Supplementary information for

In vitro structural maturation of an early stage pre-40S particle coupled with U3 snoRNA release and central pseudoknot formation

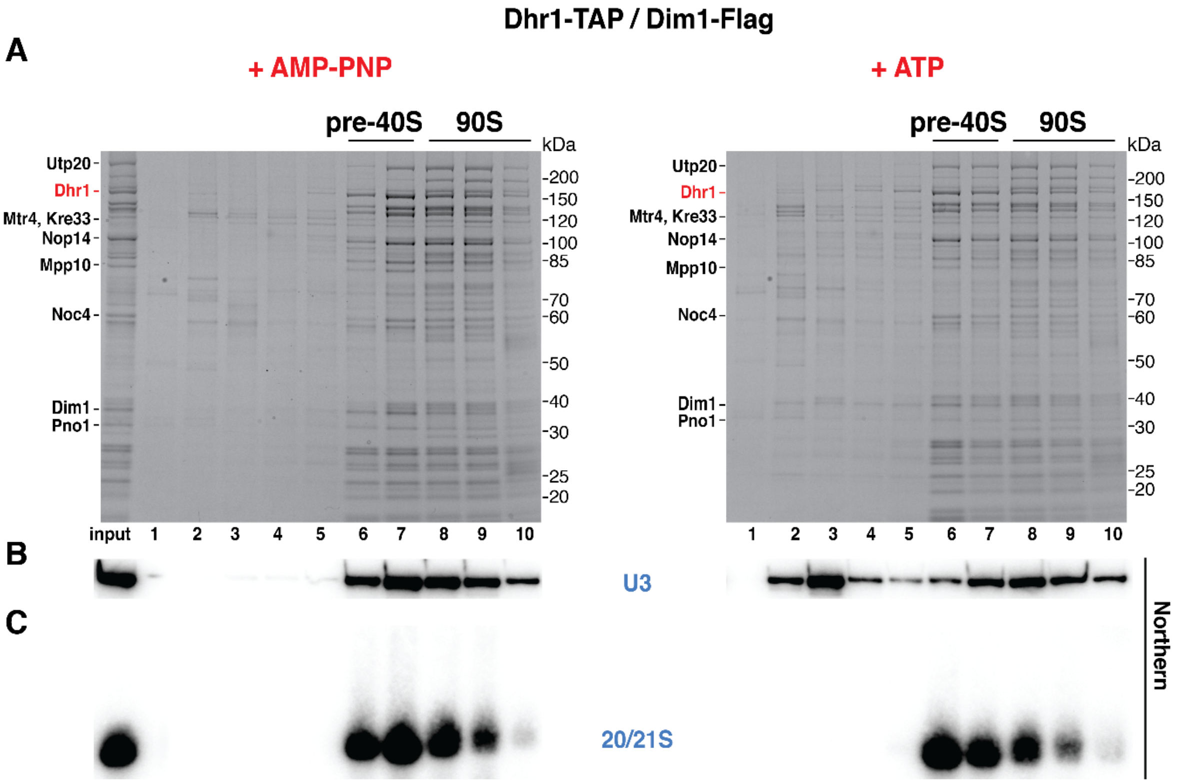
Jingdong Cheng, Giuseppe La Venuta, Benjamin Lau, Otto Berninghausen, Roland Beckmann and Ed Hurt

This supplement contains:

