RAMA UNIVERSITY

www.ramauniversity.ac.in

FACULTY OF AGRICULTURE SCIENCES AND ALLIED INDUSTRIES

(Principles of Biotechnology)

For

M.Sc. Ag (GPB)

Course Instructor

Dr Shiv Prakash Shrivastav

FASAI(Genetics and Plant Breeding)

Rama University, Kanpur

DNA AND IT' S STRUCTURE, FUNCTION, TYPES, MODES OF REPLICATION AND REPAIR

The discovery that DNA is the prime genetic molecule, carrying all the hereditary information within chromosomes, immediately had its attention focused on its structure. It was hoped that knowledge of the structure would reveal how DNA carries the genetic messages that are replicated when chromosomes divide to produce two identical copies of themselves. During the late 1940s and early 1950s, several research groups in the United States and in Europe engaged in serious efforts—both cooperative and rival—to understand how the atoms of DNA are linked together by covalent bonds and how the resulting molecules are arranged in three-dimensional space. Not surprisingly, it was feared that DNA might have very complicated and perhaps bizarre structures that differed radically from one gene to another. Great relief, if not general elation, was thus expressed when the fundamental DNA structure was found to be the double helix. It told us that all genes have roughly the same three-dimensional form and that the differences between two genes reside in the order and number of their four nucleotide building blocks along the complementary strands.

What is DNA?

The work of many scientists paved the way for the exploration of DNA. Way back in 1868, almost a century before the Nobel Prize was awarded to Watson, Crick and Wilkins, a young Swiss physician named Friedrich Miescher, isolated something no one had ever seen before from the nuclei of cells. He called the compound "nuclein." This is today called nucleic acid, the "NA" in DNA (deoxyribo-nucleic-acid) and RNA (ribo-nucleic-acid).

Two years earlier, the Czech monk Gregor Mendel, had finished a series of experiments with peas. His observations turned out to be closely connected to the finding of nuclein. Mendel was able to show that certain traits in the peas, such as their shape or colour, were inherited in different packages. These packages are what we now call genes.

For a long time the connection between nucleic acid and genes was not known. But in 1944 the American scientist Oswald Avery managed to transfer the ability to cause disease from one strain of bacteria to another. But not only that: the previously harmless bacteria could also pass the trait along to the next generation. What Avery had moved was nucleic acid. This proved that genes were made up of nucleic acid.

Solving the Puzzle

In the late 1940's, the members of the scientific community were aware that DNA was most likely the molecule of life, even though many were skeptical since it was so "simple". They also knew that DNA included different amounts of the four bases adenine, thymine, guanine and cytosine (usually abbreviated A, T, G and C), but nobody had the slightest idea of what the molecule might look like.

In order to solve the elusive structure of DNA, a couple of distinct pieces of information needed to be put together. One was that the phosphate backbone was on the outside with bases on the inside; another that the molecule was a double helix. It was also important to figure out that the two strands run in opposite directions and that the molecule had a specific base pairing.

Watson and Crick

In 1951, the then 23-year old biologist James Watson travelled from the United States to work with Francis Crick, an English physicist at the University of Cambridge. Crick was already using the process of X-ray crystallography to study the structure of protein molecules. Together, Watson and Crick used X-ray crystallography data, produced by Rosalind Franklin and Maurice Wilkins at King's College in London, to decipher DNA's structure.

This is what they already knew from the work of many scientists, about the DNA molecule:

1. DNA is made up of subunits which scientists called nucleotides.
2. Each nucleotide is made up of a sugar, a phosphate and a base.

