REVIEW

https://doi.org/10.1038/s41586-019-1517-4

Organization and regulation of gene transcription

Patrick Cramer[1]*

The regulated transcription of genes determines cell identity and function. Recent structural studies have elucidated mechanisms that govern the regulation of transcription by RNA polymerases during the initiation and elongation phases. Microscopy studies have revealed that transcription involves the condensation of factors in the cell nucleus. A model is emerging for the transcription of protein–coding genes in which distinct transient condensates form at gene promoters and in gene bodies to concentrate the factors required for transcription initiation and elongation, respectively. The transcribing enzyme RNA polymerase II may shuttle between these condensates in a phosphorylation–dependent manner. Molecular principles are being defined that rationalize transcriptional organization and regulation, and that will guide future investigations.

During the development of an organism, specific genes are expressed in distinct cells to establish different types of cell. This requires an intricate regulation of gene expression, which occurs to a large extent during the transcription of genes into RNA. Understanding gene regulation thus requires detailed knowledge of the mechanisms of transcription. Over the past decades, molecular and cellular studies have elucidated the structural organization of the factors that carry out transcription, and provided mechanistic insights into how these transcription complexes are regulated. Recent research has suggested how the transcription process could be organized by the condensation of particular factors in the nucleus of eukaryotic cells.

Transcription is carried out by RNA polymerase enzymes, which catalyse the DNA-dependent synthesis of RNA (Box 1). To initiate transcription, RNA polymerase recognizes the promoter region at the beginning of the gene (Fig. 1). The enzyme then opens the DNA duplex, starts to synthesize RNA and escapes from the promoter. The resulting elongation complex extends the RNA chain until it reaches a termination signal and releases DNA and RNA. Half a century ago, three RNA polymerases were isolated from eukaryotic cells[1]. These polymerases were later found to transcribe different classes of genes[2]. RNA polymerase (Pol) I produces the large ribosomal RNA precursor, Pol II synthesizes messenger RNAs and a variety of non-coding RNA, and Pol III produces transfer RNAs and the small ribosomal RNA. The polymerases differ in their associated factors and mechanisms of regulation (Box 1).

Here I describe our current view of how eukaryotic gene transcription is regulated. In particular, I discuss recent insights into the mechanisms of regulated transcription initiation and elongation by the three eukaryotic RNA polymerases. Furthermore, I describe recent studies that elucidate the dynamic condensation of Pol II transcription complexes in the nucleus. Finally, I integrate these studies into a model for the organization of Pol II transcription. The model postulates that an active gene is associated with separate and dynamic nuclear condensates that contain factors for transcription initiation and elongation, and that Pol II shuttles between these condensates in a phosphorylation-dependent manner.

Promoters and enhancers

For transcription to be initiated, the polymerase must first gain access to the promoter region at the beginning of a gene[3]. Promoter access is impaired by chromatin. Nucleosomes can inhibit initiation, and must be removed or shifted for transcription to occur[4–6]. Active promoters are found in nucleosome-depleted regions, which are flanked by specialized +1 and −1 nucleosomes on the downstream and upstream side of these regions, respectively[7,8]. Chromatin opening is regulated differently for distinct classes of promoters[9–11]. In the case of Pol II, one class of human promoters contains CpG islands that can impair the assembly of inhibitory nucleosomes and facilitate polymerase access[11]. Promoters such as these are often found at housekeeping genes that encode for proteins that are required in all the cell types of an organism. The activity of promoters that contain CpG islands can be altered by DNA methylation[12]. Another class of Pol II promoters contains a TATA element upstream of the transcription start site; promoters of this class are often found at genes that are cell-type-specific and regulated during differentiation[9].

Only a fraction of Pol II promoters is active in a particular cell. These promoters are activated by transcription factors that are available in the nucleus. Transcription factors bind, in a sequence-specific manner, to DNA elements and can guide polymerases to their target promoters[13–16]. Transcription factors use intrinsically disordered 'transactivation' regions that have amino acid sequences of low complexity[17–19] to recruit proteins that regulate promoter accessibility and transcription initiation[20]. About 1,600 human transcription factors are known[21]. Most of these factors bind to free DNA, but some can bind nucleosomal DNA[22]. The latter factors can act as 'pioneer' factors that open chromatin locally to enable transcription[23]. This often involves the recruitment of histone acetyltransferases and chromatin remodelling complexes that render promoters accessible to Pol II[3,4,24–27].

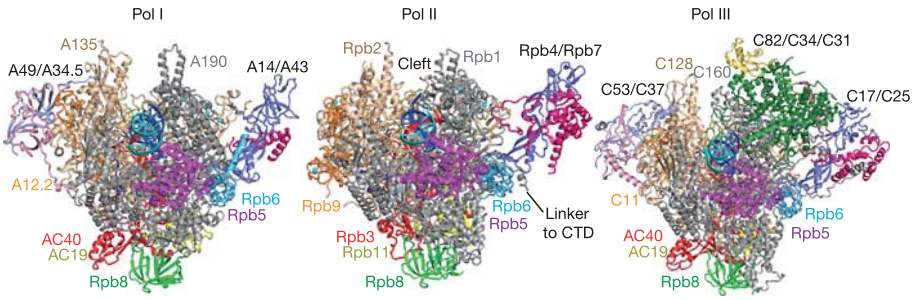
Transcription factors can bind near the promoter or at enhancers, which are distant DNA elements that regulate transcription[28,29]. Enhancers can be located far away (one million base pairs or more) from their target gene promoter[30]. Enhancers generally contain binding sites for multiple, cooperating transcription factors[31]. Active enhancers are transcribed in both directions and produce unstable enhancer RNA[32]. Such bidirectional transcription is often also observed for promoters[33]. This suggests that nucleosome depletion at promoters enables Pol II to bind in either orientation, which allows transcription to proceed in both directions. The communication of enhancers with their target gene promoters requires proximity and depends on dynamic chromatin architecture[34,35]. Enhancers usually operate within defined regions of the genome, which are known as topologically associated domains[36].

[1]Department of Molecular Biology, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany. e-mail: patrick.cramer@mpibpc.mpg.de

Box 1

RNA polymerase enzymes

The simplest RNA polymerases consist of a single polypeptide, and are encoded by bacteriophages[211]. The RNA polymerase that transcribes mitochondrial DNA resembles those of bacteriophages, but requires two additional factors for initiation and one factor for elongation[212]. Other cellular RNA polymerases also require additional factors, and are themselves multi-subunit enzymes. The multi-subunit RNA polymerase in bacteria has been extensively studied[213–216]. Eukaryotic cells contain three multi-subunit RNA polymerases, Pol I, Pol II and Pol III. The structures of these polymerases and many of their associated factors are known[85,217–222]. The figure below depicts the structures of yeast Pol I, Pol II and Pol III in the form of transcribing enzymes with DNA (template strand, blue; non-template strand, cyan) and RNA transcript (red). Pol I, Pol II and Pol III contain 14, 12 and 17 subunits, respectively. The two largest subunits are shown in silver and gold for all three RNA polymerases, and smaller subunits are shown in various other colours. The Pol II structure reveals that the two large subunits form a cleft that holds the active centre, and the smaller subunits are arrayed around the periphery[115,223]. This architecture is conserved in Pol I[224,225] and Pol III[110], both of which contain additional small subunits that cluster along the cleft (these subunits resemble Pol II initiation factors)[67,110,226]. Key elements of the polymerases are the clamp (which is mobile and can close the active centre), and the wall and dock, which bind TFIIB and its related initiation factors. The Pol II Rbp1 subunit contains a unique CTD that is composed of tandem heptapeptide repeats; this CTD is not visible in the structure, owing to its mobility. Archaea contain a single enzyme related to Pol II[227], whereas plant cells contain two additional Pol-II-like polymerases (Pol IV and Pol V), which are involved in gene silencing[228].



Promoter recognition

Promoters often contain conserved DNA sequence elements, which differ for Pol I[37], Pol II[10,38] and Pol III[39]. The RNA polymerases cannot recognize these promoter elements by themselves. Instead, promoter recognition requires transcription initiation factors, which form bridges between the polymerases and their cognate promoters (Table 1). The eukaryotic polymerases assemble with their cognate initiation factors to form specific pre-initiation complexes (PICs) on promoter DNA[40–44]. Recent structures of the PICs of Pol I[45–47], Pol II[48–57] and Pol III[58,59] elucidate how initiation factors enable promoter recognition (Fig. 2). In all three systems, initiation factors bind DNA upstream of the transcription start site, and position downstream DNA along the active centre cleft. However, although the three initiation systems are similar in their topology, they differ in other aspects.

The Pol II PIC contains general class II initiation factors[40,60–62]. These include the TATA box-binding protein TBP, which binds upstream DNA. The promoter–TBP complex then assembles with TFIIB, which binds the 'dock' and 'wall' domains of the polymerase to recruit a Pol II–TFIIF complex[51,63,64] (Fig. 2). TFIIB thus bridges between Pol II and the promoter; it also stimulates initial RNA synthesis allosterically[65]. Pol III uses a TFIIB-related factor called BRF1 and forms a PIC of similar architecture[58,59] (Fig. 2). Pol I uses the TFIIB-related factor TAF1B (or Rrn7 in yeast)[66], which, however, binds the polymerase protrusion domain[45–47]. In the Pol I PIC, upstream DNA is also bent in a different direction and DNA is loaded deeper into the cleft (Fig. 2). Thus, PIC architecture is similar for Pol II and Pol III but differs for Pol I. There are subunits in Pol I and Pol III that show distant similarities to the Pol II initiation factors TFIIE and TFIIF[67], but the three initiation systems also contain unrelated factors.

