Molecular basis of inheritance

4.1 The Discovery of DNA:

1. Early Discovery

- **1869 – Friedrich Miescher**
    - Worked on **white blood cells (WBCs)** from pus (collected from hospital bandages).
    - Used a **salt solution** to wash pus from bandages.
    - Added a **weak alkaline solution** → cells lysed → **nuclei precipitated**.
    - From nuclei, he isolated a **new chemical substance** → named it **nuclein**.

2. Properties of Nuclein

- **High phosphorus content**.
- Showed **acidic properties**.
- Later called **nucleic acid**.

3. Later Understanding (Early 1900s)

- Miescher's **nuclein** was found to be a **mixture of proteins and nucleic acids**.
- Identified **two types of nucleic acids**:
    1. **DNA** – Deoxyribonucleic Acid
    2. **RNA** – Ribonucleic Acid

4.2 The Genetic Material is DNA

1. Early Understanding (1900s)

- **Genes** were known to:
    - Control the **inheritance of traits**.
    - Be **located on chromosomes**.

- **Chromosomes** were found to be composed mainly of:
    - **DNA**
    - **Proteins**

2. Initial Assumptions

- Most geneticists believed **proteins** were the **genetic material** because:
    - Proteins are **large, complex, and highly varied molecules**.
    - Their complexity seemed sufficient to **store genetic information**.
- DNA was considered **too simple** (only 4 nucleotides) to explain the vast **variations within species**.

Experimental Proofs that DNA is the Genetic Material:

1. Griffith's Experiment (1928) – *The Transforming Principle*

- Studied **Streptococcus pneumoniae** (bacteria causing pneumonia).
- Two strains: Bacteriophage
    - **S strain (Smooth, virulent)** → caused pneumonia, killed mice.
    - **R strain (Rough, non-virulent)** → harmless.

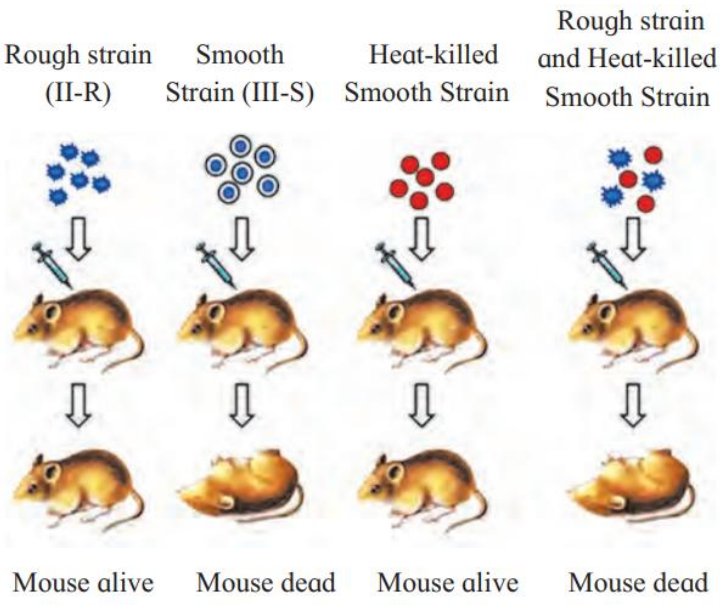


Fig. 4.1 : Griffith's experiment

Observations:

1. Live S strain → mice died.
2. Live R strain → mice lived.
3. Heat-killed S strain → mice lived.
4. Heat-killed S strain + live R strain → mice died.

Conclusion:

- Some **'transforming principle'** from the dead S strain converted the live R strain into the virulent S strain.
- The exact nature (DNA or protein) was unknown.

2. Avery, MacLeod and McCarty's Experiment (1944)

- Took **heat-killed S strain** extract and treated it with enzymes:
    - Protease → destroyed proteins (transformation still occurred).
    - RNase → destroyed RNA (transformation still occurred).
    - DNase → destroyed DNA (transformation did **not** occur).

Conclusion:

- DNA is the **'transforming principle'**.
- Provided strong evidence that **DNA is genetic material**.

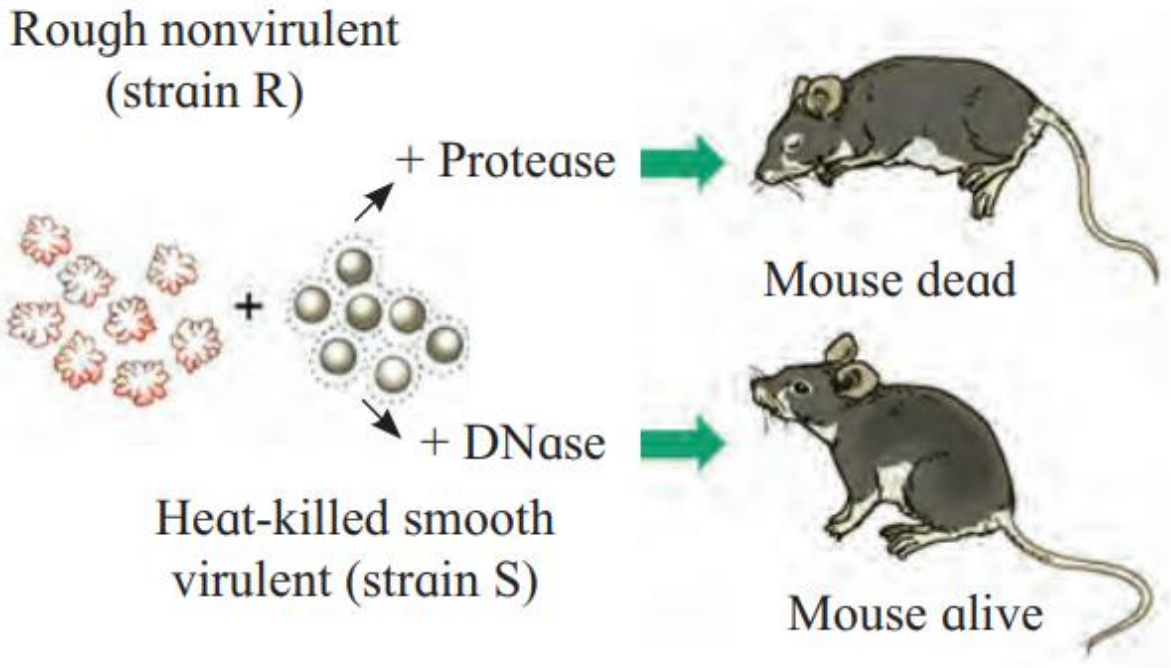
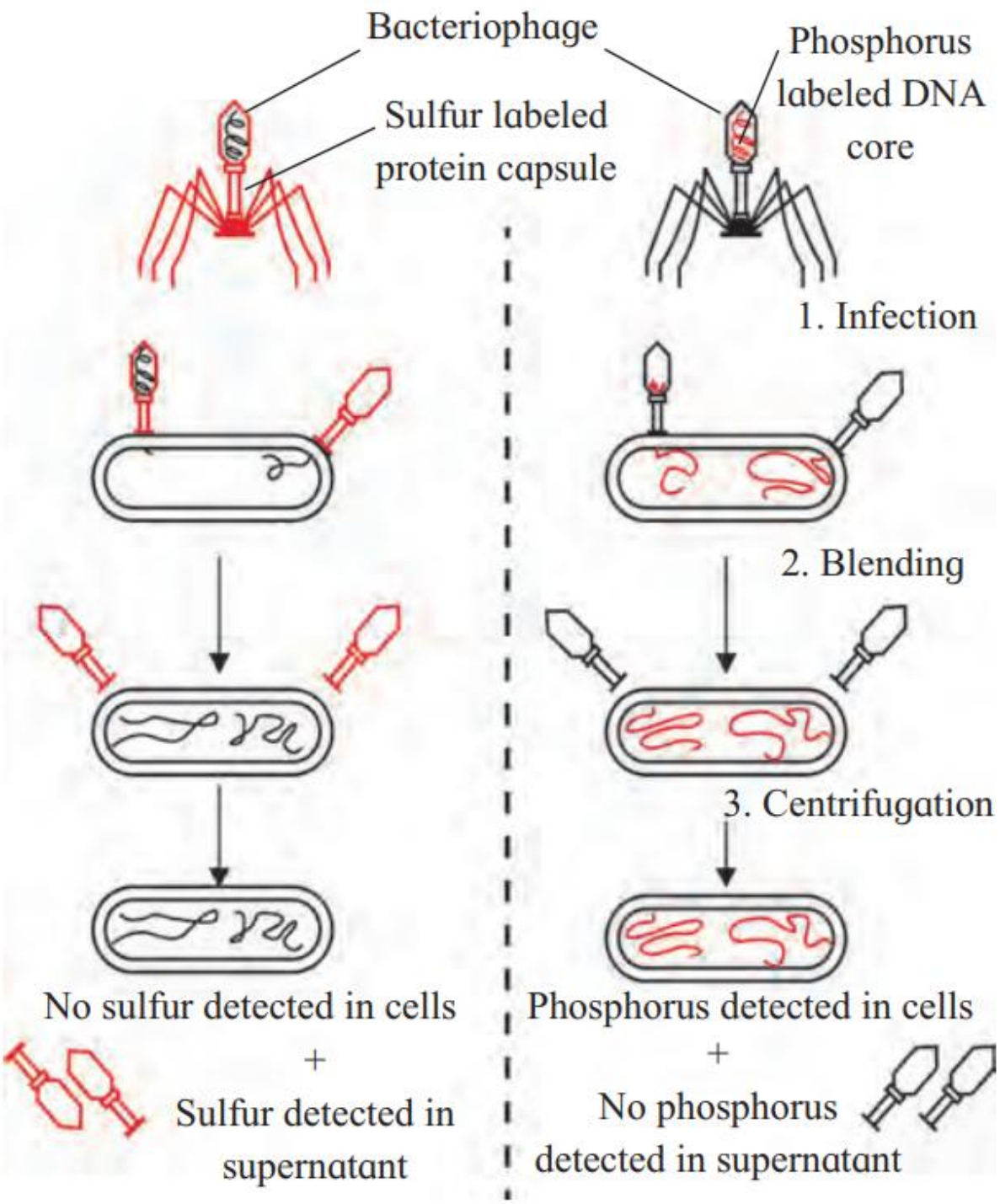


