nature communications

Article

https://doi.org/10.1038/s41467-025-57513-2

CTDP1 and RPB7 stabilize Pol II and permit reinitiation

Received: 22 May 2024

Accepted: 25 February 2025

Published online: 04 March 2025

Check for updates

Haonan Zheng[1,2], Qiqin Xu[1,2], Dexun Ji[1], Boqin Yang[1] & Xiong Ji[1]

The mechanisms governing the termination and subsequent reinitiation of RNA polymerase II (Pol II) remain poorly understood. Here we find that depletion of RPB7 leads to the destabilization of Pol II's largest subunit, RPB1. This destabilization is influenced by the loop regions of RPB7, CDK9, the C-terminal domain (CTD) of RPB1, and its linker region. The stabilization process of RPB1 is regulated by the E3 ubiquitin ligase Cullin 3. Additionally, RPB7 interacts with the phosphatase CTDP1, which is crucial for maintaining RPB1 stability. RPB7 is also vital for the reinitiation of Pol II, engages with RNA processing factors, and is localized to the RNA exit channel of the Pol II complex. The absence of RPB7 compromises RNA processing. We propose that RPB7 recruits CTDP1 to dephosphorylate Pol II, enhancing its stability and facilitating efficient reinitiation, adding an emerging dimension to transcriptional regulation.

Polymerase II (Pol II) transcription is well controlled to ensure precise gene expression[1–10]. The general transcription factor TFIIH-associated CDK7 kinase is known to phosphorylate Pol II CTD. The transition from promoter pausing to productive elongation requires the kinase activity of P-TEFb, comprised of CDK9 and CyclinT1, to phosphorylate transcription complex and the Pol II CTD[11–17]. During the termination phase, PNUTS-PP1 recognizes the poly(A) site, and dephosphorylates SPT5, leading to the deceleration of Pol II. Subsequently, the slowed Pol II becomes a target for exonuclease XRN2, which facilitates the dissociation of Pol II from the chromatin[18–21]. It has been thought that Pol II needs to be dephosphorylated after terminated from the DNA and needs to be phosphorylated again when reinitiating at the promoters to start a new round of the transcription process[22–26]. This step is the least well characterized among the processes involved in transcription because of the lack of known regulators in mammalian cells.

The stability of the Pol II complex is essential for the continuity of transcription cycles[27–29]. Pol II degradation is a common response to transcription stress. For example, treatment with the Pol II inhibitor α-amanitin, UV irradiation, or DNA lesions caused by cisplatin lead to the removal and degradation of stalled or arrested Pol II from chromatin[30–33], which is vital for halting transcription during stress and facilitating DNA damage repair. Recent studies have also shown that SPT5, a component of DSIF, stabilizes Pol II during transcription pausing, elongation, and termination[29,34,35]. However, it is unclear whether or how Pol II is stabilized after termination and released from chromatin before reinitiation.

RPB7 is a specific subunit of Pol II and forms a heterodimer with RPB4[36,37]. Previous research has shown that RPB4/7 can regulate mRNA decay and translation in the cytoplasm[38,39], interact with phosphatases SSU72 and FCP1 to facilitate the dephosphorylation of Pol II CTD[40], and promote terminator and promoter looping, but no effects on the protein levels of total RPB1[25,41]. However, all these studies were mutagenesis studies carried out in yeast or using RNA interference in *Drosophila*, and it is unknown whether the exact mechanisms are conserved in mammalian cells. We previously employed an auxin-inducible degron system to degrade Pol I, Pol II and Pol III, and observed their cross-regulation when they were close to each other[42,43]. In addition, we recently degraded 12 Pol II subunits and found that certain Pol II subunits are preferentially responsible for the proper expression of subsets of genes[44], indicating that Pol II subunits could also be regulators of Pol II activities.

Here, we found that RPB7 interacts with the phosphatase CTDP1, which is essential for maintaining the stability of Pol II and enabling its reinitiation. Further investigation revealed that the loop region of

[1]State Key Laboratory of Gene Function and Modulation Research, Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education, Beijing Advanced Center of RNA Biology (BEACON), School of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China. [2]These authors contributed equally: Haonan Zheng, Qiqin Xu. e-mail: xiongji@pku.edu.cn

RPB7, the phosphorylation state of Pol II, the CTD and linker region of RPB1, and the phosphatase activity of CTDP1 are all necessary for Pol II stability. The interaction between RPB7 and CTDP1 not only maintains Pol II stability but also dephosphorylates Pol II to prepare it for reinitiation. These findings provide informative insights into the mechanisms of transcriptional regulation.

Results

Acute RPB7 depletion causes the destabilization of RPB1

To investigate the immediate functions of RPB7 in transcription, we used previously developed auxin-inducible RPB7 degron mESCs[44]. We conducted mass spectrometry analysis of the whole-cell extracts and chromatin fractions before and after RPB7 depletion for 3 h. This approach not only enabled us to quantitatively assess the degradation of RPB7 but also offered insights into the protein levels of other Pol II subunits and associated factors. The mass spectrometry data consistently indicated RPB1 decreases more than other components of the Pol II complex, except for RPB4 (Fig. 1a, b and Supplementary Fig. 1a–c). Proteins with reduced abundance in chromatin fractions are enriched in RNA metabolism and transcription (Fig. 1b). We further examined total, serine 5, serine 2-phosphorylated RPB1, and Pol II small subunits in these cells with western blotting. The results showed that RPB7 depletion led to the destabilization of RPB1 (Fig. 1c).

We then constructed RPB4 N-terminal degron mESCs and found that the RPB4 N-terminal degron also caused the destabilization of RPB1 (Fig. 1c), which supports the conclusion that RPB4 and RPB7 have similar functions, as they form a heterodimer[36]. We previously found that inserting the degron tag into the C-terminus of RPB4 did not have apparent effects on gene expression[44]. RPB4 has splicing isoforms in mESCs (Supplementary Fig. 1d). We confirmed through genotyping experiments that both the C-terminal and N-terminal RPB4 degron cells were homozygous (Supplementary Fig. 1e). Western blot analysis revealed that the C-terminal degron cells still exhibited the endogenous RPB4 band, while the N-terminal degron cells did not (Supplementary Fig. 1f). The cell growth and affected gene expression are also different (Supplementary Fig. 1g, h). This illustrates the difference in degradation outcome between the two degron cell lines. Given the presence of two isoforms of RPB4, we believe that this inconsistency may be attributed to the C-terminal degron not completely degrading isoforms of RPB4, while the N-terminal degron affects all isoforms.

We analyzed previously published RPB4/7 ChIP-seq data by comparing their signals at the gene body with signal of RPB1. The results show that RPB4/7 is distributed throughout the entire gene body, with RPB7 indicating a relatively stronger association with chromatin in the gene body, even compared with RPB1 (Supplementary Fig. 1i). The pausing index of RPB4 and RPB7 is smaller than that of RPB1 in the ChIP-seq analysis (Supplementary Fig. 1j, k). This suggests dynamic conformational changes within the Pol II complex that may allow RPB7 to be more closely associated with the DNA during transcription elongation in the gene body in cells. Additionally, this observation may imply some non-canonical functions of RPB7, which warrant further investigations.

In contrast, we found that RPB3 depletion did not affect the protein levels of RPB1 (Fig. 1d). As Pol II and Pol III protein structures are preserved[45], we sought to determine whether the corresponding subunit of RPB7 in Pol III (RPC8) is responsible for the stability of RPC1. Surprisingly, RPC8 depletion also led to a decrease in RPC1, similar to the depletion of the RPB4-equivalent Pol III subunit, RPC9 (Fig. 1d). Knocking down RPB7 using shRNA in mammalian cells also resulted in decreased protein levels of RPB1 (Supplementary Fig. 1l). These results demonstrated that RPB7 is required for the stability of RPB1 proteins in mammalian cells, which differs from the role of RPB7 in yeast[40].

If RPB7 stabilizes the RPB1, then RPB7 depletion should lead to the repression of gene expression. We performed total RNA-seq, chromatin-associated RNA-seq (ChAR-seq), and RPB1 ChIP-seq after RPB7 depletion for 3 h with spike-in to test this prediction (Supplementary Data 1–2). The results were consistent with a global repression of gene expression (Fig. 1e–g). As a control, we also showed that RPB7 degradation globally decreased RPB7 ChIP-seq signals (Fig. 1f, g).

We next explored the molecular determinants of RPB7-mediated RPB1 stability. Pol II is known to be assembled in the cytoplasm, be imported into the nucleus, and then functions at chromatin[46,47]. We treated cells with the nuclear export inhibitor leptomycin B (LMB). We found that RPB1 dramatically decreased in the nucleus, while it accumulated in the cytoplasm after treatment with auxin and LMB (Supplementary Fig. 2a), suggesting that RPB1 is destabilized in the nucleus. The increase in cytoplasmic RPB1 is a result of LMB inhibiting the export of Pol II-importing factors, as previously reported[48,49].

The loop region of RPB7 is responsible for RPB1 protein stability

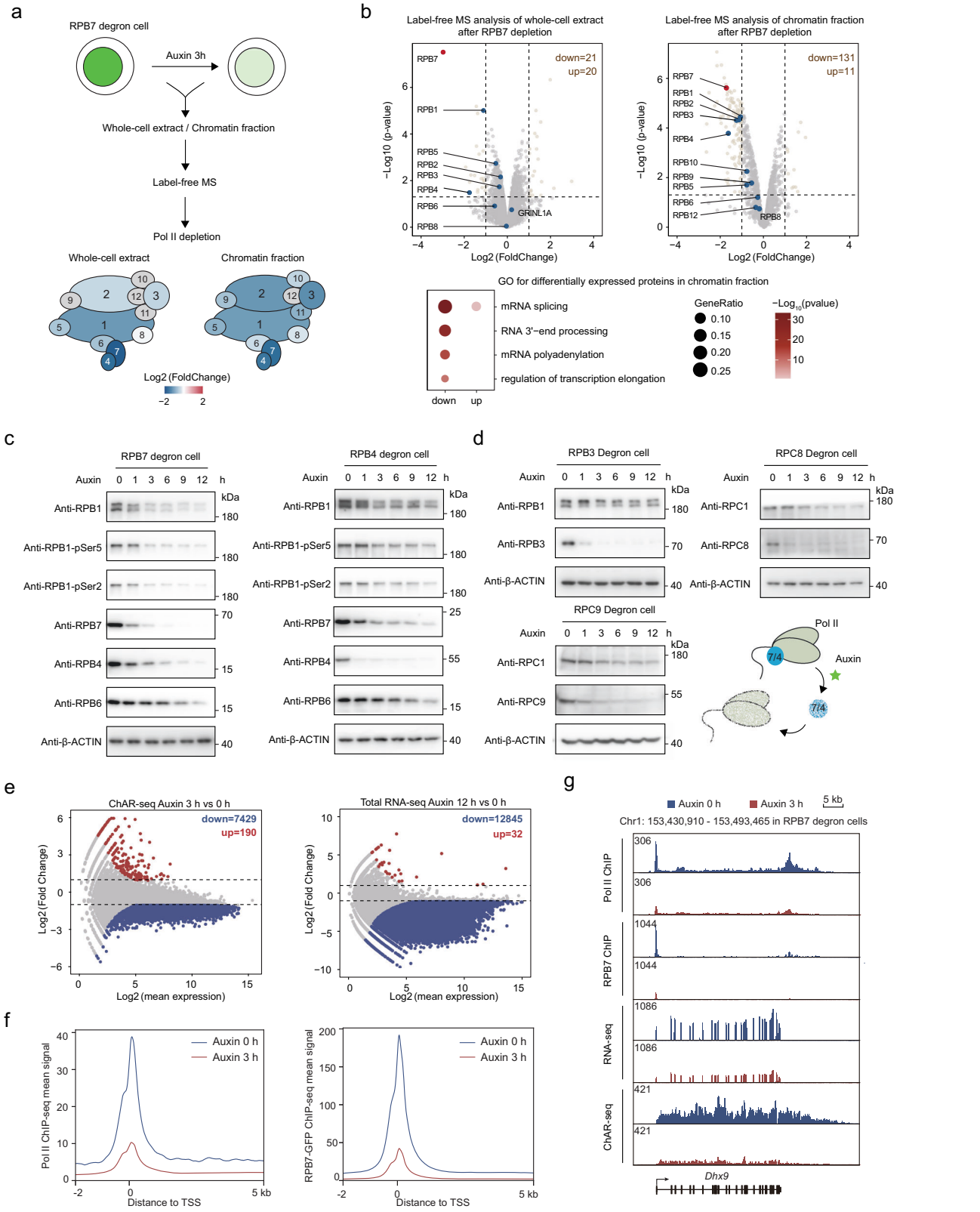
We conducted domain dissection of RPB7 and identified the regions that interact with RPB1 and RPB4. Deletions in the 36–42, 110–116, or 117–138 regions of RPB7 did not significantly affect its role in stabilizing RPB1 or the cell growth rate (Supplementary Fig. 2b, c). Additionally, these regions did not notably impact interactions with RPB1 and RPB4 (Fig. 2a). These findings align with the known structures of RPB1–RPB4–RPB7 and are consistent with these internal deletions not significantly interfering with the predicted structures of RPB7, as indicated by AlphaFold (Supplementary Fig. 2d, e). We next examined the protein structure of RPB7 in Pol II and found that a loop region comprising six amino acids (VIQPGR) in RPB7 directly interacts with RPB1[50]. Deletion of this loop region did not significantly lose its interaction with RPB4 (Fig. 2a) or affect the overall structure of RPB7(Fig. 2b). Wild-type (WT) or loop region-deleted RPB7 was knocked into the *Tigre* locus in RPB7 degron cells. These results showed that wild-type RPB7, but not loop region-deleted RPB7, rescued the protein levels of RPB1 (Fig. 2c, d). RPB7 was recently shown to mediate Pol II dimerization[51,52]. We then wondered whether the dimerization of RPB7 contributed to RPB1 protein stability. The previously described RPB7 dimerization mutant was knocked into RPB7 degron cells[51]. Interestingly, the RPB7 dimerization mutant also could not rescue the protein levels of RPB1 (Fig. 2d). We observed an increase in the levels of exogenously expressed RPB7 when the endogenous RPB7 was degraded. However, upon inhibiting newly translated RPB7 with cycloheximide, we did not observe the increase. This suggests that autoregulation occurs at the translational level (Supplementary Fig. 2f).

We next examined the cellular localization of loop region-deleted RPB7 and found that it was mostly enriched in the cytoplasm (Fig. 2e). Because RPB1 is destabilized in the nucleus, and RPB7 with cytoplasmic localization may not be functional, a loop region-deleted RPB7 fused with a nuclear localization signal was overexpressed in RPB7 degron cells. Fluorescence imaging confirmed its nuclear localization (Fig. 2e, f). However, the nuclear loop region-deleted RPB7 could still not rescue the protein levels of RPB1 (Fig. 2d), indicating that the RPB7 interaction with Pol II, rather than its nuclear localization, is required for the stability of RPB1. The loop region from yeast or the similar region in RPC8 or RPA43, the RPB7 corresponding subunit in Pol III and Pol I[45] (Fig. 2g), respectively, were used to replace the loop region of RPB7. The results showed that the loop region was functionally conserved between mouse and yeast, and the loop region from RPC8 was functionally comparable to that of RPB7 but region from RPA43 was not (Fig. 2g, h).

