Molecular Cell, Volume *83*

Supplemental information

Structural basis of transcription reduction by a promoter-proximal +1 nucleosome

Julio Abril-Garrido, Christian Dienemann, Frauke Grabbe, Taras Velychko, Michael Lidschreiber, Haibo Wang, and Patrick Cramer

Supplemental information

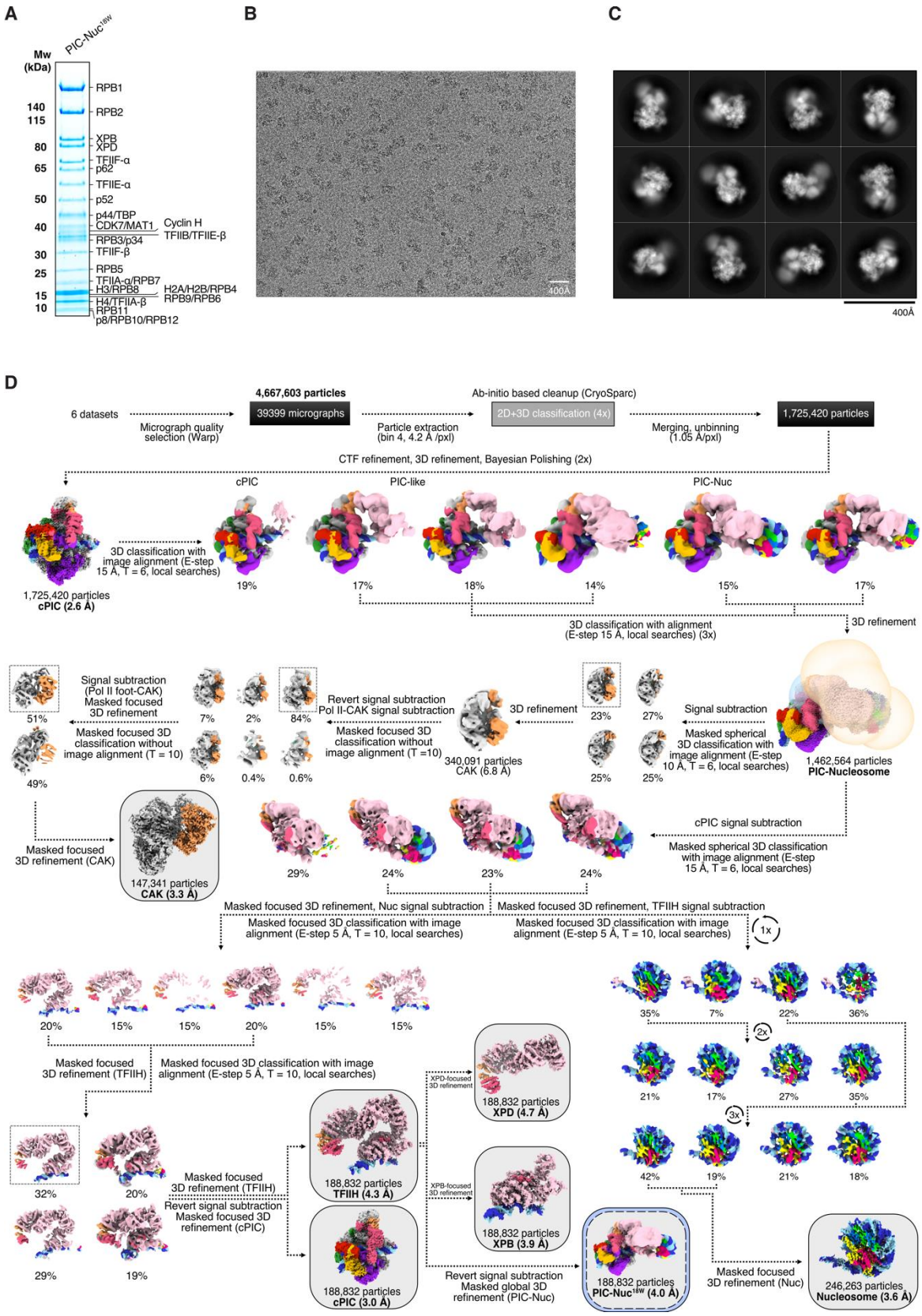


Figure S1. Sample preparation and cryo-EM processing analysis of PIC-Nucleosome^{18W}, Related to Figure 2.

(A) Sucrose gradient fraction used for cryo-EM sample preparation (analyzed by SDS-PAGE). Molecular weight is described on the left side of the gel and assembly components on the right side.

(B) Exemplary cryo-EM micrograph. A scale bar is provided at the bottom right of the figure.

(C) Representative 2D class averages depicting mammalian PIC-nucleosome^{18W}. Scale bar is placed at the bottom right of the figure.

(D) Processing classification strategy employed to sort out particles and for structure determination of PIC-Nuc^{18W}. Each step of the tree is described on the diagram and the reported resolutions are shown under the respective cryo-EM maps. Final maps used for model building are located within grey boxes, whereas consensus refinements are enclosed in grey-blue boxes.

