Paf1C regulates RNA polymerase II progression by modulating elongation rate

Liming Hou[a,b], Yating Wang[a,b], Yu Liu[c], Nan Zhang[a,b], Ilya Shamovsky[d], Evgeny Nudler[d,e], Bin Tian[c], and Brian David Dynlacht[c,1]

[a]Department of Pathology, New York University School of Medicine, New York, NY 10016; [b]Perlmutter Cancer Institute, New York University School of Medicine, New York, NY 10016; [c]Department of Microbiology, Biochemistry and Molecular Genetics, Rutgers New Jersey Medical School, Newark, NJ 07103; [d]Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY 10016; and [e]Howard Hughes Medical Institute, New York University School of Medicine, New York, NY 10016

Edited by Robert Tjian, University of California, Berkeley, CA, and approved June 6, 2019 (received for review March 12, 2019)

Elongation factor Paf1C regulates several stages of the RNA polymerase II (Pol II) transcription cycle, although it is unclear how it modulates Pol II distribution and progression in mammalian cells. We found that conditional ablation of Paf1 resulted in the accumulation of unphosphorylated and Ser5 phosphorylated Pol II around promoter-proximal regions and within the first 20 to 30 kb of gene bodies, respectively. Paf1 ablation did not impact the recruitment of other key elongation factors, namely, Spt5, Spt6, and the FACT complex, suggesting that Paf1 function may be mechanistically distinguishable from each of these factors. Moreover, loss of Paf1 triggered an increase in TSS-proximal nucleosome occupancy, which could impose a considerable barrier to Pol II elongation past TSS-proximal regions. Remarkably, accumulation of Ser5P in the first 20 to 30 kb coincided with reductions in histone H2B ubiquitylation within this region. Furthermore, we show that nascent RNA species accumulate within this window, suggesting a mechanism whereby Paf1 loss leads to aberrant, prematurely terminated transcripts and diminution of full-length transcripts. Importantly, we found that loss of Paf1 results in Pol II elongation rate defects with significant rate compression. Our findings suggest that Paf1C is critical for modulating Pol II elongation rates by functioning beyond the pause-release step as an "accelerator" over specific early gene body regions.

Paf1 complex | transcription elongation rate | RNA polymerase II | nucleosome occupancy

Recent studies have begun to focus on the rate-limiting steps that control transcriptional elongation in mammalian cells. Indeed, it is now recognized that many mammalian genes exhibit transcriptional pausing of RNA Polymerase II (Pol II) within the first 25 to 50 bp downstream of the transcription start site (TSS), demonstrated by the presence of promoter-proximal peaks of Pol II. It is also clear that the multistep regulation of elongation is critical for maintaining steady-state as well as inducible levels of gene expression (1, 2). Pausing is enforced by several multisubunit complexes, including DRB sensitivity-inducing factor (DSIF, comprised of Spt4 and Spt5) and negative elongation factor (NELF). Phosphorylation of Ser2 of Pol II, DSIF, and NELF by positive transcription elongation factor b (pTEFb) results in dissociation of NELF, which coincides with pause release and entry of Pol II into the productive elongation phase (1).

Another factor with an established role in elongation is the Paf1 complex (Paf1C). Mammalian Paf1C consists of 5 subunits (Cdc73, Ctr9, Leo1, Paf1, and Ski8) which, together with Rtf1, assembles a multifunctional platform on genic regions (3, 4). The distribution of all subunits of yeast and mammalian Paf1C on chromatin has been mapped genome-wide (5–7). Paf1C is enriched on active genes. Intriguingly, the genomic profiles for Paf1C occupancy differs in yeast and mammalian cells: whereas yeast Paf1 peaks over gene bodies and tapers off toward the TSS and transcription end site (TES), mammalian Paf1C occupancy is highest at both the TSS and TES, and in many cases, binding extends upstream of the TSS and well beyond the TES (6, 7).

Recent elegant cryo-EM studies from the Cramer laboratory have shown how mammalian Paf1C, Spt5/DSIF, and another elongation factor, Spt6, can interact with Pol II to assemble an activated elongation complex (8, 9). These structural studies suggested that Paf1C could displace NELF for pause release subsequent to phosphorylation of the latter protein by pTEFb (comprised of cyclin T/CDK9), and Spt6 might promote active Pol II elongation by opening the RNA clamp formed by DSIF. However, Spt5, Paf1C, and Spt6 associate with distinct regions of Pol II, and since neither Paf1C nor Spt6 makes contacts with the active site of Pol II, the stimulatory effect of both factors is likely to be allosteric in nature (8, 9). Despite these important conceptual advances, much remains unknown about whether or not the activities of Paf1C are mechanistically coupled with Spt5 and Spt6 and how they function during the early stages of transcription elongation. Moreover, much of our knowledge regarding these elongation factors has been obtained using yeast, and given differences between yeast and mammalian Paf1C—especially very large differences in gene lengths, which could influence elongation rates—additional analysis of the role of this elongation complex in mammalian cells is required.

Significance

The factors that regulate RNA polymerase II (Pol II) elongation rate and processivity are poorly understood. Here, we show that the Paf1 complex (Paf1C) modulates Pol II elongation rate, and loss of Paf1 results in accumulation of Ser5P in the first 20 to 30 kb of gene bodies, coinciding with reduction of histone H2B ubiquitylation specifically within this region. Moreover, reduced elongation rates provoked by Paf1 depletion resulted in defects in Pol II processivity and premature termination of transcription. Paf1 ablation did not impact the recruitment of other key elongation factors, suggesting that Paf1C function may be mechanistically distinguishable from each of these factors. Our data pinpoint Paf1C as a key modulator of Pol II elongation rates across mammalian genes.

Author contributions: L.H. and B.D.D. designed research; L.H., Y.W., and B.D.D. performed research; L.H., Y.W., Y.L., B.T., and B.D.D. contributed new reagents/analytic tools; L.H., Y.W., Y.L., N.Z., I.S., E.N., B.T., and B.D.D. analyzed data; and L.H. and B.D.D. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: All large-scale data sets have been deposited in Gene Expression Omnibus (GEO) database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE116169).

[1]To whom correspondence may be addressed. Email: brian.dynlacht@nyumc.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1904324116/-/DCSupplemental.

Published online June 28, 2019.

We have previously shown that depletion of Paf1C components leads to substantial defects in 3′ end processing and polyadenylation site (PAS) usage (7). We speculated that such processing defects could arise from Pol II stalling provoked by the loss of certain Paf1C subunits, such as Paf1, and indeed, multiple groups contemporaneously observed similar accumulations of Pol II over gene bodies upon Paf1 removal (6, 7, 10). However, despite the biochemical and genetic studies in yeast that revealed a role for the Paf1C as a positive regulator of elongation, the role of mammalian Paf1C as an elongation factor remained unclear. Notably, one group suggested a role for this factor in enforcement of Pol II pausing (10, 11), whereas another study suggested that Paf1C could, conversely, act by promoting pause release (6). The latter study attempted to link an increase in total and Ser5P Pol II to reduced enrichment of Cdk9 and Cdk12, but loss of Cdk9 or Cdk12 alone did not affect Pol II distribution or occupancy in their system (6). Moreover, the loss of Cdk9 resulted in a reduction of Paf1 recruitment, prompting a model wherein the recruitment of Paf1 is pTEFb dependent. Therefore, a mechanism that explains promoter-proximal Pol II pausing after loss of Paf1 has remained elusive.

Here, we clarify these conflicting models by delineating the role of mammalian Paf1C in transcriptional elongation and uncovering processes regulated by this complex. Using genome-wide sequencing of nascent RNA after conditional ablation of Paf1 in murine myoblasts, we show that Paf1 positively regulates elongation and Pol II progression, and in the absence of this protein, unphosphorylated Pol II accumulates around TSS-proximal pausing regions, an event that coincides with increased nucleosome occupancy in this region. Furthermore, the loss of Paf1 triggers an accumulation of Ser5 phosphorylation (Ser5P) and overt reductions in H2B ubiquitylation (H2Bub) specifically over the first 20 to 30 kb of gene bodies. In addition, we found that loss of Paf1 leads to severely reduced elongation rates and accumulation of nascent transcripts over early gene bodies. Since defects in elongation extend over gene bodies well beyond the TSS, our data suggest that Paf1 regulates Pol II distribution beyond the promoter-proximal pause point. Our results thus support a model in which Paf1C is an essential rate-limiting transcriptional elongation factor that functions within a defined window to promote Pol II processivity, elongation rate variation, and progression past polyadenylation sites in early gene body regions to prevent accumulation of immature, truncated transcripts.

