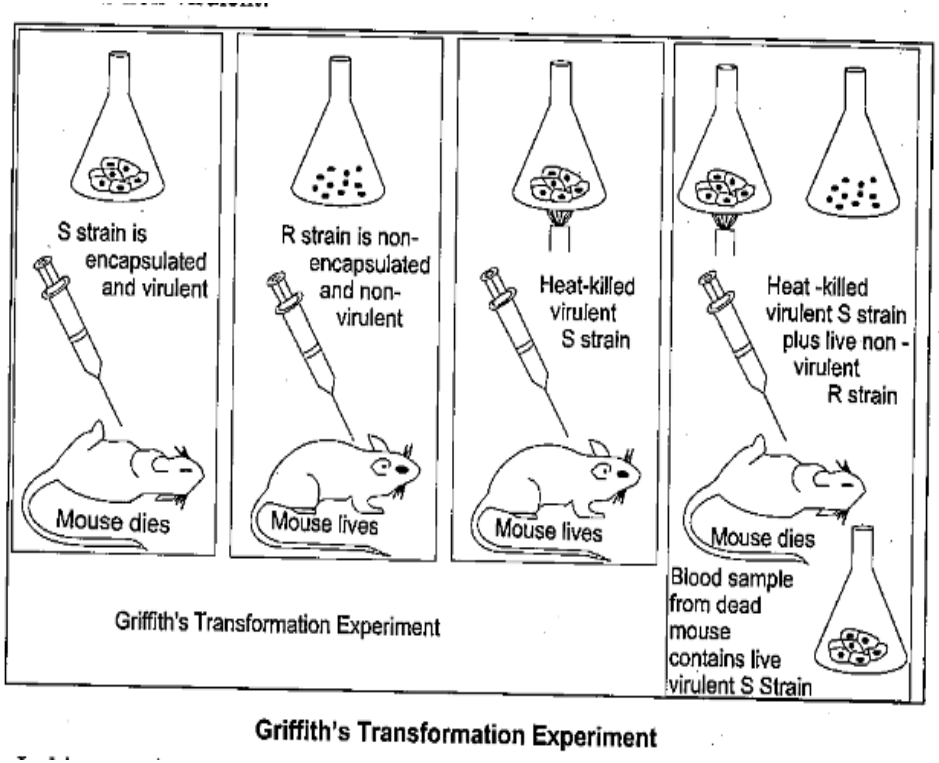
MOLECULAR BASIS OF INHERITANCE

THE DISCOVERY OF DNA

- The discovery of DNA has evolved from the discovery of nucleic acid.
- In 1869, Friedrich Miescher separated a cellular substance from the nuclei of pus cells and called it **Nuclein.**
- Chemically, nuclein has a high phosphorus content and showed acidic properties. Hence it was named as nucleic acid.
- There are two kinds of nucleic acids found in cells i.e **DNA**(deoxyribonucleic acid) and RNA (ribonucleic acid)

GRIFFITH'S EXPERIMENT

- In 1928, Frederick Griffith performed an experiment on bacterium *Streptococcus pneumoniae* that causes pneumonia in humans and other mammals.
- Griffith used two strains of *Streptococcus* i.e
  I. Virulent, Smooth, pathogenic and encapsulated **S-type**
  II. Non-virulent, Rough, non-pathogenic and non-capsulated **R-type**.

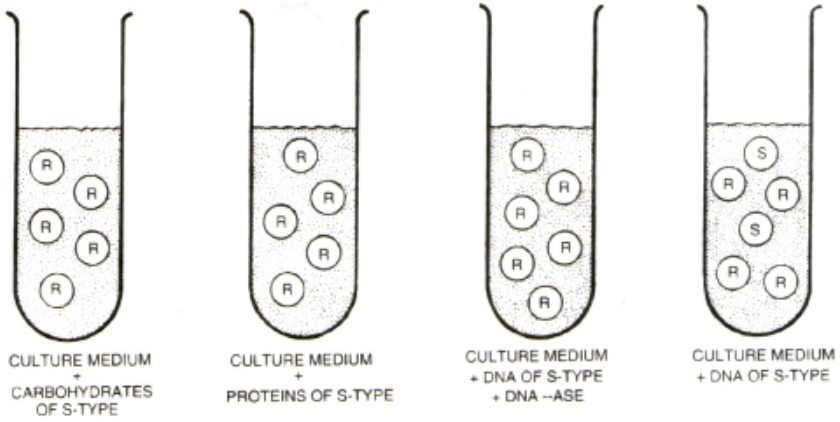


Griffith's Transformation Experiment

Griffith concluded that the R-type bacterium must have taken up, to what he called a "**transforming principle**" from the heat- killed S- bacterium which allowed R-strain to get transformed into Smooth-coated bacterium and become virulent.

AVERY, McCARTY MACLEOD'S EXPERIMENT

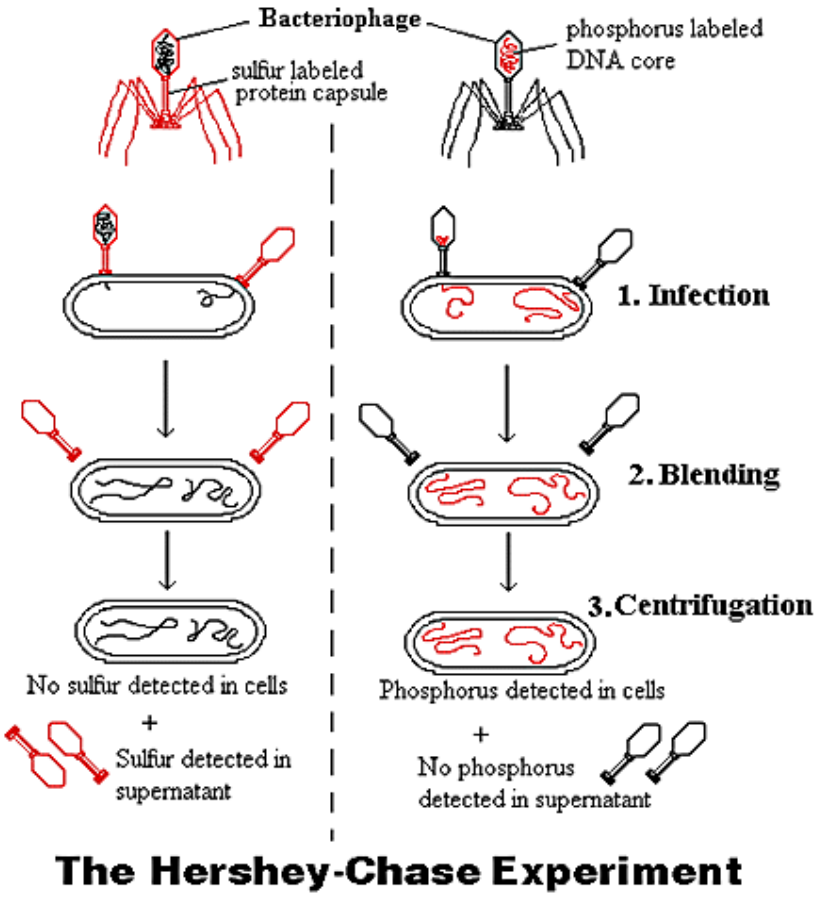
- In 1944, after some 10 years of research and experimentation, **Oswald Avery**, **Collin MacLeod** and **Maclyn McCarty** first evidenced to prove the DNA is a genetic material through their experiments.



These experiments prove that the tranforming principle is DNA but all biologists were not convinced.

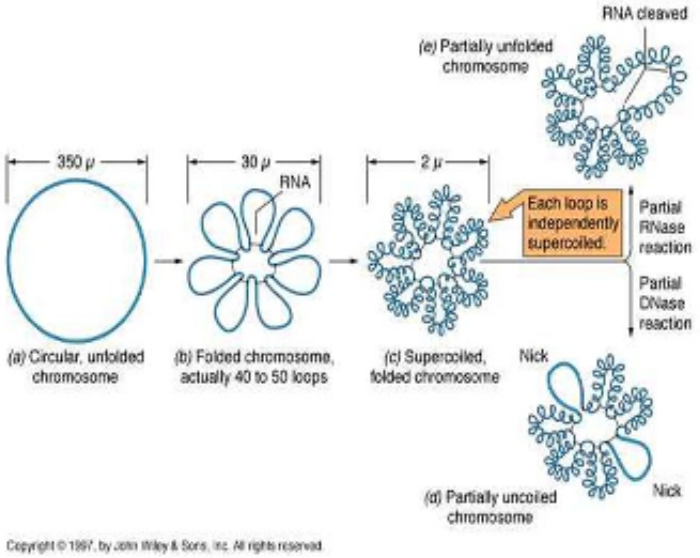
HERSHEY-CHASE EXPERIMENT

- Hershey and Chase worked with viruses that infect bacteria i.e bacteriophages, which are composed of DNA and protein. They used radioactive phosphorus ^{32}P in the medium for some viruses and radioactive sulphur ^{35}S for some others.



PACKAGING OF DNA IN PROKARYOTES

1. In prokaryotes like E.coli, cell size is almost 2-3µ long.
2. They do not have well organized nucleus.
3. It is without nuclear membrane and nucleolus.
4. The nucleoid is small, circular, highly folded naked ring of DNA which is 1100µ long in perimeter, containing about 4.6 million base pairs.



5) The 1100µ long nucleoid is to be fitted or packaged into a cell which is hardly 2-3µ long. Hence the negatively charged DNA becomes circular, reducing the size to 350µm in diameter.

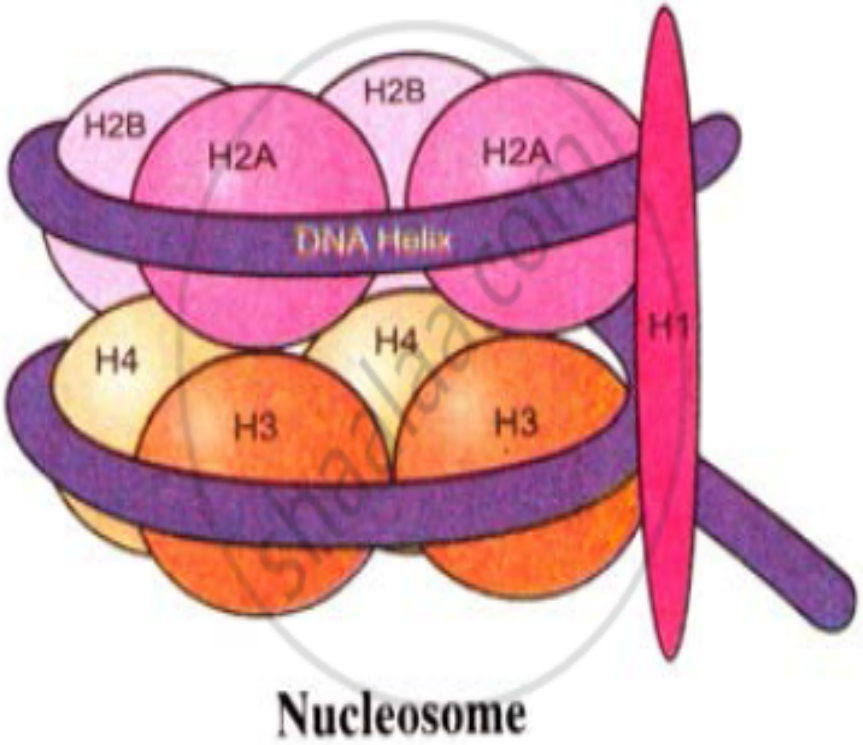
6) This is further reduced to 30µm in diameter because of folding/looping.

7) 40-50 domains(loops) are formed.

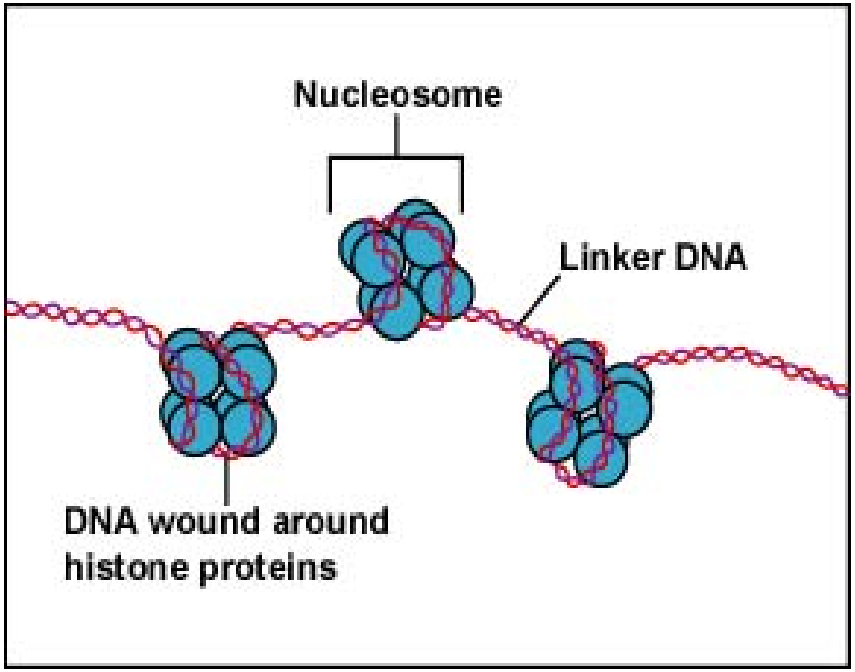
8) Each domain if further coiled and supercoiled thereby reducing the size down to 2µ in diameter.

9) This coiling is assisted by positively charged HUprotein(Histone like DNA binding proteins)and enzymes like DNA Gyrase and Topoisomerase I, for maintaining super coiled state.

PACKAGING OF DNA IN EUKARYOTES



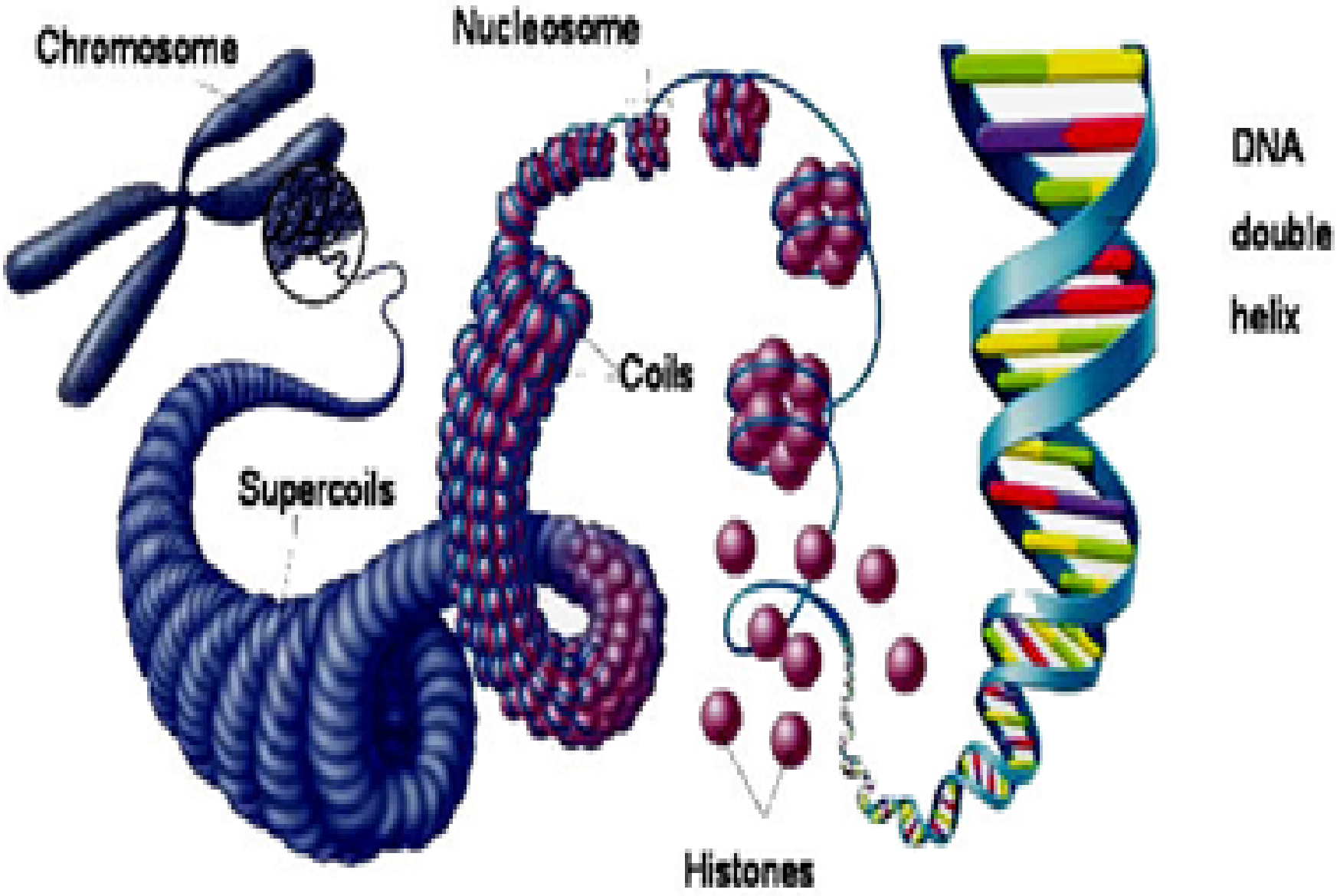
Nucleosome



Beads- on- string magnified

1. The organization of DNA is much more complex in eukaryotes.
2. Histones are required for the packaging of DNA.
3. Histones are proteins that are rich in the basic amino acid residues lysine and arginines which carry positive charge in their side chain.
4. Eight molecule of histones (two each of H_2A, H_2B, H_3 and H_4) get organized to form **histone octamer.**
5. DNA is negatively charged and it is wrapped around the positively charged **histone octamer** forming a structure known as **Nucleosome.**

6. H1 protein binds the DNA thread where it enters and leaves the Nucleosome.
7. Under the electron microscope, nucleus shows Chromatin network, the nucleosomes in Chromatin are seen as ' **Beads-on- string**'.
8. Around the octamer, DNA molecule is wrapped as 1 and 3/4th turn. This DNA is called **Core DNA** and it consists of about 146bp (base pairs).
9. Adjacent Nucleosomes are linked with small segment of DNA called **Linker DNA**; of about 54bp.
10. This 'beads-on-string' structure gets condensed into **nucleosome fiber** which is coiled like a telephone wire to make **Solenoid fiber** with diameter 30nm or 300Å.


