Module 1: Introduction and Replication of DNA

B. E. BIOTECHNOLOGY
Outcome Based Education (OBE) and Choice Based Credit System (CBCS)
SEMESTER - IV

MOLECULAR BIOLOGY

Course Code	**18BT42**	CIE Marks	40
Teaching Hours/Week (L:T:P)	(4:0:0)	SEE Marks	60
Credits	04	Exam Hours	03

Course Learning Objectives:

- To familiarize with the cell and molecular biology of both Prokaryotes and Eukaryotes.
- To acquire basic fundamental knowledge and explore skills in molecular biology
- To become aware of the complexity and harmony of the cells.
- To learn the molecular mechanism of DNA replication, repair, transcription, protein synthesis and gene regulation in various organisms.

Module-1

INTRODUCTION & REPLICATION OF DNA:
Chromosomal theory of heredity, genes and their location. Genetic code, Information flow in biological systems: central dogma, updated central dogma. Structures and forms of nucleic acids – DNA and RNA. Reversible denaturation and hyperchromic effect. Tertiary structure of DNA: DNA supercoiling. Replication of DNA, structure and function of DNA polymerases, models of replications in prokaryotes, mechanism of DNA replication and enzymes involved. Proteomics of DNA replication, Fidelity of DNA replication, Inhibitors of DNA replication

Module-2

TRANSCRIPTION:
Structure and function of RNA polymerases (prokaryotes & eukaryotes), mechanism of transcription in prokaryotes and eukaryotes, transcription factors, Fidelity of RNA synthesis, post-transcriptional processing (RNA editing, siRNA, splicing, poly A tail and 5' capping), Ribozymes, transcription inhibitors.

Module-3

TRANSLATION:
Mechanism of translation, activation of amino acid initiation, elongation and termination of protein synthesis. Post-translational modification and protein targeting, protein splicing. Differences between prokaryotic and eukaryotic protein synthesis, inhibitors of translation.

Module-4
GENE EXPRESSION IN PROKARYOTES & EUKARYOTES: Regulation of gene expression in prokaryotes: Operon model, gal, lac, trp Operons; positive versus negative regulation. Regulation of eukaryotic gene expression, transcriptional control, homeobox in the control of developments in insects and vertebrates.
Module-5
GENETIC RECOMBINATION, MUTATION & GENE MAPPING: Genetic recombination in bacteria and viruses, site specific recombination, transposons and insertion sequences; Retroviruses. DNA damage & Repair, Mutation, Role of recombination and transposition in evolution; gene mapping techniques.
Course Outcomes: At the end of the course the student will be able to: • Explain replication, transcription and translation processes with underlying differences in prokaryotic and eukaryotic systems. • Elaborate importance of genetic recombination with special reference to bacterial system. • Outline DNA damage and repair mechanisms

INTRODUCTION

Molecular Biology :

- The field of **biology** that studies the composition, structure and interactions of cellular **molecules** – such as nucleic acids and proteins – that carry out the **biological** processes essential for the cell's functions and maintenance.

Cell :

- The basic structural, functional, and biological unit of all known organisms.
- Often referred to as the "building blocks of life".

Biomolecules:

- **Biological molecule**, any of numerous substances that are produced by cells and living organisms.
- Biomolecules have a wide range of sizes and structures and perform a vast array of functions.
- The four major types of biomolecules are carbohydrates, lipids, nucleic acids, and proteins.

1. **Carbohydrates**,

   Made up primarily of molecules containing atoms of carbon, hydrogen, and oxygen,

   Are essential energy sources and structural components of all life, and they are among the most abundant biomolecules on Earth.

   They are built from four types of sugar units—monosaccharides disaccharides, oligosaccharides, and polysaccharides.

2. **Lipids:**

Fulfill a variety of roles, including serving as a source of stored energy and acting as chemical messengers.

They also form membranes, which separate cells from their environments and compartmentalize the cell interior, giving rise to organelles, such as the nucleus and the mitochondrion, in higher (more complex) organisms

3. **Proteins:**

Are major structural elements of cells.

They also serve as transporters, moving nutrients and other molecules in and out of cells, and as enzymes and catalysts for the vast majority of chemical reactions that take place in living organisms.

Proteins also form antibodies and hormones, and they influence gene activity..

4. **Nucleic Acids:**

Nucleic acids, namely DNA and RNA, have the unique function of storing an organism's genetic code

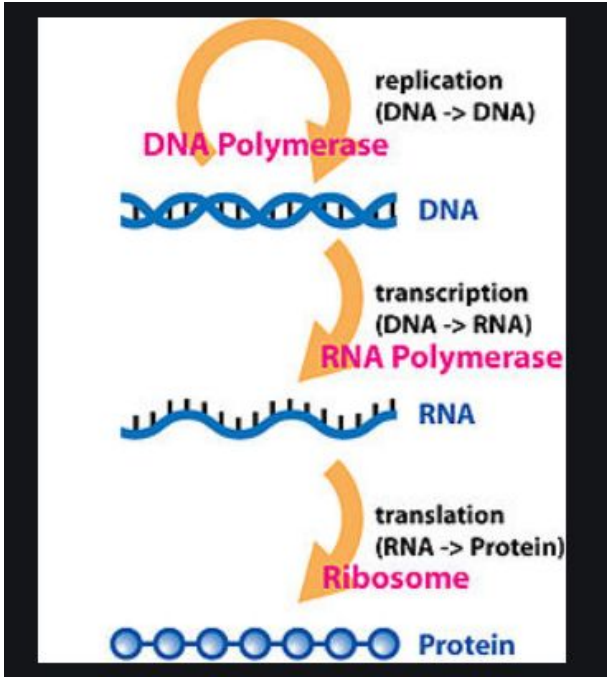
Involved in storage and transfer of genetic information

They provide biological diversity and specific characters to different organisms and also to cells

Information flow in biological systems: central dogma

- In molecular biology, central dogma illustrates the flow of genetic information from DNA to RNA to protein.
- It is defined as a process in which the information in DNA is converted into a functional product.
- It is suggested that the information present in a DNA is essential to make up all proteins and RNA acts as a messenger that carries information through the ribosomes.
- The **central dogma of molecular biology** is an explanation of the flow of genetic information within a biological system.
- It is often stated as "DNA makes RNA, and RNA makes protein"
- the transfer of information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible.

- Genetic information that is stored in DNA is finally expressed in the form of protein by a series of intermediate processes.
- **1st step in the expression of genetic information** - transfer of information present in DNA to RNA- **Transcription-** carried out by enzyme RNA polymerase
- **Next step**: RNA is transported to cytoplasm where the information is decoded and transferred to proteins-- process called as **Translation.** Site of translation is Ribosomes
- Thus the information written in the language of only 4 letters is finally translated into entirely different language with a complex alphabet of 20 amino acids.
- Transfer of information between two types of nucleic acids are reversible but once the information is transferred to protein then the transfer is irreversible.



Updated central dogma:

Special transfers of biological sequential information:

1. **Reverse transcription:** Transfer of information from RNA to DNA (the reverse of normal transcription).

   It is the process by which genetic information from RNA gets transcribed into new DNA.

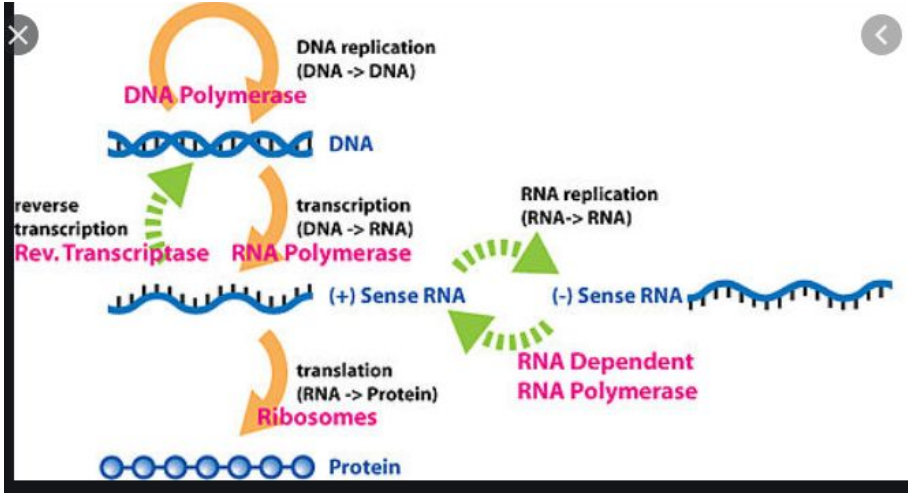
   Enzyme involved Reverse Transcriptase

2. **RNA replication**

   copying of one RNA to another.

   Many viruses replicate this way.

   Enzymes involved RNA-dependent RNA polymerases



3. **Direct translation from DNA to protein**
It has been demonstrated in a cell-free system (i.e. in a test tube), using extracts from *E. coli* that contained ribosomes, but not intact cells.

However, it was unclear whether this mechanism of translation corresponded specifically to the genetic code

Chromosomal theory of heredity

- The chromosomal theory of inheritance was given by **Walter Sutton and Theodor Boveri** in the early 1900s.
- It is the fundamental theory of genetics. According to this theory, **genes are the units of heredity and are found in the chromosomes.**
- Chromosomal Theory of Inheritance came into existence long after Mendelian genetics
- In 1902 and 1903, Sutton and Boveri published independent papers proposing what we now call the **chromosome theory of inheritance**.
- Theodor Boveri observed that proper embryonic development of Sea Urchins does not occur unless chromosomes are present
- Walter Sutton observed the separation of chromosomes into daughter cells during meiosis.
- These observations led to the development of chromosomal theory of inheritance

Observations that support the chromosome theory of inheritance include:

- Chromosomes, like Mendel's genes, come in matched (homologous) pairs in an organism. For both genes and chromosomes, one member of the pair comes from the mother and one from the father.

- The members of a homologous pair separate in meiosis, so each sperm or egg receives just one member. This process mirrors segregation of alleles into gametes in Mendel's law of segregation.

- The members of different chromosome pairs are sorted into gametes independently of one another in meiosis, just like the alleles of different genes in Mendel's law of independent assortment.

The chromosome theory of inheritance was proposed before there was any direct evidence that traits were carried on chromosomes, and **it was controversial at first.**

In the end, it was **confirmed** through the work of geneticist **Thomas Hunt Morgan and his students**, who studied the genetics of fruit flies

Morgan chose the fruit fly, *Drosophila melanogaster*, for his genetic studies.

- They're cheap, easy, and fast to grow.
- You can raise hundreds of them in a little bottle with sugar sludge at the bottom,short generation time
- Only 4 pairs of chromosomes (3 pairs of autosomes and one pair of sex chromosomes)

He found a mutation in a gene affecting fly eye color.

This mutation made a fly's eyes white, rather than their normal red.

Unexpectedly, Morgan found that the eye color gene was inherited in different patterns by male and female flies.

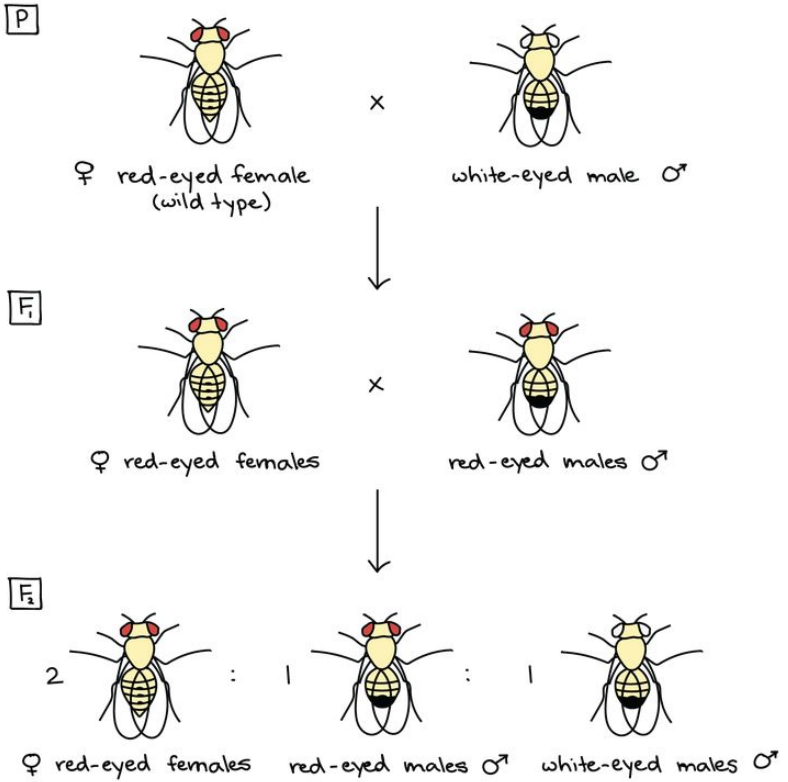
Male flies have an X and a Y chromosome (XY), while female flies have two X chromosomes (XX)

It didn't take Morgan long to realize that the eye color gene was being inherited in the same pattern as the X chromosome.

.

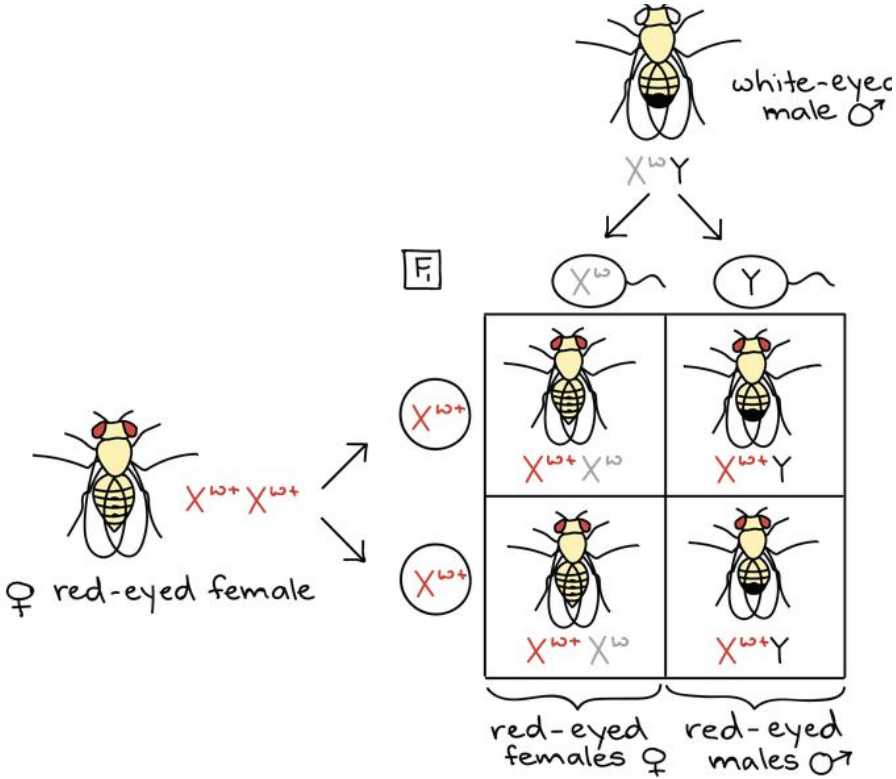
The first white-eyed fly he found was male, and when this fly was crossed with normal, red-eyed female flies, the F1 offspring were all red-eyed—telling Morgan that the white allele was recessive.

But when the F1 flies were crossed to each other,: all of the female F2 flies were red-eyed, while about half of the male F2 lies were white-eyed.