Results

Conditional Ablation of Paf1 Results in Pol II Accumulation. In an effort to analyze the impact of complete loss of Paf1 on Pol II distribution and transcriptional elongation, we used CRISPR/Cas9 to ablate the *Paf1* gene in the untransformed mouse muscle myoblast line, C2C12. We were unable to isolate stable populations of cells with homozygous *Paf1* deletions, suggesting that it is an essential gene. Therefore, we used gene editing to produce C2C12 cells with conditional expression of an sgRNA-resistant, doxycycline (Dox)-inducible Flag-Paf1 transgene in a Paf1-deletion background, allowing us to study the reversibility of phenotypes associated with loss of this factor (Fig. 1*A*). We confirmed the editing of the *Paf1* gene through sequencing (*SI Appendix*, Fig. S1*A*). Paf1 expression was reduced in these cells (termed Paf1-knockout [Paf1-KO]) by >90% within 1 d of Dox removal, and levels of exogenous Flag-Paf1 were similar to endogenous Paf1 after Dox induction (Fig. 1*B*). Consistent with its essentiality, we note that depletion of Paf1 resulted in reduced cell growth rates after Dox removal for 2 d, although it did not affect cell viability after 4 d (*SI Appendix*, Fig. S1*B*). Therefore, we restricted our analyses to a window within 2 to 4 d after Dox removal. As an additional indicator of Paf1 depletion, we showed that bulk H2B monoubiquitylation of lysine 120 (H2Bub) was

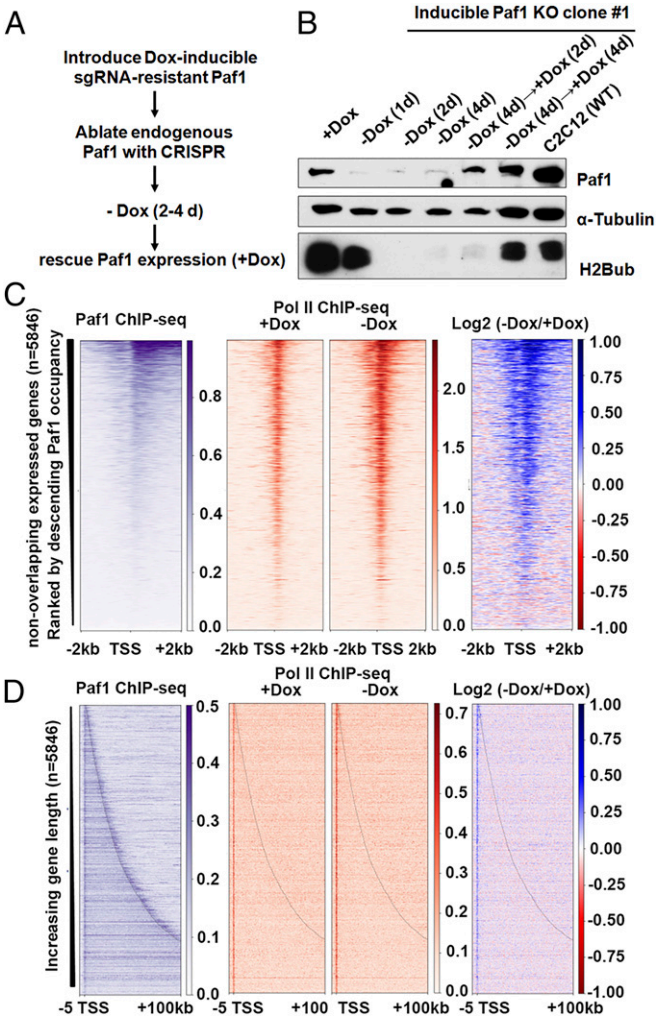

Fig. 1. Paf1 deletion through gene editing leads to increased RNA Pol II occupancy in the TSS-proximal region. (*A*) Schema for generating conditional Paf1-KO myoblasts using CRISPR/Cas9 gene editing. (*B*) Western blot showing the loss of Paf1 after Dox removal (−Dox) from 1 to 4 d. Depletion of Paf1 is accompanied by loss of H2Bub, and its level can be partially rescued by reintroducing Dox for 4 d. Note that the mobility of Dox-induced Flag-Paf1 is shifted with respect to endogenous Paf1, denoted by wild-type C2C12 (WT) in last lane. α-Tubulin was used as loading control. (*C*) Heatmaps show unphosphorylated Pol II occupancy over 2 kb upstream and downstream of TSS after ablation of Paf1. Nonoverlapping 5,846 expressed genes were sorted by descending levels of Paf1 occupancy (*Left*), and corresponding Pol II ChIP-seq read densities and $\log_2$ fold change (−Dox/+Dox) of normalized Pol II ChIP-seq read densities are shown to the *Right*. (*D*) Heatmaps were plotted as increasing gene length over regions 5 kb upstream to 100 kb downstream of the TSS. Signals are aligned according to their TSS, and TESs are indicated by dotted black lines.

erased in whole cell extracts of 2 inducible Paf1-KO clones within 2 to 4 d after Dox removal, as expected from previous studies (Fig. 1*B* and *SI Appendix*, Fig. S1*C*) (7, 12). Importantly, H2Bub levels could be partially restored by Dox treatment of Paf1-depleted cells for 4 d, confirming that this *Paf1* null phenotype could be reversed and was not due to off-target or indirect effects (Fig. 1*B*).

We and others previously observed that Pol II accumulates on genes after silencing of mammalian Paf1 (6, 7, 10). Pol II carboxyl-terminal domain (CTD) phosphorylation state is intimately linked to progression through the elongation cycle: unphosphorylated Pol II is first modified on Ser5 by cyclin

H/CDK7 before recruitment of pTEFb (cyclin T/CDK9) and cyclin K/CDK12, which promote Ser2 phosphorylation, and Ser2-phosphorylated Pol II is associated with the active elongation complex. Using an antibody (8WG16) that primarily recognizes the unphosphorylated CTD of Pol II (13), we first examined Pol II occupancy before and after Paf1 removal by performing ChIP-seq and ChIP-qPCR in our inducible Paf1-KO cell line. We found that unphosphorylated Pol II levels were markedly increased on TSS-proximal regions after depletion of Paf1 on 5846 nonoverlapping, expressed genes (fragments per kilobase per million mapped reads [FPKM] > 1, RNA-seq) within 5 kb upstream of the TSS and 10 kb downstream of TESs (Fig. 1 *C* and *D* and *SI Appendix*, Fig. S1 *D–F*), consistent with previous experiments (7). Moreover, this phenotype could be partially rescued through expression of Paf1 (*SI Appendix*, Fig. S1*E*). We found that Paf1 depletion provoked Pol II accumulation irrespective of gene length or expression levels, and all genes exhibited a narrow band of accumulated polymerase near the TSS (Fig. 1*D* and *SI Appendix*, Fig. S1*F*). These results are consistent with recent findings in yeast that Paf1 regulates transcriptional elongation on nearly all genes (14). When we overlaid Paf1 recruitment with unphosphorylated Pol II in wild-type or Paf1-KO cells genome-wide, we found that Paf1 enrichment peaked slightly downstream (*SI Appendix*, Fig. S1*G*). Since Paf1 enrichment peaked downstream of unphosphorylated Pol II, we speculated that Paf1 could act to promote elongation at a point coinciding with, or subsequent to, Pol II CTD phosphorylation (see below).

We asked whether the impact of Paf1 ablation was restricted to this subunit or whether it could be generalized to other Paf1C components. We ablated 5 subunits using siRNAs (*SI Appendix*, Fig. S1*H*). We found that depletion of multiple subunits resulted in Pol II accumulation, although loss of Paf1 or Ctr9 had the greatest impact, and Leo1 or Ski8 depletion had minimal impact (*SI Appendix*, Fig. S1*I*). These results are interesting in light of recent cryo-EM studies that demonstrate mutually exclusive binding of Ctr9 and NELF to Pol II (8, 9). Furthermore, Ski8 is positioned on the elongation complex through binding of Ctr9 to Pol II, and based on its location within the complex, Leo1 could potentially play a role in stabilizing Spt5 and DNA rewinding upstream of Pol II and thereby facilitate elongation (8, 9). Our functional data, however, suggest that Ski8 and Leo1 may not play a prominent role in promoting Pol II progression.

Overall, our results confirm the utility of our Paf1-KO cells and suggest that depletion of multiple Paf1C subunits triggers genome-wide accumulation of unphosphorylated Pol II on TSS-proximal regions. These results suggested that Paf1C loss triggers either increased initiation or defects in elongation rate, possibilities which we subsequently test below.

Loss of Paf1 Triggers Genome-Wide Changes in Nucleosome Occupancy. It is known that the TSS-proximal (+1) nucleosome is a major barrier to Pol II and that Pol II frequently pauses within a window encompassed by the first few downstream nucleosomes (15–18). Paf1C is thought to serve as a Pol II-associated platform that recruits factors required for Pol II progression, including the E3 ligase, RNF20, and other factors required for H2B ubiquitylation, a modification believed to enhance the rate of RNA Pol II progression (19, 20) and processivity (12). While the loss of Paf1 could explain the genome-wide reduction in RNF20 recruitment and erasure of H2Bub, a mechanistic explanation for the observed increase in Pol II density was lacking. We therefore investigated potential changes in the chromatin landscape by comparing nucleosome locations and density in wild-type and Paf1-depleted myoblasts. First, using MNase-seq, we examined genome-wide nucleosome occupancy on all nonoverlapping expressed genes after Paf1 depletion in Paf1-KO cells. We found that Paf1-KO cells exhibited considerably enhanced nucleosome accumulation over TSS-proximal regions (Fig. 2*A* and *SI Appendix*, Fig. S2 *A* and *B*). We performed analogous studies on cells depleted of Paf1 through siRNA-mediated knockdown (KD) and observed similar results (*SI Appendix*, Fig. S2*C*), and importantly, we showed that increased nucleosome occupancy could be largely rescued by reexpression of Paf1, confirming that this defect was indeed specifically provoked by the loss of Paf1 (Fig. 2*A* and *SI Appendix*, Fig. S2*A*). In parallel, we also performed assays for transposase-accessible chromatin (ATAC-seq) in siRNA-depleted (siPaf1) and control cells (siNS) (Fig. 2 *B* and *C* and *SI Appendix*, Fig. S2 *D* and *E*). These experiments confirmed the reduced accessibility of chromatin at TSS-proximal regions after Paf1 ablation. These findings suggested that Pol II accumulation could result from increased nucleosome occupancy around TSS-proximal regions, and loss of Paf1 could therefore impose a considerable barrier to Pol II progression.

