

DNA – Deoxyribonucleic Acid:

DNA as an acidic substance present in nucleus was first identified by Friedrich Meischer in 1869. He named it as 'Nuclein'. However, due to technical limitation in isolating such a long polymer intact, the elucidation of structure of DNA remained elusive for a very long period of time.

It was only in 1953 that **James Watson** and **Francis Crick**, based on the X-ray diffraction data produced by **Maurice Wilkins** and **Rosalind Franklin**, proposed a very simple but famous **Double Helix** model for the structure of DNA. One of the hallmarks of their proposition was base pairing between the two strands of polynucleotide chains.

➤ Wilkins, Watson and Crick were awarded Nobel Prize for the same in 1962.

Chargaff's Rules: Chargaff (1950) made observations on the base and other contents of DNA. These observations are called Chargaff's rule.

i) Purine and pyrimidine base pairs are equal in amount, that is, adenine + guanine = thymine + cytosine. $[A + G] = [T + C]$

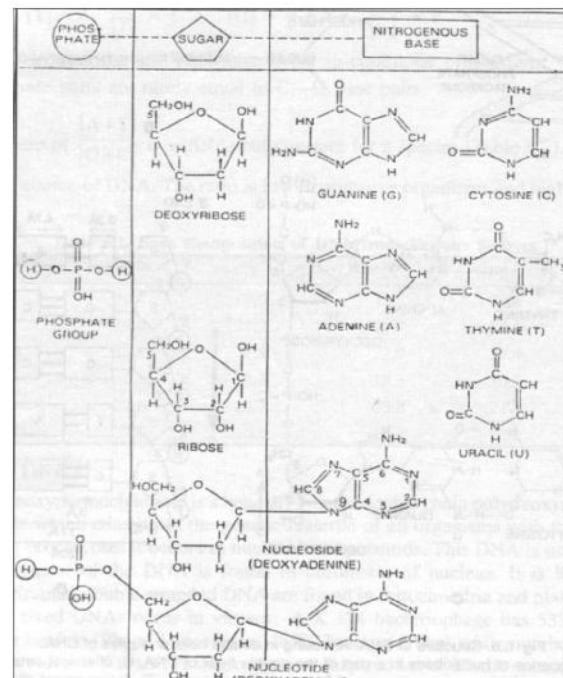
$$\text{i.e. } \frac{[A+G]}{[T+C]} = 1$$

ii) Molar amount of adenine is always equal to the molar amount of thymine. Similarly, molar concentration of guanine is equalled by molar concentration of cytosine.

$$[A] = [T], \text{i.e. } \frac{[A]}{[T]} = 1; [G] = [C], \text{i.e. } \frac{[G]}{[C]} = 1$$

iii) Sugar deoxyribose and phosphate occur in equimolar proportions.

iv) The ratio of $\frac{[A+T]}{[G+C]}$ is variable but constant for a species. It can be used to identify the source of DNA. The ratio is low in primitive organisms and higher in advanced ones.



Species	A	G	C	T	A+T/ C+G
1. Man	30.4	19.0	19.9	30.1	1.55
2. Calf	29.0	21.2	21.2	28.5	1.35
3. Wheat germ	28.1	21.8	22.7	27.4	1.25
4. Euglena	22.6	27.7	25.8	24.4	0.88
5. E. coli	24.7	26.0	25.7	23.6	0.93

Purine	Pyrimidine
1. Purines are larger-sized nitrogen containing biomolecules.	1. Pyrimidines are smaller-sized nitrogen containing biomolecules.
2. A purine is nine-membered.	2. A pyrimidine is 6-membered.
3. It is double ring.	3. It is a single ring.
4. A purine contains four nitrogen atoms at 1,3,7 and 9 positions.	4. A pyrimidine has nitrogen atoms at two places, 1 and 3-positions.
5. Purine bases are of two types – Adenine (A) and Guanine (G).	5. Pyrimidine bases are of three types- cytosine (C), Thymine(T) and Uracil (U)

Q. If the double stranded DNA has 20% of cytosine, calculate the percent of adenine in DNA.

3. If the amount of Adenine in DNA of a plant is 20% and that of guanine 30%, then amount of cytosine in the DNA of the same plant will be – (CoHSEM-2002) 1

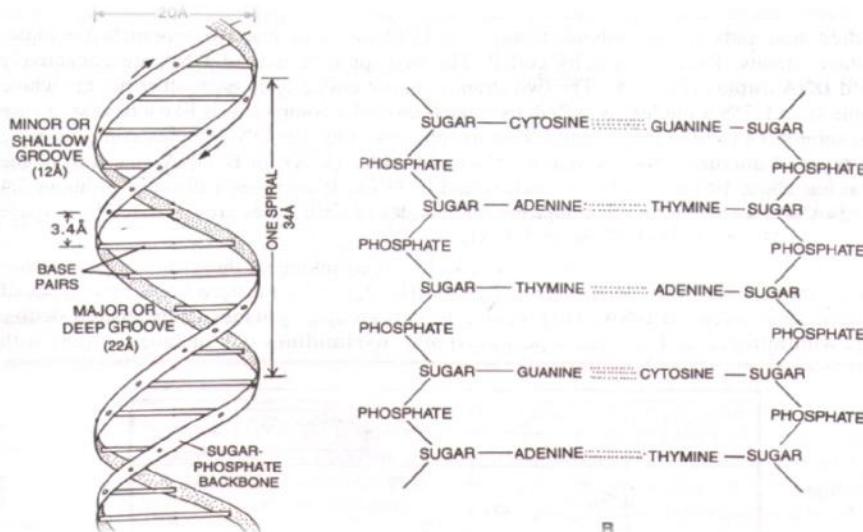
- a) 20% b) 30% c) 40% d) 60%

Structure of DNA: DNA is a helically twisted double chain polydeoxyribonucleotide macromolecule which constitutes the genetic material of all organisms with the exception of riboviruses. DNA is the largest macromolecules with a diameter of 2 nm (20 \AA^0 or $2 \times 10^{-9} \text{ m}$) and often having a length in millimeters.

Characters	A-DNA	B-DNA	C-DNA	D-DNA	Z-DNA
1. Handedness of helix	Right handed	Right handed	Right handed	Right handed	Left handed
2. Pitch of helix per turn	25 \AA^0	34 \AA^0	30 \AA^0	24 \AA^0	46 \AA^0
3. Diameter of helix	23 \AA^0 (Widest)	20 \AA^0 (Widest)	19 \AA^0 (Widest)	-----	18 \AA^0 (Widest)
4. Base pairs per turn of helix	11	10	9.33	8	12 (6 dimers)
5. Distance (vertical rise per pair) between 2 base pairs.	2.5 \AA^0	3.4 \AA^0	3.3 \AA^0	3.03 \AA^0	3.6 \AA^0

The salient features of the Double-helix structure of DNA are as follows:

- It is made of two polynucleotide chains, where the backbone is constituted by sugar-phosphate, and the bases project inside.
- The two chains have anti-parallel polarity. It means, if one chain has the polarity $5' \rightarrow 3'$, the other has $3' \rightarrow 5'$.
- The bases in two strands are paired through hydrogen bonds (H-bonds) forming base pairs (bp). Adenine forms two hydrogen bonds with Thymine from opposite strand and vice-versa. Similarly, Guanine is bonded with Cytosine with three H-bonds. As a result, always a purine comes opposite to a pyrimidine. This generates approximately uniform distance between the two strands of the helix.
- The two chains are coiled in a right-handed fashion. The pitch of the helix is 3.4 nm (a nanometre is one billionth of a metre, that is 10^{-9} m) and there are roughly 10 base pairs in each turn. Consequently, the distance between a bp in a helix is approximately equal to 0.34 nm .
- The plane of one base pair stacks over the other in double helix. This, in addition to H-bonds, confers stability of the helical structure.



27. Draw and label Watson and Crick's double helical model of DNA molecule.

3 (CoHSEM-2001)

PACKING OF DNA HELIX:

DNA Packing in Prokaryotes: In prokaryotes, such as *E. coli*, though they do not have a defined nucleus, the DNA is not scattered throughout the cell. DNA (being negatively charged) is held with some proteins (that have positive charges) in a region termed as 'nucleoid' or prochromosome. The DNA in nucleoid is organised in large loops held by proteins (non-histone proteins).

DNA Packing in Eukaryotes: In eukaryotes, this organisation is much more complex. There is a set of positively charged, basic proteins called **histones**.

Histones are rich in the basic amino acid residues lysines and arginines. Both the amino acid residues carry positive charges in their side chains. They attract negatively charged strands of DNA. Histones are organised to form a unit of eight molecules called as histone octamer. The negatively charged DNA is wrapped around the positively charged histone octamer to form a structure called nucleosome.

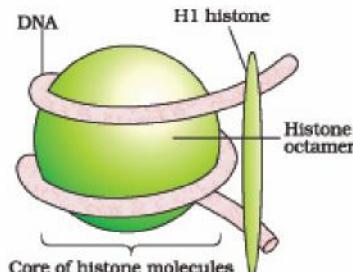


Figure 6.4a Nucleosome

NUCLEOSOME: Nucleosome is submicroscopic sub-unit of chromatin formed by wrapping of DNA over a core of histone proteins. There are five types of histone proteins – H1, H2A, H2B, H3 and H4. Four of them (H2A, H2B, H3 and H4) occur in pairs to produce histone octamer, called *nu body* or core of nucleosome. DNA connecting two adjacent nucleosomes is called linker DNA or interbread. It bears H1 histone protein molecule.

The beads-on-string structure in chromatin is packaged to form chromatin fibers that are further coiled and condensed at metaphase stage of cell division to form chromosomes. The packaging of chromatin at higher level requires additional set of proteins that collectively are referred to as Non-histone Chromosomal (NHC) proteins.

