Published in final edited form as:
Mol Cell. 2013 November 21; 52(4): . doi:10.1016/j.molcel.2013.10.001.

Stable pausing by RNA polymerase II provides an opportunity to target and integrate regulatory signals

Telmo Henriques[1,*], **Daniel A. Gilchrist**[1,*], **Sergei Nechaev**[1,*,#], **Michael Bern**[1,†], **Ginger W. Muse**[1], **Adam Burkholder**[2], **David C. Fargo**[2], and **Karen Adelman**[1]

[1]Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709, USA

[2]Integrated Bioinformatics, National Institute of Environmental Health Sciences, National Institutes of Health, Research, Triangle Park, NC 27709, USA

SUMMARY

Metazoan gene expression is often regulated after the recruitment of RNA polymerase II (Pol II) to promoters, through the controlled release of promoter-proximally paused Pol II into productive RNA synthesis. Despite the prevalence of paused Pol II, very little is known about the dynamics of these early elongation complexes or the fate of short transcription start site-associated (tss) RNAs they produce. Here, we demonstrate that paused elongation complexes can be remarkably stable, with half-lives exceeding 15 minutes at genes with inefficient pause release. Promoter-proximal termination by Pol II is infrequent and released tssRNAs are targeted for rapid degradation. Further, we provide evidence that the predominant tssRNA species observed are nascent RNAs held within early elongation complexes. We propose that stable pausing of polymerase provides a temporal window of opportunity for recruitment of factors to modulate gene expression and that the nascent tssRNA represents an appealing target for these interactions.

INTRODUCTION

At many metazoan genes, especially those in developmental and stimulus-responsive pathways, transcriptionally engaged Pol II pauses after generating a short, 20–65 nt RNA (Muse et al., 2007; Core et al., 2008; Nechaev et al., 2010; Rahl et al., 2010). The establishment of a paused polymerase involves the generation of an accessible promoter chromatin structure and recruitment of the transcription machinery, likely through the action of one or more DNA-binding transcription factors (Adelman and Lis, 2012). Pol II then initiates RNA synthesis and comes under control of two pause-inducing factors: the Negative Elongation Factor complex (NELF) and DRB-Sensitivity Inducing Factor (DSIF), which inhibit further elongation (Li et al., 2013; Yamaguchi et al., 2012). Release of paused Pol II into productive synthesis is triggered by a different class of transcription factors, exemplified by c-myc and NF-κB (Barboric et al., 2001; Blau et al., 1996; Eberhardy and

Corresponding author: Karen Adelman, adelmank@niehs.nih.gov.
*These authors contributed equally to this work
#Current address: University of North Dakota School of Medicine, Grand Forks, ND, 58202, USA
†Current address: Washington University School of Medicine, St. Louis, MO 63110, USA

ACCESSION NUMBERS

Genomic data presented in this work were deposited in GEO under accession number GSE49078.

Farnham, 2002; Rahl et al., 2010). These factors recruit the kinase Positive Transcription Elongation Factor-b (P-TEFb; Cheng and Price, 2007; Peterlin and Price, 2006), which phosphorylates the Pol II C-terminal domain as well as pause-inducing factors to dissociate NELF and stimulate productive elongation.

The existence of several independent regulatory steps in the transcription cycle that are controlled by distinct transcription factors has been suggested to enable the integrated control of gene expression, with activators that stimulate Pol II recruitment working in combination with factors that mediate pause release (Adelman and Lis, 2012; Blau et al., 1996). However, our models for such coordinated control are currently limited by our lack of knowledge about the lifetime and dynamics of promoter Pol II. For example, if the paused elongation complex were stable, it could facilitate the integration of signals and transcription factor binding events over time. In this model, the transient binding of a transcription factor that promoted recruitment and initiation of Pol II would have a long-lived effect, with the stably paused polymerase serving as a lasting consequence of the binding event. The subsequent binding of a transcription factor that mediated pause release would then be sufficient to trigger productive RNA synthesis.

As appealing as this model may be, recent work suggests that promoter-associated Pol II is unstable and susceptible to premature termination (also called abortive elongation or transcription attenuation), with Pol II undergoing many iterative cycles of: initiation, pausing and termination before proceeding productively into the gene (Brannan et al., 2012; Wagschal et al., 2012). This transcriptional ‘idling’ would present a very different regulatory framework, as it would require constant re-initiation of transcription and the continuous, simultaneous presence of multiple transcription factors to promote both initiation and productive elongation. Furthermore, iterative rounds of promoter-proximal termination would lead to the generation and release of many short tssRNA species. Such RNAs have been detected in multiple systems (Fejes-Toth, 2009; Flynn et al., 2011; Preker et al., 2008; Taft et al., 2009; Valen et al., 2011; Yus et al., 2012) and in association with a number of different epigenetic and regulatory factors (Brockdorff, 2013; Kanhere et al., 2010) leading to a great deal of interest in their biogenesis and potential functions.

Direct investigation of RNAs derived from promoter-associated Pol II is ideal for answering questions concerning Pol II dynamics and the levels of transcription termination. This strategy overcomes the limitations of assays such as ChIP-seq or Global Run-on assays (e.g. GRO-seq; Core et al., 2008) that reflect the steady-state occupancy of Pol II on the genome, but reveal little about Pol II turnover. We recently developed a highly-sensitive method for isolating the short, capped RNA species (scRNAs) generated by paused polymerase and identified their origins genome-wide using high-throughput sequencing (Nechaev et al., 2010). Notably, the scRNA-seq method gives information about the 5’ and 3’-ends of the scRNAs with single nucleotide resolution, allowing us to pinpoint both the sites of transcription initiation and the exact locations within the promoter-proximal region where Pol II pauses. Thus, the scRNA-seq method is well suited for investigation of the abundance, origin and stability of RNA species originating near transcription start sites, as well as elucidating their relationship with paused Pol II.

Herein, we characterize the dynamics of promoter-associated Pol II and the short RNAs generated during early elongation. We find that paused Pol II is extremely stable and that its release from the promoter occurs predominantly through the transition to productive synthesis. Accordingly, pausing could indeed enable a temporal integration of transcription factor binding events and other regulatory signals. Further, our data indicate that tssRNAs are long-lived when associated with paused Pol II, but are marked for rapid degradation when released through premature termination. Thus, as has recently been suggested (Ji et al.,

2013), epigenetic modifiers and transcription regulators that interact with tssRNAs may do so *in cis*, while the RNA is stably held within a paused elongation complex.

RESULTS

scRNAs generated by paused Pol II remain stably associated with chromatin

Genomic analyses of scRNAs purified from whole *Drosophila* S2 cells (Figure 1A) and isolated nuclei (Nechaev et al., 2010) revealed a strong correlation between the level of Pol II ChIP-seq signal near gene promoters and the number of scRNA reads observed. These findings suggest that promoter-proximal Pol II detected in ChIP assays is often engaged in transcription and generates short, nascent transcripts (Nechaev et al., 2010). However, our previous work did not define whether the scRNAs observed remained stably associated with paused Pol II on the DNA template, or if these RNA species were released from polymerase during termination. To address this question, we probed the cellular localization of scRNAs by performing a biochemical fractionation procedure that separated RNAs that were tightly bound to chromatin from RNA species that could be released during successive washes (Wysocka et al., 2001). It is known that productive transcription elongation complexes are very stable, withstanding wash conditions that remove nearly all other non-histone proteins (Figure S1A; Wysocka et al., 2001).

Chromatin-associated and detergent-soluble scRNA species were subjected to high-throughput sequencing from their 3’-ends and compared with scRNAs derived from whole cells. Chromatin-associated scRNA sequences mapped efficiently to the genome with a marked enrichment around TSSs. In contrast, far fewer of the soluble scRNA species were uniquely mappable or derived from known sites of transcription initiation (Supplemental Table S1), suggesting that the biogenesis of these two sets of scRNAs were different. As anticipated, short nuclear RNA species such as snRNAs and snoRNAs were significantly enriched in the soluble RNA fraction (Figure S1B; $P<0.001$ Mann-Whitney).

Normalized heat map representations of scRNA 3’-ends revealed that genes enriched in promoter Pol II signal generate scRNAs that are predominant in the chromatin fraction (Figure 1A; Figure 1B, blue lines). There was a strong correlation between the number of chromatin-associated scRNA-seq reads from a given promoter and the Pol II ChIP-seq and GRO-seq signals from that region (Figure S1D), suggesting that these scRNAs remain associated with engaged Pol II. Permanganate footprinting at several genes indicated no loss of reactivity associated with paused Pol II in the washed chromatin fraction, supporting the idea that paused polymerase remained engaged on the genome during our wash steps (Figure S1C). Moreover, chromatin-associated scRNAs exhibited a focused 3’-end distribution within the region where promoter-proximal pausing occurs (Figure 1B; 20 to 65 nt downstream of the TSS; Nechaev et al., 2010). In contrast, detergent-soluble scRNAs were far less enriched at genes with high promoter-proximal Pol II levels (Figure 1C; Figure S1E) and displayed more dispersed 3’-end positions (Figure S1F) that extended beyond the typical pausing region and into gene bodies (Supplemental Table S1B).

These findings suggested that the majority of short transcripts arising from highly paused genes remain tightly associated with chromatin, and this idea was strongly supported by Northern blot analysis (e.g. Figure S1G, *Hsp70* and *CG9008*). Moreover, in agreement with earlier work (Rasmussen and Lis, 1993), Northern blots demonstrated that the short RNA species arising from two paused genes were predominantly capped at their 5’-ends. Thus, we find that scRNAs stably retained on chromatin are likely to be associated with or derived from paused Pol II, whereas RNAs in the soluble fraction have a distinct origin.

