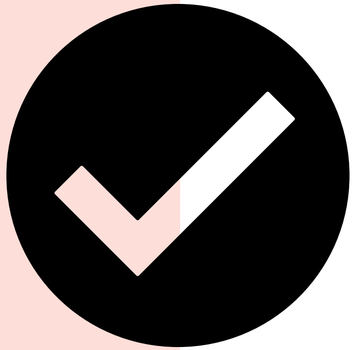
Cellular and Molecular Biology I

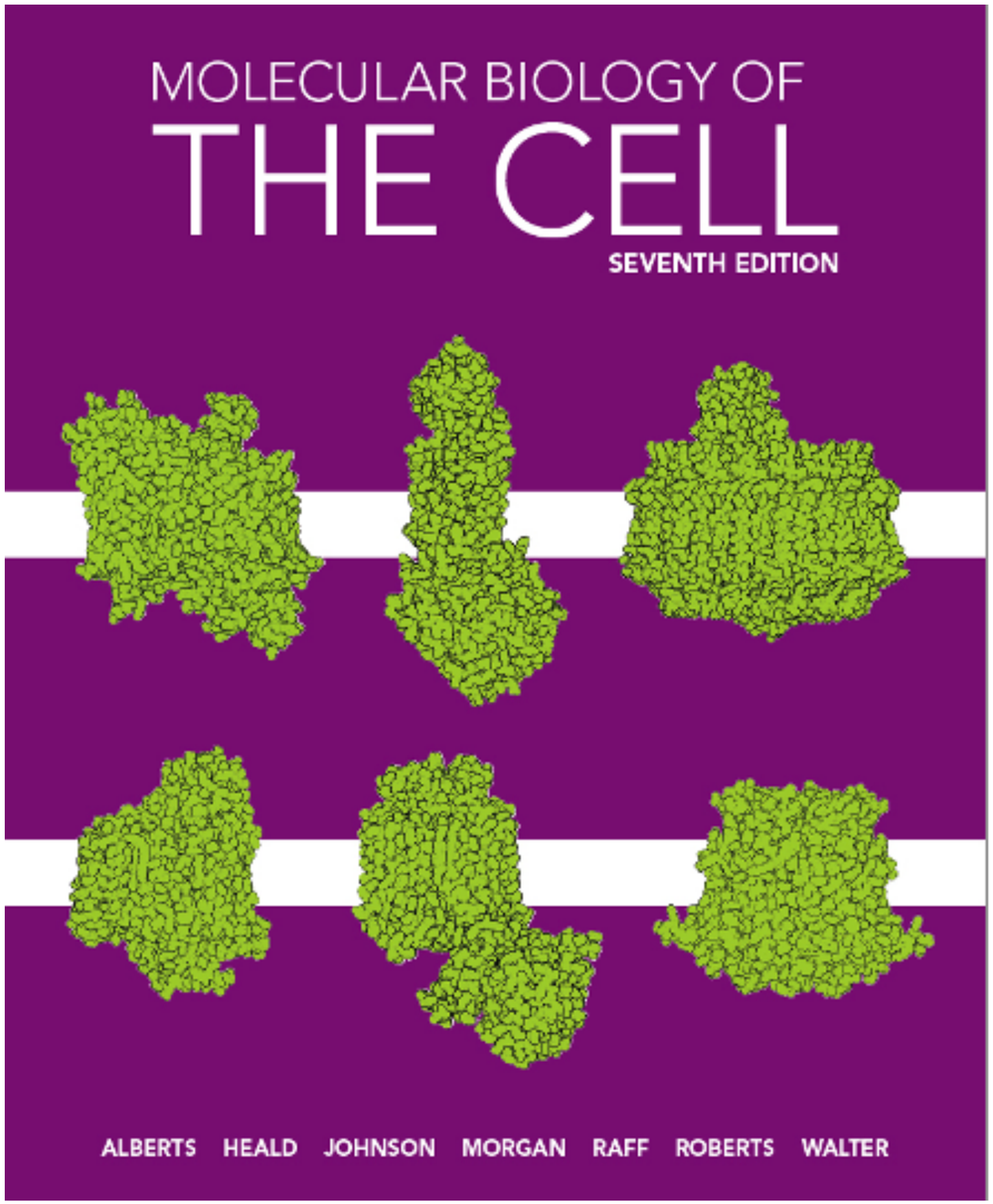
BIO-205-5

Camille Goemans

Quick recap

- Lecture 1 - Structure and packing of DNA
- Lecture 2 - Structure of chromosomes and how genomes evolve
- Lecture 3 - Mobile genetic elements and DNA replication
- Lecture 4 - DNA replication and DNA repair



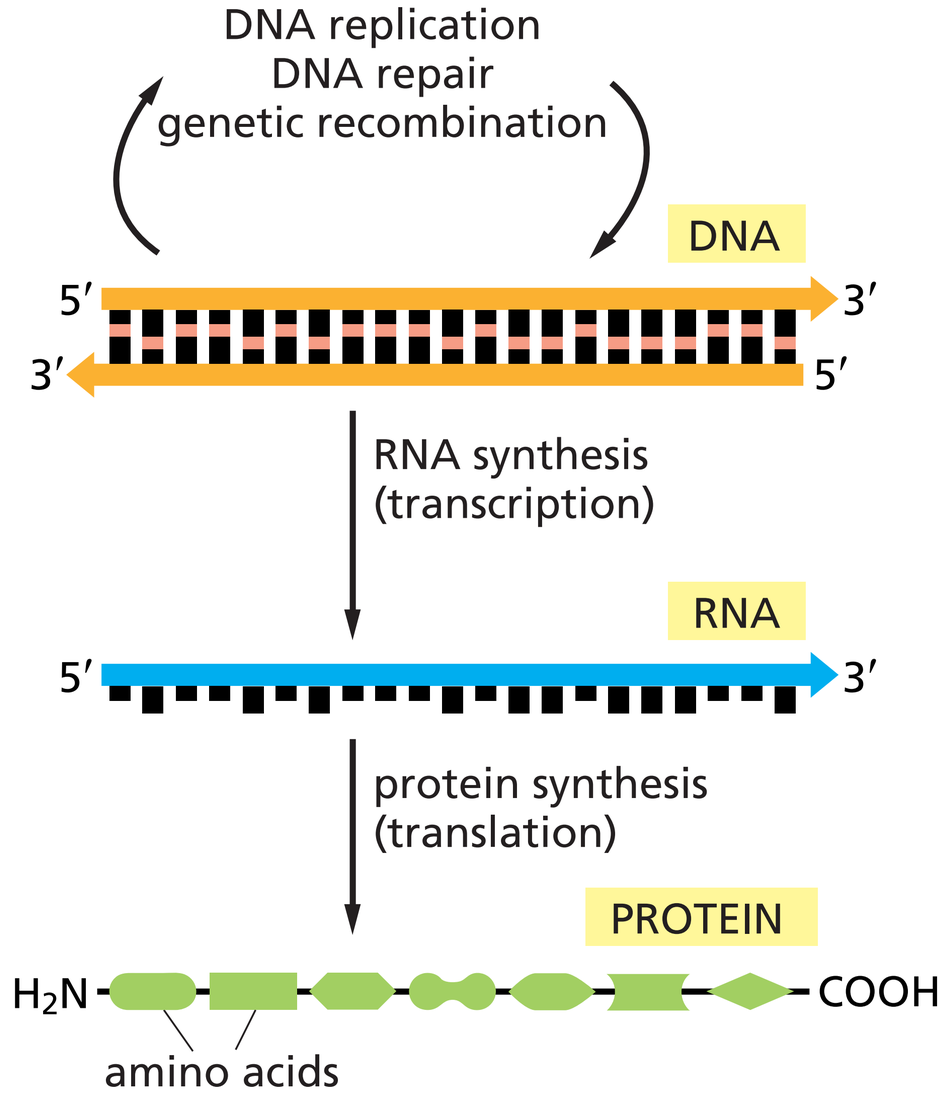


Chapter 6

How Cells Read the Genome: From DNA to Protein

Plan

- RNA
- Transcription
- Transcription initiation
- RNA processing
- Non-coding RNAs

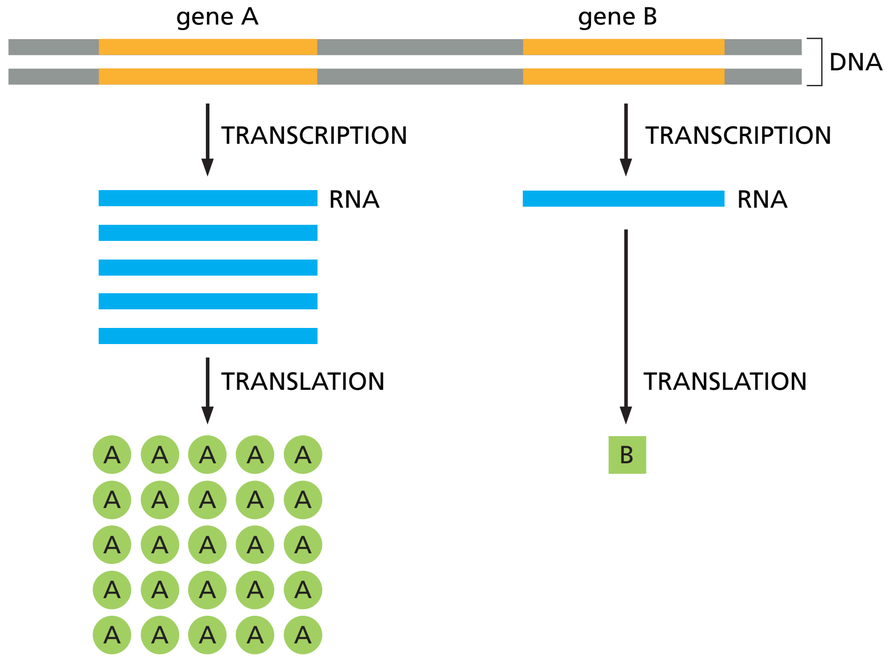


Central dogma of molecular biology

- **all cells** express their genetic information the same way (i.e. all follow the central dogma)
- important **variations** between organisms (e.g. RNA transcripts in eukaryotes are subject to a series of **processing** steps in the nucleus)
- some **RNAs are final products**: some with **3D structure** and catalytic function in the cell, other as regulators of gene expression (next chapter)
- roles of many **non-coding RNAs** are unknown

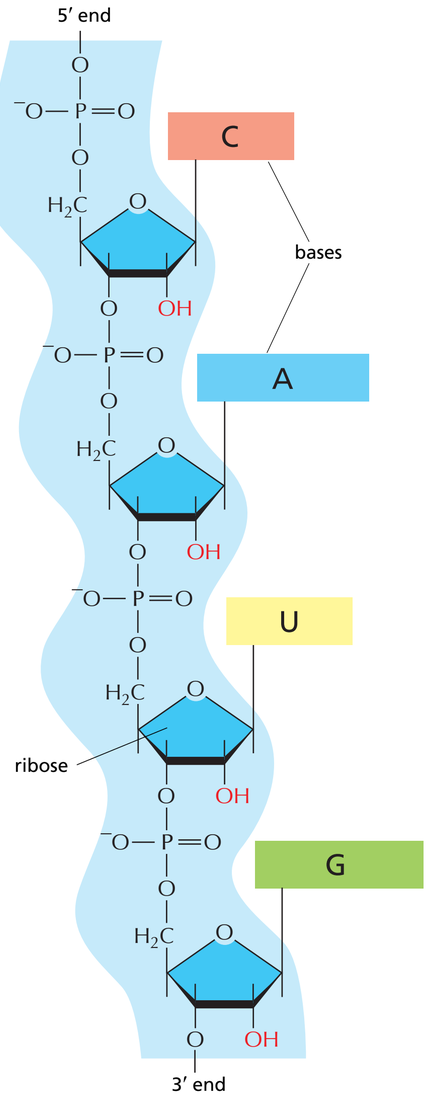
From DNA to RNA

- **Many identical RNA molecules** can be made from the same gene
- Genes can be transcribed and translated with **different efficiencies**
- Reminder: **genes** encode for **RNAs** as final products or **proteins**
- Cells can **regulate** the expression of each gene depending on **its needs**



