

Transcriptional regulation II

Eukaryotic Transcriptional Regulation

1. Level of Chromatin (DNA accessibility)

- Histone modifications
- Histone modifying enzymes & remodeling complexes
- Nucleosome composition
- DNA methylation

2. Level of DNA (Interaction with basal transcription machinery)

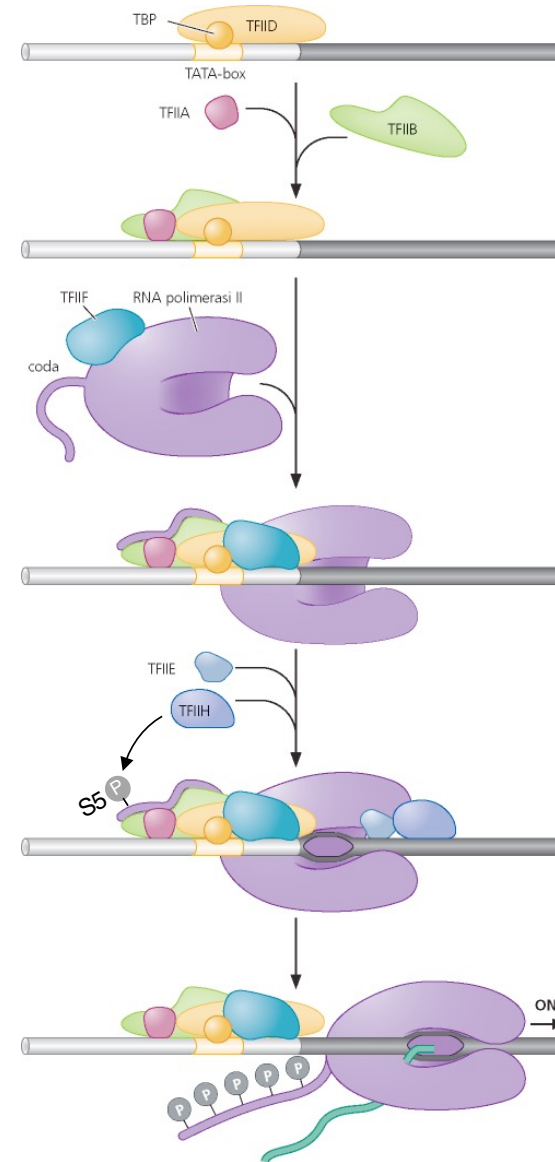
- Regulatory sequences (enhancers, silencers)
- Transcription factors (activators, repressors)

3. Level of Regulatory RNA (Interaction with DNA, RNA or protein)

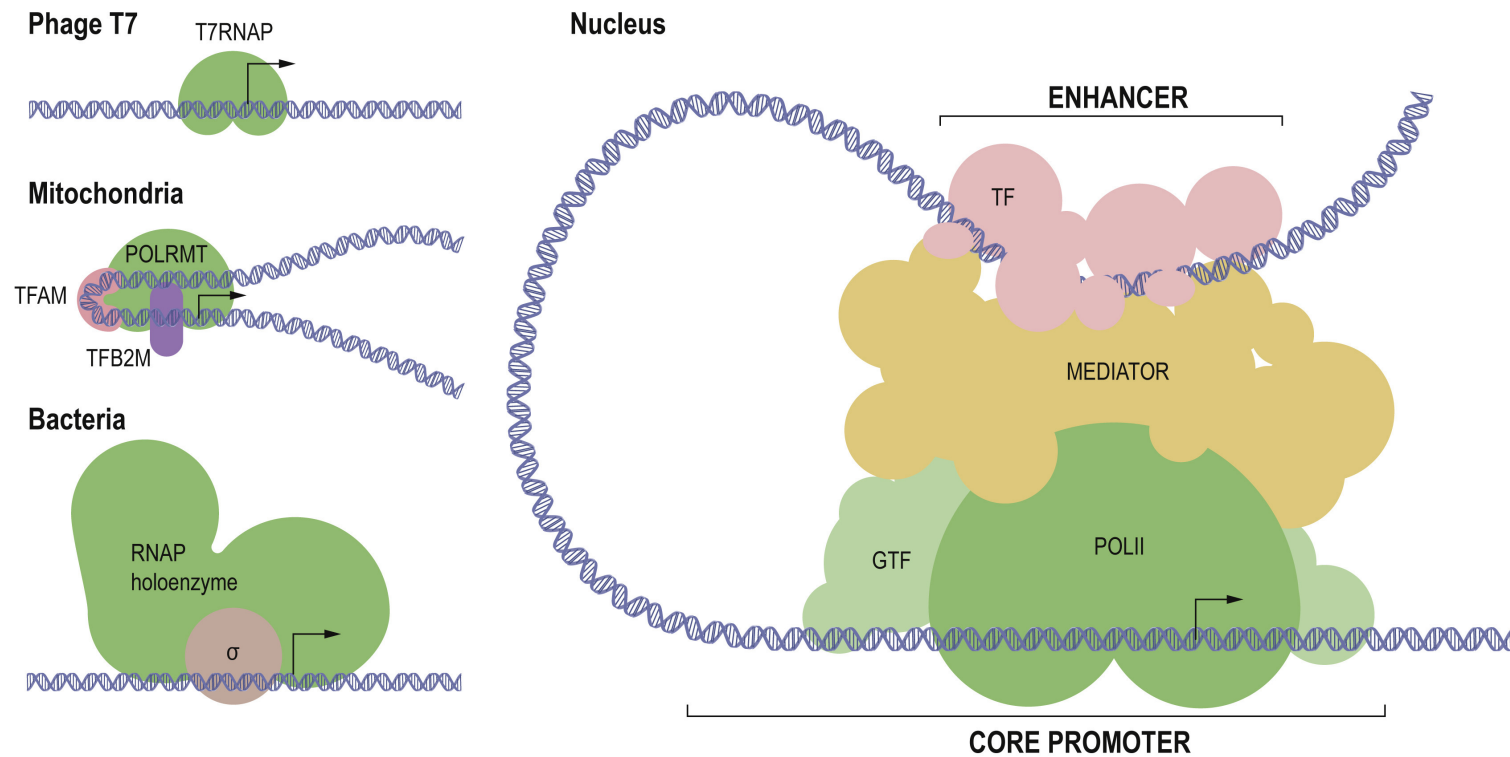
- Small and long non-coding RNAs

Transcription initiation in eukaryotes

At the transcription start site, Pol II initiation is regulated by a protein assembly known as the **pre-initiation complex (PIC)** containing **TFIIA**, **TFIIB**, **TFIID**, **TFIIE**, **TFIIF**, **TFIIH**, **Pol II** and **Mediator**

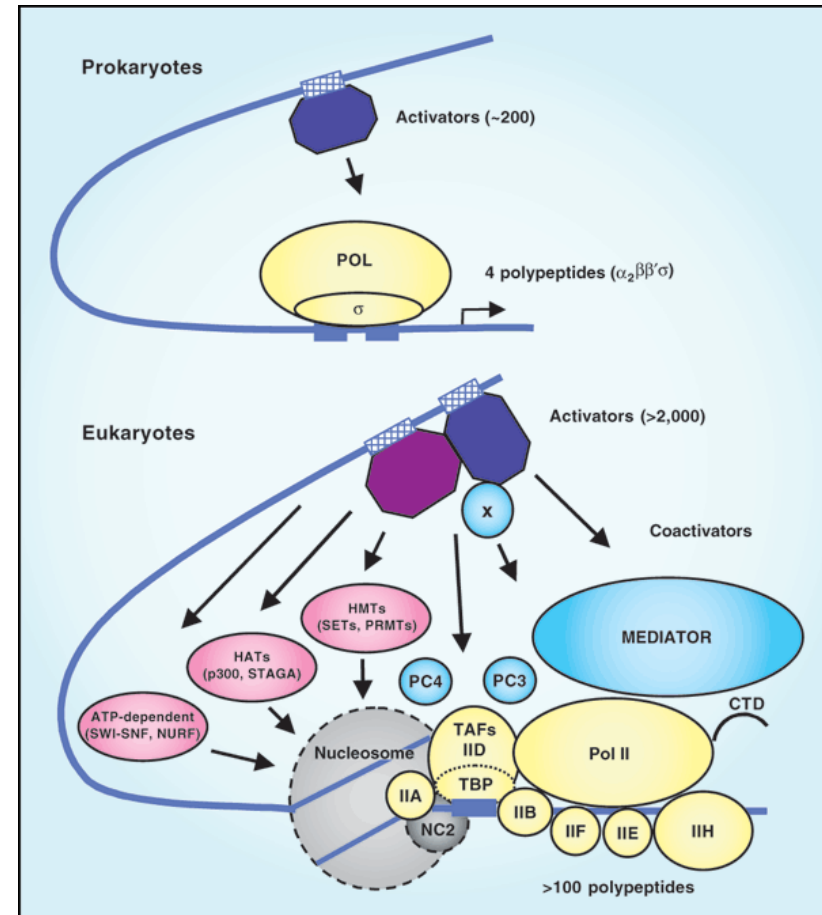


Schematic comparison of transcription initiation systems



Comparison of transcription activation mechanisms in prokaryotes and eukaryotes

The ultimate purpose of DNA-binding activators is to mediate RNA polymerase recruitment and function. But whereas activators interact directly with RNA polymerase subunits in prokaryotes, the effects of activators on RNA polymerase II in eukaryotes are mostly indirect and involve interactions with diverse cofactors that modify chromatin structure and facilitate recruitment of RNA polymerase II and general initiation factors

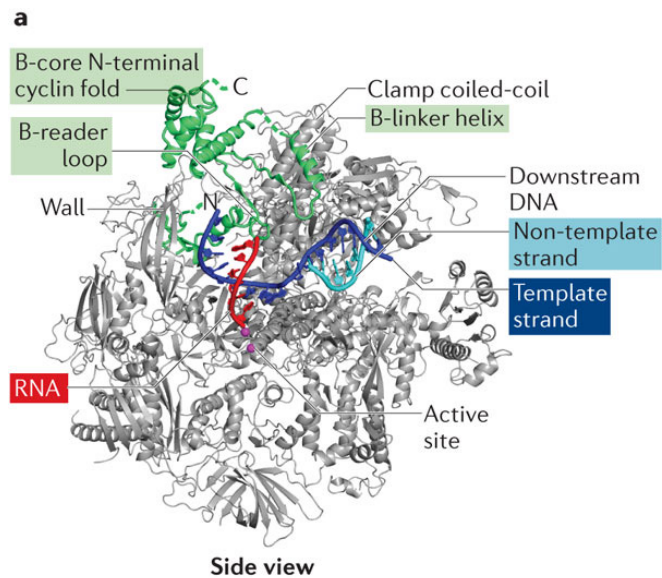


RNA transcription

Types of RNAs transcribed from active regulatory elements

Type	Acronyms	Estimated number	Half-life	Processing features
pre-messenger RNA	pre-mRNA	19,954	minutes to hours	capped spliced polyA tail
long non-coding RNA	lncRNA lincRNA	30,000	minutes to hours	capped spliced polyA tail
enhancer RNA	eRNA	40,000-65,000	minutes	capped
promoter-associated RNA	uaRNA paRNA PROMPT	Gene and lncRNA promoters	minutes	capped

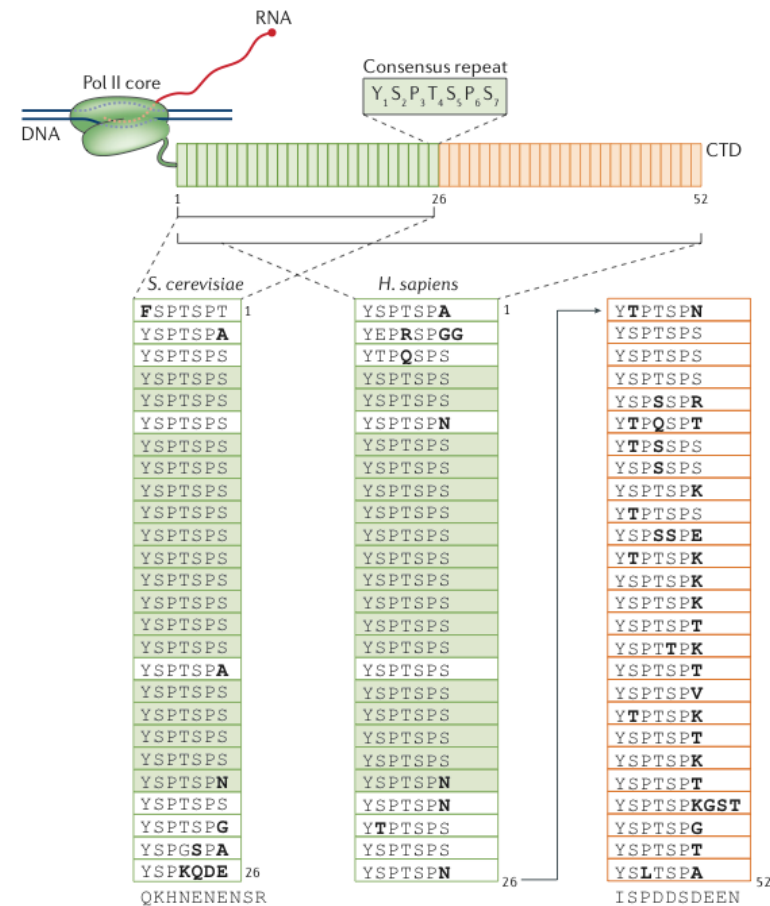
RNA Pol II



Factor	Gene name		Mass (kDa)		Uniprot accession number	
	Yeast	Human	Yeast	Human	Yeast	Human
<i>Pol II (RNAP*): transcribing enzyme</i>						
RPB1	<i>RPO21</i>	<i>POLR2A</i>	191.6	217.2	P04050	P24928
RPB2	<i>RPB2</i>	<i>POLR2B</i>	138.8	133.9	P08518	P30876
RPB3	<i>RPB3</i>	<i>POLR2C</i>	35.3	31.4	P16370	P19387
RPB4	<i>RPB4</i>	<i>POLR2D</i>	25.4	16.3	P20433	O15514
RPB5 ⁺	<i>RPB5</i>	<i>POLR2E</i>	25.1	24.6	P20434	P19388
RPB6 ⁺	<i>RPO26</i>	<i>POLR2F</i>	17.9	14.5	P20435	P61218
RPB7	<i>RPB7</i>	<i>POLR2G</i>	19.1	19.3	P34087	P62487
RPB8 ⁺	<i>RPB8</i>	<i>POLR2H</i>	16.5	17.1	P20436	P52434
RPB9	<i>RPB9</i>	<i>POLR2I</i>	14.3	14.5	P27999	P36954
RPB10 ⁺	<i>RPB10</i>	<i>POLR2L</i>	8.3	7.6	P22139	P62875
RPB11	<i>RPB11</i>	<i>POLR2J</i>	13.6	13.3	P38902	P52435
RPB12 ⁺	<i>RPB12</i>	<i>POLR2K</i>	7.7	7.0	P40422	P53803
Total (12 subunits)			513.6	516.7		

The carboxy-terminal domain (CTD)

This domain consists of heptapeptide repeats with the consensus sequence YSPTSPS and forms the **carboxy-terminal domain (CTD)** of the largest subunit of Pol II RNA polymerase B1 (RPB1). The CTD is not required for the catalytic functions of Pol II, but it is crucial for the regulation of multiple steps in transcription, as well as for the coupling of transcription to a number of co-transcriptional processes.