Many promoters lack obvious DNA sequence elements, which poses the question of how they are recognized. Initiation factors may recognize the flanking +1 nucleosome, as has been reported for the multiprotein complex TFIID[68], a factor that also contributes to the recognition of promoter elements[9,69]. Promoter recognition may also involve sensing the physical properties of promoter DNA—in particular, its 'bendability'[45] or local DNA strain[70]. Such indirect readout of features of DNA shape can explain why the architecture of the Pol II PIC is similar at promoters that differ strongly in sequence[71]. There is also evidence that Pol II PICs can differ in factor composition at different promoters[72], and this may contribute to promoter selection.

Promoter opening

A key function of the PIC is to open DNA. The mechanism of promoter opening differs between transcription systems. In the Pol I and Pol III systems, DNA is opened spontaneously with the use of binding energy alone[73–77]. It is speculated that spontaneous DNA opening requires DNA sequence features that facilitate strand separation. During evolution, such facilitated DNA 'meltability' may be maintained for the limited number of Pol I and Pol III promoters. In these systems, DNA opening is probably coupled to promoter recognition because some contacts with DNA can form only after DNA strand separation. Recognition of DNA single strands in open promoter DNA indeed occurs in the bacterial[78,79] and mitochondrial[80,81] transcription systems.

DNA opening by Pol II generally requires an additional enzyme, the DNA translocase XPB. XPB is a subunit of the general factor TFIIH[82] and binds DNA downstream of Pol II[83] (Fig. 2). XPB hydrolyses ATP to unwind DNA and propel it into the polymerase active centre[84–86]. The opening of Pol II promoters can be regulated in cells[87] and is blocked by the XPB inhibitor triptolide (Table 2). The dependence of promoters on XPB can vary[88,89]. The open promoter complex of Pol II appears to contain less-extensive protein–DNA interactions when compared to the related PIC of Pol III, which may explain why Pol II generally requires the help of XPB to open DNA whereas Pol III does not. During evolution, the Pol II PIC may have lost elements that stabilize open DNA and acquired elements that stabilize closed DNA. This may have rendered Pol II dependent on XPB and established an additional layer of transcriptional regulation.

Initiation regulation

The formation, stability and function of the PIC are regulated in all three transcription systems. The formation of the yeast Pol I PIC requires the initiation factor Rrn3 (Fig. 2). Phosphorylation of Rrn3

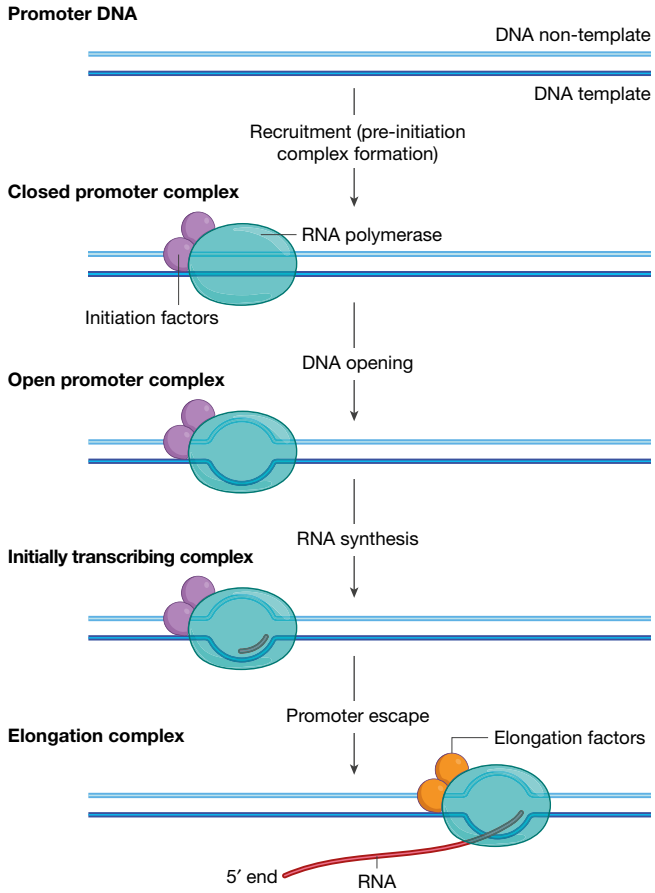


Fig. 1 | Key steps of gene transcription. The RNA polymerase enzyme associates with initiation factors to recognize promoter DNA and form a PIC. Subsequent DNA opening converts the closed promoter complex to the open promoter complex, which contains the DNA template strand in the polymerase active site. DNA-dependent RNA synthesis then generates an initially transcribing complex. When the RNA grows to a critical length, the polymerase escapes from the promoter and forms an elongation complex that can bind elongation factors. Finally, polymerase dissociation from DNA and RNA terminates the transcription cycle (not shown).

Table 1 | Selected examples of factors for human Pol II transcription

Transcription phase	Factor (no. of subunits)	Function
Initiation	TFIIB (1)	Bridges between Pol II and promoter DNA
	TFIID (14)	Contributes to promoter DNA recognition and includes TBP
	TFIIE (2)	Activates TFIIH and stabilizes the open promoter complex
	TFIIF (2)	Stabilizes TFIIB and the PIC
	TFIIH (10)	Catalyses DNA opening and Pol II CTD phosphorylation, and stimulates promoter escape
	Mediator (about 35)	Bridges between transcription factors and the PIC, stimulates CDK7 and can function in early elongation regulation
Elongation	DSIF (2)	Enables Pol II pausing and active elongation, and recruits elongation and 3′ processing factors
	Capping enzymes (3)	Catalyses 5′ RNA cap formation, and prevents pre-mRNA degradation by 5′ exonucleases
	NELF (4)	Stabilizes promoter-proximally paused Pol II
	P-TEFb (2)	Triggers activation of promoter-proximally paused Pol II by phosphorylating Pol II and elongation factors
	SEC (6)	Contains P-TEFb and ELL, a positive elongation factor
	SPT6 (1)	Recognizes phosphorylated CTD linker, and stimulates elongation
	PAF (5 or 6)	Stimulates elongation, and recruits chromatin-modifying enzymes
	CHD1 (1)	Remodels nucleosomes co-transcriptionally, and is ATP-dependent
	FACT (2)	Histone chaperone that facilitates nucleosome passage
	SET1 complex (7)	Histone methyltransferase that targets histone H3 lysine-4
	SET2 (1)	Histone methyltransferase that targets histone H3 lysine-36
	TFIIS (1)	Stimulates RNA cleavage, improves RNA proofreading and restarts arrested Pol II with backtracked RNA
Termination	CPSF (14)	Recognizes poly-adenylation sequence, cleaves pre-mRNA and dephosphorylates transcription machinery
	CStF (5)	Binds Pol II CTD and contributes to RNA binding
	XRN2 complex (3)	'Torpedo' nuclease complex that degrades cleaved, nascent RNA from the 5′ end and terminates Pol II transcription

prevents its binding to a Pol I subcomplex[90,91], represses transcription and restrains cell growth[92]. The human Rrn3 counterpart (TIF-IA) also regulates PIC assembly[93]. The formation of the Pol III PIC is also regulated by phosphorylation[94], and by the repressor protein MAF1[95]. Pol III recruitment to promoters is controlled by cell-cycle kinases and tumour-suppressor proteins that sequester the initiation factor TFIIIB[96].

Pol II initiation is regulated by the co-activator complex known as Mediator[97]. Mediator stabilizes the PIC in vitro[97], but in vivo the PIC–Mediator complex is short-lived[98,99]. Mediator contains a conserved core that comprises two modules (known as the 'head' and 'middle'), and contacts Pol II and the initiation factors TFIIB and TFIIH[50,100,101] (Fig. 2). The periphery of Mediator differs between species. Whereas the 'tail' module of Mediator binds activating transcription factors[102], the dissociable kinase module is implicated in repression[56,102,103]. Mediator stimulates phosphorylation of Pol II by the TFIIH kinase subunit CDK7[97]. CDK7 phosphorylates the C-terminal domain (CTD)—a tail-like extension from the body of Pol II[104]—to facilitate the transition to the elongation phase of transcription[97].

Elongation regulation

An elongation complex forms when the RNA grows to a critical length; this complex extends the RNA chain in a processive manner. The elongation complex of RNA polymerases generally contains one turn of a DNA–RNA hybrid duplex that is located within a DNA bubble[105–111]. To add a nucleotide to the growing RNA, the polymerase closes the active site[112,113], catalyses the formation of a phosphodiester bond using a two-metal-ion mechanism[112,114,115] and moves to the next template position[116]. Certain DNA sequences can, however, interrupt the nucleotide-addition cycle and induce transcriptional pausing. Pausing can lead to polymerase backtracking, arrest and termination[117]. Pol II can arrest in front of nucleosomes, but can be rescued by the elongation factor TFIIS[118]. TFIIS binds the Pol II funnel and pore, aligns the DNA–RNA hybrid with the active site and triggers the cleavage of backtracked RNA to restart transcription[119]. Pausing and arrest are prevented by a TFIIS-like subunit in Pol I[120] and Pol III[121].