Chromosome
Nucleosome
DNA
double
helix
Coils
Supercoils
Histones

- The packaging of chromatin at higher levels need additional set of proteins that are collectively called **Non-Histone Chromosomal** (NHC) proteins.
- A loosely packed region of chromatin that stains light, is called **Euchromatin** and densely packed region that stains dark is called **Heterochromatin**.
- Euchromatin is considered as transcriptionally active chromatin, while Heterochromatin is inactive.

DNA REPLICATION

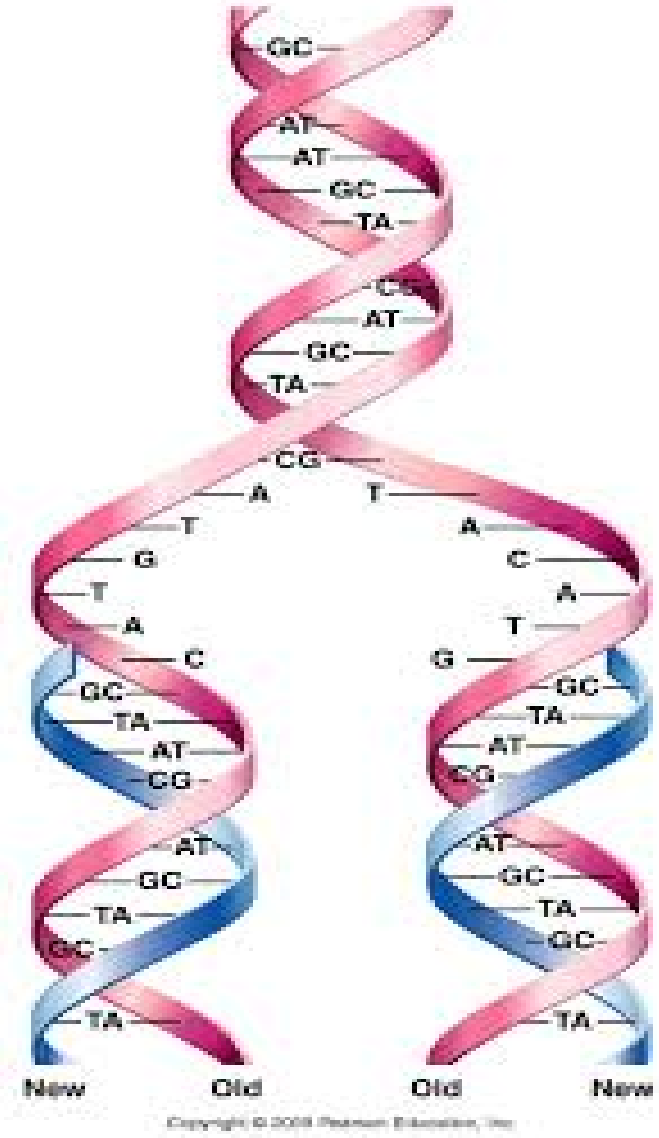
- The DNA molecule regulates and controls all the activities of the cell. As a carrier of genetic information, DNA has to perform two important functions:-

A. **Heterocatalytic function:-** when DNA directs the synthesis of chemical molecules other than itself, then such functions of DNA are called heterocatalytic functions.

Eg. Synthesis of RNA(Transcription), synthesis of protein (Translation), etc.

B. **Autocatalytic function:-** when DNA directs the synthesis of DNA itself, then such function of DNA is called autocatalytic function.

- **The process by which DNA duplicates itself is called replication. Through replication, it forms two copies that are identical to each other.**
- **In eukaryotic organisms, replication of DNA takes place only once in the cell cycle.**
- **It occurs in the S-phase of Interphase in the cell cycle.**
- **DNA replicates through semiconservative mode of replication.**
- **The model for Semiconservative replication was proposed by Watson and Crick, on the basis of antiparallel and complementary nature of DNA strands.**



- The process of semiconservative replication is as below:-

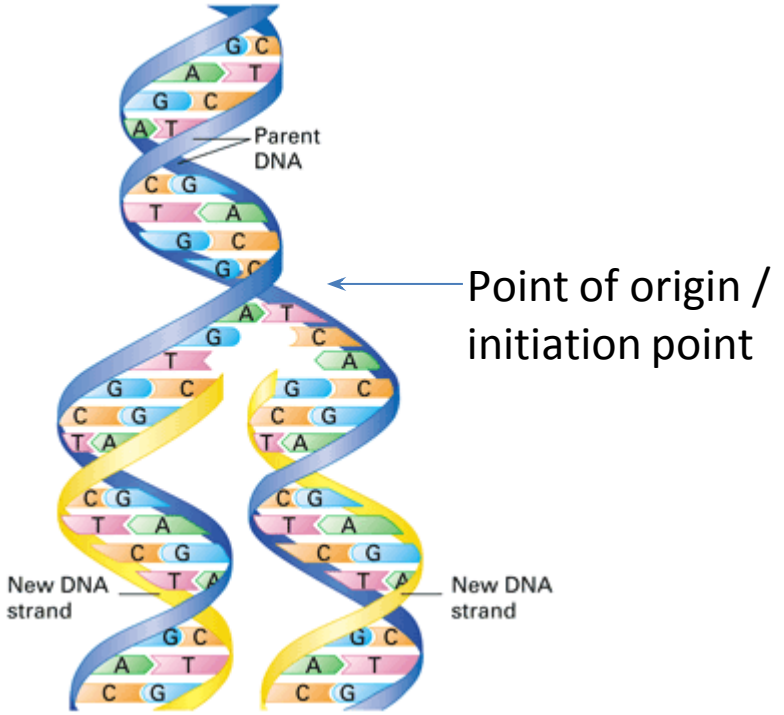
1. Activation of Nucleotides:-

- Four types of DNA nucleotides are present in the Nucleoplasm- dAMP, dGMP, dCMP and dTMP.
- They are activated into triphosphates like dATP, dGTP, dTTP and dCTP using ATP in the presence of enzyme phosphorylase. The process is known as Phosphorylation.

dAMP+ATP $\xrightarrow{\text{phosphorylase}}$ dATP + AMP

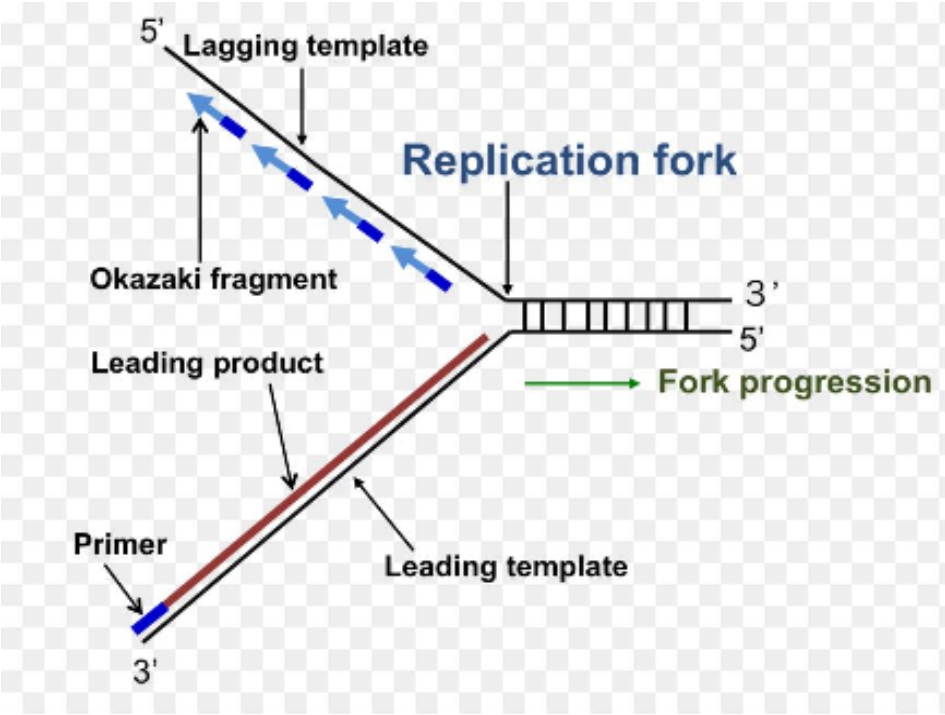
2.Point of origin or Initiation point

- DNA replication begins at certain specific sites called ‘O’-Origin.
- In prokaryotes, there is only one origin however in Eukaryotes, there are more than one origin.
- At the point ‘O’, enzyme Endonuclease nicks one of the strands of DNA, temporarily.



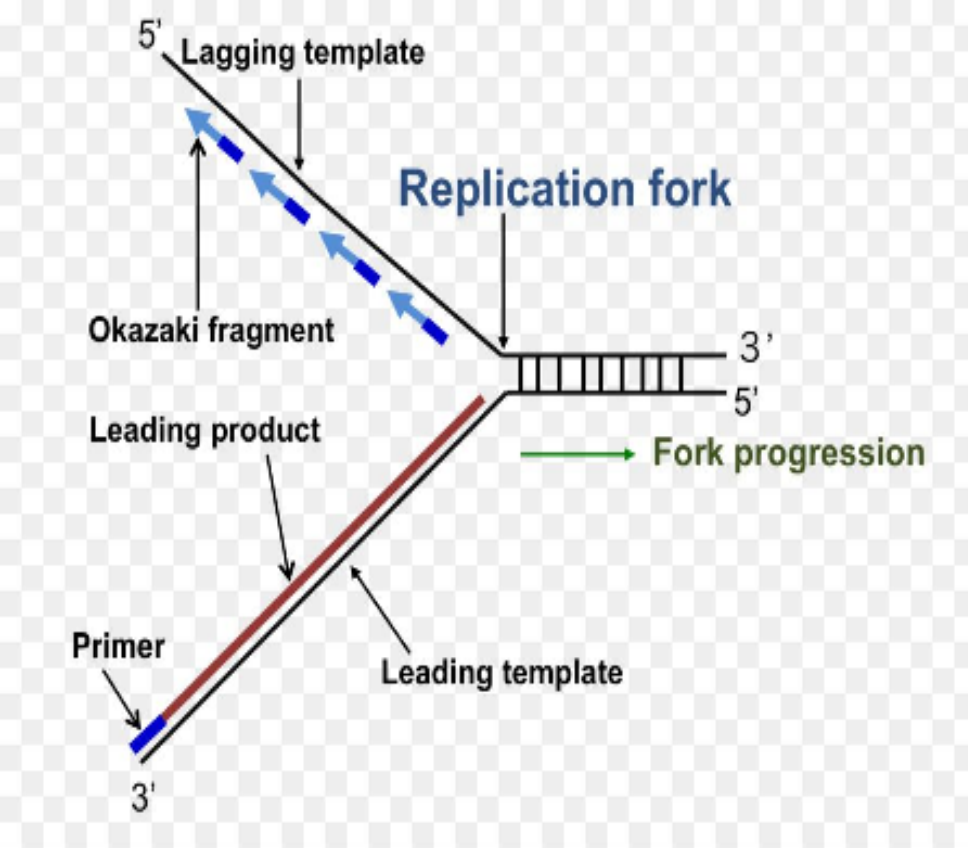
3. Unwinding of DNA molecule

- The enzyme **DNA helicase** breaks the hydrogen bonds due to which the strands of DNA separate and unwind.
- Now the DNA molecule appears as inverted ‘Y’-shaped structure called **Replication fork**.
- The separated strands are prevented from recoiling by **SSBP**(single strand DNA binding protein)



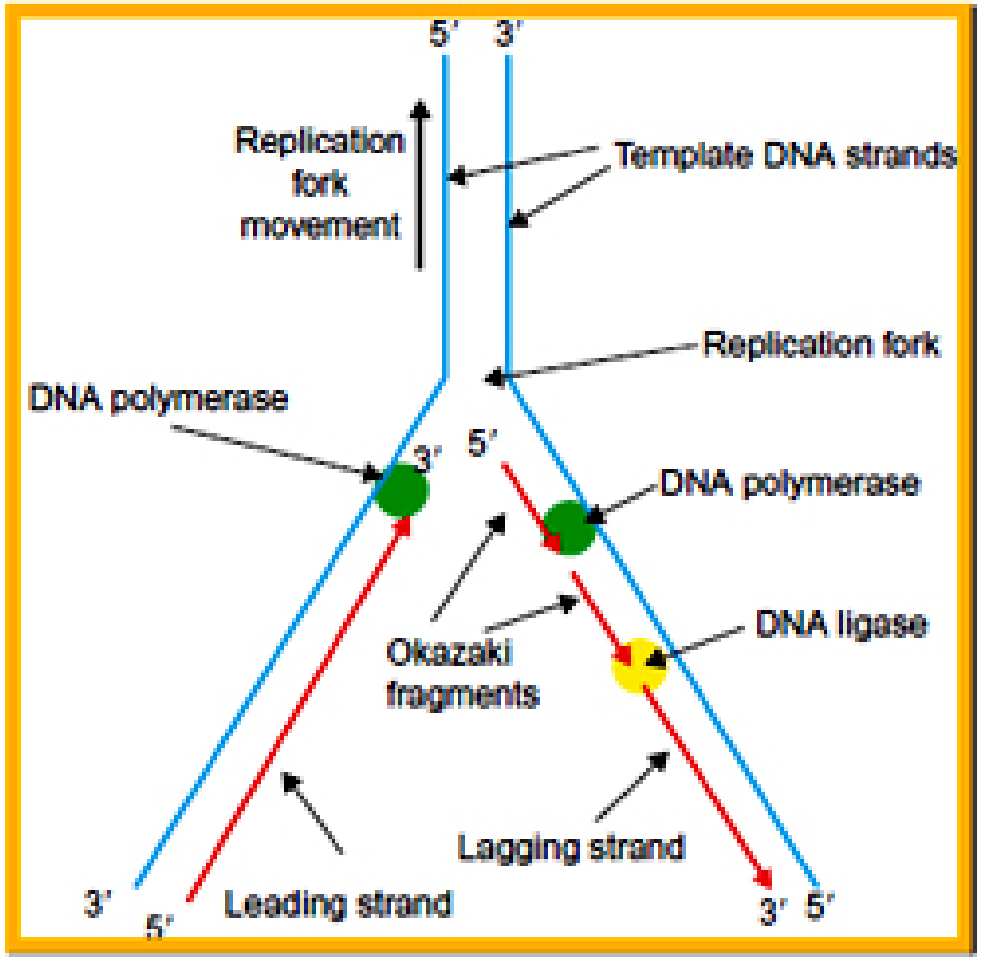
4. Synthesis of new strands

- Each separated strand acts as a template or mould for the synthesis of new complementary strand.
- It takes place with the help of a small RNA molecule called **RNA Primer**.
- **RNA Primer** get associated with the 3' end of template strand and attracts complementary nucleotides from surrounding Nucleoplasm.
- The synthesis of new complementary strand is catalyzed by DNA Polymerase.
- The new complementary strand is always formed in 5' → 3' direction.



