Widespread Backtracking by RNA Pol II Is a Major Effector of Gene Activation, 5’ Pause Release, Termination, and Transcription Elongation Rate

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Authors
Ryan M. Sheridan, Nova Fong, Angelo D’Alessandro, David L. Bentley

Correspondence
david.bentley@ucdenver.edu

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.

Highlights
- Dominant-negative TFIIS stabilizes backtracked pol II throughout human genes
- RNA cleavage by pol II is essential for rapid activation of stress-inducible genes
- Rescue from backtracking is important for escape from promoter-proximal pause sites
- Rescue from backtracking is a major stimulus of rapid transcriptional elongation

In Brief
RNA polymerase II can backtrack and resume transcription after cleavage of the nascent RNA transcript. Sheridan et al. show that in vivo human pol II backtracks frequently and that RNA cleavage enhances gene activation, 5’ pause release, normal termination, and rapid transcription elongation.
Widespread Backtracking by RNA Pol II Is a Major Effector of Gene Activation, 5′ Pause Release, Termination, and Transcription Elongation Rate

Ryan M. Sheridan, Nova Fong, Angelo D’Alessandro, and David L. Bentley

1Department Biochemistry and Molecular Genetics, RNA Bioscience Initiative, University of Colorado School of Medicine, Aurora, CO 80045, USA
2Lead Contact
*Correspondence: david.bentley@ucdenver.edu
https://doi.org/10.1016/j.molcel.2018.10.031

SUMMARY

In addition to phosphodiester bond formation, RNA polymerase II has an RNA endonuclease activity, stimulated by TFIIS, which rescues complexes that have arrested and backtracked. How TFIIS affects transcription under normal conditions is poorly understood. We identified backtracking sites in human cells using a dominant-negative TFIIS (TFIISDN) that inhibits RNA cleavage and stabilizes backtracked complexes. Backtracking is most frequent within 2 kb of start sites, consistent with slow elongation early in transcription, and in 3′ flanking regions where termination is enhanced by TFIISDN, suggesting that backtracked pol II is a favorable substrate for termination. Rescue from backtracking by RNA cleavage also promotes escape from 5′ pause sites, prevents premature termination of long transcripts, and enhances activation of stress-inducible genes. TFIISDN slowed elongation rates genome-wide by half, suggesting that rescue of backtracked pol II by TFIIS is a major stimulus of elongation under normal conditions.

INTRODUCTION

Both initiation of polymerase II (pol II) transcription and post-initiation steps in the transcription cycle—5′ pausing and release, elongation and termination—can influence the control of gene expression. The best understood post-initiation control mechanism is exerted at the promoter-proximal pause site to regulate the flux of pol II into productive elongation within gene bodies (Jonkers and Lis, 2015; Mayer et al., 2017; Nechaev and Adelman, 2011). Following transit past the 5′ pause, transcription elongation within genes proceeds at rates that differ by at least an order of magnitude between genes (Fuchs et al., 2014; Jonkers et al., 2014). For reasons that are not understood, elongation starts out slowly and then accelerates within genes (Danko et al., 2013; Jonkers et al., 2014). This variation in elongation rate is of physiological significance because co-transcriptional processes, including pre-mRNA processing and chromatin modification, are kinetically coupled to ongoing transcription (Bentley, 2014; Naftelberg et al., 2015). Furthermore, sites of termination at the ends of genes are determined by elongation rate as a result of kinetic competition between pol II and the termination factor Xrn2 (Fong et al., 2015; Hazelbaker et al., 2013; Proudfoot, 2016). The mechanisms that determine elongation rates within genes are poorly understood, but one potential rate-limiting event is pol II arrest and backtracking that can occur at positions where the enzyme mis-incorporates a base or runs into a roadblock like a DNA binding factor or a nucleosome (Bondarenko et al., 2006; Churchman and Weissman, 2011; James et al., 2017; Thomas et al., 1998).

In vitro, the transcription factor TFIIS rescues pol II elongation complexes that have backtracked by stimulating the RNA endonuclease activity of the polymerase, which cleaves the nascent RNA to produce a new 3′ end situated in the active site (Izban and Luse, 1992; Reines, 1992; Thomas et al., 1998). TFIIS-stimulated endonucleolytic cleavage releases RNA fragments 2–14 bases long (Izban and Luse, 1993). In vitro, TFIIS does not affect the pause-free velocity of yeast or human pol II and stimulates elongation rate entirely by reducing the length of pauses (Ishbashshi et al., 2014; Zhang and Burton, 2004). In yeast, TFIIS (Dst1) is normally dispensable but is essential when GMP synthesis is inhibited (Exinger and Lacroute, 1992; Gómez-Herreros et al., 2012; Malik et al., 2017; Shaw and Reines, 2000). In Adst1 cells, overexpression of a dominant-negative TFIIS mutant (TFIISDN) that inhibits RNA cleavage is lethal, implying an important function of this activity even when nucleoside triphosphate (NTP) levels are unperturbed (Sigurdsson et al., 2010). TFIISDN has mutations in two conserved residues at the tip of domain III (D190A, E191A), that projects through the side channel of the polymerase and interacts with the active site (Cheung and Cramer, 2011; Jeon et al., 1994; Kettenberger et al., 2003; Kettenberger et al., 2004). These two acidic residues form part of a conserved motif, QTRSADEP, that interacts with the pol II trigger loop and bridge helix and is identical in yeast and mammalian TFIIS (Figure S1A) (Kettenberger et al., 2003). The aspartic acid (D) and glutamic acid (E) residues in this motif coordinate a Mg2+ ion, metal ion B, and a nucleophilic water molecule, which participate in the RNA cleavage reaction (Cheung and Cramer, 2011; Kettenberger et al., 2003). Orthologous acidic residues in bacterial GreA/GreB serve the same function in stimulating transcript
cleavage by RNA polymerase (Sosunova et al., 2003). In the nucleotide addition reaction by the incoming NTP (Cramer et al., 2001; Sosunov et al., 2003). Yeast TFIISDN binding affinity for pol II is similar to wild-type (WT) TFIIS, and like WT TFIIS (Ishibashi et al., 2014; Zhang and Burton, 2004), it does not affect the pause-free rate of nucleotide addition (Sigurdsson et al., 2010), suggesting that its dominant-negative effect is exerted specifically through inhibition of RNA endonuclease activity.