3. There are 4 different bases in a DNA molecule:
   adenine (a purine)
   cytosine (a pyrimidine)
   guanine (a purine)
   thymine (a pyrimidine)
4. The number of purine bases equals the number of pyrimidine bases
5. The number of adenine bases equals the number of thymine bases
6. The number of guanine bases equals the number of cytosine bases
7. The basic structure of the DNA molecule is helical, with the bases being stacked on top of each other

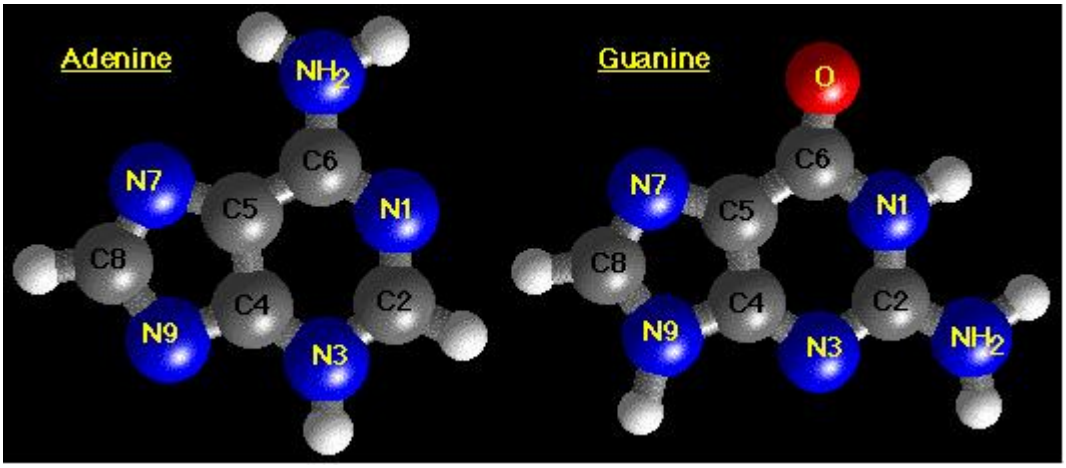
Components of DNA

DNA is a polymer. The monomer units of DNA are nucleotides, and the polymer is known as a "polynucleotide". Each nucleotide consists of a 5-carbon sugar (deoxyribose), a nitrogen containing base attached to the sugar, and a phosphate group. There are four different types of nucleotides found in DNA, differing only in the nitrogenous base. The four nucleotides are given one letter abbreviations as shorthand for the four bases.

- A is for adenine
- G is for guanine
- C is for cytosine
- T is for thymine

Purine Bases

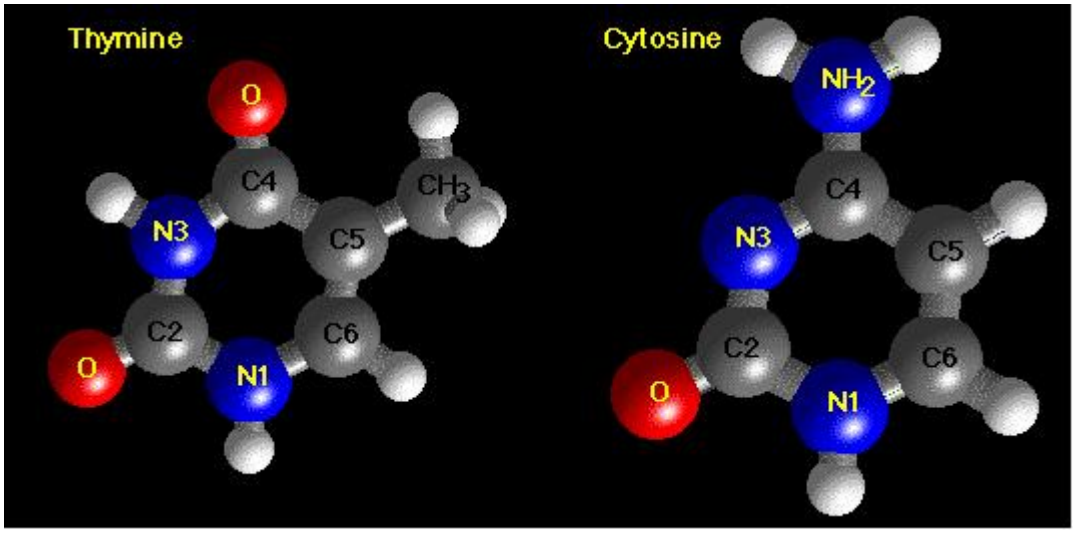
Adenine and guanine are purines. Purines are the larger of the two types of bases found in DNA. Structures are shown below:



The 9 atoms that make up the fused rings (5 carbon, 4 nitrogen) are numbered 1-9. All ring atoms lie in the same plane.

