UC San Diego
UC San Diego Electronic Theses and Dissertations

Title

Structural Basis for Lesion Recognition and Commitment to Transcription-Coupled Repair

Permalink

https://escholarship.org/uc/item/0dx3f053

Author

Sarsam, Reta D.

Publication Date

2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

Structural Basis for Lesion Recognition and Commitment to Transcription-Coupled Repair

A dissertation submitted in partial satisfaction of the

requirement for the degree Doctor of Philosophy

in

Chemistry

by

Reta D. Sarsam

Committee in charge:

      Professor Andres E. Leschziner, Chair
      Professor Mark Anthony Herzik, Co-Chair
      Professor Thomas C. Hermann
      Professor James T Kadonaga
      Professor Dong Wang

2021

The dissertation of Reta D. Sarsam is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

This work is dedicated to my husband, Kedis, who has been enduring our living apart for the sake of realizing my dream to obtain my doctorate in science. This work is also dedicated to my parents, Mr. and Mrs. Dawood Sarsam who have departed their homeland aspiring to a better future for our family.

EPIGRAPH

A dream doesn't become reality through magic;

it takes sweat, determination, and hard work.

Colin Powell

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SUPPLEMENTARY FIGURES AND TABLES

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my advisor, Prof. Andres Leschziner for the support and guidance that he has given me over the past years. You have been a truly great mentor.

Besides my advisor, I would like to acknowledge and thank Prof. Dong Wang for his collaboration and supervision of this work. Also, I would like to acknowledge the post-docs: Dr. Jun Xu and Dr. Indrajit Lahiri for all their effort they put into this research.

Finally, I am thankful to Prof. Mark Herzik, Prof. Thomas Herman and Prof. James Kadonaga for their encouragement and insightful comments.

Chapter 1, in full, has been submitted for publication of the material, Reta D Sarsam, Jun Xu, Indrajit Lahiri, Juntaek Oh, Zhen Zhou, Jenny Chong, Nan Hao, Dong Wang, and Andres E. Leschziner. The dissertation author was a co-author of this paper.

VITA

2013 Bachelor of Science, University of California Davis

2013 Junior Specialist, University of California Davis

2017 Master of Science, University of California San Diego

2021 Doctor of Philosophy, University of California San Diego

PUBLICATIONS

Yan, L., Qi, Y., Ricketson, D., Li, L., Subramanian, K., Zhao, J., Yu, C., Wu, L., **Sarsam, R**., Wong, M. and Lou, Z., 2020. Structural analysis of a trimeric assembly of the mitochondrial dynamin-like GTPase Mgm1. *Proceedings of the National Academy of Sciences*, *117*(8), pp.4061-4070.

Erickson, A.I., **Sarsam, R.D.** and Fisher, A.J., 2015. Crystal structures of Mycobacterium tuberculosis CysQ, with substrate and products bound. *Biochemistry*, *54*(45), pp.6830-6841.

Murley, A., **Sarsam, R.D.**, Toulmay, A., Yamada, J., Prinz, W.A. and Nunnari, J., 2015. Ltc1 is an ER-localized sterol transporter and a component of ER–mitochondria and ER–vacuole contacts. *Journal of Cell Biology*, *209*(4), pp.539-548.

Erickson, A., **Sarsam, R.** and Fisher, A.J., 2014. Expression, purification and preliminary crystallographic analysis of Mycobacterium tuberculosis CysQ, a phosphatase involved in sulfur metabolism. *Acta Crystallographica Section F: Structural Biology Communications*, *70*(6), pp.750-753.

ABSTRACT OF THE DISSERTATION

Structural Basis for Lesion Recognition and Commitment to Transcription-Coupled Repair

by

Reta D. Sarsam

Doctor of Philosophy in Chemistry

University of California San Diego, 2021

Professor Andres E. Leschziner, Chair

Professor Mark Anthony Herzik, Co-Chair

Transcription-coupled nucleotide excision repair (TC-NER) is a highly conserved pathway that removes bulky lesions in the transcribed genome. Cockayne syndrome B protein (CSB), or its

yeast ortholog Rad26, plays important roles in the lesion-recognition steps of TC-NER. How Rad26 distinguishes between RNA polymerase II (Pol II) stalled at a DNA lesion or other obstacles, how a lesion-arrested Pol II is committed to the recruitment of downstream repair factors, and what the fate is of a lesion-arrested Pol II remain unknown. Here, we present cryo-EM structures of Pol II-Rad26 complexes stalled at different obstacles to establish a universal mechanism for the Rad26-mediated recognition of stalled Pol II. We also present a 3.1Å cryo-EM structure of lesion-arrested Pol II-Rad26 bound to a newly identified TC-NER factor, ELOF1/Elf1, that provides insights into its role in the commitment of lesion-arrested Pol II to TC-NER. Finally, we provide biochemical data revealing how Rad26 displaces a lesion-stalled Pol II during TC-NER. These results establish the structural basis of lesion-recognition, commitment to repair, and displacement of lesion-arrested Pol II during TC-NER.

CHAPTER 1: STRUCTURAL BASIS FOR LESION RECOGNITION AND COMMITMENT TO TRANSCRIPTION-COUPLED REPAIR

INTRODUCTION

Transcription-coupled DNA nucleotide excision repair (TC-NER), a highly conserved sub-pathway of nucleotide excision repair, is the first line of defense that detects and removes a broad spectrum of transcription-blocking lesions in the transcribed genome (1-7). Cockayne syndrome group B (CSB) protein, or its ortholog Rad26 in *Saccharomyces cerevisiae*, a member of the Swi2/Snf2 family of nucleosome remodeling helicases/ATPases, plays a crucial early role in eukaryotic TC-NER *(1-7)*. During the lesion recognition steps of TC-NER, CSB/Rad26 distinguishes the lesion-arrested Pol II from other types of arrested Pol II and facilitates subsequent recruitment of downstream repair factors, including CSA, UVSSA, and TFIIH *(1, 4-6, 8)*. In addition to its role in TC-NER, CSB/Rad26 also functions as a processivity factor for Pol II arrested in the absence of DNA damage and regulates a subset of genes crucial for neurological differentiation and development (8-12). Mutations in CSB are linked to Cockayne syndrome, a severe neurodevelopmental disorder characterized by photosensitivity and premature aging (6, 13).

Several critical steps of TC-NER remain poorly understood. These include: (i) how CSB/Rad26 distinguishes between Pol II stalled at a DNA lesion or other obstacles in order to initiate TC-NER only in the presence of a bulky lesion; (ii) how a lesion-arrested Pol II-CSB/Rad26 complex is committed to the recruitment of downstream repair factors ; and (iii) how the lesion-arrested Pol II is cleared from the site to make it accessible to the repair machinery *(1, 4-6, 8)*. Here, we have used cryo-electron microscopy (cryo-EM) and functional assays to address these gaps in our understanding of TC-NER. We report high-resolution structures of Pol II stalled

at different obstacles, including a CPD lesion, showing that Rad26 uses a universal approach to recognize a stalled Pol II but establishes a pronounced interaction with it in the presence of a lesion. To understand if and how this CPD-induced structure is involved in committing a lesion-arrested Pol II-Rad26 complex to TC-NER, we solved the cryo-EM structure of this complex bound to Elf1, an elongation factor that was only recently identified as a new core Transcription-Coupled Repair (TCR) factor (14). Our structure revealed that the presence of Elf1 both strengthened the CPD-dependent interaction between Rad26 and Pol II and led to new interactions between Rad26 and Pol II absent from all other Rad26-containing structures. Finally, we provide biochemical evidence showing that Rad26, together with TFIIS, can displace a CPD-lesion arrested Pol II *in vitro*, providing a mechanistic framework for the exposure of a DNA lesion during TC-NER.

RESULTS

Rad26 uses the same mode to bind all arrested Pol II elongation complexes

To investigate whether Rad26 interacts with an arrested Pol II differently depending on whether the obstacle is a DNA lesion, we solved cryo-EM structures of Pol II-Rad26 complexes either stalled at a CPD DNA lesion (Pol II(CPD)-Rad26) or containing a transcription scaffold that mimics a backtracked state after arrest at a non-lesion site (Backtracked Pol II-Rad26) (fig. 1, figs. S1-4). In all structures, as it was the case in our previous structure of a Pol II-Rad26 complex stalled at a non-lesion site (by nucleotide deprivation) (8), Rad26 is bound behind the polymerase near the upstream fork of the transcription bubble and interacts with the protrusion and the wall domain of Rpb2, and the clamp coiled-coil of Rpb1. Similarly, the binding of Rad26 bends the upstream DNA by ~80° towards the Pol II stalk (Rpb4/7) in all cases. Thus, Rad26 has a universal mode of interacting with Pol II regardless of the type of arrest (fig. 1).

Rad26 and Rpb4/7 interact more strongly in the Pol II(CPD)-Rad26 stalled complex

Three-dimensional classification of the Pol II(CPD)-Rad26 stalled complex dataset (fig. S1) revealed two states with differences in the interaction between Rad26 and Pol II: one shows well-defined density between Rad26 and Rpb4/7, which we termed the "Strong" interaction (fig. 1E), while the other has much weaker density in the same area, a state we refer to as the "Weak" interaction (fig. 1D). The Strong interaction state is specific to Pol II(CPD)-Rad26 stalled complex and has three main structural features: Rpb4/7 has shifted towards Rad26 (relative to the core of Pol II) (fig. 2A-C); Rad26 has moved towards Rpb4/7, with a concomitant higher bending of the upstream DNA (fig. 2E); and the density connecting Rpb4/7 and Rad26 is stronger (figs. 1E, 2A,B, and fig. S5A,B). The enhanced Rad26-Rpb4/7 interaction and closer proximity of Rpb4/7 to Rad26 do not appear to be general features of all stalled Pol II complexes as we did not observe them in our previous structures of Pol II-Rad26 stalled at a non-lesion site (8) or in our new Backtracked Pol II-Rad26 complex (fig. 1B-C, fig. S5D and fig. S6). In fact, the Rad26-Rpb4/7 interaction is weakest in the Backtracked Pol II-Rad26 state and Rpb4/7 has moved further away from Rad26 (fig. 2C, fig S6).

The Rad26-Rpb4/7 interaction is not required for DNA bending

We previously showed that the upstream DNA duplex is bent by approximately $80°$ in the presence of Rad26 (8). A major question left unresolved by that structure is whether this significant bending required the interaction between Rad26 and Rpb4/7 or was induced solely by Rad26. Although we observed similar bending of the upstream DNA in both the Strong and Weak Rad26-Rpb4/7 states, the latter still shows density, albeit weak, connecting Rad26 and Rpb4/7 (fig. S5C).

To test whether the Rad26-Rpb4/7 interaction is required for the bending of the upstream DNA, we solved a structure of core Pol II(CPD)-Rad26 with 10-subunit Pol II lacking Rpb4/7 (fig. 1F, fig. S7). The DNA was bent to a similar extent in the absence of Rpb4/7, showing that the formation of this structure only requires the binding of Rad26 to DNA and Pol II (Rpb1 and 2). As is the case for the Weak interaction state, the DNA bending in the absence of Rpb4/7 was slightly less pronounced than in the Strong interaction state (fig. 2E,F). Our data showed that Rad26 becomes more ordered upon its interaction with Rpb4/7 despite the latter's intrinsic flexibility (figs. S8-10). Indeed, the local resolution of Rad26 is highest in the presence of a Strong Rad26-Rpb4/7 interaction and lowest in the absence of Rpb4/7 (figs. S2A, S4A, and S7E). Taken together, these observations suggest that the Strong Rad26-Rpb4/7 interaction seen in our Pol II(CPD)-Rad26 structure stabilizes Rad26.

Elf1 induces new interactions between Pol II(CPD) and Rad26

Given the more pronounced interaction between Rad26 and Pol II in the presence of a lesion, we hypothesized that it might be involved in the recruitment of downstream repair factors. Thus, we wondered if other early TC-NER factors might further strengthen this lesion-dependent state. Specifically, we focused on Elf1, the *S.cerevisiae* ortholog of human ELOF1, a highly conserved transcription elongation factor that was recently identified as a new core TC-NER factor (fig. S11A) (14-17). Loss of ELOF1 in humans or deletion of Elf1 in yeast leads to UV sensitivity (14). Intriguingly, in another genome wide multi-omics analysis of the UV-induced DNA damage response, human ELOF1 was found as a top interactor of human CSB [18]. Based on these studies, we hypothesized that Elf1/ELOF1 could be involved in the initiation of TC-NER by modulating the interaction between Pol II and Rad26/CSB as well as other downstream repair factors. While

yeast Elf1 and human ELOF1 share a highly conserved core domain, Elf1 contains a yeast-specific C-terminus (fig. S11A), which was not observed in a published structure of Pol II-Spt4/5-Elf1 even though full-length protein was used (17). Given this, we generated an Elf1 C-terminal truncation (Elf1ΔC) that mimics human ELOF1 (fig. S11B). Elf1ΔC and WT Elf1 behave similarly in their response to UV damage *in vivo* (fig. S11C). We began by testing whether Elf1 or Elf1ΔC interfered with the stalling of Pol II at a CPD lesion in a transcription assay. Addition of Elf1 or Elf1ΔC had no effect on the stalling pattern of Pol II (fig. 3A). We then tested whether Elf1 or Elf1ΔC have an effect on the Pol II-Rad26 interaction; here, a gel mobility assay showed that both Elf1 and Elf1ΔC promoted the formation of the Pol II(CPD)-Rad26 complex (fig. 3B). To understand the structural basis of Elf1's role in promoting the interaction between Rad26 and Pol II stalled at a CPD lesion, we solved a cryo-EM structure of a Pol II(CPD)-Rad26-Elf1 complex (fig. 3C, figs. S12 and S13). Elf1 was bound in the downstream channel, next to the lobe domain of Rpb2 and bridging the cleft, as previously reported (fig. S11D-F) (16, 17). While there was no direct interaction between Rad26 and Elf1, the presence of Elf1 resulted in a significant improvement in the local resolution of Rad26, to 4Å from 8Å in the Strong state, our second-best map. Most strikingly, the Pol II(CPD)-Rad26-Elf1 complex, which we refer to as the "Engaged state", showed new clear density at the interface between lobe 2 of Rad26 and the wall domain of Rpb2, corresponding to newly structured elements (fig. 3C-E, fig. 13H,I). The flap-loop of Rpb2, which was disordered in the other structures, is folded and interacts directly with Rad26 lobe 2 (fig. 3E). The interaction between Rad26 and Rpb4/7 seen in the Strong state is preserved in this structure (fig. 3D,E). Comparing the bending of the upstream DNA among the Weak, Strong, and Engaged states highlights how Rad26 (and the DNA to which it is bound) shifts towards Pol II as

the interactions between them become stabilized (fig. 3F). This is consistent with the idea that these structures represent steps in the commitment of a lesion-stalled Pol II to TC-NER.

Displacement of backtracked lesion-stalled Pol II requires Rad26 and TFIIS

A lesion-arrested Pol II must eventually be displaced from the DNA for repair factors to access the lesion. In *Escherichia coli*, Mfd displaces Pol II stalled at DNA lesions by directly pushing it forward using its ATP-dependent translocase activity (19, 20). Unlike Mfd, Rad26 alone could not displace Pol II from a template containing a CPD lesion (fig. 4A,B). We tested whether adding TFIIS, which cleaves 3' RNA overhangs from backtracked elongation complexes, could promote Pol II displacement by Rad26. Our data showed that the combination of Rad26 and TFIIS resulted in robust displacement of Pol II from a CPD-containing substrate both in the presence and absence of NTPs (fig. 4A-D, fig. S14). In our assay (fig. S14), Pol II displacement results in transcripts being released from the transcription scaffold and thus detected in the supernatant (the scaffold is pulled down). The supernatant contained shortened transcripts of various lengths resulting from TFIIS-stimulated cleavage, but a relatively small amount of full-length transcripts. Full-length transcripts, on the other hand, were present in the pellet, corresponding to elongation complexes that remain engaged with the DNA template. These data suggest that Pol II cannot be displaced when stalled at the CPD site. Instead, it must backtrack and the 3' end of the RNA must be cleaved before Pol II is displaced by the joint action of TFIIS and Rad26 (fig. 4B-D). Importantly, we also tested whether the translocase activity of Rad26 is necessary for its role in Pol II displacement by using an ATPase-dead Rad26 mutant (K328R). While backtracking could take place in the presence of ATPase-dead Rad26, Pol II displacement was abolished (fig. 4B-D). TFIIS alone cannot displace Pol II even after transcript cleavage, which is consistent with previous

reports (21). Finally, we compared the displacement patterns of either 12-subunit or 10-subunit (i.e. lacking Rpb4/7) Pol II. Our data revealed that Rpb4/7 also plays a role in modulating Pol II backtracking and subsequent displacement (fig. S14). Taken together, these observations suggest that Rad26 cannot displace CPD-stalled Pol II in a manner similar to that used by Mfd. Rather, CPD-arrested Pol II can only be displaced, by the joint action of Rad26 and TFIIS, after one or more backtracking events have taken place.

DISCUSSION

Similarity in binding of eukaryotic and prokaryotic TC-NER factors to arrested polymerase

The similarity among our cryo-EM structures of Rad26 bound to different types of arrested Pol II (including at a DNA lesion) showed that Rad26 uses a universal binding-and-scanning mechanism to engage Pol II regardless of the type of arrest. Rad26 binds to the upstream DNA of a stalled Pol II, where its ATP-dependent translocation pulls DNA away from Pol II, effectively biasing Pol II forward. Rad26 facilitates the transcription bypass of non-covalent barriers or small DNA lesions by Pol II, but it remains more stably bound to a Pol II arrested at a bulky DNA lesion. Intriguingly, the structures of stalled RNAP-Mfd complexes, the bacterial counterpart of arrested Pol II-Rad26, revealed a similar upstream loading for Mfd (20). Furthermore, both Rad26 and Mfd pull DNA out of the RNA polymerase to bias the polymerase forward in an ATP-dependent manner to scan for lesions. Interestingly, no structural similarity underlies this functional similarity. A major difference between Rad26 and Mfd is the mechanism by which a lesion-arrested polymerase is displaced (discussed below).

