MODES OF TRANSCRIPTIONAL REGULATION

Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans

Karen Adelman[1] *and John T. Lis*[2]

Abstract | Recent years have witnessed a sea change in our understanding of transcription regulation: whereas traditional models focused solely on the events that brought RNA polymerase II (Pol II) to a gene promoter to initiate RNA synthesis, emerging evidence points to the pausing of Pol II during early elongation as a widespread regulatory mechanism in higher eukaryotes. Current data indicate that pausing is particularly enriched at genes in signal-responsive pathways. Here the evidence for pausing of Pol II from recent high-throughput studies will be discussed, as well as the potential interconnected functions of promoter-proximally paused Pol II.

Pre-initiation complex
(PIC). An entry form of Pol II in a complex with general transcription factors in which the polymerase is bound to the promoter DNA but has not yet initiated RNA synthesis.

[1]*Laboratory of Molecular Carcinogenesis, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709, USA.*
[2]*Cornell University, Molecular Biology and Genetics, Biotechnology Building, Ithaca, New York 14853, USA.*
e-mails: adelmank@niehs.nih.gov; johnlis@cornell.edu
doi:10.1038/nrg3293

Higher organisms have evolved sophisticated mechanisms for responding in an integrated and balanced manner to various developmental, environmental and nutritional cues by precisely modulating transcription output. In response to intra- and extracellular cues, organisms must execute complex programs that require exquisite regulation of the timing and level of gene expression. These different transcriptional-regulatory programmes are orchestrated by the concerted action of sequence-specific transcription factors that recruit the transcription machinery. The enzyme that transcribes mRNA from protein-encoding genes is RNA polymerase II (Pol II). With the help of a constellation of accessory factors, Pol II executes a series of distinct steps: it binds to promoters, initiates RNA synthesis and then pauses in early transcriptional elongation. The paused Pol II remains stably associated with the nascent RNA and is fully capable of resuming elongation; however, further signals are needed to elicit the transition to a productive elongation complex. After this maturation has occurred, the polymerase processively progresses through the gene, terminates and eventually reinitiates transcription. To understand how developmental and homeostatic transcriptional programmes operate requires that we know the transcription factors that are involved and their targets. But just as important is an understanding of the mechanisms by which the interplay between Pol II and regulatory factors leads to highly specific yet readily modulated transcription profiles.

Traditional models of eukaryotic gene regulation were largely based on studies in *Saccharomyces cerevisiae* that primarily emphasized the recruitment step in the transcription cycle and assumed that little regulation occurred after the formation of a pre-initiation complex (PIC). However, recent findings in metazoan systems have revealed that much of the transcription regulation occurs well after the recruitment of Pol II and the transcription machinery to a gene promoter, through controlling pausing and the efficiency of early elongation. Thus, we are in the midst of a paradigm shift in our understanding of gene regulation as it applies to higher eukaryotic systems.

Here, we focus on the promoter-proximal pausing of Pol II and its regulated escape into productive elongation[1]. In this Review, we use the shorthand of calling promoter-proximal pausing simply Pol II pausing, although (as discussed below) there is evidence that the polymerase can pause during productive elongation as well. We describe the basic biochemical properties of paused Pol II and recent evidence from genome-wide studies, indicating that this type of regulation is widespread in metazoans[2–9]. We then discuss how Pol II pausing can influence chromatin structure at promoters to facilitate gene activity and how pausing might lead to rapid or synchronous transcriptional responses when cells are exposed to an activation signal. We also highlight how regulation of early elongation can interplay with factors that regulate Pol II recruitment and the RNA-processing machineries to finely modulate transcription in response to distinct signals that occur during development, homeostasis and disease.

Discovery of promoter-proximal Pol II pausing

Although pausing has only recently been recognized to be a prevalent regulatory strategy, evidence that transcription elongation could be a rate-limiting step in gene expression surfaced more than 30 years ago. A number of studies in mammalian cell culture in the late 1970s and the early 1980s indicated that transcription initiation did not obligatorily lead to the production of a full-length transcript[10,11]. Insight into when this post-initiation block occurred came from *in vivo* analyses of the uninduced *Drosophila melanogaster* heat shock genes (Hsp genes) in the Lis laboratory[15] using ultraviolet protein–DNA crosslinking[12], nuclear run-on assays[13], permanganate footprinting[14] and analysis of the short, capped RNAs (scRNAs) (BOX 1). These studies revealed that transcriptionally engaged polymerase accumulates just downstream of the Hsp promoters and is associated with 20–60-nucleotide-long nascent RNA[13,15]. The properties of these promoter-associated Pol IIs were strikingly similar to those ascribed by Roberts and colleagues[16] to *Escherichia coli* RNA polymerases that pause at the start of the lambda late gene transcription unit. Thus, the Lis group[17] referred to the promoter-proximal Pol II found at the Hsp genes as 'paused'.

Importantly, additional work carried out in the late 1980s and early 1990s revealed that other promoters displayed paused Pol II. In fact, a large fraction of *D. melanogaster* genes investigated in detail (6 out of 10 genes) showed characteristics of Pol II pausing[17,18]. Moreover, a handful of mammalian genes, including key cell regulatory genes, such as human *MYC* (also known as *c-MYC*) and *FOS*, showed an enrichment of engaged Pol II just downstream of the transcription start site (TSS) that was effectively indistinguishable from that seen at the *D. melanogaster* Hsp genes[19–21]. Pol II was also found to accumulate on promoters of the HIV long terminal repeat (LTR), although this regulatory system displayed several features that distinguished it from pausing at endogenous genes. First, the nascent RNA transcribed at the HIV LTR forms a functionally important unique secondary structure[4,22] and, second, the HIV LTR produces an abundant, 59-nucleotide-long RNA that results from premature termination of the early elongation complex[22]. By contrast, there is no current evidence suggesting that high levels of promoter-proximal termination by Pol II at endogenous genes. Nonetheless, these studies of multiple-gene systems provided evidence of regulation after recruitment of Pol II to a gene promoter, begging the question of how widespread these 'alternative' mechanisms of gene regulation might be.

Pol II pausing is widespread in metazoans

The findings described above, although appreciated by the field, were eclipsed by studies of transcription in the powerful yeast model system that demonstrated that recruitment of Pol II to promoters was a major mode of gene regulation and provided no compelling evidence for promoter-proximally paused Pol II[23,24]. The interest in promoter-associated polymerase was recently reignited by the ability to carry out Pol II chromatin immunoprecipitation (ChIP) assays genome-wide using ChIP followed by microarray (ChIP–chip) or ChIP followed by high-throughput sequencing (ChIP–seq) techniques (see BOX 1 on methods for detecting Pol II). These studies have provided evidence for widespread post-recruitment regulation of gene expression in metazoans.

Identifying and defining promoter-associated Pol II. The global localization of Pol II occupancy in multiple species has revealed that Pol II exhibits various distributions along genes that provide insights into the mechanics of their regulation. In yeast, Pol II usually displays an approximately uniform distribution across the transcription unit[25], as expected from models in which Pol II, after it has been recruited, experiences few regulatory barriers. By striking contrast, Pol II in *D. melanogaster*[4,6,9] and mammalian cells[26,27] is frequently non-uniformly distributed on the bodies of genes. In these higher eukaryotes, a large fraction of genes displays Pol II signal that is concentrated near TSSs, indicating that polymerases recruited to these promoters are not efficiently released downstream into the gene. However, initial genome-wide studies used Pol II ChIP to localize the polymerase, which in itself is not sufficient to distinguish between species that are paused during early elongation and those that are blocked at another post-recruitment step in the transcription cycle (FIG. 1; BOX 1; see ChIP). Therefore, elucidating the true status of poised, promoter-associated Pol II (FIG. 1) required the development and use of additional assays.

Defining the status of promoter-proximal Pol II is important for understanding the regulation of polymerase release from the promoter region into productive synthesis. For example, a recruited polymerase that fails to initiate RNA synthesis and is trapped as a PIC (FIG. 1a) would involve substantially different mechanisms for release than would a Pol II that had synthesized a short transcript but that was blocked in early elongation. Moreover, early elongation complexes that accumulate downstream of promoters can be present in several conformations that are not all competent to resume RNA synthesis. Paused Pol II can be readily induced to restart transcription (FIG. 1b), whereas arrested and terminating elongation complexes cannot (FIG. 1c,d) and require either rescue or reinitiation to generate a productive transcript.

The first genome-wide ChIP study of Pol II distribution in human primary lung fibroblasts[26] in 2005 referred to promoter-proximal accumulation of Pol II as PICs (FIG. 1a), because the peak of Pol II mapped near the TSSs and because extensive studies *in vitro* had firmly established the concept of a PIC as an intermediate that occurs early in the transcription cycle. However, in 2007, genomic analysis of Pol II in human embryonic stem cells (ESCs) revealed that the Pol II accumulation was accompanied by chromatin signatures of gene activity, suggesting that these Pol IIs had undergone transcription initiation[27]. Concurrent ChIP–chip analyses in *D. melanogaster* S2 cells and early embryos also identified a widespread accumulation of promoter-associated Pol II, and follow-up with permanganate footprinting (BOX 1) investigated whether the observed Pol II had paused during elongation through the promoter-proximal

Heat shock genes
(Hsp genes). These genes are a set of highly conserved genes that encode molecular chaperones. These genes are rapidly induced in cells or organisms in response to various cellular stresses, including a several-degree increase in temperature.

