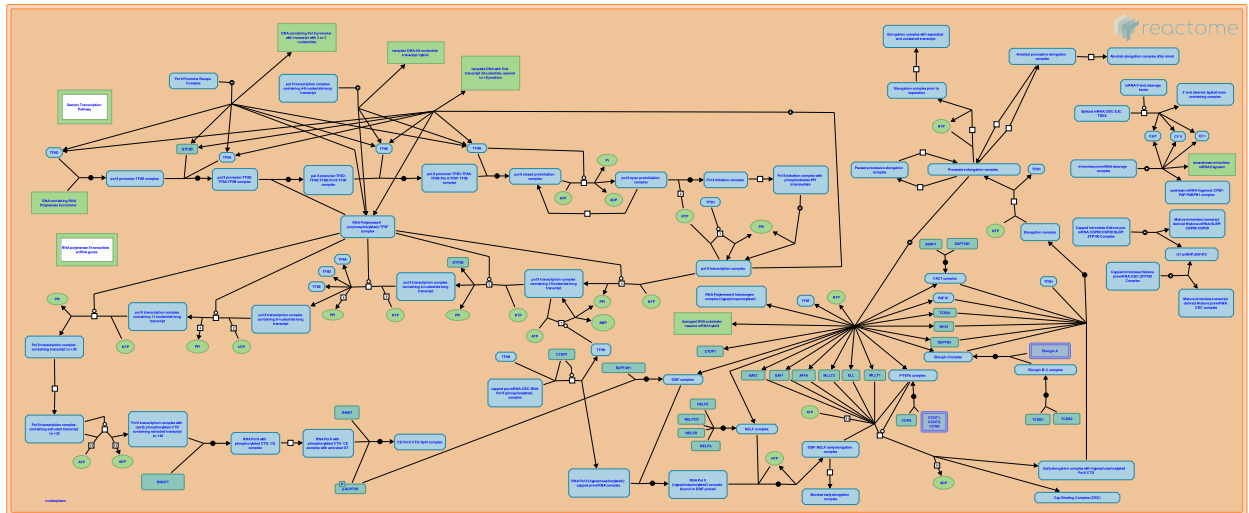


RNA Polymerase II Transcription



Conaway, JW., Conaway, RC., Freedman, LP., Gopinathrao, G., Hernandez, N., Joshi-Tope, G., May, B., Proudfoot, NJ., Reinberg, D.

European Bioinformatics Institute, New York University Langone Medical Center, Ontario Institute for Cancer Research, Oregon Health and Science University.

The contents of this document may be freely copied and distributed in any media, provided the authors, plus the institutions, are credited, as stated under the terms of [Creative Commons Attribution 4.0 International \(CC BY 4.0\) License](https://creativecommons.org/licenses/by/4.0/). For more information see our [license](https://reactome.org/about/licenses).

This is just an excerpt of a full-length report for this pathway. To access the complete report, please download it at the [Reactome Textbook](https://reactome.org/textbook).

25/04/2024

Introduction

Reactome is open-source, open access, manually curated and peer-reviewed pathway database. Pathway annotations are authored by expert biologists, in collaboration with Reactome editorial staff and cross-referenced to many bioinformatics databases. A system of evidence tracking ensures that all assertions are backed up by the primary literature. Reactome is used by clinicians, geneticists, genomics researchers, and molecular biologists to interpret the results of high-throughput experimental studies, by bioinformaticians seeking to develop novel algorithms for mining knowledge from genomic studies, and by systems biologists building predictive models of normal and disease variant pathways.

The development of Reactome is supported by grants from the US National Institutes of Health (P41 HG003751), University of Toronto (CFREF Medicine by Design), European Union (EU STRP, EMI-CD), and the European Molecular Biology Laboratory (EBI Industry program).

Literature references

- Fabregat, A., Sidiropoulos, K., Viteri, G., Forner, O., Marin-Garcia, P., Arnau, V. et al. (2017). Reactome pathway analysis: a high-performance in-memory approach. *BMC bioinformatics*, 18, 142. [↗](#)
- Sidiropoulos, K., Viteri, G., Sevilla, C., Jupe, S., Webber, M., Orlic-Milacic, M. et al. (2017). Reactome enhanced pathway visualization. *Bioinformatics*, 33, 3461-3467. [↗](#)
- Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P. et al. (2018). The Reactome Pathway Knowledgebase. *Nucleic Acids Res*, 46, D649-D655. [↗](#)
- Fabregat, A., Korninger, F., Viteri, G., Sidiropoulos, K., Marin-Garcia, P., Ping, P. et al. (2018). Reactome graph database: Efficient access to complex pathway data. *PLoS computational biology*, 14, e1005968. [↗](#)