In metazoan cells, the elongation phase of Pol II transcription is also regulated[122–124]. Pol II often pauses about 50 base pairs downstream of the transcription start site[125–127], and such promoter-proximal pausing is highly regulated[128] (Fig. 3). Recent studies have indicated the molecular mechanisms of polymerase pausing and its allosteric

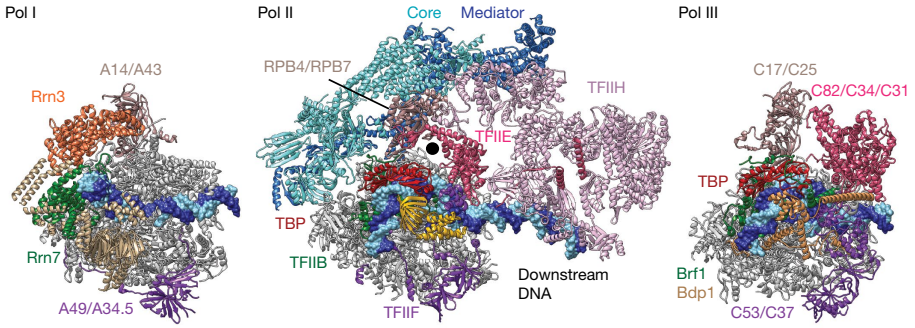


Fig. 2 | Structures of eukaryotic transcription PICs. The cryo-electron microscopy structures of transcription PICs of yeast Pol I[45], Pol II[48] and Pol III[59] are shown as ribbon models. Polymerases are in silver. DNA template and non-template strands are in blue and cyan, respectively. Upstream and downstream DNA are pointing to the left and right, respectively. Polymerases would travel to the right after promoter escape. Upstream DNA is bent in all complexes, but the point of bending and the direction of the bend differ for the Pol I complex (which also lacks TBP (red)). Related proteins are labelled and shown in the same colour. The Pol-I-specific initiation factor Rrn3 and the Pol-II-specific factors TFIIH and Mediator are indicated. For Pol III, TFIIB includes the subunits Brf1 (green) and Bdp1 (yellow). The point of CTD attachment to Pol II is indicated with a black dot.

regulation (Fig. 4). Polymerase pausing involves tilting of the DNA–RNA hybrid[129–131], which impairs nucleotide addition and pause escape[132]. Paused Pol II is stabilized by the factors DSIF and NELF[133]. DSIF binds around exiting DNA and RNA[134,135], whereas NELF binds the opposite side of Pol II, the so-called funnel[129]. NELF impairs binding of TFIIS[136] to the funnel[137], and restricts Pol II mobility[129] to suppress release from pausing[138]. Release of paused Pol II into gene bodies requires the kinase CDK9, a subunit of the positive transcription elongation factor b (P-TEFb)[139]. P-TEFb phosphorylates DSIF, NELF and the Pol II CTD[140], and triggers formation of an activated elongation complex[138]. In the activated elongation complex, the elongation factor SPT6[141] binds the phosphorylated linker to the CTD[138,142], and the PAF complex[143] binds to the funnel and competes with NELF[129,138].

Promoter-proximal pausing can limit the frequency of transcription initiation, and thereby regulate a gene by changing the amount of RNA synthesized per unit of time[144–146]. Transcription factors can target both initiation and the elongation phase[147]. For example, the oncogenic transcription factor MYC can promote release of Pol II from pausing[148]. Factors of the BET family, such as BRD4, can bind enhancers and recruit P-TEFb[149]. P-TEFb can also be recruited as part of the super elongation complex[150] (also known as TATCOM1[151]), which contains fusion partner proteins of the mixed-lineage leukaemia protein. P-TEFb is inhibited when it associates with the non-coding RNA 7SK[152,153].

The phosphorylated CTD recruits many factors that are required for Pol II elongation and for co-transcriptional events such as RNA processing, histone modification and chromatin remodelling[141,154–159]. The initial phosphorylation by CDK7 targets serine-5 residues of the heptapeptide repeats of the CTD, and leads to the recruitment of the capping enzyme; this results in the protection of the nascent RNA 5′ end with a cap structure[155]. Subsequent CTD phosphorylation by the CDK9 subunit of P-TEFb recruits positive elongation factors[141], including the histone methyltransferases SET1[160] and SET2[161]. CDK7 and CDK9 can be inhibited with small molecules (Table 2), and inhibitors of transcription elongation are currently being explored for the treatment of human cancers[162,163].

Transcription condensates

Because of the large number of factors required, the question arises of how transcription is organized in the nucleus—that is, of how the required factors are delivered quickly, and how factors for initiation and elongation are kept separate. It has long been known that Pol I transcription occurs within a membrane-less nuclear compartment (the nucleolus)[164]. On the basis of microscopy, it was suggested that Pol II transcription takes place at nuclear 'hubs'[165] or in static transcription 'factories'[166]. Live-cell super-resolution microscopy has subsequently revealed dynamic foci of Pol II, which have been referred to as hubs, clusters or condensates[167–172]. These foci have been suggested to form by liquid–liquid phase separation of proteins with disordered regions[173]. Condensates of Pol II and Mediator locate to sites of transcription[169], and condensates that contain Mediator and BRD4 have been found at clustered enhancers (known as super-enhancers)[171]. Condensates form around transcription factors[170,174,175], which have long been known to recruit transcription complexes[17–19]. Transcription factors use their disordered transactivation region to form condensates, and attract Pol II[170] and Mediator[174]. These condensates apparently serve to concentrate and deliver proteins for transcription initiation at sites of high transcription activity, and I therefore refer to them as 'promoter condensates'.

Although it is difficult to test whether the phase separation observed in vitro drives factor condensation in vivo, it is highly likely that the same intermolecular forces underlie both phenomena. Phase separation relies on multivalent and cooperative interactions between intrinsically disordered protein regions, and is known to concentrate proteins in cells[176,177]. Transcription factors can undergo phase separation[174,178] and recruit the intrinsically disordered CTD of Pol II[178]. The CTD alone can phase-separate in the presence of a crowding agent[172]. Thus, the CTD is likely to be a client of promoter condensates to enable the recruitment of Pol II to active genes[179].

CTD phosphorylation by CDK7 counteracts CTD self-association and phase separation[172]. The phosphorylated CTD can, however, be

Table 2 | Selected examples of transcription inhibitors

Inhibitor	Target	Mode of action	References
Triptolide	TFIIH subunit XPB	Inhibits promoter DNA opening	205–207
THZ1	TFIIH subunit CDK7	Inhibits CTD phosphorylation and promoter escape	208
Flavopiridol	P-TEFb subunit CDK9 and other kinases	Inhibits Pol II release from pausing and activation	207,209
5,6-Dichlorobenz-imidazole 1-β-D-ribofuranoside (DRB)	P-TEFb subunit CDK9 and other kinases	Inhibits Pol II release from pausing and activation	207,210
α-Amanitin	Pol II	Inhibits Pol II elongation by impairing translocation after nucleotide addition	116
Actinomycin D	DNA	Inhibits transcription elongation and other processes by DNA intercalation	207

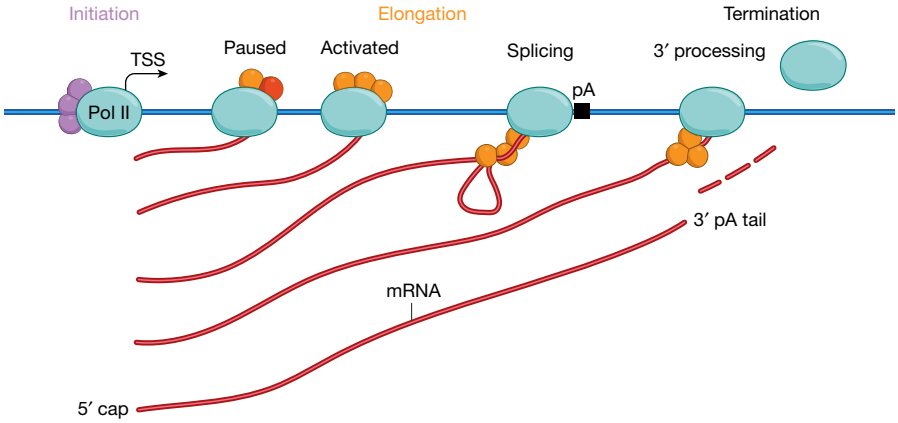


Fig. 3 | Pol II progression through the transcription cycle. Pol II (turquoise) requires different sets of protein factors during initiation (violet) and elongation (red and orange) of the RNA chain (red). The latter include factors that are required for co-transcriptional pre-mRNA processing (in particular, splicing and 3′ processing), and factors for chromatin remodelling and modification. The black arrow indicates the transcription start site (TSS), and the black box indicates the poly-adenylation (pA) site. Pol II moves from left (upstream) to right (downstream). Note that nucleosomes—the building blocks of chromatin—have been omitted for clarity.

incorporated into phase-separated droplets formed by a disordered region in P-TEFb[180]. Thus, the phosphorylated CTD can be a client for a condensate that is distinct from promoter condensates, and is formed by an elongation factor. Because the phosphorylated CTD is a hallmark of the elongation complex, it is likely there are nuclear condensates that contain phosphorylated, transcribing Pol II with nascent RNA; I refer to these as 'gene-body condensates'. Gene-body condensates may correspond to splicing speckles (which are known to contain phosphorylated Pol II)[181–184] and to paraspeckles, which can form around RNA[185]. Nascent RNA is known to bind Pol II elongation factors[186] and RNA processing factors[187]; such RNA–protein interactions may occur in gene-body condensates, because RNA can support phase separation by RNA-binding proteins[188,189] and by RNA–protein complexes[185]. The existence of distinct promoter and gene-body condensates is supported by the observation that differently phosphorylated forms of Pol II occupy similar, but adjacent, nuclear locations[190].

