

# Nucleic acids



## BIOLOGICAL FOUNDATIONS OF BIOINFORMATICS



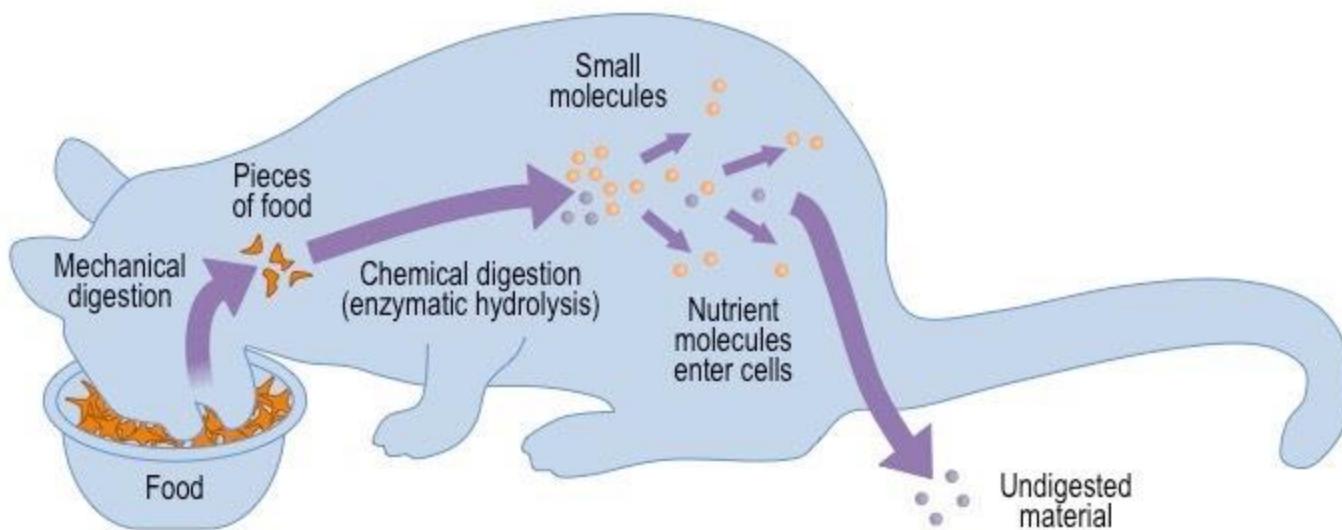
MAG. RER. NAT. ALEXANDRA HUBER

SS 2024

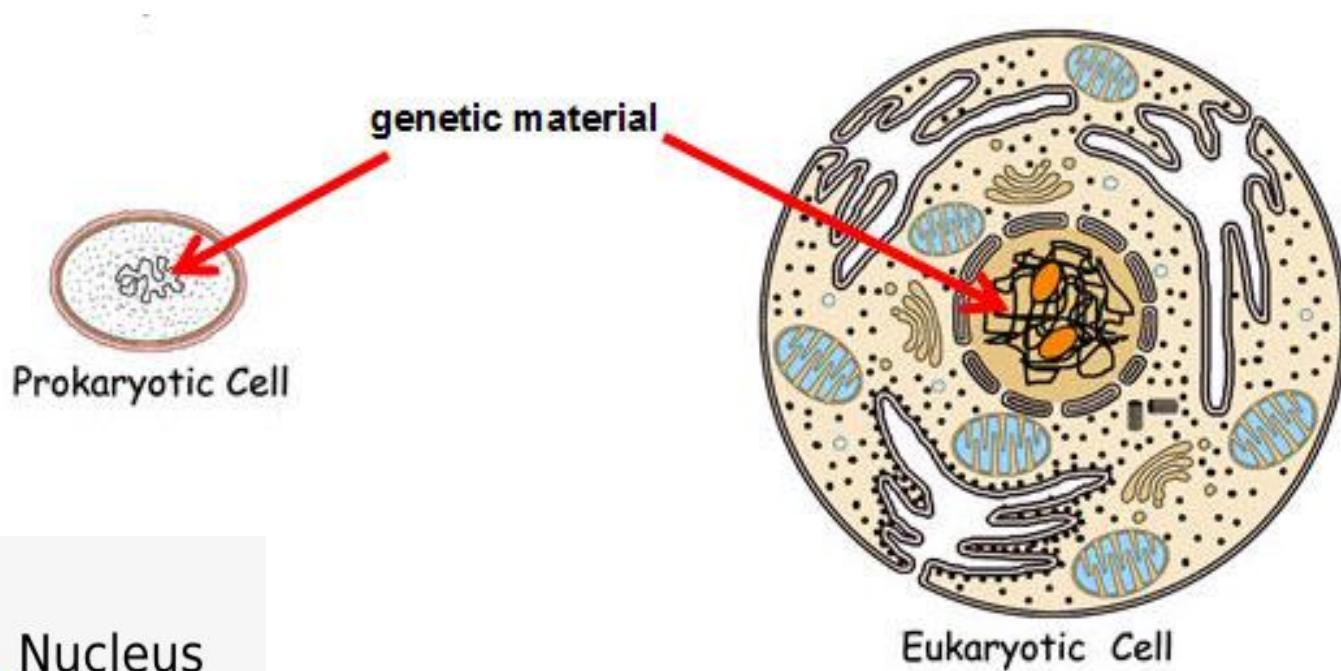
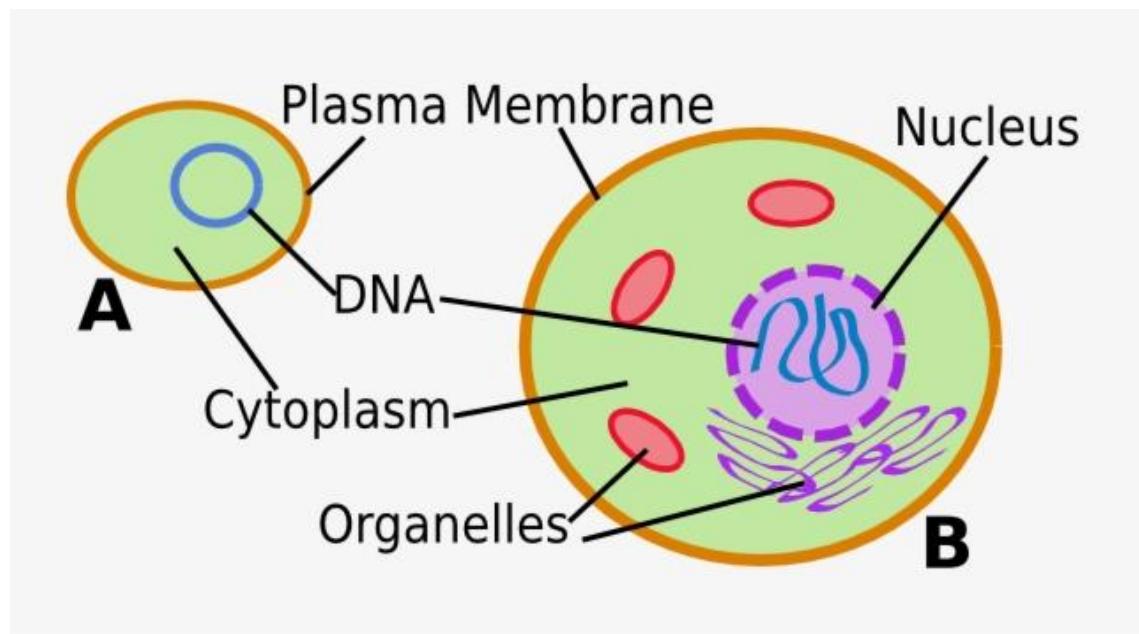


# PROTEIN EXPRESSION

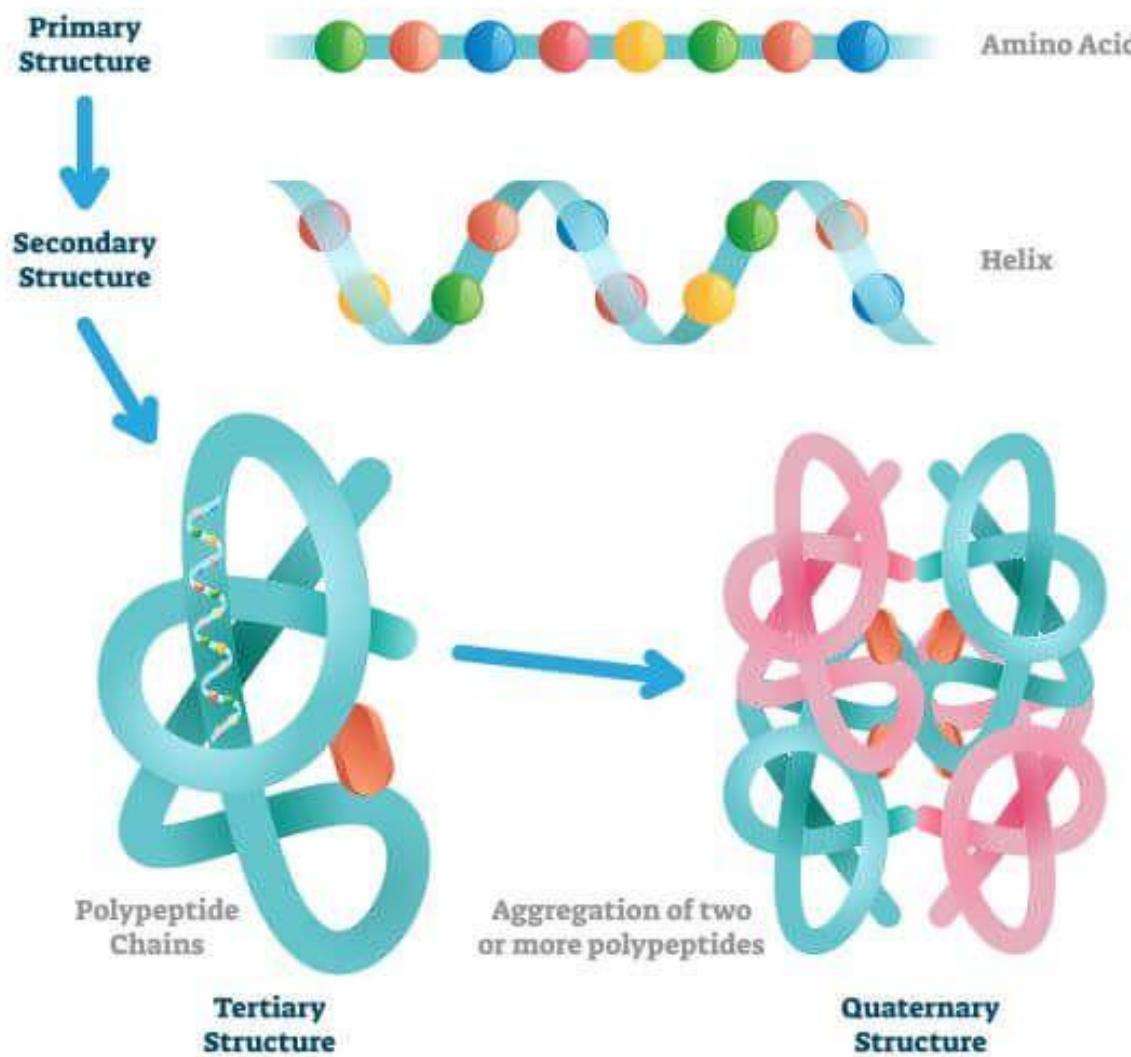
- Proteins are fundamental to life on Earth. They control all biochemical reactions, provide structure to organisms, and transport vital molecules such as oxygen and carbon dioxide, and antibodies defend the organism.
- The process of decoding the instructions stored in DNA to make RNA, which in turn is decoded to make a specific protein is known as the **central dogma of molecular biology**.



| INGESTION                    | DIGESTION                 | ABSORPTION                    | ASSIMILATION                    | ELIMINATION                 |
|------------------------------|---------------------------|-------------------------------|---------------------------------|-----------------------------|
| <i>Taking food into body</i> | <i>Breaking down food</i> | <i>Moving food into cells</i> | <i>Making food part of cell</i> | <i>Removing unused food</i> |



# PROTEIN STRUCTURE





# PROTEIN EXPRESSION

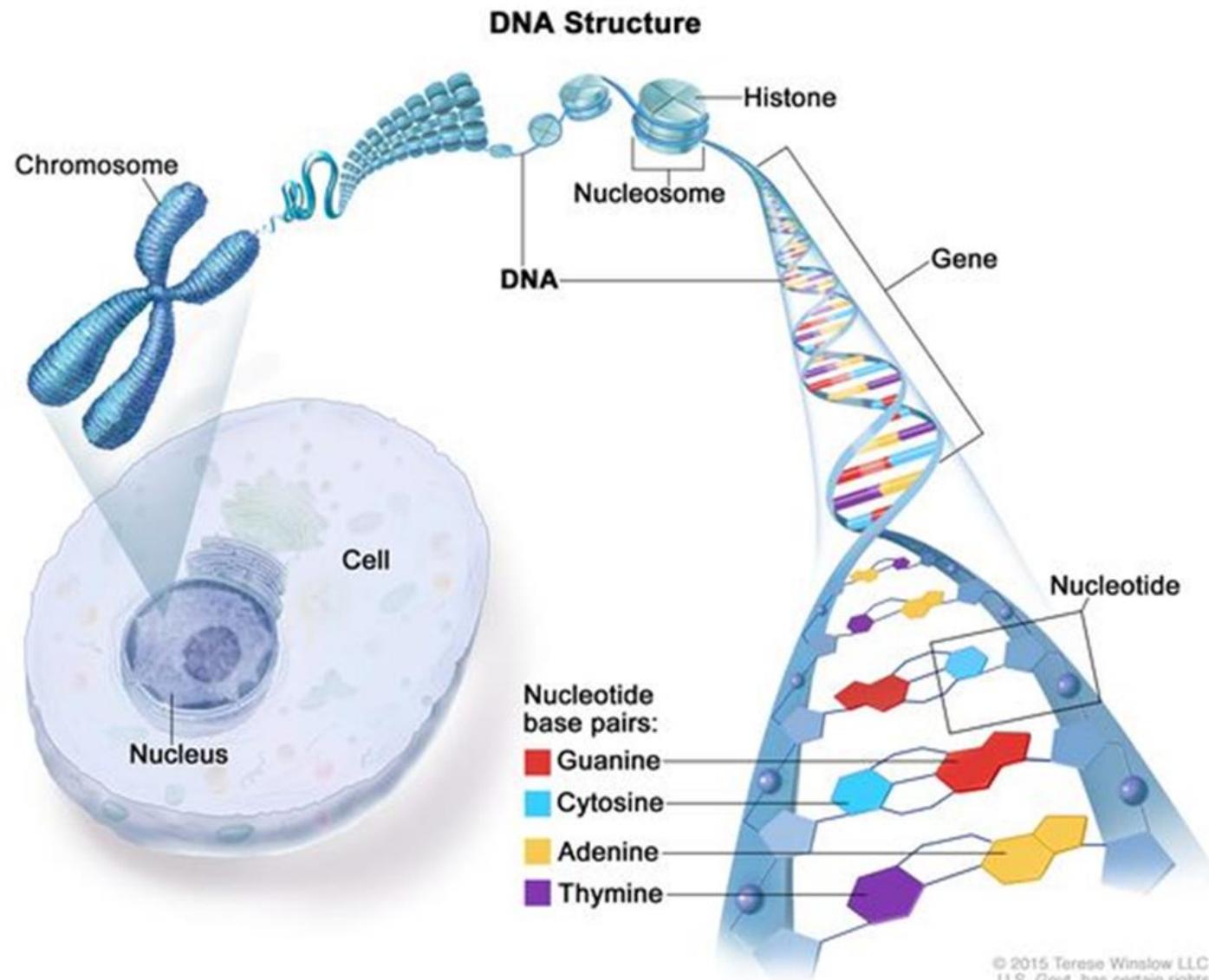
- Proteins are fundamental to life on Earth. They control all biochemical reactions, provide structure to organisms, and transport vital molecules such as oxygen and carbon dioxide, and antibodies defend the organism.
- The process of decoding the instructions stored in DNA to make RNA, which in turn is decoded to make a specific protein is known as the **central dogma of molecular biology**.

## DIFFERENTIATION BY PROTEIN EXPRESSION

- There are more than 200 different cell types in our bodies. The differences between cells in a multicellular organism arise from **differences in gene expression**, not from differences in the cells' genomes.
- During development, cells differentiate from each other. During this process, there are a number of regulatory mechanisms that switch genes on and off. As genes code for a specific protein, by switching genes on and off, the organism can control the proteins made by its different cells.
- This is very important - you don't want a muscle cell secreting amylase, and you don't want your brain cells to start creating myosin. This regulation of genes is controlled by cell-cell communications.

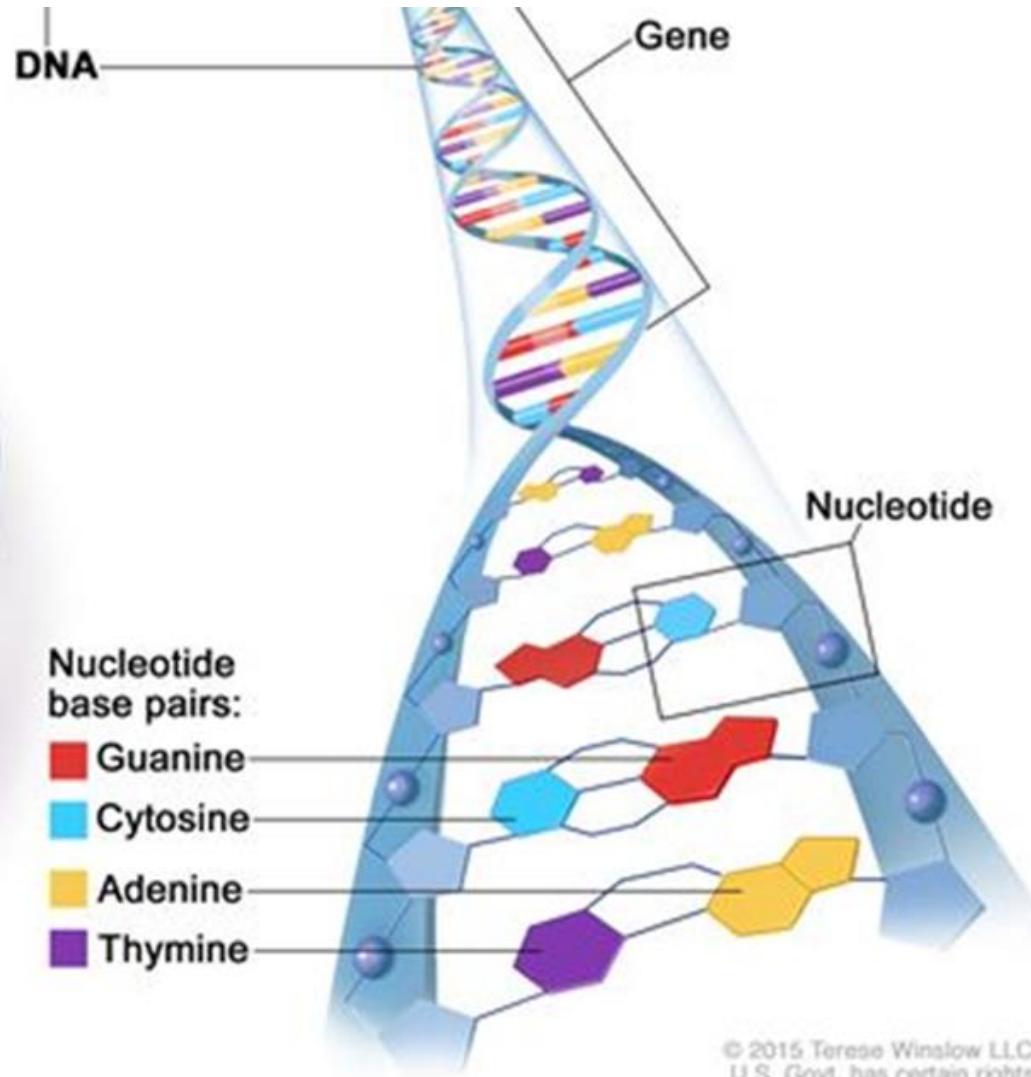
# DNA

- DNA is the chemical name for the molecule that carries genetic instructions in all living things.
- The DNA molecule consists of two strands that wind around one another to form a shape known as a double helix.
- DNA is organized structurally into chromosomes - wound around histones to form nucleosomes as part of those chromosomes.



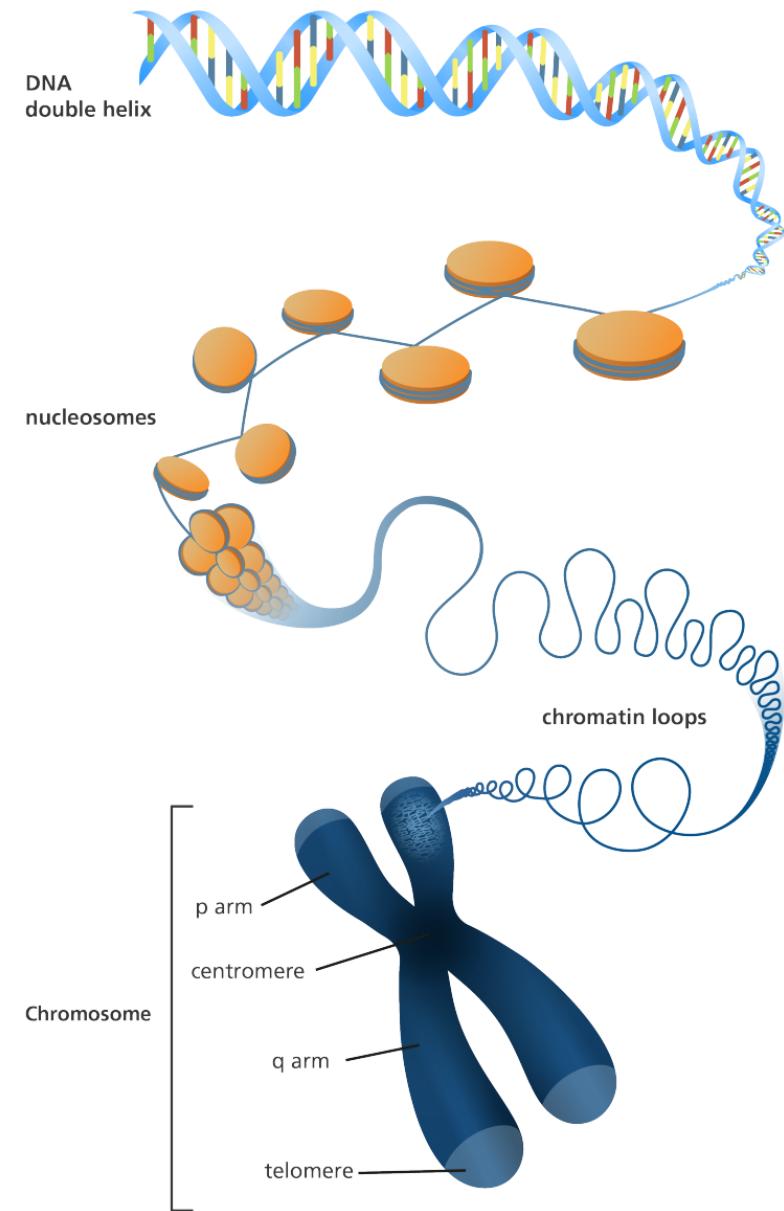
# DNA - BASES

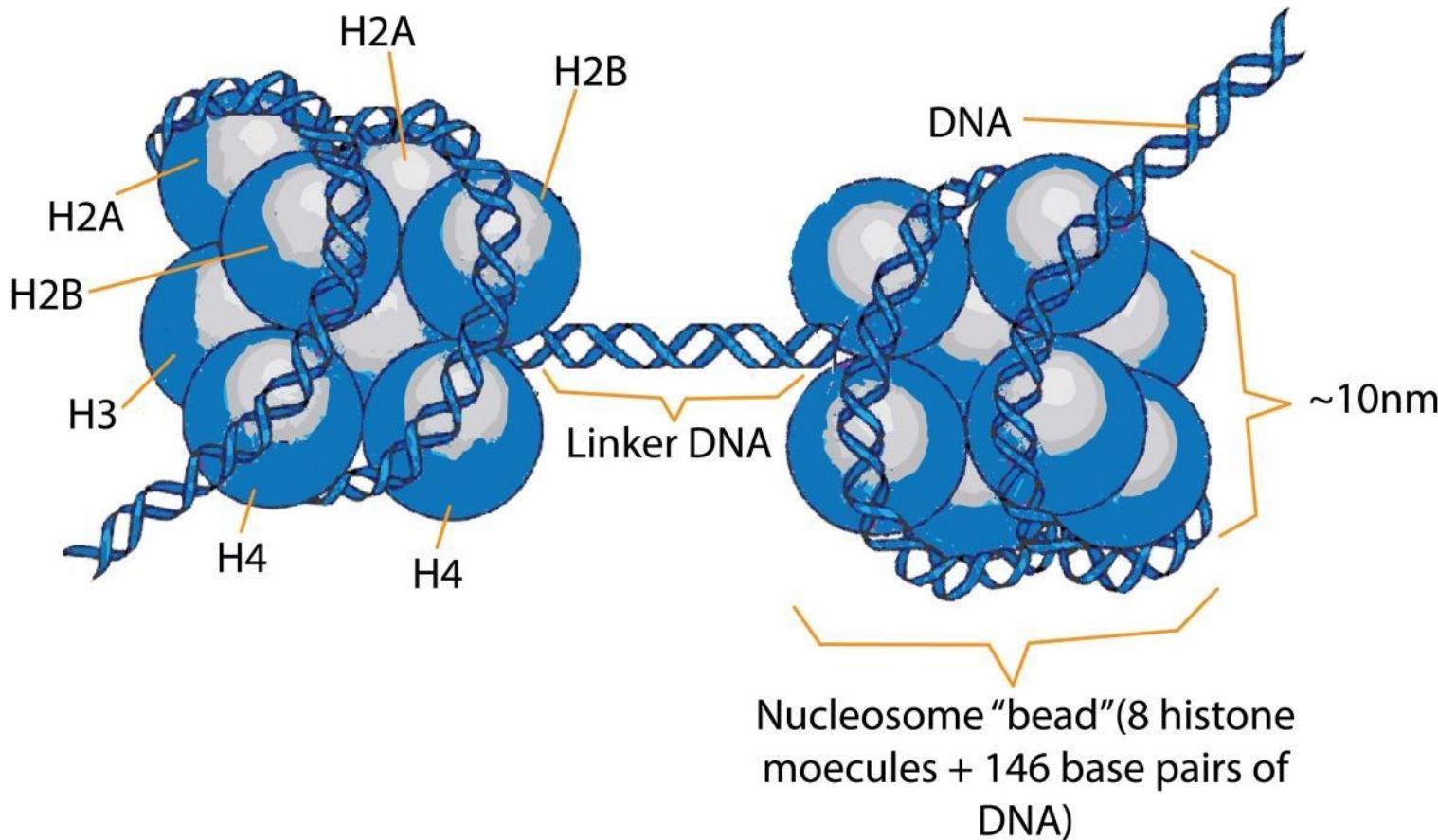
- Each strand has a backbone made of alternating sugar (deoxyribose) and phosphate groups. Attached to each sugar is one of four bases--adenine (A), cytosine (C), guanine (G), thymine (T).
- The two strands are held together by bonds between the bases; adenine bonds with thymine, and cytosine bonds with guanine.



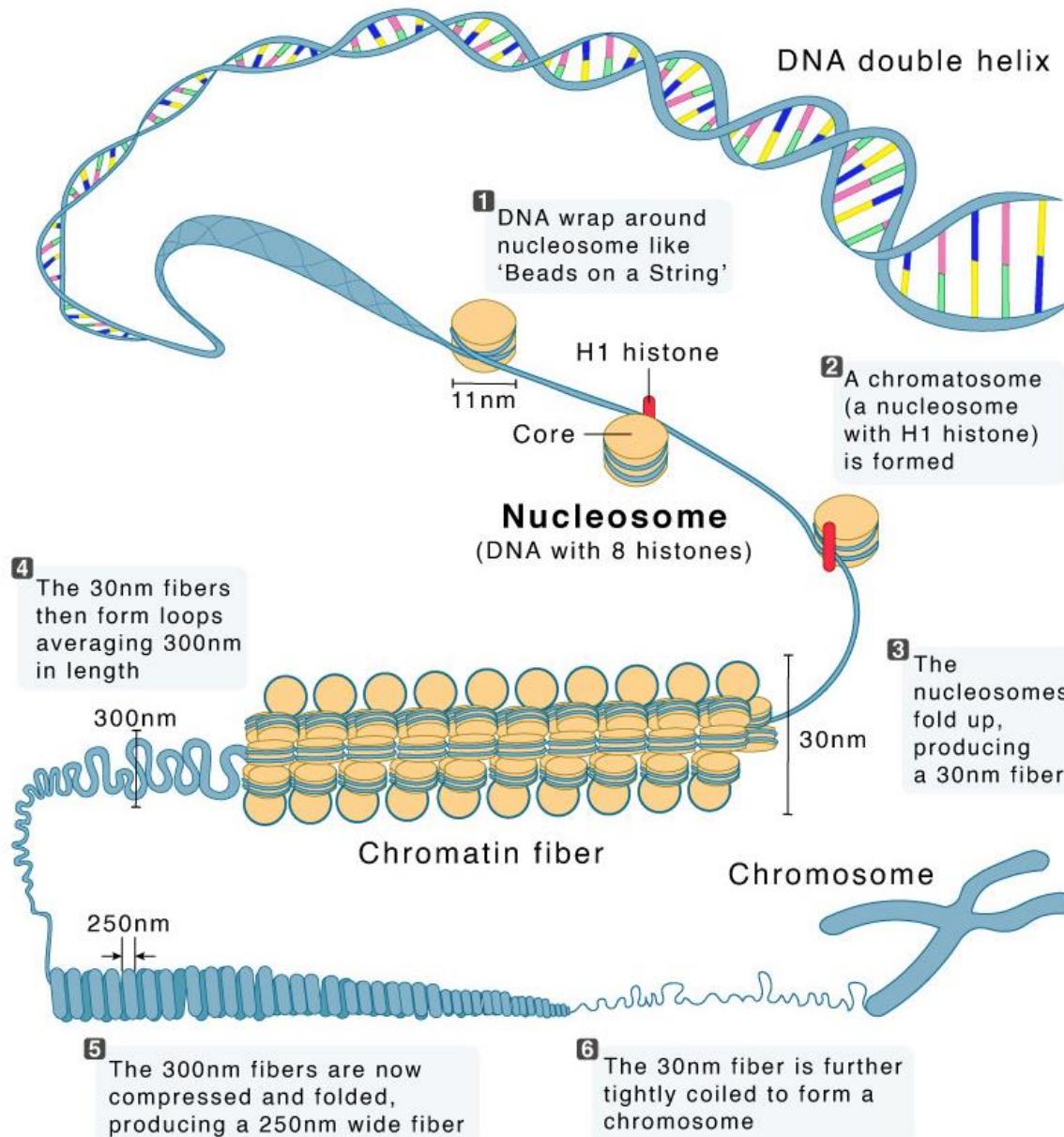
# HISTONES - NUCLEOSOMES

- Histones are proteins acting as spools around which DNA winds to create structural units called nucleosomes.
- Nucleosomes in turn are wrapped into 30-nanometer fibers that form tightly packed chromatin.
- Histones prevent DNA from becoming tangled and protect it from damage.
- In addition, histones play important roles in gene regulation and DNA replication.
- Without histones, unwound DNA in chromosomes would be very long. For example, each human cell has about 1.8 meters of DNA if completely stretched out.



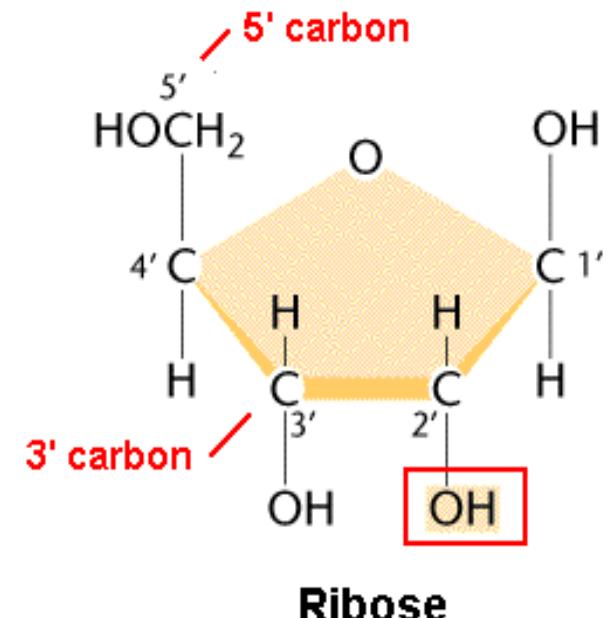
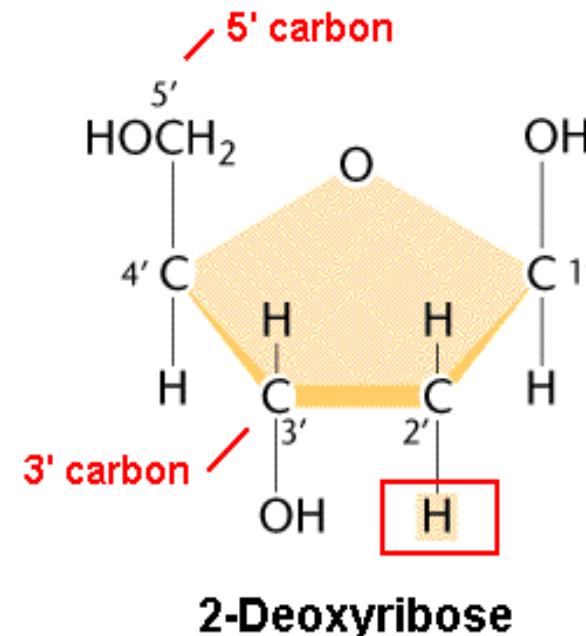


## Nucleosomes



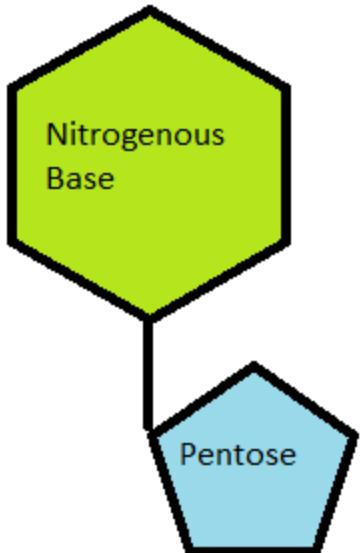
# DNA – DEOXYRIBONUCLEIC ACID

- central information storage system of organisms
- The sequence of the bases along the backbones serves as instructions for assembling RNA molecules and protein.
- The name comes from its structure, which is a sugar and phosphate backbone with bases sticking out from it.
- “Deoxyribo” refers to the sugar and nucleic acid refers to the phosphate and the bases.

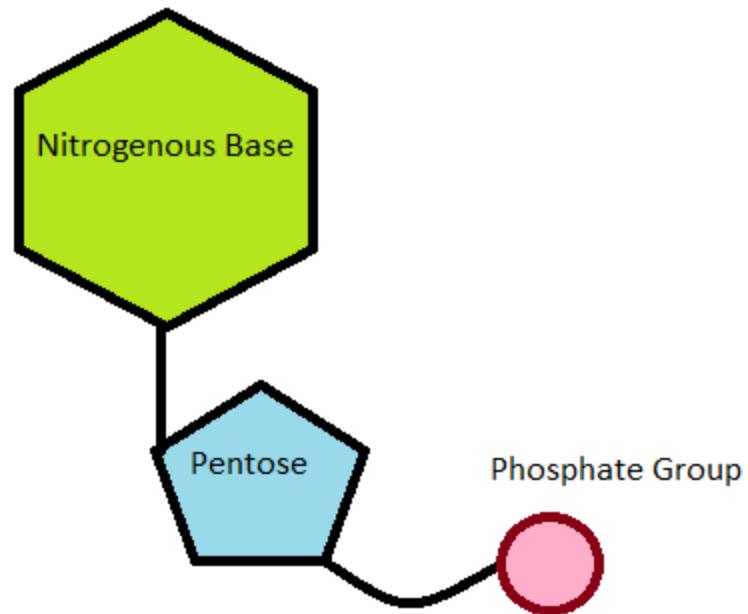


# DNA BUILDING BLOCKS

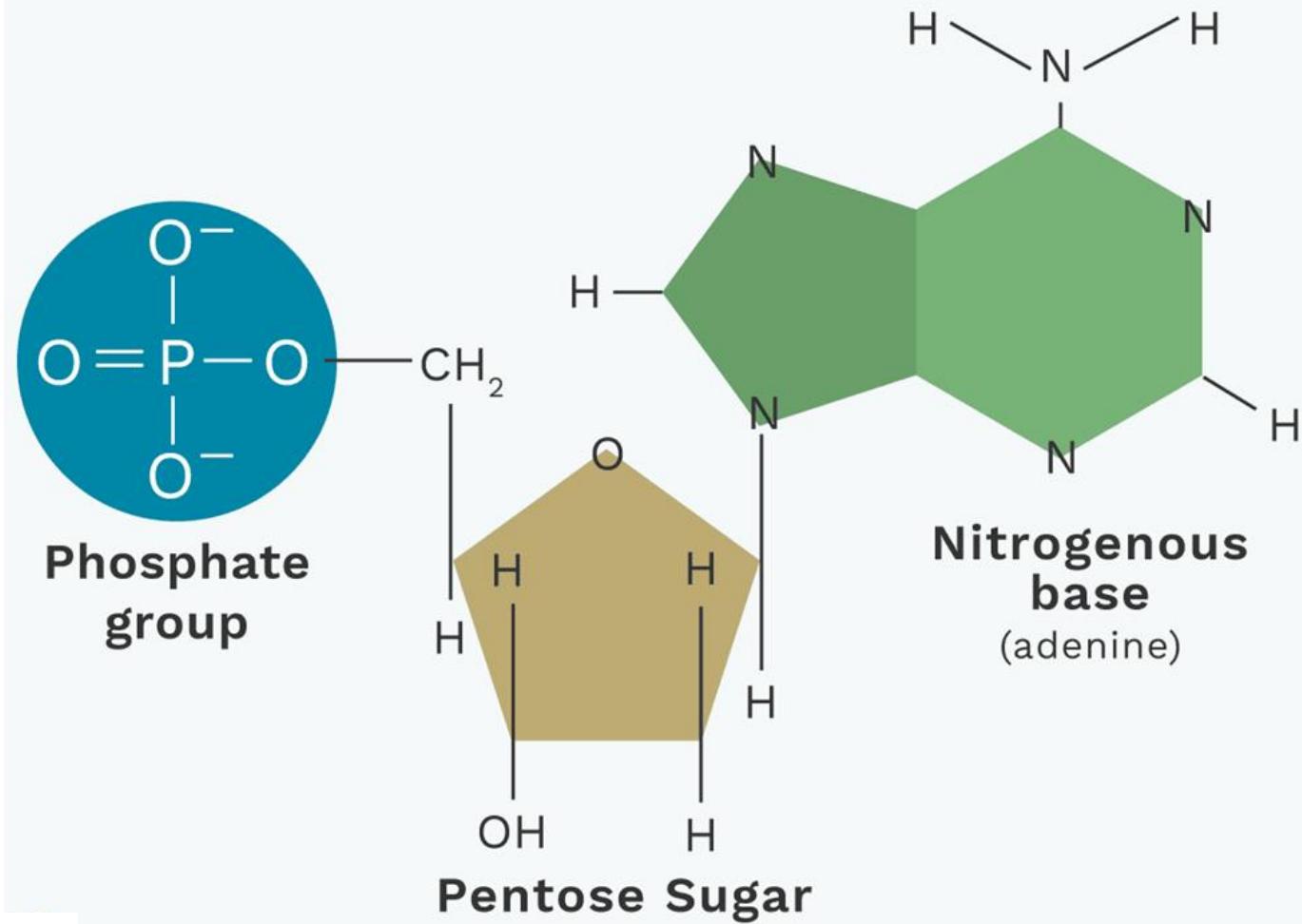
Nucleoside



Nucleotide

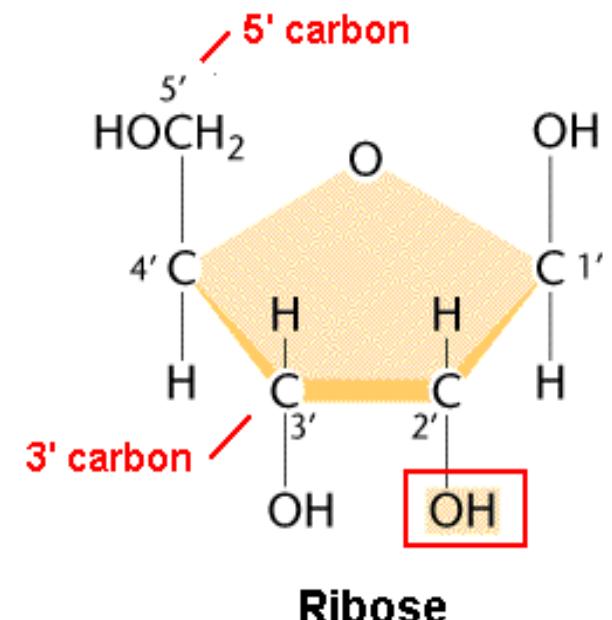
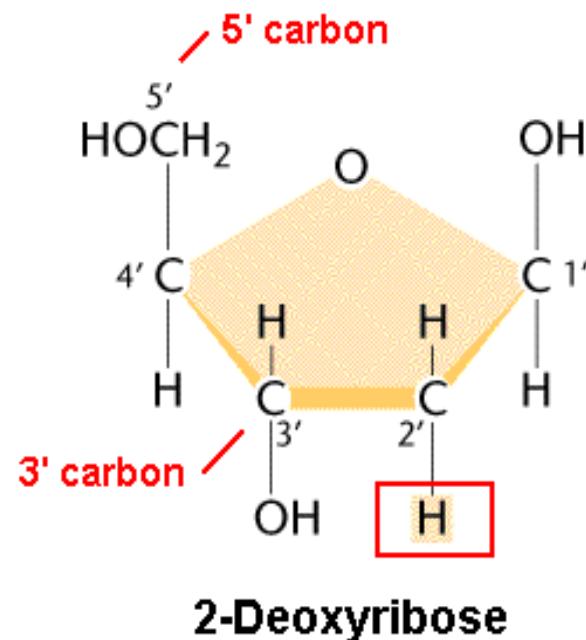


## 3 Parts of a Nucleotide



## DNA - SUGAR

- central information storage system of organisms
- The sequence of the bases along the backbones serves as instructions for assembling RNA molecules and protein.
- The name comes from its structure, which is a sugar and phosphate backbone with bases sticking out from it.
- “Desoxyribo” refers to the sugar and nucleic acid refers to the phosphate and the bases.