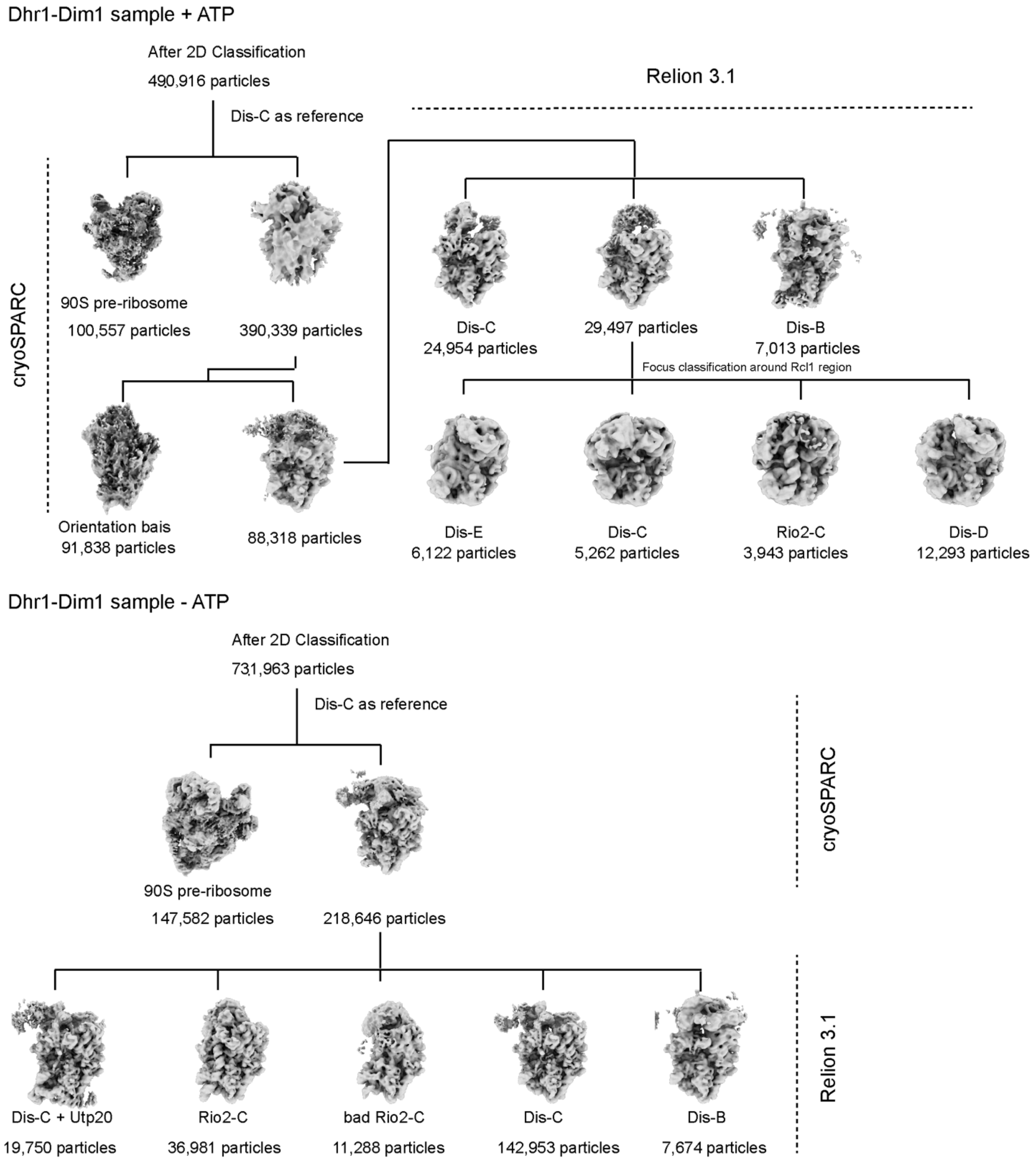
Supplementary Figures S1 to S5

Supplementary Tables S1 to S2

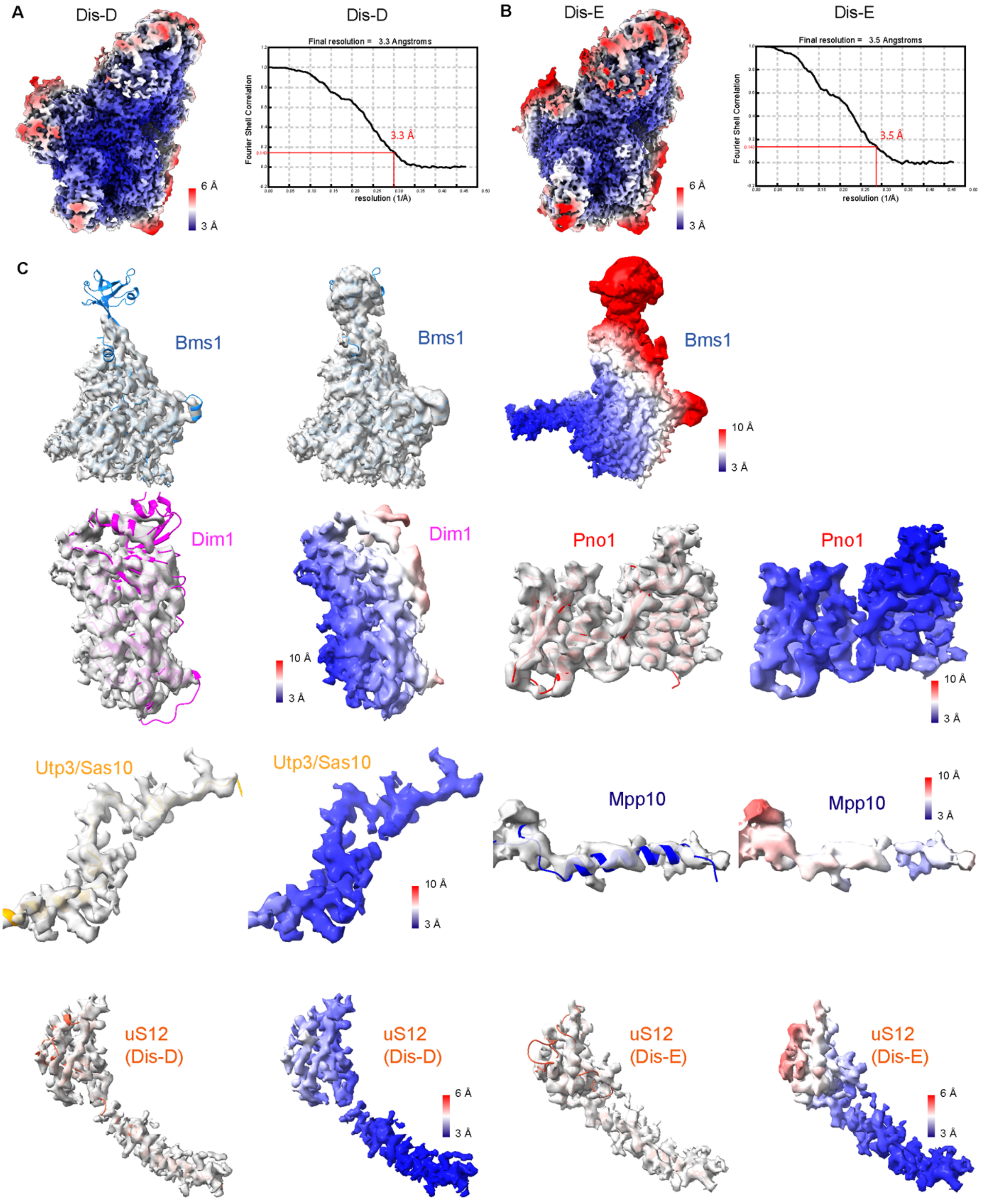
SUPPLEMENTARY FIGURES



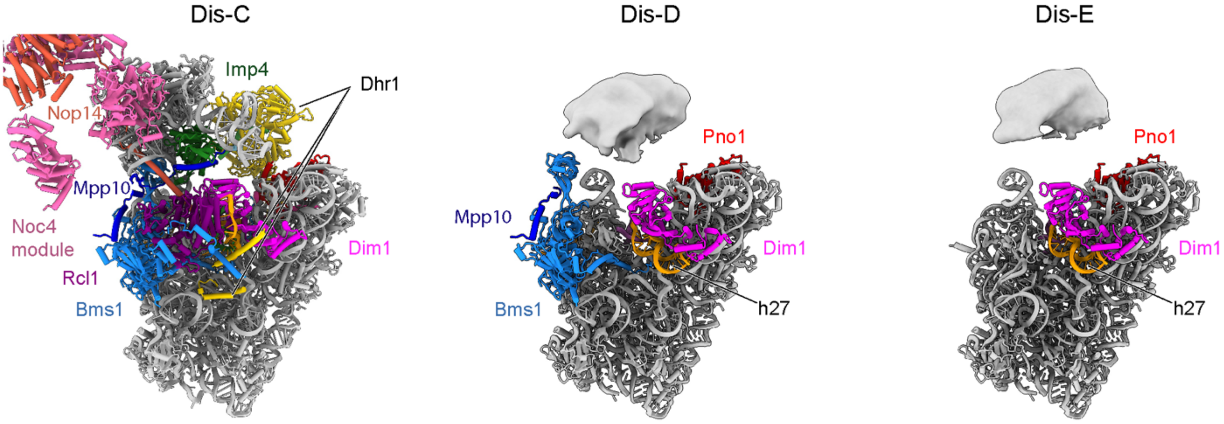
Supplementary Figure S1. ATP-induced U3 snoRNA release from isolated primordial pre-40S particles. Analysis of Dhr1-TAP–Dim-Flag affinity-purified assembly intermediates by sucrose gradient centrifugation, followed by Northern blotting or SDS–polyacrylamide gel electrophoresis and Coomassie staining. Samples were either treated with AMP-PNP (left) or ATP (right). (**A**) SDS-PAGE of sucrose gradient fractions containing the pre-40S and 90S intermediates. The indicated assembly factor bands were identified by mass spectrometry (the band marked with an asterisk was identified as Mtr4). (**B**,**C**) For Northern blotting, DNA probes for the U3 snoRNA (**b**) and the 20S, 21S, 23S pre-rRNA species (**C**) were used. Since it is known that the Dhr1-Dim1 preparation, which was employed in this experiment, contains more post-A_1 than pre-A_1 90S particles[5], the amount of 23S pre-rRNA in the Dhr1-Dim1 eluate (Supplementary Figure S1) is lower compared to the Noc4-Dhr1 preparation (Figure 1).



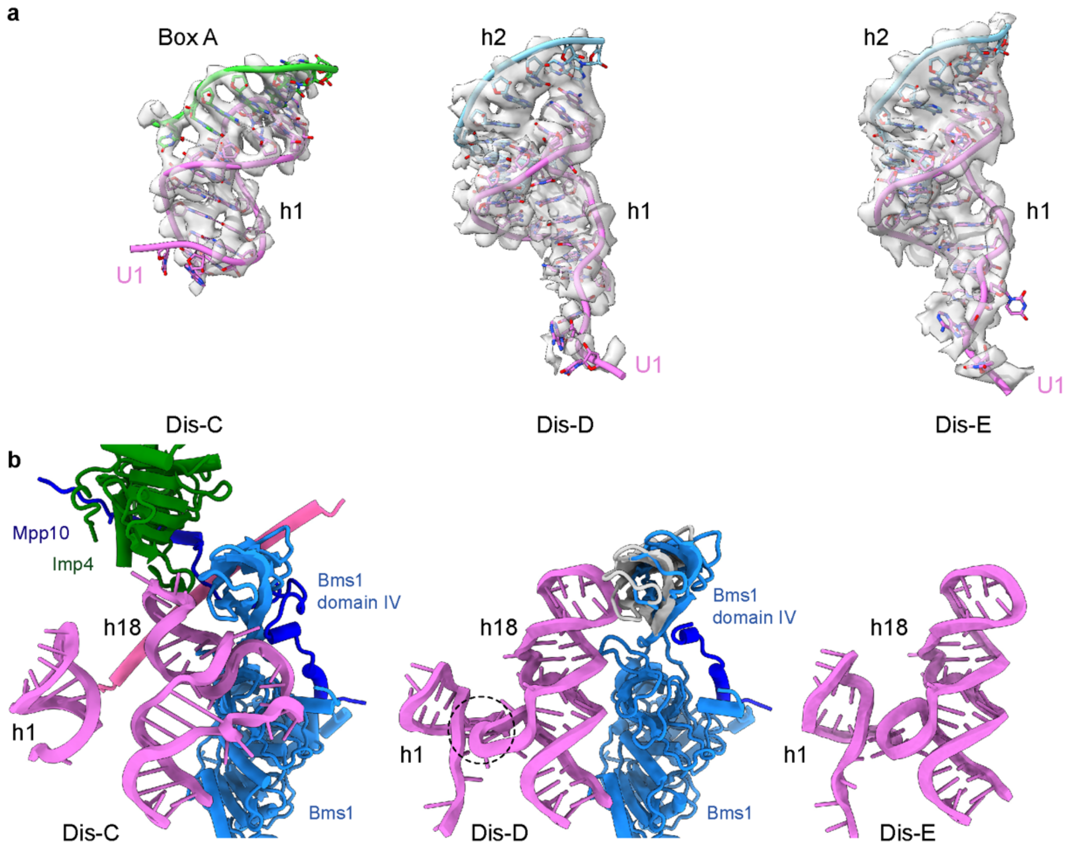
Supplementary Figure S2. The sorting scheme of the yeast Dhr1-Dim1 samples. Classification scheme of the two datasets used in this study, one treated with ATP whereas the other one served as the mock control. Classes that are not shown either could not be further refined to higher resolution or could not be further classified. The corresponding software used for the given steps is indicated.



