nature portfolio

https://doi.org/10.1038/s41586-021-03554-8

Supplementary information

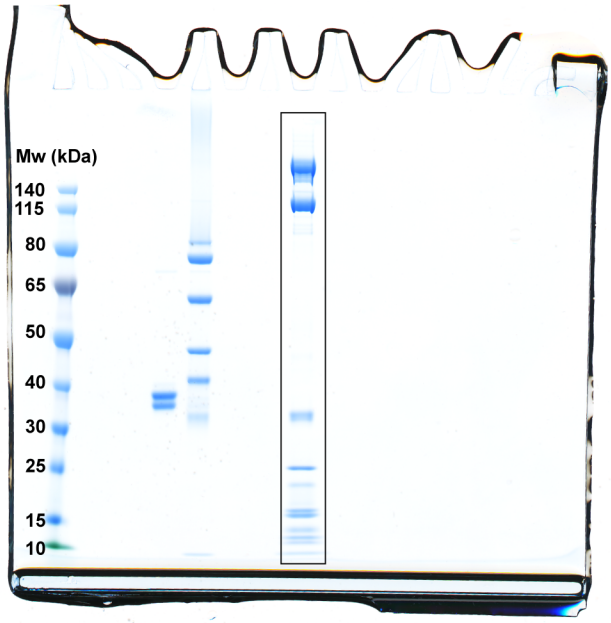
Structures of mammalian RNA polymerase II pre-initiation complexes