To determine whether the soluble scRNAs arise as degradation intermediates of longer RNA species such as mature mRNAs, we generated heatmaps of scRNA 3'-ends at *Drosophila* genes ranked from highest to lowest mRNA expression (Figure 1D). This analysis revealed that the majority of soluble scRNAs are derived from abundant transcripts, with more than 50% of all soluble scRNA reads mapping near the most highly expressed 3% of TSSs (Figure 1D, arrow head depicts 50% point). In contrast, only 9% of chromatin-associated reads come from this group of highly expressed genes. Thus, these data support a model wherein soluble scRNA species arise through degradation of longer mRNAs rather than being generated by termination of Pol II.

Exosome depletion does not lead to accumulation of scRNAs

The dearth of soluble short RNAs derived from genes with high levels of promoter Pol II raised the possibility that scRNAs released during premature termination may be rapidly degraded. To test this, we sequenced scRNAs from samples depleted of the primary 3'-5' degradation machinery, the exosome, by targeting the Rrp40 subunit with RNA interference (RNAi) (Figure S2A; Chlebowski et al., 2013; Kiss and Andrulis, 2010). To allow accurate normalization between exosome-depleted and Mock RNAi-treated samples, each cell pellet was spiked with 15 *in vitro*-synthesized and capped RNAs (see Supplemental Experimental Procedures). Notably, for both replicates of this experiment, normalization of the Mock-treated and Rrp40-depleted scRNA libraries by sequencing depth revealed nearly identical read counts for the spiked RNAs (median spike ratio =1.05), indicating that exosome depletion did not grossly alter the cellular pool of scRNAs.

Heat maps of scRNA 3'-ends from Mock-treated and Rrp40-depleted samples revealed nearly identical distributions of scRNAs (Figure 2A) and no significant differences were detected between samples in the number of scRNAs mapping near TSSs (Figure S2B, S2C). Moreover, Northern blot analysis of short transcripts generated by paused Pol II at the *Hsp70* gene showed negligible effects of Rrp40-depletion on their level, length or cellular localization (Figure 2B). Thus, the exosome degradation pathway does not appear to affect scRNA abundance, suggesting that these species are protected from 3'-to-5' exonuclease digestion through stable association with paused Pol II.

We then considered the possibility that scRNAs released from Pol II during termination are processed through a distinct pathway. For example, the addition of several non-templated 3'-end residues to structured or short RNA species has been shown to promote their degradation (e.g. oligo-adenylation or oligo-uridylation; LaCava et al., 2005; Mullen and Marzluff, 2008). Notably, since scRNAs in this study were sequenced from their 3'-ends, reads containing non-templated 3'-residues would have failed to align in our original analysis. Thus, we reanalyzed the raw sequence reads to determine whether the 3'-ends of scRNAs from mock-treated or Rrp40-depleted samples consisted of stretches of any given nucleotide. Strikingly, we found a specific enrichment in 3'-oligo-adenylated species in the exosome-depleted samples (Figure 2C; Figure S2D, S2E, note lack of detectable oligo-uridylation). The presence of a short stretch of 3'-A nucleotides suggested the involvement of an oligo-adenylase such as the TRAMP complex, which targets RNAs for degradation by the exosome through addition of 3–6 A residues (Schmidt and Butler, 2013). Taken together, these data imply that released scRNAs are marked for efficient degradation through oligo-adenylation.

To determine the origin of oligo-adenylated scRNAs, we removed the non-templated oligo-A tails from these sequences *in silico* and re-mapped the trimmed reads against the *Drosophila* genome. Remarkably, 40% of the re-mapped reads from Rrp40-depleted cells aligned within 100 nt of TSSs (as compared to 5% from Mock-treated cells), with an enrichment at genes with high levels of promoter Pol II (Figure 2D; Figure S2G). The 3'-

ends of the trimmed reads from Rrp40-depleted samples overlapped precisely with the 3'-ends from scRNA reads that were mapped directly (Figure 2E), suggesting that oligo-adenylated species were derived from scRNAs released by promoter-proximal termination.

Termination appeared to be a rare event, however, as the percentage of scRNAs that were oligo-adenylated was a consistently low level (~1%) across genes regardless of Pol II promoter signal (Figure 2F). Further, the absence of efficient pause release was not found to encourage termination as it does in bacterial systems (von Hippel and Yager, 1991): to the contrary, genes with low levels of productive elongation generated significantly fewer oligo-adenylated RNAs than highly active genes (Figure 2G; $P<0.0001$ Kruskal-Wallis; gene activity measured by GRO-seq signal within the gene body). This suggests that promoter-paused Pol II is stable while awaiting the transition to productive elongation. To validate this finding, Rrp40-depleted cells were treated with Flavopiridol, a potent inhibitor of P-TEFb kinase that prevents Pol II release into productive elongation (Chao et al., 2000 and see below). No significant increase in termination was detected in these cells as measured by the fraction of oligo-adenylated species (Figure S2F). Thus, these results support the idea that paused elongation complexes are stabilized against termination, even in the absence of productive pause release (Cheng et al., 2012).

Quantification reveals abundant, stable tssRNAs

Thus far, our genomic data provide evidence for a generally stable promoter Pol II complex. Moreover, when paused polymerase does terminate, the released RNAs appear to be specifically marked for efficient exosome degradation through oligo-adenylation (Figure 3A, scRNA reads at example gene). To confirm these findings using an independent technique, we developed a qPCR strategy (Figure 3B) based on methods for quantifying short RNA species such as miRNAs. This technique amplifies short RNAs independently of the 5'-end cap, enabling an evaluation of all RNAs under 200 nt within the cell. For this reason, we refer to the RNA species measured by this technique as short transcription start site associated RNAs, or tssRNAs. The qPCR approach also allowed independent detection of tssRNAs that are oligo-adenylated *in vivo*, since these species are specifically amplified when omitting the poly(A) tailing step prior to anchored oligo-dT reverse priming (right side of diagram, Figure 3B). The results from this assay performed on multiple genes demonstrated that scRNA-seq data (Figure 3C; Figure S3) were in remarkable agreement with qPCR-based analysis of tssRNAs (Figure 3D), providing little evidence for tssRNA accumulation in exosome-depleted cells yet supporting the presence of a low level of oligo-adenylated RNA species.

To address whether paused elongation complexes might be targeted by 5'-3' exonuclease activity, we evaluated levels of paused Pol II in S2 cells depleted of the nuclear 5'-3' exoribonuclease, Xrn2 (Figure S3C, S3D). Xrn2 has been previously implicated in premature termination at the HIV gene as well as near mammalian promoters (Brannan et al., 2012; Wagschal et al., 2012). However, it has been suggested that the primary target of Xrn2 activity near mammalian TSSs might be the upstream antisense RNAs common at many promoters (Reines, 2011). Working in *Drosophila*, which does not exhibit antisense transcription upstream of mRNA genes (Core et al., 2012; Nechaev et al., 2010), we did not observe an effect of Xrn2 knock-down on Pol II ChIP signal either promoter-proximally or downstream within active genes (e.g. Figure S3D), in contrast to data from human cells (Brannan et al., 2012). Further, we did not observe an effect of depleting Xrn2, alone or in combination with the decapping enzyme Dcp2, on levels of tssRNAs (Figure S3C). These results suggest that in the absence of antisense transcription, the exonuclease activity of Xrn2 and associated termination complex is not broadly targeted to *Drosophila* mRNA loci.

Importantly, a qPCR-based measurement of tssRNA levels enables quantification of the number of molecules of short RNA present per genomic locus *in vivo*, which has thus far remained elusive using sequencing techniques. In agreement with prior studies, our Northern blots of *Hsp70* (Figure 2B) revealed that paused Pol II and associated nascent RNA are present at nearly one molecule per gene copy, suggesting nearly complete Pol II occupancy of *Hsp70* promoters; however, this type of quantitative analysis has not yet been carried out at other genes. We find that levels of tssRNA species differ considerably from gene to gene (Figure 3D; Figure S3), with several promoters (e.g. *CG9884* or *CG2207*) displaying similar RNA abundance to *Hsp70*. Notably, this suggests that a number of paused genes can attain a relative Pol II occupancy of >50%, implying that an early elongation complex and associated tssRNA would occupy at least one allele of the gene at any given moment.

Directly measuring turnover of promoter Pol II complexes reveals remarkable stability

To directly investigate the temporal stability of promoter Pol II and tssRNAs, we used the potent inhibitor of transcription initiation, Triptolide (Trp). This chemical rapidly blocks the ATPase activity of the TFIIH helicase, thereby preventing the opening of template DNA (Titov et al., 2011; Vispe et al., 2009). Treatment of cells with Trp thus allowed us to measure the lifetime of Pol II elongation complexes present at the moment of drug treatment, analogous to a ‘chase’ or ‘decay’ experiment. Notably, Trp-dependent inhibition of transcription initiation should not affect the release of an engaged elongation complex into productive synthesis nor would it impede termination by Pol II. As such, the decay of Pol II ChIP and tssRNA signals at a given gene following Trp-treatment informs on the rate at which Pol II vacates the promoter region, either through escape into productive synthesis or through termination.

Following treatment with Trp or DMSO as a control, cells were harvested at multiple time points for short RNA qPCR and Pol II ChIP assays. In agreement with prior studies demonstrating the fast action of Trp, the tssRNA (Figure 4A) and Pol II ChIP signals (Figure 4B) at the *CG4427* gene (top row) showed a rapid decline upon Trp-treatment. The *CG4427* gene is highly active (Gene body GRO-seq signal=1638 RPKM), suggesting that the rapid loss of promoter-associated Pol II and short tssRNA signals results from efficient release of polymerase into the gene body and elongation of the nascent RNA. However, as noted above, the loss of promoter Pol II signal might also reflect termination by promoter Pol II, wherein the polymerase dissociates from the DNA and the tssRNA is rapidly degraded. To distinguish between these possibilities, we investigated the half-life of Pol II and RNA at genes across a spectrum of activity levels.