Long terminal repeat
(LTR). In HIV, this promoter resides in a region of LTRs. Transcription from this promoter produces both viral proteins and new RNA genomes.

Box 1 | **Methods used to detect paused polymerase**

Chromatin immunoprecipitation

This technique involves protein–DNA crosslinking coupled with immunoprecipitation. When an antibody that targets RNA polymerase II (Pol II) is used, chromatin immunoprecipitation (ChIP) can identify regions of DNA that are bound by Pol II. We note that several antibodies are available that recognize different phosphorylation states of the Pol II carboxy-terminal heptapeptide repeat domain (CTD), including the early elongation form characterized by predominant Ser7 and Ser5 phosphorylation and the productive elongation form that is also phosphorylated at Ser2 (REF. 91). However, because our knowledge is incomplete concerning how the many reported modifications of the CTD affect the affinity of these antibodies for their target epitope and because we do not know the complete modification status of Pol II at every step of the transcription cycle, we caution against using phospho-CTD antibodies as the sole method for establishing the presence of a paused Pol II.

Advantages. A snapshot of Pol II distribution can be achieved through rapid crosslinking of whole cells. Analysis of individual genes is straightforward using quantitative PCR. ChIP is readily adapted for high-throughput genome-wide studies, either by hybridizing immunoprecipitated DNA to an array (ChIP–chip) or through high-throughput sequencing of Pol-II-bound DNA (ChIP–seq).

Disadvantages. ChIP has a low spatial resolution and sensitivity, and signal and specificity are highly dependent on the antibody used.

Permanganate footprinting

This method detects locally melted regions of DNA, such as those that arise from paused polymerase by selectively modifying unpaired thymines in a stable, open transcription bubble. Modified thymines are then converted to strand breaks that are visualized by ligation-meditated PCR (LM-PCR).

Advantages. This method can be carried out directly on whole cells or tissues, it achieves nucleotide-level resolution for mapping paused polymerase, and it does not require antibodies.

Disadvantages. Permanganate footprinting is low-throughput, as the readout involves LM-PCR on individual genes. As a result, the application is limited to genes in which good primers for primer extension and LM-PCR can be designed, making permanganate footprinting challenging in mammalian systems.

Nuclear run-on assays and global run-on sequencing

Run-on assays detect elongation-competent RNA polymerases through their ability to incorporate a label into nascent RNA in isolated nuclei. Global run-on sequencing (GRO-seq) is a genome-wide nuclear run-on method that allows high-resolution mapping of transcriptionally engaged Pol II. Transcriptionally engaged Pol II is allowed to elongate for ~100 nucleotides in the presence of 5-bromouridine-5′-triphosphate (Br-UTP). The RNAs are then base-hydrolysed to ~100 nucleotides in length, and RNAs are affinity-purified using anti-BrUTP beads and specific linkers are added to the 5′ and 3′ ends before submitting samples to next-generation sequencing. The specific 5′ primer allows the orientation of the RNAs to be determined, whereas three affinity purifications at various points in the sample preparation provide a low background.

Advantages. These methods specifically reveal transcriptionally engaged and active polymerase and have a high sensitivity and a low background. They are adaptable for high-throughput genome-wide applications and can be used in various organisms.

Disadvantages. These methods are technically challenging and require preparation of nuclei. Resolution for mapping of paused polymerase is reduced by the necessity to allow polymerase to run-on and to incorporate labelled nucleotides into RNA.

Short, capped RNA analysis and sequencing

Short, capped RNA (scRNA) analysis involves direct isolation and identification of short RNA species derived from promoter-proximal Pol II. Initial use of this technique isolated RNAs produced at individual genes using complementary sequence-specific probes[15,98]. Extending this technique genome-wide by scRNA sequencing (scRNA-seq)[7] involves isolation of nuclei, size selection of short (<100-nucleotide) RNA species and enzymatic degradation of RNAs that lack the 5′ cap before directional linker addition and high-throughput sequencing. This strategy allows for highly sensitive detection of RNA produced by promoter-proximal Pol II, including RNA species that have been generated by Pol II that pauses only transiently or that prematurely terminates transcription.

Advantages. scRNA-seq pinpoints the start site of transcription and the final nucleotide added by paused polymerase at single-nucleotide resolution. A low-background and high-sensitivity assay is well-suited for high-throughput genome-wide applications. It does not require antibodies, cell treatment or labelling, and it can be used in various organisms.

Disadvantages. This technique is technically challenging and requires preparation of nuclei. It does not distinguish between RNA species that remain associated with paused Pol II and those that have been released through transcription termination.

Ligation-meditated PCR (LM-PCR). A technique that can be used to map the ends of DNA fragments precisely from a specific region of the genome. Small DNA linkers are added to ends of DNA samples and then primers that are complementary to this linker are combined with a sequence-specific primer to amplify the DNA of interest by PCR.

region[6,9]. This permanganate footprinting demonstrated the presence of stably melted DNA located 20–60 bases downstream of the TSSs of dozens of genes[4,6,9], and this is diagnostic of a transcriptionally engaged polymerase. In addition, depletion of negative elongation factor (NELF; discussed below), which induces pausing, released many

of these Pol II complexes from promoter regions, further supporting their designation as engaged but paused species[6]. However, it was unclear what fraction of these promoter-proximal elongation complexes was competent to resume RNA synthesis, because permanganate footprinting cannot distinguish between paused, arrested and terminating complexes (FIG. 1b–d). Thus, these Pol II species were initially referred to as 'stalled'[4,6,9], which is a general term that includes all of these different forms of engaged Pol II (FIG. 1f).

Global nuclear run-on sequencing (GRO-seq; see BOX 1) in human primary lung fibroblasts in 2008 confirmed that many of the promoter-associated Pol II molecules are indeed paused by demonstrating that they are largely capable of resuming transcription *in vitro* following treatment with the detergent sarkosyl[2]. Sarkosyl is thought to remove pause-inducing factors from the elongation complex, allowing RNA synthesis to continue. Importantly, arrested or terminating elongation complexes cannot be induced to 'run-on' in this assay, such that the peak of signal observed near promoters by GRO-seq clearly represents Pol II in a paused state[2,5,28].

Furthermore, although backtracking and arrest of early elongation complexes (FIG. 1c) were found to occur commonly *in vitro* using metazoan transcription systems[29], promoter-proximal backtracked complexes were found to be rapidly rescued from arrest *in vivo*. Indeed, genomic analyses of scRNAs generated by paused Pol II in *D. melanogaster* demonstrated that such backtracking is efficiently followed by TFIIS (also known as TCEA1)-mediated cleavage of the extruded RNA[7]. Thus, current evidence suggests that a large fraction of promoter-associated Pol II is in a stably paused state that is competent to resume RNA synthesis. However, more detailed analyses of Pol II status, and in particular the contribution of premature transcription termination to the promoter-proximal Pol II signal, will be required to address this issue conclusively.

Patterns of paused Pol II. Many metazoan genes display higher levels of Pol II on their promoters than on their gene body, but this ratio of promoter to gene body Pol II density, which is termed the pausing index, can dramatically vary among genes (examples shown in FIG. 2). Owing to this broad range of pausing indices and to the inherent difficulties in applying a discrete threshold to continuous data sets, calculations of the fraction of genes that display Pol II pausing in mouse ESCs have produced estimates ranging from ~30% to ~90%[5,8]. Therefore, rather than reflecting a biological difference, the reported differences in prevalence of pausing are probably a consequence of using different methods and different statistical criteria to define Pol II pausing. Notably, when GRO-seq and a consistent data analysis method are used to measure Pol II occupancy in human primary lung fibroblasts, mouse ESCs or *D. melanogaster* cell culture, a similar fraction of genes is found to display paused Pol II in all cases: ~30% of all genes[2,5,28]. Thus, despite the fact that the definition of what constitutes 'pausing' is highly variable among the different groups studying this phenomenon, the

