bioRχiv

Structure and function of IWS1 in transcription elongation

Della Syau[1], Felix Steinruecke[1], Sophie Roth[1], Ernst Schmid[2], Karen Adelman[2,3,4], Johannes Walter[2,5], and Lucas Farnung[1,*]

[1]Department of Cell Biology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA
[2]Department of Biological Chemistry & Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA
[3]The Eli and Edythe L. Broad Institute, Cambridge MA 02142, USAUSA
[4]Ludwig Center at Harvard, Boston, MA 02115, USA
[5]Howard Hughes Medical Institute, Boston, MA 02115, USA
*Correspondence should be addressed to L.F. (Lucas_Farnung@hms.harvard.edu)

Transcription elongation by RNA polymerase II is a tightly regulated process that requires coordinated interactions between elongation factors. IWS1 (Interacts with SPT6) has been implicated as a core elongation factor, but its molecular role remains unclear. We show that the intrinsically disordered C-terminal region of IWS1 contains short linear motifs (SLiMs) that multivalently engage the elongation machinery. Using cryo-electron microscopy, we map SLiMs in IWS1 that interact with Pol II subunits RPB1, RPB2, and RPB5, as well as elongation factors DSIF, SPT6, and ELOF1. Functional assays demonstrate that distinct IWS1 SLiMs specify IWS1 recruitment and IWS1-dependent transcription stimulation. IWS1 recruitment to the transcription elongation complex depends on association via the RPB1 jaw and binding of downstream DNA, and transcription elongation stimulation requires interactions with the RPB2 lobe and ELOF1. Based on these findings, we identify other transcription elongation factors including ELOA and RECQL5 that bind the RPB1 jaw and demonstrate that IWS1 protects the activated transcription elongation complex from RECQL5 inhibition. We also reveal the binding of the histone reader and IWS1 interactor LEDGF to a transcribed downstream nucleosome. Our findings establish IWS1 as a modular scaffold that organizes the transcription elongation complex, illustrating how disordered regions regulate transcription elongation.

Introduction

Transcription elongation by RNA polymerase II (Pol II) requires the coordinated action of elongation factors that collectively maintain the enzyme in a processive, chromatin-compatible state. Rather than operating as isolated units, elongation factors form extensive interaction networks on Pol II. Despite structural and functional advances in elucidating the architecture of elongation complexes, the principles by which individual elongation factors assemble onto Pol II to form the productive transcription elongation machinery remain incompletely understood.

IWS1 (Interacts with SPT6) is a highly conserved Pol II–associated elongation factor required for normal gene expression from yeast to mammals (1-7). Genetic ablation of IWS1 in mice causes early embryonic lethality (8,9), and depletion of IWS1 in human cells leads to Pol II accumulation at promoters with loss of transcription across gene bodies (2), underscoring its essential role in productive elongation. Beyond transcription itself, IWS1 has been implicated in co-transcriptional RNA processing (10) and chromatin regulation (11-17).

Structurally, IWS1 contains a central helical bundle formed by a HEAT repeat and a TFIIS N-terminal domain (TND) that directly bind to the transcription elongation complex (6,18). The central HEAT/TND domain is flanked by long disordered N- and C-terminal regions (Figure 1C). The N-terminal region binds histones and contains well-characterized TND-interacting motifs (TIMs) that bind transcription elongation factors such as TFIIS, ELOA, and LEDGF (2,13,19). Current structural and functional knowledge of IWS1, however, is heavily skewed toward its structured HEAT/TND core (6,18,20-22) and the N-terminal intrinsically disordered region. In contrast, the C-terminal tail is highly conserved and known to interact with DNA (13), yet no structural elements or binding motifs have been assigned. The persistence of strong evolutionary conservation in the apparently disordered C-terminal tail suggests it may encode important short linear motifs (SLiMs) for protein interactions. IDRs often serve as interaction hubs through such SLiMs, and the unexplained conservation of the IWS1 C-terminus raises the possibility that critical, yet undiscovered, interactions reside in the C-terminal tail. Addressing this knowledge gap is essential for a complete understanding of how IWS1 functions as a core elongation factor.

Using functional assays and single-particle cryo-electron microscopy (cryo-EM), we show that IWS1 directly stimulates transcription elongation, and we identify that the far C-terminus of IWS1 is constitutively required for binding the elongation complex. We find that IWS1 is an integral component of activated transcription elongation complexes that interacts via discrete SLiMs in the IWS1 C-terminus with Pol II subunits RPB1, RPB2, and elongation factors DSIF, SPT6, and ELOF1. Structure-function analyses show that mutational disruption of individual IWS1 SLiMs impairs transcription elongation *in vitro*. This work establishes IWS1's C-terminal tail as a previously unrecognized scaffold that organizes the Pol II elongation complex, illustrating a general principle whereby conserved disordered regions of elongation factors orchestrate the assembly and processivity of the transcription elongation apparatus. Together, our findings

represent the most complete structure of a transcription elongation complex to date and highlight the broader significance of intrinsically disordered protein segments in the regulation of gene expression.

Results

IWS1 stimulates transcription elongation

To assess the role of IWS1 in transcription elongation, we first used a fully reconstituted transcription elongation system to perform RNA extension assays with porcine RNA polymerase II (Pol II) and purified human transcription elongation factors (DSIF, SPT6, PAF1 complex including RTF1: PAF1c, TFIIS, ELOF1, P-TEFb, and IWS1) on a nucleosomal DNA substrate (Figure S1 and S2A, Methods). We observed that addition of IWS1 increased the fraction of full-length transcripts in a concentration-dependent manner (Figures S2B-C). To precisely evaluate the contribution of IWS1, we next omitted the strong transcription elongation stimulator RTF1 from our assay. As expected, we observed substantially less transcription in the absence of RTF1 (Figures 1A-B). Strikingly, addition of full-length IWS1 in reactions lacking RTF1, however, led to a strong increase in extended RNA products. Thus, IWS1 directly stimulates transcription elongation (Figures 1A-B).

To further test the effects of IWS1 on transcription elongation complexes on diverse DNA substrates, we performed time-course assays on linear, mono-nucleosomal, and dinucleosomal DNA substrates. Indeed, IWS1 increased full-length transcripts 2-3.5–fold on all substrates even in the presence of RTF1 (Figure S2). Additionally, IWS1 facilitated polymerase passage through nucleosome entry barriers (Figure S2K). Together, these findings establish IWS1 as a transcription elongation factor that directly promotes RNA polymerase II elongation on naked DNA and chromatinized templates.

We next addressed whether the IWS1 N- and C-terminal regions are required for transcription stimulation. Deletion of the intrinsically disordered N-terminus (IWS1 ΔN, Figure 1C) only showed a minor reduction in IWS1's stimulatory effect on elongation (Figure 1D). In contrast, deletion of the intrinsically disordered but well-conserved IWS1 C-terminus (IWS1 ΔC, Figure 1C) fully abrogated the stimulatory effect on elongation (Figure 1D). Unexpectedly, these results suggested that the C-terminal tail of IWS1 is critical for promoting transcription elongation, whereas the histone binding N-terminal domain contributes only minimally.

To determine whether the loss of stimulation in the presence of IWS1 ΔC was due to impaired association with the elongation machinery, we performed size exclusion chromatography experiments with reconstituted elongation complexes (Methods). Due to the similar molecular weight of IWS1 ΔC and RTF1 (Figure S1A), we again omitted RTF1 from our reactions to unambiguously determine the presence of IWS1 and its truncations in the complex. We first confirmed that the activated elongation complex assembled with full-length IWS1 in the absence of RTF1 (Figures 1F, S3A-B). Whereas IWS1 ΔN co-eluted with the elongation

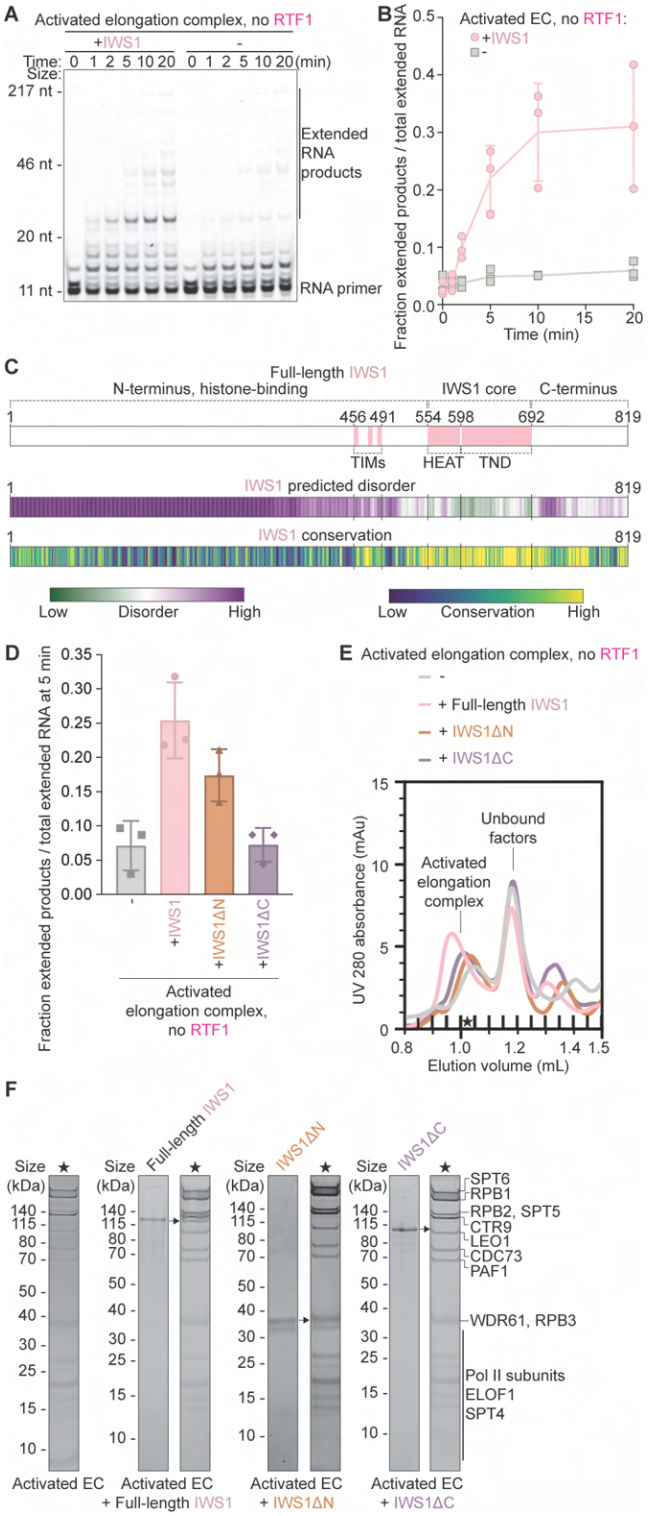

Figure 1. IWS1 stimulates transcription elongation and requires its C-terminus for elongation complex binding. (A) Time-course RNA extension assay without RTF1 and with or without IWS1 on linear DNA. Samples were separated by denaturing gel electrophoresis. RNA extension was monitored using FAM dye on RNA. Initial RNA and extended products are marked. Results are representative of at least three independent experiments. **(B)** Quantification of A from n=3 independent experiments. **(C)** Domain architecture of IWS1 with known elements labeled. Predicted disorder (IUpred score) and conservation (ConSurf score) for each residue indicated by color scheme. **(D)** Quantification of n=3 independent RNA extension assays without RTF1 quenched after 5 minutes after addition of CTP, GTP, and UTP (10 μM). Representative gel image shown in Figure 4D. **(E)** UV 280 nm size exclusion chromatogram of elongation complexes assembled without IWS1, with full-length IWS1, with IWS1ΔN (551-819), or with IWS1ΔC (1-699). RTF1 and TFIIS were excluded in complex formation. Peaks corresponding to activated elongation complex or unbound factors indicated.

(F) Representative SDS-PAGE of fraction marked with star from size exclusion complexes shown in E. Arrows between gels indicate expected migration of each IWS1 construct. Activated EC = Pol II, DSIF, SPT6, PAF1c, TFIIS, ELOF1. Individual data points are shown for B and D. Columns represent mean values across n = 3 independent experiments and error bars represent standard deviations in B and D.

complex similarly to full-length IWS1 (Figures 1F, S3C), IWS1 ΔC failed to co-elute with the complex (Figures 1F, S3D), suggesting a complete loss of stable association in the absence of the IWS1 C-terminus. Accordingly, an IWS1 truncation lacking the N- and C-termini (HEAT/TND only) also did not associate with the elongation complex (Figure S3E). We conclude that the C-terminal region, but not the N-terminus, is required for IWS1 binding to the transcription elongation complex and that the structured HEAT/TND domain alone is not sufficient to support IWS1 recruitment to the transcription elongation complex.

Cryo-EM structure of the activated transcription elongation complex with ELOF1 and IWS1

Because the intrinsically disordered C-terminal tail of IWS1 lacked structural domains but showed high levels of amino acid sequence conservation (Figure 1C), we asked whether it harbored short linear motifs (SLiMs) that interact with the elongation machinery. We prepared a cryo-EM sample containing Pol II, the kinase P-TEFb, ELOF1, IWS1, and human elongation factors DSIF, PAF1c, and SPT6 on a mononucleosomal substrate (Figure S4, Methods). Of note, we included LEDGF in the preparation because of its known interaction with IWS1 and nucleosomes (19,23-25). Pol II was positioned in front of the nucleosome and transcribed up to 4 bps in front of the nucleosome upon addition of NTPs (Figure S1C, S4). We performed size-exclusion chromatography to select for the formed complex containing IWS1, mildly cross-linked relevant fractions using glutaraldehyde, and prepared the sample for single-particle cryo-EM characterization (Figure S4B, Methods).

Analysis and classification of the cryo-EM data resulted in a reconstruction of the Pol II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1-LEDGF transcription elongation complex at a nominal resolution of 2.4 Å (FSC 0.143) from 762,523 particles (Figure 2B, Table 1, Video 1). The complex is bound to a partially unwrapped downstream nucleosome resolved at a lower resolution (see below). Further classification and masked refinements resulted in local reconstructions of different parts of the elongation complex at resolutions between 2.4 Å and 7.1 Å (Figures S5-S6, Methods). A pseudo-atomic model was built into the overall reconstruction with the aid of the locally refined cryo-EM maps and an extensive AlphaFold-Multimer prediction screen (Figure S5-S7, Methods, Table 2). The structure shows good stereochemistry (Figures S5-S6, Table 1).

Architecture of the activated transcription elongation complex with IWS1 and ELOF1

The resulting structure of the Pol II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1-LEDGF-nucleosome complex positions the polymerase 9 base pairs (bp) in front of the nucleosome (bp -9) with five nucleotides of backtracked RNA (Figure

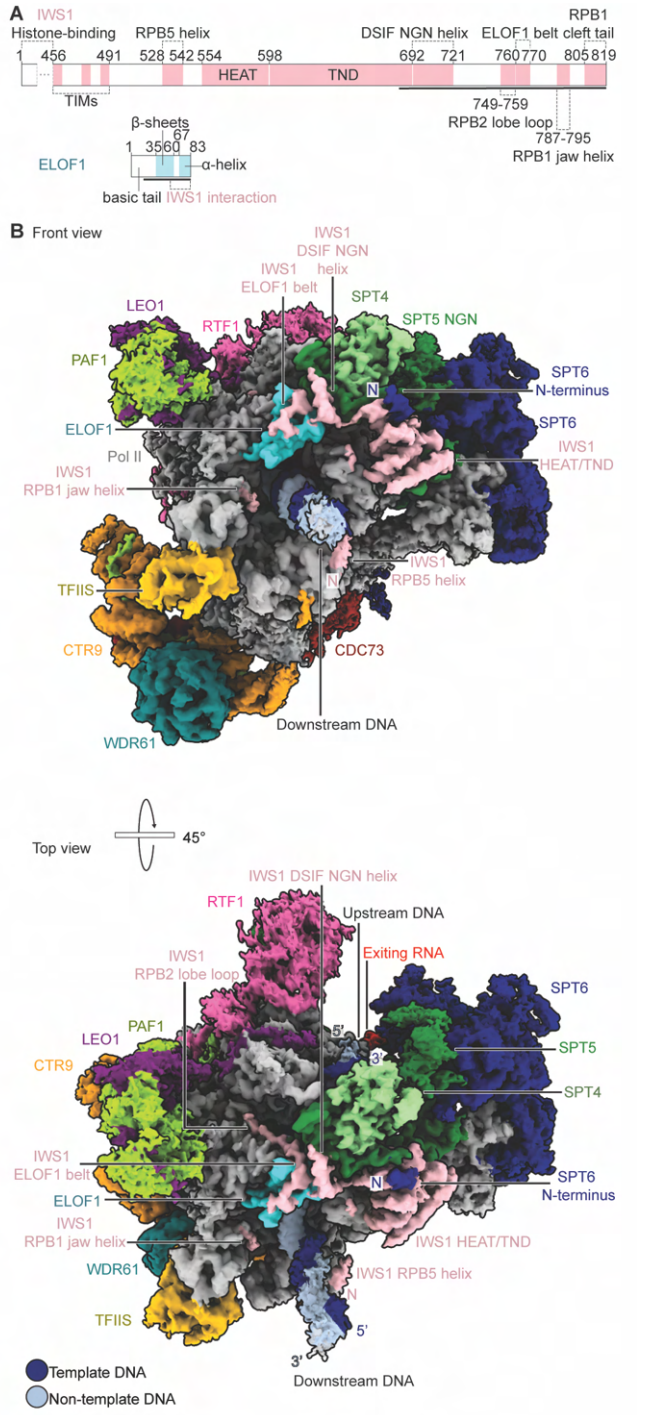

Figure 2. Structure of the RNA polymerase II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1 transcription elongation complex. (A) Domain architecture of IWS1 and ELOF1. Modelled regions are indicated as solid black lines. **(B)** Two views of the Pol II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1 transcription elongation complex on map T. Color code used throughout.

S8B-C). The complex had an architecture similar to mammalian elongation complex structures lacking ELOF1 and IWS1 (22,26-28). In addition to revealing four SLiM-like interactions confined to 10–20-residue long segments in the IWS1 C-terminal region with the elongation machinery (see below), it revealed the following new features. First, we observed a ~40 Å shift of the PAF1-LEO1 dimerization domain towards the RPB2 lobe (Figure S8M). This shift

likely reflects the conformational flexibility of the PAF1-LEO1 module previously observed across different structures (27), with the presence of IWS1 and ELOF1 potentially favoring a conformation proximal to the RPB2 lobe (Figure S8M). Second, we visualized a C-terminal helix of LEO1 (residues 528-561) next to the upstream DNA and SPT4 (Figure S8N). Third, we observed the CDC73 Ras-like domain bound to the RPB1 foot next to the SPT6 tSH2 domain (Figure S8O) (21). Fourth, our structure positions the zinc finger of TFIIS domain III within the Pol II funnel, likely favored by the backtracked configuration that we observed in the Pol II active site and funnel (Figure S8B,C,P). Finally, we visualized ELOF1. As observed in transcription-coupled nucleotide excision repair (TC-NER) complexes (29,30), ELOF1 binds the lobe of RPB2 (Figure S8Q). In our complex, ELOF1 is positioned next to the SPT5 NGN domain and contacts the IWS1 C-terminus (Figure S8Q, see below). We visualized additional residues of the ELOF1 N-terminus (residues 17-19) that stretch along the Pol II clamp towards the transcription bubble (Figure S8R). Similar to structural observations of Elf1 in yeast (21,31,32), ELOF1 is an integral component of the activated elongation complex in mammals. Together, this structure provides the most complete overview of the mammalian transcription elongation complex yet.

IWS1 structure and interactions with the activated elongation complex

In our structure, IWS1 interacted extensively with the transcription elongation machinery and stabilized transcription elongation factors including ELOF1 and SPT5. First, we observed a density positioned against RPB5 next to the downstream DNA (Figures 3C, S8S). Due to the absence of this helix in structures without IWS1 (27,28) and the high interaction score of IWS1 and RPB5 of an AlphaFold-Multimer prediction screen (Figure S7A), we docked the predicted IWS1 helix (residues 528-542) onto this density. Next, we fitted and refined the IWS1 HEAT and TND domains (residues 554-692). The HEAT/TND domains are positioned at the RPB1 clamp head above the SPT5 KOW2-3 domains as visualized in other activated transcription elongation complexes containing IWS1 or the yeast homolog Spn1 (Figures 3D, S8T) (21,22,33).

Similar to structural studies of Spn1 (18,20,21), we observed a helix from the SPT6 N-terminus (residues 207-230) that bound to the IWS1 TND (Figures 3D, S8T, S9B). Notably, casein kinase 2 was assumed to be important for mediating the SPT6-IWS1 interaction (34), but our complex formation that only used P-TEFb for phosphorylation indicates that casein kinase 2 is not constitutively required. Additionally, in proximity to the known SPT6-IWS1 interaction, we found an SPT5 helix (SPT5 residues 160-170) that contacts the IWS1 TND and the SPT6 N-terminus (Figures 3E, S8U). In the past (22), we likely did not observe these interactions between SPT6, SPT5, and IWS1 in the mammalian elongation complex due to lack of sufficient resolution or absence of ELOF1.

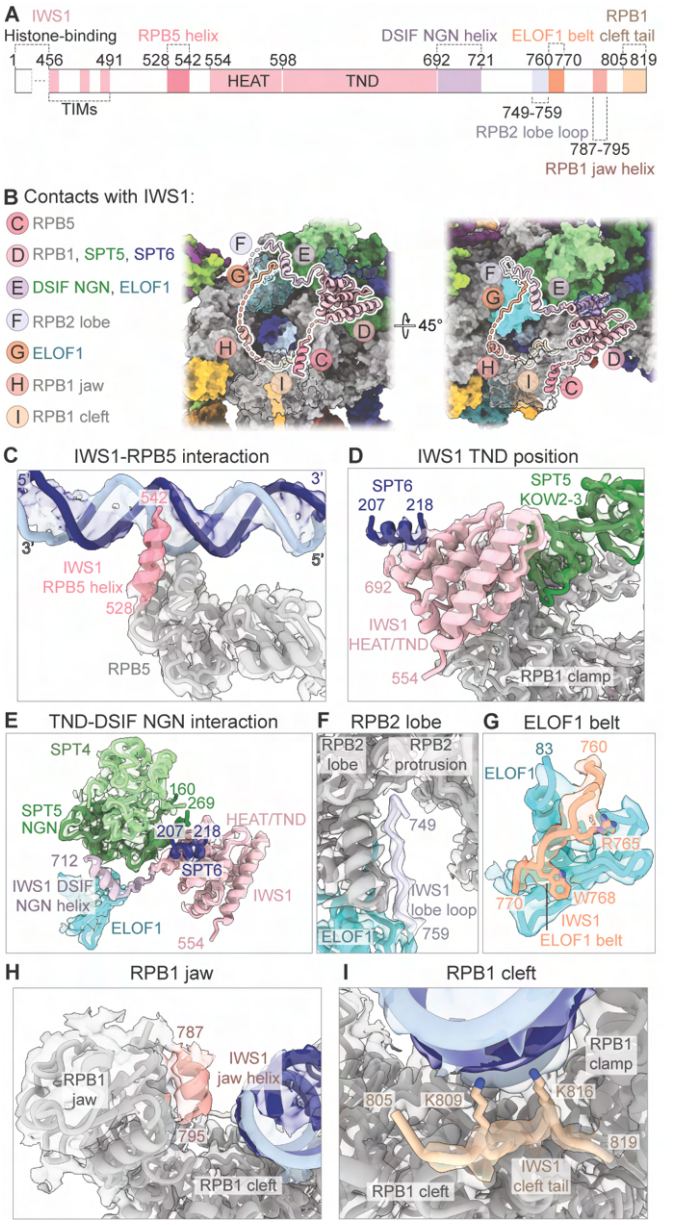

Figure 3. IWS1 interaction interfaces with RNA polymerase II and transcription elongation factors. (A) Domain architecture diagram of IWS1 with boundaries of newly modeled SLiMs indicated. Color code used throughout. **(B)** Schematic overview of IWS1 contacts with transcription elongation complex. **(C)** IWS1 interaction with RPB5 overlaid on map I. **(D)** Position of IWS1 HEAT/TND and interactions with SPT5 and SPT6 overlaid on map B. **(E)** IWS1 interaction with DSIF NGN and ELOF1 overlaid on map B. **(F)** IWS1 interaction with RPB2 lobe overlaid on map B. **(G)** IWS1 interaction with ELOF1 overlaid on map B. Residues of IWS1 ELOF1 belt with clear side-chain density shown as sticks. **(H)** IWS1 interaction with RPB1 jaw overlaid on map C. **(I)** IWS1 interaction with RPB1 cleft overlaid on map C. Positively charged residues shown as sticks.

Surprisingly, we observed multiple IWS1 SLiMs that positioned sequentially along IWS1 residues 693–819 interacting with (i) the SPT5 NGN domain, (ii) the RPB2 lobe, (iii) the elongation factor ELOF1, and (iv) the RPB1 jaw.

Residues of IWS1 (residues 693-721) immediately after the HEAT/TND domains and part of the IWS1 C-terminus (IWS1 "DSIF NGN helix") wrap around the DSIF NGN domain, mediating contacts between the DSIF NGN, IWS1, and ELOF1 (Figures 3E, S8U). IWS1 subsequently inserts

into the region between the RPB2 protrusion and the RPB2 lobe above the non-template strand (IWS1 "RPB2 lobe loop," residues 749-759, Figures 3F, S8V).

From the RPB2 lobe, IWS1 residues 760-770 wrap around the solvent-exposed surface of ELOF1 (Figures 3G, S8W). We named this segment the "ELOF1 belt" of IWS1. Our AlphaFold-Multimer predictions recapitulated these findings (Figure S9A). The extensive interaction surface of more than 2800 Å2 between IWS1, ELOF1, and DSIF may stabilize the binding of DSIF and ELOF1. Indeed, the SPT5 NGN domain and SPT4 are more stably anchored to the elongation complex compared to previous structures (27,35) (Figure S9C). This could also stabilize the observed interactions between the SPT5 and SPT6 N-terminal helices and the IWS1 HEAT/TND.

Finally, we fitted the most C-terminal helix of IWS1 (residues 787-795, "RPB1 jaw helix") into a density between the RPB1 jaw and the downstream DNA (Figures 3H, S8X). AlphaFold-Multimer also predicted that the very C-terminus of IWS1 could meander into a crevice formed between the downstream DNA, the RPB1 clamp, the RPB1 cleft, and RPB5 (Figure S7D). We were able to assign backbone and some sidechain densities of IWS1 residues 805-819 to this interaction and named it the IWS1 "RPB1 cleft tail" (Figures 3I, S8Y). These contacts between the extreme C-terminus of IWS1, RPB1, and RPB5 position positively charged residues of IWS1 (residues R787, K790, R793, K794, K809, K816) near the downstream DNA backbone (Figure S9D). Consequently, we observed that truncating the IWS1 C-terminus led to an overall reduction in IWS1 DNA binding using a fluorescence polarization assay (Methods, Figure S9E).