Strikingly, the loss of tssRNA and Pol II signals at the moderately active *CG9884* gene (middle row, Gene body GRO-seq signal= 345) was much slower than that observed at *CG4427,* with a half-life just under 10 minutes. Likewise, the minimally active *CG33174* gene (bottom row, Gene body GRO-seq signal=89) displayed a very stable elongation complex (T½ ~ 15 minutes). The longer lifetime of promoter Pol II and tssRNA at genes with lower activity levels suggested that in the absence of a signal to undergo productive synthesis, the elongation complex remains intact and stably associated with the promoter region.

Indeed, analysis of over a dozen genes revealed consistently longer half-lives for tssRNA and promoter Pol II signals at genes with low levels of activity (Figure 4C; Figure S4). This relationship suggests that the primary source of decay in tssRNA and Pol II ChIP signals following Trp-treatment is release into productive elongation and not termination. Moreover, the tight agreement between the lifetimes observed for promoter Pol II and tssRNAs at each gene (Figure 4, Figure S4) provides strong evidence that these RNA species remain associated within the early elongation complex. Thus, we find that the

paused polymerase can remain engaged and 'wait' for the signal for pause release, achieving half-lives of 15–20 minutes at genes with low activity.

The dramatic decrease of Pol II ChIP signals in Trp-treated cells is somewhat surprising, since ChIP assays should capture all forms of Pol II and Trp-treament does not prevent the formation of a Pre-initiation complex (PIC). However, we reasoned that this may reflect the low stability of the closed polymerase complex which may not crosslink efficiently to DNA. To investigate the status of the PIC at these genes, we performed ChIP on the general transcription factor TFIIA after 10 minutes of Trp-treatment. Interestingly, TFIIA levels do not decrease at gene promoters (Figure S4C) in Trp-treated cells, and are even found to increase slightly, confirming that Trp-treatment does not prevent assembly of the general transcription machinery (Titov et al., 2011).

Blocking pause release leads to accumulation of promoter Pol II

The above results imply that paused Pol II can remain stably engaged in the promoter region in the absence of a signal to elongate productively into the gene. The prediction of this model is that inhibition of pause release through use of the P-TEFb inhibitor Flavopiridol (FP) should cause an increase in promoter Pol II, including at genes that previously did not accumulate paused polymerase. To test this model, we treated cells with FP for 10 minutes to inhibit Pol II escape from promoters (Figure S5A) prior to harvesting cells for scRNA-seq. As described earlier, each cell pellet was spiked with a defined set of RNAs to enable accurate normalization between samples (see Extended Experimental Procedures).

In agreement with our model, heat map representations of scRNA 3'-ends depict a broad increase in RNA abundance including the appearance of scRNAs from genes that display low levels of promoter Pol II and scRNA reads in untreated cells (Figure 5A, note bottom portion of heat map, Figure S5B). Nearly identical results were obtained when scRNAs were isolated from FP-treated cells that had been Mock-treated with RNAi or depleted of Rrp40 (Figure S5C and S5D). The absence of exosome activity on these FP-dependent scRNA species suggests that they are protected from degradation, likely through continued association with the polymerase. Composite metagene analyses of the average scRNA-seq signal across all Pol II-occupied genes reveals a considerable increase in reads that is restricted to the region of pausing (Figure 5B). This result demonstrates that inhibition of P-TEFb effectively prevents downstream movement of paused polymerase, in agreement with a recent study (Cheng et al., 2012).

To determine which genes were particularly susceptible to accumulation of scRNAs in FP-treated cells, we separated genes into quartiles based upon the level of Pol II pausing observed at their promoters in untreated cells (horizontal quartiles, Figure 5C), and by the level of gene activity as determined by GRO-seq signal within gene bodies (vertical quartiles, Figure 5C). This created a 4 × 4 matrix where the most paused, least expressed genes are at top right. These genes would be anticipated to infrequently experience P-TEFb-mediated pause release in untreated cells. In agreement with this, they are largely unaffected by a brief FP-treatment. In contrast, genes in the bottom left corner are highly active genes that typically don't exhibit stable pausing of Pol II: these promoters presumably recruit P-TEFb efficiently so that promoter Pol II is rapidly released into productive elongation. Notably, it is this group of genes that display the strongest accumulation of scRNAs in FP-treated cells (Figure 5C; S5E and S5F). This finding suggests that by preventing P-TEFb-mediated pause release, a paused elongation complex can be trapped near the promoters of many genes, including those that normally don't exhibit characteristics of stable pausing.

To verify that the observed increase in scRNAs reflected an accumulation of paused Pol II near gene promoters, Pol II ChIP was performed in FP-treated cells. Figure 5D shows a

significant increase in the promoter Pol II ChIP signal at a number of genes with low pausing indices and high gene activity, indicative of polymerase being retained near these promoters. By comparison, highly paused genes (Figure 5D, right; genes from top row of Figure 5C), exhibited very little change in promoter Pol II levels upon FP-treatment. In summary, blocking P-TEFb activity causes an exosome-independent increase in scRNAs and accumulation of Pol II at genes that normally undergo efficient productive elongation, arguing that inhibiting pause release causes even non-paused promoters to retain promoter Pol II.

Disruption of NELF-mediated pausing destabilizes promoter-associated Pol II

The surprisingly long lifetime of paused Pol II suggests that the factors that establish pausing may also play a role in stabilization of the early elongation complexes. The Gdown1 protein has recently been shown to associate with Pol II near promoters, rendering the polymerase resistant to termination by the TTF2 protein (Cheng et al., 2012). However, it is unclear whether the NELF and DSIF factors that establish paused Pol II also play a role in maintenance of stably paused complexes. Notably, depletion of NELF in *Drosophila* cells has been shown to significantly affect gene expression and Pol II occupancy at only a subset of ~250 genes (Gilchrist et al., 2010; Gilchrist et al., 2008). The remainder of genes displayed modest decreases of Pol II near their promoters (on average ~25%), with no concomitant increases in gene body transcription as determined by expression microarray or GRO-seq assays (Core et al., 2012; Gilchrist et al., 2010). Thus, Pol II recruited to promoters in the absence of NELF does not appear to freely enter productive synthesis.

We investigated the role of NELF in defining the stability of promoter Pol II, by performing Pol II ChIP in NELF-depleted cells following the addition of Trp. Pol II ChIP samples at each time point were evaluated by qPCR using promoter-proximal primer sets (Figure 6A), focusing on genes with considerable promoter Pol II ChIP signal in untreated cells that displayed minimal reduction in promoter Pol II signal upon NELF-depletion (Gilchrist et al., 2010). As anticipated, Pol II ChIP signal at the highly active *CG4427* gene decreased rapidly upon addition of Trp in Mock-RNAi-treated cells (Figure 6A, top left) and this fast loss of Pol II from the promoter was not significantly altered by NELF-depletion. In contrast, genes with longer promoter Pol II half-lives (as shown in Figure 4) displayed a consistently faster decay of Pol II ChIP signal upon NELF-depletion (Figure 6A).

To probe the relationship between the half-life of promoter Pol II and the effect of NELF-depletion at a given gene, we extended the Pol II ChIP analysis to a total of 12 genes selected from the top quartile of pausing indices (Figure 6 and S6) with a wide range of gene activity. Consistent with a rapid transition from paused to productive elongation complexes, highly active genes have a shorter average residence time of Pol II near the promoter that is not dependent on the presence of NELF (Figure 6B, High gene activity). However, genes with intermediate and lower activity levels have more stably paused Pol II in Mock-treated cells, and exhibit significantly reduced half-lives for promoter Pol II upon depletion of NELF (Figure 6B). Thus, long-lived pausing near promoters is broadly dependent on the activity of NELF.

That NELF was found to impact the stability of promoter Pol II, even at genes where NELF-depletion caused little decrease in total Pol II ChIP signal is striking. This implies that, despite giving similar Pol II ChIP signal in untreated and NELF-depleted cells, the fundamental stability and rate of turnover of promoter-associated Pol II in the absence of NELF is markedly different. This finding extends the well-defined role of NELF in inhibiting early transcription elongation and establishing a paused complex (Li et al., 2013; Yamaguchi et al., 2012), and suggests that NELF likely works in conjunction with Gdown1 and other factors to maintain paused Pol II stably associated with the promoter region.

DISCUSSION

Our work reveals that the paused elongation complex displays remarkable stability near promoters. This finding provides a framework for understanding how multiple transcription factors can exert combinatorial control of gene expression. We propose that the transient binding of a transcription factor that facilitates recruitment and transcription initiation can be 'remembered' through the stable establishment of paused Pol II (Figure 7A). Pausing would then provide an extended window of time for binding of a second transcription factor that recruits P-TEFb and triggers productive elongation. In this way, information from multiple, temporally separate signaling and/or transcription factor binding events could be integrated through paused Pol II to cooperatively regulate transcription output. Thus, the sophisticated, multi-layered nature of signal-responsive gene expression can be reconciled with the rapid, transient dynamics of transcription factor binding.

Consistent with the stable pausing of Pol II, recent work has shown that Pol II near mRNA promoters is often associated with the Gdown1 factor, which renders the early elongation complex resistant to termination (Cheng et al., 2012). Further, work in yeast demonstrates that a Pol II CTD modification that occurs near gene promoters, Tyrosine-1 phosphorylation, impairs the recruitment of termination factors (Mayer et al., 2012). Accordingly, our data provides evidence that termination near promoters of *Drosophila* protein coding genes is rare, fitting with suggestions that termination serves primarily to impede spurious transcription and regulate production of non-coding RNAs (Almada et al., 2013; Reines, 2011). Indeed, the antisense RNAs arising from divergent transcription at mammalian promoters are subject to termination and exosome-mediated degradation (Almada et al., 2013; Flynn et al., 2011).