Pol II backtracking sites have been mapped genome-wide in budding yeast by comparing nascent RNA 3’ ends in WT and Δdst1 cells. This study showed that backtracking is ubiquitous and quite uniformly distributed throughout genes (Churchman and Weissman, 2011). TFIIS deletion has little effect on yeast under normal growth conditions, and it remains unclear whether resolution of backtracked complexes has a widespread influence on elongation rates or gene activation. In vitro, TFIIS facilitates elongation of pol II stalled at the Drosophila Hsp70 promoter, and in vivo, it stimulates Hsp70 activity and pol II recruitment to the promoter (Adelman et al., 2005). Backtracking has been directly detected in Drosophila at the +1 nucleosome (Weber et al., 2014) and near promoter proximal pause sites (Nechaev et al., 2010). It has also been inferred within 3’ flanking regions in S. pombe where TFIIS delays transcription termination (Lemay et al., 2014) and in human cells at termination sites that occur near sequence elements predicted to destabilize the DNA-RNA hybrid (Schwaib et al., 2016).

The importance of resolving backtracked pol II complexes for the transition to productive elongation at 5’ pause sites in vivo and for subsequent elongation and termination remains unclear in part because sites of backtracking have not been mapped in a multicellular organism. It remains to be determined whether backtracking occurs uniformly throughout metazoan genes or is more common at particular positions like 5’ pause sites or termination sites. Here, we used an inducible TFIISDN in human cells to identify sites of backtracking genome-wide and to investigate the role of RNA cleavage in pol II transcription under normal and stressed conditions. The results demonstrate that backtracking is ubiquitous and has widespread effects on release of paused pol II, elongation rate within genes, completion of long transcripts, termination, and gene activation.

RESULTS

Pol II RNA Cleavage Activity Promotes Escape from 5’ Pause Sites

To investigate the importance of the RNA cleavage activity that rescues backtracked pol II, we generated a HEK293 Flp-in cell line that inducibly overexpresses the mouse TFIISDN mutant with substitutions (D282A, E283A) homologous to those that inhibit cleavage activity in yeast (Figures S1A and S1B) (Sigurdsson et al., 2010). The mouse and human homologs are 96.7% identical (Figure S1A). Under the conditions of our experiments (24 hr induction), expression of TFIISDN did not detectably affect viability or proliferation as shown by flow cytometry (Figure S1C). This result is consistent with the viability of yeast co-expressing WT and mutant TFIIS (Sigurdsson et al., 2010). Because endogenous WT TFIIS will compete with TFIISDN for binding to pol II in our cells, the effects of the mutant will reflect a partial inhibition of RNA cleavage activity.

To establish whether rescue of backtracked complexes is indeed inhibited by TFIISDN in vivo, we performed global run on sequencing (GRO-seq) (Core et al., 2008), anti-pol II chromatin immunoprecipitation sequencing (ChIP-seq), and nascent RNA 3’ end mapping (mNET-seq) (Nechaev et al., 2010) before and after induction of the mutant protein. Only polymerases that can actively elongate in the run-on reaction will incorporate Br-UTP and produce GRO-seq signal; therefore, if TFIISDN stabilizes backtracked pol II, it is predicted to reduce the GRO-seq signal relative to the total amount of pol II engaged on the gene. These experiments showed that expression of TFIISDN specifically reduced the GRO-seq signal at promoter proximal regions (Figures 1A and 1D) with relatively little effect on total pol II occupancy and transcriptionally engaged pol II as measured by ChIP-seq and mNET-seq, respectively (Figures 1B–1D). These results were observed in each biological replicate (Figures S2 and S3A–S3E) and together show that TFIISDN prevents promoter proximal pol II from resuming elongation in the run-on assay, strongly suggesting that in vivo the mutant stabilizes backtracked complexes by preventing RNA cleavage as it does in vitro (Sigurdsson et al., 2010). Most importantly, the results in Figure 1 show that pol II commonly backtracks within 300 bp of the transcription start site (TSS) and that cleavage activity helps maintain 5’ paused pol II in an elongation-competent state.

The reduced GRO-seq signal at 5’ pause sites caused by TFIISDN suggests a direct effect on paused pol II complexes. We tested this idea by ChIP mapping of wild-type and dominant negative mutant TFIIS expressed in human breast cancer cells (Figure S4A). This experiment showed that TFIIS strongly localizes to 5’ pause sites on human genes (Figures 1E and 1F), consistent with previous work in mouse embryonic stem cells (ESCs) (Carriere et al., 2012). Quantitation of ChIP-seq signals relative to a spike-in control showed that WT and mutant TFIIS were recruited approximately equally well to 5’ pause sites when measured relative to pol II at these regions (Figures 1E and S4B–S4E).

Unexpectedly, the stabilization of backtracked pol II at promoter proximal pause sites did not detectably increase pol II occupancy in this region as measured by pol II ChIP-seq and mNET-seq (Figures 1B–1D and S3B–S3D), possibly because of turnover of backtracked complexes by premature termination that has been reported at 5’ pause sites (Erickson et al., 2018; Steurer et al., 2018). We also noted that TFIISDN elevated relative GRO-seq signals in the region between +1 and +3 kb, possibly reflecting a problem in the maturation of pol II complexes into a fully elongation-competent state (Figures 1A and S3A).