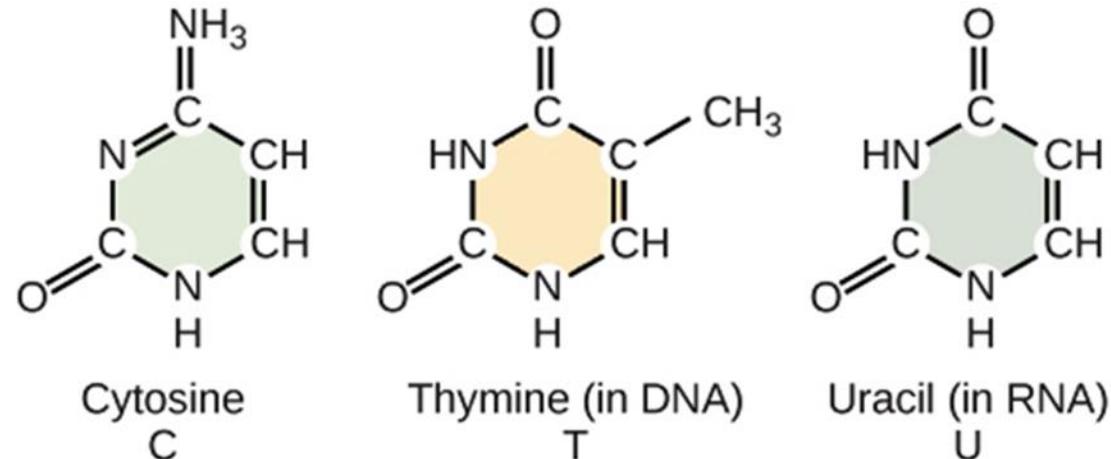


## Nitrogenous Bases

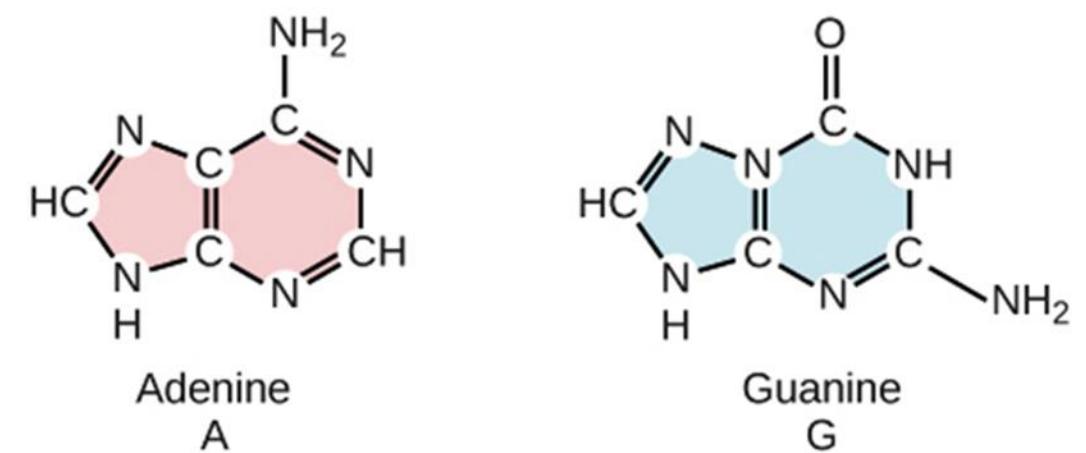
### DNA - BASES

- The bases go by the names of adenine, cytosine, thymine, and guanine, otherwise known as A, C, T and G.
- DNA is a polymer of the four bases, which allows enormous complexity to be encoded by the pattern of those bases, one after another.

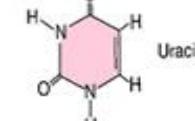
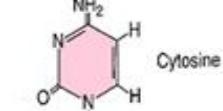
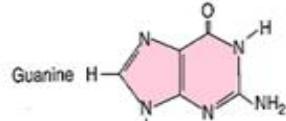
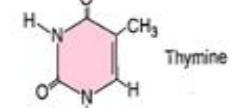
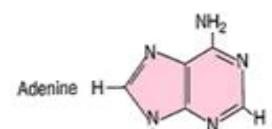
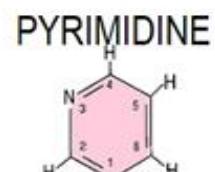
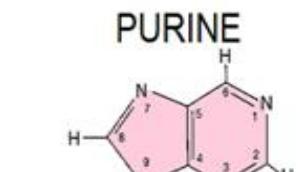
#### Pyrimidines



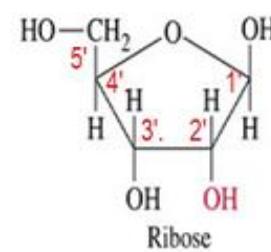
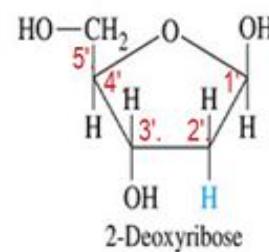
#### Purines



### Nitrogenous Base.



### Pentose Sugar

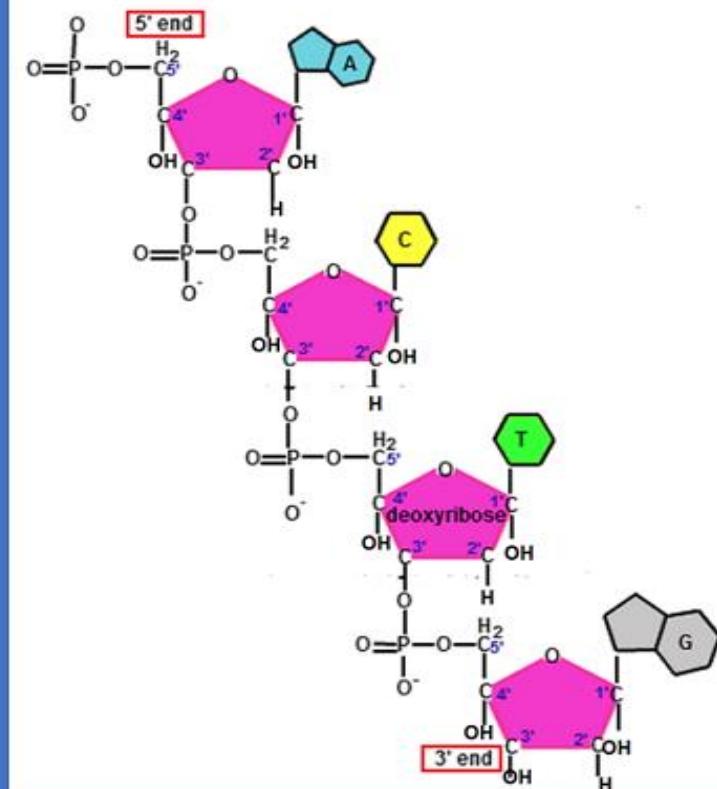


### Nucleoside

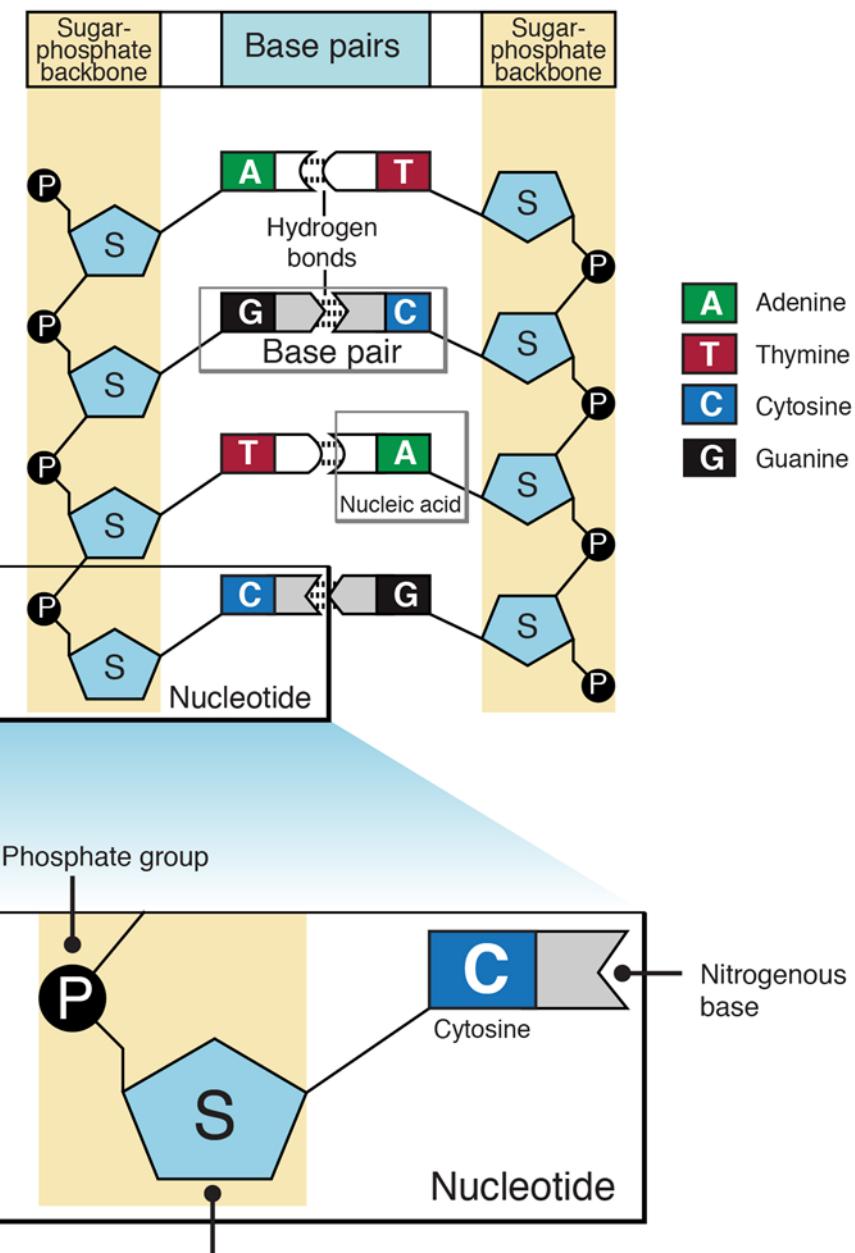


Glycosidic bond

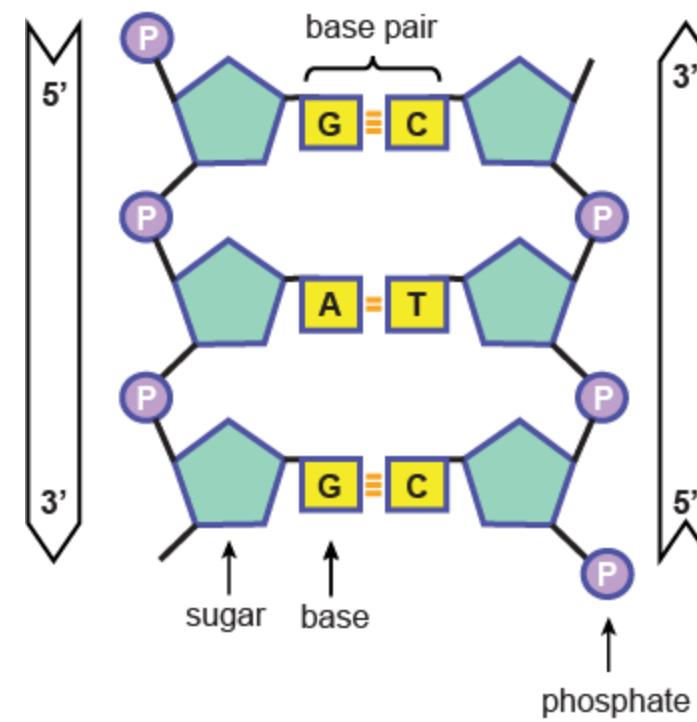
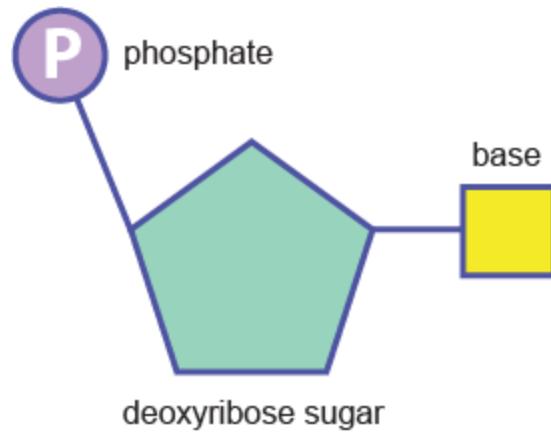
### Polynucleotide



## Deoxyribonucleic Acid (DNA)



# NUCLEOTIDES



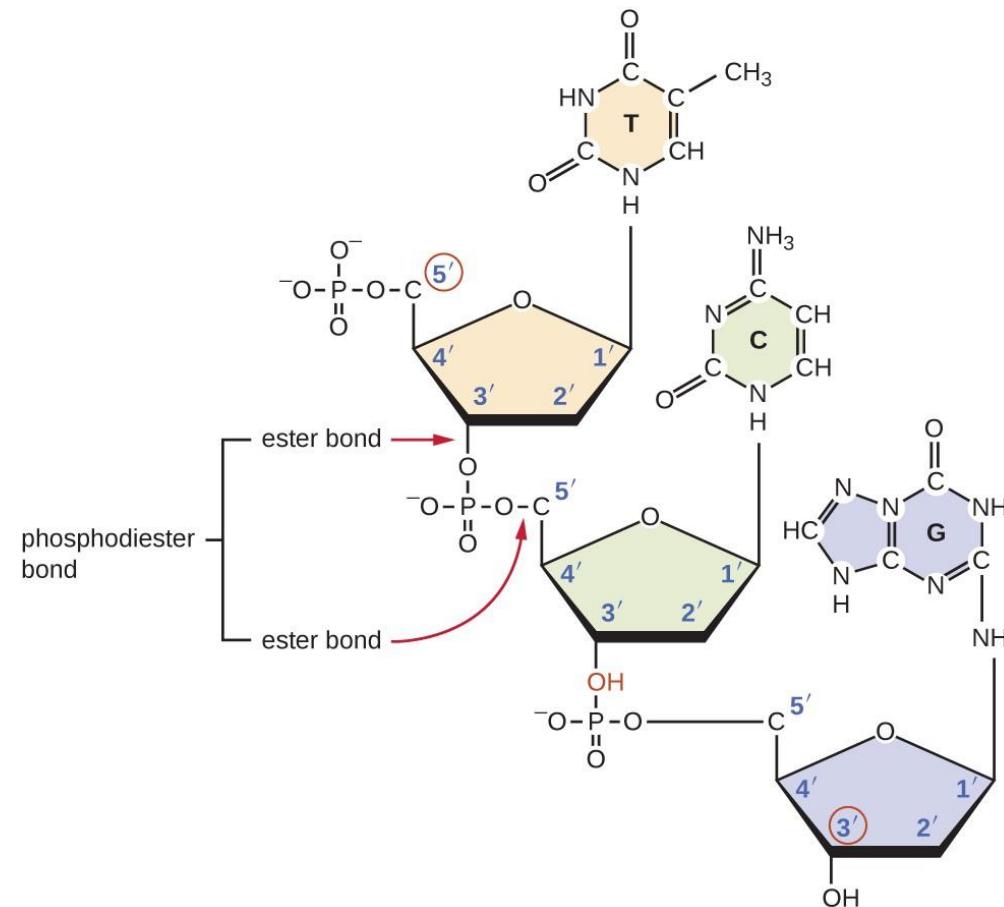
# DNA BONDS

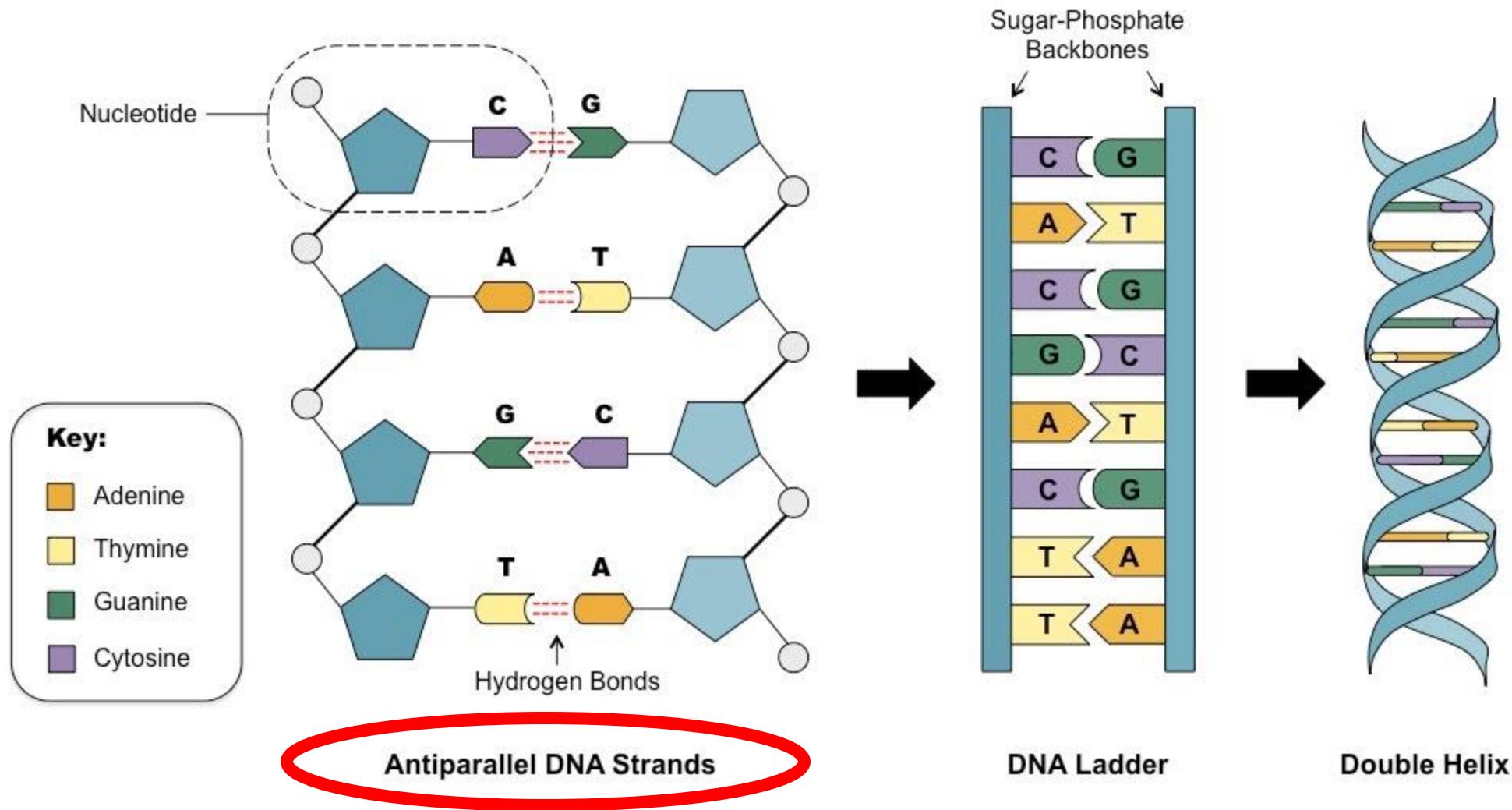
Carbons of the deoxyribose sugar are numbered clockwise, starting from the oxygen atom.

The phosphate group (attached to the 5'-C of the sugar) joins with the hydroxyl (OH) group attached to the 3'-C of the sugar.

This results in a phosphodiester bond between the two nucleotides.

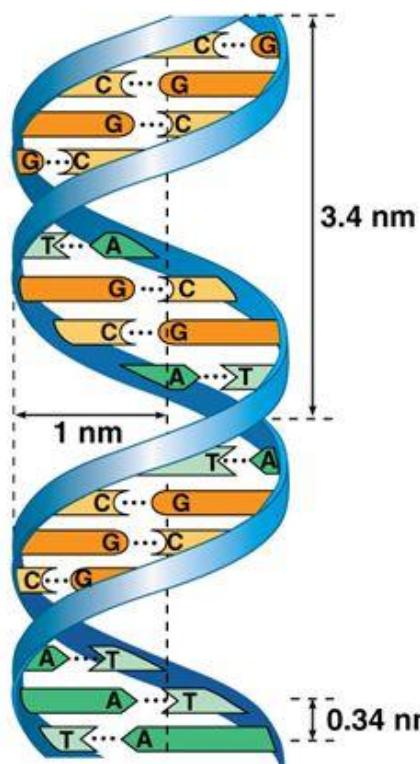
DNA polymerase always moves in a 5' to 3' direction.



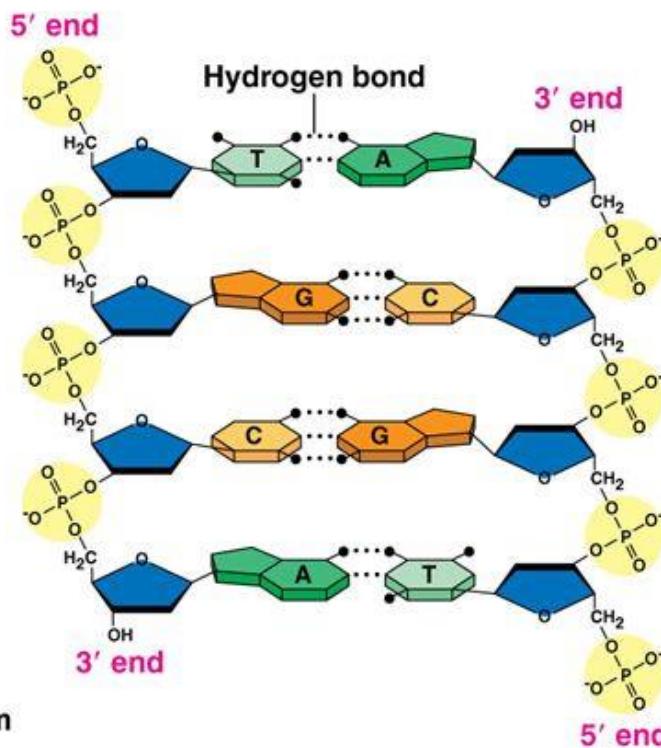


# Structure of DNA

**Antiparallel:** one strand ( $5' \rightarrow 3'$ ), other strand runs in opposite, upside-down direction ( $3' \rightarrow 5'$ )



(a) Key features of DNA structure

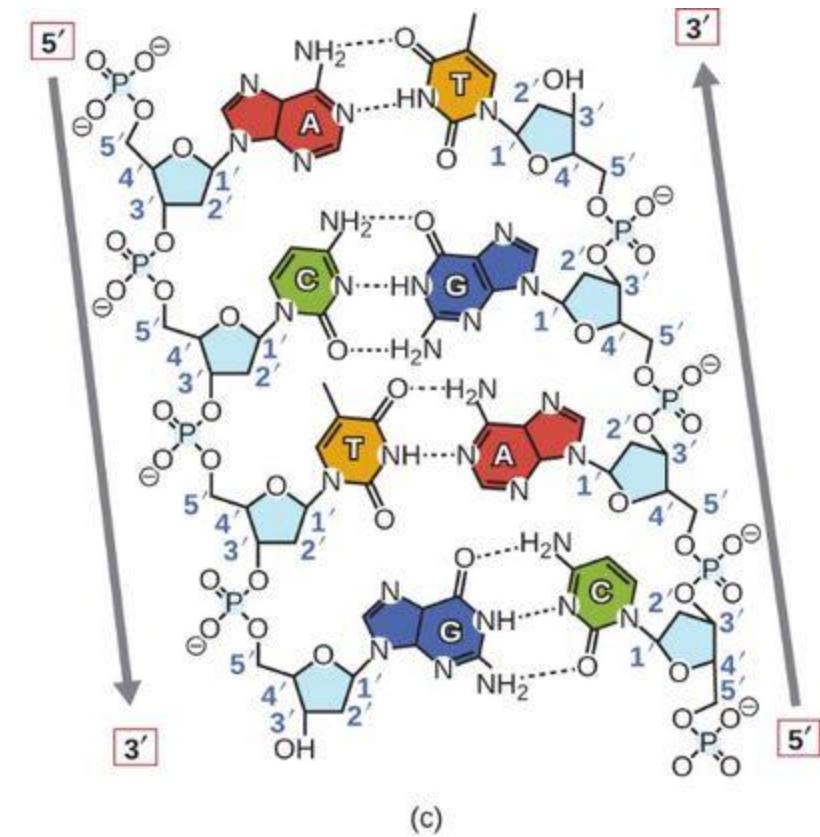
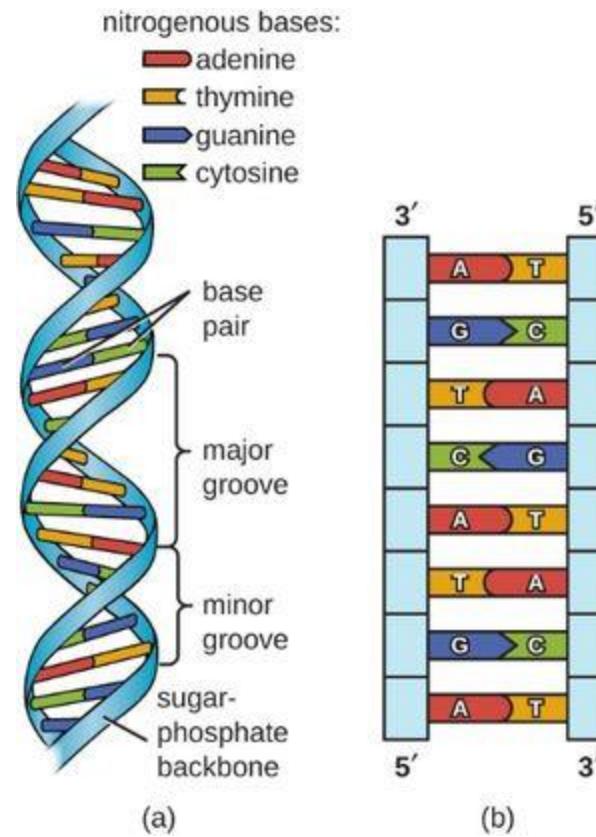


(b) Partial chemical structure



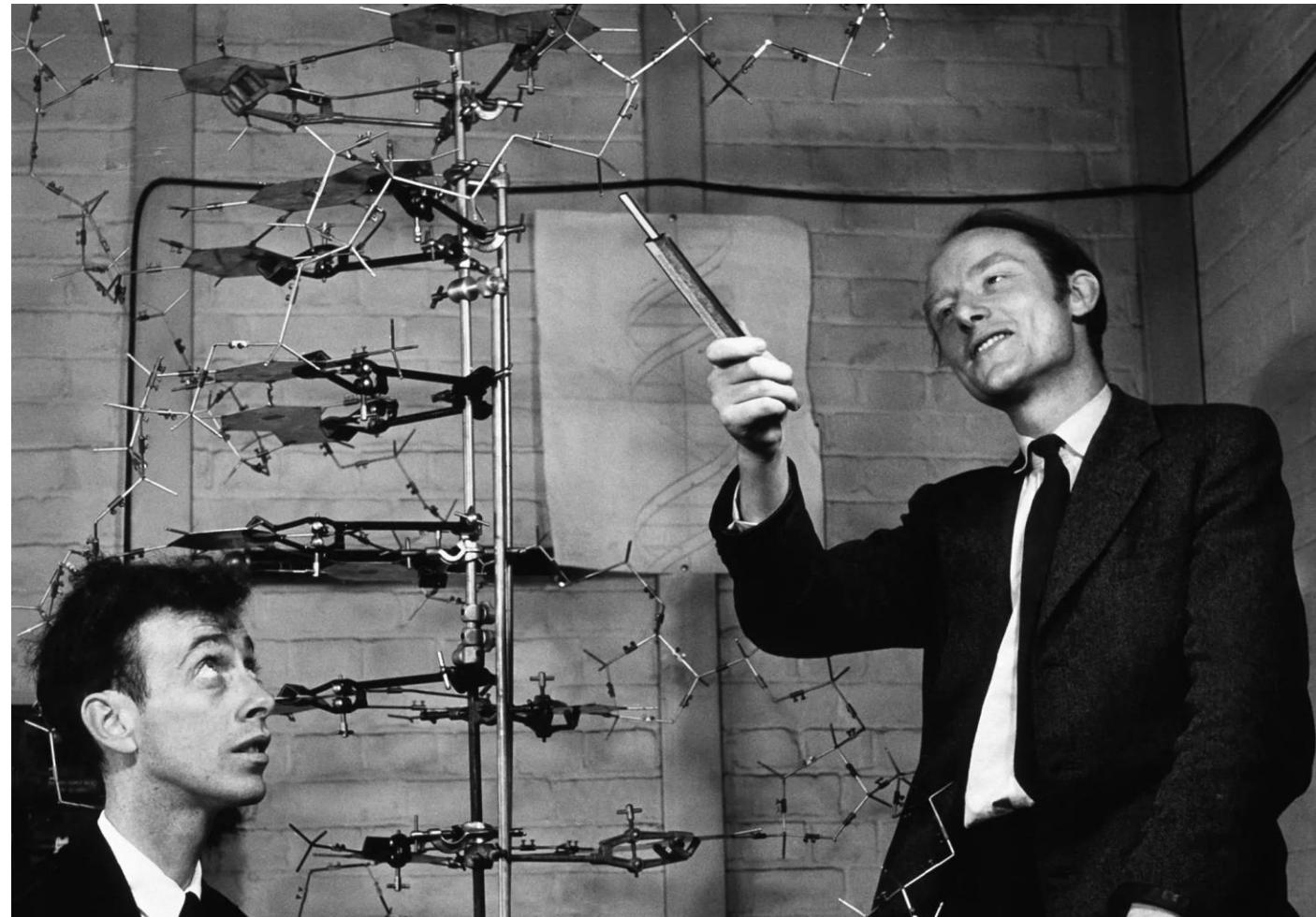
(c) Space-filling model

- The major and minor grooves are opposite each other, and each runs continuously along the entire length of the DNA molecule. They arise from the antiparallel arrangement of the two backbone strands. The grooves are actual structural features of the molecule. The grooves are important in the attachment of DNA Binding Proteins involved in replication and transcription.
- The major groove is wider than the minor groove. These grooves allow proteins to bind to and recognize DNA sequences from the outside of the helix. The grooves expose the edges of each base pair located inside the helix, which allows proteins to chemically recognize specific DNA sequences.



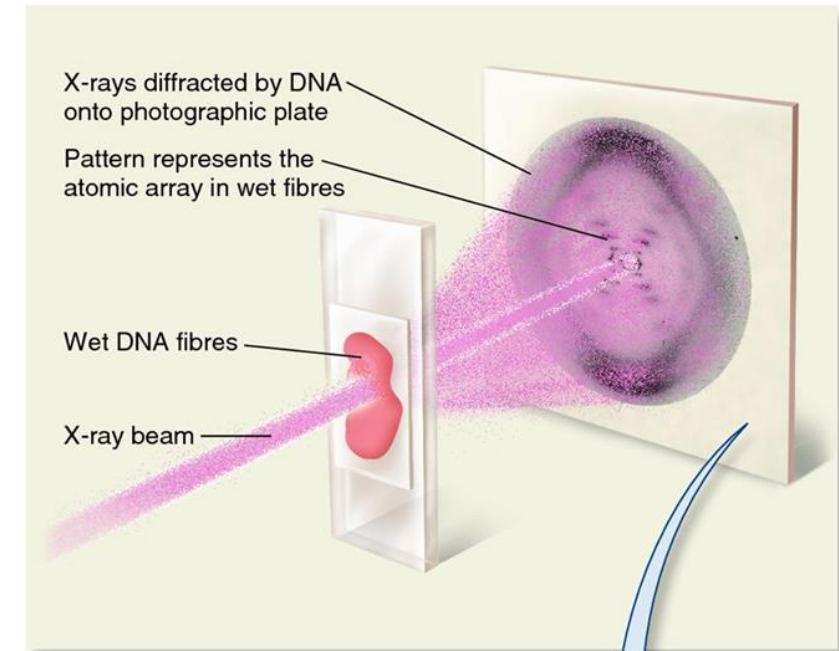
## **WATSON-CRICK-MODEL**

The Nobel Prize in Physiology or Medicine in 1962 was awarded to James Watson, Francis Crick and Maurice Wilkins for their discovery of the molecular structure of DNA, which helped solve one of the most important of all biological riddles.



# X-RAY DIFFRACTION

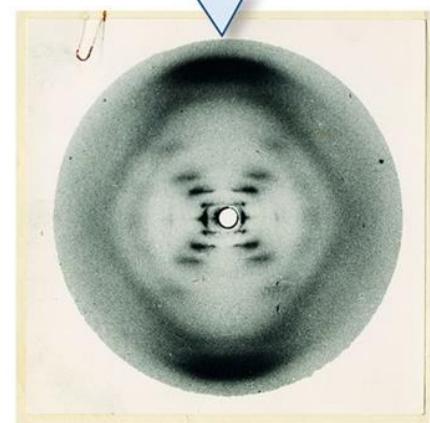
- DNA fibres produce diffraction patterns when exposed to X-rays.
- These patterns can be visualized in the form of pictures. This was a very difficult process. Rosalind Franklin started working on X-ray diffraction techniques to understand the structure of DNA. She have taken X-ray diffraction picture of DNA fibre (famously known as photograph 51) which later contributed towards the discovery of double helical structure of DNA by Watson and Crick.



(a) The method of X-ray diffraction



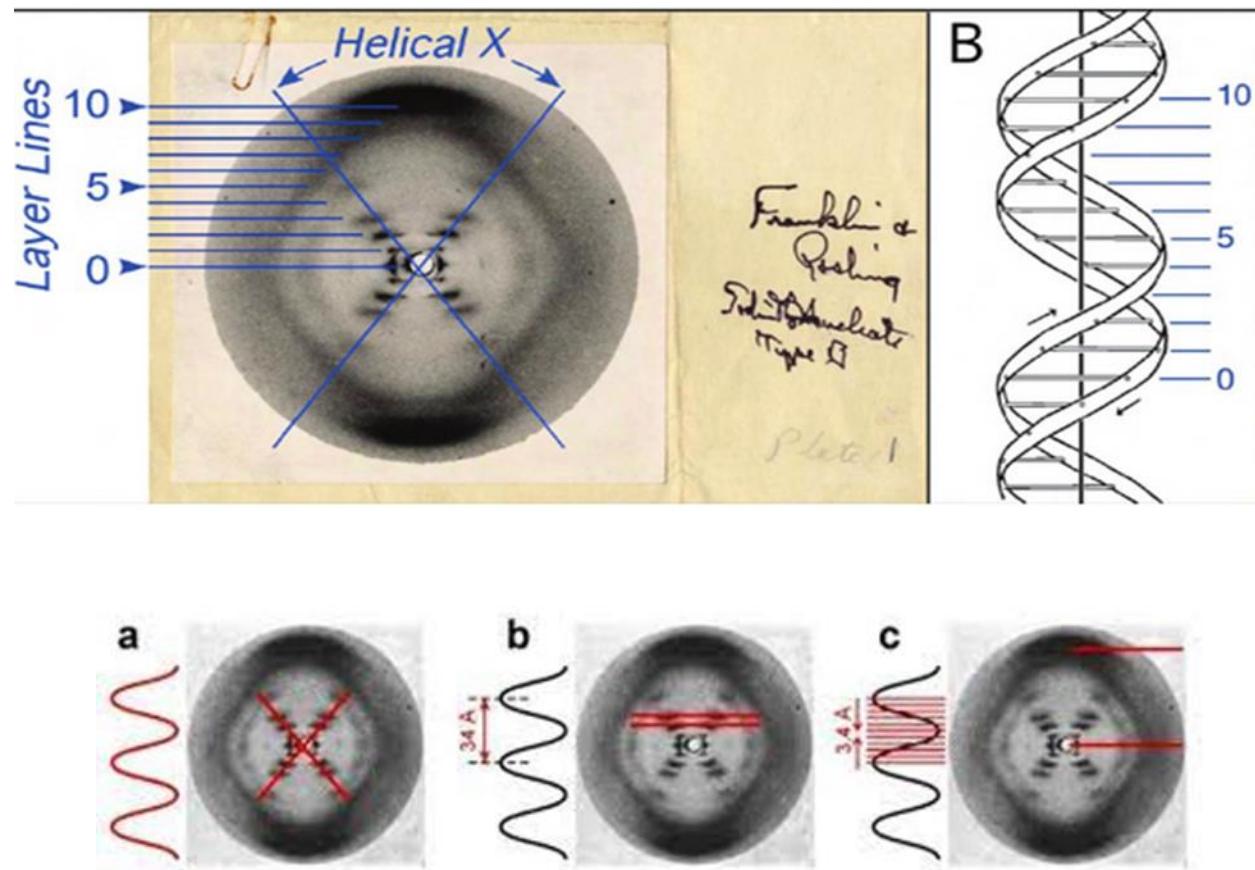
(b) Rosalind Franklin



(c) Franklin's X-ray diffraction pattern of wet DNA fibres

## PHOTOGRAPH 51

- The picture taken in 1952 showed a pattern that provides information on the position of the atoms present in DNA.