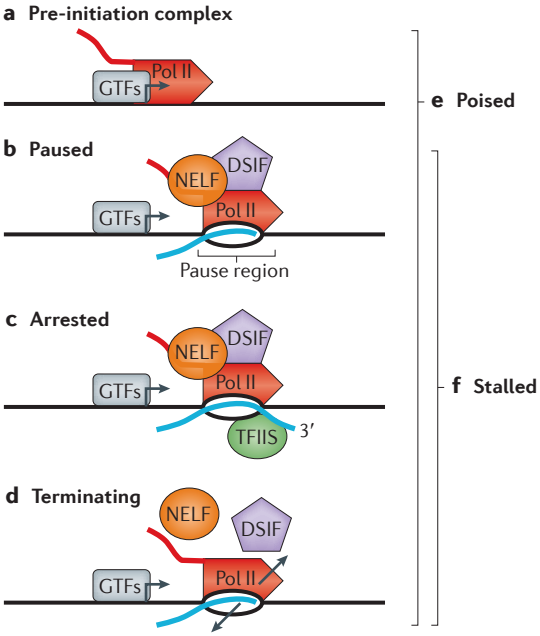
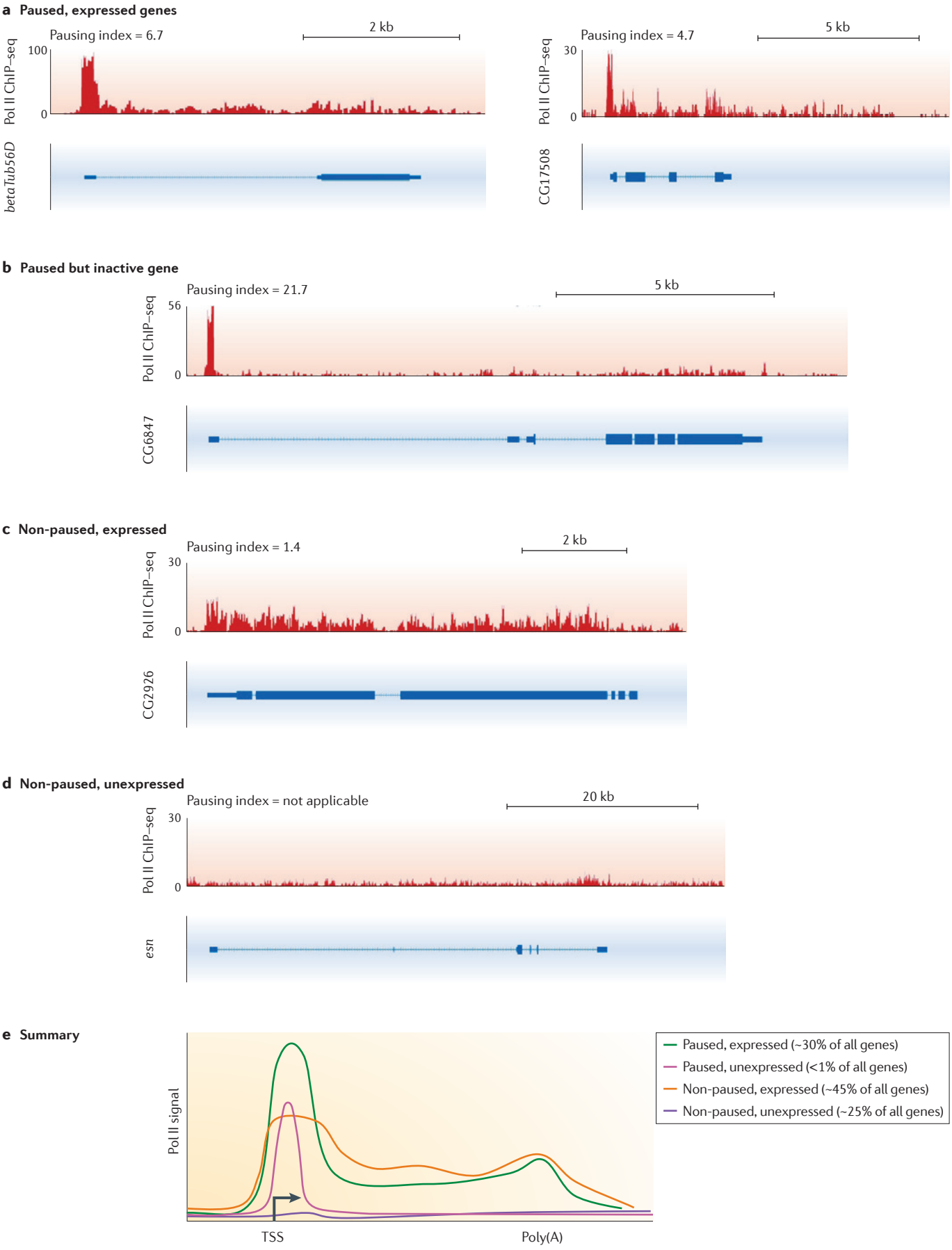


Figure 1 | **Defining the terms used to describe promoter-associated Pol II complexes.** The promoter region is depicted with the transcription start site (TSS) labelled with an arrow. RNA polymerase II (Pol II) is illustrated as a red rocket. The general transcription factors (GTFs; grey oval) are shown centred at the TSS (arrow). The pause-inducing factors negative elongation factor (NELF; orange oval), DRB-sensitivity-inducing factor (DSIF; purple pentagon) and transcript cleavage factor TFIIS (green circle) are shown. The nascent RNA transcript is shown in blue, and a bracket indicates the pausing region, usually 20–60 nucleotides downstream from the TSS. **a** | Pre-initiation complex: an entry form of Pol II in a complex with general transcription factors in which the polymerase is bound to the promoter DNA but has not yet initiated RNA synthesis. **b** | Paused: an early elongation complex that has transiently halted RNA synthesis. Paused polymerase is fully competent to resume elongation, remaining stably engaged and associated with the nascent RNA. The 3′ end of the RNA may have 'frayed' slightly from the Pol II active site in a manner that would slow further RNA synthesis, but the RNA is properly aligned with the active site. Two protein complexes, DSIF and NELF, reduce the rate of elongation and facilitate the establishment of the stably paused state. **c** | Arrested: a stably engaged elongation complex wherein the polymerase has backtracked along the DNA template, such that the RNA 3′ end is displaced from the active site. Restart of an arrested complex usually requires TFIIS, which induces Pol II to cleave the nascent RNA at the active site, creating a new 3′ end that is properly aligned with the Pol II active site and releasing a short (2–9-nucleotide) 3′ RNA. **d** | Terminating: an unstable elongation complex that is in the process of dissociating from the DNA template and releasing the nascent RNA. The released Pol II could have the potential rapidly to reinitiate transcription and to 'recycle' at the promoter. **e** | Poised: a generic term that simply indicates that Pol II is located near the TSS but does not specify anything about its transcriptional status. It can include any of the above complexes (**a**–**d**). **f** | Stalled: a term that indicates Pol II is engaged in transcription but that makes no assumptions about its ability to resume synthesis. This term includes paused, arrested and terminating complexes (**b**–**d**, above).



◀ Figure 2 | **Patterns of Pol II distribution across gene regions.** RNA polymerase II (Pol II) chromatin immunoprecipitation followed by high-throughput sequencing (ChIP–seq) signal is shown at genes that exemplify various Pol II distributions observed across metazoan genomes. Example genes are taken from *Drosophila melanogaster* S2 cells[7], and the pausing index of each gene is indicated. **a** | Two paused but active genes with differing pausing indices. **b** | A paused but inactive gene. **c** | A non-paused gene that is expressed. **d** | A non-paused, unexpressed gene. We note that a pausing index cannot reliably be calculated for genes that lack significant Pol II promoter signal. **e** | Shown are profiles of Pol II signal that exemplify the four major groups of genes. We note that the 'paused, unexpressed' group is significantly underrepresented *in vivo*, suggesting that most paused genes display some basal RNA synthesis. Approximate percentages of genes that fall into each category are given[2,5,28].

proportion of genes that exhibit Pol II pausing appears to be approximately constant across species and developmental stages reported to date.

Interestingly, in all systems evaluated thus far, genes that exhibit pausing are enriched in signal-responsive pathways, including development, cell proliferation and stress or damage responses. This enrichment is of particular interest in pluripotent cells, such as ESCs, in which pausing has been suggested to have a role in cell differentiation[27]. Underscoring that the presence and level of paused Pol II can be highly regulated, the specific genes that are paused in various cell types and under varying conditions, such as cell stress or cell cycle regulation, can differ dramatically[5,30].

Genomic analysis of Pol II distribution also indicates that pausing occurs at genes across the range of expression levels[3,7,8]. In fact, recent global analyses of Pol II distribution by GRO-seq indicate that few paused genes are transcriptionally inactive (<1%)[2,5]. This argues strongly against a common perception that Pol II pausing is predominantly a mechanism to silence gene expression[31,32]. It is consistent with previous data on paused Pol II: for example, all of the traditionally defined paused genes (such as *D. melanogaster* Hsp genes and *β-tubulin*, and mammalian *MYC* and *FOS*) exhibit considerable basal expression, and the *D. melanogaster* Hsp genes continue to undergo pausing during activation[1]. On the basis of these data, we argue that pausing should be considered to be a mechanism for tuning expression from active genes and perhaps for poising them for future changes in expression, rather than as a means of gene inactivation.

We note that Pol II can reduce its elongation velocity and/or can pause during productive synthesis as well, although the factors that are involved and the mechanisms that govern pausing within the gene appear to be distinct from those that regulate promoter-proximal Pol II[33]. Slowing of productive elongation is best characterized at the 3′ end of genes, where considerable accumulation of Pol II is observed just downstream of the poly(A) site[2,34] (FIG. 2e). This slowing of Pol II at the end of the transcription unit is thought to facilitate the coupling of RNA cleavage with transcription termination[35]. Likewise, pausing within exons has been reported[36], in which it is proposed to have a role in promoting splicing. Accordingly, evidence suggests that Pol II elongation rates can have an impact on alternative splicing, and slower elongation favours inclusion of exons with inherently weak splice sites[37]. As such, we now appreciate that gene expression can be regulated at almost every step in the transcription cycle, from PIC formation through to productive elongation and RNA processing.