In a typical nucleus, some regions of chromatin are loosely packed and stains light and are referred to as euchromatin. The chromatin that is more densely packed and stains dark are called as Heterochromatin. Euchromatin is said to be transcriptionally active chromatin, whereas heterochromatin is inactive.

THE SEARCH FOR GENETIC MATERIAL

a) Transforming Principle (Griffith's Experiment):

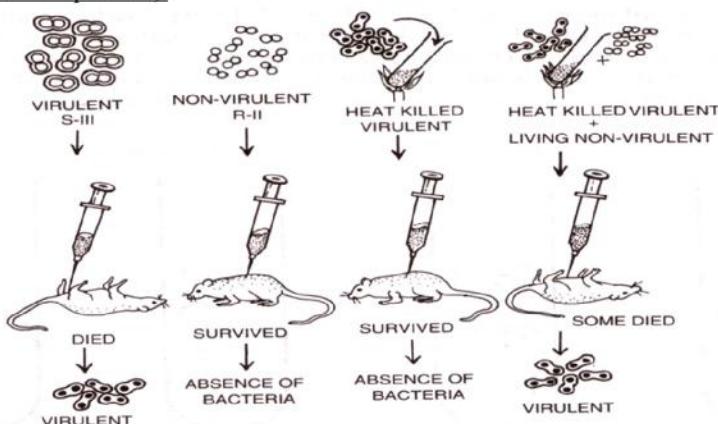


Fig: Summary of Griffith's experiments on transformation in *Diplococcus* or *Streptococcus pneumoniae*

Transformation is the change in genetic constitution/ material of an organism by picking up genes present in the remains of its dead relatives. Transformation experiment was performed by a British doctor, *Frederick Griffith* in 1928 on bacterial strains *Streptococcus pneumoniae* or *Diplococcus* or *Pneumococcus pneumoniae*. The bacterium has two strains –

- Virulent or S – strain which causes disease pneumonia.
- Non-virulent or R – strain which is unable to cause pneumonia.

Virulent bacteria are known as S – type because when grown on suitable medium they form smooth colonies. These diplococci are covered by sheath of mucilage (polysaccharide) around them. The sheath is not only the cause of toxigenicity but also protects the bacteria from phagocytes of host. The non-virulent types of bacteria do not produce the disease. They form irregular or rough colonies.

Griffith performed his experiment by injecting the above bacteria into mice and found the following results:

- a) S-strain (Virulent) bacteria were injected into mice, the mice developed pneumonia and finally died.
 b) R- strain (non-virulent) bacteria were injected into mice, the mice suffered no illness because R – strain was non- pathogenic.
 c) Heat killed S-strain bacteria were injected into mice, they did not suffer from pneumonia and thus survived.
 d) A mixture of R-strain (non-virulent) and heat killed S-bacteria were injected into mice, the mice developed pneumonia and died.

Thus some genetic factor from dead S-strain bacteria converted the live R-strain bacteria into live S – strain and later produced the disease. In short, living R-strain bacteria were somehow transformed. Griffith's effect gradually became known as transformation.

Biochemical Characterization of Transforming Principle: In 1944, Oswald Avery, Colin MacLeod and Maclyn McCarty, purified biochemicals from the heat killed S- type bacteria into 3 components – DNA, carbohydrates and protein. DNA fraction was further divided into 2 parts: One with DNase (Deoxyribonuclease) and other without it. The four components were then added to the separate culture tubes containing R – type bacteria.

They were then analyzed for bacterial population.

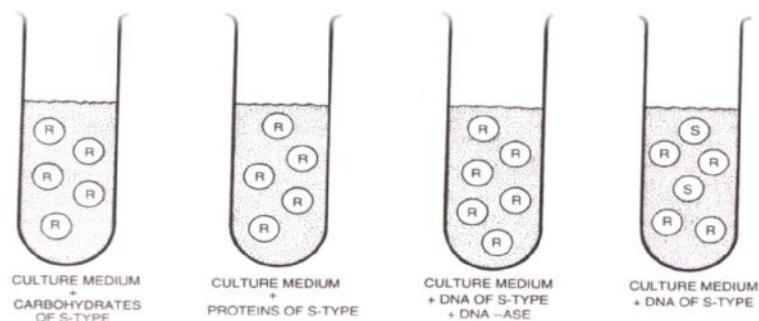
R – type + Protein S- type → R – type.

R – type + Carbohydrate S- type → R – type.

R – type + DNA S- type + DNase → R – type.

R – type + DNA S- type → S – type.

Only DNA of S – type can change R – type bacteria into S – type. Therefore, the character or gene of virulence is located in DNA. Thus they proved that the chemical to be inherited is DNA.



b) Bacteriophage Multiplication (Transduction)/ Hershey and Chase Experiment:

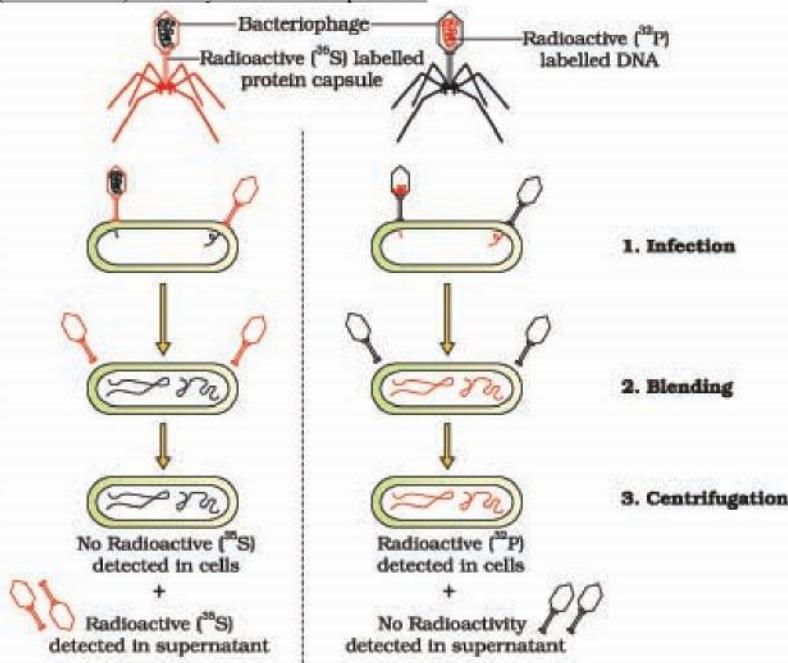


Figure: Hershey - Chase experiment

Bacteriophages are bacterial viruses. T₂ is a bacteriophage which infects *Escherichia coli*, the bacterium present as human commensal as human intestine. A.D. Hershey and Martha Chase (1952) raised T₂ bacteriophage over two different colonies, one having radioactive phosphorus ^{32}P and the other having radioactive sulphur ^{35}S . Radioactive sulphur got incorporated in the capsid proteins of bacteriophage. Radioactive phosphorus ^{32}P became component of DNA of bacteriophage. The two types of bacteriophages were taken out and allowed to infect bacteria, *E. coli* separately. Soon after infection, the cultures were gently agitated in a blender to separate the adhering protein coats of the virus from the bacterial cells. The empty phage capsids or ghosts got separated from the bacterial cells. In the experiment using ^{35}S , radioactivity remained limited to supernatant having phage ghosts. In the experiment using ^{32}P , radioactivity was absent in the supernatant having phage ghosts. The ^{32}P was present in bacteria and the phages which multiplied in them. Therefore, DNA is the genetic material which could enter the bacteria and helped in phage multiplication.

Properties of Genetic Material (DNA versus RNA)

A molecule that can act as a genetic material must fulfill the following criteria:

- (i) It should be able to generate its replica (Replication).
- (ii) It should chemically and structurally be stable.
- (iii) It should provide the scope for slow changes (mutation) that are required for evolution.
- (iv) It should be able to express itself in the form of 'Mendelian Characters'.

DNA Vs RNA as Genetic Material:

- Heat which killed bacteria in Griffith's experiment did not destroy their DNAs but RNA is labile and easily degradable.
- RNA has 2' - OH group in every nucleotide which is quite reactive. This 2' - OH group is absent in DNA.
- Uracil present in RNA is less stable as compared to thymine of DNA.
- The rate of mutation is slow in case of DNA. Being unstable, RNA mutates at a much faster rate, that is why RNA viruses have shorter life span and mutate and evolve very fast.
- The nucleotides of DNA are not exposed all the times due to presence of complementary strands. It is not so in case of RNA.

Thus DNA which is stable enough is a better genetic material for storage of genetic information.

Q. 29. Why is DNA molecule a more stable genetic material than RNA? Explain.

REPLICATION: Replication is the process of formation of carbon copies. DNA replication can occur in three possible methods – conservative, disruptive and semi-conservative.

- In conservative replication the parent structure remains intact. The replica is completely new structure.
- In disruptive or dispersive replication, the parent structure fragments and two new structures are formed afresh.
- In semi-conservative replication, one half of the parent passes into each replica while the second half is built anew.

27. What are the three possible methods of DNA replication? 3 (CoHSEM, 2014)

DNA REPLICATION IS SEMICONSERVATIVE: As proposed by Watson and Crick, DNA replication is semi-conservative (a type of replication in which one strand of the daughter duplex is derived from the parent while the other strand is formed as new).

EXPERIMENTAL PROOF:

1) **Meselson and Stahl Experiment:** It is now proven that DNA replicates semiconservatively.

