The Importance of Controlling Transcription Elongation at Coding and Noncoding RNA Loci

BENJAMIN S. SCRUGGS AND KAREN ADELMAN

Epigenetics and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709

Correspondence: adelmank@niehs.nih.gov

Here we discuss current paradigms for how transcription initiation and elongation control are achieved in mammalian cells, and how they differ at protein-coding mRNA genes versus noncoding RNA (ncRNA) loci. We present a model for the function of ncRNAs wherein the act of transcription is regulatory, rather than the ncRNA products themselves. We further describe how the establishment of transcriptionally engaged, but paused, RNA polymerase II impacts chromatin structure around divergent transcription start sites, and how this can influence transcription factor binding and mRNA gene activity in the region.

Recent advances in genomic technology have enabled new insights into gene regulation in mammalian cells. By probing chromatin structure and the locations of DNA-binding transcription factors (TFs), thousands of intergenic regions have been identified that show hypersensitivity to nucleases (e.g., DNase) and are occupied by a plethora of TFs and coactivator complexes. Further, these regions of open chromatin have been shown to broadly support assembly of the transcription machinery and initiation of RNA synthesis (Kim et al. 2010; Rhee and Pugh 2012; The ENCODE Project Consortium 2012; Andersson et al. 2014). We now recognize that the mammalian genome is not as globally repressive to Pol II initiation as was once assumed, with Pol II transcribing an abundance of noncoding RNA species (ncRNAs) in addition to protein-coding messenger RNAs (mRNAs). Notably, ncRNAs arise from within gene bodies as well as from extragenic space, with particularly large classes of ncRNAs generated from regions designated as enhancers (Kim and Shiekhattar 2015), and from upstream and antisense of mRNA genes (Core et al. 2008; Preker et al. 2008; Seila et al. 2008; Sigova et al. 2013). Whereas transcription elongation of mRNAs is highly processive and can extend tens of kilobases (kb) in length, most ncRNAs are <1 kb in length and show dispersed 3'-end locations, suggesting premature termination (Preker et al. 2008; Flynn et al. 2011; Ntini et al. 2013; Sigova et al. 2013; Kim and Shiekhattar 2015). Further, many ncRNAs are highly unstable and subject to directed exosome-mediated degradation (Flynn et al. 2011; Preker et al. 2011).

Collectively, these findings reveal that the mammalian species rely heavily on postinitiation control of RNA synthesis, by modulating the processivity of transcription elongation complexes and stability of the RNA products (Adelman and Lis 2012; Almada et al. 2013; Ntini et al. 2013). Moreover, the discovery that Pol II generates many short, unstable ncRNA transcripts raises a number of questions about the roles of these species in regulation of protein-coding genes. Herein, we present a model wherein the act of transcription elongation and promoter-proximal pausing of Pol II is the functionally relevant aspect of ncRNA analysis.

OVERVIEW OF THE TRANSCRIPTION CYCLE

Metazoan gene transcription can be regulated at several distinct steps in the transcription cycle (Adelman and Lis 2012). The promoter region surrounding the transcription start site (TSS) must first be rendered accessible through the removal of repressive nucleosomes (Fig. 1A). This typically occurs through the concerted action of DNA-binding TFs that associate with specific motifs in the DNA and cooperate to recruit chromatin-modifying enzymes. One particularly important class of TFs are the pioneer factors (PFs) (Fig. 1A; Zaret and Carroll 2011; Soufi et al. 2012), which are capable of recognizing their DNA-binding sites within a nucleosome. These pioneer TFs work together with other TFs to initiate chromatin remodeling and expose core promoter elements (Fig. 1B) around the TSS. General transcription factors and the Mediator complex are then able to bind the core elements, forming an assembly that recruits and positions Pol II for initiation of RNA synthesis (Fig. 1C).

Early elongation of the RNA is punctuated by promoter-proximal pausing, wherein the polymerase and nascent 25–60-nt-long RNA are bound by two complexes, DSIF (DRB sensitivity inducing factor) and NELF (negative elongation factor) (Fig. 1D). Notably, when bound by these factors, the paused polymerase is remarkably stable within the promoter proximal region, with a median half-life of 8–10 min (Henriques et al. 2013; Jonkers et al. 2014). Moreover, paused polymerase remains associated with the nascent RNA and competent to resume synthesis (Nechaev et al. 2010; Core et al. 2012). Thus, once Pol II is recruited and loaded onto the DNA by the activity of TFs, paused Pol II is competent to “wait” patiently for
additional signals that trigger release into productive synthesis (Henriques et al. 2013). Importantly, work in *Drosophila* revealed that this long-lived occupancy of promoters by paused elongation complexes impacts chromatin structure, with paused Pol II preventing local nucleosome assembly (Gilchrist et al. 2008, 2010, 2012). Although the role of pausing in establishing promoter nucleosome architecture in mammalian cells had not been previously established, paused polymerase was reported to influence the positioning and high turnover rate of the first downstream nucleosome (Schones et al. 2008; Skene et al. 2014).