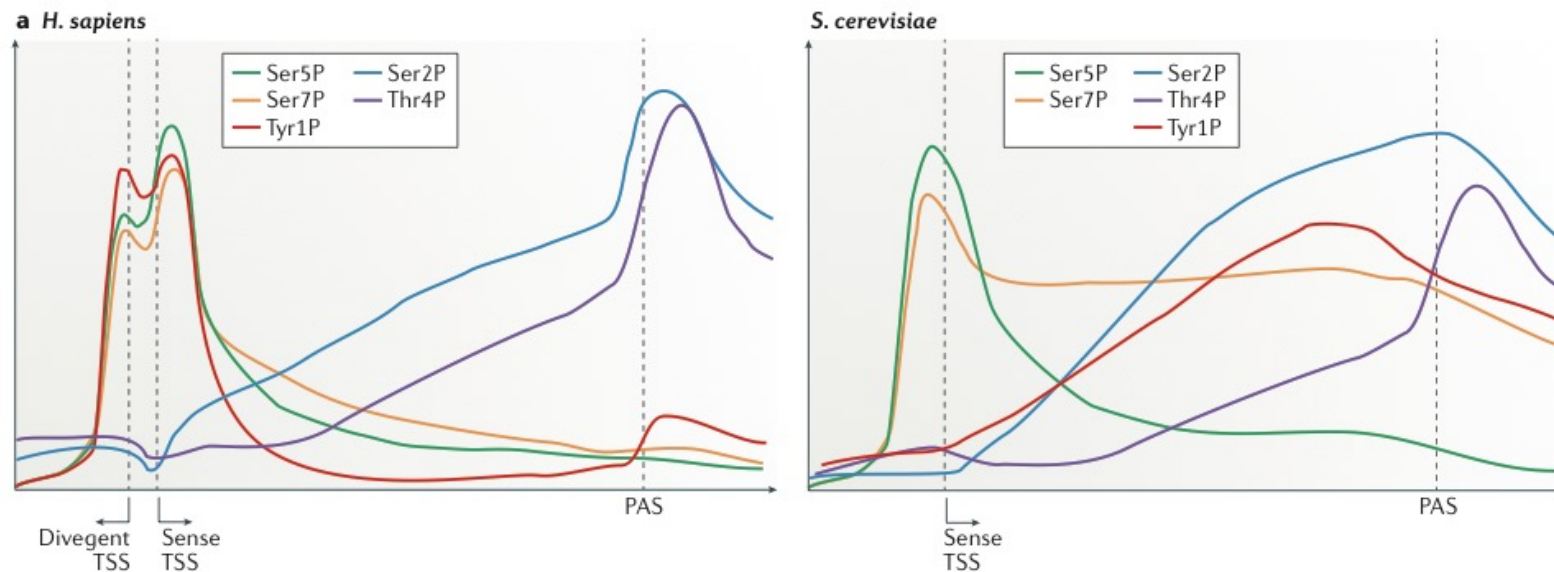
Post-translational modifications of the CTD and their associated processes

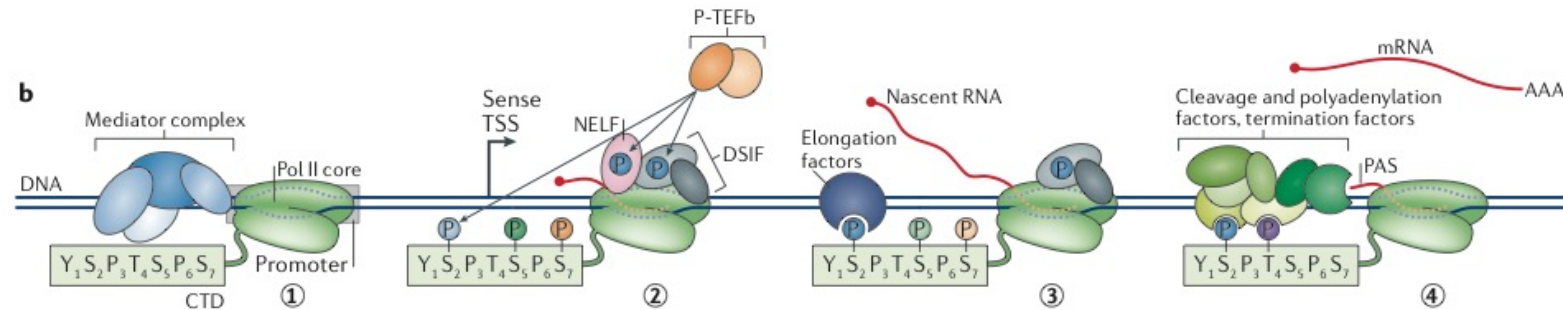
Post-translational modification	Position in the CTD	Organisms	Associated process or processes
Ser5 phosphorylation	Multiple repeats	<ul style="list-style-type: none"> • <i>Saccharomyces cerevisiae</i> • <i>Schizosaccharomyces pombe</i> • <i>Homo sapiens</i> 	Transcription initiation, mRNA capping and splicing, non-coding RNA transcription termination and chromatin modification
Ser2 phosphorylation	Multiple repeats	<ul style="list-style-type: none"> • <i>S. cerevisiae</i> • <i>S. pombe</i> • <i>H. sapiens</i> 	Transcription elongation, promoter-proximal pause and release, splicing, transcription termination and DNA topology
Ser7 phosphorylation	Multiple repeats	<ul style="list-style-type: none"> • <i>S. cerevisiae</i> • <i>S. pombe</i> • <i>H. sapiens</i> 	snRNA expression, interaction with the Integrator complex and P-TEFb recognition
Thr4 phosphorylation	Multiple repeats	<ul style="list-style-type: none"> • <i>S. cerevisiae</i> • <i>S. pombe</i> • <i>Gallus gallus</i> • <i>H. sapiens</i> 	Transcription elongation and termination, post-transcriptional splicing, processing of histone mRNA and chromatin remodelling
Tyr1 phosphorylation	Multiple repeats	<ul style="list-style-type: none"> • <i>S. cerevisiae</i> • <i>S. pombe</i> • <i>G. gallus</i> • <i>H. sapiens</i> 	Inhibition of recruitment of transcription termination factors, CTD stability, antisense and enhancer transcription
Arg methylation	Arg1,810 of human RPB1	<ul style="list-style-type: none"> • <i>Mus musculus</i> • <i>H. sapiens</i> 	snRNA and snoRNA regulation, R-loop resolution and transcription termination
Lys methylation	Lys7 in the non-consensus region of human CTD	<ul style="list-style-type: none"> • <i>H. sapiens</i> • <i>M. musculus</i> • <i>Drosophila melanogaster</i> • <i>Caenorhabditis elegans</i> 	Supports nucleosome occupancy at promoters; negatively regulates gene expression
Lys acetylation	Lys7 in the non-consensus region of murine CTD; Lys7 in repeats 39, 42, 47 and 49 of human CTD	<ul style="list-style-type: none"> • <i>H. sapiens</i> • <i>M. musculus</i> 	Induction of growth-factor response genes, transcription elongation; maintains balance between Lys methylation and acetylation and affects mRNA expression levels
O-GlcNAcylation	Ser5 and/or Ser7 in multiple repeats	<i>H. sapiens</i>	Pre-initiation complex assembly
Ubiquitylation	RPB1 Lys residues 859, 1866, 1873, 1887, 1908, 1922	<i>M. musculus</i>	RPB1 degradation

CTD, carboxy-terminal domain of RNA polymerase II; P-TEFb, positive transcription elongation factor B; RPB1, RNA polymerase B1; snRNA, small nuclear RNA; snoRNA, small nucleolar RNA.

Transcription regulation by the CTD code

Average chromatin immunoprecipitation (ChIP) profiles of phosphorylated residues of the carboxy-terminal domain (CTD) of RNA polymerase II (Pol II) across protein-coding genes in humans (*Homo sapiens*) and in budding yeast (*Saccharomyces cerevisiae*).





1. **Recruitment of the core Pol II enzyme** to the promoter with an unphosphorylated CTD that interacts with the Mediator complex. The Mediator complex has a high affinity for unphosphorylated CTD and upon phosphorylation of Ser5 of the CTD, this affinity is lost and Pol II escapes from the promoter.
2. **Pol II during promoter-proximal pausing.** At this step, Pol II is highly phosphorylated at Ser5 and Ser7, is paused downstream of the transcription start site (TSS) and is bound by **negative elongation factor (NELF)** and **DRB sensitivity-inducing factor (DSIF)**. The arrival of **positive elongation factor B (P-TEFb)** leads to the phosphorylation of NELF, DSIF and Ser2. These phosphorylation events are followed by the release of NELF and the transition into productive transcription elongation.
3. **During productive elongation**, the CTD contains lower levels of Ser5P and Ser7P and higher levels of Ser2P, which promotes the recruitment of many transcription elongation, chromatin-modifying and RNA-processing factors that regulate co-transcriptional processes.
4. **Pol II transitions from transcription elongation to termination.** Ser2P and Thr4P levels peak, thereby promoting the recruitment of cleavage and polyadenylation factors, as well as termination factors that release Pol II from the DNA.

General Transcription Factors

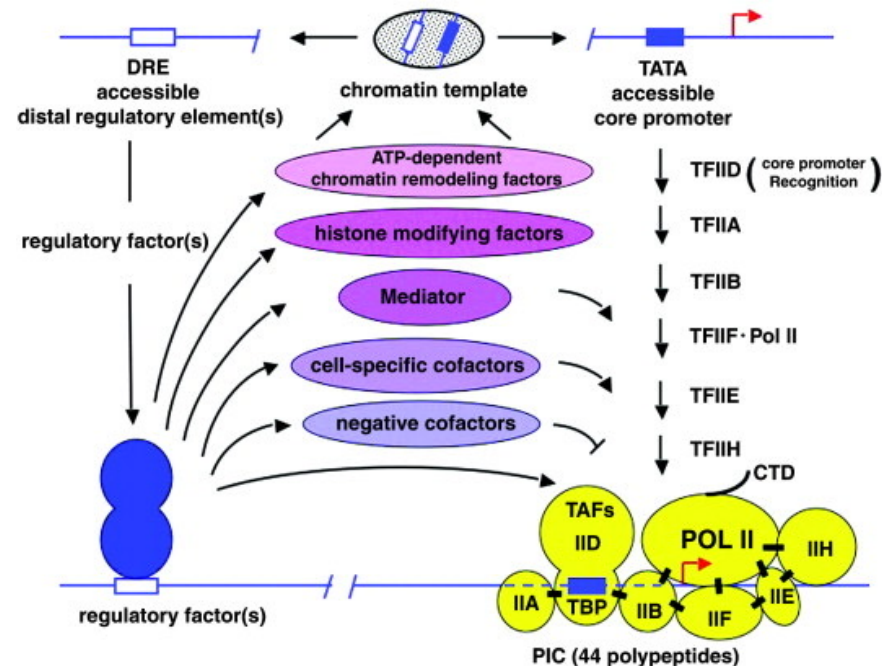
Factor	Gene name		Mass (kDa)		Uniprot accession number		Copies
	Yeast	Human	Yeast	Human	Yeast	Human	
TFIIA [§] : TBP stabilization and counteracts repressive effects of negative co-factors							
Large subunit	TOA1	GTF2A1	32.2	41.5	P32773	P52655	1
Small subunit	TOA2	GTF2A2	13.5	12.5	P32774	P52657	1
Total (2 subunits)			45.7	54.0			
TFIIB: Pol II recruitment, TBP binding and TSS selection							
TFIIB (TFB*)	SUA7	GTF2B	38.2	34.8	P29055	Q00403	1
TFIID: Pol II recruitment and promoter recognition							
TBP (TBP*): recognition of the TATA box	TBP	TBP	27.0	37.7	P13393	P20226	1
TAF1	TAF1	TAF1	120.7	212.7	P46677	P21675	1
TAF2	TAF2	TAF2	161.5	137.0	P23255	Q6P1X5	1
TAF3	TAF3	TAF3	40.3	103.6	Q12297	Q5VWG9	1
TAF4	TAF4	TAF4	42.3	110.1	P50105	O00268	2
TAF5	TAF5	TAF5	89.0	86.8	P38129	Q15542	2
TAF6	TAF6	TAF6	57.9	72.7	P53040	P49848	2
TAF7	TAF7	TAF7	67.6	40.3	Q05021	Q15545	1
TAF8	TAF8	TAF8	58.0	34.3	Q03750	Q7Z7C8	1
TAF9	TAF9	TAF9	17.3	29.0	Q05027	Q16594	2
TAF10	TAF10	TAF10	23.0	21.7	Q12030	Q12962	2
TAF11	TAF11	TAF11	40.6	23.3	Q04226	Q15544	1
TAF12	TAF12	TAF12	61.1	17.9	Q03761	Q16514	2
TAF13	TAF13	TAF13	19.1	14.3	P11747	Q15543	1
TAF14	TAF14	NA	27.4	NA	P35189	NA	3
Total (14–15 subunits)			1,200	1,300			
TFIIE: recruitment of TFIIH and open DNA stabilization							
TFIIE α (TFE*)	TFA1	GTF2E1	54.7	49.5	P36100	P29083	1
TFIIE β	TFA2	GTF2E2	37.0	33.0	P36145	P29084	1
Total (2 subunits)			91.7	82.5			

TAFs recognize DNA elements in the promoter region

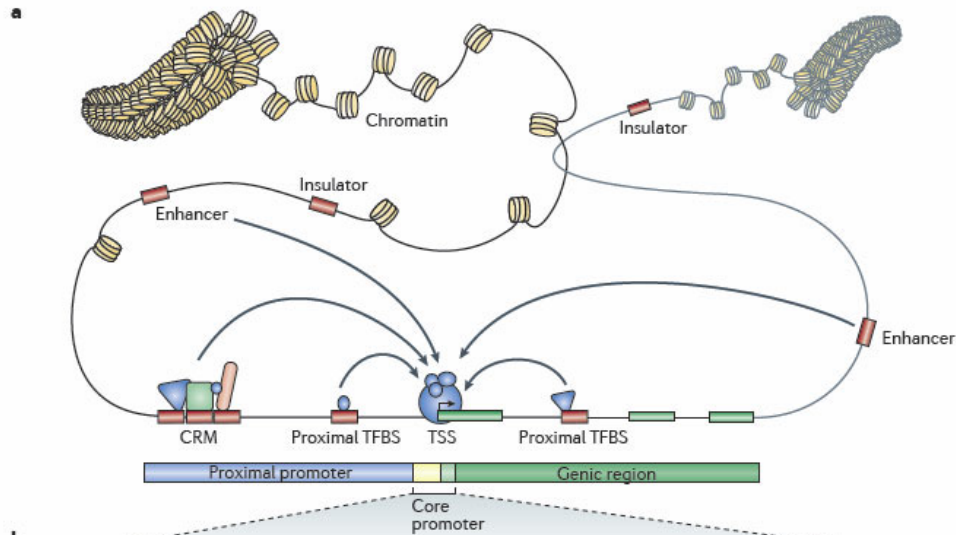
PIC assembly pathway for genes with a TATA-containing core promoter

Assembly of a PIC containing RNA polymerase II and general initiation factors (yellow) is nucleated by binding of TFIID to the core promoter.