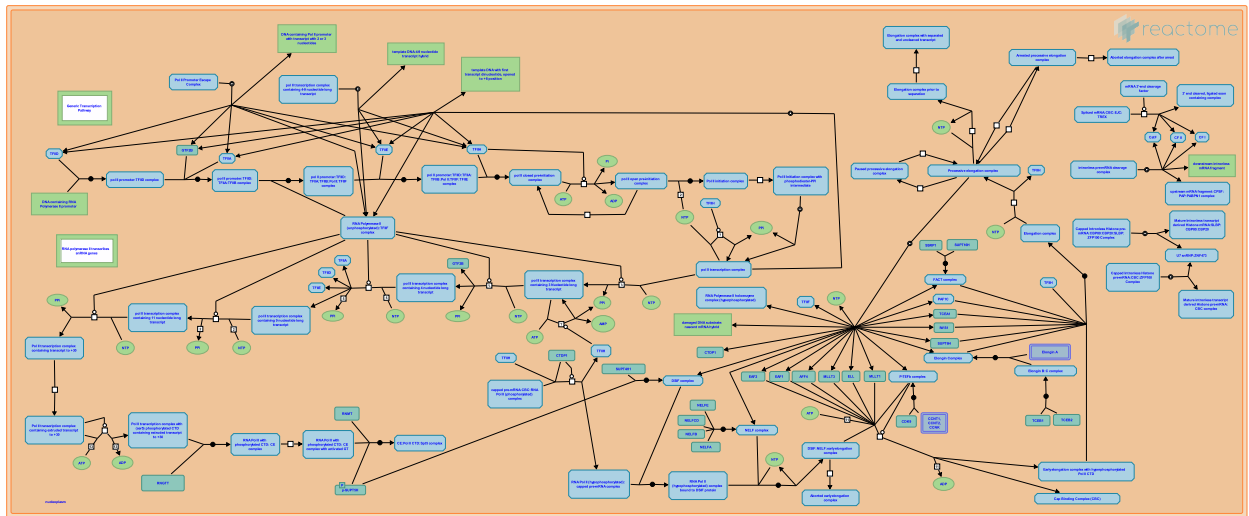
Reactome database release: 88

This document contains 9 pathways ([see Table of Contents](#))

RNA Polymerase II Transcription ↗

Stable identifier: R-HSA-73857

Compartments: nucleoplasm



RNA polymerase II (Pol II) is the central enzyme that catalyses DNA- directed mRNA synthesis during the transcription of protein-coding genes. Pol II consists of a 10-subunit catalytic core, which alone is capable of elongating the RNA transcript, and a complex of two subunits, Rpb4/7, that is required for transcription initiation. The transcription cycle is divided in three major phases: initiation, elongation, and termination. Transcription initiation include promoter DNA binding, DNA melting, and initial synthesis of short RNA transcripts. The transition from initiation to elongation, is referred to as promoter escape and leads to a stable elongation complex that is characterized by an open DNA region or transcription bubble. The bubble contains the DNA-RNA hybrid, a heteroduplex of eight to nine base pairs. The growing 3-end of the RNA is engaged with the polymerase complex active site. Ultimately transcription terminates and Pol II dissociates from the template.

Literature references

Cramer, P. (2004). Structure and function of RNA polymerase II. *Adv. Protein Chem.*, 67, 1-42. ↗

iii: They can function in either orientation in the DNA.

Combinatorial mechanisms of transcription regulation: The specific combination of TA and TR binding sites within the proximal promoter and/or distal enhancer(s) provides a "combinatorial transcription code" that mediates cell- or tissue-specific expression of the associated target gene. Each promoter or enhancer region mediates expression in a specific subset of the overall expression pattern. In at least some cases, each enhancer region functions completely independently of the others, so that the overall expression pattern is a linear combination of the expression patterns of each of the enhancer modules.

Co-Activator and Co-Repressor Complexes: DNA-bound TA and TR proteins typically recruit the assembly of specific Co-Activator (Co-A) and Co-Repressor (Co-R) Complexes, respectively, which are essential for regulating target gene transcription. Both Co-A's and Co-R's are multi-protein complexes that contain several specific protein components.

Co-Activator complexes generally contain at least one component protein that has Histone Acetyl Transferase (HAT) enzymatic activity. This functions to acetylate Histones and/or other chromatin-associated factors, which typically increases that transcription activation of the target gene. By contrast, Co-Repressor complexes generally contain at least one component protein that has Histone De-Acetylase (HDAC) enzymatic activity. This functions to de-acetylate Histones and/or other chromatin-associated factors. This typically increases the transcription repression of the target gene.

Adaptor (Mediator) complexes: In addition to the co-activator complexes that assemble on particular cell-specific TA factors, - there are at least two additional transcriptional co-activator complexes common to most cells. One of these is the Mediator complex, which functions as an "adaptor" complex that bridges between the tissue-specific co-activator complexes assembled in the proximal promoter (or distal enhancers). The human Mediator complex has been shown to contain at least 19 protein distinct components. Different combinations of these co-activator proteins are also found to be components of specific transcription Co-Activator complexes, such as the DRIP, TRAP and ARC complexes described below.

TBP/TAF complex: Another large Co-A complex is the "TBP-associated factors" (TAFs) that assemble on TBP (TATA-Binding Protein), which is bound to the TATA box present in many promoters. There are at least 23 human TAF proteins that have been identified. Many of these are ubiquitously expressed, but TAFs can also be expressed in a cell or tissue-specific pattern.

Specific Coactivator Complexes for DNA-binding Transcription Factors.

A number of specific co-activator complexes for DNA-binding transcription factors have been identified, including DRIP, TRAP, and ARC (reviewed in Bourbon, 2004, Blazek, 2005, Conaway, 2005, and Malik, 2005). The DRIP co-activator complex was originally identified and named as a specific complex associated with the Vitamin D Receptor member of the nuclear receptor family of transcription factors (Rachez, 1998). Similarly, the TRAP co-activator complex was originally identified as a complex that associates with the thyroid receptor (Yuan, 1998). It was later determined that all of the components of the DRIP complex are also present in the TRAP complex, and the ARC complex (discussed further below). For example, the DRIP205 and TRAP220 proteins were shown to be identical, as were specific pairs of the other components of these complexes (Rachez, 1999).

