Molecular Biology

MLT 344

Ilham Qattan
A. Professor of Medical Molecular Virology (MMV)
Deputy of the Medical Laboratories Technology Department
ESJ editorial board member
Board member of the Canter of genes and genetic diseases
Tel (W): +966(0) 148618888 Ext: 3609
Mob +966(0)556472725
@dr_itq
Email: iqattan@taibahu.edu.sa
Web: elitedoctorsonline.com/dr630

COURSE OUTLINE

1. Introduction to Molecular Biology
2. Gene and genomes
3. Biomolecule structures and functions
4. Replication
5. Transcription
6. Translation
7. Mutation
8. Recombinant DNA technology
9. PCR
10. DNA sequencing
11. Protein analysis methods

INTRODUCION
TO
MOLECULAR BIOLOGY

Introduction

Molecular biology (MB) is a combination of the ideas from genetics and biochemistry. Molecular biology is the study of molecular processes of replication, transcription, translation, and cell function. The central dogma of molecular biology where genetic material is transcribed into RNA and then translated into protein.

Since the late 1950s and early 1960s, molecular biologists have learned to characterize, isolate, and manipulate the molecular components of cells and organisms. These components include DNA, RNA, and proteins, the major structural and enzymatic type of molecule in cells.

Biochemistry and Genetics

- *Biochemistry* is the study of molecules (e.g. proteins). Biochemists take an organism or cell and dissect it into its molecular components, such as enzymes, lipids and DNA, and reconstitute them in test tubes (*in vitro*).

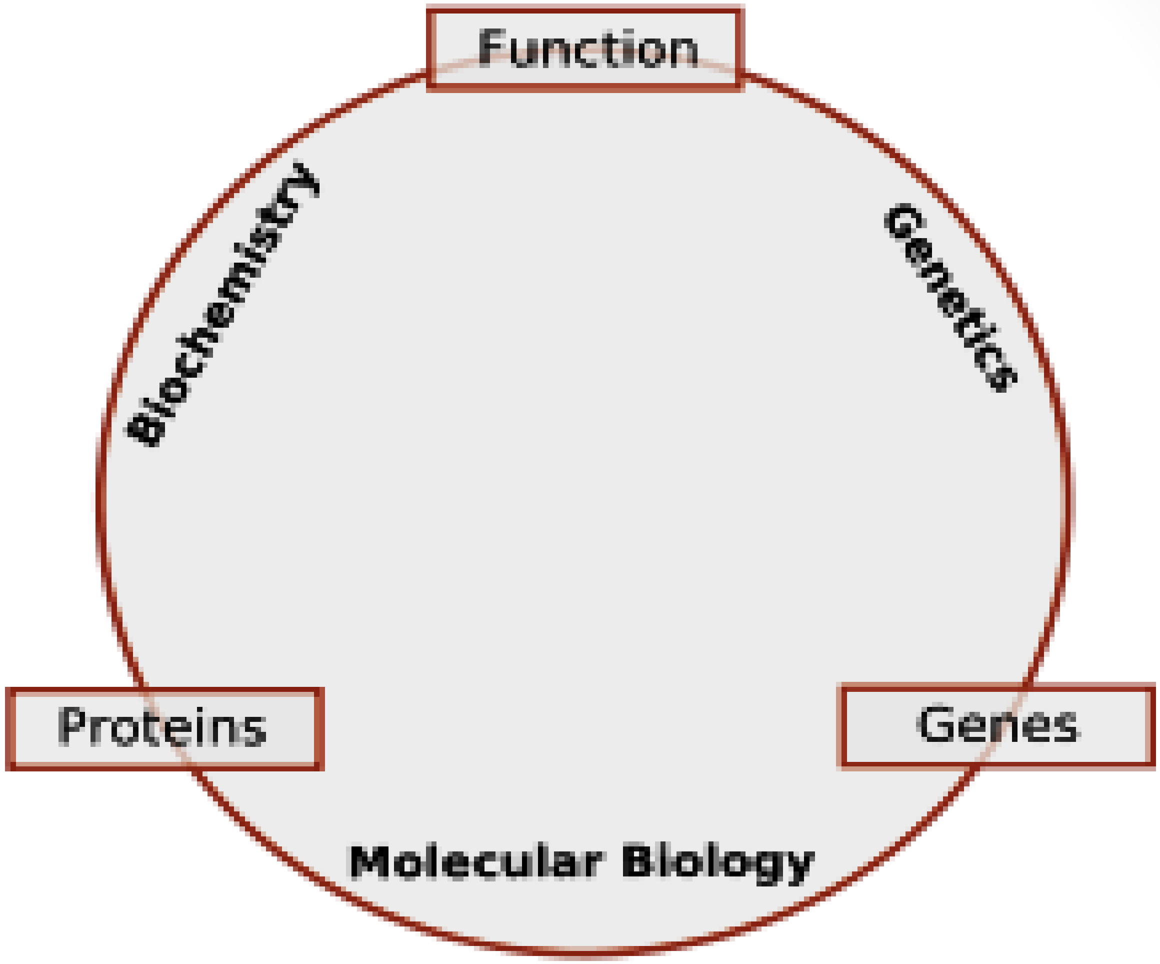
- *Genetics* is the study of the effect of genetic differences on organisms. Often this can be inferred by the absence of a normal component (e.g. one gene).

MOLECULAR BIOLOGY

- Branch of biology investigating **molecular** basis of biological function and interactions between different **Biomolecules:**
    1. DNA
    2. RNA
    3. Protein
  - Study Cell function and molecular processes:
    - Replication (DNA synthesis)
    - Transcription (mRNA transcript)
    - Translation (protein synthesis)
  - Stemmed from Genetics and Biochemistry.
  - Study of gene structure and function at a molecular
    Level.

BIOMOLECULES

- **Nucleic acid**
  - DNA
  - RNA
    - Nucleotides
- **Protein**
  - Polypeptide
    - Amino acids



Schematic relationship between biochemistry, genetics and molecular biology

MOLECULAR BIOLOGY APPLICATIONS:

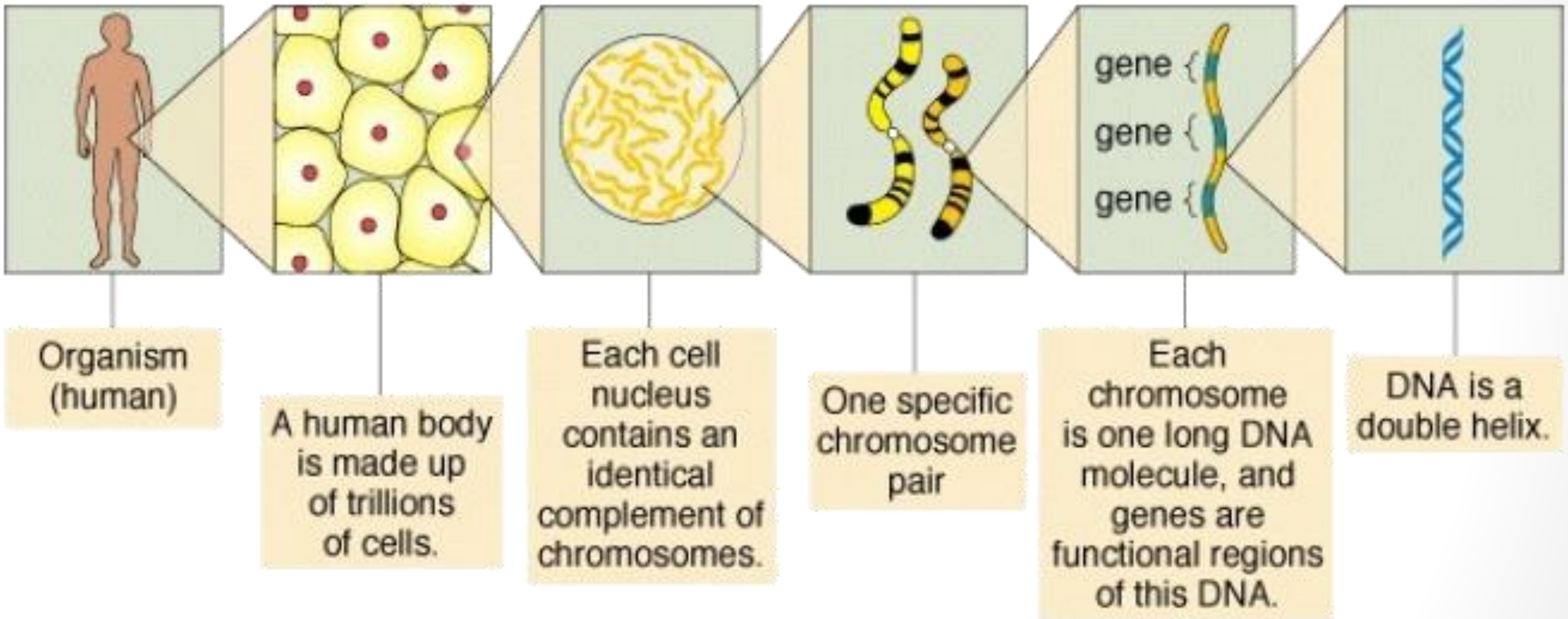
1) Molecular cloning
2) DNA Sequencing
3) Genotyping
4) Gene therapy
5) Screening for genetic diseases
6) Gene function
7) Gene knock out (disease Models)

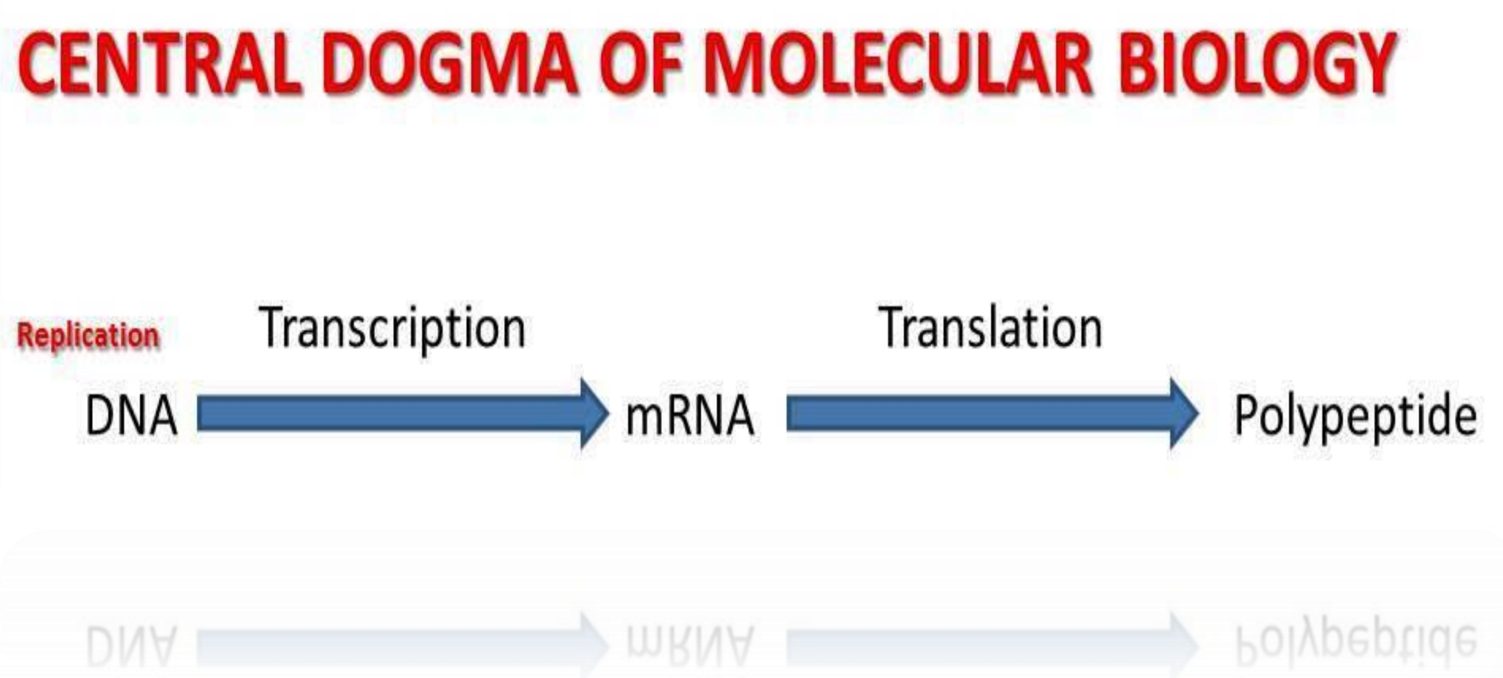
MOLECULAR STRUCTURE OF DNA & RNA

NUCLEIC ACID:

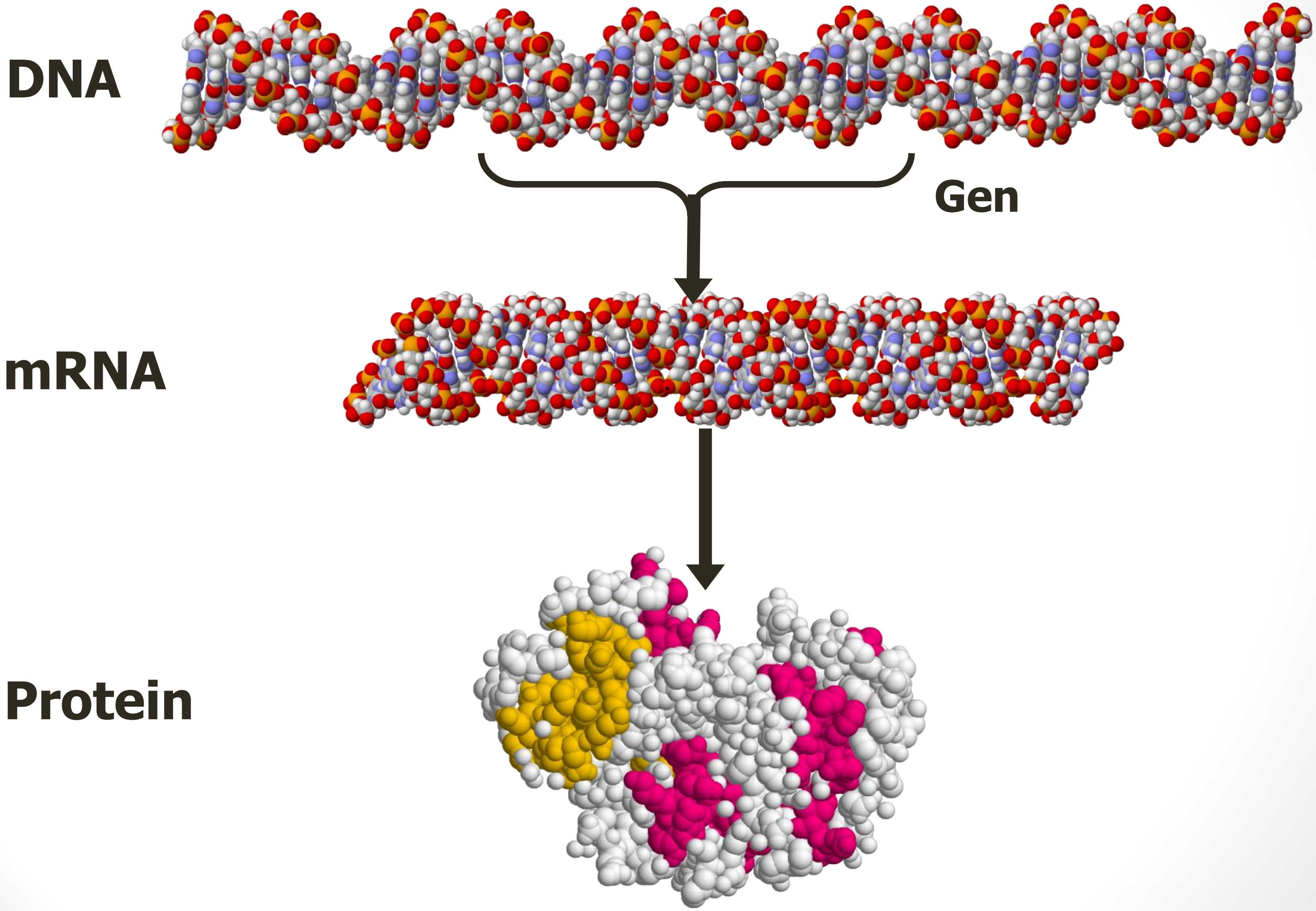
- **Molecule of life**
- Storage and transmission of genetic information
- **Two types**
  1) Deoxyribonucleic acid (**DNA**)
  2) Ribonucleic acid (**RNA**)
- Consists of long polymer of repeating units called **nucleotides**

Molecular Biology




CENTRAL DOGMA OF MOLECULAR BIOLOGY
Replication
Transcription
Translation
DNA
mRNA
Polypeptide

4m Genes 2 Proteins



The structure of DNA and RNA.

- DNA (deoxyribonucleic acid) and RNA (ribonucleic acid) are composed of ;
- Two different classes of nitrogen containing bases: the **purines** and **pyrimidines**.
- The most commonly occurring **purines** in DNA are **adenine (A)** & **guanine (G)**.
- The most commonly occurring **pyrimidines** in DNA are **cytosine (C)** & **thymine (T).**
- RNA contains the same bases as DNA with the exception of thymine. Instead, RNA contains the **pyrimidine** **uracil (U).**

- Adenine, guanine, cytosine, thymine and uracil are usually abbreviated using the single letter codes A, G, C, T and U, respectively.

- The resulting molecules are called **nucleosides** and can serve as elementary precursors for DNA and RNA synthesis, *in vivo*.

- Before a nucleoside can become part of a DNA or RNA molecule it must become complexed with a phosphate group to form a **nucleotide** (either a deoxyribonucleotide or ribonucleotide).

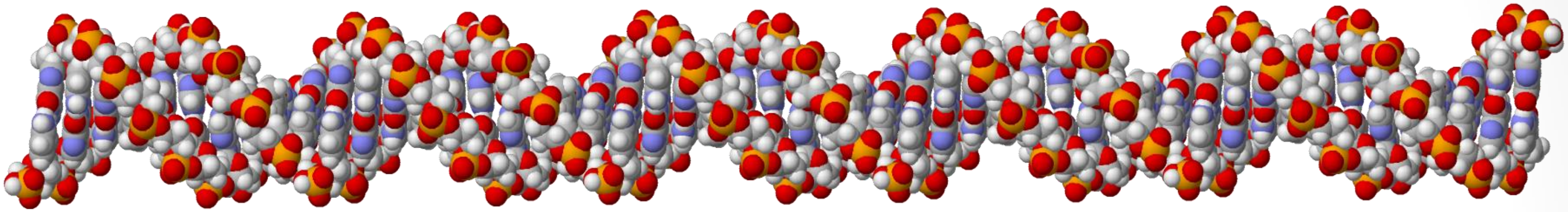
Nucleoside = Nucleobase + Pentose
Nucleotide = Nucleobase + Pentose + Phosphate Group

free base	nucleoside	nucleotide
Adenine (A)	Adenosine	Adenosine monophosphate (AMP)
Guanine (G)	Guanosin	Guanosine monophosphate (GMP
Cytosine (C)	Cytidin	Cytidine monophosphate (CMP)
Thymine (T)	Thymidin	Thymidin monophosphate (TMP)

LENGTH OF NUCLEOTIDES

- **Dinucleotide**
  - 2 nucleotides
- **Trinucleotide**
  - 3 nucleotides
- **Tetra nucleotide**
  - 4 nucleotides
- **Oligonucleotide**
  - ≤ 200 nucleotides
- **Polynucleotide chain (strand)**
  - > 200 nucleotides

DNA - Molecule



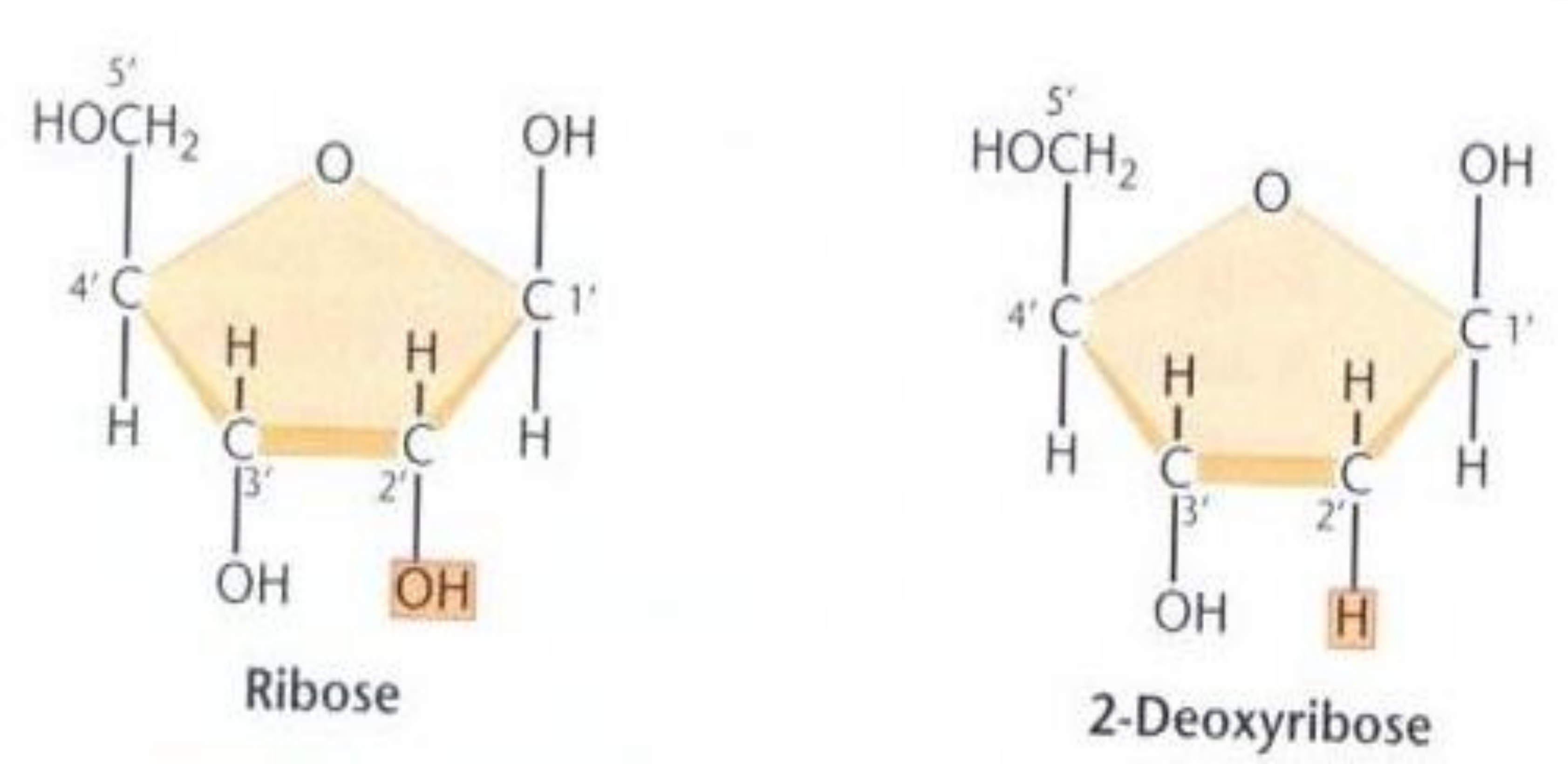
DNA-sequence (Alphabet: ATGC)

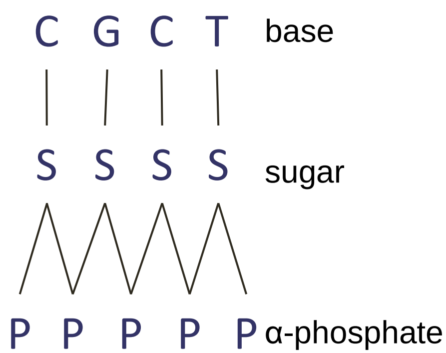
CCTAGACATTGCTTTCCCATCCTGCTACTCAATGACAGTTTCTGGTTTCACTGGG
TCACTCTCATCTTGATGCACTCCCGGGCAAGAGCTAACTGAAAGGCAGCTGCGT
AACACATACCA GACACAACAGTTTATCATGGGAGAGTGAATTAAACCAGGAA...

