NUCLEIC ACIDS – BIOCHEMISTRY MODULE 2

- **Structure and Function of Nucleic Acids**
- **Genes and Chromosomes (brief summary)**
- **DNA Replication**
- **DNA Transcription**
- **Protein Synthesis**

MAIN TOPICS

DNA Replication, and Repair: process by which DNA is copied with high fidelity.

RNA Synthesis and Processing: the process by which the DNA genetic code is ‘read’ and transcribed, forming mRNA (messenger RNA). This is the intermediate step in protein expression. Only what is required is transcribed.

Protein Synthesis: process by which the genetic code is translated into a protein sequence, the end product of gene expression.

TRANSCRIPTION

genes

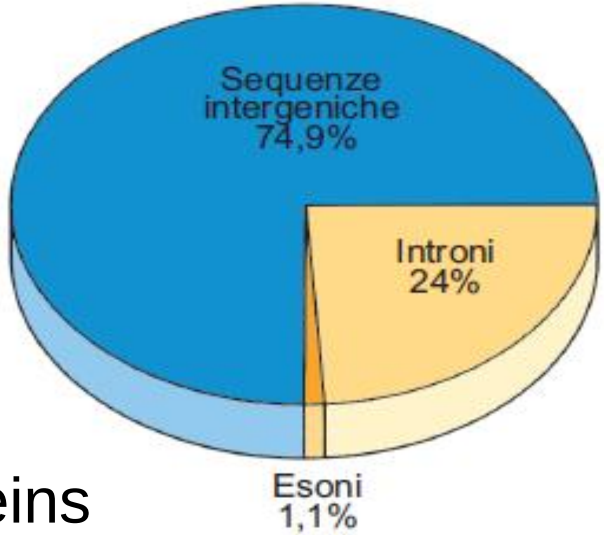
Definitions

1. **Classic definition:** Portion of DNA that determines a single phenotype
2. **A gene – a protein** (Beadle & Tatum 1940): “Each gene encodes information for a protein”
3. **Current definition:** fragment of DNA (or sometimes RNA) that contains the primary sequence to produce a biologically functional gene product (RNA, protein).

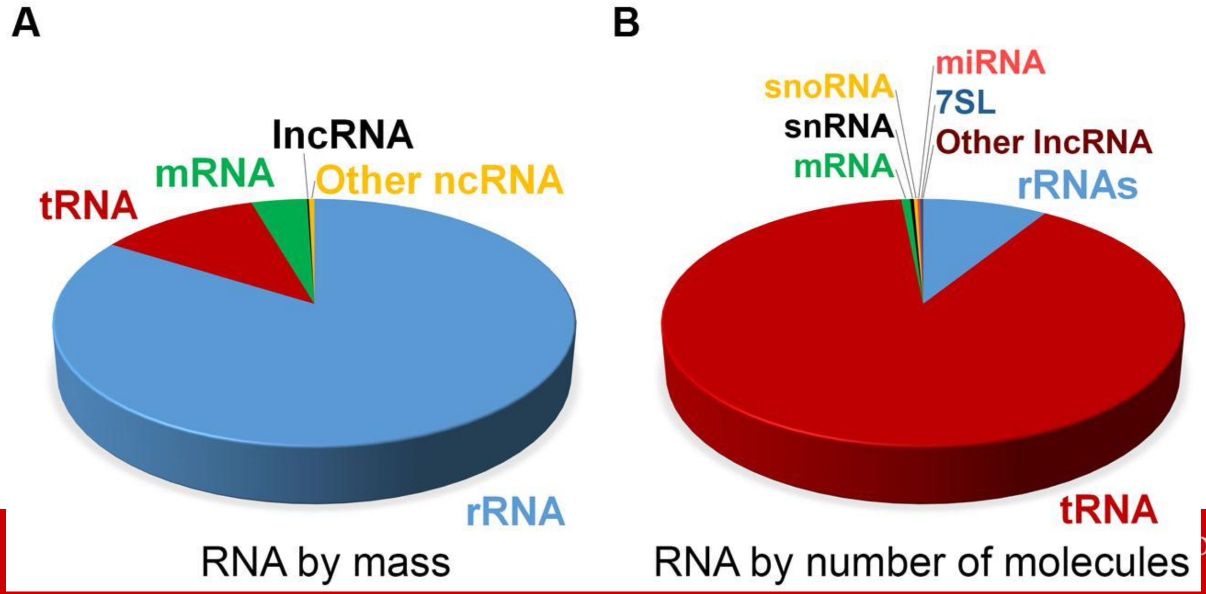
TRANSCRIPTION

Human genome

- 23 pairs of chromosomes
- 3.100.000.000 base pairs
- 20.000-30.000 genes, which correspond to mRNA that translates into proteins
- The function of about half of the discovered genes is unknown



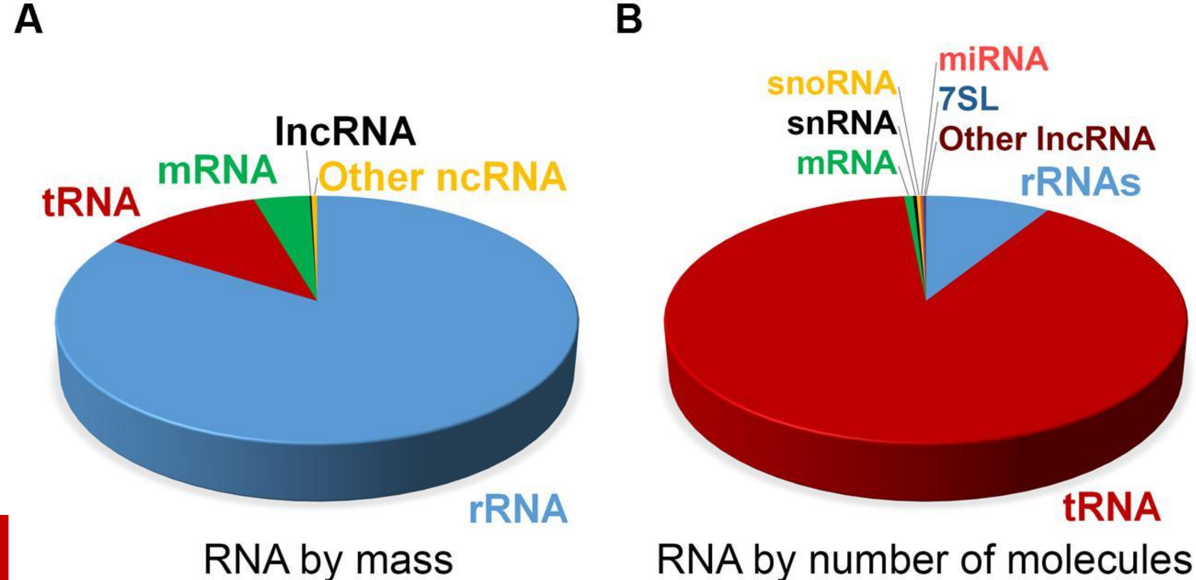
Less than 2% of the genome encodes proteins



TRANSCRIPTION

Human genome

- **Non-Coding RNA:** 80,000 non-redundant non-coding RNA genes (lncRNAs, snRNAs, and microRNAs). Non-coding RNAs are involved in regulation and other cellular processes.
- **Ribosomal RNA (rRNA):** one of the most abundant types of RNA in human cells, constituting about 80-90% of total cellular RNA. They are essential for protein synthesis.
- **Transfer RNA (tRNA):** carries amino acids in an activated form to the ribosome for peptide-bond formation, in a sequence dictated by the mRNA template. There is at least one kind of tRNA for each of the 20 amino acids.



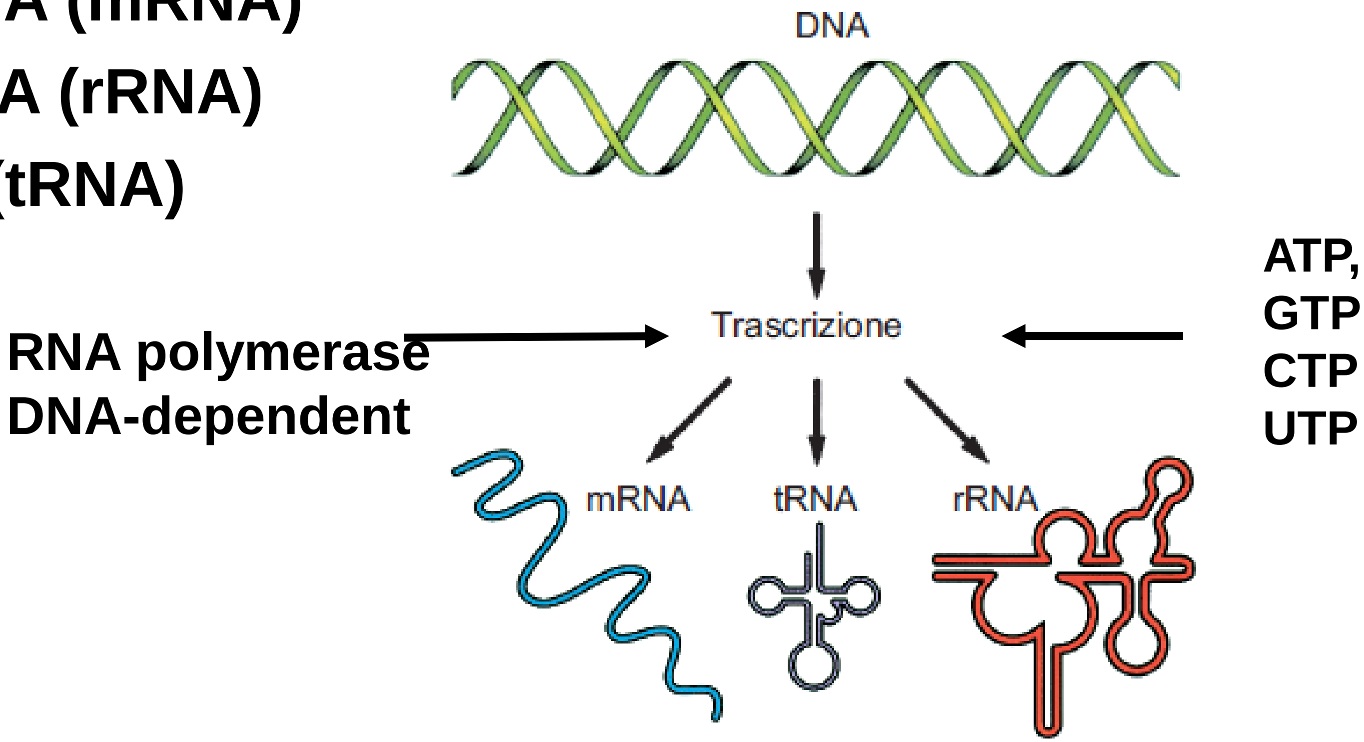
TRANSCRIPTION

RNA

RNA contains the sugar ribose, and usually uracil instead of thymine. It normally exists as a single strand.

3 main types of RNA:

- **Messenger RNA (mRNA)**
- **Ribosomal RNA (rRNA)**
- **Transfer RNA (tRNA)**



TRANSCRIPTION

RNA polimerase

✓In the cytosol for prokaryotes

✓In the nucleus for eukaryotes

in eukaryotes:

RNA polymerase I pre-rRNA (18s, 5.8s, 28s)

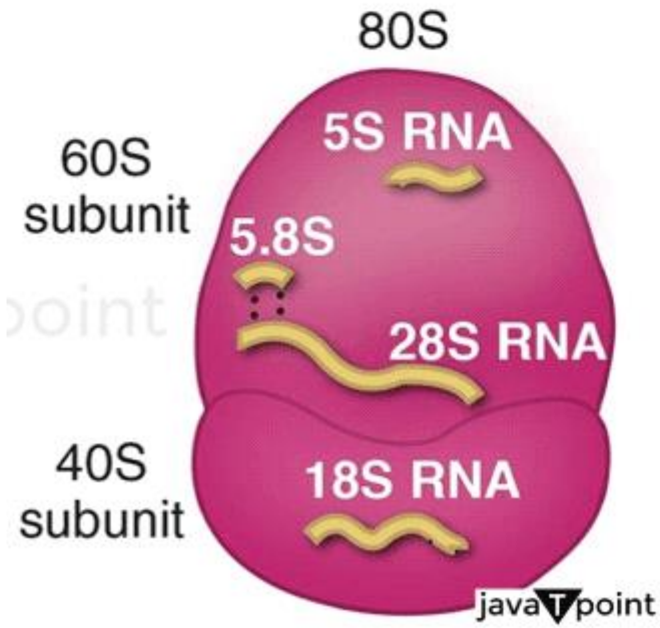
RNA polymerase II pre mRNA, mRNA and several types of non-coding RNAs.

RNA polymerase III tRNA, rRNA 5s, ..

RIBOSOMAL RNA

The main types of rRNA present in eukaryotes include:

- 18S rRNA: Part of the small subunit (40S) of the ribosome, crucial for mRNA binding and translation initiation.
- 5.8S rRNA: Found in both the small and large ribosomal subunits; it plays a role in the structure and stability of the ribosome.
- 28S rRNA: A component of the large subunit (60S) of the ribosome, involved in peptide bond formation during protein synthesis.
- 5S rRNA: Also part of the large subunit, contributing to the overall structure and function of the ribosome.

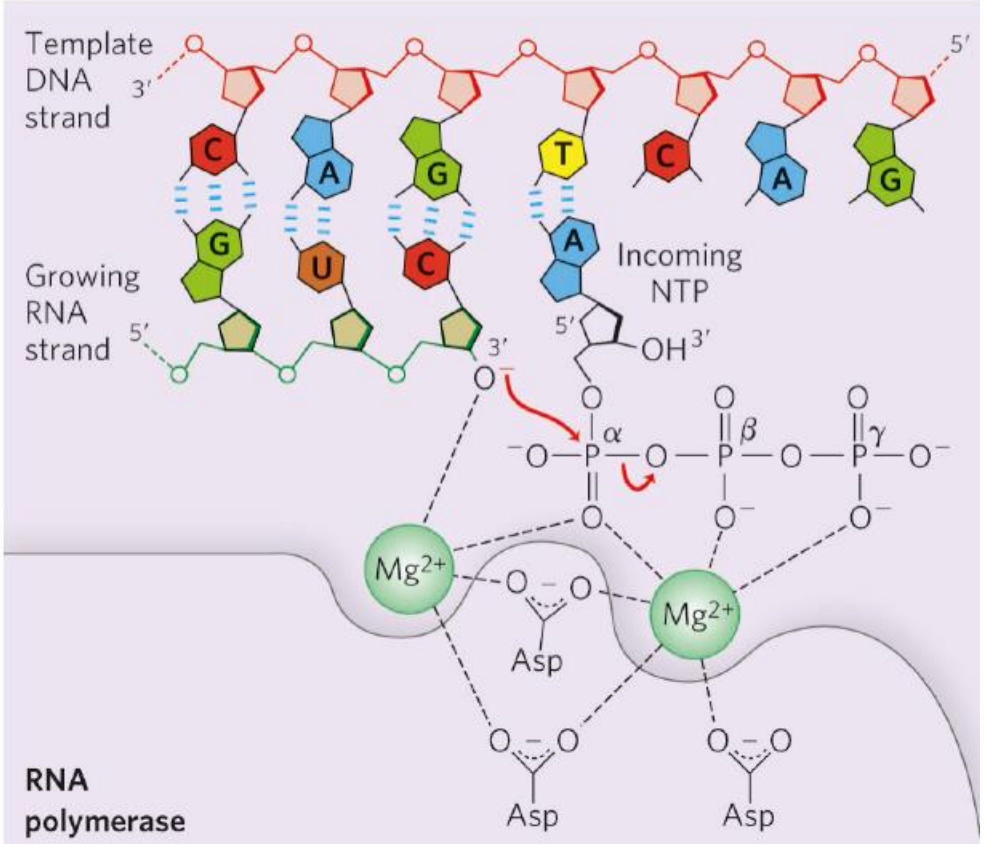


TRANSCRIPTION

Main features

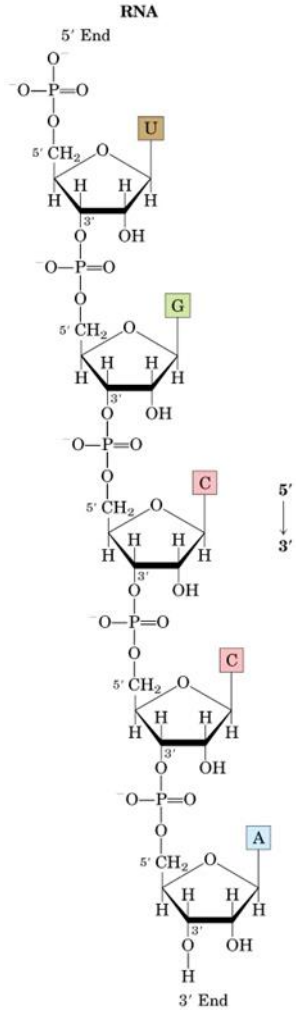
$$(NMP)_n + NTP \rightarrow (NMP)_{n+1} + PP_i$$

RNA polymerase-DNA dependent



-OH in 3' binds to phosphate α:

- Mg^{2+} ions promote the deprotonation of –OH, allowing to be more nucleophile.
- The other Mg^{2+} ion binds to the NTP and promotes the exit of the PPi group.



TRANSCRIPTION

Main features

Template	Single strand DNA
Substrate	NTP
Primer	no
Enzyme	RNA polymerase
Product	ssRNA
Base pairing	A-U, U-A, G-C, C-G

TRANSCRIPTION

Main features

The RNA molecule will be complementary to the DNA template strand and identical (except for the U replacing the T and the sugar) to the non-template DNA coding sequence.

DNA

5'------ G C A G T A C A T G T C------ 3' **coding strand**

3'------ C G T C A T G T A C A G------ 5' **template strand**

transcription

5'------ G C A G U A C A U G U C------ 3' **RNA**

RNA synthesis proceeds in the direction 5' → 3'

TRANSCRIPTION

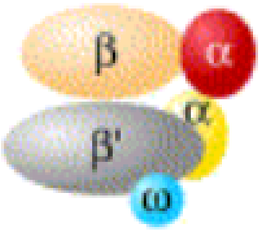
RNA polymerase

Viral RNA polymerase (1 subunit)

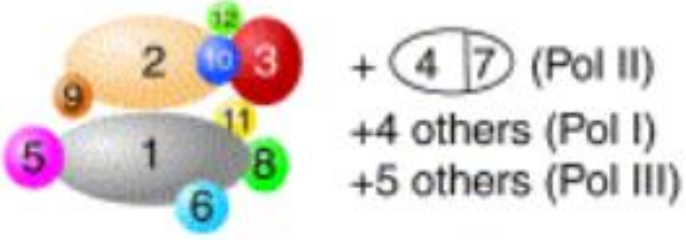
Prokaryotic RNA polymerase (5 different subunits)

Eukaryotic RNA polymerase (at least 12 subunits)

Virus



Bacteria



Eukaryotes

TRANSCRIPTION

RNA polymerase in bacteria

It has five subunits with a total molecular weight of approximately 450 kDa:

Subunit	Composition	Role
α **Subunits**	2 copies (α2)	- **Assembly**: The α subunits are crucial for the assembly of the RNAP complex. They serve as a scaffold for the binding of β and β' subunits. - **Transcription Regulation**: The C-terminal domain (αCTD) interacts with various transcription factors and upstream promoter DNA, influencing transcription regulation [1] [3].
β **Subunit**	1 copy (β)	- **Catalytic Activity**: The β subunit is involved in forming the active site for RNA synthesis. It contributes to the binding of the DNA template and ribonucleoside triphosphates (rNTPs). - **Structural Support**: It forms part of the "claw" structure of RNAP, facilitating the entry of DNA into the enzyme [1] [2] [3].
β' **Subunit**	1 copy (β')	- **Catalytic Activity**: Similar to β, the β' subunit also plays a critical role in catalyzing RNA synthesis and forms part of the active site. It is involved in binding to the DNA template and stabilizing RNA-DNA hybrids during transcription [1] [2] [3].
ω **Subunit**	1 copy (ω)	- **Stability and Assembly**: The ω subunit is thought to assist in the proper folding and stability of RNAP. Although it is not essential for transcription, it helps maintain the structural integrity of the enzyme during RNA synthesis [1] [4].

TRANSCRIPTION

RNA polymerase

RNA synthesis is catalysed by the enzyme DNA-DEPENDENT RNA POLYMERASE, which functions similarly to DNA polymerase:

1. IT ATTACKS RIBONUCLEOTIDE TRIPHOSPHATES REALISING Ppi.

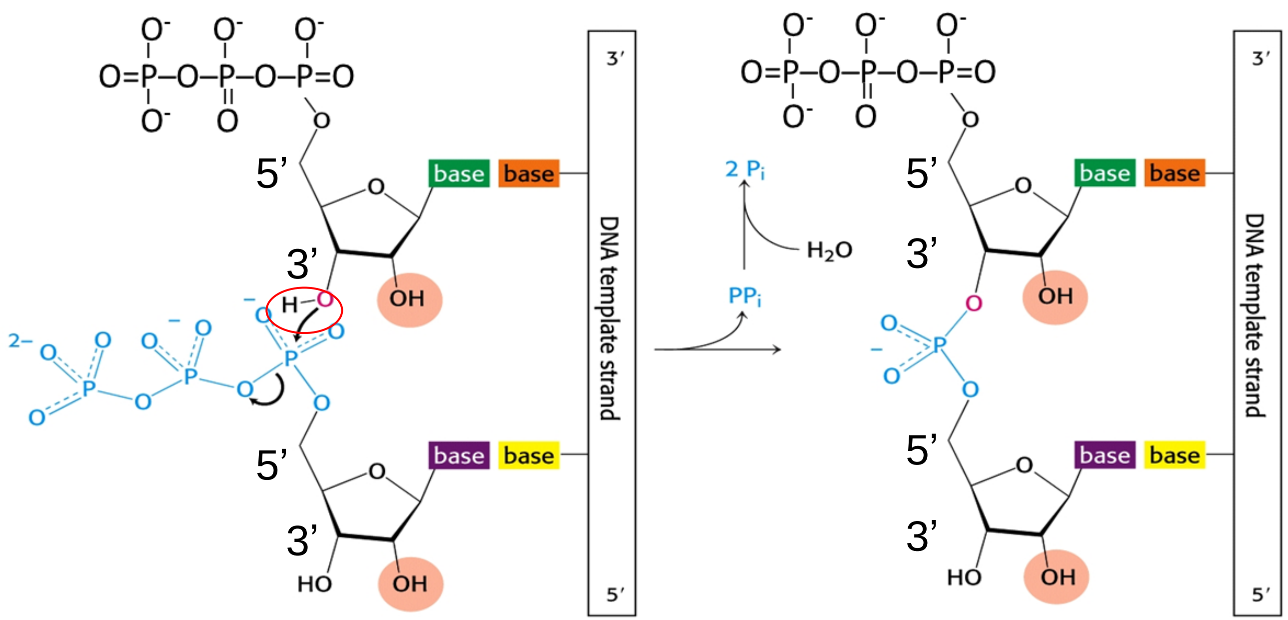
2. THE REACTION PRODUCT IS A COMPLEMENTARY COPY OF THE TEMPLATE DNA.

