Science
AAAS

Supplementary Materials for

Structural basis of nucleosome retention during transcription elongation

Martin Filipovski, Jelly H. M. Soffers, Seychelle M. Vos, Lucas Farnung

Correspondence to: Lucas_Farnung@hms.harvard.edu

This PDF file includes:

Other Supplementary Materials for this manuscript include the following:

Materials and Methods

Plasmid cloning

The following sequence containing the TspRI cut site and the modified Widom 601 sequence with an A-less cassette to nucleosomal bp +64 was synthesized and inserted into a pIDTSmart-Kan vector (IDT): 5′-ACG AAG CGT AGC ATC ACT GTC TTG TGT TTG GTG TGT CTG GGT GGT GGC CGT TTT CGT TGT TTT TTT CTG TCT CGT GCC TGG TGT CTT GGG TGT TTT CCC CTT GGC GGT TTT TTC GAA GGG GAC AGC GCG TAC GTG CGT TTA AGC GGT GCT AGA GCT GTC TAC GAC CAA TTG AGC GGC CTC GGC ACC GGG ATT CTG AT-3′ (+64 vector). Clones for expression were prepared as previously described (*21*, *29*, *30*).

Protein expression and purification

Sus scrofa RNA polymerase II, *H. sapiens* DSIF, *H. sapiens* PAF1c, *H. sapiens* RTF1, *H. sapiens* SPT6, *H. sapiens* TFIIS, and *H. sapiens* P-TEFb were purified similarly as described (*21*, *29*, *30*).

Xenopus laevis histones were expressed and purified as described previously (*31*). Purified histones were aliquoted, flash-frozen, lyophilized, and stored at −80 °C before use.

For octamer formation, lyophilized histones were resuspended in unfolding buffer (7 M guanidine hydrochloride, 20 mM Tris-HCl pH 7.5, 10 mM DTT) to a concentration of 1.5 mg ml^{-1}. H2A, H2B, H3, and H4 were then combined at a molar ratio of 1.2:1.2:1:1. The sample was incubated on ice for 30 min before it was dialyzed against 3 changes of 600 mL refolding buffer (2 M NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0, 5 mM β-mercaptoethanol) for a total of 18 h at 4 °C. Dialyzed sample was recovered and applied to a Cytiva HiLoad S200 16/600 pg size exclusion column, pre-equilibrated in refolding buffer. Peak fractions containing histone octamer were pooled and concentrated to 30 μM and stored at 4 °C prior to use.

Nucleosome preparation

The +64 vector was used as a template for a large-scale PCR reaction with primers 5′-ACG AAG CGT AGC ATC ACT GTC TTG-3′ and 5′-/Cy5/-TAC GTA TAA TGC CGT AAG ATC ACG CG-3′. The thermocycling conditions were as follows: 1. 95 ºC for 5 min, 2. 95 ºC for 1 min, 3. 60 ºC for 30 s, 4. 72 ºC for 45 s, cycle between steps 2 and 4 thirty-four times, 5. 72 ºC for 5 mins, and 6. pause at 4 ºC. The PCR products were pooled from eight 48-well PCR plates (100 μL per well, 40 mL total volume). DNA was purified using a Resource Q, 1 mL (GE Healthcare) and eluted with a 20–40% NaCl gradient of TE Buffer (10 mM Tris pH 8.0, 2 M NaCl, 1 mM EDTA pH 8.0). Peak fractions were pooled, ethanol-precipitated, resuspended in H_2O, and digested with commercial TspRI (NEB) to generate the 3'-overhangs for RNA binding. TspRI-digested DNA was purified by native PAGE using a Prep Cell apparatus (Bio-Rad).

Nucleosomes were reconstituted essentially as described before (*31*). Histone octamer and DNA were mixed at a 1:1 molar ratio in RB-High Buffer (20 mM HEPES, pH 7.4, 2 M KCl, 1 mM EDTA, pH 8.0, and 1 mM DTT), transferred to Slide-A-Lyzer Mini Dialysis Units 20K MWCO (Thermo Scientific), and gradient dialyzed against RB-Low Buffer (20 mM HEPES, pH 7.4, 30 mM KCl, 1 mM EDTA, pH 8.0, and 1 mM DTT) for 18 hours at 4 ºC. The sample was further dialyzed against RB-Low Buffer for 4 hours at 4 ºC. The sample was centrifuged for 10 min at 21,000*g* to collect precipitate. The nucleosomes were subsequently purified by native PAGE using a Prep Cell apparatus (Bio-Rad). Peak fractions were pooled and concentrated in a

10K MWCO Amicon Ultra-4 Centrifugal Filter Unit (Millipore). Nucleosome concentration was quantified by absorbance at 280 nm. The molar extinction coefficient of the nucleosome was obtained by summing the molar extinction coefficients of the octamer and the DNA components at 280 nm.

Complex preparation for cryo-EM

All concentrations refer to the final concentrations in the transcription reaction. 5,000 nM RNA (5'-/6-FAM/-UUAUCACUGUC-3') and 700 nM nucleosomes were mixed and incubated for 5 min on ice. 700 nM *S. scrofa* Pol II was added to the mixture and incubated for 5 min on ice. DSIF, SPT6, PAF1c (each 1,400 nM), 1,750 nM RTF1, 700 nM P-TEFb, Compensation Buffer (to adjust to final concentrations of buffer components), and 0.5 mM 3'-dATP were added to the reaction. The reaction was incubated for 20 min at 30 ºC. Transcription was started by the addition of CTP, GTP, and UTP (1 mM each), and TFIIS (420 nM). The final concentrations of buffer components were 75 mM NaCl, 20 mM HEPES, pH 7.4, 3 mM $MgCl_2$, 4% (v/v) glycerol, and 1 mM TCEP. The final reaction volume was 100 µL. The reaction was incubated at 30 ºC for 90 min and stopped by adding 2 µL 0.5 M EDTA, pH 8.0. The sample was centrifuged for 10 min at 21,300*g* and applied to a Superose 6 Increase 3.2/300 (Cytiva) on a Äkta pure 25 with Micro kit (Cytiva). 50 µL fractions were collected. Peak fraction samples were applied to NuPAGE 4-12% Bis-Tris, 10-well gel (Invitrogen) and ran in 1X MES buffer for 30 min at 200 V to assess complex formation. The gel was stained with One-Step Blue Protein Gel Stain (Biotum) and imaged. Denaturing gel electrophoresis was performed to assess RNA extension. Peak fraction samples were mixed with 2X STOP Buffer (6.4 M urea, 50 mM EDTA, 2X TBE), incubated with 40 µg proteinase K (NEB) at 37 ºC for 30 minutes, denatured at 95 ºC for 10 minutes, and applied to a 12% denaturing urea gel (8 M urea, 1X TBE Buffer, 12% acrylamide:bis-acrylamide). The gel was run in 1X TBE Buffer for 30 min at 300 V and imaged with a Typhoon 5 (GE), 6-FAM fluorescence at 400 PMT.

Relevant peak fractions were individually crosslinked with 0.1% (v/v) glutaraldehyde for 10 min on ice and then quenched with 8 mM aspartate and 2 mM lysine for 10 min on ice. The reactions were transferred to a Slide-A-Lyzer Mini Dialysis Unit 20 K MWCO (Thermo Scientific) and dialyzed against a buffer containing 75 mM NaCl, 20 mM Na-HEPES pH 7.4, 5 mM Tris-HCl pH 7.5, 3 mM $MgCl_2$, and 1 mM TCEP for 2 hours at 4 ºC. Complex concentrations were quantified by absorbance at 280 nm. The molar extinction coefficient of the complex was obtained by summing the molar extinction coefficient of all individual components. The fraction with highest concentration (64 nM) was selected for analysis by cryo-EM.

Quantifoil R2/1 on 200 Mesh Copper grids were glow discharged for 30 s at 15 mA with 10 s hold time using a Pelco Easiglow plasma discharge system. 2 µL of sample were applied on each side of the grid, incubated for 10 s, blotted with Ted Pella standard vitrobot filter paper for 4 s with blot force 10 and vitrified by plunging into liquid ethane using a Vitrobot Mark IV (FEI Company), operated at 4 °C and 100 % humidity. Sample application from both sites and sample incubation for 10 s has consistently resulted in better ice quality for transcription elongation complexes compared to single-sided sample application.