The destabilization of RPB1 due to RPB7 loss depends on phosphorylation

Previous studies have shown that phosphorylated Pol II is the substrate of ubiquitination and degradation in yeast[53]. CDK9 is a kinase for the Pol II CTD[54]. Therefore, we treated RPB7 degron cells with

https://doi.org/10.1038/s41467-025-57513-2



different concentrations of CDK9 inhibitors or other transcription inhibitors. Western blotting showed that CDK9 degradation by PROTAC (dCDK9) or Pol II CTD phosphorylation inhibition by flavopiridol or DRB stabilized hypophosphorylated RPB1 after RPB7 depletion[55] (Fig. 3a, b). Fluorescence imaging experiments showed that the treatment of DRB or flavopiridol did not affect the nuclear localization of RPB1 upon auxin treatment in RPB7 degron cells (Supplementary Fig. 3a), indicating the rescue of nuclear destabilization of RPB1.

CDK9 inhibition not only reduces CTD phosphorylation but also halts elongation, potentially aiding in Pol II stabilization. To explore this, we employed other transcription inhibitors, such as dBET6[56], which did not prevent RPB1 degradation subsequent to RPB7 depletion (Supplementary Fig. 3b). Additionally, the relative increase in pSer2

Fig. 1 | Rapid disruption of RPB7 triggers the destabilization of RPB1 in mESCs. **a** Schematic of the procedure of quantitative mass spectrometry analysis. Same as ChIP-seq protocol, RPB7 degron cells were first lysed by 0.05% NP-40 buffer, and then by nuclei lysis buffer to get the chromatin fraction. The bottom shows the fold change of each Pol II subunit's protein level in MS data. The unidentified subunits were shown in gray color. **b** Volcano plot showing the protein enrichment changes identified by label-free MS of whole-cell extract (left) and chromatin fraction (right) upon RPB7 degradation. The red dot indicates RPB7 and the blue dots indicate other subunits of Pol II. The differentially enriched proteins were identified by DEP (|log2FC|>1, *p* value < 0.05, two-sided Wald test in DESeq2 software). Bottom shows the GO enrichment analysis of the differentially enriched proteins in MS of chromatin fraction. **c** Western blot analyses of Pol II subunits in whole-cell lysates after time-course auxin treatment in RPB7 (left) and RPB4 (right) degron cells. β-Actin served as the loading control. **d** Same as (**c**) but for RPB3, RPC8 (RPB7-equivalent subunit in Pol III) and RPC9 (RPB4-equivalent subunit in Pol III) degron cells. The diagram on the right shows that the depletion of RPB7 or RPC8 causes the destabilization of RPB1 and RPC1, respectively. **e** MA plot displaying the gene expression changes identified by chromatin-associated RNA (left) and total RNA-seq (right) before and after RPB7 degradation. The red and blue dots represent the up- and downregulated genes identified by DESeq2 (|log2FC|>1, *p*adj < 0.05, two-sided Wald test in DESeq2 software). **f** Meta-gene plots showing the changes in Pol II (RPB1, left) and RPB7 (right) ChIP-seq signals at −2 kb to 5 kb around transcription start sites (TSS) upon RPB7 degradation. **g** Genome browser of Pol II (RPB1) and RPB7 ChIP-seq, RNA-seq, ChAR-seq tracks at the 153,430,910–153,493,465 region on chromosome 1 in RPB7 degron cells after 3 h treatment with (red) or without (blue) auxin. The *y*-axis shows the read counts, normalized to the spike-in. Only sense strand signals are presented for ChAR-seq tracks.

and pSer5 levels was observed when comparing the treatment of DRB with simultaneous RPB7 depletion to treatment with DRB (Supplementary Fig. 3c). To investigate the role of CDK7, we selected two CDK7 inhibitors: THZ1[57] and SY-5609[58]. Both inhibitors were shown to reduce the levels of pSer5 and pSer2 on RPB1, as indicated by our western blot analyses (Supplementary Fig. 3d). CDK7 is shown to preferentially phosphorylate the CTD at Ser5 and Ser7[59,60], and its loss of activity may indirectly contribute to the reduction of pSer2 levels. Initially, we treated RPB7 degron cells with a CDK7 inhibitor or Actinomycin D for 2 h before adding auxin to deplete RPB7. However, Pol II destabilization was still observed (Supplementary Fig. 3e), likely due to the persistence of the hyperphosphorylated state of RPB1 following treatment with these inhibitors. The reason we emphasize CDK9 over CDK7 is that the CDK7 inhibitors we used cannot restore the protein stability of Pol II following RPB7 depletion. We propose that RPB7 degradation may lead to severe defects in Pol II dephosphorylation, and some phosphorylation events may not be inhibited by DRB or THZ1, while we cannot exclude alternative explanations for these results.

The RPB7 ChIP-MS[44] showed a preferential interaction between RPB7 and CTDP1 compared with RPB7 and other phosphatases, namely, RPAP2, PNUTS (PPP1R10)-PP1, SSU72, and PP2AC (PPP2CA) (Fig. 3c). CTDP1 is an ortholog of FCP1. RPB4 was reported to interact with the phosphatases SSU72 and FCP1 to facilitate Pol II dephosphorylation in yeast[40]. We conducted evolutionary tree analyses for FCP1 and found both similarities and differences between CTDP1 in human or mouse, and FCP1 in yeast or *Drosophila* (Supplementary Fig. 3f). We then inserted a TurboID tag into the CTD of RPB1 in RPB7 degron cells (Fig. 3d, e, and Supplementary Fig. 3g), which allowed us to proximally label the dynamically and transiently RPB1 interacting proteins with biotin during the depletion of RPB7. The biotin-labeled proteins were analyzed with mass spectrometry as reported previously[61,62]. The RPB1 TurboID-MS results showed that RPB7 depletion caused a decrease in Pol II subunits interacting with RPB1 (Supplementary Data 3). The GO analyses showed that the differentially interacting proteins were enriched in mRNA processing, negative regulation of phosphorylation, chromatin modifications, and transcription initiation pathways (Supplementary Fig. 3h). CTDP1 was also significantly decreased in RPB1 TurboID-MS after RPB7 depletion (Fig. 3e). CTDP1 and RPB7 can be co-immunoprecipitated (Fig. 3f). The domain dissection analyses revealed that the 154–172 region of RPB7 is necessary for its interaction with CTDP1 (Fig. 3g), while it does not affect its interactions with RPB1 and RPB4. To explore the possibility that RPB4/7 can independently recruit CTDP1, we conducted a ChIP-western blotting assay in RPB1 degron cells[42] using an antibody against RPB4 to assess the interaction between RPB4 and CTDP1 in the absence of RPB1. The results revealed that the interaction between CTDP1 and RPB4 is dependent on the presence of RPB1, suggesting that RPB4/7 interacts with CTDP1 while remaining associated with Pol II (Supplementary Fig. 3i).

The phosphatase activity of CTDP1 contributes to the regulation of RPB1 stability

Then, a degron-GFP tag was inserted into the C-terminus of CTDP1 to construct a CTDP1 degron mESCs. Genotyping and western blotting analyses confirmed homozygous gene editing (Fig. 4a and Supplementary Fig. 3g). CTDP1 degradation also caused the destabilization of RPB1 (Fig. 4a), similar to what we observed after RPB7 depletion. Additionally, the acute degradation of phosphatases CTDSP1, SSU72, or TFIIF subunit RAP74 (previously known to interact with phosphatase CTDP1[63,64]), did not affect the protein abundances of RPB1 (Fig. 4a). The minor effects of these phosphatases on Pol II phosphorylation may be because of functional redundancy[65].

We next treated cells with calyculin A, an inhibitor of protein phosphatases PP1 and PP2A, and with the PP2A inhibitor okadaic acid and found that neither inhibitors had a substantial effect on the total RPB1 protein levels after the depletion of RPB7 (Supplementary Fig. 3j), which is in accordance with the strong interactions between RPB7 and CTDP1 and weaker interactions between RPB7 and the other phosphatases (Fig. 3e). Previous depletion of PNUTS (PPP1R10), a PP1 regulatory subunit, or RPAP2 also did not reveal defects of RPB1 protein stability in various biological contexts[66,67]. Thus, we did not follow up on PP1, PP2A, and RPAP2 in our study.

We next sought to investigate how CTDP1 depletion affects Pol II phosphorylation. We conducted CTDP1 degradation followed by mass spectrometry and cellular fractionation analyses. The degradation of CTDP1 resulted in the destabilization of RPB1, which was further confirmed through whole-cell extract MS analysis (Fig. 4b). Western blotting results showed that CTDP1 depletion led to an increase in hyperphosphorylated Pol II (IIO) in the soluble fraction, a decrease in chromatin fraction (Fig. 4c). We conducted CTDP1 ChIP-seq assays before and after RPB7 depletion, which demonstrated a decreased chromatin binding of CTDP1 following RPB7 degradation (Fig. 4d). The CTDP1 ChIP-seq signals upon RPB7 degradation were comparable to the input and IgG ChIP-seq signals (Fig. 4d), which are considered background signals in ChIP-seq experiments. This distinction becomes clearer when comparing the RPB1 ChIP-seq signals before and after degradation itself (Fig. 4d). We showed that the phosphatase activity and the N-terminus domain of CTDP1 are necessary for maintaining RPB1 protein stability by expressing phosphatase mutants (N-terminus domain deletion, D188E/D190E[68] and D302K[69]) of CTDP1 in CTDP1 degron cells (Fig. 4e, f). We also examined the pSer2/Total and pSer5/Total ratio upon RPB7 degradation for 1 h and noted an increase in pSer2/Total and pSer5/Total ratio in the soluble fraction following RPB7 degradation (Supplementary Fig. 3k).

The destabilization of RPB1 due to RPB7 loss depends on CTD, the linker region of RPB1

The Pol II CTD is phosphorylated during the transcription process and dephosphorylated for Pol II reinitiation[25,70]. Initially, we investigated RPB1 protein levels in both RPB1 wild-type and CTD-deleted RPB1

https://doi.org/10.1038/s41467-025-57513-2

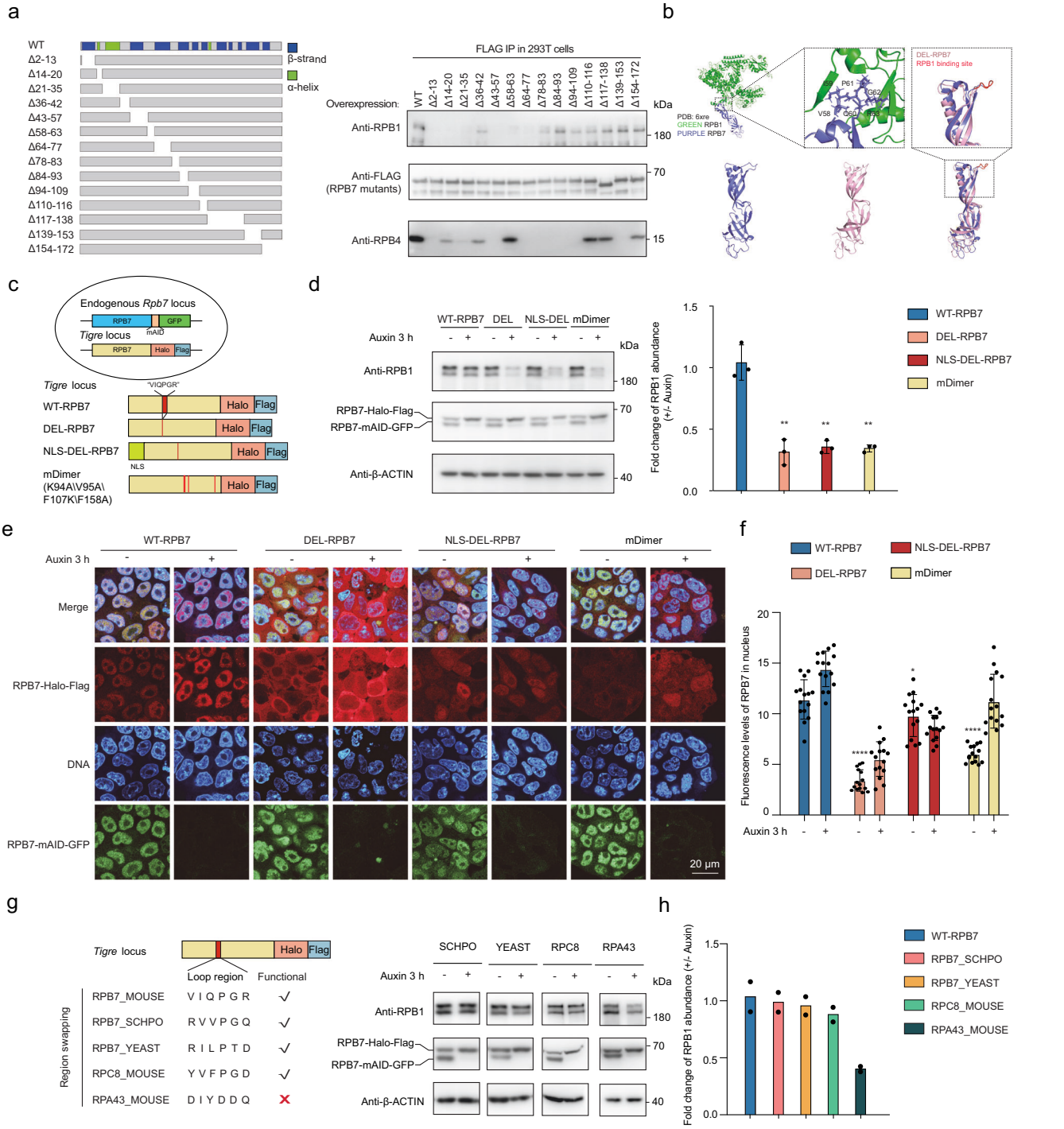


Fig. 2 | RPB7 interaction with RPB1 and its dimerization are associated with the stability of nuclear RPB1. a Schematic of the design of RPB7 mutants, based on the protein secondary structure provided by Uniport. Right: Western blot analyses of anti-FLAG immunoprecipitates collected from lysate of exogenous RPB7-FLAG-expressing 293T cells. FLAG served as the IP fraction loading control. **b** Visualization of the RPB7 structure (blue) in the Pol II complex (PDB: 6xre) and its structural alignment with DEL-RPB7 (pink), predicted by AlphaFold2[93]. **c** Schematic of exogenous expression of wild-type (WT) or mutant RPB7 at the *Tigre* locus in RPB7 degron cells is shown. The loop region was deleted in DEL-RPB7, and the NLS tag "MAPKKKRKVGIHGVPAA" was fused at its N-terminus to generate the NLS-DEL-RPB7 construct. The dimerization mutant (mDimer) was designed according to the sequence alignment of mouse RPB7 and yeast RPB7 mutant previously reported[51]. **d** Western blotting analyses for these four *Tigre* knock-in cells in the presence and absence of auxin treatment. β-Actin served as the loading control. Right: Fold changes in RPB1 protein levels ($n = 3$, biological replicates) are shown as means ± SD. Statistical significance, two-sided unpaired *t*-test; $^{**}p < 0.01$. L-R *p* value: 0.0021, 0.0015, and 0.0012. **e** Representative fluorescence signals of these four *Tigre* knock-in RPB7 (green) degron cells, which were stained with Janelia Fluor 549 HaloTag ligand (red) to visualize the exogenously expressed RPB7 and Hoechst to indicate nuclei (DNA, blue). Scale bar, 20 μm. **f** Nuclear signals of exogenous RPB7 in (**e**) ($n = 15$, biological replicates) are shown as means ± SD. Statistical analyses were performed by compared each RPB7 mutant with wild-type RPB7 under Auxin 0 h condition, two-sided unpaired *t*-test; $^{****}p < 0.0001$; $^{*}p < 0.05$. L-R *p* value: 9.242×10^{-11}, 0.0402 and 1.145×10^{-10}. **g** The loop regions from functionally equivalent subunits of RPB7 (RPC8 and RPA43) or yeast are shown. The right panel shows the western blot analyses in RPB7 mutants-expressing RPB7 degron cells. **h** Fold changes in RPB1 protein levels in (**g**) ($n = 2$, biological replicates) are shown as means.