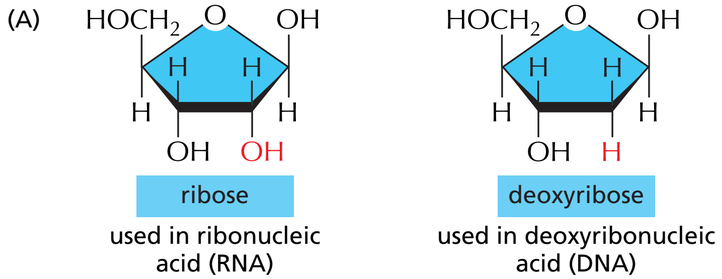
RNA molecules are single-stranded

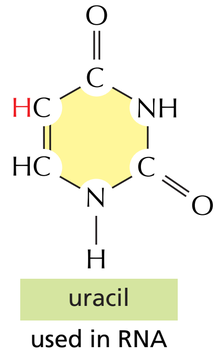
- The cell copies a portion of its DNA into an RNA nucleotide sequence = **transcription**
- RNA is a linear polymer made of **4 nucleotides** linked by **phosphodiester bonds**

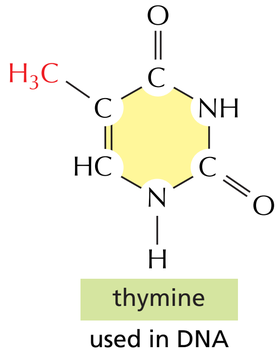


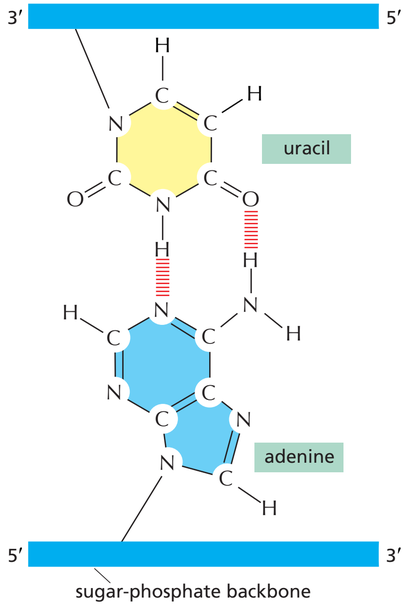
RNA molecules are single-stranded

- **RNA** is **different than DNA** in 2 aspects:
  - nucleotides are **ribonucleotides** rather than deoxyribonucleotides
  - contains **uracil** instead of thymine base (which also forms H-bonds with adenine)



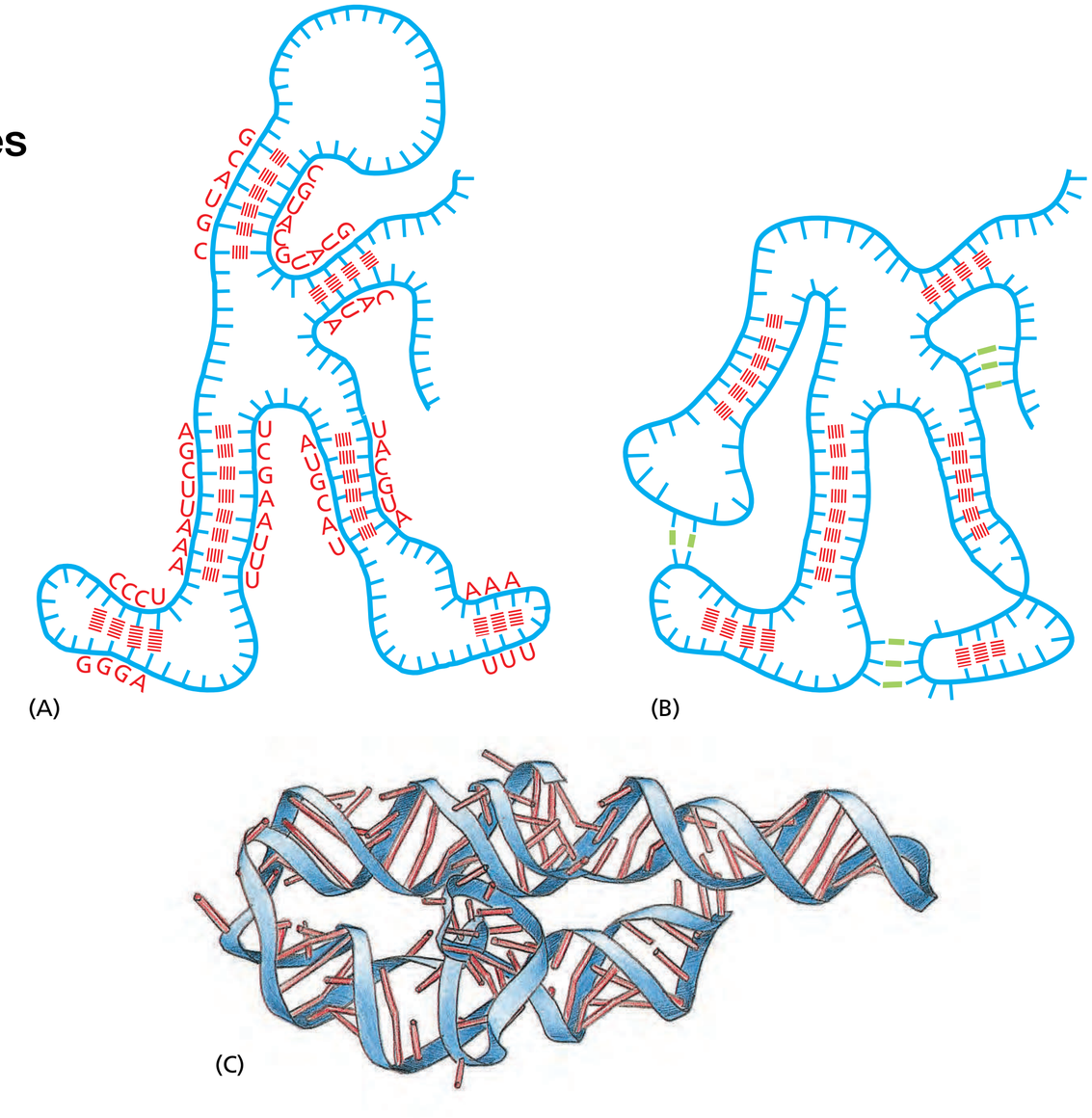






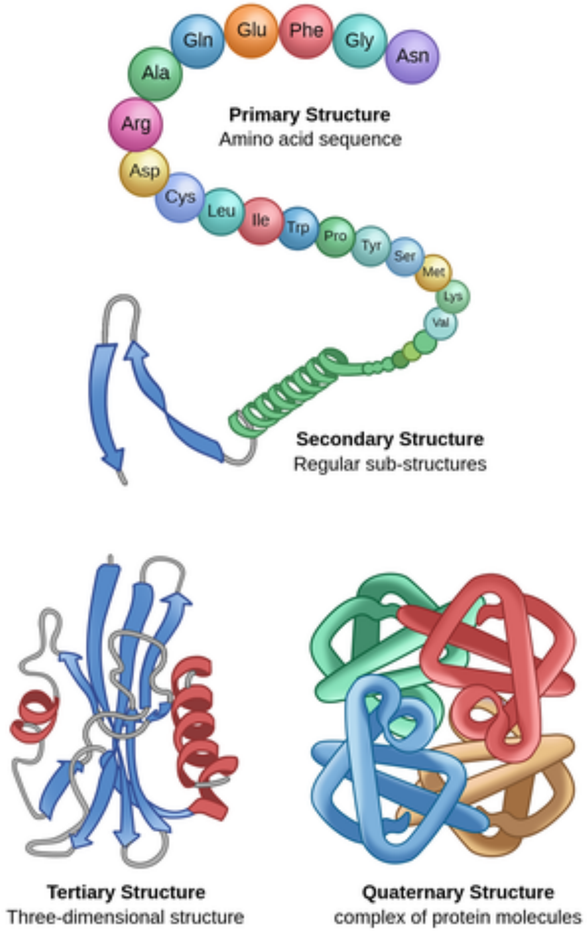
RNA molecules are single-stranded

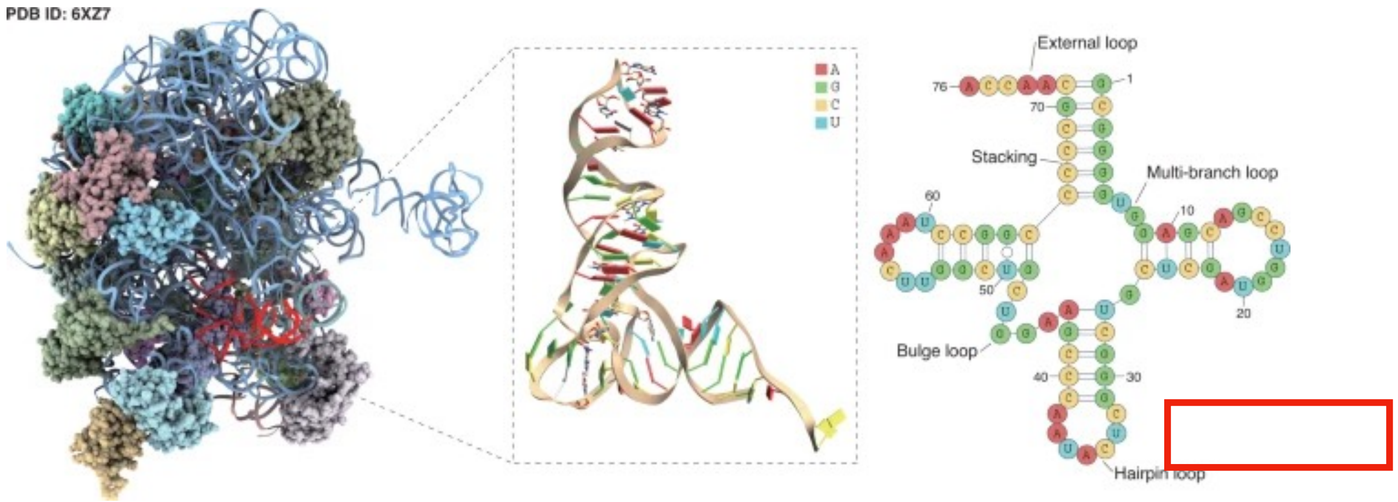
- RNA is single-stranded and **folds into particular shapes**