In the format provided by the authors and unedited

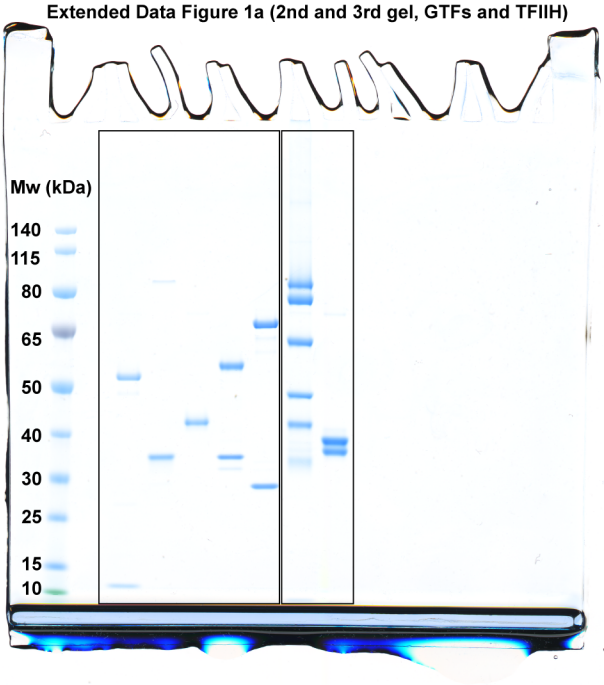
Data collection and pre-processing										
	No ADP·BeF_3	with ADP·BeF_3 (1)	with ADP·BeF_3 (2)							
Magnification	81,000x	81,000x	81,000x							
Voltage (kV)	300	300	300							
Electron exposure (e–/Å^2)	41.1	44.8	41.6							
Defocus range (μm)	0.2-3.0	0.2-3.0	0.2-3.0							
Pixel size (Å)	1.05	1.05	1.05							
Initial particle images (no.)	1,535,552	6,106,191	5,277,512							
Model/Map Name	Proximal cPIC	Distal cPIC	TFIIH (No ADP·BeF3)	OC cPIC	TFIIH (+ADP·BeF_3)	Proximal CC	Distal CC	OC	IC XPB	IC
EMDB Code	12611 (Proximal cPIC) 12620 (Consensus cPIC) 12622 (Consensus UC)	12612 (Distal cPIC)	12615 (Overall) 12624 (XPB focus) 12625 (XPD focus)	12613 (OC, cPIC) 12621 (Consensus cPIC) 12623 (Consensus UC)	12616 (Overall) 12626 (XPD focus)	12617	12618	12619	12614 (XPB focus)	12630 12629 (IC, cPIC) 12627 (Consensus cPIC) 12628 (Consensus UC)
PDB code	7NVS	7NVT	7NVW	7NVU	7NVX	7NVY	7NVZ	7NW0	7NVV	
Symmetry imposed	C1	C1	C1	C1	C1	C1	C1	C1	C1	C1
Final particle images (no.)	63,190 271,685 (Consensus)	50,177	129,156	255,110 714,395 (Consensus)	399,247	15,226	11,412	26,146	714,395	13,380 (IC) 109,841 (cPIC) 711,215 (Consensus)
Map resolution (Å) FSC threshold = 0.143	2.8 (Proximal cPIC) 2.4 (Consensus cPIC) 2.8 (Consensus UC)	2.9 (Distal CC, cPIC)	4.3 (Overall) 3.3 (XPB focus) 4.1 (XPD focus)	2.5 (OC, cPIC) 2.4 (Consensus cPIC) 2.7(Consensus UC)	3.9 (Overall) 2.9 (XPB focus) 4.0 (XPD focus)	7.3	7.2	6.6	2.9	9.1 (IC) 2.8 (IC, cPIC) 2.4 (Consensus cPIC) 2.7 (Consensus UC)
Map resolution range (Å)	2.5-6.5 (Proximal cPIC) 2.3-5.5 (Consensus cPIC) 2.5-6.5 (Consensus UC)	2.5-6.5 (Distal CC, cPIC)	4.0-6.4 (Overall) 3.2-5.6 (XPB focus) 3.9-6.3 (XPD focus)	2.2-5.4 (OC, cPIC) 2.2-5.1 (Consensus cPIC) 2.5-6.5(Consensus UC)	3.8-6.8 (Overall) 2.8-6.8 (XPB focus) 3.7-8.5 (XPD focus)	4.3-13.9	4.4-20.4	4.2-18.2	2.8-6.8	7.0-23.0
Refinement										
Initial model used (PDB code)	5FLM 5IY7 5GPY	5FLM 5IY7 5GPY	6NMI 6RO4	5FLM 5IY7 5GPY	6NMI 6RO4	7NVS 7NVW	7NVT 7NVW	7NVU 7NVX		
Model resolution (Å) FSC threshold = 0.5	2.8	2.8	7.5	2.5	7.3	9.1	9.7	9.2	3.1	
Map sharpening *B* factor (Å^2)	-16.9	-15.5	-133	-22	-105	-140	-104	-166		
Model composition										
Non-hydrogen atoms	44073	44114	23647	44240	23663	67474	67474	66935	9071	
Protein residues	5249	5249	2814	5295	2796	8063	8063	7986	975	
Nucleotides	98	100	42	90	50	128	128	134	50	
Ligands	ZN: 10 MG: 1	ZN: 10 MG: 1	ZN: 7 SF4: 1 UNK: 27	ZN: 10 MG: 1	ZN: 7 MG: 1 ADP:1 BEF:1 SF4: 1 UNK: 24	ZN: 17 MG: 1 SF4: 1 UNK: 27	ZN: 17 MG: 1 SF4: 1 UNK: 27	ZN: 15 MG: 2 ADP: 1 BEF: 1 SF4: 1 UNK: 24	MG: 1 ADP:1 BEF:1 UNK: 24	
Mean *B* factors (Å^2)										
Protein	24.69	25.07	57.35	52.65	43.28	41.31	59.62	65.61	41.66	
Nucleotides	62.46	64.86	97.98	118.51	74.26	49.64	83.10	103.80	87.36	
Ligand	40.35	39.98	84.01	74.72	55.74	72.31	120.17	84.24	41.12	
R.m.s. deviations										
Bond lengths (Å)	0.003	0.003	0.002	0.003	0.002	0.002	0.002	0.002	0.003	
Bond angles (°)	0.453	0.460	0.452	0.449	0.455	0.49	0.516	0.420	0.830	
Validation										
MolProbity score	1.31	1.40	1.63	1.28	1.68	1.53	1.56	1.39	1.31	
Clashscore	4.39	4.85	7.26	3.84	6.57	6.49	6.57	4.97	3.91	
Poor rotamers (%)	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	
Ramachandran plot										
Favored (%)	97.53	97.27	96.46	97.44	95.49	97.05	96.76	97.30	97.29	
Allowed (%)	2.47	2.83	3.54	2.56	4.47	2.95	3.24	2.69	2.71	
Disallowed (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	