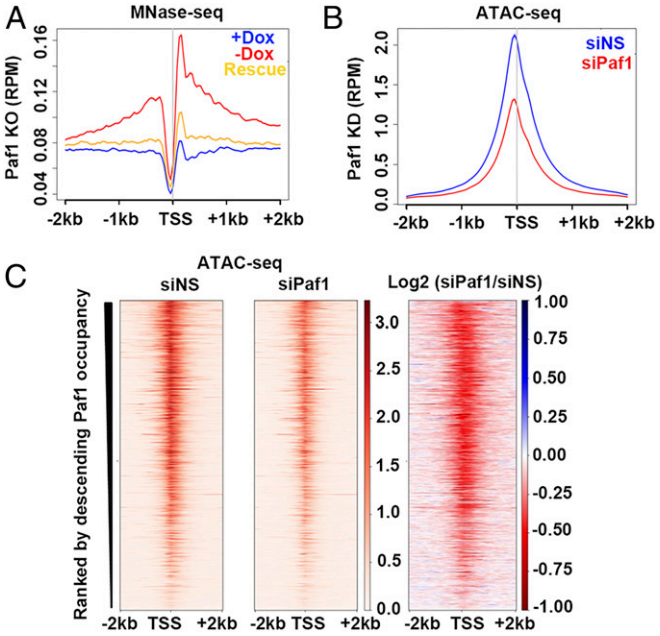

Fig. 2. Paf1-deficient cells exhibit marked alterations in nucleosome occupancy over TSS-proximal regions. (*A*) Metaplots of MNase-seq read densities around a 2-kb window flanking the TSS of all nonoverlapping expressed genes in Paf1-KO cells with (+Dox), without (−Dox) Paf1 and rescue conditions by adding Dox back. Reexpression of Paf1 partially rescues altered nucleosome occupancy in Paf1-KO cells. RPM, reads per million. (*B*) Metaplots of ATAC-seq read densities around a 2-kb region flanking TSS in Paf1-depleted (siPaf1) versus control siNS cells. (*C*) Heatmaps of ATAC-seq reads density were plotted after ranking according to Paf1 occupancy. Log_2 ratios of normalized ATAC-seq reads comparing Paf1-depleted to control cells were calculated and plotted.

Loss of Paf1 Results in Accumulation of Ser5P Pol II in the First ∼20 to 30 kb of Gene Bodies. We sought to confirm and extend our understanding of aberrant Pol II accumulation in cells lacking Paf1. Our foregoing studies (detecting unphosphorylated Pol II) were unable to fully reveal the engaged polymerase molecules in elongation complexes. To that end, we performed ChIP-seq to detect Ser5 phosphorylation of the Pol II C-terminal domain (CTD), since promoter-paused Pol II exhibits Ser5 phosphorylation (21), and compared the genome-wide profiles in the presence and absence of Paf1. Heatmaps and metagene plots of Ser5-phosphorylated Pol II (Ser5P Pol II) within ±2 kb of the TSS indicated a significant increase after Paf1 depletion (Fig. 3*A* and *SI Appendix*, Fig. S3*A*). Surprisingly, when we plotted these data as a function of gene length, we observed an accumulation of Ser5P Pol II in early gene body regions within a window ∼20 to 30 kb downstream of the TSS after

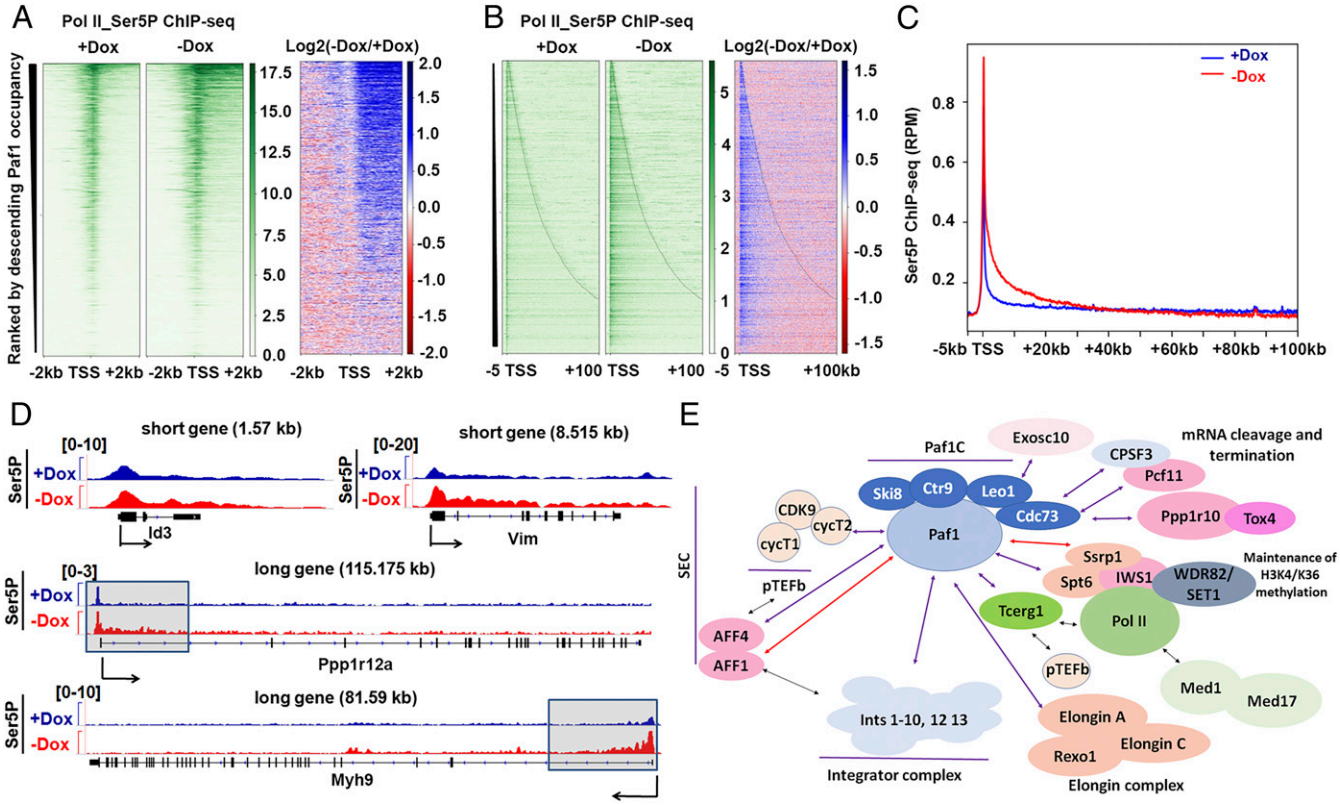

Fig. 3. Ser5 phosphorylation of Pol II accumulates in early gene bodies after Paf1 depletion. (*A* and *B*) Heatmaps of Ser5 phosphorylated Pol II (Ser5P) ChIP-seq densities were plotted based on Paf1 occupancy (*A*) or gene length (*C*) as described in the legend to Fig. 1 *C* and *D*, respectively. (*C*) A metagene plot of normalized Ser5P ChIP-seq densities over regions 5 kb upstream to 100 kb downstream of the TSS (+Dox versus −Dox). (*D*) Profiles for individual genes of various lengths (short versus long) are shown. (*E*) Proteomic analysis of chromatin-associated proteins interacting with Paf1. Purple arrows indicate associations identified in this study, red arrows indicate associations identified in one MS experiment, and black arrows indicate previously published interactions.

Paf1 depletion (Fig. 3 *B–D*). The extent of Ser5 phosphorylation downstream of the TSS was independent of gene length (Fig. 3 *B–D*). On short genes, therefore, Pol II adopts a "persistently accumulated" state throughout the entire length of the gene, whereas on most long genes, the accumulated polymerase is restricted to a narrow, early gene body region downstream of the TSS (Fig. 3 *B–D*).