Transcription assay

All concentrations refer to the final concentrations in the transcription reaction. 120 nM RNA (5'-/6-FAM/-UUAUCACUGUC-3') and 120 nM nucleosomes were mixed and incubated for 5 min on ice. 150 nM *S. scrofa* Pol II was added to the mixture and incubated for 5 min on ice. DSIF, SPT6, and PAF1c (225 nM each), 115 nM P-TEFb, Compensation Buffer, and 0.5 mM 3'-dATP were added to the reaction. The reaction was incubated for 20 min at 30 ºC.

Transcription was started by the addition of CTP, GTP, and UTP (1 mM each), and TFIIS (90 nM). The final concentrations of buffer components were 75 mM NaCl, 20 mM HEPES, pH 7.4, 3 mM $MgCl_2$, 4% (v/v) glycerol, and 1 mM TCEP. The final reaction volume was 50 µL. The reaction was incubated at 30 ºC. Reaction samples were collected after 30 s, 1 min, 2 min, 5 min, 10 min, 30 min, 60 min, and 90 min of incubation and quenched in 2X STOP Buffer. 40 µg proteinase K (NEB) was added to each sample and the reaction was incubated at 37 ºC for 30 minutes and then denatured at 95 ºC for 10 min. RNA extension products were analyzed by denaturing gel electrophoresis and visualized using a Typhoon 5, 6-FAM fluorescence at 400 PMT.

Cryo-electron microscopy and image processing

Cryo-EM data were collected on a ThermoFisher Scientific Titan Krios operated at 300 keV equipped with a Gatan BioQuantum GIF and a Gatan K3 direct electron detector. A total of two datasets were collected from two biological replicates. Data acquisition was automated using SerialEM (v3.8.6) software at a nominal magnification of 105,000×, corresponding to a pixel size of 0.83 Å in nanoprobe EFTEM mode. Movies consisting of 50 frames were collected in counted mode with 2.8 s exposure time. For dataset 1, the electron rate was 18.956 e^{-} $Å^{-2}$ s^{-1} for a total electron exposure of 53.077 e^{-} $Å^{-2}$. For dataset 2, the electron rate was 19.594 e^{-} $Å^{-2}$ s^{-1} for a total electron exposure of 54.863 e^{-} $Å^{-2}$. The two datasets were initially processed independently and merged as indicated (Fig. S2). Image processing and analysis were performed with cryoSPARC (v3.2.0) using default parameters, unless stated otherwise (*32*). Movies were aligned using patch motion correction followed by contrast transfer function (CTF) estimation in cryoSPARC. Particles were picked by blob-based automatic picking, resulting in 858,350 particles from 11,059 micrographs for dataset 1 and 1,502,664 particles from 14,116 micrographs for dataset 2. Particles were extracted with a 450 pixel box size. All classifications and refinements were conducted in cryoSPARC. Volumes employed for masking of areas of interest were generated by low-pass filtering the regions of interest to 15 Å and then using RELION to expand the volume containing the area of interest by 3-5 hard pixels and 3-7 soft pixels. The particles underwent initial two-dimensional (2D) classification and 3D classifications. A subset of particles from dataset 1 was used to generate an input model for 3D classification.

For the Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome structure stalled at +54, initial 2D classification and 3D classification for dataset 1 and dataset 2 resulted in the selection of 270,193 particles and 372,372 particles, respectively. Outputs from the initial 3D classifications were used as input for the heterogeneous 3D classifications. Classes exhibiting Pol II and nucleosome density were subjected to subsequent rounds of heterogeneous 3D classification, resulting in 178,429 particles from both datasets. The merged particle sets were subjected to two additional rounds of heterogeneous refinement in cryoSPARC. The final set of 105,420 particles were subjected to non-uniform refinement with an overall resolution of 3.0 Å (map A) (*33*). The final set of particles was additionally subjected to 3D classification with focus on the CTR9-WDR61 region, resulting in 38,693 particles. To better resolve the Pol II active site, the nucleosome, and the CTR9-WDR61 region, local refinement for Pol II active site, nucleosome, and CTR9-WDR61. This resulted in reconstructions at 3.0 Å (Pol II active site, map B), 3.8 Å (nucleosome, map C), and 6.5 Å (CTR9-WDR61, map D).

Using Frankenmap and Noise2map with default parameters and a diameter of 360 Å (*34*), maps A, C, and D were subsequently used to generate a composite map (map E). Specifically, the composite map was generated by aligning the locally refined maps of the nucleosome (map C) and CTR9-WDR61 (map D) onto map A. Map E was subsequently local-resolution filtered

and used for the refinement of the atomic model. Data processing, data quality, and metrics of the rewrapped complex are reported in Table S1.

To obtain the map of Pol II-DSIF-SPT6-TFIIS-nucleosome complex stalled 38 base pairs within the nucleosome, all particles from dataset 1 and dataset 2 were combined and subjected to initial 2D classification and 3D classification. A class of 801,646 particles exhibiting distinctive Pol II and nucleosome density was selected for subsequent rounds of 3D classification. The resulting 406,170 particles were subjected to 3D variability analysis in order to generate a family of structures with novel nucleosome conformations. Four of these structures were selected as inputs for subsequent rounds of 3D classification, resulting in 64,141 particles. The final set of particles was subjected to non-uniform refinement with an overall resolution of 3.3 Å (map F). The final set of particles was additionally subjected to 3D classification with focus on the SPT4/5 region, resulting in 16,556 particles. To better resolve the Pol II active site, the nucleosome, and SPT4/5 region, local refinements were performed for Pol II active site, nucleosome, and SPT4/5. This resulted in reconstructions at 3.1 Å (Pol II active site, map G), 5.9 Å (nucleosome, map H), and 3.3 Å (SPT4/5, map I).

Using Frankenmap and Noise2map, maps G and H were used to generate a composite map (map J). Map J was used for the refinement of the atomic model. Data processing, data quality, and metrics of the +38 complex are reported in Table S2.

Model building and refinement

Structures of the complete activated elongation complex (PDB 6TED) (*21*), nucleosome (PDB 3LZ0), and an AlphaFold generated model of TFIIS were rigid body docked into the local resolution filtered composite map, locally adjusted, and refined using Coot (v0.9.7) and ISOLDE (v1.3) (*35*, *36*). The novel conformation of the DNA was modeled by *de novo* DNA generation in UCSF ChimeraX (v 1.3) (*37*) and trajectory adjustments in Coot, ISOLDE and PyMol. Density in the active site of Pol II allowed unambiguous assignment of the DNA register by defining purine and pyrimidine bases in the DNA·RNA hybrid. Identification of the register was conducted in map B. The post-translocated state is modeled in the final submission. All PAF1c subunits, RTF1, SPT6 chains were side chain stubbed.

For the +38 complex, structures of the complete activated elongation complex (PDB 6TED) and the nucleosome (PDB 3LZ0) were rigid body docked into map J. The structures were locally adjusted and refined using Coot and ISOLDE. DNA trajectory adjustments were done in Coot, ISOLDE, and PyMol. Density in the active site of Pol II allowed unambiguous assignment of the DNA register. CTR9, WDR61, CDC73, parts of PAF1, and RTF1 were removed from the model due to low-resolution and ambiguous density. SPT4 and SPT5 NGN domains were rigid-body docked into a low-resolution filtered map of a SPT4/5-core Pol II local refinement. The SPT6 chain was side chain stubbed.

Both models were refined with ISOLDE and Coot. The models were subsequently real space refined in PHENIX (v 1.20.1) (*38*) with base-pair restraints. Refinement statistics are reported in Table S3. Additional information on input structural models and model confidence is given in Table S4.

Figure generation

Figures and movie S1 were generated in UCSF ChimeraX. Angular distribution plots and FSC curves were generated in cryoSPARC and adjusted in Adobe Illustrator 2022.