The release of paused Pol II into productive synthesis represents one of the central regulatory steps governing metazoan gene expression (Adelman and Lis 2012). Pause release involves the kinase positive transcription elongation factor (P-TEFb), which phosphorylates DSIF, dissociating NELF from the elongation complex and enabling Pol II movement into the gene body (Fig. 1E; Yamada et al. 2006; Cheng and Price 2007). P-TEFb can be recruited by a variety of interaction partners, including DNA-binding TFs, histone modifications, and the general transcription machinery (Barboric et al. 2001; Eberhardy and Farnham 2002; Jang et al. 2005; Yang et al. 2005; Takahashi et al. 2011). Notably, P-TEFb is often recruited as part of a super elongation complex (SEC) that contains additional elongation factors and chromatin modifiers (Luo et al. 2012; Yu et al. 2015). In this way, release of promoter-paused Pol II is temporally coupled with loading of factors that facilitate processive elongation and maturation of the RNA transcript (Fig. 1F).

Even after pause release, transcription elongation is not highly processive by default, and gene regulation is increasingly appreciated to involve factors that modulate the progression of polymerase down the DNA template. Indeed, Pol II appears to be prone to termination in the absence of “processivity signals,” such as encountering the 5′-splice site, which recruits the U1 snRNP through base-pairing to the U1 snRNA (Fig. 1F). Although the underlying mechanisms remain to be defined, the association of U1 snRNP with nascent RNA is thought to shield Pol II from cryptic polyadenylation and termination sites within the transcribed region (Kaida et al. 2010; Berg et al. 2012). The loading of RNA processing factors onto the elongation complex therefore influences the fate of the RNA (Moore and Proudfoot 2009; Jensen et al. 2013). Further, productive elongation is facilitated by factors that modify or remodel nucleosomes encountered as Pol II moves downstream into the gene (Venkatesh and Workman 2015; Yu et al. 2015).

The rules described above were defined primarily through study of protein-coding genes and may not uniformly apply to the diverse repertoire of ncRNA species generated by Pol II. Thus, it will be of great interest to identify key factors or interactions that distinguish coding from noncoding RNA transcription and to define functions for ncRNA transcription.

**WIDESPREAD INITIATION OF TRANSCRIPTION REDEFINES THE RULES OF ENGAGEMENT**

The term “promoter” classically refers to a set of core DNA sequence elements that direct the assembly of the
general transcription factors near the TSS (Butler and Kadonaga 2002). The core promoter elements are recognized by the general transcription factors (Fig. 2A; Butler and Kadonaga 2002). The growing scaffold of transcription factors and the Mediator complex recruits and properly positions Pol II for initiation of protein-coding genes. Elegant in vitro transcription assays established basic rules regarding characteristic sequence motifs and spatial arrangements of these DNA elements to enable optimal transcription output (Burke and Kadonaga 1996; Pal et al. 2005; Kostrewa et al. 2009). The model that arose from this work suggested that transcription initiation was highly specific and tightly regulated to avoid the formation of unwanted and potentially deleterious RNA species.

Surprisingly, systematic analyses of DNA sequences surrounding experimentally defined TSSs demonstrated that only a small fraction of genes display any given canonical promoter motif (Lenhard et al. 2012; Core et al. 2014; Forrest et al. 2014; Scruggs et al. 2015). The most common motif is the Initiator (Chen and Struhl 1985; Kadonaga 2012), with approximately one-third of genes harboring a close match to this motif (Scruggs et al. 2015). However, the majority of mRNA genes have only a short CA dinucleotide present at the TSS rather than the full Initiator consensus. Moreover, <10% of active promoters display canonical TATA or BRE motifs (Fig. 2), leaving many TSSs to be specified by a single motif or degenerate core promoter-like elements (e.g., TATA-like motifs; Basehoar et al. 2004). Overall, the limited occurrence of core promoter motifs observed around start sites used in vivo demonstrates that initiation is less dependent on previously defined core promoter elements than presumed.

The low information content observed around TSSs suggested that initiation might be promiscuous in vivo, with widespread opportunities for Pol II to engage in RNA synthesis. Indeed, next-generation sequencing techniques that allow for sensitive detection of RNAs have revealed that transcription initiates at tens of thousands of noncoding loci across metazoan genomes (Kim et al. 2010; The ENCODE Project Consortium 2012; Andersson et al. 2014). Importantly, the DNA sequence elements and factors involved in transcription initiation of mRNA or ncRNA species appear indistinguishable (Core et al. 2014; Duttke et al. 2015; Scruggs et al. 2015). Thus, despite looser “rules for engagement” of Pol II than previously appreciated, transcription initiation appears to proceed through a unified mechanism at coding and noncoding TSSs.

