Molecular Basis of Inheritance

Nucleic Acids

Nucleotides

- Nucleosides
- Phosphoric Acid

Nucleosides

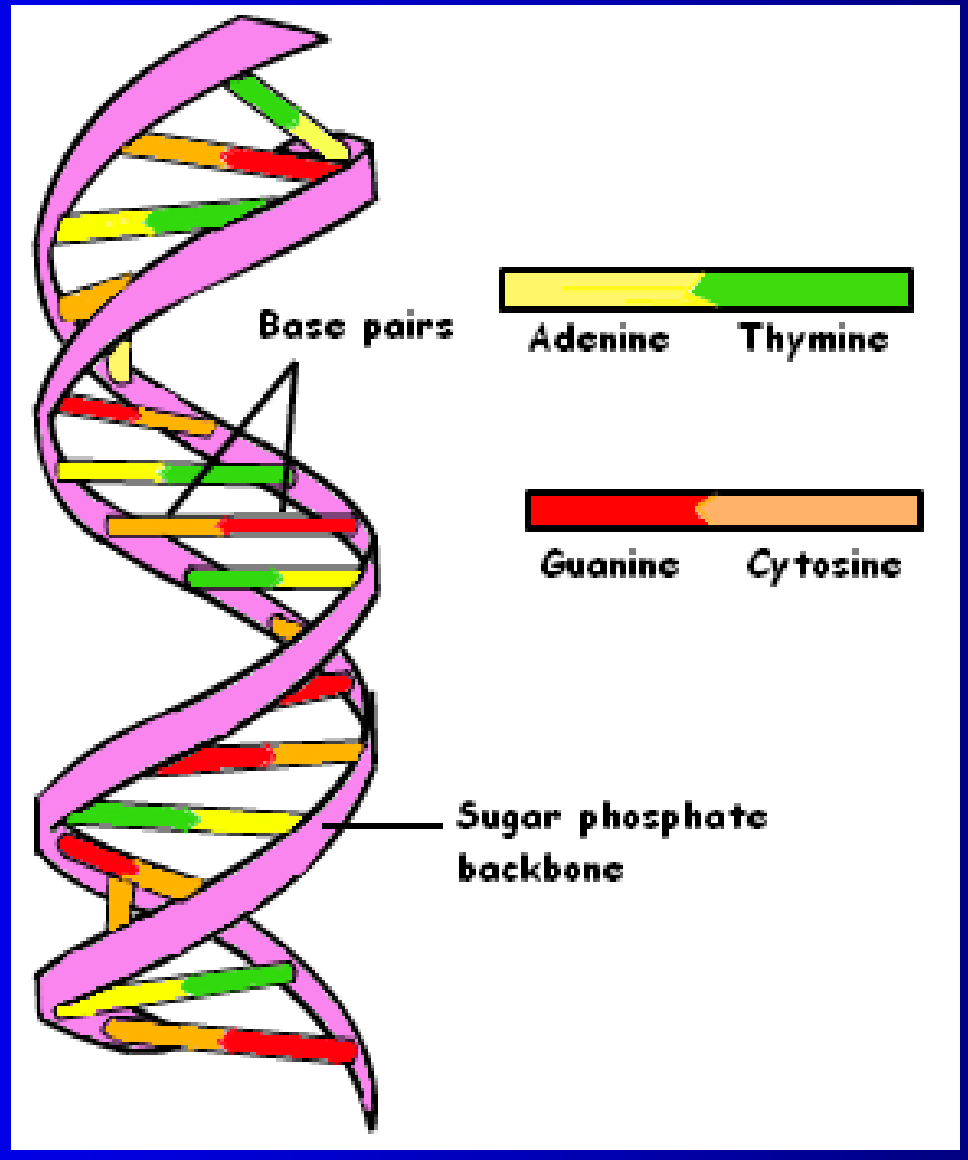
- Nitrogen bases
- Pentose sugar

Nitrogen bases

- Purines
  - Adenine
  - Guanine
- Pyrimidines
  - Cytosine
  - Thymine
  - Uracil

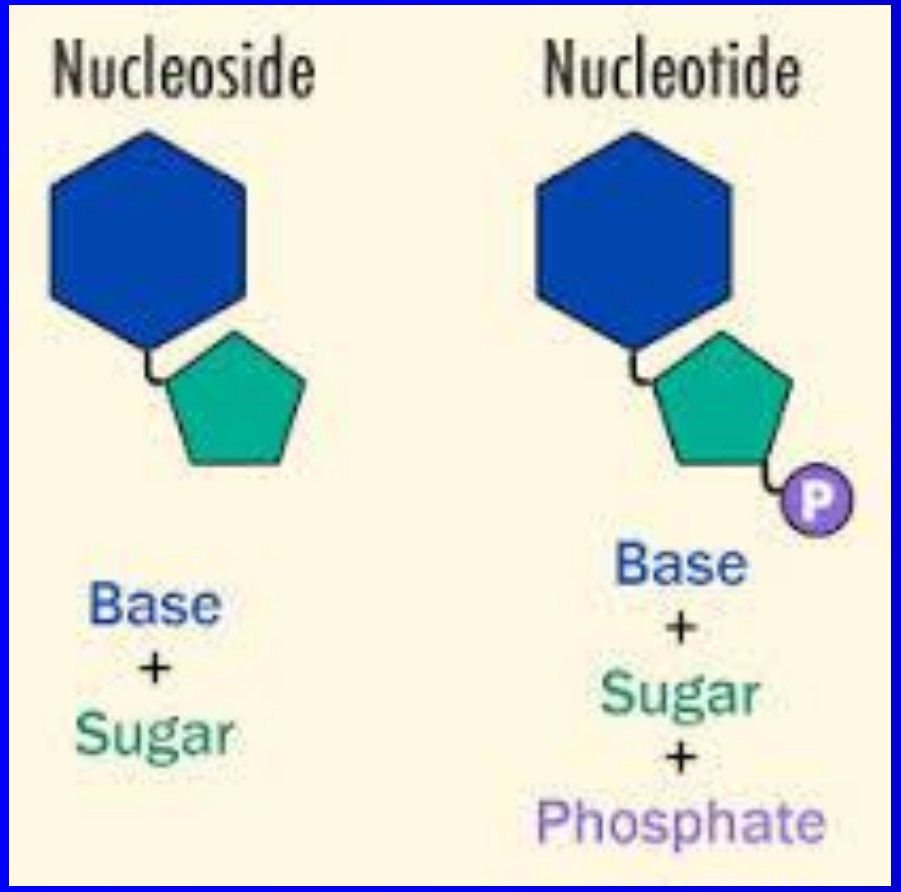
B. John Ebenezer

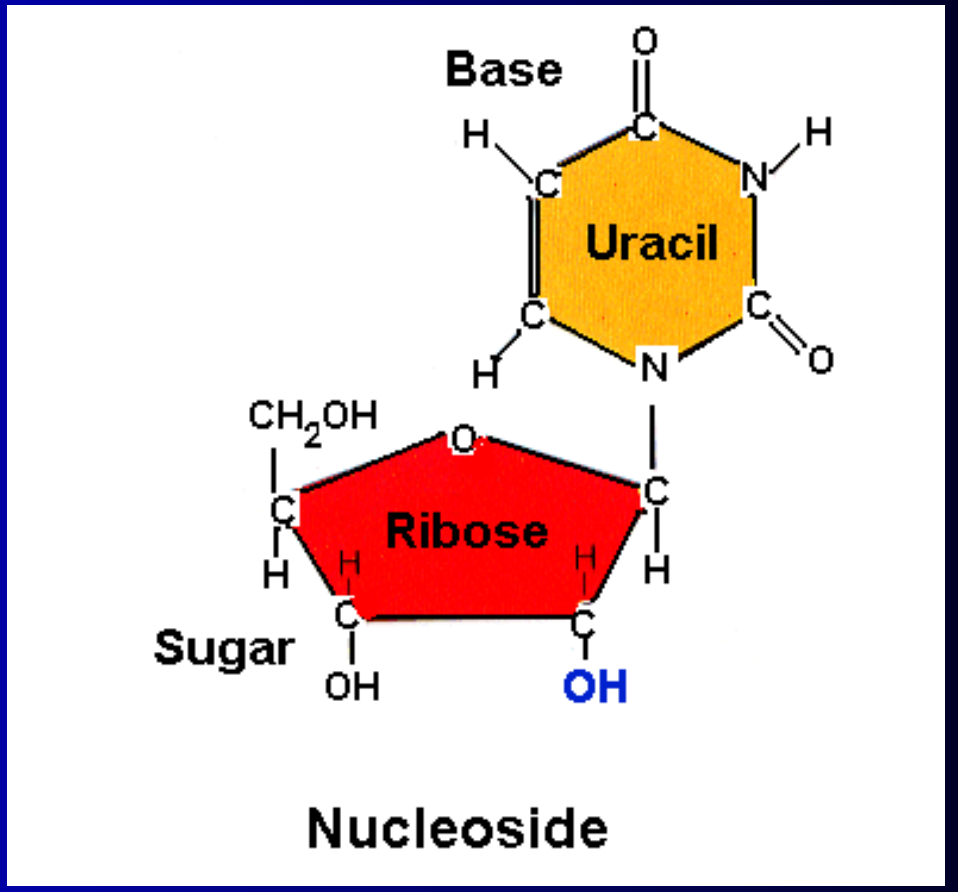
Double Stranded Helical Structure of DNA



B. John Ebenezer

Nucleoside and Nucleotide





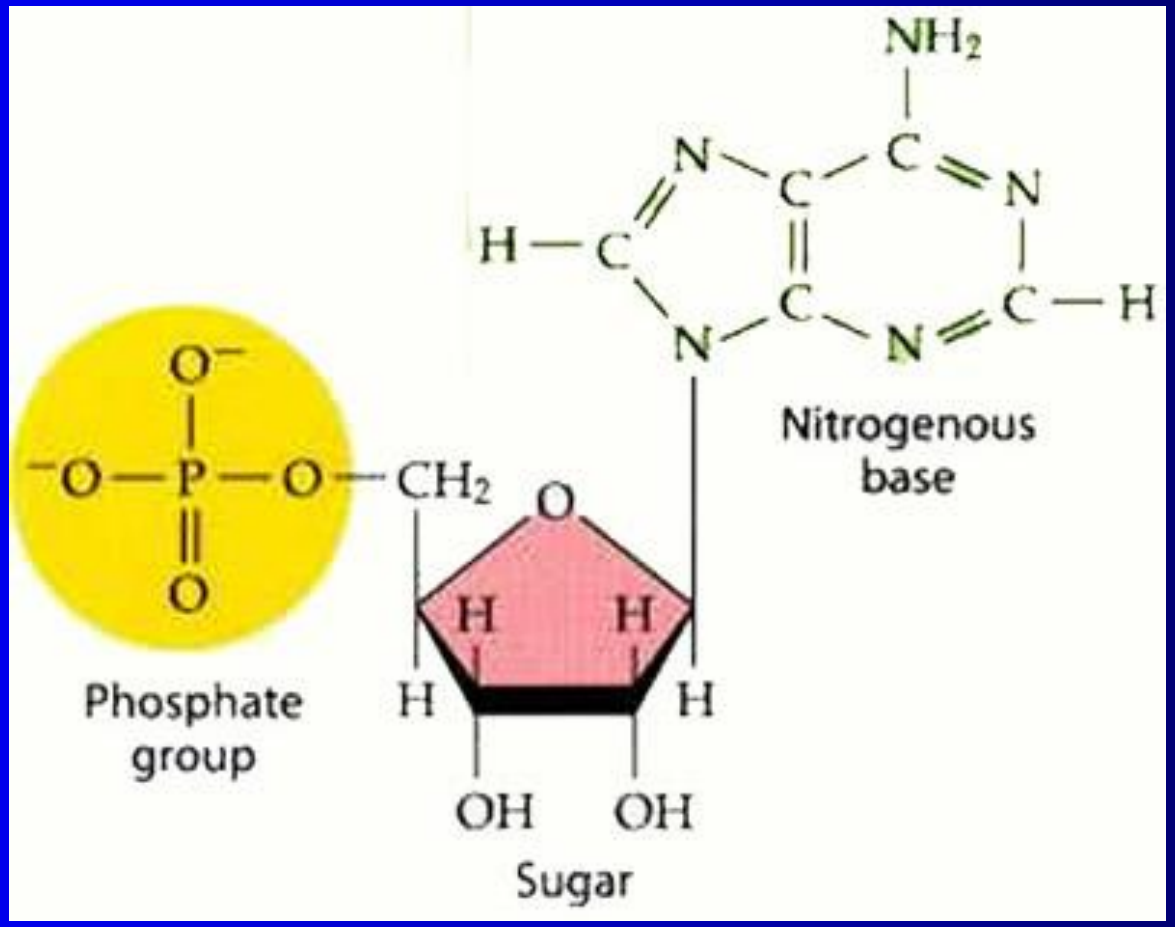
Nitrogen Base + Sugar = Nucleoside

Nitrogen Base + Sugar + Phosphate = Nucleotide

Nucleoside of RNA

B. John Ebenezer

Nucleotide

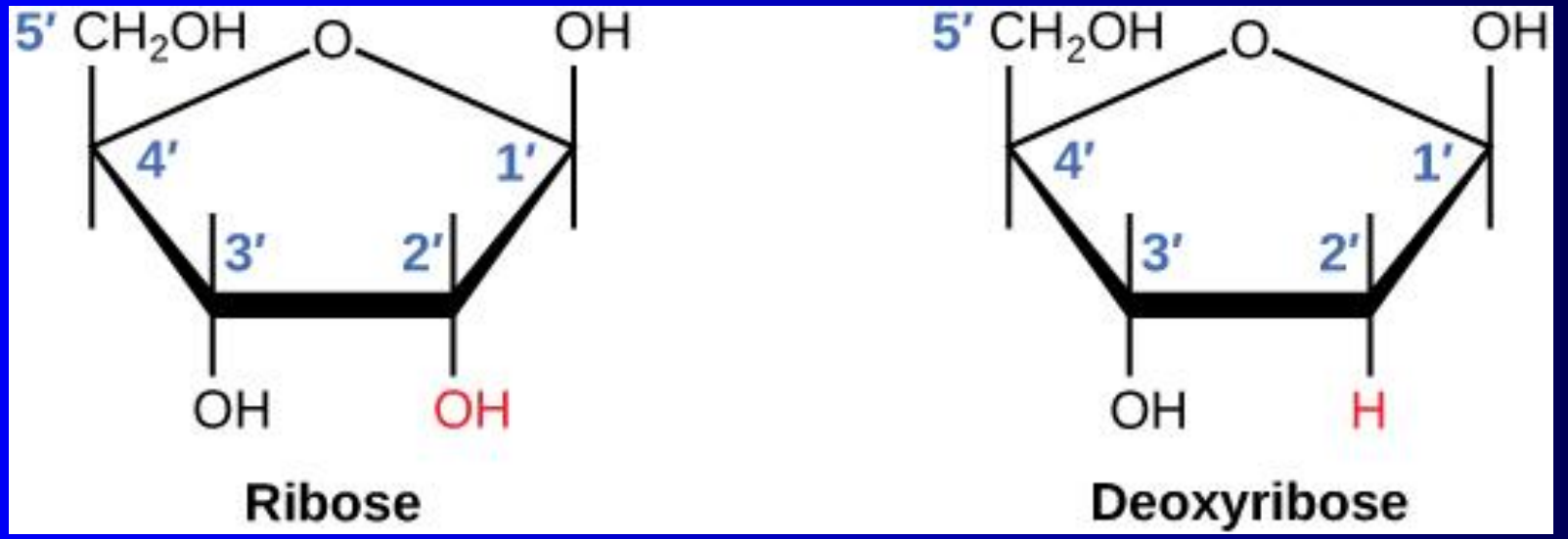


Nitrogen Base + Sugar = Nucleoside

Nucleoside + Phosphate = Nucleotide

B. John Ebenezer

Ribose and Deoxyribose Sugar



B. John Ebenezer

Nucleosides and Nucleotides

Adenosine, guanosine, cytidine thymidine, and uridine are nucleosides.

Adenylic acid, guanylic acid, cytidylic acid thymidylic acid, and uridylic acid are nucleotides.

B. John Ebenezer

The sugar found in nucleic acids is either ribose sugar or deoxyribose sugar.

A nucleic acid which contains ribose sugar is called ribonucleic acid (RNA) while that contains deoxyribose sugar is called deoxyribonucleic acid (DNA).

B. John Ebenezer

Watson & Crick Model of DNA

According to Watson and Crick DNA exists as a double helix.

The two strands of polynucleotides are antiparallel i.e., they run in the opposite direction.

The backbone is formed by the sugar-phosphate-sugar chain.

The nitrogen bases are arranged perpendicular to this backbone facing inside.

Purine pairs with pyrimidine.

B. John Ebenezer

Watson & Crick Model of DNA

DNA is a double stranded helical structure. Each strand appears like a helical staircase.

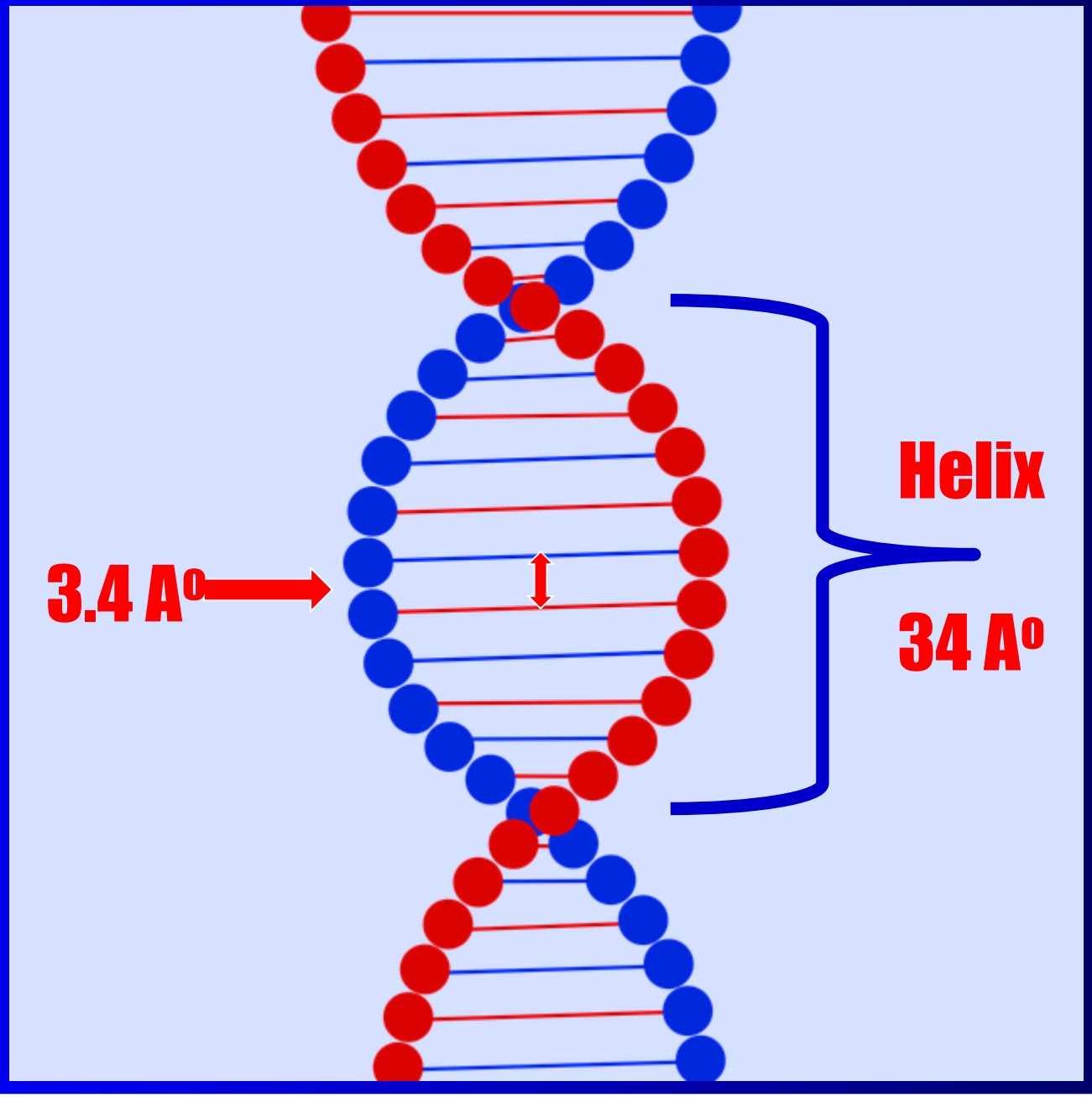
Each helix has a distance of 34A^{o}.

Each helix consists of ten nucleotides.

So the distance between any two nucleotide is 3.4A^{o}. This form of DNA is called B-DNA.

B. John Ebenezer

Sructure of DNA

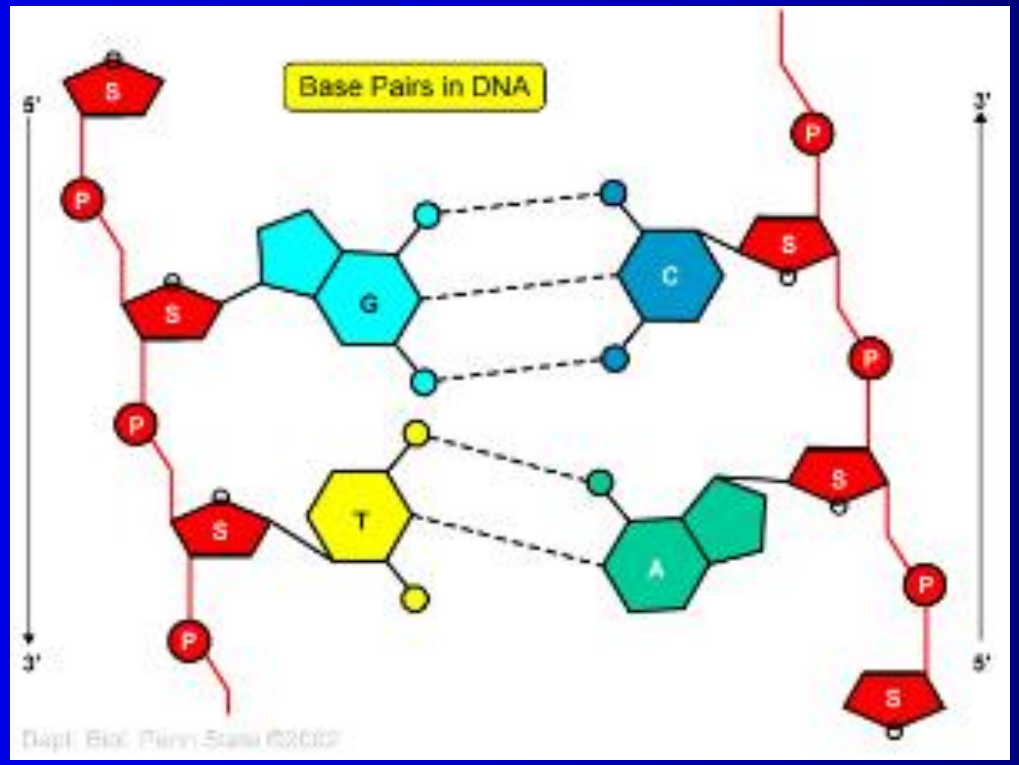


B. John Ebenezer

In DNA, Adenine pairs with Thymine with two hydrogen bonds.

Guanine pairs with Cytosine with three hydrogen bonds.

In RNA Adenine pairs with Uracil but Guanine pairs with Cytosine.



B. John Ebenezer

Phosphodiester bond

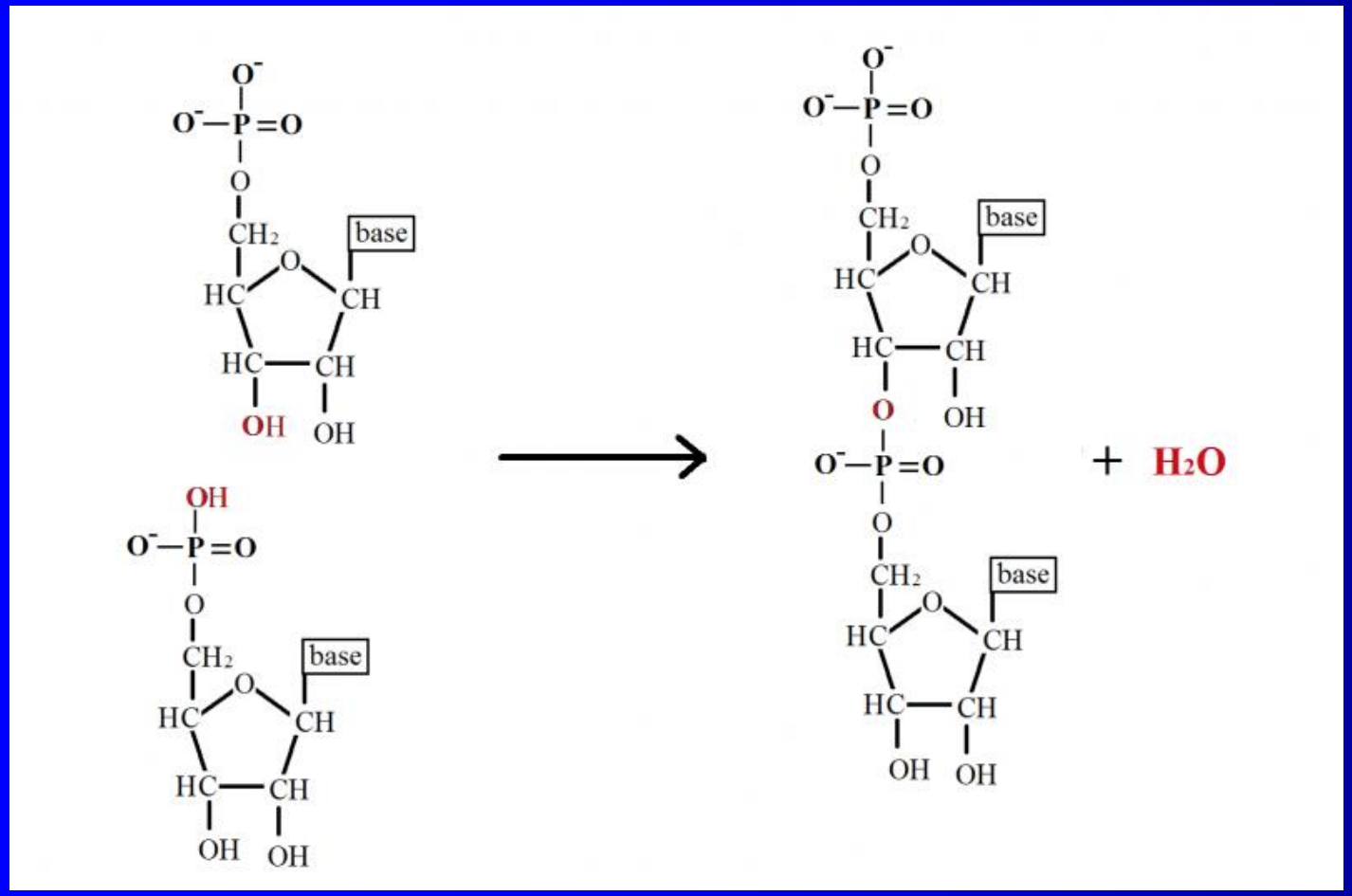
In a nucleic acid a phosphate group links the **3rd C of one sugar** of one nucleotide to the **5th C of the sugar** of the next nucleotide.

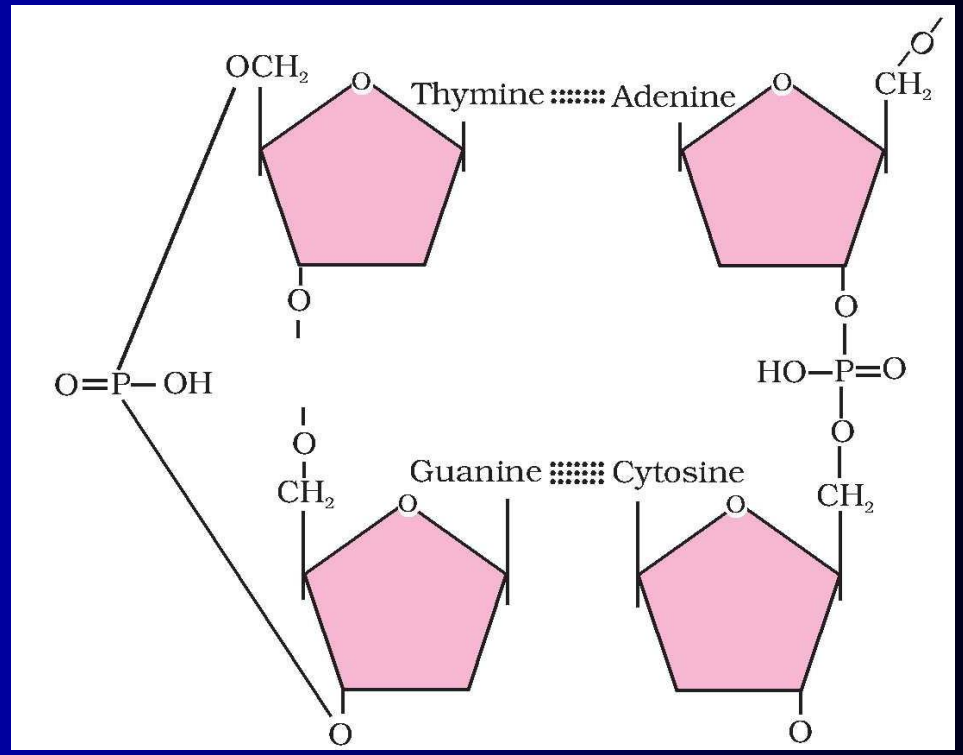
The bond formed between the phosphate and hydroxyl group of sugar is an **ester bond.**

As there are two ester bonds in a nucleic acid, it is called **phosphodiester bond.**