Supplementary Text

Displacement of the NGN and KOW1 domains could allow for the observed change in the trajectory of the upstream DNA to allow the DNA to reach towards the partially transcribed nucleosome. Second, the N-terminal part of SPT5 that connects to the NGN domain has been shown to bind the exposed proximal H2A-H2B dimer during Pol II traversal through a nucleosome (*12*). Therefore, the SPT5 N-terminus could compete with rewrapping of upstream DNA and only displacement of SPT4 and the SPT5 NGN and KOW1 domains could sufficiently disrupt this interaction and allow for DNA rewrapping. These observations suggest a hand-off mechanism where different binding partners initially stabilize the proximal H2A-H2B dimer until transcribed upstream DNA can rewrap the histone octamer to ensure nucleosome retention. Stabilization of the transferred histone octamer may, however, still require the action of histone chaperones such as FACT. It is possible that the distal H2A-H2B dimer can be lost during this process, resulting in the formation of a hexasome (*4*, *39*). This observation is consistent with *in vitro* and *in vivo* experiments that suggest nucleosome retention is dependent on survival of the proximal H2A-H2B dimer (*2*, *8*). Our findings also have implications for the role of the histone chaperone FACT during chromatin transcription. Previous structures have shown that FACT recognizes partially transcribed nucleosomes and the C-terminal domain of FACT subunit SPT16 can bind the exposed DNA interface of the H2A-H2B dimer to prevent its loss (*12*, *40*). Further transcription will then displace FACT and rewrapping of upstream DNA will stabilize the H2A-H2B dimer, as visualized in our structure. Overall, this model suggests that FACT may bind the transcribed nucleosome in a 25-30 bp window during Pol II passage through the proximal half of the nucleosome.

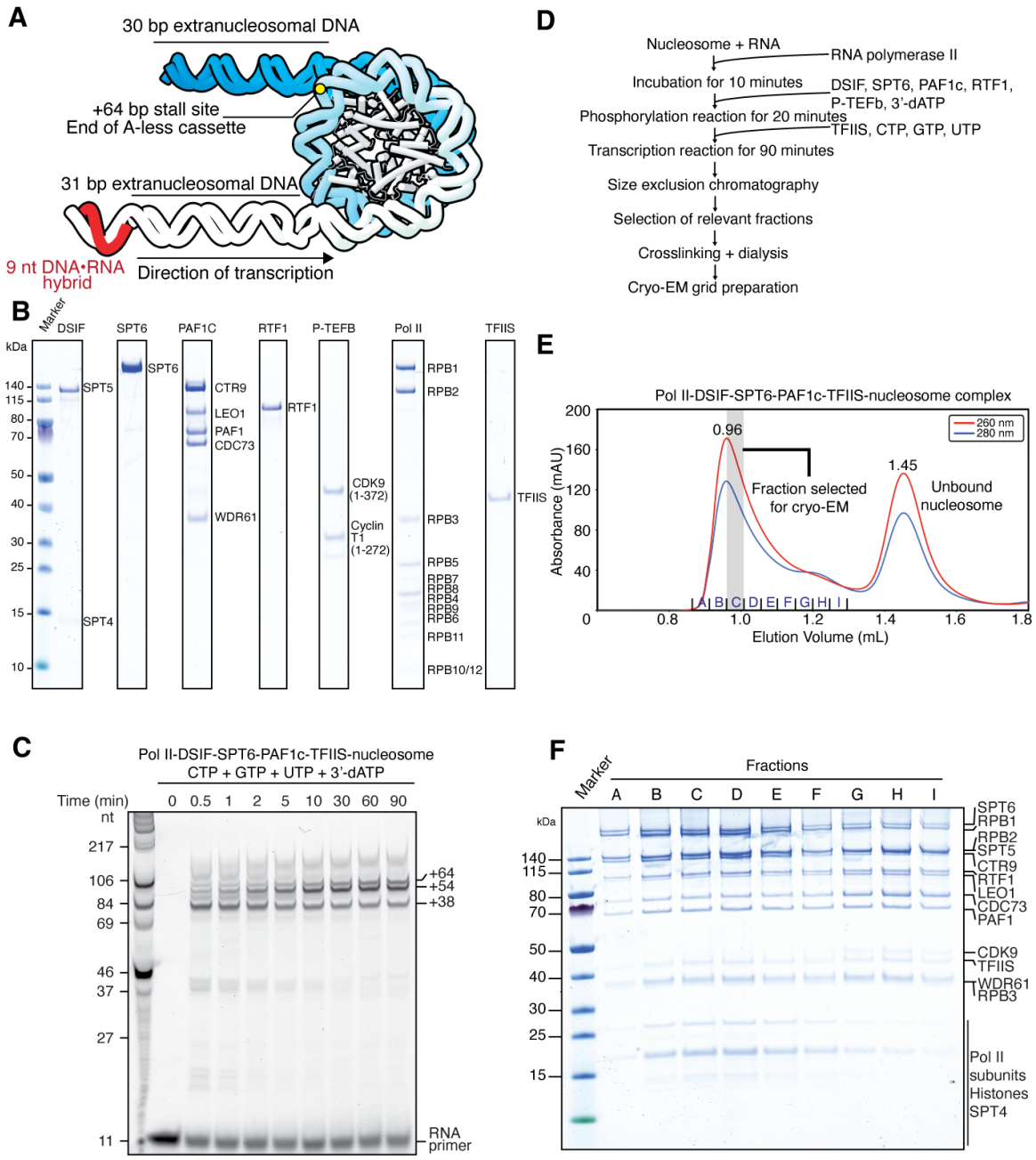

Fig. S1. Preparation of proteins, complex, and transcription assay.

(A) Schematic of nucleosome substrate used for the formation of the elongation complex. 11-nt RNA primer is shown in red. Position of +64 bp stall site and end of A-less cassette is indicated. **(B)** SDS-PAGE of purified protein components. Samples were run on a 4-12 % Bis-Tris SDS-PAGE in 1X MES Running buffer. The gel was stained with Coomassie Blue (OneStep Blue). **(C)** Time course of RNA extension reaction shows robust extension of the RNA primer into the nucleosome. Stall sites within the nucleosome are indicated with the base pair position in the nucleosome. **(D)** Flowchart of complex formation and preparation of sample for single-particle cryo-EM. **(E)** Chromatogram of size exclusion chromatography run of Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex formation. **(F)** SDS-PAGE of fractions of size exclusion run (E).