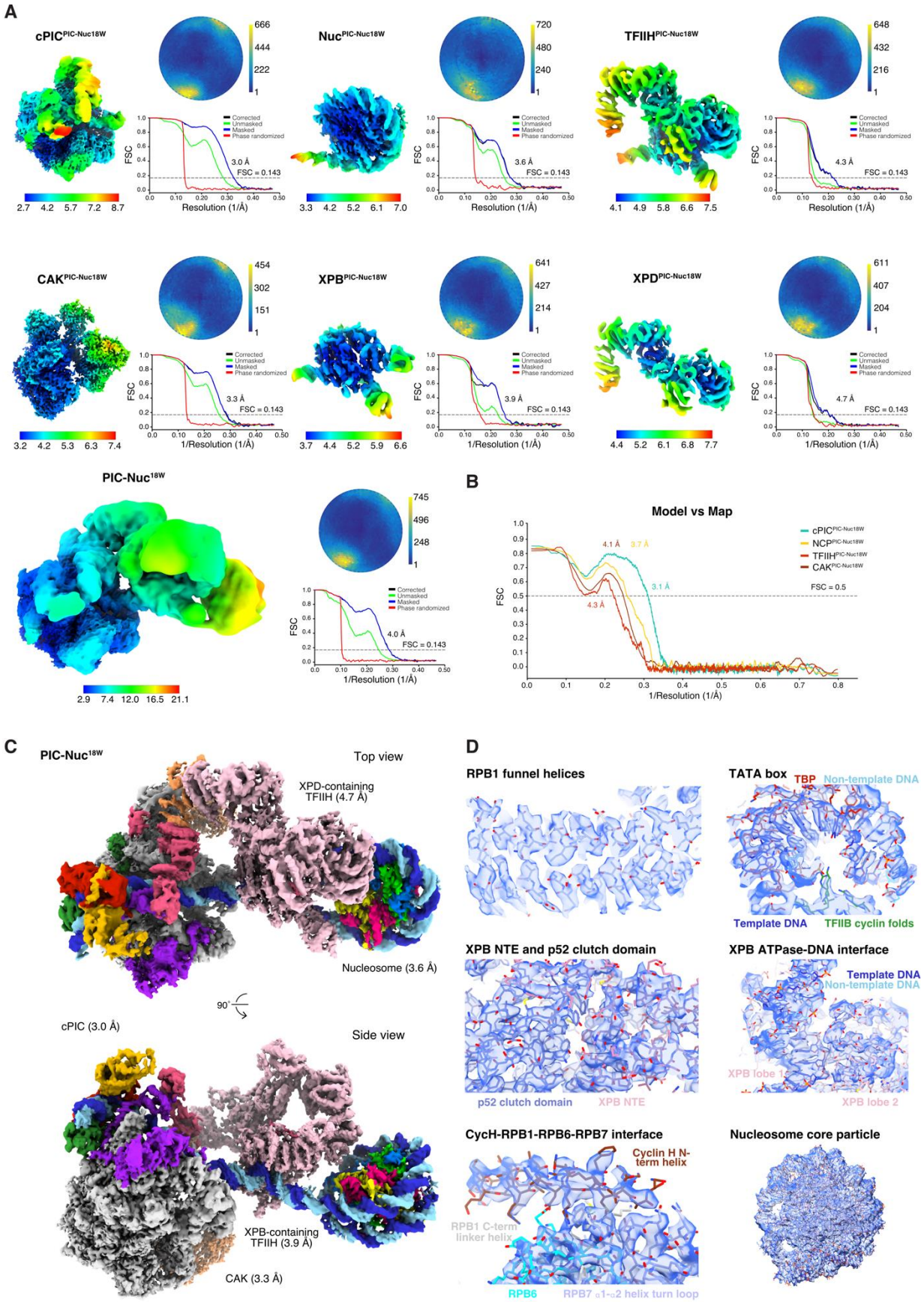

A
cPIC PIC-Nuc18W
Nuc PIC-Nuc18W
TFIIH PIC-Nuc18W
CAK PIC-Nuc18W
XPB PIC-Nuc18W
XPD PIC-Nuc18W
PIC-Nuc18W
Corrected
Unmasked
Masked
Phase randomized
FSC = 0.143
Resolution (1/Å)
1/Resolution (1/Å)
B
Model vs Map
cPIC PIC-Nuc18W
NCP PIC-Nuc18W
TFIIH PIC-Nuc18W
CAK PIC-Nuc18W
FSC = 0.5
C
PIC-Nuc18W
Top view
XPD-containing TFIIH (4.7 Å)
Nucleosome (3.6 Å)
cPIC (3.0 Å)
Side view
XPB-containing TFIIH (3.9 Å)
CAK (3.3 Å)
D
RPB1 funnel helices
TATA box
TBP
Non-template DNA
Template DNA
TFIIB cyclin folds
XPB NTE and p52 clutch domain
p52 clutch domain
XPB NTE
XPB ATPase-DNA interface
Template DNA
Non-template DNA
XPB lobe 1
XPB lobe 2
CycH-RPB1-RPB6-RPB7 interface
Cyclin H N-term helix
RPB1 C-term linker helix
RPB6
RPB7 α1-α2 helix turn loop
Nucleosome core particle

Figure S2. Map quality assessment of the PIC-nucleosome^{18W} structure, Related to Figure 2.

(A) Focused refinements and overall maps filtered and colored by estimated local resolution, shown as top views. The superscript typography of each map is used to differentiate it from the ones shown in **Figure S5**. Their angular distributions and resolution (following the FSC 0.143 cut-off criterion) are provided on the right side of each map.

(B) Model-map correlation between the individual focused refined maps, used for model building, and the respective built atomic models. Resolution following the FSC 0.5 cut-off criterion is marked with a dashed line.

(C) Composite maps showing individual local refinements in distinct views for PIC-Nuc^{18W}. The different protein and DNA components are colored according to the colors used in their structural models, depicted in **Figure 2B**.

(D) Cryo-EM density maps (shown in blue surface) fitted to the corresponding region of the modelled structures.

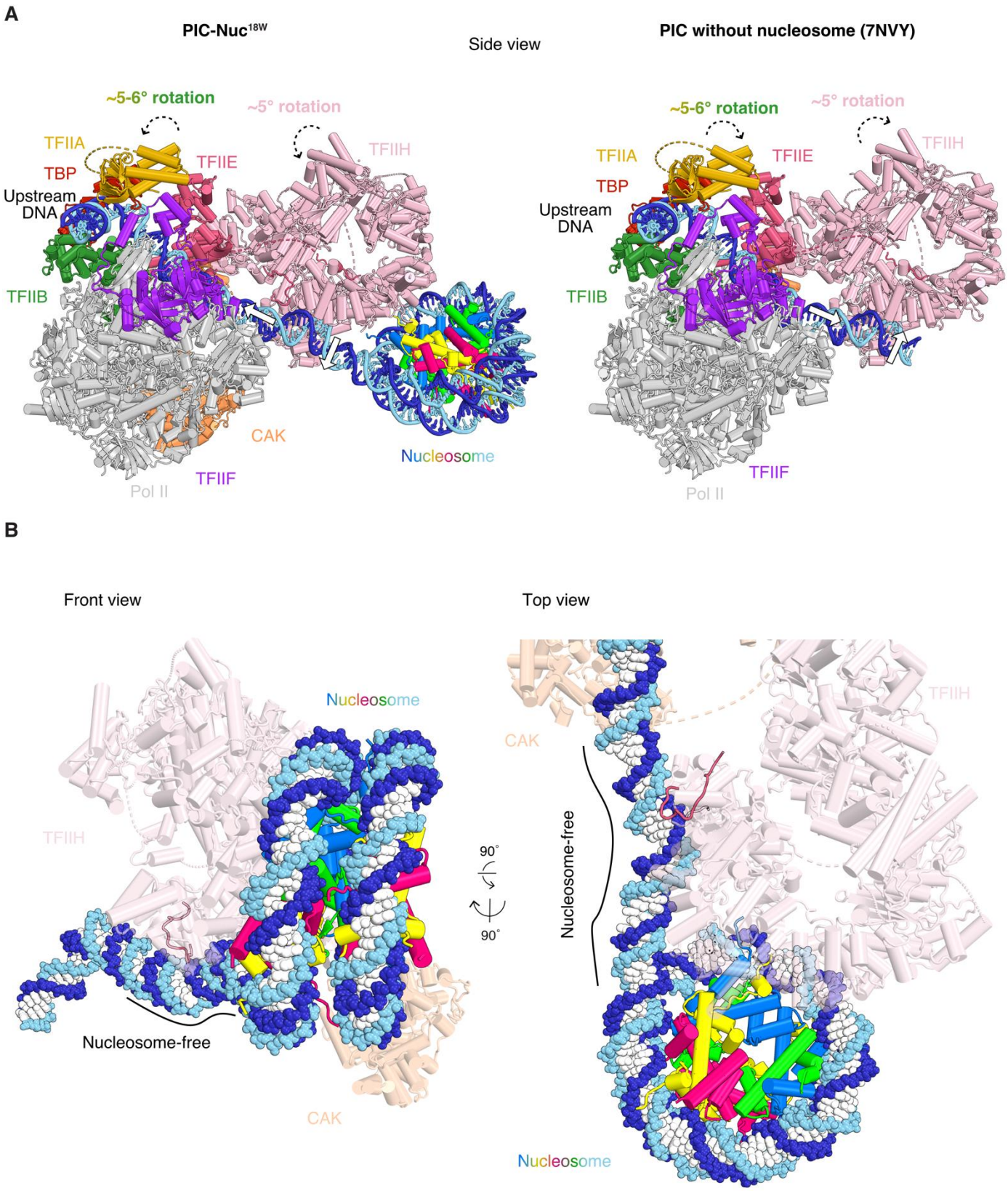

A
PIC-Nuc18W
Side view
PIC without nucleosome (7NVY)
~5-6° rotation
~5° rotation
TFIIA
TBP
Upstream DNA
TFIIE
TFIIH
TFIIB
CAK
Nucleosome
TFIIF
Pol II
B
Front view
Top view
Nucleosome
TFIIH
Nucleosome-free
CAK
90˚
90˚
Nucleosome

Figure S3. Downstream DNA is free for TFIIH binding in PIC-Nuc^{18W} and mammalian PIC without nucleosome, Related to Figure 3.