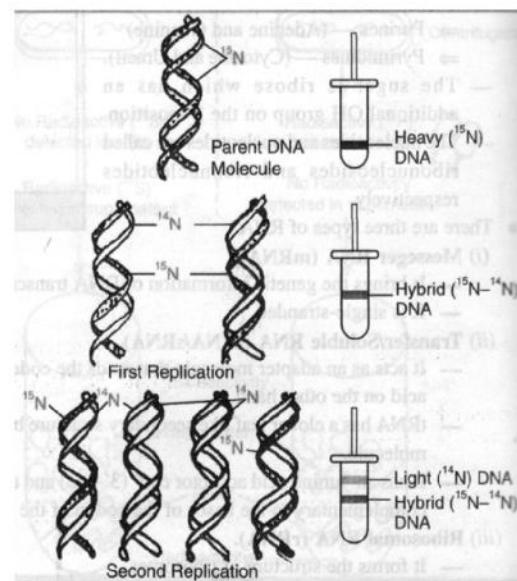
It was shown first in *Escherichia coli* and subsequently in higher organisms, such as plants and human cells. Matthew Meselson and Franklin Stahl performed the following experiment in 1958:

(i) They grew *E. coli* in a medium containing $^{15}\text{NH}_4\text{Cl}$ (^{15}N is the heavy isotope of nitrogen) as the only nitrogen source for many generations. The result was that ^{15}N was incorporated into newly synthesised DNA (as well as other nitrogen containing compounds). This heavy DNA molecule could be distinguished from the normal DNA by centrifugation in a cesium chloride (CsCl) density gradient.

(ii) Then they transferred the cells into a medium with normal $^{14}\text{NH}_4\text{Cl}$ and took samples at various definite time intervals as the cells multiplied, and extracted the DNA that remained as double-stranded helices. The various samples were separated independently on CsCl gradients to measure the densities of DNA.

(iii) Thus, the DNA that was extracted from the culture one generation after the transfer from ^{15}N to ^{14}N medium [that is after 20 minutes; *E. coli* divides in 20 minutes] had a hybrid or intermediate density.

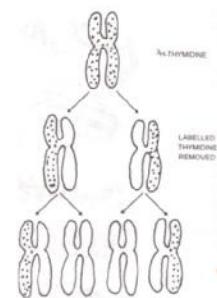
DNA extracted from the culture after another generation [that is after 40 minutes, II generation] was composed of equal amounts of this hybrid DNA and of 'light' DNA. In this way at each replication, one strand is conserved in the daughter while the second is freshly synthesized.



Q. What is semi-conservative DNA replication? How was it experimentally proved and by whom? (CBSE, 2008) 5

Q. If *E. Coli* was allowed to grow for 80 minutes then what would be the proportions of light and hybrid densities DNA molecule?

Semi-conservative Replication of Chromosomes: Taylor *et al* demonstrated the semi-conservative mode of DNA replication and chromosomes in the root tips Broad Bean (*Vicia faba*) with radioactive ^3H containing thymine instead of normal thymine. Thymine is incorporated into DNA which is the structural elements of chromosomes. Taylor *et al* found that all the chromosomes became radioactive. Labelled thymine was then replaced with normal one. Next generation came to have radioactivity in one of the two chromatids of each chromosome while in subsequent generation radioactivity was present in 50% of the chromosomes. This is possible only if out of the two strands a chromosome, one is formed afresh while the other is conserved at each replication. This is semiconservative replication.

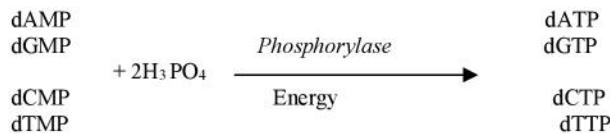


MECHNISM OF DNA REPLICATION: It occurs during S-phase of cell cycle. In prokaryotes, it begins at a particular spot called origin of replication or *Ori*. [ori C (for origin of chromosome replication)].

DNA replication takes place as follows-

1) **Activation of deoxyribonucleotides:**

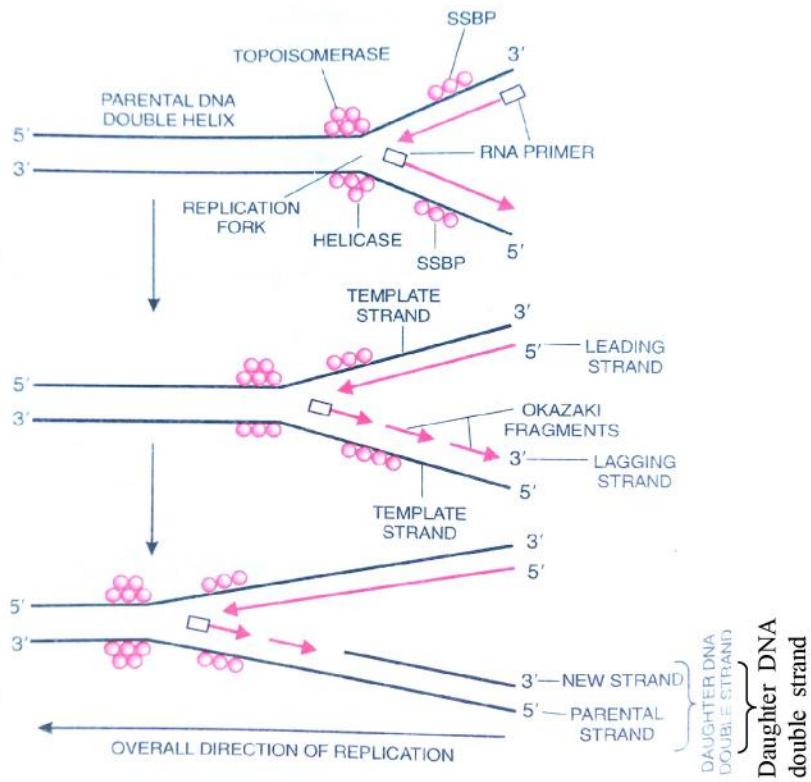
There are 4 types of deoxyribonucleotides –dAMP, dGMP, dTMP & dCMP. With the help of energy, phosphate & an enzyme – *phosphorylase*, the nucleotides are changed into triphosphate state.



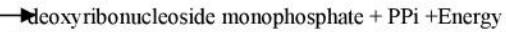
2) Exposure of DNA Strands: Enzyme *helicase* acts over the ori site & unzips the two strands of DNA. The separated strands are stabilized by means of single stranded DNA binding proteins (DBP or SSB). Unwinding creates tension which is relieved by enzymes *topoisomerases I and II*. Topoisomerase II of prokaryotes is also called *gyrase*. Unzipping of the double stranded DNA forms a Y-shaped organized called replication fork. (A branch point in a replication eye at which DNA synthesis occurs is called a replication fork).

3) Formation of RNA primer: A short chain of RNA, called RNA primer is formed on the DNA template at the 5' end with the help of enzyme called primase. Formation of RNA primer is essential because without the presence of RNA primer, DNA polymerases cannot add nucleotides.

4) Base Pairing: Deoxyribonucleotide triphosphates (dATP, dGTP, dCTP and dTTP) come to lie opposite nitrogen base of exposed DNA template- dTTP opposite A, dCTP opposite G, dATP opposite T & dGTP opposite C. Enzyme *pyrophosphatase* separates the two extra phosphates present on deoxyribonucleotides & energy is released in the process which is used in the formation of hydrogen bonds.



Pyrophosphatase



5) New strand Formation: Prokaryotes have three major types DNA synthesizing enzymes called DNA Polymerases III, II and I. They can add nucleotides in 5' → 3' direction on 3' → 5' strand. DNA polymerase III is mainly involved in DNA replication, i.e. addition and polymerization of new bases. DNA Polymerase III and δ can add nucleotide only at 3' end of the chain so that DNA strand formation occurs in 5' to 3' direction. Since the two strands of DNA run in antiparallel directions, replication over the two templates thus proceeds in opposite directions. On one strand with polarity 3' → 5', the synthesis of DNA is continuous. It is called leading strand.

Replication is discontinuous on other template strand with polarity 5' → 3' because only a short DNA segments are formed in the 5' → 3' direction due to exposure of a small stretch of template at one time. These short DNA segments are called **Okazaki fragments**. An RNA primer is required every time a new Okazaki fragment is to be built. After start of nucleotide chain formation, RNA primer is removed & the gap is filled by DNA polymerase I or more rarely RNase H. Okazaki fragments are joined together by means of enzyme, DNA ligase. DNA strand built up of Okazaki fragments is called lagging strand.

Since one strand grows continuously while the other discontinuously, DNA replication is semi-discontinuous.

6) Proof reading: Proofreading is the removal of a mismatched base immediately after it has been added. Though DNA polymerase is very accurate, a wrong nucleotide gets inserted occasionally. Mismatched nucleotides are removed and replaced by correct nucleotides with the help of enzyme DNA polymerase I (Kornberg, 1970) and III (main).

Q. 3. Name the enzyme involved in the continuous replication of DNA strand. Mention the polarity of the template strand. 1 (CBSE,2010)

Leading Strand	Lagging Strand
1. It is a replicated strand of DNA which grows continuously without any gap. 2. It does not require DNA ligase for its growth. 3. The direction of growth of leading strand is 5' → 3'. 4. Only a single RNA primer is required. 5. Formation of leading strand is quite rapid. 6. Its template opens in 3' → 5' direction. 7. Formation of leading strand begins immediately at the beginning of replication.	1. It is a replicated strand of DNA which is formed in short segments called Okazaki fragments. Its growth is discontinuous. 2. DNA ligase is required for joining Okazaki fragments. 3. The direction of growth of lagging strand is 3' → 5' though in each Okazaki fragment it is 5' → 3'. 4. Starting of each Okazaki fragment requires a new RNA primer. 5. Formation of lagging strand is slower. 6. Its template opens in 5' → 3' direction. 7. Formation of lagging strand begins a bit later than that of leading strand.

TRANSCRIPTION: The process of copying of genetic information from template or antisense strand of the DNA into RNA is called transcription. Simply, formation of RNA over DNA Template is called transcription.

It is meant for taking coded information from DNA to the site where it is required for protein synthesis.

Sense and Antisense Strands:

Out of the two strands of DNA duplex, only one carries the genetic information. It has 3' → 5' polarity and is called template or antisense strand.