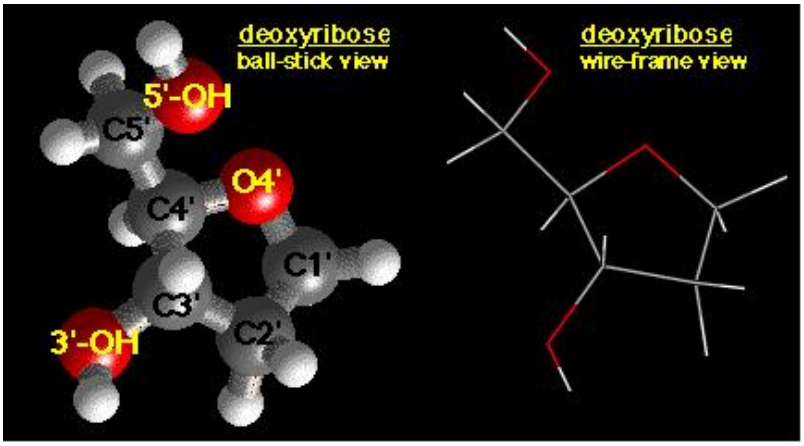
Pyrimidine Bases

Cytosine and thymine are pyrimidines. The 6 stoms (4 carbon, 2 nitrogen) are numbered 1-6. Like purines, all pyrimidine ring atoms lie in the same plane.



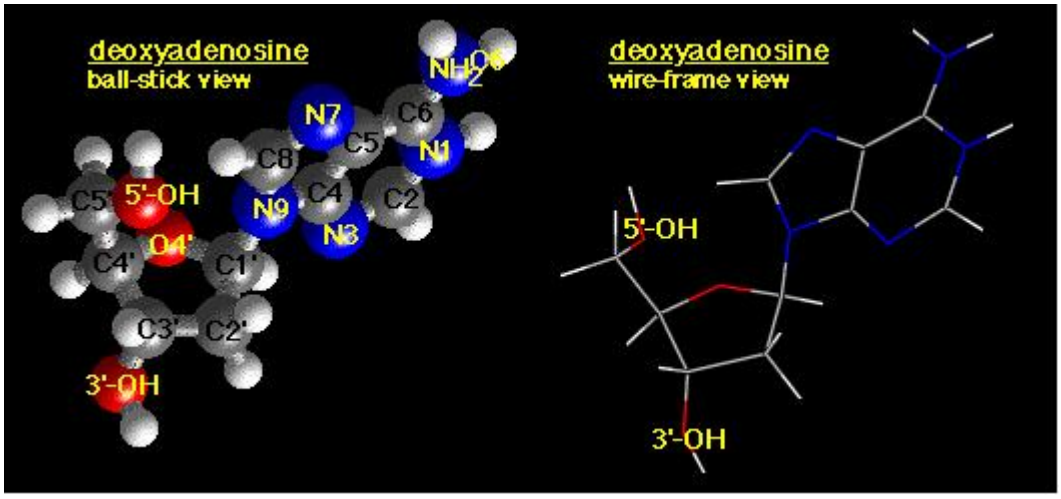
Deoxyribose Sugar

The deoxyribose sugar of the DNA backbone has 5 carbons and 3 oxygens. The carbon atoms are numbered 1', 2', 3', 4', and 5' to distinguish from the numbering of the atoms of the purine and pyrmidine rings. The hydroxyl groups on the 5'- and 3'- carbons link to the phosphate groups to form the DNA backbone. Deoxyribose lacks an hydroxyl group at the 2'-position when compared to ribose, the sugar component of RNA.



Nucleosides

A nucleoside is one of the four DNA bases covalently attached to the C1' position of a sugar. The sugar in deoxynucleosides is 2'-deoxyribose. The sugar in ribonucleosides is ribose. Nucleosides differ from nucleotides in that they lack phosphate groups. The four different nucleosides of DNA are deoxyadenosine (dA), deoxyguanosine (dG), deoxycytosine (dC), and (deoxy)thymidine (dT, or T).