https://doi.org/10.1038/s41467-025-57513-2

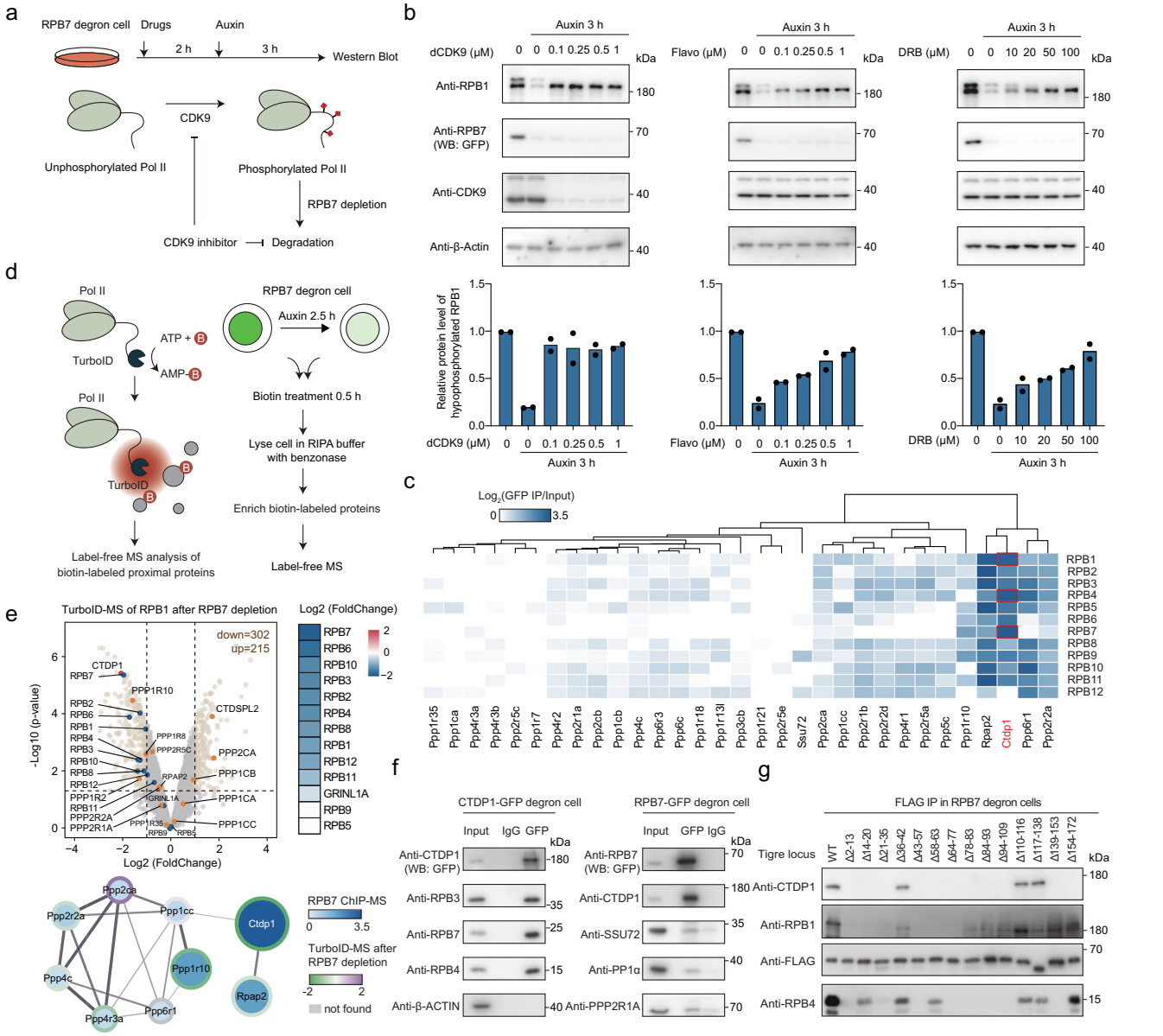


Fig. 3 | CDK9 and CTDP1 are associated with RPB7 in RPB1 stabilization regulation. a Schematic of transcription inhibitor treatment in RPB7 degron cells is shown. **b** Western blotting analyses for RPB7 degron cells, treated with the indicated range of dosages of the THAL-SNS-032 (dCDK9), flavopiridol (Flav) or DRB for 2 h and then with 500 μM auxin for 3 h. β-Actin served as a loading control. The bottom bar graph showing the relative protein level of hypophosphorylated RPB1 as means ($n = 2$, biological replicates). **c** Heatmap showing the enrichment of Pol II related phosphatases in ChIP-MS of the twelve subunits of Pol II. **d** The schematic diagram on the left shows TurboID tag at the C-terminus of RPB1 will label the proximally interacting proteins upon adding the biotin, then the biotin-labeled proteins were enriched by streptavidin beads to detected by label-free MS. The flow diagram on the right shows the TurboID-MS assay we did in RPB7 degron cells. **e** Up: Volcano plot showing the protein enrichment changes identified by TurboID-MS upon RPB7 degradation. The red dot indicates CTDP1, blue dots indicate other subunits of Pol II, and yellow dots indicate subsets of Pol II related phosphatases. The heatmap shows the $\log_2$FC of Pol II subunits. Bottom: A network showing the phosphatases enriched in RPB7 ChIP-MS. The color and size of the circles both indicate the log2(IP/GFP) enrichment score. The border color of the circles indicates the log2FC of these phosphatases in RPB1 TurboID-MS upon RPB7 degradation. The p value was analyzed with two-sided t-test with standard deviation moderated by empirical Bayes method in Limma software. **f** Co-immunoprecipitation (Co-IP) assay. Lysate of CTDP1 or RPB7 degron cells were immunoprecipitated with antibodies against either GFP ("GFP" lane) or rabbit IgG ("IgG" lane). In total, 5% of cell lysate without antibodies was loaded as a control ("Input" lane). Left: Western blot analyses of anti-GFP immunoprecipitates collected from lysate of CTDP1-GFP degron cells. Right: Reverse co-IP of RPB7 in RPB7-GFP degron cells is shown. **g** Co-IP of FLAG in knock-in RPB7-Halo-Flag cells after auxin treatment. Levels of enriched RPB1, RPB4, and CTDP1 were examined.

following RPB7 degradation. We found that even the CTD-deleted RPB1 experienced decreased protein levels, suggesting contributions from other sites to RPB1 stability (Fig. 5a–d). Previous research revealed that the linker region followed by the C-terminal domain (CTD) of RPB1 undergoes phosphorylation and interacts with SPT6. This interaction is regulated by P-TEFb-mediated phosphorylation of the RPB1 linker region[71]. Upon deleting the CTD and linker regions of RPB1, resistance to RPB7 degradation was observed (Fig. 5c, d). Moreover, deletion of the linker region or mutation of its phosphorylation sites still resulted in decreased protein stability of RPB1 subsequent to RPB7 degradation (Fig. 5c, d). Collectively, our findings indicate that the linker region, as well as the CTD of RPB1, contributes to its destabilization when RPB7 is degraded. Similar experiments shown in Fig. 5c were performed with CTDP1 degron cells,

https://doi.org/10.1038/s41467-025-57513-2

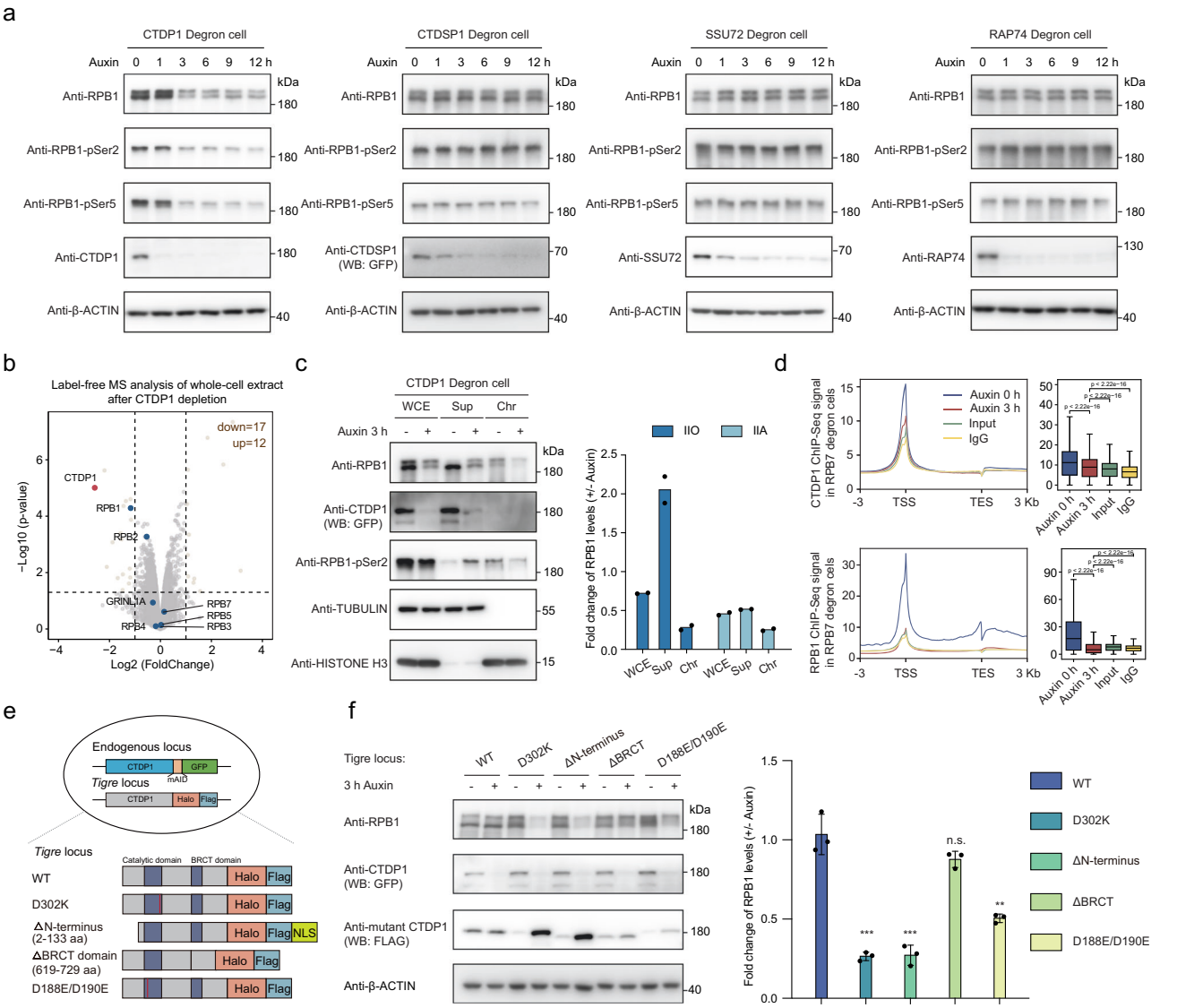


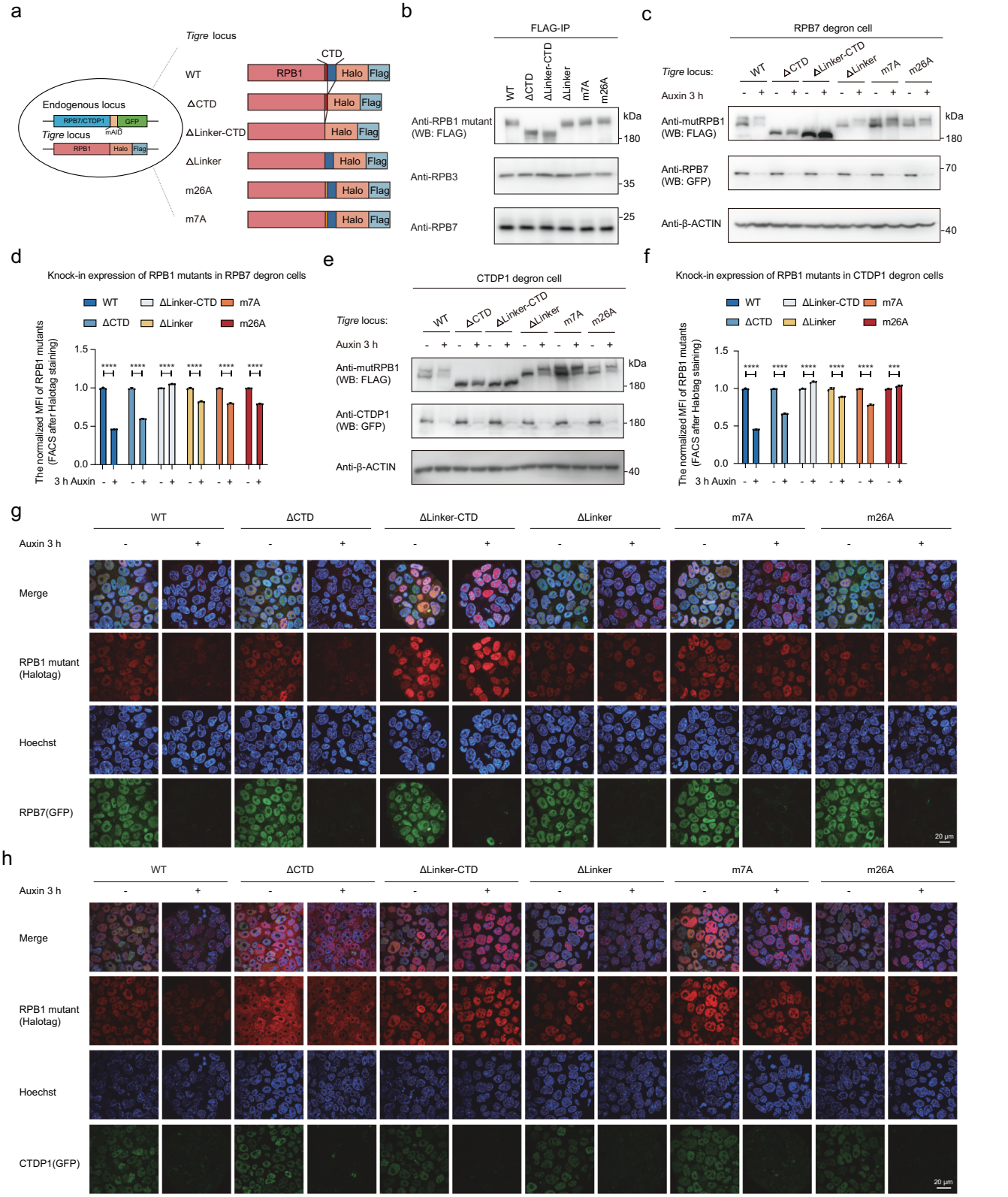
Fig. 4 | Degradation of CTDP1, but not other phosphatases, leads to the destabilization of RPB1. a Western blot analyses of the indicated protein in CTDP1 degron cells, CTDSP1 degron cells, SSU72 degron cells, and RAP74 degron cells after a time-course auxin treatment. **b** Volcano plot showing the protein enrichment changes identified by label-free MS of whole-cell lysates upon CTDP1 degradation. The red dot indicates CTDP1, blue dots indicate subunits of Pol II. The differentially enriched proteins were identified by DEP (|log2FC|> 1, p value < 0.05, analyzed with two-sided t-test with standard deviation moderated by empirical Bayes method in Limma software). **c** Western blot analyses of the indicated proteins in two subcellular fractions of CTDP1 degron cells in the presence and absence of auxin treatment. Right: The fold changes in hypophosphorylated Pol II (IIA) and hyperphosphorylated Pol II (IIO) upon depletion of CTDP1 are shown (n = 2, biological replicates). **d** Meta plots showing CTDP1 (up) and RPB1 (down) ChIP-Seq signal upon RPB7 degradation, as well as an input and IgG ChIP-Seq signal. Right shows the quantification of the meta plots at TSS region (−0.5 to +0.5 kb around TSS). The bounds of box indicate 25% (Q1) to 75% (Q3) of the data. The center of the box indicates 50% of the data. The upper and lower whiskers indicate Q3 + 1.5IQR and Q1 − 1.5IQR, respectively (IQR = Q3 − Q1). Statistics in two biological replicates were identified using two-sided Wilcoxon test. **e** Schematic of exogenous expression of wild-type or four mutated CTDP1 (D302K, Δ2–133, Δ619–729, and D188E/D190E) in CTDP1 degron cells. **f** Western blot analyses in exogenous CTDP1-expressing cells treated with or without 500 μM auxin for 3 h. GFP signal served as the degradation control, FLAG signal served as exogenous CTDP1-expressing control, and β-Actin served as the loading control. Right: Fold changes of total RPB1 signals in the left panel (n = 3, biological replicates) are shown as means ± SD. Statistical significance, two-sided unpaired t-test; ***p < 0.001; **p < 0.01; n.s., not significant; error bars, SD. L-R p value: 0.0005, 0.0007, 0.1174, and 0.0021.

