

Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome

Ryan A. Flynn^{a,1,2}, Albert E. Almada^{a,b,1}, Jesse R. Zamudio^a, and Phillip A. Sharp^{a,b,3}

^aDavid H. Koch Institute for Integrative Cancer Research, Cambridge, MA 02139; and ^bDepartment of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Phillip A. Sharp, May 12, 2011 (sent for review March 3, 2011)

Divergent transcription occurs at the majority of RNA polymerase II (RNAPII) promoters in mouse embryonic stem cells (mESCs), and this activity correlates with CpG islands. Here we report the characterization of upstream antisense transcription in regions encoding transcription start site associated RNAs (TSSa-RNAs) at four divergent CpG island promoters: *Isg2011*, *Tcea1*, *Txn1*, and *Sf3b1*. We find that upstream antisense RNAs (uaRNAs) have distinct capped 5' termini and heterogeneous nonpolyadenylated 3' ends. uaRNAs are short-lived with average half-lives of 18 minutes and are present at 1–4 copies per cell, approximately one RNA per DNA template. Exosome depletion stabilizes uaRNAs. These uaRNAs are probably initiation products because their capped termini correlate with peaks of paused RNAPII. The pausing factors NELF and DSIF are associated with these antisense polymerases and their sense partners. Knockdown of either NELF or DSIF results in an increase in the levels of uaRNAs. Consistent with P-TEFb controlling release from pausing, treatment with its inhibitor, flavopiridol, decreases uaRNA and nascent mRNA transcripts with similar kinetics. Finally, *Isg2011* induction reveals equivalent increases in transcriptional activity in sense and antisense directions. Together these data show divergent polymerases are regulated after P-TEFb recruitment with uaRNA levels controlled by the exosome.

non-coding RNA | polymerase pausing

RNA polymerase II (RNAPII) transcription is a highly regulated process controlling cell type and state. Recruitment of chromatin modifying factors and RNAPII to promoters by DNA binding transcription factors are key regulatory steps (1–3). However, genome-wide profiling of RNAPII indicates that this polymerase is bound and engaged in the early steps of transcriptional initiation at most active and many inactive genes in human embryonic stem cells suggesting postinitiation modes of regulation may occur more frequently than previously appreciated (4). Moreover, divergent transcription, defined by detection of populations of low abundance small RNAs (19–25 nts) generated by nonoverlapping (separated by approximately 250 bps) bidirectional transcription, was detected at the majority of transcriptional start sites (TSSs) in mouse embryonic stem cells (mESCs) (5). Polymerases engaged in divergent transcription near promoters were simultaneously described in human lung fibroblasts (6). Surprisingly, RNAPII only productively elongates in the protein-coding sense direction from these divergent promoters. Related results have been reported for several other eukaryotic systems (7–11). Altogether, these data suggest that control of RNAPII elongation and RNA stability may be major points of transcriptional regulation and that mechanisms controlling these processes may dictate whether a stable RNA molecule is synthesized.

In recent years it has become clear that RNAPII pausing is a major mode of transcriptional regulation (6, 12). The Negative Elongation Factor (NELF) and DRB-Sensitivity Inducing Factor (DSIF) protein complexes bind and arrest RNAPII 20–30 nts downstream of the TSS (13). Recruitment of P-TEFb to a paused RNAPII complex and subsequent phosphorylation of the RNA-

PII carboxyl-terminal domain (CTD) at serine 2, DSIF, and NELF results in the dissociation of NELF from the elongation complex and continuation of transcription (13). More recently it was recognized, in mESCs, that c-Myc stimulates transcription at over a third of all cellular promoters by recruitment of P-TEFb (12). Intriguingly in these same cells, NELF and DSIF have bimodal binding profiles at divergent TSSs. This suggests divergent RNAPII complexes might be poised for signals controlling elongation and opens up the possibility that in the antisense direction P-TEFb recruitment may be regulating release for productive elongation.