In dA and dG, there is an "N-glycoside" bond between the sugar C1' and N9 of the purine.

Nucleotides

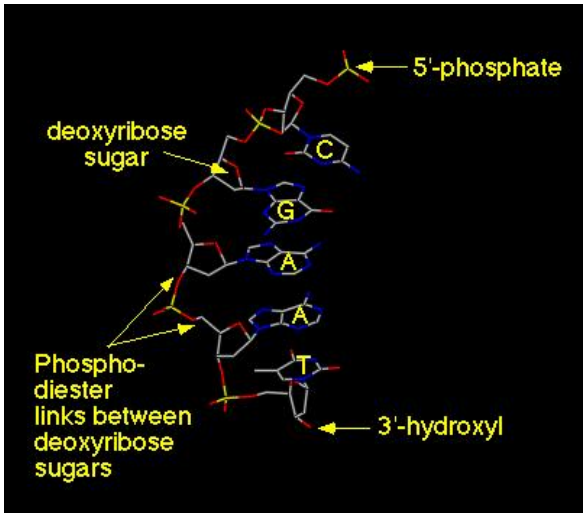
A nucleotide is a nucleoside with one or more phosphate groups covalently attached to the 3'- and/or 5'-hydroxyl group(s).

DNA Backbone

The DNA backbone is a polymer with an alternating sugar-phosphate sequence. The deoxyribose sugars are joined at both the 3'-hydroxyl and 5'-

hydroxyl groups to phosphate groups in ester links, also known as "phosphodiester" bonds.

Example of DNA Backbone: 5'-d (CGAAT)



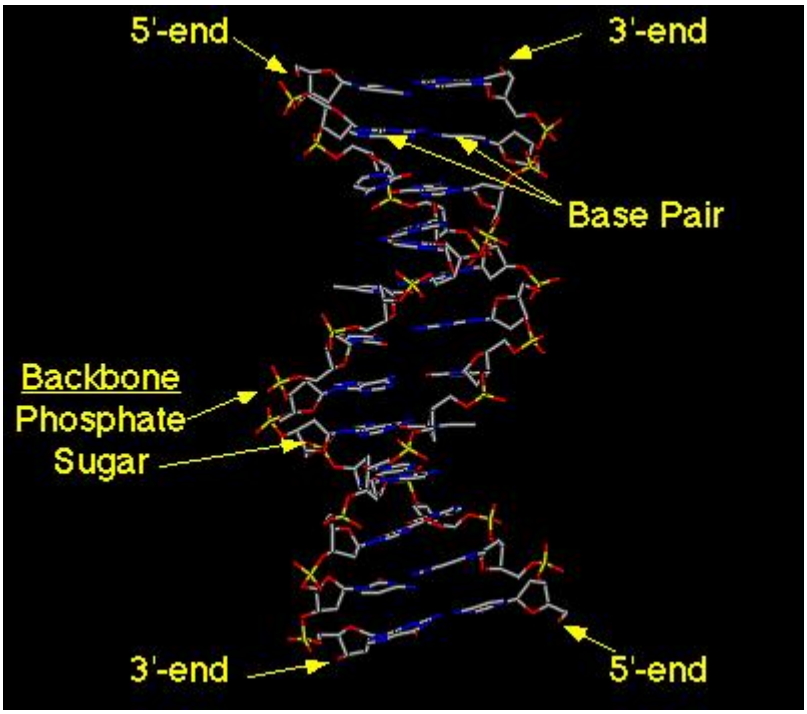
Features of the 5'-d(CGAAT) structure:

- Alternating backbone of deoxyribose and phosphodiester groups
- Chain has a direction (known as polarity), 5'- to 3'- from top to bottom
- Oxygens (red atoms) of phosphates are polar and negatively charged
- A, G, C, and T bases can extend away from chain, and stack atop each other
- Bases are hydrophobic