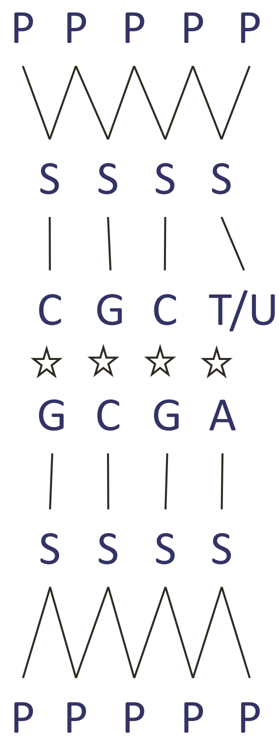
- DNA is short for **Deoxyribonucleic acid**.
- DNA carry most of the **genetic information** required to produce the **three macromolecules** required for life.
- **Eukaryotic** DNA is mainly found in **Chromosomes** inside the cell **nucleus** and is called **Genomic DNA (g-DNA)**, but can also be found in the Mitochondria and is called **Mitochondrial DNA (mt-DNA)**.
- **Prokaryotic** DNA can be found as g-DNA, but also can be freely available in the cytoplasm as **Plasmid DNA**.

PENTOSE SUGAR

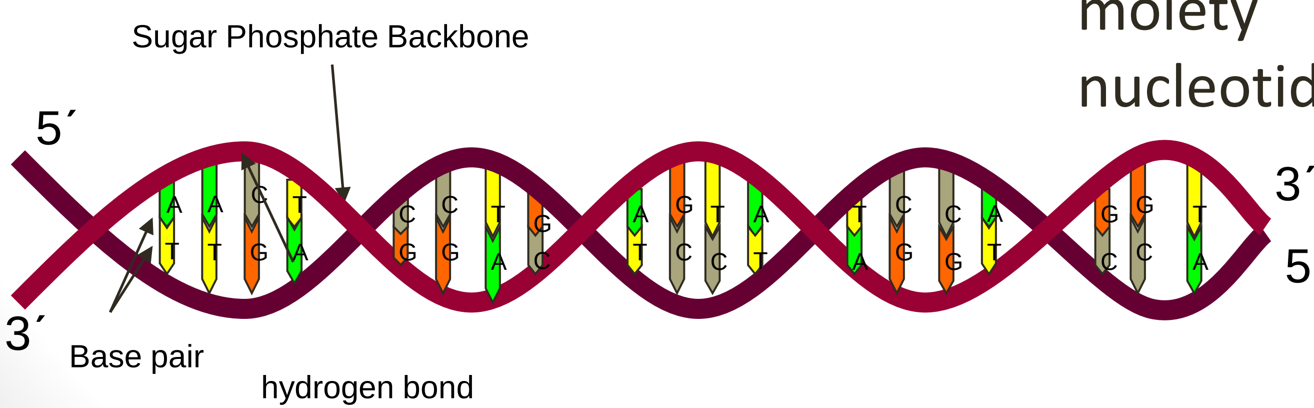
- Carbon found in sugar are numbered 1' to 5'.
- RNA contains ribose (**OH** group at C-2').
- DNA contains deoxyribose (**H** atom at C-2').





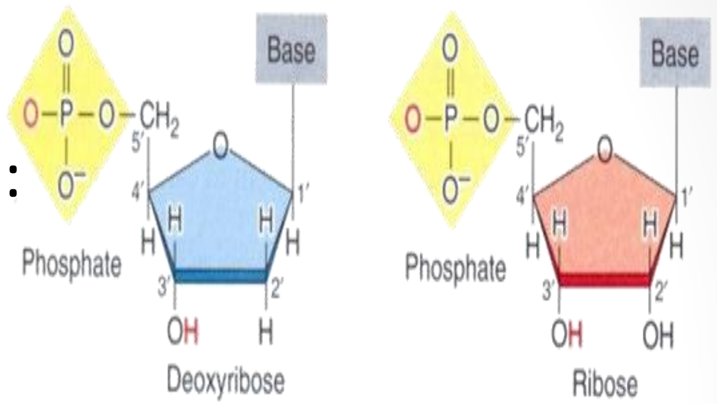


- DNA and RNA are simply long polymers of nucleotides called **nucleotides**.
- Only the **α phosphate** is included in the polymer.
- It becomes chemically bonded to the **3' carbon** of the sugar moiety of another nucleotide.



[read as **3 prime** and **5 prime**]

- Nucleic acid is composed of a long polymer of individual molecules called **nucleotides**.
- Each nucleotide is composed of a:
  1. **Nitrogenous base**
  2. **Sugar molecule**
  3. **Phosphate molecule**.

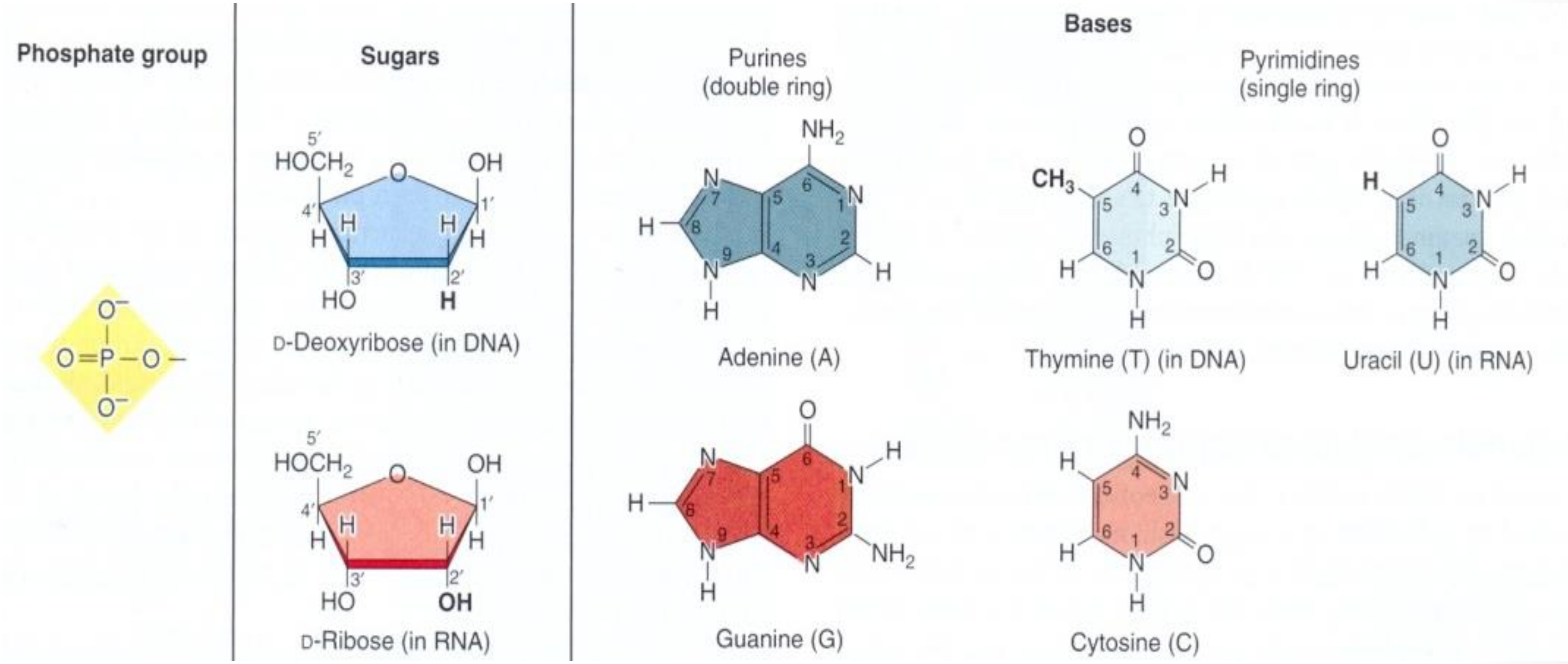


(a) Repeating unit of deoxyribonucleic acid (DNA)

(b) Repeating unit of ribonucleic acid (RNA)

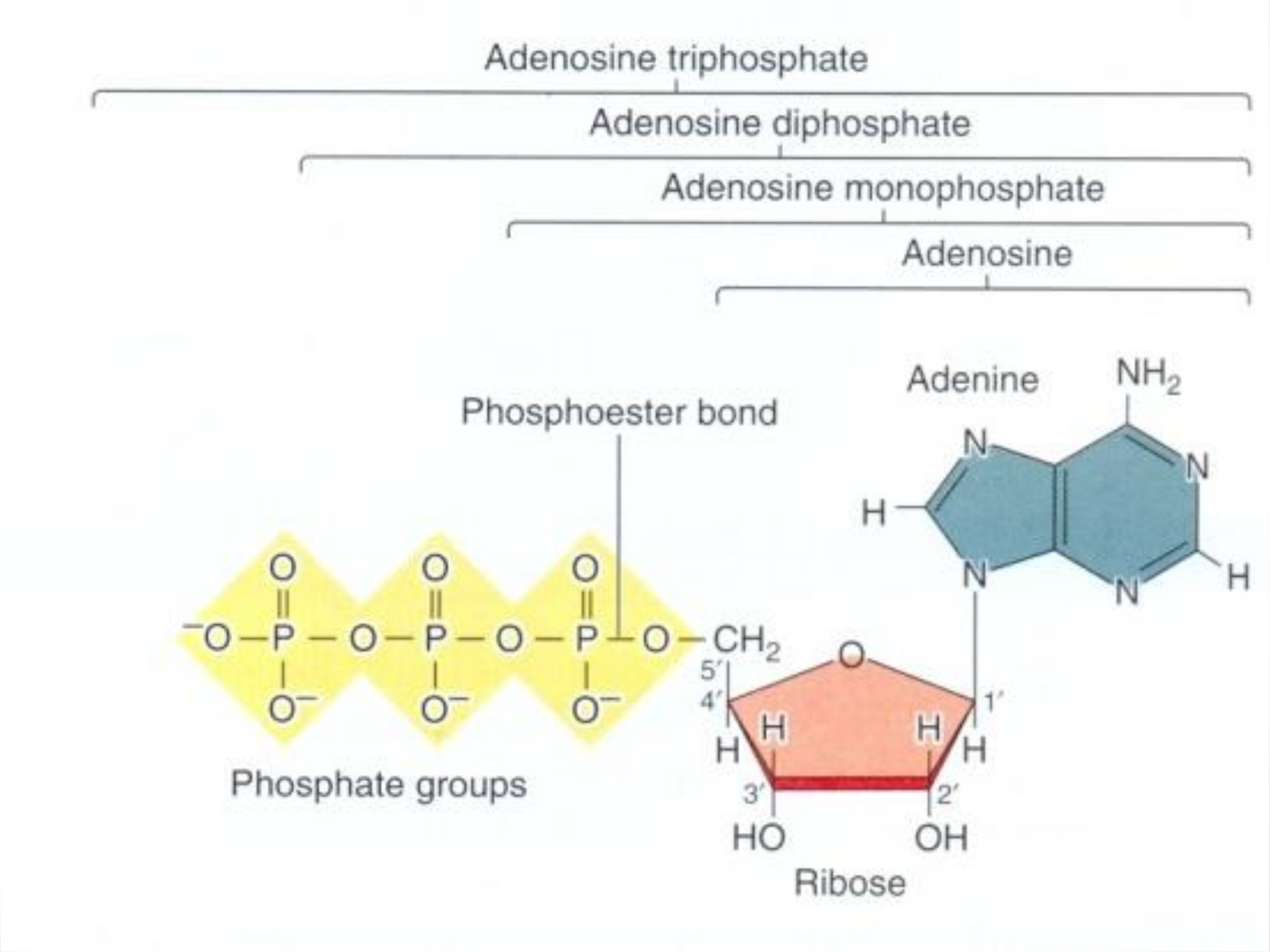
- The nitrogenous bases fall into two types:
  1. **Purines:** adenine (A), and guanine (G)
  2. **Pyrimidines:** cytosine (C), thymine (T), and uracil (U)

NUCLEOTIDE STRUCTURE

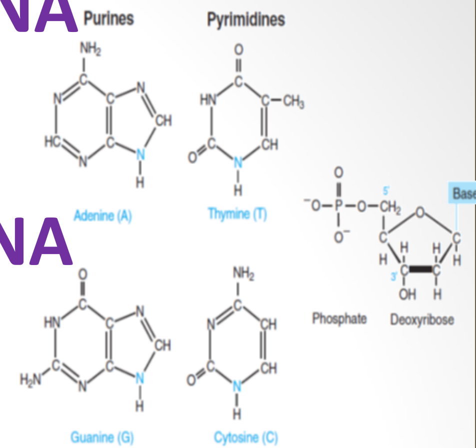


NUCLEOSIDE AND NUCLEOTIDE

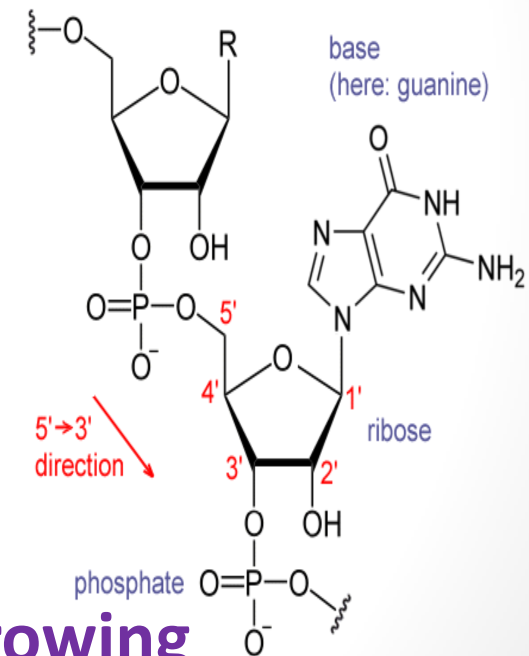
Base attached to sugar called nucleoside



- The sugar molecule composing the **DNA** strand is **Deoxyribose**.
- The sugar molecule composing the **RNA** strand is **Ribose**.



- Both sugar molecules are **pentose** sugars (five-carbon sugars)
- The Carbon **#1** is attached the **nitrogenous** base
- The Carbon **#3** is attached to the **phosphate** group
- The carbon **#5** is attached to the **growing chain** of polypeptides

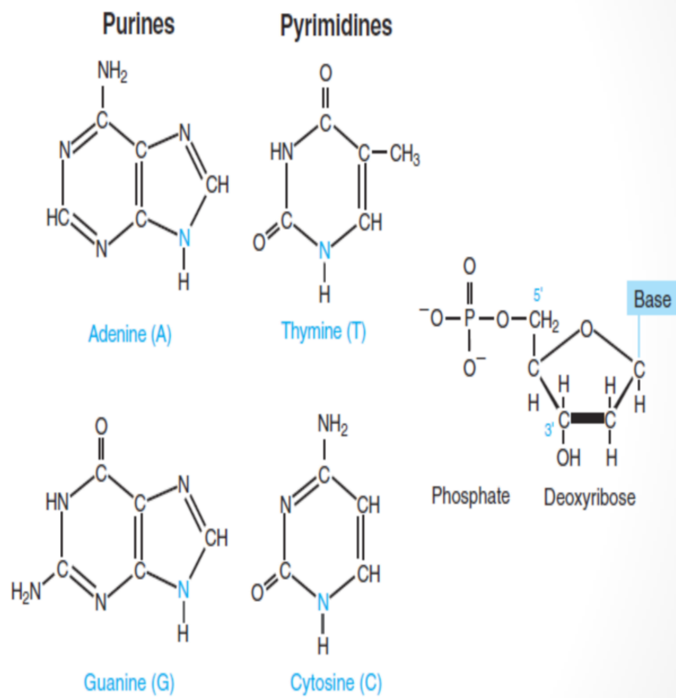


5/09/2015

elitedoctorsonline.com/dr630

Base Pairs in DNA:

The four bases of DNA and the general structure of a nucleotide in DNA. Each of the four bases bonds with deoxyribose (through the nitrogen shown in blue) and a phosphate group to form the corresponding nucleotides.

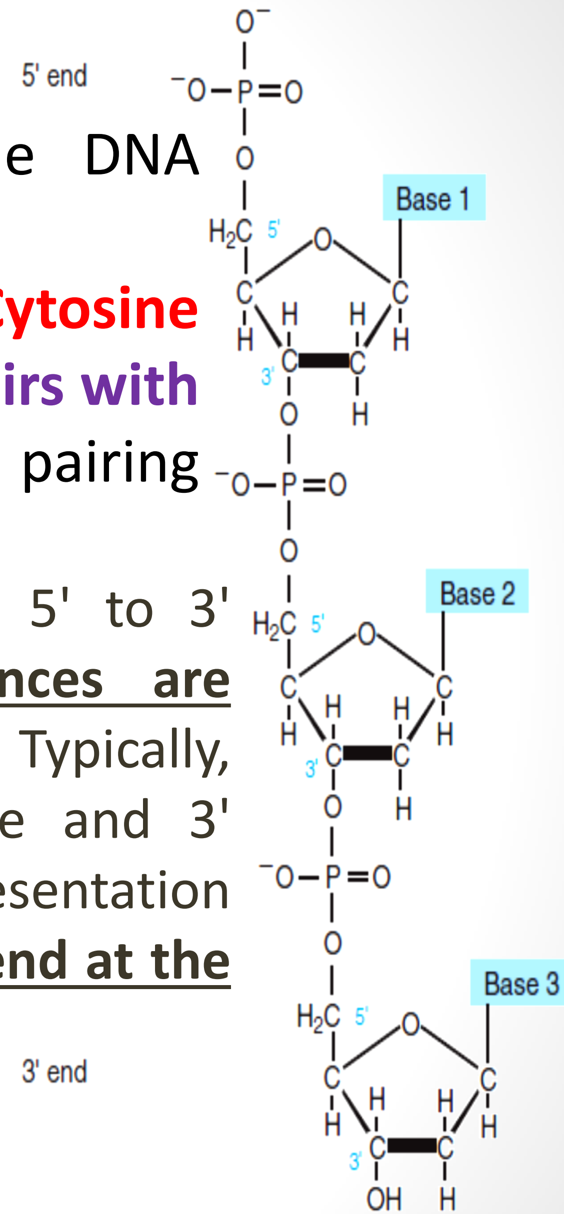


Base Pairs in RNA:

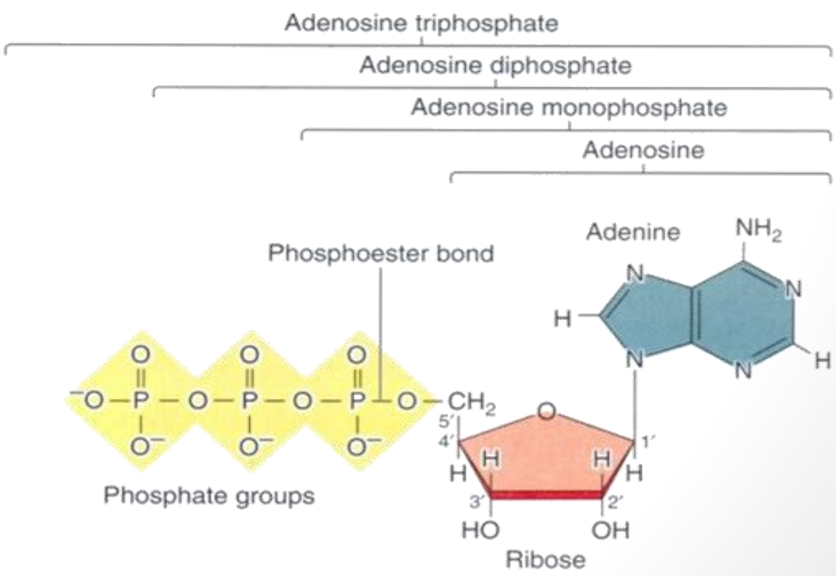
An important structural feature of RNA that distinguishes it from DNA is the presence of a hydroxyl group at the 2' position of the ribose sugar.

Nucleic Acid Strand:

- The **arrangement** of the bases in the DNA molecule is **not random**.
- **Guanine** in one chain **always pairs with Cytosine** in the other chain, and **Adenine always pairs with Thymine** (**Uracil** in RNA), so that this base pairing forms two **complementary** strands.
- Polynucleotide is connected by a series of 5' to 3' phosphate linkages. **Polynucleotide sequences are referenced in the 5' to 3' direction**. Typically, polynucleotides will contain a 5' phosphate and 3' hydroxyl terminal groups. The common representation of polynucleotides is as an **arrow** with the **5' end at the left and the 3' end at the right**.



- The **native** state of DNA, as elucidated by Watson and Crick in 1953, is a **double helix**. The **helical structure** resembles a **right-handed spiral staircase** in which its two **polynucleotide chains** run in opposite directions, held together by **hydrogen bonds** between pairs of bases.
- Because of the **complementary** nature of the two strands of DNA, knowledge of the sequence of nucleotide bases on one strand **automatically allows** us to determine the sequence of bases on the other strand.



DNA MOLECULAR STRUCTURE

Watson-Crick model

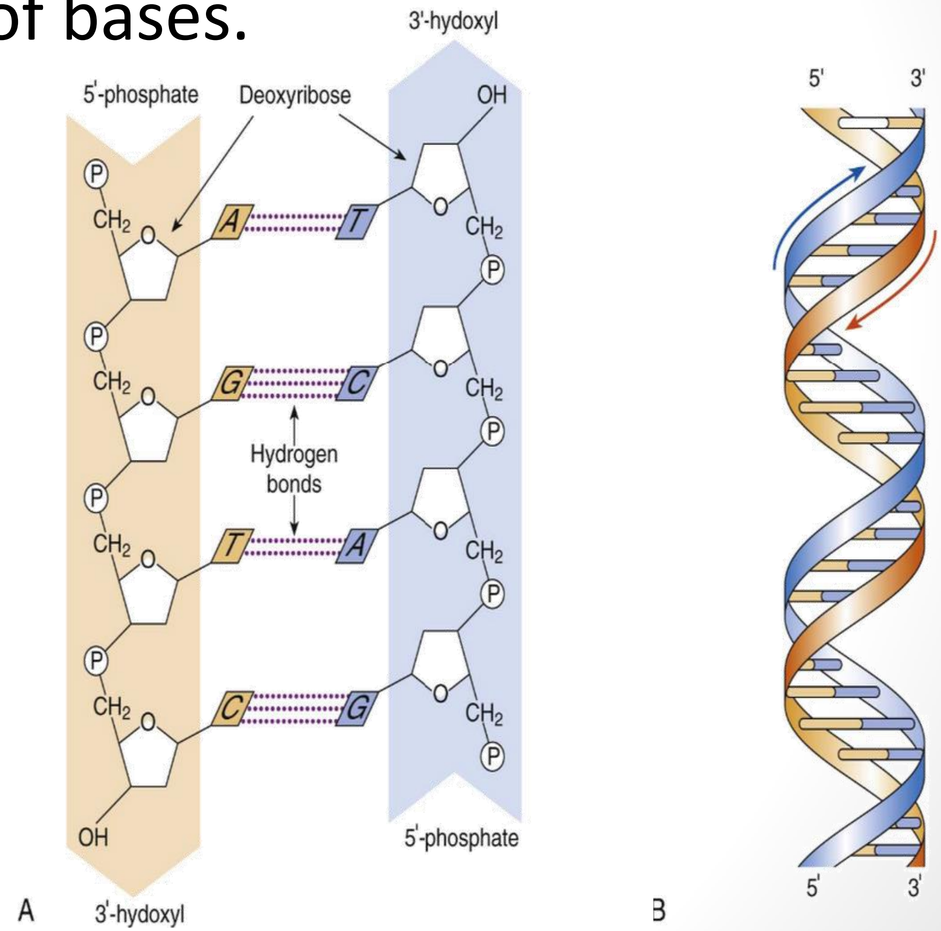
- Two DNA strands twisted together around central axis forming double helix (ds-DNA)
- Two DNA strands are antiparallel (opposite orientations)
  - One strand 5'-3'
  - Other strand 3'-5'
- Double stranded chains stabilized by hydrogen bond between opposite bases.
- Each strand of helix is complementary to the other (complementary base pairing)
- Adenine base in one strand forms 2 hydrogen bonds with Thymine (A = T) base on the opposite strand
- Guanine forms 3 hydrogen bonds with cytosine (G ≡ C)

- The **helical structure** resembles a **right-handed spiral staircase** in which its two **polynucleotide chains** runs in opposite directions, held together by **hydrogen bonds** between pairs of bases.

DNA Double Helix:

Sugar-phosphate backbone and nucleotide pairing of the DNA double helix.