# ROSALIND FRANKLIN - THE TALE OF A HISTORICAL OMISSION

In 1962, Francis Crick, James Watson and Maurice Wilkins shared the Nobel Prize in Physiology or Medicine for their discovery of the double-helix structure of DNA. This was one of the major scientific achievements of the twentieth century, a revolution in our understanding of the building blocks of biology that indisputably merited recognition at the highest levels. But recognition for whom?

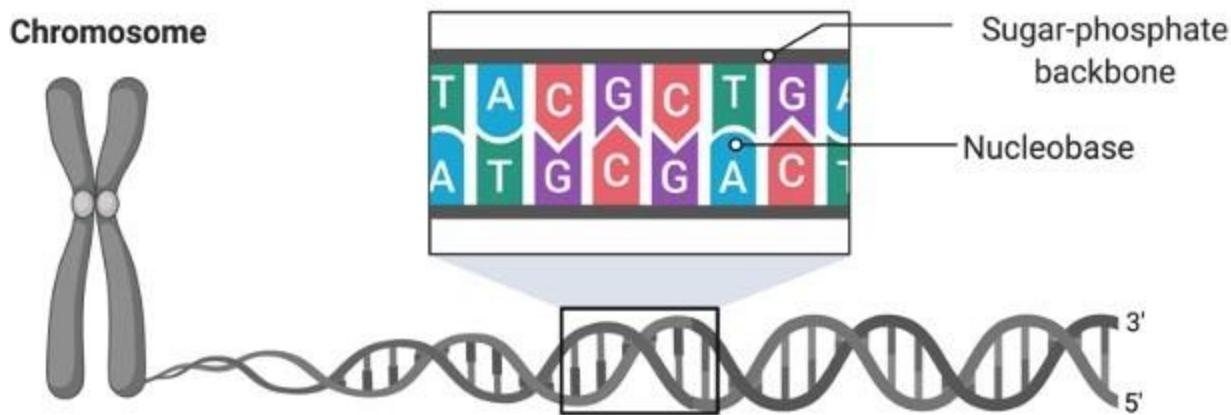
When Watson, Crick and Wilkins thanked their many collaborators from the podium, one of the names mentioned only in passing was Rosalind Franklin, a British chemist who had been Wilkins' colleague at King's College, London. **Her X-ray photographs had given Watson and Crick the final clues they needed to solve the puzzle of DNA's shape**, and the calculations from her crystallography work provided the measurements that allowed them to accurately model the atomic structures involved.

When Watson and Crick announced their discovery in 1953, the King's lab had been well on the way to solving the problem of DNA on its own. But the deep personal antipathy between Franklin and Wilkins had held back their lab's progress, and Franklin was never made fully aware of how her data had informed Watson and Crick's model. The **absence of any significant recognition for her contributions was predictable**.

Franklin and Wilkins' personal tensions, Watson and Crick's determination to win at all costs, and the lingering "boys club" mentality of post-war British labs are all ingredients in this story.

Rosalind Franklin did not share the 1962 Nobel Prize: she had died of ovarian cancer four years earlier at the age of 37 and the Nobel committee does not consider posthumous candidacies.





### Nucleobases

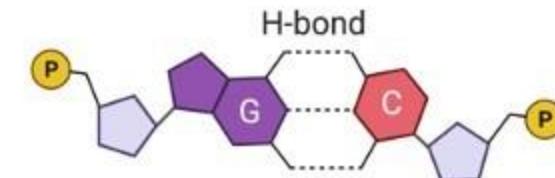
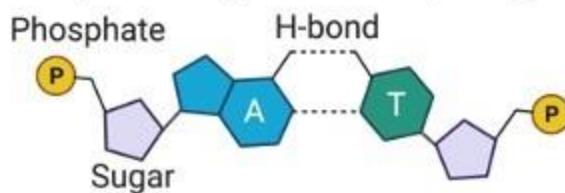
#### Purines:



#### Pyrimidines:

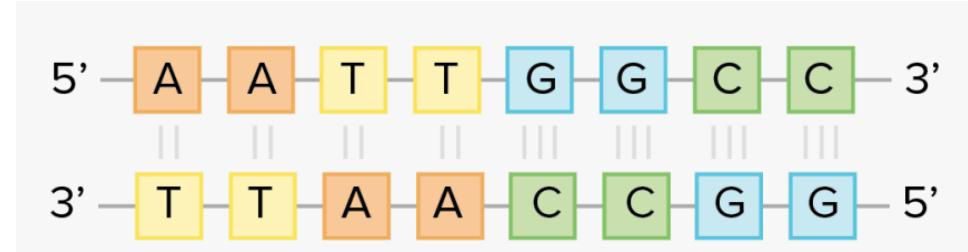


### Complementary nucleobase pairing



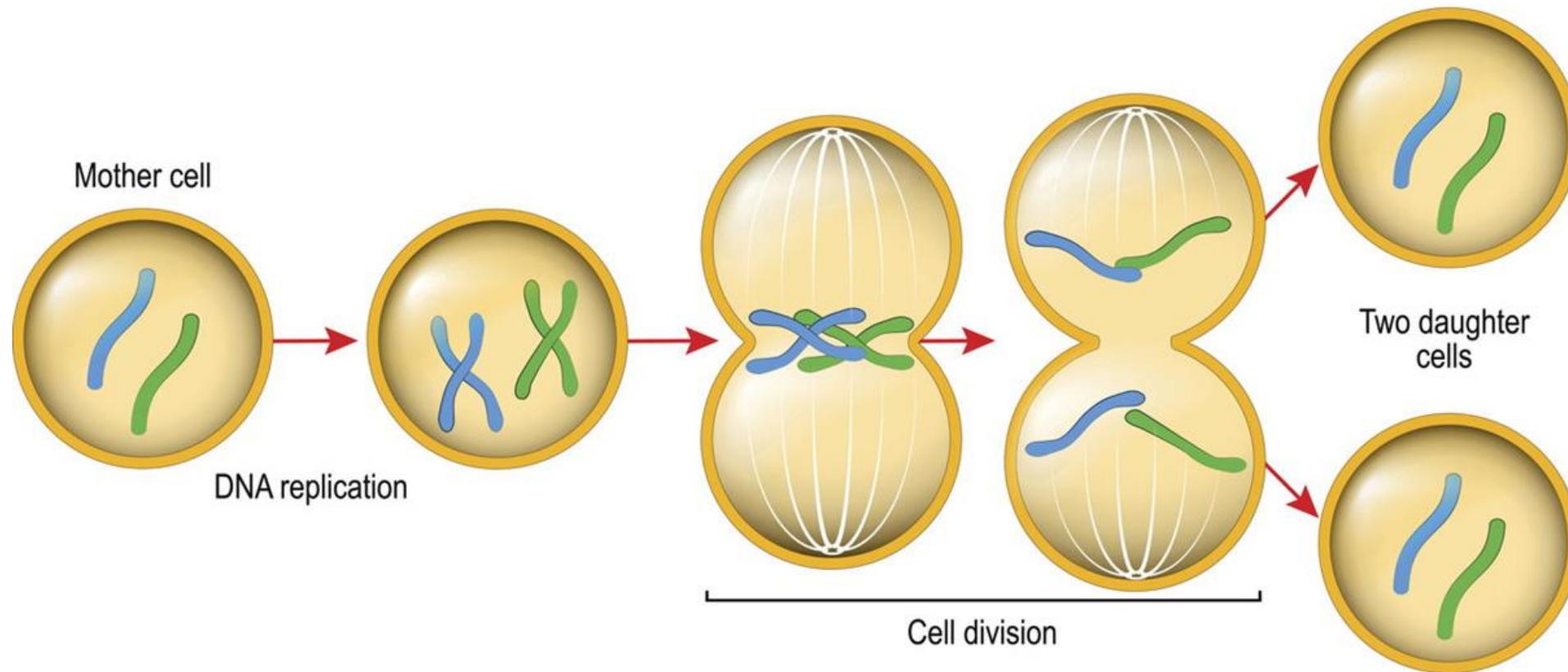
*Created with BioRender*

# COMPLEMENTARY BASE PAIRING

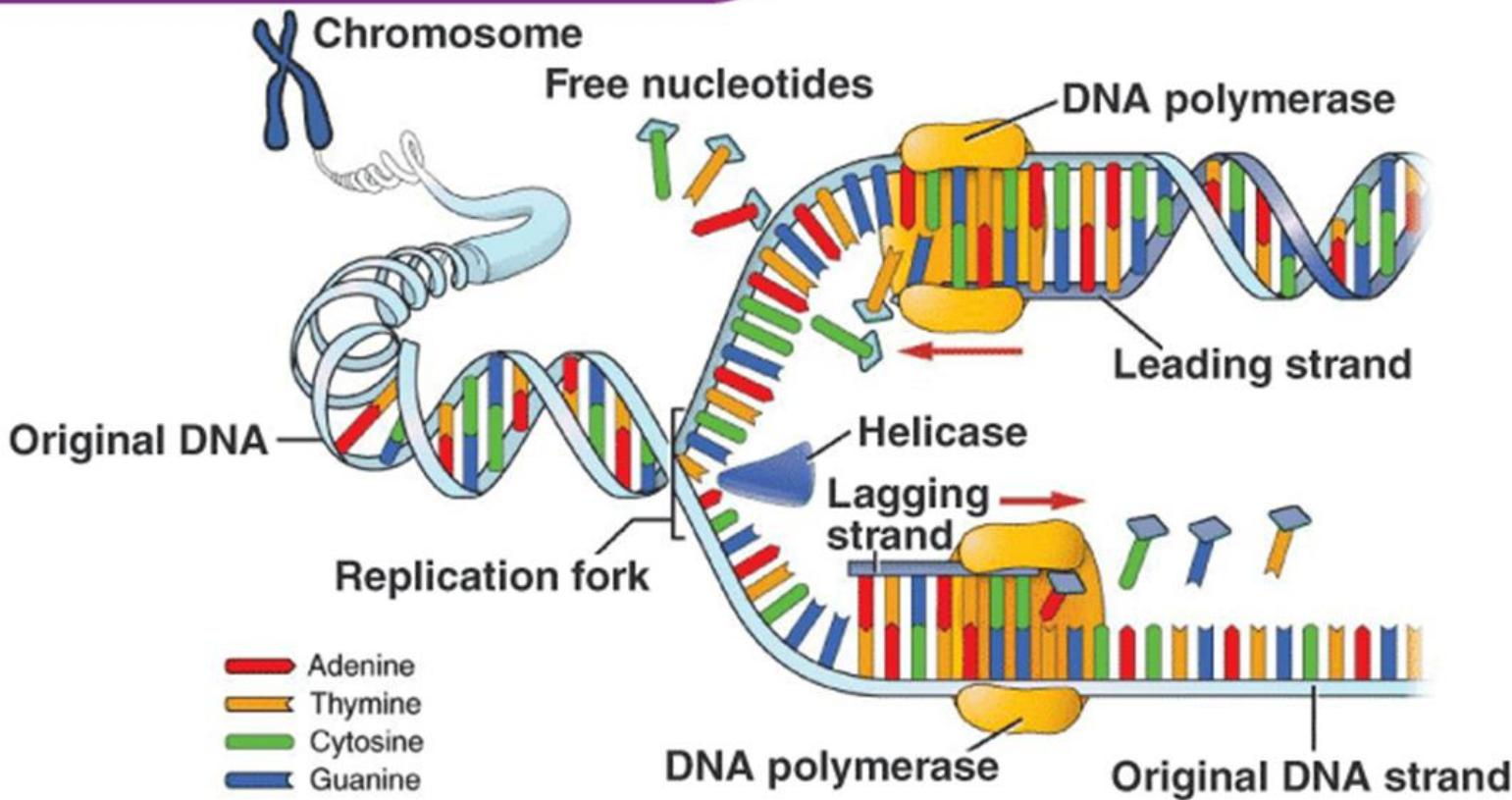


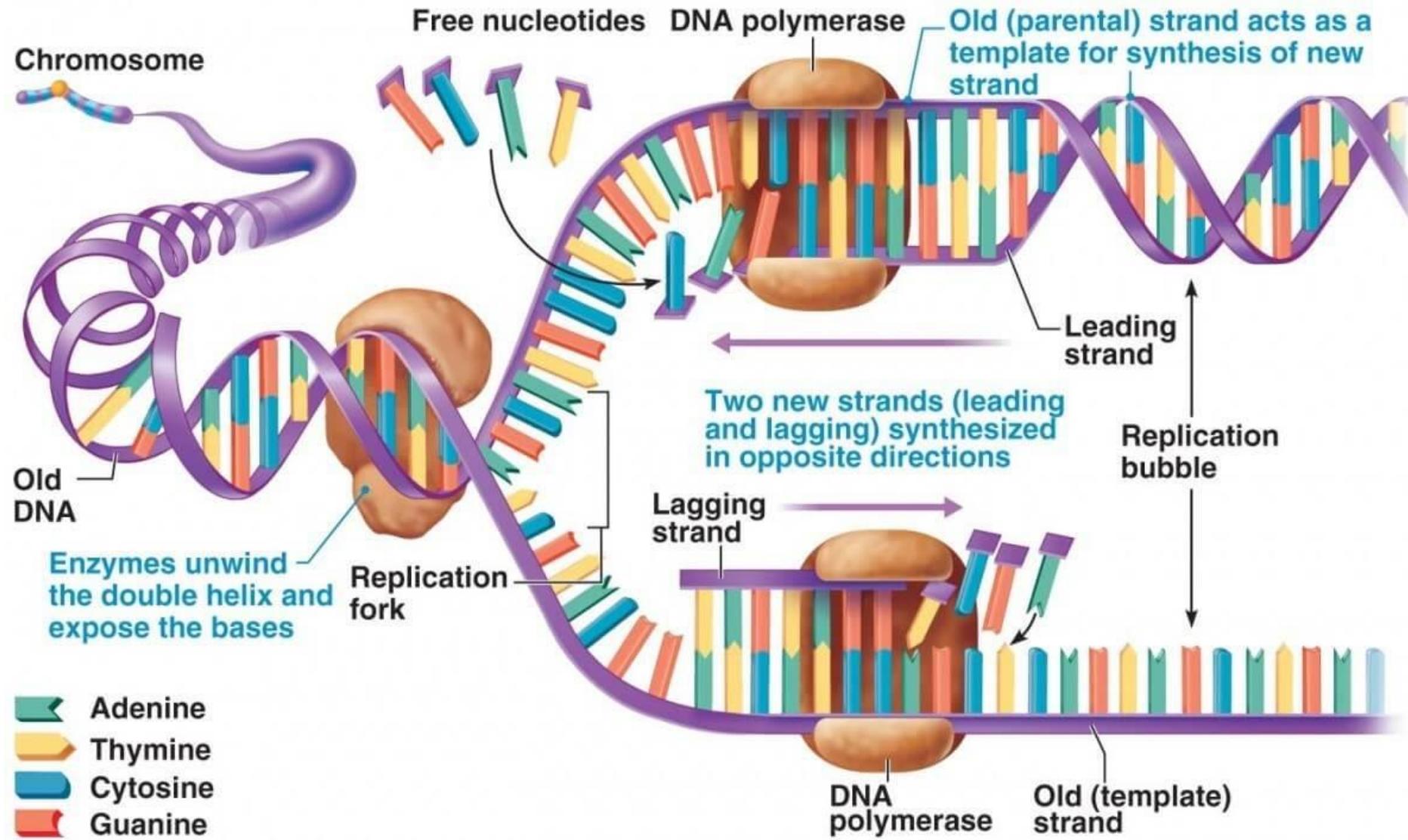
- is the phenomenon where in DNA guanine always hydrogen bonds to cytosine and adenine always binds to thymine. The bond between guanine and cytosine shares three hydrogen bonds compared to the A-T bond which always shares two hydrogen bonds.
- The human genome is made up of approximately 3 billion base pairs and is very complex, because of this we measure base pairs in Kbp (Kilobase pairs).
- Complementary base pairing is important in DNA as it allows the base pairs to be arranged in the most energetically favourable way; it is essential in forming the helical structure of DNA.
- It is also important in replication as it allows **semiconservative replication**.

# MITOSIS

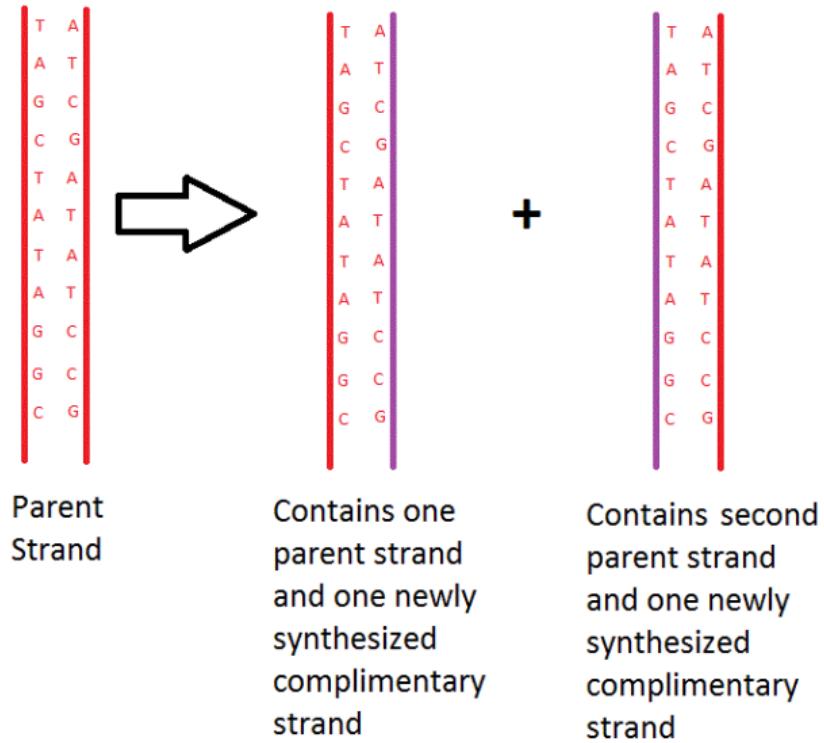
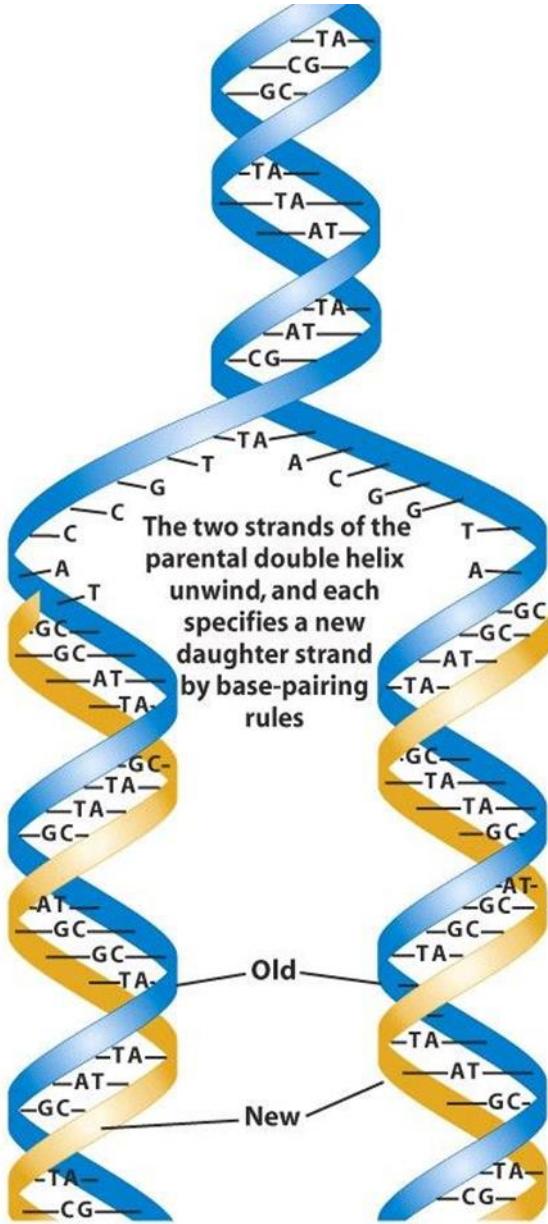


## DNA REPLICATION



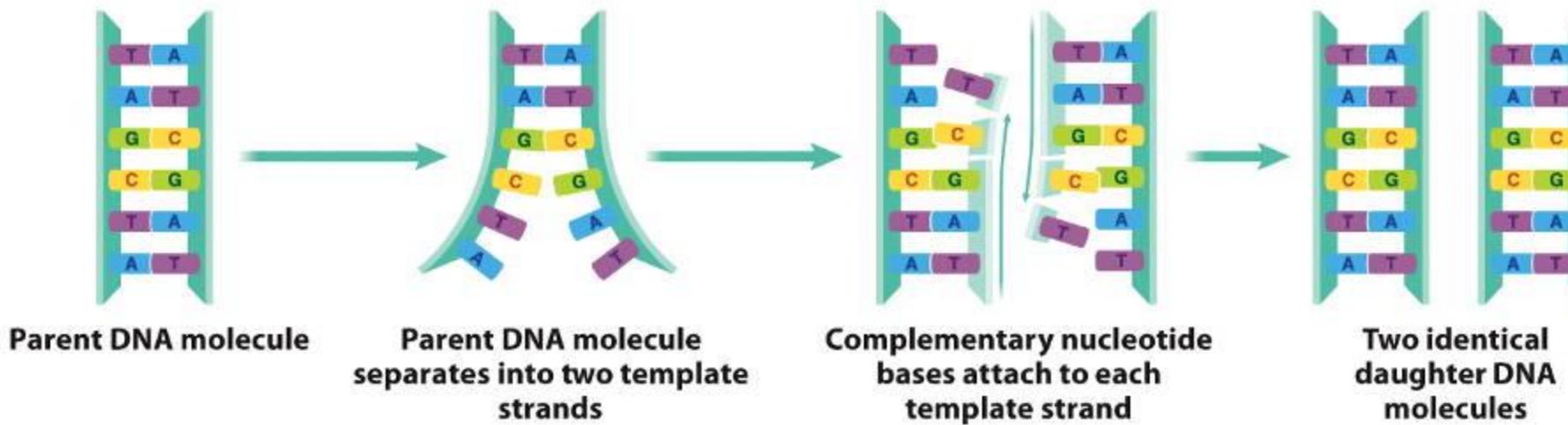


DNA replication  
results in two  
identical DNA  
strands.



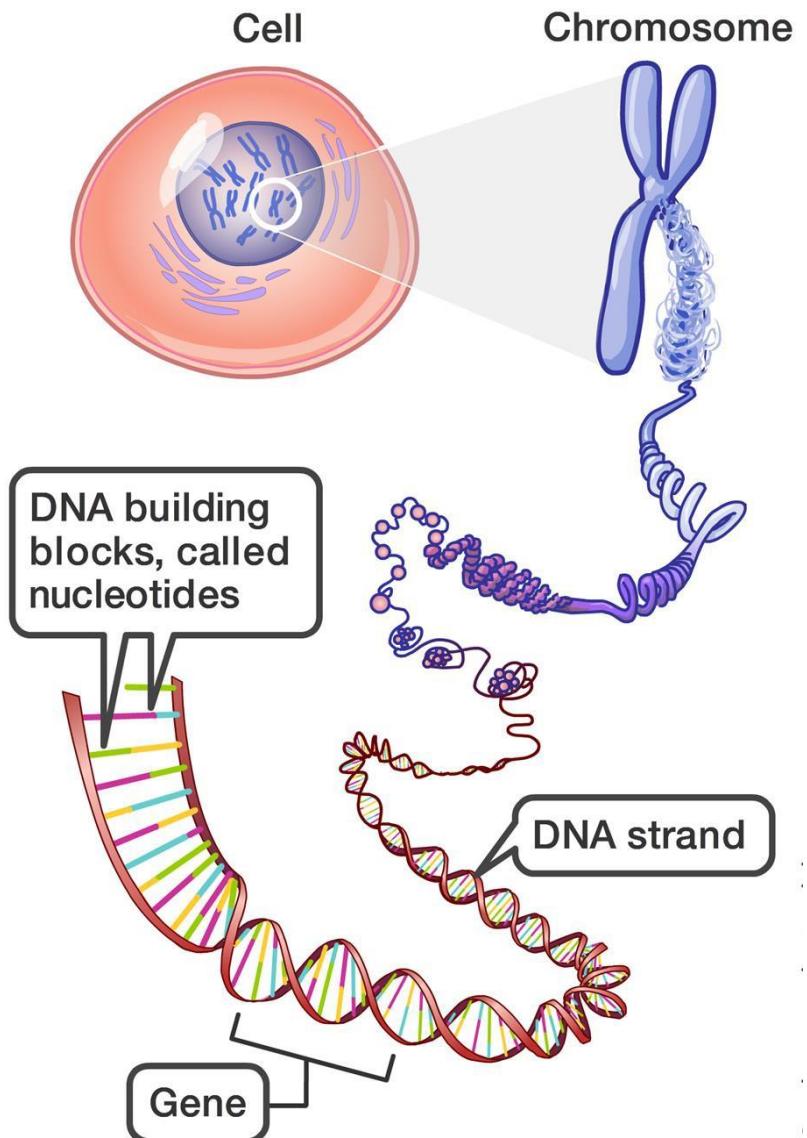
## DNA COMPLEMENTARITY

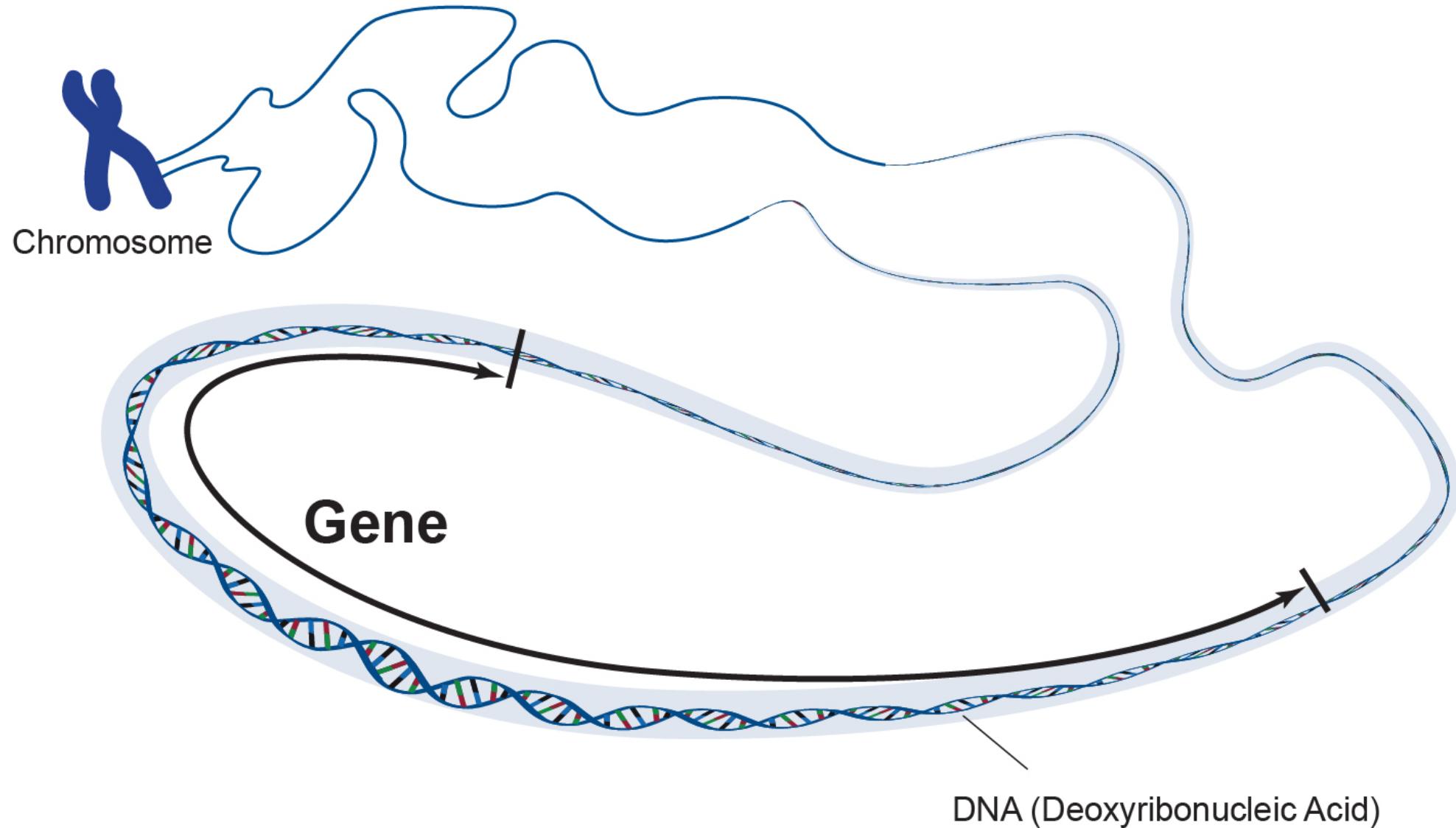
Complementary base pairing makes it possible to produce two identical strands by separating the parent molecule and using each strand as a template to build a new complementary strand.



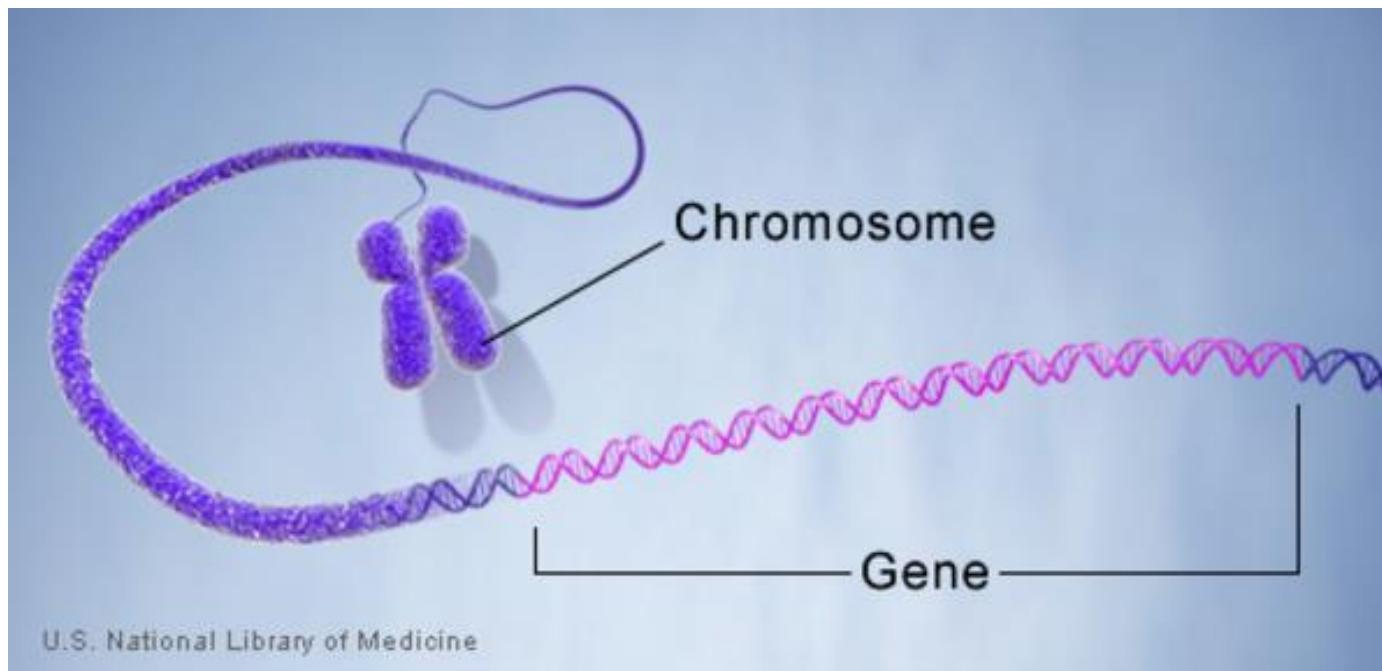
# GENES

- Functionally, DNA is organized into genes (pieces of DNA), which lead to observable traits.
- A gene is a specific sequence of bases that codes for a protein (made from amino acids) that determines a characteristic
- Those traits do not come from the DNA itself, but actually from RNA that is made from the DNA. The RNA sequence of bases determines the sequence of amino acids in proteins.
- The central dogma of molecular biology is that genes, which are made of DNA, are made into messenger RNAs, which are then made into proteins (e. g. observable traits like eye color).



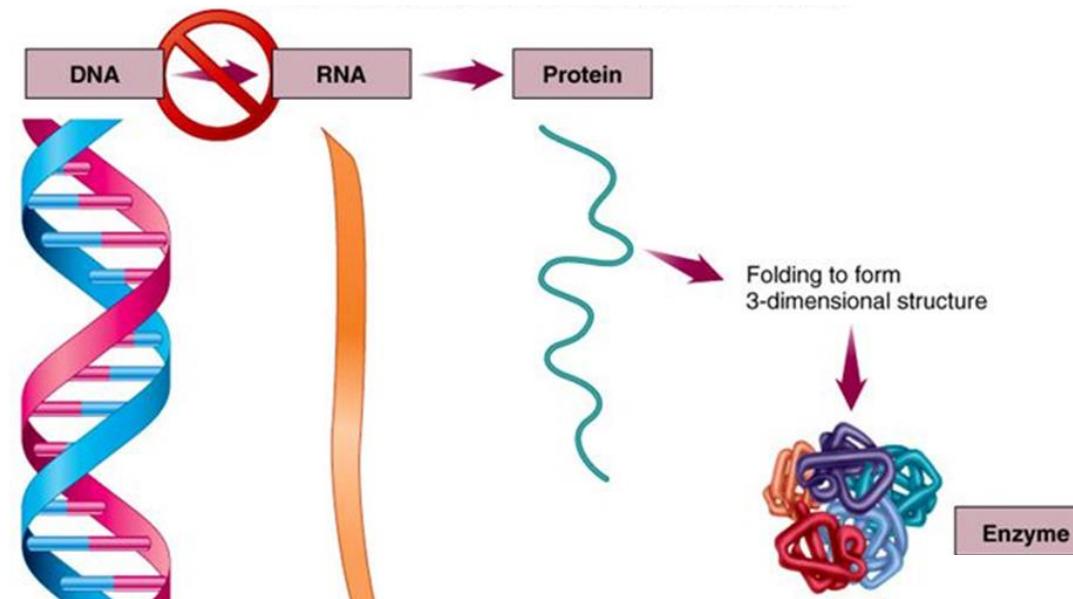


Genes are made up of DNA. Each chromosome contains many genes.

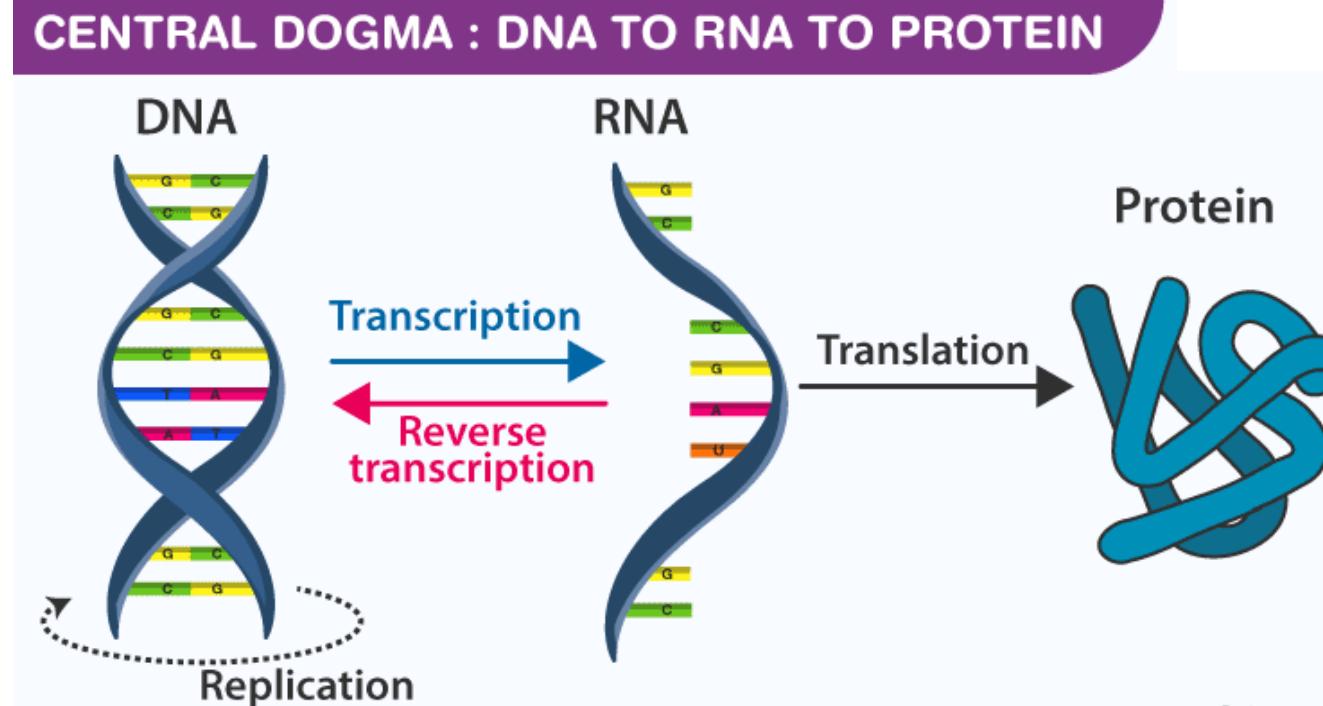


A gene is the basic physical and functional unit of heredity. Some genes act as instructions to make proteins. However, many genes do not code for proteins. In humans, genes vary in size from a few hundred DNA bases to more than 2 million bases.

# WHAT IS THE PURPOSE OF DNA? - A 'RECIPE'



# DOGMA OF MOLECULAR BIOLOGY



The flow of information from DNA to RNA to proteins is one of the fundamental principles of molecular biology.