Model for the organization of transcription

From these observations and considerations, a simplified hypothetical model emerges for the organization of Pol II transcription (Fig. 5). In this model, dynamic promoter condensates contain transcription factors, co-activators, unphosphorylated Pol II enzymes and initiation factors. Transient gene-body condensates contain phosphorylated Pol II enzymes, nascent RNA, elongation factors, RNA processing factors and elongation-specific co-activators. Promoter condensates support high rates of initiation, whereas gene-body condensates support RNA elongation and processing within chromatin. These condensates are chemically distinct, which would keep initiation- and elongation-related factors separate.

This model explains the self-organization of Pol II transcription. In the model, transcription factors recruit cofactors and Pol II, and drive the formation of a dynamic promoter condensate. This condensate supports PIC assembly, transcription initiation, RNA synthesis and Pol II phosphorylation. These events, in turn, result in the formation of a transient gene-body condensate that supports elongation and co-transcriptional RNA processing. When the polymerase reaches the end of the gene, polymerase dephosphorylation[191,192] liberates Pol II from the gene-body condensate. The CTD may then be transferred back to the promoter condensate, and Pol II could be recycled upon transcription termination. In this model, the largely unstructured condensates enable rapid formation of highly structured and functional transcription complexes on DNA (promoter condensate) or on nascent RNA (gene-body condensate). Reactions that occur within the structured complexes would thus trigger Pol II shuttling between condensates.

The model helps us to understand how cells can rapidly change their gene-expression program (for example, during differentiation). Promoter condensates are short-lived, dynamic structures that result from self-organization in the nucleus, which in turn depends on the concentration of transcription factors. Condensation by phase separation explains how different transcription factors can recruit the same general Pol II machinery to promoters. Promoter condensates are highly dynamic, and can grow, shrink and reform at alternative sites in the nucleus. Gene-body condensates are formed by transcribing polymerases and are thus a transient, downstream consequence of promoter activity. Condensate formation and dynamics can be controlled by post-translational modifications such as phosphorylation, methylation, acetylation or ubiquitination, which may alter the phase-separation properties of their substrate proteins. It is likely that condensates can be shared by several active genes, and this would explain how a single enhancer can activate two target genes[193]. It is also expected that genes with low transcription activity lack such condensates. It is further likely that additional transcription-related condensates exist (for example, to store transcription complexes).

Future perspectives

Fifty years after the isolation of the three eukaryotic RNA polymerases, much progress has been made in the study of transcription using complementary bottom-up and top-down approaches. Bottom-up studies

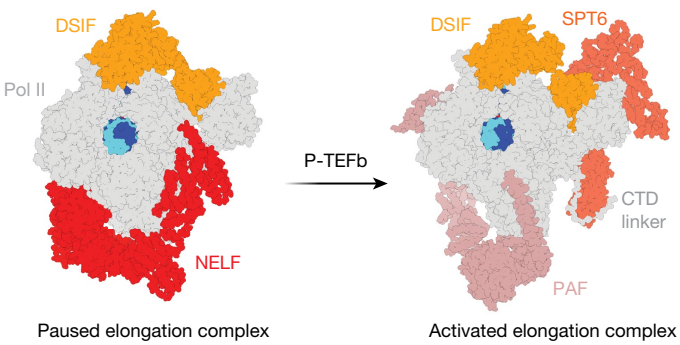


Fig. 4 | Switch from Pol II pausing to active elongation. The structures of the paused transcription elongation complex Pol II–DSIF–NELF[129] and the activated elongation complex Pol II–DSIF–PAF–SPT6 (EC*)[138]. The view is as in Box 1, along downstream DNA and looking into the polymerase cleft. The general elongation factor DSIF (orange) is present in the paused and activated elongation complex, but changes its conformation between the two complexes. The negative elongation factor NELF (red) stabilizes the paused state. In the activated elongation complex, the positive elongation factor SPT6 (light red) binds the phosphorylated linker to the Pol II CTD, and the PAF complex (pale red) binds the polymerase funnel and competes with NELF.

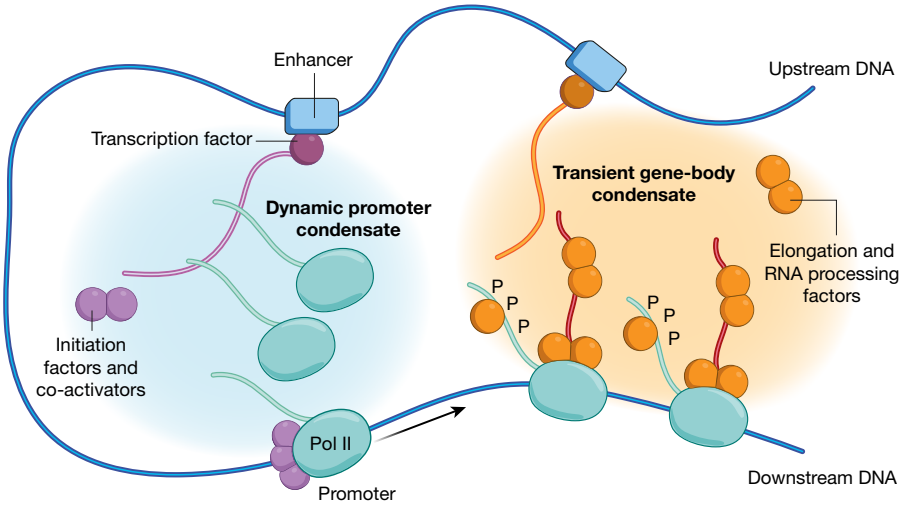


Fig. 5 | Condensate-based model of Pol II transcription. Hypothetical model of an active Pol II gene (blue) with dynamic promoter condensate (turquoise) and transient gene-body condensate (orange). The promoter condensate is established by transcription factors that bind to regulatory elements such as enhancers (blue boxes), and recruits Pol II with an unphosphorylated CTD (turquoise ovals with tail-like extension), co-activators such as Mediator, and initiation factors (violet). The gene-body condensate is formed by nascent RNA (red), elongation and RNA processing factors (orange), and phosphorylated Pol II. Some enhancers and transcription factors may contribute to the formation or stability of gene-body condensates. Pol II is predicted to shuttle between these condensates upon changes in the phosphorylation of its CTD. Condensates may be shared by several genes. Other transcription-related condensates exist.

have reconstituted transcription reactions and complexes for functional and structural analysis. Top-down studies have used microscopy and functional genomics to describe transcription in cells, and genome-wide. The challenge for future studies is to combine both approaches to obtain mechanistic insights that are relevant within the cellular context. These efforts may be guided by considering the model that transcription involves distinct nuclear condensates. We need to investigate the contents, biophysical properties, location and dynamics of nuclear condensates. The transitions of Pol II between condensates—and thus the initiation–elongation and the elongation–termination transitions—need to be characterized on a structural level, and the factors involved should be better-defined. For example, paused Pol II may be a shuttling intermediate between promoter and gene-body condensates, but it is unclear whether paused Pol II is frequently released from DNA (a mechanism that is referred to as attenuation or premature termination)[194]. How transcription is coupled to RNA splicing[181] and 3′ processing, which is coupled to termination[195], also remains to be defined at the mechanistic level.

Finally, we should investigate the mechanisms of transcriptional regulation and organization in the context of chromatin and genome architecture. For example, further study is needed regarding how pioneer transcription factors invade chromatin and how transcription condensates and chromatin influence one another, provided that histones can also undergo phase separation[196]. How Pol II initiation and pausing are influenced by the +1 nucleosome, and which molecular mechanisms position this specialized nucleosome, also require investigation. We should explore how Pol II passes through nucleosomes during elongation[197–199], which involves a mechanism that requires the ATP-dependent chromatin remodeller CHD1[200,201] and the histone chaperone FACT[202,203]. Addressing these questions promises to further unravel the complex mechanisms that underlie transcriptional regulation, and may also unveil the principles that underlie cell differentiation or the growth of cancerous cells.

Note added in proof: After this review had been accepted, a study was published[204] that supports the condensate-based model for the organization of Pol II transcription proposed here. In brief, the unphosphorylated Pol II CTD is incorporated into Mediator condensates, whereas the phosphorylated CTD is incorporated into condensates formed by splicing factors. The authors suggest that CTD phosphorylation drives the exchange of Pol II from condensates for transcription initiation to condensates for splicing, which is consistent with the proposal made here that Pol II shuttles between promoter and gene-body condensates upon its phosphorylation during the transition from transcription initiation to elongation.

Received: 8 April 2019; Accepted: 30 July 2019;
Published online 28 August 2019.