5. Leading and lagging strand

- The Template strand with free 3' end is called **Leading template** and with free 5' end is called **Lagging template**.
- As both the strands of the parental DNA are antiparallel, new strands are always formed in 5' → 3' direction.
- One of the newly synthesized strand develops continuously towards replicating fork is called **leading strand**.
- Another new strand develop discontinuously away from the replicating fork is called **Lagging strand.**
- DNA synthesis on lagging template takes place in the form of small fragments, called **Okazaki fragments**. Okazaki fragments are joined by enzyme **DNA Ligase**.
- RNA primers are removed by DNA polymerase and replaced by DNA sequence with the help of **DNA polymerase – I** in Prokaryotes and **DNA polymerase – α** in Eukaryotes.
- Finally, **DNA Gyrase** enzyme forms double helix to form daughter DNA molecules.

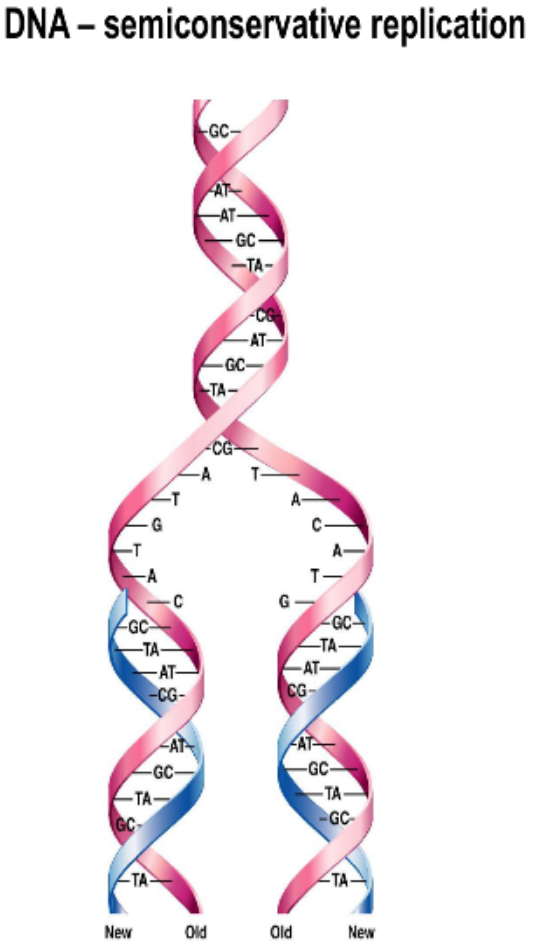


6. Formation of daughter DNA molecule

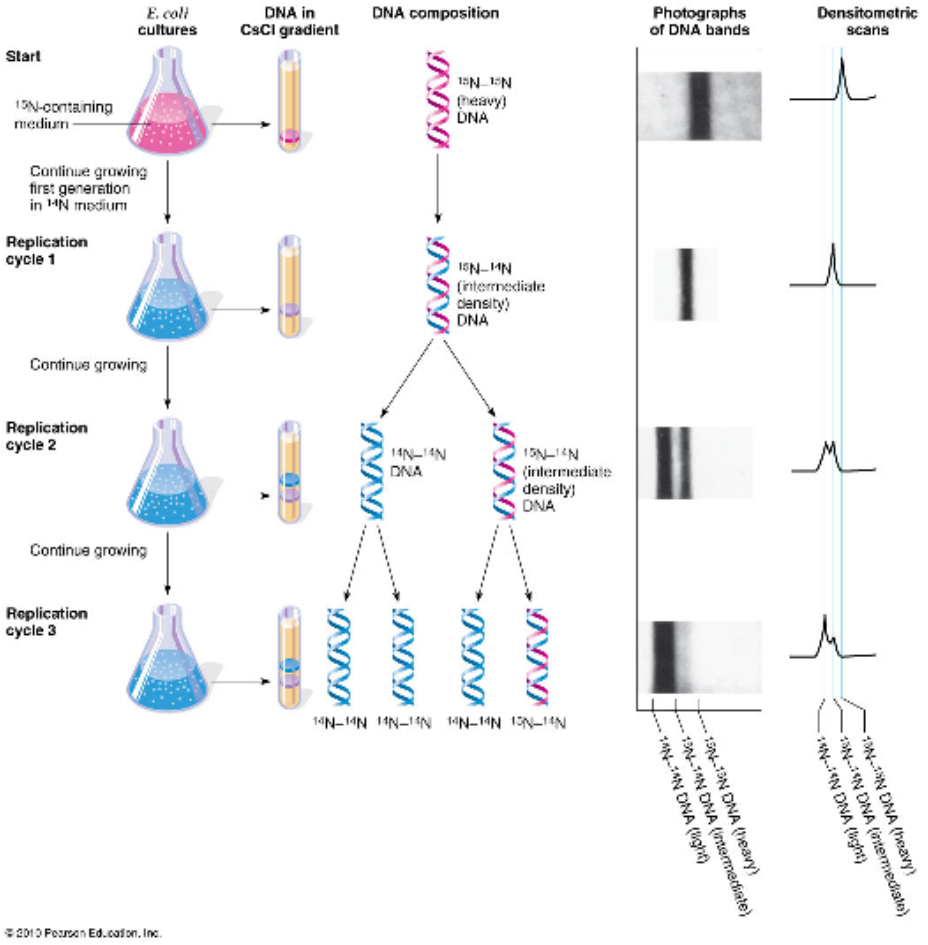
- At the end of the replication, two daughter DNA molecules are formed.
- In each daughter DNA, one strand is parental and the other one is totally newly synthesized. Thus, 50% is contributed by mother DNA.
- Hence, it is described as Semiconservative replication.

Experiment to confirm the Semiconservative nature of DNA Replication

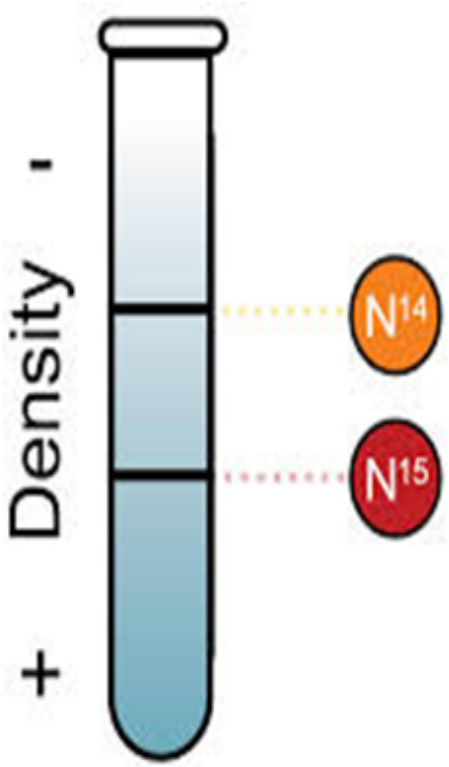
- In newly formed DNA molecule, one strand is old and other strand is newly synthesized. Thus, it is called Semiconservative mode of replication.
- It was experimentally proved by Matthew Meselson and Franklin Stahl(1958) by using **equilibrium – density – gradient – centrifugation** technique.



1. **Meselson and Stahl cultured bacteria *E.coli* in the medium containing ^{14}N(light nitrogen) and obtained equilibrium density gradient band by using 6M $CsCl_2$. The position of this band is recorded.**

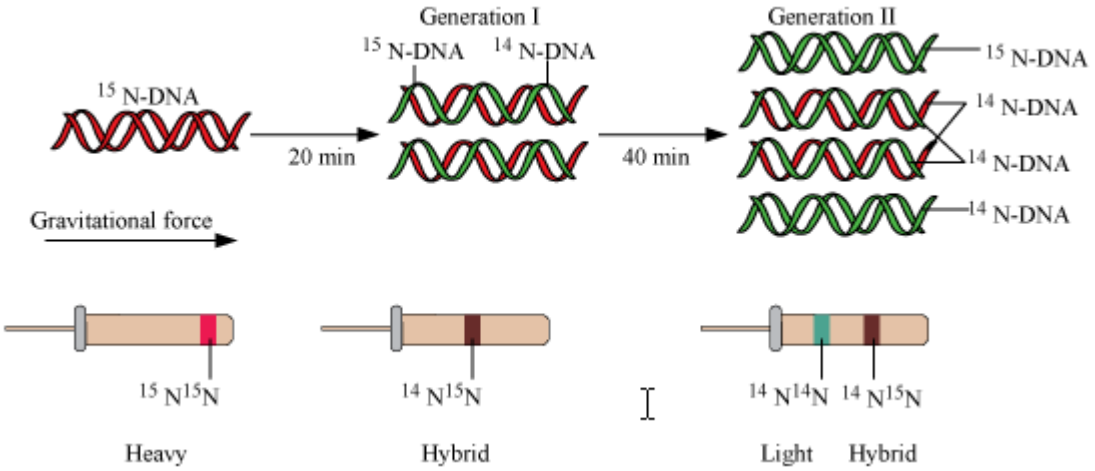


2. *E.coli* cells were then transferred to ^{15}N medium (heavy isotopic nitrogen) and allowed to replicate for several generations. At equilibrium point density gradient band was obtained, by using 6M CsCl2. The position of the band is recorded.



3. The heavy DNA (^{15}N) molecule can be distinguished from normal DNA by centrifugation in a 6M Cesium chloride ($CsCl_2$) density gradient. The density gradient value of 6M $CsCl_2$ and ^{15}N DNA is almost same. Therefore, at the equilibrium point ^{15}N DNA will form a band. In this both the strands of DNA are labelled with ^{15}N.

4. Such *E.coli* cells were then transformed to another medium containing ^{14}N i.e normal(light) nitrogen. After first generation, the density gradient band for $^{14}N\,^{15}N$ was obtained and its position was recorded. After second generation, two density gradient bands were obtained – one at $^{14}N\,^{15}N$ position and other at ^{14}N position.
5. The position of bands after two generations clearly proved that DNA replication is Semiconservative.

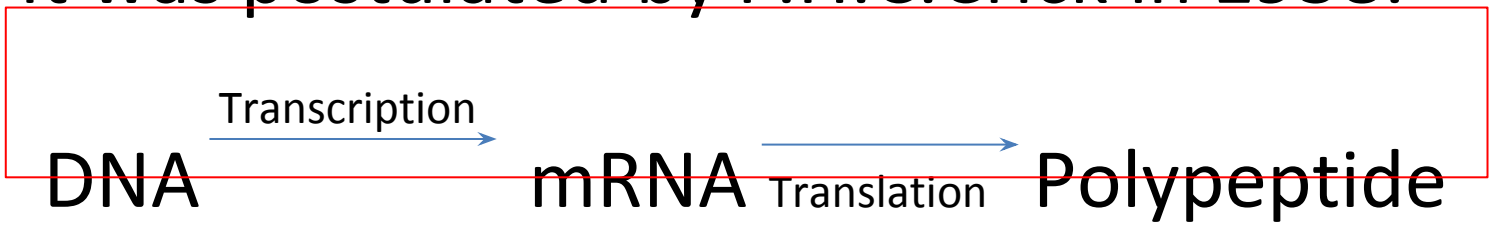


PROTEIN SYNTHESIS

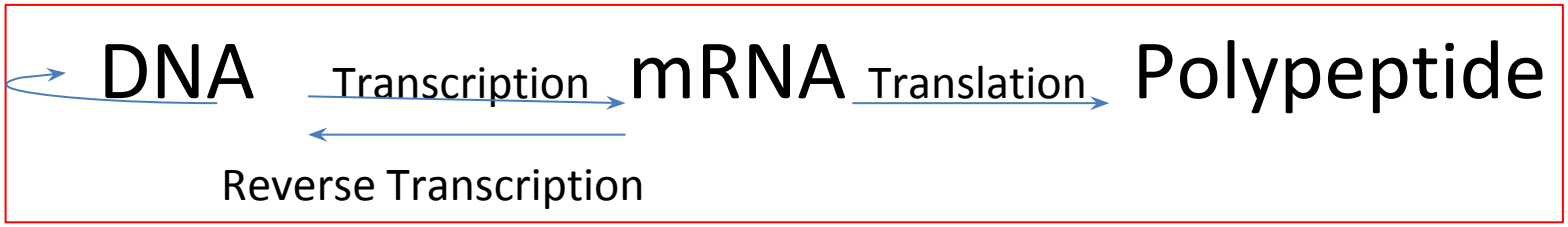
- Proteins are very important biomolecules.
- They serve as structural components, enzymes and hormones.
- The cell needs to synthesize new protein molecules.
- The process of protein synthesis includes **Trancription** and **Translation**.
- The process of copying genetic information from one strand of DNA into a single stranded RNA, is termed as **Transcription** and the process of formation of protein from RNA is called **Translation**.

CENTRAL DOGMA OF PROTEIN SYNTHESIS

- Double stranded DNA molecule gives rise to mRNA which acts as a messenger to programme the synthesis of a polypeptide chain (protein).
- This type of unidirectional flow of information from DNA to RNA to protein/proteins is referred as central dogma of protein synthesis.
- It was postulated by F.H.C.Crick in 1958.



The present concept of central dogma in retroviruses or riboviruses is given by Temin (1970) and Baltimore(1970)



Accordingly enzyme RNA dependent DNA polymerase, synthesizes DNA from RNA.

TRANSCRIPTION

- During Transcription, information of only one strand of DNA is copied into RNA with the help of RNA polymerase enzyme.
- Transcription takes place in nucleus in Eukaryotes whereas Translation occurs in cytoplasm.
- DNA transfers information to mRNA which moves to ribosomes.
- Transcription occurs in the nucleus during G_1 and G_2 phases of cell cycle.
- DNA has promotor and terminator sites. Transcription starts at promotor site and stops at terminator site.
- The process of Transcription in both Prokaryotes and Eukaryotes, involves three stages viz. Initiation, Elongation and Termination.

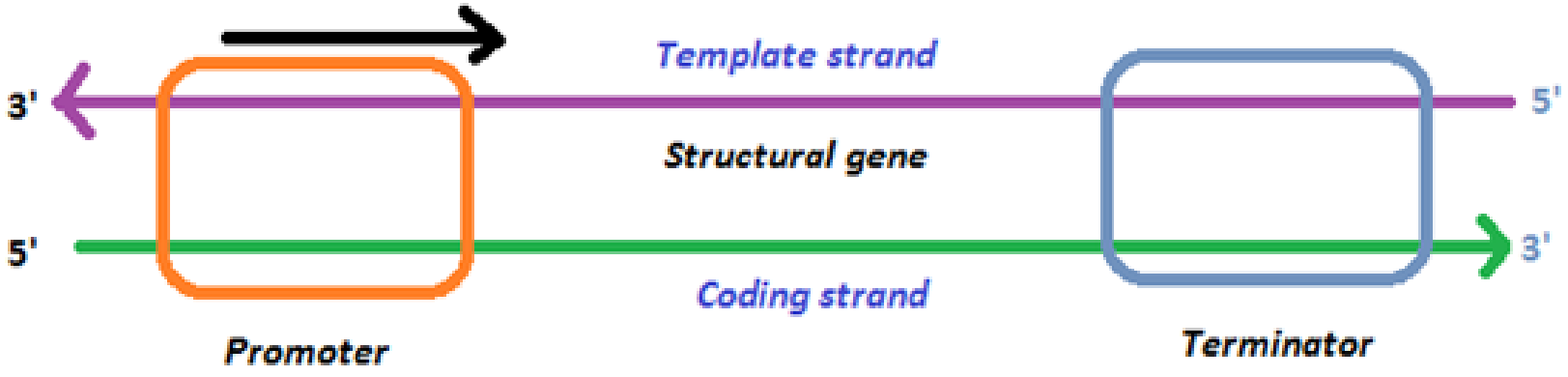
Transcription Unit

- Each transcribed segment of DNA is called transcription unit.
- It consists of (i)Promotor (ii) The structural gene (iii) Terminator.

i. **Promotor** :- It is a DNA sequence that provides binding site for enzyme RNA polymerase. RNA polymerase binds to specific Promotor. It is present towards 5' end i.e upstream. In prokaryotes, the enzyme recognizes the promotor by its sigma factor sub unit.