Pol II Cleavage Activity Antagonizes Termination at Gene 3’ Ends

In addition to promoter proximal pausing, which occurs within a short window downstream of the TSS, pol II undergoes a second major pause within a much larger window downstream of the poly(A) site preceding transcription termination (Anamika et al., 2012; Glover-Cutter et al., 2008). Poly(A) site cleavage exposes
a 5′ PO₄ that serves as a substrate for the 5′-3′ exonuclease Xm2, which “torpedoes” the polymerase and promotes termination (Connelly and Manley, 1988; Kim et al., 2004; Proudfoot, 2016; West et al., 2004). Inhibition of cleavage activity by TFIIS DN caused a reproducible upstream shift in both total pol II and the C-terminal domain (CTD) Ser2 phosphorylated isoform associated with the 3′ pause (Figures 2A–2C and S3G). A shift of roughly equal magnitude was also observed for mNET-seq signal for the region from the TSS to +300 bp. The average fold change for two (GRO-seq, mNET-seq) or three (pol II ChIP-seq) biological replicates is shown for genes >1 kb long and separated by >2 kb. p values were calculated using the Welch two-sample t-test with Bonferroni-Holm correction.

We monitored the pol II wave after DRB washout by sequencing nascent RNA pulse labeled with bromouridine (Bru-seq) (Veloso et al., 2014) and by anti-pol II ChIP-seq (Figures 3A and 3B). By these two methods, elongation rates were estimated for

**Pol II Cleavage Activity Is Required to Maintain Rapid Elongation within Genes**

The results above suggest that resolution of backtracked pol II complexes by the pol II endonuclease is important for pol II release from 5′ pause sites and proper termination at 3′ ends. Next, we investigated whether RNA cleavage also influences productive elongation within gene bodies. Elongation rates were measured using the reversible Cdk9 inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), which arrests pol II at the 5′ ends of genes while permitting it to clear from gene bodies. Relief of DRB permits the 5′ paused polymerases to be released into the gene body in a synchronous “wave” (Singh and Padgett, 2009). We monitored the pol II wave after DRB washout by sequencing nascent RNA pulse labeled with bromouridine (Bru-seq) (Veloso et al., 2014) and by anti-pol II ChIP-seq (Figures 3A and 3B). By these two methods, elongation rates were estimated for...
~600–1,200 genes over 75 kb long. In control cells, we observed a median elongation rate of ~2–3 kb/min (Figure 3C), in agreement with previous reports (Jonkers and Lis, 2015). Remarkably, expression of TFIIS DN reduced the median elongation rate to only ~1–1.5 kb/min (Figure 3C). This reduction was observed for both proximal and distal regions of genes (Figure S5A). We conclude that the RNA cleavage activity of pol II is a major positive effector of elongation rate through gene bodies.

If RNA cleavage is crucial for normal elongation within genes, one might expect that long genes would be most affected by TFIIS DN since they have more potential backtrack sites. To study this question, we measured nascent RNA synthesis by Bru-seq and identified 1,000 genes with the greatest reduction in signal after induction of TFIIS DN. Notably within this group, the longest genes showed the greatest decrease in nascent RNA levels after inhibition of pol II cleavage activity (Figure 3D). Furthermore, the most significant reductions in nascent RNA abundance caused by TFIIS DN occurred at the 3' ends of long genes for both biological replicates, with little effect at 5' ends (Figures 3E, S5B, and S5C) or in genes of average length (Figure S5D). These observations suggest that cleavage activity is important not only to maintain rapid elongation rates but also to permit completion of long transcripts, presumably by preventing premature termination.

**Pol II Pauses and Backtracks Most Frequently during Early Elongation and in the Termination Zone**

The widespread transcriptional defects caused by TFIIS DN at early and late stages of the transcription cycle suggest that
backtracking is a relatively ubiquitous event even in unstressed human cells. To determine exactly where backtracking occurs, we analyzed 30 ends of nascent transcripts in immunopurified pol II complexes by mNET-seq (Churchman and Weissman, 2011; Nojima et al., 2015). Cross-correlation analysis demonstrated a good correlation between replicate mNET-seq data-sets at the single base pair level (Figure S6A). A two-step procedure was employed to map positions of backtracking:

1. locate pause sites defined as positions with a mNET-seq signal at least three standard deviations above the mean for a 200 bp window, and
2. identify those pause sites where inhibition of RNA cleavage by TFIISDN caused an accumulation of extended nascent 3’ ends immediately downstream.

Using step 1 of this scheme, we identified 179,605 pausing events occurring between the TSS and +5 kb downstream of the poly(A) site in control cells that were common between two biological replicates. Notably, there is specific enrichment of pauses within 2 kb of the TSS and in the termination zone downstream of poly(A) sites after normalizing for total mNET-seq signal (Figures 4A and 4B).

We did not detect elevated pol II pausing near 5’ or 3’ splice sites, but consistent with co-transcriptional splicing, abundant RNA 3’ ends were reproducibly detected at splice donor sites corresponding to step 1 intermediates associated with pol II (Figures 4C and S6B, left panels). In contrast, excised lariats, which would appear as mNET-seq reads at splice acceptor

Figure 3. Pol II Cleavage Activity Is Required for Rapid Elongation through the Gene Body
(A) Inhibition of pol II cleavage activity reduces the rate of transcription elongation. Metaplots of Bru-seq signals (500 bp bins) after washout of DRB for non-overlapping genes >75 kb long.
(B) Metaplots of pol II ChIP-seq signals after washout of DRB as in (A).
(C) Transcription rates were calculated using Bru-seq data shown in (A) (left) or pol II ChIP-seq data shown in (B) (right) by comparing the 10 min and 20 min time points for genes, in which waves were detected for both time points. p values were calculated using the Welch two-sample t test.
(D) Transcription of long genes is more sensitive to inhibition of pol II RNA cleavage activity. The fold change in Bru-seq signal was plotted for the top 1,000 genes that show reduced Bru-seq signal after expression of TFIISDN, divided into quartiles based on gene length. Fold changes were calculated from two biological replicates for genes >1 kb long and separated by >2 kb. p values were calculated using the Welch two-sample t test with Bonferroni-Holm correction. ****p < 0.0001.
(E) Metaplots of Bru-seq signals for genes from the 3rd and 4th quartiles from D (n = 500, median length: 141 kb). The mean signal and SEM were plotted as in Figure 2B for two biological replicates. Negative values correspond to anti-sense signal. p values (bottom panel) were calculated by comparing the sense Bru-seq signals for uninduced and TFIISDN-expressing cells for each replicate using the Welch two-sample t test. p values were then averaged across the two biological replicates with the shaded region representing the SEM.
sites, were not present in any of our datasets, implying that they are not stably associated with pol II (Figures 4C and S6B, right panels).

