MOLECULAR BIOLOGY

بسم الله الرحمن الرحيم

FINAL – Lecture 9

Transcription (Pt.2)

﴿ وَإِن تَتَوَلَّوْا يَسْتَبْدِلْ قَوْمًا غَيْرَكُمْ ثُمَّ لَا يَكُونُوا أَمْثَالَكُم ﴾

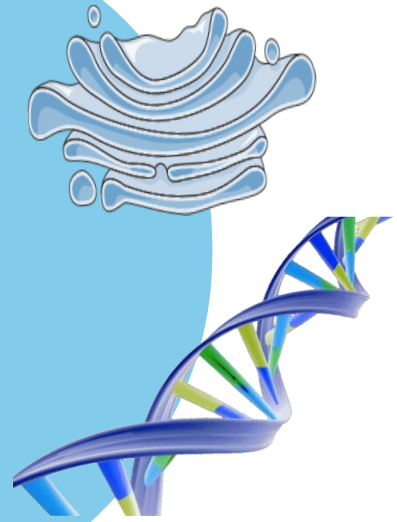
اللهم استعملنا ولا تستبدلنا

Written by :

- Mahmood Al-Absi
- Salah Budair

Reviewed by :

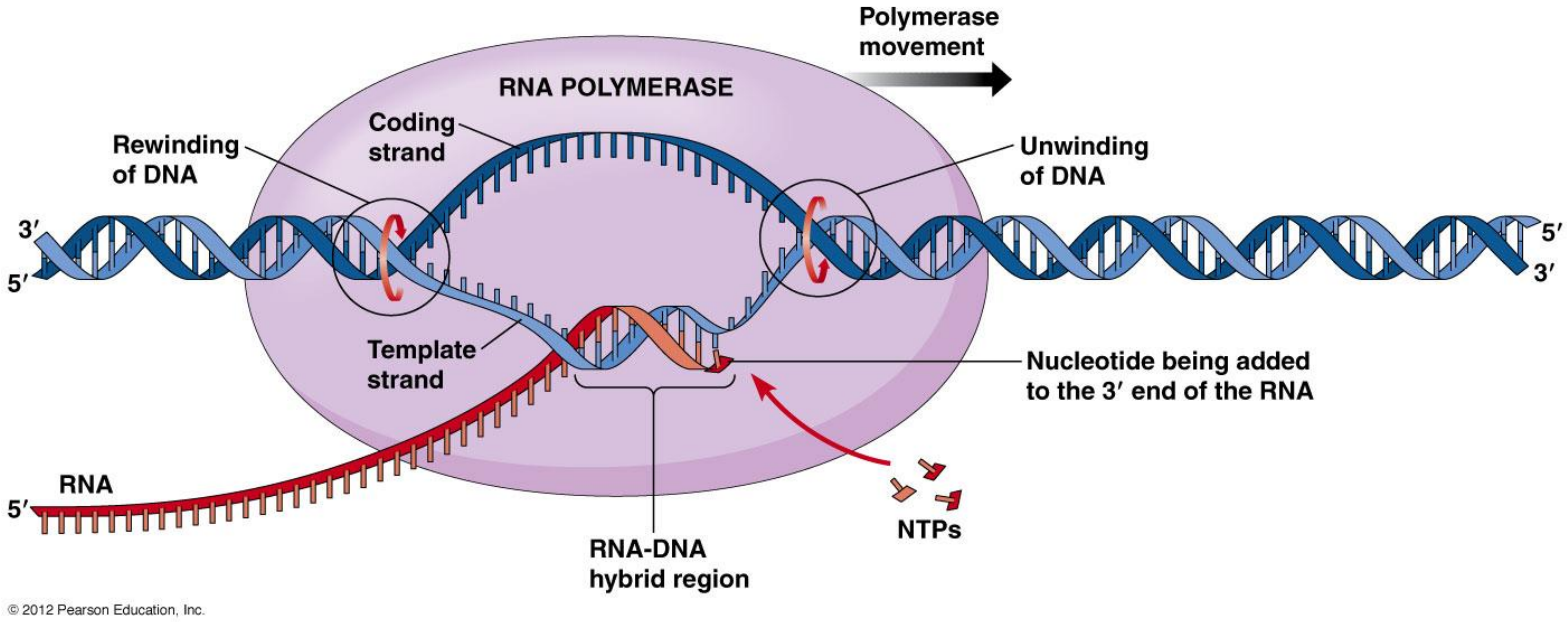
- Laith Joudeh



Quiz on the previous lecture

Mechanism of transcription (elongation)

- As the polymerase moves forward, it:
  1. Unwinds the template DNA ahead of it (like what?)
  2. Elongates the RNA
  3. Rewinds the DNA behind it
- Imagine the RNA polymerase as having six arms, two arms in front of it unwinding the DNA strand, two on the sides adding nucleotides, and two behind its back rewinding the DNA.

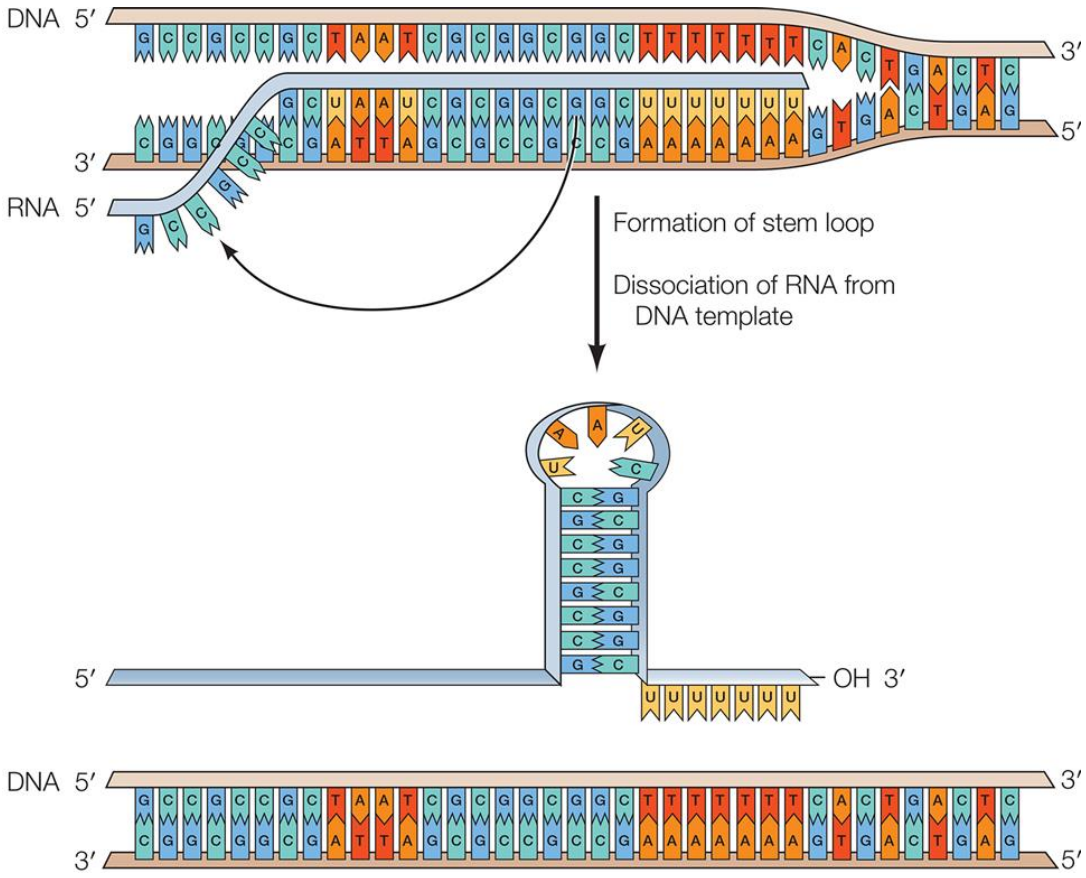


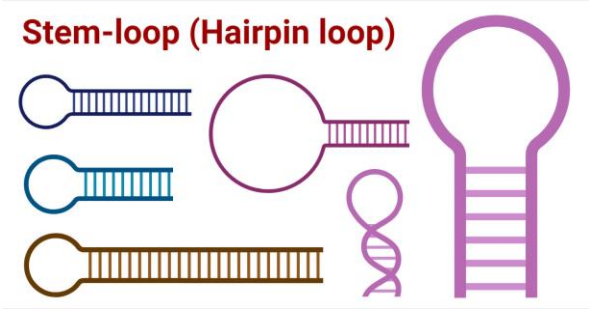
RNA synthesis continues until the polymerase encounters a termination signal (consensus sequence) where the RNA is released from the polymerase, and the enzyme dissociates from its DNA template.

Termination sequences

- The most common termination signal among genes in E. coli is a stem-loop structure that consists of a symmetrical inverted repeat of a GC-rich sequence followed by A residues (why?).
- Transcription of the GC-rich inverted repeat results in the formation of a stable stem-loop structure.
- The GC-rich sequence (a palindromic sequence) forms a stable stem-loop structure, also called a hairpin loop, in the RNA. This structure disrupts the transcription complex.
- The A-rich region in the DNA template produces a U-rich region in the RNA. These A-U base pairs are weak due to their two hydrogen bonds, causing RNA polymerase to dissociate from the DNA template.

A video is deleted from this file, as it cannot be opened because the file is in PDF format. You can retrieve the video from the original PowerPoint presentation provided by the doctor.