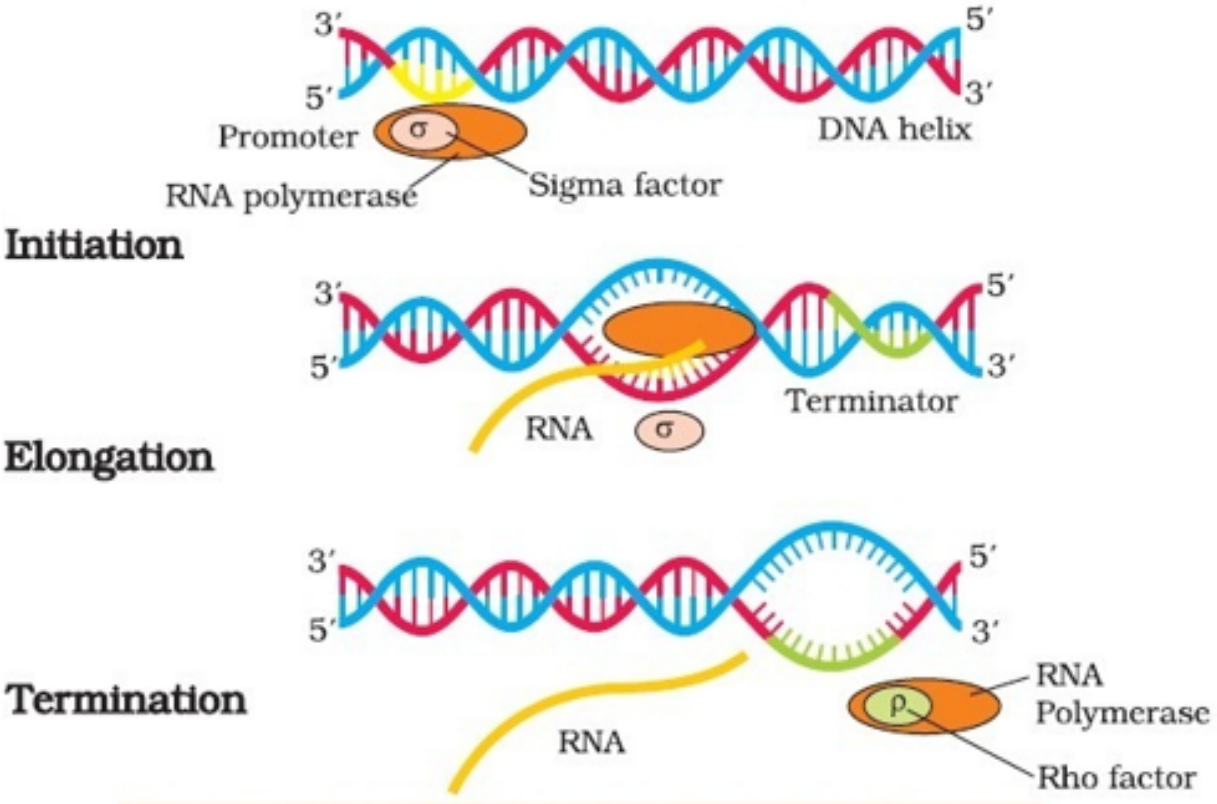
ii. Structural genes:- Structural genes are the segments of DNA, which carry codes for the synthesis of proteins. The DNA strand which is used for synthesis of RNA is called **Template strand** which is oriented in 3'⟶5' direction. It is also called **Antisense stand** while the other strand, not involved in RNA synthesis is called **Coding strand**. It is oriented in 5'⟶3'direction. It is also called **Sense strand.**

iii. **Terminator** :- It is a small DNA sequence which terminates the transcription process. It is present towards 3' end/downstream.

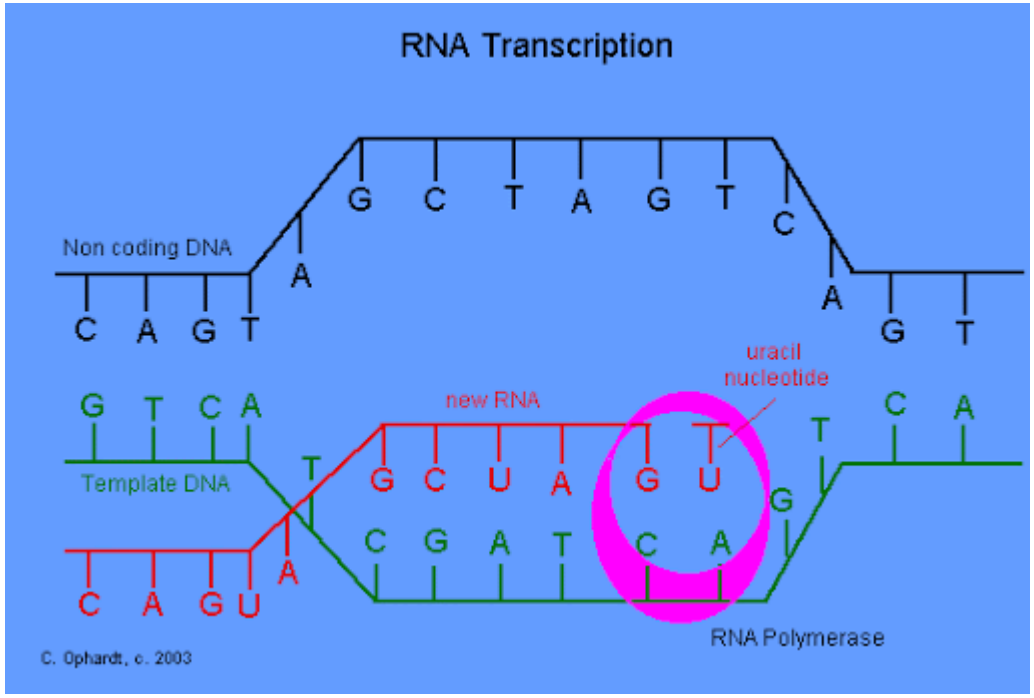


Transcription unit

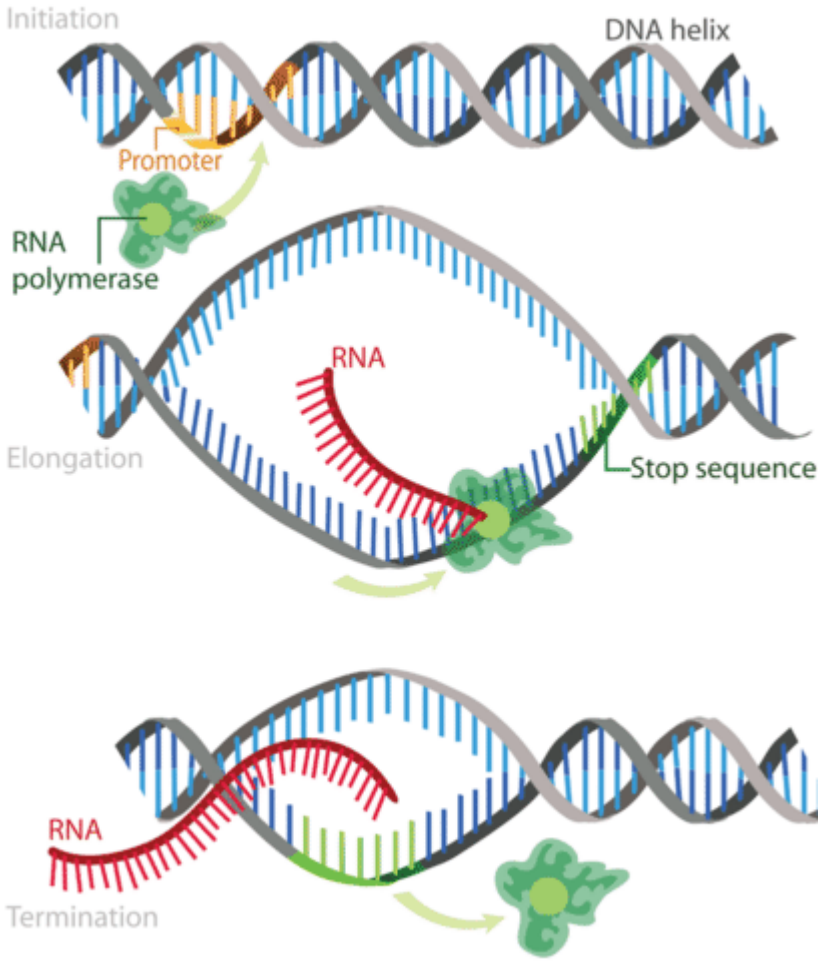
During transcription the enzyme RNA polymerase binds to the promotor site and brings about the **initiation** of the process.



The two strands of DNA separate from each other and according to the base sequence present on the template strand, the complementary RNA nucleotides are selected and joined one after the other to form the mRNA strand (**elongation**). A small part of RNA remains attached to the enzyme.



- As the enzyme reaches to the terminator region, both the enzyme and newly constructed RNA falls off. This is the **termination** of transcription.
- As the mRNA grows, the transcribed region of DNA molecule becomes spirally coiled and attains(regains) double helical form.



In bacteria, mRNA does not require any processing because it has no introns. Prokaryotes posses only one type of RNA polymerase. In Eukaryotes, there are three types of RNA polymerases i.e

RNA polymerase-I transcribes rRNA

RNA polymerase-II transcribes mRNA and hnRNA (heterogeneous nuclear RNA)

RNA polymerase-III transcribes tRNA and snRNA(small nuclear RNA)

Transcription unit and the gene

- The DNA sequence coding for mRNA/tRNA or rRNA is defined as a gene.
- Cistron is a segment of DNA coding for a polypeptide.
- A single structural gene in transcription unit is said to be **monocistronic** where as a long segment of DNA having set of various structural genes in one transcription unit is referred as **polycistronic**.
- Structural genes in eukaryotes have interrupted non-coding sequences called **introns**.
- The coding sequences or express- sequences are defined as **exons**. Only exons appear in processed mRNA in Eukaryotes.

Processing of hnRNA

- In eukaryotes, forms of RNA transcribed from DNA are called primary transcripts.
- Such transcripts undergo changes called processing or maturation before becoming functional.
- Primary transcript if non functional and contains both exons and introns.
- During processing only introns are removed by the process called **splicing**.

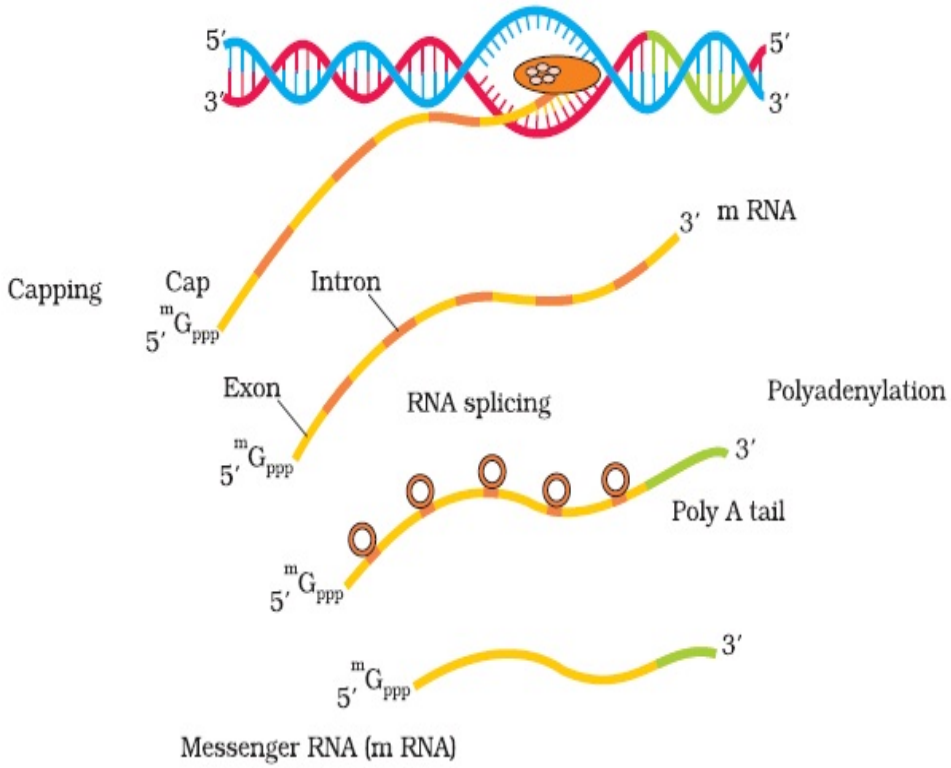


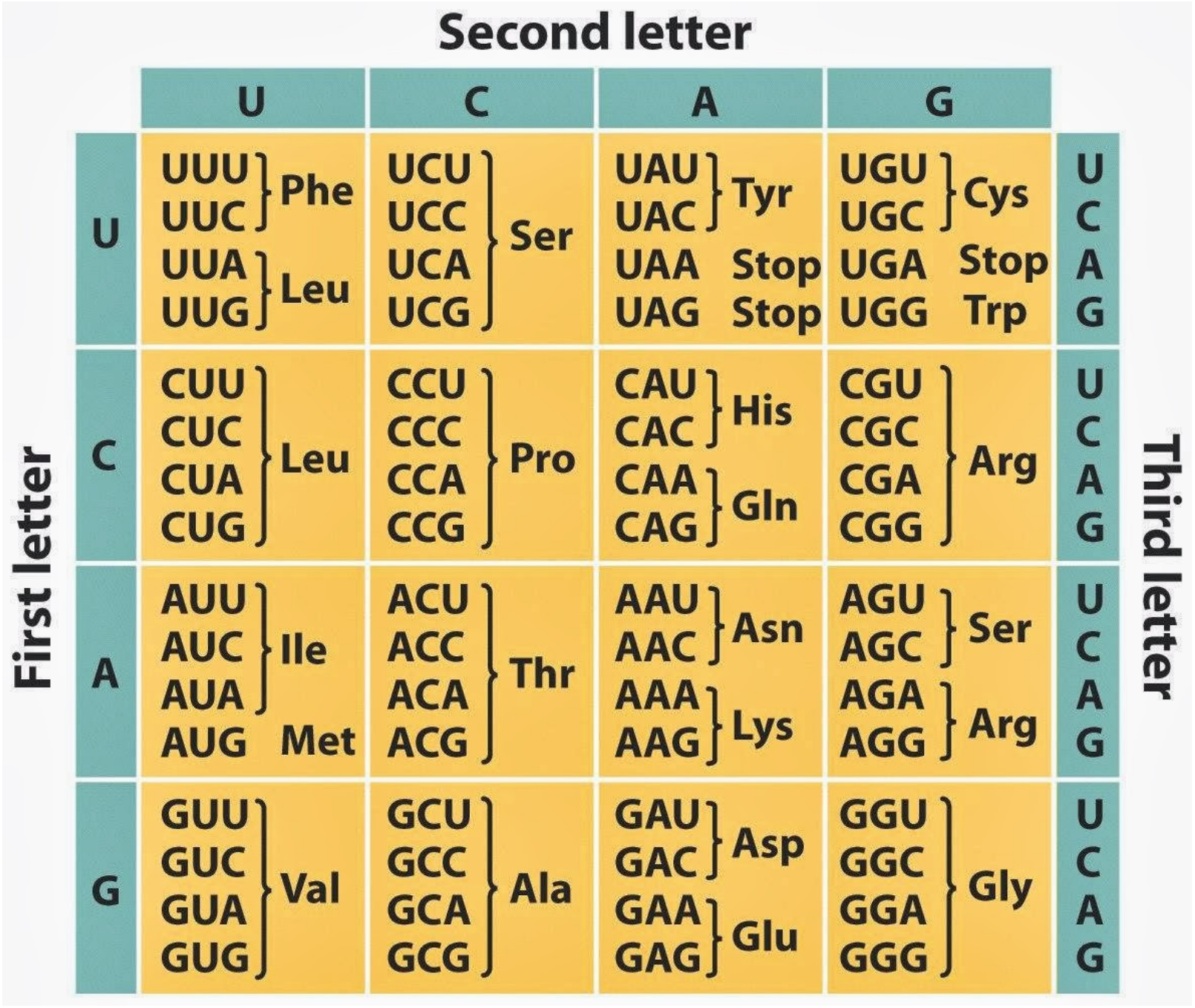
Figure 6.11 Process of Transcription in Eukaryotes

- Exons are joined in a definite sequence(order) by DNA ligase enzyme.
- Heterogeneous nuclear RNA, undergoes the process of capping and tailing.
- In **capping**, methylated guanosine triphosphate is added to 5' end of hnRNA.
- In **tailing**, polyadenylation take place at 3' end.
- It is the fully processed hnRNA, now called mRNA.