(A) Comparison of PIC-Nuc^{18W} (left panel) with PIC without nucleosome (right panel) (PDB: 7NVY[1]) in side view. PIC is found in its canonical conformation, and only minor movements in the upstream complex were observed. The downstream DNA region is slightly bent when compared to the structure of the nucleosome-free PIC (white arrow). Arrowheads indicate the local motion of different PIC subunits.

(B) Front view and top close-up view of TFIIH-Nuc^{18W}. TFIIH binds downstream DNA when the nucleosome is located at a distal position to the promoter region. Nucleosomes are shown in sphere representation and TFIIH in transparent cartoon representation. Solid lines denote nucleosome-free DNA.

Superimposition of models was performed by aligning on Pol II. Color of the different components is denoted adjacent to them.

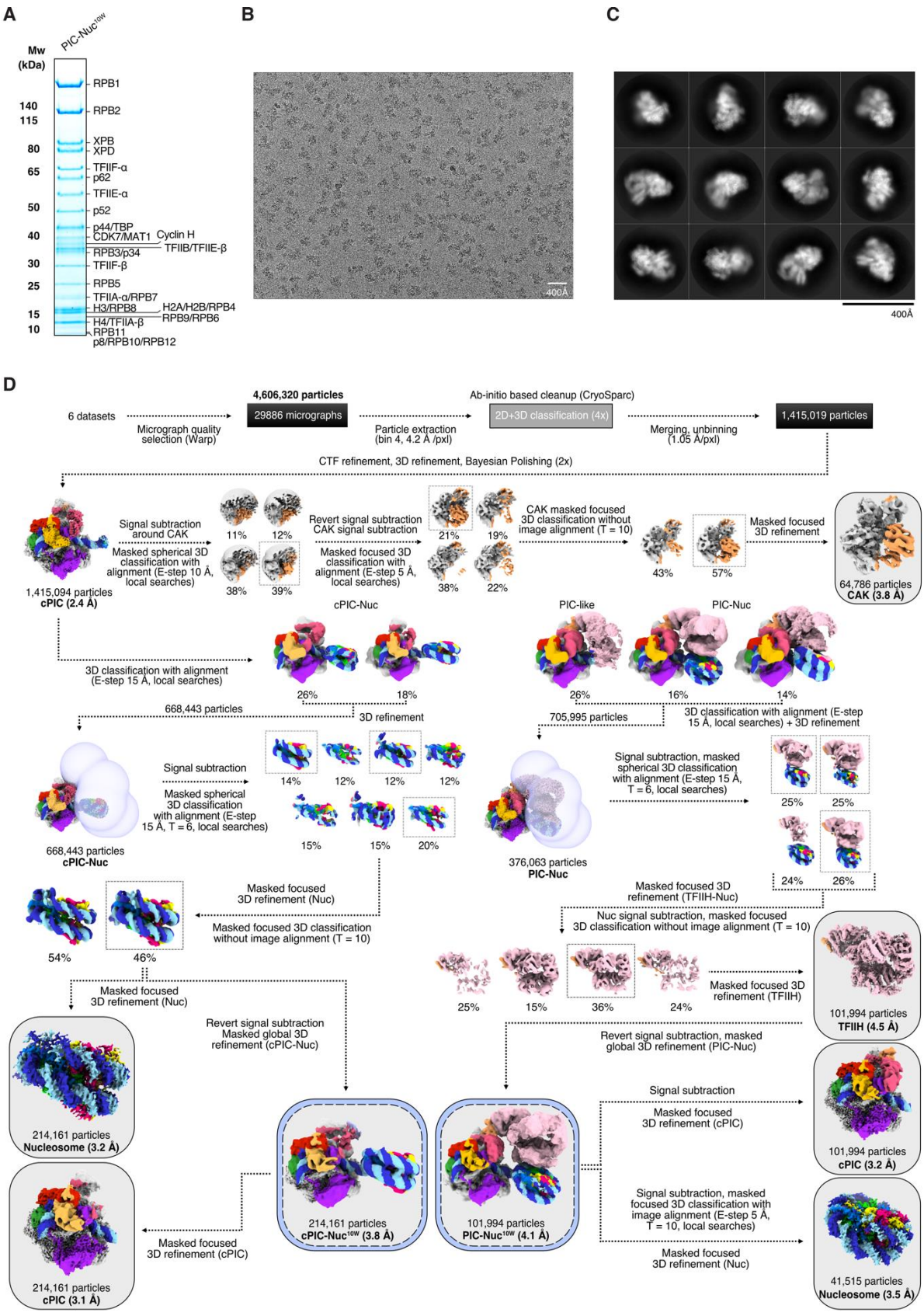

A
PIC-Nuc10W
Mw (kDa)
140
115
80
65
50
40
30
25
15
10
RPB1
RPB2
XPB
XPD
TFIIF-α
p62
TFIIE-α
p52
p44/TBP
CDK7/MAT1
Cyclin H
TFIIB/TFIIE-β
RPB3/p34
TFIIF-β
RPB5
TFIIA-α/RPB7
H3/RPB8
H2A/H2B/RPB4
RPB9/RPB6
H4/TFIIA-β
RPB11
p8/RPB10/RPB12
B
400Å
C
400Å
D
4,606,320 particles
6 datasets
Micrograph quality selection (Warp)
29886 micrographs
Particle extraction (bin 4, 4.2 Å /pxl)
Ab-initio based cleanup (CryoSparc)
2D+3D classification (4x)
Merging, unbinning (1.05 Å/pxl)
1,415,019 particles
CTF refinement, 3D refinement, Bayesian Polishing (2x)
Signal subtraction around CAK
Masked spherical 3D classification with alignment (E-step 10 Å, local searches)
11%
12%
38%
39%
1,415,094 particles
cPIC (2.4 Å)
Revert signal subtraction CAK signal subtraction
Masked focused 3D classification with alignment (E-step 5 Å, local searches)
21%
19%
38%
22%
CAK masked focused 3D classification without image alignment (T = 10)
43%
57%
Masked focused 3D refinement
64,786 particles
CAK (3.8 Å)
3D classification with alignment (E-step 15 Å, local searches)
cPIC-Nuc
26%
18%
668,443 particles
3D refinement
PIC-like
PIC-Nuc
26%
16%
14%
705,995 particles
3D classification with alignment (E-step 15 Å, local searches) + 3D refinement
Signal subtraction
Masked spherical 3D classification with alignment (E-step 15 Å, T = 6, local searches)
14%
12%
12%
12%
15%
15%
20%
668,443 particles
cPIC-Nuc
376,063 particles
PIC-Nuc
Signal subtraction, masked spherical 3D classification with alignment (E-step 15 Å, T = 6, local searches)
25%
25%
24%
26%
Masked focused 3D refinement (Nuc)
Masked focused 3D classification without image alignment (T = 10)
54%
46%
Masked focused 3D refinement (TFIIH-Nuc)
Nuc signal subtraction, masked focused 3D classification without image alignment (T = 10)
25%
15%
36%
24%
Masked focused 3D refinement (TFIIH)
101,994 particles
TFIIH (4.5 Å)
Masked focused 3D refinement (Nuc)
Revert signal subtraction Masked global 3D refinement (cPIC-Nuc)
Revert signal subtraction, masked global 3D refinement (PIC-Nuc)
214,161 particles
Nucleosome (3.2 Å)
Signal subtraction
Masked focused 3D refinement (cPIC)
101,994 particles
cPIC (3.2 Å)
214,161 particles
cPIC-Nuc10W (3.8 Å)
101,994 particles
PIC-Nuc10W (4.1 Å)
Signal subtraction, masked focused 3D classification with image alignment (E-step 5 Å, T = 10, local searches)
Masked focused 3D refinement (Nuc)
Masked focused 3D refinement (cPIC)
214,161 particles
cPIC (3.1 Å)
41,515 particles
Nucleosome (3.5 Å)