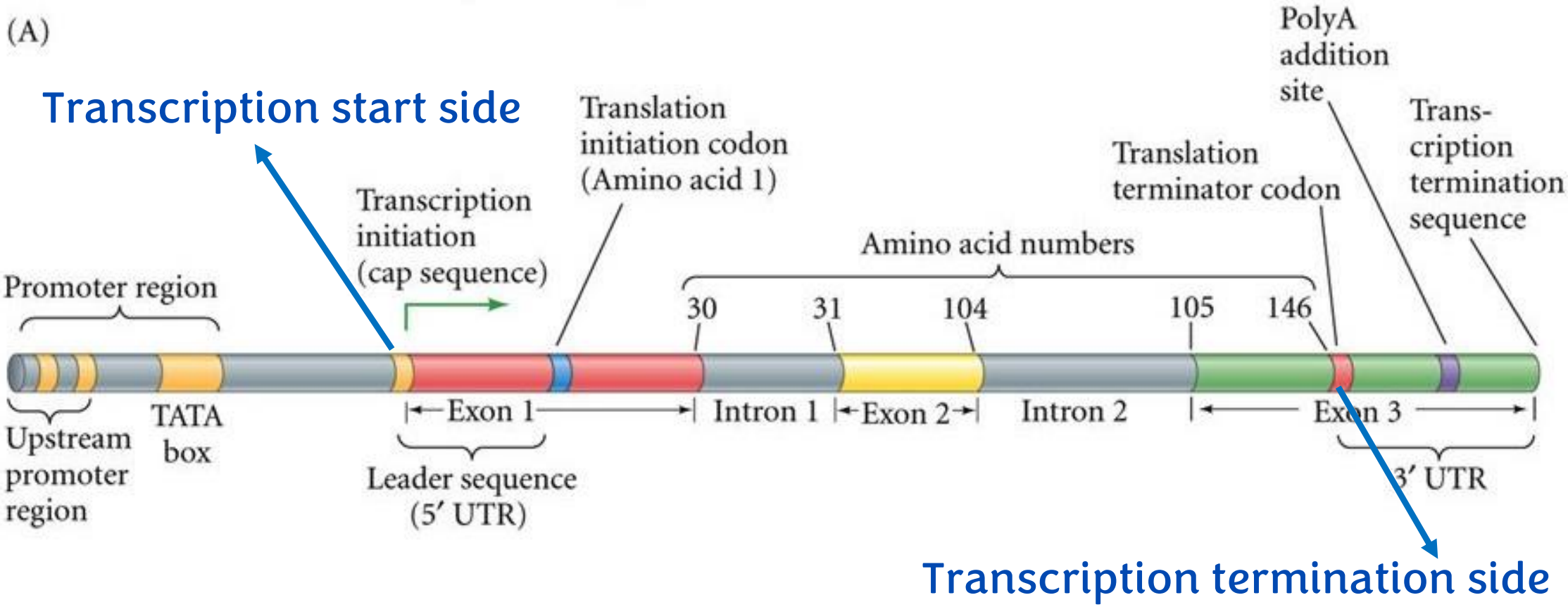




Click here to view a video explaining the termination sequence

Transcription in eukaryotes

Anatomy of a eukaryotic gene



RNA polymerases

- In contrast to bacteria, which contain a single type of RNA polymerase, eukaryotic nuclei have three, called RNA polymerase I, RNA polymerase II, and RNA polymerase III
  - RNA polymerase I transcribes rRNA genes.
  - RNA polymerase II transcribes protein-encoding genes (mRNA), long noncoding RNA (lncRNA), and microRNA (miRNA). (We will focus on this.)
  - RNA polymerase III transcribes tRNA genes and one rRNA gene.

Eukaryotic RNA polymerases

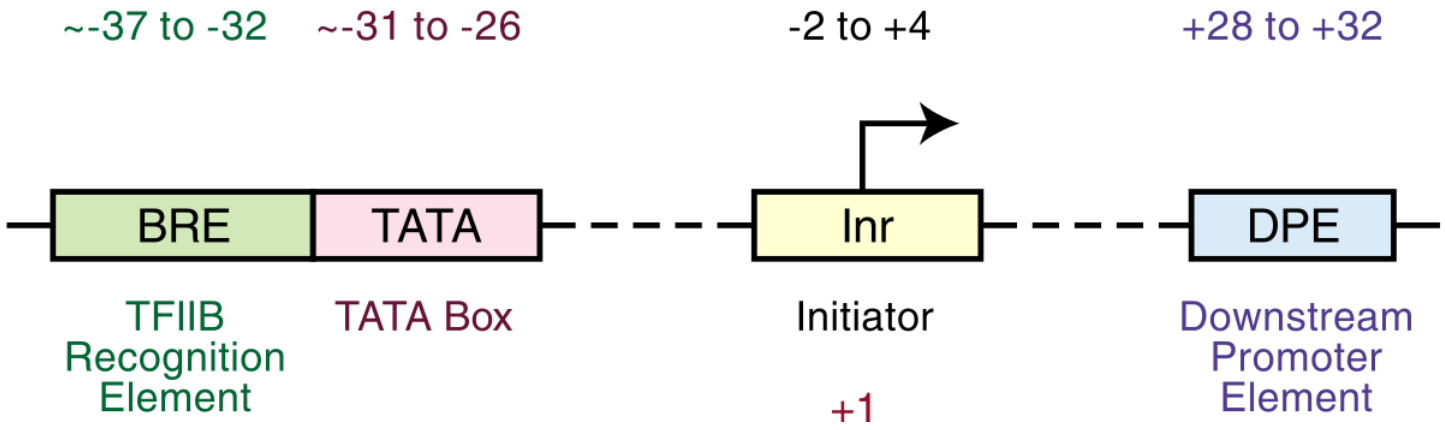
- Eukaryotic transcription initiation must deal with the packing of DNA into nucleosomes.
- While bacterial RNA polymerase is able to initiate transcription without help from any proteins, **as it does it all from recognition of the promoter to synthesis of RNA and rewinding DNA**, eukaryotic RNA polymerases require help from **general transcription factors**. **They help in transcription and are generally used for all genes.**
  - They are "general" because they assemble on all promoters used by RNA polymerase II.
  - They are designated as TFII (for transcription factor for polymerase II), and listed as TFIIA, TFIIB, and so on.

General transcription factors

- These general transcription factors **carry out the functions that prokaryotic RNA polymerase does**, such as:
  - Help position the RNA polymerase correctly at the promoter (like what? Sigma subunit of prokaryotic RNA polymerase).
    - **Help RNA polymerase in recognizing the promotor region.**
  - Aid in pulling apart the two strands of DNA to allow transcription to begin (like what? Helicase enzyme).
  - Push the RNA polymerase forward to begin transcription.

Core components of promoters

- The promoter region in eukaryotic cells is complex.
- Not all of these sequences exist at once, but genes can have a combination of these promoter elements.
- This is called the core promoter region. ➔ It is where the RNA polymerase and general transcription factors bind.



Remember:

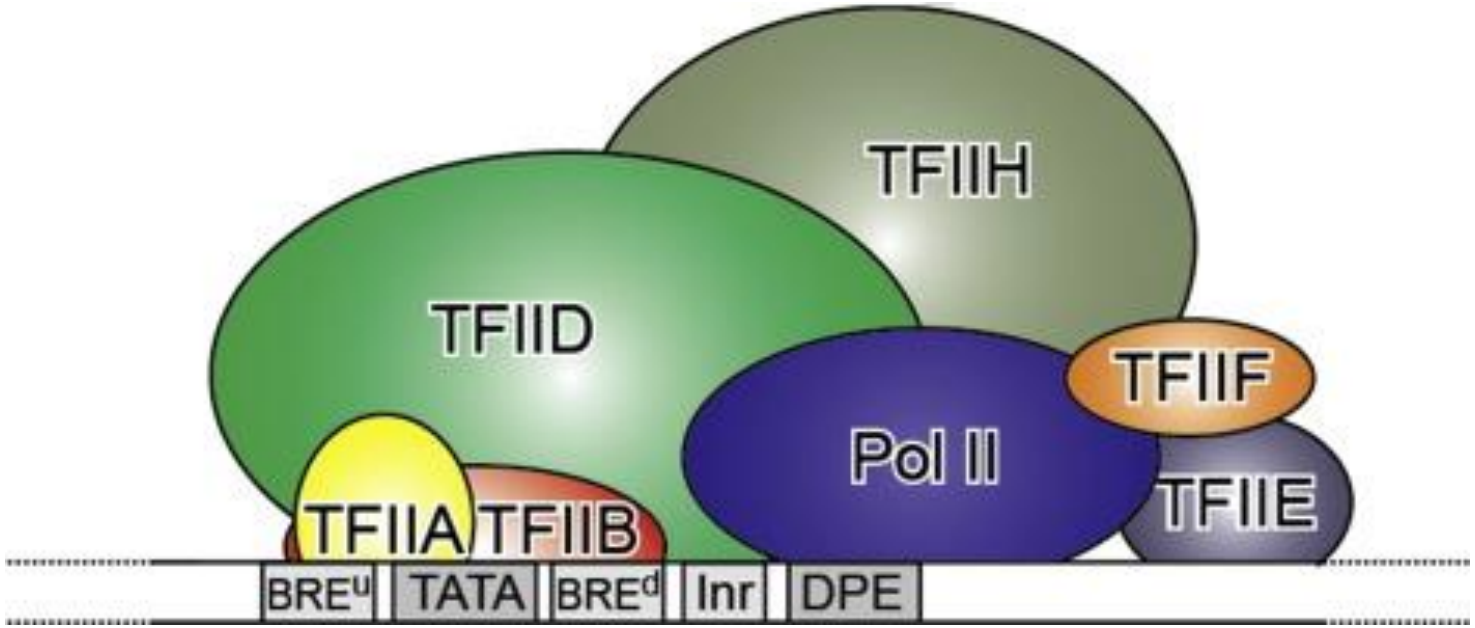
- Elements: DNA sequences
- Factors: Proteins

Promoter Elements, Simplified

- Humans have multiple promoter regions, differing in their structure and position compared to bacteria. Some genes may have a single promoter or a combination of promoter sequences. In humans, promoters are typically located upstream of the transcription start site (TSS), although regulatory elements such as enhancers can occur within genes.
- In bacteria, the promoter is upstream of the TSS and includes conserved regions, such as the -10 and -35 sequences, which guides RNA polymerase binding.
- Human promoters are more complex and contain specific elements that aid in transcription initiation:
    1. TFIIB Recognition Element (BRE): A region upstream of the TSS that binds transcription factor IIB.
    2. TATA Box: A sequence rich in thymine (T) and adenine (A) nucleotides, located upstream of the TSS, important for RNA polymerase II binding.
    3. Initiator (Inr): A core promoter element that includes the transcription start site (+1 nucleotide). It encompasses the transcription start site.
    4. Downstream Promoter Element (DPE): Found downstream of the TSS, it enhances transcription but it is not located within the coding sequence of the gene.

