

In the Name of God, the Merciful, the Compassionate

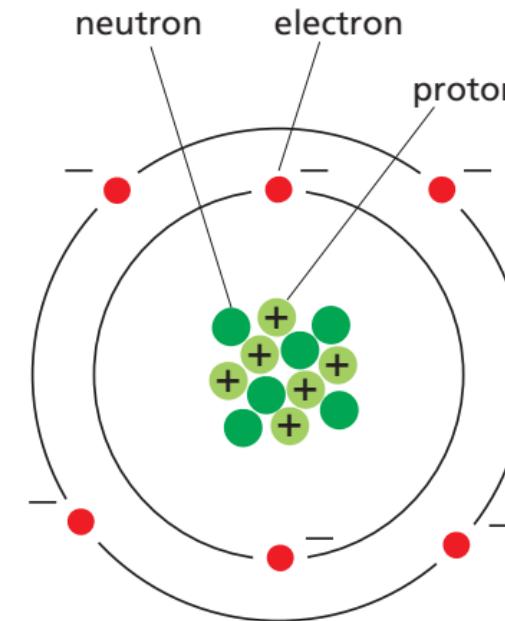
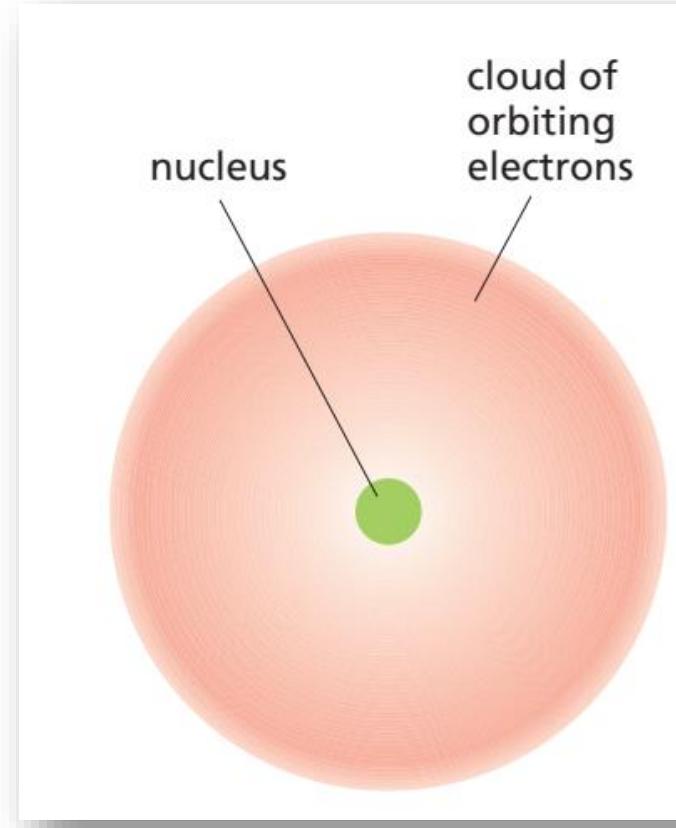
Introduction to Bioinformatics

02 - Introduction to Cell Biology

Instructor: Hossein Zeinali
Amirkabir University of Technology

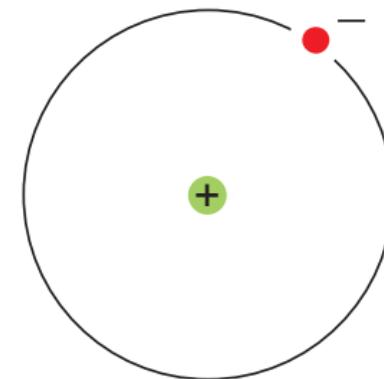


Atoms



carbon atom

atomic number = 6
atomic weight = 12

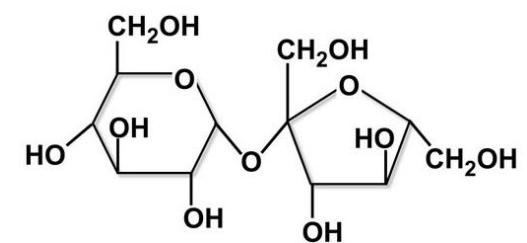
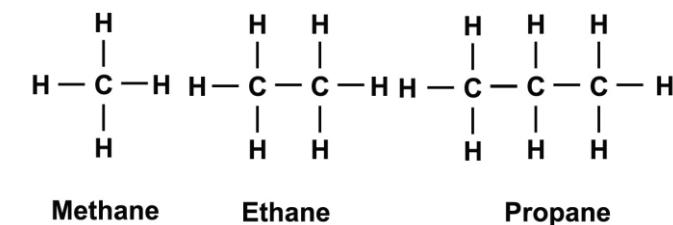
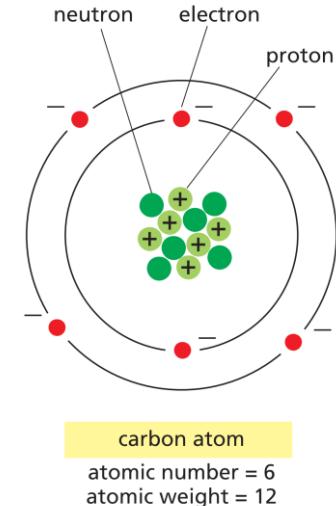


hydrogen atom

atomic number = 1
atomic weight = 1

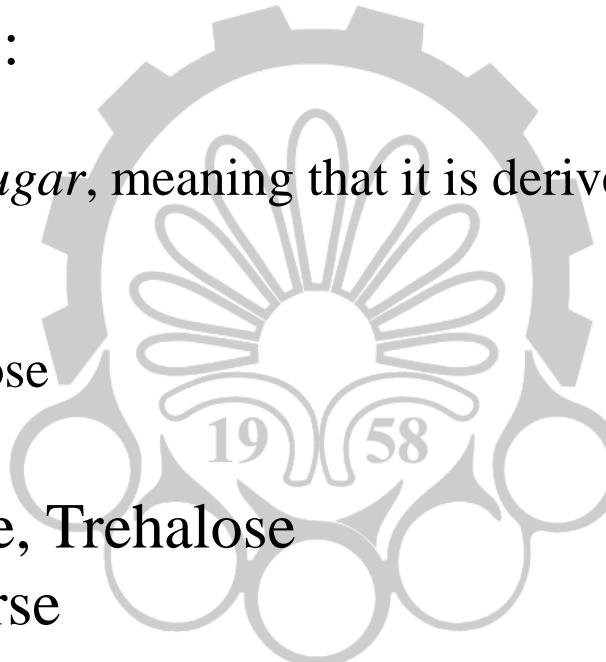
Carbon

- Carbon has the ability to bond with other nonmetallic elements and forms a vast number of compounds, more than any other element. So, it has often been referred to as the *king of the elements*.
- Carbon is the *second most abundant element* in the human body by mass after oxygen.
- Carbon 12 is the main isotope of Carbon with 6 protons and 6 neutrons.
- **Hydrocarbon:** is an organic compound consisting entirely of **hydrogen** and **carbon**.
- **Carbohydrate:** is a biomolecule consisting of **carbon**, **hydrogen** and **oxygen** atoms.



Carbohydrate Types

- *Monosaccharides:*
 - **Pentose** (5 carbon atoms):
 - **Ribose:** $C_5H_{10}O_5$
 - **Deoxyribose:** is a *deoxy sugar*, meaning that it is derived from the sugar ribose by loss of an oxygen atom.
 - Hexose (6 carbon atoms):
 - Glucose, Fructose, Galactose
- *Disaccharides:*
 - Sucrose, Lactose, Maltose, Trehalose
 - Not important in this course
- *Polysaccharides:*
 - Glycogen, Cellulose, Hydrocolloids, Starch
 - Not important in this course



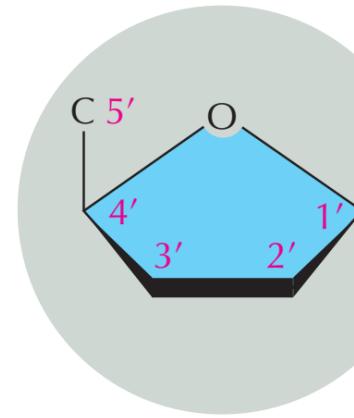
Carbohydrate Types

- *Monosaccharides:*

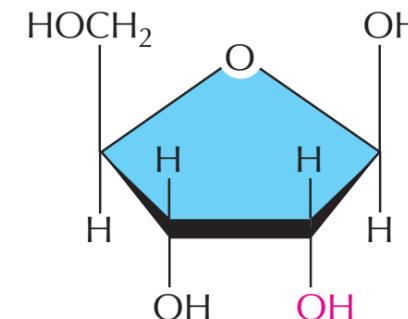
SUGARS

PENTOSE

a five-carbon sugar

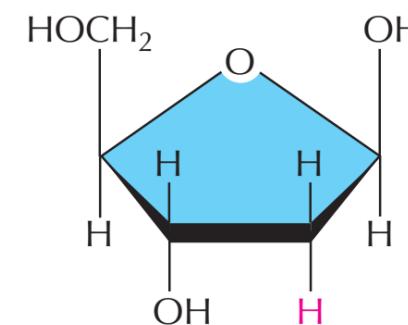


two kinds of
pentoses are used



β-D-ribose

used in ribonucleic acid (RNA)



β-D-2-deoxyribose

used in deoxyribonucleic acid (DNA)

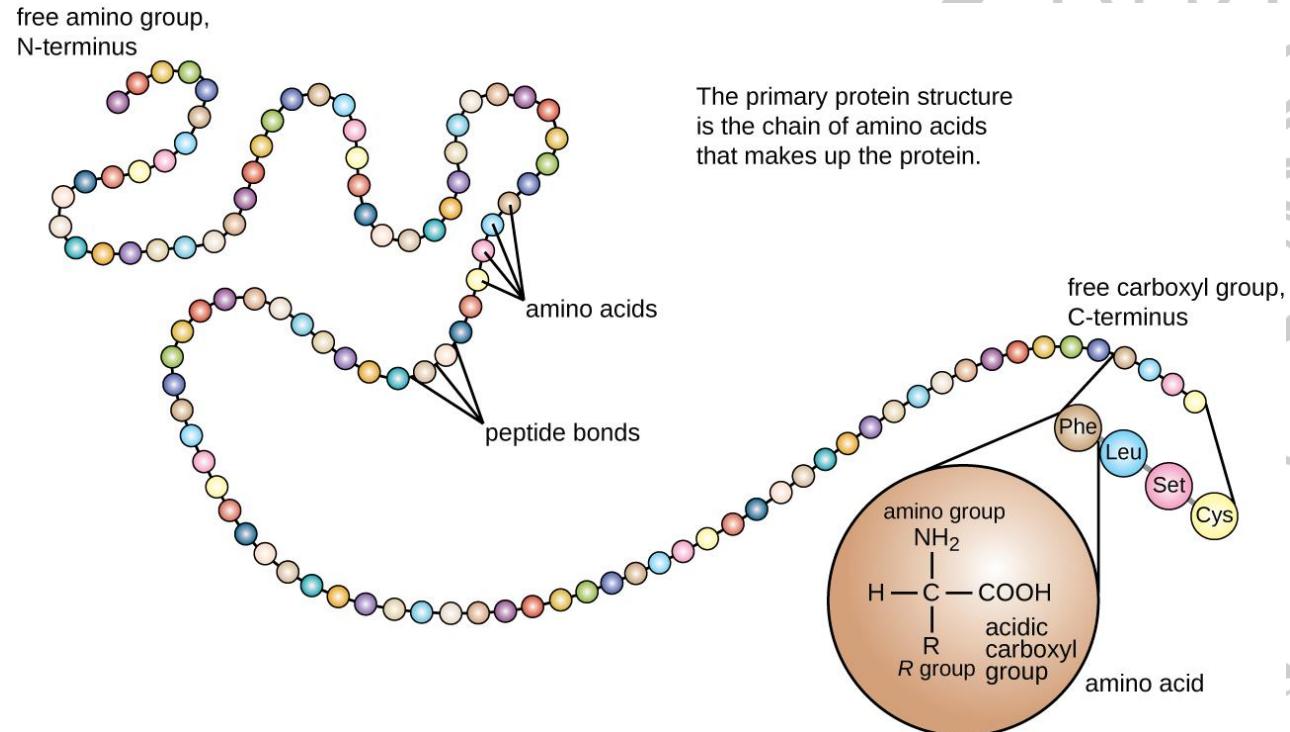
Each numbered carbon on the sugar of a nucleotide is followed by a prime mark; therefore, one speaks of the "5-prime carbon," etc.

– Not important in this course

Tehran Polytechnic

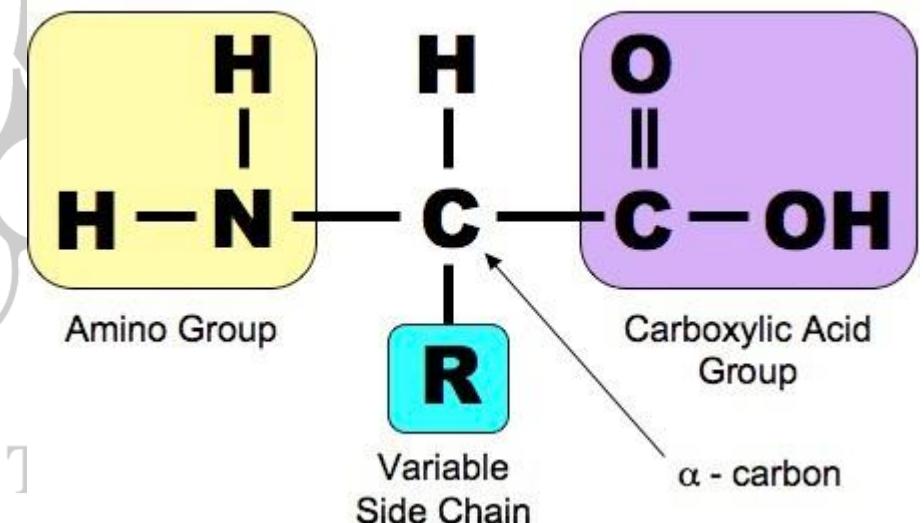
Proteins

- Proteins are linear polymers of amino acids (monomer) whose order is often called the primary structure or sequence.



The primary protein structure
is the chain of amino acids
that makes up the protein.

of]
technic)



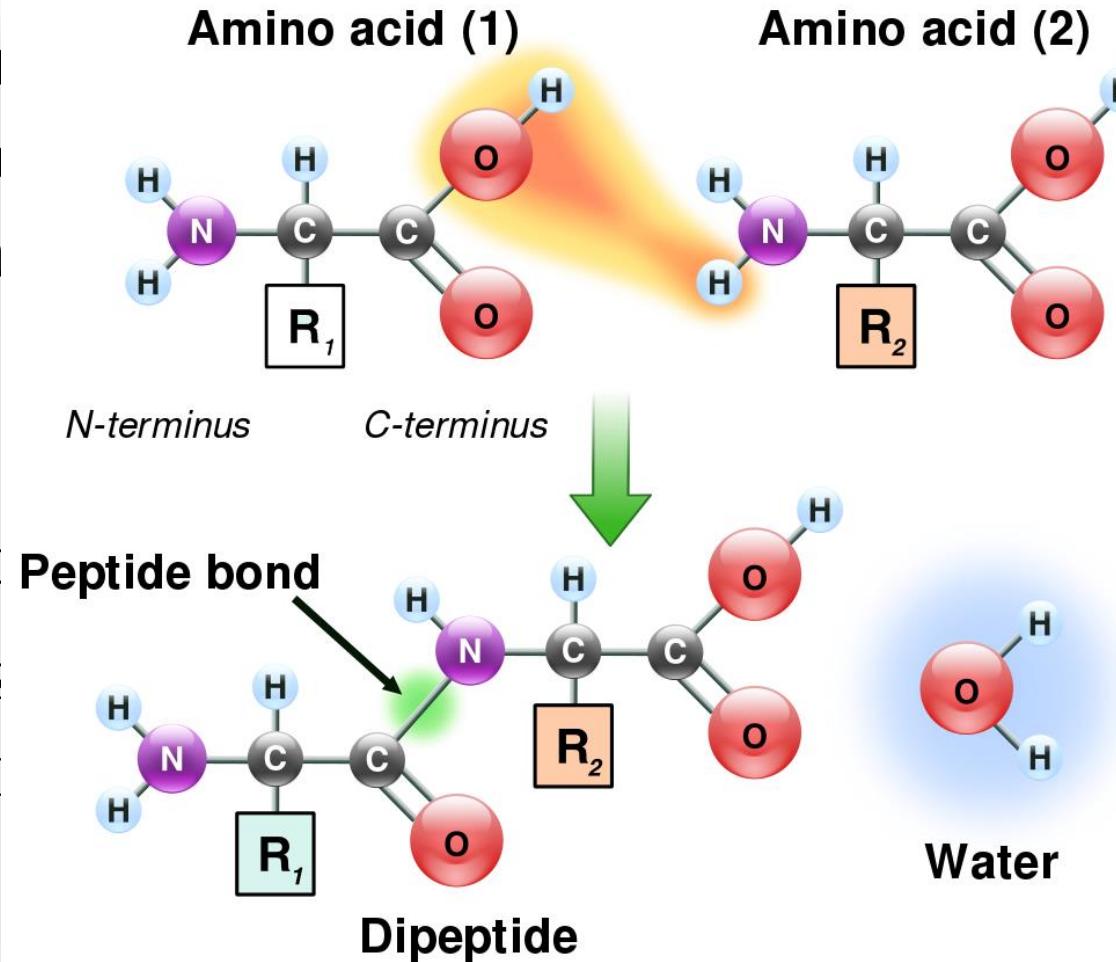
Peptide Bonds

- A **peptide bond** is a type of chemical bond linking two consecutive amino acids from C1 (carbon of Carboxylic group) of one amino acid and N2 (nitrogen of Amino group) of another along a peptide or protein chain, causing the release of a molecule of water (H_2O), hence the process is a dehydration synthesis reaction.
- **Dipeptide** is an organic compound derived from two amino acids.
- **Polypeptide** is a single linear chain of many amino acids (any length), held together by amide bonds.

Amirkabir University of Technology
(Tehran Polytechnic)

Peptide Bonds

- A **peptide bond** is formed between two amino acids from the C=O group of one acid and N-H (nitrogen) of another or protein chain. Hence the process is called condensation.
- **Dipeptide** is a peptide consisting of two amino acids.
- **Polypeptide** is a peptide consisting of many amino acids (any length), held together by peptide bonds.



two consecutive atoms of one amino acid along a peptide chain. It is formed by the loss of water (H_2O), which is bonded to the oxygen atoms of amino acids. A peptide bond consists of two amino acids (any length), held together by peptide bonds.

Enzymes

- Enzymes are both proteins and biological catalysts (biocatalysts) and accelerate chemical reactions.
- Almost all metabolic processes in the cell need enzyme catalysis in order to occur at rates fast enough to sustain life.
- Enzymes are produced inside the cells.
- An **exoenzyme**, or **extracellular enzyme**, is an enzyme that is produced by a cell and functions outside of that cell.
- Exoenzymes are produced by both prokaryotic and eukaryotic cells.

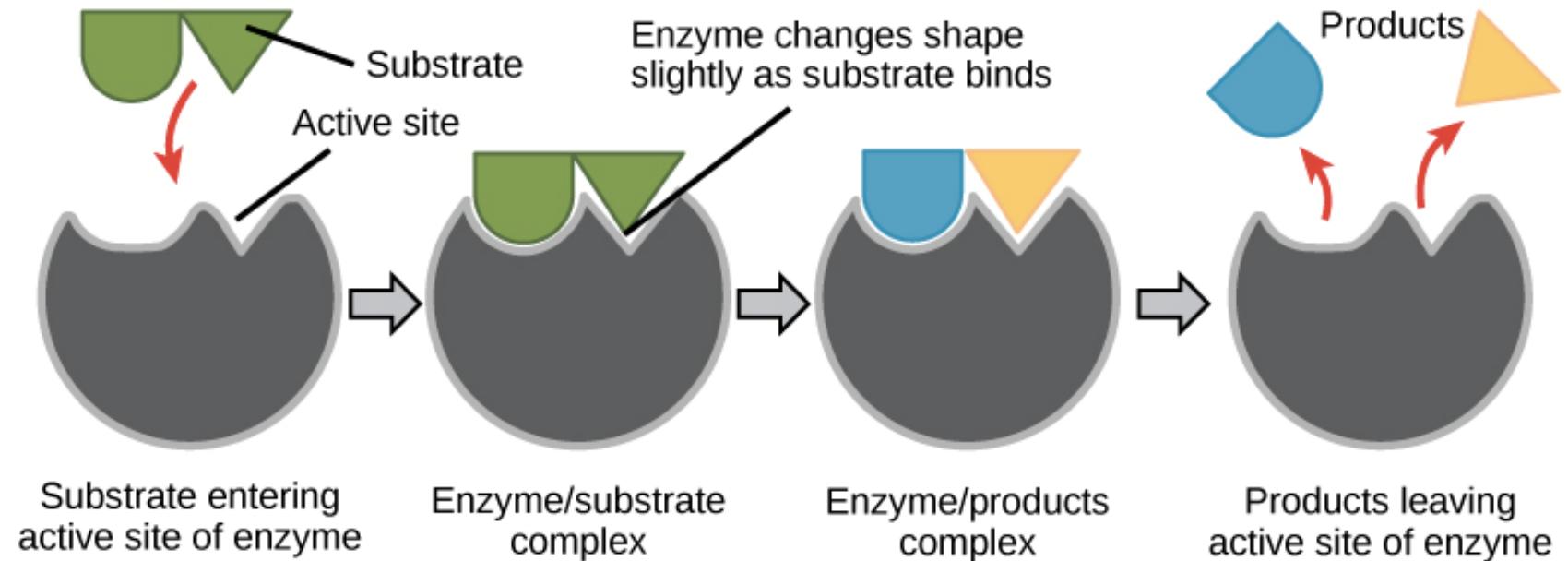
Properties of Enzymes

- Catalytic property:
 - The enzymes remain unchanged after the reaction and reused by the cells several times.
- Specificity:
 - Enzymes are very specific in their action.
- Reversibility:
 - Most of the enzymes catalyzed reactions are reversible.
- Sensitiveness to heat and temperature:
 - Enzymes will be inactive at very low temperature. The enzymatic activity increases with increase in temperature up to a certain level.
- Sensitiveness pH: **Amirkabir University of Technology
(Tehran Polytechnic)**
 - Enzymes are specific for their pH requirement.

How Do Enzymes Work?

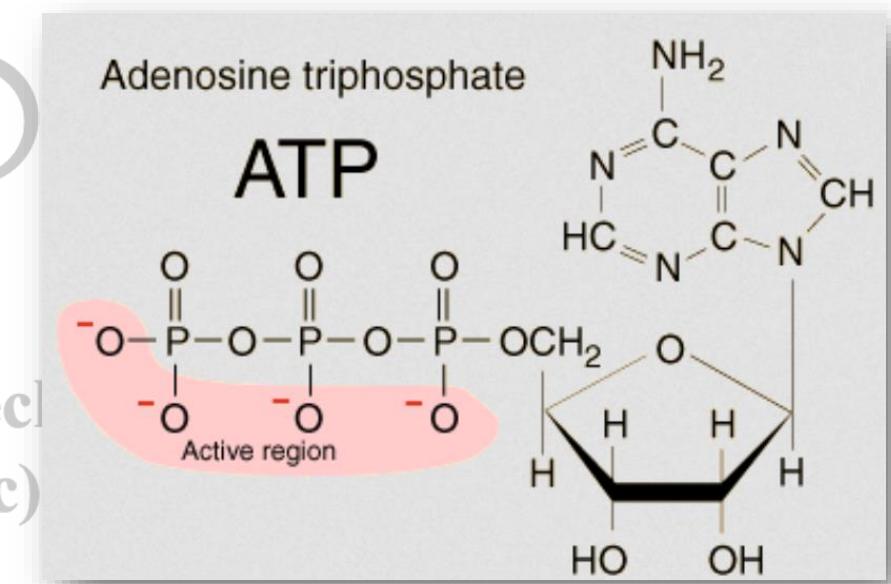
- Names:

- Protease: affects on proteins
- Carbohydrase: affect on carbohydrates
- Lipase: affect on lipids



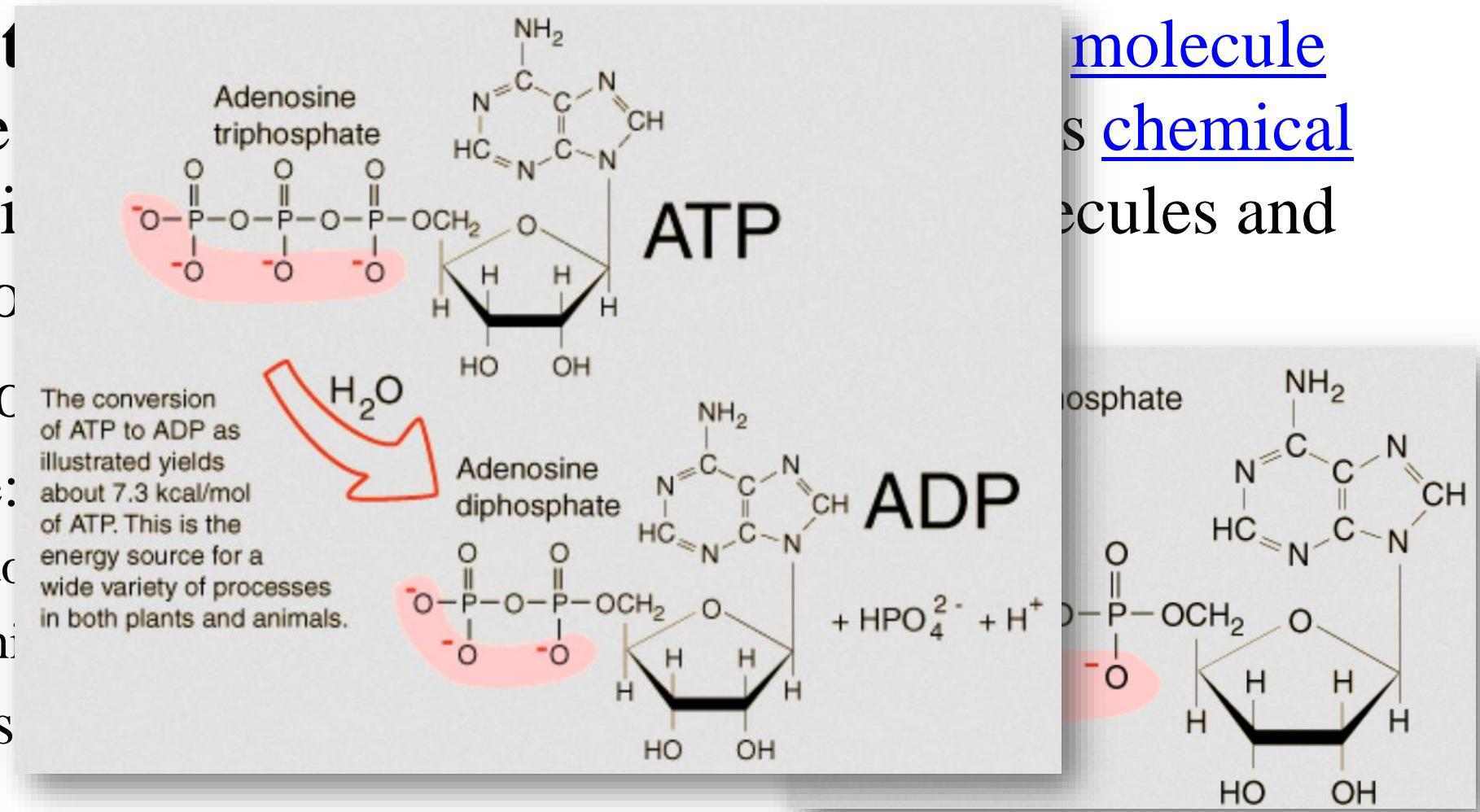
Adenosine Triphosphate (ATP)

- **Adenosine triphosphate (ATP)**, energy-carrying molecule found in the cells of all living things. ATP captures chemical energy obtained from the breakdown of food molecules and releases it to fuel other cellular processes.
- ATP has two parts:
 - Adenosine:
 - One pentose molecule (ribose)
 - One adenine (a kind of base) molecule
 - Three phosphate molecules



Adenosine Triphosphate (ATP)

- Adenosine triphosphate (ATP) is found in the cells of all living things. It stores energy obtained from food and releases it to power cellular work.
- ATP has two components:
 - Adenosine:
 - One pentose sugar
 - One adenine base
 - Three phosphate groups



Cell Types

- There are two types of cell:

- **Prokaryote:**

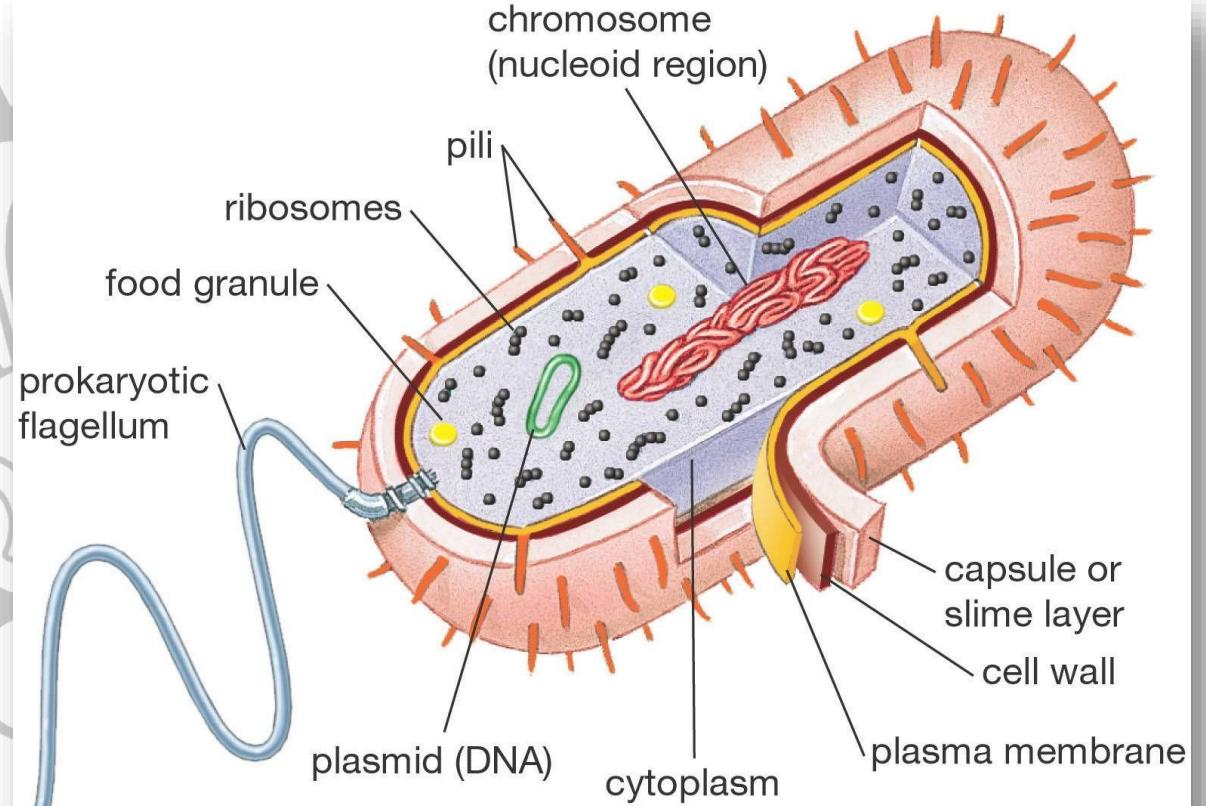
- In bacteria
 - No nucleus
 - Their genomes are circular

- **Eukaryotes:**

- In other animals and plants
 - Linear genomes with multiple chromosomes in pairs.



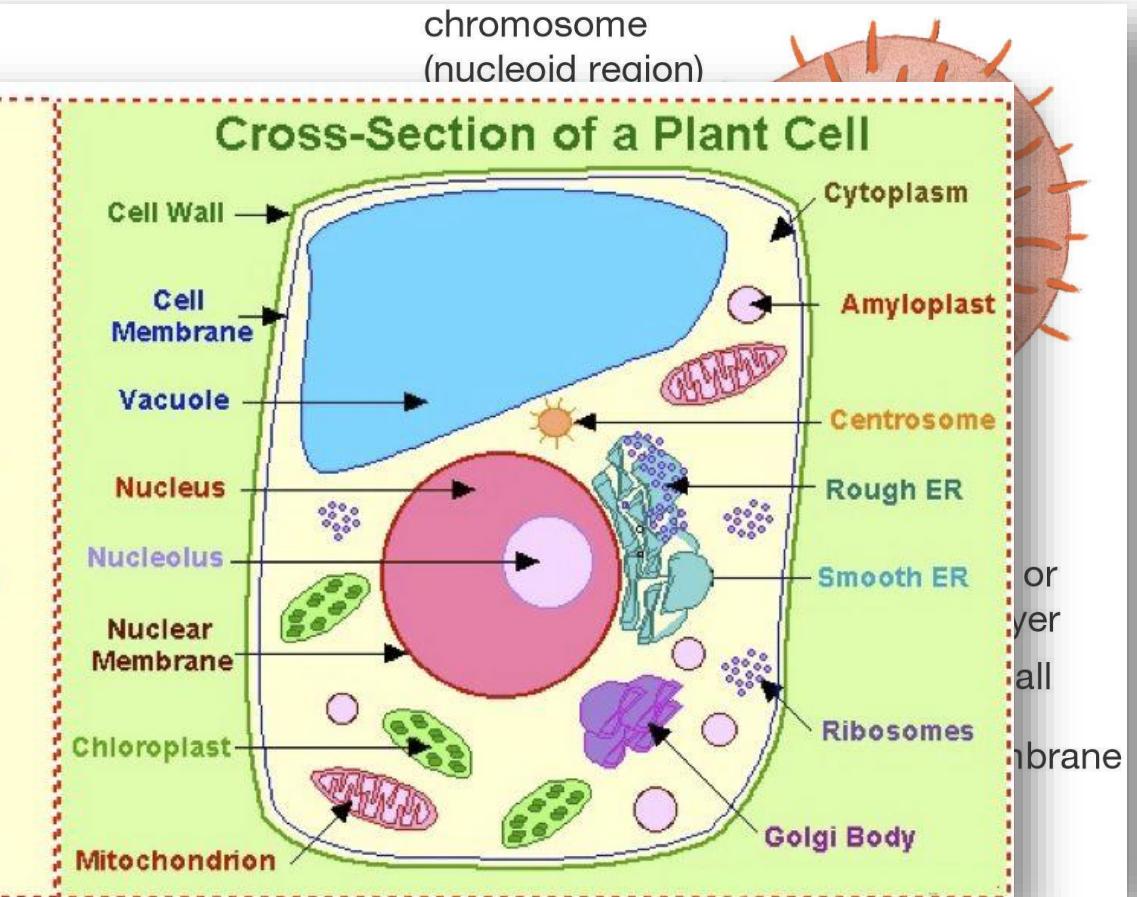
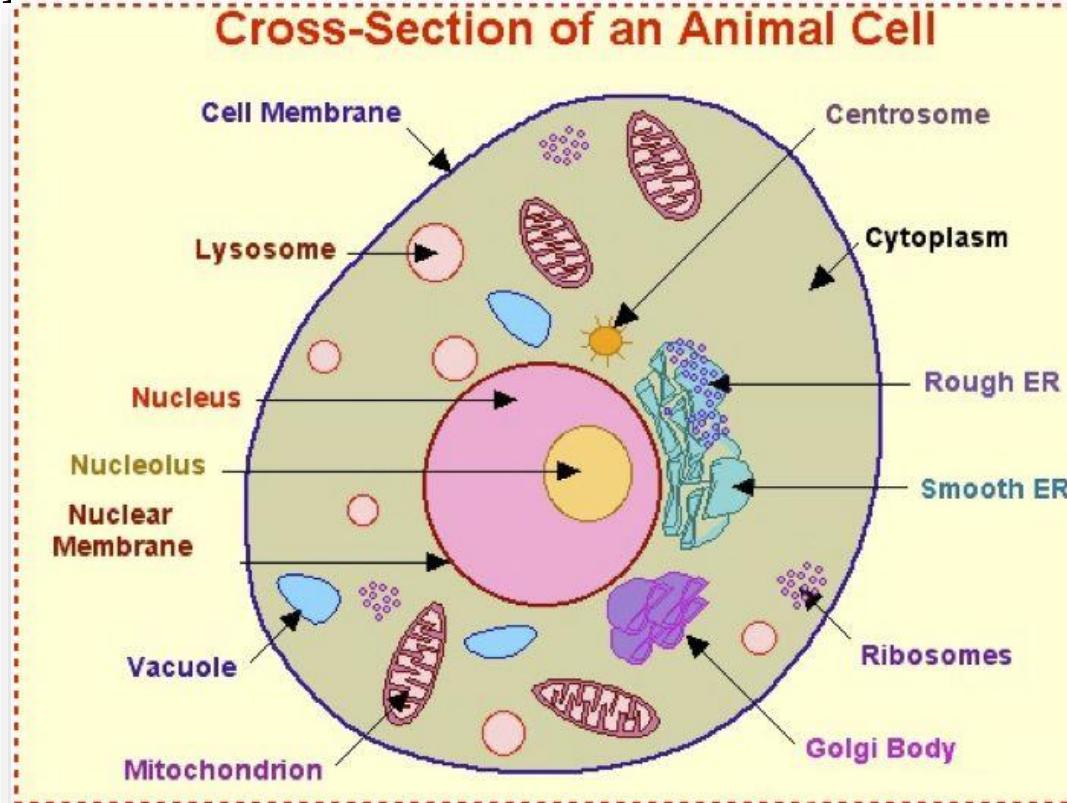
Amirkabir University
(Tehran Polytechnic)



Copyright © 2005 Pearson Prentice Hall, Inc.