Mechanistic insights into displacement of lesion-arrested Pol II during TC-NER

A lesion-arrested RNA polymerase needs to be moved away from the lesion site to allow access to repair factors for subsequent repair *(4, 5, 7, 22-24).* In prokaryotes, Mfd directly displaces the polymerase from the lesion site in an ATP-dependent manner (19, 20). In contrast, CSB fails to displace Pol II from a CPD lesion *in vitro* (25). Previous biochemical work suggested that lesion-arrested Pol II can backtrack, an intrinsic movement presumably promoted by TFIIH, which is loaded downstream of Pol II after CSB-mediated lesion detection during TC-NER (4, 5, 7, 26). The fate of this backtracked lesion-arrested Pol II has long been a puzzle in eukaryotic TC-NER. One model suggests that this backtracked lesion-arrested Pol II can remain engaged with the DNA template and can resume transcription after the lesion is repaired *(4).* However, recent work provided *in vivo* evidence against this model (27), showing that the lesion-arrested Pol II is released during TC-NER in mammalian cells. How this lesion-arrested Pol II could be released remained unknown. Here, we present biochemical data showing that a lesion-arrested Pol II can be released when located upstream of a CPD lesion, but not when at the CPD lesion site itself. Importantly, this release required the joint action of Rad26 and TFIIS, as Rad26 was not able to displace Pol II on its own. The requirement for TFIIS, and for active translocation by Rad26, suggests that displacement is the result of a combination of TFIIS-induced backtracking/cleavage and Rad26-dependent forward translocation. We propose that, after cleavage of the transcript's 3' end by TFIIS, the forward bias imposed on Pol II by Rad26's translocation activity results in annealing of the upstream fork of the transcription bubble, and therefore in shortening of the RNA-DNA hybrid (hyper-translocation) (fig. S14J). This reduces the overall stability of Pol II, leading to the release of the arrested Pol II and the RNA transcript. This is also in agreement with our observation that only backtracked Pol II (i.e. not Pol II stalled at the CPD) could be displaced.

Interactions between Rad26, Pol II, and Elf1 and implications for TC-NER

Rpb4/7 plays important roles in a number of molecular and cellular processes, including transcription and DNA repair. Rpb4/7 is a hub for interactions with several transcription factors, including Spt4/5 and Spt6 (16, 28). A previous genetic study in *S. cerevisiae* showed that Rpb4/7 promotes Rad26-dependent TC-NER while suppressing Rad26-independent TC-NER (29). The strengthened interaction between Rad26 and Rpb4/7 we observed in the Strong state of CPD-lesion-arrested Pol II is consistent with observation, suggesting that Rpb4/7 play a role in the first recognition step of TC-NER. The stable conformation we observed in the Strong Rad26-Rpb4/7 interaction would effectively compete off Spt4/5 (fig. S15), an elongation factor that binds both Rpb4/7 and the protrusion domain of Pol II (fig. S15) (16, 17, 30). Since Spt4/5 functions as an inhibitor of TC-NER (31), we propose that its steric exclusion by the lesion-induced strong interaction between Rad26 and Rpb4/7 is a major step in committing a lesion-stalled Pol II to TC-NER.

Addition of Elf1 to the Pol II (CPD)-Rad26 complex both strengthened the lesion-induced interaction between Rad26 and Pol II and led to new interactions between Rad26 and Pol II absent from all other CPD-stalled structures we have solved to date. Furthermore, the cryo-EM structure of this complex has the highest resolution (3.1Å) of all the Rad26-containing complexes, with a particularly marked improvement in the density for Rad26 itself (4Å as opposed to 8Å in the next-best structure). This is consistent with the idea that Rad26 is less flexible relative to Pol II in the presence of Elf1. This additional strengthening of the Rad26-Rpb4/7 interaction may commit the complex to TC-NER and facilitate the recruitment of downstream repair factors, suggesting that the Pol II(CPD)-Rad26-Elf1 structure represents the "engaged" complex in TC-NER. Given the

high conservation of Pol II, Rad26/CSB, ELOF1/Elf1, and Spt4/5/DSIF, we expect that similar structures would be observed for the human complexes.

Considering the data presented here, along with previous work, we propose a stepwise model for the initiation of TC-NER (fig. 4E). Initial interaction of Rad26 with an arrested Pol II results in bending of the upstream DNA towards Rpb4/7. Binding of Rad26 also leads to the displacement of Spt4/5 while its remodeler-like DNA translocation biases Pol II forward, promoting the bypass of non-lesion obstacles or small lesions (8, 11). In the case of lesion-arrested Pol II, Rad26 establishes a stronger interaction with Rpb4/7. Binding of the TC-NER factor Elf1 further strengthens the Rad26-Rpb4/7 interaction, setting the stage for the recruitment of downstream factors. Once TFIIH is recruited downstream of the lesion-arrested Pol II complex, it uses its helicases to help Pol II backtrack, allowing TFIIH and XPA to identify the DNA lesion for TC-NER (26). The backtracked Pol II can then be displaced by the joint action of Rad26/CSB and TFIIS.

METHODS

Protein expression and purification

Expression and purification of Rad26 were performed essentially as previously described (8). Briefly, recombinant Rad26 protein was expressed in *Escherichia coli* strain Rosetta 2(DE3) (Novagen) and purified by Ni-NTA agarose (Qiagen), Hi-Trap Heparin HP (GE Healthcare), and Superdex 200 10/300 GL columns (GE Healthcare). Rad26 mutants were purified in the same manner as wild-type proteins. Expression and purification of yeast TFIIS were performed as described (32). Expression and purification of yeast Elf1 and yeast Elf1ΔC (1-85) were performed essentially as previously described (16, 33). Briefly, GST-tagged Elf1 protein was expressed

in *Escherichia coli* strain Rosetta 2(DE3) (Novagen) and purified by Glutathione Sepharose 4 Fast Flow resin (GE Healthcare), and Superdex 200 10/300 GL column (GE Healthcare). Elf1ΔC (1-85) was purified in the same manner as wild-type protein. Recombinant Spt4/5 was a gift from Dr. Jianhua Fu and expressed and purified as described (34).

Saccharomyces cerevisiae 10-subunit Pol II was purified essentially as previously described (35). Briefly, Pol II (with a protein A tag in the Rpb3 subunit) was purified by an IgG affinity column (GE Healthcare), followed by Hi-Trap Heparin (GE Healthcare) and Mono Q anion exchange chromatography columns (GE Healthcare). Pol II was purified by incubating 10-subunit Pol II with 3-fold of Rpb4/7 followed by gel filtration. His6-tagged Rpb4/7 heterodimer was purified from E. coli by Ni-affinity chromatography followed by gel filtration as previously described (36).

In vitro transcription assay

Pol II elongation complexes were assembled essentially as previously described with a labeled RNA primer (8). For Pol II displacement assay, *in vitro* transcription was started by adding rNTPs mixture to a final concentration of 0.1 mM each. After 5 min, the rNTPs were removed by washing the Streptavidin magnetic beads (NEB) five times with elongation buffer (40mM HEPES, pH7.4, 5mM DTT, 5mM $MgCl_2$, and 40mM KCl). After removing rNTPs, 1μM TFIIS was added and further incubated at 30ºC for 30min to induce Pol II backtracking. To evaluate the effect of Rad26 on Pol II displacement, 200nM Rad26, elongation buffer or 0.1mM rNTPs, and 3mM dATP were added. After incubation at 30ºC for 10min, Pol II complex in the supernatant or bound to the resin were separated. Since the release of labeled RNA transcript occurs when Pol II is displaced from its template, free RNA transcripts are expected to be found in the supernatant (S)

11

while transcripts bound to the Pol II complex are found in the pellet (B). Supernatant or pellet samples were boiled for 10 min at 95°C in formamide loading buffer, and the RNA transcripts were separated by denaturing PAGE (6M urea). The gel was visualized by phosphorimaging and quantified using Image Lab software (Bio-Rad). For gel analysis, the transcripts were divided into 3 regions according to the length of RNA: region I corresponds to full-length transcripts (representing Pol II's that are stalled at the CPD lesion or uncleaved backtracked Pol II); while region II + III corresponds to shorten transcripts (i.e. Pol II's that have backtracked from the CPD lesion and have been cleaved by TFIIS). Relative displacement for each region was calculated as following: $\frac{Intensity\,in\,S\,for\,region\,I\,or\,II\,or\,III}{Total\,intensity\,S + B\,in\,region\,I + II + III}$

Note that a very small portion of sequence-specific paused Pol II proximal to the transcription start site (independent from the CPD-induced stalling) may also contribute to signals in region III, presumably due to less stable RNA:DNA hybrid. The contribution from this background is negligible.

Preparation of Pol II-Rad26 and Pol II(CPD)-Rad26-Elf1 complexes for electron microscopy

Template and non-template DNA oligonucleotides were obtained from IDT and further purified by PAGE. PAGE-purified RNA oligonucleotides were purchased from Dharmacon. HPLC-purified CPD lesion-containing template was purchased from TriLink. The RNA, template DNA (non-damaged or CPD lesion containing) and non-template DNA were annealed to form the scaffold as previously described (8).

To form the CPD-arrested Pol II complex, Pol II and three-fold excess of scaffold were mixed and further purified by gel filtration in 50mM HEPES, pH7.4, 5mM DTT, 5mM MgCl$_2$, and 40mM KCl. To form the backtracked Pol II complex, Pol II and the scaffold were incubated in

50mM HEPES, pH7.4, 5mM DTT, 5mM MgCl$_2$, and 40mM KCl. To form the backtracked Pol II-Rad26 complex, Rad26 was added to backtracked Pol II complex and incubated for 30 minutes. The final buffer was composed of 50mM HEPES, pH 7.4, 5mM DTT, 5mM MgCl$_2$, 40mM KCl, 200mM NaCl. For the Pol II(CPD)-Rad26 complex, a final concentration of 0.02% Glutaraldehyde was added after adding Rad26 and incubated for another 30 minutes. The crosslink reaction was terminated by adding 1M Tris-HCl, pH 8.0 to a final concentration of 100mM. The final concentrations of the different components were 1μM Pol II, 2μM Rad26, 1.1μM scaffold. To form the Pol II(CPD)-Rad26-Elf1 complexes, 4-fold excess of Elf1 was incubated with Pol II(CPD)-Rad26 complex. The final concentrations of the different components were 1μM Pol II, 2μM Rad26, 1.1μM scaffold, and 5μM Elf1.

The sequences used for elongation complex preparation are as follows: non-template DNA, 5′-CTAGTTGATCTCATATTTCATTCCTACTCAGGAGAAGGAGCAGAGCG-3′; template DNA, 5′-CGCTCTGCTCCTTCTCCCATCCTCTCGATGGCTATGAGATCAACTAG-3′; CPD lesion-containing template DNA, 5′-CGCTCTGCTCCTTCTCC<u>XX</u>TCCTCTCGATGGCTATGAGATCAACTAG-3′ (XX = CPD lesion); RNA (for Pol II(CPD)), 5′-AUCGAGAGGA-3′; RNA (for Backtracked Pol II), 5′-AUCGAGAGGAUGCAGAC-3′.

Electron microscopy

An aliquot of 4uL of each sample was applied to glow-discharged Quantifoil holey carbon films R1.2/13 Cu grids. The grids were blotted and plunge-frozen in liquid ethane using a Vitrobot Mark IV (FEI). Data collection was performed using Leginon[37] on an FEI Talos Arctica operated

at 200 kV, equipped with a Gatan K2 summit direct detector. For the Pol II(CPD)-Rad26 sample, 3,358 movies were recorded in counting mode at a dose rate of 11.3 electrons pixel^{-1}s^{-1} with a total exposure time of 7.05 s sub-divided into 150 ms frames, for a total of 47 frames. The images were recorded at a nominal magnification of 36,000x resulting in an object-level pixel size of 1.16 Å pixel^{-1}. For the Backtracked Pol II-Rad26 sample, 9,167 movies were recorded in counting mode at a dose rate of 6.75 electrons pixel^{-1} s^{-1} with a total exposure time of 11 s sub-divided into 200 ms frames, for a total of 55 frames. The images were recorded at a nominal magnification of 36,000x resulting in an object-level pixel size of 1.16 Å pixel^{-1}. For the Pol II(CPD)-Rad26 sample with Pol II lacking Rpb4/7, 955 movies were recorded in super-resolution mode at a dose rate of 5.34 electrons pixel^{-1} s^{-1} with a total exposure time of 13 s sub-divided into 250 ms frames, for a total of 44 frames. The images were recorded at a nominal magnification of 36,000x resulting in an object-level pixel size of 1.16 Å pixel^{-1} (0.58 Å per super-resolution pixel). For the Pol II(CPD)-Rad26-Elf1 sample, two datasets with total of 6,000 movies were recorded in counting mode at a dose rate of 6.9 electrons pixel^{-1} s^{-1} for the first dataset and 7.4 electrons pixel^{-1} s^{-1} for the second dataset with a total exposure time of 10 s sub-divided into 200 ms frames, for a total of 50 frames. The images were recorded at a nominal magnification of 36,000x resulting in an object-level pixel size of 1.16 Å pixel^{-1}. See Table S1 for details on cryo-EM data collection, refinement and validation.

Image processing

Movie frame alignment was performed using MotionCore2 (38) using the dose-weighted frame alignment option. CTF estimation was executed on the non-dose-weighted aligned micrographs using GCTF using the local defocus per particle option (39). Particle picking was

performed using FindEM [40] with 2D averages selected from the initial processing serving as templates. Motion correction, CTF estimation and particle picking were performed within the framework of Appion (41). Two-dimensional classification was performed to identify bad Pol II particles. Following 2D classification, an initial 3D classification was performed using a Pol II Elongation Complex model (PDB 1Y77) as reference. The 2D and initial 3D classifications were carried out using particles binned by 4 (4.64 Å pixel^{-1}). The detailed processing schemes for each sample are shown in figs. S1, S3, S7, S12. All initial refinements and classifications were done in Relion 3[42]. Once the final particles were selected, local and global ctf refinement were performed to further improve the resolution using cryoSPARC(43) The final map was refined in cryoSPARC using non-uniform refinement algorithm(44). The statistics for refinement of all maps are listed in Table S1.

Model building

For building the models of Pol II (CPD) Conformation 1 and 2, models of Pol II(CPD) complex (PDB accession 6O6C)(45) and Rpb4/7 of Pol II elongation complex model (PDB accession 5VVS)(8) were used as starting models for Pol II core (10 subunits) and Rpb4/7, respectively. The composite reference model of Pol II core and Rpb4/7 and the density maps were used as inputs in RosettaCM(46), in which 10 models were generated. A model with the best Rosetta energy was selected for each density map. Models were manually optimized in coot(47) and then refined using Rosetta Relax to further optimize the position and geometry of the amino acids side chains. The nucleic acids scaffold was manually built in coot. A selected model was refined using PHENIX real space refinement(48) with secondary structure restrains option followed by second round of Rosetta Relax, in which 10 models were generated. A model with the

best score function was selected as the final model. The metals were manually added to each model followed by a final run of PHENIX real space refinement. The model of backtracked Pol II complex apo was built using the same steps described above except that the Pol II(CPD) complex Conformation 1 model was used as a starting model.

For building the model of Rad26 for Pol II(CPD)-Rad26, Backtracked Pol II-Rad26 complex, and Pol II(CPD)-Rad26 with Pol II lacking Rpb4/7, the model of Pol II-Rad26 stalled (PDB accession 5VVR) (8) was used as a reference. The Rad26 starting model was rigid body docked into the density map using UCSF Chimera(49). The N-terminal helix of Rad26 was manually adjusted or deleted in coot to best fit the density map. The composite reference model of optimized Rad26 and Pol II(CPD) Conformation 1 (built as described above) was used as a starting model in RosettaCM.

To build the model Pol II(CPD)-Rad26-Elf1, the reference models for Rad26 and Elf1 were selected based on homology detection using hidden Markov model as implemented in HHpred (50). The segmented density of Rad26 and Elf1, and the references models from HHpred were used as inputs to build their models using RosettaCM. Pol II was built using the composite models of Pol II 10 subunit (from PDB: 6O6C) and Rpb4/7 (from PDB: 5VVS) as described above. Nucleic acid scaffolds for all models were built in coot. The same steps described above were performed to improve position and geometry of the amino acids side chains. FSC curves of map-to-model were calculated in Rosetta. The validation statistics for all models are shown in Table S1.

Structure analysis

16

All figures were generated using UCSF ChimeraX(51). The cryo-EM maps were first segmented using Seggar(52) as implemented in UCSF Chimera. The segmented densities were colored in ChimeraX.

To generate difference map for Strong state minus Weak state of Pol II(CPD)-Rad26, the cryo-EM maps were first low-pass filtered in SPIDER(53) with FQ operation, 'top-hat' function preserving frequencies below 0.1 (a resolution of 10Å in our maps). The difference map was generated in ChimeraX with volume operation (vol) as follows: the filtered Strong state was fitted into Weak state map with 'fitmap' command, the Strong state was resampled on the grid of Weak state map with 'vol resample' command, and the Weak state map was subtracted from the resampled Strong state map with 'vol subtract' command. The same steps were followed to generate difference map for Pol II(CPD)-Rad26 (Strong state) minus Backtracked Pol II-Rad26 and Pol II(CPD)-Rad26-Elf1 minus Pol II(CPD)-Rad26.

The consensus refinement and the masks used in multi-body refinement were prepared in Relion 3 using the default options. Multi-body refinement generated 10 structures, which describe flexibility along each eigenvector. To visualize flexibility along eigenvector 1 and 2 for Pol II(CPD)-Rad26 (Strong state), the model of Rpb4/7 was segmented out from Pol II(CPD)-Rad26 Strong Rad26-Rpb4/7 state model and was rigid-body fitted separately into each one of the ten structures from multi-body refinement. The models were fitted using 'fitmap' command in ChimeraX (see fig. S8). The same steps were followed to visualize the flexibility along eigenvector 1 and 2 for Pol II(CPD) but using the consensus refoment of Pol II(CPD) to generate masks and as an input for multi-body refinement (see fig. S9). The segmented model of Rpb4/7 chains from Pol II(CPD) (lacking Rad26) model was used for rigid-body fitting.

To obtain the cross-correlation coefficients between the Rpb4/7 model and the different cryo-EM maps shown in fig. S5, the cryo-EM maps for the Strong and Weak states of Pol II(CPD)-Rad26 were aligned with the 'fitmap' function in ChimeraX. Then, the full complex model was aligned to its corresponding cryo-EM density. To calculate cross-correlation coefficients, the model of Rpb4/7 from Pol II(CPD)-Rad26 (Strong state) was fitted into the segmented Rpb4/7 density from Weak and Strong state while disabling the options for allowing any rotations and shifts. The same steps were performed to calculate the cross-correlation for the fitting of Rpb4/7 model for Strong state of Pol II(CPD)-Rad26 into the map of Backtracked Pol II-Rad26 (fig. S6).