Formation of preinitiation complex

- Formation of a transcription complex is initiated by the binding of the transcription factor TFIID complex. **(It is the first subunit that binds).**
- Other subunits bind to the promoter, followed by binding of the polymerase.
- TFIIH then binds.

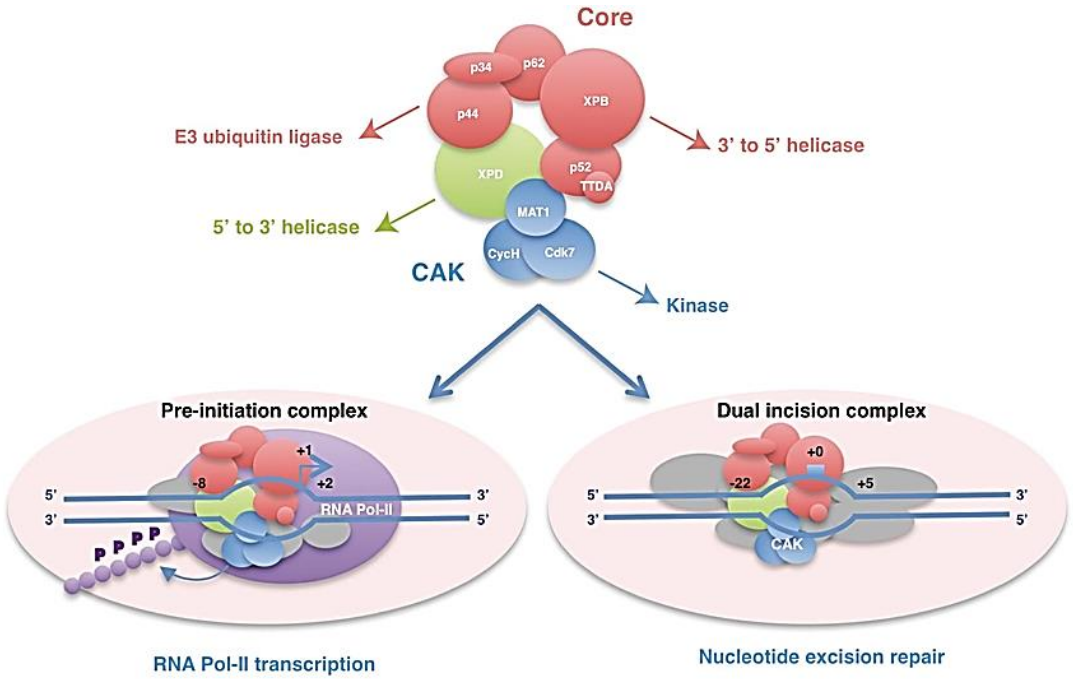


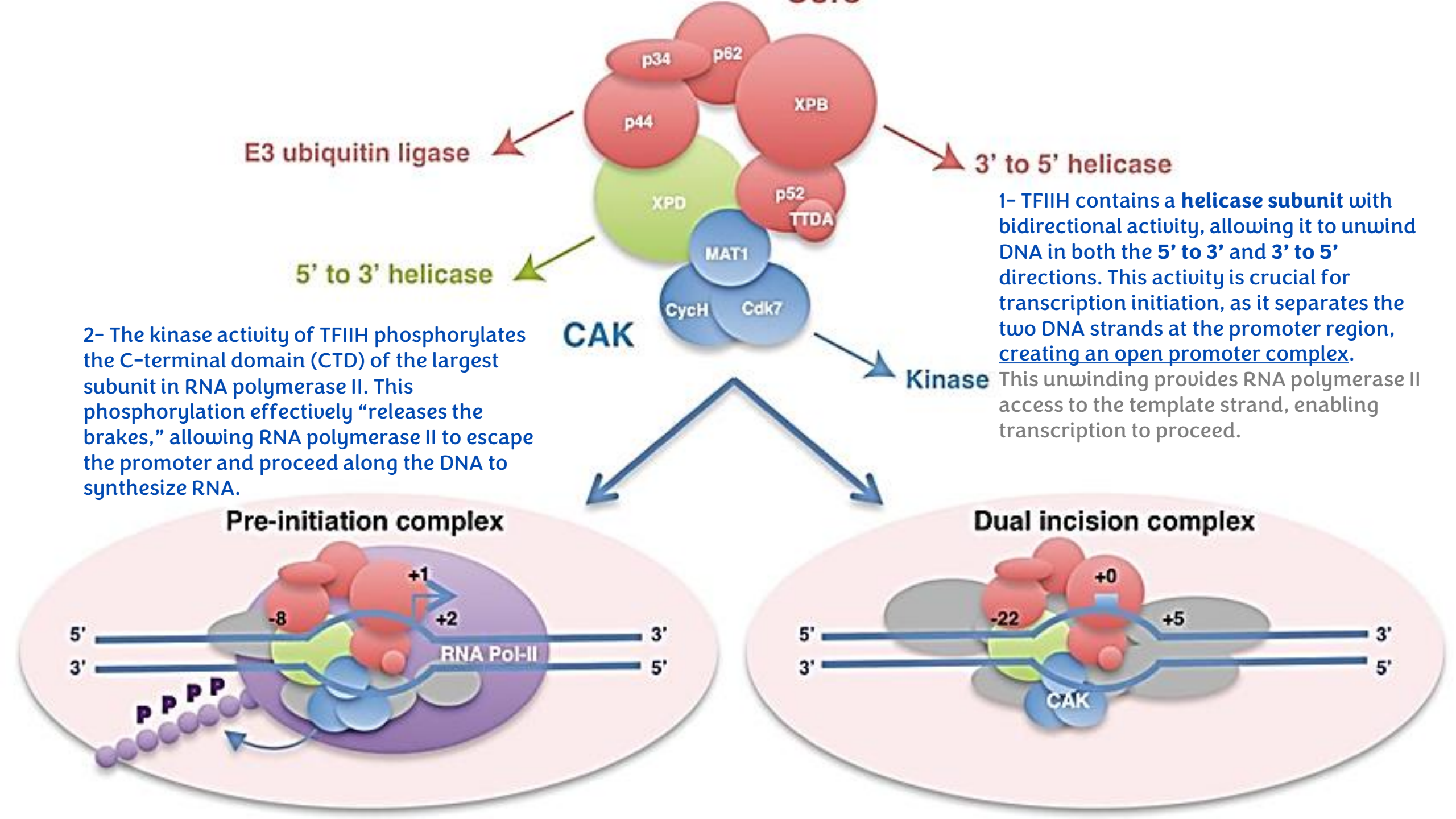
The Preinitiation Complex

- Before RNA polymerase II binds to the DNA, the **preinitiation complex (PIC)** is assembled. The PIC is a collection of proteins that interact to initiate transcription. The process begins with **transcription factor II D (TFIID)**, which is the first protein to recognize the promoter region. TFIID binds to the TATA box or other core promoter elements, signaling the presence of a gene that requires transcription.
- Once TFIID is bound, additional transcription factors and components of the PIC associate in a stepwise manner, including **TFIIB, TFIIE, TFIIF,** and others. After the assembly of these components, **RNA polymerase II** binds to the promoter, positioning itself at the transcription start site (TSS). Finally, **TFIIH** binds to the complex.

TFIIH

- TFIIH is a multi-subunit factor that has two important enzymatic activities:
    1. A helicase activity, which unwind DNA around the transcription start site and is catalyzed by two subunits (XPB and XPD proteins), which are also required for nucleotide excision DNA repair.
    2. A kinase activity that phosphorylates the C-terminal domain (CTD) of the largest subunit of RNA polymerase II.
    - Phosphorylation releases the polymerase from the preinitiation complex and leads to the initiation of transcription.