B. John Ebenezer

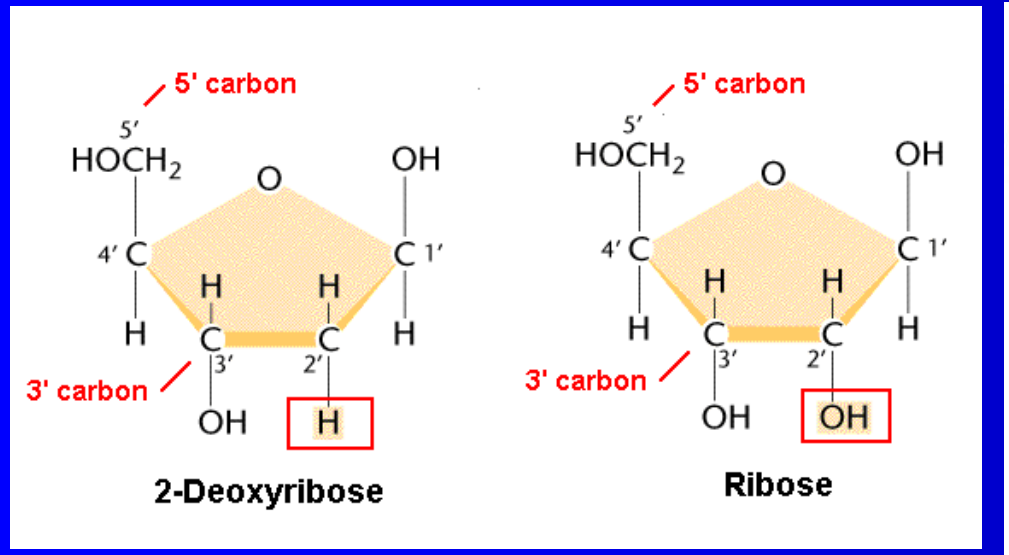
Phosphodiester bond

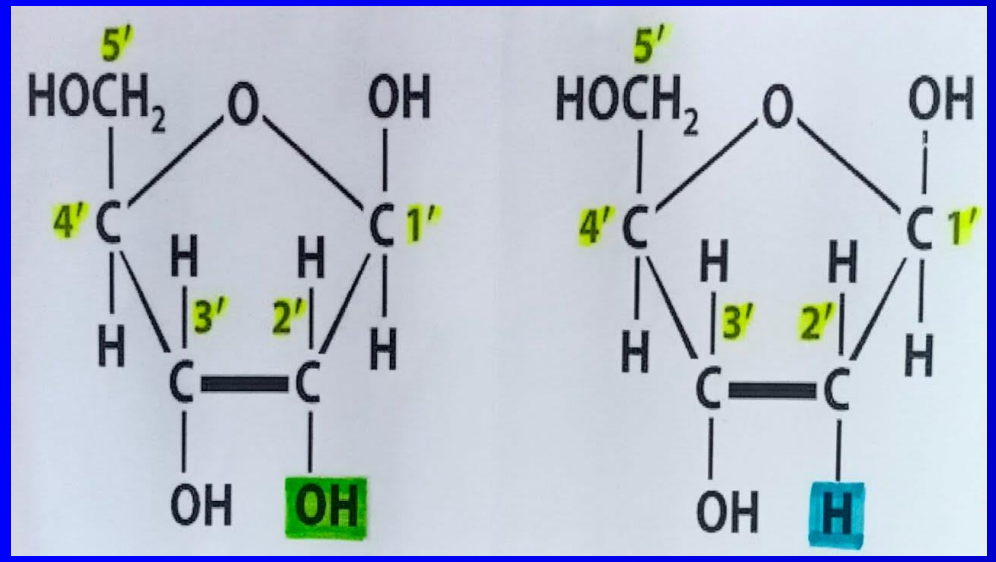


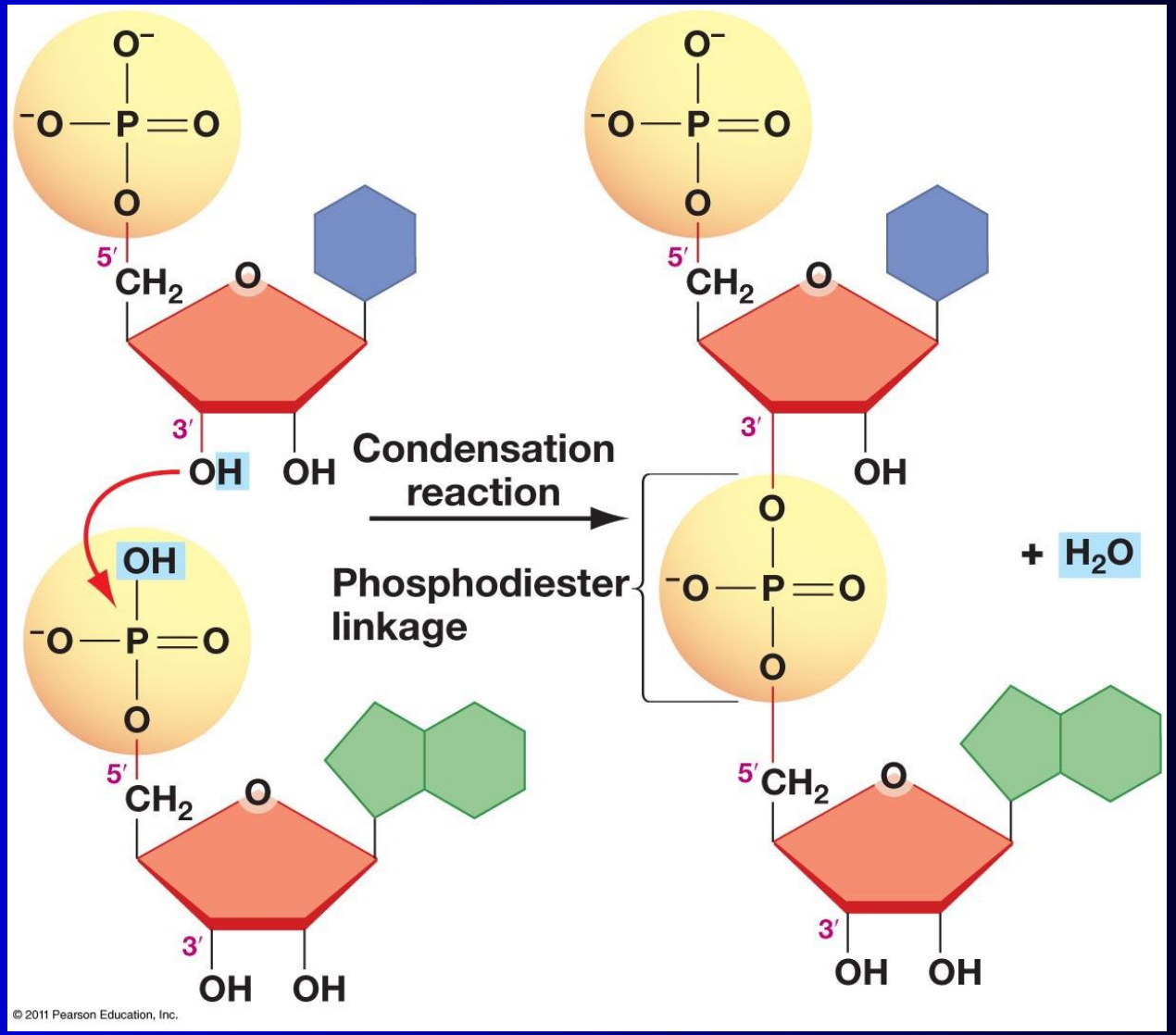


B. John Ebenezer

Phosphodiester bond

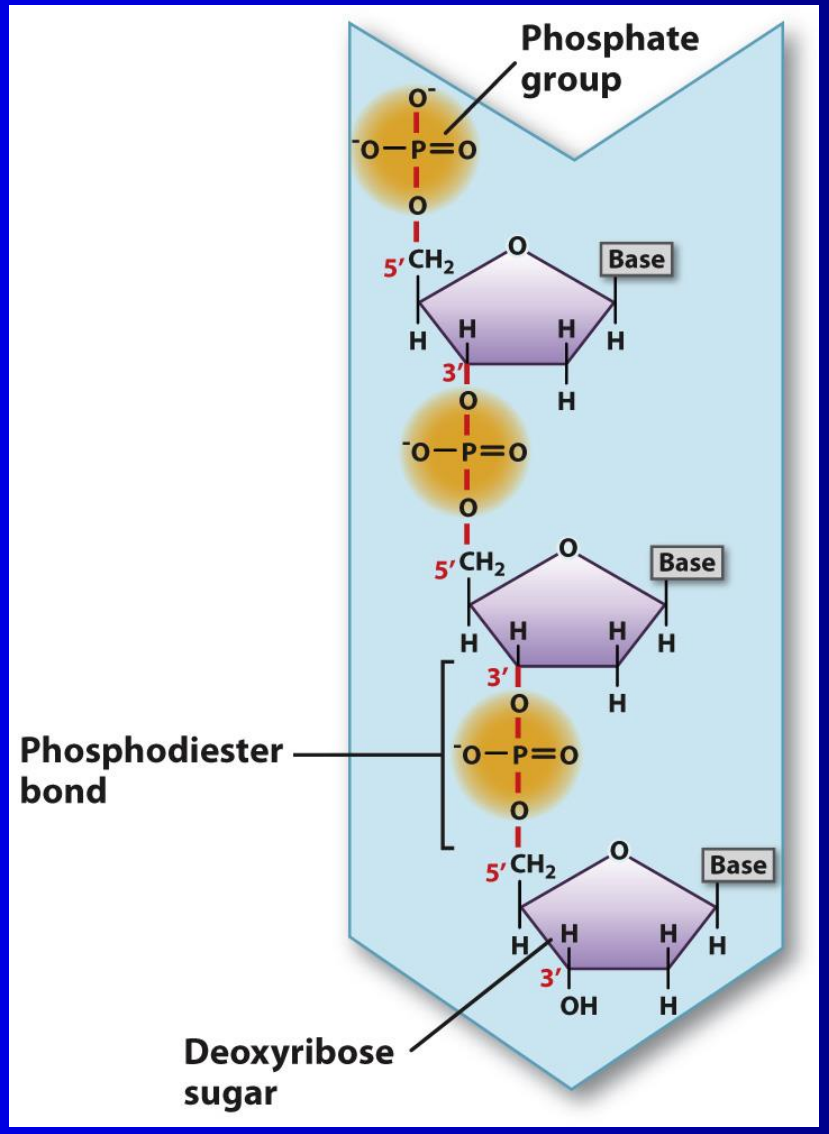






Phosphodiester bond

CH_2OH Hydroxymethyl
CH_2 Methylene
OH Hydroxyl
PO_4 Phosphate



B. John Ebenezer

The Salient Features of DNA

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

(i) It is made of two polynucleotide chains, where the backbone is made by sugar-phosphate, and the bases project inside.

(ii) The two chains have anti-parallel polarity. It means, if one chain has the polarity 5′ - 3′, the other has 3′ - 5′.

(iii) The bases in two strands are paired through hydrogen bonds (H-bonds) forming base pairs (bp).

B. John Ebenezer

The Salient Features of DNA

Adenine forms two hydrogen bonds with Thymine from opposite strand and vice-versa.

Guanine is bonded with Cytosine with three H-bonds.

As a result, always a purine comes opposite to a pyrimidine.

This makes uniform distance between the two strands of the helix.

B. John Ebenezer

The Salient Features of DNA

(iv) The two chains are coiled in a right-handed fashion.

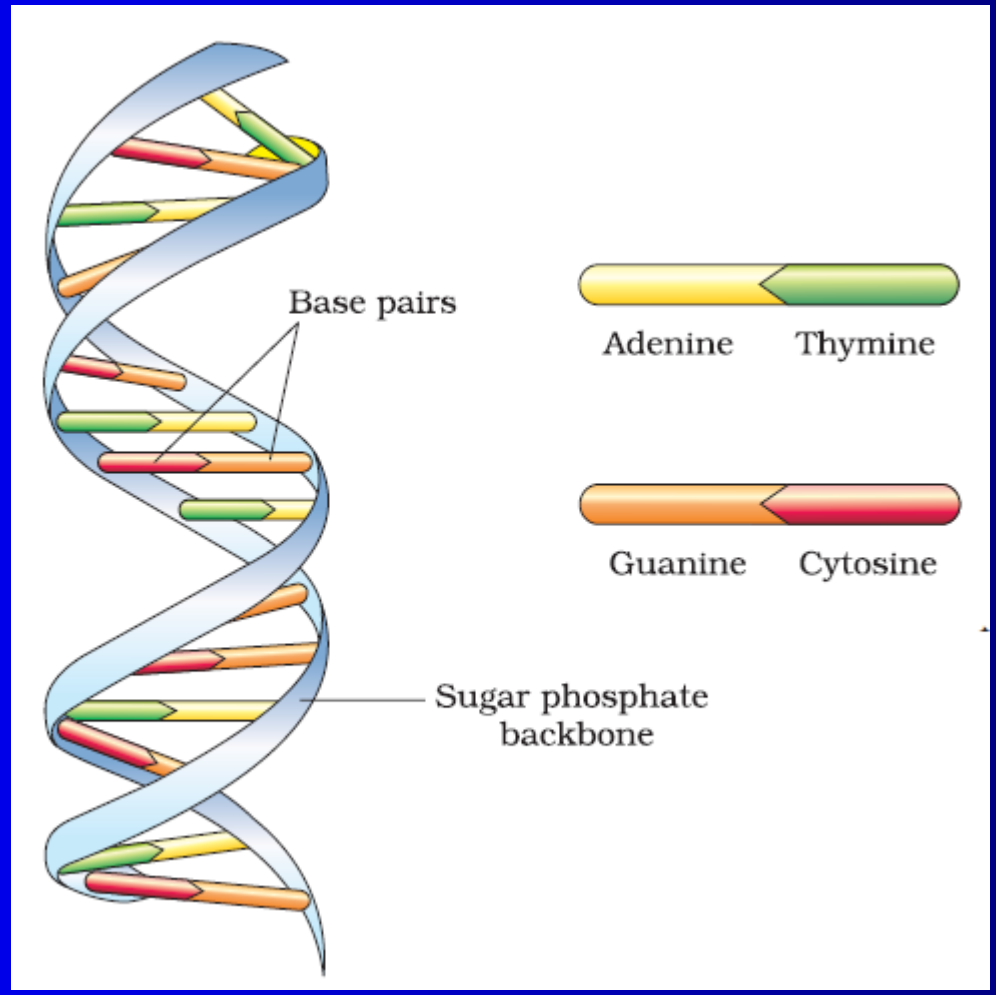
The pitch of the helix is 3.4 nm (a nanometre is one billionth of a metre, that is 10^{-9} m) and there are roughly 10 bp in each turn.

Consequently, the distance between a bp in a helix is approximately equal to 0.34 nm.

(v) 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

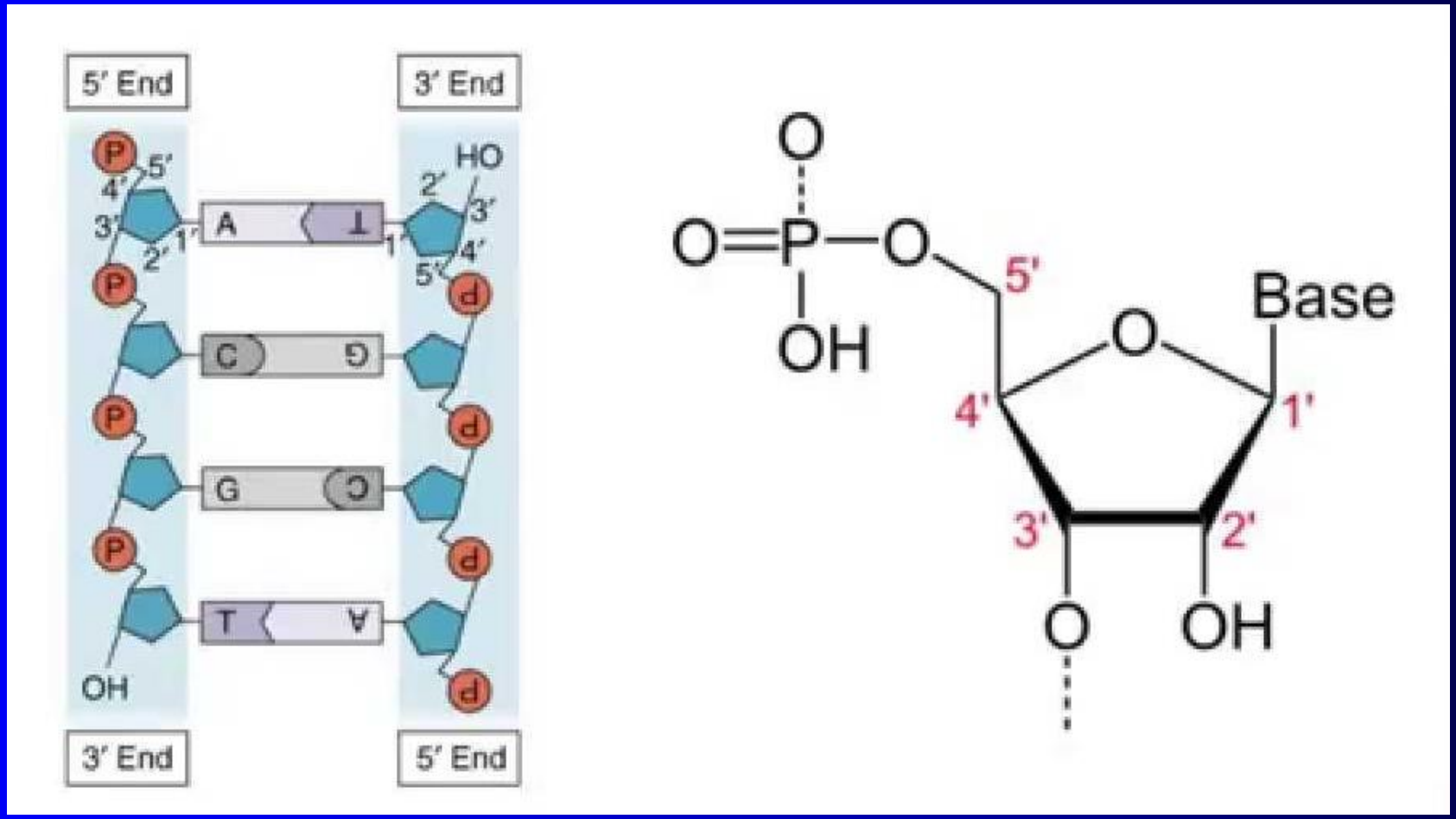
B. John Ebenezer

The DNA



B. John Ebenezer

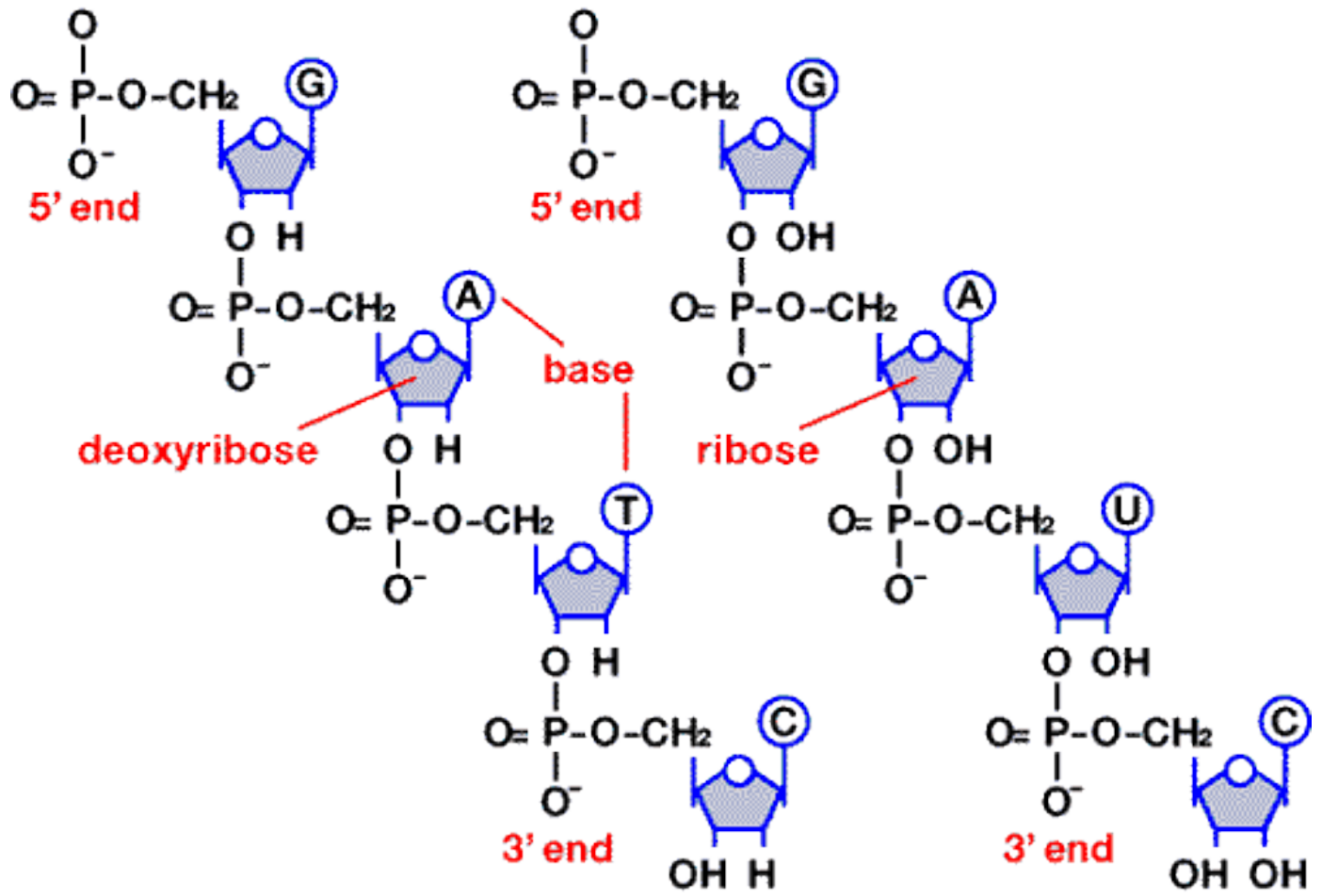
5’ and 3’ ends of DNA



B. John Ebenezer

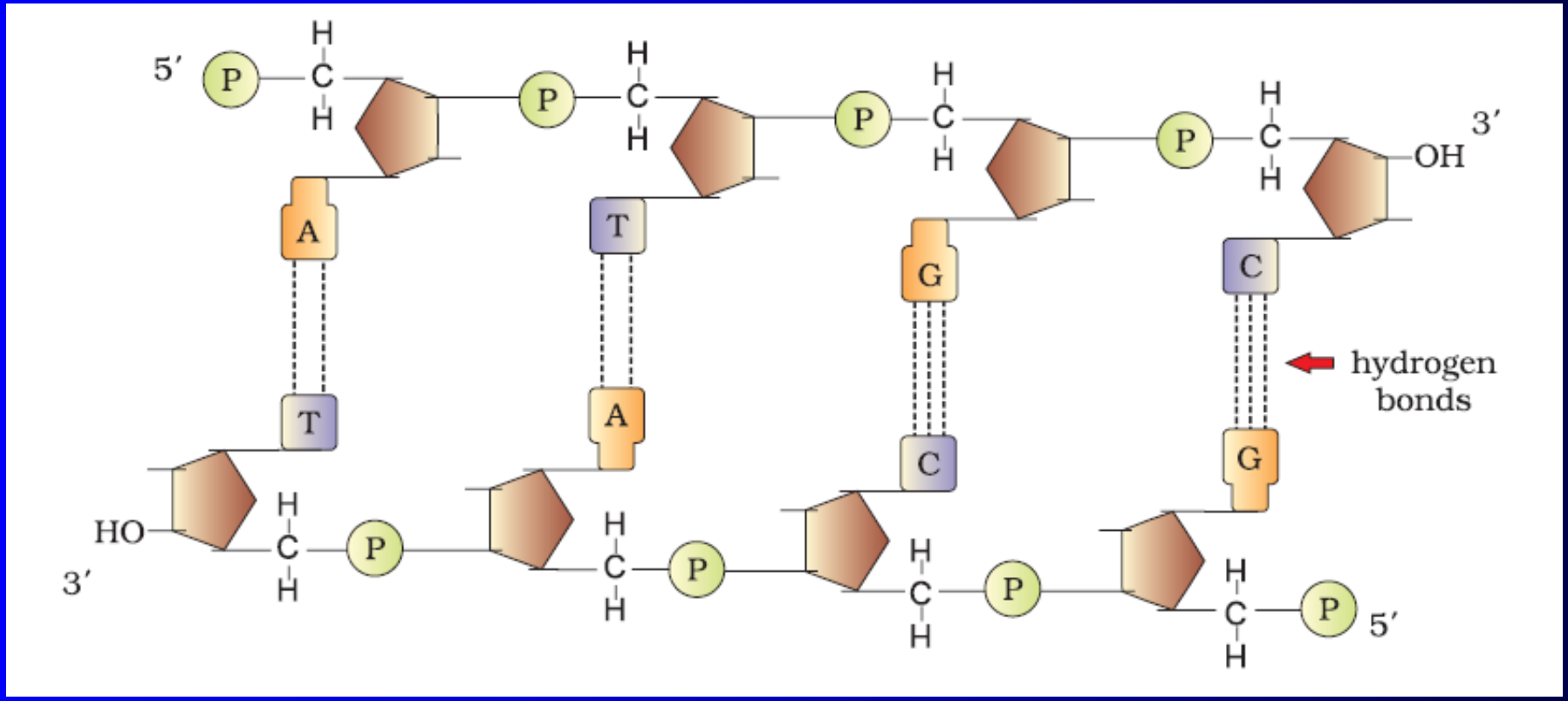
5’ and 3’ ends of DNA

Fig: H



B. John Ebenezer

The DNA



B. John Ebenezer

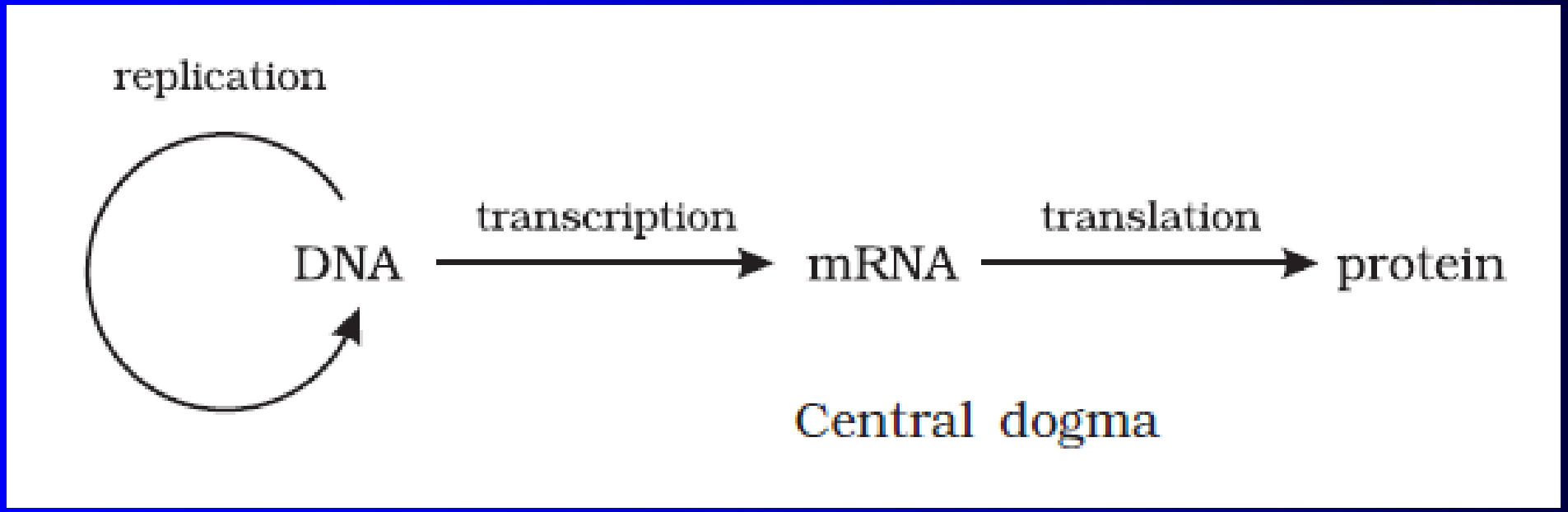
Central Dogma

Francis Crick proposed the Central dogma in molecular biology, which states that the genetic information flows from DNA to RNA and RNA to Protein.

In some viruses the flow of information is in reverse direction, that is, from RNA to DNA.

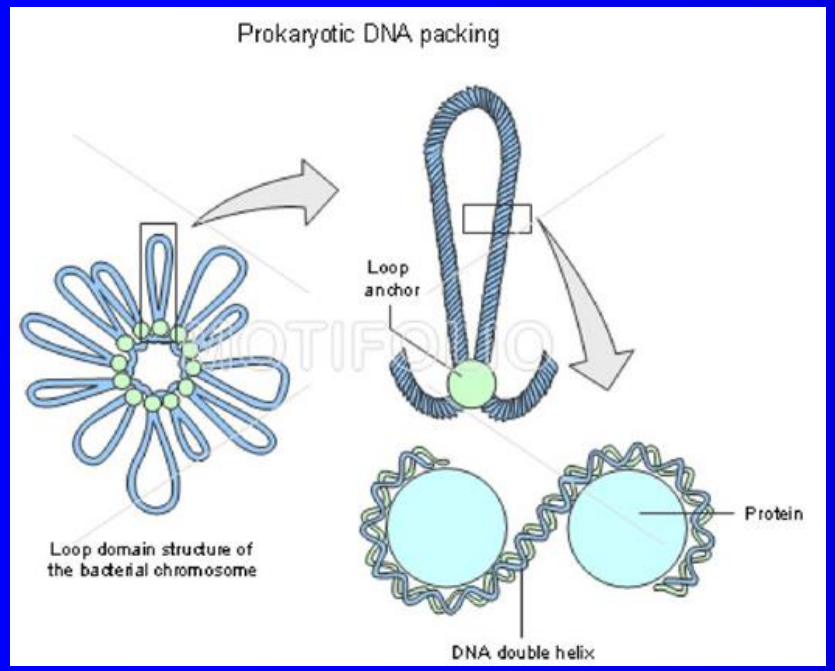
B. John Ebenezer

Central Dogma



B. John Ebenezer

Packaging of DNA in Prokaryotes



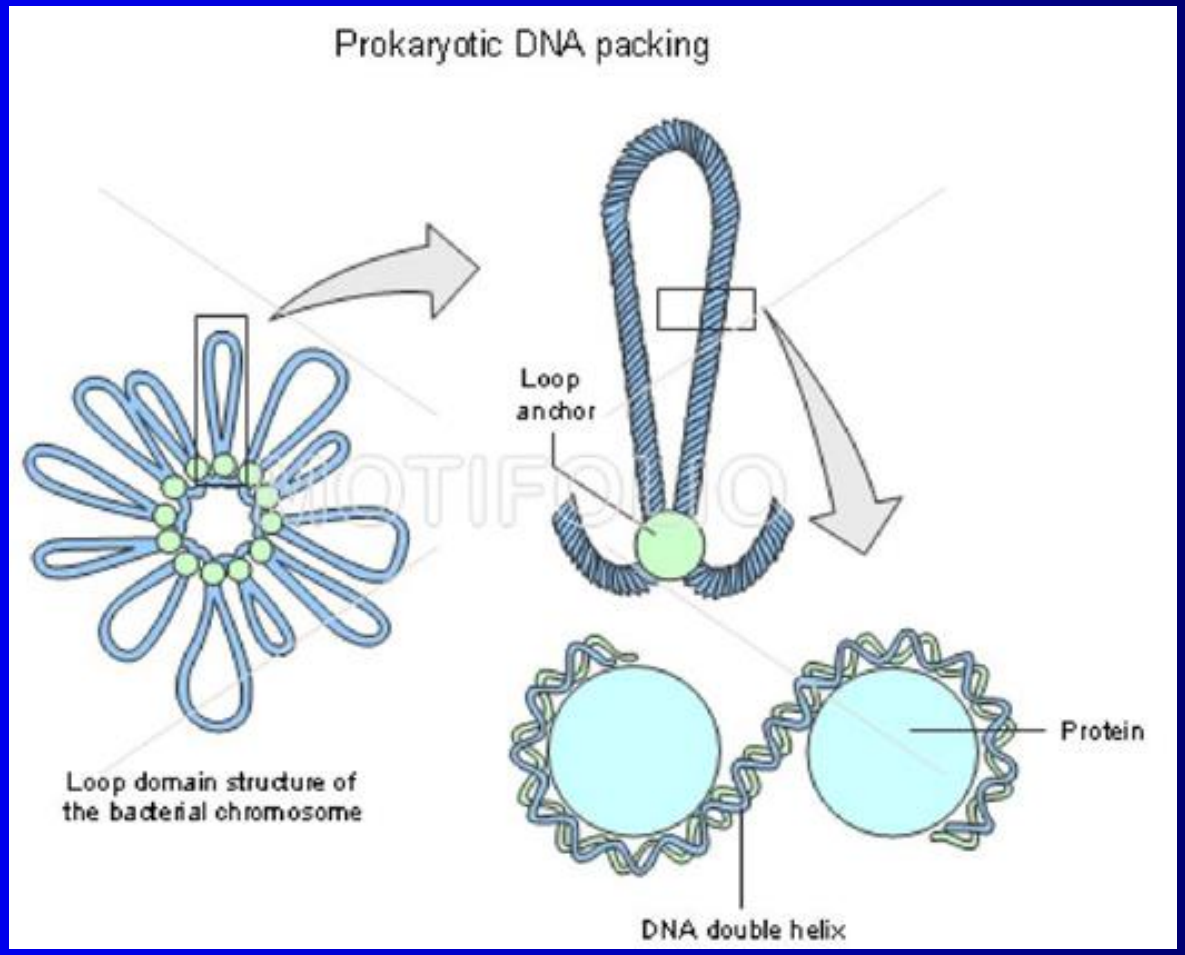
In prokaryotes, such as, *E. coli*, though they do not have a defined nucleus, the DNA is not scattered throughout the cell.

DNA being negatively charged, it is held with some proteins that have positive charges in a region termed as ‘nucleoid’.

The DNA in nucleoid is organised in large loops held by proteins.

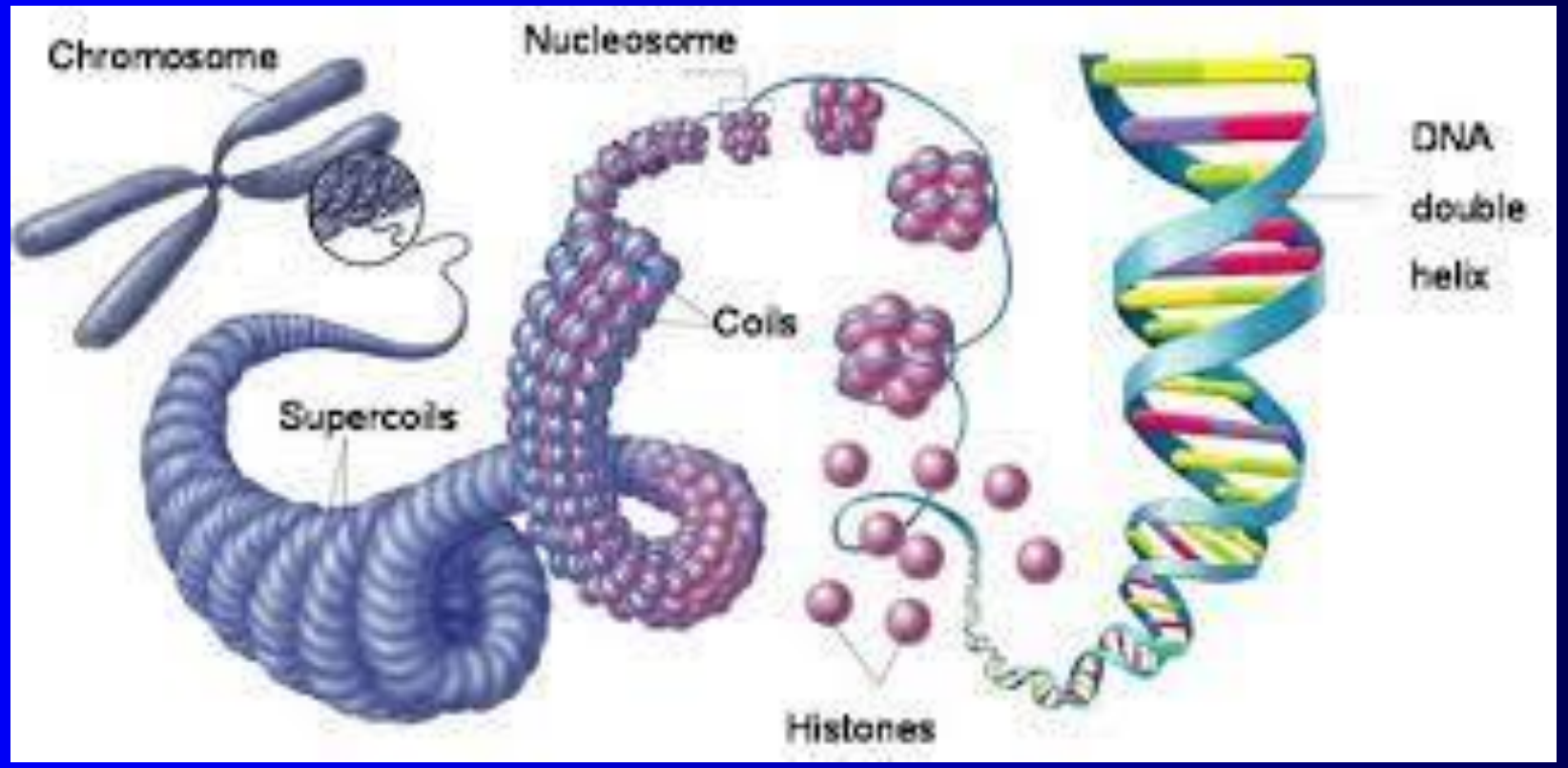
B. John Ebenezer

Packaging of DNA in Prokaryotes



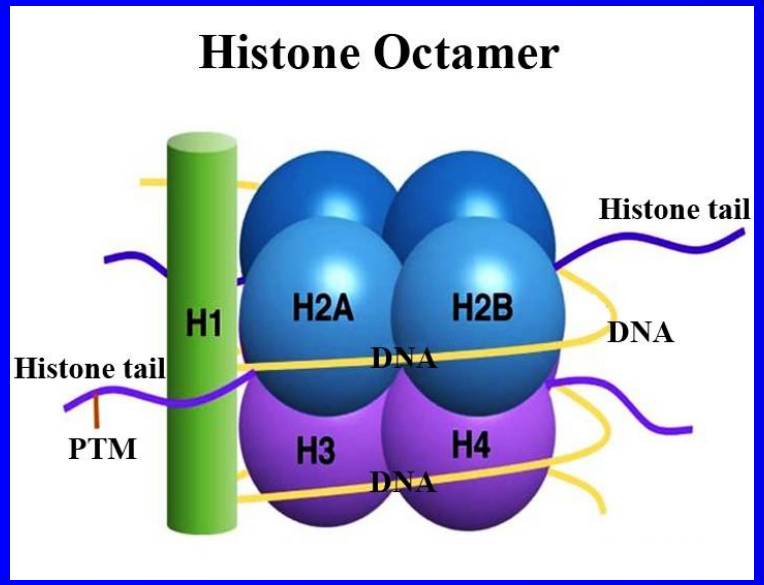
B. John Ebenezer

Packaging of DNA in Eukaryotes



B. John Ebenezer

Packaging of DNA in Eukaryotes



In eukaryotes, this organisation is much more complex.