Yeast strain construction

To make *elf1-ΔC* mutant strain, a *URA3* fragment from pRS306 was PCR amplified with Primer 1 and Primer 2 (see below) to replace the part of *ELF1* open reading frame that encodes C-terminal 60 amino acids (with the incorporation of a TAA stop codon immediately after the amino acid 58). The PCR-cassette was transformed into cells using the method described previously(54). The resulting mutant strains were further confirmed by sequencing. Primers are listed below:

Primer 1:

TGATGTATATAGTGATTGGTTTGACGCCGTCGAAGAAGTCAATTCTGGCCGTGGATA ACCTGATGCGGTATTTTCTCC.

Primer 2:

TTAAAATATAAAATATATATGACCTAAGTAAATATGGTTTTTTCTCAGGACCGGACG GCATCAGAGCAGATTGTA

All genotypes of yeast strains are listed in Table S2.

UV survival assay

Yeast cells were grown at 30°C to optical density (OD) of 3 at 600 nm and diluted to OD 0.6 in YPD medium. Cells were plotted on YPD plate with 5-fold serial dilutions. Once dried, the plates were UV irradiated with UV crosslinker (FisherBiotech® FB-UVXL-1000) in dark room and wrapped in foil after irradiation. Plates were incubated for 2 days at 30°C before imaging.

ACKNOWLEDGEMENTS

REFERENCES

1. Gregersen LH, Svejstrup JQ. The Cellular Response to Transcription-Blocking DNA Damage. Trends Biochem Sci. 2018;43(5):327-41.

2. Hanawalt PC, Spivak G. Transcription-coupled DNA repair: two decades of progress and surprises. Nat Rev Mol Cell Biol. 2008;9(12):958-70.

3. Lagerwerf S, Vrouwe MG, Overmeer RM, Fousteri MI, Mullenders LH. DNA damage response and transcription. DNA Repair (Amst). 2011;10(7):743-50.

4. Lans H, Hoeijmakers JHJ, Vermeulen W, Marteijn JA. The DNA damage response to transcription stress. Nat Rev Mol Cell Biol. 2019;20(12):766-84.

5. Wang W, Xu J, Chong J, Wang D. Structural basis of DNA lesion recognition for eukaryotic transcription-coupled nucleotide excision repair. DNA Repair (Amst). 2018;71:43-55.

6. van den Heuvel D, van der Weegen Y, Boer DEC, Ogi T, Luijsterburg MS. Transcription-Coupled DNA Repair: From Mechanism to Human Disorder. Trends Cell Biol. 2021;31(5):359-71.

7. Noe Gonzalez M, Blears D, Svejstrup JQ. Causes and consequences of RNA polymerase II stalling during transcript elongation. Nat Rev Mol Cell Biol. 2021;22(1):3-21.

8. Xu J, Lahiri I, Wang W, Wier A, Cianfrocco MA, Chong J, Hare AA, Dervan PB, DiMaio F, Leschziner AE, Wang D. Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. Nature. 2017;551(7682):653-7.

9. Selby CP, Sancar A. Cockayne syndrome group B protein enhances elongation by RNA polymerase II. Proc Natl Acad Sci U S A. 1997;94(21):11205-9.

10. Vélez-Cruz R, Egly J-M. Cockayne syndrome group B (CSB) protein: at the crossroads of transcriptional networks. Mechanisms of ageing and development. 2013;134(5-6):234-42.

11. Xu J, Wang W, Xu L, Chen JY, Chong J, Oh J, Leschziner AE, Fu XD, Wang D. Cockayne syndrome B protein acts as an ATP-dependent processivity factor that helps RNA polymerase II overcome nucleosome barriers. Proc Natl Acad Sci U S A. 2020;117(41):25486-93.

12. Wang Y, Chakravarty P, Ranes M, Kelly G, Brooks PJ, Neilan E, Stewart A, Schiavo G, Svejstrup JQ. Dysregulation of gene expression as a cause of Cockayne syndrome neurological disease. Proc Natl Acad Sci U S A. 2014;111(40):14454-9.

13. Cleaver JE, Revet I. Clinical implications of the basic defects in Cockayne syndrome and xeroderma pigmentosum and the DNA lesions responsible for cancer, neurodegeneration and aging. Mech Ageing Dev. 2008;129(7-8):492-7.

14. Olivieri M, Cho T, Alvarez-Quilon A, Li K, Schellenberg MJ, Zimmermann M, Hustedt N, Rossi SE, Adam S, Melo H, Heijink AM, Sastre-Moreno G, Moatti N, Szilard RK, McEwan A, Ling AK, Serrano-Benitez A, Ubhi T, Feng S, Pawling J, Delgado-Sainz I, Ferguson MW, Dennis JW, Brown GW, Cortes-Ledesma F, Williams RS, Martin A, Xu D, Durocher D. A Genetic Map of the Response to DNA Damage in Human Cells. Cell. 2020;182(2):481-96 e21.

15. Prather D, Krogan NJ, Emili A, Greenblatt JF, Winston F. Identification and characterization of Elf1, a conserved transcription elongation factor in Saccharomyces cerevisiae. Mol Cell Biol. 2005;25(22):10122-35.

16. Ehara H, Yokoyama T, Shigematsu H, Yokoyama S, Shirouzu M, Sekine SI. Structure of the complete elongation complex of RNA polymerase II with basal factors. Science. 2017;357(6354):921-4.

17. Ehara H, Sekine SI. Architecture of the RNA polymerase II elongation complex: new insights into Spt4/5 and Elf1. Transcription. 2018;9(5):286-91.

18. Boeing S, Williamson L, Encheva V, Gori I, Saunders RE, Instrell R, Aygun O, Rodriguez-Martinez M, Weems JC, Kelly GP, Conaway JW, Conaway RC, Stewart A, Howell M, Snijders AP, Svejstrup JQ. Multiomic Analysis of the UV-Induced DNA Damage Response. Cell Rep. 2016;15(7):1597-610.

19. Fan J, Leroux-Coyau M, Savery NJ, Strick TR. Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. Nature. 2016;536(7615):234-7.

20. Kang JY, Llewellyn E, Chen J, Olinares PDB, Brewer J, Chait BT, Campbell EA, Darst SA. Structural basis for transcription complex disruption by the Mfd translocase. Elife. 2021;10.

21. Donahue BA, Yin S, Taylor J-S, Reines D, Hanawalt P. Transcript cleavage by RNA polymerase II arrested by a cyclobutane pyrimidinedimer in the DNA template. Proc Natl Acad Sci USA. 1994;91:8502-6.

22. Geijer ME, Marteijn JA. What happens at the lesion does not stay at the lesion: Transcription-coupled nucleotide excision repair and the effects of DNA damage on transcription in cis and trans. DNA repair. 2018;71:56-68.

23. Mullenders L. DNA damage mediated transcription arrest: Step back to go forward. DNA Repair (Amst). 2015;36:28-35.

24. Steurer B, Marteijn JA. Traveling Rocky Roads: The Consequences of Transcription-Blocking DNA Lesions on RNA Polymerase II. J Mol Biol. 2017;429(21):3146-55.

25. Selby CP, Sancar A. Human transcription-repair coupling factor CSB/ERCC6 is a DNA-stimulated ATPase but is not a helicase and does not disrupt the ternary transcription complex of stalled RNA polymerase II. Journal of Biological Chemistry. 1997;272(3):1885-90.

26. Li CL, Golebiowski FM, Onishi Y, Samara NL, Sugasawa K, Yang W. Tripartite DNA Lesion Recognition and Verification by XPC, TFIIH, and XPA in Nucleotide Excision Repair. Mol Cell. 2015;59(6):1025-34.

27. Chiou YY, Hu J, Sancar A, Selby CP. RNA polymerase II is released from the DNA template during transcription-coupled repair in mammalian cells. J Biol Chem. 2018;293(7):2476-86.

28. Vos SM, Farnung L, Boehning M, Wigge C, Linden A, Urlaub H, Cramer P. Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature. 2018;560(7720):607-12.

29. Li S, Smerdon MJ. Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in Saccharomyces cerevisiae. The EMBO journal. 2002;21(21):5921-9.

30. Selvam K, Ding B, Sharma R, Li S. Evidence that Moderate Eviction of Spt5 and Promotion of Error-Free Transcriptional Bypass by Rad26 Facilitates Transcription Coupled Nucleotide Excision Repair. J Mol Biol. 2019;431(7):1322-38.

31. Li W, Giles C, Li S. Insights into how Spt5 functions in transcription elongation and repressing transcription coupled DNA repair. Nucleic Acids Res. 2014;42(11):7069-83.

32. Awrey DE, Shimasaki N, Koth C, Weilbaecher R, Olmsted V, Kazanis S, Shan X, Arellano J, Arrowsmith CH, Kane CM, Edwards AM. Yeast transcript elongation factor (TFIIS), structure and function. II: RNA polymerase binding, transcript cleavage, and read-through. J Biol Chem. 1998;273(35):22595-605.

33. Xu J, Chong J, Wang D. Opposite roles of transcription elongation factors Spt4/5 and Elf1 in RNA polymerase II transcription through B-form versus non-B DNA structures. Nucleic Acids Res. 2021.

34. Crickard JB, Lee J, Lee TH, Reese JC. The elongation factor Spt4/5 regulates RNA polymerase II transcription through the nucleosome. Nucleic Acids Res. 2017;45(11):6362-74.

35. Wang D, Bushnell DA, Huang X, Westover KD, Levitt M, Kornberg RD. Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. Science. 2009;324(5931):1203-6.

36. Armache KJ, Mitterweger S, Meinhart A, Cramer P. Structures of complete RNA polymerase II and its subcomplex, Rpb4/7. J Biol Chem. 2005;280(8):7131-4.

37. Suloway C, Pulokas J, Fellmann D, Cheng A, Guerra F, Quispe J, Stagg S, Potter CS, Carragher B. Automated molecular microscopy: the new Leginon system. J Struct Biol. 2005;151(1):41-60.

38. Zheng SQ, Palovcak E, Armache J-P, Cheng Y, Agard DA. Anisotropic correction of beam-induced motion for improved single-particle electron cryo-microscopy. bioRxiv. 2016:061960.

39. Zhang K. Gctf: Real-time CTF determination and correction. J Struct Biol. 2016;193(1):1-12.

40. Roseman A. FindEM—a fast, efficient program for automatic selection of particles from electron micrographs. Journal of structural biology. 2004;145(1-2):91-9.

41. Lander GC, Stagg SM, Voss NR, Cheng A, Fellmann D, Pulokas J, Yoshioka C, Irving C, Mulder A, Lau PW, Lyumkis D, Potter CS, Carragher B. Appion: an integrated, database-driven pipeline to facilitate EM image processing. J Struct Biol. 2009;166(1):95-102.

42. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife. 2018;7:e42166.

43. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14(3):290-6.

44. Punjani A, Zhang H, Fleet DJ. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat Methods. 2020;17(12):1214-21.

45. Lahiri I, Xu J, Han BG, Oh J, Wang D, DiMaio F, Leschziner AE. 3.1 Å structure of yeast RNA polymerase II elongation complex stalled at a cyclobutane pyrimidine dimer lesion solved using streptavidin affinity grids. Journal of structural biology. 2019;207(3):270-8.

46. Song Y, DiMaio F, Wang RY, Kim D, Miles C, Brunette T, Thompson J, Baker D. High-resolution comparative modeling with RosettaCM. Structure. 2013;21(10):1735-42.

47. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60(Pt 12 Pt 1):2126-32.

48. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol. 2018;74(Pt 6):531-44.

49. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605-12.

50. Soding J, Biegert A, Lupas AN. The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. 2005;33(Web Server issue):W244-8.

51. Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH, Ferrin TE. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. Protein Sci. 2018;27(1):14-25.

52. Pintilie GD, Zhang J, Goddard TD, Chiu W, Gossard DC. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. J Struct Biol. 2010;170(3):427-38.

53. Shaikh TR, Gao H, Baxter WT, Asturias FJ, Boisset N, Leith A, Frank J. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. Nat Protoc. 2008;3(12):1941-74.

54. Gietz RD, Schiestl RH, Willems AR, Woods RA. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast. 1995;11(4):355-60.

55. Ehara H, Kujirai T, Fujino Y, Shirouzu M, Kurumizaka H, Sekine SI. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. Science. 2019;363(6428):744-7.

56. Bernecky C, Plitzko JM, Cramer P. Structure of a transcribing RNA polymerase II-DSIF complex reveals a multidentate DNA-RNA clamp. Nat Struct Mol Biol. 2017;24(10):809-15.

FIGURES

Figure 1 | Rad26 interacts in a similar way with paused, CPD-stalled and Backtracked RNA Pol II

(A) Cartoon representation of the different complexes analyzed by cryo-EM. The black rhomboid represents the CPD lesion. **(B-F)** Cryo-EM maps of **(B)** Pol II-Rad26 at a non-lesion arrest from our previous work (5.8Å) (8), **(C)** Backtracked Pol II-Rad26 (4.4Å), **(D,E)** Pol II(CPD)-Rad26 in states showing either a "Weak" (3.7Å) (D), or "Strong" (3.5Å) (E) interaction between Rad26 and Rpb4/7 (orange arrow), and **(F)** Pol II(CPD)-Rad26 with Pol II lacking Rpb4/7 (4.7Å). The maps were filtered according to the local resolution and were segmented and colored to highlight the different components, as indicated in (B). Cartoon representations of each structure, in the same orientation, are shown next to the maps. **(G-K)** Cryo-EM densities corresponding to the DNA/RNA scaffolds in the vicinity of the active site of Pol II segmented from the maps shown in B-F. The active site Bridge helix was included as a reference point. A close-up of the cryo-EM density corresponding to the CPD lesion is shown in (J). The color scheme used throughout the paper is as follows: Pol II: grey; Rad26: orange; non-template strand: green; template strand: blue; RNA: red.

A

Rad26

non-lesion arrest   backtracked   CPD lesion arrest   CPD lesion arrest
10-subunit Pol II

B
Rad26   Rpb4/7
Pol II
TS
NTS
RNA

C

D
"Weak" state

E
"Strong" state

F

G
Downstream
DNA
Bridge helix

H
backtracked
RNA

I
CPD

J
CPD
CPD

K
CPD
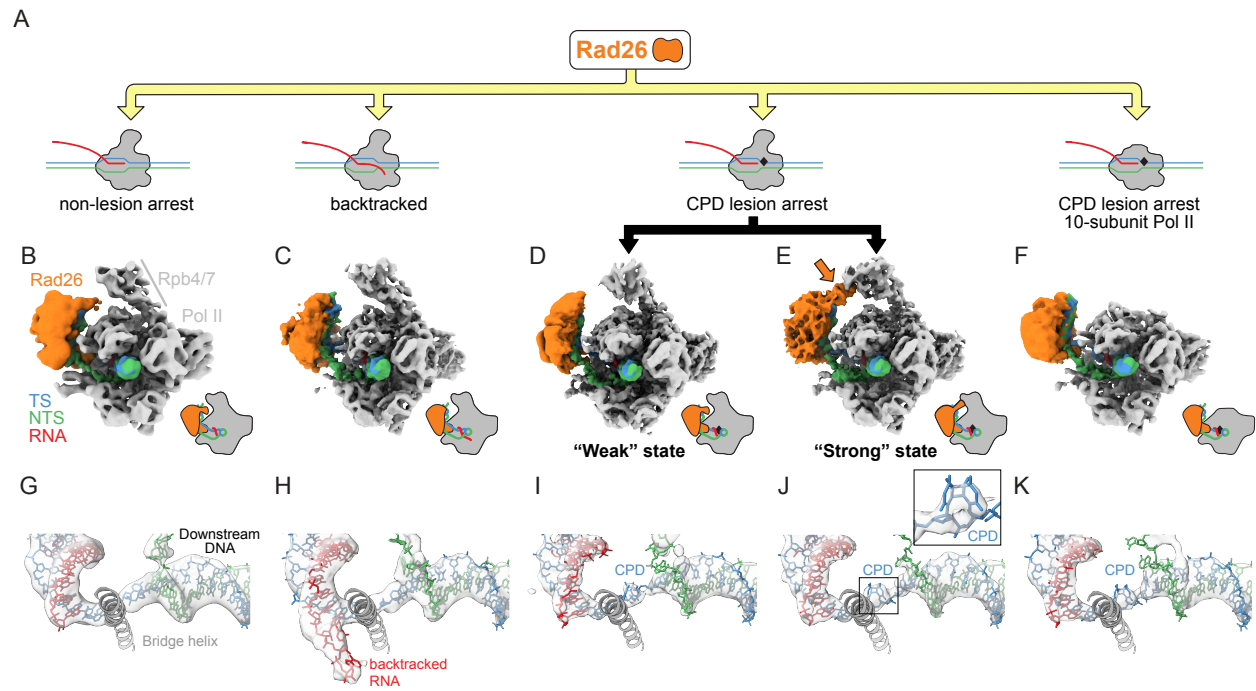
Figure 2 | A stronger Rad26-Rpb4/7 interaction is observed in the presence of a CPD lesion

(A, B) Difference density map (in blue) calculated by subtracting Pol II(CPD)-Rad26 with Weak Rad26-Rpb4/7 interaction (Fig. 1D) from Pol II(CPD)-Rad26 with Strong Rad26-Rpb4/7 interaction (Fig. 1E), displayed on either **(A)** the cryo-EM density or **(B)** the atomic model for Pol II(CPD)-Rad26 with Weak Rad26-Rpb4/7 interaction. **(C,D)** Rpb4/7 moved closer to Rad26 in the Strong Rad26-Rpb4/7 state relative to both the Weak state (C) and the Backtracked complex (D). Models were aligned to each other using the core of Pol II (i.e. excluding Rpb4/7). Zoomed-in views of Rpb4/7 are shown to the right of each set of aligned models. The distance between Cα of K212 in Rpb4 between the structures in each comparison is shown as a reference. **(E)** Superimposed cryo-EM densities for the DNA/RNA scaffolds in the Strong and Weak Rad26-Rpb4/7 states. Darker and lighter shades represent the Strong and Weak Rad26-Rpb4/7 states, respectively. The corresponding molecular models are shown on the right. **(F)** Same as in (E) but comparing the Strong Rad26-Rpb4/7 state (darker colors) with Pol II lacking Rpb4/7 (lighter colors).