Our findings also clarify the function of NELF, which is found at most, if not all, gene promoters (Gilchrist et al., 2010; Rahl et al., 2010) but which regulates the expression of only a subset of genes (Gilchrist et al., 2008; Narita et al., 2007). We demonstrate that NELF increases the stability of promoter Pol II, even at genes with similar Pol II ChIP signals in Mock-treated and NELF-depleted cells (Figure 6; Figure 7B). The more rapid loss of Pol II from promoters observed upon NELF RNAi does not lead to higher levels of productive elongation at most genes (Core et al., 2012), suggesting that polymerase that fails to pause does not proceed into processive RNA synthesis.

Consistent with this, P-TEFb activity is thought to be required for both pause release and productive elongation (Peterlin and Price, 2006). Notably, we find that inhibition of P-TEFb by FP causes accumulation of paused Pol II at a broad array of genes, indicating that the key determinant of Pol II lifetime in the promoter region is the rate of P-TEFb recruitment. In agreement with this, the level of pausing observed at a given gene is not hard-wired, but can be tuned in a context- or signal-dependent manner (Danko et al., 2013; Gilchrist et al., 2012).

Finally, our study has important implications concerning the function and localization of RNAs derived from promoter regions. Our work reveals that tssRNAs arising from highly paused genes in *Drosophila* are likely stably retained within the elongation complex and are not released as long-lived species. We note that these tssRNAs would have a very well defined origin and remain localized *in cis* as a stable landmark near the promoter (Figure 7C). Importantly, a nascent RNA-targeting mechanism has long been appreciated in the HIV system (Wei et al., 1998), and evidence suggesting recruitment of regulatory factors by nascent transcripts of cellular genes is beginning to emerge (Ji et al., 2013). Thus, we propose that nascent tssRNAs associated stably with paused Pol II provide a locus-specific

target for recruitment of chromatin modifying or transcription elongation factors that modulate gene expression.

EXPERIMENTAL PROCEDURES

Cell growth, RNAi and drug treatments

Drosophila S2 cells from the DGRC were grown in M3 media supplemented with bactopeptone, yeast extract and 10% FBS. For all experiments, cells were harvested at a consistent cell density of 4–6×10^6 cells/ml. NELF RNAi was performed for 72 hours as described previously (Gilchrist et al., 2008) and Rrp40 RNAi employed the same protocol. To measure the half-lives of promoter Pol II and associated short-RNA, S2 cells were treated with 10μM of Trp (InvivoGen) for the indicated times before harvesting for ChIP or RNA analysis. We note that the use of 10μM Trp for the short time periods employed here has no effect on levels of Pol II protein as determined by Western (data not shown). To block P-TEFb-mediated pause release, cells were treated with 500 nM Flavopiridol (Sigma) for 10 minutes.

scRNA-seq

Total RNA was extracted from nuclei, whole cells, or fractionated cells as indicated using Trizol reagent (Invitrogen) and libraries prepared for 3’ scRNA-sequencing as described in (Nechaev et al., 2010). Cell fractionation was performed as described previously (Wysocka et al., 2001) with modifications noted in Extended Experimental Procedures.

scRNA reads were trimmed to 26 nt and uniquely aligned to the *Drosophila* genome with Bowtie allowing 2 mismatches. To account for sequencing depth, all scRNA samples were normalized to 15 million mappable reads (Supplemental Table S1). Additionally, for data shown in Figures 2 and 5, 15 synthetic short capped RNAs were spiked into the Trizol RNA preparation at a specific number of copies per genome to ensure proper normalization between samples.

Short-RNA qPCR

For all short RNA quantification, cells were harvested at the time points indicated, spike-in controls added for normalization and short RNA isolated using the Qiagen miRNeasy Mini Kit. For Figure 3, total short RNA was quantified by Oligo-A tailing 200 ng of short RNA *in vitro* using 1.5U of E. coli Poly(A) Polymerase (NEB) for 20min at 37°C. This step was omitted to quantify short RNAs that were oligo-adenylated *in vivo*. Samples were reverse transcribed with an anchored oligo-dT primer (IDT) using SuperScript II Reverse Transcriptase (Invitrogen) according to the manufacturer’s instructions. To quantify all short tssRNAs (as in Figure 4 and 6), 200 ng RNA isolated as above was reverse transcribed using the Qiagen miScript Reverse Transcription Kit according to the manufacturer’s instructions. All qPCR experiments were performed using a universal reverse primer together with a forward primer complementary to the 5’-end of nascent RNA.

ChIP

Pol II ChIP was performed using the anti-Rpb3 antibody as previously described (Muse et al., 2007), except that cells treated with Trp, FP or DMSO for 10 minutes were then crosslinked with formaldehyde for 12 minutes. For Trp experiments, the percent of Pol II ChIP signal present at each time point was calculated relative to the value prior to addition of the drug. The percent of Pol II ChIP signal in FP treated samples was calculated relative to control (DMSO) values.

Half-life determination

The half-lives for both short RNA qPCR and Pol II ChIP qPCR signals (Figure 4 and 6) during the Trp-treatment time course were derived as linear fits of $\log_{10}$ transformed values. For easier visualization, the data shown are not log transformed. Thus, the lines shown are derived from a fit to single exponential decay, as this is the simplest model that can describe the data.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank Erik Andrulis (Case Western Reserve) for the kind gift of the Rrp40 antibody, Michael Marr (Brandeis) for the TFIIA antibody and the *Drosophila* Genomics Resource Center, supported by NIH grant OD010949-10 for S2 cells. This research was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences to K.A. (Z01 ES101987).

References

Adelman K, Lis JT. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. Nat Rev Genet. 2012; 13:720–731. [PubMed: 22986266]

Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature. 2013

Barboric M, Nissen RM, Kanazawa S, Jabrane-Ferrat N, Peterlin BM. NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. Mol Cell. 2001; 8:327–337. [PubMed: 11545735]

Blau J, Xiao H, McCracken S, O'Hare P, Greenblatt J, Bentley D. Three functional classes of transcriptional activation domain. Mol Cell Biol. 1996; 16:2044–2055. [PubMed: 8628270]

Brannan K, Kim H, Erickson B, Glover-Cutter K, Kim S, Fong N, Kiemele L, Hansen K, Davis R, Lykke-Andersen J, Bentley DL. mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. Mol Cell. 2012; 46:311–324. [PubMed: 22483619]

Brockdorff N. Noncoding RNA and Polycomb recruitment. Rna. 2013; 19:429–442. [PubMed: 23431328]

Chao SH, Fujinaga K, Marion JE, Taube R, Sausville EA, Senderowicz AM, Peterlin BM, Price DH. Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. J Biol Chem. 2000; 275:28345–28348. [PubMed: 10906320]

Cheng B, Li T, Rahl PB, Adamson TE, Loudas NB, Guo J, Varzavand K, Cooper JJ, Hu X, Gnatt A, et al. Functional Association of Gdown1 with RNA Polymerase II Poised on Human Genes. Mol Cell. 2012; 45:38–50. [PubMed: 22244331]

Cheng B, Price DH. Properties of RNA Polymerase II Elongation Complexes Before and After the P-TEFb-mediated Transition into Productive Elongation. J Biol Chem. 2007; 282:21901–21912. [PubMed: 17548348]

Chlebowski A, Lubas M, Jensen TH, Dziembowski A. RNA decay machines: The exosome. Biochimica et biophysica acta. 2013; 1829:552–560. [PubMed: 23352926]

Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT. Defining the status of RNA polymerase at promoters. Cell reports. 2012; 2:1025–1035. [PubMed: 23062713]

Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science. 2008; 322:1845–1848. [PubMed: 19056941]

Danko CG, Hah N, Luo X, Martins AL, Core L, Lis JT, Siepel A, Kraus WL. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. Mol Cell. 2013; 50:212–222. [PubMed: 23523369]

Eberhardy S, Farnham P. Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. J Biol Chem. 2002; 277:40156–40162. [PubMed: 12177005]

Fejes-Toth K. Post-transcriptional processing generates a diversity of 5'-modified long and short RNAs. Nature. 2009; 457:1028–1032. [PubMed: 19169241]

Flynn RA, Almada AE, Zamudio JR, Sharp PA. Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:10460–10465. [PubMed: 21670248]

Gilchrist DA, Dos Santos G, Fargo DC, Xie B, Gao Y, Li L, Adelman K. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. Cell. 2010; 143:540–551. [PubMed: 21074046]

Gilchrist DA, Fromm G, Dos Santos G, Pham LN, McDaniel IE, Burkholder A, Fargo DC, Adelman K. Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks. Genes & development. 2012; 26:933–944. [PubMed: 22549956]

Gilchrist DA, Nechaev S, Lee C, Ghosh SK, Collins JB, Li L, Gilmour DS, Adelman K. NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. Genes & development. 2008; 22:1921–1933. [PubMed: 18628398]

Ji X, Zhou Y, Pandit S, Huang J, Li H, Lin CY, Xiao R, Burge CB, Fu XD. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell. 2013; 153:855–868. [PubMed: 23663783]

Kanhere A, Viiri K, Araujo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, et al. Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. Mol Cell. 2010; 38:675–688. [PubMed: 20542000]

Kiss DL, Andrulis ED. Genome-wide analysis reveals distinct substrate specificities of Rrp6, Dis3, and core exosome subunits. Rna. 2010; 16:781–791. [PubMed: 20185544]