Cellular mechanisms for removal of improperly processed, spliced, or aberrantly transcribed products likely account for the instability of transcripts from divergent promoters. In *Saccharomyces cerevisiae* cryptic unannotated transcripts (CUTs) derived from promoter-proximal regions are stabilized in the absence of the exosome (7, 9, 14). The exosome, with 3' to 5' exonuclease activity, is a multisubunit protein complex important for degradation and processing of mRNA, rRNA, snoRNA, and tRNA (14–16). The phosphorylation state of the RNAPII CTD (17, 18) and sequence elements within the RNA can influence targeting of transcripts to the exosome (19). Upon exosome depletion in human cells, promoter upstream transcripts (PROMPTs) are stabilized farther upstream (approximately 1 kb) from antisense TSSa-RNAs and are detected in both sense and antisense orientations in the upstream promoter region (8). However, it is unclear how various promoter associated RNAs, including PROMPTs, relate to transcription from divergent mammalian promoters.

Although multiple studies have identified distinct RNA species from mammalian promoters, the precise mapping of RNAs produced from divergent CpG island promoters has not been described. In light of these questions, we sought to investigate RNAPII divergent transcription through a detailed biochemical analysis of the antisense transcripts. We have characterized upstream antisense RNAs or uaRNAs from four divergent promoters in mESCs. We show that antisense RNAs are predominantly 5' capped and have heterogeneous 3' ends ranging in size from 40–1,100 bases in length. Both sense and antisense RNAPII complexes were involved in RNAPII pausing and both depend on PTEF-b recruitment and phosphorylation for subsequent elongation. We further show that low steady-state levels of uaRNAs, at least in part, are due to their targeting by the RNA exosome. Finally, we characterize induction of antisense and sense tran-

Author contributions: R.A.F., A.E.A., J.R.Z., and P.A.S. designed research; R.A.F. and A.E.A. performed research; R.A.F., A.E.A., J.R.Z., and P.A.S. analyzed data; and R.A.F., A.E.A., J.R.Z., and P.A.S. wrote the paper.

The authors declare no conflict of interest.

¹R.A.F. and A.E.A. contributed equally to this work.

²Present address: Stanford University School of Medicine, Stanford, CA 94305.

³To whom correspondence should be addressed. E-mail: sharp@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1106630108/-DCSupplemental.

scription from the divergent promoter of the *Isg2011* gene to show that PTEF-b activation at both sense and antisense RNAPII complexes occur with similar kinetics.

Results

Divergent RNAPII Produces Low Abundant Capped Upstream Antisense RNAs (uaRNAs) with 3' Heterogeneity. To test whether short antisense TSSa-RNAs previously described (5) are derived from longer transcripts and to determine the structure of their 5' termini, we used rapid amplification of 5' complementary DNA ends (5'-RACE) to characterize divergent upstream antisense RNAs from the *Isg2011*, *Tcea1*, *Txn1*, and *Sf3b1* genes in V6.5 mESCs (Fig. 1). These genes were selected as representatives of divergent TSSs associated with CpG islands and spanning a range of expression levels. The positions of oligonucleotides for specific priming for the 5'-RACE overlapped sequences found in antisense TSSa-RNAs from each promoter. The dependence of the specific 5'-RACE products on treatment with Tobacco Acid Pyrophosphatase (TAP) indicated the presence of a capped structure (Fig. S1, lanes 2 and 3). The sequenced 5'-RACE products for the antisense TSSs revealed RNAs initiated upstream of the previously characterized antisense TSSa-RNAs for all four promoters (Fig. 1 and Fig. S24, leftward arrows). Two predominant uaRNA TSSs for *Isg2011* were identified with the most upstream