DNA Double Helix

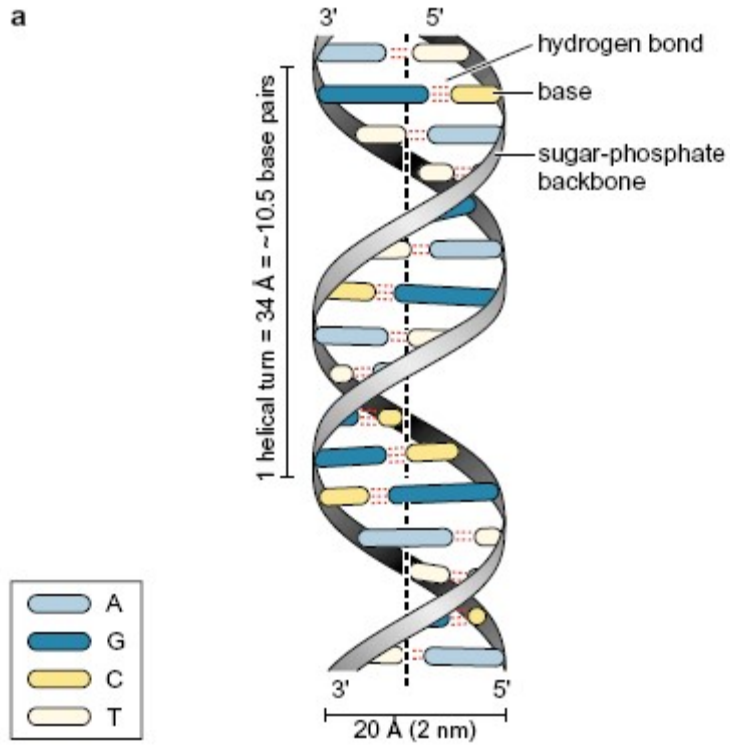
DNA is a normally double stranded macromolecule. Two polynucleotide chains, held together by weak thermodynamic forces, form a DNA molecule.

Structure of DNA Double Helix



Features of the DNA Double Helix

- Two DNA strands form a helical spiral, winding around a helix axis in a right-handed spiral
- The two polynucleotide chains run in opposite directions
- The sugar-phosphate backbones of the two DNA strands wind around the helix axis like the railing of a sprial staircase
- The bases of the individual nucleotides are on the inside of the helix, stacked on top of each other like the steps of a spiral staircase.

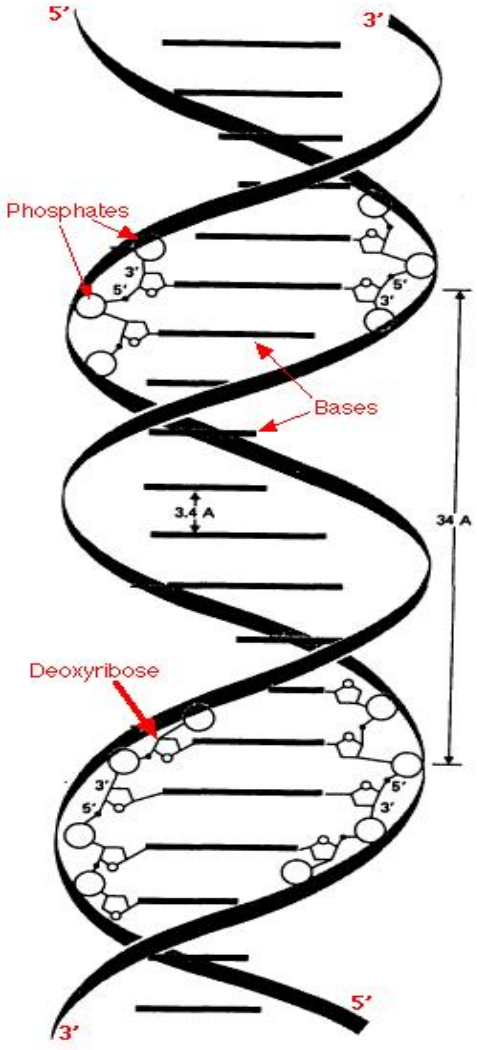


The Double Helix

The double helix of DNA has these features:

- It contains two polynucleotide strands wound around each other.
- The backbone of each consists of alternating deoxyribose and phosphate groups.
- The phosphate group bonded to the 5' carbon atom of one deoxyribose is covalently bonded to the 3' carbon of the next.
- The two strands are "antiparallel"; that is, one strand runs 5′ to 3′ while the other runs 3′ to 5′.
- The DNA strands are assembled in the 5′ to 3′ direction and, by convention, we "read" them the same way.
- The purine or pyrimidine attached to each deoxyribose projects in toward the axis of the helix.