1. Roeder, R. G. & Rutter, W. J. Multiple forms of DNA-dependent RNA polymerase in eukaryotic organisms. *Nature* **224**, 234–237 (1969). **Fifty years ago, three RNA polymerases were isolated from nuclei of eukaryotic cells**.
2. Sentenac, A. Eukaryotic RNA polymerases. *CRC Crit. Rev. Biochem*. **18**, 31–90 (1985).
3. Fuda, N. J., Ardehali, M. B. & Lis, J. T. Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**, 186–192 (2009).
4. Lorch, Y. & Kornberg, R. D. Chromatin-remodeling for transcription. *Q. Rev. Biophys*. **50**, e5 (2017).
5. Knezetic, J. A. & Luse, D. S. The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II *in vitro*. *Cell* **45**, 95–104 (1986).
6. Lorch, Y., LaPointe, J. W. & Kornberg, R. D. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell* **49**, 203–210 (1987).
7. Talbert, P. B., Meers, M. P. & Henikoff, S. Old cogs, new tricks: the evolution of gene expression in a chromatin context. *Nat. Rev. Genet*. **20**, 283–297 (2019).
8. Schones, D. E. et al. Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898 (2008).
9. Müller, F. & Tora, L. Chromatin and DNA sequences in defining promoters for transcription initiation. *Biochim. Biophys. Acta* **1839**, 118–128 (2014).
10. Vo ngoc, L., Wang, Y. L., Kassavetis, G. A. & Kadonaga, J. T. The punctilious RNA polymerase II core promoter. *Genes Dev*. **31**, 1289–1301 (2017).
11. Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev*. **25**, 1010–1022 (2011).
12. Schübeler, D. Function and information content of DNA methylation. *Nature* **517**, 321–326 (2015).
13. Dynan, W. S. & Tjian, R. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**, 79–87 (1983). **Evidence is provided that a DNA sequence-specific transcription factor can guide Pol II to a target promoter**.
14. Engelke, D. R., Ng, S. Y., Shastry, B. S. & Roeder, R. G. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell* **19**, 717–728 (1980). **Evidence is provided that a DNA sequence-specific transcription factor can guide Pol III to a target promoter**.
15. Payvar, F. et al. Purified glucocorticoid receptors bind selectively *in vitro* to a cloned DNA fragment whose transcription is regulated by glucocorticoids *in vivo*. *Proc. Natl Acad. Sci. USA* **78**, 6628–6632 (1981). **A hormone-sensitive DNA sequence-specific transcription factor can bind near its Pol II target promoter**.
16. Mulvihill, E. R., LePennec, J. P. & Chambon, P. Chicken oviduct progesterone receptor: location of specific regions of high-affinity binding in cloned DNA fragments of hormone-responsive genes. *Cell* **28**, 621–632 (1982).

17. Ptashne, M. & Gann, A. Transcriptional activation by recruitment. *Nature* **386**, 569–577 (1997).
18. Kadonaga, J. T., Courey, A. J., Ladika, J. & Tjian, R. Distinct regions of Sp1 modulate DNA binding and transcriptional activation. *Science* **242**, 1566–1570 (1988).
A transcription factor contains separate DNA binding and transactivation regions.
19. Sigler, P. B. Acid blobs and negative noodles. *Nature* **333**, 210–212 (1988).
20. Fong, Y. W., Cattoglio, C., Yamaguchi, T. & Tjian, R. Transcriptional regulation by coactivators in embryonic stem cells. *Trends Cell Biol*. **22**, 292–298 (2012).
21. Lambert, S. A. et al. The human transcription factors. *Cell* **172**, 650–665 (2018).
An inventory of human transcription factors is provided.
22. Zhu, F. et al. The interaction landscape between transcription factors and the nucleosome. *Nature* **562**, 76–81 (2018).
23. Iwafuchi-Doi, M. & Zaret, K. S. Cell fate control by pioneer transcription factors. *Development* **143**, 1833–1837 (2016).
24. Brownell, J. E. et al. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843–851 (1996).
25. Utley, R. T. et al. Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* **394**, 498–502 (1998).
26. Kraus, W. L. & Kadonaga, J. T. p300 and estrogen receptor cooperatively activate transcription via differential enhancement of initiation and reinitiation. *Genes Dev*. **12**, 331–342 (1998).
27. An, W., Palhan, V. B., Karymov, M. A., Leuba, S. H. & Roeder, R. G. Selective requirements for histone H3 and H4 N termini in p300-dependent transcriptional activation from chromatin. *Mol. Cell* **9**, 811–821 (2002).
28. Banerji, J., Rusconi, S. & Schaffner, W. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **27**, 299–308 (1981).
29. Benoist, C. & Chambon, P. *In vivo* sequence requirements of the SV40 early promotor region. *Nature* **290**, 304–310 (1981).
30. Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).
31. Reiter, F., Wienerroither, S. & Stark, A. Combinatorial function of transcription factors and cofactors. *Curr. Opin. Genet. Dev*. **43**, 73–81 (2017).
32. Core, L. J. et al. Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet*. **46**, 1311–1320 (2014).
33. Neil, H. et al. Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature* **457**, 1038–1042 (2009).
Most gene promoters in yeast give rise to bidirectional RNA synthesis.
34. Robson, M. I., Ringel, A. R. & Mundlos, S. Regulatory landscaping: how enhancer–promoter communication is sculpted in 3D. *Mol. Cell* **74**, 1110–1122 (2019).
35. van Steensel, B. & Furlong, E. E. M. The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol*. **20**, 327–337 (2019).
36. Nora, E. P. et al. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**, 381–385 (2012).
37. Sharifi, S. & Bierhoff, H. Regulation of RNA polymerase I transcription in development, disease, and aging. *Annu. Rev. Biochem*. **87**, 51–73 (2018).
38. Haberle, V. & Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol*. **19**, 621–637 (2018).
39. Dergai, O. & Hernandez, N. How to recruit the correct RNA polymerase? Lessons from snRNA genes. *Trends Genet*. **35**, 457–469 (2019).
40. Reinberg, D. et al. The RNA polymerase II general transcription factors: past, present, and future. *Cold Spring Harb. Symp. Quant. Biol*. **63**, 83–105 (1998).
41. Grummt, I. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes Dev*. **17**, 1691–1702 (2003
42. Sentenac, A. & Riva, M. Odd RNA polymerases or the A(B)C of eukaryotic transcription. *Biochim. Biophys. Acta* **1829**, 251–257 (2013).
43. Schramm, L. & Hernandez, N. Recruitment of RNA polymerase III to its target promoters. *Genes Dev*. **16**, 2593–2620 (2002).
44. Geiduschek, E. P. & Kassavetis, G. A. The RNA polymerase III transcription apparatus. *J. Mol. Biol*. **310**, 1–26 (2001).
45. Engel, C. et al. Structural basis of RNA polymerase I transcription initiation. *Cell* **169**, 120–131.e122 (2017).
This paper presents the structure of a Pol I pre-initiation complex.
46. Sadian, Y. et al. Structural insights into transcription initiation by yeast RNA polymerase I. *EMBO J*. **36**, 2698–2709 (2017).
This paper presents the structure of a Pol I pre-initiation complex.
47. Han, Y. et al. Structural mechanism of ATP-independent transcription initiation by RNA polymerase I. *eLife* **6**, e27414 (2017).
This paper presents the structure of a Pol I pre-initiation complex.
48. Schilbach, S. et al. Structures of transcription pre-initiation complex with TFIIH and Mediator. *Nature* **551**, 204–209 (2017).
This paper presents the structure of a Pol II pre-initiation complex containing TFIIH and core Mediator.
49. Plaschka, C. et al. Transcription initiation complex structures elucidate DNA opening. *Nature* **533**, 353–358 (2016).
50. Plaschka, C. et al. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **518**, 376–380 (2015).
The three-dimensional architecture of a Pol II pre-initiation complex containing core Mediator is derived.
51. Kostrewa, D. et al. RNA polymerase II–TFIIB structure and mechanism of transcription initiation. *Nature* **462**, 323–330 (2009).
52. Mühlbacher, W. et al. Conserved architecture of the core RNA polymerase II initiation complex. *Nat. Commun*. **5**, 4310 (2014).
53. Louder, R. K. et al. Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature* **531**, 604–609 (2016).
54. He, Y., Fang, J., Taatjes, D. J. & Nogales, E. Structural visualization of key steps in human transcription initiation. *Nature* **495**, 481–486 (2013).
This paper describes the architecture of a Pol II pre-initiation complex containing TFIIH.
55. He, Y. et al. Near-atomic resolution visualization of human transcription promoter opening. *Nature* **533**, 359–365 (2016).
56. Robinson, P. J. et al. Structure of a complete Mediator–RNA polymerase II pre-initiation complex. *Cell* **166**, 1411–1422.e1416 (2016).
This paper describes the overall topology of a Pol II pre-initiation complex containing TFIIH and Mediator.
57. Liu, X., Bushnell, D. A., Wang, D., Calero, G. & Kornberg, R. D. Structure of an RNA polymerase II–TFIIB complex and the transcription initiation mechanism. *Science* **327**, 206–209 (2010).
58. Vorländer, M. K., Khatter, H., Wetzel, R., Hagen, W. J. H. & Müller, C. W. Molecular mechanism of promoter opening by RNA polymerase III. *Nature* **553**, 295–300 (2018).
The structure of a Pol III pre-initiation complex is described.
59. Abascal-Palacios, G., Ramsay, E. P., Beuron, F., Morris, E. & Vannini, A. Structural basis of RNA polymerase III transcription initiation. *Nature* **553**, 301–306 (2018).
The structure of a Pol III pre-initiation complex is described.
60. Kornberg, R. D. Eukaryotic transcriptional control. *Trends Cell Biol*. **9**, M46–M49 (1999).
61. Roeder, R. G. The role of general initiation factors in transcription by RNA polymerase II. *Trends Biochem. Sci*. **21**, 327–335 (1996).
62. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. Five intermediate complexes in transcription initiation by RNA polymerase II. *Cell* **56**, 549–561 (1989).
63. Chen, H. T. & Hahn, S. Mapping the location of TFIIB within the RNA polymerase II transcription preinitiation complex: a model for the structure of the PIC. *Cell* **119**, 169–180 (2004).
64. Bushnell, D. A., Westover, K. D., Davis, R. E. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II–TFIIB cocrystal at 4.5 angstroms. *Science* **303**, 983–988 (2004).
65. Sainsbury, S., Niesser, J. & Cramer, P. Structure and function of the initially transcribing RNA polymerase II–TFIIB complex. *Nature* **493**, 437–440 (2013).
66. Knutson, B. A. & Hahn, S. Yeast Rrn7 and human TAF1B are TFIIB-related RNA polymerase I general transcription factors. *Science* **333**, 1637–1640 (2011).
67. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* **45**, 439–446 (2012).
68. Vermeulen, M. et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 (2007).
69. D'Alessio, J. A., Wright, K. J. & Tjian, R. Shifting players and paradigms in cell-specific transcription. *Mol. Cell* **36**, 924–931 (2009).
70. Levens, D., Baranello, L. & Kouzine, F. Controlling gene expression by DNA mechanics: emerging insights and challenges. *Biophys. Rev*. **8**, 259–268 (2016).
71. Pugh, B. F. & Venters, B. J. Genomic organization of human transcription initiation complexes. *PLoS ONE* **11**, e0149339 (2016).
72. Andersen, P. R., Tirian, L., Vunjak, M. & Brennecke, J. A heterochromatin-dependent transcription machinery drives piRNA expression. *Nature* **549**, 54–59 (2017).
73. Kassavetis, G. A., Blanco, J. A., Johnson, T. E. & Geiduschek, E. P. Formation of open and elongating transcription complexes by RNA polymerase III. *J. Mol. Biol*. **226**, 47–58 (1992).
74. Kato, H., Nagamine, M., Kominami, R. & Muramatsu, M. Formation of the transcription initiation complex on mammalian rDNA. *Mol. Cell. Biol*. **6**, 3418–3427 (1986).
75. Logquist, A. K., Li, H., Imboden, M. A. & Paule, M. R. Promoter opening (melting) and transcription initiation by RNA polymerase I requires neither nucleotide β,γ hydrolysis nor protein phosphorylation. *Nucleic Acids Res*. **21**, 3233–3238 (1993).
76. Gokal, P. K., Mahajan, P. B. & Thompson, E. A. Hormonal regulation of transcription of rDNA. Formation of initiated complexes by RNA polymerase I *in vitro*. *J. Biol. Chem*. **265**, 16234–16243 (1990).
77. Schnapp, A. & Grummt, I. Transcription complex formation at the mouse rDNA promoter involves the stepwise association of four transcription factors and RNA polymerase I. *J. Biol. Chem*. **266**, 24588–24595 (1991).
78. Feklistov, A. & Darst, S. A. Structural basis for promoter-10 element recognition by the bacterial RNA polymerase σ subunit. *Cell* **147**, 1257–1269 (2011).
79. Zuo, Y. & Steitz, T. A. Crystal structures of the *E. coli* transcription initiation complexes with a complete bubble. *Mol. Cell* **58**, 534–540 (2015).
80. Posse, V. & Gustafsson, C. M. Human mitochondrial transcription factor B2 is required for promoter melting during initiation of transcription. *J. Biol. Chem*. **292**, 2637–2645 (2017).
81. Hillen, H. S., Morozov, Y. I., Sarfallah, A., Temiakov, D. & Cramer, P. Structural basis of mitochondrial transcription initiation. *Cell* **171**, 1072–1081.e1010 (2017).
82. Egly, J. M. & Coin, F. A history of TFIIH: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair (Amst.)* **10**, 714–721 (2011).
83. Kim, T. K., Ebright, R. H. & Reinberg, D. Mechanism of ATP-dependent promoter melting by transcription factor IIH. *Science* **288**, 1418–1421 (2000).
Crosslinking shows that TFIIH acts on downstream DNA to open the promoter.