Fig. 4.2 : DNA transforms bacteria

3. Hershey–Chase Experiment (1952) – *The Blender Experiment*

- Worked with **bacteriophage (virus infecting bacteria)**.
- Labelled phages with:
  - **^{32}P** (radioactive phosphorus) → labels DNA.
  - **^{35}S** (radioactive sulfur) → labels protein coat.
- Allowed labelled phages to infect **E. coli**.
- After blending & centrifugation:
  - **^{32}P (DNA)** entered bacterial cells.
  - **^{35}S (protein)** remained outside.



Conclusion:

- DNA, not protein, enters bacteria and carries genetic information.
- Final proof that **DNA is the genetic material**.

Exam Tips:

- *Griffith* → discovery of **transforming principle**.
- *Avery, MacLeod & McCarty* → identified **DNA** as the transforming principle.
- *Hershey–Chase* → final proof that **DNA is genetic material**.

Table notes for DNA discovery experiments:

Scientist(s)	Year	Experiment / Material Used	Observation	Conclusion
Friedrich Miescher	1869	Collected **pus cells** (WBCs) from hospital bandages. Used **salt solution** (to wash pus) and **weak alkaline solution** (to lyse cells, precipitate nuclei).	Isolated a new chemical from nuclei → **nuclein** (acidic, high in phosphorus).	First discovery of **nucleic acid** (later known as DNA).
Frederick Griffith	1928	Studied **Streptococcus pneumoniae**: S strain (smooth, virulent) and R strain (rough, non-virulent). Injected into mice.	Heat-killed S strain + live R strain → mice died.	Proposed the presence of a **"transforming principle"** that converted R strain → S strain.
Avery, MacLeod & McCarty	1944	Treated heat-killed S strain extract with enzymes: **Protease**, **RNase**, **DNase**.	Transformation occurred with protease &	Identified **DNA** as the **transforming principle**.

			RNase, but not with DNase.	
Hershey–Chase	1952	Worked with **bacteriophage (virus)**. Labelled DNA with **^{32}P** and protein coat with **^{35}S**. Infected bacteria.	Only **^{32}P (DNA)** entered bacteria; **^{35}S (protein)** remained outside.	Confirmed that **DNA is the genetic material**.

4.3 DNA packaging:

- The length of DNA in a mammalian cell is about 2.2 meters. (multiplying the total number of base pairs by the distance between the consecutive base pairs, which is 0.34 nm).
- The approximate size of a typical nucleus is 10^{-6} m.
- Therefore, it must be condensed, coiled and supercoiled to fit inside such a small nucleus.

Packaging in Prokaryotes:

- In prokaryotes like E. coli, cell size is almost 2-3 µm.
- They do not have a well-organised nucleus.
- It is without a nuclear membrane and nucleolus.

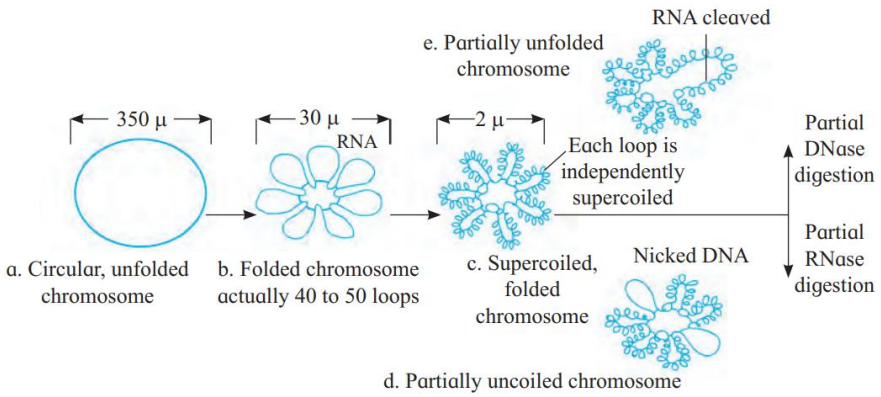


Fig. 4.4 : DNA Packaging in Prokaryotes

- The nucleoid is a small, circular, highly folded, naked ring of DNA which is 1100µm long in perimeter, containing about 4.6 million base pairs.
- The 1100P long (approximately 1.1 mm, if cut and stretched out) nucleoid is to be fitted or packaged into a cell which is hardly 2-3µm long.
- Hence the negatively charged DNA becomes circular,
- reducing the size to 350Pm in diameter.
- This is further reduced to 30Pm in diameter because of folding/ looping.
- 40-50 domains (loops) are formed. Formation of loops is assisted by RNA connectors.
- Each domain is further coiled and supercoiled, thereby reducing the size down to 2µm in diameter.
- This coiling (packaging) is assisted by positively charged HU (Histone like DNA binding proteins) proteins and enzymes like DNA gyrase and DNA topoisomerase I, for maintaining super coiled state.

Packaging in Eukaryotes:

Nucleus in Eukaryotes

- Well-organized **nucleus** with:
    - **Nuclear membrane**
    - **Nucleolus**
    - **Thread-like chromosomes**

Chromosome Structure

- DNA is associated with **histone and non-histone proteins**.
- Reported by **R. Kornberg (1974)**.
- Organization of DNA in eukaryotes is **highly complex**.

Protein Charges

- A protein's charge depends on the **abundance of amino acids with charged side chains**.
- **Histones**: proteins rich in **lysine and arginine**.

- Lysine & arginine = **basic amino acids**, carry **positive charges**.

Histones

- Positively charged, basic proteins.
- Occur along with protamines.
- Bind to negatively charged DNA (phosphate backbone) to form a stable structure.

Histone Octamer

- 8 histone molecules (H2A, H2B, H3, H4) arrange together.
- Form the core unit around which DNA is wrapped.
- This structure is the basis of **nucleosome organisation**.