Importantly, the IWS1 C-terminal SLiM interactions are conserved (Figures S9F-I, S10A-D). AlphaFold-Multimer predictions of *S. cerevisiae* Rpb1-Spn1, Rpb2-Elf1-Spn1, Rpb5-Spn1, Spt4-Spt5-Spn1, and Spt6-Spn1 captured similar interactions involving the Spn1 C-terminus (Figure S9F-I). Notably, the DSIF NGN helix and RPB1 cleft tail are not conserved in our predicted Spn1 interactions (Figures S9F-I, S10).

In summary, we observe that IWS1, ELOF1, SPT6, and DSIF form an interconnected network, favored by highly specific interactions of the IWS1 C-terminal SLiMs that contribute to enhanced stability of the elongation complex, rationalizing why IWS1 is a core transcription elongation factor.

The extreme C-terminus of IWS1 mediates association with the elongation complex

To better understand which of the newly identified C-terminal interactions are required for IWS1 recruitment to the elongation complex, we generated a series of C-terminal IWS1 truncations (Figure 4A) that removed the entire C-terminus downstream of the DSIF NGN helix (IWS1 1-721), the ELOF1 belt and extreme C-terminus (IWS1 1-759), or only the extreme C-terminus containing the RPB1 jaw helix and the RPB1 cleft tail (IWS1 1-786). Using size exclusion chromatography, we then reconstituted elongation complexes with these truncations. We found that the larger C-terminal truncations (IWS1 1-721, IWS1 1-759) did not associate with the elongation complex, and even the most minimal truncation (IWS1 1-786) failed to associate with the elongation complex (Figures 4B, S11A-C), indicating that the RPB1 jaw helix and interaction with the downstream DNA in the Pol II cleft are required for IWS1 recruitment and association with the transcription elongation complex. As expected, RNA extension assays showed that none of the truncations stimulated transcription elongation as IWS1 recruitment via the RPB1 jaw helix and interaction with downstream DNA in the far C-terminus of IWS1 is required (Figure 4D-E).

Specific regions in the C-terminus of IWS1 stimulate transcription elongation

To further investigate how the other IWS1 SLiMs influence IWS1-elongation complex binding or transcription, we generated separate glycine-serine mutations of the RPB2 lobe loop (IWS1 residues 747-756) and the ELOF1 belt (IWS1 residues 760-771). We also generated a glycine-serine mutation of the IWS1 RPB5 interacting helix (IWS1 residues 524-550). These mutants stably associated with the elongation complex, consistent with our findings that the IWS1 RPB1 jaw helix and cleft tail interactions determine elongation complex association (Figures 4C, S11D-F). Despite stable elongation complex association, RNA extension levels of the RPB2 lobe loop mutant and the ELOF1 belt mutant were significantly compromised (Figure 4D-E). The RPB5 helix mutant, however, only showed a mild reduction in transcription stimulation compared to wildtype IWS1. These mutational studies of IWS1 show that the disruption of individual SLiM interactions within the IWS1 C-terminus are detrimental for efficient transcription elongation, even if IWS1 is stably bound to the elongation complex. Interaction of IWS1 with ELOF1 and the RPB2 lobe is required for direct stimulation of transcription elongation.

In addition to disrupting IWS1 SLiM interactions by mutating the SLiMs on IWS1, we identified specific residues on ELOF1 that interact with IWS1 and mutated these residues to alanine ($ELOF1_{beltmut}$) (Figure S12A-B). Similar to our results with the IWS1 ELOF1 belt GS mutant, reconstituted elongation complexes containing $ELOF1_{beltmut}$ showed binding of wildtype IWS1 (Figure S12C) but complexes containing $ELOF1_{beltmut}$ and wildtype IWS1 had reduced RNA extension activity compared to reactions containing wildtype ELOF1 (Figure S12D-F). Together, our size exclusion chromatography and RNA extension assays results demonstrate that recruitment of IWS1 to the transcription elongation complex and transcription stimulation by IWS1 involve interactions of distinct IWS1 regions that have non-overlapping function.

Multiple elongation factors bind an overlapping surface on the RPB1 jaw

The extreme C-terminus of IWS1 is important for IWS1 association with the elongation complex. Importantly, the

Figure 4. Roles of IWS1 C-terminal SLiMs in association with elongation complex and stimulation of transcription elongation. (A) Domain architecture of IWS1 and its mutants. **(B)** UV 280 nm size exclusion chromatograms and representative SDS-PAGE of elongation complexes assembled with full-length IWS1, IWS1 (1-786), IWS1 (1-759), or IWS1 (1-721). RTF1 and TFIIS were excluded in complex formation. Peaks corresponding to activated elongation complex or unbound factors indicated. Fraction loaded on SDS-PAGE indicated by star on chromatogram. Arrows between gels indicate expected migration of each IWS1 construct. **(C)** UV 280 nm size exclusion chromatograms and representative SDS-PAGE of elongation complexes assembled with full-length IWS1, IWS1-GlySer-lobe loop, IWS1-GlySer-belt, or IWS1-GlySer-RPB5 helix. RTF1 and TFIIS were excluded in complex formation. Peaks corresponding to activated elongation complex or unbound factors indicated. Fraction loaded on SDS-PAGE indicated by star on chromatogram. Arrows between gels indicate expected migration of each IWS1 construct. **(D)** RNA extension assay without RTF1 with IWS1 C-terminal truncations on linear DNA. Reactions were quenched 5 min after addition of CTP, GTP, and UTP (10 μM). Results are representative of at least three independent experiments. **(E)** Quantification of D from n = 3 independent experiments. Individual data points shown and error bars represent standard deviations. Columns represent mean values across n = 3 experiments. EC = Pol II, DSIF, SPT6, PAF1c, TFIIS, ELOF1.

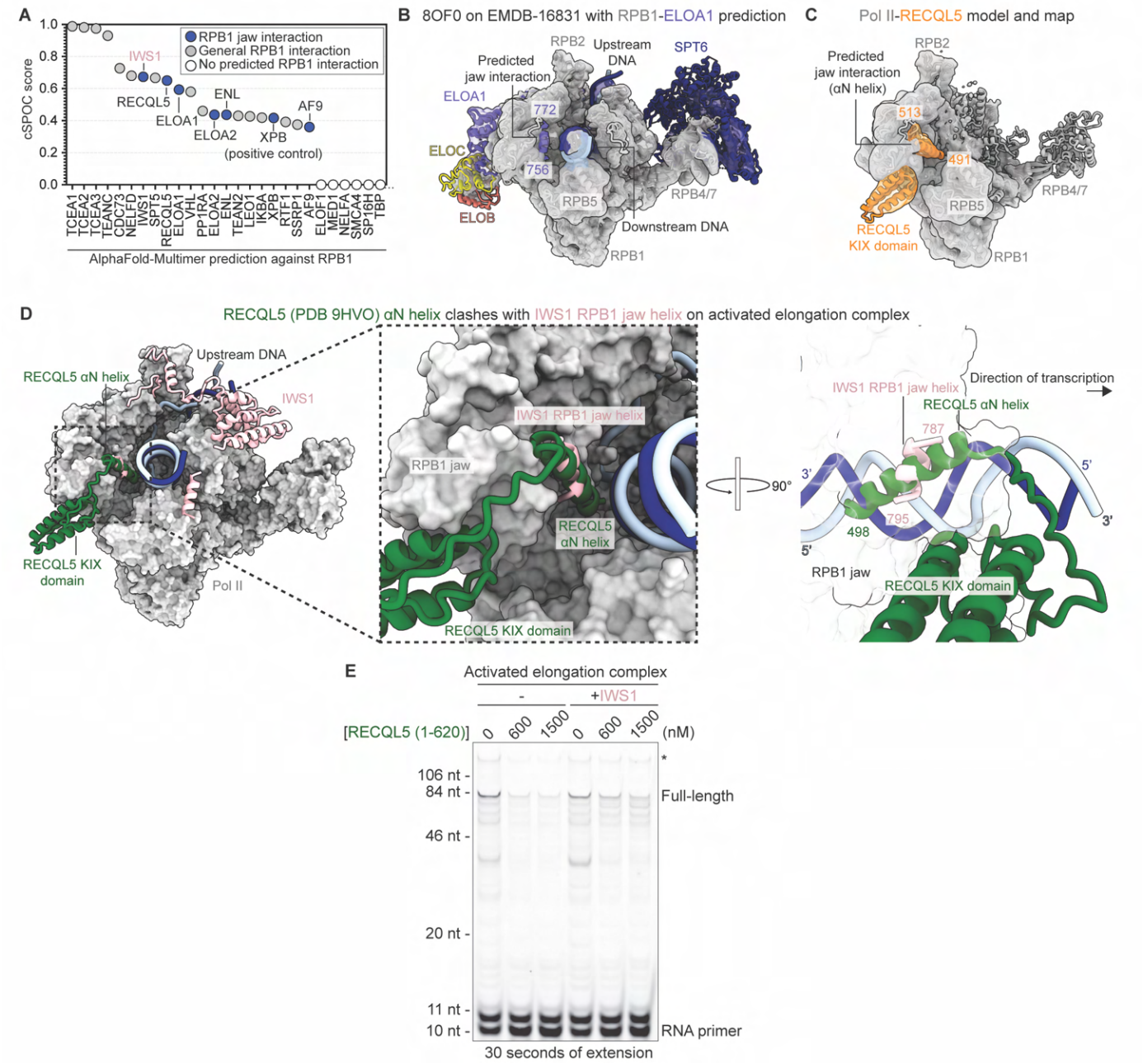

Figure 5. A protein-protein interaction screen identifies a surface of the RPB1 jaw as a binding platform for elongation factors. (A) Interaction screen using AlphaFold-Multimer. Predictions ranked by cSPOC score. Representative RPB1 jaw interactors are shown as dark blue circles, general RPB1 interactions shown as grey circles, and no predicted RPB1 interactions shown as white circles. **(B)** Fit of predicted helical element of ELOA into Pol II–ELOABC elongation complex model (PDB 8OF0) and Coulomb potential map (EMDB-16831) confirms interaction of ELOA C-terminal helix and RPB1 jaw. **(C)** Fit of RECQL5-RPB1 protein-protein interaction prediction into Pol II–RECQL5 complex Coulomb potential map (this study) confirms interaction of RECQL5 helix and RPB1 jaw. RECQL5 KIX domain is also resolved. **(D)** Overlay of RECQL5 αN helix and KIX domain from RECQL5-Pol II model (PDB 9HVO) on activated elongation complex model (this study) shows binding overlap of RECQL5 αN helix and IWS1 RPB1 jaw helix. Pol II model shown as surfaces and IWS1 model shown as cartoons. **(E)** RNA extension assay on linear substrate with increasing concentrations of RECQL5 (1-620) in the presence or absence of 300 nM IWS1. Reactions were quenched 30 min after addition of ATP, CTP, GTP, and UTP (100μM). Results are representative of two independent experiments. Activated elongation complex = Pol II, DSIF, SPT6, PAF1c including RTF1, ELOF1. Non-denatured RNA products marked with asterisk.

IWS1 RPB1 jaw binding site overlaps with the RPB1 jaw binding sites of TC-NER factor UVSSA (29,30), RNA polymerase II chaperone RPAP2 (36), and TFIIH subunit XPB (37)(Figure S13B). This suggests that the RPB1 jaw interaction could represent an important regulatory surface that is mutually exclusive to the binding of a wider variety of protein factors in initiation and elongation.

To identify additional RPB1 jaw binders, we performed an AlphaFold-Multimer pairwise interaction screen containing human RPB1 and key transcription initiation and elongation factors (2) (Figures 5A, S13A). We ranked the interactions by their SPOC score (Methods). Using this approach, we identified multiple high-confidence interactors of RPB1 (Figures 5A, S13A). We further categorized the high-

confidence interactors based on whether they are predicted to interact with the RPB1 jaw. These RPB1 jaw interactors included the transcription elongation factor Elongin A (ELOA1) and the transcription-related helicase RECQL5 (38-40)(Figure S13C). Specifically, a C-terminal helix in ELOA1 (residues 756-772) is predicted to bind the RPB1 jaw but has not been characterized. To structurally validate this prediction, we analyzed existing cryo-EM data of an Elongin ABC-containing transcription elongation complex where the C-terminal helix of ELOA1 was not modeled (41). Upon low-pass filtering the map (Methods), we observed a helical density positioned next to the RPB1 jaw which overlapped with the predicted binding site of the ELOA1 C-terminal helix (Figure 5B).

Our AlphaFold-Multimer screen also predicted that the αN "brake" helix of RECQL5 (residues 491-513) (38-40) binds to the RPB1 jaw (Figure S13C). We prepared a cryo-EM sample of a Pol II complex with RECQL5 (Figure S14A-F) and visualized the interaction between the RECQL5 αN brake helix and the RPB1 jaw, along with the RECQL5 KIX domain and the RPB1 jaw (Figure 5C). The interaction between the RECQL5 αN brake helix and the RPB1 jaw has also been observed in two recent publications at high-resolution (38,39).

Although RECQL5 was shown to bind to an elongation complex (39), its interaction with the RPB1 jaw would clash with the IWS1 binding site (Figures 5D, S14G). To test whether RECQL5 and IWS1 could compete for transcription elongation complex binding through their jaw domain interactions, we prepared RECQL5 (residues 1-620) (Figure S14H) to include in RNA extension assays (Figure S14I, Methods). We were able to recapitulate a transcriptional stalling effect (39,40) of RECQL5 lacking IWS1 (Figure 5E). Notably, the stalling effect was strongly reduced and RNA products reached their full length once IWS1 was included in the reaction, indicating that IWS1 and RECQL5 compete for binding to the transcription elongation complex via the RPB1 jaw helix.

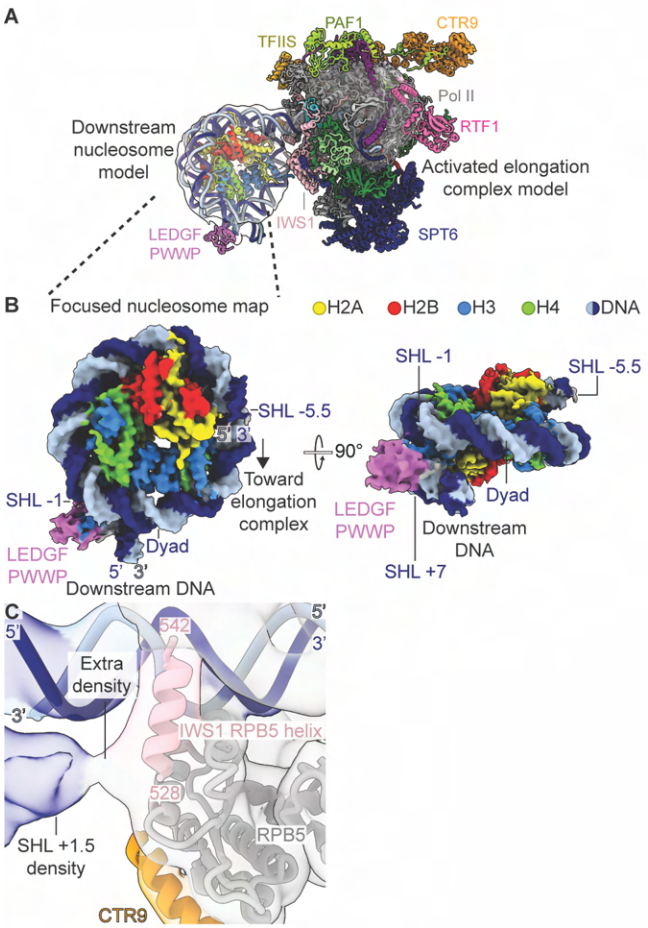

Figure 6. Interaction details of LEDGF PWWP and IWS1 with transcribed nucleosome. (A) Cryo-EM density of the nucleosome model positioned with respect to the activated elongation complex atomic model using map R. Nucleosome and activated elongation complex models shown as cartoons. **(B)** Cryo-EM density (map Q) of the nucleosome substrate. The transcribed nucleosome is unwrapped to SHL -5.5. LEDGF (orchid) binds the H3 tail at the promoter-distal side of the transcribed nucleosome and contacts both DNA gyres. **(C)** The IWS1 N-terminus reaches toward SHL +1.5 of the transcribed nucleosome after its interaction with downstream DNA and RPB5. SHL = super helical location.

LEDGF binds the distal side of the transcribed nucleosome

We also analyzed the cryo-EM densities corresponding to the nucleosomal substrate and associated factors (Figure 6). Using particle subtraction to keep the signal for the 10-subunit core Pol II and nucleosome (Figure S5, Methods), we obtained focused maps of the nucleosome (7.6 Å) and 10-subunit Pol II (8.7 Å). These maps allowed us to position the nucleosome model relative to the activated elongation complex (Figure 6A). We obtained a 3.5 Å map of the nucleosome using particle subtraction and focused refinements of the same particle set (Figure 6B). The nucleosomal DNA is unwrapped to super helical location (SHL) -5.5, exposing the promoter proximal H2A/H2B dimer (Figures 6B, S8B-D). We observed binding of the LEDGF PWWP domain only on the promoter distal side of the nucleosome because LEDGF requires interactions with two DNA gyres to engage the methylated H3 tail (Figure 6B) (24,25). The observed LEDGF-nucleosome interaction is consistent with previous observations of LEDGF binding to mononucleosomes (24,25). The IWS1 N-terminal region harbors histone chaperone activity and the IWS1 interaction with RPB5 positions it near the partially unwrapped nucleosome. Accordingly, we observed density of the IWS1 N-terminus projecting towards SHL +1.5 of the nucleosome (Figure 6C). This likely positions IWS1 to engage histones upon further unwrapping of nucleosomal DNA. We did not observe a direct interaction between LEDGF and IWS1, likely because this involves TIM3 of IWS1 (residues 482-491) and the TND of LEDGF (residues 342-426) (2,42), both of which were not visualized in our structure.

Discussion

Here we report the structure of the mammalian activated elongation complex in the presence of transcription elongation factors ELOF1 and IWS1. The IWS1 C-terminus makes extensive SLiM-like interactions with multiple core

components of the elongation complex, including Pol II subunits RPB1, RPB2, RPB5, downstream DNA, ELOF1, DSIF, and SPT6. This network of interactions may function to stabilize elongation factor binding and maintain Pol II in an elongation state. Accordingly, single-molecule studies of elongation factor binding dynamics show that yeast Spn1 (human IWS1), Spt4/5 (human DSIF), Spt6 (human SPT6), and Elf1 (human ELOF1) are stably associated with Pol II and have the longest residence times during transcription elongation (43,44). These similar binding dynamics may be enhanced by the stabilization of elongation factors on the polymerase by IWS1. Consequently, we demonstrate that IWS1 stimulates transcription elongation. In contrast to a recent publication (33), we demonstrate that the IWS1 RPB1 jaw interaction and binding of downstream DNA do not directly stimulate transcription but are critical interfaces for IWS1 association with the elongation complex. Interaction of IWS1 with the RPB2 lobe near the non-template strand of the transcription bubble and binding of ELOF1 are required for direct stimulation of transcription elongation. Our Spn1 AlphaFold-Multimer predictions (S9) and IWS1 C-terminal sequence alignments (S10) further reinforce these findings, as the RPB1 jaw helix, RPB2 lobe loop, and ELOF1 belt are well-conserved, but yeast Spn1 does not appear to contain a conserved interaction with the RPB1 cleft. Additionally, RTF1 and IWS1 appear to stimulate elongation in an additive manner that may suggest different mechanisms, consistent with genetic evidence in yeast (5,27) where a Spn1 HEAT/TND mutation (K192N) or the Spn1 HEAT/TND expressed on its own exacerbates phenotypes of RTF1 deletion.

Structural and *in silico* prediction data of RPB1 jaw interactors show that IWS1, UVSSA, RPAP2, XPB, RECQL5, ELOA1, and others bind an overlapping region on the RPB1 jaw. We show that a fully assembled activated elongation complex with IWS1 and ELOF1 protects the transcription elongation complex from inhibition by RECQL5. Together, these findings imply a regulatory mechanism where different elongation factors may compete for RPB1 jaw binding depending on the transcriptional state or cellular context. It is likely that RECQL5 association requires the partial disassembly of the activated elongation complex by the dissociation of at least IWS1, suggesting a more complex mechanism of factor exchange than previously proposed (39). Furthermore, the interaction stability of DSIF with the transcription machinery plays an important role in the initial recruitment of the TC-NER ATPase CSB to the transcription machinery during transcription-coupled nucleotide excision repair (45). Destabilization and displacement of the SPT5 NGN domain and SPT4 is a prerequisite for CSB binding. Since IWS1 and ELOF1 have a stabilizing effect on DSIF, IWS1 must also be displaced or rearranged during TC-NER. Accordingly, IWS1 is not compatible with the binding of multiple TC-NER factors including UVSSA, CSA, CSB, and STK19 (Figure S14A). Comparison of our activated elongation complex containing IWS1 and ELOF1 suggests that the positioning of IWS1 and ELOF1 are, however, compatible with the paused elongation complex (Figure S14B) (46,47).

Together, we provide the most complete model of the mammalian transcription elongation complex to date and show a modular, multivalent engagement strategy for IWS1 that integrates structural flexibility with specific molecular recognition, revealing a mechanistic principle of elongation factor function.

Limitations of the study

While we and others have shown that SPT6/Spt6 (22,48) and Spn1 N-termini (13) can bind to histones, the polymerase may need to unravel more nucleosomal DNA by transcribing further into the nucleosome for this binding to occur. The positioning of the IWS1 RPB5 interacting helix at the downstream DNA may allow the disordered IWS1 N-terminus to engage the exposed histones during transcription. How this occurs in concert with SPT6 and/or the histone chaperone FACT remains to be elucidated (15,21,22,48). Although our nucleosome substrate for our cryo-EM structure contains a T-less cassette up to nucleosome position bp -4, we observed that our complex is backtracked to bp -9 (Figure S8B-C). The stable wrapping of the DNA in our nucleosome density suggests that the nucleosome acts as a barrier to the elongation complex (35). It is possible that the polymerase adopts a backtracked state due to this barrier.

We also observed a major shift of the PAF1-LEO1 dimer on the Pol II surface. We were unable to directly attribute this architectural change to the presence of ELOF1 and IWS1, but we note that previous structures (26,27) also showed positional heterogeneity of the PAF1-LEO1 dimer. Allosteric changes introduced by IWS1 and ELOF1 could favor the observed positioning on top of RPB2.

References

1. Fischbeck, J.A., Kraemer, S.M. and Stargell, L.A. (2002) SPN1, a conserved gene identified by suppression of a postrecruitment-defective yeast TATA-binding protein mutant. Genetics, 162, 1605-1616.
2. Cermakova, K., Demeulemeester, J., Lux, V., Nedomova, M., Goldman, S.R., Smith, E.A., Srb, P., Hexnerova, R., Fabry, M., Madlikova, M. et al. (2021) A ubiquitous disordered protein interaction module orchestrates transcription elongation. Science, 374, 1113-1121.
3. Mayer, A., Lidschreiber, M., Siebert, M., Leike, K., Söding, J. and Cramer, P. (2010) Uniform transitions of the general RNA polymerase II transcription complex. Nat Struct Mol Biol, 17, 1272-1278.
4. Liu, Z., Zhou, Z., Chen, G. and Bao, S. (2007) A putative transcriptional elongation factor hIws1 is essential for mammalian cell proliferation. Biochemical and Biophysical Research Communications, 353, 47-53.
5. Zhang, L., Fletcher, A.G., Cheung, V., Winston, F. and Stargell, L.A. (2008) Spn1 regulates the recruitment of Spt6 and the Swi/Snf complex during transcriptional activation by RNA polymerase II. Mol Cell Biol, 28, 1393-1403.
6. Pujari, V., Radebaugh, C.A., Chodaparambil, J.V., Muthurajan, U.M., Almeida, A.R., Fischbeck, J.A., Luger, K. and Stargell, L.A. (2010) The transcription factor Spn1 regulates gene expression via a highly conserved novel structural motif. J Mol Biol, 404, 1-15.
7. Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R. and Hartzog, G.A. (2003) Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. Mol Cell Biol, 23, 1368-1378.
8. Orlacchio, A., Stark, A.E., Foray, C., Amari, F., Sheetz, T., Reese, E., Tessari, A., La Perle, K., Palmieri, D., Tsichlis, P.N. et al. (2018)

