MOLECULAR BASIS OF INHERITANCE

Introduction

- Name the two types of nucleic acids found in living systems. **DNA and RNA**
- Name the genetic material found in human. **DNA**
- Genetic material in some viruses is **RNA**

Introduction

- Two types of nucleic acids found in living systems are;
  1) Deoxyribonucleic Acid (DNA) And
  2) Ribonucleic acid (RNA).
- DNA acts as the genetic material in most of the organisms.
- RNA is genetic material in some viruses (Tobacco Mosaic viruses, bacteriophage, etc.).

Introduction

- RNA mostly functions as;
  - ✓messenger,
  - ✓Adapter (tRNA),
  - ✓structural, and
  - ✓catalytic molecule.

DNA

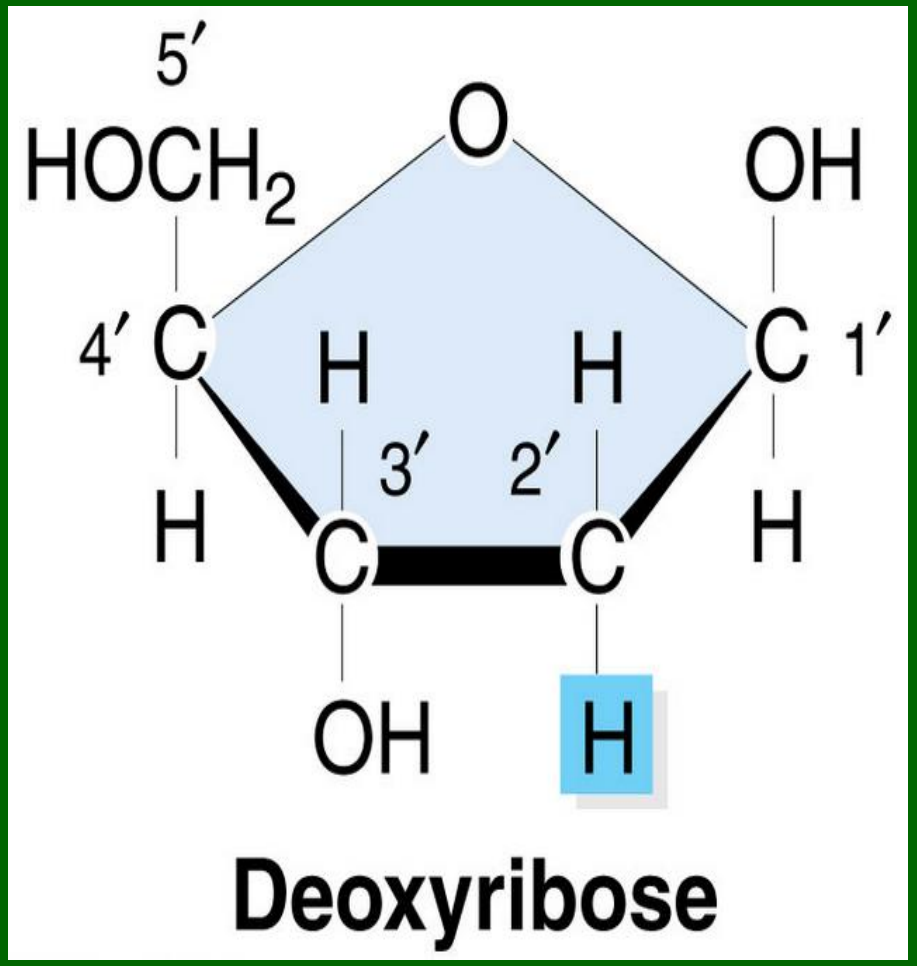
- DNA is a long polymer of deoxyribonucleotides.
- The length of DNA is usually determined by base pairs.
- Base pair is the number of nucleotides/a pair of nucleotide.
- Each organism has its unique length of DNA ;

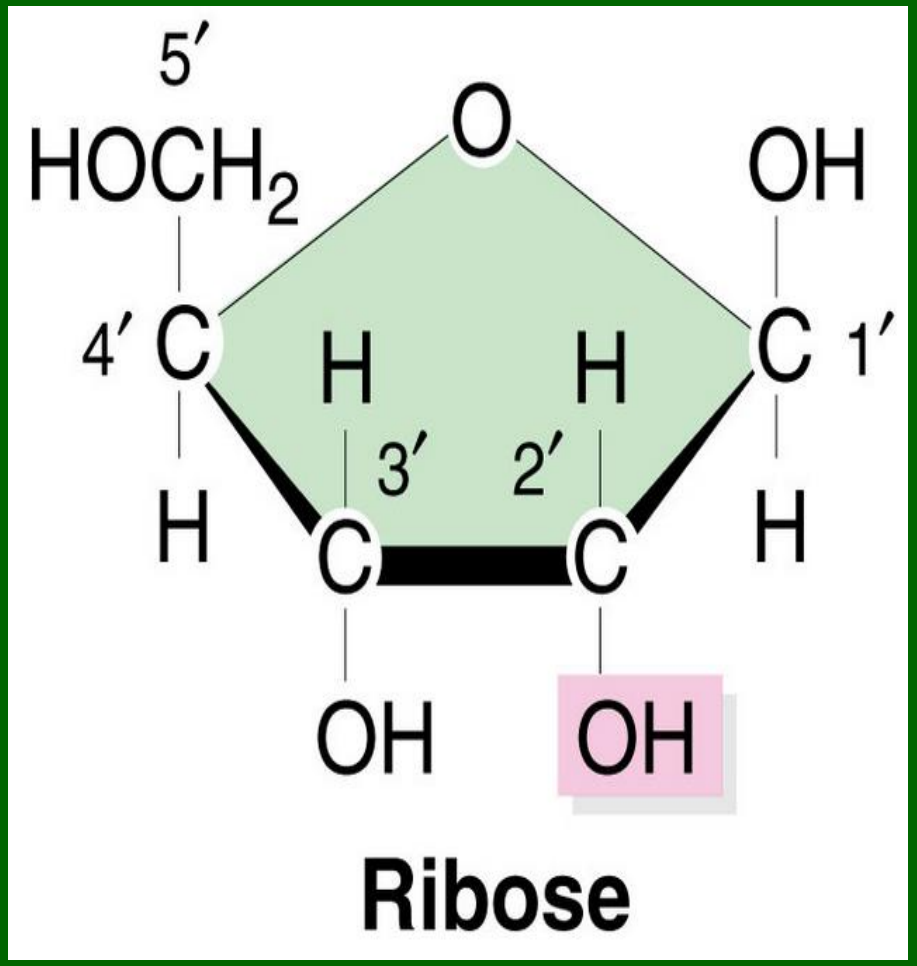
organism	length of DNA
Φ174 (bacteriophage)	5386 nucleotides
Lambda phage (Bacteriophage)	48502 base pairs (bp)
Escherichia coli	4.6×10^6 bp
human DNA (haploid content)	3.3×10^9 bp

Structure of Polynucleotide Chain

- Nucleotide has three components –
  1) a nitrogenous base,
  2) a pentose sugar (ribose in case of RNA, and deoxyribose for DNA), and
  3) a phosphate group.
- There are two types of nitrogenous bases
  1) Purines (Adenine and Guanine), and
  2) Pyrimidines (Cytosine, Uracil and Thymine).
- Cytosine is common for both DNA and RNA
- Thymine is present in DNA.
- Uracil is present in RNA at the place of Thymine.

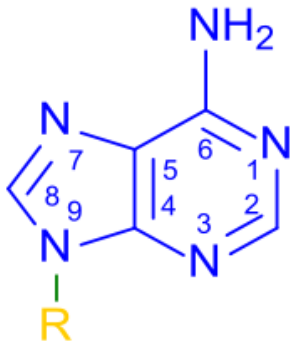
Pentose Sugar



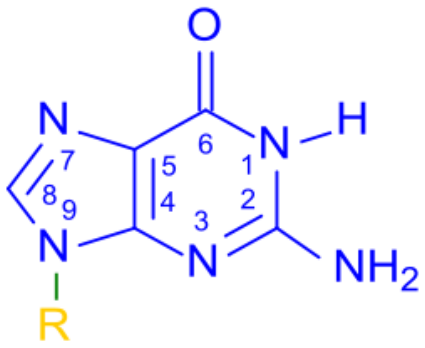


Nitrogenous Bases

Purines

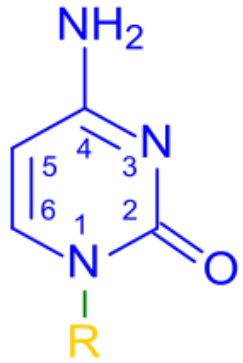


Adenine

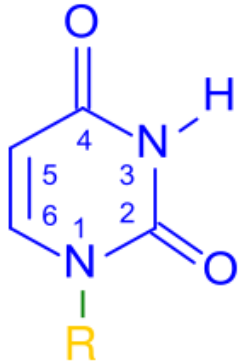


Guanine

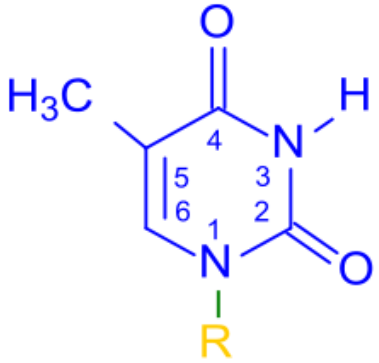
Pyrimidines



Cytosine



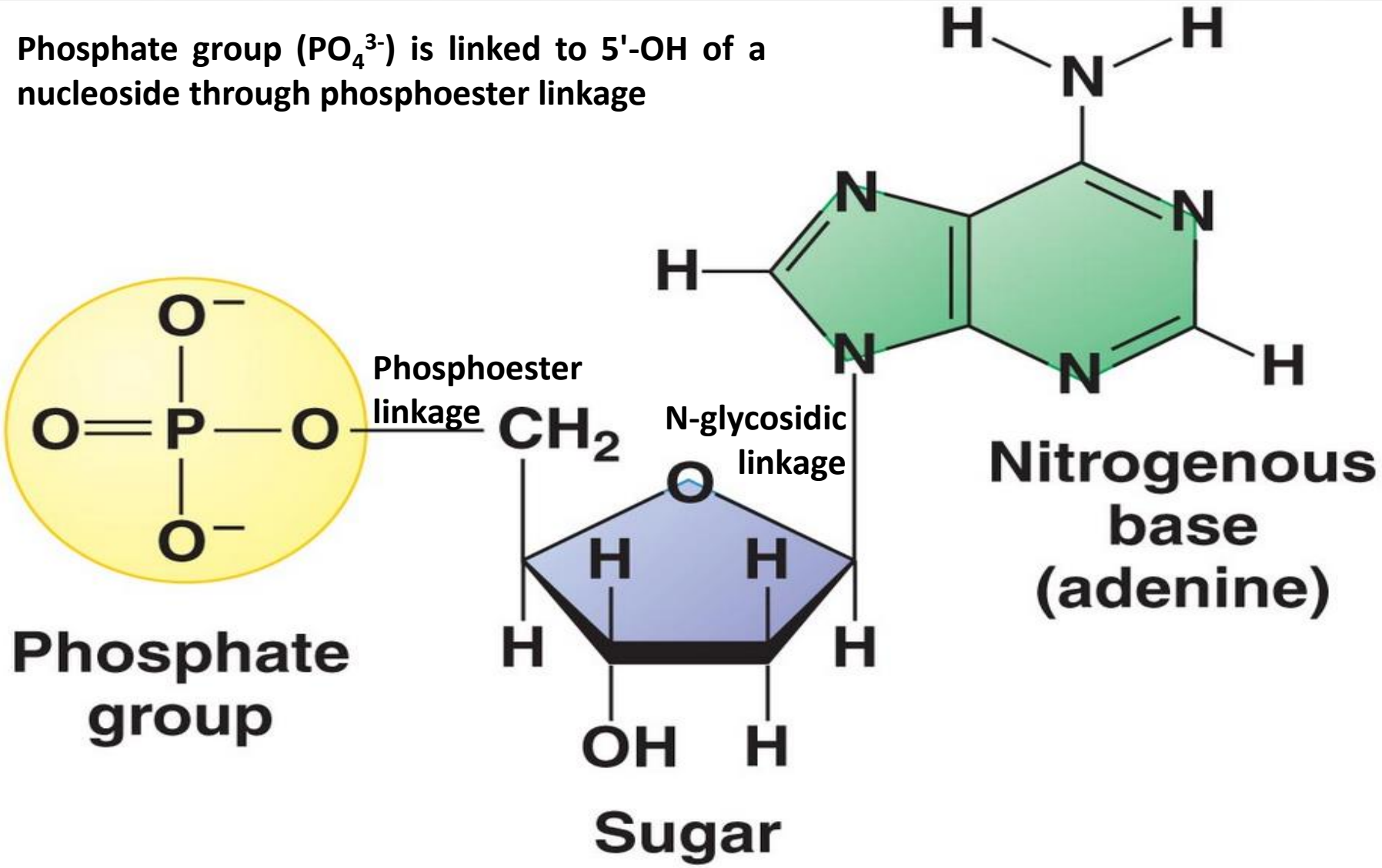
Uracil



Thymine

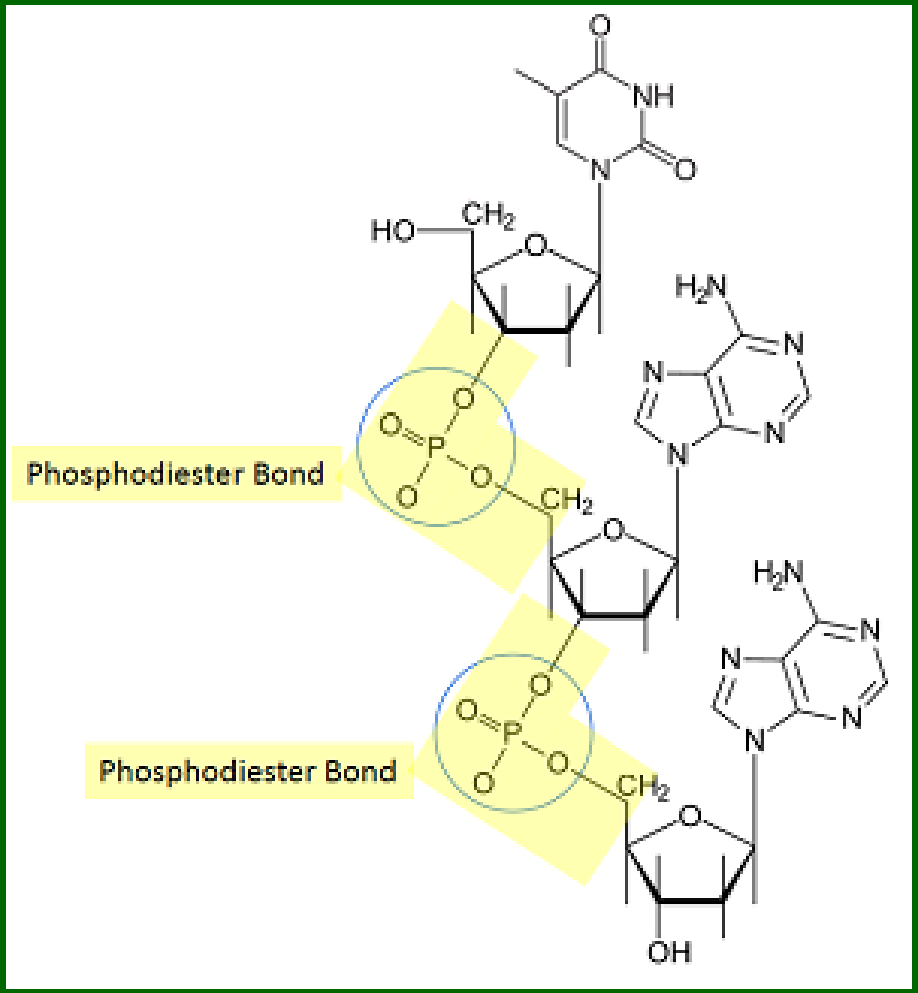
Linkages in three parts of Nucleotide

Phosphate group (PO_4^{3-}) is linked to 5'-OH of a nucleoside through phosphoester linkage

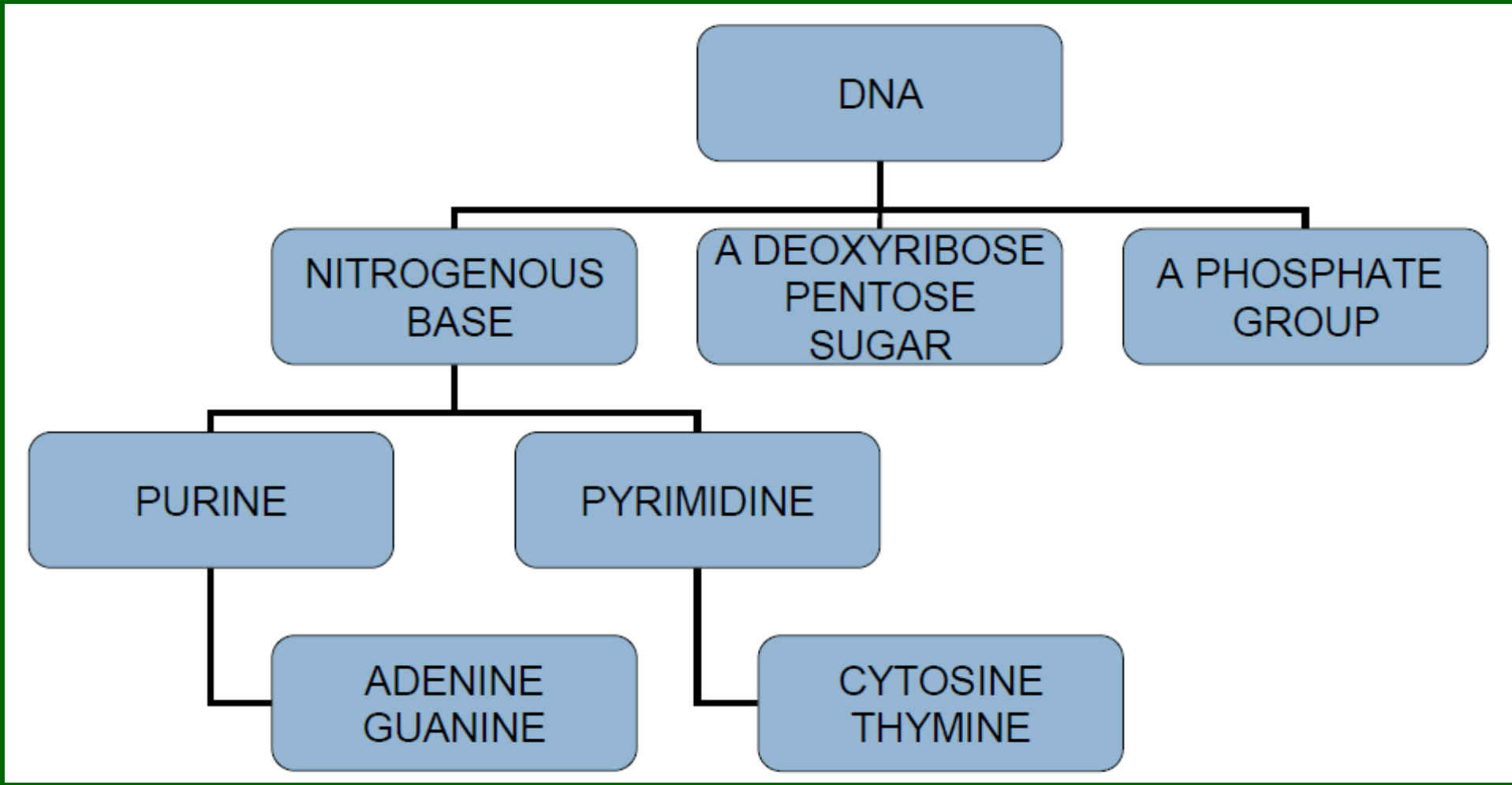


Phosphodiester linkage

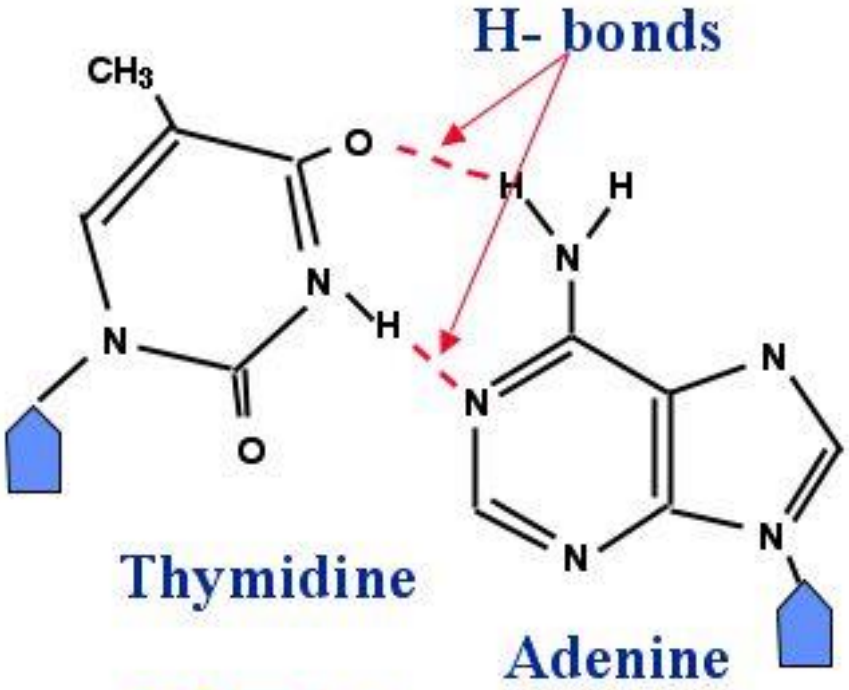
Two nucleotides are linked through 3*'-5'* *phosphodiester linkage to form a dinucleotide*



DNA (Polynucleotide)



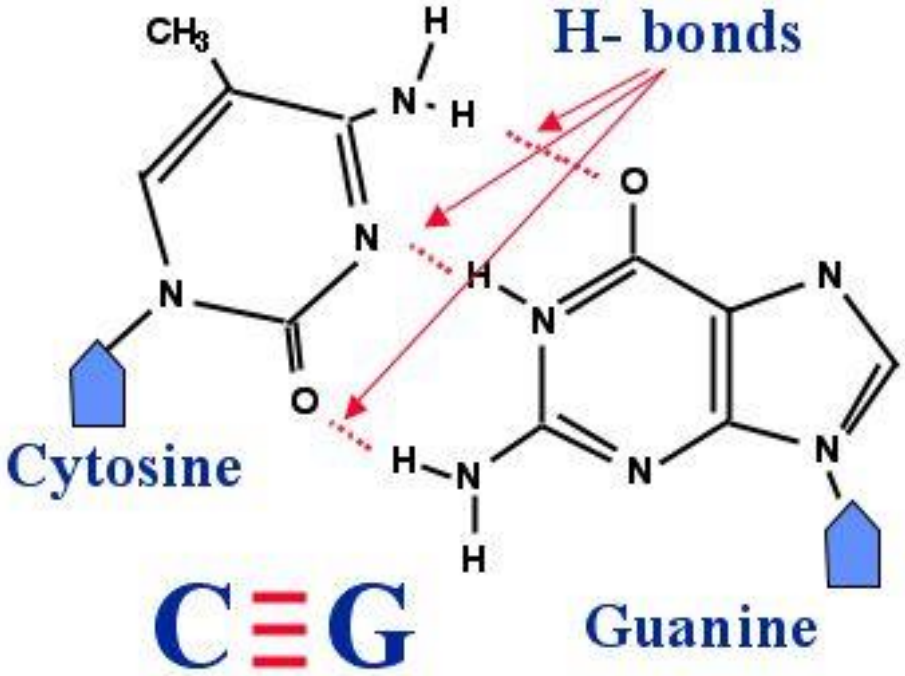
Nucleotide Pairing



Thymidine

Adenine

T = A



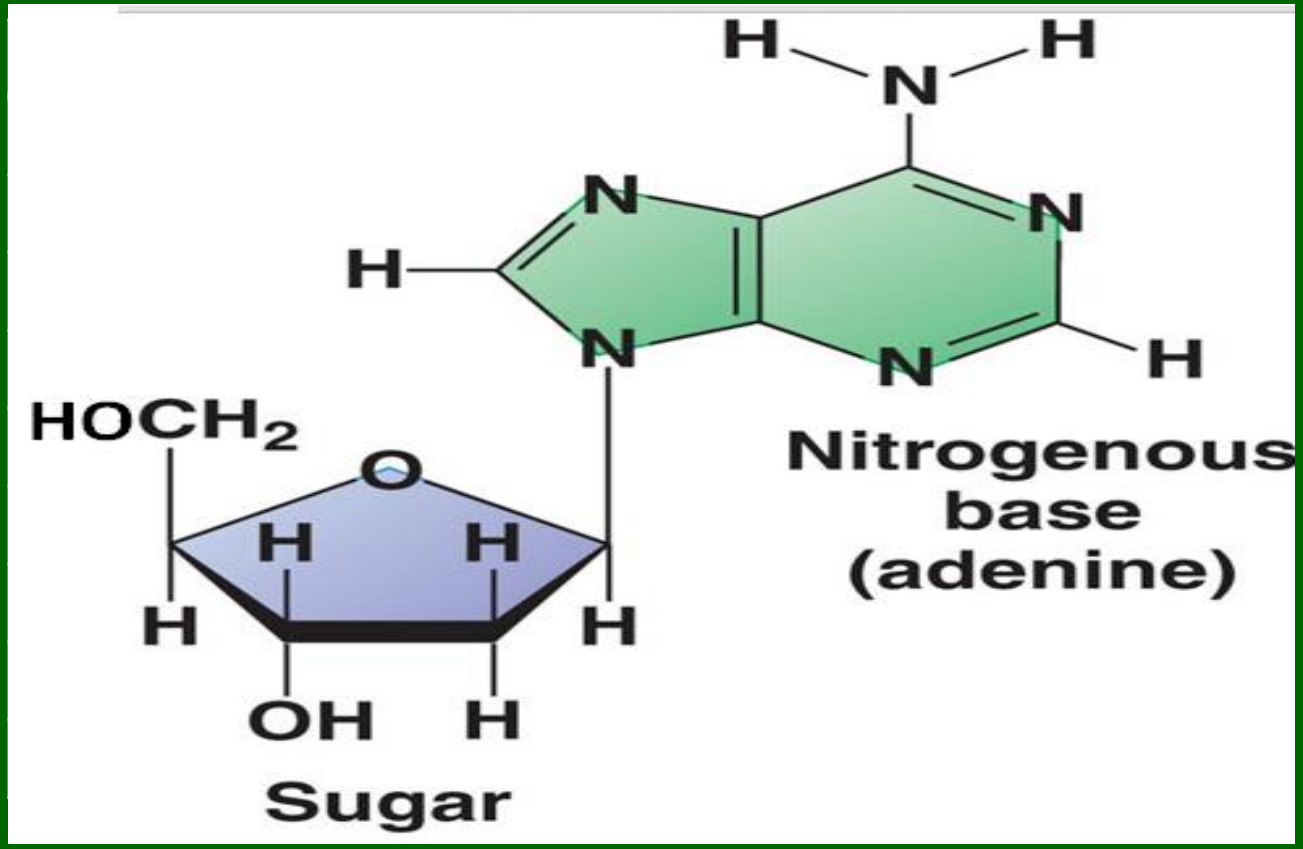
Cytosine

C ≡ G

Guanine

Nucleoside

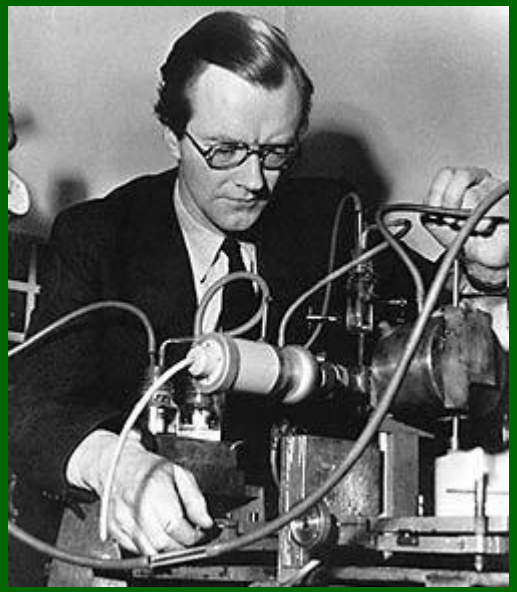
- A nitrogenous base is linked to the pentose sugar through a N-glycosidic linkage to form a nucleoside.



History: *Structure of Polynucleotide Chain*

- DNA as an **acidic substance** present in nucleus was first identified by **Friedrich Meischer** in 1869.
- He named it as '**Nuclein**'.
- Crystallised DNA **X-ray diffraction data** produced by **Maurice Wilkins** and **Rosalind Franklin**.
- In 1953, **James Watson** and **Francis Crick** proposed **Double Helix model** of DNA by using their data.

Scientists behind the **Double Helix model** of DNA



Maurice Wilkins with X-ray crystallographic equipment about 1954.

Rosalind Franklin (Female) in Paris

James Watson and Francis Crick with their DNA model at the Cavendish Laboratories in 1953

Double Helix model of DNA

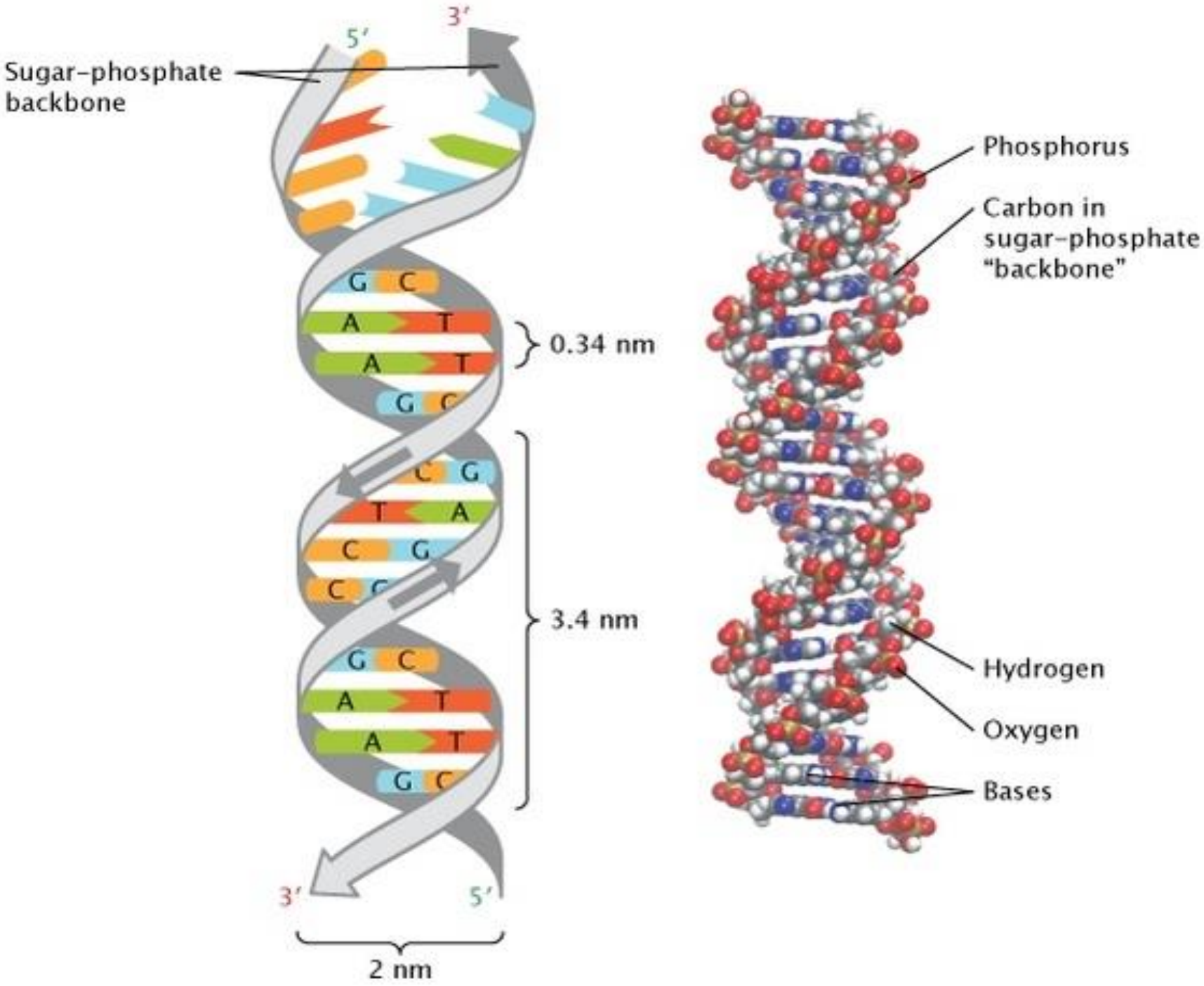
Double helix model of DNA was based on;

- **Complementary base pairing** between the two strands of polynucleotide chains.
- The observation of **Erwin Chargaff** (Chargaff rule)
  - ✓For a double stranded DNA, the ratios between **adenine and thymine and guanine and cytosine** are constant and equals one.