P: phosphate
A: adenine
T: thymine
G: guanine
C: cytosine



SINGLE STRANDED DNA (ss-DNA) or RNA STRAND

Phosphate di-ester bond

- Formed between phosphate group on one nucleotide and sugar molecule on the adjacent nucleotide.
- A phosphate group connects two sugar molecules (di-nucleotides).
- Nucleotide linked together in a linear manner to form DNA or RNA strand.
- Phosphates and sugar molecules form backbone of DNA or RNA strand.

SS-DNA or RNA STRAND

- **Bases** project from the backbone.
- Backbone is (**-**) charged (**PO_4^-**).
- Orientation of nucleotides.
- Phosphodiester bond involves phosphate attachment to the 5’ carbon in one nucleotide and to the 3’ carbon in the other
- Strand direction is based on the orientation of the sugar molecules within the strand.

SS-DNA or RNA STRAND

- Direction of strand is 5’- to -3’.
- Strand contains a specific sequence of bases
- Sequence of bases

Thymine-Adenine

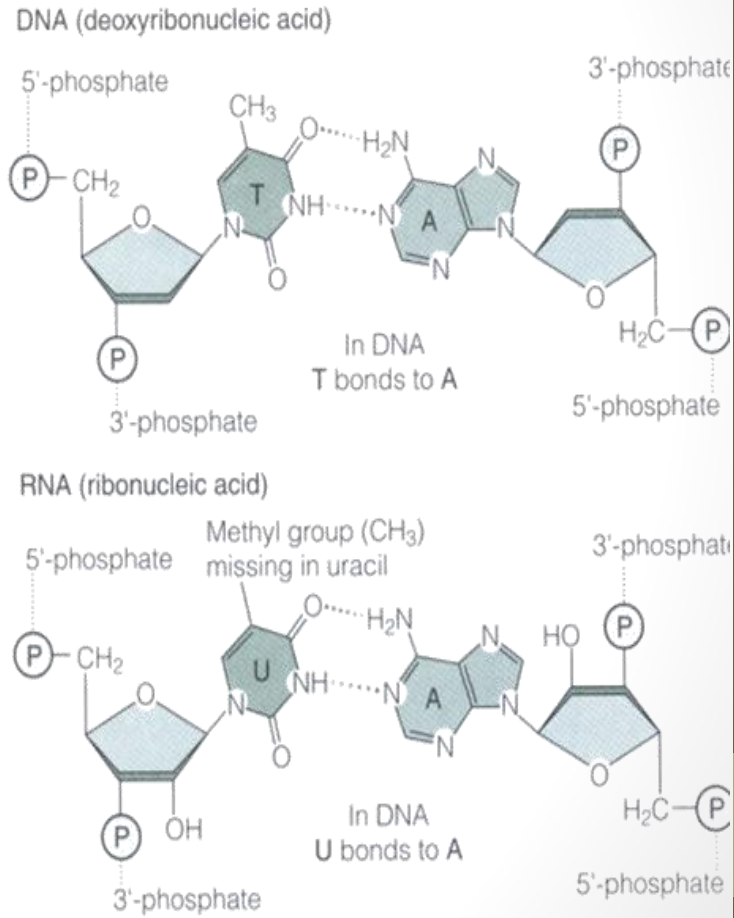
Cytosine-Guanine

abbreviated:

TACG

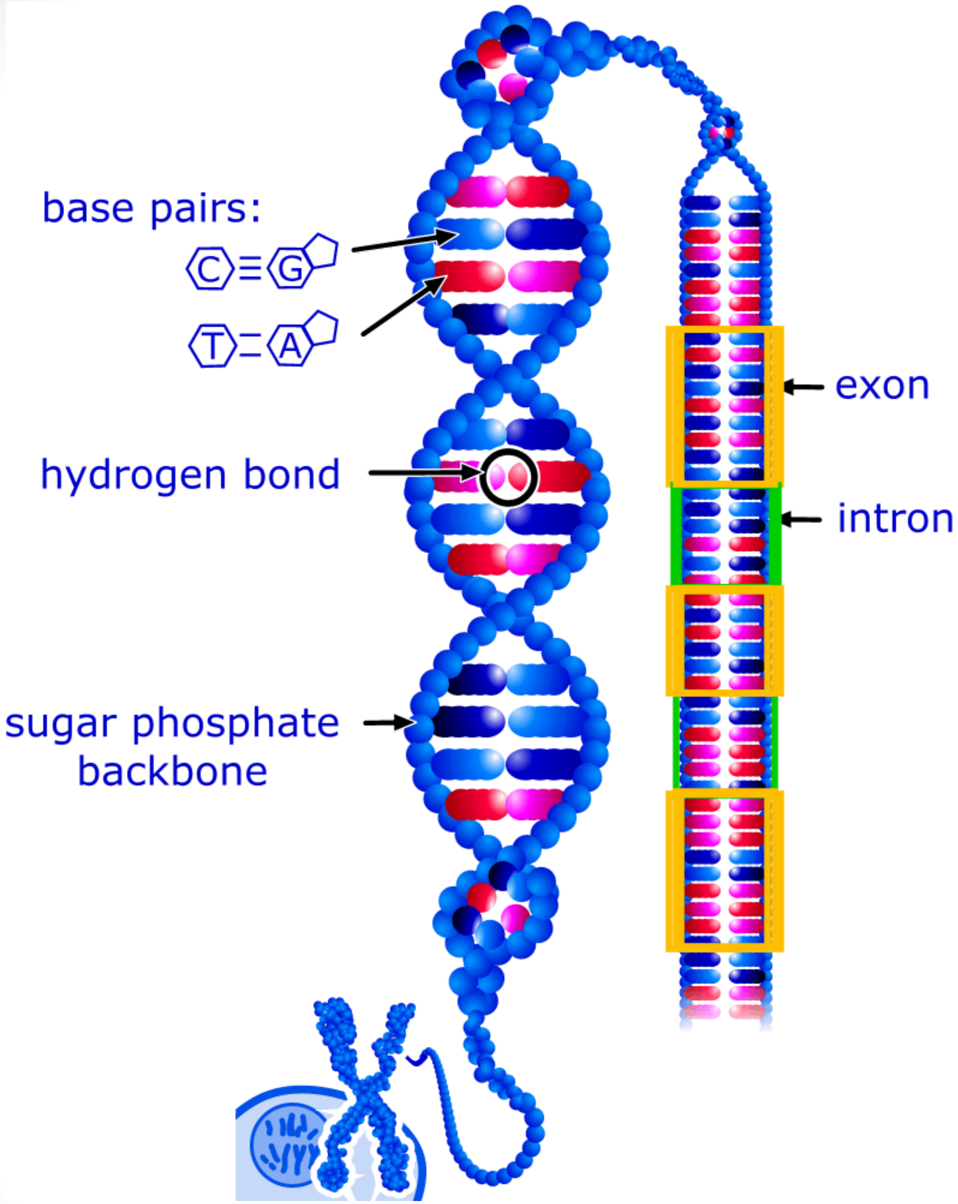
or

5’-TACG-3’



5/09/2015

elitedoctorsonline.com/dr630

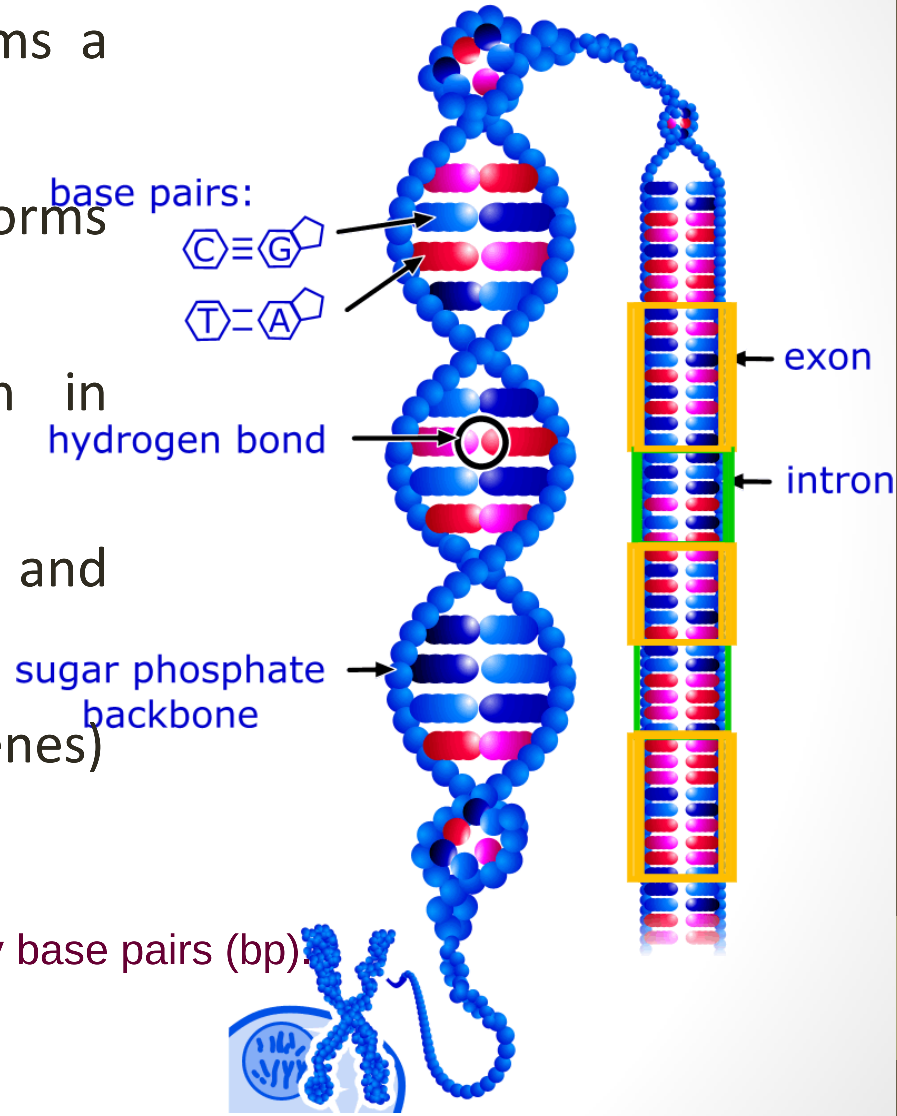


DNA

- ~3.2 billion base pairs in every cell build the human genome
- genes form only 1,5% of the human genome
- a gene is a segment of the DNA, that encodes the construction plan for a protein
- in humans there are ca. 30,000 genes only

DNA - Deoxyribonucleic acid

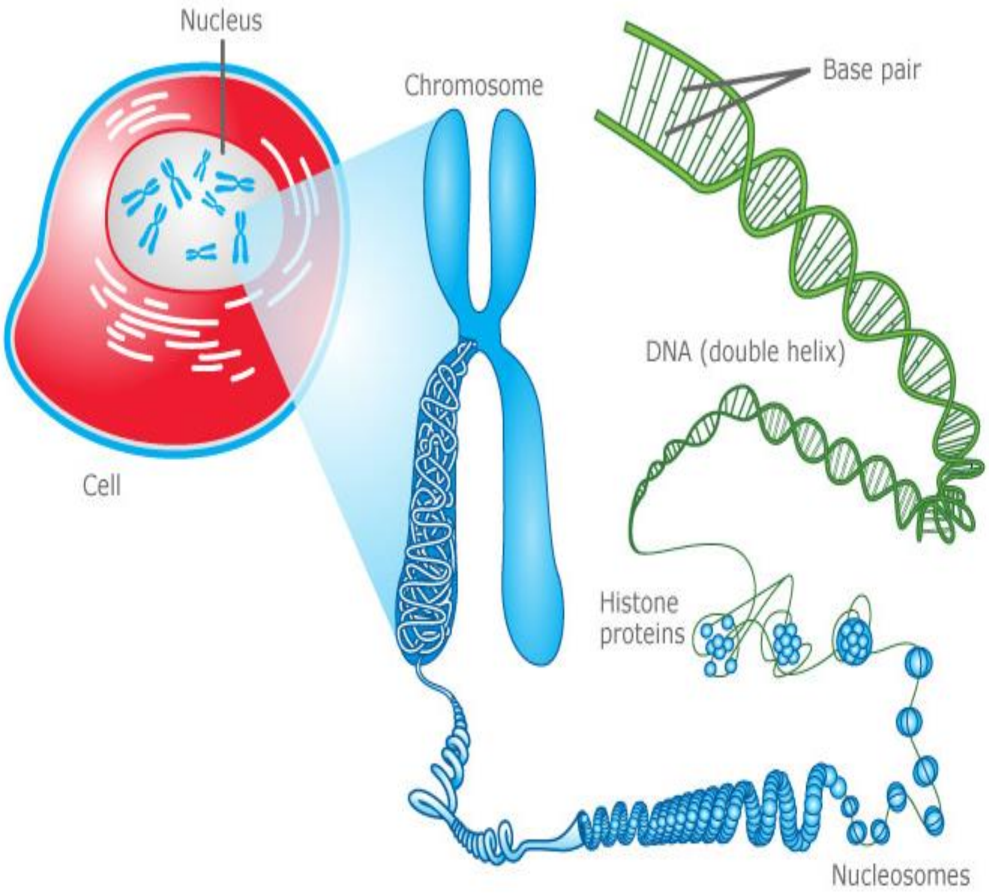
- Deoxyribonucleic acid (DNA) forms a double stranded helix.
- A sugar-phosphate backbone forms the outer shell on the helix
- The two strands of DNA run in opposite directions.
- Bases face towards each other and form hydrogen bonds
- carries the generic instructions (genes)



free Bases

Cytosine	-	C
Guanine	-	G
Adenine	-	A
Thymine	-	T

complementary base pairs (bp).



DNA

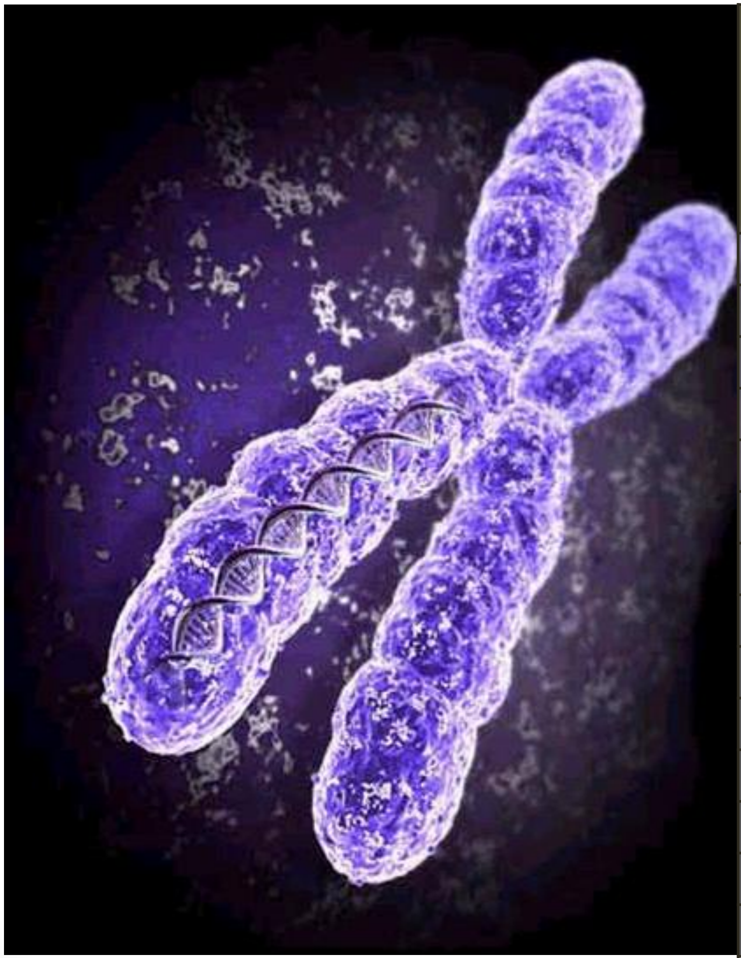
RNA

Protein

- DNA and RNA are simply long polymers of nucleotides called **polynucleotides**.
- Only the **α phosphate** is included in the polymer.
- It becomes chemically bonded to the **3' carbon** of the sugar moiety of another nucleotide.

Chromosome

A *chromosome* is a very long, continuous piece of DNA, which contains many genes, regulatory elements and other intervening nucleotide sequences.



http://www.tqnyc.org/NYC040844/Mitosis.htm

Chrom.	Genes	Bases
1	2968	245,203,898
2	2288	243,315,028
3	2032	199,411,731
4	1297	191,610,523
5	1643	180,967,295
6	1963	170,740,541
7	1443	158,431,299
8	1127	145,908,738
9	1299	134,505,819
10	1440	135,480,874
11	2093	134,978,784
12	1652	133,464,434
13	748	114,151,656
14	1098	105,311,216
15	1122	100,114,055
16	1098	89,995,999
17	1576	81,691,216
18	766	77,753,510
19	1454	63,790,860
20	927	63,644,868
21	303	46,976,537
22	288	49,476,972
X	1184	152,634,166
Y	231	50,961,097

RNA – Ribonucleic acid

In RNA the base *Thymine* (T) is replaced by *Uracil* (U). The other difference to DNA is that the sugar (*Pentose*) will be *Ribose* instead of *Deoxiribose*. Ribose has an *additional hydroxyl group*.

Bases

Cytosine - C
Guanine - G
Adenine - A
Uracil - U

RNA transmits genetic information from DNA (via transcription) into proteins (by translation).
RNA is almost exclusively found in the single-stranded form.

RNA plays several roles in biology:

- **Messenger RNA** (**mRNA**) is transcribed directly from a gene's DNA and is used to encode proteins.

- RNA genes are genes that encode functional RNA molecules; in contrast to mRNA, these RNA do not code for proteins. The best-known examples of RNA genes are **transfer RNA** (**tRNA**) and **ribosomal RNA** (**rRNA**). Both forms participate in the process of translation, but many others exist.

- RNA forms the genetic material (genomes) of some kinds of viruses.

- **Double-stranded RNA** (**dsRNA**) is used as the genetic material of some RNA viruses and is involved in some cellular processes, such as RNA interference.

Proteins

Proteins have a variety of roles that they must fulfil:

1. they are the enzymes that rearrange chemical bonds.
2. they carry signals to and from the outside of the cell, and within the cell.
3. they transport small molecules.
4. they form many of the cellular structures.
5. they regulate cell processes, turning them on and off and controlling their rates.

Proteins – Amino Acids

- There are 20 different types of amino acids (see below).
- different sequences of amino acids *fold* into different 3-D shapes.
- Proteins can range from fewer than 20 to more than 5000 amino acids in length.
- Each protein that an organism can produce is encoded in a piece of the DNA called a “gene”.
- the single-celled bacterium *E.coli* has about 4300 different genes.
- Humans are believed to have about 30,000 different genes (the exact number as yet unresolved),

PROTEINS

Polypeptide

- N-terminus ($-NH_2$)
- C-terminus (-COOH)

Protein structure

- Primary
  - Linear sequence of amino acids
- Secondary
  - Polypeptide chains arrangements in helices and non-helices
- Tertiary
  - Arrangement of helices into 3D structure
- Quaternary
  - Multiple polypeptide fold together

Proteins – Amino Acids

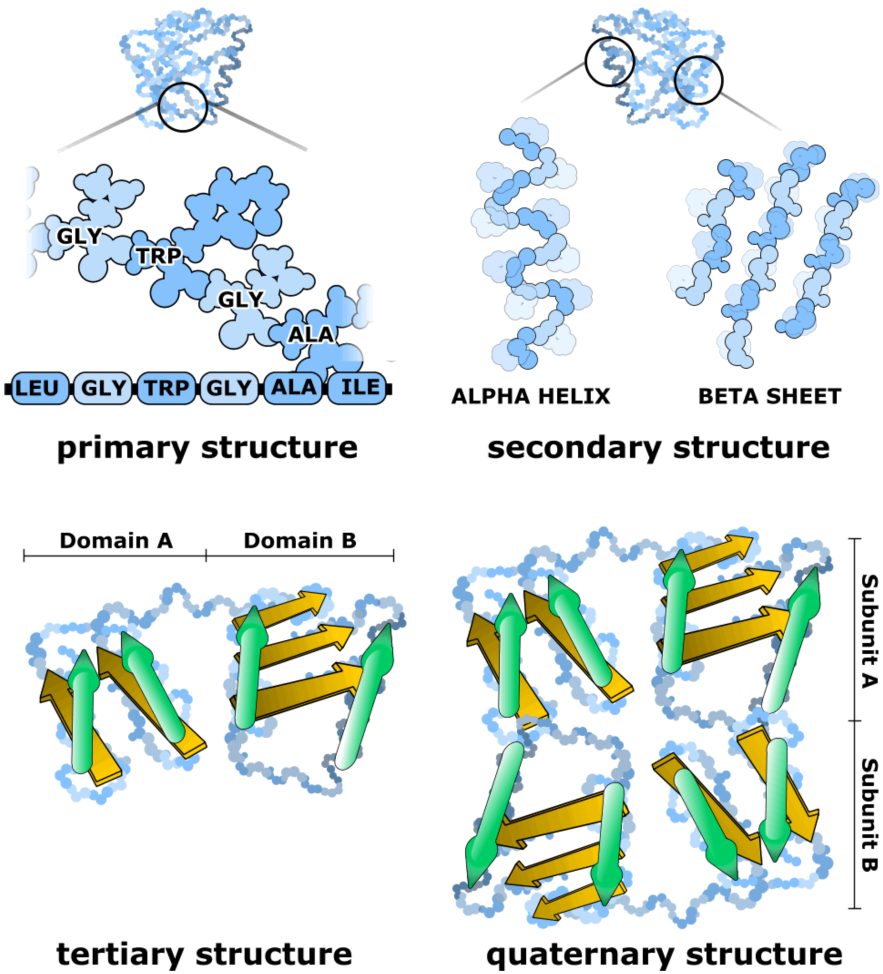
Name	1-letter code	Triplet
Glycine	G	GGT,GGC,GGA,GGG
Alanine	A	GCT,GCC,GCA,GCG
Valine	V	GTT,GTC,GTA,GTG
Leucine	L	TTG,TTA,CTT,CTC,CTA,CTG
Isoleucine	I	ATT,ATC,ATA
Histidine	H	CAT,CAC
Serine	S	TCT,TCC,TCA,TCG,AGT,AGC
Threonine	T	ACT,ACC,ACA,ACG
Cysteine	C	TGT,TGC
Methionine	M	ATG
Glutamic Acid	E	GAA,GAG
Aspartic Acid	D	GAT,GAC,AAT,AAC
Lysine	K	AAA,AAG
Arginine	R	CGT,CGC,CGA,CGG,AGA,AGG
Asparagine	N	AAT,AAC
Glutamine	Q	CAA,CAG
Phenylalanine	F	TTT,TTC
Tyrosine	Y	TAT,TAC
Tryptophan	W	TGG
Proline	P	CCT,CCC,CCA,CCG
Terminator (Stop)	*	TAA,TAG,TGA

Protein-Sequence Alphabet :-
ACDEFGHIKLMNPQRSTVWY

MENFQKVEKIGEGTYGVVY
KARNKLTGEVVALKKIRLDT
ETEGVPSTAIREISLLK...

Atypical human cell contains about 100 million proteins of about 10,000 types.

Proteins



Primary protein structure
Is the sequence of a chain of amino acids

Secondary protein structure
occurs when the sequence of amino acids are linked by hydrogen bonds.

Tertiary protein structure
occurs when certain attractions are present between alpha helices and pleated sheets.

Quaternary protein structure
Is a protein consisting of more than one amino acid chain..

Proteins - Summary

- DNA sequence determines protein sequence
- Protein sequence determines protein structure
- Protein structure determines protein folding and function.