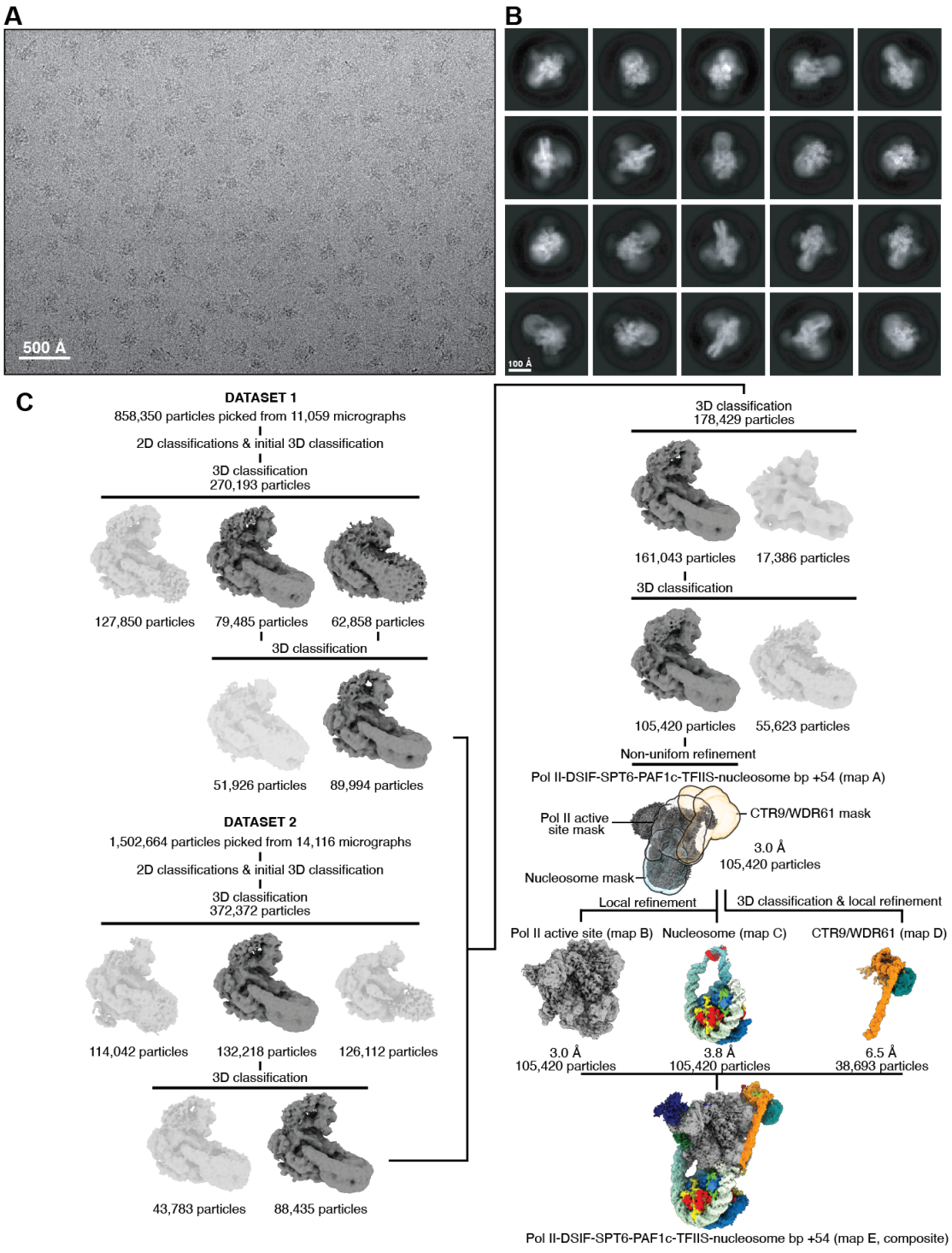

Fig. S2. Cryo-EM data processing of rewrapped complex

(A) Representative micrograph of cryo-EM data collection with scale bar (50 nm). **(B)** Representative 2D classes of mammalian Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex with scale bar (10 nm). 2D classes show Pol II-DSIF-SPT6-PAF1c-TFIIS elongation complex-like and nucleosome-like densities. **(C)** Sorting and classification tree of cryo-EM data analysis. Final maps with resolution and masks are indicated.

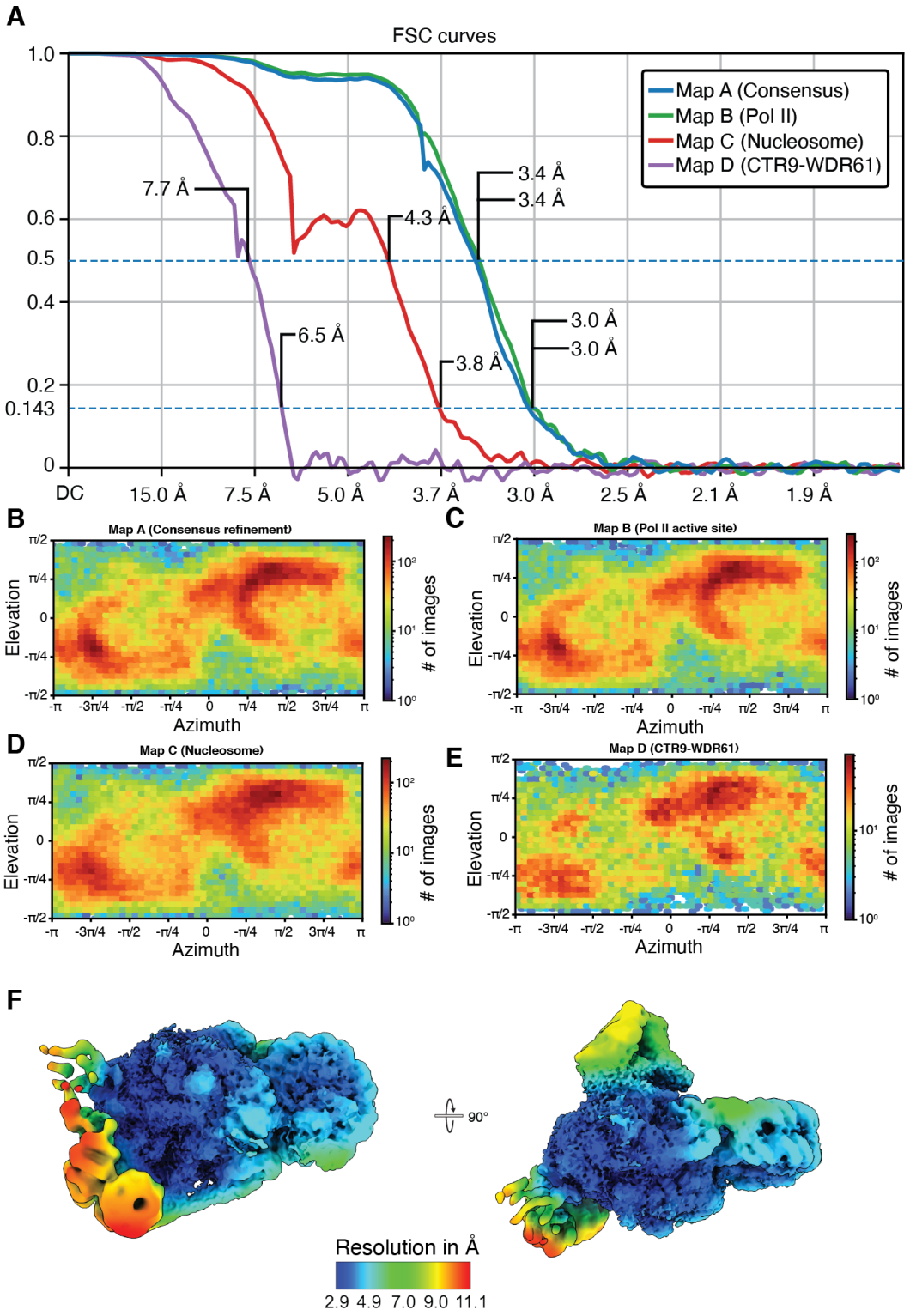

Fig. S3. Data quality and metrics of the rewrapped complex.

(A) FSC curves of maps A-D. Resolutions at the FSC threshold criterions 0.143 and 0.5 are indicated. **(B-E)** Angular distribution plot of particles employed to reconstruct maps A-D. **(F)** Local resolution of composite map E.