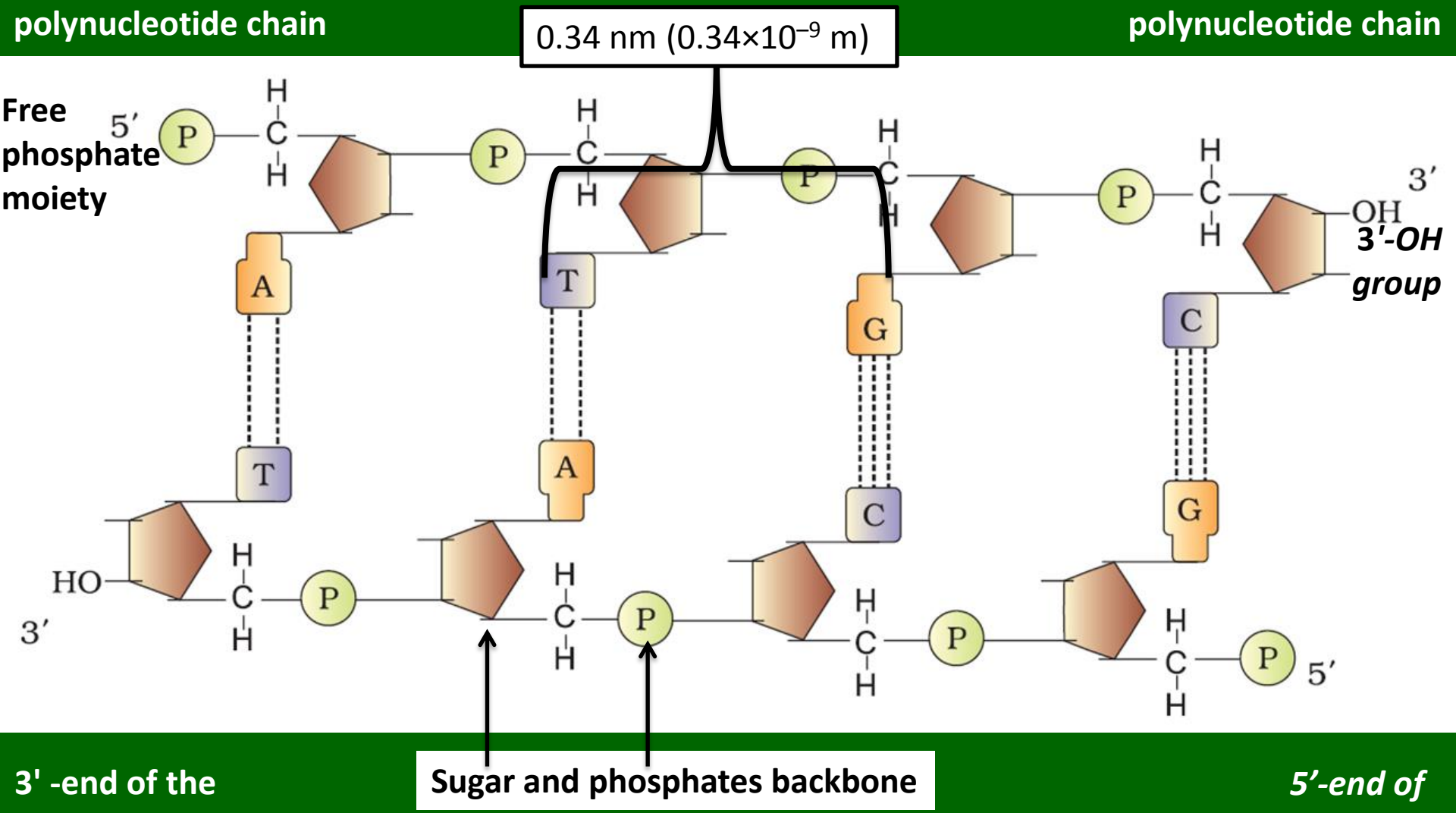
$$\frac{\mathbf{A} + \mathbf{T}}{\mathbf{G} + \mathbf{C}} = \mathbf{1}$$

Salient features of the double-helix structure of DNA proposed by Watson and Crick

- Open page number 97 and try to recall the structure of DNA
- 5 minutes

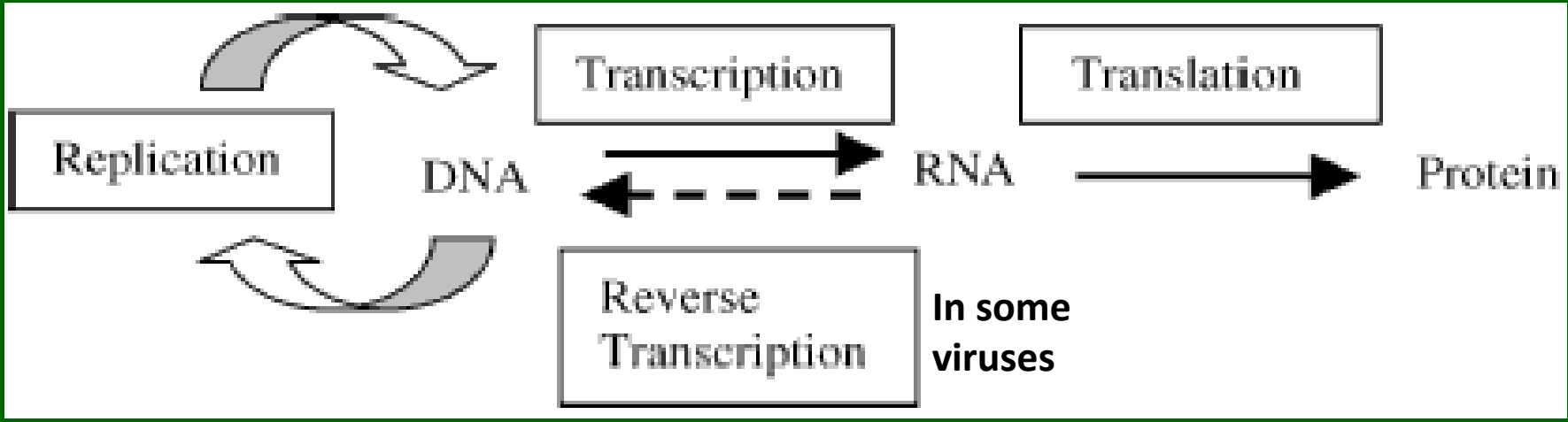

Sugar–phosphate backbone
5′
3′
G C
A T
A T
G C
0.34 nm
C G
T A
C G
C G
3.4 nm
G C
A T
A T
G C
3′
5′
2 nm
Phosphorus
Carbon in sugar–phosphate "backbone"
Hydrogen
Oxygen
Bases

Struture of Structure of Polynucleotide Chain (DNA)

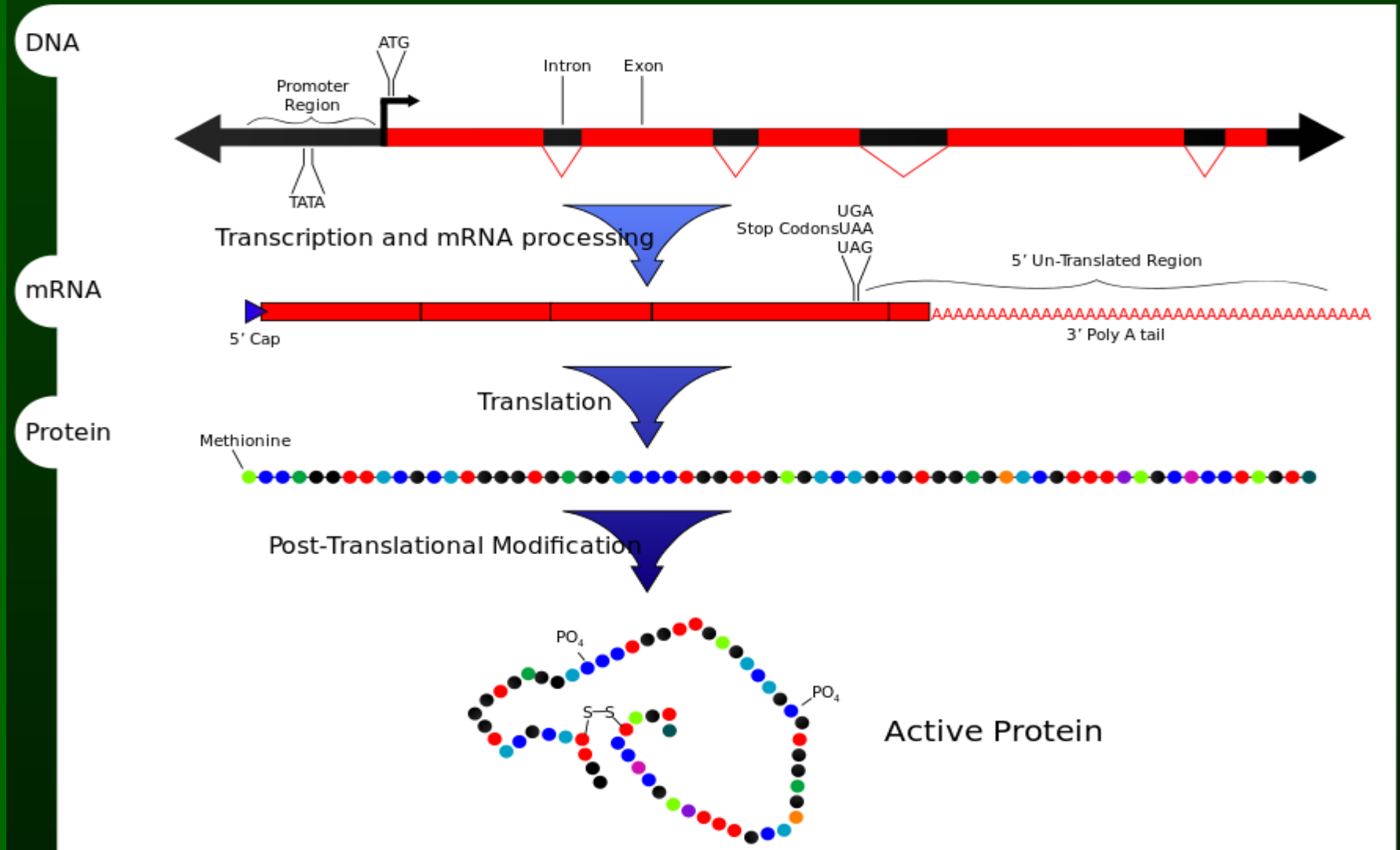


Central dogma in molecular biology

- Soon after the proposition of double helix structure for DNA, **Francis Crick** proposed the **Central dogma(doctrine/principle) in molecular biology**.
- **Central dogma in molecular biology** states that;
  - ✓ the genetic information flows from **DNA⟶RNA ⟶ Protein**.
  - ✓ In some viruses the flow of information is in reverse direction, that is, from RNA to DNA.



Central Dogma of molecular biology:Eukaryotic mode



Packaging of DNA Helix

Inclass exercise:

- Number of nitrogenous base pairs in a mammalian cell is 6.6×10^9 bp. What is the length of the DNA in meters? [Hint: *Distance between two consecutive base pairs of the DNA = 0.34×10^{-9} m/bp*),

Answer: 6.6×10^9 bp $\times$ 0.34×10^{-9} m/bp = **2.2 metres.**

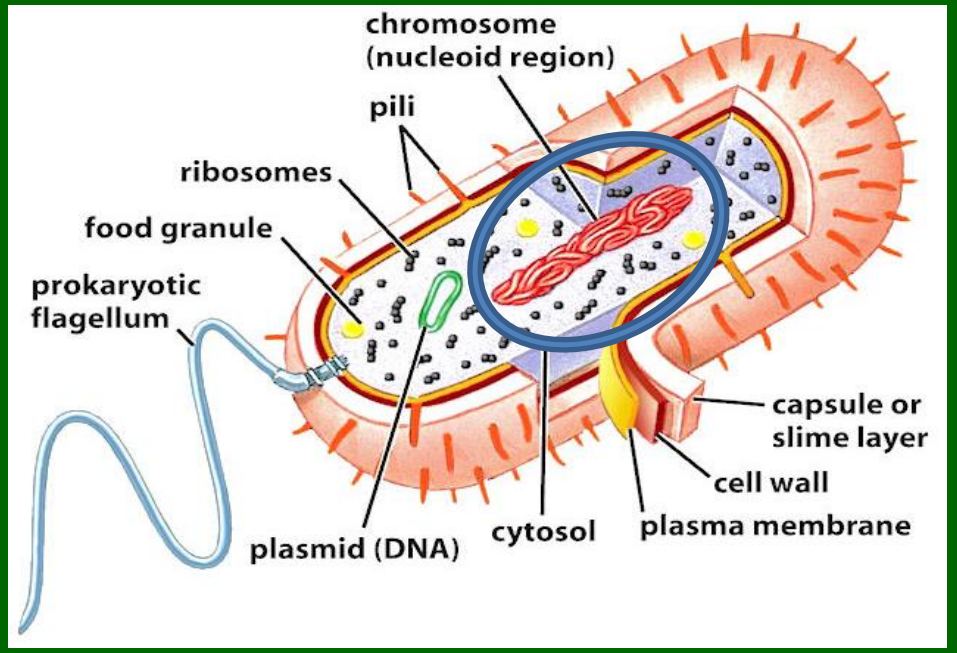
An interesting fact

Dimension of a typical nucleus = **10^{-6} m** ($<$ **2.2 m**) [approximately]

How is 2.2 metres long polymer packaged in the nucleus 10^{-6} m ($<$ 2.2 m) of cell?

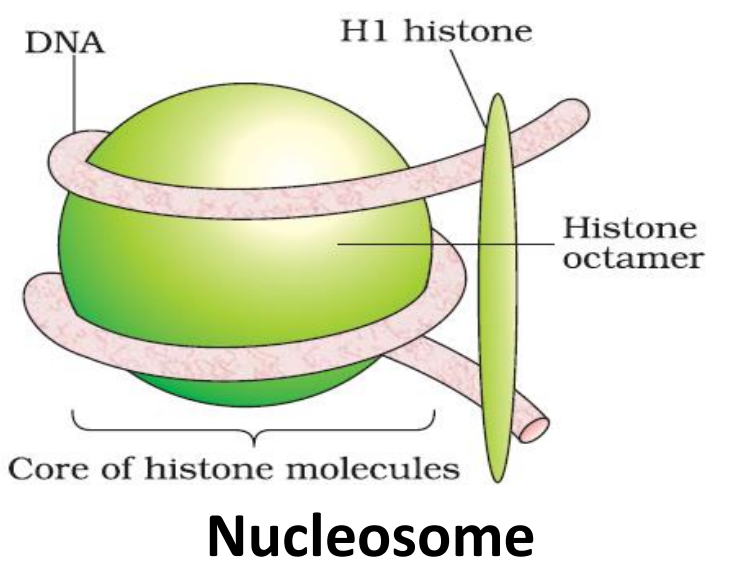
Packaging of DNA Helix In prokaryotes (like E. coli)

- Negatively charged **DNA** is held with some positively charged **non-histone proteins** (like polyamines) in the region of **nucleoid**.

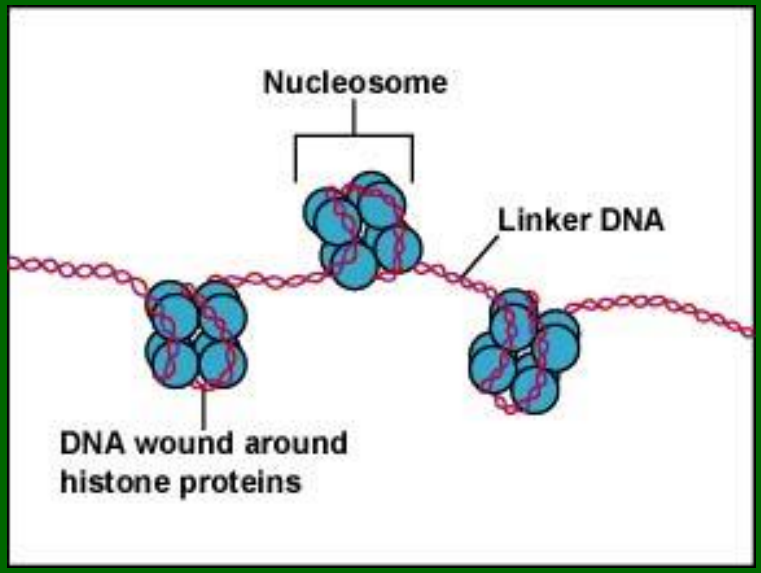


Packaging of DNA Helix In eukaryotes

- Negatively charged DNA is held with positively charged histone (octomer) proteins to form **Nucleosome**.
- **Histones are positively charged due to** rich in basic amino acid residues **lysines** and **arginines**

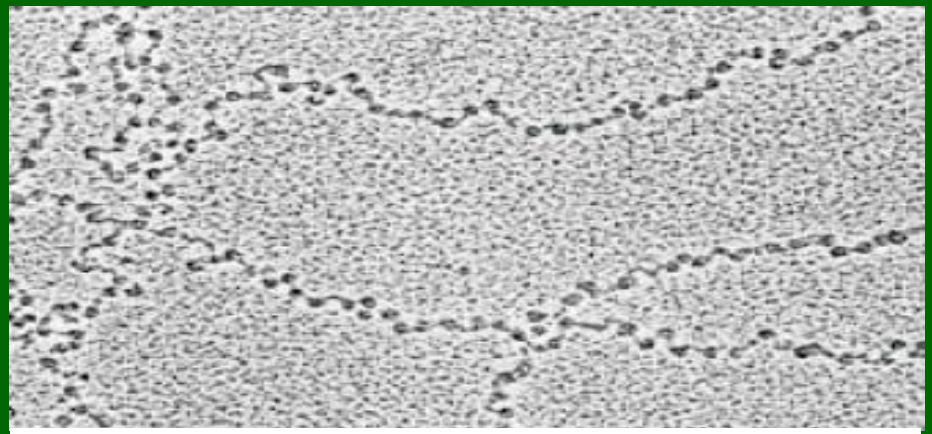


Nucleosome



Nucleosome

- The negatively charged DNA (**200 bp approx.**) is wrapped around the positively charged histone octamer to form a **nucleosome.**
- Nucleosomes repeated to form **chromatin.**
- The nucleosomes in chromatin are seen as '**beads-on-string**' structure when viewed under electron microscope



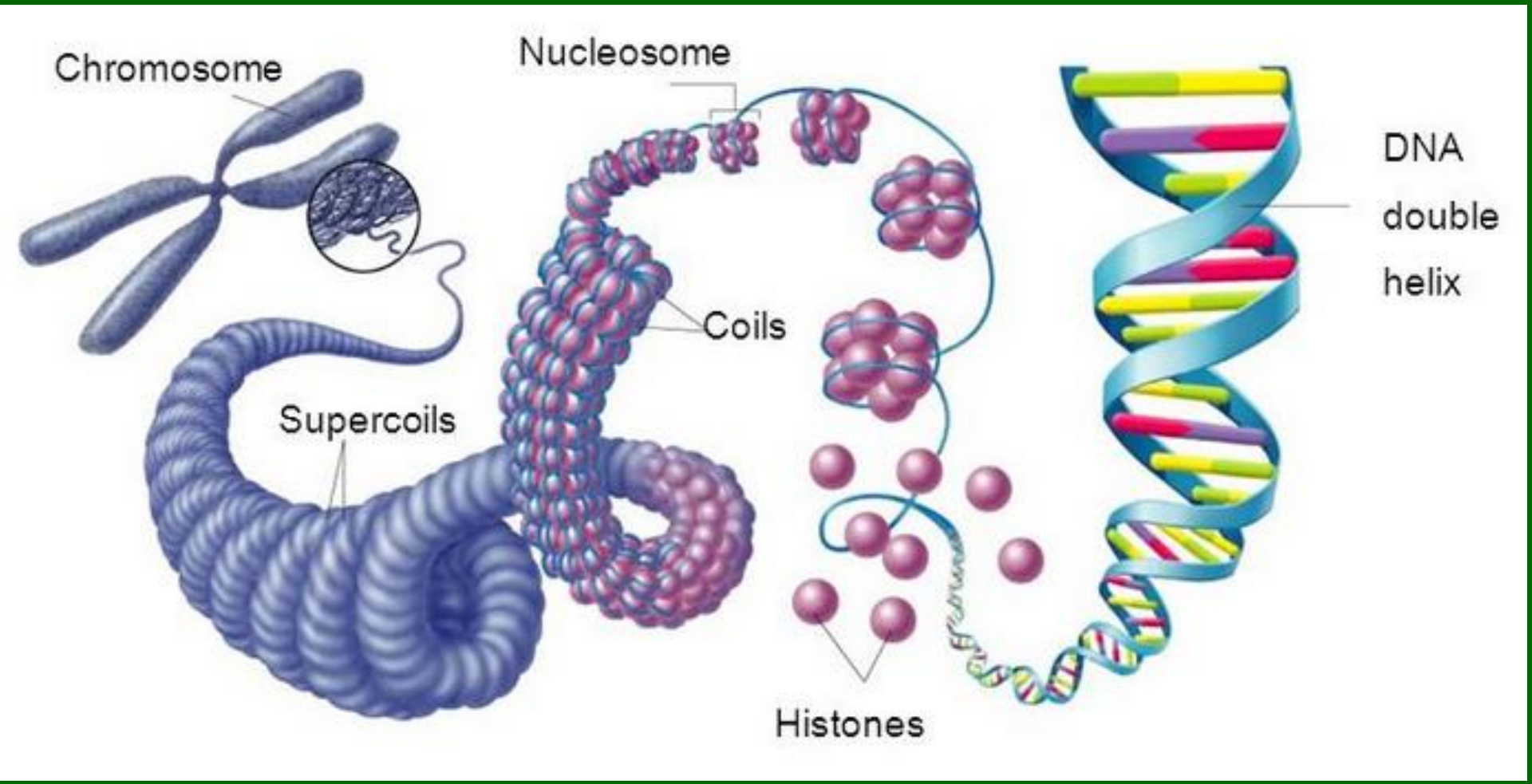
EM picture - 'Beads-on-String'

Chromatin

- The nucleosome requires **Non-histone Chromosomal (NHC) proteins for higher level packaging.**

Chromatin	
Euchromatin	**Heterochromatin**
loosely packed chromatin in the nucleus (stains light)	more densely packed chromatin (stains dark)
Transcriptionally active chromatin.	Transcriptionally inactive

Eukaryotic chromosome = DNA + Histone protein + Non-histone chromosomal protein



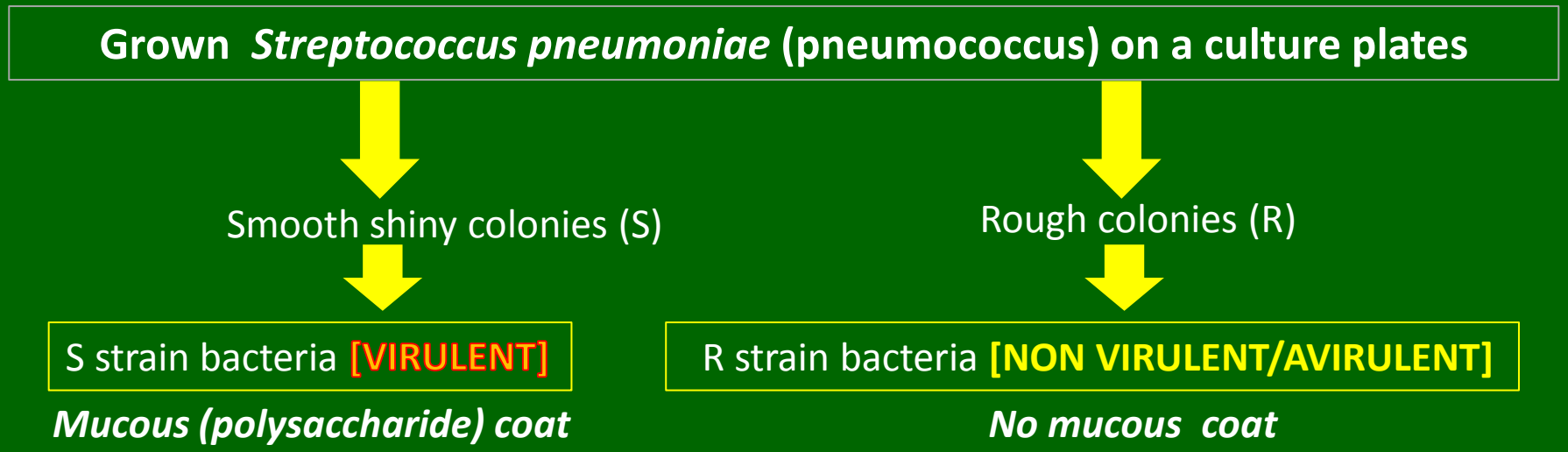
Picture comprehension

Picture comprehension

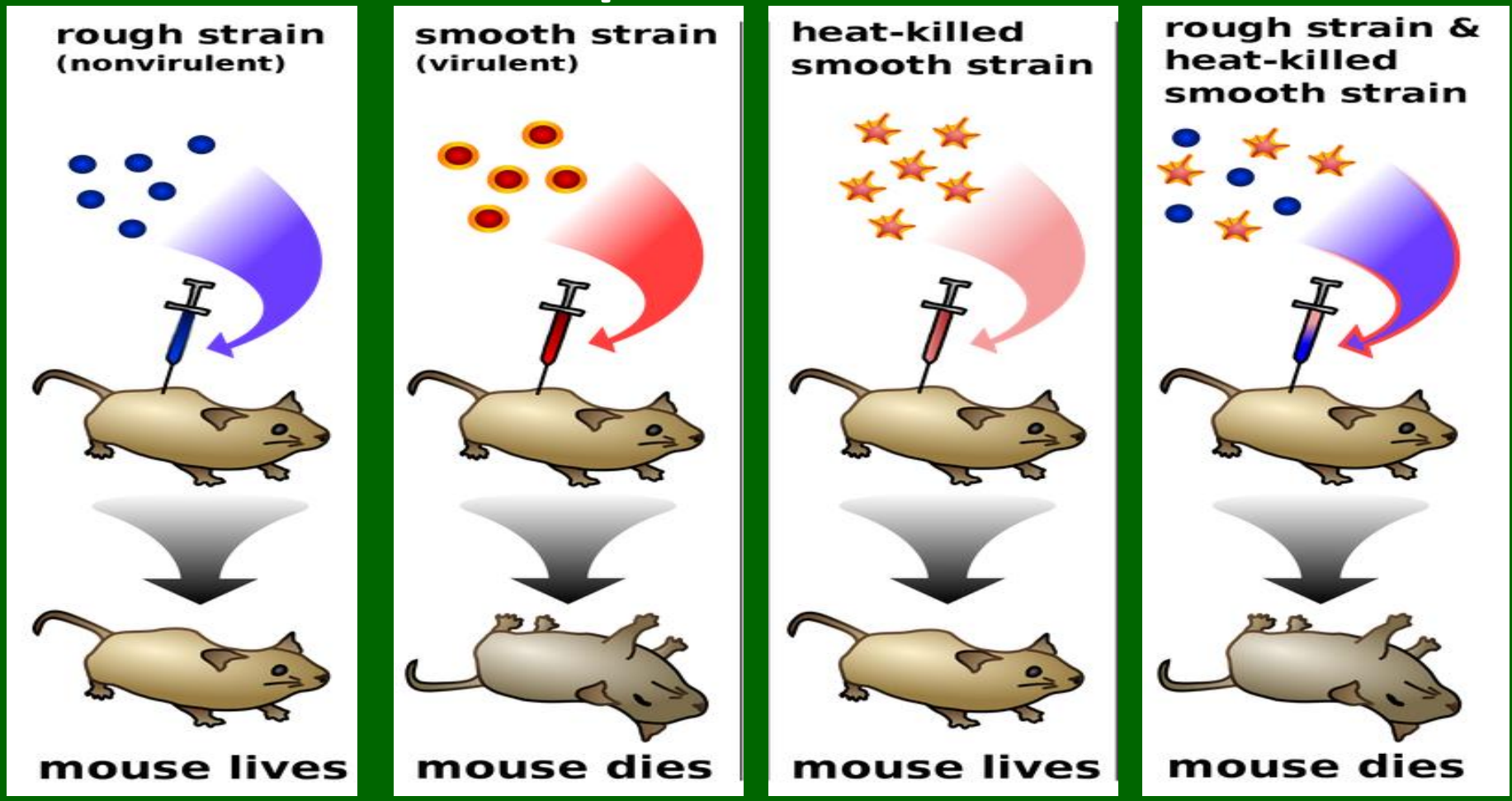
The Search For Genetic Material

Transforming Principle

- In **1928, Frederick Griffith** conducted experiments with ***Streptococcus pneumoniae*** *(bacterium responsible for pneumonia).*
- **Experiment;**



Continuation of experiment of Frederick Griffith



What conclusion can you draw from the experiment?

Transforming Principle of Frederick Griffith

Conclusion:

- R strain bacteria was **transformed by the heat-killed S strain bacteria.**
- **Some 'transforming principle'**, transferred from the heat-killed S strain, had 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, Griffith **was not able to know** the **biochemical nature** of genetic material.

Biochemical Characterisation of Transforming Principle

- Earlier, the genetic material was thought to be a protein.
- In 1933-44 , **Oswald Avery, Colin MacLeod and Maclyn McCarty** worked to determine the biochemical nature of ‘transforming principle’ in Griffith's experiment.

In search of 'Biochemical nature of Transforming Principle' by Oswald Avery, Colin MacLeod and Maclyn McCarty

Experiment

Purified biochemicals (proteins, DNA, & RNA) from the heat killed S cells

↓

Checked for the biochemicals responsible for transforming R cell into virulent S cells

Protein extract of heat killed S cells + R cells = No S cells (Not transformed)
RNA extract of heat killed S cells + R cells = No S cells (Not transformed)
DNA extract of heat killed S cells + R cells = S cells (Transformed)

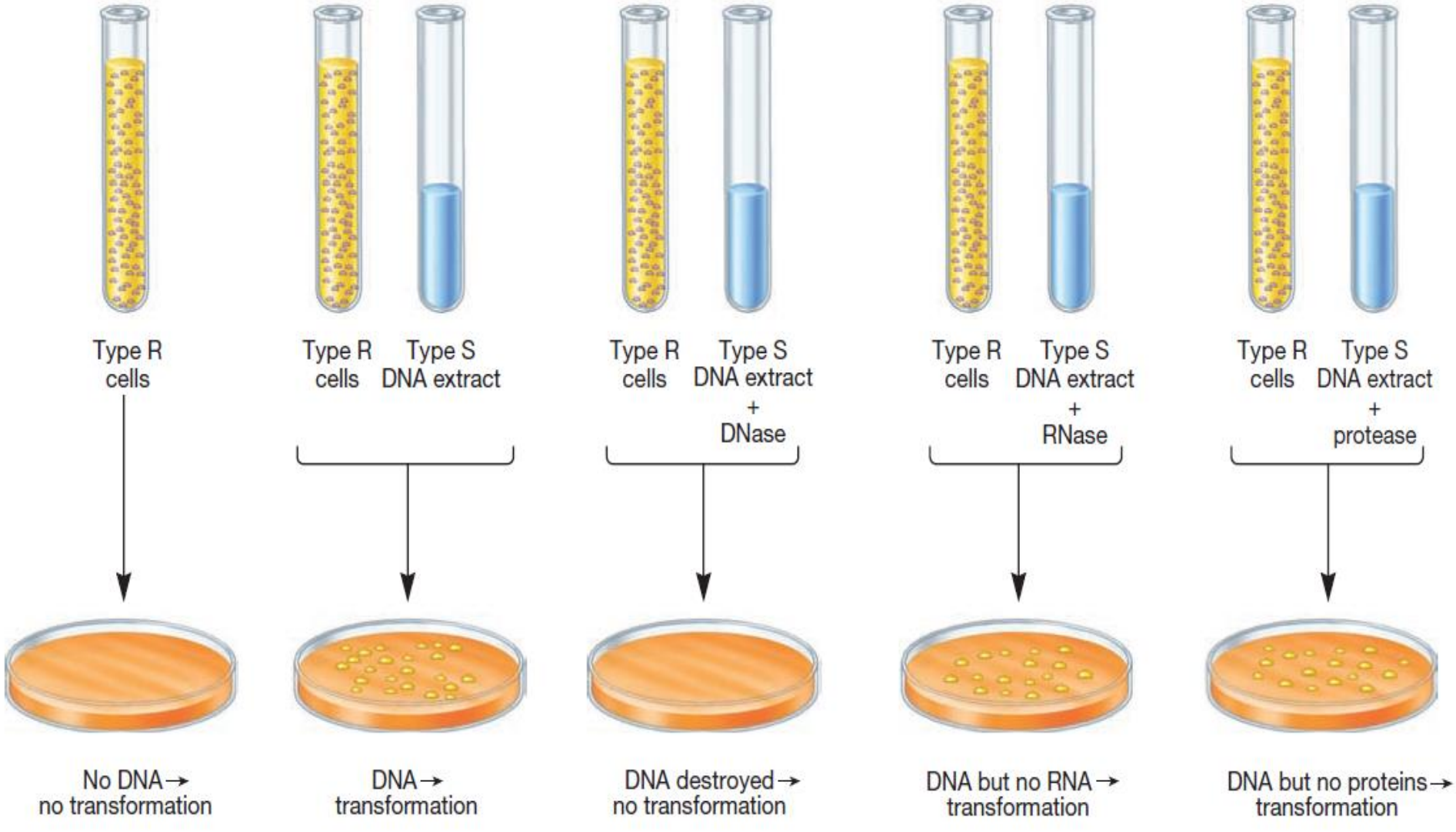
DNA extract of heat killed S cells + R cells + Proteases = S cells (Transformed)
DNA extract of heat killed S cells + R cells + RNases = S cells (Transformed)
DNA extract of heat killed S cells + R cells + DNases = No S cells (Not transformed)

↓

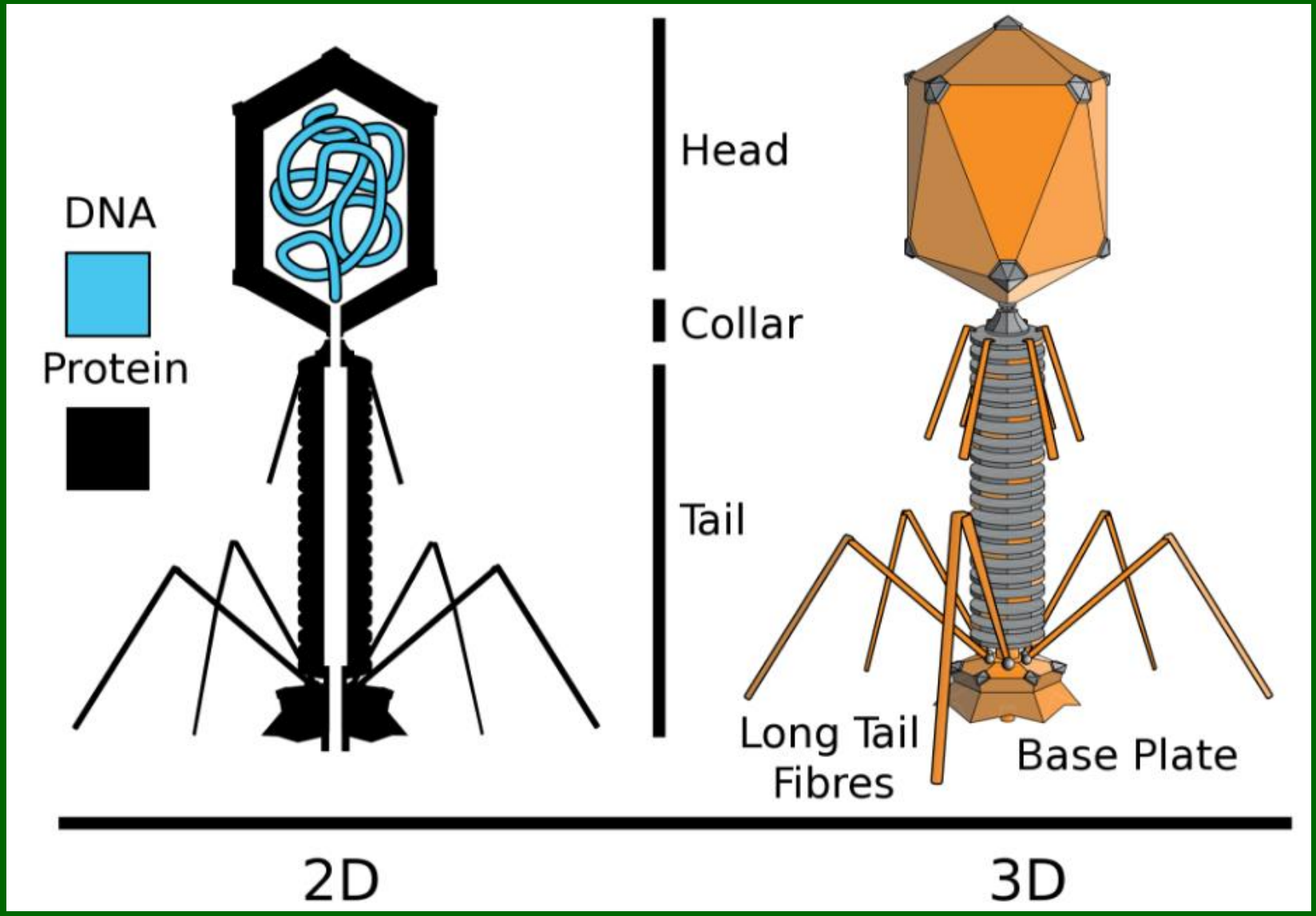
Concluded that DNA is the hereditary material

→ Not all biologists were convinced.