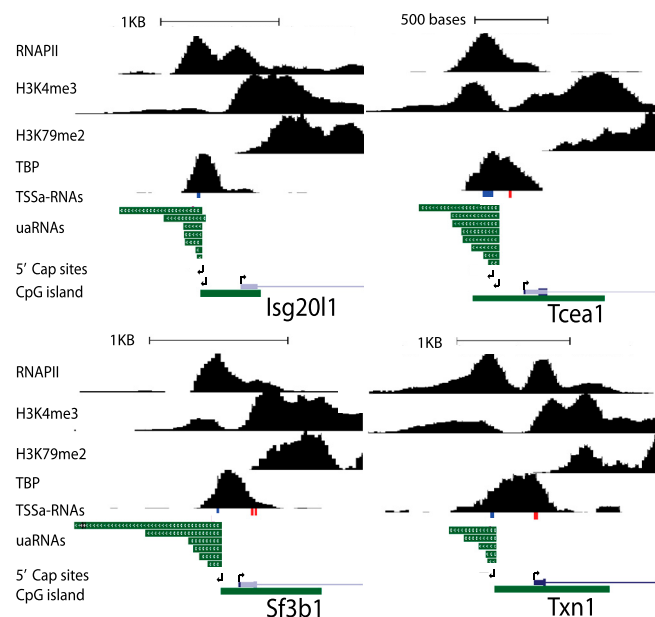


Fig. 1. Capped antisense RNA from divergent transcription initiate upstream of antisense TSSa-RNAs and display 3' heterogeneity. UCSC Genome browser view showing the location of detected 5' and 3' ends using rapid amplification of cDNA ends (RACE) at four selected CpG island divergent promoter genes: *Isg2011*, *Tcea1*, *Txn1*, and *Sf3b1*. 5'-RACE analysis was performed on upstream regions containing more than one overlapping antisense TSSa-RNA (5). Promoter regions are shown with the TSSs marked (arrows pointing in the direction of transcription). Arrows depicting antisense transcription are pointing to the left while sense TSSs are marked with arrows pointing to the right. Sense TSSs (right arrows) were labeled according to UCSC genome browser's known genes from UniProt, RefSeq, and GenBank. These were independently confirmed by 5'-RACE for *Isg2011*, *Tcea1*, and *Txn1*. 3'-RACE analysis yielded various uaRNA transcripts (green) for each divergent promoter. These range in length from approximately 50 to 1,100 nts. ChIP-seq binding profiles of RNAPII, H3K4me3, H3K79me2, TATA-binding protein (TBP), TSSa-RNA reads (antisense = blue; sense = red), and CpG island regions (green) are shown. ChIP-Seq data was obtained from the following published reports: RNAPII-8W16 (6), H3K4me3 and H3K79me2 (23), and TBP (22). We note the absence of a sense RNAPII ChIP-seq peak at the annotated TSS of *Tcea1*, likely explained in part to difficulties in mapping reads to this region since it contains high similarity (99%) with a location on chromosome 15. Scale bars are displayed at the top of each promoter region.

site 43 nts from the previously characterized clusters of antisense TSSa-RNAs. The *Tcea1* gene consists of two predominant antisense capped species with the most upstream 5' terminus 37 nts from the antisense TSSa-RNAs. One predominant capped 5' terminus was mapped for both *Txn1* and *Sf3b1* that extended 18 and 15 nts upstream from the antisense TSSa-RNA, respectively. The 5' capped termini of these uaRNAs are likely generated by RNAPII initiation events suggesting that the antisense TSSa-RNAs are products of subsequent reactions during elongation.

The 5' capped antisense RNA for each promoter align under a peak of RNAPII density near the end of the CpG island (Fig. 1). The segments between sense and antisense capped RNAs are on average 200–250 bps, comparable to the length of nucleosome-free regions associated with CpG islands (20). To test if transcripts extended beyond the sites of antisense 5' capped termini for these four genes, Reverse Transcriptase-quantitative PCR (RT-qPCR) with strand-specific RT primers was used to determine the orientation of RNA species in the upstream CpG island promoter (Fig. S3). Because these transcripts might be of low abundance, cellular RNAs were prepared with two rigorous DNase treatment steps and only signals dependent on RT were analyzed. Detectable antisense transcription at all four genes was confined to the region downstream of the mapped uaRNA 5' cap site. This provides additional evidence for initiation at these cap sites. Sense transcription within the CpG island upstream of the antisense cap site was probed for at all four regions but only detected at *Txn1* and was estimated by PCR cycles to be 80-fold less abundant than the antisense product. The inability to detect significant sense or antisense RNA signal upstream of the mapped uaRNA TSSs argues that the majority products from these regions initiate from the identified antisense TSSs.