demonstrated that both the CTD and the linker regions contribute to CTDP1-mediated RPB1 protein stability (Fig. 5e, f). We also demonstrated that RPB1 internal deletion mutants maintained interactions with other Pol II subunits and were imported into the nucleus, suggesting their incorporation into the Pol II complex (Fig. 5b, g, h). After the degradation of CTDP1, there was no significant change in the level of the m26A mutant (Fig. 5e, f), indicating that this mutant is intrinsically resistant to destabilization. In contrast, following RPB7 degradation, the level of the m26A mutant decreased (Fig. 5c, d). This suggests that RPB7, as a core component of the Pol II complex, plays a more complex regulatory role in modulating RPB1 protein levels. This finding aligns with previous reports which indicate that the ubiquitylation of RPB1 leads to the removal of RPB4/RPB7 from the Pol II complex in yeast[72].

Both the presence of RPB7 itself and the dephosphorylation activity of CTDP1 are necessary for Pol II reinitiation

Considering that Pol II accumulates at the transcription start sites after dCDK9 treatment, we wondered whether this accumulated Pol II could still associate with gene promoters in the absence of RPB7. The Pol II

https://doi.org/10.1038/s41467-025-57513-2



ChIP-seq signals were still dramatically decreased after treatment with both dCDK9 and auxin in RPB7 degron cells (Fig. 6a, b). Similar experiments were done in CTDP1 degron cells. The results revealed a decrease in Pol II chromatin binding following CTDP1 degradation; however, the effects on Pol II were less pronounced when cells were treated with dCDK9 prior to CTDP1 degradation (Fig. 6c, d). This implies that CDK9 degradation blocks transcription at the promoter-proximal pause, and that paused Pol II is intrinsically resistant to degradation. This is consistent with the RPB1 destabilization seen in SPT5-degron cells, which was prevented by CDK9 inhibition[34]. Previous studies investigating the ubiquitination of Pol II suggested that Pol II is degraded by the E3 ubiquitin ligase CUL3[34,73]. CUL3 knockout rescued the protein levels of RPB1 after RPB7 depletion (Fig. 6e). It appeared that the rescued Pol II was hyperphosphorylated. We then performed

Fig. 5 | CTD with linker region of RPB1 is required for its destabilization. **a** Schematic of exogenous expression of wild-type (WT) and mutated RPB1 are shown. The amino acid from position 1593 to the last amino acid on RPB1 represents "CTD." The amino acid from position 1486 to position 1592 on RPB1 represents the "linker region." Mutant "m26A" represents the RPB1 mutant with all serine and threonine in the linker region mutated to alanine, while "m7A" represents the RPB1 mutant with seven amino acids in the linker region mutated to alanine, which have been reported to have phosphorylation modifications[71]. **b** Western blot analyses of anti-FLAG immunoprecipitates collected from lysate of exogenous RPB1-expressing cells, which revealed their comparable interaction with other Pol II subunits, like RPB7 and RPB3 shown here. Western blot analyses of mutant RPB1 levels in whole-cell lysates after RPB7 (**c**) or CTDP1 (**e**) depletion. GFP served as the degradation control, FLAG represented the protein levels of exogenous RPB1 and β-Actin served as a loading control. The fold changes of exogenous RPB1 fluorescence intensity (median, $n = 3$, biological replicates) upon depletion of RPB7 (**d**) or CTDP1 (**f**) are shown. Exogenous RPB1-expressing cells were labeled by Janelia Fluor 549 HaloTag ligand, then detected by flow cytometer. Statistical significance, two-sided unpaired *t*-test; $^{****}p < 0.0001$; error bars, SD. **d** L-R *p* value: 1.078×10^{-8}, 9.895×10^{-9}, 1.932×10^{-6}, 1.202×10^{-6}, 1.682×10^{-7}, and 2.109×10^{-8}. **f** L-R *p* value: 7.428×10^{-9}, 2.010×10^{-8}, 2.252×10^{-5}, 3.973×10^{-5}, 2.679×10^{-7}, and 0.0006. **g**, **h** Representative fluorescence signals of exogenous RPB1-expressing cells treated with or without auxin for 3 h. GFP channel (green) represented RPB7 (**g**) or CTDP1 (**h**) and served as the degradation control. Cells were stained with Janelia Fluor 549 to visualize the exogenously expressed RPB1 (red) and Hoechst was used to indicate nuclei (DNA, blue). Scale bar, 20 μm.

cellular fractionation analyses after RPB7 degradation in CUL3 knockout cells. The results showed that the level of hyperphosphorylated Pol II (IIO) increased in the soluble fraction and decreased in the chromatin-bound fractions (Fig. 6f). The same analyses as described in Fig. 6a were also performed in CUL3 knockout RPB7 degron cells, and the Pol II ChIP-seq signals were still decreased (Fig. 6g, h), suggesting that RPB7 per se is required for the reinitiation of hypophosphorylated Pol II. The same experiments were performed by degrading CTDP1 in CUL3 knockout cells. The results indicated that Pol II signals remained decreased following CTDP1 degradation and CUL3 knockout restored RPB1 protein levels after CTDP1 depletion. Interestingly, the decrease in Pol II signals following CTDP1 degradation could be mitigated by pretreatment with dCDK9 (Supplementary Fig. 3l, m). These findings suggest that CTDP1-mediated Pol II reinitiation is primarily driven by phosphorylation events of Pol II.

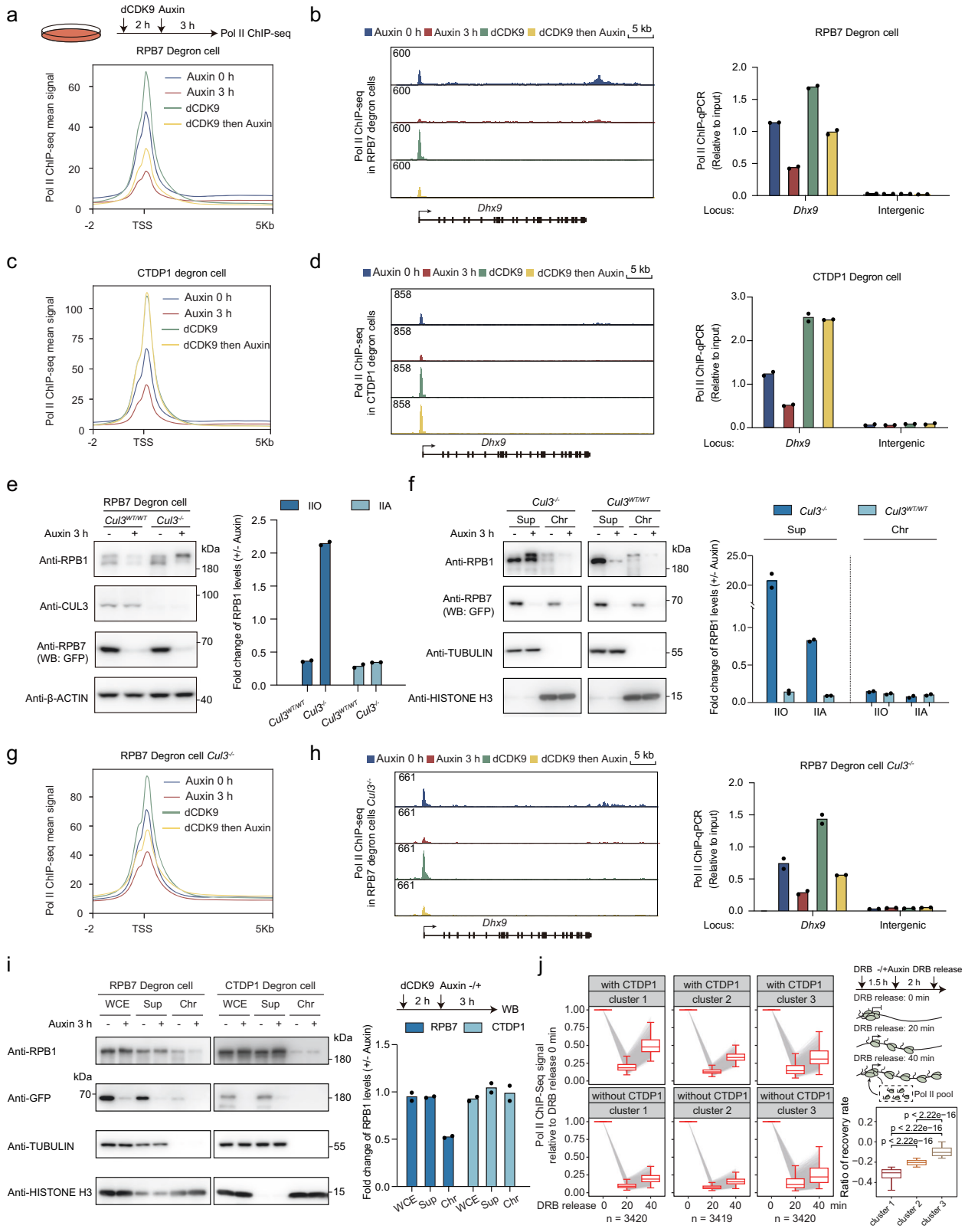
We also performed cellular fractionation analyses upon dCDK9 treatment followed by the depletion of RPB7 or CTDP1. The results showed that hypophosphorylated Pol II was still decreased in chromatin-bound fractions after RPB7 depletion but not after CTDP1 depletion (Fig. 6i). To gain insights into CUL3 knockout-rescued Pol II, we performed ChIP-seq of serine 2-phosphorylated and serine 5-phosphorylated Pol II after RPB7 depletion in CUL3 knockout cells (Supplementary Fig. 4a, b). The results showed that these Pol II ChIP-seq signals decreased across the genes. Interestingly, the smaller the gene size, the larger the decrease in Pol II ChIP-seq signals was, which was not observed for the patterns of RPB7 chromatin binding changes (Supplementary Fig. 4a, b). Then, total RNA-seq and chromatin-associated RNA sequencing were also performed after RPB7 depletion in CUL3 knockout cells, or with CTDP1 degradation cells. The results showed global suppression of gene expression, with chromatin-associated RNAs upregulated for some genes (Supplementary Fig. 4c). These results further suggested that CUL3-rescued Pol II proteins were not functional in transcription. RT-qPCR results showed that the genes dysregulated after RPB7 degradation (such as the downregulated genes, *Dhx9*, *Actb*, and *Gapdh* and the upregulated genes *Lct* and *Klhl33*) exhibited similar changes upon CTDP1 depletion (Supplementary Fig. 4d).

To specifically investigate the roles of CTDP1 in Pol II reinitiation, we next conducted CTDP1 degradation followed by Pol II ChIP-seq analyses during the DRB release process. DRB induces Pol II pausing at the transcription start site, and upon DRB washout, we observed that Pol II enters the elongation stage, with its levels at the transcription start site initially decreasing before increasing. The rising level of Pol II suggests that new Pol II is recruited to the transcription start site, which we interpret as an indication of Pol II reinitiation (Fig. 6j, upper panel). Our quantitative analyses, meta-gene and single gene example show a significant reduction in the rate of increase of Pol II at the transcription start sites following CTDP1 degradation during the DRB release process (Fig. 6j and Supplementary Fig. 4e, f). We performed an additional control experiment using transcription initiation inhibitor triptolide[2,3]. Specifically, we added 5 μM triptolide for 20 min after a 20-min withdrawal of DRB, followed by Pol II ChIP analysis. The Pol II ChIP-qPCR results confirmed our expectations: the binding of Pol II at the promoter-proximal region of the *Dhx9* gene initially decreased after the 20-min DRB withdrawal and then increased 20 min later. In contrast, the binding of Pol II in the triptolide-treated group did not increase (Supplementary Fig. 4g). These results support our conclusion that the recovery of Pol II at the promoter-proximal region is due to reinitiation. Based on these findings, we conclude that CTDP1 affects the efficiency of Pol II reinitiation. We carried out a similar Pol II ChIP experiment to degrade RPB7 during the process of DRB release. The Pol II ChIP-qPCR results indicated that after the withdrawal of DRB, the occupancy of Pol II decreased at the *Dhx9* gene locus in the absence of RPB7 at the 0-min time point of DRB release. Moreover, Pol II occupancy continued to decline at the 20-min and 40-min time points after DRB release (Supplementary Fig. 4h).

