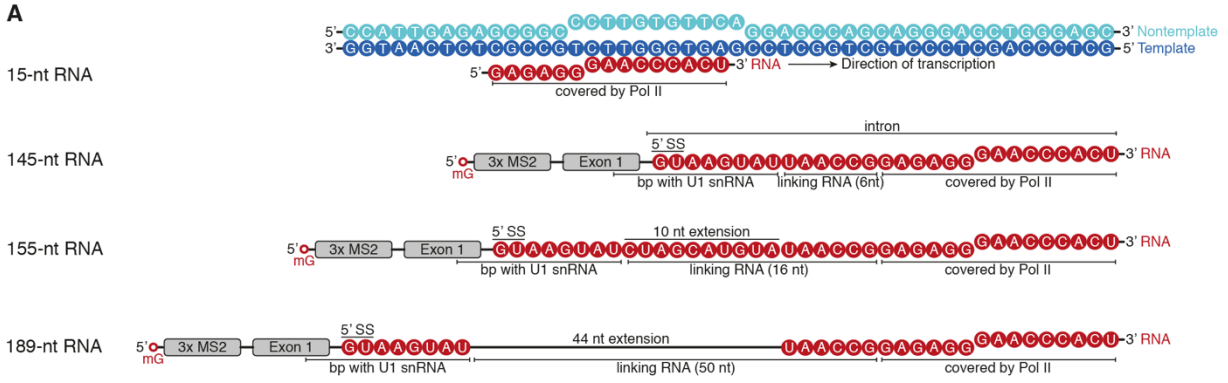
Final maps were sharpened using PHENIX AutoSharpen (42). Local resolution of the maps was estimated using RELION 3.0 (39). All resolution estimations were based on the gold-standard Fourier Shell Correlation (FSC) calculations using the FSC = 0.143 criterion. A summary of all EM reconstructions obtained in this paper is listed in **table S1**.

### Model building and refinement

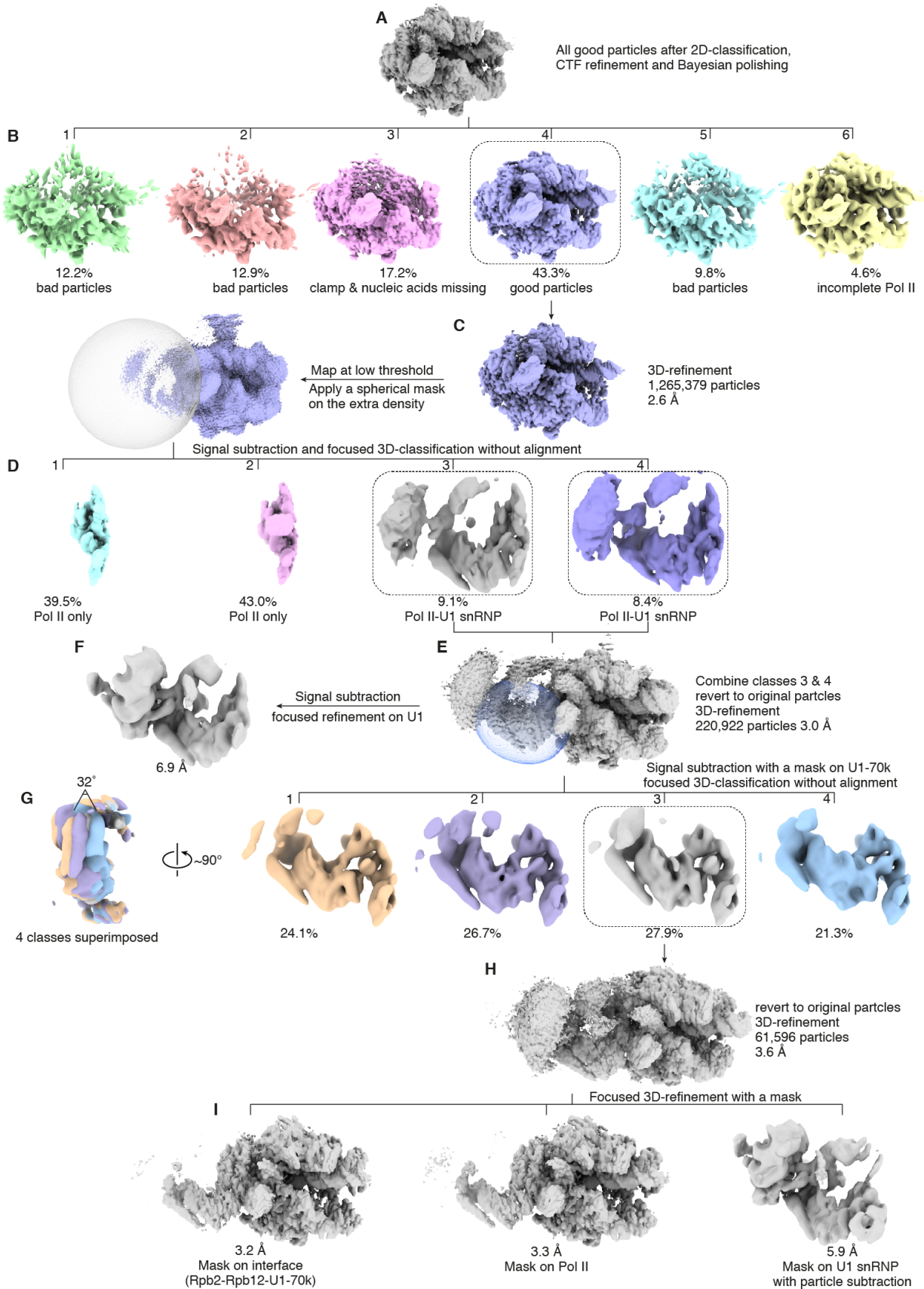
Initial models of *Sus scrofa domesticus* Pol II (PDB: 6GMH (15)) and human U1 snRNP (PDB: 3PGW (16), 4PJO and 4PKD (17)) were rigid-body fitted into the overall map in Chimera (43). The Pol II model was manually adjusted in Coot (44) using the high-resolution map of Pol II (2.6 Å) and the structure was real-space refined in PHENIX (45). The refined Pol II model was then fitted into the overall Pol-U1 snRNP map and adjusted in Coot. The U1 snRNP model was first rebuilt in the focused refined map of U1 snRNP in Coot and the structure was real-space refined in PHENIX and used later as a restraint. Both Pol II and U1 snRNP models were fitted into the interface focused refined map, and the models for RPB2, RPB12, U1-70k and stem loop I of U1 snRNA were rebuilt in Coot. The resulting complete model of Pol II-U1 snRNP was then real-space refined in the overall map using PHENIX and structure restraints of U1 snRNP.