We previously have characterized antisense RNAs from divergent TSSs by a selective enrichment protocol followed by Northern blot and observed a family of short RNAs spanning 30 to 200 nts (5). To more precisely define such RNAs, a 3'-RACE protocol was used to characterize the 3' ends of uaRNAs from the four divergent promoters. In this approach, adaptor sequences were ligated onto the free 3'-OH of large fractionated RNAs, followed by cDNA synthesis and PCR amplification using target-specific primers. The amplified products were cloned and sequenced to confirm their origin and define their 3' termini. Multiple nonpolyadenylated 3' ends were observed for uaRNAs at each TSS and were aligned to their respective promoters (Fig. 1). As few as five distinct antisense RNA 3' ends were detected for *Txn1* and as many as eight at *Tcea1*. The longest transcripts cloned were 703, 546, 415, and 1,100 nts for *Isg2011*, *Tcea1*, *Txn1*, and *Sf3b1*, respectively. However, it is likely that additional 3' ends exist because only a fraction of the 3'-RACE products were cloned and sequenced (Fig. S4). All 3'-RACE products were dependent on RT for amplification (Fig. S4). Transcripts under 100 nts were detected in the large fractionated RNA preparation. This probably reflects imperfect fractionation as similar patterns of transcripts were observed for all four genes. Because of this fractionation step, the relative levels of the various length RNAs cannot be estimated from the 3'-RACE products.

We compared the DNA sequences encompassed by uaRNAs with the location of RNAPII, TATA-binding protein (TBP) and chromatin modifications associated with active transcription determined by ChIP-seq in V6.5 mESCs (5, 21, 22) (Fig. 1). The shorter uaRNAs fell within the peak of bound RNAPII; however, the longest transcripts extended farther downstream. It is likely that the density of RNAPII in these downstream regions is below the threshold signal considered positive in the ChIP-seq analysis. Histone H3 lysine 4 trimethylation (H3K4me3) and TBP mark transcription initiation and H3 lysine 79 dimethylation (H3K79me2) correlates with elongation. The TBP density denoting the preinitiation complex was detected as a broad peak directly between the divergent RNAPII complexes. The H3K4me3 profile

generally extended the full length of the uaRNAs in the antisense direction with the exception of the longest Sf3b1 transcript. In contrast, ChIP-seq signal for H3K79me2 is absent in antisense transcribed regions for the four genes. This lack of signal might be due to limits in the sensitivity of the technique, but the same chromatin modification is clearly present in the sense direction downstream of the TSS for these four genes. This suggests differential activity of elongation complexes in the two directions.

To relate levels of uaRNAs to antisense TSSa-RNAs that were previously measured at 1 copy per 10 mESCs (5), RT-qPCR probes noted in Fig. S24 (“qPCR amplicon”) were used for absolute quantification. Copy number defined by molar equivalents as compared to a standard signal in the form of ssDNA was determined per ES cell equivalent of total RNA. The uaRNAs are present at 4.5, 1.8, 1.1, and 1.7 copies per cell for Isg2011, Tcea1, Txn1, and Sf3b1, respectively (Fig. S2B). These results indicate that uaRNAs are roughly 10-fold more abundant than previously characterized antisense TSSa-RNAs; present at approximately one copy per copy of genome sequences.

uaRNAs Are Substrates of the Exosome. Because previous studies have linked the exosome to nuclear surveillance of unannotated or cryptic transcripts, uaRNA stabilization and 3' termini were assayed upon exosome depletion. Exosc5 was targeted for knockdown with an shRNA-lentiviral delivery construct and depletion was confirmed 48 h after infection (Fig. S5). RT-qPCR was used to determine relative steady-state levels of uaRNAs (upstream antisense probe) and spliced sense mRNA (exon1-exon2 probe) between knockdown (shExosc5) and empty vector control (pLKO.1) for all four genes. Across multiple biological replicates, Exosc5 depletion led to a 2.5–3.5-fold increase in uaRNA levels, while spliced sense mRNAs were also slightly elevated yet below statistical significance (Fig. 2A). Next, we assayed for uaRNAs by DNA Southern blot of 3'-RACE products in control and exosome-depleted cells. After optimization of minimal PCR cycles and multiple probe validation of signal (Fig. S6B), the most abundant uaRNA forms for each gene were observed. In Exosc5 knockdown samples, the numbers and abundance of long RNAs increased compared to control virus infected cells further supporting uaRNAs as substrates for the exosome (Fig. 2B). These results show that upstream antisense RNAPII elongates to produce heterogeneous RNAs that are substrates for the exosome.