Cell Types

- There are two types of cell:



(Tenran Polytechnic)

Shape and Size of Cells

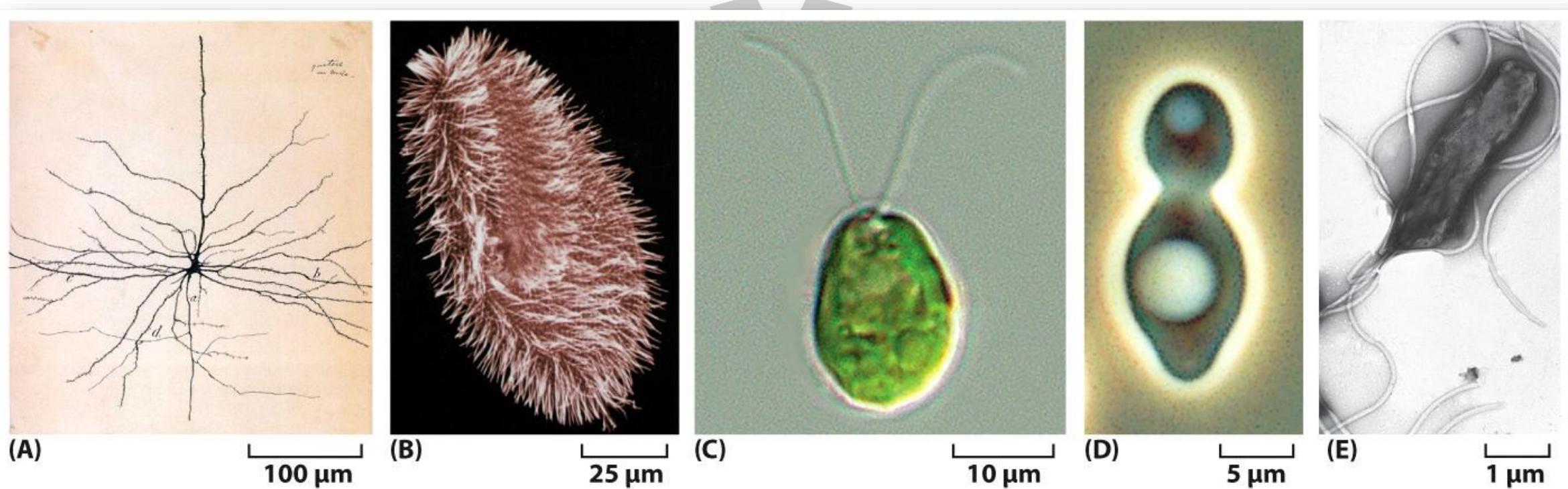
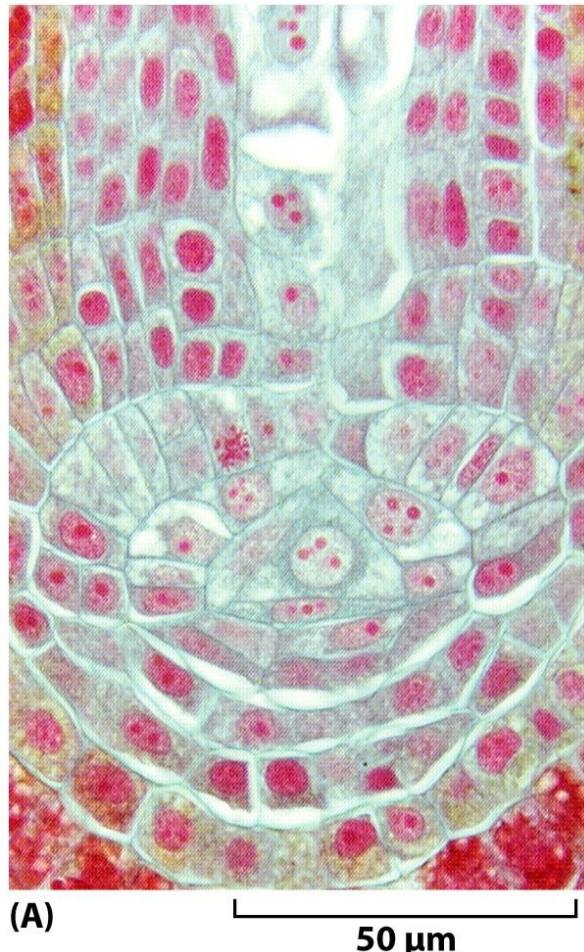


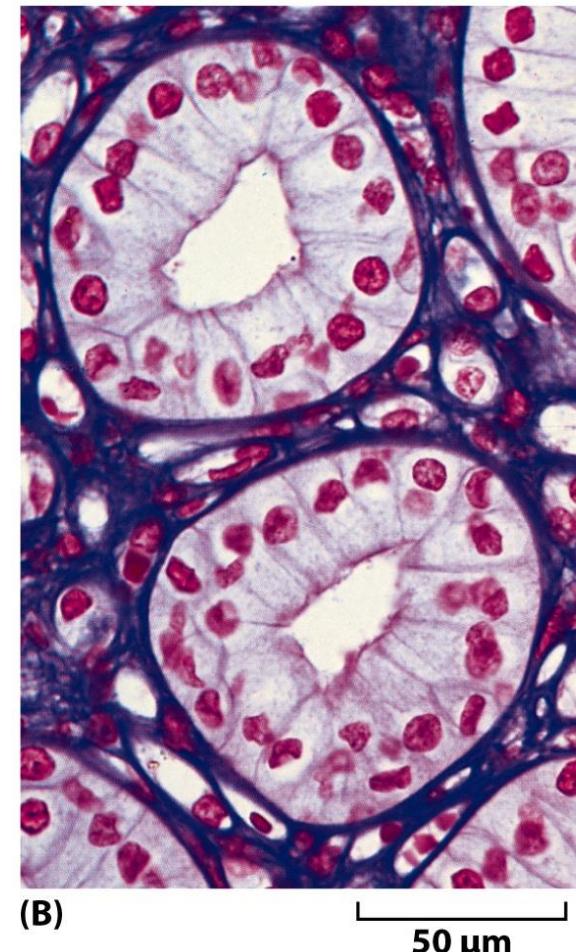
Figure 1-1 Essential Cell Biology, 4th ed. (© Garland Science 2014)

(Tehran Polytechnic)

Cells form Tissues



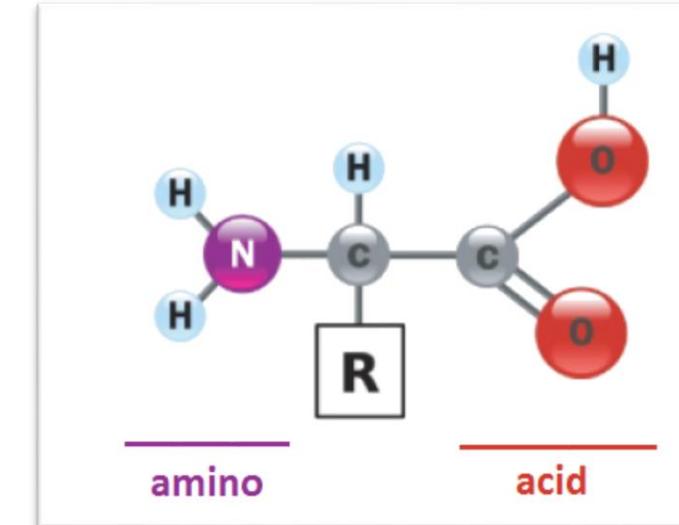
plant root tip



animal kidney tubules

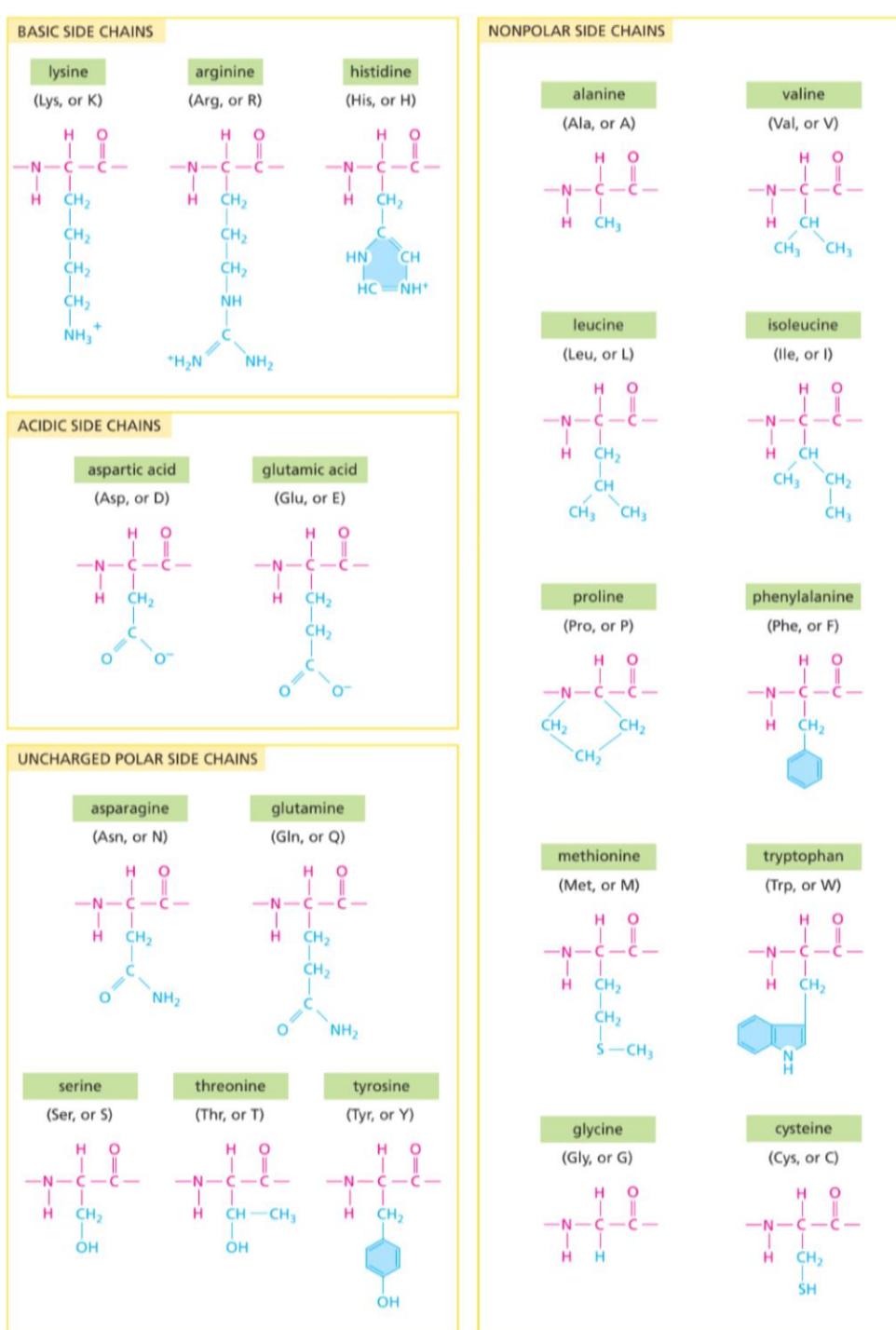
Different Types of Amino Acid

- Proteins are made up of 20 types of naturally occurring amino acids
 - Also a few other amino acids occurring infrequently.
- Amino acids can be classified into overlapping groups:
 - With nonpolar side chain (not liking water)
 - With polar side chain (liking water)
 - Charged
 - Uncharged
 - Specific amino acids
- The functional properties of proteins are almost entirely due to the side chains of the amino acids.
- As there are 20 distinct amino acids that occur in proteins, there can be 20^n different polypeptide chains of length n .

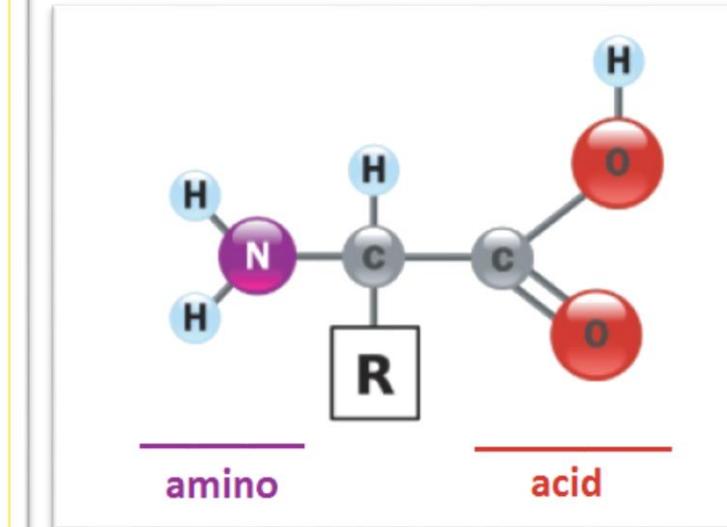


Different Types

- Proteins are made up of amino acids
 - Also a few other amino acids
- Amino acids can be
 - With nonpolar side chains
 - With polar side chains
 - Charged
 - Uncharged
 - Specific amino acids
- The functional properties of proteins depend on the side chains of the amino acids
- As there are 20 distinct amino acids, there can be 20^n different polypeptides



amino acids



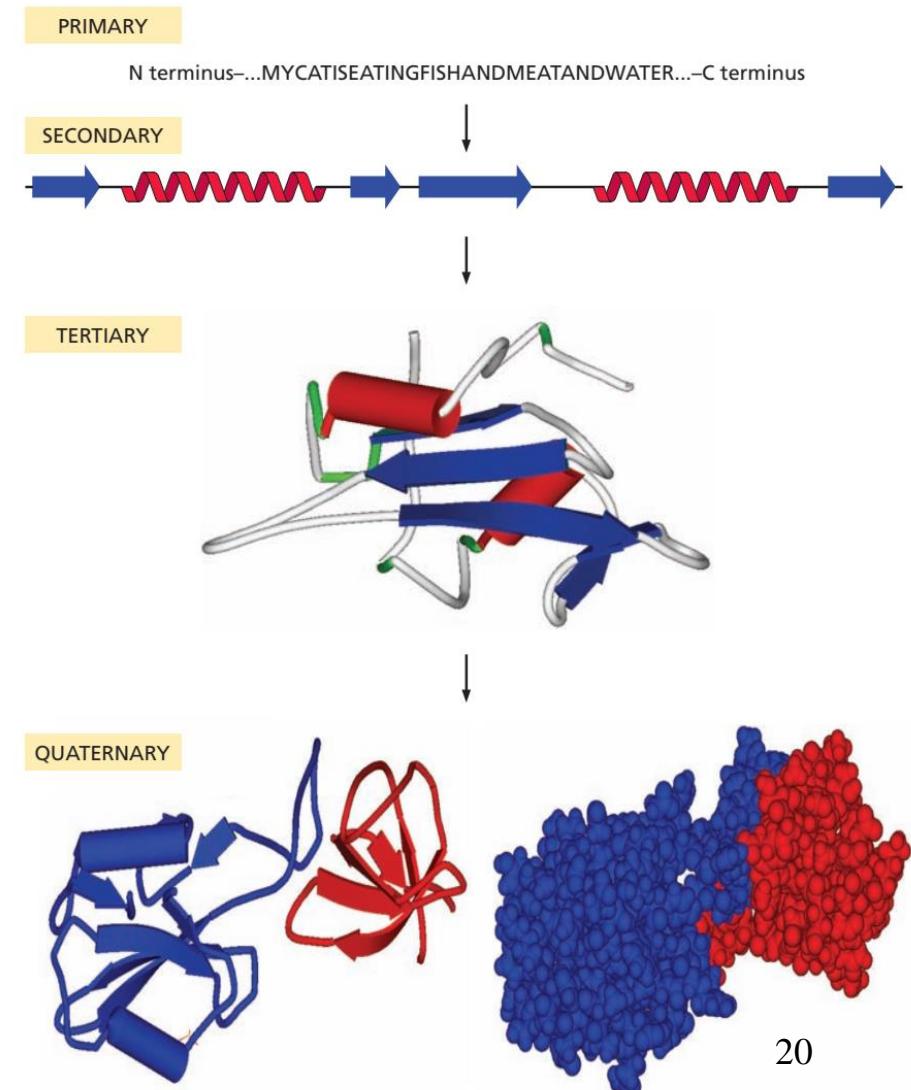
due to the side

chains, there can be 20^n

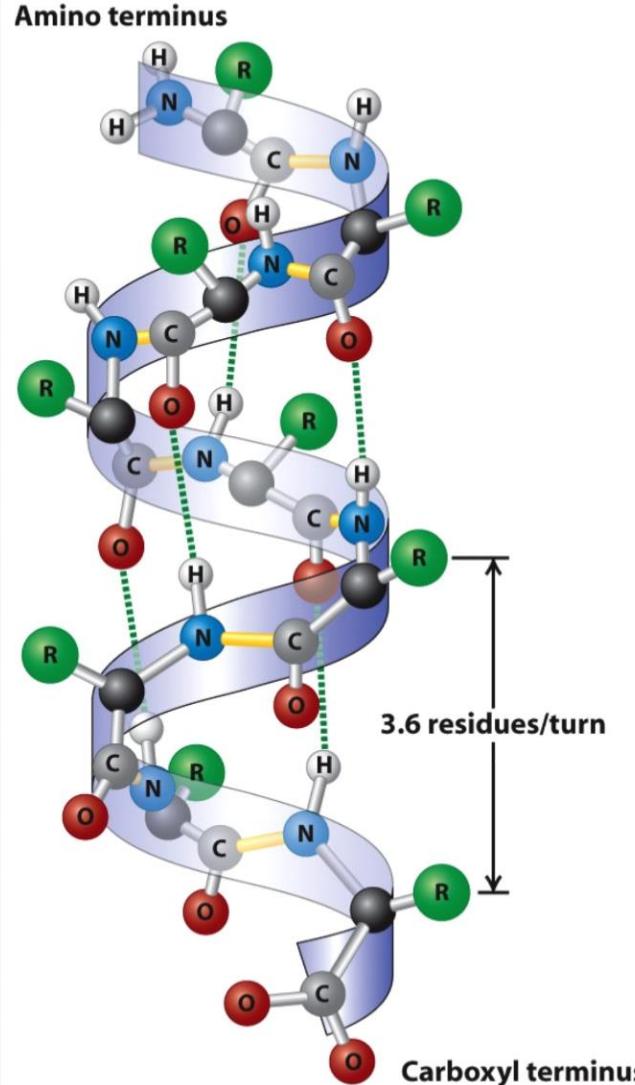
Different Levels of Protein Structure

1. **Primary structure** is the protein sequence, the types and order of the amino acids in the protein chain.
2. **Secondary structure** is the first level of protein folding, in which parts of the chain fold to form generic structures.
3. **Tertiary structure** is formed by the further folding and packing together.
4. **Quaternary structure** is a subunit composition and arrangement in a multi-subunit proteins.

Amirkabir University of Tech
(Tehran Polytechnic)



Alpha Helix (Secondary Structure)

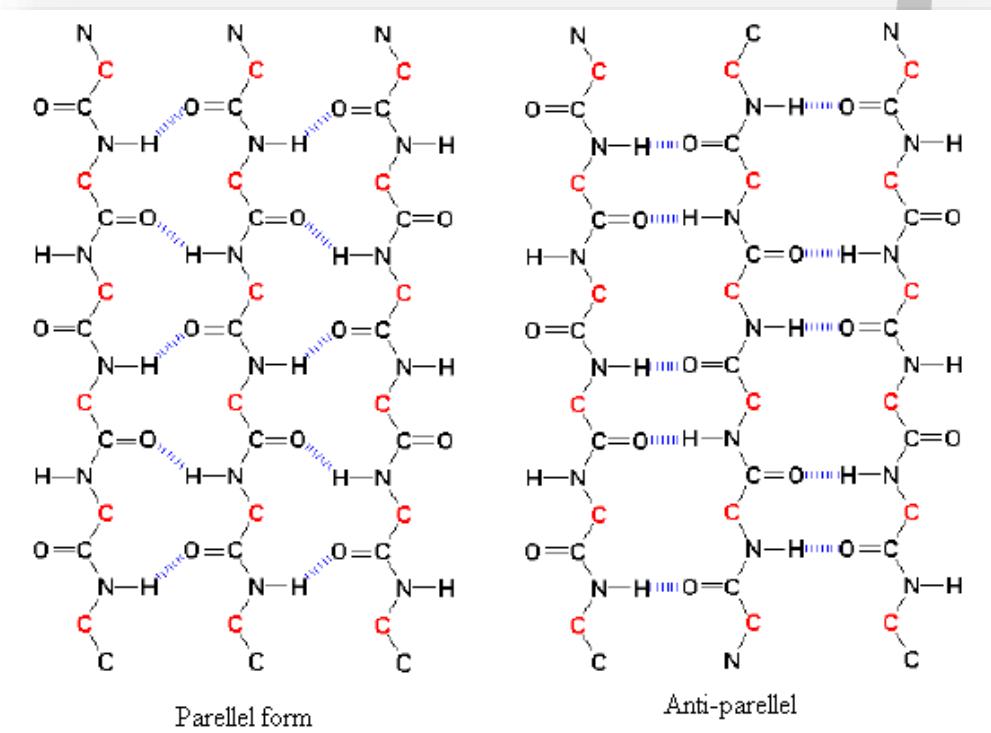


- Is a right hand-helix conformation
- Every backbone N–H group hydrogen bonds to the backbone C=O group of the amino acid located four residues earlier along the protein sequence.
- The average number of amino acids per helical turn is 3.6
- Is the most predictable from sequence, as well as the most prevalent.
- Most alpha-helices limited to a single domain, which restricts them usually to 2 nm in length.

Amin Kabil University of Technology
(Tehran Polytechnic)

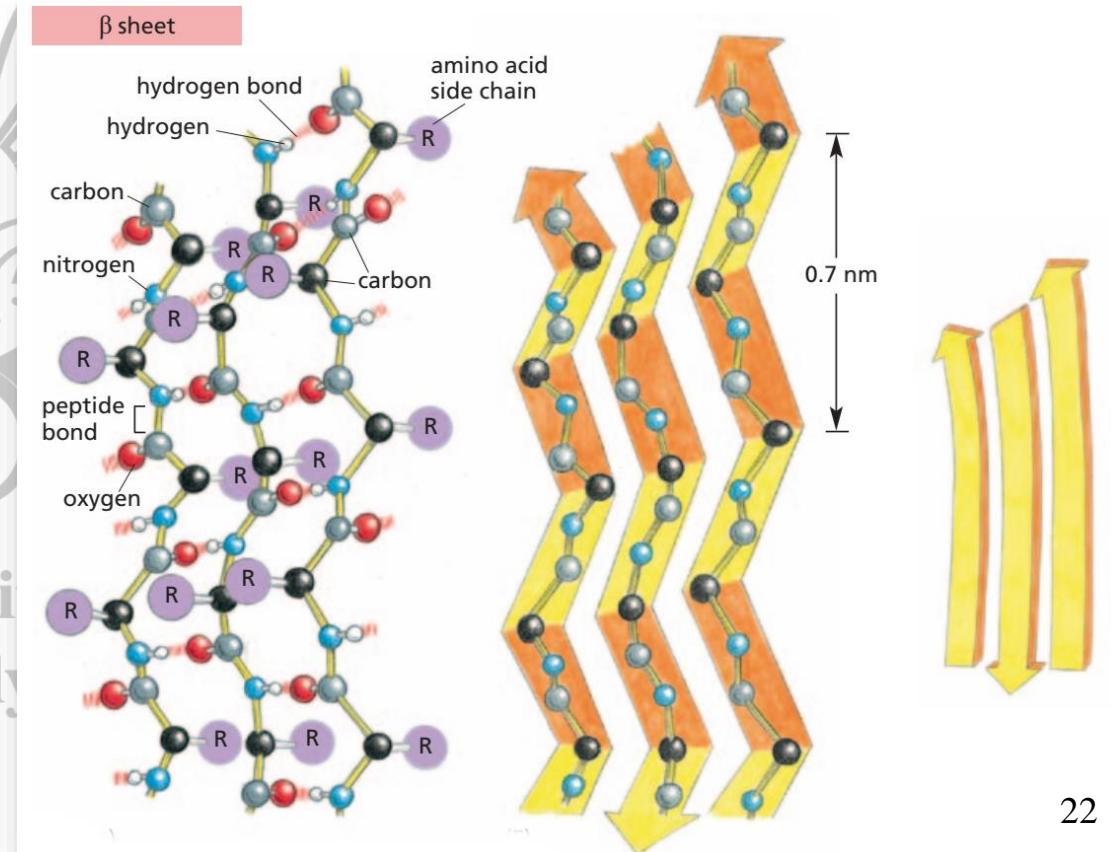
Beta Sheet (Secondary Structure)

- Beta sheets consist of *beta strands* connected laterally by at least two or three backbone hydrogen bonds, forming a generally twisted, pleated sheet.



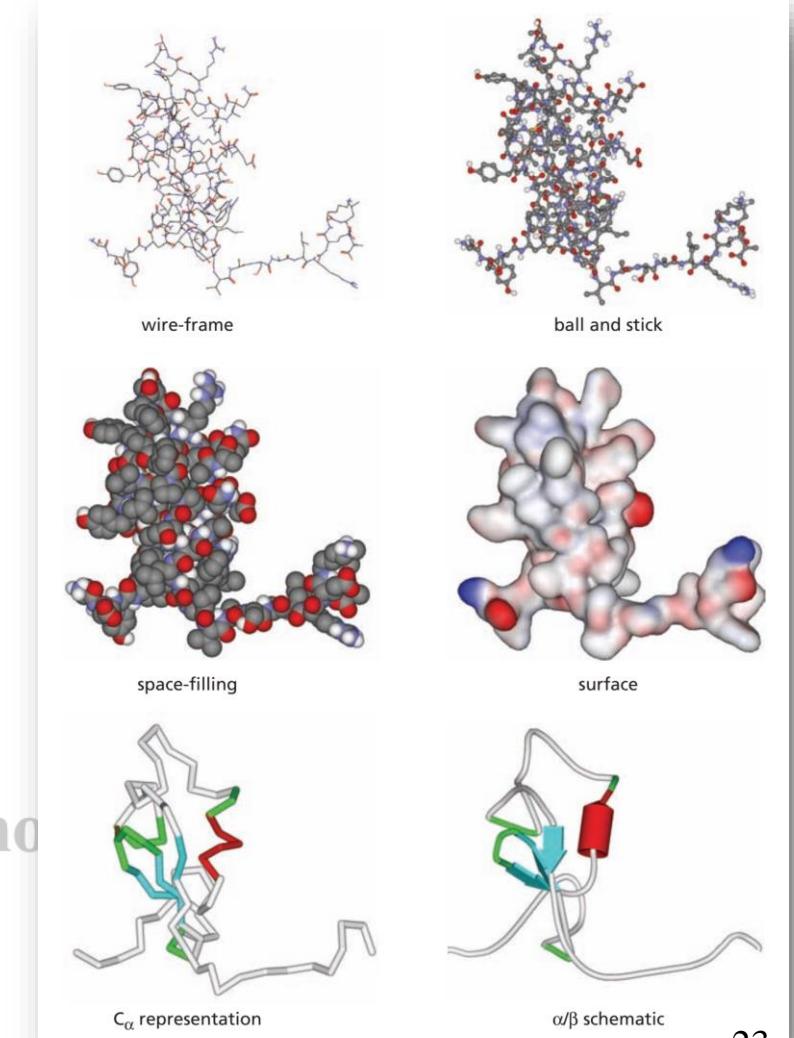
19

niversi
an Poly



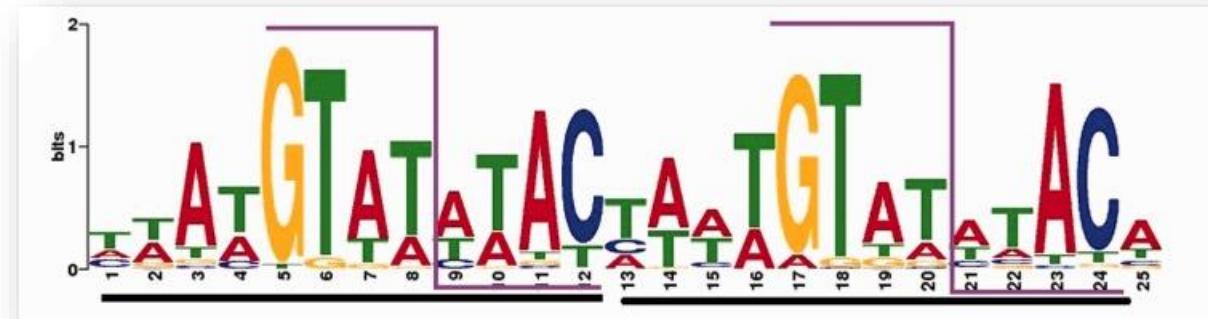
Protein Tertiary Structure

- Is a three dimensional shape of a protein.
- The tertiary structure will have a single *polypeptide chain backbone* with one or more protein secondary structures, the protein *domains*.
- The interactions and bonds of side chains within a particular protein determine its tertiary structure.
- The protein tertiary structure is defined by its atomic coordinates and save in a coordinate file.
- Visualization of protein structures: (Polytechnic)



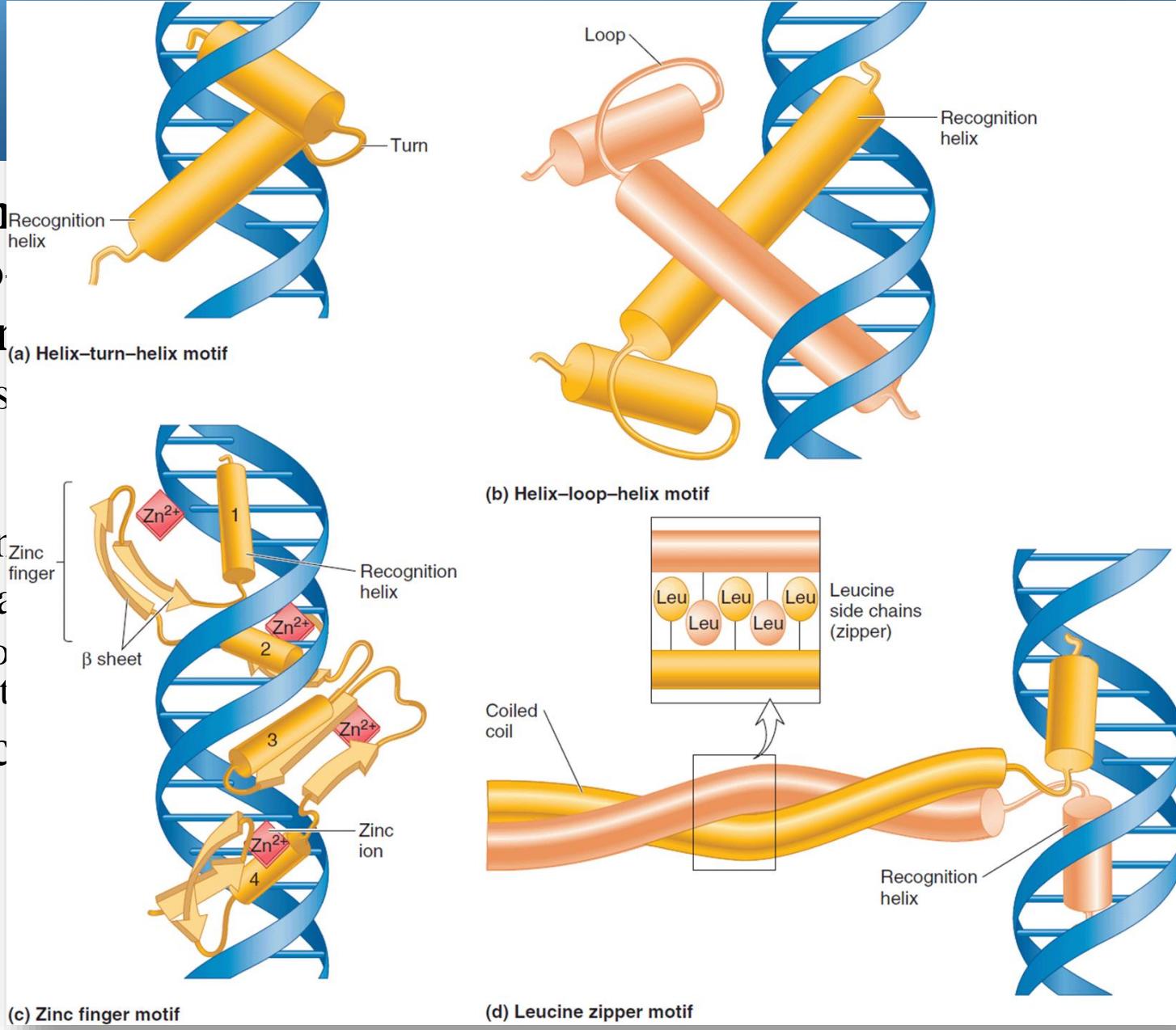
Motif

- **Sequence motif:**
 - Is a amino-acid sequence pattern that is widespread and has a biological significance.
- **Structural motif:**
 - In proteins, a structural motif describes the connectivity between secondary structural elements.
 - Structural motifs do not allow us to predict the biological functions: they are found in proteins and enzymes with dissimilar functions.
 - A structural motif does not have to be associated with a sequence motif.
 - Two biopolymers may share the same structural motif yet lack appreciable primary structure similarity.
- The existence of a sequence motif does not necessarily imply a distinctive structure.



