THE MOLECULAR BASIS OF HEREDITY AND VARIABILITY

Identification of DNA as the genetic material

In 1869, *Friedrich Miescher*, a Swiss physician, isolated a substance from cell nuclei that he called nuclein - now known as DNA.

Griffith's experiment

In 1928, *Fredrick Griffith* conducted an experiment that led to the discovery of DNA as the genetic material. Griffith studied two strains of Streptococcus pneumonia. One strain had a sugar coat and caused pneumonia. It was called smooth (S) because colonies of bacteria appear smooth. Another strain did not have a sugar coat and did not cause pneumonia. It was called rough (R) because its colonies have rough edges. When Griffith injected live S strain into a mouse, the mouse died. When Griffith injected live R strain into a mouse, the mouse did not die. Next, Griffith heated and killed the S strain. When injected, the dead S strain no longer killed the mouse. Griffith mixed the heat-killed S strain with the live R strain. When he injected the mixture into a mouse, something unexpected happened — the mouse died. Griffith studied live bacteria from the dead mouse. The smooth trait was visible. He concluded that the live R stain had changed into live S strain. Thus Griffith discovered the process of *transformation* (fig.1).

Avery, MacLeod and McCarty experiment

In 1931, *Oswald Avery*, *Colin MacLeod* and *Maclyn McCarty* identified the molecule that transformed the R strain into S strain. Avery and colleagues investigated the chemical nature of this transforming factor - more specifically, whether the transforming factor was a protein or a nucleic acid. They, like Griffith, attempted to transform the R strain into the S strain by incubating living R and heat-killed S. However, they pretreated the heat-killed S with either a protease (an enzyme that degrades proteins) or with DNAase, and enzyme that degrades DNA. They reasoned that if the transforming factor was a protein, treatment of the heat-killed S with a protease would destroy the protein and inhibit transformation and treatment with DNAase should have no effect on the transformation process. On the other hand, if DNA were the genetic material, the opposite would be true. In their experiments, Avery et al. found that protease did not affect the ability of «dead» S to

transform R but DNAase did, therefore they concluded that the genetic material in transformation is most likely DNA.

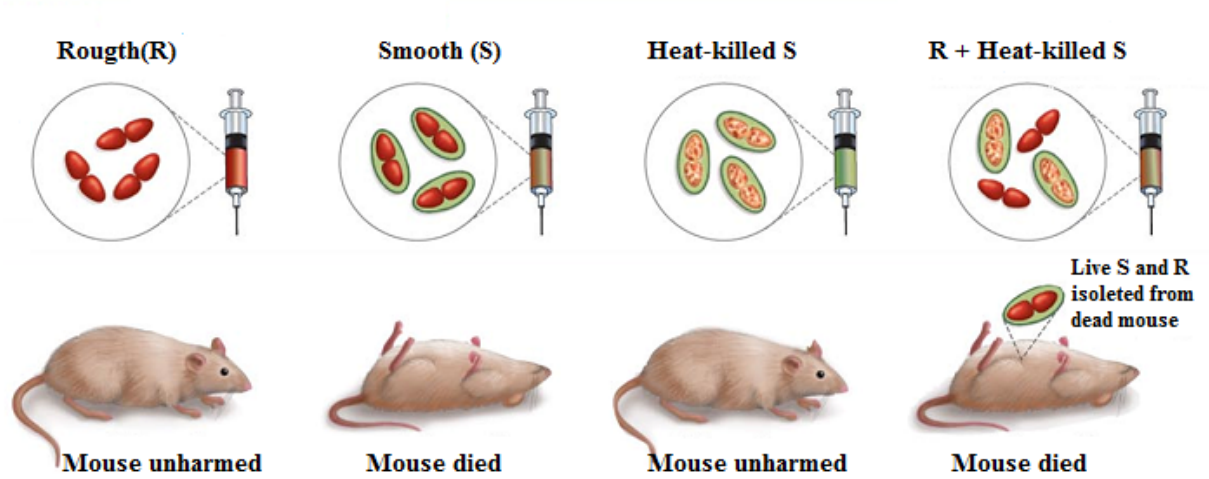

Fig.1. *Griffith's experiments on genetic transformation in pneumococcus*[1]

Hershey-Chase experiment

In 1952, *Alfred Hershey* and *Martha Chase* published results of an experiment proving that DNA was the genetic material. Hershey and Chase did an experiment with bacteriophages, a type of virus that infects bacteria. The bacteriophages were made of DNA and protein. They reproduce by attaching to and injecting their genetic material into a living bacterial cell. Hershey and Chase used radioactive phosphorus (^{32}P) to label the DNA of one set of bacteriophages. They used radioactive sulfur (^{35}S) to label the protein of a second set of bacteriophages. Hershey and Chase mixed bacteria with viruses from the two groups. The viruses injected their genetic material into the bacteria. The viruses were separated from the bacteria. Hershey and Chase found that both sets of viruses had replicated inside the bacterial cells. But only the labeled DNA had entered the bacterial cells. The labeled protein remained outside the bacterial cells. This experiment provided evidence that DNA, not protein, was the genetic material.

After the Hershey-Chase experiment, most scientists thought that DNA was the genetic material. Four scientists helped solve the structure of DNA: British scientists *Rosalind Franklin*, *Maurice Wilkins* and *Francis Crick* and American scientist *James Watson*. Franklin took a picture of DNA using X-ray diffraction, a technique that involved

[1] Copyright © The McCraw-Hill Companies, Inc.

aiming X rays at DNA. Franklin's picture showed that DNA was a double helix, with two strands of nucleotides twisted around each other like a twisted ladder. Watson and Crick saw Franklin's X-ray diffraction picture. They used Franklin's picture and data as well as other mathematical data to determine the specific structure of the DNA double helix. Watson and Crick built a model of DNA.

Structure of DNA

DNA is composed of a series of smaller molecules called ***nucleotides***. In turn, each nucleotide is made up of three primary components:

- ✓ a nitrogen-containing (nitrogenous) base;
- ✓ a carbon-based (pentose) sugar molecule called deoxyribose;
- ✓ a phosphate group attached to the sugar molecule.