RNAPII Pausing Factors Regulate uaRNA Transcription. The RNAPII pausing factors that associate with promoter proximal stalled RNAPII are composed of NELF (NELF-A,B,C/D, and E) and DSIF (Supt4h and Supt5h). In addition, RNAPII acquires phosphorylation at Ser5 on the carboxyl-terminal domain early in transcription and this modification peaks in abundance around the pause site. We first aligned ChIP-seq profiles of RNAPII-Ser5P, Supt5h, and NELF-A determined in V6.5 mESCs to uaRNA transcribed regions. For all four genes, the peaks of RNAPII-Ser5P, Supt5h, and NELF-A directly overlap the uaRNA TSS supporting postinitiation regulation by RNAPII pausing in the antisense direction (Fig. 3A). To test whether RNAPII upstream complexes are poised for transcription in both directions, we performed shRNA-mediated knockdown of NELF-A, NELF-E and Supt4h, with each providing potent targeted mRNA loss (Fig. 3B). Depletion of either NELF subunit resulted in near twofold increases in uaRNA and spliced mRNA transcripts for all four genes across six biological replicates (Fig. 3C). Supt4h knockdown also resulted in twofold increases for both uaRNAs and spliced mRNA transcripts. Together these data argue NELF and DSIF complexes are equivalently active in binding and regulating paused RNAPII complexes in both directions at divergent promoters.

P-TEFb Regulates Elongation of uaRNAs. P-TEFb promotes RNAPII elongation in the sense direction for most if not all genes; how-

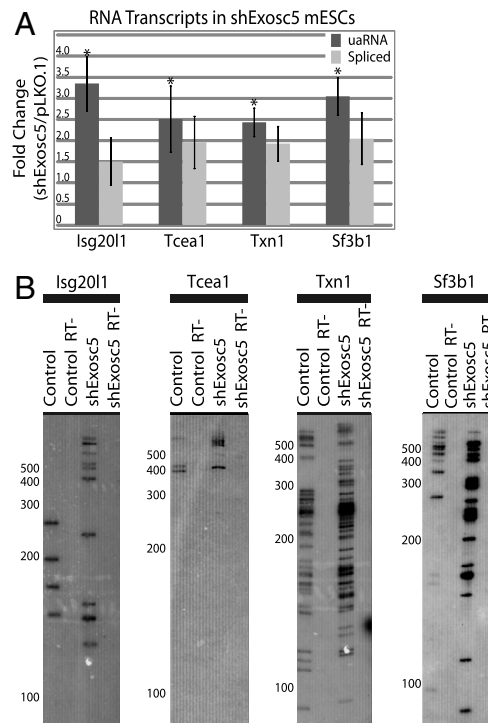


Fig. 2. uaRNA transcripts are substrates for the exosome. (A) Relative levels of uaRNA (gray amplicon in Fig. S2A, dark gray) and spliced mRNA (exon1-2 probe, light gray) transcripts in mESCs infected with virus containing an shRNA targeting Exosc5 (shExosc5) and assayed by RT-qPCR (probes shown in Table S3). Transcript levels were normalized to virus-infected cells without shRNA, empty vector (pLKO.1), and normalized to β -actin levels. Values represent four biological replicates and error represents the respective SEM. Asterisks represent significance of $p < 0.05$ in two-sided t test. (B) 3'-RACE followed by Southern blot analysis of control and shExosc5 treated mESC RNA. The Southern blots were probed with probe 1 shown in Fig. S6A. Minus RT lanes refer to 3'-RACE experiments with no reverse transcriptase added in the RT step of the procedure. Migration of 100bp molecular weight ladder (NEB) is marked on the left.

