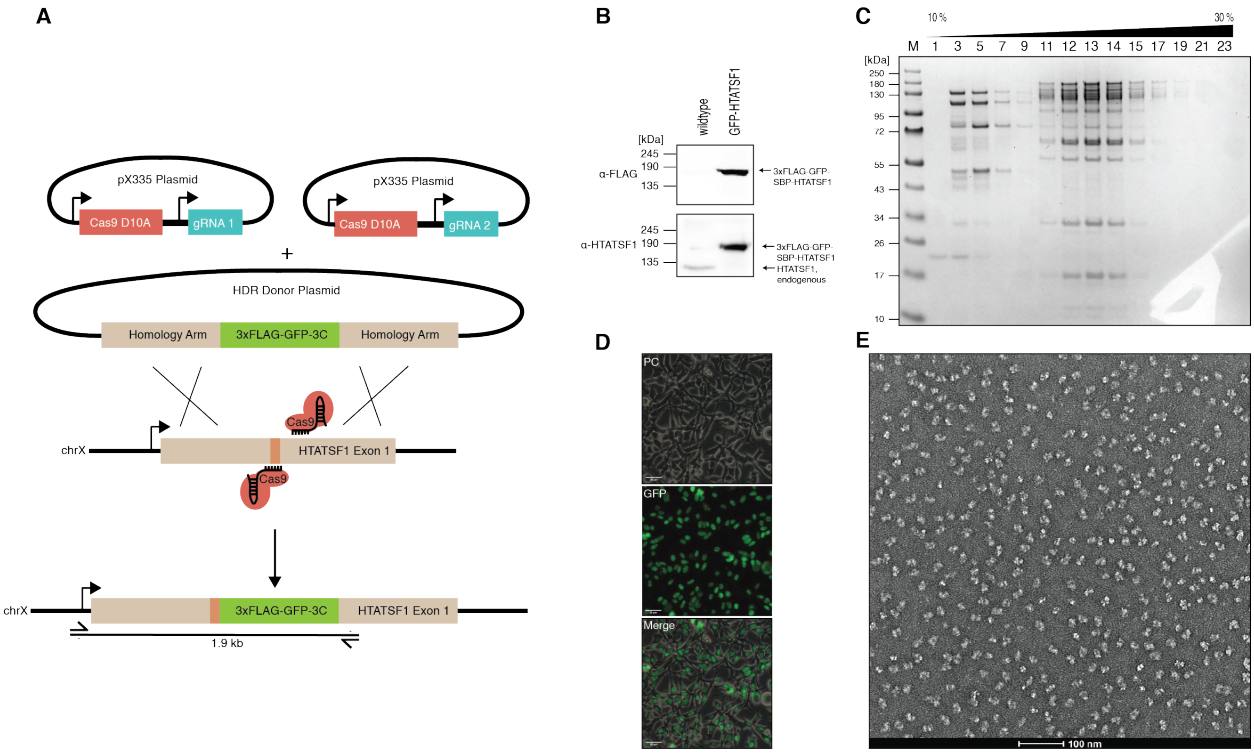


Fig. S1. Isolation of HTATSF1-bound U2 snRNP from CRISPR/Cas9 edited HEK293F cells.

(A) Schematic of CRISPR-Cas9-mediated genome editing. gRNA, guide RNA; chrX, chromosome X; HDR, homology-directed repair. **(B)** Validation of the genome editing by Western Blot. **(C)** SDS-PAGE of glycerol gradient ultracentrifugation fractions. **(D)** Expression and localization of the GFP-tagged HTATSF1 in adherently growing HEK293F cells. **(E)** Negative stain EM micrograph of the isolated sample.