Mechanisms of Pol II pausing and release

The establishment of paused polymerase requires both bringing Pol II to promoters and stably retaining the early elongation complex within the pause region. These depend on the intrinsic strength of the core promoter[38] and on specific transcription factors that recruit chromatin-remodelling proteins and the transcription machinery (FIG. 3a,b; for example, transcription factor TF1). After formation of the PIC, the promoter DNA is locally unwound, allowing the polymerase to initiate RNA synthesis and to undergo promoter escape, wherein it releases many of the contacts with promoter-bound general transcription factors (GTFs)[39]. During this process, the GTF TFIIH phosphorylates serine residues within the carboxy-terminal heptapeptide repeat domain (CTD) of the largest Pol II subunit. The early elongation complex then extends the nascent RNA as it moves downstream into the gene. However, detailed analyses of early elongation have demonstrated that this process is fraught with difficulty[29].

Work done largely in the Handa and Price laboratories[29,40,41] in the early 1990s demonstrated that Pol II elongates inefficiently through the promoter-proximal region, displaying a strong tendency to halt or to terminate within the first 100 nucleotides. These studies provided key mechanistic insights into Pol II pausing by revealing that the block in early transcription elongation results in part from the association of two pause-inducing factors with the early elongation complex (FIG. 3c). These factors, which are called DRB-sensitivity-inducing factor (DSIF; also known as SPT5–SPT4)[41] and NELF[42], are together sufficient to inhibit early elongation in a purified system, indicating that they work directly on the polymerase to help to establish the paused elongation complex. Consistent with the lack of evidence for Pol II pausing in *S. cerevisiae*, homologues of the pause-inducing NELF proteins are absent in yeast but are conserved from *D. melanogaster* to humans[43].

Despite the clear importance of DSIF and/or NELF in establishing paused polymerase, growing evidence suggests that these factors are not alone in affecting the residence time of promoter-associated Pol II. Recent *in vitro* work suggests that additional factors, such as GDOWN1 (also known as GRINL1A) and the general transcription factor TFIIF may also influence the stability or lifetime of the paused polymerase, perhaps by affecting the susceptibility of the early elongation complex to premature termination[44]. Although it remains unclear whether termination in the promoter-proximal region occurs *in vivo*, Pol II ChIP–seq studies have provided evidence for premature termination within transcribed units (that is, downstream of +500)[45], suggesting that the processivity of elongating Pol II is continually subject to regulation. Thus, much yet remains to be learned about how the efficiency of early elongation is regulated at the mechanistic and biochemical level.

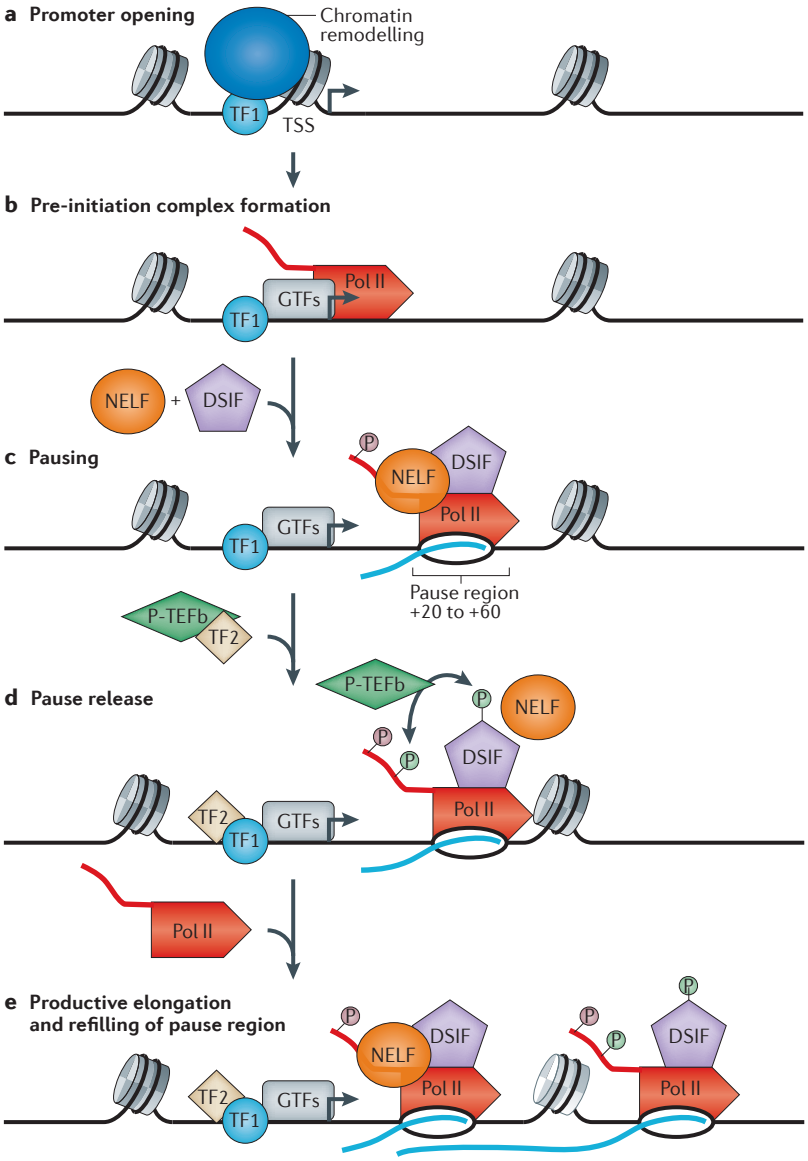


Figure 3 | **Establishment and release of paused Pol II.** The promoter region is shown with the transcription start site (TSS) labelled with an arrow. Nucleosomes are depicted in grey, and RNA polymerase II (Pol II) is illustrated as a red rocket. The nascent RNA transcript is shown in blue. Factors that are involved in the establishment or release of paused Pol II, such as DRB sensitivity-inducing factor (DSIF; purple pentagon), negative elongation factor (NELF; orange oval) and positive transcription elongation factor b (P-TEFb; green diamond) are indicated. **a** | Promoter opening often involves binding a sequence-specific transcription factor (shown here as TF1, light blue circle) that brings in chromatin remodellers (blue oval) to remove nucleosomes from around the TSS and to render the promoter accessible for recruitment of the transcription machinery. **b** | Pre-initiation complex formation involves the recruitment of a set of general transcription factors (GTFs; grey oval) and Pol II, which is also facilitated by binding specific transcription factors (also shown as TF1 for simplicity). This step precedes the initiation of RNA synthesis. **c** | Pol II pausing occurs shortly after transcription initiation and involves the association of pausing factors DSIF and NELF. The paused Pol II is phosphorylated on its carboxy-terminal heptapeptide repeat domain (CTD; shown in pink). The region in which pausing takes place is indicated on the figure. **d** | Pause release is triggered by the recruitment of the P-TEFb kinase (green diamond), either directly or indirectly by a transcription factor (shown here as TF2; beige diamond). P-TEFb kinase phosphorylates the DSIF–NELF complex to release paused Pol II and also targets the CTD (shown in green). Phosphorylation of DSIF–NELF dissociates NELF from the elongation complex and transforms DSIF into a positive elongation factor that associates with Pol II throughout the gene. **e** | In the presence of both TF1 and TF2, escape of the paused Pol II into productive elongation is rapidly followed by entry of another Pol II into the pause site, allowing for efficient RNA production. When the gene is activated, some nucleosome disruption is likely, as depicted by the lighter colouring of the downstream nucleosome.

The maturation of paused Pol II to a productively elongating form requires the kinase activity positive transcription elongation factor b (P-TEFb)[40,46,47]. P-TEFb phosphorylates the repressive DSIF–NELF complex, causing NELF to dissociate from Pol II and transforming DSIF to a state that promotes Pol II elongation[33] (FIG. 3d). P-TEFb also carries out additional phosphorylation of serine residues within the Pol II CTD, creating a platform for binding of RNA-processing factors and chromatin-modifying factors that facilitate productive RNA synthesis[33,48]. Given its key role in pause release, there is considerable interest in understanding how P-TEFb is targeted to particular gene promoters (shown in FIG. 3d as TF2). Befitting the diversity of genes that exhibit Pol II pausing, a large repertoire of factors has been reported to carry out this activity, including the acetylated histone-binding protein bromodomain-containing 4 (BRD4)[49,50], DNA-binding transcription activators such as MYC and nuclear factor-κ B (NFκB)[8,51,52] and the MED26 component of the Mediator complex[53]. Moreover, P-TEFb is found to be associated with a large number of other elongation factors and chromatin-modifying proteins in the 'super elongation complex', suggesting that these factors work together to stimulate productive elongation[54–57].