In prokaryotes, all types of RNA are synthesised by a single RNA polymerase.

In eukaryotes, the synthesis of rRNA, mRNA and small tRNA is catalysed by 3 different polymerases: polymerases I, II and III

TRANSCRIPTION

RNA polymerase reaction mechanism



1. RNA polymerase synthesizes in the 5'-3' direction (NTP added at the 3' end only)
2. -OH 3' of the chain reacts with the α-phosphate of the incoming NTP, releasing Ppi
3. The added ribonucleotide follows the Watson-Crick pairing rules
4. RNA polymerase does not require RNA primers, but takes instructions from DNA templates
5. RNA polymerase lacks intrinsic exonuclease activity (no proofreading activity), which means it cannot remove nucleotides from the RNA strand after they have been incorporated.

TRANSCRIPTION

stages

1. INITIATION: interaction with PROMOTERS

2. ELONGATION

3. TERMINATION

INITIATION:

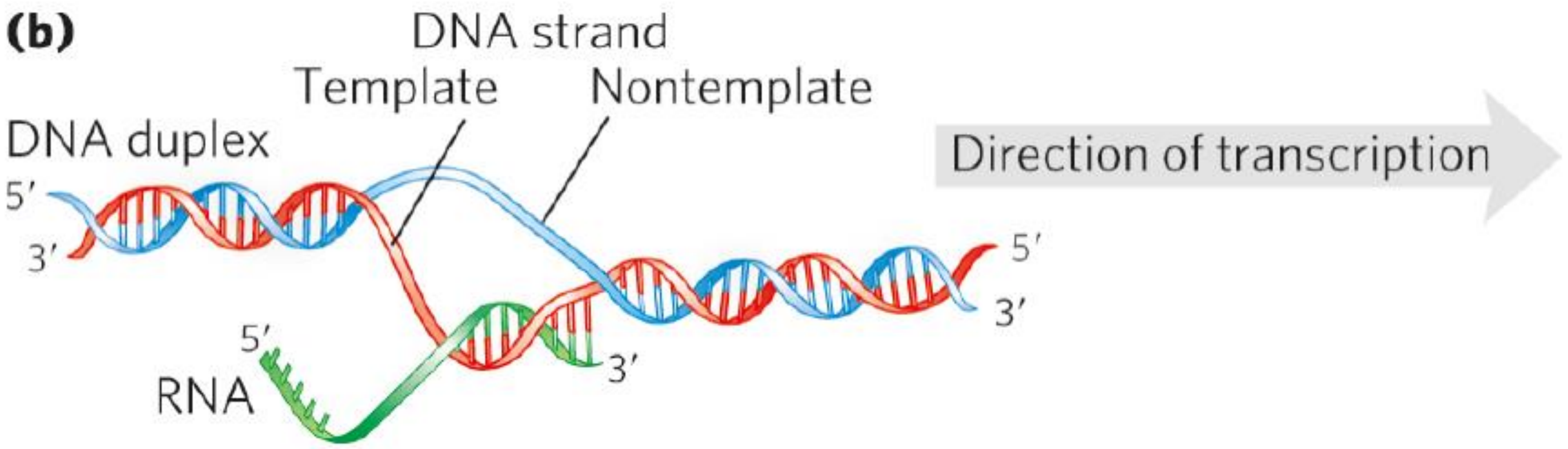
1. RNA-polymerase binds to DNA and migrates to a PROMOTER site.
2. It binds to DNA with high affinity close to the promoter.
3. The binding with the promoter induces a CLOSED COMPLEX, in which DNA is still a double strand helix.
4. RNA-polymerase decouples about 12 bp with the formation of a very stable complex: OPEN COMPLEX.

TRANSCRIPTION

stages

ELONGATION:

1. During the elongation, the RNA strand pairs with the DNA template strand to form an RNA-DNA hybrid double helix.

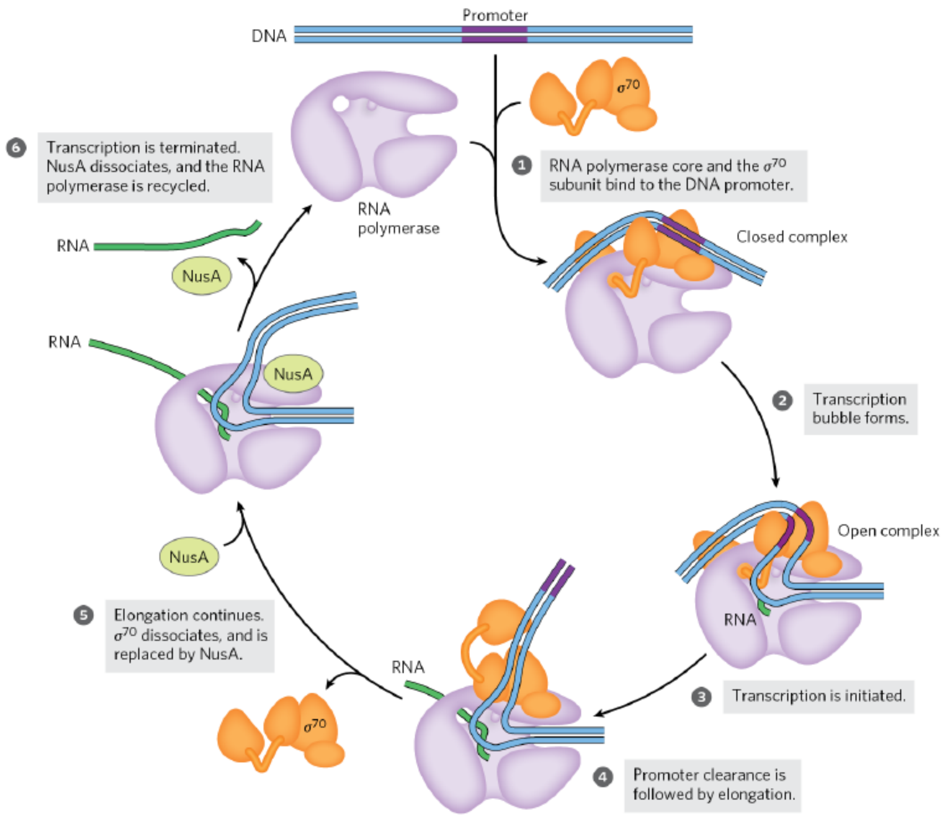


TRANSCRIPTION

stages

ELONGATION:

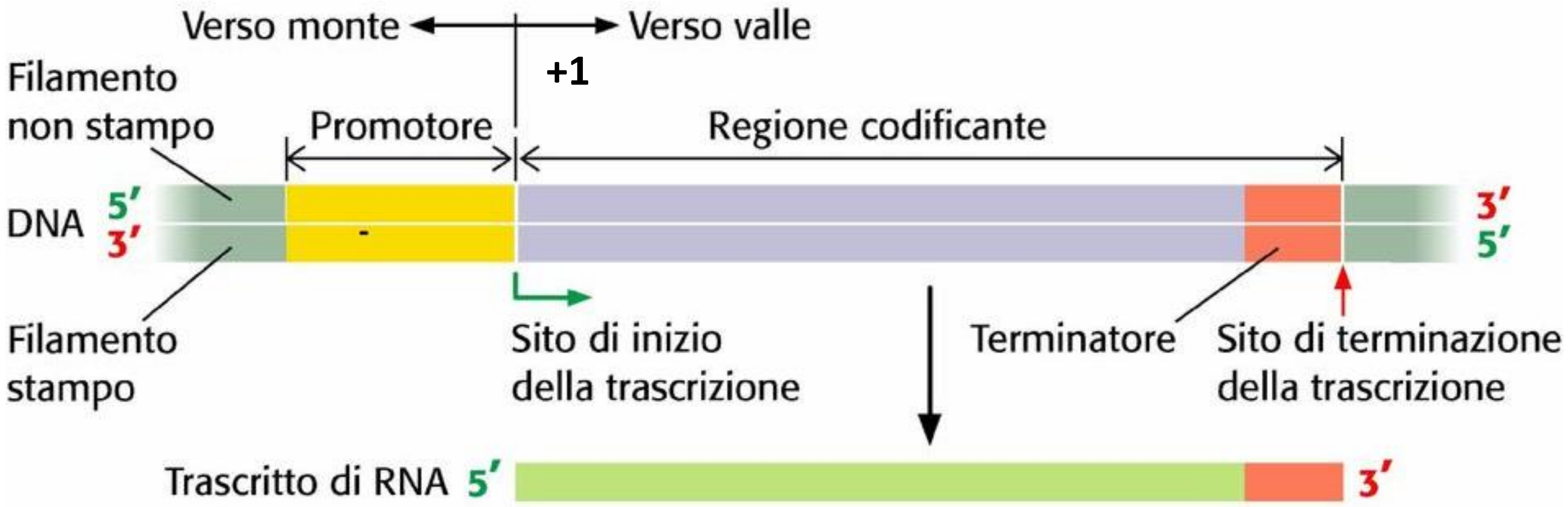
2. σ subunit dissociates to allow RNA polymerase to proceed with elongation and is replaced by the NusA protein



Sigma factors bind to the core RNA polymerase enzyme to form the holoenzyme (α2ββ′ωσ). This complex is capable of recognising and binding to specific promoter sequences, such as the -10 and -35 regions, which are crucial for initiating transcription. Once bound to the promoter, the sigma factor helps unwind the DNA at the transcription start site, allowing RNA polymerase to access the template strand.

TRANSCRIPTION

INITIATION: recognition of promoters



Promoters are specific DNA sequences that indicate the initiation of RNA synthesis to RNA polymerase. The base of the coding strand that represents the starting point of transcription is +1. This nucleotide corresponds to the first nucleotide incorporated into the RNA. Subsequent nucleotides within the transcribed region are numbered with +; non-transcribed sequences located to the left (upstream) of the start point are numbered with a negative sign -.

TRANSCRIPTION INITIATION

RECOGNITION OF THE PROMOTER: this represents the crucial and limiting step in the transcription process.

- Recognition of the promoter site of all transcribed genes involves a short common **AT consensus sequence** at -10 nt on the 5' side of the start site of transcription (+1).

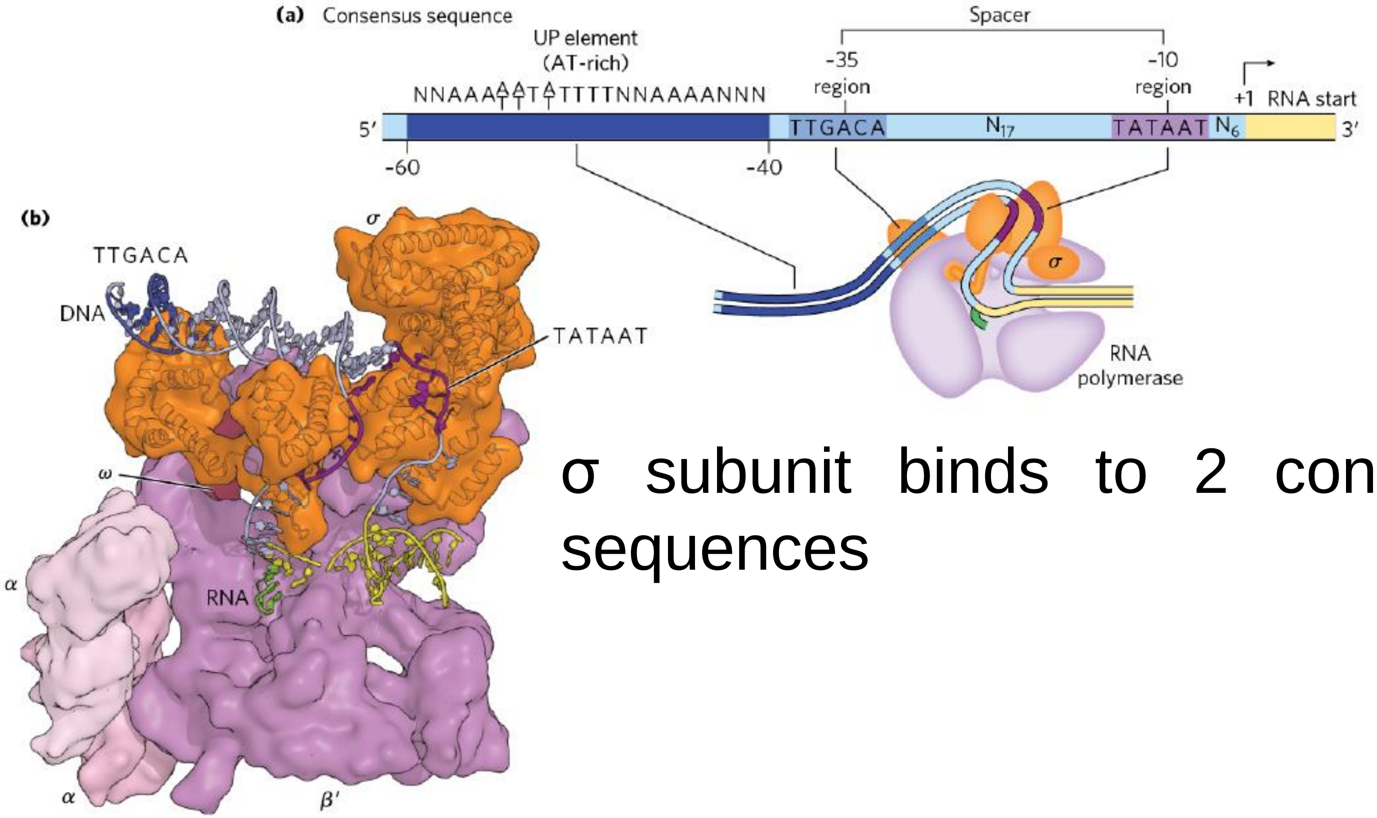
This region called Pribnow Box, is crucial for the initial binding of RNA polymerase and is recognized by the **sigma factor** of the enzyme (for E. coli = TATAAT on the non-transcribed strand).

- -35 Sequence (35 base pairs upstream of the transcription start site), the consensus sequence here is typically TTGACA. This sequence also plays a significant role in RNA polymerase binding and promoter recognition.

TRANSCRIPTION

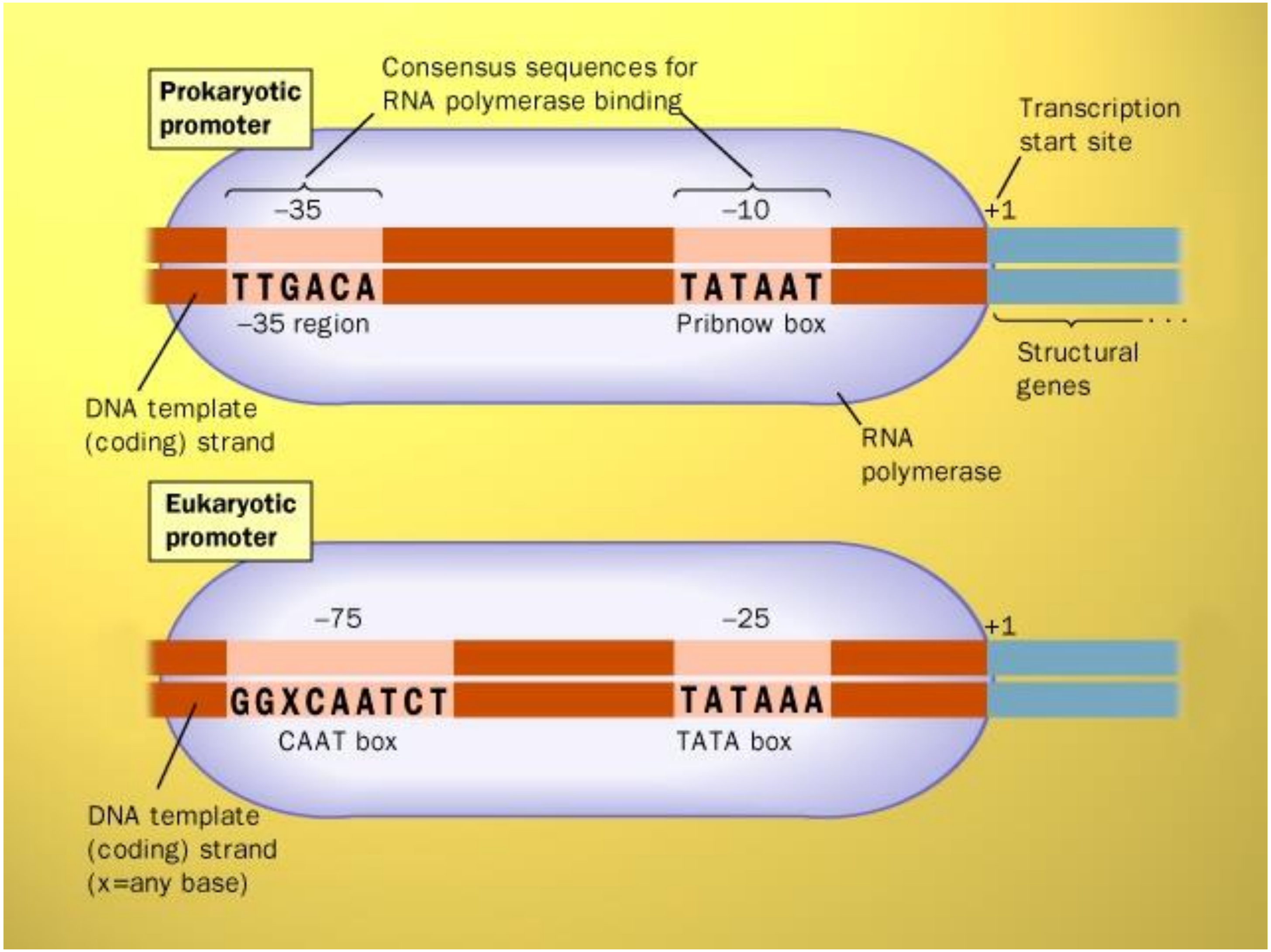
INITIATION

There is also an upstream promoter element (UP), a third AT-rich region between positions -60 and -40 to which the α subunit of RNA polymerase binds.



σ subunit binds to 2 consensus sequences

TRANSCRIPTION INITIATION



TRANSCRIPTION

INITIATION AND ELONGATION

Once made the OPEN COMPLEX, the enzyme is ready for the synthesis.

RNA-polymerase contains **two nucleotide-binding sites**:

- One is used for initiation and preferentially binds an ATP or GTP with a Km of about 100 μM
- The other is used for elongation and can bind any of the 4 ribonucleotides with a Km of about 10 μM

This results in the RNA transcription starting with a purine triphosphate base at the 5' end. After transcription of the first 10 nt, the σ subunit dissociates, making the elongation complex more stable.

Sliding of the Enzyme Core occurs in 2 patterns:

✓ Continuous sliding (1 nt at a time)

✓ Discontinuous sliding (or caterpillar movement)

TRANSCRIPTION

INITIATION AND ELONGATION

The -35 and -10 regions (consensus sequences) are highly conserved and the presence of mutations at this level influences transcription efficiency in vivo:

1. **UP-PROMOTER mutations**: make these regions more similar to consensus sequences and increase transcription rate
2. **DOWN-PROMOTER mutations**: decrease the strength of the promoter by making it less similar to the consensus sequence.

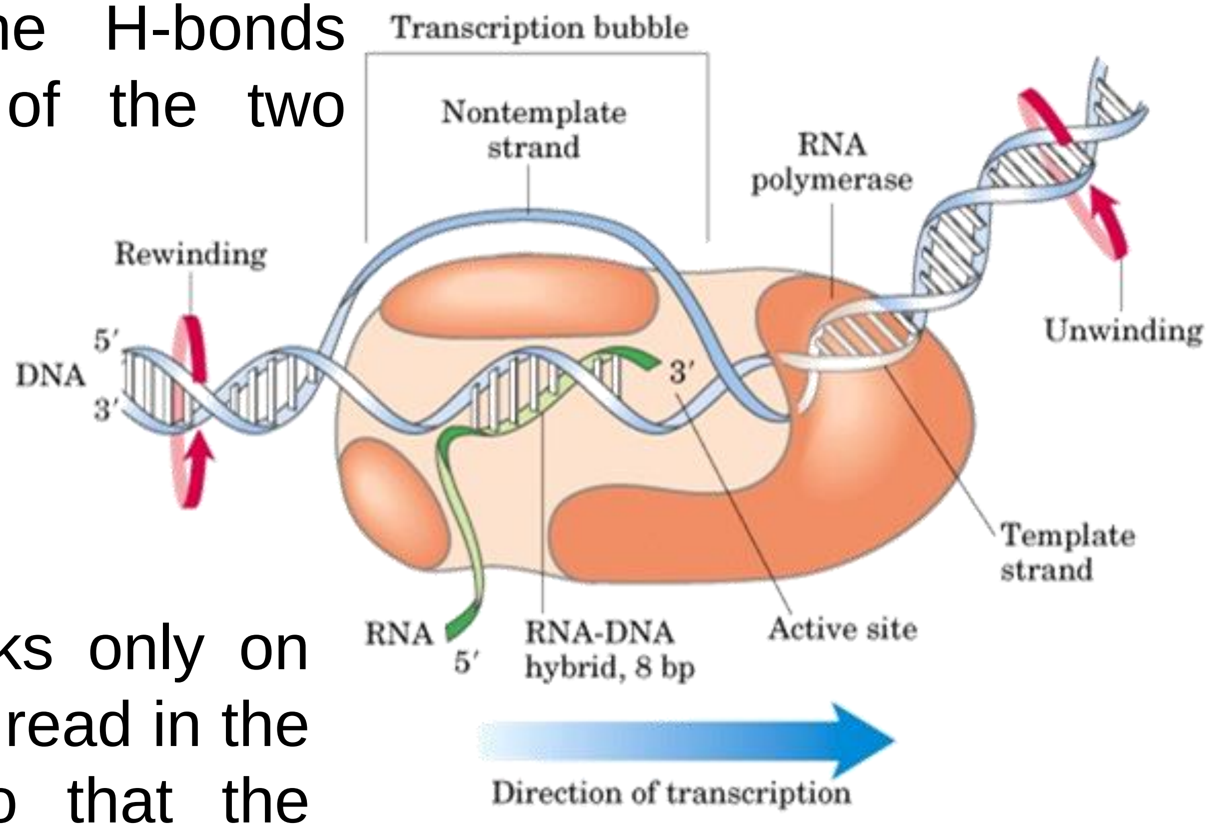
There are spacers between the -35 and -10 regions, most of which have 17 bp; spacers with 16 or 18 bp have also been found, but are less efficient.

TRANSCRIPTION

RNA polymerase action

RNA POLYMERASE unwinds the DNA and breaks the H-bonds between the bases of the two strands.

Base pairing occurs between the new-born RNA and the nt of the template DNA gene.