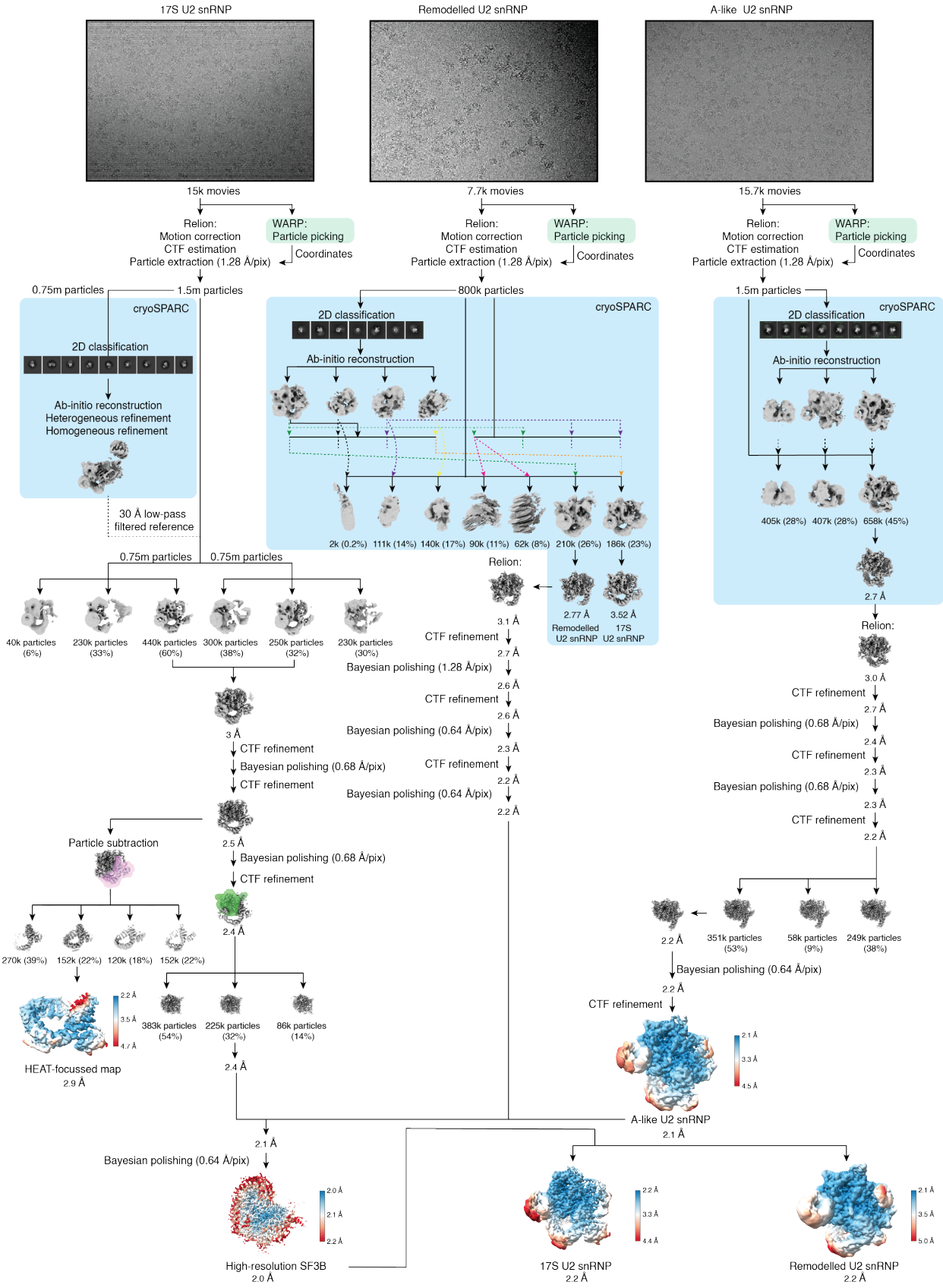

17S U2 snRNP
Remodelled U2 snRNP
A-like U2 snRNP
15k movies
7.7k movies
15.7k movies
Relion:
Motion correction
CTF estimation
Particle extraction (1.28 Å/pix)
WARP:
Particle picking
Coordinates
0.75m particles
1.5m particles
800k particles
cryoSPARC
2D classification
Ab-initio reconstruction
Heterogeneous refinement
Homogeneous refinement
30 Å low-pass filtered reference
2k (0.2%)
111k (14%)
140k (17%)
90k (11%)
62k (8%)
210k (26%)
186k (23%)
405k (28%)
407k (28%)
658k (45%)
2.77 Å
Remodelled U2 snRNP
3.52 Å
17S U2 snRNP
2.7 Å
3.1 Å
CTF refinement
2.7 Å
Bayesian polishing (1.28 Å/pix)
2.6 Å
2.6 Å
Bayesian polishing (0.64 Å/pix)
2.3 Å
2.2 Å
2.2 Å
3.0 Å
2.7 Å
2.4 Å
2.3 Å
2.3 Å
40k particles (6%)
230k particles (33%)
440k particles (60%)
300k particles (38%)
250k particles (32%)
230k particles (30%)
3 Å
2.5 Å
Particle subtraction
270k (39%)
152k (22%)
120k (18%)
152k (22%)
2.4 Å
383k particles (54%)
225k particles (32%)
86k particles (14%)
2.4 Å
HEAT-focussed map
2.9 Å
2.2 Å
351k particles (53%)
58k particles (9%)
249k particles (38%)
Bayesian polishing (0.64 Å/pix)
2.2 Å
A-like U2 snRNP
2.1 Å
2.1 Å
Bayesian polishing (0.64 Å/pix)
High-resolution SF3B
2.0 Å
17S U2 snRNP
2.2 Å
Remodelled U2 snRNP
2.2 Å

Fig. S2. Cryo-EM data processing. Processing chart for the 17S, remodelled and A-like U2 snRNP. The remodelled dataset contained particles of both the 17S and remodelled U2 snRNP which were separated by 3D classification. All three datasets were merged to obtain the high-resolution reconstruction of SF3B. The highest-resolution maps of 17S and remodelled U2 snRNP were obtained after separating the merged dataset followed by refinement. Arrows with solid lines indicate particles and dotted lines indicate volumes used as reference.