There are 4 different bases in a DNA molecule, two *purines*: ***Adenine*** (**A**), ***Guanine*** (**G**), and two *pyrimidines*: ***Thymine*** (**T**), ***Cytosine*** (**C**). Pentose sugar molecules link to phosphate groups to build DNA backbone.

When nucleotides join together in a series, they form a structure known as a ***polynucleotide***. The nucleotide attaches to another one through a connection called a ***phosphodiester bond***. DNA molecule contains two chains of polynucleotides. These two chains are helically wrapped around each other forming DNA ***double helix*** (fig.2).

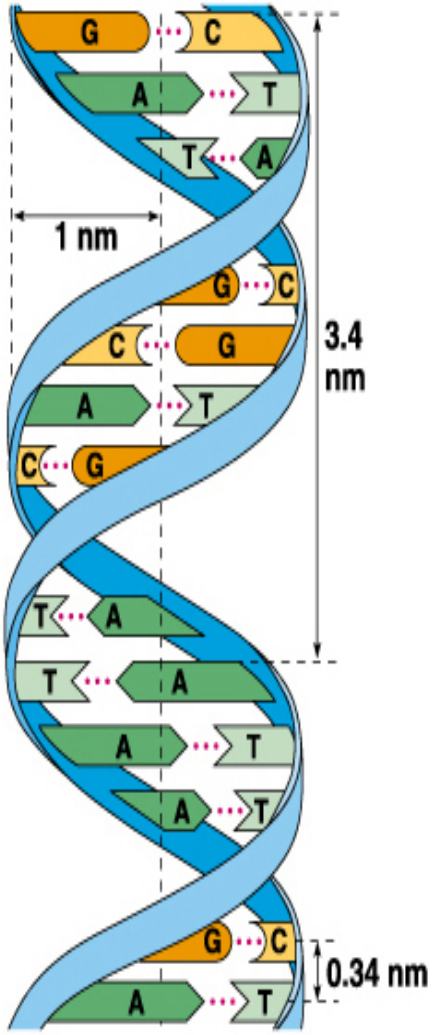

Fig.2. Model of DNA double helix[2]

Sugar-phosphate backbone lies to the outer side of helix. To the inner side, nitrogen bases from opposite stands form pairs. This base-to-base bonding is not random: adenine pairs with thymine and guanine with cytosine. These base pairs are held together by ***hydrogen bonds***. Two hydrogen bonds are formed between adenine and thymine, three bonds are formed between guanine and cytosine. The proportions of the nucleotides in DNA are the number of adenines equals the number of thymines, and the number of guanines equals the number of cytosines. In other words, A = T and G = C. It's ***The Chargaff's rule***. The double-stranded DNA that results from this pattern of bonding looks much like a ladder with sugar-phosphate side supports and base-pair rungs. The two strands run in opposite directions i.e. they are ***antiparallel*** to each other. Also they are ***complementary*** to each other: the sequence of one polynucleotide determines the sequence of the other one. The

[2] Copyright © Pearson Education, Inc., publishing as Benjamin Cummings

distance between base pairs – 0,34nm, the distance per complete turn 3,4nm, the number of base pairs per turn - 10 nucleotides.

DNA packaging[3]

During DNA packaging (in Eukaryots), long pieces of double-stranded DNA are tightly coiled, looped and folded by wrapping their DNA around special proteins called ***histones***. Together, eukaryotic DNA and the histone proteins that hold it together in a coiled form are called ***chromatin***. There are up to six levels of chromatin packaging (fig.3 a – e; fig.4)[4] to fit into the nucleus of one cell.

Nucleosome (***10nm Euchromatic Fiber***)

Nucleosome is the basic structural unit of chromatin. It has a core which is made up of eight histone protein molecules, hence is also called an ***octamer***. The ***core*** has two molecules of each H2A, H2B, H3 and H4. Around this core is wound DNA duplex making two turns which are total about 147 base pairs in length. Histone H1 molecule interacts with the core and the DNA turns (fig.3a).

[3] The haploid human genome contains approximately 3 billion base pairs of DNA packaged into 23 chromosomes. Of course, most cells in the body (except for female ova and male sperm) are diploid, with *23 pairs of chromosomes*. That makes a total of 6 billion base pairs of DNA per cell. Because each base pair is around 0.34 nanometers long (a nanometer is one-billionth of a meter), each diploid cell therefore contains about 2 meters of DNA [$(0.34 \times 10^{-9}) \times (6 \times 10^{9})$]. Moreover, it is estimated that the human body contains about 50 trillion cells—which works out to 100 trillion meters of DNA per human. Now, consider the fact that the Sun is 150 billion meters from Earth. This means that each of us has enough DNA to go from here to the Sun and back more than 300 times, or around Earth's equator 2.5 million times.

[4]

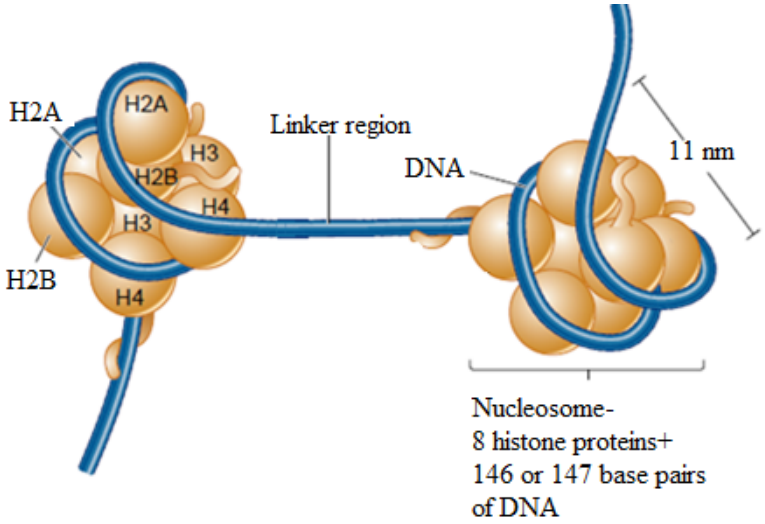

Fig. 3(a). Structure of nucleosome

Adjacent nucleosomes are connected by way of a small stretch of ***Linker DNA***. In this way, the overall structure of connected nucleosomes resembles ***beads on a string*** that shortens the length of the DNA. Whereas DNA duplex is about 2 nm in thickness, nucleosome is about ***11 nm*** in thickness.