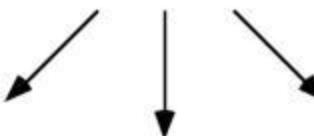
**DNA**

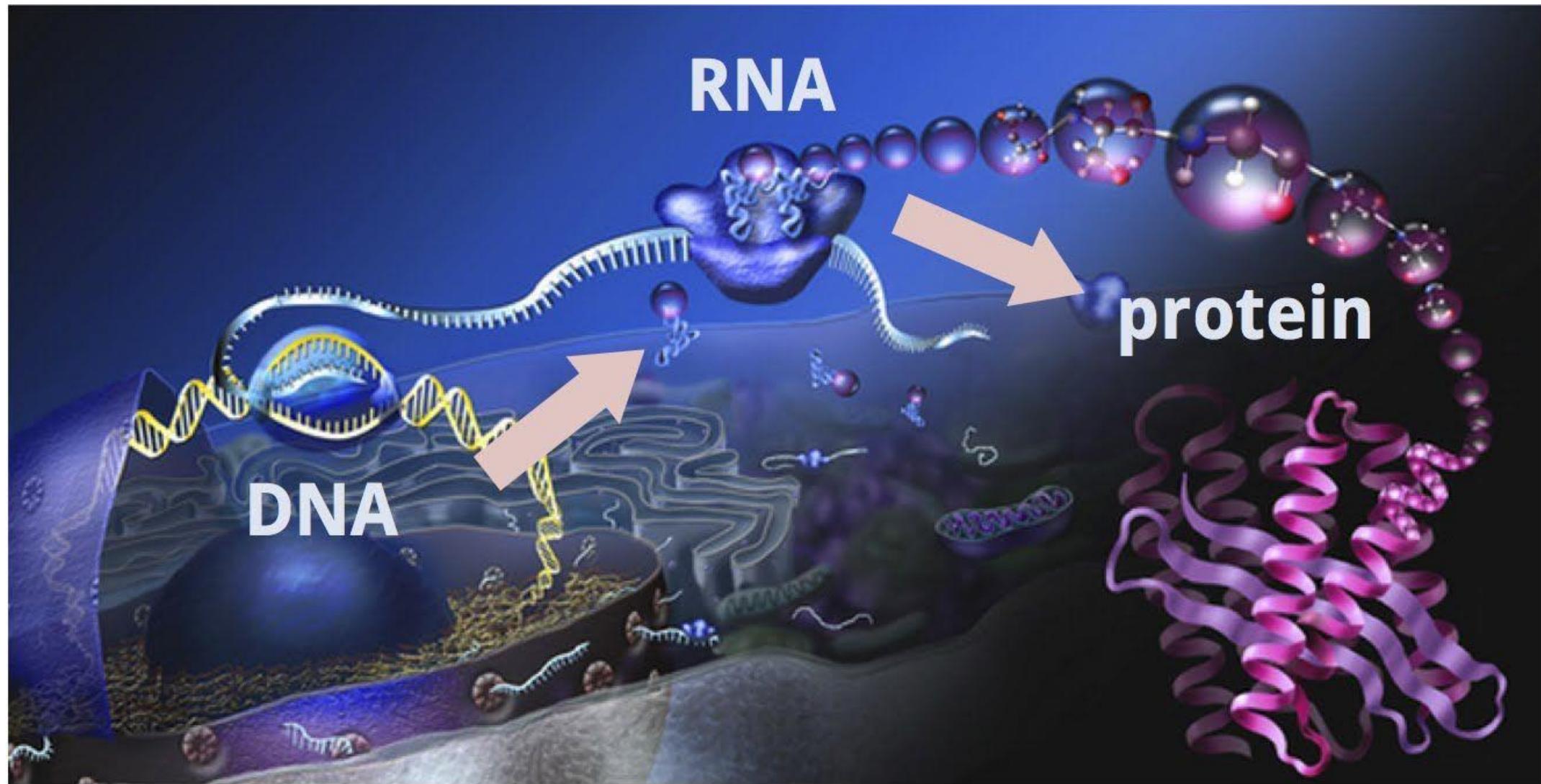


**RNA**



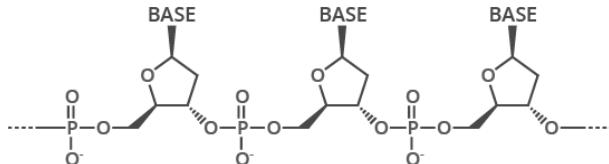
**Proteins**



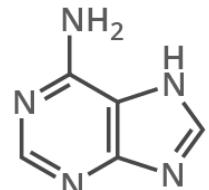


# THE CHEMICAL STRUCTURE OF DNA

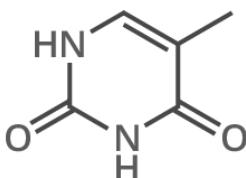
## THE SUGAR PHOSPHATE 'BACKBONE'



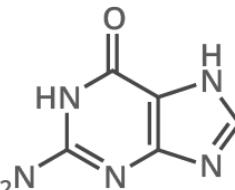
A ADENINE



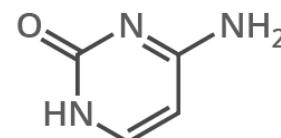
T THYMINE



G GUANINE

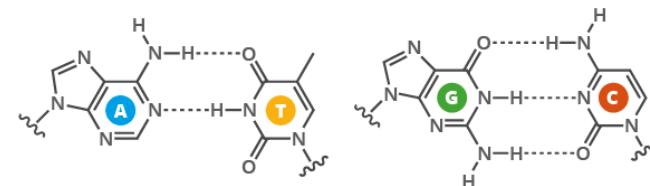


C CYTOSINE



## WHAT HOLDS DNA STRANDS TOGETHER?

DNA strands are held together by hydrogen bonds between bases on adjacent strands. Adenine (A) always pairs with thymine (T), while guanine (G) always pairs with cytosine (C). Adenine pairs with uracil (U) in RNA.



## FROM DNA TO PROTEINS

The bases on a single strand of DNA act as a code. The letters form three letter codons, which code for amino acids - the building blocks of proteins.

DNA → RNA → PROTEIN  
TRANSCRIPTION TRANSLATION

An enzyme, RNA polymerase, transcribes DNA into mRNA (messenger ribonucleic acid). It splits apart the two strands that form the double helix, then reads a strand and copies the sequence of nucleotides. The only difference between the RNA and the original DNA is that in the place of thymine (T), another base with a similar structure is used: uracil (U).

DNA SEQUENCE    T T C C T G A A C C C G T T A

mRNA SEQUENCE    U U C C U G A A C C C G U U A

AMINO ACID    Phenylalanine Leucine Asparagine Proline Leucine

In multicellular organisms, the mRNA carries genetic code out of the cell nucleus, to the cytoplasm. Here, protein synthesis takes place. 'Translation' is the process of turning the mRNA's 'code' into proteins. Molecules called ribosomes carry out this process, building up proteins from the amino acids coded for.

## FROM DNA TO PROTEINS

The bases on a single strand of DNA act as a code. The letters form three letter codons, which code for amino acids - the building blocks of proteins.

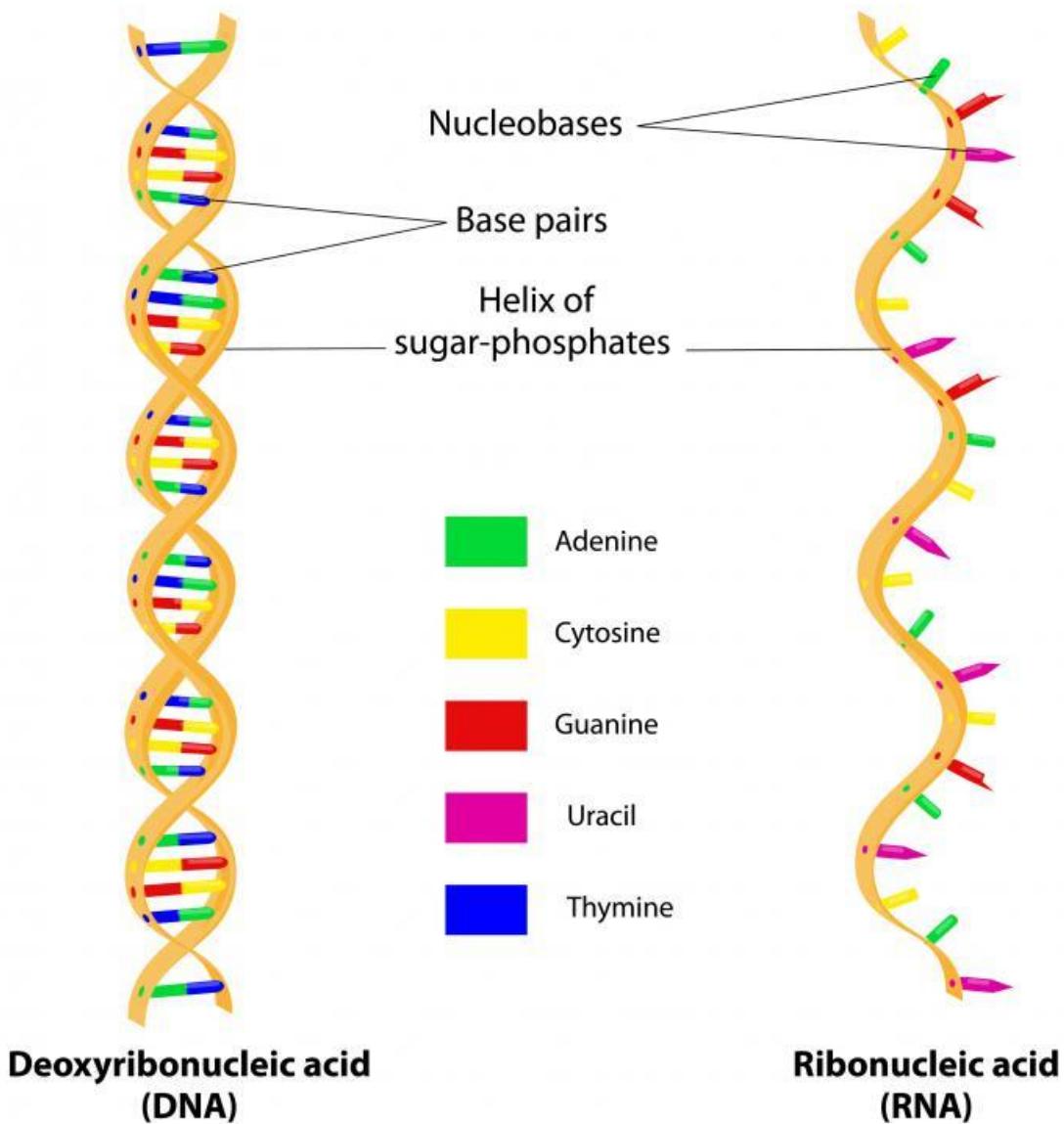


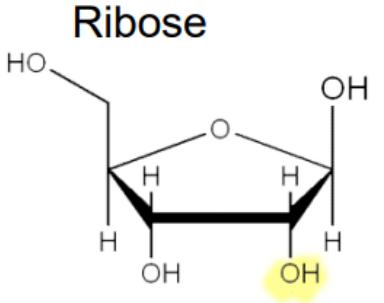
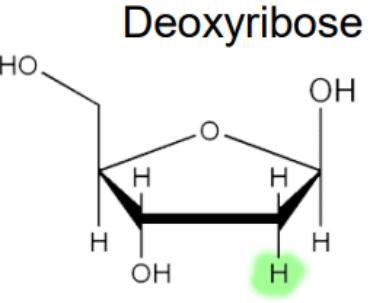
An enzyme, RNA polymerase, transcribes DNA into mRNA (messenger ribonucleic acid). It splits apart the two strands that form the double helix, then reads a strand and copies the sequence of nucleotides. The only difference between the RNA and the original DNA is that in the place of thymine (T), another base with a similar structure is used: uracil (U).



In multicellular organisms, the mRNA carries genetic code out of the cell nucleus, to the cytoplasm. Here, protein synthesis takes place. 'Translation' is the process of turning the mRNA's 'code' into proteins. Molecules called ribosomes carry out this process, building up proteins from the amino acids coded for.

# Structure of DNA & RNA



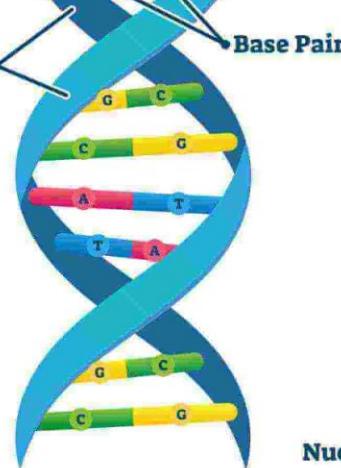
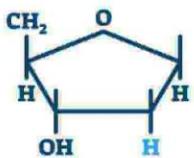
|                   | <b>RNA</b>  | <b>DNA</b>   |
|-------------------|---|--|
| Bases             | Adenine (A)<br>Guanine (G)<br><b>Uracil (U)</b><br>Cytosine (C)                               | Adenine (A)<br>Guanine (G)<br><b>Thymine (T)</b><br>Cytosine (C)                                   |
| Sugar             | Ribose<br> | Deoxyribose<br> |
| Number of strands | <b>Single stranded</b> , and often, but not always, linear in shape                           | <b>Two anti-parallel, complementary strands</b> form a double helix                                |

# DNA

Deoxyribonucleic Acid



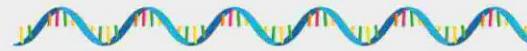
Double-Stranded  
Sugar Phosphate  
DEOXYRIBOSE



VS

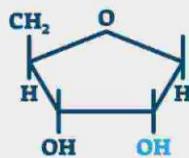
# RNA

Ribonucleic Acid

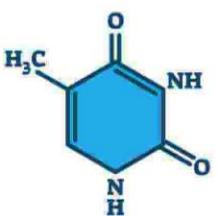


Single  
Nucleobase

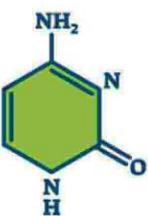
Single-Stranded  
Sugar Phosphate  
RIBOSE



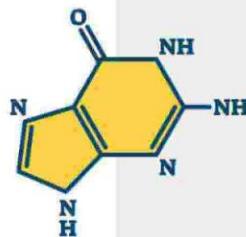
Nucleobases



Thymine



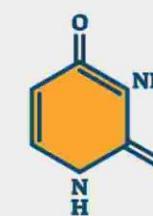
Cytosine



Guanine



Adenine



Uracil

T

C

G

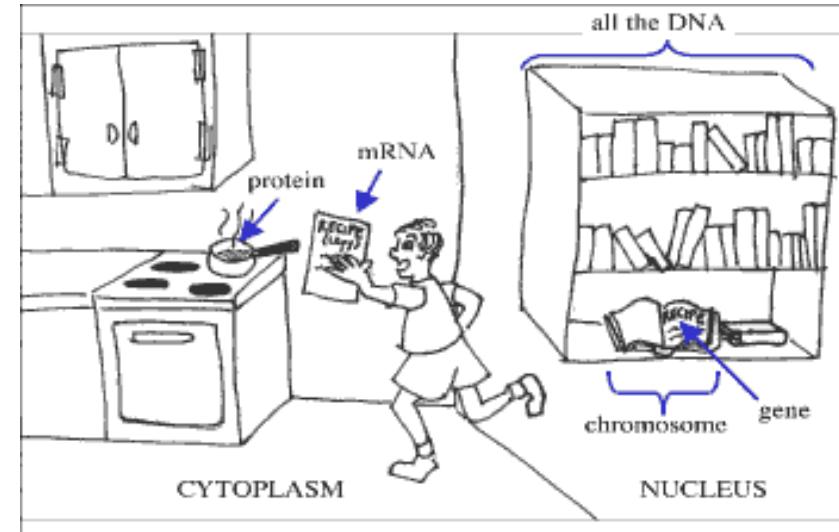
A

U



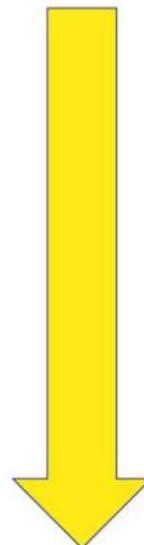
# WHAT DOES DNA ACTUALLY DO?

- DNA is sometimes referred to as the "code of life".
- This is because the sequence of bases in a DNA molecule can determine an organism's characteristics.
- Cells have specialized equipment that can read a DNA sequence and use it to piece together amino acids in order to make proteins.
- Essentially, DNA is the recipe book for our proteins.

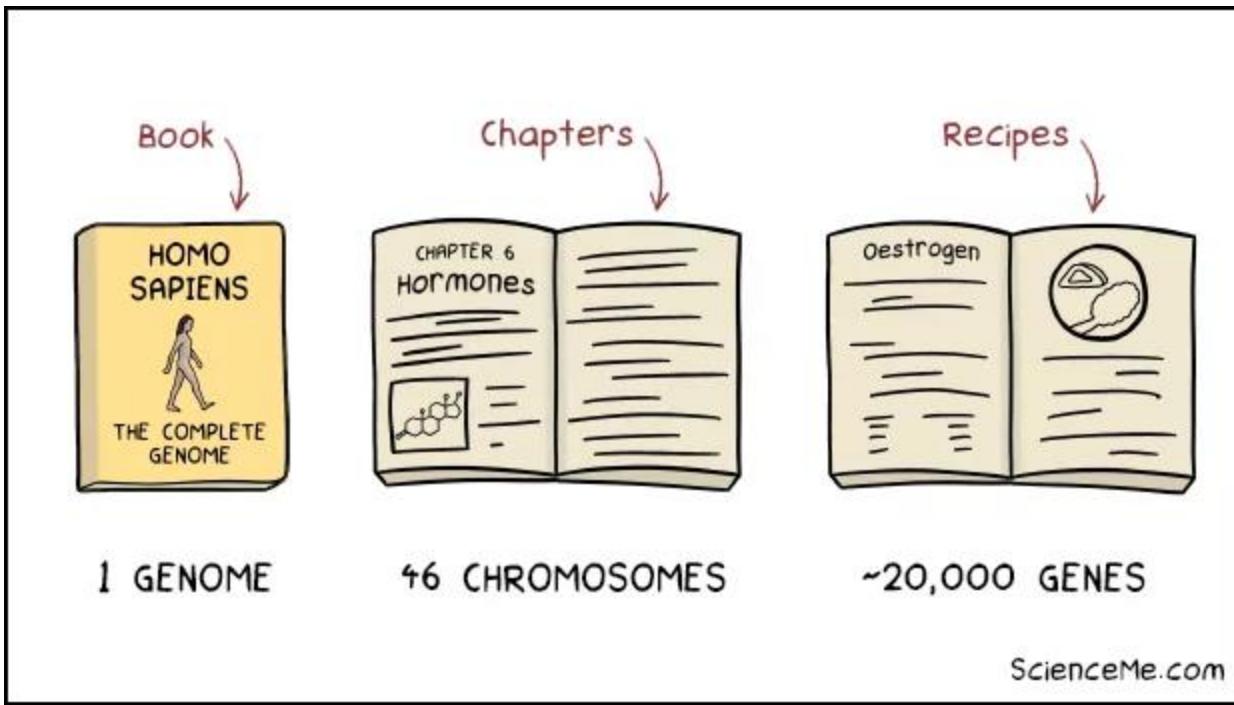


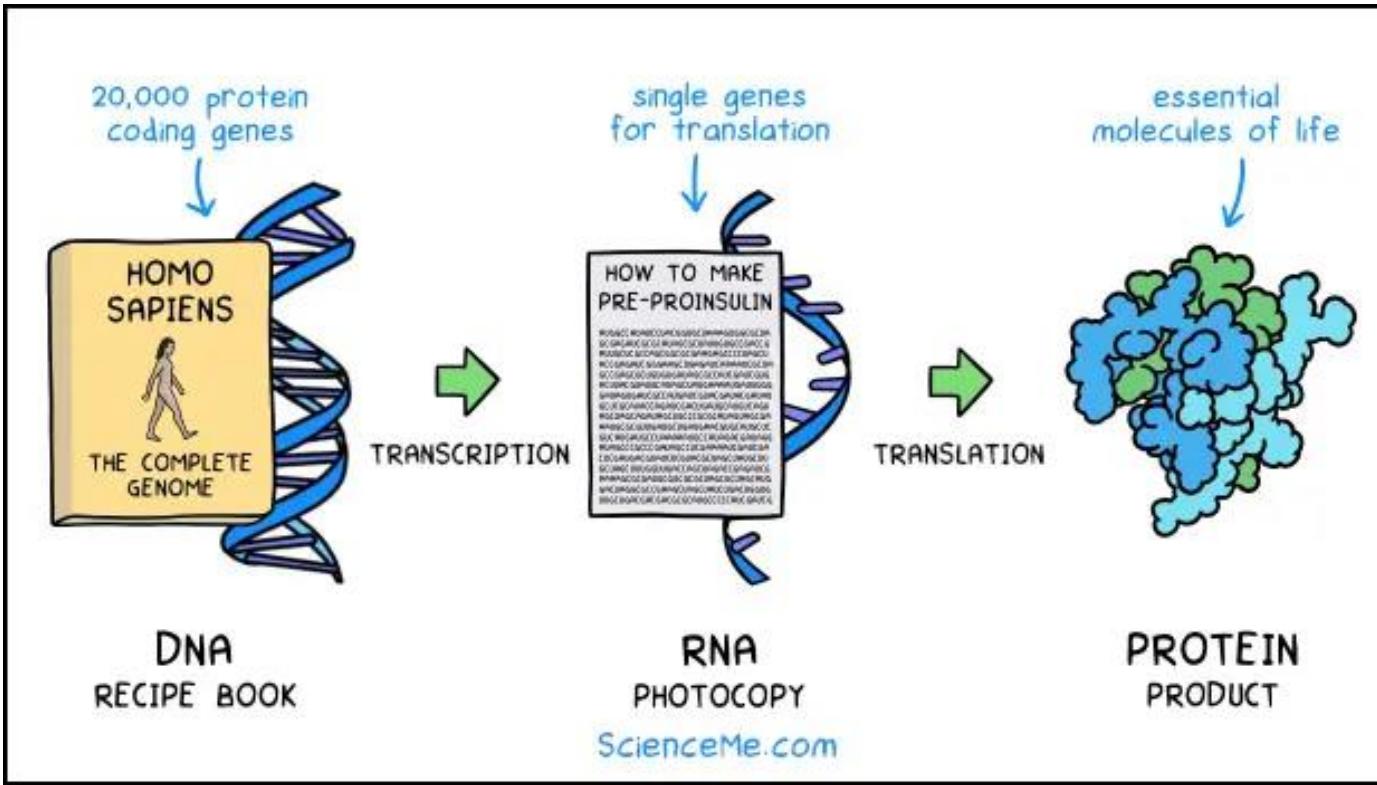
- **Imagine amino acids like letters of the alphabet...**

Amino Acids  
Polypeptides  
Protein  
Tissues  
Organs  
Organism



Letters  
Words  
Sentences  
Paragraphs  
Chapters  
Books





Just like a bunch of letters can make words,  
words put together make sentences, sentences  
put together make paragraphs, paragraphs put  
together, make chapters, and chapters put  
together make books...



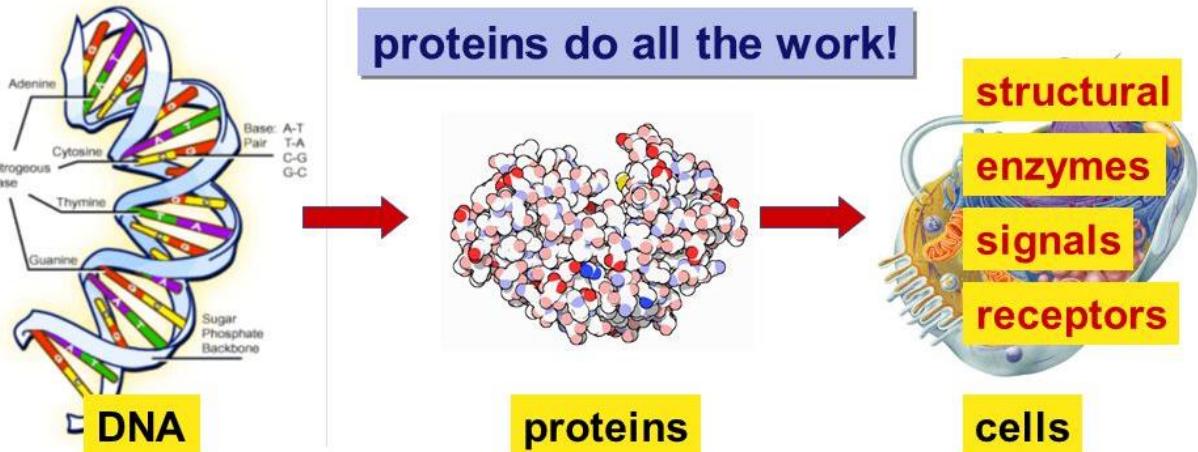
Our bodies form in a similar way...a bunch of amino acids can make a polypeptide, a bunch of polypeptides make a protein, a bunch of proteins can make tissues, a bunch of tissues can make organs, and a bunch of organs can make you!

# DNA FORMS PROTEINS

- Proteins are the moving working parts of our cells and are responsible for virtually all cellular processes. Complex cells contain various compartments (called organelles) that perform various functions from generating cellular energy to processing and packing materials.
- It is actually the different proteins within each compartment that are responsible for these cellular activities. The complex interaction of thousands of different proteins within each cell is what makes them living.

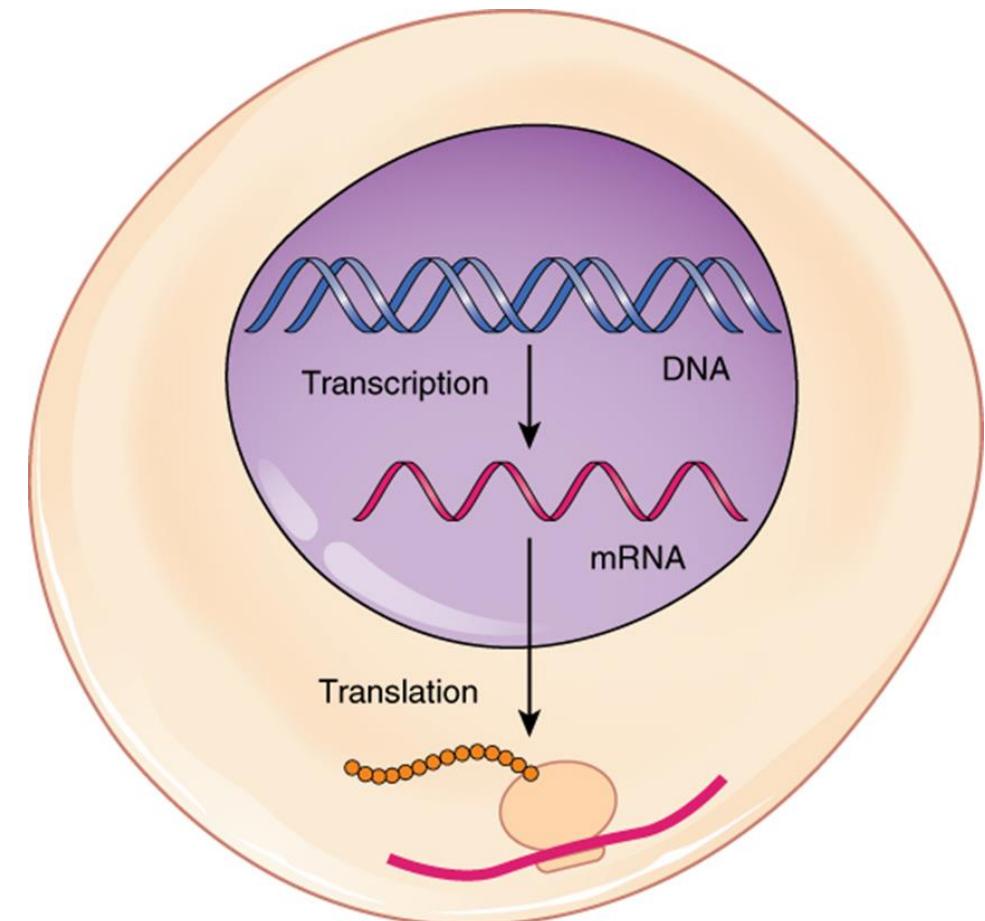
one of the major job of cells is to make proteins, because...

proteins do all the work!

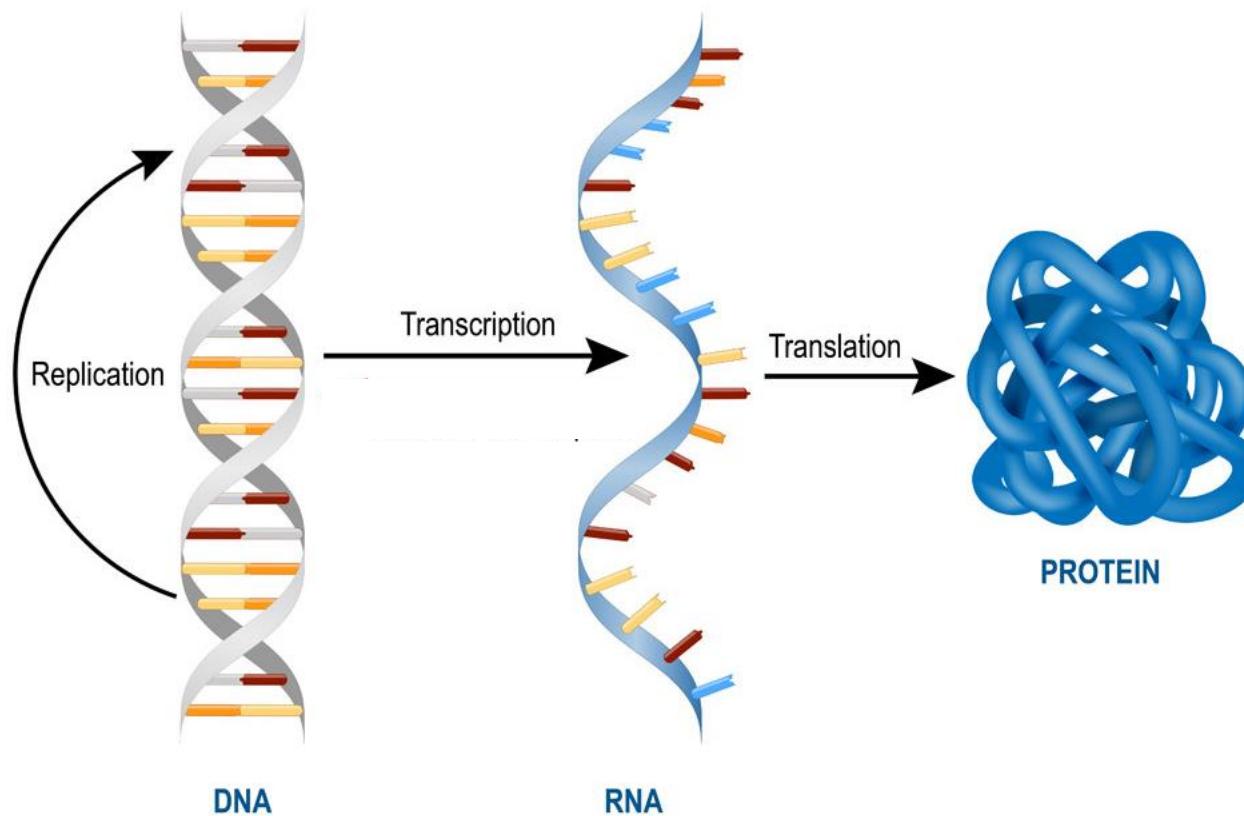


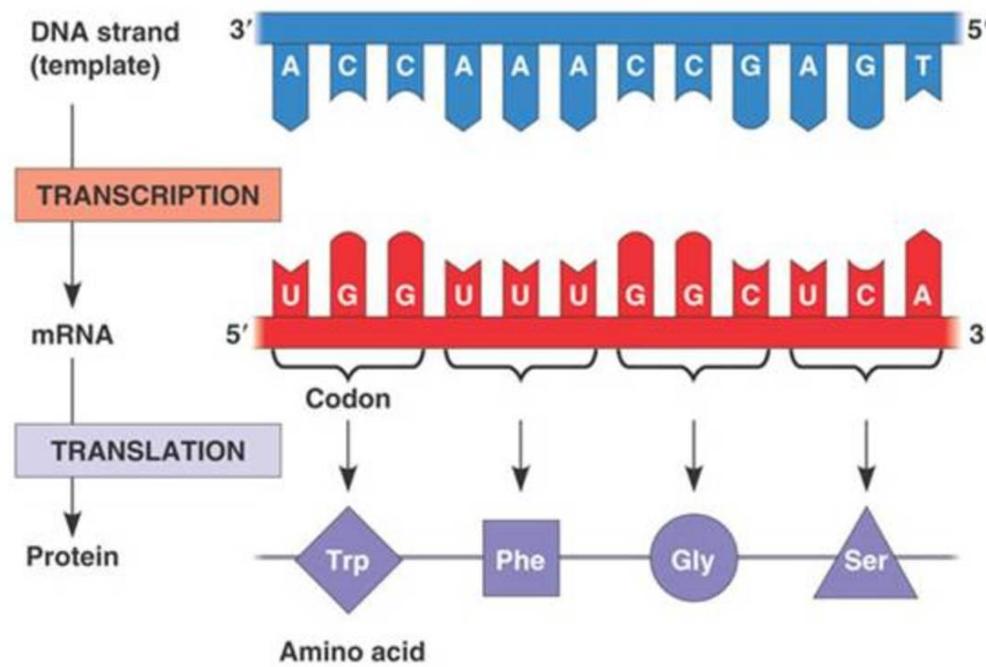
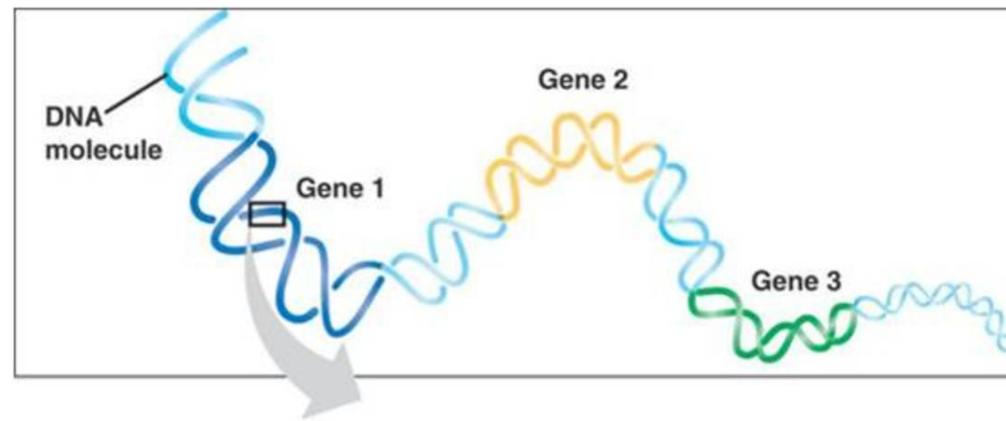
# HOW DO GENES DIRECT THE PRODUCTION OF PROTEINS?

- The journey from gene to protein is complex and tightly controlled within each cell. It consists of two major steps: transcription and translation. Together, transcription and translation are known as **gene expression**.
- During the process of **transcription**, the information stored in a gene's DNA is passed to RNA **in the cell nucleus**.
- **Translation**, the second step in getting from a gene to a protein, takes place **in the cytoplasm**. The mRNA interacts with a specialized complex called a ribosome, which "reads" the sequence of mRNA nucleotides.



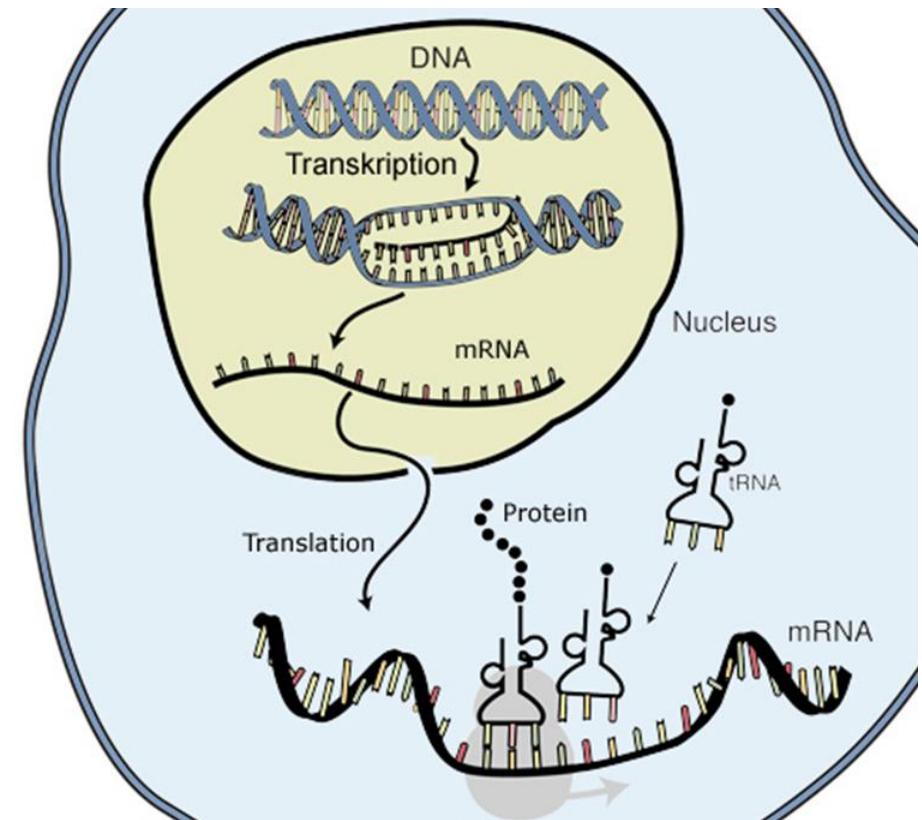
# Transcription and Translation



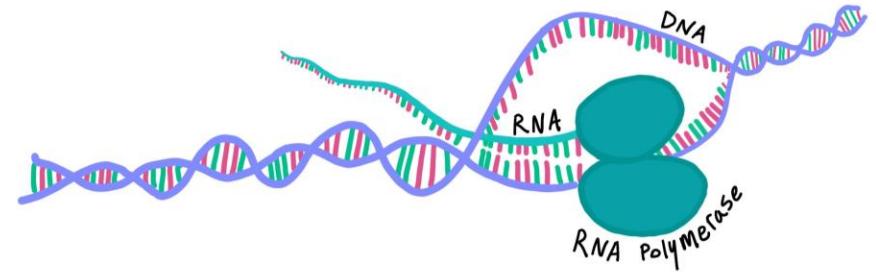


# HOW DO GENES DIRECT THE PRODUCTION OF PROTEINS?

- The type of RNA that contains the information for making a protein is called messenger RNA (mRNA) because it carries the information, or message, from the DNA out of the nucleus into the cytoplasm.
- Each sequence of three nucleotides, called a codon, usually codes for one particular amino acid. A type of RNA called transfer RNA (tRNA) assembles the protein, one amino acid at a time. Protein assembly continues until the ribosome encounters a “stop” codon (a sequence of three nucleotides that does not code for an amino acid).