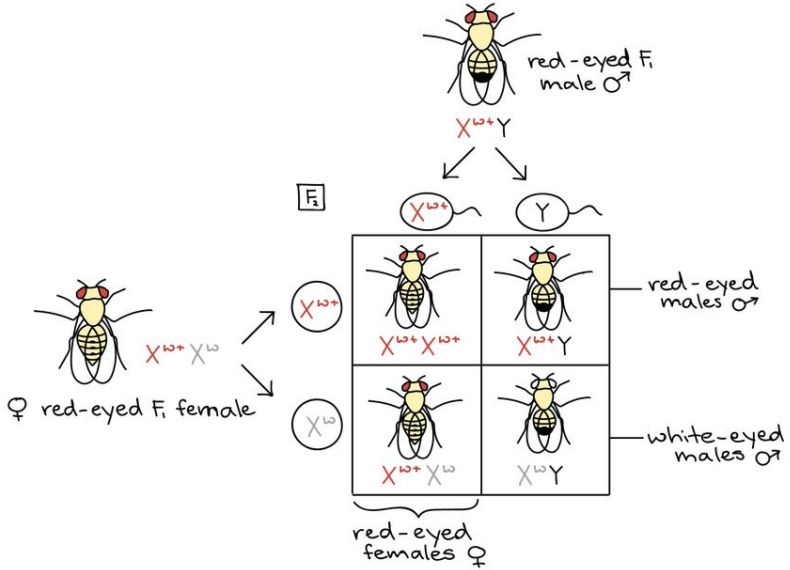


Earlier, we said that female flies have an XX genotype and male flies have an XY genotype.

If we stick the eye color gene on the X chromosome we can use a **Punnett square to show Morgan's first cross:**



when the F flies are mated to make the F2 generation:



Thus Morgan was able to **conclude that the gene for eye colour was on X chromosome**.

This explains the **behaviour of chromosomes during meiosis** by the **"principle of segregation" and "principle of independent assortment".**

Hence these principles are part of chromosomal theory of inheritance

Structure and forms of Nucleic acids

- **Nucleic acids** are the most important macromolecules that carry the genetic blueprint of a cell and carry instructions for the functioning of the cell.
- The two main types of nucleic acids are **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**.

Nucleotides

- DNA and RNA are made up of monomers known as nucleotides.
- The nucleotides combine with each other to form a polynucleotide: DNA or RNA.
- Each nucleotide is made up of three components:

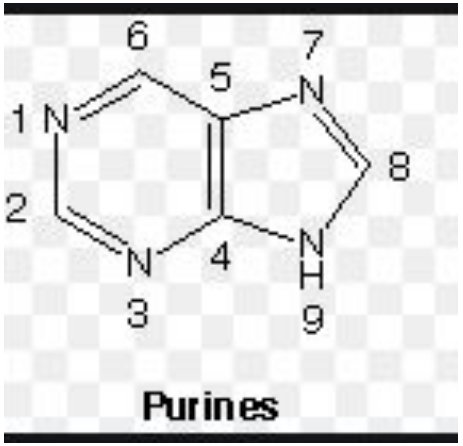
1. a nitrogenous base
2. a pentose (five-carbon) sugar
3. a phosphate group

Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.

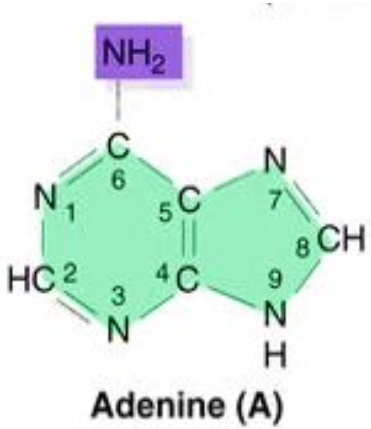
Basic components of Nucleotides

The nitrogenous bases:

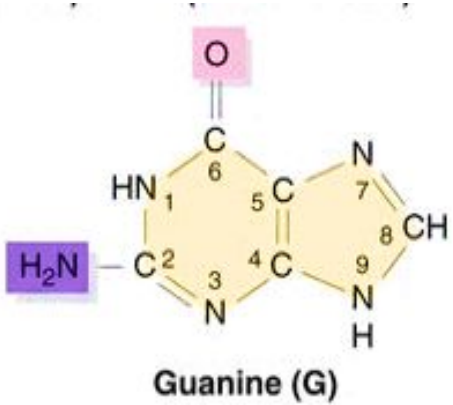
- Are organic molecules that contain carbon and nitrogen.
- Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).
- RNA nucleotides also contain one of four possible bases: adenine (A), guanine (G) cytosine (C) and uracil (U) rather than thymine.
- 2 types of nitrogenous bases: Purines and Pyrimidines
- **Purines:** The primary structure of a purine is two carbon-nitrogen rings
  - Adenine and guanine are classified as **purines**.
- **Pyrimidines:** single carbon-nitrogen ring as their primary structure
  - Cytosine, thymine, and uracil are classified as **pyrimidines**

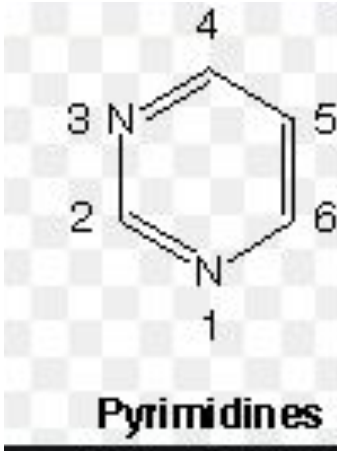


Adenine (6 amino purine)

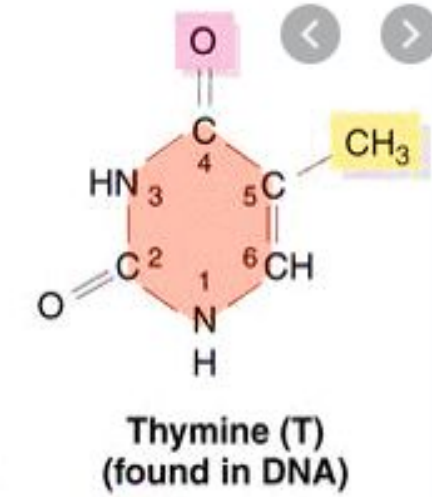


Guanine (2 amino 6 oxy purine)

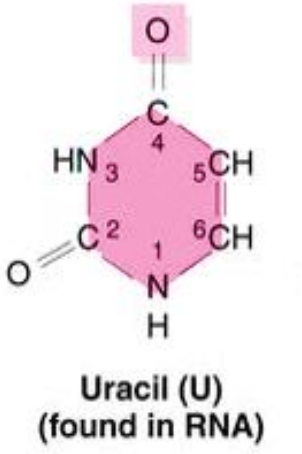




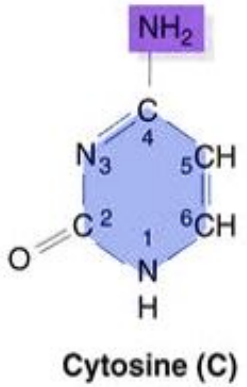
Thymine (2,4 dioxy 5 methyl pyrimidine)



Uracil (2,4 dioxy pyrimidine)

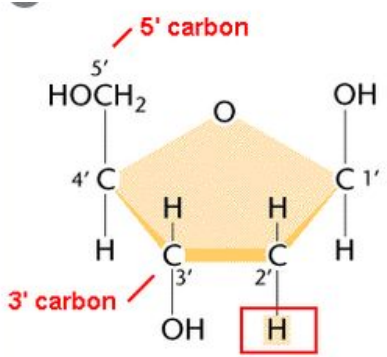


Cytosine (2 oxy 4 amino pyrimidine)

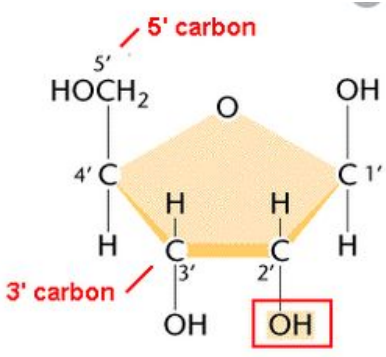


The pentose sugar:

- In DNA the sugar is deoxyribose, and in RNA, the sugar is ribose
- The difference between the sugars is the presence of the hydroxyl group on the second carbon of the ribose and hydrogen on the second carbon of the deoxyribose.
- The carbon atoms of the sugar molecule are numbered as 1′, 2′, 3′, 4′, and 5′

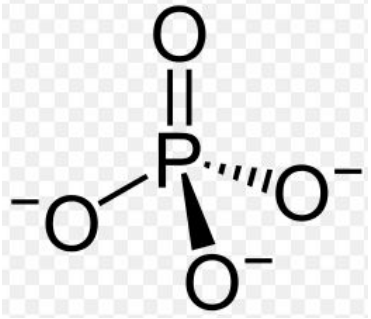


2-Deoxyribose



Ribose

The Phosphate group: H_3PO_4



Nucleoside: (base + sugar)

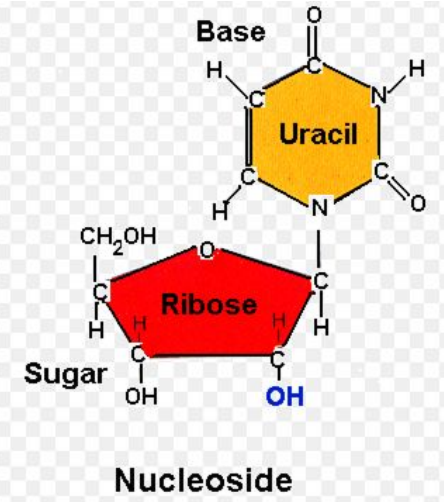
- Composed of a nucleobase, either a pyrimidine (cytosine, thymine or uracil) or a purine (adenine or guanine), a five carbon sugar which is either ribose or deoxyribose.
- **Bond: N-glycosidic bond**
- Nitrogen at position 9 of Purines and at position 1 of Pyrimidines is attached to 1st Carbon atom of the Sugar through N-glycosidic bond to form Nucleoside
- **Purine nucleoside :** have suffix "osine"

  ADENOSINE, GUANOSINE

  DEOXYADENOSINE, DEOXYGUANOSINE
- **Pyrimidine nucleoside:** have suffix "idine"

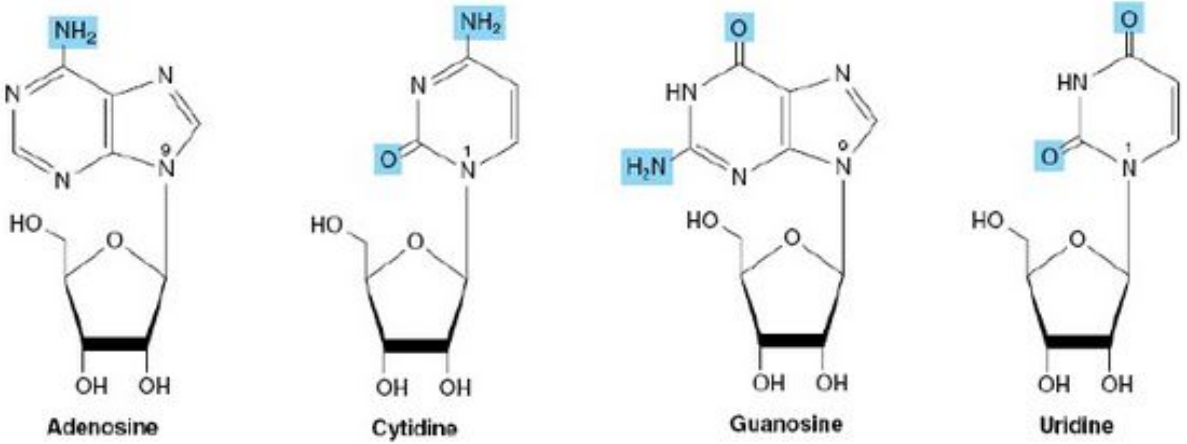
  URIDINE, CYTIDINE, THYMIDINE

  DEOXYURIDINE, DEOXYCYTIDINE,DEOXYTHYMIDINE



Nucleosides

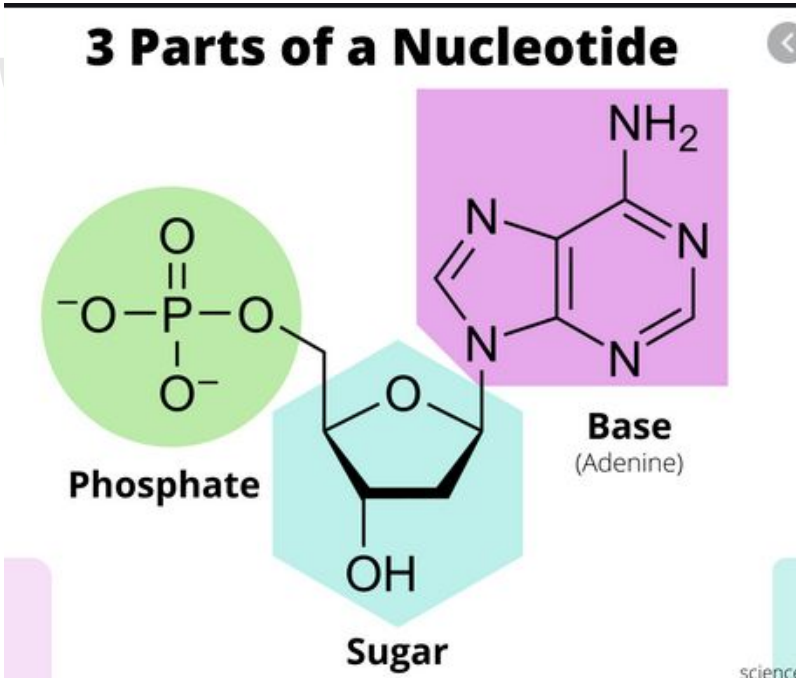
Nucleosides are derivatives of purines and pyrimidines that have a sugar linked to a ring nitrogen.

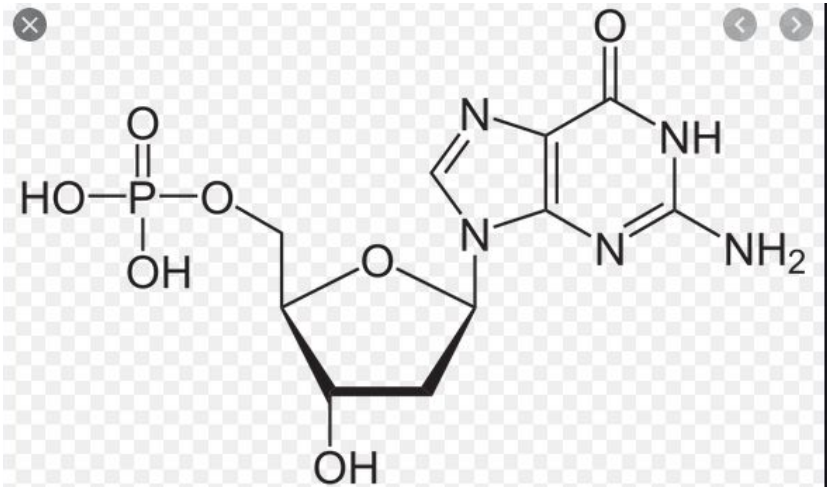


The sugar is linked to the heterocyclic base via a β-N-glycosidic bond, almost always to N-1 of a pyrimidine or to N-9 of a purine.