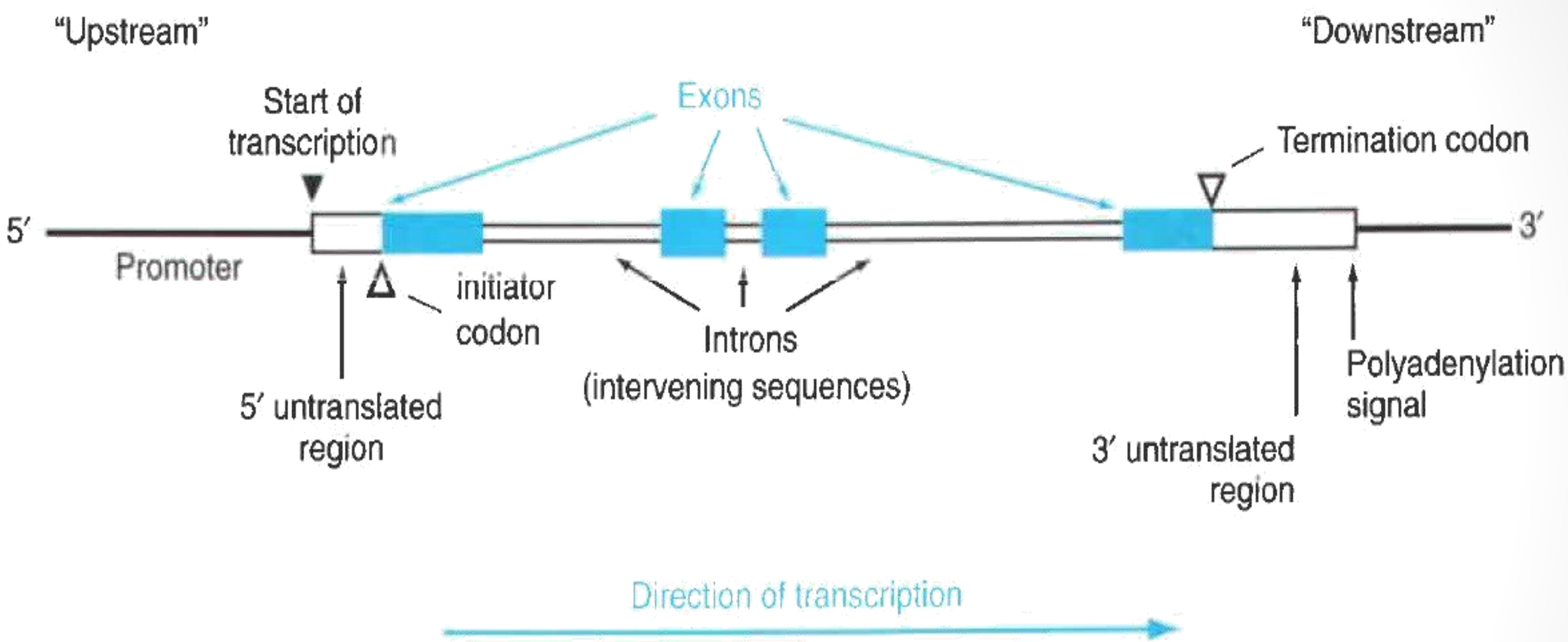
GENE

Molecular unit of **heredity**. Gene controls **trait** (eye, skin colour, height and intelligent). A trait can be governed by single gene (monogenetic, Mendlian or unifactorial) or multiple genes (polygenetic, multifactorial). Generally, genes are inherited from parent (paternal and maternal). Most genes exist in **pair** (2 alleles) and each separate from another during gamete formation (sperm and egg). The majority of genes form part of nuclear **chromosomes** inside the nucleus (Nuclear genes) and few form part of mitochondrial chromosome in the cytoplasm (Mitochondrial genes). They are mostly organized in tandem on all cellular chromosomes, but only spaced by nonfunctional or regulatory sequences on nuclear chromosomes.

GENE STRUCTURE

Each gene consists of Deoxyribonucleic acid (**DNA**) segment (sequence). **Nucleotide** is the building block of DNA. It is composed of Deoxyribose pentose sugar, phosphate molecule and nitrogenous bases (**A, T, C, and G**). The lengths of genes vary and are measured with nucleotide unit, bp or b (e.g. 1000 bp). Some genes are only a few Kbp in length, others are hundreds of Kbp (Human dystrophin gene on X chromosome, is 2.4 Mb). Most of genes are made up of varying number and lengths of **exons** and **introns** (exceptions). **Exons** (coding sequences) are DNA sequence of a gene that is expressed (transcribed) into RNA and translated into protein. Each of gene exon codes for a specific portion of the complete protein. **Eukaryotic** gene exons are interrupted by intervening sequence called **introns** (non coding sequence) that are transcribed but spliced out during post-transcriptional modification process and **exons** are joined to produce mature mRNA. The number and length of exons and introns within human genes do vary. The maximum number of **exons** in a human gene is 363 exons. The longest exon in human is 12 kbp and shortest is 2bp. The maximum number of **introns** in a single human gene is 147 introns. Lengths of introns in human genes do vary from 1 bp to 0.5 Mb.

GENE STRUCTURE



GENE EXPRESSION

All genes have regulatory regions upstream of the coding regions of genes known as **promoters**. Promoter is DNA sequence upfront of a coding gene that is recognized by the transcription machinery that promotes and regulates gene expression. The sequence of the gene encodes for the synthesis of protein or RNA. The **expression of the gene** (two steps: **transcription** generating messenger RNA (mRNA) transcript and **translation** synthesizing protein) is tightly regulated. Genes are switched on and off at specific time and with specific quantity. **House keeping genes** are expressed in all type of cells and **tissue specific genes** are selectively functional in specific cell types. During nuclear gene expression, the nuclear gene sequence is transcribed into intermediate sequence; mRNA transcript inside the nucleus and is transported into the cytoplasm where it binds to ribosome complex for translation. Every 3 nucleotides of the mRNA transcript (**codon**) codes for specific **amino acid** (building block of proteins). The sequence of the gene and its mRNA determine the linear amino acid sequence of polypeptide encoded (**colinearity**).

GENE FUNCTION

Most genes encodes for specific proteins that play a significant role in the different cellular function, structure and activity including **enzymes** that catalyze metabolic reactions, **antibodies** as component of Immunity to infection, **transcription factors**(P53) regulating gene expression, **structural protein**s providing cell support and shape (actin), **membrane Ion channels** proteins transporting Na^+ and Cl^- **ions**, Cell surface **receptors**, etc.

GENE MUTATION

Any changes in DNA sequence due to **gene mutation** can lead to production of dysfunctional proteins and mutant phenotype and subsequent appearance of genetic diseases.

GENOME

In human, it is estimated that human **genome** (all DNA sequences, 3 billion bp in the nucleus, harbors up to 30,000 genes encodes approximately 1 million different proteins; **proteomic**). **Genomic** is a pioneered field of science studying genome structure and function through comparative genomic to elucidate **gene function**.

GENOME

- Hybrid of **GENE** and **CHROMOSOME**.
- Denoted complete set of chromosomes and its genes.
- All DNA in haploid set of chromosomes.

GENOMICS

- Study of genomes of various organisms:
  - Human
  - Microorganisms
  - Plants
  - Animals

GENOMICS

Genome

Chromosomes 46

Genes 30,000

Nucleotide $3X10^9$ bp

- A
- T
- C
- G

GENOMIC GOALS

- Locating genes on chromosome (genetic map)
- Elucidating gene function and regulation.
- Identification of all proteins in genome and their functions.
- Identifying DNA polymorphism (variation).
- Comparing genes and proteins on genomes between species (**comparative genomics**)
- Databases of the genomes.

SINGLE STRANDED DNA (ssDNA) or RNA STRAND

Phosphate di-ester bond

- Formed between phosphate group on one nucleotide and sugar molecule on the adjacent nucleotide.
- A phosphate group connects two sugar molecules (**di-nucleotides**).
- Nucleotide linked together in a linear manner to form DNA or RNA **strand**.
- Phosphates and sugar molecules form **backbone** of DNA or RNA strand.

ssDNA or RNA STRAND

- **Bases** project from the backbone.
- Backbone is (**-**) charged (**PO_4^-**).
- Orientation of nucleotides.
- Phosphodiester bond involves phosphate attachment to the 5’ carbon in one nucleotide and to the 3’ carbon in the other
- Strand direction is based on the orientation of the sugar molecules within the strand.

ssDNA or RNA STRAND

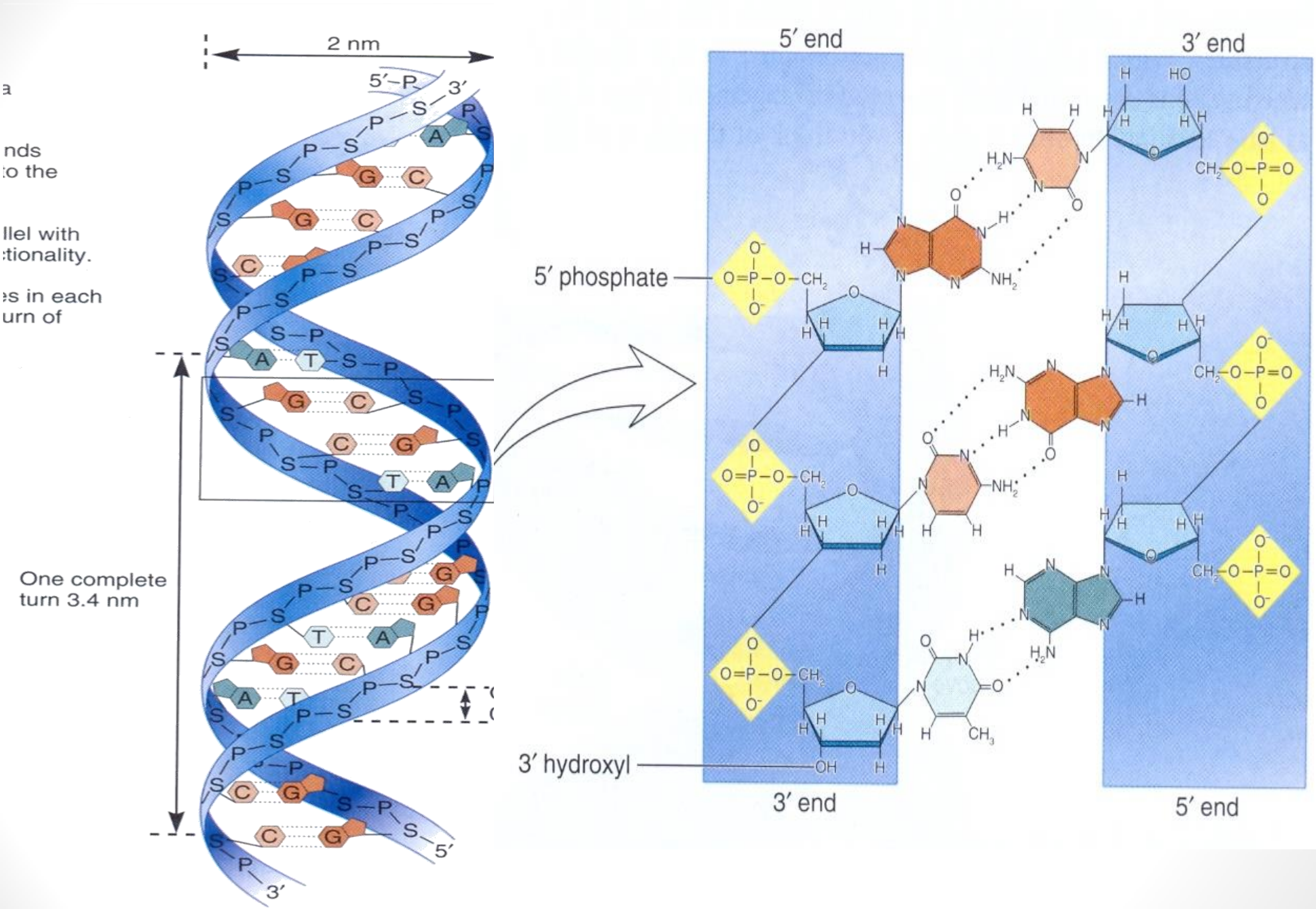
- Direction of strand is 5’- to -3’.
- Strand contains a specific sequence of bases
- Sequence of bases Thymine-Adenine -Cytosine-Guanine abbreviated:

TACG

or

5’-TACG-3’

DOUBLE HELIX STRUCTURE



COMPLEMENTARY BASE PAIRING

DNA denaturation

- Hydrogen bonds between two strands separate (denatured)
    - Higher temperature
    - pH extremes (pH >10)
    - Enzymatic

DNA renaturation (annealing)

- Gradually lowering temperature, complementary strands re-anneal or Hybridize.

DNA HYBRIDIZATION

- ss-DNA to complementary ss-DNA
- mRNA to complementary ss-DNA
- Oligonucleotide to complementary ss-DNA
- cDNA to complementary ss-DNA
- Microarray (genchip).
- Southern blot.

MOLECULAR STRUCTURE OF DNA

- DNA sequence have high proportion of
  GC contents tends to form more stable double stranded structures

  5'-TACCGCATT-3'

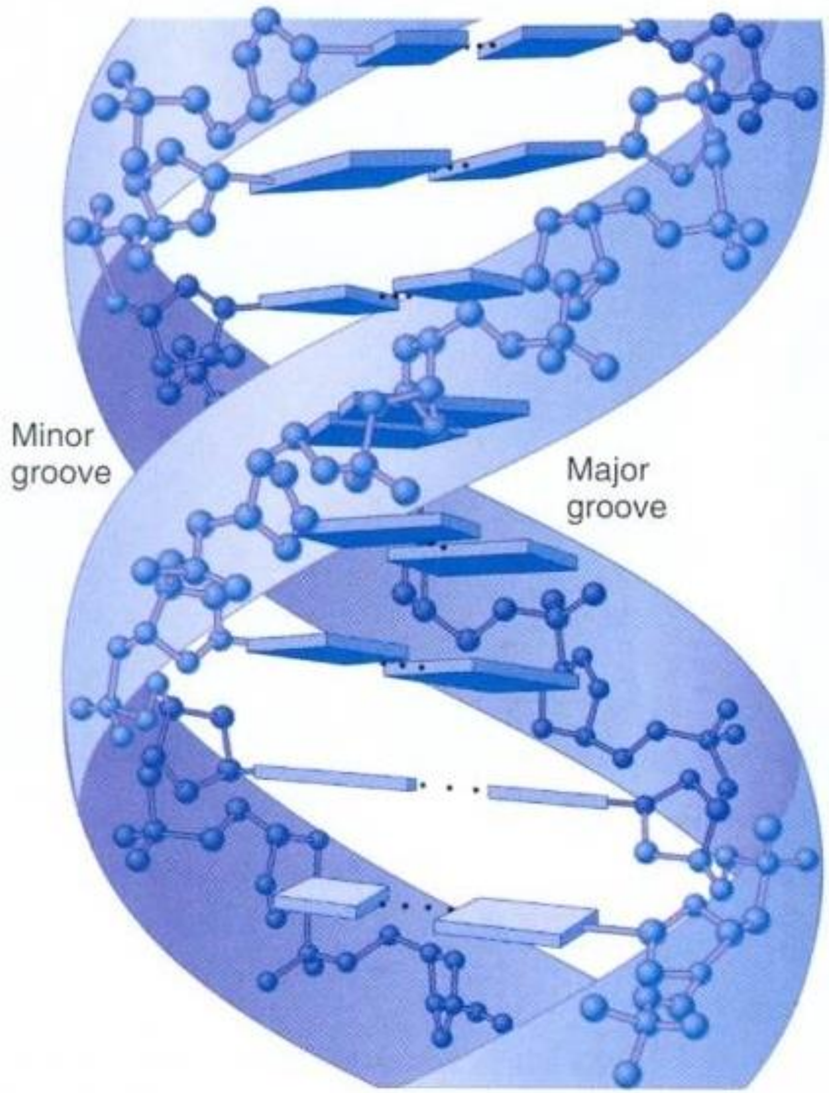
  3'-ATGGCGTAA-5'

- Right handed model

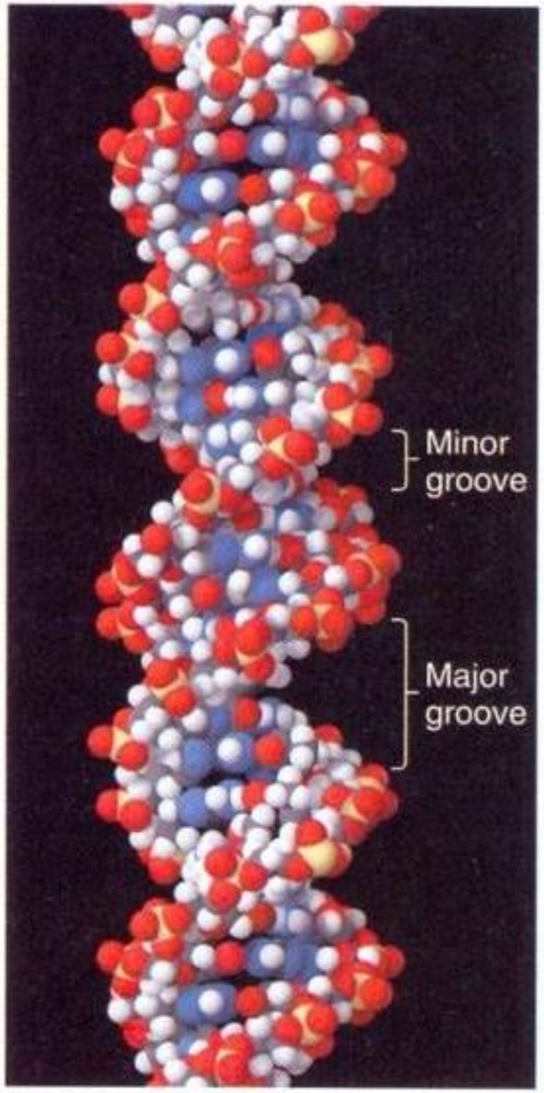
DNA Grooves

- DNA helix have two grooves
  - Minor grooves
  - Major grooves

DOUBLE HELIX MODEL

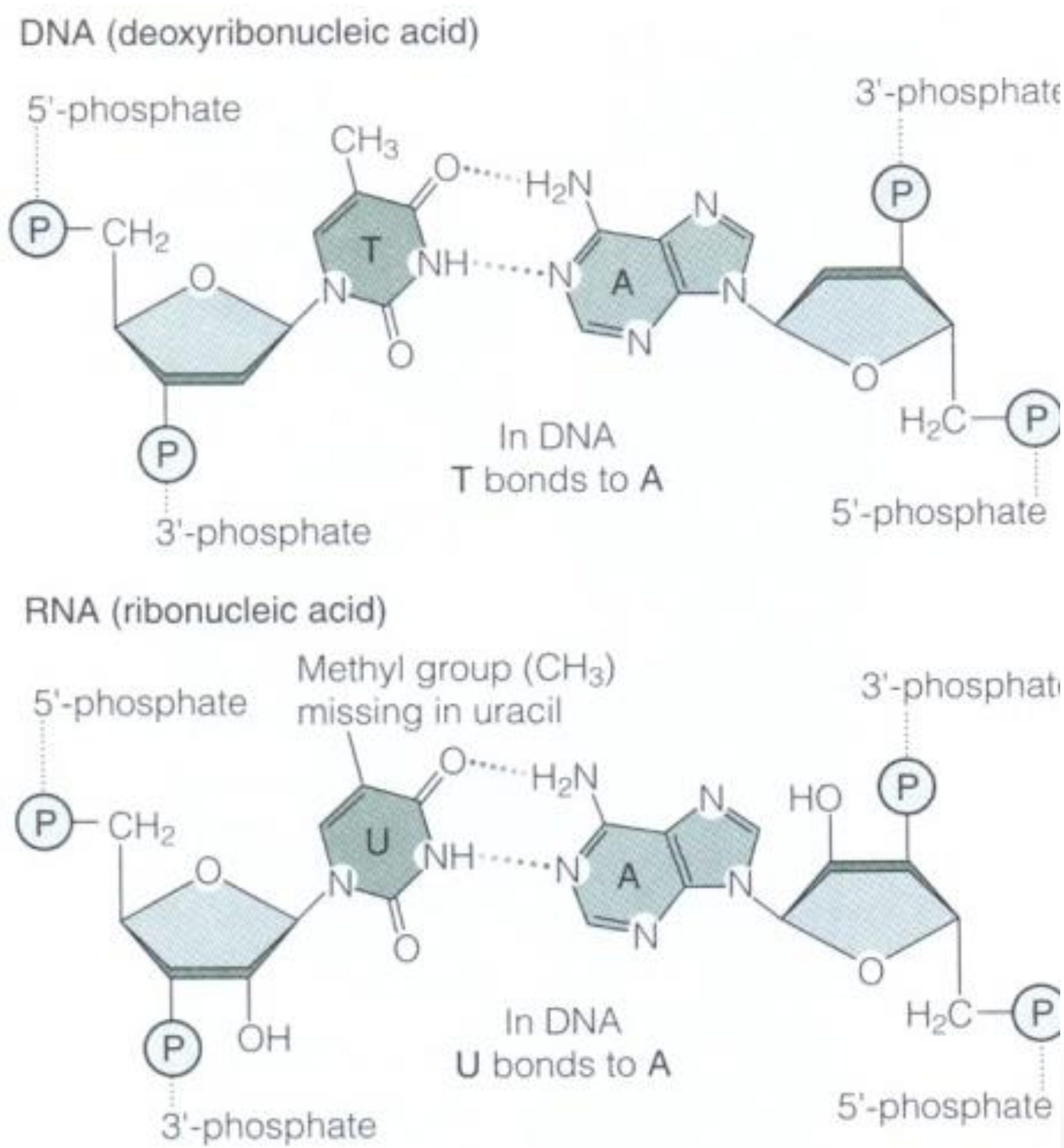


(a) Ball-and-stick model of DNA



(b) Space-filling model of DNA

BASE PAIRING



5/09/2015

elitedoctorsonline.com/dr630

DNA AND CHROMOSOMES

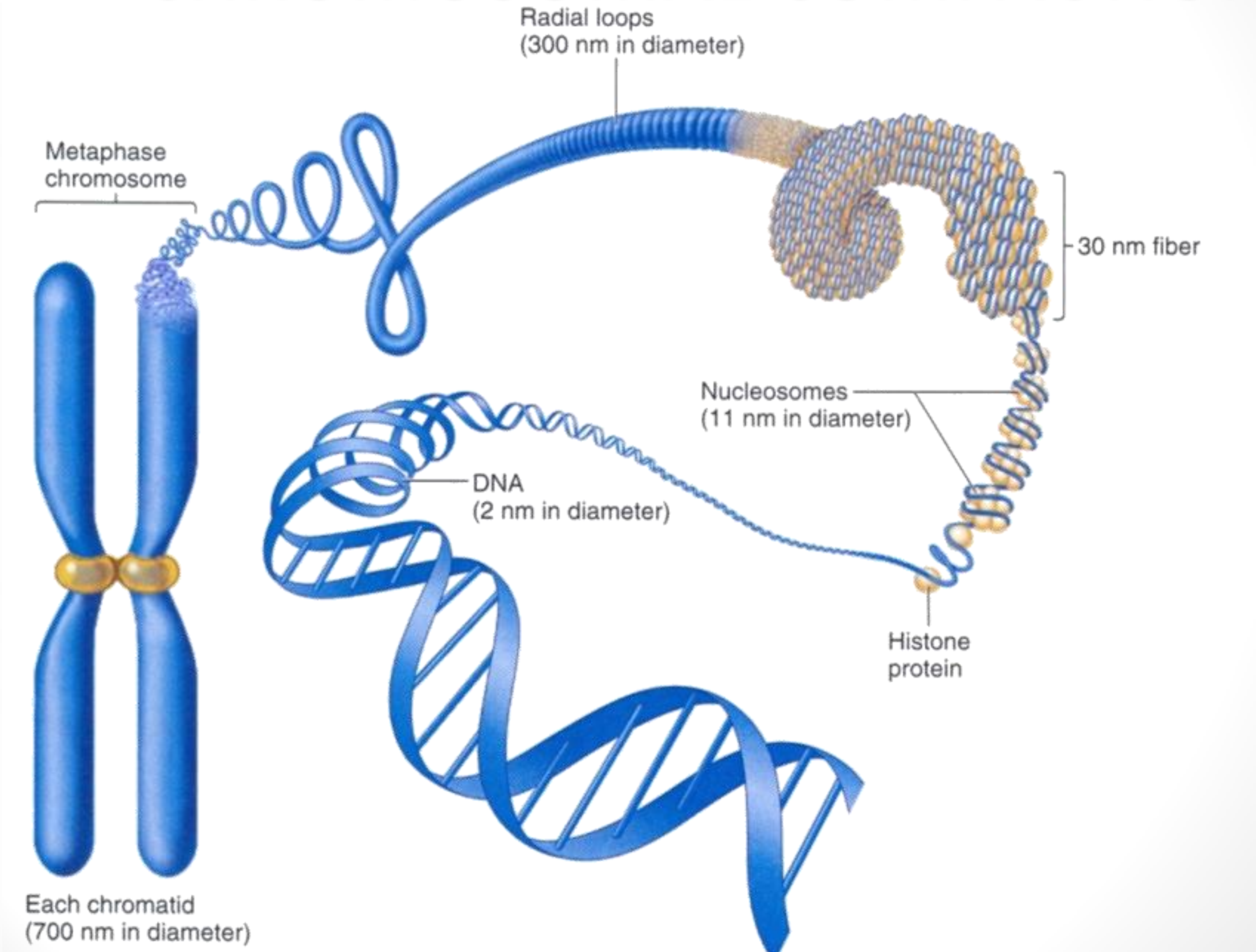
Chromosome

- Made up of linear ds-DNA wrapped around **histone** proteins to form **nucleosome**
- Nucleosome is repeating structural units associate, twist, fold and compact to form eukaryotic **chromosomes**

Prokaryote chromosome

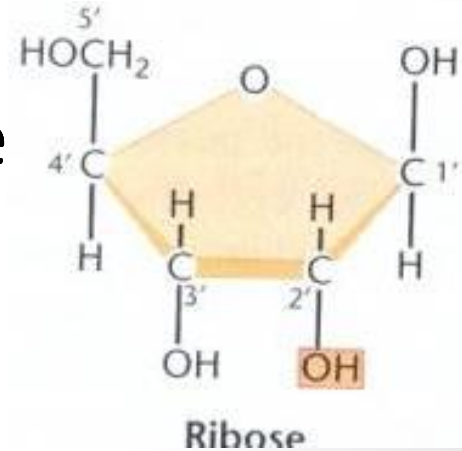
- Circular ds-DNA, no protiens.