Analysis of the sequences flanking each pause site revealed a strong preference for a T at the +1 position on the coding strand immediately downstream of the nascent RNA 3’ end and a weaker preference for Gs at the −1, −2, and −3 positions (Figure 4D) in both WT and TFIISDN datasets (Figures S6C and S6D). This bias was not observed in GRO-seq libraries made using the same method of library preparation (Figure S6E). A similar preference for T at +1 relative to pol II pause sites was also observed in yeast when TFIIS (Dst1) was deleted (Churchman and Weissman, 2011).

Next, we examined how inhibiting pol II cleavage activity affects the 3’ ends of nascent transcripts at pause sites. Notably, there is a small but reproducible increase in mNET-seq signal within 15 bp downstream of these pauses when TFIISDN was expressed (Figures 4E and S7A), indicating that impaired RNA cleavage causes a lengthening of nascent RNA 3’ ends as expected for backtrack sites. To quantify this effect, we calculated a “backtracking index,” which is the difference in mNET-seq signal between the regions 15 bp downstream and upstream of the pause, for each pause separated by at least 30 bp where there is a signal in at least one of the regions (n = 39,543/179,605 pauses). At approximately 30% of these well-separated pauses (11,166/39,543),
TFIISDN increased the backtracking index in both replicate datasets, and we identified these as high-confidence backtracking sites (Figures 5A, 5B, and S7B). This is a minimum estimate of the fraction of pauses associated with backtracking, and it therefore shows that backtracking occurs at a substantial number of pause sites throughout the length of genes and downstream of poly(A) sites. Like pause sites, backtracking sites are also enriched for a T at the +1 position just downstream of the RNA 3' end (Figure S7C). Backtracking sites are widely distributed, but after normalizing for total mNET-seq signal, they are most highly enriched within 2 kb of the TSS and in the termination zone downstream of poly(A) sites (Figure 5C). These data provide the first global picture of pol II backtracking in a metazoan and suggest that RNA cleavage by pol II is required to counteract backtracking at all post-initiation stages of the transcription cycle but especially during early elongation and in the 3' flanking region prior to termination.

Pol II Cleavage Activity Is Required for Rapid Activation of Stress-Inducible Genes

In light of how pol II cleavage activity affects pause release and elongation rate, we asked whether it influences gene activation. To investigate this question, we tested how TFIISDN impacts the upregulation of hypoxia-inducible genes in response to dimethyl sulfoxide (DMSO), which stabilizes hypoxia-inducible factor (HIF) (Semenza, 2011). TFIISDN expression prior to DMOG treatment substantially reduced nascent RNA synthesis at hypoxia-inducible genes in two biological replicates as measured by Bru-seq (Figures 6A, S8A, and S8B), as well as accumulation of mature hypoxia-inducible mRNAs as measured by qRT-PCR (Figure 6B), but had only minor effects on basal mRNA expression levels (Figure S8C). TFIISDN did not reduce transcription of HIF1A or its main binding partner, ARNT (Figure S8D).

To determine whether the reduced transcriptional response to DMOG caused by TFIISDN is physiologically significant, we assayed the metabolic response of the cells by ultra-high-pressure
liquid chromatography-mass spectrometry (UHPLC-MS). The hypoxic response involves elevated conversion of pyruvate to lactate by the HIF inducible enzyme lactate dehydrogenase encoded by the LDHA gene (Firth et al., 1995). Expression of TFIISDN prior to DMOG treatment reduced the accumulation of LDHA transcripts (Figure 6C) and, importantly, decreased lactate/pyruvate ratios relative to control cells (Figure 6D). The results therefore suggest that rescue of backtracked pol II complexes is required for cells to mount a normal metabolic response to hypoxic stress.

We also tested how inhibition of pol II cleavage activity alters the heat shock response. Expression of TFIISDN prior to heat shock at 42°C significantly reduced the accumulation of heat shock responsive transcripts but only had minor effects on basal mRNA levels (Figures 6E and S8E). Consistent with a role in activation of stress-inducible genes, TFIIS can be detected with paused pol II at the 5’ ends of these genes prior to induction (Figures S4C and S4D). Overall, the results in Figure 6 show that the ability of cells to rescue backtracked polymerases is important for rapid gene activation in response to stress stimuli.

**DISCUSSION**

In this report, we used a dominant-negative mutant TFIIS that inhibits the RNA cleavage activity of pol II (Sigurdsson et al., 2010) to identify where pol II backtracks on human genes. The results revealed important effects of RNA cleavage-mediated rescue of backtracked complexes at multiple stages of the transcription cycle (Figure 7). Inhibition of RNA cleavage by TFIISDN increased production of 3’ extended nascent RNAs within 15 bp downstream of pause sites (Figures 4E and S7A) as predicted for backtracked complexes. mNET-seq revealed approximately 180,000 pause sites and 11,000 high-confidence backtracking sites (Figures 4 and 5). Compared to yeast genes (Churchman and Weissman, 2011), which are relatively short, backtracking is less uniformly distributed along human genes, which have defined regions of high pausing and backtracking frequency within 2 kb downstream of the TSS and in the termination zone downstream of poly(A) sites (Figures 4E and S7A) as predicted for backtracked complexes. mNET-seq revealed approximately 180,000 pause sites and 11,000 high-confidence backtracking sites (Figures 4 and 5). Compared to yeast genes (Churchman and Weissman, 2011), which are relatively short, backtracking is less uniformly distributed along human genes, which have defined regions of high pausing and backtracking frequency within 2 kb downstream of the TSS and in the termination zone downstream of poly(A) sites (Figure 4A, 4B, and 5C). We did not detect elevated pausing in the vicinity of splice sites, however. The functional significance of elevated backtracking at
5’ and 3’ ends of genes is supported by our finding that TFIIS$_{DN}$ altered both escape from the 5’ pause and the position where termination occurs (Figures 1A, 2, S3A, S3F, and S3G).