Remarkably, given the low occurrence of core promoter motifs, our high-resolution mapping of TSSs in mouse macrophages and embryonic stem cells using the Start-seq strategy (Nechaev et al. 2010) indicates significant focus in initiation events (Scruggs et al. 2015; Williams et al. 2015). Start–seq identifies short (<100-nt) transcription start site–associated RNAs held within paused Pol II complexes. Accordingly, this strategy enables highly sensitive, single-nucleotide resolution detection of both mRNAs and short, unstable ncRNAs (Nechaev et al. 2010). Using this method, we found that nearly 40% of active mRNA promoters show focused initiation profiles (Fig. 3; Scruggs et al. 2015). This level of specificity is considerably greater than suggested from analysis of mature mRNA 5’ ends observed in CAGE (cap analysis gene expression) data (Rach et al. 2011; Lenhard et al. 2012). However, our findings are in agreement with a recent study of transcription initiation in human cells using GRO-cap, which is another method for defining the 5’ ends of nascent transcripts (Core et al. 2014).
Surprisingly, the presence of CpG islands did not appear to alter the focus of initiation events (Fig. 3D; Rach et al. 2011). Instead, promoters broadly displayed considerable focus, generating mRNAs with highly similar sequences and predicted RNA structures, with only a minority of genes displaying significant diversity in their 5' ends. Thus, the development of strategies to map 5' ends of nascent RNAs (Core et al. 2008, 2014; Nechaev et al. 2010) has unveiled new features of transcription initiation, revealing positional specificity of initiation events despite a dearth of traditional core promoter elements. As discussed below, this indicates that assembly of the transcription machinery to initiate at a specific location involves integration of the modest information content present within classically defined promoter motifs with other factors such as DNA structure, nucleosome positioning, and occupancy by TFs.

DIVERGENT TRANSCRIPTION INITIATION

An interesting feature of transcription initiation in many eukaryotes is divergent promoter architecture, wherein transcription originates from two nearby promoters that exist on opposite strands. The first examples of this promoter structure were defined in mammalian cells, where transcription of an mRNA gene on one strand is accompanied by synthesis of an ncRNA whose TSS is located upstream and on the opposite strand (Fig. 4A; Core et al. 2008; Preker et al. 2008; Seila et al. 2008; Flynn et al. 2011; Scruggs et al. 2015). Although the origin, prevalence, and putative functions of these upstream antisense RNAs (uaRNAs) have been debated (Wu and Sharp 2013; Duttke et al. 2015), it now seems clear that the vast majority of mammalian mRNA initiation sites (>75%) are coupled with antisense TSSs located within several hundred nucleotides upstream (Andersson et al. 2015; Scruggs et al. 2015).

ChIP-seq assays reveal two discrete complexes of Pol II and the TATA binding protein TBP, one at each of the divergent TSSs (Rhee and Pugh 2012; Core et al. 2014; Scruggs et al. 2015). Consistent with this, two distinct sets of core elements have been identified at divergent promoters, with similar motif frequency and information content at both mRNA and uaRNA loci (Duttke et al. 2015; Scruggs et al. 2015). Further, initiation is similarly focused at uaRNA and mRNA TSSs (Fig. 3D). Thus, we now appreciate that the seemingly “bidirectional” promoter structure actually arises from two distinct, oppositely oriented TSSs, each with similar core motif structure and assembly of the transcription machinery.

CONTROL OF TRANSCRIPTION ELONGATION: A CENTRAL REGULATORY STEP

Despite a seemingly unified mechanism driving the assembly of preinitiation complexes at mRNA and ncRNA loci (Core et al. 2014; Duttke et al. 2015; Scruggs et al. 2015), differences in transcriptional processivity often reveal themselves early in elongation. A striking
example of the asymmetry in elongation potential at mRNAs versus ncRNAs can be found in divergent mRNA–uaRNA pairs (Fig. 4B). We evaluated the RNA species generated at divergent TSSs in murine macrophages during an immune challenge by bacterial lipopolysaccharide (LPS). Start-seq was used to quantify the increases in short (<100-nt) nascent RNAs synthesized from early response mRNA genes within 30 min of immune stimulation, and compared with Start-RNAs observed at the coupled uaRNA TSSs. LPS treatment induced similar increases in early elongation complexes (measured by Start-RNA levels at the TSS ± 100 bp) were observed for the mRNA and uaRNAs of early response genes following 30 min activation with bacterial LPS (left). By comparison, evaluation of longer nascent RNAs (Bhatt et al. 2012; window from TSS to +250 bp) reveals that productive elongation is biased heavily toward the sense direction (right).

These findings underscore that much of the RNA repertoire within a cell is determined—not at the level of transcription initiation—but through differential elongation properties. As discussed above, engaging the splicing machinery and association of the U1 snRNP appears to play a key role in this process (Berg et al. 2012). Consistent with the fact that most mammalian mRNAs are spliced and have short 5′ untranslated regions (Bieberstein et al. 2012), U1 motifs are highly enriched downstream from sense TSSs (Fig. 5A, right; Andersen et al. 2012; Almada et al. 2013; Ntini et al. 2013; Scruggs et al. 2015) as compared with uaRNA TSSs (Fig. 5A, left). Conversely, poly(A) sequences (PASs) that trigger transcription termination and RNA polyadenylation (Proudfoot 2011) are depleted within protein-coding genes (Fig. 5B, right). In contrast, PASs occur at genomic background levels downstream from uaRNA TSSs (Fig. 5B, left; Andersen et al. 2012; Almada et al. 2013; Ntini et al. 2013), indicating that PASs have neither been selected for nor against in these regions.