ever, its role in antisense transcription at divergent promoters has not been examined. We used flavopiridol, a small molecule drug with high specificity for CDK9 inhibition to test P-TEFb's requirement for RNA synthesis at divergent TSSs. uaRNA, spliced mRNA, and nascent mRNA transcripts (exon1-intron1 probe) were measured at all four genes in mock (DMSO) or 1 μ M flavopiridol-treated mESCs. Treatment with this flavopiridol concentration for 1 h was previously shown to not affect global RNAPII-Ser5P levels while dramatically reducing RNAPII-Ser2P and Supt5h phosphorylation in these cells, indicating a block of transcriptional elongation but not initiation (12). In flavopiridol-treated cells, the nascent mRNA transcripts for all four genes were reduced to 5–12% of mock-treated controls confirming a block in elongation (Fig. 4A). Interestingly, steady-state uaRNA transcript levels decreased to 19–25% of mock-treated controls using RT-qPCR probes that require transcription of approximately 150 nts or longer from the uaRNA TSS. Spliced mRNA levels were unchanged suggesting stable transcripts over this time course. We next determined uaRNA decay rates using flavopiridol treatment over a 1 h time course (Fig. 4B). Half-lives of 27, 19, 14, and 13 min were estimated for uaRNAs from Isg2011, Tcea1, Txn1, and Sf3b1, respectively (Fig. S7).

The large decrease in uaRNA levels following loss of CDK9 activity supports P-TEFb dependent release from paused polymerase and bidirectional recruitment of P-TEFb at CpG island promoters. To confirm P-TEFb-dependent transcription, RNA produced from divergent TSSs for all four genes was measured following removal of flavopiridol. The uaRNAs (Fig. 4C, left

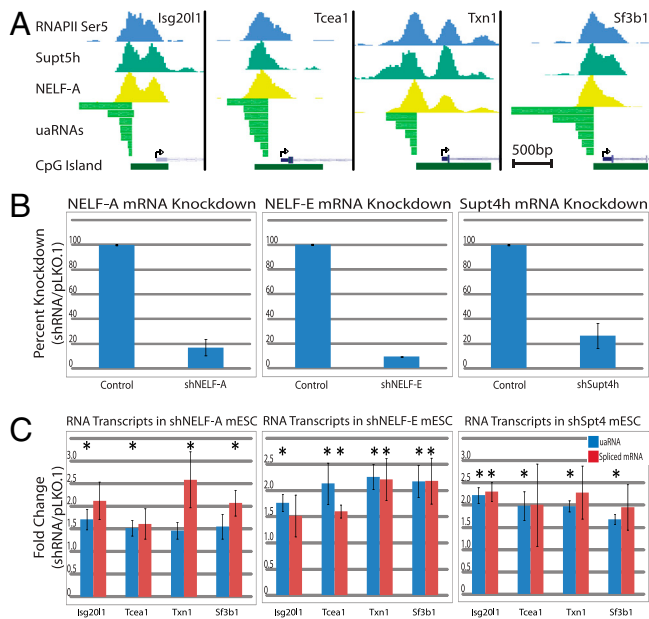


Fig. 3. Pausing factors at the antisense RNAPII complex regulate uaRNA levels. (A) UCSC Genome browser views of the four divergent promoter regions displaying ChIP-seq binding profiles (13) of RNAPII-Ser5P, Supt5h, NELF-A, and the full length antisense RNA transcripts. Each region diagrammed spans 2 kb and scale bars represent 500 bp. (B) Relative gene expression of Nelf-A, Nelf-E, and Supt4h in control and shRNA knockdown mESC after 48 h of selection measured by Taqman RT-qPCR assay. Values represent six biological replicates and error represents the respective SEM. (C) Transcript changes in shNELF-A, shNELF-E, and shSupt4h mESC lines as determined by RT-qPCR. uaRNA and spliced mRNA levels are represented by blue and red bars, respectively (probes shown in Table S3). Values represent six biological replicates and error shows the respective SEM. Asterisks represent significance of $p < 0.05$ in two-tailed *t* test.

panel) and nascent mRNA transcripts (Fig. 4C, right panel) recovered with similar kinetics and reached control steady-state levels by 30 min after flavopiridol removal. The similar recovery rates at both TSSs further supports P-TEFb acting on both divergent RNA polymerases to promote elongation.