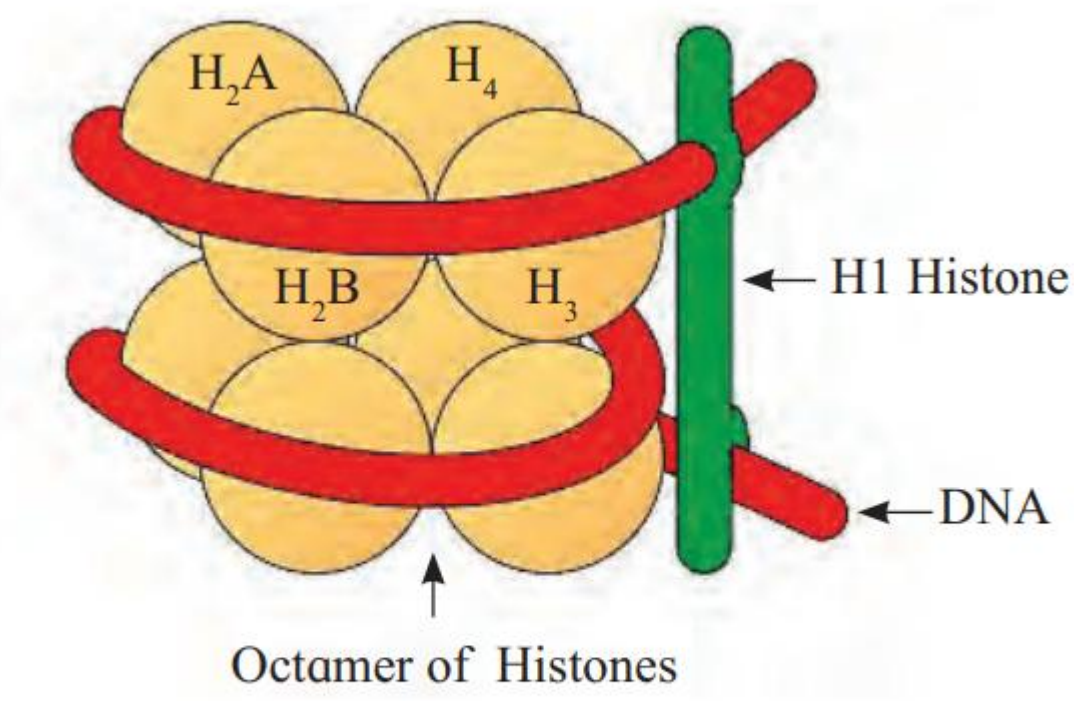


Fig. 4.5 : Nucleosome

1. Nucleosome Structure

- Negatively charged **DNA helix** wraps around positively charged **histone octamer**.
- **Histone octamer** = 8 molecules → 2 each of **H2A, H2B, H3, H4**.
- **H1 histone** → binds DNA at the entry and exit point of the nucleosome.

2. DNA Content in a Nucleosome

- Each nucleosome has about **200 base pairs (bp)** of DNA.
- **146 bp** wrapped around histone octamer (≈ 1¾ turns).
- Remaining DNA = **linker DNA** (joins adjacent nucleosomes).

3. Chromatin as Beads-on-a-String

- Nucleosomes are **repeating units of chromatin**.
- Appear like **"beads-on-string"** under an electron microscope.

4. Higher Order Packaging

- Six nucleosomes coil together → form a **solenoid structure**.
- Solenoid resembles a **coiled telephone wire**.
- Chromatin fibre thickness:
  - **10 nm fibre** (beads-on-string).
  - Coiled into a **30 nm solenoid fibre** (≈ 300 Å).
- Further **supercoiling** → looped structures → condense into **chromosomes** (during metaphase).

5. Role of Non-Histone Proteins

- **Non-Histone Chromosomal proteins (NHCs)** help in **higher-level DNA packaging** beyond the nucleosome level.

Summary for exam:

- **Nucleosome = DNA + histone octamer (H2A, H2B, H3, H4)**.
- **146 bp DNA** wraps around the histone core, with **H1** securing entry/exit.
- Packaging → Nucleosome → Solenoid (30 nm fibre) → Loops → Chromosome.

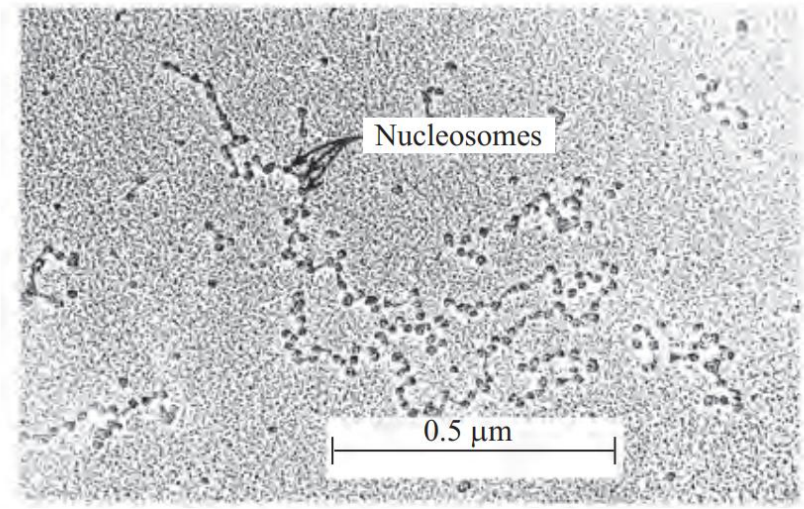
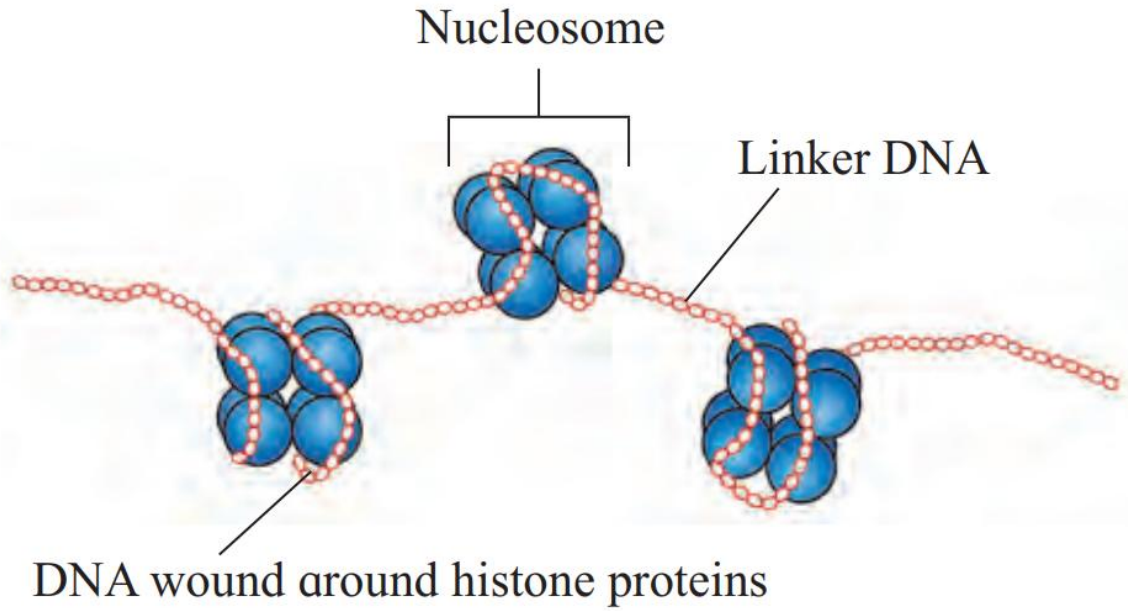
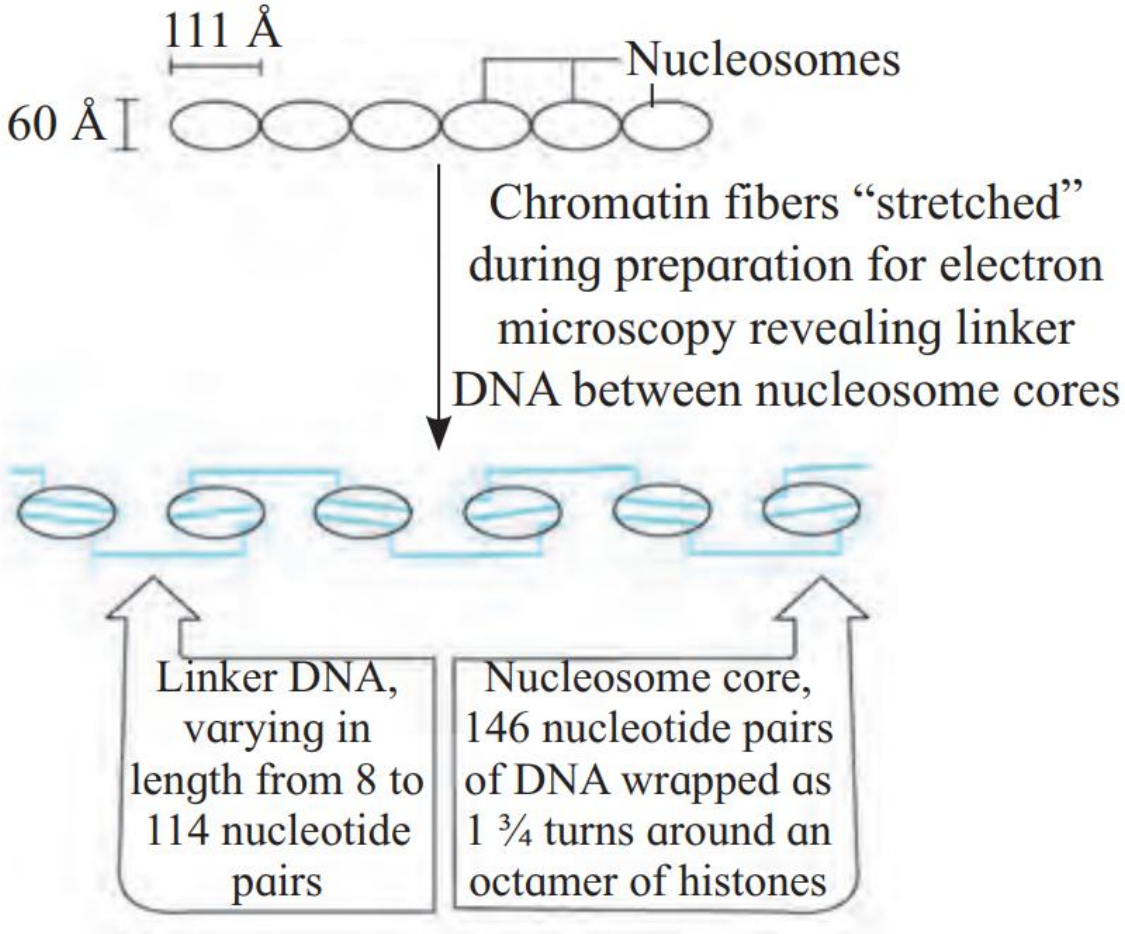