- Each base forms hydrogen bonds with the one directly opposite it, forming **base pairs** (also called nucleotide pairs).
- 3.4 Å separate the planes in which adjacent base pairs are located.
- The double helix makes a complete turn in just over 10 nucleotide pairs, so each turn takes a little more (35.7 Å to be exact) than the 34 Å shown in the diagram.
- There is an average of 25 hydrogen bonds within each complete turn of the double helix providing a stability of binding about as strong as what a covalent bond would provide.
- The diameter of the helix is 20 Å.
- The helix can be virtually any length; when fully stretched, some DNA molecules are as much as 5 cm (2 inches!) long.
- The path taken by the two backbones forms a major (wider) groove (from "34 A" to the top of the arrow) and a minor (narrower) groove (the one below).



Nucleic acids (DNA and RNA) are the polymers i.e. long chain compounds. The molecular structure of DNA has two aspects

1) its chemical sub units and

2) the way in which these chemical sub units are arranged to form a long chain molecule.

The second aspect is very significant as the accepted DNA model should be such that it explains biochemically the various aspects (function) of a gene such as stability to metabolic and external agents, the capacity for replication (self duplication) the capacity to store vast hereditary information in coded form and the capacity to express the phenotypes they control.

FUNCTIONS OF DNA

DNA carries the genetic information of a cell and consists of thousands of genes. Each gene serves as a recipe on how to build a protein molecule. Proteins perform important tasks for the cell functions or serve as building blocks. The flow of information from the genes determines the protein composition and thereby the functions of the cell.

The DNA is situated in the nucleus, organized into chromosomes. Every cell must contain the genetic information and the DNA is therefore duplicated before a cell divides (**replication**). When proteins are needed, the corresponding genes are transcribed into RNA (**transcription**). The RNA is first processed so that non-coding parts are removed (**processing**) and is then

transported out of the nucleus (**transport**). Outside the nucleus, the proteins are built based upon the code in the RNA (**translation**).

Types of DNA

DNA can be classified in various ways based on 1. number of base pair per turn. 2. coiling pattern, 3. location 4. structure, 5. nucleotide sequence and 6. number of strands.

1. Number of base per turn. Depending upon the nucleotide base per turn of the helix, tilt of the base pair and humidity of the sample, the DNA can be observed in four different forms namely A,B, C and D.

2. Coiling pattern. On the basis of coiling pattern of the helix DNA is of two types viz right handed and left handed. Most of the DNA molecules are right handed i.e. coiling of helix is in the right direction. It is also called positive coiling. All the four forms of DNA viz A, B, C and D are right handed. The Z DNA has left handed double helical structure. This DNA is considered to be associated with gene regulation.

3. Location. Based on the location in the cell DNA is of three types. Viz., chromosomal DNA cytoplasm DNA and promiscuous DNA. Chromosomal DNA is found in chromosomes. And are called as chromosomal DNA or nuclear DNA. Cytoplasmic DNA is found in the cytoplasm especially in mitochondria and chloroplasts. Such DNA plays an important role in cytoplasmic inheritance and has circular structure. Promiscuous DNA. Some DNA segments with common base sequence are found in the chloroplasts, mitochondria and nucleus. This suggests that some DNA sequences move from one organelle to other. Such DNA is referred to as promiscuous DNA.

4. Structure of RNA: It contains ribose sugar, nitrogen bases and phosphate group. The nitrogen bases include adenine, guanine, cytosine and uracil. In DNA thymine is present in place of uracil and deoxyribose sugar is found in place of ribose sugar. In RNA, the pairing occurs between adenine and uracil. It has usually single strand. However, some viruses have double stranded RNA.