Figure S4. Sample preparation and cryo-EM processing analysis of PIC-nucleosome^{10W}, Related to Figure 2.

(A) Sucrose gradient fraction used for cryo-EM sample preparation (analyzed by SDS-PAGE). Molecular weight is described on the left side of the gel and assembly components on the right side.

(B) Exemplary cryo-EM micrograph. A scale bar is provided at the bottom right of the figure.

(C) Representative 2D class averages depicting mammalian PIC-nucleosome^{10W}. Scale bar is placed at the bottom right of the figure.

(D) Processing classification strategy employed to sort out particles and for structure determination of PIC-Nuc^{10W}. Each step of the tree is described on the diagram and the reported resolutions are shown under the respective cryo-EM maps. Final maps used for model building are located within grey boxes, whereas consensus refinements are enclosed in grey-blue boxes.

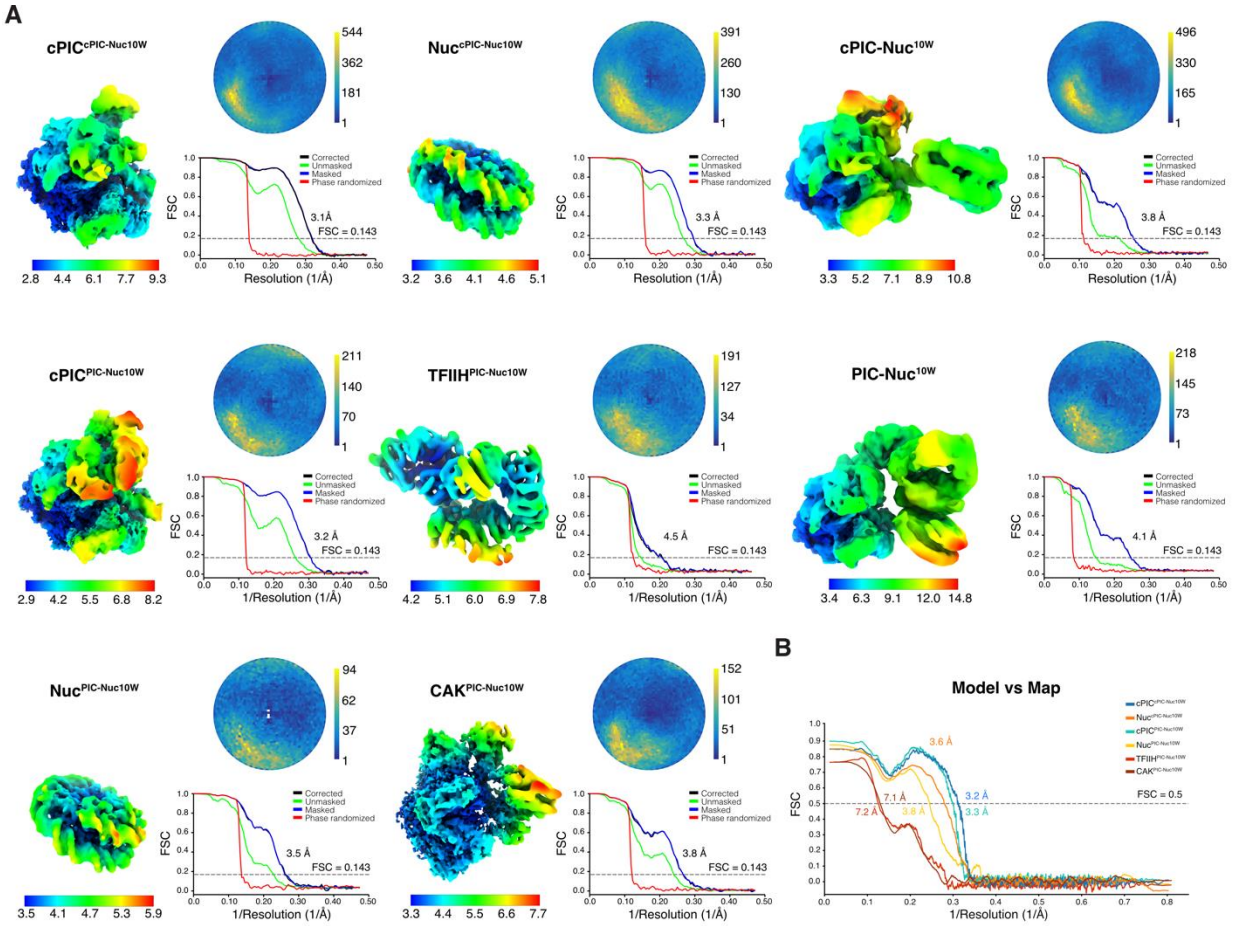

A
cPIC^cPIC-Nuc10W
Nuc^cPIC-Nuc10W
cPIC-Nuc^10W
cPIC^PIC-Nuc10W
TFIIH^PIC-Nuc10W
PIC-Nuc^10W
Nuc^PIC-Nuc10W
CAK^PIC-Nuc10W
Corrected
Unmasked
Masked
Phase randomized
FSC = 0.143
Resolution (1/Å)
1/Resolution (1/Å)
B
Model vs Map
FSC = 0.5
C
Top view
Side view
cPIC-Nuc^10W
PIC-Nuc^10W
cPIC (3.1 Å)
Nuc (3.2 Å)
TFIIH (4.5 Å)
cPIC (3.2 Å)
Nuc (3.5 Å)
CAK (3.8 Å)
90°
D
MAT1 helical region
Rpb3-11 dimerization domains
Upstream DNA
Nucleosome core particle
XPB DRD domain
MAT1 helical region
RPB3
RPB11
Non-template DNA
Template DNA
TBP
TFIIF winged helix
TFIIB cyclin folds
XPD ARCH and FeS domains
ARCH domain
FeS domain
MAT1 3-helix bundle
MAT1 RING domain

Figure S5. Map quality assessment of the PIC-nucleosome^{10W} structure, Related to Figure 2.