Supplementary Figure S3. Local resolution estimation and FSC curves. (A,B) The two maps for state Dis-D (**A**) and Dis-E (**B**) were colored according to their local resolution distribution calculated in *Relion*, with their corresponding FSC curves. (**C**) The assembly factors (Bms1, Dim1, Pno1, Sas10 and Mpp10) and uS12 in this study were shown with density maps and colored according to their local resolution distribution.



Supplementary Figure S4. The molecular models for the states Dis-C, Dis-D and Dis-E. The 40S Head in state Dis-D and Dis-E are only shown as density maps.



Supplementary Figure S5. Comparison of the rRNA helix h2 and h18 regions. (**A**) The different conformations of the 18S rRNA helix h2 region in state Dis-C (left), Dis-D (middle) and Dis-E (right) are shown as molecular models together with the transparent density maps. (**B**) Different conformations of the h1 and h18 of the 18S rRNA in state Dis-C (left), Dis-D (middle) and Dis-E (right). Interaction between helices h1 and h18 is indicated by a black circle, and domain IV of Bms1 from state Dis-C is indicated in gray.

SUPPLEMENTARY TABLES

Supplementary Table S1. Cryo-EM data collection, refinement and validation statistics.

	Dis-D	Dis-E
Data collection and processing		
Magnification	129,151	
Voltage (kV)	300	
Electron exposure (e–/Å^2)	44	
Defocus range (µm)	-0.8 to -2.5	
Pixel size (Å)	1.047	
Symmetry imposed	*C1*	
Initial particle images (no.)	552,020	
Final particle images (no.)	12,293	6,122
Map resolution (Å)	3.3	3.5
FSC threshold	0.143	0.143
Refinement		
Initial model used (PDB code)	6ZQG	
Model resolution (Å)	3.3	3.4
FSC threshold	0.5	0.5
Map sharpening *B* factor (Å^2)	-55	-55
Model composition		
Non-hydrogen atoms	49,811	44,623
Protein residues	3,298	2,651
RNA	1,102	1,106
Ligands	3	1
B factors (Å^2)	35.77	38.12
Protein	33.43	34.17
RNA	38.37	41.65
Ligand	52.59	67.23
R.m.s. deviations		
Bond lengths (Å)	0.0033	0.0060
Bond angles (°)	0.8	0.89
Validation		
MolProbity score	1.76	1.83
Clashscore	6.23	6.39
Poor rotamers (%)	0.04	0
Ramachandran plot		
Favored (%)	93.73	92.36
Allowed (%)	6.21	7.56
Disallowed (%)	0.06	0.08
PDB entry	7WTL	7WTM
EMDB entry	32790	32791

Supplementary Table S2. Molecular models of assembly factors.

Assembly factors	Full length (aa)	Modeled regions (aa)	Allowed *de novo* modelling?	Present in states
Dhr1	Not present in Dis-D and Dis-E			
Rcl1	Not present in Dis-D and Dis-E			
Bms1	1183	25-322, 349-361, 747-1045	*Yes, 25-925, from PDB:6ZQG*	Dis-D
			No, 926-1045, from PDB:6ZQG	
Mpp10	593	295-330	*No, from PDB:6ZQG*	Dis-D
Utp3	610	535-581	*Yes, from PDB:6ZQG*	Dis-D
Dim1	318	3-318	*No, from PDB:6ZQG*	Dis-D, Dis-E
Pno1	274	93-273	*Yes, from PDB:6ZQG*	Dis-D, Dis-E
uS12	145	2-145	*Yes,*	Dis-D
			Yes, except some loop regions	Dis-E