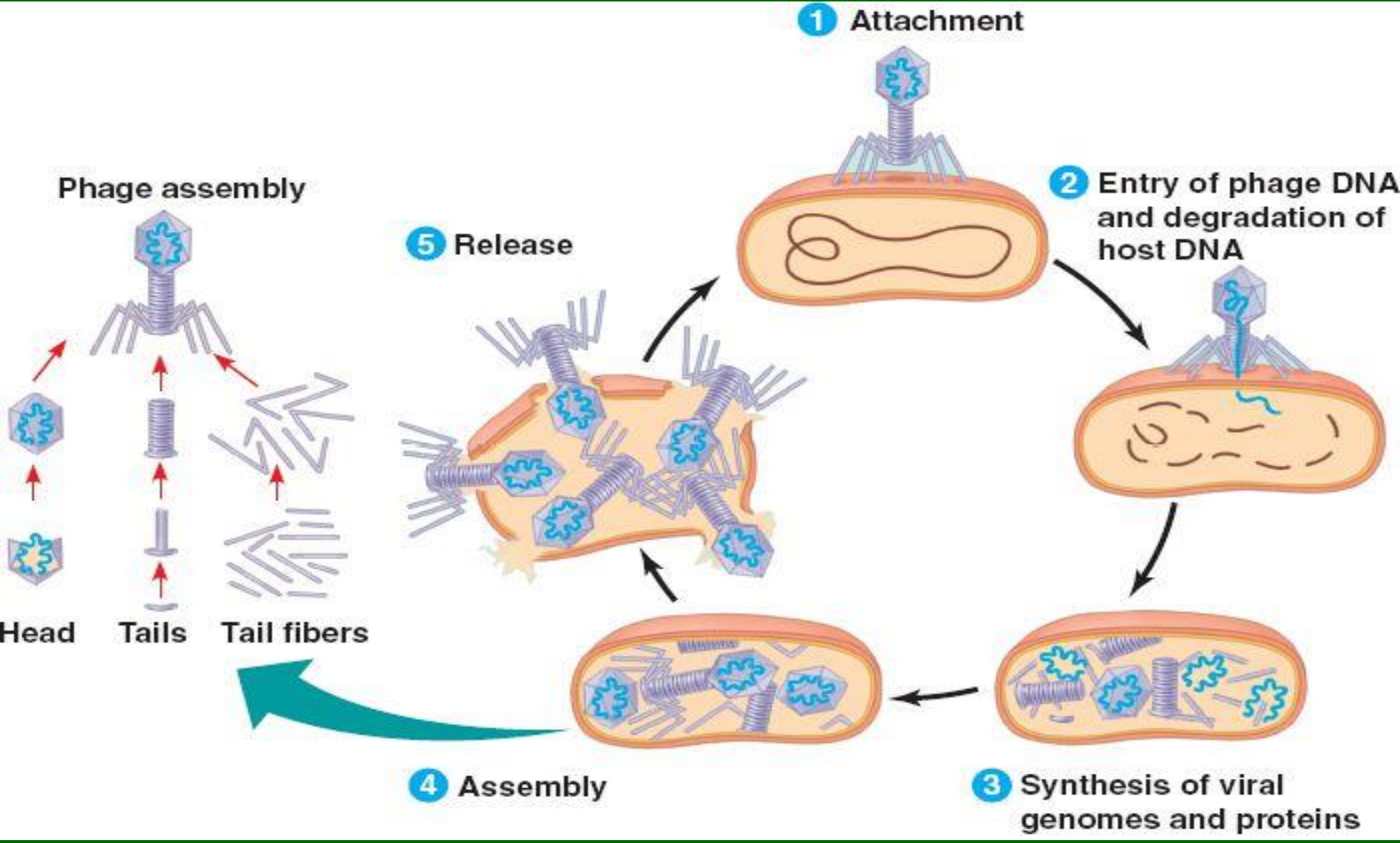
Experiment of Oswald Avery, Colin MacLeod and Maclyn McCarty on 'Biochemical nature of Transforming Principle'



BACTERIOPHAGE



The Life Cycle of Bacteriophage T4



The Genetic Material is DNA

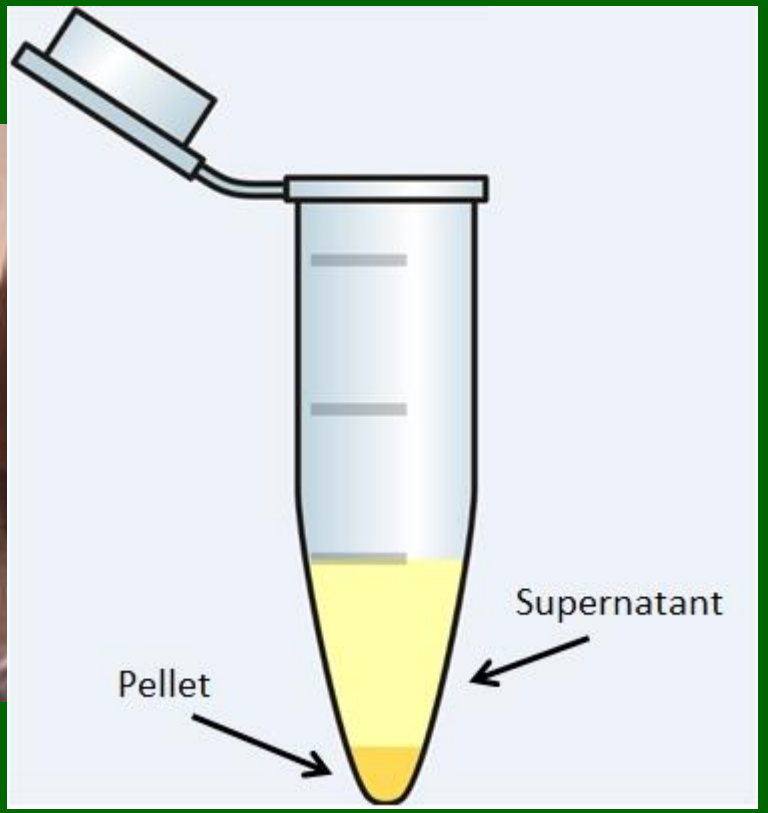
- Alfred Hershey and Martha Chase (1952) researched on **bacteriophages** and proved that **DNA is the genetic material**.
- labeled DNA and protein of bacteriophages with different radioisotopes like ^{32}P and ^{35}S respectively.
- Worked to discover whether it was protein or DNA from the viruses that entered the bacteria.
- It was found that DNA from the viruses entered the bacteria.

Martha Chase Alfred Hershey

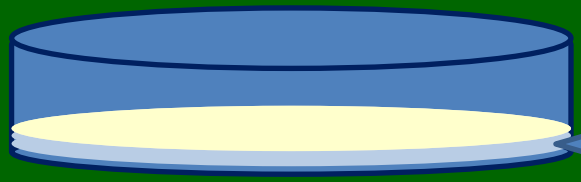
Blender

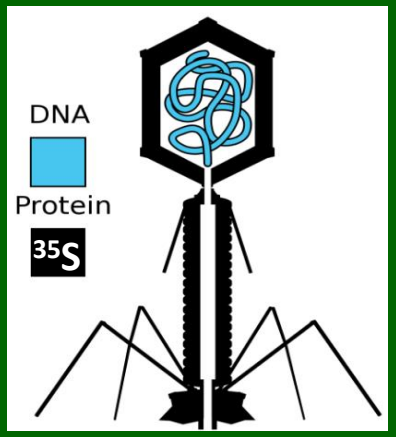
Centrifuge tubes

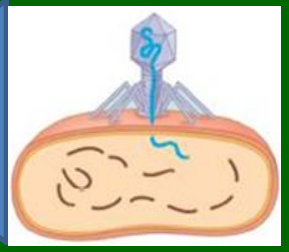
Centrifuge



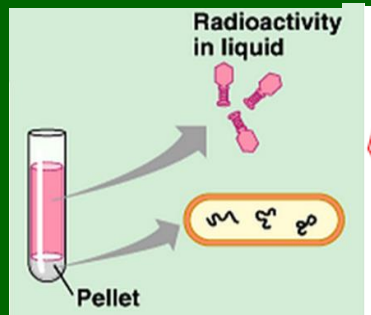
Hershey and Chase's experiment







Centrifuge

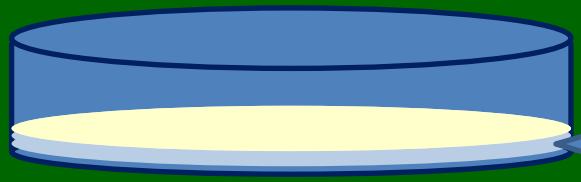


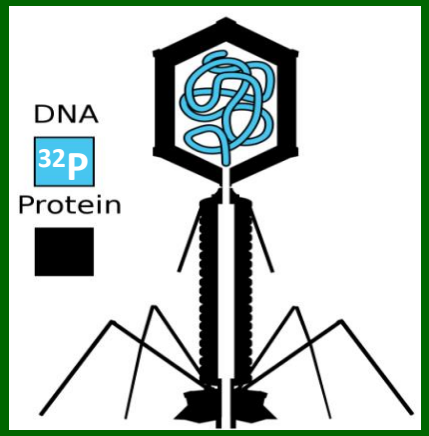
No Radioactive (^{35}S) detected in cells

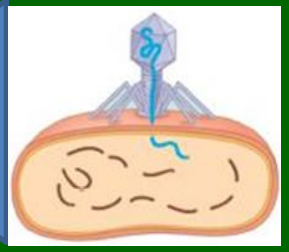
+

Radioactive (^{35}S) detected in supernatant

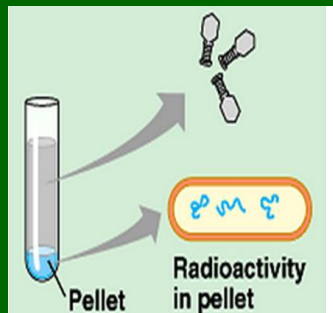
Hershey and Chase's experiment







Centrifuge



Radioactive (^{32}P) detected in cells

+

No Radioactivity detected in supernatant

Hershey and Chase's experiment

Grew some bacteriophages on medium that contained radioactive sulfur (^{35}S)

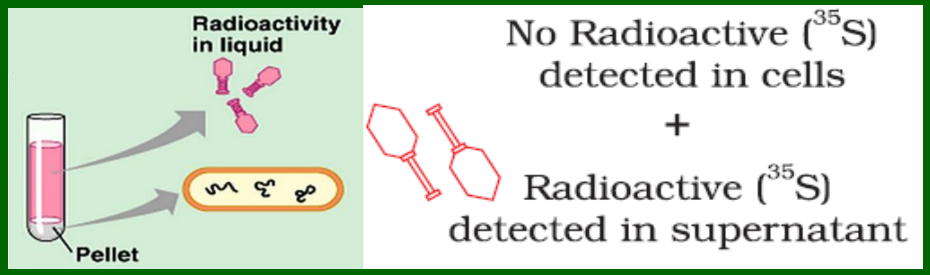
New bacteriophages contained radioactive protein but not radioactive DNA

Radioactive phages were allowed to attach to *E. Coli* and infect

Coats were removed from the bacteria by agitating them in a blender just after infection

Virus particles were separated from the bacteria by spinning them in a centrifuge

Observed and interpreted



Hershey and Chase's experiment

Grew some bacteriophages on a medium that contained radioactive phosphorus (^{32}P).

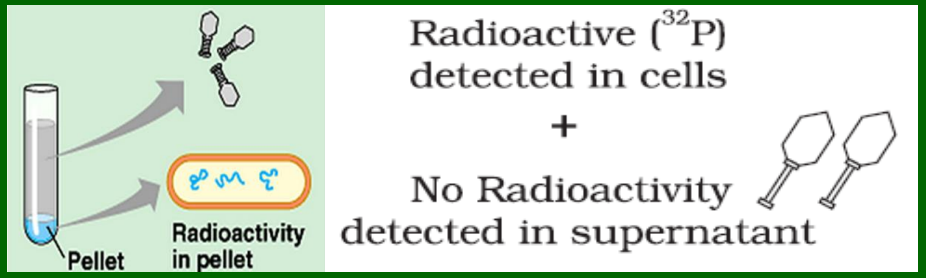
New bacteriophages contained radioactive DNA but not radioactive protein

Radioactive phages were allowed to attach to *E. Coli*

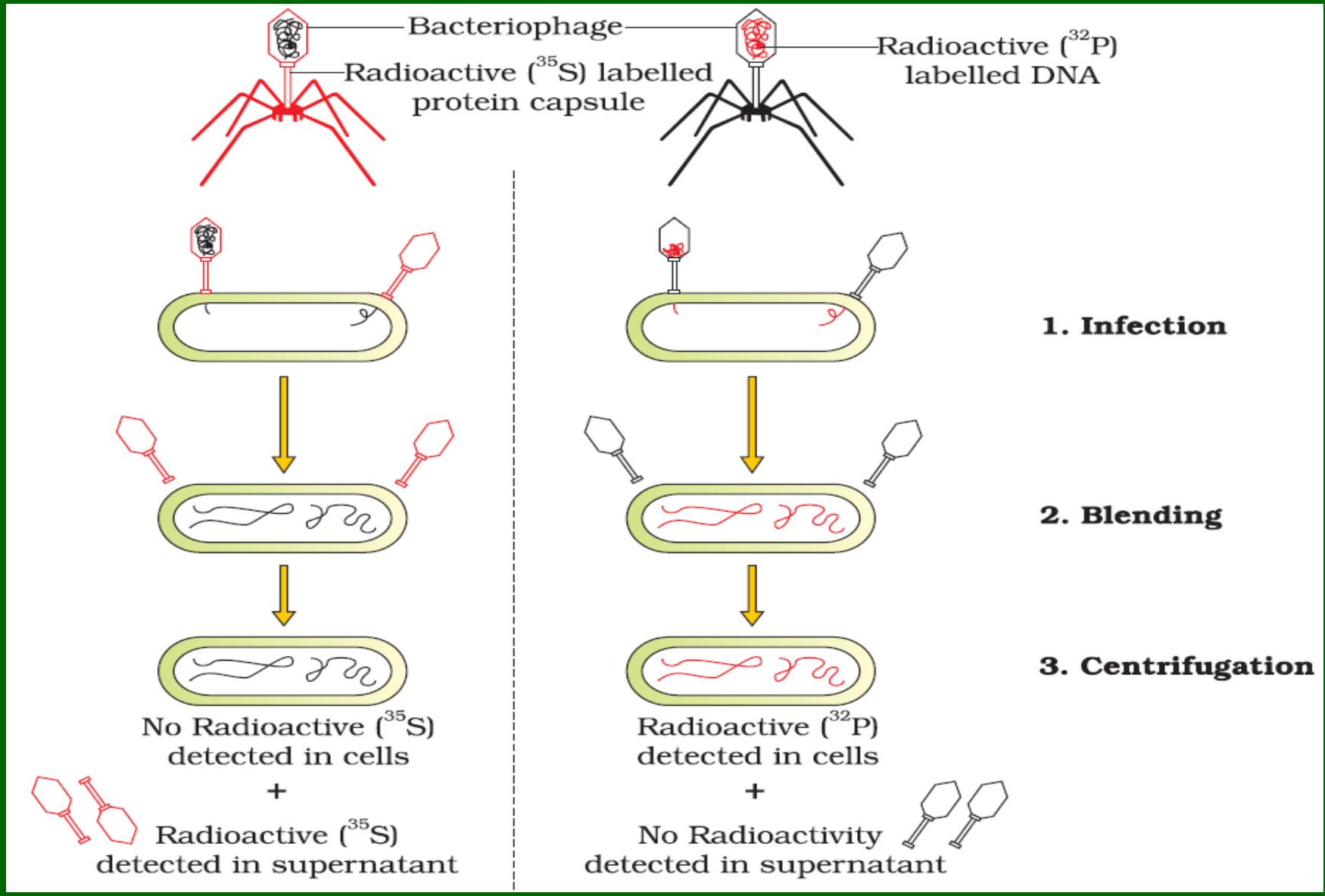
Coats were removed from the bacteria by agitating them in a blender just after infection

Virus particles were separated from the bacteria by spinning them in a centrifuge

Observed and interpreted



The Hershey-Chase experiment



Summary in search for genetic material

Years	Scientists	Test Organism	Contribution
1928	Frederick Griffith conducted experiments with	*Streptococcus pneumoniae* Mouse/Mice	Transforming principle
1933-44	Oswald Avery, Colin MacLeod and Maclyn McCarty	*Streptococcus pneumoniae*	Concluded that DNA is hereditary material or biochemical nature of 'transforming principle is DNA
1952	Alfred Hershey and Martha Chase	Bacteriophage *E. coli*	Proved that DNA is the genetic material.

Properties of Genetic Material

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

✓ It should be able to **generate its replica**(Replication).

✓ It should be **chemically and structurally stable**.

✓ It should **provide the scope for slow changes (mutation)** that are required for evolution.

✓ It should be able to **express itself in the form of 'Mendelian Characters'.**

Properties of Genetic Material (DNA versus RNA)

Reading Exercise

- Read page number 103 (4th last paragraph) and try to answer the following question
- Why DNA is a better genetic material compared to RNA?

Why DNA is a better genetic material compared to RNA?

- DNA is **chemically less reactive and structurally more stable when compared to RNA.**
- **Two strands of DNA being complementary if separated by heating come together**, when appropriate conditions are provided.
- **Presence of thymine at the place of uracil** also confers additional stability to DNA.
- In RNA, ***2‘ -OH group present* at every nucleotide is a reactive group** and makes RNA labile* and easily degradable.
- RNA is also now known to be catalytic, hence reactive.
- Therefore, among the two nucleic acids, the DNA is a better genetic material.

***Readily undergoing change**

What do you think is the first genetic material?

- RNA was the **first genetic material**
- **DNA has evolved from RNA** with chemical modifications that make it more stable.
- DNA being double stranded and having complementary strand further resists changes.

REPLICATION

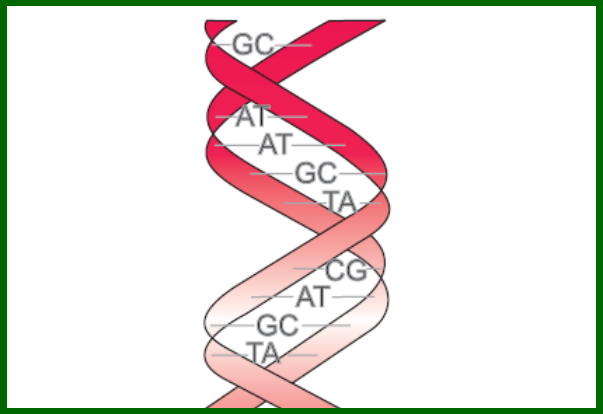
- Replication is the process of **copying DNA into DNA.**
- **Watson and Crick** had proposed a **semiconservative** replication of DNA.

"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material"

Watson and Crick, 1953

Watson-Crick model for semiconservative DNA replication

Parental DNA

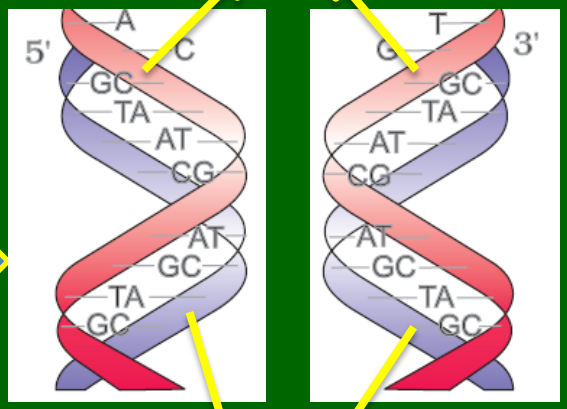


Two separate strands of DNA

Template DNA

New complementary strands

One parental DNA



Newly synthesised strands

Daughter DNA

The Experimental Proof

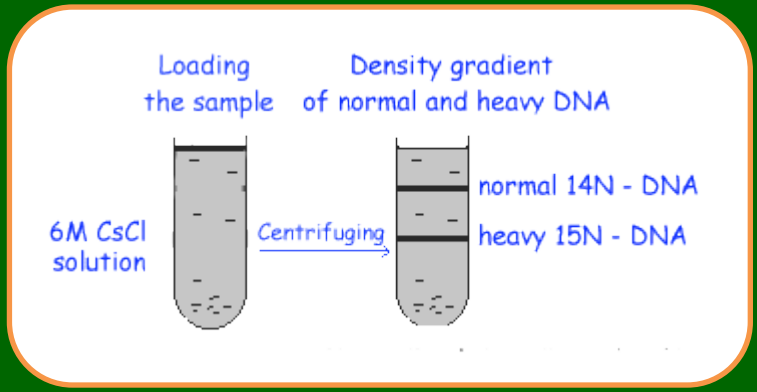
- The semiconservative replication of DNA was shown first in ***Escherichia coli*** *and subsequently in plants* and human cells.
- **Matthew Meselson and Franklin Stahl (1958)** proved that **DNA replicates semiconservatively**.

Lets Learn Together Better

- **Cesium chloride (CsCl) density gradient centrifuge**
- CsCl forms density gradient

Isotope Of Nitrogen

- *^{15}N is the heavy* isotope
- ^{14}N *is the light* isotope



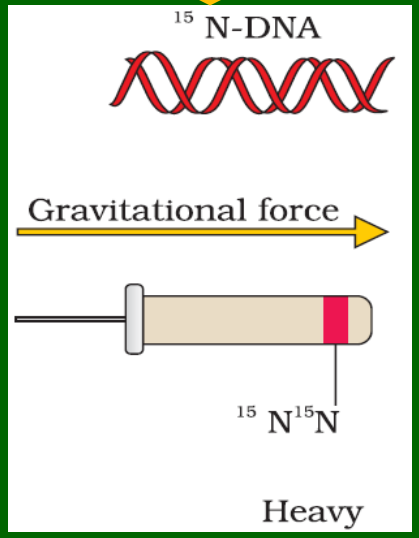
Meselson and Stahl's Experiment

Grew *E. coli* in a medium containing $^{15}NH_4Cl$ as the only nitrogen source for many generations

^{15}N was incorporated into newly synthesised DNA

Extracted the DNA

Centrifuged the extracted DNA using Cesium chloride (CsCl) density gradient centrifuge(To distinguish heavy DNA molecule from the normal DNA)

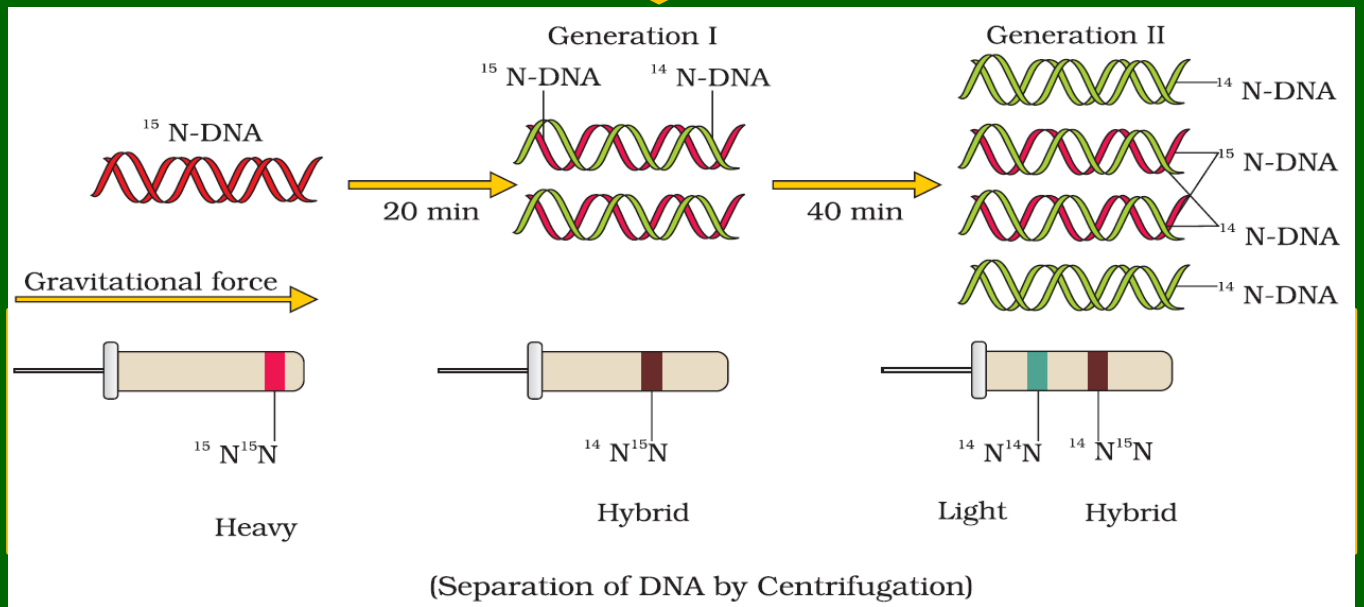


Continuation of Messelson and Stahl's Experiment

Transferred the *E. coli* cells into a medium with normal $^{14}NH_4Cl$

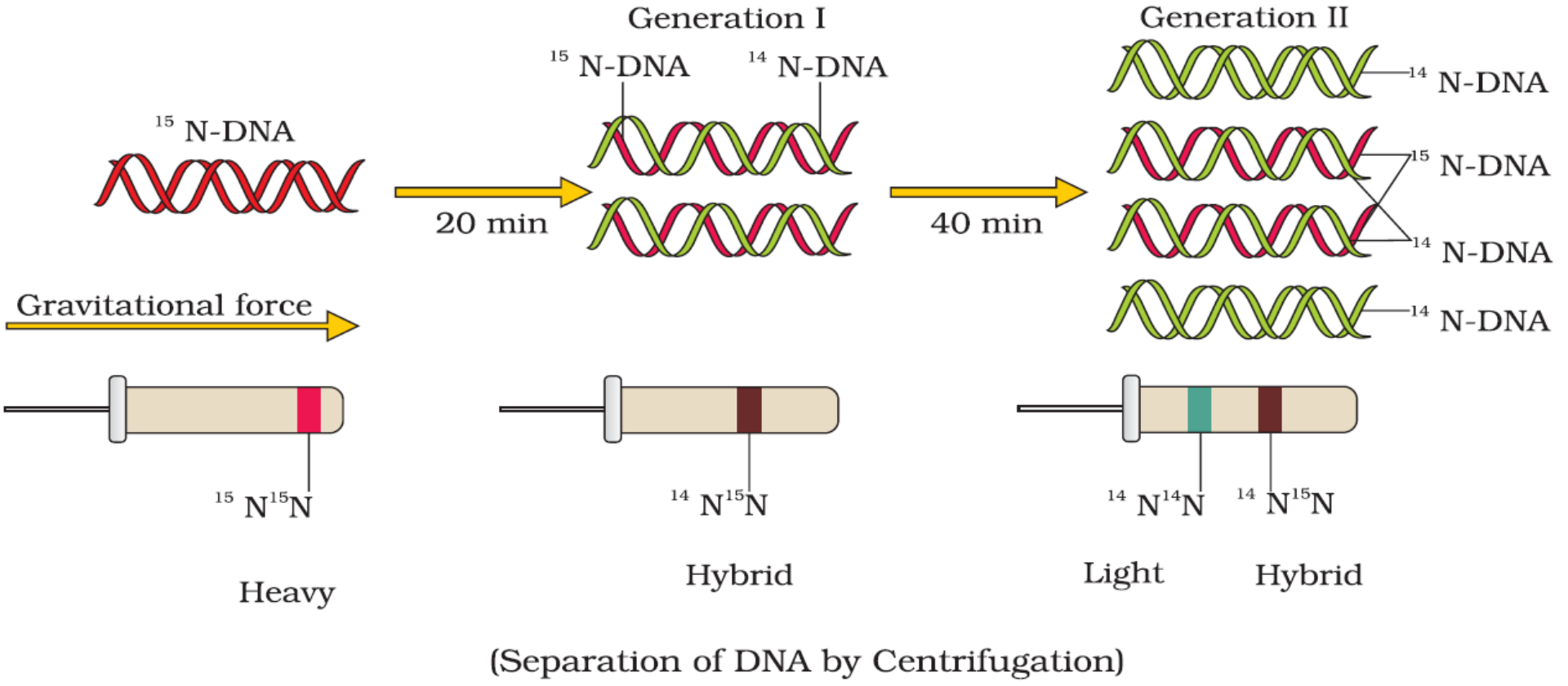
Extracted the DNA at various time intervals (20' & 40') as the cells multiplied

These samples were separated independently using CsCl gradients centrifuge to measure the densities of DNA



(Separation of DNA by Centrifugation)

Messelson and Stahl's Experiment



(Separation of DNA by Centrifugation)

Messelson and Stahl's Experiment

- Very similar experiments of Messelson and Stahl involving use of radioactive thymidine to detect distribution of newly synthesised DNA in the chromosomes was performed on ***Vicia faba (faba beans) by Taylor and et. al, 1958.***
- The experiments proved that the DNA in chromosomes also replicate semiconservatively.

Terminologies*

*Refer Figure 6.8 Replicating Fork (page no. 107) for better understanding

Term	Description
Replication fork	A small opening of the DNA helix
Helicases	Unwind double stranded DNA
Deoxyribonucleaoside triphosphate	Act as substrate and also provide energy for polymerisation of nucleotides
DNA polymerase	Catalyse the polymerisation of deoxynucleotides in 5′→3′ using DNA template
Primase	Enzyme that synthesizes primers (short RNA sequences)
Primer	Short RNA sequence that serve as a starting point for DNA synthesis
DNA ligase	Joins lagging strand (discontinuously synthesised fragment)
Lagging strand	Discontinuously synthesised fragments
Origin of replication	Region in *DNA where the replication* originates

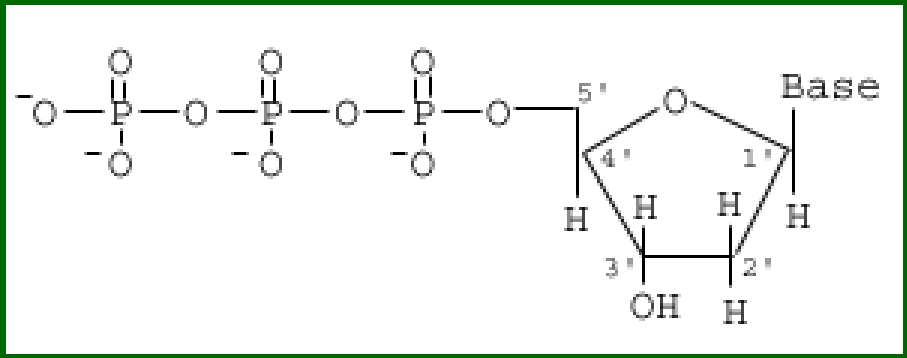


Figure: Deoxyribonucleaoside triphosphate

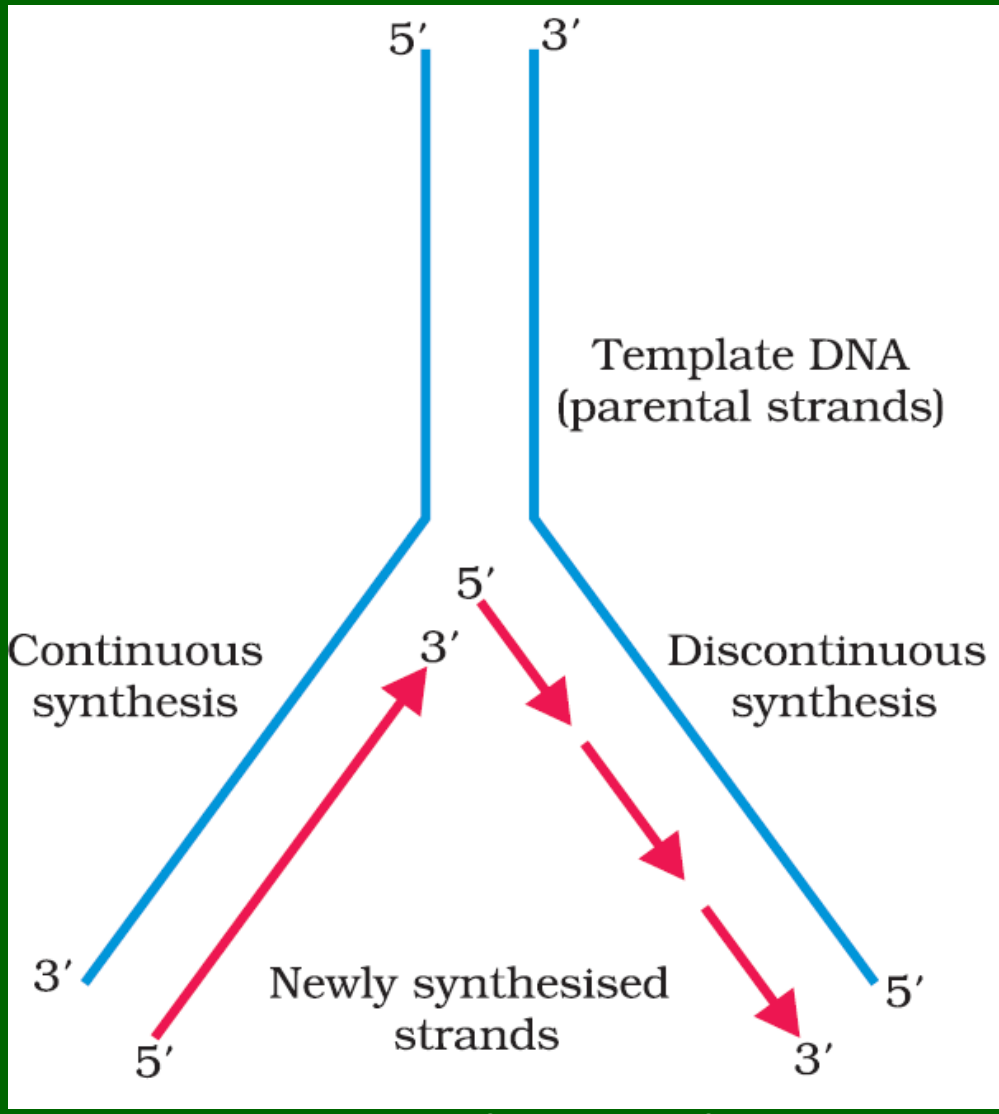


Figure: Replicating Fork

Read on page no. 106-7

- **6.4.2 The Machinery and the Enzymes**

The Machinery and the Enzymes

Pre-test

When do DNA replicates in eukaryotic cells?