The DNA molecule that Watson and Crick described was in the B form. It is now known that DNA can exist in several other forms. The primary difference between the forms is the direction that the helix spirals.

A, B, C = right-handed helix Z = left-handed helix (found in vitro under high salt)

B is the major form that is found in the cell. Z-DNA was initially found only under high salt conditions, but the cellular environment is actually a low-salt environment. The question then is whether type Z exist under cellular conditions. Several features have been discovered that can stablize Z-DNA under in a low salt environment.

Differences between DNA and RNA

S. No	Particulars	DNA	RNA
1.	Strands	Usually two, rarely one	Usually one, rarely two
2.	Sugar	Deoxyribose	Ribose
3.	Base	Adenine guanine cytosine and thymine	Adenine guanines cytosine
4.	Pairing	AT and GC	AU and GC
5.	Location	Mostly in chromosomes some in mitochondria and chloroplasts	In chromosomes and ribosomes

MODES OF REPLICATION

There are three possible modes of DNA replication:

(1) Dispersive

(2) Conservative

(3) Semiconservative

1. **In dispersive replication**, the old DNA molecule would break into several pieces, each fragment would replicate and the old and new sesgments would

recombine randomly to yield progeny DNA molecules. Each progeny molecule would have both old and new segments along its length.

2. According to the **conservative scheme**, the two newly synthesized strands (following the replication of a DNA molecule) would associate to form one double helix, while the two old strands would remain together as one double helix.

3. In contrast, in the **semi conservative mode** of DNA replication, each newly synthesized strand would remain associated with the old strand against which it was synthesized. Thus each progeny DNA molecule would consist of one old and one newly synthesized strand.

Semi Conservative Replication

The semi conservative mode of DNA replication was postulated by Watson and Crick along with the double helix model of DNA. The main features of this mode of DNA replication are as follows:

1. A progressive separation of the two strands of a DNA molecule.

2. Complementary base pairing of the bases located in the single stranded regions thus produced with the appropriate free deoxyribonulceotides.

3. Formation of phosphodiester linkages between the neighbouring deoxyribonucleotides that have base paired with the single stranded regions, thereby producing regions the new strand.

4. This ensures that the base sequence of the new strands are strictly complementary top those of the old strands.

5. The base sequence of a newly synthesized strand is dictated by the base sequence of the old strand, since the old strand serves as a template or lould for the synthesis of the new strand.

DNA Replication

It is proposed by Watson and Crick. According to this method, both the strands of parental DNA separate from one another. Each old strand

synthesizes a new strand. Thus, each of the two resulting DNA has one parental and one new strand. This method of DNA replication is universally accepted because there are several evidences in support of semi conservative method and it consists of several steps.

1. Initiation of Replication DNA replication starts at a specific point on the chromosome. This unique site is known as origin. The site of initiation differs from organism to organism. Sometime replication starts with an incision made by an incision enzyme known as endonuclease.

2. Unwinding of strands. The two stands of DNA double helix unwind. The opening of DNA stands take's places with the help of DNA unwinding protein.

3. Formation of RNA Primer. Synthesis of RNA primer is essential for initiation of DNA synthesis RNA primer is synthesized by the DNA template near the origin with the help of a special type of RNA polymerase.

4. Synthesis of DNA on primer. After formation of RNA primer, DNA synthesis starts on the RNA primer. Deoxyribose nucleotides are added to the 3e end position of RNA primer. The main DNA strand is synthesized on the DNA template with help of DNA polymerase. The DNA synthesis takes place in short pieces. Which are known as Okazaki fragments.

5. Removal of RNA Primer: DNA polymerase degrades the RNA primer 1. This enzyme also catalyzes the synthesis of short DNA segment to replace the prime. The newly synthesized segment is joined to the main DNA strand with the help of DNA ligase enzyme.

6. Union of Okazaki Fragments. The discontinuous fragment of Okazaki is joined to make continuous strands. The union of Okazaki fragments takes place with the help of a joining enzyme called polynucleotide ligase. The replication may take place either in one direction or in both the directions from the point of origin.