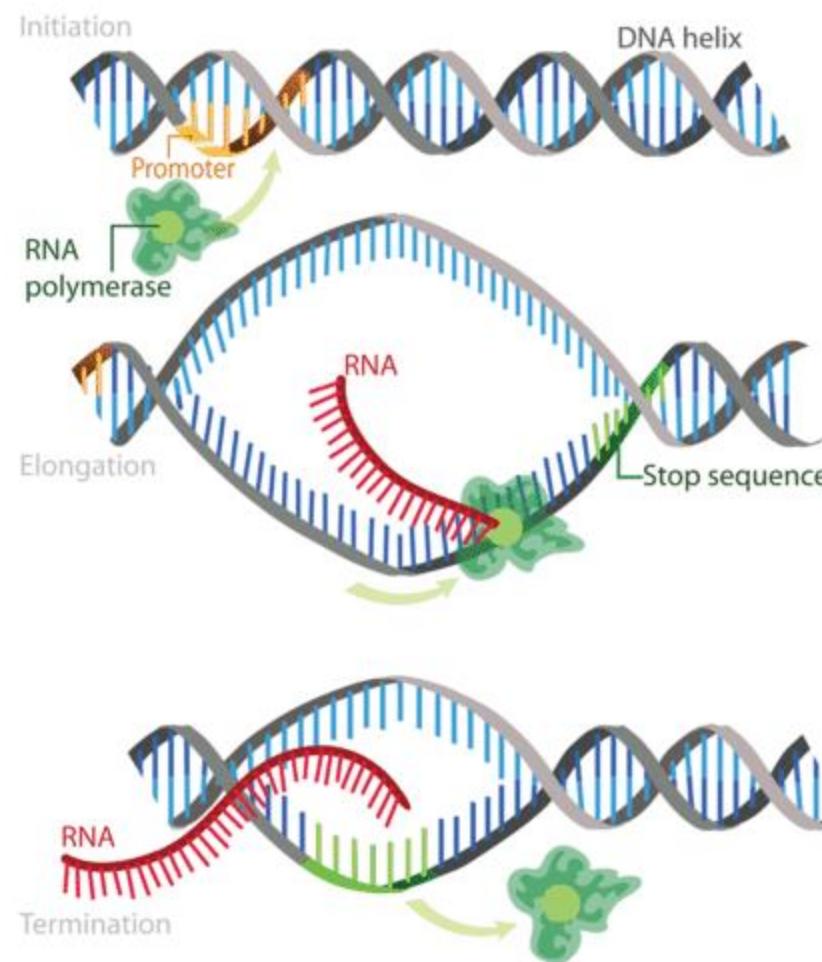


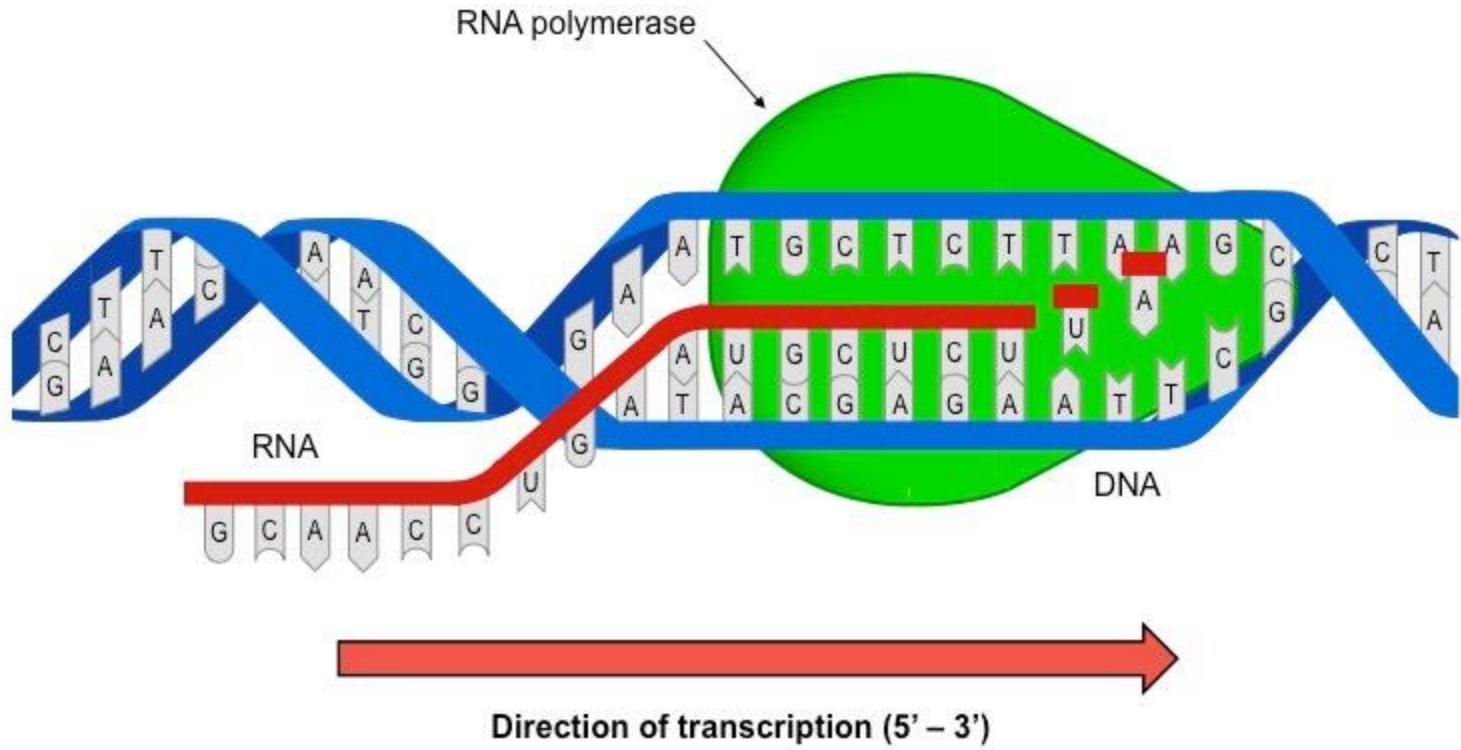
# TRANSCRIPTION



Protein production faces a number of challenges. Chief amongst these is that proteins are produced in the cytoplasm of the cell, and DNA never leaves the nucleus. To get around this problem, DNA creates a messenger molecule to deliver its information outside of the nucleus: mRNA (messenger RNA). The process of making this messenger molecule is known as transcription, and has a number of steps:

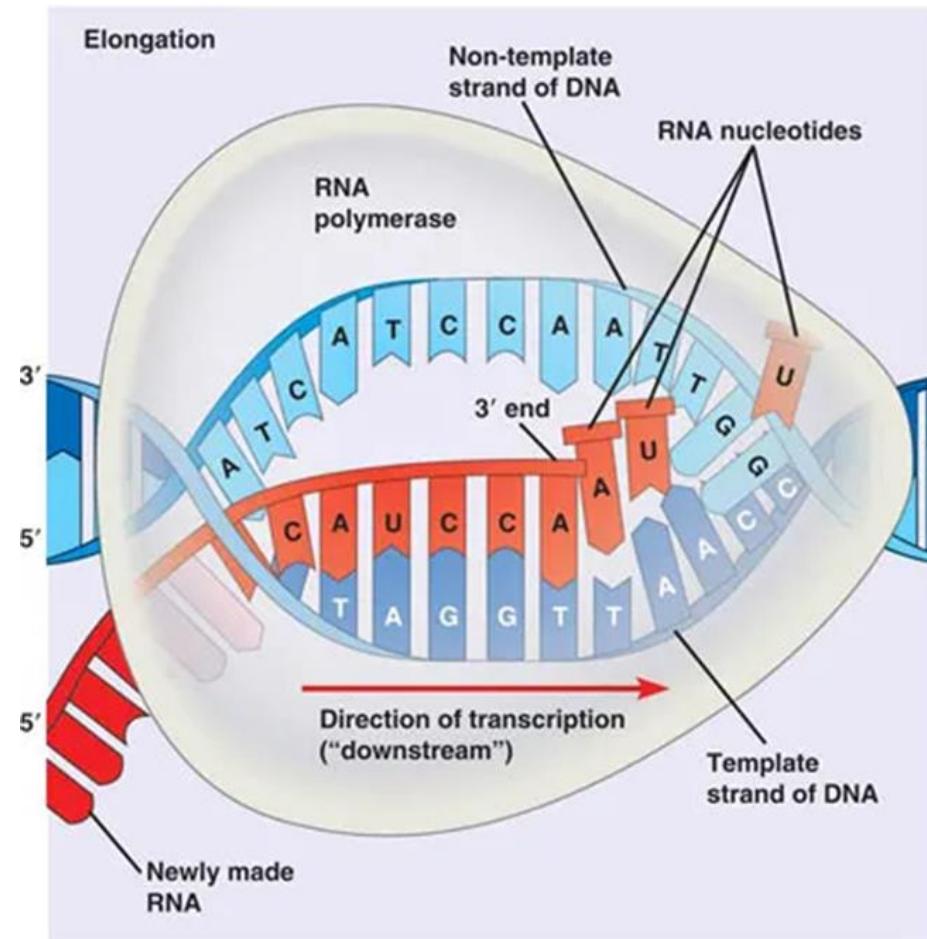
- Initiation: The double helix of the DNA is unwound by RNA Polymerase, which docks on the DNA at a special sequence of bases (promoter)
- Elongation: RNA Polymerase moves downstream unwinding the DNA. As the double helix unwinds, ribonucleotide bases (A, C, G and U) attach themselves to the DNA template strand (the strand being copied) by complementary base pairing.
- RNA Polymerase catalyses the formation of covalent bonds between the nucleotides.
- Termination: The RNA transcript is released from the DNA, along with the RNA polymerase.

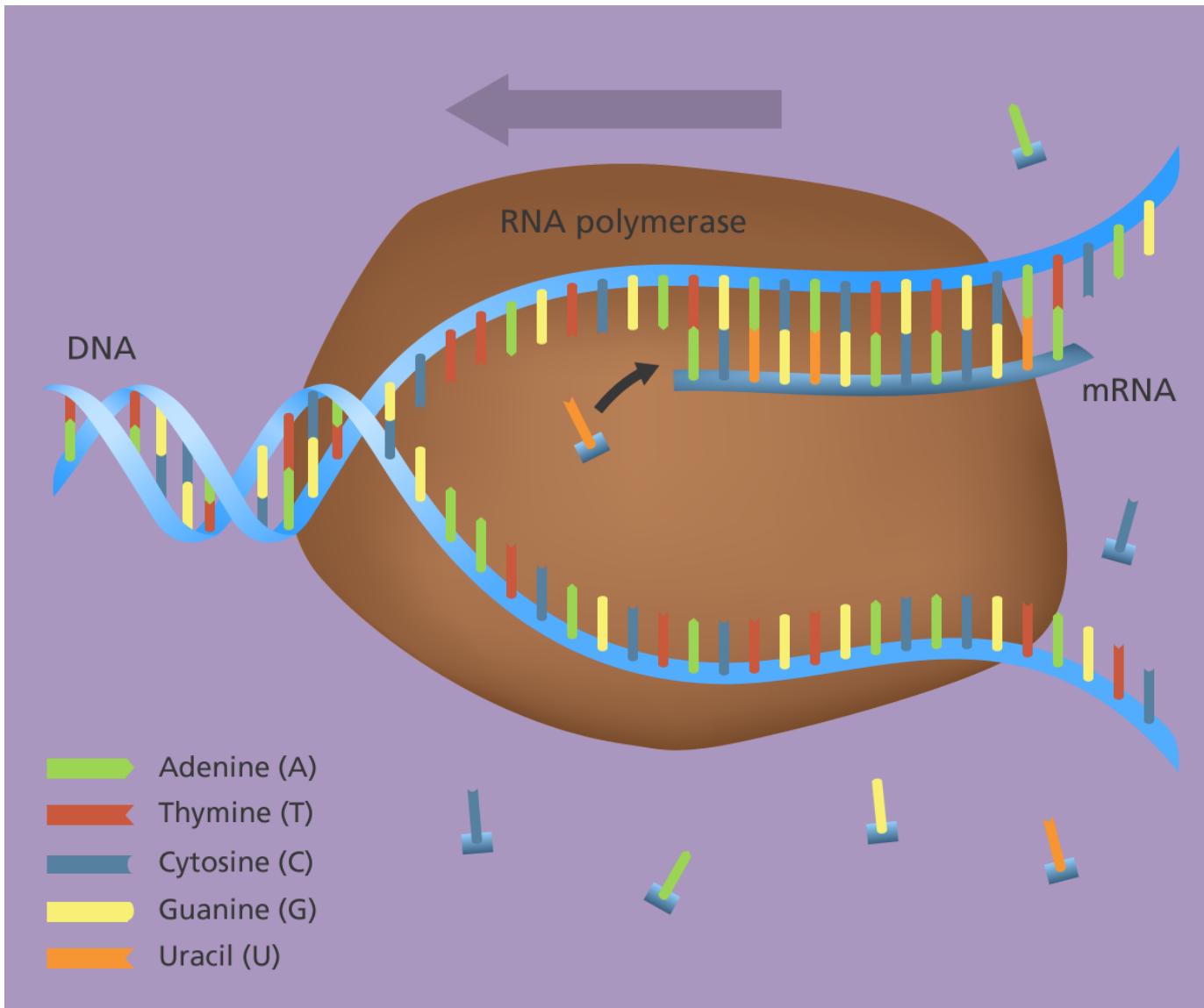




# TRANSCRIPTION

- Elongation of an RNA strand. Transcription is well underway, complementary base pairing rules dictate the sequence of bases in the growing RNA strand.



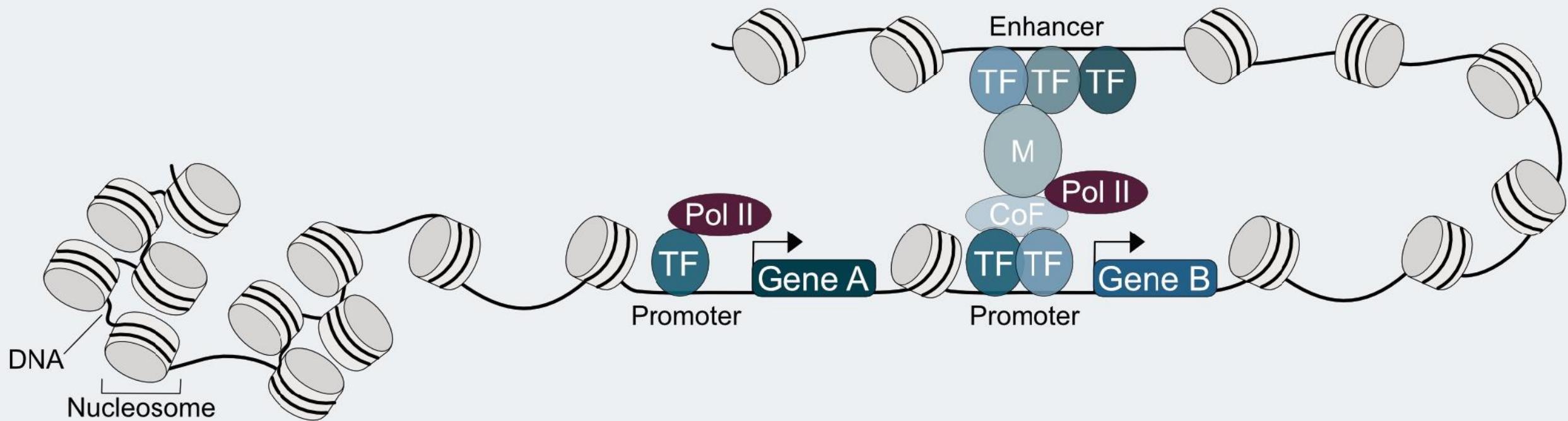


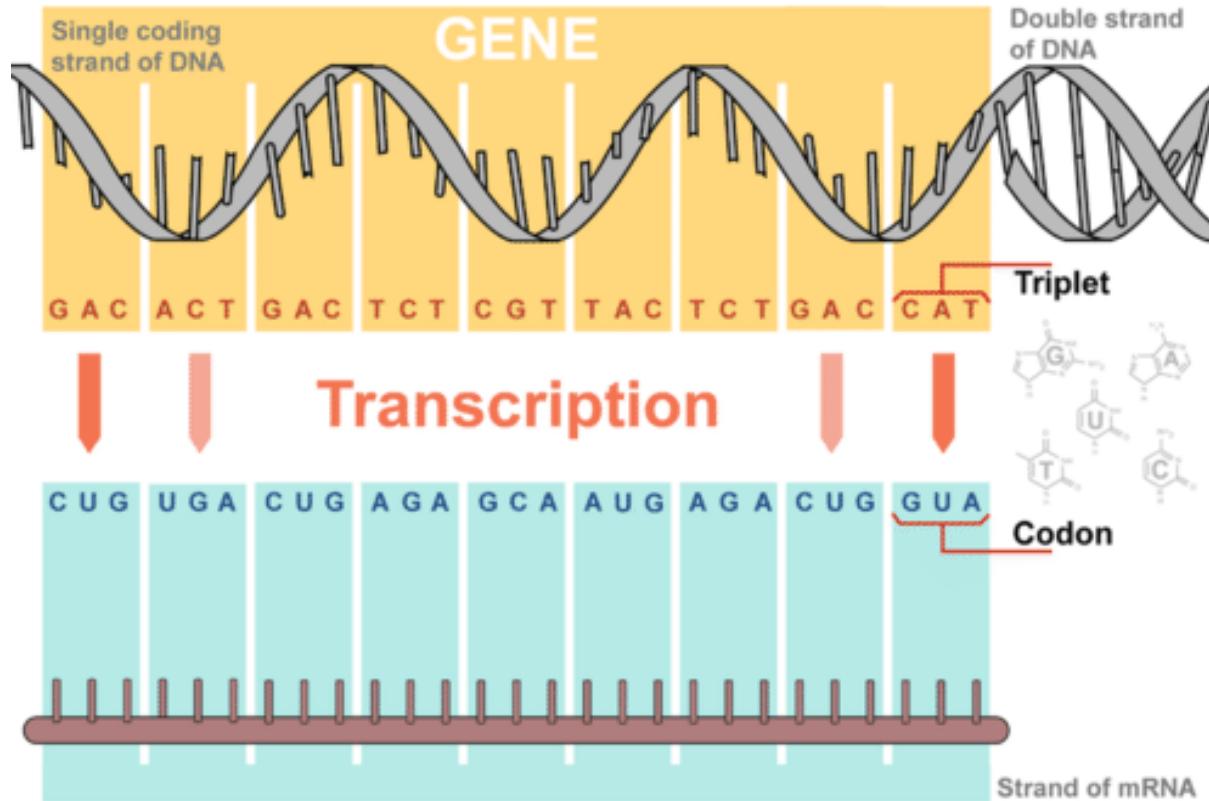
**CLOSED**

Transcription OFF

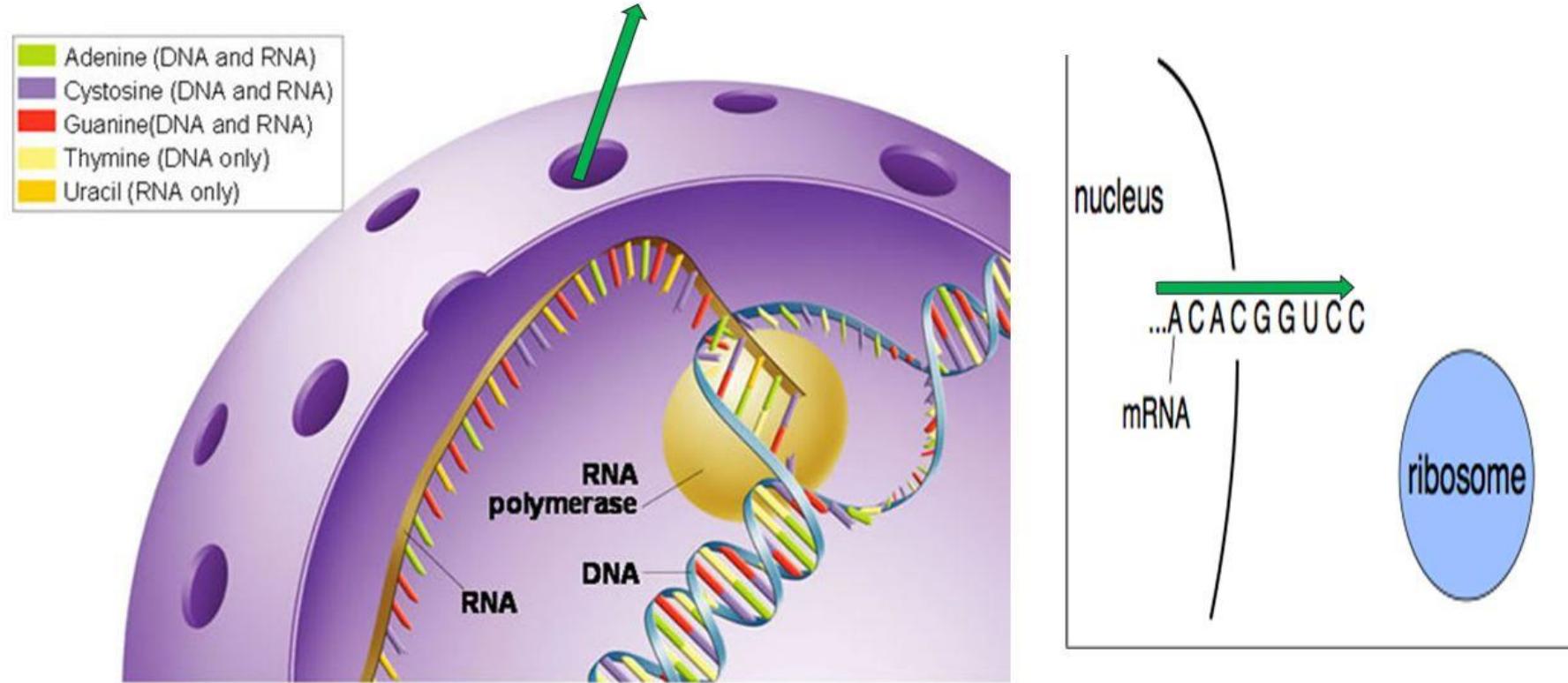
**OPEN**

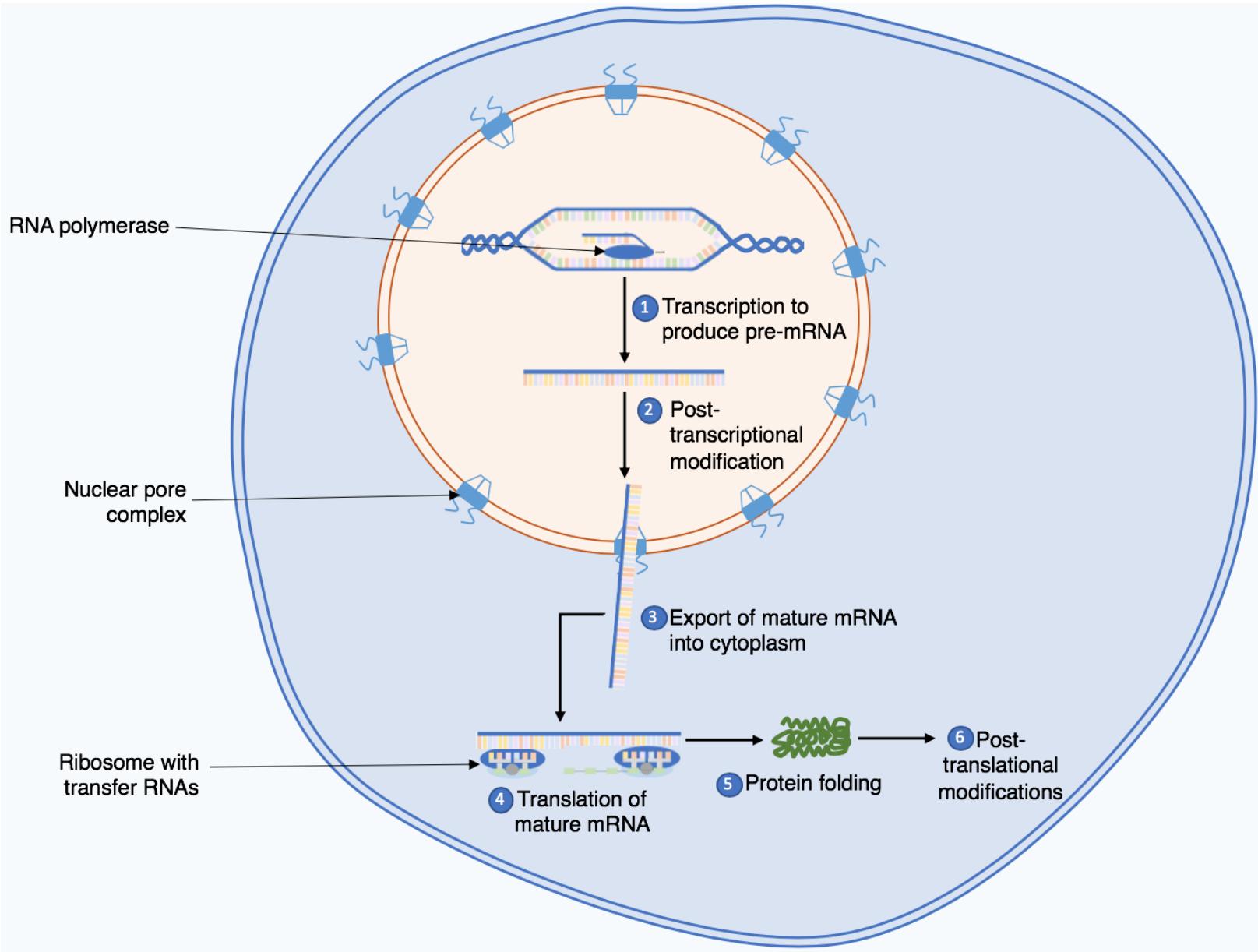
Transcription ON





## mRNA leaves nucleus through nuclear pores

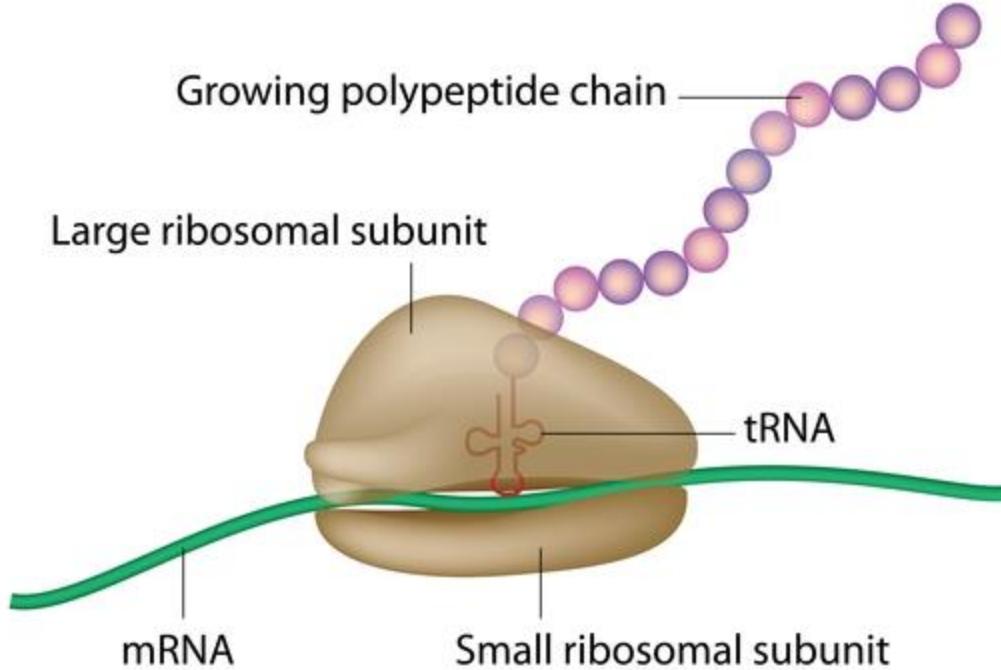




# TRANSLATION

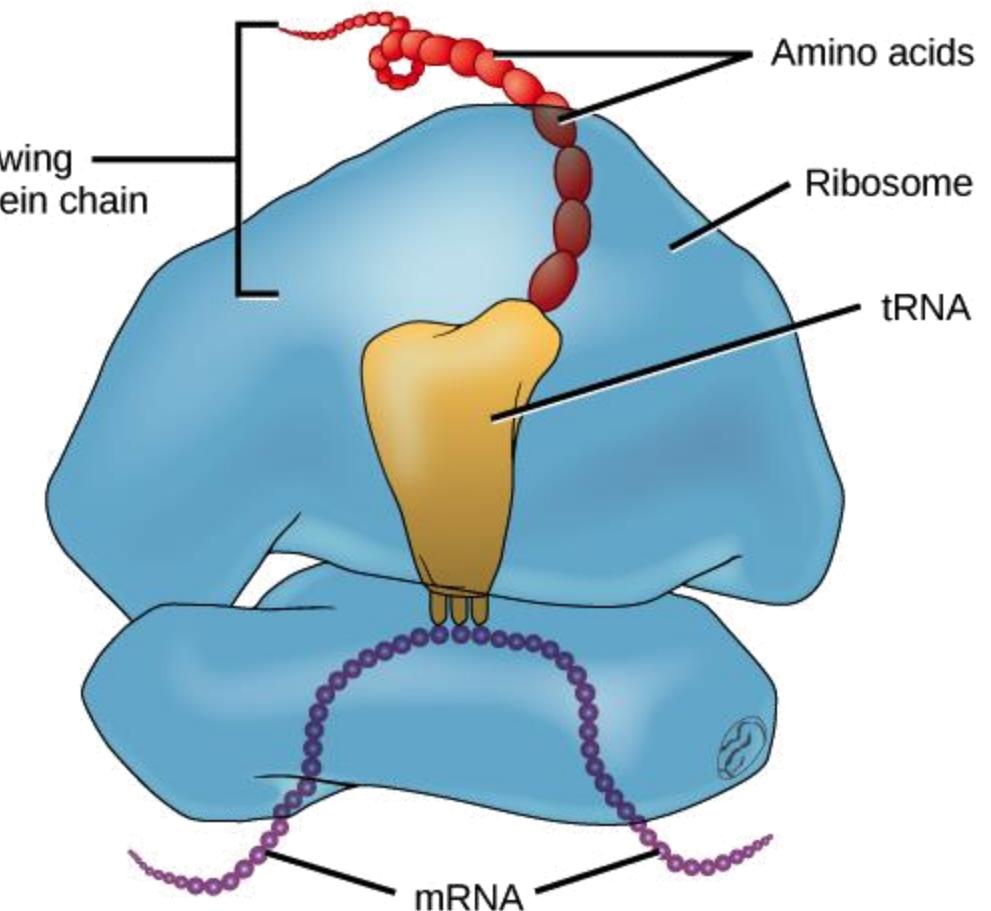
- requires some specialized equipment, a cell couldn't translate an mRNA into a protein without two pieces of molecular gear:
  - ribosomes and
  - tRNAs
- Ribosomes provide a structure in which translation can take place. They also catalyze the reaction that links amino acids to make a new protein.
- tRNAs (transfer RNAs) carry amino acids to the ribosome. They act as "bridges," matching a codon in an mRNA with the amino acid it codes for.

# RIBOSOM



# RIBOSOMES: WHERE THE TRANSLATION HAPPENS

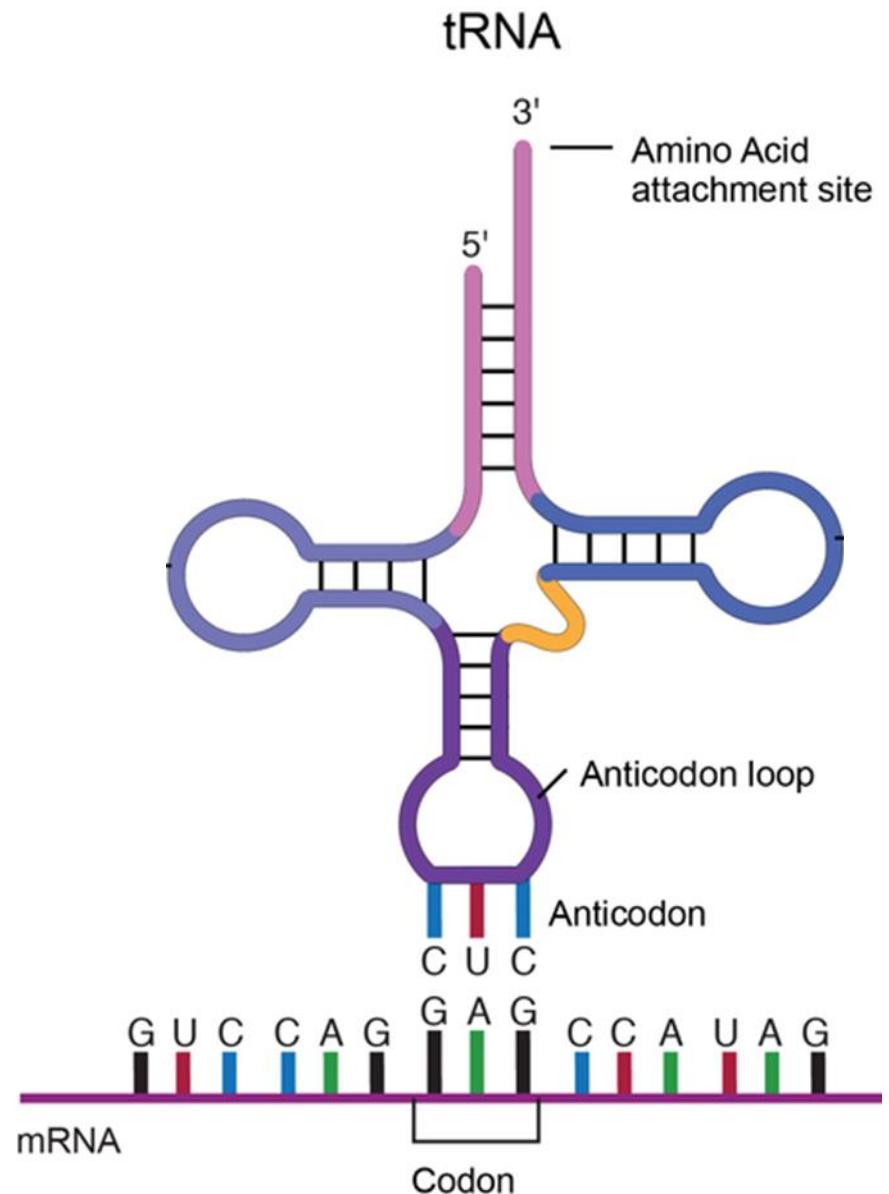
- Translation takes place inside structures called ribosomes, which are made of RNA and protein. Ribosomes organize translation and catalyze the reaction that joins amino acids to make a protein chain.
- A ribosome is made up of two basic pieces: a large and a small subunit. During translation, the two subunits come together around a mRNA molecule, forming a complete ribosome. The ribosome moves forward on the mRNA, codon by codon, as it is read and translated into a polypeptide (protein chain). Then, once translation is finished, the two pieces come apart again and can be reused.



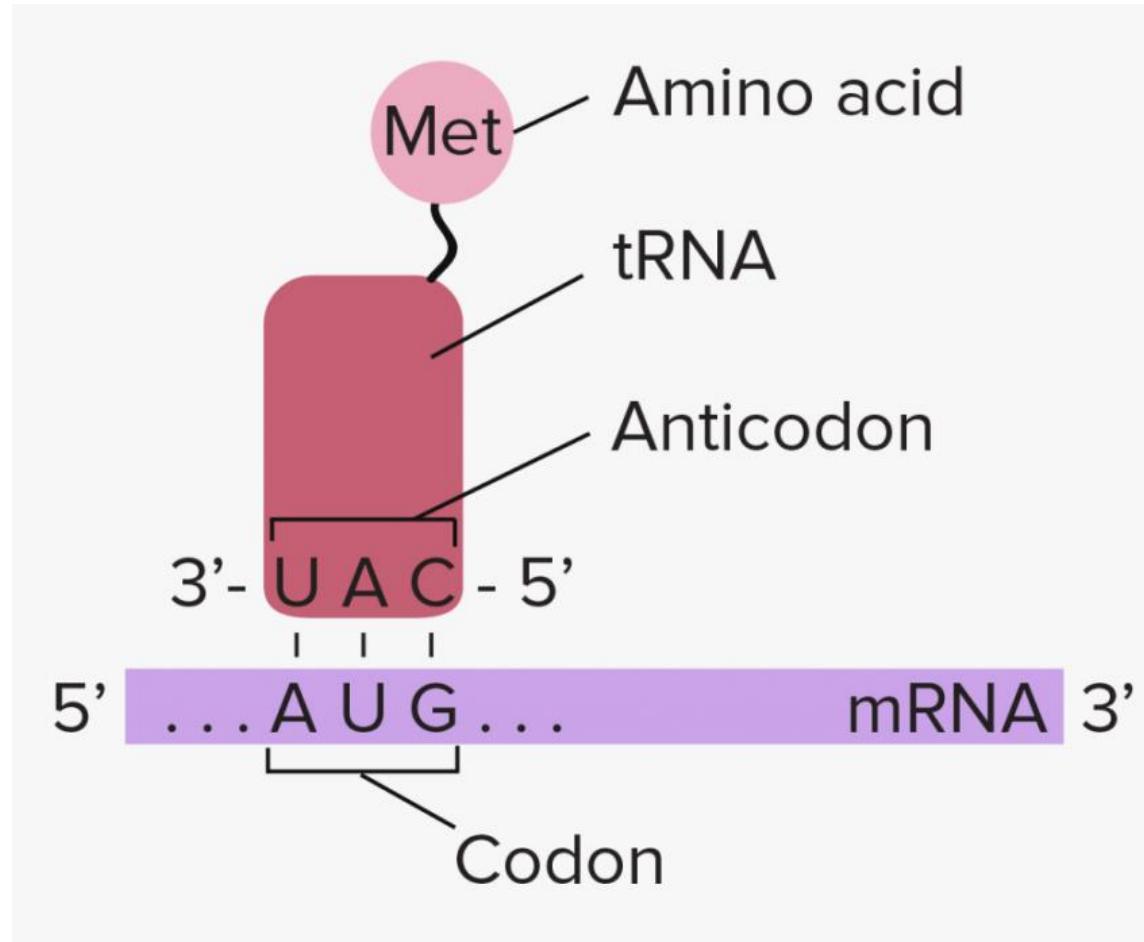
# TRANSFER RNA

- Transfer RNA (tRNA) is a small RNA molecule that participates in protein synthesis.
- Each tRNA molecule has two important areas:
  - a trinucleotide region called the anticodon and
  - a region for attaching a specific amino acid.

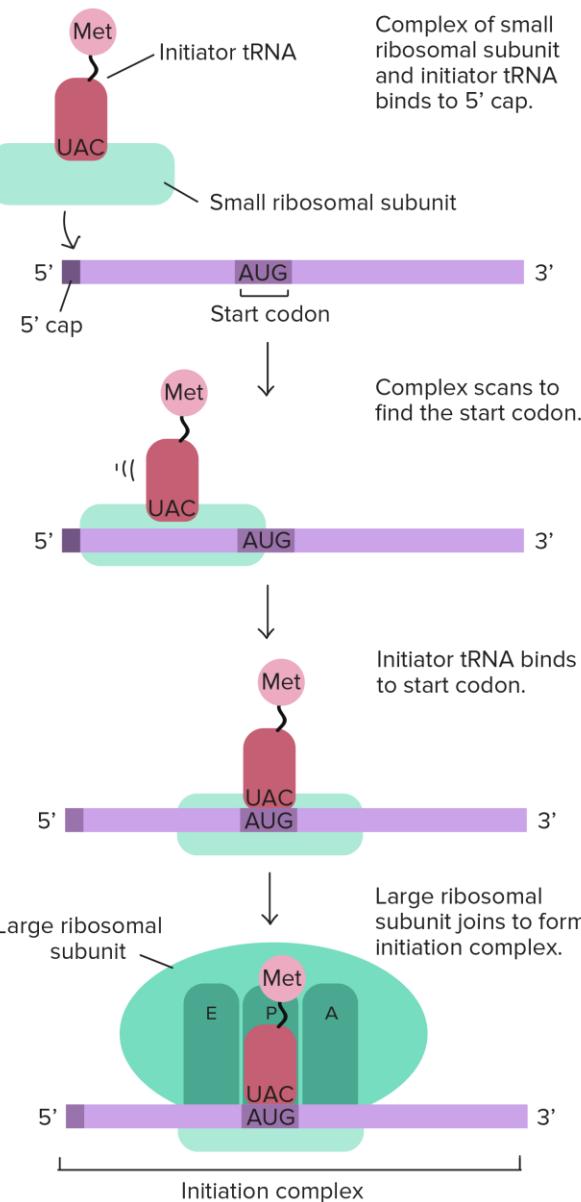
During translation, each time an amino acid is added to the growing chain, a tRNA molecule forms base pairs with its complementary sequence on the messenger RNA (mRNA) molecule, ensuring that the appropriate amino acid is inserted into the protein.