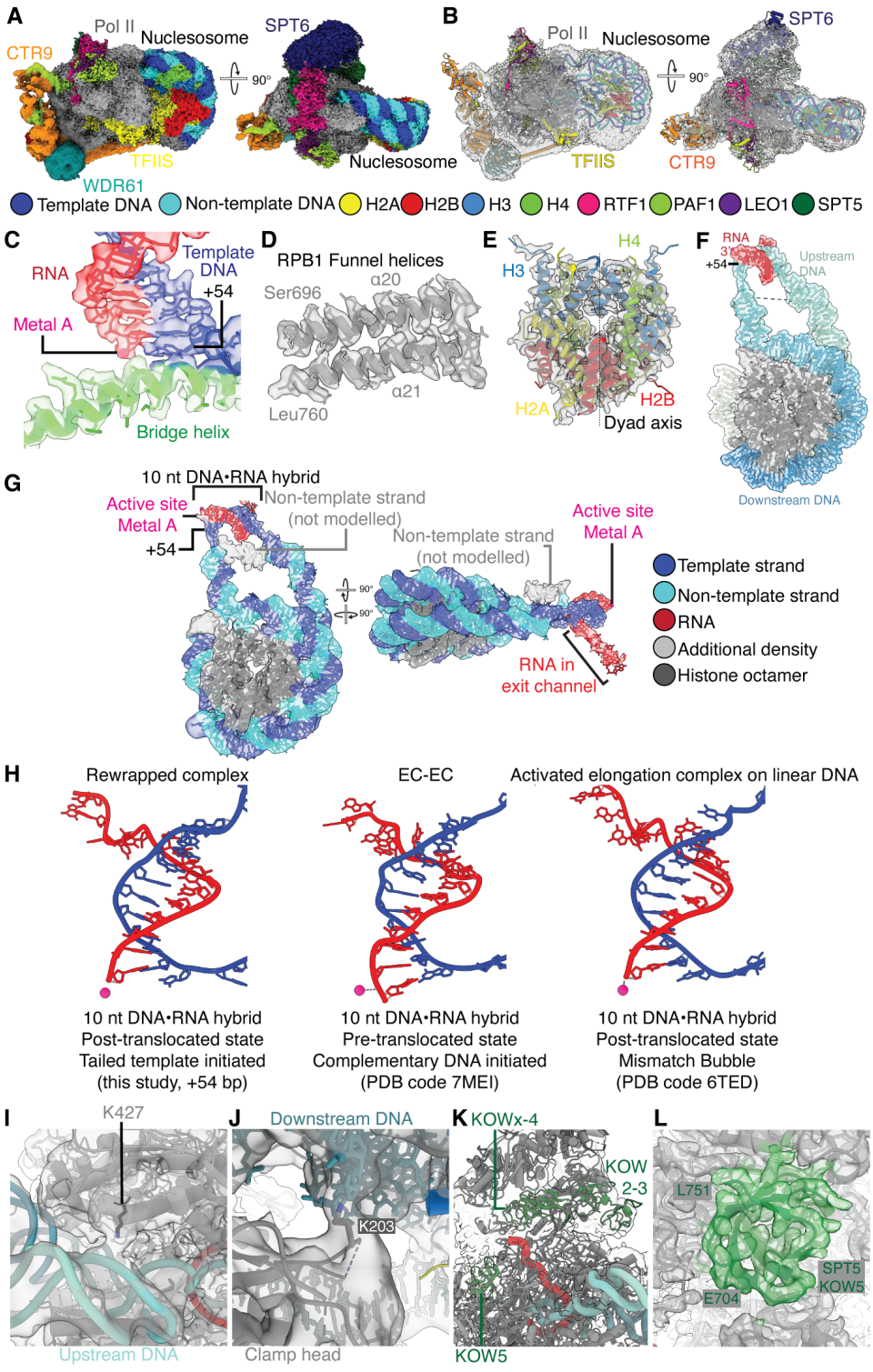

Fig. S4. Cryo-EM densities of the rewrapped complex.

(A) Cryo-EM map (map E) of Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex stalled at nucleosomal bp +54. **(B)** Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex structure with corresponding cryo-EM map (map E). Map E is shown in gray. **(C)** Active site of Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex structure with corresponding cryo-EM map (map B).

Metal A is shown in pink and the bridge helix in green. **(D)** Funnel helices of RPB1 (α20 and α21, gray) with corresponding cryo-EM map (map B). **(E)** Histone octamer with corresponding cryo-EM map (map C). **(F)** Nucleic acids with corresponding cryo-EM map (map C). DNA is colored with a gradient from white (upstream) to blue (downstream). **(G)** Nucleic acids with corresponding density reveal extra density for the non-template strand in the transcription bubble. **(H)** Comparison of DNA·RNA hybrid with DNA·RNA hybrids from other elongation complexes reveals a properly formed transcription bubble in the +54 complex. **(I)** Density for RPB2 K427 residue (map E). **(J)** Density for clamp head residue RPB1 K203 (map E). **(K)** Densities for KOW2-3, KOWx-4, and KOW5 (map E). **(L)** Densities for KOW5 (map B).

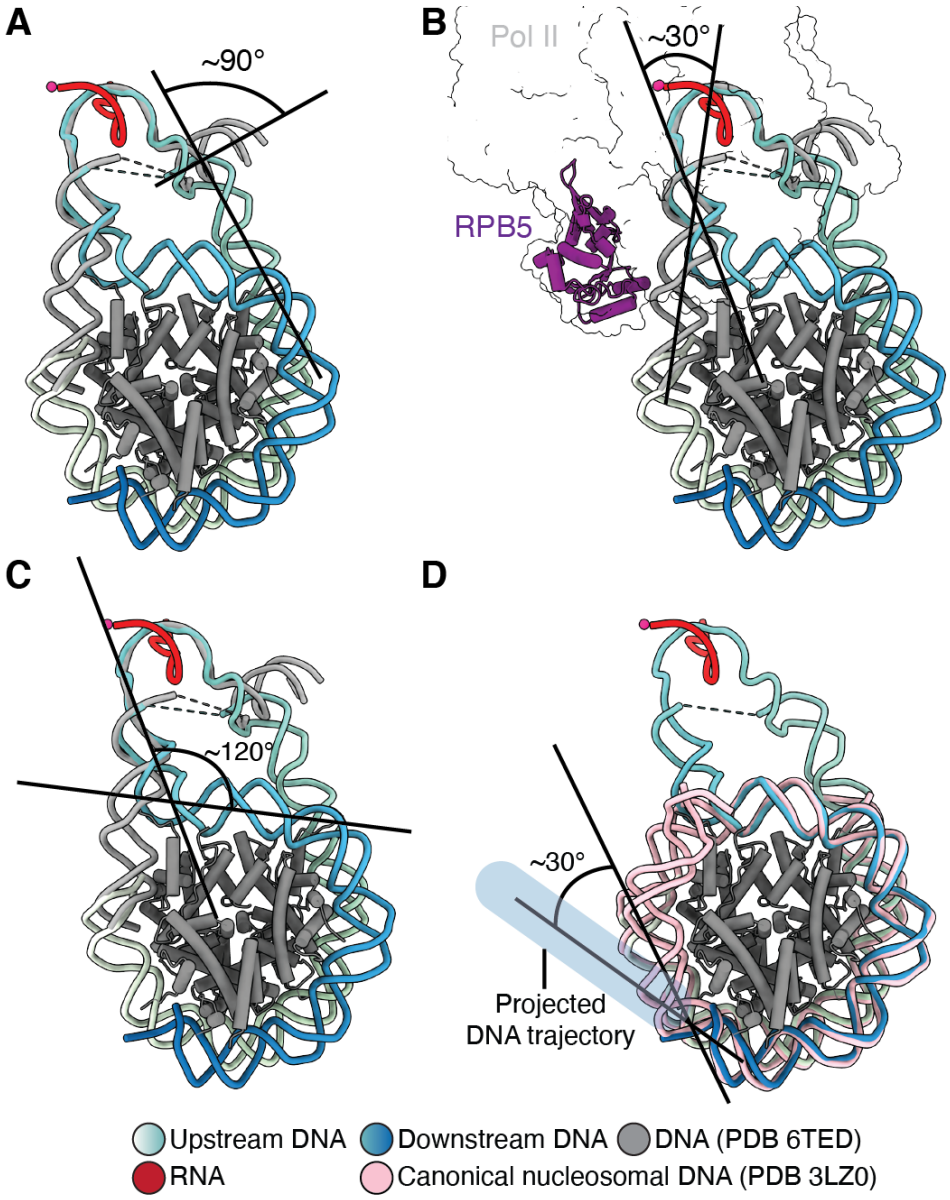

Fig. S5. Comparison of DNA trajectories.

(A) Superposition of nucleic acids from the Pol II-DSIF-SPT6-PAF1c complex on a linear DNA template (gray, PDB 6TED) and Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex (blue to white gradient, this study) reveals a change in the upstream DNA trajectory by 90°. Structures were aligned on Pol II. **(B)** Superposition of nucleic acids from the Pol II-DSIF-SPT6-PAF1c complex on a linear DNA template (gray, PDB 6TED) and Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex (blue to white gradient, this study) reveals a change in the downstream DNA trajectory by 30° with the Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex downstream DNA moving away from RPB5 (purple, PDB 6TED). Structures were aligned on Pol II. **(C)** The Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex downstream DNA (turquoise to blue gradient, this study) is contorted by 120° when it engages the nucleosome at SHL -0.5. **(D)** Comparison of canonical nucleosomal DNA (pink, PDB 3LZ0) and rewrapped

nucleosome (this study). Rewrapped nucleosome shows unwrapping from SHL +5 to SHL +7. The ~30° angle is measured between the trajectory of the canonical nucleosome and the rewrapped nucleosome at SHL+5. Linear projection of DNA from rewrapped nucleosome is indicated as a transparent blue cylinder.