We could not observe density for the exposed Pol II stalk RPB4/RPB7, as it is often the case for Pol II structures. In addition, no density corresponding to U1-A was observed, likely due to the high flexibility of the stem loop II of U1 snRNP. We therefore modelled the Pol II stalk and U1-A by superposition of previous structures of Pol II and U1 snRNP (PDB: 6GMH (15), 3PGW (16), 4PJO and 4PKD (17)). We also did not observe any density for U1-C. Comparison of the Pol II-U1 snRNP structure with the U1 snRNP crystal structure showed that when superimposed on the Sm-ring, the 5'SS-U1 duplex moves upwards, leaving a bigger gap between itself and the Sm-ring (**fig. S9D**). This may lead to dissociation of U1-C.

Figures were generated using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.), Chimera (43) and Chimera X (46). Sequence alignment was performed using Jalview (47).

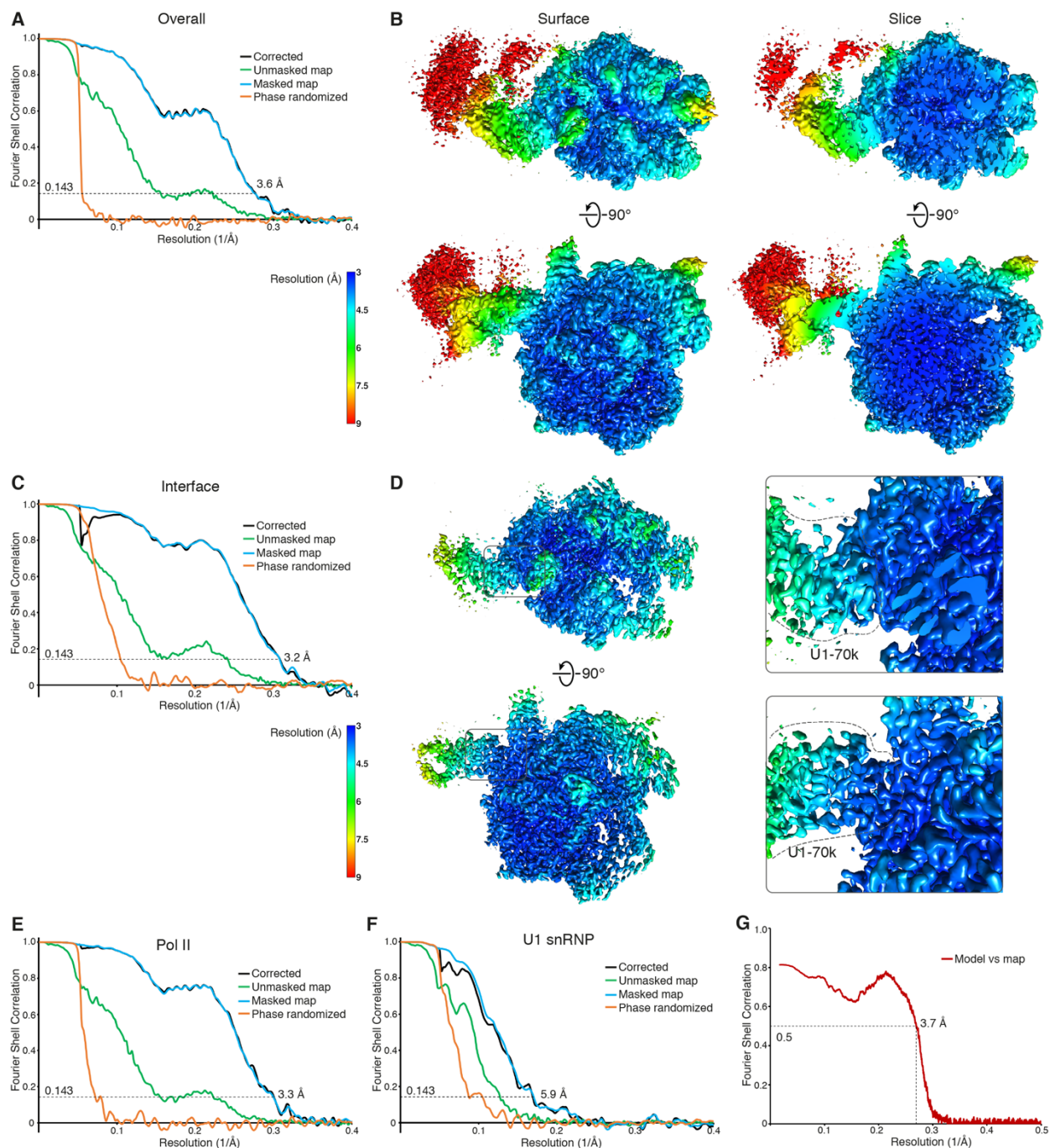


**Fig. S1. Preparation of the transcribing Pol II-U1 snRNP complex.** (A) Schematic of the DNA scaffold and the RNA constructs used for structural and biochemical analysis. The 15-nt RNA is designed to be completely covered within Pol II. The 145-nt RNA consists of three MS<sub>2</sub>-tags, a 5' exon and a truncated intron of 29 nt, of which 6 nt (linking RNA) are exposed between Pol II and U1 snRNP. The 155-nt RNA contains an extension of 10 nt within the linking RNA, while the 189-nt RNA contains a linking RNA of 50 nt, sufficient for assembly of the B complex. (B) Purified RNAs on a 20% denaturing gel (15-nt RNA, lane 1) or an 8% denaturing gel (145-nt, 155-nt and 189-nt RNAs), stained with Toluidine Blue. (C) Purified human U1 snRNP and *Sus scrofa domestica* Pol II on 4-12% NuPAGE Bis-Tris gels run in MOPS, stained with Instant Blue. The Pol II-U1 snRNP on lane 3 was used for freezing grids. (D) Representative micrograph collected on the FEI 300 kV Titan Krios with a K3 detector in electron counting mode and a 40° tilt. (E) Two-dimensional averages of the transcribing Pol II-U1 snRNP complex containing the 145-nt RNA, with a dimension of 237 Å across Pol II and U1 snRNP. (F) Angular distribution plot of the overall map, with the scale showing the number of particles assigned to a particular angle. The ring-shape of the angular distribution is resulted from the tilted data collection to resolve the preferred orientation of the complex.