84. Holstege, F. C., van der Vliet, P. C. & Timmers, H. T. Opening of an RNA polymerase II promoter occurs in two distinct steps and requires the basal transcription factors IIE and IIH. *EMBO J.* **15**, 1666–1677 (1996).
85. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16**, 129–143 (2015).
86. Grünberg, S., Warfield, L. & Hahn, S. Architecture of the RNA polymerase II preinitiation complex and mechanism of ATP-dependent promoter opening. *Nat. Struct. Mol. Biol.* **19**, 788–796 (2012).
TFIIH is found to contain a translocase that propels downstream DNA into the Pol II active centre.
87. Kouzine, F. et al. Global regulation of promoter melting in naive lymphocytes. *Cell* **153**, 988–999 (2013).
Promoter DNA opening is a regulated event in cells.
88. Dienemann, C., Schwalb, B., Schilbach, S. & Cramer, P. Promoter distortion and opening in the RNA polymerase II cleft. *Mol. Cell* **73**, 97–106.e104 (2019).
89. Alekseev, S. et al. Transcription without XPB establishes a unified helicase-independent mechanism of promoter opening in eukaryotic gene expression. *Mol. Cell* **65**, 504–514.e4 (2017).
90. Pilsl, M. et al. Structure of the initiation-competent RNA polymerase I and its implication for transcription. *Nat. Commun.* **7**, 12126 (2016).
91. Blattner, C. et al. Molecular basis of Rrn3-regulated RNA polymerase I initiation and cell growth. *Genes Dev.* **25**, 2093–2105 (2011).
92. Milkereit, P. & Tschochner, H. A specialized form of RNA polymerase I, essential for initiation and growth-dependent regulation of rRNA synthesis, is disrupted during transcription. *EMBO J.* **17**, 3692–3703 (1998).
93. Yuan, X., Zhao, J., Zentgraf, H., Hoffmann-Rohrer, U. & Grummt, I. Multiple interactions between RNA polymerase I, TIF-IA and TAF_I subunits regulate preinitiation complex assembly at the ribosomal gene promoter. *EMBO Rep.* **3**, 1082–1087 (2002).
94. Moir, R. D. & Willis, I. M. Regulation of pol III transcription by nutrient and stress signaling pathways. *Biochim. Biophys. Acta* **1829**, 361–375 (2013).
95. Pluta, K. et al. Maf1p, a negative effector of RNA polymerase III in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **21**, 5031–5040 (2001).
96. White, R. J. RNA polymerases I and III, non-coding RNAs and cancer. *Trends Genet.* **24**, 622–629 (2008).
97. Kornberg, R. D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30**, 235–239 (2005).
98. Wong, K. H., Jin, Y. & Struhl, K. TFIIH phosphorylation of the Pol II CTD stimulates mediator dissociation from the preinitiation complex and promoter escape. *Mol. Cell* **54**, 601–612 (2014).
99. Jeronimo, C. & Robert, F. Kin28 regulates the transient association of Mediator with core promoters. *Nat. Struct. Mol. Biol.* **21**, 449–455 (2014).
100. Tsai, K. L. et al. Mediator structure and rearrangements required for holoenzyme formation. *Nature* **544**, 196–201 (2017).
101. Nozawa, K., Schneider, T. R. & Cramer, P. Core Mediator structure at 3.4 Å extends model of transcription initiation complex. *Nature* **545**, 248–251 (2017).
This paper presents the crystal structure of the core Mediator coactivator complex.
102. Taatjes, D. J. Transcription factor–mediator interfaces: multiple and multi-valent. *J. Mol. Biol.* **429**, 2996–2998 (2017).
103. Jeronimo, C. & Robert, F. The mediator complex: at the nexus of RNA Polymerase II transcription. *Trends Cell Biol.* **27**, 765–783 (2017).
104. Eick, D. & Geyer, M. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chem. Rev.* **113**, 8456–8490 (2013).
105. Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* **292**, 1876–1882 (2001).
106. Nudler, E. Transcription elongation: structural basis and mechanisms. *J. Mol. Biol.* **288**, 1–12 (1999).
107. Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirov, T. H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157–162 (2007).
108. Schwinghammer, K. et al. Structure of human mitochondrial RNA polymerase elongation complex. *Nat. Struct. Mol. Biol.* **20**, 1298–1303 (2013).
109. Neyer, S. et al. Structure of RNA polymerase I transcribing ribosomal DNA genes. *Nature* **540**, 607–610 (2016).
110. Hoffmann, N. A. et al. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* **528**, 231–236 (2015).
111. Sidorenkov, I., Komissarova, N. & Kashlev, M. Crucial role of the RNA:DNA hybrid in the processivity of transcription. *Mol. Cell* **2**, 55–64 (1998).
112. Vassylyev, D. G. et al. Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163–168 (2007).
113. Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D. & Kornberg, R. D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941–954 (2006).
114. Steitz, T. A. & Steitz, J. A. A general two-metal-ion mechanism for catalytic RNA. *Proc. Natl Acad. Sci. USA* **90**, 6498–6502 (1993).
115. Cramer, P., Bushnell, D. A. & Kornberg, R. D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**, 1863–1876 (2001).
116. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by α-amanitin and implications for RNA polymerase II translocation. *Nat. Struct. Mol. Biol.* **15**, 811–818 (2008).
117. Landick, R. The regulatory roles and mechanism of transcriptional pausing. *Biochem. Soc. Trans.* **34**, 1062–1066 (2006).
118. Conaway, J. W., Shilatifard, A., Dvir, A. & Conaway, R. C. Control of elongation by RNA polymerase II. *Trends Biochem. Sci.* **25**, 375–380 (2000).
119. Cheung, A. C. & Cramer, P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**, 249–253 (2011).
120. Kuhn, C. D. et al. Functional architecture of RNA polymerase I. *Cell* **131**, 1260–1272 (2007).
121. Chédin, S., Riva, M., Schultz, P., Sentenac, A. & Carles, C. The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination. *Genes Dev.* **12**, 3857–3871 (1998).
122. Bentley, D. L. & Groudine, M. A block to elongation is largely responsible for decreased transcription of c-*myc* in differentiated HL60 cells. *Nature* **321**, 702–706 (1986).
123. Eick, D. & Bornkamm, G. W. Transcriptional arrest within the first exon is a fast control mechanism in c-myc gene expression. *Nucleic Acids Res.* **14**, 8331–8346 (1986).
124. Gariglio, P., Bellard, M. & Chambon, P. Clustering of RNA polymerase B molecules in the 5′ moiety of the adult β-globin gene of hen erythrocytes. *Nucleic Acids Res.* **9**, 2589–2598 (1981).
125. 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).
126. 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).
127. Tome, J. M., Tippens, N. D. & Lis, J. T. Single-molecule nascent RNA sequencing identifies regulatory domain architecture at promoters and enhancers. *Nat. Genet.* **50**, 1533–1541 (2018).
128. Core, L. & Adelman, K. Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* https://doi.org/10.1101/gad.325142.119 (2019).
129. Vos, S. M., Farnung, L., Urlaub, H. & Cramer, P. Structure of paused transcription complex Pol II–DSIF–NELF. *Nature* **560**, 601–606 (2018).
130. Kang, J. Y. et al. Structural basis for transcript elongation control by NusG family universal regulators. *Cell* **173**, 1650–1662.e1614 (2018).
131. Guo, X. et al. Structural basis for NusA stabilized transcriptional pausing. *Mol. Cell* **69**, 816–827.e814 (2018).
132. Saba, J. et al. The elemental mechanism of transcriptional pausing. *eLife* **8**, e40981 (2019).
133. Yamaguchi, Y., Shibata, H. & Handa, H. Transcription elongation factors DSIF and NELF: promoter-proximal pausing and beyond. *Biochim. Biophys. Acta* **1829**, 98–104 (2013).
134. Bernecky, C., Plitzko, J. M. & Cramer, P. Structure of a transcribing RNA polymerase II–DSIF complex reveals a multidentate DNA–RNA clamp. *Nat. Struct. Mol. Biol.* **24**, 809–815 (2017).
135. Ehara, H. et al. Structure of the complete elongation complex of RNA polymerase II with basal factors. *Science* **357**, 921–924 (2017).
136. Palangat, M., Renner, D. B., Price, D. H. & Landick, R. A negative elongation factor for human RNA polymerase II inhibits the anti-arrest transcript-cleavage factor TFIIS. *Proc. Natl Acad. Sci. USA* **102**, 15036–15041 (2005).
137. Kettenberger, H., Armache, K. J. & Cramer, P. Architecture of the RNA polymerase II–TFIIS complex and implications for mRNA cleavage. *Cell* **114**, 347–357 (2003).
138. Vos, S. M. et al. Structure of activated transcription complex Pol II–DSIF–PAF–SPT6. *Nature* **560**, 607–612 (2018).
The structure of a mammalian, activated Pol II elongation complex provides a model for polymerase release from promoter-proximal pausing.
139. 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).
140. Zhou, Q., Li, T. & Price, D. H. RNA polymerase II elongation control. *Annu. Rev. Biochem.* **81**, 119–143 (2012).
141. Kwak, H. & Lis, J. T. Control of transcriptional elongation. *Annu. Rev. Genet.* **47**, 483–508 (2013).
142. Sdano, M. A. et al. A novel SH2 recognition mechanism recruits Spt6 to the doubly phosphorylated RNA polymerase II linker at sites of transcription. *eLife* **6**, e28723 (2017).
143. Van Oss, S. B., Cucinotta, C. E. & Arndt, K. M. Emerging insights into the roles of the Paf1 complex in gene regulation. *Trends Biochem. Sci.* **42**, 788–798 (2017).
144. Shao, W. & Zeitlinger, J. Paused RNA polymerase II inhibits new transcriptional initiation. *Nat. Genet.* **49**, 1045–1051 (2017).
Evidence is presented that promoter-proximal pausing can regulate transcription by suppressing initiation.
145. Gressel, S. et al. CDK9-dependent RNA polymerase II pausing controls transcription initiation. *eLife* **6**, e29736 (2017).
Evidence is presented that promoter-proximal pausing can regulate transcription initiation.
146. Ehrensberger, A. H., Kelly, G. P. & Svejstrup, J. Q. Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. *Cell* **154**, 713–715 (2013).
147. Brown, S. A., Weirich, C. S., Newton, E. M. & Kingston, R. E. Transcriptional activation domains stimulate initiation and elongation at different times and via different residues. *EMBO J.* **17**, 3146–3154 (1998).
148. Rahl, P. B. et al. c-Myc regulates transcriptional pause release. *Cell* **141**, 432–445 (2010).
A transcription factor can regulate transcription elongation.
149. Li, Y., Liu, M., Chen, L. F. & Chen, R. P-TEFb: Finding its ways to release promoter-proximally paused RNA polymerase II. *Transcription* **9**, 88–94 (2018).