A **Strong - Weak**

Weak state map

B

90°

Weak state model

C

2 Å

Strong state
Weak state

D

5 Å

Strong Rad26-Rpb4/7 state
Backtracked Pol II-Rad26

E **Strong vs. Weak**

Strong: **TS** **NTS** **RNA**
Weak: **TS** **NTS** **RNA**

F **Strong vs. 10-subunit Pol II**

Strong: **TS** **NTS** **RNA**
10-subunit: **TS** **NTS** **RNA**
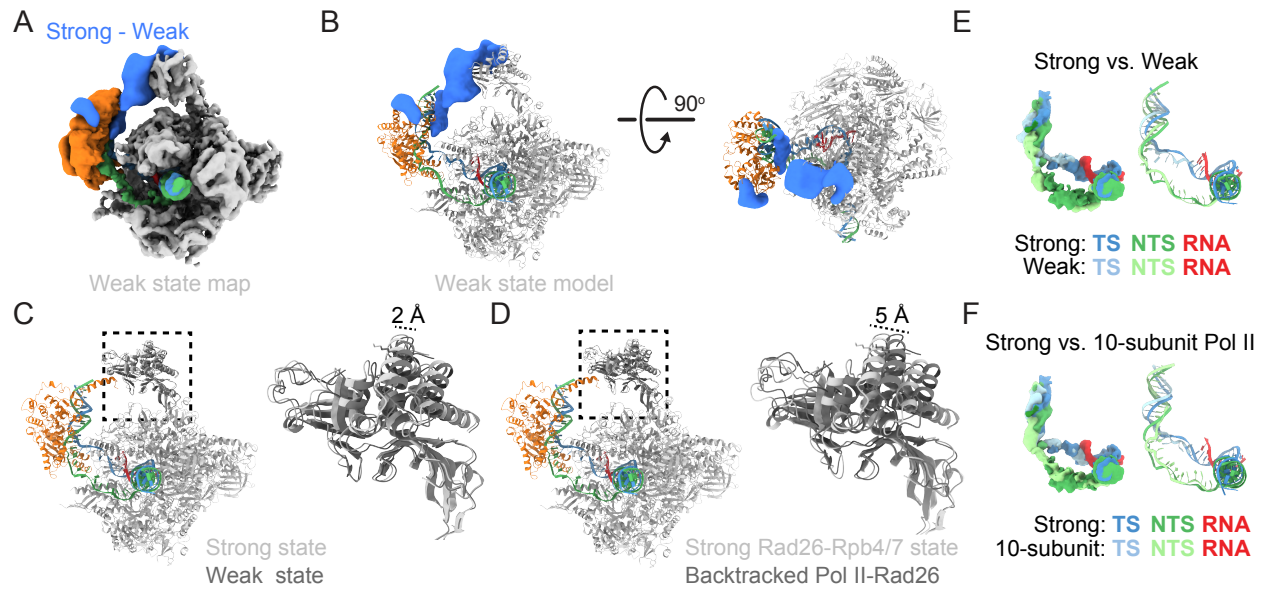
Figure 3 | Elf1 enhances interactions between Pol II and Rad26 Rpb4/7

(A) Elf1, or a C-terminally truncated version (ElfΔC) have no effect on the stalling of Pol II at a CPD lesion. **(B)** Elf1 enhances the binding of Rad26 to Pol II complex. **(C)** 3.1Å cryo-EM map of the Pol II(CPD)-Rad26-Elf1 complex, with Elf1 colored in light purple. The regions of Pol II that interact with Elf1 or Rad26 are colored in shades of purple or orange, respectively. **(D,E)** Zoomed in views of the area highlighted by the square in (C, right) for Pol II(CPD)-Rad26 without (D) or with (E) Elf1 bound. Top: Cryo-EM densities. Bottom: Corresponding models. The dashed square highlights the region where new density is present in the Pol II(CPD)-Rad26-Elf1 structure. The structural elements that become ordered in both Pol II and Rad26 are labeled (bottom). **(F)** The increase in the bend angle of the upstream DNA mirrors the stabilization of Rad26 in the cryo-EM maps. (Left) cartoon representations of the three structures being compared. (Right) The DNA/RNA scaffolds were superimposed using the downstream DNA and color-coded as shown on the left. The distances of upstream DNA shift between these states are shown.

A

Elf1

Time - WT ΔC

RO

CPD

start

B

Pol II(CPD)

Elf1: - WT ΔC

Rad26: — 10 25 50 100 — 10 25 50 100 — 10 25 50 100 (nM)

Pol II-
Rad26
(-Elf1)

C

Complex formation (%)

Elf1
-
WT
ΔC

Rad26: 10 25 50 100 (nM)

*** ns ns ns

C

Rad26   Pol II

90°

10°

70°

Elf1
TS
NTS
RNA

Rpb1: coiled-coil
(261-326)

Rpb2: lobe
(219-405)

Rpb2: protrusion
(44-218/405-465)

Rpb2: wall
(952-973)

D

Pol II(CPD)-Rad26
(Strong)

Rpb4/7

Rpb2: wall

Rad26
lobe 2

Rpb2
"flap loop"

E

Pol II(CPD)-Rad26-Elf1

Rpb4/7

Rpb2: wall

Rad26
lobe 2

Rpb2
"flap loop"

F

Pol II(CPD)-Rad26

Weak

Strong

+ Elf1

5Å 6Å

Figure 4 | Pol II can be displaced by the joint action of Rad26 and TFIIS

(A) Schematic representation of two possible scenarios for the displacement of Pol II after a lesion-induced arrest: Pol II is displaced at the CPD lesion site, or after backtracking. **(B)** Rad26 alone cannot promote the displacement of Pol II from a CPD-stalled Pol II. Denaturing PAGE showing RNA transcripts in both the Supernatant (S) and Bound fractions in the presence of either wild type (WT) or ATPase dead Rad26 (K328R). **(C)** Rad26 and TFIIS promote the displacement of Pol II from positions upstream of the CPD lesion site. Same as (B) but in the presence of TFIIS. **(D)** Quantitation of the data in (B, C) shown as mean and standard deviation (n=3). **(E)** Stepwise model for lesion recognition, reconfiguration, and Pol II displacement during TC-NER initiation in yeast.

A

Pol II
RNA
bulky
DNA
lesion
TS
NTS

Backtracked

?

?

Pol II
displacement

B

Rad26

– K328R WT

B S B S B S

RO

CPD

start

C

TFIIS
Rad26

WT K328R

NTPs: – + – +

B S B S B S B S

RO

CPD

start

D

TFIIS

– +

80

60

40

20

Pol II displacement %

WT K328R WT K328R

Rad26

E
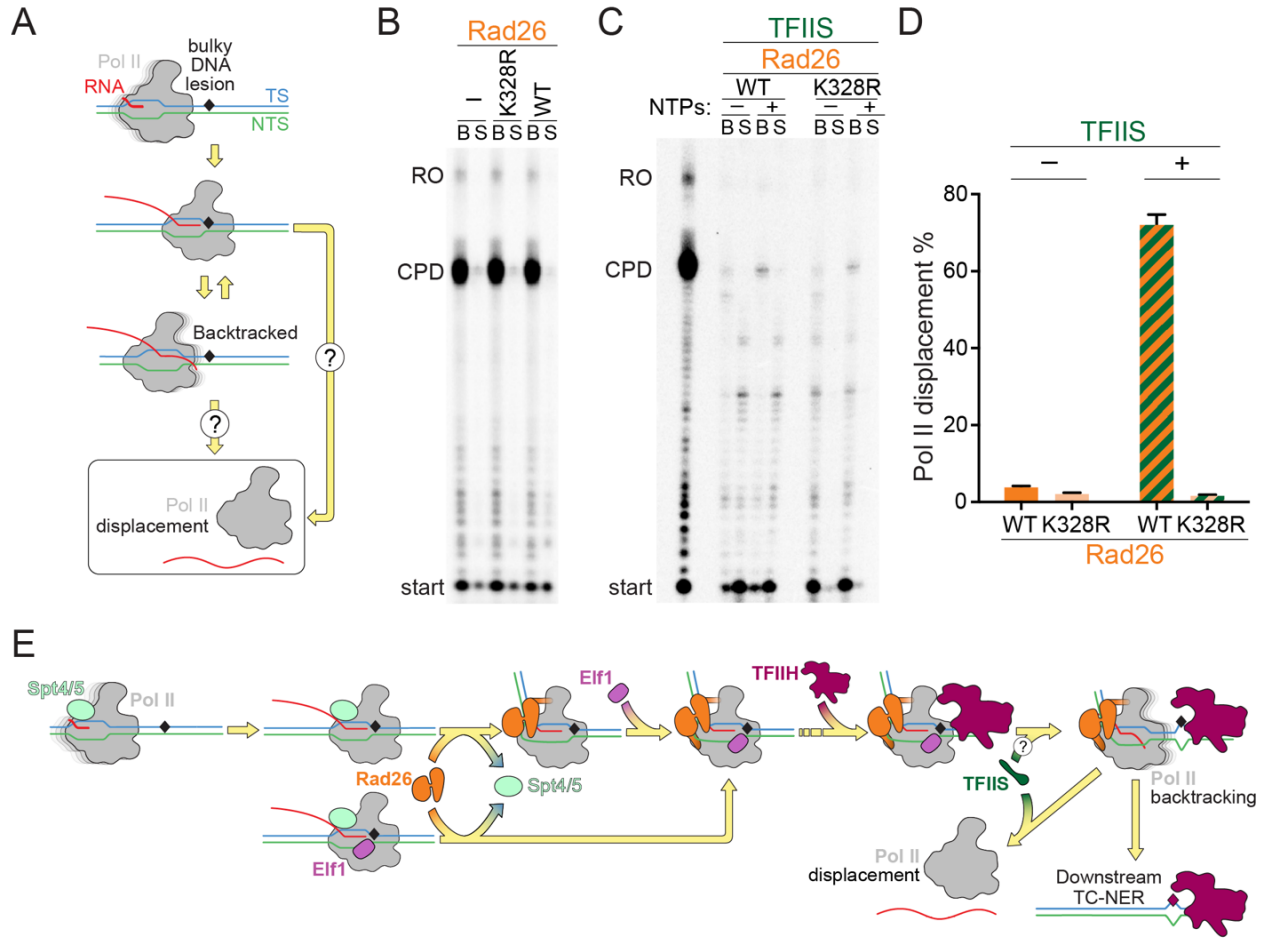
SUPPLEMENTARY FIGURES

Figure supplement 1 | Cryo-EM structure determination of the Pol II(CPD)-Rad26 and Pol II(CPD) complexes

(A-C) Representative micrograph **(A),** power spectrum **(B),** and representative 2D class averages **(C)** of Pol II(CPD)-Rad26 complexes. **(D)** Schematic of the strategy used to sort out the dataset into Pol II(CPD)-Rad26 Strong and Weak Rad26-Rpb4/7 states, and Pol II(CPD) conformations 1 and 2. Focused 3D classification was performed without alignment unless otherwise noted. The number of particles contributing to each selected structure is indicated. The percentages shown are related to the total number of particles picked from the micrographs. The indicated resolution corresponds to the 0.143 Fourier shell correlation (FSC) based on gold-standard FSC curves (see fig. S2).
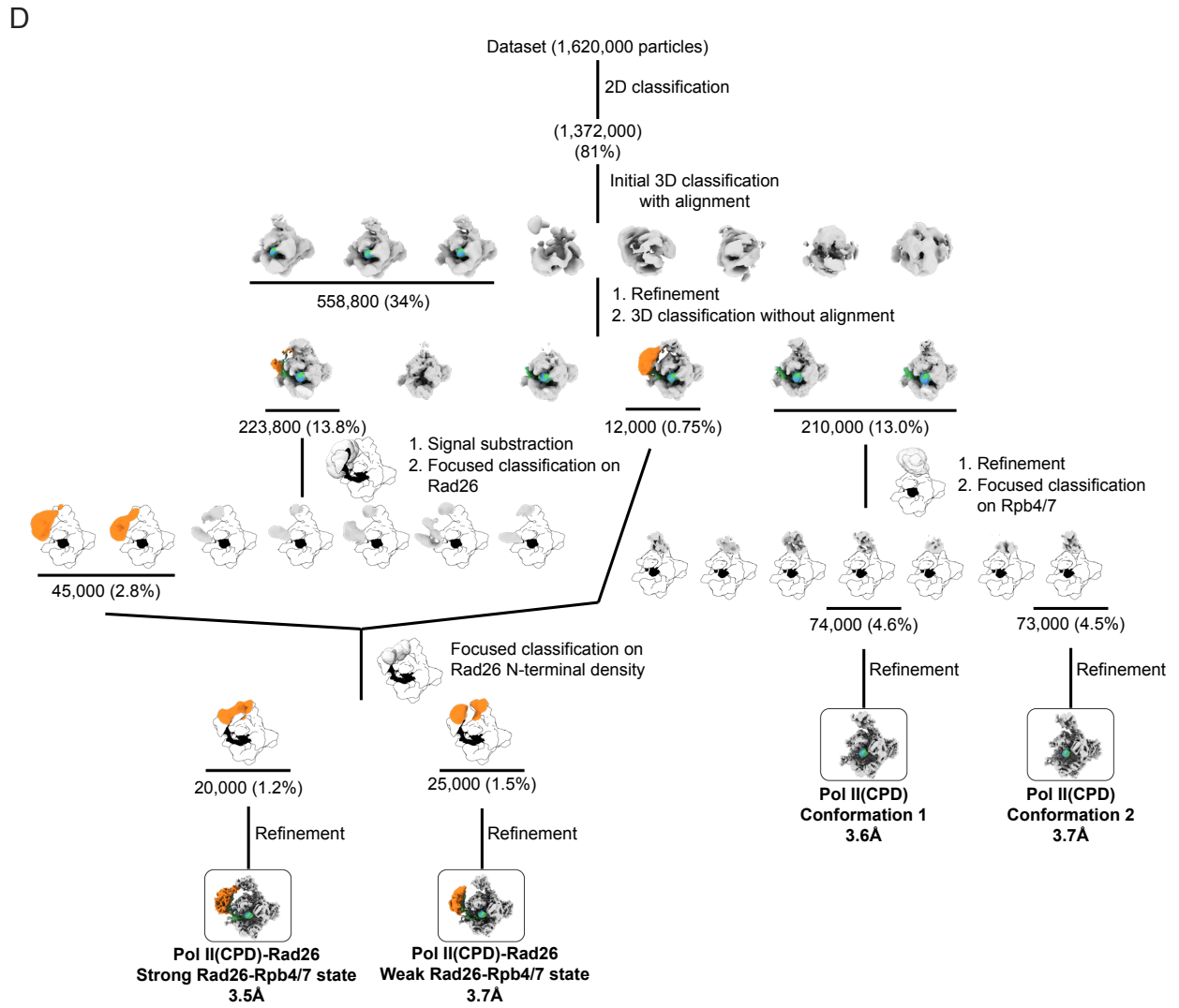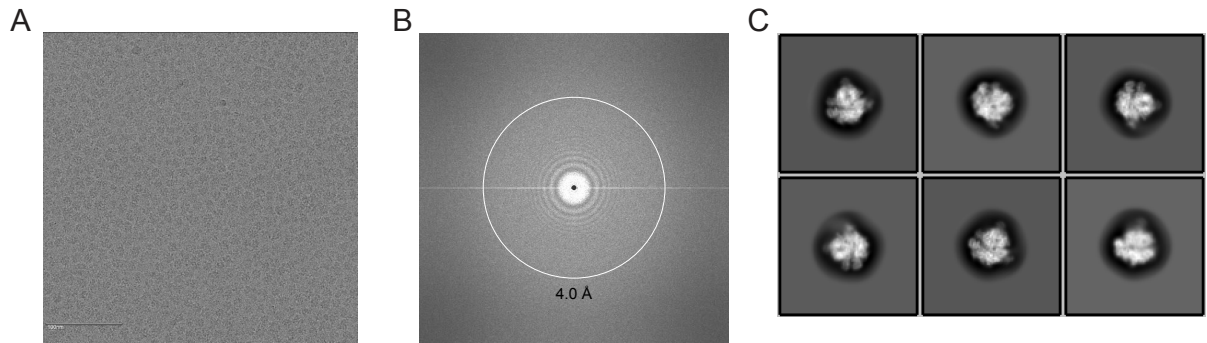
Figure supplement 2 | Analysis of the Pol II(CPD)-Rad26 and Pol II(CPD) cryo-EM maps

(A) Front and back views of locally filtered maps, colored by local resolution, of Pol II(CPD)-Rad26 Strong and Weak Rad26-Rpb4/7 states, and Pol II(CPD) conformations 1 and 2. **(B, C)** Euler angle distribution of particle images **(B)** and FSC plots **(C)** for the maps shown in (A). **(D-E)** Close-ups of the cryo-EM densities corresponding to the Rpb1 Bridge helix **(D)**, and the Rpb2/Rpb9 'Jaw' of Pol II **(E)** for the indicated structures with the models fitted in. **(F)** FSC curves for map-to-model fits for the maps shown in (A). The 0.5 FSC line is shown.