Motif

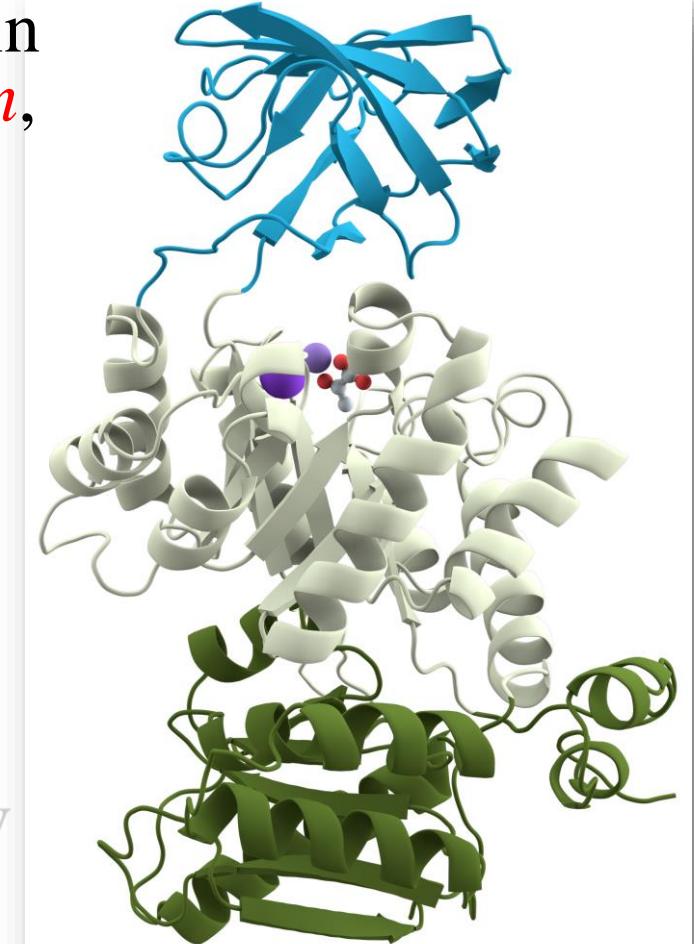
- **Sequence motif**
 - Is a amino acid sequence.
- **Structural motif**
 - In proteins and nucleic acids.
 - Structural motifs are found in proteins and nucleic acids.
 - A structural motif
 - Two biological functions with similar structures.
- The existence of a structural motif.



significance.
ary structural
are found in
ary structure
inctive

Protein Domain

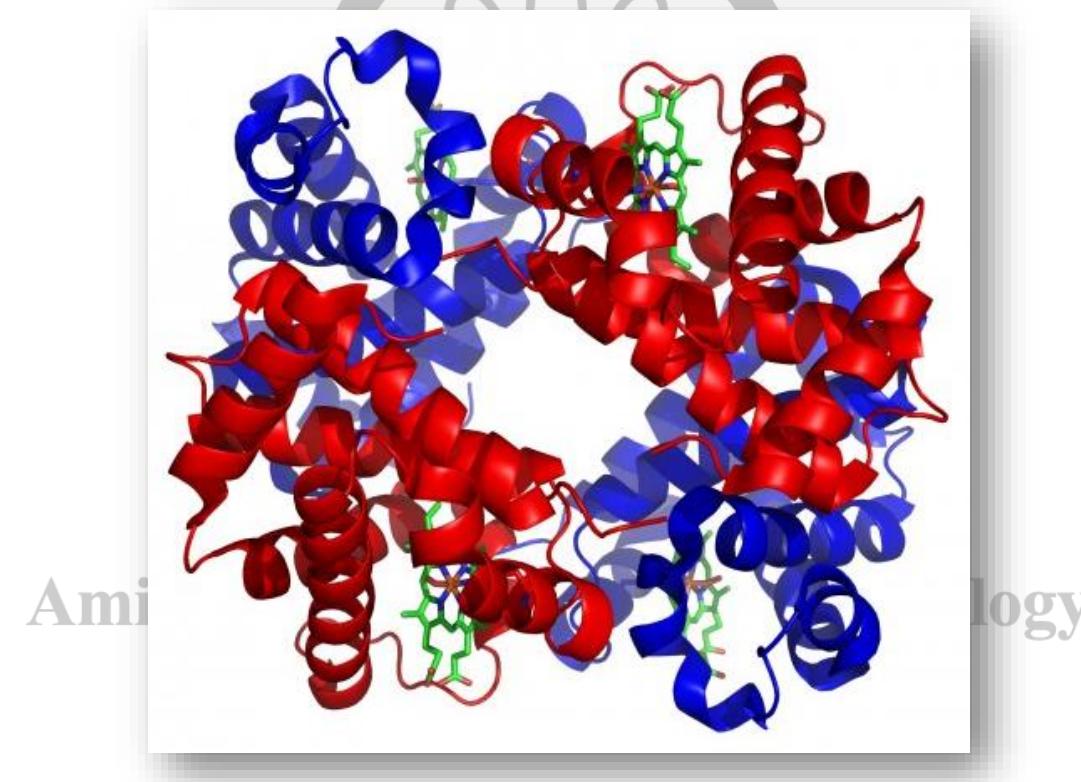
- A **protein domain** is a conserved part of a given protein sequence and tertiary structure that can *evolve, function, and exist independently* of the rest of the protein chain.
- Each domain forms a compact three-dimensional structure and often can be independently *stable*.
- One domain may appear in a variety of different proteins.
- Domains vary in length from between about 50 amino acids up to 250 amino acids in length.
- Domains often form functional units and different domains usually has different functions.
- Small proteins usually have a single domain.



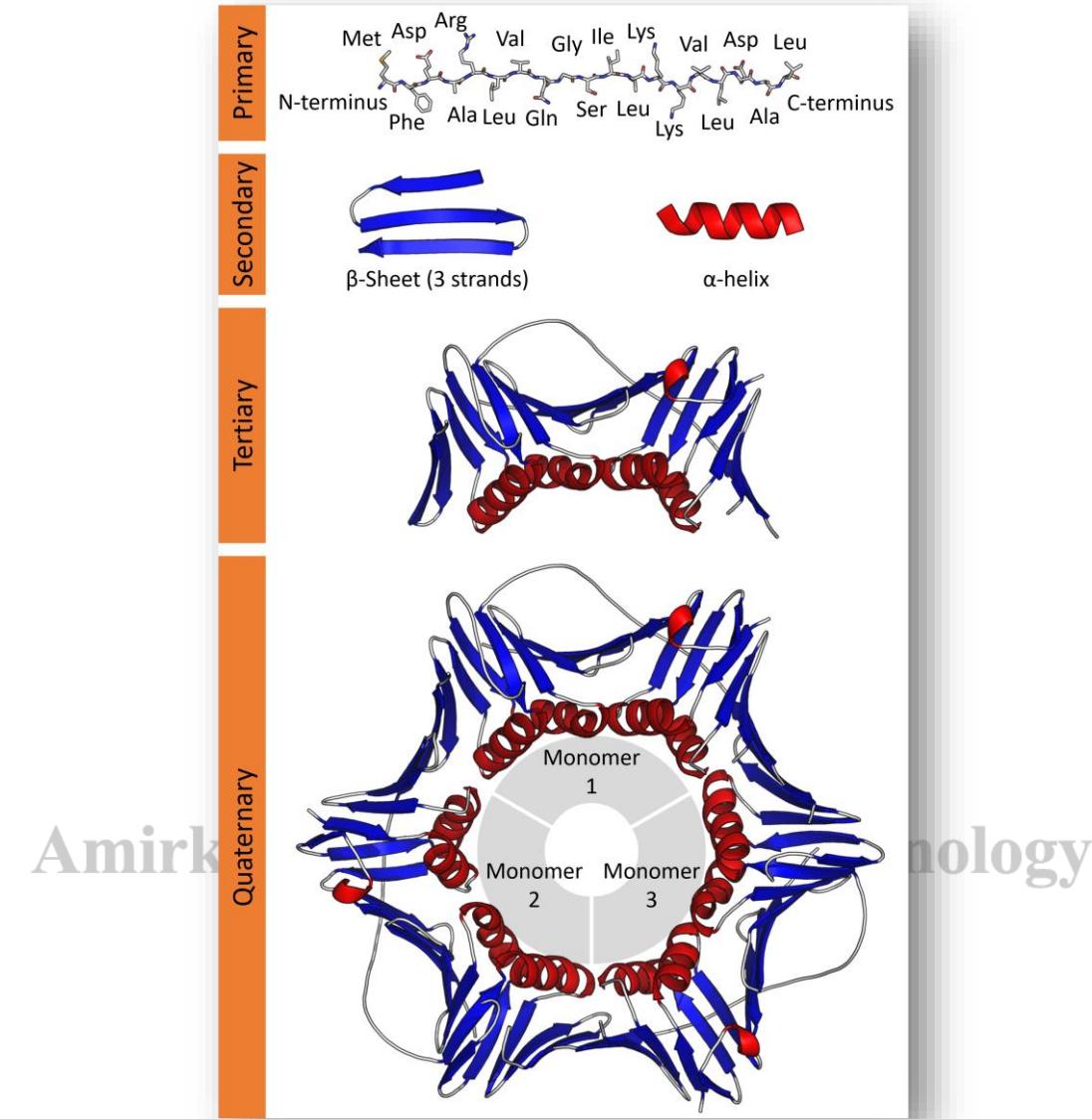
a protein with three domains

Protein Quaternary Structure

- The **quaternary structure** refers to the *number* and *arrangement* of the *protein subunits* with respect to one another.



Protein Structures: Summary



Nucleic Acids

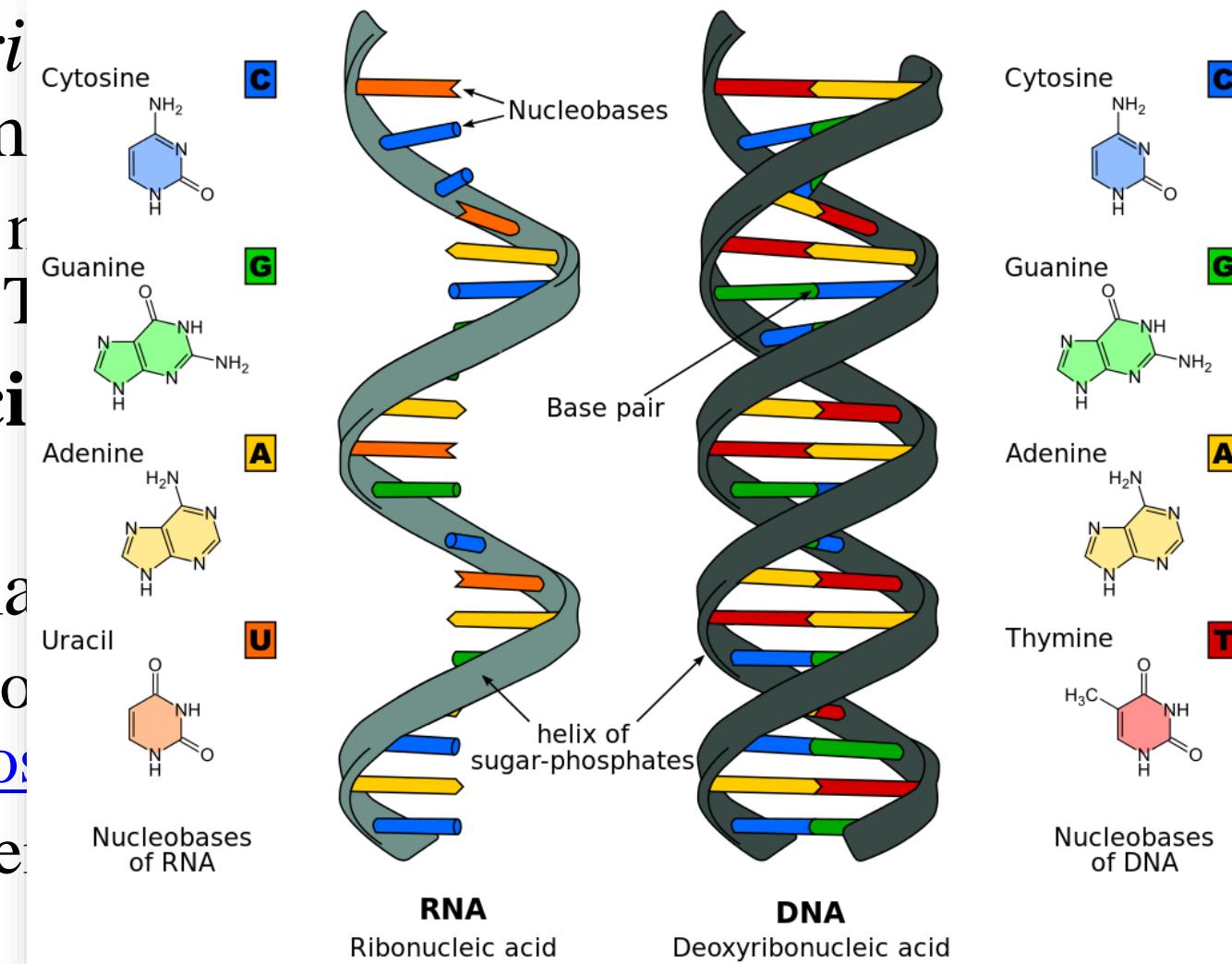
- In 1869 *Friedrich Miescher* developed ways of isolating intact *nuclei* from cells and analyzing their chemical content.
 - From the nuclei he extracted substances rich in *phosphorus* and *nitrogen*. They came to be known as "**nucleic acids**."
- **Nucleic acids** are the main *information-carrying* molecules of the cell.
- The two main classes of nucleic acids are:
 - Deoxyribonucleic acid (**DNA**): if the sugar is derived from deoxyribose
 - Ribonucleic acid (**RNA**): the sugar is a compound ribose

Amirkabir University of Technology

Tehran Polytechnic

Nucleic Acids

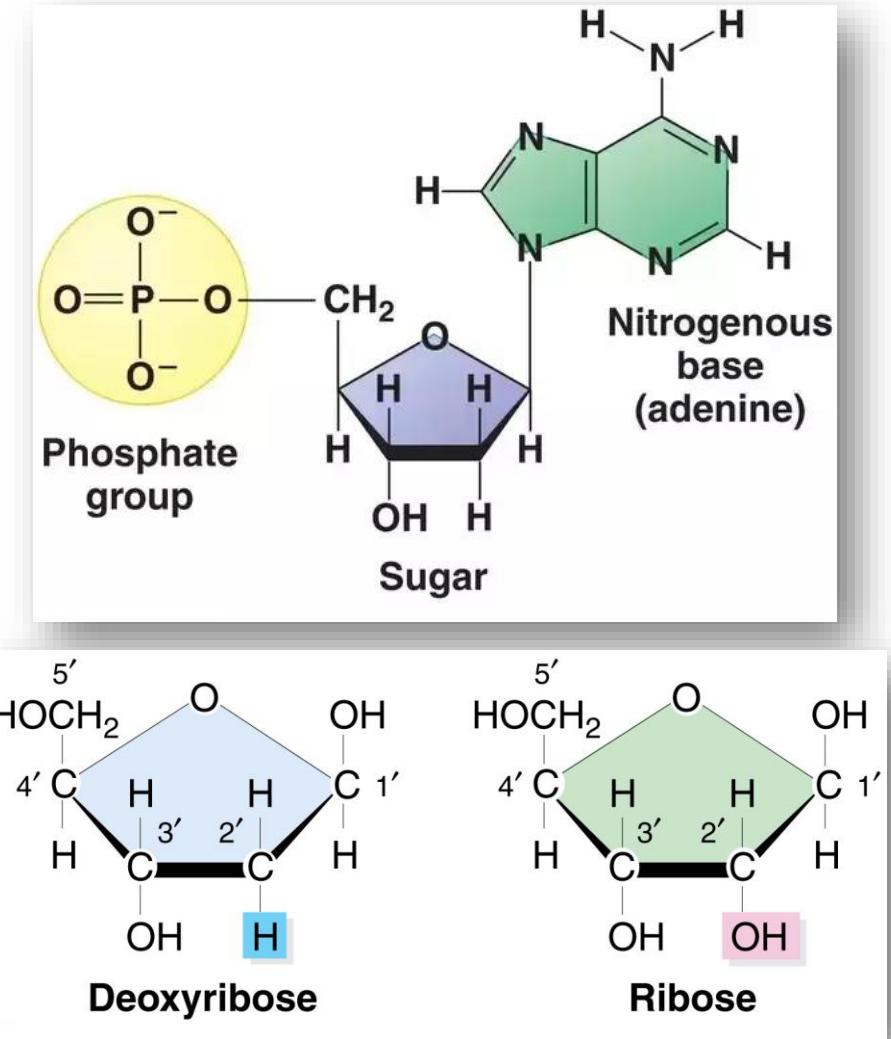
- In 1869 *Friedrich Miescher* isolated nucleic acids from white blood cells.
 - From the red blood cells he found nitrogen. This was later identified as nucleic acid.
- Nucleic acids are the genetic material of the cell.
- The two main types of nucleic acids are:
 - Deoxyribonucleic acid (deoxyribosic acid)
 - Ribonucleic acid



Nucleotide

- **Nucleotides** are molecules consisting of a *nucleoside* and a *phosphate group*. They are the basic building blocks of DNA and RNA.
- They are *organic molecules* that serve as the *monomer* units for forming the nucleic acid polymers DNA and RNA.
- They are composed of three sub unit molecules:
 - Nitrogenous base (also known as nucleobase)
 - Five-carbon sugar (ribose or deoxyribose)
 - One to three phosphate group.
- Phosphate group can bond to sugar using carbon 3' or 5'.

(Tehran Polytechnic)



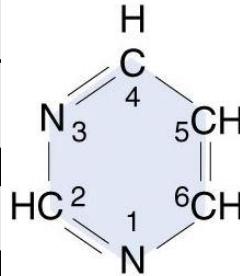
Nitrogenous Base

- A **nitrogenous base** is an *organic molecule* with a *nitrogen* atom that has the chemical properties of a *base*.
- The main biological function of a nitrogenous base is to bond nucleic acids together.
- **Pyrimidine:** single ring structure
 - Thymine (T) (**Only in DNA**), Cytosine (C), and Uracil (U) (**Only in RNA**)
- **Purine:** double ring structure
 - Adenine (A) and Guanine (G)

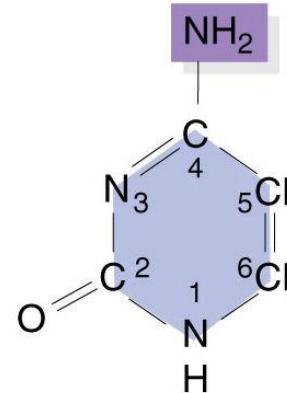
Nitrogenous Base

- A nitrogenous base is a molecule containing a nitrogen atom that is bonded to one or more hydrogen atoms.
- The main function of nitrogenous bases is to bond to phosphate groups in nucleic acids.

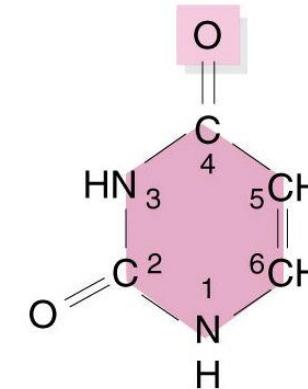
- **Pyrimidines**
 - Thymine (T) found in DNA
 - Cytosine (C) found in RNA
 - Uracil (U) found in RNA



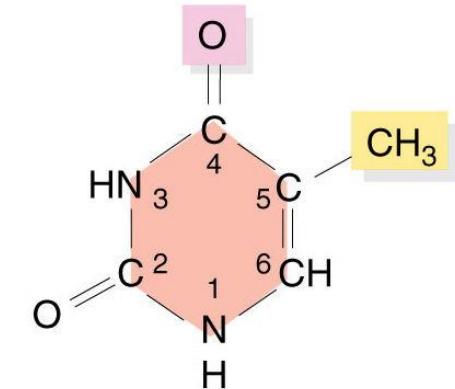
Pyrimidine



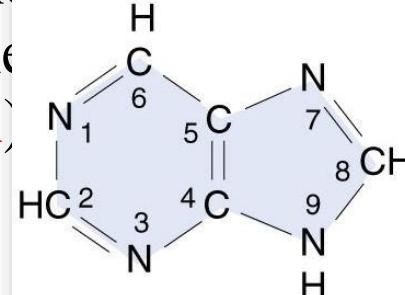
Cytosine (C)



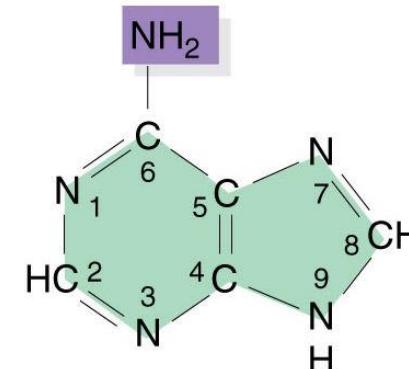
Uracil (U)
(found in RNA)



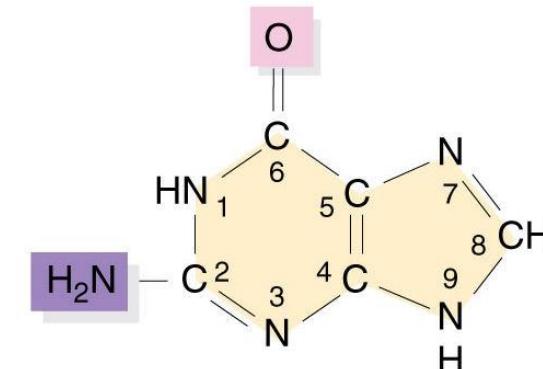
Thymine (T)
(found in DNA)



Purine



Adenine (A)

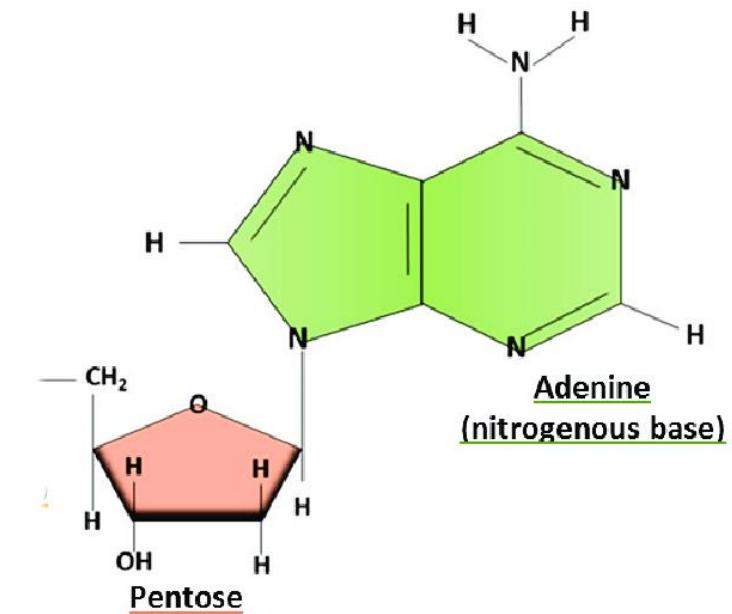


Guanine (G)

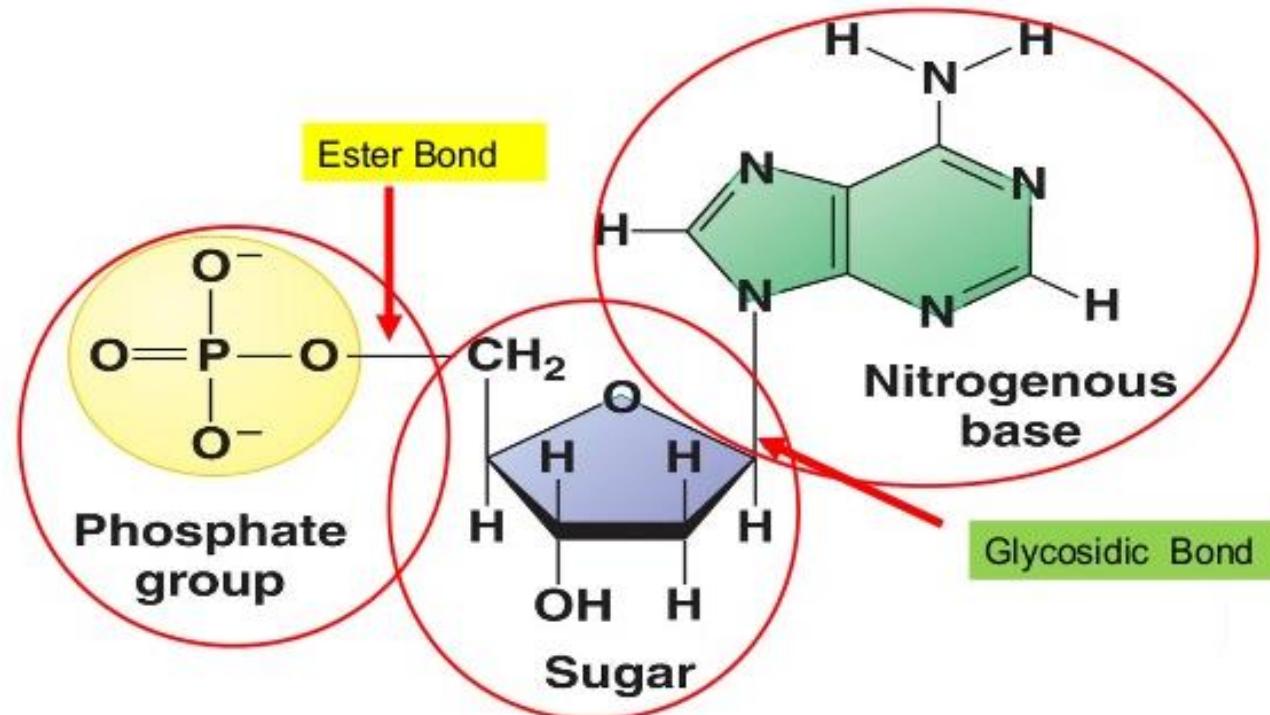
Hydrogen
to bond
structure
nine (G)

Nucleosides

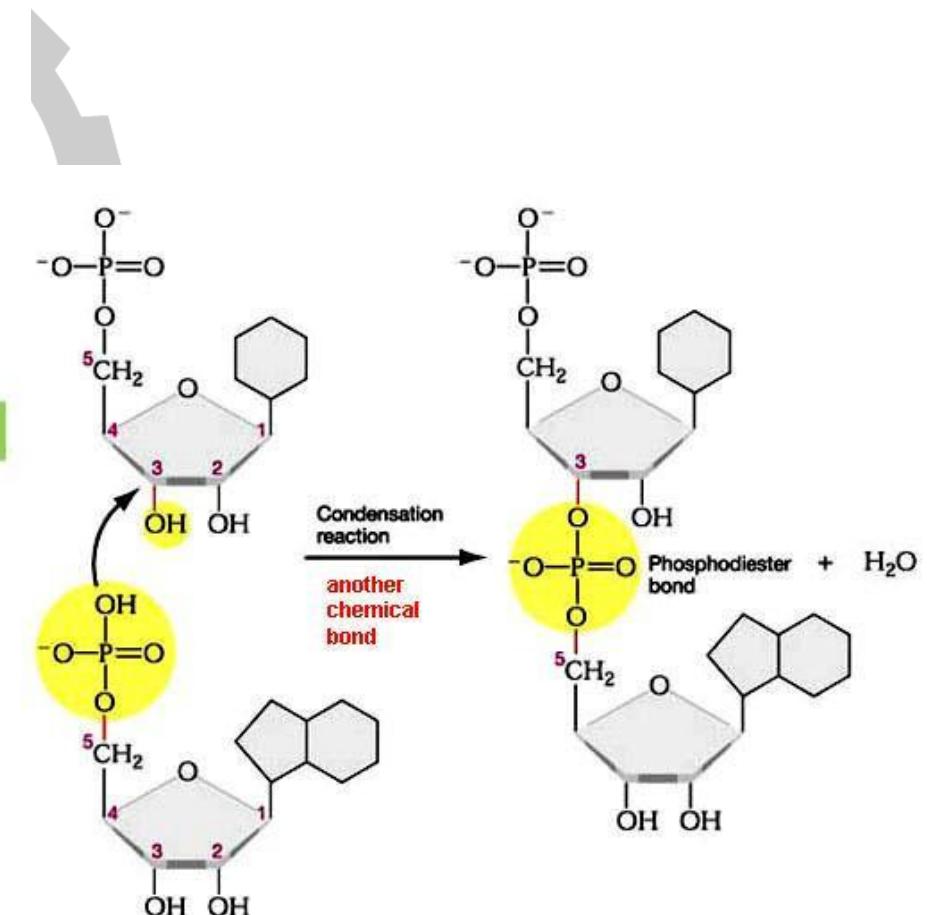
- **Nucleosides** are nucleotides without a phosphate group.
 - Consists simply of a *nitrogenous base* and a *five-carbon sugar ribose*.
- In a nucleoside, the *anomeric carbon* is linked through a glycosidic bond to the N9 of a *purine* or the N1 of a *pyrimidine*.
- **Adenosine:** composed of a molecule of *adenine base* attached to a *ribose sugar* molecule.
- **Deoxyadenosine:** composed of a molecule of *adenine base* attached to a *deoxyribose sugar* molecule.



Type of Bonds

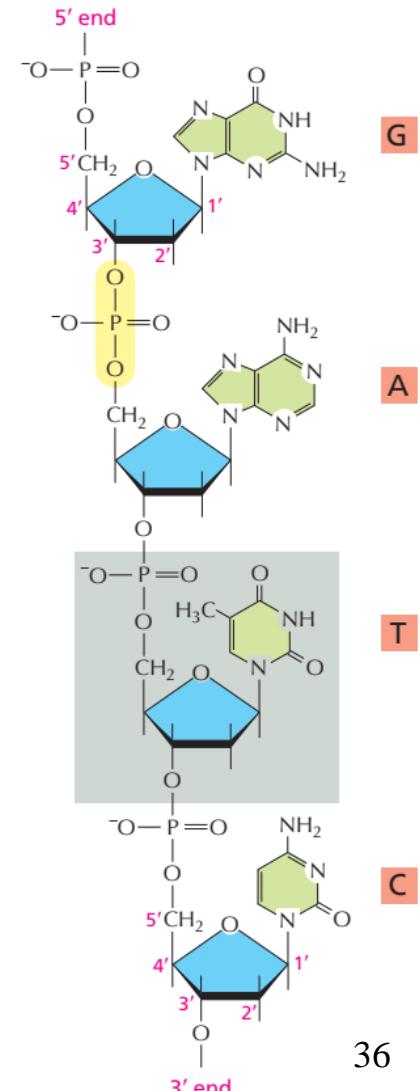


Amirkabir University
(Tehran Poly)

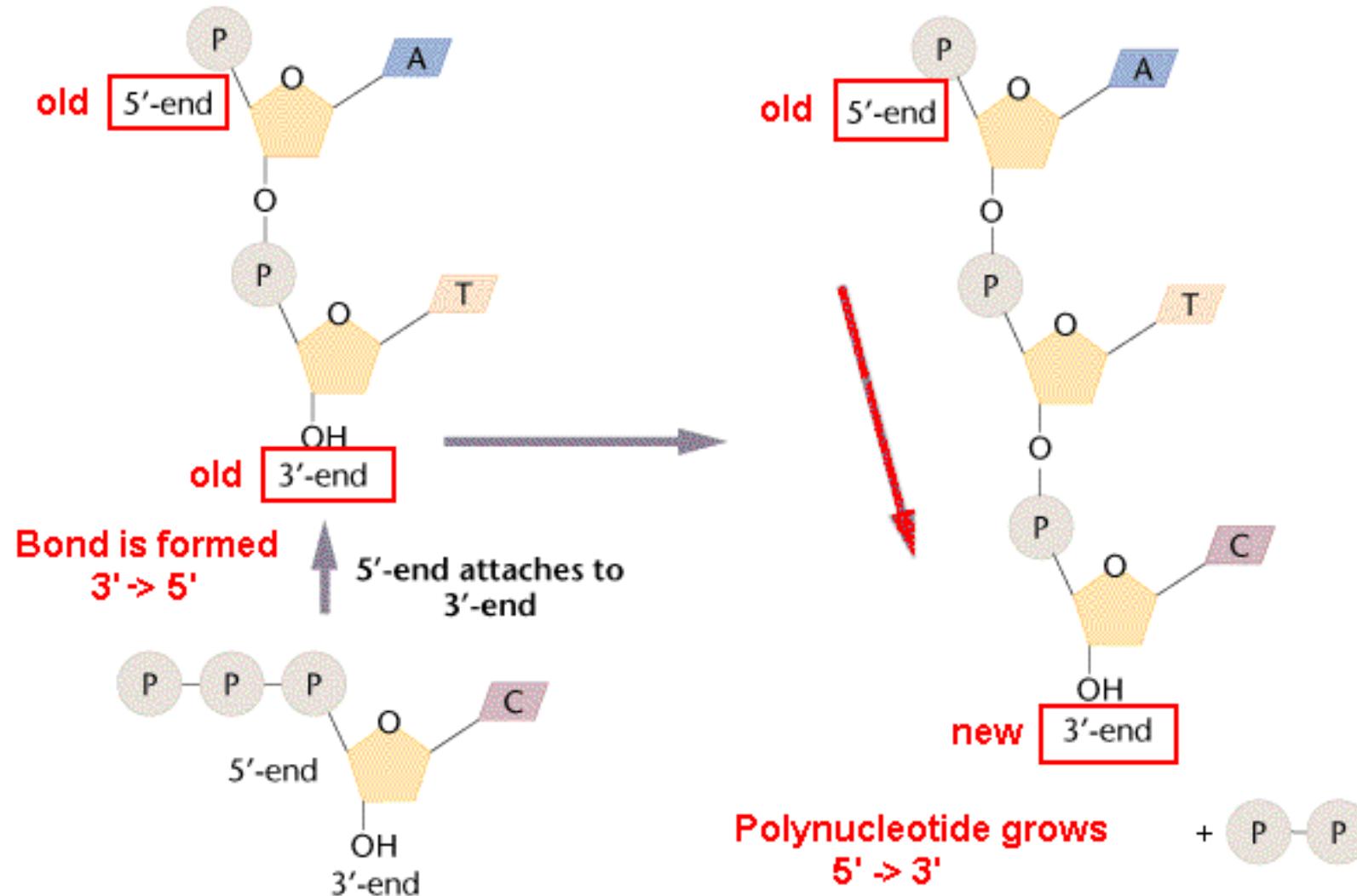


Polynucleotide

- The bonds link specific carbon atoms in the sugar ring (known as the **5' and 3'** atoms) one end of a polynucleotide chain, the 5' end, has a free phosphate group and the other, the 3' end, has a free hydroxyl group.
- The linear sequence of nucleotides in a polynucleotide chain is commonly abbreviated by a one-letter code, and the sequence is always read from the 5' end.
- In the example illustrated, the sequence is **GATC**.

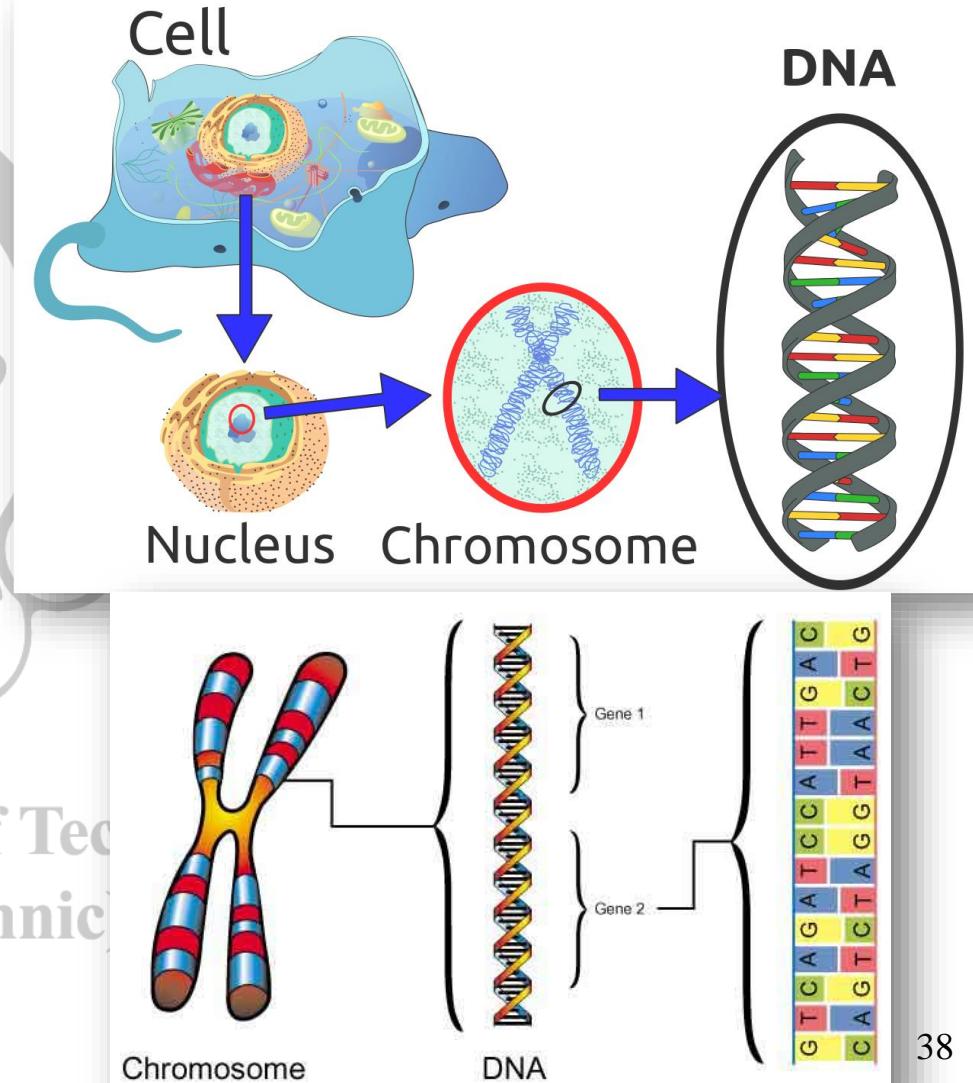


Adding a Nucleotide to Polynucleotide



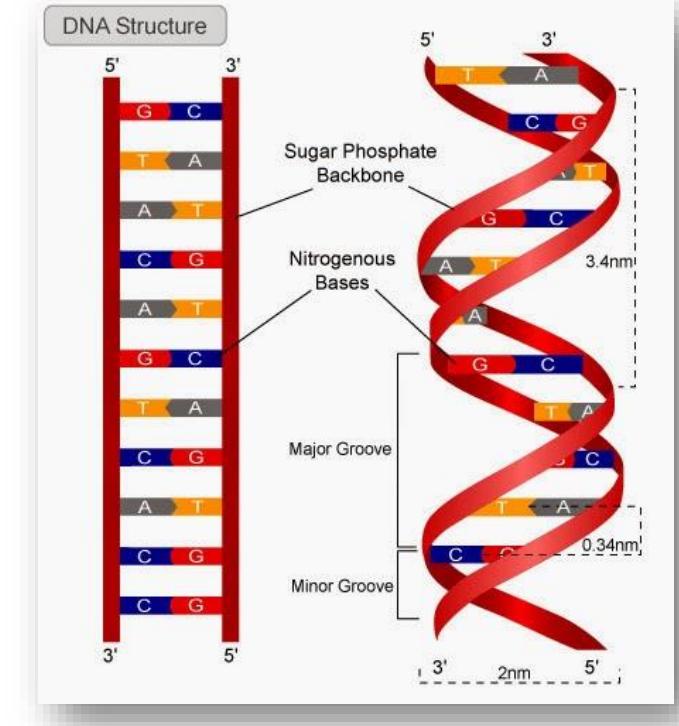
Chromosome

- A **chromosome** is a DNA molecule with part or all of the genetic material (genome) of an organism.
- A gene is a region of DNA that encodes function (genetic information).
- A human chromosome can have up to 500 million base pairs of DNA with thousands of genes.