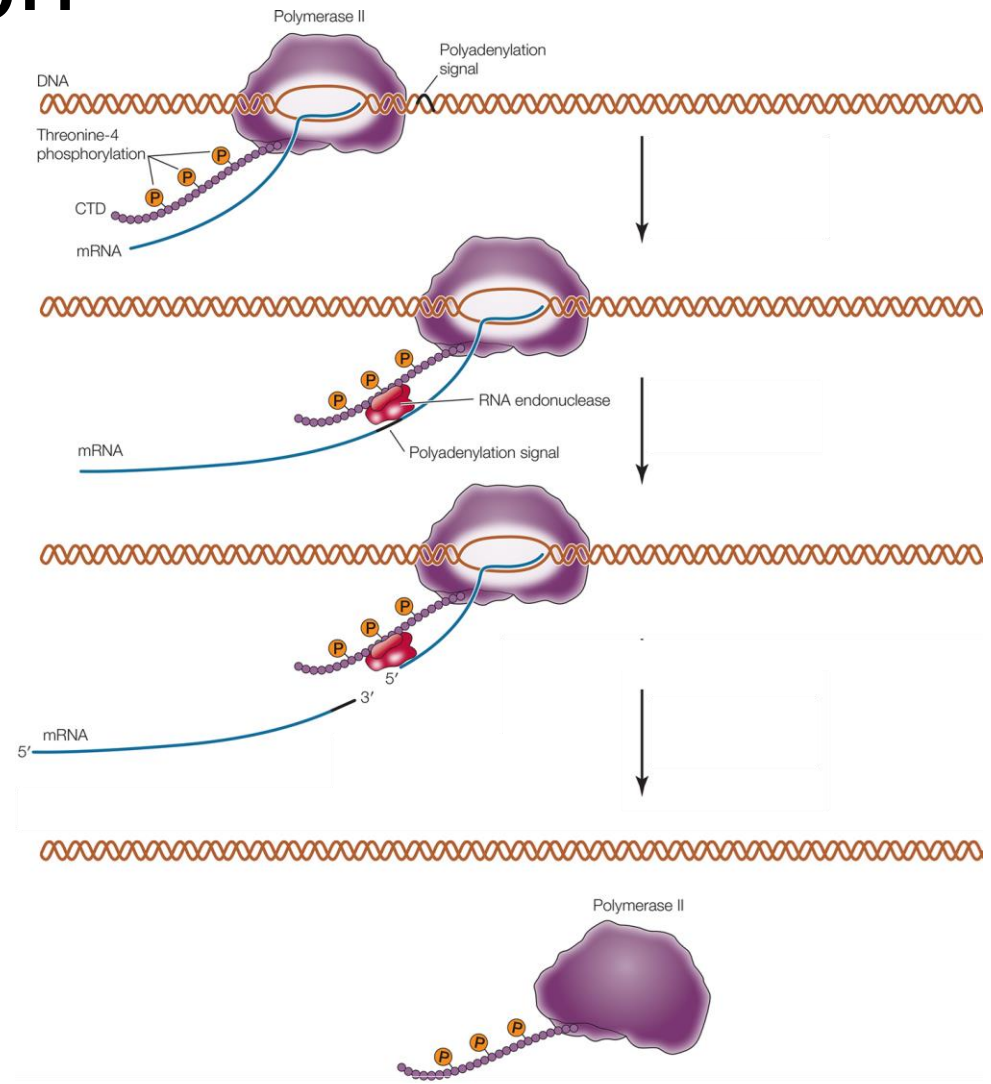


1- TFIIH contains a **helicase subunit** with bidirectional activity, allowing it to unwind DNA in both the **5' to 3'** and **3' to 5'** directions. This activity is crucial for transcription initiation, as it separates the two DNA strands at the promoter region, creating an open promoter complex. This unwinding provides RNA polymerase II access to the template strand, enabling transcription to proceed.

2- The kinase activity of TFIIH phosphorylates the C-terminal domain (CTD) of the largest subunit in RNA polymerase II. This phosphorylation effectively "releases the brakes," allowing RNA polymerase II to escape the promoter and proceed along the DNA to synthesize RNA.

Termination of Transcription

- Protein-coding genes have a polyadenylation signal at their 3' end.
- The signal is also transcribed and is recognized by an RNA endonuclease that cleaves the polyadenylation signal within the nascent transcript.
- The RNA and the polymerase are released from the DNA template.

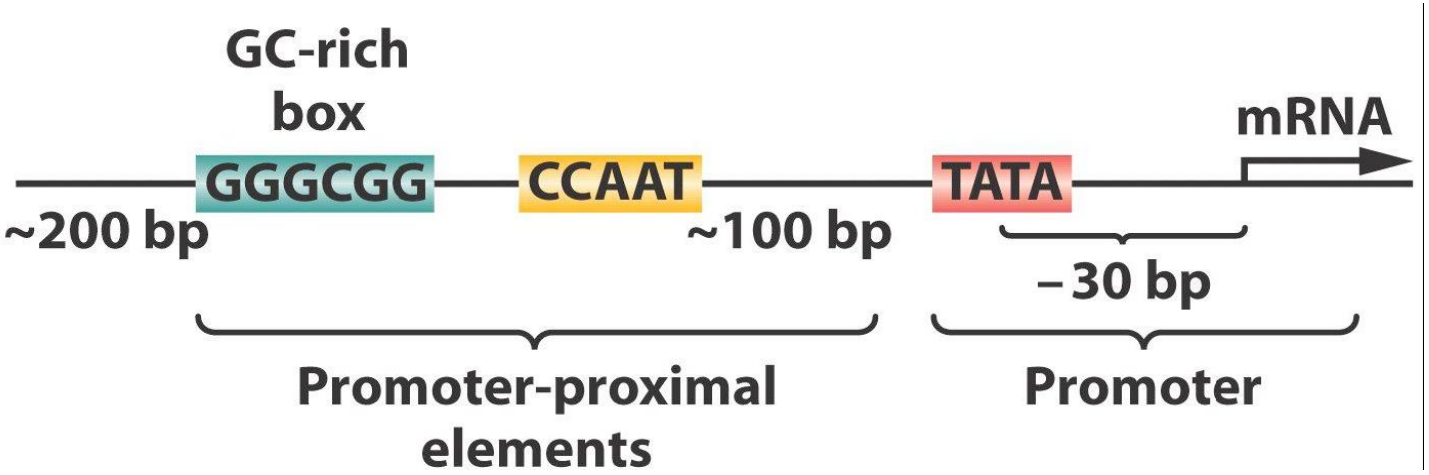


Termination of Transcription

- In eukaryotic cells, RNA elongation continues until a termination sequence is reached. Unlike in prokaryotes, where termination often involves the formation of a stem-loop structure, eukaryotes rely on a **polyadenylation sequence** to signal the end of transcription.
- The polyadenylation sequence is located at the end of the gene and directs RNA polymerase II to prepare for termination. RNA polymerase II transcribes past this sequence, but specific proteins bound to the **phosphorylated C-terminal domain (CTD)** of RNA polymerase recognize and interact with the polyadenylation sequence.
- Among these proteins, an **endonuclease** cleaves the RNA molecule at a specific site downstream of the polyadenylation signal. This cleavage marks the conclusion of transcription. The cleaved RNA is then processed further, with the addition of a **poly(A) tail** at its 3' end, which is critical for RNA stability and export from the nucleus.

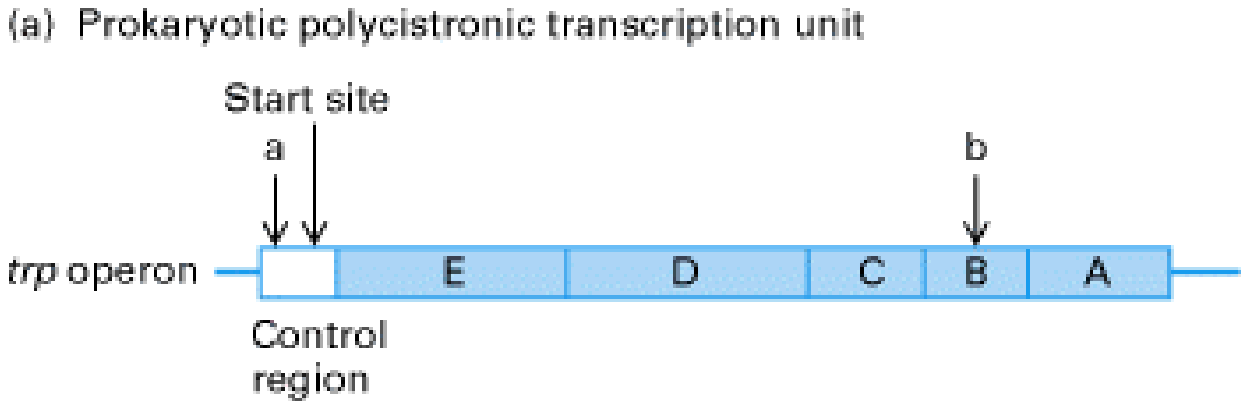
Promoter-proximal elements

- These are upstream of the core promoter region and sometime called response elements.
- They are important for strong expression (versus basal expression).
- They can be shared among different genes (gene-specific) that participate in a similar mechanism or needed for a particular purpose (example: production of enzymes for metabolism of glucose).
  - Alternative to operons!



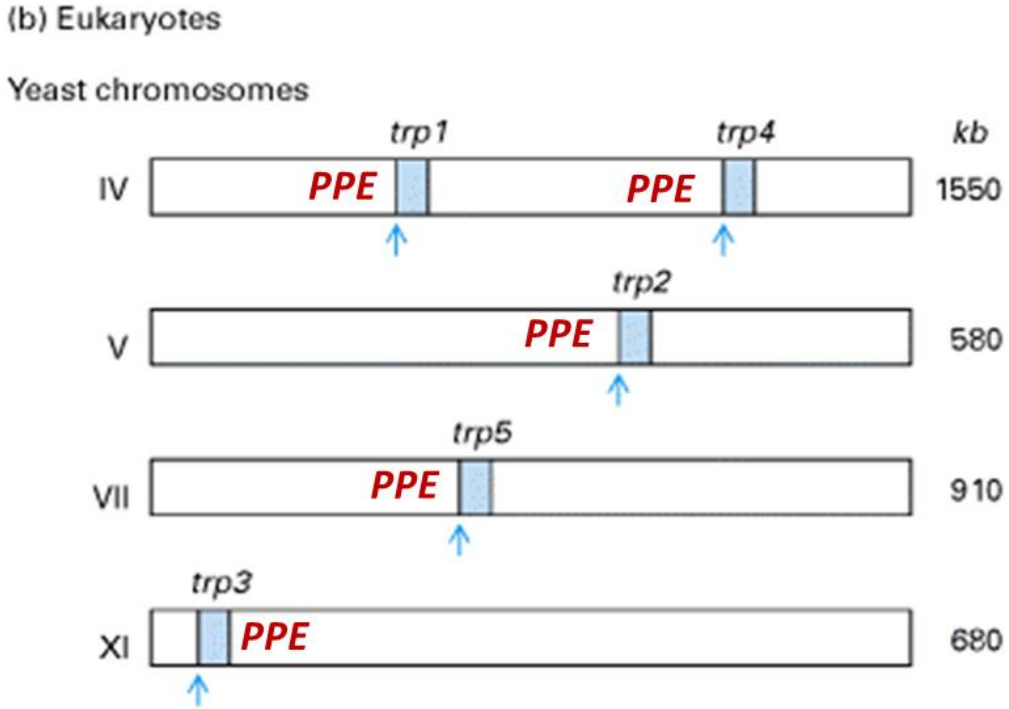
Operon vs. Proximal-promoter elements

Prokaryotes



Prokaryotes have operons, where multiple genes coding for the same purpose are transcribed together into a single mRNA at the same time. This allows for the coordinated regulation of genes involved in the same pathway.