RPB7 is linked to co-transcriptional RNA processing

We then sought to investigate the physiological roles of RPB7 and CTDP1 in Pol II stability and reinitiation. RPB7, known to contain RNA-binding domains and to interact with nascent RNA around the RNA exit channel of the Pol II complex[74,75] (Fig. 7a), was found through our Pol II subunits ChIP-MS data to preferentially interact with many RNA 3′ end processing-associated factors (Fig. 7b). To explore the potential roles of RPB7 in RNA processing, which may be connected to their transcriptional functions or its transcription-independent roles, we conducted further investigations: (1) characterization of RPB7-related read-through events: we examined the relationship between terminated Pol II and read-through events associated with RPB7. We plotted ChIP-seq signals for Pol II and pSer2 at transcription termination sites following RPB7 depletion. The results consistently showed a marked decrease in Pol II binding after RPB7 depletion (Supplementary Fig. 4i). We rescaled the pSer2 ChIP-Seq signals in Supplementary Fig. 4i at the transcription termination sites (read-through index). Our analysis revealed that the read-through index calculated from the pSer2 ChIP-Seq data significantly increased following RPB7 degradation, even after normalizing the index by the total RPB1 read-through index (Supplementary Fig. 4j). (2) Investigation of RPB7's role in RNA-binding protein interactions: given that RPB7 localizes to the Pol II holoenzyme's RNA exit channel and associates with various RNA-binding proteins, we specifically analyzed Pol II (RPB1) TurboID-MS data post-RPB7 depletion. We observed a reduction in interactions between Pol II and several RNA-binding proteins involved in 3′ end RNA processing (e.g., CPSF and CSTF factors) (Fig. 7b, c) and RNA splicing-related factors (Fig. 7d, e). These findings suggest that RPB7 might function as a hub to facilitate interactions between RNA processing factors and the Pol II complex. (3) Analysis of CUL3 and CTDP1 in 3′ end RNA processing: we performed RNA-seq on RPB7 degron, CTDP1 degron, RPB7 degron in CUL3 knockout cells, and CTDP1 degron in CUL3 knockout cells. To investigate potential transcription read-through events, we examined the transcripts downstream of the transcription end site (TES) and observed a significant decreasing trend, along with subsets of

https://doi.org/10.1038/s41467-025-57513-2



downstream genes (DoGs) that were upregulated following RPB7 degradation (Fig. 7f). We also found that CTDP1 degradation led to similar upregulation of these DoGs, which became less pronounced in CUL3 knockout cells (Fig. 7g). These results indicate that RPB7, CTDP1, and CUL3 might be interconnected in 3′ end RNA processing, potentially functioning together with Pol II as a hub for RNA processing factors while possibly also involving other unexplored mechanisms.

(4) Analysis of intron retention events: we also analyzed intron retention events using the aforementioned RNA-seq data, which indicates a defect of RNA splicing (Fig. 7h, i). The results showed that RPB7 degradation significantly impaired RNA splicing, while CUL3 knockout mitigated the effects of RPB7 degradation on splicing (Fig. 7h). Notably, the splicing events affected by RPB7 and CTDP1 exhibited limited overlap (Fig. 7i), suggesting that RPB7 and CTDP1 may have distinct

Fig. 6 | RPB7 permits Pol II reinitiation by stabilizing Pol II promoter binding. Meta-gene plots showing the mean Pol II ChIP-seq signal upon untreated (blue), auxin (red), dCDK9 (green), and auxin then dCDK9 (yellow) in *Cul3*$^{WT/WT}$ (**a**), *Cul3*$^{-/-}$ (**g**) RPB7 degron cells and CTDP1 degron cells (**c**) at −2 kb to 5 kb around TSS. **b**, **d**, **h** Left: Genome browser showing RPB1 ChIP-seq tracks in (**a**), (**c**), (**g**) at the same region shown in Fig. 1f. The *y*-axis shows the read counts normalized to the spike-in. Right: The Pol II ChIP enrichment relative to input (2.5%) at the *Dhx9* gene promoter was shown as bar graphs (means, $n = 2$, biological replicates). **e** Western blot analyses of whole-cell lysates of *Cul3*$^{WT/WT}$ or *Cul3*$^{-/-}$ RPB7 degron cells. β-Actin served as a loading control. The fold changes in hypophosphorylated Pol II (IIA) and hyperphosphorylated Pol II (IIO) are shown as means in the right panel ($n = 2$, biological replicates). **f** Western blot analyses of two subcellular fractions in *Cul3*$^{WT/WT}$ and *Cul3*$^{-/-}$ RPB7 degron mESCs with or without auxin treatment. Right: Bar graph (means) showing the fold change in IIA and IIO ($n = 2$, biological replicates). **i** Western blot analyses of three fractions of dCDK9-pretreated RPB7/CTDP1 degron cells with or without auxin. Right: Bar graph (means) showing the fold change in RPB1 in the upper panel ($n = 2$, biological replicates). **j** Left: Boxplots same as Fig. 4d, showing Pol II ChIP-Seq signal at TSS region in three clusters during DRB release with or without CTDP1. Right top: A schematic diagram showing the reinitiation of Pol II upon DRB release. At 20 min most of the Pol II at TSS was released into gene body, at 40 min Pol II TSS signal recovered due to the reinitiation of Pol II in the pool. Right down: Boxplots showing the change of recovery rate of the three clusters. Statistics, two biological replicates, two-sided Wilcoxon test.

impacts on RNA splicing, with CTDP1 potentially targeting substrates beyond the Pol II CTD. While these preliminary results show promising trends, the limited evidence prevents us from drawing definitive conclusions for RPB7's direct role in RNA processing as it stabilizes Pol II. Future studies splitting the two functions of RPB7 will be helpful.

Discussion

The mechanism that controls the stability of Pol II during termination to reinitiation has been unclear in mammalian cells. The central theme of our work posits that RPB7 and CTDP1 function to stabilize Pol II and facilitate its reinitiation. When there are defects in either RPB7 or CTDP1, hyperphosphorylated Pol II accumulates in the nucleoplasm, eventually leading to its degradation by CUL3. Furthermore, we propose that RPB7 and CTDP1 might also serve as a hub for RNA processing factors associated with Pol II, likely safeguarding co-transcriptional RNA processing events (Fig. 7j). Mechanistically, RPB7 integrates into the Pol II complex via its loop region, subsequently recruiting CTDP1, which dephosphorylates the hyperphosphorylated CTD and linker region of Pol II. RPB7 then facilitates the reinitiation of hypophosphorylated Pol II and elongates into gene body with Pol II. Regulation of Pol II termination and reinitiation is linked to the protein stability of RPB1. Hyperphosphorylated RPB1, once released from chromatin, may be susceptible to degradation by CUL3, but not for hypophosphorylated Pol II.

Previous studies have established that the phosphorylation states of Pol II are crucial for linking various co-transcriptional RNA processing events such as RNA splicing, cleavage, and polyadenylation[6,51,76–78]. We propose that RPB7, in collaboration with CTDP1, plays a significant role in the termination and reinitiation cycles of Pol II, potentially serving as a regulatory point for the completion of these co-transcriptional RNA processing activities. Once these events are completed, CTDP1 is recruited to RPB7 to dephosphorylate Pol II, thereby preparing it for a subsequent transcription cycle and simultaneously decoupling it from the previous co-transcriptional RNA processing events. This mechanism could ensure that Pol II, once reinitiated, is dephosphorylated and disengaged from any prior co-transcriptional RNA processing activities before commencing a new transcription cycle. The intricate details of these processes merit further exploration in future studies.

We found that the primary trend in the transcript downstream of the TES is a decrease after depletion of RPB7, which results from the destabilization of Pol II. It may be a little bit confusing to observe decreased Pol II signals alongside increased subsets of transcript downstream of the TES. To clarify this observation, we hypothesize that the increase in read-through transcripts results from reduced recruitment of 3′ end processing factors at TES regions, leading to insufficient RNA processing. Consequently, while the number of Pol II molecules may have decreased, the levels of RNAs exhibiting read-through (unprocessed) signals have increased for specific gene subsets. These transcripts are either less effectively degraded by nucleases or become more stable due to various indirect mechanisms. It may be useful to note that the read-through signals were observed using the RNA-seq dataset generated after 12 h of RPB7 degradation, whereas the Pol II ChIP-seq experiment was conducted after only 3 h of RPB7 degradation. The destabilization of Pol II may also lead to various feedback mechanisms or indirect dysregulations of RNA processing or RNA stability.

We found that CDK9 inhibition can rescue the protein levels of Pol II but not Pol II reinitiation after RPB7 depletion, suggesting that RPB7 per se is required for Pol II reinitiation. CUL3 knockout caused an increase in the level of phosphorylated Pol II in the soluble fractions and a decrease in the level of chromatin-bound Pol II, indicating that Pol II maybe still hyperphosphorylated after termination and release from chromatin. Phosphorylated Pol II in the nucleoplasm is prone to degradation by CUL3. Under normal conditions, RPB7 and CTDP1 dephosphorylate the Pol II to prevent its degradation. A previous study has shown that ubiquitin-ligase Asr1 ejects RPB4/RPB7 from the Pol II complex in vitro[72], indicating that RPB4/RPB7 per se could block the degradation of Pol II. RPB7 and CDK9 depletion rescued the Pol II protein level, indicating that the phosphorylation plays a role in Pol II stability control in mammalian cells, although we cannot exclude the possibility that RPB4/RPB7 themselves block the ubiquitylation sites in the Pol II complex in cells. RPB7 is known to mediate Pol II dimerization[52]. We propose that Pol II dimerization and CTD-mediated Pol II clustering via intrinsically disordered regions (IDRs) together increase the local concentrations of Pol II adjacent to gene promoters[47], which may enhance Pol II reinitiation efficiency.

When FCP1 was knocked down in yeast or *Drosophila*, it resulted in an increase in the protein level of hyperphosphorylated Pol II but a decrease in the chromatin-bound fraction[25,40], and FCP1 dephosphorylates the Pol II CTD to increase transcription in vitro[79]. In mammalian cells, depletion of CTDP1 also led to decreased chromatin-bound Pol II and increased pSer2 in the nucleoplasm. However, unlike FCP1 in *Drosophila* or yeast, CTDP1 depletion resulted in the destabilization of RPB1 in mammalian cells, potentially due to the involvement of distinct E3 ligases for RPB1. Although CTDP1 is confirmed to be a phosphatase, its molecular functions in transcription remain unclear in mammalian cells. CTDP1 is highly enriched in RPB7 IP preparations, and its depletion parallels the observation of RPB7 depletion, whereas the other phosphatases that we investigated were not. We also detected a small fraction of Ssu72 and PP1α in RPB7 IP preparations, which supports the findings of previous studies[40]. However, the predominant role of RPB7 and CTDP1 is in Pol II stability and reinitiation, which is distinct from what was found in previous studies[25,40]. Nevertheless, obtaining direct live-cell imaging evidence of RPB7 and CTDP1's involvement in Pol II stability and reinitiation presents a challenge. Furthermore, the substrate specificity of CTDP1 function in chromatin and soluble nucleoplasm remains unclear. Additionally, it is not yet understood how CUL3 distinguishes between hyperphosphorylated and hypophosphorylated Pol II, nor why perturbation of FCP1 in yeast, its homolog CTDP1 and other Pol II CTD phosphatases in mammalian cells have different effects on the protein stability of RPB1. In

https://doi.org/10.1038/s41467-025-57513-2

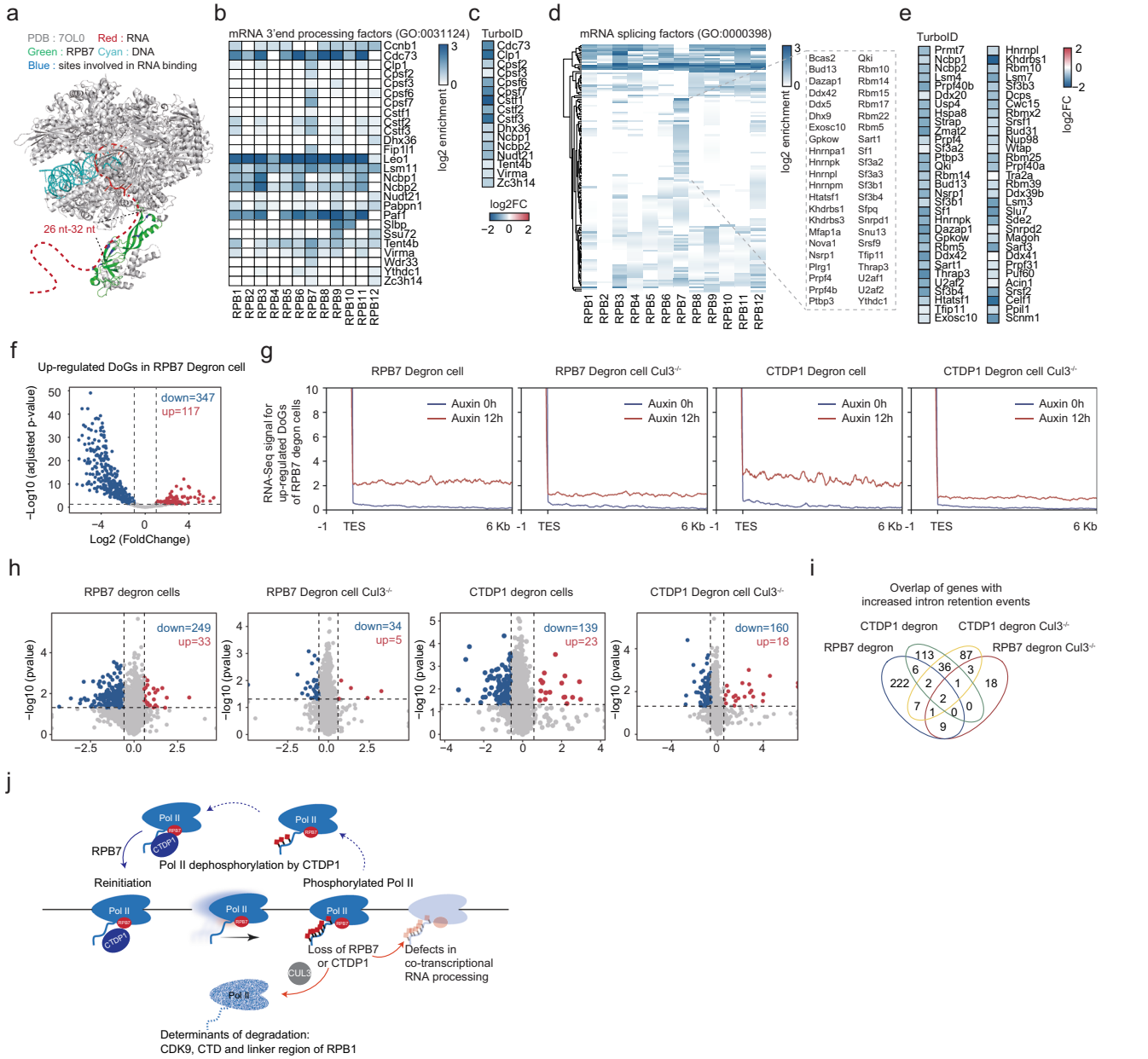


Fig. 7 | RPB7 is linked to co-transcriptional RNA processing. a Visualization of the RNA (red) in the elongating Pol II complex (PDB: 7OL0). The unresolved region in the 5′ end of RNA is indicated by dotted lines. The sites in RPB7 contributing to its RNA-binding ability are shown in blue as reported previously[74]. The region from 26 nt (3′) to 32 nt (3′) in nascent RNA is considered to interact with RPB7 as RPB7 could cross-link to it[75]. **b**–**e** Heatmap showing the enrichment in the ChIP-MS of the 12 subunits of Pol II[44] and $\log_2$ fold change of TurboID-MS after RPB7 degradation of the mRNA 3′ end processing factors (**a**, **b** GO:0031124) and splicing factors (**c**, **d** GO:0000398). **f** A Volcano plot identifying the upregulated DoGs (downstream of gene-containing transcripts, $\log_2$FC > 1, analyzed with two-sided Wald test in DESeq2 software) after RPB7 degradation. **g** RNA-Seq signal for the upregulated DoGs upon RPB7 degradation, *Cul3*$^{-/-}$ and RPB7 degradation, CTDP1 degradation and *Cul3*$^{-/-}$ and CTDP1 degradation. The RNA-Seq signals under untreated conditions are quantile normalized to obtain comparable signals. The RNA-Seq signal after IAA treatment is rescaled according to the scaling ratio of the untreated group. **h** Volcano plots showing the genes with increased intron retention events upon RPB7 degradation, *Cul3*$^{-/-}$ and RPB7 degradation, CTDP1 degradation and *Cul3*$^{-/-}$ and CTDP1 degradation. Statistics were identified using two-sided *t*-test. Genes were identified as upregulated with p value < 0.05 and $\log_2$FC (x-axis)< −1. **i** Venn diagram showing the overlap of genes with increased intron retention events upon RPB7 degradation, *Cul3*$^{-/-}$ and RPB7 degradation, CTDP1 degradation and *Cul3*$^{-/-}$ and CTDP1 degradation. **j** Model of the roles of RPB7 and CTDP1 in regulating Pol II stability as determined in this study. (Upper) After being phosphorylated by CDK9, Pol II needs to be dephosphorylated by CTDP1. Then, RPB7 further facilitates the reinitiation of hypophosphorylated Pol II at gene promoters. (Bottom) The degradation of RPB7 or CTDP1 leads to the defects in co-transcriptional RNA processing and the accumulation of hyperphosphorylated RPB1 in the soluble fraction; then, Pol II is degraded by the CUL3-dependent pathway.

addition, CTDP1 is reported to be associated with various diseases, such as congenital cataracts, facial dysmorphism, and neuropathy (CCFDN)[80–82]. Thus, further investigation of CTDP1 in these diseases would drive our understanding of the biological functions of CTDP1-mediated Pol II stability and reinitiation.