The other strand is its mirror copy with a polarity of 5' → 3'. It is called coding or sense strand. The coding strand is no less important. The code present in it is similar to that of mRNA except for the presence of thymine instead of Uracil.

- A DNA sequence is called 'sense' if its sequence is the same as that of mRNA copy that is translated into protein. The sequence on opposite strand is called the antisense sequence.

TRANSCRIPTION UNIT: The segment of DNA that takes part in transcription is called transcription unit. It has 3 components-

- A promoter region in the beginning
- A terminator region in the end and
- A structural gene in between.

Besides a promoter, eukaryotes also require an enhancer. In many cases, the promoter has AT rich sequence called TATA box or Pribnow box *after the name of discoverer*. Promoter region is the proximal area of transcription unit which provides sites for attachment of transcription factors and RNA polymerase.

Terminator region is either palindromic sequences or poly A sequences. There is a site for attachment of rho factor for release RNA polymerase.

Structural gene is the area of template strand that is involved in transcription or formation of RNA. Structural gene is component of that strand of DNA which has $3' \rightarrow 5'$ polarity (as transcription can occur only in $5' \rightarrow 3'$). This strand is also called template strand or master strand or antisense or (-) strand. The other strand which has polarity of $5' \rightarrow 3'$ is displaced during transcription. This non-template strand which does not take part is also called sense or coding strand or plus (+) strand because genetic code present in this strand is similar to genetic code (based on mRNA) except that Uracil is replaced by thymine.

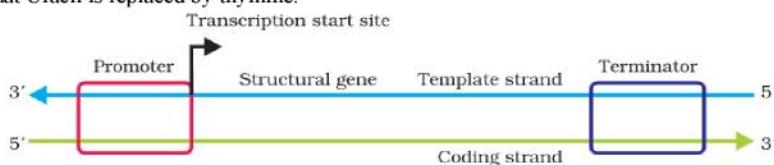


Figure: Components of a transcription unit

ENZYMES AND FACTORS: Transcription requires a DNA dependent enzyme RNA polymerase. Eukaryotes have 3 RNA polymerases, the RNA polymerase I transcribes rRNAs (28S, 18S, and 5.8S), whereas the RNA polymerase III is responsible for transcription of tRNA, 5srRNA, and snRNAs (small nuclear RNAs). The RNA polymerase II transcribes precursor of mRNA, the heterogeneous nuclear RNA (hnRNA).

Prokaryotes have only one RNA polymerase which synthesizes all types of RNAs. RNA polymerase of *E. coli* has 5 polypeptide chains: β , β' , α , ω and σ . The core enzyme is composed of five chains (2α , β , β' , and ω) and catalyses RNA synthesis. When σ factor is bound to core enzyme, the six subunit complex is termed *RNA polymerase holoenzyme*. σ (sigma) factor recognizes the start signal or promoter region (TATA box) of DNA. The part of polymerase enzyme minus σ factor is called core enzyme having 2α , β , β' chains and ω .

Termination of transcription requires a separate termination factor called *rho* (ρ) – factor.

PROCESS OF TRANSCRIPTION:

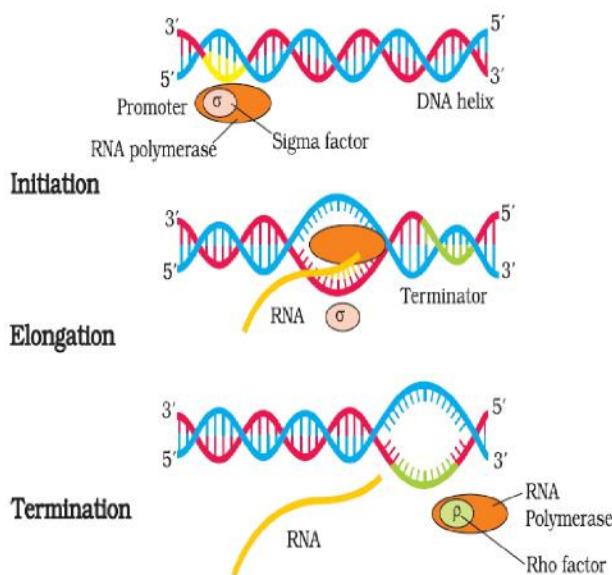


Figure: Process of Transcription in Prokaryotes

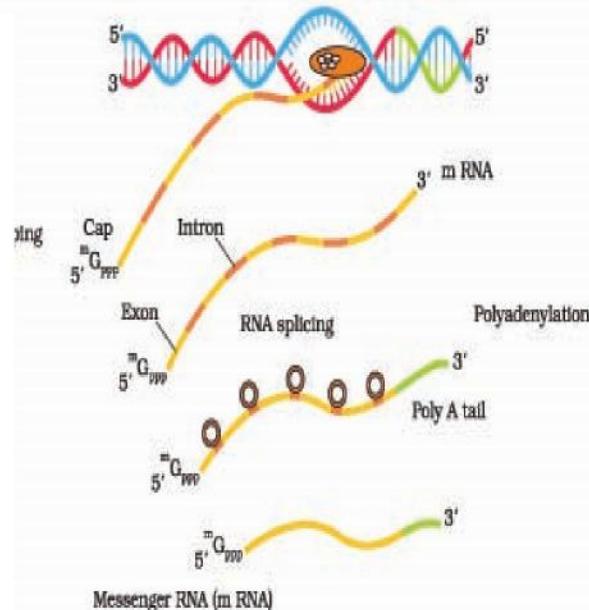
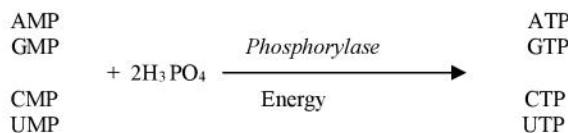


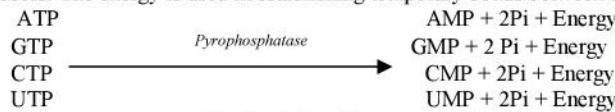
Figure: Process of Transcription in Eukaryotes

1. Activation of Ribonucleotides: Four types of ribonucleotides (AMP, GMP, UMP and CMP) take part in the synthesis of RNA over DNA. Before their incorporation, the nucleotides are converted into active state through phosphorylation. It produces triphosphates. Energy, phosphate and enzyme *phosphorylase* are required.



2. Initiation: The transcription unit begins with a promoter and ends in a terminator. RNA polymerase reaches the promoter region. Sigma (σ) factor recognizes the promoter region. As soon as the enzyme RNA polymerase gets attached, the template DNA of transcription unit begins to unzip. Core enzyme of RNA polymerase travels along the template strand.

3. Base Pairing: The activated phosphorylated ribonucleotides come to lie opposite complementary nitrogen bases of the template strand – ATP opposite T, UTP opposite A, GTP opposite C and CTP opposite G. the two extra phosphates are removed by the enzyme pyrophosphatase and energy is released in the process. The energy is used in establishing temporary bonds between complementary base pairs.



4. Formation of RNA chain: The core enzyme with the help of energy and Mg^{2+} , establishes phosphodiester bonds between adjacent ribonucleotides forming RNA chain. As the enzyme moves along the DNA template, RNA chain elongates. Synthesis continues till the enzyme reaches the termination region.

5. Chain separation: RNA polymerase is separated from the DNA template by means of rho factor (ρ). Rho factor has *ATPase* activity. This separates RNA polymerase as well as the newly built RNA strand. As soon as RNA strand separates, the sense and antisense strands of DNA re-establish hydrogen bonds between their complementary base pairs. The duplex nature is restored.

POST TRANSCRIPTION CHANGES: Freshly synthesized RNA strand is called primary transcript or heterogenous RNA (*hnRNA*, especially in case of mRNA). It is often bigger than the functional RNA. The various modifications which occur in primary transcript are as follows:

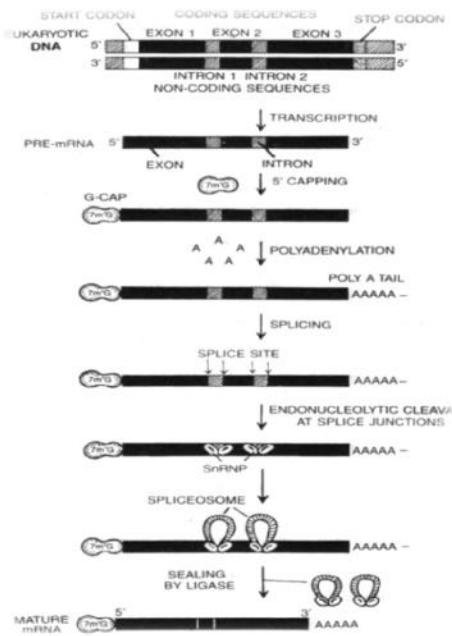
a) Splicing:

Most of the eukaryotic mRNAs contain non-coding segments called introns or intervening sequences (IVS). They are removed. The process of removal of introns through cutting and joining the essential sequences or exons is called splicing. This is carried out with the help of snRNPs (snurps). The latter are formed by association of proteins with small nuclear RNAs (snRNAs). One snRNP binds to 5' end of the intron while another attaches to 3' end of the intron. Each intron starts with dinucleotide GU and ends with AG (GU – AG rule). They are recognized by components of snRNPs.

A spliceosome develops there. With the help of energy from ATP, cuts are made at both 5' and 3' ends of introns. This releases the introns. The ends of the adjacent exons are sealed by *RNA ligase* to produce processed RNA.

b) Terminal Addition: (Capping and Tailing)

Additional nucleotides are added to the ends of RNAs for specific functions, e.g. CCA segment in tRNA, cap nucleotides at 5' end of mRNA or poly – A segments (200 – 300 residues) at 3' end of mRNA. Cap is formed by modification of GTP into 7- methyl guanosine or 7mG.