Answer: During the S-phase of cell cycle

Replication of DNA

Formation of replication fork

The DNA-dependent DNA polymerases catalyse polymerisation of DNA in 5′→*3′* direction

Synthesis of leading strand *(continuous strand)* on the DNA template with polarity 3′→*5′*

Synthesis of lagging strand (discontinuous strand) on the DNA template with polarity 5′→*3′*

DNA ligase joins lagging strand

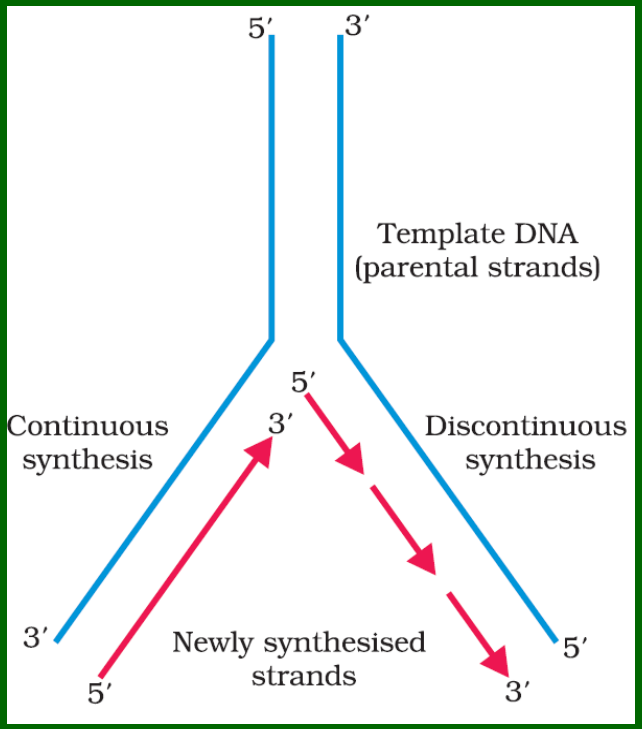


Figure: Replicating Fork

TRANSCRIPTION

- Transcription is the process of **copying genetic information from one strand of the DNA into RNA.**
- A **segment of one strand of DNA is transcribed into RNA**
- **Why both the strands are not copied during transcription?**
- **Read paragraph 2 of section 6.5 TRANSCRIPTION in page number 107**

Why both the strands of DNA are not copied during transcription?

- First, if both strands act a template, they would code for RNA molecule with different sequences, and in turn, if they code for proteins, the sequence of amino acids in the proteins would be different. Hence, **one segment of the DNA would be coding for two different proteins,** and this would **complicate the genetic information transfer machinery**.
- Second, the **two RNA molecules if produced simultaneously would be complementary to each other**, hence **would form a double stranded RNA.** This would prevent RNA from being translated into protein and the exercise of transcription would become a futile one.

Transcription Unit

- A **transcription unit in DNA** comprises of three regions:
  1) **A Promoter**
  2) **The Structural gene**
  3) **A Terminator**
- Two strands of the DNA in the structural gene of a transcription unit are;
  1) **Template strand (with polarity 3'→5‘)**
  2) **Coding strand (with polarity *5′→3‘)***

Transcription Unit

- Based on coding strand, location of promoter, structural gene and terminator are described;

Three regions	Location
Promoter	towards 5'-end (upstream) of the coding strand of structural gene
Terminator	towards 3'-end (downstream) of the coding strand
Structural gene	Promoter and terminator flank the structural gene

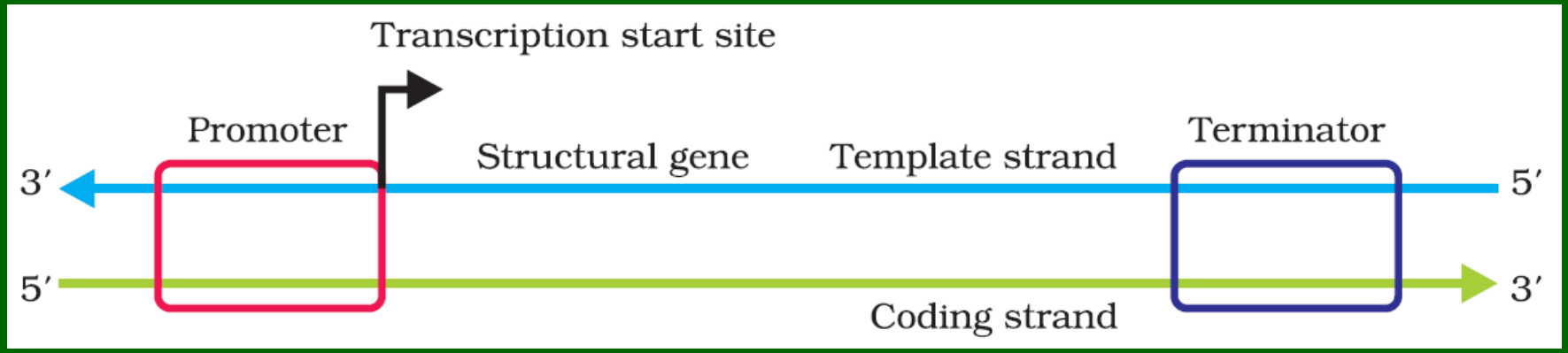
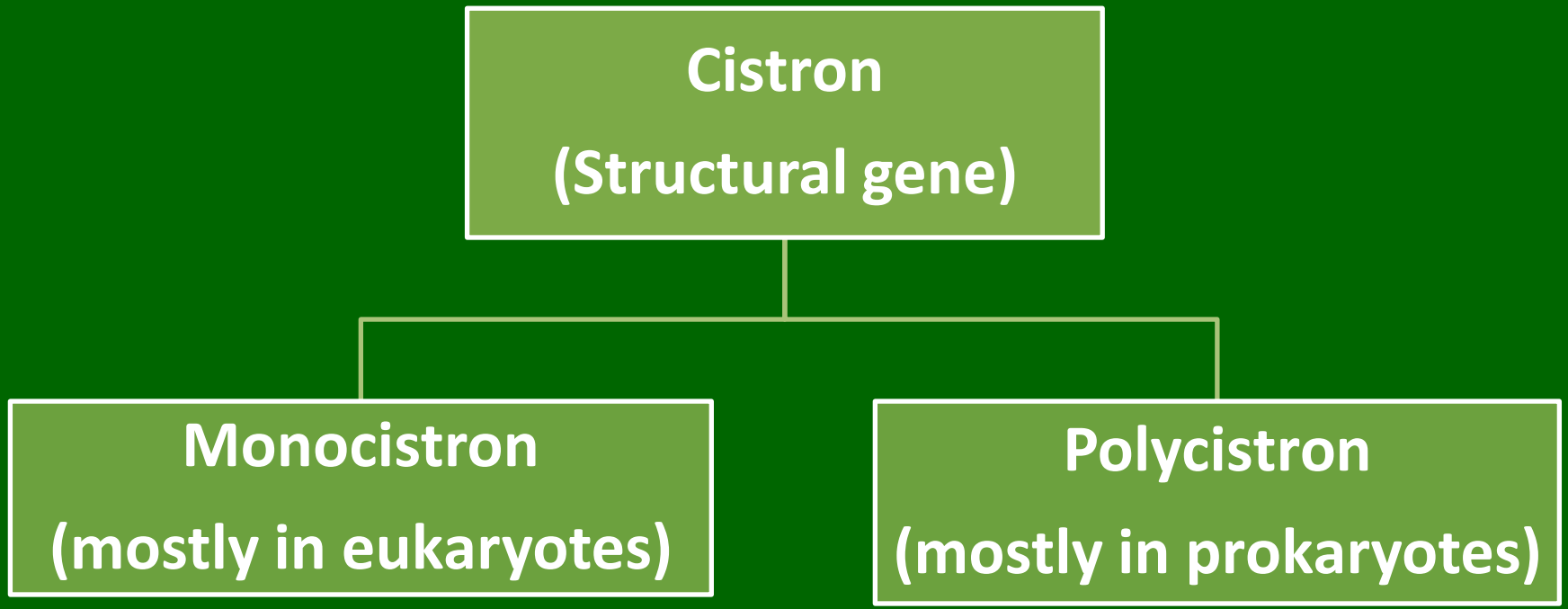


Figure: Schematic structure of a transcription unit

Transcription Unit and the Gene

- **Cistron is** a segment of DNA or RNA coding for a specific polypeptide.



Eukaryotic cistron

- Eukaryotic monocistronic structural genes have interrupted coding sequences i.e Exons are interrupted by **introns**.
- **Exons/coding sequences are** expressed sequences
- **Exons** appear in **mature or processed RNA**
- **Introns**: intervening sequences
- **Introns** do not appear in mature or processed RNA.
- **Regulatory genes** : sequence of DNA that do not code for any protein.

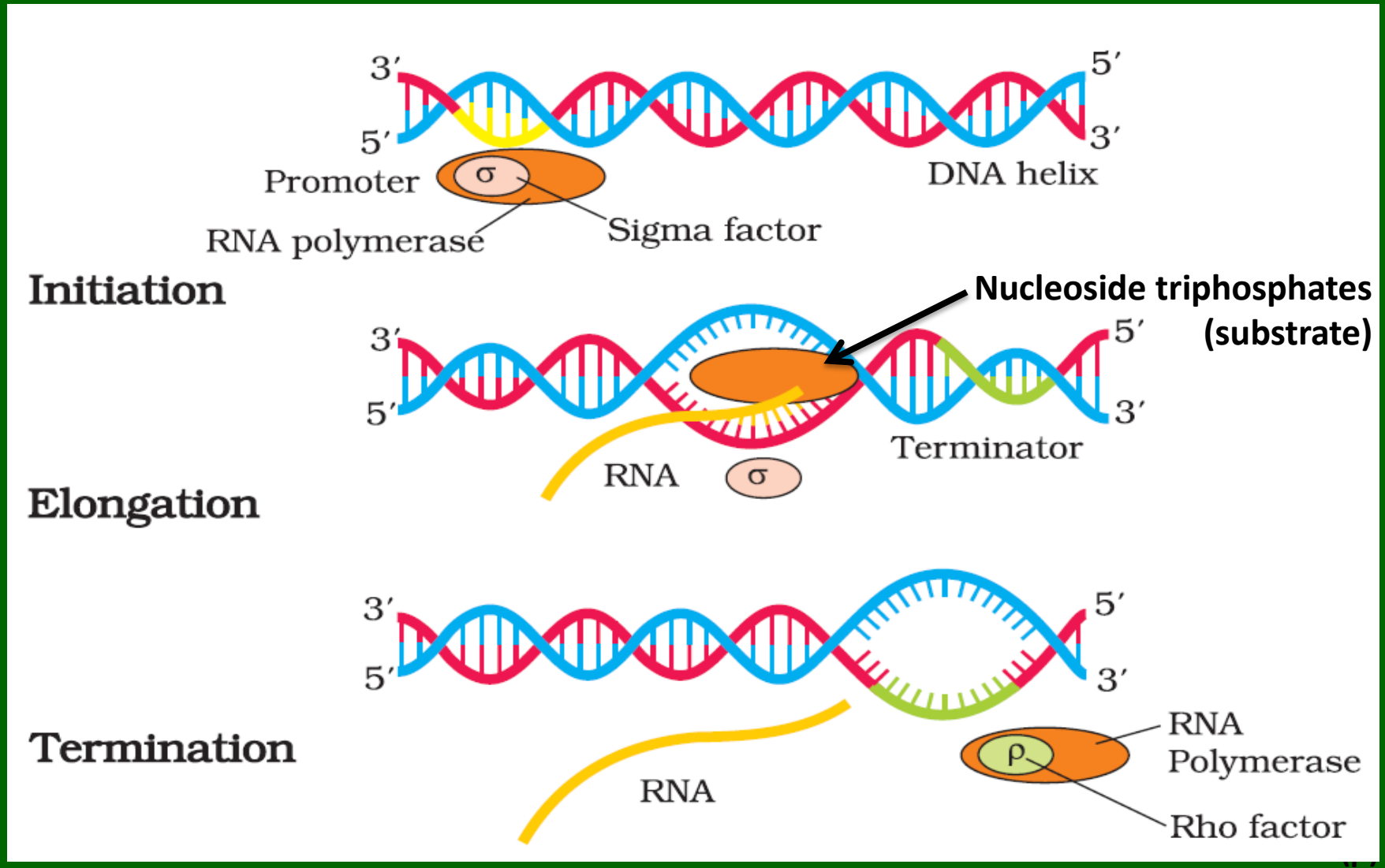
Types of RNA

- Three major types of RNAs In bacteria are :
  1) **mRNA (messenger RNA)**
  2) **tRNA (transfer RNA)**
  3) **rRNA (ribosomal RNA)**

  All these RNAs are needed to synthesise protein in a cell.

Types of RNA	Function
mRNA	provides the template for translation
tRNA	brings amino acids and reads the genetic code during translation
rRNAs	play structural and catalytic role during translation

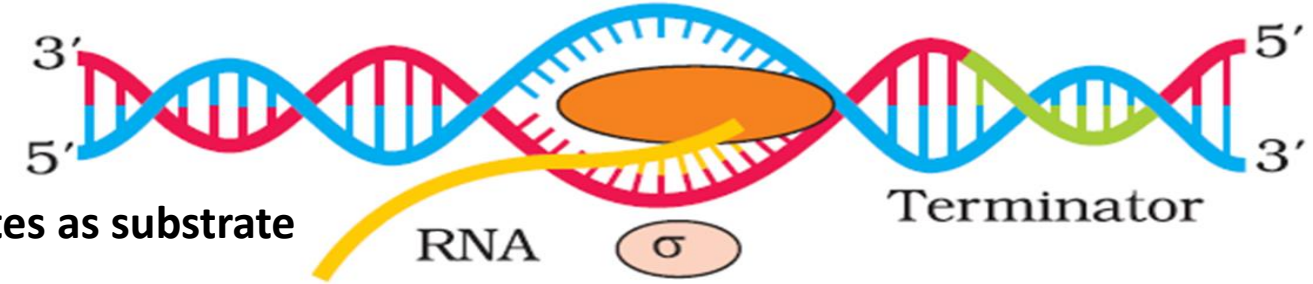
Process of Transcription in Bacteria



Process of Transcription in Bacteria

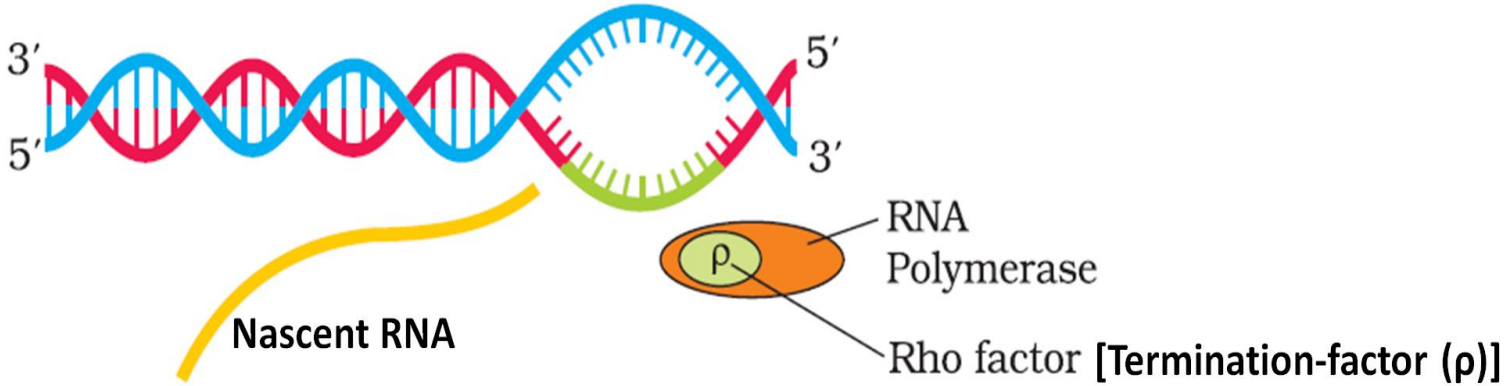
Initiation

Elongation

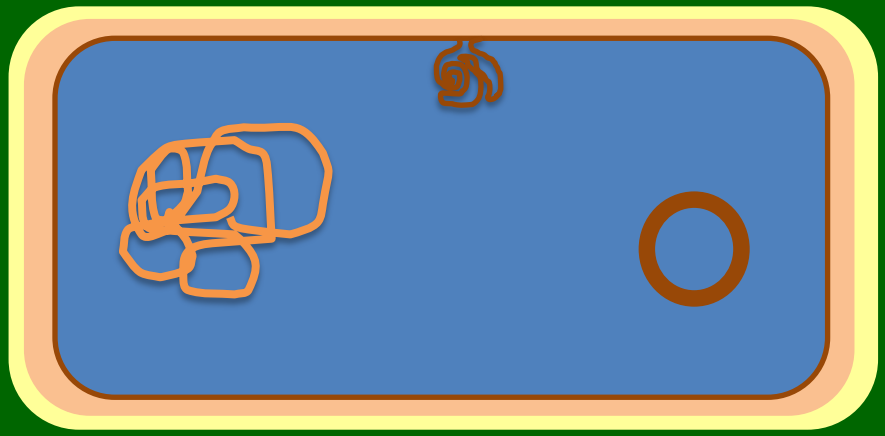


Uses nucleoside triphosphates as substrate

Termination



bacteria



In bacteria,

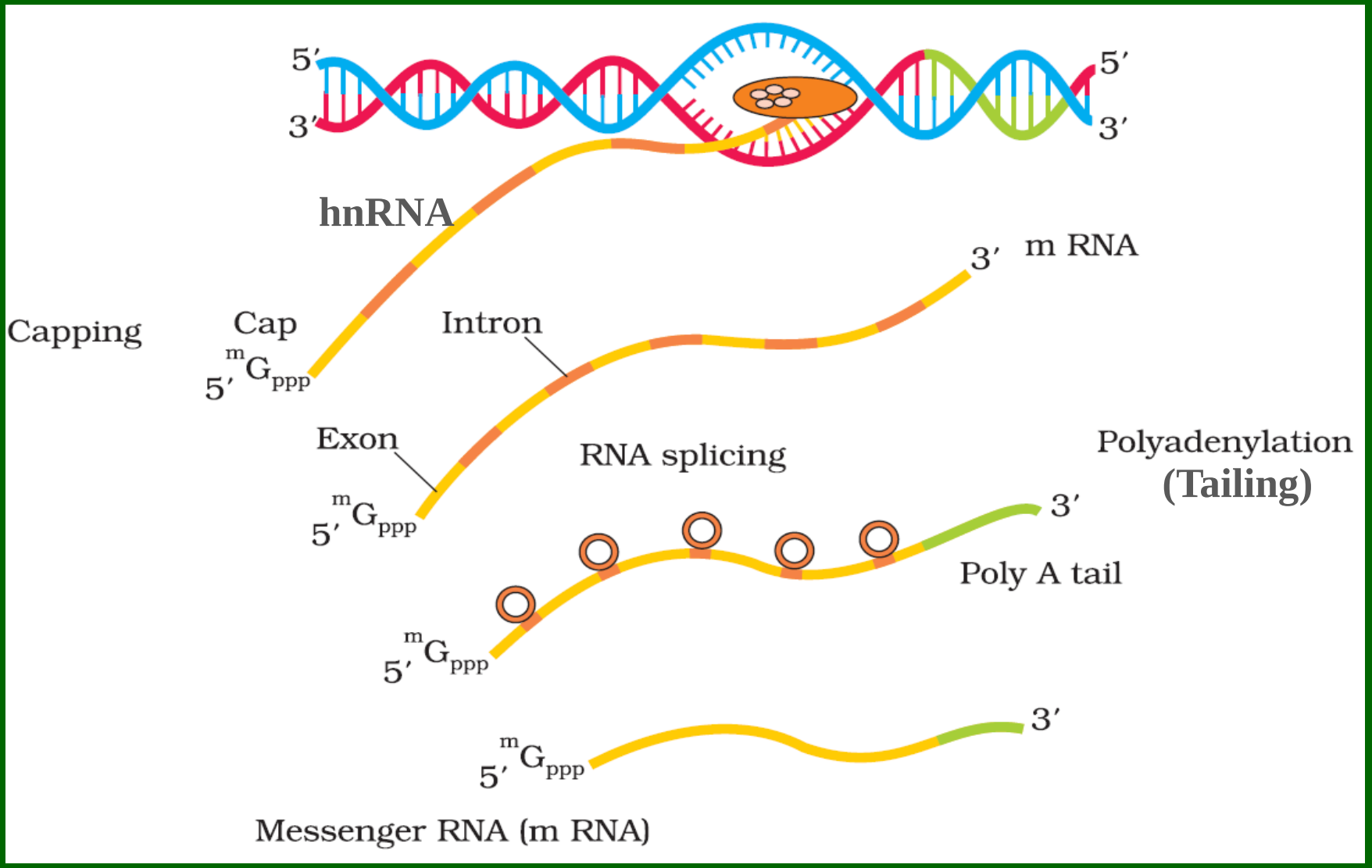
- **mRNA does not require any processing** to become active.
- **transcription and translation take place in the same compartment** i.e bacterial cell.
- The **transcription and translation can be coupled in bacteria** i.e many times the translation can begin much before the mRNA is full transcribed. .

Transcription in eukaryotes

- The three RNA polymerases found in the nucleus (*in addition to the RNA polymerase found in the organelles*);

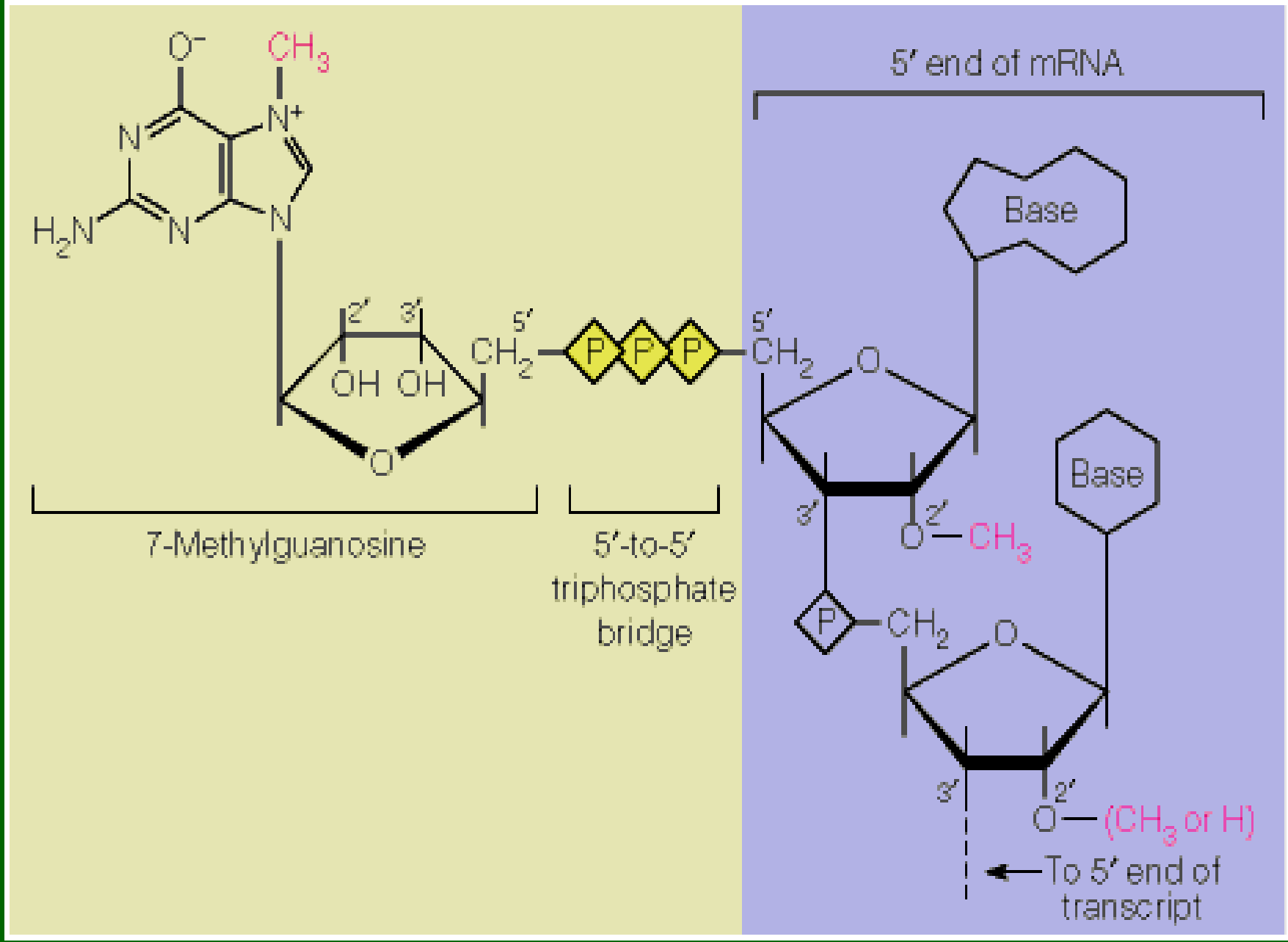
RNA polymerases	Division of labour
RNA polymerase I	**Transcribes rRNAs (28S, 18S, and 5.8S)**
RNA polymerase II	**Transcribes the heterogeneous nuclear RNA (hnRNA) [Precursor of mRNA]**
RNA polymerase III	**Transcription of tRNA, 5srRNA, and snRNAs (small nuclear RNAs)**

Processing of Transcribed RNA in Eukaryotes



Transcription in eukaryotes

- hnRNA undergo processing called as
  1) **Splicing** (*process where the introns are removed and exons are joined in a defined order*)
  2) **Capping** (*Methyl guanosine triphosphate is added to the 5'-end of hnRNA*)
  3) **Tailing/polyadenylation** [***adenylate residues (200-300) are added*** at 3'*-end in a template independent manner*]
- The fully processed hnRNA, now called mRNA, that is transported out of the nucleus for translation



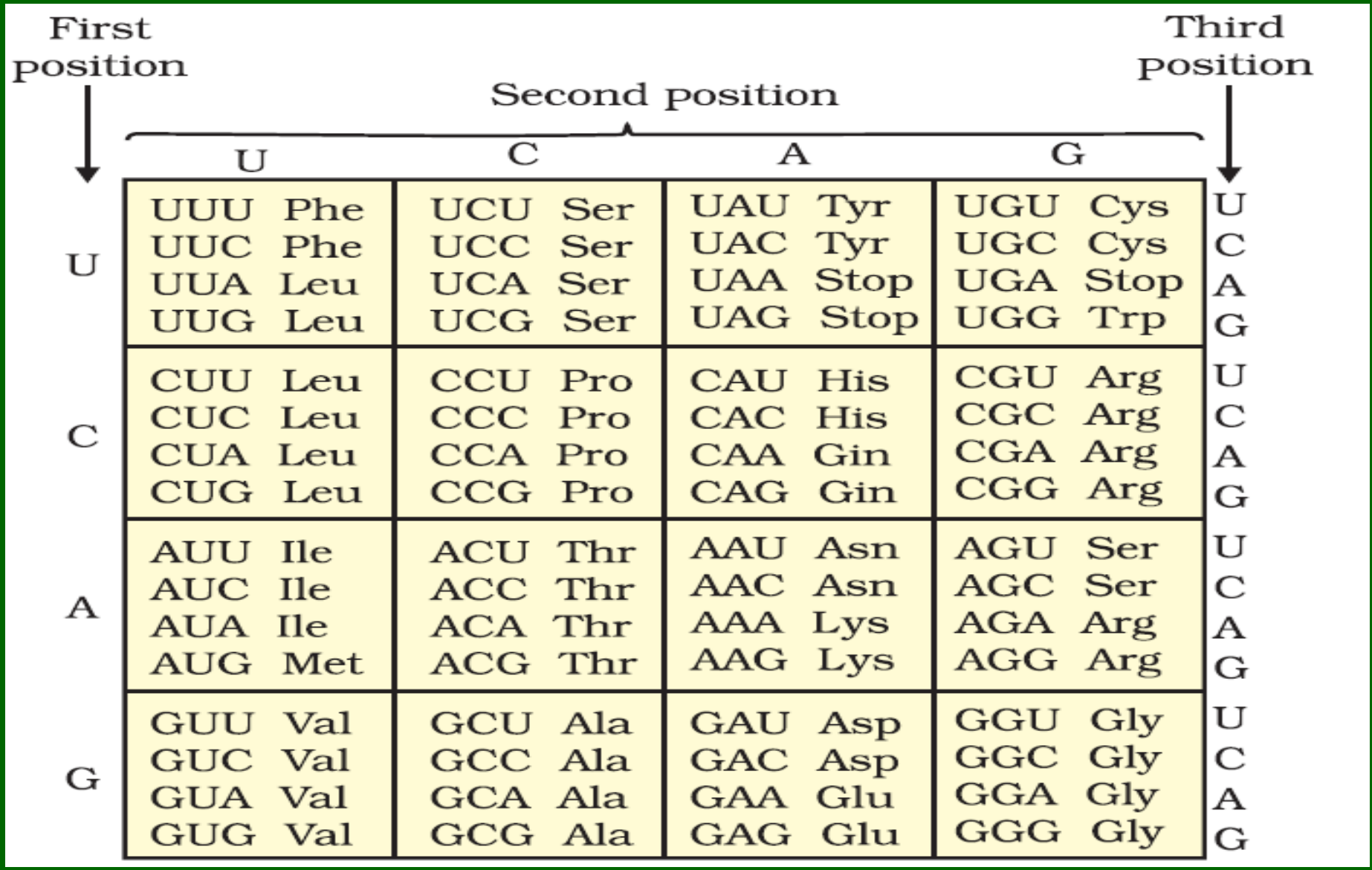
Methyl guanosine triphosphate

GENETIC CODE

Scientists	Contribution
George Gamow, a physicist	Suggested that in order to code for all the 20 amino acids, the code should be made up of three nucleotides i.e. **codons are triplet**
Har Gobind Khorana	**Developed a chemical method for synthesising RNA molecules** with defined combinations of bases (homopolymers and copolymers)
Marshall Nirenberg	**Cell-free system for protein synthesis**
Severo Ochoa	Severo Ochoa enzyme (**polynucleotide phosphorylase**) was also helpful in polymerising RNA with defined sequences in a template independent manner (**enzymatic synthesis of RNA**).
These contribution helped the code to be deciphered* and a checker-board for genetic code was prepared	

***Decoded**

The Codons for the Various Amino Acids



Salient features of genetic code

- **Read The salient features of genetic code on page number 112**

1) The codon is **triplet**. *61 codons code for amino acids and 3 codons (UAA, UAG & UGA) do not code for any amino acids, hence they function as stop codons.*
2) One codon codes for only one amino acid, hence, it is **unambiguous and specific.**
3) Some amino acids are coded by more than one codon, hence the code is **degenerate.**
4) The codon is read in mRNA in a **contiguous fashion**. There are no punctuations.
5) The code is **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.
6) **AUG has dual functions**. It codes for Methionine (met) , and it also act as **initiator codon.**

Mutation

- **Mutation** is a phenomenon which results in **alteration of DNA sequences** and **consequently results in changes in the phenotype of an organism**.
- **Gene/Point mutation is due to change in a single base pair of DNA. Example Sickle cell anemia** (GAG→GUG i.e Glutamic acid → Valine).