The emerging picture is that a “U1–PAS axis” influences the differences in Pol II processivity in mRNA versus ncRNA loci (Almada et al. 2013; Ntini et al. 2013). However, the U1 motif is not ubiquitous within several hundred nucleotides of protein coding TSSs.
(Fig. 5A), nor are PASs encountered frequently by Pol II transcribing short ncRNAs. Thus, other mechanisms and factors must help define the processivity of Pol II. Such factors likely include alternate termination machineries like the Integrator complex (Lai et al. 2015) and histone modifications such as trimethylation of histone H3 at Lysine 4 (H3K4me3). The H3K4me3 mark is considered a defining feature of mRNA promoters and is enriched on the first downstream nucleosomes within active genes. This H3K4me3 mark is thought to distinguish promoters from enhancer loci and other regions that generate ncRNAs, which are enriched in monomethylation of H3 (H3K4me1; Calo and Wysocka 2013). Interestingly, H3K4me3 is also enriched over the initial 5’ splice site (Bieberstein et al. 2012), suggesting an intriguing connection between H3K4 methylation, the recruitment of the U1 snRNP, and processivity of Pol II elongation that will be interesting to probe in future.

**POTENTIAL FUNCTIONS OF NONCODING RNA SYNTHESIS**

Quantification of numerous ncRNA species has indicated that the majority is present at <1 copy per cell (Seila et al. 2008; Wu and Sharp 2013). Therefore, despite the appealing models of ncRNA function represented by Xist ncRNA coating the inactivated X-chromosome or Malat1 ncRNA nucleating nuclear speckle formation (Lee and Bartolomei 2013; Bergmann and Spector 2014), many ncRNAs are insufficiently plentiful to play such architectural roles in gene expression. Further, ncRNAs often arise from poorly conserved regions of the genome, tend not to undergo RNA splicing or polyadenylation, and display extremely short lifetimes in vivo (Kim et al. 2010; Marques et al. 2013; Schaukowitch et al. 2014).

So what purpose might this panoply of ncRNAs serve if many of them are poorly conserved, short-lived, and of low abundance? One possibility is that some enhancer RNAs and upstream antisense RNAs could work in cis, functioning as local targets of interaction while still engaged within transcribing or paused Pol II. For example, we recently proposed that nascent RNA emanating from paused Pol II might interact with RNA binding proteins and/or epigenetic modifiers that stimulate mRNA transcription (Henriques et al. 2013). In this scenario, one molecule of RNA generated and held in place at the site of enhancer RNA transcription could be sufficient to target factors to the desired enhancer–promoter locus or to stabilize a loop structure between enhancer and promoter (Kim and Shiekhattar 2015). In addition, the act of transcription itself has been shown to modify chromatin structure. For example, the establishment of stably paused Pol II impacts the chromatin environment by preventing local assembly of nucleosomes (Gilchrist et al. 2008, 2010, 2012). Chromatin-modifying enzymes travelling with Pol II deposit histone modifications that affect nearby transcription units (Smolle et al. 2012; Venkatesh et al. 2012; Marquardt et al. 2014), and further stimulate enhancer activity (Kaikkonen et al. 2013; Schaukowitch et al. 2014). We expand on these possibilities below.
using the example of divergent mRNA–ncRNA promoters as a framework for discussion.

**DIVERGENT PROMOTERS: INTERPLAY BETWEEN CODING AND NONCODING RNA SYNTHESIS**

TSSs for uaRNAs represent discrete sites of initiation that occur at variable distances from the mRNA TSS (Fig. 6A; Scruggs et al. 2015). Distinct assemblies of the transcription machinery are present at each TSS, with evidence of paused Pol II in both directions (Fig. 6B, Pol II and TBP; Escoubet-Lozach et al. 2011). Moreover, we found that highly positioned nucleosomes are located downstream from paused Pol II at each coupled TSSs, with the area between the polymerases depleted of nucleosomes (up to \( \approx 700 \text{ bp} \), Fig. 6, MNase and FAIRE; Ostuni et al. 2013). Thus, the location of the uaRNA TSS specifically delineates the position of upstream nucleosomes and the edge of open chromatin at mRNA promoters.

Importantly, we found that promoters supporting larger regions of nucleosome depletion in resting macrophages showed higher transcription factor binding and gene activation during immune challenge. We studied NF-κB binding in macrophages stimulated with bacterial LPS, which triggers NF-κB nuclear localization and chromatin association. We discovered that early response genes with more distant uaRNA TSSs and larger regions of nucleosome depletion were bound by significantly higher levels of NF-κB immediately following immune challenge than were genes with closer or absent uaRNA TSSs. Critically, the differences in NF-κB binding upstream of mRNA TSSs correlated with the size of the open chromatin and position of uaRNA TSSs and not the total number of motifs in the region (Scruggs et al. 2015). This suggested that a functional consequence of divergent promoter architecture is to position nucleosomes and define the region of accessible DNA upstream of stimulus-responsive genes.