A model for the regulation of PIC assembly and function involves, sequentially: (i) **binding of regulatory factors to distal control elements**; (ii) **regulatory factor interactions with cofactors that modify chromatin structure** to facilitate additional factor interactions; and (iii) **regulatory factor interactions with cofactors that act after chromatin remodeling** to facilitate, through direct interactions, recruitment or function of the general transcription machinery.

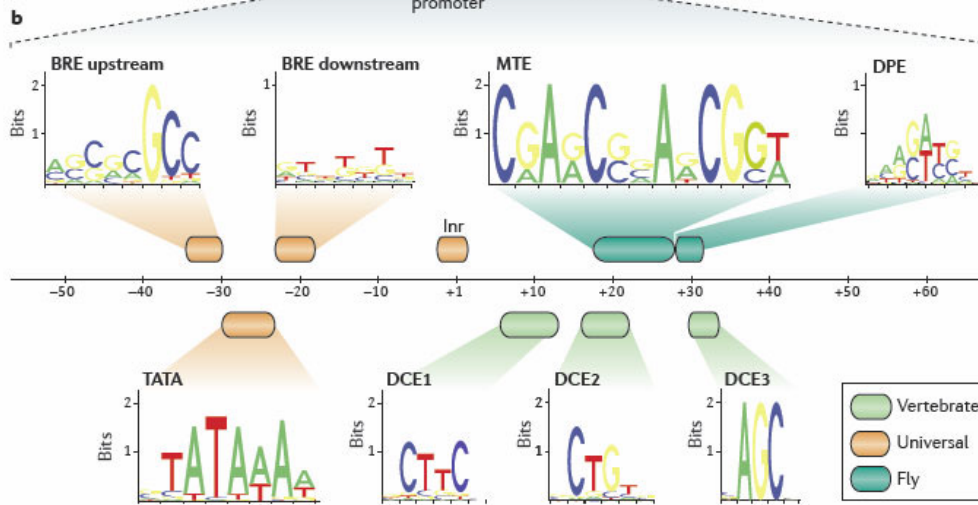


Promoter elements and regulatory signals in Metazoa



The term '**core promoter**' is often used to focus on the DNA region in the immediate vicinity of the TSS, which is assumed to dock the **pre-initiation complex (PIC)**

The core promoter consists of several interchangeable sequence elements around the TSS, which bind core components of the PIC.

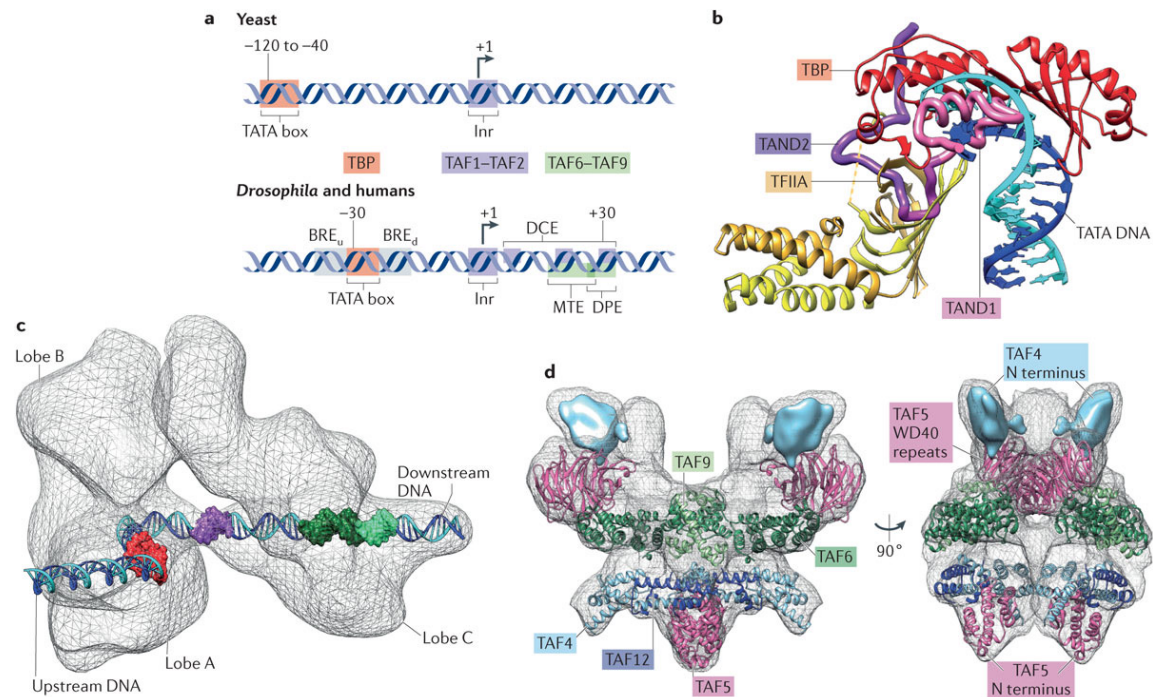


The region around the TSS has several over-represented sequence patterns; the **TATA box** and **initiator (Inr)** are the most studied and most prevalent

BRE, B recognition elements
DCE, downstream core element;
DRE, DNA recognition element;
MTE, motif ten element.

TFIID and promoter specificity

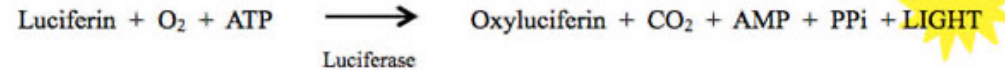
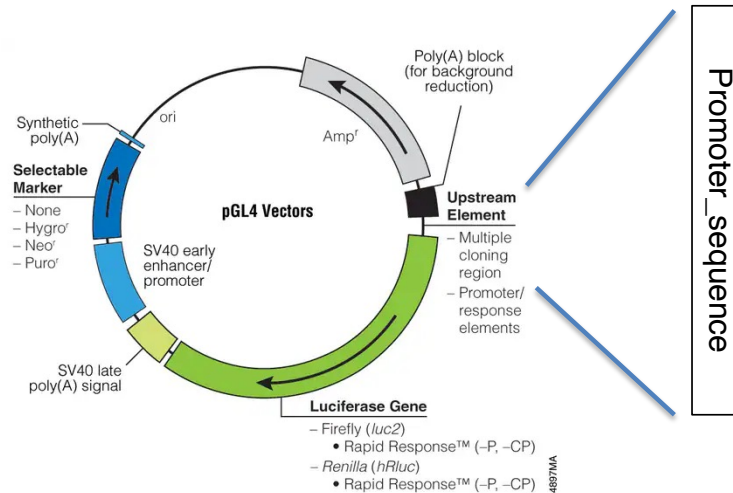
It was originally thought that TFIID was recruited to Pol II promoters through the TATA-binding activity of TBP. A conserved TATA box is, however, found only in about 15% of yeast and human promoters. Analysis of core promoter sequences led to the identification of several additional core promoter elements that are recognized by TAFs. TFIID also binds to transcription factors, which help to recruit TFIID to promoters and may facilitate enhancer–promoter interactions.



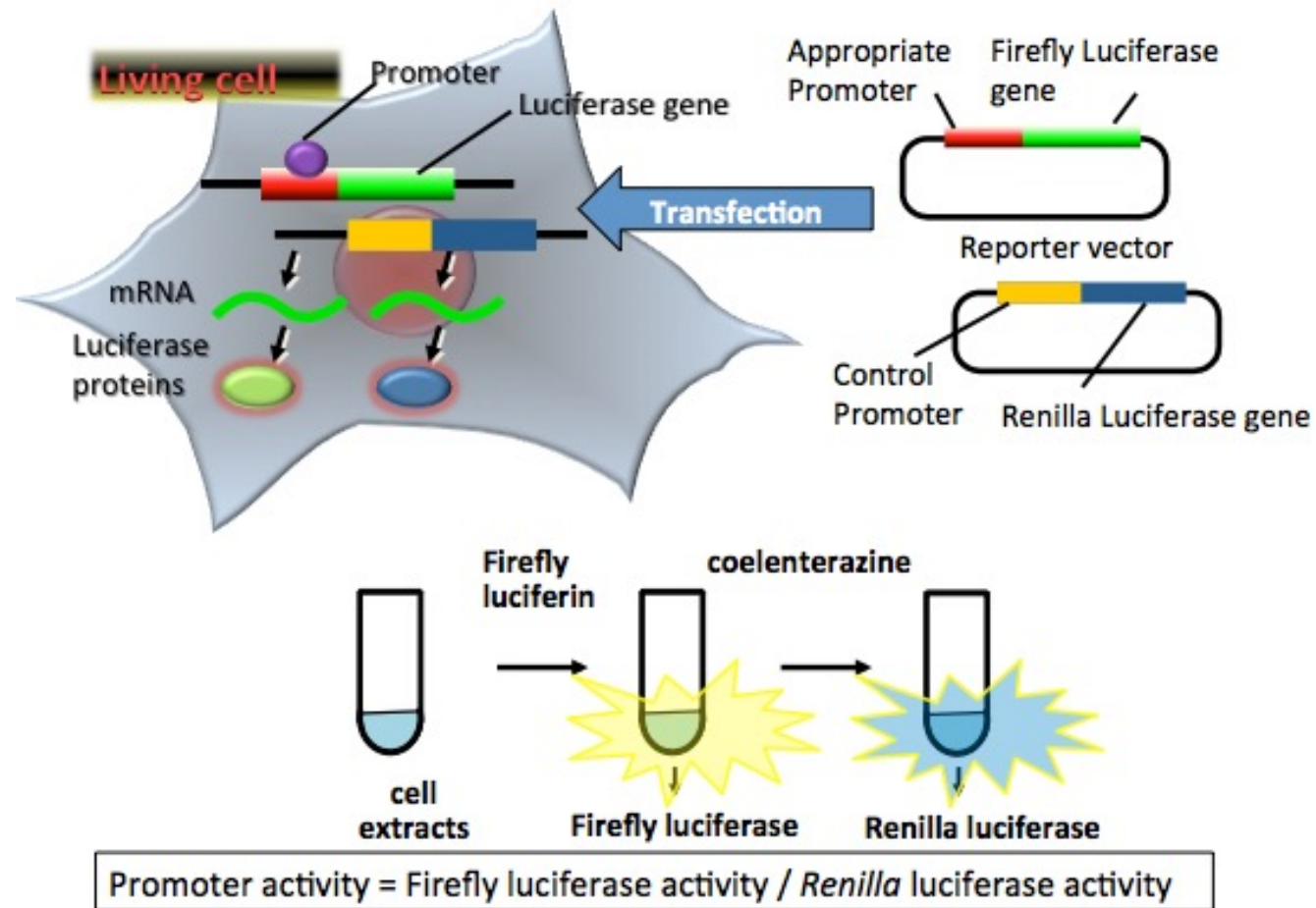
- 1. How do we identify a promoter?**
- 2. How do we study its activity?**

The Luciferase reporter assay

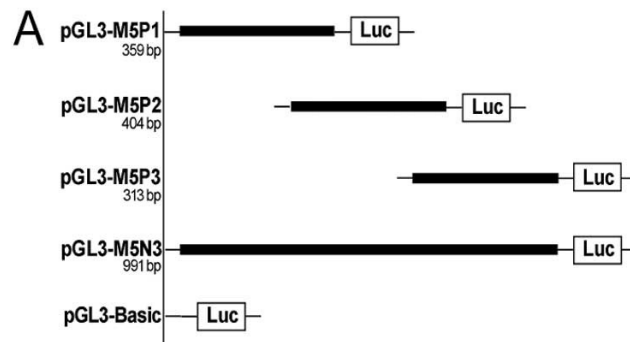
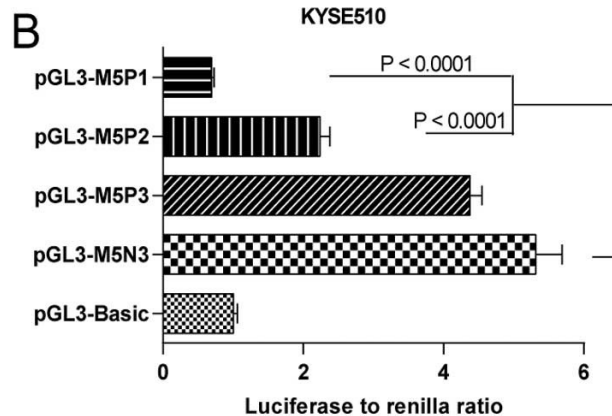
The luciferase reporter assay is commonly used as a tool to study gene expression at the transcriptional level. It is widely used because it is convenient, relatively inexpensive, and gives quantitative measurements instantaneously.