Proteomic Analysis of Paf1C Interacting Proteins. Next, we sought a mechanistic explanation for Pol II accumulation by analyzing the interactome of Paf1 using mass spectrometric (MS) sequencing. Previous studies that employed immunopurification of a Paf1C subunit (Cdc73) from human cancer cells led to the conclusion that this factor stably interacts with 3′ mRNA processing factors, specifically cleavage and polyadenylation factors, CPSF and CstF (22). On the other hand, parallel studies with antibodies directed against Paf1 and Leo1 did not precipitate these factors. In an effort to identify Paf1C-associated proteins in a nontransformed cell line, we isolated MNase-treated, soluble chromatin from our inducible KO expressing Flag-Paf1 or control non-Flag tagged wild-type C2C12 myoblast cell lines and performed Flag immunoaffinity purification and mass spectrometric sequencing (*Materials and Methods*). We identified a cadre of transcription and RNA processing and cleavage factors that were enriched in Flag-Paf1 expressing cells, with a >2-fold difference in the number of peptides recovered compared with control C2C12 myoblasts (Fig. 3*E* and Dataset S1). Surprisingly, the cohort of polypeptides that we purified were largely nonoverlapping with those derived from cancer cells (22). In support of a positive role for Paf1 in elongation of Pol II, we found that all subunits of pTEFb (CDK9, Cyclin T1/T2) copurified, together with AFF4, a component of the Super Elongator Complex (SEC). We also identified the majority (12/14) of subunits associated with Integrator, a protein complex with roles in 3′ end processing of small nuclear (snRNA) and enhancer (eRNA) RNAs (23–26). Spt6 and an associated protein, IWS1, were also identified in our proteomic data, and other proteins with roles in elongation, including the Elongin complex (Elongin A and an Elongin A-binding protein, EloA-BP1/Rexo1) and TCERG1, copurified. Interestingly, we identified several transcript processing and termination factors, including CPSF3 and Pcf11, and Exosc10, an RNA exosome component responsible for degrading aberrant transcripts and RNA surveillance (27, 28). To confirm physiological interactions between Paf1 and a subset of these proteins, we induced expression of Flag-Paf1 in Paf1-KO cells, isolated chromatin, and performed immunoprecipitation and Western blotting. These studies confirmed the interaction of Paf1 with other known Paf1C subunits, Spt6, Elongin A, Pol II, and 2 subunits of the Integrator complex (*SI Appendix*, Fig. S3*B*). The complexes associated with Paf1 differ from those obtained through purification of Cdc73, suggesting that Paf1C could participate in subunit-specific or cell type-dependent interactions. These studies provide a resource for studying novel Paf1C activities in nontransformed cells and suggested that Paf1 indeed interacts with positive elongation factors to promote transcriptional elongation.

Mechanistic Investigation of Factors that Coordinately Regulate Transcription Elongation. We were intrigued by the relatively weak association of Paf1 with kinases that phosphorylate CTD Ser2 (CDK9) (*SI Appendix*, Fig. S3*B* and Dataset S1). Previous

studies indicated that the loss of Paf1 could trigger increases in total and Ser5 phosphorylated Pol II and diminished recruitment of human CDK9 and CDK12 (6). Notably, genome-wide reductions in CDK9 recruitment after Paf1 ablation were observed only after normalization to total Pol II (6). Our attempts at detecting CDK9 by ChIP were unsuccessful, however, perhaps owing to the availability of antibodies that did not effectively recognize the mouse protein bound to chromatin. Therefore, we interrogated Ser2 phosphorylation by ChIP-seq, as it reflects the activity of both kinases. We found that this modification was not reduced after Paf1 depletion (*SI Appendix*, Fig. S3*C*). Rather, its occupancy increased, suggesting that the increase in accumulated Pol II was not due to depletion of CDK9 or CDK12 (also see below). Further, as before, the increase in Ser2 phosphorylation was independent of gene length and encompassed a region within ∼20 to 30 kb downstream of the TSS (*SI Appendix*, Fig. S3*C*).

Next, we examined 2 other elongation factors, Spt6 and FACT (Ssrp1 and Spt16), since we recovered peptides for each of these factors after Paf1 purification, and they have been shown to modulate nucleosome occupancy to facilitate elongation. Spt6 is a histone H3 chaperone that enhances elongation, and FACT dissociates nucleosomal H2A/H2B dimers to promote Pol II progression. We were particularly interested in a potential role for Spt6, since yeast loss-of-function mutations result in defects in nucleosome occupancy (29, 30). Moreover, genetic and physical interactions between Paf1C and Spt6, Spt5, and FACT have been observed in plants, yeast, and mammals (8, 9, 31–34), but the relationship between each of these proteins and Paf1C has not been explored in the context of elongation in mammalian cells. Paf1C can make direct contacts with Spt6 (9) and FACT, or it could interact through Pol II as an intermediary.

To determine whether these factors cooperate with Paf1C to suppress Pol II accumulation, we first examined their protein levels after singly depleting each of these factors. We found that depletion of Paf1 did not significantly impact the levels of Spt5, Spt6, or Spt16 or vice versa (*SI Appendix*, Fig. S4*A*). Next, we examined unphosphorylated Pol II levels after knockdown of Spt6 by ChIP-qPCR, and we found that loss of Spt6 also results in accumulation of unphosphorylated Pol II on the majority of loci we examined (*SI Appendix*, Fig. S4*B*). We therefore asked if the elongation defect provoked by Paf1 loss was due to the loss of Spt6 from chromatin. We examined the recruitment of Spt6 to several genes by ChIP-qPCR after depleting Paf1 in Paf1-KO cells or using siRNAs. Surprisingly, we observed enhanced recruitment of Spt6 in cells depleted of Paf1 through either approach (*SI Appendix*, Fig. S4 *C* and *D*). The impact of Paf1 depletion could be attributed specifically to loss of this protein since Spt6 recruitment was rescued upon restoration of Paf1 in KO cells (*SI Appendix*, Fig. S4*C*). To interrogate the genome-wide impact of Paf1 removal, we performed ChIP-seq on Spt6 after depleting Paf1. We observed dramatically increased recruitment of Spt6 after loss of Paf1 in both Paf1-KO (*SI Appendix*, Fig. S4 *E–G*) and Paf1 KD cells (*SI Appendix*, Fig. S4 *H–J*). The accumulation of Spt6 was most obvious within early gene body regions (*SI Appendix*, Fig. S4 *G* and *J*). Moreover, we also observed enhanced recruitment of both FACT subunits (Spt16 and SSRP1) after Paf1 knockdown by ChIP-qPCR (*SI Appendix*, Fig. S5 *A* and *B*). Likewise, Spt16 was not globally depleted from chromatin after removal of Paf1, and instead, we observed dramatic increases in the occupancy of this factor by ChIP-seq (*SI Appendix*, Fig. S5 *C–E*).

Finally, we examined the interplay between Paf1 and Spt5, a component of DSIF required for efficient transcriptional elongation in yeast and mammalian cells (8, 35–37), although we did not detect a stable association between Paf1 and Spt5 in our proteomic experiments. We found that depletion of Paf1 resulted in increased occupancy of Spt5 by ChIP-qPCR, and Spt5 occupancy reverted to wild-type levels upon reexpression of Paf1 (*SI Appendix*, Fig. S5*F*). Interestingly, we showed that Spt5 accumulated genome-wide using ChIP-seq, and once again, Spt5 was most highly enriched within early gene bodies after loss of Paf1 (*SI Appendix*, Fig. S5 *G–I*). When we normalized Spt6, Spt16, and Spt5 ChIP-seq data to Ser2-phosphorylated Pol II, however, the differences between wild-type and Paf1-depleted cells were largely negated, suggesting that these factors continued to travel with elongating Pol II (*SI Appendix*, Fig. S5*J*). These experiments therefore definitively rule out a role for loss of any of these elongation factors as an explanation for Pol II accumulation.

We conclude that the phenotypes associated with loss of Paf1C—accumulation of Pol II and increased nucleosome occupancy (and loss of H2Bub and reduced elongation rates; see later)—cannot be ascribed to an inability to recruit key elongation factors, FACT, Spt5, or Spt6 to chromatin. These findings are important in light of the observed, extensive interactions that Paf1C subunits make with both Spt6 and Spt5 (8, 9) and suggest that stable interactions of the latter factors with Pol II occur independently of Paf1.

The Impact of Paf1 Loss on H2B Ubiquitylation. H2B monoubiquitylation has been shown to positively correlate with RNA Pol II elongation rates in mammalian cells (20). Moreover, H2Bub levels are tied to recruitment of Rad6/RNF20, which is further linked to Paf1C and Pol II occupancy (19, 38–41). We initially investigated this modification in Paf1-KO cells using ChIP-qPCR. We found that H2Bub levels were decreased on a few expressed genes in wild-type myoblasts, and levels of this modification could be restored upon Paf1 reexpression, attesting to the conclusion that ubiquitylation of this residue depends on the presence of Paf1 (*SI Appendix*, Fig. S6*A*). We next performed ChIP-seq on H2Bub and observed reductions in H2Bub after Paf1 removal (Fig. 4 *A–E*). This reduction occurred on all expressed genes, irrespective of gene expression levels and length (Fig. 4 *A–E*). We further validated these findings in siRNA-depleted (siPaf1) cells (*SI Appendix*, Fig. S6 *B–D*).

Although the reductions in H2Bub could be anticipated based on the global loss of this modification in cell extracts (Fig. 1*B*), we found remarkable differences in the loss of this mark on 5′ versus 3′ ends of gene bodies. On long genes (>20 kb), Paf1 was required for maximal ubiquitylation over the early gene body region, after which the modification was independent of this factor (Fig. 4 *B, D*, and *E*). For genes of intermediate length (10 kb to 20 kb) and short genes (<5 kb), we uncovered a strong dependence on Paf1 throughout the entire gene body (Fig. 4 *B, D*, and *E*). Paradoxically, we note that Paf1 is recruited throughout the gene bodies of long and short genes, and genome-wide profiles show increasing enrichment from 5′ to 3′ regions, with greatest enrichment at regions surrounding and downstream of the TES (Fig. 1*D*) and (7). However, metagene profiles of H2Bub ChIP-seq data show that the reductions in H2Bub occurred predominantly in a window ∼20 to 30 kb downstream of the TSS (Fig. 4*F*). Of note, we observed a strong correlation between enhanced Ser5P Pol II occupancy and reduced H2Bub over regions within 20 kb downstream of the TSS, in striking contrast with a window spanning +20 kb to the TES (Fig. 4*G*).