Interestingly, although only a subset of genes appears to accumulate high levels of paused Pol II, most *D. melanogaster* or mammalian promoters display a detectable enrichment of polymerase near the promoter compared with the gene body[3,8]. In addition, analysis of the location of factors that regulate the establishment and release of pausing suggests that transient Pol II pausing is a general feature of the transcription cycle. For example, the vast majority of active promoters are bound by the pause-inducing factors DSIF and NELF[3,8]. The levels of DSIF and NELF at promoters correspond extremely well with total promoter Pol II, suggesting that these

factors associate with most early elongation complexes. Further, treatment of cells with the P-TEFb inhibitor flavopiridol blocks the entry of most Pol II into productive synthesis in both *D. melanogaster* and mammals[8,58], indicating that polymerase release from the promoter region typically requires the activity of P-TEFb. Taken together, these data suggest that the early elongation complex comes under the control of DSIF and NELF at most genes and that the escape of Pol II into productive elongation involves the release of this repressive complex by P-TEFb. Thus, we envision that the rate of P-TEFb recruitment would be crucial for determining both gene expression levels and the appearance of paused Pol II. At many genes, P-TEFb recruitment to promoters may immediately follow transcription initiation, leading to a rapid release of polymerase into the gene. However, at other genes, P-TEFb recruitment may be a much slower event, permitting accumulation of paused Pol II.

Functions of paused Pol II

Given the prevalence of paused Pol II at genes within important developmentally and environmentally responsive pathways, identifying the functional roles of paused Pol II has become an active topic of research. We discuss models for four functions below, some of which may be interconnected (FIG. 4).

Establishing permissive chromatin. Wrapping promoter DNA around histone proteins to form nucleosomes can present a barrier to transcription by rendering crucial recognition sequences inaccessible. As a result, remodelling promoter chromatin to remove or to displace nearby nucleosomes is often required to permit recruitment of the transcription machinery and gene expression[59] (FIG. 3a). Whereas many genes, especially those in yeast[60,61], have been shown to couple this nucleosome remodelling temporally with gene activation, genes with paused polymerase have been shown to undergo nucleosome removal to open promoters before and independently from gene activition[62,63]. Moreover, paused genes have been shown to persist in a nucleosome-deprived, regulatory-factor-accessible state that is dependent on the presence of the paused Pol II[3,64–66] (FIG. 4a).

A relationship between the paused polymerase and the lack of promoter nucleosomes is apparent at the *D. melanogaster* Hsp genes[62,63], in which the promoter regions were shown to be nucleosome-deprived even in the uninduced state. Further studies of *Hsp70* transgenes indicated that promoter-proximal mutations affecting the levels of paused Pol II also disrupted the binding of heat shock factor (HSF) to its target sites during heat shock and subsequent gene activation[64,67]. Notably, this work suggested that pausing could help to maintain an open and accessible promoter structure to facilitate binding by regulatory transcription factors as well as the transcription machinery.

The link between paused Pol II and maintenance of a nucleosome-deprived promoter has recently been demonstrated at a genome-wide level in *D. melanogater*[3]. Genes with paused polymerase were globally shown to possess low levels of promoter nucleosome occupancy, which was dependent on the presence of promoter-associated Pol II: depletion of the pause-inducing factor NELF, which considerably reduced promoter Pol II levels at highly paused genes, led to a concomitant increase in promoter nucleosome occupancy at these genes[3,65]. Thus, paused promoters display a dynamic competition for promoter binding between nucleosomes and Pol II. Importantly, genes affected in this way often decreased their expression levels following NELF depletion and loss of paused Pol II.

Interestingly, the underlying DNA sequence appears to contribute to the requirement for promoter-proximal Pol II to prevent nucleosome assembly over many TSSs. Packaging DNA into nucleosomes requires that the underlying sequences are somewhat flexible and amenable to making regular bends as they wrap around the histone proteins, and it has been shown that certain sequences are particularly well- or ill-suited for this purpose[68,69]. Strikingly, genes with high levels of paused Pol II in *D. melanogaster* possess promoter sequences that are nucleosome-friendly and that are predicted to promote chromatin assembly[3]. Genes at which less paused Pol II is present tend to disfavour nucleosome assembly[3]. Likewise, paused Pol II is enriched in mammals at CpG island promoters[2], which tend to possess open chromatin[70]. Although currently a subject of debate, evidence suggests that mammalian promoters with a moderately high CG content intrinsically favour nucleosome formation[71–73], suggesting that, like in *D. melanogaster*, the transcription machinery helps to maintain accessible chromatin architecture around these promoters.

Thus, it is tempting to speculate that highly regulated promoters, many of which exhibit paused Pol II, have evolved DNA sequences that enable a dynamic competition to occur between paused Pol II and nucleosomes[74]. For example, it has been proposed that the presence of paused Pol II poises genes in ESCs for expression during development, in part by altering promoter chromatin[27]. As such, the loss of paused Pol II later in development could permit nucleosome occlusion and permanent gene repression. The formation of repressive chromatin may be further enhanced by the recruitment of the Polycomb repressive complexes PRC1 and PRC2. Notably, in ESCs, bivalent genes[75] that contain both PRCs have substantially less paused Pol II than genes that lack Polycomb[5,76]. Likewise, in *D. melanogaster*, mutations in a key component of the PRC2 complex that would presumably create a more accessible chromatin structure allows an increase in Pol II recruitment and pausing on thousands of promoters in the early embryo[77]. Importantly, the chromatin-opening function of paused Pol II would be connected to the other potential functions of pausing. For example, the presence of paused Pol II might allow genes that are transcribed at lower basal levels to be continually accessible and primed for bursts of transcription activation in response to specific cues or for generating synchronous transcriptional responses to signalling (see below).

CpG islands
Regions of higher-than-normal CpG sequence content that are on average 1,000 base pairs in length. Such regions contain ~70% of all mammalian promoters, including both genes that are highly regulated and broadly expressed.

Polycomb
Regulate chromatin structure to contribute to epigenetic inheritance of a repressed state. They form several complexes, which are broadly defined as Polycomb repressive complexes 1 and 2 (PRC1 and PRC2), and these are thought to compact chromatin structure.

Bivalent genes
Exhibit histone modifications that are characteristic of both gene repression and activation. These genes display low levels of Pol II occupancy and activity and are hypothesized to be poised for activation during development.