Supplementary Table 1 | Cryo-EM data collection and processing information

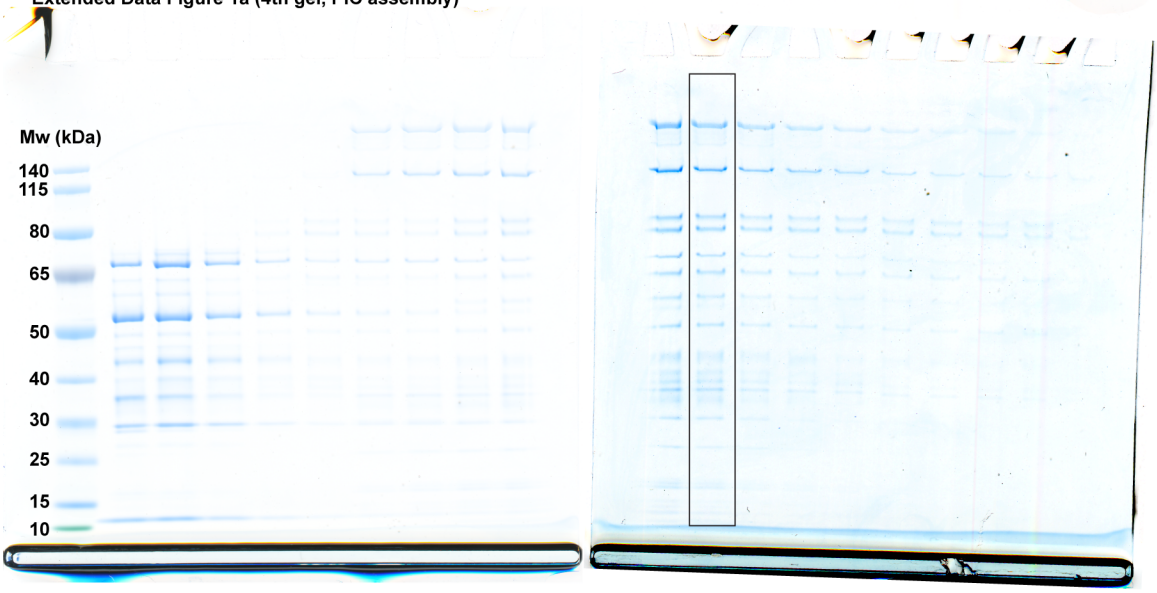
Extended Data Figure 1a (left most gel, *S. scrofa* Pol II)



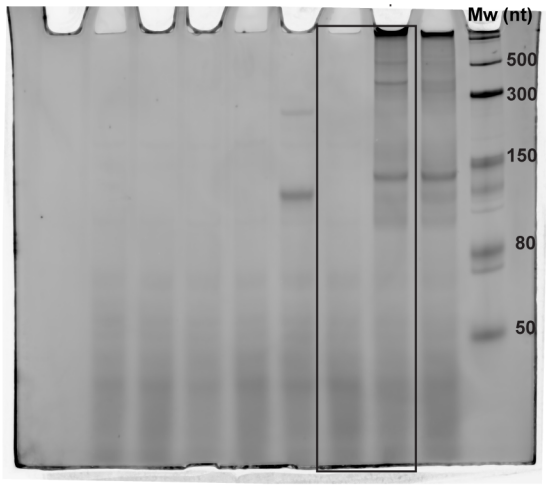
Extended Data Figure 1a (2nd and 3rd gel, GTFs and TFIIH)