LaCava J, Houseley J, Saveanu C, Petfalski E, Thompson E, Jacquier A, Tollervey D. RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. Cell. 2005; 121:713–724. [PubMed: 15935758]

Li J, Liu Y, Rhee HS, Ghosh SK, Bai L, Pugh BF, Gilmour DS. Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. Mol Cell. 2013; 50:711–722. [PubMed: 23746353]

Mayer A, Heidemann M, Lidschreiber M, Schreieck A, Sun M, Hintermair C, Kremmer E, Eick D, Cramer P. CTD tyrosine phosphorylation impairs termination factor recruitment to RNA polymerase II. Science. 2012; 336:1723–1725. [PubMed: 22745433]

Mullen TE, Marzluff WF. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. Genes & development. 2008; 22:50–65. [PubMed: 18172165]

Muse GW, Gilchrist DA, Nechaev S, Shah R, Parker JS, Grissom SF, Zeitlinger J, Adelman K. RNA polymerase is poised for activation across the genome. Nat Genet. 2007; 39:1507–1511. [PubMed: 17994021]

Narita T, Yung TM, Yamamoto J, Tsuboi Y, Tanabe H, Tanaka K, Yamaguchi Y, Handa H. NELF interacts with CBC and participates in 3' end processing of replication-dependent histone mRNAs. Mol Cell. 2007; 26:349–365. [PubMed: 17499042]

Nechaev S, Fargo DC, dos Santos G, Liu L, Gao Y, Adelman K. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in Drosophila. Science. 2010; 327:335–338. [PubMed: 20007866]

Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. Mol Cell. 2006; 23:297–305. [PubMed: 16885020]

Preker P, Nielsen J, Kammler S, Lykke-Andersen S, Christensen MS, Mapendano CK, Schierup MH, Jensen TH. RNA exosome depletion reveals transcription upstream of active human promoters. Science. 2008; 322:1851–1854. [PubMed: 19056938]

Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA. c-Myc Regulates Transcriptional Pause Release. Cell. 2010; 141:432–445. [PubMed: 20434984]

Rasmussen EB, Lis JT. In vivo transcriptional pausing and cap formation on three Drosophila heat shock genes. Proceedings of the National Academy of Sciences of the United States of America. 1993; 90:7923–7927. [PubMed: 8367444]

Reines D. Decapping goes nuclear. Mol Cell. 2011; 46:241–242. [PubMed: 22578538]

Schmidt K, Butler JS. Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity. Wiley interdisciplinary reviews. 2013; 4:217–231. [PubMed: 23417976]

Taft R, Glazov E, Cloonan N, Simons C, Stephen S, Faulkner G, Lassmann T, Forrest A, Grimmond S, Schroder K, et al. Tiny RNAs associated with transcription start sites in animals. Nat Genet. 2009; 41:572–578. [PubMed: 19377478]

Titov DV, Gilman B, He QL, Bhat S, Low WK, Dang Y, Smeaton M, Demain AL, Miller PS, Kugel JF, et al. XPB, a subunit of TFIIH, is a target of the natural product triptolide. Nature chemical biology. 2011; 7:182–188.

Valen E, Preker P, Andersen PR, Zhao X, Chen Y, Ender C, Dueck A, Meister G, Sandelin A, Jensen TH. Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes. Nat Struct Mol Biol. 2011; 18:1075–1082. [PubMed: 21822281]

Vispe S, DeVries L, Creancier L, Besse J, Breand S, Hobson DJ, Svejstrup JQ, Annereau JP, Cussac D, Dumontet C, et al. Triptolide is an inhibitor of RNA polymerase I and II-dependent transcription leading predominantly to down-regulation of short-lived mRNA. Molecular cancer therapeutics. 2009; 8:2780–2790. [PubMed: 19808979]

von Hippel PH, Yager TD. Transcript elongation and termination are competitive kinetic processes. Proceedings of the National Academy of Sciences of the United States of America. 1991; 88:2307–2311. [PubMed: 1706521]

Wagschal A, Rousset E, Basavarajaiah P, Contreras X, Harwig A, Laurent-Chabalier S, Nakamura M, Chen X, Zhang K, Meziane O, et al. Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce premature termination of transcription by RNAPII. Cell. 2012; 150:1147–1157. [PubMed: 22980978]

Wei P, Garber ME, Fang SM, Fischer WH, Jones KA. A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell. 1998; 92:451–462. [PubMed: 9491887]

Wysocka J, Reilly PT, Herr W. Loss of HCF-1-chromatin association precedes temperature-induced growth arrest of tsBN67 cells. Mol Cell Biol. 2001; 21:3820–3829. [PubMed: 11340173]

Yamaguchi Y, Shibata H, Handa H. Transcription elongation factors DSIF and NELF: Promoter-proximal pausing and beyond. Biochimica et biophysica acta. 2012

Yus E, Guell M, Vivancos AP, Chen WH, Lluch-Senar M, Delgado J, Gavin AC, Bork P, Serrano L. Transcription start site associated RNAs in bacteria. Molecular systems biology. 2012; 8:585. [PubMed: 22617959]

HIGHLIGHTS

- Pol II pausing near promoters generates stable, long-lived elongation complexes
- Disrupting pausing greatly reduces the lifetime of promoter Pol II
- Blocking pause release leads to widespread accumulation of Pol II at active genes
- Stable transcription-start site associated RNAs remain with Pol II on chromatin

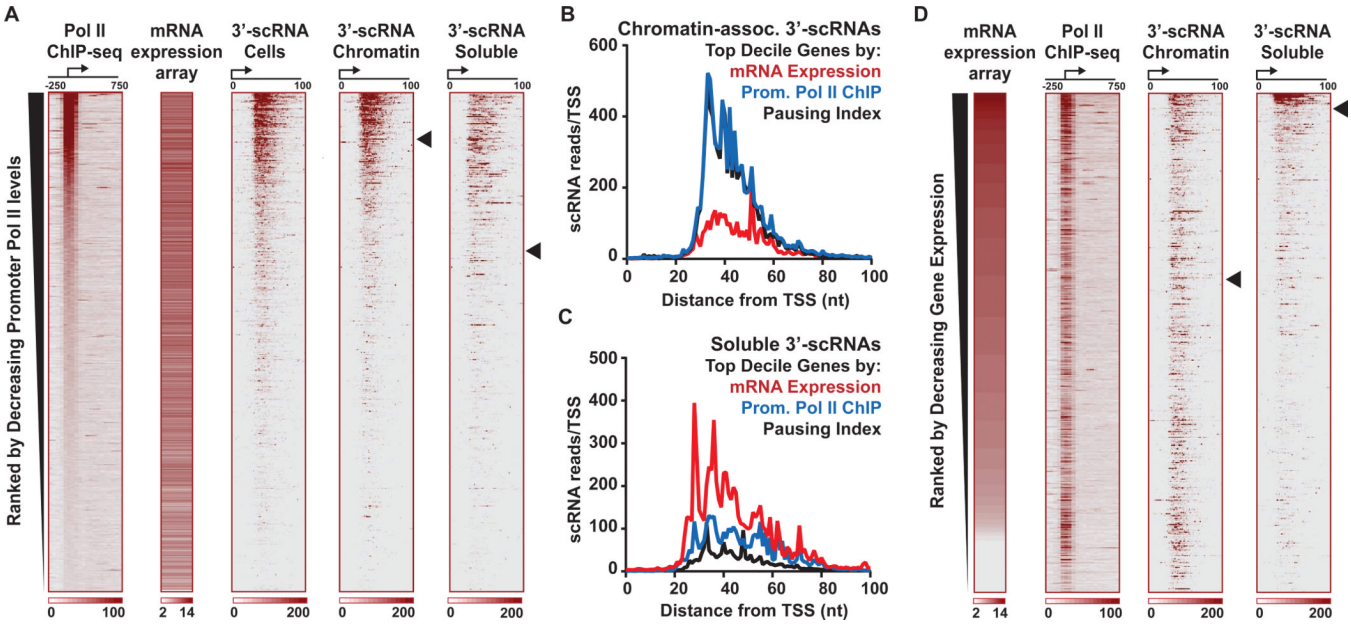

Figure 1. scRNAs co-localize with Pol II and remain stably associated with chromatin

(A) Heat map representations of Pol II ChIP-seq reads, mRNA expression levels determined by microarray and 3'-end sequencing reads of scRNAs isolated from different cellular fractions. Shown are *Drosophila* genes with significant Pol II occupancy (as determined by Pol II ChIP-seq using an α-Rpb3 antibody) and >550 nt in length (N=6,152). Genes are rank ordered by decreasing number of Pol II ChIP-seq reads observed around the transcription start sites (+/− 150 bp from TSS, depicted by arrow). Arrow heads at right of heat maps indicate the position above which 50% of the scRNA-seq reads are located. Color bars at bottom indicate range for each depth-normalized data set.

(B and C) Composite metagene analysis of scRNA reads 3'-ends in the chromatin-associated pellet (B) or detergent-soluble fraction (C). Shown is the average number and position of scRNA 3'-end reads per TSS at the top decile of genes when rank ordered by mRNA expression, promoter Pol II ChIP-seq signal or Pausing Index, as indicated.

(D) Heat maps are shown as in (A) but with genes rank ordered by decreasing mRNA expression levels. Of the Pol II-occupied TSSs shown in (A) 4,954 are represented by unique transcripts on the microarray and are included in this panel.

See also Figure S1.