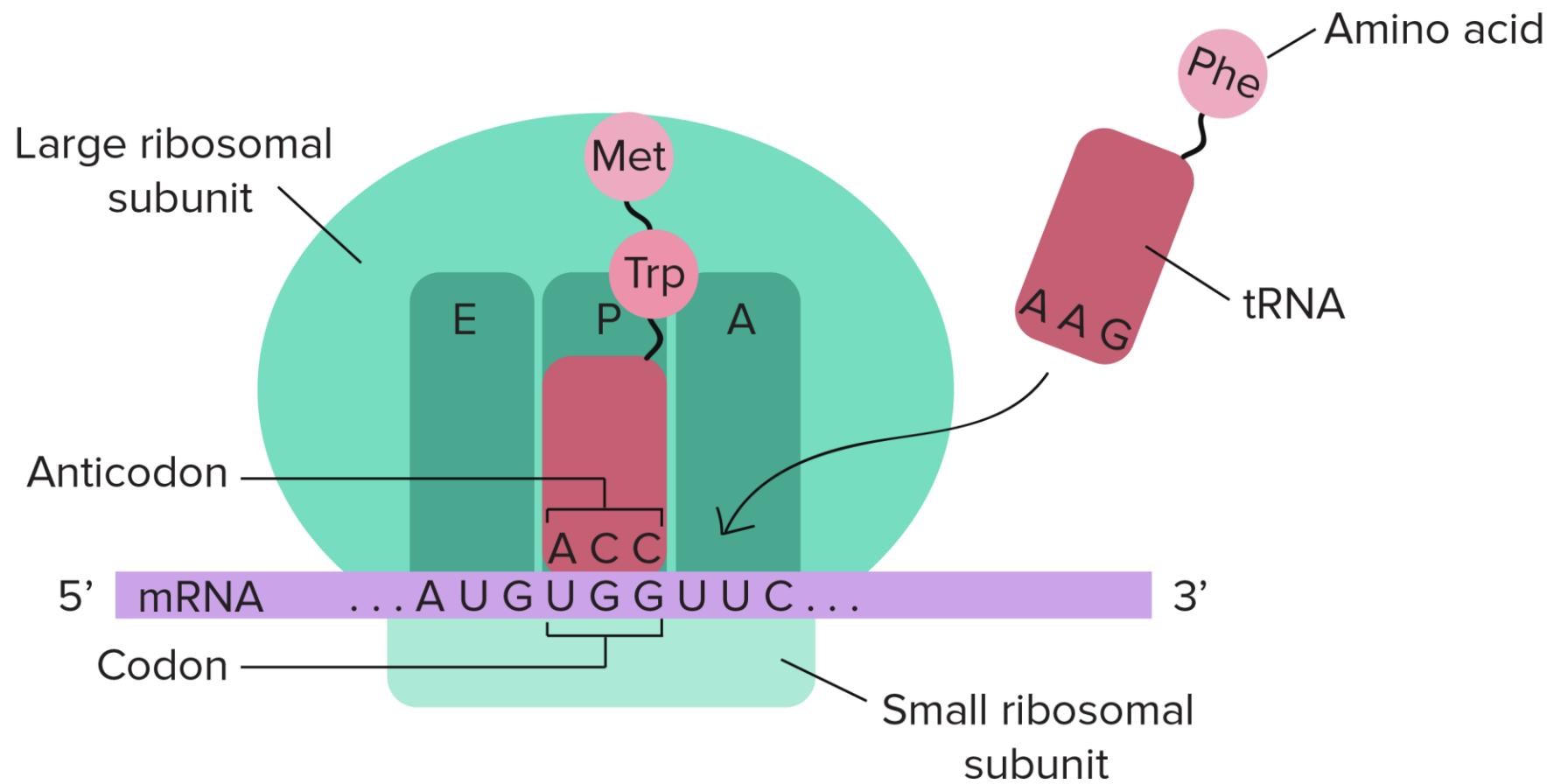
Each tRNA has an anticodon, a set of three nucleotides that binds to a matching mRNA codon through base pairing. The other end of the tRNA carries the amino acid that's specified by the codon.



## Eukaryotic translation initiation



The ribosome has three slots for tRNAs: the A site, P site, and E(xit) site.  
tRNAs move through these sites (from A to P to E) as they deliver amino acids during translation.



**E-site**      **P-site**      **A-site**

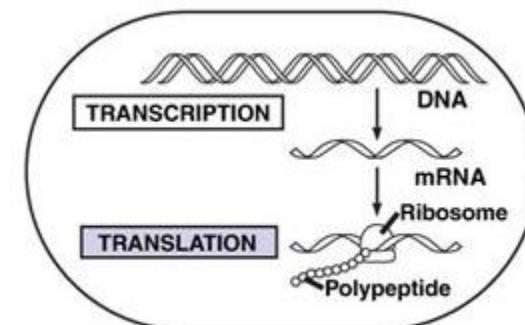
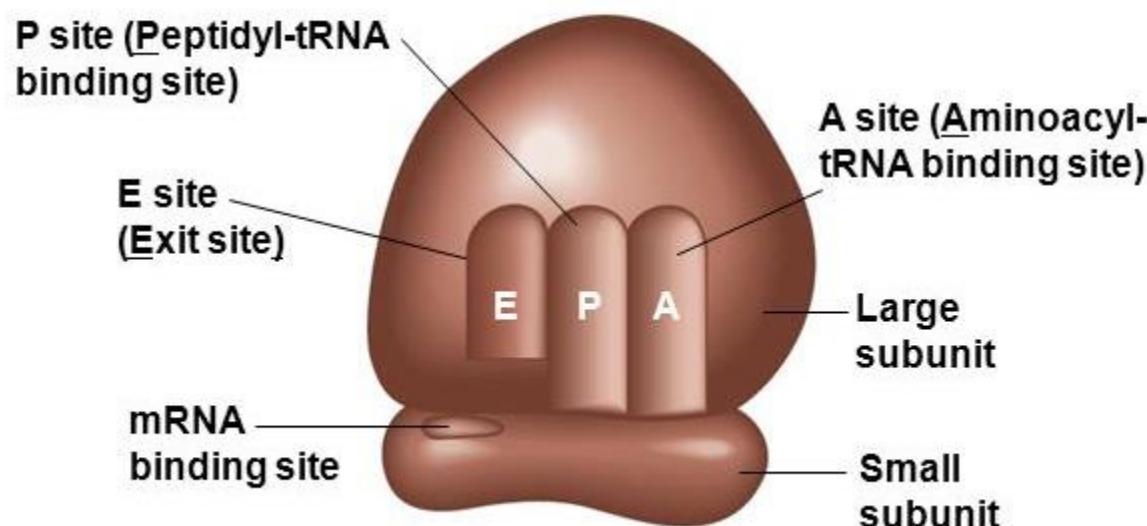
**E-site** is where  
the tRNA exits  
after it drops  
off its amino  
acid.

**A-site** is where  
the tRNA binds  
to bring the  
AMINO ACIDS



**P-site** is where the  
growing POLYPEPTIDE is  
kept.

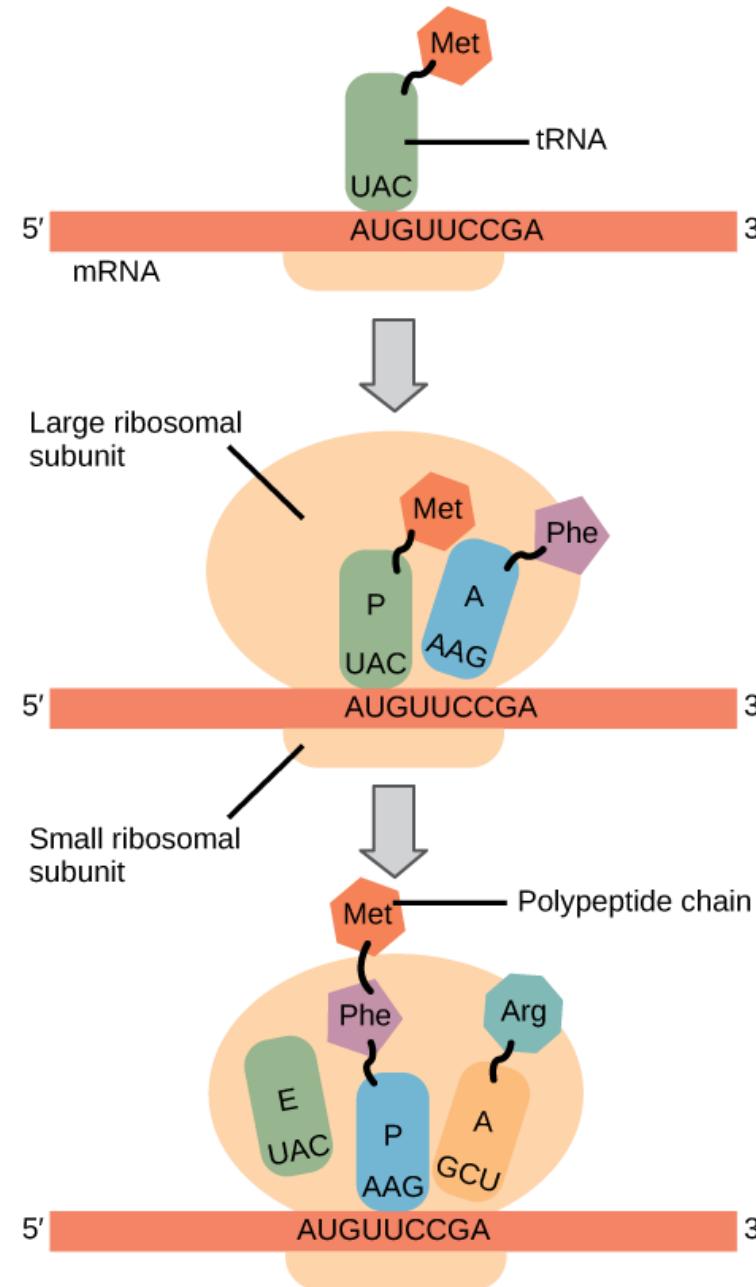
## Schematic model showing binding sites on ribosome

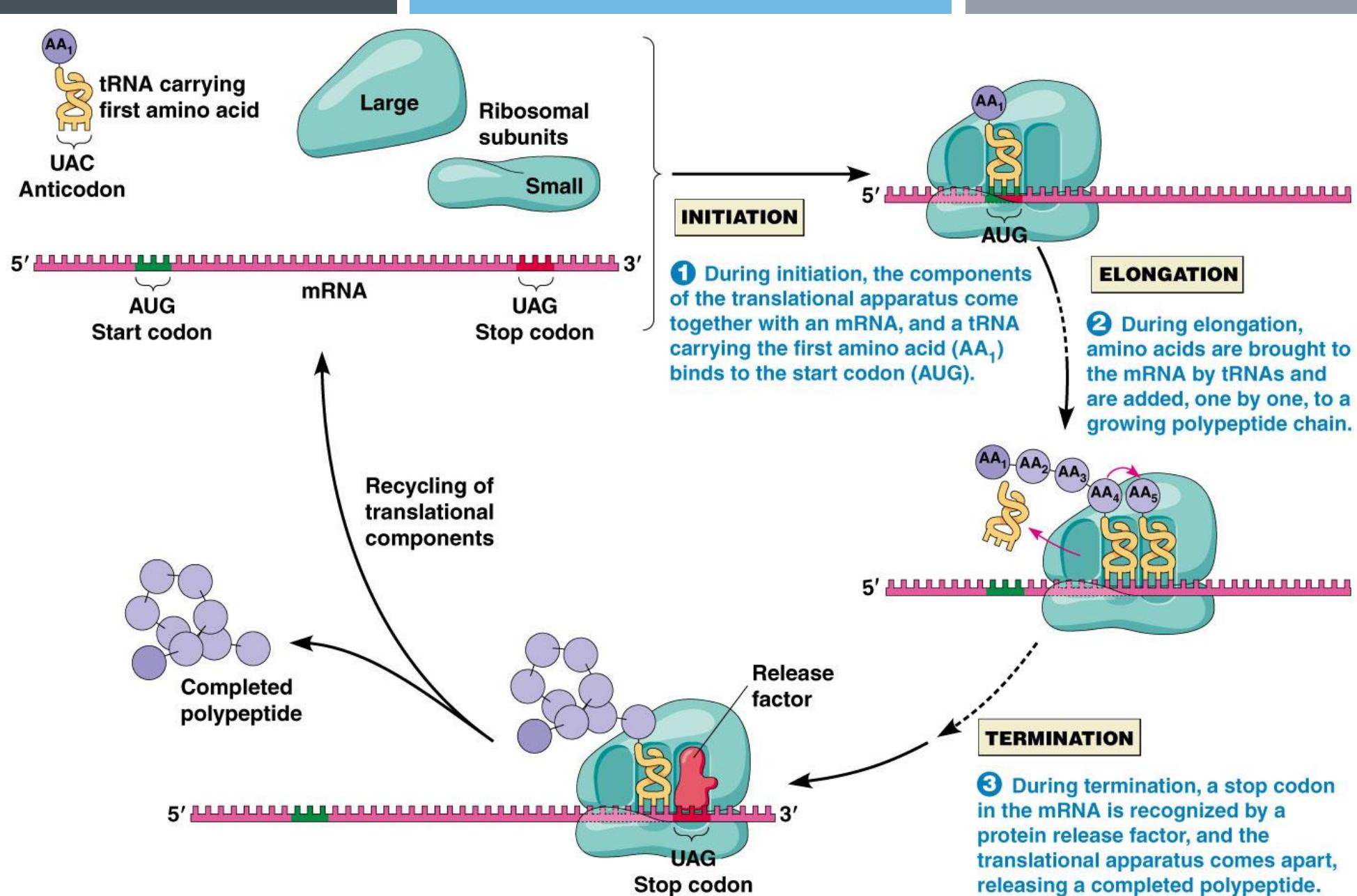


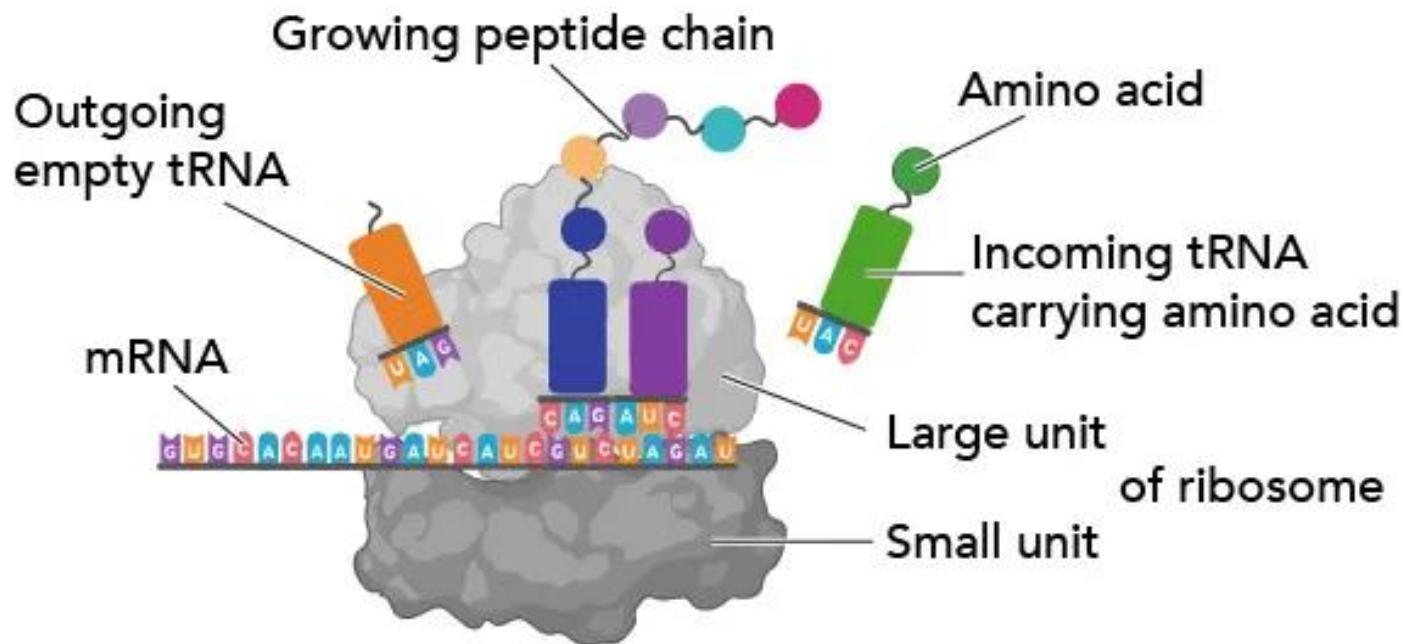
Initiation: the start codon (always AUG) is identified on the mRNA chain.  
tRNA bind to mRNA inside ribosome.

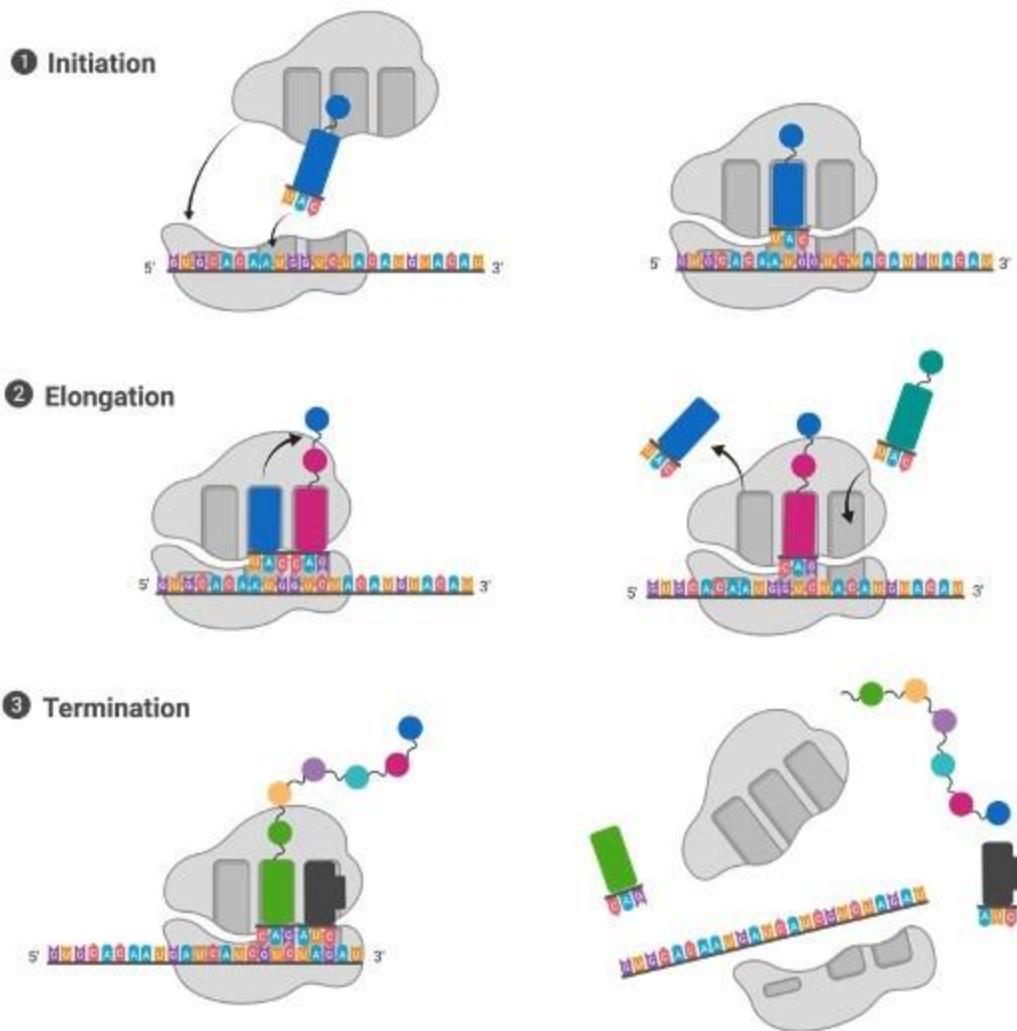
The second stage of protein translation is elongation: As tRNAs enter slots in the ribosome and bind to codons, their amino acids are linked to the growing polypeptide chain in a chemical reaction.

The result is a polypeptide whose amino acid sequence mirrors the sequence of codons in the mRNA.

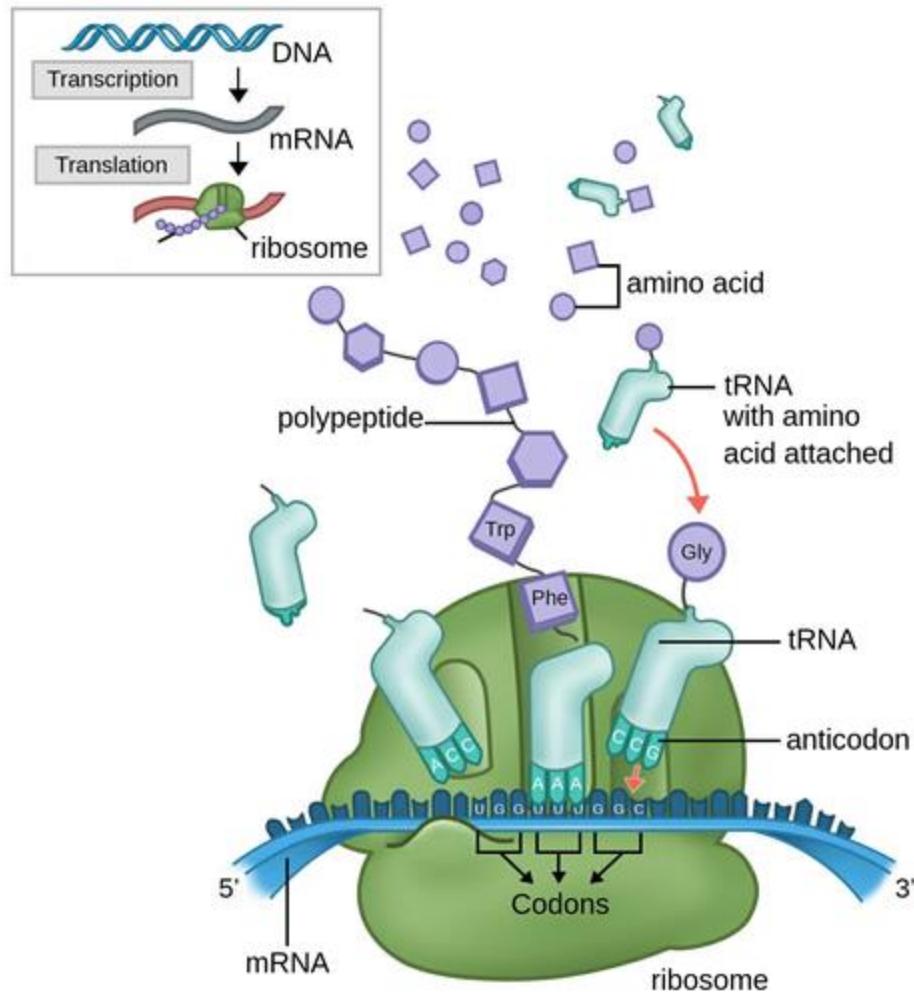


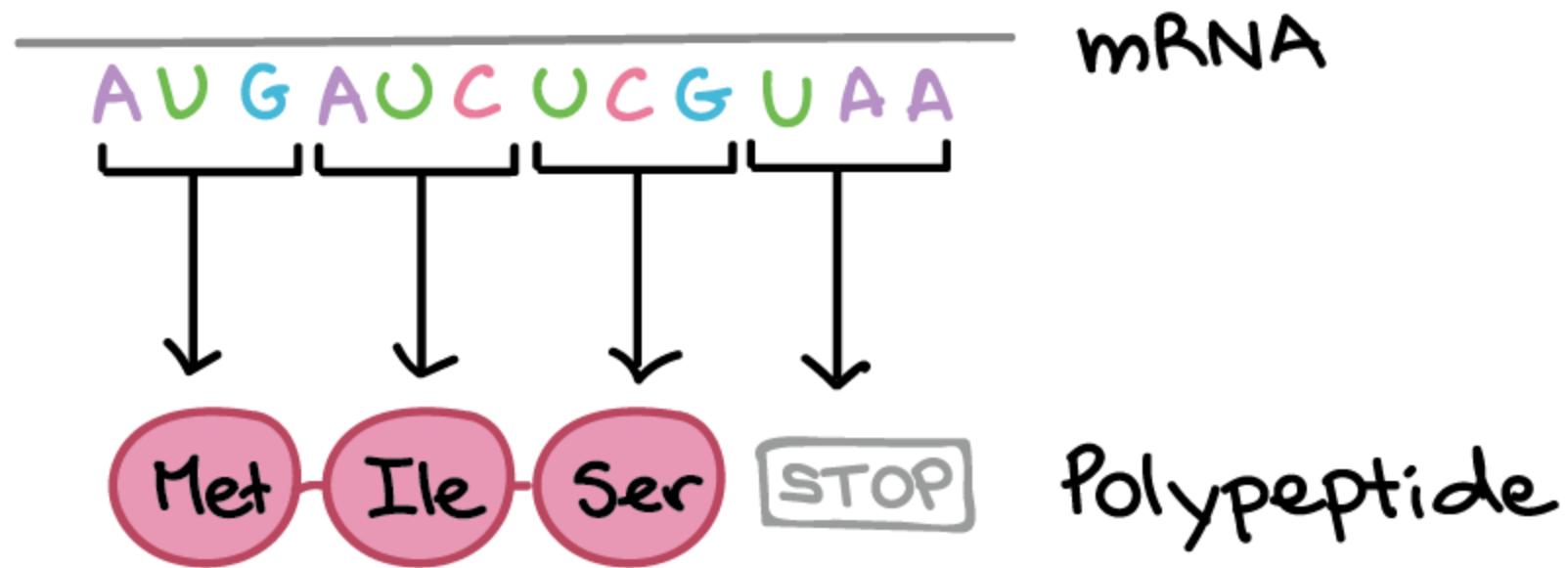






*Created with BioRender*





## What does our bodies do with proteins?

- Our bodies will break down the proteins into building blocks called amino acids and rebuild what we need...

It's like breaking apart a Lego house to rebuild something else.



Essential amino acids can be found in many different foods. The best sources of amino acids are found in animal proteins such as beef, poultry and eggs (complete proteins). Animal proteins are the most easily absorbed and used by your body.



## NON ESSENTIAL

Alanine  
Arginine  
Asparagine  
Aspartate  
Cystine  
Glutamic  
Glycine  
Ornithine  
Proline  
Serine  
Tyrosine

## ESSENTIAL

Histidine  
Isoleucine  
Leucine  
Lysine  
Methionine  
Phenylalanine  
Threonine  
Tryptophan  
Valine



### Essential amino acids

These cannot be synthesized within the body

Threonine  
Histidine  
Tryptophan  
Valine  
Leucine  
Isoleucine

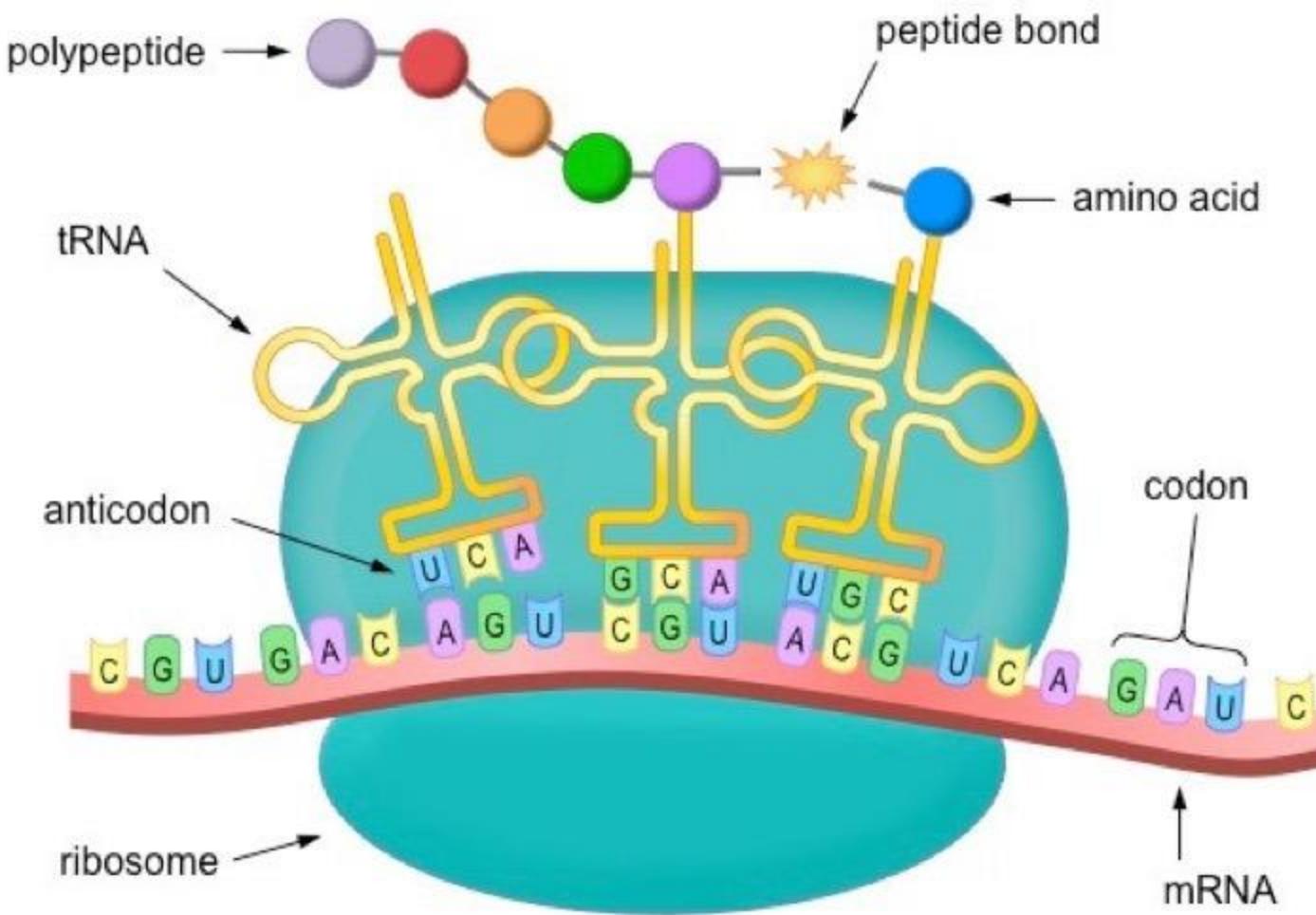
These are included in protein that forms muscles.  
They account for 30-40% of essential amino acids.

### Non-essential amino acids

These can be synthesized within the body

Alanine  
Aspartic acid  
Glycine  
Asparagine  
Serine  
Proline  
Glutamic acid  
Arginine  
Glutamine  
Cysteine  
Tyrosine

All amino acids are required for body growth.  
Since "essential amino acids" cannot be synthesized within the body, they have to be consumed in the form of food.



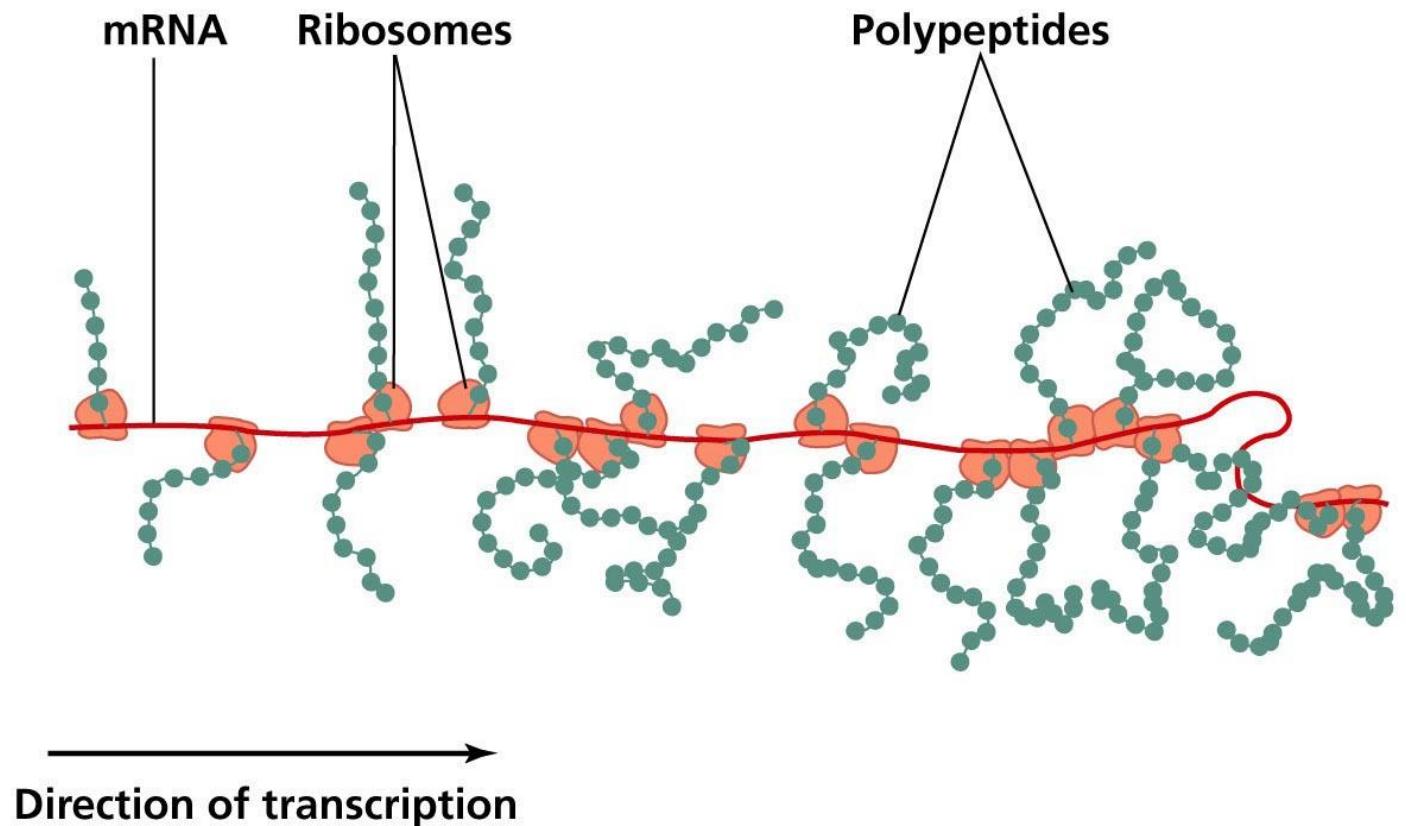
# TRANSLATION (SUMMARY)

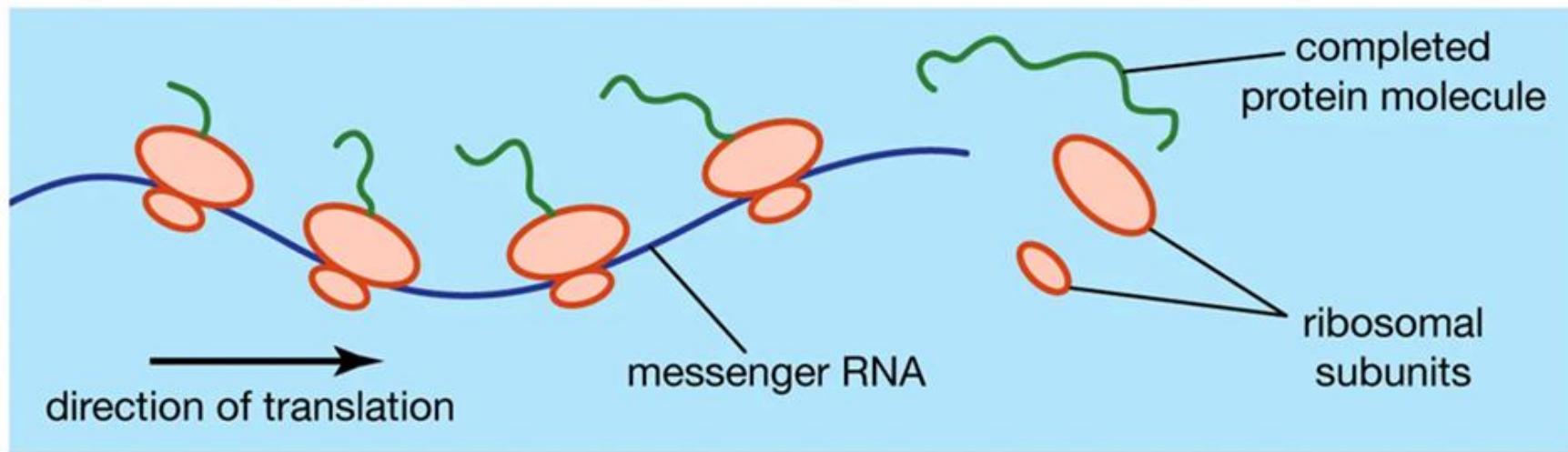
Once mRNA has left the nucleus, it is directed to a ribosome to construct a protein. This process can be broken down into 6 main stages:

- Initiation: Ribosome attaches to the mRNA molecule at the start codon. This sequence (always AUG) signals the start of the gene to be transcribed. The ribosome can enclose two codons at a time
- tRNAs (transfer RNAs) act as couriers. There are many types of tRNA, each one complementary to the 64 possible codon combinations. Each tRNA is bonded to a specific amino acid. As AUG is the start codon, the first amino acid to be 'couriered' is always Methionine.
- Elongation: The stepwise addition of amino acids to the growing polypeptide chain. The next amino acid tRNA attaches to the adjacent mRNA codon.
- The bond holding the tRNA and amino acid together is broken, and a peptide bond is formed between the adjacent amino acids.
- As the Ribosome can only cover two codons at a time, it must now shuffle down to cover a new codon. This releases the first tRNA which is now free to collect another amino acid. Steps 2-5 repeats along the whole length of the mRNA molecule.
- Termination: As the polypeptide chain elongates, it peels away from the ribosome. During this phase, the protein starts to fold into its specific secondary structure. Elongation continues (perhaps for hundreds or thousands of amino acids) until the ribosome reaches one of three possible Stop codons (UAG, UAA, UGA). At this point, the mRNA dissociates from the ribosome

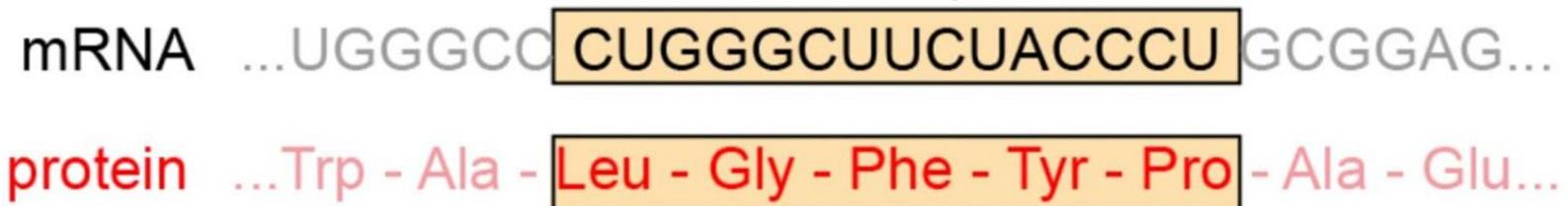
# POLYRIBOSOMS

Several ribosomes may translate the same mRNA molecule simultaneously. The complex of ribosomes along a single mRNA is called a polyribosome or polysome.





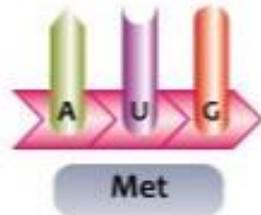
## THE GENETIC CODE



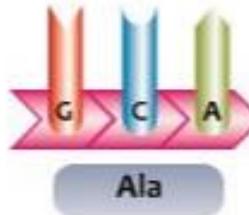
- Translation involves “decoding” a messenger RNA (mRNA) and using its information to build a chain of proteins (a polypeptide or protein).
- These relationships between mRNA codons and amino acids are known as the genetic code.
  