Transcriptional Induction of Isg2011 Similarly Increases mRNA and uaRNA Levels. Interferon-stimulated 20 kDa exonuclease-like 1, Isg2011, is one of two homologs of an apoptosis-enhancing exonuclease. To determine how divergent paused RNAPII complexes respond to gene activation, doxorubicin, a DNA intercalator and inducer of double stranded breaks (23), was used to induce apoptosis in mESCs. Treatment with 1 μ M doxorubicin for 1.5, 4, and 6 h was followed by measurement of Isg2011 uaRNA, nascent mRNA, and spliced mRNA levels. Isg2011 transcriptional output in either direction did not significantly change with 1.5 h of treatment. However, both uaRNA and spliced mRNA transcripts had equivalent induction levels of 8- and 12-fold following 4- and 6-h treatments, respectively (Fig. 5). The nascent mRNA transcripts showed only a twofold increase at 6 h of treatment indicating tightly coordinated pre-mRNA processing. Divergent TSS products for thioreductase 1, Txn1, which are not expected to respond to DNA damage, did not change with treatment and served as an additional control for transcription fidelity during cellular stress. These results support a model for gene activation at divergent CpG island promoters preceded by stimulation of elongation in both directions.

Discussion

The detection of capped 5' termini on uaRNAs for all four studied promoters strongly supports a distinct and specific initiation

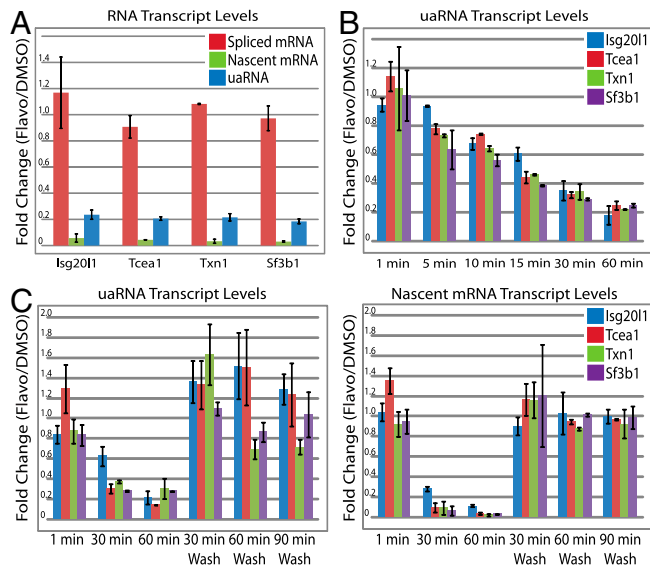


Fig. 4. uaRNAs are P-TEFb-dependent transcripts and have short half-lives. (A) Relative levels of spliced mRNA (red), nascent RNA (exon1-intron 1 probe, green), and uaRNA (blue) transcripts, as measured by RT-qPCR, after a 1 μ M flavopiridol treatment for 1 h. (B) uaRNA transcript levels assayed by RT-qPCR from amplicon shown in Fig. S2A over a 1 h time course with 1 μ M flavopiridol. (C) uaRNA (left panel) and nascent sense RNA (right panel) levels assayed by RT-qPCR over an hour 1 μ M flavopiridol treatment followed by a phosphate buffered saline (PBS) wash off of flavopiridol at the indicated times. Isg2011, Tcea1, Txn1, and Sf3b1 RNA transcripts are shown in blue, red, green, and purple, respectively. All values are relative to mock (DMSO) treated cells and normalized to β -Actin. Values represent two biological replicates with the error representing the respective SEMs. All probe sequences are shown in Table S3.