Nucleotides: Phosphorylated form of nucleoside

- Composed of three subunit molecules: a nucleobase, a five-carbon sugar (ribose or deoxyribose), and a phosphate group consisting of one to three phosphates.
- In a nucleotide, the sugar occupies a central position, with the base attached to its 1′ carbon and the phosphate group (or groups) attached to its 5′ carbon.
- Nucleotides may have a single phosphate group, or a chain of up to three phosphate groups, attached to the 5’ carbon of the sugar.
- 5’ carbon of sugar is attached to phosphate group by an **Ester bond**
- A maximum of 3 phosphates groups can be attached to nucleoside
- Nucleotides are named as: Nucleoside Monophosphate (NMP), Nucleoside diphosphate (NDP) and Nucleoside triphosphate (NTP) depending on no of PO4 group
- Nucleoside triphosphate is the basic building block of nucleic acids





deoxyguanosine monophosphate

Nucleoside triphosphate (Adenosine triphosphate)

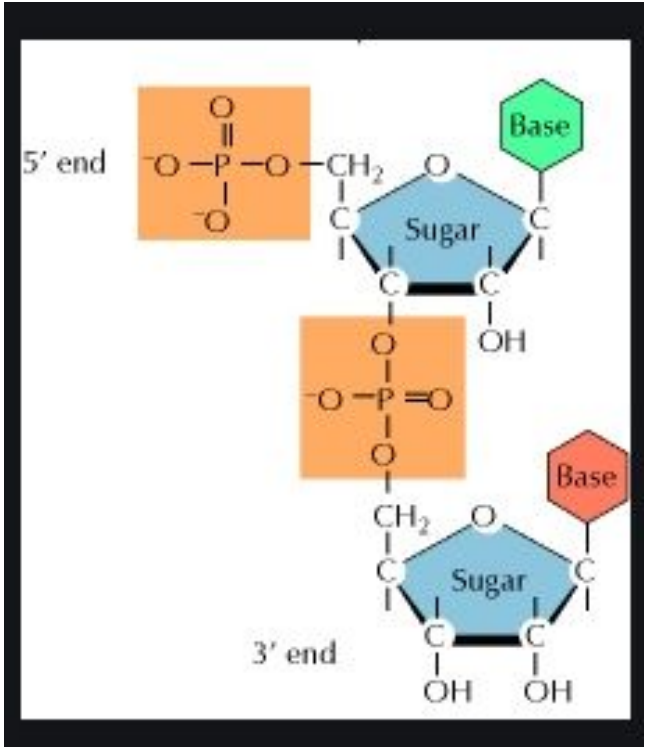
deoxyThymidine triphosphate

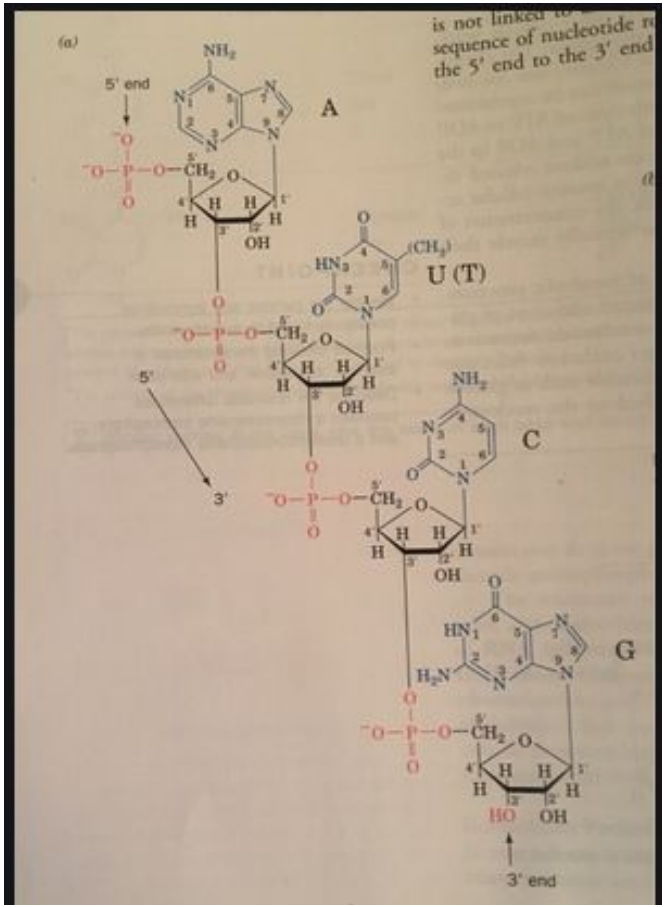
Building blocks of nucleic acids: Nuleotides-
In RNA: ATP, GTP, CTP, UTP
In DNA: dATP, dGTP, dCTP, dTTP

Nucleoside triphosphate NTP

Polynucleotide:

- Nucleic acids are polynucleotides composed of a series of nearly identical building blocks called nucleotides.
- The phosphate group connects successive sugar residues by bridging the 3′-hydroxyl group on one sugar to the 5′-hydroxyl group of the next sugar in the chain.
- These nucleoside linkages are called **phosphodiester bonds** and are the same in RNA and DNA.
- Only one phosphate is involved in this bond formation (release of inorganic pyrophosphate).
- This is repeated n no. of times to produce a polynucleotide or nucleic acid
- Thus the 1st nucleotide contains free 5’ phosphate group. No other free phosphates groups are present in nucleic acid
- This end of polynucleotide is referred to as 5’ end or “head” of nucleic acid.
- Similarly last nucleotide contains free OH group at 3’ end and is referred to as 3’ end or “tail” of nucleic acid

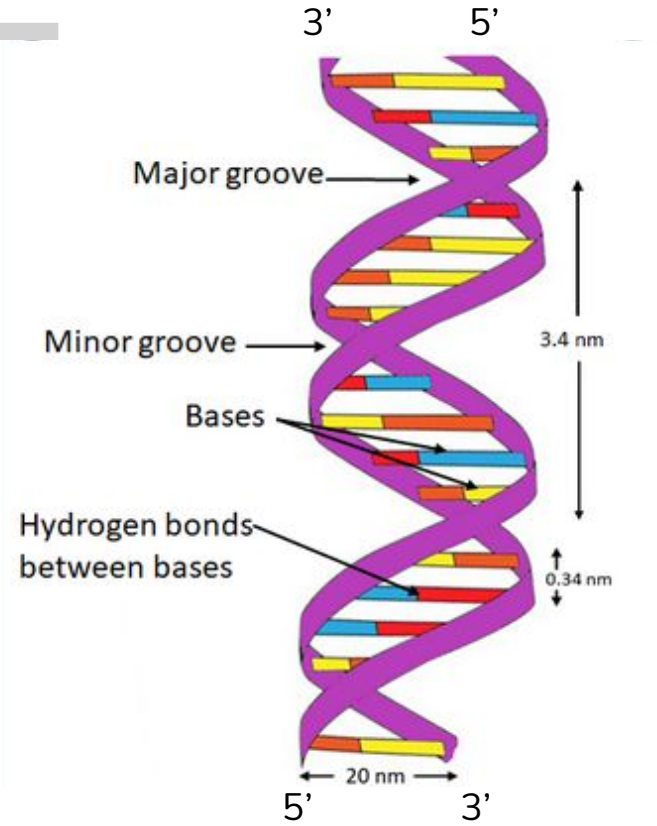

5′ end
Base
Sugar
Base
Sugar
3′ end
OH
OH
OH


5′ end
A
U (T)
5′
3′
C
G
3′ end

Structure of DNA

- In the 1950s, Francis Crick and James Watson worked together at the University of Cambridge, England, to determine the structure of DNA.
- Watson and Crick proposed that the DNA is made up of two strands that are twisted around each other to form a right-handed helix, called a double helix.
- Base-pairing takes place between a purine and pyrimidine: namely, **A pairs with T, and G pairs with C**.
- In other words, adenine and thymine are complementary base pairs, and cytosine and guanine are also complementary base pairs.
- This is the basis for Chargaff's rule; because of their complementarity, there is as much adenine as thymine in a DNA molecule and as much guanine as cytosine.
- Adenine and thymine are connected by **two hydrogen bonds**, (A=T) and cytosine and guanine are connected by **three hydrogen bonds** (G ≡ C) G:C bond is stronger than A:T bond

- The two strands are **anti-parallel** in nature i.e, one strand will have the 3′ carbon of the sugar in the "upward" position, whereas the other strand will have the 5′ carbon in the upward position.
- The diameter of the DNA double helix is uniform throughout because a purine (two rings) always pairs with a pyrimidine (one ring) and their combined lengths are always equal.
- Sugars and phosphates together form the backbone of the helix.
- The bases are present perpendicular to the axis of the helix and face each other.
- Helix has 10 bases in a turn
- Each base is thus at an angle of 36° to the adjoining base.
- The linear distance of one complete turn is 34Å or 3.4nm and therefore each base is 3.4Å or 0.34nm away from the adjoining base
- The diameter of the helix is 20Å or 2nm
- 3D structure of DNA creates two grooves of different sizes in each turn of DNA molecules
- One of these is larger and known as **Major groove**, while other is relatively smaller and known as **Minor groove**


3’
5’
Major groove
Minor groove
Bases
Hydrogen bonds
between bases
3.4 nm
0.34 nm
20 nm
5’
3’

Different forms of DNA

1. A form
2. B form
3. Z form

A DNA:

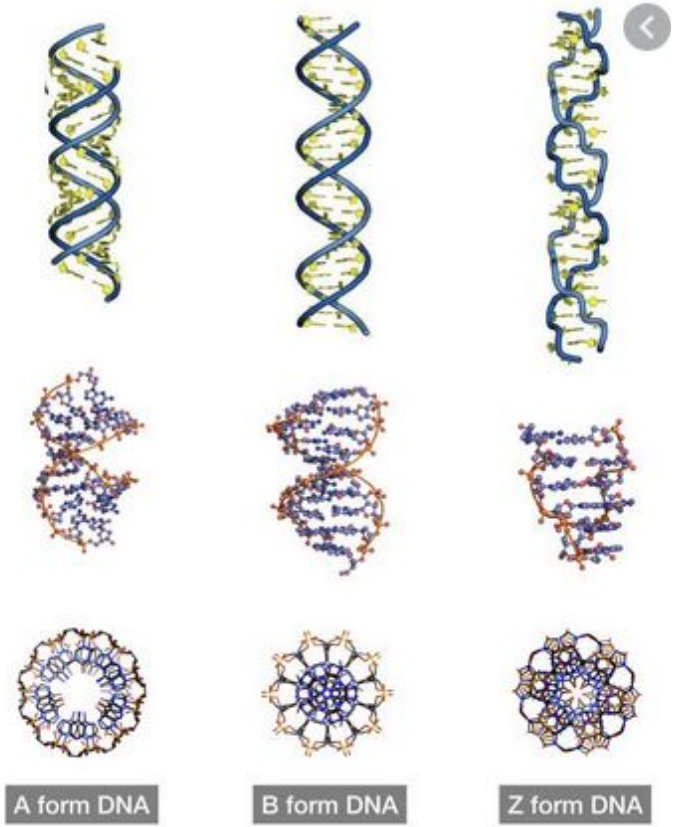
- Right-handed double helix
- Appears when the relative humidity of the environment is less than 75% it is rarely present in normal physiological condition
- Two strands of A-DNA are anti-parallel
- The molecule is asymmetrical because the glycosidic bonds of a base pair are not diametrically opposite to each other. Therefore, major grooves and minor grooves can be observed in each turn.
- One turn of the helix consists of 11 base pairs with a length of 2.86nm
- The backbone of A-DNA is formed by sugar phosphates. All the nitrogenous bases are at the core centre of the helix linked by Hydrogen bonds
- The helix width of A-DNA is 2.3nm.
- Overall, A-DNA is wider than the more commonly found B-DNA.

B DNA:

- It is the common form of DNA exists under normal physiological condition.
- Right-handed double helix, which was discovered by Watson and Crick
- The double strands of B-DNA run in opposite directions.
- The structure is asymmetrical with major grooves and minor grooves present alternatively.
- The molecule is asymmetrical is because the glycosidic bonds of a base pair are not diametrically opposite to each other
- In one turn, there are 10 base pairs with a length of 3.4nm. The distance between adjacent deoxyribonucleotides is 0.34nm
- The backbone of B-DNA is formed by sugar phosphates that are linked continuous using phosphodiester bonds and the core region consists of nitrogenous bases.
- The two strands are held together by hydrogen bonds between nitrogenous bases.
- The helix width of B-DNA is 2nm.
- B-DNA is narrower than A-DNA.

Z DNA:

- Left-handed double helix
- It has a very different structure when compared with A-DNA and B-DNA.
- The backbone : zigzag
- Same as the other forms of DNA, hydrogen bond is present to hold the two strands together.
- The helix width is 1.8nm, which is **the narrowest among the three types**.
- The structure consists of major and minor grooves.
- One turn has 12 base pairs and the length is 4.56nm.
- The distance between two adjacent deoxyribonucleotides is 0.37nm.
- Z-DNA is difficult to be observed since it is unstable.


A form DNA
B form DNA
Z form DNA

STRUCTURE OF RNA

- **RNA - Ribonucleic acid** molecule essential in various biological roles in coding, decoding, regulation and expression of genes.
- **Structurally** RNA is quite **similar to DNA.**
- However, whereas DNA molecules are typically long and double stranded, RNA molecules are much shorter and are typically single stranded.
- RNA molecules perform a variety of roles in the cell but are mainly involved in the process of **protein synthesis (translation)** and its regulation.

RNA Structure

- RNA is typically single stranded and is made of **ribonucleotides** that are linked by phosphodiester bonds.
- A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and a phosphate group.
- Even though RNA is single stranded, most types of RNA molecules show extensive intramolecular base pairing between complementary sequences within the RNA strand, creating a predictable three-dimensional structure essential for their function

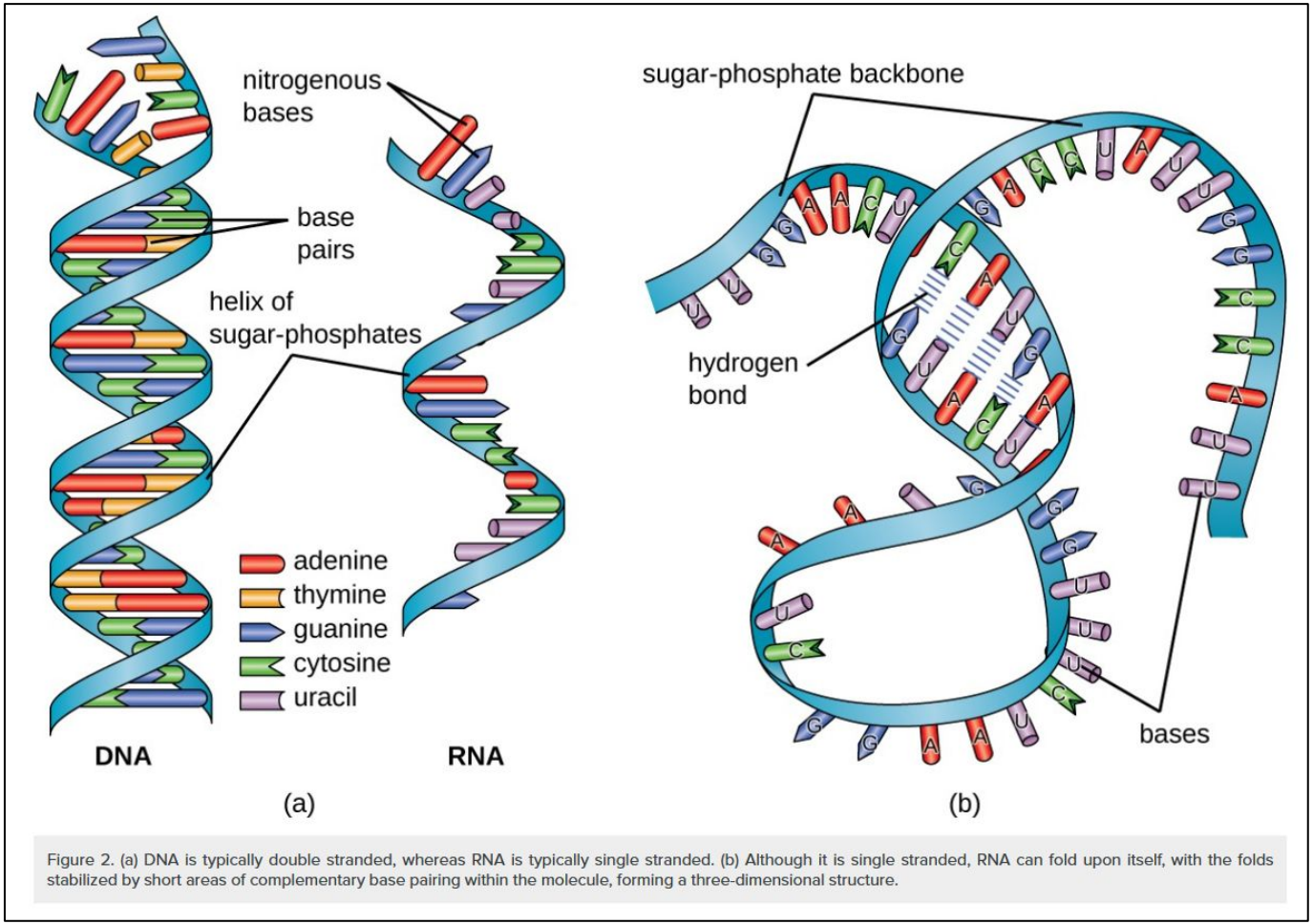


Figure 2. (a) DNA is typically double stranded, whereas RNA is typically single stranded. (b) Although it is single stranded, RNA can fold upon itself, with the folds stabilized by short areas of complementary base pairing within the molecule, forming a three-dimensional structure.