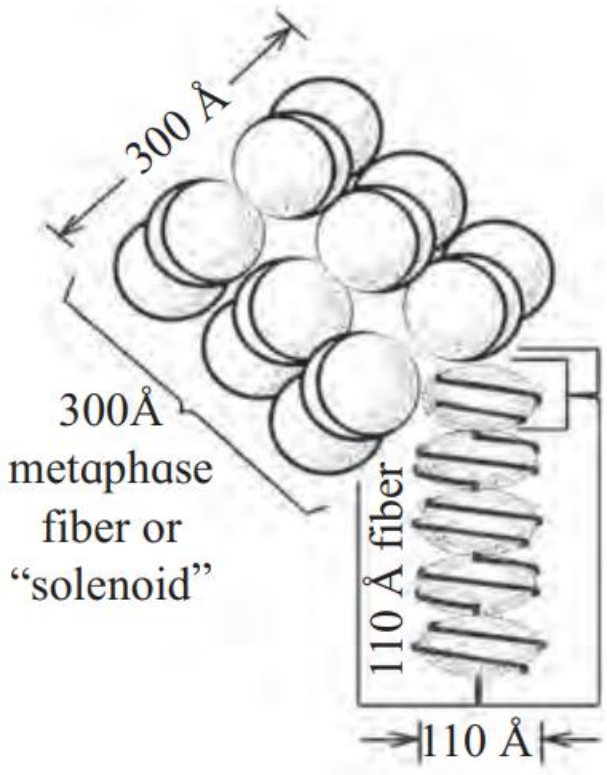
Fig. 4.6 : Chromatin showing beads-on-string



A : Beads-on-string magnified



B : Chain of nucleosomes forming 10 to 11 nm thick fibre



C : Solenoid forming 30 nm thick fibre

Fig. 4.7 : DNA packaging

Non-Histone Chromosomal Proteins (NHC): These are additional sets of proteins that contribute to the packaging of chromatin at a higher level.

Heterochromatin vs Euchromatin

1. Heterochromatin

- Proposed by **Heitz**.
- Regions of **chromonema/chromosome** that remain **condensed** during **interphase and early prophase**.
- **Location:** Common near **centromeres**, **telomeres**, and intercalated regions.
- **Properties:**
    - **Genetically inactive** (not expressed).

- **Transcriptionally almost inactive**.
- **Strongly stained**, appears **dark** under microscope.
- Contains **2–3 times more DNA** than euchromatin.

2. Euchromatin

- Regions of chromonema that are in **non-condensed (loosely packed)** state.
- **Properties:**
  - **Genetically active** → allows gene expression.
  - **Transcriptionally active** and **fast replicating**.
  - **Lightly stained**, appears **light** under microscope.

3. Key Differences (Summary Table)

Feature	Heterochromatin	Euchromatin
Packing	Condensed	Loosely packed
Location	Near centromere, telomeres	Throughout chromosome arms
Staining	Dark (strong staining)	Light (weak staining)
DNA content	Rich (2–3 times more than euchromatin)	Less
Genetic activity	Mostly inactive	Active

Exam Tip:

- Write "**Heterochromatin = condensed, dark, inactive**"
- Write "**Euchromatin = loose, light, active**"
  → Easy to score full marks.

4.4 DNA Replication :

Functions of DNA

1. Role of DNA

- DNA is the **genetic material** of the cell.
- Controls and regulates **all cellular activities**.
- Directs the synthesis of other biomolecules.
- Ensures **faithful transfer of genetic information** to daughter cells during cell division.

2. Major Functions of DNA

a. Heterocatalytic Function

- DNA directs the synthesis of **other molecules**.
- Examples:
    - **Transcription** → DNA synthesises RNA.
    - **Translation** → DNA indirectly directs protein synthesis (via mRNA).
- Essential for **gene expression** and cell functioning.

b. Autocatalytic Function

- DNA directs the synthesis of **itself**.
- Example: **Replication**.
- Ensures:
    - Formation of two identical DNA copies.
    - Equal distribution of DNA to **daughter cells** during cell division.

✅ **Exam Tip:**

- *Heterocatalytic = DNA → RNA → Protein*
- *Autocatalytic = DNA → DNA (Replication)*

Some important information:

1. **Timing of Replication**
    - DNA replicates **only once** in a cell cycle.
    - Occurs in the **S-phase (synthesis phase)** of **interphase**.
2. **Mode of Replication**
    - DNA follows **semiconservative replication**.
    - Proposed by **Watson and Crick** (1953).
3. **Basis of the Model**
    - DNA strands are **antiparallel** (run in opposite directions).
    - DNA strands are **complementary** (A–T, G–C base pairing).
4. **Semiconservative Concept**
    - During replication, the **two parental strands** separate.
    - Each serves as a **template** for the synthesis of a new complementary strand.
    - Result → Each daughter DNA molecule has:
        - **One old (parental) strand**
        - **One newly synthesised strand**

The process of semiconservative replication is as follows:

DNA Replication (7 Steps)

1. Activation of Nucleotides

- DNA has 4 types of nucleotides: **dAMP, dGMP, dCMP, dTMP**.
- These are activated by **ATP** with the help of the **phosphorylase enzyme**.
- Result → formation of **deoxyribonucleotide triphosphates**:
    - **dATP, dGTP, dCTP, dTTP**.

- Process = **Phosphorylation**.
- Activated nucleotides serve as **substrates** for DNA synthesis.

2. Point of Origin (Initiation Point)

- Replication begins at a specific site = **Origin (O)**.
- Terminates at **Termination site (T)**.
- **Replicon** = unit of DNA where replication occurs.
    - **Prokaryotes** → one replicon.
    - **Eukaryotes** → multiple replicons (in tandem).
- At the origin:
    - **Endonuclease enzyme** introduces a **nick** in one DNA strand.
    - Nick occurs in the **sugar–phosphate backbone (phosphodiester bond)**.

3. Unwinding of DNA Molecule

- **DNA helicase** → breaks weak **hydrogen bonds** at the origin.
- Strands separate and form a **Y-shaped replication fork**.
- **Unwinding is bidirectional**.
- Each separated strand acts as a **template**.
- **SSBP (Single-Strand Binding Proteins)** → bind to separated strands.
    - Prevent rejoining/recoiling.
    - Stabilize the unwound DNA for replication.

4. Replication Fork

- Formed due to **unwinding and separation** of parental DNA strands.
- Appears **Y-shaped**.