3D structures of proteins vs. RNA

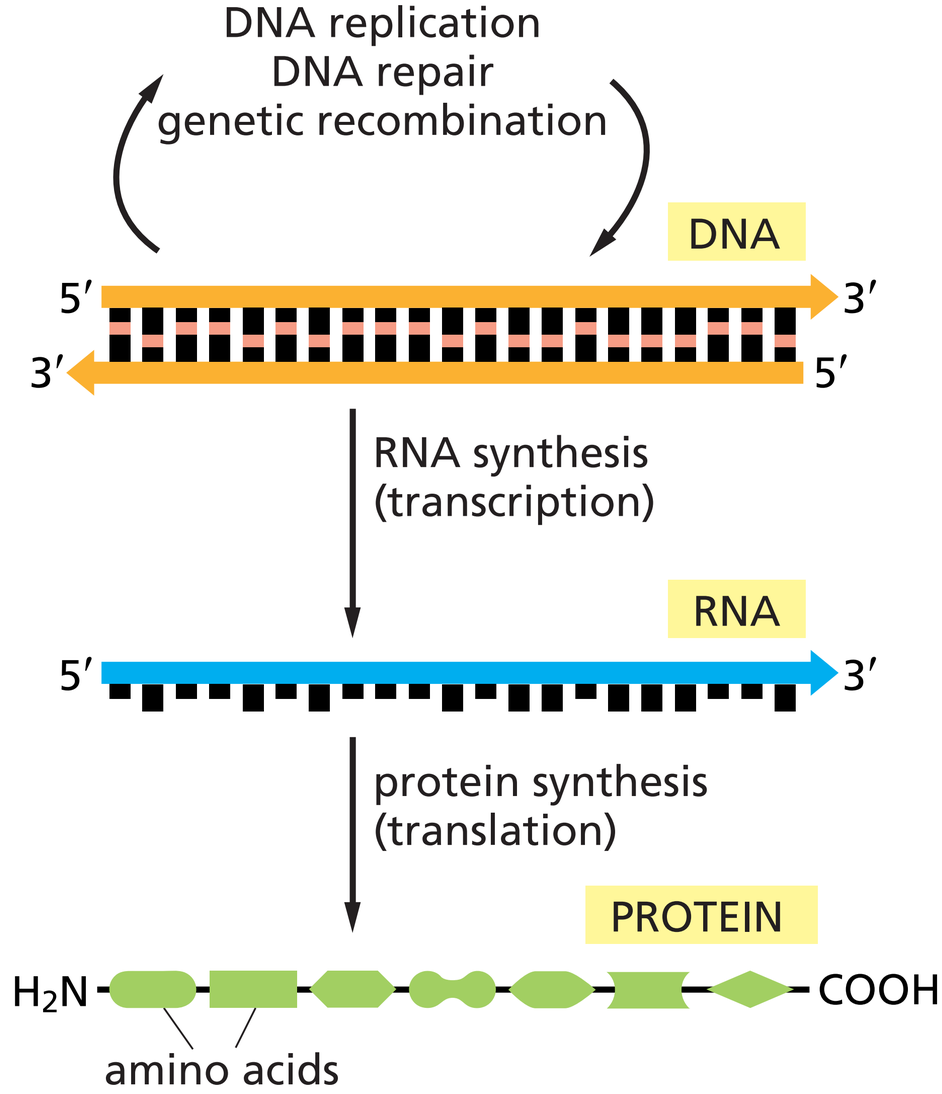
- 3D structures of **protein vs. RNA**





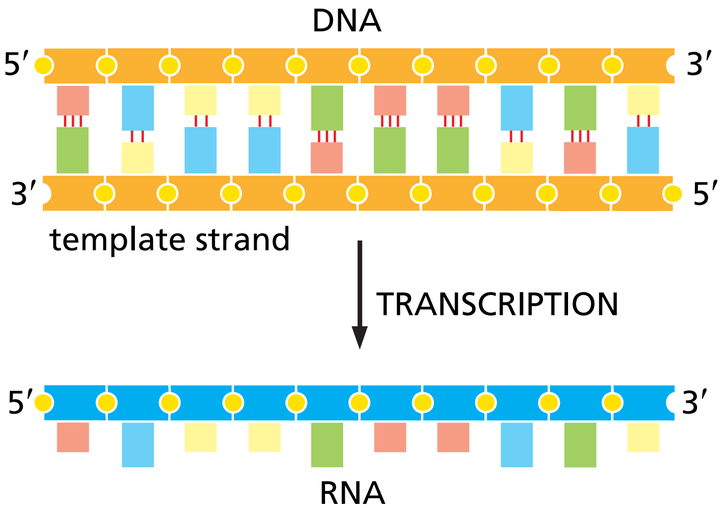
Plan

- RNA
- Transcription
- Transcription initiation
- RNA processing
- Non-coding RNAs

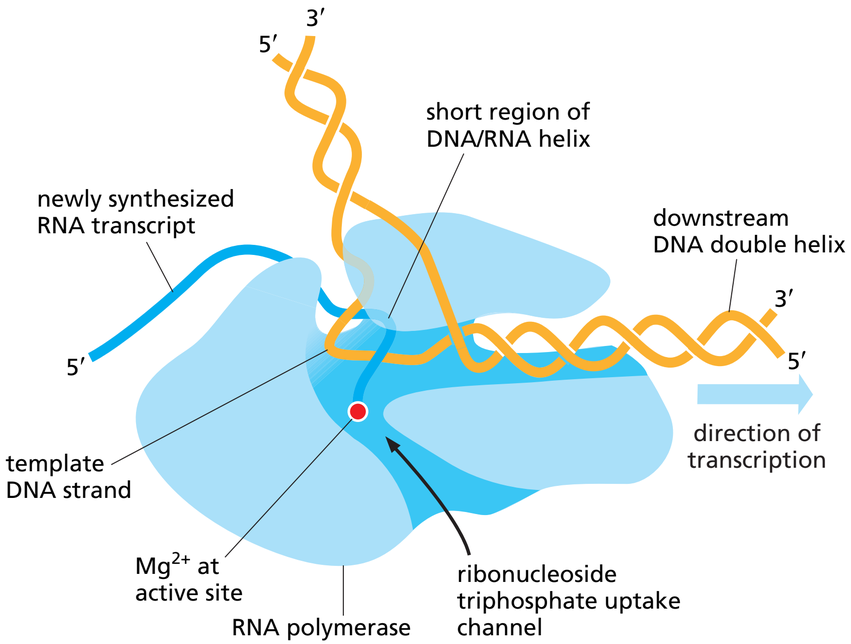


Transcription

- RNA is **complementary** to one strand of DNA
- Starts with the **opening and unwinding of DNA** to expose the bases of each strand
- One of the strands then acts as a **template** for synthesis of an RNA molecule (complementary base-pairing)
- When a good match is made, the **ribonucleotide is covalently** linked to the growing RNA = transcript



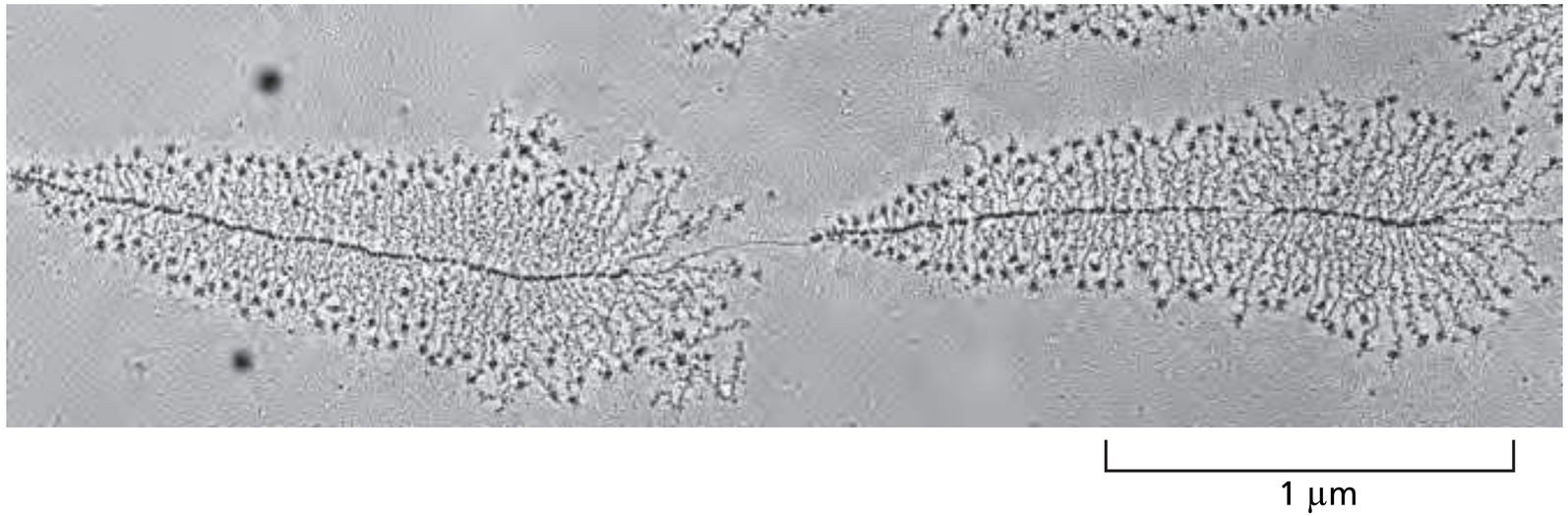
Transcription



- The growing RNA is extended one nucleotide at a time **from 5’ to 3’**
- The substrates are **ribonucleoside triphospates** (ATP, CTP, UTP, GTP)

Transcription

Example: ribosomal RNAs



- **Many RNA copies** from the same gene can be made in a short time
- New molecules can be made **before the previous one is completed**

RNA polymerase vs. DNA polymerase

- **RNA polymerase** catalyzes essentially the same chemical reaction as the **DNA polymerase**
- Differences between the enzymes:
  - links **ribonucleotides** and not **deoxyribonucleotides**
  - can start an RNA chain without a **primer** (less accurate)
  - makes 1 **mistake** every 10^4 nucleotide (as compared to 10^7 for DNA polymerase)
  - the RNA polymerase that begins an RNA molecule finishes it **without dissociating** from the DNA
- **RNA polymerase** have a modest proofreading mechanism: if an incorrect ribonucleotide is added, the active site can perform an excision reaction and the nucleoside is released
- DNA- and RNA-polymerase are quite **different structurally,** such enzymes have arisen twice during early evolution of the cells

Cells produce different categories of RNAs

TABLE 6–1 Principal Types of RNAs Produced in Cells

Type of RNA	Function
mRNAs	Messenger RNAs, code for proteins
rRNAs	Ribosomal RNAs, form the basic structure of the ribosome and catalyze protein synthesis
tRNAs	Transfer RNAs, central to protein synthesis as adaptors between mRNA and amino acids
snRNAs	Small nuclear RNAs, function in a variety of nuclear processes, including the splicing of pre-mRNA
snoRNAs	Small nucleolar RNAs, help to process and chemically modify rRNAs
miRNAs	MicroRNAs, regulate gene expression by blocking translation of specific mRNAs and cause their degradation
siRNAs	Small interfering RNAs, turn off gene expression by directing the degradation of selective mRNAs and the establishment of compact chromatin structures
piRNAs	Piwi-interacting RNAs, bind to piwi proteins and protect the germ line from transposable elements
lncRNAs	Long noncoding RNAs, many of which serve as scaffolds; they regulate diverse cell processes, including X-chromosome inactivation

Cells produce different categories of RNAs

TABLE 6–1 Principal Types of RNAs Produced in Cells	
Type of RNA	Function
mRNAs	Messenger RNAs, code for proteins

Cells produce different categories of RNAs

Non-coding RNAs

rRNAs	Ribosomal RNAs, form the basic structure of the ribosome and catalyze protein synthesis
tRNAs	Transfer RNAs, central to protein synthesis as adaptors between mRNA and amino acids
snRNAs	Small nuclear RNAs, function in a variety of nuclear processes, including the splicing of pre-mRNA
snoRNAs	Small nucleolar RNAs, help to process and chemically modify rRNAs
miRNAs	MicroRNAs, regulate gene expression by blocking translation of specific mRNAs and cause their degradation
siRNAs	Small interfering RNAs, turn off gene expression by directing the degradation of selective mRNAs and the establishment of compact chromatin structures
piRNAs	Piwi-interacting RNAs, bind to piwi proteins and protect the germ line from transposable elements
lncRNAs	Long noncoding RNAs, many of which serve as scaffolds; they regulate diverse cell processes, including X-chromosome inactivation