EUKARYOTIC METAPHASE CHROMOSOMAL COMPACTION

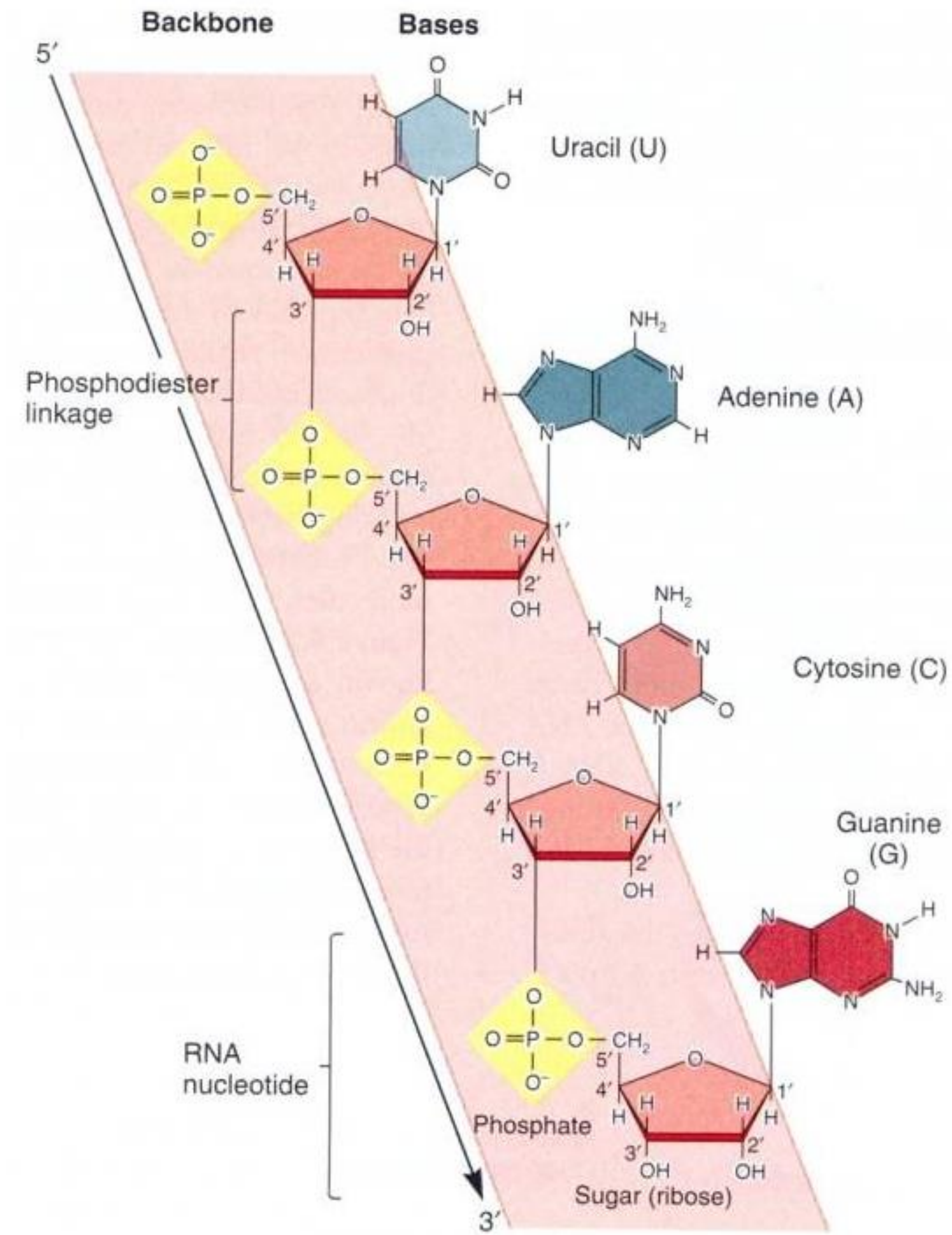


RNA STRUCTURE

- Ribonucleic acid.
- Structure similar to DNA with several exceptions
    1) Ribose sugar.
    2) Uracil nucleotide base.
    3) Single stranded.
  - Refold to form double stranded (complementary regions)
  - RNA molecules present transie degraded (unstable)



RNA STRAND



Backbone
Bases
5′
Uracil (U)
Adenine (A)
Cytosine (C)
Guanine (G)
Phosphodiester linkage
RNA nucleotide
Phosphate
Sugar (ribose)
3′

RNA HYBRIDIZATION

- mRNA to ss-DNA.
- mRNA to c-DNA.
- Oligonucleotide to mRNA.
- mRNA to anti-mRNA.
- MicroRNA.

THREE MAJOR CLASSES OF RNA MOLECULES

(1)Ribosomal RNA (rRNA)
(2)Messenger RNA (mRNA)
(3)Transfer RNA (tRNA)

1) rRNA

- Structural component of ribosomes
- **Translation**

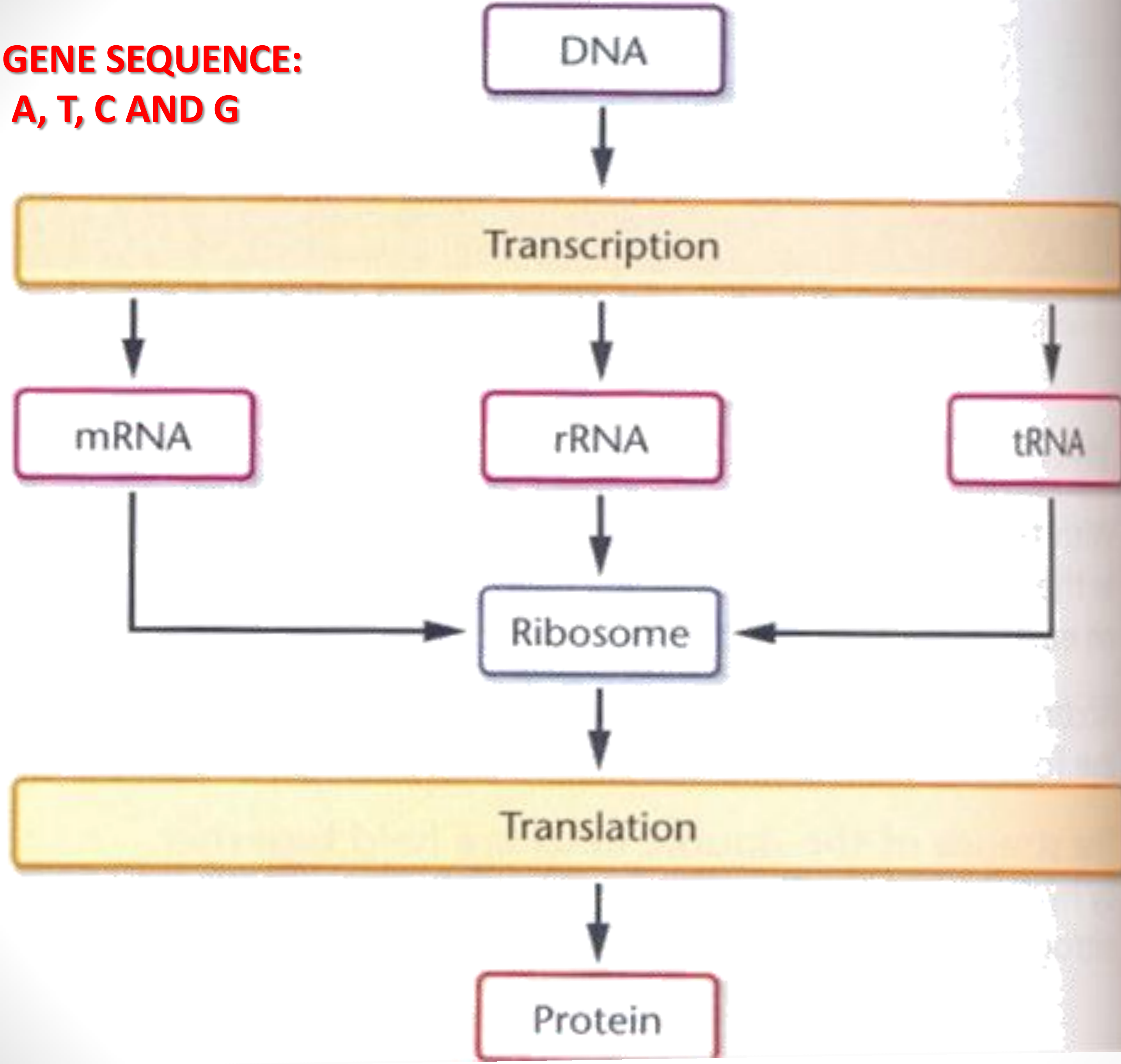
2) mRNA

- Carry genetic information from DNA of the gene
  to the ribosome for **translation**

3) TRANSFER RNA

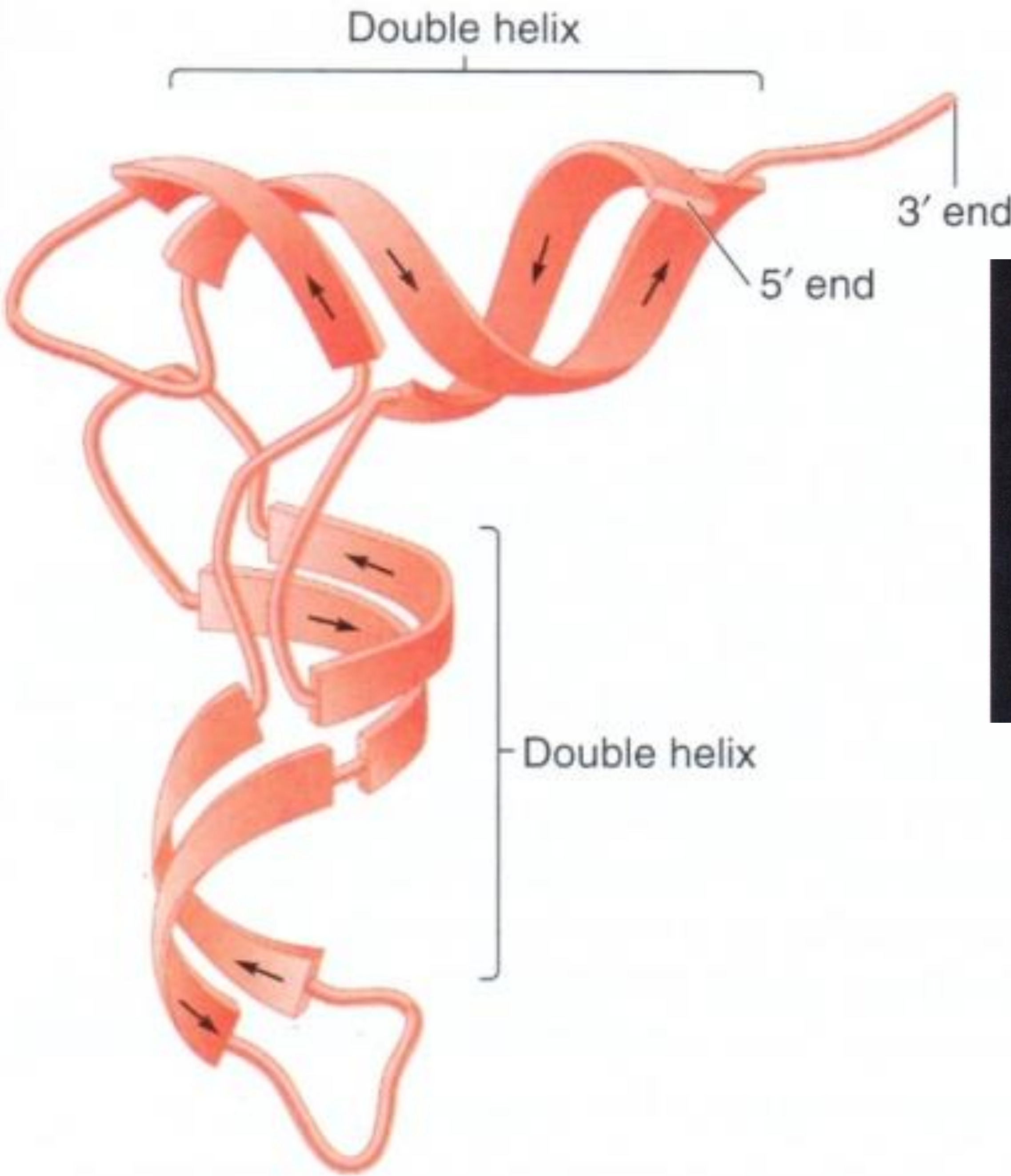
- Smallest RNA molecules.
- Carry a.a. to ribosome during **translation**

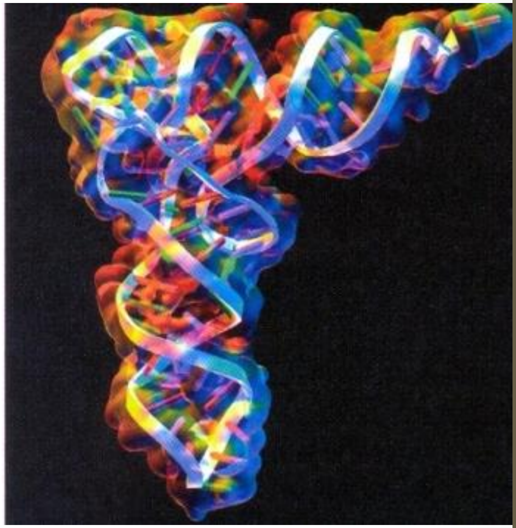
GENE SEQUENCE:
A, T, C AND G



GENE EXPRESSION

tRNA





UNIQUE RNA

1) **Small nuclear RNA (sn-RNA)**
   - Processing of mRNA.
2) **Telomerase RNA**
   - Chromosome end replication.
   - Animals and yeast.
3) **Antisense RNA**
   - Short interfering RNA (gene regulation).

PROTEIN STRUCTURE AND FUNCTION

- Long organic molecule

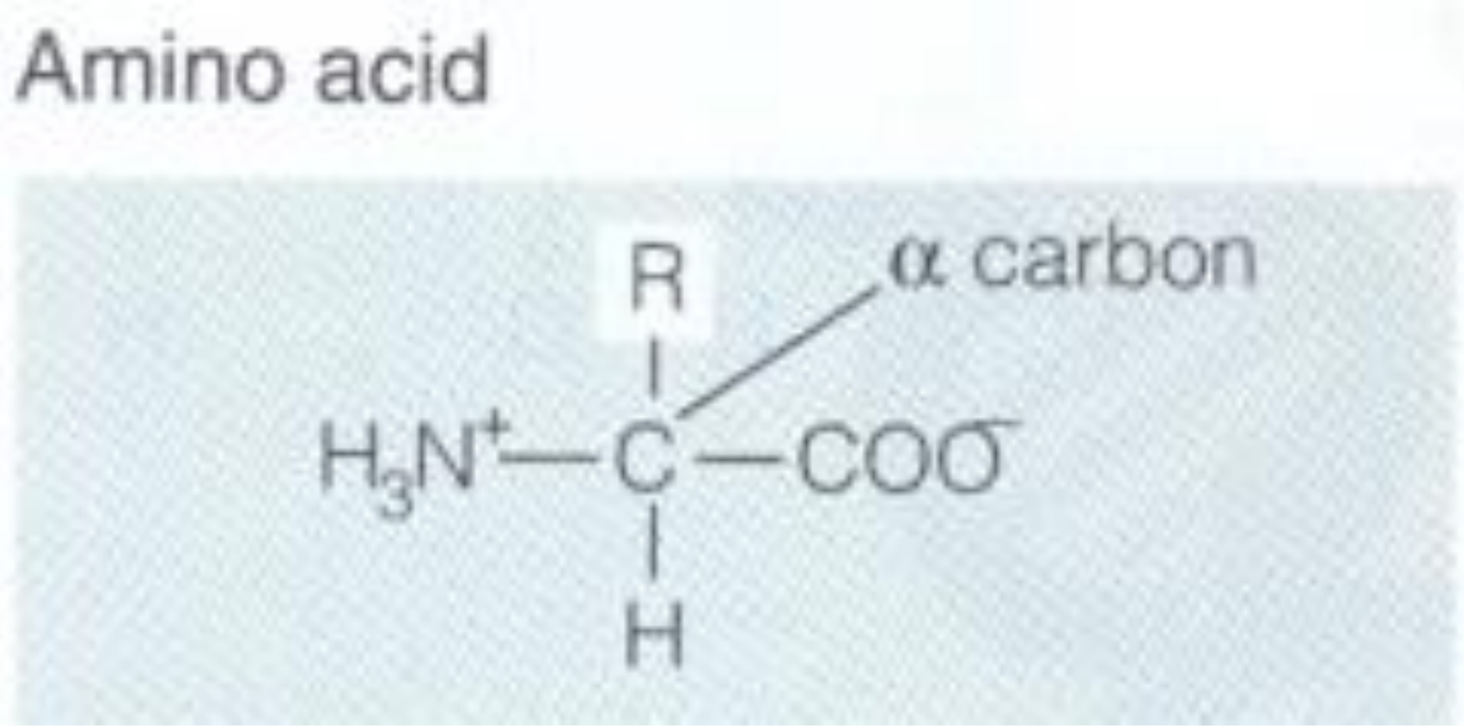
Classes

- Enzymes
- Hormones
- Antibodies
- Transcription factors
- Structural protein
- Regulating protein

PROTEIN STRUCTURE

- Proteins made of a single polypeptide
  chain or multiple **polypeptides**
- Polypeptide composed of monomers,
  amino acids
- 20 different amino acid exist
- Amino acid composed of basic backbone
  and unique side group

AMINO ACID STRCUTURE

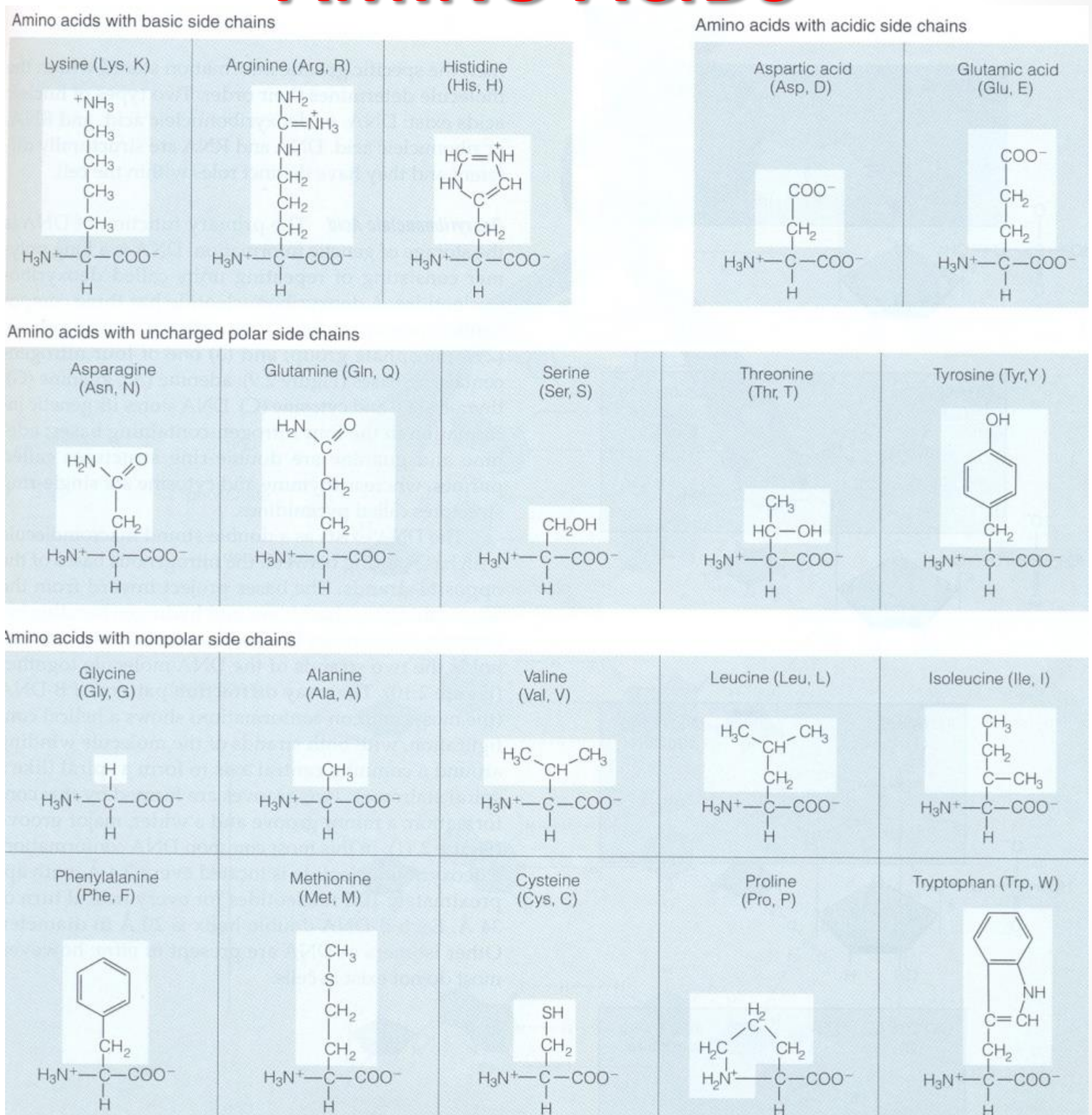


PROTEINS

Classes of amino acids

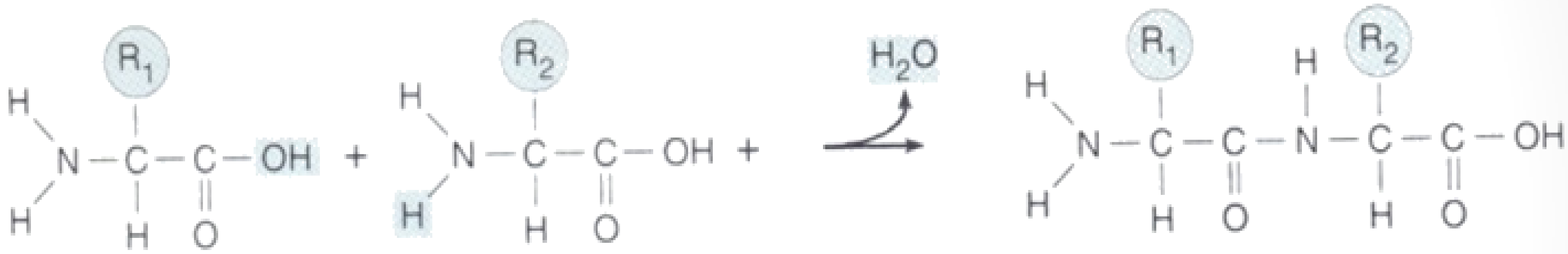
- Nonpolar
- Uncharged polar
- Negatively charged polar **(acidic)**
- Positively charged polar **(basic)**
- Amino acids joint by **peptide bond**