(A) Focused refinements and overall maps filtered and colored by estimated local resolution, shown as top views. The superscript typography of each map indicates whether those particles contained TFIIH. Their angular distributions and resolution (following the FSC 0.143 cut-off criterion) are provided on the right side of each map.

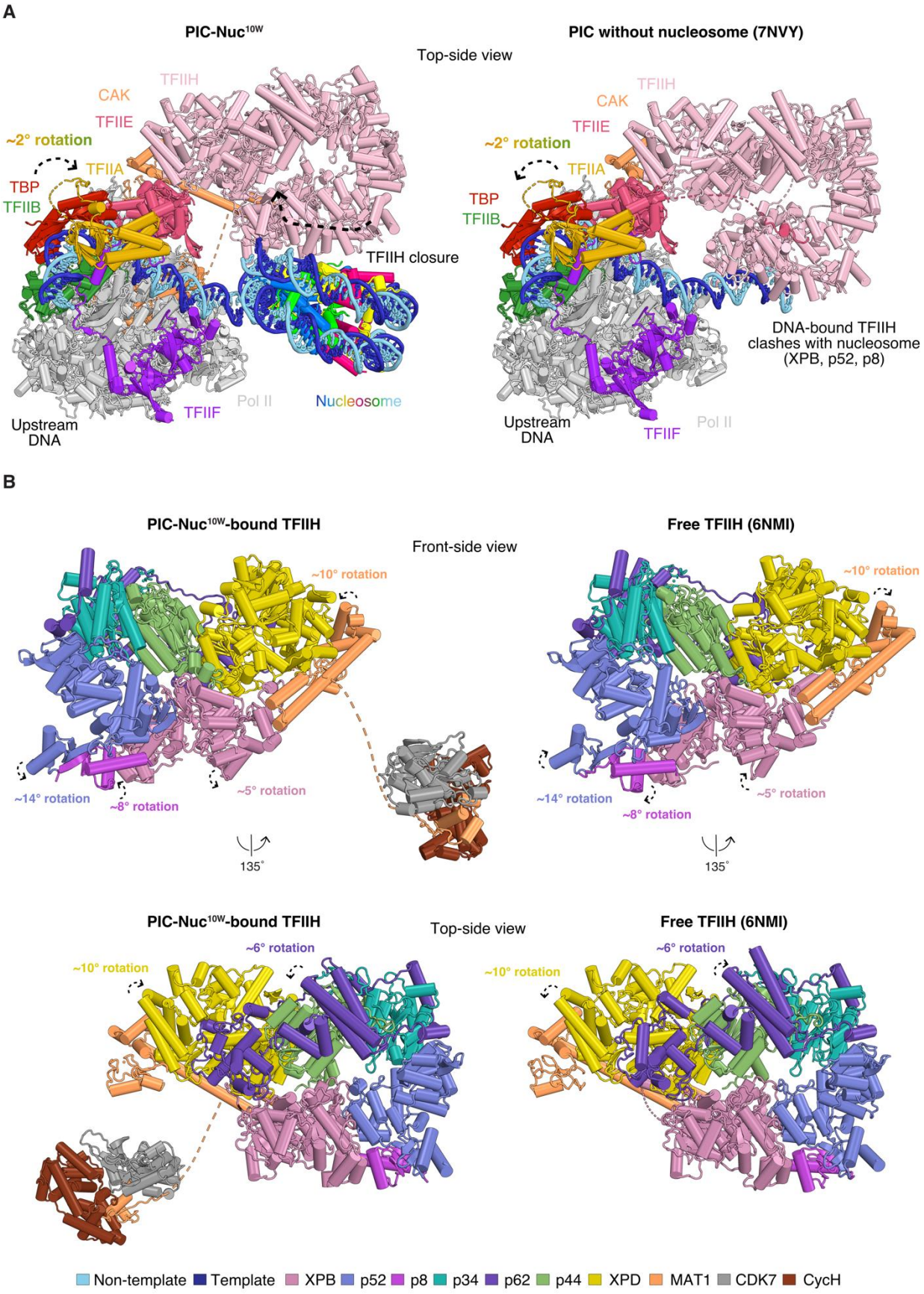

A
PIC-Nuc10W
Top-side view
PIC without nucleosome (7NVY)
TFIIH
CAK
TFIIE
~2° rotation
TFIIA
TBP
TFIIB
TFIIH closure
Pol II
Nucleosome
TFIIF
Upstream DNA
DNA-bound TFIIH clashes with nucleosome (XPB, p52, p8)
B
PIC-Nuc10W-bound TFIIH
Front-side view
Free TFIIH (6NMI)
~10° rotation
~14° rotation
~8° rotation
~5° rotation
135˚
PIC-Nuc10W-bound TFIIH
Top-side view
Free TFIIH (6NMI)
~6° rotation
Non-template
Template
XPB
p52
p8
p34
p62
p44
XPD
MAT1
CDK7
CycH

Figure S6. Comparison of PIC-Nuc^{10W} and PIC-Nuc^{10W}-bound TFIIH with previously reported structures without the nucleosome, Related to Figure 3 and 4.
(A) Comparison of PIC-nucleosome^{10W} on a previously reported nucleosome-free PIC structure (PDB: 7NVY[1]) in top-side view. TFIIH now adopts a closed conformation due to the presence of the nucleosome at the region where XPB normally engages free DNA. Arrowheads on TFIIH indicate the closure of TFIIH. Superimposition of the models was performed by aligning on Pol II. Color of the different components is denoted adjacent to them.
(B) Comparison of TFIIH from PIC-Nuc^{10W} (PIC-Nuc^{10W}) (left panel) with free TFIIH (right panel) (PDB: 6NMI[2]) in different views. TFIIH adopts different conformations due to the different position of the nucleosome. Arrowheads indicate the local motion of TFIIH subunits. Superimposition of the models was performed by aligning on TFIIH. TFIIH is depicted in cartoon representation and the color of the different components is denoted exclusively at the bottom legend.