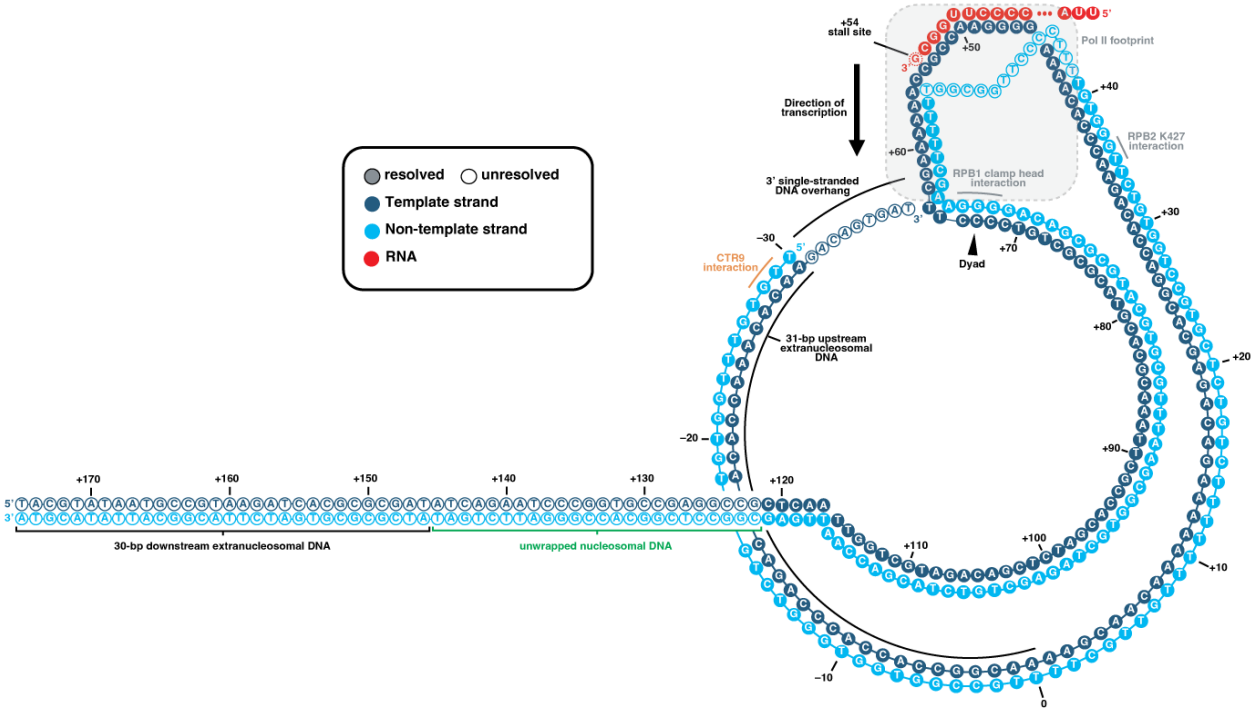

Fig. S6. DNA path overview.

Protein-nucleosome DNA contacts of Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex. Nucleotides are depicted as solid spheres (resolved) or empty spheres (not resolved and not modelled in deposited PDB structure).

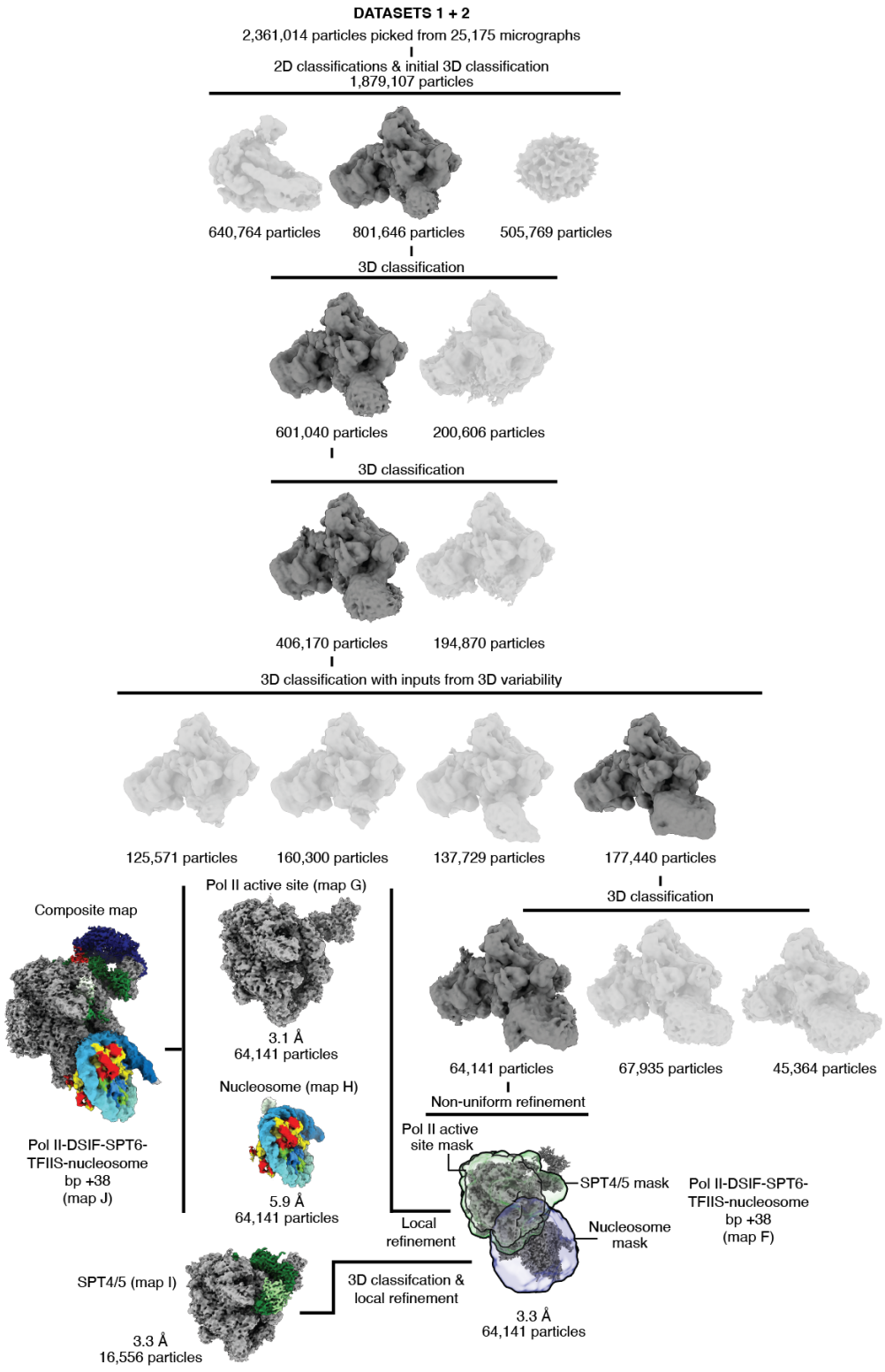

Fig. S7. Cryo-EM data processing of the +38 complex.

Sorting and classification tree of the Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex stalled at nucleosomal bp +38. Final maps with resolution and masks are indicated.