Since Paf1 removal resulted in significant diminutions in H2Bub levels over early ∼20- to 30-kb gene body regions, we asked whether ablation of RNF20, the E3 ligase responsible for ubiquitylation, could account for increased occupancy of unphosphorylated Pol II. Ablation of RNF20 completely abolished cellular H2B ubiquitylation in whole cell extracts (*SI Appendix*, Fig. S6*E*), as expected. However, using ChIP-qPCR, we found no significant differences in unphosphorylated Pol II enrichment after RNF20 silencing (*SI Appendix*, Fig. S6*F*). Importantly, these results suggest that Pol II accumulation after

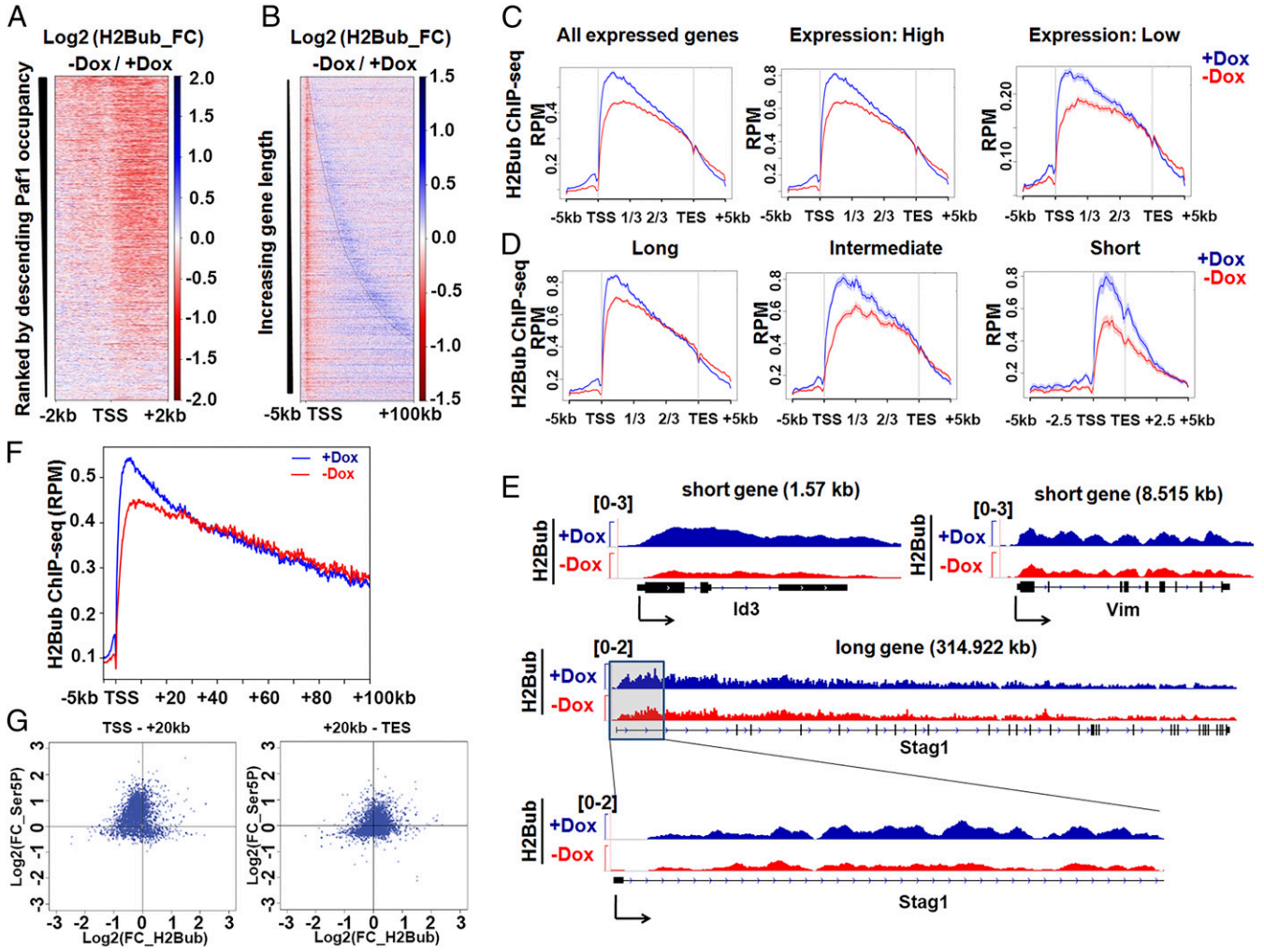

Fig. 4. Paf1-depleted cells exhibit reductions in H2Bub deposition over early gene body regions. (*A*) Heatmaps showing $\log_2$ ratio of H2Bub ChIP-seq densities at promoter-proximal regions (TSS −2 kb to TSS +2 kb) in Paf1-KO cells ranked according to decreasing Paf1 occupancy. (*B*) Heatmaps showing $\log_2$ ratio of H2Bub ChIP-seq densities were plotted based on increasing gene length. (*C* and *D*) Metagene plots of H2Bub ChIP-seq read densities in Paf1-KO cells with (+) or without (–) Dox treatment in different groups. Genes were grouped according to expression level (*C*; high [FPKM > 30]) or low (0 < FPKM < 30) expression) or length (*D*; long [>20 kb], intermediate [10 kb < gene length< 20 kb], and short [<5 kb]), as indicated. (*E*) H2Bub ChIP-seq for individual short or long genes are shown. Zoomed-in shaded rectangular area denotes region with decreased H2Bub levels. (*F*) A metagene plot of normalized H2Bub ChIP-seq densities over regions 5 kb upstream to 100 kb downstream of the TSS (+Dox versus −Dox). (*G*) Scatterplots of $\log_2$ fold changes of normalized Ser5P occupancy versus $\log_2$ fold changes of normalized H2Bub in 2 regions (TSS to +20 kb and +20 kb to TES) in Paf1-KO cells.

Paf1 removal cannot be explained solely by the loss of RNF20, which probably results in a reduction of H2Bub throughout all genic regions rather than a reduction in a specific early gene body region, as we observed after Paf1 removal.

Our results are notable because they suggest that (*i*) Paf1C plays a profound role in regulating H2B ubiquitylation over a specialized early gene body region, and Pol II progression is also intimately linked to Paf1C recruitment within this region; (*ii*) deposition and/or maintenance H2Bub levels becomes independent of Paf1C recruitment downstream of this window; and (*iii*) despite its demonstrated role in Pol II elongation, regulation of H2Bub levels via RNF20 cannot fully account for the role of Paf1 in promoting Pol II elongation.

Depletion of Paf1 Leads to Reduced Pol II Processivity. The above observations (increased nucleosome occupancy, Pol II accumulation, and reduced H2Bub levels) suggested that Paf1 loss could trigger significant changes in gene expression. Therefore, we examined steady-state gene expression levels after depleting Paf1 using RNA-seq. As shown in *SI Appendix*, Fig. S7*A*, 450 genes were up-regulated, and 447 genes were down-regulated in the absence of Paf1 (adjusted *P* value <0.05; $\log_2$ FPKM fold change >1) using DEseq2 analysis. Data from replicate experiments were highly concordant (*SI Appendix*, Fig. S7*B*). When we plotted heatmaps based on either Paf1 occupancy or gene length for these RNA-seq data, we found that there was a significant reduction in RNA-seq reads near the TSS in the sense direction, followed by increased read density in early gene body regions (*SI Appendix*, Fig. S7 *C*–*E*). Moreover, we observed a noticeable global reduction in RNA-seq reads in downstream genic body regions (*SI Appendix*, Fig. S7 *D* and *E*) (note the red signal in the last 2 panels in *SI Appendix*, Fig. S7*D*), which suggests that loss of Paf1 could result in a reduction of full-length transcripts. The increased RNA-seq reads in early gene body regions implies a defect in Pol II processivity (nucleotide additions per initiation event), accompanied by an increase in prematurely terminated shortened transcripts, as shown for individual genes (*SI Appendix*, Fig. S7*F*).

The observed increase in reads within early gene body regions, combined with decreased read counts further downstream, suggested that DEseq2 (which takes into account the entire gene

