Published in final edited form as:

Nat Rev Mol Cell Biol. 2015 March ; 16(3): 167–177. doi:10.1038/nrm3953.

Getting up to speed with transcription elongation by RNA polymerase II

Iris Jonkers¹ and John T. Lis²

¹University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands ²Department of Molecular Biology and Genetics, Cornell University, 416 Biotechnology Building, 14853, Ithaca, New York, USA

Abstract

Recent advances in sequencing techniques that measure nascent transcripts and that reveal the positioning of RNA polymerase II (Pol II) have shown that the pausing of Pol II in promoter-proximal regions and its release to initiate a phase of productive elongation are key steps in transcription regulation. Moreover, after the release of Pol II from the promoter-proximal region, elongation rates are highly dynamic throughout the transcription of a gene, and vary on a gene-by-gene basis. Interestingly, Pol II elongation rates affect co-transcriptional processes such as splicing, termination and genome stability. Increasing numbers of factors and regulatory mechanisms have been associated with the steps of transcription elongation by Pol II, revealing that elongation is a highly complex process. Elongation is thus now recognized as a key phase in the regulation of transcription by Pol II.

It has become increasingly evident that transcription elongation by RNA polymerase II (Pol II) is a highly regulated process. Regulation occurs both during early steps of elongation through Pol II pausing and after Pol II is released to enter a phase of productive elongation. During the initial steps of elongation, Pol II can pause and accumulate at very high levels in the promoter-proximal region, 30–60 nucleotides downstream of the transcription start site (TSS; reviewed in REFS 1,2) (FIG. 1). This is a key rate-limiting step for transcription that is potentially subject to regulatory control, and can act as a quality checkpoint for transcript 5'-capping and Pol II modification before productive elongation^{1,2}. Genome-wide studies have indicated that Pol II pausing is a common regulatory step in the transcription of developmental genes and of genes involved in stimulus-controlled pathways (such as heat shock protein 70 (*Hsp70*), discussed below)^{1,2}.

Pol II pausing at promoter-proximal regions depends on the core promoter features that recruit Pol II to this region^{3,4} (FIG. 1a). It involves promoter-associated transcription factors (TFs) that function with negative elongation factor (NELF) and DRB-sensitivity-inducing

Correspondence to J.T.L.johnlis@cornell.edu.

Competing interests statement

The authors declare no competing interests.

factor (DSIF)^{1,2} to stabilize the paused Pol II. In some cases, nucleosomes may also contribute to pausing^{3–5} (FIG. 1a).

Release of paused Pol II is mediated by the positive transcription elongation factor-b (P-TEFb) complex, comprising cyclin T1 and cyclin-dependent kinase 9 (CDK9)^{6,7}. P-TEFb is recruited to promoters through direct or indirect interactions with specific TFs and cofactors (FIG. 2), and phosphorylates the carboxy-terminal domain (CTD) of Pol II at Ser2, as well as NELF (which is evicted from Pol II upon phosphorylation) and DSIF (which then becomes a positive elongation factor)^{1,2,6} (FIG. 1a). The direct recruitment or artificial tethering of P-TEFb to the *Hsp70* promoter, where Pol II is normally stably paused and highly accumulated at the promoter-proximal region, leads to high levels of gene body transcription in the absence of a heat shock stimulus⁸. These and many other studies indicate that P-TEFb is the key regulator of early elongation steps^{6–8}. P-TEFb can be found in two states: as part of an inhibitory complex, or as an active complex that phosphorylates pausing factors and the Pol II CTD⁷. Thus, the emerging model is that the level of pausing depends on the balance between pausing factors (such as NELF, DSIF, the +1 nucleosome and the core promoter elements) and activating factors (discussed below) that either recruit P-TEFb to paused Pol II, or activate P-TEFb.

After Pol II is released from the promoter-proximal pause site, it commences productive elongation. Interestingly, elongation following Pol II release is more complex than initially thought. Elongation rates can vary between and within genes^{9–14}, and seem to play a part in co-transcriptional processes such as splicing and transcription termination, as well as in the maintenance of genome stability^{11,14–22}. Multiple factors can modulate elongation rates, including histone marks and features of genes such as the number of exons^{9,11–13}. The development of methods that allow genome-wide measurements of elongation rates (BOX 1; TABLE 1) now enables the study of the role and regulation of elongation rates throughout the entire transcription cycle.

As transcription initiation²³ and termination²⁴ by Pol II, and the effects of histone turnover on these processes²⁵, have been the subject of other reviews in this Focus issue, here we discuss the dynamics of elongation by Pol II. We evaluate the many factors that modulate and regulate the early phases of transcription elongation by Pol II and the later phase of productive elongation throughout a gene. We focus on the molecular mechanisms that underlie promoter-proximal pausing of Pol II, and discuss the co-transcriptional processes that can be influenced by variations in elongation rates throughout a gene. Overall, we review the evidence indicating that transcription elongation by Pol II is not a static process, but a highly dynamic and important part of the transcription cycle.

Pol II pausing is a common feature of genes

Over the years, it has been debated whether, following Pol II recruitment to promoters, transcription is rate-limited at the level of initiation or of release from promoter-proximal regions into productive elongation^{2,26}. Overall, once recruited to a promoter, Pol II initiation and gene entry seem to be highly efficient, as little Pol II is detected in a pre-initiation complex (PIC) for most genes (FIG. 1b). This was observed in a human myelogenous

leukaemia (K562) cell line by chromatin immunoprecipitation–exonuclease digestion (ChIP–exo), which maps the precise position of all DNA-associated Pol II²⁷. It has also been observed in *Drosophila melanogaster* S2 cells by quantitative comparisons of ChIP followed by sequencing (ChIP–seq) and genome-wide run-on sequencing (GRO–seq) signals (TABLE 1), which are assays that measure all DNA-associated Pol II or only the transcriptionally engaged Pol II²⁸, respectively. These assays indicated that the bulk of promoter-associated Pol II on most genes is positioned on the pause region.

P-TEFb acts at all active genes

A combination of genome-wide high-throughput sequencing methods (BOX 1; TABLE 1) and drug treatments that inhibit P-TEFb have suggested that P-TEFb-driven release of paused Pol II from promoter-proximal regions to begin productive elongation is a widespread and necessary step in transcription (FIG. 1a). Studies have shown that inhibition of P-TEFb, which prevents Pol II release, blocks almost all transcription^{9,29,30}. Thus, all active genes experience a potentially rate-limiting pausing step in the transcription cycle and require P-TEFb activity for gene body transcription. However, this pause step causes a significant accumulation of promoter-proximally paused Pol II only at a subset of active genes in untreated cells (40–70%, depending on the method and cell type)^{3,9,29,31–33}. Thus, presumably P-TEFb activity is simply not limiting on the remainder of genes that do not show accumulation of paused Pol II. This finding indicates that a pausing- and P-TEFb-dependent release step could become a rate-limiting and potentially regulatory step at all active genes (FIG. 1b).

Pausing helps to maintain active transcription

The fact that a large proportion of highly expressed genes show significant transcription pausing indicates that pausing is not a simple repressive step, but a crucial feature of active and open promoters. Indeed, when one of the main pausing factors, NELF, is knocked down, nucleosomes move towards and occupy the nucleosome-free regions at promoters at which TFs can bind efficiently, causing a decrease in the expression of these genes. Thus, pausing can function to maintain genes in an active state³⁴. Also, DNA mutations in pausing regions that disrupt pausing on *Hsp70* reduce the heat shock response³⁵ by making upstream promoter elements inaccessible to the heat shock factor (HSF) regulatory protein during heat shock³⁶. Thus, promoter-proximal pausing is a feature of active, primed and highly regulated genes and contributes to maintaining the promoter in an open state that is accessible to TFs.

Nonetheless, transcription regulation is complex, and examples in which initiation is rate-limiting have been reported (FIG. 1b). For example, a recent study in lymphocytes reported that DNA melting by TFIID was rate-limiting in resting cells when Pol II was associated with the PIC³⁷. Also, a study of human α 1 antitrypsin gene expression during enterocyte differentiation showed that Pol II accumulated at the PIC and that this step was rate-limiting until developmental regulatory signals activated gene expression³⁸. Therefore, although Pol II recruitment and pausing seem to be major rate-limiting and regulated steps in transcription, steps between recruitment and pausing can also be rate-limiting, at least in some cases.

Mechanisms that contribute to pausing

Although the importance and pervasiveness of promoter-proximal pausing in higher eukaryotes is well documented, how this rate-limiting step is regulated is still subject to debate²⁶. Promoter-proximal pausing of Pol II can depend on the rates of four reactions: recruitment of Pol II to the TSS; entry of recruited Pol II into the pause site; termination of this paused Pol II and release from the DNA; and release and progression of paused Pol II into productive elongation (FIG. 1b). Below, we discuss how, and to what extent, these individual reactions contribute to pausing and overall transcriptional output. We first describe how transcription termination (FIG. 1b) contributes to the pausing mechanisms, and then discuss how Pol II is released from the pause site to continue productive transcription elongation.