GENETIC CODE

- DNA is a master molecule of a cell that initiates, guides, regulates and controls the process of protein synthesis.
- The information for the synthesis of proteins is located in the DNA itself.
- About, 20 different types of amino acids are involved in the process of synthesis of proteins.
- DNA molecule has 4 types of nitrogen bases to identify these 20 different type of amino acids.
- But how is it possible that 20 types of amino acids are encoded by 4 types of nitrogen bases ?

- According to Crick, this information is stored in the form of coded language (Cryptogram) called genetic code, that contains code words (Codons) each one specifying (representing) specific amino acids.
- If each codon has only one nitrogen base then there will be $4^1 = 4$ codons are possible (not enough to code 20 different amino acids).
- If each codon has two nitrogen bases then there will be $4^2 = 4*4=16$ codons are possible (not enough to code 20 different amino acids).
- If each codon has three nitrogen bases then there will be $4^3 = 4*4*4=64$ codons are possible (more than enough to code 20 different amino acids).
- Hence G. Gamov (1954) suggested that in a codon, there must be combination of three consecutive nitrogen bases that will be sufficient to specify 20 different types of amino acids.


Second letter
U
C
A
G
First letter
Third letter
U
UUU Phe
UUC
UUA Leu
UUG
UCU Ser
UCC
UCA
UCG
UAU Tyr
UAC
UAA Stop
UAG Stop
UGU Cys
UGC
UGA Stop
UGG Trp
C
CUU Leu
CUC
CUA
CUG
CCU Pro
CCC
CCA
CCG
CAU His
CAC
CAA Gln
CAG
CGU Arg
CGC
CGA
CGG
A
AUU Ile
AUC
AUA
AUG Met
ACU Thr
ACC
ACA
ACG
AAU Asn
AAC
AAA Lys
AAG
AGU Ser
AGC
AGA Arg
AGG
G
GUU Val
GUC
GUA
GUG
GCU Ala
GCC
GCA
GCG
GAU Asp
GAC
GAA Glu
GAG
GGU Gly
GGC
GGA
GGG

Characteristics of Genetic code

i. **Genetic code is a triplet code :**

Sequence of three consecutive bases constitute codon, which specifies one particular amino acid. In every living organism genetic code is a triplet code.

ii. **Genetic code has distinct polarity :**

Genetic code shows definite polarity i.e. direction. It, therefore, is always read in 5' → 3' direction and not in 3' → 5' direction. Otherwise message will change.

e.g. 5' AUG 3'

iii. Genetic code is non-overlapping :

code is non overlapping i.e. each single base is a part of only one codon. Adjacent codons do not overlap. If non-overlapping, then with 6 consecutive bases only two amino acid molecules will be in the chain. Had it been overlapping type, with 6 bases, there would be 4

amino acid molecules in a chain.

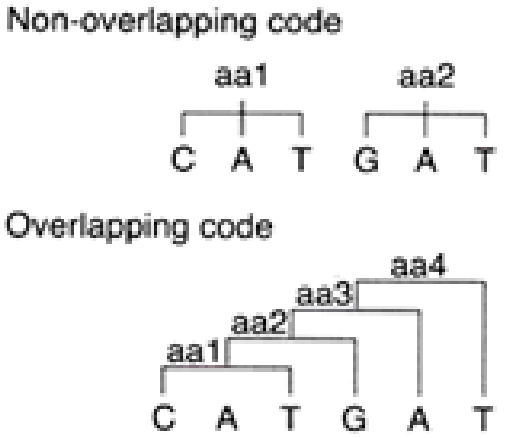
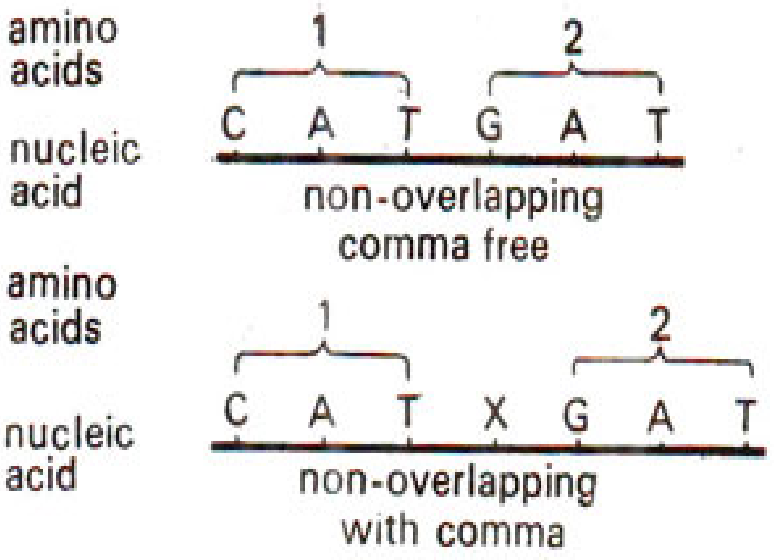
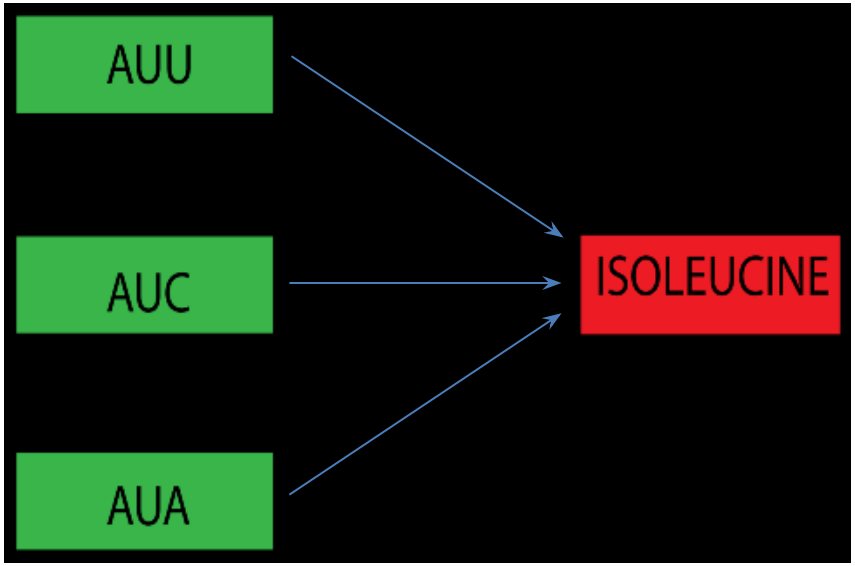


Fig. 15.3: Non-overlapping and overlapping code

iv. **Genetic code is commaless :-** there is no gap or punctuation mark between successive /consecutive codons.

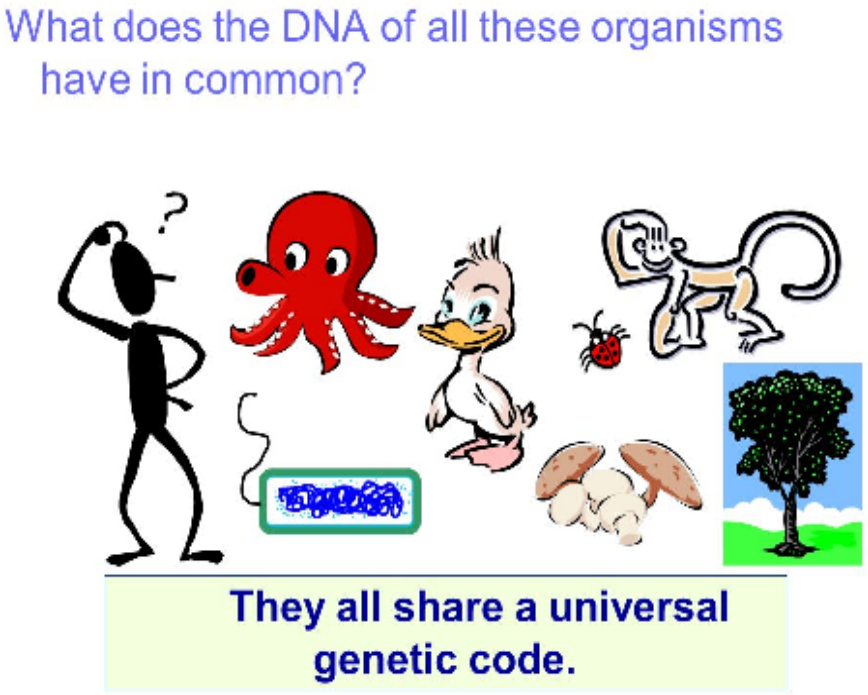
V. **Genetic code has degeneracy** :- usually single amino acid is encoded by single codon. However, some amino acids are encoded by more than one codons. E.g. Cysteine has two codons, while Isoleucine has three codons. This is called degeneracy of the code.





vi. **Genetic code is universal** :- By and large in all living organisms the specific codon specifies same amino acid. E.g. codon AUG always specifies amino acid methionine in all organisms from bacteria up to humans.

vii. **Genetic code is non-ambiguous** :- specific amino acid is encoded by a particular codon. Alternatively, two different amino acids will never be encoded by the same codon.



Non-Ambiguous

- The genetic code is non-ambiguous.
- Thus one codon cannot specify more than one amino acid.

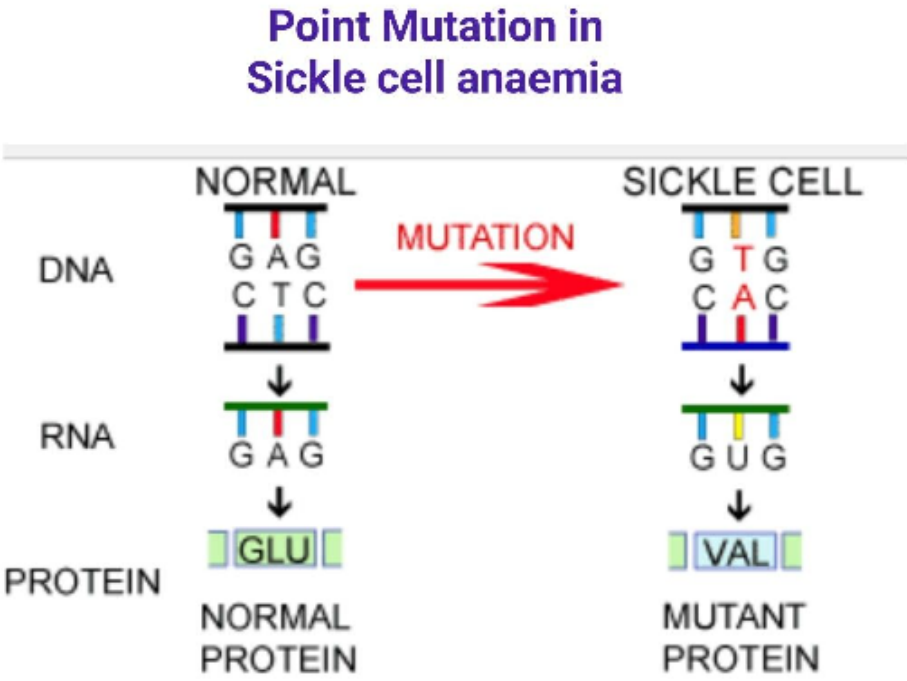
First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU Phenylalanine	UCU Serine	UAU Tyrosine	UGU Cysteine	U
	UUC Phenylalanine	UCC Serine	UAC Tyrosine	UGC Cysteine	C
	UUA Leucine	UCA Serine	UAA Stop codon	UGA Stop codon	A
	UUG Leucine	UCG Serine	UAG Stop codon	UGG Tryptophan	G
C	CUU Leucine	CCU Proline	CAU Histidine	CGU Arginine	U
	CUC Leucine	CCC Proline	CAC Histidine	CGC Arginine	C
	CUA Leucine	CCA Proline	CAA Glutamine	CGA Arginine	A
	CUG Leucine	CCG Proline	CAG Glutamine	CGG Arginine	G
A	AUU Isoleucine	ACU Threonine	AAU Asparagine	AGU Serine	U
	AUC Isoleucine	ACC Threonine	AAC Asparagine	AGC Serine	C
	AUA Isoleucine	ACA Threonine	AAA Lysine	AGA Arginine	A
	AUG Methionine; start codon	ACG Threonine	AAG Lysine	AGG Arginine	G
G	GUU Valine	GCU Alanine	GAU Aspartic acid	GGU Glycine	U
	GUC Valine	GCC Alanine	GAC Aspartic acid	GGC Glycine	C
	GUA Valine	GCA Alanine	GAA Glutamic acid	GGA Glycine	A
	GUG Valine	GCG Alanine	GAG Glutamic acid	GGG Glycine	G

viii. **Initiation codon and Termination codon** :-
AUG is always an initiation codon in any and every mRNA. AUG codes for amino acid methionine. Out of 64 codons, three codons viz. UAA, UAG and UGA are termination codons which terminate/stop the process of elongation of polypeptide chain, as they do not code for any amino acid.

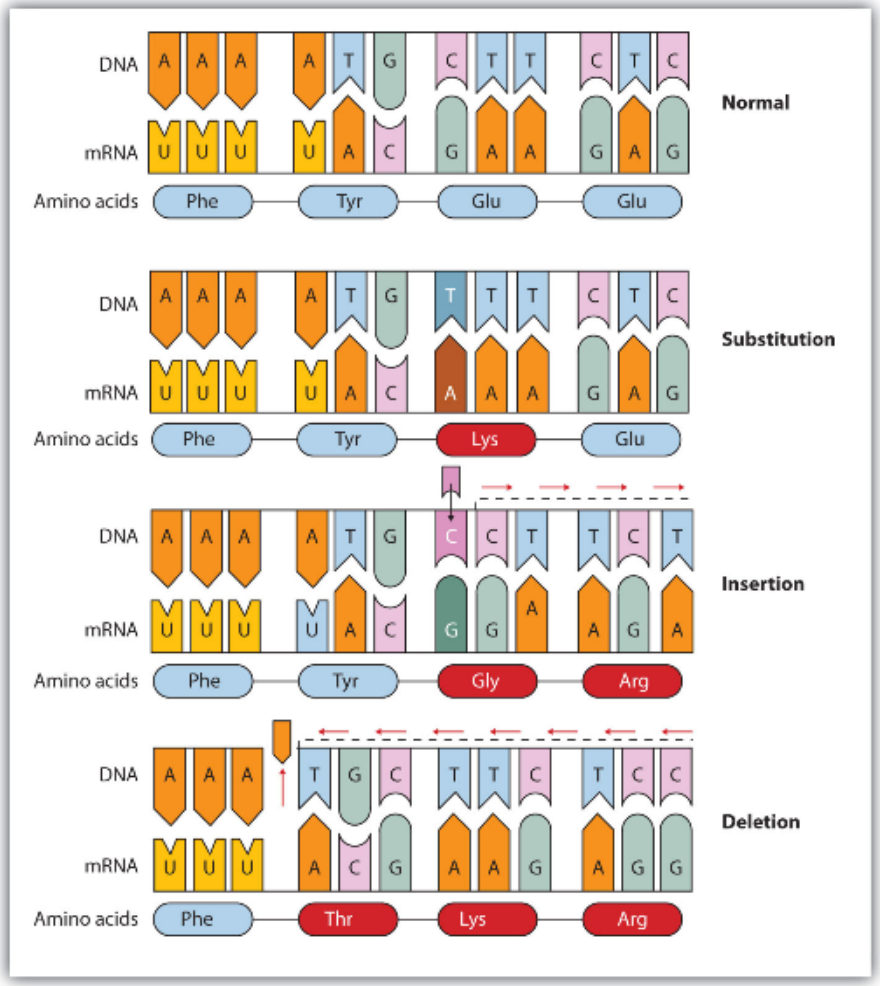
ix. **Codon and Anticodon** :- codon is a part of DNA e.g. AUG is codon. It is always represented as 5' AUG 3'. Anticodon is a part of tRNA. It is always represented as 3' UAC 5'.

Mutations and Genetic Code

- Mutation is a phenomenon in which sudden change in the DNA sequence takes place.
- It results in the change of genotype (i.e. character).
- Along with recombination, mutation is a raw material for evolution as it results in variations.
- During mutation, possibility of loss (deletion) or gain (insertion/ duplication) of a segment of DNA results in alteration in the chromosomes.
- Mutation can also occur due to change in a single base pair of DNA. This is known as point mutation. Eg. Sickle cell anaemia.