Solenoid (***Three-dimensional zigzag structure***)

The 10 nm fibers is further coiled into ***30 nm Euchromatic fibers***, where six nucleosomes make one turn, forming solenoid or three-dimensional zigzag structure (fig.3b).

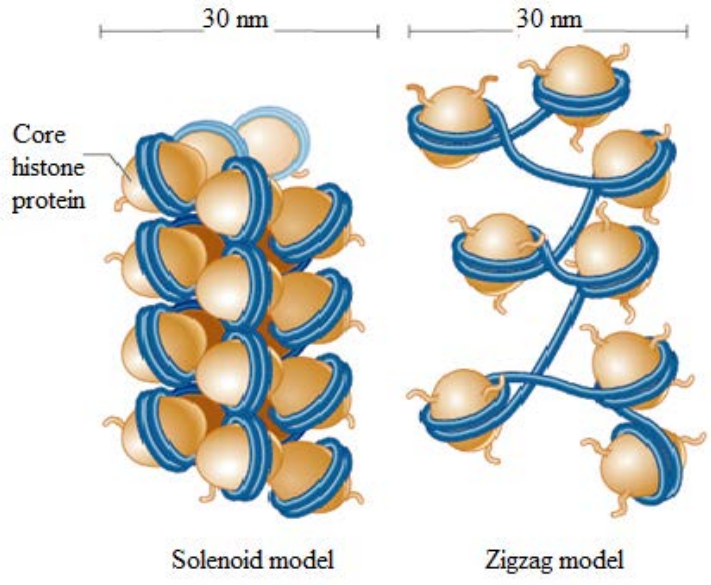

Fig. 3 (b). The 30-nm fiber

In the solenoid model, the nucleosomes are packed in a spiral configuration. In the zigzag model, the linker DNA forms a more irregular structure, and less contact occurs

between adjacent nucleosomes. The zigzag model is consistent with more recent data regarding chromatin conformation.

***Radial Loop Domains* (*300nm chromatin fiber*)**

Next level of chromatin compaction involves interactions between the 30nm fiber and the filamentous network of proteins in the nucleus –these proteins are involved in compacting the DNA into Radial Loop Domains. The distance that these loops radiate from the protein scaffold to which they connect is 300nm. Chromatin that is organized into radial loop domains is regarded as Euchromatic, for it is "loosely" organized and is therefore transcriptionally active (fig.3c).

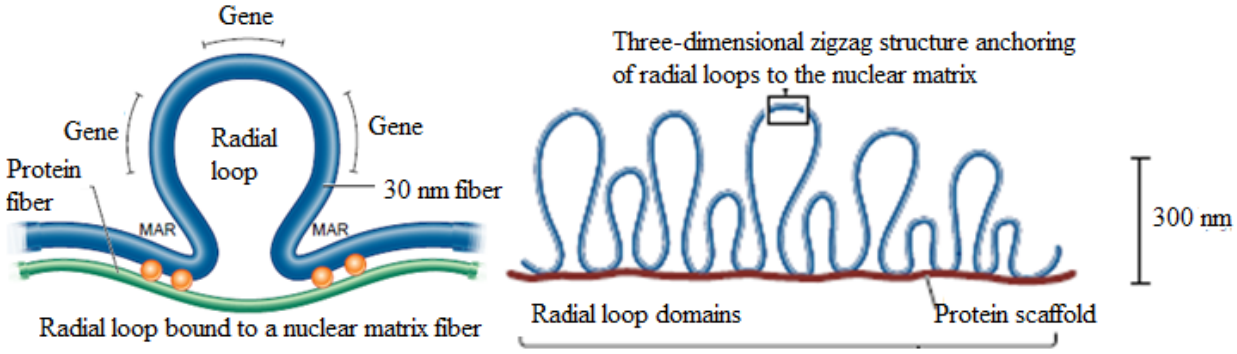

Fig. 3(c). Radial Loop Domains

***Interphase Chromosome* (*700nm*)**

In some regions, however, the chromatin is further compacted via the folding of the protein scaffold itself to form a fiber approximately 700nm in diameter. These Heterochromatic regions are composed of repetitive DNA are too "tightly" packed to be active (fig.3d).

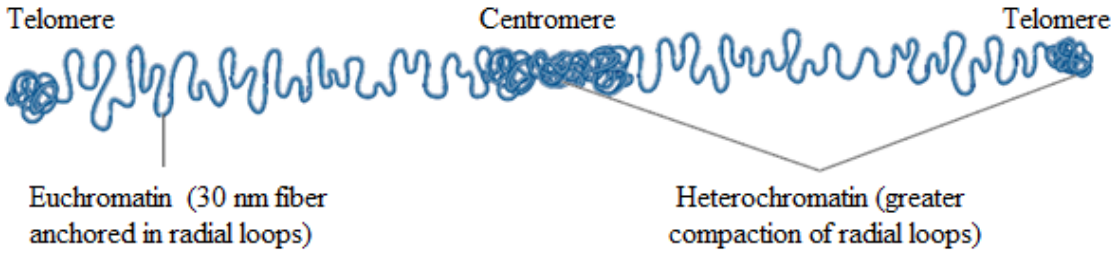

Fig. 3(d). Chromatin structure during interphase

***Metaphase Chromosome* (*1400nm*)**

Condensation process goes on to ultimately condense the entire chromatin into a highly condensed structure called Metaphase chromosome. This condensation into

chromosome takes place at the time of cell division and the maximum degree of condensation of chromosome is reached during metaphase of cell division, in which chromosome is shortest in length and becomes visible in light microscope. All the above levels of chromatin packaging are shown in picture below (fig.3e).