Eukaryotes

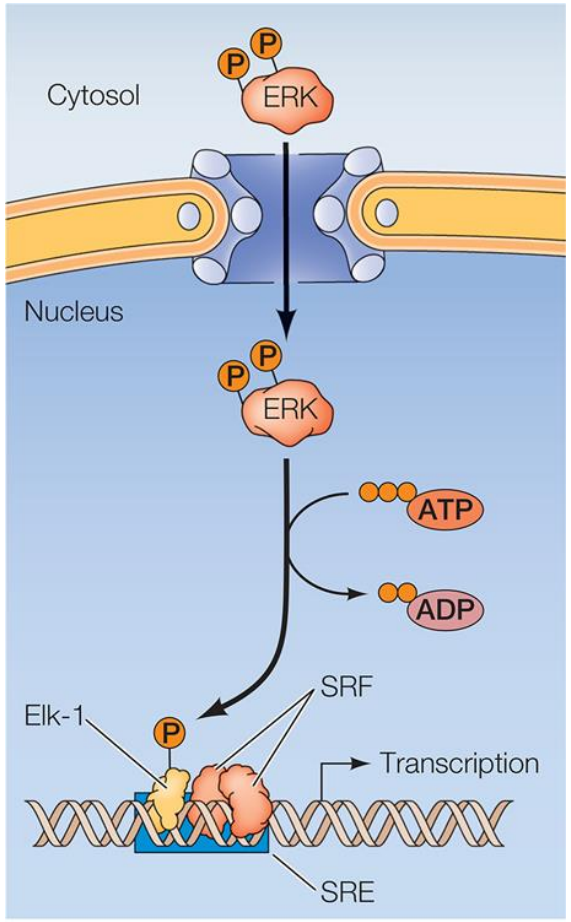


PPEs Simplified

- In eukaryotes, proximal promoter elements (PPEs) enable coordinated transcription of genes across different chromosomes, such as those involved in tryptophan synthesis. These genes share similar PPEs, which facilitate their coordinated activation.
- Specific transcription factors bind to the shared PPEs, recruiting RNA polymerase II and triggering synchronized transcription of these genes, ensuring their expression occurs together to fulfill a common function.

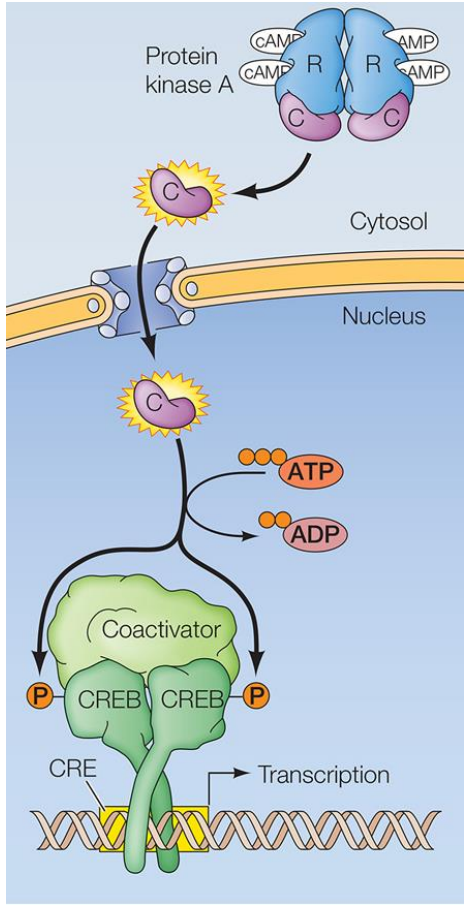
Examples of response elements

Serum response element (SRE)



Stimulation of cell proliferation induced by growth factors.

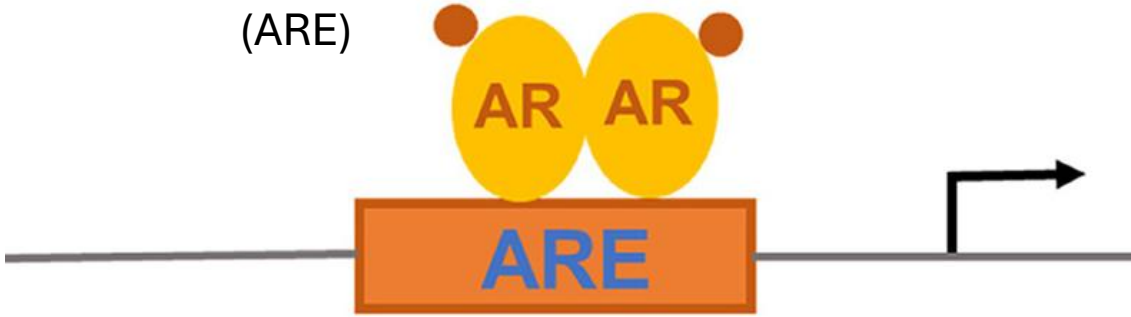
cAMP response element (CRE)



Regulation of proliferation, survival, and differentiation

The signaling pathways mentioned release intermediates, such as activated transcription factors, that bind to identical **promoter-proximal elements (PPEs)** in multiple genes. This allows for the coordinated transcription of different genes, enabling the production of proteins that collectively achieve a specific biological function.

Androgen response element (ARE)



Regulation of metabolism and differentiation

Examples of Response Elements

- **First: Serum Response Element (SRE)**

  The **serum response element (SRE)** is a promoter-proximal element (PPE) found in specific genes. Its activation begins when a growth factor binds to a receptor, typically a **receptor tyrosine kinase**, initiating a signaling cascade. For example, the **RAS-RAF-MEK pathway** activates **ERK**, which then translocates to the nucleus. Inside the nucleus, ERK activates transcription factors bound to the SRE, promoting gene transcription, cell proliferation, and division.

Examples of Response Elements

- **Second: Cyclic AMP Response Element (CRE)**

  The **cyclic AMP (cAMP) response element (CRE)** is activated when a ligand binds to a **G-protein-coupled receptor (GPCR)**, triggering the activation of **adenylyl cyclase**. This increases cAMP levels, which activate **protein kinase A (PKA)**. The catalytic subunit of PKA translocates to the nucleus, where it phosphorylates transcription factors bound to the CRE, driving the expression of genes involved in proliferation, survival, and differentiation. Additionally, PKA can phosphorylate cytosolic proteins involved in glucose metabolism.