Methods

Cell culture

Degron system was performed with V6.5 mouse embryonic stem cell line. The cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with 15% fetal bovine serum (FBS), also supplemented with

1 × nucleosides, 1 × nonessential amino acids, 1 × penicillin/ streptomycin, 1 mM Sodium Pyruvate, 0.1 mM β-mercaptoethanol, 3 μM CHIR99021 and 1 μM PD0325901, 1000 U/ml mouse leukemia inhibiting factor (mLIF). All the cells were cultured in 0.2% gelatin-coated plates at 37 °C and 5% CO_2 under humid environment. For cell passage, Trypsin-EDTA (0.25%) was used and then quenched directly with culture medium. For *Drosophila* S2 cells, a gift from Jian Zhu Lab in Peking University, Schneider's *Drosophila* Medium with 10% FBS was used, and cells were cultured at 28 °C. HEK293T cells were purchased from ATCC (Cat#CRL-3216) and cultured in DMEM containing 10% FBS and 1 × penicillin/streptomycin. All cells were tested negatively for mycoplasma contamination and then used for further experiments.

Cell lines generation

There are eight degron cell lines (RPB7, RPB4, RPC8, RPC9, CTDP1, RAP74, CTDSP1, and SSU72) generated in this paper as described before[43]. The mAID-GFP tag is fused to each gene's C-terminus, except for RPB4, which is to its N-terminus. To generate donors for gene editing, either side of homology arms (±800 bp from cleavage site) amplified from the genomic DNA was cloned into pGEM-TEasy vector using Gibson assembly. Then mAID-GFP tag was inserted between two homology arms. For sgRNA vector, BbsI-digested CRISPR/Cas9 vectors were ligated with annealed sgRNA oligos. Equal amounts of Cas9-mCherry vector with sgRNA and donor vectors were encapsulated using FuGENE HD transfection reagent according to the manufacturer's instructions and then co-transfected into the parental cell lines Tet-ON-hPGK-OsTIR1 (V6.5) mESCs which were generated previously in our lab[43]. After 24 h, cells were seeded into 10 cm^2 dishes, then fresh medium with 500 μg/ml geneticin was re-added every day to select the drug-resistant clones. The residual cells were collected and single GFP-positive cell was seeded into 96-well plates for cloning. Then the clones were further examined by genotyping to screen homozygous genome-edited clones. Further western blot validation for the effect of auxin-induced degradation was done. For TurboID assay, RPB1 was tagged with TurboID at its C-terminus in RPB7 degron cells. For *Tigre* locus knock-in editing, Cas9-GFP vector with sgRNA targeting to *Tigre* locus and donors containing CMV enhancer/chicken beta actin (CAG) promoter-RPB1\RPB7\CTDP1-Halo-Flag-PGK promoter-HygR cassette were co-transfected into RPB7 or CTDP1 degron cell by using FuGENE HD. For the CTDP1 phosphatase activity-dead mutant, the promoter was replaced with TRE promoter. Different to degron cell generation, clones were selected with Hygromycin B and mCherry-positive cells were collected after labeled with Janelia Fluor 549 HaloTag ligand. Then the clones were further examined by western blot analysis for detecting the expression of Flag-tagged RPB1 or RPB7. For *Cul3* knockout editing, RPB7/CTDP1 degron cells were co-transfected with two Cas9-GFP/sgRNA vectors targeting two sites of *Cul3*, respectively, (exons 3 and 16). After 24 h, GFP-positive cells were sorted into 96-well plates for cloning. Then the CUL3 knockout cells were further validated by western blot for CUL3 protein level. All primer sequences are provided in Supplementary Data 4. The reagents used in this study are listed in Supplementary Data 5.

Cell treatments

To deplete the specific target proteins, degron cells were pretreated with 2 μg/ml doxycycline (Dox) for 12 h to induce the OsTIR1 expression and then with 500 μM Auxin at the indicated time points. For Pol II transcription inhibition, CDK9 degrader THAL-SNS-032, DRB, dBET6, Flavopiridol, SY-5609, THZ1, Actinomycin D, Calyculin A, Okadaic Acid were added at indicated concentrations, respectively, followed by treatment with 500 μM auxin for 3 h. For the Pol II nuclear translocation inhibition, 15 nM LMB was used to pretreat the cells for 3 h. For protein synthesis inhibition, 100 μg/ml cycloheximide was used to pretreat the cells for 1 h.

Subcellular fractionation

To specifically examine the change of protein levels, two subcellular fractions were isolated according to the ChIP-seq protocol. Briefly, 10 million cells in three biological replicates were gently scraped from the plate, transferred to 1.5 ml Eppendorf tube, and collected by centrifugation (300 × *g*, 4 °C for 3 min). Cell pellets were rinsed twice with 1 ml PBS/1 mM EDTA, and then resuspended gently with 0.5 ml glycerol buffer (20 mM Tris-HCl (pH 8.0), 75 mM NaCl, 0.5 mM EDTA, 0.85 mM DTT, 50% glycerol (vol/vol)), further with 0.5 ml nuclei lysis buffer (10 mM HEPES (pH 7.6), 1 mM DTT, 7.5 mM $MgCl_2$, 0.2 mM EDTA, 0.3 M NaCl, 1 M urea, 1% Nonidet P-40) for 2 min incubation on ice. After centrifugation at 13,000 × *g* for 2 min at 4 °C, the supernatant representing the soluble fraction was collected as the fraction "Sup." Then the chromatin pellet was washed with 1 ml PBS/1 mM EDTA, sonicated in 0.5 ml PBS/1 mM EDTA using the Qsonica system (Q800R3 DNA Shearing Sonicator) with 30% amplitude, 30 s ON, 30 s OFF for 4 cycles. The homogenized pellet was collected as the fraction "Chr."

Western blotting

For western blot analyses, samples in at least two biological replicates were boiled in 100 °C with 2 × SDS Loading buffer (125 mM Tris-HCl (pH 6.8), 4% SDS (wt/vol), 20% glycerol (vol/vol) 10% β-mercaptoethanol, 0.01% bromophenol blue (wt/vol)) for 10 min and analyzed by SDS–PAGE. After loading samples and molecular weight marker, run the gel for 30 min at 80 V, then increased voltage to 120 V until the dye front reached the bottom of the gel. Then, the proteins in gel were electrophoretically transferred onto PVDF membrane under 200 mA for 4 h in ice bath. Membranes were blocked with 5% (w/v) skim milk in TBST (1‰ Tween-20 in TBS) for 0.5 h at room temperature and then incubated with the antibodies of target proteins overnight at 4 °C. After being washed three times with TBST, the corresponding secondary antibody diluted 1:5000 in TBST with 5% skim milk was added for 1 h incubation at room temperature, then washed three times with TBST again. After reacting with Enhanced/Super ECL Kit, the lighted bands were captured by Amersham™ Imager 600 system (GE Healthcare Life Sciences). Pictures were further analyzed by using Adobe Photoshop.

Imaging

Immunofluorescence (IF) assays for RPB1 localization were performed. Briefly, pretreated cells in two biological replicates were seeded on glass coverslips to grow for 3 h with or without auxin, then fixed by 4% paraformaldehyde in PBS for 10 min at room temperature. After being washed three times with PBS, cells were permeabilized with 0.2% Triton X-100 in PBS for 10 min at room temperature and washed three times with PBS again. Followed by 30 min incubation with 4% bovine serum albumin (BSA) in PBS, cells were incubated with antibody against RPB1 (1:1000 dilution in 4% BSA) overnight at 4 °C. After being washed three times with PBST (0.1%Tween-20 in PBS), cells were stained with secondary antibodies Alexa Fluor 594 Goat anti-Rabbit IgG (1:1000 dilution) for 1 h at room temperature and protected from exposure to light. Then nuclei were stained with DAPI in PBS (2 μg/ml) for 5 min at room temperature. After being washed three times with PBST again, coverslips were mounted on slides using VECTASHIELD Antifade Mounting Medium and sealed with colorless nail polish. For HaloTag imaging, the cells were seeded on glass-bottom dishes to grow with DOX for 12 h, followed by 500 μM auxin treatment for 2.5 h. Cells were further labeled with Janelia Fluor 549 HaloTag ligand and Hoechst for 30 min at a final concentration of 40 nM and 2 μg/ml, respectively, then fixed by 4% paraformaldehyde in PBS for 10 min at room temperature. Images were acquired by Nikon A1R microscope with 100× or 60×-magnification oil immersion objective lens. All the fluorescence images were post-processed and exported using NIS Elements Viewer software. Treatment and control groups were exported under the same optical parameters for later comparative

analysis. For nuclear WT or mutated RPB1/RPB7 quantification, the Hoechst stain was used to define nuclear fraction. Briefly, the exported images acquired under the same parameters were loaded in ImageJ. After being transformed to "8-bit" format, regions in Hoechst stain bigger than 1000 pixels were selected as ROI by setting "Threshold" to "Auto" in "Adjust" menu, "Fill Holes" and "Watershed" in the "binary" menu. Then the densities of corresponding RPB1\RPB7-HaloTag signals in each ROI were measured. The quantification of the cytoplasmic region was obtained by subtracting the nuclear region signal from the whole-cell region signal.

Flow cytometry

HaloTag-expressing cells with or without Auxin treatment were labeled with 40 nM Janelia Fluor 549 HaloTag ligand for 30 min. Trypsin-EDTA (0.25%) was used to acquire single-cell suspensions. The cell suspensions were filtered through a 70 μm nylon mesh to remove any undigested cells and debris. The samples were then loaded onto a BD LSRFortessa flow cytometer (BD FACSDiva™ Software v8.0). Gating strategies can be found in Supplementary Fig. 5. FlowJo (v10.8.1) and GraphPad Prism (v9.5.0) were used for data analysis.

Cell proliferation assay

After counting with a cell counter for knock-in WT or truncated RPB7 cells (15 cell lines in total), 3000 cells in three biological replicates are seeded to 96-well plates. After growing with DOX for 12 h, experimental groups were treated with 500 μM Auxin for indicated time points. Control groups were added with cell culture medium in the same volume. At indicated time points, CCK-8 reagent was added to each well in a volume ratio of 1:10 for a 2 h reaction at 37 °C. Then the absorbance at 450 nm wavelength of light was measured for each well by BioTek Cytation 5. Wells with cell culture medium but without cells were used as blank controls. For the wild-type, RPB4-N degron, and RPB4-C degron cell lines, experimental groups were treated with 500 μM auxin at specified time intervals following a 12-h induction with DOX. Control groups received an equivalent volume of cell culture medium. The application and assessment of the CCK-8 assay were conducted in accordance with the above procedures.

ChIP-seq and ChIP-qPCR

ChIP-seq was performed as described before[43]. Two biological replicates were analyzed for each sample. Briefly, 20 million cells mixed with 10% HEK293T cells were collected by trypsinization and then crosslinked by 1% formaldehyde (wt/vol) for 10 min at room temperature. Glycine was then added to 0.125 M to quench formaldehyde for 5 min at room temperature. Followed by being lysed with 0.5 ml of ice-cold NP-40 lysis buffer (10 mM Tris-HCl (pH 7.5), 150 mM NaCl and 0.05 % Nonidet P-40) on ice for 5 min, cells were transferred onto the top of 1.25 ml sucrose cushion (24% sucrose (wt/vol) in NP-40 lysis buffer), then centrifugated at 13,000 × *g* for 10 min at 4 °C. After discarding the supernatant, the nuclei pellets were washed once with 1 ml PBS/1 mM EDTA. Glycerol buffer (20 mM Tris-HCl (pH 8.0), 75 mM NaCl, 0.5 mM EDTA, 0.85 mM DTT, 50% glycerol (vol/vol)) was then used to resuspend the nuclei pellet, and equal volume of nuclei lysis buffer (10 mM HEPES (pH 7.6), 1 mM DTT, 7.5 mM $MgCl_2$, 0.2 mM EDTA, 0.3 M NaCl, 1 M urea, 1 % Nonidet P-40) was added for 2 min incubation on ice. The chromatin pellet was collected by centrifugation at 13,000 × *g* for 2 min. After being washed twice with PBS/1 mM EDTA, the pellet was resuspended with 1 ml sonication buffer (20 mM Tris-HCl (pH 8.0), 5 mM $CaCl_2$,150 mM NaCl, 2 mM EDTA (pH 8.0), 0.1% SDS, 1% Triton X-100), then incubated with 1000 U MNase for 15 min at 37 °C with 700 rpm shaking. In total, 20 μl 0.5 M EDTA and 40 μl 0.5 M EGTA were added to stop the reaction. The pellet was divided into 300 μl/1.5 ml Eppendorf tube and sonicated using the Qsonica system with 70% amplitude, 30 s ON, 60 s OFF for 20 cycles. After being centrifugated twice at 13,000 × *g* for 10 min at 4 °C, the supernatant was transferred into a new 1.5 ml Eppendorf tube, and 20 μl supernatant was collected as input. Then 1 μl GFP abs, 2 μl RPB1-NTD abs, 3 μl anti-RPB1-pSer2, and 3 μl anti-RPB1-pSer5 were added, respectively, for each ChIP assay. After being incubated overnight at 4 °C with rotation and centrifugated at 13,000 × *g* for 10 min at 4 °C at the next day, the lysate was transferred to a new tube with 30 μl pre-washed Protein G magnetic beads and incubated for another 3 h at 4 °C with rotation. After being washed once with sonication buffer, once with high-salt wash buffer (20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), once with LiCl wash buffer (10 mM Tris-HCl (pH 8.0), 250 mM LiCl, 1 mM EDTA, 1% NP-40) and three times with TE buffer (1 mM EDTA, 10 mM Tris-HCl (pH 8.0)), the beads were then incubated twice with 150 μl elution buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% SDS) at 65 °C for 30 min with 900 rpm shaking. The supernatant was collected as the eluted ChIP DNA. Together with input, the supernatant was added with 5 μl Proteinase K and incubated at 65 °C overnight. Protease K was inactivated at 80 °C for 20 min. DNA was purified through the DNA Extraction Reagent, then used for library preparation with NEBNext Ultra II DNA library prep kit according to the manufacturer's instruction and the library was sent to Novogene for Novaseq PE150 sequencing. For ChIP-qPCR, the DNA before library preparation was used as a template and run PCR with SYBR qPCR Master Mix on Bio-Rad CFX Connect™ Real-Time PCR Detection System. For DRB release/total RPB1 ChIP-seq, cells were pretreated with 100 μM DRB for 3.5 h, then rinsed by PBS twice to wash out DRB and incubated in 37 °C fresh medium for another 0, 20, and 40 min, respectively, followed by ChIP-seq protocols immediately. For CTDP1 or RPB7 depletion groups, 500 μM Auxin was added after cells had been treated with DRB for 1.5 h. Auxin was also added in the fresh medium.