Evidence for semi conservation replication

Various experiments have demonstrated the semi-conservative mode of DNA replication. Now it is universally accepted that DNA replicates in a semi-

conservative manner. There are three important experiments, which support that DNA replication is semi-conservative. These include (1) Meselson and Stahl experiment (2) Cairns experiment and (3) Taylor.s experiment.

Taylor.s experiment: Taylor (1969) conducted his experiments with root tip cells of *vicia faba*. He treated root tips with radioactive thymidine to label the DNA. The root tips were grown in the normal medium. In the first generation both chromatids were labeled. In the second generation of cell division, one chromatid of each chromosome was labeled and the other one was normal. This demonstrated semi conservative mode of chromosome replication. The DNA replication is associated with chromosome replication.

Enzymes involved in DNA / RNA replication

DNA replication involves several proteins and enzymes, which together form the multienzymes complex, rep0lication apparatus or replisome. In E coli at lest two dozen gene products are involved in DNA replication. Many of these protein were first identified through studies of mutants e.g. Genes dna E, dna N, dna x etc of E colic code for the four of the seven polypeptides of the complete DNA polymerase III enzyme, and DNA G specifies the primase enzyme. Some enzymes like ligase, DNA polymerase 1 etc were discovered biochemically.

DNA repair systems

Damages to the genetic material, i.e., DNA are taken care of by the DNA repair systems. The various damages to DNA may be grouped into the following two types:

(1) Single base changes: Such changes affect a single base of a DNA molecule they do not produce structural distortions and do not affect either replication or transcription of the affected molecules. These changes ar represented by the conversion of one base into another, eg; deamination of 5 methylcytosine results in thymine and by the covalent addition of a small group to a base which affects its pairing behavior. As a result, the affected base does not pair properly with its partner base.

(2) Structural distortations: These changes generally adversely affect the replication and or transcription of the affected DNA molecule. They are represented by a single strand nick, removal of a base, covalent links between bases in the same or in the opposite strands (eg) Pyrimidine dimmers and addition of a bulky adduct to a base which may distort the configuration of the double helix.

The repair systems recognize a variety of changes in DNA to initiate action. Each cell possesses several repair systems in order to be able to deal with the various types of DNA damage; these systems may be grouped into the following general categories

1. Direct repair
2. Excision repair
3. Mismatch repair
4. Tolerance systems
5. Retrieval systems

1. Direct repair of DNA

The reversal or simple removal of the damage to the DNA is known as direct repair, eg., removal of the covalent bonds between the two 4 and two 5 carbons of the two thymine residues participating in the formation of thymine dimmers. Thymine dimers are generally formed due to UV radiation and interfere with replication and transcription. A specific enzyme mediates the splitting of the covalent bonds between the two T residues, which specifically recognizes to thymine dimmers. The enzyme can bind to the thymine dimmers in the dark, but requires the energy from blue light for removal of the covalent bonds between the T residues; that is why this process is known as photoreactivation. The direct repair system is wide spread in nature and is especially important in plants.

2. Excision repair

In this repair pathway, the damaged or mispaired segment of the DNA strand is exercised and new stretch of DNA is synthesized in its place. The

various excision repair systems vary in their specificity. The repair process consists of the following steps:

a. Recognition and incision: The damaged section of a strand recognized by an endonuclease; this enzyme then cuts the affected strand on both the sides of damage.

b. Excision: After the incision, a 5’ to 3 ‘ exonulcease digests away the damage/ mispaired section; this generates a single stranded region in the DNA double helix.

c. Synthesis: In this step, the single stranded region produced by excision serves as a template for a DNA polymerase which synthesis the replacement for the excised segment. DNA ligase then seals the nick that remains after the synthesis of the replacement for the excised section.

3. Mismatch repair: When single bases in the DNA are mismatched, either due to alterations in the existing bases or due to errors during replication, structural distortions result in the DNA double helix.

4. Tolerance systems: These systems deal with the damages that block normal replication at the damaged sites possibly by permitting the replication of the damaged sites possibly with a high frequency of errors. These systems may be particularly important in the eukaryotes where the genome size is very large and hence a complete repair of the damage is rather unlikely.

5. Retrieval systems: These systems are also known as post replication repair or recombination repair.