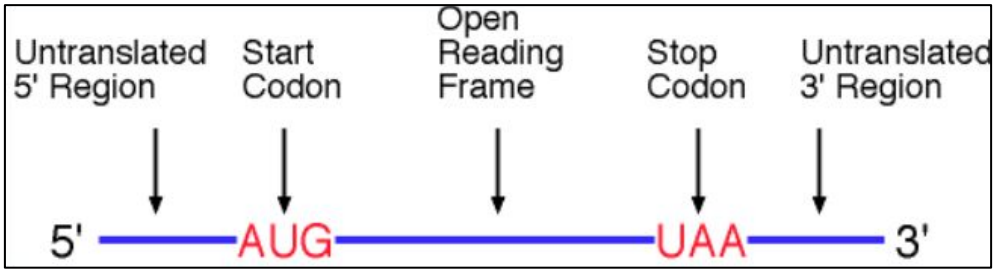
TYPES OF RNA:

- The three main types of RNA directly involved in protein synthesis are **messenger RNA (mRNA)**, **ribosomal RNA (rRNA)**, and **transfer RNA (tRNA)**.

mRNA: Messenger RNA

- mRNA carries genetic codes from the DNA in the nucleus to ribosomes, the sites of protein translation in the cytoplasm.
- Messenger RNA (mRNA) is a single-stranded RNA molecule that is complementary to one of the DNA strands of a gene.
- It acts as **Template** for protein synthesis
- Represents only 2% of total cellular RNA
- mRNA is relatively unstable and short-lived in the cell, especially in prokaryotic cells, ensuring that proteins are only made when needed.

- Prokaryotic and eukaryotic messenger RNAs (mRNAs) have different structures.
- Prokaryotic mRNAs are often **polycistronic** (that is, they carry the information for more than one protein)
- Eukaryotic mRNAs are **monocistronic** and almost always code for a single protein
- Eukaryotic mRNAs also have structural features that prokaryotic ones do not
- Eukaryotic mRNAs contain a methylated **cap** at the **5′ end** and stretch of A residues **@ 3'end** known as **polyA tail**
- **Coding region** : codes for primary structure of proteins. Begins from initiation codon AUG and ends in one of the 3 stop codons: UAA, UAG, UGA
- There are 3' and 5' non coding regions also.



rRNA: Ribosomal RNA

- Ribosomal RNA is essential for protein synthesis.
- rRNA is a major constituent of **ribosomes** (Ribosomes are composed of rRNA and protein)
- rRNA composes up to about 60% of the ribosome by mass and providing the location where the mRNA binds.
- It is the most stable and more abundant form of RNA
- It constitutes 80% of total cellular RNA
- Single stranded

tRNA: Transfer RNA

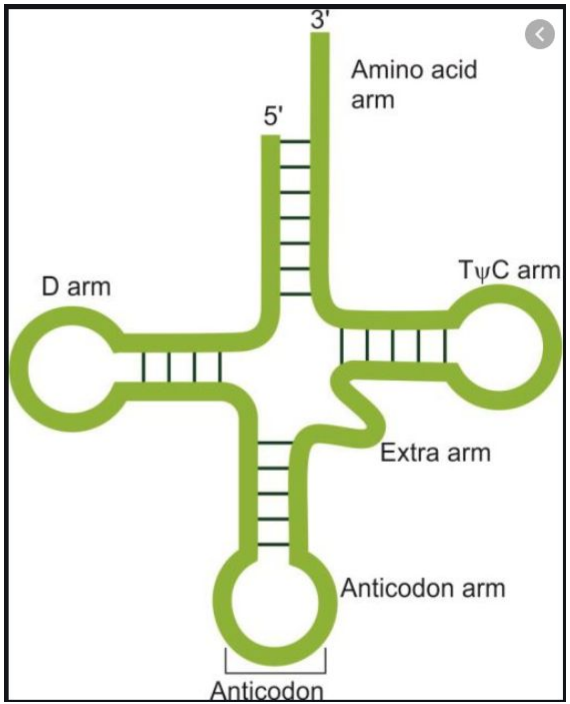
- It is an **adaptor molecule** between mRNA and amino acid
- One of the smallest, usually only 70–90 nucleotides long.
- It carries the correct amino acid to the site of protein synthesis in the ribosome.
- Constitutes 10-15% of total cellular RNA
- The common secondary structure of tRNAs is the **cloverleaf** pattern, where the 5′ and 3′ sequences are base-paired, and then the other three stem-loops of the cloverleaf are formed by intramolecular base pairs over a short distance.
- All tRNAs end in the **acceptor sequence**, CCA
- The other part of the tRNA is the **anticodon**, which forms base pairs with the trinucleotide ***codon*** sequence of the *ribosome*-bound message. The anticodon is found at the same place in each tRNA cloverleaf, away from the acceptor stem
- The acceptor stem accepts a specific amino acid, while the anticodon determines the placement of that amino acid at the correct point in the growing polypeptide chain.

Clover leaf structure: Secondary structure of tRNA:

- Comprised of three characteristic loops.
- The loop closest to the 5' end is called the **dihydrouridine arm (D arm),** because it contains dihydrouridine bases, which are unusual nucleotides common only to tRNA.
    - Stem: 4bp
    - Loop: variable size
- The loop closest to the 3' end is called the **T arm,** after its sequence of thymine-pseudouridine-cytosine (pseudouridine is also an unusual base).
    - Stem: 5bp
    - Loop: 7 bases
- The loop on the bottom of the cloverleaf contains the **anticodon**, which binds complementarily to the mRNA codon.
    - Stem: 5bp
    - Loop: 7 bases
- At the 3' end of the tRNA molecule, opposite the anticodon, extends a three nucleotide (CCA) acceptor site that includes a free -OH group. A specific tRNA binds to a specific amino acid through its **acceptor stem**.
    - Stem: 7bp

Tertiary structure of tRNA:

- The cloverleaf folds further into a tertiary structure, a sort of vague L-shape.
- At one end of the L lies the anticodon; at the other is the acceptor stem.
- The L-shaped structure simply amplifies the two active ends of tRNA: the anticodon and the acceptor stem.



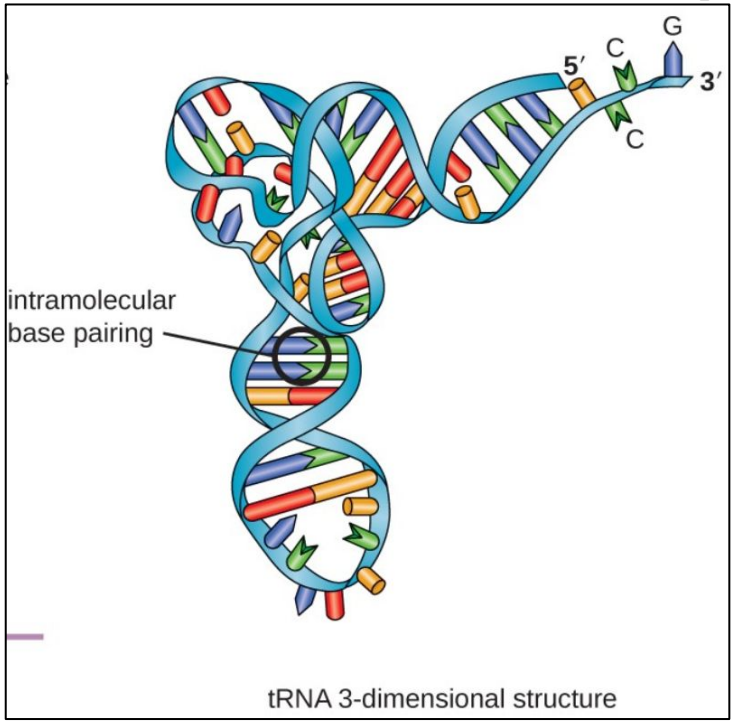


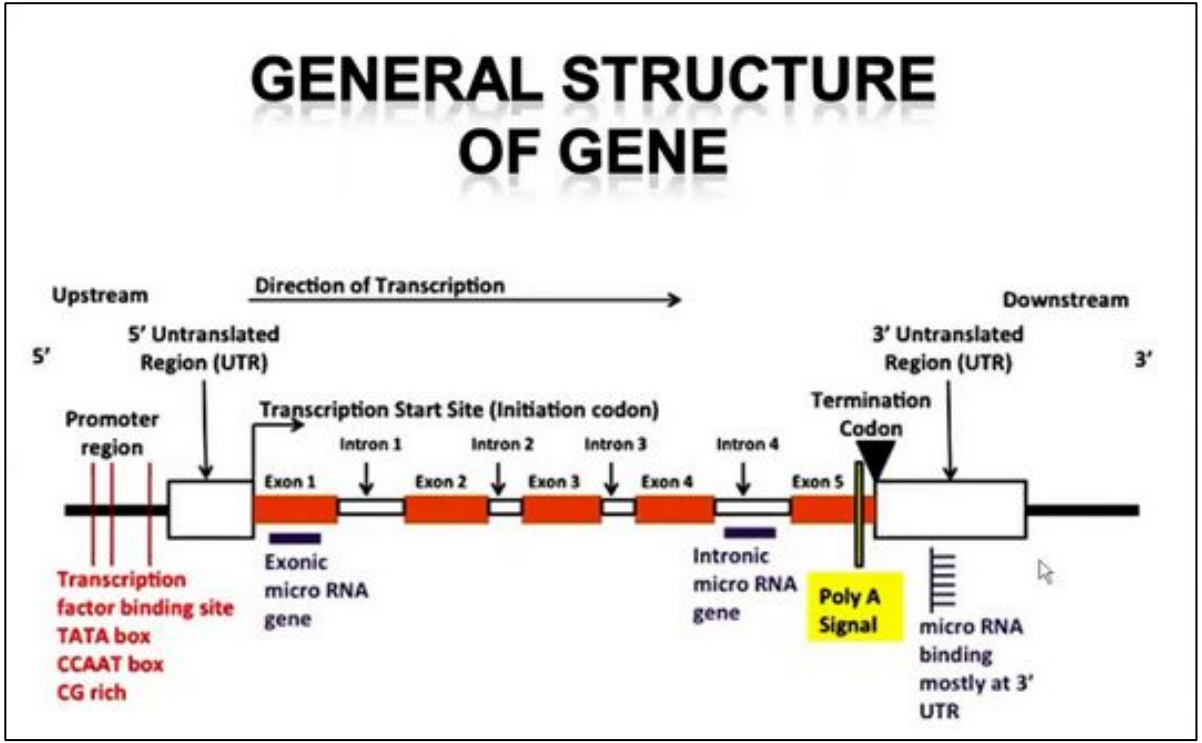
Table 1. Structure and Function of RNA

	mRNA	rRNA	tRNA
Structure	Short, unstable, single-stranded **RNA** corresponding to a gene encoded within DNA	Longer, stable RNA molecules composing 60% of ribosome's mass	Short (70-90 nucleotides), stable RNA with extensive intramolecular base pairing; contains an amino acid binding site and an mRNA binding site
Function	Serves as intermediary between **DNA** and protein; used by ribosome to direct synthesis of protein it encodes	Ensures the proper alignment of mRNA, tRNA, and ribosome during protein synthesis; catalyzes **peptide bond** formation between amino acids	Carries the correct amino acid to the site of protein synthesis in the ribosome

Structure of GENE

- A gene is a **functional- hereditary unit** made up of nucleotides which forms proteins.
- Genes are **located on chromosomes**
- Genes are actually DNA strands thus are made up of the nucleotide chain. The chemical structure of a gene comprises nucleotides.
- In general, the gene structure consists of **two types of elements:** core elements and regulatory elements.
- The core elements or transcriptional region actually take parts in protein formation. While the regulatory elements maintain gene expression.
- Transcriptional region starts from Transcription Start Site (TSS). This nucleotide is designated as 1st nucleotide.
- Nucleotides upstream of TSS (in the regulatory region) are given negative numbers.
- Much of gene structure is broadly similar between eukaryotes and prokaryotes.

- Although DNA is a double-stranded molecule, typically only one of the strands encodes information that the RNA polymerase reads to produce protein-coding mRNA or non-coding RNA. This **'sense' or 'coding' strand,** runs in the 5' to 3' direction



Regulatory region:

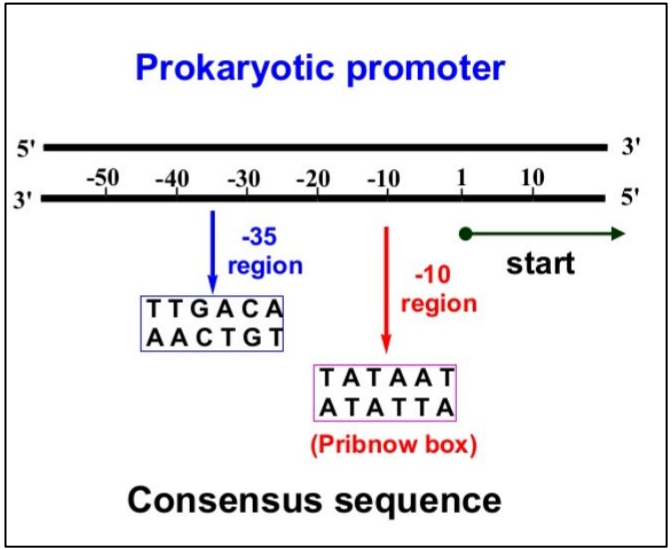
1. **PROMOTER:**

- It is the most important region
- It facilitates the binding of enzyme RNA polymerase to gene and directs it to initiate transcription
- Promoters are asymmetrical and are orientation and position specific.
- Have bipartite organization: 2 regions of highly conserved sequences
    - **1st region**: known as **Recognition site**
        - It is recognised by enzyme RNA polymerase and facilitates its binding to the template
        - It is present far from transcriptional unit
    - **2nd region:** known as **RNA polymerase binding site**
        - It helps in initiation of RNA synthesis
        - It is present near to the transcriptional unit

Prokaryotic promoter:

Ist region:

- Consists of **Hexanucleotide**
- Has consensus sequence **TTGACA**
- Present between 30-40 nucleotides upstream of TSS
- Known as **"-35 region"**



2nd region:

- Consists of **Hexanucleotide**
- Has consensus sequence **TATAAT**
- Present between -5 to -14
- Known as **"-10 region" or "Pribnow box"**
- This region is highly A-rich. Hence easier to open thermodynamically during initiation of transcription

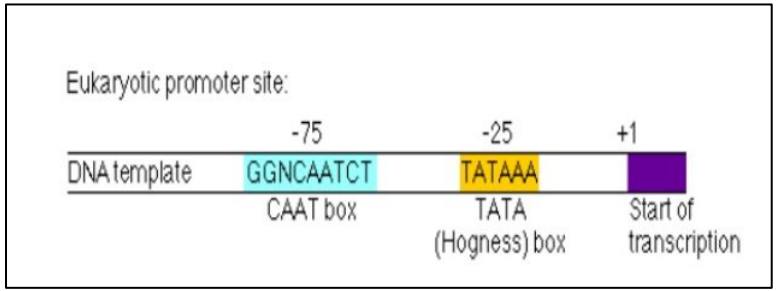
Distance between -10 and -35 region varies from promoter to promoter. Optimal distance 17bp

Eukaryotic promoter:

- It is more complex

1st region

- Consists of **Nanonucleotide**
- Has consensus sequence **GG(C/T)CAATCT**
- Present between -70 to -80
- Similar to "-35 region" of Prokaryotes
- Known as **CAAT box**



2nd region:

- Consists of **Septanucleotide**
- Has consensus sequence **TATA(A/T)A(A/T)**
- Present between -19 to -27
- Known as **"TATA box"**
- Similar to Pribnow box of Prokaryotes
- This region is highly A-rich. Hence easier to open thermodynamically during initiation of transcription

Optimal distance between CAAT box and TATA box is 48bp

2. Other regulatory elements:

- Besides promoter eukaryotic genes alo contain certain regulatory elements such as
  - Activators
  - Silencers
  - Enhancers
- All these are **cis acting elements-** present on gene itself and manifest their effect without being expressed in the form of either RNA or Protein
- Their function is to regulate the level of gene expression either by themselves or with the help of other **trans acting elements**

Transcriptional unit:

- It is the portion of the gene which starts from TSS and ends at transcriptional terminator
- Entire transcriptional unit is copied to form a RNA
- **1st element: TSS.** do not have conserved sequence
- **2nd element: spacer region** of varying length

In prokaryotes:

- **Spacer region** is followed by **Ribosome binding site (RBS)** or **Shine Dalgarno (SD)sequence**
- **RBS** is complimentary with 3'end of 16s rRNA with which it base pairs and facilitates binding of small subunit of Ribosome to the mRNA during initiation of translation.
- **4-9 bases** long

In eukaryotes:

- **RBS** is **not present**

3rd element: initiation codon

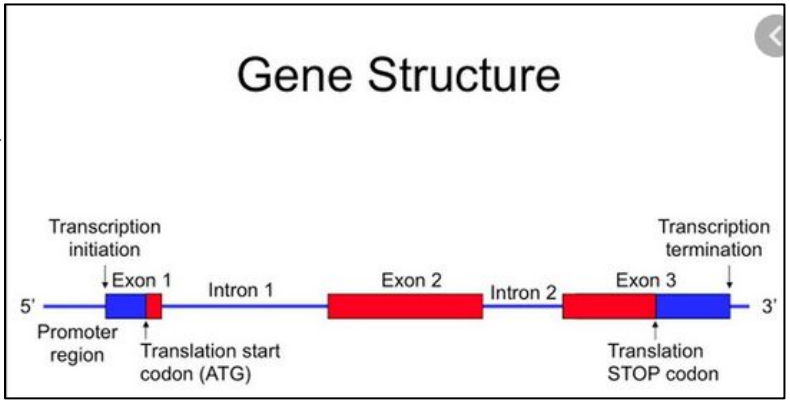
4th element: coding region.