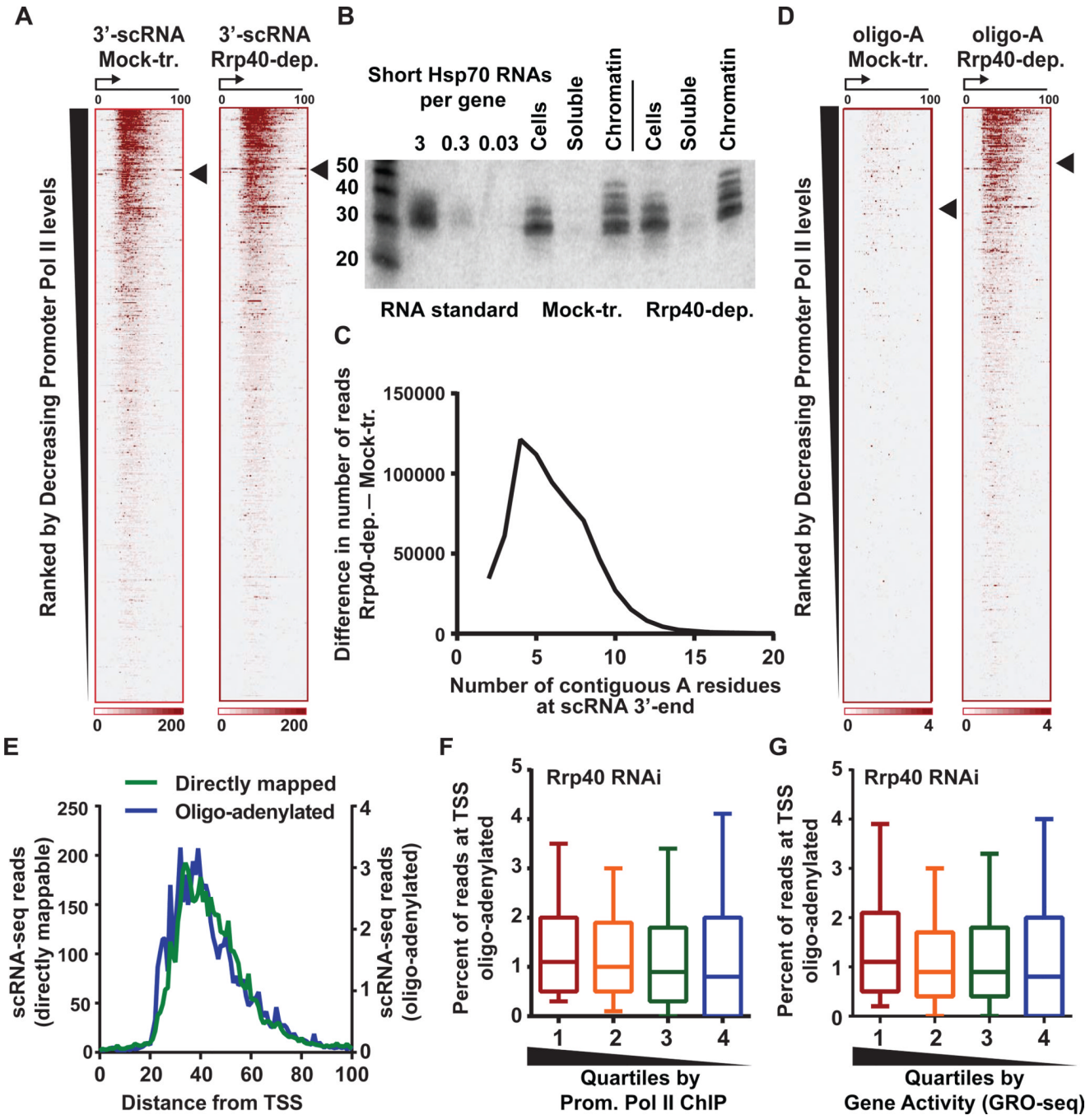

Figure 2. Disruption of exosome activity reveals a low-level of termination coupled with targeted degradation of released RNAs

(A) Heat maps showing 3'-end distributions of scRNAs isolated from Mock-treated or Exosome(Rrp40)-depleted cells at genes occupied by Pol II. Genes are rank-ordered and shown as in Figure 1A.

(B) Northern analysis of short RNAs generated by promoter Pol II at the *Hsp70* gene in Mock-treated and Rrp40-depleted cells. RNAs derived from whole cells are shown along with species present in the chromatin or soluble fractions. To quantify the abundance of these species, a 33 nt synthetic DNA corresponding to the Hsp70 5'-end was loaded over a range of molecular equivalents per gene copy in the S2 cell genome.

(C) Enrichment in 3’-oligo-adenylated species observed in Rrp40-depleted samples. Shown is the difference in number of scRNA-seq reads with 3'-A residues of the specified length in Rrp40-depleted vs. Mock-treated samples.
(D) Oligo-adenylated species identified in exosome-depleted cells are derived from near TSSs. Heat maps are depicted as in (A) for the species that mapped uniquely after trimming 3’-A residues.
(E) Comparison of 3’-end positions of directly mappable and 3’-oligo-adenylated scRNA species isolated from Rrp40-depleted samples shows that they both localize tightly within the region of promoter-proximal pausing, suggesting that they share a common origin.
(F and G) The percentage of reads bearing non-templated A-tails is consistently low across gene quartiles ranked by promoter Pol II signal (F) and gene activity (G).
See also Figure S2.

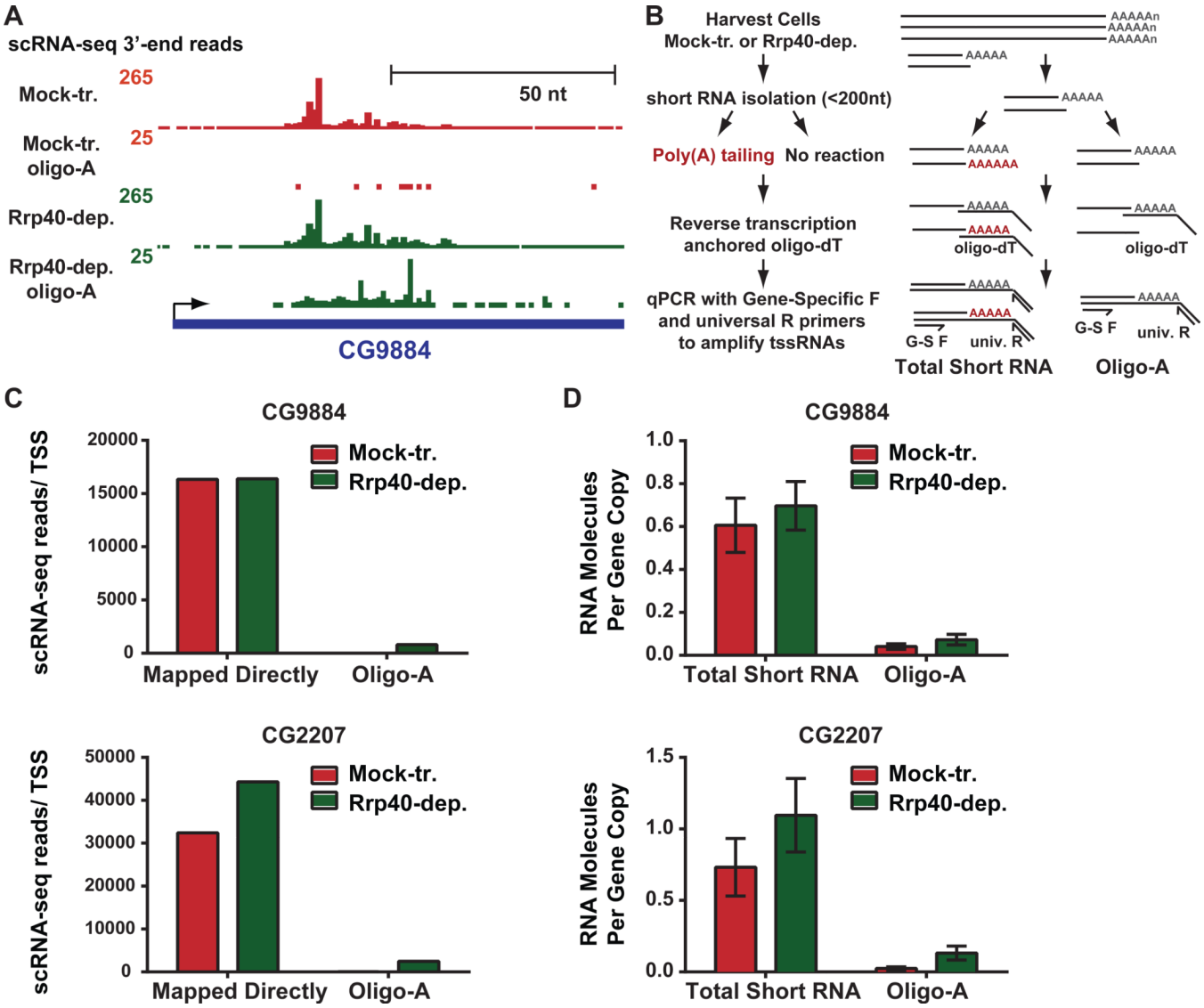

Figure 3. Quantitative PCR indicates abundant tssRNAs at paused genes, with low levels of oligo-adenylated species

(A) 3'-end locations are shown for scRNAs derived from an example gene (*CG9884*) in Mock-treated and Rrp40-depleted cells. The number of reads at each location, in single nucleotide intervals, is shown for each sample. Note the >10-fold difference in scale between reads that were mapped directly and those with A-tails at their 3'-ends (oligo-A), as well as the lack of oligo-A reads in cells with intact exosome activity. The gene model is shown below, with TSS depicted as an arrow.

(B) Schematic of the experimental protocol for amplification of tssRNA species. Following isolation of short RNAs, an *in vitro* adenylation reaction (left) allows for reverse transcription of all short RNA species using an anchored oligo-dT primer. Omitting the *in vitro* adenylation (right) allows for detection of only those species that are oligo-adenylated in vivo. The oligo-dT primer contains a sequence complementary to the universal Reverse primer (univ. R), allowing amplification of individual tssRNAs when coupled with a gene-specific Forward primer (G-S F).