Some other types of mutations

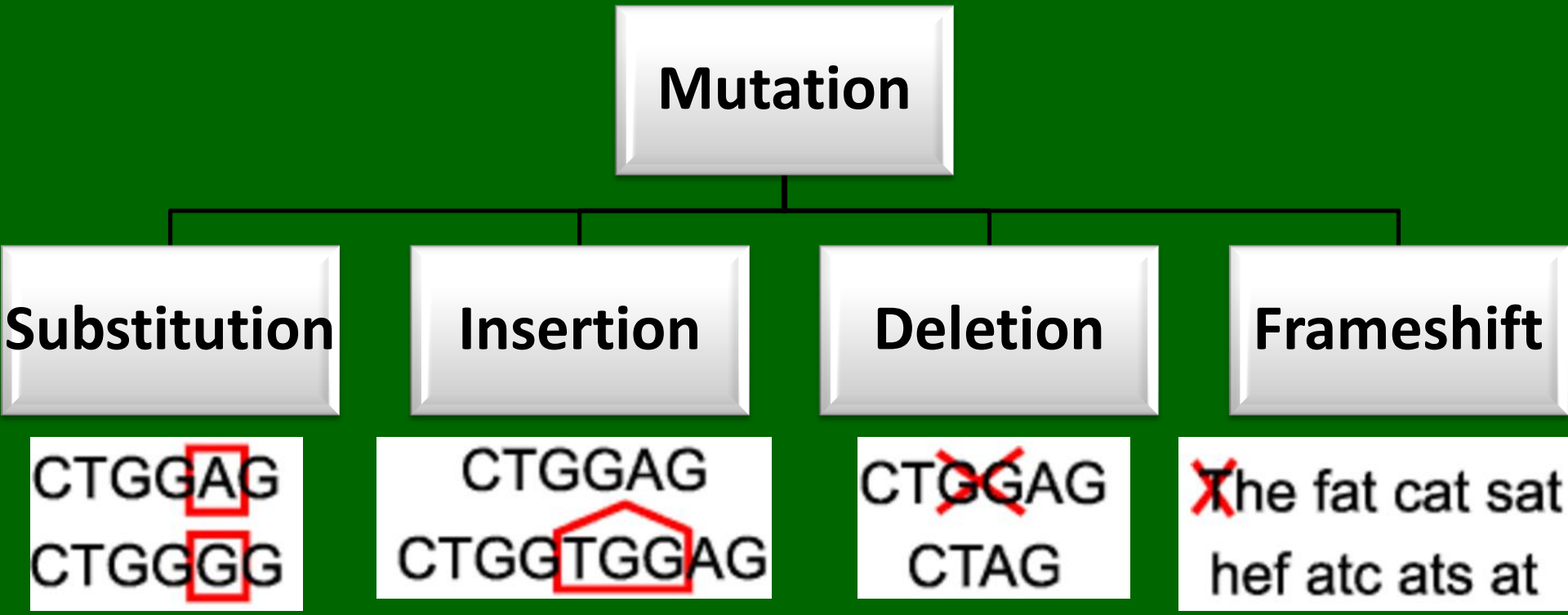
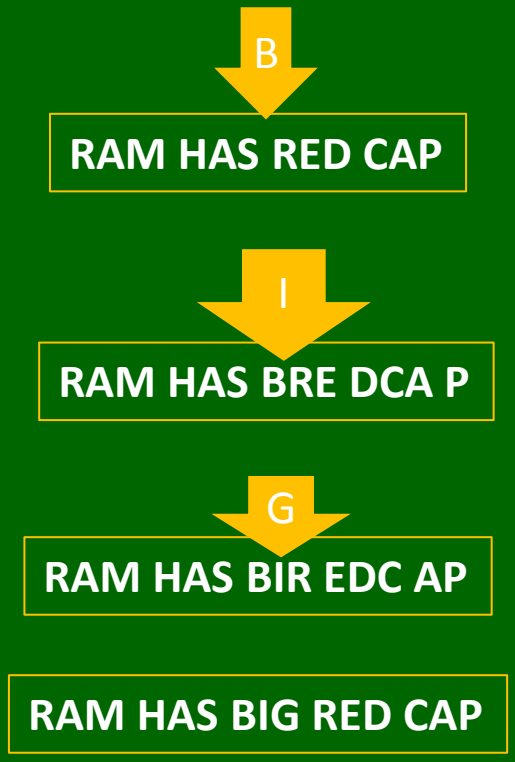
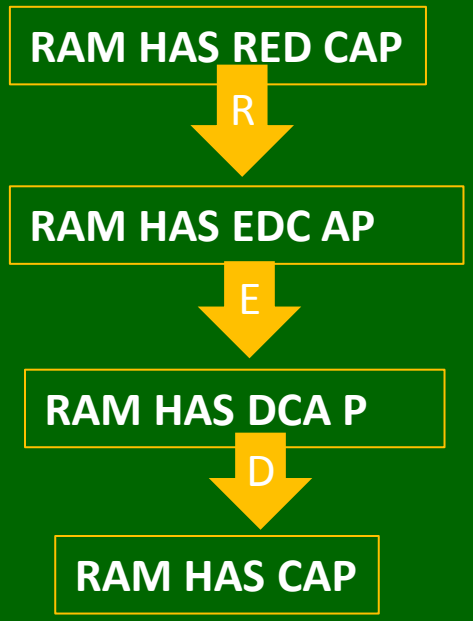


Illustration of Frame shift mutation

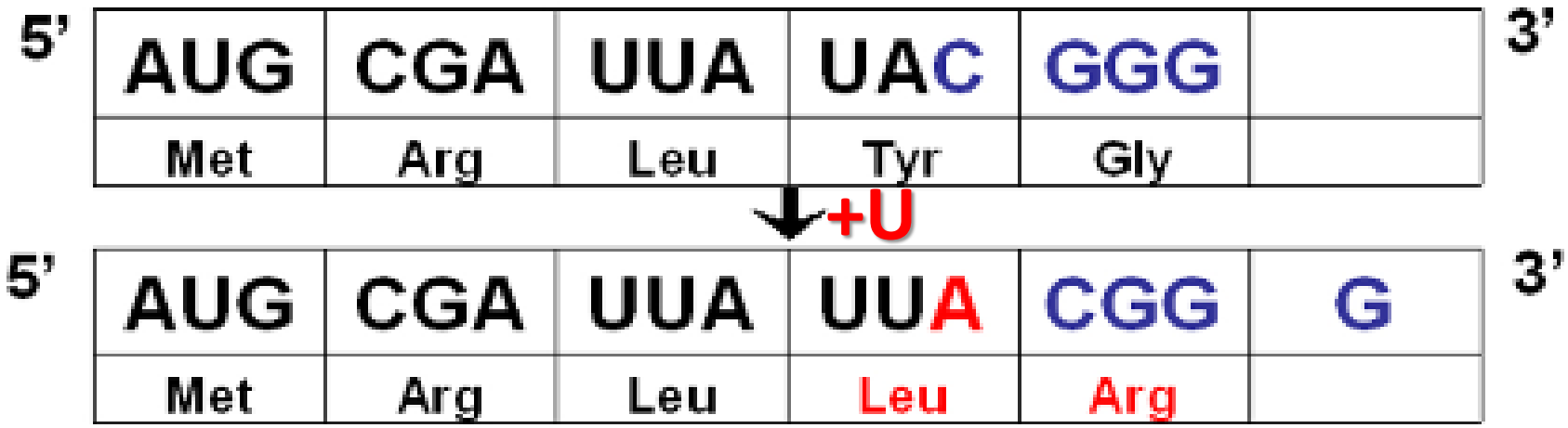
Insertion of letters B, I and G and rearranging the statement to make a triplet word



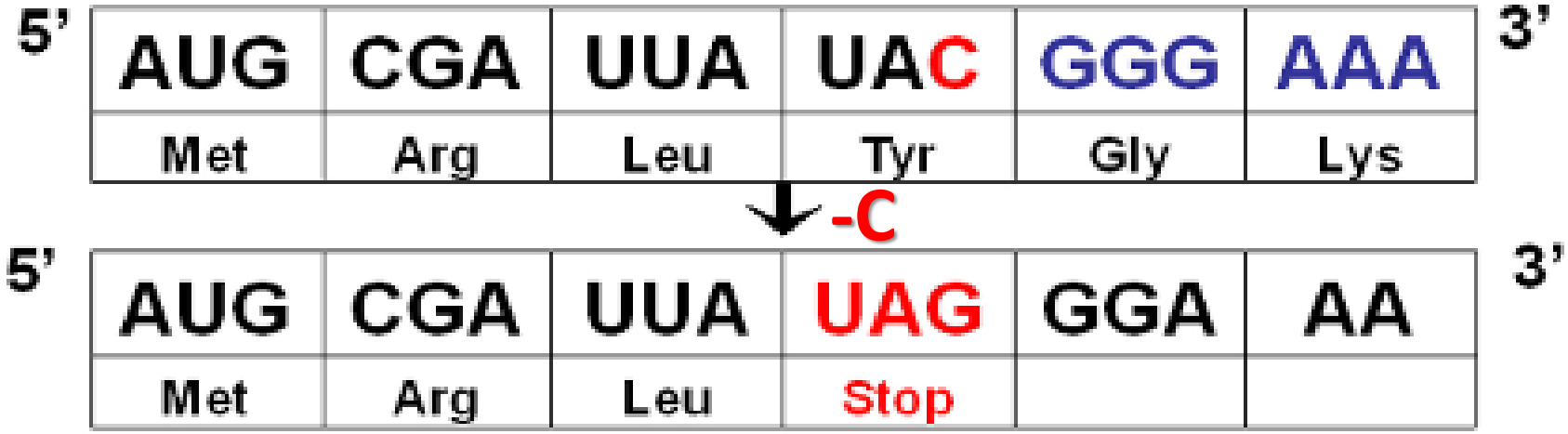
Deleting the letters R, E and D, one by one and rearranging the statement to make a triplet word



Insertion



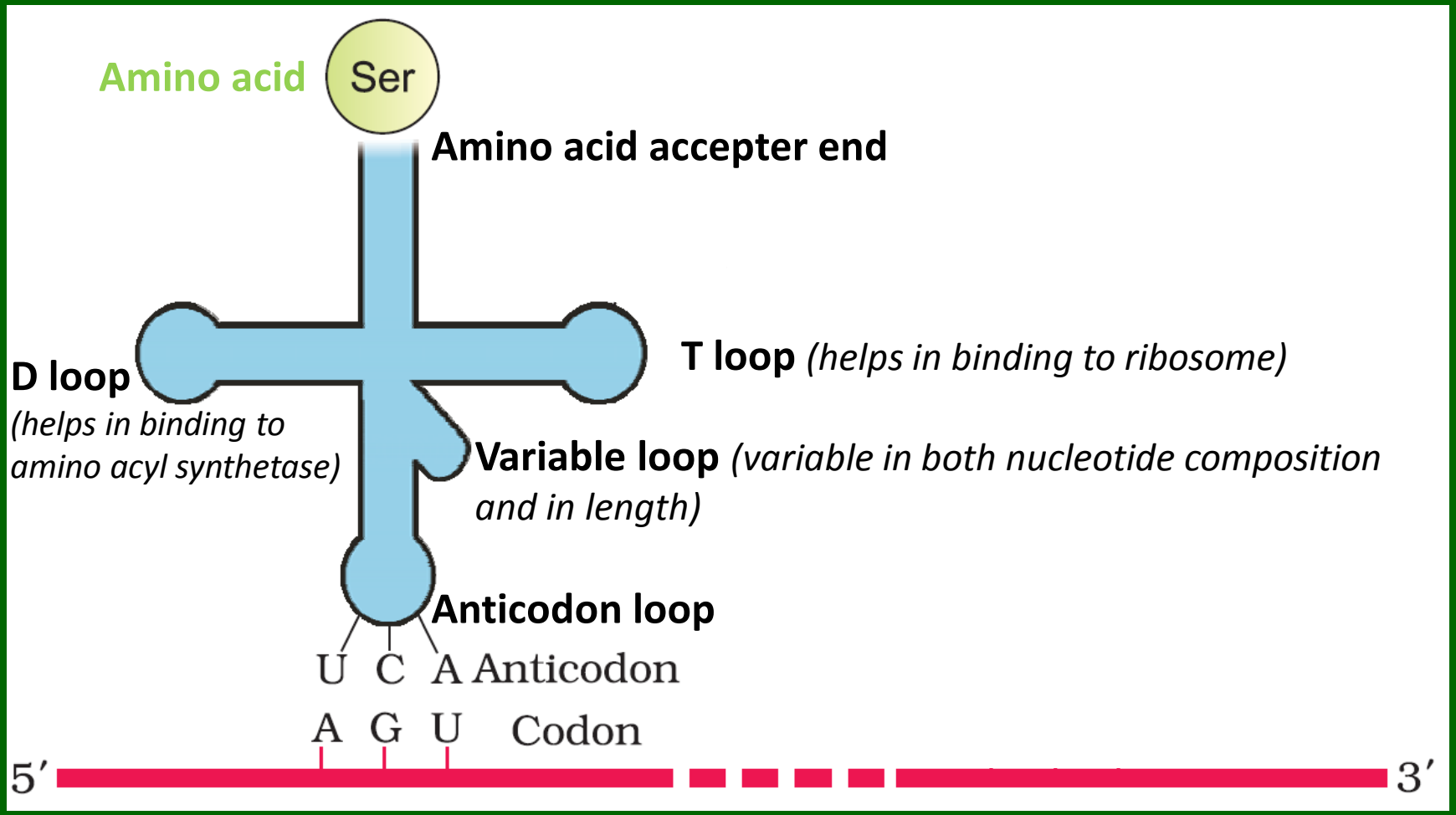
Deletion



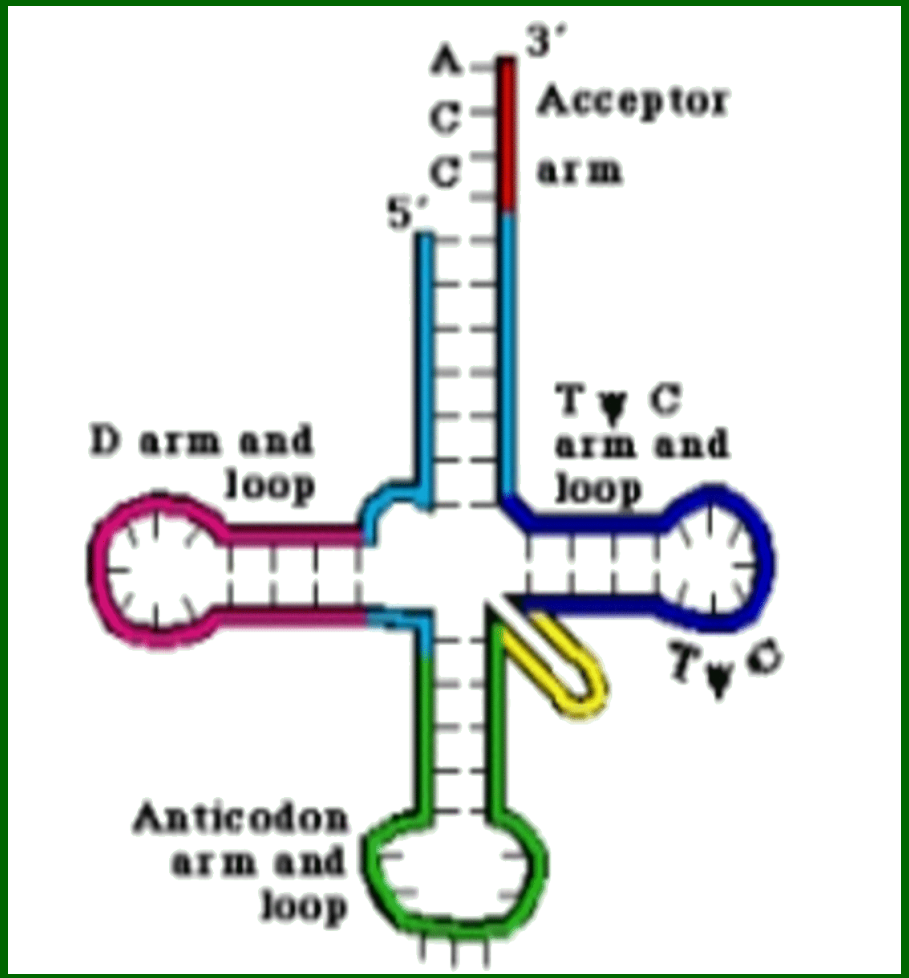
tRNA– the Adapter Molecule

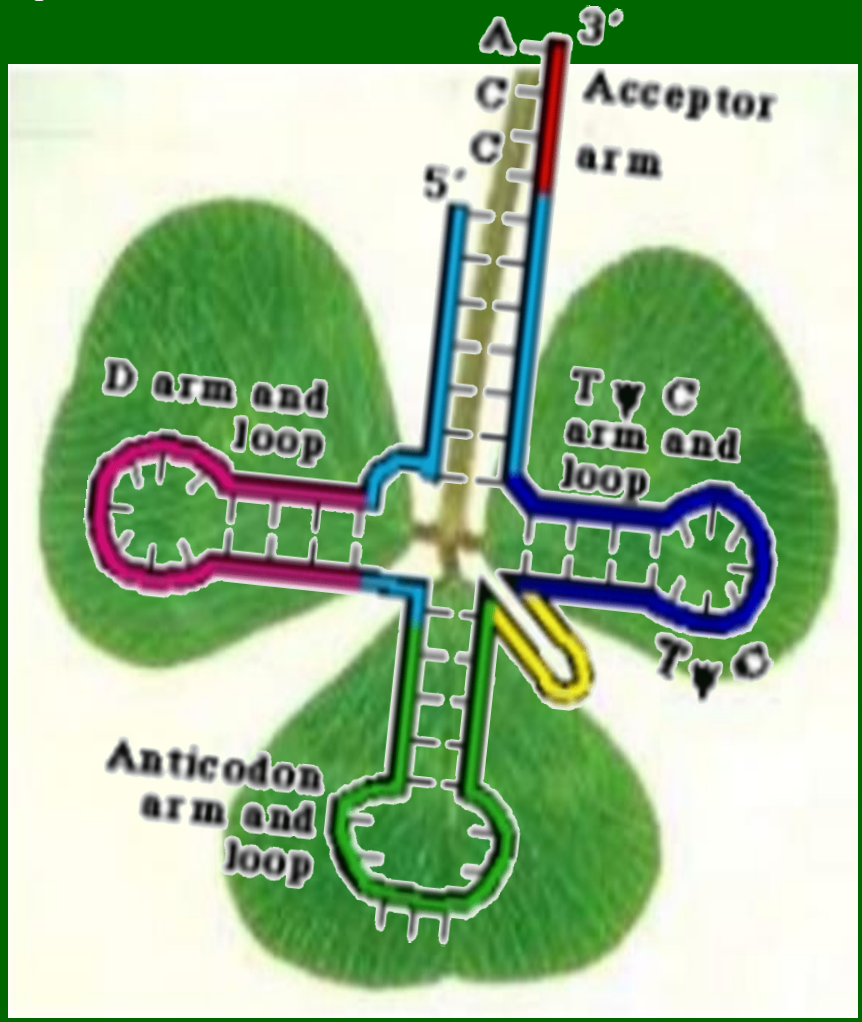
- Earlier known as **sRNA (soluble RNA)**
- Francis Crick postulated the **presence of an adapter molecule** that would on one hand **read the code** and on other hand would **bind to specific amino acids**.
- Role as an **adapter molecule** was assigned much later.

tRNA– the Adapter Molecule

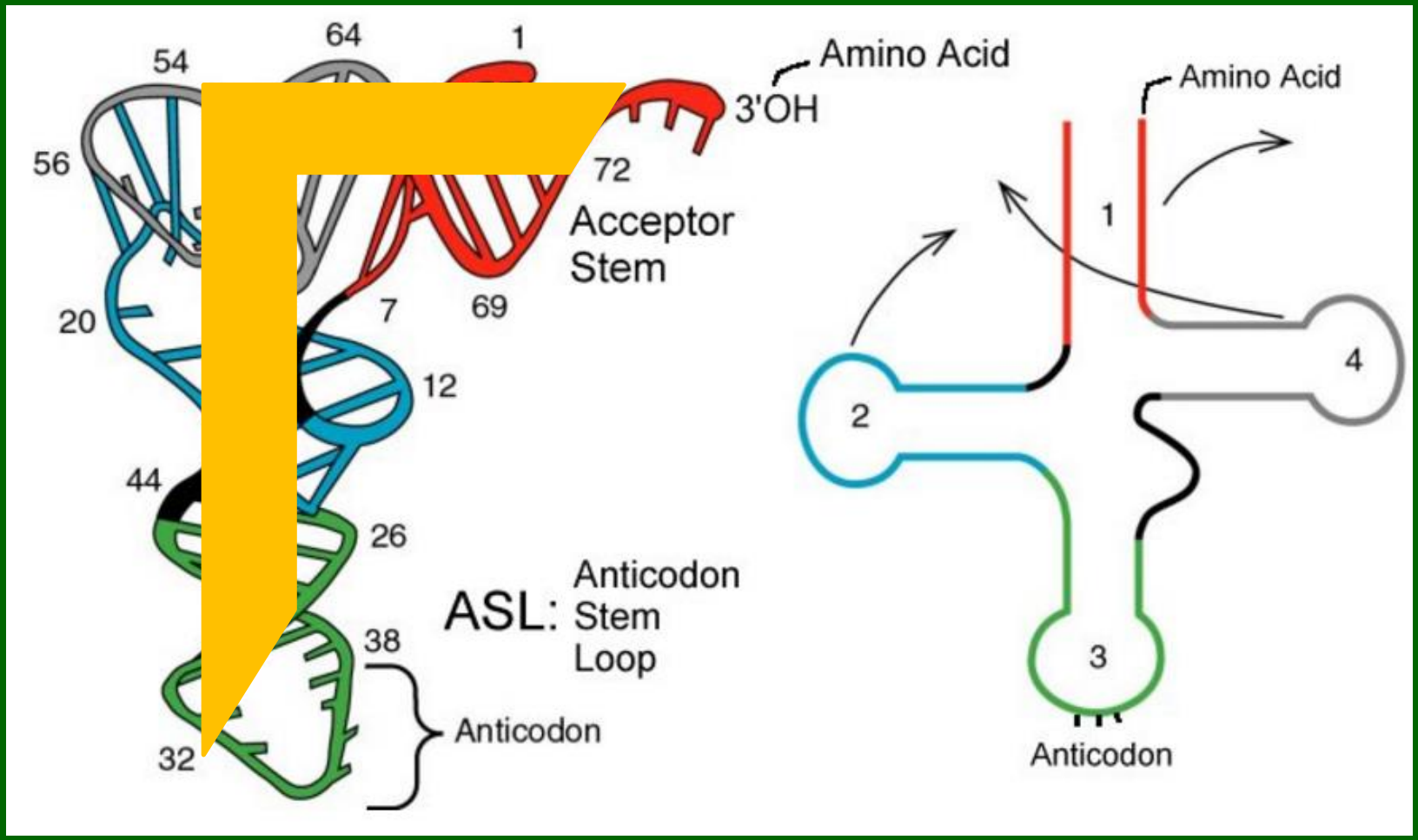


Clover leaf like Secondary structure of tRNA





tRNA is a compact, **inverted L shaped** molecule in actual structure



TRANSLATION

- Translation refers to the **process of polymerisation of amino acids to** form a **polypeptide.**
- The sequence of bases in the mRNA **defines the order and sequence of amino acids** (*joined by peptide bonds*).
- Formation of a peptide bond requires energy.

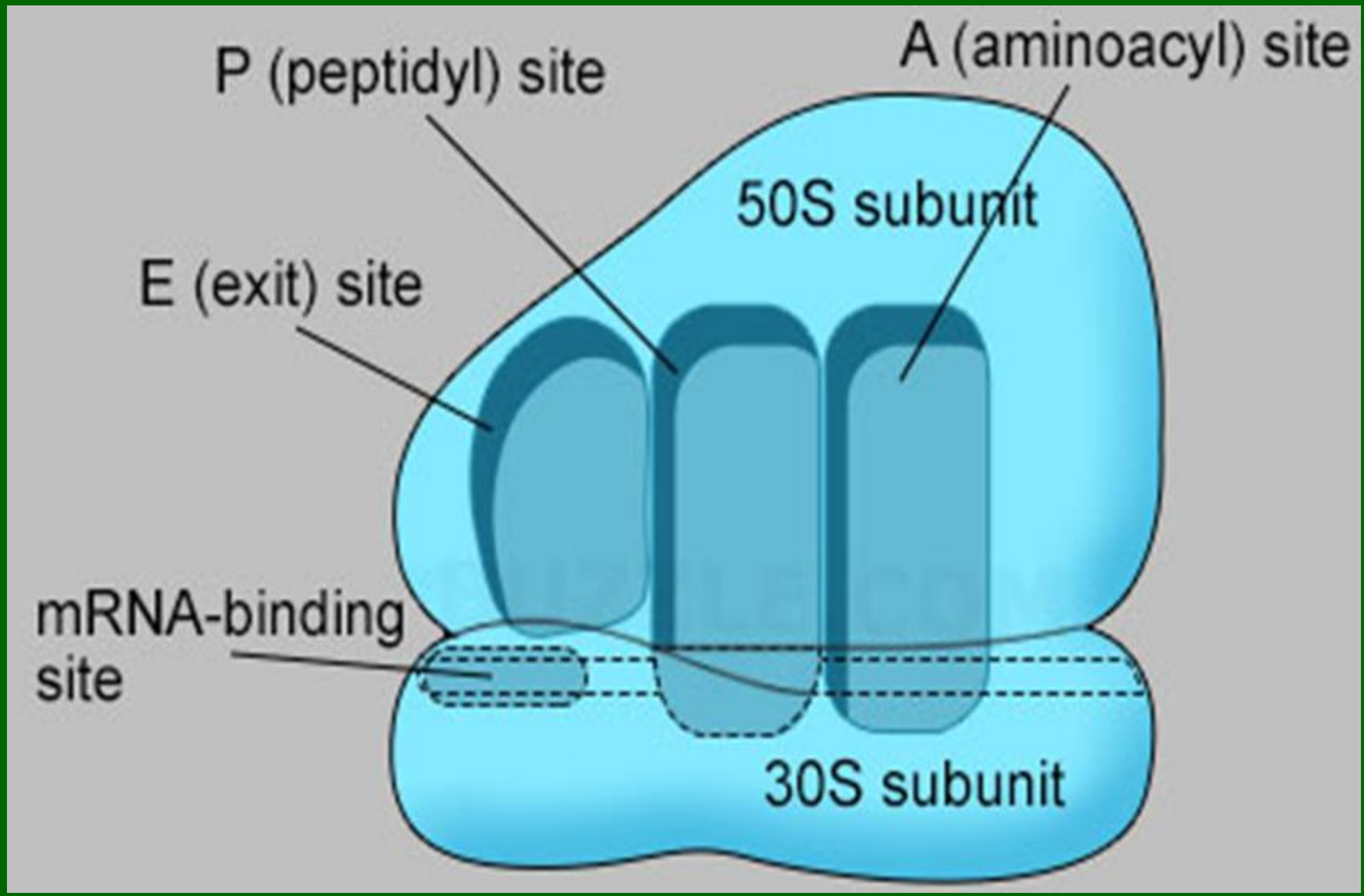
Aminoacylation of tRNA
(Charging of tRNA or formation of aminoacyl tRNA)

- In aminoacylation of tRNA, **amino acids** are **activated in the presence of ATP** and **linked to their cognate tRNA to form aminoacyl tRNA.**
- **If two such charged tRNAs** are brought close enough, the **formation of peptide bond** between them would be favoured energetically.
- The presence of a catalyst would enhance the rate of peptide bond formation.

Ribosome

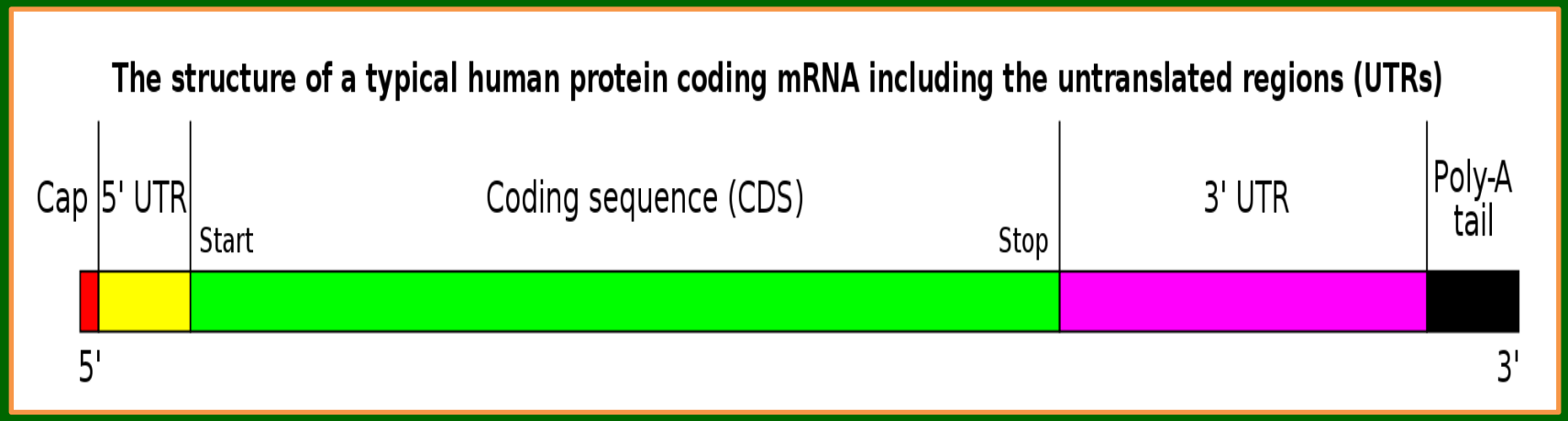
- Ribosome is the **cellular factory** responsible for synthesising proteins
- The ribosome consists of
  - ✓**structural RNAs** and
  - ✓**80 different proteins**.
- In inactive state, ribosome exists as two subunits;
  1. **a large subunit** and
  2. **a small subunit**.

Schematic diagram of Ribosome



Translational unit

- A translational unit is **present in mRNA** coding for a **polypeptide**
- It is the **sequence of RNA** that is flanked by the start codon (AUG) and the stop codon.
- Untranslated, additional sequences of mRNA that are **present** at both *5' -end (before start codon) and at 3' -end (after stop codon)* are referred as **untranslated regions (UTR).**
- **UTRs** are required for efficient translation process.