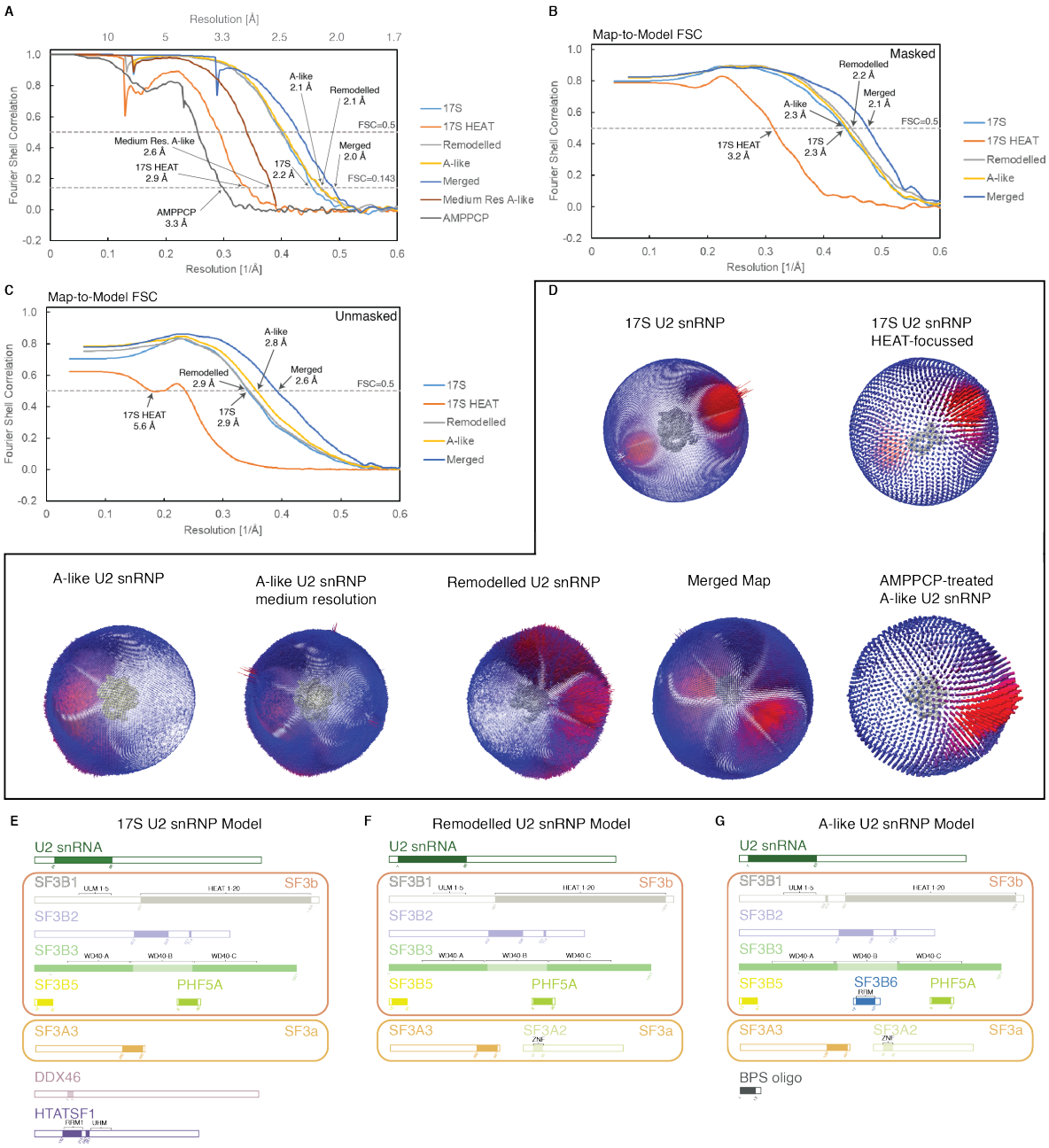


Fig. S3. FSC curves and angular distributions of the cryo-EM reconstructions. (A) Gold Standard FSC curves of masked maps determined in Relion 3.1 or, in the case of the AMP-PCP-treated A-like U2 snRNP map, in cryoSPARC v2.15; grey: AMP-PCP-treated A-like, orange: 17S HEAT, brown: medium resolution A-like, light blue: 17S consensus, yellow: A-like, light gray: remodelled, dark blue: merged map. The local minima in the FSC curves for the 17S HEAT map

and the merged SF3b map are likely caused by the tight mask for local refinement. **(B)** Masked map-to-model FSC curves. **(C)** Unmasked map-to-model FSC curves. **(D)** Angular distribution of particles used in the reconstruction of the maps. **(E)-(G)**, Overview of the modelled proteins and RNA in the structure of the 17S **(E)**, remodelled **(F)**, and A-like U2 snRNP **(G)**. Filled in rectangles denote the sequences that were modelled, opaque rectangles were modelled by fitting a crystal structure, empty rectangles are parts of the proteins and RNA that could not be assigned to the reconstructed maps. Selected domains important for this work are indicated.

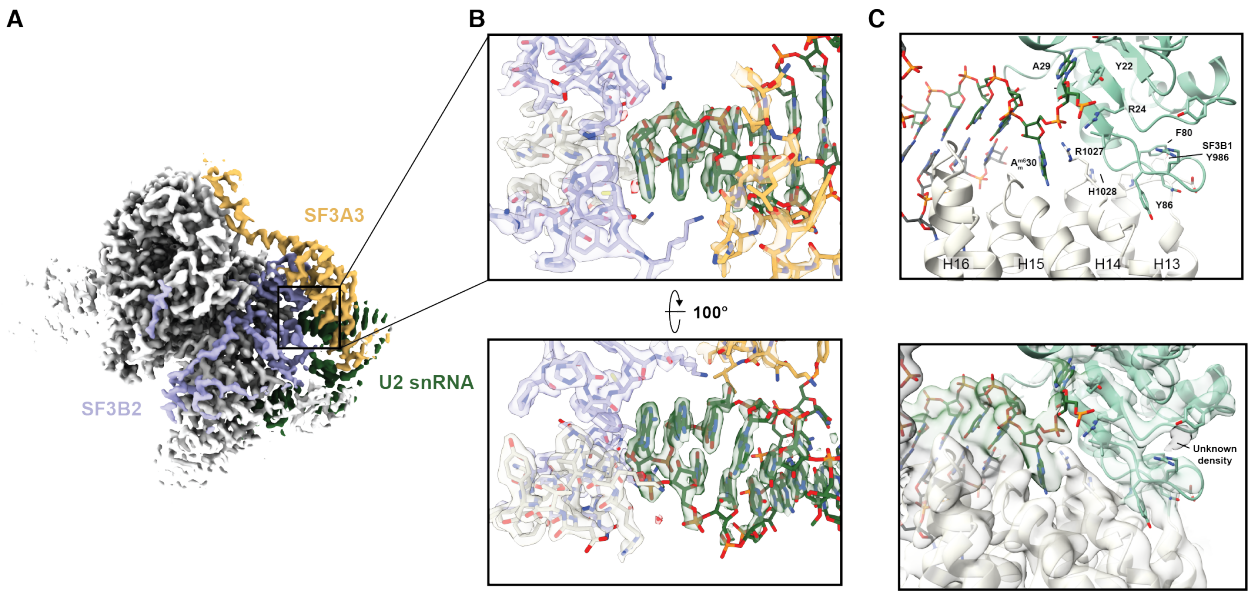


Fig. S4. High-resolution interfaces of U2 snRNA. (A) Position of SF3B2, SF3A3, and U2 snRNA in the U2 snRNP shown in the map of the remodelled U2 snRNP. **(B)** Close-up of the interactions of the stem loop IIa (SLIIa). **(C)** Upper panel: Atomic model of the interface of SF3B6 with SF3B1^{HEAT} and U2 snRNA. The methyl group modifications at the 2'-O and N6 of A30 were not modelled. Lower panel: The atomic model fit into the 0.5σ gaussian-filtered medium-resolution map of the A-like U2 snRNA.