event from antisense RNAPII complexes at divergent promoters. Further, RNAPII-Ser5P, NELF, and DSIF profiles at divergent TSSs suggest that these antisense RNAPII complexes are poised for transcription elongation. Correspondingly, we find that depletion of NELF and DSIF, factors known to promote the pausing of RNAPII, modestly increases steady-state uaRNA levels. This is consistent with the model that the two divergent and paused com-

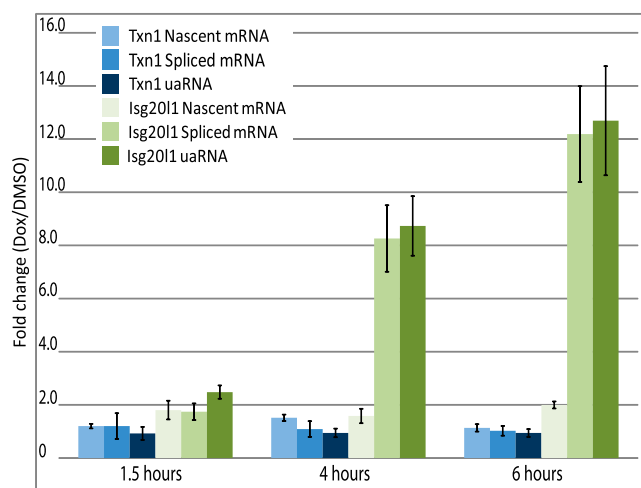


Fig. 5. Divergent sense and antisense transcription induced with similar kinetics. mESCs were treated with 1 μ M doxorubicin for 1.5, 4, and 6 h after which total RNA was collected. RT-qPCR analysis to determine the relative fold change of nascent, spliced mRNA, and uaRNA transcription followed. Changes in transcript levels for two genes, Isg2011 and Txn1, green and blue bars, respectively, are shown. Values represent biological triplicates and error of the respective SEM. All probes are shown in Table S3.

13. Peterlin BM, Price DH (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 23:297–305.
14. Houseley J, LaCava J, Tollervey D (2006) RNA-quality control by the exosome. *Nat Rev Mol Cell Biol* 7:529–539.
15. Mitchell P, Petfalski E, Shevchenko A, Mann M, Tollervey D (1997) The exosome: A conserved eukaryotic RNA processing complex containing multiple 3' → 5' exoribonucleases. *Cell* 91:457–466.
16. Schmid M, Jensen TH (2008) The exosome: A multipurpose RNA-decay machine. *Trends Biochem Sci* 33:501–510.
17. Gudipati RK, Villa T, Boulay J, Libri D (2008) Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat Struct Mol Biol* 15:786–794.
18. Vasiljeva L, Kim M, Mutschler H, Buratowski S, Meinhart A (2008) The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain. *Nat Struct Mol Biol* 15:795–804.
19. Anderson JR, et al. (2006) Sequence-specific RNA binding mediated by the RNase PH domain of components of the exosome. *RNA* 12:1810–1816.
20. Ramirez-Carrozzi VR, et al. (2009) A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* 138:114–128.
21. Kagey MH, et al. (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467:430–435.
22. Marson A, et al. (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* 134:521–533.
23. Nitiss JL (2009) Targeting DNA topoisomerase II in cancer chemotherapy. *Nat Rev Cancer* 9:338–350.
24. Gilchrist DA, et al. (2008) NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev* 22:1921–1933.
25. Gilchrist DA, et al. (2010) Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* 143:540–551.
26. Lin S, Coutinho-Mansfield G, Wang D, Pandit S, Fu XD (2008) The splicing factor SC35 has an active role in transcriptional elongation. *Nat Struct Mol Biol* 15:819–826.
27. Nechaev S, et al. (2010) Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* 327:335–338.
28. Lebreton A, Tomecki R, Dziembowski A, Seraphin B (2008) Endonucleolytic RNA cleavage by a eukaryotic exosome. *Nature* 456:993–996.
29. Wang X, et al. (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature* 454:126–130.
30. Kanhere A, et al. (2010) Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol Cell* 38:675–688.
31. Boyer LA, et al. (2006) Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 441:349–353.