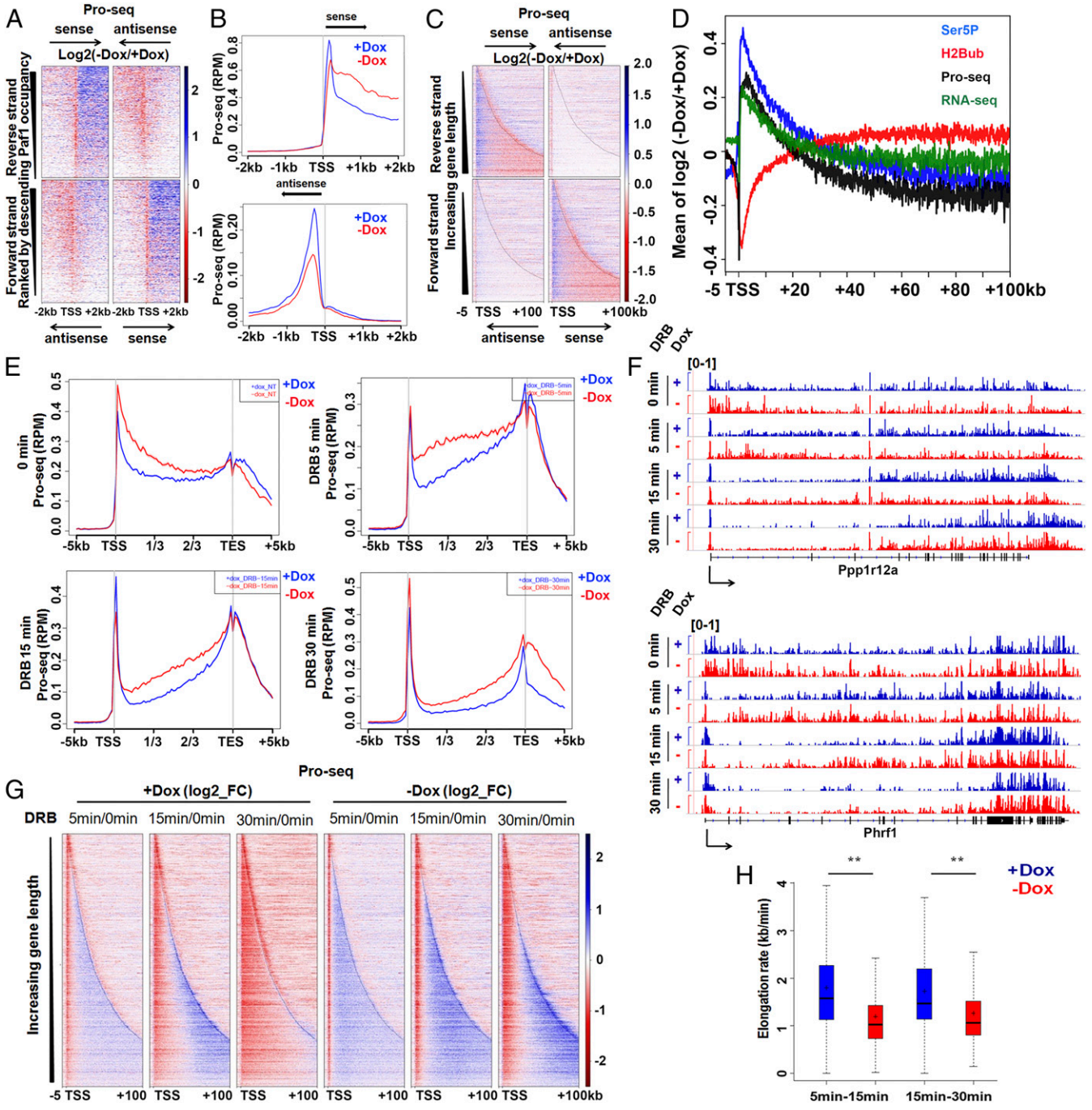

Fig. 5. Paf1 loss dramatically hampers transcriptional elongation. (*A*) Heatmaps showing $\log_2$ fold change (−Dox/+Dox) of normalized Proseq read densities for genes in both forward (n = 3,023) and reverse (n = 2,745) strands within 2-kb regions flanking the TSS before or after Paf1 depletion in Paf1-KO cell. Both sense and antisense directions for both strands are indicated. (*B*) Metagene analysis of Proseq data for both sense and antisense directions in regions within 2 kb of the TSS. (*C*) Analysis as in *A*, but plotted as a function of gene length in Paf1-KO cells expressing (+Dox) or lacking (−Dox) Paf1 expression. (*D*) Metagene plot of mean $\log_2$ ratios shows that the increase in Proseq reads correlates with the accumulation of Ser5P, RNA-seq reads, and the reduction of H2Bub over an early ∼20- to 30-kb gene body window. Data in *SI Appendix*, Fig. S7*E* were replotted here for comparison. (*E*) Metagene plots of Proseq data at each time interval after DRB treatment (0, 5, 15, and 30 min) over the entire gene body plus 5-kb flanking regions upstream of TSS and downstream of TES. RPM, reads per million. (*F*) Proseq data are shown for two individual genes before (+Dox) and after (−Dox) Paf1 ablation at each time point after DRB treatment (0, 5, 15, and 30 min). Note that the "travel distance" of engaged Pol II after DRB treatment was affected by Paf1 ablation (+Dox versus −Dox). (*G*) Heatmaps show the $\log_2$ fold change of normalized Proseq signal in DRB treatment time points (5 min, 15 min, and 30 min) compared with untreated control (0 min) in +Dox and −Dox conditions. Note that the red color indicates the "travel distance" of engaged Pol II after DRB treatment. (*H*) Elongation rates were calculated using a HMM method. Boxplots show the distribution of elongation rates in cells expressing (blue) or lacking (red) Paf1 at indicated time intervals before and after Paf1 removal. "+" sign in the boxplot denotes the mean of elongation rates. $^{**}P < 0.01$ by Wilcoxon rank-sum test. Note the compressed distribution of elongation rates after Paf1 removal.

body when calculating read counts, without distinguishing whether transcripts are prematurely terminated) could underestimate the number of differentially expressed/regulated genes in *SI Appendix*, Fig. S7*A*. Moreover, we found that genes down-regulated after Paf1 depletion tended to be highly expressed (with higher Paf1 occupancy), whereas up-regulated genes tended to be weakly expressed (with lower Paf1 occupancy), suggesting that up-regulation could result from an indirect effect of Paf1 ablation (*SI Appendix*, Fig. S7*G*). Our previous work suggested that loss of Paf1 leads to defects in PAS usage, which resulted in production of aberrant, prematurely shortened transcripts in early exons and introns (7). Our current RNA-seq data (and Proseq data below) suggest that Pol II stalling over early gene bodies is likely to explain this increased frequency of premature processing.

Loss of Paf1C Results in Significant Reductions in Elongation Rate. Given the observed changes in nucleosome density, accumulation of Ser5 phosphorylated Pol II, H2Bub reductions, and increased prematurely terminated transcripts from early gene body regions after Paf1 depletion, we investigated whether the loss of Paf1 could impinge on Pol II elongation rates. A prior study in which run-on transcription assays were coupled to genome-wide sequencing of nascent RNA suggested that, in some cell types, Paf1 could act as a negative regulator of elongation through the enforcement of pausing and that removal of Paf1 could trigger release of paused Pol II into gene bodies (10). However, one alternative explanation for these data could not be ruled out, namely, that Paf1 could function like a positive elongation factor such as Spt5. Notably, knockdown of Spt5 resulted in the increased density of Pol II across gene bodies, prompting the conclusion that loss of this factor promotes release of Pol II from promoters (42, 43). Computational modeling and further experimentation, on the other hand, suggested that the observed increases in Pol II density could stem from reduced processivity or elongation rates in early gene body regions after Spt5 removal (35, 37, 43).

To distinguish the effects on elongation in gene bodies versus promoter clearance, we performed genome-wide precision nuclear run-on and sequencing (Proseq) to analyze nascent RNA production, as well as active Pol II density, location, directionality, and rates in Paf1-KO cells (44, 45). We examined our KO cells in the presence or absence of Dox and performed a time course in which we treated cells with 5,6-dichloro-1-β-d-ribofuranosylbenzimidazole (DRB) to block CDK9 activity for 5, 15, and 30 min, enabling us to observe Pol II escape from a promoter-proximal pause and calculate elongation rates over the duration of treatment. We showed that this method was robust, and data from replicates were highly correlated (*SI Appendix*, Fig. S8*A*). Heatmap and metagene analyses show that the nascent RNA signal was significantly reduced at the TSS-proximal region in both directions, followed by an increased level of nascent RNA downstream in the sense direction, but not in the antisense direction (Fig. 5 *A* and *B*). The decreased promoter-proximal nascent RNA peak suggests a reduction in initiation (see *Discussion*). When we plotted the heatmaps based on gene length, we found that the increased Pol II density was restricted to early gene body regions (Fig. 5*C*). Remarkably, there was a near-precise overlap between these increased nascent reads and the accumulation of Ser5P, increases in prematurely terminated transcripts, and the reduction in H2Bub (Fig. 5*D*). Moreover, scatterplot analysis shows that the increased Proseq signal correlated with a decrease in H2Bub and accumulation of Ser5P Pol II over the first 20 kb, but not in the region from 20 kb downstream of the TSS to TES (*SI Appendix*, Fig. S8 *B* and *C*). Over the course of DRB treatment, we observed that the Pol II elongation rate was significantly reduced in Paf1-depleted cells, as shown in both metagene and single-gene analyses (Fig. 5 *E* and *F*). When we plotted Proseq density heatmaps showing $\log_2$ ratios of later time points (5 min, 15 min, and 30 min after DRB treatment) versus untreated cells (0 min) in the presence or absence of Paf1 over 5,846 nonoverlapping expressed genes, we found that the elongation rates were more "homogenized" or compressed after loss of Paf1, irrespective of gene length (Fig. 5*G*) (note the width of the red "band" in Fig. 5*G* after Dox removal). Next, we used these Proseq density maps to generate line plots based on the $\log_2$ ratios, which allowed us to approximate the mean "travel distance" over various time intervals, and we found the mean travel distance at the beginning of DRB treatment (0 to 5 min) was ∼12 kb in the +Dox condition, whereas the travel distance was significantly reduced after Paf1 depletion (∼5 kb) (0 to 5 min; *SI Appendix*, Fig. S8*D*). Notably, Pol II took 15 min or 30 min to travel through the first ∼12 kb or ∼28 kb, respectively, in Paf1-depleted cells, whereas Pol II required 5 min or 15 min, respectively, to travel the same distance in Paf1-replete cells (*SI Appendix*, Fig. S8*D*).