- In an mRNA, the instructions for building a polypeptide come in groups of three nucleotides called codons.
  - There are 64 different codons for amino acids
  - One codon, AUG, is a “start” signal to kick off translation (it also specifies the amino acid methionine)
  - “stop” codons mark the polypeptide as finished

A codon is made of 3 base pairs  
64 codons total

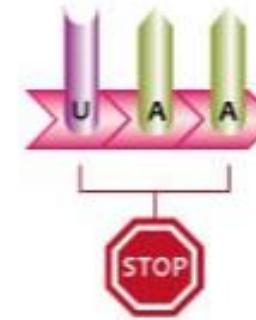
1 codon (AUG) encodes methionine and starts translation of all proteins



61 codons encode 20 amino acids (redundant code)



3 codons stop protein translation



# THE AMINO ACIDS CODON CHART

- The full set of relationships between codons and amino acids (or stop signals) is called the genetic code.
- The genetic code is often summarized in a codon chart (or codon table), where codons are translated to amino acids.

|              |   | Second letter            |            |             |             | Third letter     |
|--------------|---|--------------------------|------------|-------------|-------------|------------------|
|              |   | U                        | C          | A           | G           |                  |
| First letter | U | UUU<br>UUC<br>UUA<br>UUG | Phe<br>Ser | Tyr<br>Stop | Cys<br>Stop | U<br>C<br>A<br>G |
|              | C | CUU<br>CUC<br>CUA<br>CUG | Leu        | Pro         | His<br>Gln  | U<br>C<br>A<br>G |
|              | A | AUU<br>AUC<br>AUA<br>AUG | Ile        | Thr         | Asn<br>Lys  | U<br>C<br>A<br>G |
|              | G | GUU<br>GUC<br>GUA<br>GUG | Val        | Ala         | Asp<br>Glu  | U<br>C<br>A<br>G |

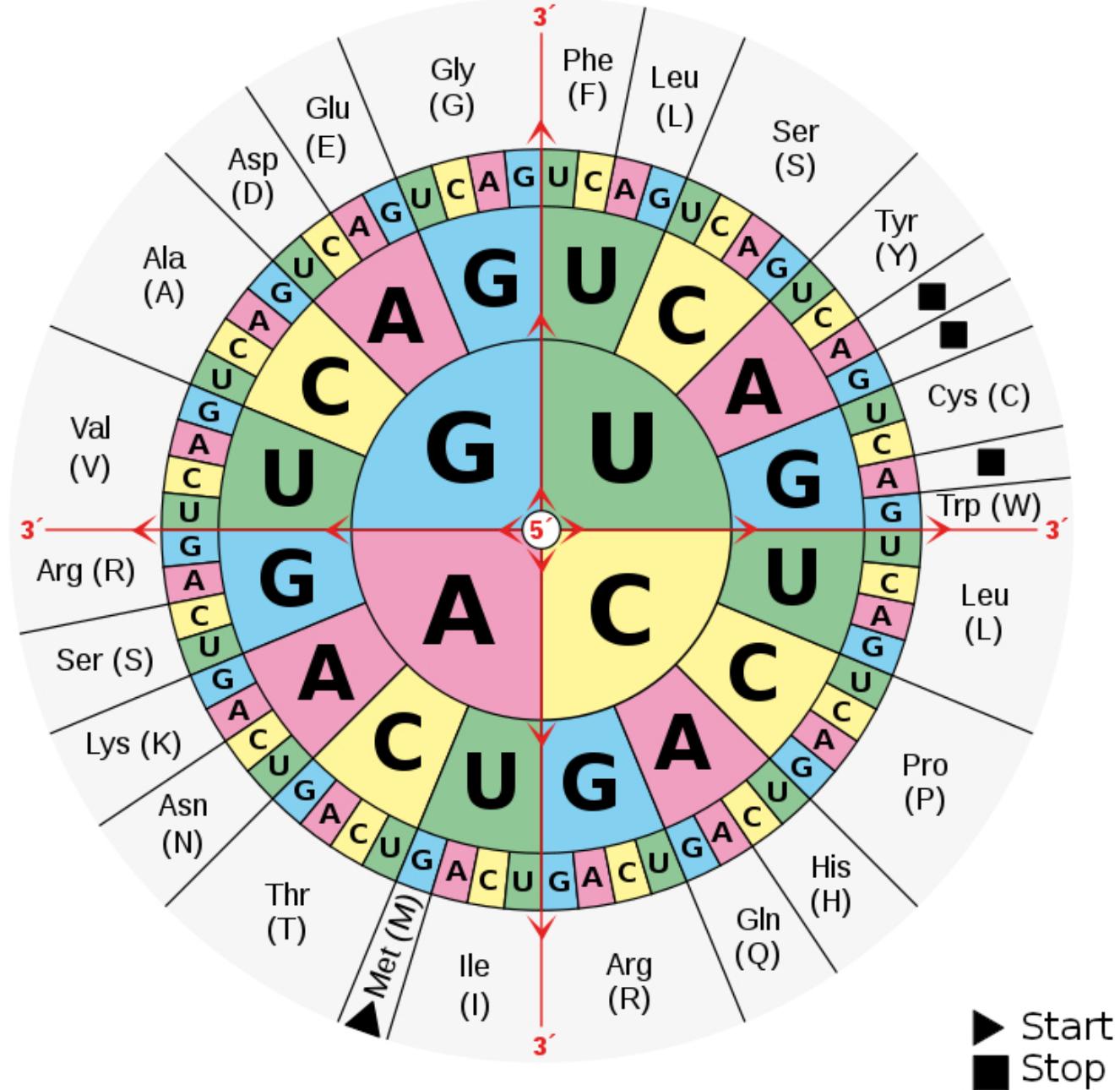
# HOW DO YOU READ THE CODON CHART?

- Let's take codon ACU as an example. If you want to know which amino acid ACU encodes, you first look at the left side of the table. Find the "A" on the axis of the left side, which refers to the first letter of the codon triplet. All these codons starting with "A" are in this row.
- Next, we look at the top of the table. This upper axis indicates the second letter of the codon triplet. Once we find "C" along the upper axis, it tells us about the column in which our codon will be found. Find the intersecting box of "A" row and "C" column in the table. You will see this box containing four codons and easily find the one you're looking for.
- In our example, ACU encodes Thr (or Threonine). You may also notice that all ACU, ACC, ACA, and ACG encode the same amino acid. Notice that many amino acids are represented in the table by more than one codon.

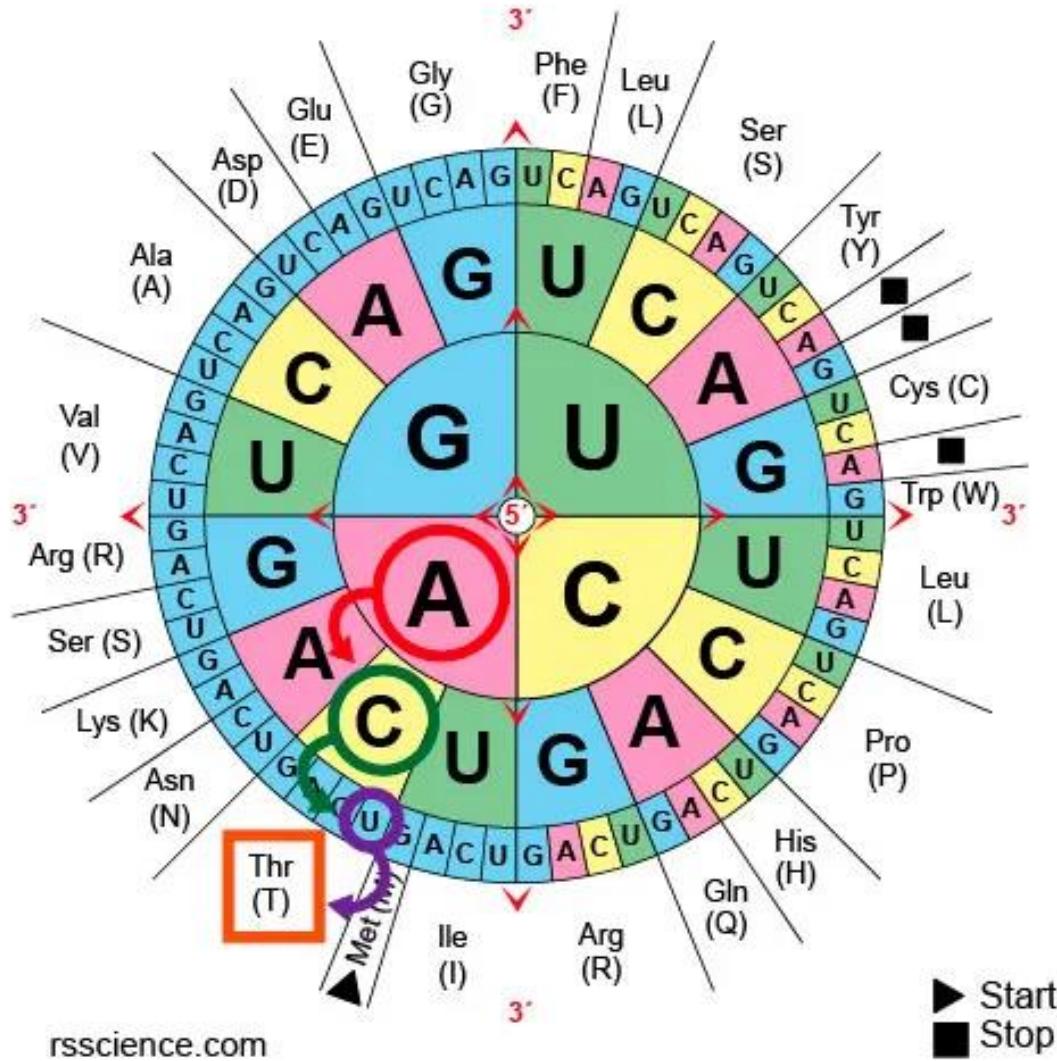
|              |   | Second letter                    |                              |  |                                       |         |  |
|--------------|---|----------------------------------|------------------------------|--|---------------------------------------|---------|--|
|             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

rsscience.com

Condon sets can also be presented as a codon wheel (sun model).



For a codon wheel,  
the rule is the same:  
start from the center  
to find the first letter  
of triplet, then move  
toward the periphery  
for 2nd and 3rd  
letters.



## EXAMPLE: GENETIC CODE

GENE 1

TACTTGTTCACATAACCTTGAAATT

Step 1 - transcribe the DNA into mRNA



Step 2 - draw lines to separate the mRNA into codons (3 bases)

- How many amino acids are present in this gene?
- List each of the amino acids in this gene:

GENE 1

TACTTGTTCACATAACCTTGAAATT

Step 1 - transcribe the DNA into mRNA

AUG|AAC|AAA|U|GU|AU|U|G|AA|ACU|U|AA

Step 2 - draw lines to separate the mRNA into codons (3 bases)

- How many amino acids are present in this gene?
- List each of the amino acids in this gene:

8

Met-Asn-Lys-Cys-Ile-Glu-Thr-Stop

## DNA, RNA and Protein Synthesis

1. Define the following terms:

a. **Replication**

b. **Transcription**

c. **Translation**

2. Break the following DNA sequence into triplets. (Draw a line to separate triplets)

**CCGATACGCGGTTTCCCAGGG CTAATTAA**

3. If the above code showed the bases on one strand of DNA, what would the complementary strand read?

4. What would the code in problem #2 be transcribed into if it was copied into mRNA? (RNA is the same as DNA except instead of Thymine it has Uracil. That means A combines with U instead of T)

What would the **amino acid sequence** be, translated from the mRNA sequence the genetic code table provided to translate)

1. Define the following terms:

a. **Replication**- The process by which DNA is duplicated before a cell divides

b. **Transcription**- The process by which a molecule of DNA is copied into a complementary strand of mRNA

**Translation**- Process in which a message carried by messenger RNA is decoded into a polypeptide chain (protein) by ribosomes in the cytoplasm

2. Break the following DNA sequence into triplets. (Draw a line to separate triplets)

**CCG|ATACG|CGG|TTT|CCC|AGGG|CTA|ATT|AA**

3. If the above code showed the bases on one strand of DNA, what would the complementary strand read?

**GGC TAT GCG CCA AAG GGT CCC GAT TAA ATT**

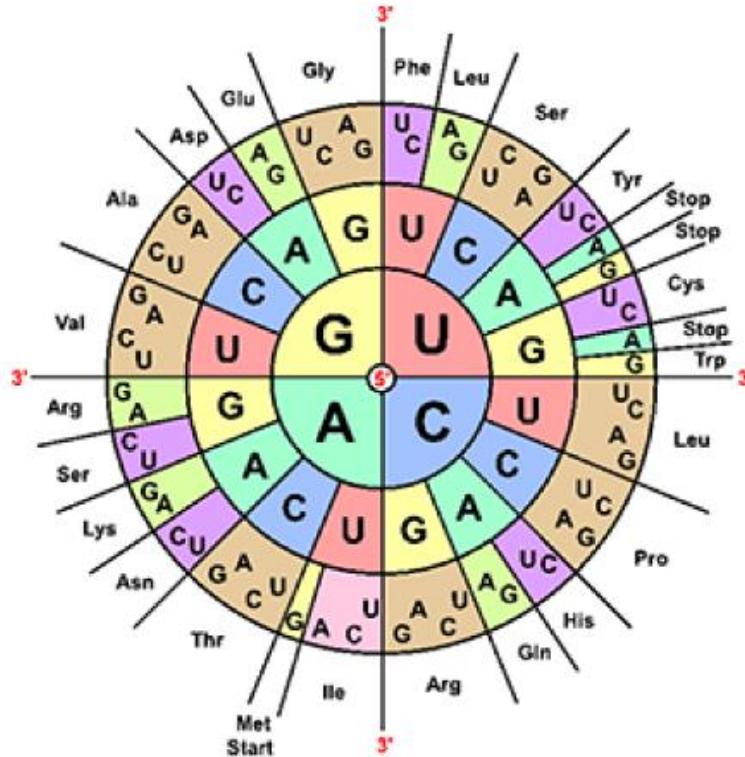
4. What would the code in problem #2 be transcribed into if it was copied into mRNA? (RNA is the same as DNA except instead of Thymine it has Uracil. That means A combines with U instead of T)

**GGC UAU GCG CCA AAG GUU CCC GAU UAA AUU**

What would the **amino acid sequence** be, translated from the mRNA sequence the genetic code table provided to translate)

**Gly-tyr-ala-pro-lys-gly-pro-asp-stop-ile**

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



1. Use the codon chart to write the amino acid that corresponds to each codon found in mRNA:

C C C \_\_\_\_\_  
C A G \_\_\_\_\_  
G A A \_\_\_\_\_  
U U U \_\_\_\_\_

A G U \_\_\_\_\_  
U A C \_\_\_\_\_  
C G U \_\_\_\_\_  
C C A \_\_\_\_\_

1. Use the codon chart to write the amino acid that corresponds to each codon found in mRNA:

C C C Pro  
C A G Gln  
G A A Glu  
U U U Phe  
A G U Ser  
U A C Tyr  
C G U Arg  
C C A Pro

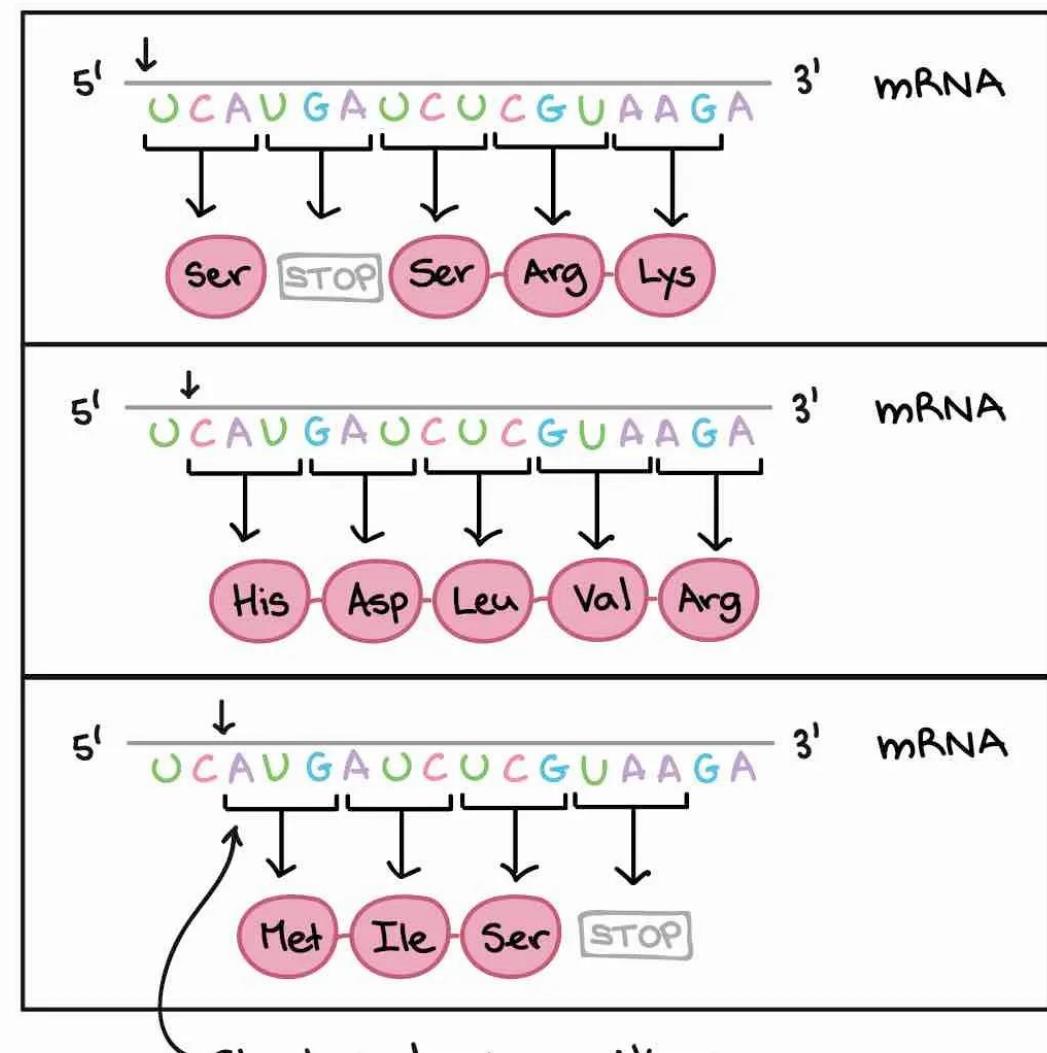
# READING FRAME

- Since the DNA sequence is read by triplets, starting from which letter (or reading frame) becomes a critical problem.
- Let's look at an example. The mRNA below can be translated into three totally different orders of amino acids, depending on the frame in which it's read. How do our cells know which of these proteins to make?
- Our cells use a very smart strategy to solve this problem – the “start codon”. Because the translation only begins at the start codon (AUG) and continues in successive groups of three, the position of the start codon ensures that the mRNA is read in the correct frame (in the example above, in Frame 3).

FRAME 1

FRAME 2

FRAME 3



Start codon's position  
ensures that this  
frame is chosen

# WHAT HAPPENS IF THE DNA SEQUENCES ARE MESSED UP – MUTATION

Mutations (changes in DNA sequences) may derail the genetic information and cause cells to make the wrong proteins. Mutations are the major cause of cancers and many genetic disorders.

Even a single base pair altered (called point mutation) can cause a significant consequence. Point mutations can have one of three effects.

- Silent mutation

First, the base substitution can be a silent mutation where the altered codon corresponds to the same amino acid. For example, changing from UCU to UCC has no effect since both codons equally encode Serine (Ser).

- Missense mutation

Second, the base substitution can be a missense mutation where the altered codon corresponds to a different amino acid. For example, changing from UCU to UGU will turn Serine (Ser) to Cysteine (Cys). If this mutation happens in the critical region (i.e., enzymatic site) of the protein, a point mutation can mess up the whole protein function.

- Nonsense mutation

Third, the base substitution can be a nonsense mutation where the altered codon becomes a stop signal. This is the worst cause because the translation will terminate too early, resulting in a truncated protein.

### Normal

Amino Acids — Ala Ile Arg Leu Gly Tyr Ser Ala Cys Ile His Val Ala Ile Arg →  
tRNA  
anticodon ...CGAUUAUUCGAUCCAAUGUCACGUACGUAGUGCAUCGUAUAGCG...  
mRNA ...GCUAUAAGGCUAGGUUACAGUGCAUGCAUACACGUAGCUUAACG...  
5' codons 3'



Protein

### Missense mutation

Amino Acids — Ala Ile Arg Leu **Ala** Tyr Ser Ala Cys Ile His Val Ala Ile Arg →  
tRNA  
anticodon ...CGAUUAUUCGAUCCGAUCGAAUGUCACGUACGUAGUGCAUCGUAUAGCG...  
mRNA ...GCUAUAAGGCUAG**C**UACAGUGCAUGCAUACACGUAGCUUAACG...  
5' codons 3'



Protein

Missense mutation

### Nonsense mutation

Amino Acids — Ala Ile Arg Leu Gly Tyr Ser Ala Cys stop →  
tRNA  
anticodon ...CGAUUAUUCGAUCCAAUGUCACGUACGAUU  
mRNA ...GCUAUAAGGCUAGGUUACAGUGCAUG**C**UAAACACGUAGCUUAACG...  
5' codons 3'



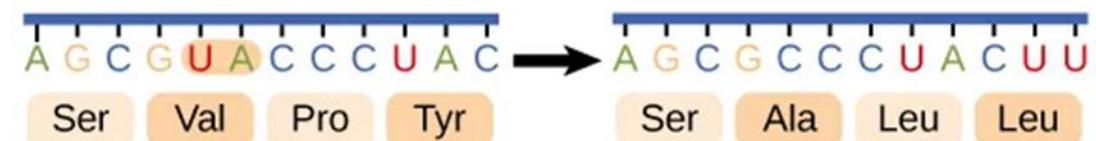
Protein

Nonsense mutation

# FRAMESHIFT MUTATIONS

- Mutations could also happen when nucleotides are inserted or deleted from the original DNA sequence. The insertion or deletion of “one or two” nucleotides can change the reading frame. A frameshift can totally mess up the amino acid orders.

## Frameshift Mutations



# GENETIC CODE

To crack the genetic code, researchers needed to figure out how nucleotides sequences in a DNA or RNA molecule could encode the sequence of amino acids. In the mid-1950s, it was predicted that the genetic code is likely composed of triplets of nucleotides – because the possible combination of duplet is not enough ( $4 \times 4 = 16$ ), and that of quadruplet is too many ( $4 \times 4 \times 4 \times 4 = 256$ ), to cover 20 kinds of amino acids.

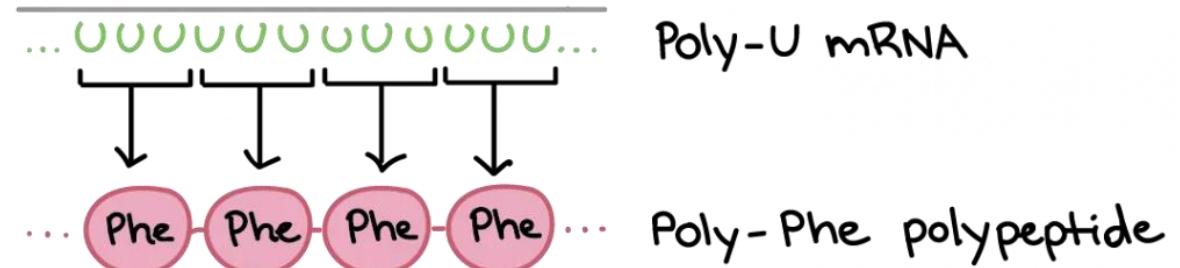
## Redundant but not Ambiguous

- Redundant- multiple codons can code for the same amino acid
- Not Ambiguous- no codon codes for more than one amino acid

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

Copyright © 2005 Pearson Education, Inc. Publishing as Pearson Benjamin Cummings. All rights reserved.

# HOW WAS THE GENETIC CODE DISCOVERED?



- The actual experiments began in 1961 by American biochemist Marshall Nirenberg. Nirenberg was able to link the relationships between nucleotide triplets to particular amino acids by two experimental innovations:
  - He synthesized artificial mRNA molecules with specific, known sequences.
  - He had a system to translate mRNAs into polypeptides outside of a cell (a “cell-free” system). Nirenberg did so in a test tube of cytoplasm from burst E. coli bacteria, which contains all the ingredients needed for translation.
- Nirenberg started with an mRNA molecule consisting only of the nucleotide Uracil (called poly-U). When he added poly-U mRNA to the cell-free system, he found that the polypeptides made consisted exclusively of the amino acid – Phenylalanine (Phe). Nirenberg concluded that UUU might code for phenylalanine. Using the same approach, he discovered triplet CCC codes for Proline (Pro).
- Following this concept, it was possible to decipher the entire genetic code.
- For their contributions Nirenberg and others received the Nobel Prize in 1968.

# WHAT IS NON-CODING DNA?

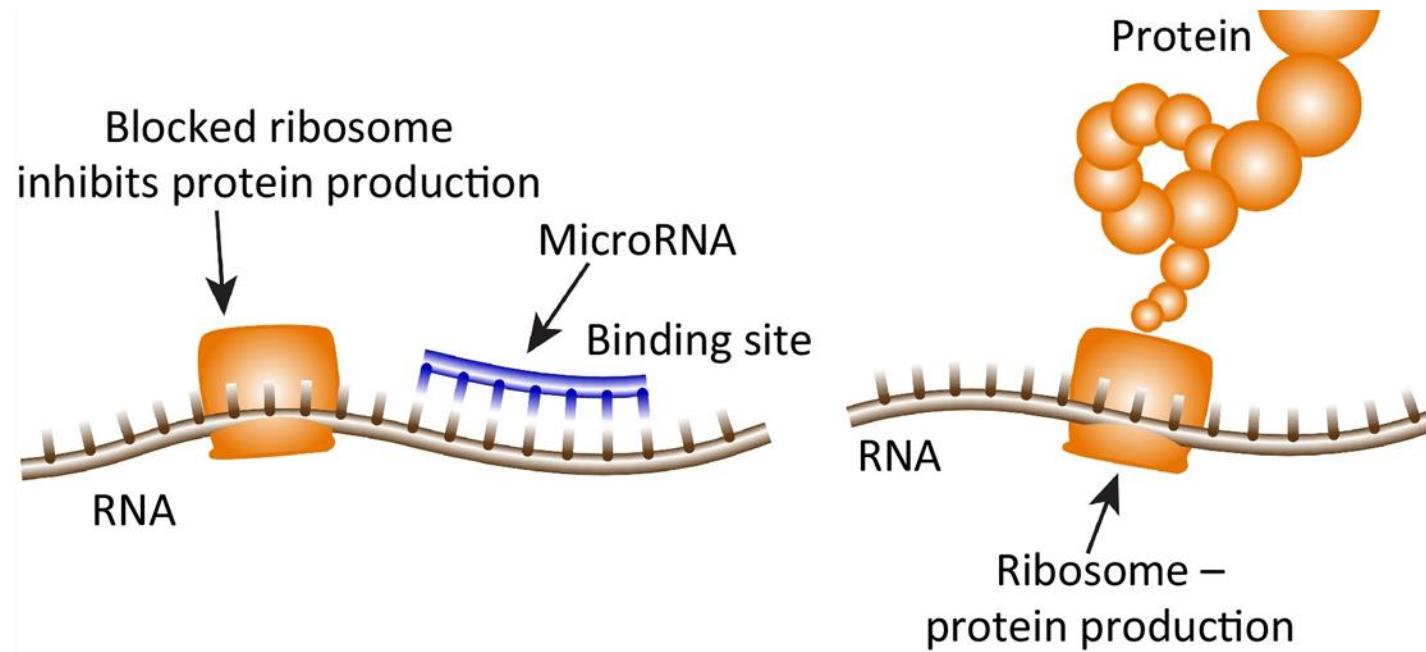
- Only about 1 percent of DNA is made up of protein-coding genes; the other 99 percent is noncoding.
- Noncoding DNA does not provide instructions for making proteins.
- Scientists once thought noncoding DNA was “junk,” with no known purpose. However, it is becoming clear that at least some of it is **integral to the function of cells, particularly the control of gene activity**.
- For example, noncoding DNA contains sequences that act as regulatory elements, determining when and where genes are turned on and off. Such elements provide sites for specialized proteins (called transcription factors) to attach (bind) and either activate or repress the process by which the information from genes is turned into proteins (transcription).

# NON-CODING DNA

Noncoding DNA contains many types of regulatory elements:

- Promoters provide binding sites for the protein machinery that carries out transcription. Promoters are typically found just ahead of the gene on the DNA strand.
- Enhancers provide binding sites for proteins that help activate transcription. Enhancers can be found on the DNA strand before or after the gene they control, sometimes far away.
- Silencers provide binding sites for proteins that repress transcription. Like enhancers, silencers can be found anywhere on the DNA strand.
- Insulators provide binding sites for proteins that control transcription in a number of ways. Some prevent enhancers from aiding in transcription (enhancer-blocker insulators). Others prevent structural changes in the DNA that repress gene activity (barrier insulators). Some insulators can function as both a blocker and barrier.

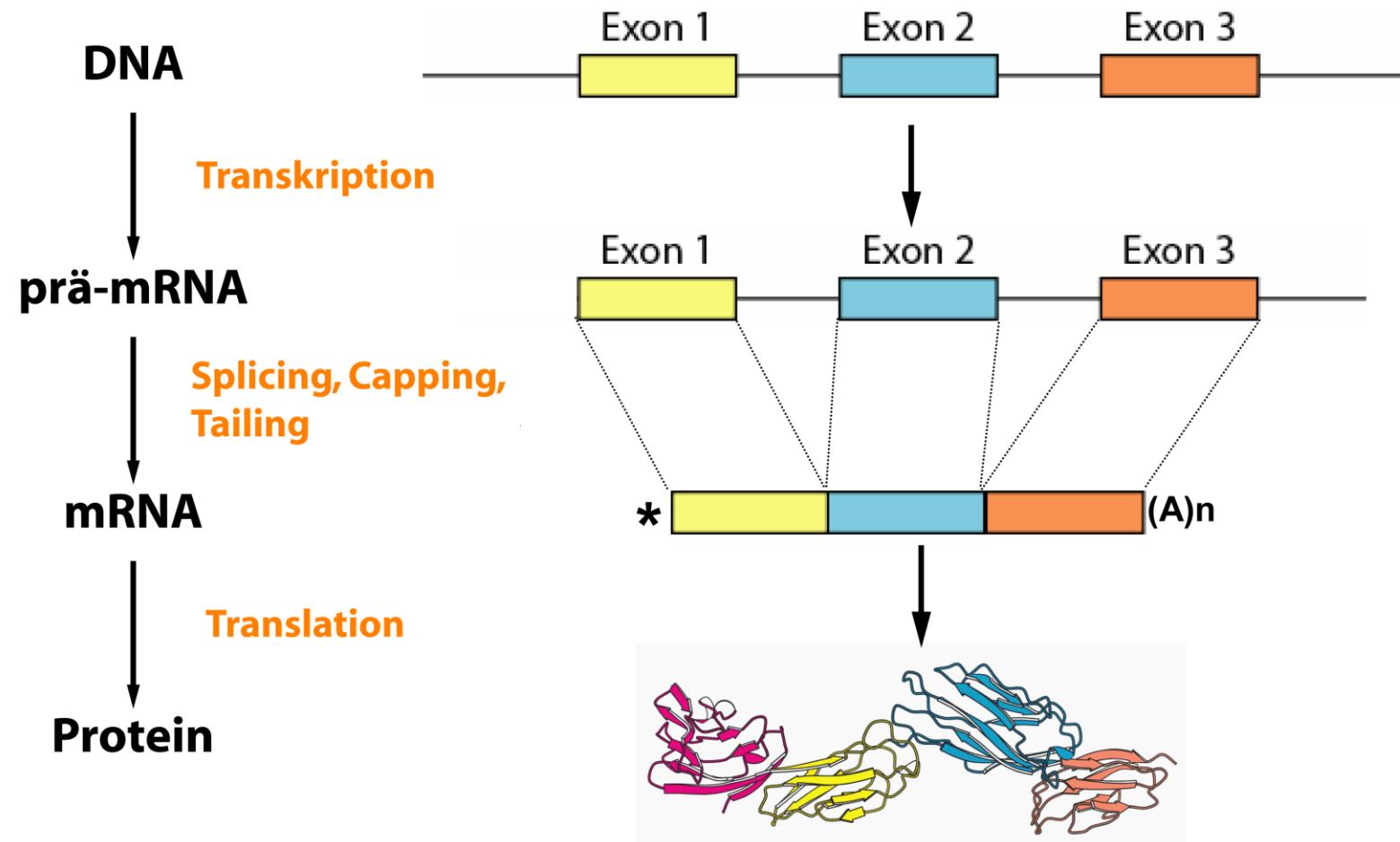
# MICRO RNA



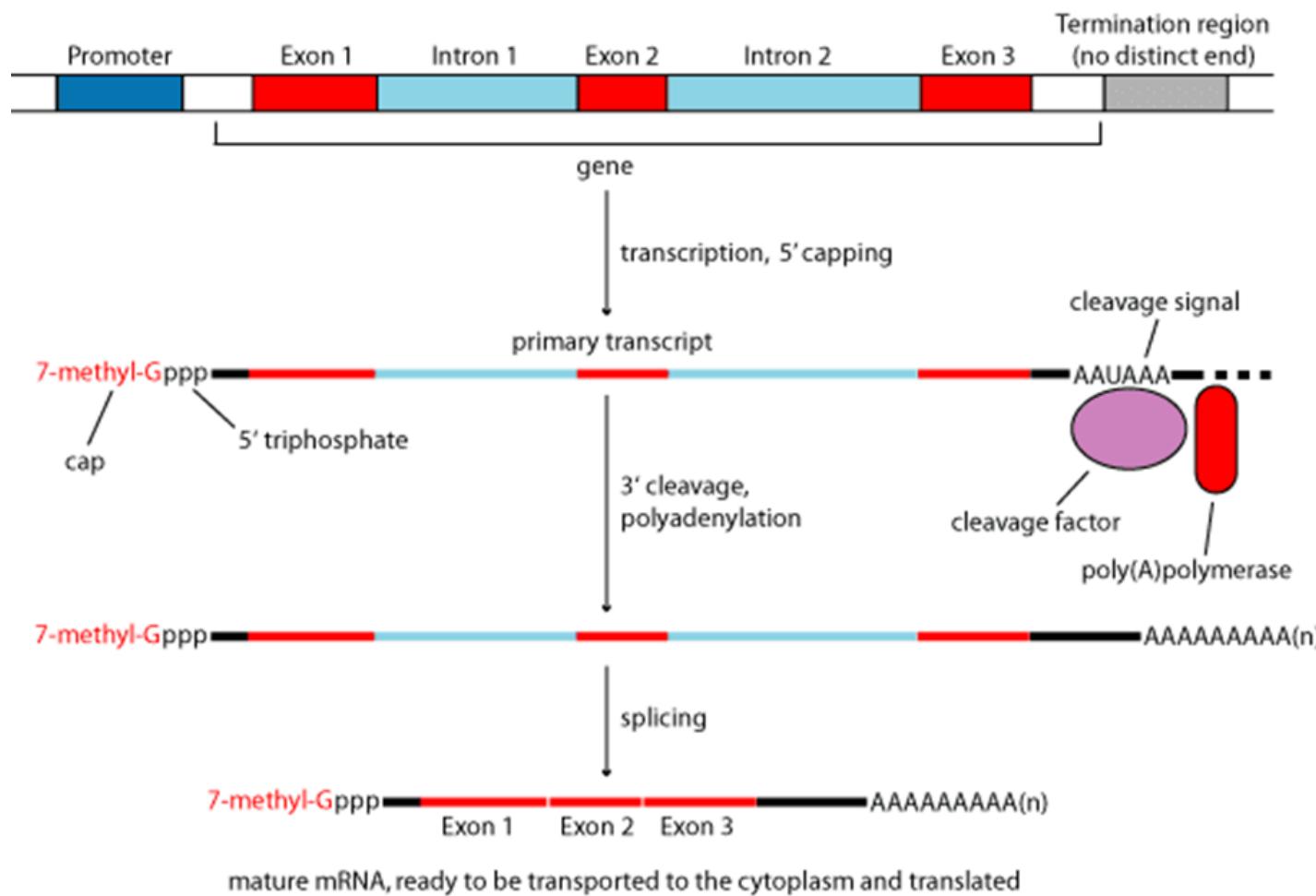
# NON-CODING RNA

- Other regions of noncoding DNA provide **instructions for the formation of certain kinds of RNA molecules**. Examples of specialized RNA molecules, e. g.
  - transfer RNAs (tRNAs) and
  - ribosomal RNAs (rRNAs),
  - microRNAs (miRNAs), which are short lengths of RNA that block the process of protein production; and
  - long noncoding RNAs (lncRNAs), which are longer lengths of RNA that have diverse roles in regulating gene activity.
- Some **structural elements of chromosomes** are also part of noncoding DNA. For example, repeated noncoding DNA sequences at the ends of chromosomes form telomeres. **Telomeres** protect the ends of chromosomes from being degraded during the copying of genetic material. Repetitive noncoding DNA sequences also form satellite DNA, which is the basis of the **centromere**, which is the constriction point of the X-shaped chromosome pair. Satellite DNA also forms heterochromatin, which is densely packed DNA that is important for controlling gene activity and maintaining the structure of chromosomes.
- Some noncoding DNA regions, called **introns**, are located within protein-coding genes but **are removed before a protein is made**. Regulatory elements, such as enhancers, can be located in introns.
- The identity of regulatory elements and other functional regions in noncoding DNA is not completely understood. Researchers are working to understand the location and role of these genetic components.

Sequences not needed to make a protein are called introns; the sequences that are expressed are called exons. The introns are cut out by various enzymes and the exons are spliced together to form a complete RNA molecule.