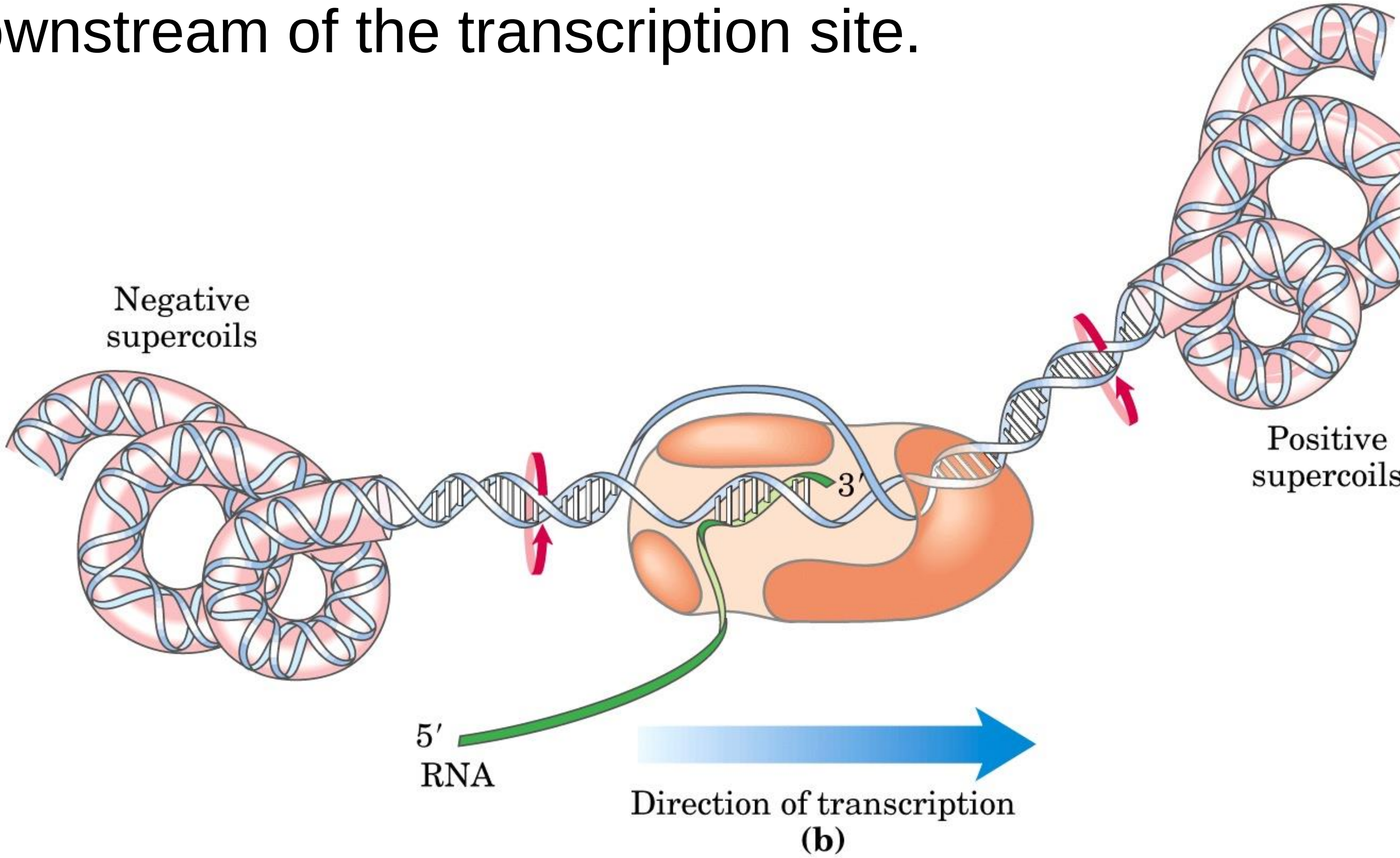


RNA polymerase works only on the DNA strand that is read in the direction 3’ → 5’ so that the transcript is synthesized in the direction 5’ → 3’

TRANSCRIPTION

RNA polymerase action

The unwinding of the DNA involves the formation of supercoils that require the intervention of topoisomerases located upstream and downstream of the transcription site.



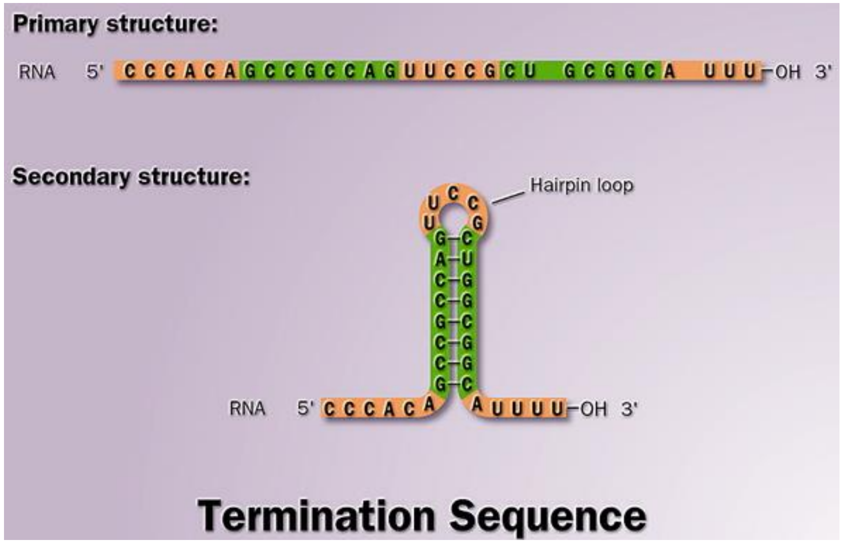
TRANSCRIPTION

termination

1. TERMINATION DEPENDS ON A PROTEIC FACTOR ρ (rho)

2. TERMINATION DOESN'T DEPEND ON ρ FACTOR

- genes with GC-rich symmetrical segments that form a hairpin structure in the transcript followed by a poly-A tail , 4-8A, transcribed as U.



EVENTS IN RHO INDEPENDENT TERMINATION:

1. RNA polymerase near the GC-rich site slows down (it is more difficult to unwind this stretch of DNA)

2. This slowdown allows the transcribed GC residues to pair up (formation of the hairpin structure), destabilizing the association of mRNA with the DNA.

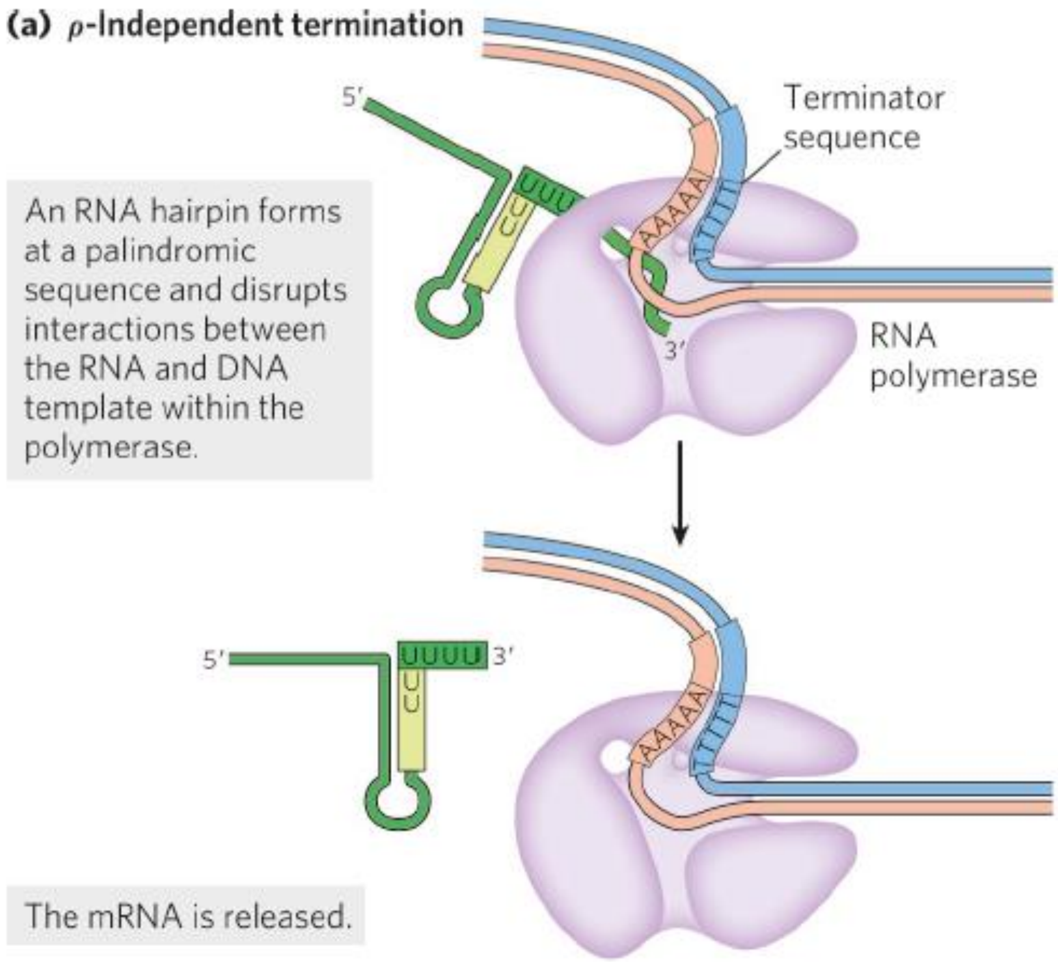
3. The formation of this structure moves the transcript away from the template (the RNA-DNA complex is less stable)

4. Further weakening of the RNA-DNA bond when the A-rich sequence is copied (the A-U pair is not very stable)

TRANSCRIPTION

termination

RHO INDEPENDENT TERMINATION:



TRANSCRIPTION

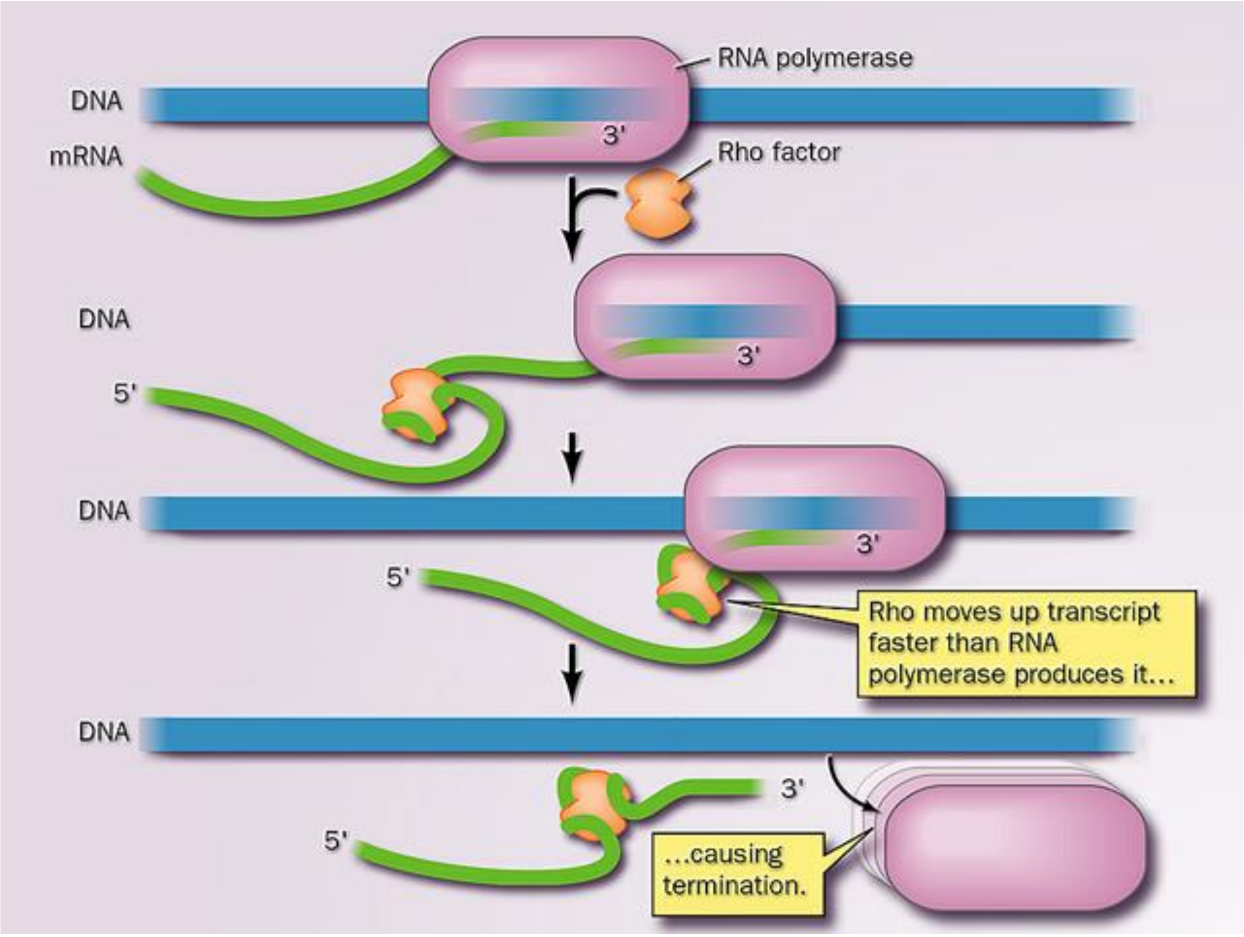
termination

RHO-DEPENDENT TERMINATION

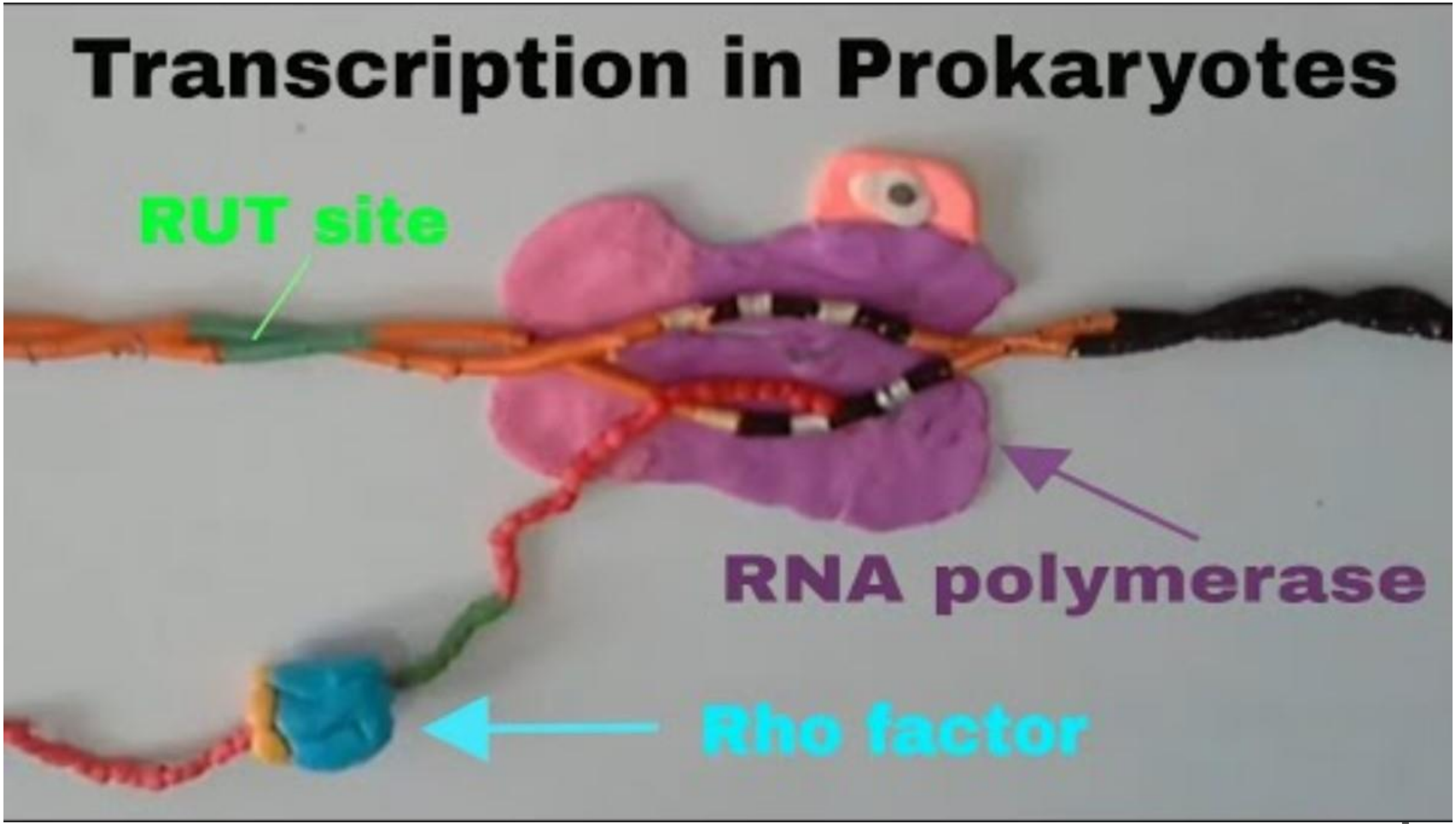
This type of termination seems less frequent and has a more complex mechanism:

- The rho factor is a protein consisting of 6 equal units and functions as a DNA-RNA-dependent helicase that unwinds the mixed RNA-DNA double helix; this process requires energy (ATP hydrolysis).
- The rho factor binds at a site near the 5‘ end of the nascent RNA strand, while the RNA polymerase is stationary, approaches the 3’ end and causes the release of the transcript. It is unclear what makes RNA polymerase stationary: the NusA protein also seems involved in this process.

RHO DEPENDENT TERMINATION

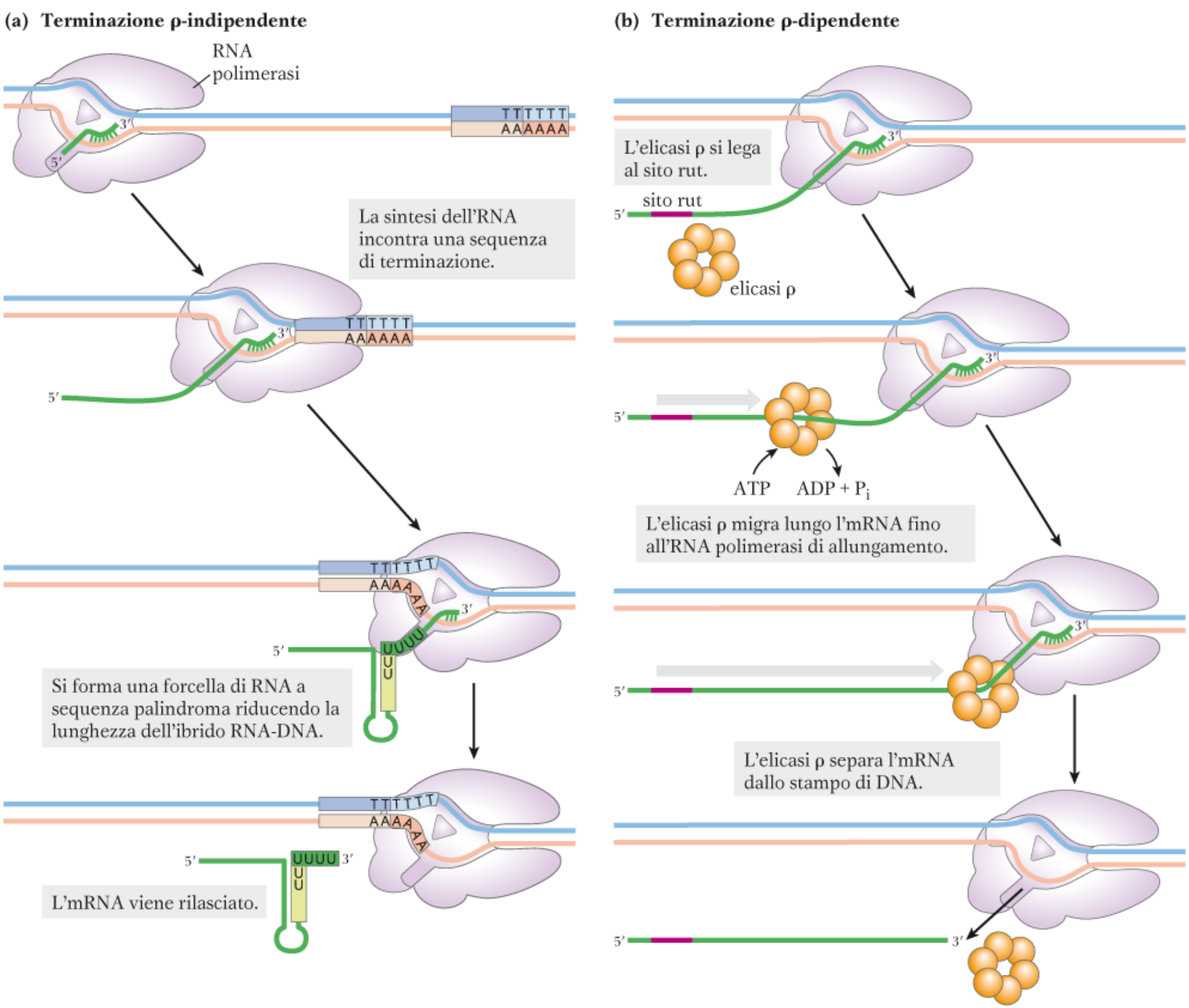


From 6:44


Transcription in Prokaryotes
RUT site
RNA polymerase
Rho factor

TRANSCRIPTION

Termination in *E. Coli*

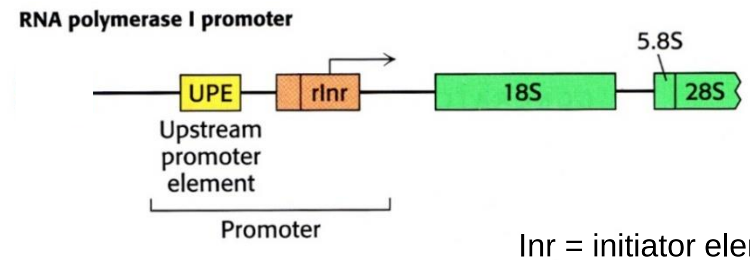


TRANSCRIPTION

eukaryotes

RNA polymerase I

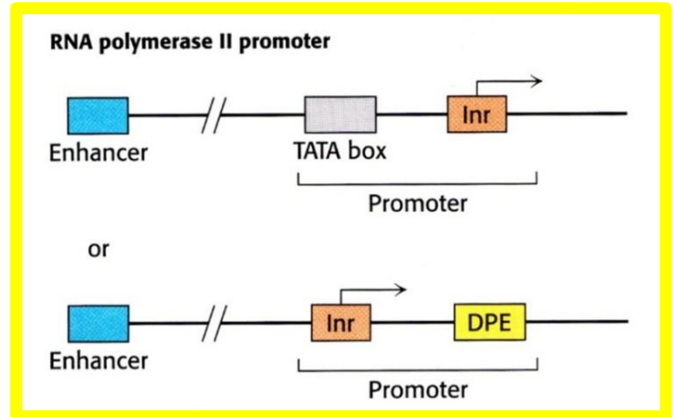
pre-rRNA (18s, 5.8s, 28s)



Inr = initiator elements

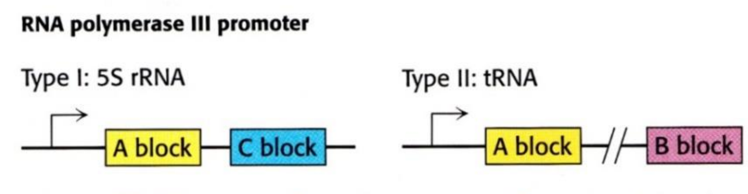
RNA polymerase II

mRNA, ..