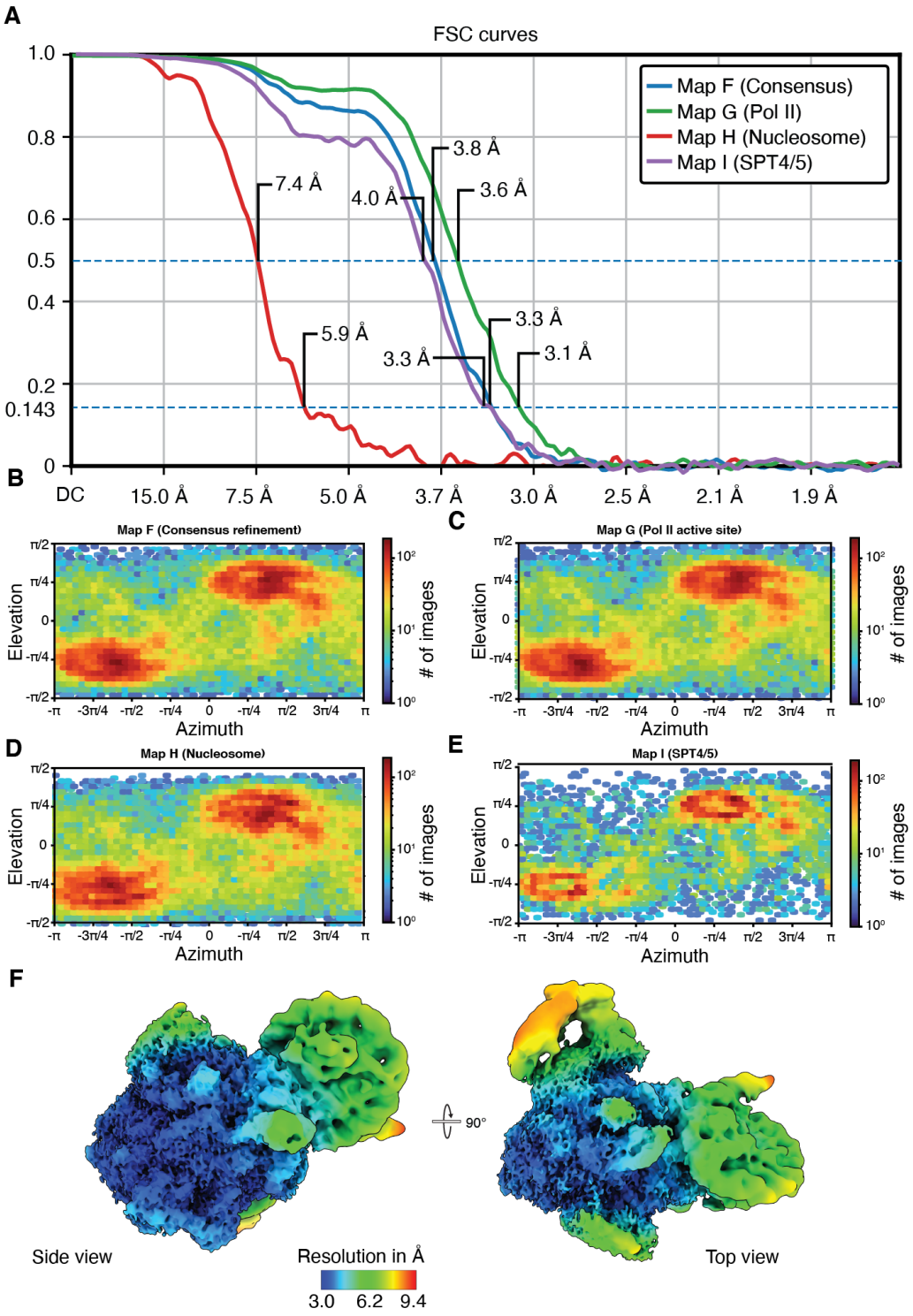

Fig. S8. Data quality and metrics of the +38 complex.

(A) FSC curves of maps F-I of Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex stalled at nucleosomal bp +38. Resolutions at the FSC threshold criterions 0.143 and 0.5 are indicated. **(B-E)** Angular distribution plot of particles employed to reconstruct maps F-I. **(F)** Local resolution of composite map (map J).

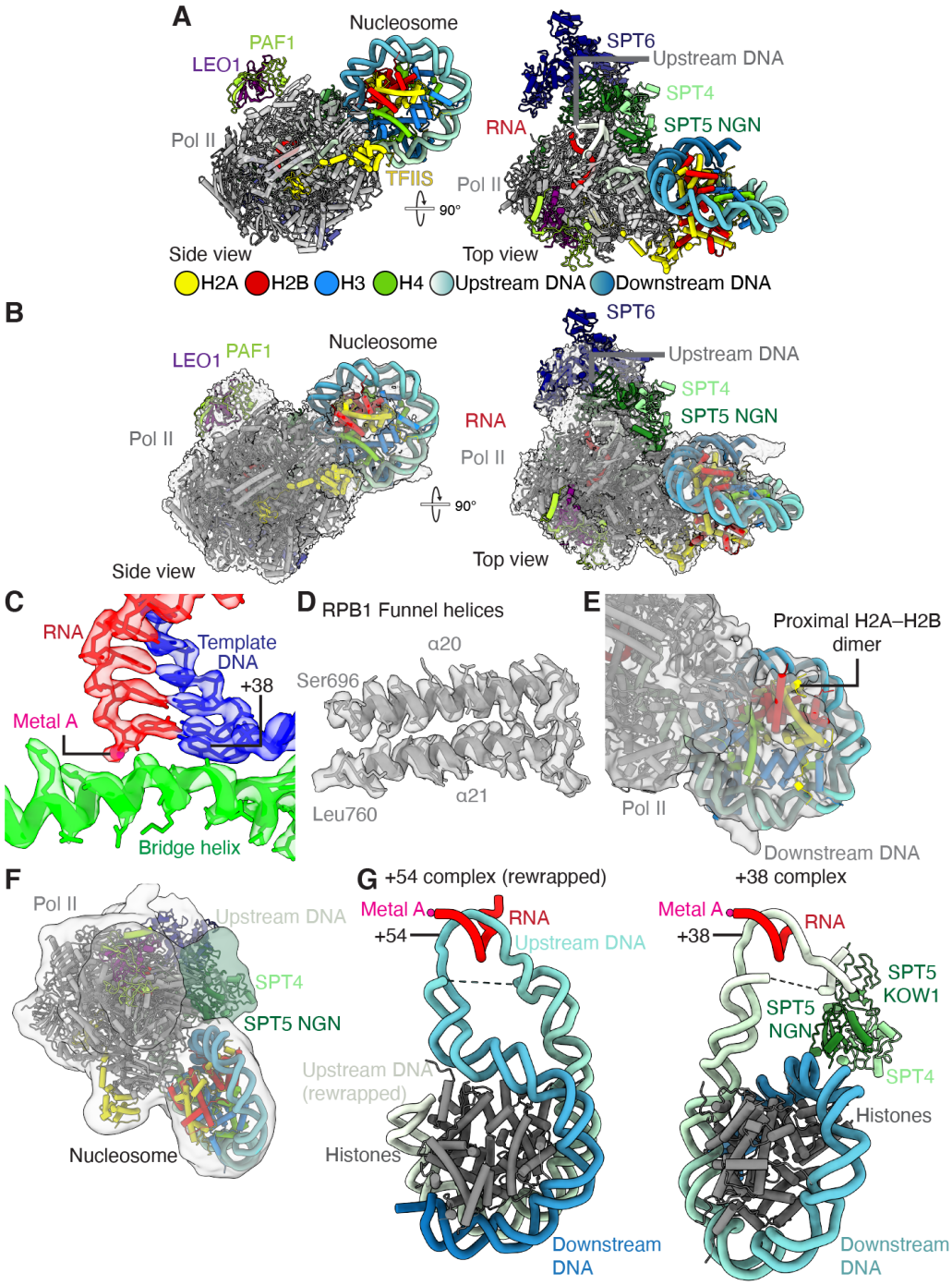

Fig. S9. Overview and cryo-EM densities of the +38 complex.

(A) Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex stalled at nucleosomal bp +38. **(B)** +38 Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome structure with corresponding cryo-EM map (map F). Cryo-EM map is shown in gray. **(C)** Active site of Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome complex stalled at nucleosomal bp +38 with corresponding density (map G). **(D)** RPB1 funnel helices with corresponding density (map G). **I** Partially unraveled nucleosome with corresponding density (map H). **(F)** SPT4 and SPT5 NGN and KOW1 domains with corresponding density in low pass filtered map I. **(G)** Comparison of rewrapped nucleosome complex stalled at +54 with complex stalled at +38.

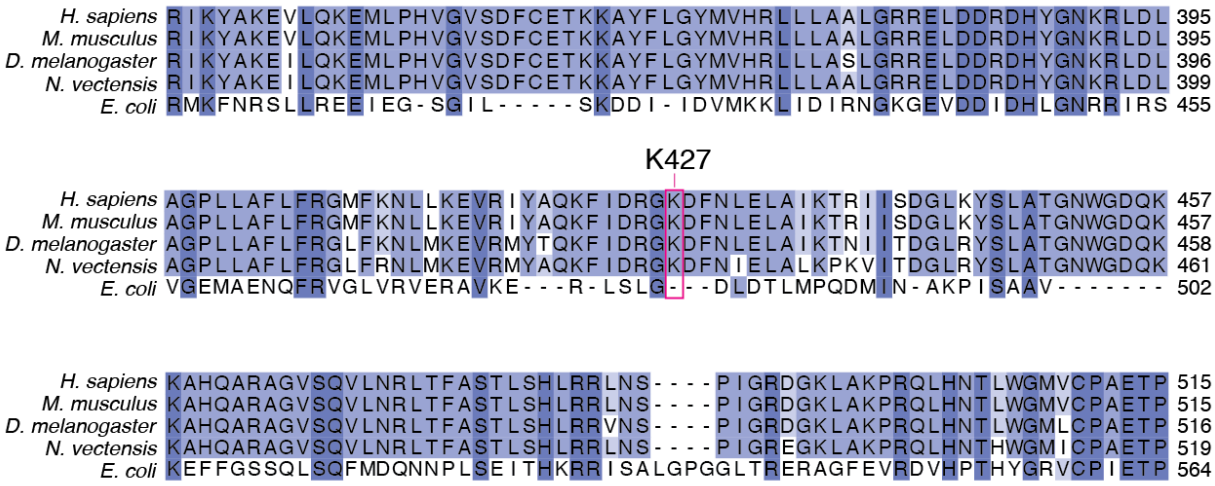

Fig. S10. Multiple sequence alignment of RPB2 K427.