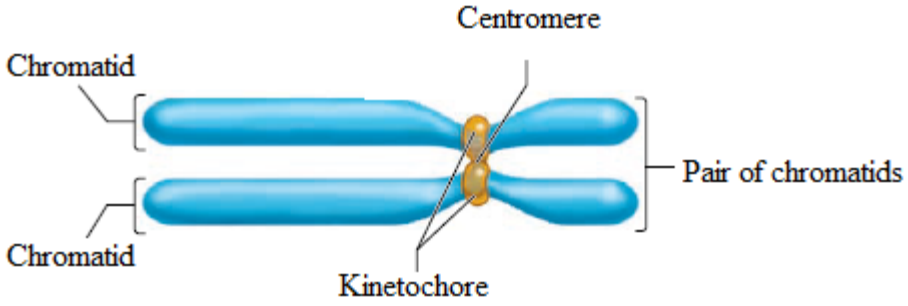

Fig. 3(e). Metaphase Chromosome

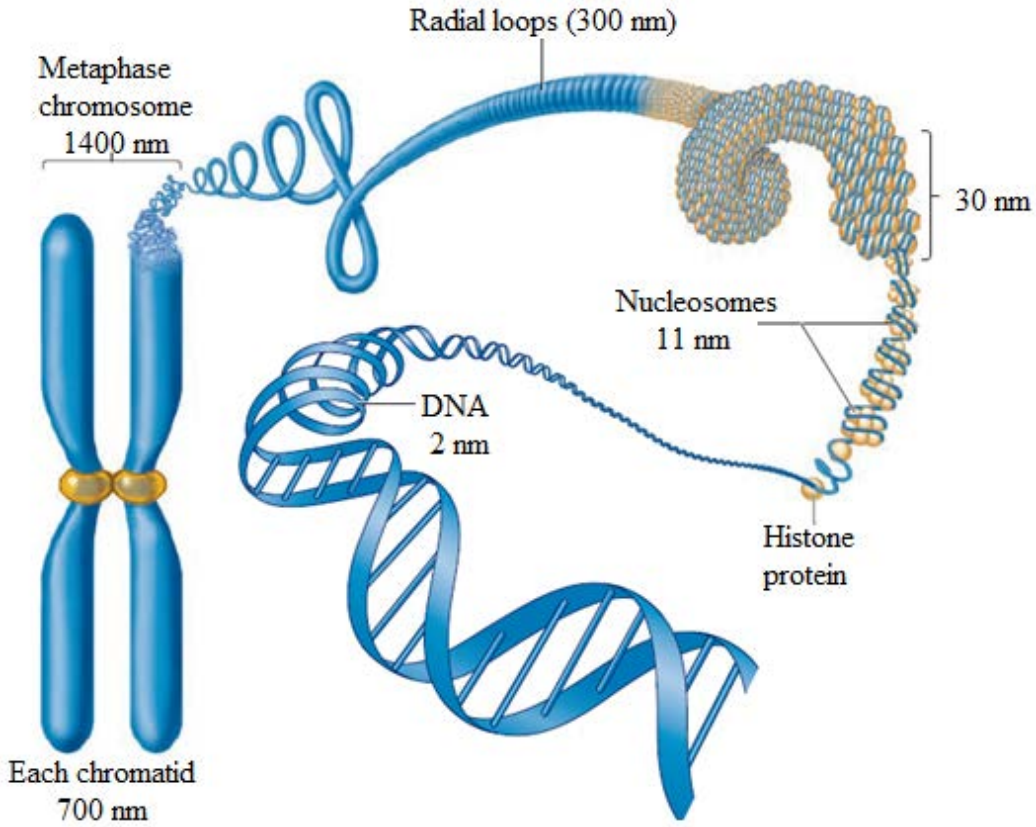

Fig.4. Levels of DNA packaging

Euchromatin and Heterochromatin

There are two varieties of chromatin. They are euchromatin and heterochromatin. These two forms are distinguished in a cytological manner dealing with how intensely each form is stained.

Euchromatin is the lightly packed material. It is less intense stained than heterochromatin.

Heterochromatin is the tightly compacted regions of chromosomes. In general, these regions of the chromosome are transcriptionally inactive.

It consists of two types:

- ✓ *Constitutive*: refers to chromosomal regions that are always heterochromatic and permanently inactive with regard to transcription. Constitutive heterochromatin usually contains highly repetitive DNA sequences, such as tandem repeats, rather than gene sequences (centromeric DNA);
- ✓ *Facultative*: refers to chromatin that can occasionally interconvert between heterochromatin and euchromatin. An example of facultative heterochromatin occurs in female mammals when one of the two X chromosomes is converted to a heterochromatic Barr body.

Chromosome Morphology

Chromosomes are complex structures located in the cell nucleus; they are composed of DNA (40%), histone (60-80%) and nonhistone proteins (15%), RNA, and polysaccharides. During cell division (at metaphase) chromosome has X-shaped structure called a *metaphase chromosome*. It consists of two sister chromatides. Centromere divides a chromosome in two arms (short arm – *p*, long arm – *q*).

Chromosome structure could be explained with the help of some terminologies as given below (fig.5):

- ✓ Chromatide
- ✓ Chromonema or Chromonemata
- ✓ Chromomeres
- ✓ Primary constriction or Centromere and Kinetochore
- ✓ Telomers
- ✓ Secondary Constriction
- ✓ Nucleolar Organizer
- ✓ Satellite