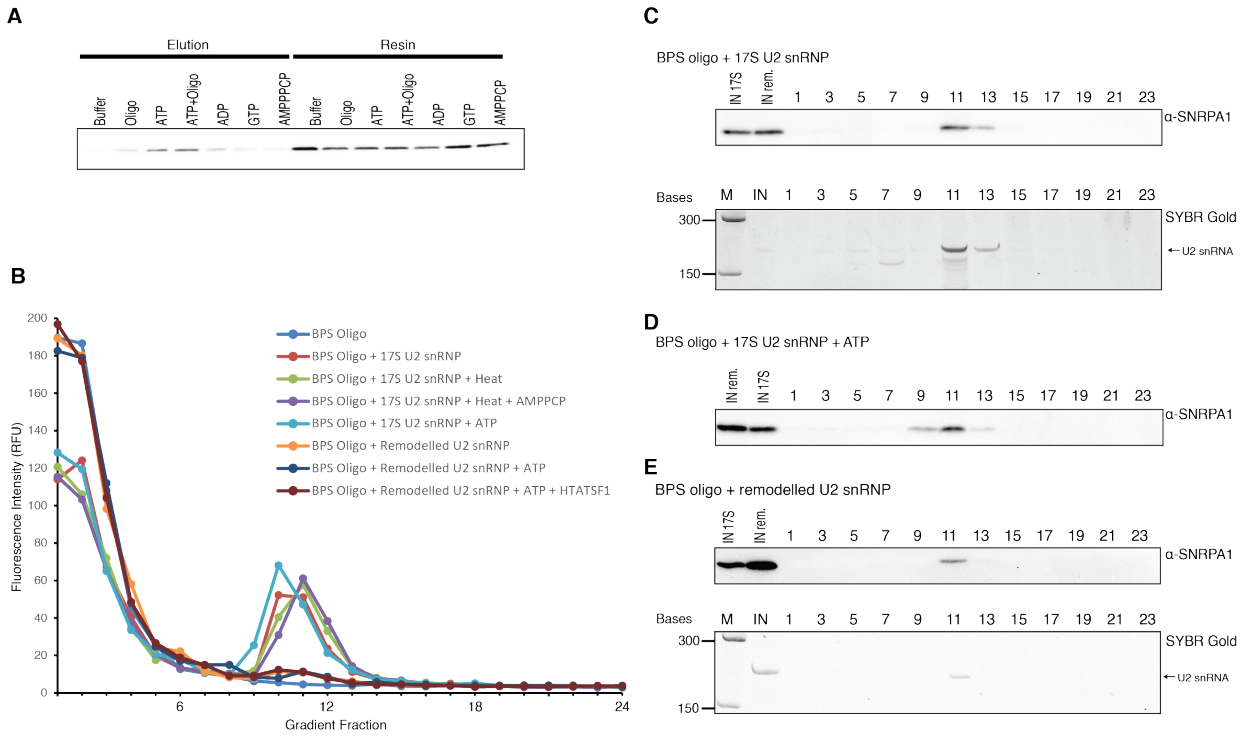


Fig. S5. Biochemical characterisation of the 17S U2 snRNP *in vitro* remodeling and engagement with BPS RNA. (A) Western blot analysis of the eluates from the GFP-HTATSF1 or GFP-DDX46 sample immobilised on the GFP nanobody resin and incubated under various conditions. The resin and eluted fractions were probed for the presence of the U2 snRNP with antibodies against the core U2 snRNP components SNRPA1 and SNRPB2. **(B)** Glycerol gradient analysis of the Cy5-labelled BPS oligonucleotide to the 17S U2 snRNP or remodelled U2 snRNP. The relative Cy5-fluorescence of each fraction of the glycerol gradients was plotted. RFU, relative fluorescence units. **(C-E)** Anti-SNRPA1 Western blot and SYBR Gold-stained Urea-PAGE analysis (to visualise RNA) of the glycerol gradients shown in (B) show that the U2 snRNP migrates at around fraction 11, the same fraction where the BPS oligo migrates when bound to the U2 snRNP. Specifically, the gradients were analysed after incubation of the BPS oligo with 17S U2 snRNP **(C)**, 17S U2 snRNP and ATP **(D)**, and with remodelled U2 snRNP **(E)**.