AMINO ACIDS



Amino acids with basic side chains
Lysine (Lys, K)
Arginine (Arg, R)
Histidine (His, H)
Amino acids with acidic side chains
Aspartic acid (Asp, D)
Glutamic acid (Glu, E)
Amino acids with uncharged polar side chains
Asparagine (Asn, N)
Glutamine (Gln, Q)
Serine (Ser, S)
Threonine (Thr, T)
Tyrosine (Tyr, Y)
Amino acids with nonpolar side chains
Glycine (Gly, G)
Alanine (Ala, A)
Valine (Val, V)
Leucine (Leu, L)
Isoleucine (Ile, I)
Phenylalanine (Phe, F)
Methionine (Met, M)
Cysteine (Cys, C)
Proline (Pro, P)
Tryptophan (Trp, W)

PEPTIDE BOND FORMATION



PROTEINS

Polypeptide

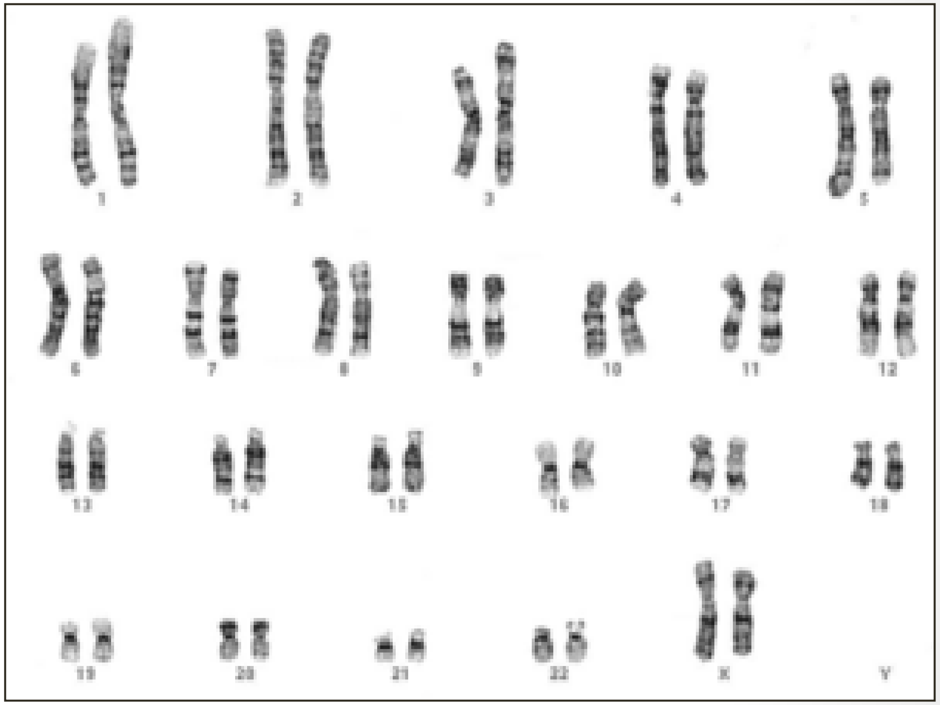
- N-terminus (-NH_2)
- C-terminus (-COOH)

Protein structure

- Primary
  - Linear sequence of amino acids
- Secondary
  - Polypeptide chains arrangements in helices and non-helices
- Tertiary
  - Arrangement of helices into 3D structure
- Quaternary
  - Multiple polypeptide fold together

Chromosome

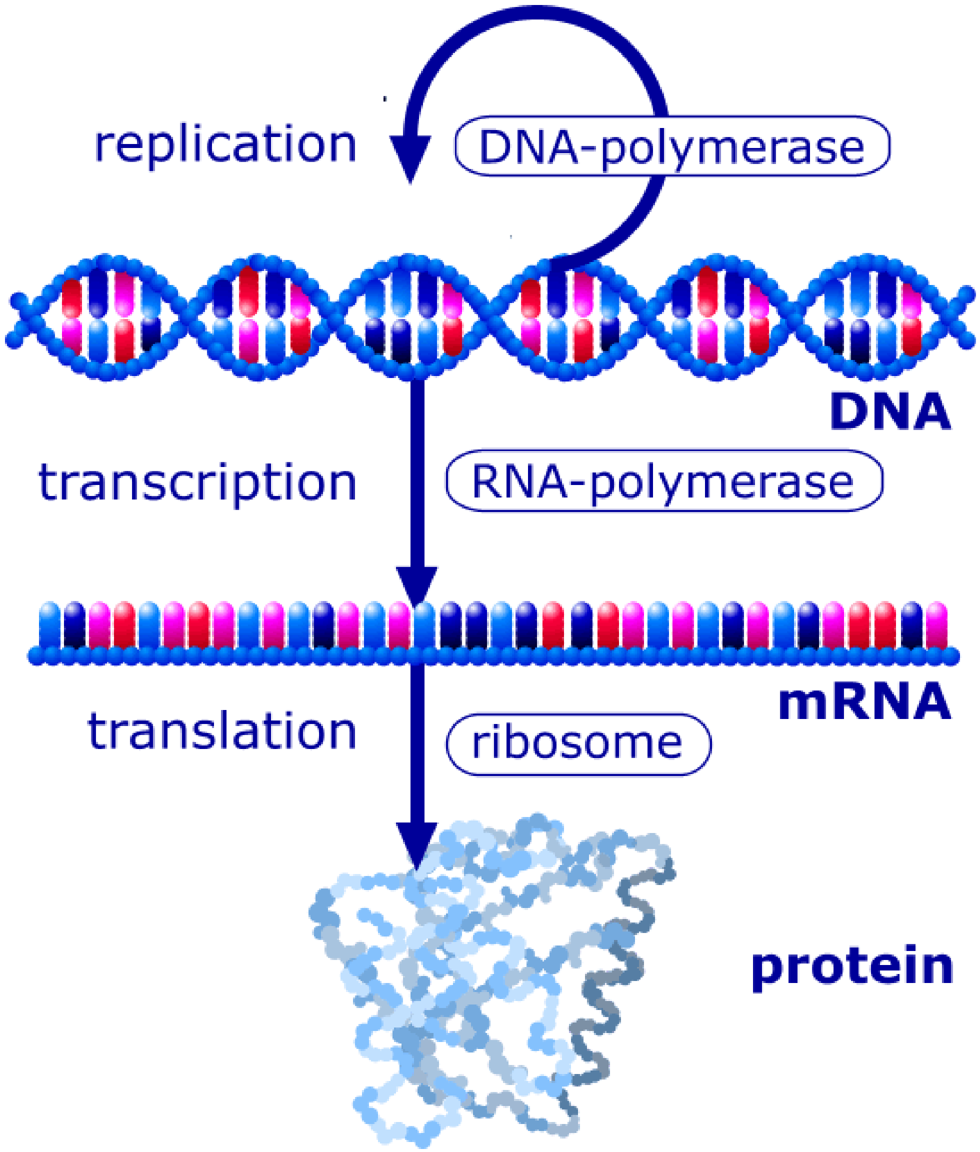
Species	# of chromosomes
Fruit Fly	8
Human	46
Rye (Roggen)	14
Ape	48
Guinea Pig	16
Sheep	54
Dove (Taube)	16
Horse	64
edible snail	24
Chicken	78
Earthworm	32
Carp (Karpfen)	104
Pig	40
Butterflies	~380
Wheat	42
Fern (Farn)	~1200



Karyogram of human female

http://www.answers.com/topic/human-karyogram-png

Gene Expression



Gene Expression - Transcription

Messenger RNA (mRNA)

Messenger RNA is RNA that carries information from DNA to the ribosome sites of protein synthesis in the cell.

Once mRNA has been transcribed from DNA, it is exported from the nucleus into the cytoplasm, where it is bound to ribosomes and translated into protein.

Non-coding RNA or "RNA genes"

RNA genes (sometimes referred to as non-coding RNA or small RNA) are genes that encode RNA that is not translated into a protein.

The most prominent examples of RNA genes are transfer RNA (tRNA) and ribosomal RNA (rRNA), both of which are involved in the process of translation.

Gene Expression - Translation

- The **genetic code** is made up of three letter 'words' (termed a codon) formed from a sequence of three nucleotides (e.g.. ACT, CAG, TTT).

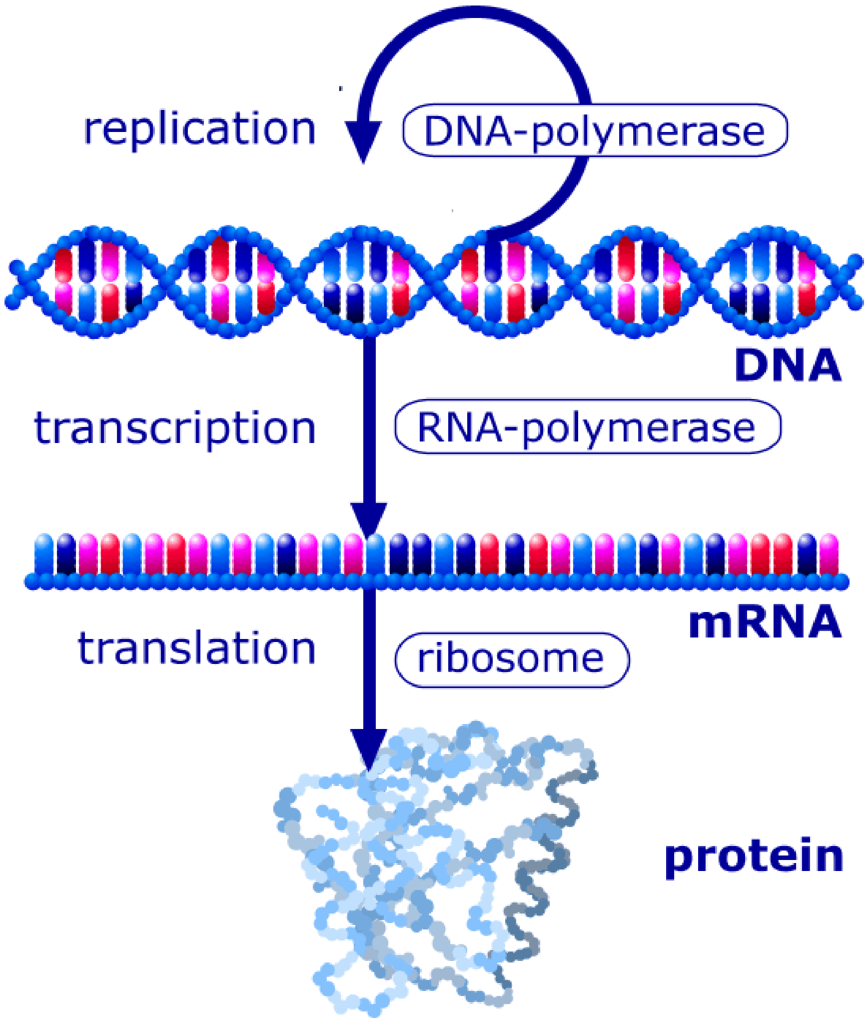
- These ***codons*** can then be translated with messenger RNA and then transfer RNA, with a codon corresponding to a particular amino acid.

- Since there are 64 possible codons, most amino acids have more than one possible codon.

- There are also three 'stop' or 'nonsense' codons signifying the end of the coding region.

Name	1-Letter Nickname	Triplet
Glycine	G	GGT,GGC,GGA,GGG
Alanine	A	GCT,GCC,GCA,GCG
Valine	V	GTT,GTC,GTA,GTG
Leucine	L	TTG,TTA,CTT,CTC,CTA,CTG
Isoleucine	I	ATT,ATC,ATA
Histidine	H	CAT,CAC
Serine	S	TCT,TCC,TCA,TCG,AGT,AGC
Threonine	T	ACT,ACC,ACA,ACG
Cysteine	C	TGT,TGC
Methionine	M	ATG
Glutamic Acid	E	GAA,GAG
Aspartic Acid	D	GAT,GAC,AAT,AAC
Lysine	K	AAA,AAG
Arginine	R	CGT,CGC,CGA,CGG,AGA,AGG
Asparagine	N	AAT,AAC
Glutamine	Q	CAA,CAG
Phenylalanine	F	TTT,TTC
Tyrosine	Y	TAT,TAC
Tryptophan	W	TGG
Proline	P	CCT,CCC,CCA,CCG
Terminator (stop)	*	TAA,TAG,TGA

A gene codes for a protein



CCTGAGCCAACTATTGATGAA

GGACUCGGUUGAUAACUACUU

PEPTIDE

DNA REPLICATION

What is DNA?

- DNA is double stranded helix, it is the genetic code and It determines the physical characteristics from hair color to what we are allergic to.
- Human DNA codes is 20 amino acids which are building the blocks of life.

What Is DNA Replication ?

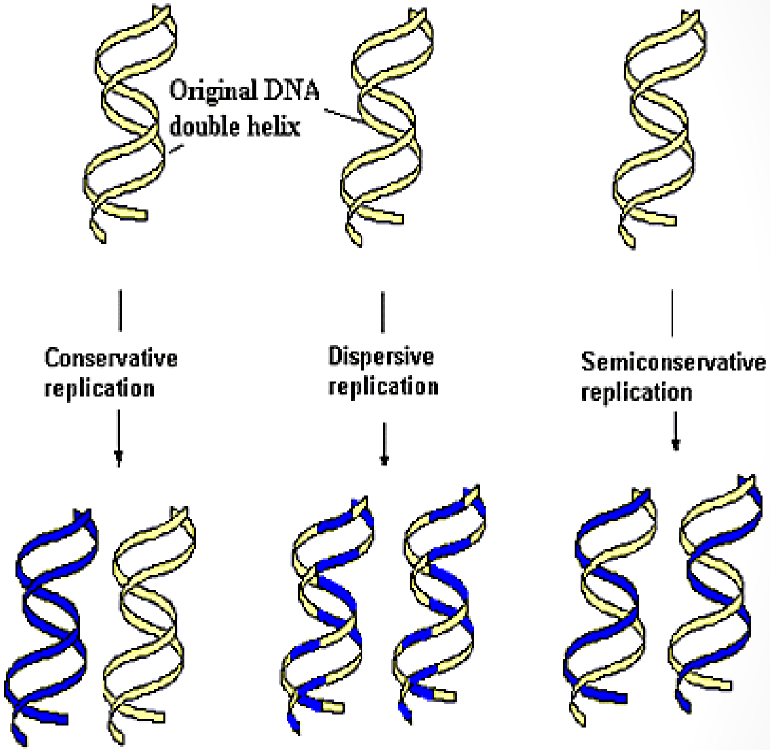
- Replication is the process in which the DNA within a cell makes an exact copy of itself. It is a Processing of synthesis in which each strand can serve as template for the new synthesis complementary strand and the entire ds-DNA is copied to produce a second, identical double helix.

The Possible Models of DNA Replication

1- Conservative- would leave the original strand intact and copy it.

2- Dispersive-would produce two DNA molecule with sections of both old and new along each strand.

3- Semiconservative –would produce DNA molecule with both one old strand and one new strand.



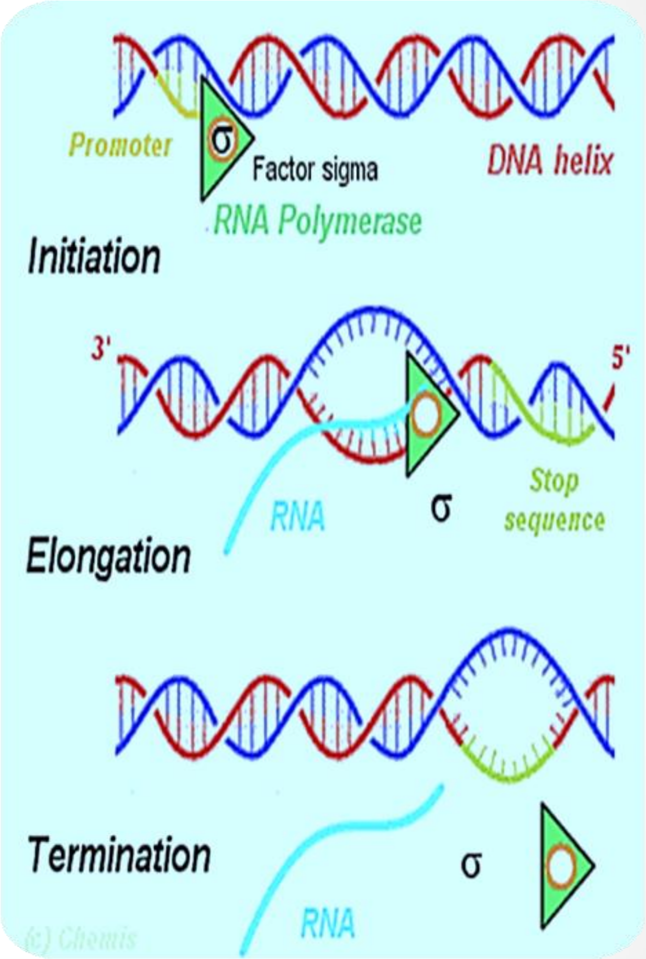
Possible Models of DNA Replication

Enzymes participate in DNA replication

Protein	Functions
DNA Polymerase	DNA synthesis, 5′ → 3′ direction of new strand; requires RNA primer
Helicase	Unwinds double-strand DNA at the replication fork
Primase	Synthesis of short RNA sequence; primer for DNA synthesis
Single-strand binding protein	Binds to single-strand DNA to keep strands from base pairing
DNA ligase	Joins DNA fragments during DNA replication

REPLICATION STAGES

1. Initiation
2. Elongation or polymerization
3. Termination



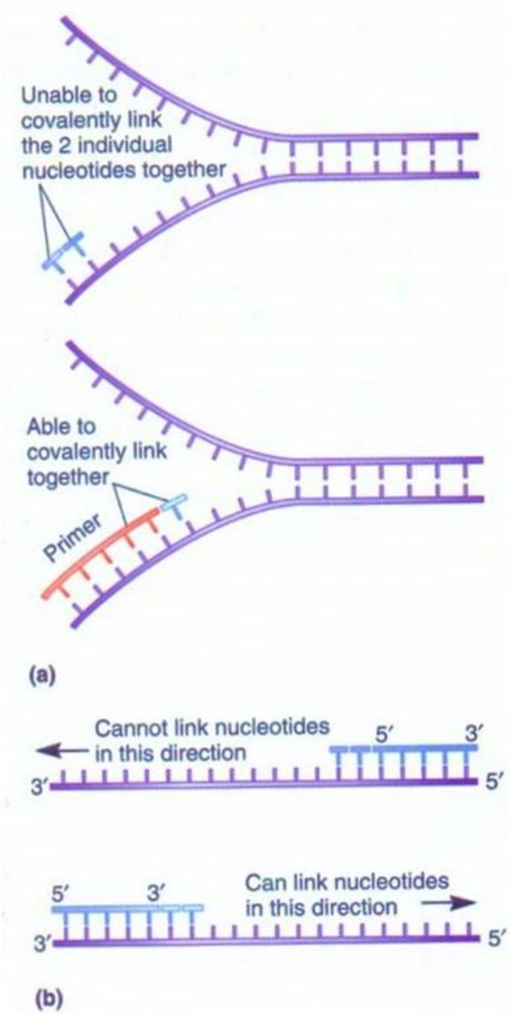
Condition of replication

- The enzymes that catalyze addition of the nucleotides are called DNA polymerases. All DNA polymerases synthesize new DNA in the 5' → 3' direction and all these enzymes can only add a nucleotide onto a preexisting 3'-OH group.

- Therefore, for a new chain to be started, there must be a primer, which is a site at which the DNA polymerase can attach the first nucleotide. In most cases this primer is a short stretch of RNA.

- When the double helix is opened up at the beginning of replication, an RNA-polymers (RNA-primer) enzyme is acting first resulting as formation of this RNA primer.
- A specific RNA-polymerizing enzyme, participates in primer synthesis by laying down a short stretch of RNA.
- At the growing end of this RNA primer is a 3'-OH group to which DNA polymerase can add the first deoxyribonucleotide.
- The continued extension of the molecule occurs as DNA rather than RNA.
- The newly synthesized molecule has a structure like that shown in. Then the primer eventually can removed.

Polymerization Features



Viruses as a model for replication
What is a virus?

Viruses are uniquely different from the many uni-cellular micro-organisms you have studied so far. Protozoa, yeasts, bacteria, mycoplasmas, rickettsia and chlamydia are all living organisms with the features in common as they are all cells, they store their genetic information as DNA/ RNA Within their cell, they contain all the organelles necessary for producing energy and synthesizing proteins, carbohydrates, cell wall structures etc .Replicate by means of binary fission.

Viruses properties.

- They are not cells.
- They are very simple structures consisting essentially of a nucleic acid genome, protected by a shell of protein.
- They are metabolically inert and can only replicate once they are inside a host cell
- .The genome consists of only one type of nucleic acid: either RNA or DNA.
- Most DNA viruses are double stranded and most RNA viruses have a single stranded (ss) genome. However, ssRNA genome may be either positive sense (this means that it can be used as mRNA to make proteins) or negative sense. Negative sense RNA is complimentary to mRNA, in other words, it has to be copied into mRNA.
- The viral genome codes only for the few proteins necessary for replication: some proteins are non-structural e.g. polymerase and some are structural, i.e. they form part of the virion structure.They have no organelles.
- They are very small, sizes range from 20 to 200 nm.

Terminology

Virion = virus particle
Capsid or Core = protein shell which surrounds and protects the genome. It is built up of multiple (identical) protein sub-units called capsomers. Capsids are either icosahedral or tubular in shape
Nucleocapsid = genome plus capsid
Envelope = lipid membrane which surrounds some viruses. It is derived from the plasma membrane of the host cell.
Peplomers = proteins found in the envelope of the virion. They are usually glycosylated and are thus more commonly known as glycoproteins.

General concept of viral replication

Viruses are the ultimate parasite they are totally dependent on a host cell to replicate (make more copies of itself). While the sequence of events varies somewhat from virus to virus, the general strategy of replication is similar:

Adsorption: The surface of the virion contains structures that interact with molecules (receptors) on the surface of the host cell. This is usually a passive reaction (not requiring energy), but highly specific. It is the specificity of the reaction between viral protein and host receptor that defines and limits the host species and type of cell that can be infected by a particular virus. Damage to the binding sites on the virion or blocking by specific antibodies (neutralization) can render virions non-infectious.

Uptake: The process whereby the virion enters the cell. It occurs either as a result of fusion of the viral envelope with the plasma membrane of the cell or else by means of endocytosis.

Uncoating: Once inside the cell, the protein coat of the virion dissociates and the viral genome is released into the cytoplasm.

Early phase: Once the genome is exposed, transcription of viral mRNA and translation of a number of non-structural ("early") proteins takes place. The function of these is to replicate the viral genome.

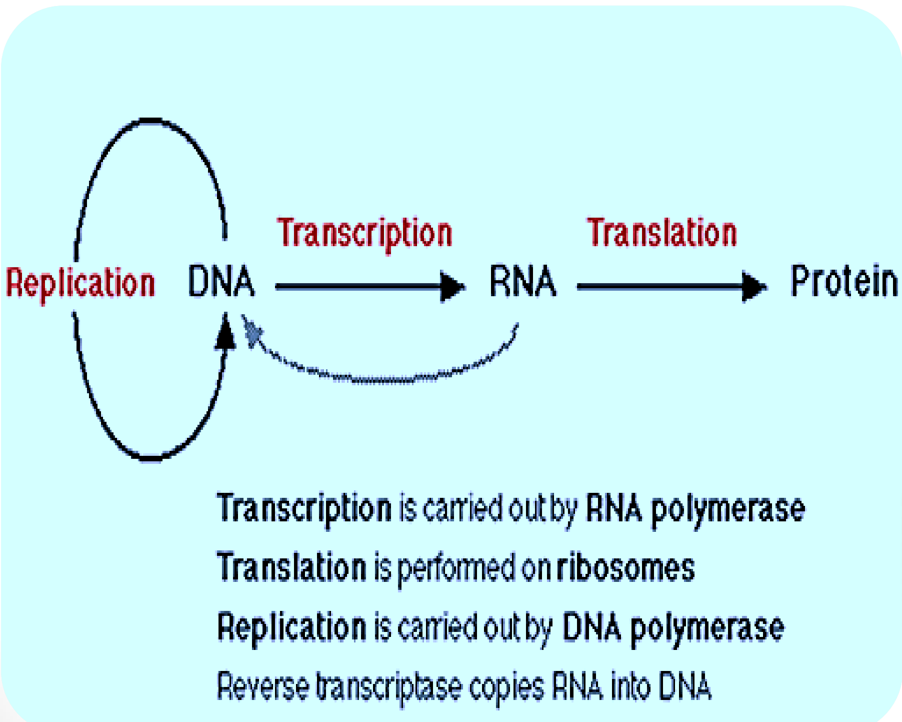
Genome replication: Multiple copies of the viral genome are synthesized by a viral polymerase (one of the "early" proteins).