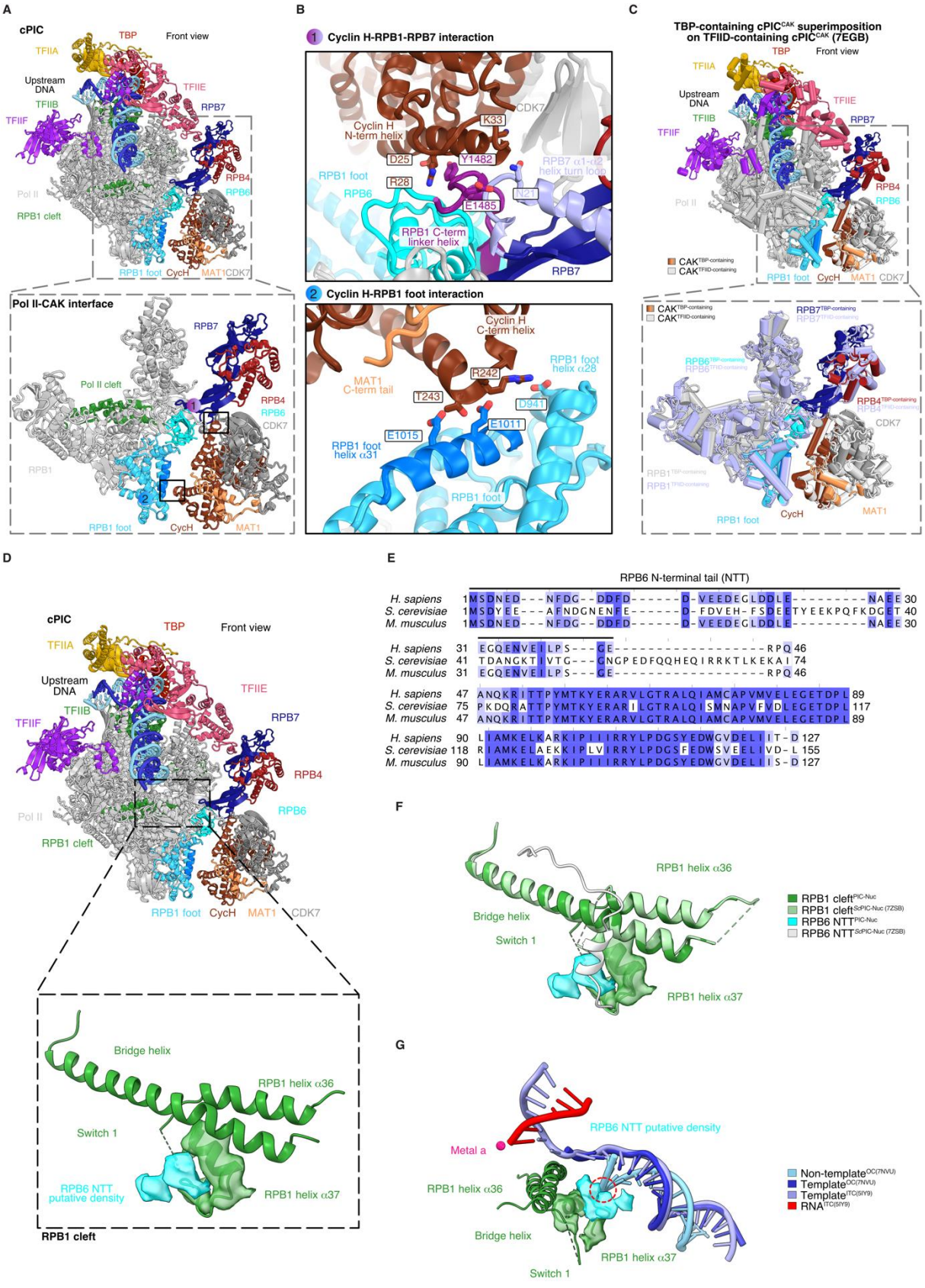

A
cPIC
Front view
TBP
TFIIA
Upstream DNA
TFIIB
TFIIE
TFIIF
RPB7
RPB4
RPB6
Pol II
RPB1 cleft
RPB1 foot
CycH
MAT1
CDK7
Pol II-CAK interface
Pol II cleft
RPB1
B
1 Cyclin H-RPB1-RPB7 interaction
Cyclin H N-term helix
K33
CDK7
D25
Y1482
RPB7 α1-α2 helix turn loop
R28
N21
E1485
RPB1 C-term linker helix
2 Cyclin H-RPB1 foot interaction
Cyclin H C-term helix
MAT1 C-term tail
R242
T243
RPB1 foot helix α28
D941
E1011
E1015
RPB1 foot helix α31
C
TBP-containing cPIC^CAK superimposition on TFIID-containing cPIC^CAK (7EGB)
CAK^TBP-containing
CAK^TFIID-containing
RPB7^TBP-containing
RPB7^TFIID-containing
RPB6^TBP-containing
RPB6^TFIID-containing
RPB4^TBP-containing
RPB4^TFIID-containing
RPB1^TBP-containing
RPB1^TFIID-containing
D
Bridge helix
RPB1 helix α36
Switch 1
RPB6 NTT putative density
RPB1 helix α37
RPB1 cleft
E
RPB6 N-terminal tail (NTT)
H. sapiens
S. cerevisiae
M. musculus
F
RPB1 cleft^PIC-Nuc
RPB1 cleft^ScPIC-Nuc (7ZSB)
RPB6 NTT^PIC-Nuc
RPB6 NTT^ScPIC-Nuc (7ZSB)
G
Metal a
Non-template^OC(7NVU)
Template^OC(7NVU)
Template^ITC(5IY9)
RNA^ITC(5IY9)

Figure S7. A PIC-bound nucleosome stabilizes TFIIH CAK and favors ordering of RPB6 NTT, Related to Figure 2.

(A) Enlarged view of the Pol II-CAK interaction interface, shown in front view and within a dashed box. CAK engages via two contact points between the Pol II stalk (RPB4/RPB7) and foot (RPB1) via its subunit Cyclin H. TFIIH and the nucleosome are omitted to facilitate visualization.

(B) Cyclin H interacts with Pol II through its N- and C-terminal regions. The N-terminal region contacts electrostatically both RPB1 C-terminus linker helix and RPB7 α1-α2 loop, whereas its C-terminus interacts with RPB1 foot helices α28 and α31. Interacting residues are enclosed in rectangular boxes.

(C) Comparison of CAK between TBP-containing (this study) and TFIID-containing PIC complexes (PDB: 7EGB[3]), shown in front view within a dashed box. There are no perceivable differences between these complexes, except for inherent protein local motion.

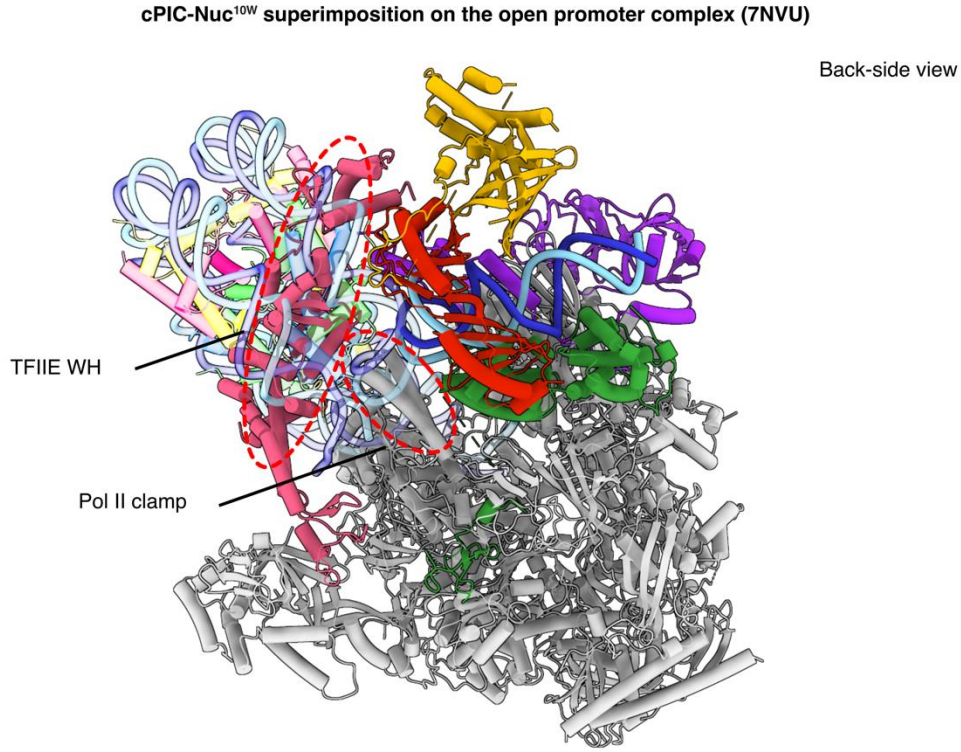
(D) Magnification of the Pol II cleft, shown in front view and within a dashed box. Putative RPB6 NTT density locates above the helix α31 of Pol II cleft. TFIIH and the nucleosome are omitted to facilitate visualization.