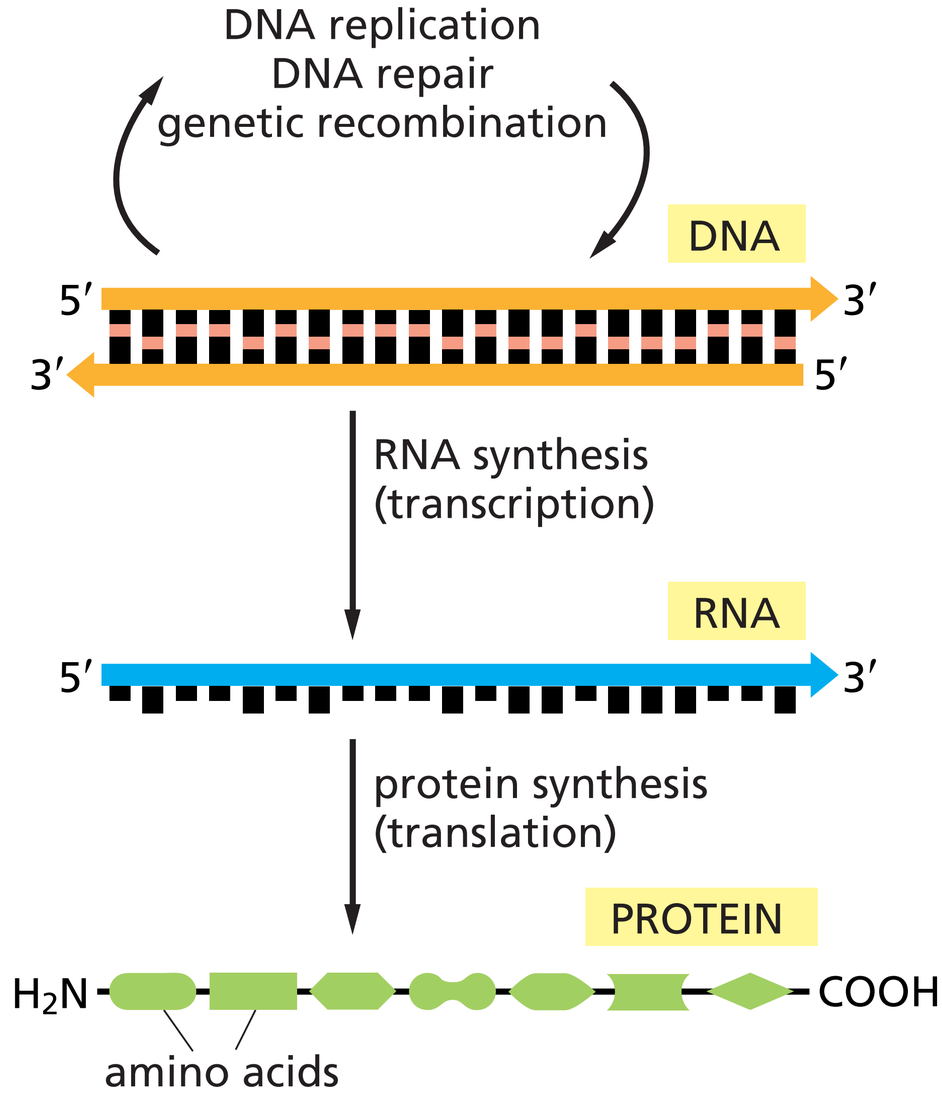
- >10.000 in humans
- enzymatic, structural or regulatory function
- many with mysterious function

Transcription unit

- Each transcribed segment of DNA is a **transcription unit**
- In Eukaryotes, it typically carries the information of **one gene**
- In bacteria, a set of **adjacent genes** is often transcribed together, the mRNA carries the information for several distinct proteins

Plan

- RNA
- Transcription
- Transcription initiation
- RNA processing
- Non-coding RNAs



Where to start?

- The RNA polymerase has to know where to start and where to end
- In **bacteria**:
    - A **sigma factor** in bacteria is a type of protein that plays a key role in initiating transcription.

Here's how it works:

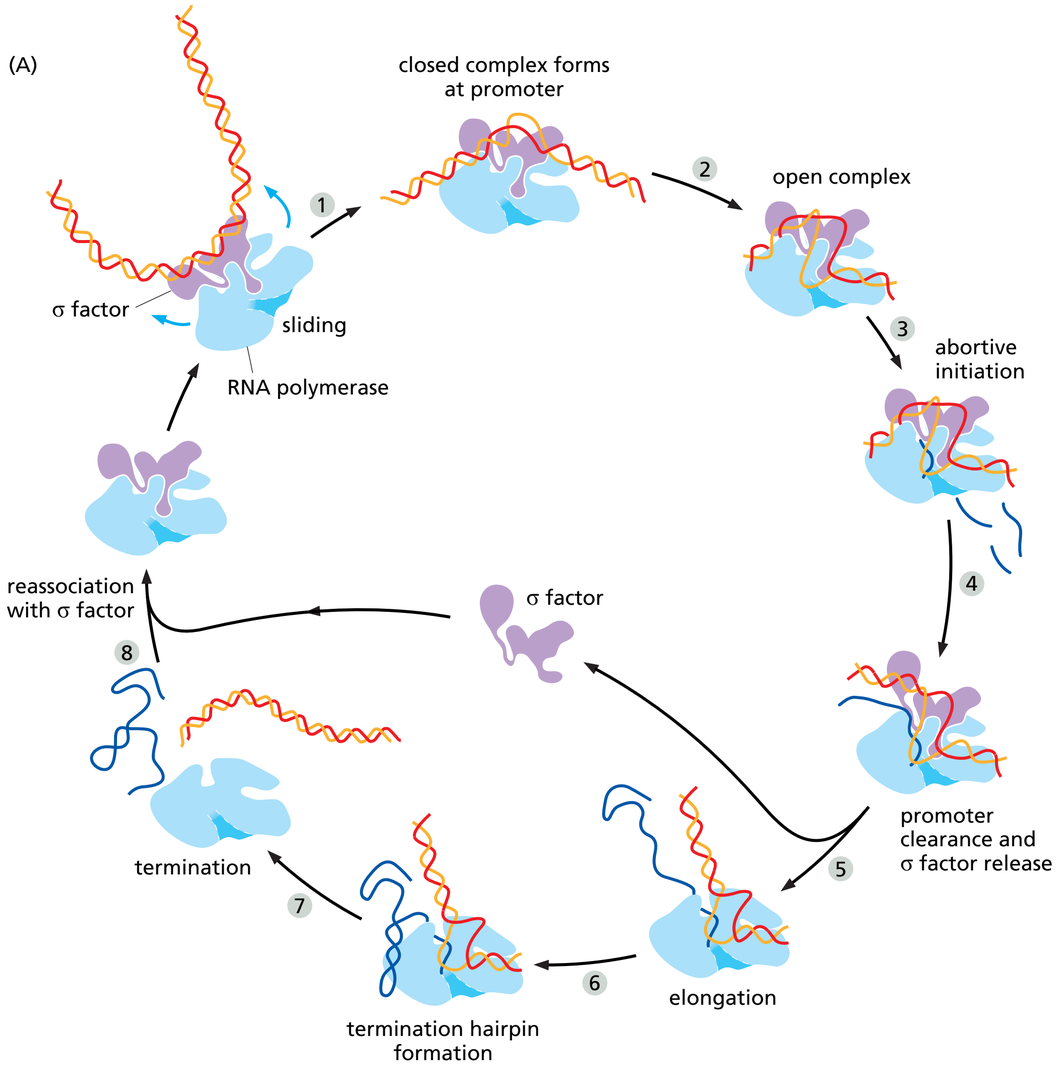
- The bacterial RNA polymerase enzyme cannot on its own recognize **where to start** transcribing a gene.
- A sigma factor binds to the RNA polymerase core enzyme and helps it specifically **recognize promoter sequences** (short stretches of DNA at the beginning of genes).
- Once transcription starts, the sigma factor often dissociates, leaving the RNA polymerase to continue elongating the RNA chain.

Types of sigma factors:

- σ^{70} (sigma-70): The "housekeeping" sigma factor in E. coli; responsible for most gene expression under normal growth conditions.
- Alternative sigma factors: Turn on in response to specific stresses or environmental conditions, such as:
- σ^{32} (sigma-32): Heat shock response.
- σ^{54} (sigma-54): Nitrogen limitation.
- σ^S (sigma-S / σ^{38}): Stationary phase and stress response.

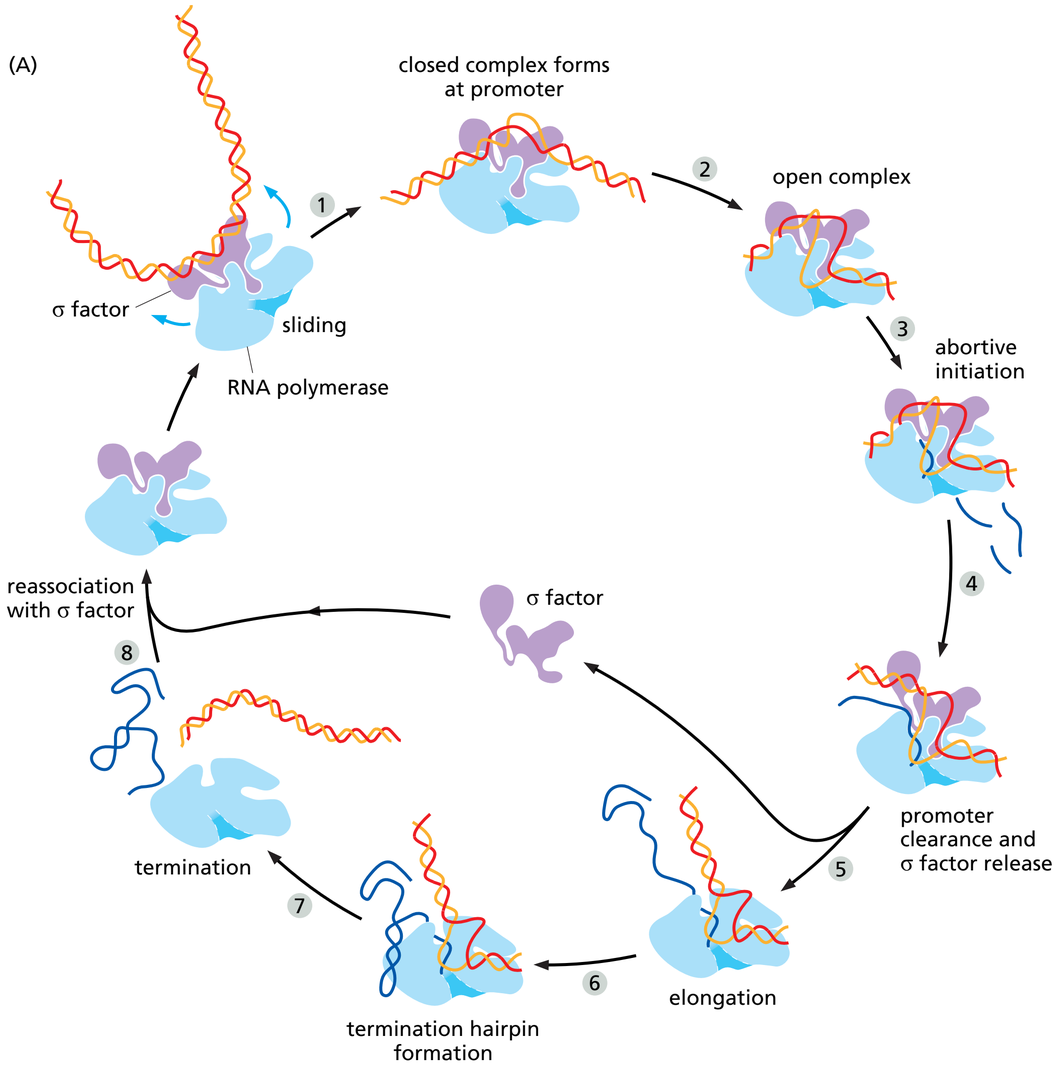
Where to start?

- The RNA polymerase has to know where to start and where to end
- In **bacteria**:
  - **sigma factor** (additional subunit of the polymerase)
  - when they find a **promoter** (=starting point of RNA synthesis), the RNA polymerase binds tightly
  - opens up the double helix =transcription bubble
  - sigma factor stabilises one of the strand and the other serves as template
  - there are **different** sigma factors that are each used in particular environmental situations (e.g. heat stress, acid stress,…) and control the expression of the genes needed to deal with the given stress



Where to start?