Prokaryotic Transcription	Eukaryotic Transcription
<ol style="list-style-type: none"> It occurs in contact with cytoplasm. There is no specific period for its occurrence. Processing of released RNA is occurs in cytoplasm. It is coupled to translation. There is only one RNA polymerase. RNA polymerase does not have separate transcription factors. 	<ol style="list-style-type: none"> It occurs inside the nucleus. Major part of transcription occurs in G₁ and G₂ phases. Processing occurs inside the nucleus. Transcription and translation are spatially separated. There are three types of RNA polymerases. Transcription factors are involved in recognition of promoter site.

Template Strand	Coding Strand
<ol style="list-style-type: none"> It is also called antisense or (-) strand or master strand. It has 3' → 5' polarity. It is that strand upon which RNA is transcribed in 5' → 3' direction. 	<ol style="list-style-type: none"> It is called sense strand or coding strand or (+) or nontemplate strand. It has 5' → 3' polarity. It has same sequence of bases found in mRNA except T at the place of U. It does not code any information.

RIBONUCLEIC ACID (RNA): RNA is a single chain polynucleotide which functions as carrier of coded genetic or hereditary information from DNA to cytoplasm for taking part in protein and enzyme synthesis. There are six types of RNAs – ribosomal, transfer, messenger, genetic (genomic), small nuclear and cytoplasmic.

1. Ribosomal RNA (rRNA): It is the most abundant RNA (70 – 80% of total) which has 3 – 4 types. As the name indicates, rRNA is a constituent of ribosomes.

Functions: i) rRNAs bind protein molecules and give rise to ribosomes.

2. Messenger RNA (mRNA): It is a long RNA which brings instructions from the DNA for the formation of particular type of polypeptide. mRNA is therefore, also called informational or genetic RNA. The instructions are present in base sequence of the nucleotides. It is called genetic code. mRNA has methylated region at the 5' terminus. It functions as cap for attachment with ribosome. Cap is followed by an initiation codon (AUG) either immediately or after a small non-coding leader region. Then there is coding region followed by termination codon (UAA, UAG or UGA). After termination codon there is a small non-coding trailer region and poly A area or tail at the 3' terminus. An mRNA may specify only a single polypeptide or a number of them. The former is called monocistronic while the latter is known as polycistronic.

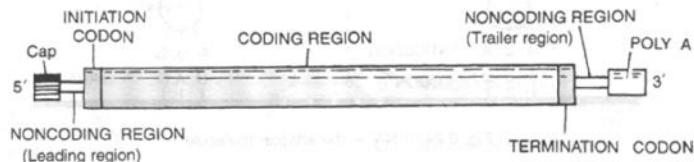


Fig. 6.25. Parts of an mRNA strand.

Functions: i) mRNA carries coded information for translation into polypeptide formation.
ii) Through reverse transcription it can form compact genes which are used in genetic engineering. The phenomenon also occurs in nature and has added certain genes in the genome.
iii) It has a cap region for attachment to ribosome.
iv) Cap protects the mRNA from degradation from enzymes.
v) mRNA has a tail region for protection from cellular enzymes and detachment from ribosome.

3) Transfer RNA (tRNA): It is also called soluble or sRNA in which form it was known before the discovery of genetic code. tRNA is the smallest RNA and sedimentation coefficient of 4S. The nitrogen bases of several of its nucleotides get modified, e.g. pseudouridine (ψ), dihydrouridine (DHU), inosine (I), ribo-thymidine (rT). This causes coiling of the otherwise single stranded tRNA into inverted L-shaped form (three dimensional, Klug 1974) or clover – leaf (two dimensional, Holley, 1965). Five regions are unpaired or single stranded – AA binding site, T ψ C loop, DHU loop, extra arm and anticodon loop.

- i) **Anticodon Loop:** It has 7 bases out of which three bases form anticodon for recognizing and attaching to the codon of mRNA.
 - ii) **AA-Binding Site:** It is amino acid binding site. The site lies at the 3' end opposite the anticodon and has CCA - OH group. The 5' end bears G. AA - binding site and anticodon are the two recognition sites of tRNA.
 - iii) **T_ΨC loop:** It has 7 bases. The loop is the site for attaching to ribosome.
 - iv) **DHU loop:** The loop contains 8 – 12 bases. It is binding site for aminoacyl synthetase enzyme.
 - v) **Extra Arm:** It is a variable side arm or loop which lies between T_ΨC loop and anticodon. It is not present in all tRNAs. The exact role of extra arm is not known.
- Functions:**
- i) tRNA is adapter molecule which is meant for transferring amino acids to ribosomes for synthesis of polypeptides.
 - ii) They hold peptidyl chains over the mRNAs.
 - iii) The initiator tRNA has dual function of initiation of protein synthesis as well as bringing in of the first amino acid.

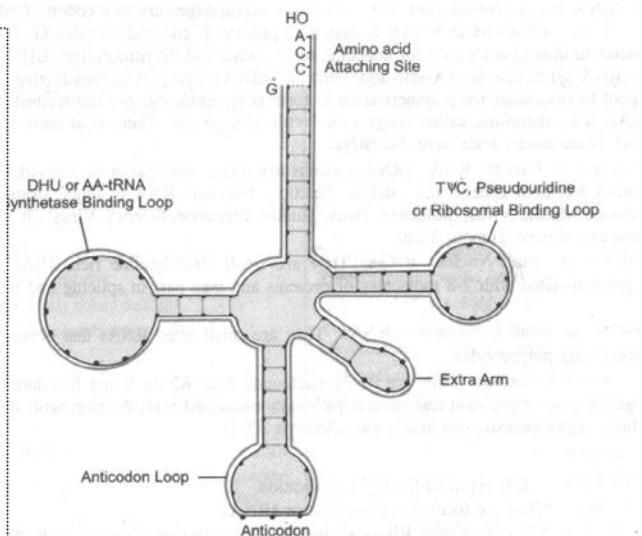


Fig. 6.20. Diagrammatic structure of a tRNA.

Codon	Anticodon
1. It is found in mRNA and DNA. 2. Codon is complementary to a triplet of template strand. 3. It determines the position of an amino acid in a polypeptide.	1. It occurs in tRNA. 2. It is complementary to a codon. 3. It helps in bringing a particular amino acid at its proper position during translation.

Genome = The total genetic content contained in a haploid set of chromosomes in eukaryotes, in a single chromosome in bacteria or in the DNA or RNA of viruses. i.e. an organism's genetic material.

RNA WORLD: RNA was the first genetic material. There is now enough evidence to suggest that essential life processes (such as metabolism, translation, splicing, etc.), evolved around RNA. RNA used to act as a genetic material as well as a catalyst (there are some important biochemical reactions in living systems that are catalysed by RNA catalysts and not by protein enzymes). But, RNA being a catalyst was reactive and hence unstable. Therefore, DNA has evolved from RNA with chemical modifications that make it more stable. DNA being double stranded and having complementary strand further resists changes by evolving a process of repair.

Q.4. Would it be possible to originate the first formed cell if enzymes were not formed before the formation of the first living cell? Give one reason of your answer. 1+1=2 (CoHSEM-2003)

Ans:- Yes, it would be possible. RNA used to act as a genetic material as well as a catalyst (there are some important biochemical reactions in living systems that are catalysed by RNA catalysts and not by protein enzymes).

GENETIC CODE: The relationship between the sequence of amino acids in a polypeptide and nucleotide sequence of DNA or mRNA is called Genetic code.

The salient features of genetic code are as follows:

1. **Triplet Code:** Genetic code is triplet code where three adjacent nitrogen bases specify one amino acid, e.g. UGG for tryptophan, AUG for methionine. 61 codons code for amino acids and 3 codons do not code for any amino acids, hence they function as stop codons.
2. **Non-Overlapping:** Three successive nucleotides or nitrogen bases code for only one amino acid. None of these nitrogen bases become part of any other codon.
3. **Commaless:** Once the reading is commenced at a specific codon, there is no punctuation between codons, and the message is read in a continuing sequence of nucleotide triplets until a translation stop codon is reached. If a nucleotide is deleted or added, the whole genetic code will read differently.
4. **Nonambiguous Codons:** It means that one codon specifies only one amino acid and not any other, and specific.

SECOND BASE				THIRD BASE (3' end)			
FIRST BASE (5' end)							
U	U	C	A	G			
	UUU	UCU	UAU	Try	UGU		U
	UUC	UCC	UAC		UGC	Cys	C
	UUA	UCA	UAA	Stop (ochre)	UGA	Stop (opal)	A
C	UUG	UCG	UAG	Stop (amber)	UGG	Trp	G
	CUU	CCU	CAU	His	CGU		U
	CUC	CCC	CAC		CGC		C
	CUA	CCA	CAA	Gln	CGA	Arg	A
A	CUG	CCG	CAG		CGG		G
	AUU	CCU	CAU	His	CGU		U
	AUC	CCC	CAC		CGC		C
	AUA	CCA	CAA	Gln	CGA	Arg	A
G	AUG	CCG	CAG		CGG		G
	Ile	Leu	Pro	Arg			
	AUU	CUU	CAU	His	CGU		U
	AUC	CUC	CAC		CGC		C
A	AUA	CUA	CAA	Gln	CGA	Arg	A
	AUG	CUG	CAG		CGG		G
	Ile	Leu	Pro	Arg			
	AUU	CUU	CAU	His	CGU		U
C	AUC	CUC	CAC		CGC		C
	AUA	CUA	CAA	Gln	CGA	Arg	A
	AUG	CUG	CAG		CGG		G
	Met	Leu	Pro	Arg			
G	GUU	GCU	GAU	Asp	GGU		U
	GUC	GCC	GAC		GGC		C
	GUA	GCA	GAA	Glu	GGA	Gly	A
	GUG	GCG	GAG		GGG		G
A	Val	Ala					
	GUU	GCU	GAU	Asp	GGU		U
	GUC	GCC	GAC		GGC		C
	GUA	GCA	GAA	Glu	GGA	Gly	A
C	GUG	GCG	GAG		GGG		G
	Val	Ala					

5. Degeneracy of code: The code is degenerate. For a particular amino acid more than one codon can be used. One amino acid has more than one code triplet. e.g. Phenylalanine has two codons i.e. UUC and UUU. Similarly Arginine has 6 codons.