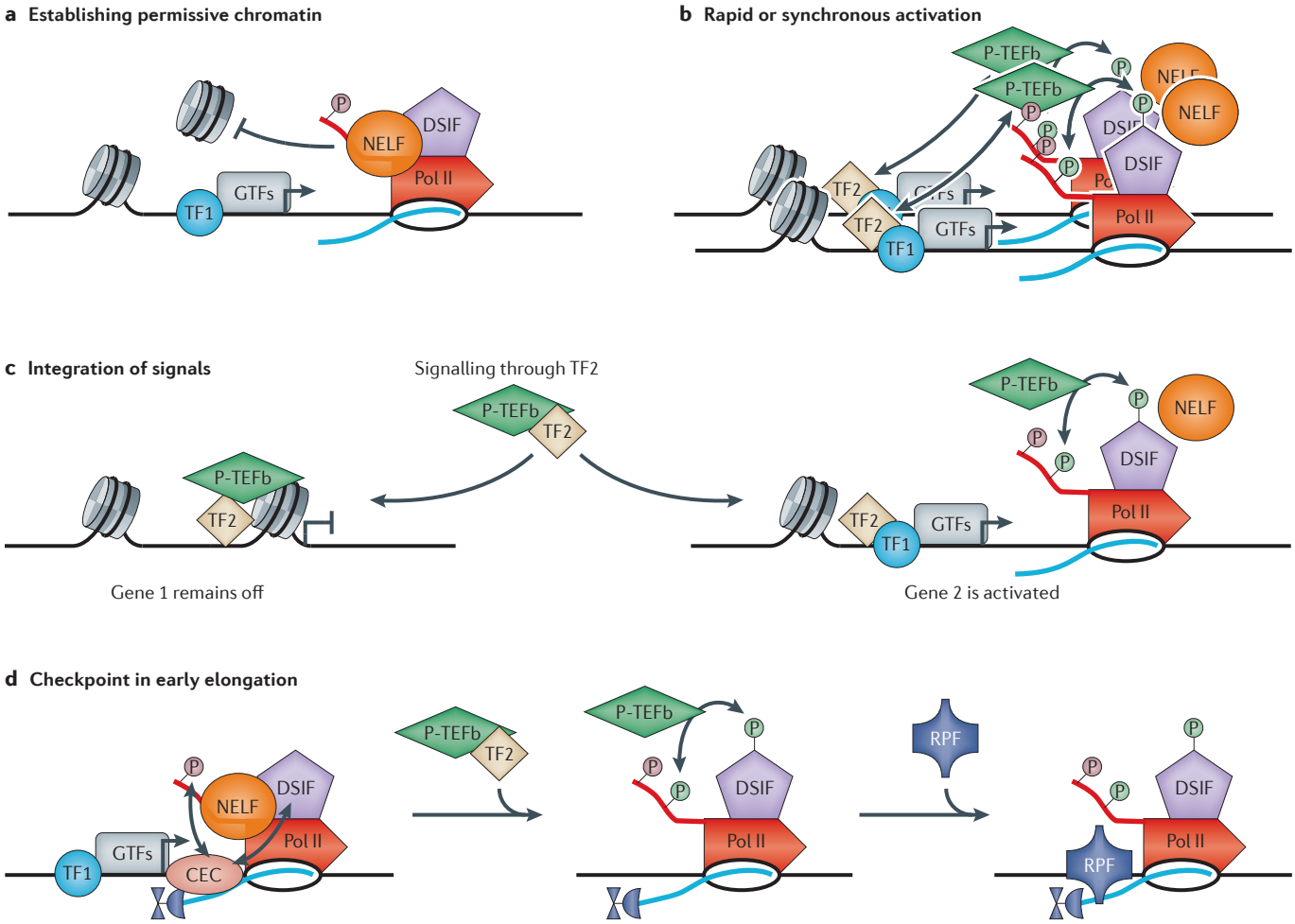


Figure 4 | **Illustrations of the main hypotheses for the functions of Pol II pausing.** The promoter region is depicted with the transcription start site (TSS) labelled with an arrow. Nucleosomes are shown in grey, and RNA polymerase II (Pol II) is illustrated as a red rocket. The general transcription factors (GTFs; grey oval) are shown centred at the TSS. The pause-inducing factors negative elongation factor (NELF; orange oval) and DRB-sensitivity-inducing factor (DSIF; purple pentagon) are shown. The nascent RNA transcript is shown in blue. **a** | Establishing permissive chromatin. After nucleosomes have been remodelled, paused Pol II helps to maintain the nucleosome-deprived structure by blocking nucleosome assembly over promoter sequences. Pausing would thus keep the promoter region accessible for activator and transcription factor binding. **b** | Rapid or synchronous gene activation. At a gene with paused Pol II, gene activation could proceed simply through recruitment of positive transcription elongation factor b (P-TEFb), thereby triggering the rapid release of paused Pol II into productive elongation. If a number of genes that harbour paused Pol II were all activated by the same signal and associated transcription factor (shown as TF2), then these genes could be activated simultaneously. **c** | Integrating multiple regulatory signals. Pausing represents a separate step in the transcription cycle for factors to act and allows for combinatorial control between transcription factors that recruit the transcription machinery (TF1) and those that trigger pause release (TF2), where both would be necessary for gene activation. In this example, signalling through TF2 such that it binds DNA and recruits P-TEFb would not lead to activation of a gene that was not paused (that is, a gene that lacks TF1) but would stimulate synthesis from gene 2 that was loaded with paused polymerase. **d** | A checkpoint in early elongation. On the left, arrows depict interactions between the capping enzyme complex (CEC) and DSIF–NELF as well as the Ser5 phosphorylation of the carboxy-terminal heptapeptide repeat domain (CTD) of Pol II, which is thought to stimulate capping activity. The hat represents the 5′ RNA cap. In the centre, P-TEFb-dependent phosphorylation events release paused Pol II and create a platform for binding of RNA-processing factors (RPFs) on the Ser2-phosphorylated CTD of Pol II, as shown on the right.

A pausing framework for rapid and/or synchronous activation. Although rapid gene activation at many genes involves mechanisms that are independent of pausing[78,79], the presence of Pol II is an appealing way to generate an accessible promoter region that can be quickly bound by activators and co-activators. Importantly, the presence of paused polymerase would allow a promoter to be readily switched from experiencing long-lived pausing to undergoing productive elongation simply through binding transcription activators that associate

with P-TEFb (FIG. 4b), bypassing a number of potentially slow or stochastic steps that are involved in PIC formation. The open promoter and scaffold of GTFs that remain after Pol II escape[80] may ensure continuous rapid entry of a succession of Pol II complexes on the activated gene (FIG. 3b). Moreover, the nucleosome-deprived status of paused promoters is likely to facilitate transcription factor binding, resulting in more efficient, reliable activation[81].

In support of a role for paused Pol II in rapid activation, pausing has been observed at *D. melanogaster* genes that are rapidly induced, such as the Hsp genes[1] and a number of genes that are involved in early embryonic development[6,9,82], leading to the idea that pausing facilitates synchronous changes in gene expression[83]. Consistent with this, many mammalian genes with paused Pol II (*MYC, FOS, JUNB* and *TNF-alpha*) have fast, transient expression kinetics[20,84,85].

However, not all rapidly induced genes display paused Pol II before activation[86–88], nor are most paused genes highly inducible. In fact, recent work that surveyed the prevalence of paused genes across several signal transduction networks in *D. melanogaster* and murine ESCs revealed that pausing was more enriched at promoters that encode the constitutively expressed components of signal transduction pathways (for example, receptors, kinases and transcription factors) than at the inducible downstream targets of these pathways[89]. Moreover, pausing was shown to regulate network activity largely through affecting the basal expression of signal transduction machineries[89]. Thus, the role of pausing in stimulus-responsive networks is not limited to poising inducible genes for activation. Instead, Pol II pausing can regulate the expression of key molecules, such as transcription factors and signalling proteins, thereby tuning cellular responsiveness to external cues.

Integrating multiple regulatory signals. Pausing represents an additional regulatory step in the transcription cycle beyond Pol II recruitment. Accordingly, this could allow activators that influence pause release to work together with factors that stimulate recruitment to exert combinatorial control of transcription levels[30,90] (FIG. 4c). Indeed, most promoters contain binding sites for multiple transcription activators. Importantly, some activators specifically function to recruit general transcription factors (GTFs) or to establish a paused Pol II (for example, transcription factors SP1 (REF. 90) and GAGA factor[4]; shown as TF1 in FIG. 4c), some factors bring P-TEFb to the promoter (for example, MYC and HIV TAT[8,52,90]; shown as TF2), and others appear both to recruit and to release paused Pol II (for example, NFκB and herpesvirus VP16 protein[51,90]). Thus, the particular combination of transcription activators that bind near any promoter would determine the rates of Pol II recruitment and pause release, thereby defining the rate-limiting step for transcription. In this way, cellular events that altered the levels or activity of individual transcription factors could be integrated on a gene-by-gene basis, depending on the sequence context and associated factors on the promoter and enhancer regions.

Checkpoint for coupling elongation and RNA processing. Pol II coordinates the efficient processing of nascent RNA: adding a cap to the 5′ end, coupling splicing events to transcription and facilitating the 3′ end processing of RNAs. By coupling RNA processing to the status and activity of Pol II itself, the cell ensures that nascent RNA is properly protected from degradation and efficiently matures into a functional mRNA. Pol II is phosphorylated on its CTD at various positions, providing a binding platform to recruit an entourage of protein factors that can execute both early and later events of RNA processing[91]. Phosphorylation of Ser5 within the CTD creates a binding platform for interaction with the 5′ capping enzyme (FIG. 4d) and stimulates the activity of this enzyme[92]. *In vivo*, 5′ capping occurs as the nascent RNA is extended from 20 to 30 nucleotides in length, and the bulk of RNAs associated with paused Pol II is capped[7,15]. This was initially determined by detailed analysis of the Hsp genes[15] and extended by recent global analyses[7] in *D. melanogaster*. Interactions have also been reported between the RNA-capping machinery and the pause-regulatory factor DSIF[93,94]. Thus, pausing may provide both a kinetic 'window of opportunity' as well as an interaction surface to facilitate addition of the 5′-methyl cap to the nascent RNA before the transition to productive elongation.

As mentioned above, phosphorylation of paused Pol II by P-TEFb provides a binding platform for complexes that carry out 3′ end processing[95]. As such, the requirement for P-TEFb activity to phosphorylate the DSIF–NELF complex and to trigger pause release may also ensure that Pol II does not proceed into the gene before it is appropriately modified for binding by the RNA-processing factors (FIG. 4d). Although rigorous tests of pausing as an obligatory checkpoint for Pol II CTD modification are lacking, the fact that the P-TEFb kinase phosphorylates both DSIF–NELF and Pol II might functionally couple pause release to this Pol II modification.

Conclusions and perspectives

In the past few years, a new picture of transcription regulation has emerged: genome-wide data in metazoans now point to the widespread importance of Pol II pausing in transcription regulation. Indeed, the escape of paused Pol II into productive elongation is regulated during environmental stress[6], immunological signalling[85] and development[96].

Studies of pausing over the decades coupled with an explosion of interest in recent years have led to considerable understanding of the characteristics and function of paused Pol II. Nonetheless, three major categories of questions remain. The first concerns the pervasiveness and patterns of pausing in eukaryotes. Studies underway in many laboratories will sample a broad swath of additional cell types and organisms in addition to *D. melanogaster*, mice and humans studied thus far. These studies should identify common features of genes regulated by this mechanism, as well as revealing cell-type-specific or condition-specific patterns of paused Pol II. Quantitative genome-wide studies should also assess whether paused polymerases constitute nearly all

promoter-associated Pol II or whether there are promoters with substantial amounts of other forms: for example, PICs or arrested Pol II, which are indicative of alternative modes of regulation (FIG. 1).

The second category deals with mechanistic questions that address pausing in molecular terms. We know several factors[41,42] and DNA elements[7,82] that are involved in stabilizing the paused state, but the full repertoire of factors and their interactions remains to be determined. Moreover, we know very little about how these factors interact to mediate efficient pausing. How stable are paused Pol II complexes, what are the relative levels of termination and escape to productive elongation and how might this balance be controlled? It will also be important to elucidate further how P-TEFb is either directly or indirectly targeted to promoters and how its kinase activity is regulated. Several mechanisms for P-TEFb recruitment have been documented, but surely more are to be discovered[33]. Future work should also elucidate exactly how the Pol II paused complex is transformed into a productively elongating machine. These events need to be examined in living cells with optical and biochemical methods that provide detailed information on the position and dynamics of paused Pol II and the accompanying protein and DNA interactions. Improvements in inhibitor discovery and in the already powerful molecular (BOX 1) and microscopic technologies[97] provoke optimism that these challenging mechanistic goals will be achieved.