(E) Protein sequence alignment of RPB6 orthologs in *H. sapiens*, *S. cerevisiae* and *M. musculus*. The alignment was performed with Jalview 2.11.2.4[4] using the T-Coffee algorithm. The RPB6 NTT is denoted with a solid black line. The color code indicates the conservation scoring across different organisms.

Overlay of the putative RPB6 NTT density of mammalian PIC-nucleosome (this study) on a yeast PIC-nucleosome complex (PDB: 7ZSB[5]) (**F**), a nucleosome-free mammalian PIC complex with opened promoter DNA (PDB: 7NVU[1]) and a human PIC-initial transcribing complex (ITC, PDB: 5IY9[6]) (**G**). The putative NTT clashes with the loaded DNA of opened promoter DNA complexes, suggesting the inhibitory function of the nucleosome favors the ordering of such tail.

Colors of the different components in all panels is denoted adjacent to their corresponding subunits and by legends. Superimposition of models shown was performed by aligning on Pol II.

A


cPIC-Nuc10W superimposition on the open promoter complex (7NVU)
Back-side view
TFIIE WH
Pol II clamp

B

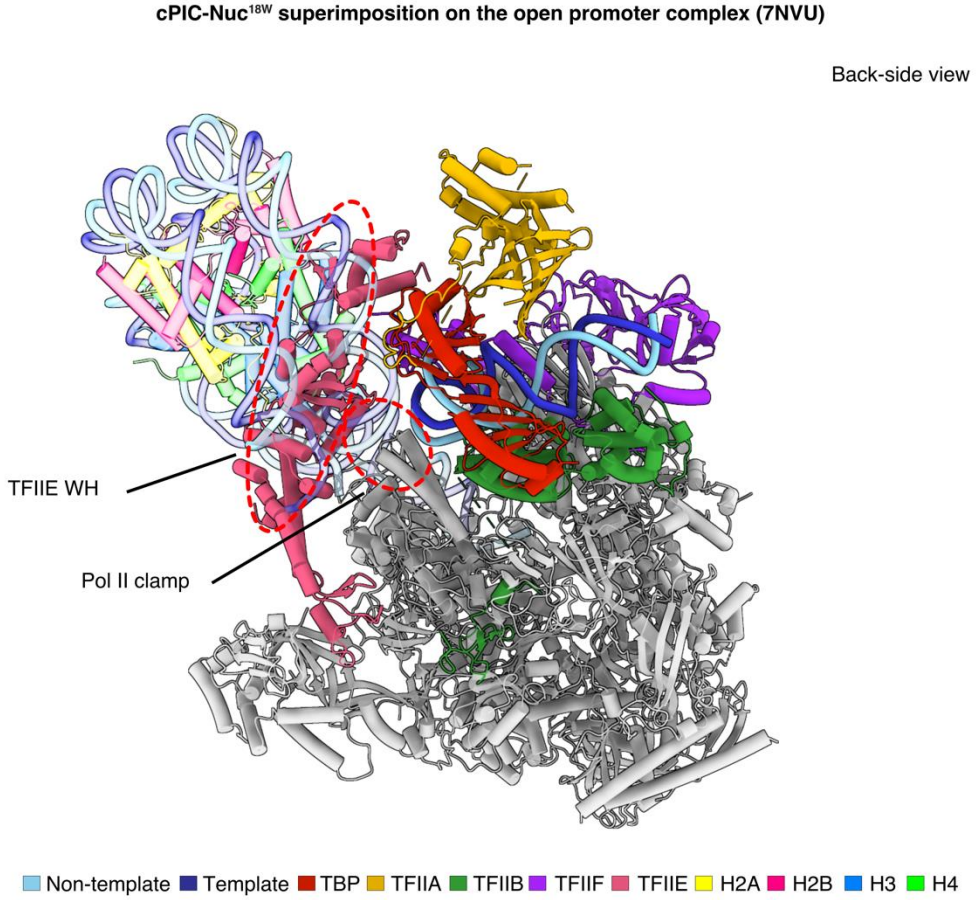

cPIC-Nuc18W superimposition on the open promoter complex (7NVU)
Back-side view
TFIIE WH
Pol II clamp
Non-template
Template
TBP
TFIIA
TFIIB
TFIIF
TFIIE
H2A
H2B
H3
H4

Figure S8. Core PIC-nucleosome complexes are incompatible with DNA opening, Related to Figure 1 and 2.

(A) **(B**) Modelling of our PIC-nucleosome structures missing TFIIH (cPIC) on an open promoter complex (OC, PDB: 7NVU[1]), shown in back-side view. TFIIH-independent spontaneous promoter opening would not be compatible on a nucleosomal template since the nucleosome would clash with the Pol II clamp and TFIIE winged helix (WH). Superimposition of the models was performed by aligning on Pol II and the DNA. Steric clashes are indicated with arrowheads and red dashed ovals, proteins are depicted in cartoon representation and the color of the different components is denoted exclusively at the bottom legend.