- In **bacteria**:
  - **abortive initiation** when the RNA polymerase remains bound to the promotor, which creates a stress
  - RNA polymerase **releases** itself from the promoter and discards the sigma factor
  - RNA polymerase **moves down along DNA** and synthesizes RNA
  - RNA polymerase encounters a **terminator**, it releases the RNA molecule and the DNA template
  - In bacteria, a **terminator** is a string of A-T pairs preceded by a twofold symmetric DNA sequence which forms a hairpin when transcribed into RNA



Where to start?

- In **bacteria**:
  - **Abortive initiation** in bacteria is a phenomenon that happens at the very start of transcription, when RNA polymerase is just beginning to make RNA.

    1. Closed complex formation: RNA polymerase (with a sigma factor) binds to the promoter DNA
    2. Open complex formation: The DNA strands near the promoter are melted (opened) so transcription can begin.
    3. Initial transcription: RNA polymerase starts linking together the first few ribonucleotides.
    4. Abortive initiation: Instead of moving forward into productive transcription, the enzyme frequently releases these short RNA fragments (usually 2–9 nucleotides long) and tries again.

    ➡ This cycle of repeatedly making and releasing short transcripts is called abortive initiation.

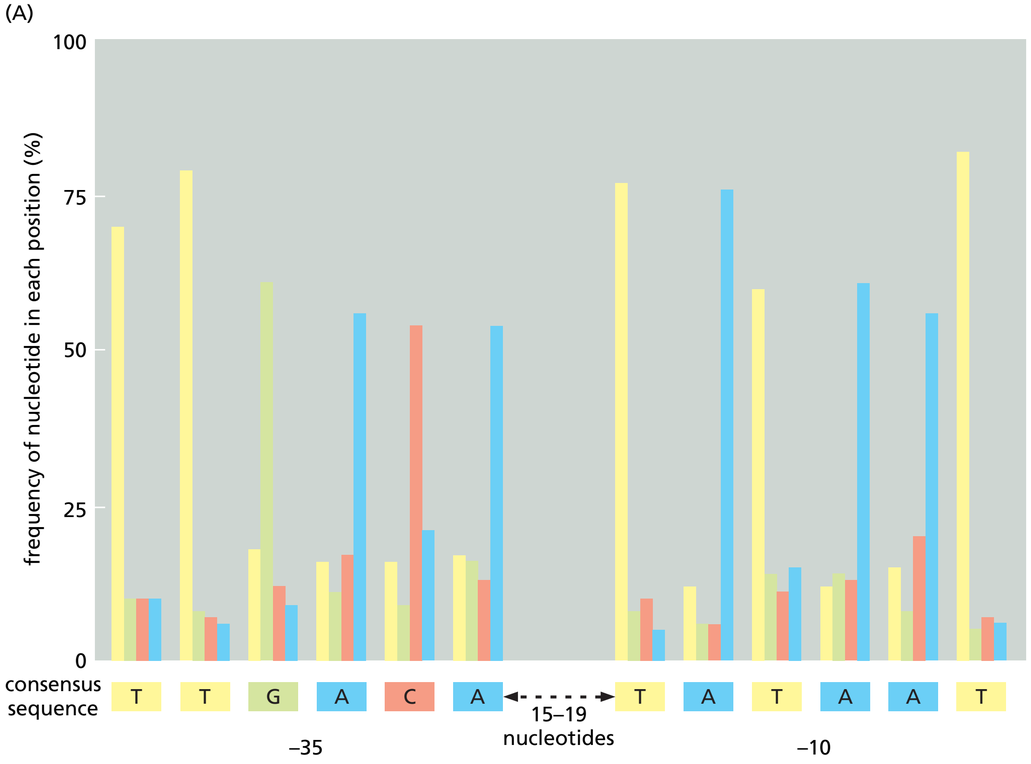
    ➡ Promoter escape: Eventually, RNA polymerase clears the promoter, the sigma factor often loosens or falls off, and the enzyme transitions into the elongation phase, making full-length RNA.

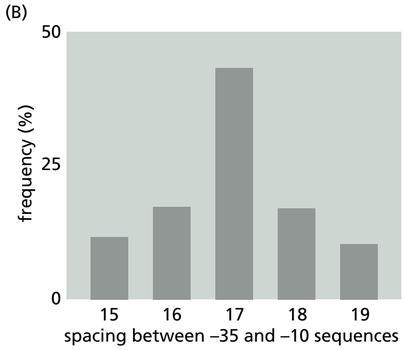
Why does abortive initiation happen?

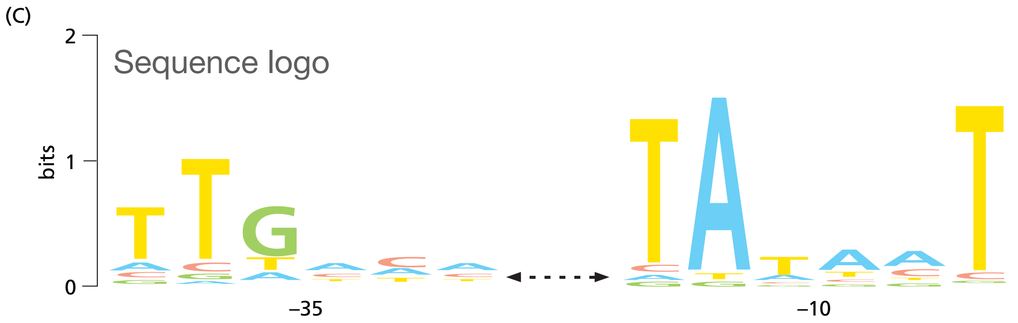
- It allows RNA polymerase to “test” its ability to escape the promoter.
- It may help ensure transcription starts only when conditions are correct.
- Mechanistically, it occurs because the RNA polymerase is initially tethered at the promoter and must break contacts with promoter DNA and sigma factor to move forward.

Where to start?

- **Initiation and termination signals** are sometimes difficult to recognise
- The bacterial promoters are made of **-35 and -10 sequences**
- Comparison of many **bacterial promoters** shows a high degree of variation although they are recognised by the same sigma factor (-10 and -35 sequences)
- **Consensus sequence** = common features

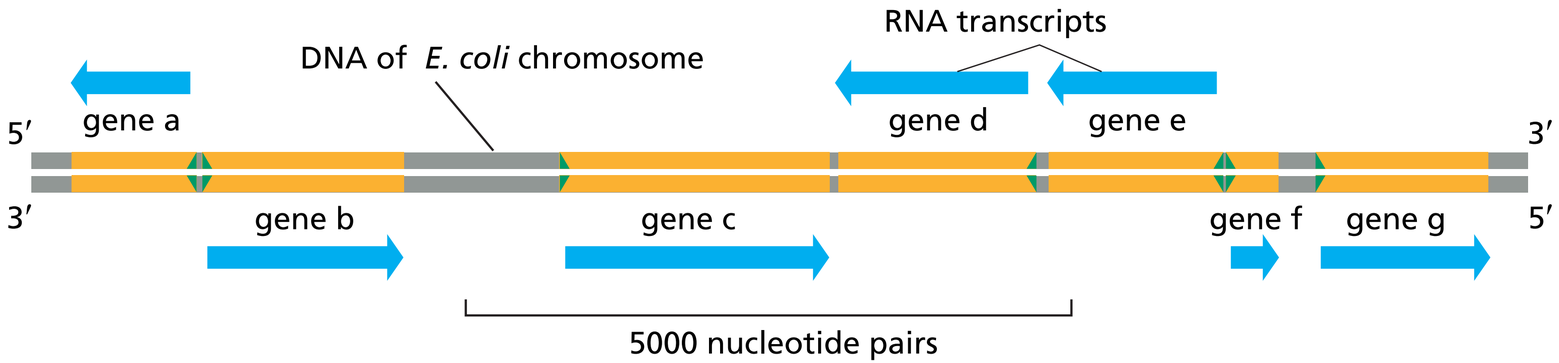






Where to start?

- **Promoter sequences are asymmetric** so the RNA polymerase can only bind in one direction
- The RNA polymerase only synthesises RNA in the **5’ to 3’ direction**
- The promoter specifies which **strand** to take as template



And in eukaryotes?

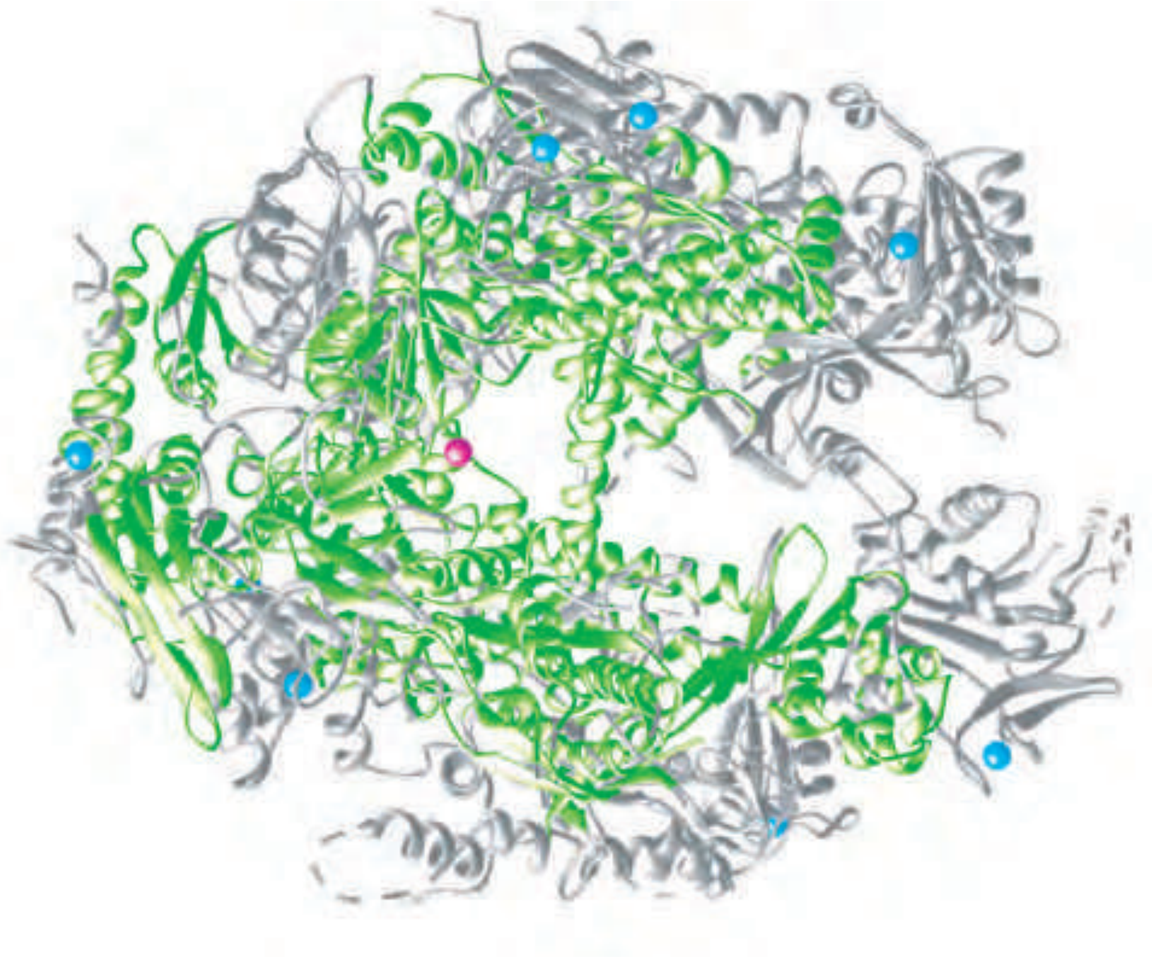
- **Three types** of RNA polymerase (I, II and III)
- **Structurally** similar but transcribe different **categories of genes**

TABLE 6–2 The Three RNA Polymerases in Eukaryotic Cells

Type of polymerase	Genes transcribed
RNA polymerase I	5.8S, 18S, and 28S rRNA genes
RNA polymerase II	All protein-coding genes, plus snoRNA genes, miRNA genes, siRNA genes, lncRNA genes, and most snRNA genes
RNA polymerase III	tRNA genes, 5S rRNA genes, some snRNA genes, and genes for other small RNAs

The rRNAs were named according to their "S" values, which refer to their rate of sedimentation in an ultracentrifuge. The larger the S value, the larger the rRNA.