Translation

- Three steps of translation;
  1. Initiation
  2. Elongation
  3. Termination

Translation

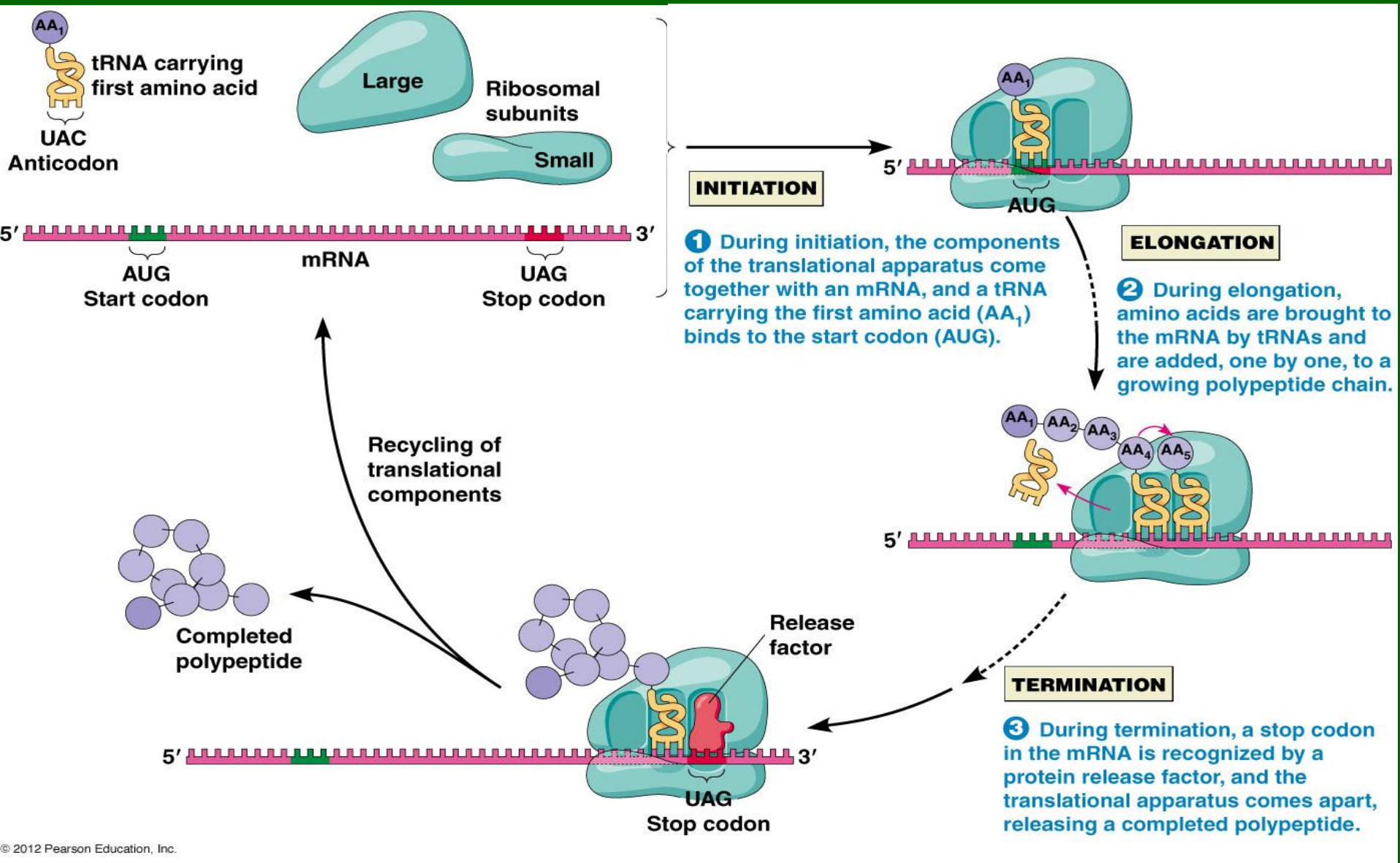
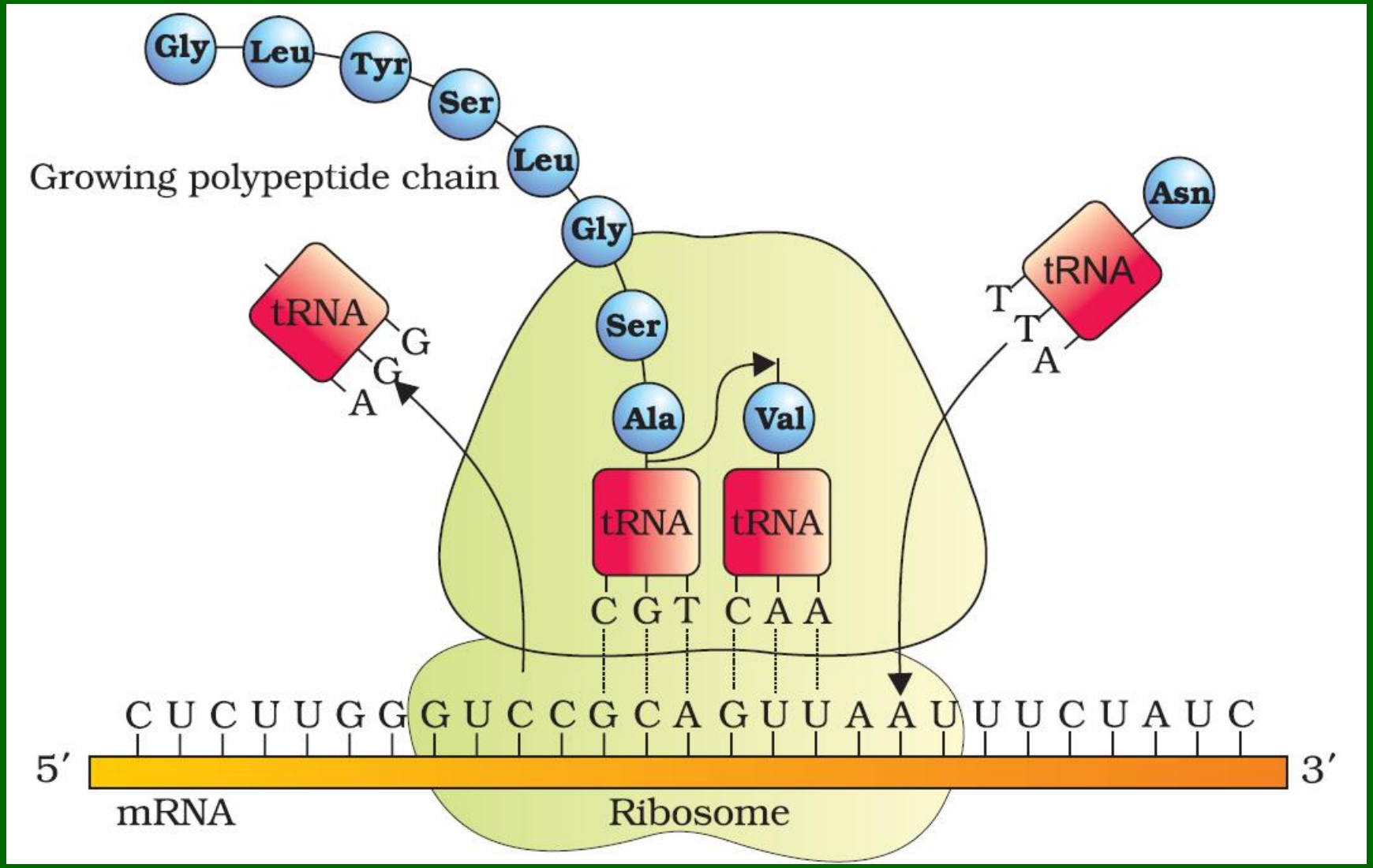
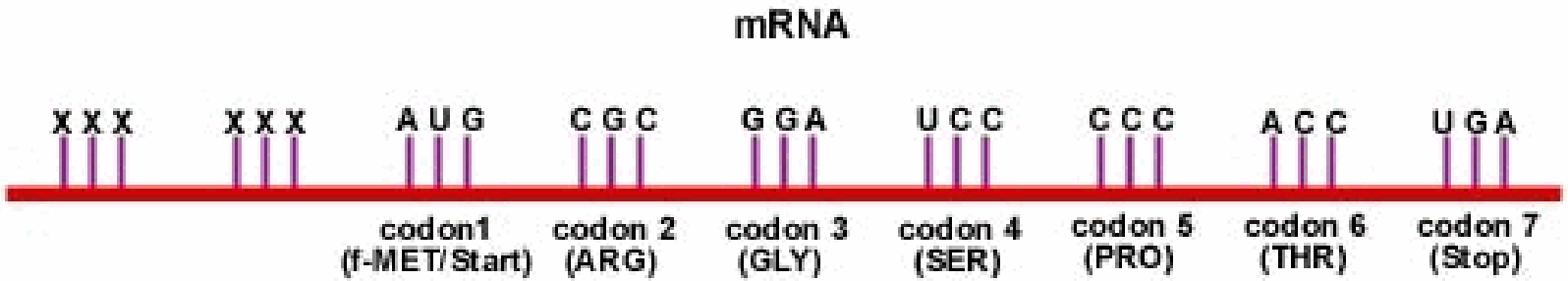


Figure: Translation



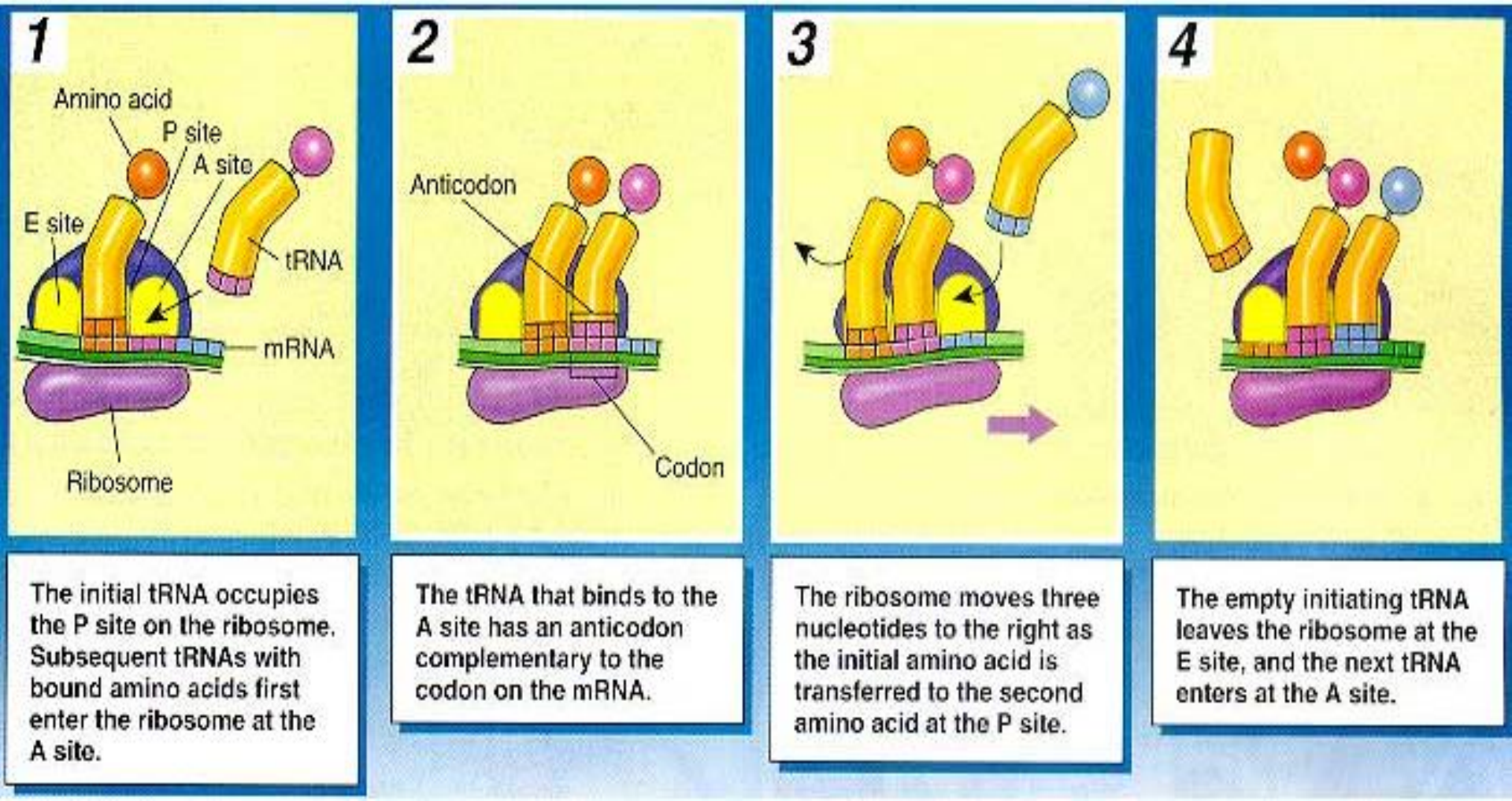
Animation: Translation



Translation

- Ribosome binds to the **mRNA at the start codon (AUG)**
- Recognised by the **initiator tRNA**
- **Aminoacylated tRNA** sequentially bind to the appropriate codon in mRNA by forming complementary base pairs with the tRNA anticodon while ribosome moves from codon to codon along the mRNA
- The Amino acids are added one by one, translated into Polypeptide sequences dictated by DNA and represented by mRNA.
- At the end, a **release factor binds to the stop codon (**UAG, UAA, or UGA**), terminating translation and releasing the** complete polypeptide from the ribosome.

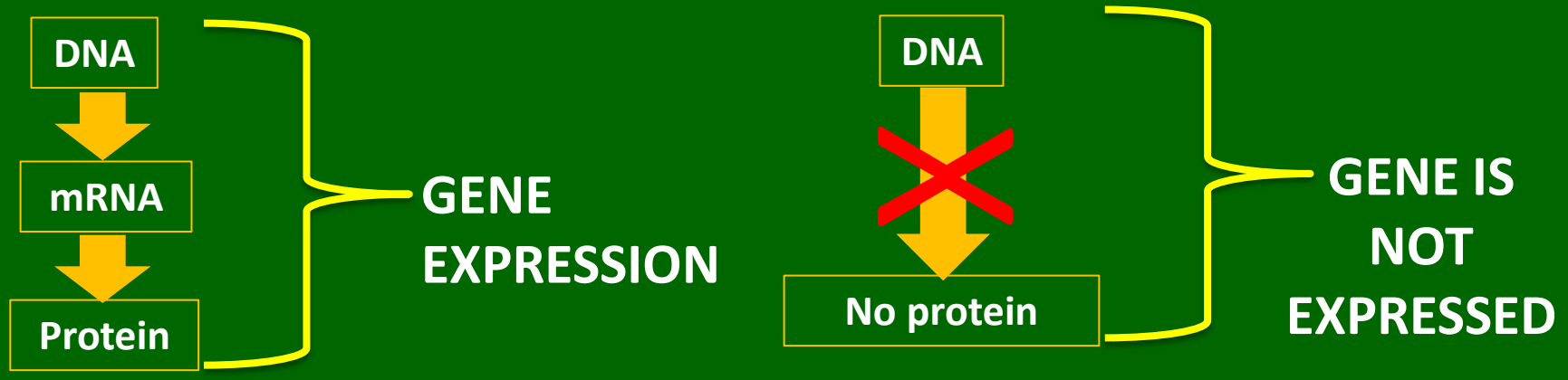
Translation



REGULATION OF GENE EXPRESSION

GENE EXPRESSION

- Gene expression results in the formation of a polypeptide (protein).



REGULATION OF GENE EXPRESSION

- Gene expression is **regulated at various levels.**
- In eukaryotes, the regulation of gene expression could be exerted at
  1. **transcriptional level** (formation of primary transcript)
  2. **processing level** (regulation of splicing),
  3. **transport of mRNA from nucleus to the cytoplasm**,
  4. **translational level**.

REGULATION OF GENE EXPRESSION

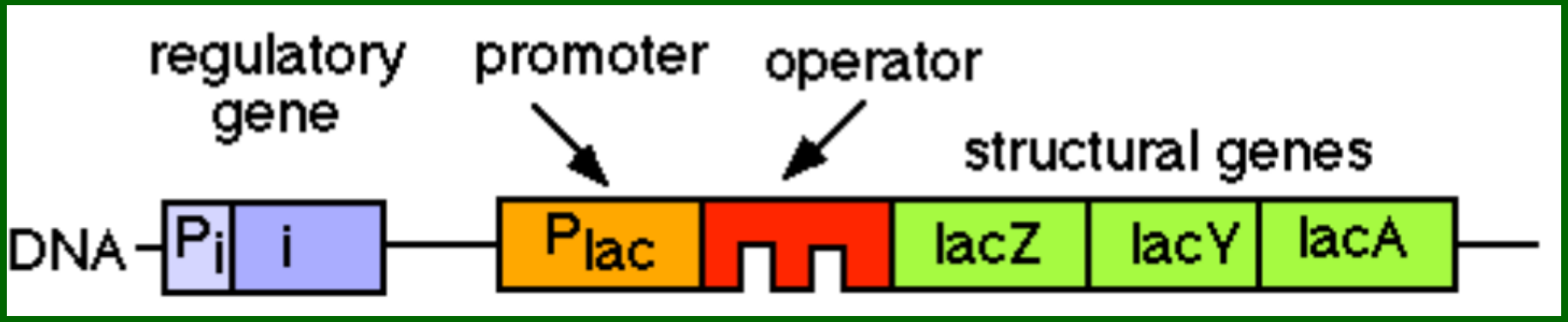
- Genes in a cell are expressed to perform a particular function or set of functions.
- Expression of genes is regulated by;
  - ✓ **metabolic conditions,**
  - ✓ **physiological conditions or**
  - ✓ **environmental conditions**

Regulation Of Gene Expression In Prokaryotes

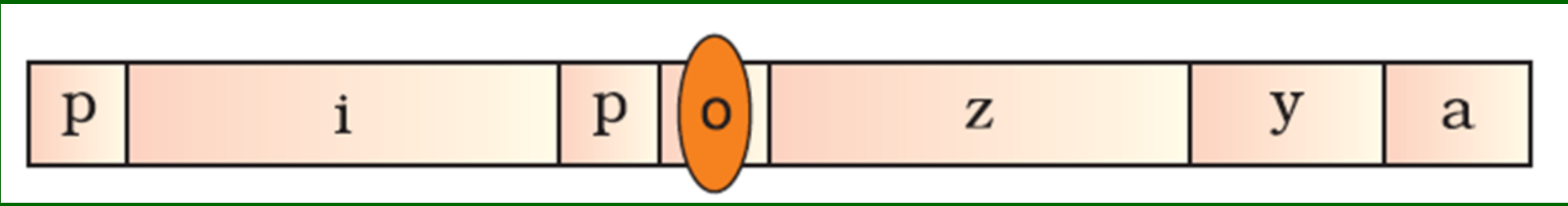
- In prokaryotes, **transcriptional initiation** is the predominant site (i.e. **promoter**) for control of gene expression.
- In a transcriptional unit, the activity of RNA polymerase at a given promoter is regulated by interaction with **Regulatory proteins**, which affect its ability to recognise start sites.
- Regulatory proteins can act both
  - ✓ **positively (activators)** and
  - ✓ **negatively (repressors).**

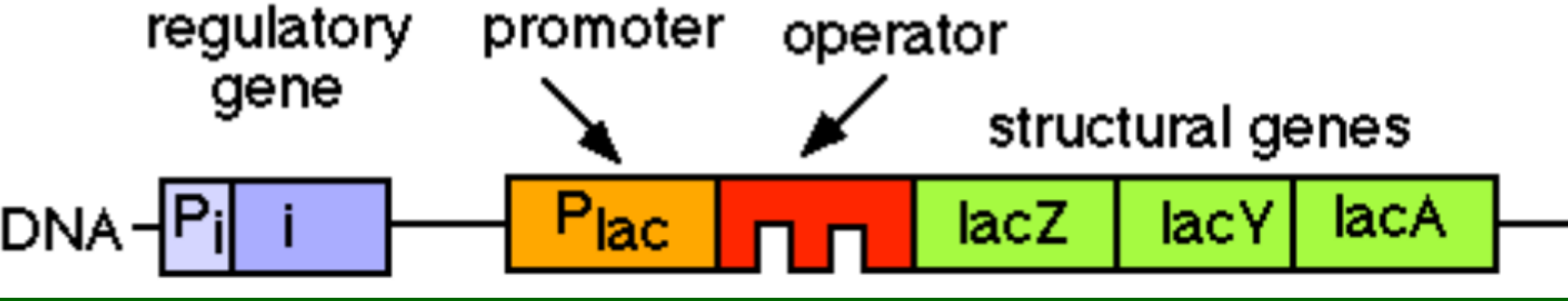
Important terms

- **Operon:** sequence of DNA containing a cluster of genes under the control of a single promoter.
- Examples of bacterial **operon are:** ***lac*** operon, *trp operon, ara operon, his operon, val operon, etc.*
- **lac operon** (lac refers to lactose) is a polycistronic structural gene is regulated *by a* common promoter and regulatory genes.
- **Inducer:** Bio-molecule that favours gene expression
- **Repressor:** Bio-molecule that doesn't allow gene expression



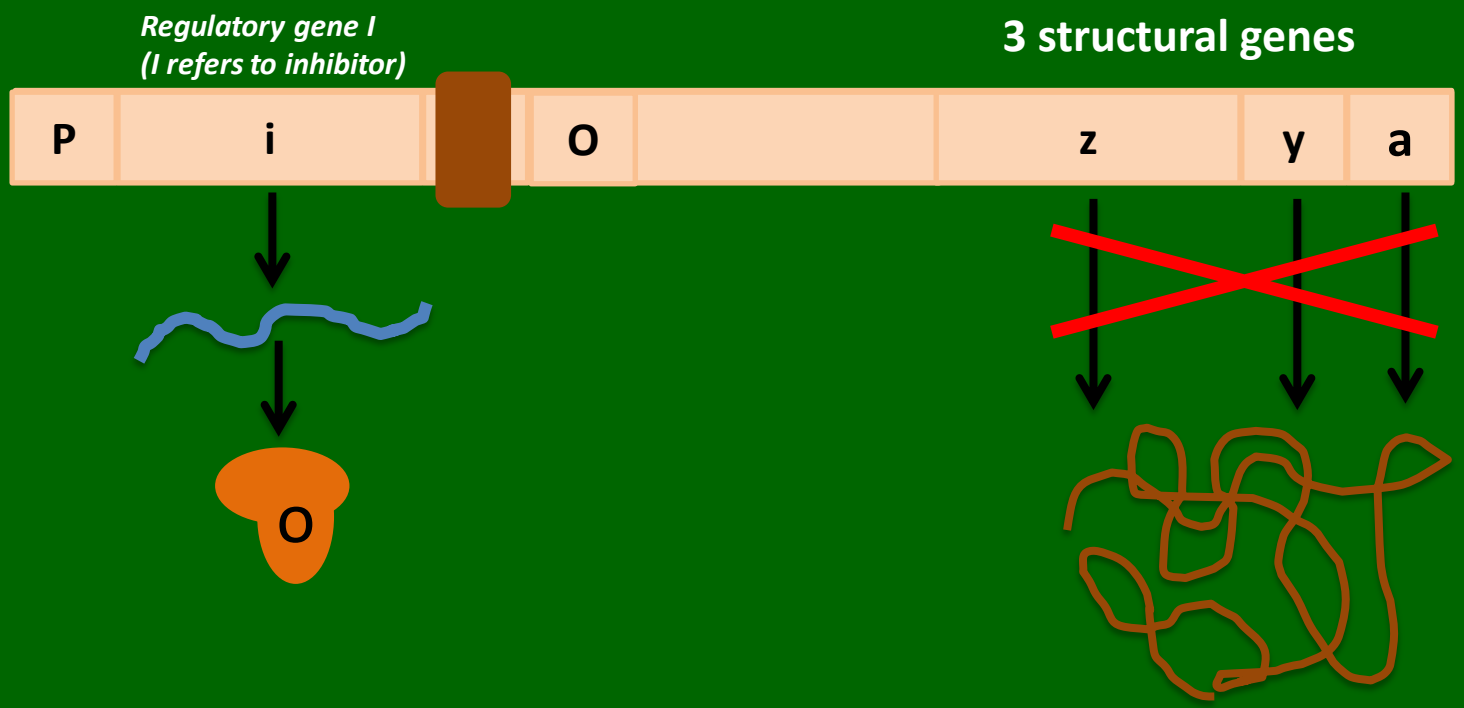
Lac operon


p
i
p
o
z
y
a


regulatory gene
promoter
operator
structural genes
DNA
Pi
i
Plac
lacZ
lacY
lacA

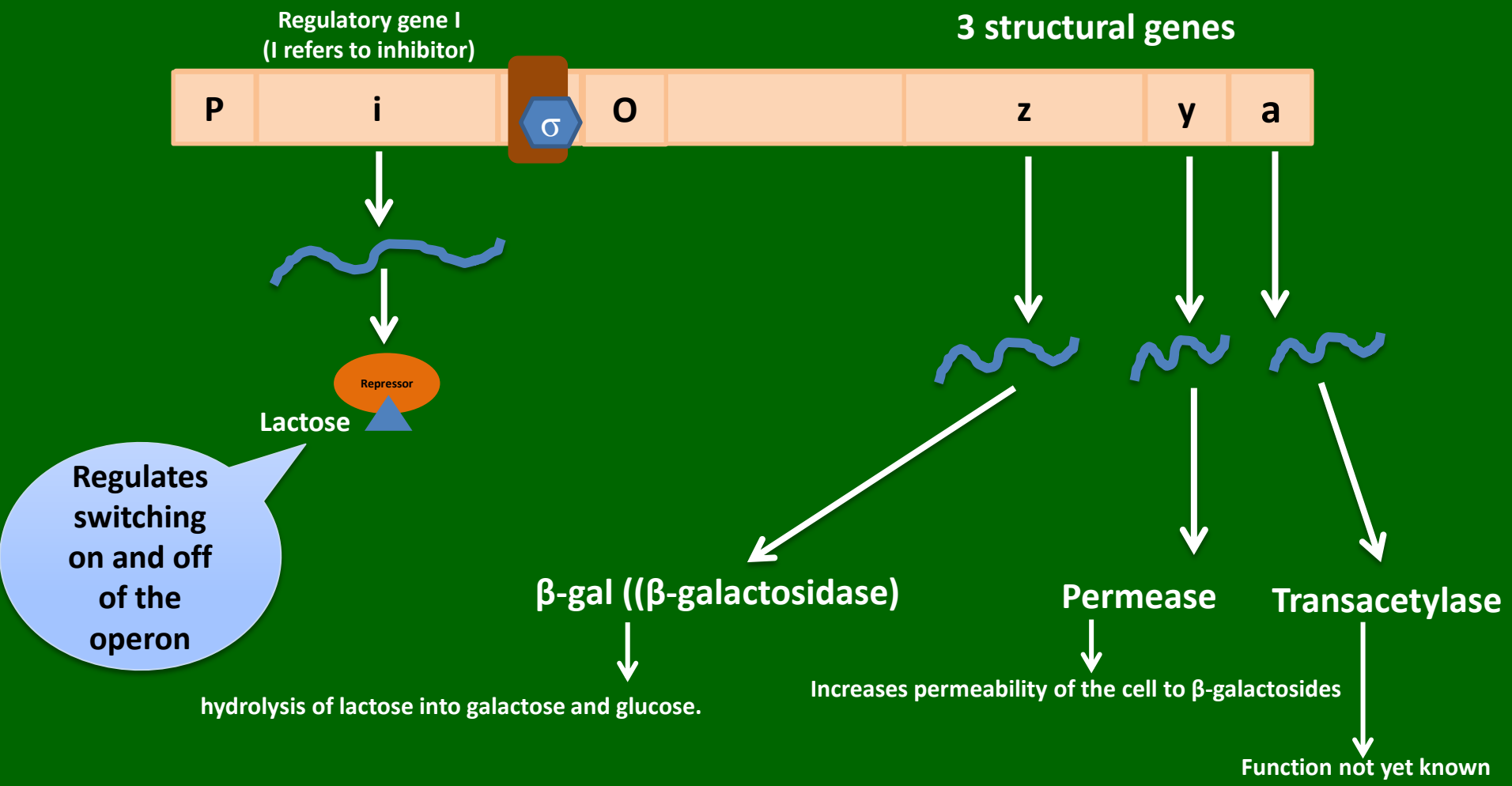
The *Lac operon:* *elucidated by* **Francois Jacob** (geneticist) and **Jacque Monod** (biochemist)

In the absence of inducer (Lactose/Allolactose)



The *Lac operon:* *elucidated by* Francois Jacob (geneticist) and Jacque Monod (biochemist)

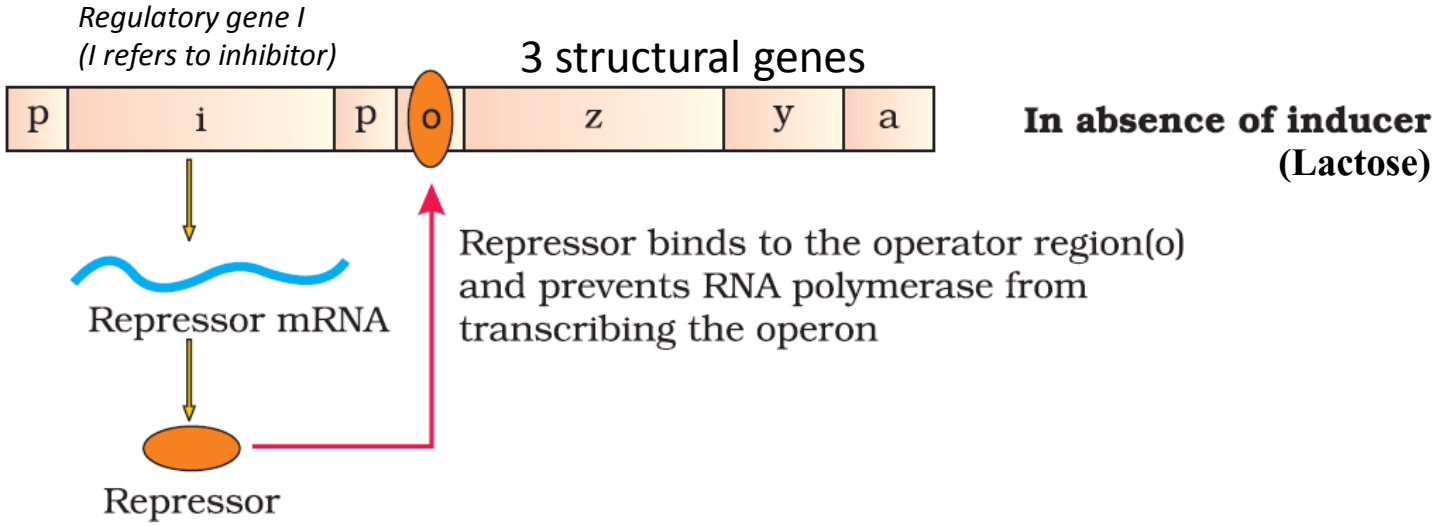
In the presence of inducer (Lactose/Allolactose)



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

The *Lac operon:* *elucidated by* **Francois Jacob** (geneticist) and **Jacque Monod** (biochemist)

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



Home work

- Read the goals of HGP from page numbers 118-121

Regulation Of Gene Expression In Prokaryotes

- The **accessibility of promoter regions** of prokaryotic DNA for RNA polymerase is regulated by the interaction of proteins with **operators**.
- The operator region is adjacent to the promoter elements in most operons
- In most cases the sequences of the operator bind a repressor protein.
- Each operon has its specific operator and specific repressor.
- For example, lac operator is present only in the lac operon and it interacts specifically with lac repressor only.

The *Lac operon*

- The *lac operon consists of one regulatory gene (I gene* –term *I refer to inhibitor)* and three structural genes (*z, y, and a).*
- *The I gene codes for the* ***repressor*** of the *lac operon.*
- *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 β-galactosides.
- The *a gene encodes a transacetylase.*
- *All the three gene products in lac operon Beta-galactosidase,* permease, *transacetylase are required for metabolism of lactose.*

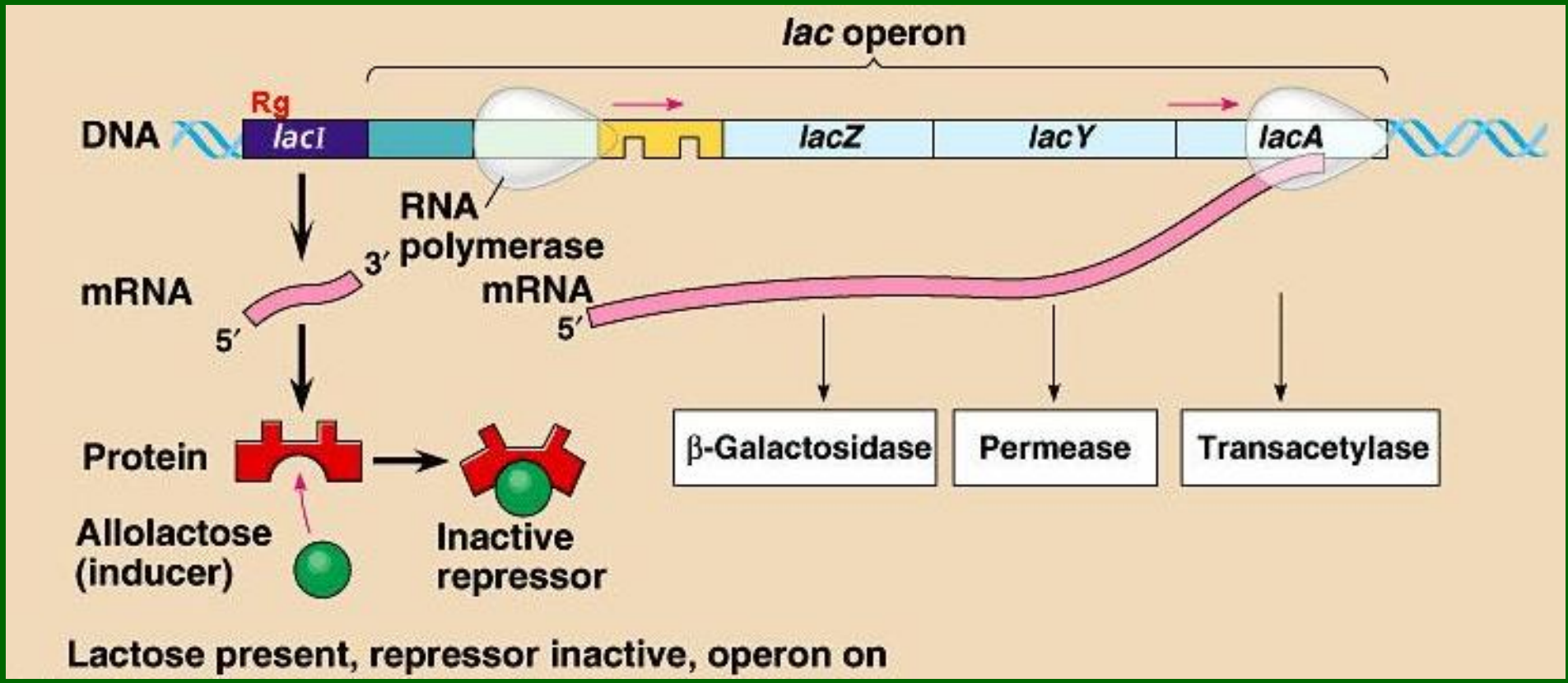
The *Lac operon*

- Lactose is the substrate for the enzyme beta-galactosidase
- Lactose regulates switching on and off of the operon. Hence, it 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 (Remember, 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 *Lac operon*

- 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.***
- ***The repressor protein binds to the operator region of the*** operon and prevents RNA polymerase from transcribing the operon.
- 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. Essentially, regulation of *lac operon can also be visualised as regulation* of enzyme synthesis by its substrate. *Remember, glucose or galactose cannot act as inducers for lac operon.*
- Regulation of *lac operon by repressor is referred to as* ***negative*** **regulation.**

When Lactose is present



Pre-test: Human Genome Project

- What was aim the of human genome project?
- Genetic make-up of an organism lies in the DNA sequences _____(True/False).
- If two individuals differ, then their DNA sequences should also be _____, at least at some places.
- What are the reasons for possibility of launching HGP in the year 1990?
- Why Human Genome Project (HGP) was called a mega project?
- Enlist some non-human organisms whose DNA sequence have also been sequenced?

Goals of HGP

- In which years HGP was started and ended?

Answer: **Started in 1990** and **completed in 2003**

- What were the departments which coordinated the HGP?

Answer: Coordinated by the **U.S. Department of Energy** and the **National Institute of Health**.

- Which Trust became major partner during the early years of the HGP?

Answers: The **Wellcome Trust (U.K.)** became a major partner

- Name some of the contributing countries?

Answer: **Japan, France, Germany, China and others**

Just for information

- DNA of **bacteria, yeast, *Caenorhabditis elegans*** *(a free living* non-pathogenic nematode), ***Drosophila*** *(the fruit fly),* ***plants*** *(rice and Arabidopsis), etc., have also been sequenced.*
- Advantages of learning about non-human organisms DNA sequences;

✓ providing clues to **understanding human biology**

✓ can lead to an understanding of their natural capabilities that can be applied toward **solving challenges in health care, agriculture, energy production, environmental remediation.**

Why HGP?

- Knowledge about the **effects of DNA variations among individuals** can lead to revolutionary new ways
  - **to diagnose,**
  - **treat and**
  - **prevent thousands of disorders that affect human beings.**
- Aimed **to find complete DNA sequence of human genome**

Human Genome Project

- Assumptions leading to the quest of finding complete DNA sequence of human genome
  - ✓**Genetic make-up of an organism lies in the DNA sequences.**
  - ✓**If two individuals differ, then their DNA sequences should also be different, at least at some places**.

Human Genome Project

- Very ambitious HGP (launched in the year 1990) was possible because;
  - ✓**Availability of genetic engineering techniques**
  - ✓**Possibility to isolate and clone any piece of DNA**
  - ✓**Availability of simple and fast techniques for determining DNA sequences**

Some goals of HGP

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

Why HGP is a MEGA PROJECT?

- Human genome is said to have approximately 3 x 10^9 bp, and its estimated cost was approximately 9 billion US dollars *i.e if the cost of sequencing required is US $ 3 per bp (the estimated cost in the beginning*.
- HGP **require the expertise** and **creativity of tens of thousands of scientists** from varied disciplines in both the public and private sectors worldwide.
- The **enormous amount of data* generated** necessitated the **use of high speed computational devices** for data **storage and retrieval, and analysis**.
- HGP was closely associated with the rapid development of **Bioinformatics** (i.e a new area in biology).

If the obtained sequences were to be stored in typed form in books, and if each page of the book contained 1000 letters and each book contained 1000 pages, then 3300 such books would be required to store the information of DNA sequence from a single human cell.

Methodologies of HGP

- The methods involved two major approaches.