In addition, these various transcription co-activator proteins identified in mammalian cells were found to be the orthologues or homologues of the Mediator ("adaptor") complex proteins (reviewed in Bourbon, 2004). The Mediator proteins were originally identified in yeast by Kornberg and colleagues, as complexes associated with DNA polymerase (Kelleher, 1990). In higher organisms, Adapter complexes bridge between the basal transcription factors (including Pol II) and tissue-specific transcription factors (TFs) bound to sites within upstream Proximal Promoter regions or distal Enhancer regions (Figure 1). However, many of the Mediator homologues can also be found in complexes associated with specific transcription factors in higher organisms. A unified nomenclature system for these adaptor / co-activator proteins now labels them Mediator 1 through Mediator 31 (Bourbon, 2004). For example, the DRIP205 / TRAP220 proteins are now identified as Mediator 1 (Rachez, 1999), based on homology with yeast Mediator 1.

Example Pathway: Specific Regulation of Target Genes During Notch Signaling:

One well-studied example of cell-specific regulation of gene transcription is selective regulation of target genes during Notch signaling. Notch signaling was first identified in *Drosophila*, where it has been studied in detail at the genetic, molecular, biochemical and cellular levels (reviewed in Justice, 2002; Bray, 2006; Schweisguth, 2004; Louvri, 2006). In *Drosophila*, Notch signaling to the nucleus is thought always to be mediated by one specific DNA binding transcription factor, Suppressor of Hairless. In mammals, the homologous genes are called CBF1 (or RBPJkappa), while in worms they are called Lag-1, so that the acronym "CSL" has been given to this conserved

transcription factor family. There are at least two human CSL homologues, which are now named RBPJ and RBPJL.

In *Drosophila*, Su(H) is known to be bifunctional, in that it represses target gene transcription in the absence of Notch signaling, but activates target genes during Notch signaling. At least some of the mammalian CSL homologues are believed also to be bifunctional, and to mediate target gene repression in the absence of Notch signaling, and activation in the presence of Notch signaling.

Notch Co-Activator and Co-Repressor complexes: This repression is mediated by at least one specific co-repressor complexes (Co-R) bound to CSL in the absence of Notch signaling. In *Drosophila*, this co-repressor complex consists of at least three distinct co-repressor proteins: Hairless, Groucho, and dCtBP (*Drosophila* C-terminal Binding Protein). Hairless has been shown to bind directly to Su(H), and Groucho and dCtBP have been shown to bind directly to Hairless (Barolo, 2002). All three of the co-repressor proteins have been shown to be necessary for proper gene regulation during Notch signaling *in vivo* (Nagel, 2005).

In mammals, the same general pathway and mechanisms are observed, where CSL proteins are bifunctional DNA binding transcription factors (TFs), that bind to Co-Repressor complexes to mediate repression in the absence of Notch signaling, and bind to Co-Activator complexes to mediate activation in the presence of Notch signaling. However, in mammals, there may be multiple co-repressor complexes, rather than the single Hairless co-repressor complex that has been observed in *Drosophila*.

During Notch signaling in all systems, the Notch transmembrane receptor is cleaved and the Notch intracellular domain (NICD) translocates to the nucleus, where it then functions as a specific transcription co-activator for CSL proteins. In the nucleus, NICD replaces the Co-R complex bound to CSL, thus resulting in de-repression of Notch target genes in the nucleus (Figure 2). Once bound to CSL, NICD and CSL proteins recruit an additional co-activator protein, Mastermind, to form a CSL-NICD-Mam ternary co-activator (Co-A) complex. This Co-R complex was initially thought to be sufficient to mediate activation of at least some Notch target genes. However, there now is evidence that still other co-activators and additional DNA-binding transcription factors are required in at least some contexts (reviewed in Barolo, 2002).

Thus, CSL is a good example of a bifunctional DNA-binding transcription factor that mediates repression of specific target genes in one context, but activation of the same targets in another context. This bifunctionality is mediated by the association of specific Co-Repressor complexes vs. specific Co-Activator complexes in different contexts, namely in the absence or presence of Notch signaling.

Literature references

- Washburn, MP., Swanson, SK., Tomomori-Sato, C., Sato, S., Yao, T., Florens, LA. et al. (2005). The mammalian Mediator complex. *FEBS Lett*, 579, 904-8. ↗
- Roeder, RG. (2005). Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett*, 579, 909-15. ↗
- Malik, S., Roeder, RG. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci*, 30, 256-63. ↗
- Barolo, S., Posakony, JW. (2002). Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling. *Genes Dev*, 16, 1167-81. ↗
- Kadonaga, JT. (2004). Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors. *Cell*, 116, 247-57. ↗