The third category contains questions addressing the function of this regulation. In this Review, we emphasize varying levels of evidence for four potential roles of pausing (FIG. 4). These proposed functions will be clarified by further rigorous tests that include global studies as well as targeted analysis of specific genes and phenotypic analysis following systematic disruption of pausing. It will be interesting to define whether pausing serves different roles at different functional classes of genes and how these putative roles are interconnected. For example, the transient checkpoint established by pausing could be particularly useful at highly active genes to ensure that the nascent RNA is properly processed. However, the opening of chromatin structure by paused Pol II could both fine-tune the basal expression of signalling proteins[89] and facilitate a rapid transcriptional response[64]. In this Review, we highlight our current but incomplete understanding of Pol II pausing at promoters and its role in gene regulation. After decades of research and numerous cycles of simplifying and confounding theories and observations, we now have a framework and many of the tools needed to understand mechanistically transcription and its regulation genome-wide.

1. Lis, J. Promoter-associated pausing in promoter architecture and postinitiation transcriptional regulation. *Cold Spring Harb. Symp. Quant. Biol.* **63**, 347–356 (1998).
2. Core, L. J., Waterfall, J. J. & Lis, J. T. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848 (2008).
GRO-seq maps the position, amount and orientation of transcriptionally engaged RNA polymerases genome-wide and shows peaks of promoter-proximal polymerase residing on ~30% of human genes.
3. Gilchrist, D. A. *et al.* Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143**, 540–551 (2010).
Global analyses of Pol II pausing and nucleosome occupancy reveal that Pol II and nucleosomes compete for promoter occupancy to regulate gene expression coordinately.
4. Lee, C. *et al.* NELF and GAGA factor are linked to promoter-proximal pausing at many genes in *Drosophila*. *Mol. Cell. Biol.* **28**, 3290–3300 (2008).
Comprehensive analysis of promoter-associated Pol II in *D. melanogaster* using ChIP–chip and permanganate footprinting demonstrates that NELF-mediated pausing of Pol II is common in *D. melanogaster*.
5. Min, I. M. *et al.* Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.* **25**, 742–754 (2011).
6. Muse, G. W. *et al.* RNA polymerase is poised for activation across the genome. *Nature Genet.* **39**, 1507–1511 (2007).
Genome-wide Pol II ChIP–chip assays, coupled with permanganate footprinting and genetic manipulation of NELF, indicated that Pol II pausing is widespread in *D. melanogaster*.
7. Nechaev, S. *et al.* Global analysis of short RNAs reveals widespread promoter-proximal stalling and arrest of Pol II in *Drosophila*. *Science* **327**, 335–338 (2010).
8. Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
Investigation of transcription factor MYC reveals its important role in releasing paused Pol II through recruitment of P-TEFb.
9. Zeitlinger, J. *et al.* RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nature Genet.* **39**, 1512–1516 (2007).
Genome-wide Pol II ChIP–chip assays indicated that Pol II pausing is widespread during early embryonic development in *D. melanogaster*.
10. Fraser, N. W., Sehgal, P. B. & Darnell, J. E. DRB-induced premature termination of late adenovirus transcription. *Nature* **272**, 590–593 (1978).
11. Gariglio, P., Bellard, M. & Chambon, P. Clustering of RNA polymerase B molecules in the 5′ moiety of the adult beta-globin gene of hen erythrocytes. *Nucleic Acids Res.* **9**, 2589–2598 (1981).
12. Gilmour, D. S. & Lis, J. T. RNA polymerase II interacts with the promoter region of the noninduced hsp70 gene in *Drosophila melanogaster* cells. *Mol. Cell. Biol.* **6**, 3984–3989 (1986).
13. Rougvie, A. E. & Lis, J. T. The RNA polymerase II molecule at the 5′ end of the uninduced *hsp70* gene of *D. melanogaster* is transcriptionally engaged. *Cell* **54**, 795–804 (1988).
Nuclear run-on assays demonstrate that the Pol II complex associated with the *Hsp70* promoter is transcriptionally engaged but is unable to penetrate further into the gene without heat shock induction.
14. Giardina, C., Perez-Riba, M. & Lis, J. T. Promoter melting and TFIID complexes on *Drosophila* genes *in vivo*. *Genes Dev.* **6**, 2190–2200 (1992).
15. Rasmussen, E. B. & Lis, J. T. *In vivo* transcriptional pausing and cap formation on three *Drosophila* heat shock genes. *Proc. Natl Acad. Sci. USA* **90**, 7923–7927 (1993).
16. Grayhack, E. J., Yang, X. J., Lau, L. F. & Roberts, J. W. Phage lambda gene Q antiterminator recognizes RNA polymerase near the promoter and accelerates it through a pause site. *Cell* **42**, 259–269 (1985).
17. Rougvie, A. E. & Lis, J. T. Postinitiation transcriptional control in *Drosophila melanogaster*. *Mol. Cell. Biol.* **10**, 6041–6045 (1990).
18. Law, A., Hirayoshi, K., O'Brien, T. & Lis, J. T. Direct cloning of DNA that interacts *in vivo* with a specific protein: application to RNA polymerase II and sites of pausing in *Drosophila*. *Nucleic Acids Res.* **26**, 919–924 (1998).
19. Strobl, L. J. & Eick, D. Hold back of RNA polymerase II at the transcription start site mediates down-regulation of c-Myc *in vivo*. *EMBO J.* **11**, 3307–3314 (1992).
20. Krumm, A., Meulia, T., Brunvand, M. & Groudine, M. The block to transcriptional elongation within the human c-myc gene is determined in the promoter-proximal region. *Genes Dev.* **6**, 2201–2213 (1992).
21. Plet, A., Eick, D. & Blanchard, J. M. Elongation and premature termination of transcripts initiated from c-Fos and c-Myc promoters show dissimilar patterns. *Oncogene* **10**, 319–328 (1995).
22. Kao, S. Y., Calman, A. F., Luciw, P. A. & Peterlin, B. M. Anti-termination of transcription within the long terminal repeat of HIV-1 by *Tat* gene product. *Nature* **330**, 489–493 (1987).
23. Stargell, L. A. & Struhl, K. Mechanisms of transcriptional activation *in vivo*: two steps forward. *Trends Genet.* **12**, 311–315 (1996).
24. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).
25. Steinmetz, E. J. *et al.* Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol. Cell* **24**, 735–746 (2006).
26. Kim, T. H. *et al.* A high-resolution map of active promoters in the human genome. *Nature* **436**, 876–880 (2005).
27. Guenther, M. G., Levine, S. S., Boyer, L. A., Jaenisch, R. & Young, R. A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77–88 (2007).
28. Larschan, E. *et al.* X chromosome dosage compensation via enhanced transcriptional elongation in *Drosophila*. *Nature* **471**, 115–118 (2011).
29. Kephart, D. D., Marshall, N. F. & Price, D. H. Stability of *Drosophila* RNA polymerase II elongation complexes *in vitro*. *Mol. Cell. Biol.* **12**, 2067–2077 (1992).
30. Nechaev, S. & Adelman, K. Promoter-proximal Pol II: when stalling speeds things up. *Cell Cycle* **7**, 1539–1544 (2008).
31. Hargreaves, D. C., Horng, T. & Medzhitov, R. Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* **138**, 129–145 (2009).
32. Schones, D. E. *et al.* Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898 (2008).
33. Peterlin, B. M. & Price, D. H. Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell* **23**, 297–305 (2006).