6. Universal: The code is nearly universal, i.e. a codon specifies the same amino acid from bacteria to a tree or human being. E.g. bacteria to human UUU would code for Phenylalanine (phe). Some exceptions to this rule have been found in mitochondrial codons, and in some protozoans.

7. Initiation Codon: Polypeptide synthesis is signaled by an initiation codon – commonly AUG or methionine codon and rarely GUG or valine codon. AUG has dual functions- It codes for Methionine (met) , and it also acts as *initiator codon*.

8. Termination codon: Polypeptide chain termination is signaled by any of the three termination codons – UAA, UAG and UGA. They do not specify amino acid and hence also called nonsense codons.

Initiation Codon	Termination codon
1. These codons are found at 5' end of mRNA. 2. Mostly AUG (occasionally, GUG, UUG or CUG) is the initiation codon. 3. It starts the initiation of protein synthesis.	1. These are found at 3' end of mRNA. 2. UAA, UAG and UGA are three termination codons and only one is present at 3' end. 3. It stops the process of protein synthesis.

PROTEIN SYNTHESIS (Translation)

Machinery for Protein Synthesis: It consists of ribosomes, amino acids, mRNAs and *amino acyl tRNA synthetase*. mRNA functions as a template having genetic information. Ribosome is the site of protein synthesis.

1. Ribosomes: Protein synthesis occurs over the ribosomes. Ribosomes are, therefore, called protein factories. Each ribosome has two unequal parts, small and large. The different parts of a ribosome connected with protein synthesis are:

- i) A tunnel for mRNA. It lies between two subunits.
- ii) A groove for the passage of newly synthesized polypeptide. The groove is part of the larger subunit.
- iii) There are three reactive sites – P (D), A and E. P site (peptidyl transfer or donor site) is jointly contributed by the two ribosomal subunits. A – site (amino-acyl or acceptor site) is situated on larger subunit of ribosome. It faces the tunnel between the two subunits. E or exit site is part of larger subunit facing the tunnel.
- iv) Enzymic *peptidyl transferase* is ribozyme. It is component of larger subunit of ribosome (23S rRNA in prokaryotes).
- v) Smaller subunit of ribosome has a point for recognizing mRNA and binding area for initiation factors.

Ribosomes usually form rosette or helical groups during active protein synthesis. They are known as polyribosomes or polysomes.

2. Amino Acids: There are some 20 amino acids and amides which constitute building blocks or monomers of proteins. They occur in cellular pool.

3. mRNA:

4. tRNA:

5. Amino-Acyl tRNA-Synthetase: It is the enzyme that helps in combining amino acid to its particular tRNA. The enzyme is specific for each amino acid. It is also called *aa-activating enzyme*.

MECHANISM OF PROTEIN SYNTHESIS

1. Activation of Amino acids: In presence of Mg^{2+} and ATP, an amino acid gets attached to a specific enzyme *amino acyl tRNA synthetase*. The enzyme is specialized to recognize only one specific amino acid. It produces *aminoacyl-adenylic acid – enzyme complex*. Pyrophosphate is released. The latter breaks up to release energy.



2. Charging of tRNA: The complex reacts with tRNA specific for the amino acid to form aminoacyl-tRNA complex. Enzyme and AMP are released. The tRNA complexed with amino acid is sometimes called *charged tRNA*. The amino acid is linked to 3'-OH end of tRNA through its –COOH group.



Aminoacyl adenylate enzyme

3. Initiation: It requires factors called **initiation factors**. mRNA attaches itself to smaller subunit of ribosome in the region of its cap. The attachment is such that initiation codon of mRNA (AUG or GUG) comes to lie at P-site. Initiation factor already present smaller subunit catalyses the reaction (eIF2 in eukaryotes and IF3 in prokaryotes). Aminoacyl tRNA complex specific for the initiation codon (methionine-tRNA or valine-tRNA) reaches the P-site. Anticodon (e.g. UAC of tRNA^{Met}) establishes temporary hydrogen bonds with the initiation codon (e.g. AUG of mRNA. The codon anticodon reaction occurs in presence of initiation factor eIF₃ in eukaryotes and IF₂ in prokaryotes.

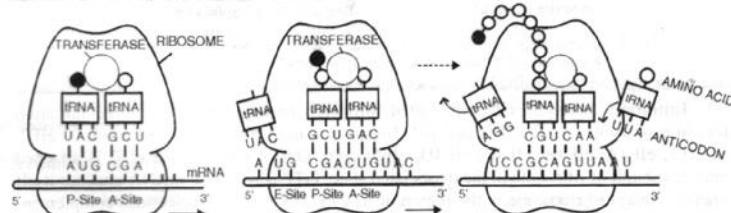


Fig: Elongation of the polypeptide chain.

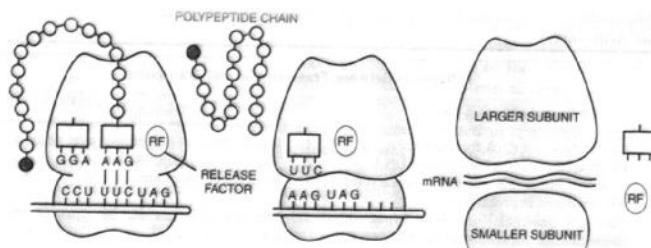


Fig: Termination of polypeptide chain.

4. Elongation: An aminoacyl tRNA complex reaches the A – site and attaches to mRNA codon next to initiation codon with the help of its anticodon. The step requires GTP and an elongation factor (eEF1 in eukaryotes and EF – Tu as well as EF-Ts in prokaryotes). A peptide bond ($-CO - NH -$) is established between the carboxyl group ($-COOH$) of amino acid attached to tRNA at P-site and amino group ($-NH_2$) group of another amino acid attached to tRNA at A-site. The reaction is catalysed by the enzyme peptidyl transferase (an RNA enzyme). In the process, the connection between tRNA and amino acid at the P-site breaks. The freed tRNA of the P-site slips to E-site and from there to the outside of ribosome with the help of G-factor. The A – site carries peptidyl tRNA complex. Soon after the establishment of first peptide linkage, the ribosome or mRNA rotates slightly. The process is called translocation. It requires a factor called translocase (EF-G in prokaryotes and eEF2 in eukaryotes) and energy from GTP. As a result of translocation, the freed tRNA of P-site is slipped to outside and the A-site codon alongwith peptidyl-tRNA complex reaches the P-site.

A new codon is exposed at the A-site. It attracts a new aminoacyl tRNA complex. The process of bond formation and translocation is repeated. One by one all the codons of mRNA are exposed at the A-site and get decoded through incorporation of amino acids in the peptide chain. The peptide chain elongates.

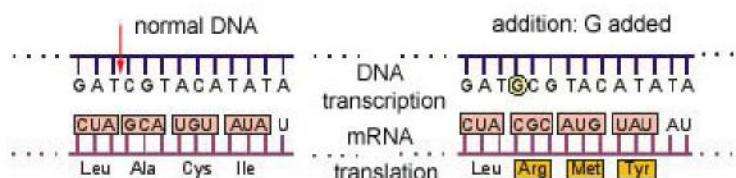
5. Termination: Polypeptide synthesis is terminated when a termination codon (UAA or UAG or UGA) of mRNA reaches the A-site. These codons are not recognized by any tRNAs. Therefore, no more aminoacyl tRNA reaches the A-site. The P-site tRNA is hydrolysed and the completed polypeptide is released in the presence of GTP-dependent release factor. The two subunits of ribosome separate or undergo dissociation in the presence of dissociation factor.

Transcription	Translation
1. It is the formation of RNA from DNA. 2. The template is antisense strand of DNA. 3. It occurs inside the nucleus in eukaryotes and cytoplasm in prokaryotes. 4. The raw materials are four types of ribonucleotide triphosphates – ATP, GTP, CTP and UTP. 5. Transcription requires RNA polymerases and some transcription factor.	1. It is the synthesis of polypeptide over ribosome. 2. The template is mRNA. 3. It occurs in cytoplasm. 4. The raw materials are 20 types of amino acids. 5. All the three types of RNAs take part in translation.

Mutations and Genetic Code: Mutation is a phenomenon which results in alteration of DNA sequences and consequently results in changes in the genotype and the phenotype of an organism. In addition to recombination, mutation is another phenomenon that leads to variation in DNA.

Loss (deletions) or gain (insertion/duplication) of a segment of DNA, results in alteration in chromosomes. Since genes are known to be located on chromosomes, alteration in chromosomes results in abnormalities or aberrations. Chromosomal aberrations are commonly observed in cancer cells.

In addition to the above, mutation also arise due to change in a single base pair of DNA. This is known as *point mutation*. A classical example of such a mutation is sickle cell anemia. Deletion and insertion of base pairs of DNA, causes frame-shift mutations. There are many chemical and physical factors that induce mutations. These are referred to as mutagens. UV radiations can cause mutations in organisms – it is a mutagen.



Gene: The term gene was introduced by **Johannsen** in 1909. **Johannsen** defined gene as an elementary unit of inheritance which can be assigned to a particular trait. Presently, a gene is defined as a unit of inheritance composed of a segment of DNA or chromosome situated at a specific locus (gene locus) which carries coded information associated with a specific function and can undergo crossing over as well as mutation.

Cistron: A cistron is a segment of the DNA molecule carrying information for the production of one polypeptide, or a functional RNA molecule. Currently such a gene is called structural gene.

Or, A **cistron** can be defined as a segment of DNA coding for a polypeptide, the structural gene in a transcription unit could be said as **monocistronic** (mostly in eukaryotes) or **polycistronic** (mostly in bacteria or prokaryotes).