A Pol II(CPD)-Rad26 (Strong state) · Pol II(CPD)-Rad26 (Weak state) · Pol II(CPD) (Conformation 1) · Pol II(CPD) (Conformation 2) — Resolution (Å): 3 4 5 6 7 8 10 12 14

B Angular distribution plots (Azimuth vs Elevation; # of images)

C FSC curves (FSC vs Spatial frequency 1/Å)
- Pol II(CPD)-Rad26 (Strong state): No Mask (7.9Å), Masked (3.3Å), Corrected (3.3Å), FSC = 0.143
- Pol II(CPD)-Rad26 (Weak state): No Mask (8.6Å), Masked (3.6Å), Corrected (3.5Å), FSC = 0.143
- Pol II(CPD) (Conformation 1): No Mask (7.2Å), Masked (3.6Å), Corrected (3.6Å), FSC = 0.143
- Pol II(CPD) (Conformation 2): No Mask (7.5Å), Masked (3.7Å), Corrected (3.7Å), FSC = 0.143

D Pol II(CPD)-Rad26 (Strong state) · Pol II(CPD)-Rad26 (Weak state) · Pol II(CPD) Conformation1 · Pol II(CPD) Conformation2

E Pol II(CPD)-Rad26 (Strong state) · Pol II(CPD)-Rad26 (Weak state) · Pol II(CPD) Conformation1 · Pol II(CPD)Conformation2

F Map-to-Model FSC (FSC vs Spatial frequency 1/Å)
- Pol II(CPD)-Rad26 (Strong state): Map-to-Model FSC (3.6Å), FSC = 0.5
- Pol II(CPD)-Rad26 (Weak state): Map-to-Model FSC (3.6Å), FSC = 0.5
- Pol II(CPD) (Conformation 1): Map-to-Model FSC (3.8Å), FSC = 0.5
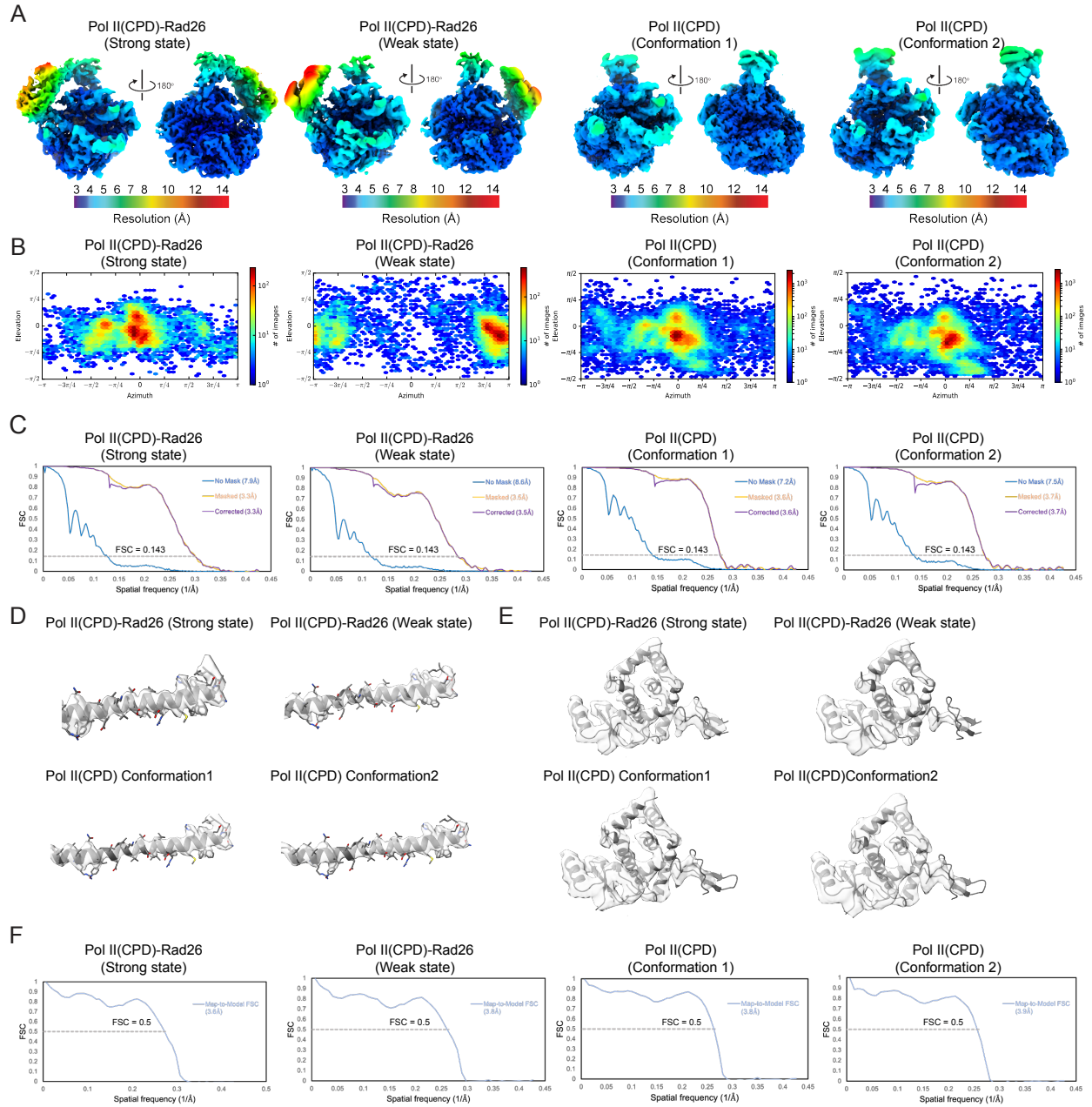- Pol II(CPD) (Conformation 2): Map-to-Model FSC (3.9Å), FSC = 0.5

36

Figure supplement 3 | Cryo-EM structure determination of the Backtracked Pol II-Rad26 and Backtracked Pol II complexes

(A-C) Representative micrograph **(A),** Power spectrum **(B),** and representative 2D class averages **(C)** of Backtracked Pol II-Rad26 complexes. **(D)** Schematic representation of the strategy used to sort out the dataset into Backtracked Pol II-Rad26 and Backtracked Pol II. Focused 3D classification was performed without alignment unless otherwise noted. The number of particles contributing to each selected structure is indicated. The percentages shown are related to the total number of particles picked from micrographs. The indicated resolution corresponds to the 0.143 Fourier shell correlation (FSC) based on gold-standard FSC curves (see fig. S4).
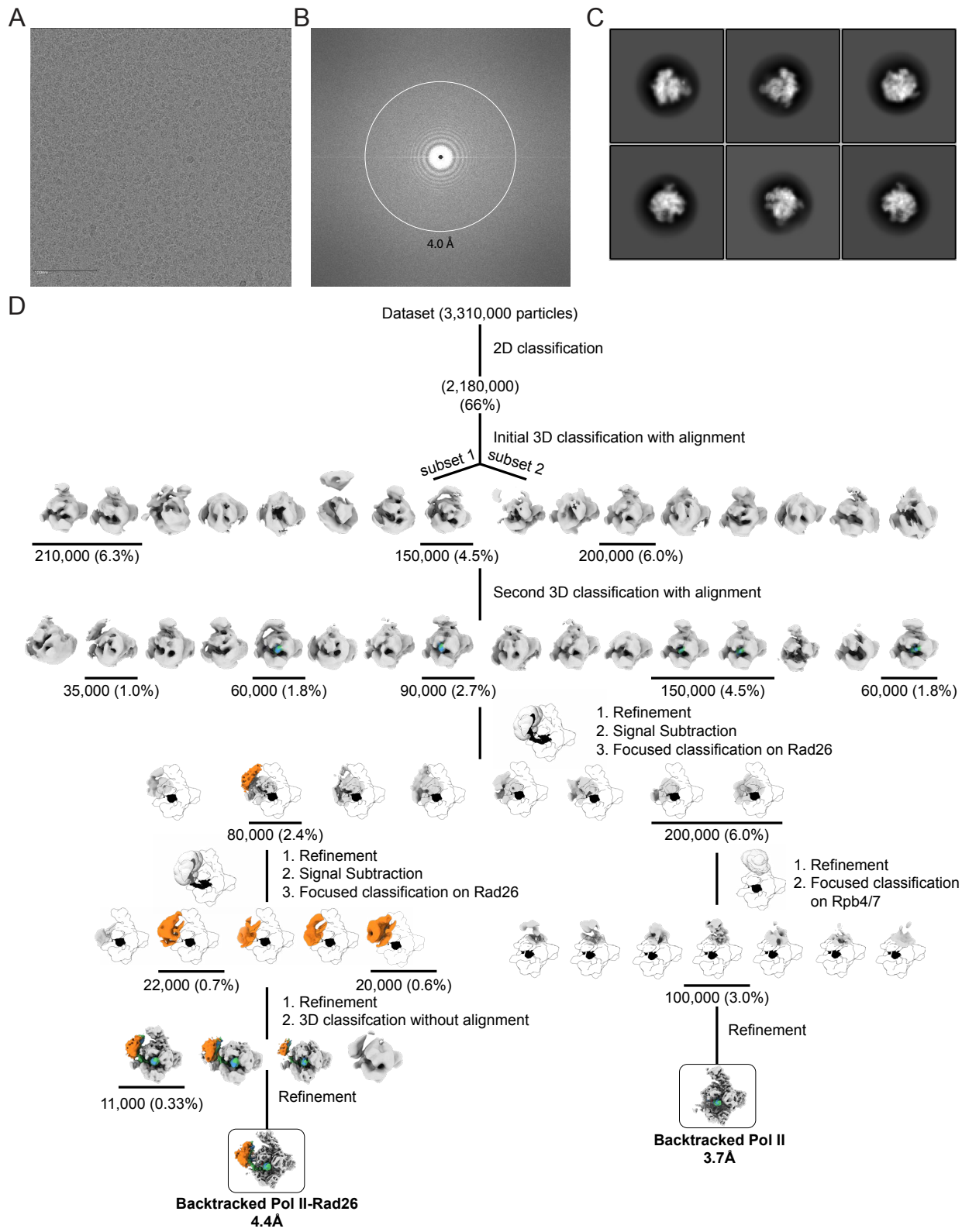
A

B

4.0 Å

C

D

Dataset (3,310,000 particles)

2D classification

(2,180,000)
(66%)

Initial 3D classification with alignment

subset 1   subset 2

210,000 (6.3%)   150,000 (4.5%)   200,000 (6.0%)

Second 3D classification with alignment

35,000 (1.0%)   60,000 (1.8%)   90,000 (2.7%)   150,000 (4.5%)   60,000 (1.8%)

1. Refinement
2. Signal Subtraction
3. Focused classification on Rad26

80,000 (2.4%)   200,000 (6.0%)

1. Refinement
2. Signal Subtraction
3. Focused classification on Rad26

1. Refinement
2. Focused classification
on Rpb4/7

22,000 (0.7%)   20,000 (0.6%)   100,000 (3.0%)

1. Refinement
2. 3D classifcation without alignment

11,000 (0.33%)

Refinement

Refinement

Backtracked Pol II-Rad26
4.4Å

Backtracked Pol II
3.7Å

38

Figure supplement 4 | Analysis of the Backtracked Pol II-Rad26 and Backtracked Pol II cryo-EM maps

(A) Front and back views of locally filtered maps, colored by local resolution, of Backtracked Pol II-Rad26 and Backtracked Pol II. **(B, C)** Euler angle distribution of particle images **(B)** and FSC plots **(C)** for the maps shown in (A). **(D-E)** Close-ups of the cryo-EM densities corresponding to the Rpb1 Bridge helix **(D)**, and the Rpb2/Rpb9 'Jaw' of Pol II **(E)** for the indicated structures with the models fitted in. **(F)** FSC curves for map-to-model fits for the maps shown in (A). The 0.5 FSC line is shown.
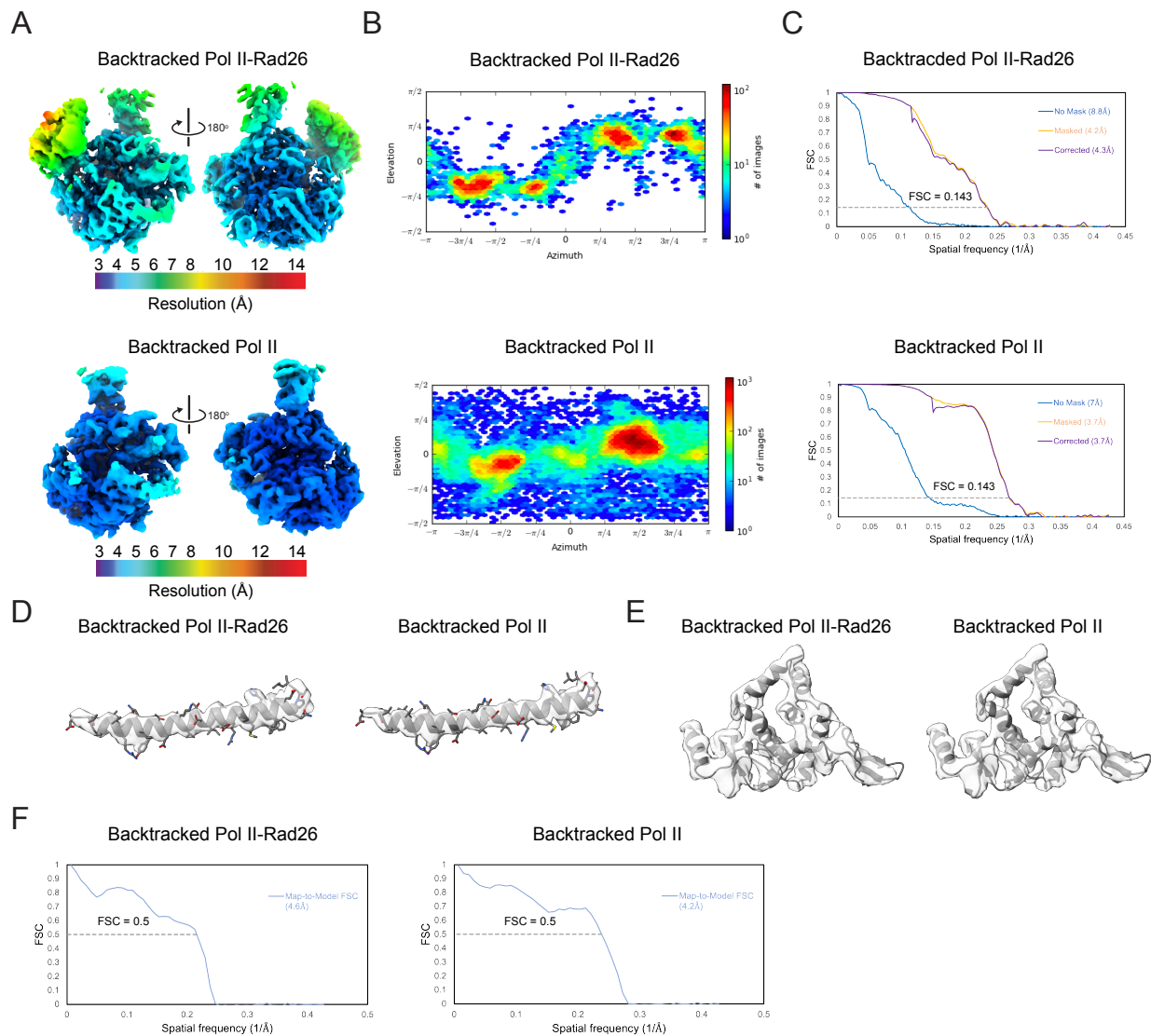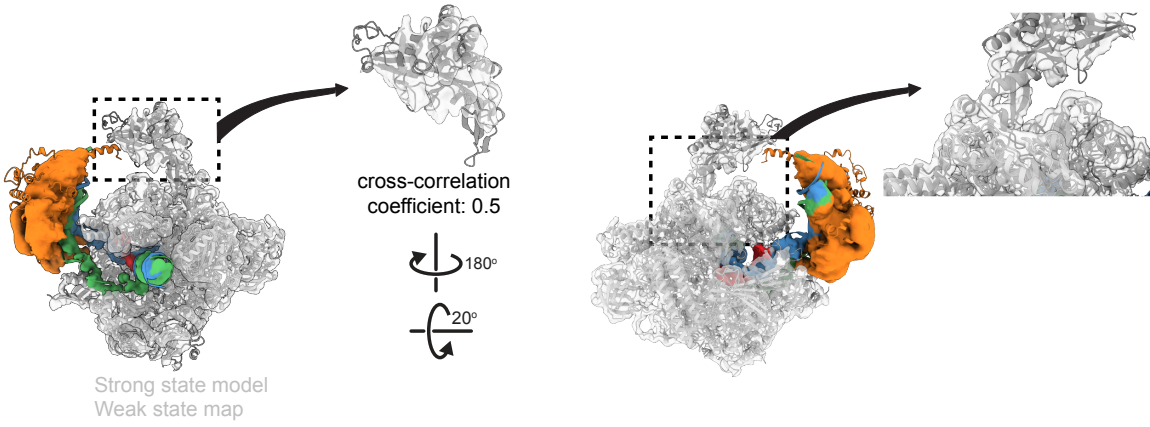
A
Backtracked Pol II-Rad26

Resolution (Å)
3 4 5 6 7 8 10 12 14

Backtracked Pol II

Resolution (Å)
3 4 5 6 7 8 10 12 14

B
Backtracked Pol II-Rad26

Backtracked Pol II

C
Backtracded Pol II-Rad26

Backtracked Pol II

D
Backtracked Pol II-Rad26   Backtracked Pol II

E
Backtracked Pol II-Rad26   Backtracked Pol II

F
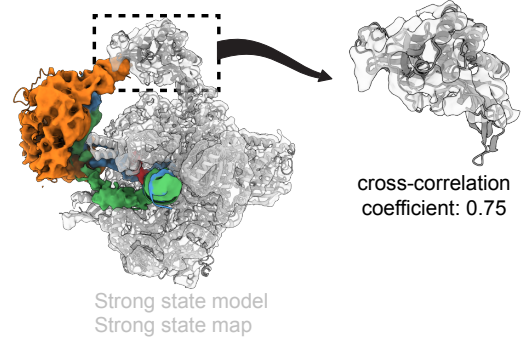Backtracked Pol II-Rad26   Backtracked Pol II

Figure supplement 5 | Structural analysis of Pol II-Rad26 and Pol II(CPD)-Rad26 complexes

(A) Two views are shown of the model for the Strong Rad26-Rpb4/7 state of Pol II(CPD)-Rad26 fitted into the cryo-EM map of the Weak Rad26-Rpb4/7 state of Pol II(CPD)-Rad26, with zoomed-in view of the cryo-EM density of Rpb4/7 with the model fitted in shown to their right. Fitting of the model into the map was driven by the core of Pol II. The cross-correlation coefficient for the fitting of the Rpb4/7 model for the Strong state into the map of the Weak state was 0.5 as reported by Fit-in-Map in ChimeraX. **(B)** Model for the Strong Rad26-Rpb4/7 state fitted into the cryo-EM map for the same state. The cross-correlation coefficient for the fitting of the Rpb4/7 model for the Strong state into the map of the Strong state was 0.75 as reported by Fit-in-Map in ChimeraX. **(C)** Cryo-EM map of the Weak Rad26-Rpb4/7 state of Pol II(CPD)-Rad26 shown at lower threshold, where the interaction between Rad26 and Rpb4/7 becomes apparent. **(D)** Superposition of models for Pol II(CPD)-Rad26 (Strong state) and Pol II-Rad26 (no lesion). The models were aligned using the core of Pol II. Two zoomed-in views of Rpb4/7 from the two models are shown to the right.
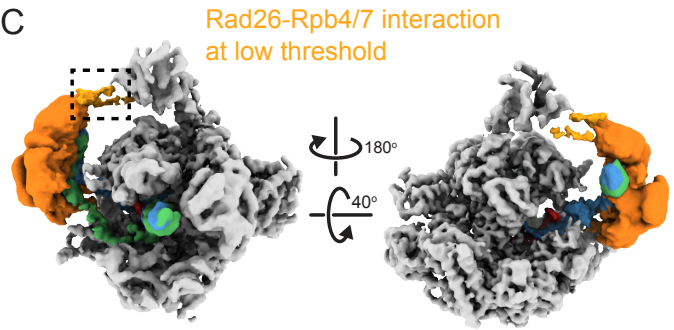
A