(C and D) Example quantification of the scRNA read counts near promoters (C, from the TSS to 100 nt downstream) and qPCR results (D) for tssRNA species generated by the

CG9884 gene (top), and *CG2207* (below). Values represent the averages of ≥3 biological replicates ± SEM. The small differences in directly mappable scRNAs or total short RNAs observed at individual genes were not statistically significant.
See also Figure S3.

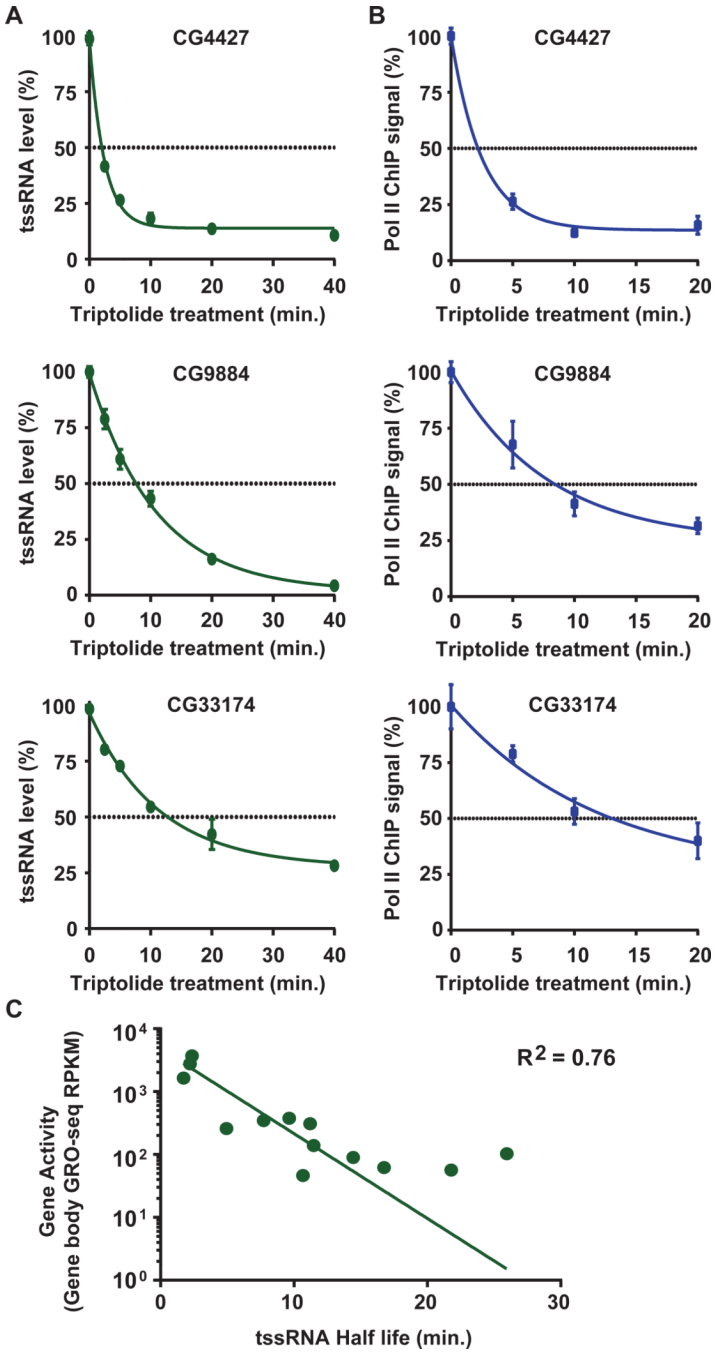

Figure 4. Measurement of the turnover of tssRNA and promoter Pol II indicates a long-lived elongation complex whose residence time is dictated by the rate of pause release

(A and B) Quantification of the short tssRNA levels (A, green) and promoter Pol II ChIP signal (B, blue) during a time course of treatment with Trp. Plots show the decay of signal from genes with varying levels of gene activity as measured by gene body GRO-seq signal: *CG4427* (top) is highly active, *CG9884* (middle) has moderate activity and *CG33174* (bottom) exhibits very low activity. Signals at each time point were quantified by qPCR relative to DMSO-treated cells at t = 0. For tssRNA, each point represents the mean of three biological replicates ± SEM. ChIP samples show average values and range of two replicates.

(C) The relationship between gene activity and the half-lives measured for tssRNAs is shown at thirteen genes, revealing that the lifetime of early elongation complexes is inversely correlated with the rate of release into productive elongation.
See also Figure S4.

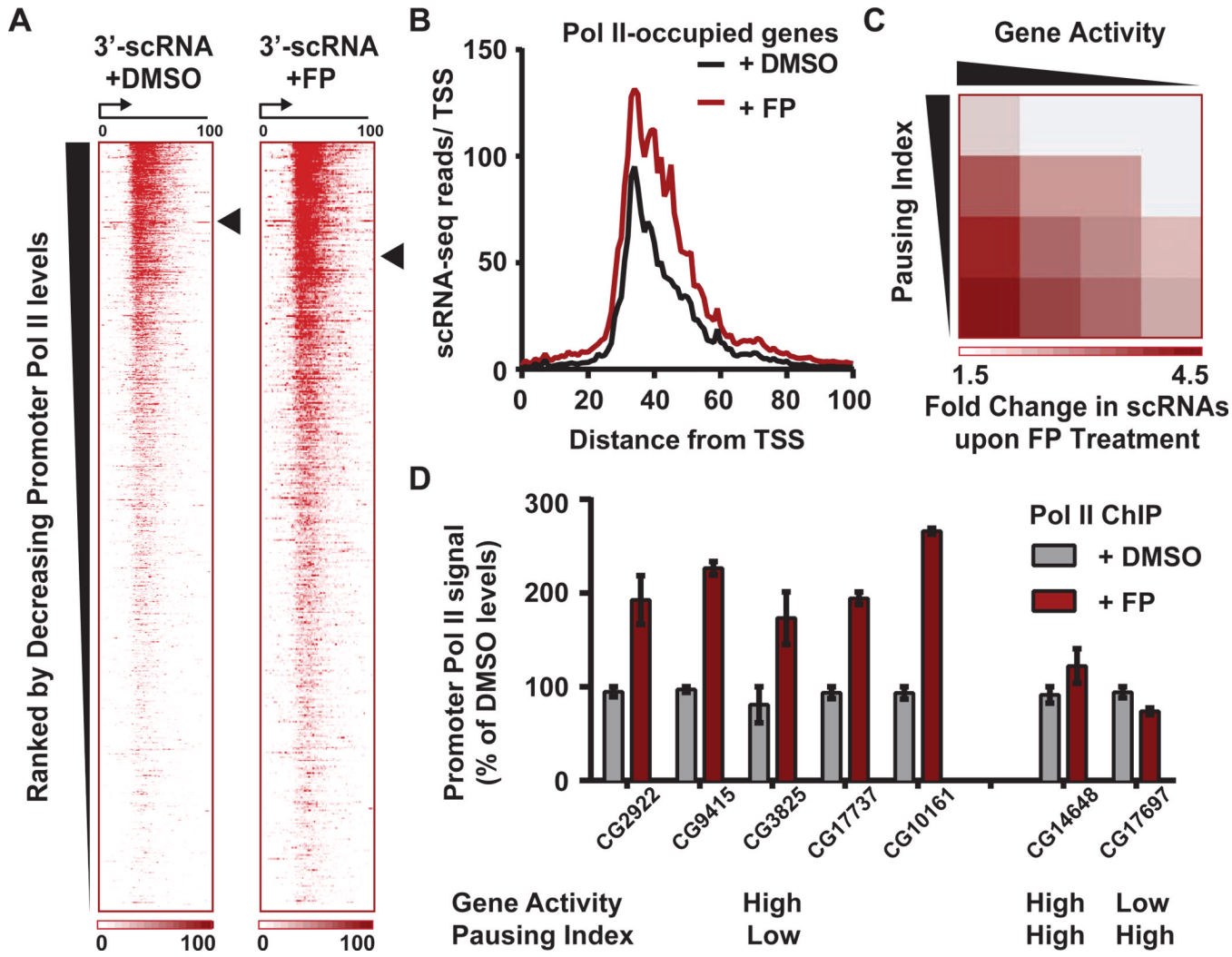

Figure 5. Flavopiridol causes broad accumulation of RNAs and Pol II near promoters

(A) Heat maps showing 3'-end distributions of scRNAs isolated from DMSO-treated or FP-treated cells at genes occupied by Pol II. Genes are rank-ordered and shown as in Figure 1A. Arrow heads at right of heat maps indicate the position above which 50% of the scRNA-seq reads are located, revealing a shift in FP-treated cells.

(B) Composite metagene analysis of scRNA reads 3'-ends detected in the DMSO-treated or FP-treated cells. Shown is the average number and position of scRNA 3'-end reads per TSS.

(C) Heatmap showing the Fold change in scRNA reads upon FP-treatment for genes divided into groups based on their Pausing index (y-axis) and Gene activity level (Gene body GRO-seq signal, x-axis). Color bars at bottom indicate range.

(D) Quantification of the change in promoter Pol II ChIP signal at genes following FP treatment. Promoter Pol II signal was quantified relative to DMSO-treated cells (set as 100%) using primer pairs centered near the TSSs for each gene. The relative pausing index and gene activity level for each gene is shown. Data represents the mean of two biological replicates ± range.

See also Figure S5.