RNA-seq and RT-qPCR

RPB7/CTDP1 Degron cells with or without CUL3 expression were treated with 2 μg/ml Dox for 12 h, then 500 μM Auxin was added into the medium for 12 h. Then 0.5 ml TRIzol reagent was used to lyse each sample (one well on a 6-well plate) according to the manufacturer's instruction, with 10% *Drosophila* S2 cells for spike-in normalization. Two biological replicates were analyzed for each sample. Total RNA was then delivered to Novogene (Beijing) for poly(A) RNA library construction and sequenced by Novaseq 6000. Two biological replicates were performed. To study the functions of site-specific mutations of RPB7, 500 μM Auxin treatment was extended to 24 h. For RT-qPCR assay, total RNA was extracted by using FastPure Cell/Tissue Total RNA Isolation Kit, and then equal amount of RNA was converted to cDNA by using HiScript III RT SuperMix for qPCR. Then, Quantitative RT-PCR was conducted same as ChIP-qPCR described above.

Chromatin-associated RNA-seq (ChAR-seq)

Chromatin-associated RNA was isolated according to the protocol published previously[83]. Two biological replicates were analyzed for each sample. Briefly, RPB7 Degron cells were pretreated with Dox for 12 h, and then 500 μM Auxin for another 3 h. About 10 million cells were used and spiked-in with 10% *Drosophila* S2 cells. Then the chromatin pellets were extracted by using the protocol same as ChIP-seq. After being washed with PBS, the chromatin pellets were lysed by TRIzol reagent. The extracted RNA was sent to Novogene for non-coding RNA library construction and sequencing.

Co-immunoprecipitation analyses

About 10 million cells in two biological replicates were gently scraped, transferred to a centrifuge tube, and lysed by 1 ml western or IP Lysis buffer on ice for 15 min. After centrifugation at 12,000 × *g* for 10 min at 4 °C, the supernatant was collected in a new Eppendorf tube and sampled 50 μl as "Input." One microgram GFP abs or 1 μg rabbit IgG control was added to the rest, respectively, and incubated overnight at

4 °C in rotation. The next day, 30 µl Protein G magnetic beads were added and incubated for another 3 h. For FLAG-IP, anti-Flag magnetic beads were added to the supernatant and also incubated overnight at 4 °C. All the beads were washed on the magnetic stand with lysis buffer three times, each time for 5 min. Then the enriched proteins were eluted in 2 × SDS Loading buffer at 100 °C for 10 min. For RPB4-ChIP Western Blot, lysate of RPB1-N degron cells with or without Auxin treatment were collected as ChIP-seq protocol described above.

Label-free MS and TurboID-MS

The quantitative mass spectrometry analyses of the whole-cell extract and chromatin fractions before and after RPB7 depletion for 3 h were done. For whole-cell extract, cells in 6-well plate were lysed in 200 µl 2 × SDS loading buffer directly and incubated at 100 °C for 10 min to release proteins. For chromatin fractions, same as ChIP-seq protocol, 10 million cells were first lysed by 0.05% NP-40 buffer, and then by lysis buffer with 0.5 M urea and 0.5% NP-40 for 2 min, the pellet was the chromatin fraction. The cells for TurboID-MS of RPB1 were cultured in LB-2i medium for 1 week as described before[62]. After being labeled with 20 µM biotin for 30 min, cells were lysed by RIPA buffer on ice for 15 min, then 1 µl Benzonase was added to release the chromatin fraction at 4 °C for 30 min. After centrifugation at 12,000 × *g* for 10 min at 4 °C, the supernatant was incubated with Dynabeads™ M-280 Streptavidin beads overnight at 4 °C. On the next day, the beads were pelleted by magnetic rack, and the supernatant was discarded, then the beads were washed at room temperature twice with RIPA buffer (1 ml, 2 min), once with 1 M KCl (1 ml, 2 min), once with 0.1 M Na_2CO_3 (1 ml, ~10 s), once with 2 M urea in 10 mM Tris-HCl (pH 8.0) (1 ml, ~10 s), and once with RIPA lysis buffer (1 ml, 2 min at RT) in order. Then the beads were resuspended in 2 × SDS Loading buffer (with 2 mM biotin and 20 mM DTT) to elude the enriched proteins at 100 °C for 10 min. Samples in three biological replicates were prepared. After performing SDS–PAGE electrophoresis and Coomassie blue staining, enriched proteins in gels were excised and subjected to in-gel digestion. Following dehydration with acetonitrile, proteins were subjected to reduction with dithiothreitol (DTT) and alkylation with iodoacetamide (IAA), and then digested with trypsin (Promega, Cat#V5111) at 37 °C overnight. The resulting peptides were extracted using a solution of 0.1% (v/v) formic acid in acetonitrile, subsequently dried in a vacuum centrifuge concentrator, and re-dissolved in 0.1% (v/v) formic acid (in H_2O). The samples were then loaded onto a C18 Reversed Phase HPLC column on Easy-nLC 1200 system (Thermo Fisher Scientific), with a 60 min LC gradient applied at a flow rate of 300 nL/min. HPLC mobile phase composition: Buffer A consisted of 0.1% (v/v) formic acid in H_2O, and Buffer B consisted of 0.1% (v/v) formic acid in 80% acetonitrile. The gradient was programmed as follows: 4%–8% B over 2 min; 8%–25% B over 37 min; 25%–35% B over 11 min; 35%–95% B over 7 min; 95% B for 3 min.

Proteomic analysis was conducted using a Thermo Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific). The MS1 full scan was performed at a resolution of 60,000, covering a mass range of 350–1500 m/z. MS2 scans were generated by HCD fragmentation (30% power) at a resolution of 15,000. Data-dependent acquisition was employed. The fixed first mass for the MS2 spectrum was set at 110.0 m/z, with an isolation window of 1.6 m/z. The LC–MS/MS data were searched against the Uniprot *Mus musculus* [10090] _UP000000589 database using Proteome Discoverer 2.2 software. Trypsin was specified as the enzyme, with a maximum of two missed cleavage sites allowed. Carbamidomethylation (C) was set as a fixed modification, while oxidation (M) and acetylation (protein N-term) were considered as variable modifications. The precursor mass tolerance was set to 10 ppm, and the fragment ion mass tolerance was set to 0.02 Da. The false discovery rates (FDR) were controlled at 1% for both peptide and protein levels.

Structure alignment

The structure prediction of DEL-RPB7, Δ36–42-RPB7, Δ110–116-RPB7, Δ117–138-RPB7, and Δ153–172-RPB7 were done by Alphafold2. After running the run_alphafold.sh file according to the guidance in Github, the most credible model was uploaded into Pymol software for alignment with wild-type RPB7 reported previously (PDB: 6xre).

shRNA knockdown

The shRNA sequences targeting RPB7 were determined by BLOCK-iT™ RNAi Designer. The oligos were annealed and then ligated into the pLKO.1-cloning vector which had been digested with EcoRI and AgeI. After screening for inserts, the positive plasmid DNA was transfected into HEK293T cells for transient knockdown of RPB7 expression by using Lipo2000 Transfection Reagent.

ChIP-seq alignment

Each experiment was performed with two replicates. Raw ChIP-seq reads with spike-in were first trimmed by Cutadapt software (v3.4)[84] and then mapped to mm10 and hg19 genome by Bowtie2 software (v2.3.5.1)[85] in end-to-end mode. Only uniquely and concordantly mapped reads were kept for further analysis. The bigwig files were generated using bamCoverage from deepTools software (v3.5.1)[86] normalized by the reads mapped to hg19 genome. The meta plots were plotted using deepTools software with the bigwig files merged from the two repeats. The active genes were defined as genes whose RNA-seq mapped reads in two repeats of RPB7 untreated degron cells were larger than 3.

MS data analysis

Each experiment was performed with three replicates. The peptides were mapped using Proteome Discoverer software. The abundance for each protein was calculated as the sum of the abundance of its related peptide groups. Proteins containing missing values were discarded. For MS performed in chromatin fraction or RPB1 TurboID-MS, proteins classified as nuclear proteins based on the subcellular location information provided by UniProt database were kept for further analysis. The abundances of the proteins were normalized by the library size of each sample. The differentially enriched proteins in TurboID-MS were identified using the R package DEP (v1.8.0)[87] with a threshold of p value < 0.05 and log2FC > 1. To generate a network of phosphatases enriched in RPB7 ChIP-MS, we selected the transcription-related phosphatases based on a review[7]. Then the phosphatases enriched in the RPB7 ChIP-MS (log2 IP/GFP value > 0, obtained from a previous publication[44]) were selected as inputs to generate a network using String database in Cytoscape software (v3.8.2)[88]. The abundance changes of the phosphatases in RPB1 TurboID-MS upon RPB7 degradation were shown as the border color of the circles.

RNA-seq alignment and differential analysis

Each experiment was performed with two replicates. Raw RNA-seq reads with spike-in were first trimmed by Cutadapt software (v3.4)[84] and then mapped to mm10 and dm6 genome by STAR software (v2.7.10a)[89]. Reads that were not primarily mapped, RCR duplicates and rRNA reads were discarded for further analysis. The bigwig files were generated using bamCoverage from Deeptools software (v3.5.1)[86] normalized by the reads mapped to dm6 genome. Reads that mapped to each gene were counted by featureCounts software (v2.0.1)[90]. The differentially expressed genes were identified by the R package DESeq2 (v1.24.0)[91]. Size factors of the analysis in RPB7 degron cells upon RPB7 degradation were set as the number of the reads mapped as spike-in. Size factors of the analysis in RPB7 mutated cells with auxin treatment compared with WT mESCs were set as the number of reads mapped to *Gapdh*. GO enrichment analysis was performed using the R package clusterProfiler (v3.12.0)[92].

The DoGs found by RNA-Seq were first identified by DoGFinder for each replicate of the samples before and after RPB7 degradation. Reproducible DoGs for each treatment were then merged to determine upregulated DoGs after RPB7 degradation by DESeq2. DoGs were identified as upregulated with *p*adj <0.05 and log2FC > 1. The increased intron retention events were defined as genes with downregulated percent of exon reads. The percent of exon reads were counted as the reads at exons over the reads at the whole transcript including exons and introns. Statistics were identified using *t*-test. Genes were identified as upregulated with p value < 0.05 and log2FC < −1.

ChAR-seq alignment and differential analyses

Each experiment was performed with two replicates. Raw ChAR-seq reads with spike-in were first trimmed by Cutadapt software (v3.4)[84]. rRNA reads were discarded, the remaining reads were then mapped to mm10 and dm6 genome by Bowtie2 software (v2.3.5.1)[85] in end-to-end mode. Only uniquely and concordantly mapped reads were kept for further analysis. The strand-specific bigwig files were generated using bamCoverage from Deeptools software (v3.5.1)[86] normalized by the reads mapped to dm6 genome. To obtain the newly transcribed reads at gene body regions, genes shorter than 1 kb were first removed, then reads that mapped to gene body regions (+300 to TES) of each gene were counted by featureCounts software (v2.0.1)[90]. The differentially expressed genes of ChAR-seq were identified by the R package DESeq2 (v1.24.0)[91]. Size factors were set as the number of the reads mapped as spike-in.

Statistics and reproducibility

The statistical tests used are specified in the figure legends, with the n values denoting the number of independent biological replicates. In cases where no n value is indicated and specific instructions, the experiments were independently repeated three times. No statistical method was used to predetermine the sample size. No data were excluded from the analyses. The experiments were not randomized. The Investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequencing data generated in this study have been deposited in the GEO database under accession code: GSE223475. The mass spectrometry dataset can be found in ProteomeXchange database: PXD048909. Microscopy images data can be found in Mendeley data: https://doi.org/10.17632/vfw72cj42z.4. Source data are provided with this paper.