And in eukaryotes?



- **Structurally** similar to bacterial RNA polymerase
  - BUT requires many initiation factors (not just a sigma factor) = **general transcription factors**
  - BUT transcription initiation takes place on DNA that is **packaged into nucleosomes and chromatin**

General transcription factors

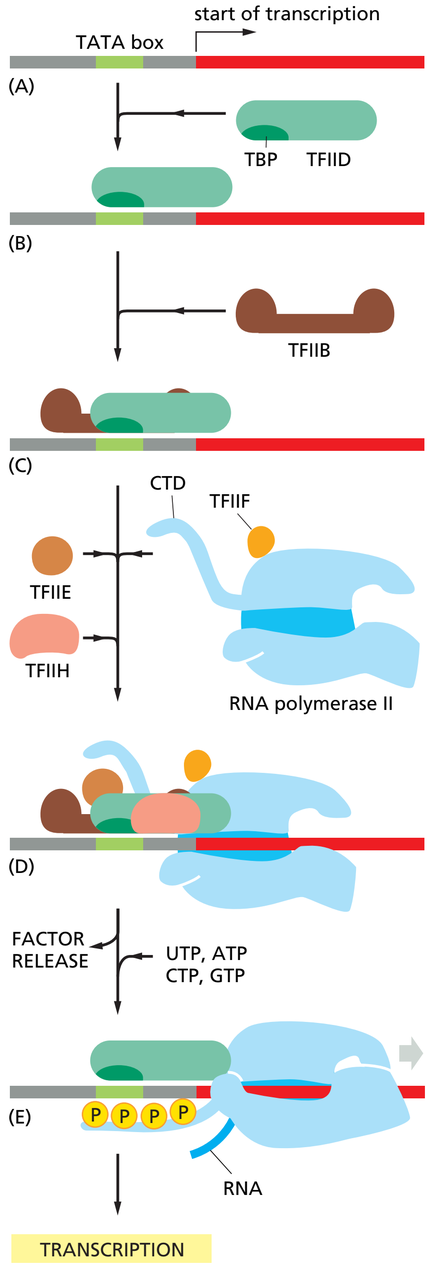
- Help **position** eukaryotic RNA polymerase correctly at the promoter
- Aid in **pulling the two strands** of DNA apart
- **Release** the RNA polymerase from the promoter to start its elongation mode
- Are needed at **nearly all promoters** used by RNA polymerase II
- Denoted TFIIA, B, C, … (for transcription factor for polymerase II)

Transcription initiation

TABLE 6–3 The General Transcription Factors Needed for Transcription Initiation by Eukaryotic RNA Polymerase II

Name	Number of subunits	Roles in transition initiation
TFIID TBP subunit TAF subunits	 1 ~11	 Recognizes TATA box Recognizes other DNA sequences near the transcription start point; regulates DNA-binding by TBP
TFIIB	1	Recognizes BRE element in promoters; accurately positions RNA polymerase at the start site of transcription
TFIIF	3	Stabilizes RNA polymerase interaction with TBP and TFIIB; helps attract TFIIE and TFIIH
TFIIE	2	Attracts and regulates TFIIH
TFIIH	9	Unwinds DNA at the transcription start point, phosphorylates Ser5 of the RNA polymerase CTD; releases RNA polymerase from the promoter

TFIID is composed of TBP and ~11 additional subunits called TAFs (TBP-associated factors); CTD, C-terminal domain.



Transcription initiation

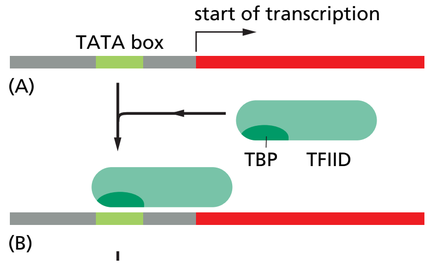
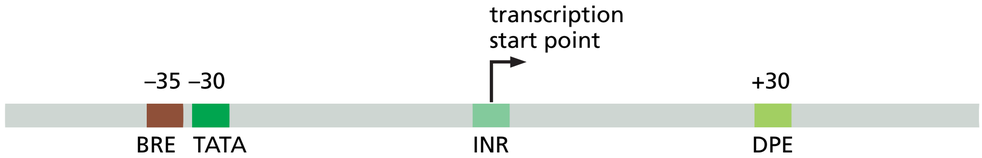


TABLE 6–3 The General Transcription Factors Needed for Transcription Initiation by Eukaryotic RNA Polymerase II

Name	Number of subunits	Roles in transition initiation
TFIID TBP subunit TAF subunits	 1 ~11	 Recognizes TATA box Recognizes other DNA sequences near the transcription start point; regulates DNA-binding by TBP



- **TATA box**: DNA sequence of T and A; typically 25 nt upstream of the transcription start
- **most important one** for RNA polymerase II

element	consensus sequence	general transcription factor
BRE	G/C G/C G/A C G C C	TFIIB
TATA	T A T A A/T A A/T	TBP subunit of TFIID
INR	C/T C/T A N T/A C/T C/T	TFIID
DPE	A/G G A/T C G T G	TFIID

Transcription initiation

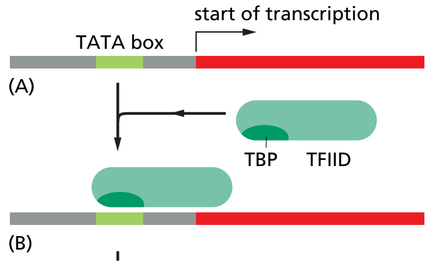
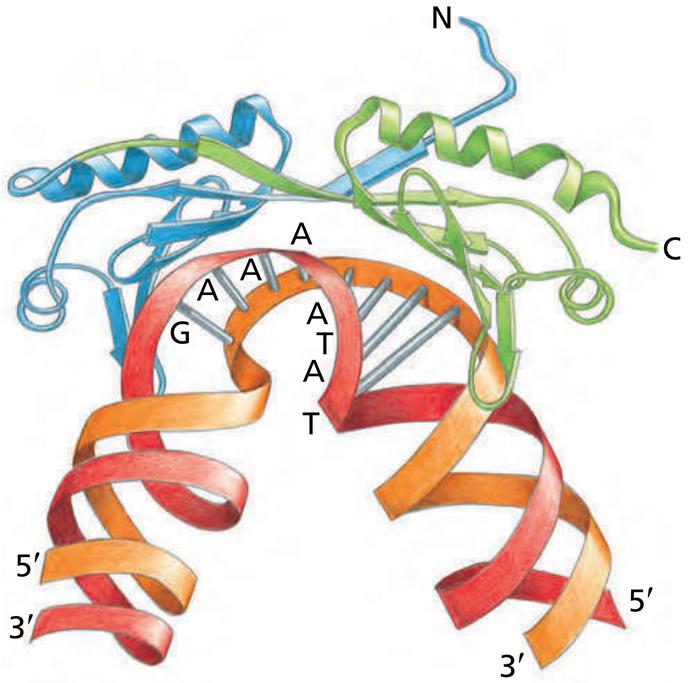


TABLE 6–3 The General Transcription Factors Needed for Transcription Initiation by Eukaryotic RNA Polymerase II

Name	Number of subunits	Roles in transition initiation
TFIID TBP subunit TAF subunits	 1 ~11	 Recognizes TATA box Recognizes other DNA sequences near the transcription start point; regulates DNA-binding by TBP



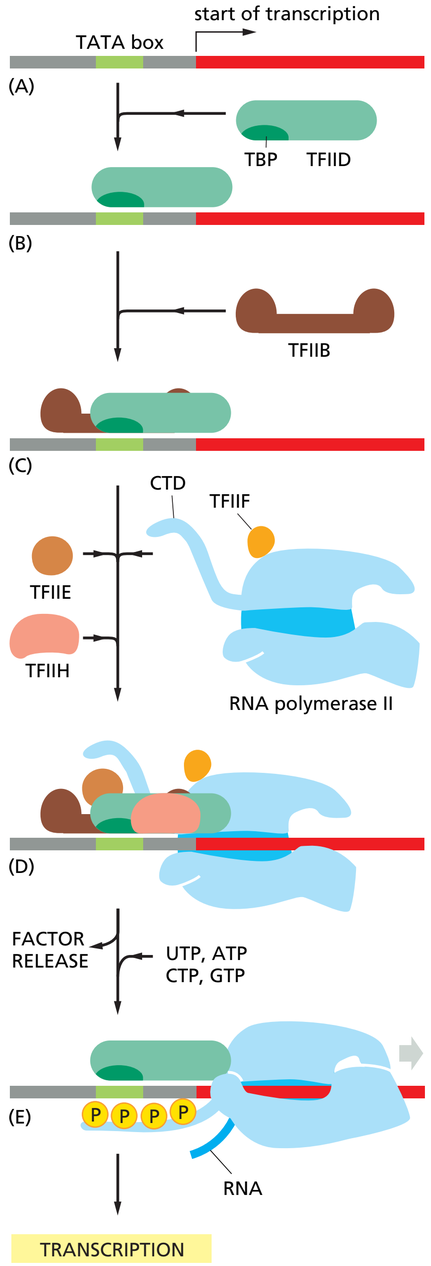
- TBP binding create a **distortion in the DNA** of the TATA box, which allows the assembly of other factors (to form a complete **transcription initiation complex**)

Transcription initiation

TABLE 6–3 The General Transcription Factors Needed for Transcription Initiation by Eukaryotic RNA Polymerase II

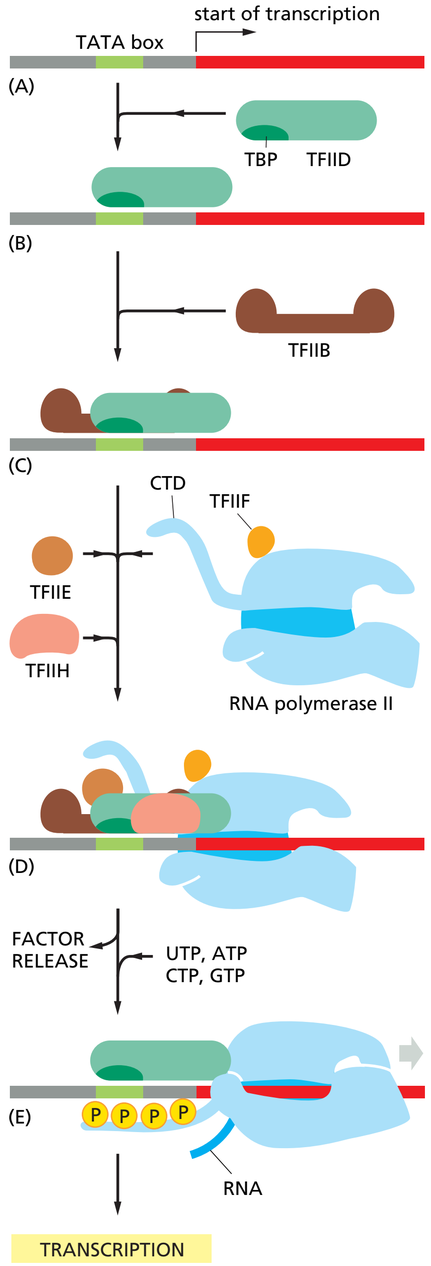
Name	Number of subunits	Roles in transition initiation
TFIID TBP subunit TAF subunits	 1 ~11	 Recognizes TATA box Recognizes other DNA sequences near the transcription start point; regulates DNA-binding by TBP
TFIIB	1	Recognizes BRE element in promoters; accurately positions RNA polymerase at the start site of transcription
TFIIF	3	Stabilizes RNA polymerase interaction with TBP and TFIIB; helps attract TFIIE and TFIIH
TFIIE	2	Attracts and regulates TFIIH
TFIIH	9	Unwinds DNA at the transcription start point, phosphorylates Ser5 of the RNA polymerase CTD; releases RNA polymerase from the promoter

TFIID is composed of TBP and ~11 additional subunits called TAFs (TBP-associated factors); CTD, C-terminal domain.