1. **Expressed Sequence Tags (ESTs):** focused on identifying all the genes that expressed as RNA**.**
2. **Sequence Annotation:** blind approach of simply **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.**

Protocol of HGP

Total DNA from a cell is isolated

↓

Converted into random fragments of relatively smaller sizes

↓

Cloned* in suitable host using specialised vectors; BAC (bacterial artificial chromosomes), and YAC (yeast artificial chromosomes).

↓

Amplified DNA are sequenced using automated DNA sequencers i.e. specialised computer based programs .

*Cloning resulted into amplification of each piece of DNA fragment

Automated DNA sequencers works on the principle of a method developed by Frederick Sanger. He is also credited for developing method for determination of amino acid sequences in proteins.

Methodologies of HGP

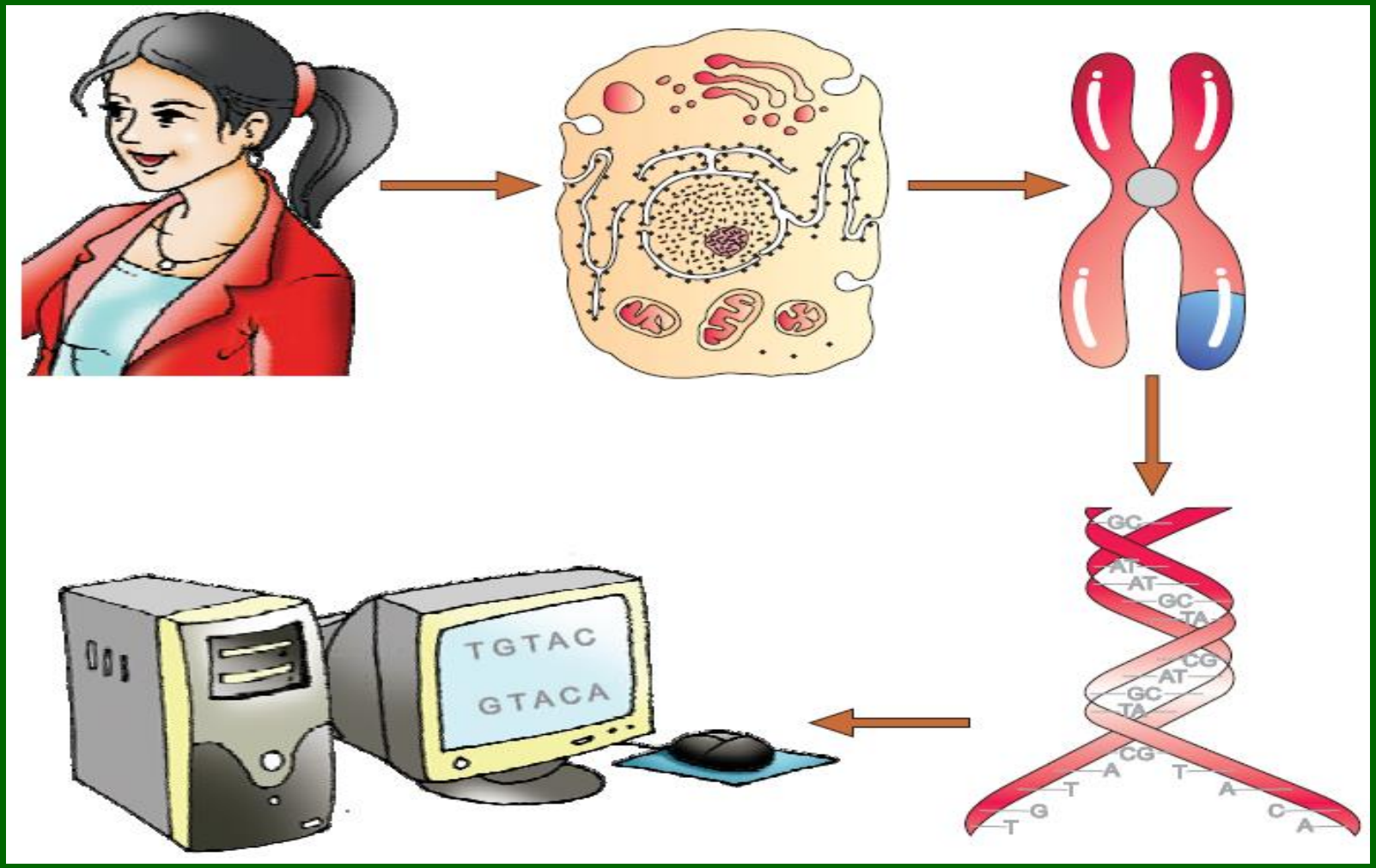


Figure A representative diagram of human genome project

Sequencing DNA fragments; A step of Method in HGP

- These sequences were arranged based on some overlapping regions present in them.
- This generation of overlapping fragments is required for sequencing.
- Alignment of these sequences was humanly not possible.
- Specialised computer based programs were developed.
- These sequences were subsequently annotated and were assigned to each chromosome.
- The sequence of chromosome 1 was completed only in May 2006 (this was the last of the 24 human chromosomes – 22 autosomes and X and Y – to be sequenced).
- Another challenging task was assigning the **genetic and physical maps** on the genome.
- This was generated using information on polymorphism of restriction endonuclease recognition sites, and microsatellite DNA (repetitive DNA sequences).

Some Salient Features of Human Genome

1. The human genome **contains 3164.7 million nucleotide bases.**
2. The average gene consists of 3000 bases, but sizes vary greatly, with the **largest known human gene being dystrophin at 2.4 million bases.**
3. The total number of genes is estimated at 30,000 much lower than previous estimates of 80,000 to 1,40,000 genes.
4. Almost all **(99.9 per cent) nucleotide bases are exactly the same in all people.**
5. The **functions are unknown for over 50 per cent of discovered genes.**

Some Salient Features of Human Genome

6. Less than **2 per cent of the genome codes for proteins**.
7. **Repeated sequences make up very large portion** of the human genome.
8. Repetitive sequences are stretches of DNA sequences that are **repeated many times** (hundred to thousand times). They are thought to have **no direct coding functions**, but they shed light on **chromosome structure, dynamics and evolution**.
9. Chromosome 1 has most genes (2968), and the Y has the fewest (231).
10. Scientists have identified about **1.4 million locations where single base DNA differences (SNPs – single nucleotide polymorphism,** pronounced as 'snips') occur in humans. This information promises to **revolutionise the processes of finding chromosomal locations for disease-associated sequences and tracing human history.**

Applications and Future Challenges

- Deriving meaningful knowledge from the DNA sequences will **help in understanding of biological systems**.
- Complete sequence of HG sequence when properly studied, will **enable radically new approach to biological research.**
- **Can be used to study all the genes in a genome**, for example,
  - all the transcripts in a particular tissue or organ or tumor, or
  - how tens of thousands of genes and proteins work together in interconnected networks to orchestrate the chemistry of life.

Home work

- Read **DNA FINGERPRINTING** from page numbers 121-122

Why we are different?

Differences in base sequences of human genome (i.e. 3 × 10^9 bp) make every individual unique in their phenotypic appearance.

DNA FINGERPRINTING

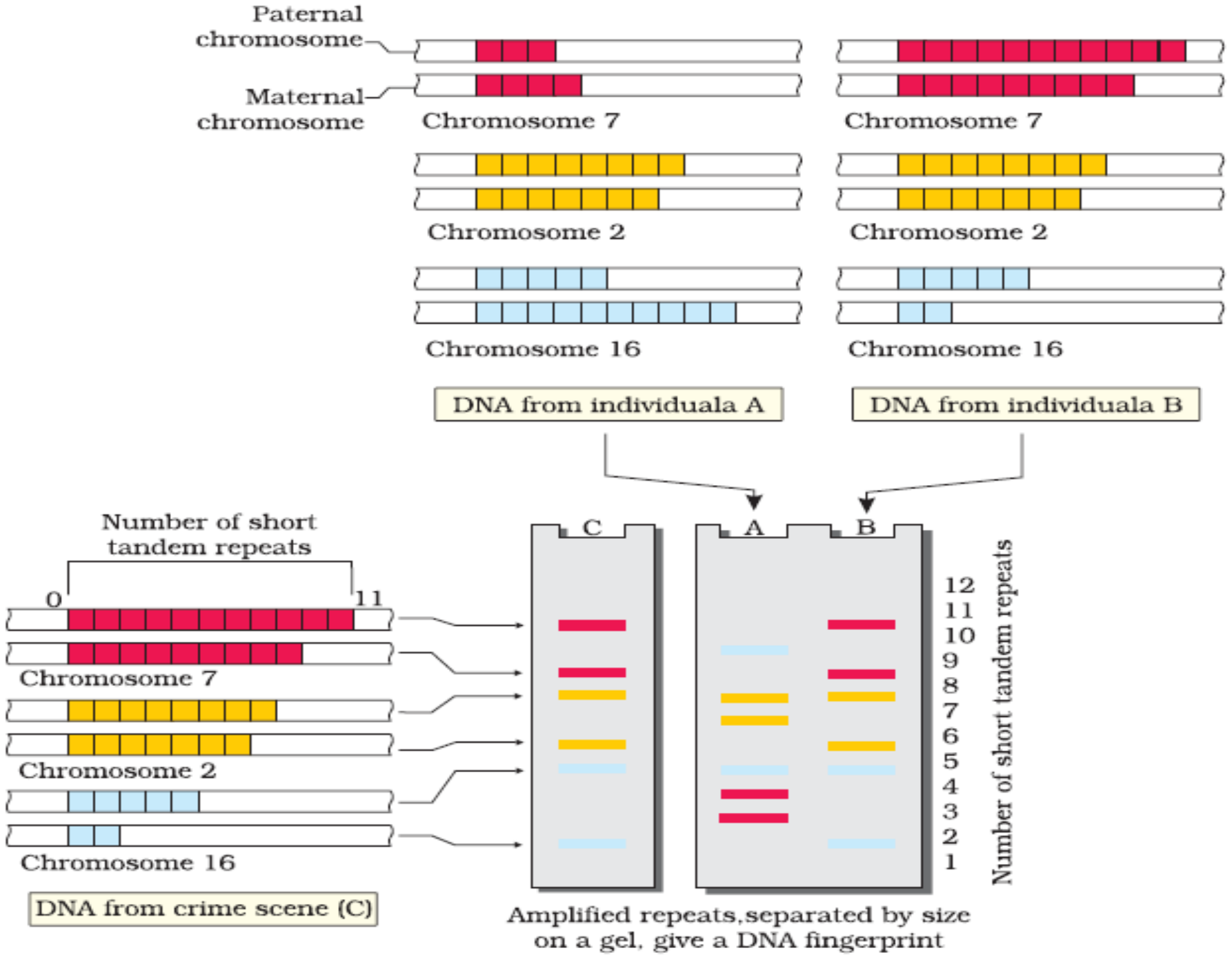
- DNA fingerprinting is a **very quick way to compare the DNA sequences of any two individuals or among individuals of a population.**
- DNA fingerprinting was **developed by Alec Jeffreys.**
- Uses of DNA fingerprinting
  - ✓ **Forensic applications**
  - ✓ **Used in paternity testing**
  - ✓ **Used in study of genetic biodiversity**
  - ✓ **Used in evolutionary biology**
- DNA fingerprinting **involves identifying differences in repetitive DNA***.

A small stretch of non coding,* shows *very high degree of polymorphism, extensively repeating DNA, form a large portion of human genome

Protocol of DNA Fingerprinting

1. **Isolation of DNA**
2. Digestion of DNA by **restriction endonucleases**
3. Separation of DNA fragments by **electrophoresis**
4. Transferring (blotting) of separated DNA fragments to synthetic membranes, such as **nitrocellulose or nylon (i.e. Southern blotting)**
5. **Hybridisation** using **radio labelled VNTR* probe**
6. Detection of hybridised DNA fragments (*usually appears dark bands*) by **autoradiography**

***Variable Number of Tandem Repeats (VNTR)** -a mini-satellite DNA


Paternal chromosome
Maternal chromosome
Chromosome 7
Chromosome 7
Chromosome 2
Chromosome 2
Chromosome 16
Chromosome 16
DNA from individuala A
DNA from individuala B
Number of short tandem repeats
0
11
Chromosome 7
Chromosome 2
Chromosome 16
DNA from crime scene (C)
C
A
B
12
11
10
9
8
7
6
5
4
3
2
1
Number of short tandem repeats
Amplified repeats,separated by size on a gel, give a DNA fingerprint

Satellite /Repetitive DNA

Depending on base composition *(A:T rich or G:C rich), length of segment, and number of repetitive units*

Mini-satellites

Micro satellites

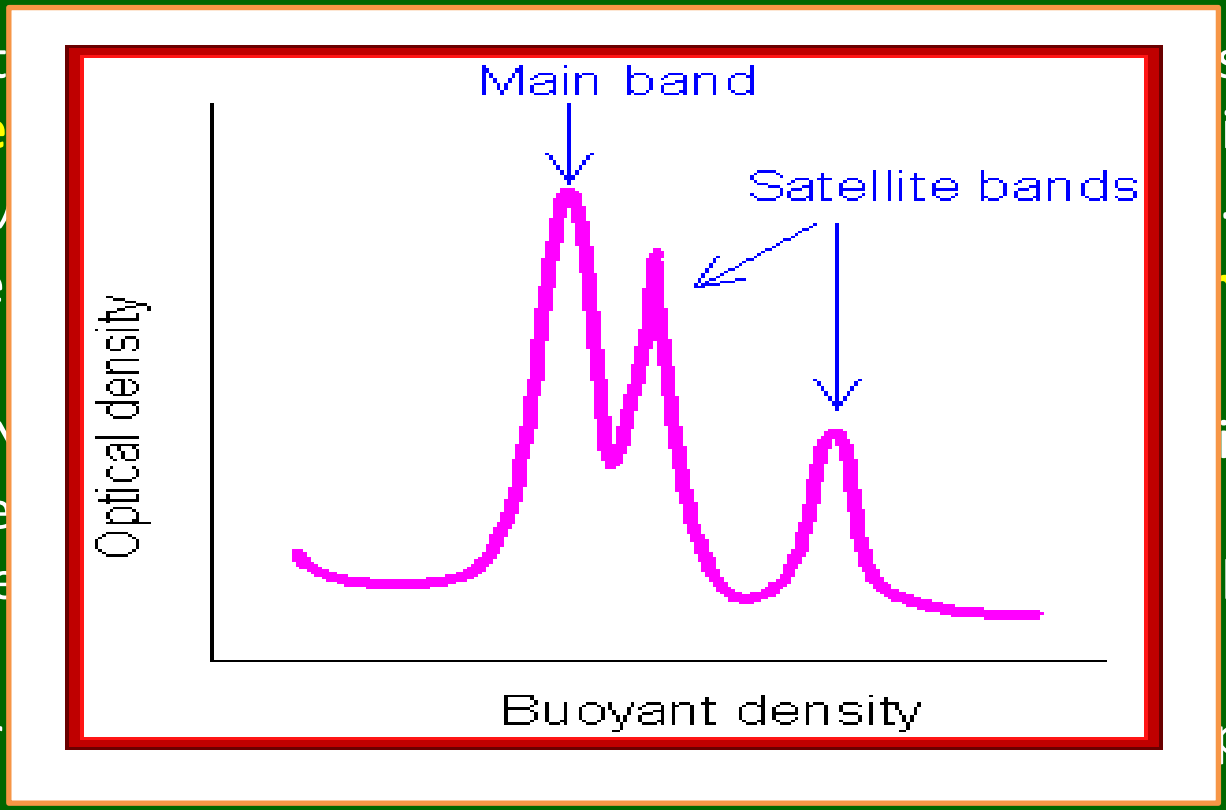
Variable Number of Tandem Repeat (VNTR)

(Small DNA sequence arranged tandemly in many copy number, copy number varies from chromosome to chromosome in an individual. The numbers of repeat show very high degree of polymorphism.

Read paragraph 2 of DNA fingerprinting on page number 121

With an example explain repetitive DNA

- The **repetitive DNA** form different peaks unlike **bulk genomic DNA** during **density gradient centrifugation.**
- The **bulk DNA** forms a **major peak** and the other **small peaks** are referred to as **satellite DNA.**
- The sat… such as **micro satellite**… ion *(A:T rich or G:C rich)*…
- Satellite… **m** and form the basis of…
- Since DN… in, bone, saliva, sperm e… polymorphism, these be… lications.
- As the… children, DNA fingerpr… putes.



Write a note on DNA Polymorphism

- Polymorphism is variation at genetic level
- Polymorphism in DNA sequence is the basis of **genetic mapping of human genome** and **DNA fingerprinting.**
- Polymorphism **arises due to mutations**.
- New mutations may arise in an individual either in
  - ✓**somatic cells or**
  - ✓**germ cells** *(cells that generate gametes in sexually reproducing organisms)*.
- If a germ cell mutation is inherited to successive generations successfully, then, it can spread to the other members of population (through sexual reproduction).

DNA Polymorphism

- **Allelic sequence variation** has traditionally 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.
- If an **inheritable mutation** is observed in a population at high frequency, it is referred to as **DNA polymorphism.**
- The probability of such variation to be observed in non coding DNA sequence would be higher as mutations in these 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.
- DNA polymorphisms may be **single nucleotide change**, very **large scale changes** etc.
- For evolution and speciation, such polymorphisms play very important role.

DNA Fingerprinting technique

- **Alec Jeffreys used a satellite DNA** as probe.
- Earlier, the technique involved **Southern blot hybridisation** using **radio labelled VNTR** as a probe.
- **Variable Number of Tandem Repeats (VNTR)** belongs to a class of satellite DNA referred to as **mini-satellite.**

What is VNTR?

- The VNTR belongs to a class of satellite DNA referred to as mini-satellite.
- VNTR is small DNA sequence arranged tandemly in many copy numbers.
- The copy number varies from chromosome to chromosome in an individual.
- The numbers of repeat show very high degree of polymorphism.
- Size of VNTR varies from 0.1 to 20 kb.

What is VNTR?

- 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
- DNA from a single cell is enough to perform DNA fingerprinting analysis.
- Currently, many different probes are used to generate DNA fingerprints.

Chapter Ends with two Question???

What is STR?

Why STR is used in DNA fingerprinting these days?

- End of chapter…….

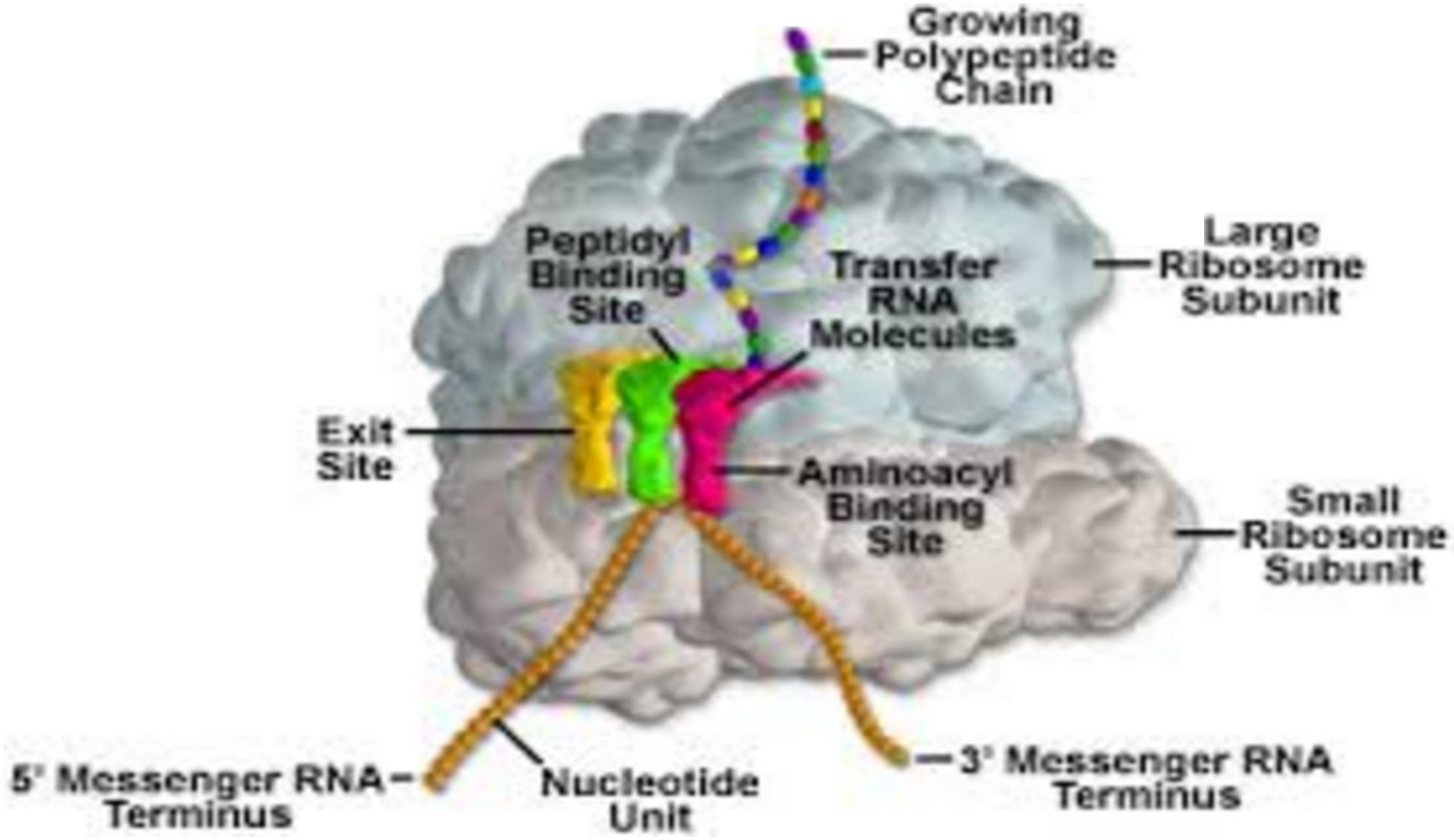
Satellite DNAs consist of short sequences about five to a few hundred base pairs in up to several million base pairs of DNA.

Types of satellite DNA.

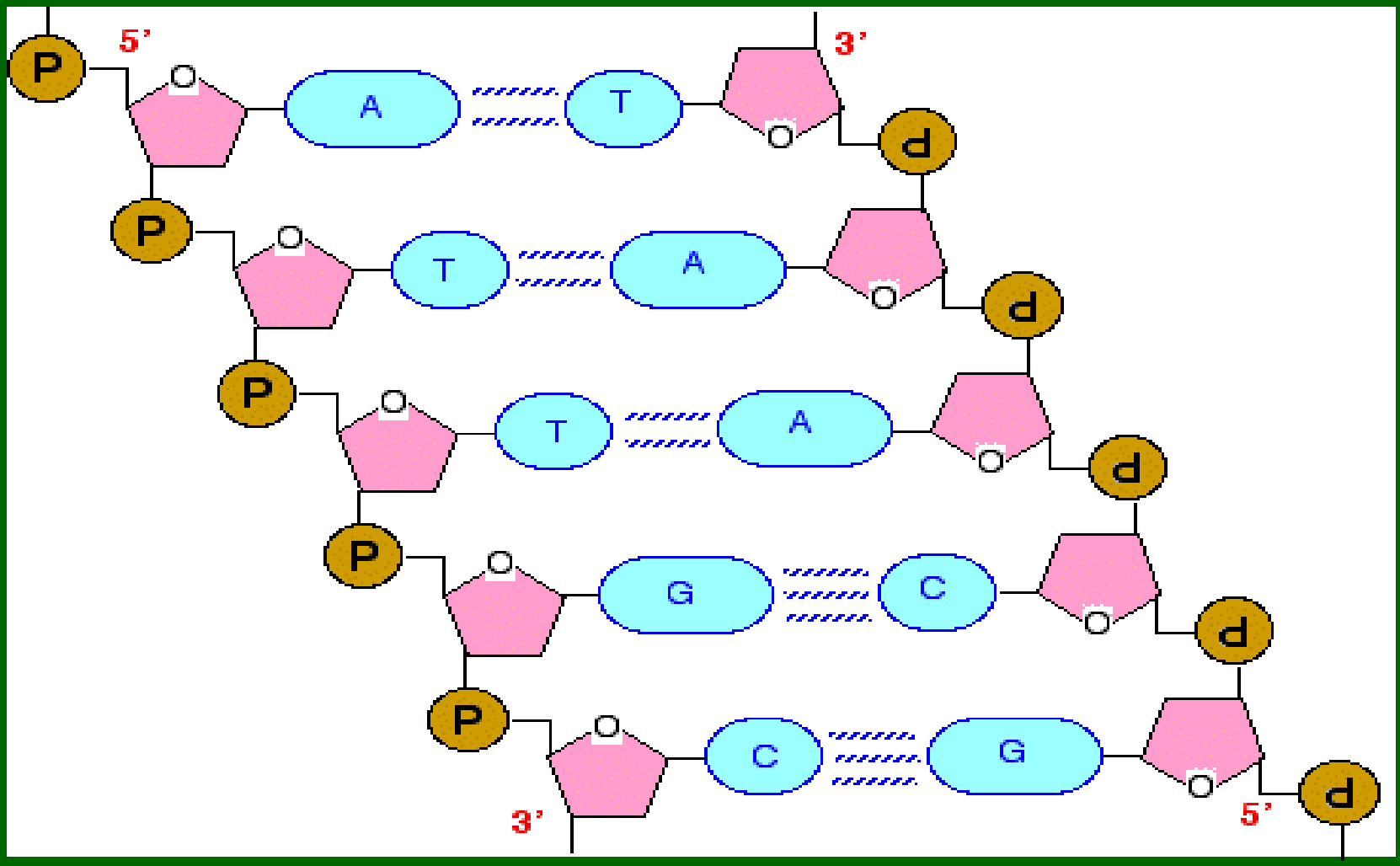
A.Minisatellite DNA - Minisatellite DNA or VNTR (D1S7) sequence range from about 10 to 100 base pairs and found in sizeable cluster containing as many as 3000 repeats.

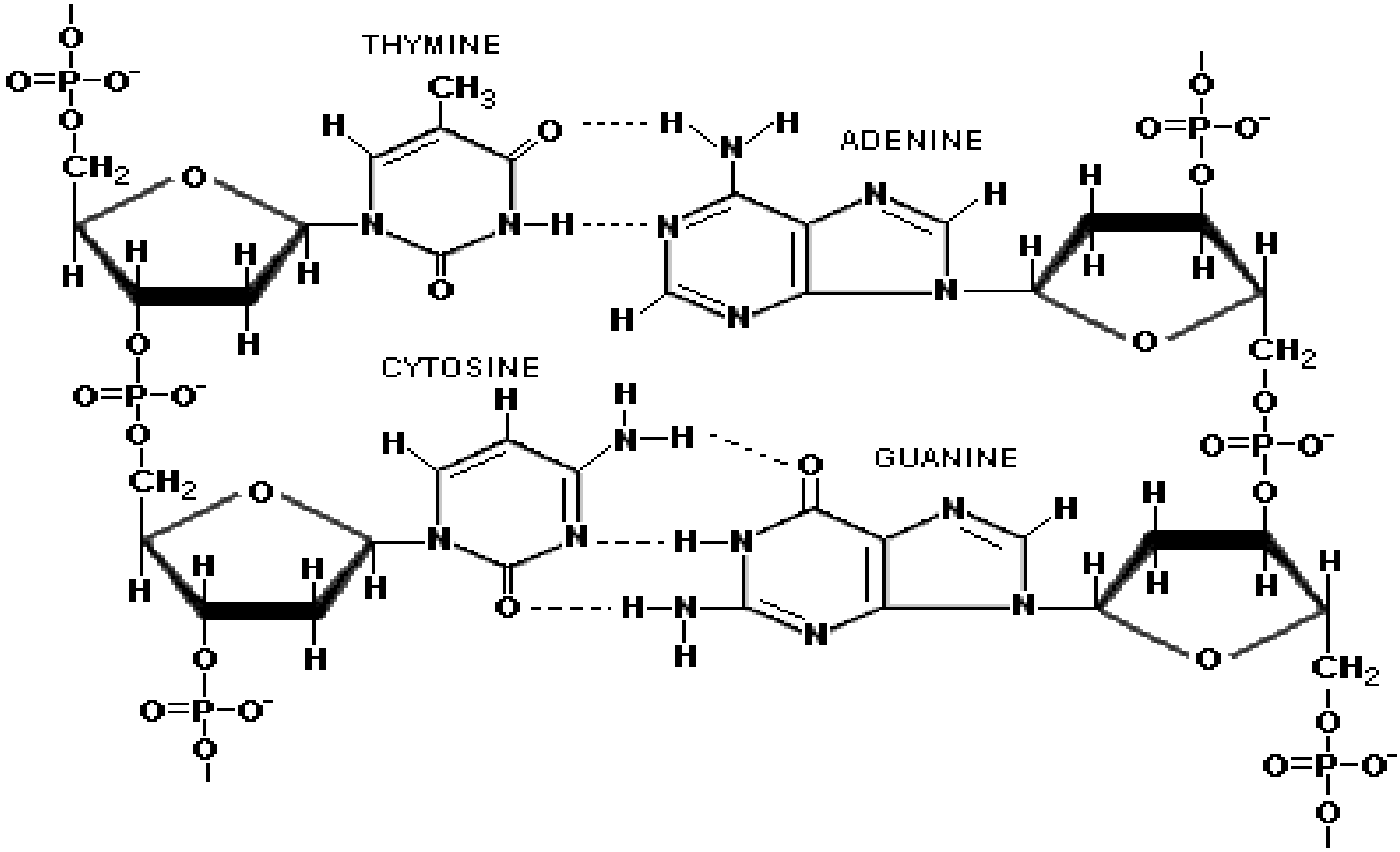
B.Microsatellite DNA- Microsatellite DNA or STR are the shortest sequence (1 to 5 base pair long) and are typically present in small clusters of about 10 to 40 base pair in length

Structure of Ribosome



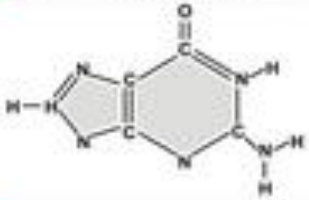
Struture of DNA



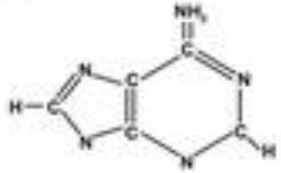

THYMINE
ADENINE
CYTOSINE
GUANINE

Cytosine

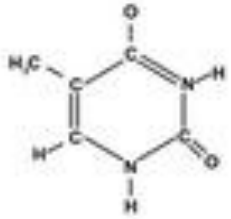
Guanine



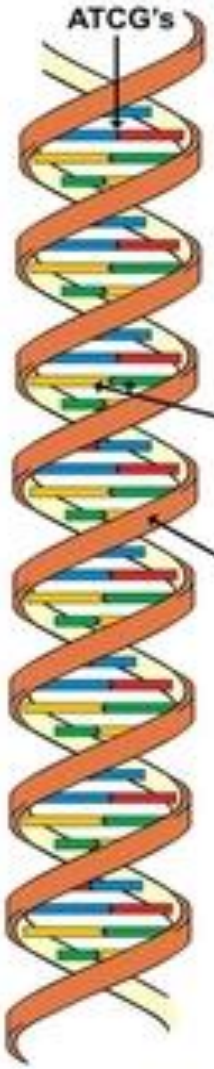
Adenine



Thymine

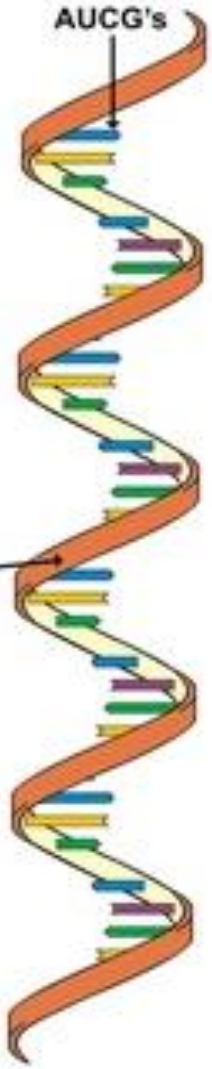


Nitrogenous Bases



DNA

Deoxyribonucleic Acid



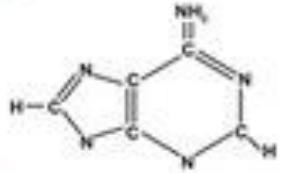
RNA

Ribonucleic Acid

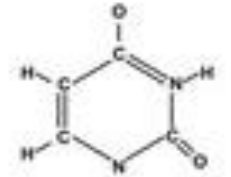
Cytosine

Guanine

Adenine

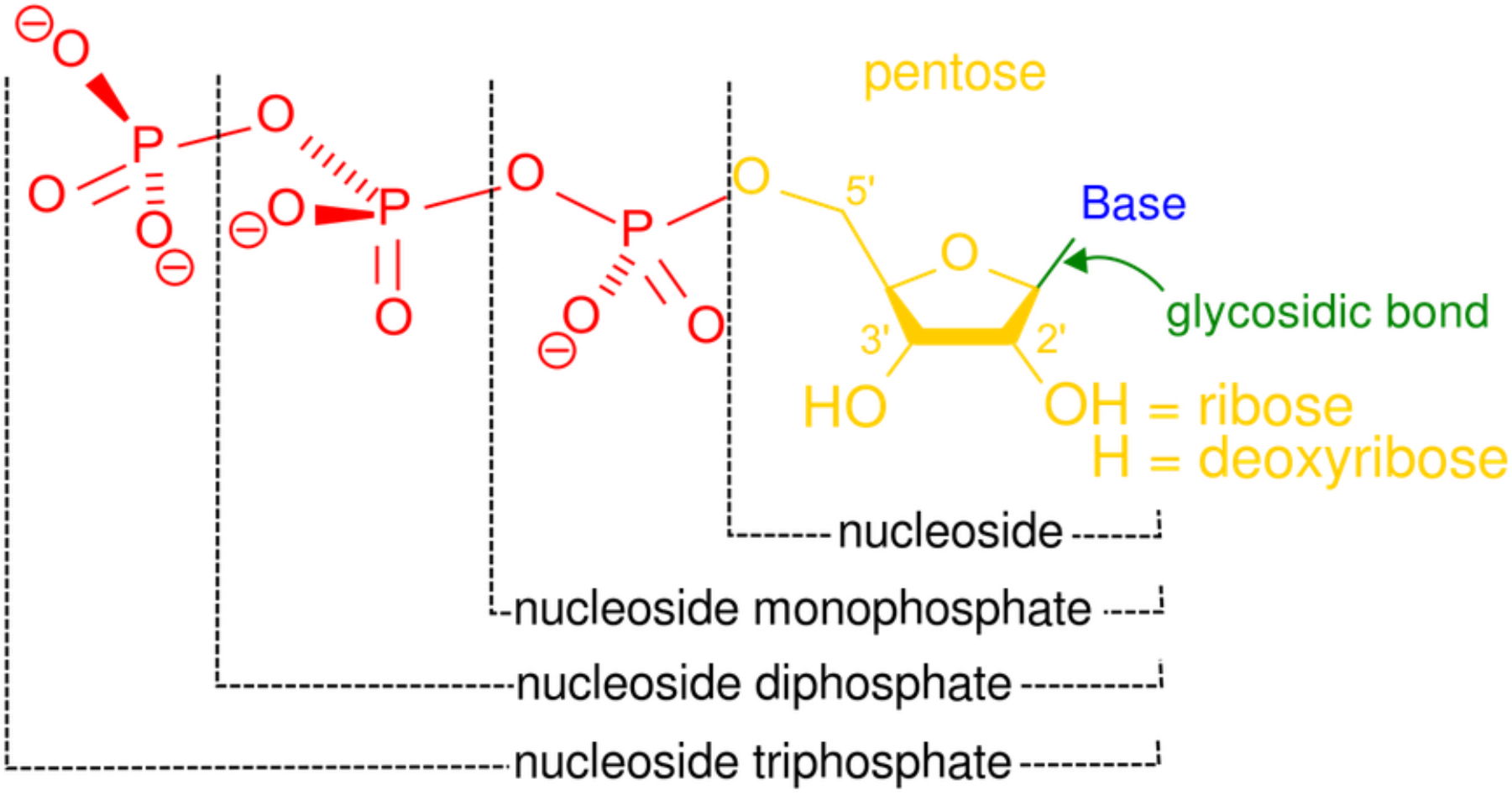


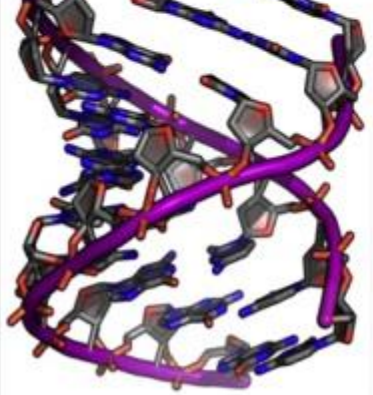
Uracil



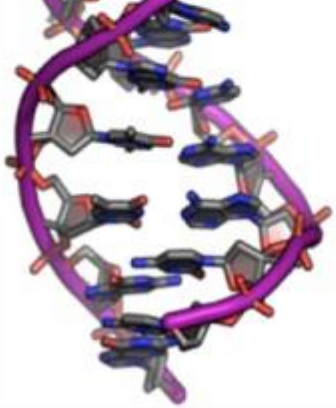
Replaces Thymine in RNA

Nitrogenous Bases

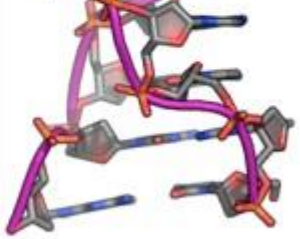

pentose
Base
glycosidic bond
5'
3'
2'
HO
OH = ribose
H = deoxyribose
nucleoside
nucleoside monophosphate
nucleoside diphosphate
nucleoside triphosphate