Multiple sequence alignment of RPB2 protrusion domain reveals conservation of residue K427 in a broad range of eukaryotes, but not in *E. coli*.

Table S1. Cryo-EM data collection, refinement, and validation statistics for the rewrapped complex.

Data collection and Processing				
Microscope	Thermo Fisher Titan Krios			
Voltage (keV)	300			
Camera	Gatan K3			
Magnification	105,000			
Pixel size at detector (Å per pixel)	0.83			
Total electron exposure (e^{-} Å^{-2})	53.73, 54.8			
Exposure rate (e^{-} Å^{-2} s^{-1})	18.956, 19.594			
Number of frames collected during exposure	50			
Defocus range (μm)	0.6-1.8			
Automation software	SerialEM			
Energy filter slit width (eV)	20			
Micrographs collected (no.)	25,175			
Micrographs used (no.)	24,964			
Total extracted particles (no.)	2,361,014			
	Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome, bp +54, composite (Map A)	Pol II (Active site), local refinement (Map B)	Nucleosome, local refinement (Map C)	CTR9-WDR61 local refinement (Map D)
	EMD-26620	EMD-26620	EMD-26620	EMD-26620
Refined particles (no.)	1,879,107	1,879,107	1,879,107	1,879,107
Final particles (no.)	105,420	105,420	105,420	38,693
Resolution (global, Å)				
FSC 0.5 unmasked	4.6	7.4	8.2	12.2
FSC 0.5 masked	3.4	3.4	4.3	7.7
FSC 0.143 unmasked	3.8	4.1	6.5	8.7
FSC 0.143 masked	3.0	3.0	3.8	6.5
Map sharpening *B* factor (Å^2)	66.6	65.7	107.7	278.9
3DFSC sphericity	0.664	0.901	0.909	0.884
Map sharpening methods	cryoSPARC	cryoSPARC	cryoSPARC	cryoSPARC

Table S2. Cryo-EM data collection, refinement, and validation statistics for the +38 complex.

Data collection and Processing				
Microscope	Thermo Fisher Titan Krios			
Voltage (keV)	300			
Camera	Gatan K3			
Magnification	105,000			
Pixel size at detector (Å per pixel)	0.83			
Total electron exposure (e^{-} Å^{-2})	53.73, 54.8			
Exposure rate (e^{-} Å^{-2} s^{-1})	18.956, 19.594			
Number of frames collected during exposure	50			
Defocus range (μm)	0.6-1.8			
Automation software	SerialEM			
Energy filter slit width (eV)	20			
Micrographs collected (no.)	25,175			
Micrographs used (no.)	24,964			
Total extracted particles (no.)	2,361,014			
	Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome, bp +38, composite (Map F)	Pol II (Active site), local refinement (Map G)	Nucleosome, local refinement (Map H)	SPT4/5, local refinement (Map I)
	EMD-26621	EMD-26621	EMD-26621	EMD-26621
Refined particles (no.)	1,879,107	1,879,107	1,879,107	1,879,107
Final particles (no.)	64,142	64,142	64,142	16,556
Resolution (global, Å)				
FSC 0.5 unmasked	8.4	8.2	9.8	10.8
FSC 0.5 masked	3.8	3.6	7.4	4.0
FSC 0.143 unmasked	6.0	4.4	7.8	7.4
FSC 0.143 masked	3.3	3.1	5.9	3.3
Map sharpening *B* factor (Å^{2})	43.8	64.4	406.3	36.4
3DFSC sphericity	0.915	0.901	0.863	0.758
Map sharpening methods	cryoSPARC	cryoSPARC	cryoSPARC	cryoSPARC

Table S3. Model composition, refinement, and validation for the rewrapped and the +38 complex.

	PolII-DSIF-SPT6-PAF1c-TFIIS-nucleosome, bp +54 PDB: 7UNC EMD-26620	Pol II-DSIF-SPT6-PAF1c-TFIIS-nucleosome, bp +38 PDB: 7UND EMD-26621
Model composition		
Non-hydrogen atoms (no.)	57,638	52,909
Protein residues (no.)	7,938	6,430
Ligands (no.)	10	10
Nucleotide residues (no.)	308	247
Model Refinement		
Initial models used (PDB #)	3LZ0, 6TED	3LZ0, 6TED
Refinement packages	Coot, ISOLDE, PHENIX real space	Coot, ISOLDE, PHENIX real space
Model-Map scores		
Cross-correlation coefficient	0.73	0.68
Model resolution (Å)	3.0	3.4
FSC threshold	0.143	0.143
Mean *B* factors (Å^2)		
Protein residues (no.)	189.36	162.48
Ligands (no.)	187.46	189.70
Nucleotide residues (no.)	177.12	426.05
R.m.s. deviations from ideal values		
Bond lengths (Å)	0.007	0.004
Bond angles (°)	1.167	0.857
Validation		
MolProbity score	1.80	2.15
CaBLAM outliers (%)	2.82	3.28
Clashscore	6.01	7.85
Poor rotamers (%)	1.07	2.08
C-beta deviations (%)	0.00	0.00
EMRinger score	2.51	2.06
Ramachandran plot		
Favored (%)	92.89	92.09
Allowed (%)	7.05	7.50
Outliers (%)	0.06	0.41

5

Table S4. Input structural models and model confidence

Complex/domain	Chain id(s)	Input model	Level of confidence	Complex (PDB)
Pol II	A-L	6TED	Atomic	7UNC
SPT5 KOW5	Z	6TED	Atomic	7UNC
SPT5 KOW2-3, KOWx-4	Z	6TED	Rigid body fitting	7UNC
SPT6	M	6TED	Rigid body fitting, stubbed	7UNC
PAF1	V	6TED	Rigid body fitting, stubbed	7UNC
LEO1	U	6TED	Rigid body fitting, stubbed	7UNC
CTR9	Q	6TED	Rigid body fitting, stubbed	7UNC
CDC73	X	6TED	Rigid body fitting, stubbed	7UNC
WDR61	W	6TED	Rigid body fitting, stubbed	7UNC
RTF1	R	6TED	Rigid body fitting, stubbed	7UNC
TFIIS	O	AlphaFold	Rigid body fitting, stubbed	7UNC
Nucleosome	a-h	3LZ0	Atomic	7UNC
DNA/RNA	N, P, T	*De novo*/6TED	Atomic	7UNC
Pol II	A-L	6TED	Atomic	7UND
SPT5 KOW5	Z	6TED	Atomic	7UND
SPT5 KOW2-3, KOWx-4	Z	6TED	Rigid body fitting	7UND
SPT5 NGN, KOW1	Z	6TED	Rigid body fitting	7UND
SPT4	Y	6TED	Rigid body fitting	7UND
SPT6	M	6TED	Rigid body fitting, stubbed	7UND
PAF1	V	6TED	Rigid body fitting, stubbed	7UND
LEO1	U	6TED	Rigid body fitting, stubbed	7UND
TFIIS	O	AlphaFold	Rigid body fitting	7UND
Nucleosome	a-h	3LZ0	Atomic	7UND
DNA/RNA	N, P, T	*De novo*/6TED	Atomic	7UND

Movie S1.

Mammalian RNA polymerase II−DSIF−SPT6-PAF1c-TFIIS-nucleosome structure with rewrapped upstream DNA and density (local resolution filtered composite map).