Next, we employed a hidden Markov model (HMM) to identify the position of the "wave front" of engaged Pol II on a subset of genes longer than 70 kb and to infer Pol II elongation rates during the DRB time course (*Materials and Methods*). We calculated average elongation rates of 1.80 kb/min (5- to 15-min interval) and 1.72 kb/min (15- to 30-min interval) in the presence of Paf1 (Fig. 5*H*), which are comparable to average elongation rates of ∼2 kb/min deduced using other mammalian cell lines with related computational methods (2, 46, 47). In contrast, we found that Paf1 removal led to substantially reduced elongation rates, with average elongation rates of 1.20 kb/min (5- to 15-min interval) and 1.26 kb/min (15- to 30-min interval) in Paf1-KO cells lacking Paf1, and the impact of Paf1 loss was evident at both early (5 to 15 min) and late (15 to 30 min) time intervals (Fig. 5*H*; $P < 0.01$).

Since Pol II elongation rates were compressed into a very narrow range in the absence of Paf1, compared with Paf1-replete cells at all time intervals (Fig. 5 *G* and *H*), we conclude that Paf1 plays a rate-limiting role in Pol II elongation through early gene bodies (Fig. 6). Our studies also unify prior findings and suggest that the removal of Paf1 triggers significant genome-wide reductions in Pol II elongation rates, reinforcing the role of Paf1C as an essential elongation factor that functions beyond the pause-release step.

Discussion

Experiments in cancer cells have suggested that loss of Paf1 results in pause release (10), a finding at variance with the studies presented here and with those in yeast, which

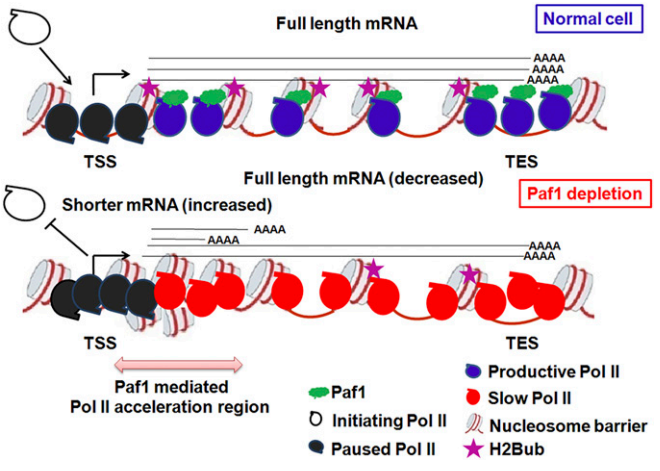

Fig. 6. Model indicating roles for Paf1 in controlling productive transcription elongation over a specialized window downstream of the TSS.

support a positive role for Paf1 in transcriptional elongation. In contrast, we found that silencing Paf1 results in Pol II accumulation owing to a reduced elongation rate. It is possible that differences in cell types could explain this discrepancy (10). If true, this would suggest that Paf1C plays opposite roles in nontransformed and transformed cells. However, we favor an alternative, parsimonious explanation that comports with mathematical modeling of other elongation factors (43), namely, that Paf1C is a positive elongation factor that promotes efficient elongation, allowing Pol II to attain optimal speed and processivity in the first 20–30 kb of gene bodies (Fig. 6). The altered distribution of Pol II observed here is distinct from previously studied promoter-proximal pausing, which occurs ∼25 to 50 bp from the TSS (1). Indeed, upon Paf1 depletion, Pol II accumulation occurred over a much greater distance from the TSS, ∼20 to 30 kb downstream of TSS. This suggests that the increased Pol II accumulation around the promoter-proximal region upon loss of Paf1 could also result from the reduced elongation rate provoked by loss of Paf1. Since the aberrant Pol II accumulation expands to a region well beyond the TSS-proximal Pol II pause site, we propose that Paf1 regulates Pol II distribution beyond the promoter-proximal pause point. Our findings therefore suggest a revised model for Paf1 function that explains previously divergent conclusions (Fig. 6) (6, 10).

In addition, H2Bub levels have been shown to strongly correlate with Pol II elongation rates in multiple cell lines (20). Although genic elongation rates appear highly variable, Pol II was shown to systematically accelerate as it progressed through early gene bodies in human cells (2, 46). Our studies indicate that Paf1-mediated H2B ubiquitylation could play a particularly important role in regulating early elongation, where diminution of H2Bub over Paf1-depleted genes strongly correlates with a reduction in Pol II elongation rates (Fig. 5*D*). This finding stands in stark contrast to another histone modification, H3K36me3, which is aberrantly shifted toward 5′ regions of genes with slow Pol II (48). Interestingly, we were able to reproduce this finding using our own data (*SI Appendix*, Fig. S8*E*), and these findings suggest that the observed elevation in H3K36me3 triggered by increased Pol II dwell time is not restricted to use of slow Pol II mutants described previously (48). Notably, we also found that in the absence of Paf1, H2Bub was most dramatically compromised over the first ∼20 to 30 kb of gene bodies, whereas distal gene body regions were less significantly impacted. This result suggests mechanistic differences in the deposition and/or maintenance of this histone modification as a function of gene length, and future studies will be required to identify the molecular basis of this observation. Intriguingly, we note that mammalian Spt5 plays a role over a critical window ∼15 to 20 kb downstream of the TSS, whereas this window is condensed to ∼500 bp in yeast, which has considerably shorter genes (35, 37). Future studies will be required to determine whether Spt5 and Paf1C are mechanistically linked within this window, although our data clearly show that recruitment of these factors is not obligatorily connected (*SI Appendix*, Fig. S5).

It has been shown that elongation is normally slower in early gene body regions compared with downstream genic regions in normal conditions, with Pol II gradually accelerating through the gene body until it reaches maximum elongation speed (2, 46). We propose that this may explain why Pol II tends to slow down more obviously in this region after loss of Paf1. Further, slow polymerase may also lead to the accumulation of paused Pol II around promoter-proximal regions and thereby inhibit transcriptional initiation, consistent with our Proseq data (Fig. 5 *A* and *B*) and recent reports (49, 50).

Another interesting finding from our studies is that ablation of Paf1 compressed the elongation rates of genes globally to a considerably narrower range within the first ∼28 kb (Fig. 5*G* and *SI Appendix*, Fig. S8*D*). It is known that elongation rates vary widely across the transcriptome (46). Previous studies in mammalian cells indicated that elongation rates could vary by up to 5-fold, ranging from 0.5 to 2.4 kb/min (46). In addition, these studies indicated significant variation within each group of elongation rates, on the order of ∼0.4 to 0.8 kb/min. Thus, it is interesting that removal of Paf1 not only reduces overall Pol II elongation speeds but also results in a more homogeneous elongation rate. Our data suggest that Paf1C could be a fine-tuning, "rheostat-like" factor whose recruitment helps dictate the observed variance in elongation speeds by accelerating Pol II progression in a yet-to-be-discovered manner. It will be interesting to determine the factors other than Paf1C that contribute to elongation rate variation.

Furthermore, we observed a striking increase in nucleosome occupancy after Paf1 removal, concomitant with increased TSS-proximal Pol II pausing and reduced initiation and elongation rates. It is well documented that the first few downstream nucleosomes are a major barrier to Pol II progression (15–18). The increased nucleosome density observed after loss of Paf1 could thus account for the accumulation of TSS-proximal unphosphorylated Pol II and decreased initiation rates observed in our Proseq experiments. However, others have observed distinct profiles: RNA Pol II pausing at highly regulated genes was shown to be characterized by reduced nucleosome occupancy within the early transcribed region and nucleosomal occlusion of their promoters, while housekeeping genes exhibited the opposite pattern (51). Further studies are needed to elucidate the interplay between Pol II pausing and nucleosome occupancy around promoters and regions downstream of the TSS in mammalian cells.

Importantly, our findings lead to several major conclusions and a revised model for Paf1C function: 1) Reminiscent of findings with Spt5 (35, 37), there is a transcriptional window within early gene body regions essential for control of Pol II elongation rates; 2) Paf1C controls the rate of Pol II elongation and plays an instrumental role in elongation throughout this window; and 3) Paf1C could play a more persistent role within this region as a polymerase "accelerator" by functioning after pause release (Fig. 6). Indeed, 2 of our findings strongly suggest a role for Paf1C beyond promoter-proximal pause release (6). First, pause release occurs as a byproduct of CDK9 phosphorylation of Ser2, but we did not see a global reduction in phosphorylated Ser2 of Pol II CTD after Paf1 reduction. On the contrary, we observed an increase in this form of modified Pol II (*SI Appendix*, Fig. S3*C*). Second, we observed that DRB treatment was able to prevent pause release in Proseq experiments, both in sense and antisense directions, even after removal of Paf1 (*SI Appendix*, Fig. S8*F*). This suggests that CDK9-mediated pause release may still be enforced in Paf1-depleted cells. Thus, we propose that Paf1C function is crucial for Pol II elongation over a specialized region within ∼20 to 30 kb downstream of the TSS (Fig. 6). Future studies will be required to reveal the entire complement of factors required to promote elongation through this window.

Finally, we propose that drastic reductions in Pol II elongation rates and accumulation of Pol II over the first ∼20- to 30-kb region could explain the usage of alternative polyA sites and transcript shortening at early introns that we observed previously (7). This phenomenon, whereby slow Pol II promotes alternative polyA site usage in the 3′ UTR, has been observed in flies and human cells (52, 53). Our findings suggest that an imbalance in Paf1C activity could trigger the production of immature nascent transcripts.