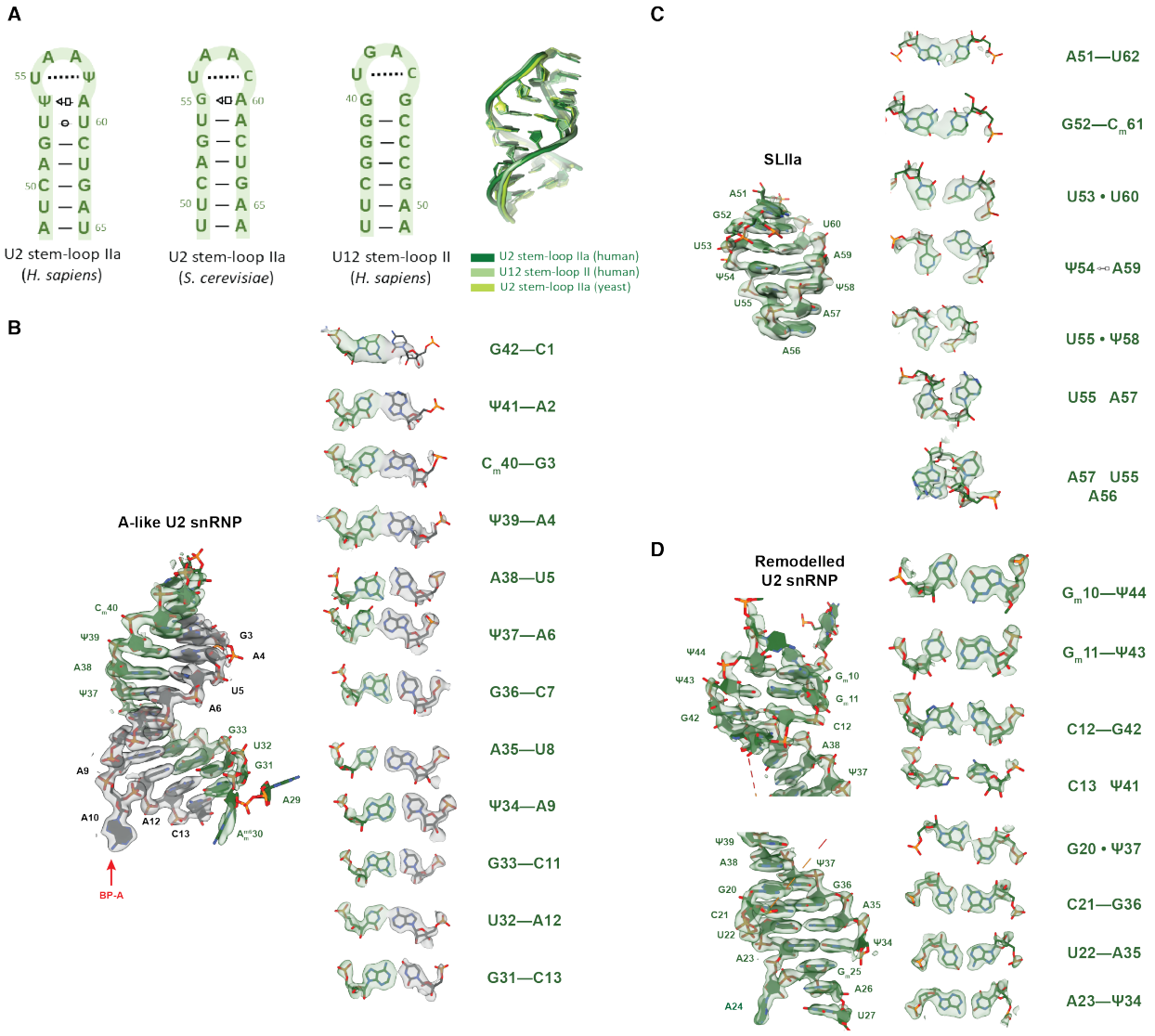


Fig. S6. High resolution modeling of the U2 snRNA. (A) Comparison of the structure of the SLIIa in human and yeast U2 snRNA, and human U12 snRNA. U2 snRNA residues U53 and U60 form trans-Watson-crick/Watson-crick interaction, while Ψ54 and A59 form a trans-Hoogsteen/Sugar Edge base-pair. Additionally, U55 forms a hydrogen bond with the phosphate backbone of the apical loop. This interaction cannot be achieved by any other nucleotide at this position, hence its evolutionary conservation. Despite significant differences in the base-pairing pattern between yeast and human SLIIa, both structures end up with nearly identical geometry (backbone RMSD=0.76 Å) **(B)** The atomic model of the branch helix fit into the 0.5σ gaussian-

filtered high-resolution map of the A-like U2 snRNP. **(C)** The atomic model of SLIIa fit into the map of the remodelled U2 snRNP. **(D)** The atomic model of the BMSL fit into the 1σ gaussian-filtered high-resolution map of the remodelled U2 snRNP.

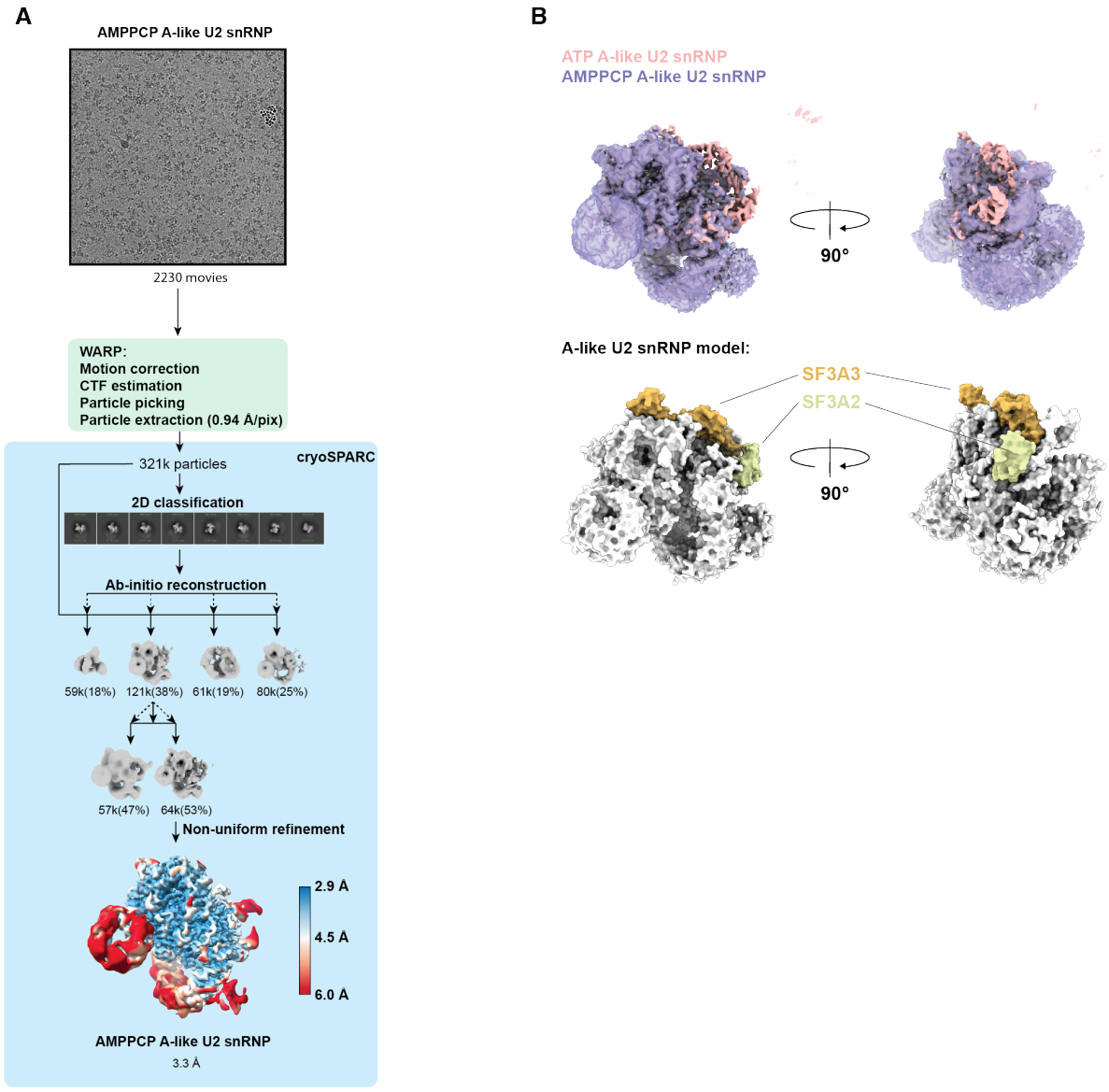


Fig. S7. Structure of the A-like U2 snRNP assembled in the presence of AMP-PCP. (A) Processing chart of 200kV/Falcon 3EC dataset of the A-like U2 snRNP assembled in the presence of AMP-PCP. **(B)** Superposition of the map of the A-like complex assembled in the presence of ATP (salmon) or AMP-PCP (purple). Comparison with the surface model of the A-like U2 snRNP shows that there is very little signal for SF3A3 and SF3A2 in the AMP-PCP A-like U2 snRNP when compared to the complex assembled in the presence of ATP. It is possible that ATP

hydrolysis and displacement of HTATSF1 facilitate docking of the SF3A complex to the A-like U2 snRNP.