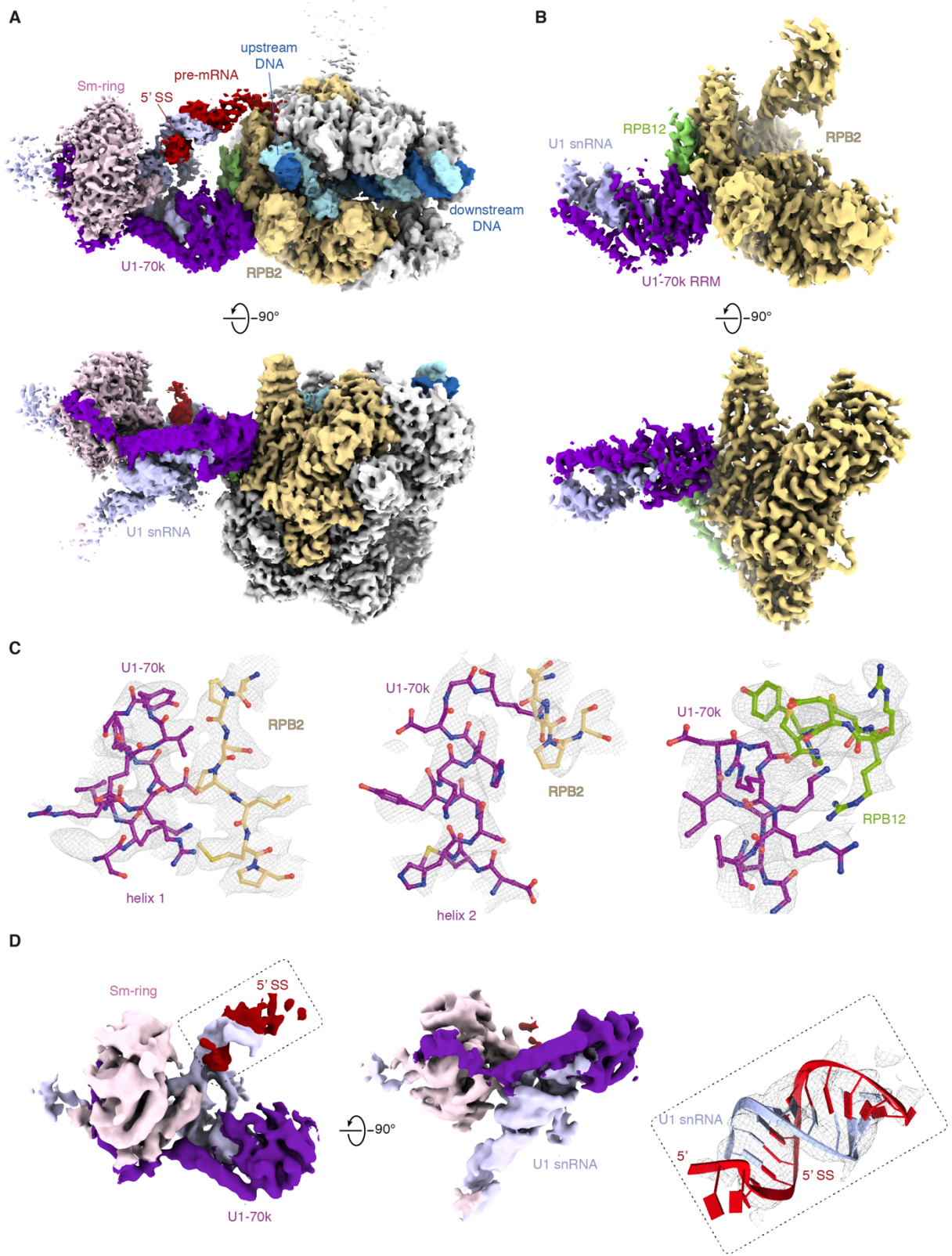
**Fig. S2. Cryo-EM data processing.** After removing bad particles and U1 snRNP alone particles with two-dimensional classification, particles were combined and 3D-refined, followed by CTF refinement and Bayesian polishing in RELION 3.0 (A). The particles were subsequently divided into six classes by 3D-classification with a fine-angle local search and the remaining bad particles as well as incomplete Pol II particles were removed (B). Good particles were 3D-refined to obtain a high-resolution map of Pol II (C). Following signal subtraction of Pol II and focused 3D-classification without alignment of the subtracted particles with a large spherical soft mask onto the extra density, the classes containing densities corresponding to U1 snRNP were selected and combined (D). This resulted in a reconstruction of Pol II-U1 snRNP at 3.0 Å resolution (E), with a resolution of 6.9 Å for U1 snRNP (F). The movement of U1 snRNP relative to Pol II was further separated by signal subtraction of Pol II and focused 3D-classification without alignment with a soft spherical mask on U1-70k (G). The final reconstruction has an overall resolution of 3.6 Å (H) and focused 3D-refinement on Pol II, the Pol II-U1 snRNP interface and U1 snRNP improved the resolution of respective regions and aided model building (I).





**Fig. S3. Local resolution maps and FSC curves.** (A) Gold-standard Fourier Shell Correlation (FSC) curves of the overall map of Pol II-U1 snRNP. (B) Local resolution estimations of the overall map in both surface and slice views showed that the Pol II core is beyond 3 Å (deep blue), while the Pol II-U1 snRNP interface is around 4-4.5 Å (cyan) and the rest of U1 snRNP in the range of 6-9 Å (green-red). (C) FSC curves of the focused refined map at the Pol II-U1 snRNP interface. (D) Local resolution estimation of the interface focused refined map showed improved

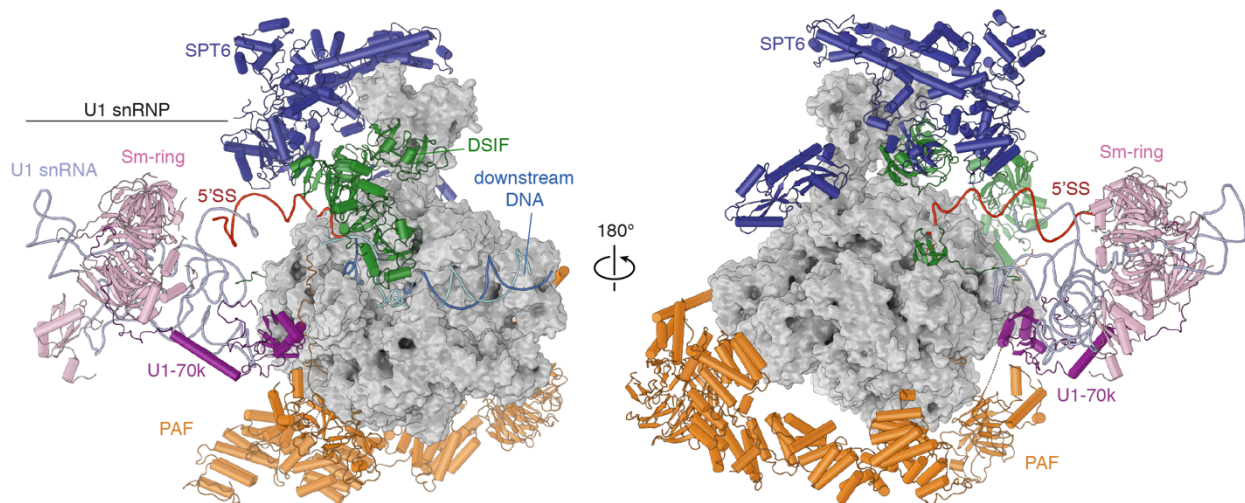
densities for U1-70k RRM at the interface (3-4 Å), where densities for bulky side chains could be observed. **(E and F)** FSC curves of the focused refined maps on Pol II (E) and U1 snRNP (F), respectively. **(G)** Model versus map FSC for the overall map using the FSC standard of 0.5.