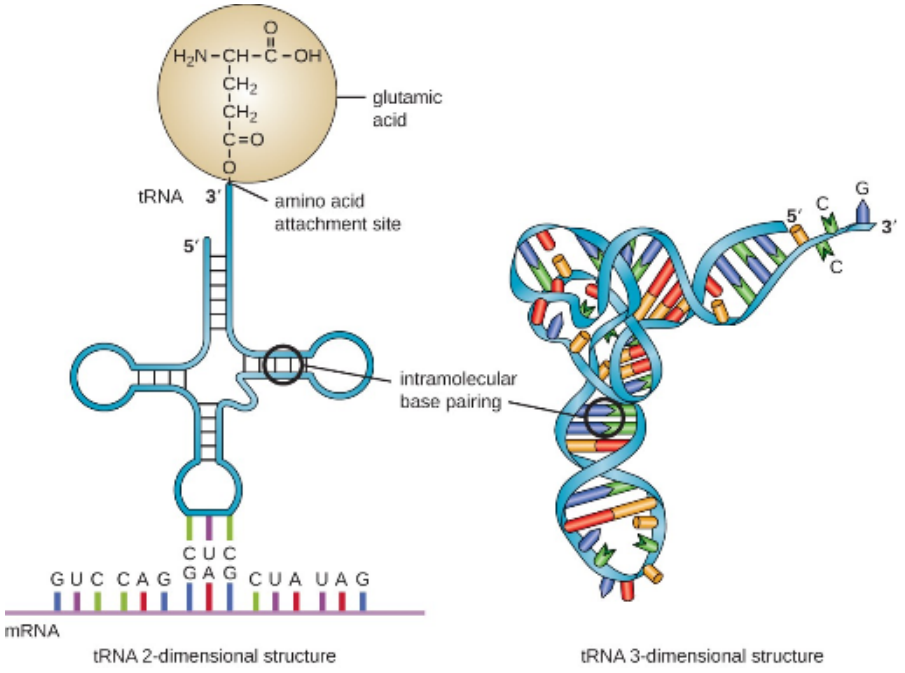


- Deletion or insertion of base pairs of DNA causes frame – shift mutations or deletion mutation.
- Insertion or deletion of one or two bases changes the reading frame from the point of insertion or deletion.
- Insertion or deletion of three or multiples of three bases (insert or delete) results in insertion or deletion of amino acids and reading frame remains unaltered from that point onwards.



t-RNA- the adapter molecule

- Scientists considered that there has to be a mechanism in which tRNA will read the codon and also simultaneously binds the amino acid as amino acid does not have any special capacity to read the codon. So tRNA is considered as an adapter molecule.
- Cloverleaf structure of tRNA possess an anticodon loop that has bases complementary to the codon. It is called anticodon.
- It shows amino acid acceptor end (3' end) having unpaired CCA bases to which amino acid binds.
- For every amino acid, there is specific tRNA. Initiator tRNA is specific to methionine. There is no tRNA's for stop codons.
- In the actual structure, the tRNA molecule looks like inverted L (3 dimensional structure)



tRNA 2-dimensional structure

tRNA 3-dimensional structure

Translation – protein synthesis

- **Translation** :- It is the process in which the sequence of codons on the mRNA strand is read/ decoded and accordingly the amino acids are joined to each other to form a polypeptide chain that makes protein.

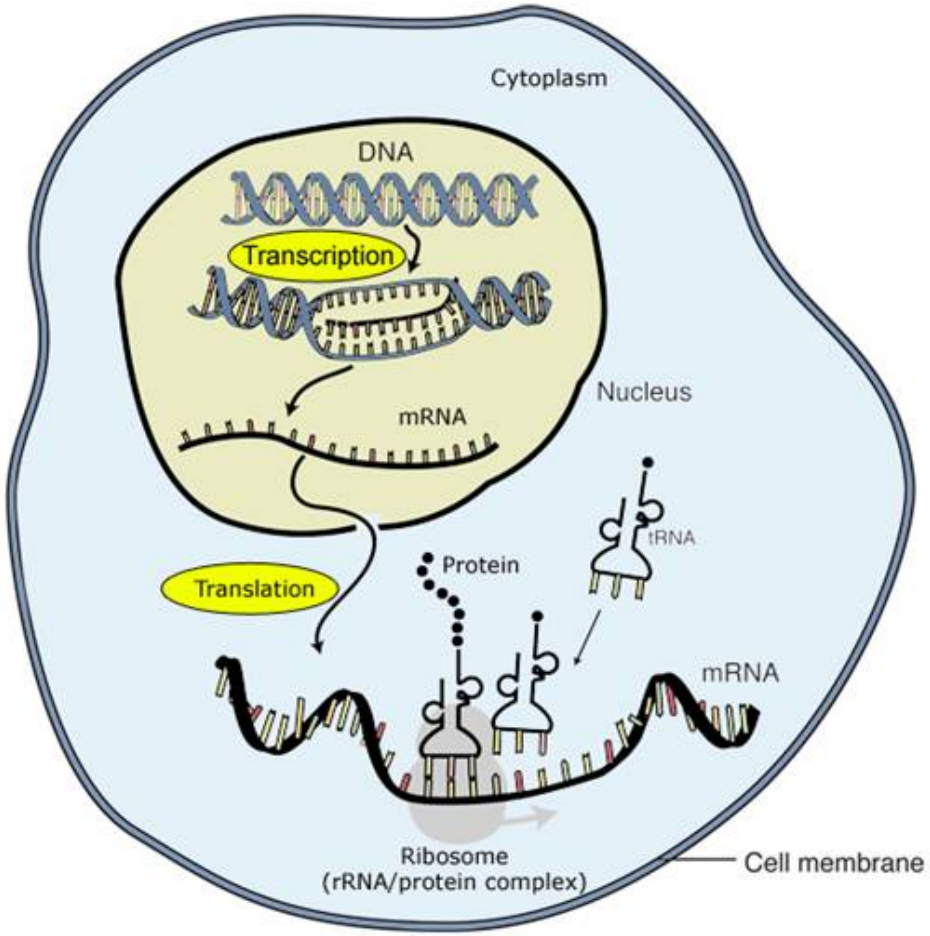


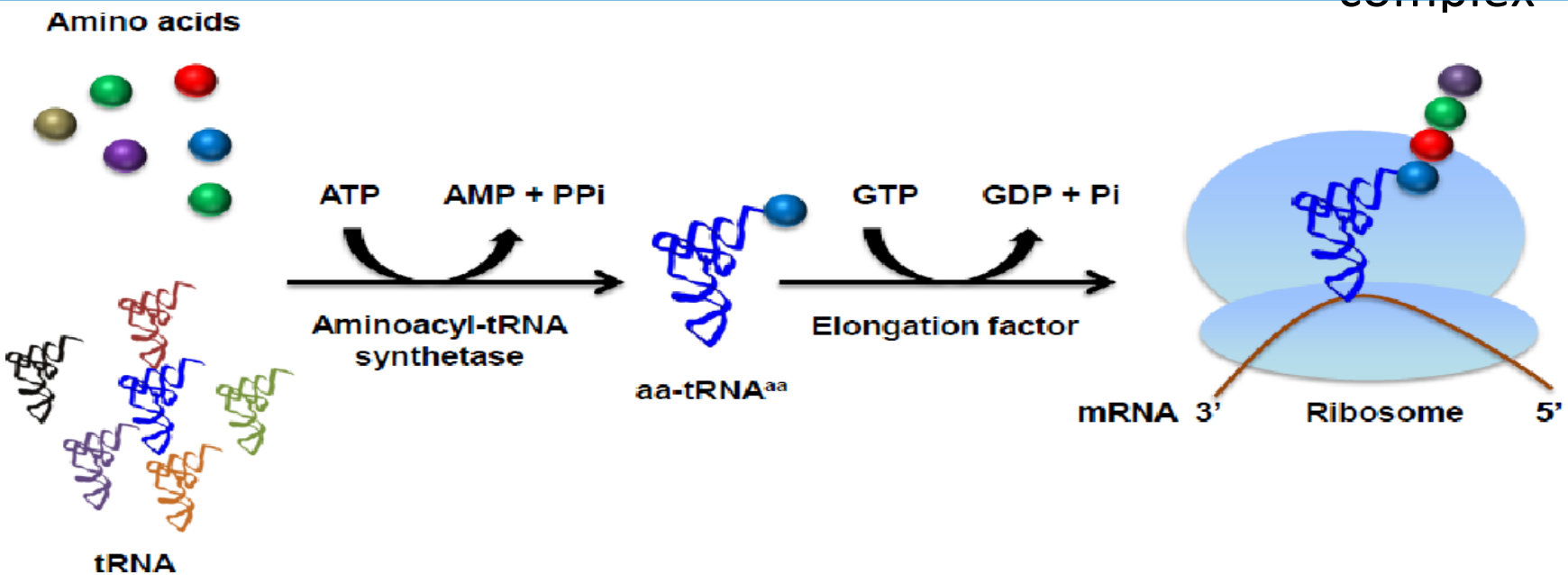
Image adapted from: National Human Genome Research Institute.

Process of Translation involves following steps

(A) Activation of amino acids and formation of AA- tRNA complex :-

In presence of an enzyme aminoacyl tRNA synthetase the amino acid (AA) molecule is activated and then each amino acid is attached to specific tRNA molecule at 3'/CCA end to form aminoacyl-tRNA complex. The reaction needs ATP. This process is called charging of tRNA or aminoacylation of tRNA.

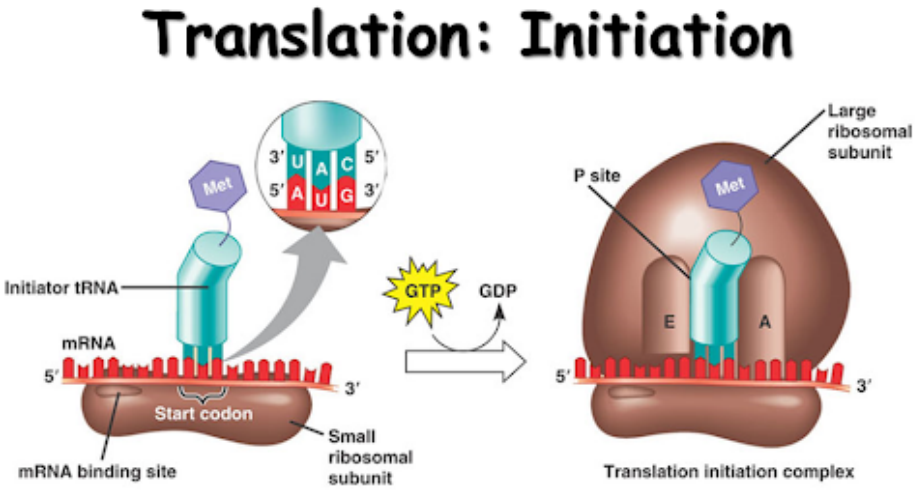
Amino acid(AA)+ATP+tRNA $\xrightarrow{\text{Aminoacyl tRNA synthetase}}$ aminoacyl-tRNA complex



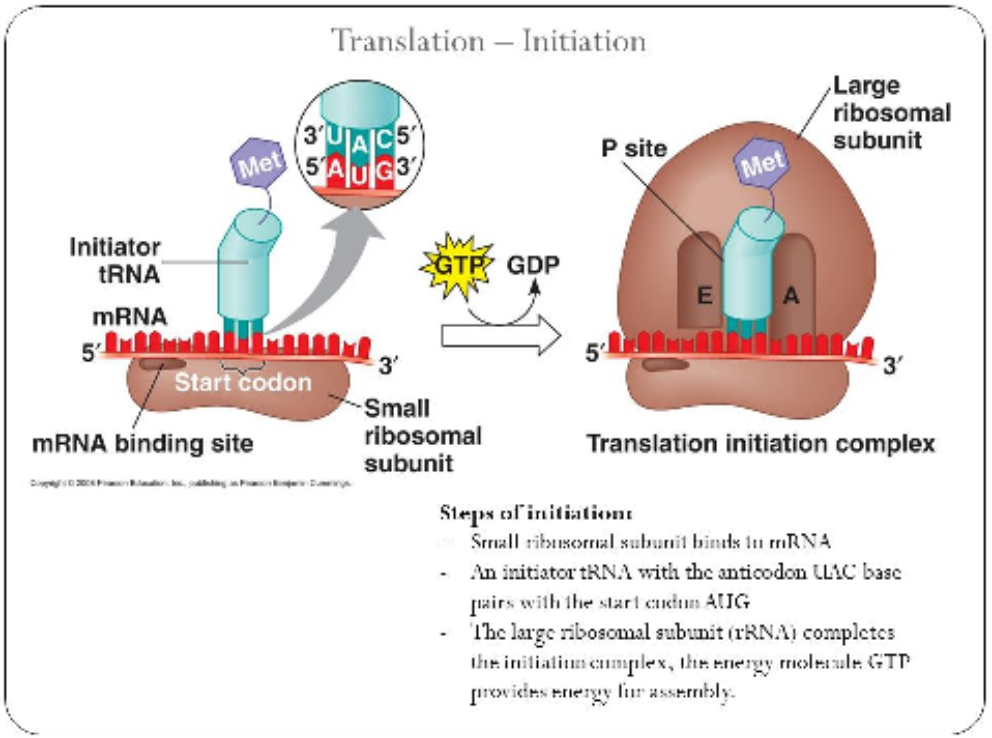
(B) **Formation of the polypeptide chain** :- it is the actual translation process which involves following steps:

i. Initiation :-

- It begins with the formation of initiation complex which requires the mRNA having codons for a polypeptide, the smaller (30s) and larger (50s) sub-units of ribosomes, the initial AA_1 -tRNA complex and ATP and GTP as source of energy.
- The process starts with the binding of mRNA on the smaller 30s sub-unit of ribosome.
- AUG is present on mRNA which initiates the process of protein synthesis.
- Initiator tRNA binds with initiation codon (AUG) by its anticodon (UAC) through hydrogen bonds. It carries activated amino acid methionine (*met*) in Eukaryotes or formyl methionine (*f-met*) in prokaryotes.