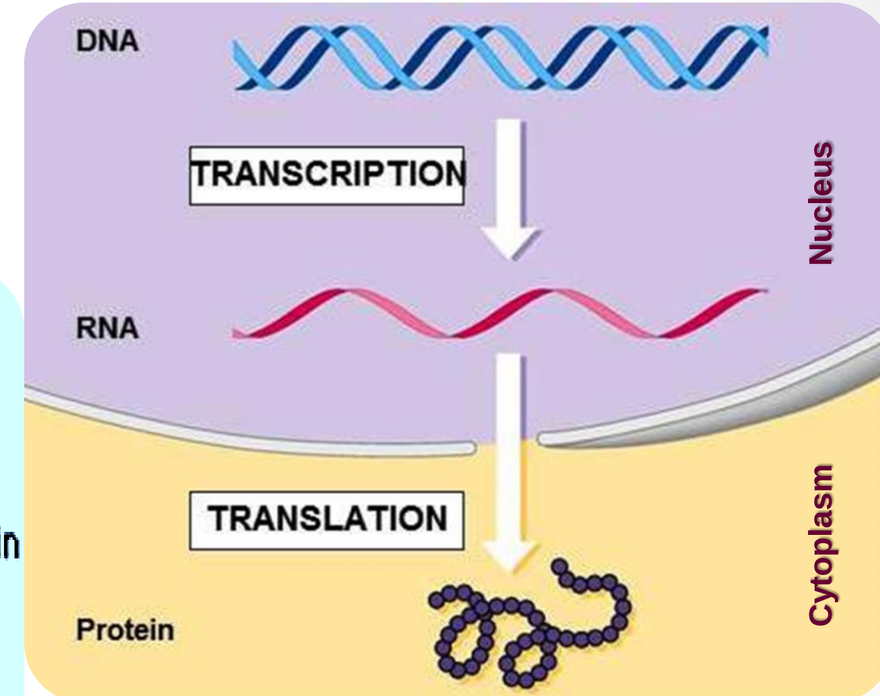
Late phase: Transcription and translation of viral mRNA and synthesis of the structural ("late") proteins which are needed to make new virions .

Assembly of new virions: Assembly of new viral capsids takes place either in the nucleus (e.g. herpes viruses) or in the cytoplasm (e.g. poliovirus) of the cell, or sometimes, just beneath the cell surface (e.g. budding viruses such as influenza). The proteins self-assemble and a genome enters each new capsid.

Release of progeny virions: Release of new infectious virions is the final stage of replication. This may occur either by budding from plasma membrane or else by disintegration (lysis) of the infected cell. Some viruses use the secretory pathway to exit the cell: virus particles enclosed in Golgi-derived vesicles are released to the outside of the cell when a transport vesicle fuses with the cell membrane.

Virus as module for replication

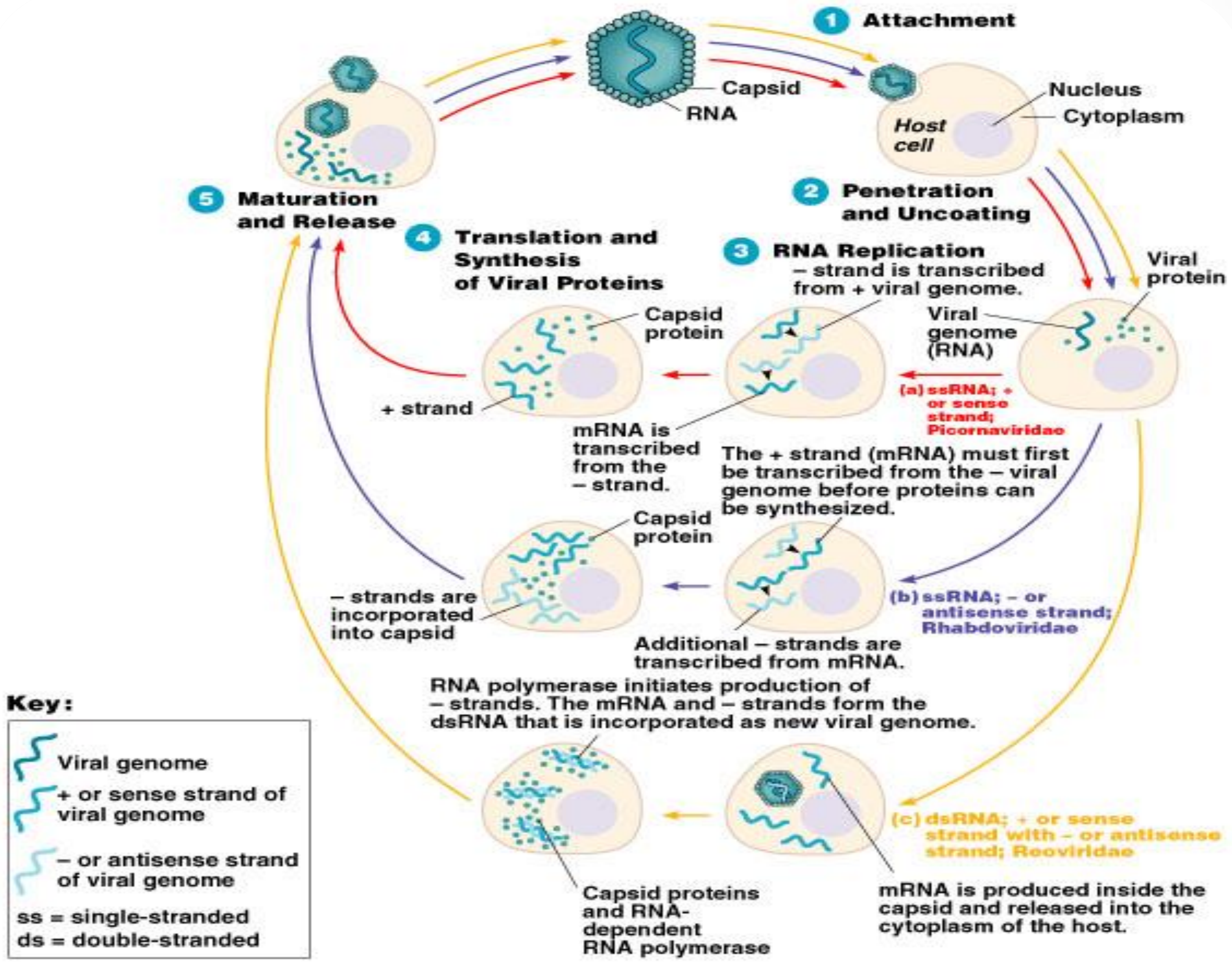




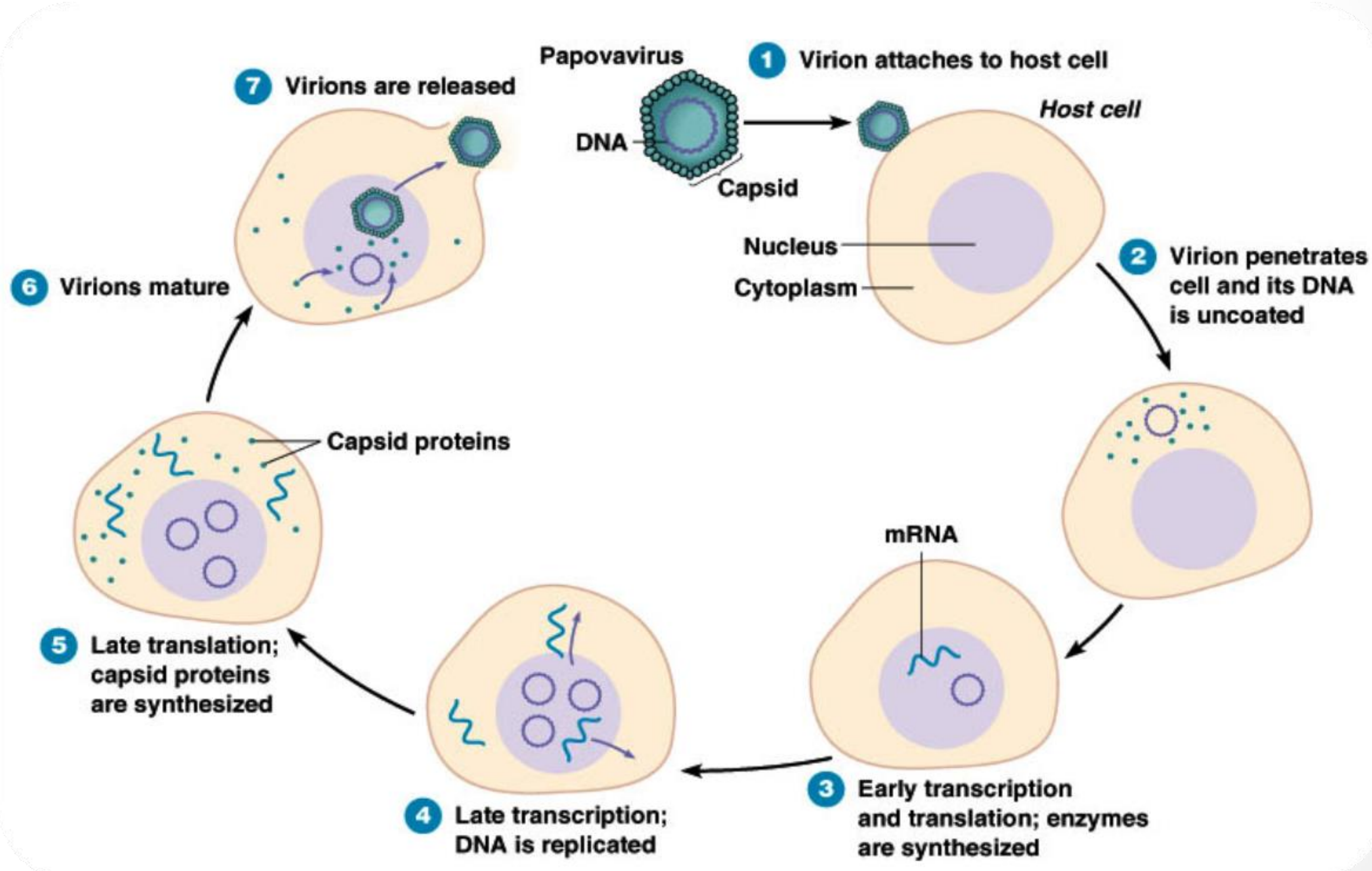
Types and function of RNA

Type of RNA	Functions in	Function
Messenger RNA (mRNA) 	Nucleus, migrates to ribosomes in cytoplasm	Carries DNA sequence information to ribosomes
Transfer RNA (tRNA) 	Cytoplasm	Provides linkage between mRNA and amino acids; transfers amino acids to ribosomes
Ribosomal RNA (rRNA) 	Cytoplasm	Structural component of ribosomes

RNA Replication



DNA Replication

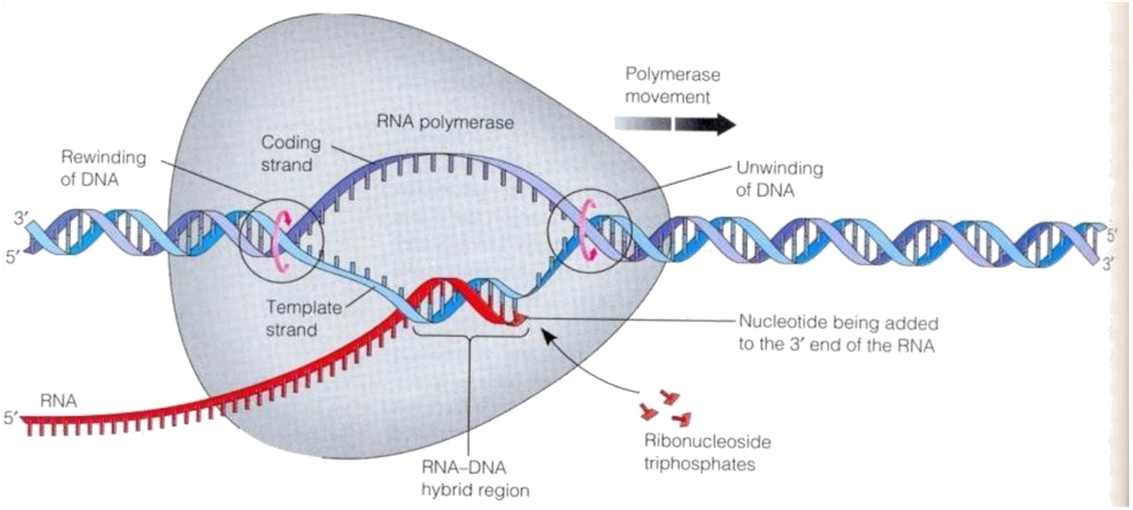


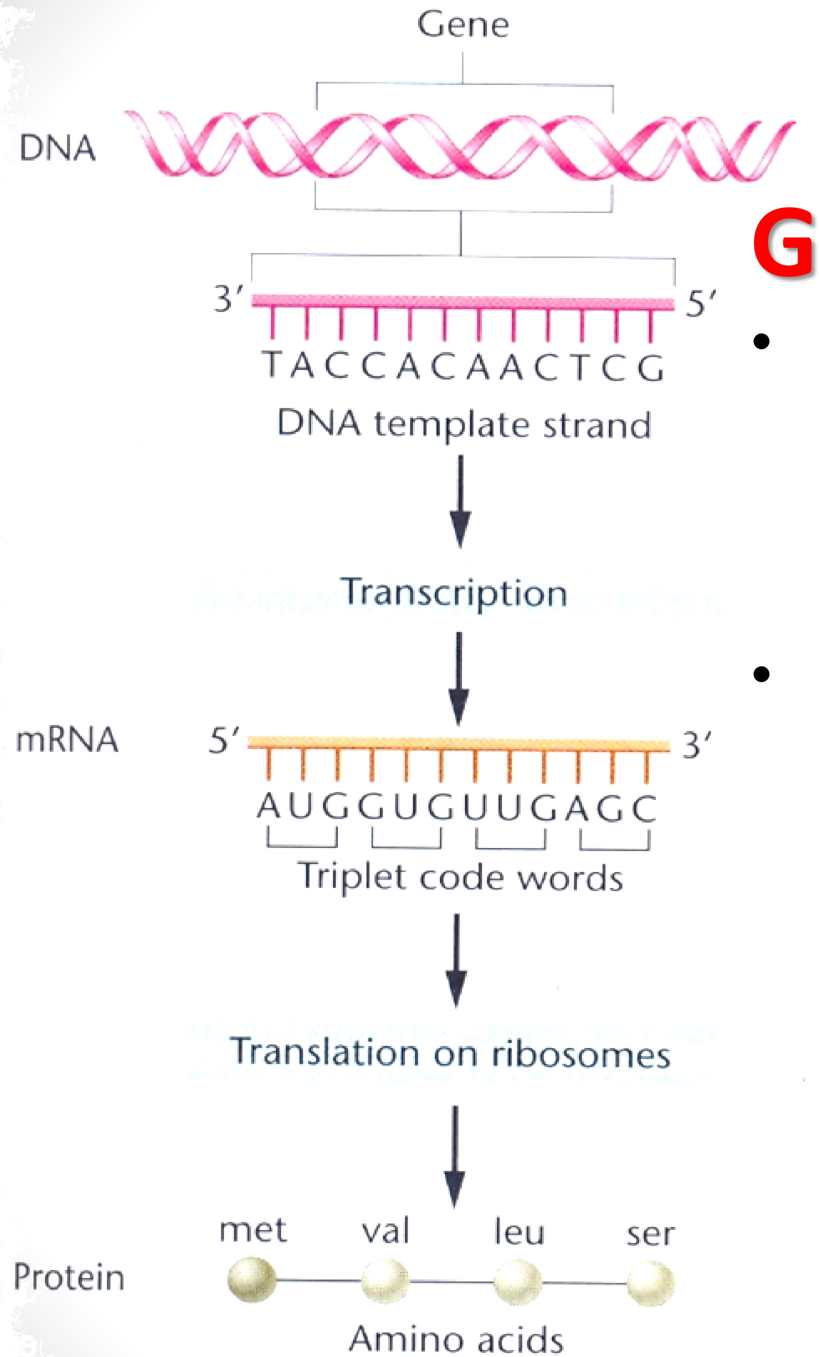
Production of mRNA by various viral types

Genome		Transcription Enzyme (for mRNA production)	Initial event in cell	Example
RNA [RNA viruses with RNA intermediates]	(+) single stranded	none	Translation*	• Picornaviruses
	(-) single stranded	RNA pol II (viral)	Transcription	• Rhabdoviruses : Rabies
	double stranded	RNA pol II (viral)	Transcription	• Reoviruses
DNA [DNA viruses with RNA intermediates]	single stranded	RNA pol II (host)	Primary Transcription	• Parvoviridae : Human Parvovirus B19
	Double stranded	RNA pol II (host)	Primary Transcription	• Hepadnaviruses: Hepatitis B virus • Herpesviruses: Varicella Zoster virus (Chickenpox, measles) • Adenoviruses
Retroviruses [RNA viruses with DNA intermediates]	single stranded RNA	Reverse Transcriptase (RT) i.e. DNA-dependent RNA polymerase	Reverse Transcription	• Human Immunodeficiency virus (HIV) • Rous sarcoma virus (RSV)

RNA pol II : RNA polymerase II (same as RNA- dependent RNA polymerase)
*the initial event in the cell is translation; therefore no transcription enzyme is required. But for subsequent transcription of (-) strands for nucleic acid synthesis, RNA pol II is used.

GENE EXPRESSION TRANSCRIPTION





GENE EXPRESSION

- Two molecular cellular processes:
  - Transcription
  - Translation
- Gene is transcribed into mRNA and translated into protein

DNA STRAND TRANSCRIPTION
Coding and non coding strand

5’-A T C G T C A G T G A T C A -3’

3’-T A G C A G T C A C T A G T -5’

DNA **CODING** strand	5’-A T C G T C A G T G A T C A -3’
mRNA	5’-**AUCGUCAGUGAUCA**-3’
DNA **TEMPLATE** strand	3’-T A G C A G T C A C T A G T -5

TRANSCRIPTION

- First step in **gene expression**.
- A process of **RNA** synthesis from **DNA** template sequence of a **gene**.
- DNA strand that is transcribed called **coding** strand and other strand is used as **template** to synthesize RNA transcript.
- Transcription involves DNA-protein interactions.
- Proteins have a role in transcription.
  - **RNA polymerase** (**RNA P**).
  - Transcription factors (**T.Fs.**).
    - General transcription factors
    - Tissue specific transcription factors
  - Termination factor

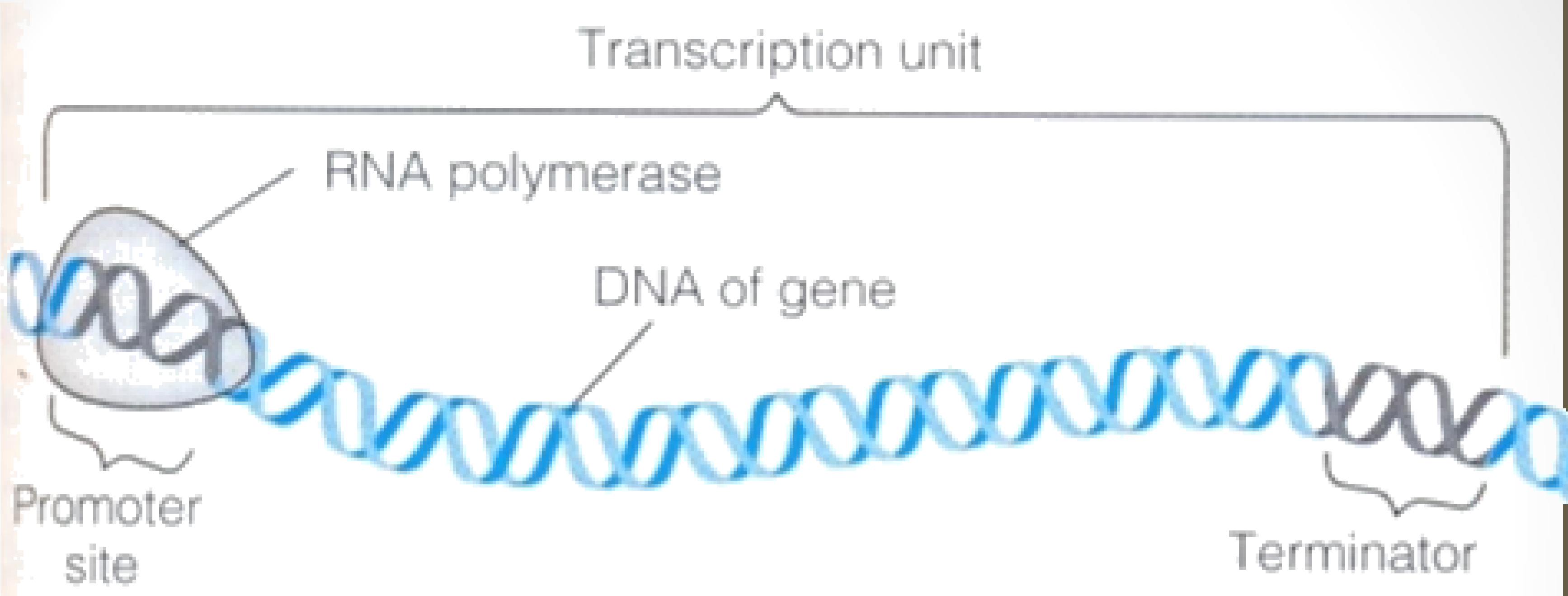
TRANSCRIPTION STAGES

- Transcription occurs in 3 stages:
    1. Initiation (Recognition)
    2. Elongation (**RNA** transcript synthesis)
    3. Termination
- Gene contains sequences and sites that determines
  beginning and end of transcription
    - **Promoter**
    - **Transcription start site**
    - **Terminator**

Transcription

1) Prokaryotic
2) Eukaryotic

TRANSCRIPTIONAL INITIATION



PROMOTER

RNA polymerase binds to promoter region to start transcription.

PROKARYOTIC TRANSCRIPTION

1) INITIATION STAGE

- A recognition step.
- Initiated by binding of RNA polymerase (RNA P) at **promoter** sequence to start transcription.

Promoter

- Promotes gene expression
- Determines transcription orientation and initiation site

  Transcription start site

  - Base position where transcription starts
- DNA sequence located at the 5’ end (**upstream**) region from transcription start site (**Gene**) and play an important role in initiation RNA transcription

PROKARYOTES PROMOTER

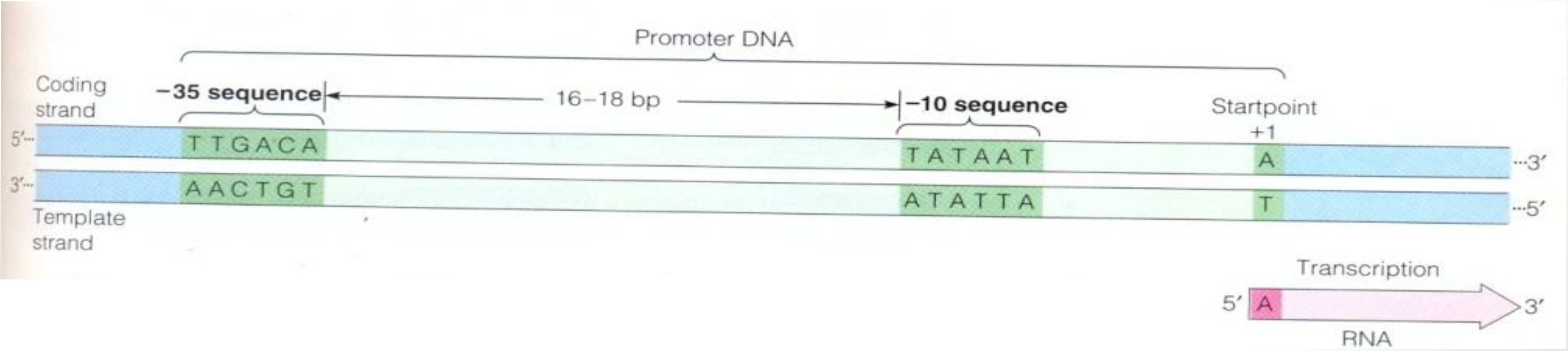
- Promoter base sequence is numbered relative to the **transcription start site**.
- Transcription start site is designated (+1)
- **Upstream** of base of transcription site is denoted (-1).