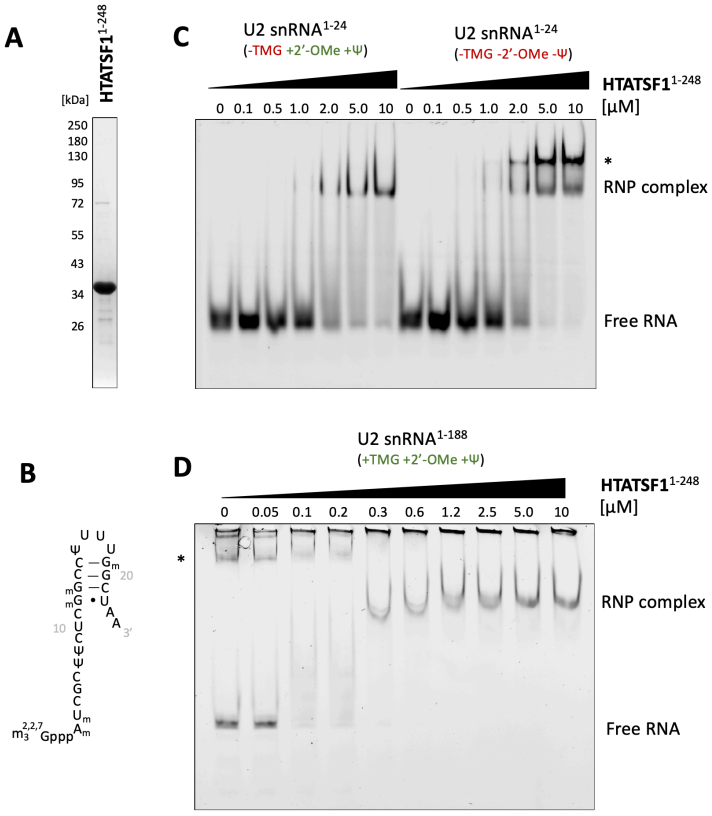
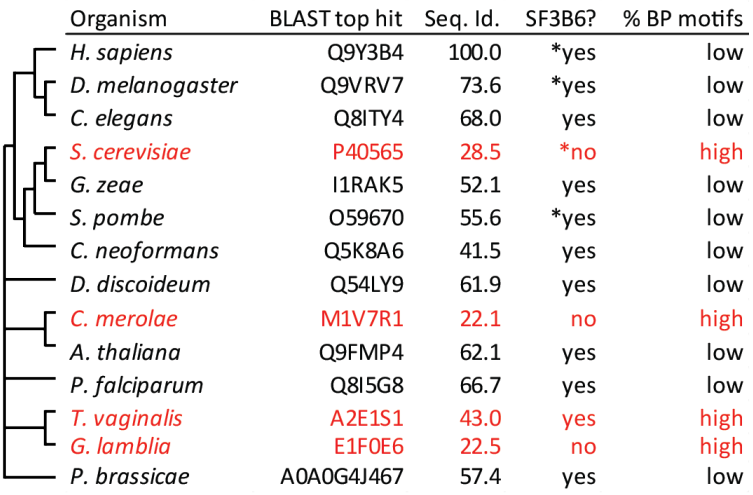


Fig. S8. Electrophoretic Mobility Shift Assay (EMSA) with HTATSF1$^{RRM/LH}$ and U2 snRNA. **(A)** SDS-PAGE of the recombinant HTATSF1$^{RRM/LH}$ sample expressed in *E.coli* and used for EMSA experiments. **(B)** Secondary structure of the U2 snRNA fragment used in the EMSA experiment. **(C)** Direct titration EMSA experiment for U2 snRNA 5'-end fragments (residues 1-24) synthesized with or without naturally occurring modifications (pseudouridylation and 2'-O methylation). **(D)** Direct titration EMSA experiment using full-length U2 snRNA purified from HEK293F cells.

A



Organism	BLAST top hit	Seq. Id.	SF3B6?	% BP motifs
H. sapiens	Q9Y3B4	100.0	*yes	low
D. melanogaster	Q9VRV7	73.6	*yes	low
C. elegans	Q8ITY4	68.0	yes	low
S. cerevisiae	P40565	28.5	*no	high
G. zeae	I1RAK5	52.1	yes	low
S. pombe	O59670	55.6	*yes	low
C. neoformans	Q5K8A6	41.5	yes	low
D. discoideum	Q54LY9	61.9	yes	low
C. merolae	M1V7R1	22.1	no	high
A. thaliana	Q9FMP4	62.1	yes	low
P. falciparum	Q8I5G8	66.7	yes	low
T. vaginalis	A2E1S1	43.0	yes	high
G. lamblia	E1F0E6	22.5	no	high
P. brassicae	A0A0G4J467	57.4	yes	low