Is productive transcription regulated by termination?

Two transcription termination-based mechanisms that regulate Pol II pausing have been proposed^{39,40}: decapping of the RNA protruding from paused Pol II, and transcription termination by XRN2-mediated RNA degradation and termination factor TTF2 (REF. 39); or the combined action of XRN2 and the Microprocessor complex⁴⁰. Moreover, GDOWN1 (also known as GRINL1A) is thought to increase pausing by blocking the interaction of TFIIF with Pol II, but also by inhibiting termination of Pol II by TTF2 (REFS 41–43). However, genome-wide measurements of the stability of paused Pol II in *D. melanogaster* cells³⁰ and mouse embryonic stem cells⁹ indicate that Pol II has a stable average half-life of ~7 minutes, or almost 1 hour in a small subset of genes⁴⁴. This makes it difficult to simulate the observed dramatic changes in transcriptional output by varying termination within the observed range at the promoter-proximal pause site (FIG. 1b).

In the highly paused *Hsp70* gene, transcription termination by paused Pol II does not decrease when the expression of the gene is induced by heat shock. Instead, termination increases as the overall initiation and release from pausing, and entry into a productive elongation increases more than 100-fold as a result of heat shock, indicating that the amount of terminating Pol II appears not to be regulated but simply proportional to concentration of promoter-proximal Pol II on *Hsp70* (REF. 45). Although this is one example in which termination is relatively constant despite changes in transcription regulated at the level of paused Pol II, *Hsp70* is a good model for the regulation of pause set-up and release, as the half-life of paused Pol II at this gene is similar to the half-lives of mammalian and *D. melanogaster* paused Pol II^{9,30,45}.

It cannot be excluded that the regulation of paused Pol II occurs at the level of termination in a subset of genes. However, it has been observed that patterns of Pol II density on promoters and gene bodies change in response to heat shock, hormones, and developmental signals^{10,29,46–50}, indicating that the two prominent regulatory steps during early elongation are Pol II recruitment to the promoter and the release of paused Pol II to productive elongation (FIG. 1b).

Tethering Pol II to pause sites

The set-up of the pause site can depend on the core promoter elements, which recruit and tightly interact with general transcription factors (GTFs) (FIG. 1a). These GTFs subsequently attract Pol II and possibly form an ‘anchor’ (depending on the presence and positioning of core promoter elements) that prevents efficient release of Pol II from the promoter-proximal region^{1,3}. In this model, Pol II is recruited, rapidly initiates and transcribes for a short distance (FIG. 1; FIG. 2c), ‘scrunching up’ the DNA⁵¹ without breaking its contacts with the core promoter. This scrunching action was thought to have a role during early elongation and promoter escape, but before Pol II reaches the pause region⁵². However, a recent study suggested that in λ -bacteriophage there is scrunching following pausing at the promoter region⁵³. This extended pause is an attractive model for a class of *D. melanogaster* promoters that display strong, proximal and tightly clustered pausing³. Indeed, core promoter elements such as the downstream promoter element (DPE) and the pause button are associated with significantly paused genes in *D. melanogaster*⁵⁴, and capture Pol II efficiently to produce proximal-clustered pausing when these and other core elements, such as TATA, are properly positioned at promoters³.

Core promoter elements tend to be less well-defined in mammals, and elements that may facilitate pausing have not been identified. It is possible that mammalian genes may use DNA-binding TFs that interact with GTFs, thereby indirectly promoting the anchor strength⁵⁵ (FIG. 2c,d). Examples of such factors are described below. Alternatively, or in addition to, the +1 nucleosome could participate in pausing by creating an energy barrier for elongating Pol II (FIG. 1a; FIG. 2c,d). Indeed, paused Pol II seems to push on the +1 nucleosome, which is indicative of a transcriptional block.^{4,5} Moreover, the position of the active site of paused Pol II seems to depend on the position of the +1 nucleosome downstream of the TSS, whereas the level of pausing is influenced by histone H2A.Z, as enrichment of H2A.Z seems to decrease pausing⁵. Conversely, paused Pol II may be preventing nucleosomes from entering the nucleosome-free region of open promoters. Indeed, knockdown of NELF and consequent reduced Pol II pausing decreases gene expression, because nucleosomes occupy the promoter regions and thus block the expression of these genes by preventing TF and GTF binding^{28,34}.

Finally, a balance of pausing factors like NELF and DSIF, as well as factors that promote the release of paused Pol II by recruiting P-TEFb⁵⁶, may determine the level of pausing and productively elongating Pol II (FIG. 1; FIG. 2c–e). Paused Pol II is not likely to be fixed in one position but seems to undergo persistent rounds of transcription, pausing and backtracking, in which TFIIS cleaves the RNA of the backtracked and paused Pol II to realign the Pol II active site, ready to be released into productive elongation upon P-TEFb kinase activity^{5,33}. Although there is currently no strong evidence that the levels of NELF and DSIF on a promoter are regulated independently of the amount of paused Pol II, the recruitment of TFs and cofactors, and the activity of P-TEFb, can vary greatly upon activation of signalling pathways⁵⁷. Moreover, the release of paused Pol II could be mediated by 3D DNA looping, which brings together promoters and enhancers⁵⁸ of paused genes; enhancers bind factors such as Mediator⁵⁹, which can attract P-TEFb⁶⁰ and thus promote paused Pol II release (FIG. 2e). Indeed, looping interactions between enhancers and

promoters in the developing *Drosophila* embryo are associated with promoters showing high levels of paused Pol II⁶¹.

Overall, it seems that many regulated TFs and cofactors can either increase pausing or stimulate paused Pol II release (FIG. 2), indicating that paused and productively transcribing Pol II levels are determined by the balance of factors regulating promoter entry and pausing, and of factors that positively influence paused Pol II release.

Factors involved in setting up pausing

In addition to the NELF and DSIF complexes, which colocalize with and stabilize promoter-paused Pol II on most or all genes, an increasing number of TFs have been shown to have a role in pausing Pol II in a gene- and sequence-specific way. In *D. melanogaster*, GAGA factor (GAF) strongly promotes pausing at many genes, including *Hsp70*, potentially by remodelling nucleosomal architecture at promoters and allowing Pol II access to promoters^{62,63}, or by interacting with the GTF TFIID^{4,64} (FIG. 2a,b) or recruiting NELF⁵⁶. Another pausing factor in *D. melanogaster*, M1BP, preferentially binds near the TSSs of constitutively expressed metabolic genes that contain moderate levels of paused Pol II. M1BP is often the only known TF associated with these genes, and it frequently binds promoters that do not bind GAF, indicating that these factors have a similar role at different genes⁴.

In mammals, SP1, myoblast determination protein 1 (MYOD1), CCAAT box-binding transcription factor (CTF) and adenovirus E1A are considered to be DNA sequence-specific TFs that predominantly work by recruiting Pol II to the promoter region without stimulating the release of paused Pol II, thereby increasing the levels of paused Pol II^{65,66} (FIG. 2a,b). The activation domains of other TFs, including viral protein 16 (VP16), CCAAT/enhancer-binding protein (C/EBP) and p53, have dual roles in setting up and releasing paused Pol II^{65,66}, and thereby in strongly activating productive transcription. Intriguingly, a mechanism that could explain such a dual role was recently proposed for tripartite motif-containing protein 28 (TRIM28): unphosphorylated TRIM28 stimulated pausing, and pause release was facilitated by its phosphorylation⁶⁷.

Other factors that do not bind specific DNA elements can also influence pausing of Pol II and its release. GDOWN1, which tightly interacts with Pol II, may increase pausing by interfering with the binding of the positive elongation factor TFIIF^{41–43}. Inhibiting P-TEFb activity can also increase pausing by decreasing P-TEFb-mediated Pol II release. Polo-like kinase 1 (PLK1), a cell cycle regulatory factor that is most active during mitosis, was recently found to phosphorylate cyclin T1, thereby inhibiting CDK9 and potentially linking promoter-proximal pausing to the cell cycle⁶⁸. Thus, both sequence-specific DNA-binding factors and other cofactors can influence the formation of paused Pol II.

Factors involved in the release of paused Pol II

The most important factor in the release of paused Pol II is P-TEFb, which functions as part of an inhibitory complex together with McCPE, LARP7, HEXIM1 or HEXIM2 and the small nuclear ribonucleoprotein (snRNP) 7SK, a non-coding RNA (ncRNA)⁷, or functions

as an active complex that interacts with TFs and cofactors and phosphorylates pausing factors and the CTD of Pol II⁷. Thus, regulation of paused Pol II release primarily takes place through the recruitment and activation of this kinase complex.