Split Gene: Split genes are those genes which possess extra or nonessential regions interspersed with essential or coding parts. The nonessential parts are called introns, spacer DNA or intervening sequences (IVS). Essential or coding parts are called exons. Split genes are characteristic of eukaryotes.

Transposons: They are segments of DNA that can jump or move from one place in the genome to another. Transposons were first discovered by Mc Clintock (1951) in case of Maize when she found that a segment of DNA moved into gene coding for pigmented kernels and produced light coloured kernels.

One Gene One Enzyme Hypothesis: In 1948, *Beadle and Tatum* proposed “One gene one enzyme hypothesis” which states that a gene controls (metabolic machinery) a structural or functional trait through controlling the synthesis of a specific protein or enzyme. This laid the foundation of biochemical genetics. Beadle and Tatum were awarded Nobel Prize in 1958.

One Gene One Polypeptide Hypothesis: Enzyme *tryptophan synthetase* of *E. coli* is made up of two polypeptides α and β , controlled by two different genes (*trp A* and *trp B*). Haemoglobin is similarly made of two α and two β polypeptides synthesized through two separate genes.

Therefore, Yanofsky et. al (1965) proposed one gene one polypeptide hypothesis (*a gene is a unit of hereditary material that specifies the synthesis of a single polypeptide*). The hypothesis states that a structural gene specifies the synthesis of a single polypeptide.

Central Dogma: It is the flow of information from DNA to mRNA (transcription) and then decoding the information present in mRNA in the formation of polypeptide chain or protein (translation).

The concept of central dogma was advanced by Crick in 1958. It proposes unidirectional flow of information from DNA to RNA and then to protein.

Reverse transcription (Teminism): Temin (1970) and Baltimore (1970) reported that retroviruses operate a central dogma reverse (inverse flow of information). RNA of these viruses first synthesizes DNA through reverse transcription or **teminism**. DNA then transfers information to RNA which takes part in translation to form polypeptide.

REGULATION OF GENE EXPRESSION: Regulation of gene expression refers to a very broad term that may occur at various levels. Considering that gene expression results in the formation of a polypeptide, it can be regulated at several levels. In eukaryotes, the regulation could be exerted at

- (i) transcriptional level (formation of primary transcript),
- (ii) processing level (regulation of splicing),
- (iii) transport of mRNA from nucleus to the cytoplasm,
- (iv) translational level.

OPERON SYSTEM: An operon is a part of genetic material (or DNA) which acts as a single regulated unit having one or more structural genes, an operator gene, a promoter gene, a regulator gene, a repressor and an inducer or corepressor (from outside). The first operon *lac* – operon was discovered by Jacob and Monod (1961). Later on a number of such operons were discovered, e.g. *trp* – operon, *ara* – operon, etc. Operons are of two types, inducible and repressible.

Inducible Operon system – Lac Operon: An inducible system is regulated unit of genetic material which is switched on in response to the presence of a chemical. It consists of following parts –

1) **Structural genes:** They are those genes which actually synthesize mRNAs. The lactose or *lac* – operon of *Escherichia coli* contains three structural genes (*lacZ*, *lacY* and *lacA*).

- i) The *lacZ* gene codes for beta-galactosidase (β -galactosidase), which is primarily responsible for the hydrolysis of the disaccharide, lactose into its monomeric units, galactose and glucose.
- ii) The *lacY* gene codes for permease, which increases permeability of the cell to β -galactosides.
- iii) The *lacA* gene encodes a transacetylase. Hence, all the three gene products in *lac* operon are required for metabolism of lactose.

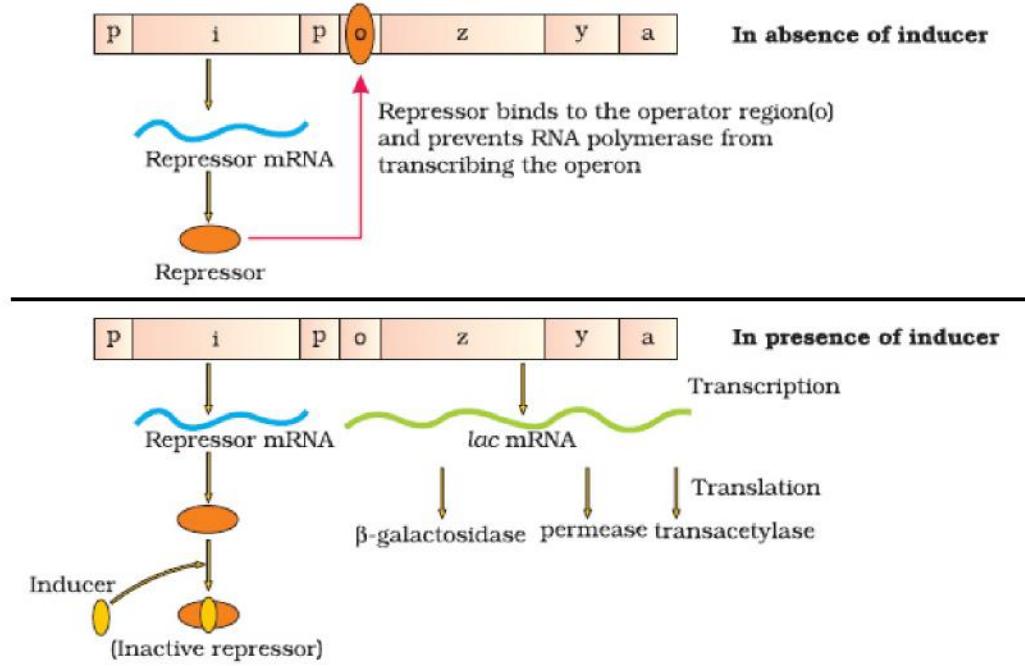


Diagram: Jacob and Monod's model of an inducible Operon. Upper – switched off; Lower- switched on

2) **Operator Gene:** It is a gene which directly controls the synthesis of mRNAs over the structural genes. It is switched off by the presence of repressor. An inducer can take away the repressor and switch on the gene. This gene then directs the structural genes to transcribe.

3) **Promoter Gene:** This gene possesses the site for RNA polymerase attachment.

4) **Regulator Gene (*lac i*-Gene):** In *lac*-operon, it is called *lac i*-gene because it produces an inhibitor or repressor. The *i* gene codes for the repressor of the *lac* operon. The repressor of the operon is synthesised (all-the-time – constitutively) from the *i* gene. The repressor protein binds to the operator gene of the operon and prevents RNA polymerase from transcribing the operon (stops the working of RNA polymerase).

Repressor protein + operator gene \rightarrow Switched off.

Repressor: It is a proteinaceous substance synthesized in all time (constitutively) by the regulator gene. Repressor is meant for blocking the operator gene so that the structural genes are unable to form mRNAs. It has two allosteric sites, one for attaching to operator gene and second for binding to inducer. After coming in contact with inducer the repressor undergoes conformational change in such a way that it is unable to combine with operator.

Inducer: It is a chemical (substrate, hormone or some other metabolite) which after coming in contact with the repressor, changes the latter into non-DNA binding state so as to free the operator gene. If lactose is provided in the growth medium of the bacteria, the lactose is transported into the cells through the action of **permease** (a very low level of expression of *lac* operon has to be present in the cell all the time, otherwise lactose cannot enter the cells). The inducer for *lac*-operon of *Escherichia coli* is lactose.

Inducer (Lactose) + Repressor \rightarrow Switched on

(Glucose or galactose cannot act as inducers for lac operon. The inducer for *lac*-operon of *Escherichia coli* is lactose).

Q. How long the lac operon would be expressed in the presence of lactose?

Q. In the medium where *Escherichia coli* was growing, lactose was added which induced the lac operon. But why does lac-operon shut down sometime after addition of lactose in the medium.

Ans:- *Lac operon* is switched on adding lactose in the medium. Due to this switch on of *lac operon* system, β -galactosidase is formed which converts lactose into glucose and galactose. As soon as lactose is consumed, repressor again become active and cause switch off (shut down) of system.

Q. In Lac operon, what will happen when all lactose molecules are completely being used up?

HUMAN GENOME PROJECT (HGP): Genetic make-up of an organism or an individual lies in the DNA sequences. If two individuals differ, then their DNA sequences should also be different, at least at some places. These assumptions led to the quest of finding out the complete DNA sequence of human genome. With the establishment of genetic engineering techniques where it was possible to isolate and clone any piece of DNA and availability of simple and fast techniques for determining DNA sequences, a very ambitious project of sequencing human genome was launched in the year 1990.

Human genome is said to have approximately 3×10^9 bp, and if the cost of sequencing required is US \$ 3 per bp (the estimated cost in the beginning), the total estimated cost of the project would be approximately 9 billion US dollars. Further, if the obtained sequences were to be stored in typed form in books, and if each page of the book contained 1000 letters and each book contained 1000 pages, then 3300 such books would be required to store the information of DNA sequence from a single human cell. The enormous amount of data expected to be generated also necessitated the use of high speed computational devices for data storage and retrieval, and analysis. HGP was closely associated with the rapid development of a new area in biology called as **Bioinformatics**.

Human Genome Project was a project of sequencing the entire human genome and find out its characteristic features. HGP is called mega project because -

- i) It involved many countries (USA, UK, Japan, France, Germany, China) for determining the nucleotide sequences of genes.
- ii) It required bioinformatics data basing and other high speed computational devices for analysis, storage and retrieval of information.
- iii) As a mega project, it was to sequence 3×10^9 base pairs costing 9 billion US dollars.

Goals: HGP had set the following goals

1. Determine the sequence and number of all the base pairs in the human genome.
2. Identify all the genes present in human genome.
3. Determine the functions of all the genes.
4. Identify the various genes that cause genetic disorders.
5. Store the information in data bases.
6. Improve tools for data analysis.
7. Find out possibilities of transfer of technology developed during HGP to industry.
8. Determine proneness and immunity to various disorders.
9. The project may result in many ethical, legal and social issues (ELSI) which must be addressed and solved.