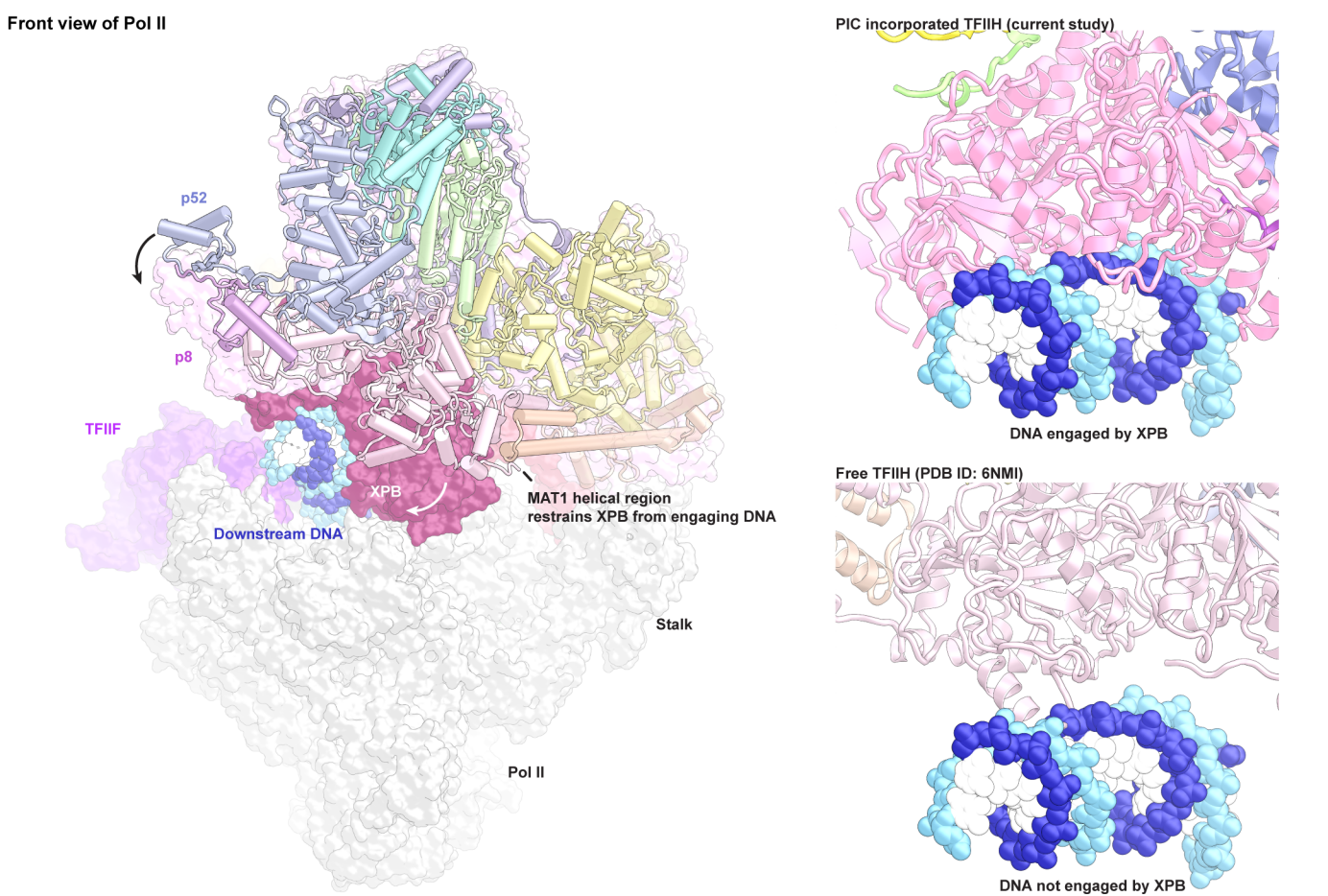
Extended Data Figure 1a (4th gel, PIC assembly)