Editions

2008-02-26

Reviewed

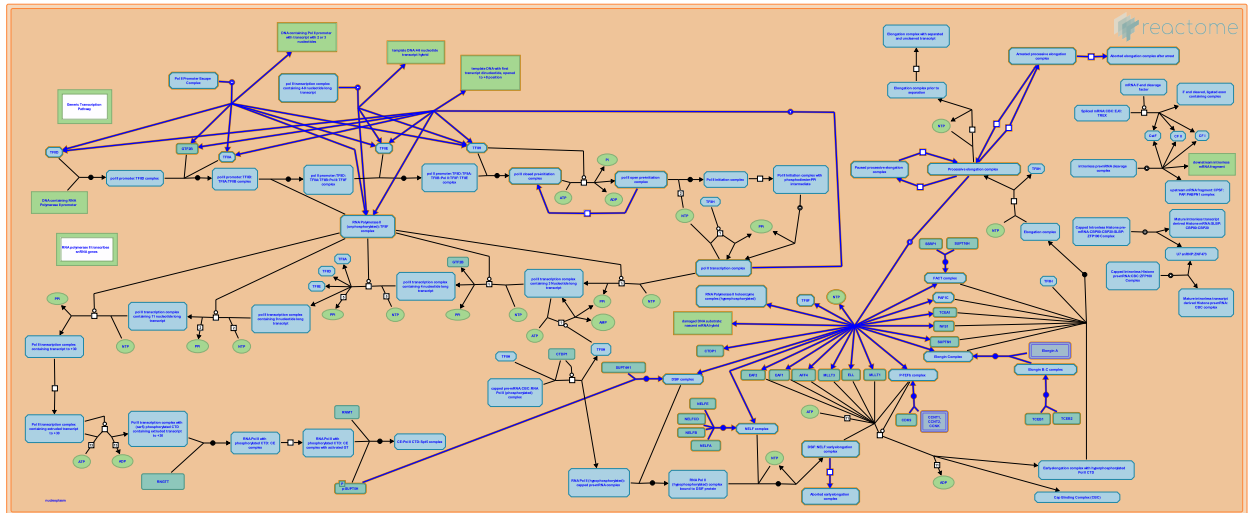
Freedman, LP.

RNA Polymerase II Pre-transcription Events ↗

Location: RNA Polymerase II Transcription

Stable identifier: R-HSA-674695

Compartments: nucleoplasm



For initiation, Pol II assembles with the general transcription factors TFIIB, TFIID, TFIIE, TFIIIF and TFIIH, which are collectively known as the general transcription factors, at promoter DNA to form the pre-initiation complex (PIC). Until the nascent transcript is about 15 nucleotides long, the early transcribing complex is functionally unstable. In the beginning, short RNAs are frequently released and Pol II has to restart transcription (abortive cycling). There is a decline in the level of abortive transcription when the RNA reaches a length of about four nucleotides, and this transition is termed escape commitment

Literature references

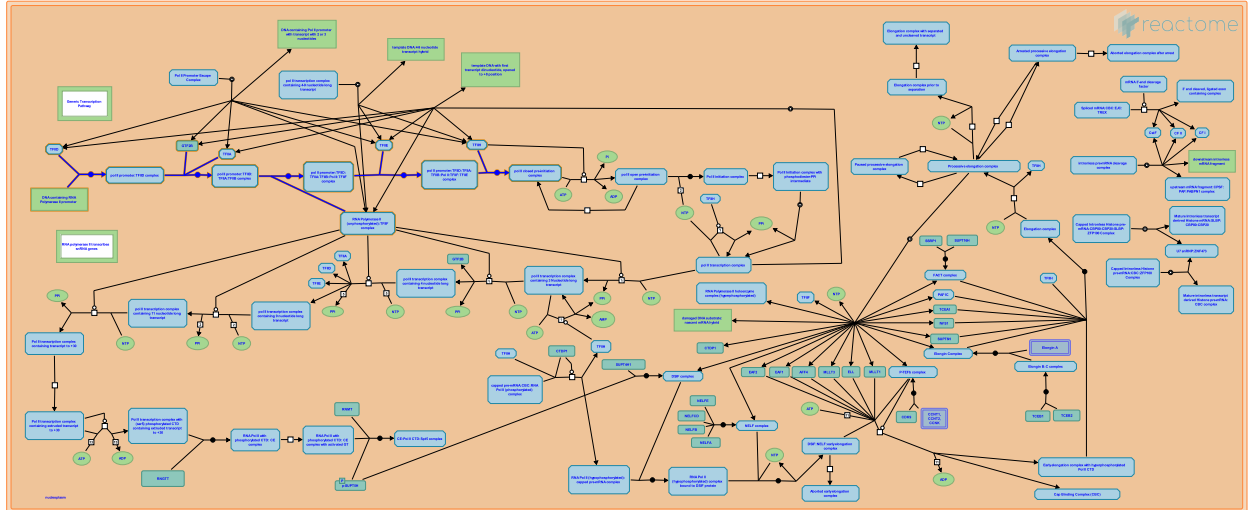
Cramer, P. (2004). Structure and function of RNA polymerase II. *Adv. Protein Chem.*, 67, 1-42. ↗

RNA Polymerase II Transcription Pre-Initiation And Promoter Opening ↗

Location: RNA Polymerase II Transcription

Stable identifier: R-HSA-73779

Compartments: nucleoplasm



Formation of the pre-initiation complex proceeds in five steps, recognition and binding of core promoter elements by TFIID, binding of TFIIA and TFIIB to the pol II promoter:TFIID complex, recruitment of RNA Polymerase II Holoenzyme by TFIIF to the pol II promoter:TFIID:TFIIA:TFIIB complex, binding of TFIIE to the growing preinitiation complex, and formation of the closed pre-initiation complex (Orphanides et al. 1997).