The primary factors that form the activating complex with P-TEFb are bromodomain-containing protein 4 (BRD4) and the larger super elongation complex (SEC)⁶⁹ (FIG. 2d,e). BRD4, which competes for binding to P-TEFb with the P-TEFb inhibitory complex HEXIM–7SK, is recruited to TSSs via histone acetylation and in specific cases by interacting with acetylated NF- κ B^{70–73}. The SEC — a large multisubunit complex that interacts directly with P-TEFb — comprises elongation factors such as eleven-nineteen Lys-rich leukaemia (ELL), eleven-nineteen leukaemia (ENL) and AF4/FMR2 family member 4 (AFF4)^{69,74–77}, and can vary in composition depending on cellular context and gene target specificity⁷⁸. The SEC can interact with a subset of co-activators such as Mediator, polymerase-associated factor 1 (PAF1) and Integrator, the latter of which is a complex that interacts with the CTD of Pol II^{60,79–82}. Moreover, the SEC can also colocalize with BRD4 at gene promoters⁸³. The SEC and BRD4 predominantly mediate the recruitment P-TEFb, but their regulatory importance and composition with regard to paused Pol II release seems to vary across different genes, cell types and stimuli. Ultimately, recruitment of these cofactors depends on the TFs that bind to promoters or enhancers (FIG. 2d,e).

Well-known DNA-binding TFs that interact with P-TEFb to release paused Pol II include MYC, a TF associated with cancer²⁹, and NF- κ B^{10,84–86}. Indeed, MYC not only mediates pause release²⁹, but the *MYC* gene is itself also regulated by release of Pol II into productive elongation. Therefore, drugs that indirectly inhibit MYC-mediated pause release (such as JQ1, which inhibits BRD4 binding to acetylated histones) are attractive therapeutic interventions for cancer^{87–89}. NF- κ B — the final TF of a pathway that is activated following, for example, stimulation of immune cells and stress response⁹⁰ — interacts with P-TEFb either directly after phosphorylation of Ser276 of the NF- κ B subunit RELA^{85,91}, or indirectly through BRD4 after Lys310 acetylation of RELA^{72,73} (FIG. 2d). Moreover, protein phosphatase 1G (PPM1G) may remove a phosphate group from the inactive form of P-TEFb to facilitate its release and activation from the inhibitory HEXIM–7SK complex after induction of the NF- κ B pathway through an interaction with RELA⁹². Finally, genes encoding proteins at important regulatory nodes within the NF- κ B pathway are often controlled by a rate-limiting pause step⁹³, indicating that both the direct response to NF- κ B activation and the modulation and fine-tuning of the components of the NF- κ B pathway are regulated at the level of promoter-proximal pausing.

In addition to canonical TFs and elongation factors, the splicing factor SRSF2 has recently been shown to mediate the release of paused Pol II. SRSF2 interacts with the inactive P-TEFb complex via HEXIM–7SK and is recruited to genes that encode an exon-splicing enhancer (ESE). SRSF2 binds to the RNA protruding from paused Pol II at these loci and activates P-TEFb, thereby stimulating the release of paused Pol II^{94,95}. Interestingly, this links the recruitment of co-transcriptional splicing machinery to the regulation of early elongation steps in the transcription cycle, supporting the hypothesis that pausing is not only a regulatory point for controlling transcription levels, but that this regulation may be coupled

to a quality checkpoint to ensure the proper assembly of the transcription and co-transcriptional processing machineries^{2,95}.

Box 1

Methods to measure elongation rates

Transcription elongation has long been understudied in the transcriptional research field, partly because elongation rates are difficult to measure. Measuring elongation rate requires the quantification of the distance travelled by nascent transcribing polymerases as a function of elapsed time (TABLE 1). Nascent transcription can be measured on a gene-by-gene basis by monitoring the production of intronic RNA¹²⁷, or by assessing the levels and position of chromatin-bound RNA¹²⁸. Methods such as genome-wide run-on sequencing (GRO-seq) provide snapshots of the position and quantity of engaged RNA polymerase II (Pol II) genome-wide by quantification of short run-on RNA that can be isolated with biotin- or bromo-tagged nucleotides that are incorporated during run-on^{3,9,12,13,31}. GRO-seq combines nuclear run-ons, in which engaged Pol II incorporates bromo-tagged nucleotides after removal of impediments, and is followed by RNA isolation and next-generation sequencing. GRO-seq can be converted to precision run-on sequencing (PRO-seq) with single nucleotide resolution by using biotin-tagged nucleotides.

Furthermore, Pol II can be tagged with fluorescent proteins, and the rate of fluorescence recovery after photobleaching (FRAP) at a single locus or at multiple loci provides information on elongation rates^{45,106,109,129}. Chromatin immunoprecipitation followed by sequencing (ChIP-seq) of Pol II is also indicative of nascent transcription, as most chromatin-bound Pol II is transcriptionally engaged²⁸. However, ChIP-seq of Pol II is often associated with high background levels within gene bodies, and its use is restricted to highly expressed genes. The monitoring of elapsed time necessary to measure elongation rates is achieved by inhibiting Pol II release from pause (with flavopiridol or DRB)⁹ (FIG. 3b), or by inducing transcription in a highly synchronous and timed manner (with, for example, tumour necrosis factor or oestradiol, or wash-out of inhibitory drugs)^{10,12–14} (FIG. 3c). The distance Pol II travels after inhibition or induction in a defined time period is a direct measure of elongation rates at either single genes, or throughout the genome. Because the resolution of the ‘wave’ of retreating or emerging Pol II (FIG. 3b,c) is a composite profile of a large population of molecules, the leading edges may be smoothed, which limits the use of these methods to relatively long genes. Regardless, the timed inhibition or induction of transcription followed by the detection of nascent transcription is a vast improvement over methods that were previously used to measure elongation rates.

3D chromatin looping modulates pause release

Although regulation of paused Pol II release occurs predominantly at the promoter-proximal region, it is becoming evident that 3D promoter–enhancer interactions also play an important part in promoter-proximal pause release. Indeed, promoter–enhancer looping, which seems to be relatively stable throughout development, has been associated with

paused Pol II in *D. melanogaster*⁶¹. Moreover, NF- κ B is inducibly associated with ‘super-enhancers’ (REFS 96,97), immediately followed by BRD4 recruitment to the enhancer and BRD4-dependent release of paused Pol II at target genes⁹⁸ (FIG. 2d,e).

Enhancers and super-enhancers are thought to regulate transcription by interacting with genes through a looping mechanism that seems to depend on Mediator^{96,97}. Interestingly, Mediator interacts with the SEC, providing a potential mechanism by which enhancers can facilitate the release of paused Pol II at genes⁶⁰ (FIG. 2e). Moreover, ncRNAs can activate transcription by inducing Mediator-dependent chromatin looping, indicating that ncRNAs other than the 7SK snRNP may have roles in paused Pol II release⁹⁹ (FIG. 2e). This is further supported by the finding that some enhancer RNAs (eRNAs), which are a mark of active enhancers⁵⁵, form a decoy for the RNA recognition motif of NELF. NELF is thought to increase pausing by interacting with RNAs protruding from paused Pol II¹⁰⁰, and eRNAs seem to facilitate NELF eviction from the paused Pol II complex¹⁰¹. However, it is unclear how this novel concept should be reconciled with the known dominant role of P-TEFb kinase activity in pause release to productive elongation at virtually all genes^{9,29}.

Finally, looping at paused genes could be mediated by cohesin, a protein complex that generates chromatin loops usually associated with enhancers¹⁰². Cohesin interacts with genes that are paused and active, and its knockdown reduces the release of paused Pol II into the gene body¹⁰², presumably by interfering with looping between enhancers and gene promoters. Overall, the regulation of the early steps of elongation seems to be increasingly sophisticated and involves the interplay of DNA-binding TFs, chromatin looping, large multisubunit cofactor complexes, post-translational modifications and possibly ncRNAs to mediate the release of transcriptionally competent Pol II into the gene body (FIG. 2).

Moving on to productive elongation

After Pol II is released from the promoter-proximal region, productive elongation commences. This process is variable, as elongation rates differ between genes by as much as threefold^{9,10,12–14}. This variation can have considerable effects on mRNA accumulation, which could affect the timing of developmental regulatory processes¹⁰³.