At 5’ ends, inhibition of RNA cleavage activity impaired the ability of 5’ paused polymerases to resume elongation in the run-on (GRO-seq) assay without inhibiting recruitment of pol II to 5’ ends as measured by anti-pol II ChIP-seq and mNET-seq (Figures 1 and S3). These results indicate that pol II cleavage activity is widely required to prepare promoter proximally paused pol II for release into the gene body by maintaining it in an elongation-competent state. This idea is supported by a recent structural study showing that NELF and TFIIS occupy overlapping sites within the polymerase, suggesting that NELF could promote 5’ pausing by competing with TFIIS and preventing the rescue of backtracked pol II (Vos et al., 2018).

Following release from the promoter proximal pause, elongation efficiency improves over the following 2 kb as indicated by gradually reduced levels of pausing and backtracking (Figures 4B and 5C). The high frequency of pausing and backtracking within the first part of the gene body provides a likely explanation for the finding that transcription elongation is slower within more proximal gene regions than further downstream (Danko et al., 2013; Jonkers et al., 2014). This effect could be due to a local property of the template in the first part of the gene that promotes pausing and backtracking. Alternatively, the pol II transcription complex may become intrinsically more resistant to pausing and backtracking later during elongation as a result of an unknown maturation process possibly caused by modification of pol II or associated factors. For example, it is possible that backtracking is suppressed by changes in phosphorylation of the CTD (Buratowski, 2009; Heidemann et al., 2013) or the elongation factor Spt5 (Bourgeois et al., 2002; Yamada et al., 2006) that occur as pol II traverses the first part of the gene.

Pausing and backtracking, as well as TFIIS occupancy, are elevated downstream of poly(A) sites, and furthermore, stabilization of backtracked pol II by TFIIS$_{DN}$ shifted the termination zone more proximally (Figures 2, 4B, and 5C). Backtracking is therefore common at 3’ termination sites where it facilitates early termination. The position where termination occurs is dictated by a kinetic competition between two enzymes moving 5’-3’: the Xrn2 RNA exonuclease “torpedo” and the pol II molecule that it is chasing (Fong et al., 2015). The implication of proximal termination caused by TFIIS$_{DN}$ is that backtracked pol II complexes are susceptible to Xrn2-mediated termination. This conclusion is consistent with previous work showing that an A/U-rich sequence element predicted to promote backtracking by forming an unstable RNA-DNA hybrid coincides with transcription termination (Schwalb et al., 2016).

Early termination like that induced by TFIIS$_{DN}$ is also associated with TFIIS deletion in S. pombe (Lemay et al., 2014). The susceptibility of backtracked pol II to termination is suggested by the fact that backtracking by more than one base induces a conformation with an unfolded trigger loop (Cheung and Cramer, 2011; Zhang et al., 2010) that may be more prone to opening of the clamp that stabilizes the ternary complex (Sekine et al., 2015). It is also possible that frequent backtracking near promoter proximal pause sites could sensitize pol II to premature termination, leading to turnover of paused complexes (Erickson et al., 2018; Steurer et al., 2018). This idea is suggested by the unexpected observation that inhibition of elongation at the promoter proximal pause by TFIIS$_{DN}$ did not increase pol II density there (Figures 1B–1D and S3B–S3D). Suppression of premature termination by TFIIS is also suggested by the fact that it is specifically required for completion of long transcripts as shown by the observation that TFIIS$_{DN}$ preferentially inhibited transcription of the 3’ ends of the longest genes (Figures 3D, 3E, S5B, and S5C).

Remarkably, inhibition of pol II cleavage activity by TFIIS$_{DN}$ caused an ~50% reduction in the elongation rate within numerous genes (Figures 3A–3C and S5A). The implication of this finding is that even when the pol II complex is expected to be at its most productive within the body of genes, it still relies heavily on its RNA cleavage activity to move at its normal speed of ~2 kb/min. These results suggest that TFIIS-stimulated rescue of backtracked pol II has a remarkably profound effect on average elongation rates throughout the genome under normal growth conditions. The slowing of elongation by mutant TFIIS is most likely caused by inhibition of the strong TFIIS-activated RNA cleavage activity of pol II, but inhibition of the weak intrinsic TFIIS-independent cleavage activity (Sigurdsson et al., 2010) might also contribute. We cannot completely eliminate the possibility that, in vivo, the TFIIS D282A, E283A double mutant slows
elongation by an additional unknown mechanism, but this scenario seems unlikely because: (1) the metal ion B Mg2+ coordinated by D282 and E283 does not participate in the nucleotide addition reaction, (2) neither WT yeast nor human TFIIS affects the pause-free rate of elongation in vitro (Ishibashi et al., 2014; Zhang and Burton, 2004), and (3) yeast TFIISDN does not affect the elongation reaction in vitro (Sigurdsson et al., 2010).

The major influence of TFIIS-stimulated RNA cleavage on elongation under normal conditions suggests that physiological modulation of TFIIS activity could regulate elongation rates. In this context, it is of interest that Ccr4-Not and DSIF/NELF can positively and negatively regulate TFIIS activity in vitro (Dutta et al., 2015; Palangat et al., 2005), and NELF likely competes with TFIIS for a common binding site on pol II (Vos et al., 2018). In summary, the widespread effects of TFIISDN on multiple steps in the transcription cycle show that rescue of backtracked pol II complexes is essential for the success of most rounds of transcription in vivo under normal conditions rather than being a special requirement under stress conditions. Our results also suggest that expression of TFIISDN may be of value more generally as a means of experimentally or therapeutically manipulating pol II elongation rate.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Human cell lines
- METHOD DETAILS
  - Antibodies
  - Immunoblotting
  - Flow cytometry
  - ChIP-seq
  - Bru-seq
  - Calculation of elongation rates
  - mNET-seq
  - Identification of pol II pause sites
  - Cross correlation analysis
  - GRO-seq
  - Quantifying fold change in TSS signal
  - 5' Metaplots
  - 3' Metaplots
  - qRT-PCR
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures and two tables and can be found with this article online at https://doi.org/10.1016/j.molcel.2018.10.031.