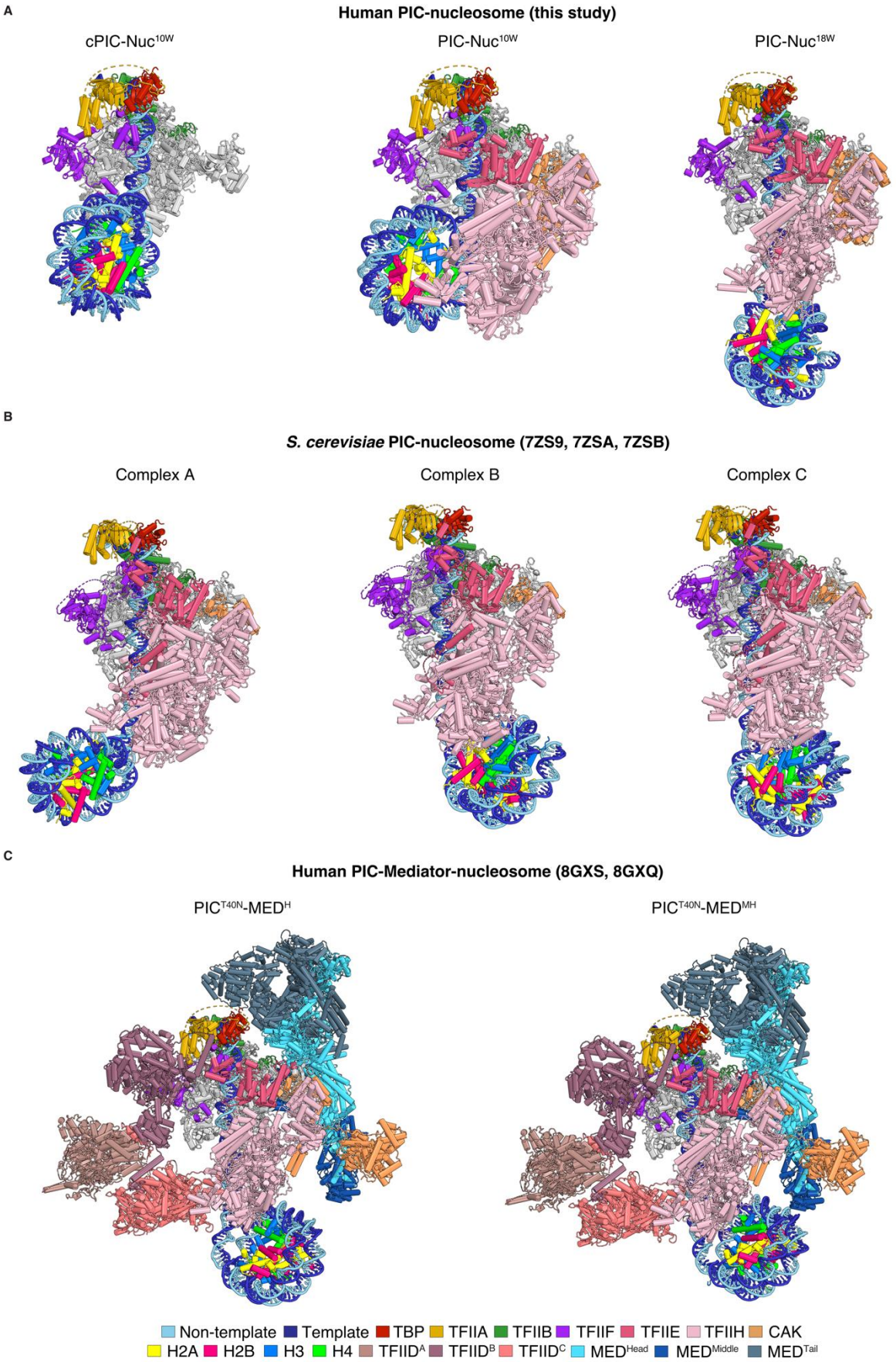

A
Human PIC-nucleosome (this study)
cPIC-Nuc10W
PIC-Nuc10W
PIC-Nuc18W
B
S. cerevisiae PIC-nucleosome (7ZS9, 7ZSA, 7ZSB)
Complex A
Complex B
Complex C
C
Human PIC-Mediator-nucleosome (8GXS, 8GXQ)
PICT40N-MEDH
PICT40N-MEDMH
Non-template
Template
TBP
TFIIA
TFIIB
TFIIF
TFIIE
TFIIH
CAK
H2A
H2B
H3
H4
TFIIDA
TFIIDB
TFIIDC
MEDHead
MEDMiddle
MEDTail

Figure S9. Comparison of human and yeast PIC-nucleosome complexes, related to Figure 7.

Models of the PIC-nucleosome complexes determined in (**A**) this study, (**B**) in yeast (PDB: 7ZS9, 7ZSA, 7ZSB[5]) and (**C**) in human (PDB: 8GXS, 8GXQ[7]) Superimposition of the models was performed by aligning on Pol II and the representations shown in the same view. Color of the different components is denoted exclusively at the bottom legend.

Table S1. Cryo-EM data acquisition, processing and refinement statistics, Related to Figure 2.

Model name PDB code Map code	cPIC-nucleosome^{10W} (PDB: 8BZ1) (EMDB: EMD-16335)	PIC-nucleosome^{10W} (PDB: 8BYQ) (EMDB: EMD-16331)	PIC-nucleosome^{18W} (PDB: 8BVW) (EMDB: EMD-16274)
Data collection and processing			
Magnification	81,000x	81,000x	81,000x
Voltage (kV)	300	300	300
Electron exposure (e–/Å^2)	50.45	50.45	41.58
Defocus range (μm)	0.6 – 1.3	0.6 – 1.3	0.5 – 1.5
Pixel size (Å)	1.05	1.05	1.05
Initial particle images (no.)	4,606,320	4,606,320	4,667,603
Symmetry imposed	C1	C1	C1
Final particle images (no.)	214,161	101,994	188,832
Map resolution (Å) at FSC = 0-143	3.8	4.1	4.0
Map resolution range (Å)	2.8 – 7.7	2.9 – 8.2	2.7 – 8.7
Map sharpening *B* factor (Å^2)	-90	-150	-114
Refinement			
Initial models used (PDB code)	7NVS, 7OHC	7NVS, 7OHC, 6NMI, 6XBZ, 7EGB	7NVS, 7OHC, 7ZSB, 7NVW, 6XBZ, 7EGB
Model composition			
Non-hydrogen atoms	53411	86010	84811
Protein residues	5667	9687	9510
Nucleotides	396	396	412
Ligands	ZN: 9 MG: 1	ZN: 16 MG: 1 SF4: 1 UNK: 154	ZN: 17 MG: 1 SF4: 1 UNK: 27
Mean B factors (Å^2)			
Protein	117.33	176.57	125.41
Nucleotides	145.93	212.46	212.92
Ligand	146.29	223.85	136.37
R.m.s. deviations			
Bond lengths (Å)	0.004	0.005	0.006
Bond angles (°)	0.737	0.860	0.896
Validation			
MolProbity score	1.22	1.47	1.46
Clashscore	4.46	6.52	5.48
Poor rotamers (%)	0.00	0.06	0.04
Ramachandran plot			
Favored (%)	98.19	97.44	96.93
Allowed (%)	1.81	2.56	3.07
Disallowed (%)	0.00	0.00	0.00

PDB: Protein Data Bank; EMD: Electron Microscopy Data Bank; FSC: Fourier Shell Correlation; R.m.s.: root-mean-square

Supplemental references

1. Aibara, S., Schilbach, S., and Cramer, P. (2021). Structures of mammalian RNA polymerase II pre-initiation complexes. Nature *594*, 124-128. 10.1038/s41586-021-03554-8.
2. Greber, B.J., Toso, D.B., Fang, J., and Nogales, E. (2019). The complete structure of the human TFIIH core complex. Elife *8*. 10.7554/eLife.44771.
3. Chen, X., Qi, Y., Wu, Z., Wang, X., Li, J., Zhao, D., Hou, H., Li, Y., Yu, Z., Liu, W., et al. (2021). Structural insights into preinitiation complex assembly on core promoters. Science *372*. 10.1126/science.aba8490.
4. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191. 10.1093/bioinformatics/btp033.
5. Wang, H., Schilbach, S., Ninov, M., Urlaub, H., and Cramer, P. (2022). Structures of transcription preinitiation complex engaged with the +1 nucleosome. Nat Struct Mol Biol. 10.1038/s41594-022-00865-w.
6. He, Y., Yan, C., Fang, J., Inouye, C., Tjian, R., Ivanov, I., and Nogales, E. (2016). Near-atomic resolution visualization of human transcription promoter opening. Nature *533*, 359-365. 10.1038/nature17970.
7. Chen, X., Wang, X., Liu, W., Ren, Y., Qu, X., Li, J., Yin, X., and Xu, Y. (2022). Structures of +1 nucleosome-bound PIC-Mediator complex. Science *378*, 62-68. 10.1126/science.abn8131.