- Unwinding creates **strain** → relieved by **super-helix relaxing enzymes** (Topoisomerase).
-

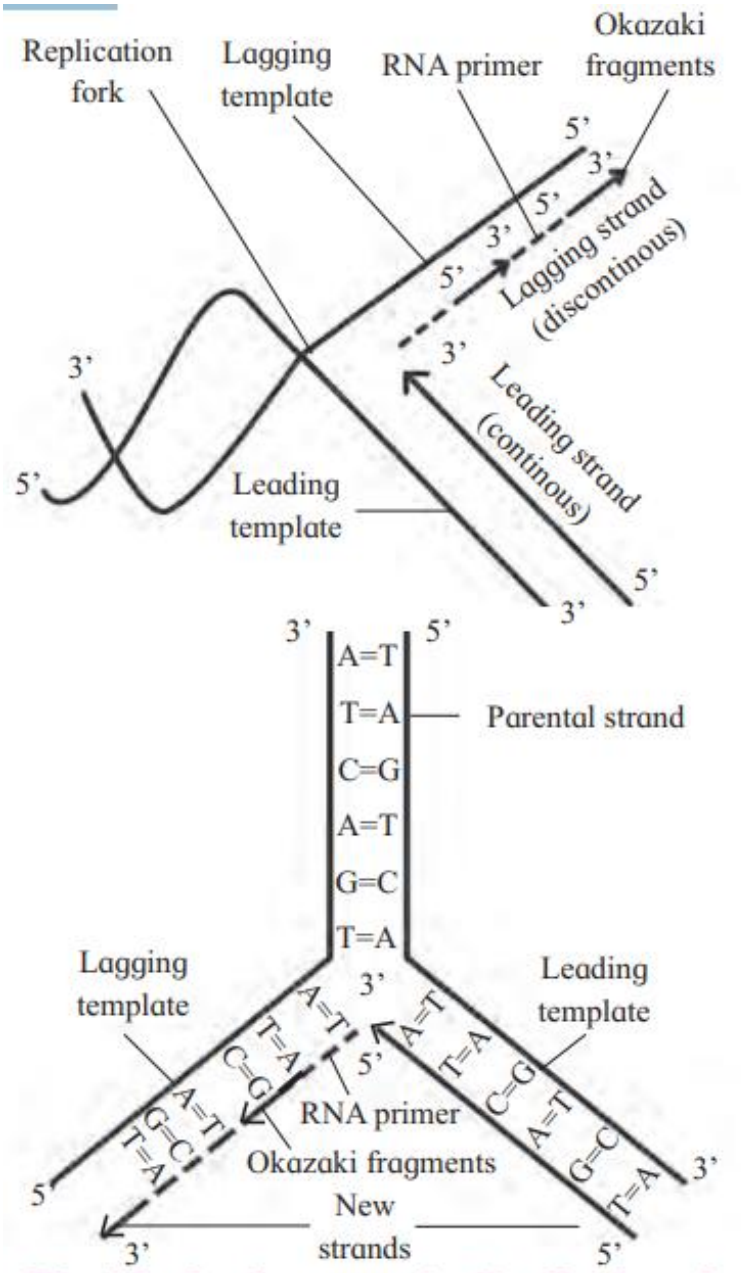


Fig. 4.8 : Semiconservative Replication of DNA

5. Synthesis of New Strands

- Each parental strand acts as a **template** (mould).

- Initiation requires a short **RNA primer**.
    - Primer attaches at the **3' end** of template strand.
    - Helps attract complementary **nucleotides** from nucleoplasm.
- Complementary base pairing occurs:
    - A ↔ T, G ↔ C.
- Nucleotides are joined by **phosphodiester bonds**.
- **Enzyme: DNA polymerase** catalyses the process.
- **Direction:** New strand synthesis always occurs in **5′ → 3′ direction**.

6. Leading and Lagging Strands

- Parental DNA strands are **antiparallel** → replication is different on each:
    - **Leading strand**: Synthesised **continuously** in the direction of the replication fork.
    - **Lagging strand**: Synthesised **discontinuously** in short fragments away from the fork.
- **Okazaki fragments**: Small fragments formed on the lagging strand.
- **Joining of fragments:**
    - Enzyme **DNA ligase** joins Okazaki fragments.
    - **RNA primers** are later removed and replaced with DNA:
        - By **DNA polymerase I** in prokaryotes.
        - By **DNA polymerase-α** in eukaryotes.
- **DNA gyrase (topoisomerase):** Rewinds and stabilises newly formed DNA double helix.

7. Formation of Daughter DNA Molecules

- At the end of replication → **two daughter DNA molecules** are formed.
- Each daughter DNA has:

- **One parental (old) strand**
- **One newly synthesized strand**

- Thus replication is **semiconservative** (50% parental contribution).

Experimental Confirmation of Semiconservative DNA Replication

1. Concept

- In **semiconservative replication**:
  - Each new DNA molecule has **one old (parental) strand + one new strand**.
- Experimentally proved by **Matthew Meselson & Franklin Stahl (1958)**.
- Used **equilibrium density gradient centrifugation** with **Cesium chloride ($CsCl_2$)**.

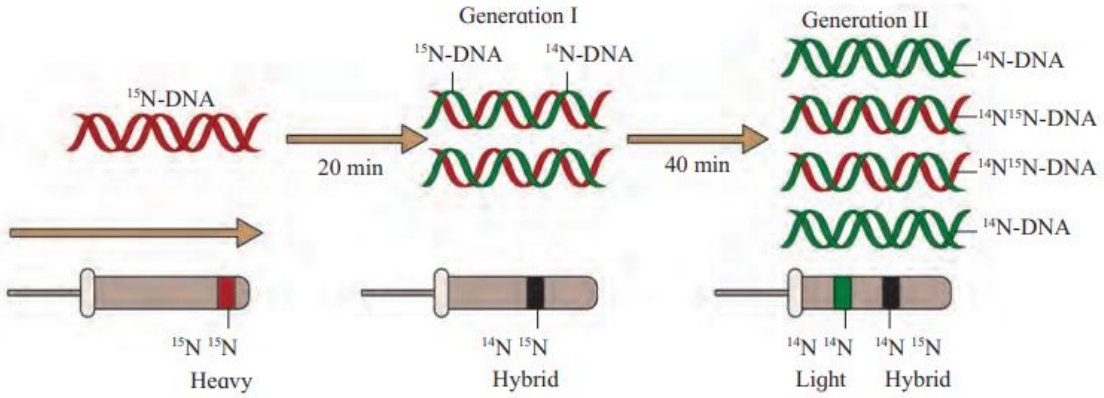


Fig. 4.9 : Interpretation of results of Meselson's experiment on the separation of DNA by equilibrium density gradient centrifugation

2. Experiment Steps

1. **Bacterial Growth in ^{15}N (Heavy Nitrogen):**
   - *E. coli* cells were grown in medium containing **^{15}N isotope**.
   - Their DNA became **heavy DNA** (both strands labelled with ^{15}N).
   - Centrifugation → formed a **single heavy band**.

2. **Transfer to ^{14}N (Light Nitrogen):**
    - Bacteria were then transferred to medium containing **^{14}N isotope**.
    - DNA synthesized after transfer contained **light nitrogen**.
3. **Results:**
    - **After 1 generation**:
        - DNA showed a **single intermediate band** (hybrid: one old ^{15}N strand + one new ^{14}N strand).
    - **After 2 generations**:
        - DNA showed **two bands**:
            - One **intermediate band (^{15}N–^{14}N hybrid)**.
            - One **light band (^{14}N–^{14}N)**.

3. Conclusion

- The appearance of **hybrid DNA after one generation** and **both hybrid + light DNA after two generations** proved that:
    - DNA replication is **semiconservative**.
    - Each daughter DNA receives **one old strand + one new strand**.

✅ **Exam Tip:**

- Name: **Meselson & Stahl (1958)**.
- Technique: **Density gradient centrifugation using $CsCl_2$**.
- Result: **Hybrid band (1st gen), Hybrid + Light bands (2nd gen)** → confirms **semiconservative replication**.

4.5 Protein Synthesis

1. Importance of Proteins

- Proteins are essential **biomolecules**.
- Functions:
    - **Structural components** (e.g., cytoskeleton, membranes).
    - **Enzymes** (biocatalysts).
    - **Hormones** (regulators of physiology).
- Cells must constantly **synthesize new proteins** for growth and functioning.

2. Steps in Protein Synthesis

Protein synthesis occurs in **two major steps**:

a. Transcription

- Process of **copying genetic information** from one DNA strand (template strand) into **RNA**.
- Produces a **single-stranded RNA transcript** (mRNA).
- Base pairing rule:
    - A pairs with **U** (instead of T).
    - T → A, G → C, C → G.
- Location: **Nucleus** (in eukaryotes).

b. Translation

- Process where **mRNA directs protein synthesis**.
- Occurs on **ribosomes** in cytoplasm.
- mRNA sequence is read in **triplets (codons)** → specifies amino acids.
- tRNA brings amino acids to ribosome → polypeptide chain is formed.