Genetic ablation of interacting with Spt6 (Iws1) causes early embryonic lethality. PLoS One, 13, e0201030.
9. Oqani, R.K., Lin, T., Lee, J.E., Kang, J.W., Shin, H.Y. and Il Jin, D. (2019) Iws1 and Spt6 Regulate Trimethylation of Histone H3 on Lysine 36 through Akt Signaling and are Essential for Mouse Embryonic Genome Activation. Sci Rep, 9, 3831.
10. Yoh, S.M., Cho, H., Pickle, L., Evans, R.M. and Jones, K.A. (2007) The Spt6 SH2 domain binds Ser2-P RNAPII to direct Iws1-dependent mRNA splicing and export. Genes Dev, 21, 160-174.
11. Reim, N.I., Chuang, J., Jain, D., Alver, B.H., Park, P.J. and Winston, F. (2020) The conserved elongation factor Spn1 is required for normal transcription, histone modifications, and splicing in Saccharomyces cerevisiae. Nucleic Acids Res, 48, 10241-10258.
12. Li, S., Almeida, A.R., Radebaugh, C.A., Zhang, L., Chen, X., Huang, L., Thurston, A.K., Kalashnikova, A.A., Hansen, J.C., Luger, K. et al. (2018) The elongation factor Spn1 is a multi-functional chromatin binding protein. Nucleic Acids Res, 46, 2321-2334.
13. Li, S., Edwards, G., Radebaugh, C.A., Luger, K. and Stargell, L.A. (2022) Spn1 and Its Dynamic Interactions with Spt6, Histones and Nucleosomes. Journal of Molecular Biology, 434, 167630.
14. Viktorovskaya, O., Chuang, J., Jain, D., Reim, N.I., López-Rivera, F., Murawska, M., Spatt, D., Churchman, L.S., Park, P.J. and Winston, F. (2021) Essential histone chaperones collaborate to regulate transcription and chromatin integrity. Genes Dev, 35, 698-712.
15. López-Rivera, F., Chuang, J., Spatt, D., Gopalakrishnan, R. and Winston, F. (2022) Suppressor mutations that make the essential transcription factor Spn1/Iws1 dispensable in Saccharomyces cerevisiae. Genetics, 222.
16. Tonsager, A.J., Zukowski, A., Radebaugh, C.A., Weirich, A., Stargell, L.A. and Ramachandran, S. (2025) The histone chaperone Spn1 preserves chromatin protections at promoters and nucleosome positioning in open reading frames. G3 (Bethesda), 15.
17. Chen, B., Dronamraju, R., Smith-Kinnaman, W.R., Peck Justice, S.A., Hepperla, A.J., MacAlpine, H.K., Simon, J.M., Mosley, A.L., MacAlpine, D.M. and Strahl, B.D. (2025) Spt6-Spn1 interaction is required for RNA Polymerase II association and precise nucleosome positioning along transcribed genes. Journal of Biological Chemistry, 108436.
18. Diebold, M.L., Koch, M., Loeliger, E., Cura, V., Winston, F., Cavarelli, J. and Romier, C. (2010) The structure of an Iws1/Spt6 complex reveals an interaction domain conserved in TFIIS, Elongin A and Med26. Embo j, 29, 3979-3991.
19. Tesina, P., Čermáková, K., Hořejší, M., Procházková, K., Fábry, M., Sharma, S., Christ, F., Demeulemeester, J., Debyser, Z., Rijck, J. et al. (2015) Multiple cellular proteins interact with LEDGF/p75 through a conserved unstructured consensus motif. Nat Commun, 6, 7968.
20. McDonald, S.M., Close, D., Xin, H., Formosa, T. and Hill, C.P. (2010) Structure and biological importance of the Spn1-Spt6 interaction, and its regulatory role in nucleosome binding. Mol Cell, 40, 725-735.
21. Ehara, H., Kujirai, T., Shirouzu, M., Kurumizaka, H. and Sekine, S.-i. (2022) Structural basis of nucleosome disassembly and reassembly by RNAPII elongation complex with FACT. Science (American Association for the Advancement of Science), 377, 1169-eabp9466.
22. Markert, J.W., Soffers, J.H. and Farnung, L. (2025) Structural basis of H3K36 trimethylation by SETD2 during chromatin transcription. Science, 387, 528-533.
23. Gérard, A., Ségéral, E., Naughtin, M., Abdouni, A., Charmeteau, B., Cheynier, R., Rain, J.-C. and Emiliani, S. (2015) The Integrase Cofactor LEDGF/p75 Associates with Iws1 and Spt6 for Postintegration Silencing of HIV-1 Gene Expression in Latently Infected Cells. Cell host & microbe, 17, 107-117.
24. Wang, H., Farnung, L., Dienemann, C. and Cramer, P. (2020) Structure of H3K36-methylated nucleosome–PWWP complex reveals multivalent cross-gyre binding. Nature Structural & Molecular Biology, 27, 8-13.
25. Koutná, E., Lux, V., Kouba, T., Škerlová, J., Nováček, J., Srb, P., Hexnerová, R., Šváchová, H., Kukačka, Z., Novák, P. et al. (2023) Multivalency of nucleosome recognition by LEDGF. Nucleic Acids Research, gkad674.
26. Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H. and Cramer, P. (2018) Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature (London), 560, 607-612.
27. Vos, S.M., Farnung, L., Linden, A., Urlaub, H. and Cramer, P. (2020) Structure of complete Pol II-DSIF-PAF-SPT6 transcription complex reveals RTF1 allosteric activation. Nat Struct Mol Biol, 27, 668-677.
28. Filipovski, M., Soffers, J.H.M., Vos, S.M. and Farnung, L. (2022) Structural basis of nucleosome retention during transcription elongation. Science, 376, 1313-1316.
29. Kokic, G., Yakoub, G., van den Heuvel, D., Wondergem, A.P., van der Meer, P.J., van der Weegen, Y., Chernev, A., Fianu, I., Fokkens, T.J., Lorenz, S. et al. (2024) Structural basis for RNA polymerase II ubiquitylation and inactivation in transcription-coupled repair. Nat Struct Mol Biol.
30. Mevissen, T.E.T., Kümmecke, M., Schmid, E.W., Farnung, L. and Walter, J.C. (2024) STK19 positions TFIIH for cell-free transcription-coupled DNA repair. Cell.
31. Ehara, H., Kujirai, T., Fujino, Y., Shirouzu, M., Kurumizaka, H. and Sekine, S.-i. (2019) Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. Science, 363, 744-747.
32. Akatsu, M., Ehara, H., Kujirai, T., Fujita, R., Ito, T., Osumi, K., Ogasawara, M., Takizawa, Y., Sekine, S.I. and Kurumizaka, H. (2023) Cryo-EM structures of RNA polymerase II-nucleosome complexes rewrapping transcribed DNA. J Biol Chem, 299, 105477.
33. Zheenbekova, A., Walshe, J.L., Ochmann, M., Bäuerle, M., Neef, U., Maier, K.C., Rus, P., Yan, Y., Urlaub, H., Cramer, P. et al. (2025) IWS1 positions downstream DNA to globally stimulate Pol II elongation. Nat Commun, 16, 7747.
34. Krogan, N.J., Kim, M., Ahn, S.H., Zhong, G., Kobor, M.S., Cagney, G., Emili, A., Shilatifard, A., Buratowski, S. and Greenblatt, J.F. (2002) RNA polymerase II elongation factors of Saccharomyces cerevisiae: a targeted proteomics approach. Mol Cell Biol, 22, 6979-6992.
35. Farnung, L., Ochmann, M., Garg, G., Vos, S.M. and Cramer, P. (2022) Structure of a backtracked hexasomal intermediate of nucleosome transcription. Mol Cell, 82, 3126-3134.e3127.
36. Wang, X., Qi, Y., Wang, Z., Wang, L., Song, A., Tao, B., Li, J., Zhao, D., Zhang, H., Jin, Q. et al. (2022) RPAP2 regulates a transcription initiation checkpoint by inhibiting assembly of pre-initiation complex. Cell Rep, 39, 110732.
37. Schilbach, S., Aibara, S., Dienemann, C., Grabbe, F. and Cramer, P. (2021) Structure of RNA polymerase II pre-initiation complex at 2.9 Å defines initial DNA opening. Cell, 184, 4064-4072.e4028.
38. Florez Ariza, A.J., Lue, N.Z., Grob, P., Kaeser, B., Fang, J., Kassube, S.A. and Nogales, E. (2025) Structural insights into transcriptional regulation by the helicase RECQL5. Nat Struct Mol Biol.
39. Zhang, L., Gordiyenko, Y., Morgan, T., Franco, C., Tufegdžić Vidaković, A. and Zhang, S. (2025) Structural basis of RECQL5-induced RNA polymerase II transcription braking and subsequent reactivation. Nat Struct Mol Biol.
40. Kassube, S.A., Jinek, M., Fang, J., Tsutakawa, S. and Nogales, E. (2013) Structural mimicry in transcription regulation of human RNA polymerase II by the DNA helicase RECQL5. Nat Struct Mol Biol, 20, 892-899.
41. Chen, Y., Kokic, G., Dienemann, C., Dybkov, O., Urlaub, H. and Cramer, P. (2023) Structure of the transcribing RNA polymerase II-Elongin complex. Nat Struct Mol Biol, 30, 1925-1935.
42. Sharma, S., Čermáková, K., De Rijck, J., Demeulemeester, J., Fábry, M., El Ashkar, S., Van Belle, S., Lepšík, M., Tesina, P., Duchoslav, V. et al. (2018) Affinity switching of the LEDGF/p75 IBD interactome is governed by kinase-dependent phosphorylation. Proceedings of the National Academy of Sciences, 115, E7053-E7062.
43. Ling, Y.H., Liang, C., Wang, S. and Wu, C. (2024) Live-cell single-molecule dynamics of eukaryotic RNA polymerase machineries. bioRxiv, 2024.2007.2008.602478.
44. Joo, Y.J., Ficarro, S.B., Chun, Y., Marto, J.A. and Buratowski, S. (2019) *In vitro* analysis of RNA polymerase II elongation complex dynamics. Genes Dev, 33, 578-589.
45. Kokic, G., Wagner, F.R., Chernev, A., Urlaub, H. and Cramer, P. (2021) Structural basis of human transcription–DNA repair coupling. Nature, 598, 368-372.
46. Su, B.G. and Vos, S.M. (2024) Distinct negative elongation factor conformations regulate RNA polymerase II promoter-proximal pausing. Mol Cell, 84, 1243-1256.e1245.

47. Vos, S.M., Farnung, L., Urlaub, H. and Cramer, P. (2018) Structure of paused transcription complex Pol II-DSIF-NELF. Nature, 560, 601-606.
48. Warner, J.L., Lux, V., Veverka, V. and Winston, F. (2024) The histone chaperone Spt6 controls chromatin structure through its conserved N-terminal domain. bioRxiv, 2024.2011.2025.625227.
49. Markert, J.W., Vos, S.M. and Farnung, L. (2023) Structure of the complete Saccharomyces cerevisiae Rpd3S-nucleosome complex. Nat Commun, 14, 8128.
50. Farnung, L., Vos, S.M., Wigge, C. and Cramer, P. (2017) Nucleosome-Chd1 structure and implications for chromatin remodelling. Nature, 550, 539-542.
51. Simon, M.D., Chu, F., Racki, L.R., de la Cruz, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J. and Shokat, K.M. (2007) The site-specific installation of methyl-lysine analogs into recombinant histones. Cell, 128, 1003-1012.
52. Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods, 14, 290-296.
53. Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E. and Scheres, S.H. (2018) New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife, 7.
54. Tegunov, D., Xue, L., Dienemann, C., Cramer, P. and Mahamid, J. (2021) Multi-particle cryo-EM refinement with M visualizes ribosome-antibiotic complex at 3.5 Å in cells. Nature Methods, 18, 186-193.
55. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H. and Ferrin, T.E. (2023) UCSF ChimeraX: Tools for structure building and analysis. Protein Sci, 32, e4792.
56. Croll, T.I. (2018) ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr D Struct Biol, 74, 519-530.
57. PyMOL. The PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.
58. Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr, 66, 213-221.
59. Pei, J., Kim, B.H. and Grishin, N.V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res, 36, 2295-2300.
60. Consortium, U. (2025) UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res, 53, D609-D617.
61. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A. et al. (2021) Highly accurate protein structure prediction with AlphaFold. Nature, 596, 583-589.
62. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A. et al. (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Research, 50, D439-D444.
63. Bryant, P., Pozzati, G. and Elofsson, A. (2022) Improved prediction of protein-protein interactions using AlphaFold2. Nature Communications, 13, 1265.
64. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B.J., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., Zhang, F. et al. (2021) The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. Protein Sci, 30, 187-200.
65. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J. et al. (2022) Protein complex prediction with AlphaFold-Multimer. bioRxiv, 2021.2010.2004.463034.
66. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S. and Steinegger, M. (2022) ColabFold: making protein folding accessible to all. Nature Methods, 19, 679-682.
67. Schmid, E.W. and Walter, J.C. (2025) Predictomes, a classifier-curated database of AlphaFold-modeled protein-protein interactions. Molecular Cell, 85, 1216-1232.e1215.
68. Steinegger, M. and Söding, J. (2017) MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nature Biotechnology, 35, 1026-1028.
69. Lim, Y., Tamayo-Orrego, L., Schmid, E., Tarnauskaite, Z., Kochenova, O.V., Gruar, R., Muramatsu, S., Lynch, L., Schlie, A.V., Carroll, P.L. et al. (2023) *In silico* protein interaction screening uncovers DONSON's role in replication initiation. Science, 381, eadi3448.
70. Okonechnikov, K., Golosova, O., Fursov, M. and team, U. (2012) Unipro UGENE: a unified bioinformatics toolkit. Bioinformatics, 28, 1166-1167.

Acknowledgements

We thank all members of the Farnung lab for discussions. We thank J. W. Markert for support with data processing. We thank The Harvard Cryo-EM Center for Structural Biology at Harvard Medical School for support with data collection. We thank S.M. Vos for critical reading and input. L.F. is supported by The Smith Family Awards Program for Excellence in Biomedical Research, the Damon Runyon Rachleff Innovator Award Program, and the NIH Director's New Innovator Award (DP2-ES036404). J.C.W. was supported by NIH grant HL098316. He is an American Cancer Society Research Professor and a member of the Howard Hughes Medical Institute. K. A. was supported by NIH grant R01GM139960. F. S. and E.W.S. were supported by the National Science Foundation (DGE 2140743). This paper was typeset with the bioRxiv word template by @Chrelli: www.github.com/chrelli/bioRxiv-word-template

Author contributions

D.S. conducted all biochemical experiments unless stated otherwise. D.S. prepared the IWS1 complex for cryo-EM. D.S. collected and analyzed IWS1 cryo-EM data. F.S. prepared RECQL5, the RECQL5-RNA polymerase II complex, and conducted cryo-EM analysis of RECQL5. S.R. prepared the RECQL5 (residues 1-620) protein. D.S. and S.R. purified RNA polymerase II. D.S., J.C.W, K.A., and L.F. conceived the transcription elongation *in silico* prediction screen. E.S. conducted the *in silico* prediction screen and provided analysis. L.F. designed and supervised research. D.S. and L.F. wrote the manuscript with input from all authors.

Competing interest statement

J.C.W. is a co-founder of MOMA Therapeutics, in which he has a financial interest. K.A. is a consultant to Syros Pharmaceuticals and Odyssey Therapeutics, is on the SAB of CAMP4 Therapeutics, and received research funding from Novartis not related to this work. Readers are welcome to comment on the online version of the paper.

Data availability

Correspondence and requests should be addressed to Lucas Farnung (Lucas_Farnung@hms.harvard.edu). Materials are available from Lucas Farnung upon request and a completed materials transfer agreement with Harvard University.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT in order to revise text for clarity and conciseness. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Quantification and statistical analysis

No statistical methods were used to predetermine sample size. The experiments were not randomized, and the investigators were not blinded to allocation during experiments and outcome assessment.

Materials and Methods

Cell culture

Sf9 cells (Expression Systems, Cat# 94-001S), *T.ni* (Hi5) cells (Expression Systems, Cat# 94-002S), and Sf21 cells (Expression Systems, Cat# 94-003S) were cultured in ESF 921 insect cell culture medium (Expression Systems, Cat# 96-001-01) at 27°C for baculovirus production and protein expression, unless otherwise indicated. *E. coli* DH5α (New England Biolabs (NEB) Cat# C2987H), DH10EMBacY (Geneva Biotech), and BL21-CodonPlus (DE3)-RIL (Agilent Cat# 230245) were cultured in LB broth (EMD-Millipore Cat# 71753-5) at 37°C for plasmid production and protein expression, unless otherwise indicated.

Cloning

The following sequence containing the TspRI cut site and the modified Widom 601 sequence with an A-less cassette to nucleosomal bp -4 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 AAT TTT CGA TGT ATA TAT CTG ACA CGT GCC TGG AGA CTA GGG AGT AAT CCC CTT GGC GGT TAA AAC GCG 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 ATA TCG CGC GTG ATC TTA CGG CAT TAT ACG TA-3' (-4 vector) using the same approaches as previously described (28). The bp +64 vector and bp +115 vector were generated as previously described (22,28).

The dinucleosome vector was generated by assembling the reverse Widom 601 sequence and the Widom 601 sequence with a 15 bp linker between the two sequences: 5'-GAT CAG ACT CGA CTA-3' into a pIDTSmart-Kan vector (IDT).

The H3K36C histone H3 variant was generated by introducing the lysine-to-cysteine mutation to a vector encoding X. laevis H3 C110A using 'Round-The-Horn cloning (49).

Plasmid constructs for *H. sapiens* DSIF, *H. sapiens* SPT6, *H. sapiens* PAF1 complex, *H. sapiens* RTF1, *H. sapiens* TFIIS, *H. sapiens* IWS1, *H. sapiens* LEDGF, and *H. sapiens* PTEF-b were previously generated (22,24,28).

A gene-optimized *H. sapiens* ELOF1 gBlock was synthesized and cloned into a 1-B vector using ligation-independent cloning (LIC). Similarly, a gene-optimized *H. sapiens* ELOF1 gBlock containing E55L, E79L, and N82A mutations was synthesized and cloned into a 1-B vector using LIC.

IWS1 (1-786), IWS1 (1-759), IWS1 (1-721), IWS1ΔC (1-699), IWS1ΔN (551-819), and IWS1 HEAT/TND (551-699) were generated using 'Round-The-Horn cloning. IWS1-GlySer-loop (747-756), IWS1-GlySer-belt (760-771), and IWS1-GlySer-RPB5helix (524-550) were generated by amplifying the IWS1 438-C vector backbone followed by Gibson assembly with gBlocks (IDT) encoding GlyGlyGlyGlySer repeats of equivalent length to the segment of IWS1 being replaced.

RECQL5 was amplified from cDNA and cloned into the 438-C vector using LIC. RECQL5 (1-620) was amplified from the full-length RECQL5 vector and cloned into the 1-C vector using LIC.

Protein expression

H. sapiens DSIF, *H. sapiens* TFIIS, and X. laevis histones were expressed in *E. coli*. *H. sapiens* SPT6, *H. sapiens* PAF1 complex, *H. sapiens* RTF1, *H. sapiens* IWS1 and mutants, *H. sapiens* LEDGF, and *H. sapiens* PTEF-b were expressed in Hi5 insect cells.

H. sapiens ELOF1 and ELOF1beltmut were expressed in BL21 (DE3) RIL cells at 37 °C until an OD of 0.6 at 600 nm was reached. Expression was induced with the addition of 1 mM IPTG and cells were grown at 18 °C overnight. Cells were centrifuged and pellets were resuspended in lysis buffer (20 mM Na·HEPES pH 7.4, 300 mM NaCl, 30 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM b-mercaptoethanol, 0.284 µg mL–1 leupeptin, 1.37 µg mL–1 pepstatin A, 0.17 mg mL–1 PMSF, and 0.33 mg mL–1 benzamidine), then snap-frozen in liquid nitrogen and stored at -80 °C until purification.

Protein purification

S. scrofa RNA polymerase II, *H. sapiens* DSIF, *H. sapiens* SPT6, *H. sapiens* RTF1, *H. sapiens* TFIIS, *H. sapiens* IWS1, *H. sapiens* LEDGF, and *H. sapiens* PTEF-b were purified as previously described (22,24,26,27,47). X. laevis histones were purified as previously described (50).

H. sapiens ELOF1 with an N-terminal 6xHis tag and TEV protease cleavage site was purified at 4°C. Cells expressing ELOF1 were resuspended in lysis buffer and subsequently lysed by sonication. The lysate was centrifuged and cleared by ultra-centrifugation. The supernatant containing ELOF1 was subsequently filtered using 5 mm and 0.45 mm syringe filters. The filtered supernatant was applied to a HisTrap HP 5 mL (Cytiva). The column was subsequently washed with 5 CV lysis buffer, 5 CV high salt buffer (20 mM Na·HEPES pH 7.4, 1000 mM NaCl, 30 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM b-mercaptoethanol), and 5 CV lysis buffer. ELOF1 was eluted from the HisTrap column using a 10 CV 0-100% gradient of nickel elution buffer (20 mM Na·HEPES pH 7.4, 300 mM NaCl, 500 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM b-mercaptoethanol). The elution was fractionated and analyzed using SDS-PAGE. Fractions containing ELOF1 were pooled and applied to 3.5k MWCO SnakeSkin dialysis tubing (Thermo Scientific) in lysis buffer at 4°C overnight. 1.5 mg of TEV protease was added to the sample prior to dialysis to remove the N-terminal His6 tag. The dialyzed sample was applied to a HisTrap HP 5 mL, pre-equilibrated in lysis buffer. The flow-through containing ELOF1 was collected and subsequently concentrated using an Amicon 3,000 MWCO centrifugal filter unit (Millipore). The concentrated sample was applied to a HiLoad 16/600 Superdex 75 pg (Cytiva), equilibrated in gel filtration buffer (20 mM Na·HEPES pH 7.4, 300 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP). The elution was fractionated and analyzed by SDS-PAGE. Sample containing ELOF1 was concentrated using an Amicon 3,000 MWCO centrifugal filter unit (Millipore).

H. sapiens PAF1 complex containing untagged PAF1, LEO1, CDC73, WDR61, and CTR9 with a C-terminal MBP, 6xHis tag, and TEV protease cleavage site was purified at 4°C. Cells expressing the PAF1 complex were resuspended in lysis buffer and lysed by douncing using a B-type douncer. The lysate was centrifuged and cleared by ultra-centrifugation. The supernatant containing the PAF1 complex was subsequently filtered using 5 mm and 0.45 mm syringe filters. The filtered supernatant was applied to a HisTrap HP 5 mL (Cytiva). The column was subsequently washed with 9 CV lysis buffer. A self-packed XK column (Cytiva) with 15 mL of Amylose resin (New England Biolabs) was attached to the HisTrap column. PAF1 complex was eluted from the HisTrap column using nickel elution buffer. The HisTrap column was removed, and the amylose column was washed with 5 CV lysis buffer. PAF1 complex was eluted from the amylose column with 5 CV amylose elution buffer (20 mM Na·HEPES pH 7.4, 300 mM NaCl, 30 mM imidazole pH 8.0, 116.9 mM maltose, 10% (v/v) glycerol, 5 mM b-mercaptoethanol). The elution was fractionated and analyzed using SDS-PAGE. Fractions containing PAF1 complex were pooled and applied to 3.5k MWCO SnakeSkin dialysis tubing (Thermo Scientific) in dialysis buffer (20 mM Na·HEPES pH 7.4, 150 mM NaCl, 30 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM b-mercaptoethanol) at 4°C overnight. 1.5 mg of TEV protease was added to the sample prior to dialysis to remove the C-terminal MBP-His6 tag. The dialyzed sample was applied to a HiTrap Q HP 5 mL (Cytiva), pre-equilibrated in dialysis buffer. The column was subsequently washed with 10 CV dialysis buffer. PAF1 complex was eluted from the HiTrap Q column with a 10 CV 0-100% gradient of high salt buffer (20 mM Na·HEPES pH 7.4, 1000 mM NaCl, 30 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM b-mercaptoethanol). The elution was fractionated and analyzed by SDS-PAGE. Sample containing the PAF1 complex was collected and subsequently concentrated using an Amicon 30,000 MWCO centrifugal filter unit (Millipore). The concentrated sample was applied to a Superose 6 Increase 10/300 GL (Cytiva), equilibrated in gel filtration buffer. The elution was fractionated and analyzed by SDS-PAGE. Sample containing the PAF1 complex was concentrated using an Amicon 30,000 MWCO centrifugal filter unit (Millipore).

H. sapiens IWS1 and IWS1 constructs (1-786), (1-759), (1-721), ΔC (1-699), GlySer-lobe loop, GlySer-belt, and GlySer-RPB5 helix were purified as described for full-length IWS1 (22).

H. sapiens IWS1 HEAT/TND (551-699) with an N-terminal 6xHis-MBP tag and TEV protease cleavage site was purified at 4°C. Cells expressing IWS1 HEAT/TND were resuspended in lysis buffer and subsequently lysed by sonication. The lysate was centrifuged and cleared by ultra-centrifugation. The supernatant containing IWS1 HEAT/TND was subsequently filtered using 5 mm and 0.45 mm syringe filters. The filtered supernatant was applied to a HisTrap HP 5 mL (Cytiva). The column was subsequently washed with 9 CV lysis buffer. A self-packed XK column (Cytiva) with 15 mL of Amylose resin (New England Biolabs) was attached to the HisTrap column. IWS1 HEAT/TND was eluted from the HisTrap column using nickel elution buffer. The HisTrap column was removed, and the amylose column was washed with 5 CV lysis buffer. IWS1 HEAT/TND was eluted from the amylose column with 5 CV amylose elution buffer (20 mM Na·HEPES pH 7.4, 300 mM NaCl, 30 mM imidazole pH 8.0, 116.9 mM maltose, 10% (v/v) glycerol, 5 mM b-mercaptoethanol). The elution was fractionated and analyzed using SDS-PAGE. Fractions containing IWS1 HEAT/TND were pooled and applied to 3.5k MWCO SnakeSkin dialysis tubing (Thermo Scientific) in lysis buffer at 4°C overnight. 1.5 mg of TEV protease with an N-terminal MBP tag was added to the sample prior to dialysis to remove the N-terminal His6-MBP tag. The dialyzed sample was applied to a HisTrap HP 5 mL, pre-equilibrated in lysis buffer. The flow-through containing IWS1 HEAT/TND was collected and subsequently applied to a self-packed XK column (Cytiva) with 15 mL of Amylose resin

(New England Biolabs) pre-equilibrated in lysis buffer. The flow-through was concentrated using an Amicon 3,000 MWCO centrifugal filter unit (Millipore). The concentrated sample was applied to a Superose 6 Increase 10/300 GL (Cytiva), equilibrated in gel filtration buffer (20 mM Na·HEPES pH 7.4, 300 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP). The elution was fractionated and analyzed by SDS-PAGE. Sample containing IWS1 HEAT/TND was concentrated using an Amicon 3,000 MWCO centrifugal filter unit (Millipore).

H. sapiens IWS1 ΔN (551-819) with an N-terminal 6xHis-MBP tag and TEV protease cleavage site was purified at 4°C. Cells expressing IWS1 ΔN were resuspended in lysis buffer and subsequently lysed by sonication. The lysate was centrifuged and cleared by ultra-centrifugation. The supernatant containing IWS1 ΔN was subsequently filtered using 5 mm and 0.45 mm syringe filters. The filtered supernatant was applied to a HisTrap HP 5 mL (Cytiva). The column was subsequently washed with 9 CV lysis buffer. A self-packed XK column (Cytiva) with 15 mL of Amylose resin (New England Biolabs) was attached to the HisTrap column. IWS1 ΔN was eluted from the HisTrap column using nickel elution buffer. The HisTrap column was removed, and the amylose column was washed with 10 CV low salt buffer (20 mM Na·HEPES pH 7.4, 150 mM NaCl, 30 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM b-mercaptoethanol). A HiTrap Q HP 5 mL (Cytiva) column was attached to the amylose column, and IWS1 ΔN was eluted from the amylose column with 15 CV low salt amylose elution buffer (20 mM Na·HEPES pH 7.4, 150 mM NaCl, 30 mM imidazole pH 8.0, 116.9 mM maltose, 10% (v/v) glycerol, 5 mM b-mercaptoethanol). The amylose column was removed, and the HiTrap Q HP column was eluted with a 10 CV gradient of high salt buffer (20 mM Na·HEPES pH 7.4, 1000 mM NaCl, 30 mM imidazole pH 8.0, 10% (v/v) glycerol, 5 mM b-mercaptoethanol). The elution was fractionated and analyzed using SDS-PAGE. Fractions containing IWS1 ΔN were pooled and applied to 3.5k MWCO SnakeSkin dialysis tubing (Thermo Scientific) in low salt buffer at 4°C overnight. 1.5 mg of TEV protease with an N-terminal MBP tag was added to the sample prior to dialysis to remove the N-terminal His6-MBP tag. The dialyzed sample was applied to a 15 mL amylose column with a HiTrap SP HP 5 mL (Cytiva) column attached in tandem. The columns were washed with 8 CV low salt buffer. The amylose column was detached, then the SP column was eluted with a 12 CV gradient of high salt buffer. The elution was fractionated and analyzed using SDS-PAGE. Fractions containing IWS1 ΔN were pooled and concentrated using an Amicon 10,000 MWCO centrifugal filter unit (Millipore). The concentrated sample was applied to a Superdex 200 Increase 10/300 (Cytiva), equilibrated in gel filtration buffer (20 mM Na·HEPES pH 7.4, 300 mM NaCl, 10% (v/v) glycerol, 1 mM TCEP). The elution was fractionated and analyzed by SDS-PAGE. Sample containing IWS1 ΔN was concentrated using an Amicon 10,000 MWCO centrifugal filter unit (Millipore).

All concentrated samples were subsequently aliquoted, flash-frozen, and stored at -80°C prior to future use.

RECQL5 expression and purification

Hi5 cells (600 mL) grown in ESF-921 media (Expression Systems) were infected with 300 μL of V1 virus for protein expression. The cells were grown for 48–72 h at 27 °C. Cells were harvested by centrifugation (238g, 4 °C, 30 min) and resuspended in lysis buffer. The cell resuspension was snap-frozen and stored at -80 °C.