RNA polymerase III

tRNA, rRNA 5s.

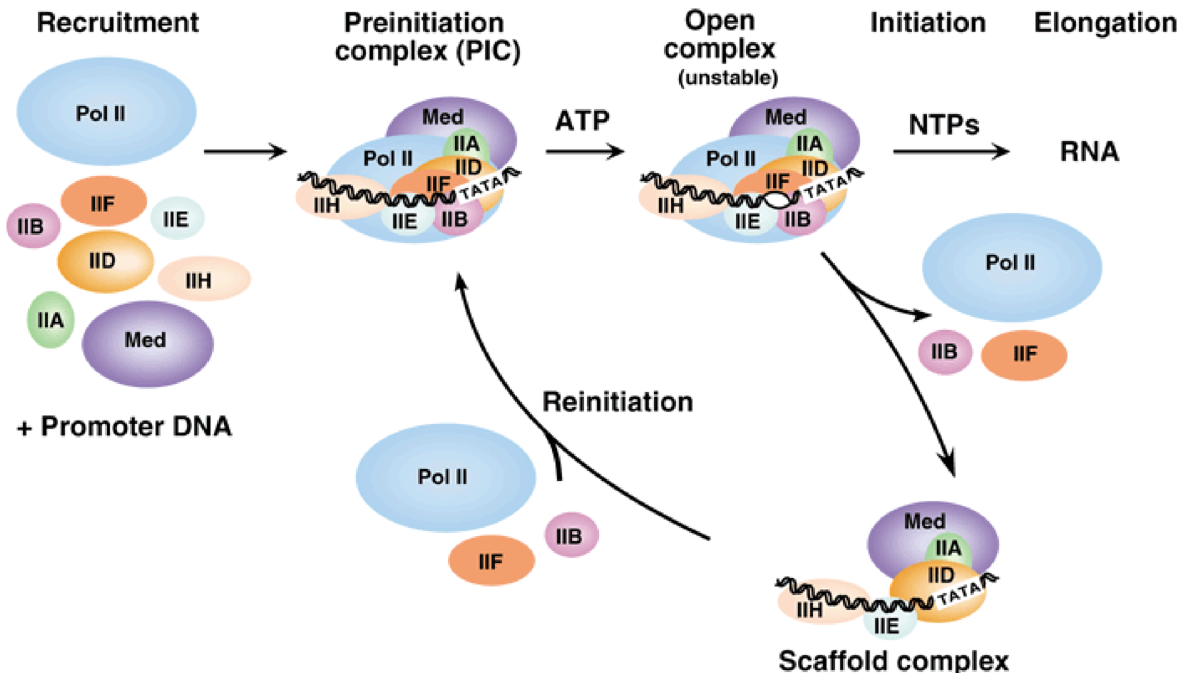


Each enzyme is a complex of 10 or more subunits. Some subunits are unique to each kind of polymerase, whereas other subunits appear in all three polymerases. The eukaryotic gene has two major parts: **the structural gene itself**, which is transcribed into RNA, and a **regulatory portion** that controls the transcription (a promoter or an enhancer).

TRANSCRIPTION

eukaryotes

Transcription factors recognize and bind to the TATA box, facilitating the recruitment of RNA polymerase II.



Nature Structural & Molecular Biology **11**,394–403 (2004)

Eukaryotic promoters are ‘dumb’ and only the assembly of TFs allows the action of RNA pol II

TRANSCRIPTION

eukaryotes

table 26–1

Proteins Required for Transcription at the RNA Polymerase II Promoters of Eukaryotes

Transcription factor	Number of subunits	Subunit M_r	Functions
Initiation			
RNA polymerase II	12	10,000–220,000	Catalyzes RNA synthesis
TBP (TATA-binding protein)	1	38,000	Specifically recognizes the TATA box
TFIIA	3	12,000, 19,000, 35,000	Stabilizes binding of TFIIB and TBP to the promoter
TFIIB	1	35,000	Binds to TBP; recruits RNA polymerase–TFIIF complex
TFIID	12	15,000–250,000	Interacts with positive and negative regulatory proteins
TFIIE	2	34,000, 57,000	Recruits TFIIH; ATPase and helicase activities
TFIIF	2	30,000, 74,000	Binds tightly to RNA polymerase II; binds to TFIIB and prevents binding of RNA polymerase to nonspecific DNA sequences
TFIIH	12	35,000–89,000	Unwinds DNA at promoter; phosphorylates RNA polymerase; recruits nucleotide-excision repair complex
Elongation*			
ELL†	1	80,000	
P-TEFb	2	43,000, 124,000	
SII (TFIIS)	1	38,000	
Elongin (SIII)	3	15,000, 18,000, 110,000	

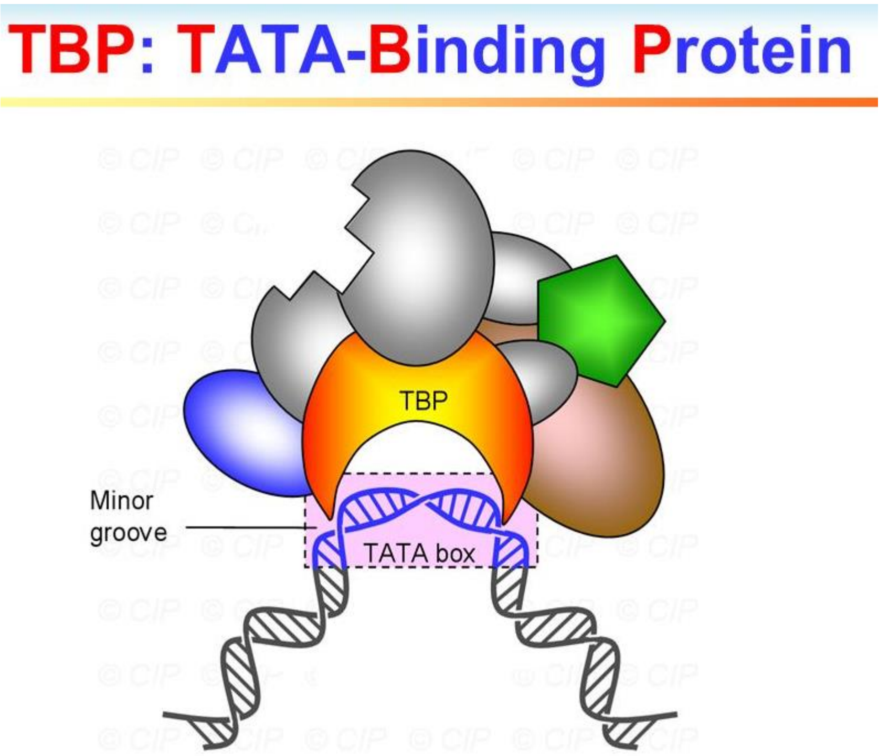
*All elongation factors suppress the pausing or arrest of transcription by the RNA polymerase II – TFIIF complex.

†The name is derived from the term *e*leven-nineteen *l*ysine-rich *l*eukemia. The gene for the factor ELL is the site of chromosomal recombination events frequently associated with the cancerous condition known as acute myeloid leukemia.

TRANSCRIPTION

Pre-initiation complex (PIC)

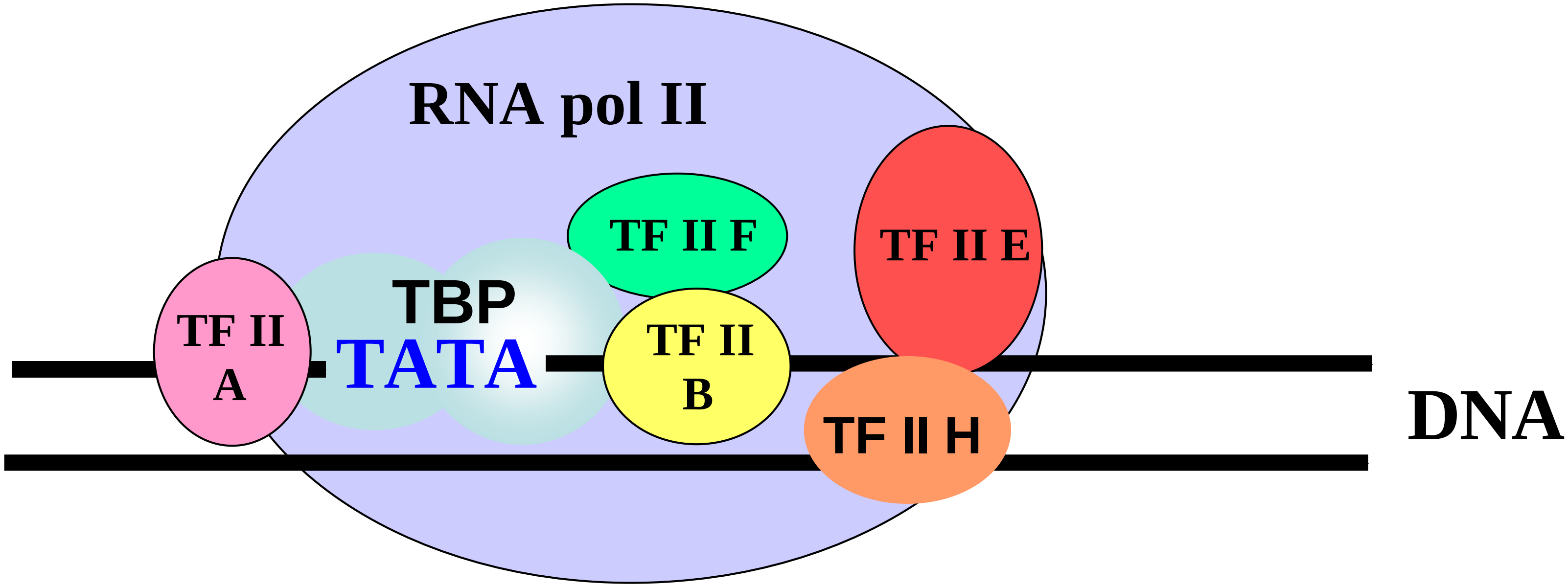
Eukaryotic RNA-pol II does NOT bind the promoter directly and associates with 6 transcription factors (TFs).
TBP binds the TATA-box forming the TFIID complex (general transcription factor D for RNA pol II)



TBP has a saddle-like shape. The "brackets" of the saddle are rich in phenylalanine residues, which contain aromatic rings. These rings resemble nitrogenous bases in size and shape, enabling TBP to interact effectively with the DNA. After binding to the TATA box, TBP helps recruit additional general transcription factors (such as TFIIB, TFIIF, TFIIE, and TFIIH) and RNA polymerase II itself.This assembly is necessary for forming the pre-initiation complex (PIC), which is essential for starting transcription.

TRANSCRIPTION

Pre-initiation complex (PIC)

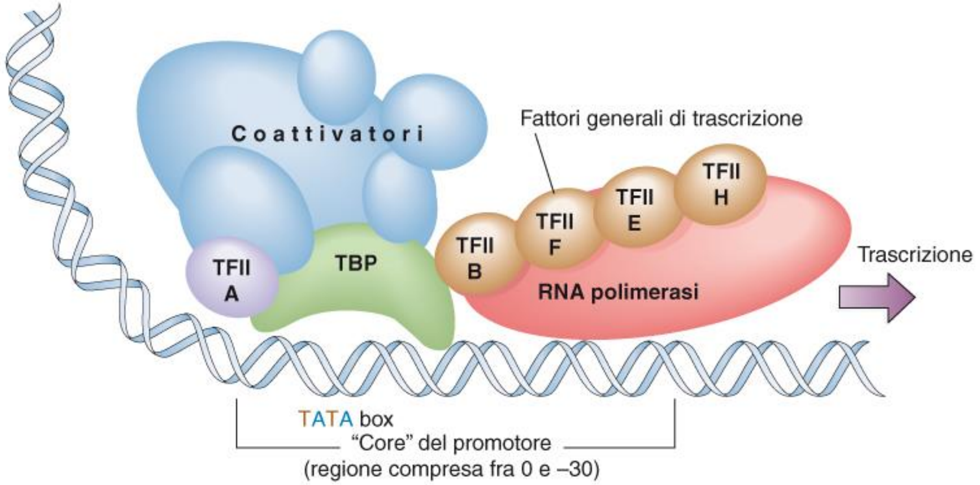


TRANSCRIPTION

Transcription factors

General transcription factors (TFII): recruitment of RNA polymerase II (low transcription rate)

Specific transcription factors: +/- transcription rate (coactivators)

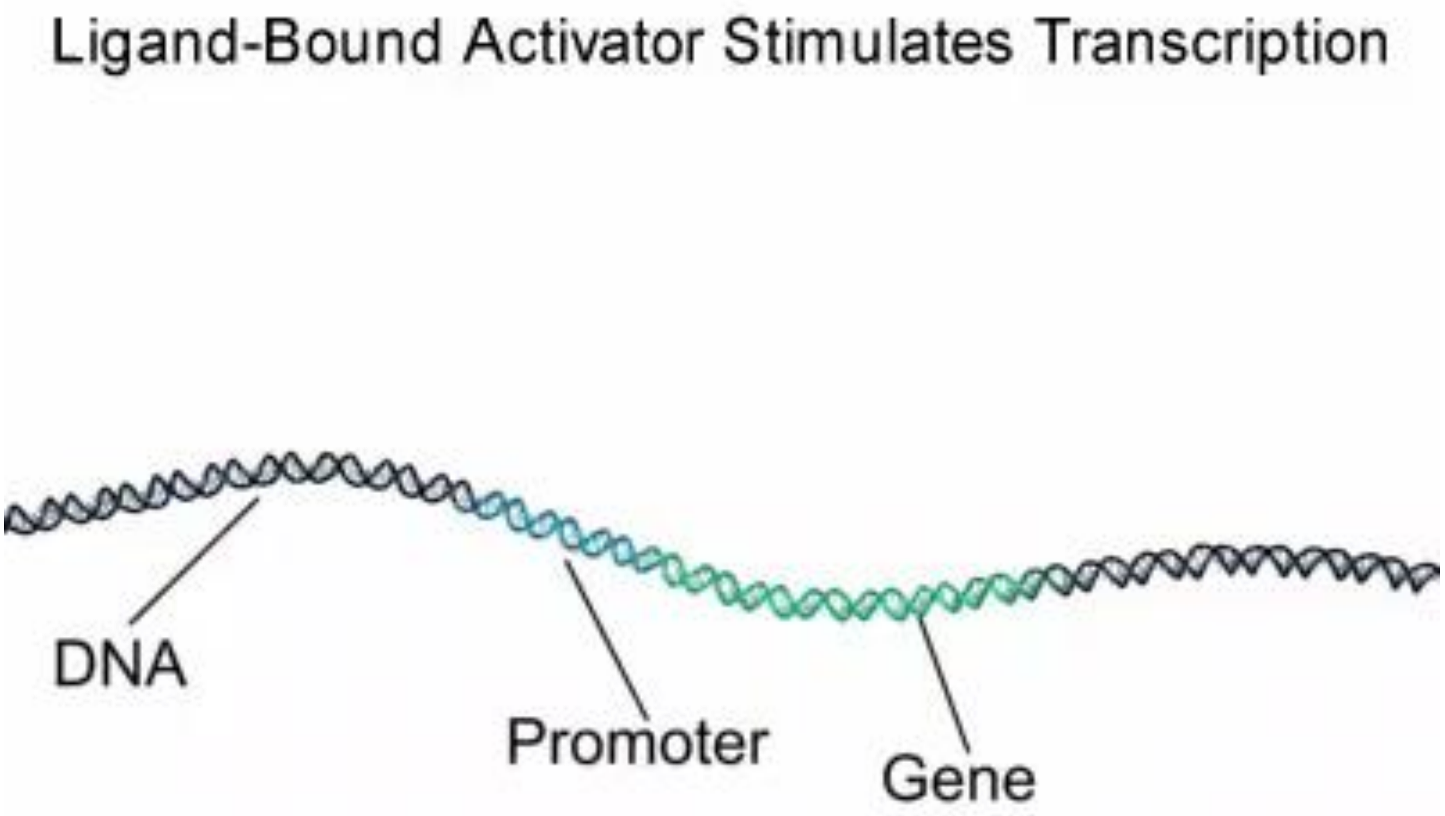


M. Lieberman **Marks Biochimica medica** Copyright 2010 C.E.A. Casa Editrice Ambrosiana

SPECIFIC TRANSCRIPTION FACTORS BIND TO REGULATORY SEQUENCES THAT ARE LOCATED UPSTREAM (6-20 BP). THESE FACTORS ARE DIMERS AND INTERACT WITH THE PIC AND CAN INCREASE THE FREQUENCY OF TRANSCRIPTION (ACTIVATORS) OR DECREASE IT (REPRESSORS). THEY ARE SPECIFIC TO A GENE OR GROUP OF GENES.

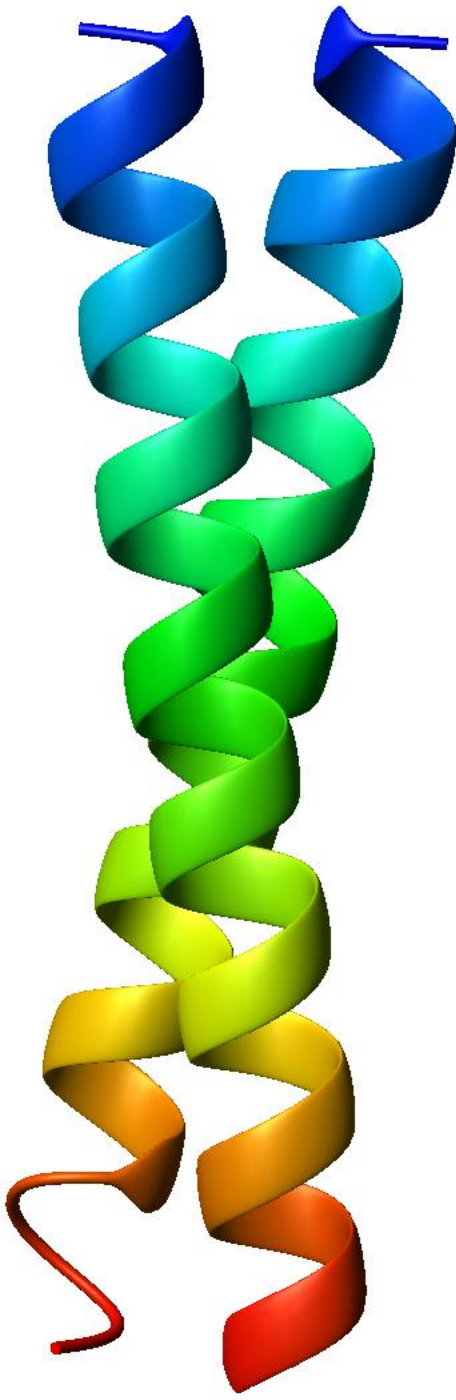
TRANSCRIPTION

Transcription factors



TRANSCRIPTION

Transcription factors

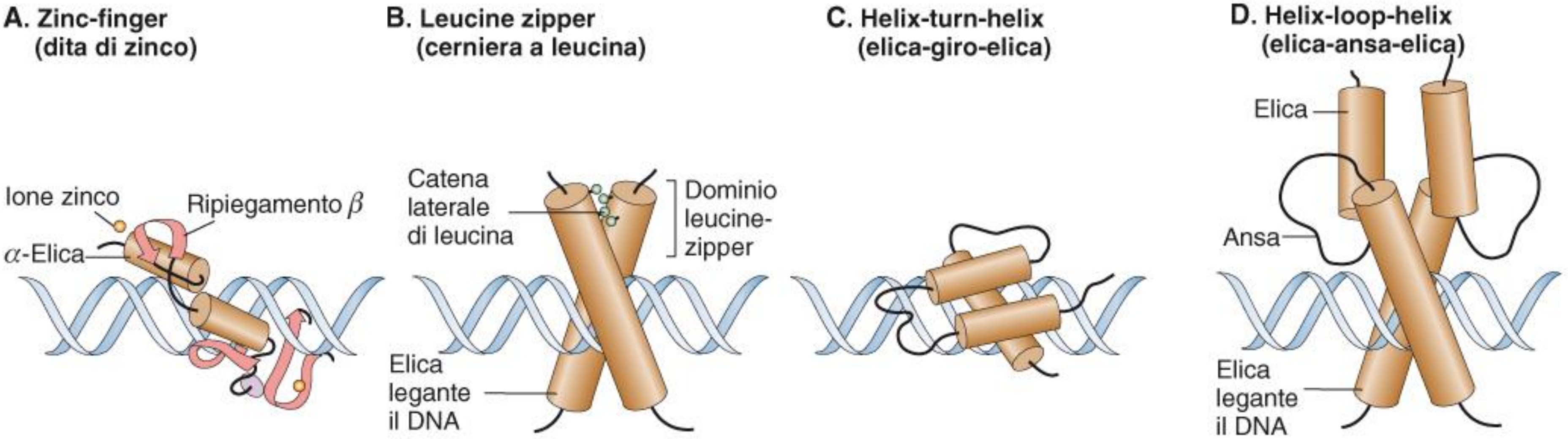


- A common pattern in DNA-binding proteins is the presence of α-helical segments that interact directly with the major groove of the B-form of DNA

- Proteins can recognise specific sites in DNA

TRANSCRIPTION

Transcription factors



Structural motifs of transcription factors for DNA recognition (α-helix in the major DNA groove).

Most binding sites (response elements) for specific factors are 15-20 bp long. Many transcription factors respond to hormones, nutrients or second messengers.