3. Central Dogma of Molecular Biology

- Proposed by **F.H.C. Crick (1958)**.
- Describes the **flow of genetic information**:

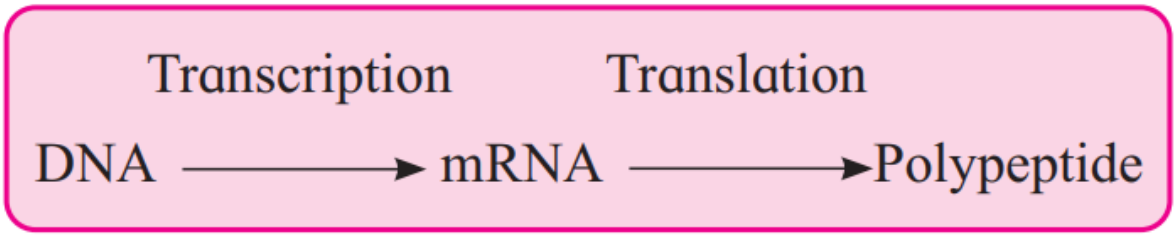
DNA ⟶ RNA ⟶ Protein

Meaning:

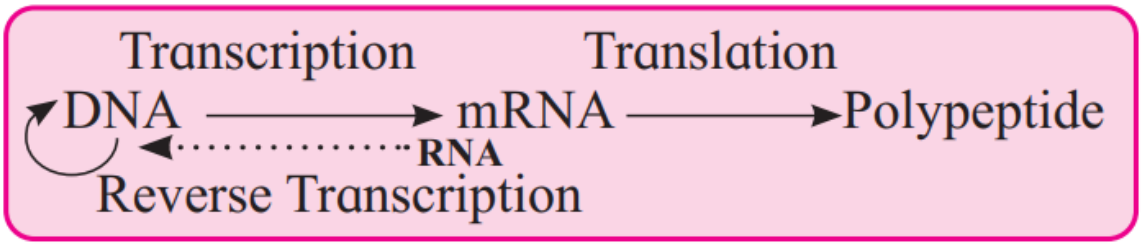
  - DNA → **mRNA** (via transcription).
  - mRNA → **Protein** (via translation).
- Unidirectional transfer of information (DNA → RNA → Protein).

✅ **Exam Tips:**

- Always mention **Crick (1958)** for Central Dogma.
- Use a simple **flow chart**: DNA → mRNA → Protein.
- Difference: *Transcription (copying message)* vs *Translation (building protein)*.



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



A. Transcription

1. Definition

- Process of **copying genetic information** from **one DNA strand** into **RNA transcript**.
- Enzyme: **RNA polymerase**.
- Product: **mRNA**, which carries genetic code from DNA to ribosome.

2. Location

- **Prokaryotes:** DNA in nucleoid → transcription in nucleoid.
- **Eukaryotes:** DNA in nucleus → transcription in **nucleus**.
- Translation occurs in **cytoplasm** (after mRNA leaves nucleus).

3. Timing (Cell Cycle)

- Transcription occurs during **G_1 and G_2 phases** of the cell cycle.

4. Key Elements of Transcription

- **Promoter**
  - DNA sequence at **5′ end (upstream)** of structural gene.
  - Provides a **binding site** for RNA polymerase.
  - In prokaryotes → recognized by the **sigma (σ) factor** of RNA polymerase.
- **Structural gene**
  - DNA sequence that codes for RNA.
  - Only **one strand (template strand)** is copied into RNA.
- **Terminator**

- DNA sequence that **signals the end** of transcription.

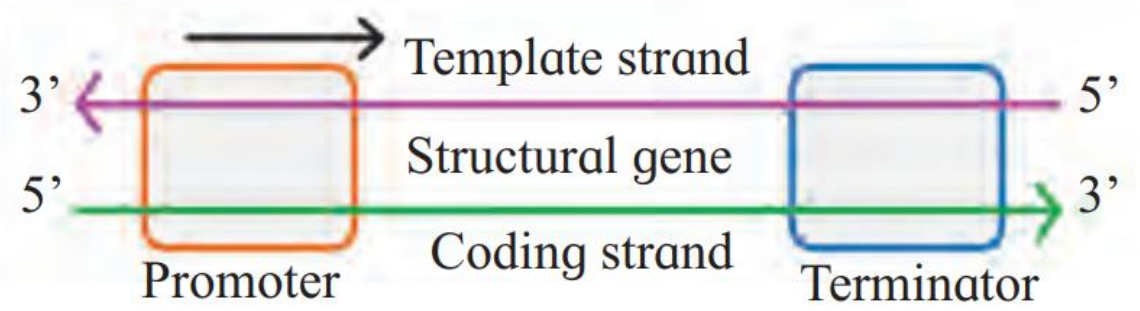


Fig. 4.10 : Transcription unit

6. Stages of Transcription

1. **Initiation**
   - RNA polymerase binds to the promoter region.
   - In prokaryotes, → σ-factor helps in promoter recognition.
2. **Elongation**
   - RNA polymerase moves along the template strand.
   - Adds ribonucleotides complementary to DNA bases (A–U, T–A, G–C, C–G).
   - RNA grows in **5′ → 3′ direction**.
3. **Termination**
   - When **RNA polymerase** reaches the **terminator signal** on DNA → it **detaches** from DNA.
   - The **newly formed mRNA (primary transcript)** is released.
   - The transcribed DNA region **rewinds** into the **double helix** again.

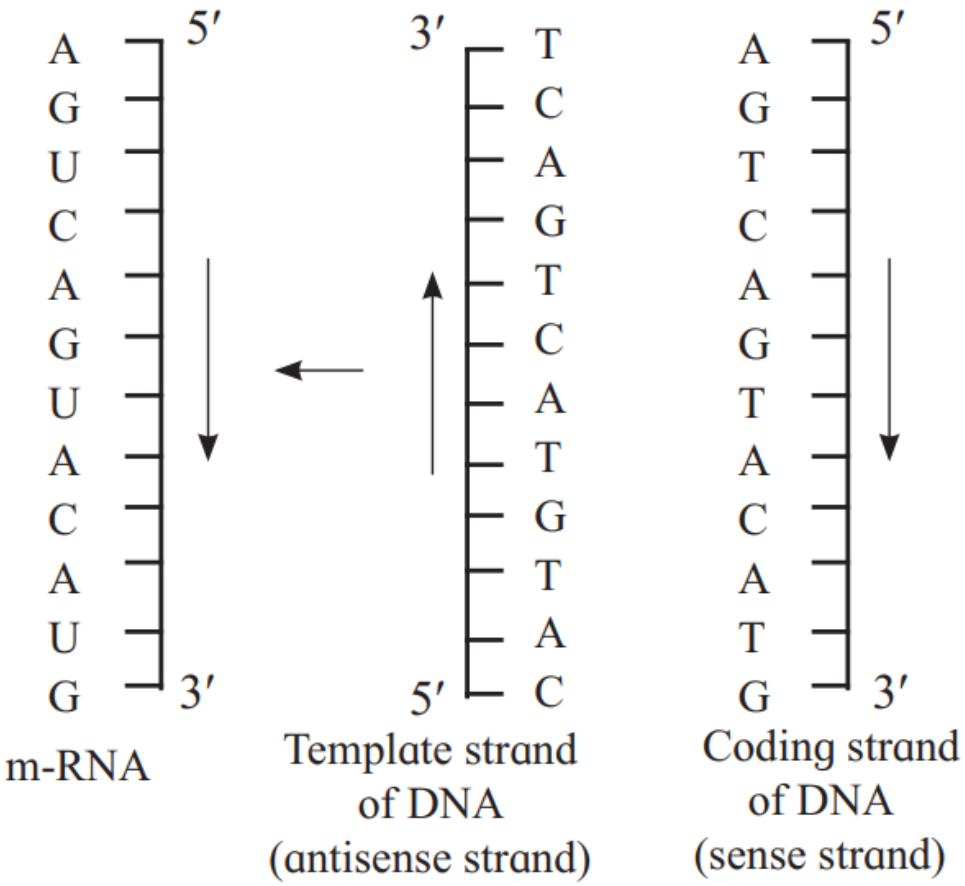


Fig. 4.11 : Formation of Template and Coding strand of DNA

✅ **Exam Tip:**

- Always mention **three stages (Initiation, Elongation, Termination)**.
- For prokaryotes → highlight **sigma factor** at promoter.
- For eukaryotes → transcription in nucleus, translation in cytoplasm.

Structural Genes and Transcription

1. Polarity of DNA Strands

- DNA strands are **antiparallel** → one runs **3′ → 5′**, other **5′ → 3′**.
- **RNA polymerase** synthesizes RNA in **5′ → 3′ direction**.

2. Template Strand (Antisense Strand)

- Strand with **3′ → 5′ polarity**.
- Serves as **template** for RNA synthesis.
- Complementary base pairing occurs (A–U, T–A, G–C, C–G).
- Directly guides the sequence of **mRNA**.