Protein purification was performed at 4 °C. Frozen cell pellets were thawed and lysed by sonication. Lysates were cleared by centrifugation (18,000g, 4 °C, 30 min) and ultracentrifugation (235,000g, 4 °C, 60 min). The supernatant containing RECQL5 was filtered using 0.8-μm syringe filters (Millipore) and applied onto a HisTrap HP 5 ml (Cytiva), pre-equilibrated in lysis buffer. After sample application, the column was washed with 10 CV lysis buffer, 5 CV high salt buffer (1 M NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT, 30 mM imidazole pH 8.0, 0.284 μg mL−1 leupeptin, 1.37 μg mL−1 pepstatin A, 0.17 mg mL−1 PMSF, 0.33 mg mL−1 benzamidine), and 5 CV lysis buffer. The protein was eluted with a gradient of 0–100% elution buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT, 500 mM imidazole pH 8.0, 0.284 μg mL−1 leupeptin, 1.37 μg mL−1 pepstatin A, 0.17 mg mL−1 PMSF, 0.33 mg mL−1 benzamidine). Peak fractions were pooled and dialyzed for 16 h against 600 mL dialysis buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT, 30 mM imidazole) in the presence of 2 mg His6-TEV protease. The dialyzed sample was applied to a HisTrap HP 5 ml (Cytiva). The flow-through containing RECQL5 was concentrated using an Amicon 50,000 MWCO centrifugal filter unit (Millipore) and applied to a Superdex 200 16/600 pg (Cytiva) size exclusion column, pre-equilibrated in gel filtration buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT). Peak fractions were concentrated to ~100 μM, aliquoted, flash frozen, and stored at -80 °C.

RECQL5 (1-620) expression and purification

H. sapiens RECQL5 (1-620) was expressed in BL21 (DE3) RIL cells at 37 °C until an OD of 0.6 at 600 nm was reached. Expression was induced with the addition of 1 mM IPTG and cells were grown at 18 °C overnight. Cells were centrifuged and pellets were resuspended in lysis buffer, then snap-frozen in liquid nitrogen and stored at -80 °C until purification.

Protein purification was performed at 4 °C. Frozen cell pellets were thawed and lysed by sonication. Lysates were cleared by centrifugation (47,800g, 4 °C, 30 min). The supernatant containing RECQL5 (1-620) was filtered using 0.8-μm syringe filters (Millipore) and applied onto a HisTrap HP 5 mL (Cytiva), pre-equilibrated in lysis buffer. After sample application, the column was washed with 6 CV lysis buffer, 5 CV high salt buffer (1 M NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT, 30 mM imidazole pH 8.0, 0.284 μg mL−1 leupeptin, 1.37 μg mL−1 pepstatin A, 0.17 mg mL−1 PMSF, 0.33 mg mL−1 benzamidine), and 5 CV lysis buffer. The protein was eluted with a 13 CV gradient of 0–100% elution buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT, 500 mM imidazole pH 8.0, 0.284 μg mL−1 leupeptin, 1.37 μg mL−1 pepstatin A, 0.17 mg mL−1 PMSF, 0.33 mg mL−1 benzamidine). Peak fractions were pooled and dialyzed for 16 h against 500 mL dialysis buffer (150 mM NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT, 30 mM imidazole) in the presence of 2 mg His6-TEV protease. The dialyzed sample was applied to a HisTrap HP 5 mL (Cytiva) with HiTrap Heparin HP 5 mL (Cytiva) connected in tandem. The columns were washed with 6 CV of dialysis buffer. The HisTrap HP 5 mL was removed, and the HiTrap Heparin HP 5 mL was eluted with a 13 CV gradient of 0–100% high salt buffer. The elution was fractionated and analyzed by SDS-PAGE, and fractions containing RECQL5 (1-620) were concentrated using an Amicon 30,000 MWCO centrifugal filter unit (Millipore) and applied to a Superdex 200 Increase 10/300 (Cytiva) column, pre-equilibrated in gel filtration buffer (300 mM NaCl, 20 mM Na·HEPES pH 7.4, 10% (v/v) glycerol, 1 mM DTT). The elution was fractionated and analyzed by SDS-PAGE. Sample containing RECQL5 (1-620) was concentrated using an Amicon 30,000 MWCO centrifugal filter unit (Millipore), aliquoted, flash frozen, and stored at -80 °C.

H3K36cme3 octamer formation

Alkylation was used to install a trimethyl lysine analog on purified H3 containing the K36C mutation as previously described (51). Octamers with H3K36Cme3 were formed as previously described (28).

Nucleosome preparation

The bp -4, bp +64, bp +115, and dinucleosome DNA vectors were used as templates for a large-scale PCR reaction as previously described (28). The PCR products were pooled, then purified using a HiTrap Q HP 5 mL (Cytiva) 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, and digested with TspRI (NEB) as previously described (28). An equal volume of phenol:chloroform:isoamyl alcohol 25:24:1 (Sigma-Aldrich) was added to TspRI-digested DNA, then the DNA was applied to a MaXtract High Density (Qiagen) 1.5 mL tube and centrifuged to separate the two phases. The aqueous phase containing DNA was further extracted using an equal volume of chloroform, then applied to another MaXtract High Density tube and centrifuged again. TspRI-digested DNA was purified from the aqueous phase using ethanol precipitation.

Nucleosomes containing H3K36Cme3-modified octamers and the TspRI-digested -4 DNA were reconstituted as described (50). Similarly, nucleosomes containing unmodified octamers and the TspRI-digested +115 DNA were reconstitued. Histone octamer and DNA were mixed at a 1:1 molar ratio in RB-High Buffer (20 mM Na·HEPES, pH 7.4, 2 M KCl, 1mM 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 Na·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 3 hours at 4 ºC. The sample was centrifuged for 5 min at 21,000 x 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 30K 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. Dinucleosomes containing unmodified octamers and the TspRI-digested dinucleosome DNA were reconstitued similarly, except octamer and DNA were mixed at a 2:1 molar ratio.

In vitro RNA extension assay

All concentrations refer to final reaction conditions. 120 nM of RNA (IDT) (5'-/6-FAM/-UUAUCACUGUC-3') was incubated on ice with 120 nM of DNA (+64 linear DNA, +115 nucleosome, or dinucleosome) with a 3' TspRI-digested overhang complementary to the RNA sequence. The RNA-DNA hybrid was incubated with 150 nM of Pol II on ice. A factor mix of 300 nM IWS1, 300 nM SPT6, 300 nM DSIF, 250 nM P-TEFb, 300 nM ELOF1, 300 nM PAF1c, and 300 nM RTF1 (if included) was added to the Pol II-RNA-DNA mixture. Buffer was added to adjust the final reaction conditions to 20 mM Na·HEPES pH 7.4, 65 mM NaCl, 10% glycerol, and 1 mM TCEP. 3 mM MgCl2 and 0.5 mM ATP was added to the sample, then the sample was incubated at 30° C for 30 minutes. Transcription was started after the addition of 90 nM TFIIS, 0.01 mM CTP, 0.01 mM GTP, and 0.01 mM UTP. 3 μL reactions were quenched with 3 μL 2x STOP buffer (6.4 M urea, 50 mM EDTA, 2x TBE) at various timepoints (+64 linear DNA with RTF1: 0 sec, 10 sec, 30 sec, 45 sec, 60 sec, 90 sec. +115 nucleosome with RTF1: 0 min, 0.5 min, 1 min, 1.5 min, 2 min, 4 min. Dinucleosome with RTF1: 0 min, 1 min, 2 min, 3 min, 4 min, 6 min. +64 linear DNA without RTF1: 0 min, 1 min, 2 min, 5 min, 10 min, 20 min.). Samples were incubated with 30 mg Proteinase K (NEB) at 37° C for 25 minutes, then treated with 36 μL formamide (molecular biology grade, Millipore) at 95° C for 5 minutes. Samples for the +64 linear DNA or +115 nucleosome substrates were loaded onto 20% denaturing TBE-urea gels (8 M urea, 1x TBE, 10% formamide, 20% Bis-Tris acrylamide 19:1) and run with 1x TBE at 300 V for 90 minutes. Samples for the dinucleosome substrate were loaded onto 5% denaturing TBE-urea gels (8 M urea, 1x TBE, 10% formamide, 5% Bis-Tris acrylamide 19:1) and run with 1x TBE at 300 V for 20 minutes. The gels were imaged on an Amersham Typhoon (GE) using the 6-FAM fluorescence at 650 PMT.

Gel images were quantified using GelAnalyzer (GelAnalyzer 23.1.1 (available at www.gelanalyzer.com) by Istvan Lazar Jr., PhD and Istvan Lazar Sr., PhD, CSc). Original scanned TIFFs were opened in the GelAnalyzer application and bands were recognized as dark spots on a light background. Lanes were detected automatically with a threshold of 3, tilted, and equal width. Manual adjustment of lane boundaries was performed if lanes were curved and bands were not contained within the automatically defined lanes. Baseline was set with the rolling ball method and 10% peak width tolerance, which was applied to all lanes on the same gel image. Total extended RNA raw volume was measured for each lane by setting pixel boundaries in the region between the top of the primer band (bottom) and the edge of the well (top). Full-length product or strong band raw volume was measured for each lane by setting pixel boundaries in the region between the bottom of the full-length band (bottom) and the edge of the well (top). The full-length/strong band raw volume value was divided by the total extended RNA raw volume value for each lane, then imported into GraphPad Prism for mean and standard deviation calculations as well as plotting.

Barrier indices for nucleosome entry sites were calculated as previously described (35), except using GelAnalyzer. Raw volume values were taken for the first nucleosome entry barrier band and second nucleosome entry barrier band setting pixel boundaries around each band. Raw volume values for the upper bands were taken by setting pixel boundaries in the region between the top of the barrier band (bottom) and the edge of the well (top). The raw volume values were used for the barrier index calculation, defined as barrier index = $\text{Intensity}_{\text{barrier band}}$ / ($\text{Intensity}_{\text{barrier band}}$ + $\text{Intensity}_{\text{upper band}}$) (35). Calculated barrier indices were imported into GraphPad Prism for mean and standard deviation calculations as well as plotting.

RECQL5-IWS1 RNA extension assay

All concentrations refer to final reaction conditions. Template DNA (IDT) (5'-TCT GCT GAG TAG CTC ATG TCG TAC GTC GTC AGC TCA TAC GAT GCT CTC AGC TCA CTG CTA GCT ACG AGT GGG TTC TGC CGC TCT CAA TGG-3') was combined with RNA primer (IDT) (5'-/6-FAM/-AGAACCCACU-3') in 20 mM Na·HEPES pH 7.5, 100 mM NaCl, 10% glycerol and 3 mM MgCl2 and annealed by heating at 95 ºC for 5 minutes, then subsequently stepping down to 30 ºC at a rate of -2 ºC/minute. 120 nM of the RNA-template hybrid was incubated with 150 nM of Pol II at 30 ºC. 150 nM of nontemplate DNA (IDT) (5'- CCA TTG AGA GCG GCA GAA CCC ACT CGT AGC TAG CAG TGA GCT GAG AGC ATC GTA TGA GCT GAC GAC GTA CGA CAT GAG CTA CTC AGC AGA-3') was added to the RNA-DNA-Pol II mix and incubated at 30 ºC for 10 minutes. A factor mix of 300 nM SPT6, 300 nM DSIF, 250 nM P-TEFb, 300 nM ELOF1, 300 nM PAF1c, and 300 nM RTF1 was added to the Pol II-RNA-DNA mixture. Buffer was added to adjust the final reaction conditions to 20 mM Na·HEPES pH 7.4, 60 mM NaCl, 10% glycerol, and 1 mM TCEP. 3 mM MgCl2 and 1 mM ATP was added to the sample, then the sample was incubated at 30° C for 30 minutes. The Pol II-factor mix was divided and 300 nM of IWS1 or buffer was added at 30° C for 10 minutes. The mix was further divided and RECQL5 (1-620) at 0 nM (buffer), 600 nM, or 1500 nM was added and incubated at 30° C for 10 minutes. Transcription was started after the addition of 0.1 mM ATP, 0.1 mM CTP, 0.1 mM GTP, and 0.1 mM UTP. 3 μL reactions were quenched with 3 μL 2x STOP buffer (6.4 M urea, 50 mM EDTA, 2x TBE) after 30 seconds. Samples were incubated with 30 mg Proteinase K (NEB) at 37° C for 25 minutes, then treated with 36 μL formamide (molecular biology grade, Millipore) at 95° C for 5 minutes. Samples were loaded onto 20% denaturing TBE-urea gels (8 M urea, 1x TBE, 10% formamide, 20% Bis-Tris acrylamide 19:1) and run with 1x TBE at 300 V for 90 minutes. The gels were imaged on an Amersham Typhoon (GE) using the 6-FAM fluorescence at 650 PMT.

Complex formation for size exclusion chromatography

All concentrations refer to final reaction conditions. 120 nM of RNA (IDT) (5'-/6-FAM/-UUAUCACUGUC-3') was incubated on ice with 120 nM of the +64 linear DNA substrate with a 3' TspRI-digested overhang complementary to the RNA sequence. The RNA-DNA hybrid was incubated with 150 nM of Pol II on ice. A factor mix of 600 nM IWS1, 300 nM SPT6, 300 nM DSIF, 250 nM P-TEFb, 300 nM ELOF1, and 300 nM PAF1c was added to the Pol II-RNA-DNA mixture. Buffer was added to adjust the final reaction conditions to 20 mM Na·HEPES pH 7.4, 65 mM NaCl, 10% glycerol, 3 mM MgCl2, and 1 mM TCEP. 1 mM ATP was added to the sample, then the sample was incubated at 30° C for 30 minutes. The sample was centrifuged for 5 minutes at 13,000 x g and applied to a Superose 6 Increase 3.2/300 (Cytiva) column equilibrated in 20 mM Na·HEPES pH 7.4, 65 mM NaCl, 5% glycerol, 3 mM MgCl2, and 1 mM TCEP. Sample was detected at both 260 and 280 nm on an Äkta pure 25 with MicroKit. Corresponding fractions were analyzed for protein content using NuPAGE 4-12% Bis-Tris (Invitrogen) run in 1X MES buffer, stained with SyproRuby (Invitrogen), and detected using a SyproRuby fluorescence scan on an Amersham Typhoon (GE).

Fluorescence anisotropy assay

A FAM-labeled 45mer double-stranded DNA was synthesized (IDT): 5'-/6-FAM- GCC GCG TAT AGG GTC CAT CAG AAT TCG GAT GAA CTC GGT GTG AAG-3'. All concentrations refer to final reaction conditions. 20 microliter reactions were set up on ice containing 10 nM FAM-labeled 45mer double-stranded DNA, wild-type IWS1 or IWS1ΔC at 0, 10, 100, 1000, 2000, 3000, 4000, 5000, 6000, 7000, or 8000 nM, 20 mM Na·HEPES pH 7.4, 50 mM NaCl, 4% glycerol, 1 mM TCEP, 3 mM MgCl2, and 10 μg/mL BSA. 18 microliters of each reaction were transferred to a medium-binding 384-well microplate (Greiner 784076), then incubated on ice for 20 minutes. Plates were subsequently imaged on a Tecan Spark Multimode Microplate Reader using 470 nm excitation, 518 emission, gain high, and 100 flashes/read. A blank buffer well with no DNA or IWS1 was used for normalizing Z-height. A DNA only well was used as a blank.

RNA polymerase II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1-LEDGF-nucleosome complex cryo-EM sample preparation

All concentrations refer to final reaction conditions. RNA (5'-/6-FAM/-UUAUCACUGUC-3') (1400 nM) and the H3K36Cme3 bp -4 nucleosome (1000 nM) were incubated for 5 minutes on ice. Pol II (700 nM) was further added for 5 minutes on ice. DSIF (1400 nM), SPT6 (1400 nM), PAF1c (1512 nM), RTF1 (2100 nM), P-TEFb (700 nM), IWS1 (3500 nM), ELOF1 (3500 nM), LEDGF (1750 nM), and MgCl2 were mixed and dialyzed into 400 mL

of transcription buffer (50 mM NaCl, 20 mM Na·HEPES pH 7.4, 3 mM MgCl2, 4% (v/v) glycerol and 1 mM TCEP pH 8) at 4° C for 2 hours. Next, 3'-dATP (1 mM) was added for 45 minutes at 30° C to initiate phosphorylation by P-TEFb. Transcription was initiated by addition of CTP, GTP, UTP (1 mM each), and TFIIS (420 nM) at 30° C for 30 minutes. The sample was centrifuged for 5 minutes at 21,300 x g and applied to a Superose 6 Increase 3.2/300 (Cytiva) column equilibrated in transcription buffer. Sample was detected at both 260 and 280 nm on an Äkta pure 25 with MicroKit. Corresponding fractions were analyzed for RNA extension products and protein content.

For RNA extension analysis, 5 µL of each fraction was mixed with 2x STOP buffer (6.4 M urea, 50 mM EDTA, 2X TBE). 40 µg proteinase K (NEB) was added to each sample and incubated at 37° C for 30 minutes. 49.6 µL of formamide (molecular biology grade, Millipore) was added to each sample, then denatured at 95° C for 15 minutes. 5 µL samples were separated by a 12% denaturing urea gel (8 M urea, 1X TBE, 12% acrylamide:bis-acrylamide, 10% formamide) and RNA product was detected using a 6-FAM fluorescence scan on a Amersham Typhoon (GE) with 500 PMT. Proteins were detected using a NuPAGE 4-12% Bis-Tris (Invitrogen) ran in 1X MES buffer and stained with One-Step Blue Protein Gel Stain (Biotium).

Fractions containing complex were crosslinked with 0.1 % (v/v) glutaraldehyde for 10 minutes on ice followed by quenching with 2.4 mM aspartate and 2 mM lysine for 10 minutes on ice. Samples were dialyzed for three hours in cryo-EM buffer (20 mM Na·HEPES pH 7.4, 1 mM TCEP pH 8, 50 mM NaCl, and 3 mM MgCl2). Samples were centrifuged for 5 minutes at 21,300 x g following dialysis.

Complex concentration was determined by measuring absorbance at 280 nm. The molar extinction coefficient of the complex was obtained by summing the molar extinction coefficients of all protein and nucleic acid components at 280 nm.

UltrAuFoil R 2/2 on 200 mesh gold grids were glow-discharged for 120 s at 30 mA using a Pelco Easiglow plasma discharge system. 4 µL of dialyzed sample was applied to grids for 8 s, blotted for 3 s with a blot force of 8, and vitrified by plunging into liquid ethane using a Vitrobot Mark IV (FEI) at 4° C and 100 % humidity.

RNA polymerase II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1-LEDGF-nucleosome complex cryo-EM data collection and image processing

Cryo-EM data were collected on a ThermoFisher Scientific Titan Krios operated at 300 keV equipped with Falcon4i direct electron detector and a Selectris energy filter. Data collection was automated using EPU software. A single dataset was collected at a pixel size of 1.19 Å with a defocus range of -1.8 to -0.6 µm. The dataset yielded 30,316 micrographs. The dataset was collected with 48 movie frames at an exposure time of 4.21 s with an electron flux of 14.04 e- Å-2 s-1 for a total exposure of 37 e- Å-2.

Initial image processing was conducted in cryoSPARC (52).Movies were aligned using cryoSPARC Live Patch Motion Correction followed by CTF estimation. The cryoSPARC Blob picker was used to pick 4,597,368 particles. Particles were extracted at a box size of 460 pixels and downsampled to 400 pixels. Initial junk particles were removed, and then the remaining 1,691,625 particles that revealed clear density for Pol II and elongation factors were subjected to 3D classification using a mask encompassing the IWS1 HEAT/TND, DSIF NGN, and ELOF1. From this 3D classification, a class of 777,108 particles containing clear density for the IWS1 HEAT/TND was selected for downstream processing steps.

The 777,108 particles were re-extracted with a box size of 460 pixels and subjected to homogeneous refinement. Local CTF correction, global CTF correction, removal of duplicate particles, and reference motion correction were then followed by another round of homogeneous refinement. The corresponding 2.4 Å map (map A) revealed high resolution features for Pol II, the IWS1 HEAT/TND, DSIF NGN, and ELOF1, but low-resolution features for all other components. To visualize the other components, the 762,523 particles from map A were subjected to various rounds of classification and refinement. All mask generation, local refinements, and 3D classifications were performed in cryoSPARC unless stated otherwise.

Iterative 3D classifications using a mask encompassing the IWS1 HEAT/TND, ELOF1, and DSIF NGN domains ("IWS1+ELOF1 mask") was performed to enrich for the IWS1 HEAT/TND, ELOF1, and DSIF NGN. After performing 3D classification, 378,540 particles were subjected to non-uniform refinement followed by local refinement using the focus mask to produce a IWS1+ELOF1 map of 2.6 Å (map B). A similar processing workflow was used for the IWS1 C-term map ("IWS1 C-term mask," 2.4 Å, 156,410 particles, map C), TFIIS map ("TFIIS mask," 4.3 Å, 36,198 particles, map D), DSIF map ("DSIF mask," 2.7 Å, 79,487 particles, map E), RTF1 map ("RTF1 mask," 2.8 Å, 27,130 particles, map F), 10-subunit Pol II map (mask encompassing Pol II excluding RPB4/7 "10-subunit Pol II mask," 2.5 Å, 115,139 particles, map G), non-template DNA map ("Non-template DNA mask," 2.6 Å, 113,665 particles, map H), RPB5+DNA map ("RPB5+DNA mask," 2.7 Å, 59,649 particles, map I), CDC73 Ras-like domain map ("CDC73 mask," 3.4 Å, 43,339 particles, map J), PAF1+LEO1 map ("PAF1+LEO1 mask," 3.2 Å, 28,545 particles, map M), and SPT6 map ("SPT6 mask," 2.7 Å, 90,114 particles, map O).

A mask encompassing the CTR9 TPRs and WDR61 ("CTR9+WDR61 mask") was used to perform iterative 3D classifications of the 762,523 particles from map A. To enrich for the CTR9 N-terminus, another mask encompassing CTR9 residues 1-447 ("CTR9 TPR mask") was used to perform further 3D classifications. The resulting 83,886 particles were subjected to non-uniform refinement followed by local refinement using the CTR9+WDR61 mask to produce a CTR9+WDR61 map (3.8 Å, 63,837 particles, map K), as well as local refinement using the CTR9 TPR mask to produce a CTR9 TPR map (2.6 Å, 63,837 particles, map L).

The 28,545 particles in the PAF1+LEO1 map were subjected to additional 3D classification using a mask encompassing the LEO1 C-terminus ("LEO1 C-term mask") to enrich for the LEO1 C-terminal helix. The resulting 6,771 particles were subjected to non-uniform refinement followed by local refinement using the focus mask to produce a LEO1 C-term map (4 Å, 6,771 particles, map N).

The 90,114 particles in the SPT6 map were subjected to local refinement using the IWS1+ELOF1 mask to improve resolution of the SPT6-IWS1 interaction. This produced a SPT6 N-term map (2.9 Å, 90,114 particles, map P).

While map A revealed clear density for all elongation factors, we were unable to see the downstream nucleosome. Instead, all we could visualize in our maps was significant amounts of noise at low contour levels where we expected the downstream nucleosome, likely due to flexibility between the Pol II and the downstream nucleosome interface. Consequently, we created a mask encompassing the visible low-resolution features ("Nucleosome mask") and performed particle subtraction in RELION (v4.0) (53) using map A particles. Subtracted particles were imported back into cryoSPARC for iterative rounds of 2D classification and heterogeneous refinement to obtain a distinct nucleosome class of 85,098 particles. 3D classification using a mask encompassing the entire density was performed to obtain a 33,160 particle class containing density for the LEDGF PWWP. To further enrich the PWWP density, 3D classification using a mask encompassing the PWWP binding site for the distal H3 ("PWWP mask") was performed. The resulting class of 19,562 particles was subjected to non-uniform refinement followed by local refinement of the entire density to produce a nucleosome map of 3.5 Å (map Q).

We were unable to simultaneously visualize both Pol II and the nucleosome in the original data set (particles and processing corresponding to map A). Therefore, we reverted the 19,562 particle class from the nucleosome map Q to their original particles in RELION (v5.0) (53).We performed particle subtraction of the reverted particles in RELION (v5.0) using a map encompassing the downstream nucleosome, DNA, and Pol II without RPB4/7 ("10-subunit Pol II with nucleosome mask"). Subtracted particles were imported back into cryoSPARC for ab-initio reconstruction, non-uniform refinement, and local refinements with focus masks of the nucleosome or 10-subunit pol II to produce a nucleosome placement map (7.6 Å, 19,562 particles, map R) and a local 10-subunit Pol II map (8.8 Å, 19,562 particles, map S). Map S was aligned to map A using Align 3D Maps in cryoSPARC, then map R was aligned to the aligned map S.

To generate a composite map, maps B-P were aligned to map A using Align 3D Maps in cryoSPARC. The Frankenmap package from Warp (54)was then used to produce the composite map T using the aligned maps B-P.

RNA polymerase II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1-LEDGF-nucleosome complex model building and refinement

All fitting was performed in UCSF ChimeraX v1.7-1.9 (55) and refinements were performed in ISOLDE (56). An initial activated elongation complex model was generated where Pol II, DSIF, SPT6, PAF1c, and RTF1 from PDB

6TED were modeled into map A. The HEAT/TND of IWS1 and TFIIS from PDB 9EGZ were also modeled into map A. PAF1 and LEO1 from PDB 6TED were replaced with an AlphaFold-Multimer generated model of PAF1 and LEO1 with RPB2. Similarly, SPT6 was replaced with an AlphaFold-Multimer model of SPT6-RPB4-RPB7. AlphaFold-Multimer generated models of ELOF1-RPB2, IWS1-RPB5, IWS1-SPT4-SPT5, IWS1-SPT6, IWS1-RPB2, IWS1-ELOF1, and IWS1-RPB1 were also used to build initial models of ELOF1 bound to the elongation complex, the IWS1 RPB5 helix (residues 528-542), the IWS1-NGN interaction (IWS1 residues 693-721), the SPT6 N-terminal helix (residues 207-230) interaction with the IWS1 HEAT/TND, the IWS1 RPB2 lobe loop (IWS1 residues 749-760), the IWS1-ELOF1 belt (IWS1 residues 760-770), and the IWS1-RPB1 jaw and cleft interactions (IWS1 residues 787-819), respectively.

Maps A and G were used to refine all subunits of Pol II. Missing residues in the RPB2 protrusion domain from PDB 6TED were also built. Map B was used to refine IWS1 residues 554-712 and 749-770. Gaussian-filtered map B to 1.5 standard deviations provided sufficient density to place IWS1 residues 713-721 based on the AlphaFold model. Map C was used to refine IWS1 residues 805-819. IWS1 residues 787-795 from the IWS1-RPB1 AlphaFold prediction were docked into the Map C density. Map D was used to refine TFIIS. The Zn finger domain of TFIIS was rigid-body built into Gaussian-filtered map D. Map E was used to refine SPT4 and SPT5. Map F was used to refine the pincer helices and fastener of RTF1 and rigid-body fit the RTF1 linker and Plus3 domain. Map H was used to unambiguously assign the register as bp -9 by defining purine and pyrimidine bases in the DNA · RNA hybrid. Map H was also used to build the backtracked RNA strand using PyMOL v2.5.2 (57), and model the backbone of the single-stranded nontemplate DNA in the transcription bubble. Map I was used to rigid-body fit the IWS1 RPB5 helix (residues 528-542). Map J was used to rigid-body fit an AlphaFold model of the CDC73 Ras-like domain (residues 360-531). Map J was also Gaussian-filtered to 2 standard deviations to rigid-body fit the tSH2 domain. Map K was used to refine WDR61, CTR9 residues 440-810, and CDC73 residues 218-260. Map K was also used to rigid-body fit CTR9 residues 3-439 and PAF1 residues 25-113. Map L was used to rigid-body fit the CTR9 trestle helix. Map M was used to rigid-body fit PAF1 and LEO1. Map N was used to rigid-body fit the LEO1 C-terminal helix. Map O was used to refine RPB4 and RPB7 and fit SPT6. Map P was used to refine SPT6 residues 207-218, then map P Gaussian-filtered to 2 standard deviations was used to rigid-body fit SPT6 residues 219-230 using the IWS1-SPT6 AlphaFold model.