ACKNOWLEDGMENTS

We thank T. Blumenthal, J. Hesselberth, S. Ramachandran, N. Mukherjee, R. Landick, and members of our lab for helpful discussions, S. Bustos and R. Cohrs for help with extract sonication, and K. Helm and the U. Colorado Cancer Center Support grant (P30CA046934) for flow cytometry. We thank K. Diener, B. Gao, and the UC Denver Sequencing facility for sequencing. R.M.S. is a scholar of the UC Denver RNA Bioscience Initiative. This work was supported by a Boettcher Webb-Waring early career award to A.D. and NIH grant R35GM118051 to D.L.B.

AUTHOR CONTRIBUTIONS

R.M.S. and D.L.B. designed experiments and wrote the manuscript. R.M.S. performed experiments and all computational analysis. N.F. constructed the inducible TFIISWT and TFIISDN cell lines. A.D. designed and carried out metabolomic analysis.

DECLARATION OF INTERESTS

Authors declare no competing interests.

RECEIVED: April 24, 2018
REvised: August 10, 2018
ACCEPTED: October 17, 2018
PUBLISHED: November 29, 2018

REFERENCES


STAR METHODS

KEY RESOURCES TABLE

<table>
<thead>
<tr>
<th>REAGENT or RESOURCE</th>
<th>SOURCE</th>
<th>IDENTIFIER</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antibodies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rabbit anti-pan pol II CTD</td>
<td>Schroeder et al., 2000</td>
<td>N/A</td>
</tr>
<tr>
<td>Mouse 8WG16 anti-pol II</td>
<td>Thompson et al., 1990</td>
<td>N/A</td>
</tr>
<tr>
<td>Rabbit anti-TFIIS</td>
<td>This paper</td>
<td>N/A</td>
</tr>
<tr>
<td>Rabbit anti-avi-tag</td>
<td>GenScript</td>
<td>A00674</td>
</tr>
<tr>
<td>Rabbit anti-Cstf77</td>
<td>Glover-Cutter et al., 2008</td>
<td>N/A</td>
</tr>
<tr>
<td>Mouse B44 anti-BrU</td>
<td>Gratzner, 1982</td>
<td>N/A</td>
</tr>
<tr>
<td>Chemicals, Peptides, and Recombinant Proteins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5,6-dichlorobenzimidazole-1-ribofuranoside</td>
<td>Sigma</td>
<td>D1916</td>
</tr>
<tr>
<td>Deposited Data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All sequencing datasets</td>
<td>This paper</td>
<td>GEO: GSE120201</td>
</tr>
<tr>
<td>Experimental Models: Cell Lines</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HEK293 Flp-in T-Rex pcDNA5 avi-tag TFIIS\textsubscript{DN}</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>HMLE-RAS pCW57bla avi-tag TFIIS\textsubscript{WT}</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>HMLE-RAS pCW57bla avi-tag TFIIS\textsubscript{DN}</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Oligonucleotides</td>
<td>See table S2 for oligonucleotide sequences</td>
<td>N/A</td>
</tr>
<tr>
<td>Software and Algorithms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RStudio version 3.5.0</td>
<td><a href="https://www.rstudio.com/">https://www.rstudio.com/</a></td>
<td></td>
</tr>
<tr>
<td>R version 1.1.453</td>
<td><a href="https://www.r-project.org/">https://www.r-project.org/</a></td>
<td></td>
</tr>
<tr>
<td>Picard Tools version 2.18.7</td>
<td><a href="https://broadinstitute.github.io/picard/">https://broadinstitute.github.io/picard/</a></td>
<td></td>
</tr>
<tr>
<td>Samtools version 1.5</td>
<td>Li et al., 2009</td>
<td><a href="http://www.htslib.org/">http://www.htslib.org/</a></td>
</tr>
<tr>
<td>DESeq2 version 1.10.1</td>
<td>Love et al., 2014</td>
<td><a href="https://bioconductor.org/packages/release/bioc/html/DESeq2.html">https://bioconductor.org/packages/release/bioc/html/DESeq2.html</a></td>
</tr>
</tbody>
</table>

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, David Bentley (david.bentley@ucdenver.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human cell lines

Flip-In-293 TREX cells (Female, Invitrogen) expressing inducible dominant negative mouse TFIIS (TCEA1) (D282A, E283A) with a N-terminal Avitag (TFIIS\textsubscript{DN}) were generated through integration of a pcDNA5/FRT/Avi-TFIIS\textsubscript{DN} construct using Fip recombinase-mediated site-specific recombination. The cells were maintained in DMEM supplemented with 10% FBS, 200 μg/mL hygromycin B, 6.5 μg/mL blasticidin, and penicillin/streptomycin. All experiments were performed after induction with 2 μg/mL doxycycline for 24 hr.

HMLE-RAS cells (Elenbaas et al., 2001) expressing doxycycline inducible mouse Avi-TFIIS\textsubscript{DN} and Avi-TFIIS\textsubscript{WT} were made by lentiviral infection and blasticidin selection.
METHOD DETAILS

Antibodies
Rabbit anti-TFIIS was raised against recombinant full-length mouse GST-tagged TFIIS and affinity purified. Mouse TFIIS is 96.7% identical to its human homolog and this antibody is therefore expected to react equally well with both. Rabbit anti-pol II pan CTD, rabbit anti-CstF77, 8WG16 anti-pol II CTD, and B44 anti-BrdU have been described (Glover-Cutter et al., 2008; Gratzner, 1982; Thompson et al., 1990). Rabbit anti-Avi-tag was purchased from GenScript (A00674). Rat anti-CTD Ser2-P (3E10) was purchased from Chromotek.

Immunoblotting
Immunoblots were developed with HRP conjugated swine anti-rabbit secondary antibody (DAKO, P0217) and ECL Plus (Perkin Elmer, NEL103E001EA).