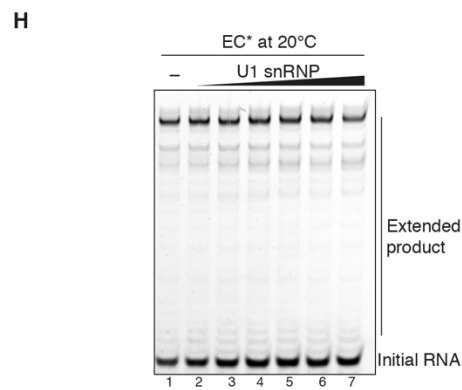
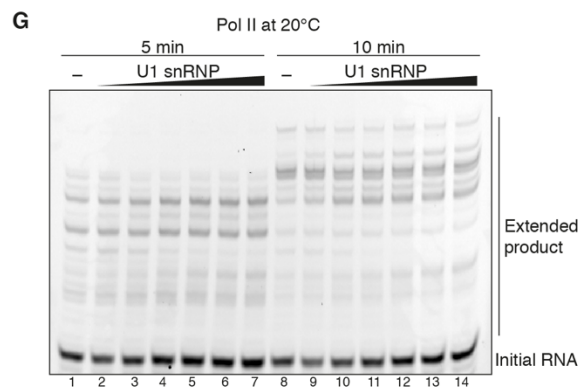
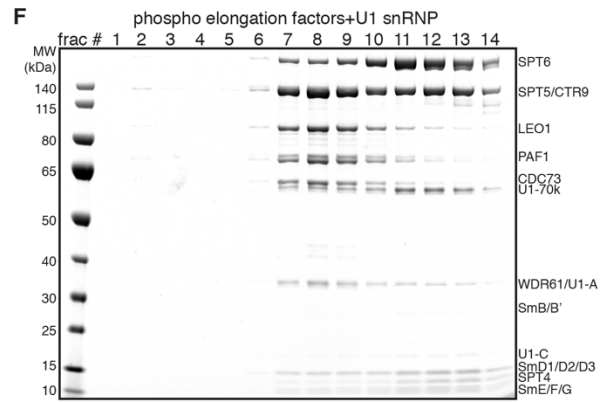
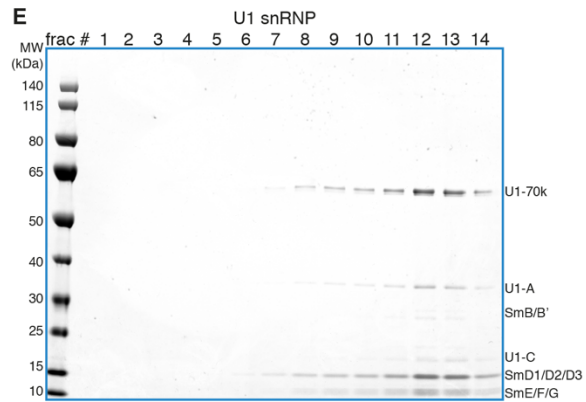
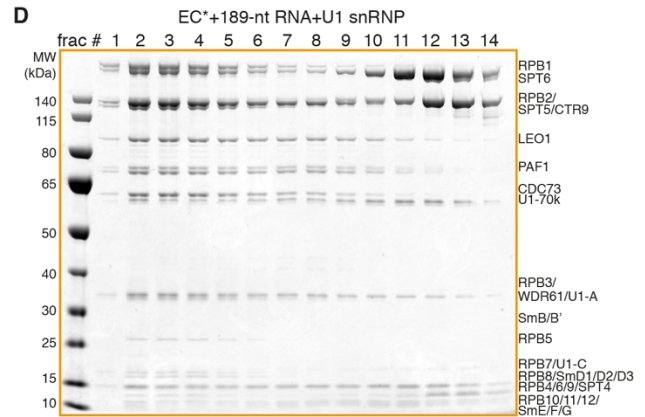
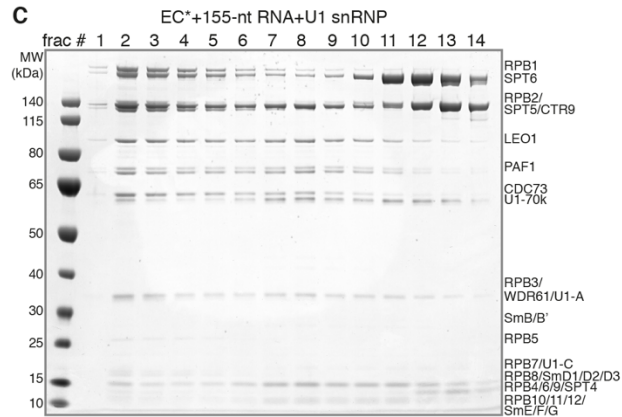
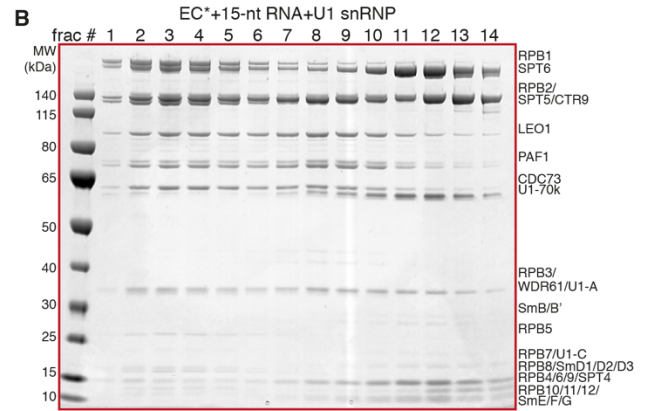
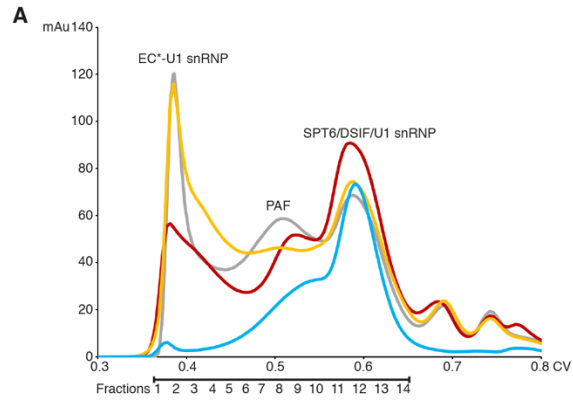
**Fig. S4. Cryo-EM densities of Pol II-U1 snRNP.** (A) Cryo-EM densities of the overall map, where weak densities were observed for the linking RNA and the Pol II stalk. The views and colors are the same as in Fig. 1, B and C. (B) Cryo-EM densities of the focused refined map at the interface region, with densities for U1-70k colored in purple, RPB2 in gold and RPB12 in green. (C) Cryo-EM densities at the Pol II-U1 snRNP interface shown in grey mesh with models of U1-70k, RPB2 and RPB12 in cartoon. Densities for bulky side chains at the interface were clearly resolved. (D) Cryo-EM densities of the focused refined map for U1 snRNP in two views, with improved densities for U1 snRNA and Sm-ring. Densities for the 5'SS-U1 snRNA duplex are shown in grey mesh with the RNA model fitted.