The nucleosome model was built separately from the activated elongation complex model. PDB 6S01 was fit into map Q, then the downstream DNA was extended in PyMOL v2.5.2 (57). Regions of the DNA from PDB 6S01 which did not have sufficient density in map Q were deleted.

To position the nucleosome relative to the activated elongation complex, the nucleosome model was fit into the aligned map R in UCSF ChimeraX v1.9 (55).

Due to low resolution, side-chains were removed for the following regions: TFIIS (residues 250-301), CDC73 (360-531), SPT6 (219-230, 332-339, 426-430, 483-488, 696-704, 759-763, 775-808, 815-882, 1075-1135, 1172-1175, 1227-1322,1323-1514), PAF1 (45-113, 151-371), LEO1 (356-512, 529-555), CTR9 (3-439, 811-892), IWS1 (528-542, 712-721, 787-795), SPT5 (270-397, 524-645, 753-781), RTF1 (353-502, 597-624), and LEDGF (1-91). Nucleotide bases were removed for the single-stranded non-template DNA.

The model was subsequently real space refined in PHENIX (58) with secondary restraints.

RNA polymerase II-RECQL5 complex formation and cryo-EM analysis and image processing

ECs were formed on a bubble scaffold with the following nucleic acid sequence: template DNA (IDT) (5′ -GCT CCC AGC TCC CTG CTG GCT CCG AGT GGG TTC TGC CGC TCT CAA TGG-3′), non-template DNA (IDT) (5′ - CCA TTG AGA GCG GCC CTT GTG TTC AGG AGC CAG CAG GGA GCT GGG AGC-3′) and RNA (IDT) (5'-/6-FAM/ UUA AGG AAU UAA GUC GUG CGU CUA AUA ACC GGA GAG GGA ACC CAC U-3′) (26). The scaffold contains 9 bps of RNA–DNA hybrid, 37 nt of exiting RNA, 11 nt of bubble, 14 bps of upstream DNA and 23 bps of downstream DNA. RNA and template DNA were mixed in equimolar ratios and annealed by incubating the nucleic acids at 95 °C for 5 min and then decreasing the temperature in steps of 1 °C min−1 to a final temperature of 30 °C in a thermocycler. All concentrations refer to the final concentrations used in the assay. S. scrofa Pol II (300 nM) and the RNA–template hybrid (500 nM) were incubated for 5 min on ice. The non-template DNA (667 nM) was added and the sample was incubated for another 5 min. RECQL5 (500 nM) and ADP-BeF3 (1 mM) were added in a final buffer containing 150 mM NaCl, 20 mM Na·HEPES pH 7.4, 3 mM MgCl2, 4% (v/v) glycerol and 1 mM TCEP pH 8. The sample was incubated for 30 min on ice.

The sample was mildly cross-linked using glutaraldehyde. The sample was dialyzed into dialysis buffer (150 mM NaCl, 20 mM Na·HEPES pH 7.4 at 25 °C, 3 mM MgCl2, 1 mM TCEP pH 8) for 3 h.

Quantifoil R2/1 on 200 Mesh copper grids were glow discharged for 30 s at 15 mA using a Pelco Easiglow plasma discharge system. 4 μL of dialyzed sample was applied to grids for 8 s, blotted for 5 s with a blot force of 8, and vitrified by plunging into liquid ethane using a Vitrobot Mark IV (FEI) at 5 °C and 100 % humidity.

Cryo-EM data was collected on a ThermoFisher Scientific Talos Arctica operating at 200 kV equipped with a Gatan K3 direct electron detector. Data collection was automated using SerialEM software. Data was collected at a pixel size of 1.1 Å with a defocus range of 1 to 2.2 μm. The data set yielded 227 micrographs with 50 movie frames at an exposure time of 3.99 s with an electron flux of 13.045 e− per A2 per s for a total exposure dose of 52.11 e− per Å2.

Initial data processing was performed in cryoSPARC (52). Movie alignment was performed using patch motion correction followed by patch CTF estimation. Blob picker was then used and 52,539 particles were extracted at a box size of 350 pixels. Two rounds of heterogenous refinement were performed to enrich for Pol II. Remaining particles were subjected to non-uniform refinement. To enrich for the RECQL5 KIX domain, we performed masked classification without image alignments in RELION (v4.0) (53). After classification, particles were re-imported to cryoSPARC and subjected to non-uniform refinement (map U).

Fit of predicted ELOA helix in coulomb potential map

PDB 8OF0 and EMD-16831 were opened in UCSF ChimeraX v1.8 (55) and EMD-16831 was Gaussian-filtered to 1 standard deviation. RPB1 from a RPB1-ELOA1 AlphaFold prediction was aligned to the RPB1 model in PDB 8OF0, placing the predicted RPB1 jaw-interacting ELOA1 helix (residues 756-772) from the AlphaFold model into the unassigned density in EMD-16831.

Sequence alignments

Sequence alignments were performed using PROMALS3D (59) with structure inputs. IWS1 homolog sequences were fetched from UniProt (60). IWS1 homolog structures were fetched from the AlphaFold Protein Structure Database (61,62) using their UniProt ID or predicted using AlphaFold2 (63) when not available in the database.

Selection of Candidates for AlphaFold-Multimer (AF-M) Screen

Human proteins with the GO term "transcription elongation from RNAP2 promoters" (GO:0006368) were selected to screen against RPB1 (2). To identify other interactors of RPB1 beyond those involved in elongation, "top" interactors with more than 5 evidence types listed for the *H. sapiens* RPB1 identifier on BioGRID (64) were selected. A total of 80 proteins were selected for the screen. All proteins were screened against the *H. sapiens* RPB1 (Uniprot ID P24928) residues 1-1487.

AF-M Structure Generation

Unless stated otherwise, all interactions were predicted using AlphaFold-Multimer (65) (AF-M) via an in-house modified version of ColabFold 1.52 (66,67) . AF-M was run with v2.3 weights, 1 ensemble, 3 recycles, templates enabled, dropout disabled, and maximum Multiple Sequence Alignments (MSA) depth settings (max_seq = 508, max_extra_seq = 2048). MSAs (paired and unpaired) were fetched from a remote server via the MMSeqs2 (68) API that is integrated into the ColabFold (66) pipeline. To reduce computational burden during the interaction screens we only ran 3 models (1,2, and 4) and did not run any pairs with total amino acid lengths exceeding 3,600 residues (GPU memory limit). The generation time of each structure prediction as well as the run settings have been recorded as REMARK lines in the PDB output files.

AF-M Prediction Analysis

Python scripts were utilized to analyze structural predictions generated by AF-M as previously described (67). Briefly, confident interchain residue contacts were extracted from AF-M structures by identifying proximal residues (heavy atom distance <5 Å) where both residues have pLDDT values >50 and PAE score <15 Å. All downstream analysis of interface statistics (average pLDDT, average PAE) were calculated on data from these selected inter-residue pairs (contacts). Average interface pLDDT values above 70 are generally considered confident. The average models score was calculated by averaging the number of independent AF-M models that predicted a specific inter-residue contact across all unique pairs in all models. This number was additionally normalized by dividing by the number of models run to produce a final average model score that ranges from 0 (worst) to 1 (best). An average models value above 0.5 is generally considered confident. pDockQ (63) estimates of interface accuracy scores were calculated independently of the contact analysis described above using code that was adapted from the original implementation (69). pDockQ values above 0.23 are considered confident.

SPOC Interaction Analysis

The random forest classifier algorithm (Structure Prediction and Omics Classifier) SPOC (67) was used to score the binary interaction predicted by via AF-M. Briefly, this classifier was trained to distinguish biologically relevant interacting protein pairs from non-biologically likely interaction pairs in AF-M interaction screens. SPOC assigns each interaction a score that ranges from 0 (worst) to 1 (best). Higher scores indicate that AF-M interface data and several types of externally sourced omics datasets support the potential existence of the binary interaction produced by AF-M. The highest confident interactions are generally associated with SPOC scores above 0.5.

Figure generation

All figures were generated using Adobe Illustrator, GraphPad Prism, and UCSF ChimeraX v1.7-1.9 (55). Sequence alignments were formatted in Unipro UGENE (70), then exported to .svg format and imported into Adobe Illustrator.

Local resolution estimations for each map were generated using the Local Resolution Estimation job in cryoSPARC. Local resolution maps were generated in ChimeraX using Surface Color of the refinement volume by the volume data of the local resolution output. FSC resolution values were obtained using the following calculation: Resolution (Å) = (box size * pixel size) / wave number. Wave number values were generated in cryoSPARC after FSC-mask auto-tightening. FSC curves were plotted in GraphPad Prism using the FSC tightmask values and calculated resolutions.

Table 1. Cryo-EM data collection, refinement and validation statistics

	#1 composite map T RNA polymerase II-DSIF-SPT6-PAF1c-TFIIS-IWS1-ELOF1	#2 activated elongation complex (consensus) map A	#3 IWS1+ELOF1 map B	#4 IWS1 C-term map C	#5 TFIIS map D
Data collection and processing					
Magnification	105,000	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e–/Å^2)	37	37	37	37	37
Defocus range (μm)	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6
Pixel size (Å)	1.19	1.19	1.19	1.19	1.19
Symmetry imposed	C1	C1	C1	C1	C1
Initial particle images (no.)	4,597,368	4,597,368	762,523	762,523	762,523
Final particle images (no.)	762,523	762,523	378,540	156,410	36,198
Map resolution (Å)	2.4	2.4	2.6	2.4	4.3
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.4-10	2.4-10	2.6-10	2.4-10	4.3-10
Refinement					
Initial model used (PDB code)	6TED,9EGZ				
Model resolution (Å)	2.4				
FSC threshold	0.143				
Model resolution range (Å)	2.4-3.1				
Map sharpening *B* factor (Å^2)	63.2				
Model composition					
Non-hydrogen atoms	60482				
Protein residues	8304				
Ligands	11				
B factors (Å^2)					
Protein	97.29				
Ligand	149.34				
R.m.s. deviations					
Bond lengths (Å)	0.007				
Bond angles (°)	0.813				
Validation					
MolProbity score	2.27				
Clashscore	12.82				
Poor rotamers (%)	2.31%				
Ramachandran plot					
Favored (%)	94.47%				
Allowed (%)	5.38%				
Disallowed (%)	0.15%				

	#6 DSIF map E	#7 RTF1 map F	#8 10-subunit Pol II map G	#9 Non-template DNA map H	#10 RPB5+DNA map I
Data collection and processing					
Magnification	105,000	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e–/Å^2)	37	37	37	37	37
Defocus range (μm)	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6
Pixel size (Å)	1.19	1.19	1.19	1.19	1.19
Symmetry imposed	C1	C1	C1	C1	C1
Initial particle images (no.)	762,563	762,523	762,523	762,523	762,523
Final particle images (no.)	79,487	27,130	115,139	113,665	59,649
Map resolution (Å)	2.7	2.8	2.5	2.6	2.7
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.7-10	2.8-10	2.5-10	2.6-10	2.7-10
Refinement					
Initial model used (PDB code)					
Model resolution (Å)					
FSC threshold					
Model resolution range (Å)					
Map sharpening *B* factor (Å^2)					
Model composition					
Non-hydrogen atoms					
Protein residues					
Ligands					
B factors (Å^2)					
Protein					
Ligand					
R.m.s. deviations					
Bond lengths (Å)					
Bond angles (°)					
Validation					
MolProbity score					
Clashscore					
Poor rotamers (%)					
Ramachandran plot					
Favored (%)					
Allowed (%)					
Disallowed (%)					

	#11 CDC73 map J	#12 CTR9+WDR61 map K	#13 CTR9 TPR map L	#14 PAF1+LEO1 map M	#15 LEO1 C-term map N
Data collection and processing					
Magnification	105,000	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e–/Å^2)	37	37	37	37	37
Defocus range (μm)	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6
Pixel size (Å)	1.19	1.19	1.19	1.19	1.19
Symmetry imposed	C1	C1	C1	C1	C1
Initial particle images (no.)	762,563	762,523	762,523	762,523	762,523
Final particle images (no.)	43,339	63,837	63,837	28,545	6,771
Map resolution (Å)	3.4	3.8	2.6	3.2	4.0
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	3.4-10	3.8-10	2.6-10	3.2-10	4.0-10
Refinement					
Initial model used (PDB code)					
Model resolution (Å)					
FSC threshold					
Model resolution range (Å)					
Map sharpening *B* factor (Å^2)					
Model composition					
Non-hydrogen atoms					
Protein residues					
Ligands					
B factors (Å^2)					
Protein					
Ligand					
R.m.s. deviations					
Bond lengths (Å)					
Bond angles (°)					
Validation					
MolProbity score					
Clashscore					
Poor rotamers (%)					
Ramachandran plot					
Favored (%)					
Allowed (%)					
Disallowed (%)					

	#16 SPT6 map O	#17 SPT6 N-term map P	#18 LEDGF+ nucleosome map Q	#19 Nucleosome placement map R	#20 Local 10-subunit Pol II map S
Data collection and processing					
Magnification	105,000	105,000	105,000	105,000	105,000
Voltage (kV)	300	300	300	300	300
Electron exposure (e–/Å^2)	37	37	37	37	37
Defocus range (μm)	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6	-1.8 to -0.6
Pixel size (Å)	1.19	1.19	1.19	1.19	1.19
Symmetry imposed	C1	C1	C1	C1	C1
Initial particle images (no.)	762,523	762,523	762,523	19,562	19,562
Final particle images (no.)	90,114	90,114	19,562	19,562	19,562
Map resolution (Å)	2.7	2.9	3.5	7.6	8.7
FSC threshold	0.143	0.143	0.143	0.143	0.143
Map resolution range (Å)	2.7-10	2.9-10	3.5-10	7.6-20	8.7-20
Refinement					
Initial model used (PDB code)			6S01		
Model resolution (Å)			3.5		
FSC threshold			0.143		
Model resolution range (Å)			3.5-4.2		
Map sharpening *B* factor (Å^2)			48.0		
Model composition					
Non-hydrogen atoms			12505		
Protein residues			853		
Ligands			0		
B factors (Å^2)					
Protein			144.86		
Ligand			-		
R.m.s. deviations					
Bond lengths (Å)			0.007		
Bond angles (°)			0.980		
Validation					
MolProbity score			2.14		
Clashscore			15.50		
Poor rotamers (%)			1.83		
Ramachandran plot					
Favored (%)			96.27		
Allowed (%)			3.73		
Disallowed (%)			0.00		

	#21 RECQL5-RNA polymerase II map U
Data collection and processing	
Magnification	36,000
Voltage (kV)	200
Electron exposure (e–/Å^2)	52.11
Defocus range (μm)	1 to 2.2
Pixel size (Å)	1.1
Symmetry imposed	C1
Initial particle images (no.)	52,539
Final particle images (no.)	15,305
Map resolution (Å)	3.1
FSC threshold	0.143
Map resolution range (Å)	3.1-10
Refinement	
Initial model used (PDB code)	
Model resolution (Å)	
FSC threshold	
Model resolution range (Å)	
Map sharpening *B* factor (Å^2)	
Model composition	
Non-hydrogen atoms	
Protein residues	
Ligands	
B factors (Å^2)	
Protein	
Ligand	
R.m.s. deviations	
Bond lengths (Å)	
Bond angles (°)	
Validation	
MolProbity score	
Clashscore	
Poor rotamers (%)	
Ramachandran plot	
Favored (%)	
Allowed (%)	
Disallowed (%)	

Table 2. Input structural models and model confidence

Complex/domain	Chain id(s)	Input model	Level of confidence
RNA Polymerase II	A-L	6TED	Atomic
SPT6	M	6TED, Alphafold	Atomic (residues 207-218, 266-331, 340-425, 431-482, 512-695, 705-758, 809-814, 883-1074, 1136-1171), Rigid body fitting (residues 219-230, 332-339, 426-430, 483-488, 696-704, 759-763, 775-808, 815-882, 1075-1135, 1172-1175, 1227-1322,1323-1514)
DNA	N,T	6TED, 6S01, *de novo*	Atomic except for single-stranded non-template DNA (residues -18--8) on 9MLC
ELOF1	O	Alphafold	Atomic
RNA	P	6TED, *de novo*	Atomic
CTR9	Q	6TED	Atomic (residues 440-810), Rigid body fitting (residues (3-439, 811-892)
RTF1	R	6TED	Atomic (residues 503-596), Rigid body fitting (residues 353-502, 597-624)
TFIIS	S	9EGZ	Atomic (residues 132-249), Rigid body fitting (residues 250-301)
LEO1	U	Alphafold	Atomic (residues 513-528), Rigid body fitting (residues 356-512, 529-555)
PAF1	V	Alphafold	Atomic (residues 120-149), Rigid body fitting (residues 45-113, 151-371)
WDR61	W	Alphafold	Atomic
CDC73	X	6TED, Alphafold	Atomic (residues 218-260), Rigid body fitting (residues 360-531)
SPT4	Y	6TED	Atomic
SPT5	Z	6TED	Atomic (residues 160-269, 417-523, 702-752), Rigid body fitting (residues 270-397, 524-645, 753-781)
H3	a,e	6S01	Atomic
H4	b,f	6S01	Atomic
H2A	d,h	6S01	Atomic
H2B	c,g	6S01	Atomic
IWS1	i	9EGZ, Alphafold	Atomic (residues 554-711, 749-770, 787-795, 805-819), Rigid body fitting (residues 528-542, 712-721, 787-795)
LEDGF PWWP	I	6S01	Rigid body fitting

Figure S1

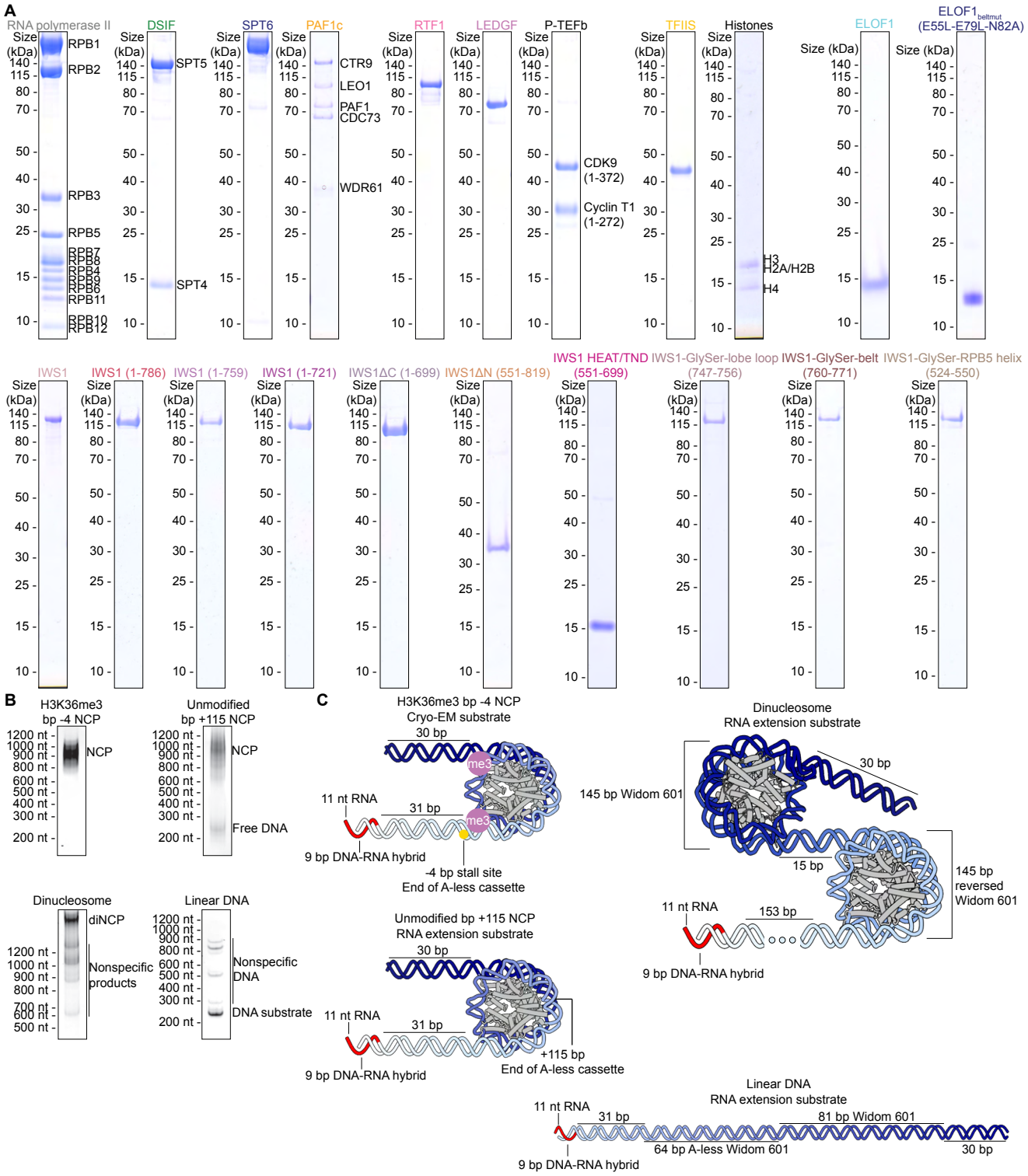
A
RNA polymerase II
RPB1
RPB2
RPB3
RPB5
RPB7
RPB8
RPB4
RPB9
RPB6
RPB11
RPB10
RPB12
DSIF
SPT5
SPT4
SPT6
PAF1c
CTR9
LEO1
PAF1
CDC73
WDR61
RTF1
LEDGF
P-TEFb
CDK9 (1-372)
Cyclin T1 (1-272)
TFIIS
Histones
H3
H2A/H2B
H4
ELOF1
ELOF1 beltmut (E55L-E79L-N82A)
Size (kDa)
140
115
80
70
50
40
30
25
15
10
IWS1
IWS1 (1-786)
IWS1 (1-759)
IWS1 (1-721)
IWS1ΔC (1-699)
IWS1ΔN (551-819)
IWS1 HEAT/TND (551-699)
IWS1-GlySer-lobe loop (747-756)
IWS1-GlySer-belt (760-771)
IWS1-GlySer-RPB5 helix (524-550)
B
H3K36me3 bp -4 NCP
Unmodified bp +115 NCP
1200 nt
1000 nt
900 nt
800 nt
600 nt
500 nt
400 nt
300 nt
200 nt
NCP
Free DNA
Dinucleosome
diNCP
Nonspecific products
Linear DNA
Nonspecific DNA
DNA substrate
C
H3K36me3 bp -4 NCP Cryo-EM substrate
30 bp
me3
11 nt RNA
31 bp
9 bp DNA-RNA hybrid
-4 bp stall site
End of A-less cassette
Dinucleosome RNA extension substrate
145 bp Widom 601
30 bp
15 bp
145 bp reversed Widom 601
153 bp
Unmodified bp +115 NCP RNA extension substrate
+115 bp
End of A-less cassette
Linear DNA RNA extension substrate
81 bp Widom 601
64 bp A-less Widom 601

Figure S1. Constructs and substrates. (A) SDS-PAGE of purified protein components. Samples were loaded onto 4-12% Bis-Tris gels, run in 1x MES buffer, and stained with OneStep Blue. **(B)** Native gels of substrates used. Samples were run on TBE native gels and stained with Sybr Gold. **(C)** Schematic of substrates used. DNA is colored from light to dark blue in the direction of transcription. RNA colored in red. Histones colored in gray.