150. Smith, E., Lin, C. & Shilatifard, A. The super elongation complex (SEC) and MLL in development and disease. *Genes Dev.* **25**, 661–672 (2011).
151. 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).
152. Yang, Z., Zhu, Q., Luo, K. & Zhou, Q. The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature* **414**, 317–322 (2001).
153. Nguyen, V. T., Kiss, T., Michels, A. A. & Bensaude, O. 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature* **414**, 322–325 (2001).
154. Buratowski, S. Progression through the RNA polymerase II CTD cycle. *Mol. Cell* **36**, 541–546 (2009).
155. Bentley, D. L. Coupling mRNA processing with transcription in time and space. *Nat. Rev. Genet.* **15**, 163–175 (2014).
156. Shilatifard, A., Conaway, R. C. & Conaway, J. W. The RNA polymerase II elongation complex. *Annu. Rev. Biochem.* **72**, 693–715 (2003).
157. Becker, P. B. & Workman, J. L. Nucleosome remodeling and epigenetics. *Cold Spring Harb. Perspect. Biol.* **5**, a017905 (2013).
158. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18**, 407–422 (2017).
159. Chen, F. X., Smith, E. R. & Shilatifard, A. Born to run: control of transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **19**, 464–478 (2018).
160. Shilatifard, A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* **81**, 65–95 (2012).
161. McDaniel, S. L. & Strahl, B. D. Shaping the cellular landscape with Set2/SETD2 methylation. *Cell. Mol. Life Sci.* **74**, 3317–3334 (2017).
162. French, C. A. Small-molecule targeting of BET proteins in cancer. *Adv. Cancer Res.* **131**, 21–58 (2016).
163. Bradner, J. E., Hnisz, D. & Young, R. A. Transcriptional addiction in cancer. *Cell* **168**, 629–643 (2017).
164. Ditlev, J. A., Case, L. B. & Rosen, M. K. Who's in and who's out–compositional control of biomolecular condensates. *J. Mol. Biol.* **430**, 4666–4684 (2018).
165. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. & de Laat, W. Looping and interaction between hypersensitive sites in the active β-globin locus. *Mol. Cell* **10**, 1453–1465 (2002).
166. Papantonis, A. & Cook, P. R. Transcription factories: genome organization and gene regulation. *Chem. Rev.* **113**, 8683–8705 (2013).
167. Cisse, I. I. et al. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* **341**, 664–667 (2013).
Live-cell imaging visualizes Pol II clusters and their dynamics in human nuclei.
168. Buckley, M. S. & Lis, J. T. Imaging RNA polymerase II transcription sites in living cells. *Curr. Opin. Genet. Dev.* **25**, 126–130 (2014).
169. Cho, W. K. et al. Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* **361**, 412–415 (2018).
Imaging reveals nuclear condensates for Pol II transcription.
170. Chong, S. et al. Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* **361**, eaar2555 (2018).
Imaging reveals nuclear condensates for Pol II transcription.
171. Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).
Imaging reveals nuclear condensates for Pol II transcription.
172. Boehning, M. et al. RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* **25**, 833–840 (2018).
Imaging reveals nuclear condensates for Pol II transcription.
173. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. A phase separation model for transcriptional control. *Cell* **169**, 13–23 (2017).
This article presents the hypothesis that transcription involves phase-separated nuclear condensates.
174. Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855.e1816 (2018).
175. Nair, S. J. et al. Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. *Nat. Struct. Mol. Biol.* **26**, 193–203 (2019).
176. Kato, M. & McKnight, S. L. A solid-state conceptualization of information transfer from gene to message to protein. *Annu. Rev. Biochem.* **87**, 351–390 (2018).
177. Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).
178. Kwon, I. et al. Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell* **155**, 1049–1060 (2013).
179. Lu, F., Portz, B. & Gilmour, D. S. The C-terminal domain of RNA polymerase II is a multivalent targeting sequence that supports *Drosophila* development with only consensus heptads. *Mol. Cell* **73**, 1232–1242.e1234 (2019).
180. Lu, H. et al. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**, 318–323 (2018).
181. Herzel, L., Ottoz, D. S. M., Alpert, T. & Neugebauer, K. M. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat. Rev. Mol. Cell Biol.* **18**, 637–650 (2017).
182. Bregman, D. B., Du, L., van der Zee, S. & Warren, S. L. Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J. Cell Biol.* **129**, 287–298 (1995).
183. Mortillaro, M. J. et al. A hyperphosphorylated form of the large subunit of RNA polymerase II is associated with splicing complexes and the nuclear matrix. *Proc. Natl Acad. Sci. USA* **93**, 8253–8257 (1996).
184. Misteli, T. & Spector, D. L. RNA polymerase II targets pre-mRNA splicing factors to transcription sites *in vivo*. *Mol. Cell* **3**, 697–705 (1999).
185. Van Treeck, B. & Parker, R. Emerging roles for intermolecular RNA–RNA interactions in RNP assemblies. *Cell* **174**, 791–802 (2018).
186. Battaglia, S. et al. RNA-dependent chromatin association of transcription elongation factors and Pol II CTD kinases. *eLife* **6**, e25637 (2017).
187. Lewis, J. D. & Tollervey, D. Like attracts like: getting RNA processing together in the nucleus. *Science* **288**, 1385–1389 (2000).
188. Molliex, A. et al. Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell* **163**, 123–133 (2015).
189. Castello, A. et al. Comprehensive identification of RNA-binding domains in human cells. *Mol. Cell* **63**, 696–710 (2016).
190. Ghamari, A. et al. *In vivo* live imaging of RNA polymerase II transcription factories in primary cells. *Genes Dev.* **27**, 767–777 (2013).
191. Proudfoot, N. J. Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. *Science* **352**, aad9926 (2016).
192. Parua, P. K. et al. A Cdk9–PP1 switch regulates the elongation–termination transition of RNA polymerase II. *Nature* **558**, 460–464 (2018).
193. Fukaya, T., Lim, B. & Levine, M. Enhancer control of transcriptional bursting. *Cell* **166**, 358–368 (2016).
An enhancer is shown to be able to activate two target genes.
194. Kamieniarz-Gdula, K. & Proudfoot, N. J. Transcriptional control by premature termination: a forgotten mechanism. *Trends Genet.* **35**, 553–564 (2019).
195. Porrua, O., Boudvillain, M. & Libri, D. Transcription termination: variations on common themes. *Trends Genet.* **32**, 508–522 (2016).
196. Gibson, B. A. et al. Organization and regulation of chromatin by liquid–liquid phase separation. Preprint at https://www.biorxiv.org/content/10.1101/523662v1 (2019).
Histones are shown to undergo phase separation.
197. Farnung, L., Vos, S. M. & Cramer, P. Structure of transcribing RNA polymerase II–nucleosome complex. *Nat. Commun.* **9**, 5432 (2018).
Cryo-electron microscopy provides the structure of a Pol II–nucleosome complex.
198. Ehara, H. et al. Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* **363**, 744–747 (2019).
199. Kujirai, T. et al. Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* **362**, 595–598 (2018).
Cryo-electron microscopy provides the structures of several Pol II–nucleosome complexes.
200. Skene, P. J., Hernandez, A. E., Groudine, M. & Henikoff, S. The nucleosomal barrier to promoter escape by RNA polymerase II is overcome by the chromatin remodeler Chd1. *eLife* **3**, e02042 (2014).
201. Smolle, M. et al. Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nat. Struct. Mol. Biol.* **19**, 884–892 (2012).
202. Hsieh, F. K. et al. Histone chaperone FACT action during transcription through chromatin by RNA polymerase II. *Proc. Natl Acad. Sci. USA* **110**, 7654–7659 (2013).
203. Orphanides, G., LeRoy, G., Chang, C. H., Luse, D. S. & Reinberg, D. FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92**, 105–116 (1998).
204. Guo, Y. E. et al. Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* **572**, 543–548 (2019).
205. Chen, F., Gao, X. & Shilatifard, A. Stably paused genes revealed through inhibition of transcription initiation by the TFIIH inhibitor triptolide. *Genes Dev.* **29**, 39–47 (2015).
206. Titov, D. V. et al. XPB, a subunit of TFIIH, is a target of the natural product triptolide. *Nat. Chem. Biol.* **7**, 182–188 (2011).
207. Bensaude, O. Inhibiting eukaryotic transcription: which compound to choose? How to evaluate its activity? *Transcription* **2**, 103–108 (2011).
208. Kwiatkowski, N. et al. Targeting transcription regulation in cancer with a covalent CDK7 inhibitor. *Nature* **511**, 616–620 (2014).
209. Chao, S. H. et al. Flavopiridol inhibits P-TEFb and blocks HIV-1 replication. *J. Biol. Chem.* **275**, 28345–28348 (2000).
210. Zhu, Y. et al. Transcription elongation factor P-TEFb is required for HIV-1 tat transactivation *in vitro*. *Genes Dev.* **11**, 2622–2632 (1997).
211. Jeruzalmi, D. & Steitz, T. A. Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme. *EMBO J.* **17**, 4101–4113 (1998).
212. Hillen, H. S., Temiakov, D. & Cramer, P. Structural basis of mitochondrial transcription. *Nat. Struct. Mol. Biol.* **25**, 754–765 (2018).
213. Zhang, G. et al. Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* **98**, 811–824 (1999).
214. Nudler, E. RNA polymerase active center: the molecular engine of transcription. *Annu. Rev. Biochem.* **78**, 335–361 (2009).
215. Ray-Soni, A., Bellecourt, M. J. & Landick, R. Mechanisms of bacterial transcription termination: all good things must end. *Annu. Rev. Biochem.* **85**, 319–347 (2016).
216. Feng, Y., Zhang, Y. & Ebright, R. H. Structural basis of transcription activation. *Science* **352**, 1330–1333 (2016).
217. Martinez-Rucobo, F. W. & Cramer, P. Structural basis of transcription elongation. *Biochim. Biophys. Acta* **1829**, 9–19 (2013).