B

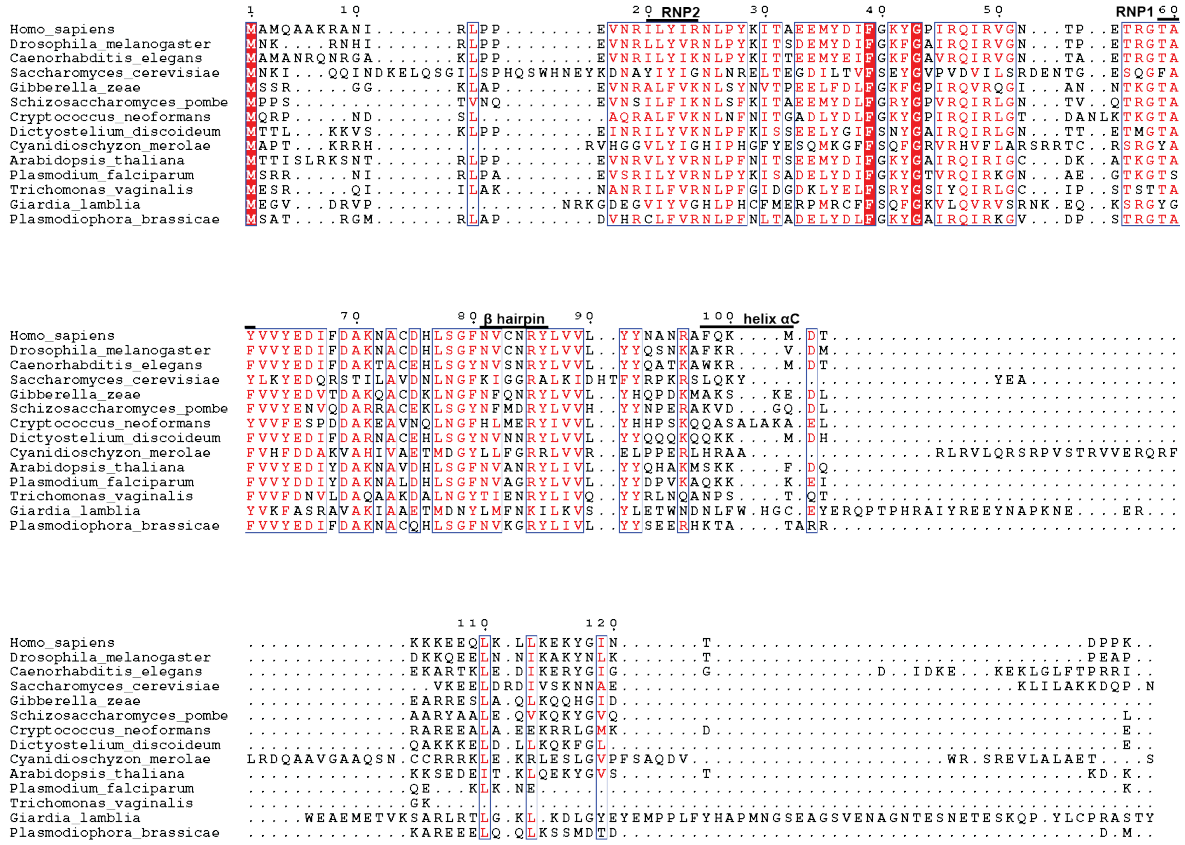


Fig. S9. SF3B6 sequence conservation. (A) Evolutionary co-occurrence analysis of SF3B6 and branch site sequence motifs. For a number of eukaryotic organisms, SF3B6 homologues were

identified by using Blastp against a database of all non-redundant protein sequences (2021/08/31) (*66*). Listed are the UniProt ID and sequence identity to human SF3B6. Based on the T-Coffee multiple sequence alignment (*67*, *68*) it was decided whether the BLAST hit is a likely SF3B6 homologue. Specifically, the conservation of the interface with SF2B1^{HEAT} (e.g. Y86), with U2 snRNA (12-16) and with SF3B1^{Nterm} was considered (around F98). * indicates that biochemical evidence for or against the SF3B6 homologue in this organism is available. The percentage of introns that contain a consensus branch sequence (% BP motifs) was determined previously (*69*). Organisms with strong consensus sequences are coloured in red. Many organisms with consensus branch sequences do not have an SF3B6 homologue. T. vaginalis may be an exemption, though its homologue differs from other SF3B6 homologues in the missing C-terminus, including the helix αC, which is involved in SF3B1^{Nterm} binding. **(B)** Multiple sequence alignment for SF3B6 from species listed in (A).

Table S1. Cryo-EM data collection, refinement, and validation statistics (part 1)

	17S U2 snRNP core PDB 7Q3L EMD-13793	17S U2 snRNP HEAT EMD-13810	**A-like U2 snRNP** PDB 7Q4O EMD-13811	A-like U2 snRNP Medium Res EMD-13813
Data collection and processing				
Microscope	TFS Krios		TFS Krios	
Voltage (keV)	300		300	
Camera	Gatan Quantum-K3		Gatan Quantum-K3	
Magnification	130kx		130kx	
Pixel size at detector (Å/pixel)	0.64		0.64	
Total electron exposure (e^{-}/Å^2)	53.45		52.06	
Exposure rate (e-/pixel/sec)	21.9		21.9	
Number of frames	40		40	
Defocus range (μm)	-0.8 to -1.8		-0.8 to -1.8	
Automation software	SerialEM		SerialEM	
Energy filter slit width	20 eV		20 eV	
Micrographs collected (no.)	15,531		15,681	
Micrographs used (no.)	15,479		15,598	
Total extracted particles (no.)	1,500,165		1,470,005	
For each reconstruction:				
Final particles (no.)	225,934	152,253	249,011	658,325
Point-group	C1	C1	C1	C1
Resolution (global, Å)				
FSC 0.5 (unmasked/masked)	(3.4/2.5)	(6.8/3.4)	(3.3/2.5)	(3.7/2.9)
FSC 0.143 (unmasked/masked)	(2.8/2.2)	(3.8/2.9)	(2.7/2.2)	(3.1/2.6)
Resolution range (local, Å)	2.2-7.6	2.9-5.9	2.1-3.3	2.6-8.3
Map sharpening *B* factor (Å^2) (main maps are unsharpened)	-44	-47	-40	-71
Model composition				
Protein (aa)	1903	579	2564	
Ligands (no.)	4	-	5	
RNA (bases)	41	-	48	
Waters (no.)	195	-	285	
Model refinement				
Refinement package	REFMAC5	REFMAC5	REFMAC5	
- real or reciprocal space	Reciprocal	Reciprocal	Reciprocal	
- resolution cutoff (Å)	2.3	3.0	2.2	
- Average FSC	0.7567	0.6331	0.7698	
B factors (Å^2)				
Protein residues	109	137	119	
Ligands	100	-	71	
RNA/DNA	254	-	99	
Waters	57	-	48	
R.m.s. deviations from ideal values				
Bond lengths (Å)	0.0063	0.0053	0.0053	
Bond angles (°)	1.6368	1.6856	1.4646	
Validation				
MolProbity score	1.43	1.11	1.91	
CaBLAM outliers	30 (1.7%)	7 (1.2%)	34 (1.4%)	
Clashscore	1.69	1.27	5.06	
Poor rotamers	40 (2.46%)	11 (2.21%)	97 (4.50%)	
C-beta deviations >0.25Å (%)	1.26	0.18	0.81%	
EMRinger score	4.41	3.30	4.17	
Ramachandran plot				
Favored (%)	96.71	98.78	97.22	
Outliers (%)	0.05	0.00	0.00	
Res. at FSC(map,model map)=0.5 (Å)	(2.9/2.3)	(5.6/3.2)	(2.8/2.3)	