Moreover, productive elongation is not very efficient within the first kilobase of mammalian genes (FIG. 3a), as indicated by the increase in Pol II elongation rates from ~0.5 kb per minute within the first few kilobases, to 2–5 kb per minute after ~15 kb^{9,10,12–14}. This may be due to the gradual accumulation and modification of components of the transcription machinery. For instance, phosphorylation of Ser2 of the Pol II CTD does not reach a maximum until a few kilobases into the gene body¹⁰⁴. Furthermore, pausing factors, such as NELF, DSIF and GDOWN1 (REFS 1,41–43), may be gradually transformed or removed from the transcriptional machinery. This gradual ‘maturation’ of the transcription machinery could facilitate the recruitment of factors that are important for co-transcriptional events such as splicing.

It has become clear that the transition from paused to productively elongating Pol II is not an ON–OFF switch. Even after Pol II has transcribed an extensive part of the gene body, further impediments may still affect transcription elongation rates. Slowdown of Pol II can

be caused by the presence of exons^{3,9,11–13}, and by mRNA cleavage and polyadenylation sites^{3,20–22,31,32,105} (FIG. 3a), probably through distinct mechanisms, and may be necessary to complete co-transcriptional mRNA splicing and 3'-end processing. Indeed, co-transcriptional splicing has been estimated to take 20–30 seconds *in vivo*, which is consistent with the delay in Pol II elongation caused by an exon^{9,106}. Thus, transcription through the gene body is subject to dynamic changes in elongation rate from start to finish (FIG. 3a).

Variations in transcription elongation rate could be caused by different, not necessarily mutually exclusive, factors: histone marks can tighten or loosen DNA binding around nucleosomes, restricting or facilitating Pol II transcription efficiency, respectively; elongation factors, histone chaperones and nucleosome remodellers can facilitate Pol II movement through chromatin by alleviating pausing and stalling during elongation; and some DNA sequences may be more difficult to transcribe than others owing to their DNA topology (for example, G-rich DNA that could result in R-loops).

Some of the histone marks associated with the gene body are ubiquitylation of histone H2B (mediated by the E3 ligase ring finger 20 (RNF20)), acetylation of H3 at Lys56 (H3K56; mediated by histone acetyltransferase RTT109), trimethylation of H3 at Lys36 (H3K36me3; mediated by the methyltransferase SET2) and dimethylation of H3 at Lys79 (H3K79me2; mediated by the methyltransferase DOT1).^{25,107} These histone marks some removal and restoration in the wake of elongating Pol II^{1,25,107}. Examples of such histone chaperones are FACT, SPT6, ASF1, nucleosome assembly protein 1 (NAP1) and chromatin-remodelling complexes such as chromatin-helicase DNA-binding protein 1 (CHD1)^{1,25}. SPT6 depletion decreases elongation rates at *Hsp70* (REF. 109), but whether other histone chaperones directly affect elongation rates remains unclear.

Recent genome-wide elongation rate studies (BOX 1) have uncovered that H3K79me2, H2B ubiquitylation and H4K20me1 positively correlate with elongation rates^{9,12,13}. Interestingly, RNF20 was recently shown to facilitate stabilization of the early elongation complex, through interaction with the H4K16 histone acetyltransferase MOF, P-TEFb and the elongation factor PAF1 in *D. melanogaster*¹¹⁰. This is in line with the preferential H2B ubiquitylation of the intron 1 region of genes in mammals¹¹¹, which is also the region where Pol II elongation is still slow and possibly requires the presence of elongation factors and histone marks that could increase elongation efficiency^{9,10,12–14}. Many of these important elongation and chromatin maintenance factors have been discussed in a recent review²⁵.

Interestingly, the helicase RECQL5 modulates elongation rates seemingly without altering nucleosomes. RECQL5 directly interacts with Pol II and seems to inhibit elongation, as knockdown of this helicase increases elongation rates. However, RECQL5 knockdown also increases deletions in the genome and rearrangements, suggesting that well-controlled elongation is important for genome stability.¹⁴ Other factors that affect elongation are elongin and TFIIS, which reduce or resolve co-transcriptional stalling¹¹².

Elongation through exons

Low-complexity DNA sequences and relatively low CG content are associated with higher elongation rates^{9,10,12,13}. By contrast, exons have the strongest negative effect on elongation

rates, delaying transcription through a gene by 20–30 seconds per exon^{9,106}. It is difficult to tease apart what causes the delay at exons, as exons have numerous characteristics that may affect elongation rates. For example, CG content is increased at exons¹¹³, as is nucleosome occupancy^{114,115}. Interestingly, AT-rich regions surrounding exons also seem to negatively affect elongation rates, and can increase subsequent splicing. This can be countered by the DBIRD complex (formed by DBC1 (deleted in breast cancer 1) and ZNF326 (zinc-finger protein 326; also known as ZIRD)), which interacts directly with Pol II and facilitates elongation through AT-rich exons, thereby increasing exon skipping¹¹⁶.

Moreover, exons are enriched for specific histone marks that are associated with reduced elongation rates (H3K36me3 and H4K20me1), whereas other marks are mostly enriched within the first intron (H2B ubiquitylation and H3K79me2) and associated with higher elongation rates^{111,117}. Finally, increased H3K9me3 and its associated phosphorylated heterochromatin protein 1 γ (HP1 γ) at exons can promote the inclusion of alternatively spliced exons by decreasing Pol II elongation rates¹¹⁸. Overall, the observed slowdown of Pol II at exons could increase splice site recognition and facilitate co-transcriptional splicing. Indeed, slowing down Pol II (either through mutation of the active site, or chemically) decreases exon skipping in humans and yeast^{11,17,119,120}. Alternatively, splicing itself may influence the local elongation rate and account for a slowdown of Pol II that occurs precisely at the 3' splice sites of *D. melanogaster* introns³, possibly via a mechanism reminiscent of that in *Escherichia coli*, where nascent RNA duplex formation near the RNA exit channel directly causes RNA polymerase to pause transiently¹²¹.

Although Pol II slowdown near exons seems to promote splicing in many cases, it was recently reported that slow elongation rates can increase exon skipping, by recruiting splice repressor ETR3 (also known as CELF2)¹⁵. Thus, elongation rates have a more complex role in splicing than originally thought and probably affect the interplay of competing and cooperative interactions of secondary RNA structures and RNA-binding factors.

Elongation during the termination process

Pol II pauses after transcription through the poly(A) site and before termination (as further discussed in REF. 24), as evidenced by increased Pol II density around the poly(A) site^{3,20–22,31,32,105}. The pausing of Pol II is thought to increase transcription termination efficiency by allowing the RNA exonuclease XRN1 (known as Rat1 in yeast) — which degrades the unprotected 5' end of the RNA after poly(A) cleavage — to 'overtake' and displace Pol II²². Recently, a model to describe termination-associated pausing was proposed, in which transcription-dependent R-loops attract the Dicer–Argonaute RNA interference machinery. This machinery promotes the deposition of H3K9me2 over termination pause sites, increasing pausing and therefore termination^{21,122}. Thus, H3K9 methylation seems to be a histone mark that can interfere with efficient elongation at both exons and termination sites.

Conclusion and perspectives

The regulation of promoter pausing and Pol II speed during productive elongation have a key role in regulating the level and timing of RNA production. The challenge is to

understand how, in molecular terms, the numerous potential players discussed in this Review execute both the regulation and mechanics of elongation. The emergence of high-resolution methods that measure both the density of elongating Pol II and elongation rates across the genome (BOX 1; FIG. 3) provides unprecedented opportunities to monitor these processes, both in normal conditions and under conditions in which candidate factors or features of the genome are perturbed, as exemplified by a recent study¹⁴. The power of studying elongation rates of many genes simultaneously with statistical significance and the ability to correlate these with genome-wide data on histone marks and TF location has already provided insights^{9,10,12–14}. However, a full understanding of elongation and its regulation requires systematically disrupting elongation factors, histone marks and features of the genes that have been implicated in these processes and assessing the consequences. The use of highly specific drugs to rapidly disrupt the function of specific factors, together with high-resolution genome-wide methods (BOX 1; FIG. 3b,c), should allow a mechanistic assessment of the immediate (first-order) role of factors *in vivo*. These drugs could be either small molecules, which have been used successfully to inhibit catalytic activities such as that of P-TEFb^{9,14,29,30}, or larger RNA and protein aptamers¹²³ that disrupt macromolecular protein–protein, protein–DNA or protein–RNA interactions that are crucial for transcription regulation.

Finally, as more and more factors are implicated in early and productive transcription elongation, the connection to human disease is increasingly being recognized (reviewed in REF. 124). Already, BRD4 has been identified as a potential therapeutic target in cancer¹²⁵, and the SEC is known to form MLL–ELL fusions after translocations in leukaemia^{69,77}. Inhibition of BRD4 with small drugs, such as JQ1, has shown promising results in cancer, by targeting *MYC* expression^{87,89}, or in cancer and inflammation, by disrupting BRD4 occupancy at super-enhancers and thereby blocking transcription^{88,98,126}. Promoter-proximal pausing and elongation mechanisms will almost certainly have a role in diseases beyond cancer, and these too need investigation.