References

1. Roeder, R. G. 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* **26**, 783–791 (2019).
2. Vo Ngoc, L., Wang, Y. L., Kassavetis, G. A. & Kadonaga, J. T. The punctilious RNA polymerase II core promoter. *Genes Dev.* **31**, 1289–1301 (2017).
3. Field, A. & Adelman, K. Evaluating enhancer function and transcription. *Annu. Rev. Biochem.* **89**, 213–234 (2020).
4. Osman, S. & Cramer, P. Structural biology of RNA polymerase II transcription: 20 years on. *Annu. Rev. Cell Dev. Biol.* **36**, 1–34 (2020).
5. Ferrie, J. J., Karr, J. P., Tjian, R. & Darzacq, X. "Structure"-function relationships in eukaryotic transcription factors: the role of intrinsically disordered regions in gene regulation. *Mol. Cell* **82**, 3970–3984 (2022).
6. Bentley, D. L. Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* **15**, 163–175 (2014).
7. Cossa, G., Parua, P. K., Eilers, M. & Fisher, R. P. Protein phosphatases in the RNAPII transcription cycle: erasers, sculptors, gatekeepers, and potential drug targets. *Genes Dev.* **35**, 658–676 (2021).
8. Morgan, M. A. J. & Shilatifard, A. Reevaluating the roles of histone-modifying enzymes and their associated chromatin modifications in transcriptional regulation. *Nat. Genet.* **52**, 1271–1281 (2020).
9. Chen, X. et al. Structural insights into preinitiation complex assembly on core promoters. *Science* **372**, https://doi.org/10.1126/science.aba8490 (2021).
10. Nojima, T. & Proudfoot, N. J. Mechanisms of lncRNA biogenesis as revealed by nascent transcriptomics. *Nat. Rev. Mol. Cell Biol.* **23**, 389–406 (2022).
11. Maldonado, E. & Reinberg, D. News on initiation and elongation of transcription by RNA polymerase II. *Curr. Opin. Cell Biol.* **7**, 352–361 (1995).
12. Serizawa, H. et al. Association of Cdk-activating kinase subunits with transcription factor TFIIH. *Nature* **374**, 280–282 (1995).
13. Fujinaga, K., Huang, F. & Peterlin, B. M. P-TEFb: the master regulator of transcription elongation. *Mol. Cell* https://doi.org/10.1016/j.molcel.2022.12.006 (2022).
14. Zhou, Q., Li, T. & Price, D. H. RNA polymerase II elongation control. *Annu. Rev. Biochem.* **81**, 119–143 (2012).
15. Rice, A. P. Roles of CDKs in RNA polymerase II transcription of the HIV-1 genome. *Transcription* **10**, 111–117 (2019).
16. Bacon, C. W. & D'Orso, I. CDK9: a signaling hub for transcriptional control. *Transcription* **10**, 57–75 (2019).
17. Eick, D. & Geyer, M. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem. Rev.* **113**, 8456–8490 (2013).
18. Cortazar, M. A. et al. Control of RNA Pol II speed by PNUTS-PP1 and Spt5 dephosphorylation facilitates termination by a "Sitting Duck Torpedo" mechanism. *Mol. Cell* **76**, 896–908.e894 (2019).
19. Parua, P. K. et al. A Cdk9-PP1 switch regulates the elongation-termination transition of RNA polymerase II. *Nature* **558**, 460–464 (2018).
20. Tellier, M. et al. CDK9 and PP2A regulate RNA polymerase II transcription termination and coupled RNA maturation. *EMBO Rep.* **23**, e54520 (2022).
21. Rodríguez-Molina, J. B., West, S. & Passmore, L. A. Knowing when to stop: transcription termination on protein-coding genes by eukaryotic RNAPII. *Mol. Cell* **83**, 404–415 (2023).
22. Kecman, T. et al. Elongation/termination factor exchange mediated by PP1 phosphatase orchestrates transcription termination. *Cell Rep.* **25**, 259–269.e255 (2018).
23. Chen, Z. et al. Transcription recycling assays identify PAF1 as a driver for RNA Pol II recycling. *Nat. Commun.* **12**, 6318 (2021).
24. Chen, Z. et al. Phosphorylated MED1 links transcription recycling and cancer growth. *Nucleic Acids Res.* **50**, 4450–4463 (2022).
25. Fuda, N. J. et al. Fcp1 dephosphorylation of the RNA polymerase II C-terminal domain is required for efficient transcription of heat shock genes. *Mol. Cell. Biol.* **32**, 3428–3437 (2012).
26. Lei, L., Ren, D., Finkelstein, A. & Burton, Z. F. Functions of the N- and C-terminal domains of human RAP74 in transcriptional initiation, elongation, and recycling of RNA polymerase II. *Mol. Cell. Biol.* **18**, 2130–2142 (1998).
27. Noe Gonzalez, M., Blears, D. & Svejstrup, J. Q. Causes and consequences of RNA polymerase II stalling during transcript elongation. *Nat. Rev. Mol. Cell Biol.* **22**, 3–21 (2021).
28. Nakazawa, Y. et al. Ubiquitination of DNA damage-stalled RNAPII promotes transcription-coupled repair. *Cell* **180**, 1228–1244.e1224 (2020).
29. Tufegdžić Vidaković, A. et al. Regulation of the RNAPII pool is integral to the DNA damage response. *Cell* **180**, 1245–1261.e1221 (2020).

30. Bregman, D. B. et al. UV-induced ubiquitination of RNA polymerase II: a novel modification deficient in Cockayne syndrome cells. *Proc. Natl. Acad. Sci. USA* **93**, 11586–11590 (1996).
31. Ratner, J. N., Balasubramanian, B., Corden, J., Warren, S. L. & Bregman, D. B. Ultraviolet radiation-induced ubiquitination and proteasomal degradation of the large subunit of RNA polymerase II. Implications for transcription-coupled DNA repair. *J. Biol. Chem.* **273**, 5184–5189 (1998).
32. Lee, K. B., Wang, D., Lippard, S. J. & Sharp, P. A. Transcription-coupled and DNA damage-dependent ubiquitination of RNA polymerase II in vitro. *Proc. Natl. Acad. Sci. USA* **99**, 4239–4244 (2002).
33. Nguyen, V. T. et al. In vivo degradation of RNA polymerase II largest subunit triggered by alpha-amanitin. *Nucleic Acids Res.* **24**, 2924–2929 (1996).
34. Aoi, Y. et al. SPT5 stabilization of promoter-proximal RNA polymerase II. *Mol. Cell* **81**, 4413–4424.e4415 (2021).
35. Song, A. & Chen, F. X. The pleiotropic roles of SPT5 in transcription. *Transcription* **13**, 53–69 (2022).
36. Todone, F., Brick, P., Werner, F., Weinzierl, R. O. & Onesti, S. Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. *Mol. Cell* **8**, 1137–1143 (2001).
37. Calvo, O. RNA polymerase II phosphorylation and gene looping: new roles for the Rpb4/7 heterodimer in regulating gene expression. *Curr. Genet.* **66**, 927–937 (2020).
38. Goler-Baron, V. et al. Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes. *Genes Dev.* **22**, 2022–2027 (2008).
39. Harel-Sharvit, L. et al. RNA polymerase II subunits link transcription and mRNA decay to translation. *Cell* **143**, 552–563 (2010).
40. Allepuz-Fuster, P. et al. Rpb4/7 facilitates RNA polymerase II CTD dephosphorylation. *Nucleic Acids Res.* **42**, 13674–13688 (2014).
41. Allepuz-Fuster, P. et al. RNA polymerase II plays an active role in the formation of gene loops through the Rpb4 subunit. *Nucleic Acids Res.* **47**, 8975–8987 (2019).
42. Jiang, Y. et al. Cross-regulome profiling of RNA polymerases highlights the regulatory role of polymerase III on mRNA transcription by maintaining local chromatin architecture. *Genome Biol.* **23**, 246 (2022).
43. Jiang, Y. et al. Genome-wide analyses of chromatin interactions after the loss of Pol I, Pol II, and Pol III. *Genome Biol.* **21**, 158 (2020).
44. Li, Y. et al. Targeted protein degradation reveals RNA Pol II heterogeneity and functional diversity. *Mol. Cell* **82**, 3943–3959.e3911 (2022).
45. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* **45**, 439–446 (2012).
46. Wild, T. & Cramer, P. Biogenesis of multisubunit RNA polymerases. *Trends Biochem. Sci.* **37**, 99–105 (2012).
47. Wang, H. et al. The transcriptional coactivator RUVBL2 regulates Pol II clustering with diverse transcription factors. *Nat. Commun.* **13**, 5703 (2022).
48. Boulon, S. et al. HSP90 and its R2TP/Prefoldin-like cochaperone are involved in the cytoplasmic assembly of RNA polymerase II. *Mol. Cell* **39**, 912–924 (2010).
49. Forget, D. et al. The protein interaction network of the human transcription machinery reveals a role for the conserved GTPase RPAP4/GPN1 and microtubule assembly in nuclear import and biogenesis of RNA polymerase II. *Mol. Cell. Proteom.* **9**, 2827–2839 (2010).
50. Liou, S. H., Singh, S. K., Singer, R. H., Coleman, R. A. & Liu, W. L. Structure of the p53/RNA polymerase II assembly. *Commun. Biol.* **4**, 397 (2021).
51. Carminati, M., Rodríguez-Molina, J. B., Manav, M. C., Bellini, D. & Passmore, L. A. A direct interaction between CPF and RNA Pol II links RNA 3′ end processing to transcription. *Mol. Cell* **83**, 4461–4478.e4413 (2023).
52. Aibara, S., Dienemann, C. & Cramer, P. Structure of an inactive RNA polymerase II dimer. *Nucleic Acids Res.* **49**, 10747–10755 (2021).
53. Somesh, B. P. et al. Multiple mechanisms confining RNA polymerase II ubiquitylation to polymerases undergoing transcriptional arrest. *Cell* **121**, 913–923 (2005).
54. Clopper, K. C. & Taatjes, D. J. Chemical inhibitors of transcription-associated kinases. *Curr. Opin. Chem. Biol.* **70**, 102186 (2022).
55. Robb, C. M. et al. Chemically induced degradation of CDK9 by a proteolysis targeting chimera (PROTAC). *Chem. Commun.* **53**, 7577–7580 (2017).
56. Winter, G. E. et al. BET bromodomain proteins function as master transcription elongation factors independent of CDK9 recruitment. *Mol. Cell* **67**, 5–18.e19 (2017).
57. Kwiatkowski, N. et al. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014).
58. Marineau, J. J. et al. Discovery of SY-5609: a selective, noncovalent inhibitor of CDK7. *J. Med. Chem.* **65**, 1458–1480 (2022).
59. Akhtar, M. S. et al. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. *Mol. Cell* **34**, 387–393 (2009).
60. Glover-Cutter, K. et al. TFIIH-associated Cdk7 kinase functions in phosphorylation of C-terminal domain Ser7 residues, promoter-proximal pausing, and termination by RNA polymerase II. *Mol. Cell. Biol.* **29**, 5455–5464 (2009).
61. Cho, K. F. et al. Proximity labeling in mammalian cells with TurboID and split-TurboID. *Nat. Protoc.* **15**, 3971–3999 (2020).
62. Tian, K., Wang, R., Huang, J., Wang, H. & Ji, X. Subcellular localization shapes the fate of RNA polymerase III. *Cell Rep.* **42**, 112941 (2023).
63. Archambault, J. et al. FCP1, the RAP74-interacting subunit of a human protein phosphatase that dephosphorylates the carboxyl-terminal domain of RNA polymerase IIO. *J. Biol. Chem.* **273**, 27593–27601 (1998).
64. Kamada, K., Roeder, R. G. & Burley, S. K. Molecular mechanism of recruitment of TFIIF- associating RNA polymerase C-terminal domain phosphatase (FCP1) by transcription factor IIF. *Proc. Natl. Acad. Sci. USA* **100**, 2296–2299 (2003).
65. Zheng, H. et al. Identification of Integrator-PP2A complex (INTAC), an RNA polymerase II phosphatase. *Science* **370**, https://doi.org/10.1126/science.abb5872 (2020).
66. Landsverk, H. B. et al. WDR82/PNUTS-PP1 prevents transcription-replication conflicts by promoting RNA polymerase II degradation on chromatin. *Cell Rep.* **33**, 108469 (2020).
67. Wang, X. et al. RPAP2 regulates a transcription initiation checkpoint by inhibiting assembly of pre-initiation complex. *Cell Rep.* **39**, 110732 (2022).
68. Hausmann, S. & Shuman, S. Characterization of the CTD phosphatase Fcp1 from fission yeast. Preferential dephosphorylation of serine 2 versus serine 5. *J. Biol. Chem.* **277**, 21213–21220 (2002).
69. Hu, W. F. et al. CTDP1 regulates breast cancer survival and DNA repair through BRCT-specific interactions with FANCI. *Cell Death Discov.* **5**, 105 (2019).
70. Kamenski, T., Heilmeier, S., Meinhart, A. & Cramer, P. Structure and mechanism of RNA polymerase II CTD phosphatases. *Mol. Cell* **15**, 399–407 (2004).
71. Vos, S. M. et al. Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature* **560**, 607–612 (2018).
72. Daulny, A. et al. Modulation of RNA polymerase II subunit composition by ubiquitylation. *Proc. Natl. Acad. Sci. USA* **105**, 19649–19654 (2008).
73. Ribar, B., Prakash, L. & Prakash, S. ELA1 and CUL3 are required along with ELC1 for RNA polymerase II polyubiquitylation and

degradation in DNA-damaged yeast cells. *Mol. Cell. Biol.* **27**, 3211–3216 (2007).
74. Meka, H., Werner, F., Cordell, S. C., Onesti, S. & Brick, P. Crystal structure and RNA binding of the Rpb4/Rpb7 subunits of human RNA polymerase II. *Nucleic Acids Res.* **33**, 6435–6444 (2005).
75. Ujvári, A. & Luse, D. S. RNA emerging from the active site of RNA polymerase II interacts with the Rpb7 subunit. *Nat. Struct. Mol. Biol.* **13**, 49–54 (2006).
76. Xiang, K. et al. Crystal structure of the human symplekin-Ssu72-CTD phosphopeptide complex. *Nature* **467**, 729–733 (2010).
77. Hsin, J. P., Xiang, K. & Manley, J. L. Function and control of RNA polymerase II C-terminal domain phosphorylation in vertebrate transcription and RNA processing. *Mol. Cell. Biol.* **34**, 2488–2498 (2014).
78. Tian, B. & Manley, J. L. Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* **18**, 18–30 (2017).
79. Cho, H. et al. A protein phosphatase functions to recycle RNA polymerase II. *Genes Dev.* **13**, 1540–1552 (1999).
80. Kalaydjieva, L. & Chamova, T. In *GeneReviews®* (eds Adam, M. P. et al.) (University of Washington, Seattle Copyright© 1993-2023, University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle. All rights reserved., 1993).
81. Lassuthova, P. et al. Congenital cataract, facial dysmorphism and demyelinating neuropathy (CCFDN) in 10 Czech Gypsy children-frequent and underestimated cause of disability among Czech Gypsies. *Orphanet. J. Rare Dis.* **9**, 46 (2014).
82. Varon, R. et al. Partial deficiency of the C-terminal-domain phosphatase of RNA polymerase II is associated with congenital cataracts facial dysmorphism neuropathy syndrome. *Nat. Genet.* **35**, 185–189 (2003).
83. Bhatt, D. M. et al. Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150**, 279–290 (2012).
84. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
85. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
86. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–165, (2016).
87. Zhang, X. et al. Proteome-wide identification of ubiquitin interactions using UbIA-MS. *Nat. Protoc.* **13**, 530–550 (2018).
88. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
89. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
90. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* **30**, 923–930 (2014).
91. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
92. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS J. Integr. Biol.* **16**, 284–287 (2012).
93. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

Acknowledgements

This work was supported by funds from the National Natural Science Foundation of China and the National Key R&D Program of China (Grants 32350003, 2024YFA1106903, 32170569) and the Qidong-SLS Innovation Fund, Innovation Fund of Chengdu Research Institute, and by grants from the Peking-Tsinghua Center for Life Sciences, the Beijing Advanced Center of RNA Biology (BEACON), the State Key Laboratory of Gene Function and Modulation Research, the Key Laboratory of Cell Proliferation and Differentiation of the Ministry of Education at Peking University School of Life Sciences to X.J. We thank the National Center for Protein Sciences at Peking University in Beijing, China, for Ms. H.Y. for help with flow cytometry sorting, D.L. and Q.Z. for help with mass spectrometry, G.L.L. for help with biochemical assays, and L.Q.F. for help with confocal fluorescence imaging.

Author contributions

X.J. conceived and supervised the project. H.N.Z. performed most of the experiments and B.Q.Y. helped to generate the truncated RPB7 expressing cell. D.X.J. helped to validate the RPB4-N and RPB4-C degron cells, IgG ChIP-seq and RPB7 degron followed cellular fractionation analyses. Q.Q.X. performed most of the bioinformatic analyses. X.J. wrote the manuscript with input from H.N.Z., Q.Q.X., and D.X.J.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-57513-2.

Correspondence and requests for materials should be addressed to Xiong Ji.

Peer review information *Nature Communications* thanks Giacomo Cossa and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.