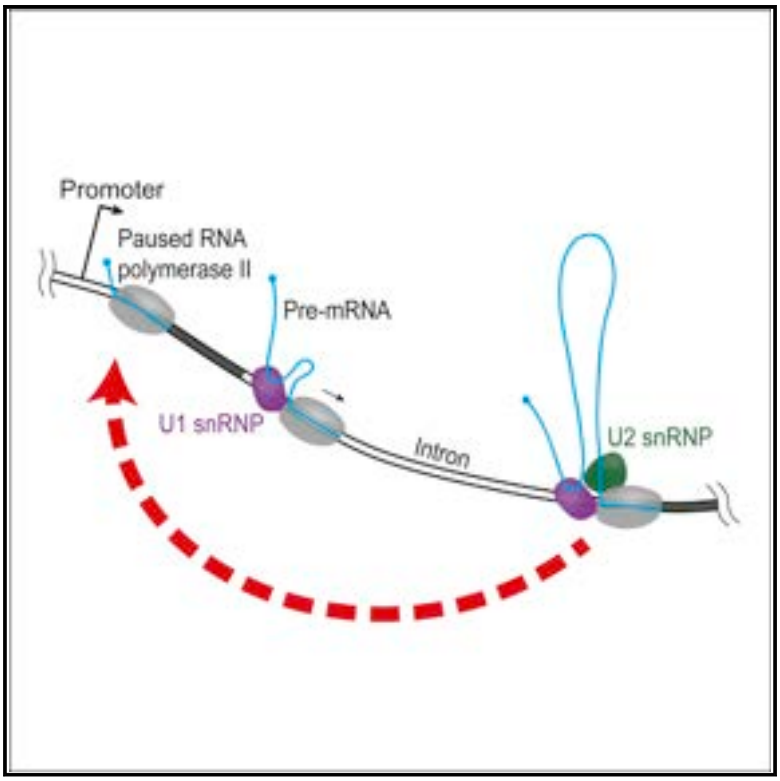
Molecular Cell

Article

Efficient RNA polymerase II pause release requires U2 snRNP function

Graphical Abstract



Authors

Livia Caizzi, Sara Monteiro-Martins, Björn Schwalb, ..., Ying Chen, Michael Lidschreiber, Patrick Cramer

Correspondence

patrick.cramer@mpibpc.mpg.de

In brief

Caizzi et al. show that the spliceosome component U2 snRNP is important for efficient release of promoter-proximally paused RNA polymerase II into productive transcription elongation. This indicates that mammalian mRNA biogenesis relies on positive feedback from the splicing to the transcription machinery.

Highlights

- U2 snRNP has a positive effect on RNA polymerase II transcription
- U2 snRNP function is required for efficient RNA polymerase II pause release
- U2 snRNP influences Pol II elongation velocity at the beginning of genes

Caizzi et al., 2021, Molecular Cell *81*, 1920–1934
May 6, 2021
https://doi.org/10.1016/j.molcel.2021.02.016

CellPress

CellPress

Molecular Cell

Article

Efficient RNA polymerase II pause release requires U2 snRNP function

Livia Caizzi,[1,2] Sara Monteiro-Martins,[1,2] Björn Schwalb,[1] Kseniia Lysakovskaia,[1] Jana Schmitzova,[1] Anna Sawicka,[1] Ying Chen,[1] Michael Lidschreiber,[1] and Patrick Cramer[1,3,*]

[1]Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

[2]These authors contributed equally

[3]Lead contact

*Correspondence: patrick.cramer@mpibpc.mpg.de

https://doi.org/10.1016/j.molcel.2021.02.016

SUMMARY

Transcription by RNA polymerase II (Pol II) is coupled to pre-mRNA splicing, but the underlying mechanisms remain poorly understood. Co-transcriptional splicing requires assembly of a functional spliceosome on nascent pre-mRNA, but whether and how this influences Pol II transcription remains unclear. Here we show that inhibition of pre-mRNA branch site recognition by the spliceosome component U2 snRNP leads to a widespread and strong decrease in new RNA synthesis from human genes. Multiomics analysis reveals that inhibition of U2 snRNP function increases the duration of Pol II pausing in the promoter-proximal region, impairs recruitment of the pause release factor P-TEFb, and reduces Pol II elongation velocity at the beginning of genes. Our results indicate that efficient release of paused Pol II into active transcription elongation requires the formation of functional spliceosomes and that eukaryotic mRNA biogenesis relies on positive feedback from the splicing machinery to the transcription machinery.

INTRODUCTION

In eukaryotic cells, RNA polymerase II (Pol II) transcribes protein-coding genes to synthesize pre-messenger RNAs (pre-mRNAs) that undergo co-transcriptional processing to yield mature mRNAs. Co-transcriptional pre-mRNA processing includes 5′ capping, splicing, 3′ cleavage, and polyadenylation of the pre-mRNA (Bentley, 2014). Co-transcriptional splicing of introns in pre-mRNAs underlies the production of different mRNA isoforms during alternative splicing. Co-transcriptional splicing results from an extensive crosstalk between the transcription machinery, the spliceosome, and the chromatin template (Bentley, 2014; Hsin and Manley, 2012; Kornblihtt et al., 2013).

Indeed, transcription and splicing are intimately linked both physically and functionally (Custódio and Carmo-Fonseca, 2016; Herzel et al., 2017; Neugebauer, 2019; Saldi et al., 2016; Tellier et al., 2020; Wallace and Beggs, 2017). The phosphorylated C-terminal domain (CTD) of Pol II recruits splicing factors (Misteli and Spector, 1999; Mortillaro et al., 1996; Nojima et al., 2018; Vincent et al., 1996) and stimulates splicing *in vitro* (Hirose et al., 1999) and *in vivo* (Fong and Bentley, 2001; McCracken et al., 1997). Mutation of the CTD impairs recruitment of the splicing factor U2AF65 and causes a slower elongation rate in a reporter gene (Gu et al., 2013). Pol II interacts with the catalytically active spliceosome (Harlen et al., 2016; Nojima et al., 2018). Mutations in the elongation factor Spt5 affect splicing (Burckin et al., 2005; Diamant et al., 2012; Lindstrom et al., 2003; Liu et al., 2012; Xiao et al., 2005) and can modulate spliceosome assembly (Maudlin and Beggs, 2019). Splicing occurs in the range of minutes and is limited by the rate of Pol II elongation (Wachutka et al., 2019). Changes in Pol II elongation velocity can alter splicing patterns (Aslanzadeh et al., 2018; Braberg et al., 2013; Fong et al., 2014; Howe et al., 2003; Maslon et al., 2019; de la Mata et al., 2003) and splicing fidelity (Aslanzadeh et al., 2018; Fong et al., 2014). Differences in promoter strength influence alternative splicing (Cramer et al., 1997; Fededa et al., 2005; Kadener et al., 2001). Also, a high Pol II occupancy correlates with increased intron retention (Braunschweig et al., 2014).

Coupling between transcription and splicing is thought to be initiated by the synthesis of nascent RNA that contains elements for spliceosome assembly. After Pol II passes a splice site sequence, the corresponding elements will appear in the nascent RNA that emerges from the Pol II surface and can trigger the binding of small nuclear RNA-protein particles (snRNPs) and assembly of the spliceosome. Passage of Pol II over a 5′ splice site (5′ SS) leads to binding of the U1 snRNP, whereas passage over the branch site (BS) and 3′ splice site (3′ SS) enables binding of U2 snRNP (Kotovic et al., 2003; Lacadie and Rosbash, 2005). Recently, it was shown that transcribing Pol II can directly bind U1 snRNP and retain the 5′ SS near the exit site for nascent pre-mRNA, suggesting the formation of a growing intron loop on Pol II that supports scanning for the BS and the 3′ SS (Zhang et al., 2020a). Subsequent recruitment of U2 snRNP may then lead to formation of the spliceosomal A complex, or pre-spliceosome, before downstream assembly of the complete spliceosome and splicing.

There is also increasing evidence that splicing can influence transcription, but the physical and kinetic basis for this remains unclear. It was long known that the presence of introns can

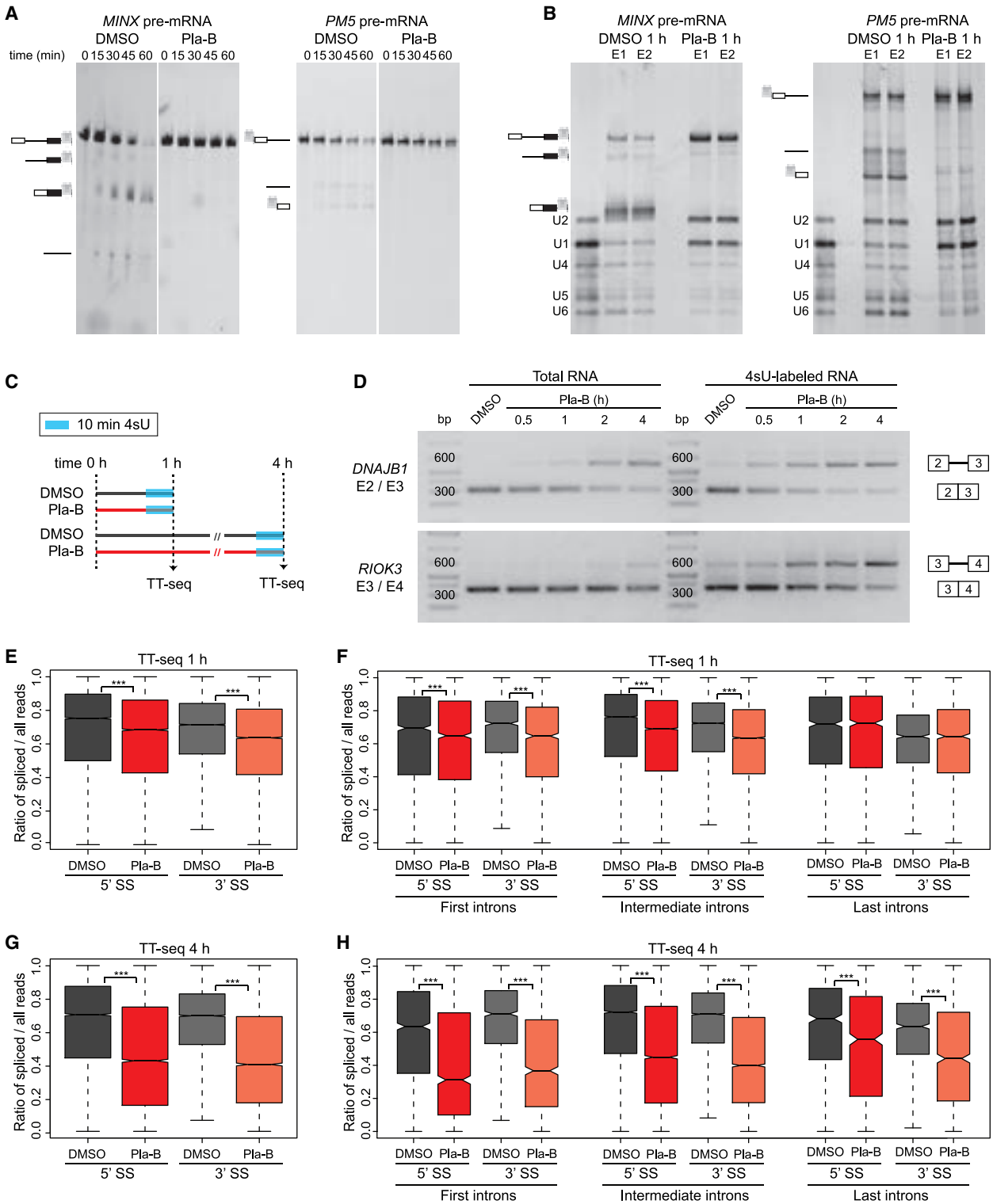


Figure 1. Rapid splicing inhibition in human cells

(A) Tris-Bis PAGE analysis of Cy5-labeled *MINX* and *PM5* pre-mRNAs from *in vitro* splicing reaction in HeLa S3 nuclear extract treated with DMSO or 1 μM Pla-B. The pre-mRNAs and splicing intermediates are indicated on the left. Figures are composites. Splicing reactions were run together with other Pla-B concentrations, but only DMSO and 1 μM Pla-B are shown (see original images in Mendeley repository data).

(legend continued on next page)

increase transcriptional efficiency (Brinster et al., 1988; Furger et al., 2002). Furthermore, the presence of a 5′ SS in the proximity of the promoter can enhance transcription in yeast and mammals (Furger et al., 2002). Insertion of a 5′ SS or a complete intron into an intronless gene enhances mRNA production, possibly because of the recruitment of transcription initiation factors (Damgaard et al., 2008). Deletion of an intron or splicing inhibition with the small-molecule inhibitor spliceostatin A (SSA) (Kaida et al., 2007) lead to an overall reduction of transcriptional output and the transcription-associated trimethylation of histone H3 residue lysine-4 (H3K4me3) (Bieberstein et al., 2012). Splicing inhibition also reduces transcription from nearby promoters and affects transcription start site (TSS) usage (Fiszbein et al., 2019).

Certain splicing factors stimulate transcription elongation, suggesting a spliceosome-dependent checkpoint for elongation. U2 snRNP stimulates elongation *in vitro* by binding of the elongation factor TAT-SF1 that interacts with the positive transcription elongation factor b (P-TEFb) (Fong and Zhou, 2001). The splicing factor SKIP enhances HIV-TAT transactivation through P-TEFb interaction (Brès et al., 2005). The splicing factor SC35 (or SRSF2) influences transcription elongation and P-TEFb recruitment at selected genes (Lin et al., 2008). Depletion of SRSF2 causes Pol II accumulation at gene promoters and impairs P-TEFb recruitment (Ji et al., 2013). Binding of the pre-exon-junction complex component Mago to nascent RNA facilitates Pol II pausing in *Drosophila* (Akhtar et al., 2019). In a yeast strain with a BS mutation, Pol II accumulates over introns (Chathoth et al., 2014). Inhibition of the splicing-associated NUAK1 kinase increases Pol II occupancy near the TSS and at the first exon-intron border (Cossa et al., 2020). The U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation (Kaida et al., 2010), hence determining mRNA length (Berg et al., 2012; Oh et al., 2017). Functional disruption of U1 snRNP increases promoter-proximal RNA cleavage (Almada et al., 2013; Chiu et al., 2018), and promoter-proximal polyadenylation sites reduce transcription activity (Andersen et al., 2013).

Despite these studies, it has not been investigated whether a rapid inhibition of spliceosome function leads to changes in cellular RNA synthesis activity and whether such changes are due to alterations in transcription initiation or elongation. Here we use three different approaches to inhibit U2 snRNP function in human cells and monitor RNA synthesis and splicing genome-wide using transient transcriptome sequencing (TT-seq) (Schwalb et al., 2016; Wachutka et al., 2019). We further monitor changes in Pol II occupancy with mammalian native elongating transcript sequencing (mNET-seq) (Nojima et al., 2015). Together with further experiments, our results show that inhibition of U2 snRNP function decreases RNA synthesis genome-wide, increases Pol II pause duration near the promoter, and decreases Pol II elongation velocity at the beginning of genes. Our results suggest that efficient P-TEFb recruitment and release of paused Pol II into active transcription elongation is stimulated by co-transcriptional formation of functional spliceosomes.

RESULTS

Rapid splicing inhibition in human cells

To specifically inhibit spliceosome assembly and function at an early stage genome-wide in human cells, we used the splicing inhibitor pladienolide B (Pla-B) (Kotake et al., 2007). Pla-B binds the spliceosomal U2 snRNP component SF3B (Cretu et al., 2018; Effenberger et al., 2016; Yokoi et al., 2011), which is essential for BS recognition (Gozani et al., 1996, 1998; Krämer et al., 1999), and inhibits usage of the BS (Cretu et al., 2018; Effenberger et al., 2016), thereby impairing co-transcriptional splicing (Drexler et al., 2020) and perturbing occupancy of genes with transcriptionally engaged Pol II (Nojima et al., 2015, 2018; Schlackow et al., 2017). Pla-B binds to a hinge of the SF3B subunit 1 (SF3B1), blocking SF3B1 in an open conformation that prevents the accommodation and stabilization of the U2 snRNA/BS duplex (Cretu et al., 2018). Consequently, Pla-B leads to a stalled "A-like" spliceosomal complex containing U1 and U2 snRNPs, impairing further spliceosome assembly *in vitro* (Cretu et al., 2018; Effenberger et al., 2016). Consistent with these *in vitro* studies, Pla-B treatment arrests spliceosome assembly *in vivo*, causing cellular mobilization of the later stage spliceosome component U5, but not of U1 and only partially of U2 (Tresini et al., 2015).

Consistent with published results (Cretu et al., 2018; Effenberger et al., 2016), we found that a concentration of 1 μM Pla-B inhibited splicing prior to the first catalytic step (Figure 1A) and inhibited spliceosome assembly after U1 and U2 binding *in vitro* (Figure 1B). To monitor changes in RNA synthesis after spliceosome inhibition, we performed TT-seq and RNA sequencing (RNA-seq) in K562 cells treated with 1 μM Pla-B or

(B) SYBR-gold staining of nuclear acids isolated from pull-down of MS2-MBP spliceosome assembled on *MINX* and *PM5* pre-mRNAs treated with DMSO or 1 μM Pla-B. E1, E2, elutions from amylose. snRNAs marker is indicated on the left. Reactions were run in the same gel, and they share the same snRNAs marker (*MINX* pre-mRNA samples were vertically flipped from the original image; see Mendeley repository data).