TRANSCRIPTION

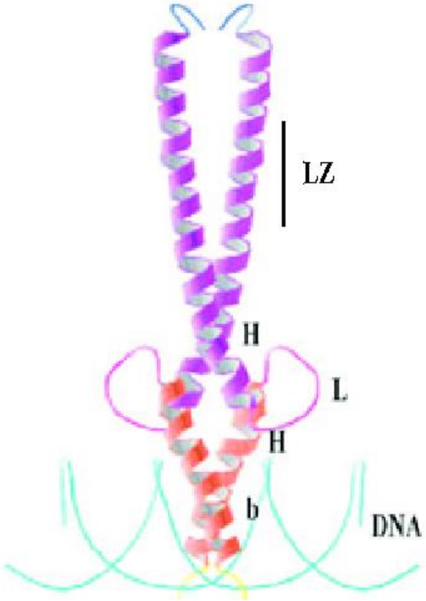
Transcription factors

The crucial point is the BINDING between protein residues, and the skeleton of DNA : the DNA-binding domain

- Most contacts occur at the level of the major groove of the B-DNA
- 80% of regulatory proteins can be assigned to one of these three classes: helix-turn-helix (helix-turn-helix or 'HTH'), zinc finger (zinc finger or 'Zn-finger') and leucine zip (leucine zip or 'bZIP')
- In addition to the DNA-binding domain, these proteins also have other domains for interaction with other proteins

TRANSCRIPTION
Transcription factors

- The **helix-turn-helix motif** was first identified in three prokaryotic proteins:
- All of these proteins bind as dimers to symmetrical sites on the DNA all contain two α-helices separated by a β-turn loop
- The C-terminal helix fits into the major groove of the DNA;
- the N-terminal helix stabilises it through hydrophobic interactions

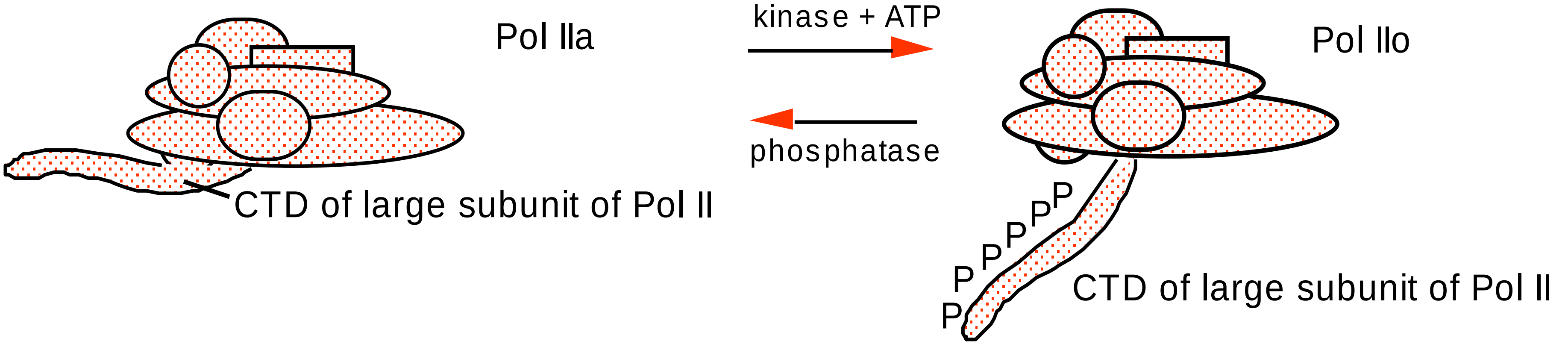


TRANSCRIPTION

RNA polymerase II phosphorylation

TF II H

Eukaryotic RNA polymerase II



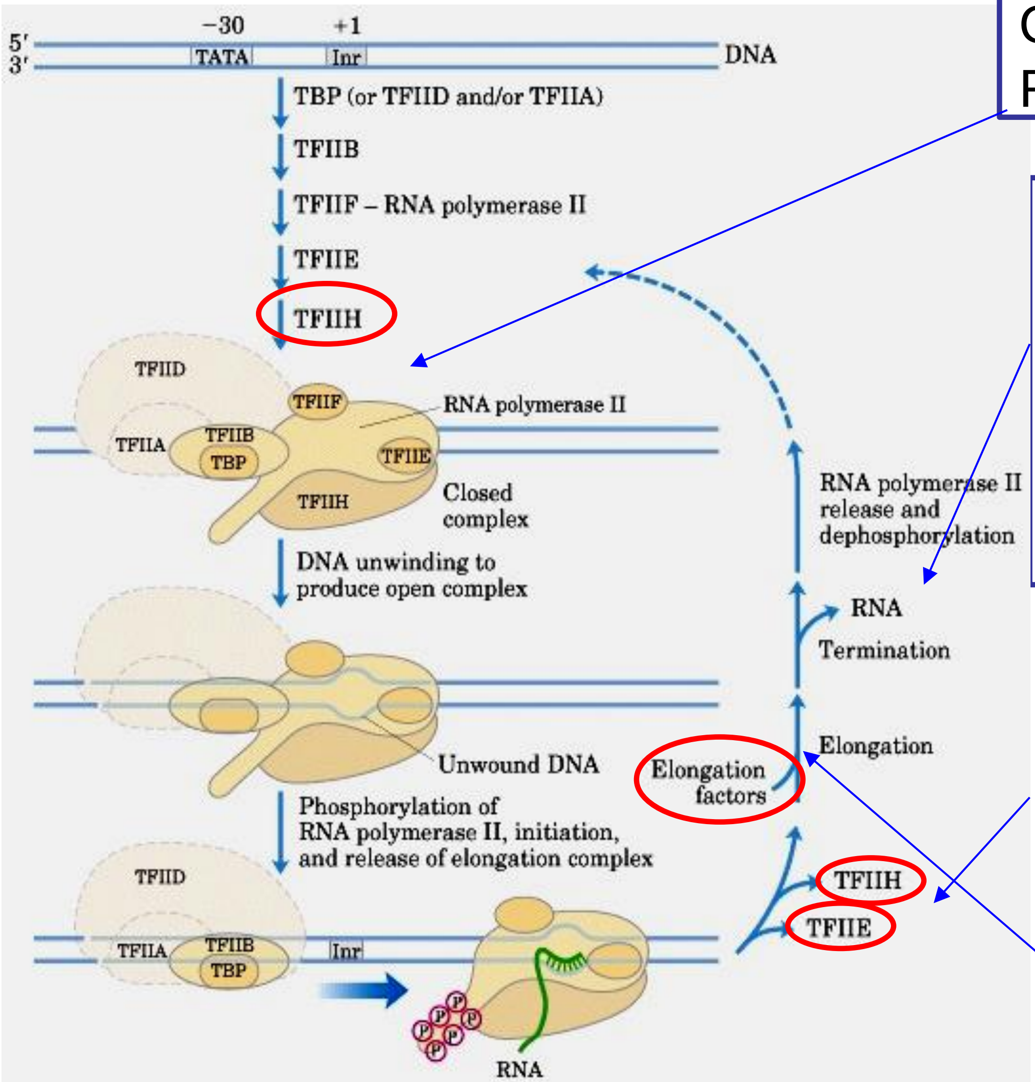
CTD contains the repetitive sequence (YSPTSPS)$_{26\text{-}50}$

Tyr Ser Pro Thr Ser Pro Ser

The phosphorylation of RNA polymerase II (Pol II) by the transcription factor TFIIH plays a key role in transitioning Pol II from the initiation phase to the elongation phase of transcription. Certain elongation factors bind to the phosphorylated C-terminal domain (CTD) and increase the elongation rate.

TRANSCRIPTION

Elongation and termination



Only NON-PHOSPHORYLATED RNA polymerase II enters the PIC.

Once the RNA has been transcribed, the elongation factors dissociate, the CTD is dephosphorylated by the action of certain termination factors and the polymerase is released.

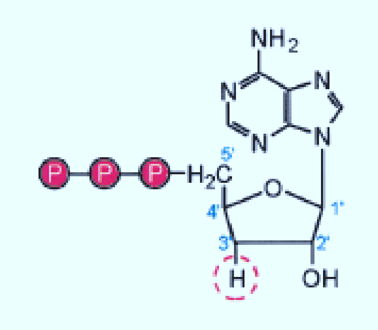
Release of TFIIE and then TFIIH after the synthesis of the first 60-70nt.

ELONGATION FACTORS come into play.

TRANSCRIPTION

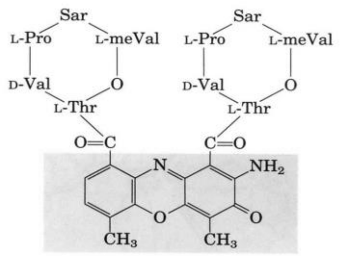
RNA polymerase as a therapeutic target

3'-deoxyadenosine



Nucleotide analogue, inhibits transcription. Antifungal activity

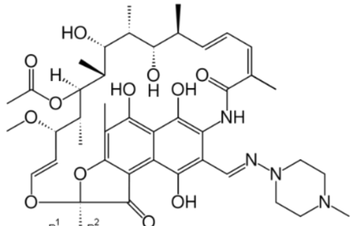
Actinomicin D



Actinomycin D

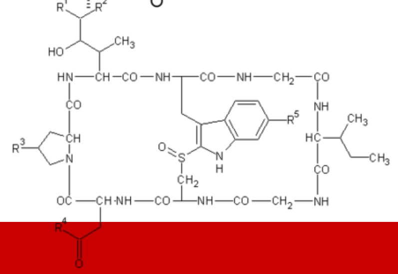
Actinomycin D can interfere with the stability and assembly of PIC by disrupting interactions between DNA and transcription factors, ultimately affecting the recruitment of RNA polymerase II.Antitumour activity

Rifamycin



Rifamycin binds to bacterial RNA polymerase, preventing the start of transcription. Used in the treatment of tuberculosis and meningitis.

α-Amanitin



Cyclic polypeptide contained in fungi (amanita phalloides). RNA polymerase II inhibitor, lethal.

TRANSCRIPTION

eukaryotic mRNA maturation

Introns: non-coding RNA fragments that interrupt coding regions of mRNA (in human genes there are on average 8 introns per gene)

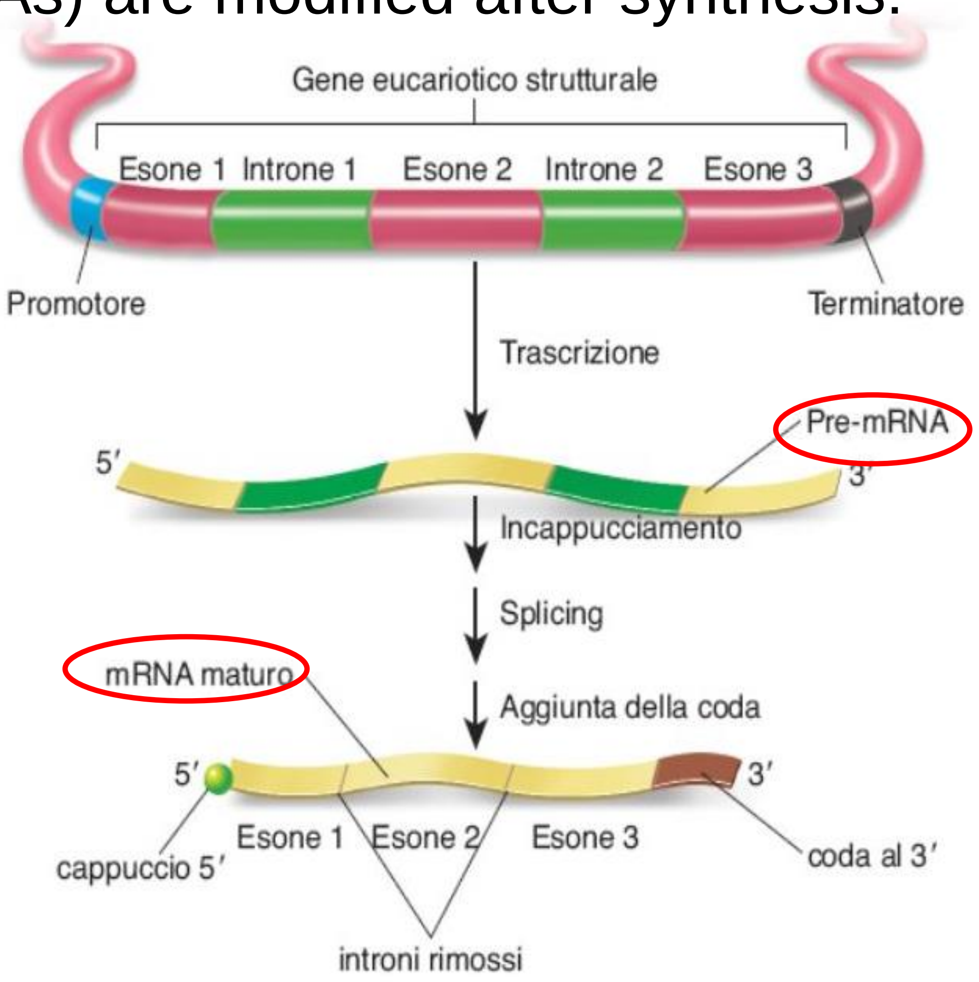
Exons: coding mRNA

TRANSCRIPTION

eukaryotic mRNA maturation

Many RNAs (all eukaryotic RNAs) are modified after synthesis:

- protection from nucleases
- formation of bonds with regulatory proteins
- coordination between transcription and translation

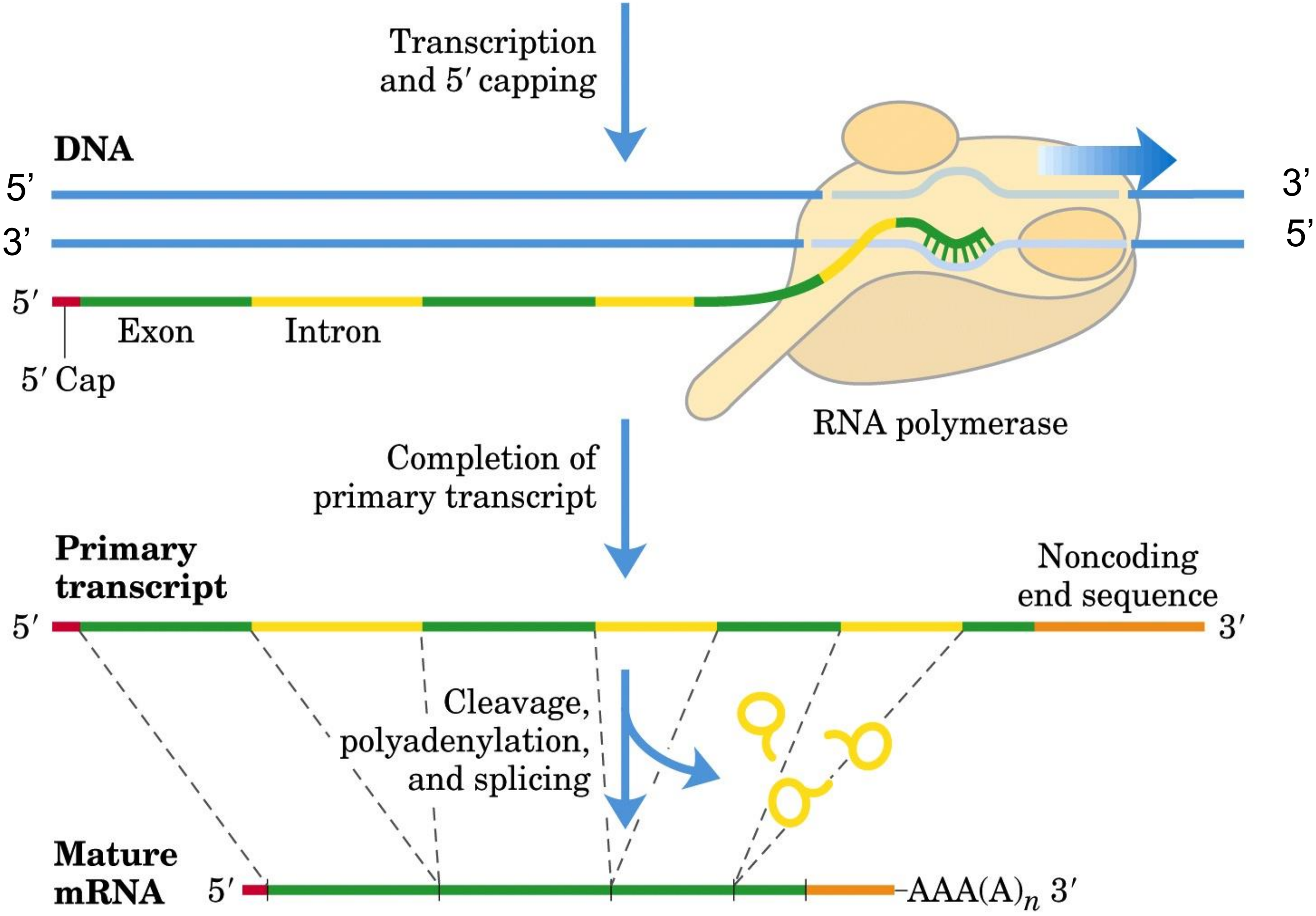


Stages of maturation are:

- Capping (or hooding)
- Polyadenylation
- Splicing

TRANSCRIPTION

eukaryotic mRNA maturation

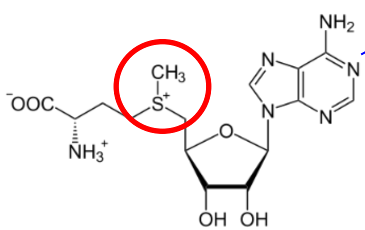


TRANSCRIPTION

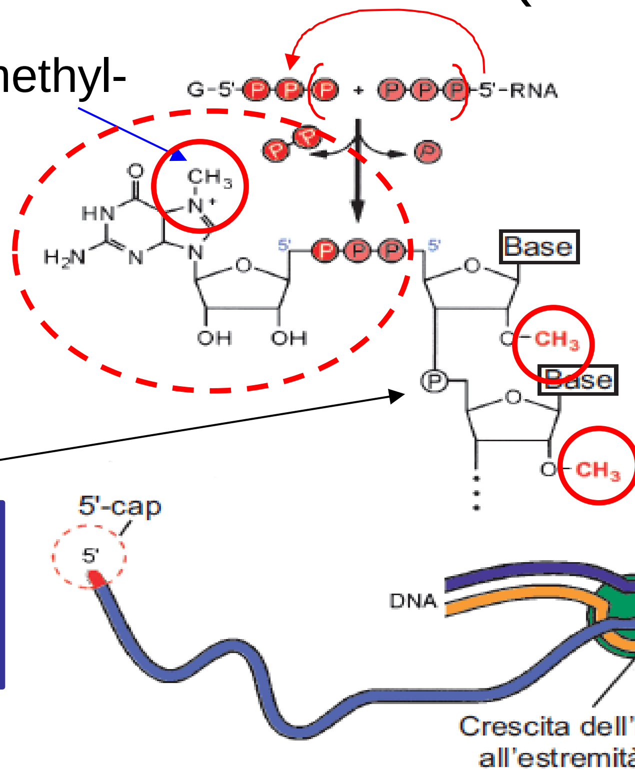
eukaryotic mRNA maturation: capping

The 5' end of the nascent RNA is sealed by binding with a **guanosine triphosphate (GTP).** A cap is formed at the 5' end that protects the RNA from rapid **degradation** and is a **recognition** signal for the ribosomes (translation initiation).

S-adenosyl methionine (SAM)



7 methyl-

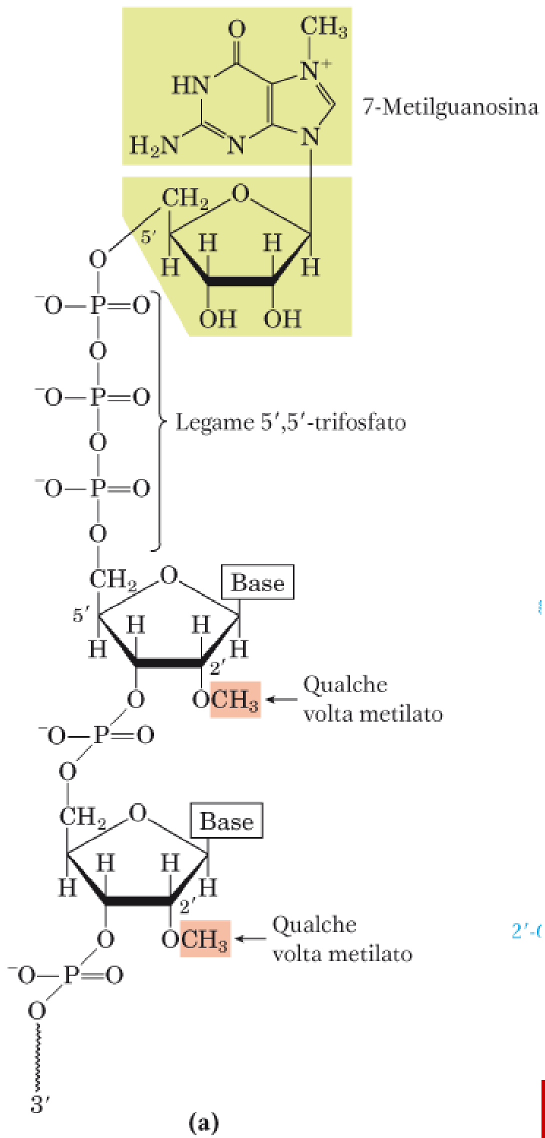


The first step involves the action of <u>RNA triphosphatase</u>, which removes one of the terminal phosphate groups from the 5' end of the nascent RNA.
Next, RNA <u>guanylyltransferase</u> catalyzes the transfer of a GMP from GTP to this diphosphate end, forming a unique 5' to 5' triphosphate bond. guanine is methylated in N-7, through <u>guanine-N7 methyltransferase</u>.

Other $-CH_3$ groups can be added to –OH 2' of 2 adjacent nucleotides

TRANSCRIPTION

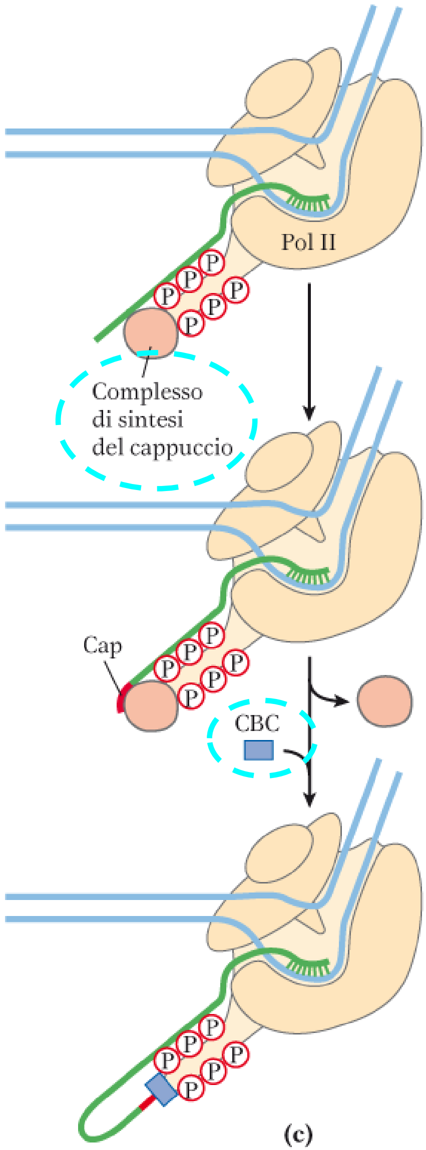
eukaryotic mRNA maturation: capping



The capping process begins shortly after transcription starts, typically when about 25-30 nucleotides of RNA have been synthesized.