PROKARYOTE TRANSCRIPTION

Promoter

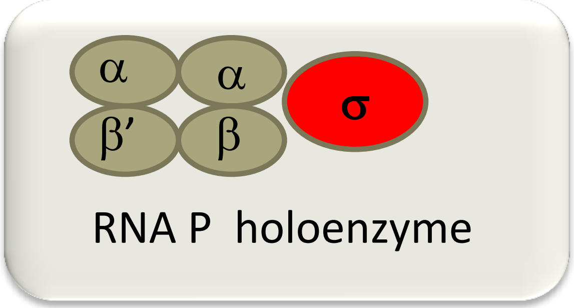
-35 sequence TTGACA

-10 sequence TATAAT



PROKARYOTIC RNA POLYMERASE

- Enzyme catalyzes RNA synthesis
- **RNA** polymerase consists of 5 subunits:
  - ➢ 2α, β, β', σ
  - ➢ Active form of enzyme is (α_2 ββ'σ)called **holoenzyme**



Subunit functions

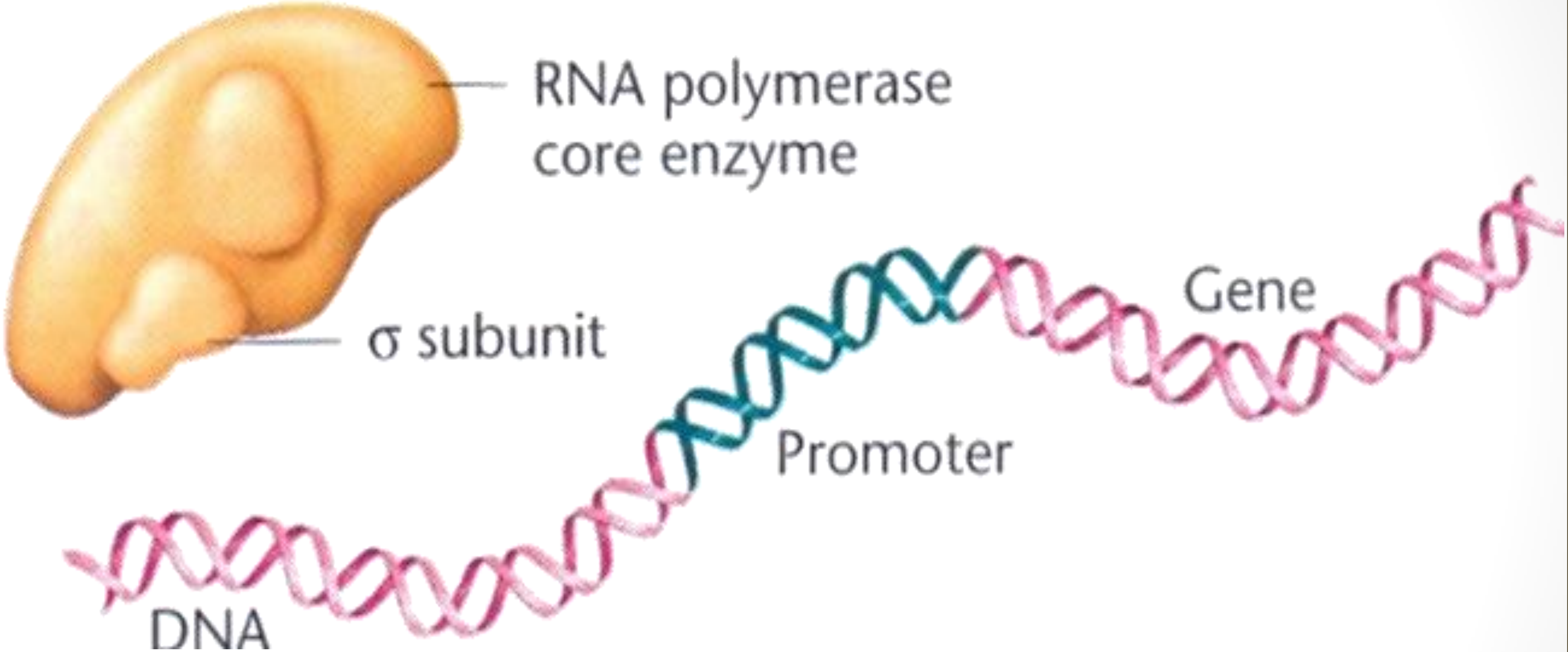
- **σ**-(Sigma subunit of RNA polymerase)
  - ➢ Recognizes **promoter** and initiates transcription
- α-Binds polymerase to DNA
- **ββ'**- Synthesis **RNA** (transcription)

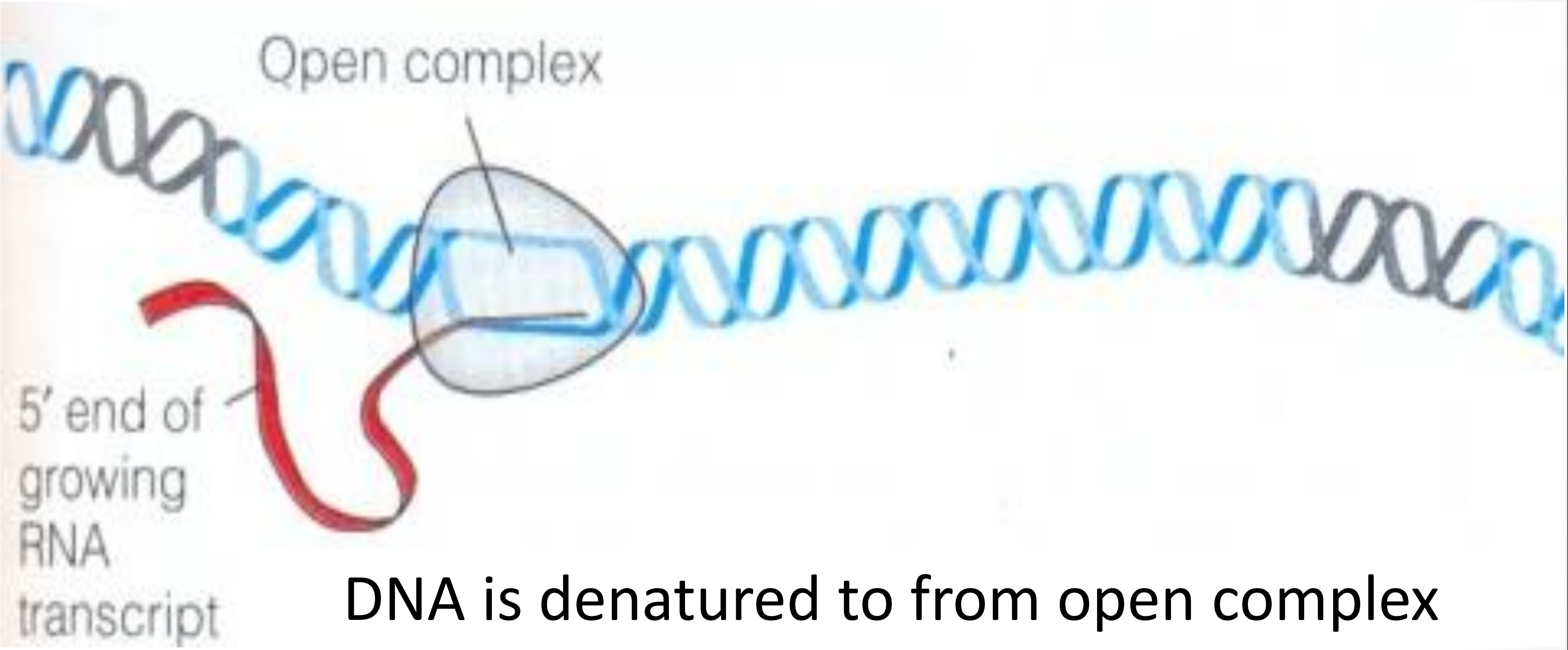
PROKARYOTIC

1) TRANSCRIPTION INITIATION

- Following assembly of RNA polymerase holoenzyme into 5 subunits.
- Sigma factor recognize **promoter** elements, binds to DNA to form closed complex
- DNA double helix is unwinded into an open complex
- Transcription is **initiated** and **sigma** factor dissociate from the holoenzyme.
- RNA transcript synthesized from **5'-3'** direction

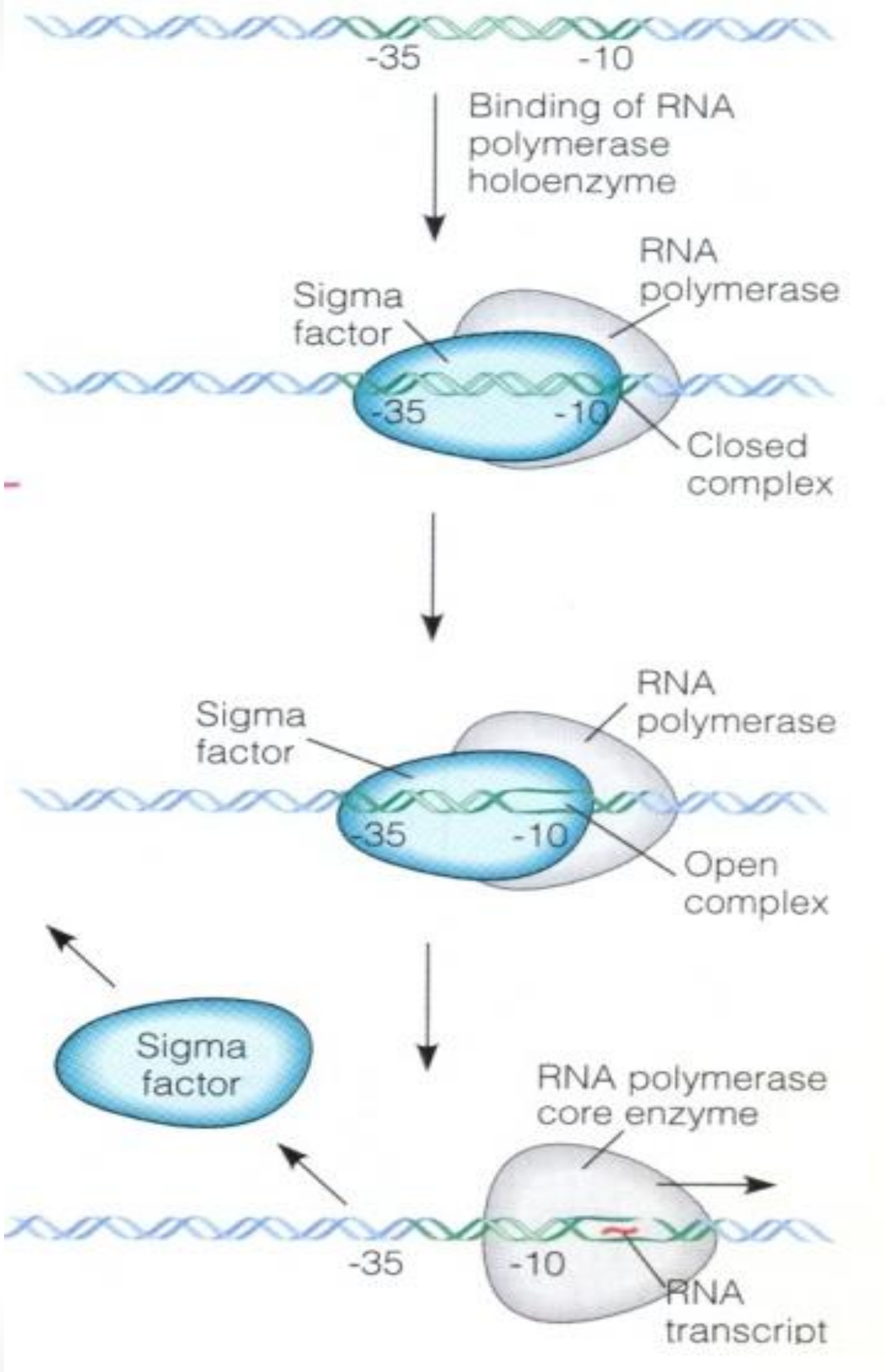
TRANSCRITPION INITIATION





DNA is denatured to from open complex

TRANSCRIPTION INITIATION



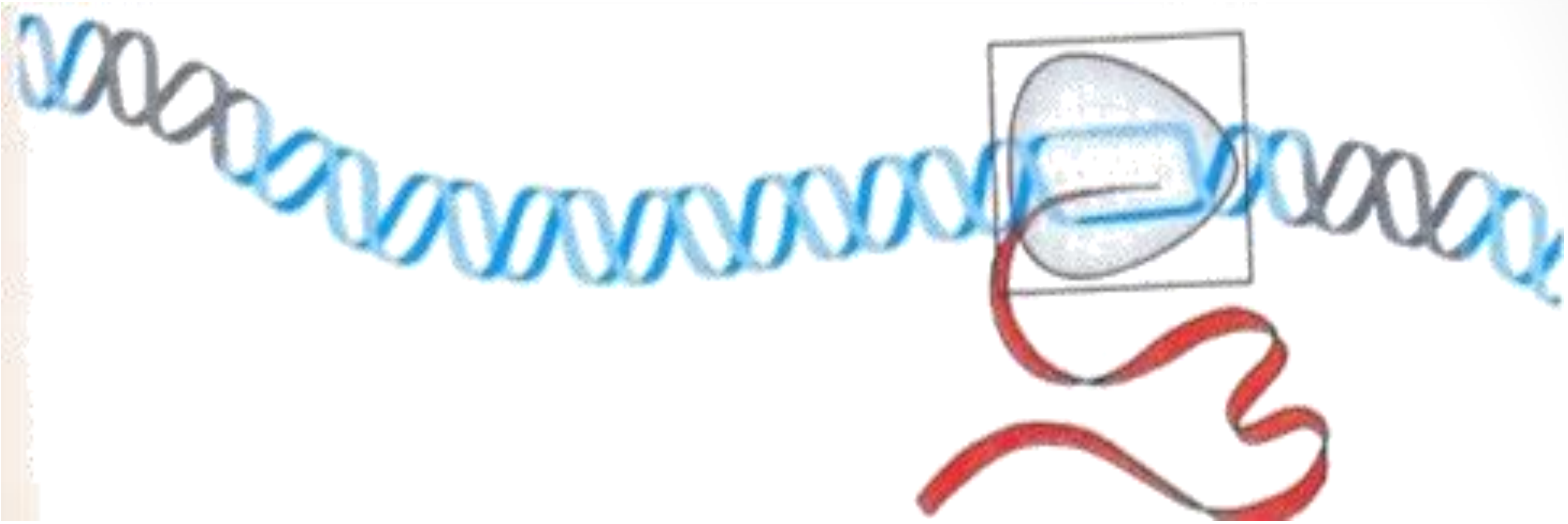
PROKARYOTIC TRASNCRIPTION

2) TRANSCRIPTION ELONGATION

- RNA polymerase moves along the DNA non coding **template** strand and catalyzing RNA synthesis by addition of NTPs to 3' end of newly synthesized RNA.

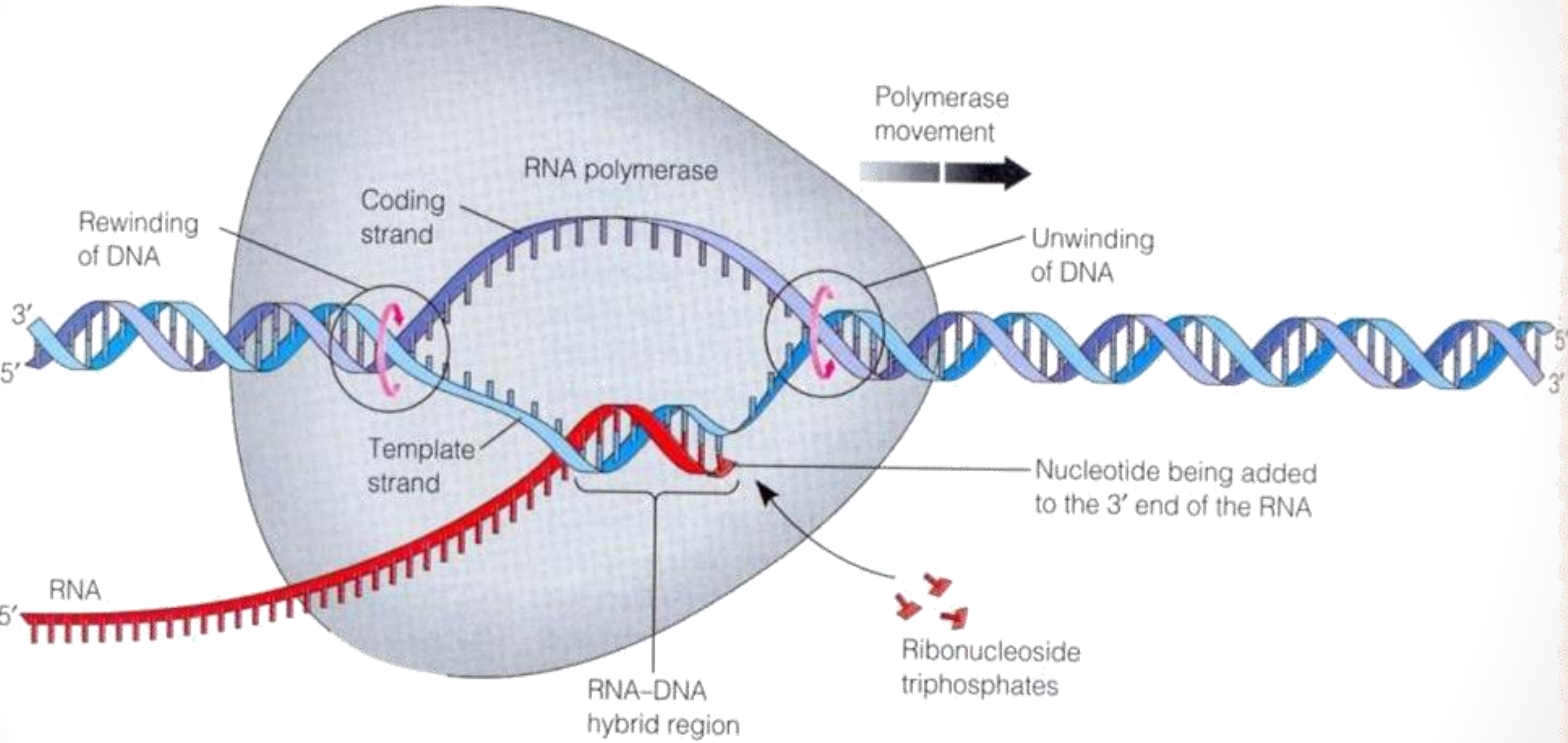
- RNA molecule synthesized complementary to the **template** strand and **U** replaces T.

TRANSCRIPTIONAL ELONGATION



RNA polymerase slides along the DNA in an open complex to catalyze RNA synthesis

TRANSCRIPTION ELONGATION

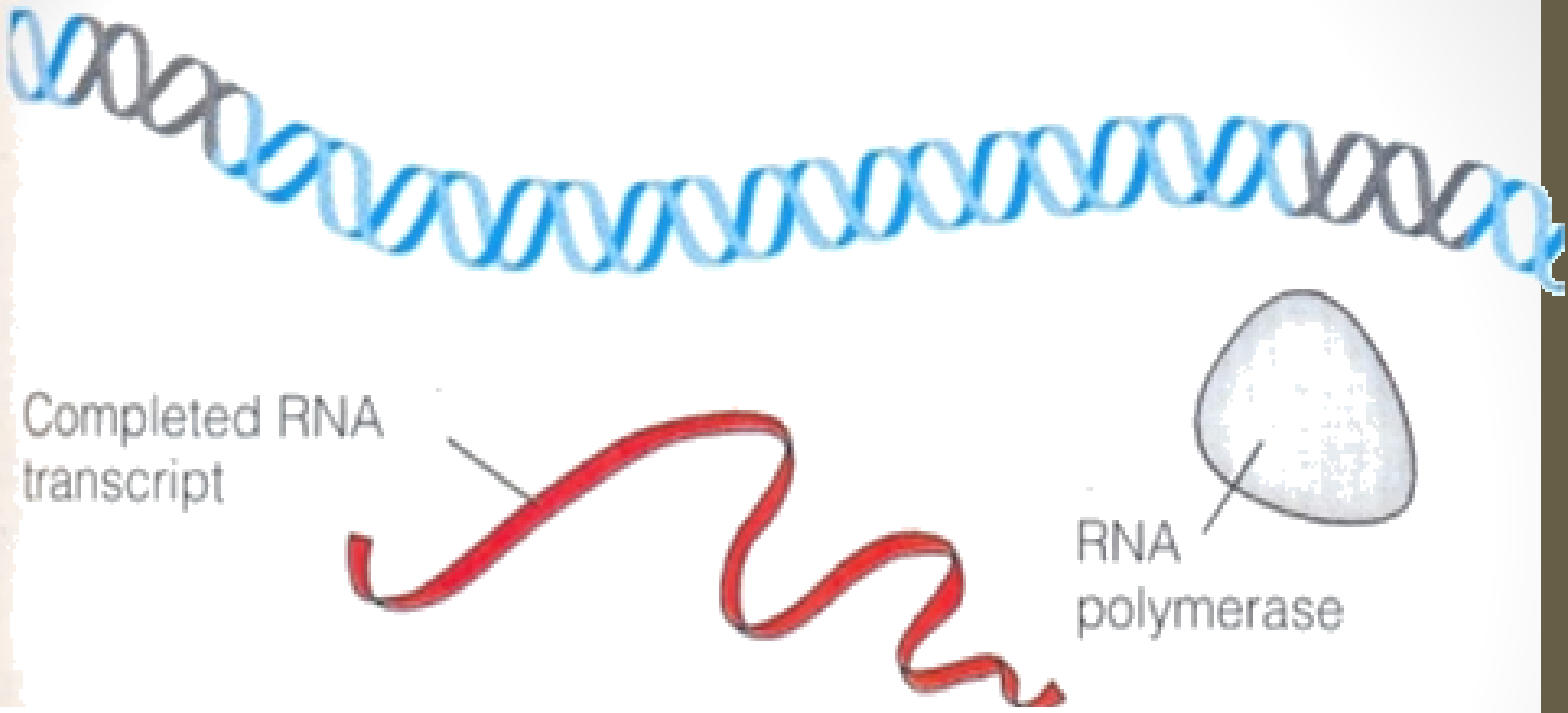


PROKARYOTIC

TRANSCRIPTIONAL TERMINATION

- RNA polymerase transcribes the gene sequence until encountering termination signal (specific nucleotide sequence).
- Termination factor rho protein interacts with RNA transcripts and cause transcription termination.
- RNA transcript and RNA core polymerase enzyme released from DNA template.

TRANSCRIPTIONAL TERMINATION



Termination sequence at the end of the gene causes the RNA polymerase and RNA transcript to dissociate from DNA.

EUKARYOTIC TRANSCRIPTION

PROMOTER

TRANSCRIPTION FACTORS

- Proteins interacts to recruit RNA polymerase to core promoter region initiate and regulate transcription.

CATEGORIES OF T.FS

1-General transcription factor

2-Specific transcription factor

HISTONES

- Eukaryotic genome contains histones that has to be either loosen or displaced to permit recognition.

RNA polymerases

- 3 nuclear eukaryotic RNA polymerase
  1) RNA Poly I (transcribes rRNA).
  2) RNA Poly II (transcribes mRNA)
  3) RNA Poly III (transcribes tRNA)

Transcription stages

- Initiation
- Elongation
- Termination

m-RNA MODIFICATIONS

- Eukaryotic mRNA is transcribed as precursor (**pre-mRNA**) in the nucleus, processed and released into cytoplasm as **mature mRNA**

mRNA processing

1. Trimming
2. Capping
3. Tailing
4. Splicing

RNA MODIFICATION