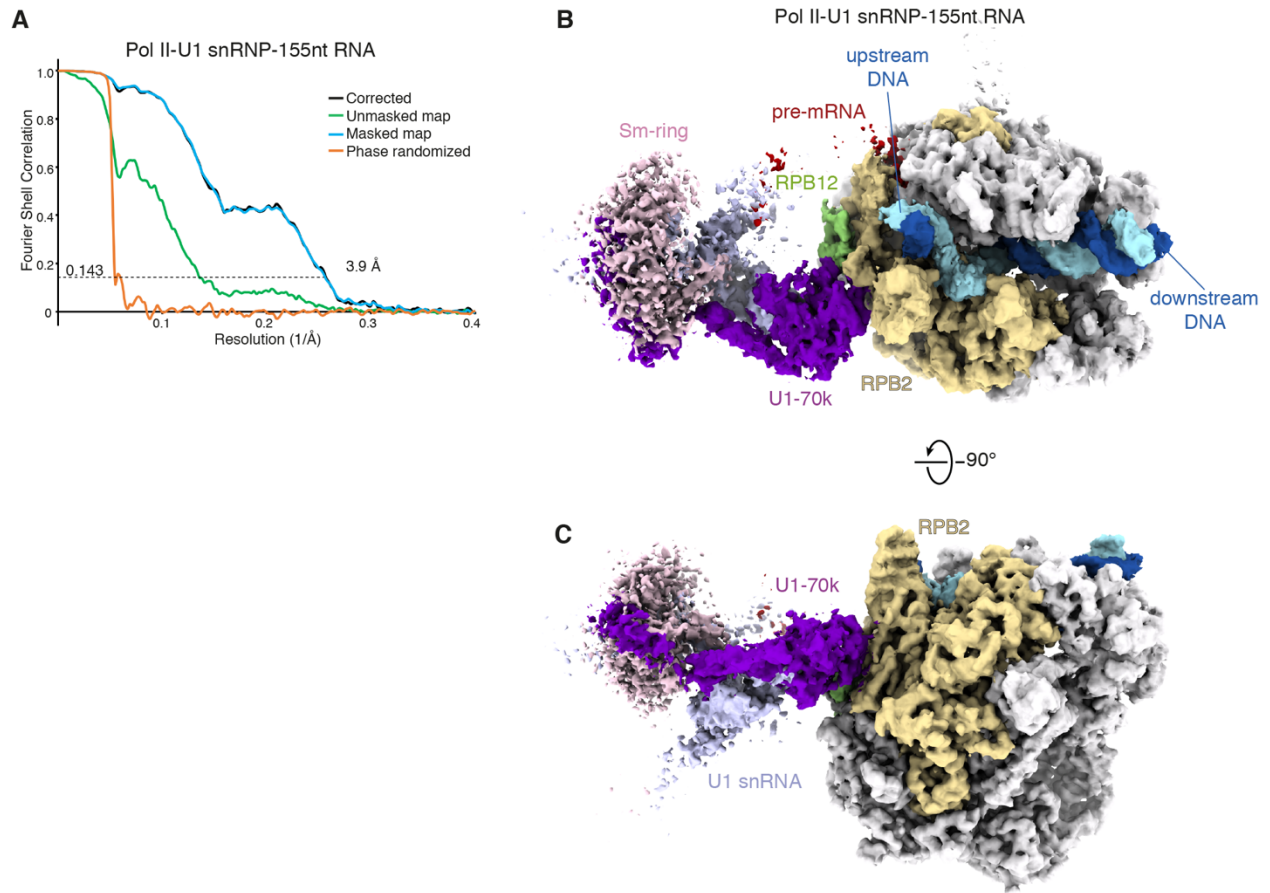


**Fig. S6. Superposition with activated elongation complex EC\*.** Superposition of the Pol II-U1 snRNP complex with the activated transcription elongation complex EC\* (PDB: 6GMH (15)). Pol II is shown in surface representation in grey, while U1 snRNP is colored in pink except for U1-70k in purple. Template DNA is depicted as blue ribbon and non-template DNA in cyan. RNA is shown as red ribbon. The elongation factors SPT6 (dark blue), DSIF (green) and PAF (orange) are compatible with U1 snRNP on the Pol II surface. The view in the left panel is the same as Fig. 1B.