Literature references

Orphanides, G., Reinberg, D., Lagrange, T. (1997). The general transcription factors of RNA polymerase II. *Genes Dev*, 10, 2657-83. ↗

Editions

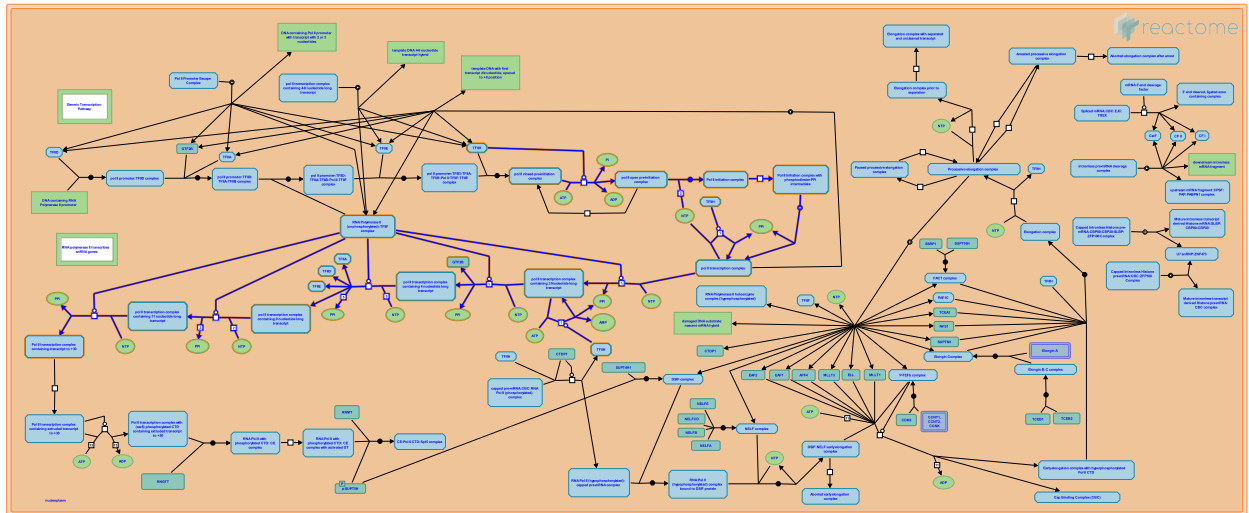
2003-09-11	Authored	Reinberg, D.
2024-03-06	Edited	Joshi-Tope, G.

RNA Polymerase II Transcription Initiation And Promoter Clearance ↗

Location: RNA Polymerase II Transcription

Stable identifier: R-HSA-76042

Compartments: nucleoplasm



The transcription cycle is divided in three major phases: initiation, elongation, and termination. Transcription initiation include promoter DNA binding, DNA melting, and initial synthesis of short RNA transcripts. Many changes must occur to the RNA polymerase II (pol II) transcription complex as it makes the transition from initiation into transcript elongation. During this intermediate phase of transcription, contact with initiation factors is lost and stable association with the nascent transcript is established. These changes collectively comprise promoter clearance.

Literature references

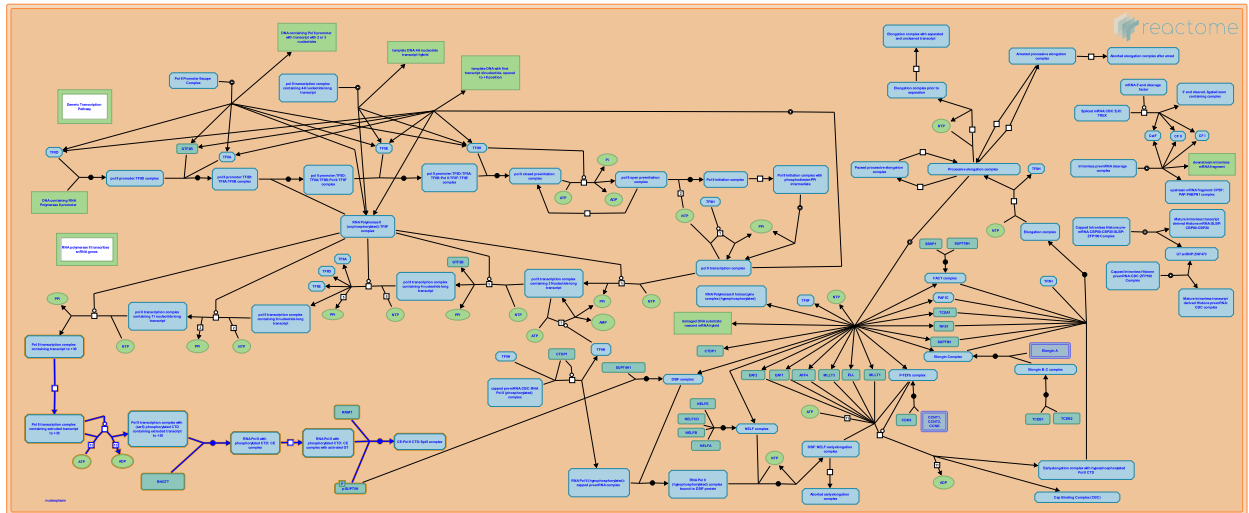
Luse, DS. (2013). Promoter clearance by RNA polymerase II. *Biochim. Biophys. Acta*, 1829, 63-8. ↗

RNA Pol II CTD phosphorylation and interaction with CE ↗

Location: RNA Polymerase II Transcription

Stable identifier: R-HSA-77075

Compartments: nucleoplasm



To facilitate co-transcriptional capping, and thereby restrict the cap structure to RNAs made by RNA polymerase II, the capping enzymes bind directly to the RNA polymerase II. The C-terminal domain of the largest Pol II subunit contains several phosphorylation sites on its heptapeptide repeats. The capping enzyme guanylyltransferase and the methyltransferase bind specifically to CTD phosphorylated at Serine 5 within the CTD. Kinase subunit of TFIIF, Cdk7, catalyzes this phosphorylation event that occurs near the promoter. In addition, it has been shown that binding of capping enzyme to the Serine-5 phosphorylated CTD stimulates guanylyltransferase activity in vitro.