Critically, the increase in NF-κB binding at genes with distant uaRNA TSSs translated into higher levels of gene activation, with the strongest induction and NF-κB occupancy at genes with uaRNA TSSs several hundred base pairs upstream (Scruggs et al. 2015). Hence, we found that mRNA promoters with distant antisense TSSs have larger nucleosome-depleted promoter regions, more accessible TF motifs, elevated stimulus-dependent TF binding, and greater transcription activation. We therefore suggest that the distance between divergent TSSs established in resting macrophages impacts promoter structure, which can alter profiles of stimulus-dependent gene induction.

**WHAT DEFINES THE LOCATION OF UPSTREAM ANTISENSE TSSs?**

The dearth of canonical core promoter elements around TSSs used in vivo suggests that multiple forms of information content are used for combinatorial control of transcription initiation. Initiator motifs are often present near TSSs (Fig. 7A; Inr motifs shown for each strand); however, we find a plethora of Initiator elements across the genome. In fact, nearly all mRNA TSSs have a close match to an Initiator motif present within 500 nt upstream, regardless of whether an antisense TSS is observed (Fig. 7C). So what distinguishes active sites of

![Figure 6](image-url) Variable promoter chromatin architectures reflect variable distances between divergent TSSs. Modified from Scruggs et al. 2015. (A) Start-RNA reads from mouse macrophages are shown in the antisense (purple) and sense (green) direction, both aligned with respect to sense TSSs. Genes are rank ordered by the distance between sense and antisense TSSs. (B) Pol II and TBP ChIP-seq, MNase-seq, and FAIRE-seq are shown for TSSs organized as in A.
transcription initiation from other possible sites? To identify information content outside of the canonical core promoter elements that define where uaRNAs initiate in mouse macrophages, we investigated locations of a panel of vertebrate TF motifs (JASPAR database). We found an enrichment of TF motifs within the nucleosome-deprived region defined by mRNA–uaRNA TSS pairs (Fig. 7A), with a peak of TF motifs just upstream of each divergent TSS (Fig. 7B).

We therefore propose that direct proximity between core promoter elements and transcription factor motifs provides the synergy necessary for establishing specific sites of transcription initiation. We note that transcription factors that are constitutively active in resting macrophages would be better suited for this task than signal-dependent factors like NF-κB that only bind their cognate sites under specific circumstances. Consistent with this hypothesis, a search for motifs that are enriched directly upstream of divergent TSSs observed in macrophages uncovered binding sites for PU.1, a master regulator of the myeloid lineage (Scruggs et al. 2015). PU.1 ChIP-seq confirms high level occupancy next to both mRNA and uaRNA TSSs (Scruggs et al. 2015). PU.1 is particularly appealing candidate for establishing the locations of divergent TSSs in macrophages because it is has properties of a pioneer factor (Nerlov and Graf 1998). Thus, we envision that cell type–specific patterns of transcription factor binding help Pol II navigate the sea of degenerate core promoter elements and guide the transcription machinery to specific sites for transcription initiation of both mRNAs and uaRNAs.

**MODEL FOR ESTABLISHMENT OF DIVERGENT PROMOTER ARCHITECTURE**

We propose the following model to illustrate how transcription factors, core promoter elements and paused
Pol II work synergistically to establish divergent promoter architecture: Lineage-specific transcription factors such as PU.1 act as pioneers to transiently expose transcription factor motifs and core promoter elements near sense and antisense TSSs (Fig. 8A). Additional TFs interact cooperatively with lineage-specific PFs to establish regions of factor enrichment that recruit chromatin remodelers necessary for nucleosome removal (Fig. 8B). However, most transcription factors bind and unbind their cognate motifs on the order of seconds (Hager et al. 2009), and so this initial collection of TFs and open chromatin is subject to collapse in the absence of some form of stabilization.

We hypothesize that the assembly of the transcription machinery and pausing of Pol II represents the formation of a longer-term scaffold, given that paused polymerase can remain stably engaged for tens of minutes in the absence of P-TEFb and productive pause release (Fig. 8C). Early elongation and pausing of Pol II would also position nucleosomes downstream from each TSS and divergent elongation could create negative supercoiling between the two start sites, disfavoring nucleosome

---

**Figure 8.** Establishment of divergent promoters is a multistep process involving coordination between transcription factors and Pol II. (A) Pioneer factors (PF) recognize their DNA-binding motifs near nucleosome-occluded TSSs. (B) Cooperation between TFs opens chromatin and recruits remodelers, exposing core promoter elements. (C) Initiation and pausing of Pol II helps maintain open chromatin, establishing a nucleosome-depleted region between divergent TSSs. (D) Chromatin-modifying factors associated with elongating Pol II establish distinct epigenetic landscapes downstream from divergent TSSs. (E) Maintenance of open chromatin renders motifs for binding of signal-dependent transcription factors (SDF, orange) accessible for immediate binding during gene activation.
occupancy (Seila et al. 2008; Teves andHenikoff 2014). Following elongation of Pol II in either direction, additional Pol II is recruited and pausing to prevent reformation of nucleosomes. Thus, once the appropriate chromatin architecture is established, the transcriptional machinery and TFs could function cooperatively to maintain this state (Fig. 8D).