The Dual Luciferase reporter assay



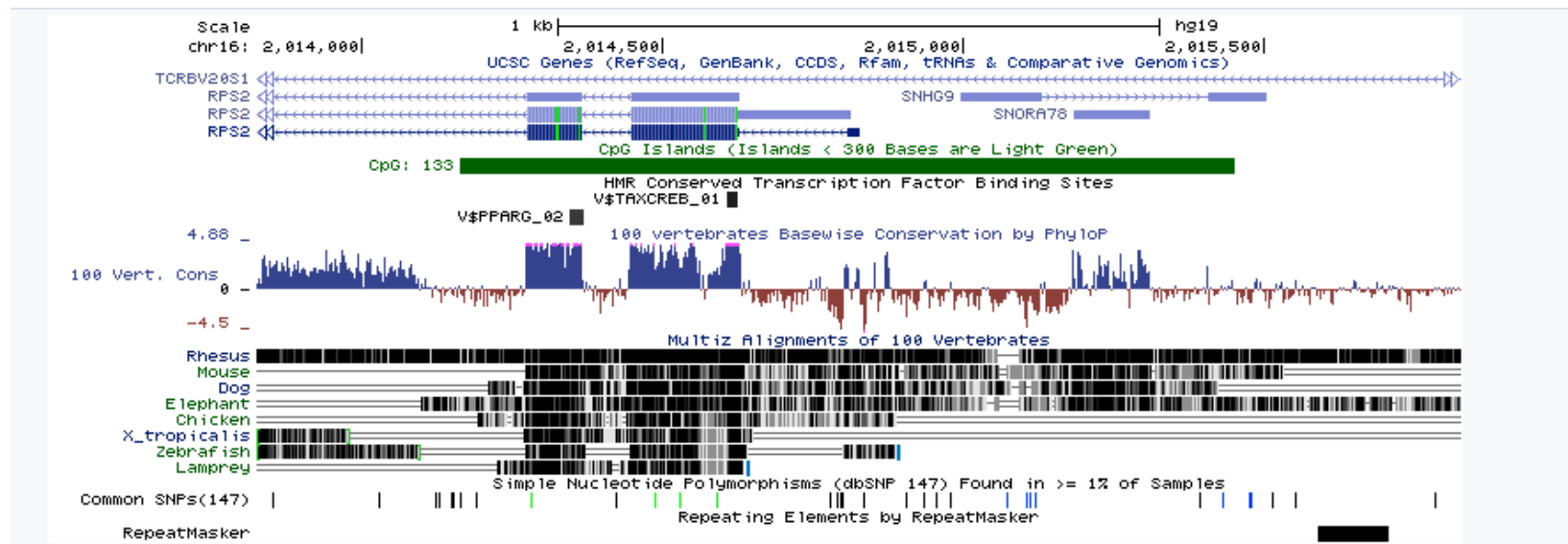
The Luciferase reporter assay

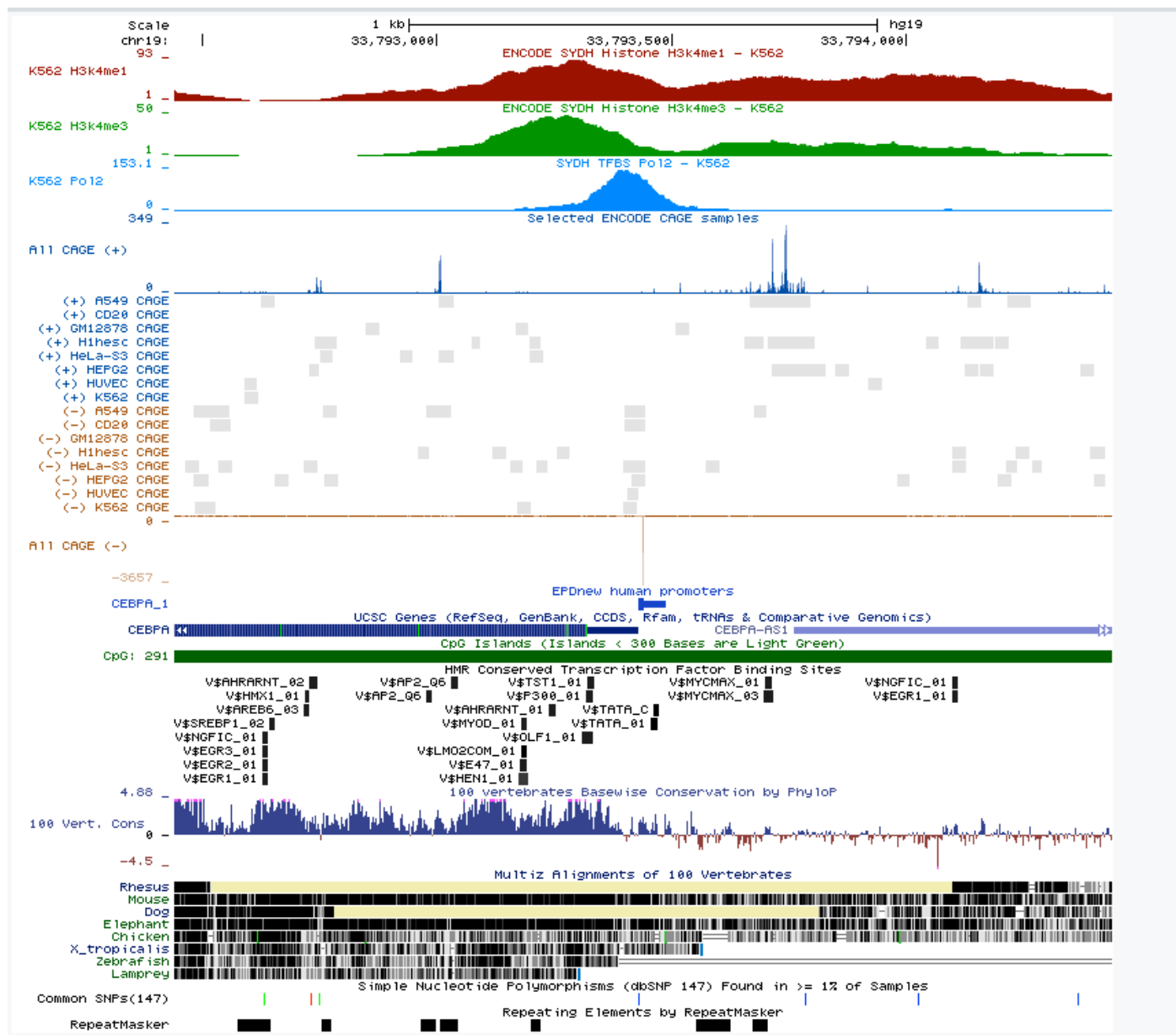


C M5N3 - GenBank KF765385
chr2:202,122,236-202,123,227 (hg19)

```
GGGTCTAGGGCTCAGAGCTTTGGAGAACAGACCTCAGTAGCACCAACACTCCAGGAT
CAATGCTACAAAGACACGGGTTACAACATAAACTGGAGAACATGGCCAAGGATGGGAA
CTCAGCCTGAGCAGGGCTGAGCCGAGCAGGGCTAAGCCAAGTAGGGCTGAGCCAGAA
CACTTCCTCCTTTTTTCTGAACAATCTACCTACATTTTCAGCTACAGGGCTGGCTTTAC
CCAGTCCGGCGGGAGGGAGGAGAGGGCTGGTCTGTGACTTCAGTGCTGAGGTTTGAT
CAAGGCAAAGGGAACTTCCTATTCCCAGACCCTTTGCAAGAAAGAATGGCATATTA
CTTGCCACCGACAGGGGTTATTATTACTAAATGGAGTCAGTATAAATGCTTTCAAT
AAAGCATGTCCAGCGCTCGGGCTTTAGTTTGCACGTCCATGAATTGTCTGCCACATCC
CTCTTCTGAATGGTTGGAAATTGGGCATCTGTTCTTTAAACAGGAAACATTTCTTG
TTCGAGTGAGTCATCTCTGTTCTGCTTTAGGAGTAAAGTTTACCCTGCAGTTCCTTCT
GTGGTGAAGTTTTCTCTTTCTCTCGGAGACCAGATTCTGCCTTTCTGCTGGAGGGAA
GTGTTTTACAGGTTCTCCTCCTTTTATCTTTTGTGTTTTTTTTCAAGCCCTGCTGAA
TTTGCTAGTCAACTCAACAGGAAGTGAGGCCATGGAGGGAGGCAGAAGAGCCAGGGT
GGTTATTGAAAGTAGAAGAACTTCTTCTGGGAGCCTTTCCACCCCTTCCCTGCT
GAGCACGTGGAGTTAGGCAGGTTAGGGGACTCGGAGACTGCGATGGTGCCAGGAAAG
GGTGGAGCGGGTGAGTGCTGTTGCCAAGGTGGCCTCTTCAACAGGAAACCACAATA
TTTTTGTTTCTTGACTTGCTCTAGAAACAGGGCTGTGGGGGTGGGAAGCAACTTGG
ATCTGCCCTTCTG
```

Any promoter should be study in its genomic context!





Promoter types in Metazoa

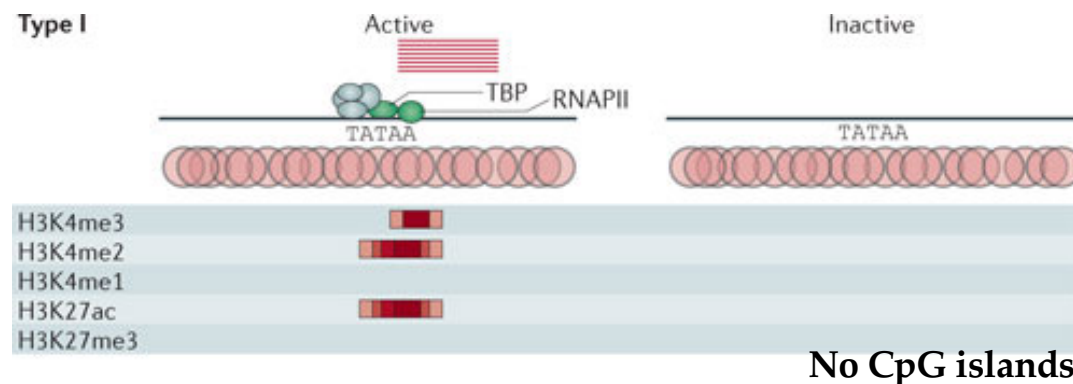
Promoter type	Dominant gene function	Common properties	Vertebrate-specific
<i>Major promoters</i>			
Type I ('adult')	Tissue-specific expression in adult peripheral tissues	Sharp ('focused') TSS, TATA-box enrichment, disordered nucleosomes	Mostly no CpG islands
Type II ('ubiquitous')	Broad expression throughout organismal cycle	Broad ('dispersed') TSS, ordered nucleosome configuration	CpG islands, TATA-depleted
Type III ('developmentally regulated')	Differentially regulated genes, often regulators in multicellular development and differentiation	Polycomb repression-regulated genes, broad H3K27me3 marks	Large CpG islands extending into the body of gene
<i>Minor promoters</i>			
TCT promoter	Highly expressed genes of translational apparatus	Sharp, pyrimidine-stretch ('TCT') initiator sequence, often full TATA box, ubiquitous-promoter-like nucleosome configuration	CpG island overlapping

DPE, downstream promoter element; DRE, DNA recognition element; H3K27me3, histone H3 lysine 27 trimethylation; TSS, transcription start site.

Promoter types in Metazoa

Promoter type	Dominant gene function	Common properties	Vertebrate-specific
<i>Major promoters</i>			
Type I ('adult')	Tissue-specific expression in adult peripheral tissues	Sharp ('focused') TSS, TATA-box enrichment, disordered nucleosomes	Mostly no CpG islands

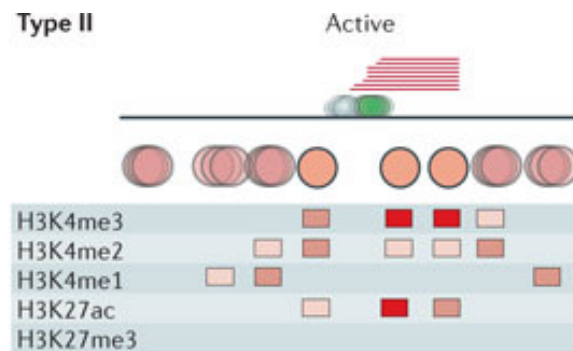
Genes that are specifically expressed in peripheral terminally differentiated tissues (e.g. liver or skeletal muscle) have **type I promoters** with a pattern of histone modification that is distinct from that of most other genes. H3K4me3 is generally only present downstream of the TSS, and there is no RNAPII binding at these promoters when the genes are not active.



Promoter types in Metazoa

Promoter type	Dominant gene function	Common properties	Vertebrate-specific
Type II ('ubiquitous')	Broad expression throughout organismal cycle	Broad ('dispersed') TSS, ordered nucleosome configuration	CpG islands, TATA-depleted

Ubiquitously expressed genes have H3K4me3 throughout their **type II promoters** across all tissues. Across classes of genes in vertebrates, H3K4me3 distribution is almost identical with the span of CpG islands. Ubiquitously expressed genes generally have short CpG islands, and the H3K4me3 mark and CpG island typically only overlap the 5' end of the gene.

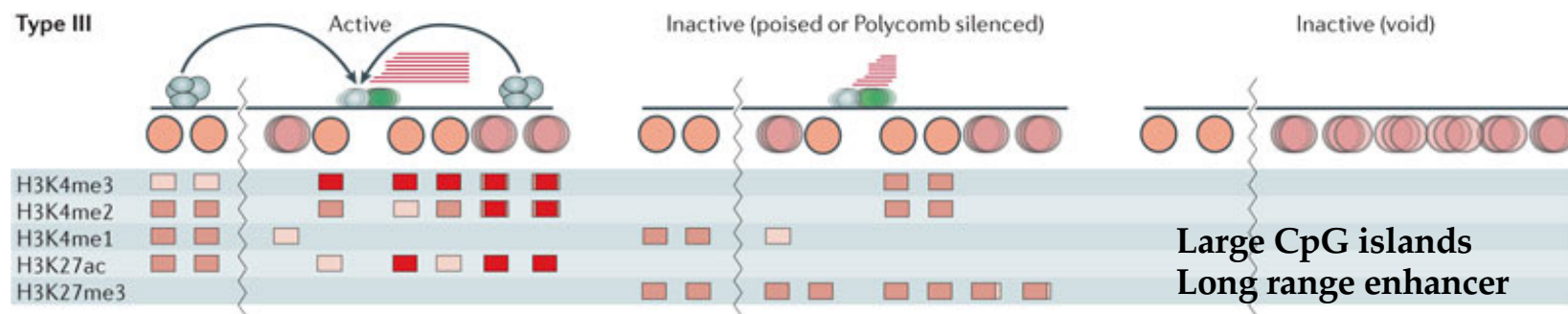


Short CpG islands

Promoter types in Metazoa

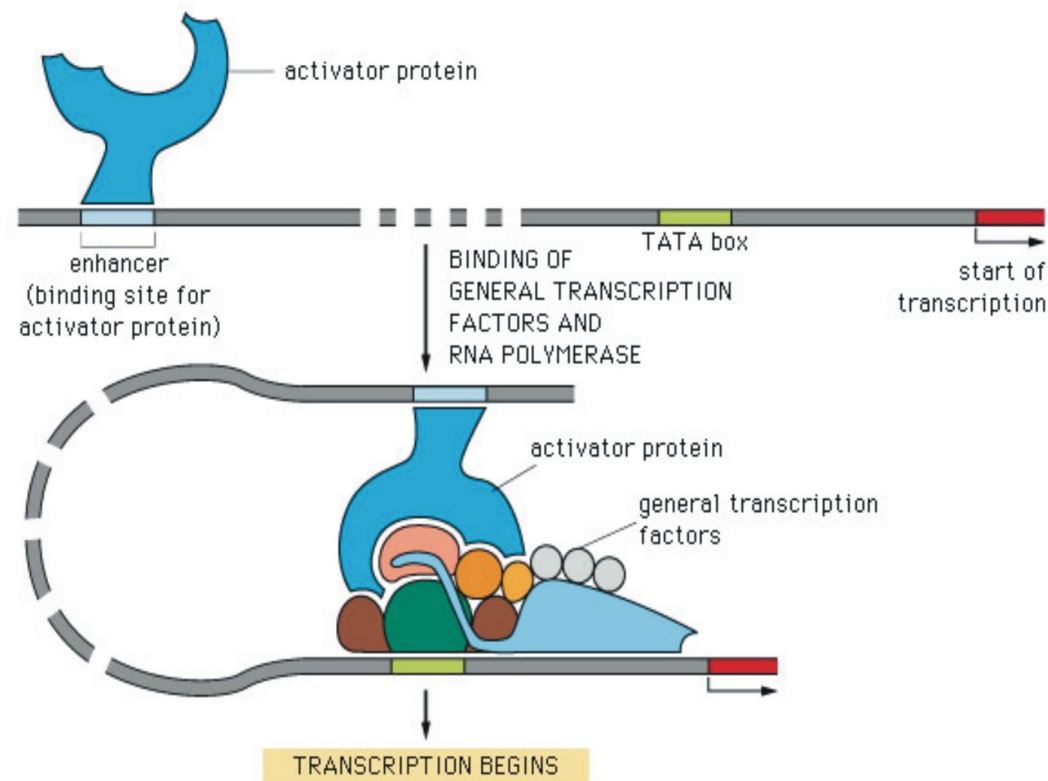
Promoter type	Dominant gene function	Common properties	Vertebrate-specific
Type III (‘developmentally regulated’)	Differentially regulated genes, often regulators in multicellular development and differentiation	Polycomb repression-regulated genes, broad H3K27me3 marks	Large CpG islands extending into the body of gene