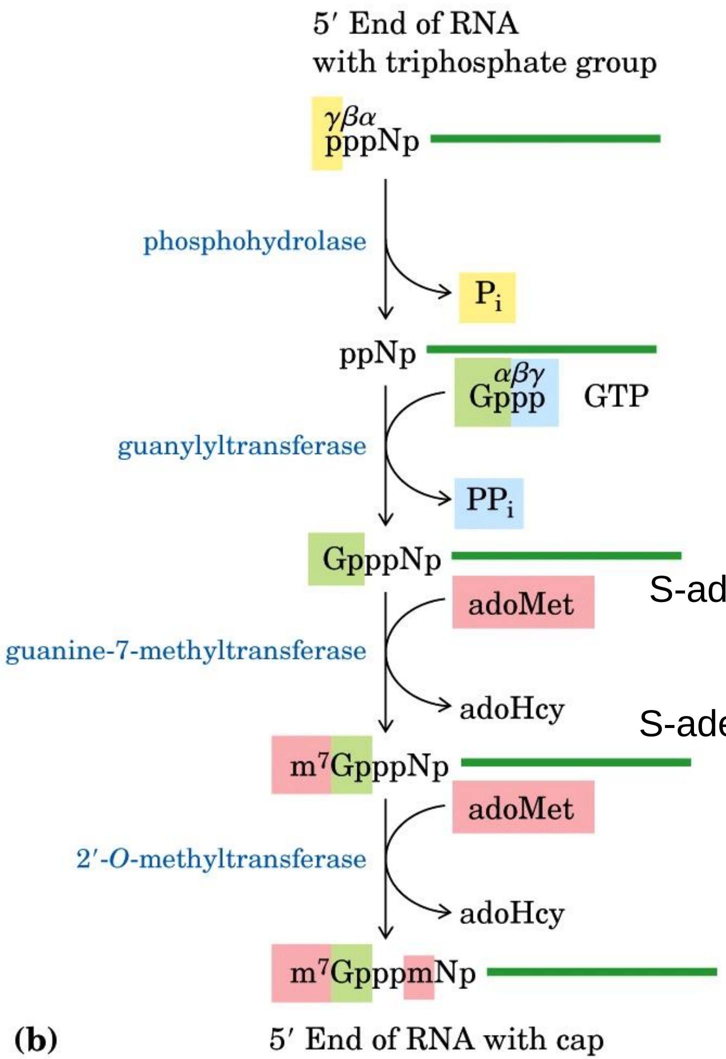
The cap binds to the CBC which anchors it to the tail of the phosphorylated CTD

CBC: Cap Binding Complex



TRANSCRIPTION

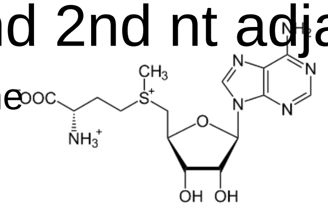
eukaryotic mRNA maturation: capping



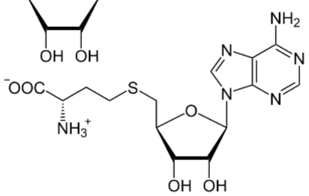
Capping MECHANISM

1. Condensation of 1 GTP with 2 phosphates (P) at the 5' end of mRNA (5' to 5' triphosphate bond).
2. G is methylated in N7
3. -CH3 groups are added at C-2' on the 1st and 2nd nt adjacent to the capping

S-adenosylmethionine

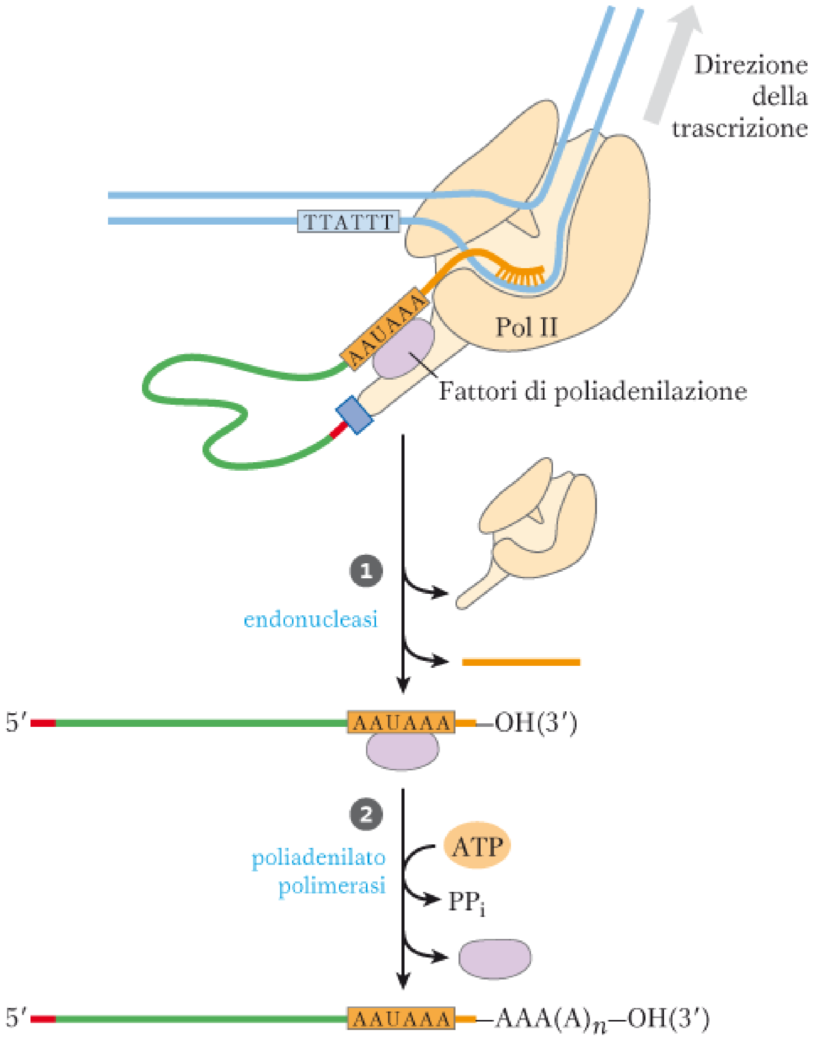


S-adenosylomocysteine



TRANSCRIPTION

mRNA maturation: polyadenilation of 3’



The addition of a poly(A) tail to the 3' end of the pre-mRNA (protection from enzymatic degradation).

1. **Cleavage**: a specific sequence in the pre-mRNA signals is synthetized at the 3' end, typically downstream of the polyadenylation signal (AAUAAA). A complex of proteins known as the Cleavage and Polyadenylation Complex (CPC) recognizes this signal and cleaves the pre-mRNA .
2. **Addition of the Poly(A) Tail**: Following cleavage, poly(A) polymerase catalyzes the addition of adenine nucleotides to the newly formed 3' end, creating a poly(A) tail, ranging from 50 to 250 adenine residues in mammals.

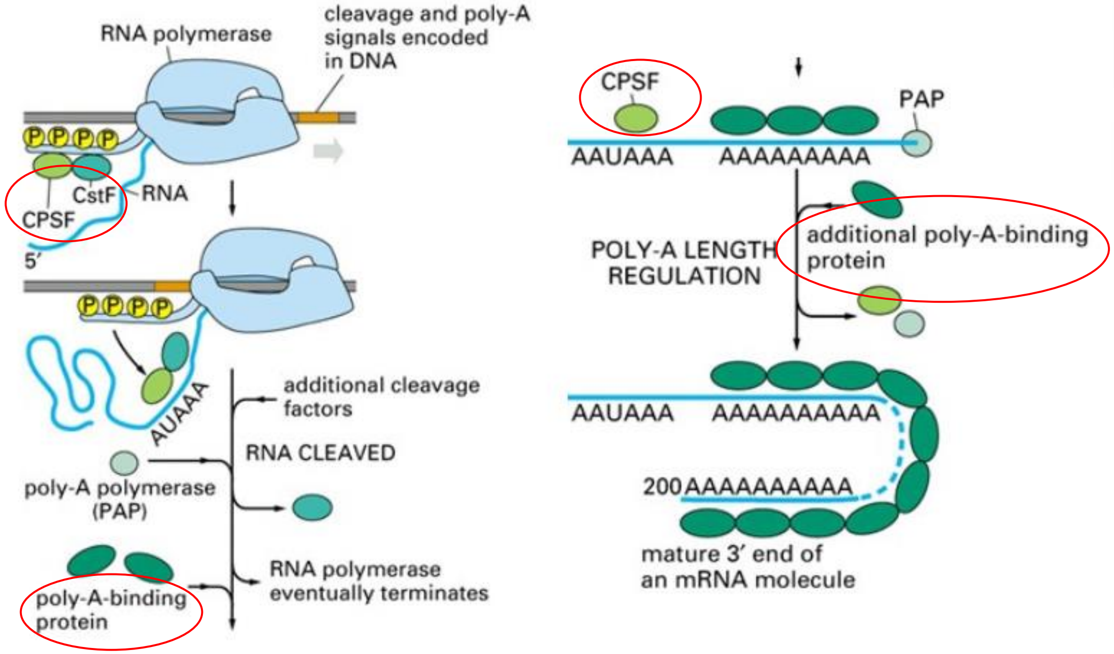
TRANSCRIPTION

mRNA maturation: polyadenilation of 3’

Poly-adenylate (poly(A)) polymerase (PAP) adds adenyl nucleotides to mRNA, forming a segment of poly-A residues.

CPSF: cleavage and polyadenylation specificity factor

CstF: cleavage stimulation factor



- Initially, about 10 nucleotides are added.
- Subsequently, the binding of Poly(A) Binding Proteins (PABPs) to the tail enhances the activity of poly(A) polymerase and stabilizes the mRNA.This stimulates poly(A) polymerase to add further A nucleotides.

TRANSCRIPTION

mRNA maturation: polyadenilation of 3’

(poli-(A)) polymerase reaction

$$\mathbf{RNA} + n\mathbf{ATP} \rightarrow \mathbf{RNA} - (\mathbf{AMP})_n + n\mathbf{PP_i}$$

n = 50-250

The poly(A) tail has essential functions:

Stability: It protects mRNA from degradation by exonucleases, thereby extending its lifespan within the cell.

Translation Efficiency: The presence of a poly(A) tail is crucial for efficient translation initiation by helping recruit ribosomal components.

Nuclear Export: The poly(A) tail is involved in transporting processed mRNA from the nucleus to the cytoplasm (where it is a bit shortened), ensuring that mRNA is available for translation.

mRNA maturation: polyadenilation of 3’

mRNAs have a half-life that depends on the cell type and is related to the function of the encoded protein. The degradation occurs via ribonucleases in eukaryotes (active in the direction 3’ → 5’), which remove the poly(A) tail and the capping. In bacteria, polynucleotide phosphorylases are the main players: these enzymes act on the 5' end, hydrolysing the phosphodiester bond.

Prokaryotes		
Degradation at the 5' end begins immediately		
Eukaryotes		
cFOS	Half-life: 10-30 min	poly-A tail short
Hemoglobin	Half-life: 24 ore	poly-A tail long

TRANSCRIPTION

mRNA maturation: splicing

Exons are the coding regions.

Introns are 'intermediate' sequences not present in mature RNA and can account for up to 90% of the primary transcript sequences.

Introns are removed and different combinations of exons form different mRNAs that will express different protein products from the same, single gene.

Humans have approx. 30,000 genes capable of producing approx. 100,000 proteins.

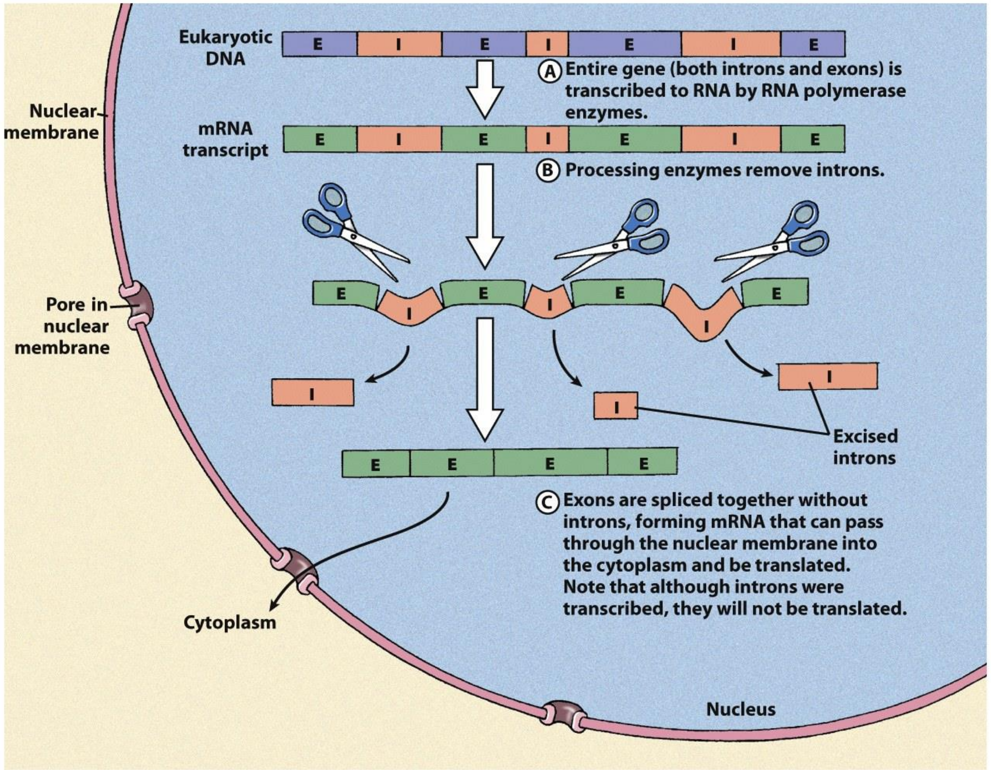
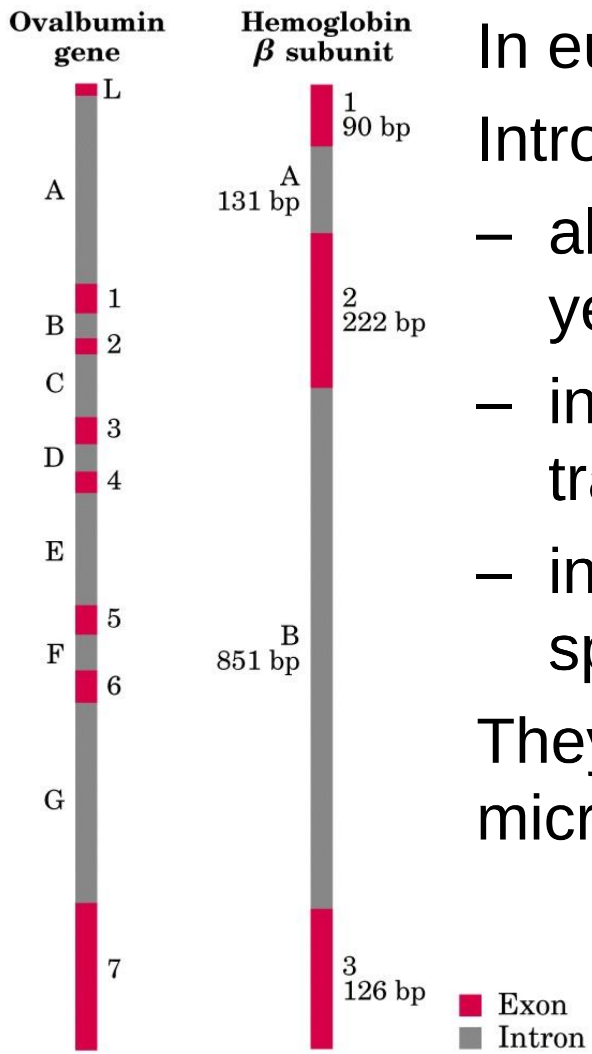


Figure 7-6 Microbiology, 7/e

TRANSCRIPTION

mRNA maturation: splicing



In eukaryotes, exons are < 1000 nt

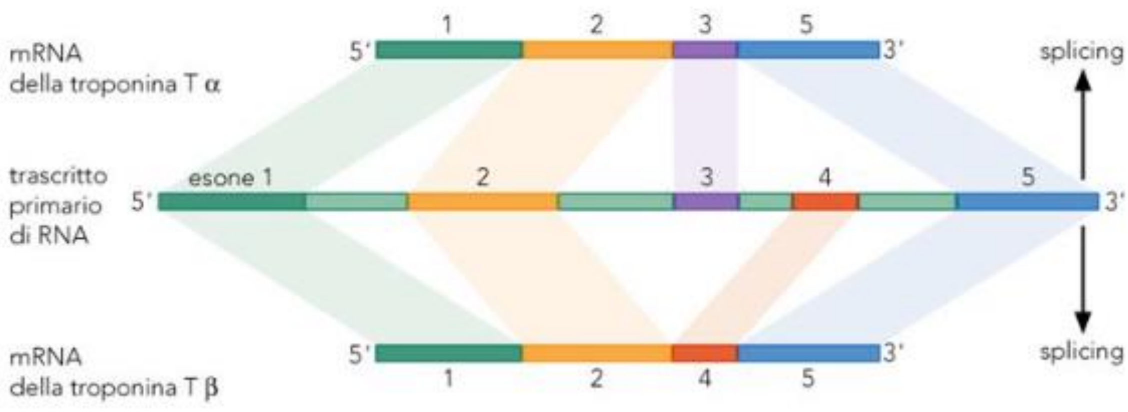
Introns are

- almost non-existent in prokaryotes and yeasts
- increase the energy cost of replication and transcription
- increase the flexibility of genes (alternative splicing)

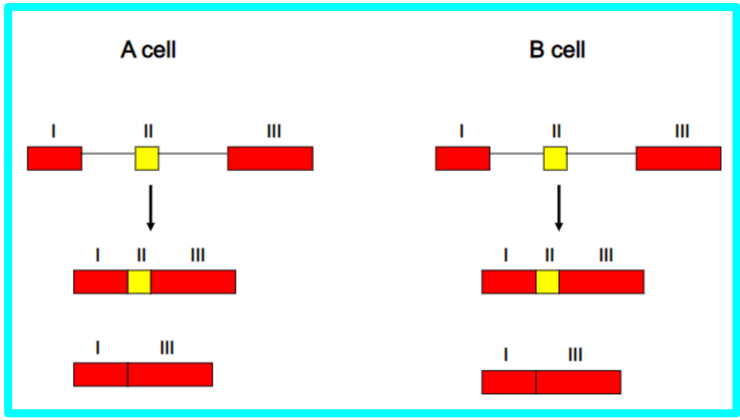
They can generate **miRtrons** (intron-derived microRNAs) with regulatory functions

TRANSCRIPTION

mRNA maturation: alternative splicing



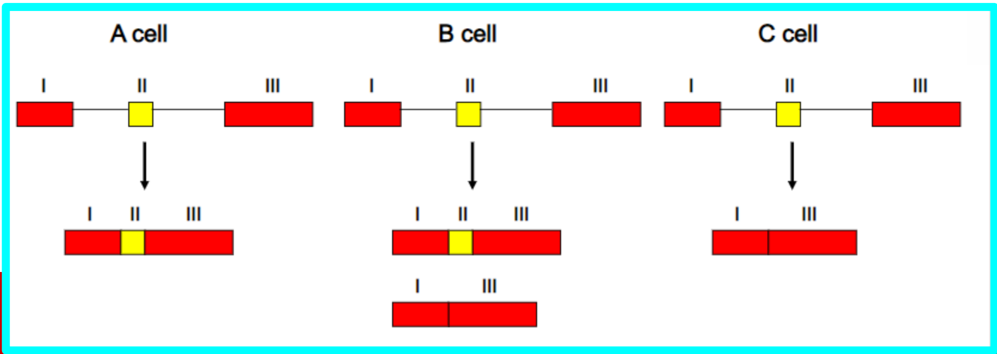
E.g. the same primary transcript relative to the gene for troponin (muscle protein) can generate two different mRNAs, one in which exon 3 is present and one in which exon 4 is present



Alternative splicing can be:

Constitutive: standard process by which all exons of a gene are joined together in a fixed order, while introns are removed.

Regulated: allows for selective inclusion or exclusion of certain exons in response to specific signals or conditions. This results in different mRNA transcripts being produced from the same gene, leading to diverse protein isoforms.