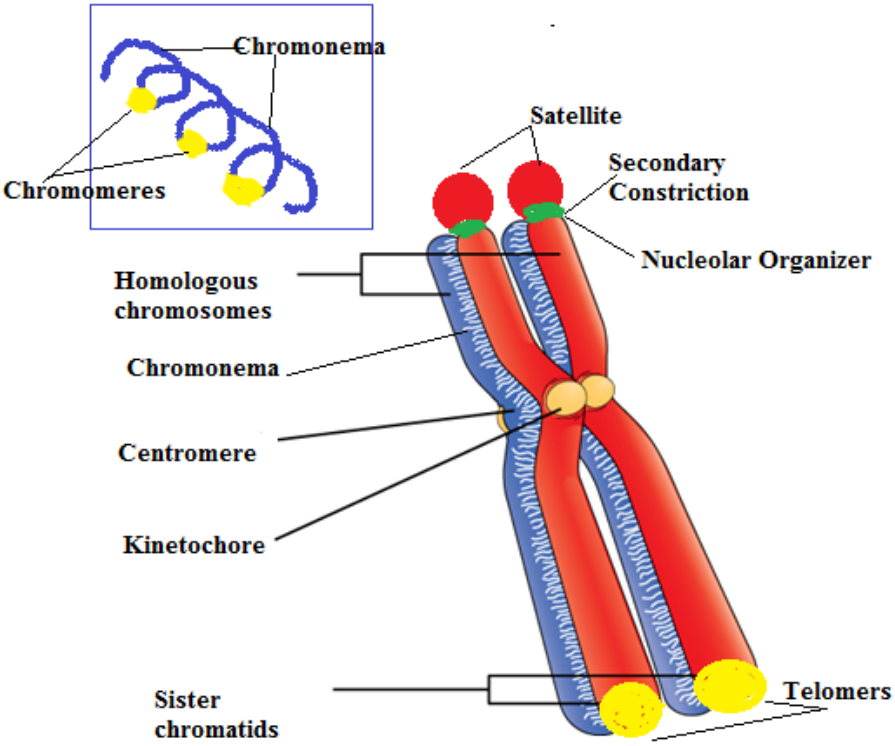

Fig. 5. Morphology of chromosome

Chromatid

Chromosome has two symmetrical structures at mitotic metaphase, these are called as chromatids**.**

Chromonema or ***Chromonemata***

Chromosomes appear as very thin filaments during mitotic prophase, these are called as chromonema or chromonemata, these are chromatids in the early stage of condensation. These are present in coiled and twisted form in a chromatid during all stages of *mitosis*. Genes are located on this structure. Chromatid may contain one or more chromonema.

Chromomeres

These are regions of tightly folded *DNA*, like small bead visible when chromosomes are relatively uncoiled.

Primary constriction or ***Centromere*** and ***Kinetochore***

Centromere is a region of DNA / chromosome (contains highly repetitive DNA) appear mostly near the middle of a chromosome. It lies within primary constriction. Fibers of mitotic spindle attaches to this during mitosis. Kinetochore is plate or cup like structure (0.20 to 0.25nm) situated upon centromere. It divides the chromosome in to two arms (*p* and *q*).

Telomers

Telomers (non-coding DNA) are present at the ends of chromosome. They have polarity, because of this chromosome segments do not fuse with others. If chromosome breaks, the broken ends do not contain telomers, so they can stick with each other. Telomers also assist in the pairing of homologous chromosomes and crossing over.

Secondary Constriction

Secondary constrictions are present at any point on chromosome. They may arise because the rRNA genes are transcribed very actively and thus interfering with chromosomal condensation. Secondary constrictions are constant in their position hence prove useful to identify a particular chromosome in a set.

Nucleolar Organizer

These are some parts of secondary constrictions which contain the *genes* coding for rRNA and that induce the formation of nucleolus.

Satellite

These are appendages of chromosome either round, elongated or knob like. Satellite is produced if secondary constriction is present in distal region of chromosome arm. It keeps connection with rest of the chromosome by a thin chromatin filament. Shape and size of satellite remains constant.

Types of chromosomes:

Chromosomes are classified according to their shape, which is determined by the position of the centromere (fig.6):

Metacentric chromosomes have short and long arms of roughly equal length with the centromere in the middle.

Submetacentric chromosomes have short and long arms of unequal length with the centromere more towards one end.

Acrocentric chromosomes have a centromere very near to one end and have very small short arms. They frequently have secondary constrictions on the short arms that connect very small pieces of DNA, called stalks and satellites, to the centromere. The stalks contain genes which code for ribosomal RNA.

Telocentric chromosomes have centromere located at the terminal end. Humans do not possess telocentric chromosomes.

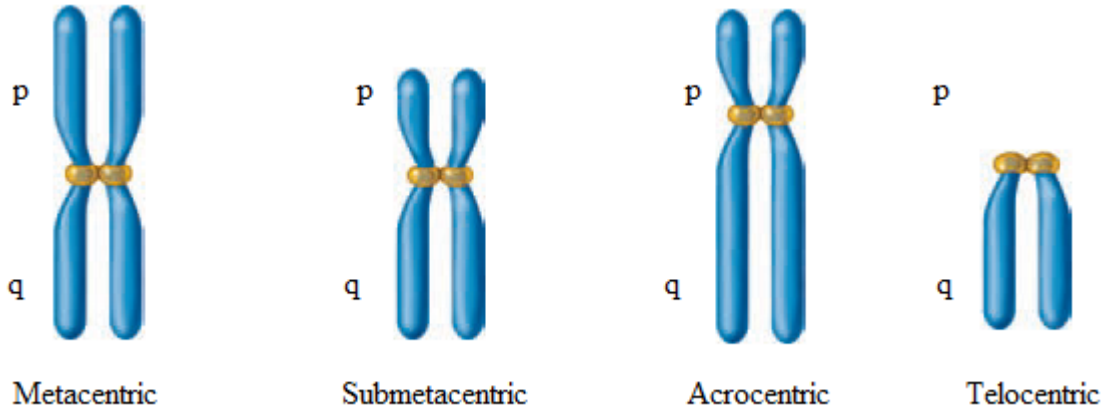

Fig.6 Morphology of chromosome

DNA Replication

Genetic material is transmitted from parent to offspring and from cell to cell. For transmission to occur, the genetic material must be copied. During process, known as ***DNA replication,*** the original DNA strands are used as templates for the synthesis of new DNA strands.

Models of DNA replication

There are three possible models of DNA replication (fig.7):

- ✓ Semi conservative
- ✓ Conservative
- ✓ Dispersive

Conservative model

According to this hypothesis, both strands of parental DNA remain together following DNA replication. In this model, the original arrangement of parental strands is completely conserved, while the two newly made daughter strands also remain together following replication.

Semiconservative model

In this mechanism, the double stranded DNA is half conserved following the replication process. In other words, the newly made double-stranded DNA contains one parental strand and one daughter strand.

Dispersive model

In this model proposes that segments of parental DNA and newly made DNA are interspersed in both strands following the replication process.