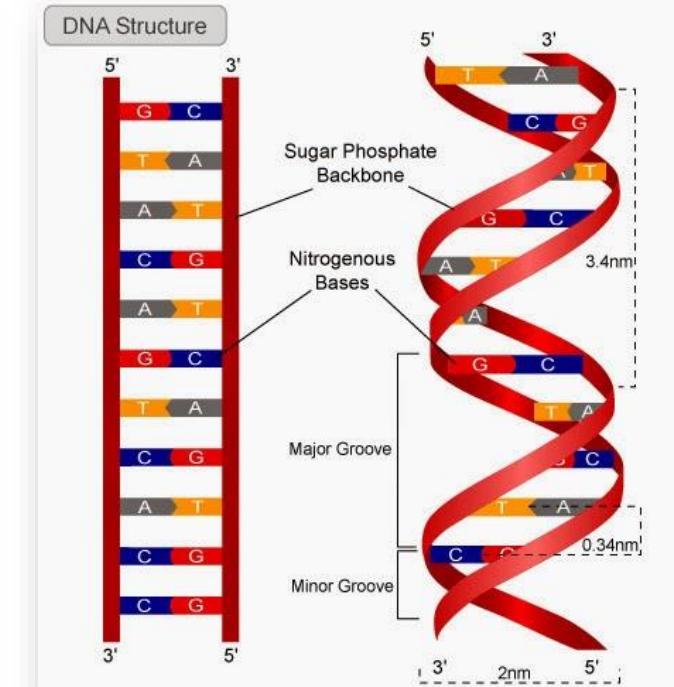
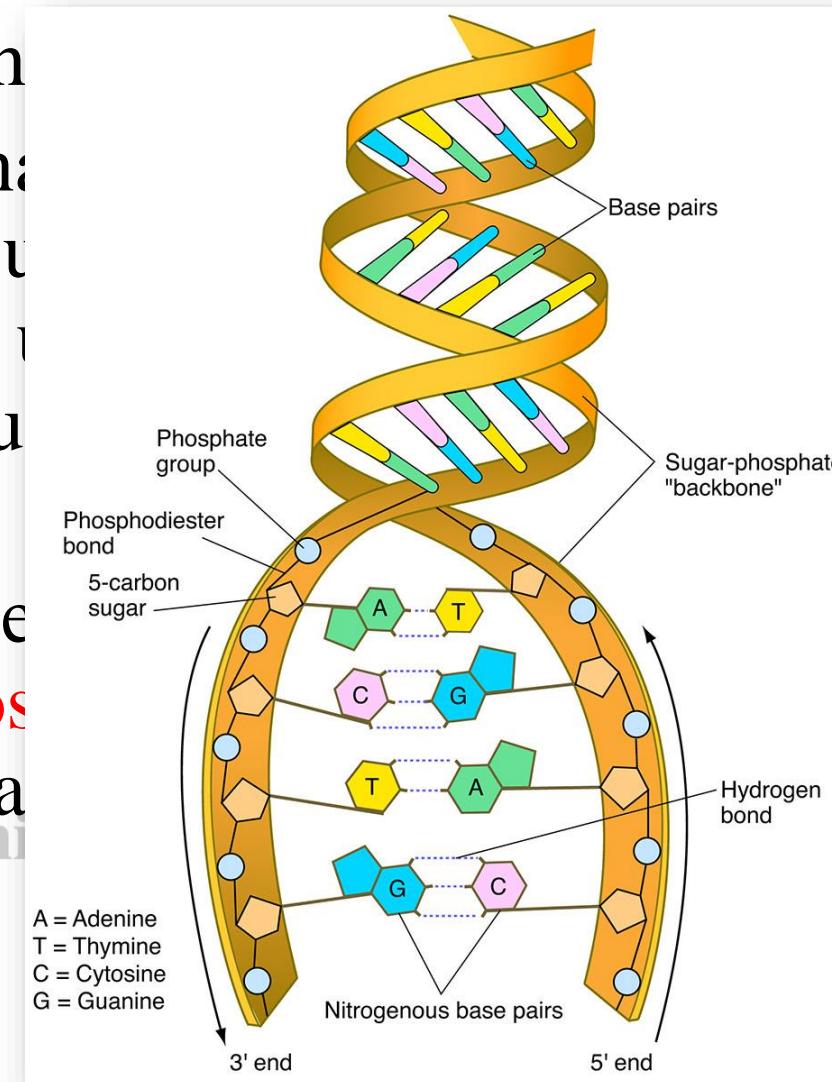
Watson and Crick's Model of DNA

- DNA has a double helix model.
- Chargaff's rule: in natural DNA the number of guanine units equals the number of cytosine units and the number of adenine units equals the number of thymine units.
- In Watson–Crick there are two base pairs: **guanine–cytosine**, and **adenine–thymine** which always appear in front of each other.



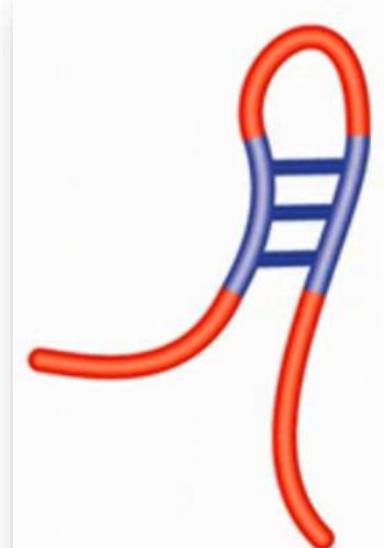
Watson and Crick's Model of DNA

- DNA has a double helix structure.
- Chargaff's rule: in naturally occurring DNA, the number of guanine units equals the number of cytosine units, and the number of adenine units equals the number of thymine units.
- In Watson–Crick the base pairs are: **guanine–cytosine**, **adenine–thymine** which always pair with each other.

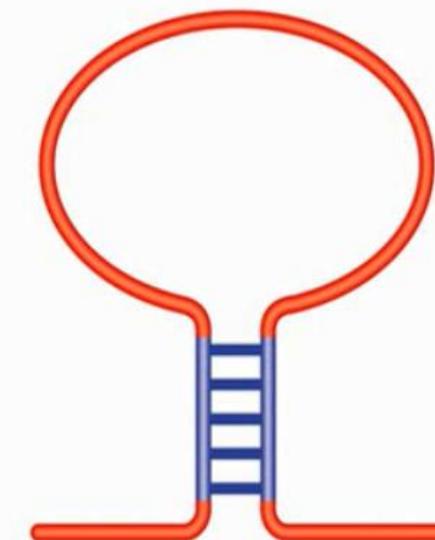
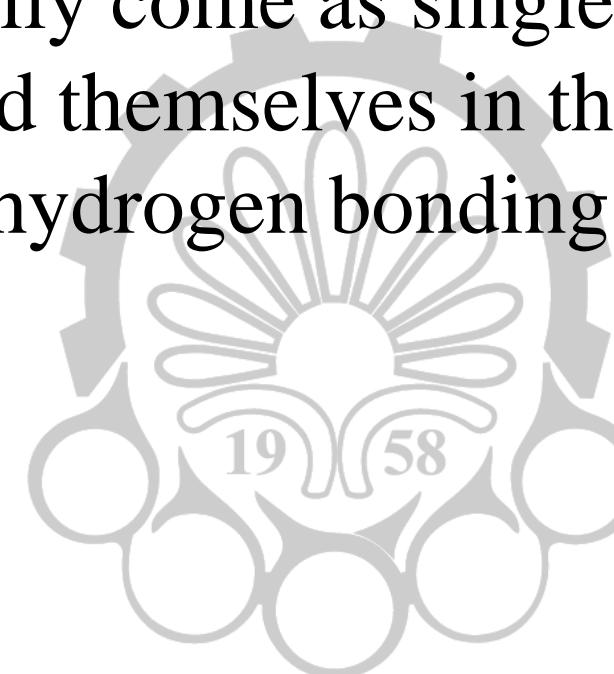


RNA Secondary Structure

- RNA molecules usually come as single strands but left in their environment they fold themselves in their *tertiary structure* because of the same hydrogen bonding mechanism.



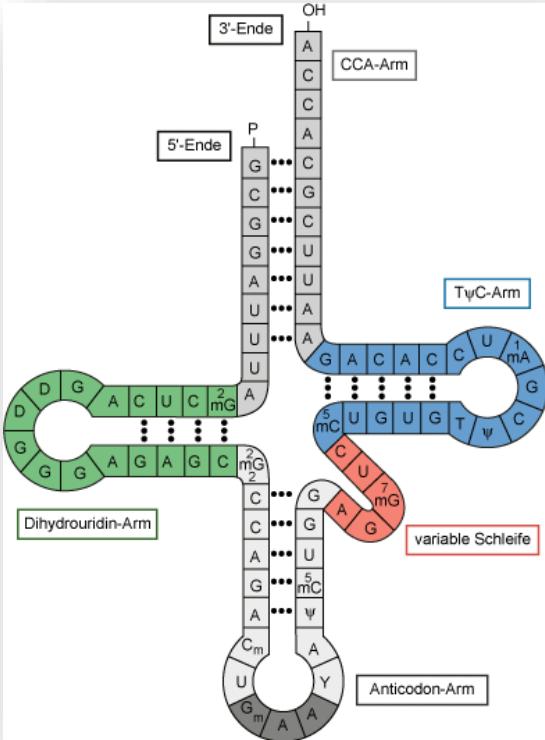
Hairpin



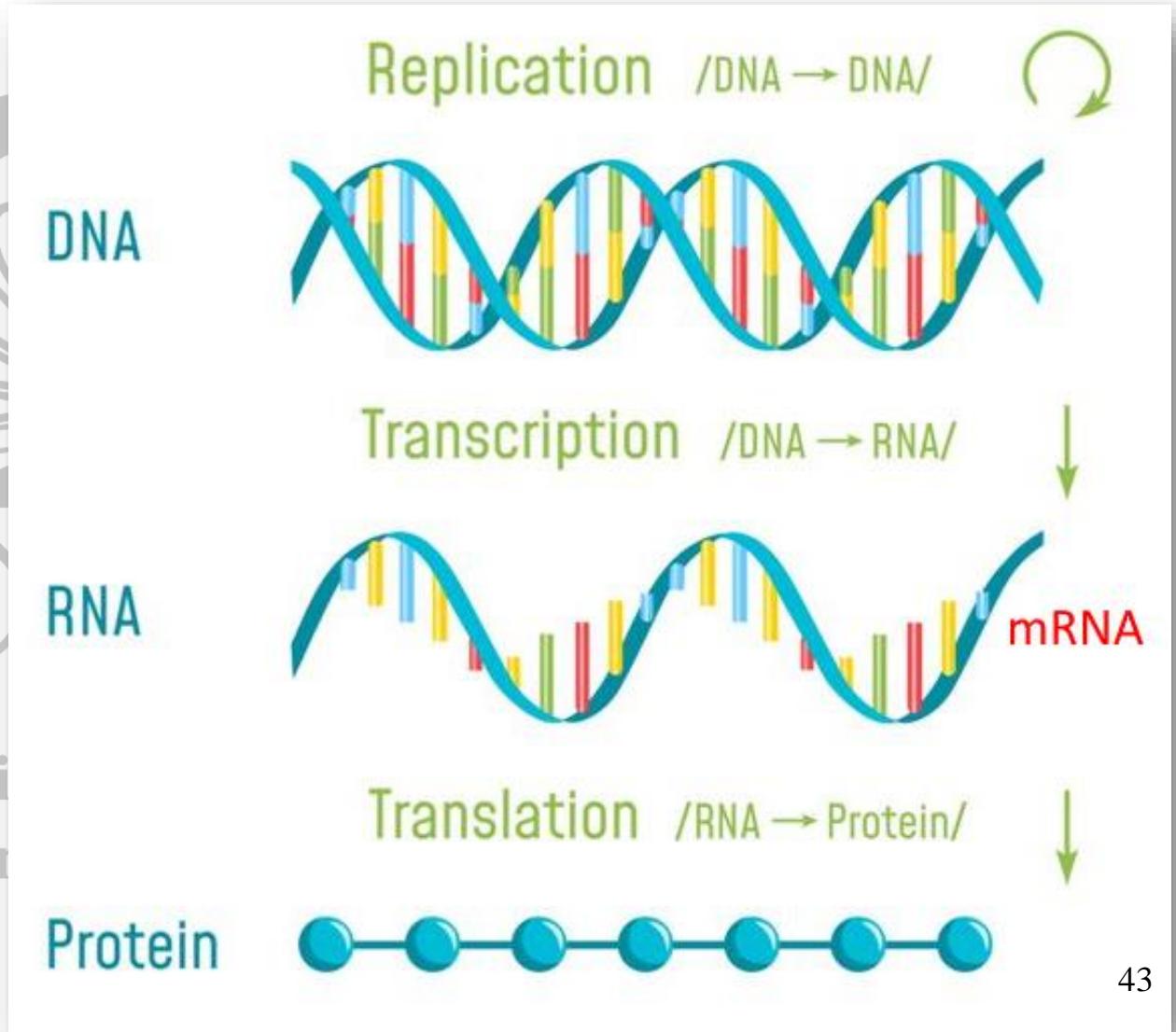
Stem-loop

Main RNA Types

- **mRNA:** Messenger RNA
- **tRNA:** Transfer RNA
- **rRNA:** Ribosomal RNA

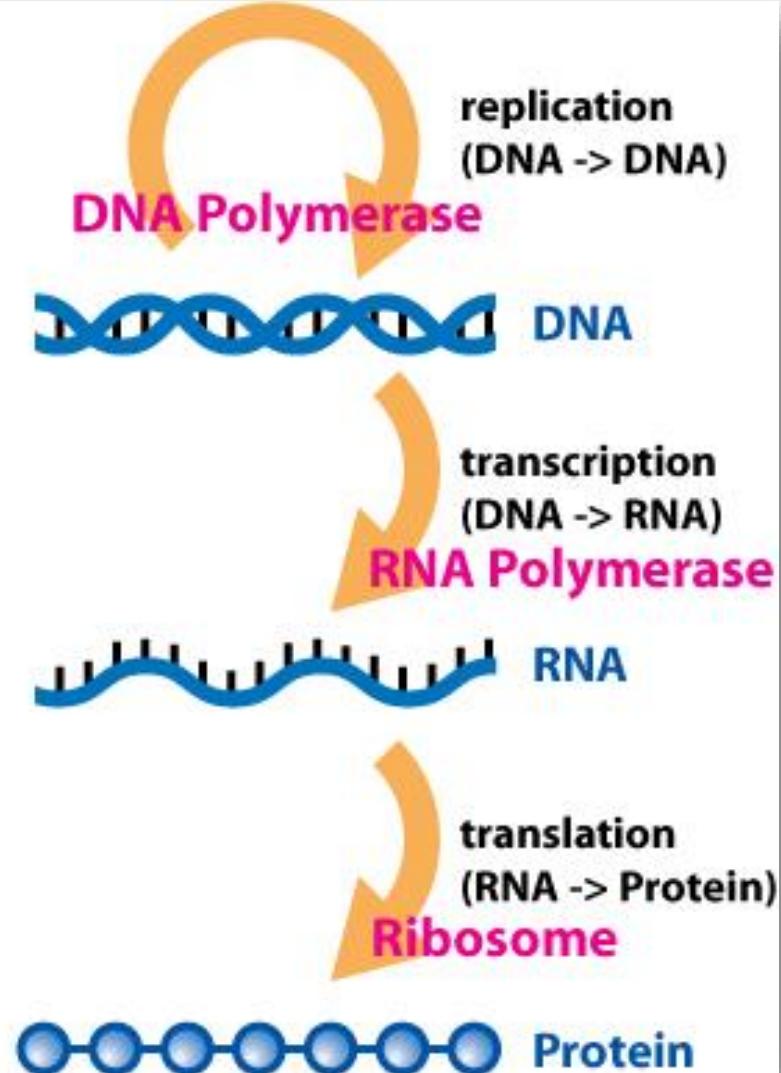
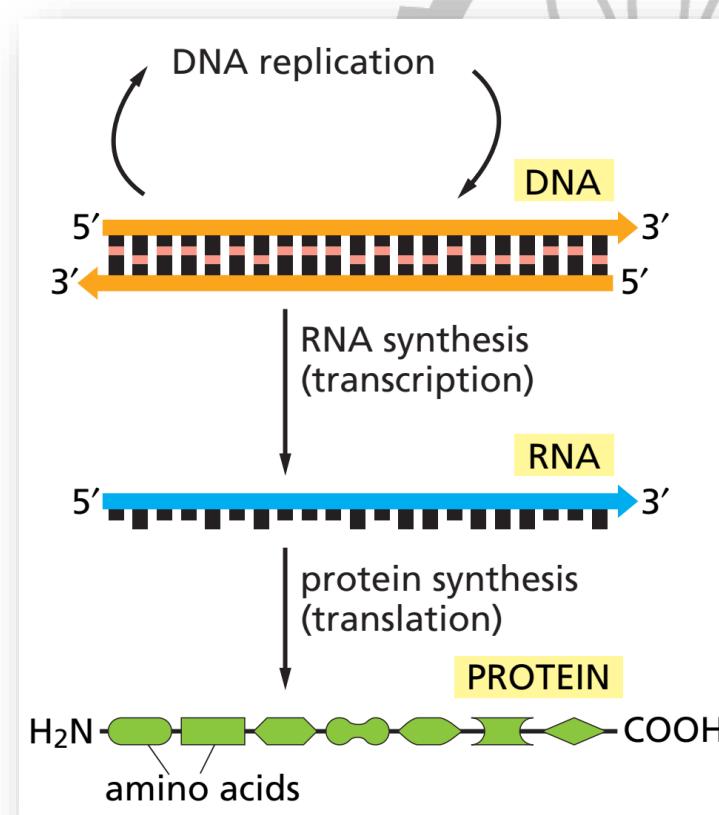


Amirkabir Uni
(Tehran)



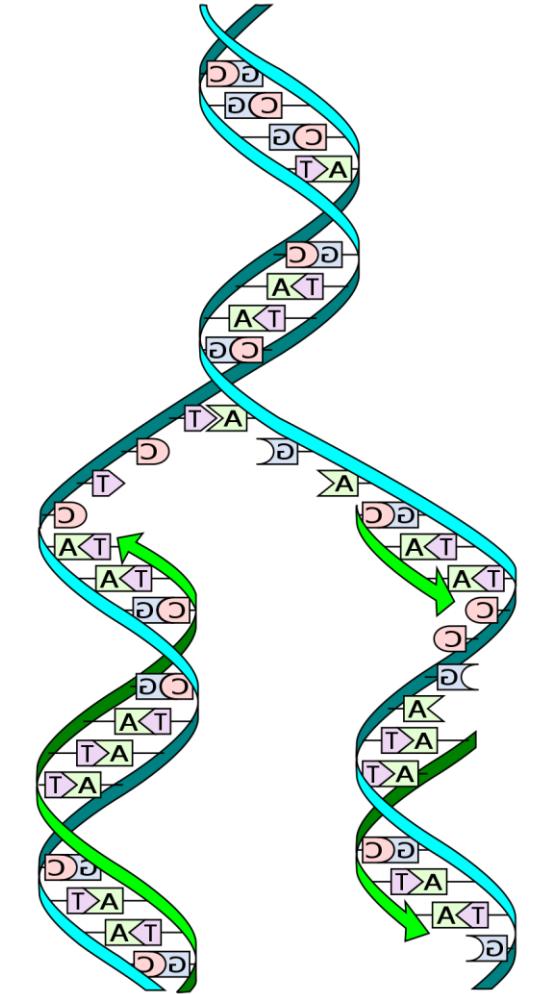
Central Dogma

- The **central dogma** is an explanation of the flow of genetic information within a biological system.

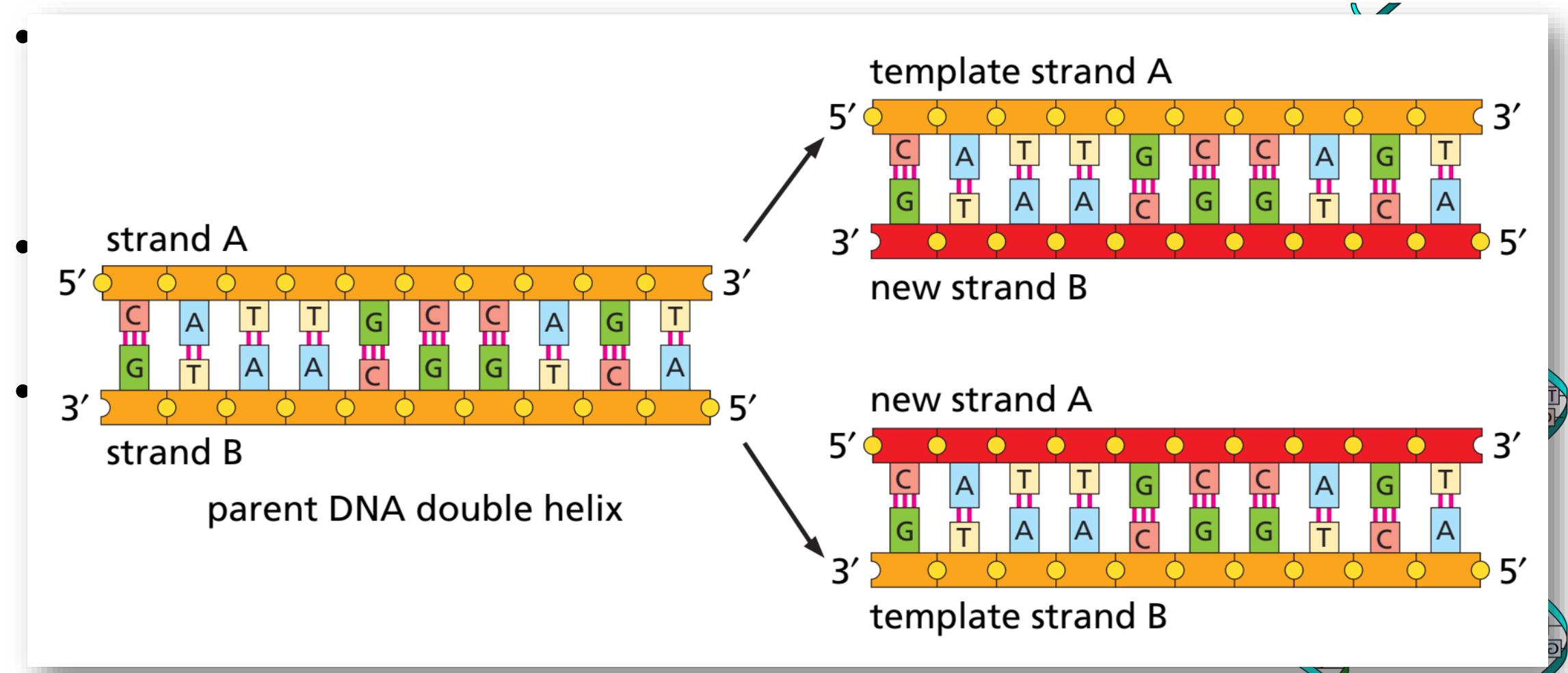


DNA Replication

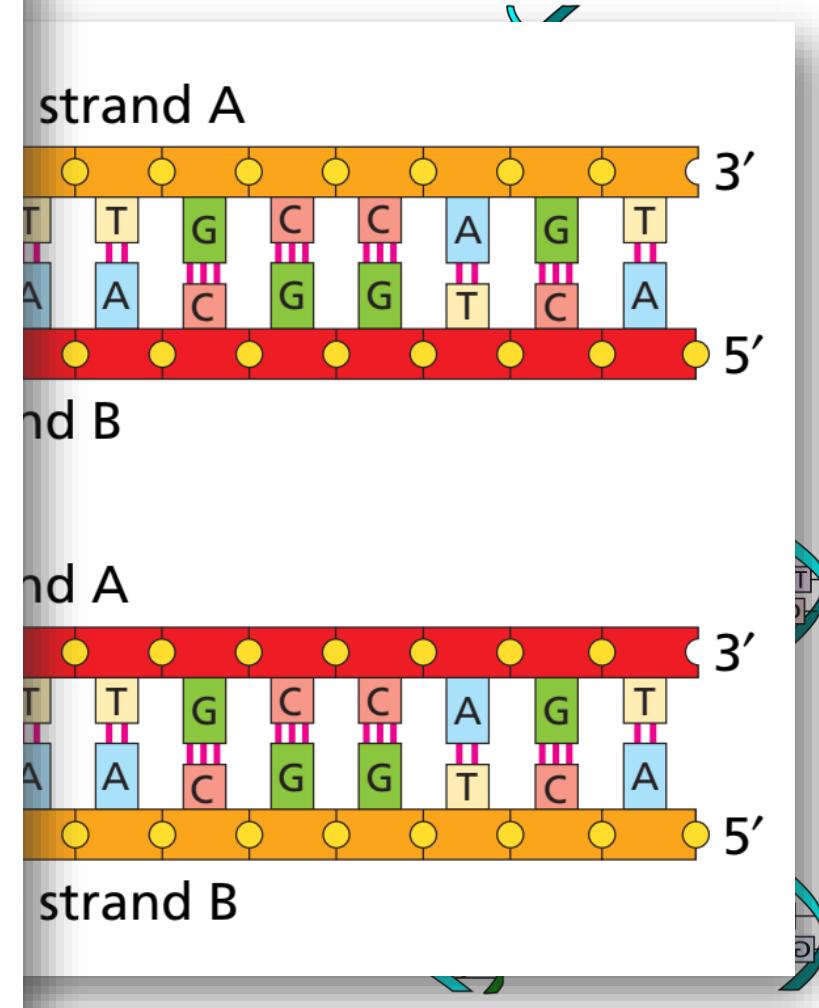
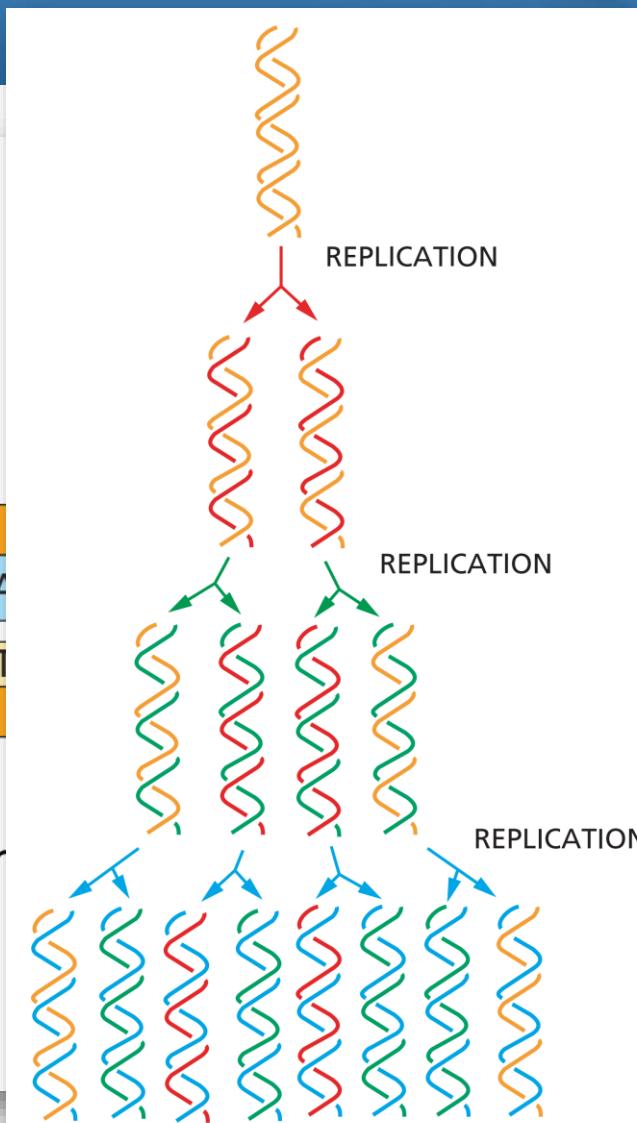
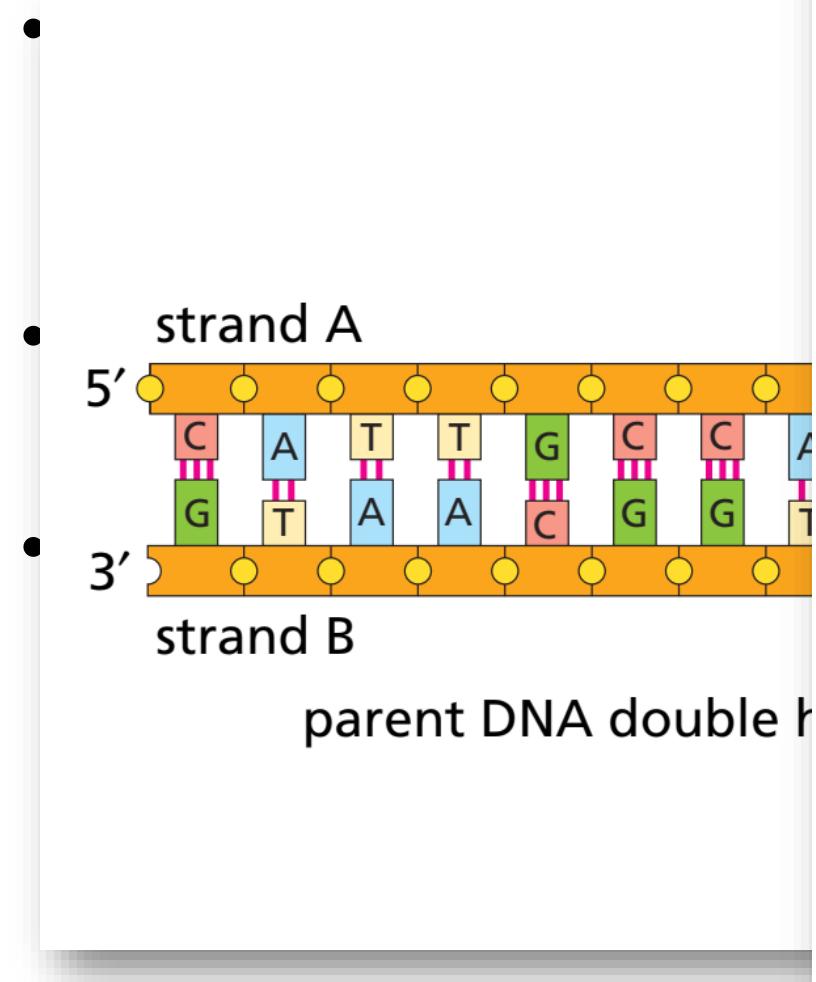
- **DNA replication** is the *biological process* of producing *two identical replicas* of DNA from one original DNA molecule.
- DNA replication occurs in all *living organisms* acting as the basis for *biological inheritance*.
- *Semiconservative replication*: two complementary strands of a DNA are separated. Each strand of the original DNA molecule then serves as a template for the production of its counterpart.



DNA Replication

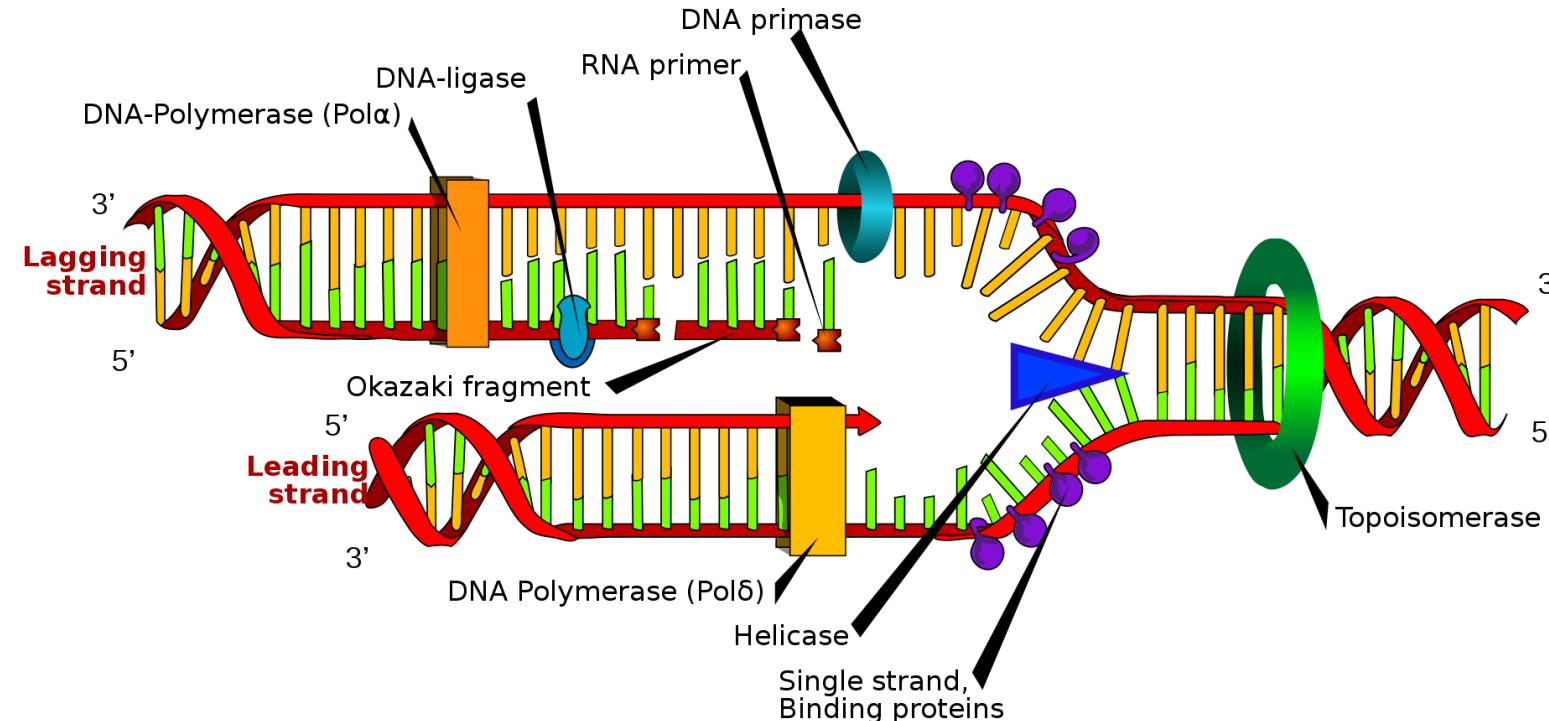


DNA Replication



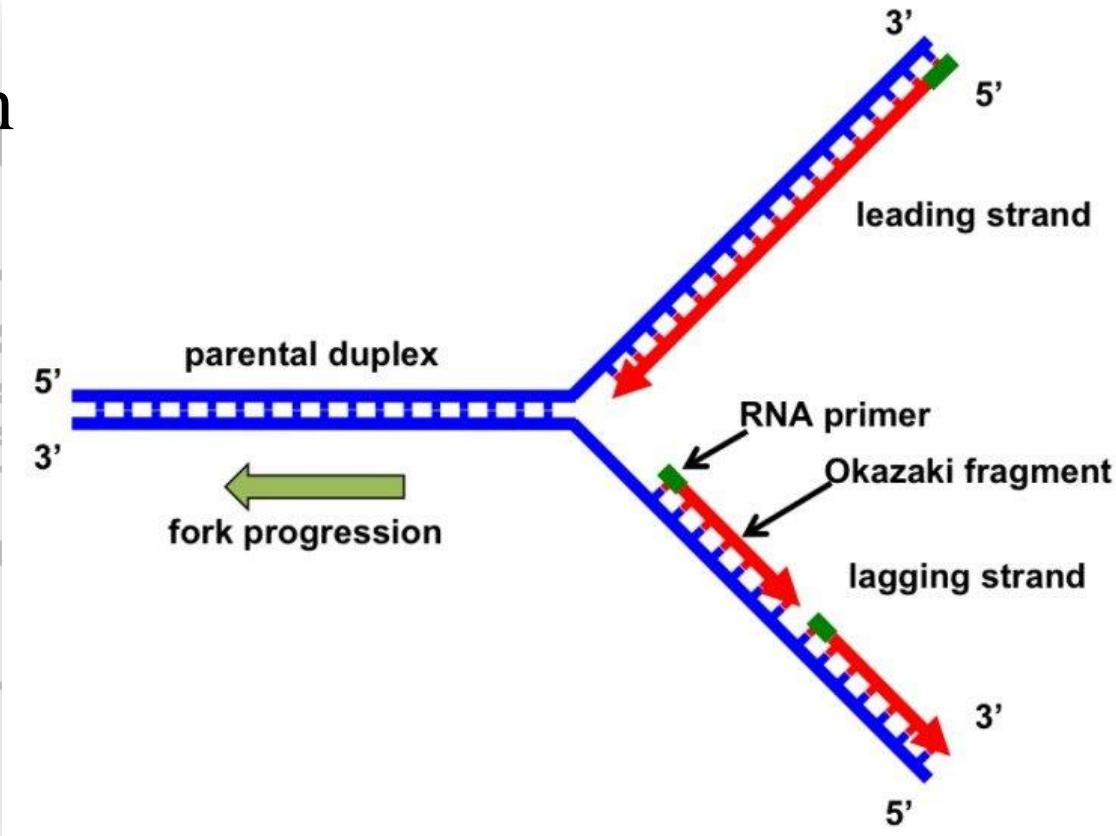
Helicases and DNA-Polymerase Enzymes

- **Helicase** breaks the hydrogen bonds holding the two DNA strands together in the helix.
- **DNA polymerases** are a family of enzymes that carry out all forms of DNA replication and extend an existing DNA or RNA strand.