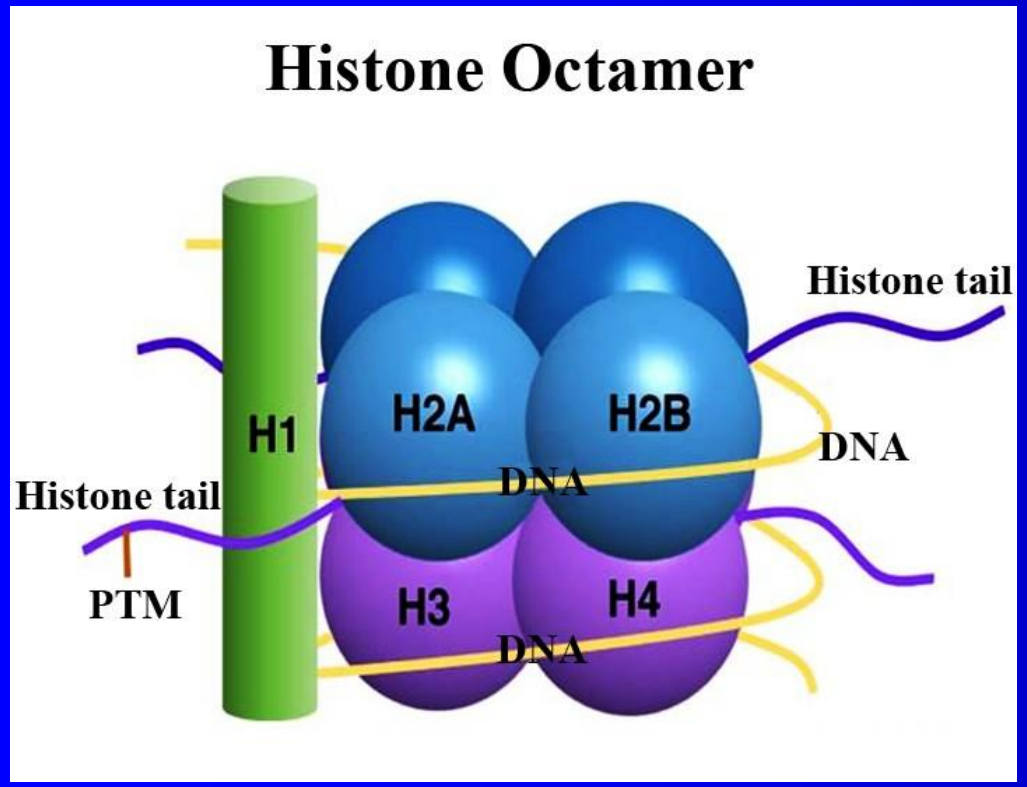
There is a set of **positively charged, basic proteins** called **histones**.

A protein acquires charge depending upon the abundance of **amino acids** with **charged side chains**.

Histones are rich in the basic amino acids **lysines** and **arginines.**

B. John Ebenezer

Packaging of DNA in Eukaryotes

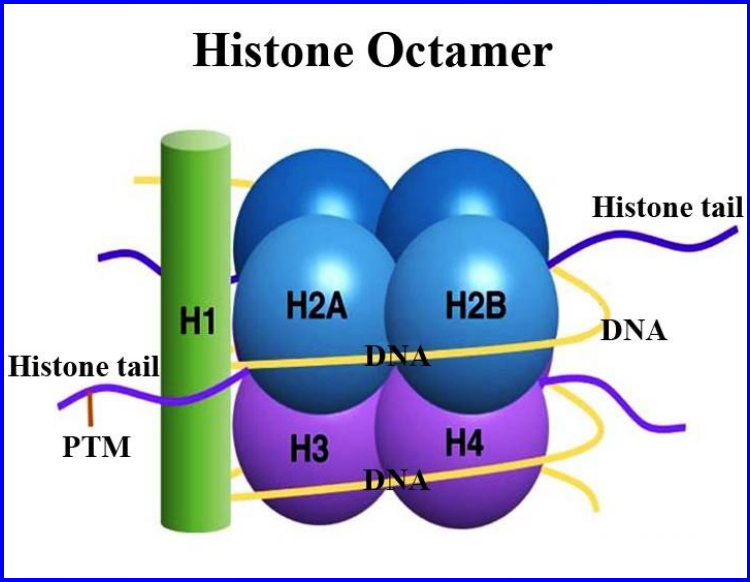


Both the amino acids carry **positive charges** in their side chains.

Histones are organised to form a unit of eight molecules called as **histone octamer**.

B. John Ebenezer

Packaging of DNA in Eukaryotes



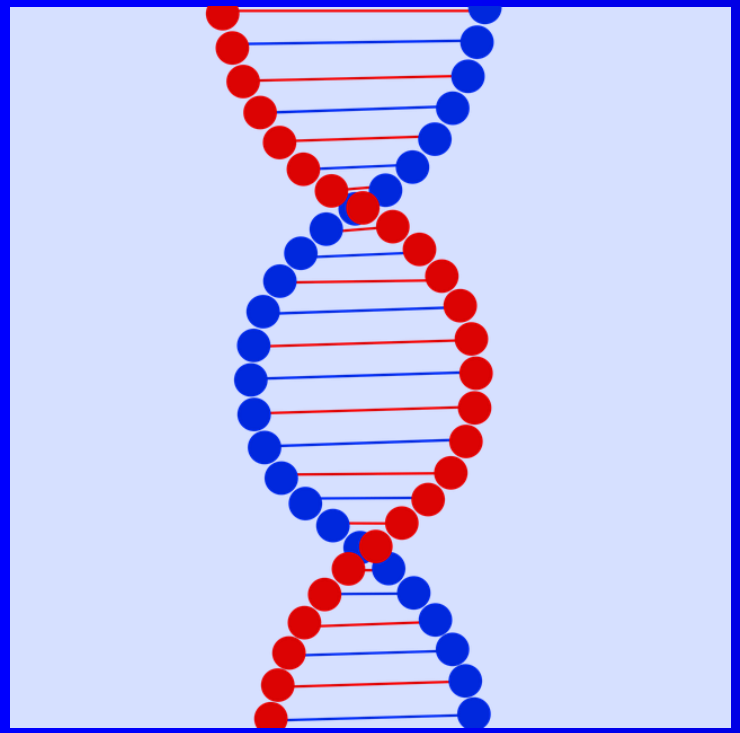
The **negatively charged DNA** is wrapped around the **positively charged histone** octamer to form a structure called **nucleosome**.

A typical nucleosome contains 200 bp of DNA helix.

Nucleosomes form the repeating unit of a structure in nucleus called **chromatin,** thread-like stained (coloured) bodies seen in nucleus.

B. John Ebenezer

Packaging of DNA in Eukaryotes



The nucleosomes in chromatin are seen as ‘beads-on-string’ structure when viewed under electron microscope .

The beads-on-string structure in chromatin is packaged to form chromatin fibers.

They are further coiled and condensed at metaphase stage of cell division to form chromosomes.

B. John Ebenezer

Packaging of DNA in Eukaryotes

The packaging of chromatin at higher level requires additional set of proteins that are collectively referred as **Non-histone Chromosomal (NHC) proteins**.

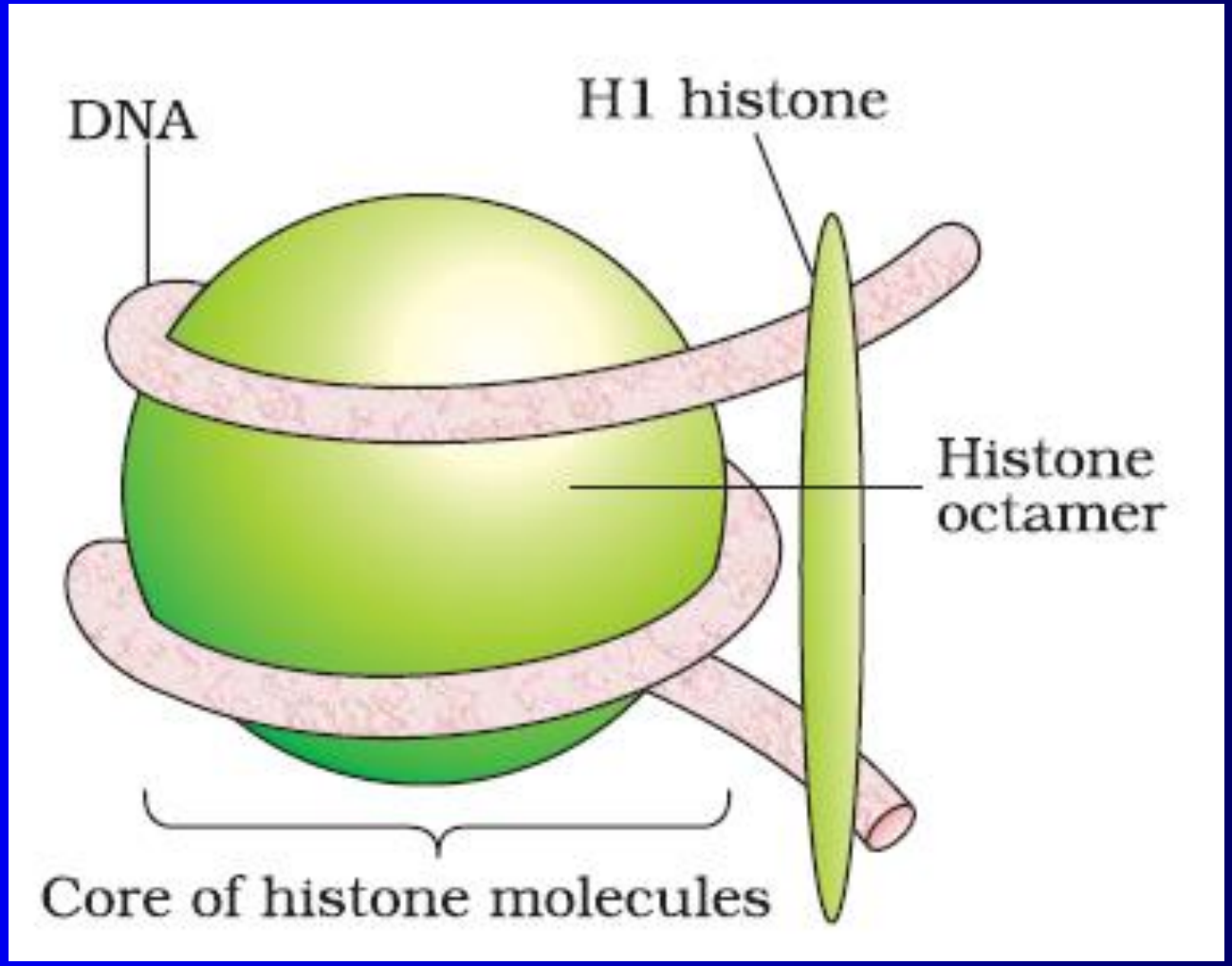
In a typical nucleus, some region of chromatin are **loosely packed** and **stains light** and are referred to as **euchromatin**.

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.

B. John Ebenezer

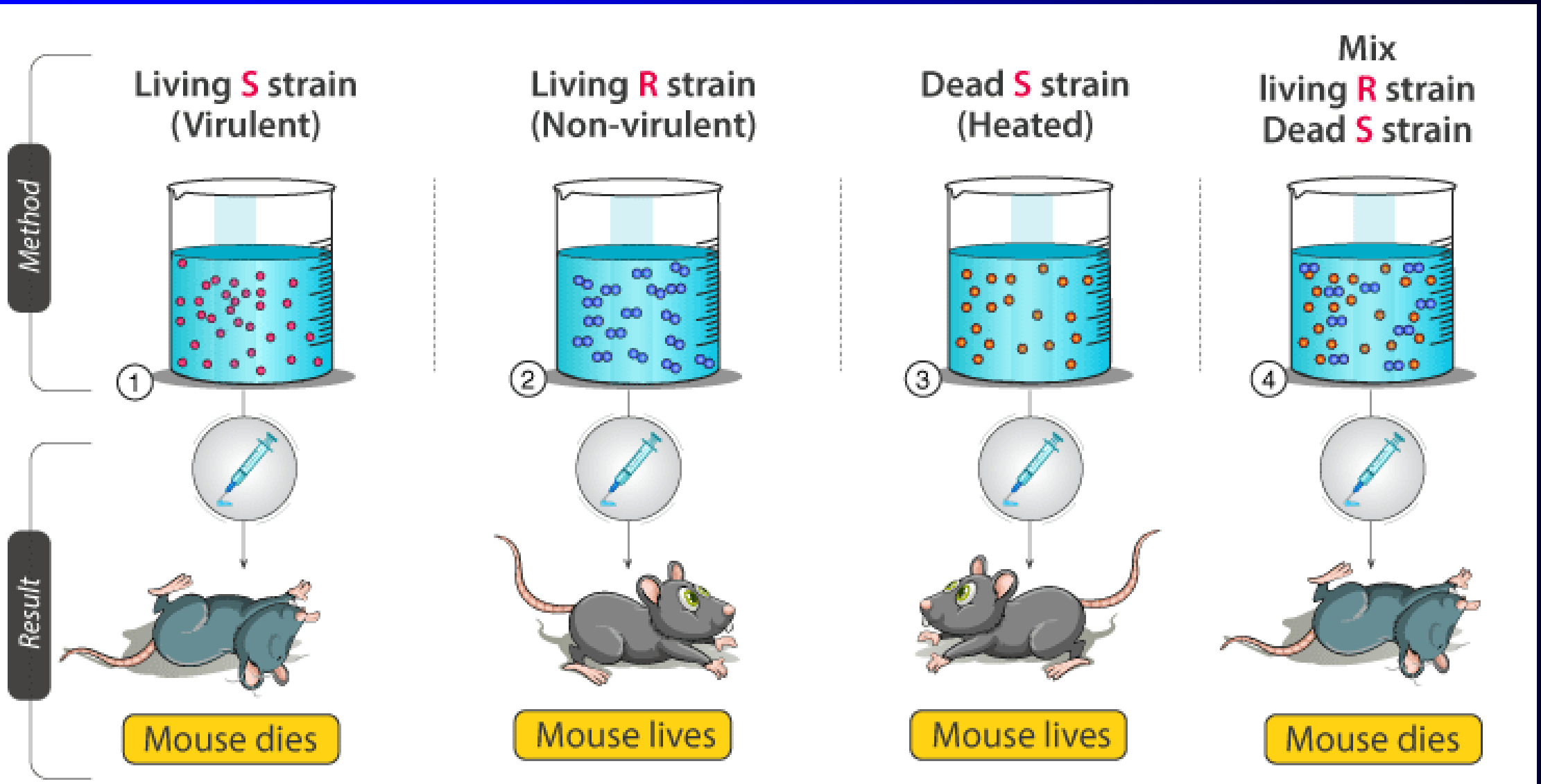
DNA Packaging in Eukaryotes



B. John Ebenezer

Griffith Experiment

Transformation of Genetic Material



B. John Ebenezer

Transformation of Genetic Material

In 1928, Frederick Griffith, in a series of experiments with *Streptococcus pneumoniae* (bacterium responsible for pneumonia), witnessed a miraculous transformation in the bacteria.

During the course of his experiment, bacteria had changed in physical form.

When *Streptococcus pneumoniae* (pneumococcus) bacteria are grown on a culture plate, some produce smooth shiny colonies (S) while others produce rough colonies (R).

B. John Ebenezer

Transformation of Genetic Material

This is because the S strain bacteria have a mucous (polysaccharide) coat, while R strain does not have a mucous coat.

Mice infected with the S strain (virulent) die from pneumonia infection but mice infected with the R strain do not develop pneumonia.

Griffith was able to kill bacteria by heating them.

He observed that heat-killed S strain bacteria injected into mice did not kill them.

When he injected a mixture of heat-killed S and live R bacteria, the mice died.

B. John Ebenezer

Transformation of Genetic Material

Moreover, he recovered living S bacteria from the dead mice.

He concluded that the heat-killed S strain bacteria transformed R strain bacteria.

Some ‘transforming principle’, transferred from the heat-killed S strain, enabled the R strain to synthesise a smooth polysaccharide coat and become virulent.

This must be due to the transfer of the genetic material.

However, the biochemical nature of genetic material was not defined from his experiments.

B. John Ebenezer

The Biochemical Characterization of Transforming Principle

Protein —Proteases→ **Transformation occurs**

So the transforming substance was not a protein.

RNA —Ribonucleases (RNAses)→ **Transformation occurs**

So the transforming substance was not RNA.

DNA —Deoxyribonucleases (DNAses)→ **Transformation did not occur**

So the transforming substance was DNA.

They concluded that DNA is the genetic material.

B. John Ebenezer

The Biochemical Characterization of Transforming Principle

Prior to the work of **Oswald Avery, Colin MacLeod and Maclyn McCarty (1933-44),** the genetic material was thought to be a protein.

They worked to determine the biochemical nature of ‘transforming principle’ in Griffith's experiment.

They purified biochemicals (proteins, DNA, RNA, etc.) from the heat-killed S cells to see which ones could transform live R cells into S cells.

They discovered that DNA alone from S bacteria caused R bacteria to become transformed.

B. John Ebenezer

The Biochemical Characterization of Transforming Principle

They also discovered that protein-digesting enzymes (proteases) and

RNA-digesting enzymes (RNAses) did not affect transformation, so the transforming substance was not a protein or RNA.

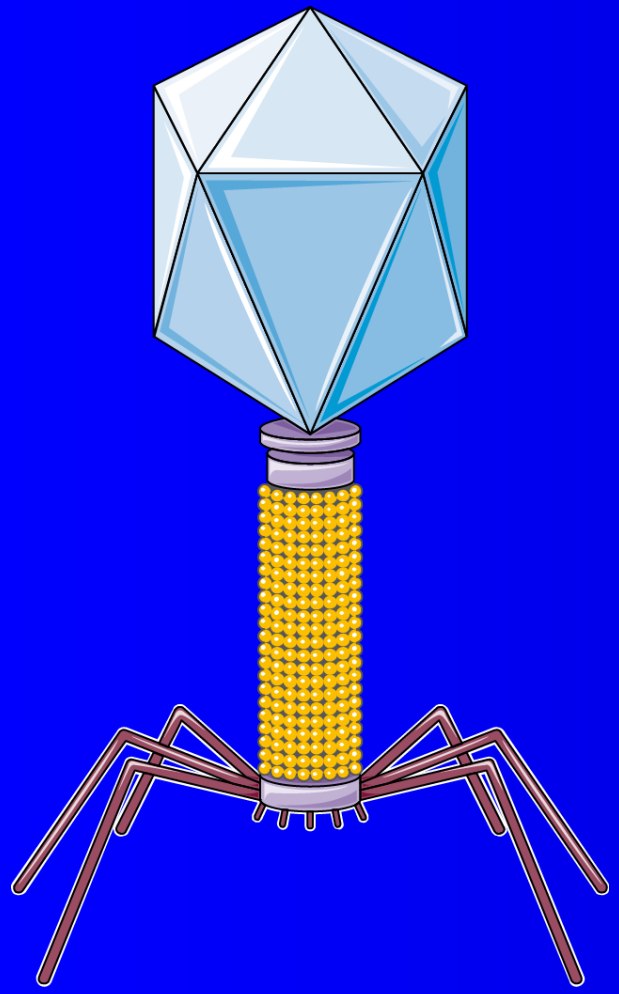
Digestion with DNAses inhibit transformation, suggesting that the DNA caused the transformation.

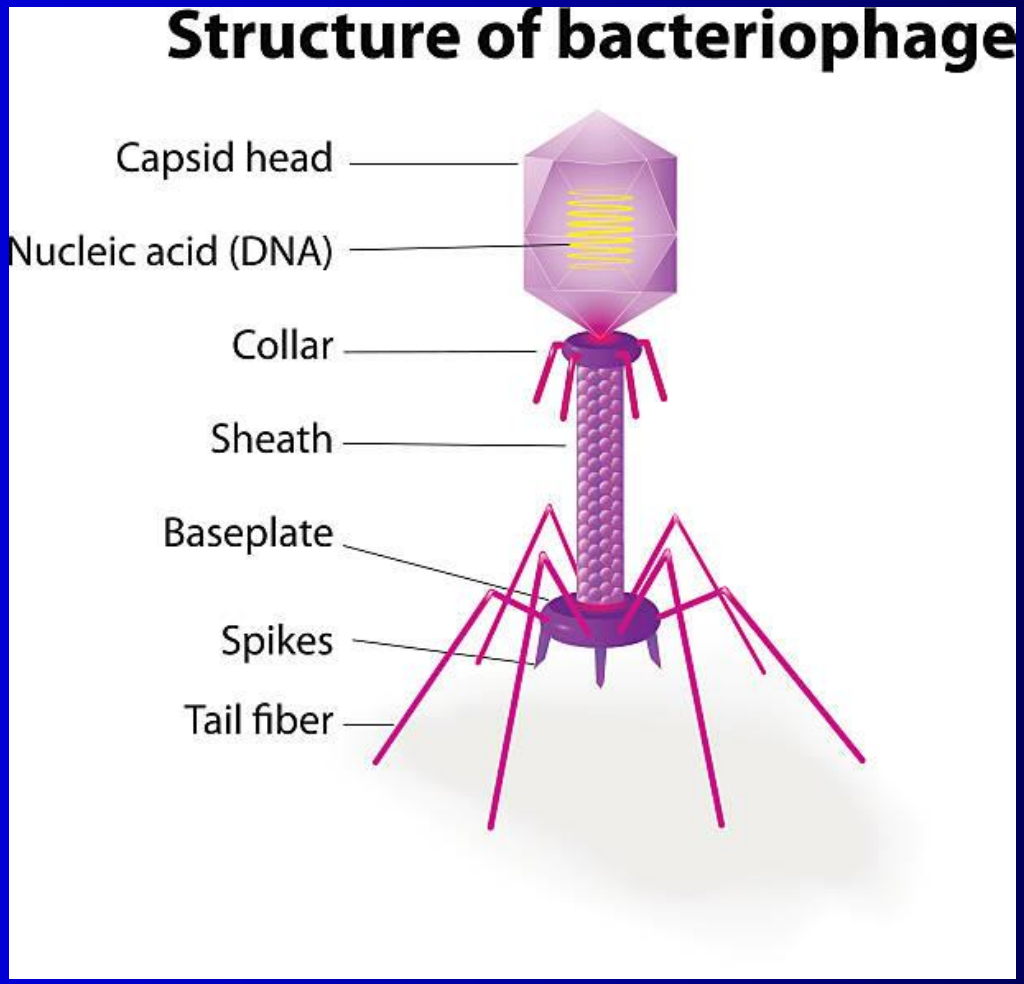
They concluded that DNA is the genetic material.

B. John Ebenezer

Hershey and Chase Experiment

Hershey and Chase Experiment to prove DNA is the Genetic Material





B. John Ebenezer

Hershey and Chase Experiment to prove DNA is the Genetic Material

Bacteriophage
Protein Coat labelled with S^{35}

↓ Infects

E. Coli Bacterium

↓ Blending

Viral Protein Coat Removed from Bacteria

↓ Centrifugation

Viral Particle separated from Bacteria

↓

Radioactive S^{35} Found in Protein Coat
But not found in Bacterial Cell

↓

Protein is not the Genetic Material

Bacteriophage
DNA Labelled with P^{32}

↓ Infects

E. Coli Bacterium

↓ Blending

Viral Protein Coat Removed from Bacteria

↓ Centrifugation

Viral Particle separated from Bacteria

↓

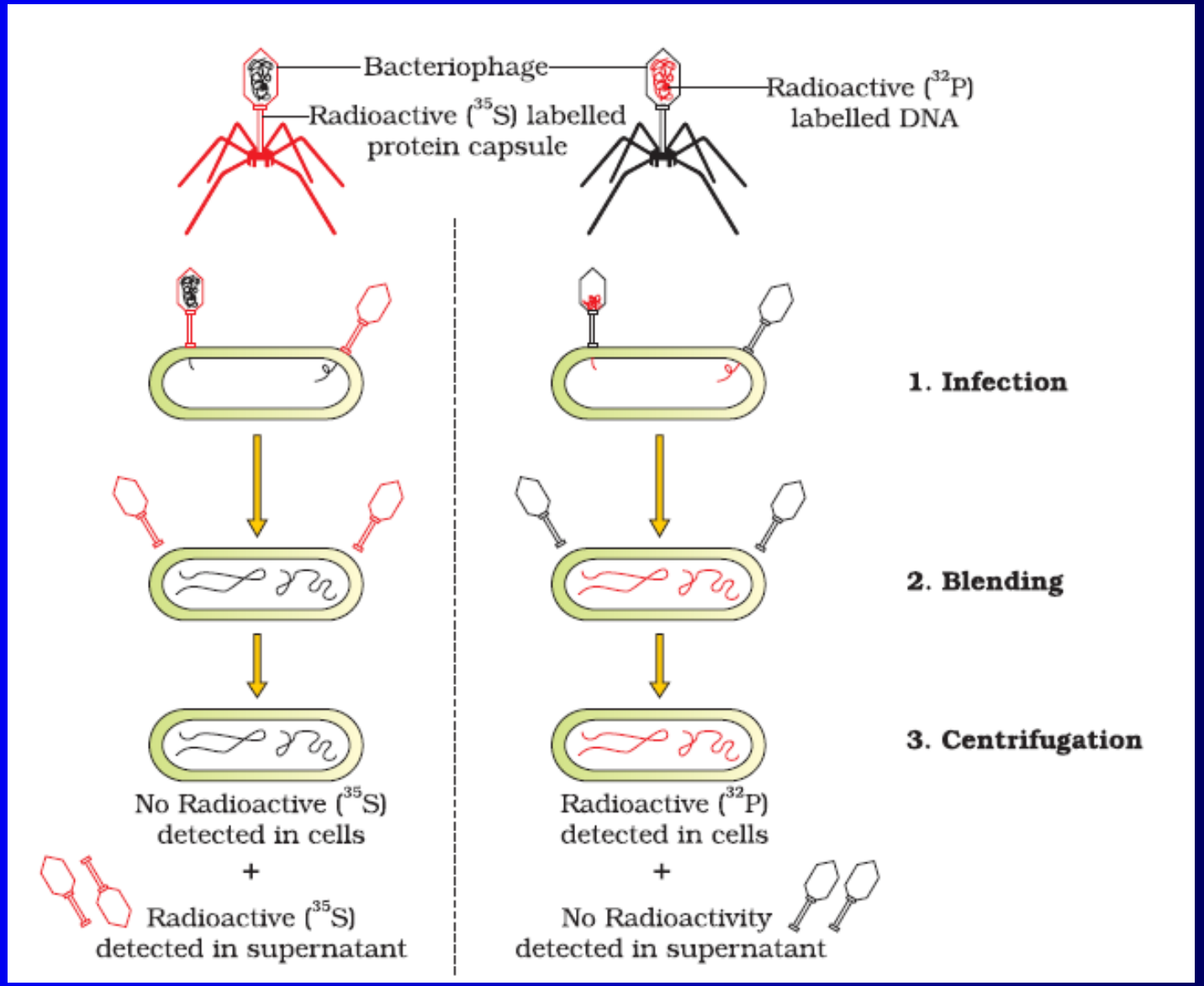
Radioactive P^{32} Found in Bacterial Cell
But not found in Protein Coat

↓

DNA is the Genetic Material

B. John Ebenezer

Hershey and Chase Experiment to prove DNA is the Genetic Material



B. John Ebenezer

Hershey and Chase Experiment to prove DNA is the Genetic Material

The evidence that DNA is the genetic material came from the experiments of **Alfred Hershey** and **Martha Chase** (1952).

They worked with viruses that infect bacteria called bacteriophages.

The bacteriophage attaches to the bacteria and its genetic material then enters the bacterial cell.

The bacterial cell treats the viral genetic material as if it was its own and subsequently manufactures more virus particles.

B. John Ebenezer

Hershey and Chase Experiment to prove DNA is the Genetic Material

Hershey and Chase worked to discover whether it was protein or DNA from the viruses that entered the bacteria.

They grew some viruses on a medium that contained radioactive phosphorus and some others on medium that contained radioactive sulfur.

Viruses grown in the presence of radioactive phosphorus contained radioactive DNA but not radioactive protein because DNA contains phosphorus but protein does not.

B. John Ebenezer

Hershey and Chase Experiment to prove DNA is the Genetic Material

Similarly, viruses grown on radioactive sulfur contained radioactive protein but not radioactive DNA because DNA does not contain sulfur.

Radioactive phages were allowed to attach to *E. coli* bacteria.

Then, as the infection proceeded, the viral coats were removed from the bacteria by agitating them in a blender.

The virus particles were separated from the bacteria by spinning them in a centrifuge.

B. John Ebenezer

Hershey and Chase Experiment to prove DNA is the Genetic Material

Bacteria which was infected with viruses that had radioactive DNA were radioactive, indicating that DNA was the material that passed from the virus to the bacteria.

Bacteria that were infected with viruses that had radioactive proteins were not radioactive.

This indicates that proteins did not enter the bacteria from the viruses. DNA is therefore the genetic material that is passed from virus to bacteria.

B. John Ebenezer

Characteristics of Genetic Material

Characteristics of a Genetic Material

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

(i) It should be able to replicate.

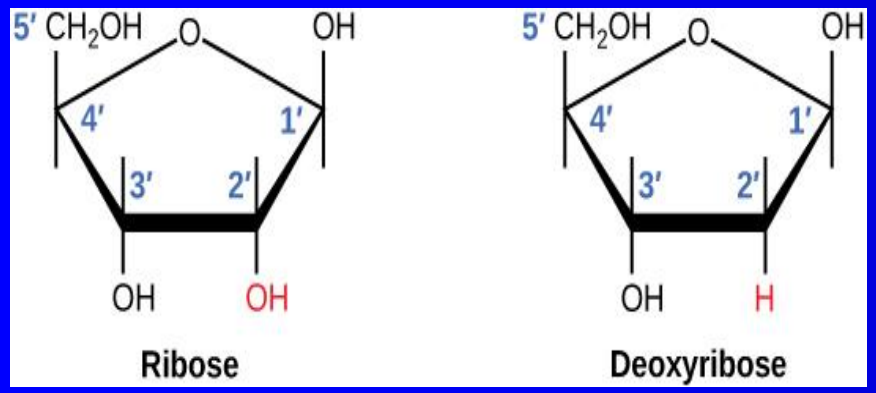
(ii) It should chemically and structurally be stable.

(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'.

B. John Ebenezer

Nucleic Acids



In Deoxyribose sugar there is no Oxygen in **2′** position

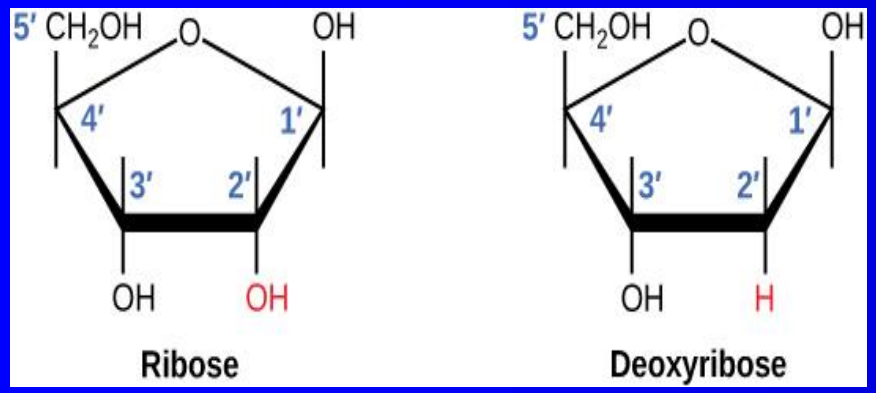
The two strands of DNA being complementary, if separated by heating come together, when appropriate conditions are provided.

Further, **2′-OH group** present at every nucleotide in RNA is a **reactive group** and makes RNA labile (easily altered) and easily degradable.

RNA is also known to be catalytic, hence reactive.

B. John Ebenezer

Nucleic Acids



DNA is chemically less reactive and structurally more stable when compared to RNA.

Therefore, among the two nucleic acids, the DNA is a better genetic material.

B. John Ebenezer

Nucleic Acids

The presence of thymine at the place of uracil also confers additional stability to DNA.

Both DNA and RNA are able to mutate.

RNA being unstable, mutate at a faster rate.

Viruses having RNA genome and having shorter life span mutate and evolve faster.

B. John Ebenezer

Nucleic Acids

RNA can directly code for the synthesis of proteins, hence can easily express the characters.

DNA is dependent on RNA for synthesis of proteins.

The protein synthesising machinery has evolved around RNA.

Both RNA and DNA can function as genetic material, but DNA being more stable is preferred for storage of genetic information.

For the transmission of genetic information, RNA is better.

B. John Ebenezer

DNA is a better genetic material.

DNA	RNA
DNA is chemically less reactive and structurally more stable when compared to RNA.	The 2′-OH group present at every nucleotide in RNA is a reactive group and makes RNA labile (easily altered) and easily degradable.
The presence of thymine at the place of uracil also confers additional stability to DNA.	RNA is also known to be catalytic, hence reactive.
DNA being stable mutate at a slower rate.	RNA being unstable, mutate at a faster rate. Viruses having RNA genome and having shorter life span mutate and evolve faster.

B. John Ebenezer

DNA is a better genetic material.

DNA	RNA
DNA is dependent on RNA for synthesis of proteins.	RNA can directly code for the synthesis of proteins, hence can easily express the characters.
	The protein synthesising machinery has evolved around RNA.
DNA being more stable is preferred for storage of genetic information.	RNA being less stable is not preferred for storage of genetic information.

B. John Ebenezer

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

But, RNA being a catalyst was reactive and hence unstable.

B. John Ebenezer

RNA World

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.

B. John Ebenezer

DNA Replication

DNA Replication

Watson and Crick proposed a scheme for replication of DNA in 1953.

The scheme suggested that the two strands would separate and act as a template for the synthesis of new complementary strands.

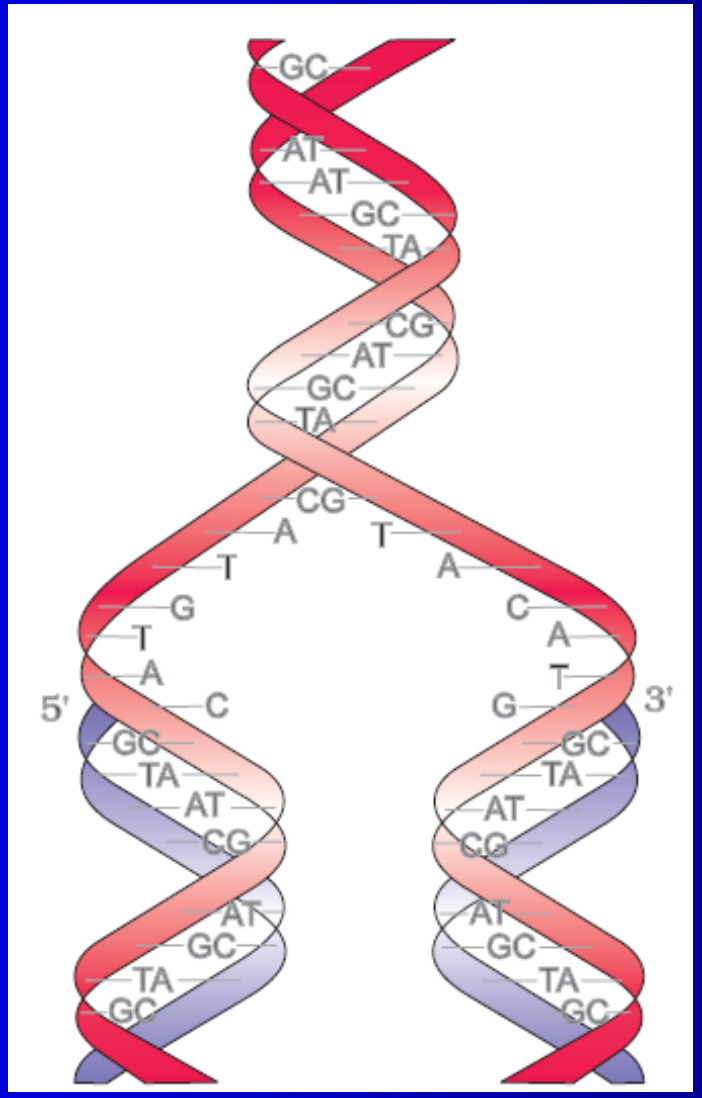
After the completion of replication, each DNA molecule would have one parental and one newly synthesised strand.