218. Nogales, E., Patel, A. B. & Louder, R. K. Towards a mechanistic understanding of core promoter recognition from cryo-EM studies of human TFIID. *Curr. Opin. Struct. Biol.* **47**, 60–66 (2017).
219. Khatter, H., Vorländer, M. K. & Müller, C. W. RNA polymerase I and III: similar yet unique. *Curr. Opin. Struct. Biol.* **47**, 88–94 (2017).
220. Kornberg, R. D. The molecular basis of eukaryotic transcription. *Proc. Natl Acad. Sci. USA* **104**, 12955–12961 (2007).
221. Engel, C., Neyer, S. & Cramer, P. distinct mechanisms of transcription initiation by RNA polymerases I and II. *Annu. Rev. Biophys.* **47**, 425–446 (2018).
222. Bieniossek, C. et al. The architecture of human general transcription factor TFIID core complex. *Nature* **493**, 699–702 (2013).
223. Cramer, P. et al. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**, 640–649 (2000).
224. Engel, C., Sainsbury, S., Cheung, A. C., Kostrewa, D. & Cramer, P. RNA polymerase I structure and transcription regulation. *Nature* **502**, 650–655 (2013).
225. Fernández-Tornero, C. et al. Crystal structure of the 14-subunit RNA polymerase I. *Nature* **502**, 644–649 (2013).
226. Jasiak, A. J., Armache, K. J., Martens, B., Jansen, R. P. & Cramer, P. Structural biology of RNA polymerase III: subcomplex C17/25 X-ray structure and 11 subunit enzyme model. *Mol. Cell* **23**, 71–81 (2006).
227. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **9**, 85–98 (2011).
228. Haag, J. R. & Pikaard, C. S. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. *Nat. Rev. Mol. Cell Biol.* **12**, 483–492 (2011).

Acknowledgements I would like to thank past and present members of the laboratory. I apologize to those colleagues whose work could not be cited owing to space restraints. I am supported by the Deutsche Forschungsgemeinschaft (SFB860, SPP1935 and EXC 2067/1- 390729940), the European Research Council (Advanced Investigator Grant TRANSREGULON, grant agreement No 693023), and the Volkswagen Foundation.

Competing interests The author declares no competing interests.

Additional information
Reprints and permissions information is available at http://www.nature.com/reprints.
Correspondence and requests for materials should be addressed to P.C.
Peer review information *Nature* thanks Dylan J. Taatjes and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.