Developmentally regulated genes (with **type III promoters**) have a number of features that are associated with repression by **polycomb group proteins (PcG)**. These features include multiple large CpG islands, wide distribution of bound PcG proteins and both H3K27me3 and H3K4me3. Because of the presence of both of these marks, which are associated with repression and activation, respectively, these are described as **bivalent promoters**. The large CpG islands often extend into the body of the genes and are closely tracked by H3K4me3, which is thus not restricted to promoter regions in developmental genes.



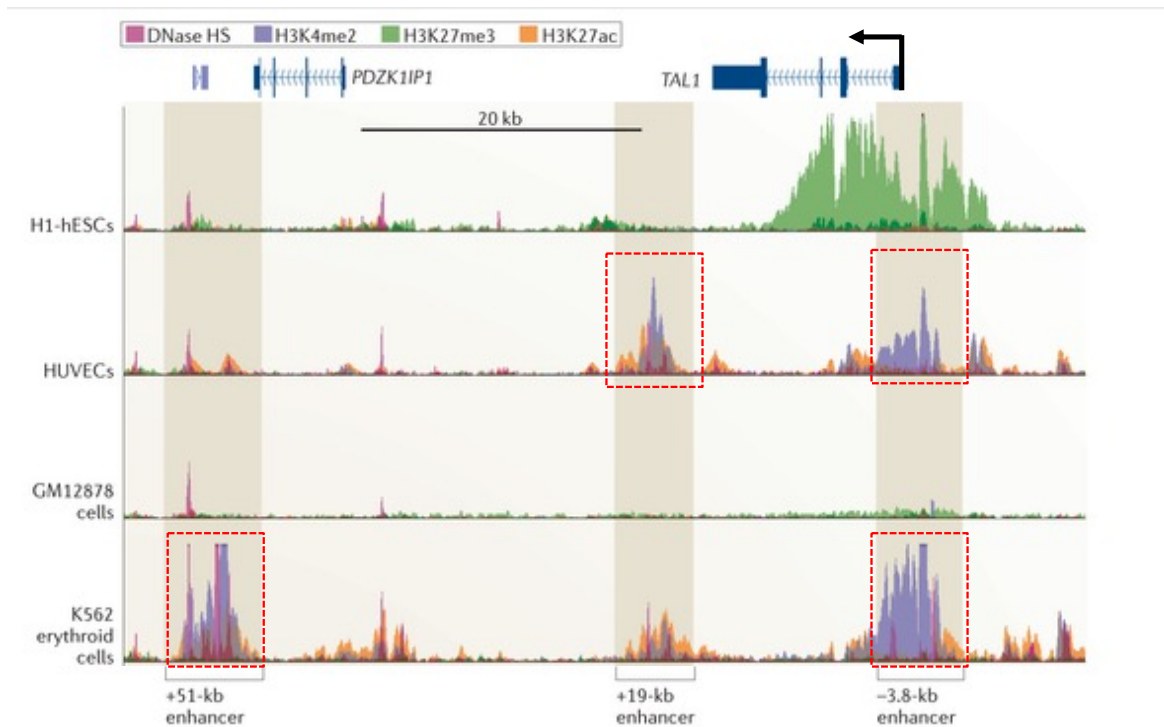
Enhancers

Enhancer is a short (50-1500 bp) region of DNA that can be bound by proteins (transcriptional activators) to increase transcription of a particular gene.



Enhancers

Genomic regions that function as transcriptional enhancers are enriched in closely spaced recognition motifs for sequence-specific transcription factors. Cell type-specific enhancers are marked by specific epigenomic features and chromatin accessibility. In particular, enhancers display enrichment of H3K4me1 or H3K4me2 and depletion of H3K4me3 compared with promoters



~10,000–50,000
putative enhancers
can be identified in
a given cell type

Enhancers

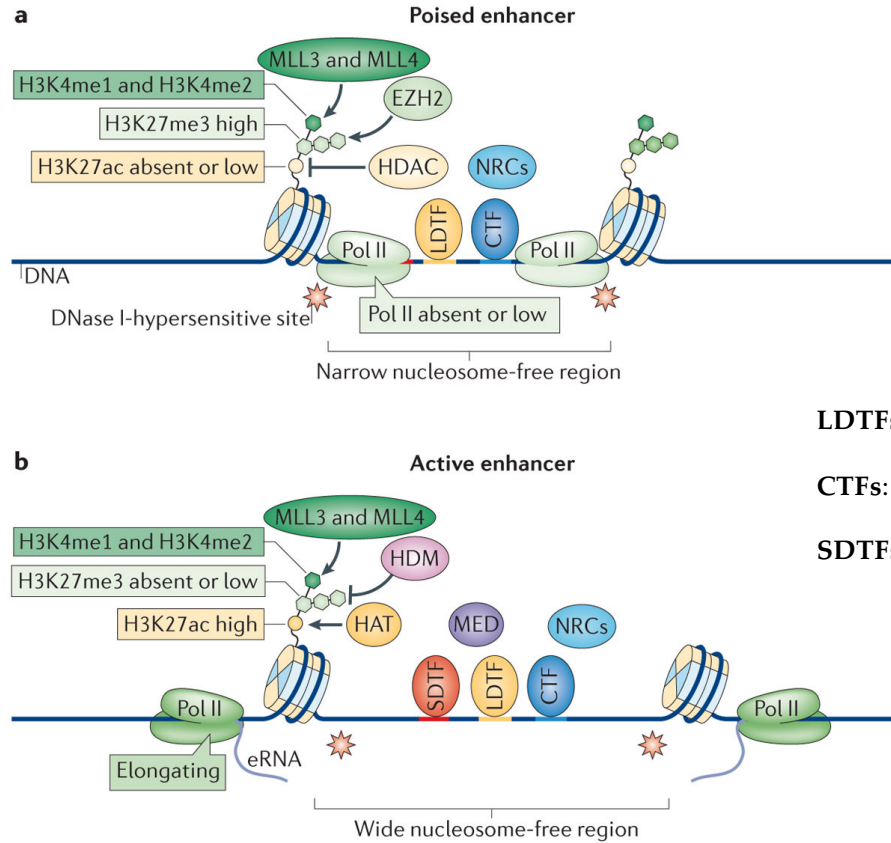
Enhancer activation begins with the binding of transcription factors and local nucleosome remodelling. Enhancer states can broadly be classified as:

Inactive enhancers: are essentially buried in compact chromatin and are devoid of transcription factor binding and histone modifications.

Primed/Poised enhancers: **Primed enhancer** are characterized by closely bound sequence-specific transcription factors that establish a and nucleosome-free region of open chromatin. However, they may require additional cues to accomplish their function, which may include signal-dependent activation, the recruitment of additional transcription factors and the eventual recruitment of co-activators that lead to enhancer activation. **Poised enhancers** can be defined as primed enhancers that also contain repressive epigenetic chromatin marks

Active enhancers: present co-activator complexes containing histone demethylase (HDM) complexes that remove H3K27me3 marks, histone acetyltransferases (HATs) that deposit H3K27 acetylation (H3K27ac) marks, and the **Mediator complex (MED)**. The transformation to elongating Pol II results in bidirectional transcription — a hallmark of active enhancers — and the generation of **enhancer RNAs (eRNAs)**, which is closely coupled to enhancer activity.

Enhancers

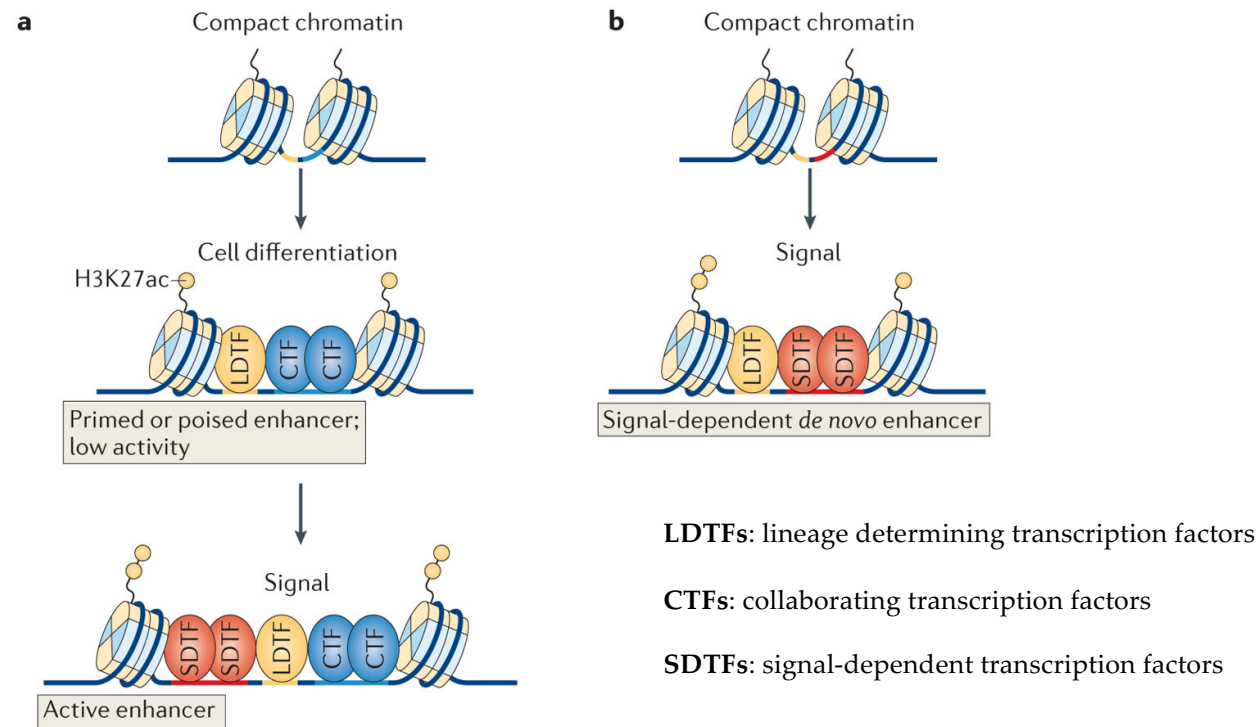


LDTFs: lineage determining transcription factors

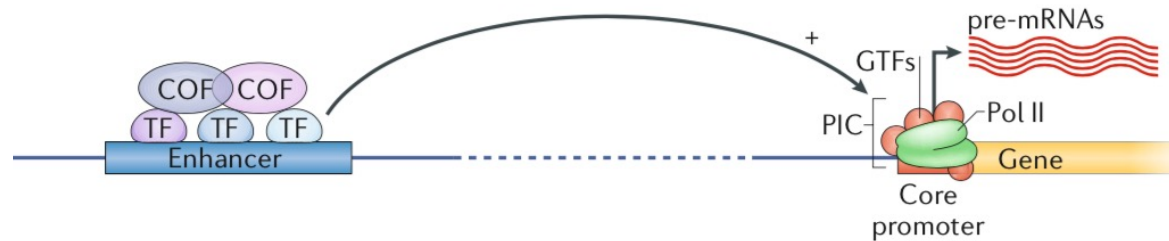
CTFs: collaborating transcription factors

SDTFs: signal-dependent transcription factors

Cell type-specific enhancer selection and activation

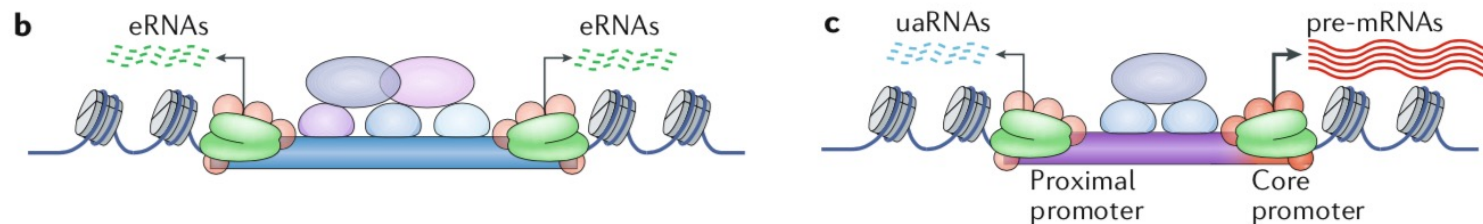


Widespread transcription of mammalian genome



The traditional view of transcription initiation

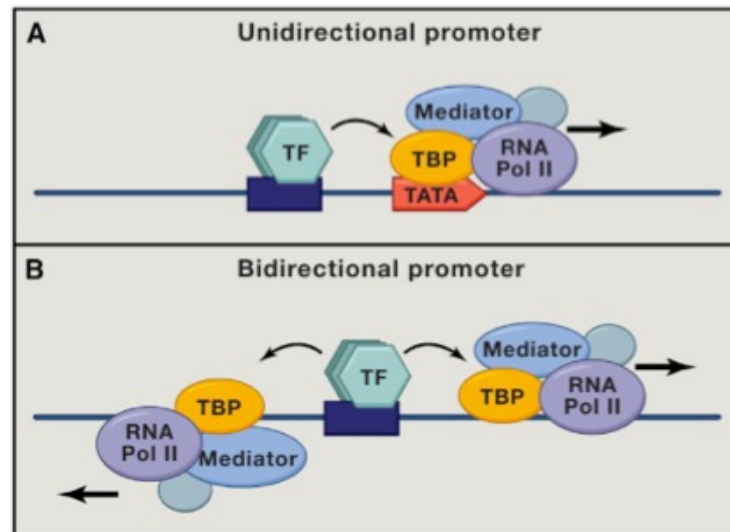
The traditional view of gene expression and the role of enhancers and core promoters has been challenged by the observation that many genomic positions outside annotated gene starts initiate transcription, including positions within enhancers.