This scheme was termed as **semiconservative** DNA replication

B. John Ebenezer

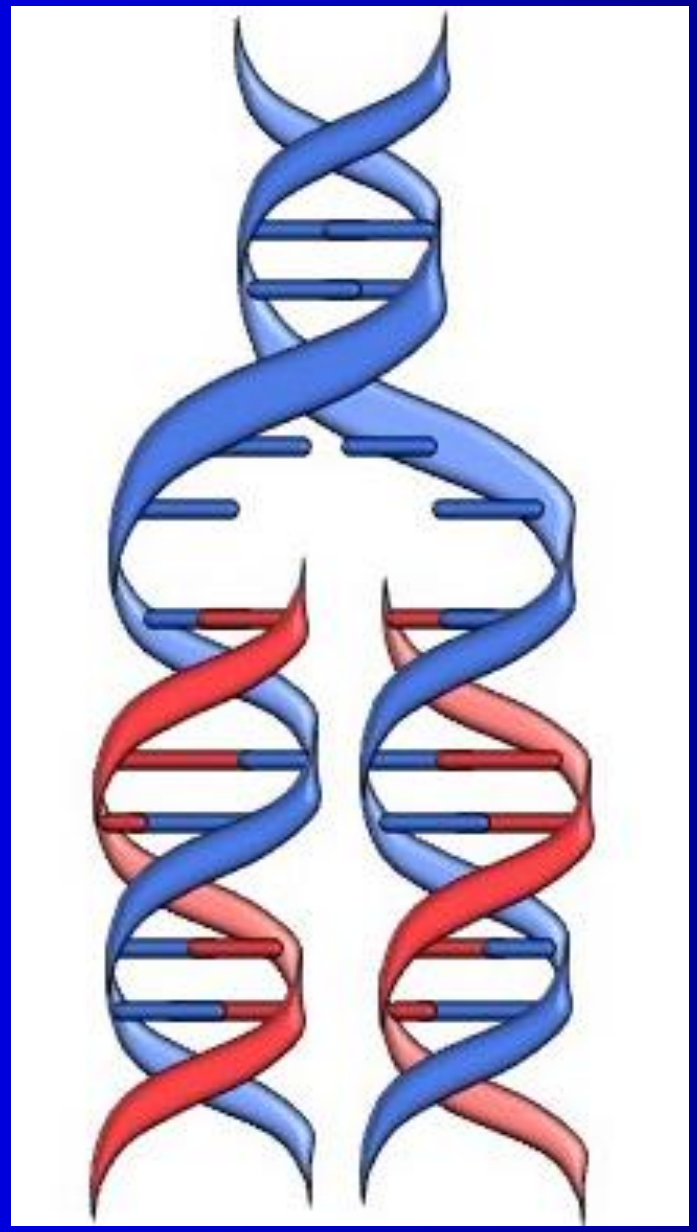
DNA Replication

Watson-Crick model for semiconservative DNA replication



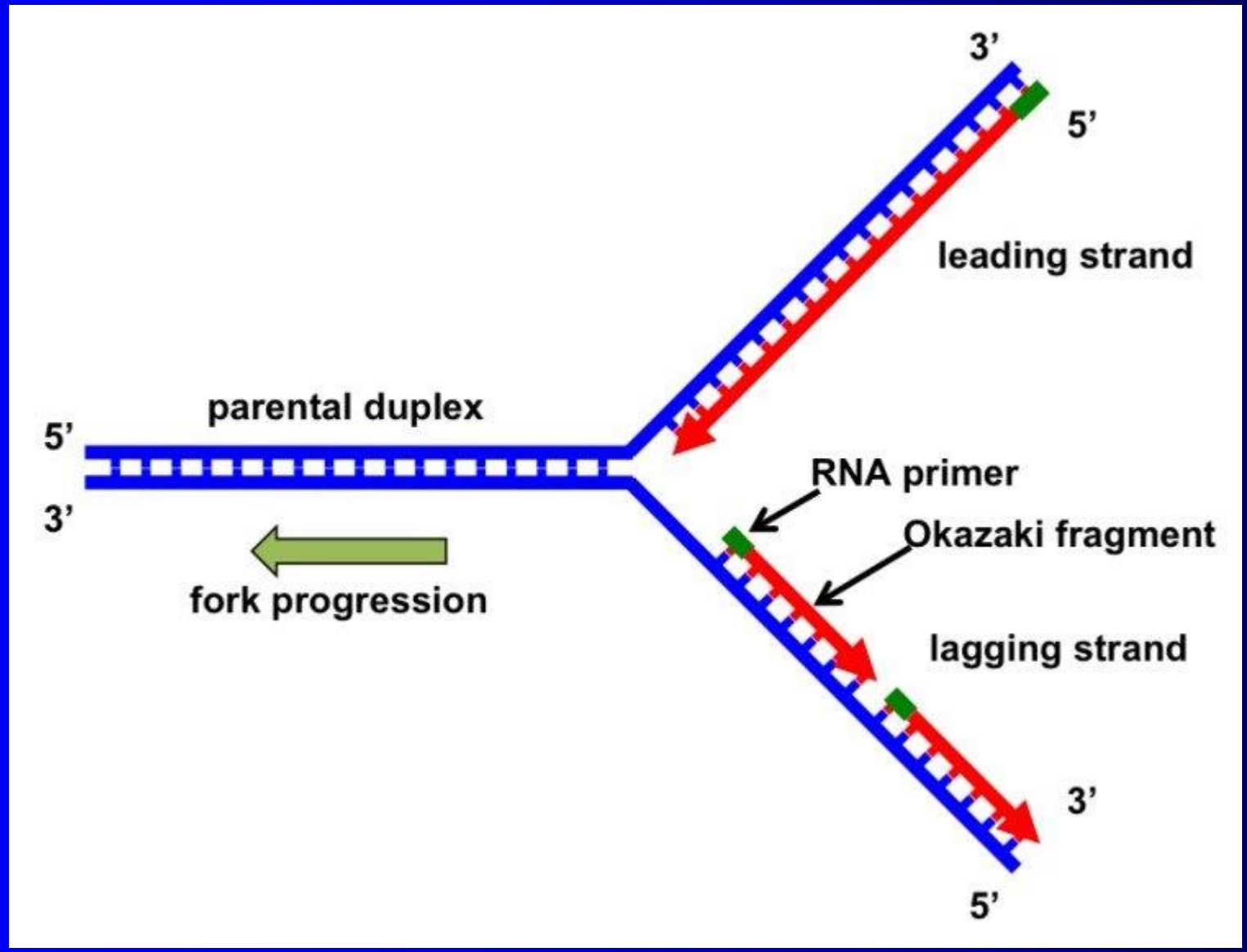
B. John Ebenezer

DNA Replication



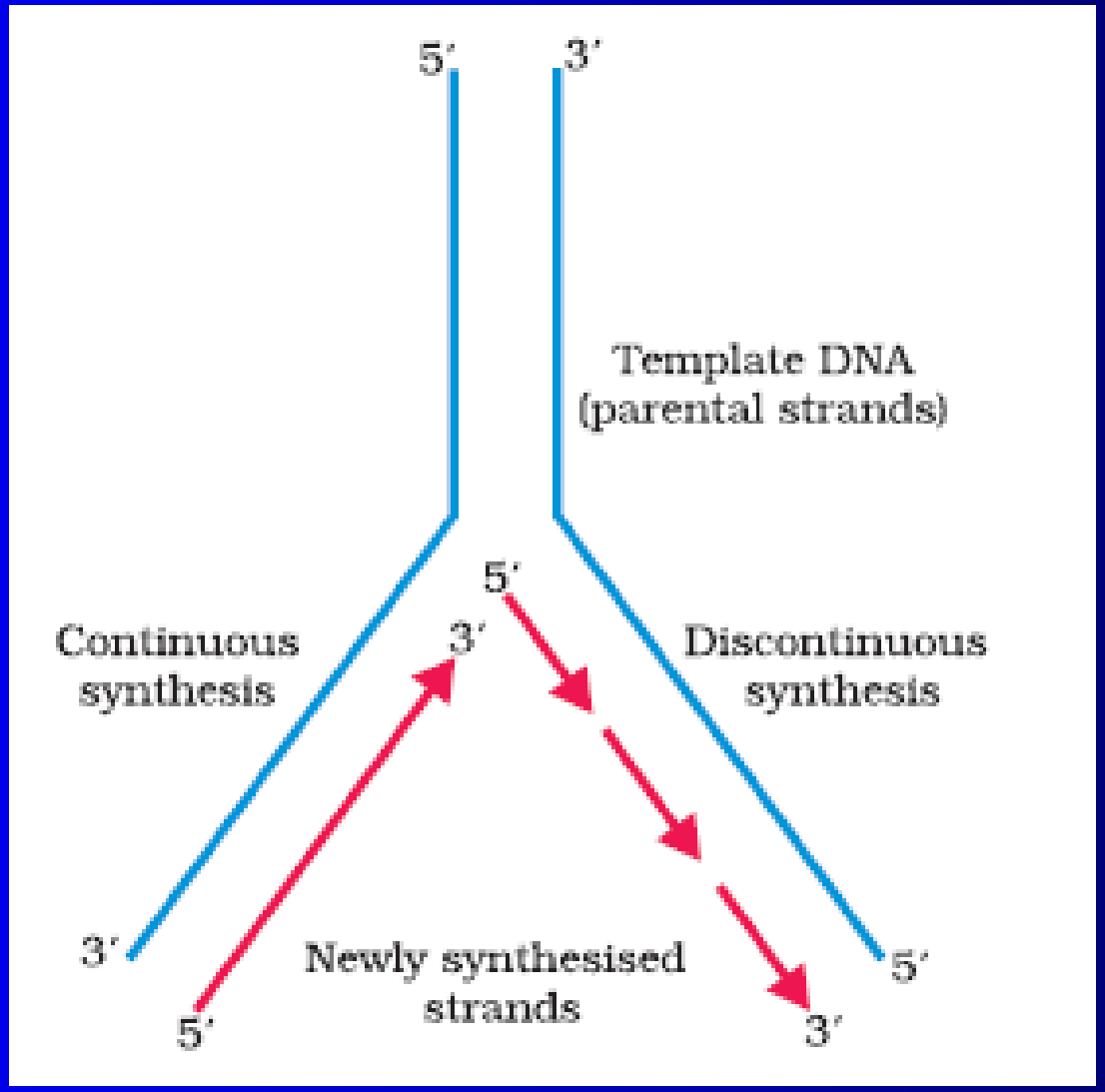
B. John Ebenezer

Replication Fork



B. John Ebenezer

Replication Fork



B. John Ebenezer

Requirements of DNA Replication

In living cells, such as *E. coli*, the process of replication requires a set of catalysts (enzymes).

The main enzyme is referred to as DNA-dependent **DNA polymerase**, since it uses a DNA template to catalyse the polymerisation of deoxynucleotides.

These enzymes are highly efficient enzymes as they have to catalyse polymerisation of a large number of nucleotides in a very short time.

B. John Ebenezer

Requirements of DNA Replication

Any mistake during replication would result into mutations.

Replication is a very expensive process.

Deoxyribonucleoside triphosphates serve dual purposes.

In addition to **acting as substrates**, they **provide energy** for polymerisation reaction (the two terminal phosphates in a deoxynucleoside triphosphates are high-energy phosphates, same as in case of ATP).

B. John Ebenezer

Requirements of DNA Replication

In addition to DNA-dependent DNA polymerases, many additional enzymes are required to complete the process of replication with high degree of accuracy.

For long DNA molecules, since the two strands of DNA cannot be separated in its entire length (due to very high energy requirement), the replication occurs within a small opening of the DNA helix, referred to as **replication fork**.

B. John Ebenezer

Rate of Polymerisation

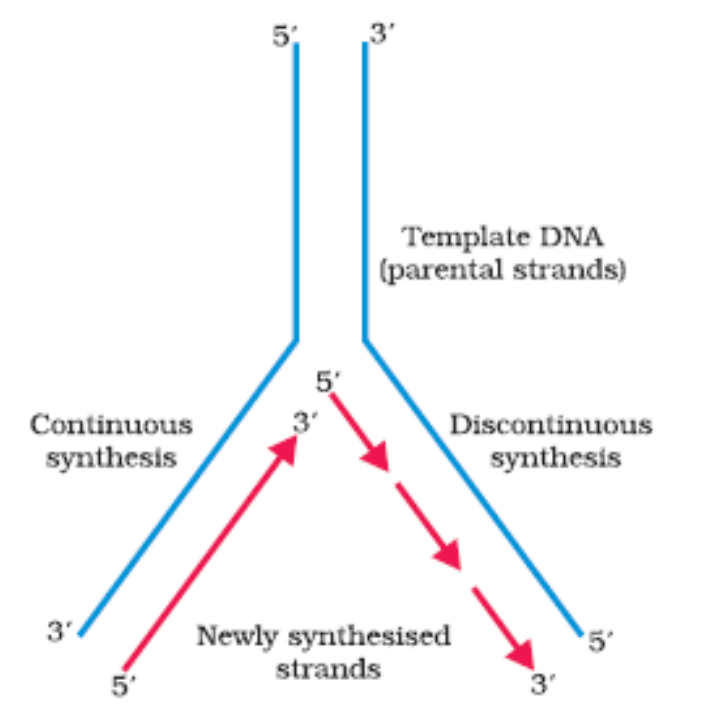
E. coli has only 4.6 $\times 10^6$ bp (compare it with human whose diploid content is 6.6 $\times$ 10^9 bp), completes the process of replication within 38 minutes.

It means the average rate of polymerisation is approximately 2000 bp per second.

Not only these polymerases have to be fast, but they also have to catalyse the reaction with high degree of accuracy.

B. John Ebenezer

DNA Replication



The DNA-dependent DNA polymerases catalyse polymerisation only in one direction, that is 5‘ to 3′ .

This creates some additional complications at the replicating fork.

On Leading strand (the template strand with polarity 3‘ to 5′), the replication is **continuous**.

On Lagging strand (the template strand with polarity 5‘ to 3′), it is **discontinuous**.

The discontinuously synthesised fragments are later joined by the enzyme **DNA ligase**

B. John Ebenezer

DNA Replication

The DNA polymerases on their own cannot initiate the process of replication.

Also the replication does not initiate randomly at any place in DNA.

There is a definite region in *E. coli* DNA where the replication originates. Such regions are termed as **origin of replication**.

It is because of the requirement of the origin of replication that a piece of DNA if needed to be propagated during recombinant DNA procedures, requires a vector.

The vectors provide the origin of replication.

B. John Ebenezer

DNA Replication

In eukaryotes, the replication of DNA takes place at S-phase of the cell-cycle.

The replication of DNA and cell division cycle should be highly coordinated.

A failure in cell division after DNA replication results into polyploidy (a chromosomal anomaly).

B. John Ebenezer

Unwinding of DNA Strands and Formation of Replication Fork

↓

DNA Polymerase Initiates Polymerisation in Leading Strand

↓

Nitrogen Bases are added from 5' to 3' direction continuously

↓

Polimerisation starts with Okazaki Fragments in Lagging strand

↓

Nitrogen Bases are added discontinuously, later joined by DNA Ligase

↓

Accuracy of polimerisation is maintained by DNA Polymerase

B. John Ebenezer

Meselson and Stahl's Experiment

Meselson and Stahl's Experiment

Meselson and Stahl performed the experiment in 1958 on *E. coli* to prove that DNA replication is semiconservative.

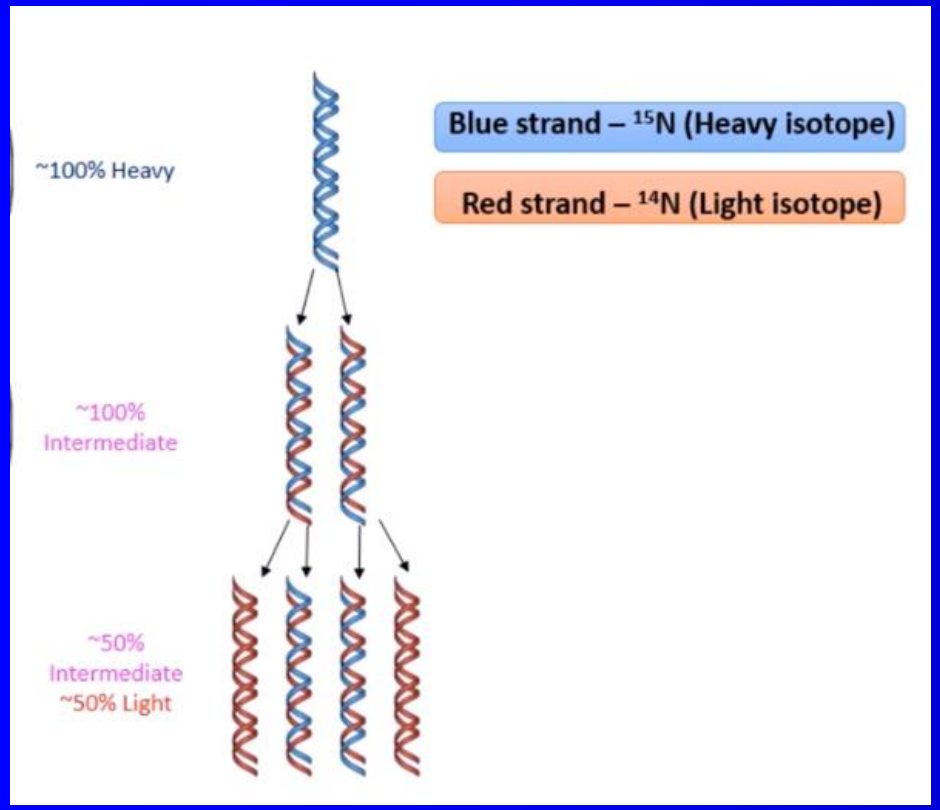
E. coli was grown in $^{15}NH_4CI$ for many generations.

N^{15} was incorporated into newly synthesised DNA.

This heavy DNA could be differentiated from normal DNA by centrifugation in **cesium chloride (CsCl)** density gradient.

B. John Ebenezer

Meselson and Stahl's Experiment



Then they transferred these *E.coli* into medium with normal $^{14}NH_4Cl$.

First generation

After 20 minutes, it was found that all the **DNA molecules of daughter cells were hybrid-First generation.**

Second generation

After 40 minutes, it was found that **50% DNA molecules were hybrid** and **50% were normal-Second generation.**

B. John Ebenezer

Meselson and Stahl's Experiment

DNA replication is semiconservative

E. coli was grown in $^{15}NH_4Cl$ for many generations

↓

N^{15} was incorporated into newly synthesized DNA

↓

This heavy DNA could be differentiated from normal DNA
by centrifugation in **cesium chloride** (CsCl) density gradient

↓

Then they transferred these *E.coli* into medium with normal $^{14}NH_4Cl$

↓

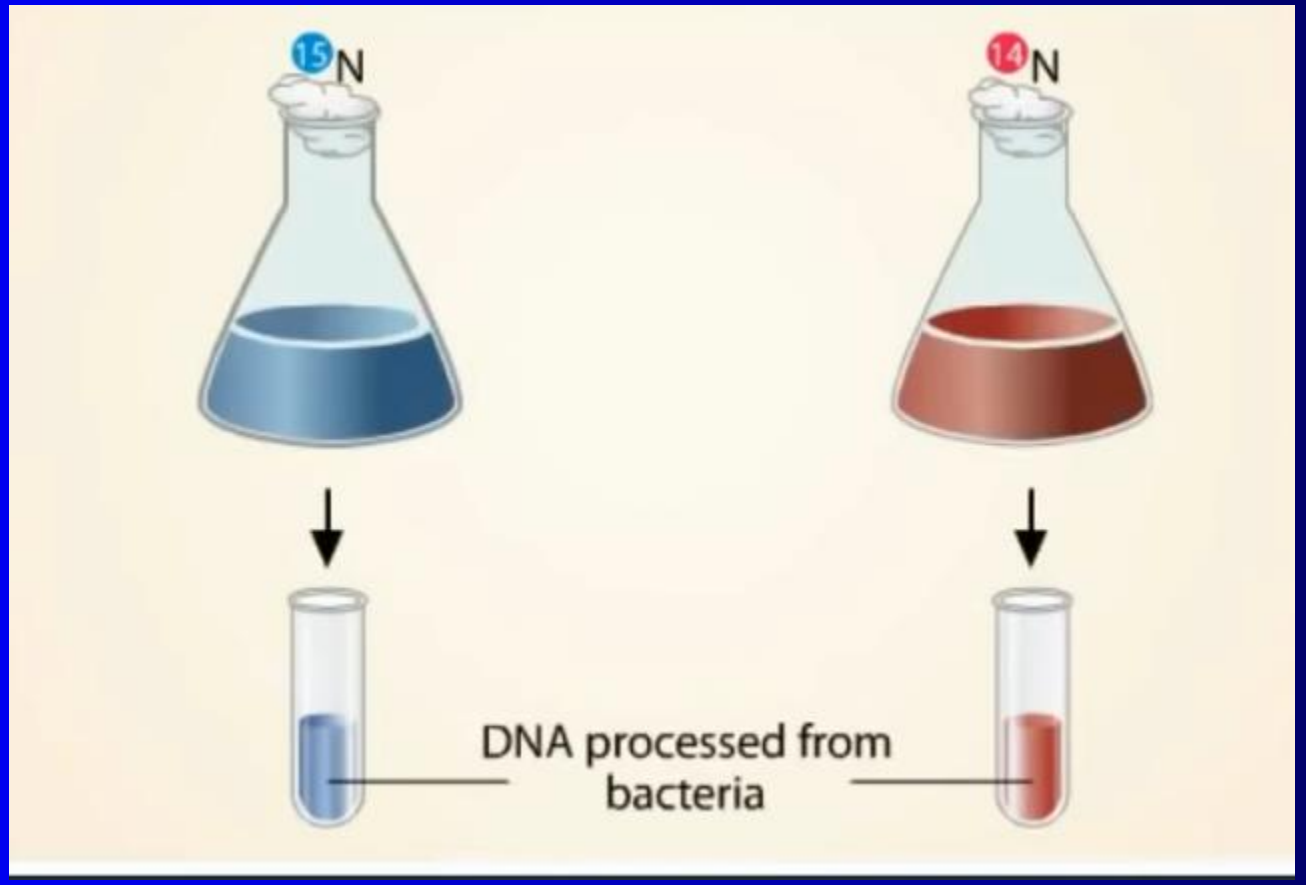
After 20 minutes, it was found that
All the DNA molecules of daughter cells were **hybrid - First generation**

↓

After 40 minutes, it was found that
50% DNA molecules were hybrid and
50% were normal - Second generation

B. John Ebenezer

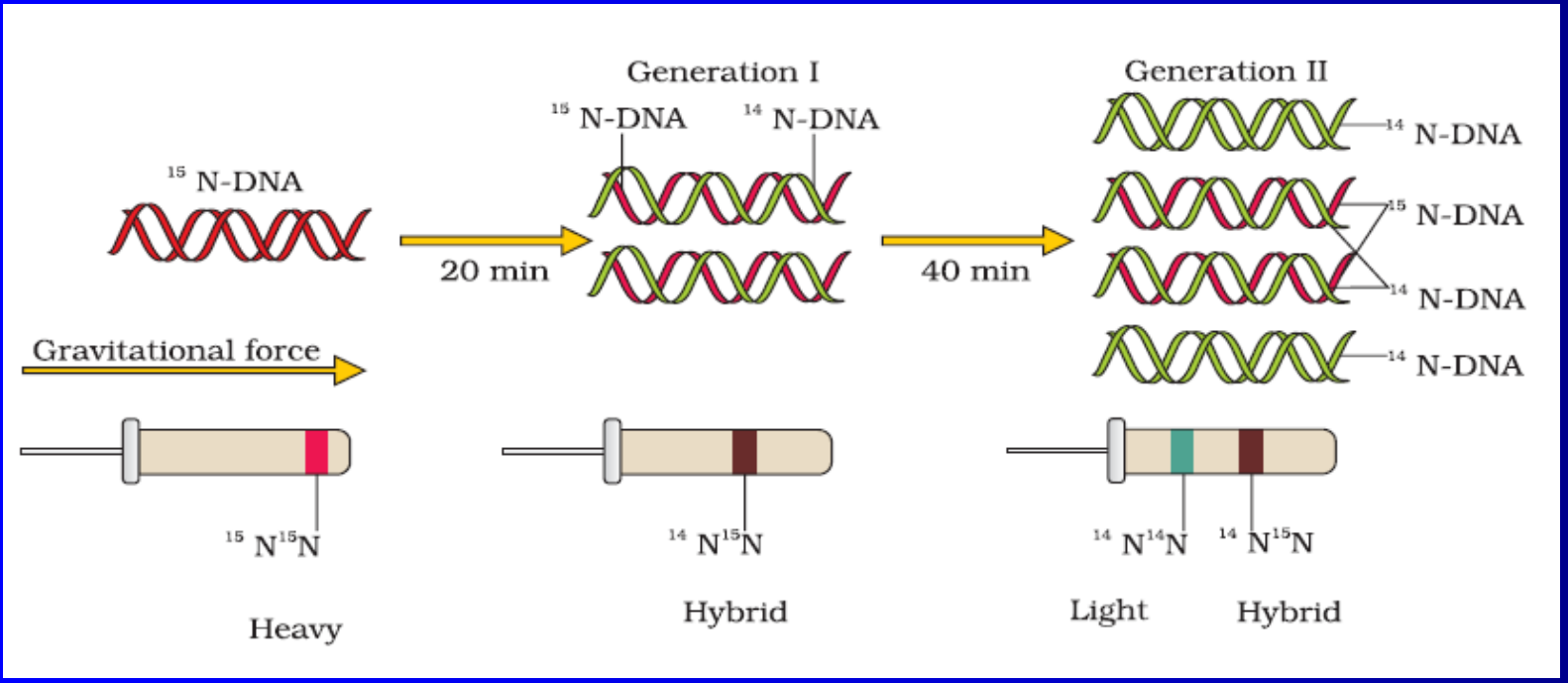
Meselson and Sahl's Experiment



B. John Ebenezer

Meselson and Sahl's Experiment

Separation of DNA by Centrifugation



Red strand ^{15}N Heavy isotope

Green strand ^{14}N Light isotope

Red & Green strands ^{15}N ^{14}N Hybrid

B. John Ebenezer

Central Dogma

DNA acts as the director of protein synthesis. It gives command for protein synthesis in the form of triplet codes.

DNA is located in the nucleus but protein synthesis occurs in the ribosome present in cytoplasm.

So the command of DNA for protein synthesis is taken to the cytoplasm by an RNA called messenger RNA or mRNA.

B. John Ebenezer

Transfer RNAs or tRNAs decode the triplet codes found in the mRNA using their anticodon sites and transfer specific aminoacids to the ribosome.

Ribosome present in the cytoplasm acts as the factory of protein synthesis. It is made up of rRNAs (ribosomal RNAs) and proteins.

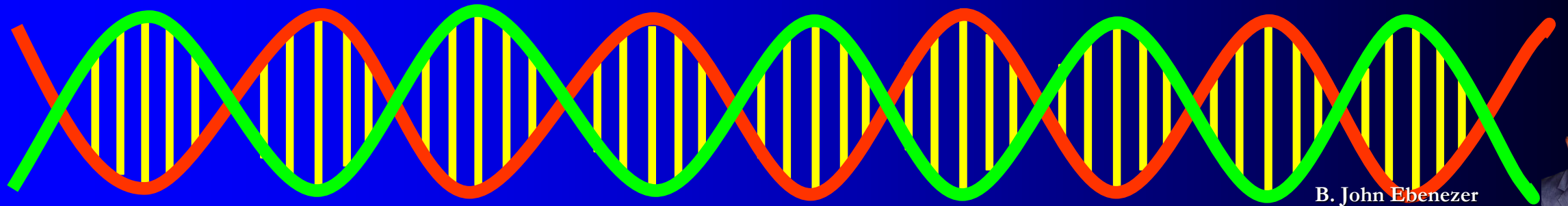
Ribosome consists of a smaller sub unit and a larger sub unit. The smaller sub unit reads the mRNA and the larger sub unit joins the aminoacids to form a polypeptide.

B. John Ebenezer

Central Dogma

Central Dogma states the transfer of genetic information from DNA to mRNA and from mRNA to protein synthesis.

DNA —Transcription→ mRNA —Translation→ Protein



B. John Ebenezer

Central Dogma

Central Dogma includes two processes namely transcription and translation.

Transcription

The synthesis of mRNA from DNA is called transcription.

This takes place in the nucleus.

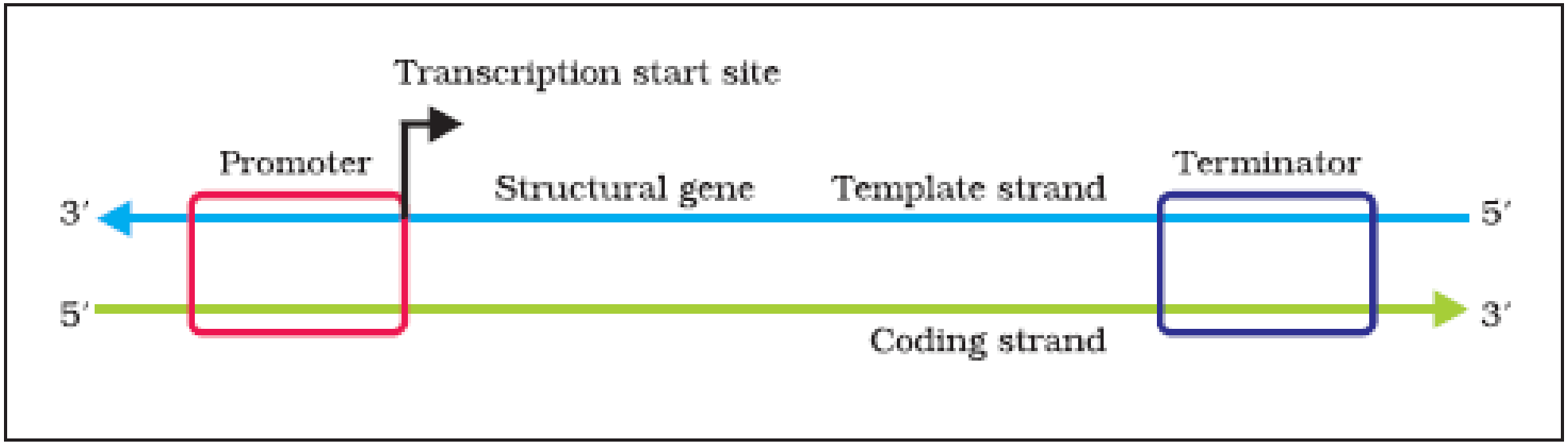
Translation

The synthesis of protein from mRNA is called translation.

This takes place in the cytoplasm.

B. John Ebenezer

Schematic structure of a transcription unit



B. John Ebenezer

DNA is a Double Stranded Helical Structure

DNA strand with 3′ to 5′ polarity is called Template strand

DNA strand with 5′ to 3′ polarity is called Coding strand

ATG	CAT	GCA	TGC	ATG	CAT	GCA	TGC

3′ Template Strand (DNA) 5′

TAC	GTA	CGT	ACG	TAC	GTA	CGT	ACG

5′ Coding Strand (DNA) 3′

B. John Ebenezer

Central Dogma

The newly synthesized mRNA is complementary to the Template strand

UAC	GUA	CGU	ACG	UAC	GUA	CGU	ACG

5′ mRNA 3′

ATG	CAT	GCA	TGC	ATG	CAT	GCA	TGC

3′ Template Strand (DNA) 5′

TAC	GTA	CGT	ACG	TAC	GTA	CGT	ACG

5′ Coding Strand (DNA) 3′

B. John Ebenezer

Central Dogma

But the newly synthesized mRNA is similar to the coding strand except thymine is replaced by uracil

UAC	GUA	CGU	ACG	UAC	GUA	CGU	ACG

5′ mRNA 3′

TAC	GTA	CGT	ACG	TAC	GTA	CGT	ACG

5′ Coding Strand (DNA) 3′

B. John Ebenezer

Only one of the DNA strands act as template

- During transcription, only one of the DNA strands acts as template for mRNA synthesis for the following reasons.

- If both the strands act as template, they would code for two different mRNA molecules and two different proteins would be formed. Hence the genetic information transfer would become complicated.

B. John Ebenezer

Central Dogma

- If two mRNA molecules were produced, they would be complementary to each other and form a double stranded mRNA.
- This would prevent the translation of mRNA into protein.

B. John Ebenezer

Transcription in Prokaryotes

B. John Ebenezer

Transcription in Prokaryotes

The requirements of Transcription are:

- Promoter
- Structural Gene
- Terminator
- Initiation factor
- Termination factor
- DNA dependent RNA Polymerase Enzyme

B. John Ebenezer

Promoter

- It is a sequence of DNA which provides binding site for RNA Polymerase.
- It is present at the 5′ end of DNA. It is also known as upstream end.

Terminator

- It is a sequence of DNA which provides binding site for Termination factor.
- It is present at the 3′ end of DNA. It is also called downstream end.

B. John Ebenezer

Structural Gene

- The gene to be transcribed is called structural gene

Cistron

- A segment of DNA coding for a polypeptide (protein) is called cistron.

B. John Ebenezer

Polycistronic

The structural gene in prokaryotes is polycistronic, because a single mRNA codes for more than one polypeptide (protein).

Monocistronic

The structural gene in eukaryotes is monocistronic, because a single mRNA codes for only one polypeptide (protein).

B. John Ebenezer

Initiation Factor

- The Initiation factor of transcription is Sigma factor (σ)

Termination Factor

- The Termination factor of transcription is Rho factor (ρ)

B. John Ebenezer

Initiation

- The DNA dependent RNA polymerase binds to the promoter along with the sigma factor and begins transcription. This is known as initiation.

- A single RNA polymerase catalyzes the transcription of all types of RNAs in bacteria.

B. John Ebenezer

Elongation

- RNA polymerase keeps adding nucleotides (polymerization) following the rule of complementarity.
- The length of mRNA keeps increasing.