TRANSCRIPTION

mRNA maturation: splicing

INTRONS:

- GROUP I
- GROUP II

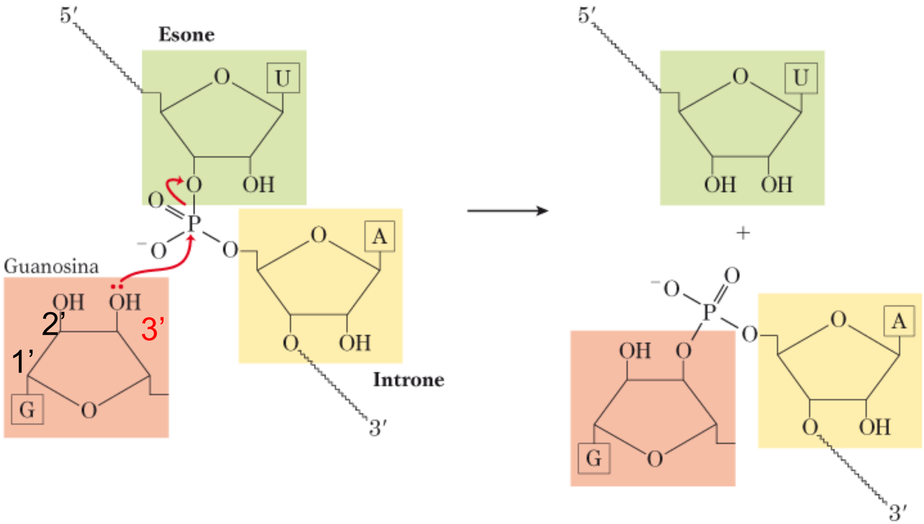
SELF-SPLICING (no ATP)

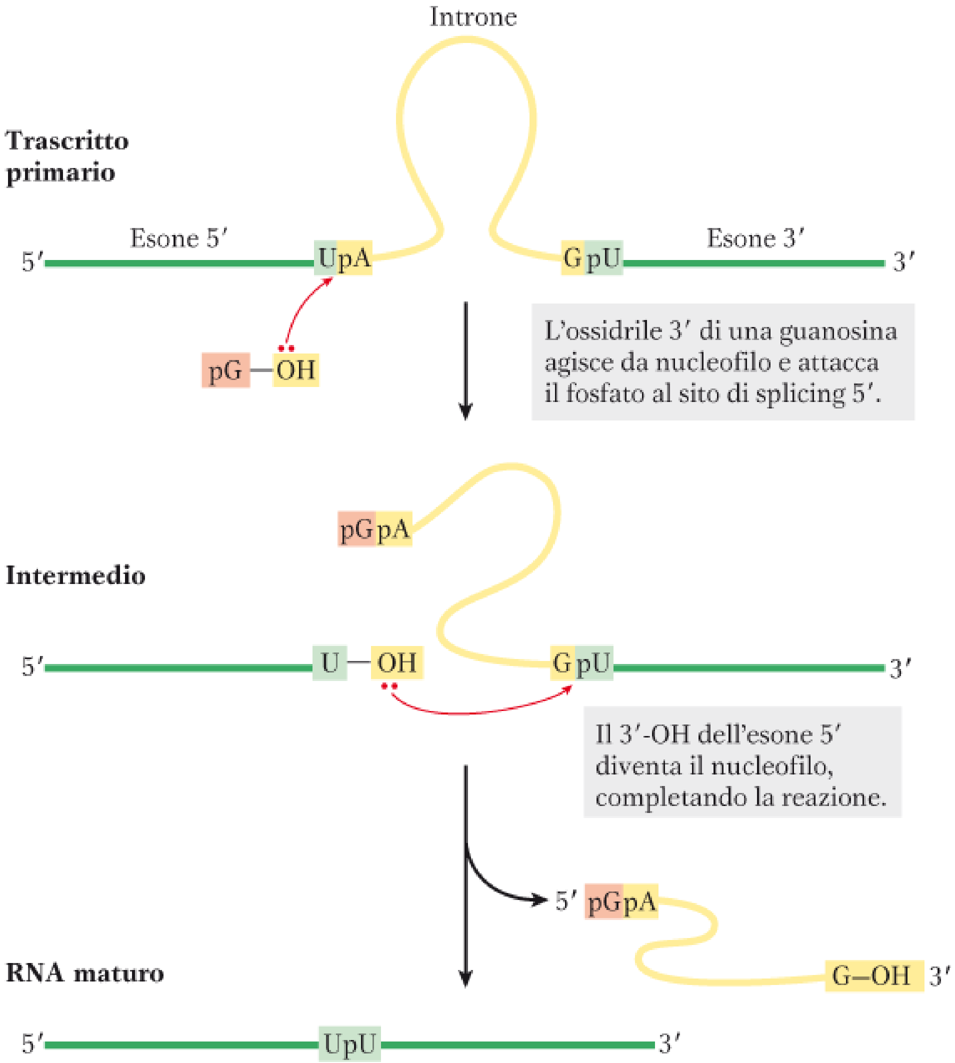
- SPLICEOSOMAL INTRONS
- INTRONS IN tRNA removed by an endonuclease and ATP

TRANSCRIPTION

mRNA maturation: splicing

Transesterification reaction for GROUP I INTRONES

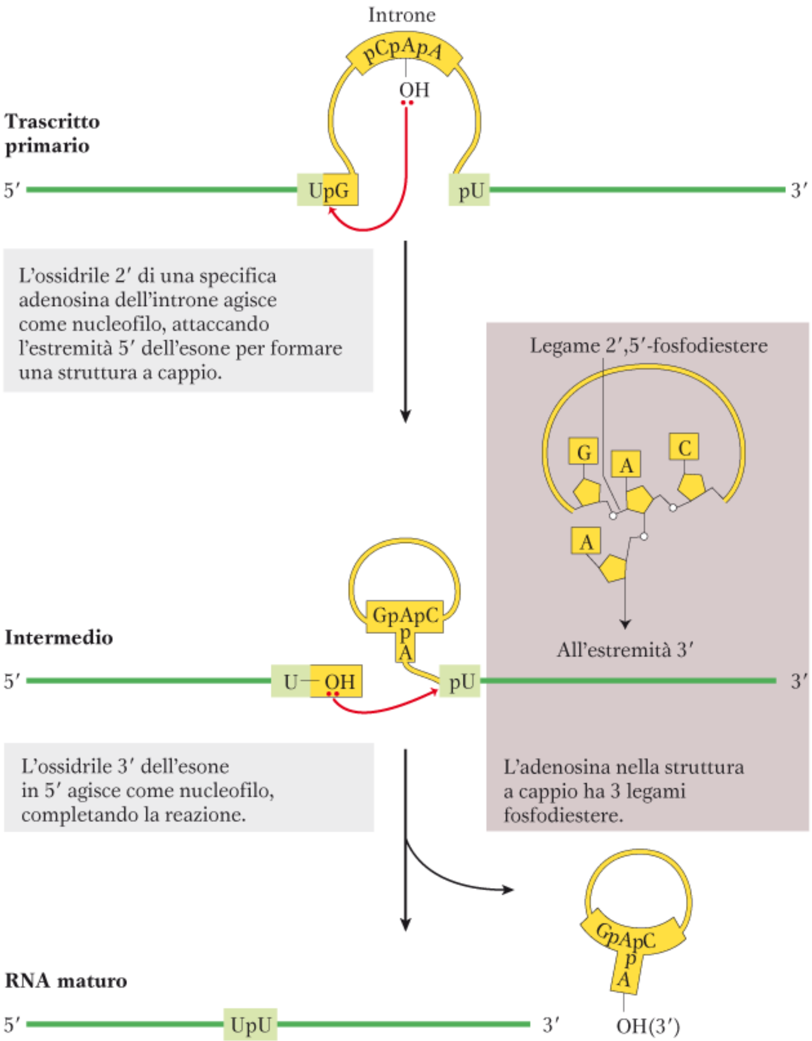




A guanine nucleoside is needed (guanosine/GMP/GDP/GTP)

TRANSCRIPTION

mRNA maturation: splicing



GROUP II INTRONS self-splicing mechanism

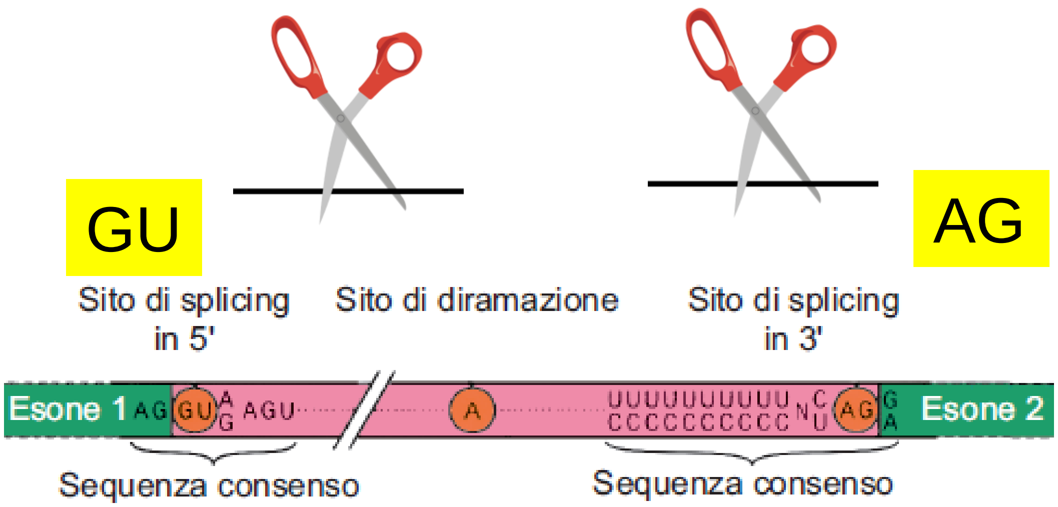
A nucleophilic attack occurs from the -OH in 2' of an intron-internal ADENOSIN; a loop is formed.

TRANSCRIPTION

mRNA maturation: splicing

SPLICEOSOMAL INTRONS: are excised by the action of the spliceosome (eukaryotes). ATP hydrolysis is necessary for the catalytic activation of the spliceosome.

It consists of RNA-protein complexes, called small nuclear ribonucleoproteins (snRNPs), called 'snurps'.
Each snRNP consists of a U-rich 100-200 nt RNA (snRNA) and 8 proteins that recognise consensus sequences.



The GU and AG sequences indicate the cleavage sites.

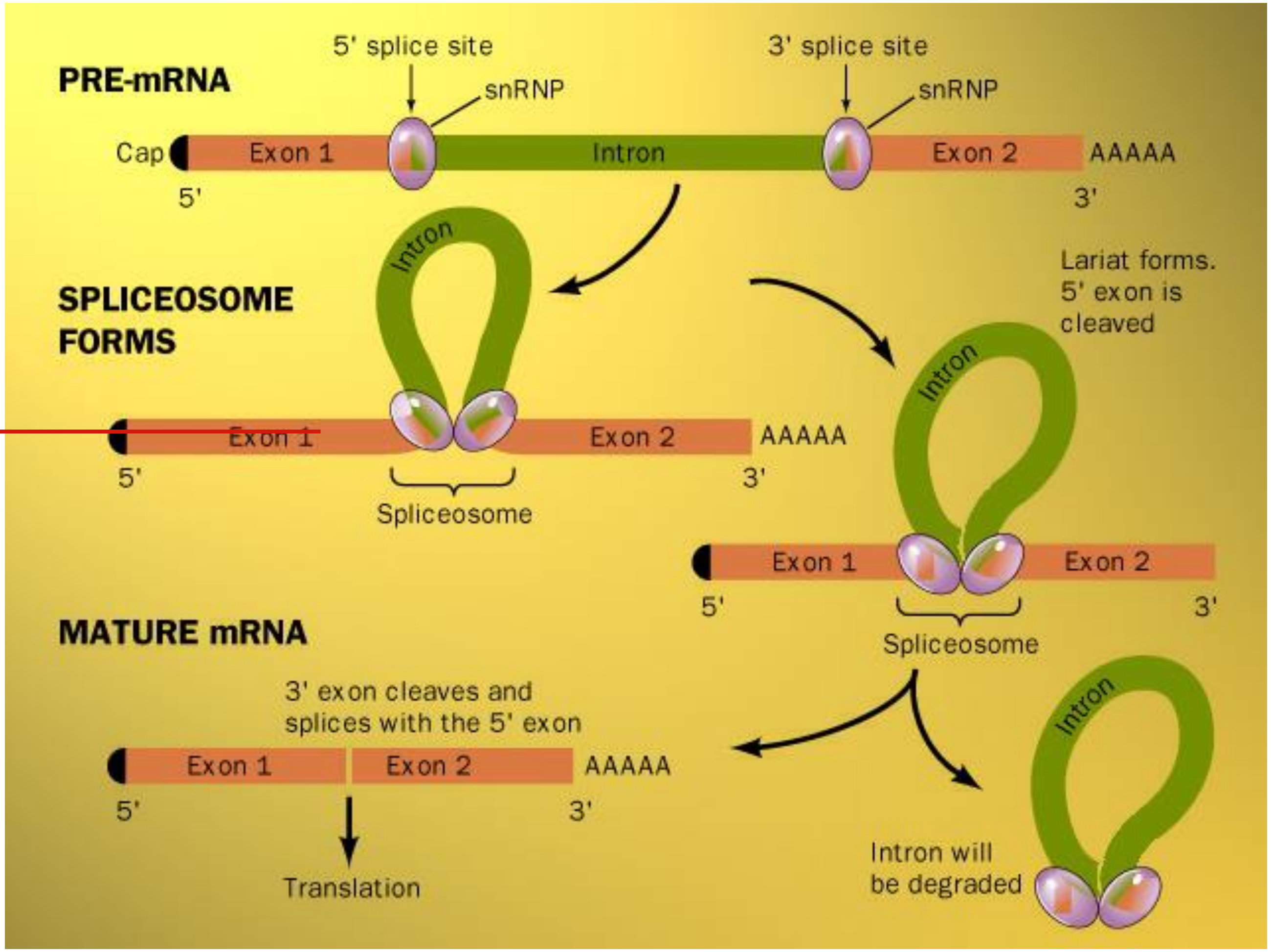
TRANSCRIPTION

mRNA maturation: splicing

2'–OH of A attacks 5'-P on the 5' end of the intron

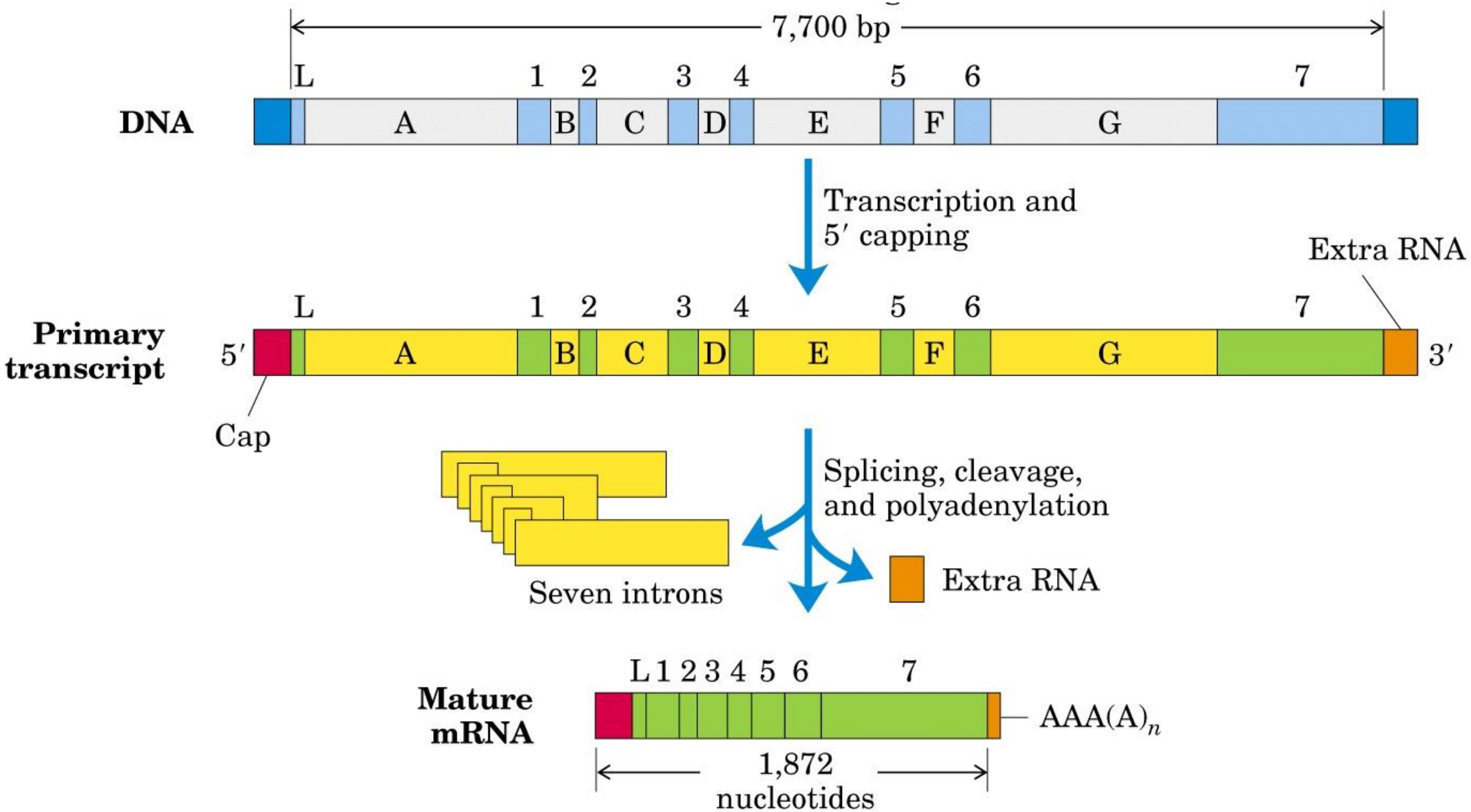
High energy cost: While the initial commitment complex formation can occur in an ATP-independent manner, the recruitment of snRNP and the formation of the active spliceosome, are ATP-dependent.

After splicing, ATP is neede for the disassembly of the spliceosome through the action of helicases.



TRANSCRIPTION

pre-mRNA processing in mRNA



From Lehninger: Principles of Biochemistry, 7th Ed.

TRANSCRIPTION

pre-mRNA processing in mRNA

The presence of introns is associated with gene flexibility: several products can be obtained from the same gene.

Examples:

- ✓ Different versions of the same protein (isoforms) in different tissues;
- ✓ Isoforms with different intracellular locations (cytosol or mitochondria)

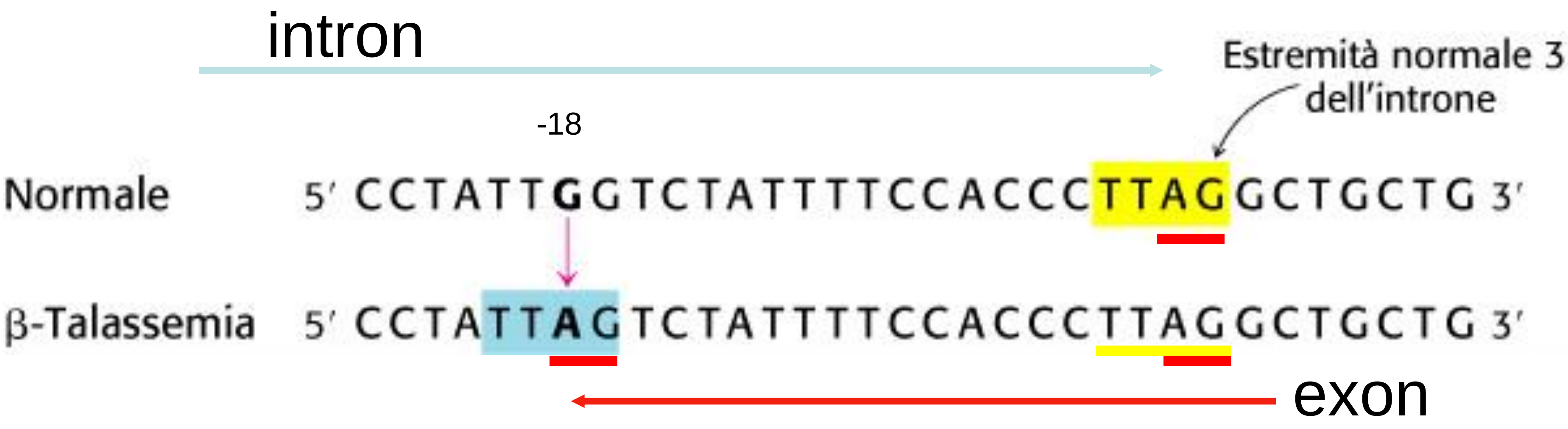
Proteins that originate from different genes but share some functional similarities (calcitonin - kidney and CGRP - intestine induce reduction of [Ca])

TRANSCRIPTION

pre-mRNA processing in mRNA

Mutations in introns can cause functional damage and disease; Beta-thalassemia, a genetic blood disorder characterized by reduced or absent production of beta-globin chains

G→A mutation in an intron of Hemoglobin



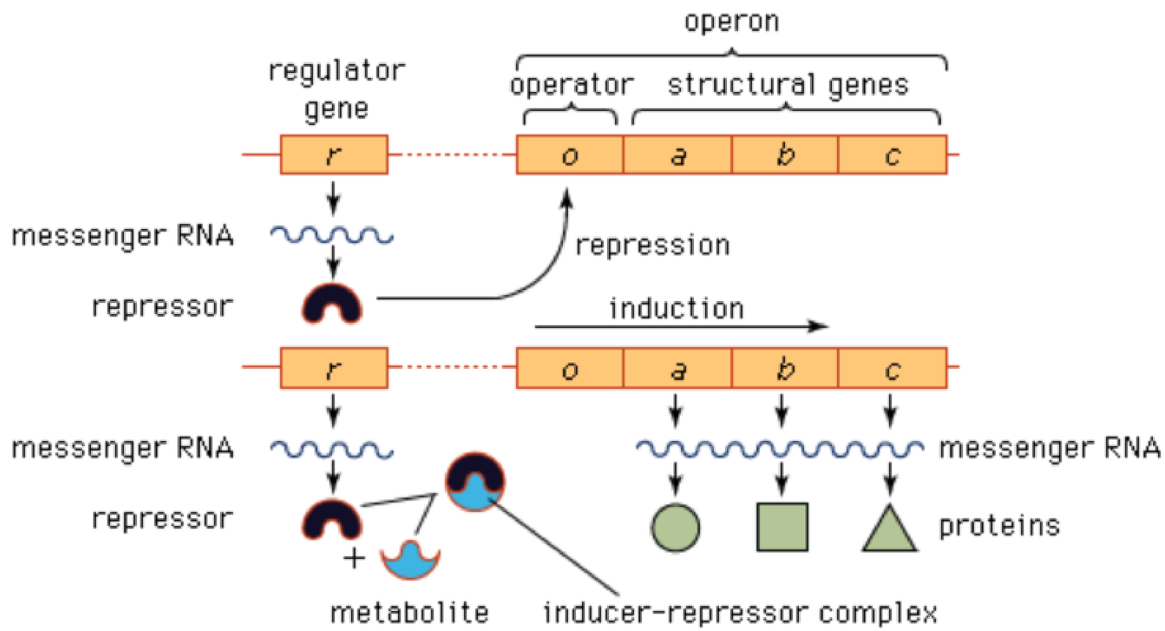
The exon downstream of the intron is longer, generating a pathological haemoglobin, beta thalassaemia.

TRANSCRIPTION

Gene expression regulation in eukaryotes

•Genes for enzymes of a metabolic pathway are clustered on the chromosome (**operon**), a phenomenon observed in both prokaryotes and eukaryotes.

A regulatory sequence adjacent to this unit determines how it should be transcribed: this sequence is called an 'operator'. Regulatory proteins act on operators to control gene transcription.

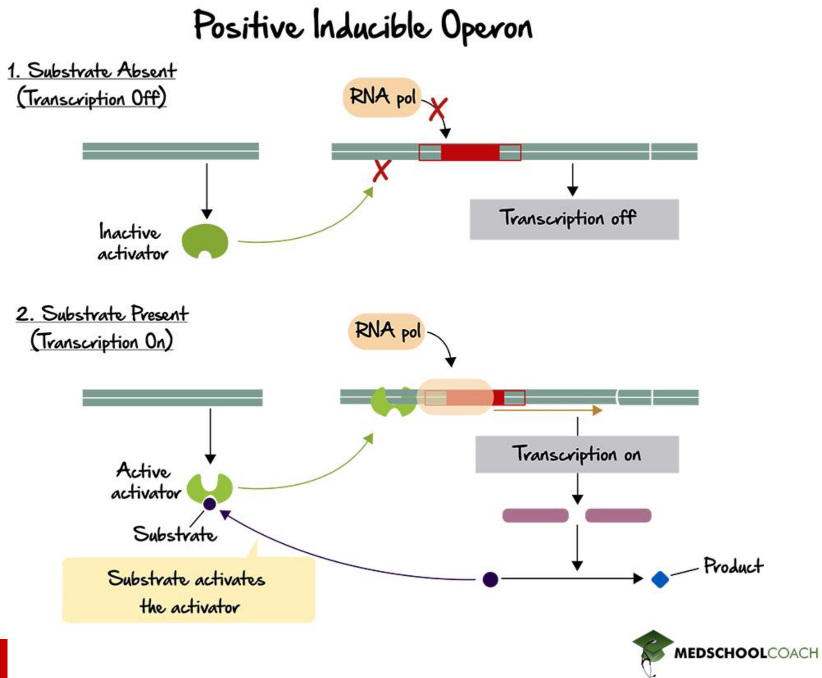


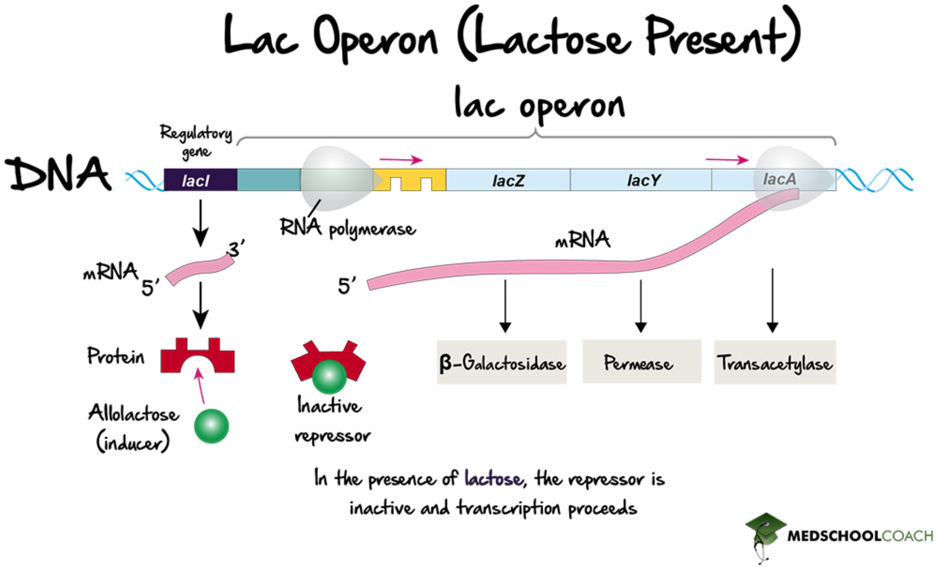
<u>Mutations on a regulatory gene affects all proteins encoded by the structural genes</u>

TRANSCRIPTION

Gene expression regulation in eukaryotes

- The increase in transcription of genes in response to a metabolite is termed ‘induction’ (inducers).
- Decreased transcription in response to a metabolite is termed ‘repression’ (repressor).
- *Free inducers* refer to small molecules that can regulate gene expression by interacting with specific proteins, such as repressors or activators.