(C) Experimental design. TT-seq was performed on K562 cells after 1 and 4 h treatments with DMSO (solvent control) or 1 μM Pla-B using a 10 min 4-thiouridine (4sU) labeling time.

(D) Agarose gel showing spliced and unspliced RT-PCR products for total and 4sU-labeled enriched RNAs upon 30 min, 1 h, 2 h, and 4 h of 1 μM Pla-B and 4 h DMSO treatments for regions spanning exons 2 and 3 of *DNAJB1* (unspliced, 597 bp; spliced, 302 bp) and exons 3 and 4 of *RIOK3* (unspliced, 650 bp; spliced, 352 bp). See also Table S1.

(E) Ratio of spliced reads over total unspliced and spliced reads upon 1 h of DMSO or 1 μM Pla-B treatment.

(F) Ratio of spliced reads over total unspliced and spliced reads upon 1 h of DMSO or 1 μM Pla-B treatment in first, intermediate (non-first and non-last), and last introns. A total of 4,698 major isoforms containing at least four exons were considered in the analysis.

(G) Ratio of spliced reads over total unspliced and spliced reads as in (E) upon 4 h of DMSO or 1 μM Pla-B treatment.

(H) Ratio of spliced reads over total unspliced and spliced reads as in (F) upon 4 h of DMSO or 1 μM Pla-B treatment. Black bars represent the median values for each group. Lower and upper boxes are the first and third quartiles, respectively. The ends of the whiskers extend the box by 1.5 times the interquartile range. Outliers are not drawn.

***$p < 2.2 \times 10^{-16}$ by Wilcoxon signed rank test.

with dimethyl sulfoxide (DMSO) as solvent control for 1 or 4 h (Figure 1C). Cellular proliferation was not affected upon rapid Pla-B treatment (Figure S1A, right). Both Pla-B and DMSO were added to cell media at 1:20,000 dilutions to prevent side effects of DMSO (Verheijen et al., 2019). RNA labeling with 4-thiouridine (4sU) was carried out for 10 min. After 1 h of Pla-B treatment, unspliced RT-PCR products and intron retention were observed (Figure 1D; Figure S1B).

TT-seq monitors splicing inhibition genome-wide

TT-seq and RNA-seq data were generated from two independent biological replicates. On average, we obtained 40 million paired-end reads per sample for TT-seq and 25 million for RNA-seq. TT-seq and RNA-seq data were mapped against the hg38 (GRCh38) genome assembly (STAR methods). The experiments were highly reproducible (Spearman correlation coefficients of 0.9 and 1) (Figures S1C and S1D). TT-seq and RNA-seq data were globally normalized using spike-ins (STAR methods). Because splicing inhibitors may vary in activity at different introns and can cause intron retention or alternative splicing (Corrionero et al., 2011; Teng et al., 2017), we included in our analysis only major isoforms with 70% or higher prevalence per gene in both DMSO and Pla-B conditions (STAR methods). For further analysis, we considered 5,535 major isoforms of protein-coding genes that showed read counts per kilobase (RPK) $\geq$ 50 in TT-seq 1 h solvent control replicates.

We analyzed our TT-seq data for the occurrence of reads spanning exon-intron junctions (5′ SS) and intron-exon junctions (3′ SS) (STAR methods) for a total of 15,551 5′ SS and 14,840 3′ SS after 1 h treatment and for 10,041 5′ SS and 9,550 3′ SS after 4 h treatment. The splicing ratio was calculated by dividing the number of spliced reads by the total amount of spliced and unspliced reads (STAR methods). In accordance with our observations (Figure 1D), the splicing ratio decreased slightly but significantly ($p < 2.2 \times 10^{-16}$, Wilcoxon signed rank test) already after 1 h of Pla-B treatment (Figure 1E). Splicing inhibition was readily visible at the first introns (Figure 1F) and was observed at all introns genome-wide upon 4 h of Pla-B treatment (Figures 1G and 1H), consistent with earlier observations (Effenberger et al., 2014; Nojima et al., 2015). These results show that splicing was rapidly inhibited genome-wide under our experimental conditions.

Inhibition of U2 snRNP function decreases RNA synthesis

Following these findings, we concentrated our analysis on studying the earliest effects of splicing inhibition. We sequenced each sample treated with Pla-B for 1 h at a depth of up to 100 million paired-end reads. Metagene analysis showed that TT-seq coverage across the gene body was strongly reduced (Figures 2A and 2C; Figure S2A). The level of this reduction was not related to the expression level of genes (Figure S2B), arguing against a direct effect of Pla-B on Pol II transcription. Consistent with this, we found that Pla-B did not exhibit effects on Pol II activity *in vitro* (Figure 2F), excluding Pla-B as a Pol II inhibitor. TT-seq coverage was still strongly reduced when the experiment was repeated with only 100 nM Pla-B (Figure S1A, left; Figures S2C–S2F), a concentration that was recently reported to globally diminish co-transcriptional splicing (Drexler et al., 2020). A less pronounced decrease in RNA synthesis was observed at intronless protein-coding genes (Figure S2G), supporting a splicing-independent role of SF3B1 (Van Nostrand et al., 2016; Wang et al., 2019). The drastic effects of Pla-B on RNA synthesis were observed only by TT-seq, not by RNA-seq (Figures S2H–S2J), which only detected the splicing defect through an accumulation of unspliced introns (Figure S2J).

Metagene plots of the TT-seq signal upon 4 h of Pla-B treatment showed a change in the slope toward the 3′ regions of genes (Figures 2B and 2C; Figure S2K), indicating defects in Pol II elongation or processivity. As long human genes can take 1 h or longer to be transcribed, RNA synthesis activity in the 3′ regions of long genes should be less affected upon 1 h of Pla-B inhibition and more affected after 4 h. To investigate this, we divided the major isoforms in different length classes and plotted the ratio of Pla-B to DMSO on the last exons after 1 and 4 h of treatment. As expected, major isoforms longer than 52 kbp showed a smaller change in transcription after 1 h of treatment compared with shorter major isoforms (Figure 2D, left; Figure S2L). In contrast, we observed that RNA synthesis is strongly defective in major isoforms of all lengths after 4 h of Pla-B treatment (Figure 2D, right; Figure S2M). Long major isoforms are more affected than short major isoforms after 4 h of treatment, explaining the slope toward the 3′ region observed in the metagene plot (Figure 2B). To investigate whether the effects of Pla-B on transcription are related to splicing inhibition, we defined two groups of transcripts in which splicing of the first intron was affected or unaffected, on the basis of DESeq2 (Love et al., 2014) analysis (STAR methods). Indeed, RNA synthesis was significantly more decreased for genes where splicing of the first intron was affected (Figure 2E). Together, these results indicate that splicing has a direct positive effect on transcription.

To further support our findings, we performed TT-seq in K562 cells upon splicing inhibition using a different chemical inhibitor, SSA. Similar to Pla-B, SSA targets SF3B1 and inhibits splicing *in vitro* and *in vivo*, interfering with spliceosome assembly after A complex formation (Corrionero et al., 2011; Kaida et al., 2007; Martins et al., 2011; Roybal and Jurica, 2010). Cellular proliferation was not affected upon rapid SSA treatment, similarly to Pla-B treatment (Figure S3A). RT-PCR analysis showed increased intron retention already after 1 h of 30 ng/mL SSA (Figure S3B). TT-seq data were generated from two independent biological replicates (Spearman correlation coefficient = 1; Figure S3C) and globally normalized using spike-ins (STAR methods). On average, 50 million paired-end reads were obtained per sample. We found that SSA inhibited splicing genome-wide after 1 h of treatment (Figure 3A). As for Pla-B treatment, metagene analysis showed that new RNA synthesis was impaired genome-wide (Figures 3B and 3C) and that the decrease in RNA synthesis was significantly higher for genes where splicing was more affected (Figure 3F).

Inhibition of BS recognition decreases RNA synthesis

To exclude that the decrease in transcription after Pla-B or SSA treatments was due to a stress response pathway triggered by

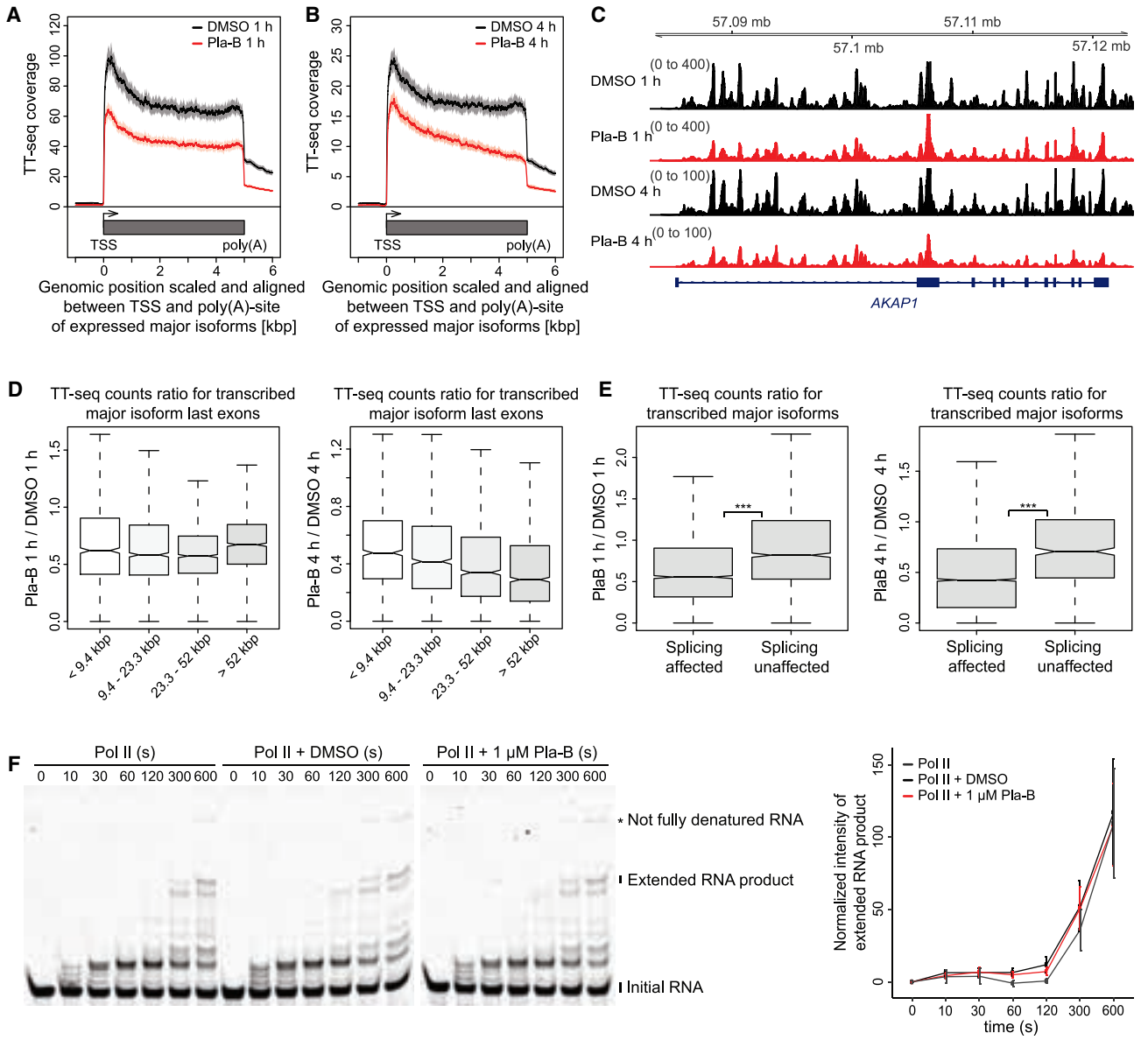


Figure 2. Inhibition of the U2 snRNP factor SF3B1 decreases RNA synthesis

(A and B) Metagene analysis comparing TT-seq signal between cells treated with DMSO or 1 μM Pla-B for 1 h (A) and 4 h (B). TT-seq coverage was averaged for 5,535 major isoforms scaled between TSS and poly(A)-site. See also Figures S2A and S2K. Solid lines represent the averaged signal, and the shaded area represents the 95% confidence interval of the mean (bootstrap).

(C) TT-seq coverage track of *AKAP1* gene upon 1 h and 4 h of DMSO or 1 μM Pla-B treatment (GViz R package).

(D) Ratio of 1 μM Pla-B to DMSO antisense bias-corrected TT-seq read counts for last exons upon 1 h (left) and 4 h (right) treatments in four different major isoform length quartiles.

(E) Ratio of 1 μM Pla-B to DMSO antisense bias-corrected TT-seq read counts for splicing-affected and unaffected major isoforms upon 1 h (left) and 4 h (right) treatments. Outliers are not drawn. ***$p < 2.2 \times 10^{-16}$ by Wilcoxon rank-sum test.

(F) RNA extension assay comparing the transcription activity of Pol II alone, Pol II plus DMSO, and Pol II in the presence of 1 μM Pla-B over a time course of 10 min (left). Figures are composites. Images used for these figures were taken from gels of replicate 2 (see original images in Mendeley repository data). For each replicate, RNA extension assays were performed at the same time, and gels were run in parallel. See also Table S2. Time-series plot comparing the normalized intensity of RNA product (right). Geometric line represents the mean of four biological replicates, and error bars represent the standard deviation, n.s., $p > 0.99$ by two-way ANOVA.

the chemical compounds, we performed TT-seq after treating cells with an antisense morpholino oligo targeting the U2 snRNA (U2 AMO) or with a control oligo (Ctr AMO). We used a U2 AMO that specifically blocks the RNA-RNA interactions between U2 snRNAs and pre-mRNA (Matter and König, 2005). RT-PCR showed that splicing was inhibited after treatment of cells with

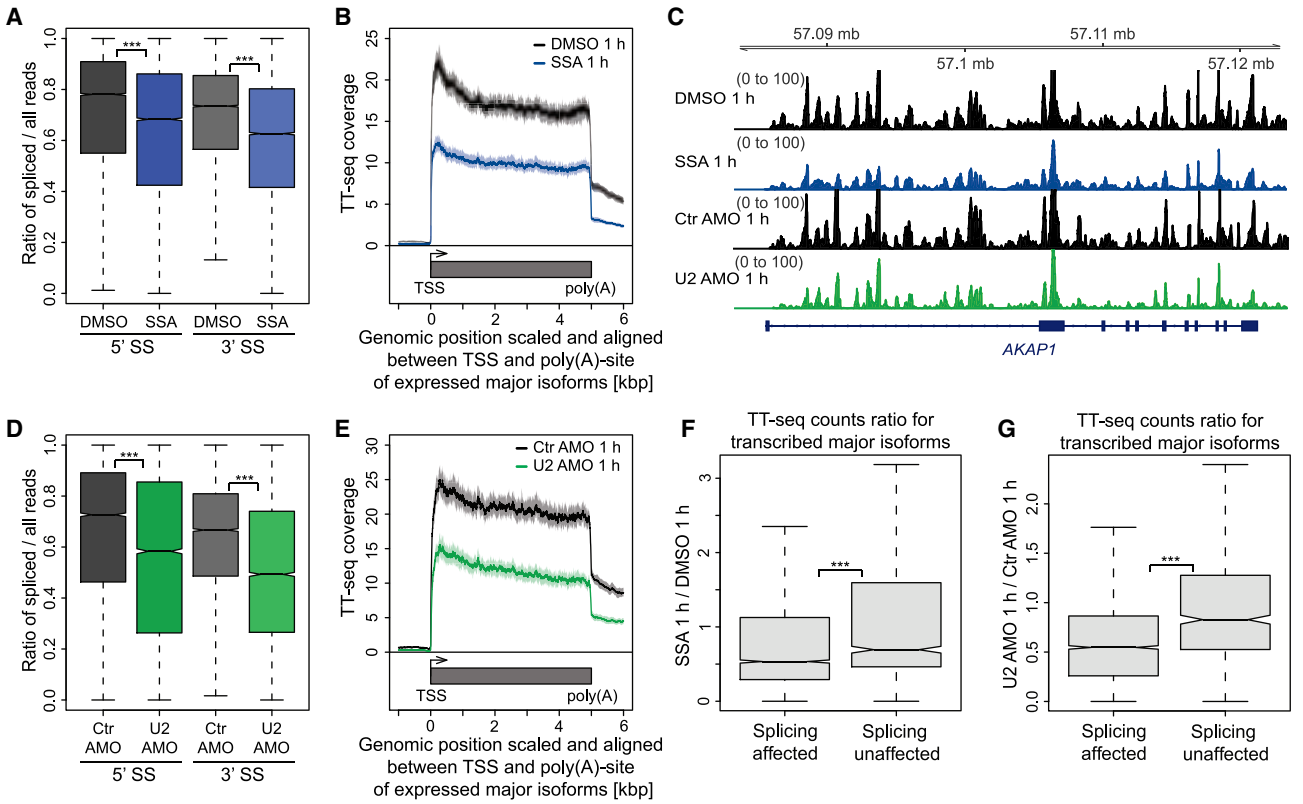


Figure 3. SSA and U2 AMO treatments decrease new RNA synthesis

(A and D) Ratio of spliced reads over total unspliced and spliced reads upon 1 h DMSO or 30 ng/mL SSA (A) and upon 1 h 75 μM Ctr AMO or 75 μM U2 AMO (D). See also Figures S3B and S3D.