Editions

2003-10-15

Authored

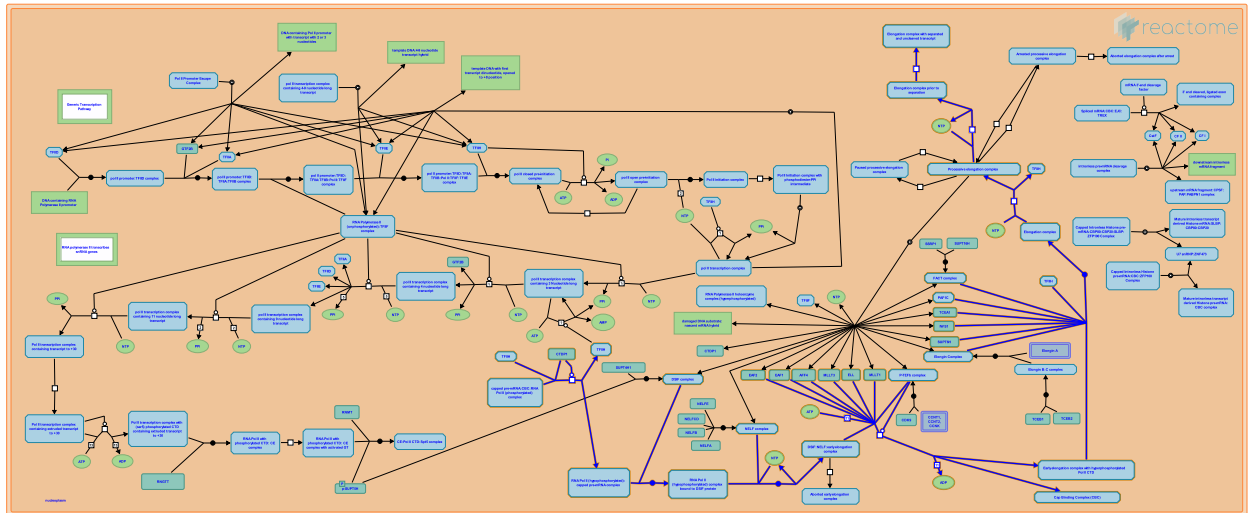
Gopinathrao, G.

RNA Polymerase II Transcription Elongation ↗

Location: RNA Polymerase II Transcription

Stable identifier: R-HSA-75955

Compartments: nucleoplasm



The mechanisms governing the process of elongation during eukaryotic mRNA synthesis are being unraveled by recent studies. These studies have led to the expected discovery of a diverse collection of transcription factors that directly regulate the activities of RNA Polymerase II and unexpected discovery of roles for many elongation factors in other basic processes like DNA repair, recombination, etc. The transcription machinery and structural features of the major RNA polymerases are conserved across species. The genes active during elongation fall under different classes like, housekeeping, cell-cycle regulated, development and differentiation specific genes etc. The list of genes involved in elongation has been growing in recent times, and include: -TFIIS, DSIF, NELF, P-Tefb etc. that are involved in drug induced or sequence-dependent arrest - TFIIF, ELL, elongin, elongator etc. that are involved in increasing the catalytic rate of elongation by altering the K_m and/or the V_{max} of Pol II -FACT, Paf1 and other factors that are involved chromatin modification - DNA repair proteins, RNA processing and export factors, the 19S proteasome and a host of other factors like Spt5-Spt5, Paf1, and NELF complexes, FCP1P etc. (Arndt and Kane, 2003). Elongation also represents processive phase of transcription in which the activities of several mRNA processing factors are coupled to transcription through their binding to RNA polymerase (Pol II). One of the key events that enables this interaction is the differential phosphorylation of Pol II CTD. Phosphorylation pattern of CTD changes during transcription, most significantly at the beginning and during elongation process. TFIIF-dependent Ser5 phosphorylation is observed primarily at promoter regions while P-Tefb mediated Ser2 phosphorylation is seen mainly in the coding regions, during elongation. Experimental evidence suggests a dynamic association of RNA processing factors with differently modified forms of the polymerase during the transcription cycle. (Komarnitsky et al., 2000). [Komarnitsky et al 2000, Arndt & Kane 2003, Shilatifard et al 2003]

Literature references

- Buratowski, S., Cho, EJ., Komarnitsky, P. (2000). Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*, 14, 2452-60. ↗
- Conaway, JW., Shilatifard, A., Conaway, RC. (2003). The RNA polymerase II elongation complex. *Annu Rev Biochem*, 72, 693-715. ↗
- Kane, CM., Arndt, KM. (2003). Running with RNA polymerase: eukaryotic transcript elongation. *Trends Genet*, 19, 543-50. ↗