Salient Features of Human Genome (Revelation of HGP):

- (i) The human genome contains 3164.7 million nucleotide bases.
- (ii) The average gene consists of 3000 bases, but sizes vary greatly. The largest gene is that of Duchenne Molecular Dystrophy on X-chromosome at 2.4 million bases.
- (iii) The total number of genes is estimated at 30,000—much lower than previous estimates of 80,000 to 1,40,000 genes. Almost all (99.9 per cent) nucleotide bases are exactly the same in all people. [Only 0.1% of human genome with some 3.2 million nucleotides represents the variability observed in human beings].
- (iv) The functions are unknown for over 50 per cent of discovered genes.
- (v) Less than 2% of the genome represents structural genes that code for proteins.
- (vi) Chromosome 1 has most genes (2968), and the Y has the fewest (231). They are the maximum and minimum genes for the human chromosomes.
- (vii) Repetitive sequences are stretches of DNA sequences that are repeated many times, sometimes hundred to thousand times. They are thought to have no direct coding functions, but they shed light on chromosome structure, dynamics and evolution.
- (viii) Repeated sequences make up very large portion of the human genome. [Approximately 1 million copies of short 5 – 8 base pairs repeated sequences are clustered around centromeres and near the ends of chromosomes. They represent *junk DNA*].

Applications and Future challenges (Prospects and Implications of Human Genome Projects):

1. Disorder: More than 1200 genes are responsible for common human cardiovascular diseases, endocrine diseases (like diabetes), neurological disorders (like Alzheimer's disease), cancers and many more.
2. Cancer: Efforts are in progress to determine genes that will change cancerous cells to normal.
3. Health Care: It will indicate prospects for a healthier living, designer drugs, genetically modified diets and finally our genetic identity.
4. Study of Tissues: All the genes or transcripts in a particular tissue, organ or tumour can be analysed to know the cause of effect produced in it.

Q. Why is HGP called a mega project?

Satellite DNA: It is that part of repetitive DNA which has long repetitive nucleotide sequences in tandem and forms a separate fraction on density ultracentrifugation. Depending upon the number of base pairs involved in repeat regions, satellite DNA is of two types, **microsatellite** sequences (1 – 6 bp repeat units flanked by conserved sequences) and **minisatellite** sequences (11 – 60 bp flanked by conserved restriction sites). The latter are hypervariable and are specific for each individual. They are being used for DNA matching or finger printing as first found out by Jaffreys *et al* (1985).

Satellites are present in many eukaryotic genomes and represent about 5% of total DNA.

Functions: i) It has some structural function in the chromosome.

ii) Minisatellites are useful for genetic mapping (Jaffrey *et al* 1998) because of high variability.

DNA PROFILING (FINGERPRINTING):

DNA fingerprinting or DNA profiling is the technique used for determining nucleotide sequences of certain areas of DNA which are unique to each individual. DNA fingerprinting was discovered by Alec Jeffreys at University of Leicester. **Dr. V.K. Kashyab** and **Dr. Lalji Singh** started the DNA finger printing technology in India at CCMB (Centre for Cell and Molecular Biology) Hyderabad.

Principle: 0.1% of genome or 3×10^6 differences occur in the base sequences of human beings. They make every individual unique. The differences occur not only in genes but also in repetitive DNA. It is also called *satellite DNA*. On the basis of composition (A:T rich, C:G rich), length and number of repetitive units, satellite DNA is of several types. One of them is hypervariable repeats minisatellite sequences commonly called *variable number tandem repeats* or VNTRs. Each VNTR has 11 – 60 base pairs surrounded by conserved restriction sites. VNTRs differ in families due to some small deletions, insertions and mutations. Each child will receive 50% of DNA from father parent and 50% DNA from mother parent. Therefore, the number of VNTRs in a particular area of the two homologous chromosomes or DNA molecules are likely to be different. As a result each individual comes to have a distinct combination of VNTRs which is peculiar to that person only.

TECHNIQUE:

Sources of DNA: White blood corpuscles of blood, sperm, vaginal swabs, skin cells, cells from hair root, etc. are used as a source of DNA for fingerprinting.

Extraction of DNA: DNA source is exposed to high speed refrigerated ultracentrifugation. It separates the nuclei and breaks the same to free DNA.

DNA Amplification: If the amount of DNA is small, the extracted DNA is amplified through PCR. It produces numerous copies of DNA.

Fragmentation: With the help of site of recognizing restriction enzymes, DNA is cut to separate VNTRs.

Separation of VNTRs: Cut DNA is exposed to electrophoresis over agarose polymer gel. It separates DNA fragments. The separated VNTRs can be recognized by staining them with dye that becomes fluorescent when illuminated with UV radiations. The separated segments are called *restriction fragment length polymorphism (RFLP)*.

Single Stranded DNA: VNTRs are treated with alkaline chemicals to split them into single stranded DNAs.

Southern Blotting: The separated VNTR single stranded sequences are transferred to nitrocellulose or nylon membrane placed over a gel. The procedure is called Southern blotting after the name of its inventor, E.M. Southern.

DNA Probes: They are small radioactive synthetic DNA segments of known sequences of nitrogen bases.

Hybridisation: Nylon sheet or nitrocellulose is immersed in a bath to which DNA probes are added. The probes get attached to single stranded VNTRs having complementary nucleotide sequences.

Exposure to X-ray Films: The nylon membrane containing the radioactive DNA probes and VNTRs is exposed to X-ray film. The hybridized radioactive VNTRs appear as dark bands. The film gives us DNA prints or DNA profiles. The prints are compared and relationship found.

APPLICATIONS:

- 1) Identification: DNA fingerprinting is a sure method of identification of criminals involved in various types of crime including rape and murder.
- 2) Paternity – Maternity Disputes: The method can provide reliable information as to real biological father, mother or offspring.
- 3) Close Relations: DNA fingerprinting can establish the closeness of relation of an intending immigrant.
- 4) Human Lineage: It provides as information as to human lineage and relationship with various apes.

Q.5. A Policeman finds a cluster of hairs with follicle cells from the site of crime and takes it to the forensic department. Another team of policemen captured two suspected persons including the real culprit and sent them to court. But the court could not finalize the case by legal procedure. In what way a biologist can solve the problem? Explain it in brief. 3 (CoHSEM-2010)

Ans:- DNA is isolated from the hair follicle cells. With the help of restriction enzymes, DNA is cut to separate VNTRs. Cut DNA is exposed to Gel electrophoresis to separate DNA fragments. The separated VNTRs can be recognized by staining them with dye. VNTRs are treated with alkaline chemicals to split them into single stranded VNTRs. The separated single stranded VNTRs sequences are transferred to nitrocellulose or nylon membrane placed over a gel. Nylon sheet or nitrocellulose is immersed in a bath to which DNA probes are added. The DNA probes get attached to single stranded VNTRs having complementary nucleotide sequences. The nylon membrane containing the radioactive DNA probes and VNTRs is exposed to X-ray film. The hybridized radioactive VNTRs appear as dark bands. The film gives the DNA prints of real culprit.

The same procedure is repeated for those two suspected persons. And DNA fingerprints are compared. The DNA fingerprint taken from hair follicle cells will be exactly identical to that of real culprit. Thus the real criminal can be identified.

30. Two blood samples A and B picked up from the crime scene were handed over to the forensic department for genetic fingerprinting. Describe how the technique of genetic fingerprinting is carried out. How will it be confirmed whether the samples belonged to the same individual or two different individuals? 5 (CBSE,2010)

VNTR	Probe
1. It is a natural small sequence of DNA. 2. VNTR is nonradioactive. 3. VNTRs help in identification of a person.	1. It is synthetic DNA fragment. 2. It is a radioactive. 3. Probes help in identification of VNTRs.

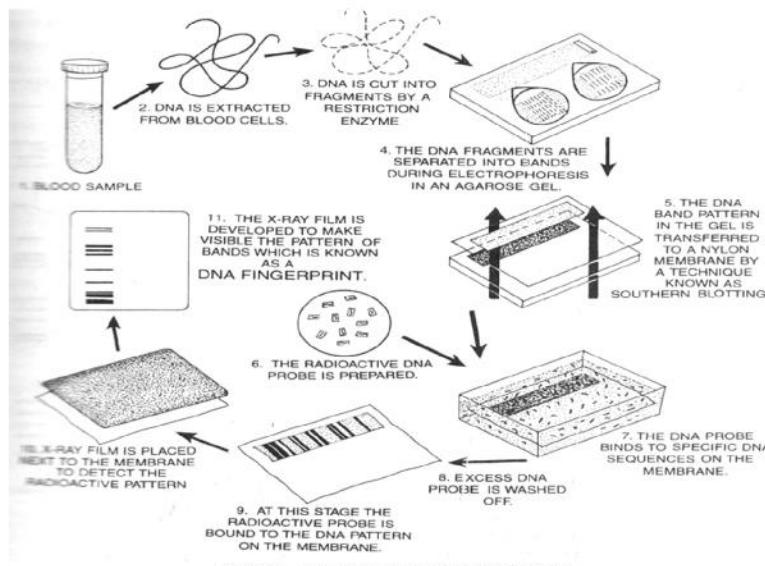


Fig. 6.40. The DNA Fingerprinting Process.

Q. A short length of DNA molecule has 80 thymine and 80 guanine bases. The total number of nucleotides in the DNA fragment is 1
 A) 40 B) 80 C) 160 D) 320

Q. Recall the experiments done by Frederick Griffith, Avery, Mac Leod and Mc Carty, where DNA was speculated to be the genetic material. If RNA, instead of DNA was the genetic material, would the heat killed strain of *Streptococcus* have transformed R-strain into virulent strain? Explain. 1

Ans:- No. Since RNA is labile (not thermostable), it would be disintegrated when heated.