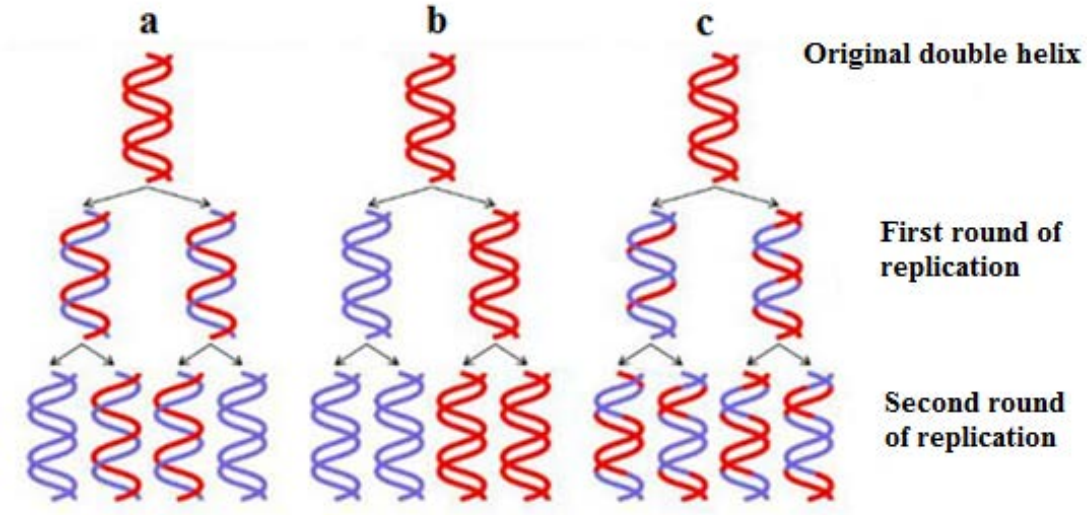

Fig.7. Models of DNA replication

a- Semiconservative, b – Conservative, c - Dispersive

DNA replicates by the ***semiconservative model of replication***.

The mechanism of DNA replication as originally proposed by Watson and Crick. DNA replication relies on the complementarity of DNA strands according to the AT/GC rule. During the replication process, the two complementary strands of DNA come apart and serve as ***template strands*** for the synthesis of two new strands of DNA.

Enzymes of DNA Replication

- ✓ *DNA helicase* breaks the hydrogen bonds between the DNA strands.
- ✓ *Topoisomerase* alleviates positive supercoiling.
- ✓ *Single-strand binding proteins* keep the parental strands apart.
- ✓ *Primase* synthesizes an RNA primer.
- ✓ *DNA polymerase III* (***DNA polymerase delta*** In Eukaryotes) synthesizes adaughter strand of DNA.

- ✓ *DNA polymerase* I (*Flap* e***xonuclease*** In Eukaryotes) excises the RNA primers and fills in with DNA
- ✓ *DNA ligase* covalently links the Okazaki fragments together.

DNA replication proceeds in three steps: initiation, elongation and termination.

Initiation

Replication starts at a specific site on the DNA known as the ***origin of replication*** (prokaryotes have only a single origin on their chromosome while eukaryotes have several on each of their chromosomes). Specific protein binds to the strand and recruits ***DNA helicase***. The function of DNA helicase is to break the hydrogen bonds between base pairs and thereby unwind the strands. This action generates positive supercoiling ahead of each replication fork. An enzyme known as a ***topoisomerase*** (***DNA gyrase*)**, travels in front of DNA helicase and alleviates positive supercoiling. In addition, a set of ***single-stranded binding proteins (SSB proteins)*** binds to the strands of parental DNA and prevent them from re-forming a double helix.

The synthesis of new daughter strands is initiated within the origin and it proceeds in bothdirections, or *bidirectionally* (fig.8). This means that two ***replication forks*** move in opposite directions outward from the origin. A replication fork is the site where the parental DNA strands have separated and new daughter strands are being made. Eventually, these replication forks meet each other on the opposite side of the chromosome to complete the replication process.

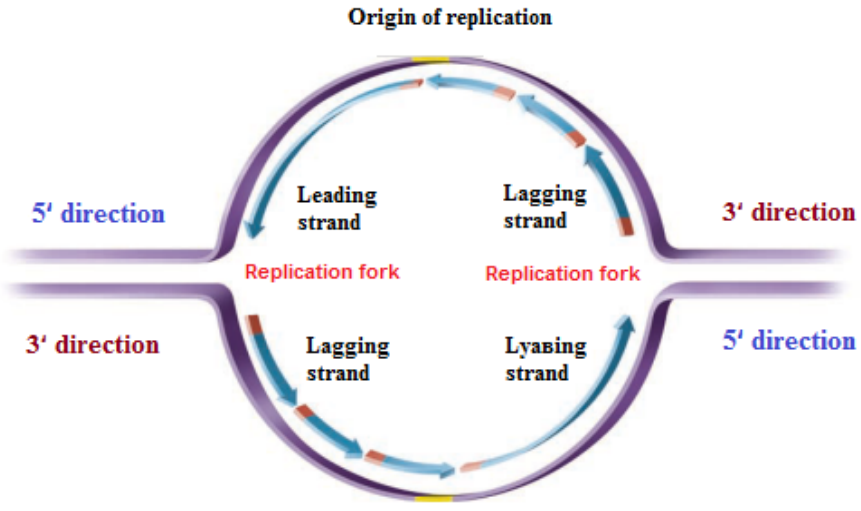

Fig.8. Bidirectional replication[5]

DNA polymerase builds new strand in 5'-3' direction, whereas the DNA template is read in 3' to 5' direction. Thus, DNA polymerase requires a primer with a free 3' OH group to operate. So, the next event in DNA replication involves the synthesis of short strands of RNA (rather than DNA) called ***RNA primers*** with 3' OH group. These strands of RNA are synthesized by enzyme known as ***primase.*** Primers help to start replication (fig.9).