(B and E) Metagene analysis comparing TT-seq signal between cells treated with DMSO or 30 ng/mL SSA (B) and 75 μM Ctr AMO or 75 μM U2 AMO (E) for 1 h. TT-seq coverage was averaged for 5,535 major isoforms scaled between TSS and poly(A)-site. Solid lines represent the averaged signal, and the shaded area represents the 95% confidence interval of the mean (bootstrap).

(C) TT-seq coverage track of *AKAP1* gene upon 1 h DMSO or 30 ng/mL SSA and upon 1 h 75 μM Ctr AMO or 75 μM U2 AMO (GViz R package).

(F and G) Ratio of 30 ng/mL SSA to DMSO (F) and 75 μM Ctr AMO to 75 μM U2 AMO (G) antisense bias-corrected TT-seq read counts for splicing-affected and unaffected major isoforms upon 1 h. Outliers are not drawn.

***$p < 2.2 \times 10^{-16}$ by Wilcoxon rank-sum test.

75 μM U2 AMO for 1 h (Figure S3D), indicating rapid splicing inhibition. TT-seq data were generated from two biological replicates after treatment with U2 AMO or Ctr AMO (Spearman correlation coefficient = 1; Figure S3E) and globally normalized using spike-ins (STAR methods). On average, 40 million paired-end reads were obtained per sample. The data showed that 1 h treatment with U2 AMO causes a decrease in new RNA synthesis across the gene body (Figures 3C and 3E) and inhibits splicing genome-wide (Figure 3D). Again, the decrease in newly synthesized RNA was significantly higher for genes where splicing was more affected (Figure 3G).

These results suggested that the function of U2 snRNP in BS recognition is required for normal RNA synthesis. U2 snRNP interacts with the BS and this relies on binding of the accessory factor U2AF to the 3′ SS (Ruskin et al., 1988; Valcárcel et al., 1996; Zamore and Green, 1989). Mutation of the 3′ SS consequently impairs U2 snRNP binding and spliceosome assembly (Bindereif and Green, 1986; Frendewey and Keller, 1985; Lamond et al., 1987; Ruskin and Green, 1985). To test if mutation of the BS and 3′ SS leads to a decrease in RNA synthesis, we used a plasmid-based assay (STAR methods). We transfected K562 cells with a pCMV plasmid carrying the intron-containing *GLuc 2* model gene with either wild-type consensus sequences for BS and 3′ SS (WT), a mutated BS (mut BS), a mutated 3′ SS (mut 3′ SS), or mutated BS and 3′ SS (mut BS-3′ SS) (Figures S3F and S3G). We found that mutation of the BS and 3′ SS decreased splicing as expected, as measured using RT-PCR (Figures S3F and S3G). Mutation of only the BS showed a weaker effect on splicing, apparently due to the use of a cryptic BS (Ruskin et al., 1985). Strikingly, mutations of the BS and 3′ SS affected not only splicing but also transcription of the model gene, as seen from a decrease in RNA signal from the flanking exons (Figure S3G). The decrease in transcription was higher for mutations that were more affected in splicing. Taken together, these results suggested that U2 snRNP binding to the BS in pre-mRNA is required for efficient Pol II transcription, in particular for transcription in the region upstream of the intron.

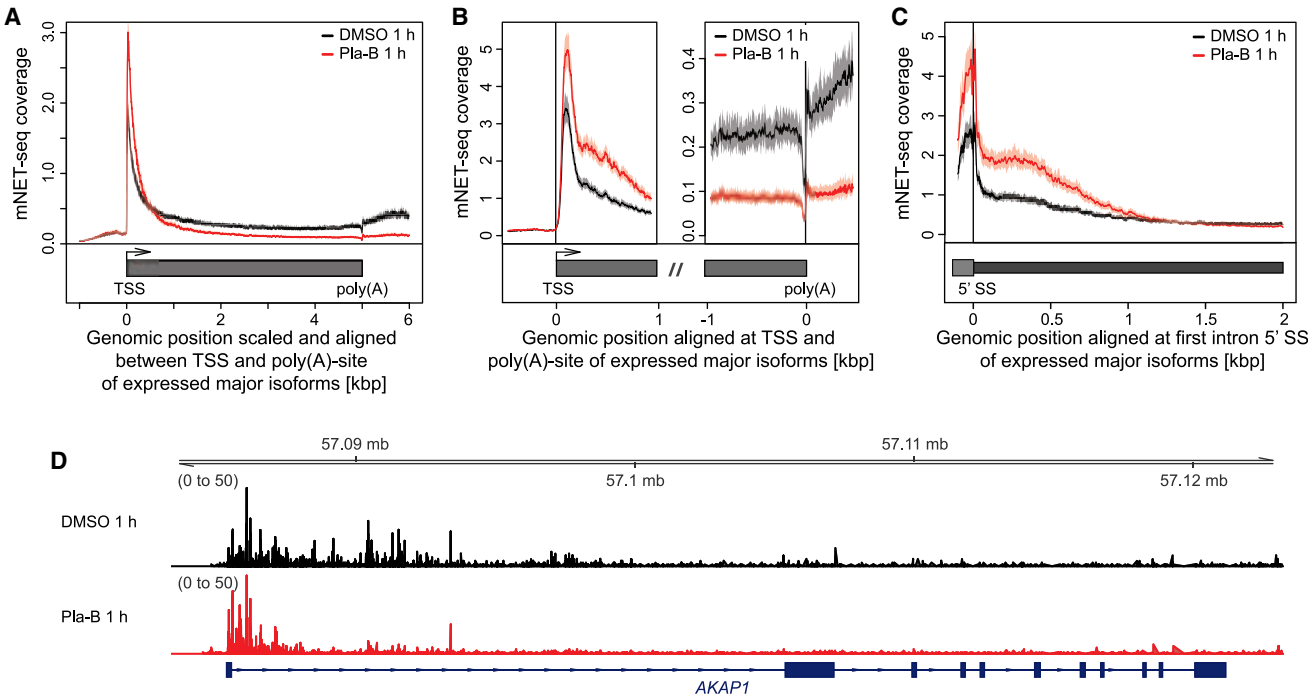


Figure 4. Inhibition of U2 snRNP function impairs early transcription elongation

(A) Metagene analysis comparing the mNET-seq signal between cells treated with DMSO or 1 μM Pla-B for 1 h. mNET-seq coverage was averaged for 5,535 major isoforms scaled between TSS and poly(A)-site. Solid lines represent the averaged signal, and the shaded area represents 95% confidence interval of the mean (bootstrap). See also Figure S4B.

(B) Metagene analysis comparing the mNET-seq signal upon 1 h DMSO or 1 μM Pla-B-treated cells aligned at TSS (left) and poly(A)-site (right). mNET-seq signal was averaged for 5,465 major isoforms that exceed 1 kbp in length.

(C) Metagene analysis comparing mNET-seq signal upon 1 h DMSO or 1 μM Pla-B-treated cells aligned at the first 5′ SS. mNET-seq signal was averaged for 3,449 first introns that exceed 2 kbp in length. See also Figure S4D.

(D) mNET-seq coverage track of *AKAP1* gene upon 1 h of DMSO or 1 μM Pla-B treatment (GViz R package).

Inhibition of U2 snRNP function impairs early transcription elongation

Our results indicated that the observed reduced RNA synthesis activity resulted from a decreased number of Pol II enzymes entering productive and processive elongation of genes. To investigate this, we performed mNET-seq of total Pol II in K562 cells treated for 1 h with 1 μM Pla-B or DMSO. The detergent Empigen was used to prevent co-immunoprecipitation of spliceosome-associated RNAs (Nojima et al., 2018; Schlackow et al., 2017). mNET-seq data were generated from two independent biological replicates and were globally normalized using a reference (spike-ins) *S. cerevisiae* genome (STAR methods). Between 114 million and 138 million paired-end reads were obtained per sample. mNET-seq reads were mapped to the human hg38 (GRCh38) genome (STAR methods). Data were highly reproducible (Spearman correlation coefficient = 1) (Figure S4A).

Metagene plots of the mNET-seq signals over major isoforms revealed a sharp peak of Pol II occupancy just downstream of the TSS, indicative of promoter-proximally paused Pol II (Figure 4A, black line). Upon 1 h of Pla-B treatment, the peak for paused Pol II strongly increased, and the signal in the gene body decreased (Figure 4A, red line; Figure 4D). We observed the same trend when we used published mNET-seq data obtained for ser5-phosphorylated (ser5P) Pol II in HeLa cells after 4 h of Pla-B treatment (Nojima et al., 2015) (Figure S4B). At intronless genes, Pol II occupancy increased also, but to a lower extent (Figure S4C). These changes in the mNET-seq signal indicate an increase in promoter-proximal Pol II pausing and a defect in the release of Pol II into elongation.

More detailed analysis of the mNET-seq data showed that Pol II accumulated not only in the promoter-proximal region but also further downstream in the 5′ region of the first intron (Figures 4B and 4C). At about 1.3 kbp downstream of the first 5′ SS, the mNET-seq signal, however, dropped below the level observed without Pla-B treatment (Figures 4A and 4C). These results are consistent with a strong defect in productive and processive Pol II elongation. Changes in Pol II occupancy were by far strongest in the beginning of the first intron (Figure 4C) and were independent of the length of the first intron (Figure S4D). These results argue against a simple model that Pol II that is paused around the 3′ end of the first intron because of accumulation of elongating Pol II complexes upstream all the way back to the promoter. Instead, our findings are consistent with a role of the functional spliceosome in facilitating the release of promoter-proximally paused Pol II into active elongation.

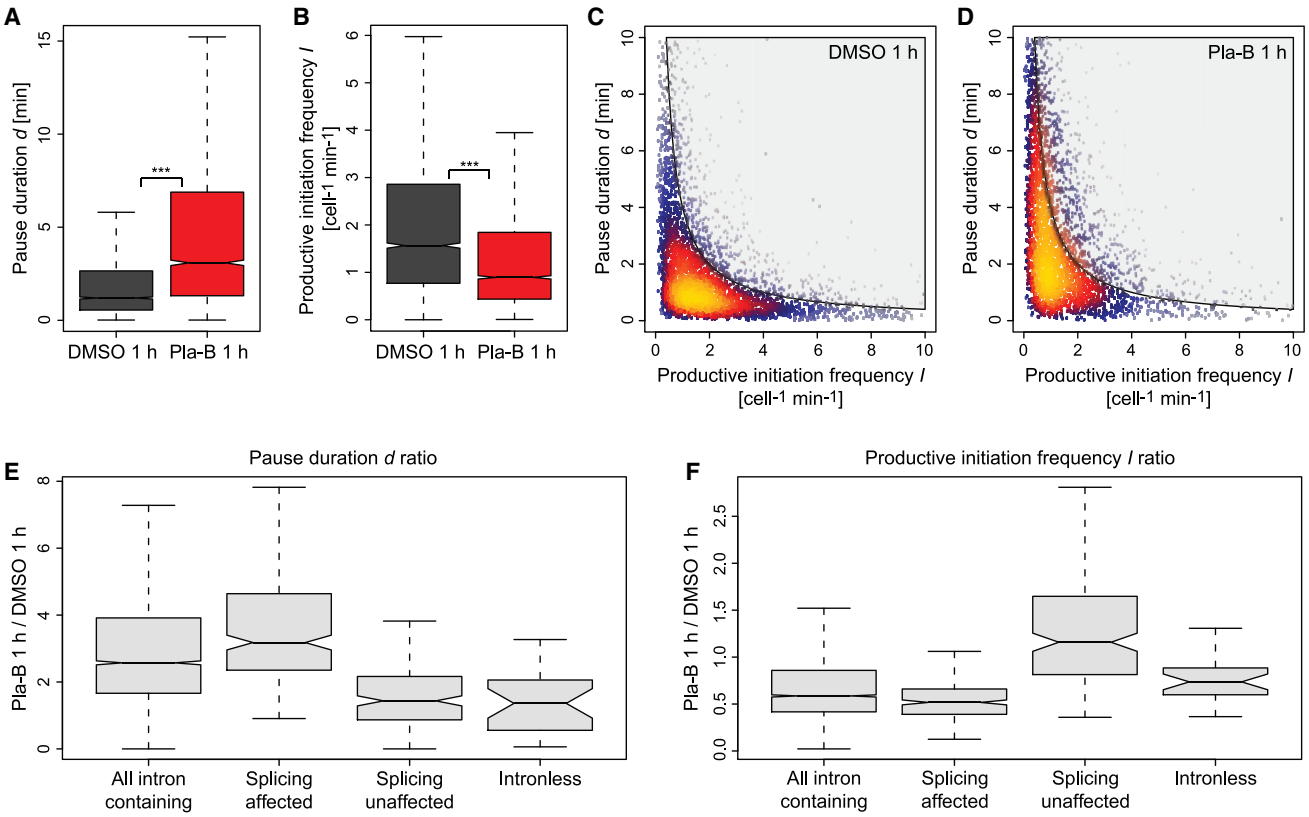


Figure 5. Inhibition of U2 snRNP function increases Pol II pause duration

(A and B) Pause duration *d* (A) and productive initiation frequency *I* (B) calculated for 3,636 intron-containing major isoforms that exceed 10 kbp in length and had an identified polymerase pause position (STAR methods) for both 1 h DMSO and 1 μM Pla-B. Outliers are not drawn. ***$p < 2.2 \times 10^{-16}$ by Wilcoxon signed rank test.

(C and D) Correlation between the productive initiation frequency *I* and pause duration *d* upon 1 h DMSO (C) and 1 μM Pla-B (D) treated cells. The gray shaded area depicts impossible combinations of *I* and *d* considering the footprint of the transcribing elongation complexes a limit distance between the active sites of initiating and paused Pol II (Ehrensberger et al., 2013).

(E and F) Ratio of Pla-B to DMSO for pause duration *d* (E) and productive initiation frequency *I* (F) upon 1 h of DMSO or 1 μM Pla-B treatment calculated for 3,636 intron-containing, 329 splicing-affected, 355 splicing-unaffected, and 29 intronless transcripts. Note that not all genes fit the criteria for pausing and productive initiation frequency calculation (STAR methods).

Inhibition of U2 snRNP function increases Pol II pause duration

We next investigated whether Pla-B treatment increased the duration of promoter-proximal pausing by Pol II. We defined Pol II pause sites as the positions of the major mNET-seq peaks with a signal at least 5 times above the median coverage in the first exon of the intron-containing major isoforms or in the first 100 bp of the intronless genes (STAR methods; Gressel et al., 2017). The distribution of genomic positions of these pause sites was not altered upon Pla-B treatment (Figure S5A). We then used our previously described approach that combines multiomics with kinetic modeling (Gressel et al., 2019; Gressel et al., 2017) to estimate the pause duration *d* as the ratio of the mNET-seq signal over the productive initiation frequency *I* in a pause window of ±100 bp (intron-containing major isoforms) or ±50 bp (intronless genes) around the Pol II pause site (STAR methods).

This quantitative kinetic analysis showed that the pause duration *d* was strongly increased genome-wide after Pla-B treatment (Figure 5A). Thus, Pol II pauses longer in the promoter-proximal region when U2 snRNP function is compromised. Furthermore, the productive initiation frequency *I* decreased after Pla-B treatment (Figure 5B), confirming the transcription defect. Plotting *d* and *I* for all expressed intron-containing major mRNA isoforms exceeding 10 kbp in length showed a shift to longer pause durations and to lower productive initiation frequencies upon 1 h of Pla-B treatment (Figures 5C and 5D). These observations can be explained by the existence of the previously described "pause-initiation limit" in mammalian cells that sets a limit to the maximum Pol II initiation frequency at a given pause duration (Gressel et al., 2017; Shao and Zeitlinger, 2017). Consistent with our previous results, splicing-affected genes showed significantly more increased pause durations and more decreased initiation frequencies than splicing-unaffected genes or intronless genes (Figures 5E and 5F; $p < 0.001$, Wilcoxon rank-sum test; STAR methods). In summary, Pla-B treatment leads to an increase of the duration of promoter-proximal Pol II pausing and to lower productive initiation frequencies at many genes.

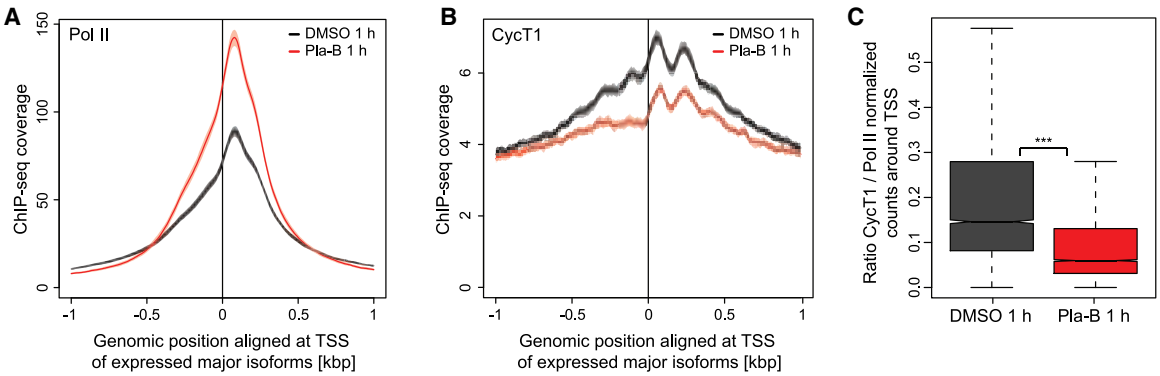


Figure 6. Inhibition of U2 snRNP function impairs P-TEFb recruitment

(A and B) Metagene analysis comparing ChIP-seq for Pol II (A) and CycT1 (B) between cells treated with DMSO and 1 μM Pla-B for 1 h. ChIP-seq coverage was aligned at the TSS and averaged for 5,465 major isoforms that exceed 1 kbp in length. Solid lines represent the averaged signal, and the shaded area represents the 95% confidence interval of the mean (bootstrap).