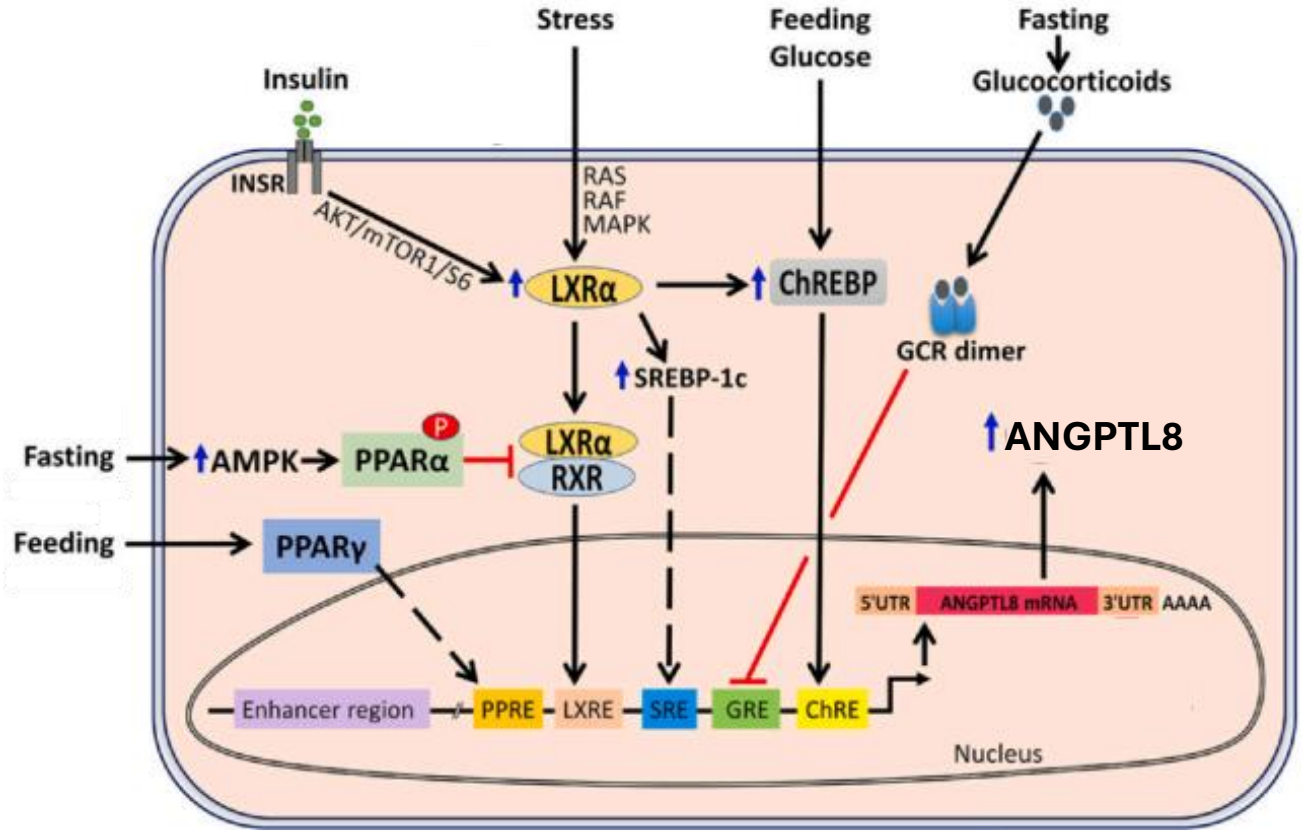
Examples of Response Elements

- **Third: Hormone Response Element (HRE)**

  The **hormone response element (HRE)** responds to steroid and thyroid hormones (e.g., androgen, estrogen, glucocorticoid, cortisol, thyroid hormones). When a hormone such as testosterone binds to its **androgen receptor (AR)** inside the cell, the receptor dimerizes and translocates to the nucleus. There, it binds to the androgen response element in the promoter regions of target genes, coordinating their expression to achieve a unified physiological outcome. For example, androgen regulates male reproductive functions and lipid and steroid metabolism, while estrogen influences the female reproductive system.

A single gene, multiple PPEs

- Angiopoietin-like protein 8 (ANGPTL8) is important for the metabolism of lipoproteins and triglycerides, and it is an inhibitor of lipoprotein lipase (LPL).
- Inducers of anabolism such as feeding, glucose, insulin stimulate transcription factors such as SREBP-1c and ChREBP.
- Inducers of catabolism such as fasting (and glucocorticoids) suppress the gene expression of ANGPTL8.
- *Note: not all PPEs have a positive effects.*



SREBP-1c: Steroid response element binding protein
ChREBP: Cholesterol response element binding protein

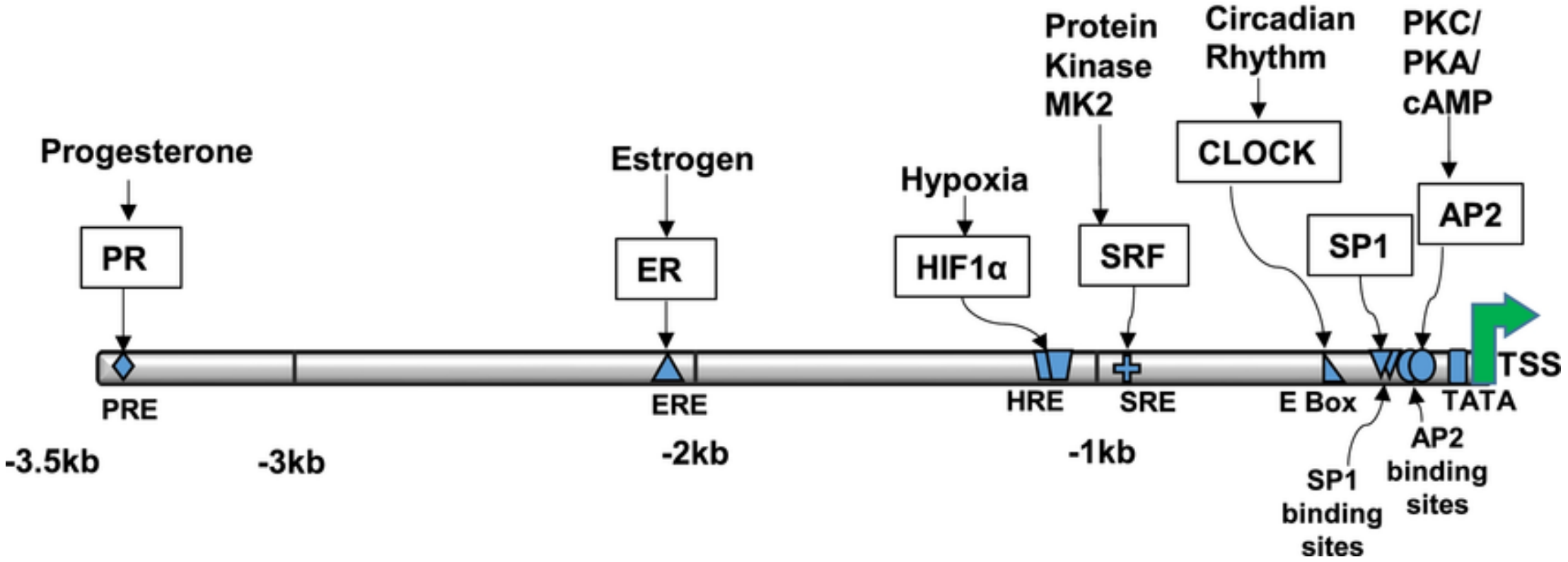
Single promoter, multiple PPEs

- A single gene can possess multiple **promoter-proximal elements (PPEs)** to allow for precise regulation. For example, the gene encoding **Angiopoietin-like protein 8 (ANGPTL8)**, which plays a crucial role in lipid metabolism by regulating lipoprotein lipase (LPL) activity and triglyceride levels, contains several PPEs.
- **Lipoprotein lipase (LPL)** hydrolyzes triglycerides in circulating lipoproteins, releasing free fatty acids that can be taken up by tissues. In adipose tissue, these fatty acids are re-esterified into triglycerides, promoting lipid storage and anabolic processes. ANGPTL8 inhibits LPL activity, reducing triglyceride breakdown and favoring lipid storage. Its expression is upregulated during conditions of high insulin levels—indicating sufficient glucose and lipid availability. Insulin activates transcription factors, such as **SREBP1c**, which bind to PPEs on the ANGPTL8 gene, promoting its transcription and enabling lipid synthesis.
- Conversely, during fasting, insulin levels drop, and the transcription factors activated by insulin are inhibited, suppressing ANGPTL8 transcription. Additionally, **glucocorticoids**, which are elevated during fasting, promote catabolic processes and may contribute to the suppression of ANGPTL8 expression. This coordinated regulation shifts the body away from lipid storage and toward lipid breakdown.
- Some PPEs may influence transcription positively by either initiating the process or enhancing its rate. Conversely, others may negatively regulate transcription by suppressing its initiation or reducing its activity.

Harmony

- The regulation of cellular processes, including their activation and inhibition, is influenced by the body's overall state and nutrient availability. These processes operate in a coordinated manner to maintain harmony within the system.
- How is this harmony achieved? When a ligand binds to its receptor, it activates multiple signaling pathways.
  - In the cytosol, rapid processes such as phosphorylation events occur. These can activate or inactivate specific enzymes or signaling molecules.
  - In the nucleus, slower processes involving transcription factors are initiated. These include gene transcription and mRNA translation, which require more time to exert their effects.
- This coordination ensures metabolic harmony. For example, it would be counterproductive to synthesize glycogen while simultaneously breaking down fatty acids. Instead, signaling pathways are regulated to favor one process over the other based on the cell's energy demands and metabolic state.

6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3 (PFKFB3)