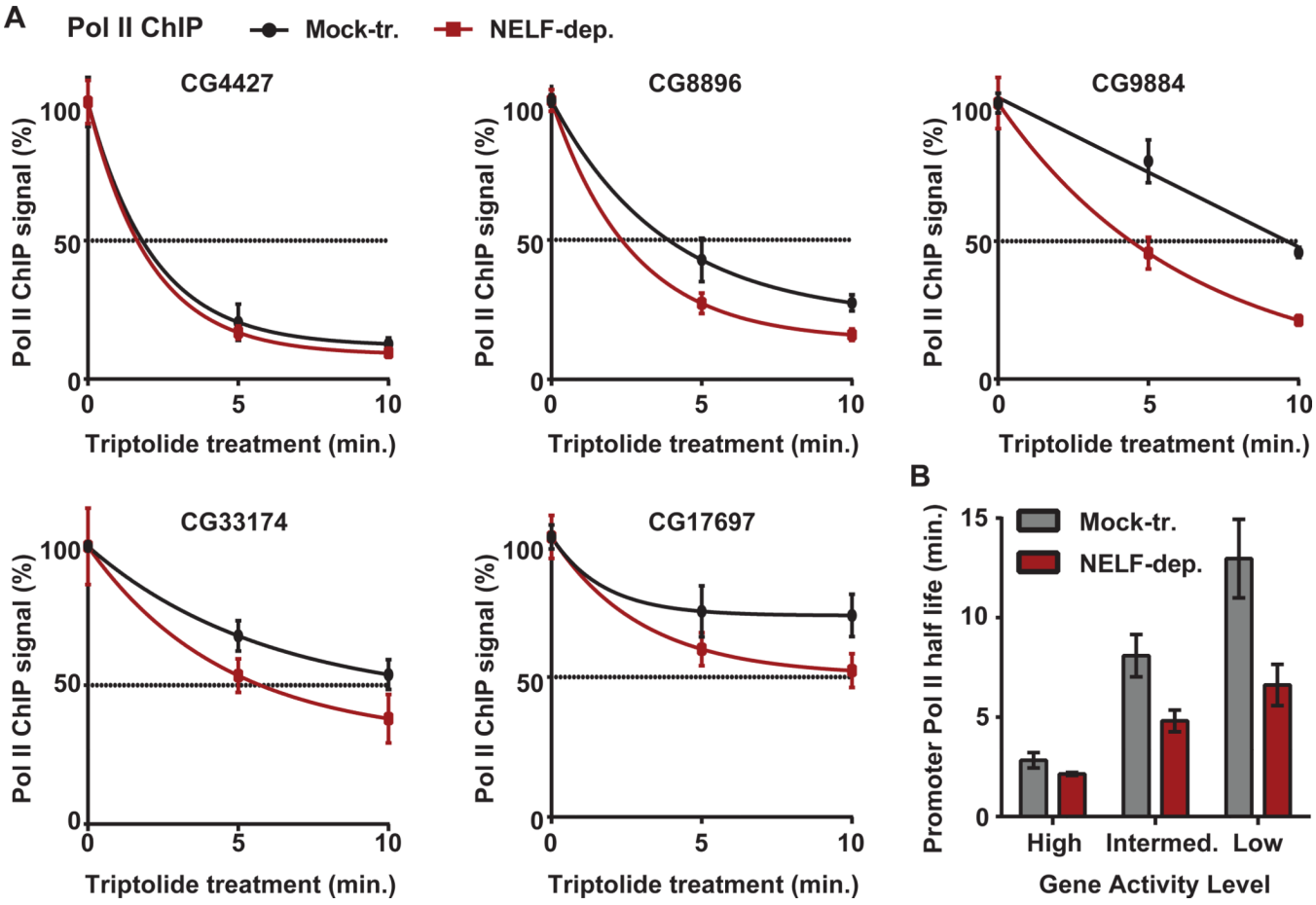

Figure 6. Generation of stable promoter-associated Pol II complexes involves the pause-inducing factor NELF

(A) Promoter Pol II ChIP signal was measured following Trp-treatment in cells that were Mock-treated or NELF-depleted. Genes shown display High (e.g. *CG4427*), intermediate (e.g. *CG8896* and *CG9884*) and Low (*CG33174* and *CG17697*) levels of gene activity as determined by GRO-seq. The percent input of promoter Pol II at each time point was compared to the signal present at t=0 (before adding Trp), and is reported as the percentage of ChIP signal remaining at each time. Values represent averages of ≥3 biological replicates ± SEM.

(B) Difference in Pol II half-life between Mock-treated and NELF-depleted cells. The twelve genes analyzed as in (A) were separated into three groups based on their gene activity levels in untreated S2 cells. Whereas highly active genes with short-lived promoter Pol II occupancy were not significantly affected by NELF-depletion (High P=0.155, N=3), genes with Intermediate and Low activity displayed more long-lived promoter occupancy in untreated cells that was significantly affected by NELF RNAi (Intermediate P=0.025, N=5; Low P=0.029, N=4). Values are mean half-lives for each group of genes ± SEM.

See also Figure S6.

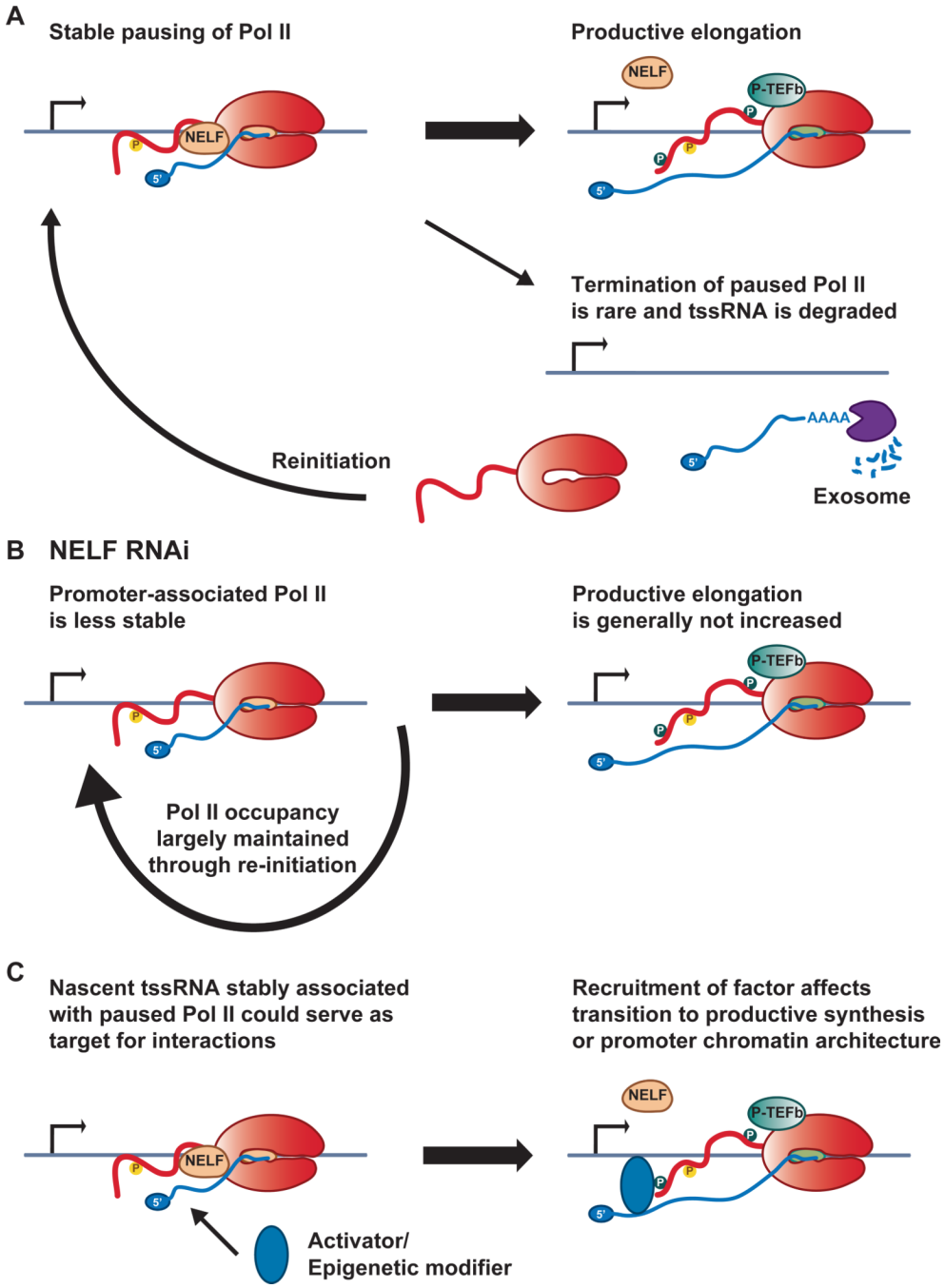

Figure 7. Stable pausing by Pol II provides a framework for regulation

(A) Model: Pol II (red) initiates transcription and is induced to pause through the activity of NELF and other factors (not shown). Paused Pol II and nascent tssRNA (blue) remain stably associated with the promoter while awaiting the recruitment of P-TEFb (green) which triggers productive elongation. Most promoter-associated Pol II can transition to productive elongation (thick arrow); however, we detect a low level of promoter-proximal termination (small arrow), with the exosome (purple) rapidly degrading the released tssRNAs.

(B) Upon depletion of NELF, elongation complexes no longer stably pause and exhibit decreased half-lives near promoters. Many genes can support rapid re-initiation of Pol II to

maintain similar Pol II occupancy, but genes with inefficient re-initiation display decreased expression.

(C) We propose that the tssRNAs produced by, and stably associated with, paused Pol II could play an important role in transcription activation *in cis* through interactions with factors that affect productive elongation and promoter chromatin.