Strong state model
Weak state map

cross-correlation
coefficient: 0.5

180°

20°

B

Strong state model
Strong state map

cross-correlation
coefficient: 0.75

C

Rad26-Rpb4/7 interaction
at low threshold

180°

40°

D

Pol II-Rad26 non-lesion
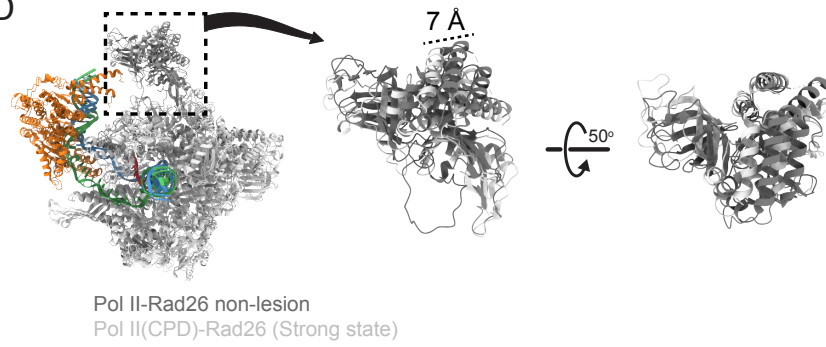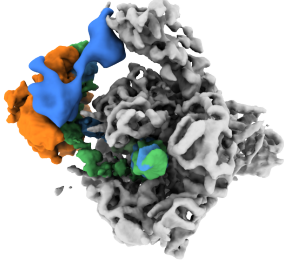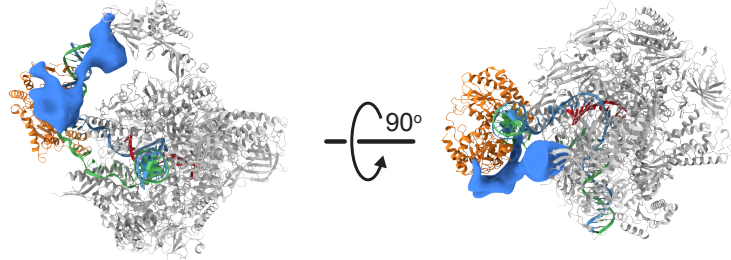Pol II(CPD)-Rad26 (Strong state)

7 Å

50°

Figure supplement 6 | The Rad26-Rpb4/7 interaction is weakest in Backtracked Pol II-Rad26

(A, B) Difference map (in blue) calculated by subtracting Backtracked Pol II-Rad26 from Pol II(CPD)-Rad26 (Strong state), displayed on either **(A)** the cryo-EM density or **(B)** the atomic model for Backtracked Pol II-Rad26. **(C)** Two views are shown of the model for Pol II(CPD)-Rad26 (Strong state) fitted into the cryo-EM map of the Backtracked Pol II-Rad26, with zoomed-in views of the cryo-EM density of Rpb4/7 with the model fitted in shown to their right. The cross-correlation coefficient for the fitting of the Rpb4/7 model for the Strong state into the map of Backtracked Pol II-Rad26 was 0.5 as reported by Fit-in-Map in ChimeraX. **(D)** same as (C), but with the model for Pol II(CPD)-Rad26 (Weak state) fitted into the cryo-EM map of the Backtracked Pol II-Rad26. The cross-correlation coefficient for the fitting of the Rpb4/7 model for the Weak state into the map of Backtracked Pol II-Rad26 was 0.5 as reported by Fit-in-Map in ChimeraX.

A
Strong - Backtracked

Backtracked Pol II-Rad26 map

B
Strong - Backtracked

90°

Backtracked Pol II-Rad26 model

C

cross-correlation
coefficient: 0.5

180°

20°

Pol II(CPD)-Rad26 (Strong state) model
Backtracked Pol II-Rad26 map

D

cross-correlation
coefficient: 0.5

180°

20°

Pol II(CPD)-Rad26 (Weak state) model
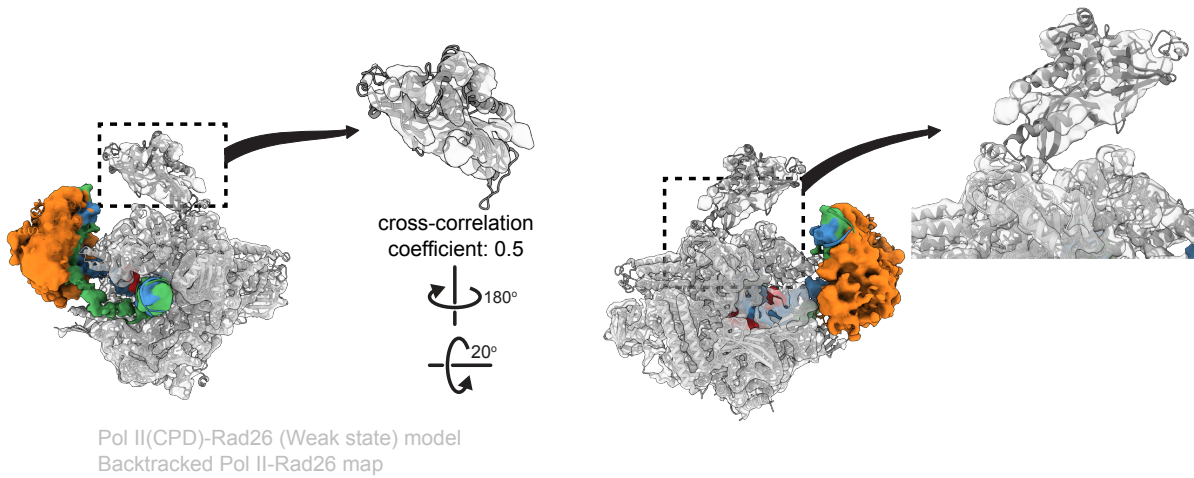Backtracked Pol II-Rad26 map
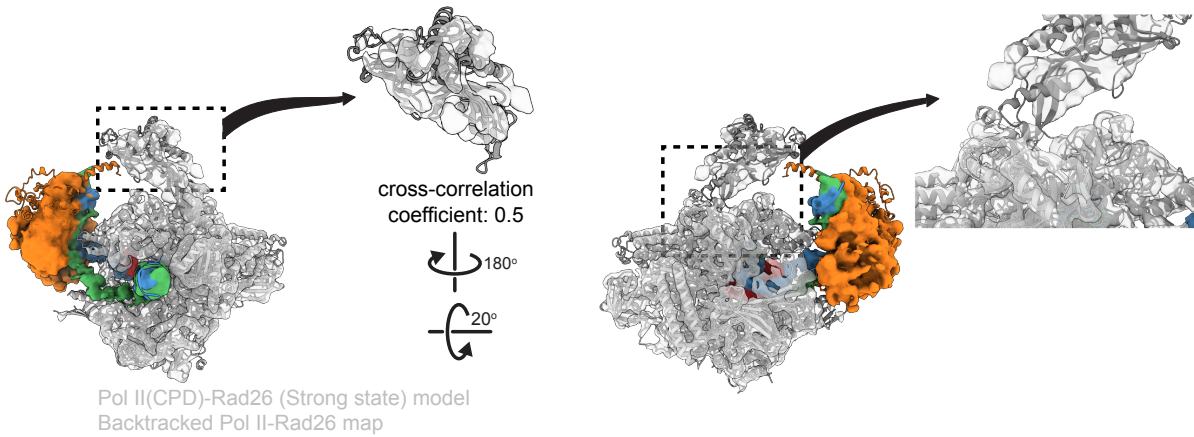
Figure supplement 7 | Cryo-EM structure determination and analysis of the Pol II(CPD)-Rad26 complex with Pol II lacking Rpb4/7

(A-C) Representative micrograph **(A)**, power spectrum **(B),** and representative 2D class averages **(C)** of Pol II(CPD)-Rad26 with Pol II lacking Rpb4/7. **(D)** Schematic representation of the strategy used to sort out the complex particles. The number of particles contributing to each selected structure is indicated. The percentages shown are related to the total number of particles picked from micrographs. The indicated resolution corresponds to the 0.143 Fourier shell correlation (FSC) based on gold-standard FSC curves. **I** Front and back views of locally filtered maps, colored by local resolution, of Pol II(CPD)-Rad26 with Pol II lacking Rpb4/7. **(F-H)** Euler angle distribution of particle images **(F)**, FSC plot **(G)** and FSC curve for the map-to-model fit **(H)** for the map shown in (E). **(I,J)** Close-ups of the cryo-EM densities corresponding to the Rpb1 Bridge helix **(I)**, and the Rpb2/Rpb9 'Jaw' of Pol II **(J)** for the indicated structure with the model fitted in.

A

B

4.0 Å

C

D

Dataset (334,000 particles)

2D classification

(329,000)
98%

Initial 3D classification
with alignment

62,000
(18.6%)

Second 3D classification
with alignment

22,000 (6.6%)

Non-uniform Refinement

Pol II(CPD)-Rad26 with Pol II lacking Rpb4/7
4.6Å

E

180°

4 5 6 7 8 9 11 13 15
Resolution (Å)

F

π/2

π/4

0

−π/4

−π/2

−π   −3π/4 −π/2 −π/4 0 π/4 π/2 3π/4 π

Azimuth

Elevation

# of images

10^2

10^1

10^0

G

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45

Spatial frequency (1/Å)

FSC

FSC = 0.143

No Mask (8.5Å)
Masked (4.6Å)
Corrected (4.6Å)

H

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

0 0.1 0.2 0.3 0.4 0.5

Spatial frequency (1/Å)

FSC

FSC = 0.5

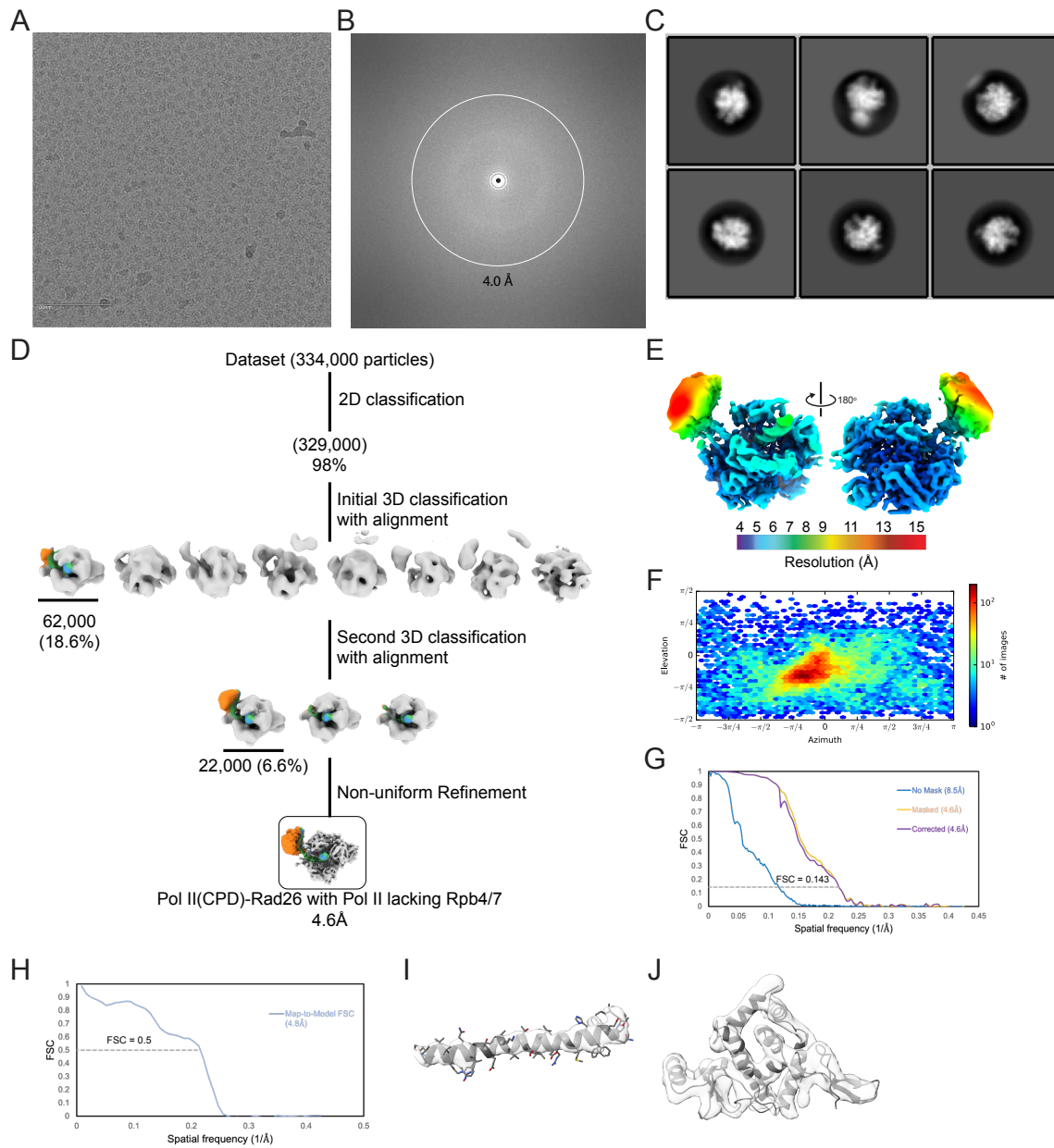Map-to-Model FSC
(4.8Å)

I

J

46

Figure supplement 8 | Multi-body refinement of Pol II(CPD)-Rad26 complex

(A) A consensus refinement of Pol II(CPD)-Rad26 in the Strong Rad26-Rpb4/7 state was used to generate three masks (Pol II core, Rpb4/7, and Rad26) and as the starting model for multi-body refinement. **(B)** Maps generated from multi-body refinement, showing improved resolutions. These maps were used to build models and to make segmented maps. **(C)** Multi-body refinement was performed with the same map used in (A) but with only two masks (Rbp4/7 and Pol II + Rad26) to focus the analysis on the flexibility in Rpb4/7. **(D)** Principal component analysis of the multi-body refinement was performed in Relion-3 (this analysis was the source of the data shown in fig. S10**).**
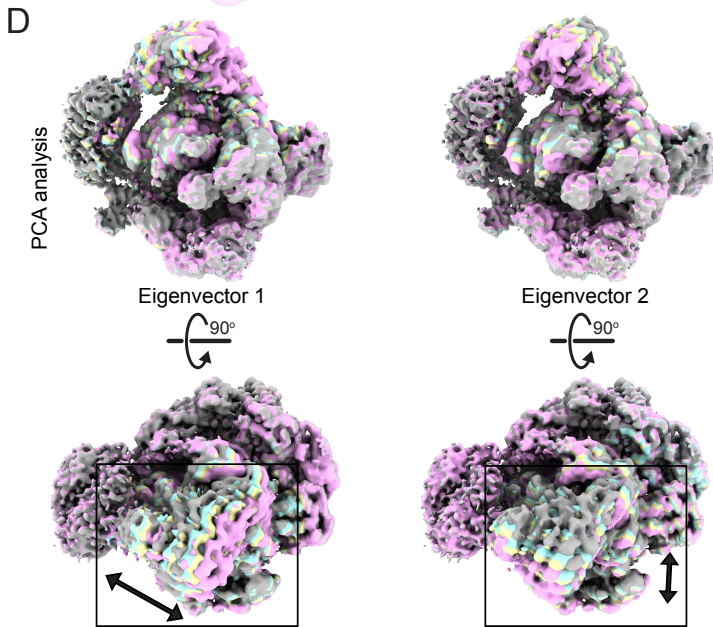
A — Pol II core, Rpb4/7, Rad26