Figure S2

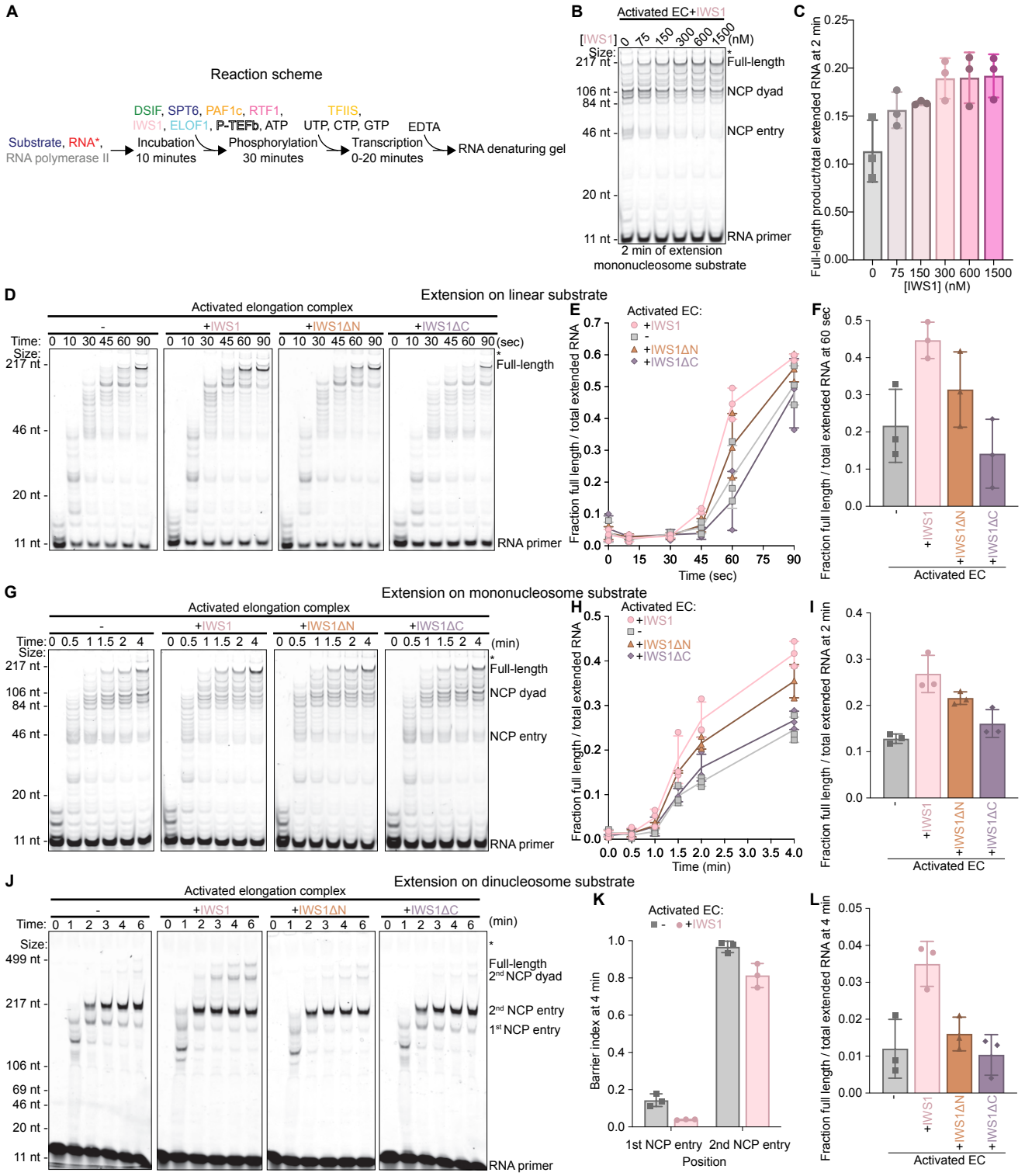
A
Reaction scheme
Substrate, RNA*, RNA polymerase II
Incubation 10 minutes
DSIF, SPT6, PAF1c, RTF1, IWS1, ELOF1, P-TEFb, ATP
Phosphorylation 30 minutes
TFIIS, UTP, CTP, GTP
Transcription 0-20 minutes
EDTA
RNA denaturing gel
B
Activated EC+IWS1
[IWS1] 0 75 150 300 600 1500 (nM)
Size:
217 nt
106 nt
84 nt
46 nt
20 nt
11 nt
Full-length
NCP dyad
NCP entry
RNA primer
2 min of extension
mononucleosome substrate
C
Full-length product/total extended RNA at 2 min
[IWS1] (nM)
D
Extension on linear substrate
Activated elongation complex
Time: 0 10 30 45 60 90 (sec)
Full-length
RNA primer
E
Activated EC:
+IWS1
-
+IWS1ΔN
+IWS1ΔC
Fraction full length / total extended RNA
Time (sec)
F
Fraction full length / total extended RNA at 60 sec
Activated EC
G
Extension on mononucleosome substrate
Time: 0 0.5 1 1.5 2 4 (min)
H
Time (min)
I
Fraction full length / total extended RNA at 2 min
J
Extension on dinucleosome substrate
Time: 0 1 2 3 4 6 (min)
499 nt
69 nt
2nd NCP dyad
2nd NCP entry
1st NCP entry
K
Barrier index at 4 min
1st NCP entry
2nd NCP entry
Position
L
Fraction full length / total extended RNA at 4 min

Figure S2. RNA extension assays on different substrates. (A) Reaction scheme of RNA extension assay. **(B)** RNA extension assay with IWS1 titration on a mononucleosomal substrate. The samples on the gel were quenched 2 min after addition of CTP, GTP and UTP (10 μM). Samples were separated by denaturing gel electrophoresis. RNA extension was monitored using FAM dye on RNA. Initial RNA and extended products are marked. Results are representative of at least three independent experiments. **(C)** Quantification of B from n = 3 independent experiments. **(D)** RNA extension assays on linear DNA with no IWS1, full-length IWS1, IWS1ΔN, and IWS1ΔC. Results are representative of at least three independent experiments. **(E)** Quantification of D from n = 3 independent experiments. Individual data points and standard deviation are shown. **(F)** Quantification of 60 sec timepoint from D from n = 3 independent experiments. Individual data points and standard deviation are shown. **(G)** RNA extension assays on mononucleosome with no IWS1, full-length IWS1, IWS1ΔN, and IWS1ΔC. Results are representative of at least three independent experiments. **(H)** Quantification of G from n = 3 independent experiments. Individual data points and standard deviation are shown. **(I)** Quantification of 2 min timepoint from G from n = 3 independent experiments. Individual data points and standard deviation are shown. **(J)** RNA extension assays on dinucleosome substrate with no IWS1, full-length IWS1, IWS1ΔN, and IWS1ΔC. Results are representative of at least three independent experiments. **(K)** Quantification of nucleosome barrier index at two NCP entry sites from n = 3 independent RNA extension assays. **(L)** Quantification of 4 min timepoint from J from n = 3 independent experiments. Individual data points and standard deviation are shown. Activated EC = Pol II, DSIF, SPT6, PAF1c including RTF1, TFIIS, ELOF1. Columns represent mean values across n = 3 independent experiments and error bars represent standard deviations for C, E, F, H, I, K, and L.

Figure S3

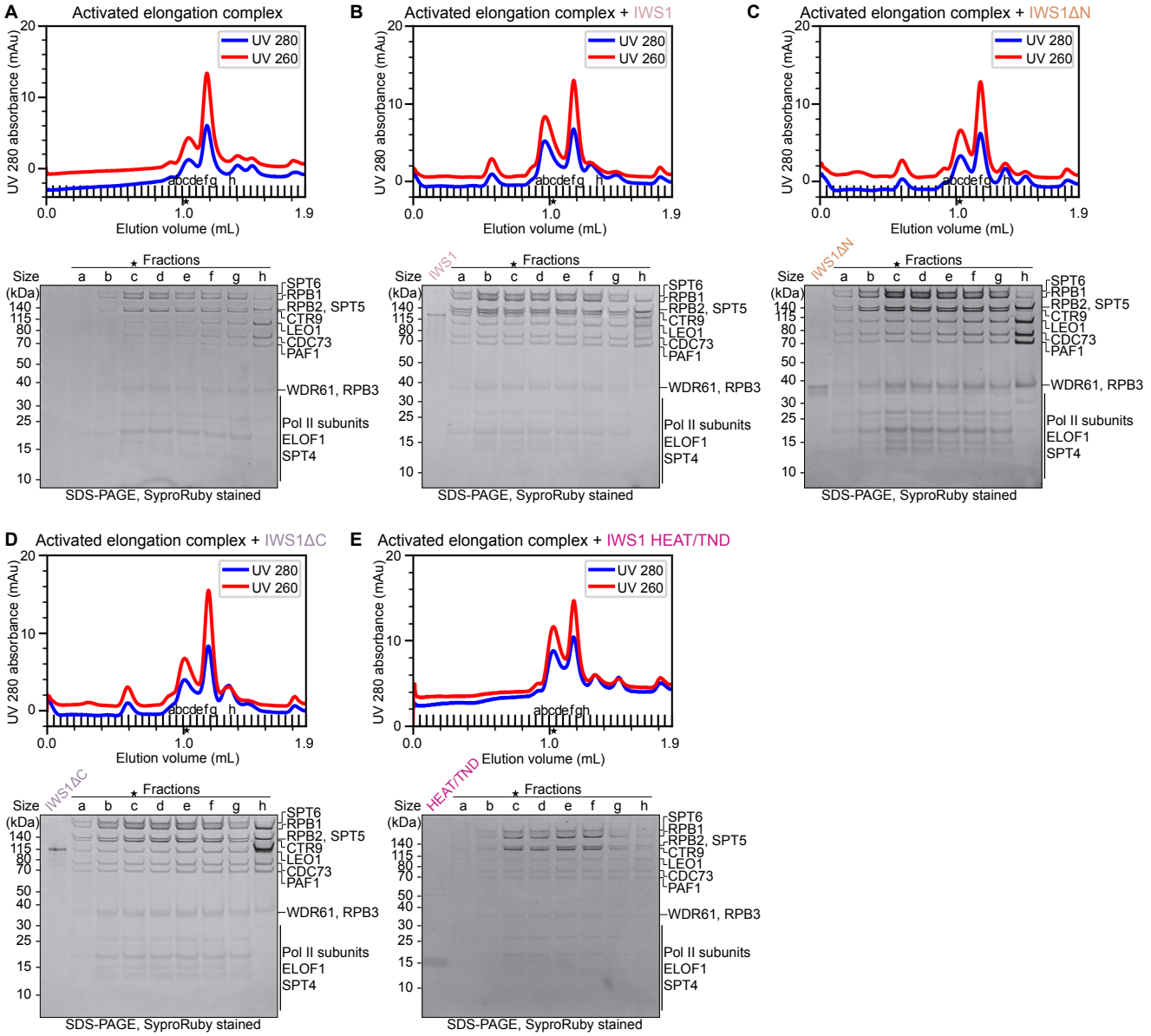
A
Activated elongation complex
B
Activated elongation complex + IWS1
C
Activated elongation complex + IWS1ΔN
D
Activated elongation complex + IWS1ΔC
E
Activated elongation complex + IWS1 HEAT/TND
UV 280
UV 260
UV 280 absorbance (mAu)
Elution volume (mL)
Fractions
Size (kDa)
SPT6
RPB1
RPB2, SPT5
CTR9
LEO1
CDC73
PAF1
WDR61, RPB3
Pol II subunits
ELOF1
SPT4
SDS-PAGE, SyproRuby stained

Figure S3. Full size exclusion chromatograms and SDS-PAGE related to Figure 1. UV 280 nm and UV 260 nm traces for size exclusion chromatograms and corresponding SDS-PAGE of elongation complexes assembled **(A)** without IWS1, **(B)** with full-length IWS1, **(C)** with IWS1ΔN, **(D)** with IWS1ΔC, and **(E)** with IWS1 HEAT/TND (551-699). Activated elongation complex = Pol II, DSIF, SPT6, PAF1c without RTF1, ELOF1. Fractions were loaded onto 4-12% Bis-Tris gels, run in 1x MES buffer, and stained with SyproRuby. Fraction displayed in Figure 1 indicated with star.

Figure S4

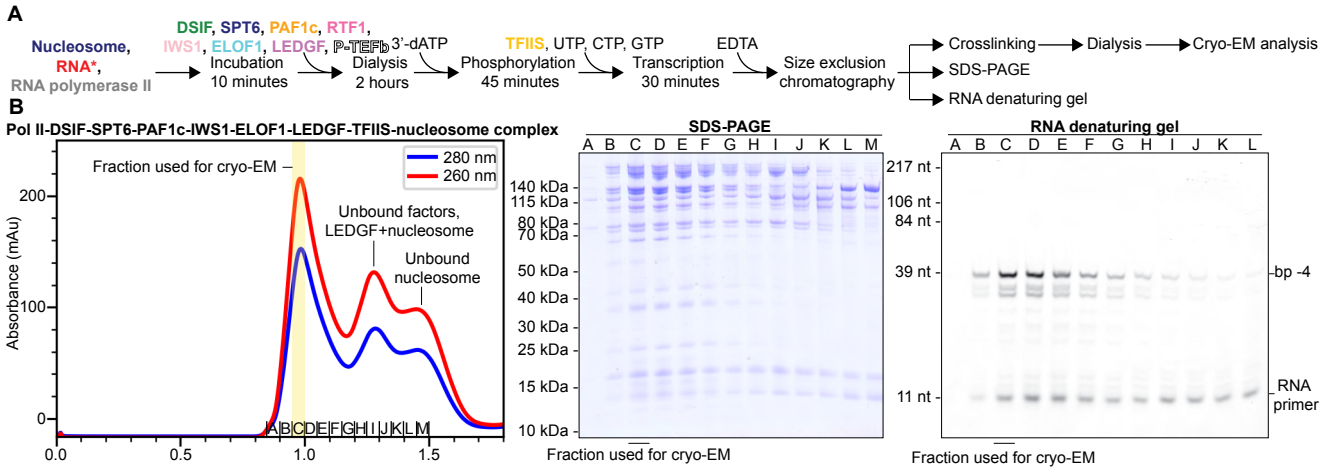
A
Nucleosome,
RNA*,
RNA polymerase II
DSIF, SPT6, PAF1c, RTF1,
IWS1, ELOF1, LEDGF, P-TEFb
Incubation
10 minutes
3'-dATP
Dialysis
2 hours
TFIIS, UTP, CTP, GTP
Phosphorylation
45 minutes
EDTA
Transcription
30 minutes
Size exclusion
chromatography
Crosslinking
Dialysis
Cryo-EM analysis
SDS-PAGE
RNA denaturing gel
B
Pol II-DSIF-SPT6-PAF1c-IWS1-ELOF1-LEDGF-TFIIS-nucleosome complex
Fraction used for cryo-EM
280 nm
260 nm
Unbound factors,
LEDGF+nucleosome
Unbound
nucleosome
Absorbance (mAu)
200
100
0
0.0
0.5
1.0
1.5
A B C D E F G H I J K L M
SDS-PAGE
A B C D E F G H I J K L M
140 kDa
115 kDa
80 kDa
70 kDa
50 kDa
40 kDa
30 kDa
25 kDa
15 kDa
10 kDa
Fraction used for cryo-EM
RNA denaturing gel
A B C D E F G H I J K L
217 nt
106 nt
84 nt
39 nt
11 nt
bp -4
RNA
primer
Fraction used for cryo-EM

Figure S4. Complex formation for cryo-EM. (A) Reaction scheme for complex formation for cryo-EM sample. **(B)** Size-exclusion chromatogram, SDS-PAGE, and TBE-urea denaturing PAGE of Pol II-DSIF-SPT6-PAF1c-IWS1-ELOF1-LEDGF-TFIIS-nucleosome complex. Fractions represented by the letters A-M. Fraction used for cryo-EM indicated by yellow bar on chromatogram and black bar on gels.

Figure S5

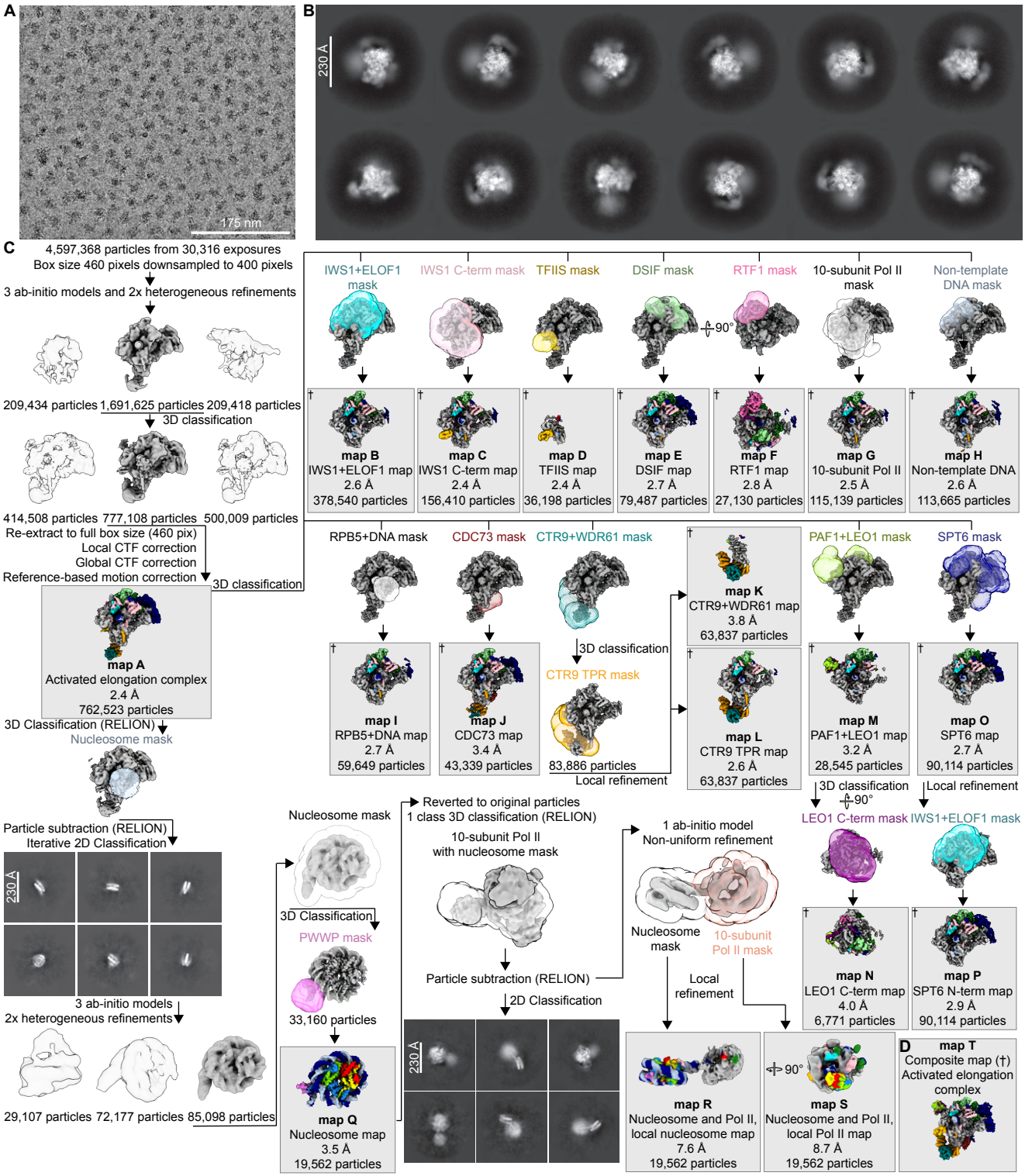
A
175 nm
B
230 Å
C
4,597,368 particles from 30,316 exposures
Box size 460 pixels downsampled to 400 pixels
3 ab-initio models and 2x heterogeneous refinements
209,434 particles 1,691,625 particles 209,418 particles
3D classification
414,508 particles 777,108 particles 500,009 particles
Re-extract to full box size (460 pix)
Local CTF correction
Global CTF correction
Reference-based motion correction
3D classification
map A
Activated elongation complex
2.4 Å
762,523 particles
3D Classification (RELION)
Nucleosome mask
Particle subtraction (RELION)
Iterative 2D Classification
230 Å
3 ab-initio models
2x heterogeneous refinements
29,107 particles 72,177 particles 85,098 particles
IWS1+ELOF1 mask
IWS1 C-term mask
TFIIS mask
DSIF mask
RTF1 mask
10-subunit Pol II mask
Non-template DNA mask
90°
map B
IWS1+ELOF1 map
2.6 Å
378,540 particles
map C
IWS1 C-term map
2.4 Å
156,410 particles
map D
TFIIS map
2.4 Å
36,198 particles
map E
DSIF map
2.7 Å
79,487 particles
map F
RTF1 map
2.8 Å
27,130 particles
map G
10-subunit Pol II
2.5 Å
115,139 particles
map H
Non-template DNA
2.6 Å
113,665 particles
RPB5+DNA mask
CDC73 mask
CTR9+WDR61 mask
map K
CTR9+WDR61 map
3.8 Å
63,837 particles
PAF1+LEO1 mask
SPT6 mask
3D classification
CTR9 TPR mask
map I
RPB5+DNA map
2.7 Å
59,649 particles
map J
CDC73 map
3.4 Å
43,339 particles
83,886 particles
Local refinement
map L
CTR9 TPR map
2.6 Å
63,837 particles
map M
PAF1+LEO1 map
3.2 Å
28,545 particles
map O
SPT6 map
2.7 Å
90,114 particles
3D classification
90°
Local refinement
LEO1 C-term mask
IWS1+ELOF1 mask
map N
LEO1 C-term map
4.0 Å
6,771 particles
map P
SPT6 N-term map
2.9 Å
90,114 particles
Nucleosome mask
3D Classification
PWWP mask
33,160 particles
map Q
Nucleosome map
3.5 Å
19,562 particles
Reverted to original particles
1 class 3D classification (RELION)
10-subunit Pol II
with nucleosome mask
Particle subtraction (RELION)
2D Classification
230 Å
1 ab-initio model
Non-uniform refinement
Nucleosome mask
10-subunit Pol II mask
Local refinement
map R
Nucleosome and Pol II, local nucleosome map
7.6 Å
19,562 particles
90°
map S
Nucleosome and Pol II, local Pol II map
8.7 Å
19,562 particles
D
map T
Composite map (†)
Activated elongation complex

Figure S5. Cryo-EM data processing of Pol II-DSIF-SPT6-PAF1c-IWS1-ELOF1-LEDGF-TFIIS-nucleosome complex. (A) Representative micrograph from data collection with scale bar (175 nm). **(B)** Representative 2D classes with scale bar (230 Å). **(C)** Sorting and classification of cryo-EM data. Resolution and particle count indicated for each final map. Resolutions reported for tight mask values after FSC-mask auto-tightening. All processing performed in cryoSPARC unless indicated. Densities colored according to Figure 2. Maps contributing to composite map indicated with dagger symbol (†). **(D)** Composite map of Pol II-DSIF-SPT6-PAF1c-IWS1-ELOF1-TFIIS activated elongation complex.

Figure S6

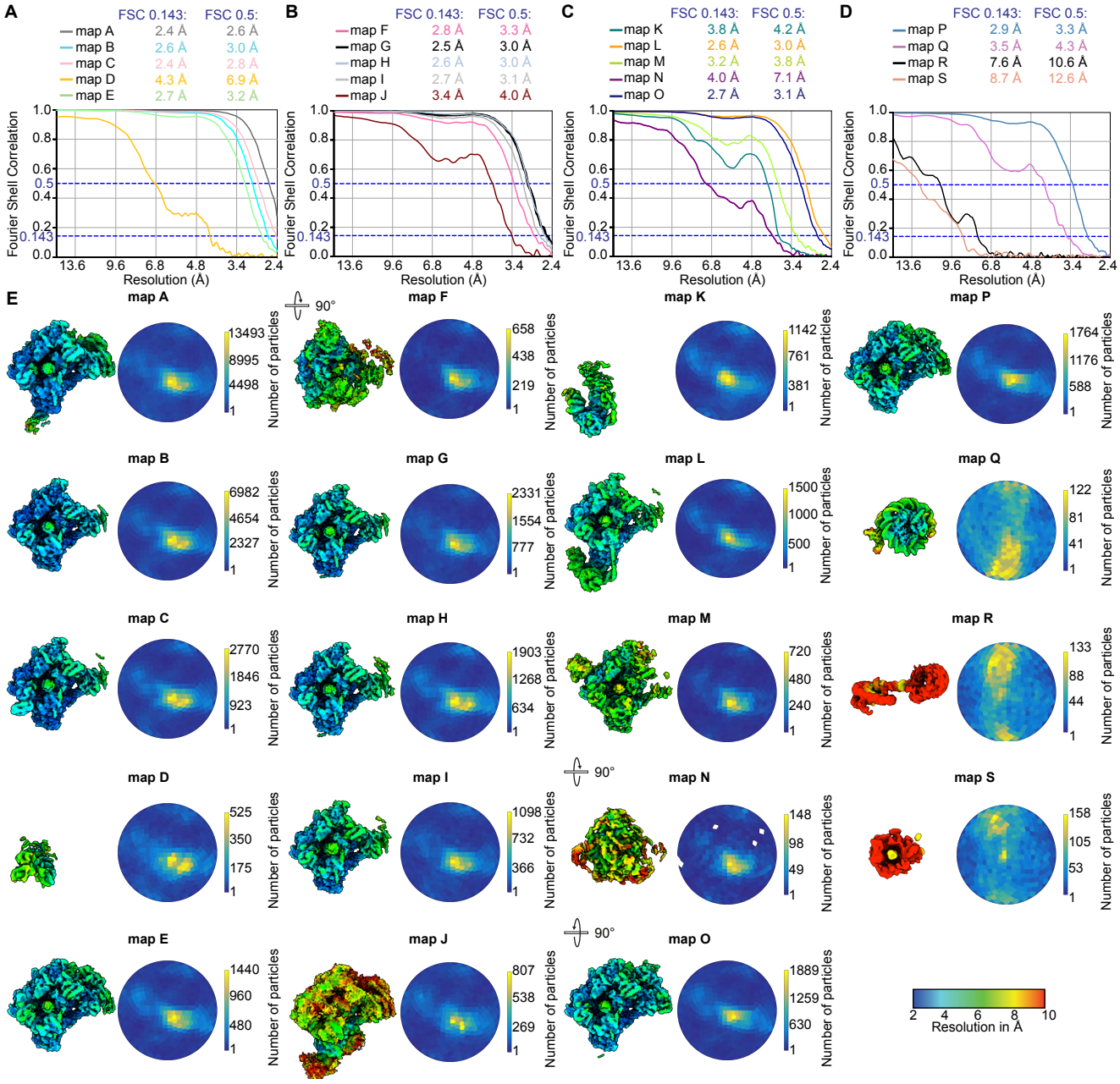
A
FSC 0.143: FSC 0.5:
map A 2.4 Å 2.6 Å
map B 2.6 Å 3.0 Å
map C 2.4 Å 2.8 Å
map D 4.3 Å 6.9 Å
map E 2.7 Å 3.2 Å
B
FSC 0.143: FSC 0.5:
map F 2.8 Å 3.3 Å
map G 2.5 Å 3.0 Å
map H 2.6 Å 3.0 Å
map I 2.7 Å 3.1 Å
map J 3.4 Å 4.0 Å
C
FSC 0.143: FSC 0.5:
map K 3.8 Å 4.2 Å
map L 2.6 Å 3.0 Å
map M 3.2 Å 3.8 Å
map N 4.0 Å 7.1 Å
map O 2.7 Å 3.1 Å
D
FSC 0.143: FSC 0.5:
map P 2.9 Å 3.3 Å
map Q 3.5 Å 4.3 Å
map R 7.6 Å 10.6 Å
map S 8.7 Å 12.6 Å
Fourier Shell Correlation
Resolution (Å)
E
Number of particles
90°
Resolution in Å

Figure S6. Data quality and metrics of Pol II-DSIF-SPT6-PAF1c-IWS1-ELOF1-LEDGF-TFIIS-nucleosome complex. FSC curves for maps **(A)** A-E, **(B)** F-J, **(C)** K-O, and **(D)** P-S. Resolution at FSC 0.5 and 0.143 reported for tight mask values after FSC-mask auto-tightening and color-coded according to map. **(D)** Local resolution estimation and particle angular distribution plots for each map.