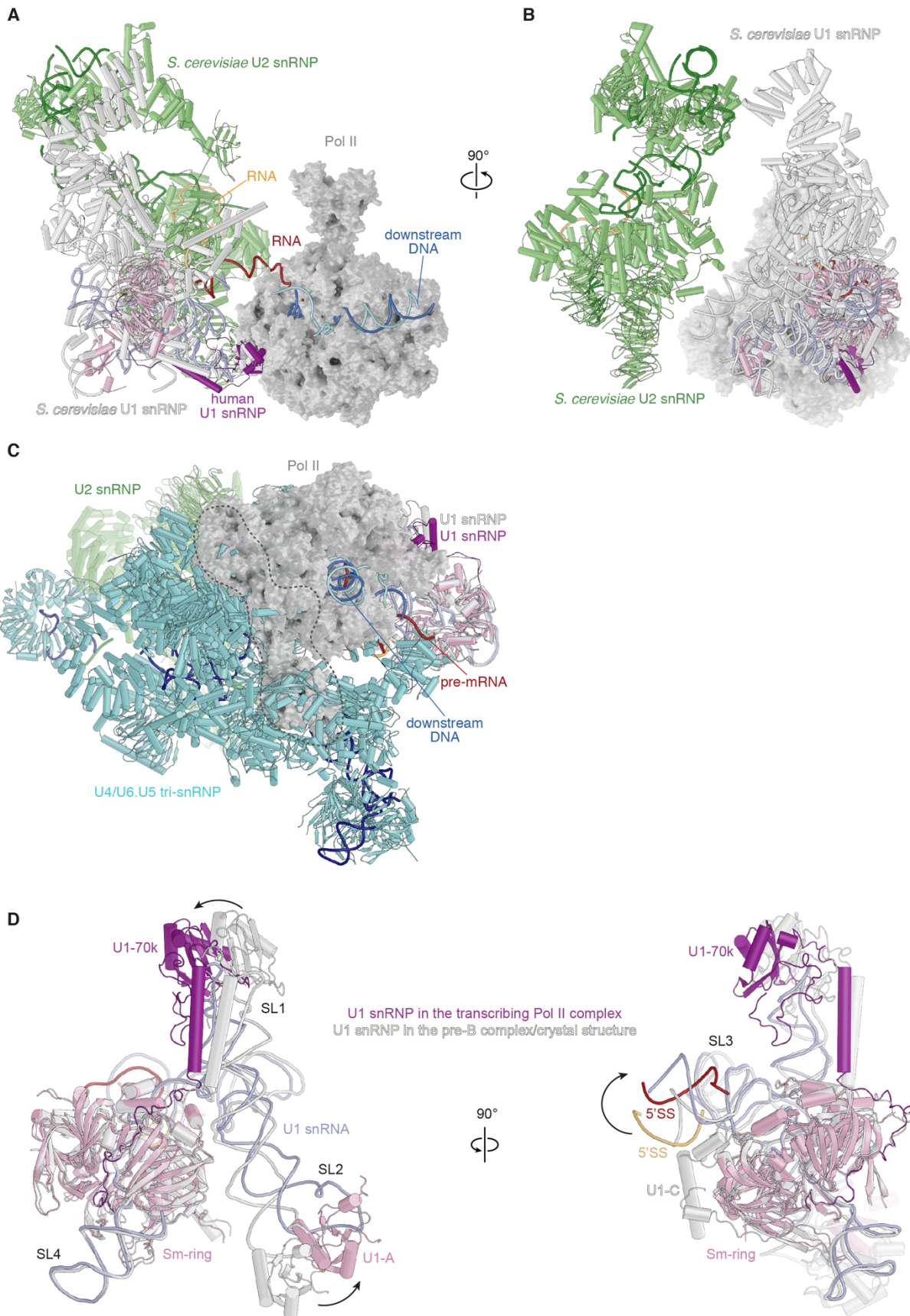


**Fig. S7. Analytical size exclusion assay for EC\*-U1 snRNP complex formation and RNA extension assays.** (A) Size exclusion chromatograms of complex formation between the activated transcription elongation complex (EC\*) and U1 snRNP using three different RNA constructs. Profiles for EC\*-U1 snRNP complex formation with 15-nt RNA (red), 155-nt RNA (grey) and 189-nt RNA (orange) as well as a control profile of U1 snRNP alone (blue) are shown. Fractions used for analysis on SDS-PAGE are marked. (B to F) Fractions of the analytical size exclusion assays were analyzed on a 4-12% NuPAGE Bis-Tris gel for EC\*-U1 snRNP complex formation with 15-nt RNA (B), 155-nt RNA (C) and 189-nt RNA (D), as well as for the control runs with U1 snRNP alone (E) and U1 snRNP together with phosphorylated elongation factors (F). Both the 155-nt and 189-nt RNAs enabled EC\*-U1 snRNP complex formation. The minimal 15-nt RNA that lacks the 5'SS still allowed U1 snRNP binding, although at apparently sub-stoichiometric levels. (G and H) RNA extension assays showed that U1 snRNP binding does not affect the transcription speed for either phosphorylated Pol II alone (G) or EC\* (H). A titration of U1 snRNP from 0 to 10x molar excess to Pol II was used. All assays were repeated in triplicate.





**Fig. S8. Cryo-EM analysis of alternative Pol II-U1 snRNP complex with 155-nt RNA.** (A) FSC curves of the overall map of the Pol II-U1 snRNP complex with 155-nt RNA (10-nt extension in the linking RNA). (B and C) Cryo-EM densities of the Pol II-U1 snRNP complex with 155-nt RNA, where the linking RNA became more mobile compared to 145-nt RNA. The same protein interface among U1-70k, RPB2 and RPB12 was observed. The views and colors are the same as in Fig. 1, B and C, and fig. S4A.



**Fig. S9. Superposition with spliceosome structures.** (A and B) Superposition of the mammalian Pol II-U1 snRNP structure with the yeast pre-spliceosome or A complex (PDB: 6G90 (21)) on the U1 snRNP Sm-ring revealed that A complex formation can happen on the surface of Pol II (grey surface). *S. cerevisiae* U1 snRNP is in white and *S. cerevisiae* U2 snRNP in light green. The pre-mRNA of the Pol II-U1 snRNP complex is colored in red, while that of the yeast A complex in orange. (C) Superposition of the Pol II-U1 snRNP structure with the human pre-B complex (PDB: 6QX9 (23)) on the U1 snRNP Sm-ring showed some clashes (highlighted with dashed lines) between Pol II (grey surface) and tri-snRNP (cyan). However, the flexibility of the pre-B complex may enable structural rearrangements (compare main text). (D) Comparison of U1 snRNP structure as observed in our Pol II-U1 snRNP complex (colored the same as Fig. 1B) with the U1 snRNP crystal structure (white). Superposition on the Sm-ring showed a movement of SL1-3 with associated proteins. In particular, the 5'SS-U1 duplex was moved away from the Sm-ring, leaving a larger gap at the binding site for U1-C (5'SS of Pol II-U1 snRNP in red, 5'SS of crystal structure in orange). The views are related by a 90° rotation as indicated.

**Table S1. Statistics of cryo-EM reconstructions and structural model.**

	Transcribing Pol II-U1 snRNP complex			
	145-nt RNA			155-nt RNA
	Overall (EMD-11972)	Interface (EMD-11973)	U1 snRNP (EMD-11974)	Overall (EMD-11975)
<b>Data collection and processing</b>				
Microscope	FEI Titan Krios			FEI Titan Krios
Voltage (kV)	300			300
Camera	Gatan K3			Gatan K3
Magnification	81 000 x			81 000 x
Pixel size (Å/pixel)	1.05			1.05
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	40.4			40.4
Exposure rate (e <sup>-</sup> /Å <sup>2</sup> /frame)	1.01			1.01
Number of frames per movie	40			40
Defocus range (µm)	0.5-2.0			0.5-2.0
Automation software	SerialEM			SerialEM
Tilt angle (°)	40			40
Energy filter slit width (eV)	20			20
Symmetry imposed	C1			C1
Initial particle numbers	1,265,379			320,847
Final particle numbers	61,596			14,497
Map sharpening <i>B</i> factor (Å <sup>2</sup> )	-72.73	-69.66	-360.19	-76.06
Map resolution (Å, FSC=0.143)	3.6	3.2	5.9	3.9
<b>Refinement</b>				
Initial models used (PDB code)	6GMH, 3PGW, 4PJO, 4PKD			
Model resolution (Å)	3.7			
Model composition				
Non-hydrogen atoms	43,792			
Protein residues	4,745			
Nucleic acid residues	274			
Ligands	Zn:8 Mg:1			
<i>B</i> factors (Å <sup>2</sup> )				
Protein	74.31			
Nucleotide	230.98			
Ligand	72.36			
R.m.s. deviations				
Bond lengths (Å)	0.002			
Bond angles (°)	0.416			
Validation				
MolProbity score	1.23			
Clash score	4.27			
Poor rotamers (%)	0.02			
CB deviations (%)	0			
Ramachandran plot				
Favored (%)	97.89			
Allowed (%)	2.11			
Outliers (%)	0.0			
PDB code	7B0Y			

**Movie S1. Cryo-EM densities of transcribing Pol II-U1 snRNP complex.** This movie shows the cryo-EM densities of the overall map, the focused refined map of U1 snRNP and the focused refined map of the Pol II-U1 snRNP interface.