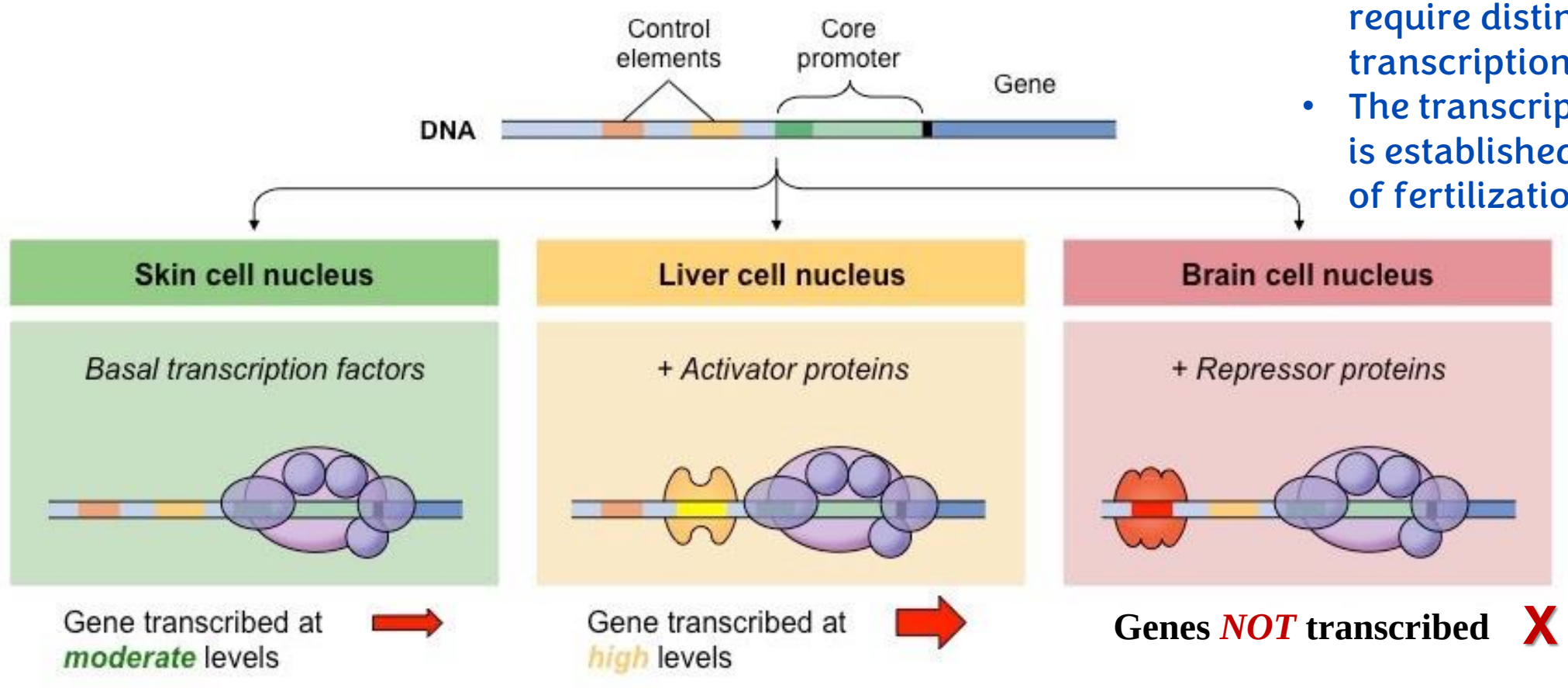


Another Example

- The enzyme responsible for synthesizing Fructose 2,6-bisphosphate, a key regulatory molecule in glycolysis and gluconeogenesis, is encoded by a gene with a promoter region containing multiple cis-regulatory elements. These include response elements sensitive to progesterone, estrogen, hypoxia, and circadian rhythm.
- For instance, during sleep, circadian rhythms regulate energy metabolism by modulating pathways that prepare the body for energy demands upon waking, such as glycogen synthesis and glucose homeostasis.
- These regulatory elements function collectively to fine-tune gene expression in response to environmental and physiological signals. While not all elements are active or necessary at a given time, their presence enhances transcriptional control and allows the body to efficiently adapt to changing conditions.

Tissue-specific transcription factors

- Tissue-specific transcription factors exist because different tissues require distinct types of transcription factors.
- The transcription program is established at the onset of fertilization.



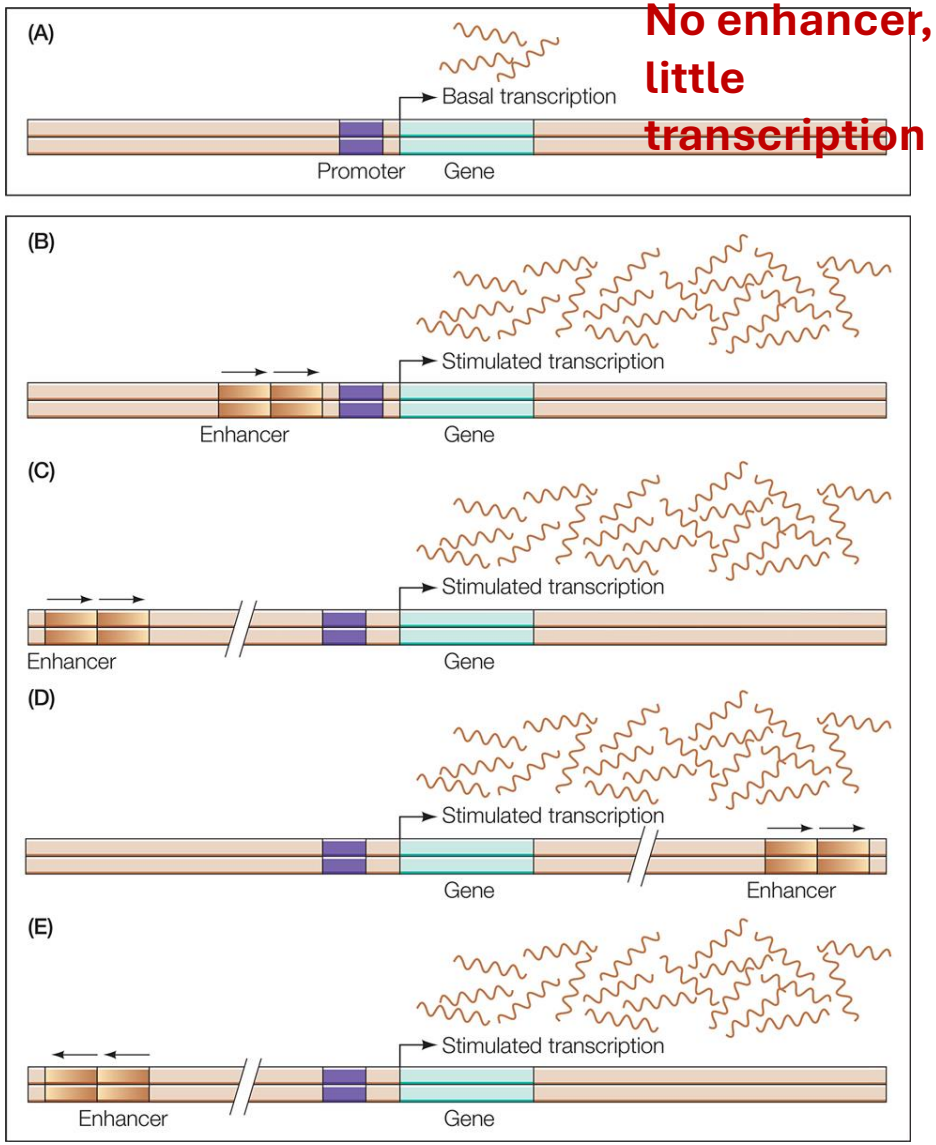
Differential expression of transcription factors (tissue-specific transcription factors) determines gene expression.

Tissue-specific Transcription Factors

- Tissue-specific transcription factors regulate transcription of genes in certain tissues.
- For example: Insulin affects specific cells by inducing transcription factors, but not all cells possess the necessary transcription factors for insulin-mediated regulation. This explains why insulin activity is tissue-specific. For instance, insulin is produced in pancreatic beta cells but not in neurons, as the transcription factors required for insulin gene regulation are present in beta cells but absent in neurons.
- Consider these examples of transcription regulation:
    1. **Skin cells**: Only basal transcription factors are present, leading to minimal transcription and low levels of protein synthesis. Enzymes like those involved in insulin signaling are not needed in high amounts here.
    2. **Liver cells**: Both basal transcription factors and activators are present, resulting in higher transcription levels. This supports liver functions like glucose metabolism.
    3. **Brain cells**: Inhibitory transcription factors suppress certain genes entirely, such as those involved in insulin production, ensuring these proteins are not produced where unnecessary.
- In summary, transcription is tissue-specific, allowing proteins and receptors to be synthesized only where needed. Cells adjust protein production based on their metabolic needs, such as responding to high sugar levels by increasing transcription of glucose-metabolism genes.

Enhancers

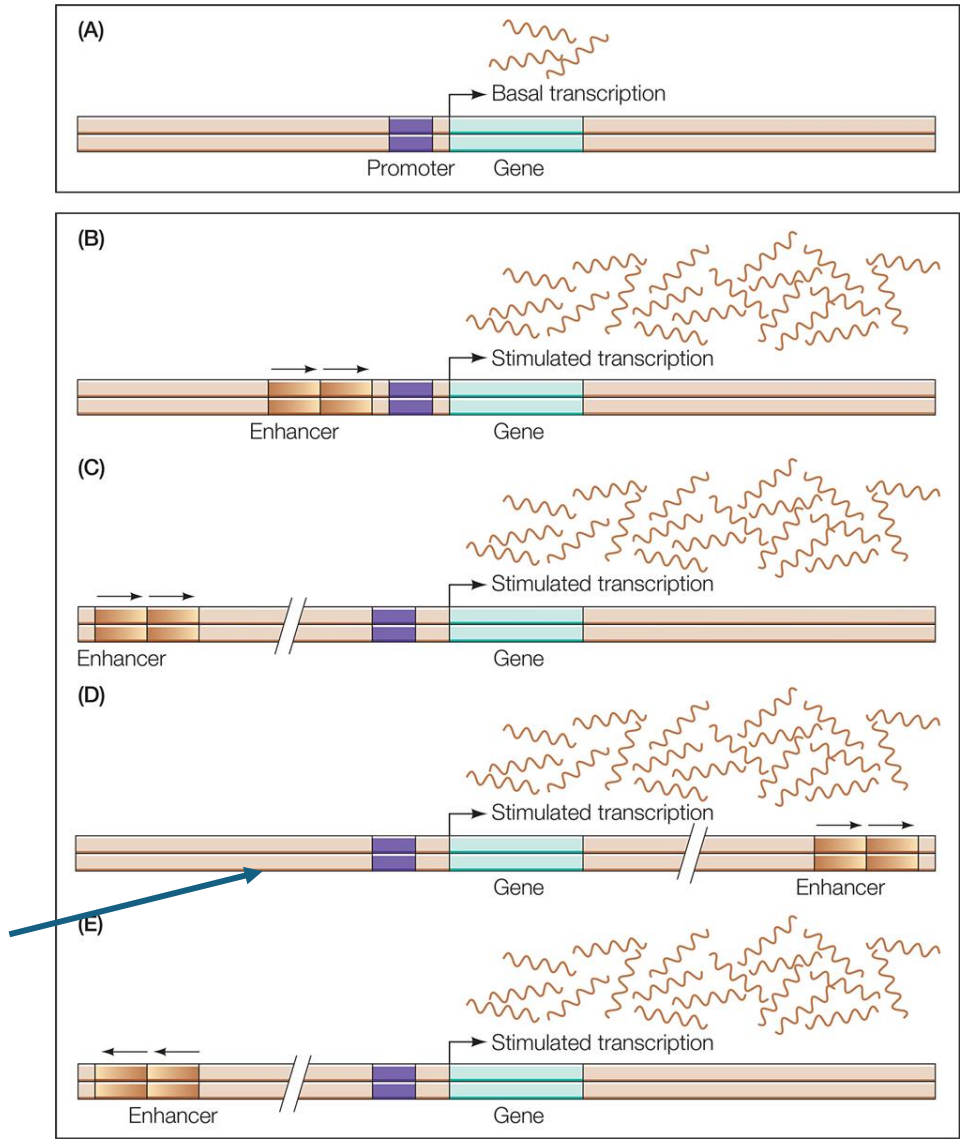
- Many genes are regulated by regulatory sequences called enhancers, which are binding sites for specialized, gene-specific, cell-specific, regulatory transcription factors that regulate RNA polymerase II.
- They can regulate transcription regardless of orientation or location due to DNA looping.
- There are 500,000 to over 1 million enhancers in the human genome, accounting for 10% or more of total genomic DNA.