Histone-modifying factors that travel with elongating Pol II could then contribute to the epigenetic landscape, lying down an extended “memory” of transcriptional activity (Kaikkonen et al. 2013; Marquardt et al. 2014; Venkatesh and Workman 2015). There is an asymmetry of histone modifications downstream from divergent TSSs, with H3K4me3 predominating over mRNAs and uaRNAs being modified with H3K4me1, characteristic of enhancers (Scruggs et al. 2015). We suggest that this asymmetry results from the recruitment of different chromatin-modifying factors to the divergent polymersomes (Fig. 8D), and might reflect the differential processivity in each direction.

Finally, the chromatin organization established around divergent TSSs in resting cells could prime promoter units for signal responsiveness: Motifs for signal-dependent TFs (SDFs) positioned between mRNA and uaRNA TSSs would be readily bound following stimulation (Fig. 8E), whereas those outside this region are refractory to immediate occupancy. In this way, we envision that a promoter could be tuned for a certain level of responsiveness through the establishment of the uaRNA TSS at a closer versus more distant location.

**CONCLUSION**

Studies of coordination between coding and noncoding transcription are uncovering new paradigms and mechanisms for gene regulation (Rinn and Chang 2012; Sigova et al. 2013; Kim and Shiekhattar 2015; Scruggs et al. 2015). Although many sites of noncoding transcription generate unstable RNAs that are rapidly degraded, we propose that the recruitment of transcription machinery to these locations can be nonetheless instrumental to maintain appropriate chromatin architecture in the region. As an example, we describe how establishment of paused Pol II upstream and antisense of protein-coding genes delineates the size of the nucleosome-depleted promoter region. Further, we provide evidence that the extent of open promoter chromatin impacts gene activity by tuning signal-dependent transcription factor binding. Therefore, we envision broadly that targeted recruitment of the transcription machinery and Pol II pausing can be useful to regulate chromatin architecture, regardless of the nature of the ncRNA generated.

**ACKNOWLEDGMENTS**

This work was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences to K.A. (Z01 ES101987).

**REFERENCES**

Adelman K, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: Emerging roles in metazoans. *Nat Rev Genet* 13: 720–731.

Almada AE, Wu X, Kriz AJ, Burge CB, Sharp PA. 2013. Promoter directionality is controlled by U1 snRNP and polyadenylation signals. *Nature* 499: 360–363.

Andersen PK, Lykke-Andersen S, Jensen TH. 2012. Promoter-proximal polyadenylation sites reduce transcription activity. *Genes Dev* 26: 2169–2179.

Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T, et al. 2014. An atlas of active enhancers across human cell types and tissues. *Nature* 507: 455–461.

Andersson R, Chen Y, Core L, Lis JT, Sandelin A, Jensen TH. 2015. Human gene promoters are intrinsically bidirectional. *Mol Cell* 60: 346–347.

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

Basehoar AD, Zanton SJ, Pugh BF. 2004. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* 116: 699–709.

Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaido D, Zhang Z, Cho S, Sherrill-Mix S, Wan L, et al. 2012. U1 snRNP determines mRNA length and regulates isoform expression. *Cell* 150: 53–64.

Bergmann JH, Spector DL. 2014. Long non-coding RNAs: Modulators of nuclear structure and function. *Curr Opin Cell Biol* 26: 10–18.

Bhatt DM, Pandya-Jones A, Tong A-J, Barozzi I, Lissner MM, Nissen RM, Kanazawa S, Spann NJ, Crotti A, Stender J, Ghisletti S, Reichart D, Andersen PK, Lykke-Andersen S, Jensen TH. 2012. Analysis of nascent RNA identifies a unified architectural principle in metazoans. *Nat Rev Genet* 13: 720–731.

Burke TW, Kadonaga JT. 1996. *Drosophila* TFIIID binds to a conserved downstream basal promoter element that is present in many TATA-box-deficient promoters. *Genes Dev* 10: 711–724.

Butler JEF, Kadonaga JT. 2002. The RNA polymerase II core promoter: A key component in the regulation of gene expression. *Genes Dev* 16: 2583–2592.

Calo E, Wysocka J. 2013. Modification of enhancer chromatin: What, how, and why? *Mol Cell* 49: 825–837.

Chen W, Struhl K. 1985. Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. *EMBO J* 4: 3273–3280.

Cheng B, Price DH. 2007. Properties of RNA polymerase II elongation complexes before and after P-TEFb-mediated transition into productive elongation. *J Biol Chem* 282: 21901–21912.