Acknowledgements

The authors thank C. Danko, F. Duarte and D. Mahat for their critical evaluation of the manuscript. J.T.L. was supported by NIGMS (National Institute of General Medical Sciences) from the US National Institutes of Health under award GM25232. I.J. was supported by a European Research Council Advanced Grant (ERCadv-671274). The content is solely the responsibility of the authors and does not necessarily represent the official views of the US National Institutes of Health or the European Research Council.

Glossary

DRB	A small drug that inhibits P-TEFb kinase activity. It is used to characterize pausing and elongation complexes, and to measure elongation rates genome-wide
Carboxy-terminal domain (CTD) of Pol II	The CTD of Pol II, which is positioned at the end of the largest Pol II subunit, is an unstructured, yet evolutionarily conserved, domain that comprises many tandem copies of the consensus heptapeptide

	YSPTSPS. Phosphorylation of these repeats is crucial for the regulation of Pol II function
+1 nucleosome	The first well-positioned nucleosome downstream of the transcription start site, which can form a barrier for elongating Pol II and might increase Pol II promoter-proximal pausing. The position of the +1 nucleosome depends on transcription, nucleosome remodelling, and DNA sequences
Pre-initiation complex (PIC)	A complex consisting of general transcription factors and Pol II that binds at the transcription start site, before DNA melting and transcription initiation
Open promoters	Promoters that are nucleosome-free and easily accessible to transcription factors and Pol II. These promoters are primed for, or undergo, active transcription
DNA melting	The process of unwinding and ‘opening’ double-stranded DNA at the transcription start site by general transcription factors to form a transcription bubble, which allows initiation of Pol II activity
General transcription factors (GTFs)	Factors that bind the core promoter region, facilitate DNA melting and transcription bubble formation, and position Pol II to initiate transcription and escape the promoter region
Enhancers	Regulatory regions that bind sequence-specific TFs and have potential transcription start sites and can interact with gene promoters three dimensionally to regulate gene expression
Mediator	A multisubunit co-activator complex that can interact with TFs, GTFs and Pol II and is essential for transcription. Mediator has been shown to mediate interaction between enhancers and gene promoters, for example at super-enhancers
Enhancer RNAs (eRNAs)	RNAs that derive from the transcription of enhancers. Some of these enhancer-derived RNAs contribute to enhancer function
Histone chaperones	Proteins that facilitate the movement of Pol II through chromatin by loosening the nucleosome–DNA interactions and then restoring these in the wake of Pol II
R-loops	An RNA–DNA hybrid structure formed during the transcription of a sequence with high GC-content that has the potential to pause Pol II. R-loops are associated with transcription termination and genome instability
Exon skipping	A form of alternative splicing, in which an exon is ‘skipped’ and removed as part of the flanking introns during transcription
MLL–ELL fusions	A fusion formed between the <i>MLL</i> gene (which encodes mixed-lineage leukaemia) and the <i>ELL</i> gene (which encodes eleven-

nineteen Lys-rich leukaemia) that greatly increases the leukaemogenic potential of a cell

References

1. Kwak H, Lis JT. Control of transcriptional elongation. *Annu. Rev. Genet.* 2013; 47:483–508. [PubMed: 24050178]
2. Adelman K, Lis J. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nature Rev. Genet.* 2012; 13:720–731. [PubMed: 22986266]
3. Kwak H, Fuda NJ, Core LJ, Lis JT. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science.* 2013; 339:950–953. In this study, pausing was mapped at the genome-wide level with base-pair resolution, showing the dependency of strong promoter-proximal pausing on core promoter elements. [PubMed: 23430654]
4. Li J, Gilmour D. Distinct mechanisms of transcriptional pausing orchestrated by GAGA factor and M1BP, a novel transcription factor. *EMBO J.* 2013; 32:1829–1841. [PubMed: 23708796]
5. Weber CM, Ramachandran S, Henikoff S. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell.* 2014; 53:819–830. [PubMed: 24606920]
6. Peterlin B, Price D. Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell.* 2006; 23:297–305. [PubMed: 16885020]
7. Zhou Q, Li T, Price DH. RNA polymerase II elongation control. *Annu. Rev. Biochem.* 2012; 81:119–143. [PubMed: 22404626]
8. Lis J, Mason P, Peng J, Price D, Werner J. P-TEFb kinase recruitment and function at heat shock loci. *Genes Dev.* 2000; 14:792–803. [PubMed: 10766736]
9. Jonkers I, Kwak H, Lis JT. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife.* 2014; 3:e02407. This study measures elongation rates genome-wide and shows that the half-lives of paused Pol II complexes on 3,181 genes are uniformly long with an average of 7 minutes. [PubMed: 24843027]
10. Danko C, et al. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol. Cell.* 2013; 50:212–222. [PubMed: 23523369]
11. Alexander R, Innocente S, Barrass J, Beggs J. Splicing-dependent RNA polymerase pausing in yeast. *Mol. Cell.* 2010; 40:582–593. [PubMed: 21095588]
12. Veloso A, et al. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res.* 2014; 24:896–905. [PubMed: 24714810]
13. Fuchs G, et al. 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biol.* 2014; 15:R69. [PubMed: 24887486]
14. Saponaro M, et al. RECQL5 controls transcript elongation and suppresses genome instability associated with transcription stress. *Cell.* 2014; 157:1037–1049. This manuscript documents that RECQL5 slows down transcript elongation and suppresses genome rearrangements at common fragile sites. [PubMed: 24836610]
15. Dujardin G, et al. How slow RNA polymerase II elongation favors alternative exon skipping. *Mol. Cell.* 2014; 54:683–690. [PubMed: 24793692]
16. Schor I, Fiszbein A, Petrillo E, Kornblihtt A. Intragenic epigenetic changes modulate NCAM alternative splicing in neuronal differentiation. *EMBO J.* 2013; 32:2264–2274. [PubMed: 23892457]
17. de la Mata M, et al. A slow RNA polymerase II affects alternative splicing *in vivo*. *Mol. Cell.* 2003; 12:525–532. [PubMed: 14536091]
18. Moehle EA, Braberg H, Krogan NJ, Guthrie C. Adventures in time and space: splicing efficiency and RNA polymerase II elongation rate. *RNA Biol.* 2014; 11:313–319. [PubMed: 24717535]
19. Plant K, Dye M, Lafaille C, Proudfoot N. Strong polyadenylation and weak pausing combine to cause efficient termination of transcription in the human gamma-globin gene. *Mol. Cell. Biol.* 2005; 25:3276–3285. [PubMed: 15798211]