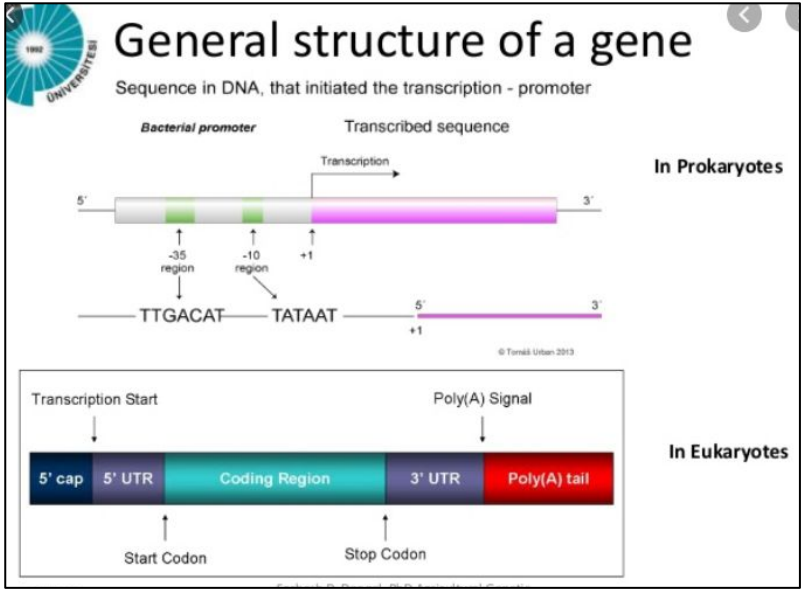
- It is the region followed by initiation codon
- It is a continuous stretch **in prokaryotes** and a single ORF is present which ends in Transcriptional stop codon
- **Eukaryotic genes** have a number of intervening sequences interspread in between coding region. These sequences are known as **Introns- do not code for any polypeptide.**
- **Coding region ends** in one of the 3 **Stop codons**

Transcriptional unit ends in transcriptional stop signal

Eukaryotes:

- There is **downstream regulatory sequence (DRS)**
- Present 10-20 nucleotides upstream of transcription terminator
- This has a consensus sequence **AAATAA**
- It is called as **Polyadenylation site**
- It also provides signal for the termination of trancsription





LOCATION OF GENES:

- Genes are contained in chromosomes, which are in the cell nucleus.
- A chromosome contains hundreds to thousands of genes.
- Every normal human cell contains 23 pairs of chromosomes, for a total of 46 chromosomes.
- Humans have about 20,000 to 23,000 genes.
- The genes on each chromosome are arranged in a particular sequence, and each gene has a particular location on the chromosome (called its locus).
- In order to make proteins, the gene from the DNA is copied by each of the chemical bases into messenger RNA (ribonucleic acid) or mRNA. The mRNA moves out of the nucleus and uses cell organelles in the cytoplasm called ribosomes to form the polypeptide or amino acid that finally folds and configures to form the protein.

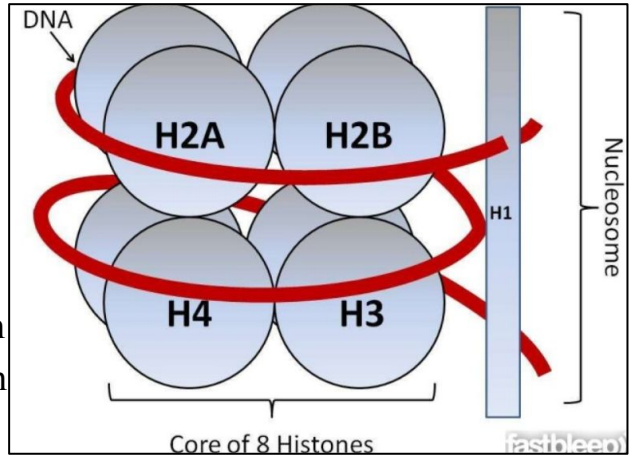
DNA Packaging

DNA packaging is an important process in living cells. Without it, a cell is not able to accommodate the large amount of DNA that is stored inside.

The coiling of DNA in the chromosomes is in the form of highly organized structure called chromatin

Nucleosome model of chromosome

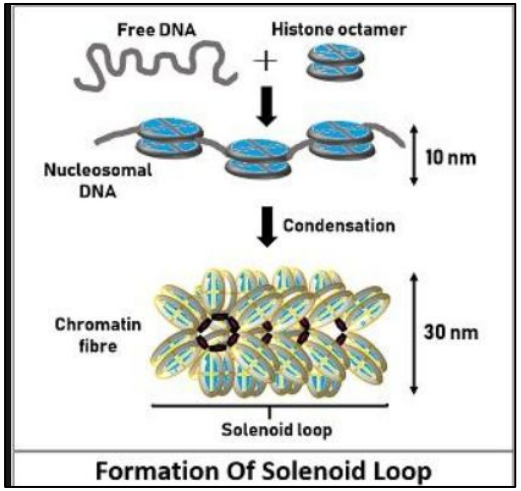
- Nucleosomes are the basic unit of chromatin structure
- Nucleosome model explains the organization of DNA and associated proteins in the chromosome.
- In eukaryotes, DNA is tightly bound to an equal mass of histones, which serve to form a repeating array of DNA-protein particles, called nucleosomes.
- Histones play a crucial role in packing this very long DNA molecule in an orderly way (i.e., nucleosome) into nucleus only a few micrometers in diameter.
- Each nucleosome is a disc-shaped particle with a diameter of about 11 nm containing 2 copies of each 4 nucleosome histones–H2A, H2B, H3, and H4
- This histone octamer forms a protein core around which the double-stranded DNA helix is wound 1¾ time containing 146 base pairs.
- Thus, nucleosomes are the fundamental packing unit particles of the chromatin and give chromatin a “beads-on-a-string” appearance in electron micrographs



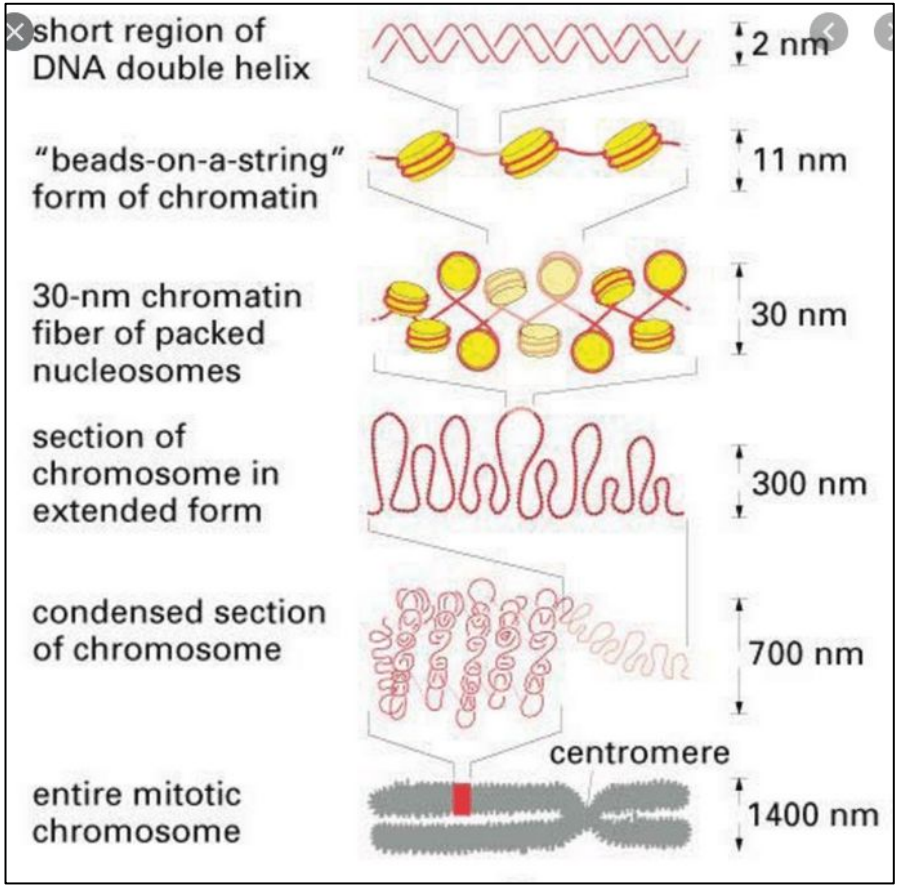
- In chromatin, the DNA extends as a continuous thread from nucleosome to nucleosome.
- Each nucleosome bead is separated from the next by a region of linker DNA which is generally 54 base pair long and contains single H1 histone protein molecule.
- **“Beads-on-a-string” represents first level of chromosomal DNA packaging**
- **2nd degree: 30nm fibre:**

  Nucleosomes are further organized in a highly coiled, lumpy fibres which have an average diameter of about 30nm.
- **3rd degree: scaffold structure**

  Condensation of DNA takes place with the help of non histone proteins resulting in the formation of scaffold structure.
- After these levels of organization there is a 10^4 fold condensation of DNA which is arranged to form chromosomes



Formation Of Solenoid Loop


short region of DNA double helix
2 nm
"beads-on-a-string" form of chromatin
11 nm
30-nm chromatin fiber of packed nucleosomes
30 nm
section of chromosome in extended form
300 nm
condensed section of chromosome
700 nm
centromere
entire mitotic chromosome
1400 nm

Euchromatin	Heterochromatin
Lightly staining regions	Darkly staining
Less tightly packed chromatin fibers therefore non condensed	Tightly packed chromatin fibers therefore condensed
Not visible – light microscope, undergo regular changes in morphology with cell division	Visible, remain highly condensed in all stages
Genetically active regions	Genetically inactive regions – either they lack genes/ contain genes that are not expressed
Replicates earlier during S phase	Replicates later during S phase
GC rich	AT rich

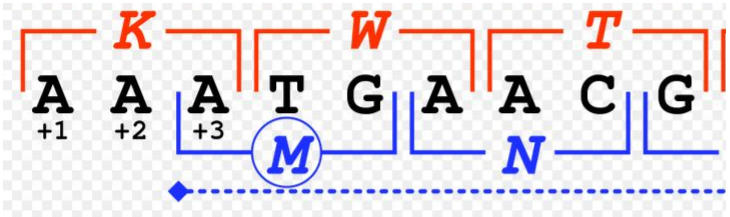
Genetic code

- The **genetic code** is the set of rules used by living cells to translate information encoded within genetic material into proteins.
- **Genetic code**, the sequence of nucleotides in DNA and RNA that determines the amino acid sequence of proteins.
- Three adjacent nucleotides constitute a unit known as the codon, which codes for an amino acid.
- For example, the sequence AUG is a codon that specifies the amino acid methionine.
- There are 64 possible codons, three of which do not code for amino acids but indicate the end of a protein. The remaining 61 codons specify the 20 amino acids that make up proteins.

- There are 20 amino acids that make up proteins and there are 4 bases present in mRNA/ DNA
- If each base had coded for one amino acid then it could code for only 4 amino acids.
- If 2 bases were to code for one a.a, then it could code for 4X4=16 a.a. 4 a.a will be left uncoded
- If 3 bases code for one a.a, then 4X4X4= 64 a.a can be coded.
- Thus a minimum of 3 nucleotides are required to form a single codon.

Reading frame

- A reading frame is defined by the initial triplet of nucleotides from which translation starts.
- It sets the frame for a run of successive, non-overlapping codons, which is known as an "open reading frame" (ORF).
- For example, the string **5'-AAATGAACG-3'**
  - if read from the first position, contains the codons AAA, TGA, and ACG ;
  - if read from the second position, it contains the codons AAT and GAA ; and
  - if read from the third position, it contains the codons ATG and AAC.
- Every sequence can, thus, be read in its 5' → 3' direction in three reading frames, each producing a possibly distinct amino acid sequence
- in the given example
  - 1st frame: Lys (K)-Trp (W)-Thr (T)
  - 2nd frame: Asn (N)-Glu (E)
  - 3rd frame: Met (M)-Asn (N)
- When DNA is double-stranded, six possible reading frames are defined, three in the forward orientation on one strand and three reverse on the opposite strand



Salient feature or characteristics of genetic code:

1. The codon is **triplet**. 61 codons code for amino acids and 3 codons do not code for any amino acids hence they function as stop codons.
2. One codon codes for only one amino acid hence it is **unambiguous and specific**.
3. Some amino acids are coded by more than one codon hence the code is **degenerate**.
4. The codon is read in mRNA in a contiguous fashion. There are **no punctuations**.
5. The code is nearly **universal**: for example from bacteria to human UUU would code for Phenylalanine phe. Some exceptions to this rule have been found in mitochondrial codons and in some protozoans.