**Movie S2. Structure of transcribing Pol II-U1 snRNP complex.** This movie shows the overall structure and a close-up view of the Pol II-U1 snRNP interface.

## References and Notes

1. K. M. Kotovic, D. Lockshon, L. Boric, K. M. Neugebauer, Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Mol. Cell. Biol.* **23**, 5768–5779 (2003). [doi:10.1128/MCB.23.16.5768-5779.2003](https://doi.org/10.1128/MCB.23.16.5768-5779.2003) [Medline](#)
2. S. A. Lacadie, M. Rosbash, Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. *Mol. Cell* **19**, 65–75 (2005). [doi:10.1016/j.molcel.2005.05.006](https://doi.org/10.1016/j.molcel.2005.05.006) [Medline](#)
3. I. Listerman, A. K. Sapra, K. M. Neugebauer, Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat. Struct. Mol. Biol.* **13**, 815–822 (2006). [doi:10.1038/nsmb1135](https://doi.org/10.1038/nsmb1135) [Medline](#)
4. E. W. J. Wallace, J. D. Beggs, Extremely fast and incredibly close: Cotranscriptional splicing in budding yeast. *RNA* **23**, 601–610 (2017). [doi:10.1261/rna.060830.117](https://doi.org/10.1261/rna.060830.117) [Medline](#)
5. G. Bird, D. A. Zorio, D. L. Bentley, RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. *Mol. Cell. Biol.* **24**, 8963–8969 (2004). [doi:10.1128/MCB.24.20.8963-8969.2004](https://doi.org/10.1128/MCB.24.20.8963-8969.2004) [Medline](#)
6. M. Tellier, I. Maudlin, S. Murphy, Transcription and splicing: A two-way street. *WIREs RNA* **11**, e1593 (2020). [doi:10.1002/wrna.1593](https://doi.org/10.1002/wrna.1593) [Medline](#)
7. P. J. Curtis, N. Mantei, C. Weissmann, Characterization and kinetics of synthesis of 15S beta-globin RNA, a putative precursor of beta-globin mRNA. *Cold Spring Harb. Symp. Quant. Biol.* **42**, 971–984 (1978). [doi:10.1101/SQB.1978.042.01.098](https://doi.org/10.1101/SQB.1978.042.01.098) [Medline](#)
8. M. J. Dye, N. Gromak, N. J. Proudfoot, Exon tethering in transcription by RNA polymerase II. *Mol. Cell* **21**, 849–859 (2006). [doi:10.1016/j.molcel.2006.01.032](https://doi.org/10.1016/j.molcel.2006.01.032) [Medline](#)
9. L. E. Giono, A. R. Kornblihtt, Linking transcription, RNA polymerase II elongation and alternative splicing. *Biochem. J.* **477**, 3091–3104 (2020). [doi:10.1042/BCJ20200475](https://doi.org/10.1042/BCJ20200475) [Medline](#)
10. M. de la Mata, C. R. Alonso, S. Kadener, J. P. Fededa, M. Blaustein, F. Pelisch, P. Cramer, D. Bentley, A. R. Kornblihtt, A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell* **12**, 525–532 (2003). [doi:10.1016/j.molcel.2003.08.001](https://doi.org/10.1016/j.molcel.2003.08.001) [Medline](#)
11. B. J. Natalizio, N. D. Robson-Dixon, M. A. Garcia-Blanco, The Carboxyl-terminal Domain of RNA Polymerase II Is Not Sufficient to Enhance the Efficiency of Pre-mRNA Capping or Splicing in the Context of a Different Polymerase. *J. Biol. Chem.* **284**, 8692–8702 (2009). [doi:10.1074/jbc.M806919200](https://doi.org/10.1074/jbc.M806919200) [Medline](#)
12. C. L. Will, R. Lührmann, Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* **3**, a003707 (2011). [doi:10.1101/cshperspect.a003707](https://doi.org/10.1101/cshperspect.a003707) [Medline](#)
13. S. M. Mount, I. Pettersson, M. Hinterberger, A. Karmas, J. A. Steitz, The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. *Cell* **33**, 509–518 (1983). [doi:10.1016/0092-8674\(83\)90432-4](https://doi.org/10.1016/0092-8674(83)90432-4) [Medline](#)
14. C. Boesler, N. Rigo, M. M. Anokhina, M. J. Tauchert, D. E. Agafonov, B. Kastner, H. Urlaub, R. Ficner, C. L. Will, R. Lührmann, A spliceosome intermediate with loosely

- associated tri-snRNP accumulates in the absence of Prp28 ATPase activity. *Nat. Commun.* **7**, 11997 (2016). [doi:10.1038/ncomms11997](https://doi.org/10.1038/ncomms11997) [Medline](#)
15. S. M. Vos, L. Farnung, M. Boehning, C. Wigge, A. Linden, H. Urlaub, P. Cramer, Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature* **560**, 607–612 (2018). [doi:10.1038/s41586-018-0440-4](https://doi.org/10.1038/s41586-018-0440-4) [Medline](#)
  16. G. Weber, S. Trowitzsch, B. Kastner, R. Lührmann, M. C. Wahl, Functional organization of the Sm core in the crystal structure of human U1 snRNP. *EMBO J.* **29**, 4172–4184 (2010). [doi:10.1038/emboj.2010.295](https://doi.org/10.1038/emboj.2010.295) [Medline](#)
  17. Y. Kondo, C. Oubridge, A. M. van Roon, K. Nagai, Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition. *eLife* **4**, e04986 (2015). [doi:10.7554/eLife.04986](https://doi.org/10.7554/eLife.04986) [Medline](#)
  18. D. A. Pomeranz Krummel, C. Oubridge, A. K. Leung, J. Li, K. Nagai, Crystal structure of human spliceosomal U1 snRNP at 5.5 Å resolution. *Nature* **458**, 475–480 (2009). [doi:10.1038/nature07851](https://doi.org/10.1038/nature07851) [Medline](#)
  19. B. Spiluttini, B. Gu, P. Belagal, A. S. Smirnova, V. T. Nguyen, C. Hébert, U. Schmidt, E. Bertrand, X. Darzacq, O. Bensaude, Splicing-independent recruitment of U1 snRNP to a transcription unit in living cells. *J. Cell Sci.* **123**, 2085–2093 (2010). [doi:10.1242/jcs.061358](https://doi.org/10.1242/jcs.061358) [Medline](#)
  20. K. Bertram, D. E. Agafonov, O. Dybkov, D. Haselbach, M. N. Leelaram, C. L. Will, H. Urlaub, B. Kastner, R. Lührmann, H. Stark, Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. *Cell* **170**, 701–713.e11 (2017). [doi:10.1016/j.cell.2017.07.011](https://doi.org/10.1016/j.cell.2017.07.011) [Medline](#)
  21. C. Plaschka, P. C. Lin, C. Charenton, K. Nagai, Prespliceosome structure provides insights into spliceosome assembly and regulation. *Nature* **559**, 419–422 (2018). [doi:10.1038/s41586-018-0323-8](https://doi.org/10.1038/s41586-018-0323-8) [Medline](#)
  22. S. Sharma, S. P. Wongpalee, A. Vashisht, J. A. Wohlschlegel, D. L. Black, Stem-loop 4 of U1 snRNA is essential for splicing and interacts with the U2 snRNP-specific SF3A1 protein during spliceosome assembly. *Genes Dev.* **28**, 2518–2531 (2014). [doi:10.1101/gad.248625.114](https://doi.org/10.1101/gad.248625.114) [Medline](#)
  23. C. Charenton, M. E. Wilkinson, K. Nagai, Mechanism of 5' splice site transfer for human spliceosome activation. *Science* **364**, 362–367 (2019). [doi:10.1126/science.aax3289](https://doi.org/10.1126/science.aax3289) [Medline](#)
  24. X. Zhan, C. Yan, X. Zhang, J. Lei, Y. Shi, Structures of the human pre-catalytic spliceosome and its precursor spliceosome. *Cell Res.* **28**, 1129–1140 (2018). [doi:10.1038/s41422-018-0094-7](https://doi.org/10.1038/s41422-018-0094-7) [Medline](#)
  25. B. L. Robberson, G. J. Cote, S. M. Berget, Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**, 84–94 (1990). [doi:10.1128/MCB.10.1.84](https://doi.org/10.1128/MCB.10.1.84) [Medline](#)
  26. J. Ule, B. J. Blencowe, Alternative Splicing Regulatory Networks: Functions, Mechanisms, and Evolution. *Mol. Cell* **76**, 329–345 (2019). [doi:10.1016/j.molcel.2019.09.017](https://doi.org/10.1016/j.molcel.2019.09.017) [Medline](#)