- Enhancers play a crucial role in gene expression. Without enhancers, genes are expressed only at basal levels, resulting in minimal mRNA and protein production. In contrast, the presence of enhancers significantly increases gene expression.
- **Differences Between Enhancers and Promoter Proximal Elements:**

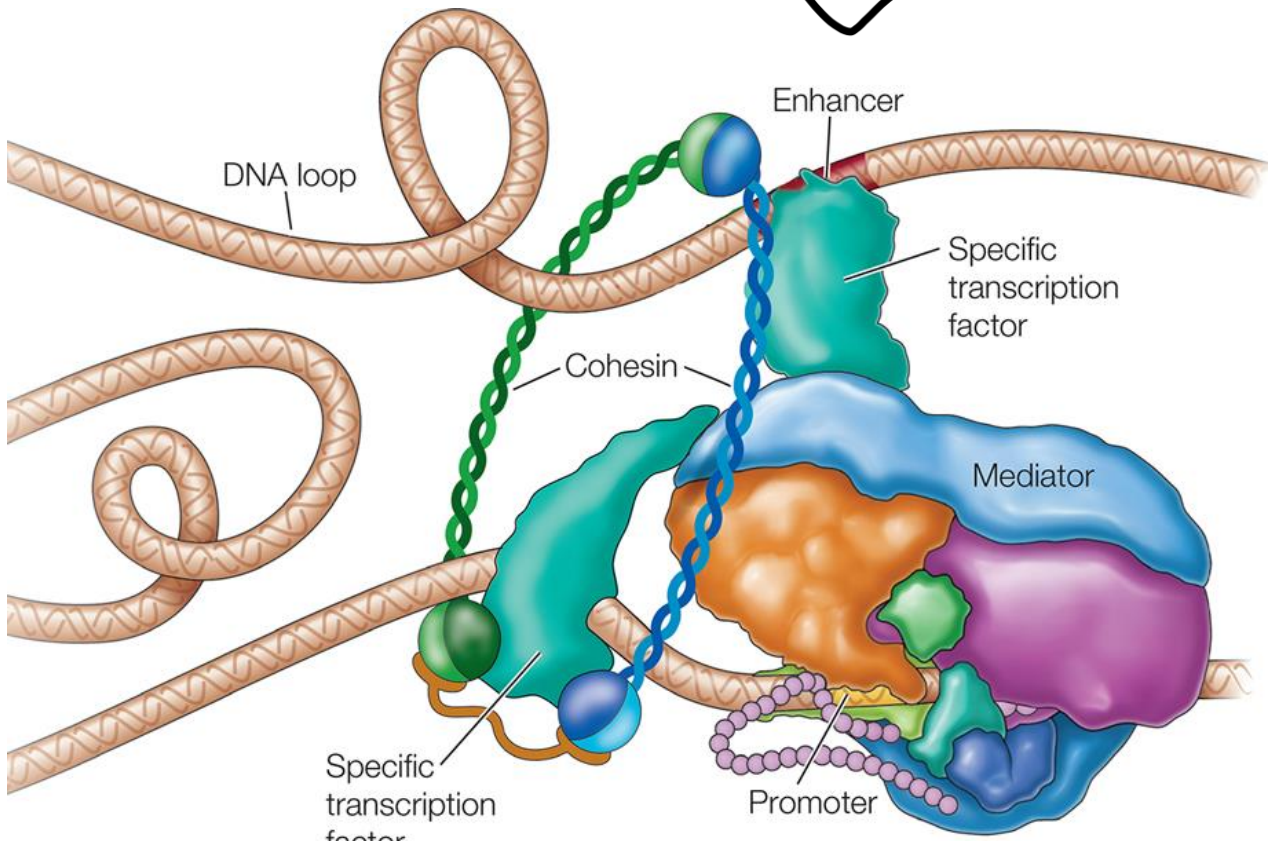
    1. **Location**: Promoter proximal elements must be located near the promoter region, while enhancers can function at variable distances. Enhancers can be **close** to the promoter, **far** away (even kilobases apart), **within** the gene itself, or **downstream** at the gene's end.
    2. **Orientation Independence**: Enhancers remain functional even if their sequence is flipped or inverted; because of DNA looping.
    3. **Specificity**: Promoter proximal elements regulate specific genes, whereas enhancers can regulate multiple genes.

- Enhancers provide versatility and flexibility in gene regulation, enabling the precise control of gene expression necessary for proper cellular function and development.



DNA looping and Cohesin

- Transcription factors bound at an enhancer can interact with a protein called (Mediator) or general transcription factors at the promoter.
- This is due to the ability of DNA to loop.
- DNA looping is stabilized by a protein called <u>cohesin</u>.

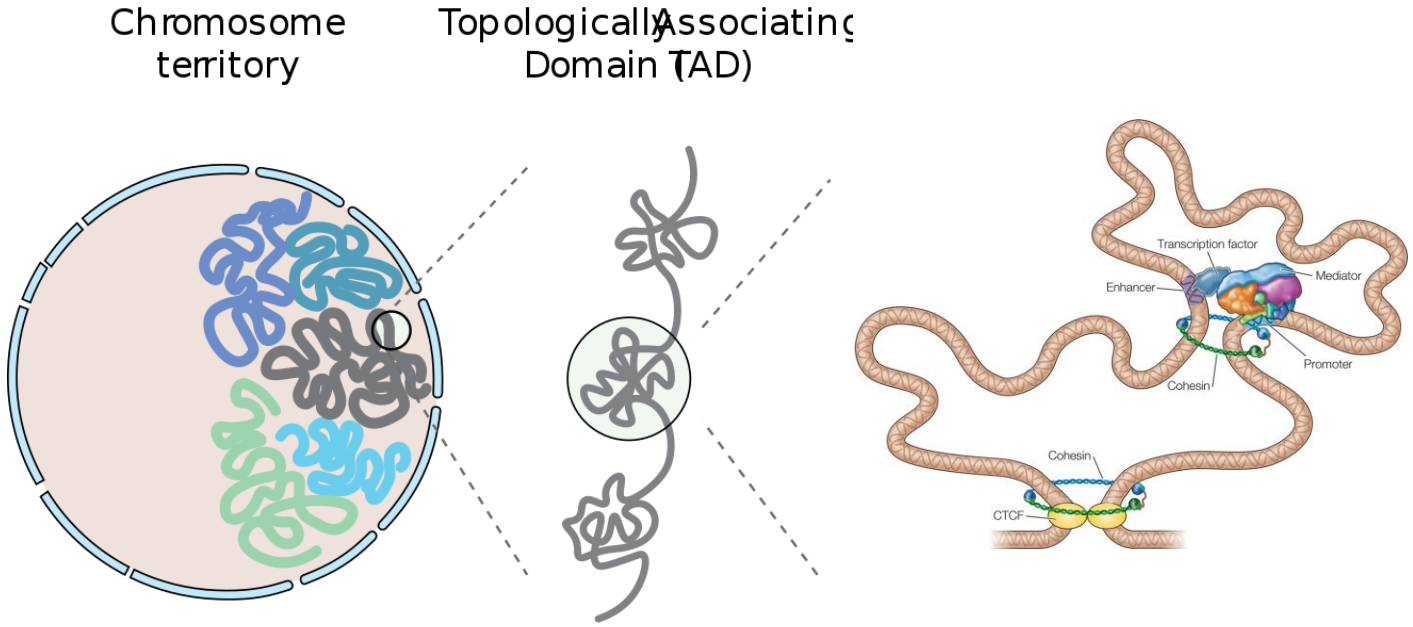


DNA Looping and Cohesin

- Enhancers are **dynamic**, meaning they can function even if their position within the genome shifts. This flexibility is due to their ability to form DNA loops, enabling interaction with the promoter region despite changes in spatial positioning.
- At the pre-initiation complex, a **mediator protein complex** facilitates the interaction between proteins bound to enhancers and those bound to the promoter region of the gene. This interaction plays a crucial role in stimulating gene expression.
- The interaction between enhancers and the promoter region is further stabilized by **cohesin**, a protein that helps maintain the DNA loop structure.
- Both DNA looping and mediation rely on **protein-protein** and **protein-DNA interactions**, which are non-covalent in nature. These interactions ensure precise communication between regulatory elements and the transcription machinery.

Enhancers, insulators, TADs, and CTCF

- DNA sequences or elements known as insulators divide the genome into topologically associating domains (TADs) forming loops. The chromatic arrangement in the nucleus is not random.
- The boundaries of the loops are stabilized and defined by cohesin and CTCF proteins.
- Enhancers are restricted to interacting with promoters in the same domain; therefore, they cannot be located extremely far away.



For any feedback, scan the code or click on it.

Corrections from previous versions:

Versions	Slide # and Place of Error	Before Correction	After Correction
V0 → V1	1. #20 2. #21 3. #27	1. These genes share the same PPEs and core promoter sequences, allowing them to be activates simultaneously. 2. cAMP Response Elements (SRE) 3. These include gene transcription, ..., protein synthesis, and protein processing, ..., effects.	1. These genes share similar **PPEs** only, which facilitate their coordinated activation. 2. cAMP Response Elements (CRE) 3. These include gene transcription, ... effects.