(C) Ratio of CycT1 to Pol II normalized counts around TSS (from −100 to +200 bp) comparing 1 h DMSO- and 1 μM Pla-B-treated samples. Outliers are not drawn. ***$p < 2.2 \times 10^{-16}$ by Wilcoxon signed rank test.

Inhibition of U2 snRNP function impairs recruitment of pause release factor P-TEFb

We next investigated the reasons for the observed increase in the duration of promoter-proximal Pol II pausing. Release of paused Pol II into active elongation requires P-TEFb, which contains the kinase CDK9 and the cyclin T1 (CycT1) (Fujinaga et al., 2004; Marshall et al., 1996; Marshall and Price, 1995; Vos et al., 2018; Wei et al., 1998). We therefore investigated whether increased Pol II proximal-promoter pausing after Pla-B treatment may be accompanied by a defect in P-TEFb recruitment. We used chromatin immunoprecipitation followed by sequencing (ChIP-seq) to measure genome occupancy with total Pol II and with the P-TEFb subunit CycT1 in K562 cells treated for 1 h with 1 μM Pla-B or DMSO. ChIP-seq data were generated from two independent biological replicates and globally normalized using a reference (spike-ins) *D. melanogaster* genome (STAR methods). On average, 30 million to 50 million paired-end reads were obtained per sample. ChIP-seq data were mapped against the hg38 (GRCh38) genome assembly (STAR methods). Data were highly reproducible (Spearman correlation coefficient = 0.99) (Figure S6A).

Analysis of these ChIP-seq data showed that Pla-B treatment resulted in increased Pol II occupancy just downstream of the TSS of protein-coding genes (Figure 6A). This is consistent with the increase in promoter-proximally paused Pol II that we had observed by mNET-seq (Figure 4). In contrast, occupancy with the P-TEFb subunit CycT1 was decreased in the same region (Figures 6B and 6C), indicating that the recruitment of P-TEFb to promoter-proximally paused Pol II is impaired. Consistent with our previous evidence, intronless genes showed similar effects, but they were much less pronounced (Figures S6B and S6C). These data indicate that U2 snRNP inhibition impairs P-TEFb recruitment to the promoter-proximal region, thereby impairing release of paused Pol II, increasing Pol II pause duration, and decreasing the productive initiation frequency at affected genes.

Inhibition of U2 snRNP function alters Pol II elongation velocity

To investigate in more detail how U2 snRNP inhibition influences Pol II elongation downstream of the promoter-proximal region and downstream of the first intron, we estimated the Pol II elongation velocity at all major isoforms exceeding 10 kbp in length. We obtained the elongation velocity *v* from the ratio of the number of Pol II enzymes released into elongation, as measured using TT-seq at each genomic location, over the Pol II occupancy, as measured using mNET-seq (Figure 7A). Metagene analysis of the cells treated with DMSO showed that the elongation velocity is low in the region close to the TSS and then increases, reaching a maximum of around 3 kbp/min (Figure 7B, black line). This agrees with previous estimates for Pol II velocity (Fuchs et al., 2014; Gressel et al., 2017; Jonkers et al., 2014; Saponaro et al., 2014; Veloso et al., 2014), supporting our approach.

After 1 h of Pla-B treatment, we observed a decrease in the elongation velocity in the promoter-proximal region and in the beginning of the first intron, whereas regions further downstream were less affected (Figure 7B; Figure S7A). This is consistent with an unaltered elongation velocity for Pol II enzymes that were released before Pla-B treatment. In accordance with this and with our previous observations, short genes were already affected throughout their entire length after 1 h of treatment (Figure S7B). We also found that the average elongation velocity in the first exon was much lower than in the first intron (Figure S7C), reflecting Pol II pausing in the promoter-proximal region. Elongation velocity was also affected upon Pla-B inhibition in intronless genes, but to a much lower extent (Figure S7D). After 4 h of Pla-B treatment, the elongation velocity decreased in all genes, as seen from a corresponding multiomics analysis with the use of published mNET-seq data (Figure 7C). In summary, these results argue against a direct effect of Pla-B on released and processive Pol II elongation complexes. Instead, the results are consistent with a direct effect of spliceosome function on Pol II release from the promoter-proximal region into efficient downstream elongation.

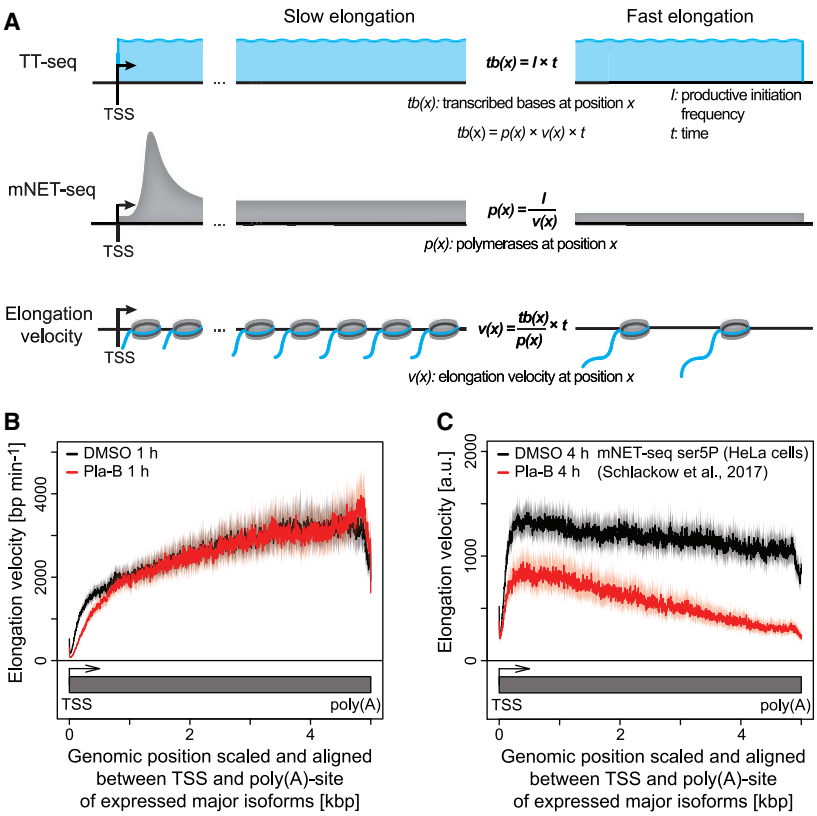


Figure 7. Inhibition of U2 snRNP function alters Pol II elongation velocity

(A) Schematic representation of Pol II elongation velocity *v*. The TT-seq signal (top) corresponds to the productive initiation frequency *I*. The mNET-seq signal (middle) corresponds to the ratio of the productive initiation frequency *I* to the elongation velocity *v*. Therefore, *v* (bottom) can be calculated by dividing the TT-seq signal over the mNET-seq signal.

(B) Metagene analysis comparing the elongation velocity in cells treated with DMSO or 1 μM Pla-B for 1 h. Elongation velocity was averaged for 5,535 major isoforms scaled between TSS and poly(A)-site. Solid lines represent the averaged signal, and the shaded area represents the 95% confidence interval of the mean (bootstrap). See also Figures S7B and S7C.

(C) Metagene analysis comparing the elongation velocities calculated in cells treated with DMSO or 1 μM Pla-B for 4 h as in (B). Datasets for mNET-seq against CTD ser5P in HeLa cells upon 4 h DMSO and 1 μM Pla-B treatments were obtained from Schlackow et al. (2017).

DISCUSSION

Here we rapidly inactivated U2 snRNP function in human cells and monitored Pol II activity, pause duration, and elongation velocity genome-wide. We found that U2 snRNP function is important for efficient release of paused Pol II into genes for active transcription elongation. Although previous studies had already shown that splicing inhibition has consequences for Pol II occupancy and transcription, our data now define these consequences genome-wide and at the level of transcription kinetics. Our results suggest that U2 snRNP function is important for the establishment of complete, RNA processing-competent, active Pol II elongation complexes in the promoter-proximal region of transcribed human genes. We refer to this process as "U2 snRNP-dependent Pol II elongation activation."

The mechanism underlying U2 snRNP-dependent Pol II elongation activation remains to be elucidated. However, recent structural studies suggest a possible model for this process. It was recently shown that U1 snRNP directly interacts with elongating Pol II, and this interaction requires neither nascent RNA nor a 5′ SS (Zhang et al., 2020a). Modeling further showed that U2 snRNP can be accommodated in addition to U1 snRNP to form the spliceosomal A complex on the Pol II surface, whereas conversion to the splicing-active B complex would separate Pol II and the spliceosome (Zhang et al., 2020a). How such separation may activate further release of paused Pol II far upstream remained unclear. Here we speculate that Pla-B arrests co-transcriptional spliceosome assembly at the stage of A complex formation on Pol II and impairs the separation of transcribing Pol II and the spliceosome.

Previous results suggest that U2 snRNP-dependent Pol II elongation activation may involve elongation factor TAT-SF1. TAT-SF1 can bind the U2 snRNP factor SF3B1 and stimulate elongation *in vitro* (Chen et al., 2009; Fong and Zhou, 2001; Loerch et al., 2019). TAT-SF1 also associates with Pol II and SPT5 (Kim et al., 1999). TAT-SF1 further interacts with CycT1 (Fong and Zhou, 2001) and co-immunoprecipitates with U2 snRNP components *in vivo* (Abramczuk et al., 2017) and *in vitro* (Agafonov et al., 2011). The yeast TAT-SF1 homolog Cus-2 interacts with PRP11, a subunit of yeast SF3a (Yan et al., 1998) and SF3b1 (Talkish et al., 2019). Consistent with our results, mutation of the BS causes Cus-2-dependent Pol II accumulation over introns (Chathoth et al., 2014). It was recently shown that TAT-SF1 binds SF3B1 in the open conformation and that displacement of TAT-SF1 must occur to allow stable U2 snRNP-BS interaction (Zhang et al., 2020b). U2 snRNP-dependent Pol II pause release may additionally involve the splicing factor SRSF2 (or SC35) that is required for ATP-dependent interaction between U2 snRNP and the BS (Fu and Maniatis, 1992) and for normal Pol II occupancy and P-TEFb recruitment (Ji et al., 2013; Lin et al., 2008).

Our results further showed that inhibition of U2 snRNP function also decreased RNA synthesis from intronless genes, albeit to a lower extent. Although U1 snRNP is poorly detectable at selected intronless genes (Listerman et al., 2006), it can be recruited to transcription units of splicing-deficient reporter genes (Spiluttini et al., 2010). Moreover, direct binding of U1 snRNP to the Pol II elongation complex does not require the presence of a 5′ SS in nascent pre-mRNA (Zhang et al., 2020a). Furthermore, U2 snRNP can also be recruited to intronless genes (Van Nostrand et al., 2016; Wang et al., 2019) and has a role in 3′-RNA

processing at intronless histone genes (Friend et al., 2007; Kyburz et al., 2006). Thus, our results may be explained with the presence of low levels of U1 and U2 snRNPs at intronless genes. However, we cannot entirely exclude indirect effects, as splicing inhibitors such Pla-B and SSA may induce alternative splicing (Corrionero et al., 2011; Teng et al., 2017). Nevertheless, the rapid inhibition of U2 snRNP function and fast transcription readout in our experiments is expected to avoid such indirect effects. In conclusion, we provide strong evidence that efficient Pol II elongation activation requires U2 snRNP function and speculate that this is due to co-transcriptional assembly of functional spliceosomes.

Limitations

Our study is based on K562 cells, a human immortalized myelogenous leukemia cell line. We cannot exclude that we miss in our analysis signals for some RNA isoforms that are not covered by the RefSeq annotation. This might affect the definition of our major isoforms and intronless genes and the assessment of genomic positions, including TSS, 5′ SS, 3′ SS, and poly(A)-site. However, it is not expected to alter any of our conclusions because we selected only major isoforms with $\geq$70% prevalence per gene, and we excluded intronless genes with UTRs longer than 100 bp. Pla-B could in principle have effects on transcription, but this is unlikely because Pol II activity is not compromised by Pla-B *in vitro* and because orthogonal *in vivo* approaches argue against this. Nevertheless, we cannot exclude possible alternative Pla-B targets. Pla-B treatment will also inhibit the formation of later stages of spliceosome assembly. Although cellular proliferation was not influenced, we cannot entirely exclude that Pla-B has a minor effect on cell viability. With respect to the TT-seq method, we cannot entirely exclude the presence of a small amount of unlabeled RNAs. However, we used spike-ins to estimate and correct for minor cross-contamination by pre-existing RNAs. mNET-seq involves immunoprecipitation, and the quality of the results relies on the specificity of the antibody used. To investigate the mechanism of U2 snRNP-dependent Pol II elongation activation, future studies should test whether and how splicing and elongation factors contribute to P-TEFb recruitment.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Employed cell lines
  - Cell culture treatments
- METHOD DETAILS
  - Total and 4sU-labeled RNA extraction and semiquantitative-PCR
  - *In vitro* splicing assay
  - MS2 affinity purification of spliceosomal complexes
  - TT-seq and RNA-seq
  - Proliferation assay
  - Morpholino inhibition
  - RNA extension assay
  - Plasmid construction for reporter gene splicing assay
  - Reporter gene splicing assay
  - mNET-seq
  - RNA spike-ins for mNET-seq
  - ChIP-seq
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - TT-seq and RNA-seq data preprocessing and normalization
  - Major isoform annotation
  - Intronless genes annotation
  - Splicing ratio
  - Identification of splicing-affected and unaffected transcripts
  - RNA amount per cell
  - Detection of Pol II pause sites
  - Productive initiation frequency
  - mNET-seq preprocessing and normalization
  - Pause duration
  - ChIP-seq data preprocessing and normalization
  - Elongation velocity estimation

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.molcel.2021.02.016.

ACKNOWLEDGMENTS

We thank Kerstin Maier and Petra Rus for sequencing. We thank Suyang Zhang and Kristina Žumer for critical reading of the manuscript. We are deeply grateful to Vladimir Pena and Kristina Žumer for stimulating and insightful discussions. We thank Vladimir Pena for providing SSA. We thank Leonhard Wachutka and Julien Gagneur for some initial data analysis. L.C. was supported by a European Molecular Biology Organization (EMBO) Long-Term Postdoctoral Fellowship (ALTF-1261-2014). S.M.-M. and K.L. were supported by the International Max Planck Research School for Genome Science (S.M.-M.) and the International Max Planck Research School for Molecular Biology (K.L.), part of the Göttingen Graduate School for Neurosciences, Biophysics, and Molecular Biosciences. A.S. was supported by an EMBO Long-Term Postdoctoral Fellowship (ALTF-519-2019). P.C. was supported by Deutsche Forschungsgemeinschaft (SFB860, SPP1935, EXC 2067/1-390729940).

AUTHOR CONTRIBUTIONS

Conceptualization, L.C. and P.C.; Methodology, L.C. and B.S.; Formal Analysis, S.M.-M.; Investigation, L.C., K.L., J.S., A.S., and Y.C.; Data Curation, S.M.-M.; Writing – Original Draft, L.C., S.M.-M., and P.C.; Writing – Review & Editing, L.C., S.M.-M., and P.C.; Visualization, S.M.-M., L.C., and P.C.; Supervision, L.C., B.S., M.L., and P.C.; Project Administration, P.C.; Funding Acquisition, L.C., A.S., and P.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: April 17, 2020
Revised: January 7, 2021
Accepted: February 10, 2021
Published: March 8, 2021

REFERENCES

Abramczuk, M.K., Burkard, T.R., Rolland, V., Steinmann, V., Duchek, P., Jiang, Y., Wissel, S., Reichert, H., and Knoblich, J.A. (2017). The splicing co-factor Barricade/Tat-SF1, is required for cell cycle and lineage progression in Drosophila neural stem cells. Development *144*, 3932–3945.

Agafonov, D.E., Deckert, J., Wolf, E., Odenwälder, P., Bessonov, S., Will, C.L., Urlaub, H., and Lührmann, R. (2011). Semiquantitative proteomic analysis of the human spliceosome via a novel two-dimensional gel electrophoresis method. Mol. Cell. Biol. *31*, 2667–2682.

Akhtar, J., Kreim, N., Marini, F., Mohana, G., Brüne, D., Binder, H., and Roignant, J.-Y. (2019). Promoter-proximal pausing mediated by the exon junction complex regulates splicing. Nat. Commun. *10*, 521.