3 **STOP CODONS:**

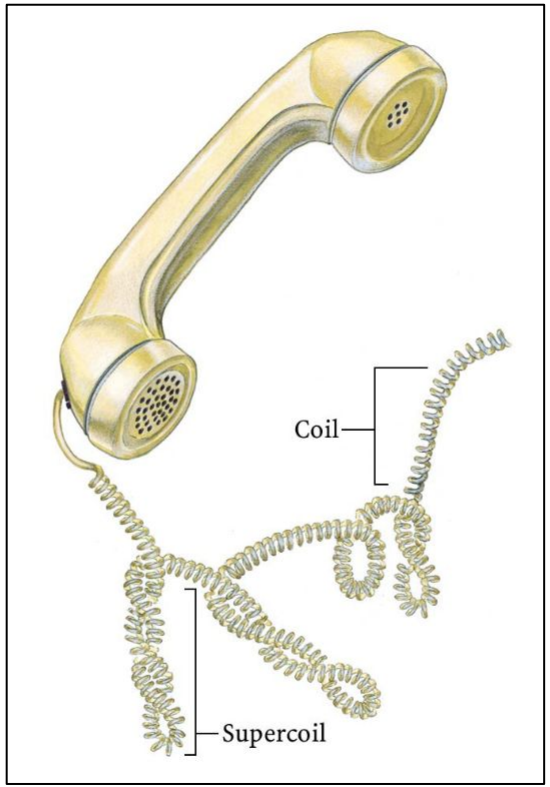
UAA: OCHRE

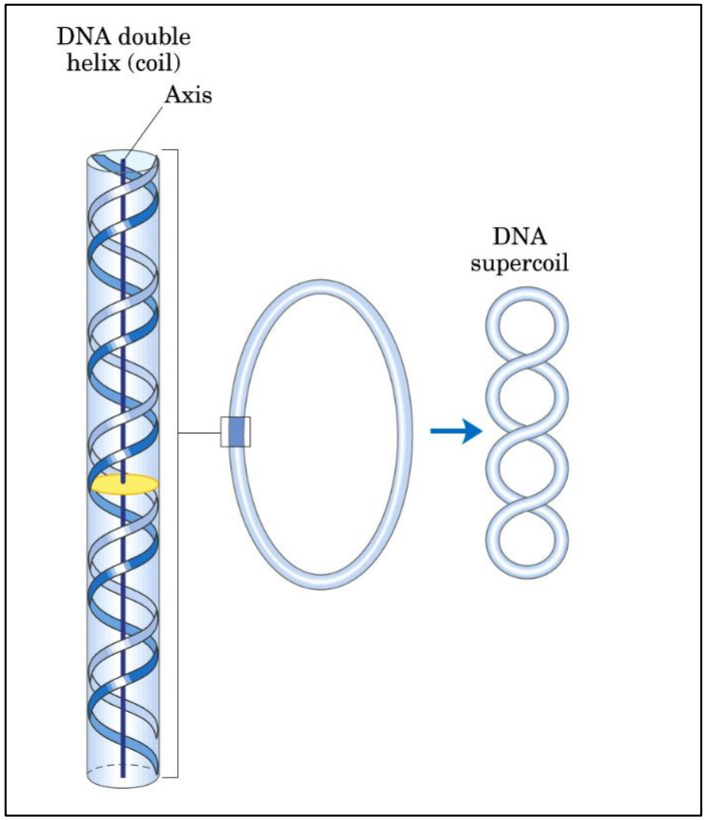
UAG: AMBER

UGA: OPAL

Tertiary structure of DNA: DNA supercoiling

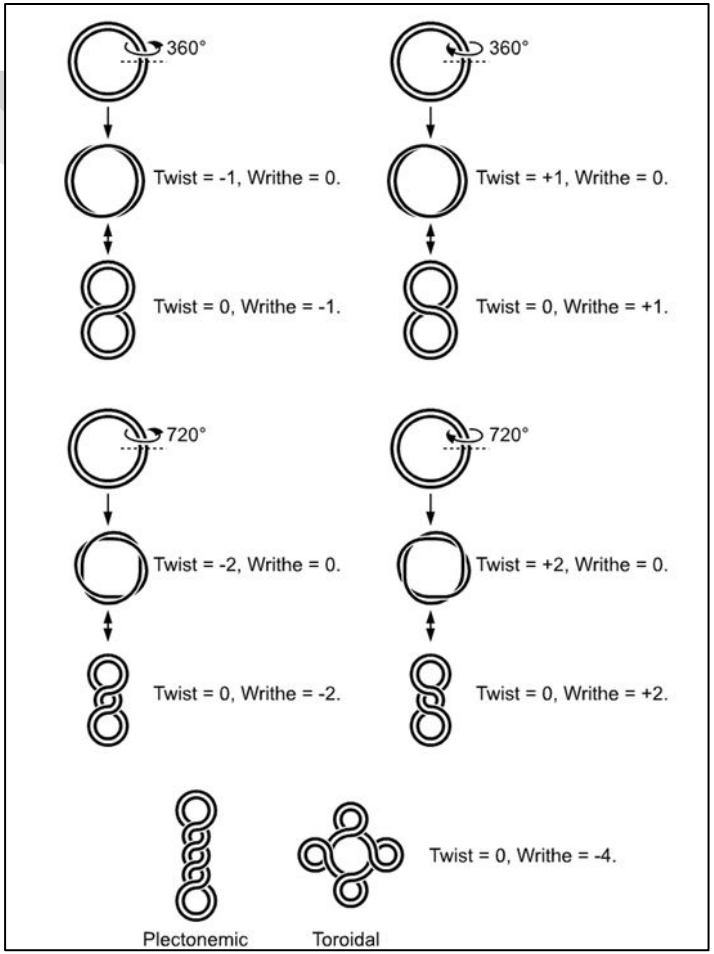
- DNA supercoiling describes a higher-order DNA structure.
- DNA supercoiling refers to the over- or under-winding of a DNA strand, and is an expression of the strain on that strand.
- Supercoiling is important in a number of biological processes, such as compacting DNA.
- Additionally, certain enzymes such as topoisomerases are able to change DNA topology to facilitate functions such as DNA replication or transcription.
- DNA supercoiling is a special property of circular, double-stranded DNA that is topological in origin. It confers new structural and energetic properties.
- The simple figure **eight (8)** is the simplest supercoil, and is the shape a circular DNA assumes to accommodate one too many or one too few helical twists.
- The two lobes of the figure eight will appear rotated either clockwise or counterclockwise with respect to one another, depending on whether the helix is over or underwound.

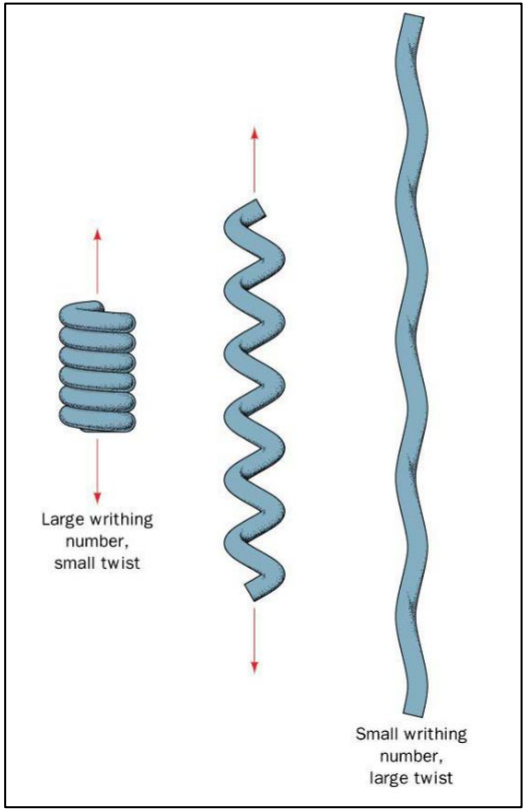




A. Right handed supercoiling = negative supercoiling (underwinding)

B. Left handed supercoiling = positive supercoiling


360°
Twist = -1, Writhe = 0.
Twist = 0, Writhe = -1.
360°
Twist = +1, Writhe = 0.
Twist = 0, Writhe = +1.
720°
Twist = -2, Writhe = 0.
Twist = 0, Writhe = -2.
720°
Twist = +2, Writhe = 0.
Twist = 0, Writhe = +2.
Twist = 0, Writhe = -4.
Plectonemic
Toroidal


Large writhing number, small twist
Small writhing number, large twist

- "Supercoiling" is an abstract mathematical property representing the **sum of twist and writhe**.
- The twist is the number of helical turns in the DNA and the writhe is the number of times the double helix crosses over on itself (these are the supercoils).
- Numerical expression for the degree of supercoiling: **Lk=Tw+Wr**
- **Lk, the linking number:** fundamental property of DNA supercoiling-- The number of times one strand winds about other strand. It is always an integer.

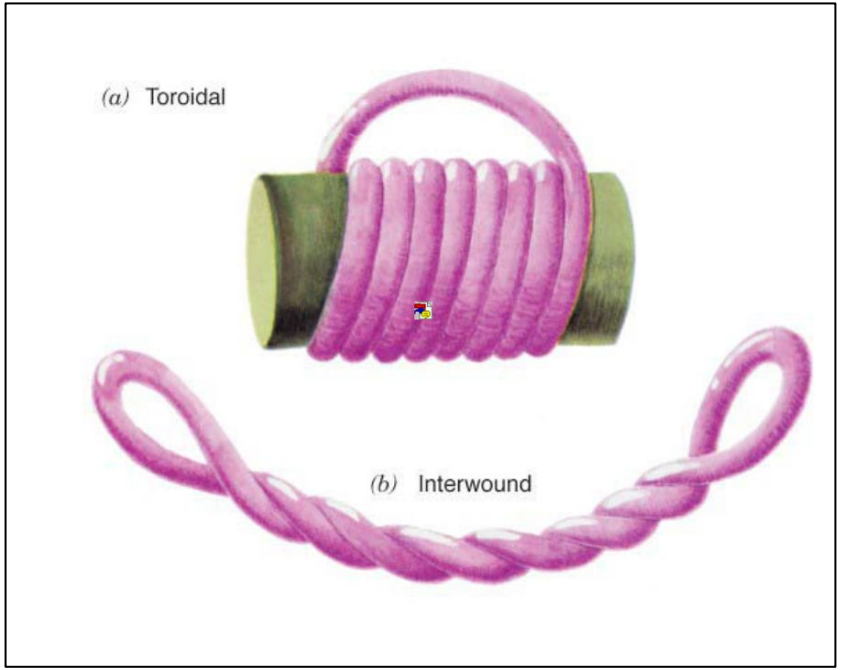
- When the DNA helix has the normal number of base pairs per helical turn it is in the **relaxed** state.
- If DNA is in the form of a circular molecule, or if the ends are rigidly held so that it forms a loop, then overtwisting or undertwisting leads to the supercoiled state.
- **Supercoiling** occurs when the molecule relieves the helical stress by twisting around itself. Overtwisting leads to **postive supercoiling**, while undertwisting leads to **negative supercoiling**
- Twist can be altered in a circular model by breaking the circle, over or undertwisting and then reconnecting the ends.

DNA compaction requires special form of supercoiling

A. Interwound: supercoiling of DNA in solution

B. Toroidal- tight left handed turns, packing of DNA

both forms are interconvertible



Methods for measuring supercoiling - based on how compact the DNA is

A. Gel electrophoresis
   i. 1 dimensional
   ii. 2 dimensional
B. Density sedimentation

. Topoisomerases are required to relieve torsional strain

A. Topoisomerases I :
   breaks only one strand
B. Topoisomerase II :
   breaks both strands

A. Topoisomerases I - breaks only one strand
   i. monomeric protein
   ii. after nicking DNA the 5'-PO4 is covalently linked to enzyme (prokaryotes)
   or the 3' end is linked to the enzyme (eukaryotes)
   iii. evidence is the formation of catenates
   iv. E. coli Topo I relaxes negatively supercoiled DNA
   v. introduces a change of increments of 1 in writhe

B. Topoisomerase II - breaks both strands
   i. supercoils DNA at the expense of ATP hydrolysis
   ii. two subunits: (alpha)2 and (beta)2
   iii. becomes covalently linked to the alpha subunit
   iv. relaxes both negative and positively supercoiled DNA
   v. introduces a change in increments of 2 in writhe.

Reversible denaturation and hyperchromic effect.

- **DNA denaturation** is a process of separating dsDNA into single strands
- **Different ways of denaturing DNA:**
  - Heat
  - Organic solvents such as dimethyl sulfoxide and formamide
  - high pH
  - Low salt concentration
- When a DNA solution is heated enough, the double-stranded DNA unwinds and the hydrogen bonds that hold the two strands together weaken and finally break.
- The process of breaking double-stranded DNA into single strands is known as DNA denaturation, or DNA denaturing.
- The temperature at which the DNA strands are half denatured, meaning half double-stranded, half single-stranded, is called the **melting temperature(Tm)**.

- The T_m of a DNA depends on
  - Base composition: G-C base pairs are stronger than A-T base pairs; therefore, DNAs with a high G+C content have a higher T_m than do DNAs with a higher A+T content.
  - Solvent composition: High ionic strength—for example, a high concentration of NaCl—promotes the double-stranded state (raises the T_m)
  - How well its bases match up: A synthetic DNA double strand made with some mismatched base-pairs has a lower T_m compared to a completely double-stranded DNA.

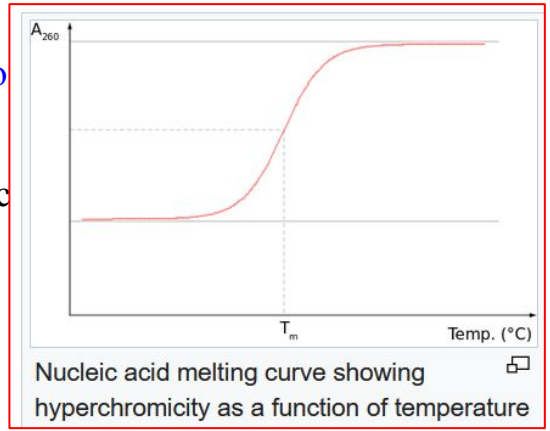
- The amount of strand separation, or melting, is measured by the absorbance of the DNA solution at 260nm.
- Nucleic acids absorb light at this wavelength because of the electronic structure in their bases, but when two strands of DNA come together, the close proximity of the bases in the two strands quenches some of this absorbance.
- When the two strands separate, this quenching disappears and the absorbance rises 30%-40%.This is called **Hyperchromicity.**
- The Hypochromic effect is the effect of stacked bases in a double helix absorbing less ultra-violet light.

- The DNA denaturation process is **reversible** under controlled conditions of pH and ionic strength.
- If the temperature is slowly decreased in the solution where the DNA had been denatured, the DNA chains will spontaneously reanneal and the original double helix structure is restored.
- The DNA renaturation resulting from slow cooling is called ***reannealing***. When the complementary strands meet, they completely reconstitute the double helix.
- The rate of renaturation depends on the structure of DNA.
    - When a given DNA has segments with the same sequence (repetitive sequences), the annealing time is shorter because the chance that one chain meets a complementary one is greater.
    - In contrast, DNA sections with unique sequences require a longer time to find its complementary strand to reform the double helix.

HYPERCHROMICITY:

- **Hyperchromicity** is the increase of absorbance (*optical density*) of a material.
- Hyperchromicity of DNA occurs when the DNA duplex is denatured.
- The UV absorption is increased when the two single DNA strands are being separated, either by heat or by addition of denaturant or by increasing the pH level.
- The opposite, a decrease of absorbance is called **hypochromicity**.
- Heat denaturation of DNA, also called melting, causes the double helix structure to unwind to form single stranded DNA. When DNA in solution is heated above its melting temperature (usually more than 80 °C), the double-stranded DNA unwinds to form single-stranded DNA.
- The bases become unstacked and can thus absorb more light. In their native state, the bases of DNA absorb light in the 260-nm wavelength region. When the bases become unstacked, the wavelength of maximum absorbance does not change, but the amount absorbed increases by 37%.

- Hyperchromicity can be used to track the condition of DNA as temperature changes.
- The **hyperchromic effect** is the striking increase in absorbance of DNA upon denaturation.
- The two strands of DNA are bound together mainly by the stacking interactions, hydrogen bonds and hydrophobic effect between the complementary bases.
- The hydrogen bond limits the resonance of the aromatic ring so the absorbance o the sample is limited as well.
- When the DNA double helix is treated with denatured agents, the interaction forc holding the double helical structure is disrupted.
- The double helix then separates into two single strands which are in the random coiled conformation.
- At this time, the base-base interaction will be reduced, increasing the UV absorbance of DNA solution because many bases are in free form and do not form hydrogen bonds with complementary bases.
- As a result, the absorbance for single-stranded DNA will be 37% higher than that for double stranded DNA at the same concentration.



Nucleic acid melting curve showing hyperchromicity as a function of temperature

REPLICATION OF DNA

- **DNA replication** is the biological process of producing two identical replicas of DNA from one original DNA molecule.
- DNA replication occurs in all living organisms acting as the most essential part for biological inheritance.
- This is essential for cell division during growth and repair of damaged tissues, while it also ensures that each of the new cells receives its own copy of the DNA
- DNA replication is **semiconservative**. Each strand in the double helix acts as a template for synthesis of a new, complementary strand.
- New DNA is made by enzymes called **DNA polymerases**, which require a template and a **primer** (starter) and synthesize DNA in the 5' to 3' direction.
- During DNA replication, one new strand (the **leading strand**) is made as a continuous piece. The other (the **lagging strand**) is made in small pieces.
- DNA replication requires other enzymes in addition to DNA polymerase, including **DNA primase**, **DNA helicase**, **DNA ligase**, and **topoisomerase**.

Structure and Function of DNA polymerases

DNA polymerase

- One of the key molecules in DNA replication is the enzyme **DNA polymerase**.
- DNA polymerases are responsible for synthesizing DNA: they add nucleotides one by one to the growing DNA chain, incorporating only those that are complementary to the template.
- Catalyses the synthesis of DNA from **5' to 3' direction.** It cannot act in opposite direction.