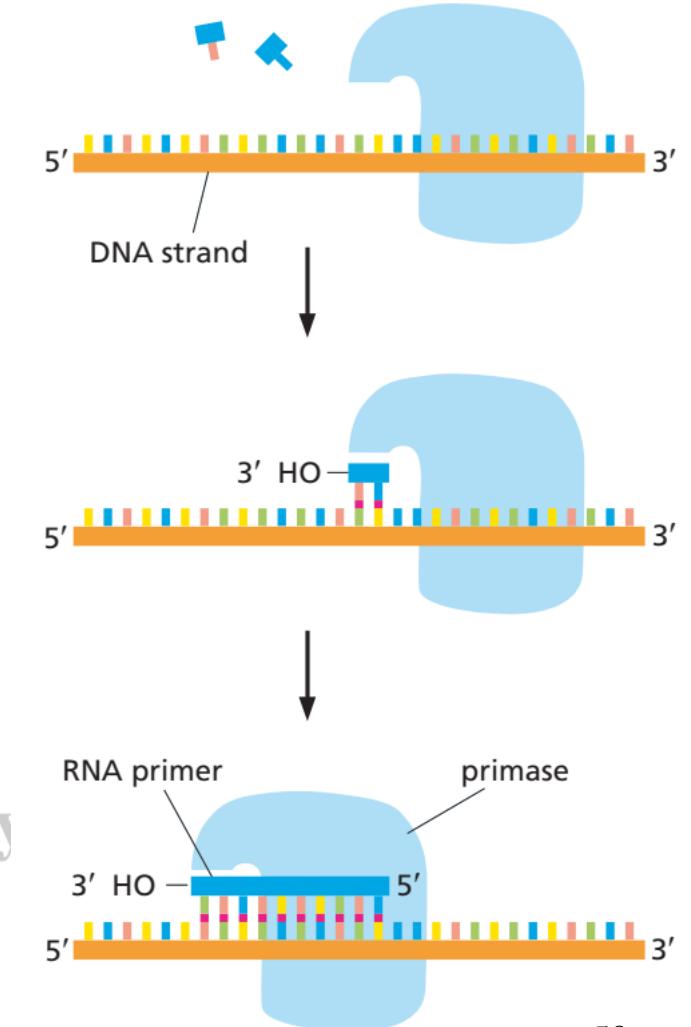
Leading and Lagging Strands

- **Leading strand** is the strand of nascent DNA which is synthesized in the *same direction* as the growing replication fork. This sort of DNA replication is *continuous*.
- **Lagging strand** is the strand of nascent DNA whose direction of synthesis is *opposite* to the direction of the growing replication fork. It is synthesized in short, separated segments and is *discontinuous*.



RNA Primer

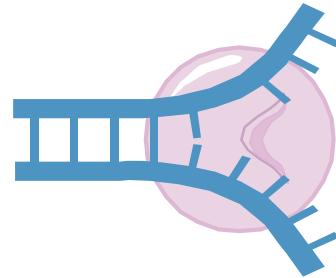
- To begin synthesis, a short fragment of RNA, called a **primer**, must be created and paired with the template DNA strand.
- Primase** catalyzes the synthesis of a short RNA segment called a primer.
- Primase is activated by the helicase where it then synthesizes a short RNA primer approximately 11 ± 1 nucleotides.
- DNA polymerase** attaches onto the 3' end of the RNA primer.



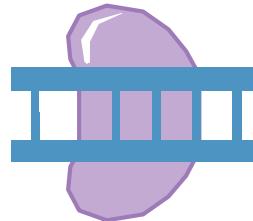
Enzymes in DNA Replication

- *Helicase*: enzyme that breaks H-bonds
- *DNA Polymerase*: enzyme that catalyzes connection of nucleotides to form complementary DNA strand in 5' to 3' direction (reads template in 3' to 5' direction)
- *DNA Primase*: enzyme that catalyzes formation of RNA starting segment (RNA primer)
- *DNA Ligase*: enzyme that catalyzes connection of two Okazaki fragments.

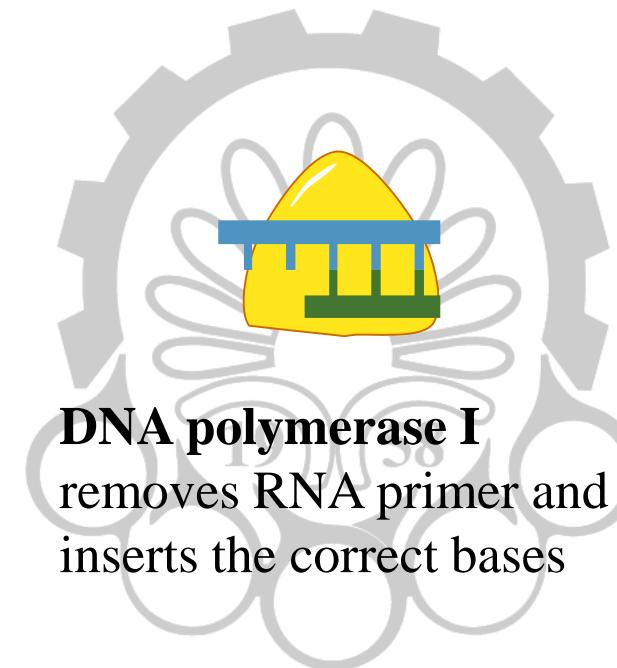
Enzymes in DNA Replication



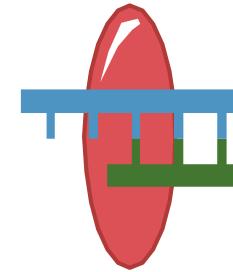
Helicase unwinds
parental double helix



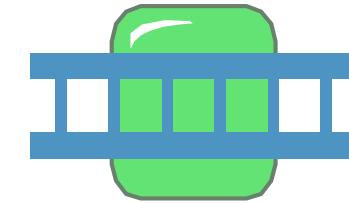
DNA polymerase binds
nucleotides to form new
strands



DNA polymerase I
removes RNA primer and
inserts the correct bases

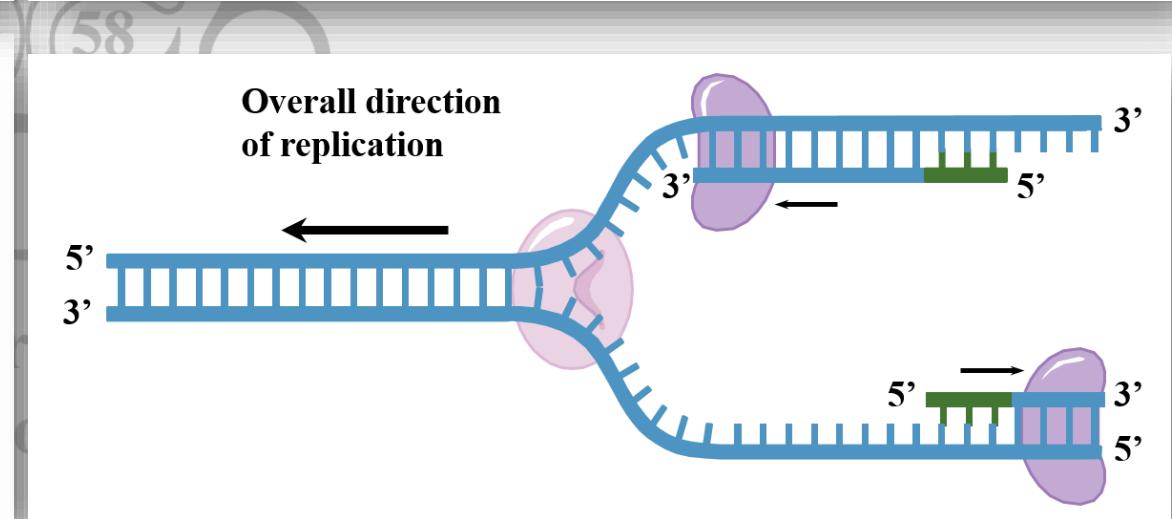
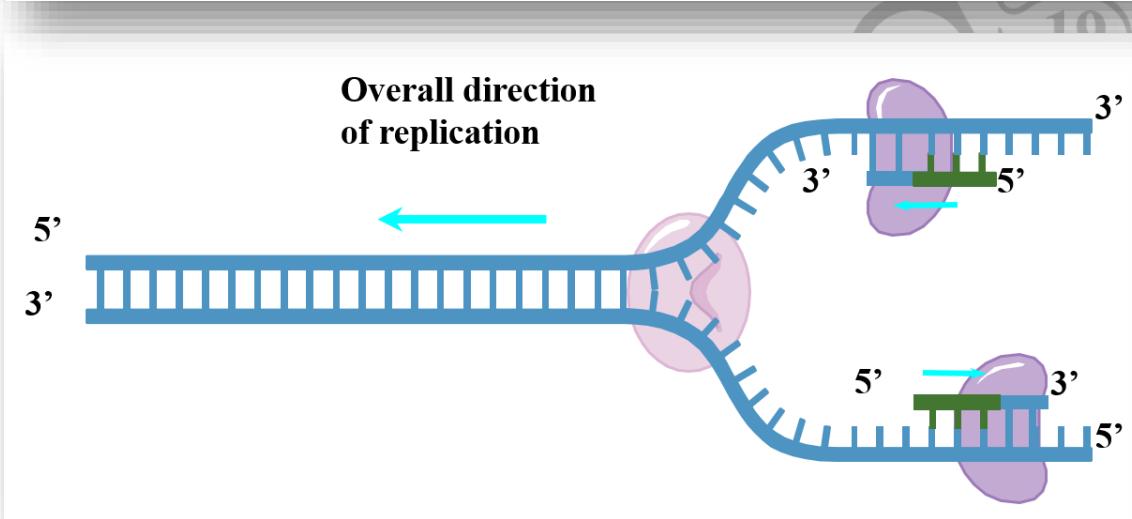
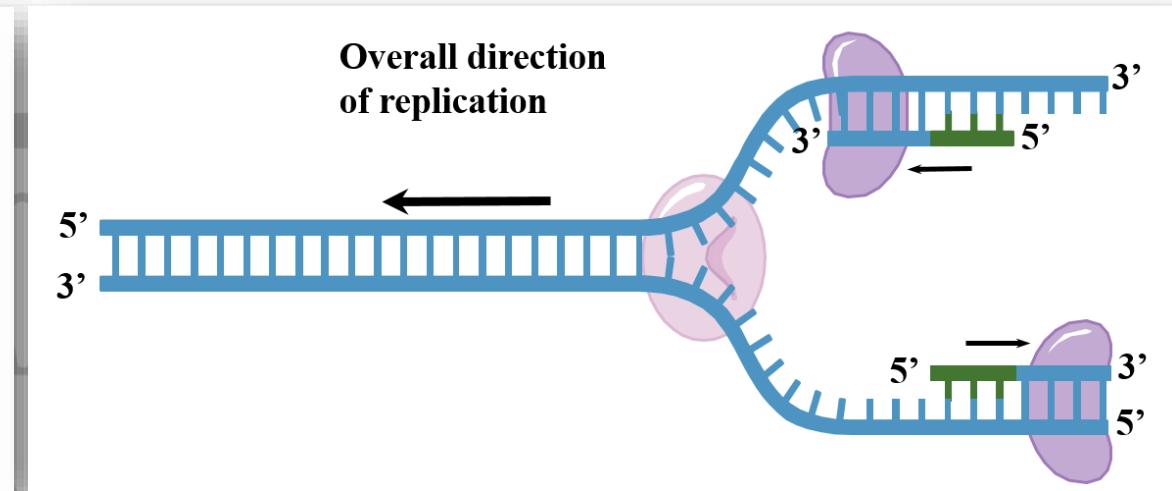
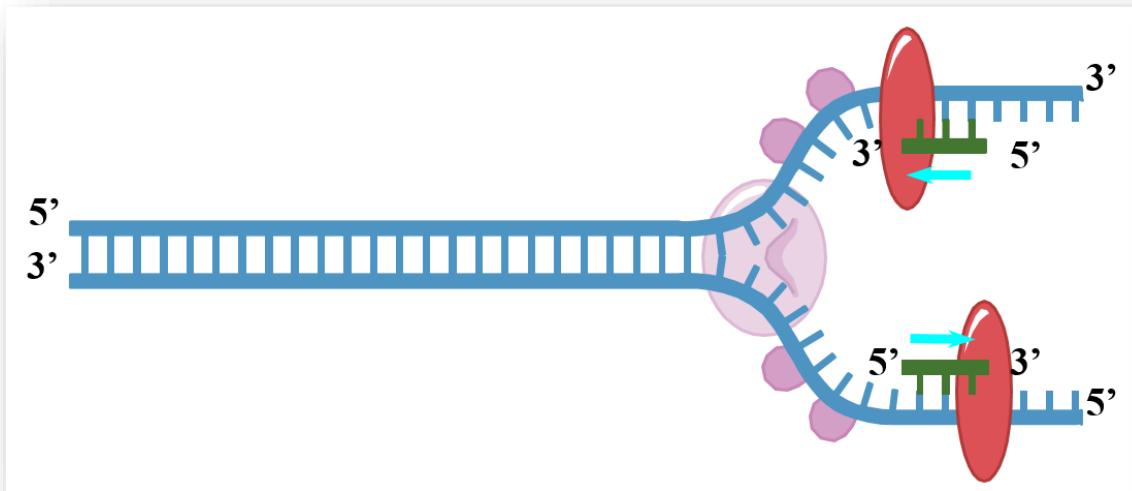


Primase adds short primer
to template strand

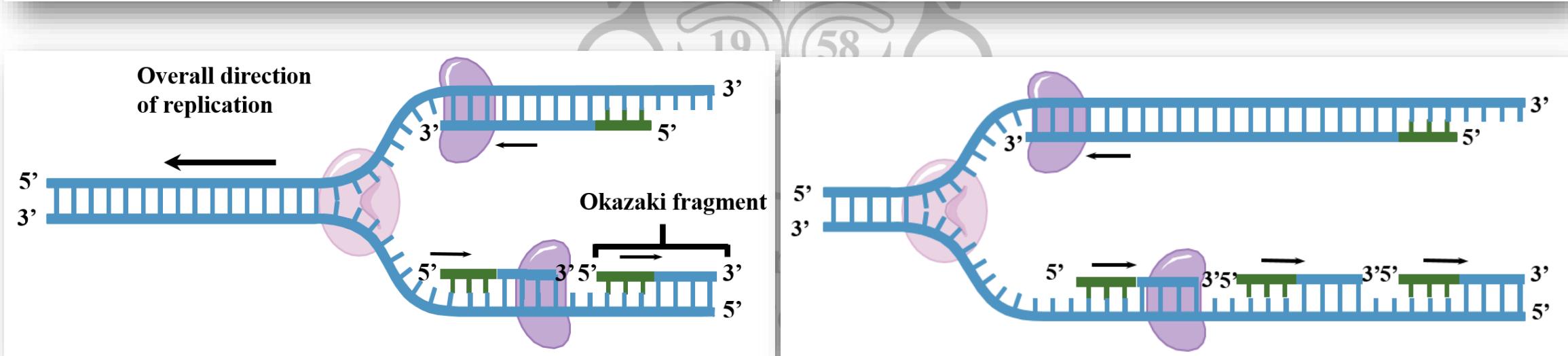
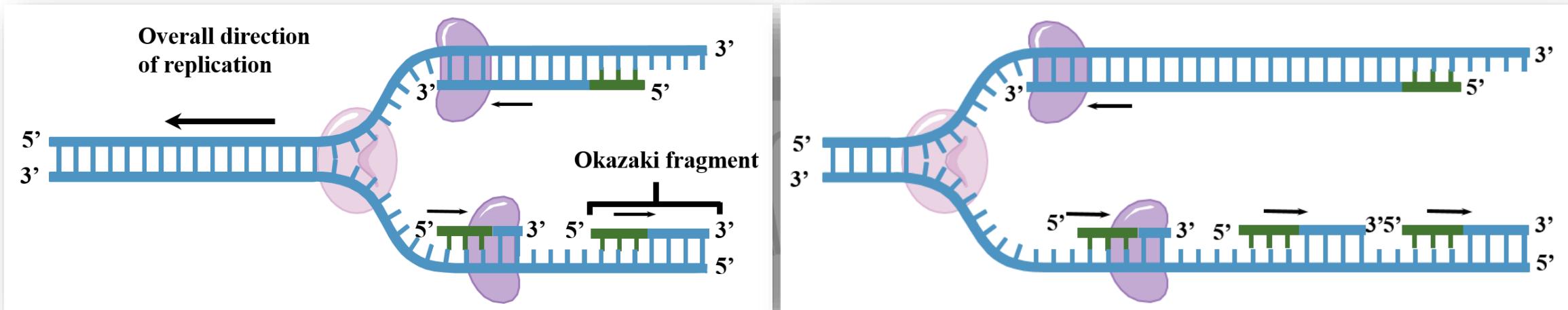


Ligase joins Okazaki fragments
and seals other nicks in sugar-
phosphate backbone

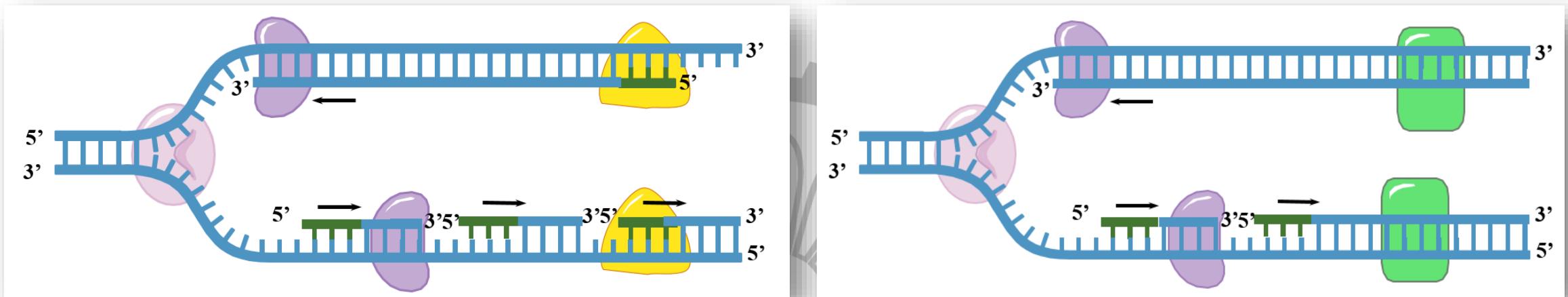
DNA Replication Steps



DNA Replication Steps (Cont.)



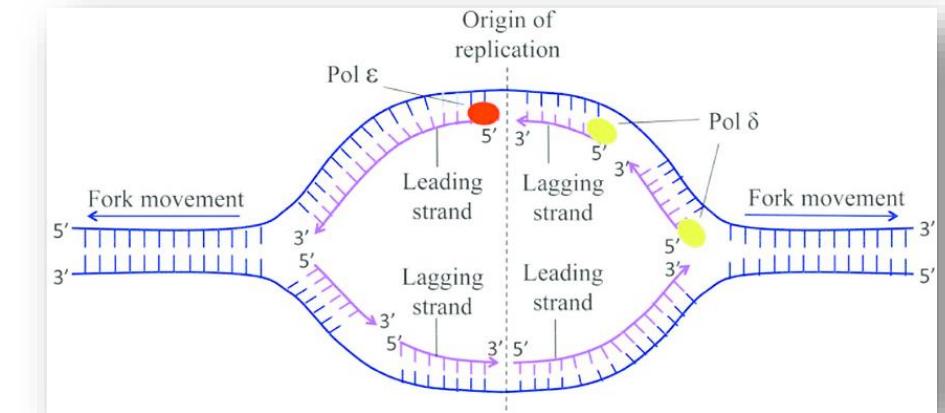
DNA Replication Steps (Cont.)



Amirkabir University of Technology
(Tehran Polytechnic)

Origin and Direction of Replication

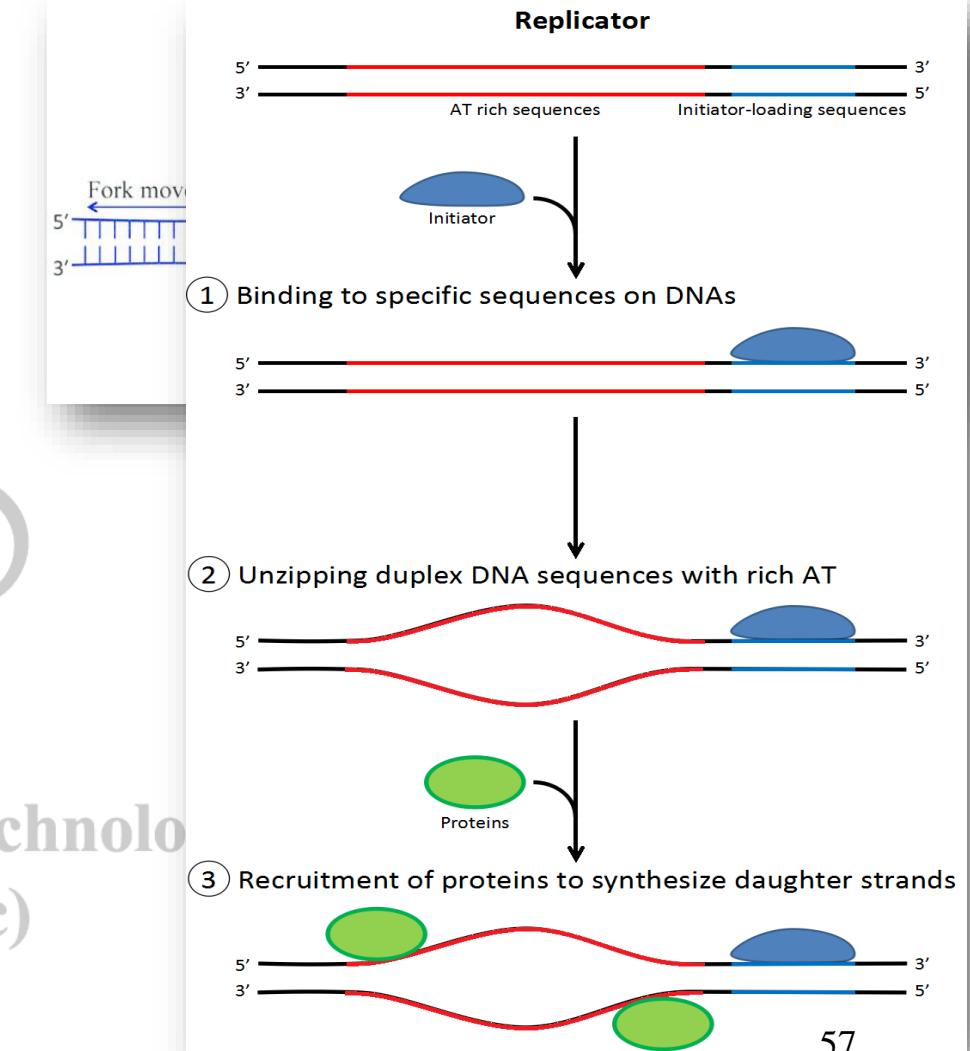
- **Origin of replication:** a specific location in the genome where DNA replication is initiated.
- Replication forks growing **bidirectionally** from the origin.
- Sequences used by initiator proteins tend to be **AT-rich** (adenine and thymine), because A-T base pairs have two hydrogen bonds.



Origin and Direction of Replication

- **Origin of replication:** a specific location in the genome where DNA replication is initiated.
- Replication forks growing **bidirectionally** from the origin.
- Sequences used by initiator proteins tend to be **AT-rich** (adenine and thymine), because A-T base pairs have two hydrogen bonds.

Amirkabir University of Technology
(Tehran Polytechnic)

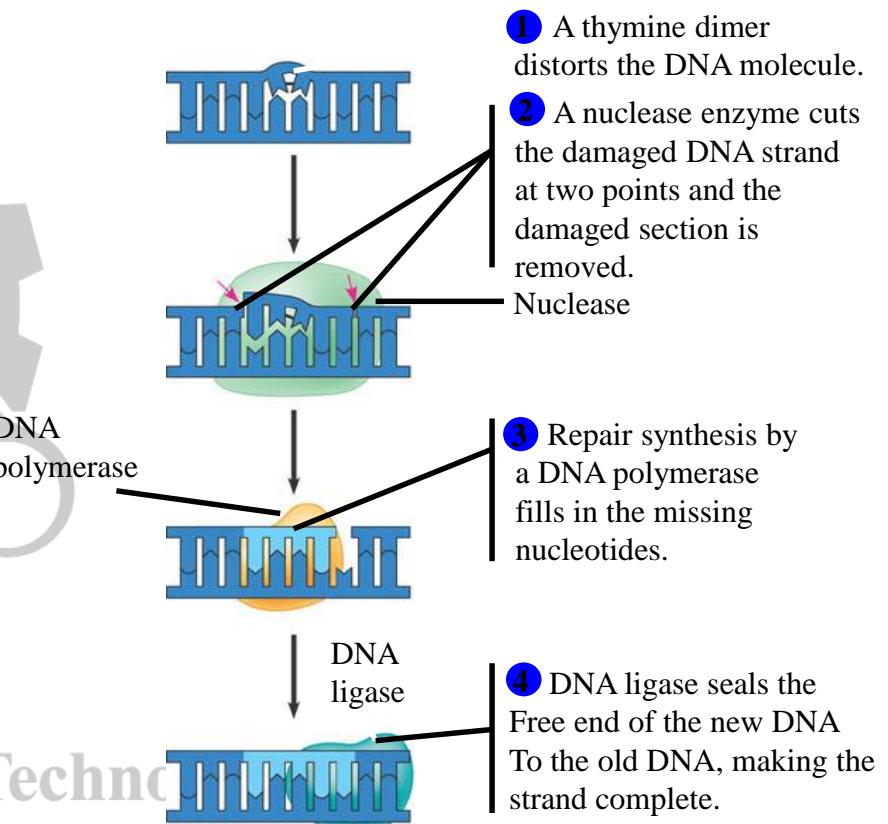


Proofreading

- DNA must be faithfully replicated, but mistakes occur
 - DNA polymerase inserts a wrong nucleotide base in 1/10,000 bases
 - DNA polymerase has a proofreading capability and can correct errors
 - Mismatch repair: ‘wrong’ inserted base can be removed
 - Excision repair: DNA may be damaged by chemicals, radiation, etc.
Mechanism to cut out and replace with correct bases.

Proofreading and Repairing DNA

- DNA polymerases proofread newly made DNA, replacing any incorrect nucleotides
- In mismatch repair of DNA, repair enzymes correct errors in base pairing
- In nucleotide excision DNA repair nucleases cut out and replace damaged stretches of DNA

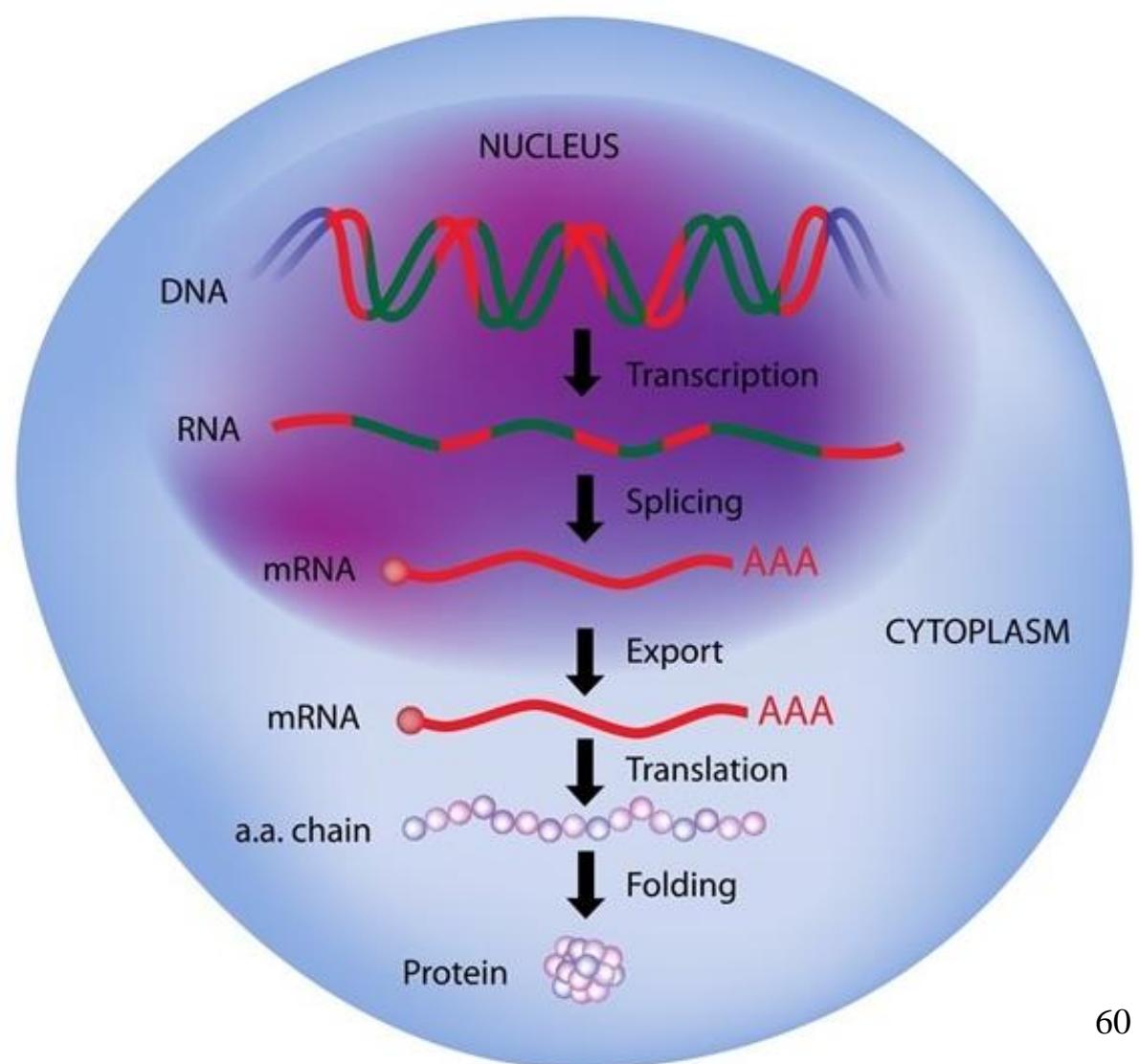


Amirkabir University of Technology
(Tehran Polytechnic)

Transcription (RNA Synthesis)

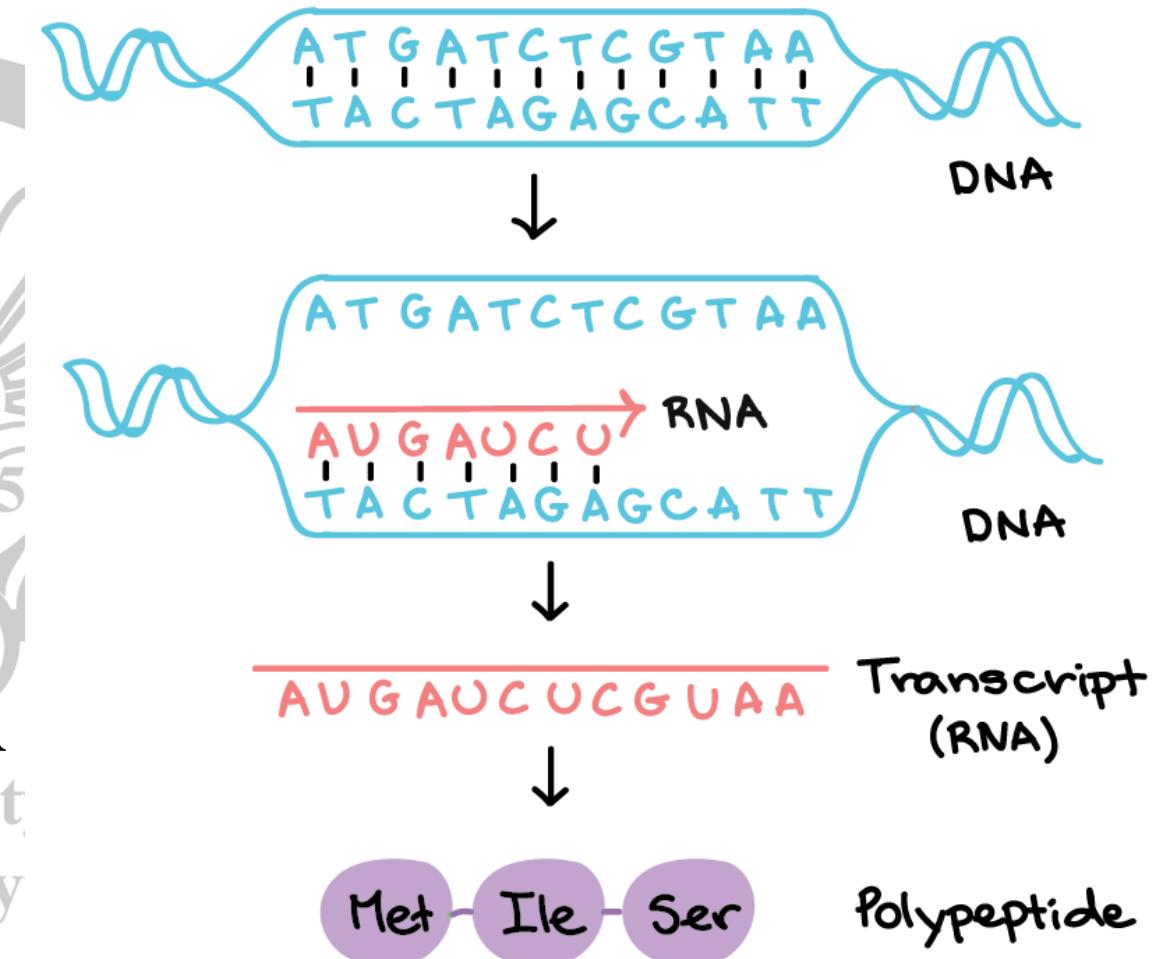
- The genetic information flows from DNA to mRNA and then to the protein synthesizing machinery.