B — Multibody refinement

C — Pol II + Rad26, Rpb4/7

D — PCA analysis

Eigenvector 1   90°

Eigenvector 2   90°

Figure supplement 9 | Multi-body refinement of Pol II(CPD) complex

(A-C) Two conformations of Pol II(CPD) were solved through focused 3D classification on Rpb4/7 (see fig. S1). **(A)** Model of Conformation 1 fitted into its density. **(B)** Model of Conformation 2 fitted into its density. **(C)** The two models were aligned by the core of Pol II and close-ups of Rpb4/7 are shown to the right. **(D)** A consensus refinement of Pol II(CPD) (with the particles for Conformations 1 and 2 combined) was used to generate two masks (Pol II core and Rpb4/7) and as the starting model for multi-body refinement. **(E)** Principal Component Analysis of the multi-body refinement was performed in Relion-3 (this analysis was the source of the data shown in fig. S10).
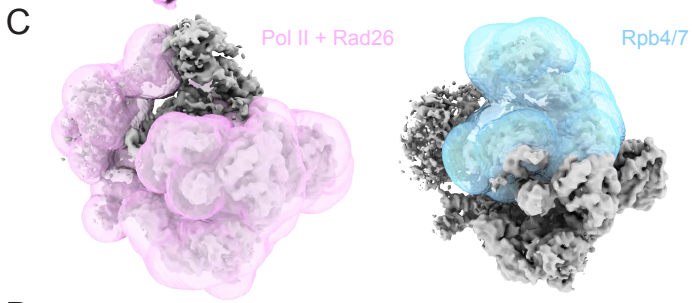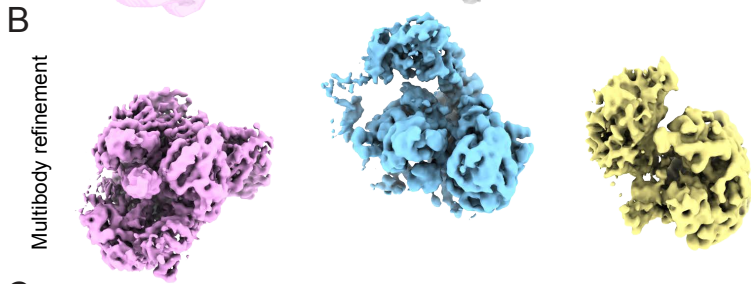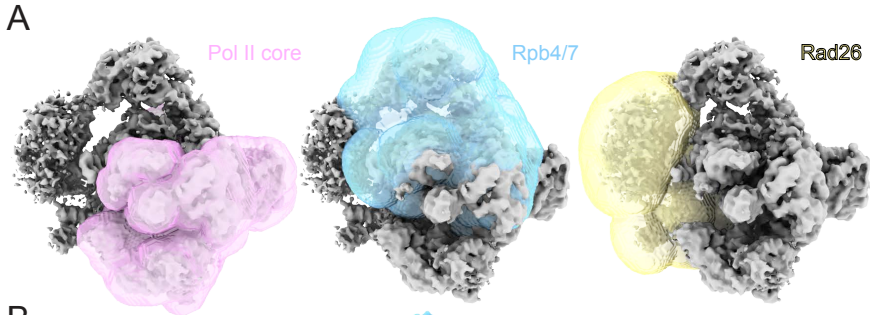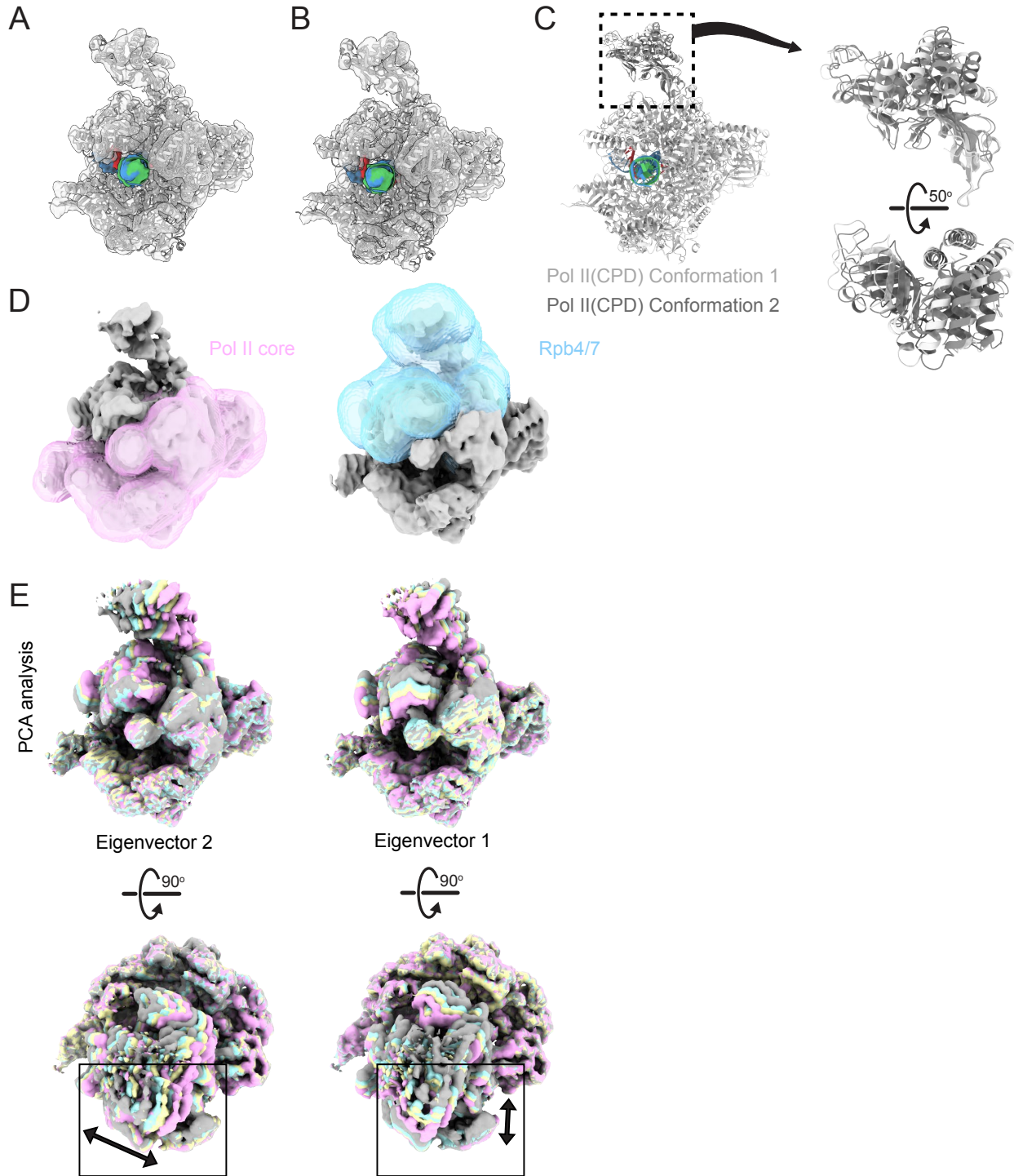
A

B

C

Pol II(CPD) Conformation 1
Pol II(CPD) Conformation 2

$-50°$

D

Pol II core

Rpb4/7

E

PCA analysis

Eigenvector 2

Eigenvector 1

$-90°$

$-90°$

Figure supplement 10 | Rpb4/7 is conformationally flexible and its main axis of motion points towards Rad26

(A) Multi-body refinement (and Principal Component Analysis) was performed on the data for the Strong Rad26-Rpb4/7 state of Pol II(CPD)-Rad26, with one body corresponding to Rpb4/7 and a second to the rest of the complex (see fig. S8). Models were generated for the ten maps representing motion along the top two eigenvectors. The ten models are superimposed in this panel, with the extreme positions indicated in dark and light purple and the intermediate ones in grey. The overall complex is shown on the left, with two zoomed-in views of Rpb4/7 (indicated by the box) shown to the right. The arrows on the rightmost panels indicate the direction of motion described by the eigenvectors. **(B)** Same as in (A) with Pol II(CPD) (i.e. without Rad26) (see fig. S9).
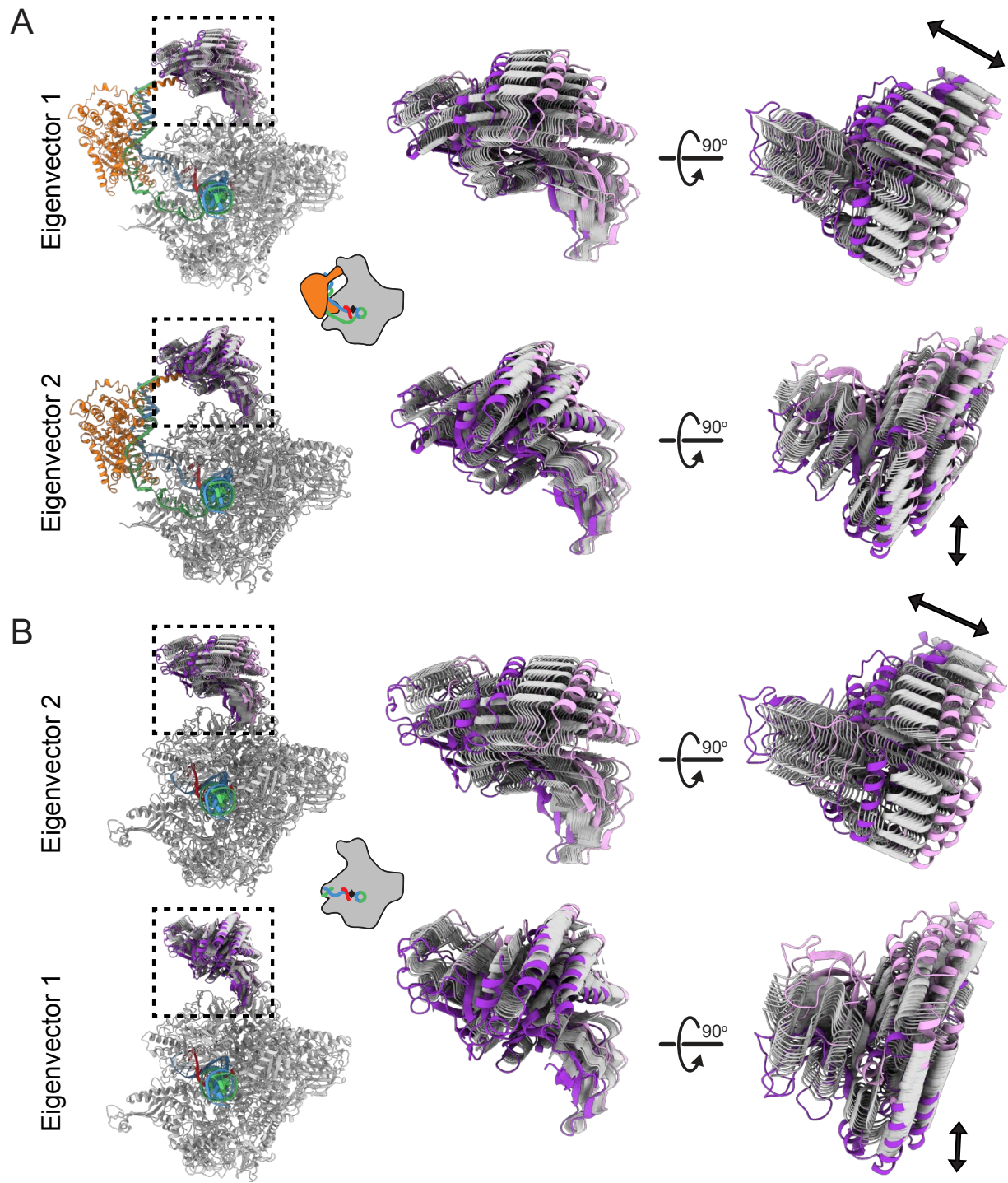
Figure supplement 11 | Purification of Elf1 and comparison of our structure of Pol II(CPD)-Rad26-Elf1 with the published structure of Pol II-Spt4/5-Elf1

(A) Sequence alignment of Elf1 orthologs from *S.cerevisiae* (*Sc*), *S.pombe* (*Sp*), humans (*h*) and *C. elegans* (*Ce*). **(B)** SDS-PAGE of purified Elf1 and Elf1ΔC (1-85), shown schematically at the top. **(C)** Deletion of Elf1 leads to UV sensitivity and can be rescued by Elf1ΔC (1-85). **(D-E)** Structures of **(D)** Pol II(CPD)-Rad26-Elf1 (this work) and **(E)** Pol II-Spt4/5-Elf1 (PDB: 6J4Y)(55). **(F)** Superimposition of the two models in (D) and (E).

A

| | 1 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 |

*Sc*Elf1 MGKRKKSTRKPTKR.LVQKLDTKFNCLFCNHEKSVSCTLDKKNSIGTLSCKICGQSFQTRINSLSQPVDVYSDWFDAVEEVNSGRGSDTDDGD...EGSD
*Sp*Elf1 MGKRKAKAKVKPKR.RAPPLDTFTCLFCNHEKSVSCSLDKQSGVGNLHCKICGQSHQCLITALSAPIDVYSDWIDACDAVANQAKEVDNINDQEAYSPE
*h*ELOF1 MGRRKSKRKPPPKKKMTGTLETQFTCPFCNHEKSCDVKMDRARNTGVISCTVCLEEFQTPITYLSEPVDVYSDWIDACEAANQ...............
*Ce*Elf1 MGKRKSKRKAPTKAKAVMPLDTQFNCPFCNHERVCEVKMDREKNVGYISCRVCSEDFQTNINYLSEPIDVYSDWVDACEQANNA...............

| | 100 | 110 | 120 | 130 | 140 |

*Sc*Elf1 SDYESDSEQDAKTQNDGEIDSDEEEVDSDEERIGQVKRGRGALVDSDDE
*Sp*Elf1 AAYEEEKE...
*h*ELOF1 ....
*Ce*Elf1 ....

B

Elf1

ELOF1

C

unirradiated   20 J/m²   30 J/m²   40 J/m²

WT
*rad16*Δ
*rad16*Δ *elf1*Δ
*rad16*Δ*elf1*ΔC(1-85)

D   E   F

Rad26   Spt4/5

Elf1   Elf1

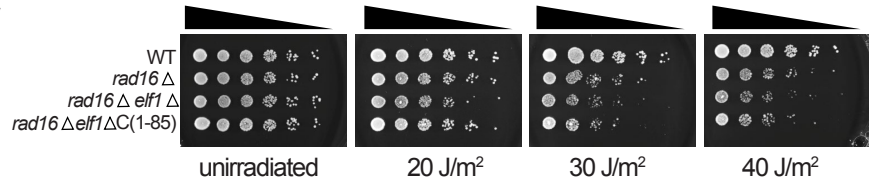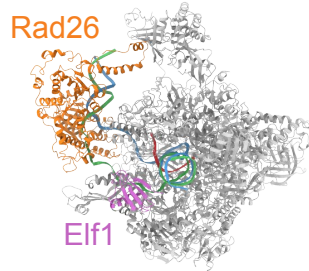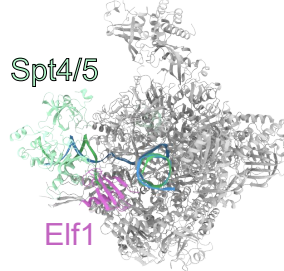**Figure supplement 12 | Cryo-EM structure determination of the Pol II(CPD)-Rad26-Elf1 complex**

(A-C) Representative micrograph **(A),** power spectrum **(B),** and representative 2D class averages **(C)** of the Pol II(CPD)-Rad26-Elf1 complex. **(D)** Schematic of the strategy used to sort out the dataset. Focused 3D classification was performed without alignment unless otherwise noted. The number of particles contributing to each selected structure is indicated. The percentages shown are related to the total number of particles picked from the micrographs. The indicated resolution corresponds to the 0.143 Fourier shell correlation (FSC) based on gold-standard FSC curves (see fig. S13).

A

B

4.0 Å

C

D

Dataset 1 (1,820,000 particles)

2D classification

(1,580,000)
(87%)

Initial 3D classification
with alignment

255,000 (8.5%)   210,000 (7%)

Dataset 2 (1,180,000 particles)

2D classification

(988,000)
(84%)

Initial 3D classification
with alignment

243,000 (8.1%)   105,000 (3.5%)

1. Refinement
2. Signal substraction
3. Focused classification on Rad26

19,000 (0.6%)   114,000 (3.8%)   71,000 (2.4%)   23,000 (0.77%)   30,000 (0.1%)   27,000 (0.9%)

1. Refinement
2. Signal substraction
3. Focused classification on Rad26

50,000
(1.7%)

40,000
(1.3%)

27,400
(0.9%)

35,000
(1.2%)

Refinement   Refinement   Refinement   Refinement
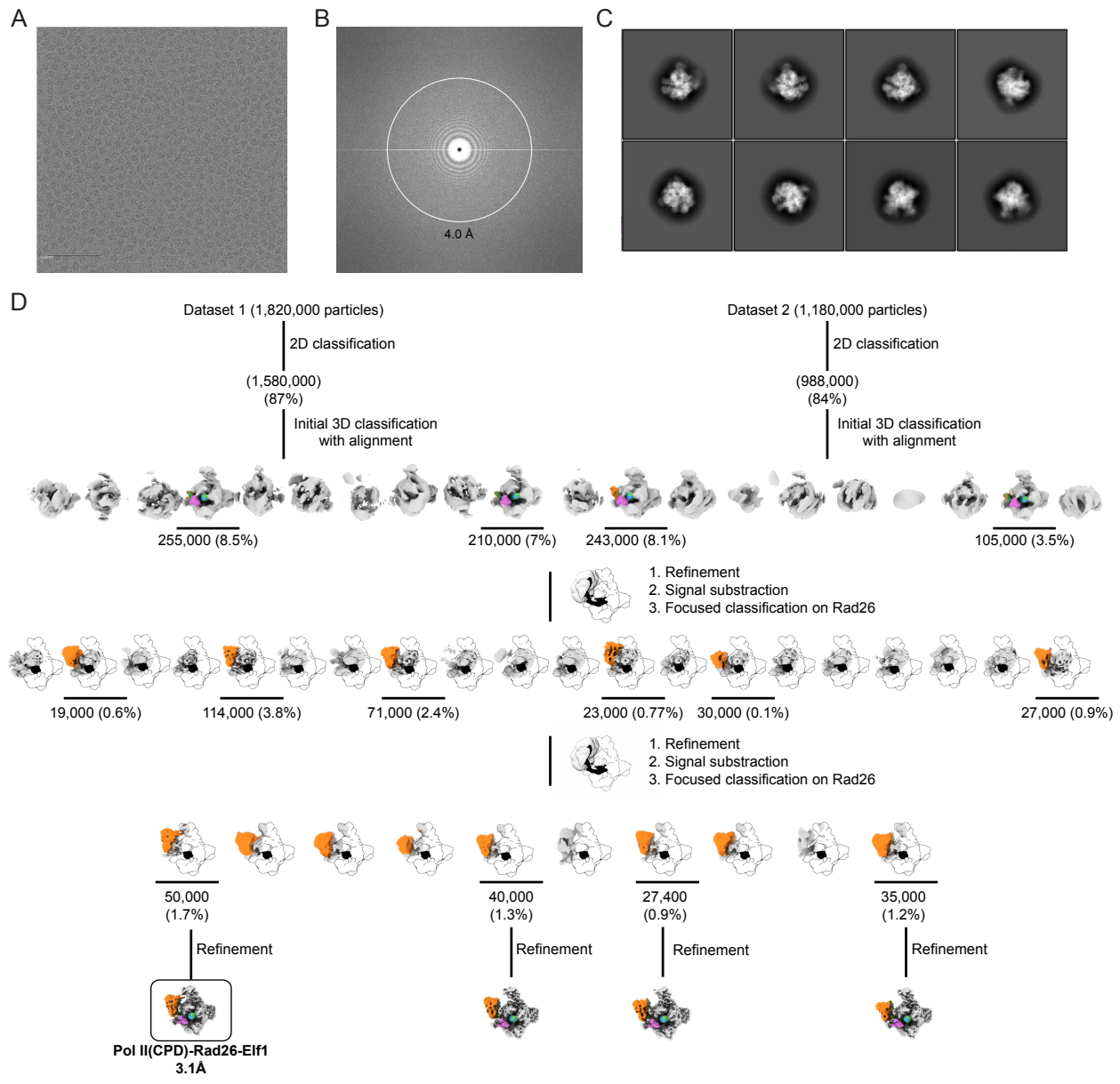
Pol II(CPD)-Rad26-Elf1
3.1Å

Figure supplement 13 | Analysis of the Pol II(CPD)Rad26-Elf1 cryo-EM map

(A) Front and back views of locally filtered maps, colored by local resolution. **(B, C)** Euler angle distribution of particle images **(B)** and FSC plots **(C)** for the map shown in (A). **(D-F)** Close-ups of the cryo-EM densities corresponding to the Rpb1 Bridge helix **(D)**, the Rpb2/Rpb9 'Jaw' of Pol II **(E)** and the Rad26 HD2-1 'wedge' **(F)** for the indicated structures with the models fitted in. **(G)** FSC curves for map-to-model fit for the map shown in (A). The 0.5 FSC line is shown. **(H, I)** Difference map (in blue) calculated by subtracting Pol II(CPD)Rad26 (Strong state) from Pol II(CPD)-Rad26-Elf1, displayed on either **(H)** the cryo-EM density or **(I)** the atomic model for Pol II-CDP-Rad26 (Strong state).