Key features of DNA polymerases:

- They always need a template
- They can only add nucleotides to the 3' end of a DNA strand
- They can't start making a DNA chain from scratch, but require a pre-existing chain or short stretch of nucleotides called a **primer**
- They **proofread**, or check their work, removing the vast majority of "wrong" nucleotides that are accidentally added to the chain

DNA polymerase has **3 distinct catalytic properties**:

1. **5'-3' polymerase activity**
2. **5'-3' exonuclease activity**
3. **3'-5' exonuclease activity**

5'-3' polymerase activity:

- For esterification of nucleotides to form DNA chain
- Causes addition of new nucleotides at 3'OH group of existing oligonucleotides
- For this activity it requires ssDNA called **template**. Hence it is called **DNA dependent DNA polymerase**
- It cannot initiate the synthesis of new DNA chain. It can only add new nucleotides to the existing molecule called **Primer.**
- DNA or RNA can act as primer for DNA Polymerase

5'-3' exonuclease activity:

- Used for removal of nucleotides from 5'end of DNA chain
- Primarily responsible for removal of primer from 5'end of newly synthesized DNA chain

3’-5’ exonuclease activity:

- Used for removal of nucleotides from 3’end of DNA chain
- This reaction is opposite to the polymerase function
- Mainly used for **proofreading**- removing the "wrong" nucleotides that are accidentally added to the chain

Structure of DNA Polymerase:

In prokaryotes

Types:

1. DNA Polymerase I
2. DNA Polymerase II
3. DNA Polymerase III

DNA Polymerase I:

- Not the primary enzyme for DNA synthesis
- Does not play important role in replication
- Its basic responsibility is in proof reading and DNA repair

DNA Polymerase II:

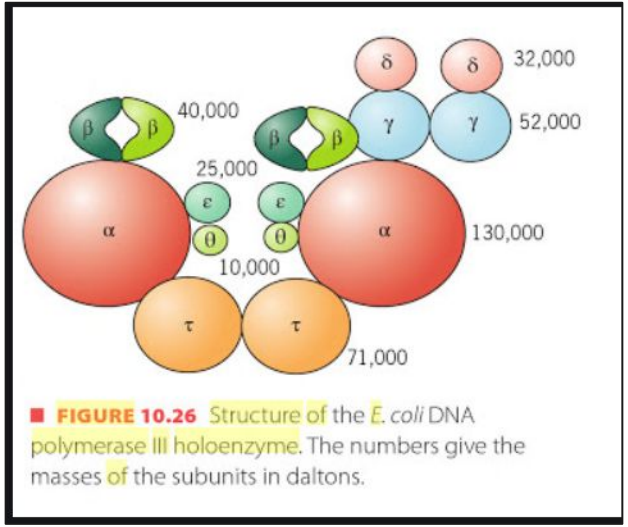
- Molecular weight 120kD
- Function is not well understood
- Should be involved in DNA repair

DNA Polymerase III:

- It is a hetero multimer of several subunits
- Has a catalytic core of 160kD. Which is made up of 3 subunits **α** (alpha), **ε**(epsilon) and *θ*(theta)
- **α** subunit: 130 kD: used for DNA synthesis
- **ε**: 25kD.
- *θ* subunit: 10kD: both are involved in proof reading activities
- 2 molecules of core enzyme condense together with the help of τ subunit which is essential for dimer formation
- Dimer is the active molecule referred as DNA Pol III of mol wt 470 kD
- To one of the **α** subunits, 2 γ (gamma) subunits and 2 □ (delta) subunits gets associated to form a asymmetric complex which can bind to template DNA
- **β** (Beta) subunit binds to asymmetric complex and assemble holoenzyme which is the complete pol III enzyme and can carry out all the functions of the enzyme

Role of different subunits of DNA POLYMERASE III:

- Alpha **α**: polymerizing activity
- Theta θ: assembly of DNA pol
- Epsilon **ε**: proof reading activity
- Tow τ: dimerization
- Gamma γ: template binding
- Delta □: increase processivity
- Beta **β** : template association



■ **FIGURE 10.26** Structure of the *E. coli* DNA polymerase III holoenzyme. The numbers give the masses of the subunits in daltons.

In eukaryotes:

3 types of DNA Polymerase III

- DNA pol **α**: main enzyme present in the nucleus, involved in replication
- DNA pol **β**: DNA Repair
- DNA pol γ: mitochondrial enzyme for replication of mitochondrial DNA

Other enzymes involved in DNA replication

DNA GYRASE or TOPOISOMERASE:

- Causes negative supertwist in DNA molecule
- It is used to unwind DNA so that 2 strands can open and form ss regions diring DNA replication
- It belongs to general class of type III DNA Topoisomerases
- Topoisomerases interconvert DNA to different topoisomers
- **Topoisomers** are different physical forms of same DNA which differ only in their linking number

2 types of Topoisomerases:

- **Type 1 Topoisomerases:**
- Can cut only one strand of DNA
- Create **Nick** in one strand by breaking phosphodiester bond anywhere in one strand
- Relieve supercoiling
- Does not require ATP

- **Type 2 Topoisomerases:**
- Can cut both the strands of DNA
- Can relieve supercoiling and untangle linked DNA helicases
- Requires ATP
- It makes covalent attachment to both DNA strands so that two circular DNA double helix are separated

DNA Ligases:

- It joins two pieces of DNA by the formation of phosphodiester bond between those molecules
- For this, they should have a free 3'OH group and a 5'PO4 groups respectively
- Hence DNA ligase can only close a **nick** but cannot incorporate new nucleotide to fill the gap
- A ***nick*** is a break in a phosphodiester bond in one strand of a double-helix.
- It is used to join polynucleotide fragments in **lagging strand** during DNA replication

DNA Primases:

- It is a specific DNA dependent RNA polymerase
- It is used for the synthesis of sequence specific RNA molecule which serve as **primer** to initiate DNA synthesis

Helicase:

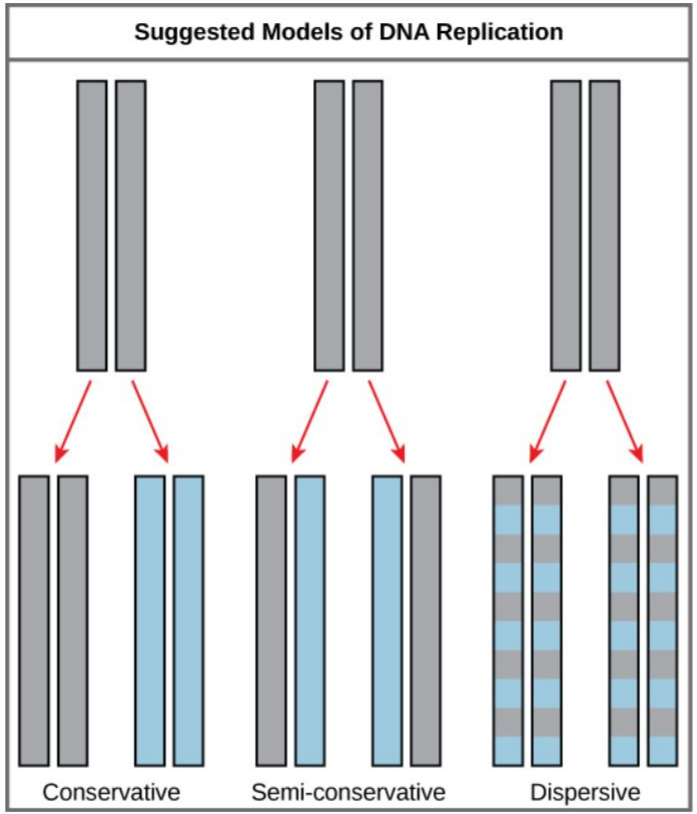
- It is referred as **unwinding protein**
- Used for melting of DNA and formation of ssDNA at the beginning of replication fork
- It is ATP dependent

SSB: Single Stranded DNA Binding Proteins:

- They are monomeric proteins
- They help DNA to remain single stranded after helicase had unwound DNA to ss from ds
- DNA will be staightened

Models of Replication in Prokaryotes

- There were **three models** suggested for DNA replication: **conservative, semi-conservative, and dispersive.**
- The conservative method of replication suggests that parental DNA remains together and newly-formed daughter strands are also together.
- The semi-conservative method of replication suggests that the two parental DNA strands serve as a template for new DNA and after replication, each double-stranded DNA contains one strand from the parental DNA and one new (daughter) strand.
- The dispersive method of replication suggests that, after replication, the two daughter DNAs have alternating segments of both parental and newly-synthesized DNA interspersed on both strands.
- Meselson and Stahl, using *E. coli* DNA made with two nitrogen istopes (^{14}N and ^{15}N) and density gradient centrifugation, determined that DNA replicated via the semi-conservative method of replication.



Suggested Models of DNA Replication: The three suggested models of DNA replication. Grey indicates the original parental DNA strands or segments and blue indicates newly-synthesized daughter DNA strands or segments.

- In conservative replication,
    - the two original DNA strands, known as the parental strands, would re-basepair with each other after being used as templates to synthesize new strands
    - the two newly-synthesized strands, known as the daughter strands, would also basepair with each other
    - one of the two DNA molecules after replication would be “all-old” and the other would be “all-new”.
- In semi-conservative replication,
    - each of the two parental DNA strands would act as a template for new DNA strands to be synthesized
    - after replication, each parental DNA strand would basepair with the complementary newly-synthesized strand just synthesized, and
    - both double-stranded DNAs would include one parental or “old” strand and one daughter or “new” strand.
- In dispersive replication,
    - after replication both copies of the new DNAs would somehow have alternating segments of parental DNA and newly-synthesized DNA on each of their two strands.

- To determine which model of replication was accurate, a seminal experiment was performed in 1958 by two researchers: Matthew Meselson and Franklin Stahl.
- Meselson and Stahl grew *E. coli* for several generations in a medium containing a "heavy" isotope of nitrogen (^{15}N) that is incorporated into nitrogenous bases and, eventually, into the DNA.
- The *E. coli* culture was then shifted into medium containing the common "light" isotope of nitrogen (^{14}N) and allowed to grow for one generation.
- The cells were harvested and the DNA was isolated.
- The DNA was centrifuged at high speeds in an ultracentrifuge in a tube in which a cesium chloride density gradient had been established.
- Some cells were allowed to grow for one more life cycle in ^{14}N and spun again.
- During the density gradient ultracentrifugation, the DNA was loaded into a gradient and spun at high speeds of 50,000 to 60,000 rpm. In the ultracentrifuge tube, the cesium chloride salt created a density gradient, with the cesium chloride solution being more dense the farther down the tube you went.

HYPOTHESIS: DNA replicates semiconservatively.

METHOD

1. Grow bacteria in ^{15}N (heavy) medium.
2. Transfer some bacteria to ^{14}N (light) medium; bacterial growth continues.
3. Take samples after 0 minutes, 20 minutes (after one round of replication), and 40 minutes (two rounds of replication).
4. Before the bacteria reproduce for the first time in the light medium (at 0 minutes), all DNA (parental) is heavy.

Sample at 0 minutes

Sample after 20 minutes

Sample after 40 minutes

$^{14}N/^{14}N$ (light) DNA

$^{14}N/^{15}N$ (intermediate) DNA

$^{15}N/^{15}N$ (heavy) DNA

Parent (all heavy)

First generation (all intermediate)

Second generation (half intermediate, half light)

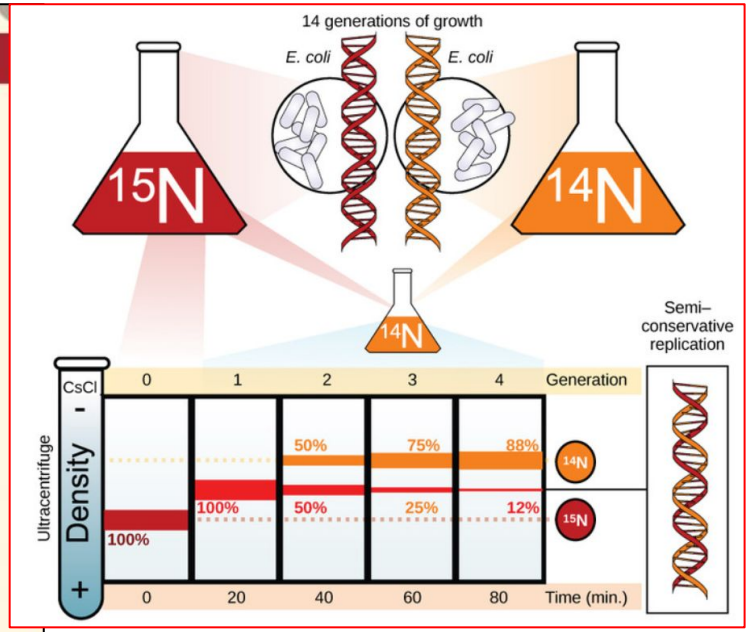
RESULTS

After 2 generations, half the DNA was intermediate and half was light only; there was no heavy-only DNA.

Parent strand ^{15}N

New strand ^{14}N

CONCLUSION: This pattern could only have been observed if each DNA molecule contains a template strand from the parental DNA; thus DNA replication is semiconservative.

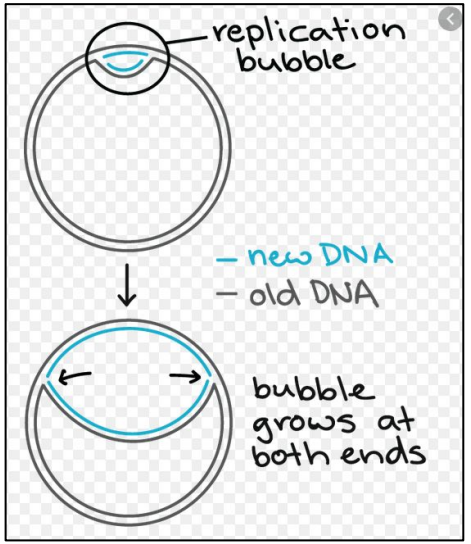


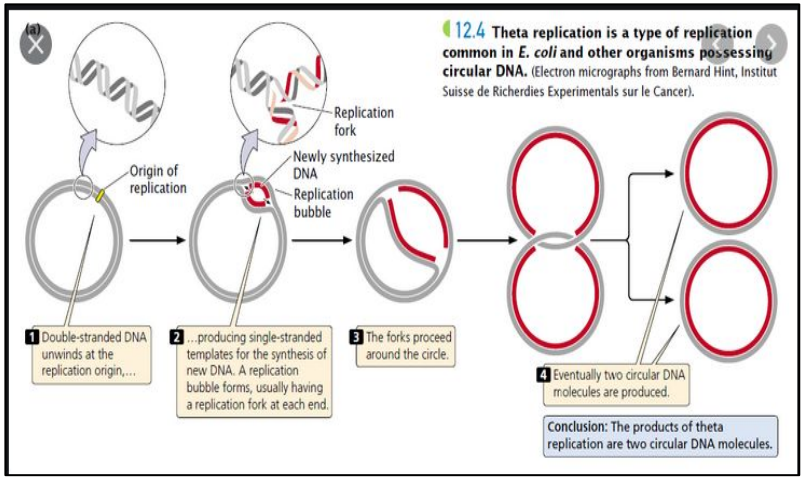
Prokaryotic DNA Replication:

- The process by which a prokaryote duplicates its DNA into another copy that is passed on to daughter cells.