Editions

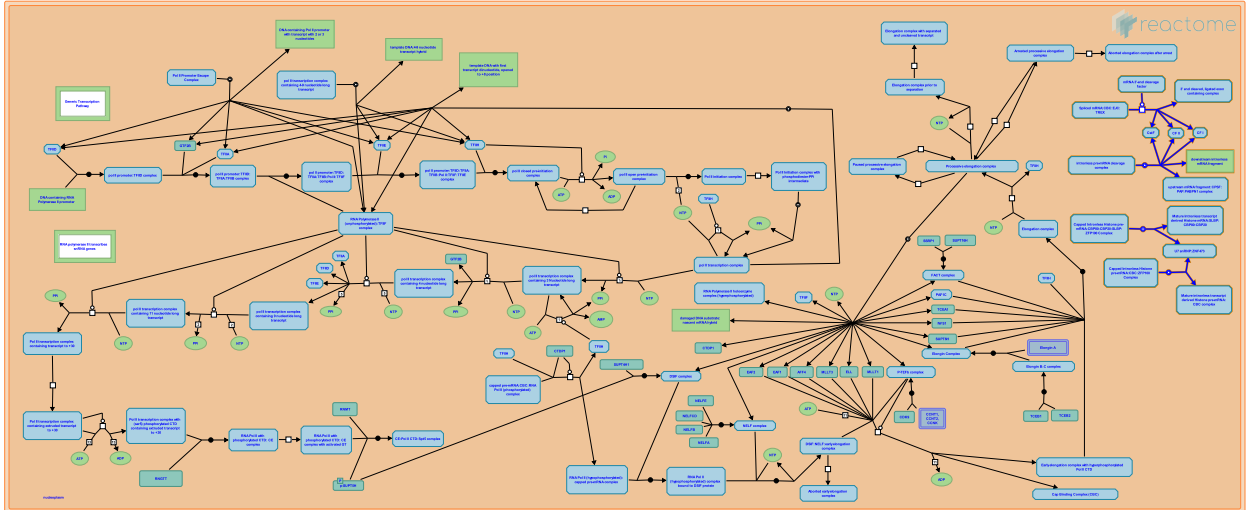
2003-09-11	Authored	Conaway, RC., Conaway, JW.
2024-03-06	Edited	Joshi-Tope, G.

RNA Polymerase II Transcription Termination ↗

Location: RNA Polymerase II Transcription

Stable identifier: R-HSA-73856

Compartments: nucleoplasm



This section includes the cleavage of both polyadenylated and non-polyadenylated transcripts.

In the former case polyadenylation has to precede transcript cleavage, while in the latter case there is no polyadenylation.

Editions

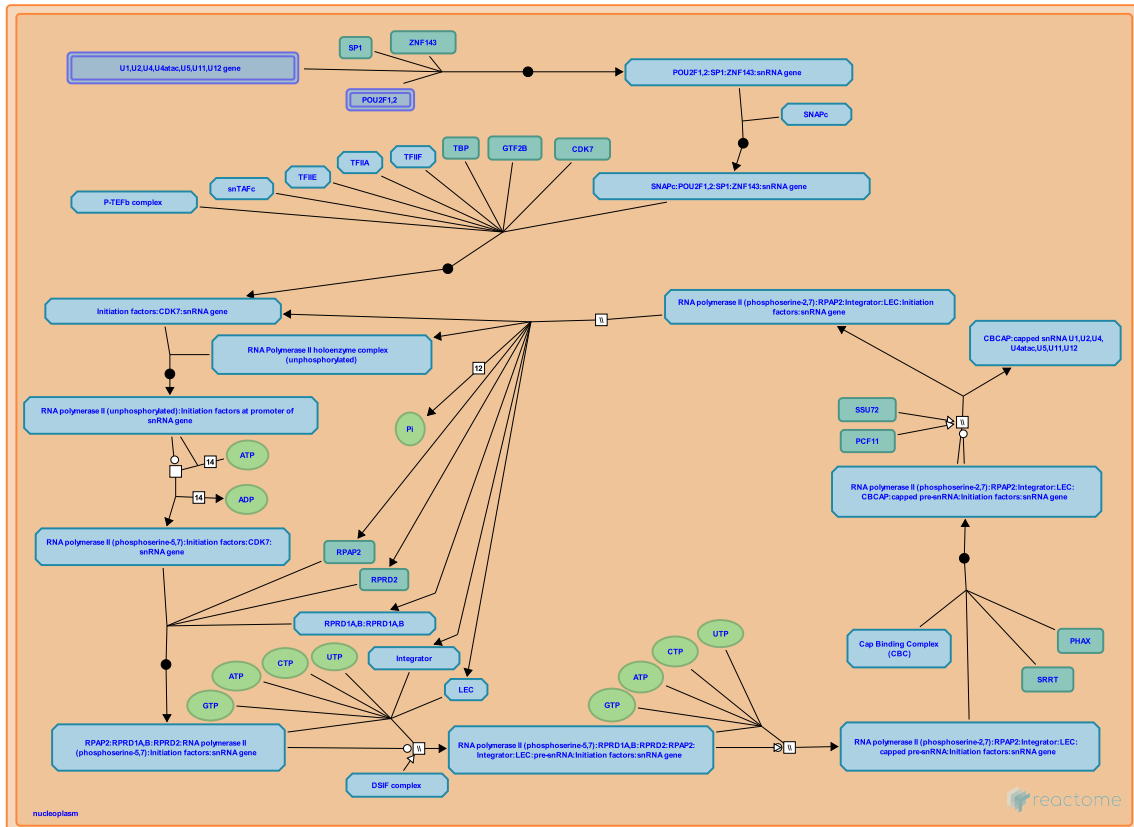
2003-09-11	Authored	Proudfoot, NJ.
2024-03-06	Edited	Joshi-Tope, G.