Materials and Methods

All siRNA sequences and antibodies used in this study are listed in *SI Appendix*, Tables S1 and S2, respectively. See *SI Appendix, Materials and*

Methods for full details of all experimental procedures. A detailed description of cell lines, vector construction, biochemical methods, mass spectrometry, RNAi, ChIP, RNA-seq, ATAC-seq, ChIP-seq, MNase-seq, Proseq, all subsequent analyses of these genome-wide data, and derivation of the HMM can be found in *SI Appendix*. All software and algorithms used in this study are listed in *SI Appendix*, Table S3. Data were deposited in GEO, accession no. GSE116169.

ACKNOWLEDGMENTS. We thank the New York University Genome Technology and Proteomics cores, particularly S. Nayak, T. Ueberheide, P. Meyn, and A. Heguy. We thank M. Hoque (Rutgers New Jersey Medical School) for assistance with library preparation. We thank Y. Yang for H2Bub ChIP-seq data. We thank D. Reinberg and G. LeRoy for the generous gift of antibodies. I.S. and E.N. were supported in part by NIH R01 GM126891, the Blavatnik Family Foundation, and by the Howard Hughes Medical Institute. B.D.D. was supported by NIH Grant R01GM122395.

1. K. Adelman, J. T. Lis, Promoter-proximal pausing of RNA polymerase II: Emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731 (2012).
2. C. G. Danko *et al*., Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol. Cell* **50**, 212–222 (2013).
3. J. A. Jaehning, The Paf1 complex: Platform or player in RNA polymerase II transcription? *Biochim. Biophys. Acta* **1799**, 379–388 (2010).
4. S. B. Van Oss, C. E. Cucinotta, K. M. Arndt, Emerging insights into the roles of the Paf1 complex in gene regulation. *Trends Biochem. Sci.* **42**, 788–798 (2017).
5. A. Mayer *et al*., Uniform transitions of the general RNA polymerase II transcription complex. *Nat. Struct. Mol. Biol.* **17**, 1272–1278 (2010).
6. M. Yu *et al*., RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. *Science* **350**, 1383–1386 (2015).
7. Y. Yang *et al*., PAF complex plays novel subunit-specific roles in alternative cleavage and polyadenylation. *PLoS Genet.* **12**, e1005794 (2016).
8. S. M. Vos, L. Farnung, H. Urlaub, P. Cramer, Structure of paused transcription complex Pol II-DSIF-NELF. *Nature* **560**, 601–606 (2018).
9. S. M. Vos *et al*., Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. *Nature* **560**, 607–612 (2018).
10. F. X. Chen *et al*., PAF1, a molecular regulator of promoter-proximal pausing by RNA polymerase II. *Cell* **162**, 1003–1015 (2015).
11. F. X. Chen *et al*., PAF1 regulation of promoter-proximal pause release via enhancer activation. *Science* **357**, 1294–1298 (2017).
12. L. Wu, L. Li, B. Zhou, Z. Qin, Y. Dou, H2B ubiquitylation promotes RNA Pol II processivity via PAF1 and pTEFb. *Mol. Cell* **54**, 920–931 (2014).
13. M. Patturajan *et al*., Growth-related changes in phosphorylation of yeast RNA polymerase II. *J. Biol. Chem.* **273**, 4689–4694 (1998).
14. H. Fischl, F. S. Howe, A. Furger, J. Mellor, Paf1 has distinct roles in transcription elongation and differential transcript fate. *Mol. Cell* **65**, 685–698.e8 (2017).
15. V. A. Bondarenko *et al*., Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. *Mol. Cell* **24**, 469–479 (2006).
16. L. S. Churchman, J. S. Weissman, Nascent transcript sequencing visualizes transcription at nucleotide resolution. *Nature* **469**, 368–373 (2011).
17. C. M. Weber, S. Ramachandran, S. Henikoff, Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell* **53**, 819–830 (2014).
18. T. Kujirai *et al*., Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* **362**, 595–598 (2018).
19. R. Pavri *et al*., Histone H2B monoubiquitination functions cooperatively with FACT to regulate elongation by RNA polymerase II. *Cell* **125**, 703–717 (2006).
20. G. Fuchs, D. Hollander, Y. Voichek, G. Ast, M. Oren, Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. *Genome Res.* **24**, 1572–1583 (2014).
21. A. K. Boehm, A. Saunders, J. Werner, J. T. Lis, Transcription factor and polymerase recruitment, modification, and movement on dhsp70 in vivo in the minutes following heat shock. *Mol. Cell. Biol.* **23**, 7628–7637 (2003).
22. O. Rozenblatt-Rosen *et al*., The tumor suppressor Cdc73 functionally associates with CPSF and CstF 3′ mRNA processing factors. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 755–760 (2009).
23. D. Baillat *et al*., Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**, 265–276 (2005).
24. D. Baillat, E. J. Wagner, Integrator: Surprisingly diverse functions in gene expression. *Trends Biochem. Sci.* **40**, 257–264 (2015).
25. B. Stadelmayer *et al*., Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat. Commun.* **5**, 5531 (2014).
26. F. Lai, A. Gardini, A. Zhang, R. Shiekhattar, Integrator mediates the biogenesis of enhancer RNAs. *Nature* **525**, 399–403 (2015).
27. P. Preker *et al*., RNA exosome depletion reveals transcription upstream of active human promoters. *Science* **322**, 1851–1854 (2008).
28. C. Kilchert, S. Wittmann, L. Vasiljeva, The regulation and functions of the nuclear RNA exosome complex. *Nat. Rev. Mol. Cell Biol.* **17**, 227–239 (2016).
29. C. D. Kaplan, L. Laprade, F. Winston, Transcription elongation factors repress transcription initiation from cryptic sites. *Science* **301**, 1096–1099 (2003).
30. C. M. DeGennaro *et al*., Spt6 regulates intragenic and antisense transcription, nucleosome positioning, and histone modifications genome-wide in fission yeast. *Mol. Cell. Biol.* **33**, 4779–4792 (2013).
31. N. J. Krogan *et al*., RNA polymerase II elongation factors of Saccharomyces cerevisiae: A targeted proteomics approach. *Mol. Cell. Biol.* **22**, 6979–6992 (2002).
32. S. L. Squazzo *et al*., The Paf1 complex physically and functionally associates with transcription elongation factors in vivo. *EMBO J.* **21**, 1764–1774 (2002).
33. Y. Liu *et al*., Phosphorylation of the transcription elongation factor Spt5 by yeast Bur1 kinase stimulates recruitment of the PAF complex. *Mol. Cell. Biol.* **29**, 4852–4863 (2009).
34. W. Antosz *et al*., The composition of the Arabidopsis RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors. *Plant Cell* **29**, 854–870 (2017).
35. A. Shetty *et al*., Spt5 plays vital roles in the control of sense and antisense transcription elongation. *Mol. Cell* **66**, 77–88.e5 (2017).
36. C. Baejen *et al*., Genome-wide analysis of RNA polymerase II termination at protein-coding genes. *Mol. Cell* **66**, 38–49.e6 (2017).
37. J. Fitz, T. Neumann, R. Pavri, Regulation of RNA polymerase II processivity by Spt5 is restricted to a narrow window during elongation. *EMBO J.* **37**, e97965 (2018).
38. V. Vethantham *et al*., Dynamic loss of H2B ubiquitylation without corresponding changes in H3K4 trimethylation during myogenic differentiation. *Mol. Cell. Biol.* **32**, 1044–1055 (2012).
39. H. H. Ng, R. M. Xu, Y. Zhang, K. Struhl, Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *J. Biol. Chem.* **277**, 34655–34657 (2002).
40. T. Xiao *et al*., Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Mol. Cell. Biol.* **25**, 637–651 (2005).
41. B. Zhu *et al*., The human PAF complex coordinates transcription with events downstream of RNA synthesis. *Genes Dev.* **19**, 1668–1673 (2005).
42. P. B. Rahl *et al*., c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
43. A. H. Ehrensberger, G. P. Kelly, J. Q. Svejstrup, Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. *Cell* **154**, 713–715 (2013).
44. H. Kwak, N. J. Fuda, L. J. Core, J. T. Lis, Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953 (2013).
45. D. B. Mahat *et al*., Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). *Nat. Protoc.* **11**, 1455–1476 (2016).
46. I. Jonkers, H. Kwak, J. T. Lis, Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* **3**, e02407 (2014).
47. M. B. Ardehali, J. T. Lis, Tracking rates of transcription and splicing in vivo. *Nat. Struct. Mol. Biol.* **16**, 1123–1124 (2009).
48. N. Fong, T. Saldi, R. M. Sheridan, M. A. Cortazar, D. L. Bentley, RNA Pol II dynamics modulate co-transcriptional chromatin modification, CTD phosphorylation, and transcriptional direction. *Mol. Cell* **66**, 546–557.e3 (2017).
49. W. Shao, J. Zeitlinger, Paused RNA polymerase II inhibits new transcriptional initiation. *Nat. Genet.* **49**, 1045–1051 (2017).
50. S. Gressel *et al*., CDK9-dependent RNA polymerase II pausing controls transcription initiation. *eLife* **6**, e29736 (2017).
51. D. A. Gilchrist *et al*., Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143**, 540–551 (2010).
52. X. Liu *et al*., Transcription elongation rate has a tissue-specific impact on alternative cleavage and polyadenylation in *Drosophila melanogaster*. *RNA* **23**, 1807–1816 (2017).
53. N. Fong *et al*., Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. *Mol. Cell* **60**, 256–267 (2015).