3. Coding Strand (Sense Strand)

- Strand with **5′ → 3′ polarity**.
- Its sequence is **identical to mRNA** (except thymine **T** is replaced by uracil **U**).
- Not used as template, but carries the same genetic "code" as the transcript.
- Called the **sense strand**.

4. Terminator

- Located at the **3′ end of coding strand (downstream)**.
- Defines the **end of transcription**.

5. Mechanism Summary

- RNA polymerase binds promoter → **unwinds DNA locally**.
- **Template strand (antisense)** is read → ribonucleotides added.
- As RNA elongates, **DNA–RNA hybrid** forms temporarily.
- RNA transcript (mRNA) eventually **dissociates** and becomes free.

✅ **Exam Tip (easy marks):**

- **Template strand = antisense strand (3′ → 5′)**.
- **Coding strand = sense strand (5′ → 3′, same as mRNA sequence with U instead of T)**.

Termination of Transcription

RNA Polymerase Types

- **Prokaryotes:**
    - Only **one type of RNA polymerase** → synthesizes all types of RNA (mRNA, rRNA, tRNA).
- **Eukaryotes:**
    - **RNA polymerase I** → transcribes **rRNA**.
    - **RNA polymerase II** → transcribes **mRNA (primary transcript)** and **hnRNA** (heterogeneous nuclear RNA).
    - **RNA polymerase III** → transcribes **tRNA** and **snRNA** (small nuclear RNA).

Transcription Unit and Gene

- **Gene**: DNA sequence coding for mRNA, tRNA, or rRNA.
- **Cistron**: Segment of DNA coding for one **polypeptide**.
- **Monocistronic gene**:
    - Single structural gene in a transcription unit.
    - **Eukaryotes** usually have **monocistronic genes**.
- **Polycistronic gene**:
    - Multiple structural genes in one transcription unit.
    - Common in **prokaryotes**.

Structure of Eukaryotic Structural Genes

- Genes are **interrupted** with:

  - **Exons** = coding sequences (expressed).
  - **Introns** = non-coding sequences (intervening).
- Only **exons** appear in the final processed mRNA.

RNA Processing (in Eukaryotes)

- The initial RNA product = **primary transcript (hnRNA)**.
- Primary transcript is **non-functional** (contains both exons + introns).
- Processing / maturation includes:
  1. **Splicing** → Removal of **introns**; joining of **exons**.
  2. (Later steps, though not in your text but important for exams):
     - **Capping** at 5′ end.
     - **Polyadenylation** at 3′ end.
- Result = **functional mRNA**, ready for translation.

✅ **Exam Tip:**

- Prokaryotes → mRNA **does not require processing** (no introns).
- Eukaryotes → mRNA requires **processing (splicing, capping, poly-A tail)**.

RNA Processing in Eukaryotes

1. Splicing

- Primary transcript (**hnRNA**) contains both **introns (non-coding)** and **exons (coding)**.
- **Introns are removed** and **exons are joined** in a definite sequence.

- Enzyme involved: **Spliceosome complex** (not DNA ligase — correction for students: DNA ligase joins DNA fragments, splicing is carried out by spliceosomes).
- Result: continuous coding sequence.

2. Capping (5′ End Modification)

- At the **5′ end** of hnRNA → a **methylated guanosine triphosphate cap (m^7G cap)** is added.
- Functions:
  - Protects mRNA from degradation.
  - Helps ribosome recognize mRNA during translation.

3. Tailing (3′ End Modification)

- At the **3′ end** of hnRNA → a **poly-A tail** (sequence of adenine nucleotides) is added.
- Process = **Polyadenylation**.
- Functions:
  - Increases stability of mRNA.
  - Aids in transport and translation efficiency.

4. Formation of Mature mRNA

- After **splicing, capping, and tailing**, hnRNA becomes fully **processed mRNA**.
- Mature mRNA is then **transported out of the nucleus** through **nuclear pores** into the cytoplasm.
- Ready for **translation (protein synthesis)** at ribosomes.

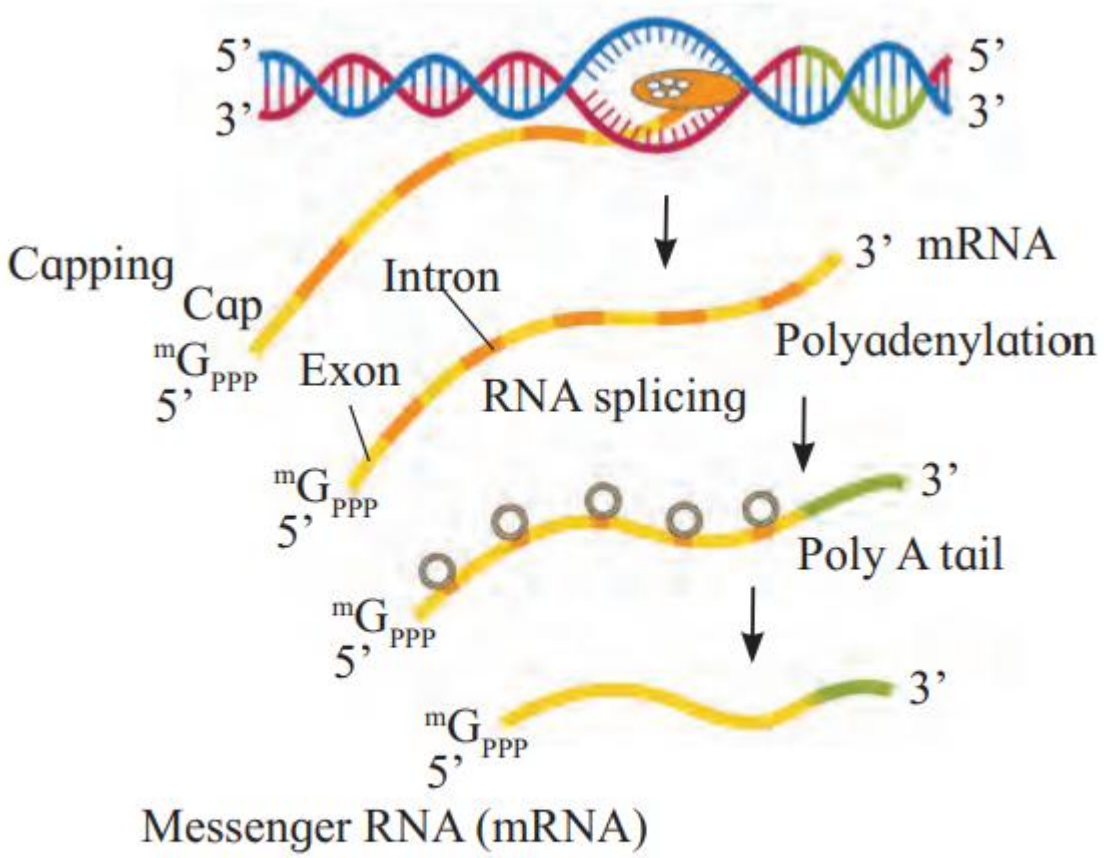


Fig. 4.12 : Transcription and Processing of hnRNA to mRNA in Eukaryotes

✅ **Exam Tip:**

- *Processing = Splicing (remove introns) + Capping (5′ m^7G cap) + Tailing (3′ poly-A tail).*
- Only **exons** appear in mature mRNA.

Genetic Code

1. Definition

- The **genetic code** is a set of rules by which the sequence of **nucleotides (bases) in DNA/mRNA** determines the sequence of **amino acids in proteins**.
- First evidenced by **Yanofsky & Sarabhai (1964)**.
- Concept of genetic code proposed by **F.H.C. Crick (1950s)** as a "cryptogram" of life.

2. The Problem

- **DNA bases = 4 types** (A, T, G, C).
- **Proteins = 20 amino acids**.
- Question: How can 4 bases specify 20 amino acids?

3. Codon Hypothesis

- **Singlet code (1 base per codon)** → 4 combinations (not enough).
- **Doublet code (2 bases per codon)** → 16 combinations (still less than 20).
- **Triplet code (3 bases per codon)** → 64 combinations (more than enough).
- Suggested by **George Gamow (1954)** → codon must be **triplet**.

4. Proof for Triplet Code

- **Crick (1961)** → Frame-shift mutation experiment showed **triplet nature** of codons.
- Total possible codons = **64** ($4^3 = 64$).
- Out of 64:
    - 61 codons code for amino acids.

- 3 codons act as stop signals (UAA, UAG, UGA).