1. **Theta mode of replication**

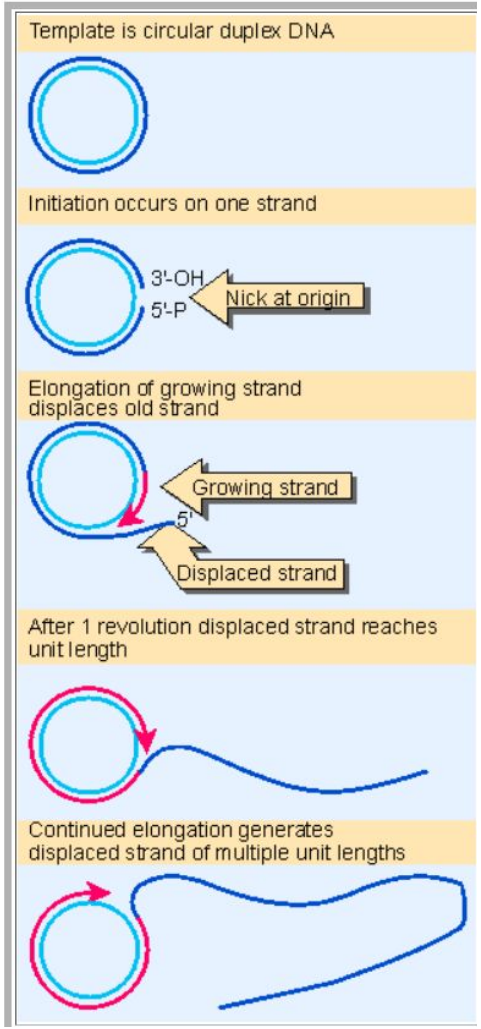
- Replication is bi-directional and originates at a single origin of replication (OriC).
- It consists of three steps: Initiation, elongation, and termination
- A replication bubble is formed at the OriC.
- The synthesis starts at both direction from the point
- Thus the shape of DNA looks similar to the latin letter theta **θ** at the beginning of replication
- This region continues to grow until entire DNA is duplicated
- This mode of replication of circular DNA is known as **theta mode of replication**





2. Rolling circle mode of replication

- Rolling circle is a mode of replication in which a replication fork proceeds around a circular template for an indefinite number of revolutions
- the DNA strand newly synthesized in each revolution displaces the strand synthesized in the previous revolution, giving a tail containing a linear series of sequences complementary to the circular template strand.
- Replication of only one strand is used to generate copies of some circular molecules.
- A nick opens one strand, and then the free 3′–OH end generated by the nick is extended by the DNA polymerase.
- The newly synthesized strand displaces the original parental strand.
- This particular process **occurs** in plasmid and virus's genome and in certain phages like M13 and ΦX174



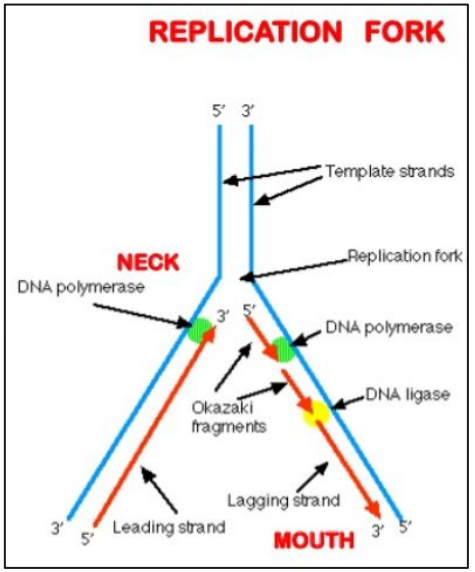
Mechanism of DNA Replication

- DNA replication employs a large number of proteins and enzymes, each of which plays a critical role during the process.
- One of the key players is the enzyme DNA polymerase, which adds nucleotides one by one to the growing DNA chain that are complementary to the template strand.
- The addition of nucleotides requires energy; this energy is obtained from the nucleotides that have three phosphates attached to them, similar to ATP which has three phosphate groups attached. When the bond between the phosphates is broken, the energy released is used to form the phosphodiester bond between the incoming nucleotide and the growing chain.
- **DNA replication occurs in three main stages**: initiation, elongation, and termination.
- **Site** of DNA replication**: origin of replication (ORI)**
    - Eukaryotes: many ORI
    - Prokaryotes: only one ORI
    - Sequences at ORI are A:T rich to make melting of DNA easier

Stage 1: Initiation

1. Formation of Replication Fork:

- It involves many protein factors
- Factor **DnaA** binds to ORI
- It may be preceded by binding of pre-priming factor **DnaT**
- Enzyme **Helicase** binds to ORI causing melting of small stretch of DNA by breaking H bonds
- Therefore DNA opens up in this region and 2 strands separate to form **Y-shaped** structure known as **Replication Fork**
- **SSB** proteins bind to separated strands to maintain the single stranded structure and prevent the reformation of dsDNA.
- SSB proteins are also aided by protein factor **DnaB**



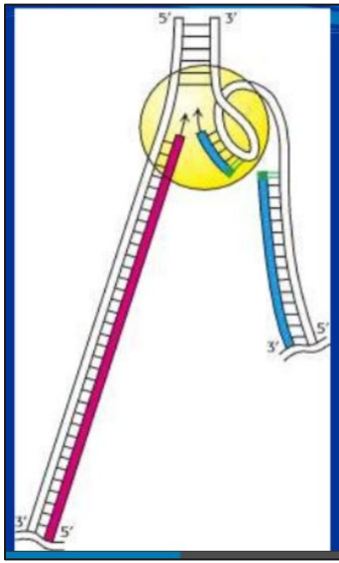
2.Primer binding:

- Since DNA polymerase can only add nucleotides to 3'end of an existing strand, the process requires **primers.**
- The enzyme **primase** synthesizes RNA primer which is complementary to 3'end of the template strand.
- Typical primer: **5 to 10 nucleotides long**
- The primer 'primes' the DNA synthesis i.e it gets it started

Thus a no. of factors get associated to DNA to initiate the replication. This entire complex is referred to as **Primosome.**

Stage 2: Elongation

- Once initiation takes place DNA polymerase adds more nucleotides to the primer- one at a time and continues to extend the DNA chain
- During this entire process, Helicase keeps binding ahead of the fork so that there is forward movement of the fork.
- The new nucleotide - ancillary factor complex of a growing chain is referred as **Replisome**
- DNA polymerase III can only make DNA in 5'-3' direction.
- One new strand, which runs 5'-3' direction towards the Replication fork is made continuously, because DNA polymerase is moving in the same direction as the replication fork. This strand is called: **Leading strand.**
- The other new strand which runs 5'-3' away from the replication fork is made in fragments. These small fragments are known as **Okazaki fragments.**
- This strand is made in discontinuous manner. It is called as **Lagging strand.**
- Enzyme DNA polymerase III has 2 catalytic cores: one core catalyses the leading strand synthesis and other core catalyses the lagging strand synthesis.
- At the point of enzyme action lagging strand folds back which changes the physical direction of the strand.
- Thus 2 strands grow in same physical direction but different biological direction and maintain 5'-3' orientation.



- **Leading strand** is extended from **one primer,** whereas **lagging strand** needs a new primer for each short okazaki fragments.
- Separate **primases** are required for leading and lagging strands. Hence **leading strand** is produced from **one primosome** and **lagging strand** is produced from **multiple primosomes.**
- **Okazaki fragments** are then joined together by the enzyme **DNA Ligase** to form a continuous strand.

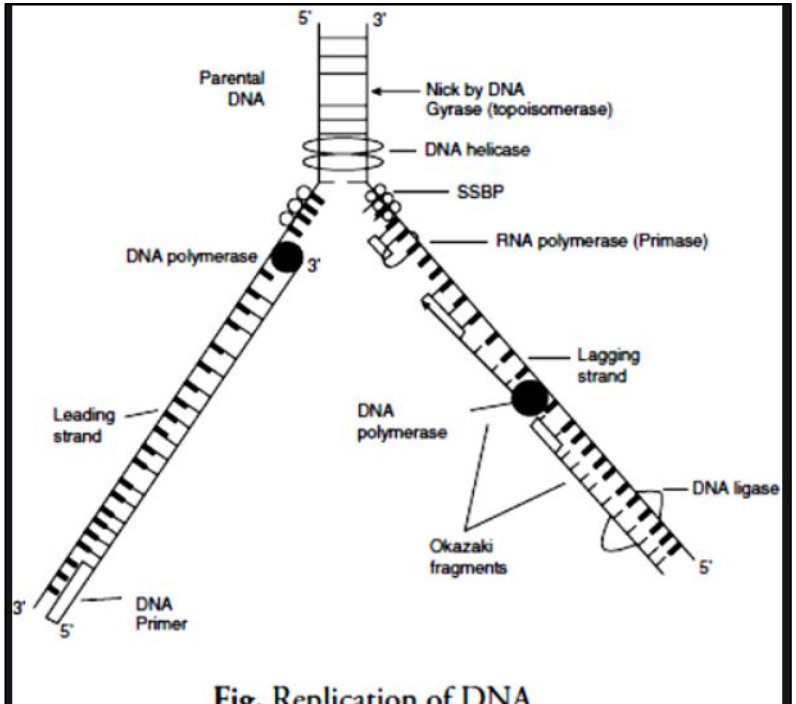
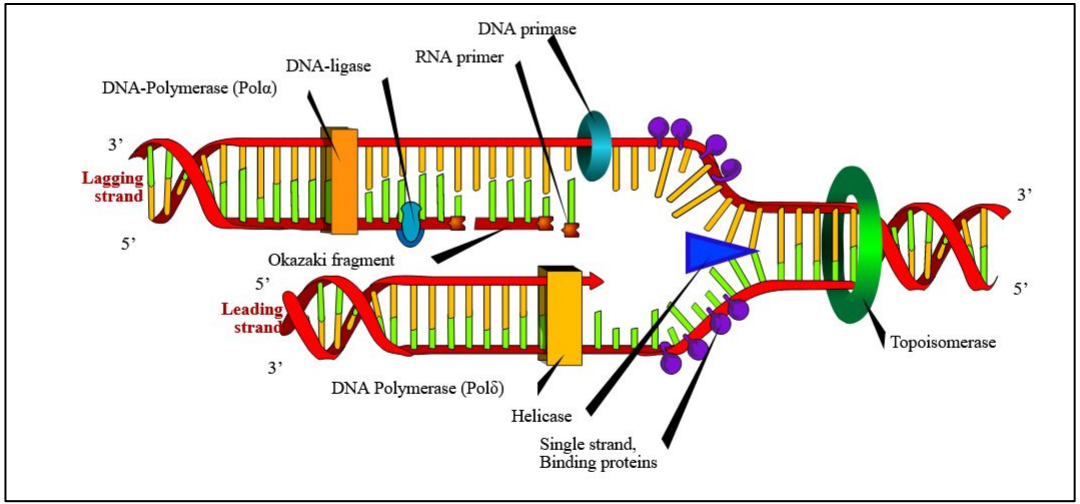
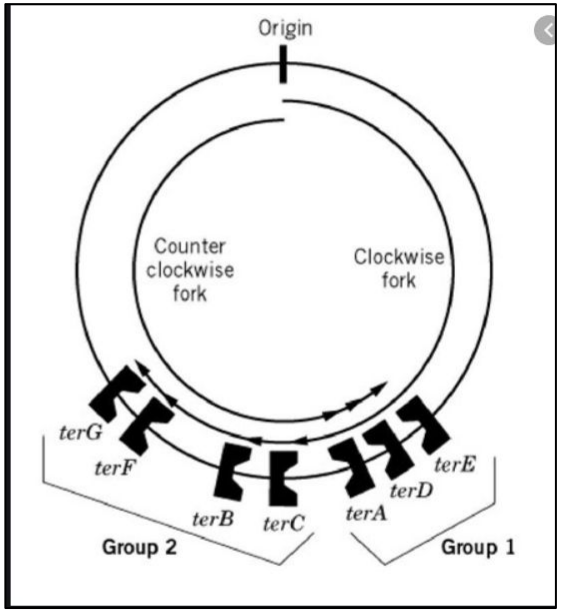


Fig. Replication of DNA



Stage 3: Termination

- **In prokaryotes** the process of replication is bidirectional from one origin
- It is terminated when the two replication forks moving in opposite directions from the origin meet at one point called ter 32.
- Ter-binding proteins - will recognizes the Termination sequences and helps to achieve the termination process.
- All the primers will be removed, and all the fragments will be connected by DNA-pol I and ligase.
- In E.coli replication of circular DNA takes about 30 minutes.



- **In eukaryotes** replication is terminated when entire DNA is duplicated in S phase of cell cycle.
- Eukaryotic chromosomes have multiple origins of replication, which initiate replication almost simultaneously. Each origin of replication forms a bubble of duplicated DNA on either side of the origin of replication.
- DNA polymerase halts when it reaches a section of DNA template that has already been replicated. However, DNA polymerase cannot catalyze the formation of a phosphodiester bond between the two segments of the new DNA strand, and it drops off.

- Once all the template nucleotides have been replicated, the replication process is not yet over. RNA primers need to be replaced with DNA, and nicks in the sugar-phosphate backbone need to be connected.
- Once all the bases are copied, the **5'-3' exonuclease activity of enzyme DNA polymerase removes the primer**. The gap where primers were, are filled by complimentary nucleotides.
- In the final stage of DNA replication, the **enyzme ligase joins the sugar-phosphate backbones at each nick site**. After ligase has connected all nicks, the new strand is one long continuous DNA strand, and the daughter DNA molecule is complete.
- The **new strand is proof read by 3'-5' exonuclease activity of DNA polymerase** to make sure there are no mistakes in the new DNA sequence

Fidelity of DNA Replication

- The fidelity of a DNA polymerase refers to its ability to accurately replicate a template.
- It was estimated that both in prokaryotic and eukaryotic cells, DNA is replicated with the very high fidelity with one wrong nucleotide incorporated once per 10^8–10^{10} nucleotides polymerized.
- The fidelity of DNA replication relies on
  - Nucleotide selectivity of replicative DNA polymerase (Choosing the correct nucleotide)
  - Exonucleolytic proofreading, and
  - Post replicative DNA mismatch repair (MMR).
- The rate of misincorporation (incorporating the incorrect nucleotide) is known as the polymerase's "error rate".
- Proofreading activity is responsible for removal of incorrectly incorporated nucleotides from the primer terminus before further primer extension. It is estimated that proofreading improves the fidelity by a 2–3 orders of magnitude.
- An appropriate level of fidelity during DNA replication ensures the ability of organisms to transfer genetic information from one generation to the next and contributes to the diversity of life.
- It is essential that the semiconservative duplication of DNA gives a nearly perfect end product, otherwise important genes might carry mutations that lead to disease or cell death.

- High fidelity DNA synthesis is beneficial for maintaining genetic information over many generations and for avoiding mutations that can initiate and promote human diseases such as cancer and neurodegenerative diseases.
- Low fidelity DNA synthesis is beneficial for the evolution of species, for generating diversity leading to increased survival of viruses and microbes when subjected to changing environments, and for the development of a normal immune system.

Proteomics of DNA replication

https://www.slideshare.net/sarathy4/enzymes-and-proteins-in-dna-replication

Inhibitors of DNA replication

- Replication of DNA is vital for life; inhibition of replication prevents cell division
- Inhibitors of eukaryotic replication can be used as **anti-cancer drugs** They will check the multiplication of cancer cells :
  - Cisplatin produces intra-strand purine cross-links and prevents replication
  - Mitomycin C produces intra-strand G−G and inter-strand C−G cross-links, and blocks replication
  - Daunorubicin intercalates in DNA and prevents access to DNA polymerase
- Inhibitors acting selectively in prokaryotes can be used as **antibiotics.** Antibiotics exploit the differences in prokaryotic and eukaryotic machinery. If a compound inhibits a prokaryotic enzyme but not its eukaryotic counterpart, it can be used as an antibiotic
- Eg: Floxacin family of drugs (e.g. norfloxacin, ciprofloxacin etc) inhibits DNA gyrase but not DNA topoisomerase II.
- DNA gyrase is present in prokaryotes only. The corresponding human enzyme is DNA topo-isomerase II
- Topoisomerase II (DNA gyrase) inhibitor
  - Novobiocin– prevents ATP binding to gyrase
  - Nalidixic acid & Ciprofloxacin- interfere with the breakage and rejoining of DNA chains
  - Camptothecin– inhibits human topoisomerase I

- Nucleotide analogues
  - 2,3 dioxyinosine, Cytarabine Zidovudine, Acyclovir- terminate DNA chain elongation
- Inhibitors of DNA; Prevents un-winding of DNA. E.g. actinomycin, mitomycin
- Inhibitors of deoxy-ribonucleotides; E.g. Anti-folates [inhibits Purine Pyrimidine synthesis]
- Inhibitors of replicative enzymes; E.g. norflox [inhibit DNA gyrase] ciploflox