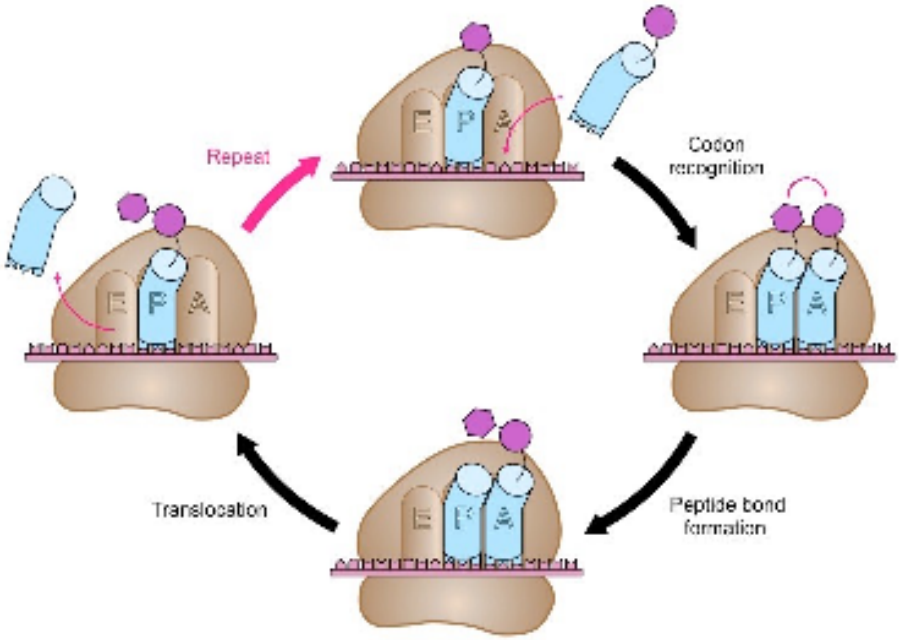


- Now the large subunit of ribosome joins with the smaller subunits, that requires Mg^{++} ions.
- The ribosome has three sites namely : aminoacyl (A) site, peptidyl site (P) and exit (E) site.
- The empty tRNA leaves from E site. Only the AA_1-tRNA complex binds at P site directly while all the other incoming tRNA complexes get attached first at A site and then are shifted to P site.



ii) **Elongation** :-

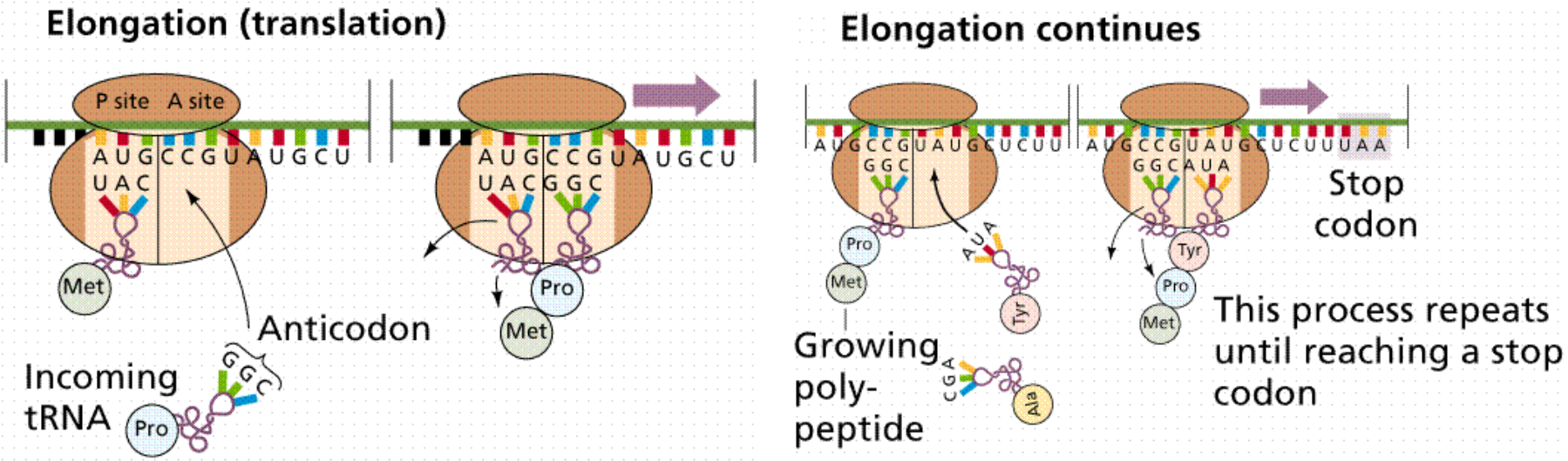
- During this process, activated amino acids are added one by one to first amino acid (methionine). This amino acid binds with amino acid binding site of tRNA. This result in formation of tRNA amino acid complex.



Addition of amino acid occurs in 3 steps cycle-

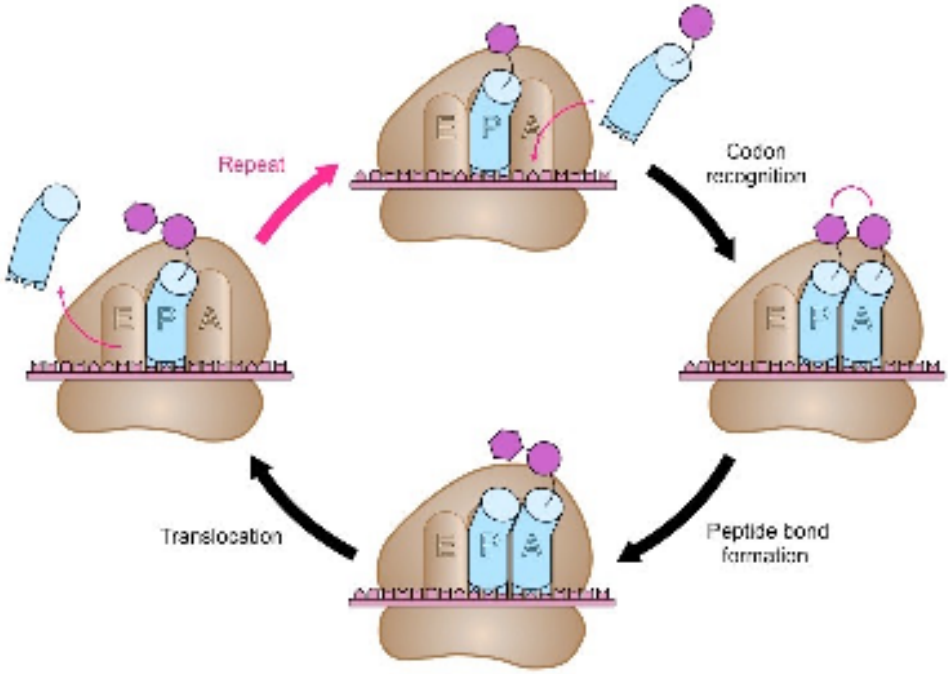
a) **<u>Codon recognition</u>**:-Amino acyl tRNA molecule enters the ribosome at A-site. Anticodon binds with the codon by hydrogen bonds.

b) Amino acid on the first initiator tRNA at P-site and amino acid on tRNA at A-site join by peptide bond. At this time first tRNA at 'P' site is kicked off.



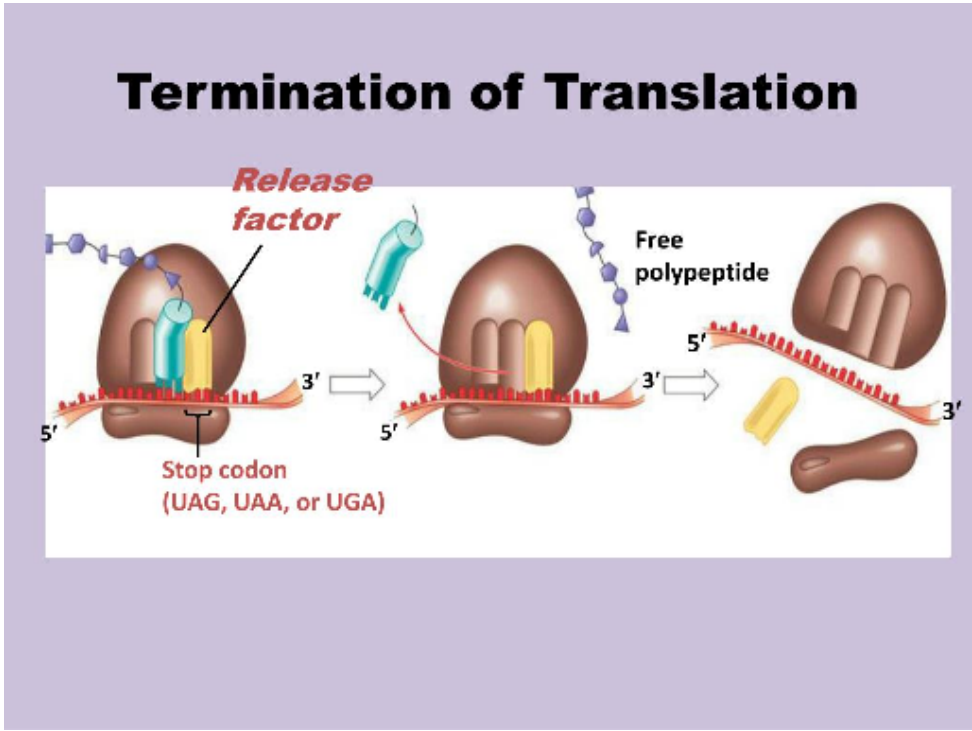
c) <u>Translocation</u> :-The tRNA at A-site carrying a dipeptide at A-site moves to the P-site. This process is called **translocation**. In translocation, both the subunits of ribosome move along in relation to tRNA and mRNA. Hence, tRNA carrying dipeptide now gets positioned at 'P' site of ribosome, making 'A' site vacant. At this site, then next charged tRNA molecule carrying amino acid will be received. During this process, first uncharged tRNA is discharged from E-site.

This process is repeated as amino acids are added to polypeptide. It takes less than 0.1 second for formation of peptide bond.



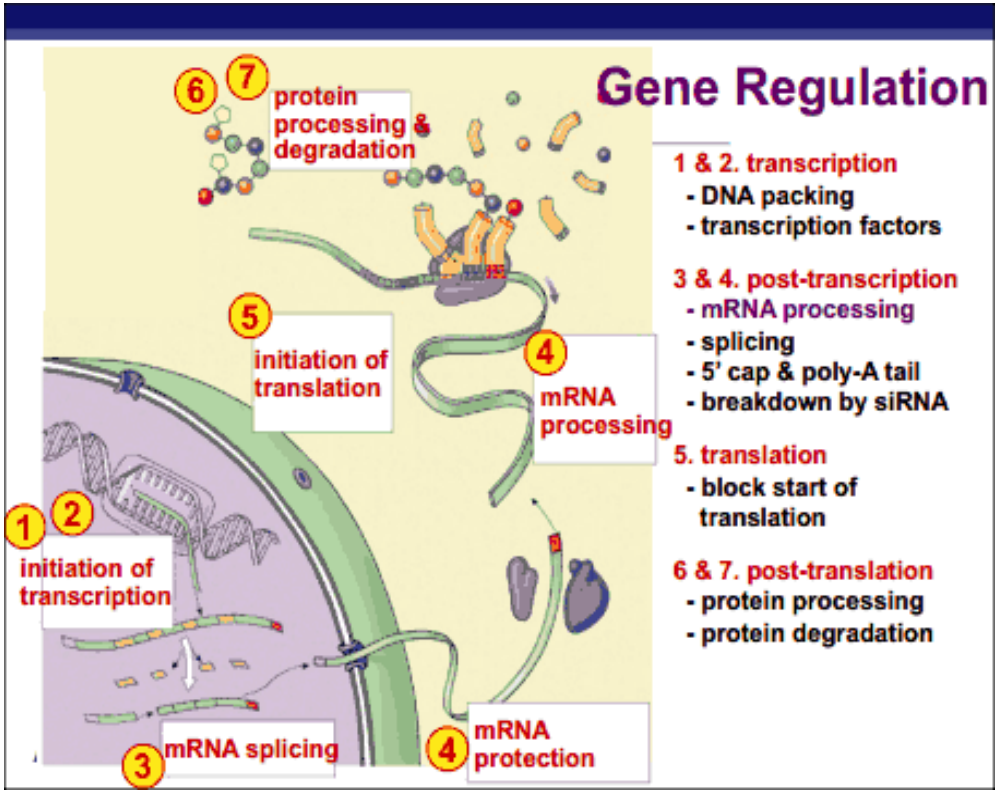
iii) Termination and release of polypeptide :-

- At the end of mRNA, there is a stop codon (UAA/UAG/UGA). It is exposed at the A-site. It is not read and joined by anticodon of any tRNA.
- The release factor binds to the stop codon, thereby terminating the translation process. The polypeptide is now released in the cytoplasm.
- Two subunits of ribosome dissociate and last tRNA is set free in the cytoplasm.
- Finally mRNA is also released in the cytoplasm. It get denatured by nucleases immediately. Hence mRNA is short lived.



Regulation of gene expression

- It is a multistep process by which a gene is regulated and its product is synthesized. Thus, gene expression results in the formation of a Polypeptide.
- Genes of a cell are expressed to perform different functions. For eg. An enzyme β-galactosidase is synthesized by E.coli. It is used for hydrolysis of lactose into glucose and galactose.
- Certain bacteria like E.coli adapt to their chemical environment by synthesizing certain enzymes depending upon the substrate present. Such adaptive enzyme is called **inducible enzyme**. A set of genes will be switched on when there is necessity to metabolise a new substrate. This phenomenon is called **induction** and small molecule responsible for this, is known as **inducer**. It is positive control.



Structure of the operon

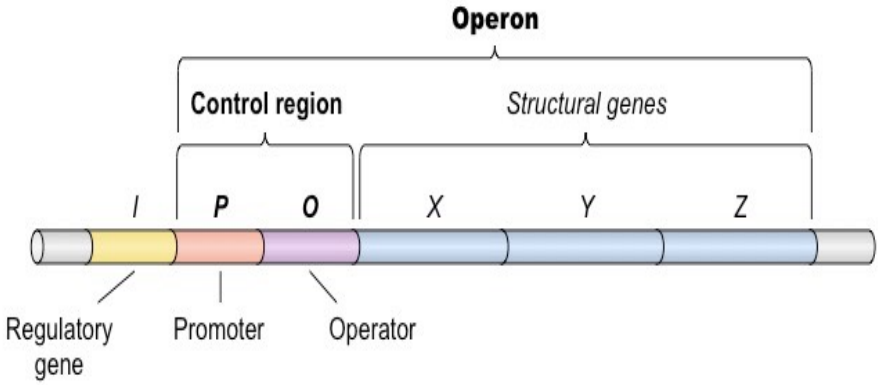
- The clusters of gene with related functions are called **operons**. They usually transcribe single mRNA molecule.
- In *E.coli* some 260 genes are grouped in 75 different operons.
- Each operon is a unit of gene expression and regulation which includes the **structural genes** and their **control elements** *(promoters and operators).*

i. The structural genes code for proteins, rRNA and tRNA required by the cell.

ii. Promoters are the signal sequences in DNA that start RNA synthesis. These are the sites where the *RNA polymerases* are bound during transcription.

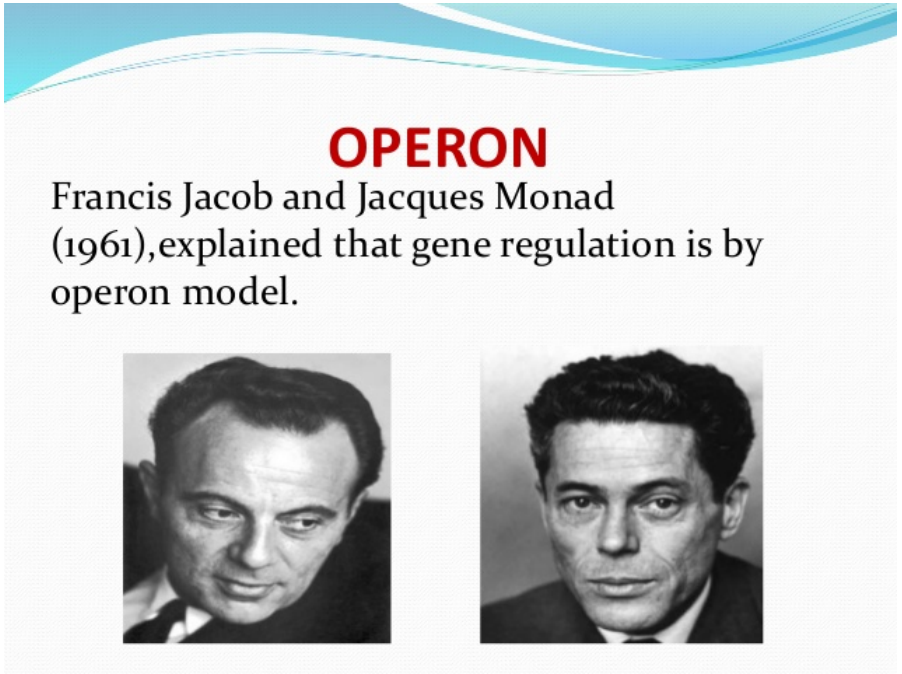
iii. The operators are present between the promoters and structural genes. The **repressor protein binds to the operator region of the operon**.

Regulatory genes are responsible for formation of repressors which interact with operators.