The modern view of transcription initiation

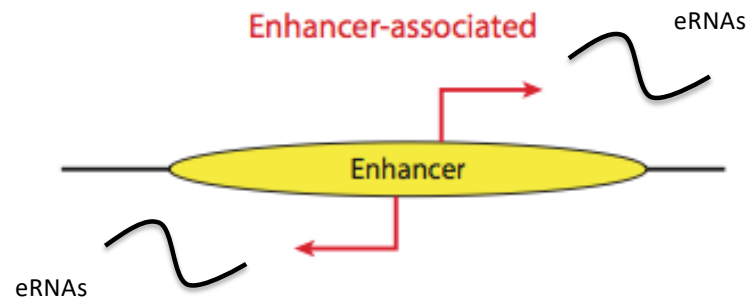
Divergent transcription from promoter and enhancer produces non coding RNAs

Most mammalian promoters lack a TATA element (TATA-less). For these promoters, TBP is recruited through sequence-specific transcription factors (such as Sp1) and components of the TFIID complex that have little sequence specificity. Thus, in the absence of strong TATA elements, TBP-complexes are recruited on both sides of the transcription factors to form preinitiation complexes in both orientations. Since transcription factors with transcription activation domains also bind at enhancer sites, these are also sites of divergent transcription.



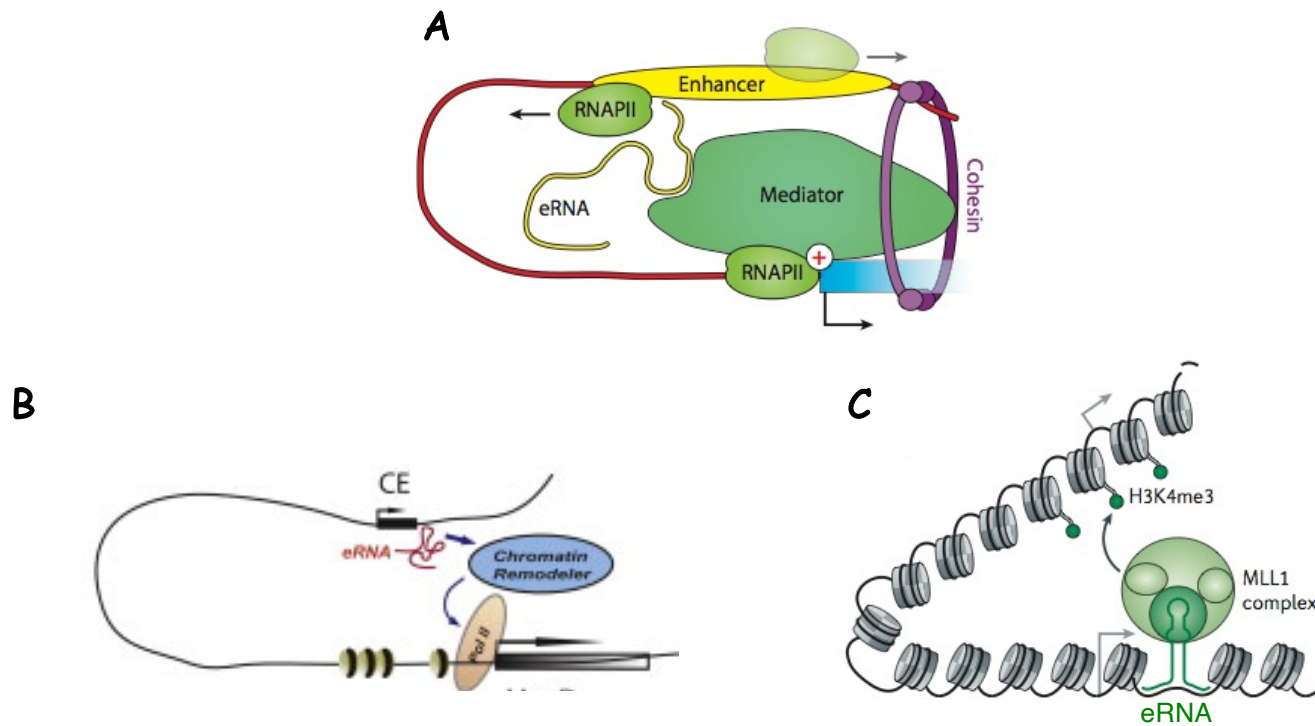
Enhancer RNAs (eRNAs)

- eRNAs are produced from transcriptional enhancers
- eRNA transcription correlates with the activation of the neighboring protein-coding genes (*cis*-acting)
- The majority of eRNAs are reported to be monoexonic and not polyadenylated (2–5 kb in length), but few cases of spliced and polyadenylated eRNAs have been described.



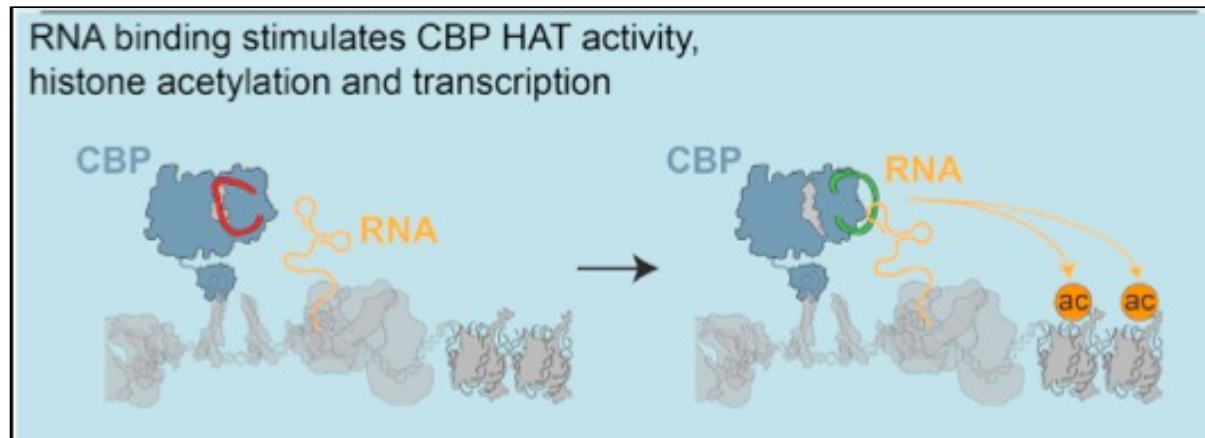
Enhancer RNAs

eRNAs **cooperate with the the Mediator and Cohesin complexes (A)** to promote enhancer–promoter interactions (chromatin looping) or they **recruit chromatin remodeling (B) or chromatin modifying complexes (C)** to the targeted promoters.



eRNA Binding to CBP Stimulates Histone Acetylation and Transcription

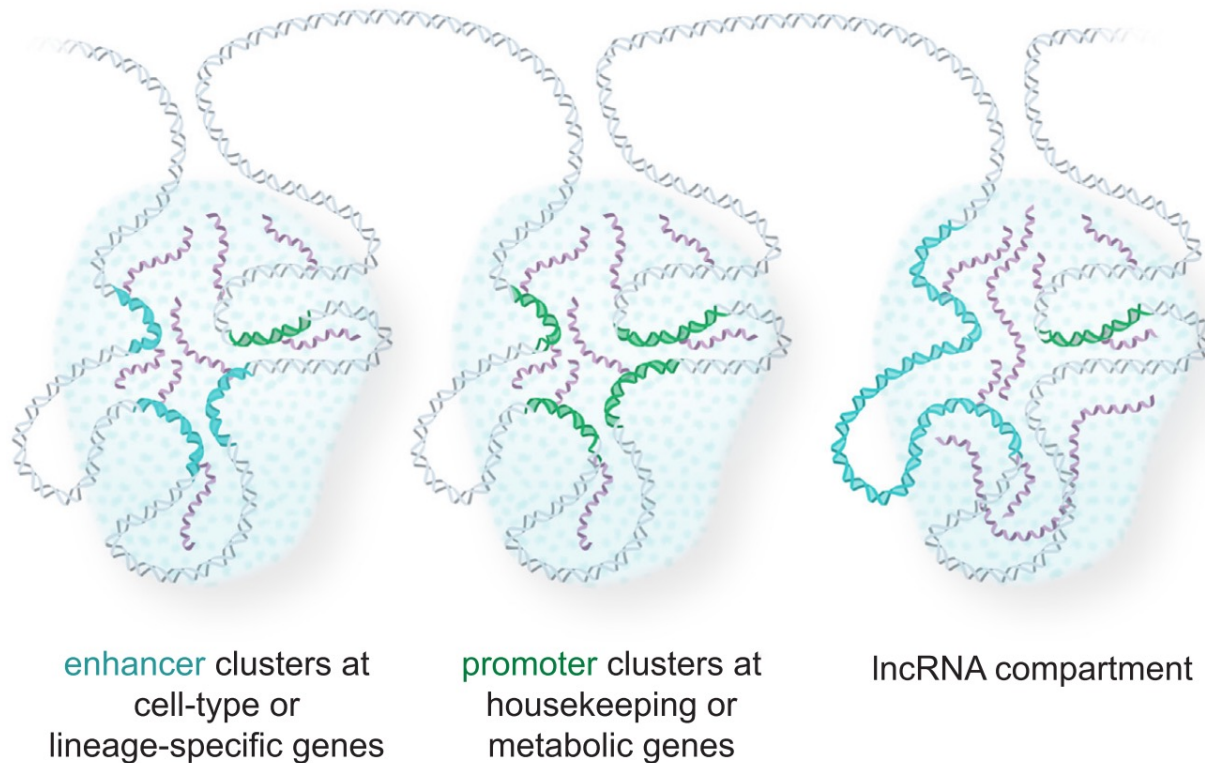
CBP/p300 are histone acetyltransferase (HAT) whose binding is a signature of enhancers. eRNA binding region in the HAT domain of CBP allows RNA to stimulate CBP's HAT activity. At enhancers where CBP interacts with eRNAs, stimulation manifests in RNA-dependent changes in the histone acetylation mediated by CBP, such as H3K27ac, and by corresponding changes in gene expression. By interacting directly with CBP, eRNAs contribute to the unique chromatin structure at active enhancers, which, in turn, is required for regulation of target genes.



Bose et al., 2017. Cell

Three-dimensional spatial compartments in transcription

Active enhancers and promoters assemble large amounts of transcription apparatus, forming transcriptional compartments that have been called **condensates**. Many condensates harbour RNA components and these can influence the formation and stability of the assembly.



The mediator complex

Mediator is a multiprotein complex that is implicated in regulating many fundamental processes involved in transcription including:

- transcription initiation,
- transcription elongation,
- chromatin architecture,
- enhancer-promoter gene looping

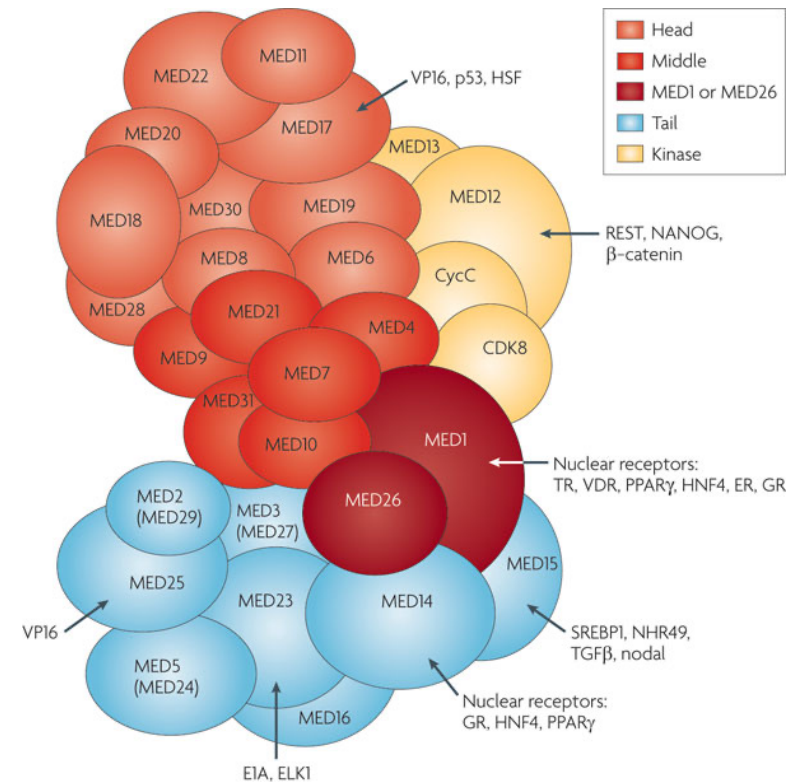
The mediator complex

The human Mediator complex has 30 subunits and is 1.2 MDa in size.