Transcription initiation

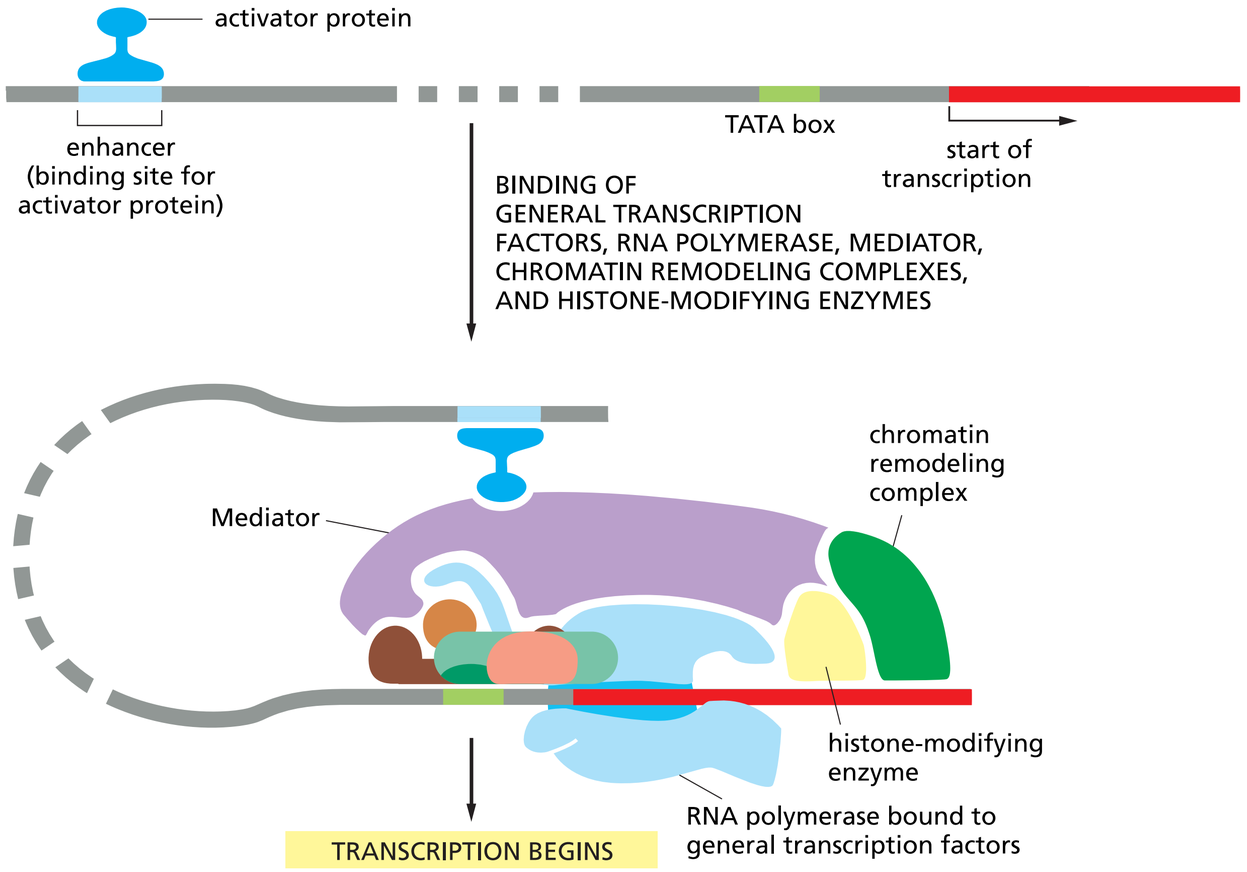
- RNA polymerase II must have **access to the template strand** at the transcription start point
- TFIIH contains a DNA helicase, which hydrolyzes ATP and unwinds DNA
- As for bacteria, RNA polymerase II remains at the **promoter**, synthesizing short RNAs until it goes through **conformational changes** and starts the **elongation phase** of transcription
- A key step in this transition is the addition of **phosphate groups** to the tail of the RNA polymerase (C-terminal domain)
- Once the RNA polymerase II starts elongation, most **general transcription factors are released**



Prokaryote vs. Eukaryote promoters

Feature	Prokaryotes	Eukaryotes
Core promoter elements	-35 and -10 regions	TATA box, Inr, BRE, DPE (variable)
RNA polymerase recognition	Direct via σ factor	Indirect, via transcription factors + RNA pol II
Complexity	Simple, short	Complex, long, variable
Regulation	Local repressors/ activators	Chromatin remodeling, enhancers, silencers
Number of RNA polymerases	Single RNA polymerase	Multiple (Pol I, II, III with distinct promoters)

Activator, mediator and chromatin-modifying proteins



- DNA in eukaryotic cells is **packaged into nucleosomes and complex chromatin structure**
- Need of **additional proteins**

1. **Transcriptional activators** are protein that bind to specific DNA sequences called **enhancers** that help attract the RNA polymerase (discussed later)

2. **Transcriptional repressors** bind DNA sequences called **silencers**

3. **Mediator** is a large protein complex that allows the **communication** between the activator and RNA polymerase complex

4. **Chromatin-remodelling complex** and **histone-modifying enzymes**

Elongation factors

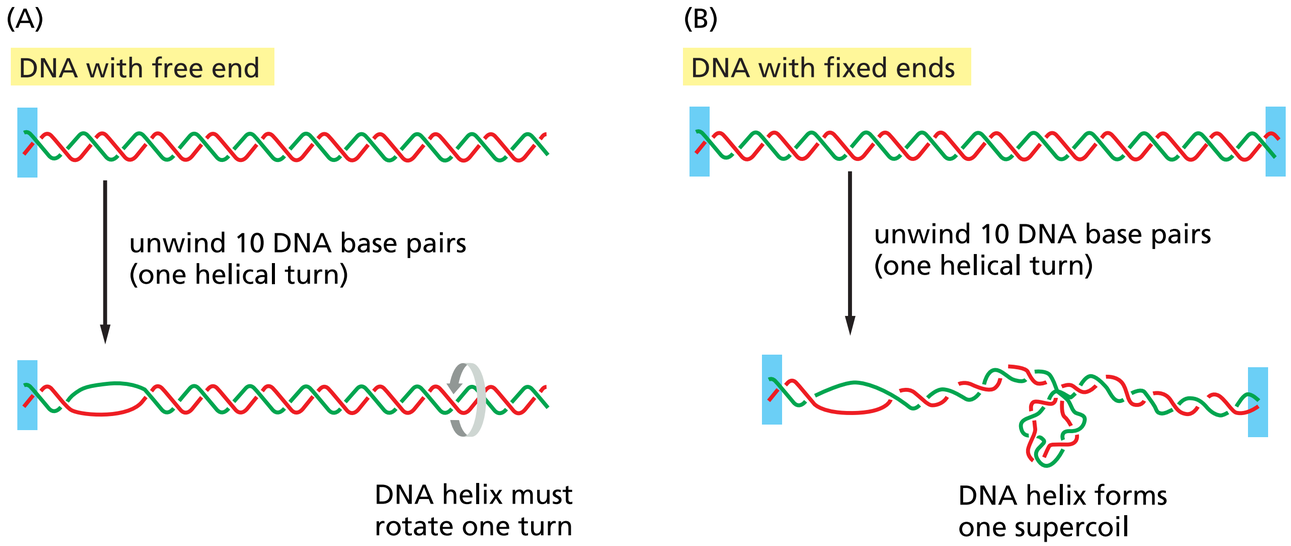
- Once it starts elongating, RNA polymerases are associated with a series of **elongation factors** (both in bacteria and eukaryotic cells)

- These factors are proteins that **decrease the likelihood that RNA polymerase dissociates** from the DNA before it reaches the end of the gene

- Some of these proteins also help **with chromatin remodelling and histone modifications** in eukaryotes

Superhelical tension

- **DNA supercoiling** is the conformation that DNA adopts in response to superhelical tension
- RNA polymerase creates superhelical tension as the ends of the stretch of DNA are anchored
- In eukaryotes, this might help with **DNA unwrapping** from nucleosome
- Resolved by **DNA topoisomerases (DNA gyrase)**



Main actors for Eukaryotes (summary)

1. RNA Polymerases

- RNA polymerase I: Transcribes most rRNA genes (except 5S rRNA).
- RNA polymerase II: Transcribes all mRNAs (protein-coding genes), snRNAs, some miRNAs, and lncRNAs.
- RNA polymerase III: Transcribes tRNAs, 5S rRNA, and some small RNAs.

2. General Transcription Factors (GTFs)

Required for initiation by RNA polymerase II:

- TFIID: recognizes core promoter (TATA box)
- TFIIA: Stabilizes TFIID binding to DNA
- TFIIB: Helps recruit RNA polymerase II and defines transcription start site
- TFIIF: Escorts RNA polymerase II to the promoter
- TFIIE: Recruits TFIIH
- TFIIH: Helicase (unwinds DNA) + kinase (phosphorylates CTD of RNA Pol II to start elongation)

Main actors for Eukaryotes (summary)

3. Promoter and Regulatory DNA Elements

- Core promoter: TATA box, initiator (Inr), BRE, DPE elements.
- Enhancers/silencers: Distal DNA sequences bound by activators or repressors to regulate transcription.

4. Transcription Activators & Repressors

- Activators: Bind enhancers; stimulate transcription
- Repressors: Bind silencers; inhibit transcription

5. Co-activators & Mediator Complex

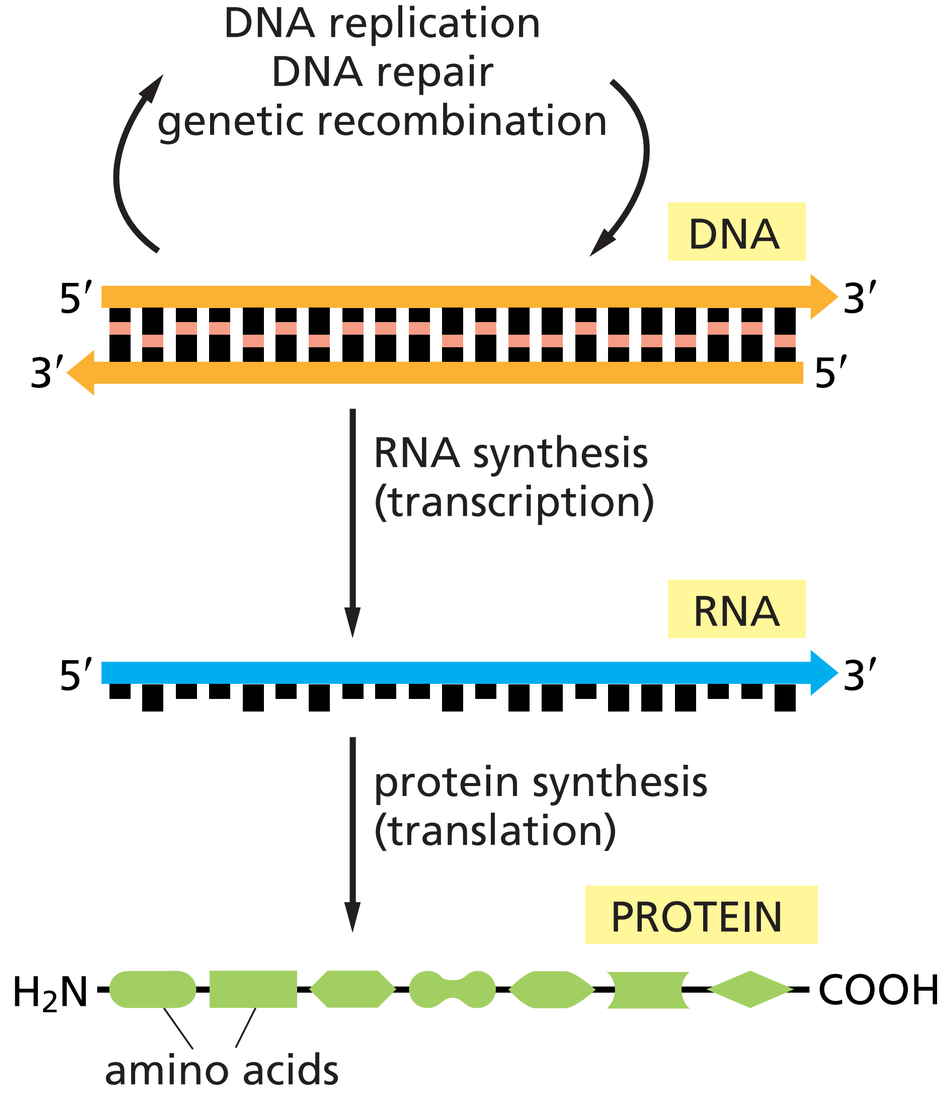
- Mediator: Large multiprotein complex bridging activators/repressors with RNA Pol II.
- Chromatin remodelers: reposition nucleosomes for access.