20. Gromak N, West S, Proudfoot N. Pause sites promote transcriptional termination of mammalian, RNA polymerase II. *Mol. Cell. Biol.* 2006; 26:3986–3996. [PubMed: 16648491]
21. Skourti-Stathaki K, Kamieniarz-Gdula K, Proudfoot NJ. R-loops induce repressive chromatin marks over mammalian gene terminators. *Nature.* 2014; 516:436–439. [PubMed: 25296254]
22. Hazelbaker D, Marquardt S, Wlotzka W, Buratowski S. Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Mol. Cell.* 2013; 49:55–66. [PubMed: 23177741]
23. Sainsbury S, Bernecky S, Cramer P. Structural basis of transcription initiation by RNA polymerase II. *Nature Rev. Mol. Cell. Biol.* 2015; 16:129–143. [PubMed: 25693126]
24. Porrua O, Libri D. Transcription termination and the control of the transcriptome: why, where and how to stop. *Nature Rev. Mol. Cell. Biol.* 2015; 16:190–202. [PubMed: 25650800]
25. Venkatesh SS, Workman JL. Histone exchange, chromatin structure and the regulation of transcription. *Nature Rev. Mol. Cell. Biol.* 2015; 16:178–189. [PubMed: 25650798]
26. Ehrensberger AH, Kelly GP, Svejstrup JQ. Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. *Cell.* 2013; 154:713–715. [PubMed: 23953103]
27. Venters B, Pugh B. Genomic organization of human transcription initiation complexes. *Nature.* 2013; 502:53–58. [PubMed: 24048476]
28. Core LJ, et al. Defining the status of RNA polymerase at promoters. *Cell Rep.* 2012; 2:1025–1035. [PubMed: 23062713]
29. Rahl P, et al. c-Myc regulates transcriptional pause release. *Cell.* 2010; 141:432–445. [PubMed: 20434984]
30. Henriques T, et al. Stable pausing by RNA polymerase II provides an opportunity to target and integrate regulatory signals. *Mol. Cell.* 2013; 52:517–528. This study reports that promoter-paused elongation complexes are highly stable, with half-lives of minutes in *D. melanogaster*. [PubMed: 24184211]
31. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science.* 2008; 322:1845–1848. [PubMed: 19056941]
32. Min IM, et al. Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.* 2011; 25:742–754. [PubMed: 21460038]
33. Nechaev S, et al. Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science.* 2010; 327:335–338. [PubMed: 20007866]
34. Gilchrist D, et al. Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell.* 2010; 143:540–551. [PubMed: 21074046]
35. Lee H, Kraus K, Wolfner M, Lis J. DNA sequence requirements for generating paused polymerase at the start of hsp70. *Genes Dev.* 1992; 6:284–295. [PubMed: 1737619]
36. Shopland L, Hirayoshi K, Fernandes M, Lis J. HSF access to heat shock elements *in vivo* depends critically on promoter architecture defined by GAGA factor, TFIID, and RNA polymerase II binding sites. *Genes Dev.* 1995; 9:2756–2769. [PubMed: 7590251]
37. Kouzine F, et al. Global regulation of promoter melting in naive lymphocytes. *Cell.* 2013; 153:988–999. This genome-wide analysis of resting lymphocytes identifies promoter melting as a third major rate-limiting step in transcription (following PIC formation and pause release). [PubMed: 23706737]
38. Soutoglou E, Talianidis I. Coordination of PIC assembly and chromatin remodeling during differentiation-induced gene activation. *Science.* 2002; 295:1901–1904. [PubMed: 11884757]
39. Brannan K, et al. mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol. Cell.* 2012; 46:311–324. [PubMed: 22483619]
40. Wagschal A, et al. Microprocessor, Setx, Xrn2, and Rrp6 co-operate to induce premature termination of transcription by RNAPII. *Cell.* 2012; 150:1147–1157. [PubMed: 22980978]
41. Cheng B, et al. Functional association of Gdown1 with RNA polymerase II poised on human genes. *Mol. Cell.* 2012; 45:38–50. [PubMed: 22244331]
42. Jishage M, et al. Transcriptional regulation by Pol II(G.) Involving mediator and competitive interactions of Gdown1 and TFIIF with pol II. *Mol. Cell.* 2012; 45:51–63. [PubMed: 22244332]

43. Davis M, Guo J, Price D, Luse D. Functional interactions of the RNA polymerase II-interacting proteins Gdown1 and TFIIF. *J. Biol. Chem.* 2014; 289:11143–11152. [PubMed: 24596085]
44. Chen F, Gao X, Shilatfard A. Stably paused genes revealed through inhibition of transcription initiation by the TFIIF inhibitor triptolide. *Genes Dev.* 2015; 29:39–47. [PubMed: 25561494]
45. Buckley MS, Kwak H, Zipfel WR, Lis JT. Kinetics of promoter Pol II on Hsp70 reveal stable pausing and key insights into its regulation. *Genes Dev.* 2014; 28:14–19. [PubMed: 24395245]
46. Saunders A, Core L, Sutcliffe C, Lis J, Ashe H. Extensive polymerase pausing during *Drosophila* axis patterning enables high-level and pliable transcription. *Genes Dev.* 2013; 27:1146–1158. [PubMed: 23699410]
47. Rougvie AE, Lis JT. The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell.* 1988; 54:795–804. [PubMed: 3136931]
48. Guenther M, Levine S, Boyer L, Jaenisch R, Young RA. Chromatin landmark and transcription initiation at most promoters in human cells. *Cell.* 2007; 130:77–88. [PubMed: 17632057]
49. Zeitlinger J, et al. RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nature Genet.* 2007; 39:1512–1516. [PubMed: 17994019]
50. Lagha M, et al. Paused pol II coordinates tissue morphogenesis in the *Drosophila* embryo. *Cell.* 2013; 153:976–987. [PubMed: 23706736]
51. Kapanidis AN, et al. Initial transcription by RNA polymerase proceeds through a DNA-scrunching mechanism. *Science.* 2006; 314:1144–1147. [PubMed: 17110578]
52. Pal M, Ponticelli A, Luse D. The role of the transcription bubble & TFIIB in promoter clearance by RNA polymerase II. *Mol. Cell.* 2005; 19:101–110. [PubMed: 15989968]
53. Strobel E, Roberts J. Regulation of promoter-proximal transcription elongation: enhanced DNA scrunching drives λ Q antiterminator-dependent escape from a σ 70-dependent pause. *Nucleic Acids Res.* 2014; 42:5097–5108. [PubMed: 24550164]
54. Hendrix DA, Hong J-WW, Zeitlinger J, Rokhsar DS, Levine MS. Promoter elements associated with RNA pol II stalling in the *Drosophila* embryo. *Proc. Natl Acad. Sci. USA.* 2008; 105:7762–7767. [PubMed: 18505835]
55. Core L, et al. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nature Genet.* 2014; 46:1311–1320. [PubMed: 25383968]
56. Li J, et al. Kinetic competition between elongation rate and binding of NELF controls promoter-proximal pausing. *Mol. Cell.* 2013; 50:711–722. [PubMed: 23746353]
57. Hargreaves D, Horng T, Medzhitov R. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell.* 2009; 138:129–145. [PubMed: 19596240]
58. Heinz S, Romanoski CE, Benner C, Glass CK. The selection and function of cell type-specific enhancers. *Nature Rev. Mol. Cell. Biol.* 2015; 16:144–154. [PubMed: 25650801]
59. Allen BL, Taatjes DJ. The Mediator complex: a central integrator of transcription. *Nature Rev. Mol. Cell. Biol.* 2015; 16:155–166. [PubMed: 25693131]
60. Takahashi H, et al. Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell.* 2011; 146:92–104. [PubMed: 21729782]
61. Ghavi-Helm Y, et al. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature.* 2014; 512:96–100. [PubMed: 25043061]
62. Lee C, et al. NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol. Cell. Biol.* 2008; 28:3290–3300. [PubMed: 18332113]
63. Farkas G, Leibovitch B, Elgin S. Chromatin organization and transcriptional control of gene expression in *Drosophila*. *Gene.* 2000; 253:117–136. [PubMed: 10940549]
64. Chopra VS, et al. Transcriptional activation by GAGA factor is through its direct interaction with dmTAF3. *Dev. Biol.* 2008; 317:660–670. [PubMed: 18367161]
65. Blau J, et al. Three functional classes of transcriptional activation domain. *Mol. Cell. Biol.* 1996; 16:2044–2055. [PubMed: 8628270]
66. Krumm A, Hickey LB, Groudine M. Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation. *Genes Dev.* 1995; 9:559–572. [PubMed: 7698646]

67. Bunch H, et al. TRIM28 regulates RNA polymerase II promoter-proximal pausing and pause release. *Nature Struct. Mol. Biol.* 2014; 21:876–883. [PubMed: 25173174]
68. Jiang L, et al. Polo-like kinase 1 inhibits the activity of positive transcription elongation factor of RNA Pol II b (P-TEFb). *PLoS ONE.* 2013; 8:e72289. [PubMed: 23977272]
69. Smith E, Lin C, Shilatifard A. The super elongation complex (SEC) and MLL in development and disease. *Genes Dev.* 2011; 25:661–672. [PubMed: 21460034]
70. Itzen F, Greifenberg A, Bosken C, Geyer M. Brd4 activates P-TEFb for RNA polymerase II CTD phosphorylation. *Nucleic Acids Res.* 2014; 42:7577–7590. [PubMed: 24860166]
71. Jang M, et al. The bromodomain protein Brd4 is a positive regulatory component of P-TEFb and stimulates RNA polymerase II-dependent transcription. *Mol. Cell.* 2005; 19:523–534. [PubMed: 16109376]
72. Zou Z, et al. Brd4 maintains constitutively active NF- κ B in cancer cells by binding to acetylated RelA. *Oncogene.* 2014; 33:2395–2404. [PubMed: 23686307]
73. Huang B, Yang X-DD, Zhou M-MM, Ozato K, Chen L-FF. Brd4 coactivates transcriptional activation of NF- κ B via specific binding to acetylated RelA. *Mol. Cell. Biol.* 2009; 29:1375–1387. [PubMed: 19103749]
74. Luo Z, et al. The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. *Mol. Cell. Biol.* 2012; 32:2608–2617. [PubMed: 22547686]
75. Lin C, et al. Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC). *Genes Dev.* 2011; 25:1486–1498. [PubMed: 21764852]
76. Smith E, et al. The little elongation complex regulates small nuclear RNA transcription. *Mol. Cell.* 2011; 44:954–965. [PubMed: 22195968]
77. Lin C, et al. AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell.* 2010; 37:429–437. [PubMed: 20159561]
78. Luo Z, Lin C, Shilatifard A. The super elongation complex (SEC) family in transcriptional control. *Nature Rev. Mol. Cell. Biol.* 2012; 13:543–547. [PubMed: 22895430]
79. Gardini A, et al. Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell.* 2014; 56:128–139. [PubMed: 25201415]
80. Kim J, Guermah M, Roeder RG. The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS. *Cell.* 2010; 140:491–503. [PubMed: 20178742]
81. Wier A, Mayekar M, Héroux A, Arndt K, VanDemark A. Structural basis for Spt5-mediated recruitment of the Paf1 complex to chromatin. *Proc. Natl Acad. Sci. USA.* 2013; 110:17290–17295. [PubMed: 24101474]
82. He N, et al. Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. *Proc. Natl Acad. Sci. USA.* 2011; 108:E636–645. [PubMed: 21873227]
83. Flajollet S, et al. The elongation complex components BRD4 and MLLT3/AF9 are transcriptional coactivators of nuclear retinoid receptors. *PLoS ONE.* 2013; 8:e64880. [PubMed: 23762261]
84. Diamant G, Dikstein R. Transcriptional control by NF- κ B: elongation in focus. *Biochim. Biophys. Acta.* 2013; 1829:937–945. [PubMed: 23624258]
85. Nowak D, et al. RelA Ser276 phosphorylation is required for activation of a subset of NF- κ B-dependent genes by recruiting cyclin-dependent kinase 9/cyclin T1 complexes. *Mol. Cell. Biol.* 2008; 28:3623–3638. [PubMed: 18362169]
86. Barboric M, Nissen RM, Kanazawa S, Jabrane-Ferrat N, Peterlin BM. NF- κ B binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell.* 2001; 8:327–337. [PubMed: 11545735]
87. Mertz J, et al. Targeting MYC dependence in cancer by inhibiting BET bromodomains. *Proc. Natl Acad. Sci. USA.* 2011; 108:16669–16674. [PubMed: 21949397]
88. Lovén J, et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell.* 2013; 153:320–334. This study shows that exceptionally high levels of the co-activators Mediator