Almada, A.E., Wu, X., Kriz, A.J., Burge, C.B., and Sharp, P.A. (2013). Promoter directionality is controlled by U1 snRNP and polyadenylation signals. Nature *499*, 360–363.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Andersen, P.R., Domanski, M., Kristiansen, M.S., Storvall, H., Ntini, E., Verheggen, C., Schein, A., Bunkenborg, J., Poser, I., Hallais, M., et al. (2013). The human cap-binding complex is functionally connected to the nuclear RNA exosome. Nat. Struct. Mol. Biol. *20*, 1367–1376.

Aslanzadeh, V., Huang, Y., Sanguinetti, G., and Beggs, J.D. (2018). Transcription rate strongly affects splicing fidelity and cotranscriptionality in budding yeast. Genome Res. *28*, 203–213.

Bentley, D.L. (2014). Coupling mRNA processing with transcription in time and space. Nat. Rev. Genet. *15*, 163–175.

Berg, M.G., Singh, L.N., Younis, I., Liu, Q., Pinto, A.M., Kaida, D., Zhang, Z., Cho, S., Sherrill-Mix, S., Wan, L., and Dreyfuss, G. (2012). U1 snRNP determines mRNA length and regulates isoform expression. Cell *150*, 53–64.

Bessonov, S., Anokhina, M., Will, C.L., Urlaub, H., and Lührmann, R. (2008). Isolation of an active step I spliceosome and composition of its RNP core. Nature *452*, 846–850.

Bieberstein, N.I., Carrillo Oesterreich, F., Straube, K., and Neugebauer, K.M. (2012). First exon length controls active chromatin signatures and transcription. Cell Rep. *2*, 62–68.

Bindereif, A., and Green, M.R. (1986). Ribonucleoprotein complex formation during pre-mRNA splicing in vitro. Mol. Cell. Biol. *6*, 2582–2592.

Braberg, H., Jin, H., Moehle, E.A., Chan, Y.A., Wang, S., Shales, M., Benschop, J.J., Morris, J.H., Qiu, C., Hu, F., et al. (2013). From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. Cell *154*, 775–788.

Braunschweig, U., Barbosa-Morais, N.L., Pan, Q., Nachman, E.N., Alipanahi, B., Gonatopoulos-Pournatzis, T., Frey, B., Irimia, M., and Blencowe, B.J. (2014). Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res. *24*, 1774–1786.

Brès, V., Gomes, N., Pickle, L., and Jones, K.A. (2005). A human splicing factor, SKIP, associates with P-TEFb and enhances transcription elongation by HIV-1 Tat. Genes Dev. *19*, 1211–1226.

Brinster, R.L., Allen, J.M., Behringer, R.R., Gelinas, R.E., and Palmiter, R.D. (1988). Introns increase transcriptional efficiency in transgenic mice. Proc. Natl. Acad. Sci. U S A *85*, 836–840.

Burckin, T., Nagel, R., Mandel-Gutfreund, Y., Shiue, L., Clark, T.A., Chong, J.-L., Chang, T.-H., Squazzo, S., Hartzog, G., and Ares, M., Jr. (2005). Exploring functional relationships between components of the gene expression machinery. Nat. Struct. Mol. Biol. *12*, 175–182.

Chathoth, K.T., Barrass, J.D., Webb, S., and Beggs, J.D. (2014). A splicing-dependent transcriptional checkpoint associated with prespliceosome formation. Mol. Cell *53*, 779–790.

Chen, Y., Yamaguchi, Y., Tsugeno, Y., Yamamoto, J., Yamada, T., Nakamura, M., Hisatake, K., and Handa, H. (2009). DSIF, the Paf1 complex, and Tat-SF1 have nonredundant, cooperative roles in RNA polymerase II elongation. Genes Dev. *23*, 2765–2777.

Chiu, A.C., Suzuki, H.I., Wu, X., Mahat, D.B., Kriz, A.J., and Sharp, P.A. (2018). Transcriptional pause sites delineate stable nucleosome-associated premature polyadenylation suppressed by U1 snRNP. Mol. Cell *69*, 648–663.e7.

Conrad, T., and Ørom, U.A. (2017). Cellular fractionation and isolation of chromatin-associated RNA. Methods Mol. Biol. *1468*, 1–9.

Corrionero, A., Miñana, B., and Valcárcel, J. (2011). Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. Genes Dev. *25*, 445–459.

Cossa, G., Roeschert, I., Prinz, F., Baluapuri, A., Vidal, R.S., Schülein-Völk, C., Chang, Y.-C., Ade, C.P., Mastrobuoni, G., Girard, C., et al. (2020). Localized inhibition of protein phosphatase 1 by NUAK1 promotes spliceosome activity and reveals a MYC-sensitive feedback control of transcription. Mol. Cell *77*, 1322–1339.e11.

Cramer, P., Pesce, C.G., Baralle, F.E., and Kornblihtt, A.R. (1997). Functional association between promoter structure and transcript alternative splicing. Proc. Natl. Acad. Sci. U S A *94*, 11456–11460.

Cretu, C., Agrawal, A.A., Cook, A., Will, C.L., Fekkes, P., Smith, P.G., Lührmann, R., Larsen, N., Buonamici, S., and Pena, V. (2018). Structural basis of splicing modulation by antitumor macrolide compounds. Mol. Cell *70*, 265–273.e8.

Custódio, N., and Carmo-Fonseca, M. (2016). Co-transcriptional splicing and the CTD code. Crit. Rev. Biochem. Mol. Biol. *51*, 395–411.

Damgaard, C.K., Kahns, S., Lykke-Andersen, S., Nielsen, A.L., Jensen, T.H., and Kjems, J. (2008). A 5′ splice site enhances the recruitment of basal transcription initiation factors in vivo. Mol. Cell *29*, 271–278.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative splicing in vivo. Mol. Cell *12*, 525–532.

Diamant, G., Amir-Zilberstein, L., Yamaguchi, Y., Handa, H., and Dikstein, R. (2012). DSIF restricts NF-κB signaling by coordinating elongation with mRNA processing of negative feedback genes. Cell Rep. *2*, 722–731.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. *11*, 1475–1489.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dölken, L., Ruzsics, Z., Rädle, B., Friedel, C.C., Zimmer, R., Mages, J., Hoffmann, R., Dickinson, P., Forster, T., Ghazal, P., and Koszinowski, U.H. (2008). High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. RNA *14*, 1959–1972.

Drexler, H.L., Choquet, K., and Churchman, L.S. (2020). Splicing kinetics and coordination revealed by direct nascent RNA sequencing through nanopores. Mol. Cell *77*, 985–998.e8.

Effenberger, K.A., Anderson, D.D., Bray, W.M., Prichard, B.E., Ma, N., Adams, M.S., Ghosh, A.K., and Jurica, M.S. (2014). Coherence between cellular responses and in vitro splicing inhibition for the anti-tumor drug pladienolide B and its analogs. J. Biol. Chem. *289*, 1938–1947.

Effenberger, K.A., Urabe, V.K., Prichard, B.E., Ghosh, A.K., and Jurica, M.S. (2016). Interchangeable SF3B1 inhibitors interfere with pre-mRNA splicing at multiple stages. RNA *22*, 350–359.

Ehrensberger, A.H., Kelly, G.P., and Svejstrup, J.Q. (2013). Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. Cell *154*, 713–715.

Fededa, J.P., Petrillo, E., Gelfand, M.S., Neverov, A.D., Kadener, S., Nogués, G., Pelisch, F., Baralle, F.E., Muro, A.F., and Kornblihtt, A.R. (2005). A polar mechanism coordinates different regions of alternative splicing within a single gene. Mol. Cell *19*, 393–404.

Fiszbein, A., Krick, K.S., Begg, B.E., and Burge, C.B. (2019). Exon-mediated activation of transcription starts. Cell *179*, 1551–1565.e17.

Fong, N., and Bentley, D.L. (2001). Capping, splicing, and 3′ processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. Genes Dev. *15*, 1783–1795.

Fong, Y.W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. Nature *414*, 929–933.

Fong, N., Kim, H., Zhou, Y., Ji, X., Qiu, J., Saldi, T., Diener, K., Jones, K., Fu, X.D., and Bentley, D.L. (2014). Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. Genes Dev. *28*, 2663–2676.

François, B.L., Zhang, L., Mahajan, G.J., Stockmeier, C.A., Friedman, E., and Albert, P.R. (2018). A novel alternative splicing mechanism that enhances human 5-HT1A receptor RNA stability is altered in major depression. J. Neurosci. *38*, 8200–8210.

Frendewey, D., and Keller, W. (1985). Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. Cell *42*, 355–367.

Friend, K., Lovejoy, A.F., and Steitz, J.A. (2007). U2 snRNP binds intronless histone pre-mRNAs to facilitate U7-snRNP-dependent 3′ end formation. Mol. Cell *28*, 240–252.

Fu, X.D., and Maniatis, T. (1992). The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3′ splice site. Proc. Natl. Acad. Sci. U S A *89*, 1725–1729.

Fuchs, G., Hollander, D., Voichek, Y., Ast, G., and Oren, M. (2014). Cotranscriptional histone H2B monoubiquitylation is tightly coupled with RNA polymerase II elongation rate. Genome Res. *24*, 1572–1583.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T., and Peterlin, B.M. (2004). Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Mol. Cell. Biol. *24*, 787–795.

Furger, A., O'Sullivan, J.M., Binnie, A., Lee, B.A., and Proudfoot, N.J. (2002). Promoter proximal splice sites enhance transcription. Genes Dev. *16*, 2792–2799.

Gozani, O., Feld, R., and Reed, R. (1996). Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. Genes Dev. *10*, 233–243.

Gozani, O., Potashkin, J., and Reed, R. (1998). A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. Mol. Cell. Biol. *18*, 4752–4760.

Gressel, S., Schwalb, B., Decker, T.M., Qin, W., Leonhardt, H., Eick, D., and Cramer, P. (2017). CDK9-dependent RNA polymerase II pausing controls transcription initiation. eLife *6*, e29736.

Gressel, S., Schwalb, B., and Cramer, P. (2019). The pause-initiation limit restricts transcription activation in human cells. Nat. Commun. *10*, 3603.

Gu, B., Eick, D., and Bensaude, O. (2013). CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. Nucleic Acids Res. *41*, 1591–1603.

Harlen, K.M., Trotta, K.L., Smith, E.E., Mosaheb, M.M., Fuchs, S.M., and Churchman, L.S. (2016). Comprehensive RNA polymerase II interactomes reveal distinct and varied roles for each phospho-CTD residue. Cell Rep. *15*, 2147–2158.

Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nat. Rev. Mol. Cell Biol. *18*, 637–650.

Hirose, Y., Tacke, R., and Manley, J.L. (1999). Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. Genes Dev. *13*, 1234–1239.

Howe, K.J., Kane, C.M., and Ares, M., Jr. (2003). Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. RNA *9*, 993–1006.

Hsin, J.-P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev. *26*, 2119–2137.

Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C.Y., Xiao, R., Burge, C.B., and Fu, X.-D. (2013). SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. Cell *153*, 855–868.

Jonkers, I., Kwak, H., and Lis, J.T. (2014). Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. eLife *3*, e02407.

Kadener, S., Cramer, P., Nogués, G., Cazalla, D., de la Mata, M., Fededa, J.P., Werbajh, S.E., Srebrow, A., and Kornblihtt, A.R. (2001). Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing. EMBO J. *20*, 5759–5768.

Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., et al. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat. Chem. Biol. *3*, 576–583.

Kaida, D., Berg, M.G., Younis, I., Kasim, M., Singh, L.N., Wan, L., and Dreyfuss, G. (2010). U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. Nature *468*, 664–668.

Kim, J.B., Yamaguchi, Y., Wada, T., Handa, H., and Sharp, P.A. (1999). Tat-SF1 protein associates with RAP30 and human SPT5 proteins. Mol. Cell. Biol. *19*, 5960–5968.

Kornblihtt, A.R., Schor, I.E., Alló, M., Dujardin, G., Petrillo, E., and Muñoz, M.J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat. Rev. Mol. Cell Biol. *14*, 153–165.

Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., and Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat. Chem. Biol. *3*, 570–575.

Kotovic, K.M., Lockshon, D., Boric, L., and Neugebauer, K.M. (2003). Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. Mol. Cell. Biol. *23*, 5768–5779.

Krämer, A., Grüter, P., Gröning, K., and Kastner, B. (1999). Combined biochemical and electron microscopic analyses reveal the architecture of the mammalian U2 snRNP. J. Cell Biol. *145*, 1355–1368.

Kyburz, A., Friedlein, A., Langen, H., and Keller, W. (2006). Direct interactions between subunits of CPSF and the U2 snRNP contribute to the coupling of pre-mRNA 3′ end processing and splicing. Mol. Cell *23*, 195–205.

Lacadie, S.A., and Rosbash, M. (2005). Cotranscriptional spliceosome assembly dynamics and the role of U1 snRNA:5'ss base pairing in yeast. Mol. Cell *19*, 65–75.

Lamond, A.I., Konarska, M.M., and Sharp, P.A. (1987). A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. Genes Dev. *1*, 532–543.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. PLoS Comput. Biol. *9*, e1003118.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lin, S.-L., Chang, D., Wu, D.-Y., and Ying, S.-Y. (2003). A novel RNA splicing-mediated gene silencing mechanism potential for genome evolution. Biochem. Biophys. Res. Commun. *310*, 754–760.

Lin, S., Coutinho-Mansfield, G., Wang, D., Pandit, S., and Fu, X.-D. (2008). The splicing factor SC35 has an active role in transcriptional elongation. Nat. Struct. Mol. Biol. *15*, 819–826.

Lindstrom, D.L., Squazzo, S.L., Muster, N., Burckin, T.A., Wachter, K.C., Emigh, C.A., McCleery, J.A., Yates, J.R., 3rd, and Hartzog, G.A. (2003). Dual roles for Spt5 in pre-mRNA processing and transcription elongation revealed by identification of Spt5-associated proteins. Mol. Cell. Biol. *23*, 1368–1378.

Listerman, I., Sapra, A.K., and Neugebauer, K.M. (2006). Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. Nat. Struct. Mol. Biol. *13*, 815–822.

Liu, C.-R., Chang, C.-R., Chern, Y., Wang, T.-H., Hsieh, W.-C., Shen, W.-C., Chang, C.-Y., Chu, I.C., Deng, N., Cohen, S.N., and Cheng, T.H. (2012).

Spt4 is selectively required for transcription of extended trinucleotide repeats. Cell *148*, 690–701.

Loerch, S., Leach, J.R., Horner, S.W., Maji, D., Jenkins, J.L., Pulvino, M.J., and Kielkopf, C.L. (2019). The pre-mRNA splicing and transcription factor Tat-SF1 is a functional partner of the spliceosome SF3b1 subunit via a U2AF homology motif interface. J. Biol. Chem. *294*, 2892–2902.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Marshall, N.F., and Price, D.H. (1995). Purification of P-TEFb, a transcription factor required for the transition into productive elongation. J. Biol. Chem. *270*, 12335–12338.

Marshall, N.F., Peng, J., Xie, Z., and Price, D.H. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J. Biol. Chem. *271*, 27176–27183.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. *17*, 10–12.

Martins, S.B., Rino, J., Carvalho, T., Carvalho, C., Yoshida, M., Klose, J.M., de Almeida, S.F., and Carmo-Fonseca, M. (2011). Spliceosome assembly is coupled to RNA polymerase II dynamics at the 3′ end of human genes. Nat. Struct. Mol. Biol. *18*, 1115–1123.

Maslon, M.M., Braunschweig, U., Aitken, S., Mann, A.R., Kilanowski, F., Hunter, C.J., Blencowe, B.J., Kornblihtt, A.R., Adams, I.R., and Cáceres, J.F. (2019). A slow transcription rate causes embryonic lethality and perturbs kinetic coupling of neuronal genes. EMBO J. *38*, e101244.

Matter, N., and König, H. (2005). Targeted 'knockdown' of spliceosome function in mammalian cells. Nucleic Acids Res. *33*, e41.

Maudlin, I.E., and Beggs, J.D. (2019). Spt5 modulates cotranscriptional spliceosome assembly in *Saccharomyces cerevisiae*. RNA *25*, 1298–1310.

McCracken, S., Fong, N., Yankulov, K., Ballantyne, S., Pan, G., Greenblatt, J., Patterson, S.D., Wickens, M., and Bentley, D.L. (1997). The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. Nature *385*, 357–361.

Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. Mol. Cell *3*, 697–705.

Mortillaro, M.J., Blencowe, B.J., Wei, X., Nakayasu, H., Du, L., Warren, S.L., Sharp, P.A., and Berezney, R. (1996). A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. Proc. Natl. Acad. Sci. U S A *93*, 8253–8257.

Neugebauer, K.M. (2019). Nascent RNA and the coordination of splicing with transcription. Cold Spring Harb. Perspect. Biol. *11*, a032227.

Nojima, T., Gomes, T., Grosso, A.R.F., Kimura, H., Dye, M.J., Dhir, S., Carmo-Fonseca, M., and Proudfoot, N.J. (2015). Mammalian NET-seq reveals genome-wide nascent transcription coupled to RNA processing. Cell *161*, 526–540.

Nojima, T., Gomes, T., Carmo-Fonseca, M., and Proudfoot, N.J. (2016). Mammalian NET-seq analysis defines nascent RNA profiles and associated RNA processing genome-wide. Nat. Protoc. *11*, 413–428.