6. Elongation & Processing Factors

- Elongation factors: Ensure processivity of RNA Pol II and overcome nucleosome barriers.
- CTD(=Ct domain of RNA pol II)-binding proteins: Recruit capping enzymes, splicing machinery, and polyadenylation factors (see next)

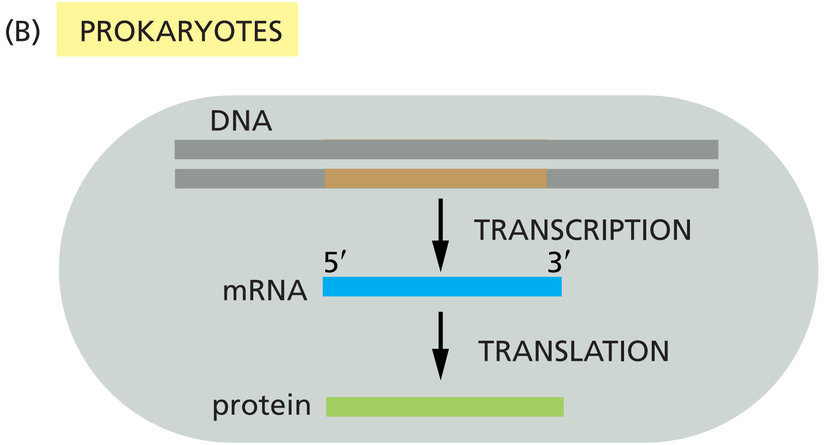
Plan

- RNA
- Transcription
- Transcription initiation
- RNA processing
- Non-coding RNAs

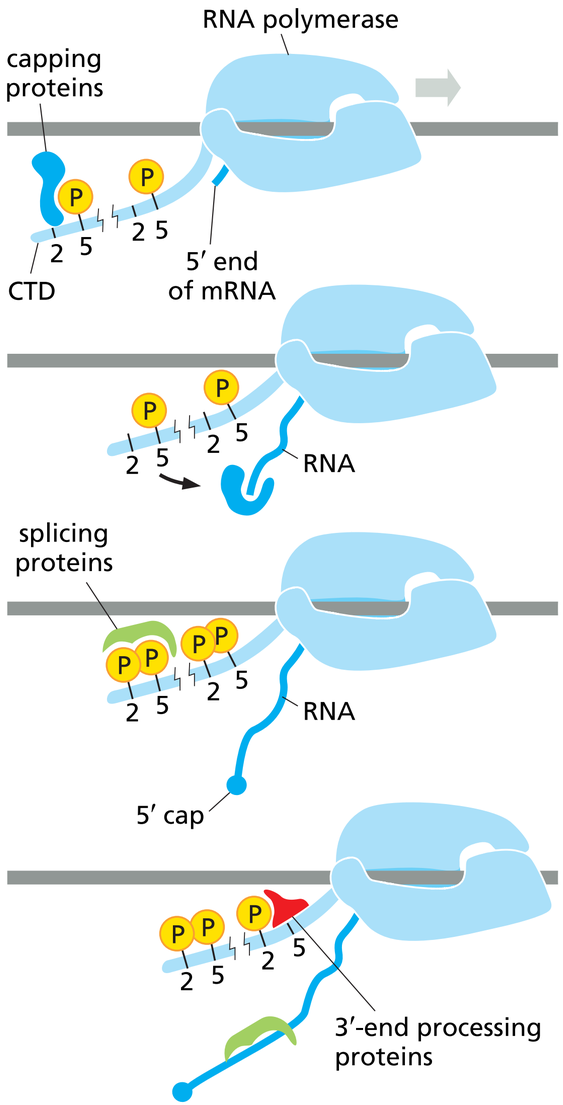


RNA processing

- In **bacteria**, mRNAs are synthesized by the RNA polymerase starting and stopping at specific spots

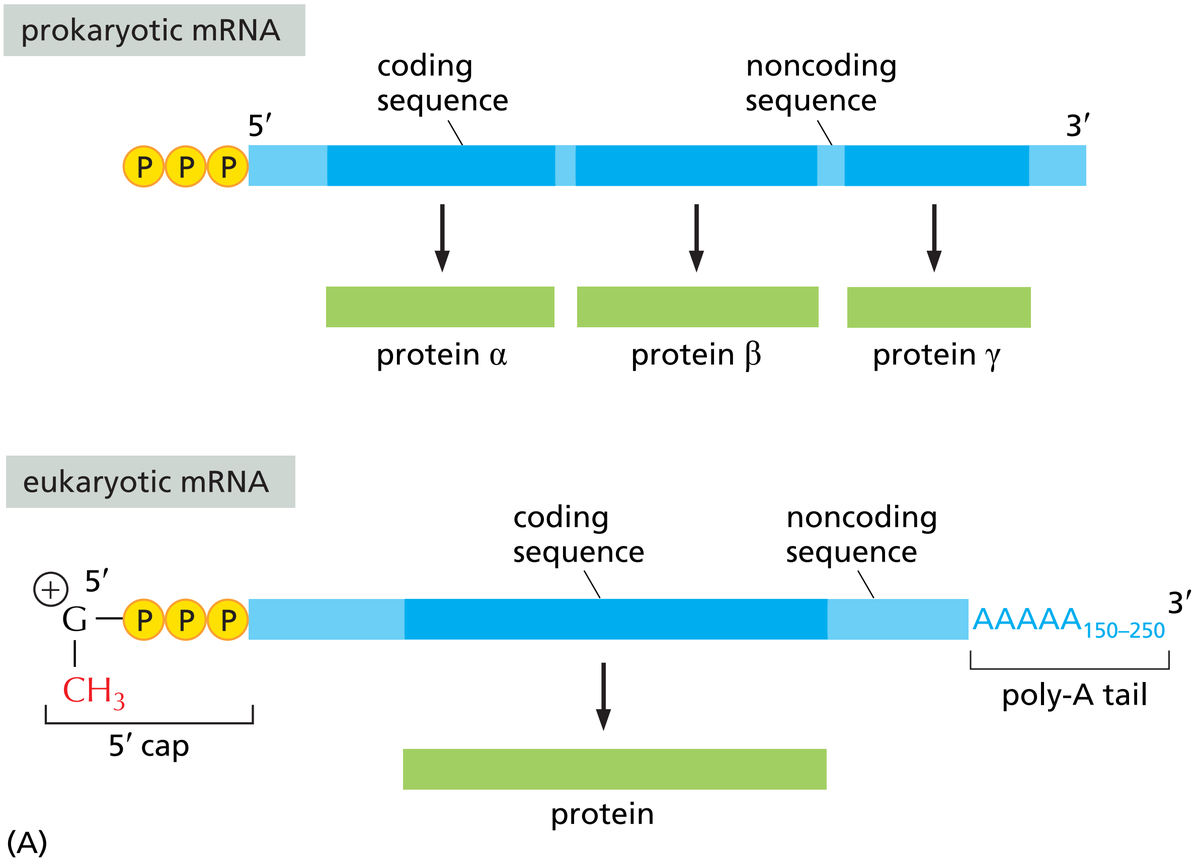


RNA processing



- In **Eukaryotes**, transcription is only the **first step** of producing a mature mRNA
- Other steps are **modification of the ends of the RNA** and the removal of introns sequences (**RNA splicing**)
- All **coupled** with transcription elongation
- During transcription initiation, the polymerase tail is **phosphorylated** at its Ct (=C terminal) domain
- This allows the association of a **new set of proteins** to the RNA polymerase

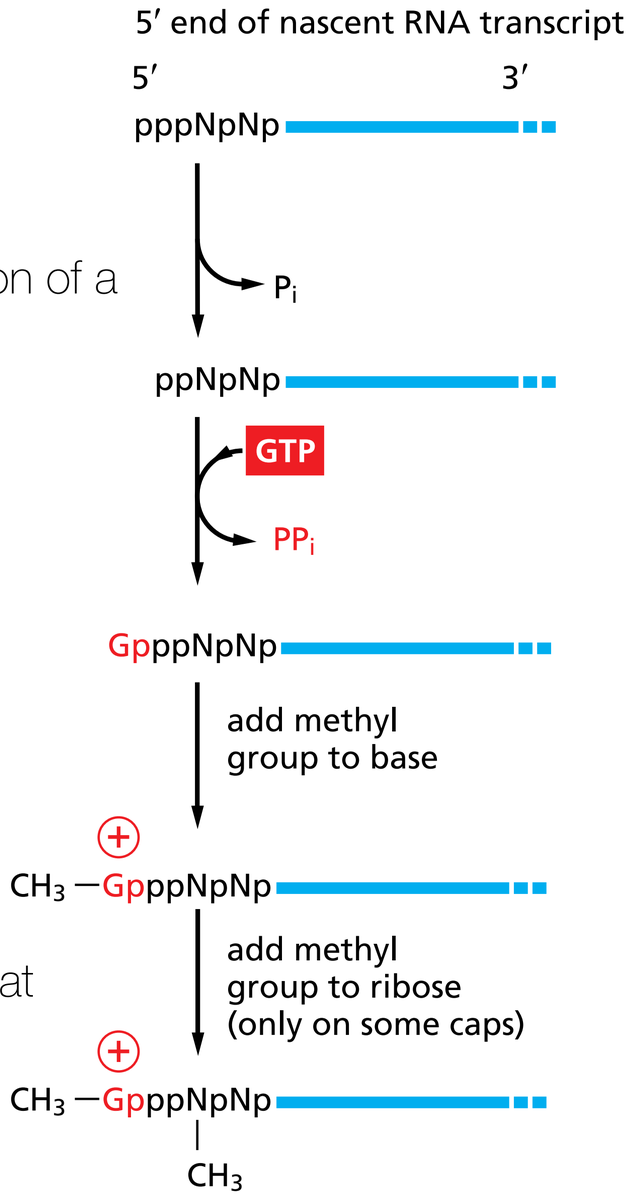
RNA ends modifications



- **Both ends** are modified
  - **capping** of the 5’ end
  - **polyadenylation** of the 3’ end
- Allows the cell to assess if both ends are there and the **mRNA is intact** before being exported out of the nucleus for translation

RNA capping

- **First modification** of eukaryotic pre-mRNAs
- After ~25 nt of RNA are produced, the 5' end of the RNA is modified by addition of a **cap**
- The cap consists of a **modified guanine nucleotide**
- **3 enzymes:**
    1. **Phosphatase** removes a phosphate from the 5' end of the RNA
    2. A **guanyl transferase** adds a GMP in a reverse linkage (5' ot 5')
    3. A **methyl transferase** adds a methyl group to the guanosine
- All bind to the **RNA polymerase phorsphorylated at Ser5** (a modification that happens during transcription initiation)
- The cap helps the cell to **distinguish** between mRNAs and other RNAs



Have a nice day!

No lecture next week, so enjoy your holiday!