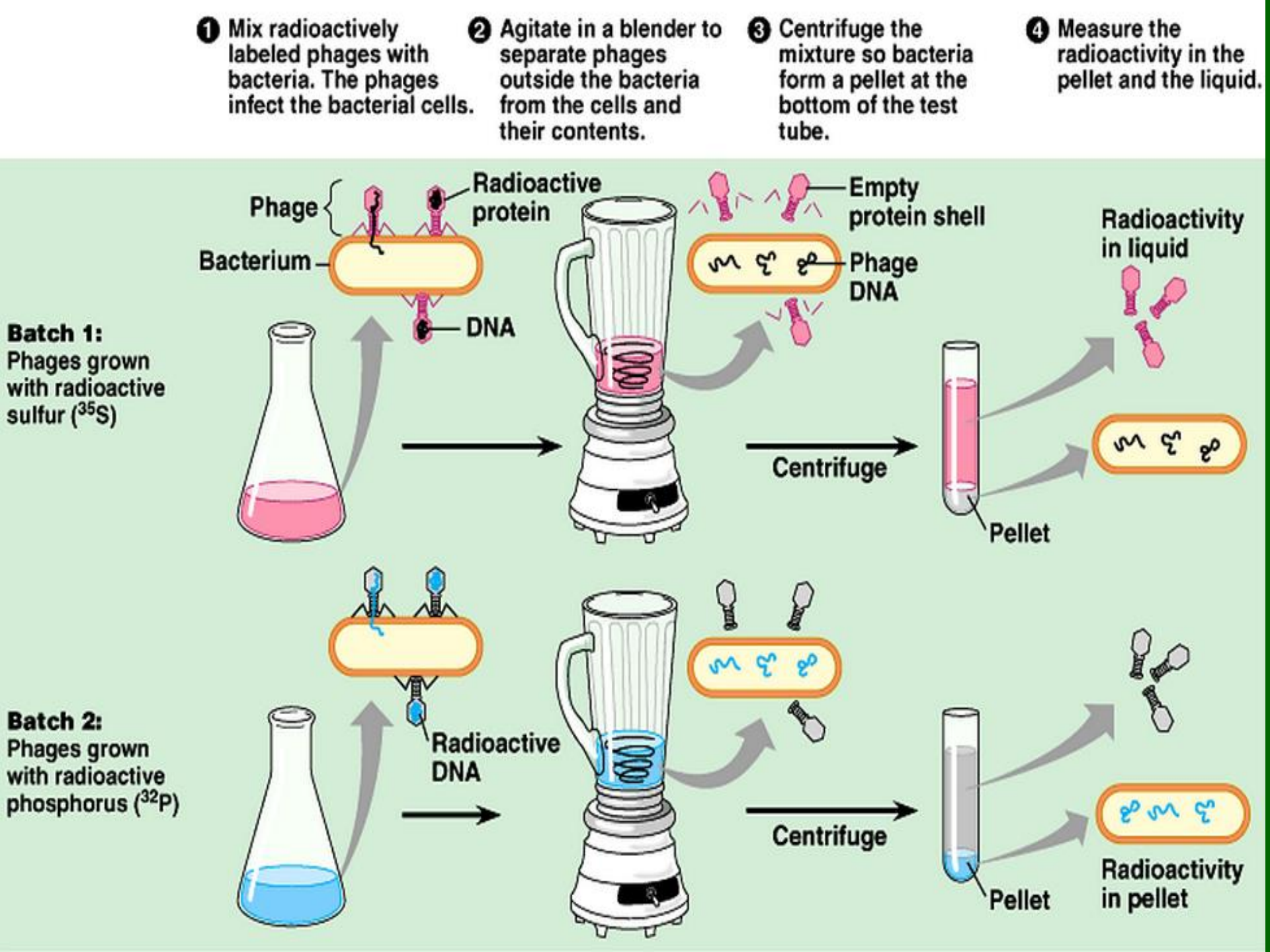
Twist = 33°
Rise = 2.56 Å
Roll = 6°
Inclin. = 21°
x-Dis. = -4.5 Å
P-Tw = -7.5°

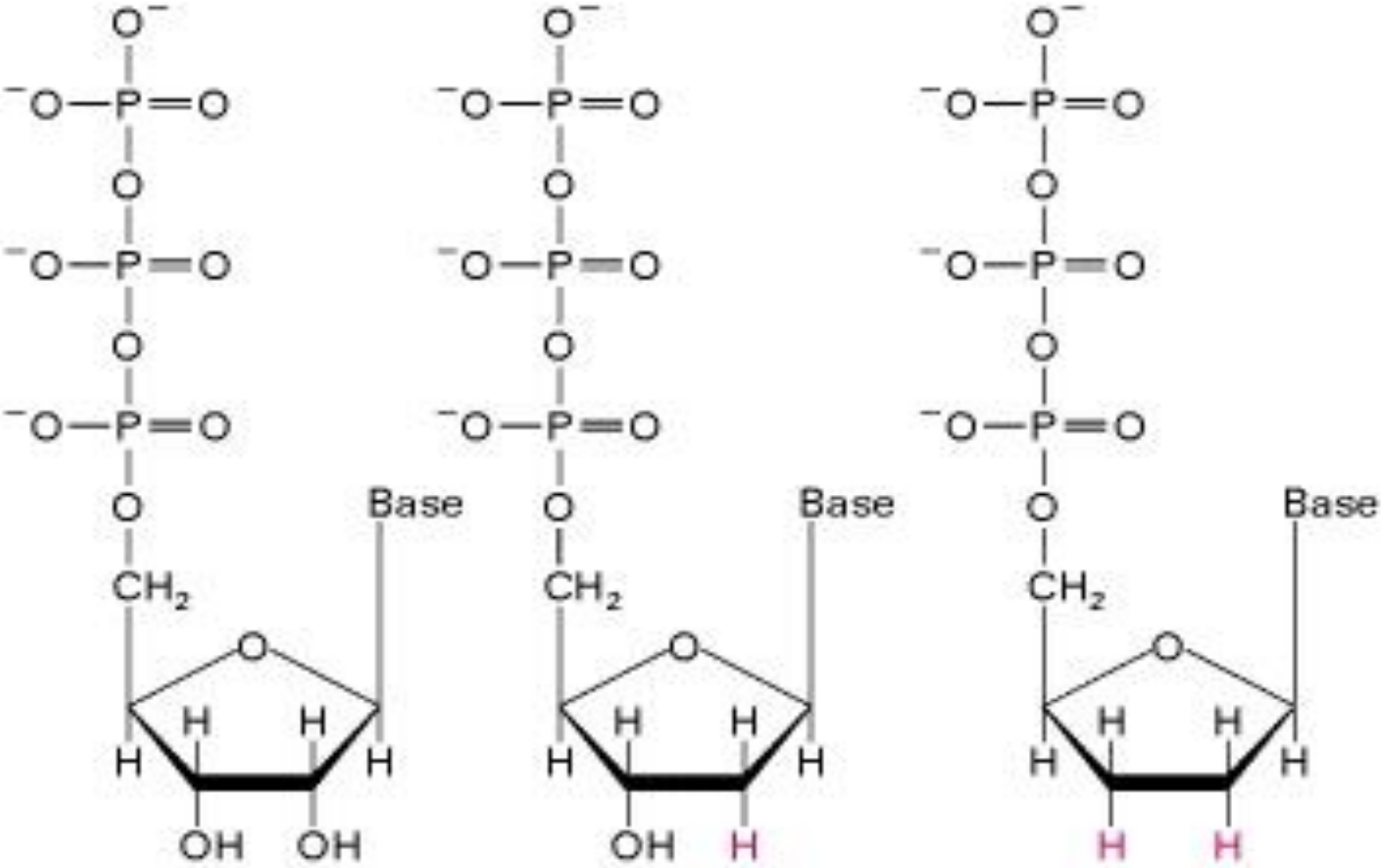


Twist = 36°
Rise = 3.38 Å
Roll = 0°
Inclin. = -6.0°
x-Dis. = 0.23 Å
P-Tw = -4.4°



<Twist
<Rise>
<Roll>
<Inclin
<*x*-Dis.
<P-Tw


1 Mix radioactively labeled phages with bacteria. The phages infect the bacterial cells.
2 Agitate in a blender to separate phages outside the bacteria from the cells and their contents.
3 Centrifuge the mixture so bacteria form a pellet at the bottom of the test tube.
4 Measure the radioactivity in the pellet and the liquid.
Phage
Radioactive protein
Bacterium
DNA
Empty protein shell
Phage DNA
Radioactivity in liquid
Batch 1:
Phages grown with radioactive sulfur (^{35}S)
Centrifuge
Pellet
Radioactive DNA
Batch 2:
Phages grown with radioactive phosphorus (^{32}P)
Centrifuge
Pellet
Radioactivity in pellet

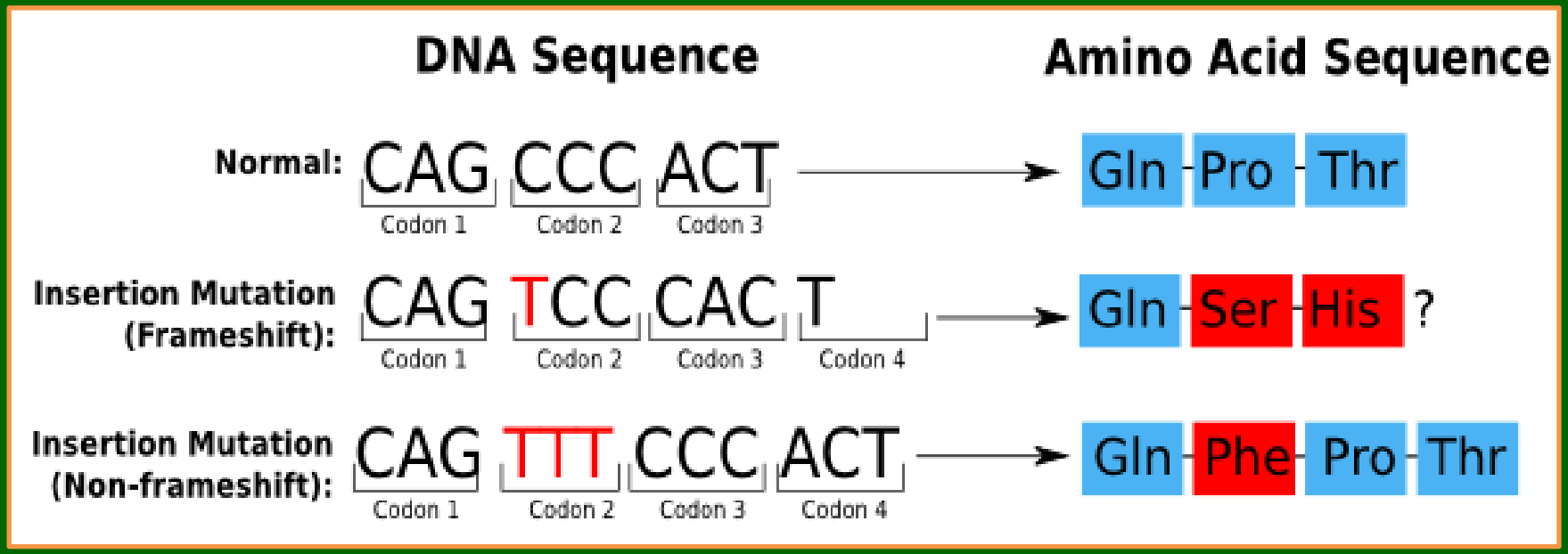


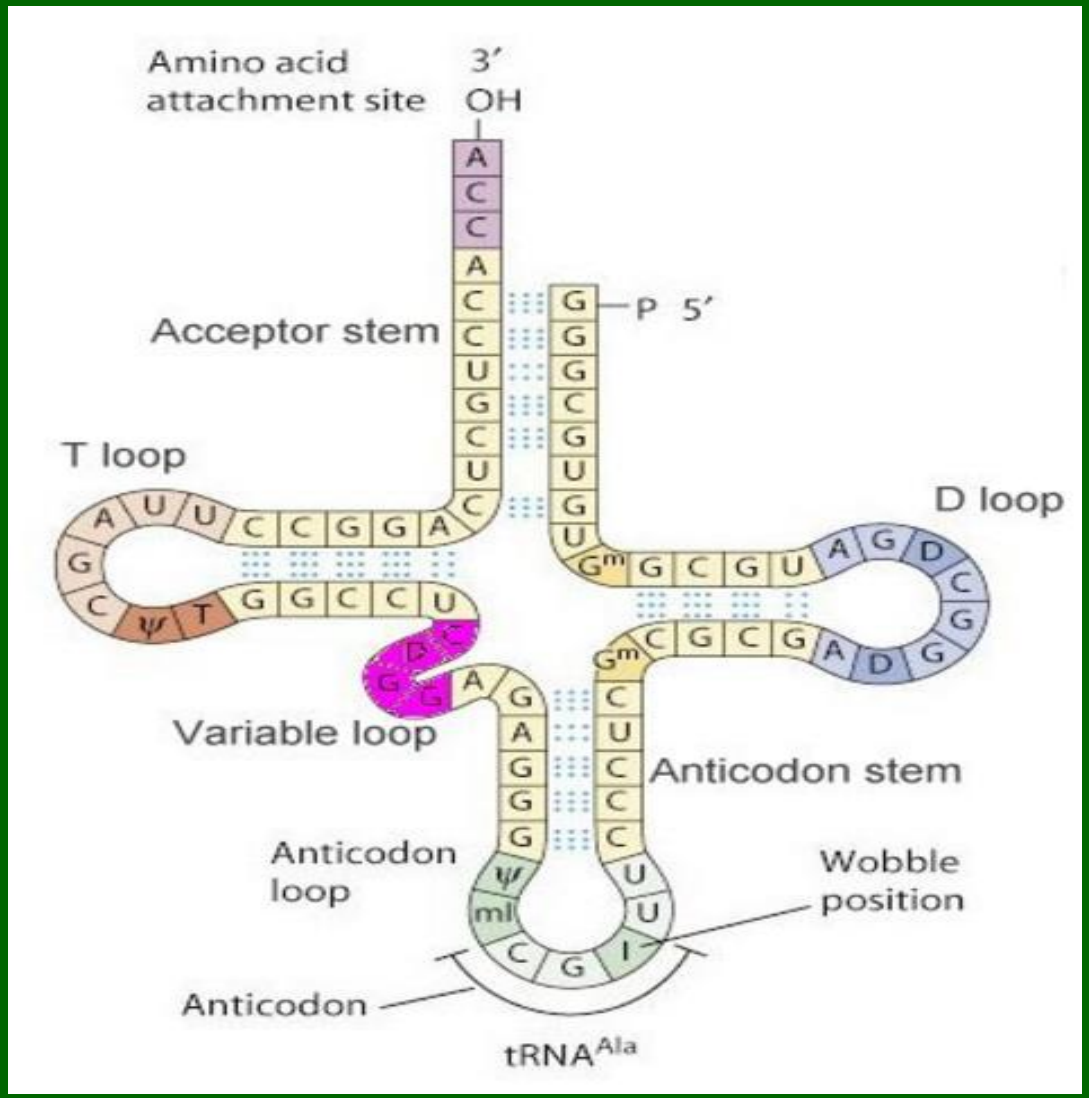
Ribonucleoside triphosphate (NTP)

Deoxyribonucleoside triphosphate (dNTP)

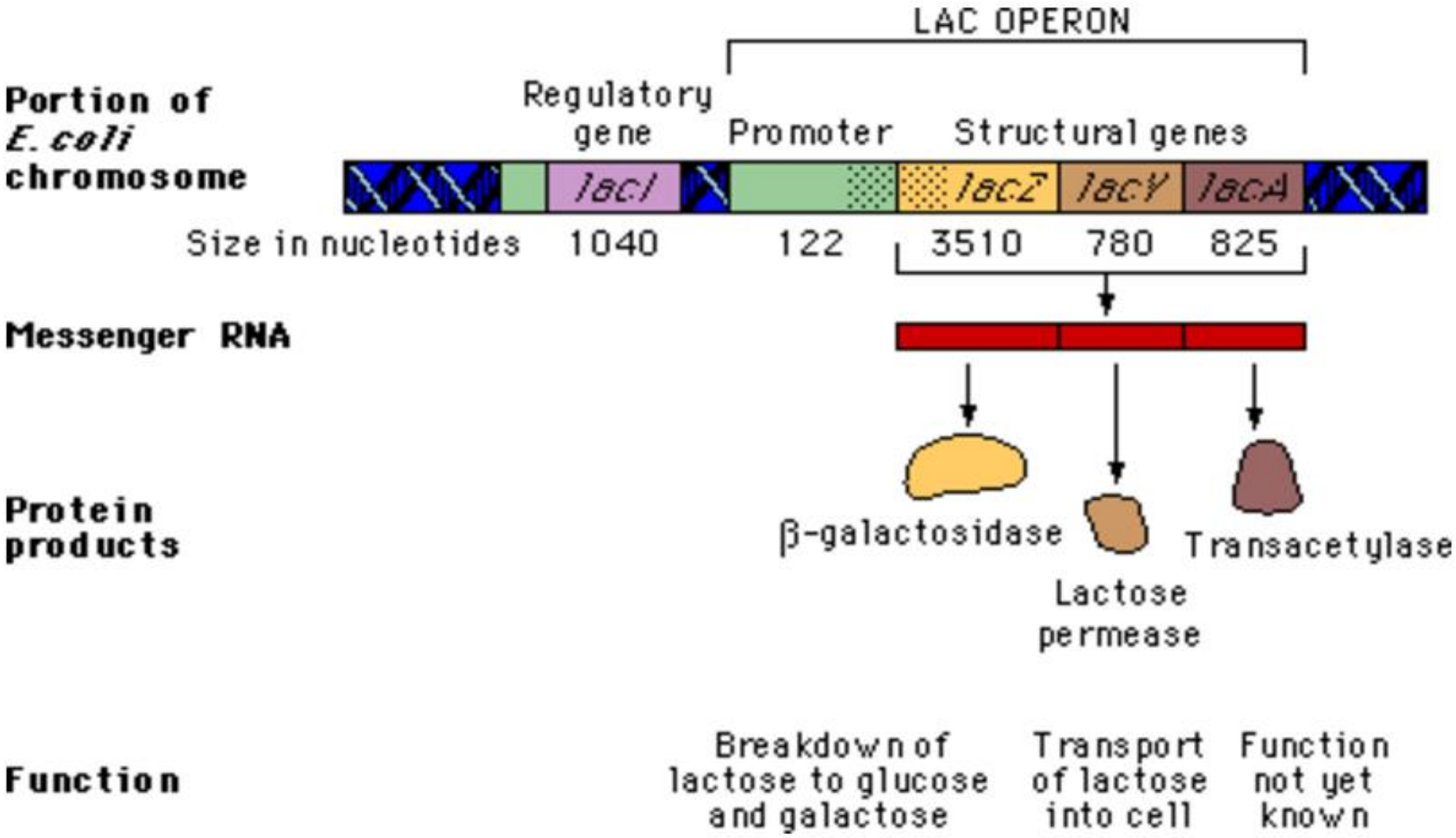
Dideoxyribonucleoside triphosphate (ddNTP)

Mutation

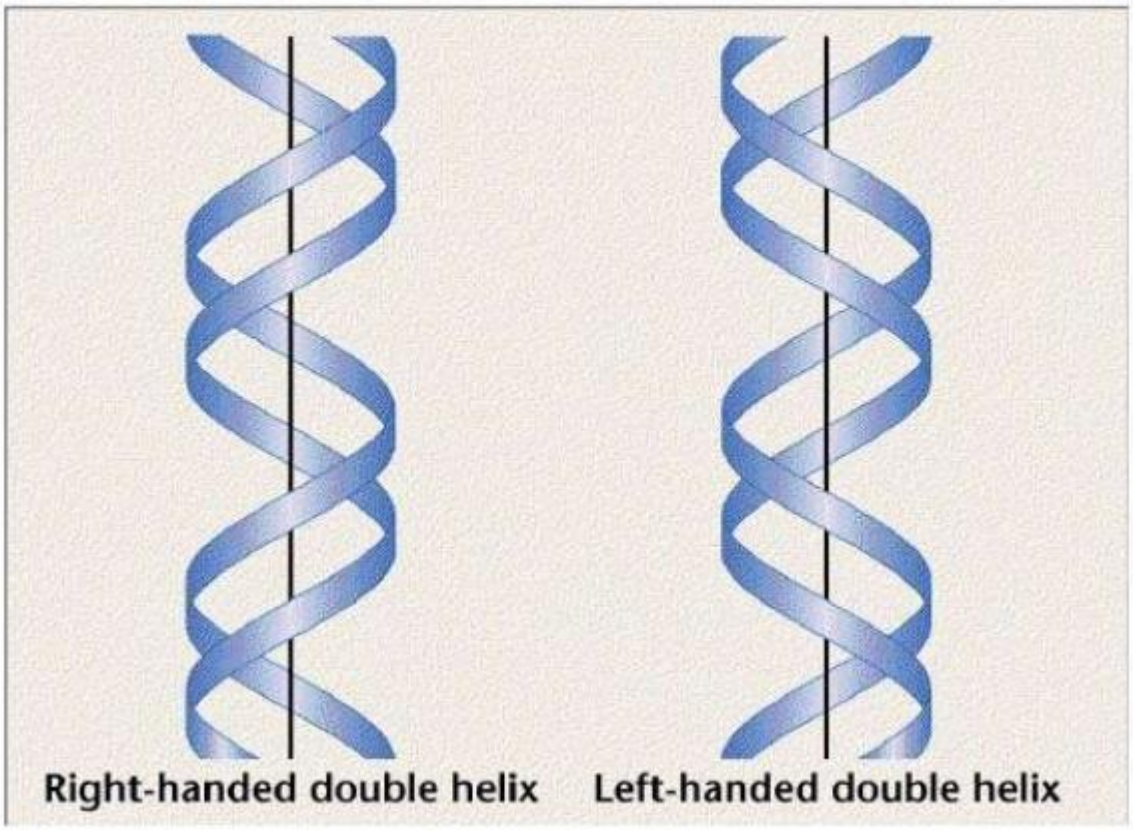

DNA Sequence
Amino Acid Sequence
Normal:
CAG
CCC
ACT
Codon 1
Codon 2
Codon 3
Gln
Pro
Thr
Insertion Mutation (Frameshift):
CAG
TCC
CAC
T
Codon 1
Codon 2
Codon 3
Codon 4
Gln
Ser
His
?
Insertion Mutation (Non-frameshift):
CAG
TTT
CCC
ACT
Codon 1
Codon 2
Codon 3
Codon 4
Gln
Phe
Pro
Thr

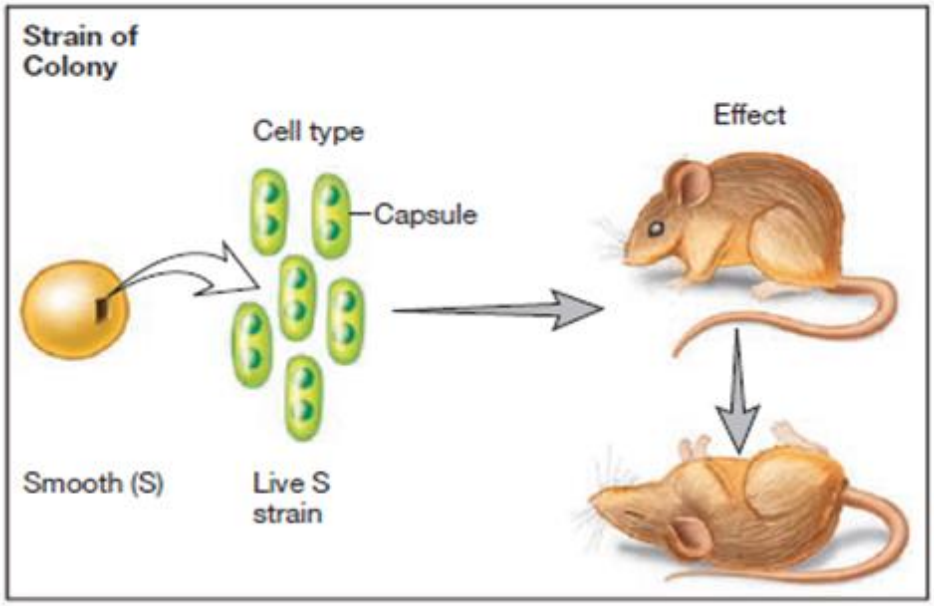

Amino acid attachment site
3′ OH
Acceptor stem
P 5′
T loop
D loop
Variable loop
Anticodon stem
Anticodon loop
Wobble position
Anticodon
tRNA^Ala

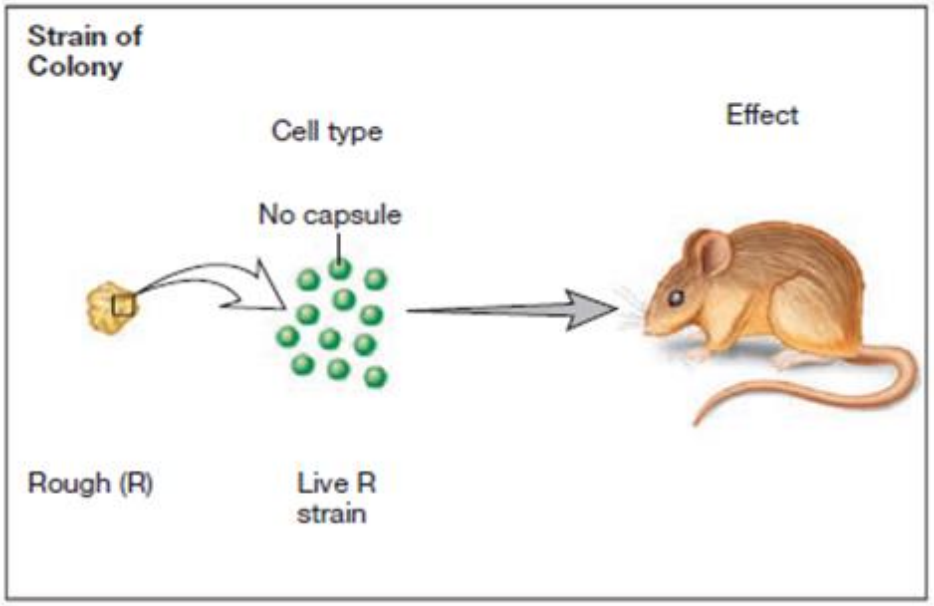
Lac operon

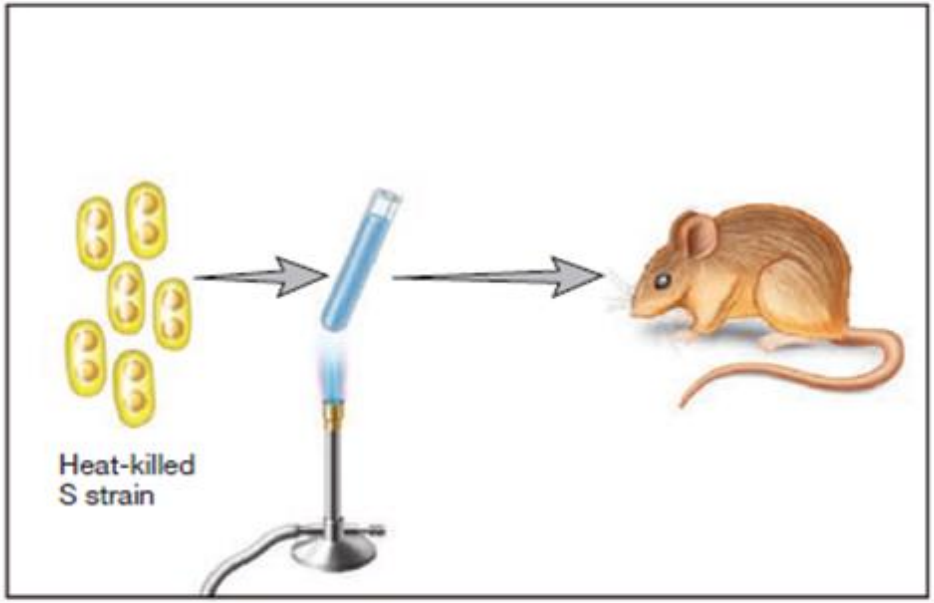

LAC OPERON
Portion of E. coli chromosome
Regulatory gene
Promoter
Structural genes
lacI
lacZ
lacY
lacA
Size in nucleotides
1040
122
3510
780
825
Messenger RNA
Protein products
β-galactosidase
Lactose permease
Transacetylase
Function
Breakdown of lactose to glucose and galactose
Transport of lactose into cell
Function not yet known

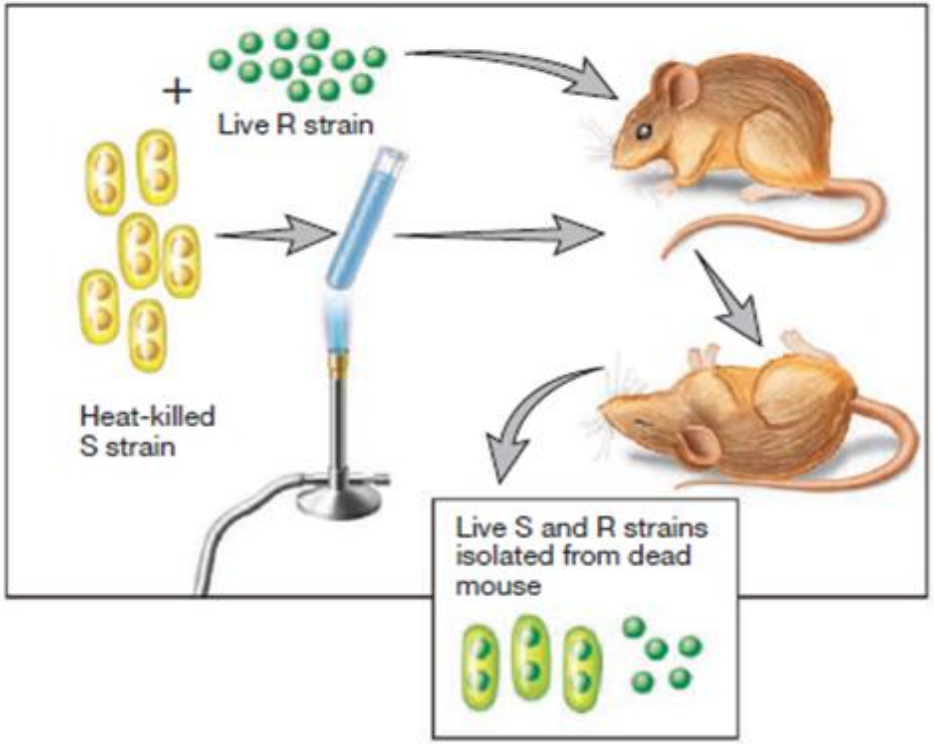
Helix Is Right-Handed




Strain of Colony
Cell type
Effect
Capsule
Smooth (S)
Live S strain


Strain of Colony
Cell type
Effect
No capsule
Rough (R)
Live R strain


Heat-killed S strain


+
Live R strain
Heat-killed S strain
Live S and R strains isolated from dead mouse