Flow cytometry
HEK293 TFIISdn cells were collected + and − 24 hr induction with doxycycline (2 µg/ml) and were incubated in Krishan stain for 72 hours before performing flow cytometry using a Beckman Coulter Gallios 561. ModFit LT software (Verity Software House, Topsham, ME) was used for cell cycle analysis.

ChIP-seq
ChIP of human extracts has been described previously (Glover-Cutter et al., 2008). 1 mg of each sonicated lysate was pre-cleared with 30 µl protein A Sepharose beads (Life technologies, 101042) for 1 hr at 4°. Pre-cleared lysates were mixed with 30 µl protein A Sepharose beads (blocked with 1 mg/ml BSA and 0.5 mg/ml E. coli tRNA) and 15 µL of rabbit anti-total pol II CTD serum, 5 µg anti-Avi tag antibody, or 25 µL rat anti-Ser2-PO4-CTD antibody. For Ser2-PO4 IPs, used 4.5 µL rabbit anti-rat IgG secondary (Jackson labs, 312-005-003).

Libraries were sequenced on an Illumina Hi-seq 4000 (1X50) or NovaSeq 6000 (2x150). Adapters were trimmed using cutadapt version 1.16 and reads were mapped to the hg19 UCSC human genome (Feb. 2009) or sacCer3 UCSC yeast genome (April 2011), where appropriate with Bowtie2 version 2.3.2. PCR duplicates were removed using Picard Tools version 2.18.7. Read coordinates were collapsed to a single basepair and centered assuming a fragment size of 180 bp.

For Figures 1E, 1F, and S4, 100 µg crosslinked S. cerevisiae lysate expressing Avi-tagged Rpb3 was mixed with 1 mg cross linked lysate prior to sonication. Each lysate was divided into two IPs using either 15 µL of rabbit anti-total pol II CTD serum or 5 ug anti-Avi-tag antibody (GenScript, A00674). For metaplots, read counts in each bin were divided by the size of the bin and the number of mapped sacCer3 spike-in reads. To calculate TFIIS/pol II ratios for each gene, the normalized Avi-tag ChIP-seq signal was divided by the normalized pol II ChIP-seq signal for each bin. For bins that lacked pol II ChIP-seq signal, a pseudo count (smallest pol II ChIP-seq value for the gene / 2) was used. The TFIIS/pol II ratios for each bin were then averaged across the plotted genes. The 5’ metaplots included genes > 2 kb long, separated by > 2 kb, containing at least one mapped read in the region, and were in common between all ChIP-seq datasets. The 3’ metaplots included genes > 2 kb long, separated by > 5 kb, containing at least one mapped read in the region, and were in common between all ChIP-seq datasets.

Bru-seq
Bru-seq was performed as described (Paulsen et al., 2014) with minor modifications. Cells were incubated with 2 mM Bromouridine for 30 min and labeled RNA was immunopurified with B44 monoclonal anti-BrdU (Gratzner, 1982). All libraries were generated using the KAPA biosciences stranded RNA-seq kit (KIK8401) and sequenced on an Illumina Hi-Seq 4000 (1x50 or 1x150). After filtering out rRNA, reads were mapped to the hg19 UCSC human genome (Feb. 2009) with Bowtie2 version 2.3.2. Adapters were trimmed using cutadapt version 1.16 and PCR duplicates were removed using Picard Tools version 2.18.7. Read coordinates were collapsed to a single basepair and centered assuming a fragment size of 180 bp. For all metaplots, read counts were normalized by the size of the region and the library size. For metaplots in Figure 3E, 6A, and S5B-D, signal was plotted for 200 bp bins for the regions from −5 kb to the TSS and from the pAS to +5 kb, and 30 bins of variable length for region from the TSS to the pAS. For Figure 3D, fold changes in Bru-seq signal were calculated by comparing the number of reads mapping to each gene using DESeq2 version 1.10.1 (Love et al., 2014). For Figure 6A, genes upregulated after treatment with 2 mM DMOG for 16 hr were identified using DESeq2 version 1.10.1 (p < 0.01, n = 288).

Calculation of elongation rates
Cells were treated with 100 µM DRB for 3 hr. For DRB Bru-seq experiments, the cells were washed twice with PBS and incubated in -DRB medium containing 2 mM bromouridine for 5, 10, or 20 min, before harvesting in with Trizol (Thermo Fisher). For DRB pol II ChIP-seq experiments (Figure 2C and 2D), the cells were washed twice with PBS and incubated in -DRB media for 5, 10, or 20 min before crosslinking. The pol II wave was identified by first calculating the mean signal for the region from +70 kb to +75 kb downstream of the TSS, which we defined as background. We then identified the position where signal rose 8 standard deviations above the mean background signal. We only included genes where the front edge of the wave was called at a position downstream of the...
wave for the previous time point. For Bru-seq DRB washout metaplots, signal was plotted for 500bp bins for non-overlapping genes > 75 kb long. For pol II ChIP-seq DRB washout metaplots, signal was plotted for 100 bp bins for the region from the TSS to +500 bp and 500 bp bins for the region from +500 bp to +60 kb for non-overlapping genes > 75 kb long.

mNET-seq

mNET-seq was performed as previously described using the 8WG16 anti-pol II CTD antibody (Nojima et al., 2016). Each mNET-seq library was made from 8-12 15 cm dishes of cells grown to 80% confluency. Nuclei were extracted with NUN1 and NUN2 buffers as described (Nojima et al., 2016) and washed pellets were digested in 100 μL with 2 μL MNase (NEB M0247) for 2 min at 37° in a thermomixer. Reactions were stopped by adding 10 μL 250 mM EGTA and pooled supernatants (400-600 μl) were diluted 10X with NET-2 buffer and immunoprecipitated with 400-600 μL antibody-conjugated protein A/G dynabeads (Invitrogen 10015D). Beads were washed 8x with 1 mL of NET-2 buffer. RNA was extracted with Trizol, treated with T4 PNK (‘3’ phosphatase minus, NET M0236) and DNasel (NEB, M0303) and libraries were generated using the Lexogen small RNA-seq kit (052). After filtering out rRNA, reads were then mapped to the hg19 UCSC human genome (Feb. 2009) with Bowtie2 version 2.3.2. Adapters were trimmed using cutadat version 1.16 and read coordinates were collapsed to a single basepair coordinate corresponding to the RNA 3’ end. PCR duplicates were filtered using a custom script that removes reads having the same 5’ and 3’ ends. 64–75% of reads were between 20 and 30 bases long and were assumed to correspond to pol II footprints. We allowed up to 5 copies of reads < 30 bp in length to be kept, for reads > 30 bp in length only a single copy of each read was kept.