34. Gromak, N., West, S. & Proudfoot, N. J. Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol. Cell. Biol.* **26**, 3986–3996 (2006).
35. Proudfoot, N. J. Ending the message: poly(A) signals then and now. *Genes Dev.* **25**, 1770–1782 (2011).
36. Carrillo Oesterreich, F., Preibisch, S. & Neugebauer, K. M. Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol. Cell* **40**, 571–581 (2010).
37. de la Mata, M. *et al.* A slow RNA polymerase II affects alternative splicing *in vivo*. *Mol. Cell* **12**, 525–532 (2003).
38. Juven-Gershon, T. & Kadonaga, J. T. Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev. Biol.* **339**, 225–229 (2010).
39. Roeder, R. G. Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett.* **579**, 909–915 (2005).
40. Marshall, N. F. & Price, D. H. Control of formation of two distinct classes of RNA polymerase II elongation complexes. *Mol. Cell. Biol.* **12**, 2078–2090 (1992).
41. Wada, T. *et al.* DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. *Genes Dev.* **12**, 343–356 (1998).
42. Yamaguchi, Y. *et al.* NELF, a multisubunit complex containing RD, cooperates with DSIF to repress RNA polymerase II elongation. *Cell* **97**, 41–51 (1999).
Biochemical assays reveal the presence and identity of the NELF complex and elucidate its role in inhibiting early transcription elongation.
43. Narita, T. *et al.* Human transcription elongation factor NELF: identification of novel subunits and reconstitution of the functionally active complex. *Mol. Cell. Biol.* **23**, 1863–1873 (2003).
44. Cheng, B. *et al.* Functional association of Gdown1 with RNA polymerase II poised on human genes. *Mol. Cell* **45**, 38–50 (2012).
45. Brannan, K. *et al.* mRNA decapping factors and the exonuclease Xrn2 function in widespread premature termination of RNA polymerase II transcription. *Mol. Cell* **46**, 311–324 (2012).
46. Marshall, N. F. & Price, D. H. Purification of P-TEFb, a transcription factor required for the transition into productive elongation. *J. Biol. Chem.* **270**, 12335–12338 (1995).
This pioneering paper describes the purification and characterization of the kinase P-TEFb.
47. Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D. & Handa, H. Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription *in vitro*. *EMBO J.* **17**, 7395–7403 (1998).
48. Czudnochowski, N., Bosken, C. A. & Geyer, M. Serine-7 but not serine-5 phosphorylation primes RNA polymerase II CTD for P-TEFb recognition. *Nature Commun.* **3**, 842 (2012).
49. 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* **19**, 523–534 (2005).
50. Yang, Z. *et al.* Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell* **19**, 535–545 (2005).
51. Barboric, M., Nissen, R. M., Kanazawa, S., Jabrane-Ferrat, N. & Peterlin, B. M. NF-κB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol. Cell* **8**, 327–337 (2001).
52. Eberhardy, S. & Farnham, P. Myc recruits P-TEFb to mediate the final step in the transcriptional activation of the cad promoter. *J. Biol. Chem.* **277**, 40156–40162 (2002).
53. Takahashi, H. *et al.* Human mediator subunit MED26 functions as a docking site for transcription elongation factors. *Cell* **146**, 92–104 (2011).
54. Mueller, D. *et al.* A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. *Blood* **110**, 4445–4454 (2007).
55. 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* **37**, 429–437 (2010).
56. Sobhian, B. *et al.* HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol. Cell* **38**, 439–451 (2010).
57. He, N. *et al.* HIV-1 Tat and host AFF4 recruit two transcription elongation factors into a bifunctional complex for coordinated activation of HIV-1 transcription. *Mol. Cell* **38**, 428–438 (2010).
58. Ni, Z. *et al.* P-TEFb is critical for the maturation of RNA polymerase II into productive elongation *in vivo*. *Mol. Cell. Biol.* **28**, 1161–1170 (2008).
59. Workman, J. L. Nucleosome displacement in transcription. *Genes Dev.* **20**, 2009–2017 (2006).
60. Boeger, H., Griesenbeck, J., Strattan, J. S. & Kornberg, R. D. Nucleosomes unfold completely at a transcriptionally active promoter. *Mol. Cell* **11**, 1587–1598 (2003).
61. Reinke, H. & Horz, W. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Mol. Cell* **11**, 1599–1607 (2003).
62. Wu, C. The 5′ ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature* **286**, 854–860 (1980).
63. Costlow, N. & Lis, J. T. High-resolution mapping of DNase I-hypersensitive sites of *Drosophila* heat shock genes in *Drosophila melanogaster* and *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **4**, 1853–1863 (1984).
64. Shopland, L. S., Hirayoshi, K., Fernandes, M. & Lis, J. T. 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.* **9**, 2756–2769 (1995).
65. Gilchrist, D. A. *et al.* NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev.* **22**, 1921–1933 (2008).
66. Leibovitch, B. A. *et al.* GAGA factor and the TFIID complex collaborate in generating an open chromatin structure at the *Drosophila melanogaster* hsp26 promoter. *Mol. Cell. Biol.* **22**, 6148–6157 (2002).
67. Lee, H., Kraus, K. W., Wolfner, M. F. & Lis, J. T. DNA sequence requirements for generating paused polymerase at the start of hsp70. *Genes Dev.* **6**, 284–295 (1992).
68. Iyer, V. & Struhl, K. Poly(dA:dT), a ubiquitous promoter element that stimulates transcription via its intrinsic DNA structure. *EMBO J.* **14**, 2570–2579 (1995).
69. Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366 (2009).
70. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Rev. Genet.* **13**, 484–492 (2012).
71. Tillo, D. *et al.* High nucleosome occupancy is encoded at human regulatory sequences. *PLoS ONE* **5**, e9129 (2010).
72. Valouev, A. *et al.* Determinants of nucleosome organization in primary human cells. *Nature* **474**, 516–520 (2011).
73. Ramirez-Carrozzi, V. R. *et al.* A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* **138**, 114–128 (2009).
74. Gilchrist, D. A. & Adelman, K. Coupling polymerase pausing and chromatin landscapes for precise regulation of transcription. *Biochim. Biophys. Acta* **1819**, 700–706 (2012).
75. Bernstein, B. E. *et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315–326 (2006).
76. Mikkelsen, T. S. *et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553–560 (2007).
77. Chopra, V. S. *et al.* The polycomb group mutant esc leads to augmented levels of paused Pol II in the *Drosophila* embryo. *Mol. Cell* **42**, 837–844 (2011).
78. Weake, V. & Workman, J. Inducible gene expression: diverse regulatory mechanisms. *Nature Rev. Genet.* **11**, 426–437 (2010).
79. Bryant, G. O. & Ptashne, M. Independent recruitment *in vivo* by Gal4 of two complexes required for transcription. *Mol. Cell* **11**, 1301–1309 (2003).
80. Yudkovsky, N., Ranish, J. A. & Hahn, S. A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408**, 225–229 (2000).
81. Bai, L., Charvin, G., Siggia, E. D. & Cross, F. R. Nucleosome-depleted regions in cell-cycle-regulated promoters ensure reliable gene expression in every cell cycle. *Dev. Cell* **18**, 544–555 (2010).
82. Hendrix, D. A., Hong, J. W., Zeitlinger, J., Rokhsar, D. S. & Levine, M. S. Promoter elements associated with RNA Pol II stalling in the *Drosophila* embryo. *Proc. Natl Acad. Sci. USA* **105**, 7762–7767 (2008).
83. Boettiger, A. N. & Levine, M. Synchronous and stochastic patterns of gene activation in the *Drosophila* embryo. *Science* **325**, 471–473 (2009).
84. Aida, M. *et al.* Transcriptional pausing caused by NELF plays a dual role in regulating immediate-early expression of the *junB* gene. *Mol. Cell. Biol.* **26**, 6094–6104 (2006).
85. Adelman, K. *et al.* Immediate mediators of the inflammatory response are poised for gene activation through RNA polymerase II stalling. *Proc. Natl Acad. Sci. USA* **106**, 18207–18212 (2009).
86. Nissen, R. M. & Yamamoto, K. R. The glucocorticoid receptor inhibits NFκB by interfering with serine-2 phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **14**, 2314–2329 (2000).
87. Kininis, M. *et al.* Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters. *Mol. Cell. Biol.* **27**, 5090–5104 (2007).
88. Lin, C. *et al.* Dynamic transcriptional events in embryonic stem cells mediated by the super elongation complex (SEC). *Genes Dev.* **25**, 1486–1498 (2011).
89. Gilchrist, D. A. *et al.* Regulating the regulators: the pervasive effects of Pol II pausing on stimulus-responsive gene networks. *Genes Dev.* **26**, 933–944 (2012).
90. Blau, J. *et al.* Three functional classes of transcriptional activation domain. *Mol. Cell. Biol.* **16**, 2044–2055 (1996).
91. Egloff, S., Dienstbier, M. & Murphy, S. Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet.* **28**, 333–341 (2012).
92. Ghosh, A., Shuman, S. & Lima, C. D. Structural insights to how mammalian capping enzyme reads the CTD code. *Mol. Cell* **43**, 299–310 (2011).
93. Mandal, S. S. *et al.* Functional interactions of RNA-capping enzyme with factors that positively and negatively regulate promoter escape by RNA polymerase II. *Proc. Natl Acad. Sci. USA* **101**, 7572–7577 (2004).
94. Moore, M. J. & Proudfoot, N. J. Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell* **136**, 688–700 (2009).
95. Buratowski, S. Progression through the RNA polymerase II CTD cycle. *Mol. Cell* **36**, 541–546 (2009).
96. Levine, M. Paused, R. N. A. Polymerase II as a developmental checkpoint. *Cell* **145**, 502–511 (2011).
97. Gould, T. J., Verkhusha, V. V. & Hess, S. T. Imaging biological structures with fluorescence photoactivation localization microscopy. *Nature Protoc.* **4**, 291–308 (2009).
98. Rasmussen, E. B. & Lis, J. T. Short transcripts of the ternary complex provide insight into RNA polymerase II elongational pausing. *J. Mol. Biol.* **252**, 522–535 (1995).

Acknowledgements
We thank the members of the Lis and Adelman laboratories for their helpful discussions on this Review. Funding for this work was provided by US National Institutes of Health (NIH) grant GM25232 to J.L. and the Intramural Research Program of the NIH National Institute of Environmental Health Sciences (Z01 ES101987) to K.A.

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Karen Adelman's homepage: http://www.niehs.nih.gov/research/atniehs/labs/lmc/tre/index.cfm
***Nature Reviews Genetics* Series on Modes of transcriptional regulation:** http://www.nature.com/nrg/series/transcriptionalregulation/index.html

ALL LINKS ARE ACTIVE IN THE ONLINE PDF