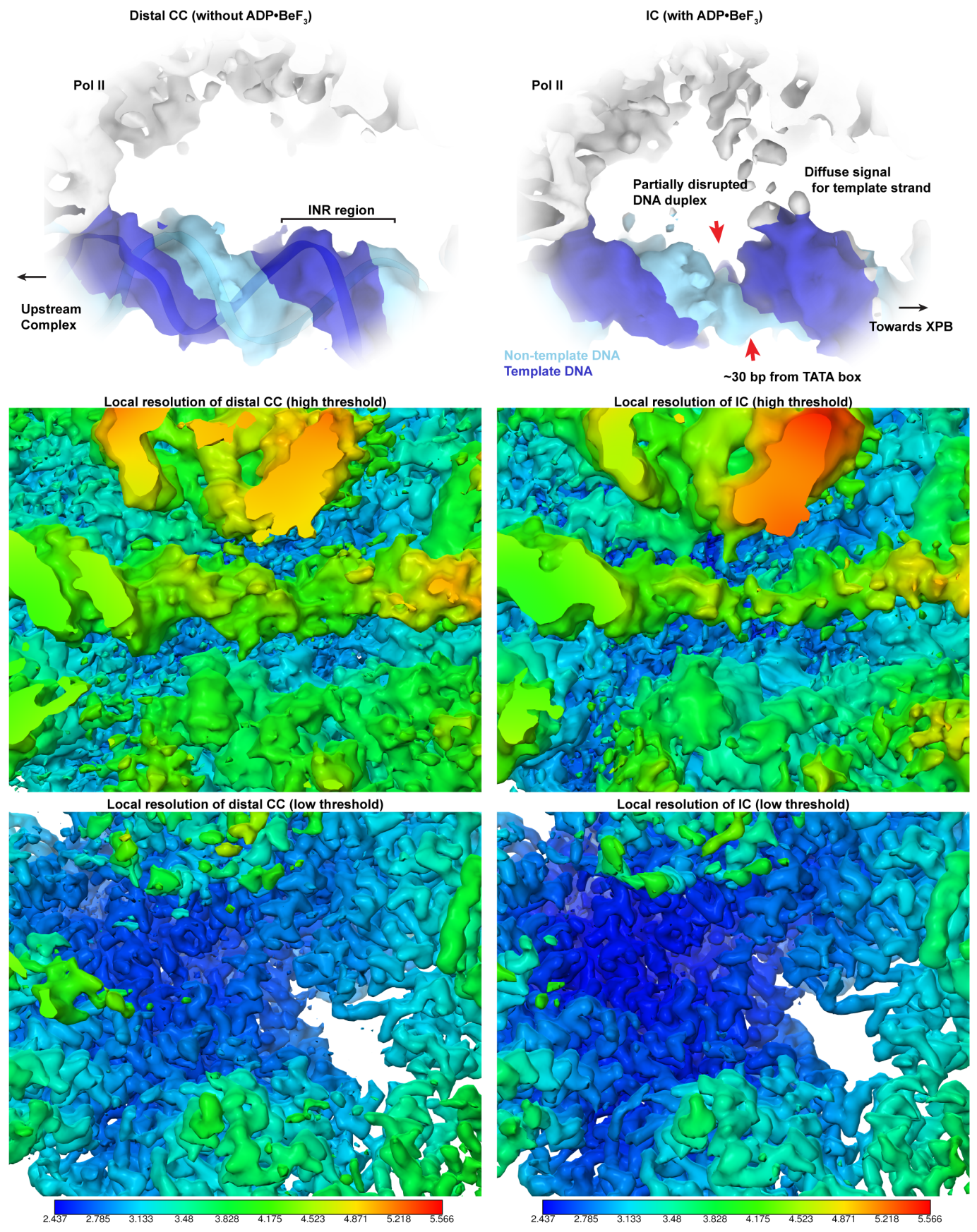
Extended Data Figure 1e (right most gel, PIC transcription assay)



Supplementary Figure 1 | Raw, uncropped images of gels
Uncropped images of all gels from SDS-PAGE analysis presented in this study. Molecular weight markers are indicated and the boxed area represent where the gel was cropped for presentation in the manuscript.