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

Core LJ, Waterfall JJ, Gilchrist DA, Fargo DC, Kwak H, Adelman K, Lis JT. 2012. Defining the status of RNA polymerase at promoters. *Cell Rep* 2: 1025–1035.

Core LJ, Martins AL, Danko CG, Waters CT, Siepel A, Lis JT. 2014. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat Genet* 46: 1311–1320.

Dutkewych SC, Lacadie SA, Ibrahim MM, Glass CK, Corcoran DL, Benner C, Heinz S, Kadonaga JT, Ohler U. 2015. Human promoters are intrinsically directional. *Mol Cell* 57: 674–684.

Eberhardt SR, Farnham PJ. 2002. Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cod promoter. *J Biol Chem* 277: 40156–40162.

Escoubet-Lozach L, Benner C, Kaikkonen MU, Lozach J, Heinz S, Spann NJ, Crotti A, Stender J, Ghisletti S, Reichart D,
et al. 2011. Mechanisms establishing TLR4-responsive activation states of inflammatory response genes. PLoS Genet 7: e1002401.

Flynn RA, Almada AE, Zamudio JR, Sharp PA. 2011. Antisense RNA polymerase II divergent transcripts are P-TEFb dependent and substrates for the RNA exosome. Proc Natl Acad Sci 108: 10460–10465.

Forrest ARR, Kawai H, Rehli M, Baillie JK, de Hoon MJL, Haberle V, Lassmann T, Kulakovskiy IV, Lizio M, Ittoh M, et al. 2014. A promoter-level mammalian expression atlas. Nature 507: 462–470.

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

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

Gilchrist DA, Fromm G, dos Santos G, Pharm LN, McDaniel IE, Burkholler A, Fargo DC, Adelman K. 2012. Regulating the regulators: The pervasive effects of Pol II pausing on stimulus-responsive gene networks. Genes Dev 26: 933–944.

Hager GL, McNally JG, Misteli T. 2009. Transcription dynamics. Mol Cell 35: 741–753.

Henriques T, Gilchrist DA, Nechaev S, Bern M, Muse GW, Burkholler A, Fargo DC, Adelman K. 2013. Stable pausing by RNA polymerase II provides an opportunity to target and integrate regulatory signals. Mol Cell 52: 517–528.

Jang MK, Mochizuki K, Zhou M, Jeong HS, Brady JN, Ozato K. 2005. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. Mol Cell 19: 523–534.

Jensen TH, Jacquier A, Libri D. 2013. Dealing with pervasive transcription. Mol Cell 52: 473–484.

Jones PA. 2012. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. Nat Rev Genet 13: 484–492.

Jonkers I, Kwak H, Lis JT. 2014. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. Elife 3: e02407.

Kadonaga JT. 2012. Perspectives on the RNA polymerase II core promoter. Wiley Interdisciplinary Rev Dev Biol 1: 40–51.

Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G. 2010. U1 snRNP proteins are mRNAs from premature cleavage and polyadenylation. Nature 468: 664–668.

Kaikonen MU, Spann NJ, Heinz S, Romanski CE, Allison KA, Stender JD, Chun BH, Tough DF, Prinjha RK, Benner C, et al. 2013. Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. Mol Cell 51: 310–325.

Kim T-K, Shiekhattar R. 2015. Architectural and functional commonalities between enhancers and promoters. Cell 162: 948–959.

Kim T-K, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, Harmin DA, Laptevicz M, Barbara-Haley K, Kuersten S, et al. 2010. Widespread transcription at neuronal activity-regulated enhancers. Nature 465: 182–187.

Kostrewa D, Zeller ME, Armauc K-J, Seizl M, Leike K, Thomm M, Cramer P. 2009. RNA polymerase II-TFIIB structure and mechanism of transcription initiation. Nature 462: 323–330.

Lai F, Gardini A, Zhang A, Shiekhattar R. 2015. Integrator mediates the biogenesis of enhancer RNAs. Nature 525: 399–409.

Lee JT, Bartolomei MS. 2013. X-inactivation, imprinting, and long noncoding RNAs in health and disease. Cell 152: 1308–1323.

Lenhard B, Sandelin A, Carninci P. 2012. Metazoan promoters: Emerging characteristics and insights into transcriptional regulation. Nat Rev Genet 13: 233–245.

Luo Z, Lin C, Shilatifard A. 2012. The super elongation complex (SEC) family in transcriptional control. Nat Rev Mol Cell Biol 13: 543–547.

Marquardt S, Escalante-Chong R, Pho N, Wang J, Churchman LS, Springer M, Buratowski S. 2014. A chromatin-based mechanism for limiting divergent noncoding transcription. Cell 157: 1712–1723.

Marques AC, Hughes J, Graham B, Kowalczyk MS, Higgs DR, Ponting CP. 2013. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. Genome Biol 14: R131.

Moore MJ, Proudfoot NJ. 2009. Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 136: 688–700.

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

Nerlov C, Graf T. 1998. PU.1 induces myeloid lineage commitment in multipotent hematopoietic progenitors. Genes Dev 12: 2403–2412.