Figure S7

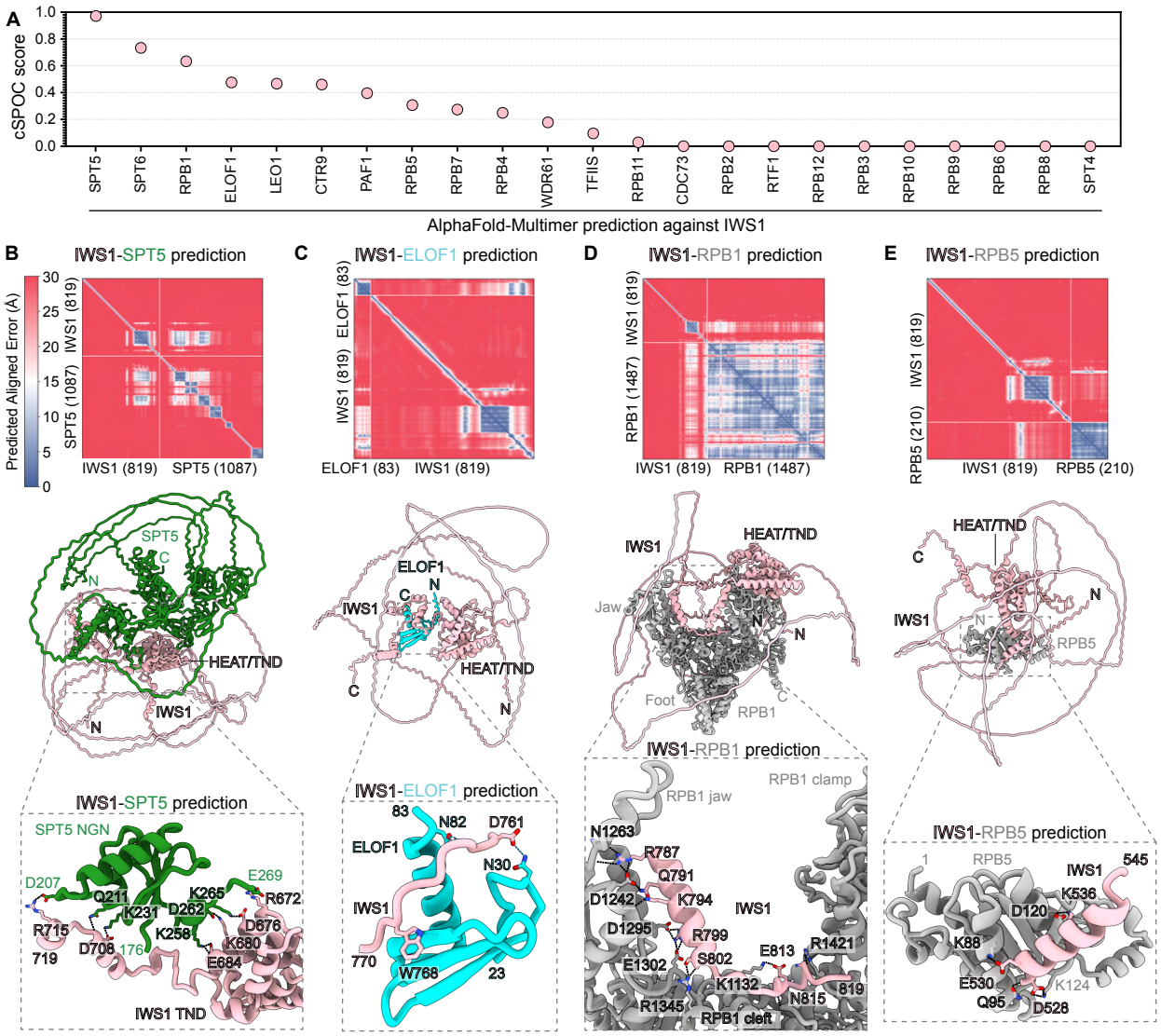
A
cSPOC score
1.0
0.8
0.6
0.4
0.2
0.0
SPT5
SPT6
RPB1
ELOF1
LEO1
CTR9
PAF1
RPB5
RPB7
RPB4
WDR61
TFIIS
RPB11
CDC73
RPB2
RTF1
RPB12
RPB3
RPB10
RPB9
RPB6
RPB8
SPT4
AlphaFold-Multimer prediction against IWS1
B
IWS1-SPT5 prediction
Predicted Aligned Error (Å)
IWS1 (819)
SPT5 (1087)
C
IWS1-ELOF1 prediction
ELOF1 (83)
D
IWS1-RPB1 prediction
RPB1 (1487)
E
IWS1-RPB5 prediction
RPB5 (210)
SPT5
HEAT/TND
IWS1
ELOF1
Jaw
Foot
RPB1
RPB5
SPT5 NGN
IWS1 TND
RPB1 jaw
RPB1 clamp
RPB1 cleft

Figure S7. AlphaFold-Multimer screen of IWS1 with core elongation factors. (A) Interaction screen using AlphaFold-Multimer. Predictions ranked by cSPOC score. **(B)** PAE plot of IWS1-SPT5 prediction and representative predicted model with IWS1-SPT5 NGN contacts shown. Interacting atoms shown as sticks and H-bonds shown as dotted lines. **(C)** PAE plot of IWS1-ELOF1 prediction and representative predicted model shown. Interacting atoms shown as sticks and H-bonds shown as dotted lines. **(D)** PAE plot of IWS1-RPB1 prediction and representative predicted model with IWS1-RPB1 jaw and cleft contacts shown. Interacting atoms shown as sticks and H-bonds shown as dotted lines. **(E)** PAE plot of IWS1-RPB5 prediction and representative predicted model shown. Interacting atoms shown as sticks and H-bonds shown as dotted lines.

Figure S8

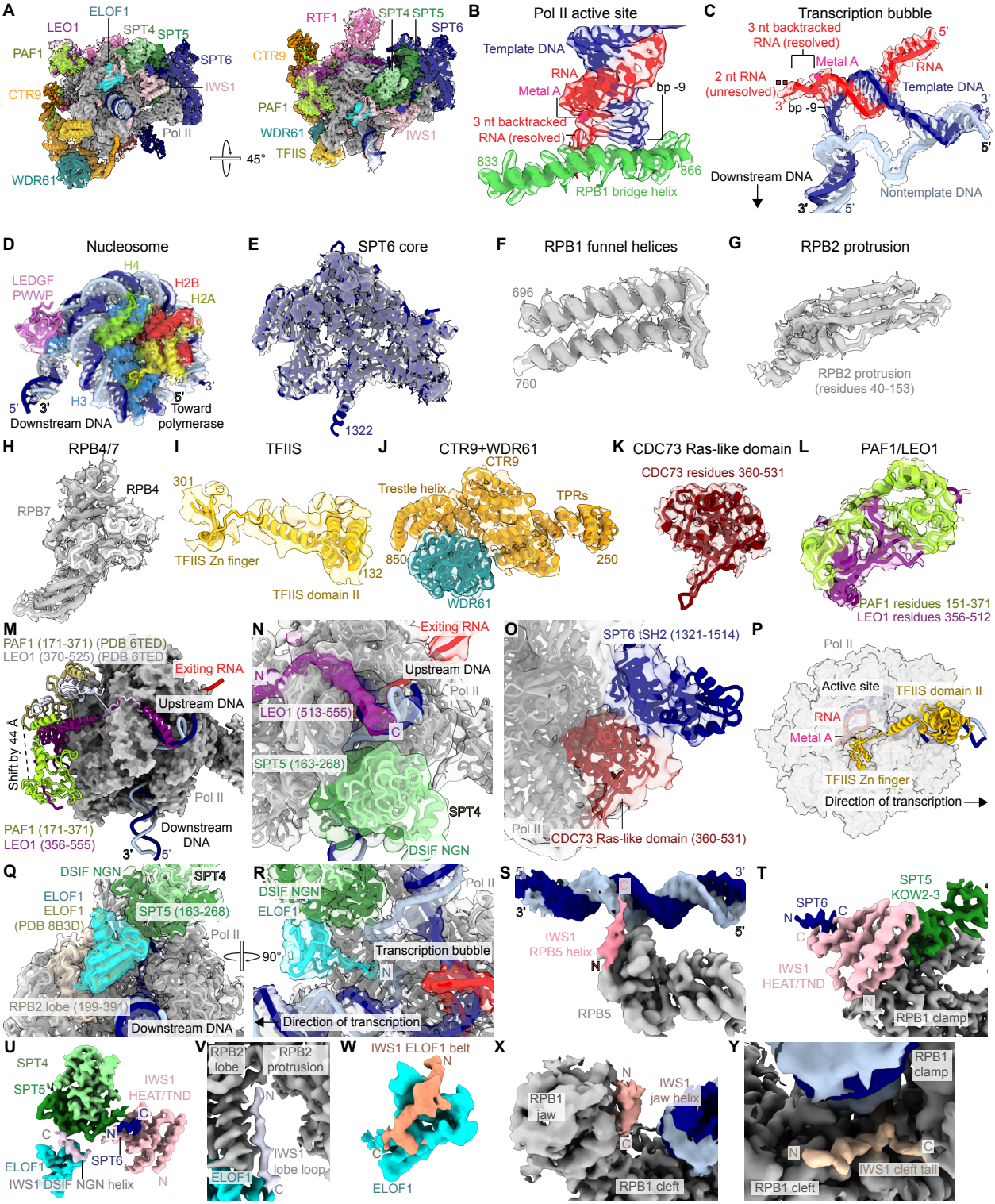
A
ELOF1
LEO1
SPT4
SPT5
PAF1
SPT6
CTR9
IWS1
Pol II
WDR61
RTF1
SPT4 SPT5
SPT6
CTR9
PAF1
WDR61
IWS1
TFIIS
45°
B
Pol II active site
Template DNA
RNA
Metal A
bp -9
3 nt backtracked
RNA (resolved)
833
866
RPB1 bridge helix
C
Transcription bubble
3 nt backtracked
RNA (resolved)
Metal A
RNA
2 nt RNA
(unresolved)
Template DNA
3'
bp -9
5'
3'
5'
Downstream DNA
Nontemplate DNA
3'
5'
D
Nucleosome
H4
LEDGF
PWWP
H2B
H2A
H3
5'
3'
5'
3'
Downstream DNA
Toward
polymerase
E
SPT6 core
1322
F
RPB1 funnel helices
696
760
G
RPB2 protrusion
RPB2 protrusion
(residues 40-153)
H
RPB4/7
RPB4
RPB7
I
TFIIS
301
TFIIS Zn finger
132
TFIIS domain II
J
CTR9+WDR61
CTR9
Trestle helix
TPRs
850
250
WDR61
K
CDC73 Ras-like domain
CDC73 residues 360-531
L
PAF1/LEO1
PAF1 residues 151-371
LEO1 residues 356-512
M
PAF1 (171-371) (PDB 6TED)
LEO1 (370-525) (PDB 6TED)
Exiting RNA
Upstream DNA
Shift by 44 Å
Pol II
PAF1 (171-371)
LEO1 (356-555)
Downstream
DNA
3'
5'
N
Exiting RNA
Upstream DNA
Pol II
LEO1 (513-555)
C
SPT5 (163-268)
SPT4
DSIF NGN
O
SPT6 tSH2 (1321-1514)
Pol II
CDC73 Ras-like domain (360-531)
P
Pol II
Active site
TFIIS domain II
RNA
Metal A
TFIIS Zn finger
Direction of transcription
Q
DSIF NGN
SPT4
ELOF1
ELOF1
(PDB 8B3D)
SPT5 (163-268)
Pol II
RPB2 lobe (199-391)
Downstream DNA
R
DSIF NGN
Pol II
ELOF1
90°
Transcription bubble
N
Direction of transcription
S
5'
3'
3'
5'
IWS1
RPB5 helix
N
RPB5
T
SPT5
KOW2-3
SPT6
N
C
C
IWS1
HEAT/TND
N
RPB1 clamp
U
SPT4
SPT5
IWS1
HEAT/TND
C
C
N
ELOF1
SPT6
IWS1 DSIF NGN helix
N
V
RPB2
lobe
RPB2
protrusion
N
IWS1
lobe loop
ELOF1
C
W
IWS1 ELOF1 belt
N
C
ELOF1
X
N
RPB1
jaw
IWS1
jaw helix
C
RPB1 cleft
Y
RPB1
clamp
N
C
IWS1 cleft tail
RPB1 cleft

Figure S8. Structural validation and cryo-EM densities of Pol II-DSIF-SPT6-PAF1c-IWS1-ELOF1-LEDGF-TFIIS-nucleosome complex. (A) Two views of transparent composite map of Pol II-DSIF-SPT6-PAF1c-IWS1-ELOF1-TFIIS activated elongation complex overlaid with atomic model. **(B)** RPB1 bridge helix, RNA, and template DNA of Pol II active site atoms overlaid on map H. Pol II is stalled at bp -9 with 3 nucleotides of backtracked RNA resolved. **(C)** RNA, template, and nontemplate DNA cartoons and atoms overlaid on map H. Single-stranded nontemplate backbone shown as cartoon. 3 nucleotides of backtracked RNA are resolved with 2 nucleotides of backtracked RNA unresolved. **(D)** LEDGF PWWP (cartoons) bound to promoter distal H3 and transcribed nucleosome (atoms) overlaid on map Q. **(E)** SPT6 core (residues 267-1322) cartoons overlaid on map O. **(F)** RPB1 funnel helices (residues 696-760) cartoons and atoms overlaid on map A. **(G)** β-sheets of RPB2 protrusion domain (residues 40-153) cartoons and atoms overlaid on map E. **(H)** RPB4 and RPB7 cartoons and atoms overlaid on map O. **(I)** TFIIS domain II, helix, and Zn finger shown as cartoons overlaid on map D. **(J)** CTR9 residues 250-850 encompassing TPRs and part of the trestle helix and WDR61 cartoons overlaid on map K. **(K)** CDC73 Ras-like domain cartoons overlaid on map J. **(L)** PAF1/LEO1 dimer cartoons overlaid on map M. **(M)** PAF1 and LEO1 from activated elongation complex model (PDB 6TED) overlaid on PAF1 and LEO1 model (this study) shown as cartoons. Pol II shown in surface representation. Distance of 44 Å measured between β carbon of Asp 285 on PAF1 models. **(N)** LEO1 C-terminal helix positioned next to upstream DNA and DSIF NGN shown as cartoons. Model overlaid on map N which was gaussian-filtered to 2 standard deviations. **(O)** Positions of SPT6 tSH2 domain and CDC73 Ras-like domain with respect to Pol II shown as cartoons. Model overlaid on map J which was gaussian-filtered to 2 standard deviations. **(P)** TFIIS domain II, helix, and Zn finger position with respect to Pol II active site. Pol II shown as transparent surface. **(Q)** ELOF1 from transcription coupled nucleotide excision repair complex (PDB 8B3D) overlaid on ELOF1 model (this study) and map B. ELOF1 position relative to DSIF NGN and RPB2 lobe shown. **(R)** ELOF1 N-terminus reaches toward transcription bubble. Model overlaid on map B. **(S)** IWS1 interaction with RPB5 colored on map I. **(T)** IWS1 HEAT/TND and interactions with SPT5 and SPT6 colored on map B. **(U)** IWS1 interaction with DSIF NGN and ELOF1 colored on map B. **(V)** IWS1 interaction with RPB2 lobe colored on map B. **(W)** IWS1 interaction with ELOF1 colored on map B. **(X)** IWS1 interaction with RPB1 jaw colored on map C. **(Y)** IWS1 interaction with RPB1 cleft colored on map C.

Figure S9

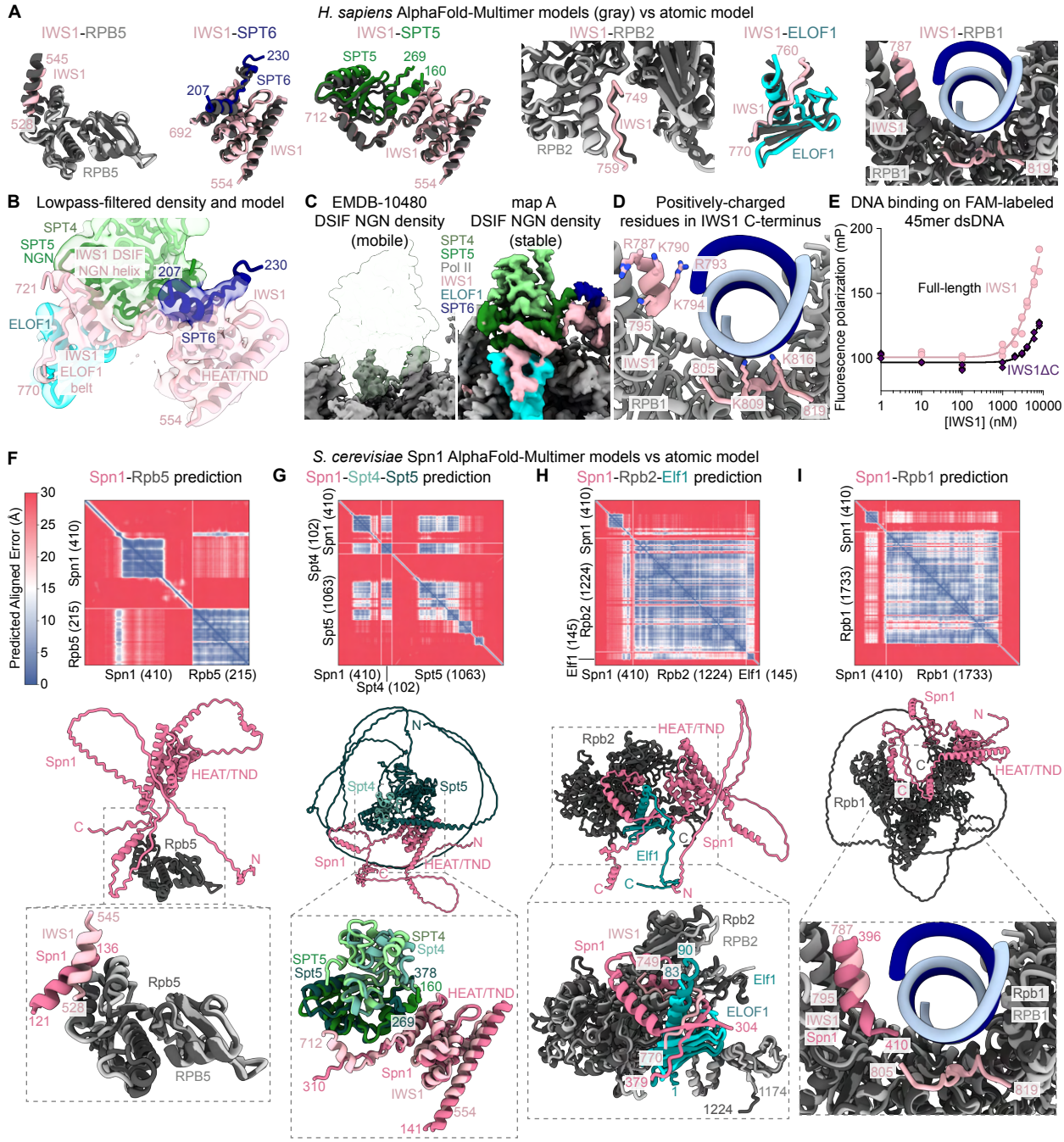
A
H. sapiens AlphaFold-Multimer models (gray) vs atomic model
IWS1-RPB5
IWS1-SPT6
IWS1-SPT5
IWS1-RPB2
IWS1-ELOF1
IWS1-RPB1
B Lowpass-filtered density and model
C EMDB-10480
DSIF NGN density (mobile)
map A
DSIF NGN density (stable)
D Positively-charged residues in IWS1 C-terminus
E DNA binding on FAM-labeled 45mer dsDNA
Fluorescence polarization (mP)
Full-length IWS1
IWS1ΔC
[IWS1] (nM)
F
S. cerevisiae Spn1 AlphaFold-Multimer models vs atomic model
Spn1-Rpb5 prediction
G Spn1-Spt4-Spt5 prediction
H Spn1-Rpb2-Elf1 prediction
I Spn1-Rpb1 prediction
Predicted Aligned Error (Å)
Spn1 (410)
Rpb5 (215)
Spt4 (102)
Spt5 (1063)
Rpb2 (1224)
Elf1 (145)
Rpb1 (1733)

Figure S9. Comparisons of AlphaFold models to atomic model and additional structural details. (A) AlphaFold-Multimer predicted models (gray) for IWS1-RPB5, IWS1-SPT6, IWS1-SPT5, IWS1-RPB2, IWS1-ELOF1, and IWS1-RPB1 interactions overlaid on atomic model. **(B)** Atomic model overlaid on map B gaussian-filtered to 1.5 standard deviation. Density from gaussian-filtered map allows placement of SPT6 residues 219-230 and IWS1 residues 713-721. **(C)** DSIF NGN density comparison between Pol II-DSIF-PAF-SPT6 activated elongation complex EMDB-10480 and map A. Density threshold set based on similarity of visible features of RNA polymerase II. **(D)** Positively charged residues (R787, K790, R793, K794, K809, K816) in IWS1 RPB1 jaw helix and RPB1 cleft tail shown as sticks. **(E)** Fluorescence DNA polarization assays of wild-type IWS1 and IWS1ΔC. Fluorescence polarization (mP) was measured for a 10 nM FAM-labeled 45mer dsDNA substrate. Nonlinear regression curve fit to data points. IC50 for IWS1 and IWS1ΔC calculated as 6580 and 7398, respectively. **(F)** PAE plot of *S. cerevisiae* Spn1-Rpb5 AlphaFold-Multimer prediction shown and representative predicted model overlaid with atomic model of human IWS1 and RPB5 from this study. **(G)** PAE plot of *S. cerevisiae* Spn1-Spt4-Spt5 AlphaFold-Multimer prediction shown and representative predicted model overlaid with atomic model of human IWS1, SPT4, and SPT5 from this study. **(H)** PAE plot of *S. cerevisiae* Spn1-Rpb2-Elf1 AlphaFold-Multimer prediction shown and representative predicted model overlaid with atomic model of human IWS1, RPB2, and ELOF1 from this study. **(I)** PAE plot of *S. cerevisiae* Spn1-Rpb1 AlphaFold-Multimer prediction shown and representative predicted model overlaid with atomic model of human IWS1 and RPB1 from this study.

Figure S10

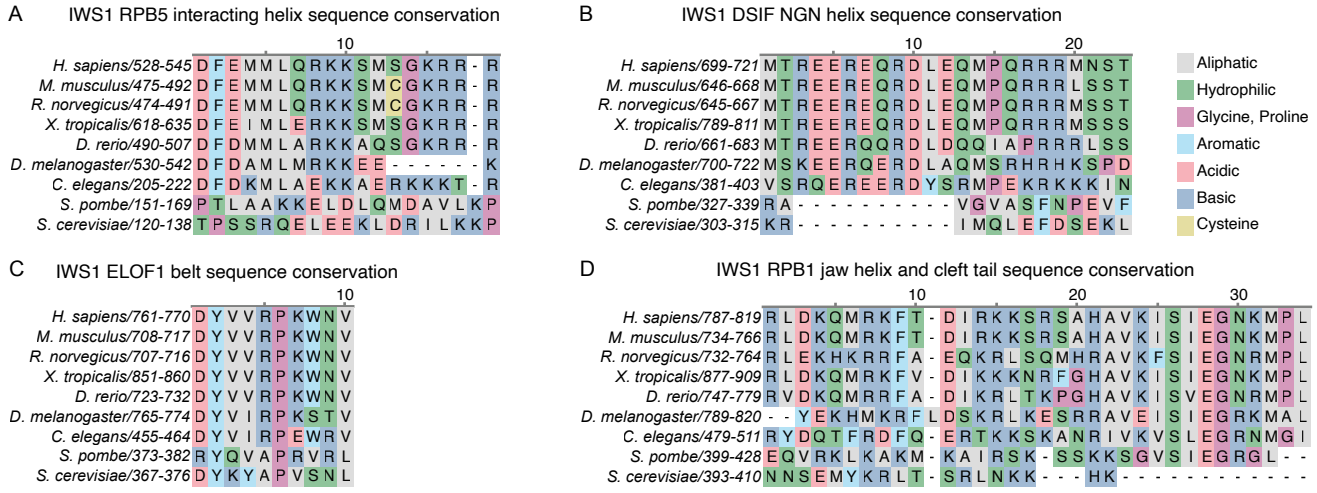
A
IWS1 RPB5 interacting helix sequence conservation
10
H. sapiens/528-545 D F E M M L Q R K K S M S G K R R - R
M. musculus/475-492 D F E M M L Q R K K S M C G K R R - R
R. norvegicus/474-491 D F E M M L Q R K K S M C G K R R - R
X. tropicalis/618-635 D F E I M L E R K K S M S G K R R - R
D. rerio/490-507 D F D M M L A R K K A Q S G K R R - R
D. melanogaster/530-542 D F D A M L M R K K E E - - - - - - K
C. elegans/205-222 D F D K M L A E K K A E R K K K T - R
S. pombe/151-169 P T L A A K K E L D L Q M D A V L K P
S. cerevisiae/120-138 T P S S R Q E L E E K L D R I L K K P
B
IWS1 DSIF NGN helix sequence conservation
10
20
H. sapiens/699-721 M T R E E R E Q R D L E Q M P Q R R R M N S T
M. musculus/646-668 M T R E E R E Q R D L E Q M P Q R R R L S S T
R. norvegicus/645-667 M T R E E R E Q R D L E Q M P Q R R R M S S T
X. tropicalis/789-811 M T R E E R E Q R D L E Q M P Q R R R M S S S
D. rerio/661-683 M T R E E R Q Q R D L D Q Q I A P R R R L S S
D. melanogaster/700-722 M S K E E R Q E R D L A Q M S R H R H K S P D
C. elegans/381-403 V S R Q E R E E R D Y S R M P E K R K K K I N
S. pombe/327-339 R A - - - - - - - - - - V G V A S F N P E V F
S. cerevisiae/303-315 K R - - - - - - - - - - I M Q L E F D S E K L
Aliphatic
Hydrophilic
Glycine, Proline
Aromatic
Acidic
Basic
Cysteine
C
IWS1 ELOF1 belt sequence conservation
10
H. sapiens/761-770 D Y V V R P K W N V
M. musculus/708-717 D Y V V R P K W N V
R. norvegicus/707-716 D Y V V R P K W N V
X. tropicalis/851-860 D Y V V R P K W N V
D. rerio/723-732 D Y V V R P K W N V
D. melanogaster/765-774 D Y V I R P K S T V
C. elegans/455-464 D Y V I R P E W R V
S. pombe/373-382 R Y Q V A P R V R L
S. cerevisiae/367-376 D Y K Y A P V S N L
D
IWS1 RPB1 jaw helix and cleft tail sequence conservation
10
20
30
H. sapiens/787-819 R L D K Q M R K F T - D I R K K S R S A H A V K I S I E G N K M P L
M. musculus/734-766 R L D K Q M R K F T - D I R K K S R S A H A V K I S I E G N K M P L
R. norvegicus/732-764 R L E K H K R R F A - E Q K R L S Q M H R A V K F S I E G N R M P L
X. tropicalis/877-909 R L D K Q M R K F V - D I K K K N R F G H A V K I S I E G N K M P L
D. rerio/747-779 R V D K Q M R R F A - D I K R L T K P G H A V K I S V E G N R M P L
D. melanogaster/789-820 - - Y E K H M K R F L D S K R L K E S R R A V E I S I E G R K M A L
C. elegans/479-511 R Y D Q T F R D F Q - E R T K K S K A N R I V K V S L E G R N M G I
S. pombe/399-428 E Q V R K L K A K M - K A I R S K - S S K K S G V S I E G R G L - -
S. cerevisiae/393-410 N N S E M Y K R L T - S R L N K K - - - H K - - - - - - - - - - - -

Figure S10. IWS1 sequence alignments. PROMALS3D multiple sequence alignments demonstrate high conservation of **(A)** IWS1 RPB5 helix (residues 528-545), **(B)** IWS1 DSIF NGN helix (699-721), **(C)** IWS1 ELOF1 belt (residues 761-770), and **(D)** IWS1 RPB1 jaw helix and cleft tail (residues 787-819).