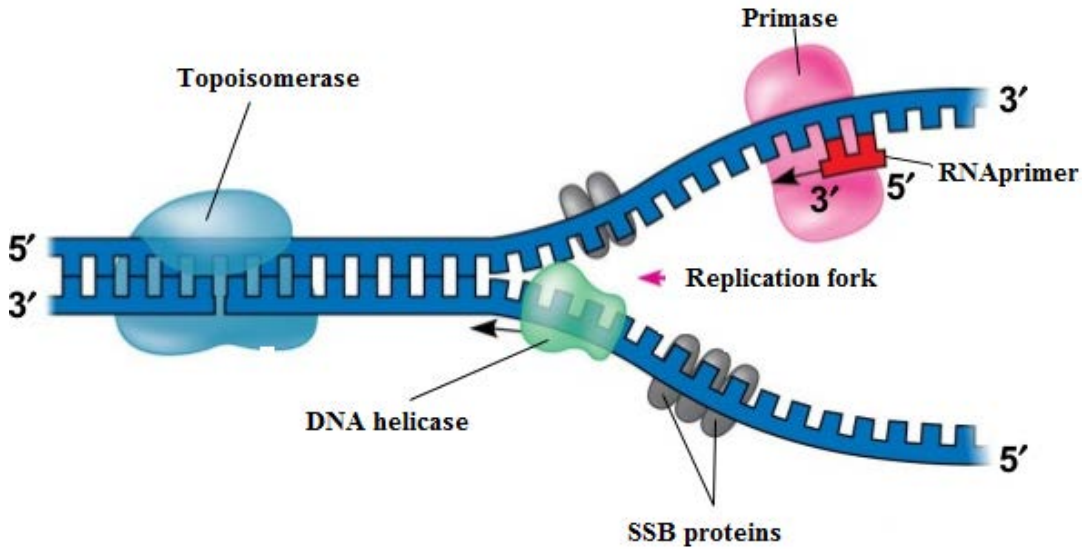

Fig. 9. Initiation[6]

Elongation

Synthesis of two strands of DNA occurs, one from them is ***the leading strand*** and another is ***the lagging strand***. The leading strand is formed in the form of very long continuous fragment. A lagging strand is formed in a series of short (approximately 1500 nucleotides) ***Okazaki fragments***. The direction of a lagging strand is opposite to the direction of movement corresponding to the replication fork.

A type of enzyme known as ***DNA polymerase*** is responsible for elongation. DNA polymerase slides along the template strand as it synthesizes a new strand by connecting

[5] Copyright © The McCraw-Hill Companies, Inc.2012.

[6] 2014 Pearson Education. Inc.

adjacent nucleotides in a 5′ to 3′ direction. Both prokaryotes and eukaryotes have many types of DNA polymerases. In Prokaryotes there are five proteins function as DNA polymerases: polymerase I, II, III, IV, and V. DNA polymerases I and III are involved in normal DNA replication, whereas DNA polymerases II, IV, and V play a role in DNA repair and the replication of damaged DNA. DNA polymerase III is responsible for most of the DNA replication.

To complete the synthesis of Okazaki fragments within the lagging strand, three additional events must occur: removal of the RNA primers, synthesis of DNA in the area where the primers have been removed, and the covalent attachment of adjacent fragments. Bacterial RNA primers are removed by ***DNA polymerase I***. A DNA polymerase enzyme does not play this role In Eukaryotes. Instead, an enzyme called ***flap endonuclease*** is primarily responsible for RNA primer removal. DNA polymerase I removes the RNA primer from the first Okazaki fragment and then synthesize DNA in the vacant region by attaching nucleotides to the 3′ end of the second Okazaki fragment. After the gap has been completely filled in, a covalent bond is still missing between the last nucleotide added by DNA polymerase I and the adjacent DNA strand that had been previously made by DNA polymerase III. An enzyme known as ***DNA ligase*** catalyzes a covalent bond between adjacent fragments to complete the replication process in the lagging strand.

Termination

Replication is terminated when the replication forks meet at the termination sequences - ***ter sequences*** In Prokaryotes. In Eukaryotes DNA replication ends at the *telomere* regions which are at both ends of linear eukaryotic chromosomes. The *telomere* regions are a 3′ moderately repetitive sequences of nucleotides. DNA polymerase synthesizes DNA only in a 5′ to 3′ direction and it is unable to replicate the 3′ ends of DNA strands. Therefore, if this problem were not solved, the chromosome would become progressively shorter with each round of DNA replication. Thus, the ends of eukaryotic chromosomes are replicated by ***telomerase***. It binds to the 3′ region of the telomere and synthesizes additional repeats of telomeric sequences to prevent the loss of genetic information.

All the above steps of DNA replication are shown in picture below (fig.10).

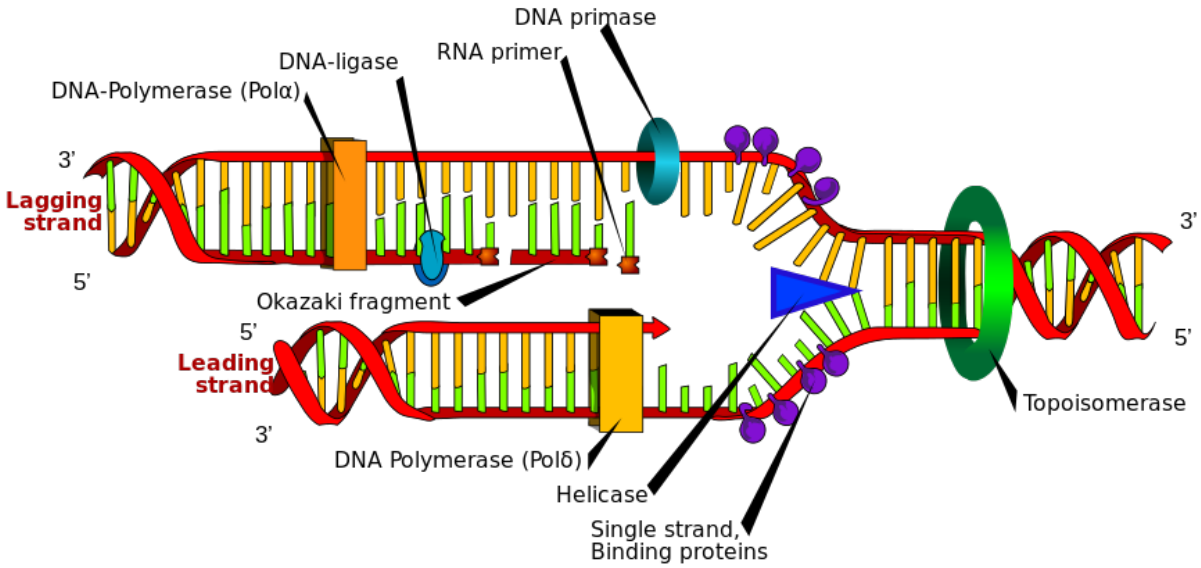

Fig.10. Scheme of DNA replication[7]

DNA Repair

All species have a variety of DNA repair systems to avoid the harmful effects of mutations. Common Types of DNA Repair Systems are shown in table below (tabl.1).