Nojima, T., Rebelo, K., Gomes, T., Grosso, A.R., Proudfoot, N.J., and Carmo-Fonseca, M. (2018). RNA polymerase II phosphorylated on CTD serine 5 interacts with the spliceosome during co-transcriptional splicing. Mol. Cell *72*, 369–379.e4.

Oh, J.-M., Di, C., Venters, C.C., Guo, J., Arai, C., So, B.R., Pinto, A.M., Zhang, Z., Wan, L., Younis, I., and Dreyfuss, G. (2017). U1 snRNP telescripting regulates a size-function-stratified human genome. Nat. Struct. Mol. Biol. *24*, 993–999.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419.

Roybal, G.A., and Jurica, M.S. (2010). Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. Nucleic Acids Res. *38*, 6664–6672.

Ruskin, B., and Green, M.R. (1985). Specific and stable intron-factor interactions are established early during in vitro pre-mRNA splicing. Cell *43*, 131–142.

Ruskin, B., Greene, J.M., and Green, M.R. (1985). Cryptic branch point activation allows accurate in vitro splicing of human β-globin intron mutants. Cell *41*, 833–844.

Ruskin, B., Zamore, P.D., and Green, M.R. (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell *52*, 207–219.

Saldi, T., Cortazar, M.A., Sheridan, R.M., and Bentley, D.L. (2016). Coupling of RNA polymerase II transcription elongation with pre-mRNA splicing. J. Mol. Biol. *428*, 2623–2635.

Saponaro, M., Kantidakis, T., Mitter, R., Kelly, G.P., Heron, M., Williams, H., Söding, J., Stewart, A., and Svejstrup, J.Q. (2014). RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. Cell *157*, 1037–1049.

Schlackow, M., Nojima, T., Gomes, T., Dhir, A., Carmo-Fonseca, M., and Proudfoot, N.J. (2017). Distinctive patterns of transcription and RNA processing for human lincRNAs. Mol. Cell *65*, 25–38.

Schwalb, B., Michel, M., Zacher, B., Frühauf, K., Demel, C., Tresch, A., Gagneur, J., and Cramer, P. (2016). TT-seq maps the human transient transcriptome. Science *352*, 1225–1228.

Shao, W., and Zeitlinger, J. (2017). Paused RNA polymerase II inhibits new transcriptional initiation. Nat. Genet. *49*, 1045–1051.

Spiluttini, B., Gu, B., Belagal, P., Smirnova, A.S., Nguyen, V.T., Hébert, C., Schmidt, U., Bertrand, E., Darzacq, X., and Bensaude, O. (2010). Splicing-independent recruitment of U1 snRNP to a transcription unit in living cells. J. Cell Sci. *123*, 2085–2093.

Talkish, J., Igel, H., Hunter, O., Horner, S.W., Jeffery, N.N., Leach, J.R., Jenkins, J.L., Kielkopf, C.L., and Ares, M. (2019). Cus2 enforces the first ATP-dependent step of splicing by binding to yeast SF3b1 through a UHM-ULM interaction. RNA *25*, 1020–1037.

Tellier, M., Maudlin, I., and Murphy, S. (2020). Transcription and splicing: a two-way street. Wiley Interdiscip. Rev. RNA *11*, e1593.

Teng, T., Tsai, J.H.C., Puyang, X., Seiler, M., Peng, S., Prajapati, S., Aird, D., Buonamici, S., Caleb, B., Chan, B., et al. (2017). Splicing modulators act at the branch point adenosine binding pocket defined by the PHF5A-SF3b complex. Nat. Commun. *8*, 15522.

Tresini, M., Warmerdam, D.O., Kolovos, P., Snijder, L., Vrouwe, M.G., Demmers, J.A.A., van IJcken, W.F., Grosveld, F.G., Medema, R.H., Hoeijmakers, J.H.J., et al. (2015). The core spliceosome as target and effector of non-canonical ATM signalling. Nature *523*, 53–58.

Valcárcel, J., Gaur, R.K., Singh, R., and Green, M.R. (1996). Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. Science *273*, 1706–1709.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods *13*, 508–514.

Veloso, A., Kirkconnell, K.S., Magnuson, B., Biewen, B., Paulsen, M.T., Wilson, T.E., and Ljungman, M. (2014). Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. Genome Res. *24*, 896–905.

Verheijen, M., Lienhard, M., Schrooders, Y., Clayton, O., Nudischer, R., Boerno, S., Timmermann, B., Selevsek, N., Schlapbach, R., Gmuender, H., et al. (2019). DMSO induces drastic changes in human cellular processes and epigenetic landscape in vitro. Sci. Rep. *9*, 4641.

Vincent, M., Lauriault, P., Dubois, M.-F., Lavoie, S., Bensaude, O., and Chabot, B. (1996). The nuclear matrix protein p255 is a highly phosphorylated form of RNA polymerase II largest subunit which associates with spliceosomes. Nucleic Acids Res. *24*, 4649–4652.

Vos, S.M., Farnung, L., Boehning, M., Wigge, C., Linden, A., Urlaub, H., and Cramer, P. (2018). Structure of activated transcription complex Pol II-DSIF-PAF-SPT6. Nature *560*, 607–612.

Vos, S.M., Farnung, L., Linden, A., Urlaub, H., and Cramer, P. (2020). Structure of complete Pol II-DSIF-PAF-SPT6 transcription complex reveals RTF1 allosteric activation. Nat. Struct. Mol. Biol. *27*, 668–677.

Wachutka, L., Caizzi, L., Gagneur, J., and Cramer, P. (2019). Global donor and acceptor splicing site kinetics in human cells. eLife *8*, e45056.

Wallace, E.W.J., and Beggs, J.D. (2017). Extremely fast and incredibly close: cotranscriptional splicing in budding yeast. RNA *23*, 601–610.

Wang, K., Yin, C., Du, X., Chen, S., Wang, J., Zhang, L., Wang, L., Yu, Y., Chi, B., Shi, M., et al. (2019). A U2-snRNP-independent role of SF3b in promoting mRNA export. Proc. Natl. Acad. Sci. U S A *116*, 7837–7846.

Wei, P., Garber, M.E., Fang, S.-M., Fischer, W.H., and Jones, K.A. (1998). 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 *92*, 451–462.

Xiao, Y., Yang, Y.H., Burckin, T.A., Shiue, L., Hartzog, G.A., and Segal, M.R. (2005). Analysis of a splice array experiment elucidates roles of chromatin elongation factor Spt4-5 in splicing. PLoS Comput. Biol. *1*, e39.

Yan, D., Perriman, R., Igel, H., Howe, K.J., Neville, M., and Ares, M., Jr. (1998). CUS2, a yeast homolog of human Tat-SF1, rescues function of misfolded U2 through an unusual RNA recognition motif. Mol. Cell. Biol. *18*, 5000–5009.

Yokoi, A., Kotake, Y., Takahashi, K., Kadowaki, T., Matsumoto, Y., Minoshima, Y., Sugi, N.H., Sagane, K., Hamaguchi, M., Iwata, M., and Mizui, Y. (2011). Biological validation that SF3b is a target of the antitumor macrolide pladienolide. FEBS J. *278*, 4870–4880.

Zamore, P.D., and Green, M.R. (1989). Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. Proc. Natl. Acad. Sci. U S A *86*, 9243–9247.

Zhang, S., Aibara, S., Vos, S.M., Agafonov, D.E., Lührmann, R., and Cramer, P. (2020a). Structure of a transcribing RNA polymerase II–U1 snRNP complex. bioRxiv. https://doi.org/10.1101/2020.10.19.344200.

Zhang, Z., Will, C.L., Bertram, K., Dybkov, O., Hartmuth, K., Agafonov, D.E., Hofele, R., Urlaub, H., Kastner, B., Lührmann, R., and Stark, H. (2020b). Molecular architecture of the human 17S U2 snRNP. Nature *583*, 310–313.

Zillmann, M., Zapp, M.L., and Berget, S.M. (1988). Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. Mol. Cell. Biol. *8*, 814–821.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal against RNA polymerase II	Diagenode	Cat#C15200004; RRID:AB_2728744
Rabbit monoclonal against Cyclin T1	Cell Signaling	Cat#D1B6G; RRID:AB_2799973
Chemicals, peptides, and recombinant proteins		
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	Cat#D2438
Pladienolide-B	Santa Cruz	Cat#sc-391691
Spliceostatin-A	Gift from Vladimir Pena	N/A
4-thiouridine	Carbosynth	Cat#NT06186
Empigen	Sigma-Aldrich	Cat#30326
Formaldehyde 16% concentrate stock methanol-free	Thermo Fisher Scientific	Cat#28908
Critical commercial assays		
Plasmo Test Mycoplasma Detection Kit	InvivoGen	Cat#rep-pt1
CellTiter 96® AQueous One Solution Cell Proliferation Assay	Promega	Cat#G3580
NEON system	Invitrogen	Cat#MPK10025
Ovation Universal RNA-seq System	NuGEN	Cat#0343-32
TruSeq Small RNA Library Kit	Illumina	Cat#RS-200-0048
NEBNext® Ultra II DNA Library Prep Kit	NEB	Cat#E7645S
Deposited data		
Raw sequencing data	This paper	GEO: GSE148433
Re-analyzed mNET-seq data	Schlackow et al., 2017	GEO: GSE81662
Original data	This paper	Mendeley Data: https://doi.org/10.17632/77k5xyhtjs.1
Experimental models: cell lines		
Human: K562 (female)	DSMZ	ACC-10; RRID:CVCL_0004
Human: HeLa S3 (female)	GBF (Helmholtz Center for Infection Research, Brunswick)	RRID:CVCL_0058
Saccharomyces cerevisiae strain: BY4741	Euroscarf	ACC-Y00000
Drosophila melanogaster: Schneider 2 (S2)	DSMZ	ACC-130; RRID:CVCL_Z232
Oligonucleotides		
Ctr AMO: CCTCTTACCTCAGTTAC AATTTATA	GeneTools	N/A
U2 AMO: TGATAAGAACAGATACTAC ACTTGA	GeneTools	N/A
Recombinant DNA		
pCMV-GLuc 2 control plasmid	NEB	Cat#N8081
pGL4.10[luc2] vector	Promega	Cat#E6651
pSV40-CLuc control plasmid	NEB	Cat#N0318S
pCMV-GLuc 2 intron inserted plasmid	This paper	N/A
Software and algorithms		
STAR (2.6.0)	Dobin et al., 2013	https://github.com/alexdobin/STAR; RRID:SCR_015899

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Samtools (1.3.1)	Li et al., 2009	https://sourceforge.net/projects/samtools/files/samtools/1.3.1/; RRID:SCR_002105
HTSeq (0.6.1.p1)	Anders et al., 2015	https://htseq.readthedocs.io/en/master/; RRID:SCR_005514
Salmon (0.13.1)	Patro et al., 2017	https://salmon.readthedocs.io/en/latest/salmon.html; RRID:SCR_017036
Cutadapt (1.18)	Martin, 2011	https://cutadapt.readthedocs.io/en/v1.18/installation.html; RRID:SCR_011841
Bowtie 2 (2.3.5)	Langmead and Salzberg., 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml; RRID:SCR_005476
DESeq2 (1.24.0)	Love et al., 2014	https://bioconductor.org/packages/release/bioc/html/DESeq2.html; RRID:SCR_015687
Scripts and data analysis	This paper	https://github.com/cramerlab/Efficient-RNA-polymerase-II-pause-release-requires-U2-snRNP-function_2021

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be sent to the lead contact, Patrick Cramer (patrick.cramer@mpibpc.mpg.de).

Materials availability

Unique reagents generated in this study are available from the lead contact without restriction.

Data and code availability

The accession number for the sequencing data and processed files reported in this paper is GEO: GSE148433. Original data were deposited to Mendeley Data: https://doi.org/10.17632/77k5xyhtjs.1.

The scripts generated during this study are available at https://github.com/cramerlab/Efficient-RNA-polymerase-II-pause-release-requires-U2-snRNP-function_2021.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Employed cell lines

Human K562 cells (female) were obtained from DSMZ (DSMZ no.: ACC-10, RRID:CVCL_0004) and cultured in antibiotic-free RPMI 1640 medium (Thermo Fisher Scientific, 31870–074) supplemented with 10% heat-inactivated fetal bovine serum (Thermo Fisher Scientific, 10500–064) and 2 mM GlutaMAX (Thermo Fisher Scientific, 35050087) at 37°C and 5% CO_2. K562 cells were authenticated at the DSMZ Identification Service according to standards for STR profiling (ASN-0002). Biological replicates were grown independently. Human HeLa S3 cells (female, used for nuclear extracts preparation) were obtained from GBF (Helmholtz Center for Infection Research, Brunswick, RRID:CVCL 0058) and cultured in antibiotic-free Ham's F-12K medium (Thermo Fisher Scientific, 21127077) supplemented with 10% heat-inactivated fetal bovine serum (Thermo Fisher Scientific, 10500–064) and 2 mM GlutaMAX (Thermo Fisher Scientific, 35050087) at 37°C and 5% CO_2. Cells were verified to be free of mycoplasma contamination using Plasmo Test Mycoplasma Detection Kit (InvivoGen, rep-pt1). *Drosophila melanogaster* S2 (Schneider-2) cells (used for ChIP-seq spike-ins) were obtained from DSMZ (DSMZ no.: ACC-130, RRID:CVCL_Z232) and cultured in Schneider's *Drosophila* medium (Biowest, L0207) supplemented with 10% heat-inactivated fetal bovine serum (Thermo Fisher Scientific, 10500–064) at 27°C without CO_2. Yeast *S. cerevisiae* cells genotype BY4741 (used for mNET-seq spike-ins) were obtained from Euroscarf (ACC-Y00000) and cultured in YP medium supplemented with 2% glucose to OD_{600} 0.5 at 30°C.

Cell culture treatments

Cells were plated the day before the experiment at a confluency of 300,000 cell / mL to obtain approximately 500,000-700,000 cells / mL on the next day. For pladienolide B (Pla-B) experiments, cells were treated with dimethyl sulfoxide (DMSO, 1:20,000 dilution, Sigma-Aldrich, D2438) as solvent control or with 1 μM Pla-B from a DMSO-resuspended 20 mM stock (Santa Cruz, sc-391691). For Spliceostatin-A (SSA) experiments, cells were treated with DMSO (1:17,000 dilution, Sigma-Aldrich, D2438) as solvent control or with SSA from a DMSO-resuspended 1 mM stock (gift from Vladimir Pena).

METHOD DETAILS

Total and 4sU-labeled RNA extraction and semiquantitative-PCR

Total RNA was extracted using QIAzol (QIAGEN, 79306) manufacturer's instruction. 4sU-labeled RNA was extracted and purified as described for TT-seq experiment (see below) except for the sonication of 4sU-labeled RNA which was omitted to avoid introns fragmentation. Isolated RNA was treated with DNase (QIAGEN, 79254) to avoid genomic DNA contamination. 400 ng of total or 140 ng of purified 4sU-labeled RNA were reverse transcribed (Thermo Fisher Scientific, EP0751) using random hexamer primers. PCR was performed with Phusion High-Fidelity DNA Polymerase (NEB, M0530L) using the following parameters:

DNAJB1: 30 s (sec) at 98°C, 28 cycles of 5 s at 98°C – 10 s at 68°C – 15 s at 72°C, and 5 minutes (min) at 72°C.
RIOK3: 30 s at 98°C, 28 cycles of 5 s at 98°C – 10 s at 59°C – 15 s at 72°C, and 5 min at 72°C.

Sequences of primers used in this assay are listed in Table S1. Amplified cDNA was loaded in a 2% agarose gel. Band intensity was analyzed with ImageJ (Fiji).

In vitro splicing assay

Two m7G(5′)ppp(5′)G-capped pre-mRNAs, *MINX* (Zillmann et al., 1988) and *PM5* (Bessonov et al., 2008) were body labeled with UTP-Cy5 by *in vitro* run-off transcription using T7 polymerase and purified with MEGAClear™ Kit (Life Technologies, AM1908). HeLa S3 nuclear extract was prepared as described in (Dignam et al., 1983). The standard *in vitro* splicing reactions (3 mM $MgCl_2$, 65 mM KCl, 20 mM HEPES–KOH pH 7.9, 2 mM ATP, and 20 mM creatine phosphate) with 40% dialyzed HeLa S3 nuclear extract was incubated with DMSO or 1 μM Pla-B for 20 min at 30°C before 10-30 nM of Cy5-labeled pre-mRNA was added. RNA was then recovered with phenol-chloroform-isoamylalcohol extraction and analyzed on denatured 4%–12% gradient PAGE (Bis-Tris NuPAGE, Invitrogen) and visualized with Typhoon™ FLA 9500 (GE Healthcare). For each pre-mRNA tested, splicing reactions were performed at the same time and gels were run in parallel.