and BRD4 are associated with super-enhancers that drive the expression of key oncogenes. [PubMed: 23582323]

89. Delmore J, et al. BET bromodomain inhibition as a therapeutic strategy to target c-Myc. *Cell*. 2011; 146:904–917. [PubMed: 21889194]
90. Oeckinghaus A, Hayden M, Ghosh S. Crosstalk in NF- κ B signaling pathways. *Nature Immunol*. 2011; 12:695–708. [PubMed: 21772278]
91. Fang L, et al. ATM regulates NF- κ B-dependent immediate-early genes via RelA Ser 276 phosphorylation coupled to CDK9 promoter recruitment. *Nucleic Acids Res*. 2014; 42:8416–8432. [PubMed: 24957606]
92. McNamara RP, McCann JL, Gudipaty SA, D’Orso I. Transcription factors mediate the enzymatic disassembly of promoter-bound 7SK snRNP to locally recruit P-TEFb for transcription elongation. *Cell Rep*. 2013; 5:1256–1268. [PubMed: 24316072]
93. Gilchrist D, et al. Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks. *Genes Dev*. 2012; 26:933–944. [PubMed: 22549956]
94. Ji X, et al. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell*. 2013; 153:855–868. This study implicates an RNA-binding protein that was traditionally thought to function in splicing in the regulated release of paused Pol II to productive elongation. [PubMed: 23663783]
95. Barboric M, et al. 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc. Natl Acad. Sci. USA*. 2009; 106:7798–7803. [PubMed: 19416841]
96. Hnisz D, et al. Super-enhancers in the control of cell identity and disease. *Cell*. 2013; 155:934–947. [PubMed: 24119843]
97. Whyte W, et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell*. 2013; 153:307–319. [PubMed: 23582322]
98. Brown JD, et al. NF- κ B directs dynamic super enhancer formation in inflammation and atherogenesis. *Mol. Cell*. 2014; 56:219–231. [PubMed: 25263595]
99. Lai F, et al. Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature*. 2013; 494:497–501. [PubMed: 23417068]
100. Yamaguchi Y, Inukai N, Narita T, Wada T, Handa H. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase, II complex and RNA. *Mol. Cell. Biol*. 2002; 22:2918–2927. [PubMed: 11940650]
101. Schaukowitz K, et al. Enhancer RNA facilitates NELF release from immediate early genes. *Mol. Cell*. 2014; 56:29–42. [PubMed: 25263592]
102. Schaaf CA, et al. Genome-wide control of RNA polymerase II activity by cohesin. *PLoS Genet*. 2013; 9:e1003382. [PubMed: 23555293]
103. Thummel CS, Burtis KC, Hogness DS. Spatial and temporal patterns of E74 transcription during *Drosophila* development. *Cell*. 1990; 61:101–111. [PubMed: 1690603]
104. Heidemann M, Hintermair C, Voß K, Eick D. Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. *Biochim. Biophys. Acta*. 2013; 1829:55–62. [PubMed: 22982363]
105. Glover-Cutter K, Kim S, Espinosa J, Bentley D. RNA polymerase II pauses and associates with pre-mRNA processing factors at both ends of genes. *Nature Struct. Mol. Biol*. 2007; 15:71–78. [PubMed: 18157150]
106. Martin R, Rino J, Carvalho C, Kirchhausen T, Carmo-Fonseca M. Live-cell visualization of pre-mRNA splicing with single-molecule sensitivity. *Cell Rep*. 2013; 4:1144–1155. [PubMed: 24035393]
107. Zentner G, Henikoff S. Regulation of nucleosome dynamics by histone modifications. *Nature Struct. Mol. Biol*. 2013; 20:259–266. [PubMed: 23463310]
108. Bintu L, et al. Nucleosomal elements that control the topography of the barrier to transcription. *Cell*. 2012; 151:738–749. In this study, optical tweezers were used to measure the movement of individual transcribing Pol II complexes through nucleosomes in real-time and thereby describes the energetic barriers in nucleosomes that could contribute to pausing. [PubMed: 23141536]

109. Ardehali MB, et al. Spt6 enhances the elongation rate of RNA polymerase II *in vivo*. EMBO J. 2009; 28:1067–1077. [PubMed: 19279664]
110. Wu L, Li L, Zhou B, Qin Z, Dou Y. H2B ubiquitylation promotes RNA pol II processivity via PAF1 and pTEFb. Mol. Cell. 2014; 54:920–931. [PubMed: 24837678]
111. Jung I, et al. H2B monoubiquitylation is a 5'-enriched active transcription mark and correlates with exon-intron structure in human cells. Genome Res. 2012; 22:1026–1035. [PubMed: 22421545]
112. Shilatifard A, Conaway RC, Conaway JW. The RNA polymerase II elongation complex. Annu. Rev. Biochem. 2003; 72:693–715. [PubMed: 12676794]
113. Amit M, et al. Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. Cell Rep. 2012; 1:543–556. [PubMed: 22832277]
114. Tilgner H, et al. Nucleosome positioning as a determinant of exon recognition. Nature Struct. Mol. Biol. 2009; 16:996–1001. [PubMed: 19684599]
115. Schwartz S, Meshorer E, Ast G. Chromatin organization marks exon-intron structure. Nature Struct. Mol. Biol. 2009; 16:990–995. [PubMed: 19684600]
116. Close P, et al. DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. Nature. 2012; 484:386–389. [PubMed: 22446626]
117. Huff J, Plocik A, Guthrie C, Yamamoto K. Reciprocal intronic and exonic histone modification regions in humans. Nature Struct. Mol. Biol. 2010; 17:1495–1499. [PubMed: 21057525]
118. Saint-André V, Batsché E, Rachez C, Muchardt C. Histone H3 lysine 9 trimethylation and HP1 γ favor inclusion of alternative exons. Nature Struct. Mol. Biol. 2011; 18:337–344. [PubMed: 21358630]
119. Schor IE, Rascovan N, Pelisch F, Alló M, Kornblihtt AR. Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. Proc. Natl Acad. Sci. USA. 2009; 106:4325–4330. [PubMed: 19251664]
120. Ip J, et al. Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. Genome Res. 2011; 21:390–401. [PubMed: 21163941]
121. Hein PP, et al. RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. Nature Struct. Mol. Biol. 2014; 21:794–802. [PubMed: 25108353]
122. Skourti-Stathaki K, Proudfoot N, Gromak N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. Mol. Cell. 2011; 42:794–805. [PubMed: 21700224]
123. Ozer A, Pagano JM, Lis JT. New technologies provide quantum changes in the scale, speed, and success of SELEX methods and aptamer characterization. Mol. Ther. Nucleic Acids. 2014; 3:e183. [PubMed: 25093707]
124. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. Cell. 2013; 152:1237–1251. [PubMed: 23498934]
125. Zuber J, et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature. 2011; 478:524–528. [PubMed: 21814200]
126. Dawson M, et al. Inhibition of BET recruitment to chromatin as an effective treatment for MLL-fusion leukaemia. Nature. 2011; 478:529–533. [PubMed: 21964340]
127. Singh J, Padgett R. Rates of *in situ* transcription and splicing in large human genes. Nature Struct. Mol. Biol. 2009; 16:1128–1133. [PubMed: 19820712]
128. Churchman L, Weissman J. Nascent transcript sequencing visualizes transcription at nucleotide resolution. Nature. 2011; 469:368–373. [PubMed: 21248844]
129. Yao J, Munson K, Webb W, Lis J. Dynamics of heat shock factor association with native gene loci in living cells. Nature. 2006; 442:1050–1053. [PubMed: 16929308]
130. Boireau S, et al. The transcriptional cycle of HIV-1 in real-time and live cells. J. Cell Biol. 2007; 179:291–304. [PubMed: 17954611]
131. Brody Y, et al. The *in vivo* kinetics of RNA polymerase II elongation during co-transcriptional splicing. PLoS Biol. 2011; 9:e1000573. [PubMed: 21264352]
132. Darzacq X, et al. *In vivo* dynamics of RNA polymerase II transcription. Nature Struct. Mol. Biol. 2007; 14:796–806. [PubMed: 17676063]