Mediator is universally required for the expression of almost all genes

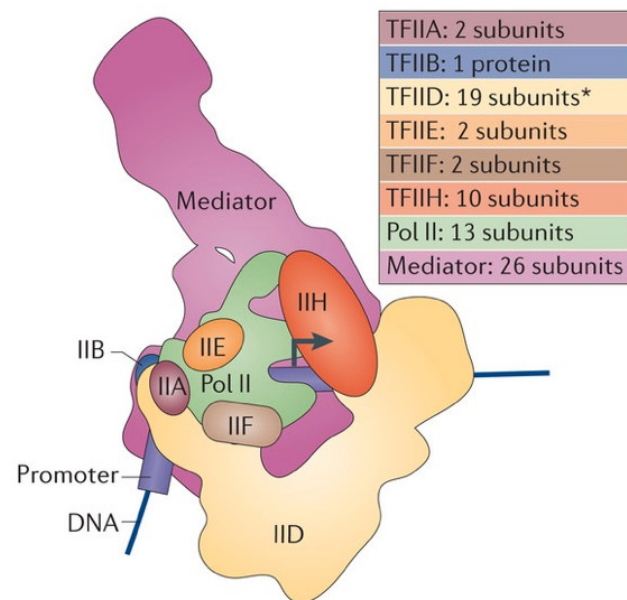
A fundamentally important characteristic of Mediator is that its subunit composition can change; subunits can be lost or added to affect its biological function. individual subunits reside in one of four distinct modules designated '**head**', '**middle**', '**tail**' and '**kinase**'

Whereas the head, middle and tail modules form a relatively stable 'core' structure, the kinase module associates reversibly with the Mediator complex



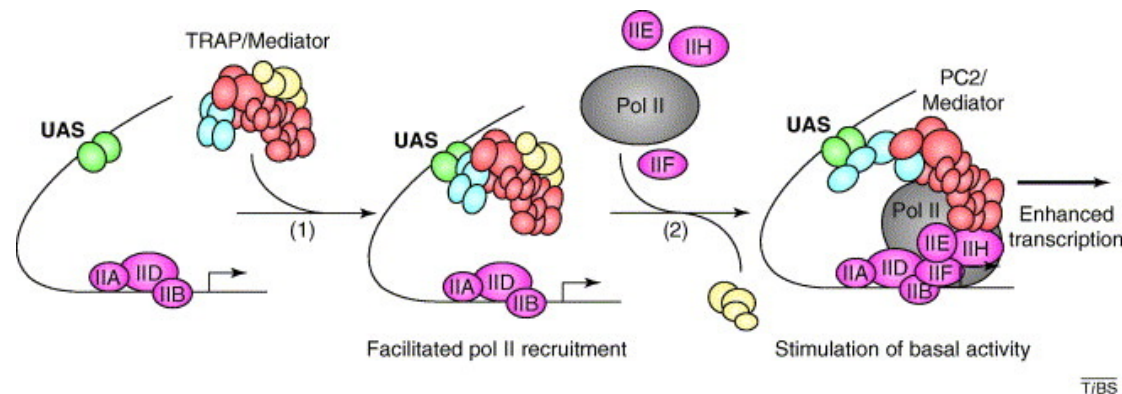
Mediator has key roles in the assembly of the PIC

For gene-specific transcription, Pol II must be recruited to specific sites on the genome. This is generally controlled by sequence-specific, DNA-binding TFs. Although TFs do not directly bind to Pol II, one mechanism by which they can promote Pol II recruitment is by binding to the Mediator complex. Mediator enables Pol II recruitment via interaction with the CTD of the Pol II. The large size of Mediator is likely to promote stable PIC formation by allowing the complex to directly interact with multiple PIC factors. Moreover, Mediator helps to regulate the recruitment and/or the activity of the PIC components.



Model for the Mediator-dependent transcriptional activation pathway

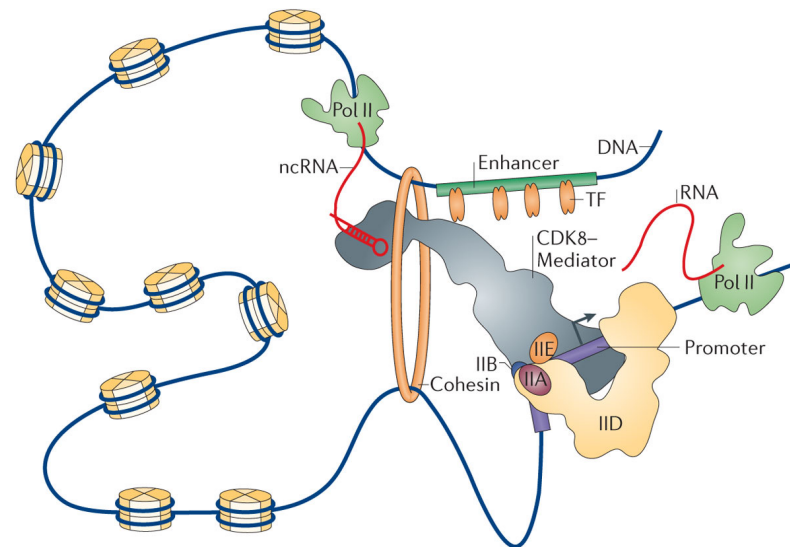
Mediator function is proposed to be manifested at two distinct levels. (1) First, in concert with the activators, **Mediator** might promote PIC assembly (**recruitment**). (2) Second, following structural or topological conversion, Mediator could modulate the (basal) function of pol II in the PIC (**post-recruitment**).



By inserting itself between the activators and the general transcription factors, the Mediator affords additional opportunities to fine-tune the diverse regulatory inputs received both from the DNA-binding factors and, most likely, from other signals and to present an appropriately calibrated output to the pol II machinery.

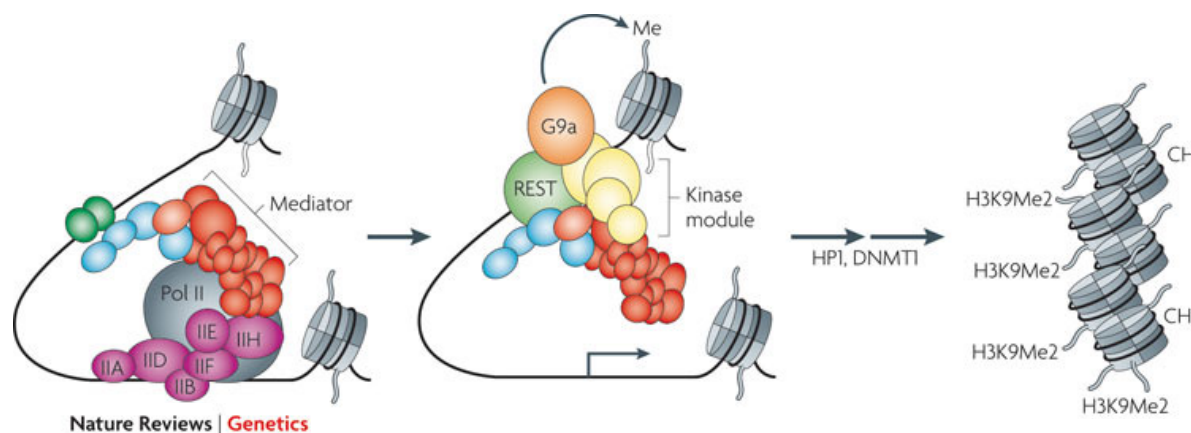
Mediator and chromatin architecture

The 3D organization of the genome has basic roles in coordinating gene expression programmes. A common theme is the formation of **DNA loops** that enable interaction between linearly separated DNA sequences. Mediator has a central role in the formation and/or stabilization of looped DNA structures in eukaryotes. In mammalian cells, long-range looping interactions between enhancer and promoter sequences appear to be important for driving high-level and cell type-specific gene expression.



The mediator complex can also contribute to epigenetic silencing

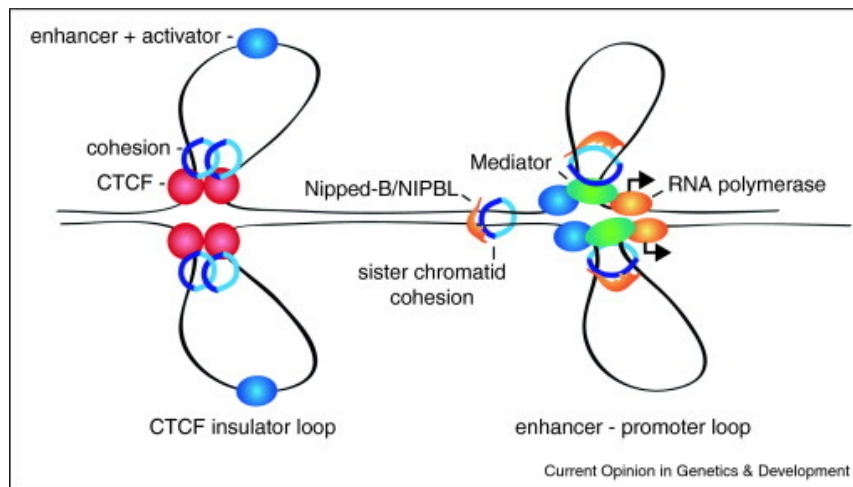
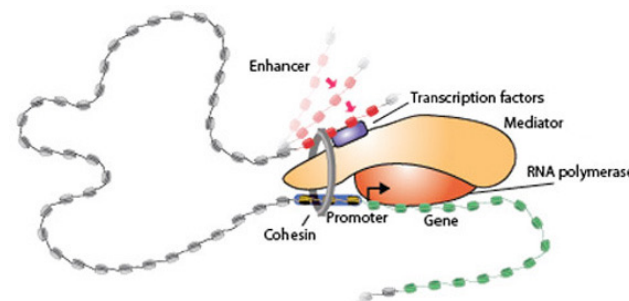
Mediator can contribute to **long-term silencing** of neuronal genes in extraneuronal cells. Starting with an active **PIC**, **REST** binding to the cognate site commits the gene to a heterochromatin fate as development proceeds. A ternary complex containing REST, the histone methyltransferase G9a and the intact Mediator is first assembled. The multipartite interactions are anchored through **the MED12 subunit of the kinase module**. Following G9a action and dimethylation of histone H3 lysine 9 (H3K9) to H3K9me2, heterochromatin protein 1 (HP1) and DNA methyltransferase 1 (DNMT1) are recruited to the site. Ultimately, the gene is embedded in transcriptionally inert heterochromatin marked by H3K9me2 and methylated DNA (CH₃). Other factors (for example, HDAC and KDM) that contribute to silencing are not shown.



Cohesin connects gene expression and chromatin architecture

Recruitment of **Mediator** to the enhancer, or to other upstream elements, results in the formation of a chromatin loop that brings together the enhancer and the promoter of the gene to be transcribed. Subsequent recruitment of **cohesin** could potentially stabilize enhancer–promoter interactions by embracing the base of the chromatin loop.

Although **cohesin** is recruited to active promoters, it also becomes associated with the DNA-binding factor **CTCF**, which has been implicated in the formation of insulator elements. Thus, cohesin is thought to have roles in transcription activation at some genes and in silencing at others.



CTCF links the architecture of the genome to its function

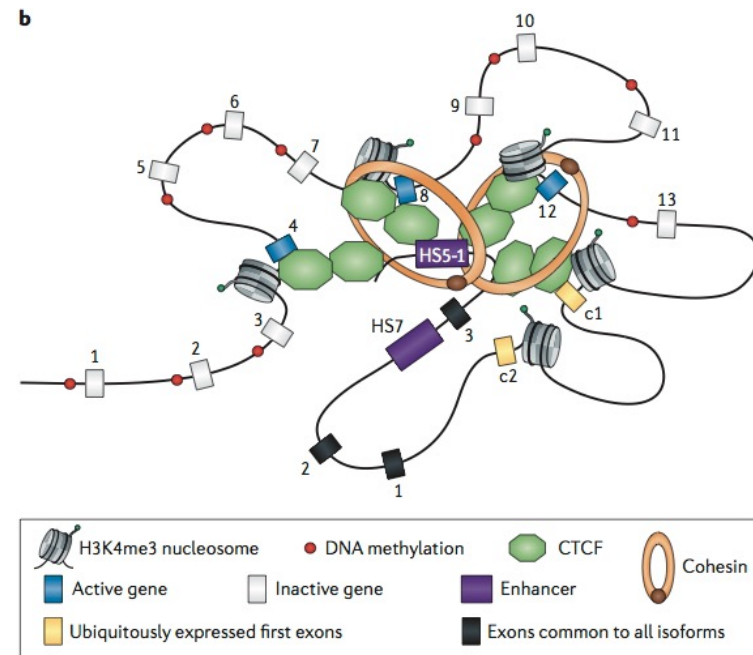
The eukaryotic genome is organized in the three-dimensional nuclear space in a specific manner that is both a cause and a consequence of its function. This organization is partly established by a special class of architectural proteins, of which **CCCTC-binding factor (CTCF)** is the best characterized.

CTCF creates boundaries between topologically associating domains in chromosomes and, within these domains, facilitates interactions between transcription regulatory sequences.

CTCF links the architecture of the genome to its function

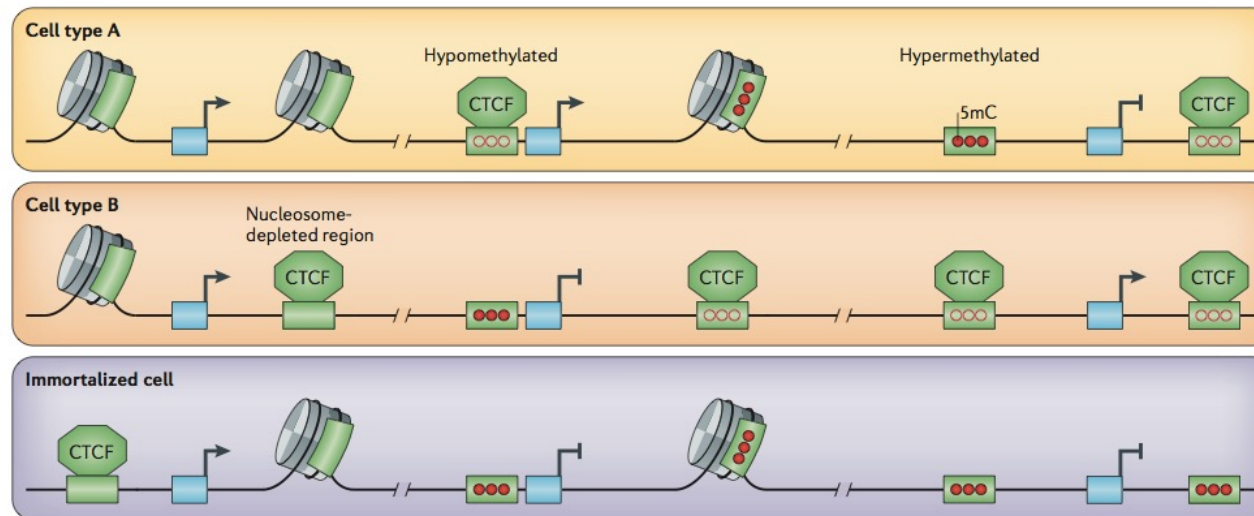
CTCF was initially classified as **insulators** -that is, a protein with the ability to interfere with enhancer–promoter communication.