Amirkabir Univ
(Tehran)

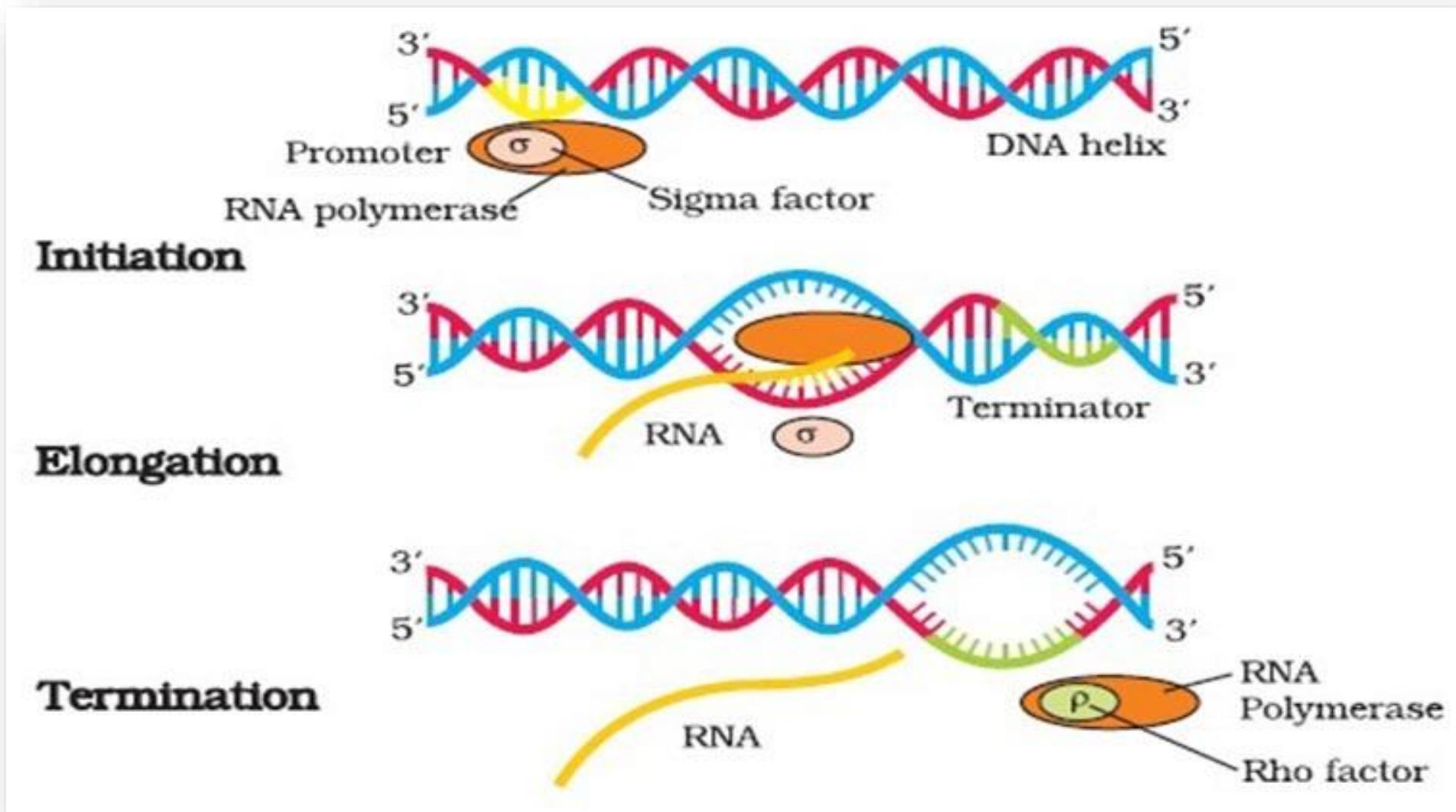


Transcription

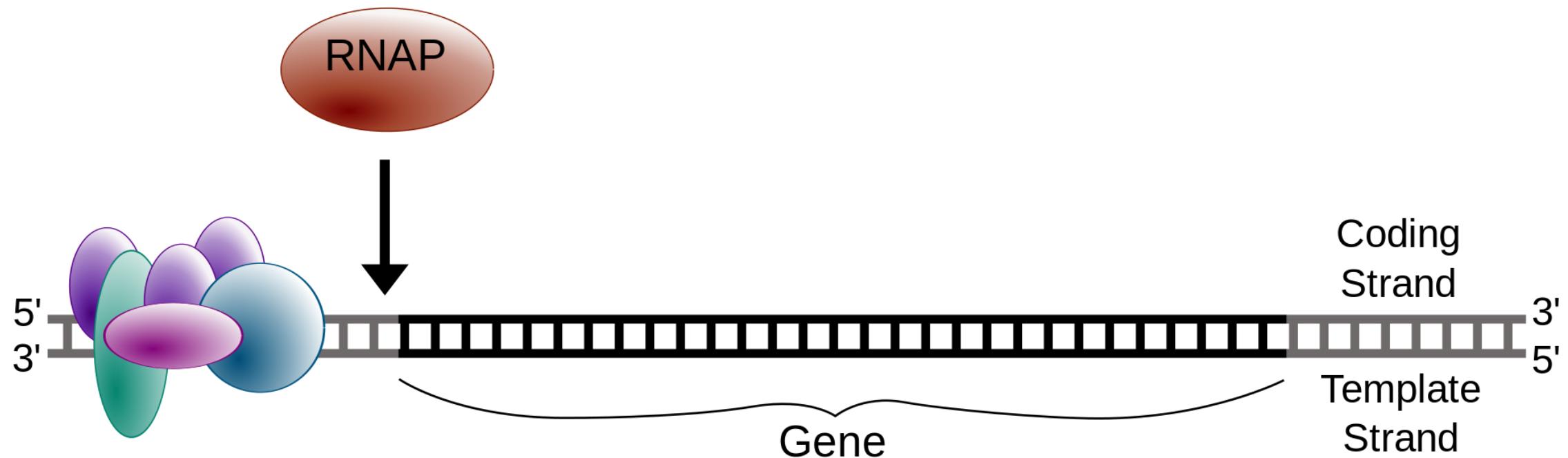
- **Transcription** is the first of several steps of DNA based **gene expression** in which a particular segment of DNA is copied into RNA (especially mRNA) by the enzyme **RNA polymerase**.
- During transcription, a DNA sequence is read by an RNA polymerase, which produces a complementary, **antiparallel** RNA strand called a **primary transcript**.



Steps in Transcription

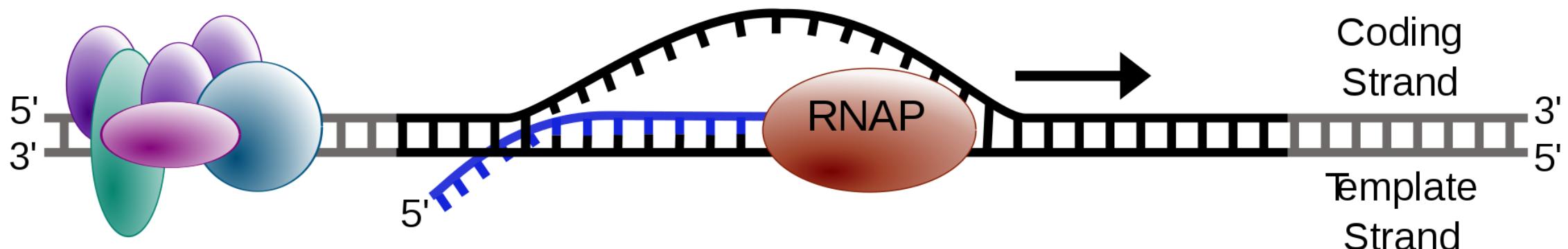


Steps in Transcription: (1)



- RNA polymerase, together with one or more general transcription factors, binds to promoter DNA.
- RNA polymerase creates a transcription bubble, which separates the two strands of the DNA helix.

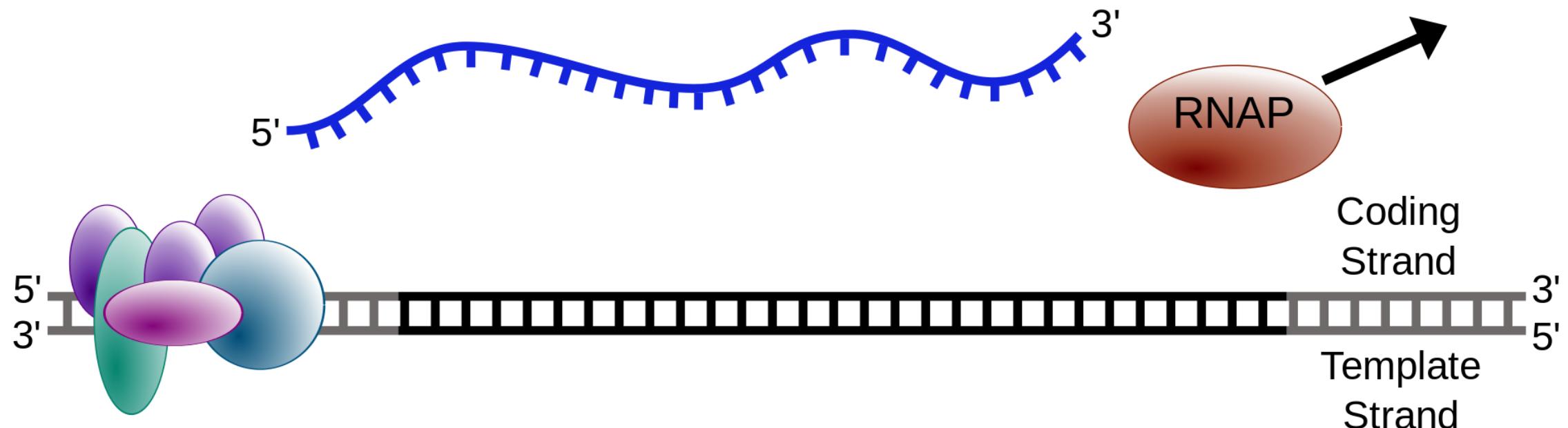
Steps in Transcription: (2)



- RNA polymerase adds RNA nucleotides (which are complementary to the nucleotides of one DNA strand).
- RNA sugar-phosphate backbone forms with assistance from RNA polymerase to form an RNA strand.

(Tehran Polytechnic)

Steps in Transcription: (3)



- Hydrogen bonds of the RNA–DNA helix break, freeing the newly synthesized RNA strand.
- The RNA may remain (in the nucleus or) exit to the cytoplasm through the nuclear pore complex.

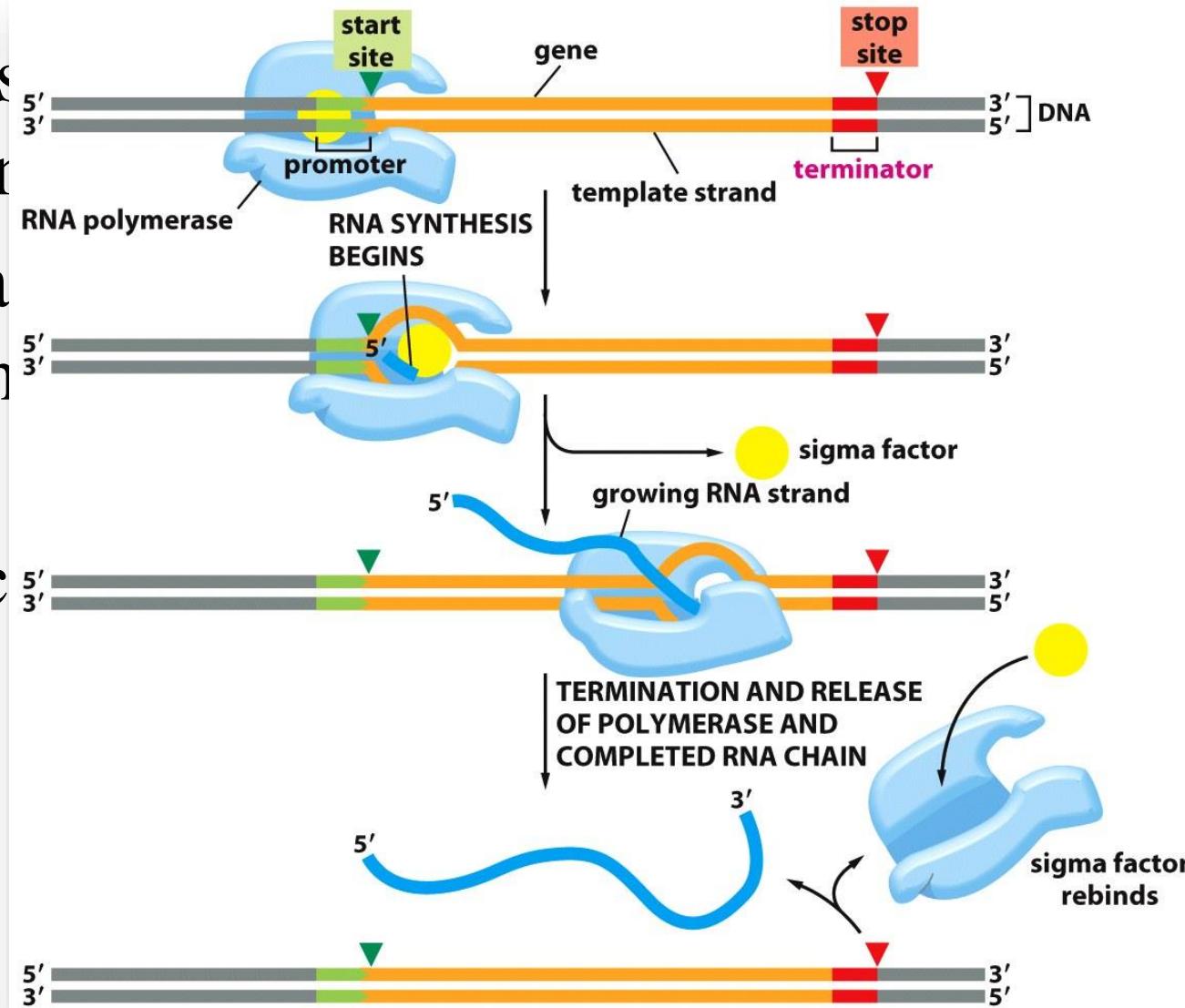
Promoter

- **Promoter** is a region of DNA that leads to initiation of transcription of a particular gene.
- Promoters are located near the transcription start sites of genes, upstream on the DNA (towards the 5' region of the sense strand).
- Promoters can be about 100–1000 base pairs long.

Amirkabir University of Technology
(Tehran Polytechnic)

Promoter

- Promoter is the site of transcription initiation.
- Promoters are located upstream on the template strand.
- Promoters control the rate of transcription.

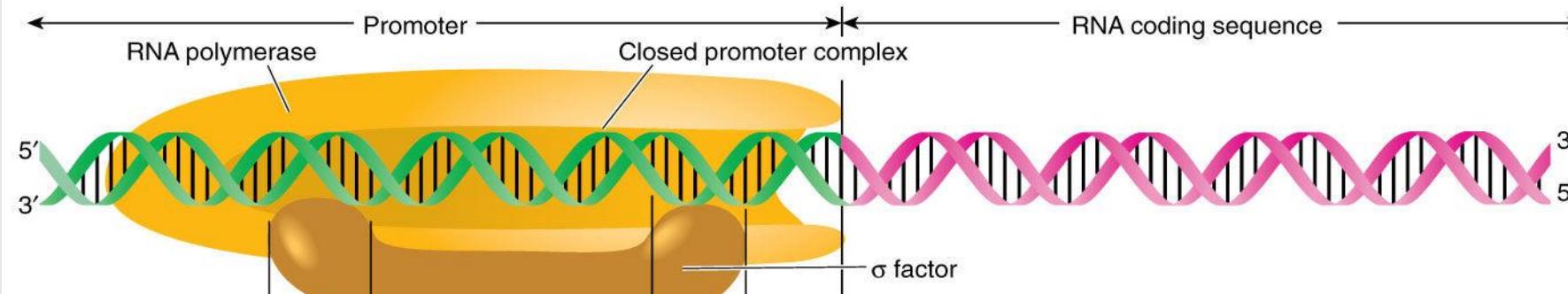


Sigma Factor

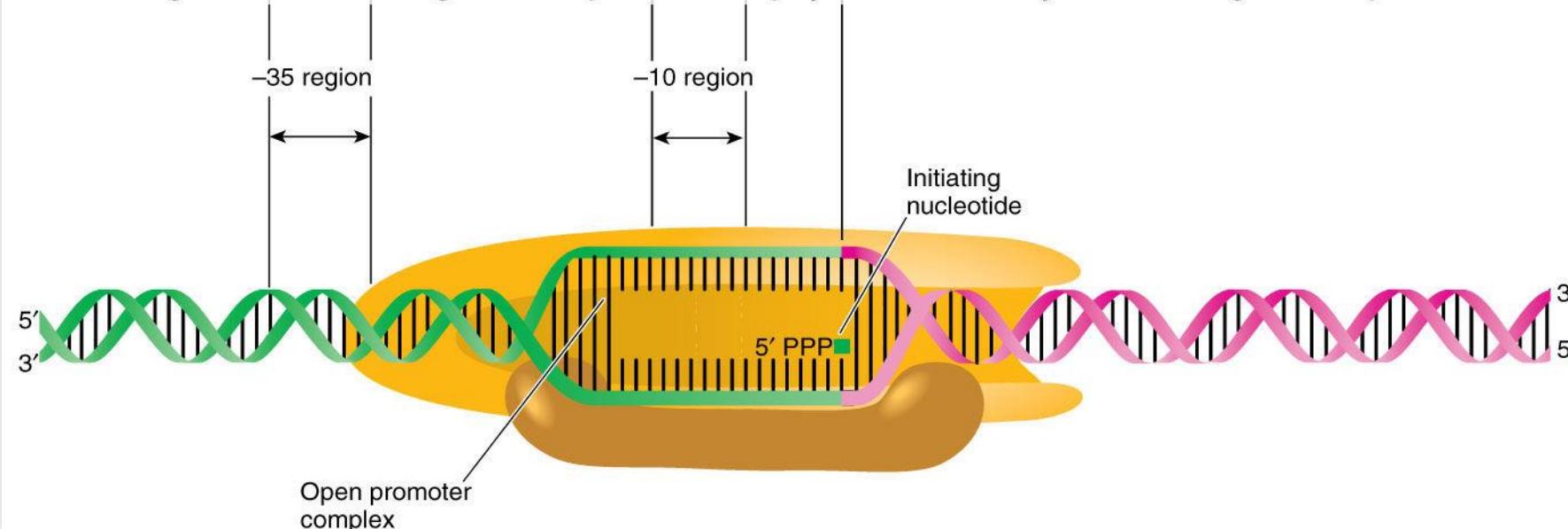
- **Sigma factor** is a bacterial transcription initiation factor that enables specific binding of RNA polymerase (RNAP) to gene promoters.
- Selection of promoters by RNA polymerase is dependent on the sigma factor that associates with it.
- The sigma factor, together with RNA polymerase, is known as the RNA polymerase *holoenzyme*.
- After binding polymerase to the promoter and producing around 10 nucleotides of RNA the sigma factor is released.
- At the end, the polymerase reassociates with a free sigma factor and searches for another promoter to begin the process again.

Initiation of Transcription in Prokaryotes

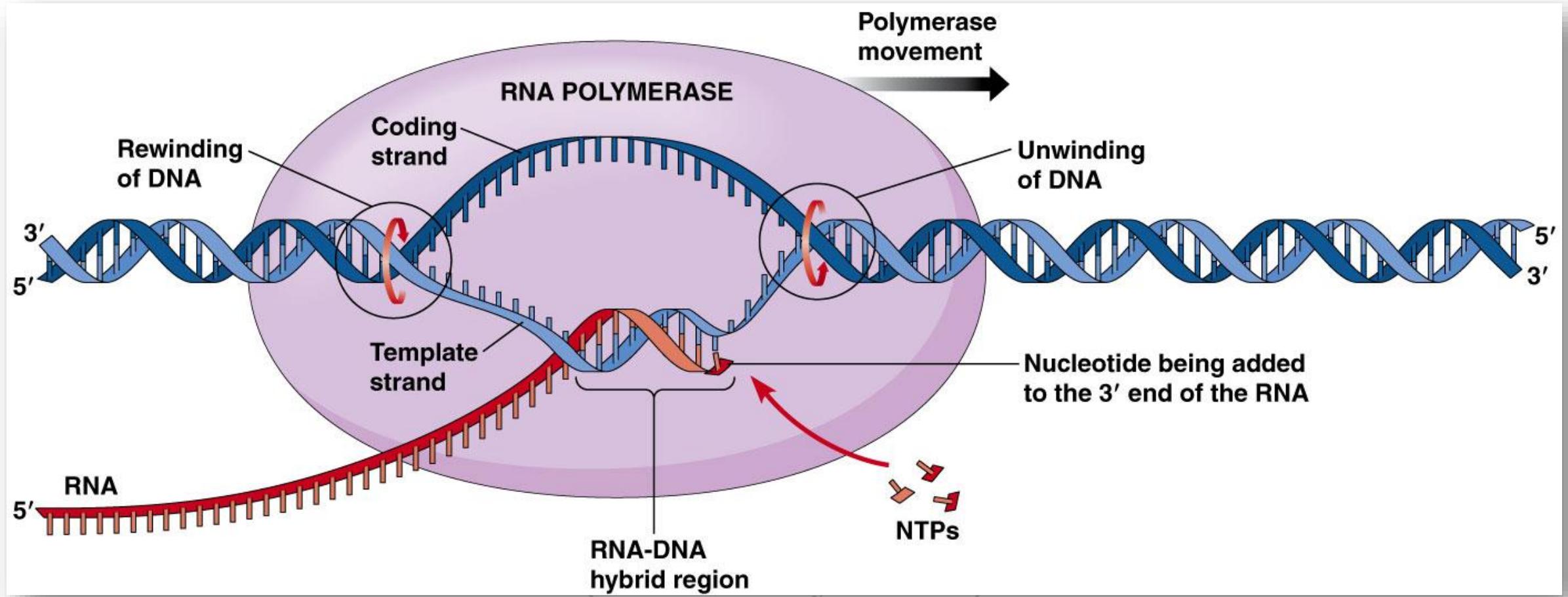
a) In initiation, the RNA polymerase holoenzyme first recognizes the promoter at the -35 region and binds to the full promoter.



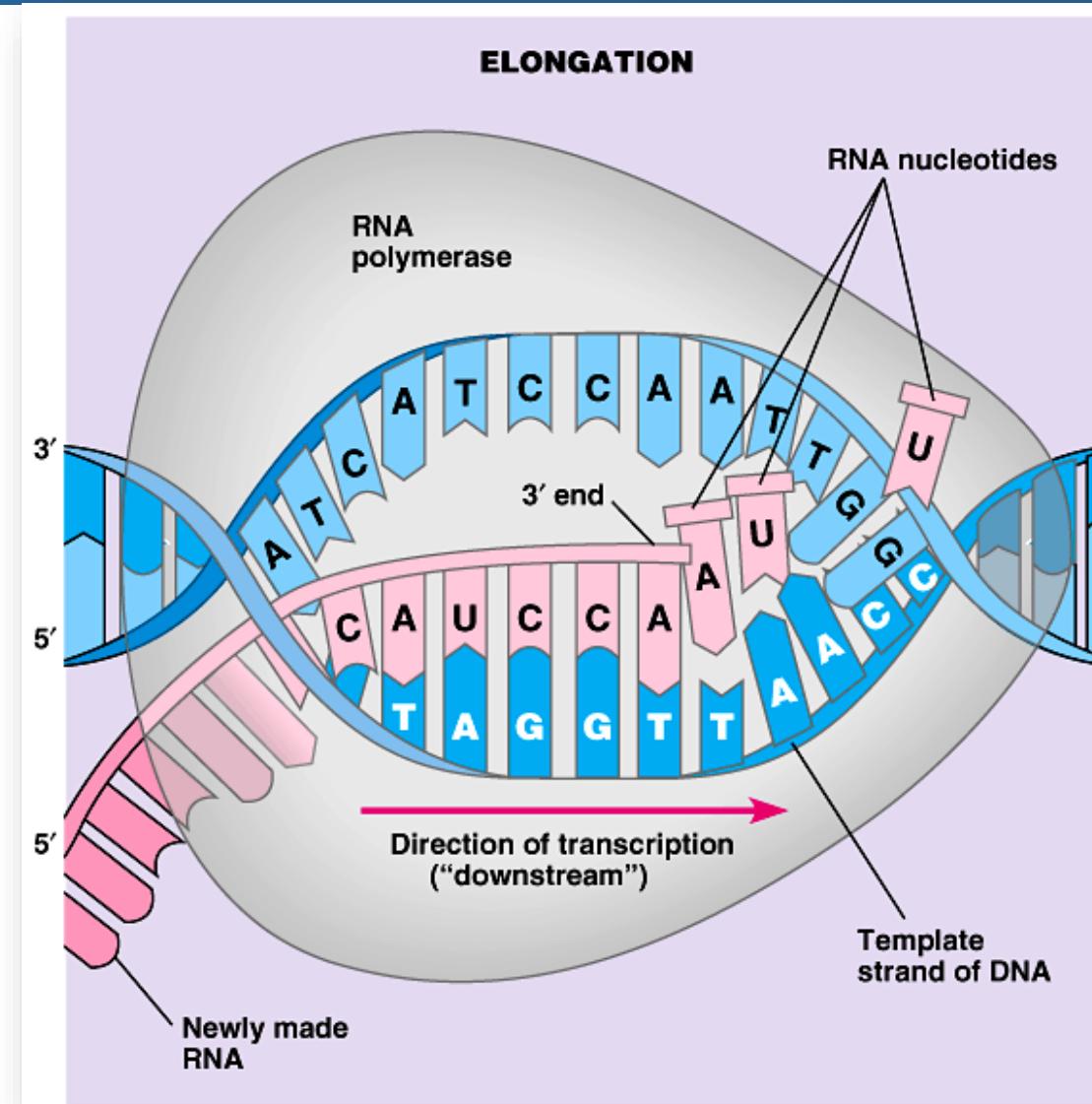
b) As initiation continues, RNA polymerase binds more tightly to the promoter at the -10 region, accompanied by a local untwisting of the DNA in that region. At this point, the RNA polymerase is correctly oriented to begin transcription at +1.



Elongation Step of Transcription



Transcription Bubble

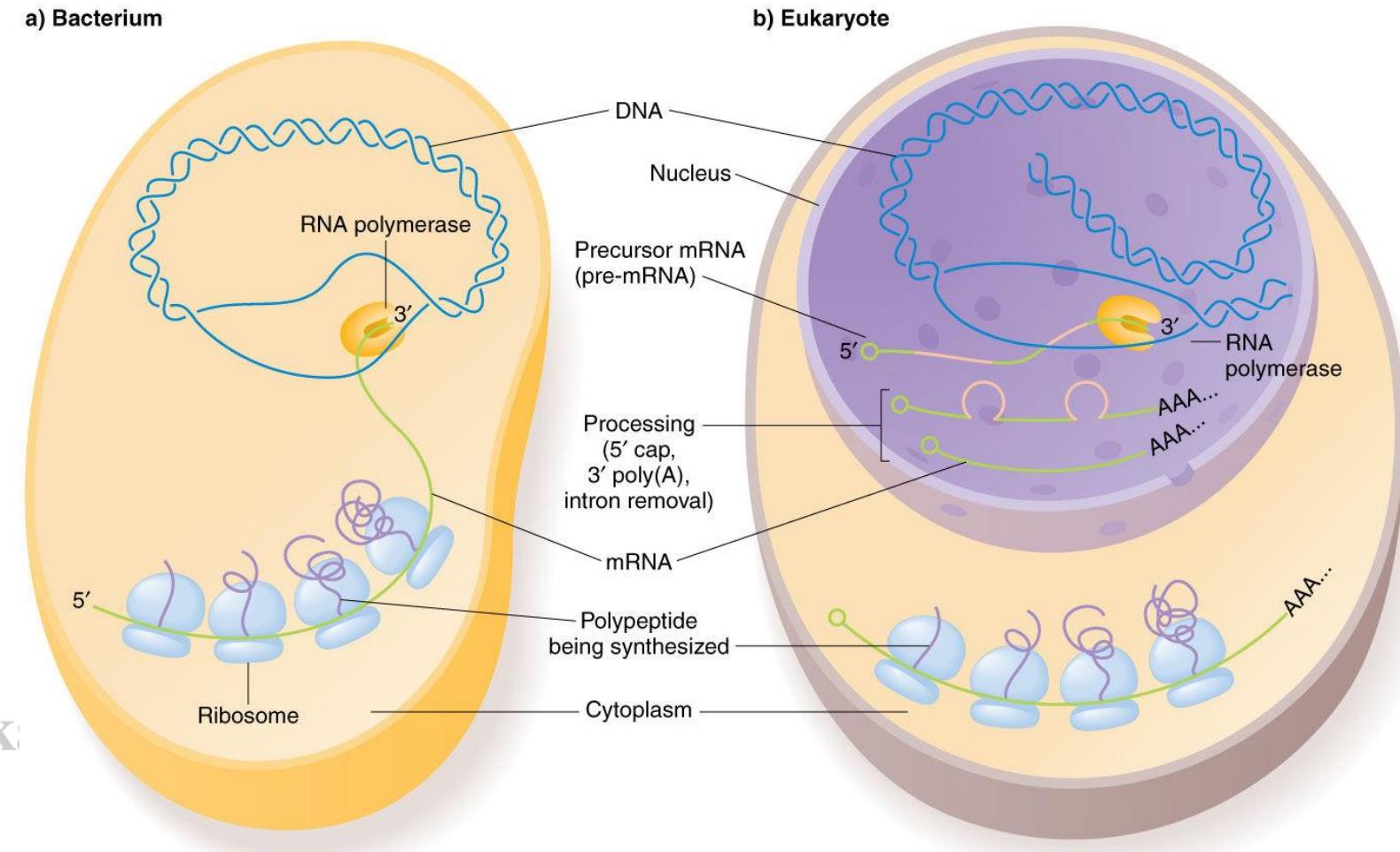


Differences between Replication and Transcription

- Ribonucleotides are used in RNA synthesis rather than deoxy ribonucleotides
- DNA helicase vs RNA polymerase for unbinding helix
- Transcription occurs along only one strand of DNA.
- U replaces T as the complementary base pair for A in RNA
- A primer is not involved in RNA synthesis
- Transcription is unidirectional while replication usually is bi-directional
- Only a portion of the genome is transcribed or copied into RNA, whereas the entire genome must be copied during DNA replication
- There is no proofreading function during RNA transcription.

Transcription in Prokaryotes vs Eukaryotes

- In prokaryotes transcription and translation can occur simultaneously and both are happened in the cytoplasm.
- In eukaryotes transcription is happened in the nucleus, and proteins are made (translation) in the cytoplasm.



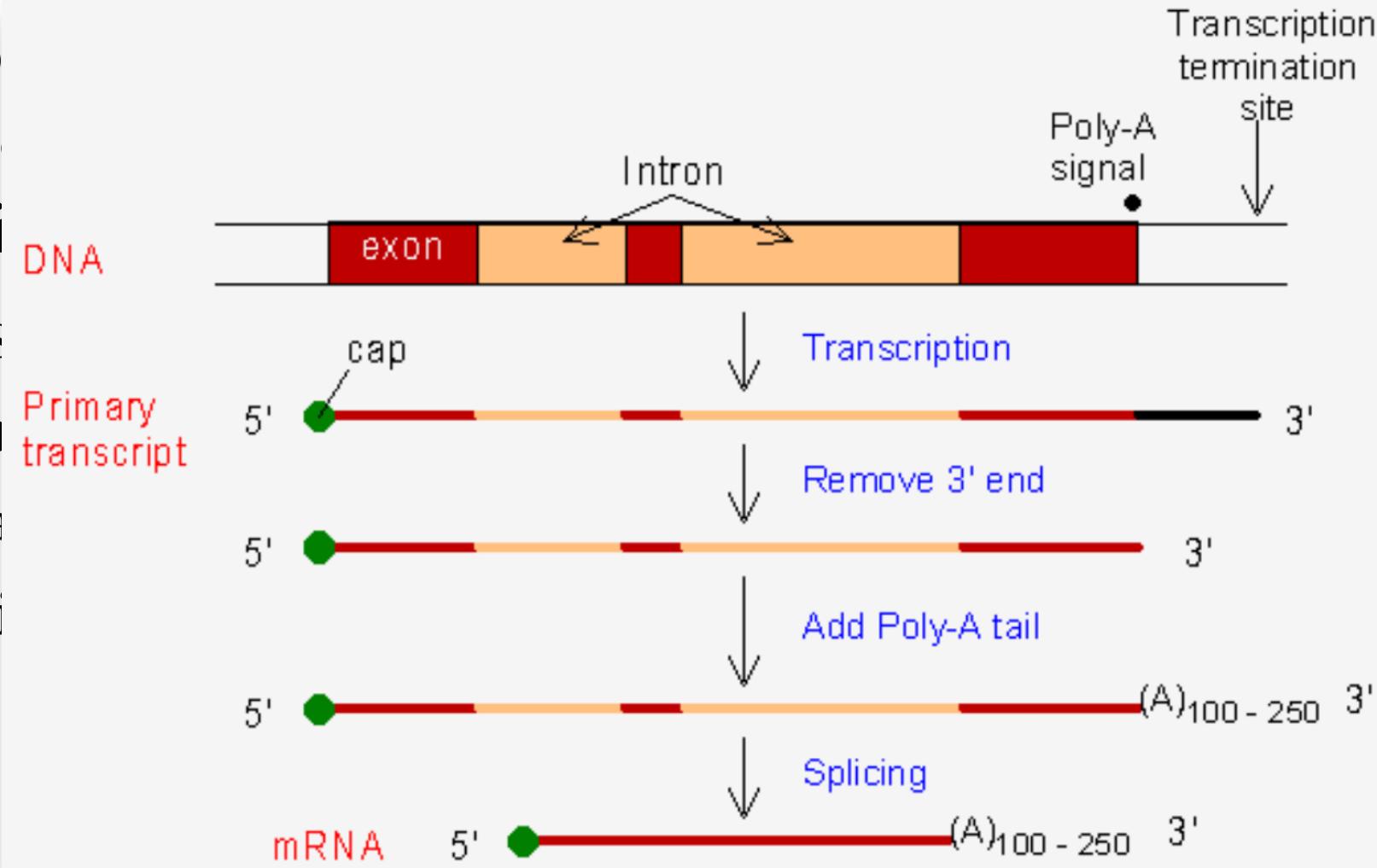
RNA Processing

- **RNA processing** is to generate a *mature mRNA* (for protein genes) or a functional tRNA or rRNA from the primary transcript (pre-mRNA).
- Processing of pre-mRNA involves the following steps:
 - **Capping:** add 7-methylguanylate (m^7G) to the 5' end.
 - **Polyadenylation:** - add a poly-A tail to the 3' end.
 - **Splicing:** remove introns and join exons.