133. Tennyson CN, Klamut HJ, Worton RG. The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nature Genet.* 1995; 9:184–190. [PubMed: 7719347]
134. Mason P, Struhl K. Distinction & relationship between elongation rate & processivity of RNA polymerase II *in vivo*. *Mol. Cell.* 2005; 17:831–840. [PubMed: 15780939]
135. Ameer A, et al. Total RNA sequencing reveals nascent transcription and widespread co-transcriptional splicing in the human brain. *Nature Struct. Mol. Biol.* 2011; 18:1435–1440. [PubMed: 22056773]

Figure 1. RNA polymerase II recruitment, initiation and gene entry, pausing and release

a | RNA polymerase II (Pol II) is associated with promoters at, and just downstream of, the transcription start site (TSS). The transcriptional state, position and composition of Pol II are variable and depend on factors that contribute to recruitment, initiation, pausing and release of Pol II. Recruitment of Pol II by general transcription factors (GTFs) results in the formation of a pre-initiation complex (PIC). After rapid Pol II initiation and entry into the pause site, Pol II pausing by negative elongation factor (NELF) and DRB-sensitivity-inducing factor (DSIF) occurs, facilitated by the core promoter elements and the +1 nucleosome. Positive transcription elongation factor-b (P-TEFb) mediates the release of paused Pol II by phosphorylating NELF, DSIF and the carboxy-terminal domain (CTD) of Pol II. DSIF becomes a positive elongation factor after phosphorylation. **b** | The transcription cycle is predominantly regulated near the TSS, at the steps of recruitment of Pol II to promoters, and release from the promoter-proximal pause site. These steps are most variable in terms of rate (as indicated by the dark blue shading of the boxes defining the steps). Other steps, such as transcription initiation and entry to the pause site, as well as transcription termination from the pause site, seem not to be as variable in rate and less subject to regulation (as indicated by the lighter blue shading of the boxes).

Figure 2. The cascade of events that precede productive elongation by RNA polymerase II

The regulatory steps in the transcription cycle are mediated by a large number of factors, cofactors and regulatory mechanisms that cooperatively mediate the recruitment of RNA polymerase II (Pol II) and release of paused Pol II at both enhancer regions and gene start sites. **a** | Inactive transcription start sites (TSSs) at enhancers and genes are often occupied by nucleosomes. **b** | Nucleosomes are pushed aside (indicated by arrows) in response to external signalling events and the recruitment of transcription factors (TFs) that work with nucleosome remodellers and modifiers. **c** | Upon binding of the general transcription machinery at the TSSs, Pol II is recruited, initiated and phosphorylated at Ser5, which promotes transcription initiation. Subsequently, Pol II advances to the pause site, where it is stabilized by pausing factors (shown in red), such as negative elongation factor (NELF) and DRB-sensitivity-inducing factor (DSIF). **d** | Further signalling events lead to the recruitment of additional TFs, such as nuclear factor- κ B (NF- κ B) or MYC, which can recruit positive transcription elongation factor-b (P-TEFb) directly (not shown), or indirectly via co-activators such as bromodomain-containing protein 4 (BRD4). **e** | Finally, the cumulative recruitment of TFs results in binding of co-activators like BRD4, attracts other co-activators and elongation factors (shown in different shades of blue) such as the super elongation complex (SEC) and Mediator, potentially causing enhancer–promoter interactions in cooperation with non-coding RNAs (ncRNAs). This results in the activation of P-TEFb,

which phosphorylates pausing factors and the carboxy-terminal domain (CTD) of Pol II, leading to Pol II pause release and progression into a phase of productive elongation. This results in mRNA synthesis at the gene, and production of enhancer RNAs (eRNAs) at the enhancer (not shown).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 3. RNA polymerase II elongation rate throughout a gene

a | The density of RNA polymerase II (Pol II) and the elongation rates it achieves are not constant, but can vary throughout the gene, as exemplified by a schematic representation of a hypothetical gene. The changes in density of engaged Pol II, and the relative Pol II elongation rates, are indicated on the y-axis and the green shading. Pol II is slowest at the promoter-proximal pause site, and speeds up after release from the pause site over a region of approximately 15kb. Further variations in elongation rates are measured around exons (blue bars), where Pol II slows down, and at the site of termination, where Pol II pauses to enable efficient termination. Factors that affect the rate of transcription by Pol II are listed. Slowdown around exons might be orchestrated by a combination of CG content, nucleosome occupancy, histone marks and splicing factors that are enriched at exons. Pausing of Pol II near the polyadenylation site at the 3' ends of genes could be caused by a combination of R-loops, histone marks and termination factors. **b** | Elongation rates can be measured by determining nascent transcription after a block of release of new Pol II into productive elongation in a timed manner; the distance travelled by the 'retreating wave' of Pol II that is already productively elongating within a period of time is indicative of the elongation rate. **c** | Elongation rates can also be measured by determining nascent transcription after wash-out of inhibitory drugs; the distance travelled by the 'emerging wave' of newly released productive Pol II within a period of time is indicative of the elongation rate.

Table 1

Methods to detect elongation rates on a gene by gene or genome-wide basis

Method	Principle	Advantages	Disadvantages	Refs
<i>In vivo</i> imaging by FRAP	Quantifies GFP-tagged Pol II at specific loci following photobleaching in real time	<ul style="list-style-type: none"> Performed <i>in vivo</i> Allows single cell measurements 	<ul style="list-style-type: none"> Limited to multiple aligned genes or gene arrays Uses ectopically (over)expressed protein Low resolution 	45,106, 109, 129–132
RT-qPCR	Intron detection after release from DRB-mediated elongation block	<ul style="list-style-type: none"> Can be performed at different locations within a gene Can measure co-transcriptional splicing 	<ul style="list-style-type: none"> Gene-by-gene basis only Can be performed only on long genes* 	127,133
ChIP-seq (genome-wide) and qPCR (gene-specific)	Immunoprecipitation of Pol II	<ul style="list-style-type: none"> Can be performed at different locations within a gene Potentially genome-wide 	Relatively low resolution and high background noise	11,134
GRO-seq (~50 bp resolution) and PRO-seq (single bp resolution)	Directly measures nascent RNA production (instead of steady-state RNA levels) at specific sites in the genome after induction via signalling pathways or elongation block with small inhibitory drugs like DRB	<ul style="list-style-type: none"> Can be performed at different locations within a gene Potentially genome-wide Highly sensitive 	Can be performed only on long genes*	9,10,14
BruDRB-seq and 4sUDRB-seq	Similar to GRO-seq, use a 'pulse' with bromo-labelled UTP or 4sU combined with a DRB wash-out 'chase'	<ul style="list-style-type: none"> Can be performed at different locations within a gene Potentially genome-wide Highly sensitive 	Can be performed only on long genes*	12,13
Total poly(A) RNA-seq	Uses the density gradient of intronic reads as relative measure of elongation rates	Applicable to many existing data sets	<ul style="list-style-type: none"> Poor sensitivity Measures only relative elongation rates High sequencing depth needed Long introns only 	9,135

4sU, 4-thiouridine; ChIP-seq, chromatin immunoprecipitation followed by sequencing; FRAP, fluorescence recovery after photobleaching; GRO-seq, genome-wide run-on sequencing; Pol II, RNA polymerase II; PRO-seq, precision run-on sequencing; RT-qPCR, reverse transcription quantitative PCR.

* This method offers insufficient resolution on short genes.