CTCF mediates both inter- and intra-chromosomal interactions between distant sites in the genome. As a result of these interactions, CTCF elicits specific functional outcomes that are context dependent, determined by the nature of the two sequences brought together and by the proteins with which they interact ('architectural' protein)



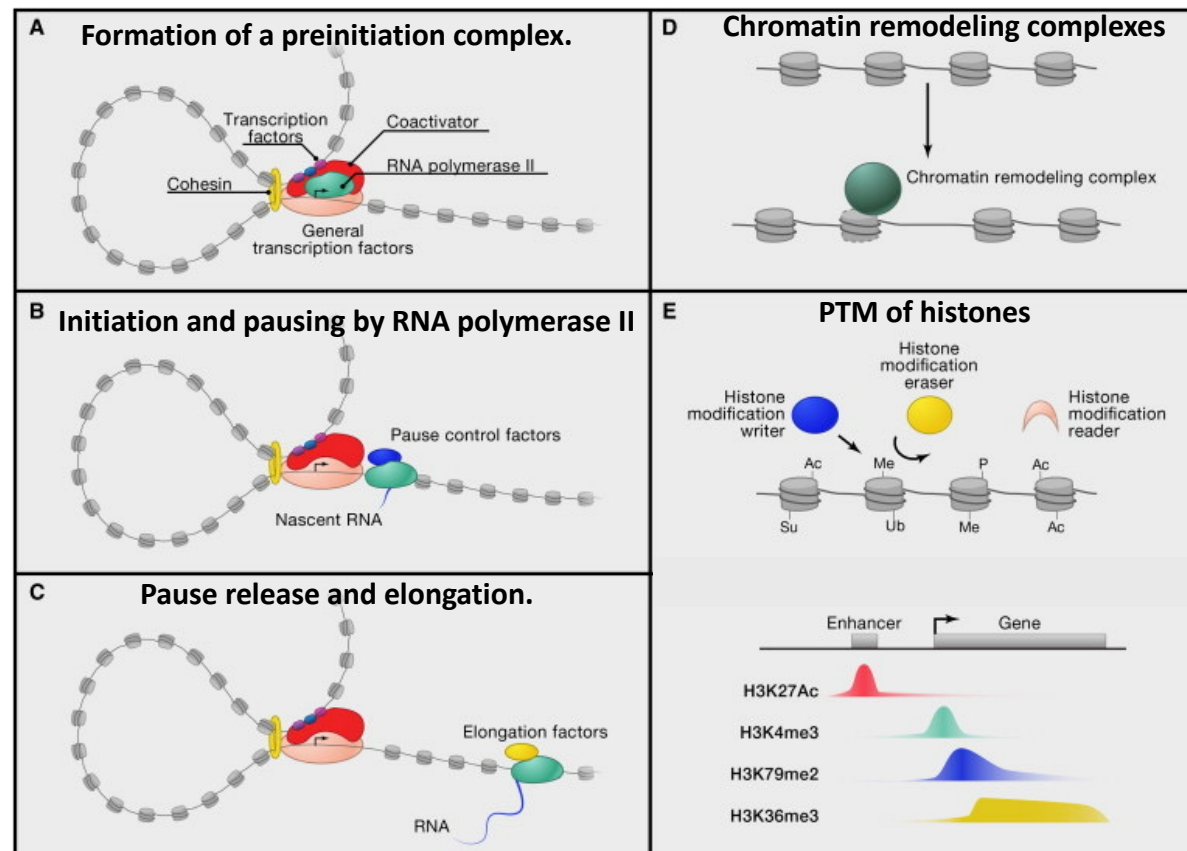
Regulation of CTCF binding to DNA

Constitutive binding sites of CTCF are present in non-methylated and nucleosome-free regions. Cell-type-specific CTCF binding is partly regulated by differential DNA methylation and nucleosome occupancy across different cell types. This suggests that cells can use ATP-dependent chromatin remodelling complexes to regulate nucleosome occupancy at specific CTCF-binding sites and control the interaction of this protein with DNA. In addition, the methylation status of cell-type-specific CTCF-binding sites may be determined by a combination of activities of *de novo* methyltransferases and TET enzymes that regulate the presence and levels of 5mC at specific sites. Immortalized cancer cell lines contain high levels of 5mC at CTCF-binding sites, which correlates with the low CTCF occupancy in these cells.



Transcriptional Control of Genes

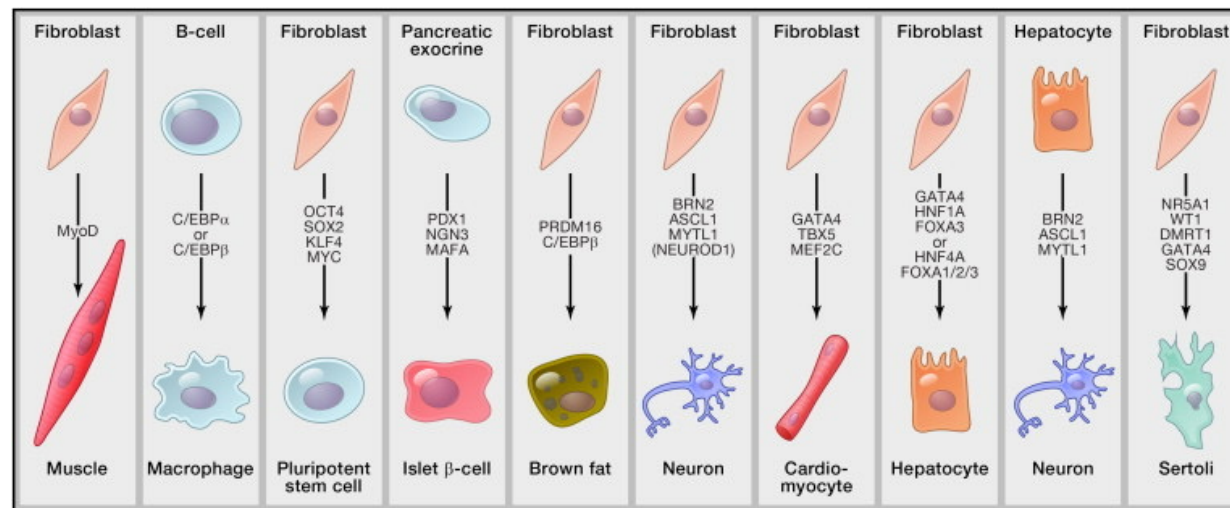
Transcriptional regulation occurs at two interconnected levels: the first involves **transcription factors and the transcription apparatus (A-C)**, and the second involves **chromatin and its regulators (D-E)**.



Master Transcriptional Regulators

The set of genes that are transcribed largely defines the cell. The gene expression program of a specific cell type includes RNA species from genes that are active in most cells (**housekeeping genes**) and genes that are active predominantly in one or a limited number of cell types (**cell-type-specific genes**).

Transcription factors that have dominant roles in the control of specific cell states and that are capable of reprogramming cell states when ectopically expressed in various cell types.

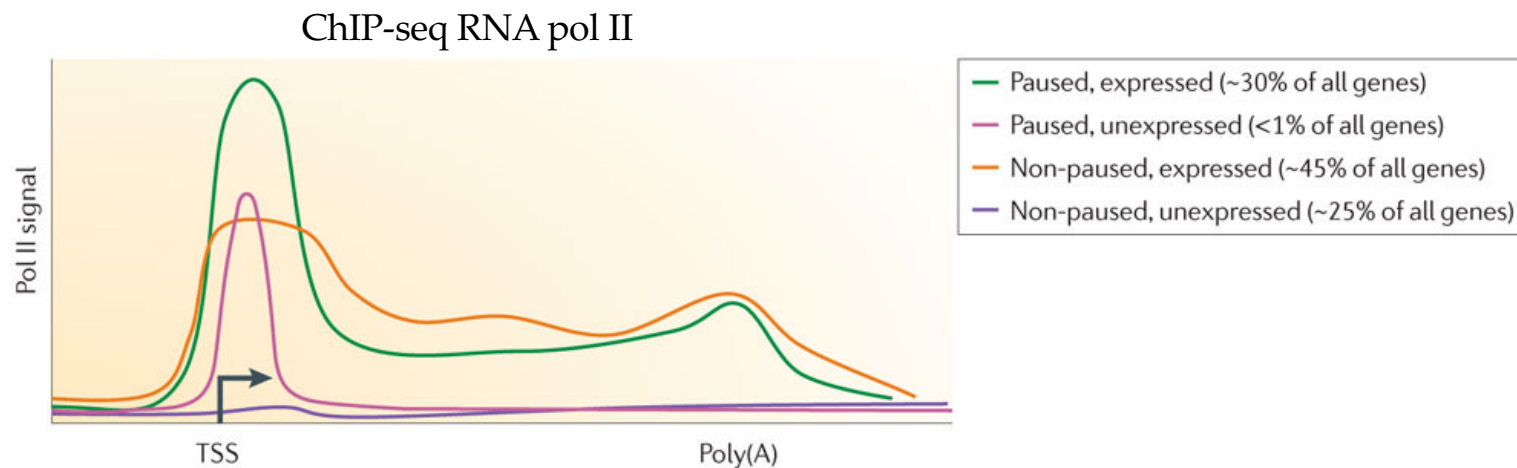


Transcription factors

- Transcription factors can be separated into two classes based on their regulatory responsibilities: **control of initiation** versus **control of elongation**. This distinction is not absolute, as some transcription factors may contribute to control of both initiation and elongation.
- Transcription factors typically bind cofactors, which are protein complexes that contribute to activation (**coactivators**) and repression (**corepressors**).
- These coactivators include **histone modifiers** (e.g. HAT) and the **Mediator complex**, among others.

Transcriptional control by RNA pol II pausing

Whereas traditional models focused solely on the events that brought RNA 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.

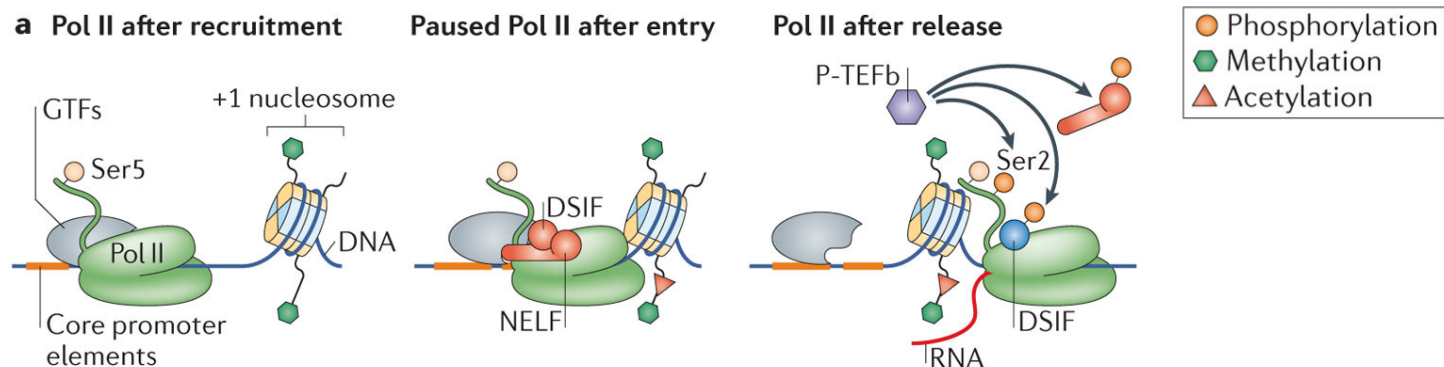


Transcriptional control by RNA pol II pausing

Once the recruited RNA Pol II molecules initiate transcription, they generally transcribe a short distance, typically 20–60 bp, and then pause. This process is controlled by the factors **DSIF** and **NELF**, which are physically associated with the paused RNA Pol II. The paused polymerases may transition to active elongation through pause release, or they may ultimately terminate transcription with release of the small RNA species.

Pause release and subsequent elongation occur through recruitment and activation of **positive transcription elongation factor b (P-TEFb)**, which phosphorylates the paused polymerase and its associated pause control factors.

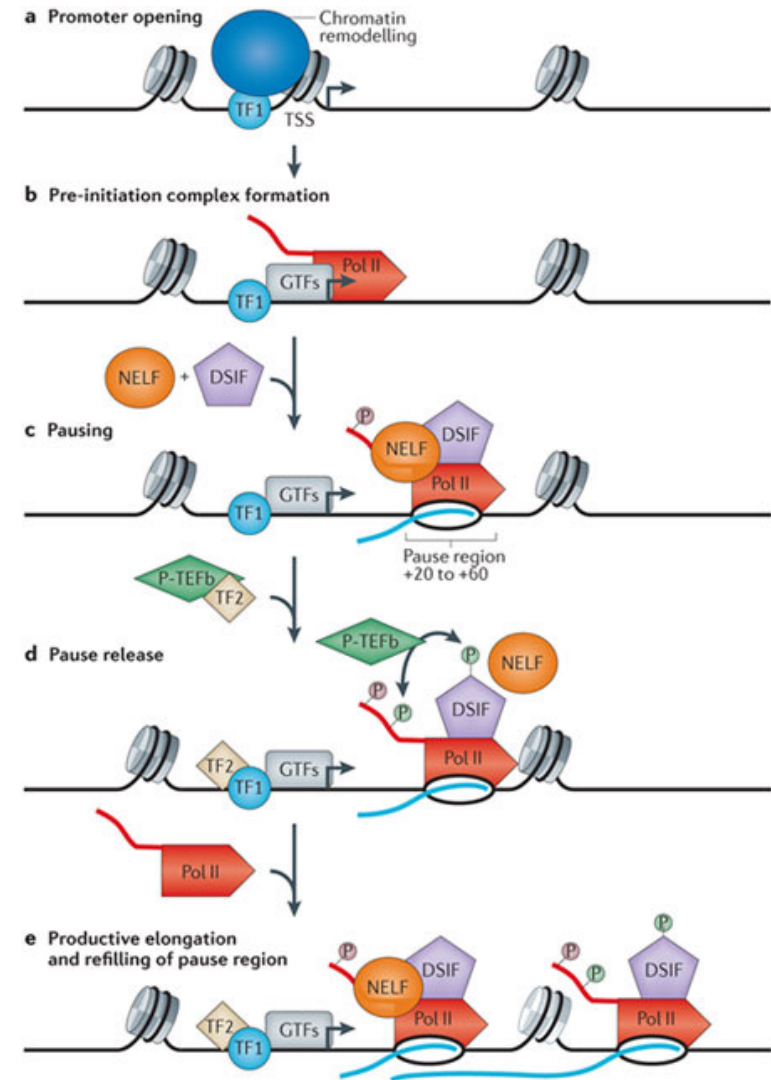
Transcription factors may stimulate P-TEFb-mediated release of RNA polymerase II from these pause sites and thus contribute to the control of transcription elongation



Establishment and release of paused Pol II

a. Promoter opening often involves binding of transcription factor (TF) that brings in chromatin remodellers 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 and Pol II. 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 CTD domain. **d. Pause release** is triggered by the recruitment of the P-TEFb kinase, either directly or indirectly by a transcription factor (TF2). P-TEFb kinase phosphorylates the DSIF–NELF complex to release paused Pol II. **e. 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.



7SK RNA regulated p-TEFb activity

7SK RNA forms the core of a small nuclear ribonucleoprotein (snRNP) complex that minimally contains P-TEFb, and the proteins HEXIM1/2 and LARP7, which control the activity of P-TEFb and stabilize 7SK RNA. The P-TEFb kinase is inactive when present in the 7SK snRNP complex and becomes active when it dissociates from the complex.