Amirkabir University of Technology
(Tehran Polytechnic)

RNA Processing

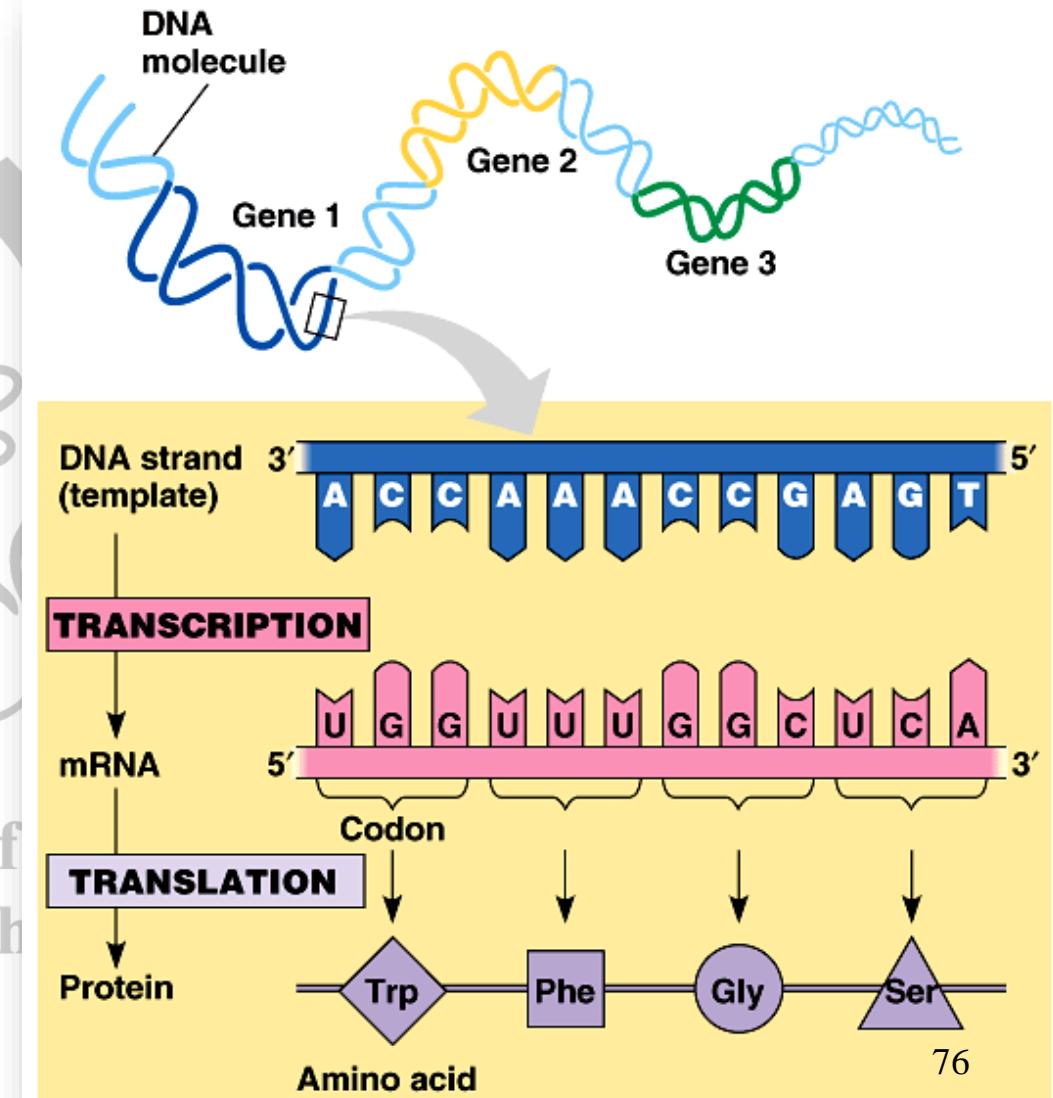
- RNA processing (genes) → protein
- Processed mRNA
 - Capping
 - Polyadenylation
 - Splicing



Translation

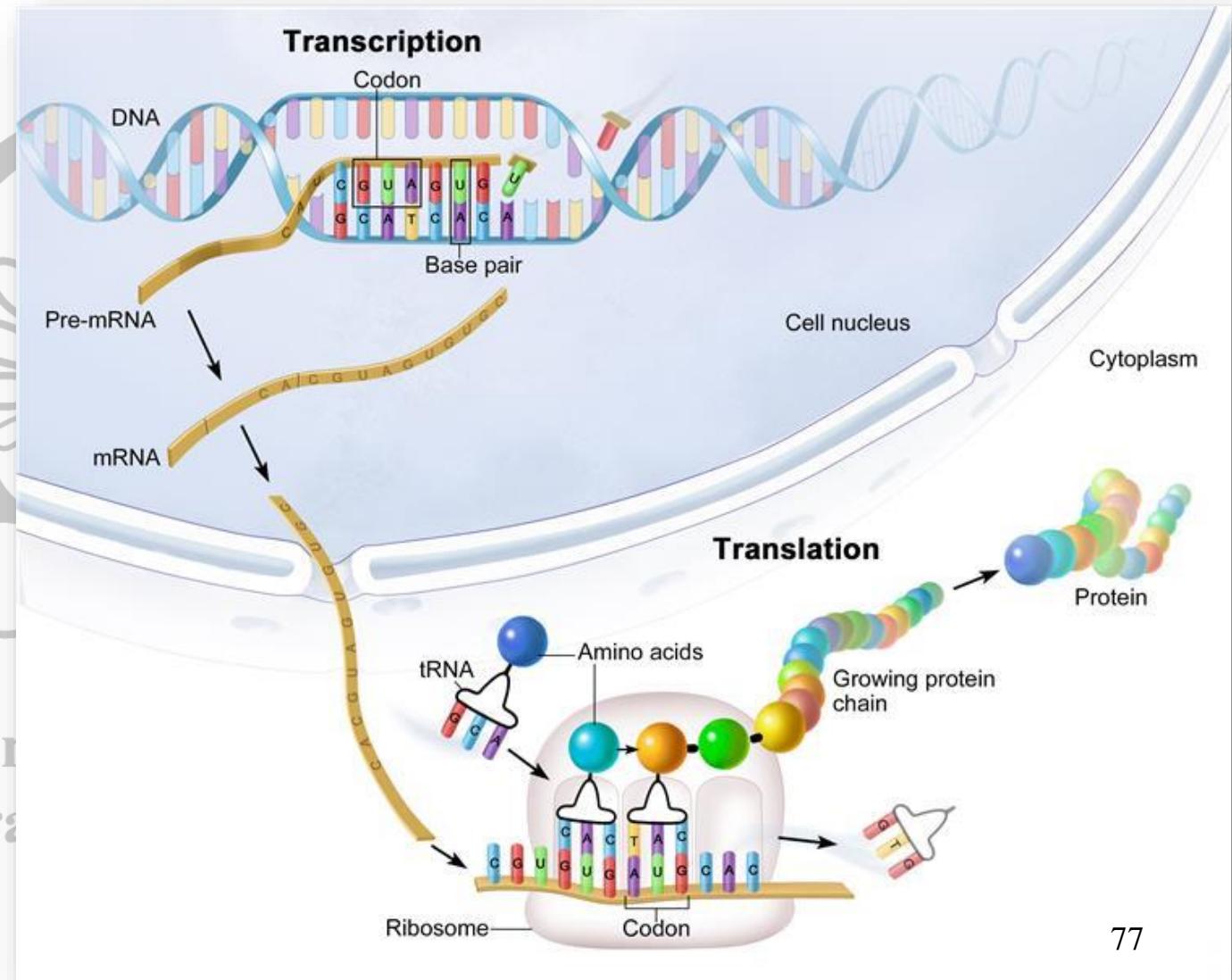
- In translation, mRNA is decoded in the ribosome decoding center to produce a specific amino acid chain, or polypeptide.
- Blocks of 3 nucleotides decoded into the sequence of amino acids.

Amirkabir University of
(Tehran Polytech)



Translation (Cont.)

- Protein is created from mRNA
- Translation from protein's language to nucleotide's language



Amirkabir University of Technology
(Tehran, Iran)

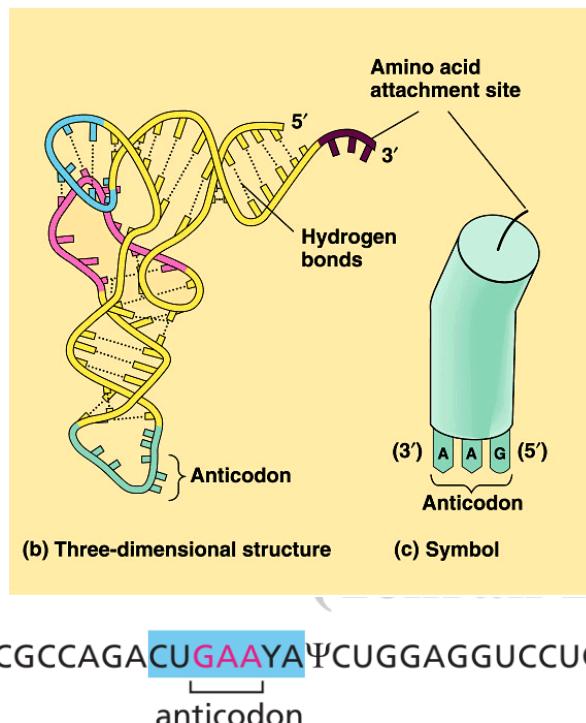
The 3 Main Types of RNA

- **Messenger RNA (mRNA)** transcribes the genetic code from DNA into a form that can be read and used to make proteins. mRNA carries genetic information from the nucleus to the cytoplasm of a cell.
- **Ribosomal RNA (rRNA)** is located in the cytoplasm of a cell, where ribosomes are found. rRNA directs the translation of mRNA into proteins.
- **Transfer RNA (tRNA)** is located in the cellular cytoplasm and is involved in protein synthesis. tRNA brings or transfers amino acids to the ribosome that corresponds to each *three-nucleotide codon* of rRNA. The amino acids then can be joined together and processed to make polypeptides and proteins.

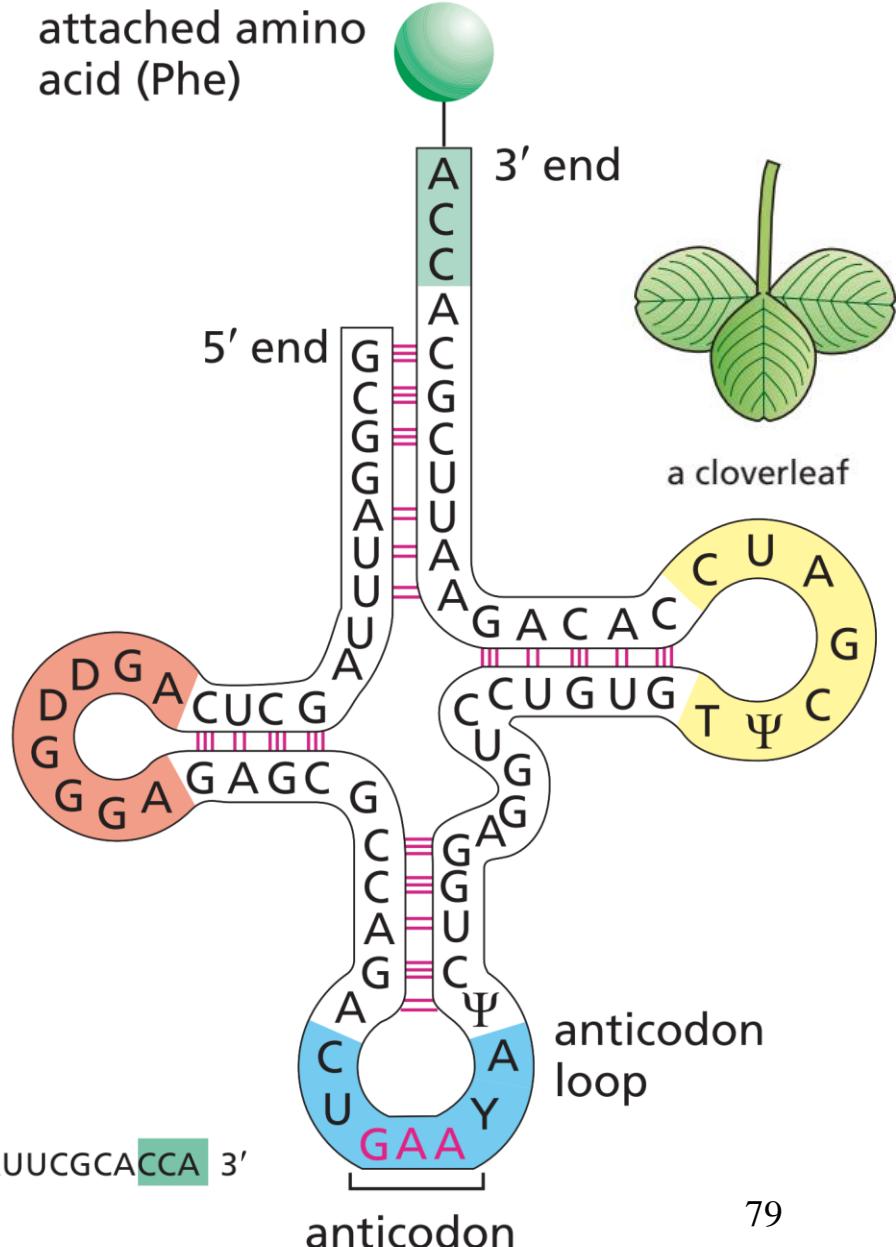
Amirkabir University of Technology
(Tehran Polytechnic)

tRNA Structure

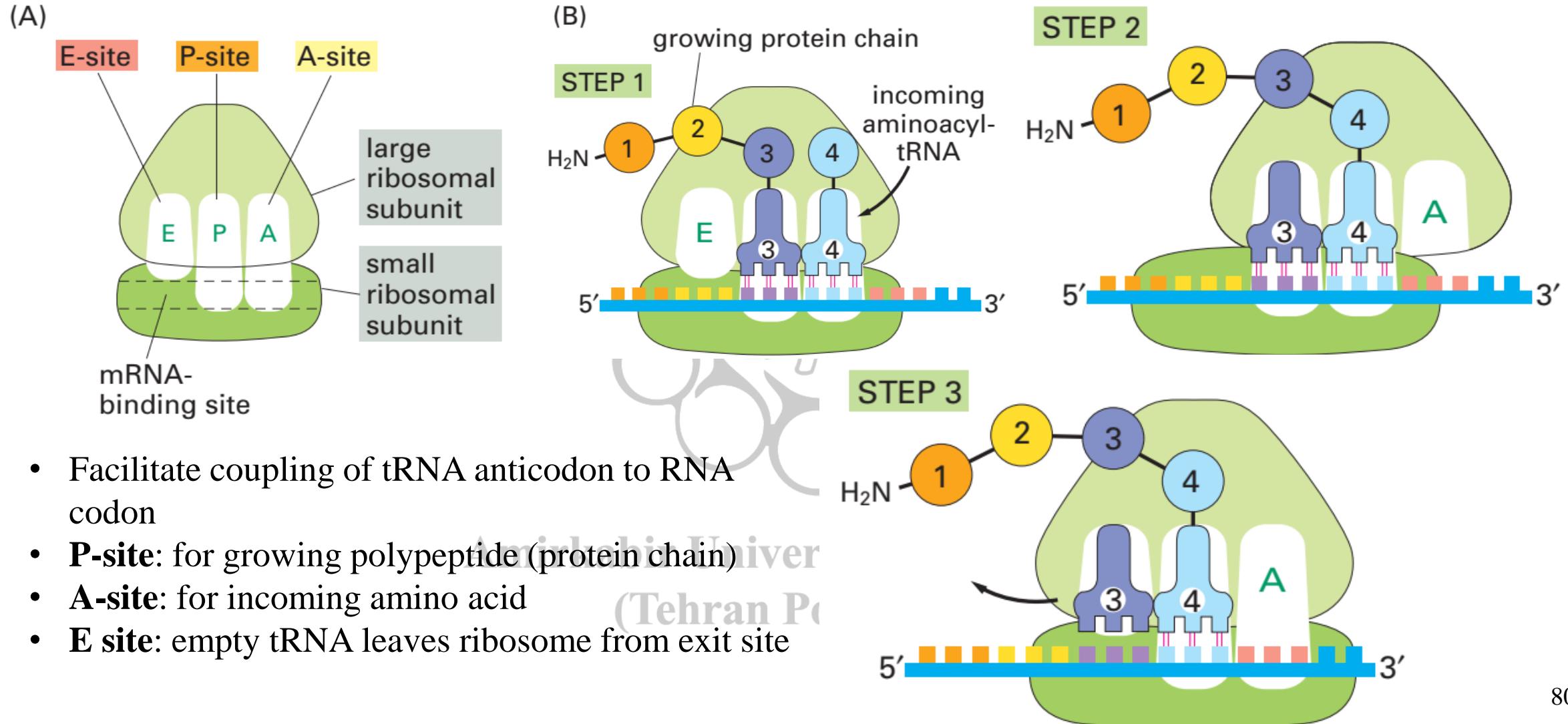
- It has 76 nucleotides.
- Anticodon define which amino acid should be connected to tRNA



rsity of Tec
'olytechnic)



Ribosome



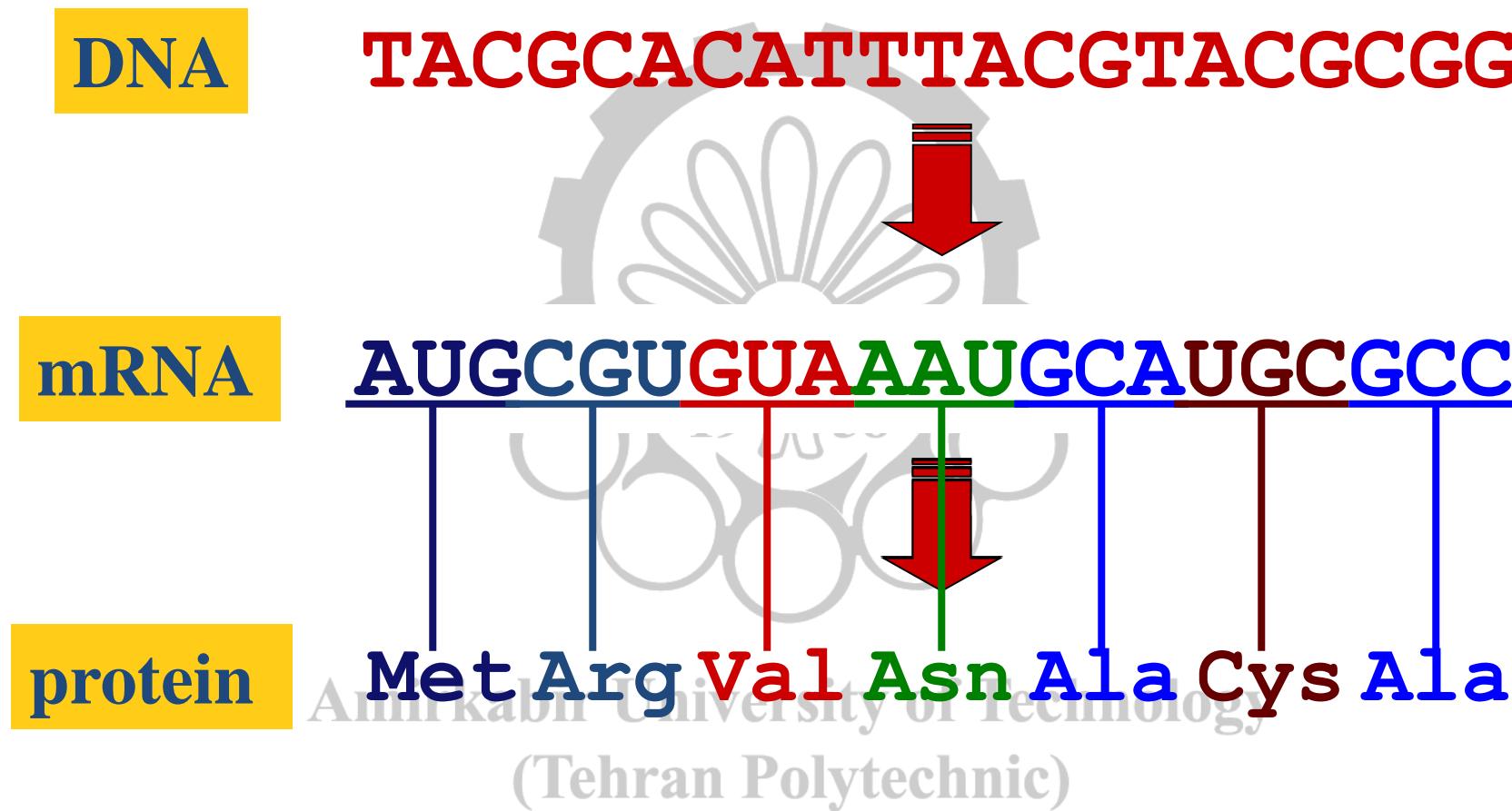
Genetic Codes

- Code is redundant
 - Several codons for each amino acid
- Start codon
 - AUG
 - Methionine
- Stop codons
 - UGA, UAA, UAG
 - There is no any anticodon for them
 - Without translation

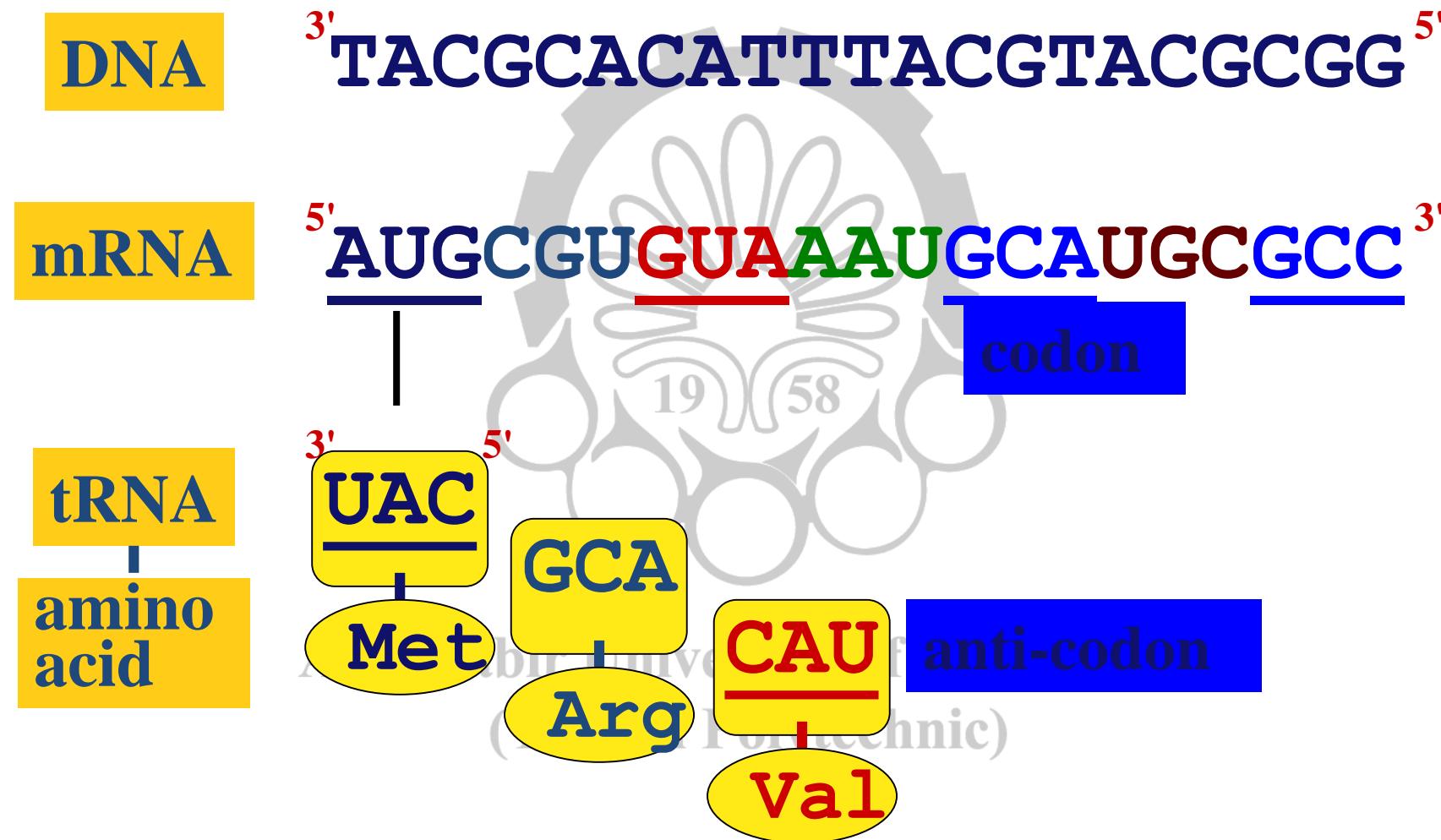
Amirkabir University
(Tehran Poly)

	Second base				
	U	C	A	G	
U	UUU Phe UUC UUA Leu UUG	UCU UCC UCA UCG	UAU Tyr UAC UAA Stop UAG Stop	UGU Cys UGC UGA Stop UGG Trip	U C A G
C	CUU CUC Leu CUA CUG	CCU CCC CCA CCG	CAU His CAC CAA Gln CAG	CGU CGC CGA CGG	U C A G
A	AUU AUC Ile AUA AUG Met or start	ACU ACC ACA ACG	AAU Asn AAC Thr AAA Lys AAG	AGU Ser AGC AGA Arg AGG	U C A G
G	GUU GUC Val GUA GUG	GCU GCC GCA GCG	GAU Asp GAC Ala GAA Glu GAG	GGU GGC GGA Gly GGG	U C A G
First base (5' end) Third base (3' end)					

mRNA Codes for Proteins in Triplets

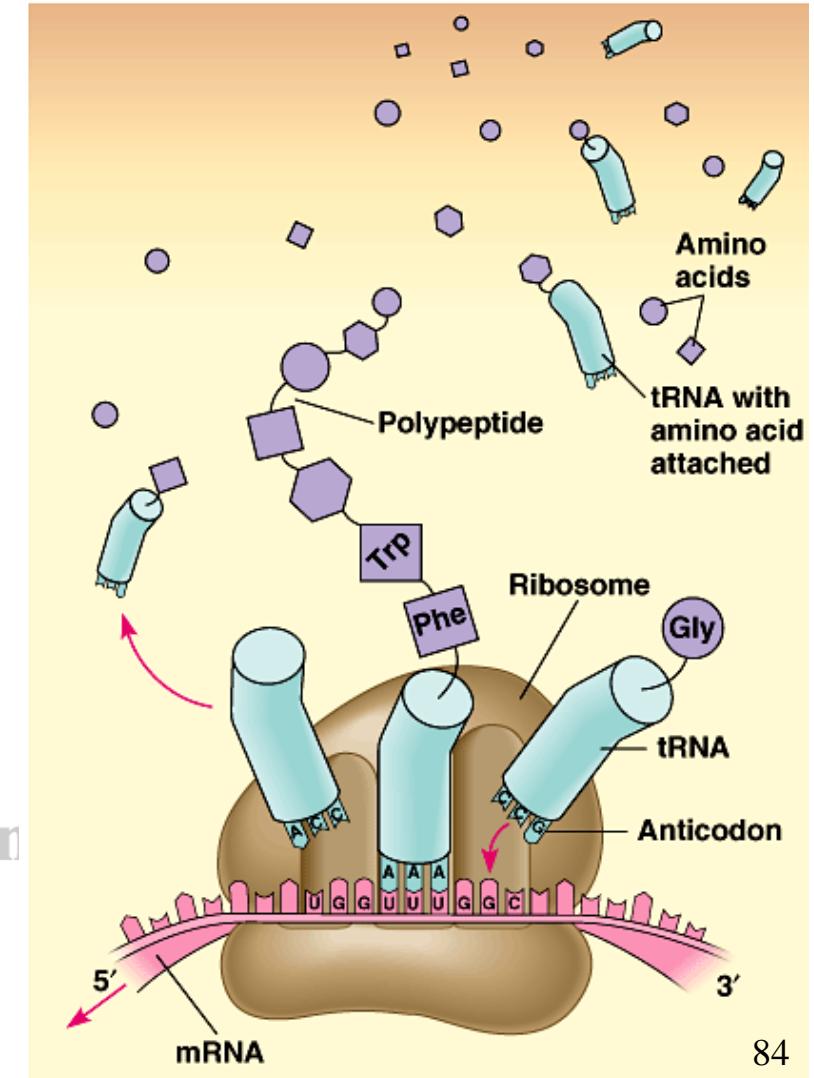
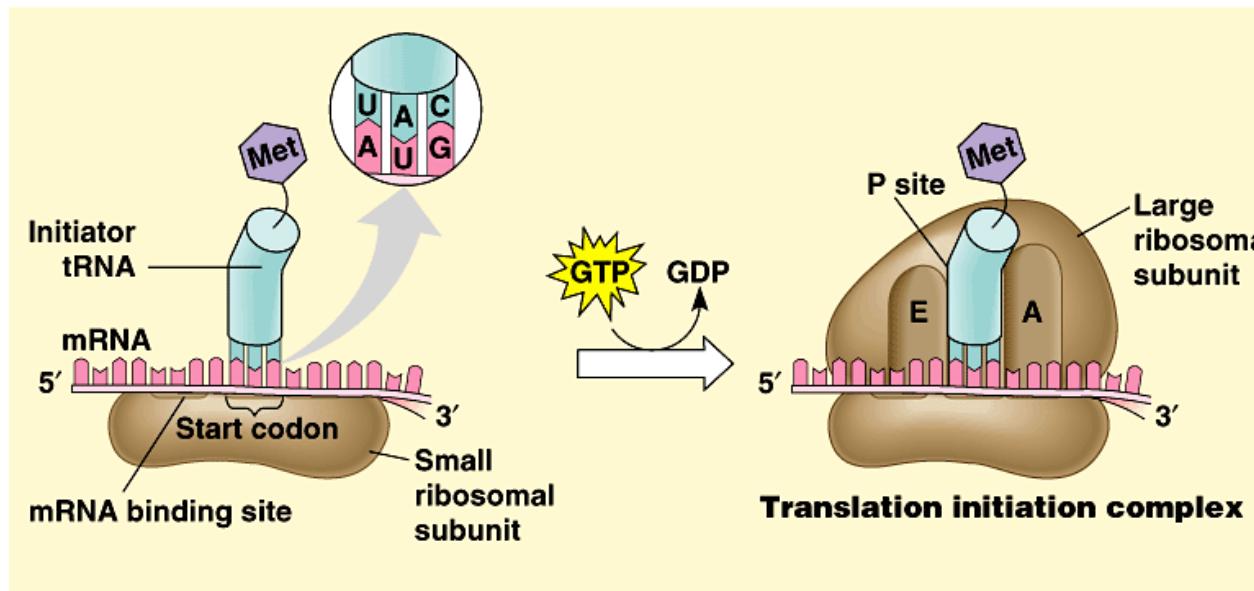


How are the codons matched to amino acids?

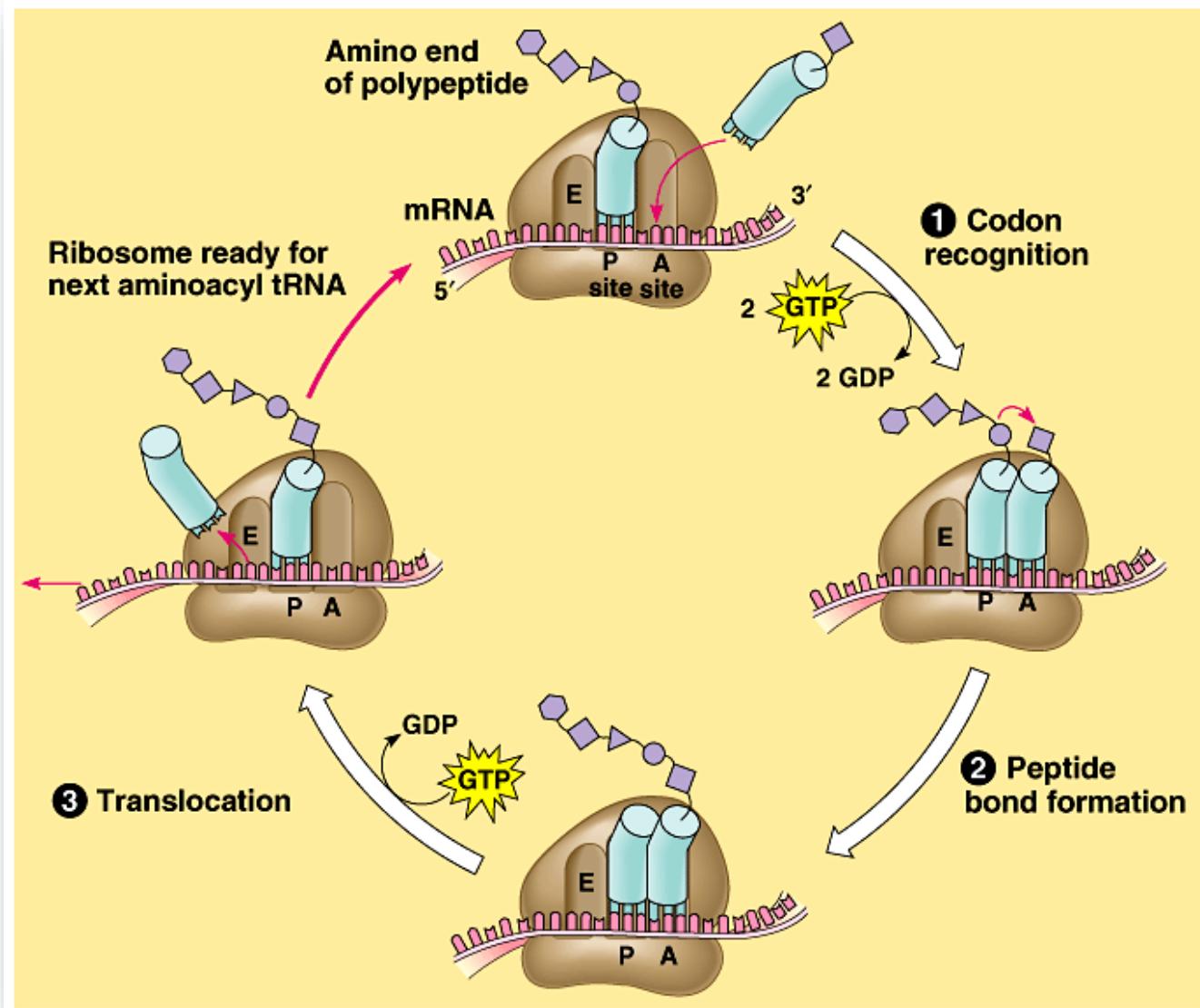


Building a Polypeptide

- Initiation
 - Brings together mRNA, ribosome subunits, proteins & initiator tRNA
- Elongation
- Termination