Ntini E, Järvelin AI, Bornholdt J, Chen Y, Boyd M, Jørgensen M, Andersen R, Hoof I, Schein A, Andersen PR, et al. 2013. Polyadenylation site—induced decay of upstream transcripts enforces promoter directionality. Nat Struct Mol Biol 20: 923–928.

Ostuni R, Piccolo V, Barozzi I, Polletti S, Terramani A, Bonifacio S, Curina A, Prosperi E, Ghisletti S, Natali G. 2013. Latent enhancers activated by stimulation in differentiated cells. Cell 152: 229–241.

Pal M, Ponticelli AS, Luse DS. 2005. The role of the transcription bubble and TFII B in promoter clarity by RNA polymerase II. Mol Cell 19: 101–110.

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

Preker P, Almvig K, Christensen MS, Valen E, Mapendano CK, Sandelin A, Jensen TH. 2011. PROMoter uPstream Transcripts share characteristics with mRNAs and are produced upstream of all three major types of mammalian promoters. Nucleic Acids Res 39: 7179–7193.

Proudfoot NJ. 2011. Ending the message: Poly(A) signals then and now. Genes Dev 25: 1770–1782.

Rach EA, Winter DR, Krijger LM, Corcoran DL, Ni T, Zhu J, Ohler U. 2011. Transcription initiation patterns indicate divergent strategies for gene regulation at the chromatin level. PLoS Genet 7: e1001274.

Rhee HS, Pugh BF. 2012. Genome-wide structure and organization of euchromatic pre-initiation complexes. Nature 483: 295–301.

Rinn JL, Chang HY. 2012. Genome regulation by long noncoding RNAs. Annu Rev Biochem 81: 145–166.

Schaukowitch K, Joo J-Y, Liu X, Watts JK, Martinez C, Kim T-K. 2014. Enhancer RNA facilitates NELF release from immediate early genes. Mol Cell 56: 29–42.

Schones DE, Cui KR, Cuddahap S, Roh T-Y, Barski A, Wang Z, Wei G, Zhao K. 2008. Dynamic regulation of nucleosome positioning in the human genome. Cell 132: 887–899.

Scruggs BS, Gilchrist DA, Nechaev S, Muse GW, Burkholler A, Fargo DC, Adelman K. 2015. Bidirectional transcription arises from two distinct hubs of transcription factor binding and active chromatin. Mol Cell 58: 1101–1112.

Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, Flynn RA, You Y, Nature RA, Sharp PA. 2008. Divergent transcription from active promoters. Science 322: 1849–1851.

Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, Guenther MG, Almada AE, Lin C, Sharp PA, Gisslaurakis CC, et al. 2013. Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. Proc Natl Acad Sci 110: 2876–2881.
Skene PJ, Hernandez AE, Groudine M, Henikoff S. 2014. The nucleosomal barrier to promoter escape by RNA polymerase II is overcome by the chromatin remodeler Chd1. *Elife* 3:e02042.

Smolle M, Venkatesh S, Gogol MM, Li H, Zhang Y, Florens L, Washburn MP, Workman JL. 2012. Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nat Struct Mol Biol* 19:884–892.

Soufi A, Donahue G, Zaret KS. 2012. Facilitators and impediments of the pluripotency reprogramming factors’ initial engagement with the genome. *Cell* 151:994–1004.

Takahashi H, Parmely TJ, Sato S, Tomomori-Sato C, Banks CAS, Kong SE, Szutorisz H, Swanson SK, Martin-Brown S, Washburn MP, et al. 2011. Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* 146:92–104.

Teves SS, Henikoff S. 2014. Transcription-generated torsional stress destabilizes nucleosomes. *Nat Struct Mol Biol* 21:88–94.

The ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74.

Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* 16:178–189.

Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, Workman JL. 2012. Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature* 489:452–455.

Williams LH, Fromm G, Gokey NG, Henriques T, Muse GW, Burkholder A, Fargo DC, Hu G, Adelman K. 2015. Pausing of RNA polymerase II regulates mammalian developmental potential through control of signaling networks. *Mol Cell* 58:311–322.

Wu X, Sharp PA. 2013. Divergent transcription: A driving force for new gene origination? *Cell* 155:990–996.

Yamada T, Yamaguchi Y, Inukai N, Okamoto S, Mura T, Handa H. 2006. P-TEFb-mediated phosphorylation of hSpt5 C-terminal repeats is critical for processive transcription elongation. *Mol Cell* 21:227–237.

Yang Z, Yik JHN, Chen R, He N, Moon KJ, Ozato K, Zhou Q. 2005. Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol Cell* 19:535–545.

Yu M, Yang W, Ni T, Tang Z, Nakadai T, Zhu J, Roeder RG. 2015. RNA polymerase II-associated factor 1 regulates the release and phosphorylation of paused RNA polymerase II. *Science* 350:1383–1386.

Zaret KS, Carroll JS. 2011. Pioneer transcription factors: Establishing competence for gene expression. *Genes Dev* 25:2227–2241.