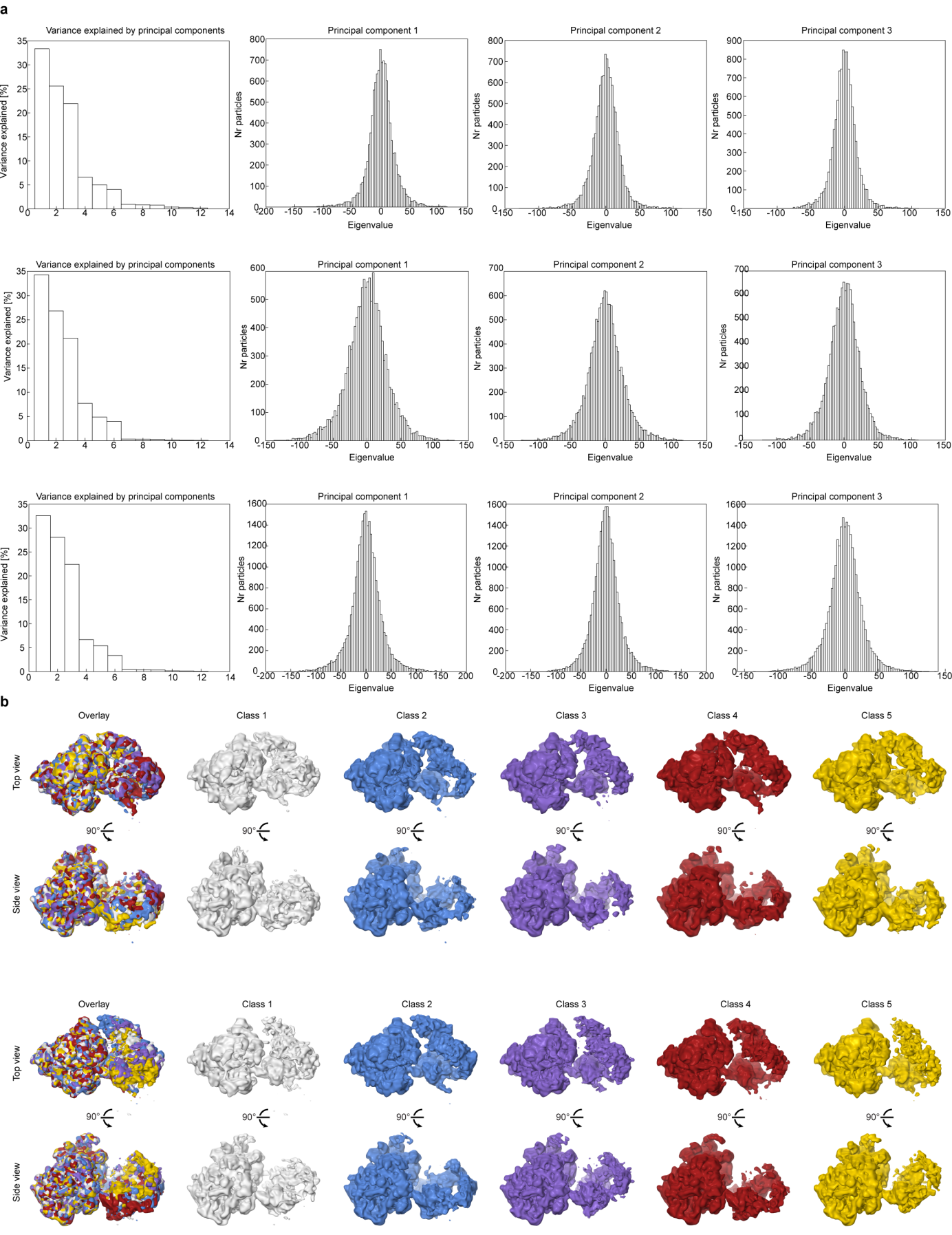


Supplementary Figure 2 | Further analysis of free-TFIIH compared to the PIC incorporated form
Superposition of the free-TFIIH over the PIC shows that our XPB domain is placed further away than in the conformation observed in the apo-TFIIH
For the XPB to make the conformational change required to engage the DNA in its current position, the MAT1 helical region must let go of the DRD domain.



Supplementary Figure 3 | Local resolution analysis of the distal CC and IC DNA binding cleft

Local resolution analysis of the DNA binding cleft of Pol II show a comparable distribution of local resolution at high and low thresholds, indicating that the differences in density intensity around the DNA are not due to resolution loss.



Supplementary Figure 4 | Multibody analysis and further 3D classification of cPIC-TFIIH

a. Multibody analysis shows that motion present between the cPIC and TFIIH are Gaussian distributions and no notable alternative conformational minima are present.

b. Further 3D classification of IC (top) and OC (bottom) classes show a roughly equal distribution of classes with varying positions of TFIIH with respect to the cPIC.