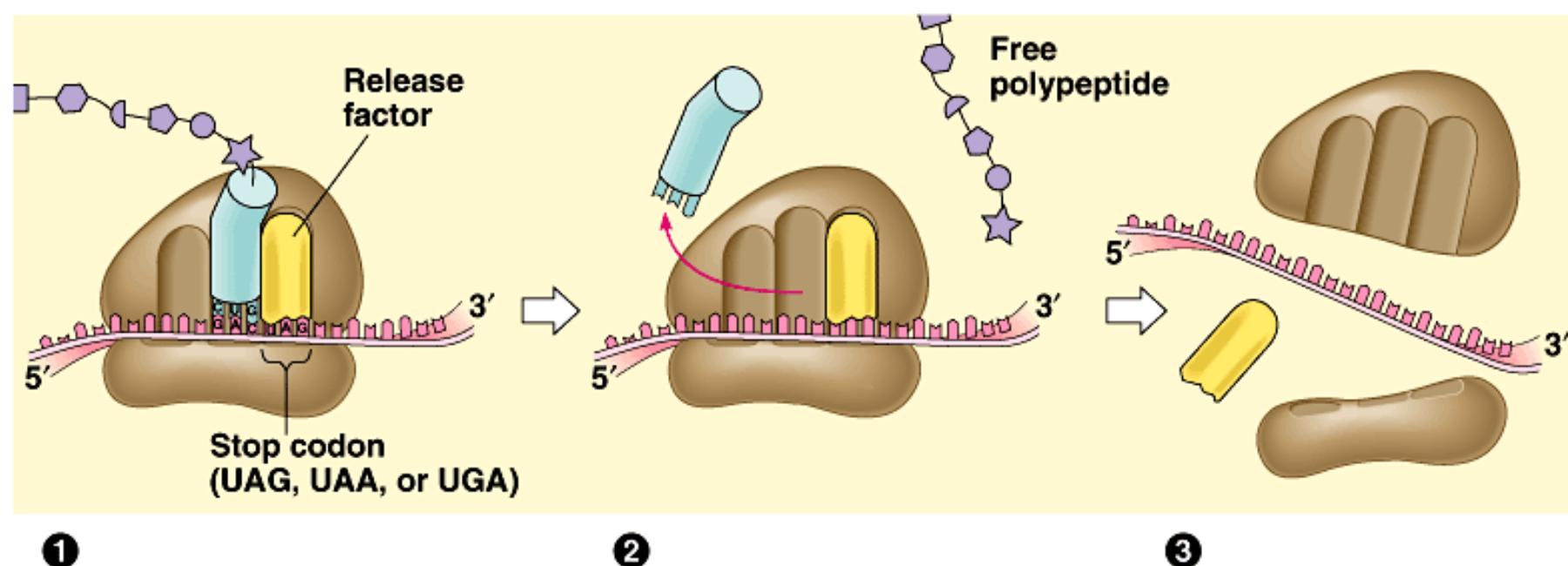
Elongation: Growing a Polypeptide



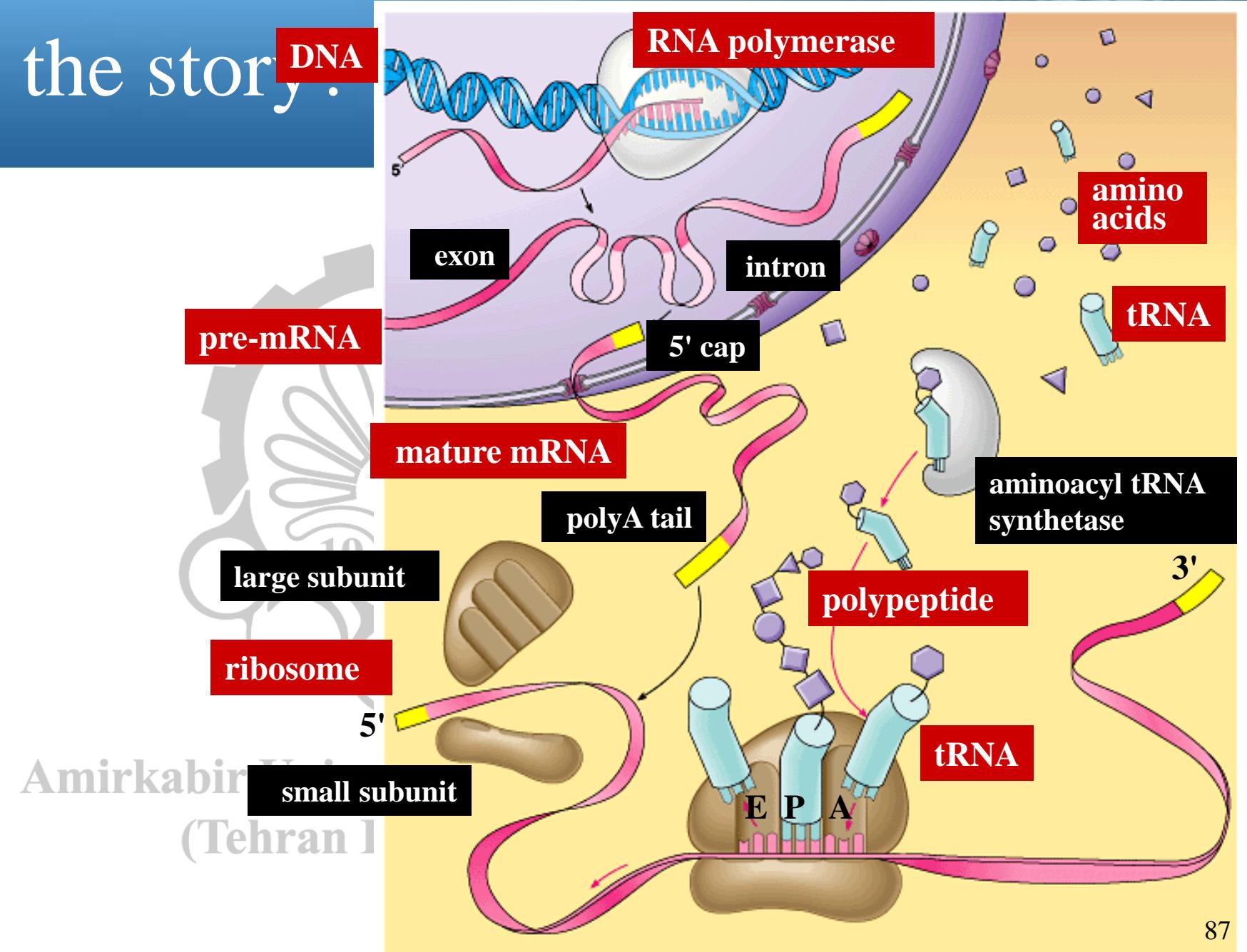
Termination: Release Polypeptide

- Release factor
 - “release protein” bonds to A site
 - bonds water molecule to polypeptide chain

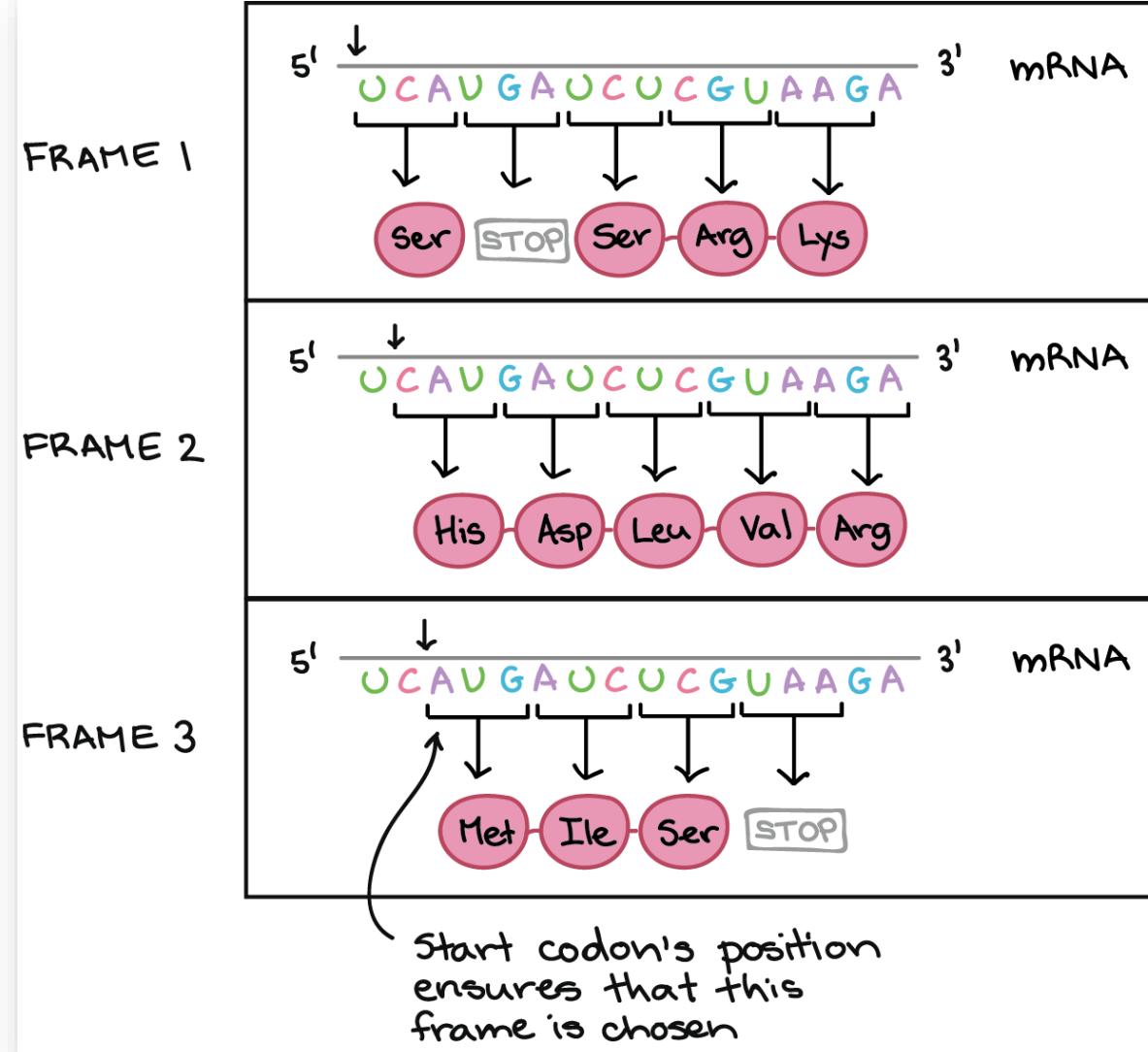
Now what happens to the polypeptide?



Can you tell the story?



mRNA Reading Frames

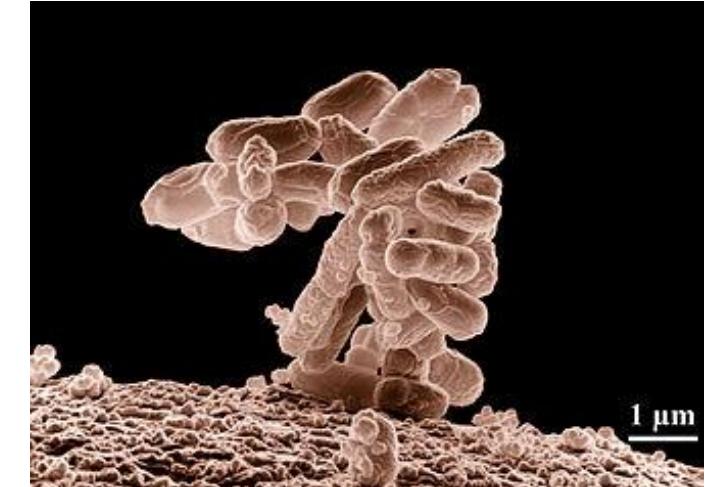


Gene Expression



Genome

- The term **genome** refers to the total genetic information contained in a cell.
- The bacteria E.coli contains about 4,400 genes present on a single chromosome.
- The genome of humans is more complex, with 23 pairs of chromosome containing 6 billion base pairs of DNA, with an estimated 30,000 –40,000 genes.

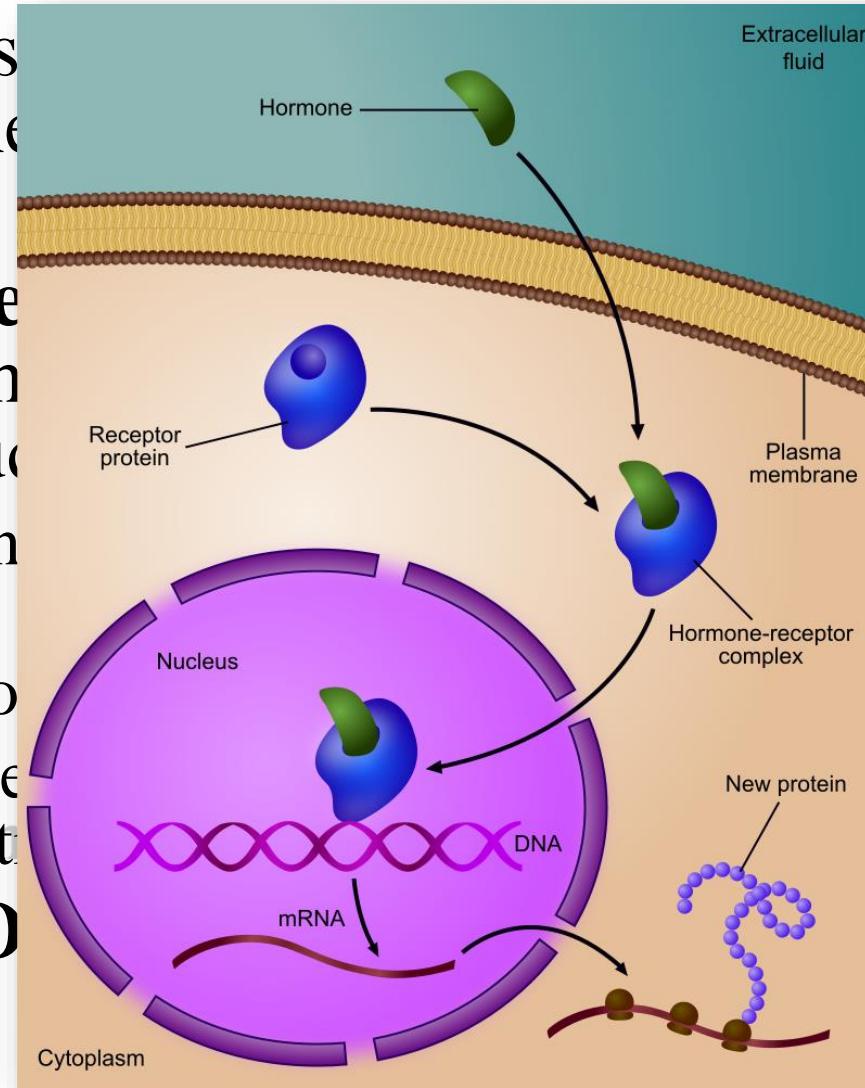


Regulation of Gene Expression

- **Gene expression** is the process by which information from a gene is used in the synthesis of a functional gene product (protein or RNA).
- **Regulation of gene expression**, or **gene regulation**, includes a wide range of mechanisms that are used by cells to **increase** or **decrease** the production of specific gene products.
 - Because the protein synthesis is expensive for a cell, we need gene regulation.
 - Virtually any step of gene expression can be modulated.
 - Gene regulation is essential for *viruses, prokaryotes and eukaryotes* as it increases the versatility and adaptability of an organism
- **Expressed gene (On)**: a used gene

Regulation of Gene Expression

- **Gene expression** is the process by which information from a gene is used in the synthesis of RNA or protein.
- **Regulation of gene expression** refers to a wide range of mechanisms that **increase** or **decrease** the production of gene products.
 - Because the protein product is the target of regulation.
 - Virtually any step of gene expression can be regulated.
 - Gene regulation is especially important because it increases the versatility of gene expression.
- **Expressed gene (Operon)**



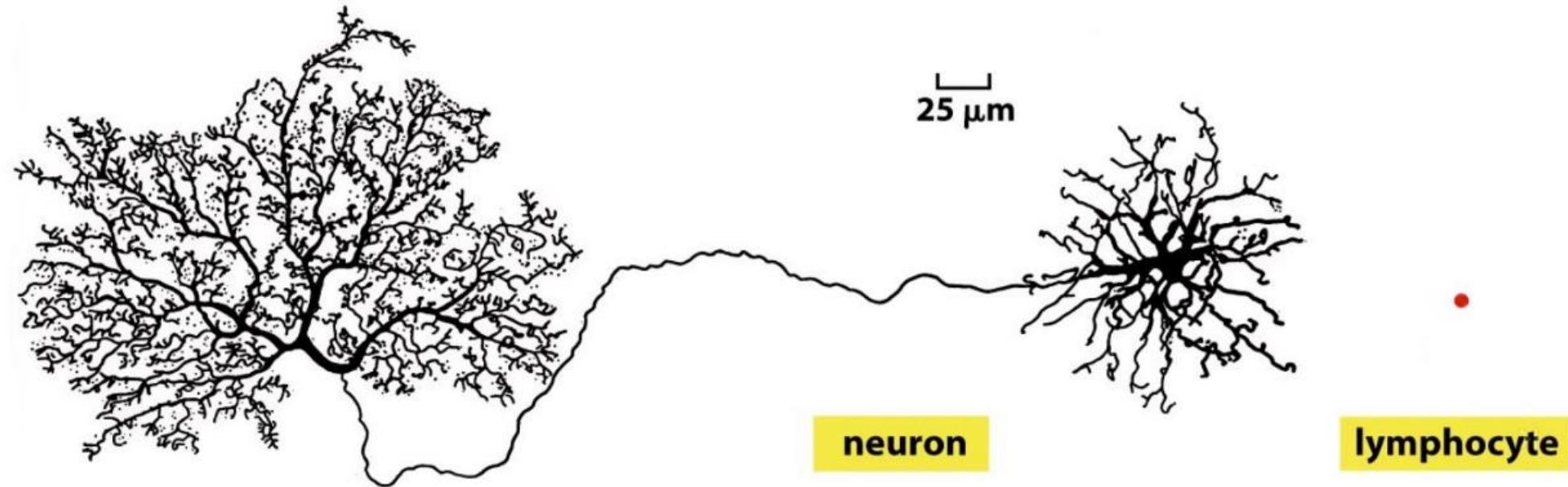
Information from a gene can be used in the synthesis of RNA or protein (protein or RNA).

Regulation, includes a variety of mechanisms to **increase** or **decrease** the production of gene products.

In order to produce a useful protein, we need gene expression to be precisely controlled.

Regulation of gene expression is called gene regulation. It is found in prokaryotes and eukaryotes as it is a common mechanism for controlling gene expression.

Gene Expression

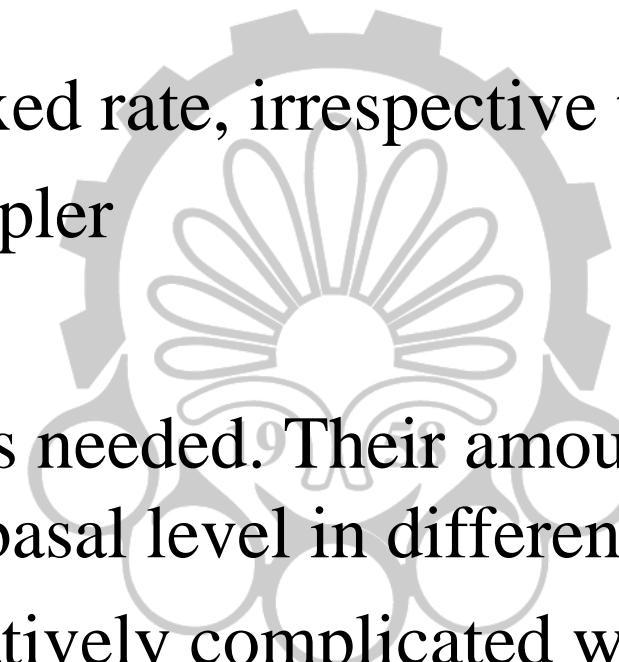


- Both of these cells have the same genome, but they express different RNAs and proteins, and they have very diverse functions.

Amirkabir University of Technology
(Tehran Polytechnic)

Classification of Genes Respect to Expression

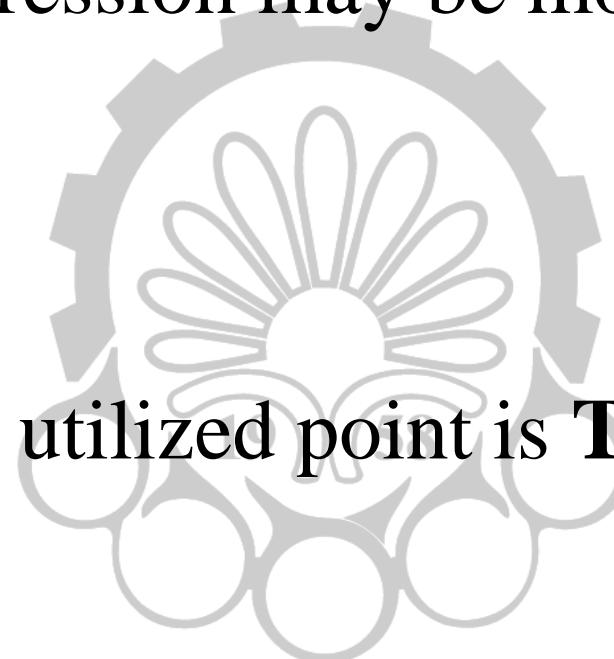
- **Constitutive** genes:
 - Are expressed at a fixed rate, irrespective to the cell condition.
 - Their structure is simpler
- **Controllable** genes:
 - Are expressed only as needed. Their amount may increase or decrease with respect to their basal level in different condition.
 - Their structure is relatively complicated with some response elements



Amirkabir University of Technology
(Tehran Polytechnic)

Regulated Stages of Gene Expression in Prokaryotes

- Any step of gene expression may be modulated:
 - Transcription
 - Translation
 - After translation
- The most extensively utilized point is **Transcription.**

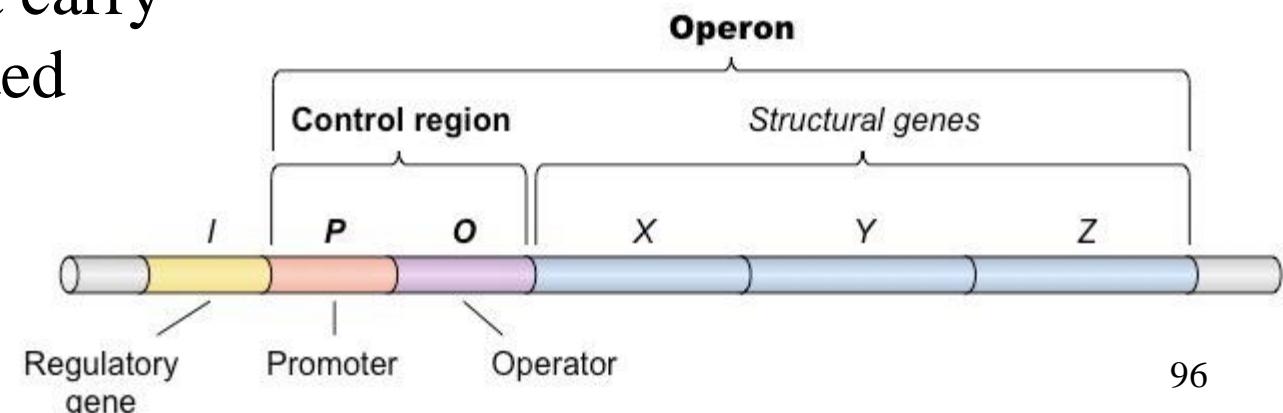


Amirkabir University of Technology
(Tehran Polytechnic)

Operon

- An **operon** is a functioning unit of DNA containing *a cluster of genes* under the control of a single *promoter*.
- An operon is made up of 3 basic DNA components:
 - **Promoter**: a nucleotide sequence that enables a gene to be transcribed.
 - **Operator**: a segment of DNA to which a repressor (to prevent transcription) binds.
 - **Structural genes**: the genes that carry the codons which can be translated into proteins are *co-regulated* by the operon.

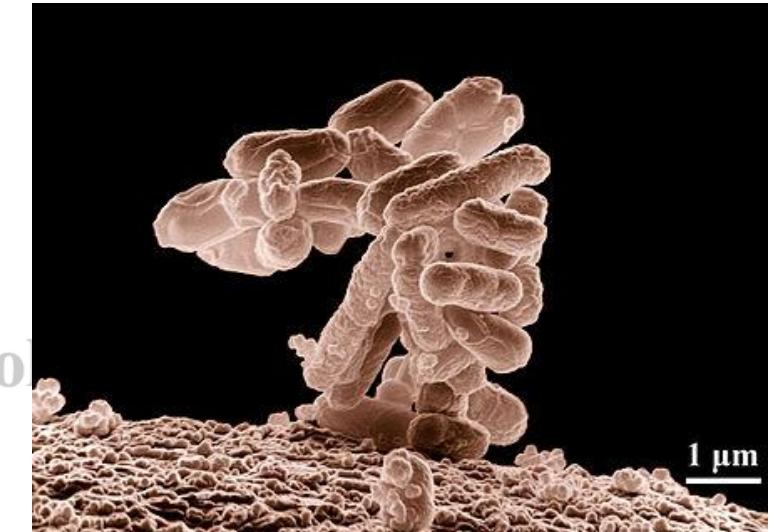
Aminoddini
(Tehran)



Escherichia coli or E. coli

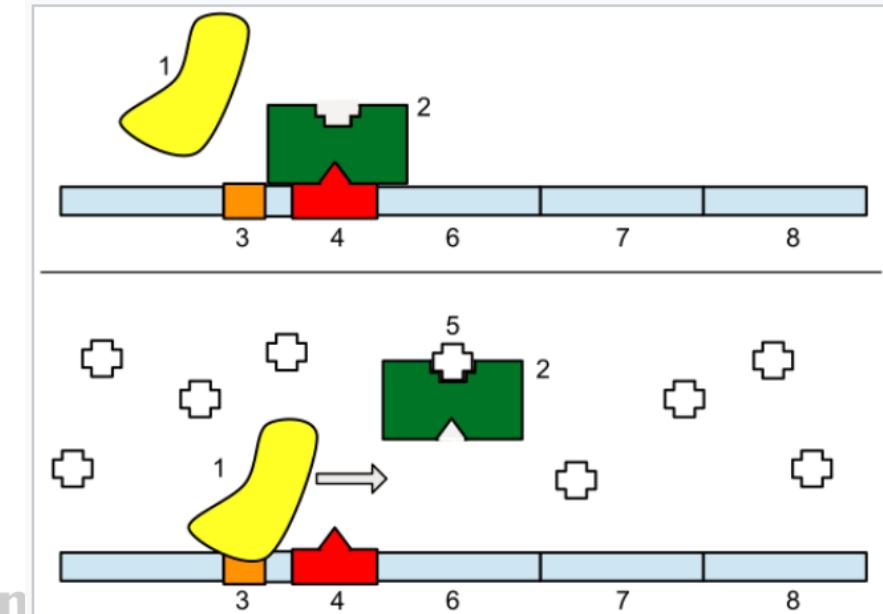
- **Escherichia coli**, also known as E. coli is a type of bacteria that normally lives in your intestines. It's also found in the gut of some animals.
- Most E. coli strains are harmless, but some serotypes can cause serious food poisoning in their hosts.
- In the absence of **glucose** as the carbon source, they eat **lactose** (milk's sugar).

Amirkabir University of Technology
(Tehran Polytechnic)



Lac Operon

- The **lac operon (lactose operon)** is an operon required for the transport and metabolism of lactose in Escherichia.
- Top:**
 - There is no lactose to inhibit the repressor
 - The gene is essentially turned off.
 - The repressor binds to the operator which obstructs the RNA polymerase from binding
- Bottom:**
 - The gene is turned on.
 - Lactose is inhibiting the repressor



1: RNA Polymerase, 2: Repressor, 3: Promoter,
4: Operator, 5: Lactose, 6: lacZ, 7: lacY, 8: lacA.

Gene Expression in Eukaryotes

- Cannot be the same as that in the prokaryotes because of:
 - Genome of the eukaryotes is much larger
 - Human genome consists of 23 pairs of chromosomes containing 1 billion base pairs of DNA
 - Proteins like the repressor proteins influence gene expression.
 - There is no operons here.
- We have more options here for gene expression because transcription and translation are done in separate places.

(Tehran Polytechnic)

Gene Expression in Eukaryotes (Cont.)

- Can be done on following stages:
 - Before transcription
 - Transcription
 - Post-transcriptional modification
 - RNA transport
 - Translation
 - After translation
- Here the most extensively utilized point is Transcription Initiation.

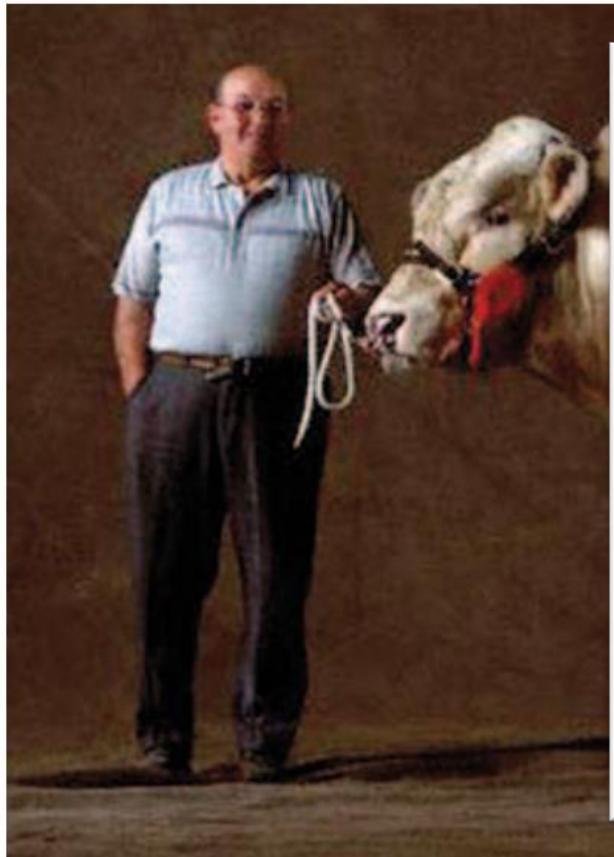


Amirkabir University of Technology
(Tehran Polytechnic)

Mutation



Mutation



(A)



(B)

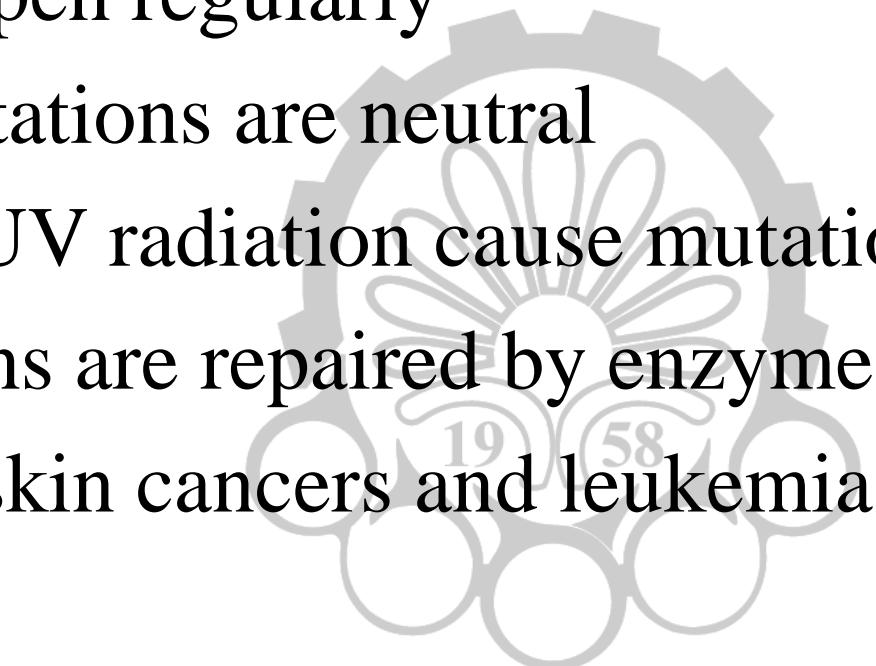


Mutation

- A **mutation** is the *alteration* of the nucleotide sequence of the genome of an organism, virus, or DNA.
- May occur in **somatic** cells (aren't passed to offspring)
- May occur in **gametes** (eggs & sperm) and be passed to offspring
- There are two ways in which DNA can become mutated:
 - Mutations can be inherited.
 - Parent to child
 - Mutations can be acquired.
 - Environmental damage
 - Mistakes when DNA is copied

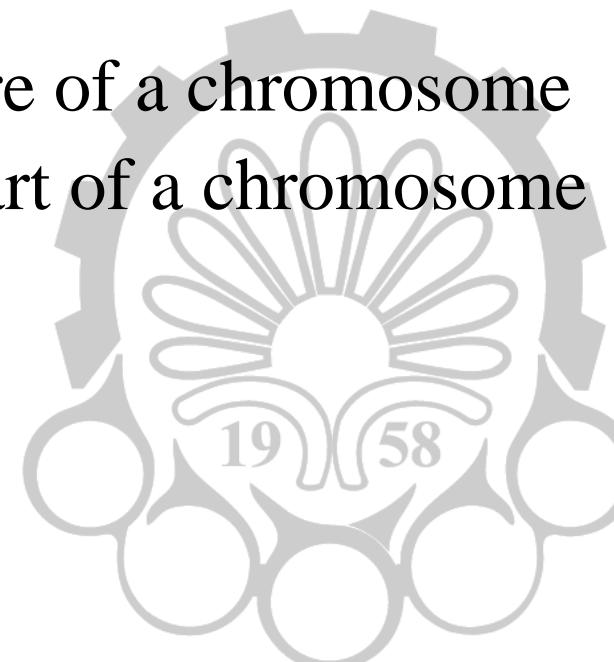
Are Mutations Helpful or Harmful?

- Mutations happen regularly
- Almost all mutations are neutral
- Chemicals & UV radiation cause mutations
- Many mutations are repaired by enzymes
- Some type of skin cancers and leukemia result from somatic mutations
- Some mutations may improve an organism's survival (beneficial)



Chromosome Mutations

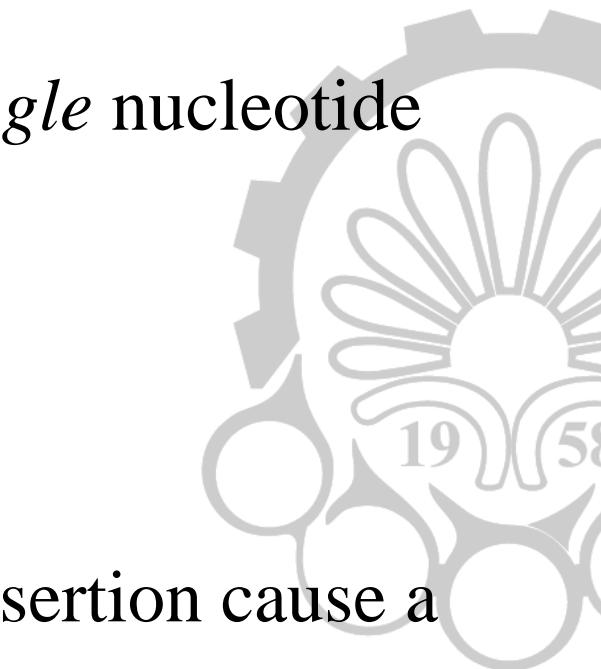
- May Involve:
 - Changing the structure of a chromosome
 - The loss or gain of part of a chromosome
- Five types exist:
 - Deletion
 - Inversion
 - Translocation
 - Nondisjunction
 - Duplication



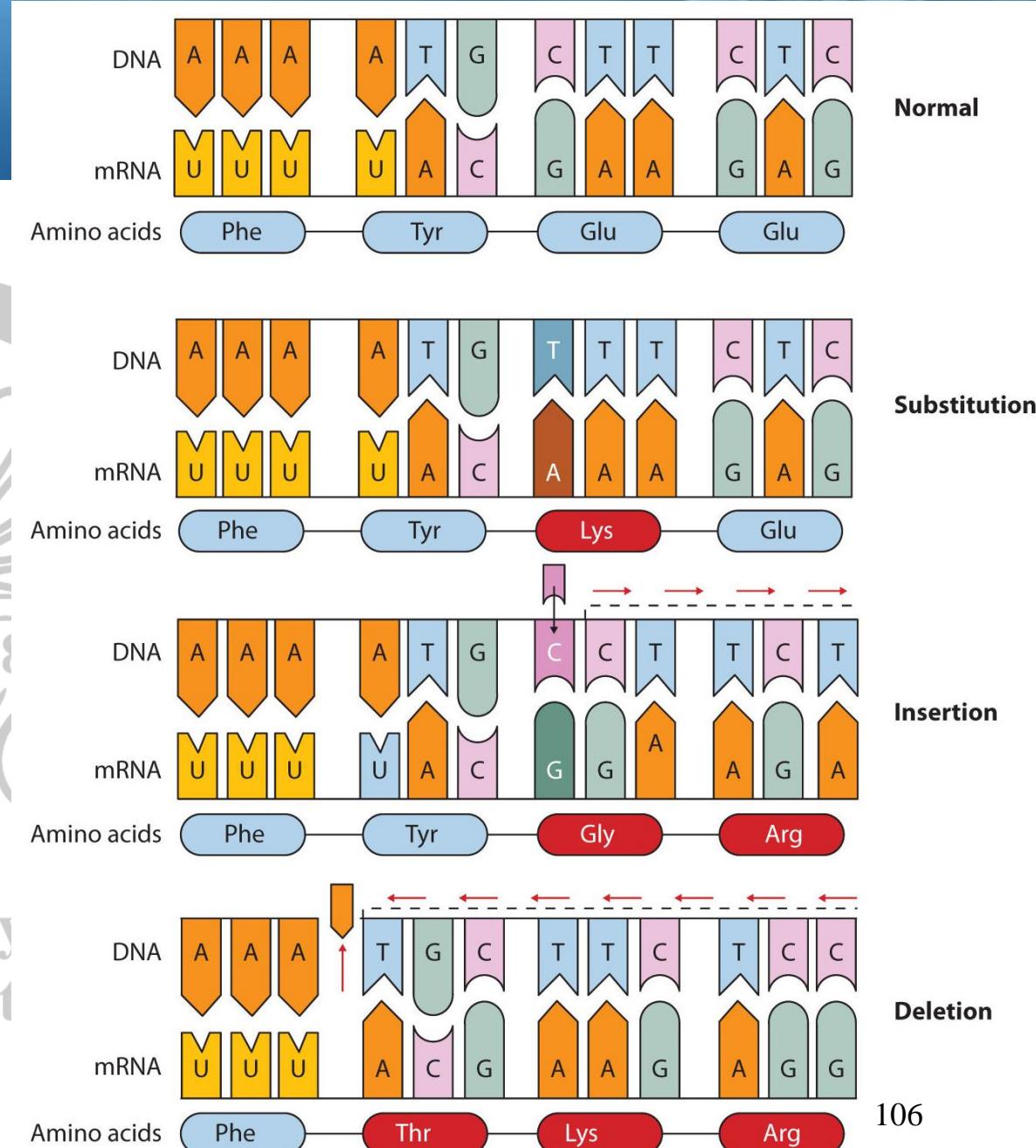
Amirkabir University of Technology
(Tehran Polytechnic)

Types of Gene Mutations

- Point Mutations
 - Change of a *single* nucleotide
 - Includes
 - *Deletion*
 - *Insertion*
 - *Substitution*
 - Deletion and Insertion cause a frame shift in the gene.



Amirkabir University
(Tehran Poly)



More Examples



More Examples



Thanks for your attention