B. John Ebenezer

Termination

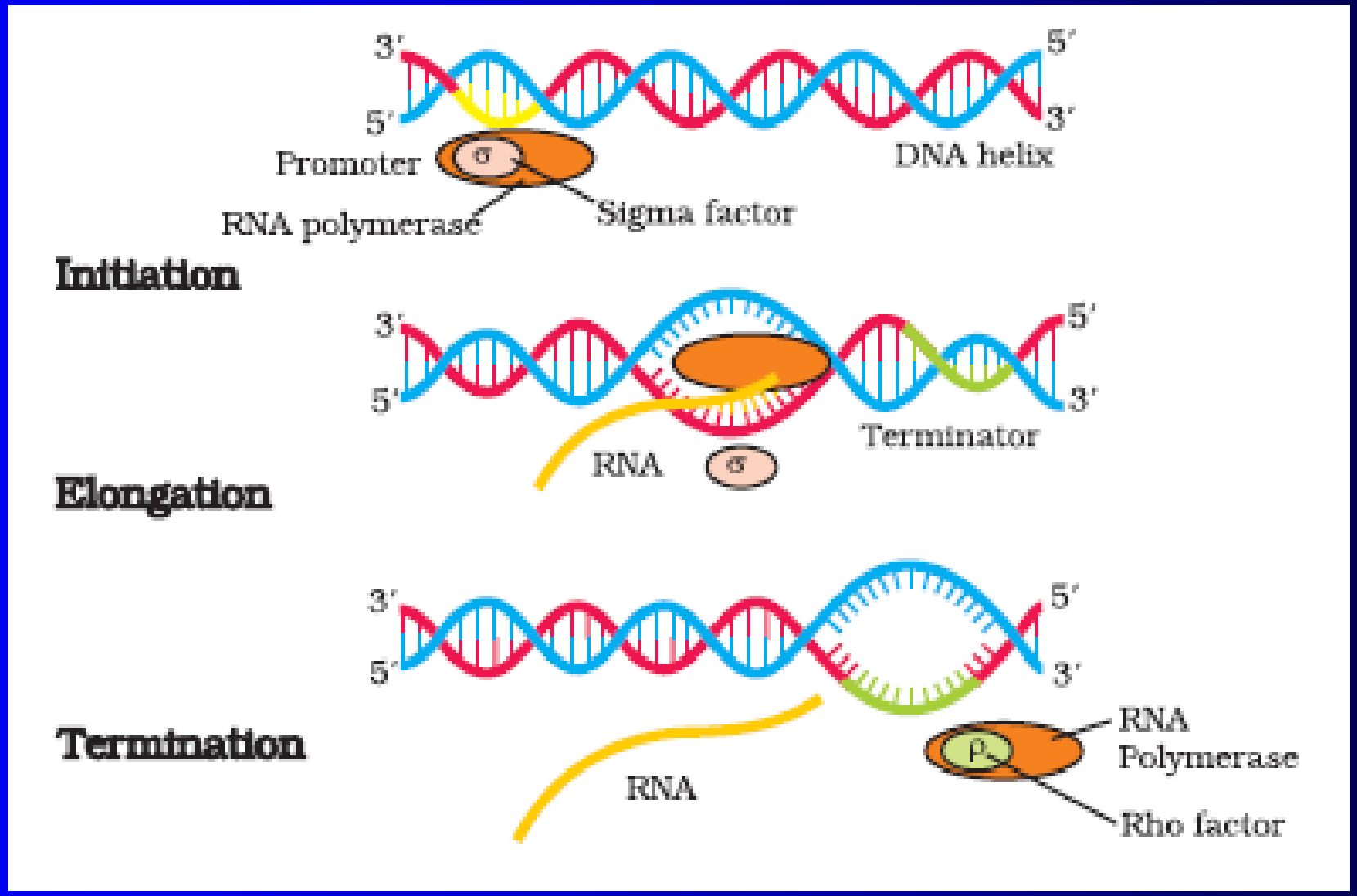
- When polymerase reaches the terminator region, the Rho factor combines with the polymerase and terminates transcription.
- The RNA polymerase and the mRNA get separated.

B. John Ebenezer

- The RNA polymerase is only capable of catalyzing the process of elongation.

- But it starts transcription by joining with the **initiation-factor** (σ) and stops transcription by joining with **termination-factor** (ρ).

B. John Ebenezer



B. John Ebenezer

- The prokaryotic mRNA does not require any processing to become active.
- Since transcription and translation take place in the same compartment (there is no separation of cytosol and nucleus in bacteria), many times translation begins much before the mRNA is fully transcribed.
- The transcription and translation can be coupled in bacteria.

B. John Ebenezer

Transcription in Eukaryotes

B. John Ebenezer

Transcription in Eukaryotes

- Transcription in eukaryotes is a complex process because it requires three different polymerases to synthesize different RNAs.
- Sigma and Rho factors are not involved in eukaryotic transcription.
- Several transcription factors like TATA binding protein (TBP), Transcription Factor IID (TFIID), Transcription Factor IIA (TFIIA) are involved in initiation of Eukaryotic Transcription.

B. John Ebenezer

- Moreover the mRNA synthesized as a result of transcription is immature known as heterogeneous nuclear RNA (hnRNA) because it contains non coding introns and coding exons.

- The non coding introns have to be removed and coding exons have to be joined together.

- It also requires capping and tailing.

B. John Ebenezer

RNA polymerase

The three types of RNA polymerase enzymes are;

- **RNA polymerase I** : It is used for the synthesis of **rRNA**.
- **RNA polymerase II** : It is used for the synthesis of precursor of **mRNA** or hnRNA.
- **RNA polymerase III** : It is used for the synthesis of **tRNA**, 5s rRNA, and snRNA (small nuclear RNAs).

B. John Ebenezer

Processing of hnRNA

B. John Ebenezer

Steps in processing of hnRNA

- The processing of hnRNA or immature mRNA occurs in three steps namely Splicing, Capping and Tailing.

- **Splicing**: It is a process by which non coding introns are removed from coding exons with the help of a protein complex called Spliceosome.

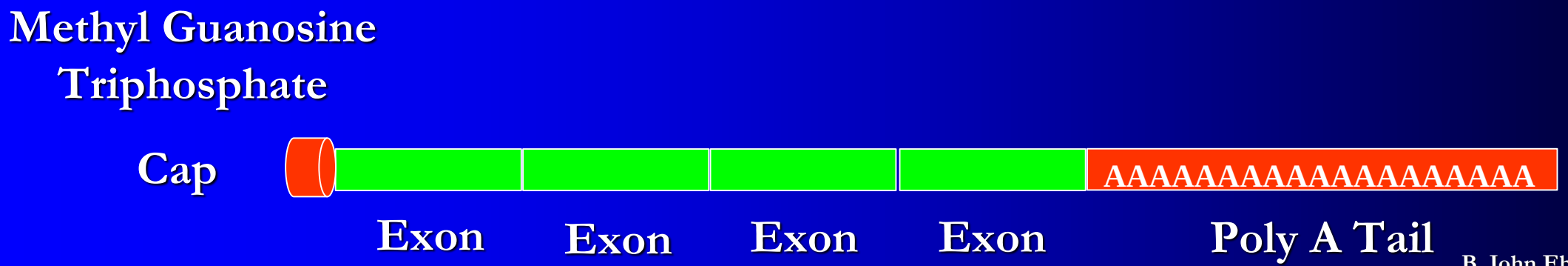
Exon Intron Exon Intron Exon Intron Exon Intron

B. John Ebenezer

Steps in processing of hnRNA

Capping: It is a process of adding a nucleotide **Methyl Guanosine Triphosphate** at the 5′ end of hnRNA.

Tailing: It is a process of adding poly adenine nucleotides **(Adenylate residues)** at the 3′ end of hnRNA. About 200-300 **adenine nucleotides** are added.



B. John Ebenezer

Why is capping and tailing required?

- Transcription occurs in nucleus. After processing, the mRNA has to be transported to cytoplasm for translation.

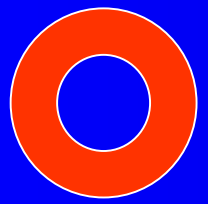
- But there are certain enemies in the cytoplasm to attack the mRNA. They are exonucleases which can remove the nucleotides either from 5′ end or from 3′ end of mRNA.

- So capping and tailing protects mRNA from being attacked by the exonucleases.

B. John Ebenezer

What is Spliceosome?

- It is a ribonucleoprotein complex present in nucleus which removes introns from hnRNA and joins exons to form an active or matured mRNA.



Exon Intron Exon Intron Exon Intron Exon Intron

B. John Ebenezer

Why is Splicing required?

- The mRNA produced in prokaryotes is a matured mRNA. It does not require any processing. It can directly get involved in protein synthesis.

- But the mRNA produced in eukaryotes is not a matured mRNA which is known as hnRNA. It consists of non coding introns and coding exons.

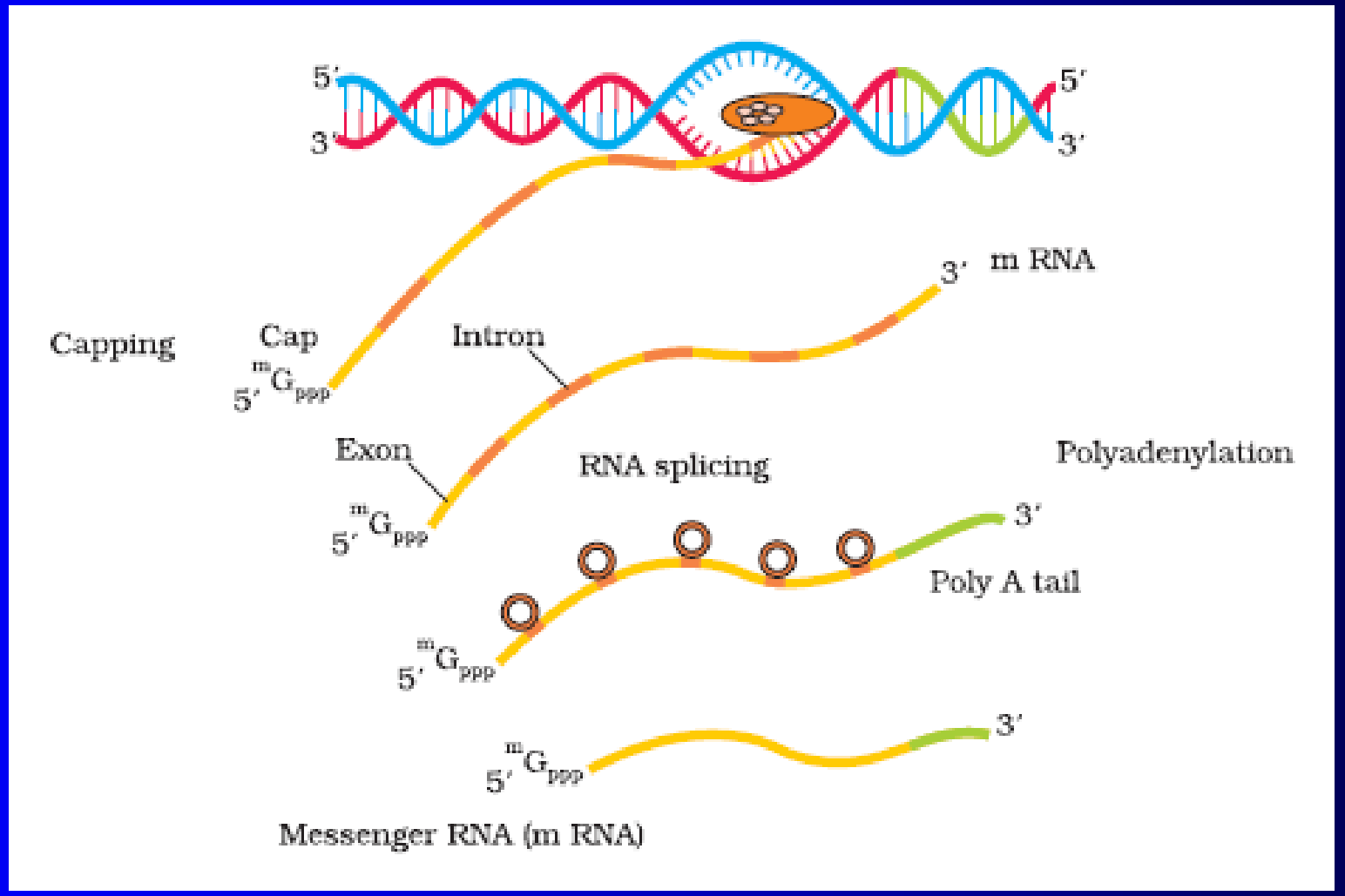
B. John Ebenezer

Why is Splicing required?

- Therefore non coding introns have to be removed from hnRNA and coding exons have to be joined together to form a matured or active mRNA.

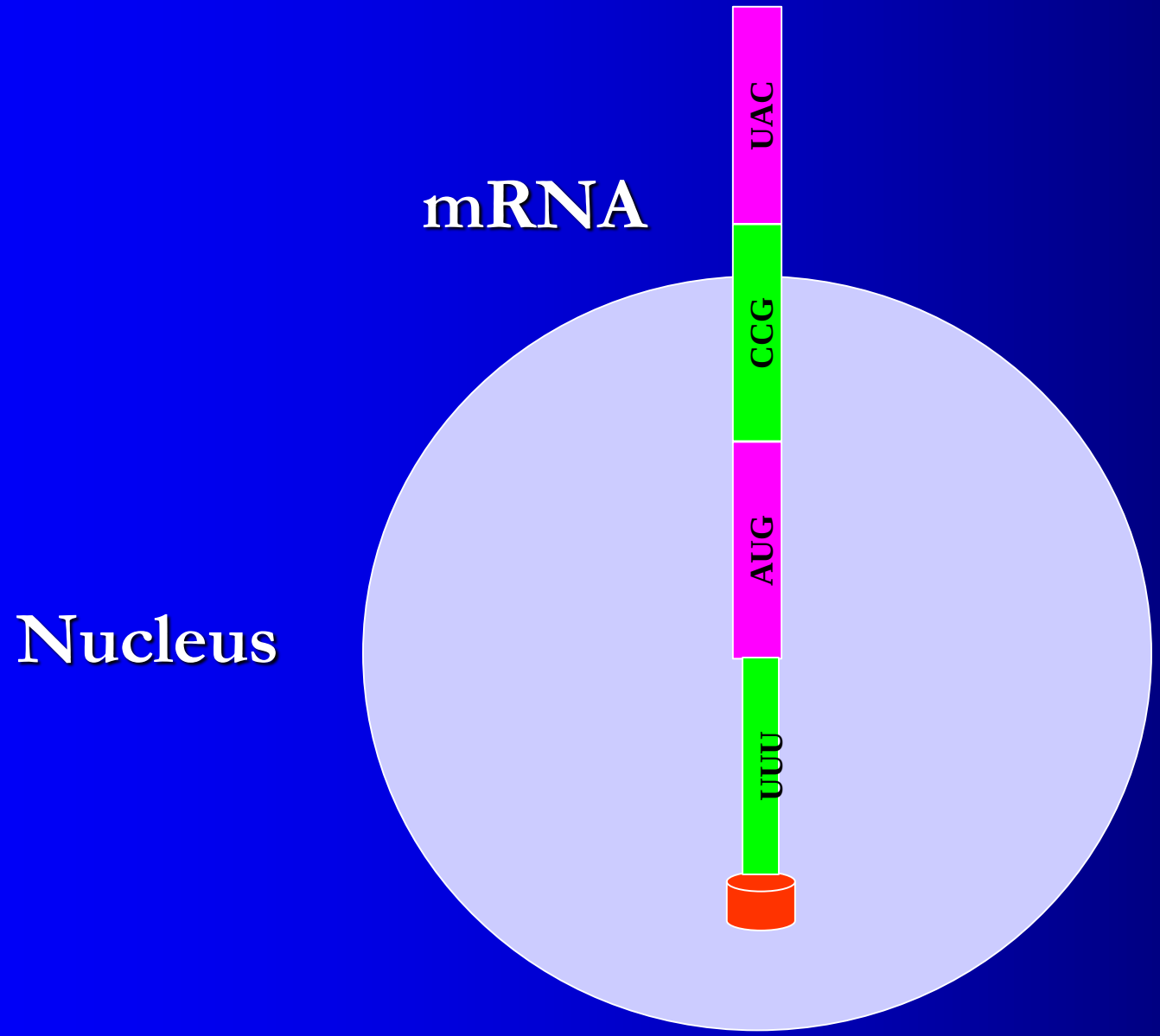
B. John Ebenezer

Process of Transcription in Eukaryotes

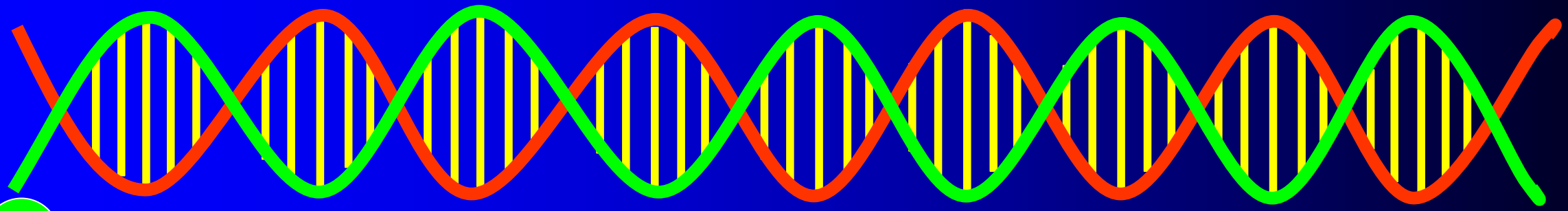


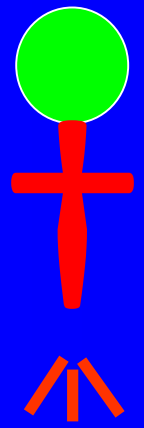
B. John Ebenezer

Transport of mRNA from nucleus to the cytoplasm



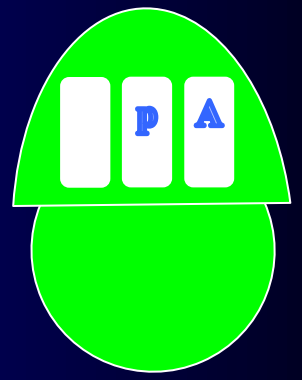
B. John Ebenezer

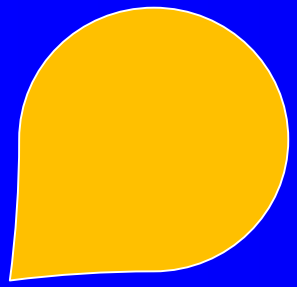


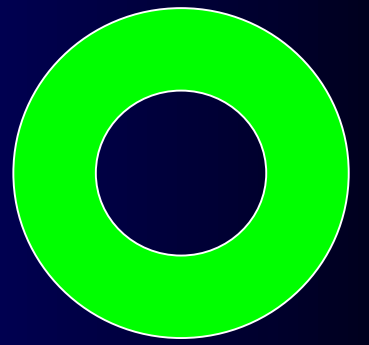


UAC

Translation







5′

3′

UUU	AUG	CCG	UAC	GAU	CCG	UAC	GAU	AUA	ACA	CAG	UAG	UGA	UA

B. John Ebenezer

Translation or Protein Synthesis

The synthesis of protein (polypeptide) from mRNA is called translation.

The order and sequence of amino acids are defined by the sequence of triplet codes in the mRNA.

The amino acids are joined together by peptide bonds.

Peptide bonds are formed between amino group of one amino acid and carboxyl group of next amino acid.

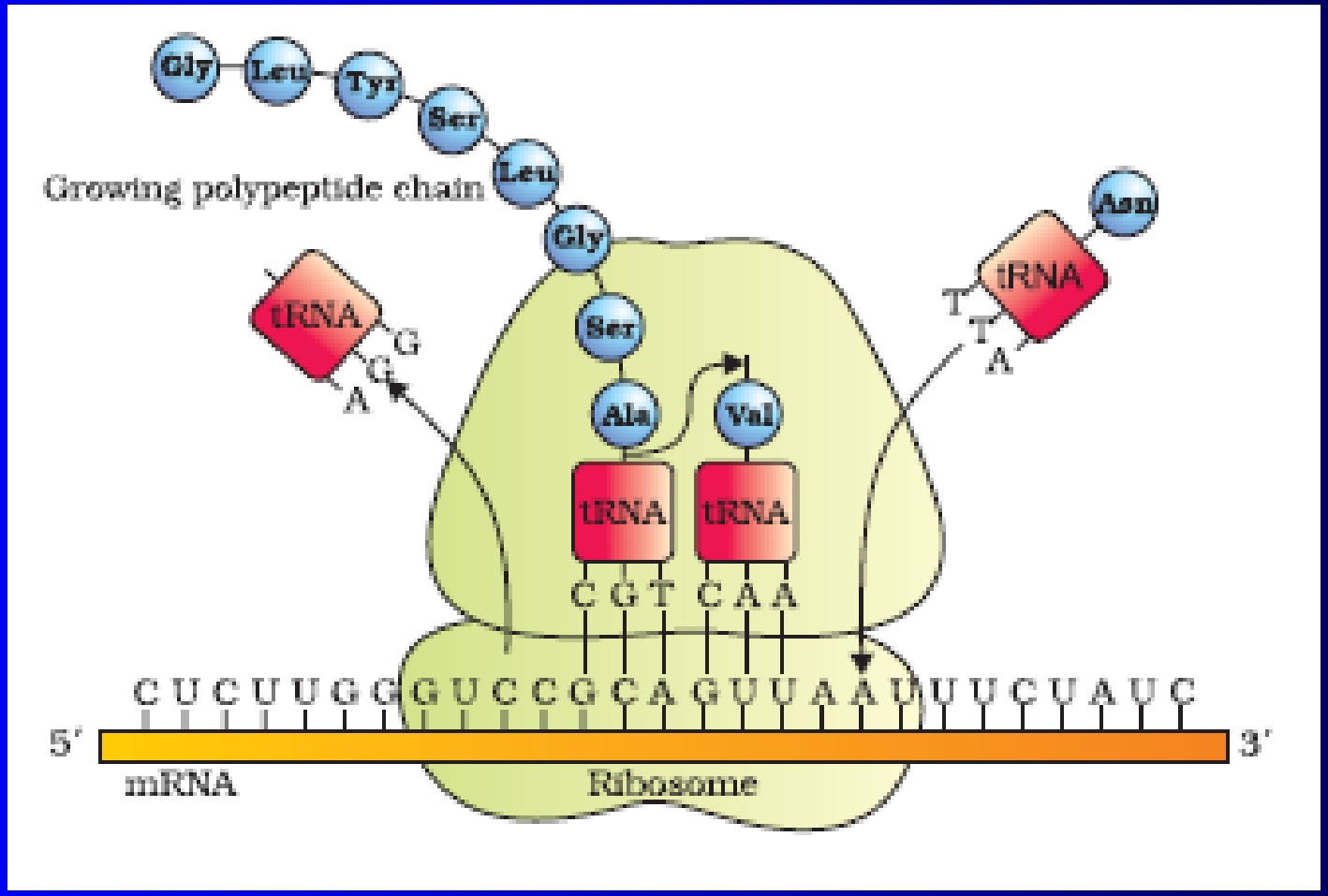
B. John Ebenezer

Requirements of translation or Protein synthesis

- Director
- Representative
- Industry
- Raw Materials
- Transport
- Link
- Catalysts
- Energy source
- Product
- DNA
- mRNA
- Ribosome
- Aminoacids
- tRNA
- Peptide bonds
- Polymerase Enzyme
- ATP
- Protein

B. John Ebenezer

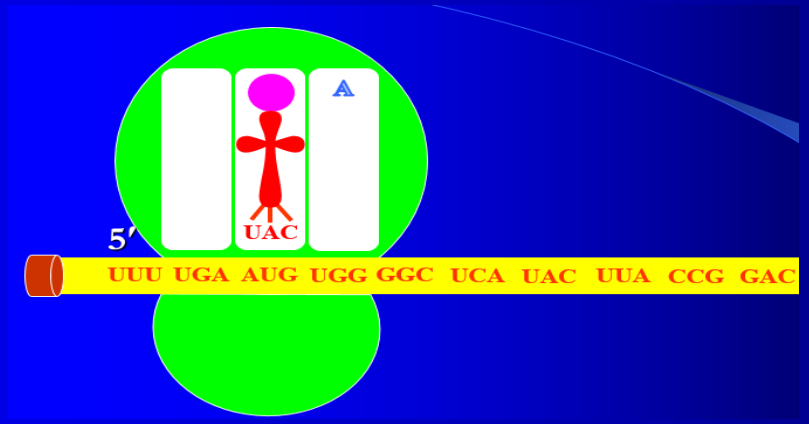
Phases of Translation



B. John Ebenezer

Phases of Translation

Translation includes three major phases such as Initiation, Elongation and Termination



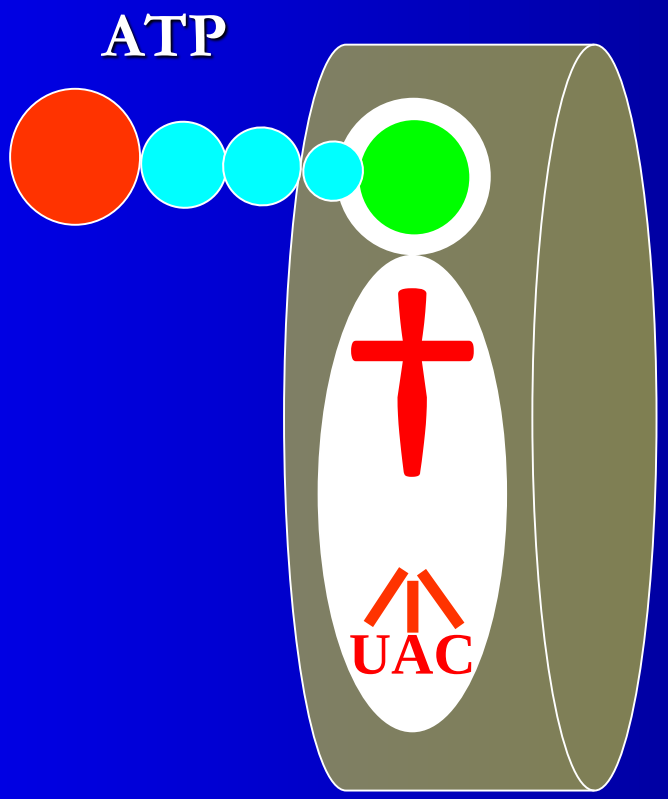
B. John Ebenezer

Activation of tRNA (Charging of tRNA or Aminoacylation of tRNA)

- The process of formation of peptide bonds requires energy.
- So the amino acids are activated in the presence of ATP and get linked to their respective tRNAs.
- This process is commonly called as **activation of tRNA** or **charging of tRNA.**
- If two such charged tRNAs are brought together, formation of peptide bond occurs between the charged aminoacids.

B. John Ebenezer

Activation of Amino acid



Aminoacyl tRNA synthetase Enzyme

B. John Ebenezer

Ribosome

- Ribosome acts as cellular factory for protein synthesis.
- The ribosome consists of structural RNAs and about 80 different proteins.
- Ribosome consists of two subunits; a smaller subunit and a larger subunit in its inactive state.
- When the smaller subunit joins with the mRNA, translation begins.

B. John Ebenezer

Initiation

- There are two sites in the larger subunit of ribosome;
- P-Site or Peptidyl Site
- A-site, Accepter Site or Aminoacyl Site.
- The smaller subunit with methionyl tRNA (Met-tRNA) attaches to the larger subunit in such a way that the start codon (AUG) comes in the P-site.

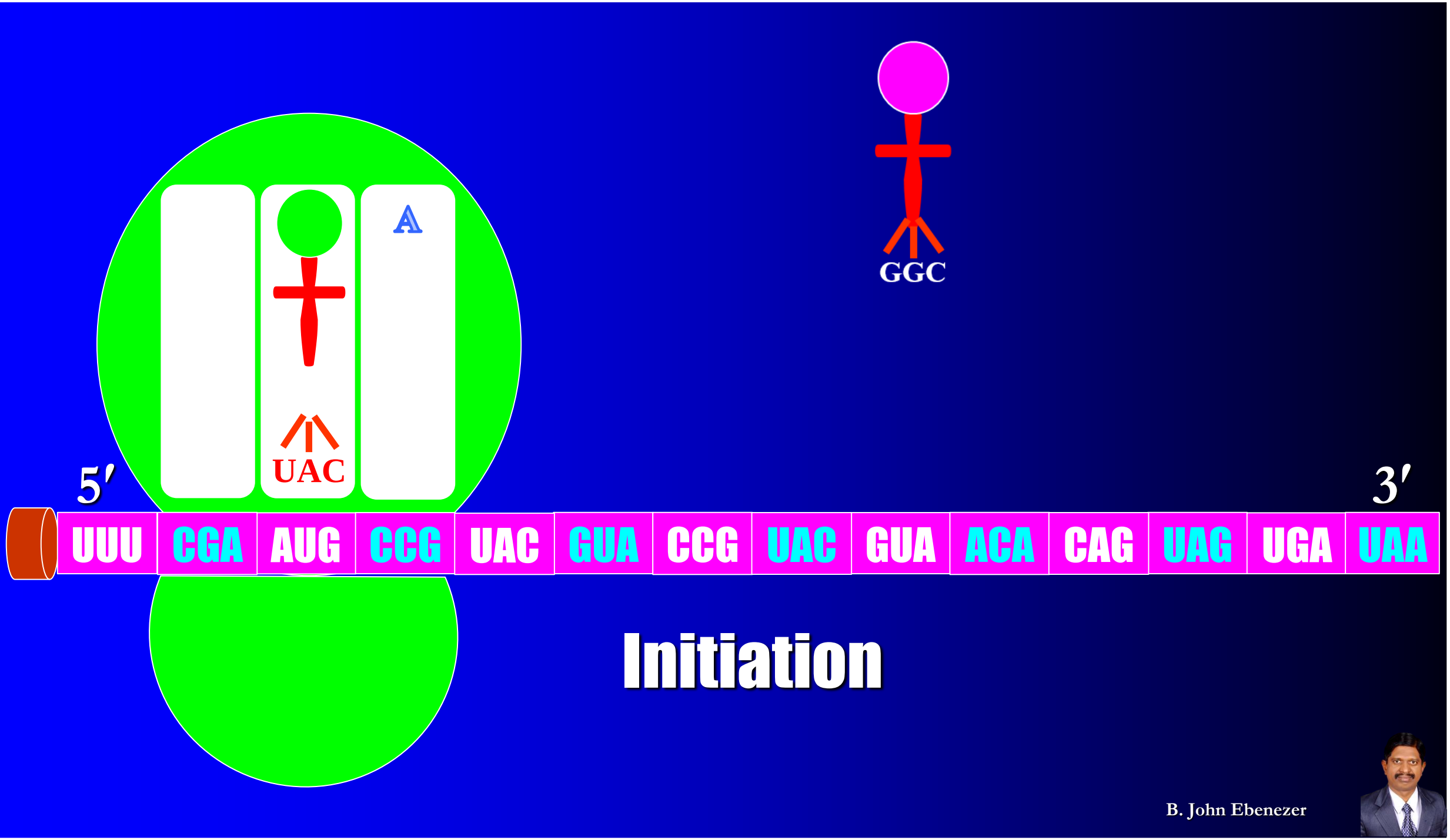
B. John Ebenezer

Initiation

- An mRNA also has some additional sequences that are not translated and are known as **untranslated regions** (**UTR**).
- The untranslated regions are present at both 5′ end (before start codon) and 3′ end (after stop codon).
- They are required for efficient translation process.

(The smaller sub unit reads the mRNA and the larger sub unit joins the aminoacids to form a polypeptide)

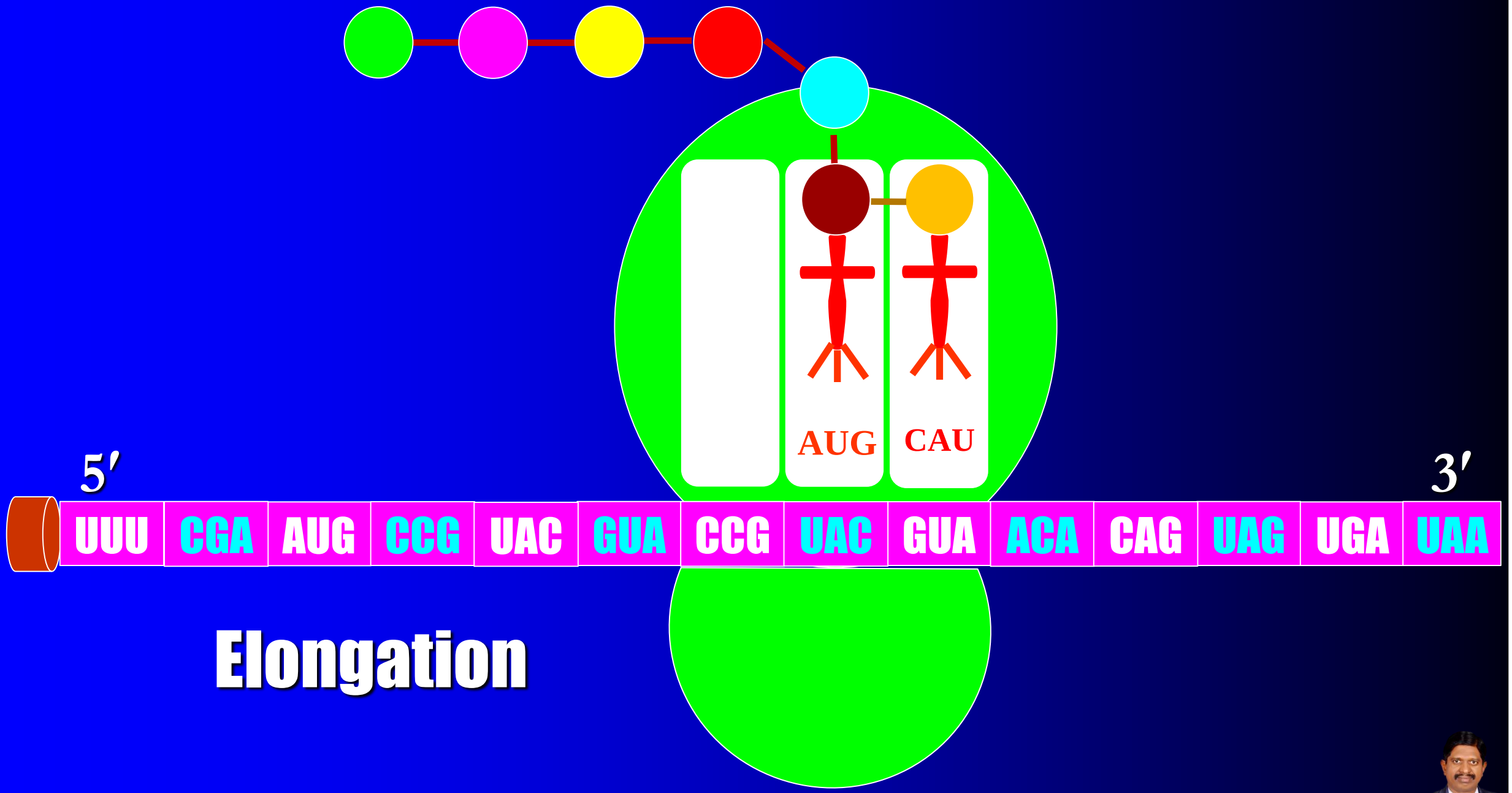
B. John Ebenezer


A
UAC
GGC
5′
3′
UUU
CGA
AUG
CCG
UAC
GUA
CCG
UAC
GUA
ACA
CAG
UAG
UGA
UAA
Initiation
B. John Ebenezer

Elongation

- The ribosome moves from codon to codon along the mRNA in 5′ to 3′ direction.
- tRNAs transfer amino acids to the A- site of ribosome as per the triplet codes in the mRNA.
- Peptide bond is formed between carboxyl group of aminoacid at P-Site and amino group of aminoacid at A-Site by the enzyme Peptidyl transferase. The polypeptide grows in length.
- When the tRNA reaches the E-Site, it gets released.