Table S1. Cryo-EM data collection, refinement, and validation statistics (part 2)

	Remodelled U2 snRNP PDB 7Q4P EMD-13812	Merged Dataset EMD-13815	AMP-PCP A-like U2 snRNP EMD-13814
Data collection and processing			
Microscope	TFS Krios	TFS Krios	TFS Glacios
Voltage (keV)	300	300	200
Camera	Gatan Quantum-K3	Gatan Quantum-K3	Falcon 3EC
Magnification	130kx	130kx	150kx
Pixel size at detector (Å/pixel)	0.64	0.64	0.94
Total electron exposure (e^{-}/Å^2)	50.76	~50-54	47
Exposure rate (e-/pixel/sec)	21.9	21.9	
Number of frames	40	40	40
Defocus range (μm)	-0.8 to -1.8	-0.8 to -1.8	-1.0 to -2.5
Automation software	SerialEM	SerialEM	EPU
Energy filter slit width	20 eV	20 eV	
Micrographs collected (no.)	7,809	39,021	2,230
Micrographs used (no.)	7,691	38,768	
Total extracted particles (no.)	800,711	3,770,881	320,883
For each reconstruction:			
Final particles (no.)	158,286	633,240	63,915
Point-group	C1	C1	C1
Resolution (global, Å)			
FSC 0.5 (unmasked/masked)	(3.5/2.5)	(2.9/2.3)	(8.9/3.9)
FSC 0.143 (unmasked/masked)	(2.8/2.2)	(2.4/2.1)	(4.1/3.3)
Resolution range (local, Å)	2.1-16.1	2.02-3.6	2.9-47.0
Map sharpening *B* factor (Å^2)	-38	-48	-108
(main maps are unsharpened)			
Model composition			
Protein (aa)	1580		
Ligands (no.)	241		
RNA (bases)	45		
Waters (no.)	5		
Model refinement			
Refinement package	REFMAC5		
- real or reciprocal space	Reciprocal		
- resolution cutoff (Å)	2.2		
- Average FSC	0.7871		
B factors (Å^2)			
Protein residues	58		
Ligands	66		
RNA/DNA	95		
Waters	46		
R.m.s. deviations from ideal values			
Bond lengths (Å)	0.0077		
Bond angles (°)	1.7742		
Validation			
MolProbity score	1.47		
CaBLAM outliers	23 (1.5%)		
Clashscore	3.45		
Poor rotamers	26 (1.95%)		
C-beta deviations >0.25Å (%)	2.93%		
EMRinger score	5.51		
Ramachandran plot			
Favored (%)	97.42		
Outliers (%)	0.00		
Res. at FSC(map,model map)=0.5 (Å)	(2.9/2.2)		

Table S2. Summary of modelled proteins and RNA (part 1).

				17S U2 snRNP PDB 7Q3L			**Remodelled U2 snRNP** PDB 7Q4P		
Molecule	Chain ID	Uniprot ID	Total Res.	Modelled Res.	Template Used	Modelling Approach	Modelled Residues	Template Used	Modelling Approach
U2 snRNA	2		188	24-66	-	De novo	7-65	-	De novo
BPS oligo	h		17	-			-		
SF3B1/ Hsh155	A	O75533	1304	491-1304	6EN4	Docked & adjusted	491-1304	17S U2 snRNP	Docked & adjusted
SF3B2/ Cus1	B	Q13435	895	452-598	6Y5Q	Docked & adjusted	452-598	17S U2 snRNP	Docked & adjusted
				701-714	-	De novo	701-714	17S U2 snRNP	Docked & adjusted
SF3B3/ Rse1	C	Q15393	1217	1-1217	6EN4	Docked & adjusted	1-1217	17S U2 snRNP	Docked & adjusted
SF3B5/ Ysf3	E	Q9BWJ5	86	12-81	6EN4	Docked & adjusted	12-81	17S U2 snRNP	Docked & adjusted
SF3B6/ -	F	Q9Y3B4	125	-			-		
PHF5A/ Rds3	G	Q7RTV0	110	8-90	6EN4	Docked & adjusted	8-90	17S U2 snRNP	Docked & adjusted
SF3A2/ Prp11	1	Q15428	464	-			42-85	7ABH	Docked & adjusted
SF3A3/ Prp9	9	Q12874	501	392-493	6Y5Q	Docked & adjusted	392-493	17S U2 snRNP	Docked & adjusted
DDX46/ Prp5	p	Q7L014	1031	-			-		
HTATSF1/ Cus2	q	O43719	755	132-215	-	Predicted & Adjusted	-		
				240-251	-	Predicted & Adjusted			

Table S2. Summary of modelled proteins and RNA (part 2).

				A-like U2 snRNP PDB 7Q4O		
Molecule	Chain ID	Uniprot ID	Total Res.	Modelled Res.	Template Used	Modelling Approach
U2 snRNA	2		188	26-65	-	De novo
BPS oligo	h		17	1-13	-	De novo
SF3B1/ Hsh155	A	O75533	1304	394-415	3LQV	Docked & adjusted
				491-1304	Remodelled U2 snRNP	Docked & adjusted
SF3B2/ Cus1	B	Q13435	895	452-598	Remodelled U2 snRNP	Docked & adjusted
				701-714	Remodelled U2 snRNP	Docked & adjusted
SF3B3/ Rse1	C	Q15393	1217	1-1217	Remodelled U2 snRNP	Docked & adjusted
SF3B5/ Ysf3	E	Q9BWJ5	86	12-81	Remodelled U2 snRNP	Docked & adjusted
SF3B6/ -	F	Q9Y3B4	125	12-101	3LQV	Docked & adjusted
PHF5A/ Rds3	G	Q7RTV0	110	8-90	Remodelled U2 snRNP	Docked & adjusted
SF3A2/ Prp11	1	Q15428	464	42-85	Remodelled U2 snRNP	Docked & adjusted
SF3A3/ Prp9	9	Q12874	501	392-493	Remodelled U2 snRNP	Docked & adjusted
DDX46/ Prp5	p	Q7L014	1031	-		
HTATSF1/ Cus2	q	O43719	755	-		