27. S. Vagner, U. Rügsegger, S. I. Gunderson, W. Keller, I. W. Mattaj, Position-dependent inhibition of the cleavage step of pre-mRNA 3'-end processing by U1 snRNP. *RNA* **6**, 178–188 (2000). [doi:10.1017/S1355838200991854](https://doi.org/10.1017/S1355838200991854) [Medline](#)
28. D. Kaida, M. G. Berg, I. Younis, M. Kasim, L. N. Singh, L. Wan, G. Dreyfuss, U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. *Nature* **468**, 664–668 (2010). [doi:10.1038/nature09479](https://doi.org/10.1038/nature09479) [Medline](#)
29. A. E. Almada, X. Wu, A. J. Kriz, C. B. Burge, P. A. Sharp, Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* **499**, 360–363 (2013). [doi:10.1038/nature12349](https://doi.org/10.1038/nature12349) [Medline](#)
30. P. K. Andersen, S. Lykke-Andersen, T. H. Jensen, Promoter-proximal polyadenylation sites reduce transcription activity. *Genes Dev.* **26**, 2169–2179 (2012). [doi:10.1101/gad.189126.112](https://doi.org/10.1101/gad.189126.112) [Medline](#)
31. M. Zillmann, M. L. Zapp, S. M. Berget, Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* **8**, 814–821 (1988). [doi:10.1128/MCB.8.2.814](https://doi.org/10.1128/MCB.8.2.814) [Medline](#)
32. A. Yuryev, M. Patturajan, Y. Litingtung, R. V. Joshi, C. Gentile, M. Gebara, J. L. Corden, The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 6975–6980 (1996). [doi:10.1073/pnas.93.14.6975](https://doi.org/10.1073/pnas.93.14.6975) [Medline](#)
33. T. Nojima, T. Gomes, A. R. F. Grosso, H. Kimura, M. J. Dye, S. Dhir, M. Carmo-Fonseca, N. J. Proudfoot, Mammalian NET-Seq Reveals Genome-wide Nascent Transcription Coupled to RNA Processing. *Cell* **161**, 526–540 (2015). [doi:10.1016/j.cell.2015.03.027](https://doi.org/10.1016/j.cell.2015.03.027) [Medline](#)
34. T. Misteli, D. L. Spector, RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell* **3**, 697–705 (1999). [doi:10.1016/S1097-2765\(01\)80002-2](https://doi.org/10.1016/S1097-2765(01)80002-2) [Medline](#)
35. K. M. Harlen, K. L. Trotta, E. E. Smith, M. M. Mosaheb, S. M. Fuchs, L. S. Churchman, Comprehensive RNA Polymerase II Interactomes Reveal Distinct and Varied Roles for Each Phospho-CTD Residue. *Cell Rep.* **15**, 2147–2158 (2016). [doi:10.1016/j.celrep.2016.05.010](https://doi.org/10.1016/j.celrep.2016.05.010) [Medline](#)
36. C. J. David, A. R. Boyne, S. R. Millhouse, J. L. Manley, The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev.* **25**, 972–983 (2011). [doi:10.1101/gad.2038011](https://doi.org/10.1101/gad.2038011) [Medline](#)
37. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005). [doi:10.1016/j.jsb.2005.07.007](https://doi.org/10.1016/j.jsb.2005.07.007) [Medline](#)
38. D. Tegunov, P. Cramer, Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019). [doi:10.1038/s41592-019-0580-y](https://doi.org/10.1038/s41592-019-0580-y) [Medline](#)
39. S. H. Scheres, Processing of Structurally Heterogeneous Cryo-EM Data in RELION. *Methods Enzymol.* **579**, 125–157 (2016). [doi:10.1016/bs.mie.2016.04.012](https://doi.org/10.1016/bs.mie.2016.04.012) [Medline](#)



40. S. H. Scheres, A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415**, 406–418 (2012). [doi:10.1016/j.jmb.2011.11.010](https://doi.org/10.1016/j.jmb.2011.11.010) [Medline](#)
41. S. H. Scheres, RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012). [doi:10.1016/j.jsb.2012.09.006](https://doi.org/10.1016/j.jsb.2012.09.006) [Medline](#)
42. T. C. Terwilliger, O. V. Sobolev, P. V. Afonine, P. D. Adams, Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr. D* **74**, 545–559 (2018). [doi:10.1107/S2059798318004655](https://doi.org/10.1107/S2059798318004655) [Medline](#)
43. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). [doi:10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084) [Medline](#)
44. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010). [doi:10.1107/S0907444910007493](https://doi.org/10.1107/S0907444910007493) [Medline](#)
45. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019). [doi:10.1107/S2059798319011471](https://doi.org/10.1107/S2059798319011471) [Medline](#)
46. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018). [doi:10.1002/pro.3235](https://doi.org/10.1002/pro.3235) [Medline](#)
47. A. M. Waterhouse, J. B. Procter, D. M. Martin, M. Clamp, G. J. Barton, Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009). [doi:10.1093/bioinformatics/btp033](https://doi.org/10.1093/bioinformatics/btp033) [Medline](#)