B. John Ebenezer



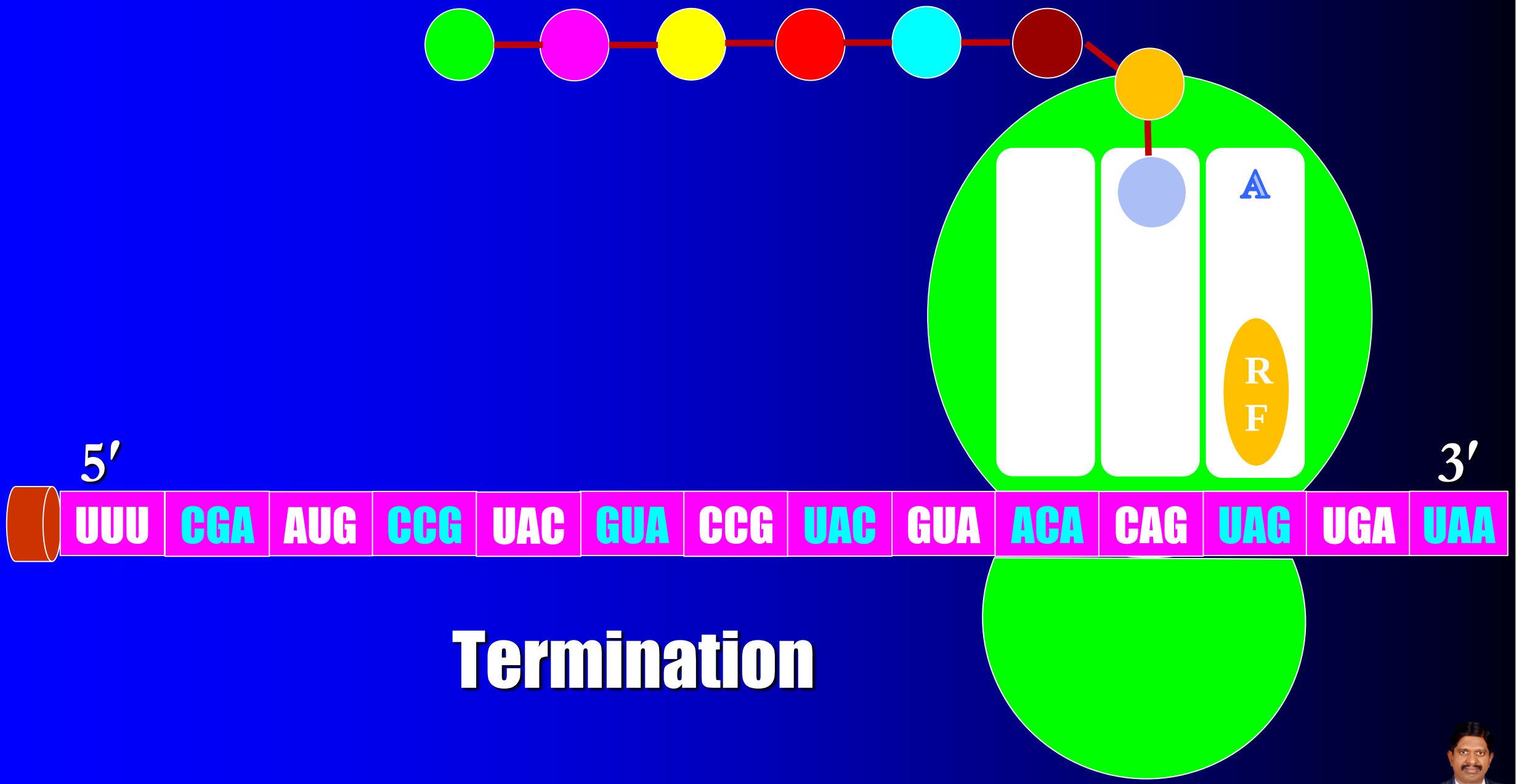
Elongation

B. John Ebenezer

Termination

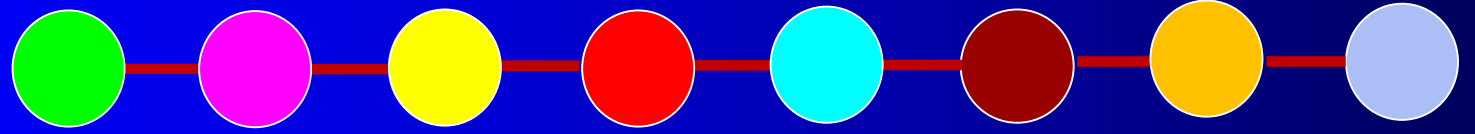
- When the ribosome reaches the stop codon, termination of polypeptide occurs as the **stop codon doesn't code for any amino acid.**
- The release factor enters the A-Site and releases the polypeptide from the ribosome.
- The smaller and larger sub units of ribosome and mRNA dissociate.

B. John Ebenezer



Termination

B. John Ebenezer



Release of Polypeptide

B. John Ebenezer

Lac Operon

Lac Operon

A cluster of genes involved in lactose metabolism in E. coli is called Lac operon.

The elucidation of the *lac* operon was also a result of a close association between a geneticist, Francois Jacob and a biochemist, Jacque Monod.

They were the first to elucidate a transcriptionally regulated system.

B. John Ebenezer

Lac Operon

In *lac* operon (here *lac* refers to lactose), a polycistronic structural gene is regulated by a common promoter and regulatory genes.

Such arrangement is very common in bacteria and is referred to as **operon**.

A few such examples are *lactose* operon, *tryptophan* operon, *arabinose* operon, *histidine* operon, *valine* operon, etc.

B. John Ebenezer

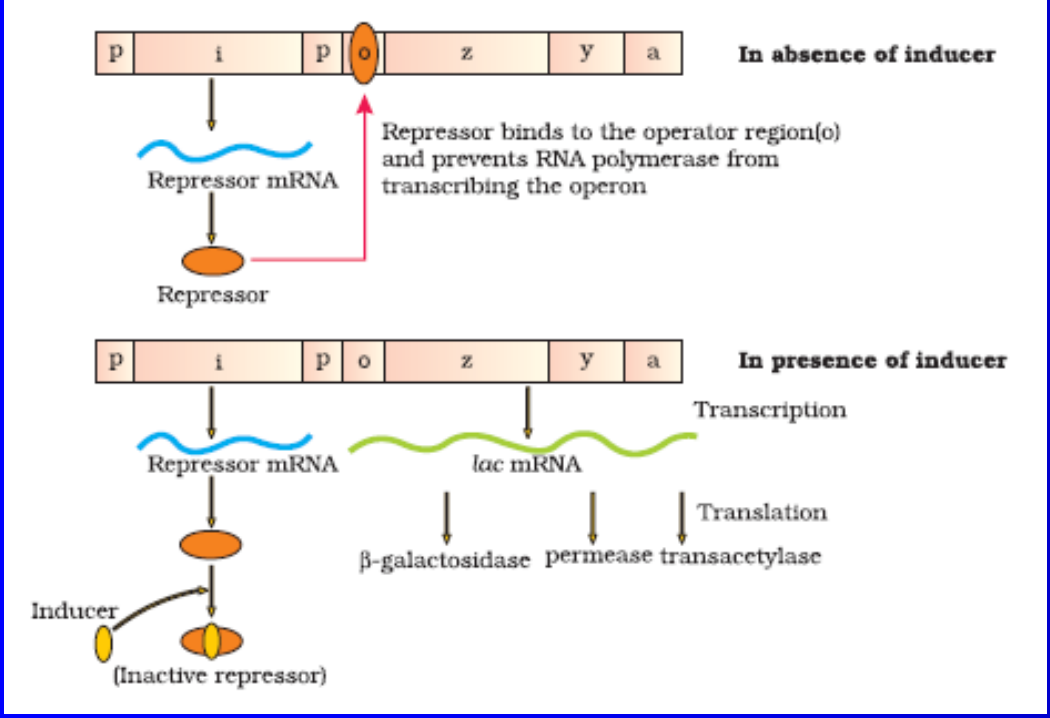
Lac Operon

All the three gene products in *lac* operon are required for metabolism of lactose.

In most other operons as well, the genes present in the operon are needed together to function in the same or related metabolic pathway.

B. John Ebenezer

Lac Operon-Regulator and Structural Genes



Genes of Lac Operon:

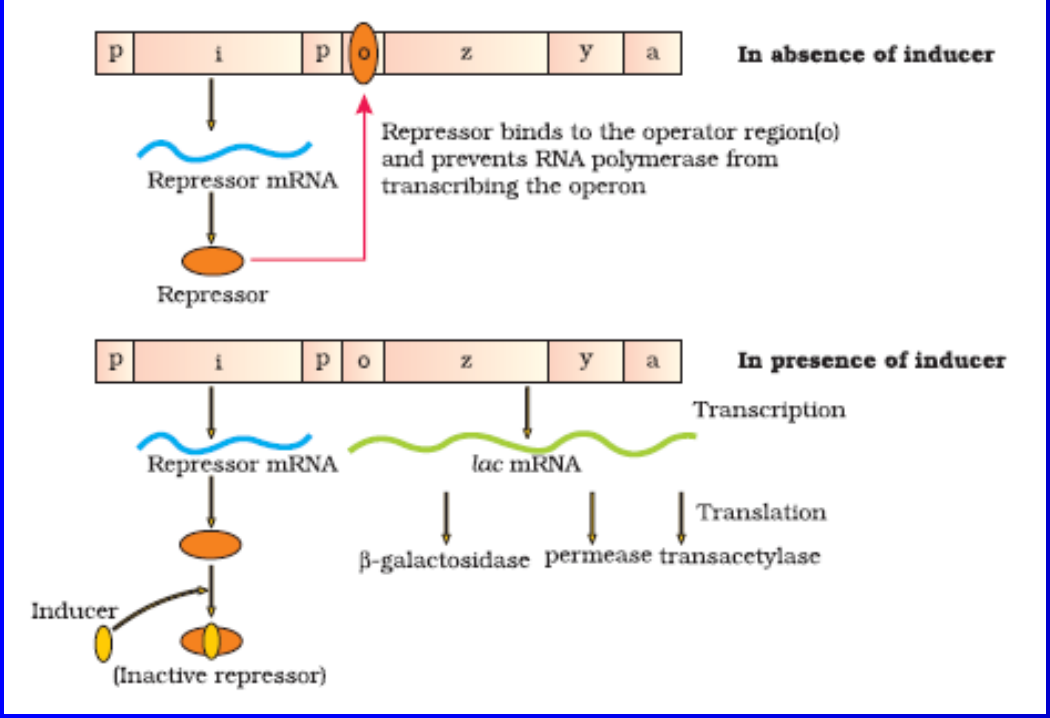
A cluster of genes involved in lactose metabolism in E. coli is called Lac operon.

Regulator Gene (*i*): The *lac* operon consists of one regulatory gene (the *i* gene)

Here the term *i* does not refer to inducer, rather it is derived from the word inhibitor) ***i*** **= inhibitor**

B. John Ebenezer

Lac Operon-Regulator and Structural Genes

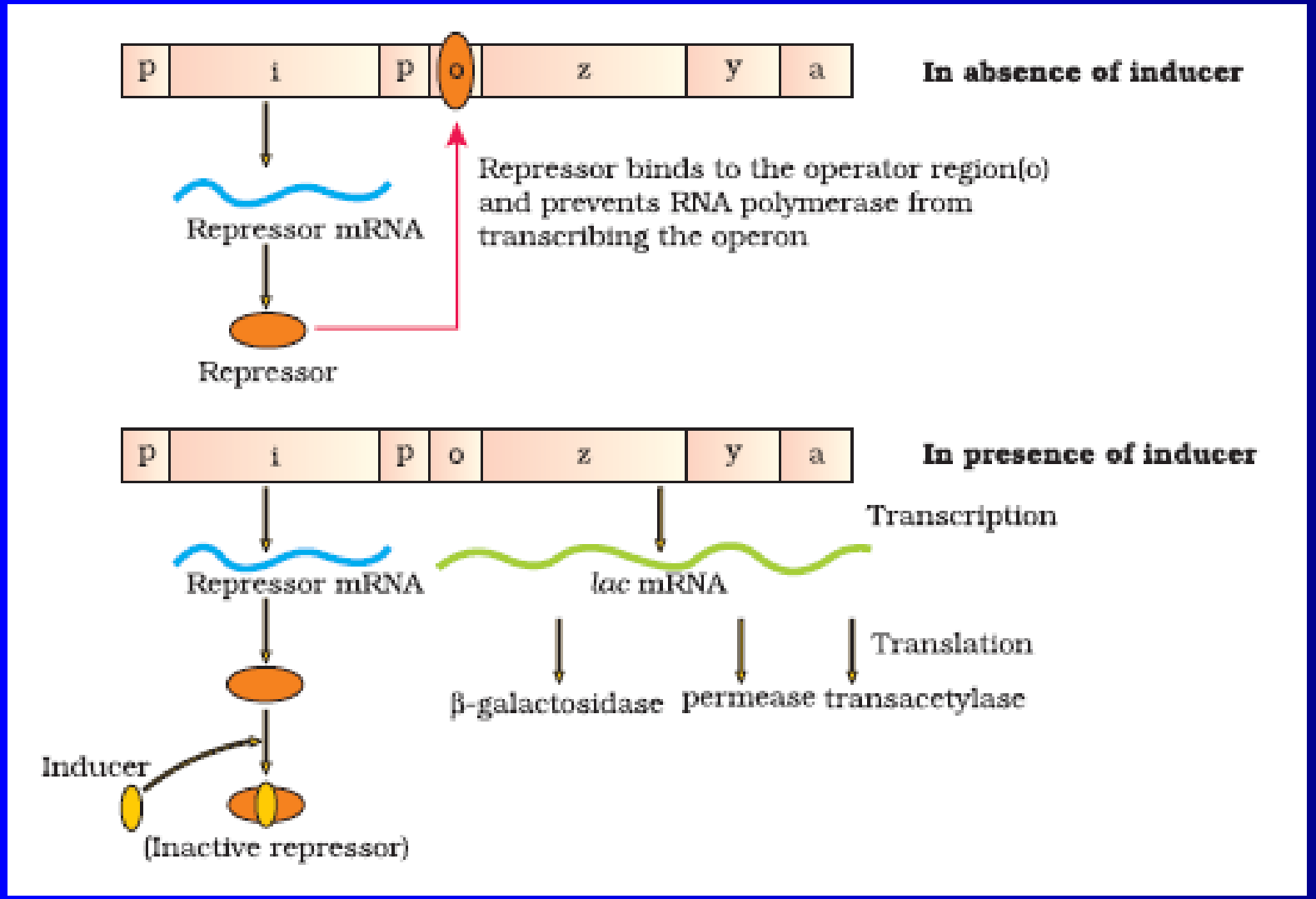


Structural Genes: The *lac* operon consists of three structural genes (*z*, *y*, and *a*).

Repressor Gene: The *i* gene codes for the repressor of the *lac* operon.

B. John Ebenezer

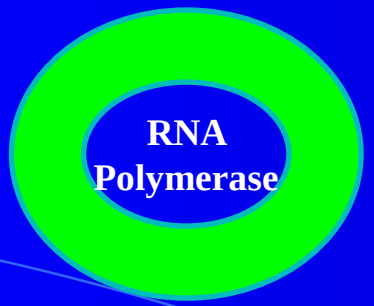
The lac Operon



The repressor protein binds to the operator region of the operon and prevents RNA polymerase from transcribing the operon

B. John Ebenezer

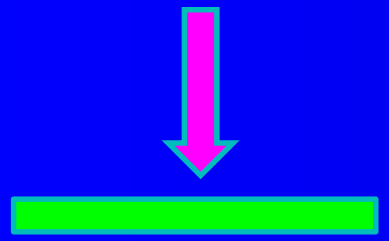
Lac Operon



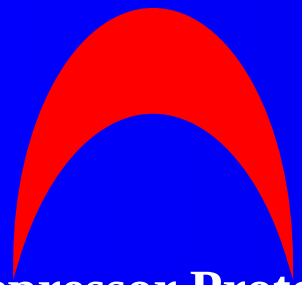
RNA Polymerase

Structural genes

Regulator Promoter Operator



mRNA



Repressor Protein

Lac Operon-Lactose is the Substrate

Substrate:

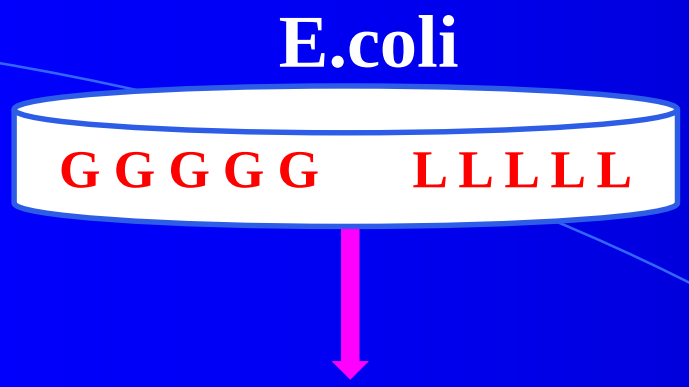
Lactose is the substrate for the enzyme beta-galactosidase and it regulates switching on and off of the operon.

Hence, lactose is termed as **inducer**.

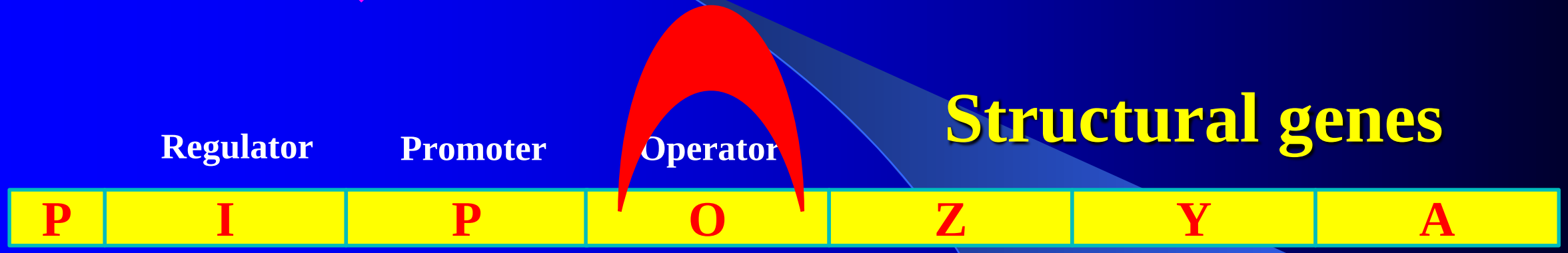
In the absence of a preferred carbon source such as glucose, if lactose is provided in the growth medium of the bacteria, the lactose is transported into the cells through the action of permease.

B. John Ebenezer

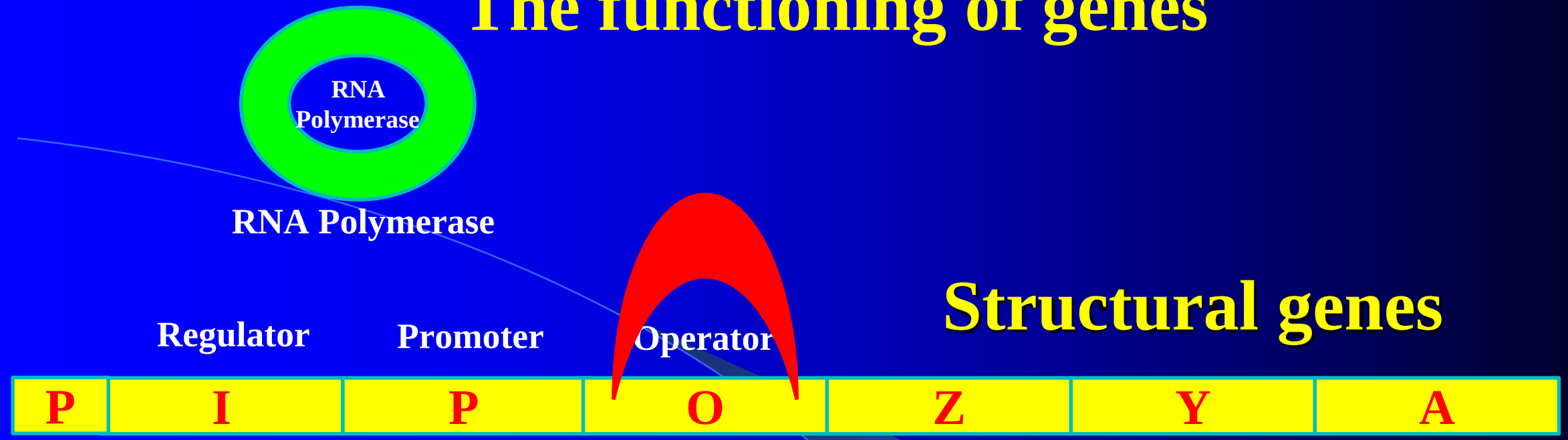
Lac Operon



Lac Operon is switched off



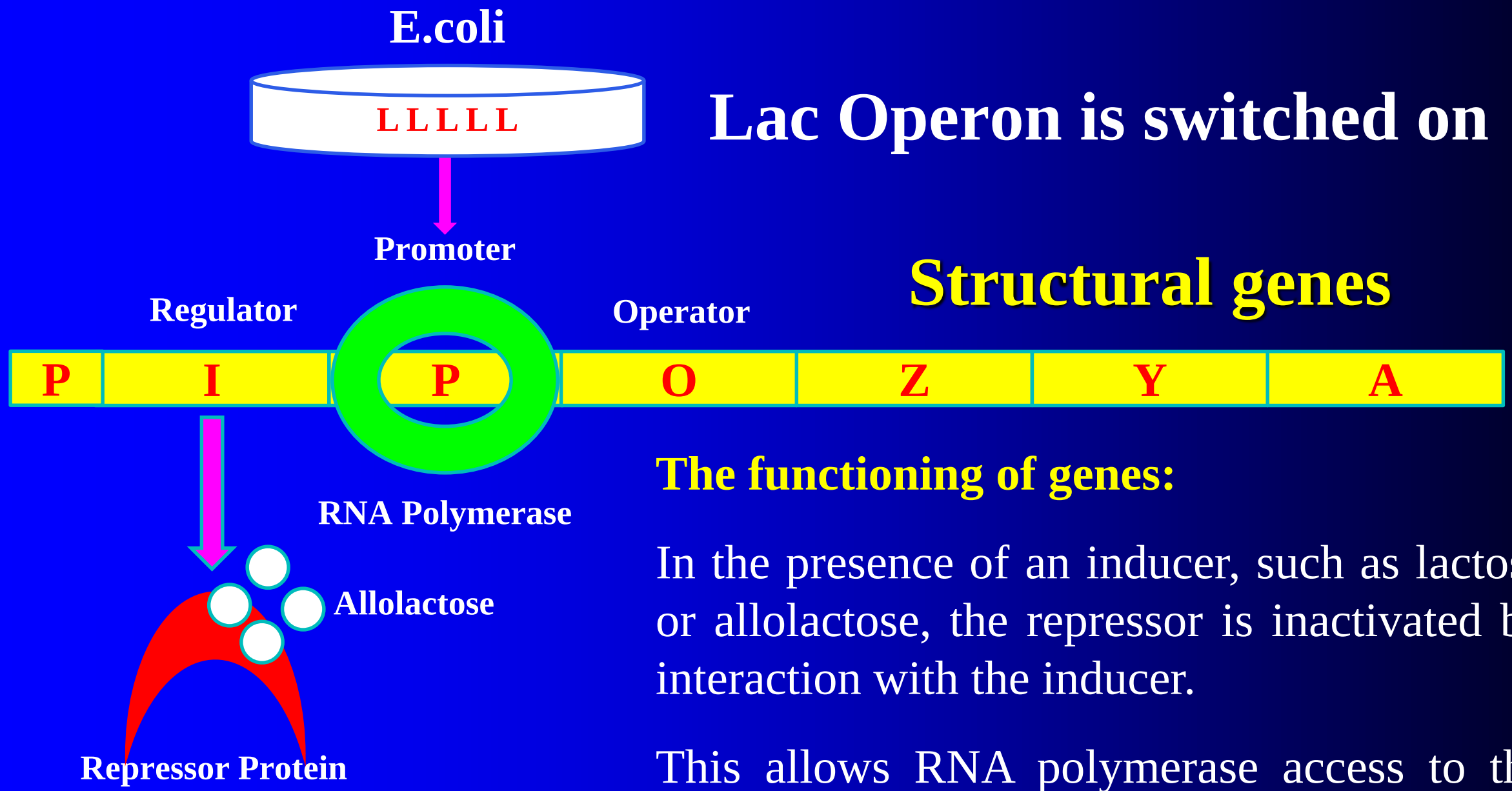
The functioning of genes



The functioning of genes:

The repressor protein binds to the operator region of the operon and prevents RNA polymerase from transcribing the operon.

B. John Ebenezer



Lac Operon is switched on

The functioning of genes:

In the presence of an inducer, such as lactose or allolactose, the repressor is inactivated by interaction with the inducer.

This allows RNA polymerase access to the promoter and transcription proceeds.

Lac Operon - Structural Genes

Transcription and Translation:

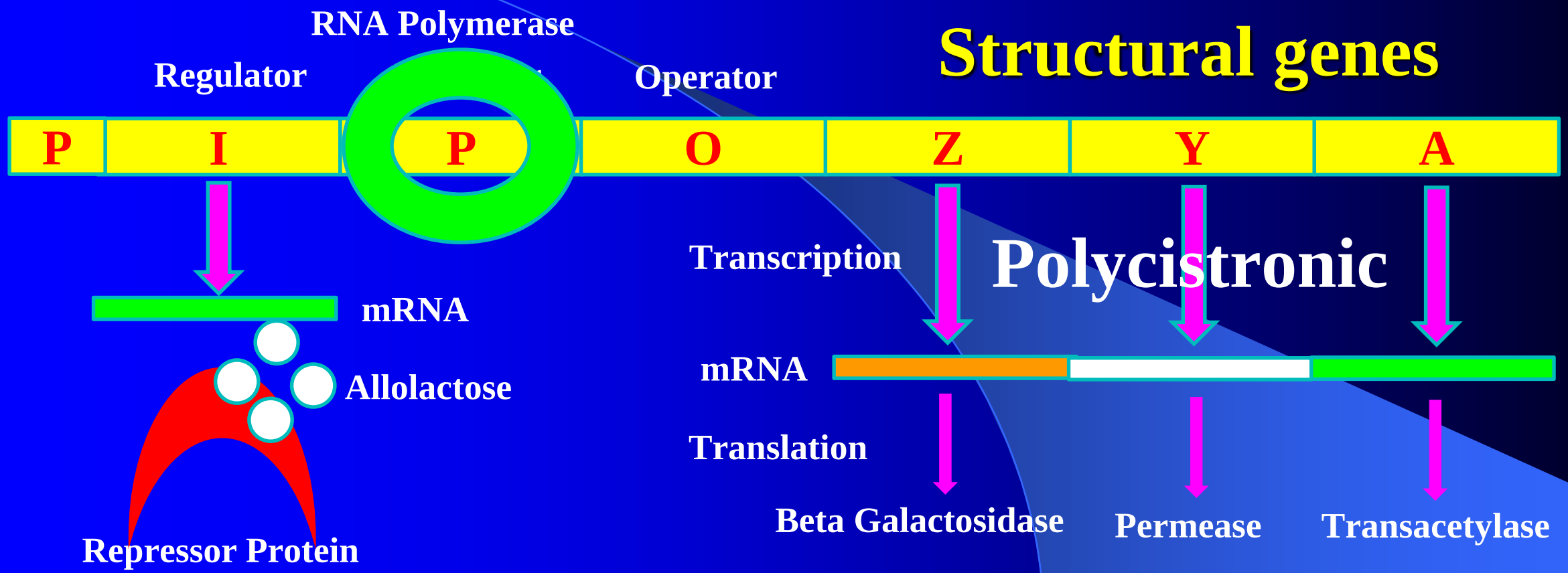
The *z* gene codes for beta-galactosidase (β-gal), which is primarily responsible for the hydrolysis of the disaccharide, lactose into its monomeric units, galactose and glucose.

The *y* gene codes for permease, which increases permeability of the cell to transport of lactose into the bacterial cell.

The *a* gene encodes a transacetylase. It protects the cells from building of toxic products. It acts as a detoxification enzyme.

B. John Ebenezer

Lac Operon


RNA Polymerase
Regulator
Operator
Structural genes
P
I
P
O
Z
Y
A
Transcription
Polycistronic
mRNA
Allolactose
mRNA
Translation
Repressor Protein
Beta Galactosidase
Permease
Transacetylase

Lac Operon - Structural Genes

Duration of Gene Expression:

The Lactose operon expresses as long as the Lactose is present.

When all lactose is converted into glucose and galactose, the reaction stops.

B. John Ebenezer

Lac Operon

Regulation of *lac* operon can also be visualised as regulation of enzyme synthesis by its substrate.

(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 lactose then induces the operon in the following manner.

The repressor of the operon is synthesised all the time constitutively from the ***i*** gene.

B. John Ebenezer

Lac Operon

Glucose or Galactose cannot act as inducers for lac operon. (Only Lactose or Allolactose acts as inducer)

Regulation of lac operon by repressor is referred to as **negative regulation**.

It is referred to as negative regulation because the repressor protein inhibits gene expression (Transcription and Translation) when lactose is absent.

The repressor protein inhibits gene expression by binding with the operator and preventing the binding of polymerase with promotor.

Lac operon is under the control of **positive regulation as well**.

B. John Ebenezer

Lac Operon

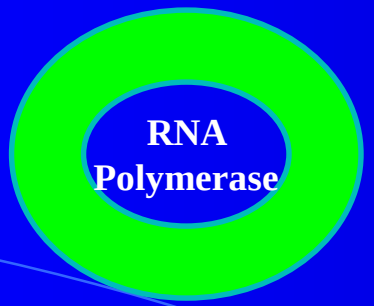
How long the lac operon would be expressed in the presence of lactose?

The Lactose operon expresses as long as the Lactose is present.

When all lactose is converted into glucose and galactose, the reaction stops.