A

2.8 3.5 4   5   6   8
Resolution (Å)

B

Elevation — Azimuth — # of images

C

FSC
No Mask (7.1Å)
Masked (3.1Å)
Corrected (3.1Å)
FSC = 0.143
Spatial frequency (1/Å)

D

E

F

G

FSC
Map-to-Model FSC (3.0Å)
FSC = 0.5
Spatial frequency (1/Å)

H

Pol II(CPD)-Rad26-Elf1 - Pol II(CPD)-Rad26 (Strong State)

160°
10°

Pol II(CPD)-Rad26 (Strong state) map

I

90°

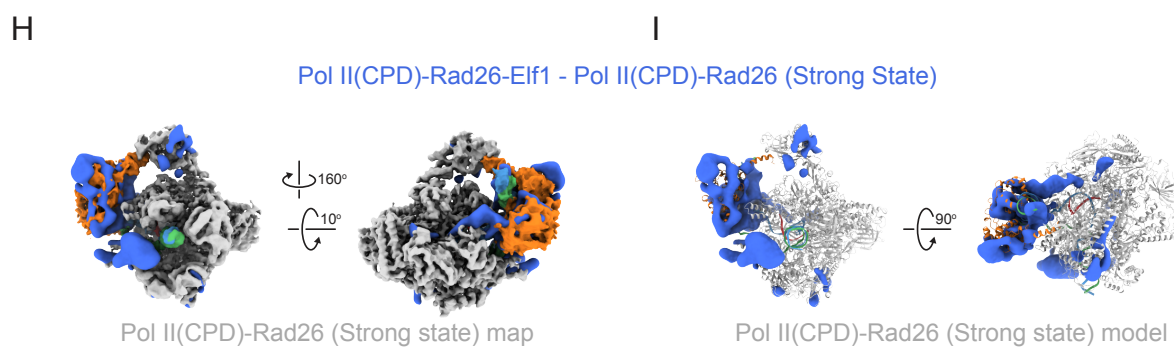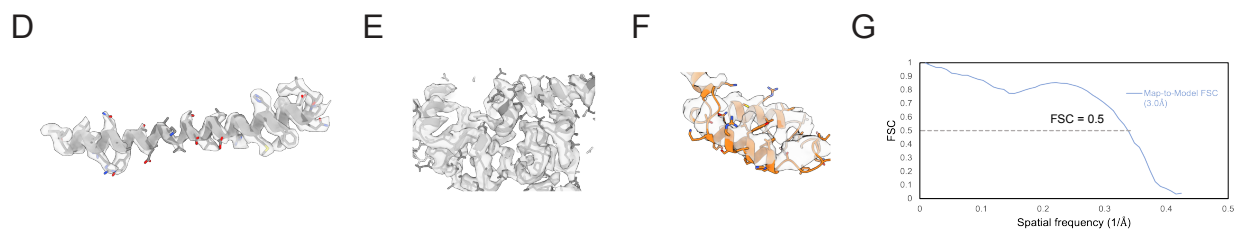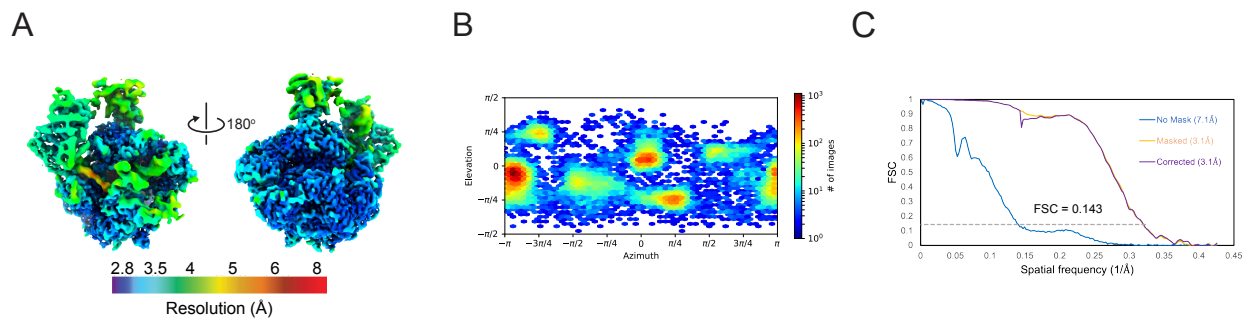Pol II(CPD)-Rad26 (Strong state) model

58

Figure supplement 14 | Rad26 and TFIIS promote the displacement of backtracking Pol II from the upstream of a CPD lesion

(A) Schematic representation of the experimental design. **(B)** Denaturing PAGE shows RNA transcripts in the Supernatant (S) and Bound to beads (B) in the absence of TFIIS. The experiment was performed with either 12-subunit or 10-subunit Pol II and in the presence of either wild type or ATPase dead Rad26 (K328R). For the analysis, we divided the transcripts as arising from three regions: Pol II stalled at the CPD (region I), and backtracked Pol II products that are proximal (region II) or distal to the CPD (region III). Region III was further defined as that portion of the transcription scaffold where Pol II could be displaced in the absence of TFIIS. **(C)** Line scans of the gel in (B) show RNA transcripts in region 1 are not released. **(D)** Quantification of relative displacement of Pol II for regions I, II and III in the absence of TFIIS. **(E, F)** Denaturing PAGE shows RNA transcripts in the Supernatant (S) and Bound to beads (B) in the presence of TFIIS with 10-subunit E or 12-subunit (F). **(G)** Line scans of the gel in (B) show RNA transcripts in the presence (top chart) and absence (bottom chart) of NTPs. **(H)** Quantification of relative displacement of Pol II for regions I, II and III. **(I)** Same as (G) but with ATPase dead Rad26 mutant (K328R). The assays included dATP to support Rad26's translocation activity. * $p < 0.05$. **(J)** Proposed model for the displacement of Pol II by the combined action of Rad26 and TFIIS. Initial backtracking of Pol II is followed by TFIIS cleavage and Rad26-aided Pol II hypertranslocation. These combined activities lead to release of Pol II and eventually to DNA repair (downstream TC-NER).
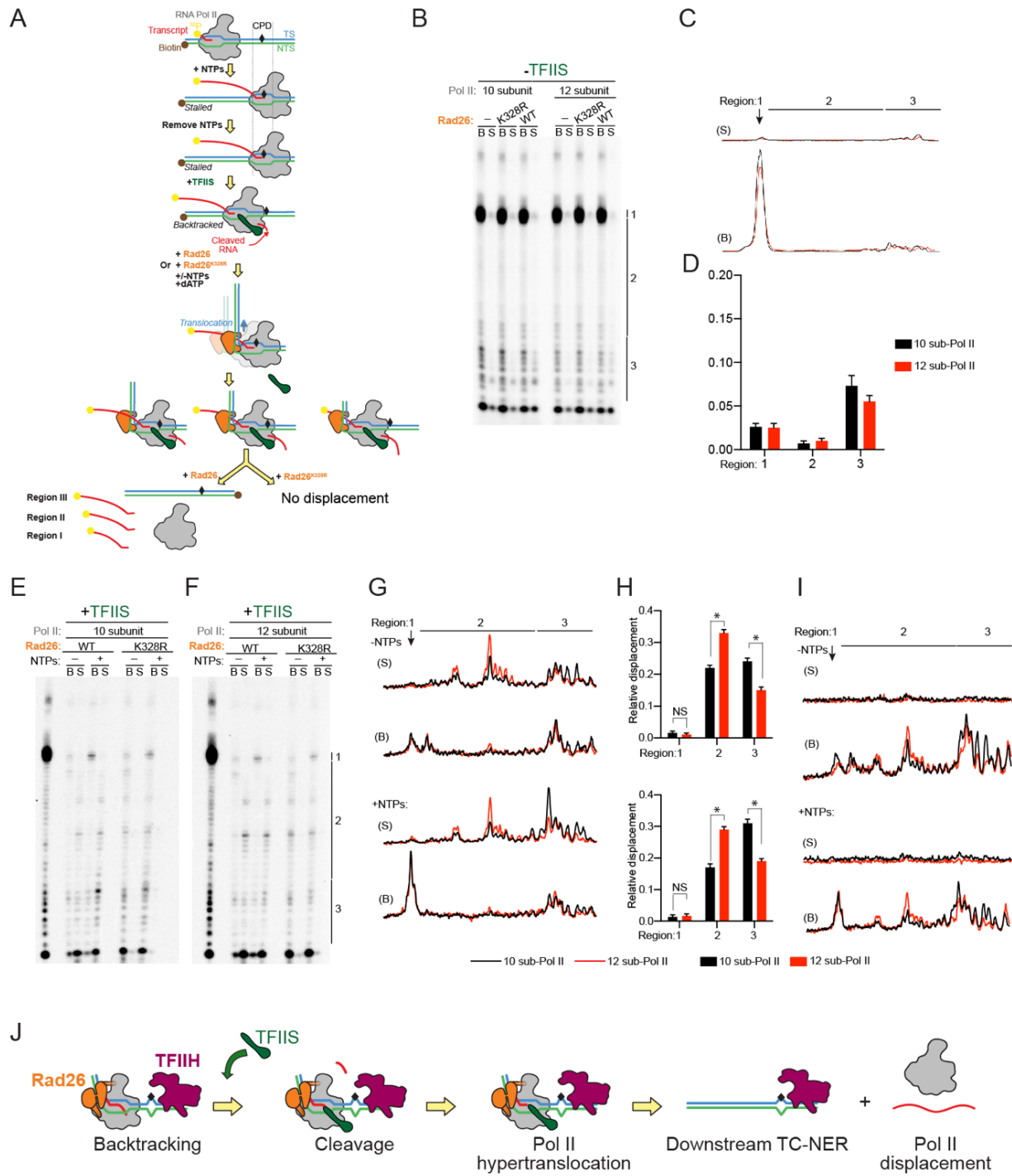
Figure supplement 15 | Rad26 can rescue Pol II-Spt4/5 from an arrest induced by Py-Im but not from one induced by a CPD-lesion

(A) Structures of Pol II(CPD)-Rad26 (this work) and Pol II-Spt4/5 (PDB: 5OIK) (56). **(B)** Rad26 cannot rescue a CPD-lesion arrested Pol II. Run-off transcript (RO) was obtained from transcription on a non-lesion template as a control. **(C)** Rad26 can rescue Pol II-Spt4/5 from an arrest induced by Py-Im.

A

Spt4/5

Rad26

B

DNA lesion
[Pyrimidine dimer (CPD)]

Time

Buffer   Spt4/5   Rad26

RO

start   CPD

C

Py-Im

Spt4/5   Rad26
+Spt4/5

Time

RO

start
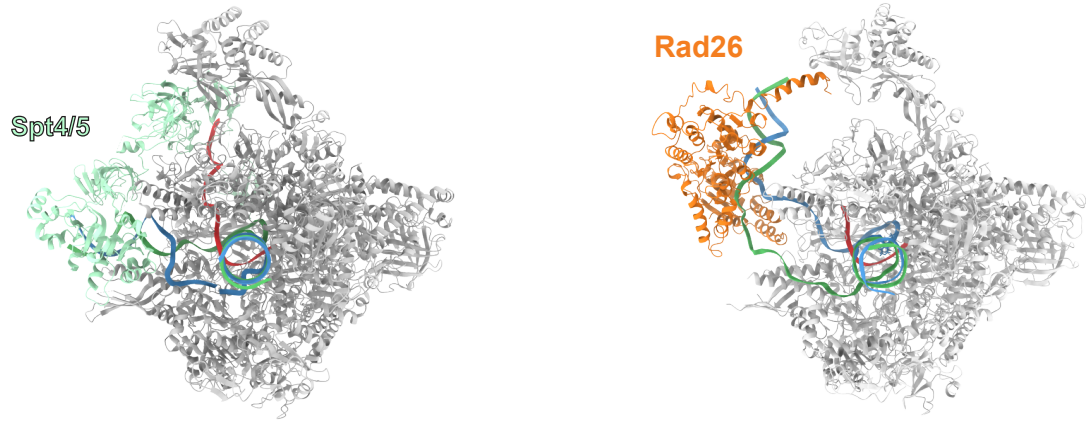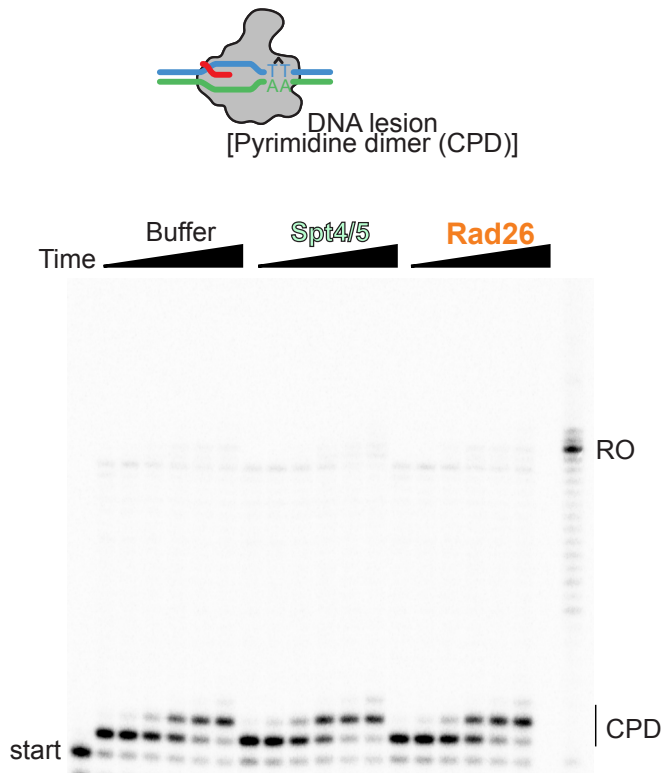
SUPPLEMENTARY TABLES

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

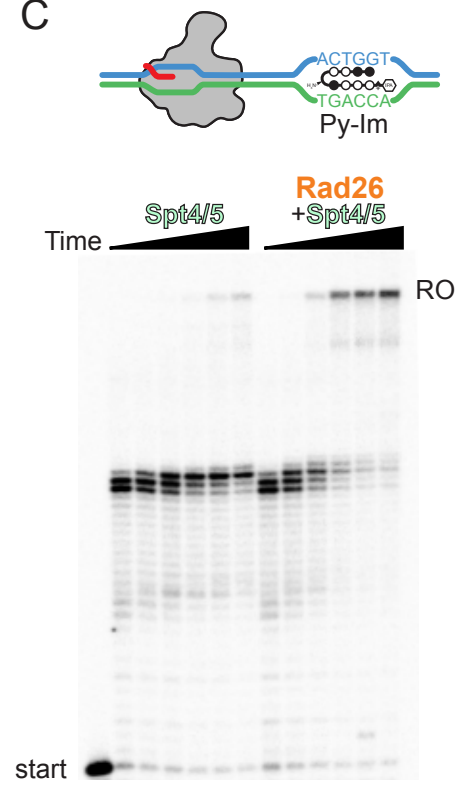|	Pol II(CPD)-Rad26 Dataset				10-subunit
	Strong state	**Weak state**	**Pol II(CD) Conf1**	**Pol II(CPD) Conf2**	
Data Collection					
Microscope			Talos Arctica		Talos Arctica
Camera			K2 Summit		K2 Summit
Camera Mode			Counting		Super-Res
Voltage (kV)			200		200
Magnification			36,000		36,000
Pixel Size (Å/pixel)			1.16		1.16
Dose rate (e-/Å^2 second)			8.4		4
Total dose (e-/Å^2)			59		52
Number of frames			47		52
Defocus range (µm)			0.6-2.5		0.6-2.5
Micrographs collected (no.)			3,358		955
Initial particle (no.)			1,620,000		334,000
Final particle (no.)	20,000	25,000	74,000	73,000	22,000
Refinement					
Initial model used			1Y77		1Y77
Final resolution (Å) (0.143 FSC threshold)	3.5	3.7	3.6	3.7	4.6
Map sharpening *B* factor (Å^2)	-73	-68	-87	-101	-153
Model Refinement					
Map-to-model resolution (Å) (0.5 FSC threshold)	3.6	3.8	3.8	3.9	4.8
Model Composition					
Nonhyrogen atoms	71,184	69,002	61,459	60,948	64,987
Protein residues	4,251	4,182	3,748	3,748	3,869
Nucleotides	103	103	56	56	100
Ligands	9	9	9	9	9
B factor (Å^2)	244	205	123	170	283
R.m.s. deviations					
Bond length (Å)	0.003	0.003	0.003	0.003	0.003
Bond angle (°)	0.56	0.54	0.59	0.60	0.644
Validation					
MolProbity score	1.7	1.8	1.7	1.7	2.1
Clash score	6.7	8.6	7.0	9.2	14.9
Poor rotamers (%)	0	0	0	0	0
Ramachandran					
Favored (%)	95.53	95.60	96.0	95.82	93.79
Allowed (%)	4.4	4.3	4.0	4.1	6.1
Disfavored (%)	0.07	0.1	0.05	0.08	0.11

Table supplement 1 | Cryo-EM data collection, refinement and validation statistics (continued)

	Backtracked Pol II-Rad26 Dataset		Pol II(CPD) Rad26-Elf1
	Backtracked Pol II-Rad26	**Backtracked Pol II**	
Data Collection			
Microscope		Talos Arctica	Talos Arctica
Camera		K2 Summit	K2 Summit
Camera Mode		Counting	Counting
Voltage (kV)		200	200
Magnification		36,000	36,000
Pixel Size (Å/pixel)		1.16	1.16
Dose rate (e-/Å^2 second)		5	5/5.5
Total dose (e-/Å^2)		55	50/55
Number of frames		55	50
Defocus range (μm)		0.6-2.5	0.6-2.5
Micrographs collected (no.)		9,167	8,000
Initial particle (no.)		3,310,000	3,000,000
Final particle (no.)	11,000	100,000	50,000
Refinement			
Initial model used	1Y77	1Y77	1Y77
Final resolution (Å) (0.143 FSC threshold)	4.4	3.7	3.1
Map sharpening *B* factor (Å^2)	-92	-117	-85
Model Refinement			
Map-to-model resolution (Å) (0.5 FSC threshold)	4.6	4.2	3.0
Model Composition			
Nonhyrogen atoms	70,048	61,419	77,243
Protein residues	4,182	3,747	4,701
Nucleotides	104	56	103
Ligands	9	9	9
B factor (Å^2)	266	149	89
R.m.s. deviations			
Bond length (Å)	0.003	0.003	0.003
Bond angle (°)	0.553	0.555	0.594
Validation			
MolProbity score	1.8	1.7	1.9
Clash score	7.9	8.1	7.9
Poor rotamers (%)	0	0	0
Ramachandran			
Favored (%)	95.36	95.88	92.49
Allowed (%)	4.52	4.07	7.45
Disfavored (%)	0.12	0.05	0.06

Table supplement 2 | *Saccharomyces cerevisiae* strains

Strain	Genotype	Background	Source
NH0256	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	S288c BY4741	ATCC
DDY4765	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rad16::NATMX*	S288c BY4741	Ref. 14
DDY4776	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 elf1::KANMX rad16::NATMX*	S288c BY4741	Ref. 14
NH1356	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 rad16::NATMX elf1-85::URA3*	S288c BY4741	This study