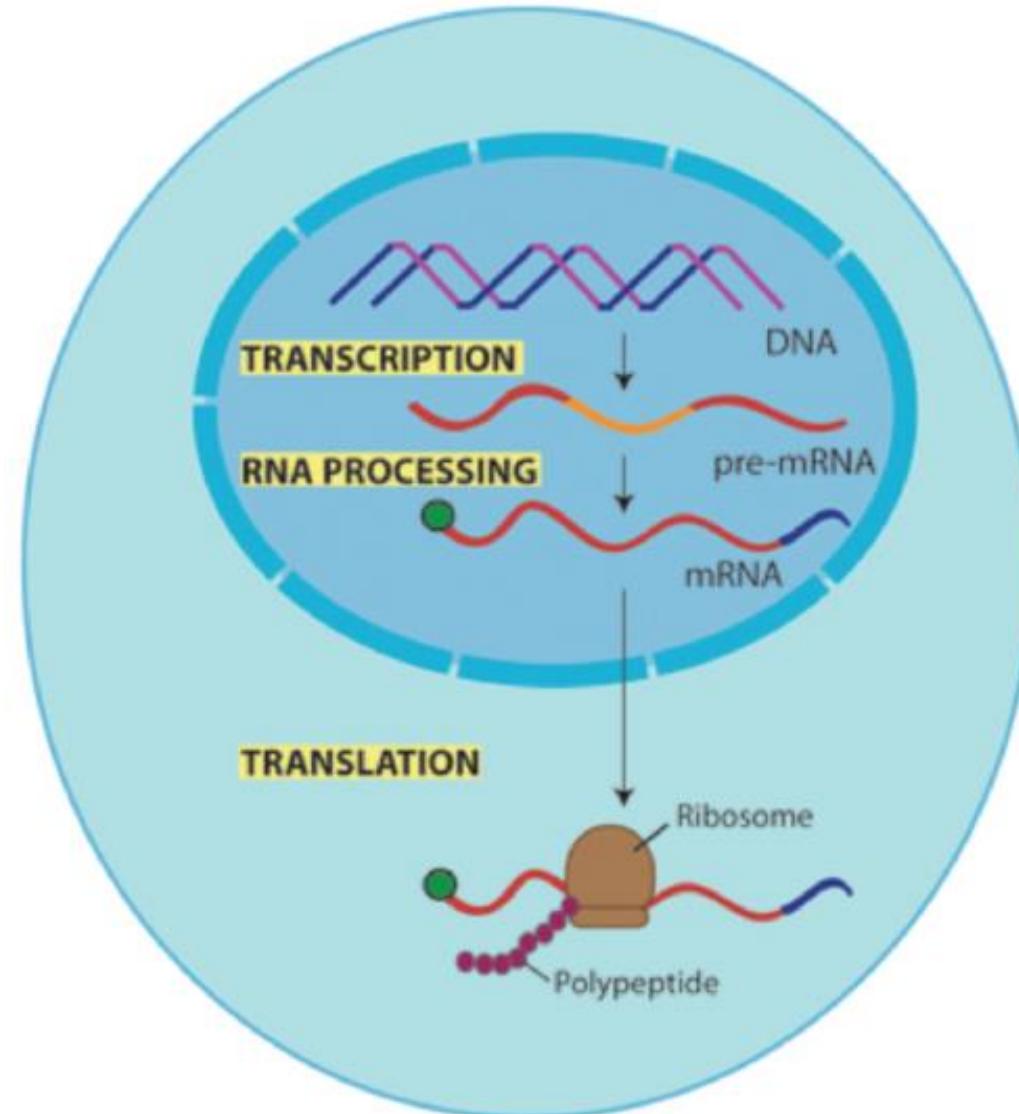


# CAPPING-TAILING-SPLICING

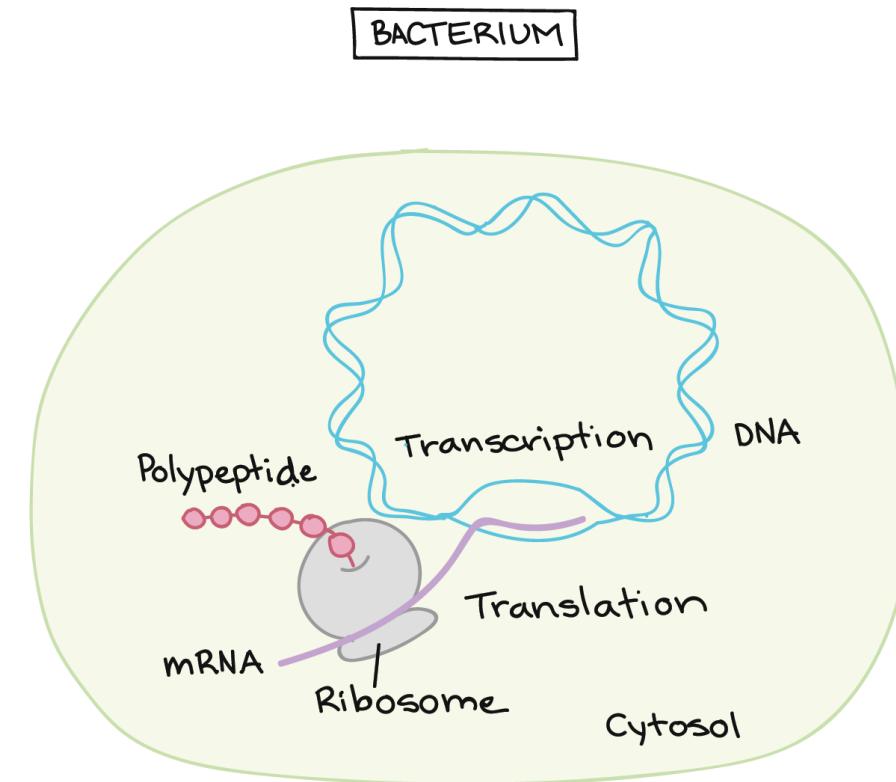
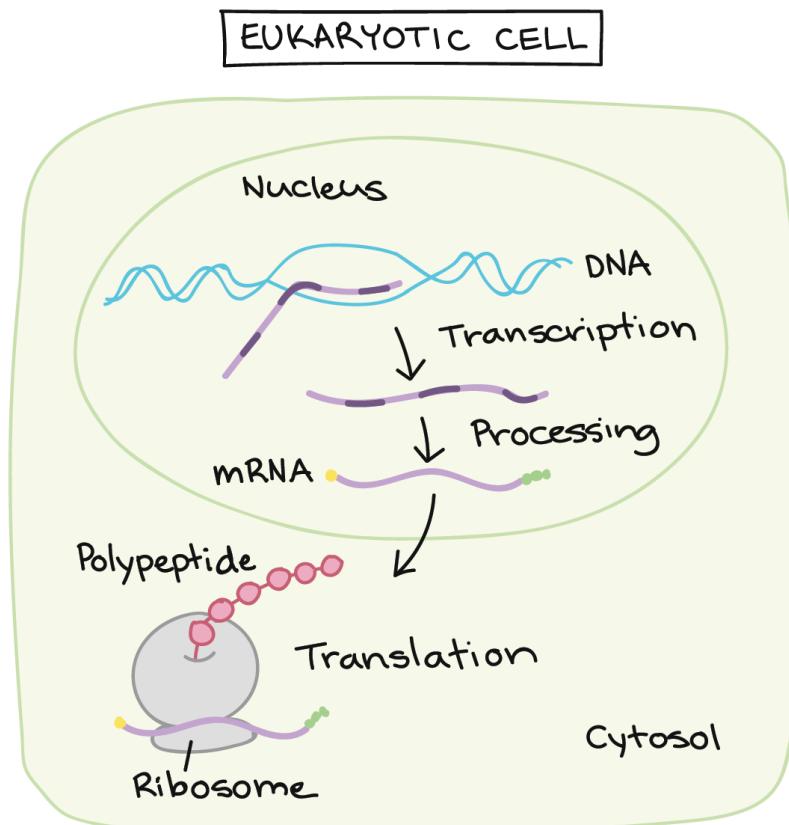


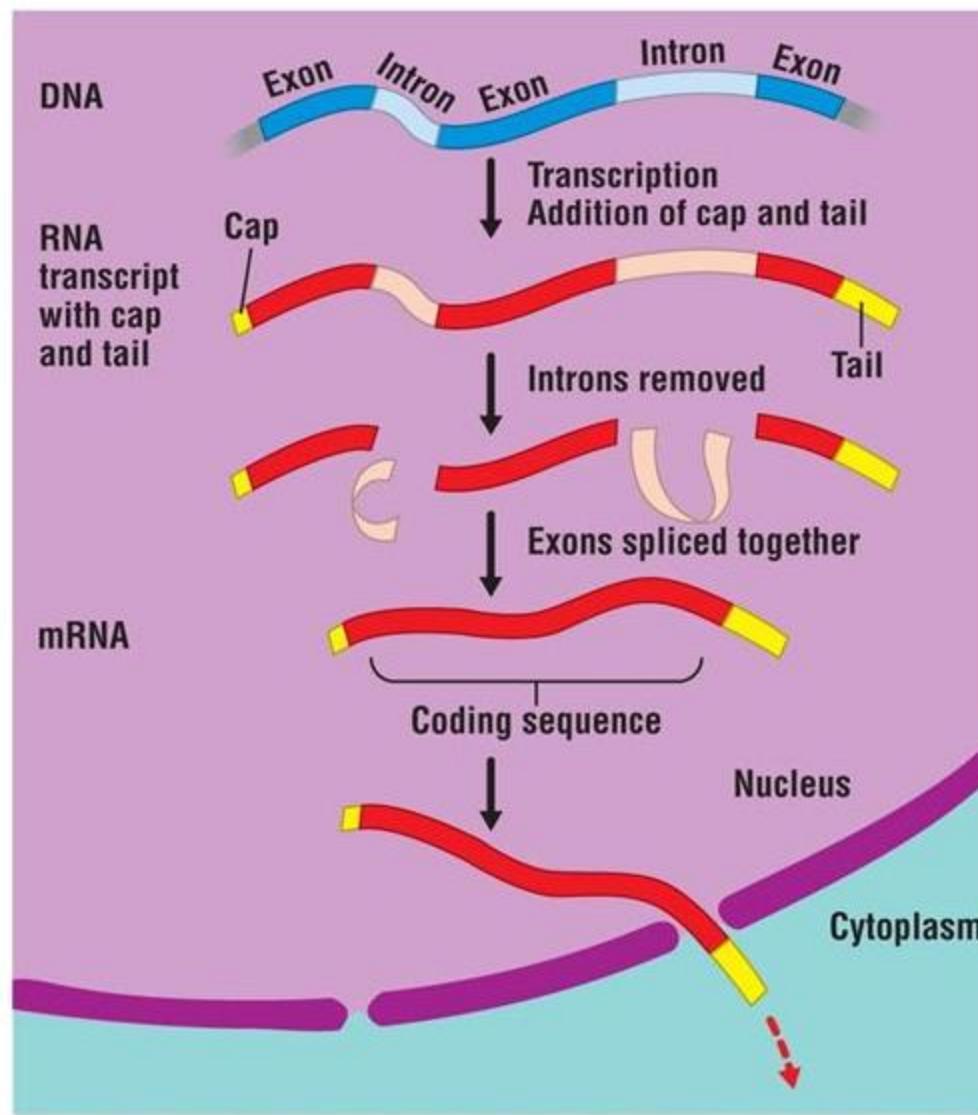
# RNA PROCESSING

Takes place in eucaryotic cells before the mRNA leaves the nucleus.



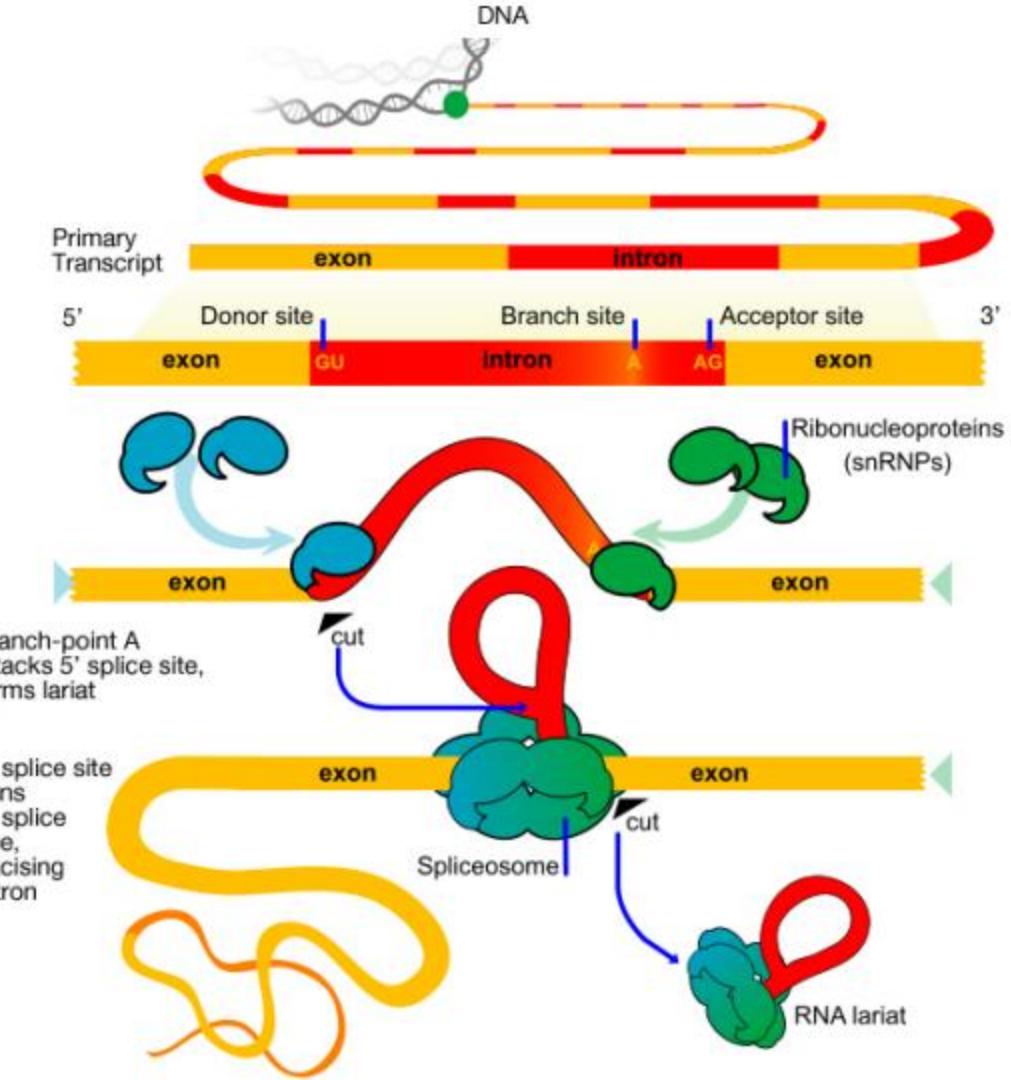
In bacterial cells, the mRNA is translated directly as it comes off the DNA template. In eukaryotic cells, RNA synthesis, which occurs in the nucleus, is separated from the protein synthesis machinery, which is in the cytoplasm. In addition, eukaryotic genes have noncoding regions.





# RNA PROCESSING

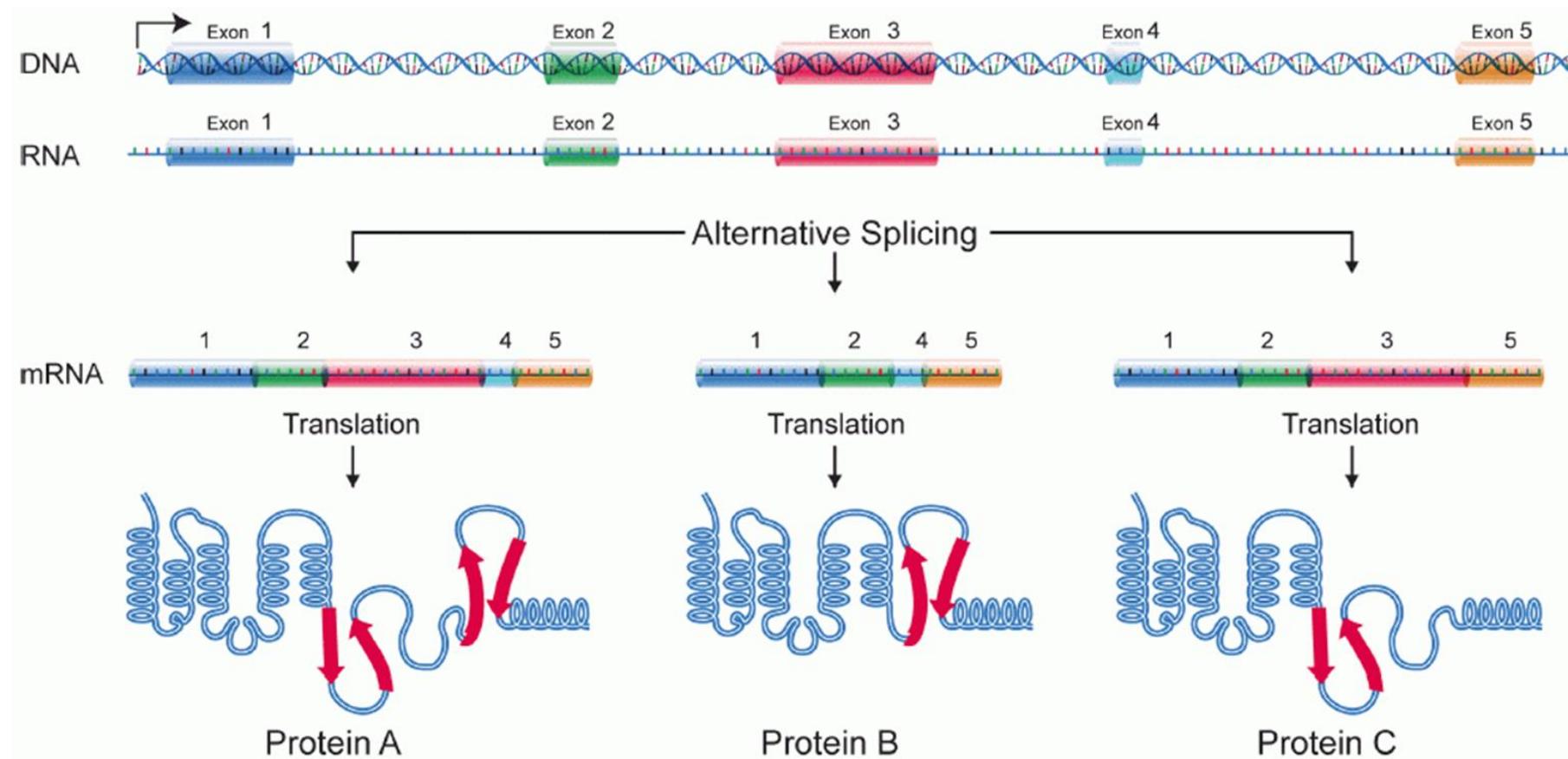
- Once a mature mRNA transcript is made it is transported to the cytoplasm for translation into protein.
- Most eukaryotic genes and their pre-mRNA transcripts contain noncoding stretches of nucleotides or regions that are not meant to be made into protein. These noncoding segments are called introns and must be removed before the mature mRNA can be transported to the cytoplasm and translated into protein. The stretches of DNA that do code for amino acids in the protein are called exons.
- During the process of splicing, introns are removed from the pre-mRNA by the spliceosome and exons are spliced back together. If the introns are not removed, the RNA would be translated into a nonfunctional protein. Splicing occurs in the nucleus before the RNA migrates to the cytoplasm.
- Once splicing is complete, the mature mRNA (containing uninterrupted coding information), is transported to the cytoplasm where ribosomes translate the mRNA into protein.



# SPLICING

- The RNA molecule contains sections that are not needed as part of the protein code that need to be removed.
- A number of genes can give rise to several different proteins, depending on which sections are treated as exons - this is known as alternative RNA splicing. This allows a relatively small number of genes to create a much larger number of different proteins. Humans have just under twice as many genes as a fruit fly, and yet can make many times more protein products.

# ALTERNATIVE SPlicing

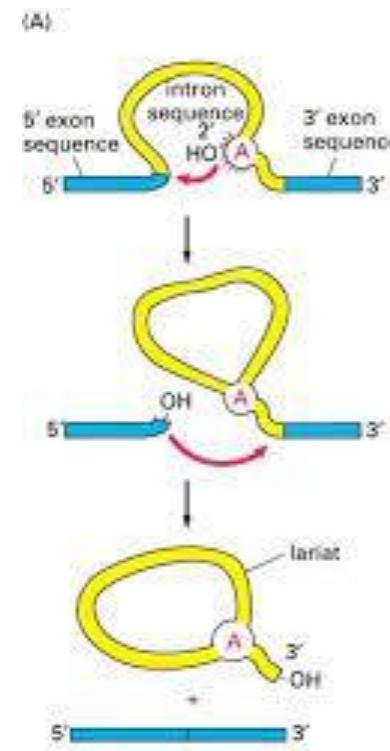


# CAN GENES BE TURNED ON AND OFF IN CELLS?

- Each cell expresses, or turns on, only a fraction of its genes at any given time. The rest of the genes are repressed or turned off. The process of turning genes on and off is known as gene regulation. Gene regulation is an important part of normal development.
- Genes are turned on and off in different patterns during development to make a brain cell look and act different from a liver cell or a muscle cell, for example. Gene regulation also allows cells to react quickly to changes in their environments. Although we know that the regulation of genes is critical for life, this complex process is **not yet fully understood**.
- Gene regulation can occur at any point during gene expression, but most commonly occurs at the level of transcription (when the information in a gene's DNA is passed to mRNA). Signals from the environment or from other cells activate proteins called transcription factors. These proteins bind to regulatory regions of a gene and increase or decrease the level of transcription. By controlling the level of transcription, this process can determine when and how much protein product is made by a gene.

# INTRONS AND EXONS

- Eukaryotic genes were long thought to be like those of prokaryote continuous sequences.
- This paradigm, changed dramatically when it was discovered that (at least some) eukaryotic protein coding genes were interrupted by non-coding sequences and eliminated from the mature (translated) mRNA before translation.
- the concept of exons and introns was coined



## SPLICING

- There are no doubts that this discovery was a revolution in genetics. It not only challenged the previous definition of what a “gene” is, but led to discoveries and concepts such as splicing, alternative splicing, or to the great discovery that some RNAs, once transcribed, can eliminate introns by themselves, a mechanism known as **autosplicing**.
- The discovery of autosplicing not only reinforced the idea of the “RNA world” but eliminated, forever, the time unanimous idea that proteins were the only catalytic molecules.

# GLOSSARY

Codon - a sequence of three organic bases in a nucleic acid that code for a specific amino acid

Exon - Coding region of eukaryotic gene. Parts of the gene that are expressed

Gene - a length of DNA made up of a number of codons; codes for a specific protein

Intron - Non coding region of a gene that separates exons

Polypeptide - a chain of amino acids joined by a peptide bond

Ribosome - a cellular organelle that functions as a protein-making workbench.

RNA - Ribonucleic Acid; a nucleic acid that acts as a messenger, carrying information from the DNA to the Ribosomes

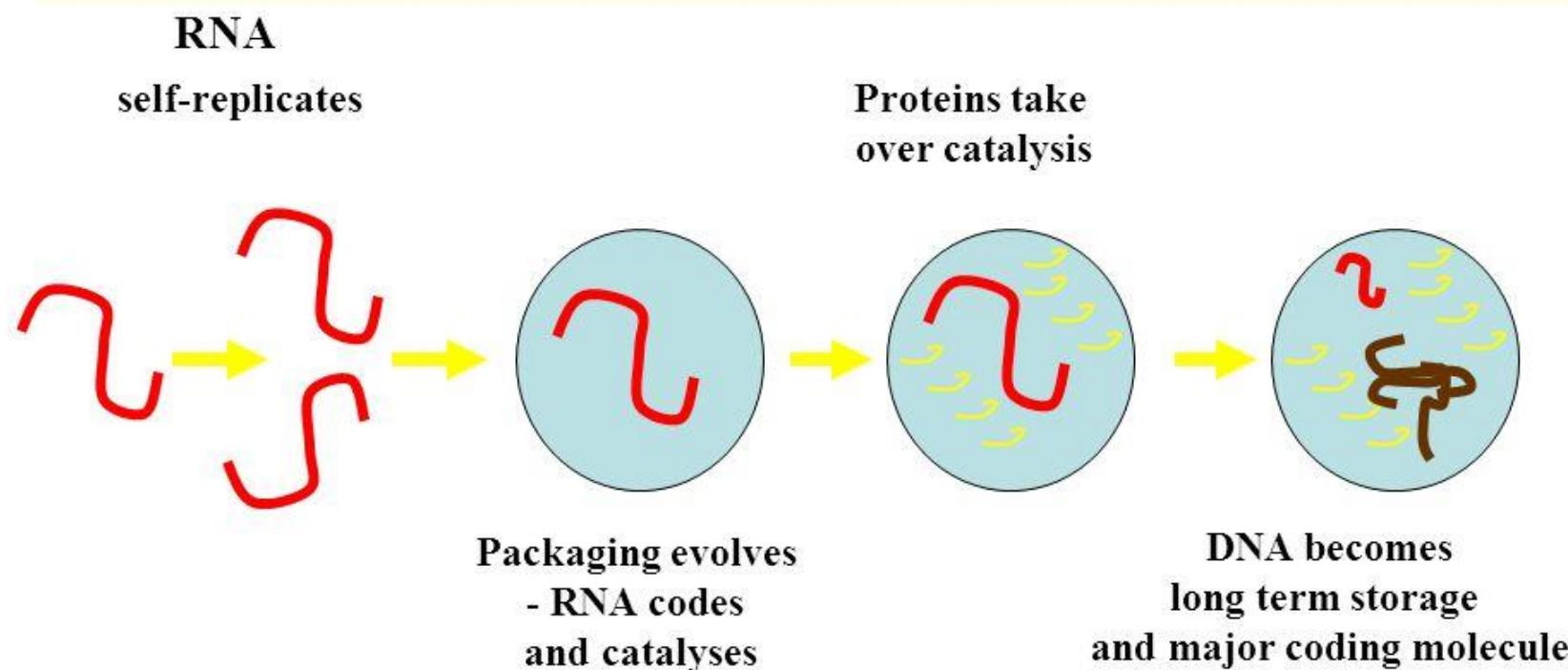
Gene expression (Protein biosynthesis) - the process of creating protein molecules in biological systems, it involves transcription and translation.

- Additional material: the RNA - hypothesis

## The RNA World

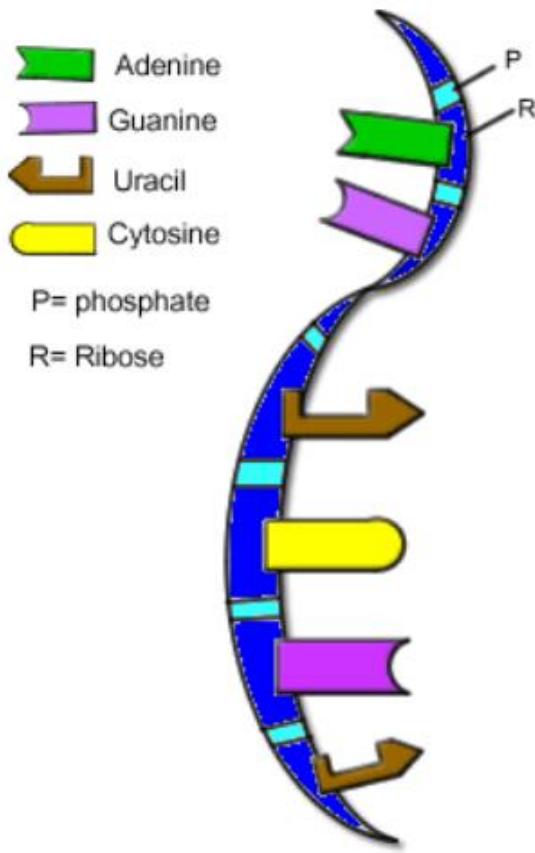
The current “most accepted” theory of life evolving hypothesizes an RNA world

RNA in the early world would have functioned as a self replicating molecule, eventually developing a number of minimal catalytic properties

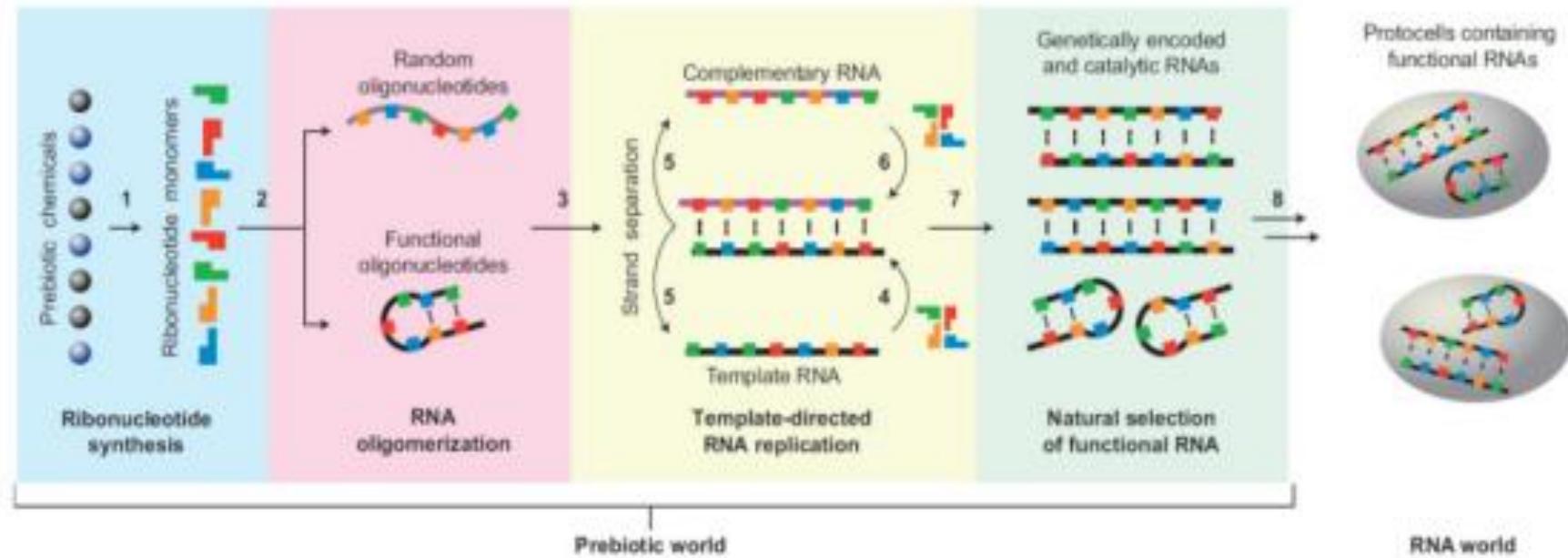


# RNA WORLD HYPOTHESIS

- The **RNA world hypothesis** proposes that self-replicating ribonucleic acid (RNA) molecules were precursors to current life.
- RNA stores genetic information like DNA, and catalyzes chemical reactions like an enzyme protein.
- Many viruses also store and transmit RNA



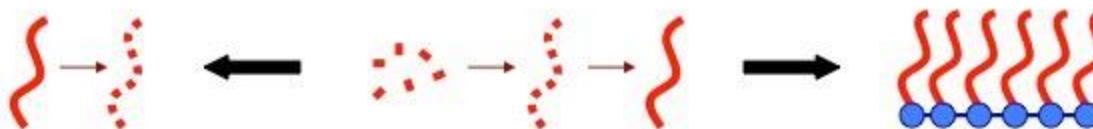
## *The RNA world*



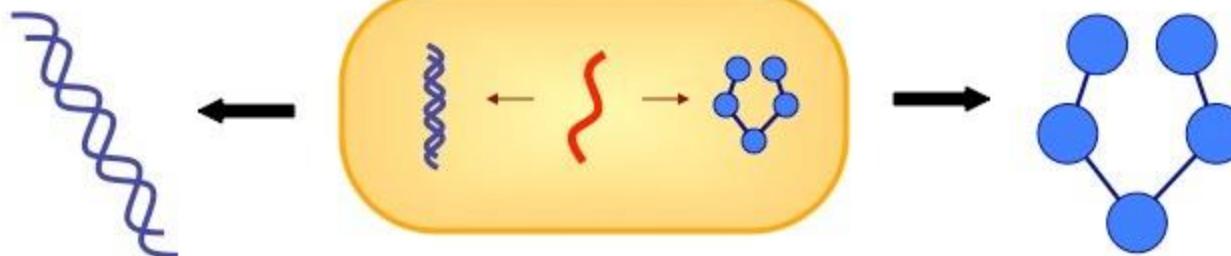
**Step 2:**  
RNA self-replicates (via  
ribozymes)

**Step 1:**  
RNA forms from  
inorganic sources

**Step 3:**  
RNA catalyses protein  
synthesis



**Step 4:**  
Membrane formation  
changes internal chemistry,  
allowing new functionality



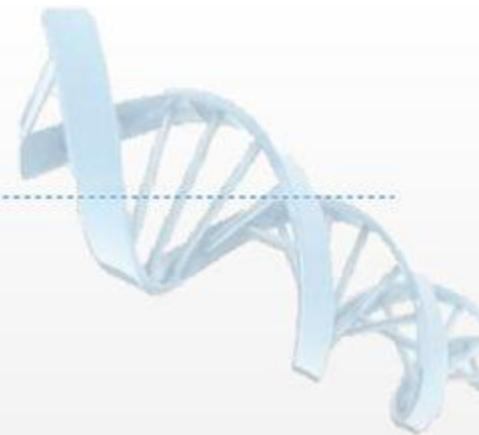
**DNA** becomes master  
template

**Step 5:**  
RNA codes both DNA  
and protein

**Proteins** catalyse  
cellular activities

## The RNA World

- ▶ The **RNA world hypothesis** was first proposed as a stage in evolution. The hypothesis describes a living system (or set of living systems) based on RNA.
- ▶ In this system, a variety of RNA enzymes could catalyze all of the reactions needed to synthesize the molecules required for life from simpler molecules available in the environment.
- ▶ The “RNA organism,” out of equilibrium with its surroundings, would have to be defined by a boundary.



• • • • • • •  
RNA is the only currently used macromolecule that is both a carrier of genetic information and an enzyme.

# RNA WORLD HYPOTHESIS

- The ocean water or primordial soup of primitive Earth (about 4 billion years ago) gave rise to a number of complex organic molecules from simple inorganic ones – one of these happened to be RNA.
- RNA molecules are capable of both storing information and performing metabolic activities. In present day cells DNA stores information and proteins perform catalysis, with RNA as the intermediate between DNA and protein. One can imagine a time when there was no DNA or protein, just RNA performing both functions: this is the RNA World hypothesis.

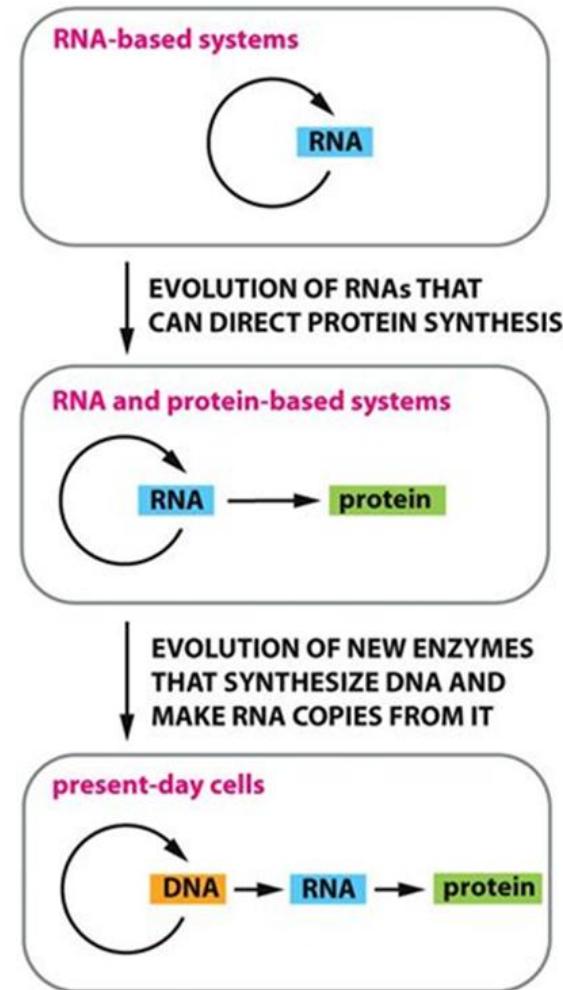


Figure 7-46 Essential Cell Biology 3/e (© Garland Science 2010)

# RNA – THE MOLECULE OF LIFE

- Ribose has one additional oxygen atom compared to Deoxyribose which makes RNA rather instabel compared to DNA. But Ribose is also more active and RNA can fold in various forms (many more than DNA).
- RNA can also build double helices (a bit wider than DNA) and it can also be folded in many different other structures, which makes this molecule an ideal molculre for the origin of life in the course of evolution.
- This leads to the RNA world hypothesis postulating that RNA was the first molecul which was metabolically active and could store inforamtion at the same time.
- RNA has the ability to fold, the inforamtion in RNA is stored exactly the same way as it is stored in DANN, i. e. in the sequence of nucleotide in a long chain. This sequence also includes the information for folding in a two and three dimensional way. The folded structures makes it possibel to control and steer chemical processes.
- In the search for the origin of life there has been speculation which molecule was there first: proteins or DNA. Proteins are necessary to control metabolism and to build up proteins, DNA contains the information to form proteins – a classic chicken egg dilemma. What was first? Protein or DNA? Neither nor. It was RNA, because it is chicken and egg in one. In its primary structure it can store information (as can DNA) and due to its ability to fold and build threedimensional structures it catylyses reactions (as proteins do).
- Information and metabolism both in one molecule – the molecule of life – the RNA

# THE CHICKEN – EGG – DILEMMA

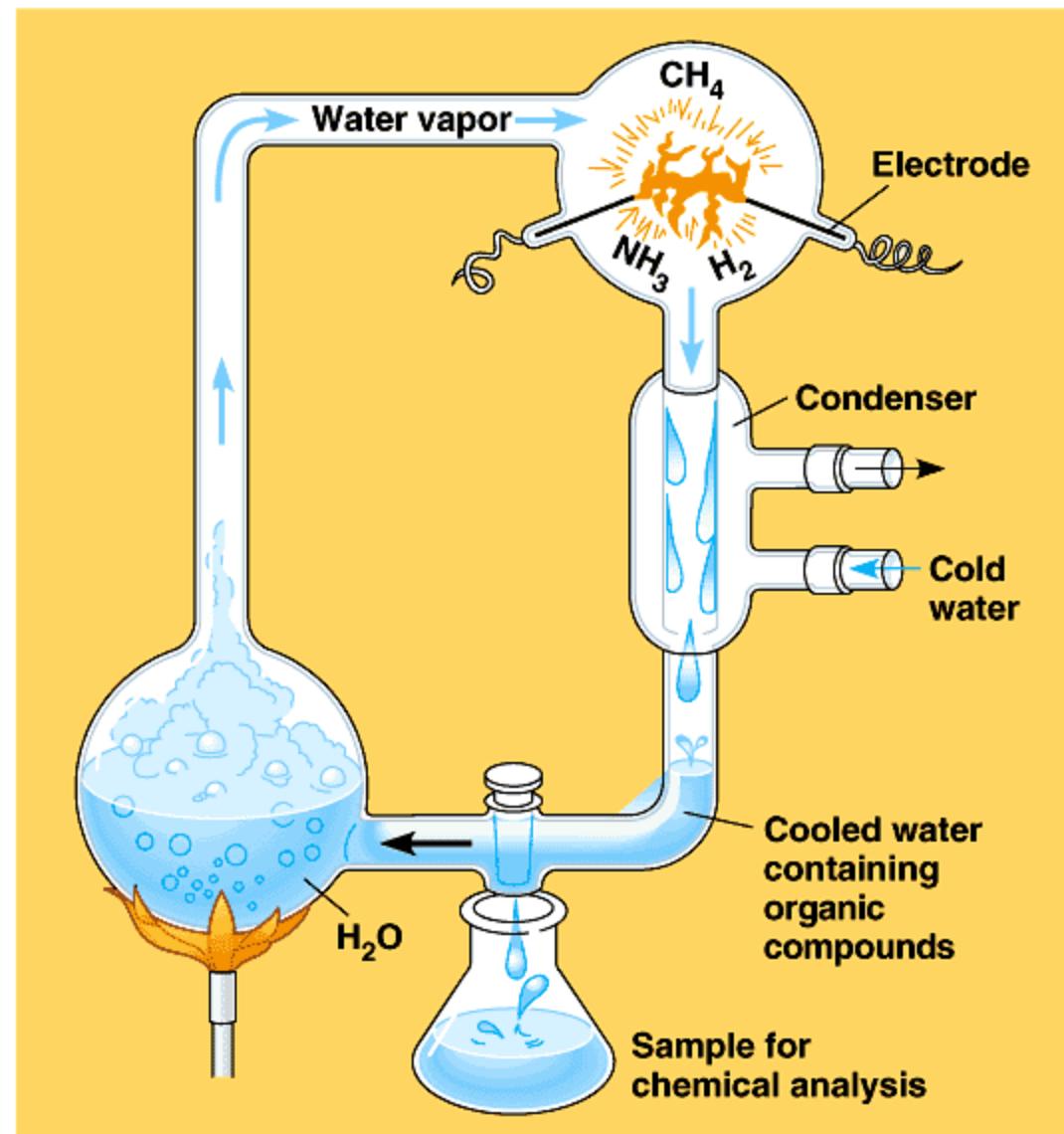
RNA autocatalyse enables RNA both to:

- store information and
- catalyse chemical reactions



## MILLER-UREY EXPERIMENT

- was the first attempt to scientifically explore ideas about the origin of life.
- Stanley Miller simulated conditions thought to be common on the ancient Earth. The purpose was to test the idea that the complex molecules of life (RNA, amino acids) could have arisen on our planet through simple, natural chemical reactions.
- The experiment was a success in that amino acids, the building blocks of life, were produced during the simulation.



Copyright © Pearson Education, Inc., publishing as Benjamin Cummings.

## MILLER-UREY EXPERIMENT

Many questions about the origin of life remain to be answered but these findings give strong support to the idea that the first living cells on Earth may have emerged from natural chemical reactions.