B. John Ebenezer

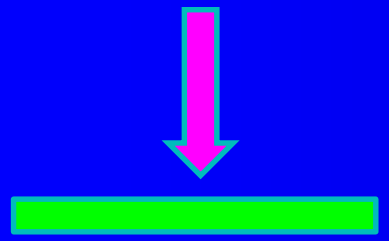
Lac Operon



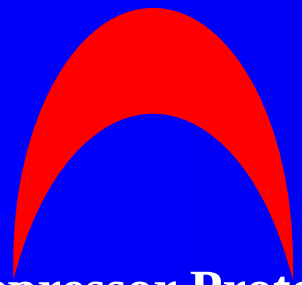
RNA Polymerase

Structural genes

Regulator Promoter Operator

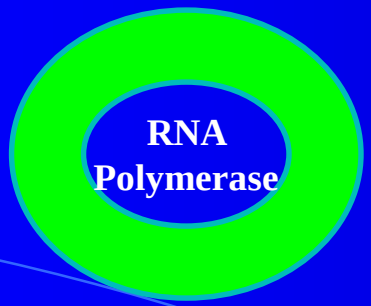


mRNA



Repressor Protein

Lac Operon



RNA Polymerase

Regulator

Promoter

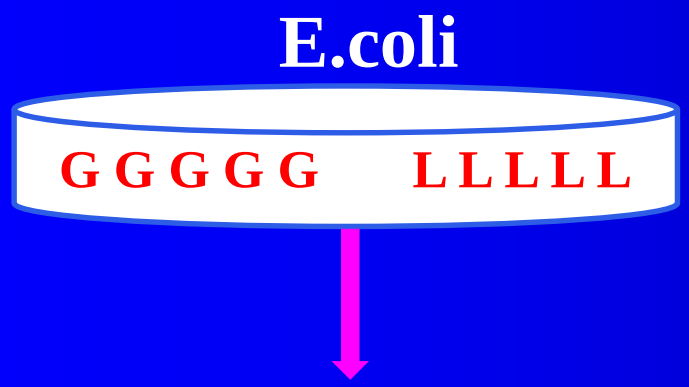
Operator

Structural genes

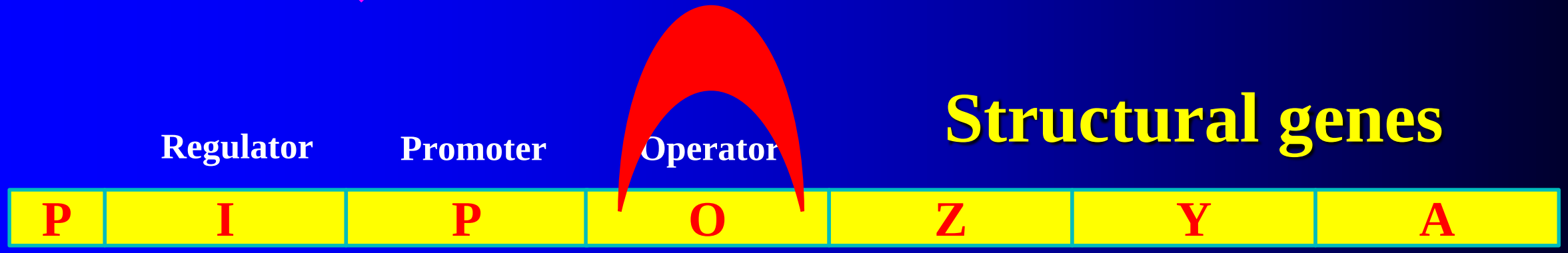
P I P O Z Y A

B. John Ebenezer

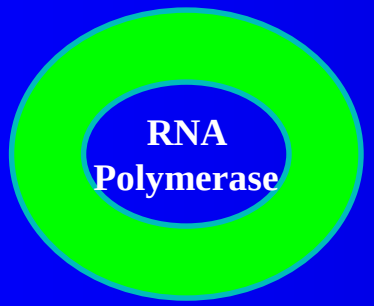
Lac Operon



Lac Operon is switched off



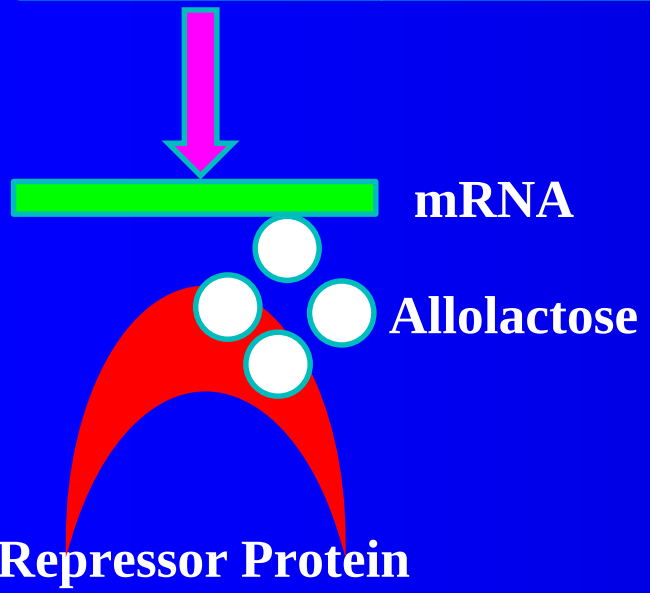
Lac Operon



RNA Polymerase

Regulator Promoter Operator Structural genes

P I P O Z Y A



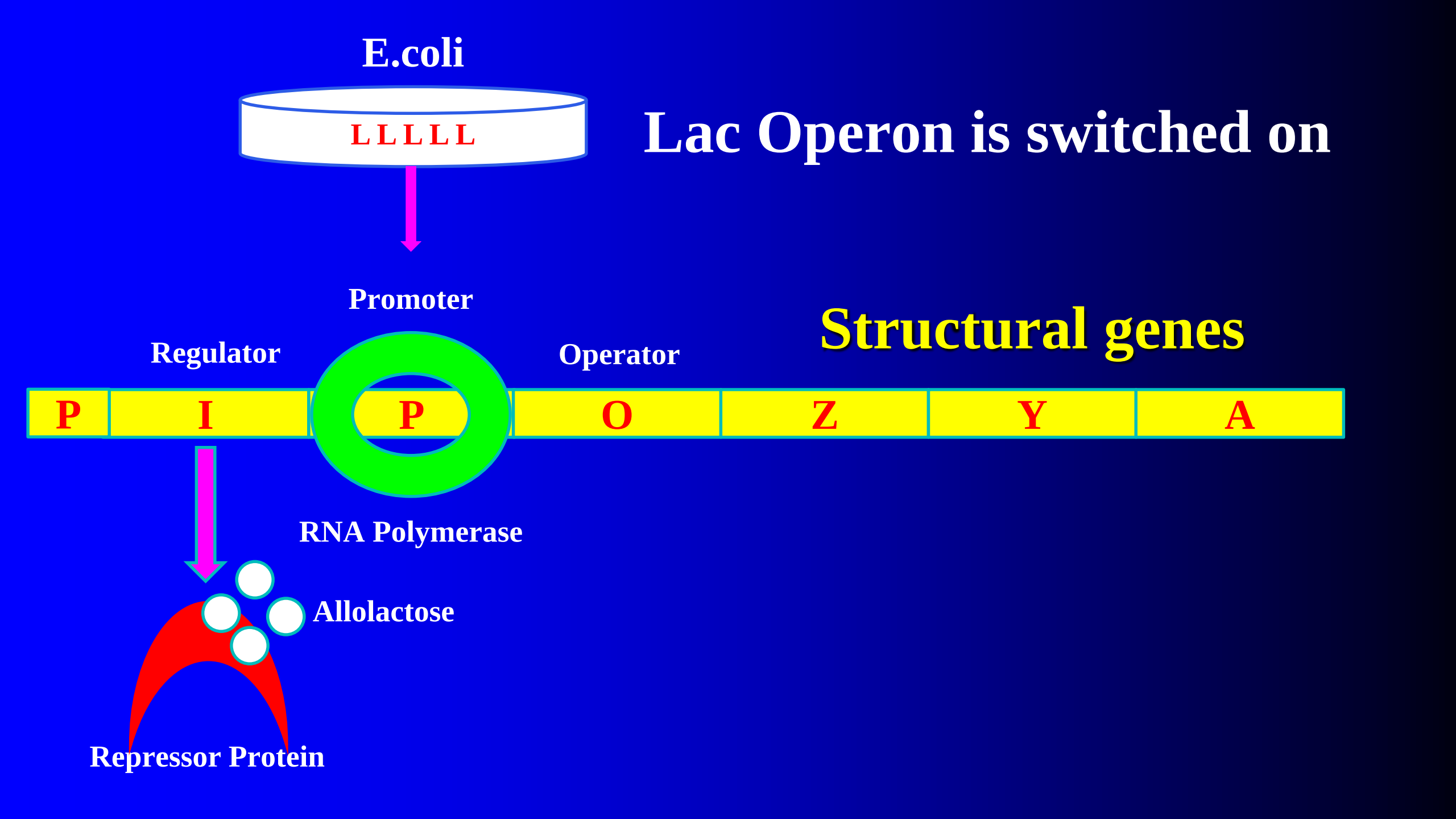
Transcription

Polycistronic

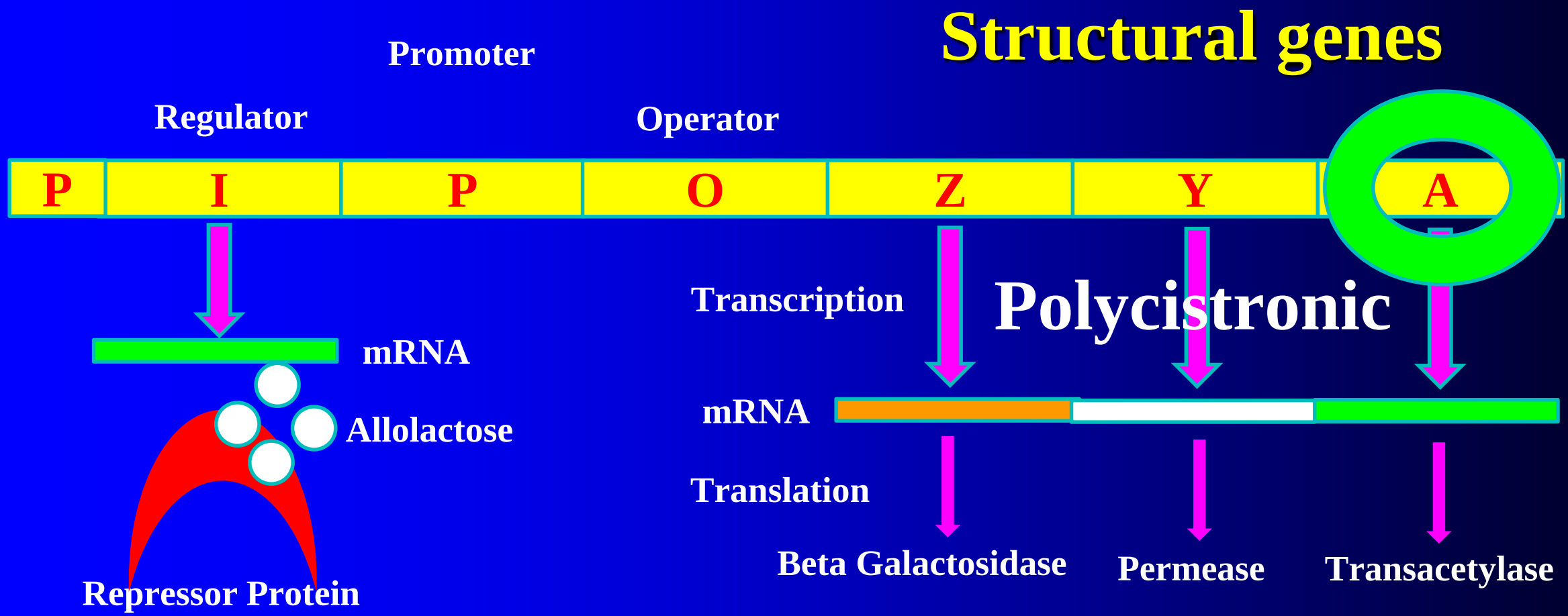
mRNA

Translation

Beta Galactosidase Permease Transacetylase


E.coli
L L L L L
Lac Operon is switched on
Promoter
Regulator
Operator
Structural genes
P
I
P
O
Z
Y
A
RNA Polymerase
Allolactose
Repressor Protein

Lac Operon

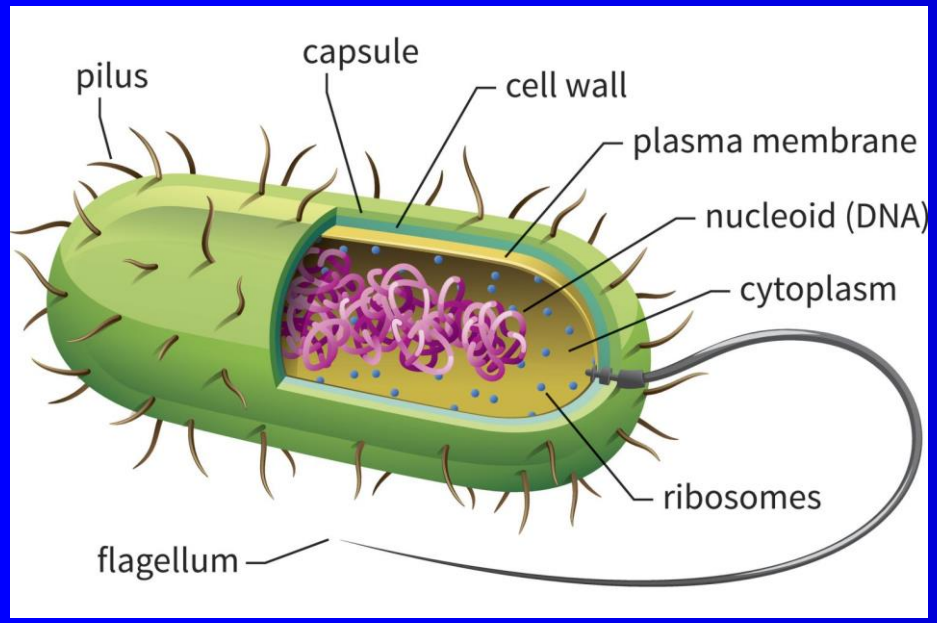

Structural genes
Promoter
Regulator
Operator
P
I
P
O
Z
Y
A
Transcription
Polycistronic
mRNA
Allolactose
mRNA
Translation
Repressor Protein
Beta Galactosidase
Permease
Transacetylase

Regulator Promoter Operator

P	I	P	O	Z	Y	A

Polycistronic

mRNA



Beta Galactosidase → Lactose → Glucose + Galactose

Permease → Increases permeability and enhances glucose uptake

Transacetylase → Adds acetyl group to Beta Galactosidase and activates it

Genetic Code

Genetic Code

It was George Gamow, a physicist, who argued that since there are only 4 bases and as they have to code for 20 amino acids, the code should have a combination of bases.

He suggested that in order to code for all the 20 amino acids, the code should be made up of three nucleotides.

4^3 ($4 \times 4 \times 4$) would generate 64 codons.

B. John Ebenezer

The Codons for the Various Amino Acids

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

B. John Ebenezer

Salient Features of Genetic Code

The codon is triplet.

61 codons code for amino acids.

3 codons do not code for any amino acids, hence they function as stop codons.

B. John Ebenezer

Salient Features of Genetic Code

Stop codons or Non-sense codons:

UAG, UGA and UAA act as stop codons or non sense codons.

They do not code for any amino acids because there are no tRNAs for stop codons.

They are involved in the termination of translation or termination of protein synthesis.

Protein synthesis cannot be terminated without stop codons.

B. John Ebenezer

Salient Features of Genetic Code

Genetic code is unambiguous:

One codon codes for only one amino acid, hence, it is unambiguous and specific.

Genetic code is degenerate:

Some amino acids are coded by more than one codon, hence the code is degenerate.

Genetic code is comma less (No Punctuations):

The codon is read in mRNA in a contiguous fashion. There are no punctuations.

B. John Ebenezer

Salient Features of Genetic Code

Genetic code is universal:

The code is nearly universal: for example, from bacteria to human **UUU would code for Phenylalanine** (phe).

Some exceptions to this rule have been found in mitochondrial codons, and in some protozoans.

AUG has dual functions:

AUG codes for Methionine (met) and also acts as initiator codon.

B. John Ebenezer

tRNA - The Adapter Molecule

tRNA the Adapter Molecule

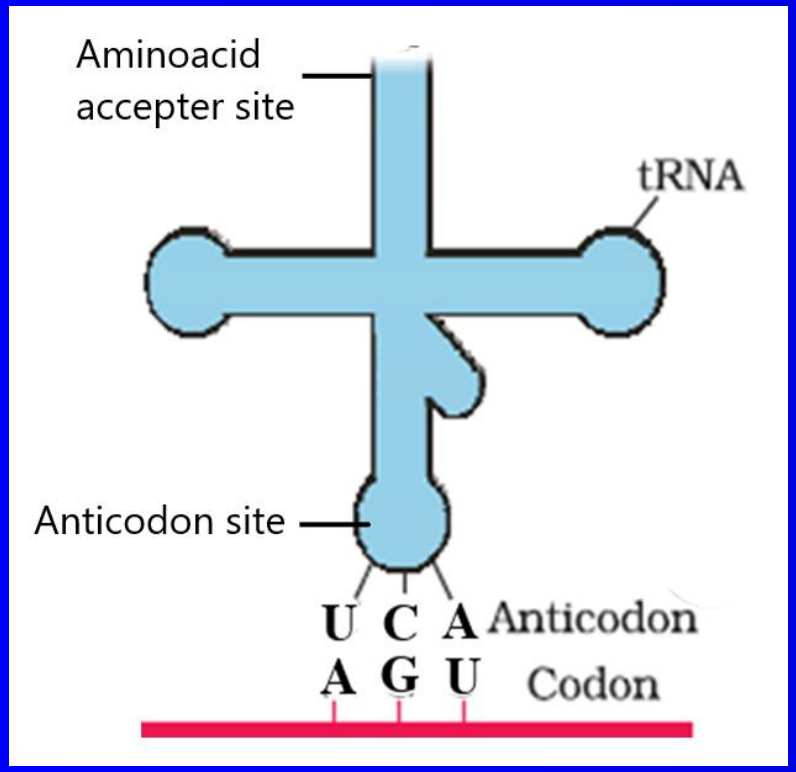
Francis Crick postulated the presence of an adapter molecule that would on one hand read the code and on other hand would bind to specific amino acids.

The tRNA, then called sRNA (soluble RNA), was known before the genetic code was postulated.

However, its role as an adapter molecule was assigned much later

B. John Ebenezer

tRNA the Adapter Molecule



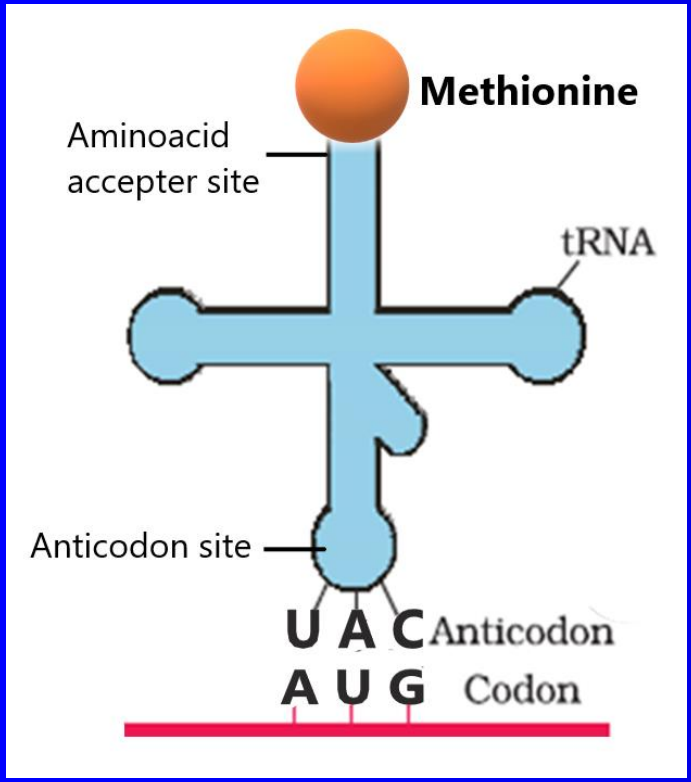
tRNAs are specific for each amino acid.

Initiator tRNA also known as $tRNA^{Met}$ is involved in initiation of translation.

There are no tRNAs for stop codons.

B. John Ebenezer

tRNA the Adapter Molecule



tRNA looks like a clover leaf.

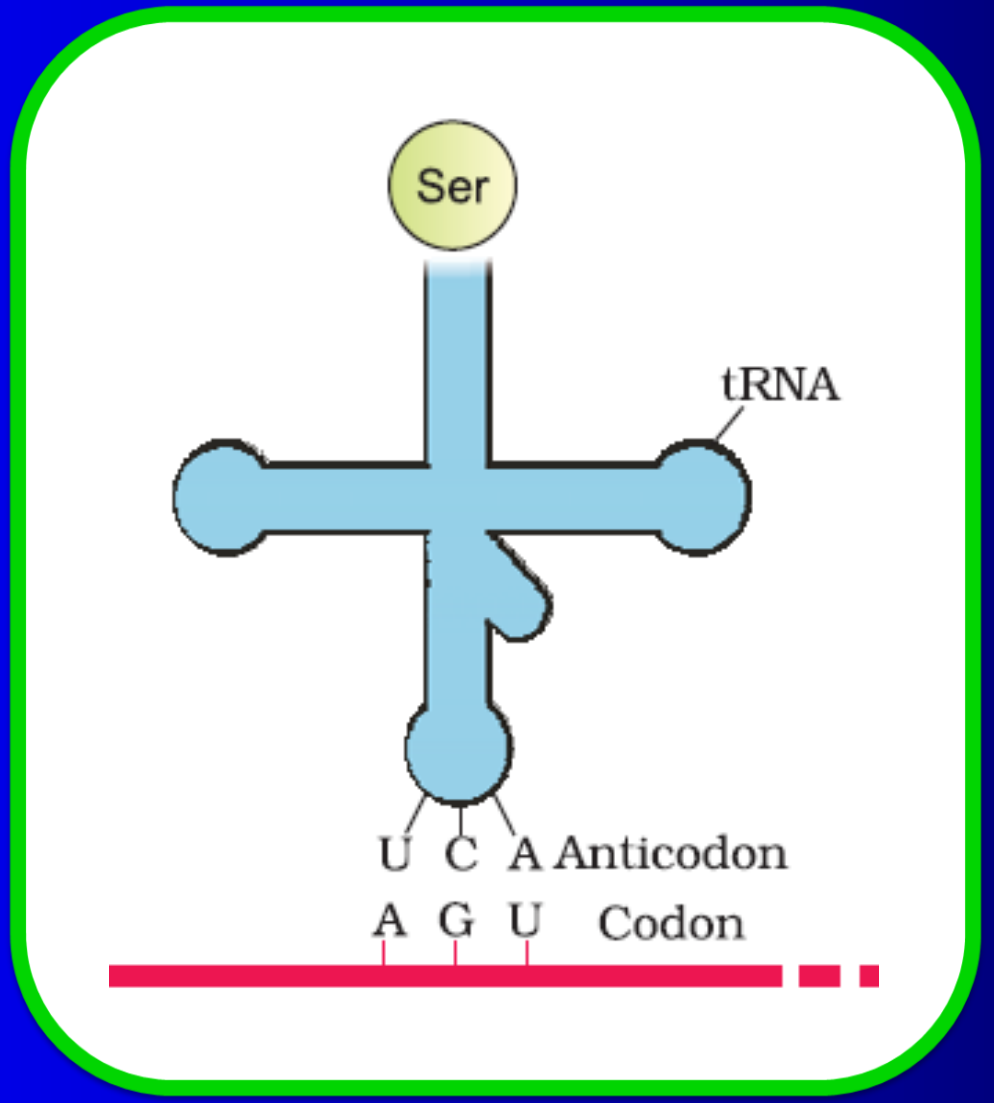
The tRNA is a compact molecule which looks like inverted L.

tRNA has an anticodon loop or anticodon site that has bases complementary to the triplet code.

It also has an amino acid accepter end or amino acid accepter site which binds with amino acid.

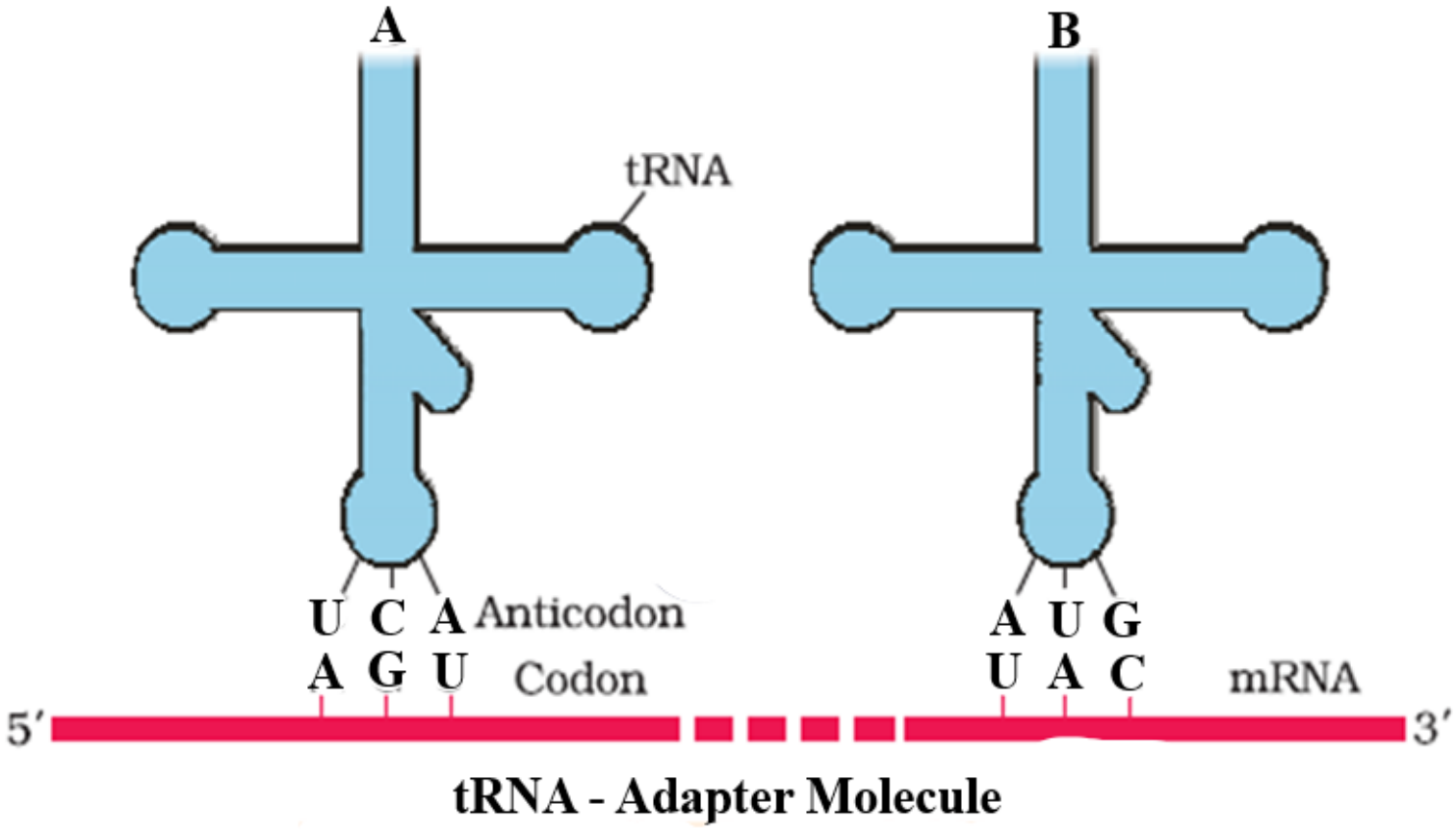
B. John Ebenezer

tRNA the Adapter Molecule



B. John Ebenezer

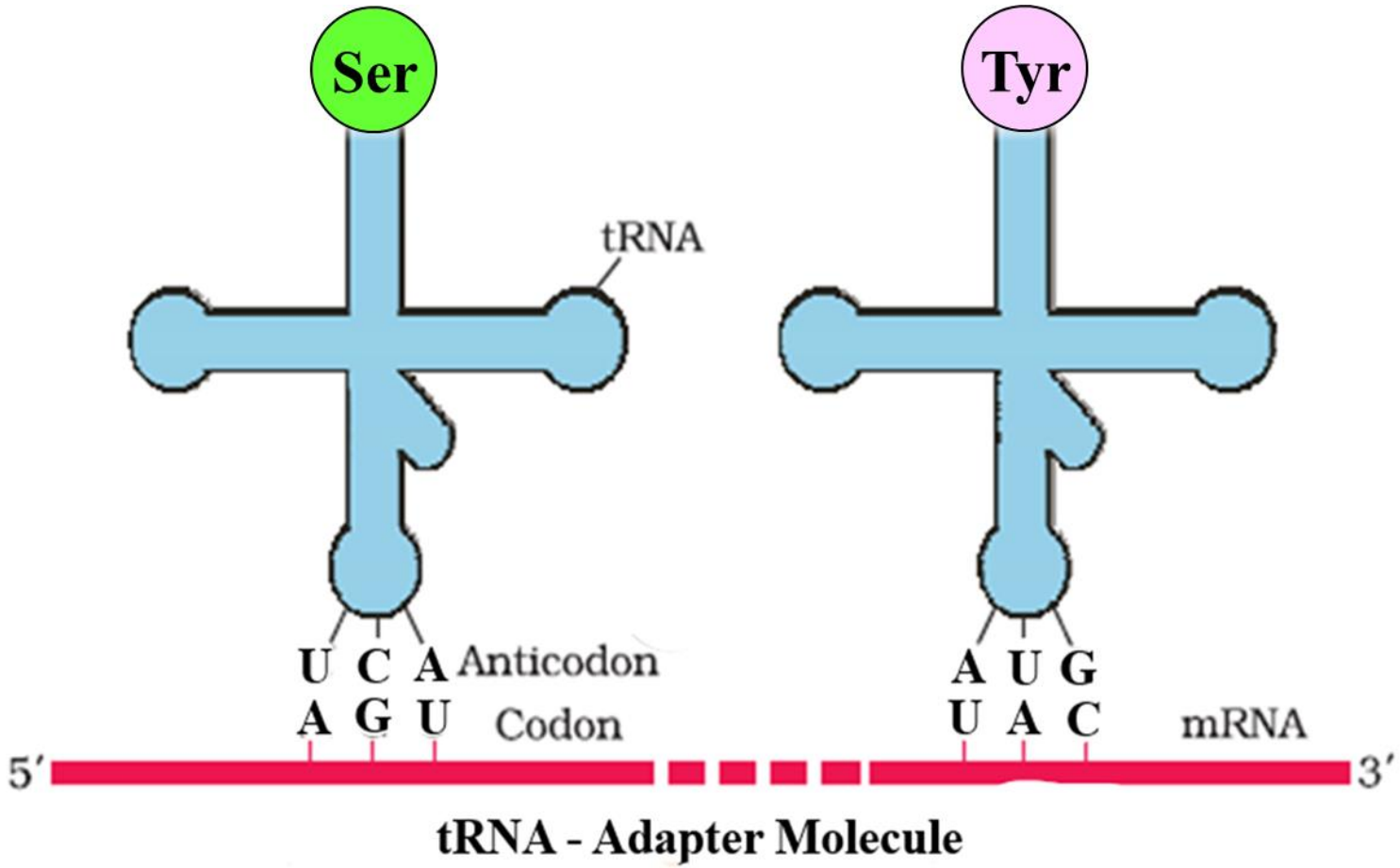
tRNA the Adapter Molecule



tRNA - Adapter Molecule

B. John Ebenezer

tRNA the Adapter Molecule

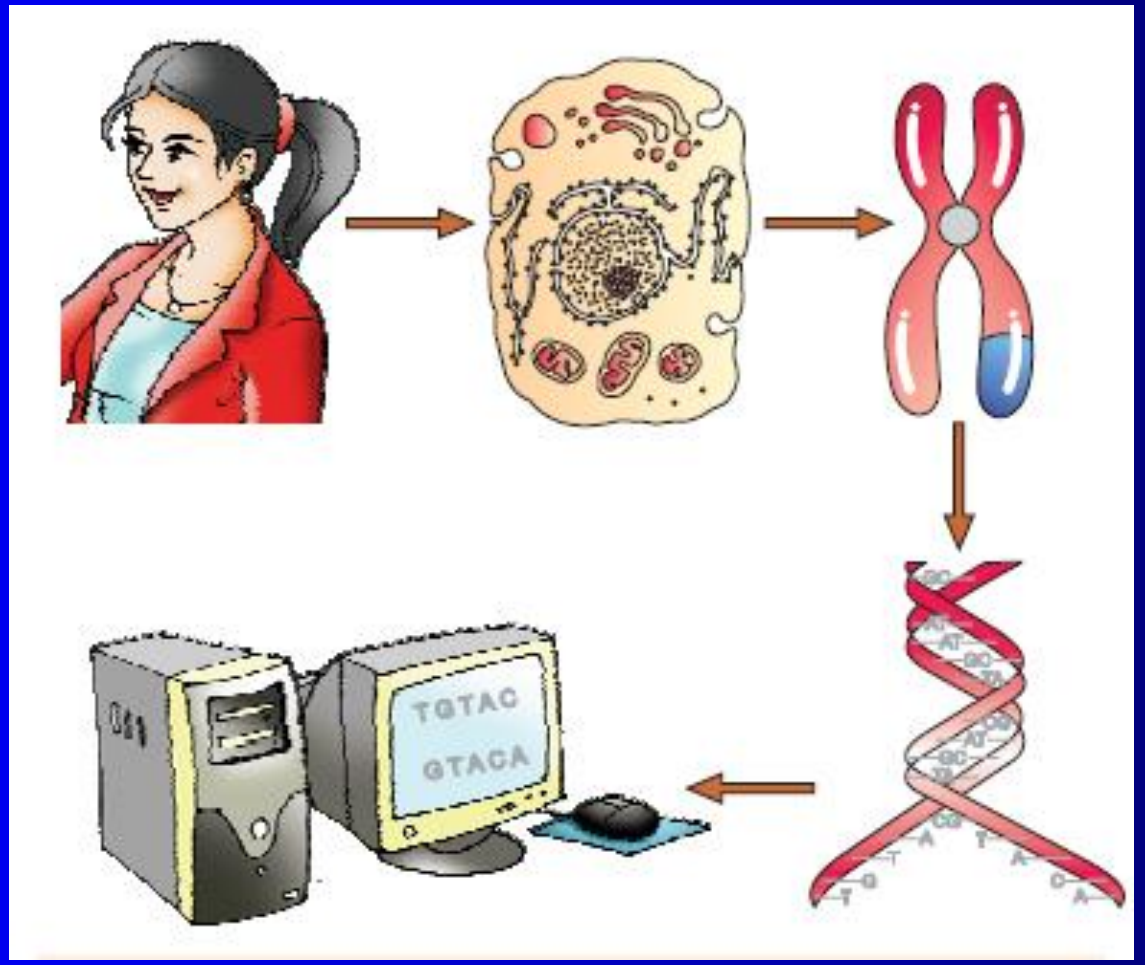


tRNA - Adapter Molecule

B. John Ebenezer

Human Genome Project

Human Genome Project



B. John Ebenezer

Human Genome Project

Human Genome Project (HGP) was called a mega project.

Human genome has approximately 3×10^9 bp.

If the cost of sequencing is US $ 3 per bp, the total estimated cost of the project would be approximately 9 billion US dollars.

B. John Ebenezer

Human Genome Project

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 are required to store the information of DNA sequence from a single human cell.

The enormous amount of data required the use of high speed computers for data storage, retrieval, and analysis.

B. John Ebenezer

Human Genome Project

The Human Genome Project was a 13-year project coordinated by the U.S. Department of Energy and the National Institute of Health.

During the early years of the HGP, the Wellcome Trust (U.K.) became a major partner.

Additional contributions came from Japan, France, Germany, China and others.

The project was completed in 2003.

B. John Ebenezer

Human Genome Project

Knowledge about the effects of DNA variations among individuals can lead to new ways to diagnose, treat and prevent the thousands of disorders that affect human beings.

B. John Ebenezer

Goals of Human Genome Project

- Identify all the approximately 20,000-25,000 genes in human DNA.
- Determine the sequences of the 3 billion chemical base pairs that make up human DNA.
- Store this information in databases.
- Improve tools for data analysis.
- Transfer related technologies to other sectors, such as industries.
- Address the ethical, legal, and social issues (ELSI) that may arise from the project.

B. John Ebenezer

Methodologies of Human Genome Project

The methods involved two major approaches.

One approach is identifying all the genes that expressed as RNA (referred to as **Expressed Sequence Tags** (ESTs).

The other method is sequencing the whole set of genome that contained all the coding and non-coding sequence, and

later assigning different regions in the sequence with functions (a term referred to as **Sequence Annotation**).

B. John Ebenezer

Methodologies of Human Genome Project

Methodologies of HGP

Expressed Sequence Tags (ESTs)

Identifying all the genes that expressed as RNA

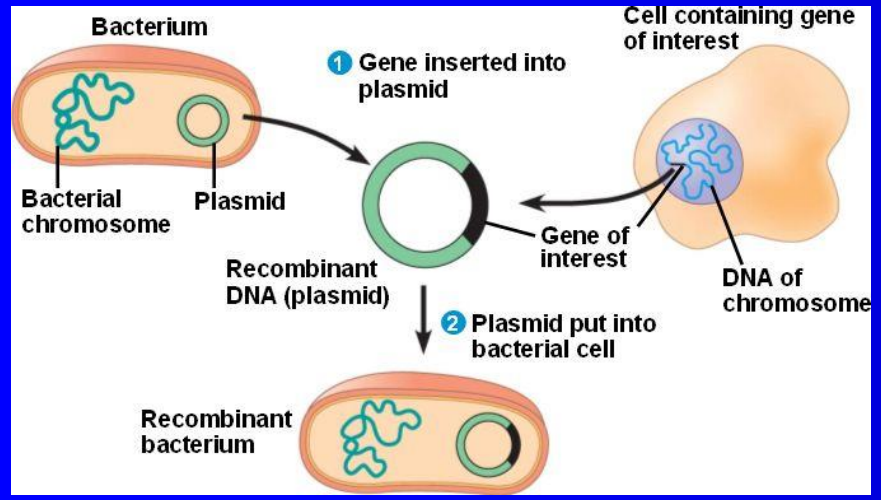
Sequence Annotation

Sequencing the whole set of genome that contained all the coding and non-coding sequence

Later assigning different regions in the sequence with functions

B. John Ebenezer

Methodologies of Human Genome Project



The total DNA from a cell is **isolated and converted into fragments** of smaller sizes and cloned in suitable host using specialised vectors.

The cloning resulted into amplification of each piece of DNA fragment so that it subsequently could be sequenced with ease.

The commonly used hosts were bacteria and yeast.

The vectors are called as **BAC** (bacterial artificial chromosomes) and

YAC (yeast artificial chromosomes)

B. John Ebenezer

Methodologies of Human Genome Project

DNA Isolation

↓

Fragments

↓

Cloning and Amplification
of DNA fragments in a suitable host

↓

Sequencing of DNA fragments
using automated DNA sequencers

↓

Arrangement based on overlapping regions

↓

Annotation and Assigning
to each chromosome

Using information from
Polymorphism and Microsatellites

Assigning
Genetic and Physical Maps
on the genome

The commonly used hosts were bacteria and yeast.

The vectors are called as BAC (bacterial artificial chromosomes) and

YAC (yeast artificial chromosomes)

B. John Ebenezer

Methodologies of Human Genome Project

The fragments were sequenced using automated DNA sequencers that worked on a method developed by Frederick Sanger.

These sequences were then arranged based on some overlapping regions present in them.

This required generation of overlapping fragments for sequencing.

B. John Ebenezer

Methodologies of Human Genome Project

Alignment of these sequences was humanly not possible.

Therefore, specialised computer based programs were developed.

These sequences were subsequently annotated and were assigned to each chromosome.