MS2 affinity purification of spliceosomal complexes

The spliceosomal complexes assembled on unlabeled *MINX* or *PM5* pre-mRNAs in presence of DMSO or 1 μM Pla-B were purified by MS2-MBP pulldown assay. Briefly, 10 nM of m^7G(5′)ppp(5′)G-capped pre-mRNA was mixed with 20-fold molar excess of MBP-MS2 fusion protein. A 10 mL standard splicing reaction (3 mM $MgCl_2$, 65 mM KCl, 20 mM HEPES–KOH pH 7.9, 2 mM ATP, and 20 mM creatine phosphate) with 40% dialyzed HeLa S3 nuclear extract was incubated with DMSO or 1 μM Pla-B for 20 min at 30°C before MBP-MS2 bound pre-mRNA was added. The assembly of the spliceosomes was started with ATP and performed for 60 min at 30°C. The reactions were centrifuged at 4000 rpm for 5 min and supernatant was incubated for 1 hour at 4°C with 250 μL amylose beads. After washing with G150 buffer (20 mM HEPES, pH 7.9, 150 mM NaCl, 1.5 mM $MgCl_2$), the complexes were eluted with 3 mM maltose in G150 buffer. The RNA, recovered from the eluted complexes, was analyzed with denatured 4%–12% gradient PAGE (Bis-Tris NuPAGE, Invitrogen), stained with SYBR gold and visualized with Typhoon ™ FLA 9500 (GE Healthcare).

TT-seq and RNA-seq

TT-seq was performed as described (Schwalb et al., 2016; Wachutka et al., 2019) with minor modifications. Specifically, 5×10^7 cells from two biological replicates were either treated for with 1 μM of Pla-B (Santa Cruz, sc-391691) or 30 ng / mL SSA (gift from Vladimir Pena) or DMSO (Sigma-Aldrich, D2438) control at times indicated. Cells were exposed the last 10 min of the treatment time to 500 μM of 4-thiouridine (4sU, Carbosynth, NT06186) at 37°C and 5% CO_2. 300 ng of RNA spike-ins mix were added to each sample after cell lysis in 10 mL of QIAzol (QIAGEN, 79306). Spike-ins sequences and production are described in (Wachutka et al., 2019). 150 μg of RNAs were sonicated to obtain fragments of < 6 kb using AFA micro tubes in a S220 Focused-ultrasonicator (Covaris). The quality of RNAs and the size of fragmented RNAs were checked using a Fragment Analyzer. 1 μg of each of the sonicated RNAs was stored at −80°C as total RNA (RNA-seq). 4sU-labeled RNAs were purified from 300 μg of each of the fragmented RNAs. Biotinylation and purification of 4sU-labeled RNAs was performed as described (Dölken et al., 2008; Wachutka et al., 2019). 100 ng of input RNA was used for strand-specific library preparation according to the Ovation Universal RNA-seq System (NuGEN, 0343-32). Libraries were prepared using random hexamer priming only. The size-selected libraries were analyzed on a Fragment Analyzer before sequencing on the Illumina NextSeq500 (2 × 75 and 2 × 150 base paired-end for shallow and deep sequencing, respectively).

Proliferation assay

Proliferation assay was performed using CellTiter 96® AQueous One Solution Cell Proliferation Assay (Promega, G3580) in four biological replicates according to the manufacturer's instructions. Briefly, 5,000 cells were plated into each well of a 96-well plate the day before the experiment. For the background measurement, the medium was added to the empty wells. Next morning, the treatment with DMSO versus 100 nM / 1 μM Pla-B or DMSO versus 30 ng / mL SSA was started in a time course of 0, 1, 4, 8, 12 and 24 h. Equal amounts of DMSO or Pla-B / SSA were added to the control wells with medium only. Afterward, CellTiter 96® AQueous One Solution Reagent was added to the wells of the corresponding time point and the plate was incubated for 1.5 h in a humidified, 5% CO_2 incubator. The absorbance was measured at 490 nm using Infinite M1000 Pro Plate Reader (Tecan). For the calculation, background

values (the wells with media only) were subtracted from the actual values (the wells with cells). The average signal and standard deviation were calculated for each condition of the corresponding time point. A two-way ANOVA test was performed using 4 replicates.

Morpholino inhibition

3.2×10^6 of K562 cells were transfected by electroporation with 7.5 nmol / 100 μL (75 μM) of antisense morpholino oligo against a control region (Ctr AMO, CCTCTTACCTCAGTTACAATTTATA) or U2 snRNA (U2 AMO, TGATAAGAACAGATACTACACTTGA) in a transfection volume of 100 μL using the NEON system (Invitrogen, MPK10025). The following transfection parameters were used: pulse: 1,350, pulse width: 10, pulse number: 4. Transfected cells were transferred to 5 mL of fresh culture medium, and harvested at the indicated times. For semiquantitative PCR analysis, RNA was isolated and treated as previously described (see "Total and 4sU-labeled RNA extraction and semiquantitative-PCR"). For TT-seq experiment, cells were exposed the last 10 min of the 1 h transfection time to 500 μM of 4-thiouridine. RNA was isolated and treated as previously described (see "TT-seq and RNA-seq").

RNA extension assay

RNA extension assays were performed with a set of A-less *JUNB* scaffold containing template DNA (*JUNB*-A-less-T-DNA), 5′-biotin-labeled non-template DNA (JUNB-A-less-NT-DNA) and a short 5′-FAM labeled RNA primer (*JUNB* U-RNA) (see Table S2 for oligonucleotides sequences). Final transcription reaction contained 75 nM Pol II, 50 nM *JUNB* scaffold, 100 mM NaCl, 20 mM HEPES 7.5, 3 mM $MgCl_2$, 1 mM DTT, 4% glycerol and 0.01 mM NTP (a mixture of CTP, GTP and UTP), in the absence or presence of 2% DMSO or 1 μM Pla-B. Time course experiments with time points 0, 10, 30, 60, 120, 300 and 600 s were performed in a similar procedure as previously described (Vos et al., 2020), with an additional incubation step at 30°C for 20 min with DMSO or Pla-B, prior the starting of transcription by addition of NTPs. RNA extension products were analyzed on 20% acrylamide-urea gels. The gels were run at 300 Volts for about 90 min in 0.5 × TBE buffer. RNA signal was detected by scanning fluorescence of the FAM label on primer RNA with Typhoon™ FLA 9500 (GE Healthcare). The experiments were repeated 4 times on different days. The RNA extension products were quantified with ImageJ (Fiji). To quantify the extension products, the overall gel background was first subtracted. The integrated intensity of the extended products (Int E) was measured. To reduce the error from unequal loading, a constant RNA band in the synthesized RNA primer, the 6th band counting from the bottom of the gel, was chosen as a normalization standard RNA intensity. The normalized intensity of extended RNA product ($Int_{t=x}$) was calculated by dividing the integrated intensity of extended RNA ($Int\ E_{t=x}$) with the intensity of the normalization standard RNA ($Int\ N_{t=x}$) and multiplying 100 at a certain time point (t = x) ($Int_{t=x} = Int\ E_{t=x} / Int\ N_{t=x} \times 100$). The normalized intensity at time point 0 ($Int_{t=0}$) was further subtracted from the normalized intensity ($Int\ final\_{t=x} = Int\ E_{t=x} / Int\ N_{t=x} \times 100 - Int\ E_{t=0} / Int\ N_{t=0} \times 100$). See also Mendeley repository data for original images and quantification. A two-way ANOVA test was performed using 4 replicates. For each replicate, RNA extension assays were performed at the same time and gels were run in parallel.

Plasmid construction for reporter gene splicing assay

GLuc 2 coding sequence (NEB, N8081) was cloned into the pGL4.10 vector backbone (Promega, E6651) with pCMV promoter (NEB, N8081). An intron sequence of 1,982 bp originating from the human genome position chr3: 89596735-89598673 (genome version GRCh38.p12) was flanked by "GTAAGAGT" sequence at its 5′ end (necessary for 5′ splice site (5′ SS) recognition) as well as "TACTAACTCTTTTTTTTTTTTAAGCTTGCAGGC" sequence at the 3′ end containing branch site (BS) sequence, polypyrimidine track sequence, HindIII recognition sequence as well as 3′ splice site (3′ SS) sequence (designed based on (Lin et al., 2003)). Human genome sequence constituting the intron was chosen based on its lack of regulatory elements reported by the ENCODE Project. All T residues located within a 45 bp window upstream of the BS sequence were changed to A residues. BS sequence mutation (AGGC > GATC) and 3′ SS sequence mutation (TACTAAC > CGGCUTG) were generated based on (Wachutka et al., 2019).

Reporter gene splicing assay

3.2×10^6 of K562 cells were transfected by electroporation with 5 μg of the plasmid encoding either wild-type or mutated GLuc 2 transcript as well as 1 μg of the pSV40-CLuc control plasmid (NEB, N0318S) (transfection control) in a transfection volume of 100 μL using the NEON system (Invitrogen, MPK10025). The following transfection parameters were used: pulse: 1,350, pulse width: 10, pulse number: 4. Transfected cells were transferred to 5 mL of fresh culture medium. The experiment was performed in 3 biological replicates. Cells were harvested after 24 h and RNA was isolated using QIAzol reagent (QIAGEN, 79306). RNA was treated with TURBO DNase I (Thermo Fisher Scientific, AM1907). 1 μg of DNA-free RNA was reverse transcribed in presence of random hexamers using Maxima reverse transcriptase (Thermo Fisher Scientific, EPO752). cDNA was subjected to real-time PCR analysis using SYBR™ Select Mastermix (Applied biosystems, 4472908) and qTOWER v2.2 thermal cycler (Analytik Jena AG). The analysis was performed in 3 technical replicates for each of the 3 biological replicates. Raw Ct values were extracted and the abundance of the PCR product was calculated based on a standard curve generated for each primer pair. The estimated PCR product amount was normalized for total RNA level (*GAPDH* PCR product) as well for the transfection efficiency (*CLuc* PCR product). Sequences of primers used in this assay are listed in Table S3.

mNET-seq

Experiments were performed as previously described (Gressel et al., 2019; Nojima et al., 2015, 2016; Schlackow et al., 2017) with minor modifications. All buffers for the cellular fractionation and immunoprecipitation (IP) were supplied with protease inhibitor

cocktail (Sigma-Aldrich, P8340) and phosphatase inhibitors (Sigma-Aldrich, 4906837001). In detail, two biological replicates of 1 × 10^8 K562 cells were treated for 1 h with 1 μM of Pla-B (Santa Cruz, sc-391691) or DMSO (Sigma-Aldrich, D2438) solvent. Cellular fractionation was performed according to the previously published protocol (Conrad and Ørom, 2017) using 1 × 10^7 cells per one reaction. Isolated chromatin was subjected to micrococcal nuclease (mNase, NEB, M0247S) digestion at 37°C and 1,400 rpm for 2 min. Afterward, soluble chromatin fractions were pooled together for each sample and diluted 8-fold with IP buffer (50 mM Tris HCl pH 7.4, 150 mM NaCl, 0.05% (vol / vol) NP-40, 0.3% (vol / vol) empigen BB (Sigma-Aldrich, 30326)). For each sample, 30 μg of RNA polymerase II (Pol II) antibody (Diagenode, C15200004, RRID:AB_2728744) were coupled to Dynabeads M-280 Sheep Anti-Mouse IgG (Invitrogen, 11201D) prior to IP step. Diluted chromatin was mixed with Pol II antibody-beads complexes and subjected to IP performed on a rotating wheel at 4°C for 1 h. Beads were washed seven times with IP buffer (50 mM Tris HCl pH 7.4, 150 mM NaCl, 0.05% (vol / vol) NP-40, 0.3% (vol / vol) empigen BB (Sigma-Aldrich, 30326)) and one time with PNKT buffer (1 × T4 PNK buffer (NEB, M0236L), 0.1% vol / vol Tween-20). For the RNA phosphorylation, beads were resuspended in PNK reaction mixture containing 1 × T4 PNK buffer (NEB, M0236L), 0.1% vol / vol Tween-20, 1 mM ATP (Cell Signaling Technology, 9804S), T4 Polynucleotide Kinase (phosphatase minus) (NEB, M0236L) and incubated at 37°C and 800 rpm for 10 min. Beads were washed once with IP buffer and resuspended in TRIzol reagent (Invitrogen, 15596026). At this point, 5 ng of RNA spike-ins were added to each sample. RNA was extracted from the beads and precipitated with GlycoBlue coprecipitant in 100% ethanol at −20°C. Next day, RNA was size-selected (25-110 nt) using a denaturing 6% (wt/vol) polyacrylamide gel containing 7 M urea. RNA was extracted from the gel by incubation with the elution buffer (1 M NaOAc pH 5.5, 1 mM of EDTA pH 8.0) and precipitated with GlycoBlue coprecipitant (Invitrogen, AM9515) in 100% ethanol at −20°C. RNA size distribution and concentration were estimated using Fragment Analyzer. RNA samples were used for the library preparation in equal amounts. Libraries were prepared using TruSeq Small RNA Library Kit (RS-200-0048) according to the manual and as described (Nojima et al., 2015). Purity and size distribution of the libraries were estimated using Fragment Analyzer. Libraries were sequenced on Illumina NEXTseq 550 (2 × 42 base paired-end).

RNA spike-ins for mNET-seq

50 mL of *S. cerevisiae* cells (wild-type strain BY4741) were grown overnight at 30°C. Cells were spin down, and pellet was washed 1 × with 25 mL milli-Q water. Cells from 25 mL volume pellet were lysed in 1.5 mL of TSNTE buffer (2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-HCl pH 8.0, 1 mM EDTA pH 8.0). One volume of acidic phenol (CarlRoth, A980.2) and 1 volume glass beads acid washed (Sigma-Aldrich, G8772) were added to 1 volume of the lysed cells and TSNTE buffer mixture. Cell lysate was further homogenized in FastPrep (8 × 40 s with 1 min on ice between each step). Cell lysate was centrifuged at high speed for 5 min. One volume of chloroform was added to 1 volume of supernatant. 0.3 M sodium acetate was added and RNA was precipitated overnight in 100% ethanol. RNA pellet was washed with 75% ethanol and resuspended in 900 μL of RNase-free water. Total RNA quality was checked in Fragment Analyzer. 200 μg of total RNA were subjected to DNase digestion (TURBO DNA-free DNase treatment, Thermo Fisher Scientific, AM1907) following manufacturer's instructions. RNA was purified with chloroform and precipitated overnight in 100% ethanol at −20°C. RNA pellet was washed with 75% ethanol, and pellet was resuspended in 50 μL of RNase-free water. Quality of total RNA DNase-treated was checked in Fragment Analyzer. 75 μg of total RNA DNase treated were subjected to mRNA purification using Dynabeads mRNA Purification kit (Invitrogen, cat n. 61006) following manufacturer's instructions. Two rounds of mRNA purification were performed to eliminate possible rRNA contaminants. Approximately 0.4% of original amount of total RNA was recovered. 250 ng of purified mRNA was fragmented for 5 min at 94°C using NEBNext® Magnesium RNA Fragmentation Module Protocol (NEB, E6150). 0.3M sodium acetate and 1 μL of Glycoblue was added to the fragmented mRNA. mRNA was precipitated overnight in 100% ethanol. RNA was washed with 75% ethanol, and pellet was resuspended in 10 μL of RNase-free water. End-repair was performed following NEB manufacturer's instruction. In details, reaction was incubated for 30 min at 37°C (9 μL of fragmented mRNA, 5.5 μL of RNase-free water, 0.5 μL of RNase OUT, 2 μL of T4 PNK buffer 10X (NEB, B0201), 1 μL of T4 PNK enzyme (NEB, M0201). 2 μL of 10 mM ATP were added to the mixture and the reaction was further incubated for other 30 min at 37°C. 0.3 M sodium acetate and 1 μL of Glycoblue was added to the reaction and RNA was precipitated overnight in 100% ethanol. RNA was washed in 75% ethanol and pellet resupended in 20 μL of RNase-free water. RNA spike-ins quality and size distribution were checked in Fragment Analyzer (miRNA kit) and in 6% urea gel, as described for mNET-seq protocol. Size distribution ranged between 30 nt and 150 nt, similar to mNET-seq samples.