Identification of pol II pause sites

Pol II pause sites were located using a custom script to identify positions where there were at least 5 mapped reads and mNET-seq signal was > 3 standard deviations above the mean signal (not including pauses) for the surrounding 200 bp. This was accomplished by performing successive rounds of “searching” where the pause sites identified during the previous rounds were excluded from the calculations for the current round. This was performed until no additional pause sites were identified. Pauses that mapped to splice sites or snoRNAs were excluded from all additional analysis except for Figure 4C. For analysis of signals and sequences surrounding pause sites (Figures 4D, 4E, S6C, S6D, and S7) and backtracking analysis (Figures 5B and S7B) pauses were only included that were separated by > 30 bp and contained signal within 15 bp upstream or downstream of the pause site. Pause density and backtrack density (Figures 4B and 5C) were calculated by binning the sites into 100 bp windows and normalizing by window size. Pause sites were only used that were shared between the two uninduced datasets (179,605 pauses) and that mapped between the TSS and +5 kb downstream of the poly(A) site. To normalize the site density, the mNET-seq signal within each bin for each gene was averaged for the biological replicates. The site density along each gene was then divided by the mNET-seq signal within the bin. For bins with no mapped mNET-seq reads, a pseudo count (smallest mNET-seq value for the gene / 2) was used. The normalized density was then averaged across non-overlapping genes > 5 kb long for the region downstream of the TSS or it was averaged across genes > 2 kb long and > 5 kb separated for the region around the poly(A) site. Genes were only included that contained at least one pause or backtrack site in the region plotted. For Figure 4D nucleotide frequency was calculated for the region ± 5 bp around the indicated sites and sequence logos were generated using the R package, ggsqlogo (https://github.com/omarwagih/ggsqlogo). For Figures S6C, S6D, S6E, and S7C, sequence logos were generated using WebLogo 3 (http://weblogo.threeplusone.com/create.cgi).

Cross correlation analysis

The mNET-seq signal at each bp was computed for each gene. R was then used to calculate the Pearson’s correlation coefficient for each gene after shifting one of the replicate datasets by the amount indicated on the x axis. The correlation coefficients were then averaged over the entire gene list. Genes were only included if they contained at least one of the pause sites used in Figure 4E.

GRO-seq

GRO-seq was performed as previously described with BrU labeling for 5 minutes (Core et al., 2008). Libraries were generated using the Lexogen small RNA-seq kit (052) and sequenced on an Illumina HiSeq 4000 (1x50 or 1x150). After filtering out rRNA, reads were mapped to the hg19 UCSC human genome (Feb. 2009) with Bowtie2 version 2.3.2. Adapters were trimmed using cutadat version 1.16 and PCR duplicates were removed using Picard Tools version 2.18.7. Read coordinates were collapsed to a single basepair and centered assuming a fragment size of 80 bp.

Quantifying fold change in TSS signal

The GRO-seq signal for each ChIP-seq replicate was calculated as the number of reads (normalized for library size) that mapped between the TSS to +300bp for all genes longer than 1 kb, separated by > 2 kb, contained at least one read in the region, and were in common between all ChIP-seq, GRO-seq, and mNET-seq datasets. Fold changes were calculated after expression of TFIIS_DN for each biological replicate. The fold changes for the replicates were averaged in Figure 1D.

5’ Metaplots

GRO-seq, pol II ChIP-seq, and mNET-seq 5’ metaplots from Figures 1A–1C, 1F, and S3A–S3C include all genes longer than 5 kb, separated by > 2 kb, containing at least one read in the region, and were in common between all ChIP-seq, GRO-seq and mNET-seq
datasets. Read counts were normalized by the size of the bin. To calculate the relative signal for each gene, the signal for each bin was divided by the total signal within the region plotted. The mean relative signal was then calculated by averaging the relative signal across the genes plotted for each individual replicate. The replicates were then averaged with the shaded region representing the standard error of the means (SEM) for two (GRO-seq and mNET-seq) or three (pol II ChIP-seq) replicates.

3' Metaplots
Pol II ChIP-seq and mNET-seq 3' metaplots from Figures 2, S3F, and S3G include all genes longer than 1 kb, separated by > 5 kb, containing at least one read in the region and were in common between all ChIP-seq and mNET-seq datasets. Read counts for each bin were normalized by the size of the bin and the total number of mapped reads. The mean signal was calculated by averaging the normalized read counts across the genes plotted for each individual replicate. The replicates were then averaged with the shaded region representing the standard error of the means (SEM) for two (mNET-seq) or three (pol II ChIP-seq) replicates.

qRT-PCR
Quantitative RT-PCR was performed on 2-3 independent biological replicates of uninduced and TFIISDN-expressing cells (induced with 2 μg/ml doxycycline for 24 hr) treated ± 2mM DMOG for 2, 8, or 16 hr or incubated at 42 °C for 30 min or 1 hr. Random primed cDNA was synthesized with SuperScript III (Invitrogen) and amplified using primers shown in Table S1. Fold change in mRNA abundance was calculated using the ΔΔCt method using β-Actin as a reference gene.

QUANTIFICATION AND STATISTICAL ANALYSIS
Statistical details including the number of replicates (N) are provided in figure legends. Significance was defined by p values determined by using the Welch t test. When appropriate, p values were corrected for multiple testing using the Bonferroni-Holm method.

DATA AND SOFTWARE AVAILABILITY
The accession number for the ChIP-Seq, Gro-seq, mNET-seq, and Bru-seq datasets reported in this paper is GEO GSE120201. All custom software is available at https://github.com/sheridar/TFIISdn_paper.