B. John Ebenezer

Methodologies of Human Genome Project

Assigning the genetic and physical maps on the genome.

This was made possible by using information on polymorphism of

restriction endonuclease recognition sites,

and

some repetitive DNA sequences known as microsatellites

B. John Ebenezer

Salient Features of Human Genome Project

Salient Features of Human Genome Project

The **human genome contains 3164.7 million nucleotide bases.**

The **average gene** consists of **3000 bases**.

Size of genes vary greatly, with the largest known human gene being **Dystrophin at 2.4 million bases**.

The total **number of genes estimated is 30,000**

It is much lower than previous estimates of 80,000 to 1,40,000 genes.

B. John Ebenezer

Salient Features of Human Genome Project

Almost all (99.9 per cent) nucleotide bases are exactly the same in all people.

The functions are unknown for over 50 per cent of discovered genes.

Less than 2 per cent of the genome codes for proteins.

B. John Ebenezer

Salient Features of Human Genome Project

Repeated sequences make up very large portion of the human genome.

Repetitive sequences are stretches of DNA sequences that are repeated many times, sometimes hundred to thousand times.

They have no direct coding functions, but they show chromosome structure, dynamics and evolution.

Chromosome 1 has most genes (2968)

Chromosome Y has the fewest genes (231)

B. John Ebenezer

Salient Features of Human Genome Project

Scientists have identified about **1.4 million locations** where single base DNA differences (SNPs - single nucleotide polymorphism known as ‘snips’ occur in humans.

The occurrence of two or more different forms or morphs in the population of a species is referred to as polymorphism.

In simplest term, polymorphism is a process where two or more possibilities of a trait are found on one gene.

B. John Ebenezer

DNA Finger Printing

DNA Finger Printing

99.9 per cent of base sequence among humans is the same.

If one aims to find out genetic differences between two individuals or among individuals of a population, sequencing the DNA every time would be a daunting and expensive task.

Imagine trying to compare two sets of 3×10^6 base pairs.

DNA fingerprinting is a very quick way to compare the DNA sequences of any two individuals.

B. John Ebenezer

DNA Finger Printing

DNA from every tissue such as blood, hair-follicle, skin, bone, saliva, sperm of an individual show the same degree of polymorphism.

Hence they become very useful identification tool in forensic applications.

As the polymorphisms are inheritable from parents to children, DNA fingerprinting is the basis of paternity testing, in case of disputes.

B. John Ebenezer

DNA Finger Printing

Mutation occurs in the germ cell is inherited to the offsprings through sexual reproduction.

Variation in allelic sequence has been described as a DNA polymorphism if more than one variant (allele) at a locus occurs in human population with a frequency greater than 0.01.

B. John Ebenezer

DNA Finger Printing

An **inheritable mutation** observed in a population at high frequency is referred to as **DNA polymorphism**.

The probability of heritable variation occur in noncoding DNA sequence would be higher.

Mutations in the noncoding DNA sequences may not have any immediate effect in an individual's reproductive ability.

These mutations keep on accumulating generation after generation and form one of the basis of variability/polymorphism.

B. John Ebenezer

DNA Finger Printing

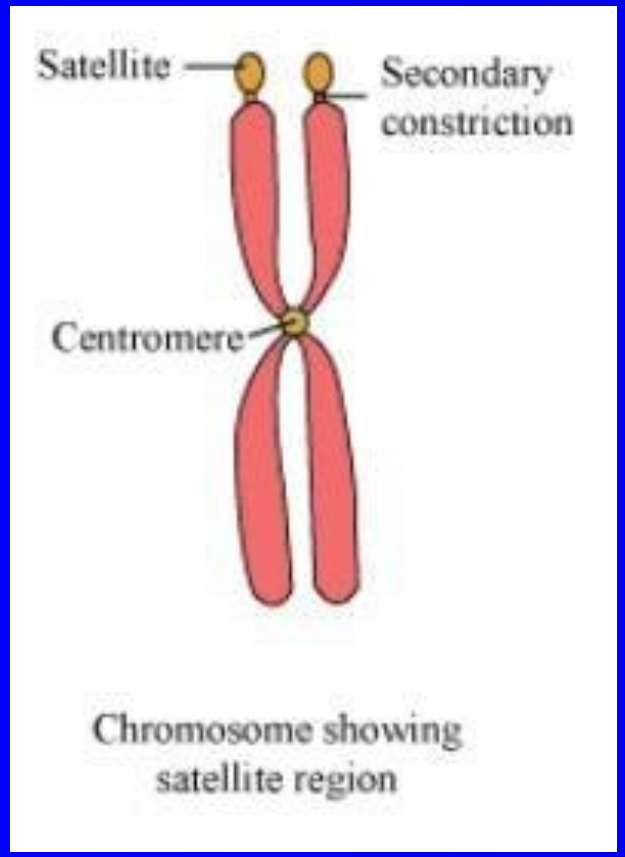
There is a variety of different types of polymorphisms ranging from single nucleotide change to very large scale changes.

Such polymorphisms play very important role in evolution and speciation,

B. John Ebenezer

Polymorphism for easy understanding

Satellite Chromosomes



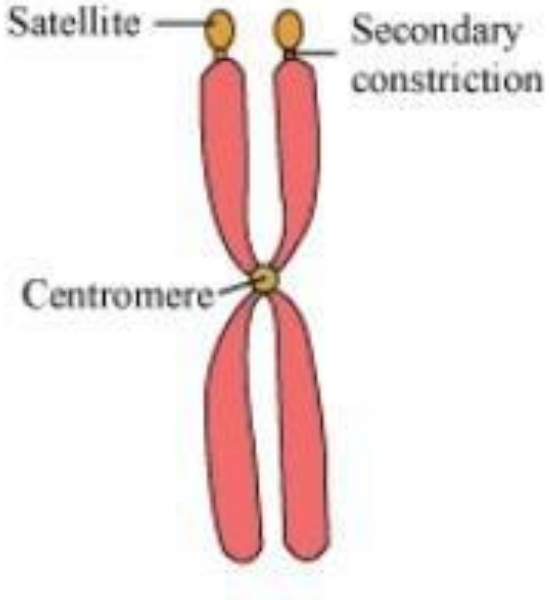
Chromosome showing satellite region

The chromosomes which have a small segment separated from remaining part of the chromosome by secondary constriction are called satellite chromosomes.

Satellite is found only in some chromosomes.
In humans, the satellite is associated with a short arm of acrocentric chromosomes such as in chromosomes 13, 14, 15 and 22.

B. John Ebenezer

Satellite Chromosomes



Chromosome showing satellite region

The SAT chromosomes are involved in the formation of nucleolus and consist of nucleolar organizer (NOR).

NOR is a region containing multiple copies of the 18S and 28S ribosomal genes that synthesize ribosomal RNA required by ribosomes.

B. John Ebenezer

Satellite Chromosomes

A **small segment** of chromosome separated from remaining part of the chromosome by secondary constriction is called a **satellite.**

A microsatellite is a short tandem repeat (STR) with base pairs ranging in length from 1-9 base pairs, repeated five to fifty times.

A minisatellite is a variable number tandem repeat (VNTR) with base pairs ranging in length from 10 to 60 base pairs, repeated five to fifty times.

The size of VNTR varies in size from 0.1 to 20 kb.

B. John Ebenezer

Satellite Chromosomes

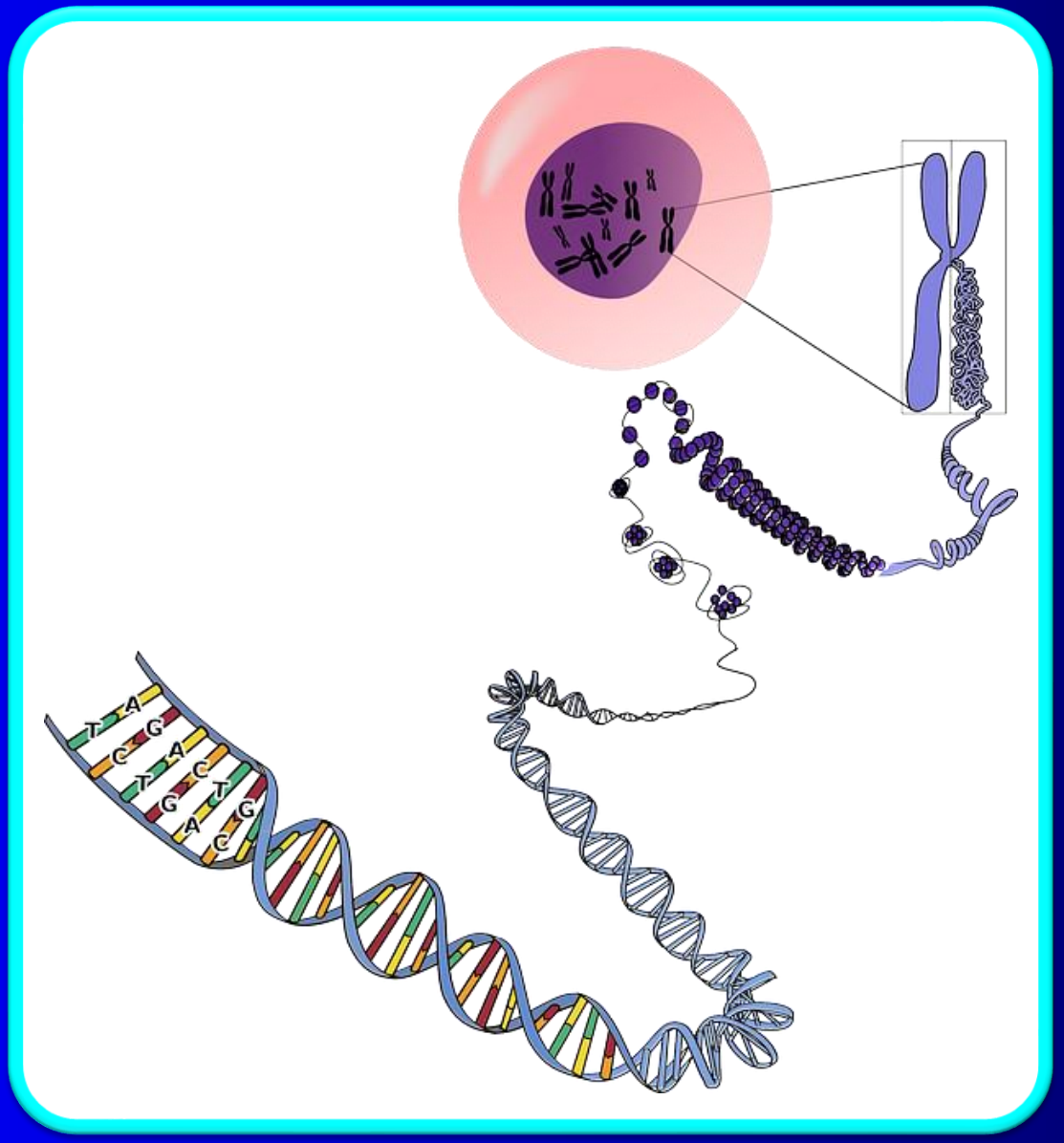
The two types of tandem repeats are STR and VNTR.

The tandem repeats which consist of 1-9 base pairs is STR.

The tandem repeats which consist of 10-60 base pairs is VNTR.

B. John Ebenezer

Steps of DNA Finger Printing



B. John Ebenezer

Microsatellites and Minisatellites

Microsatellites (STR)	Minisatellites (VNTR)
Microsatellites are short tandem repeats that consist of 1 to 9 base pairs	Minisatellites are long tandem repeats that consist of 10 to 60 base pairs
They are also known as short sequence repeats (SSR) or simple tandem repeats (STR).	They are also known as variable number tandem repeats (VNTR).
Microsatellites are rich with A and T bases.	Minisatellites are rich with G and C bases.

B. John Ebenezer

Significance of Tandem Repeats

VNTR and STR are two types of tandem repeats.

They are composed of repetitive sequences, arranging adjacent to each other in an array.

They are structural regions of the eukaryotic genome.

They consist of noncoding DNA.

They are inherited from parents.

They produce genetic polymorphism.

They are used as genetic markers in forensic genetics.

B. John Ebenezer

Significance of Tandem Repeats

Mutations in both VNTR and STR lead to genetic diseases.

VNTR form the DNA bands during DNA finger printing and help in identification.

B. John Ebenezer

Polymorphisms

Polymorphism is the difference in the nucleotide sequence between individuals.

These differences can be single base pair changes, deletions, insertions, or even changes in the number of copies of a given DNA sequence.

SNPs (single nucleotide polymorphisms) are the most common type of DNA polymorphism in humans.

B. John Ebenezer

Polymorphisms

An example of an SNP is if cytosine (C) nucleotide occurs at a particular locus in one person's DNA, thymine (T) nucleotide occurs at the same locus in another person's DNA.

Around 90% of all human genetic variation is due to SNPs.

High polymorphism in DNA creates great variation from person to person and can be easily identified.

B. John Ebenezer

DNA Finger Printing

Polymorphism in DNA sequence is the basis of genetic mapping of human genome as well as of DNA fingerprinting.

Polymorphism (variation at genetic level) arises due to mutations.

New mutations may arise in an individual either in somatic cells or in the germ cells (cells that generate gametes in sexually reproducing organisms).

B. John Ebenezer

Polymorphism

The occurrence of two or more different forms or morphs in the population of a species is referred to as polymorphism.

In simplest term, polymorphism is a process where two or more possibilities of a trait are found on one gene.

Polymorphism

B. John Ebenezer

SNIPS

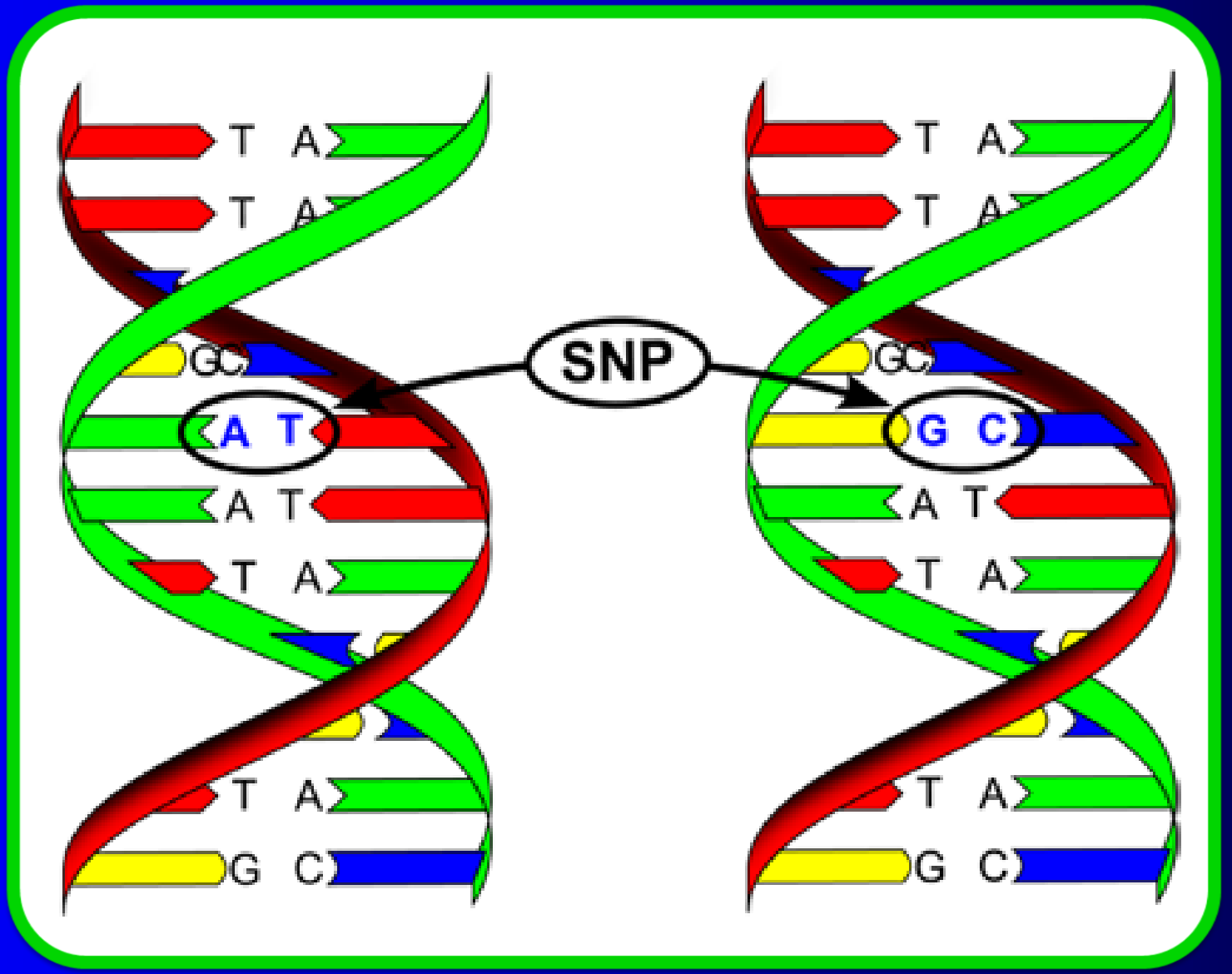
Single nucleotide polymorphisms is called SNPs (pronounced "snips"), are the most common type of genetic variation among people.

Each SNP represents a difference in a single nucleotide.

For example, a SNP may replace the nucleotide cytosine (C) with the nucleotide thymine (T) in a certain stretch of DNA.

B. John Ebenezer

SNIPS



B. John Ebenezer

DNA Finger Printing Procedure

DNA Finger Printing

The technique of DNA Fingerprinting was initially developed by Alec Jeffreys.

He used a satellite DNA as probe that shows very high degree of polymorphism.

It was called as **Variable Number of Tandem Repeats** (VNTR).

The technique involved is Southern blot hybridisation using radiolabelled VNTR as a probe.

B. John Ebenezer

DNA Finger Printing

DNA from a single cell is enough to perform DNA fingerprinting analysis.

Currently, many different probes are used to generate DNA fingerprints.

B. John Ebenezer

Applications of DNA Finger Printing

DNA fingerprinting is used in forensic science to

- Identify the criminals from blood, hair follicle, skin, bone, saliva, sperm.
- Determine paternity.
- Determine population and genetic diversities.
- Verify whether an immigrant is really a close relative of the resident.

B. John Ebenezer

Steps of DNA Finger Printing

(i) Isolation of DNA.

(ii) Digestion of DNA by restriction endonucleases.

(iii) Separation of DNA fragments by electrophoresis.

(iv) Transferring (Southern blotting) of separated DNA fragments to synthetic membranes, such as nitrocellulose or nylon.

(v) Hybridisation using labelled VNTR probe.

(vi) Detection of hybridised DNA fragments by autoradiography.

B. John Ebenezer

Steps of DNA Finger Printing

After hybridisation with VNTR probe, the autoradiogram gives many bands of differing sizes.

These bands give a characteristic pattern for an individual DNA.

It differs from individual to individual in a population except in the case of monozygotic (identical) twins.

The sensitivity of the technique has been increased by use of polymerase chain reaction.

B. John Ebenezer

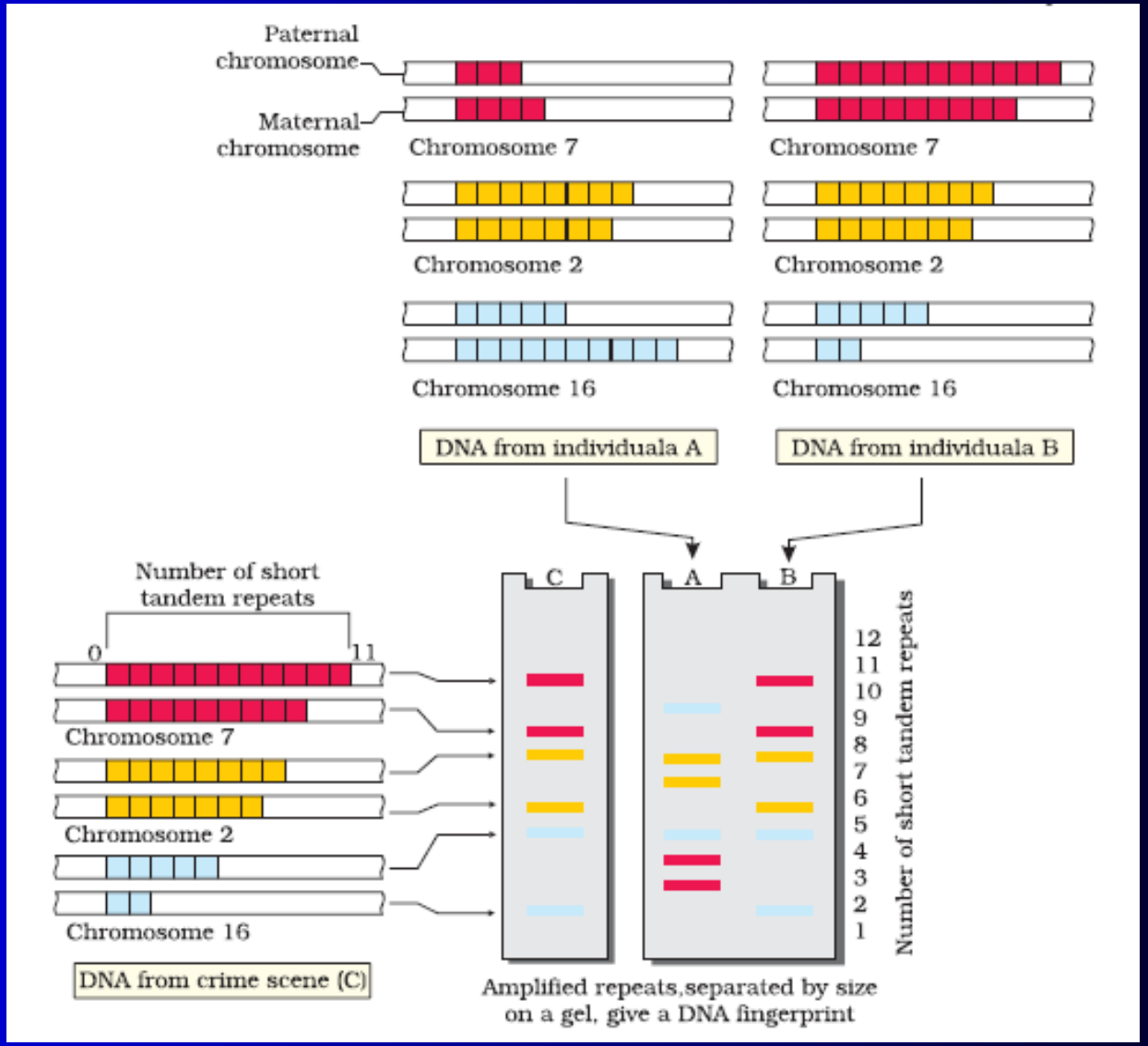
Schematic Representation of DNA Fingerprinting

Few representative chromosomes have been shown to contain different copy number of VNTR.

For the sake of understanding different colour schemes have been used to trace the origin of each band in the gel.

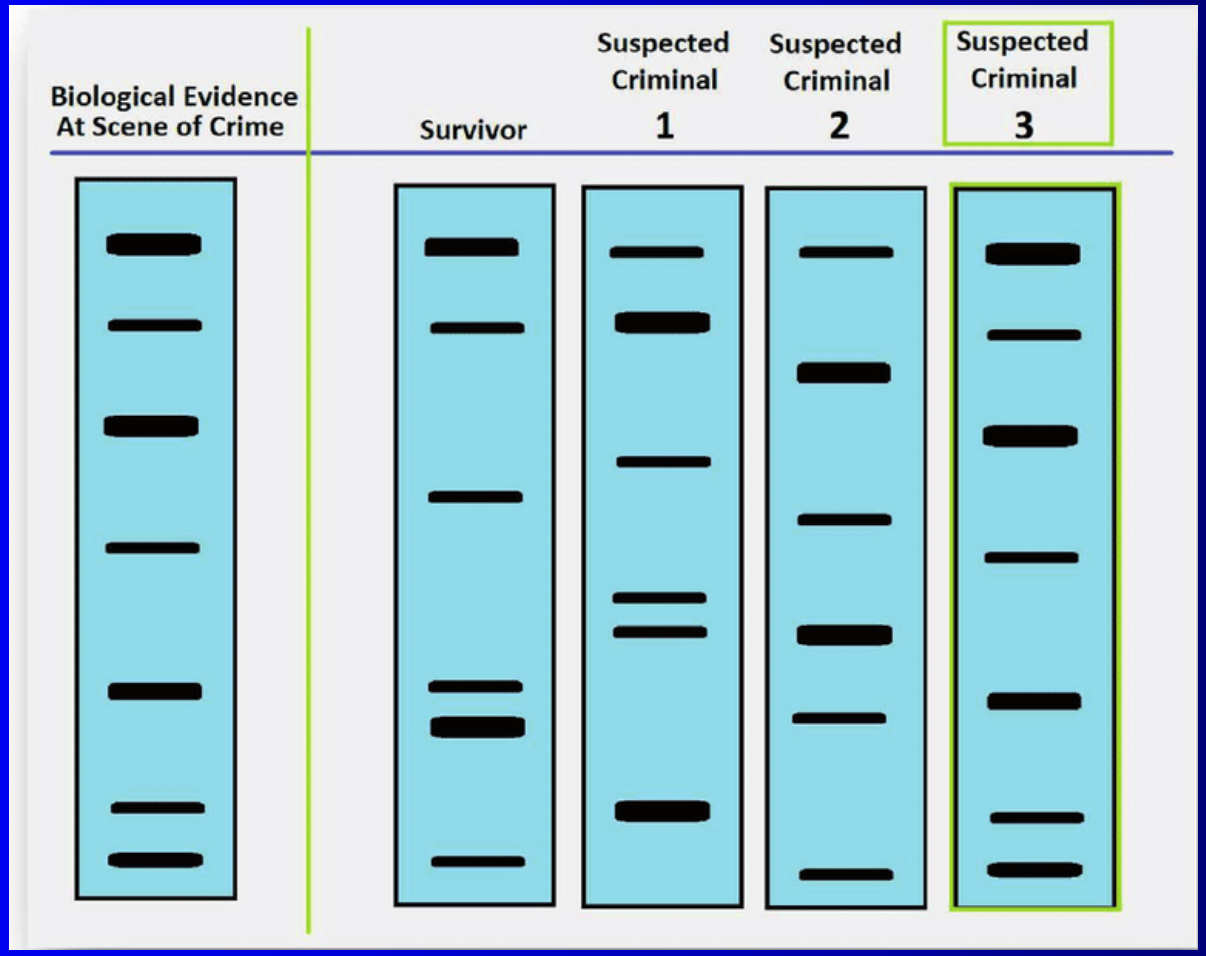
The two alleles (paternal and maternal) of a chromosome also contain different copy numbers of VNTR.

It is clear that the banding pattern of DNA from crime scene matches with individual B, and not with A.



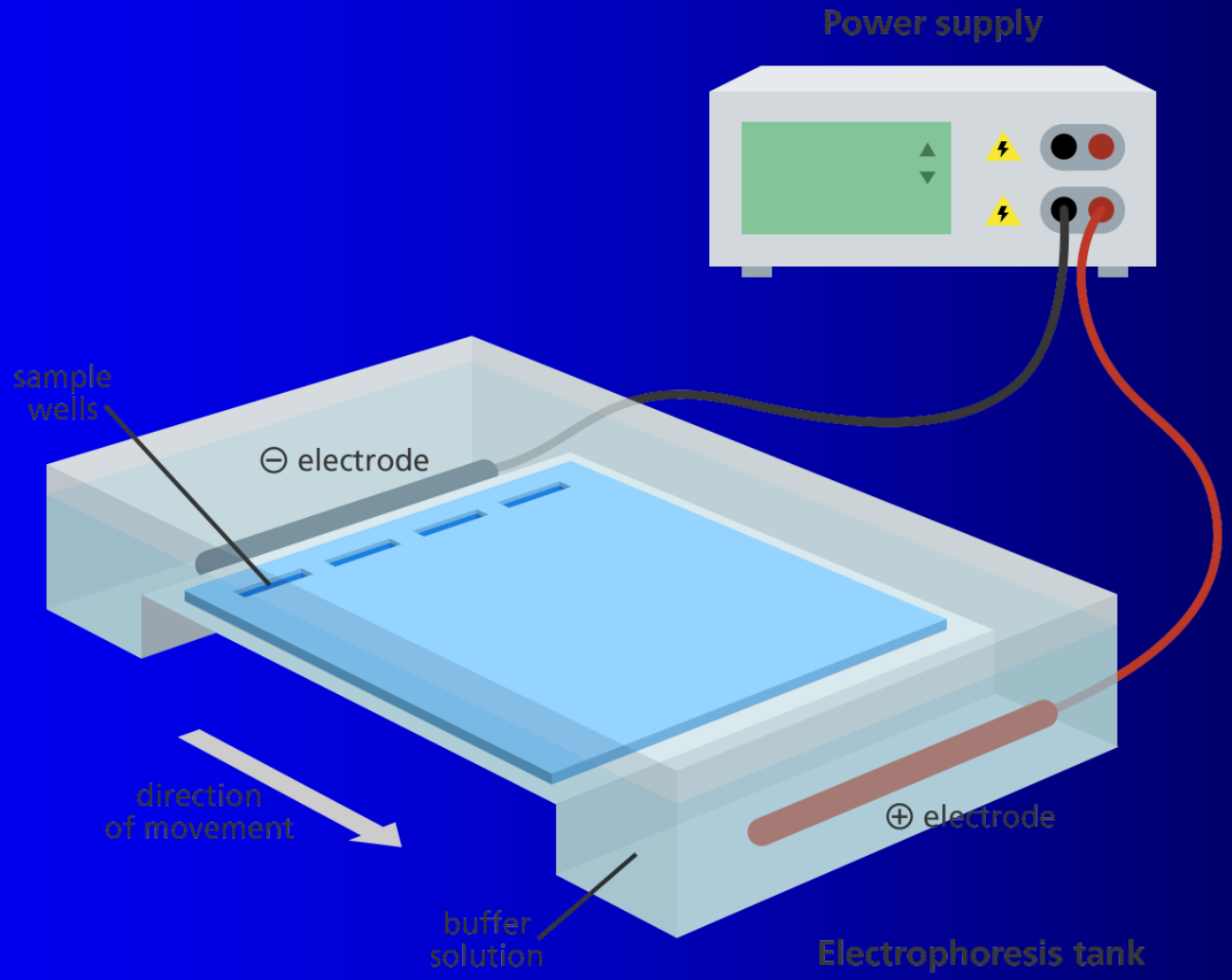
B. John Ebenezer

DNA Bands



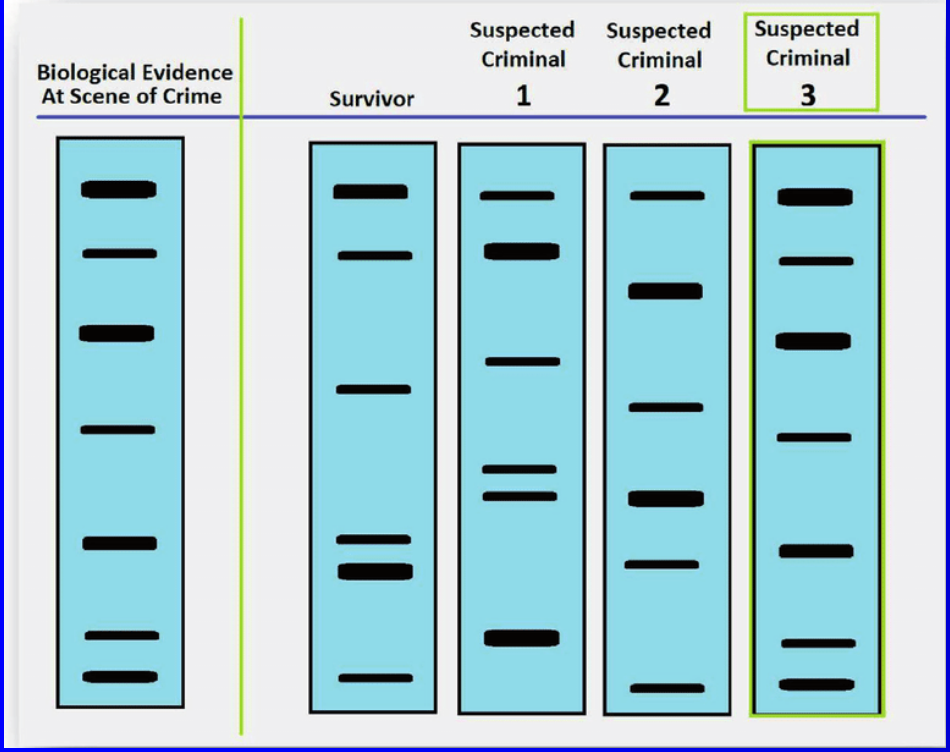
B. John Ebenezer

Electrophoresis



B. John Ebenezer

A crime occurred at a park. Police collected blood and other samples from the crime spot and sent it to the forensic department. The report of the forensic department is as follows.



(a) Identify the criminal from the above suspect criminal list.

(b) Who invented DNA Finger Printing Technology?

(c) What are the other applications of this technology?

B. John Ebenezer

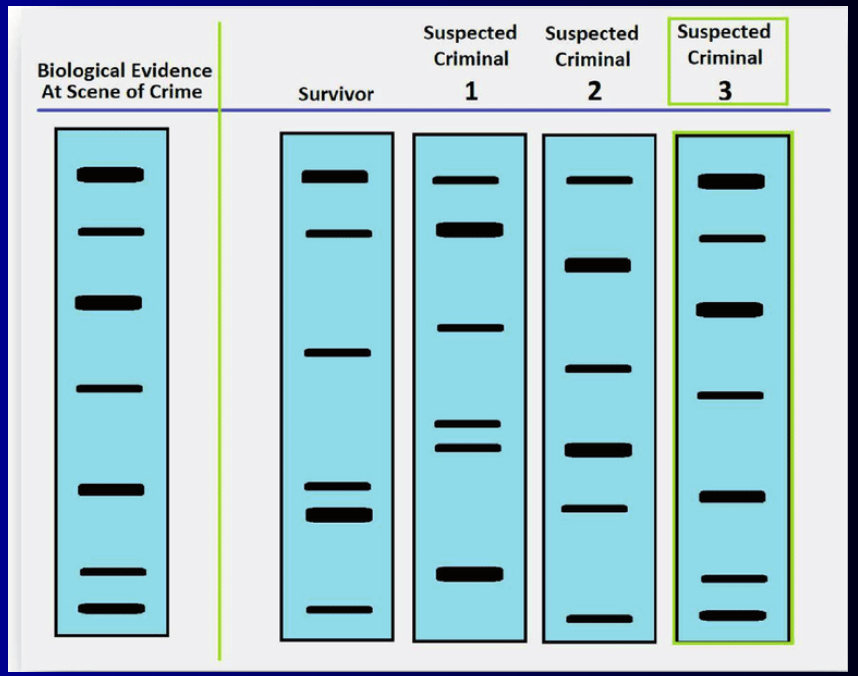
(a) Suspect 3 is the criminal

(b) Alec Jeffreys

(c) Determine paternity.

Determine population and genetic diversities.

Verify whether an immigrant is really a close relative of the resident.



B. John Ebenezer


God Bless You!