RNA polymerase II transcribes snRNA genes ↗

Location: RNA Polymerase II Transcription

Stable identifier: R-HSA-6807505

Compartments: nucleoplasm



Small nuclear RNAs (snRNAs) play key roles in splicing and some of them, specifically the U1 and U2 snRNAs, are encoded by multicopy snRNA gene clusters containing tandem arrays of genes, about 30 in the RNU1 cluster (Bernstein et al. 1985) and about 10-20 in the RNU2 cluster (Van Ardsell and Weiner 1984). Whereas U6 snRNA genes are transcribed by RNA polymerase III, U1,U2, U4, U4atac, U5, U11, and U12 genes are transcribed by RNA polymerase II. Transcription of the U1 and U2 genes has been most extensively studied and the other snRNA genes as well as other genes with similar promoter structures, for example the SNORD13 gene, are inferred to be transcribed by similar reactions. The snRNA genes transcribed by RNA polymerase II are distinguished from mRNA-encoding genes by the presence of a proximal sequence element (PSE) rather than a TATA box and the presence of the Integrator complex rather than the Mediator complex (reviewed in Egloff et al. 2008, Jawdeker and Henry 2008).

The snRNA genes are among the most rapidly transcribed genes in the genome. The 5' transcribed region of the U2 snRNA gene is largely single-stranded during interphase and metaphase (Pavelitz et al. 2008) and chromatin within the transcribed region is cleared of nucleosomes (O'Reilly et al. 2014). Transcriptional activation of the RNA polymerase II transcribed snRNA genes begins with binding of transcription factors to the distal sequence element (DSE) of the promoter (reviewed in Hernandez 2001, Egloff et al. 2008, Jawdeker and Henry 2008). The factors, which include POU2F1 (Oct-1), POU2F2 (Oct-2), ZNF143 (Staf) and Sp1, promote binding of the SNAPc complex (also known as PTF and PBP) to the PSE. SNAPc helps clear the gene of nucleosomes (O'Reilly et al. 2014) and recruits initiation factors (TFIIA, TFIIB, TFIIE, TFIIIF, and snTAFc:TBP) which recruit RNA polymerase II. Phosphorylation of the C-terminal domain (CTD) of RNA polymerase II (reviewed in Egloff and Murphy 2008) by CDK7 recruits RPA2 and the Integrator complex, which is required for later processing of the 3' end of the pre-snRNA transcript (reviewed in Chen and Wagner 2010, Baillat and Wagner 2015). The Little Elongation Complex (LEC) also appears to bind around the time of transcription initiation (Hu et al. 2013). As transcription proceeds, RPA2 dephosphorylates serine-5 and P-TEFb phosphorylates serine-2 of the CTD. As transcription reaches the end of the snRNA gene serine-7 of the CTD is phosphorylated. These marks serve to bind protein complexes and are required for 3' processing of the pre-snRNA (reviewed in Egloff and Murphy 2008). After transcription proceeds through the conserved 3' processing sequence of the pre-snRNA the Integrator complex cleaves the pre-snRNA. Transcription then terminates downstream in a less well characterized reaction that requires elements of the polyadenylation system.

Literature references

- Pavelitz, T., Elco, CP., Bailey, AD., Weiner, AM. (2008). Human U2 snRNA genes exhibit a persistently open transcriptional state and promoter disassembly at metaphase. *Mol. Cell. Biol.*, 28, 3573-88. [↗](#)
- Laitem, C., Murphy, S., Dienstbier, M., O'Reilly, D., Zaborowska, J., Kuznetsova, OV. (2014). Human snRNA genes use polyadenylation factors to promote efficient transcription termination. *Nucleic Acids Res.*, 42, 264-75. [↗](#)
- Eissenberg, JC., Smith, ER., Lin, C., Hu, D., Shilatifard, A., Gao, X. et al. (2013). The little elongation complex functions at initiation and elongation phases of snRNA gene transcription. *Mol. Cell*, 51, 493-505. [↗](#)
- Murphy, S., O'Reilly, D., Egloff, S. (2008). Expression of human snRNA genes from beginning to end. *Biochem. Soc. Trans.*, 36, 590-4. [↗](#)
- Van Arsdell, SW., Weiner, AM. (1984). Human genes for U2 small nuclear RNA are tandemly repeated. *Mol. Cell. Biol.*, 4, 492-9. [↗](#)

Editions

2015-11-02	Authored, Edited	May, B.
2016-02-19	Reviewed	Hernandez, N.

Table of Contents

Introduction	1
RNA Polymerase II Transcription	2
Generic Transcription Pathway	3
RNA Polymerase II Pre-transcription Events	6
RNA Polymerase II Transcription Pre-Initiation And Promoter Opening	7
RNA Polymerase II Transcription Initiation And Promoter Clearance	8
RNA Pol II CTD phosphorylation and interaction with CE	9
RNA Polymerase II Transcription Elongation	10
RNA Polymerase II Transcription Termination	11
RNA polymerase II transcribes snRNA genes	12
Table of Contents	14