CHROMATIN

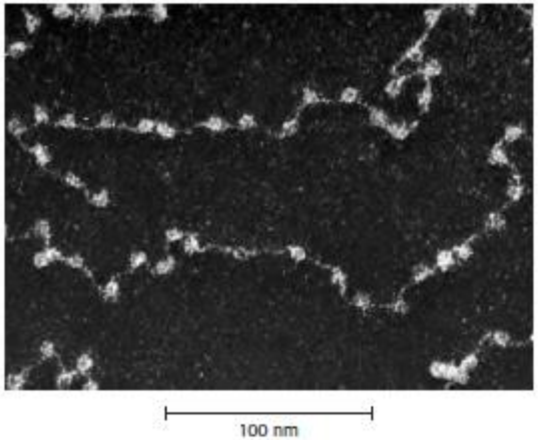


Figure 32.2 Chromatin structure. An electron micrograph of chromatin showing its "beads on a string" character. The beads correspond to DNA complexes with specific proteins. [Courtesy of Dr. Ada Olins and Dr. Donald Olins.]

Chromatin consists of **protein and DNA** in approximately equal amounts, plus a significant portion of associated **RNA**.

EUCHROMATIN: Lightly packed form of chromatin that is transcriptionally ACTIVE.
HETEROCHROMATIN: Tightly packed form of chromatin that is transcriptionally INACTIVE).

Chromatin DNA interacts very closely with proteins called **HISTONES**, which condense and sort the DNA into structural units called **NUCLEOSOMES**

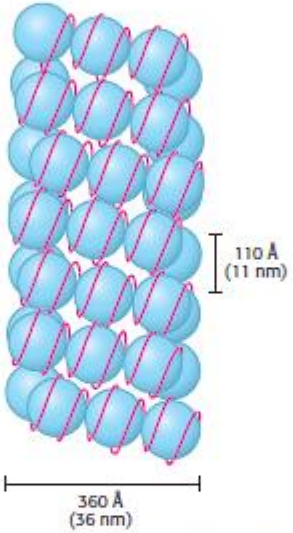
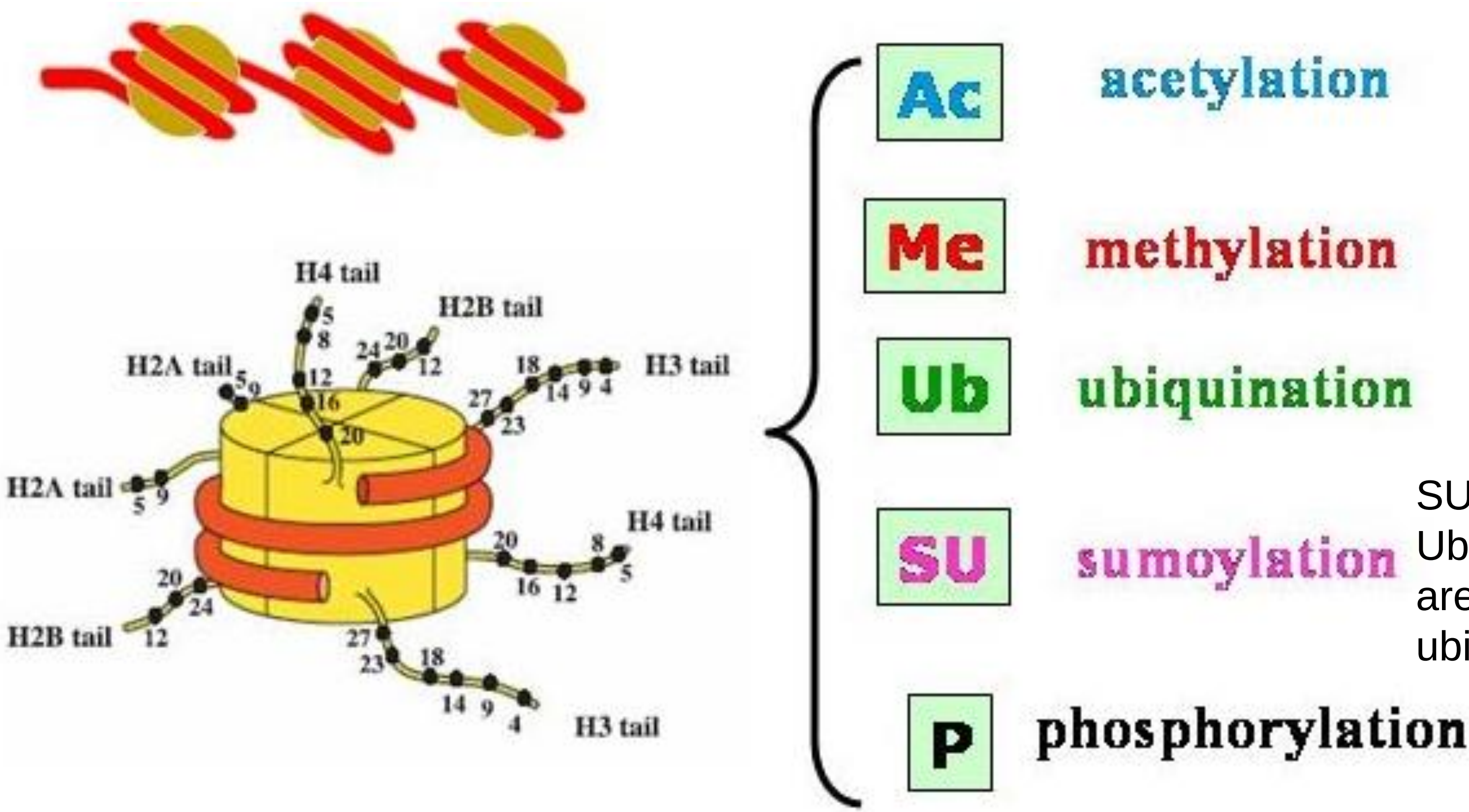


Figure 32.5 Higher-order chromatin structure. A proposed model for chromatin arranged in a helical array consisting of six nucleosomes per turn of helix. The DNA double helix (shown in red) is wound around each histone octamer (shown in blue). [After J. T. Finch and A. Klug. *Proc. Natl. Acad. Sci. U. S. A.* 73:1897–1901, 1976.]

Changes in the degree of chromatin packing occur during interphase in eukaryotic chromosomes. During gene activation, tightly packed chromatin must be converted to the open conformation to allow and facilitate the transcription process.

Regulation: histone modification

Many enzymes can induce chromatin remodelling by SHIFTING, REPLACING, MODIFYING histones (particularly the lysine-rich N-terminal tails

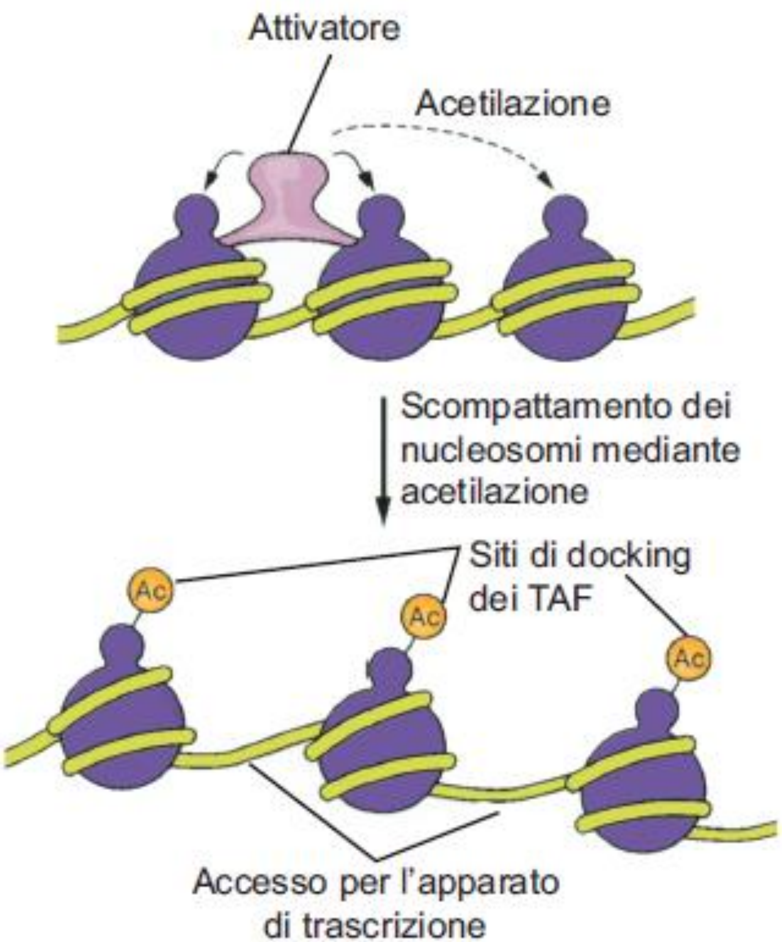


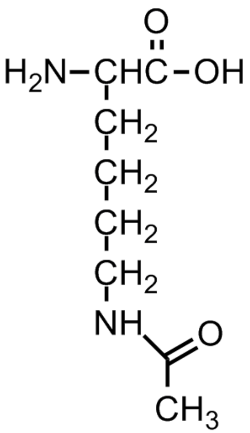
SUMOs (Small Ubiquitin-like Modifier) are small proteins like ubiquitin.

TRANSCRIPTION

Histone modifications: acetylation

Histone acetyl-transferase (HAT) → Acetylation of Lys residues



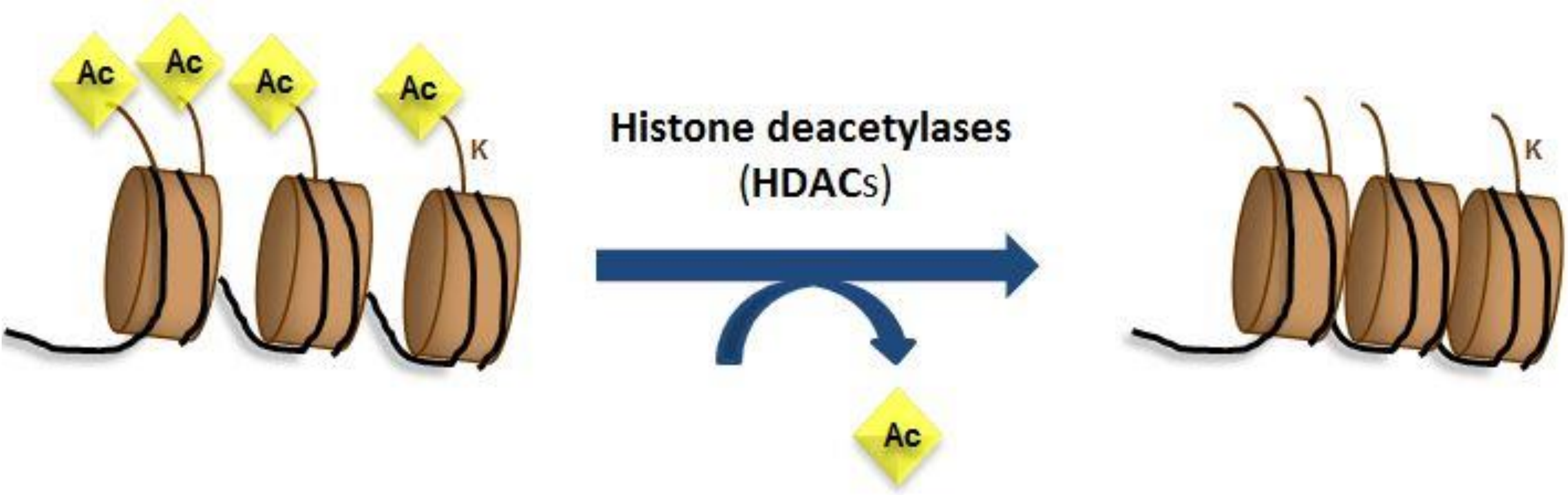


Docking sites of general transcription factors (TAFs)

↓

The gene sequence becomes more accessible

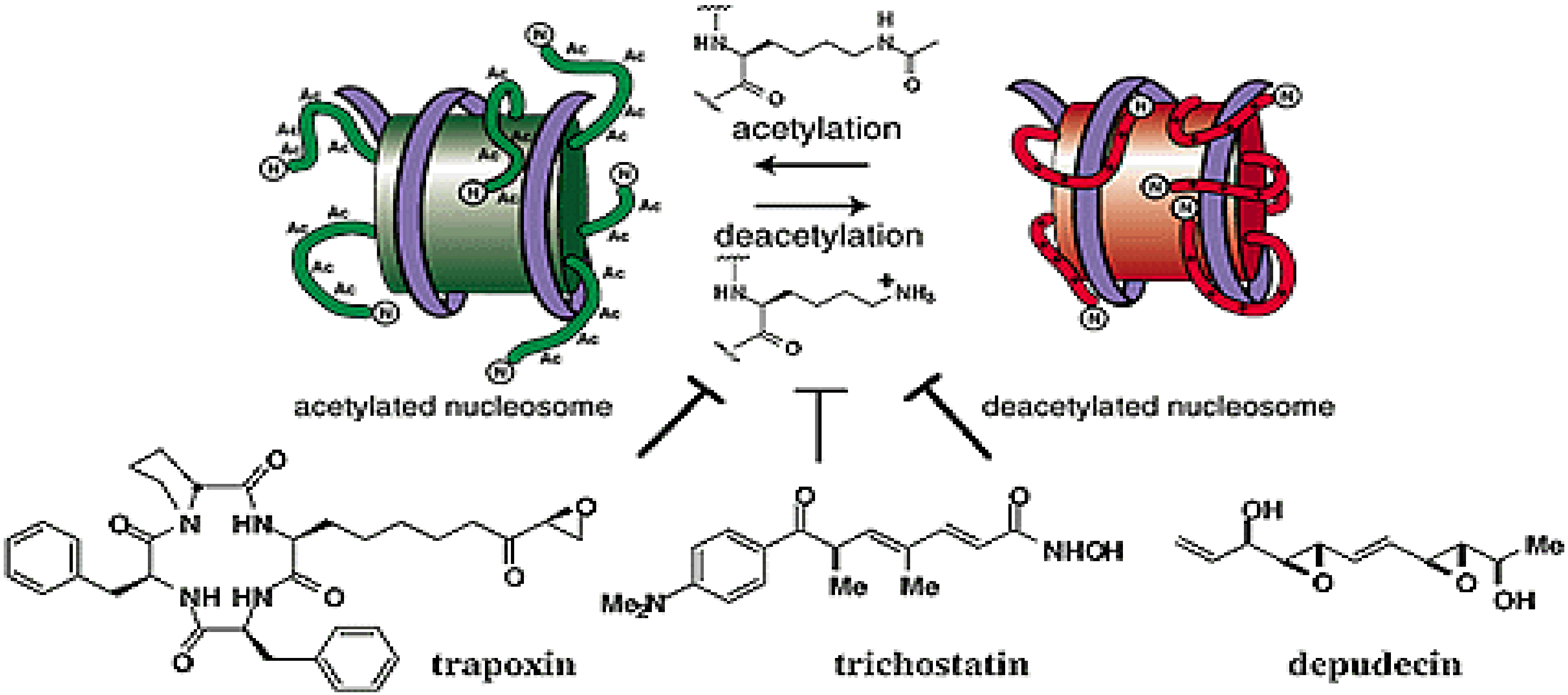
Histone modifications: deacetylation



In contrast, histone deacetylase enzymes (HDACs) promote histone deacetylation, hence gene silencing.

TRANSCRIPTION

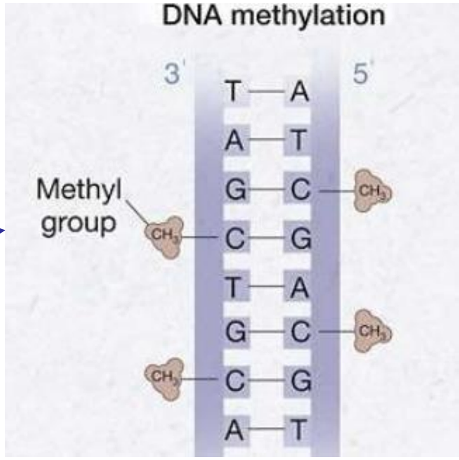
Histone modifications: deacetylation



HDAC inhibitors are being studied as anti-tumour agents (they can induce apoptosis)

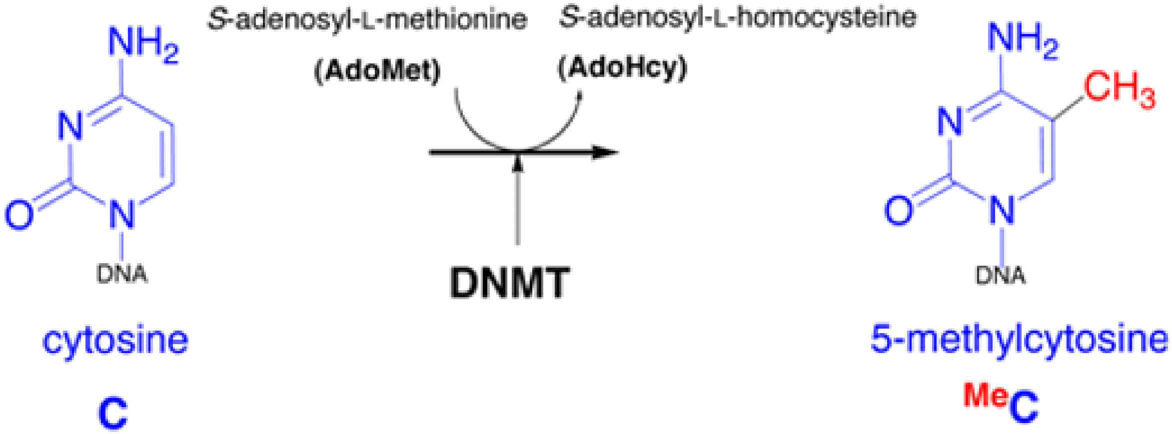
DNA methylation

DNA methylation is an essential epigenetic mechanism that involves the addition of a methyl group to the DNA molecule:



- **C5-METHYL-CYTHOSINE**
- **N4-METHYL-CYTHOSINE**
- **N6-METHYL-ADENINE**

On cytosines followed by guanines: **CpG ISLANDS**



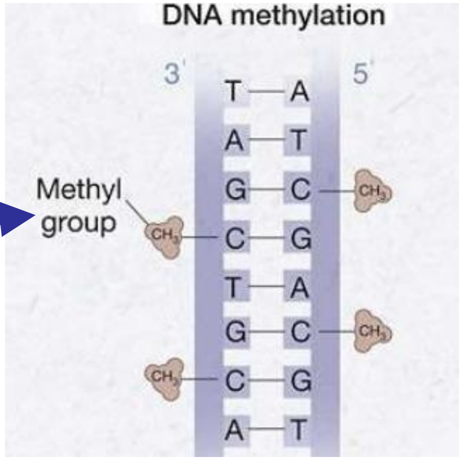
TRANSCRIPTION

DNA methylation

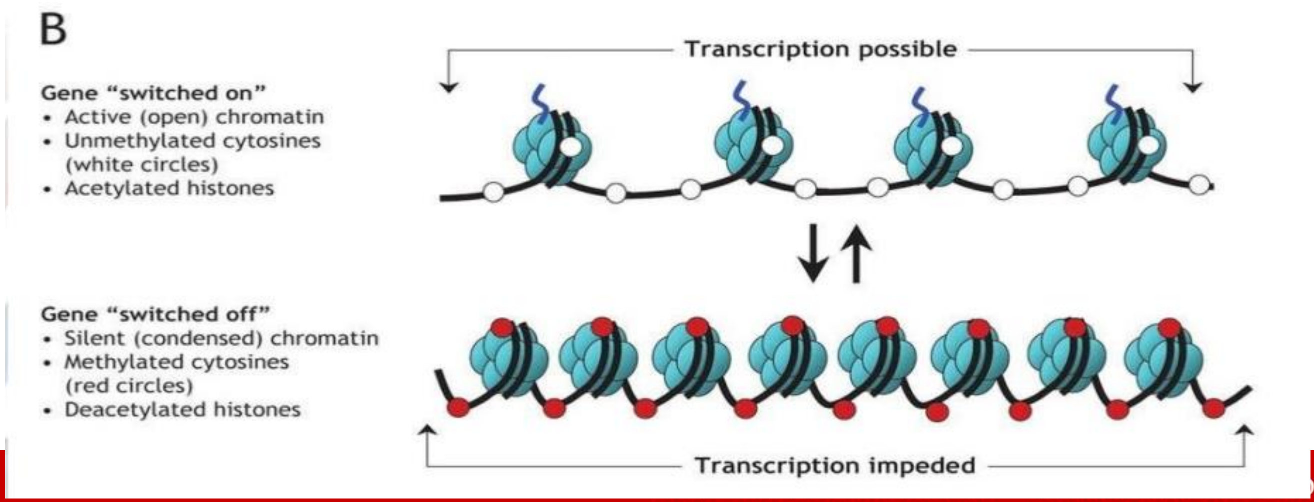
DNA methylation is an essential epigenetic mechanism that involves the addition of a methyl group to the DNA molecule:

- **C5-METHYL-CYTHOSINE**
- **N4-METHYL-CYTHOSINE**
- **N6-METHYL-ADENINE**

On cytosines followed by guanines: **CpG ISLANDS**



Methylation at the CpG level in the promoter region leads to GENETIC SILENCING



Kahoot!