Figure S11

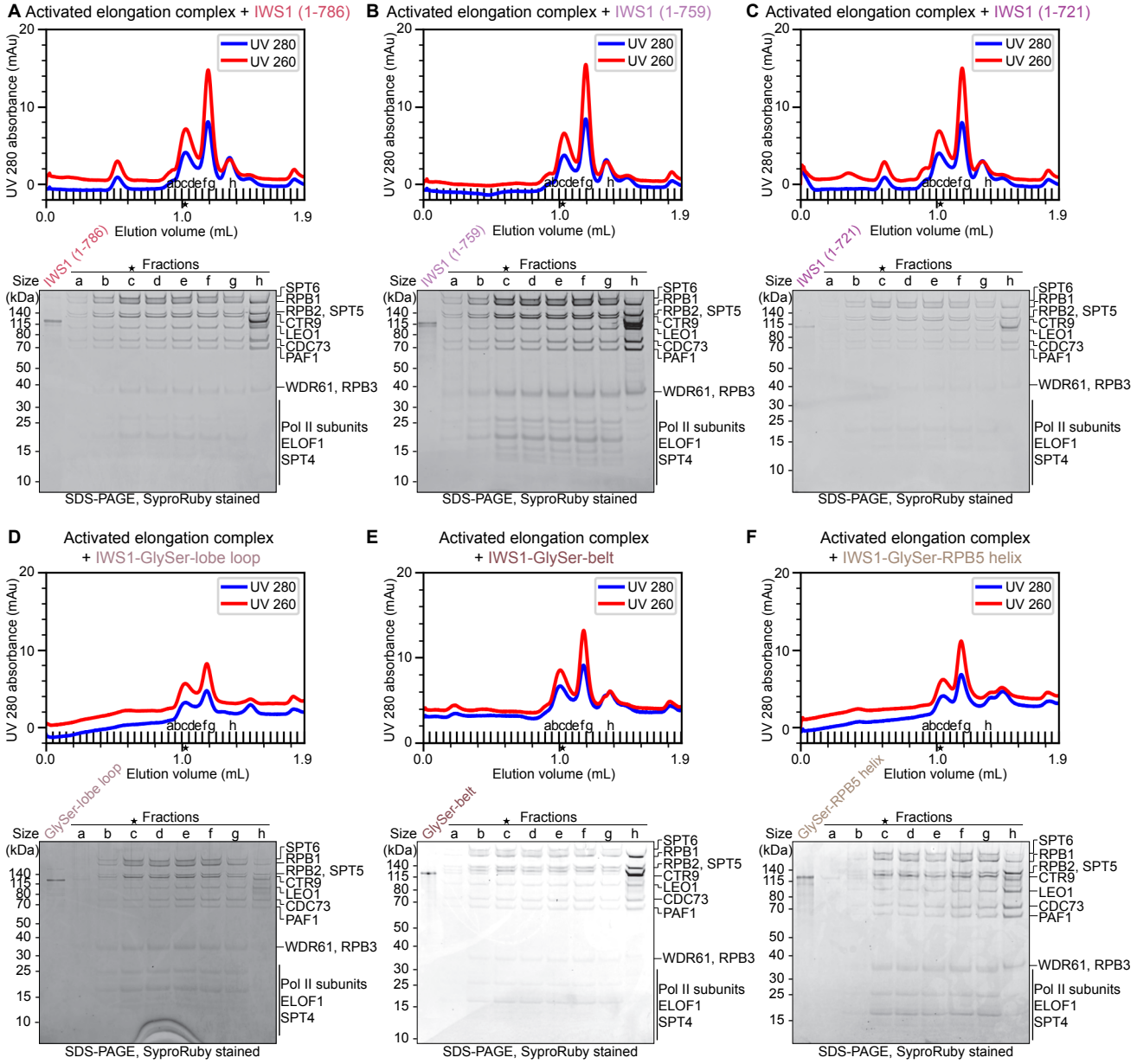
A Activated elongation complex + IWS1 (1-786)
B Activated elongation complex + IWS1 (1-759)
C Activated elongation complex + IWS1 (1-721)
D Activated elongation complex + IWS1-GlySer-lobe loop
E Activated elongation complex + IWS1-GlySer-belt
F Activated elongation complex + IWS1-GlySer-RPB5 helix
UV 280
UV 260
UV 280 absorbance (mAu)
Elution volume (mL)
abcdefg h
Fractions
Size (kDa)
a b c d e f g h
SPT6
RPB1
RPB2, SPT5
CTR9
LEO1
CDC73
PAF1
WDR61, RPB3
Pol II subunits
ELOF1
SPT4
SDS-PAGE, SyproRuby stained

Figure S11. Full size exclusion chromatograms and SDS-PAGE related to Figure 4. UV 280 nm and UV 260 nm traces for size exclusion chromatograms and corresponding SDS-PAGE of elongation complexes assembled with **(A)** IWS1 (1-786), **(B)** IWS1 (1-759), **(C)** IWS1 (1-721), **(D)** IWS1-GlySer-lobe loop, **(E)** IWS1-GlySer-belt, and **(F)** IWS1-GlySer-RPB5 helix. Activated elongation complex = Pol II, DSIF, SPT6, PAF1c without RTF1, ELOF1. Fractions were loaded onto 4-12% Bis-Tris gels, run in 1x MES buffer, and stained with SyproRuby. Fraction displayed in Figure 4 indicated with star.

Figure S12

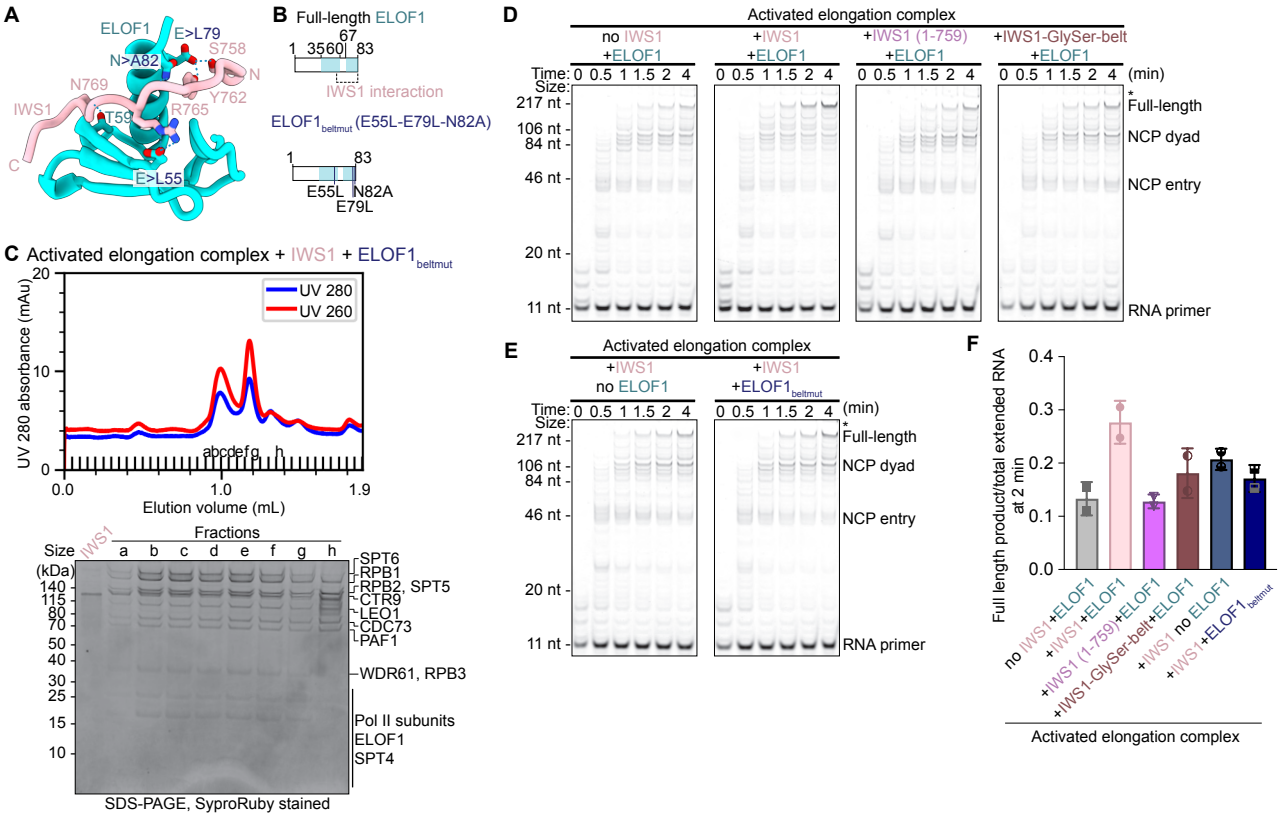
A
ELOF1
E>L79
N>A82
S758
N
N769
Y762
IWS1
T59
R765
C
E>L55
B Full-length ELOF1
67
1 35 60 83
IWS1 interaction
ELOF1beltmut (E55L-E79L-N82A)
1 83
E55L N82A
E79L
C Activated elongation complex + IWS1 + ELOF1beltmut
UV 280
UV 260
UV 280 absorbance (mAu)
Elution volume (mL)
abcdefg h
Fractions
Size (kDa)
IWS1
a b c d e f g h
SPT6
RPB1
RPB2, SPT5
CTR9
LEO1
CDC73
PAF1
WDR61, RPB3
Pol II subunits
ELOF1
SPT4
SDS-PAGE, SyproRuby stained
D
Activated elongation complex
no IWS1 +ELOF1
+IWS1 +ELOF1
+IWS1 (1-759) +ELOF1
+IWS1-GlySer-belt +ELOF1
Time: 0 0.5 1 1.5 2 4 (min)
Size:
217 nt
106 nt
84 nt
46 nt
20 nt
11 nt
Full-length
NCP dyad
NCP entry
RNA primer
E
Activated elongation complex
+IWS1 no ELOF1
+IWS1 +ELOF1beltmut
F
Full length product/total extended RNA at 2 min
no IWS1+ELOF1
+IWS1+ELOF1
+IWS1 (1-759)+ELOF1
+IWS1-GlySer-belt+ELOF1
+IWS1 no ELOF1
+IWS1+ELOF1beltmut
Activated elongation complex

Figure S12. Size exclusion chromatography and RNA extension assays with ELOF1$_{beltmut}$. (A) Interaction between IWS1 ELOF1 belt and ELOF1. Residues mutated in ELOF1$_{beltmut}$ indicated. Interacting residues shown as sticks and H-bonds shown as dotted lines. **(B)** Domain architecture of ELOF1 and ELOF1$_{beltmut}$. **(C)** UV 280 nm and UV 260 nm traces for size exclusion chromatograms and corresponding SDS-PAGE of elongation complexes assembled with full-length IWS1 and ELOF1$_{beltmut.}$ Complex does not contain RTF1 or TFIIS. Fractions were loaded onto 4-12% Bis-Tris gels, run in 1x MES buffer, and stained with SyproRuby. **(D)** RNA extension assays of IWS1 mutants on mononucleosome. Samples were separated by denaturing gel electrophoresis. RNA extension was monitored using FAM dye on RNA. Initial RNA and extended products are marked. Results are representative of at least three independent experiments. **(E)** RNA extension assays with full-length IWS1 and no ELOF1 or ELOF1$_{beltmut}$ on mononucleosome. Samples were separated by denaturing gel electrophoresis. RNA extension was monitored using FAM dye on RNA. Initial RNA and extended products are marked. Results are representative of at least three independent experiments. **(F)** Quantification of D and E from n = 2 independent experiments. Individual data points are shown. Activated elongation complex in D, E, and F = Pol II, DSIF, SPT6, PAF1c including RTF1, TFIIS. Columns represent mean values across n = 2 independent experiments and error bars represent standard deviations.

Figure S13

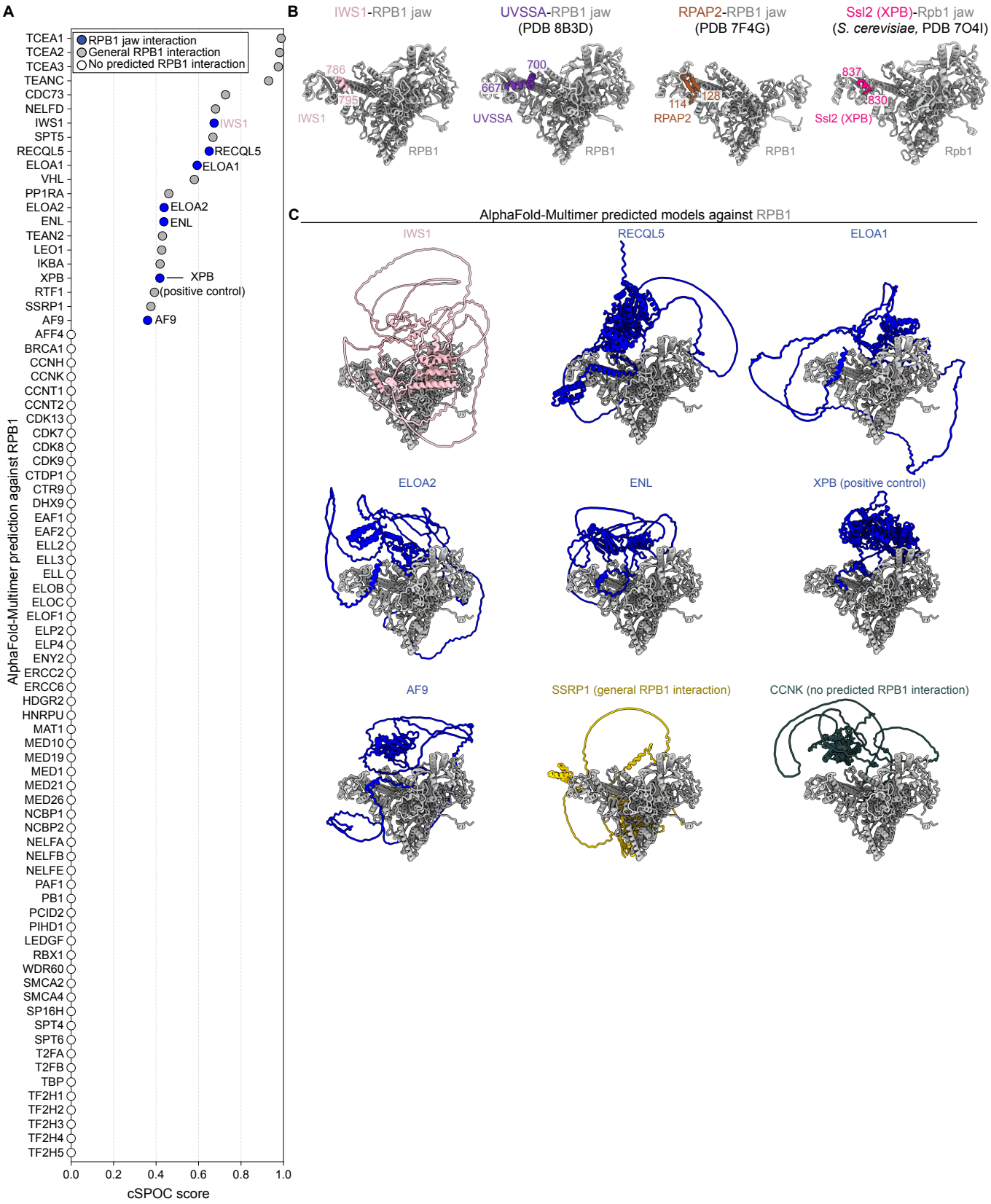
A
RPB1 jaw interaction
General RPB1 interaction
No predicted RPB1 interaction
TCEA1
TCEA2
TCEA3
TEANC
CDC73
NELFD
IWS1
SPT5
RECQL5
ELOA1
VHL
PP1RA
ELOA2
ENL
TEAN2
LEO1
IKBA
XPB
RTF1
SSRP1
AF9
AFF4
BRCA1
CCNH
CCNK
CCNT1
CCNT2
CDK13
CDK7
CDK8
CDK9
CTDP1
CTR9
DHX9
EAF1
EAF2
ELL2
ELL3
ELL
ELOB
ELOC
ELOF1
ELP2
ELP4
ENY2
ERCC2
ERCC6
HDGR2
HNRPU
MAT1
MED10
MED19
MED1
MED21
MED26
NCBP1
NCBP2
NELFA
NELFB
NELFE
PAF1
PB1
PCID2
PIHD1
LEDGF
RBX1
WDR60
SMCA2
SMCA4
SP16H
SPT4
SPT6
T2FA
T2FB
TBP
TF2H1
TF2H2
TF2H3
TF2H4
TF2H5
(positive control)
AlphaFold-Multimer prediction against RPB1
0.0
0.2
0.4
0.6
0.8
1.0
cSPOC score
B
IWS1-RPB1 jaw
UVSSA-RPB1 jaw
(PDB 8B3D)
RPAP2-RPB1 jaw
(PDB 7F4G)
Ssl2 (XPB)-Rpb1 jaw
(S. cerevisiae, PDB 7O4I)
786
795
IWS1
RPB1
700
667
UVSSA
RPB1
114
128
RPAP2
RPB1
837
830
Ssl2 (XPB)
Rpb1
C
AlphaFold-Multimer predicted models against RPB1
IWS1
RECQL5
ELOA1
ELOA2
ENL
XPB (positive control)
AF9
SSRP1 (general RPB1 interaction)
CCNK (no predicted RPB1 interaction)

Figure S13. RPB1 interaction screen. (A) All proteins screened against RPB1 using AlphaFold-Multimer. Predictions ranked by cSPOC score. Representative RPB1 jaw interactors are shown as dark blue circles, general RPB1 interactions shown as grey circles, no predicted RPB1 interaction shown as white circles. **(B)** Models of known RPB1 jaw interactors, including IWS1 (this study), UVSSA (PDB 8B3D), RPAP2 (PDB 7F4G), and yeast Ssl2 (human XPB) (PDB 7O4I). **(C)** AlphaFold-Multimer predicted models of representative candidates with RPB1. RPB1 jaw interactors colored in blue with IWS1 colored in pink. Representative general RPB1 interaction SSRP1 colored in gold. Representative non-interacting example (CCNK) colored in dark slate gray.

Figure S14

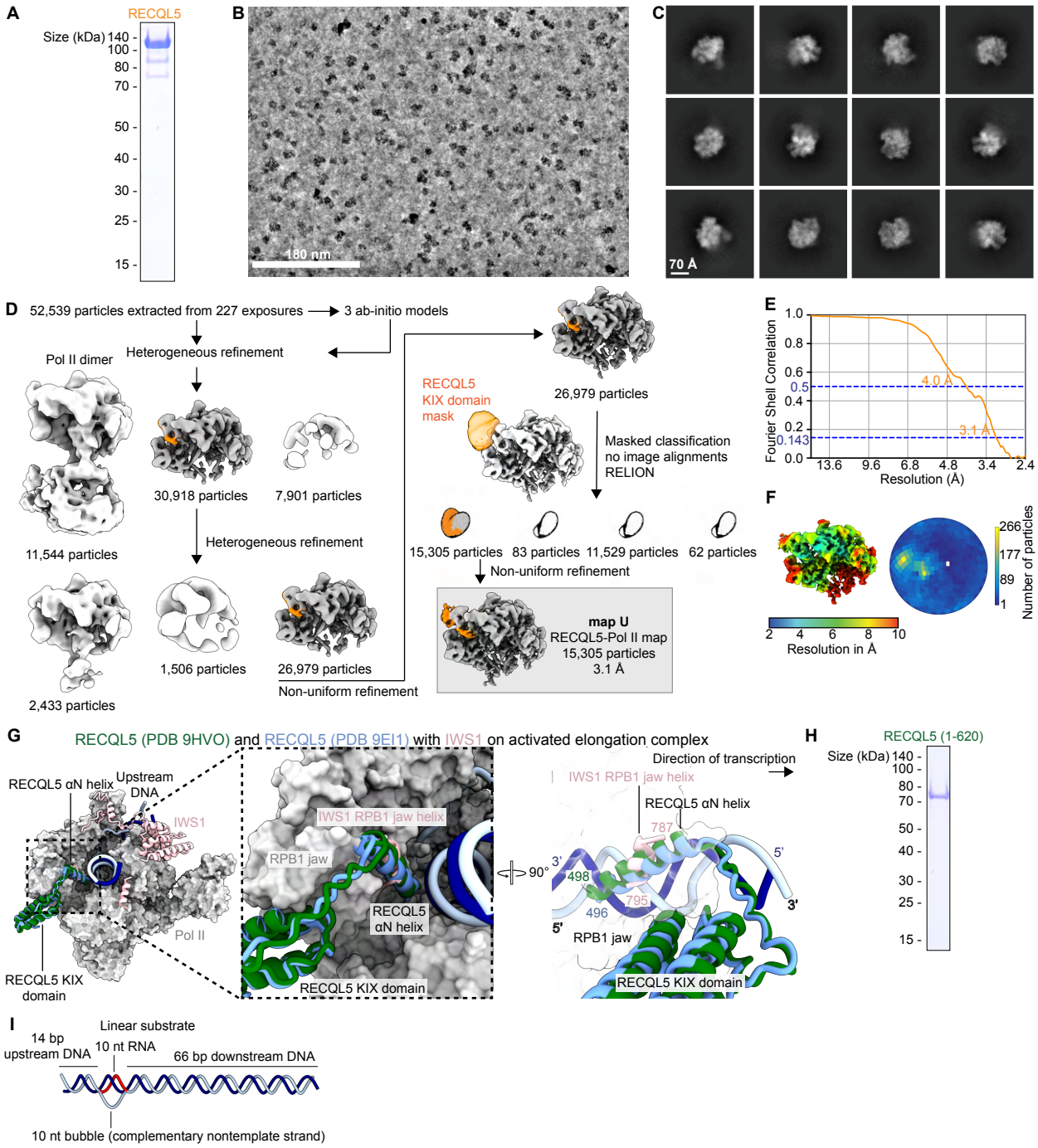
A
RECQL5
Size (kDa) 140 -
100 -
80 -
70 -
50 -
40 -
30 -
25 -
15 -
B
180 nm
C
70 Å
D
52,539 particles extracted from 227 exposures
3 ab-initio models
Heterogeneous refinement
Pol II dimer
11,544 particles
30,918 particles
7,901 particles
Heterogeneous refinement
2,433 particles
1,506 particles
26,979 particles
Non-uniform refinement
26,979 particles
RECQL5
KIX domain
mask
Masked classification
no image alignments
RELION
15,305 particles
83 particles
11,529 particles
62 particles
Non-uniform refinement
map U
RECQL5-Pol II map
15,305 particles
3.1 Å
E
Fourier Shell Correlation
1.0
0.8
0.6
0.5
0.4
0.2
0.143
0.0
4.0 Å
3.1 Å
13.6
9.6
6.8
4.8
3.4
2.4
Resolution (Å)
F
266
177
89
1
Number of particles
2
4
6
8
10
Resolution in Å
G
RECQL5 (PDB 9HVO) and RECQL5 (PDB 9EI1) with IWS1 on activated elongation complex
RECQL5 αN helix
Upstream
DNA
IWS1
Pol II
RECQL5 KIX
domain
IWS1 RPB1 jaw helix
RPB1 jaw
RECQL5
αN helix
RECQL5 KIX domain
90°
Direction of transcription
IWS1 RPB1 jaw helix
RECQL5 αN helix
787
3'
498
795
496
5'
5'
3'
RPB1 jaw
RECQL5 KIX domain
H
RECQL5 (1-620)
Size (kDa) 140 -
100 -
80 -
70 -
50 -
40 -
30 -
25 -
15 -
I
Linear substrate
14 bp
upstream DNA
10 nt RNA
66 bp downstream DNA
10 nt bubble (complementary nontemplate strand)

Figure S14. RECQL5-Pol II complex cryo-EM data processing and constructs used in RECQL5+IWS1 RNA extension assay. (A) SDS-PAGE of purified RECQL5. Gel stained with OneStep Blue. **(B)** Representative denoised micrograph from RECQL5-Pol II sample cryo-EM data collection. Scale bar of 180 nm. **(C)** Representative 2D classes from data processing with scale bar of 70 Å. **(D)** Sorting and classification of cryo-EM data. All processing performed in cryoSPARC unless indicated. Densities colored according to model in Figure 5C. **(E)** FSC curve for RECQL5-Pol II map. Resolution at FSC 0.5 and 0.143 indicated. **(F)** Local resolution estimation and angular distribution plot for RECQL5-Pol II map. **(G)** Overlay of RECQL5 αN helix and KIX domain from RECQL5-Pol II models (PDB 9HVO, PDB 9EI1) on activated elongation complex model (this study) shows binding overlap of RECQL5 αN helix and IWS1 RPB1 jaw helix. Pol II model shown as surfaces and IWS1 model shown as cartoons. **(H)** SDS-PAGE of purified RECQL5 (1-620). Gel stained with OneStep Blue. **(I)** Schematic of linear substrate used in RECQL5+IWS1 RNA extension assay. Template strand colored in navy, non-template strand colored in light steel blue, and 10 nt RNA primer colored in red.

Figure S15

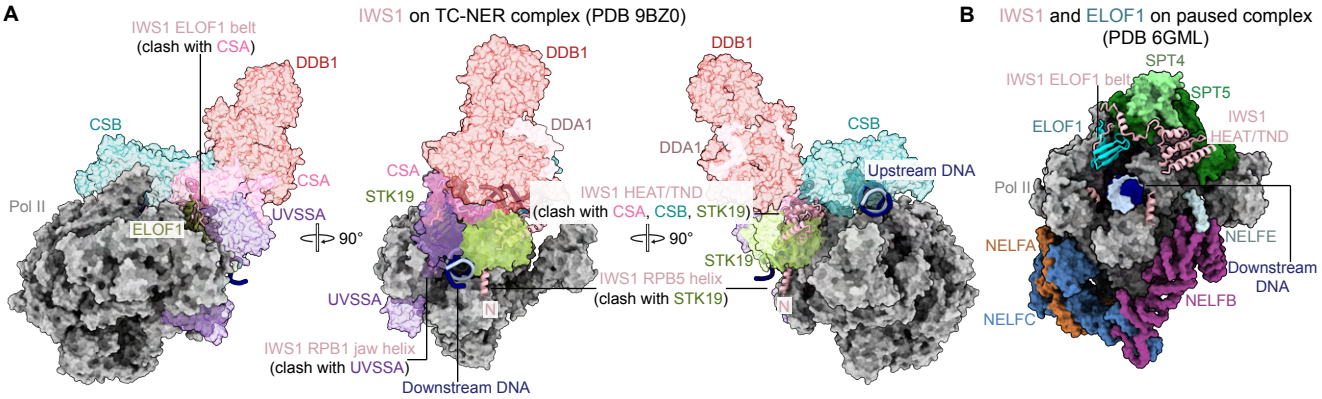
A
IWS1 ELOF1 belt
(clash with CSA)
DDB1
CSB
CSA
Pol II
ELOF1
UVSSA
90°
IWS1 on TC-NER complex (PDB 9BZ0)
DDB1
DDA1
CSA
STK19
IWS1 HEAT/TND
(clash with CSA, CSB, STK19)
UVSSA
N
IWS1 RPB1 jaw helix
(clash with UVSSA)
Downstream DNA
90°
DDB1
DDA1
CSB
Upstream DNA
STK19
IWS1 RPB5 helix
(clash with STK19)
N
B
IWS1 and ELOF1 on paused complex
(PDB 6GML)
SPT4
IWS1 ELOF1 belt
SPT5
ELOF1
IWS1
HEAT/TND
Pol II
NELFA
NELFE
Downstream
DNA
NELFB
NELFC

Figure S15. IWS1 and ELOF1 compatibility with known transcription complexes. **(A)** Three views of IWS1 model shown as cartoons overlaid on TC-NER+STK19 complex (PDB 9BZ0). 9BZ0 ELOF1 shown in cartoons and the remaining TC-NER+STK19 factors shown as transparent surfaces. Clashes are indicated: IWS1 RPB5 helix with STK19, IWS1 HEAT/TND with CSA, CSB, and STK19, IWS1 ELOF1 belt with CSA, and IWS1 RPB1 jaw helix with UVSSA. **(B)** IWS1 and ELOF1 models shown as cartoons and overlaid on mammalian paused elongation complex (PDB 6GML). Positions of IWS1 and ELOF1 do not clash with NELF.