Table 1

Common Types of DNA Repair Systems

Repair System	**Description**
Direct repair	An enzyme recognizes an incorrect alteration in DNA structure and directly converts it back to a correct structure.
Base excision repair (nucleotide excision) repair	An abnormal base or nucleotide is first recognized and removed from the DNA. A segment of DNA is excised, and then the complementary DNA strand is used as a template to synthesize normal DNA.
Mismatch repair	Similar to excision repair except the DNA defect is a base pair mismatch, not an abnormal nucleotide. The mismatch is recognized, and a segment of DNA is removed. The parental strand is used to synthesize a normal daughter strand of DNA.
Homologous recombination repair	Occurs at double-strand breaks or when DNA damage causes a gap in synthesis during DNA replication. he strands of a normal chromatid are used to repair a damaged chromatid.

[7]This work has been released into the *public domain* by its author, *LadyofHats*. This applies worldwide.

Nonhomologous end joining repair	Occurs at double-strand breaks. The broken ends are recognized by proteins that keep the ends together; the broken ends are eventually.

Base excision repair

This repair system recovers most of the DNA damage in the cell. For excision repair needed a second (complementary) DNA strand. The excision repair is usually attended by several enzymes, and the process involves not only damaged site, but also its neighboring nucleotides.

The main steps of excision repair (fig.11):

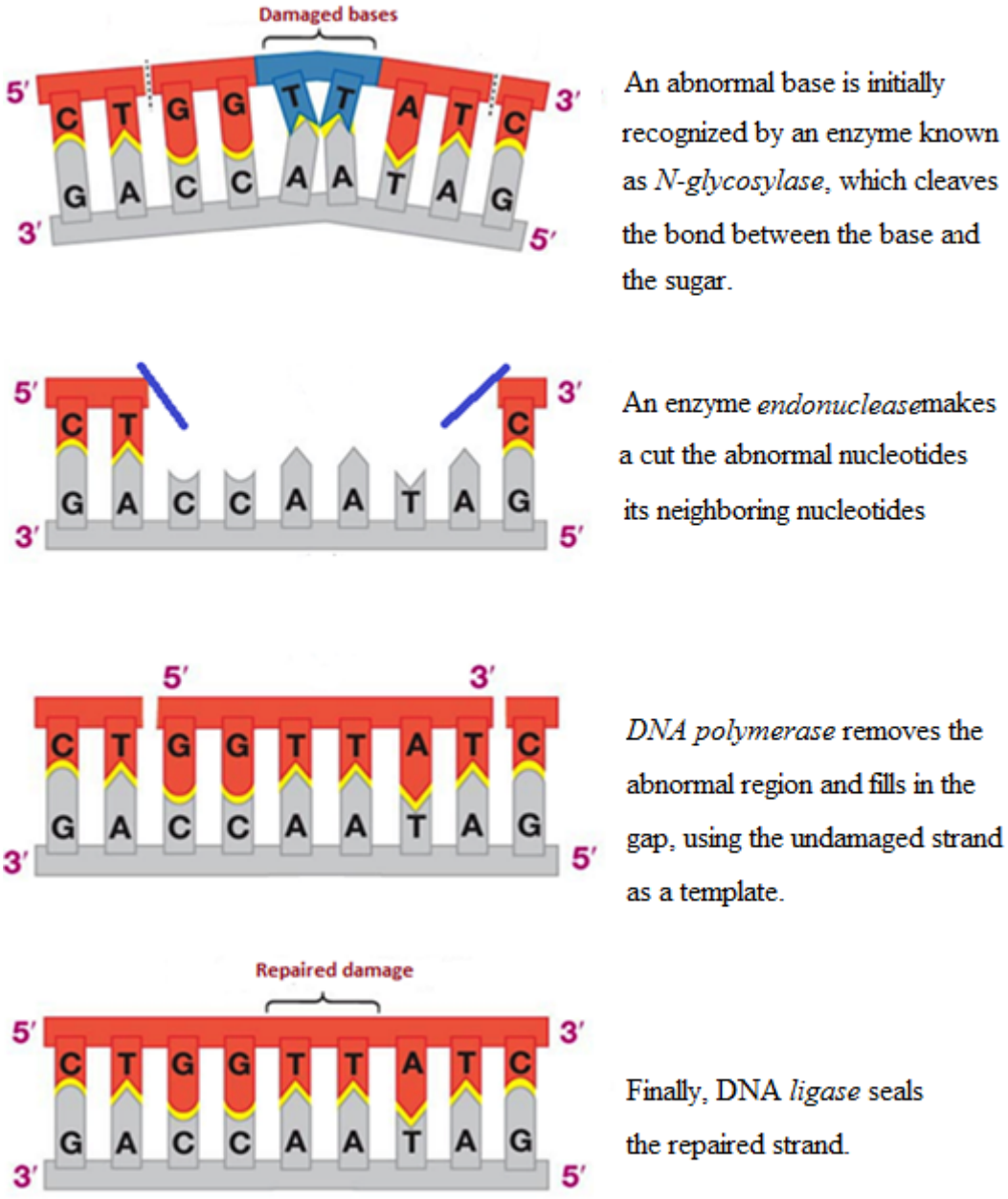

Fig. 11. Steps of excision repair