ChIP-seq

6 × 10^7 cells from two biological replicates were treated for 1 h with 1 μM of Pla-B (Santa Cruz, sc-391691) or DMSO (1:20,000 dilution, Sigma-Aldrich, D2438). Formaldehyde (16% concentrate stock methanol-free, Thermo Fisher Scientific 28908) was directly added to the media to a final concentration of 1% and incubated for 8 min. 125 mM Glycine (final concentration) was added to quench the reaction for 5 min. Cells were spin down and pellet was washed twice with PBS at 4°C. Protease (Merk Millipore, 5892970001) and phosphatase inhibitors (Merk Millipore, 4906845001) were added to all the buffers. A pellet from 3 × 10^7 cells was lysed for 10 min on ice with Farnham Lysis buffer (5 mM Pipes pH 8, 85 mM KCl, 0.5% NP-40). Pellet was centrifuged for 5 min at 1,700 g at 4°C. Pellet was resuspended with 1 mL of sonication buffer (10 mM Tris-HCl 7.5 pH, 1 mM EDTA, 0.4% SDS) and incubate on ice for 10 min and transfer to AFA milliTube. Sonication was performed with a S220 Focused-ultrasonicator (Covaris) with the following parameters: duty cycle 5%, peak incident power 140 W, cycle per burst 200, processing time 1,080 s, degassing mode continuous, water run level 8. Sheared chromatin was centrifuged at 10,000 g for 15 min at 4°C. 15 μL of samples were de-crosslinked overnight at

65°C and size distribution was checked in 1% agarose gel. 15 μg of RNA polymerase II (Pol II) antibody (Diagenode, C15200004, RRID:AB_2728744), 20 μg of Cyclin T1 (CycT1) antibody (Cell Signaling, D1B6G, RRID:AB_2799973) or 1 μg of histone H2Av (Active Motif, 39715) were coupled to Dynabeads Protein G (Thermo Fisher Scientific, 10009D) for 2 h at room temperature for each sample. 100 μg of chromatin was used for each IP. 200 ng of *Drosophila S2* sheared crosslinked-chromatin (Covaris S200 parameters: duty cycle 5%, peak incident power 140 W, cycle per burst 200, processing time 1,800 s, degassing mode continuous, water run level 8) were added to 100 μg of chromatin as spike-ins control. Chromatin was diluted with IP buffer (50 mM HEPES pH 7.9, 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Sodium-deoxycholate) to obtain a 0.05% final concentration of SDS. 1% of diluted chromatin was kept as input at 4°C. Diluted chromatin was mixed with antibody-beads complexes and subjected to IP performed on a rotating wheel at 4°C overnight. Beads were washed 5 times with IP wash buffer (100 mM Tris HCl pH 7.5, 500 mM LiCl, 1% NP-40, 1% Sodium-deoxycholate) and one time with TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA). Immuno-bound chromatin was eluted at 70°C for 10 min with elution buffer (0.1M $NaHCO_3$, 1%SDS) and de-crosslinked overnight at 65°C. After RNase A treatment at 37°C for 1.5 h and proteinase K treatment at 45°C for 2 h, DNA was extracted with one volume phenol:chloroform:isoamyl alcohol 25:24:1 (Sigma-Aldrich, P2069) and precipitated for 30 min at −80°C with 200 mM NaCl and 100% ethanol. Pellet was washed with 70% ethanol and resuspended in TE buffer. DNA quality and size distribution were checked on Fragment Analyzer. 10 ng of DNA was used for library preparation according to NEBNext® Ultra II DNA Library Prep Kit (NEB, E7645S). Purity and size distribution of the libraries were estimated using Fragment Analyzer. Size-selected libraries were sequenced on Illumina NEXTseq 550 (2 × 75 and 2 × 42 base paired-end for Pol II and CycT1, respectively).

QUANTIFICATION AND STATISTICAL ANALYSIS

TT-seq and RNA-seq data preprocessing and normalization

Paired-end 75 and 150 bp reads with additional 6 bp of barcodes were obtained for each group of samples. Reads were aligned to the hg38 (GRCh38) genome assembly (Human Genome Reference Consortium) using STAR (2.6.0, RRID:SCR_015899) (Dobin et al., 2013), with the following specifications: outFilterMismatchNmax 2, outFilterMultimapScoreRange 0 and alignIntronMax 500000. Bam files were filtered with Samtools (1.3.1, RRID:SCR_002105) (Li et al., 2009) to remove alignments with MAPQ smaller than 7 (-q 7) and only proper pairs (-f2) were selected. Read counts for different features were calculated with HTSeq (0.6.1.p1, RRID:SCR_005514) (Anders et al., 2015). Further data processing was carried out using the R / Bioconductor environment. Expressed major isoforms were defined as possessing more than 50 read counts per kilobase (RPK) in two summarized replicates of TT-seq solvent control (DMSO). Prior to quantification, data was normalized by using added RNA spike-ins as described previously (Schwalb et al., 2016):

Antisense-bias ratios were calculated for each sample *j* according to

$$c_i = median_i\left(\frac{k_{ij}^{antisense}}{k_{ij}^{sense}}\right)$$

with read counts k_{ij} for all available spike-ins *i* in sample *j*.

Sequencing depth were calculated for each sample *j* according to

$$\sigma_j = median_i\left(\frac{k_{ij}}{l_i}\right)$$

with read counts k_{ij} for all available spike-ins *i* in sample *j* for the RNA-seq samples and for the labeled spike-ins *i* in sample j for the TT-seq samples.

The cross contamination ϵ_j rate was calculated for each sample *j* according to

$$\epsilon_j = \frac{median_i\left(\frac{k_{ij}}{l_i}\right)}{\sigma_j}$$

with read counts k_{ij} for all the labeled spike-ins *i* in sample *j* for the TT-seq samples and was set to 1 for the RNA-seq samples.

Major isoform annotation

Salmon (0.13.1, RRID:SCR_017036) (Patro et al., 2017) was used in order to select the major isoforms present in our dataset. RNA-seq samples for 1 h DMSO or 1 μM Pla-B treatments were mapped against curated RefSeq annotated isoforms (UCSC RefSeq GRCh38, downloaded in April 2019). For each gene, the major isoform was determined as the one with maximum mean Transcripts Per Million (TPM) value across all RNA-seq samples. Major isoforms were excluded from further analysis if they represented less than 70% of the gene isoforms based on the calculated mean TPM value. Additionally, major isoforms associated with overlapping genes as well as isoforms located on chromosomes X, Y and M were discarded from further analysis. The final major isoform annotation includes 6,694 isoforms containing 65,976 exons and 59,282 introns.

A total of 5,535 major transcript isoforms of protein-coding genes with RPK $\geq$ 50 of TT-seq solvent control (DMSO) were included in the analysis.

Intronless genes annotation

Intronless genes were defined as RefSeq annotated genes comprising one single isoform with one single exon. To avoid effects from neighboring intron-containing genes, only intronless genes at least 1 kb distant from the neighboring intron-containing annotated transcripts (strand independent) were included in the analysis. Moreover, because a long 3′ UTR has been recently reported to be alternatively spliced in an intronless gene (François et al., 2018), only intronless genes with UTRs $\leq$ 100 bp were included. A total of 51 expressed protein-coding intronless genes were included in the analysis.

Splicing ratio

An exon-based splice junction analysis was performed. Exons from major isoforms containing a first exon > 100 bp were included, and major isoforms with a single annotated exon were excluded. A window of ± 4 bp around the splice junction (2 bp in the exon and 2 bp in the intron) was defined to investigate spliced and unspliced reads by means of the *findOverlaps* function from the GenomicRanges R package (Lawrence et al., 2013). Because our preliminary analysis showed that the transcription levels were decreased upon Pla-B treatment, we considered only exon-based splice junctions that were covered by at least 30 total reads (spliced and unspliced) in Pla-B, SSA or U2 AMO treated samples. For 1 h 1 μM Pla-B treatment 15,551 exon-based 5′ SS and 14,840 exon-based 3′ SS were included, corresponding to 2,309 first exons, 13,242 intermediate exons and 1,598 last exons. For 4 h 1 μM Pla-B treatment, 10,041 exon-based 5′ SS and 9,550 exon-based 3′ SS were included, corresponding to 1,658 first exons, 8,383 intermediate exons and 1,167 last exons. For 1 h 100 nM Pla-B treatment, 9,759 exon-based 5′ SS and 9,845 exon-based 3′ SS were included, corresponding to 1,226 first exons, 8,533 intermediate exons and 1,312 last exons. For 1 h 30 ng / mL SSA treatment, 16,459 exon-based 5′ SS and 16,178 exon-based 3′ SS were included, corresponding to 1,989 first exons, 14,4470 intermediate exons and 1,708 last exons. For 1 h 75 μM U2 AMO treatment, 8,928 exon-based 5′ SS and 8,916 exon-based 3′ SS were included, corresponding to 1,245 first exons, 7,683 intermediate exons and 1,233 last exons. Unspliced reads were defined as the ones overlapping at least 3 bp of the defined window and total reads (spliced and unspliced) the ones spanning at least 2 bp of the defined window. Afterward, the number of spliced reads was calculated by the difference between total and unspliced reads. The splicing ratio was calculated by dividing the number of spliced reads by the total amount of spliced and unspliced reads.

Identification of splicing-affected and unaffected transcripts

A differential gene expression analysis on exon-based splice junction was conducted using DESeq2 R package (Love et al., 2014) with design ~assay + condition + assay:condition, where assay corresponds to the number of spliced and total amount of spliced and unspliced reads, and condition corresponds to DMSO and Pla-B treatment. The analysis was focused on the 5′ SS of first exons and 3′ SS of second exons in order to flank the first intron. To select splicing-affected exons DESeq2 results (using altHypothesis = "less") were filtered with padj < 0.05 and log2FC < −1. For 1 h 1 μM Pla-B treatment, 220 first exons affected on 5′ SS and 186 s exons affected on 3′ SS were selected. In total, 329 protein-coding major isoform transcripts, comprising 4,326 exons, contained a first exon affected on 5′ SS and / or a second exon affected on the 3′ SS. For 4 h 1 μM Pla-B treatment, 536 first exons affected on 5′ SS and 428 s exons affected on 3′ SS were selected. In total, 784 protein-coding major isoform transcripts, comprising 9,771 exons, contained an affected first intron. For 1 h SSA 30 ng / mL treatment, 398 first exons affected on 5′ SS and 415 s exons affected on 3′ SS were selected. In total, 606 protein-coding major isoform transcripts, comprising 7,078 exons, contained an affected first intron. For 1 h 75 μM U2 AMO treatment, 180 first exons affected on 5′ SS and 276 s exons affected on 3′ SS were selected. In total, 381 protein-coding major isoform transcripts, comprising 5,058 exons, contained an affected first intron. To select splicing-unaffected exons DESeq2 results (using altHypothesis = "lessAbs" and lfcThreshold = 1.5) were filtered with padj < 0.05. For 1 h 1 μM Pla-B treatment 316 first exons affected on 5′ SS and 287 s exons affected on 3′ SS were selected. In total, 355 protein-coding major isoform transcripts, comprising 2,756 exons, contained a first exon unaffected on 5′ SS and / or a second exon unaffected on the 3′ SS. For 4 h 1 μM Pla-B treatment, 187 first exons affected on 5′ SS and 157 s exons affected on 3′ SS were selected. In total, 194 protein-coding major isoform transcripts, comprising 1,082 exons, contained a first exon unaffected on 5′ SS and / or a second exon unaffected on the 3′ SS. For 1 h 30 ng / mL SSA treatment, 402 first exons affected on 5′ SS and 372 s exons affected on 3′ SS were selected. In total, 449 protein-coding major isoform transcripts, comprising 2,842 exons, contained a first exon unaffected on 5′ SS and / or a second exon unaffected on the 3′ SS. For 1 h 75 μM U2 AMO treatment, 147 first exons affected on 5′ SS and 145 s exons affected on 3′ SS were selected. In total, 177 protein-coding major isoform transcripts, comprising 1,164 exons, contained a first exon unaffected on 5′ SS and / or a second exon unaffected on the 3′ SS.

RNA amount per cell

The RNA amount per cell was calculated using the RNA spike-ins as described previously (Gressel et al., 2019; Gressel et al., 2017). The number of spike-in molecules per cell $N[\text{cell}^{-1}]$ was calculated as

$$N = \frac{m}{M_n} N_A$$

with 50 × 10^{-9} g of spike-ins m, 5 × 10^{7} cells n the Avogadro number N_A and the molar-mass of the spike-ins M calculated as

$$M = A_n \times 329.2 + 0.9 \times U_n \times 306.2 + C_n \times 305.2 + G_n \times 345.2 + 0.1 \times U_n \times 322.26 + 159$$

Where A_n,U_n, C_n and G_n are the number of each respective nucleotide within each spike-in polynucleotide. 159 corresponds to the molecular weight of a 5′ triphosphate. The conversion factor to RNA amount per cell k [cell^{-1}] was then calculated as

$$k = mean\left(median_i\left(\frac{tb_i}{L_i \cdot N}\right)\right)$$

For all labeled spike-ins i with length L_i.

Detection of Pol II pause sites

Pause sites were defined as described previously (Gressel et al., 2019; Gressel et al., 2017) for intron-containing major isoforms exceeding 10 kbp in length. The pause site was calculated separately for DMSO and Pla-B treated samples using mNET-seq signal within a window from the TSS to the end of the first exon. For intronless genes, the pause site was calculated within a 100 bp window starting at the TSS. Pause sites between conditions exceeding 10 bp difference were excluded. The pause site m was determined via the maximization of the function

$$\rho_i = max_m P_{im}$$

where ρ_i needed to exceed 5 times the median of the signal strength P_{im} for all non-negative antisense bias corrected mNET-seq coverage values (Nojima et al., 2015).

Productive initiation frequency

The initiation frequency was calculated as described previously (Gressel et al., 2019; Gressel et al., 2017) for intron-containing major isoforms exceeding 10 kbp in length. The productive initiation frequency I_i was calculated for each major isoform i for both DMSO and Pla-B treated samples using antisense bias corrected TT-seq coverage on major exons, excluding the first exon, as

$$I_i = \frac{1}{k} \cdot \frac{tb_i^{Control}}{t \cdot L_i}$$

with labeling time t = 10 and length L. For intronless genes, the productive initiation frequency was calculated within a window from the defined pause site on DMSO treated samples + 50 bp until the end of the transcript. Genes with a calculation window inferior to 100 bp were excluded from the analysis.

mNET-seq preprocessing and normalization

Paired-end 42 bp reads with additional 6 bp of barcodes were obtained for each of the samples. Reads were trimmed for adaptor content with Cutadapt (1.18, RRID:SCR_011841) (Martin, 2011) with -O 12 -m 25 -a TGGAATTCTCGG -A GATCGTCGGACT. mNET-seq data was normalized using *S. cerevisiae* RNA spike-ins. To this end, a combined genome was generated using the Ensemble genome assembly for both human hg38 (GRCh38) and *S. cerevisiae* (R64-1-1), against which the reads were mapped using STAR (2.6.0, RRID:SCR_015899) (Dobin et al., 2013). Around 80% and 20% of reads mapped to the human and *S. cerevisiae* genomes, respectively. Bam files were filtered with Samtools (1.3.1, RRID:SCR_002105) (Li et al., 2009) to remove alignments with MAPQ smaller than 7 (-q 7) and only proper pairs (-f2) were selected. Read counts for different features were calculated with HTSeq (0.6.1.p1, RRID:SCR_005514) (Anders et al., 2015). Further data processing was carried out using the R / Bioconductor environment. Antisense bias ratio was determined using positions in regions without antisense annotation with a coverage of at least 100 according to the defined major isoforms. Data was normalized using added *S. cerevisiae* RNA spike-ins. mNET-seq coverage was normalized with a median of ratios method (Love et al., 2014) using the antisense corrected counts for *S. cerevisiae* transcripts with an RPK of 100 or higher in two summarized replicates of mNET-seq solvent control (DMSO). Datasets for mNET-seq against CTD ser5P in HeLa cells upon 4 h DMSO or Pla-B treatments were obtained from (Schlackow et al., 2017), GEO:GSE81662. Empigen treatment during IP was not performed in this publicly available dataset.

Pause duration

For each condition (DMSO and Pla-B treated samples), the pause duration $d_i^{condition}$ was calculated in a window of ± 100 bp around the pause site m as previously described (Gressel et al., 2019; Gressel et al., 2017) for intron-containing major isoforms exceeding 10 kbp in length. For intronless genes, the pause duration $d_i^{condition}$ was calculated in a window of ± 50 bp around the pause site m. The pause duration $d_i^{condition}$ was calculated using the previously calculated initiation frequency $I_i^{condition}$ and the antisense bias corrected mNET-seq coverage values $P_i^{condition}$ as

$$d_i^{condition} = s \cdot \frac{\sum_{\pm 100} P_i^{condition}}{I_i^{condition}}$$

The median pause duration (DMSO) was calibrated to resemble pause duration of previous experiments in K562 cells with a calibration factor *s* obtained in (Gressel et al., 2019; Gressel et al., 2017).

ChIP-seq data preprocessing and normalization

Paired-end 42 or 75 bp reads with additional 6 bp of barcodes were obtained for each of the samples. Reads were aligned using Bowtie 2 (2.3.5, RRID:SCR_005476) (Langmead and Salzberg, 2012) to both human hg38 (GRCh38) and *D. melanogaster* (BDGP6.28). Bam files were filtered with Samtools (1.3.1, RRID:SCR_002105) (Li et al., 2009) to remove alignments with MAPQ smaller than 7 (-q 7) and only proper pairs (-f 2) were selected. Further data processing was carried out using the R/Bioconductor environment. ChIP-seq coverages were obtained from piled-up counts for every genomic position, using physical coverage, that is, counting both sequenced bases covered by reads and unsequenced bases spanned between proper mate-pair reads. Data was normalized using added *D. melanogaster* RNA spike-ins. Normalization factors were obtained by dividing the total *D. melanogaster* read counts for each sample by the total read counts of the sample with the lowest read counts. ChIP-seq coverages were divided by the respective normalization factors.

Elongation velocity estimation

For each condition (DMSO and Pla-B treated samples), each feature (exon and intron) in all intron-containing major isoforms exceeding 10 kbp in length, or intronless genes, the elongation velocity $v_i^{condition}$ was calculated as

$$v_i^{condition} = \frac{s}{t \cdot k} \cdot \frac{\sum_{feature} tb_i^{condition}}{\sum_{feature} P_i^{condition}}$$

with calibration factor *s* (please see Pause duration). For meta-gene plotting, the formula was altered with a smoothing approach as follows

$$v_i^{condition} = \frac{s}{t \cdot k \cdot 200} \cdot \frac{\sum_{\pm 100} tb_i^{condition}}{\sum_{\pm 5} w_i \cdot P_i^{condition}}$$

with calibration factor *s* and a weighting vector $w_i = (1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1)$.

Datasets for mNET-seq against CTD ser5P in HeLa cells upon 4 h DMSO and Pla-B treatments were obtained from (Schlackow et al., 2017), GEO:GSE81662. Empigen treatment during IP was not performed in this publicly available dataset.