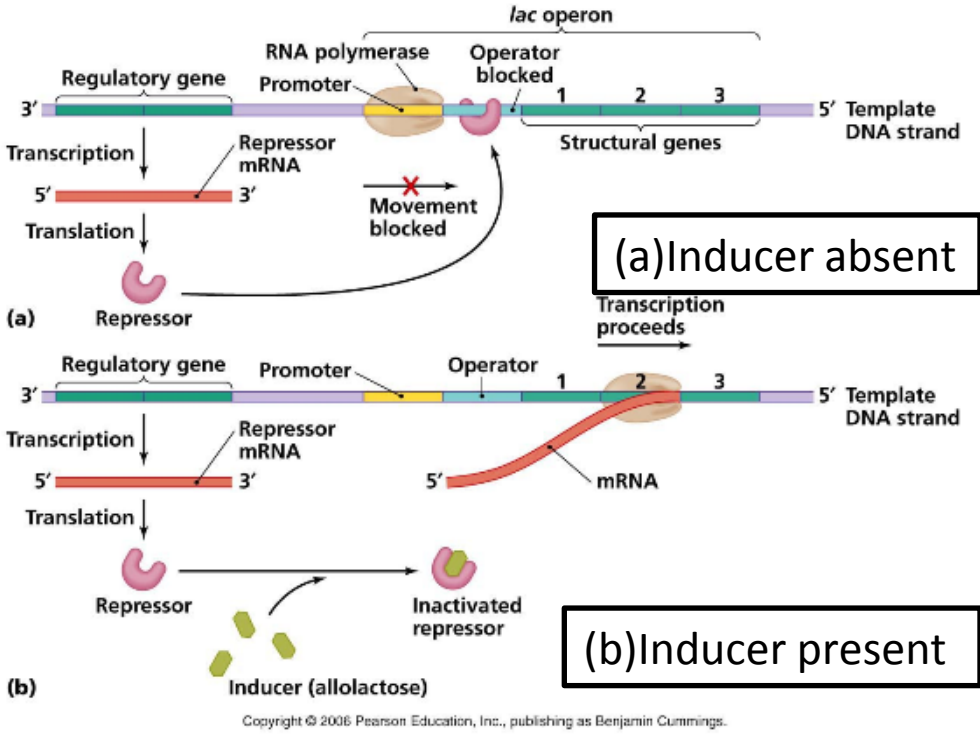


Lac operon

- Francois Jacob and J. Monod, proposed the classical model of Lac operon which can explain properly gene expression and regulation in *E. coli*.
- Lactose or Lac operon of *E.coli* is inducible operon. The operon is switched on when a chemical inducer – lactose is present in the medium.
- The metabolism of lactose in cells requires three enzymes- Permease, β-galactosidase and transacetylase. The enzyme permease is needed for entry of lactose in the cell, β-galactosidase brings about hydrolysis of lactose into glucose and galactose, while transacetylase transfers an acetyl group from acetyl Co-A to galactoside.



- The *Lac operon* has promoter sites (p), regulatory site (i) and operator site (o). Besides this it has three structural genes namely z, y and a. the 'z' gene codes for β-galactosidase, 'y' gene codes for permease and 'a' gene codes for transacetylase.
- In Lac operon a polycistronic structural gene is regulated by common promoter and regulatory genes.
- When the cell is using its normal energy source glucose, the '**i**' gene transcribes a repressor mRNA, after translation of which, a **repressor protein** is produced. It binds the operator region of the operon and prevents the RNA polymerase from transcribing the operon. As a result β-galactosidase is not produced.
- In absence of glucose, if lactose is available as energy source for the bacteria then following events occur in the cell. The lactose enters the cell as a result of the activity of permease enzyme. It act as inducer and interacts with the repressor to inactivate it.
- As the repressor is inactivated, the RNA polymerase can bind itself to the operator site and transcribe the operon to produce lac mRNA which enables formation of all the required three enzymes needed for lactose metabolism.
- this regulation of lac operon by the repressor is an example of negative control of transcription initiation.



Genomics

- The term **Genome** (introduced by H. Winker in 1920) is the total genetic constitution of an organism. Alternatively, it is a complete copy of genetic information (DNA) or one complete set of chromosomes (monoploid or haploid) of an organism.
- The term **Genomics** (term coined by T.H. Roderick in 1986) is the study of genomes through analysis, sequencing and mapping of genes along with the study of their functions.
- The sequencing of Yeast, Drosophila and mouse genome was done in order to facilitate comparative studies between humans and other organisms commonly used for genetic studies, in laboratory.

Genomics study may be classified into two types:

a. **Structural genomics:** It involves mapping,sequencing and analysis of genome.
b. **Functional genomics:** It deals with the study of functions of all gene sequences and their expression in organisms.

Application of genomics

Structural and functional genomics is used for different purposes in the improvement of crop plant, human health and livestock. The knowledge and understanding acquired from genomics research can be applied in a number of different sectors, including medicine, biotechnology and social sciences. It helps in the treatment of genetic disorders through gene therapy.

1. Genomics is used in agriculture to develop transgenic crops having more desirable characters.
2. Genetic markers developed in genomics, have applications in forensic analysis.
3. Genomics can lead to introduce new gene in microbes to produce enzymes, therapeutic proteins and even biofuels.

Human Genome Project

- The human genome project was initiated in 1990. This project was co-ordinated by the US department of Energy and National institute of health.
- The Human Genome Project began in 1990 and was completed in 2003.
- The Human Genome Project is a multinational research project to determine the genomic structure of humans.

The main aims of project are:-

I. Mapping the entire human genome at the level of nucleotide sequences.
II. To store the information collected from the project in databases.
III. To develop tools and techniques for analysis of the data.
IV. Transfer of the related technologies to the private sectors, such as industries.
V. Taking care of the legal, ethical and social issues which may arise from project.

- HGP (Human Genome Project) was closely associated with rapid development of a new area in biology, called **Bioinformatics**.
- The work of human genome project has allowed researches to begin to understood the blueprint in building and constructing the human genome.
- As researches learn more about the functions of genes and proteins, this knowledge will have a major impact in fields like Medicine, Biotechnology and the life sciences. Therefore HGP is important.
- Human Genome Project was to provide a complete and accurate sequence of the 3 billion DNA base pairs that make up the humane genome and to find out the estimated number of human genes. Now about 33000 genes have been estimated to be present in humans.
- The secret of our complexity may not lie not in the number of our genes but how we use them. It will lead to the understanding of gene structure and function in other species. Since we posses many of the genes same as flies, round worm and mice, such studies will lead t

Table 1. Comparative genome sizes of humans and other model organisms

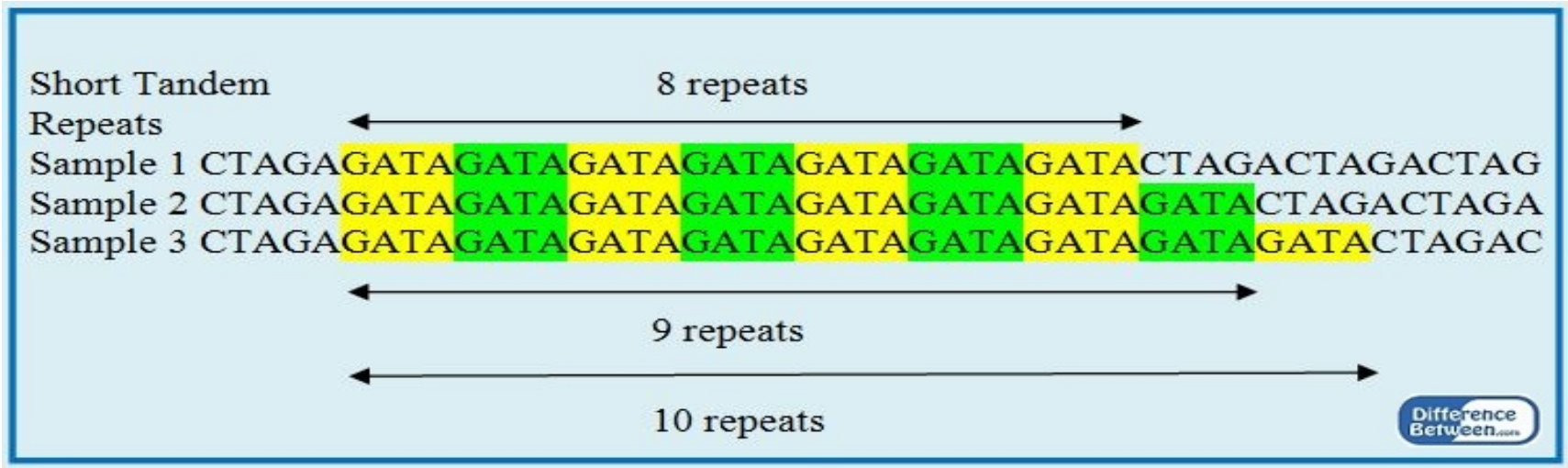
Organism	Estimated size (base pairs)	Chromosome number	Estimated gene number
Human (*Homo sapiens*)	3 billion	46	~25,000
Mouse (*Mus musculus*)	2.9 billion	40	~25,000
Fruit fly (*Drosophila melanogaster*)	165 million	8	13,000
Plant (*Arabidopsis thaliana*)	157 million	10	25,000
Roundworm (*Caenorhabditis elegans*)	97 million	12	19,000
Yeast (*Saccharomyces cerevisiae*)	12 million	32	6,000
Bacteria (*Escherichia coli*)	4.6 million	1	3,200

DNA Fingerprinting

- Every individual has its unique genetic make-up, which may be called its **Fingerprint**.
- The technique developed to identify a person with the help of DNA restriction analysis, is known as **DNA profiling** or **DNA fingerprinting**.
- The technique of fingerprinting was first given by British geneticist, Dr. Alec Jeffreys in 1984.

- DNA fingerprinting technique is based on identification of nucleotide sequence present in this wonder molecule.
- About 99.9% of nucleotide sequence in all persons, is same. Only some short sequences of nucleotides differ from person to person.

- In the population, every person shows unusual sequences of 20-100 base pairs, which are repeated several times. They are termed as Variable Number of Tandem Repeats(VNTRs)
- The length of the regions having VNTRs is different in each individual and hence is the key factor in DNA Profiling.

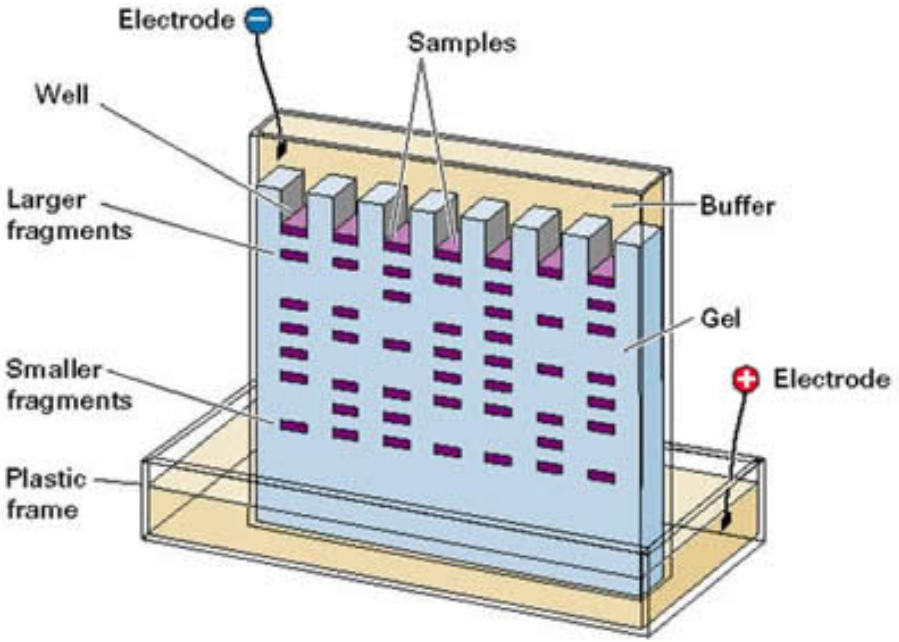


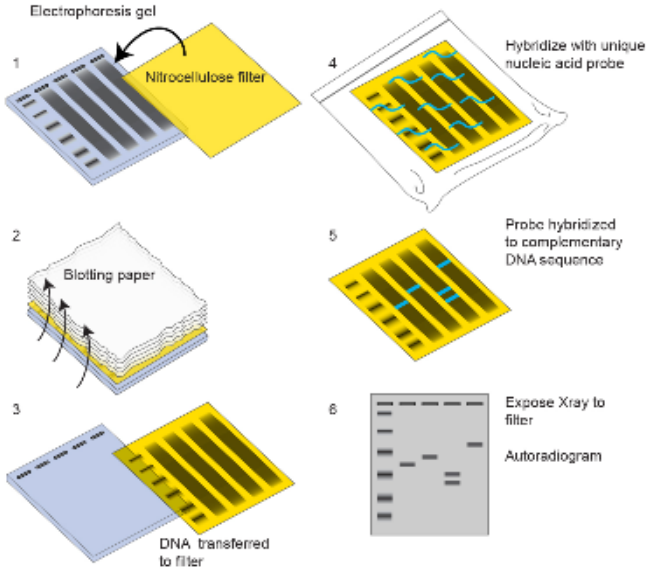
Steps involved in DNA finger printing

1. **Isolation of DNA** : the DNA must be recovered from the cells or tissues of the body (host). Only small amount of tissue like blood, hair roots, skin, etc. is required.
2. **Restriction digestion** : the isolated DNA is treated with restriction enzymes. The restriction enzymes cut the DNA into small fragments having variable lengths. This phenomenon is called **Restriction Fragment Length polymorphism** (RFLP).

3. **<u>Gel electrophoresis</u>** : The DNA samples are loaded for agarose gel electrophoresis under an electric influence. The DNA fragments, which are negatively charged move to the positive pole. The movement of these fragments depends on length of the fragments. This results in formation of bands. dsDNA splits into ssDNA by alkali treatment.

4. **<u>Southern blotting</u>** : The separated DNA fragments are transferred to a nylon membrane or a nitrocellulose filter paper by placing it over the gel and soaking them with filter paper overnight.

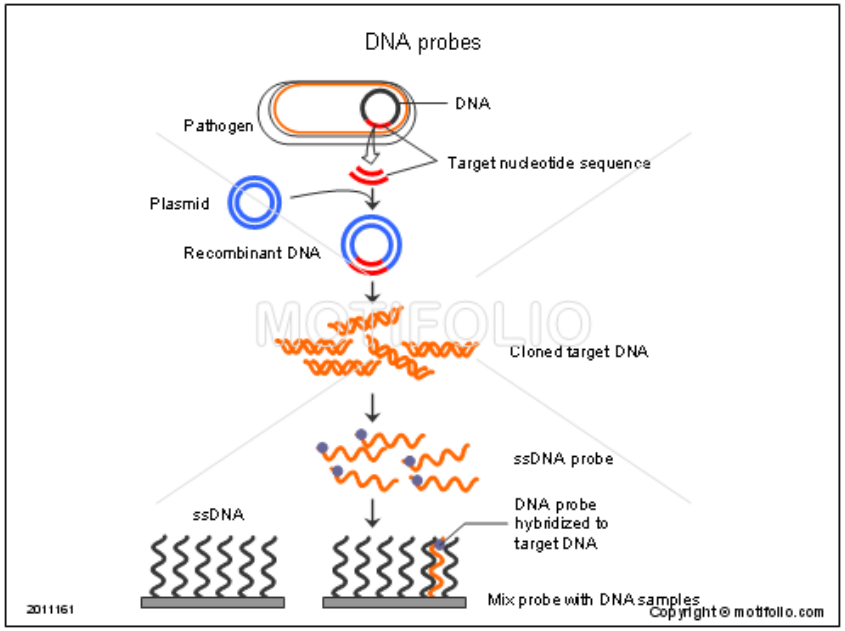


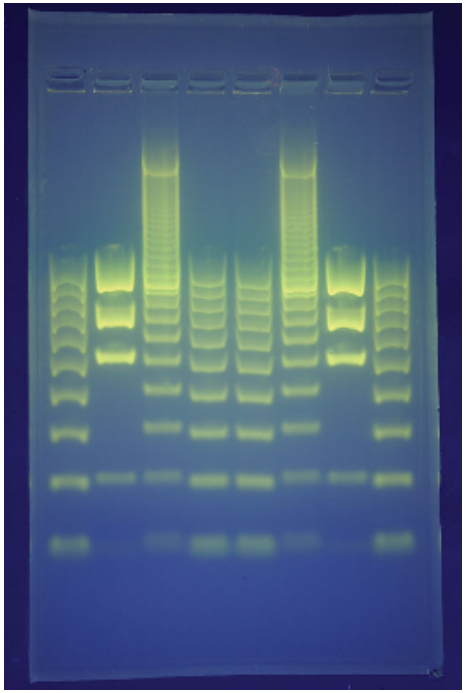


5. **Selection of DNA Probe** : A known sequence of single- stranded DNA is prepared. It is called DNA Probe. DNA Probe is obtained from organisms or prepared by cDNA preparation method. The DNA Probe is labelled with radioactive isotopes.

6. **Hybridization** :Probe DNA is added to the nitrocellulose filter paper containing host DNA. The single-stranded DNA probe pairs with the complementary base sequence of the host DNA strand. As a result DNA-DNA hybrids are formed on the nitrocellulose filter paper. Remaining single stranded DNA probe fragments are washed off.

7. **Photography** : the nitrocellulose filter paper is photographed on an X-ray film by autoradiography. The film is analysed to determine the presence of hybrid DNA.





Application of DNA fingerprinting

1. In forensic science, DNA fingerprinting is used to solve problems of rape and some complicated murder cases.
2. DNA finger printing is used to find out the biological father or mother or both, of the child, in case of disputed parentage.
3. DNA finger printing is used in pedigree analysis in cats, dogs, horses and humans.