5. Experimental Deciphering of the Code

a. Nirenberg & Matthaei (1961)

- Synthesized artificial RNA (poly-U = UUUUUU…).
- In cell-free system → produced a polypeptide of **phenylalanine**.
- Conclusion: **UUU = Phenylalanine**.
- First codon deciphered.

b. Har Gobind Khorana

- Developed techniques to synthesize RNA with **repeated nucleotide sequences**.
- Example: **CUC UCU CUC UCU** → alternating amino acids **Leucine and Serine**.
- **CUA CUA CUA…** → only **Leucine** produced.
- Proved codon sequence determines amino acid sequence.

c. Severo Ochoa

- Discovered enzyme **polynucleotide phosphorylase**.
- Helped synthesize RNA in a **template-independent manner**.
- Used to prepare RNA with defined sequences.

d. Final Contribution

- **Nirenberg, Matthaei, Khorana, and Ochoa** together **deciphered all 64 codons** of the genetic code.

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

Fig. 4.13 : Dictionary of genetic code

6. Summary

- **Genetic Code** = triplet code written in mRNA as **codons**.
- Provides universal language of life → DNA → RNA → Protein.
- Basis of **Central Dogma**.

Characteristic of the Genetic Code:

i. Genetic Code is a Triplet Code

- A **codon** = a sequence of **3 consecutive nitrogen bases**.
- Each codon specifies **one particular amino acid**.
- Codon sequence is always read in **5′ → 3′ direction**.
- Present in **all living organisms**.

ii. Genetic Code has Distinct Polarity

- Code has a **directionality**.
- Always read in **5′ → 3′ direction**.
- Reading in reverse would change the message.
    - Example: **5′ AUG 3′** = start codon (methionine).
    - If read 3′ GUA 5′ → codes for valine (different meaning).

iii. Genetic Code is Non-Overlapping

- Each base in mRNA is part of **only one codon**.
- Codons are **read one after another**, without overlapping.
- Example: 6 consecutive bases = **2 codons → 2 amino acids**.
- If overlapping, same 6 bases could give 4 codons → 4 amino acids.
- Experiments confirm code is **non-overlapping**.

iv. Genetic Code is Commaless

- Codons are **continuous**.
- There is **no gap or comma** between codons.
- Reading frame is maintained from the **start codon to stop codon**.

v. Genetic Code has Degeneracy

- Some amino acids are encoded by **more than one codon**.
    - Example: Cysteine → 2 codons, Isoleucine → 3 codons.
- This property is called **degeneracy of the code**.
- Explained by **Wobble Hypothesis**:
    - The first two bases in a codon are usually fixed.

- The **third base can vary**, still coding for the same amino acid.

vi. Genetic Code is Universal

- Same codon specifies the **same amino acid** in almost all organisms (from bacteria to humans).
- Example: **AUG = Methionine** in all living organisms.

vii. Genetic Code is non-ambiguous

- A specific codon codes for **only one amino acid**.
- Two different amino acids are **never coded by the same codon**.

viii. Initiation Codon and Termination Codons

- **Initiation codon:**
  - **AUG** → always the start codon.
  - Codes for **Methionine**.
- **Termination codons (Stop codons):**
  - **UAA, UAG, UGA**.
  - Do not code for any amino acid → signal end of polypeptide chain.

ix. Codon and Anticodon

- **Codon:**
  - Triplet of bases in **mRNA** (originally copied from DNA).
  - Written in **5′ → 3′ direction**.
  - Example: **5′ AUG 3′**.
- **Anticodon:**
  - Complementary triplet on **tRNA**.
  - Written in **3′ → 5′ direction**.

- Example: **3′ UAC 5′**.

✅ **Summary for Exams:**

- Genetic code is **triplet, universal, non-overlapping, commaless, degenerate, non-ambiguous**.
- Initiation codon = **AUG**, Stop codons = **UAA, UAG, UGA**.
- Codon (mRNA) pairs with Anticodon (tRNA) during translation.

Mutations and Genetic Code

1. Definition of Mutation

- **Mutation** = sudden **change in DNA sequence**.
- Leads to a change in **genotype**, which may be expressed in **phenotype** (character).
- Along with **recombination**, mutation is an important **raw material for evolution** because it generates variation.

2. Types of Mutations (based on DNA changes)

a. Chromosomal Mutations

- Involve **loss (deletion)** or **gain (insertion/duplication)** of DNA segments.
- Cause alteration in chromosome structure.

b. Point Mutation

- Involves **change in a single base pair** of DNA.
- Example: **Sickle cell anaemia** (substitution mutation in β-globin gene → Glutamic acid replaced by Valine).

c. Frame-Shift Mutations

- Caused by **insertion or deletion** of base pairs.

- **If 1 or 2 bases inserted/deleted** → reading frame shifts from that point onwards → produces a completely different amino acid sequence.
- **If 3 or multiples of 3 bases inserted/deleted** → adds or removes amino acid(s), but **reading frame remains unchanged**.

3. Relation of mutation to Genetic Code

- Genetic code is **read in triplets (codons)**.
- Mutations that alter the **triplet reading frame** (frameshift) → most drastic effects.
- Point mutations change only **one codon**, but can still alter protein structure/function.

✅ **Exam Tip:**

- *Point mutation → one base pair change (e.g., sickle cell anaemia).*
- *Frame-shift mutation → insertion/deletion of 1 or 2 bases (changes reading frame).*
- *Deletion/Insertion of 3 bases → no frame shift, only loss/gain of amino acid(s).*

tRNA – The Adapter Molecule

1. Concept

- Proposed because **amino acids cannot directly read codons**.
- **tRNA (transfer RNA)** acts as an **adapter**:
    - **Reads codons** on mRNA (via anticodon).
    - **Carries specific amino acids** to ribosome.

2. Cloverleaf Structure (2D Model)

- **Anticodon Loop**:
    - Contains a triplet sequence (**anticodon**).
    - Complementary to codon on mRNA.
- **Amino Acid Acceptor End (3′ end):**
    - Always ends with **CCA bases** (unpaired).
    - Serves as the **amino acid binding site**.
- Other loops: D-loop, TΨC loop, variable loop → help in stability and recognition.

3. Specificity of tRNA

- Each amino acid has a **specific tRNA**.
- Example: **Initiator tRNA** is specific for **Methionine (AUG start codon)**.
- There are **no tRNAs for stop codons** (termination is handled by release factors).

4. 3D Structure

- Actual tRNA folds into an **inverted L-shape**.
- Structure ensures **anticodon loop** is positioned to pair with codon, while **amino acid end** binds to ribosome.

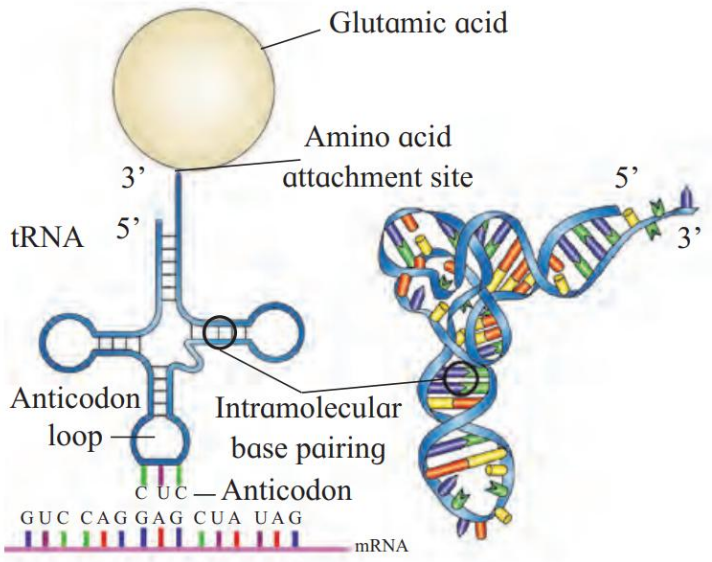


Fig. 4.14: t-RNA - the adapter molecule

✅ **Exam Tip:**

- *tRNA = Adapter molecule* (term by Crick).
- **Anticodon** pairs with codon (mRNA).
- **CCA at 3′ end** = amino acid attachment site.
- Special note: **No tRNA for stop codons**.

B. Translation – Protein Synthesis

1. Definition

- **Translation** = Process in which **codons of mRNA** are read and translated into a **specific sequence of amino acids** forming a polypeptide (protein).
- Occurs on **ribosomes**.
- Proteins are synthesized **intracellularly** within the cell.

2. Requirements for Translation

- **Amino acids** – raw materials (~20 types found in cytoplasm).
- **mRNA** – carries codon sequence from DNA.
- **tRNA** – adapter molecules carrying amino acids.

- **Ribosomes** – site of protein synthesis (large + small subunit).
- **ATP & Mg^{2+} ions** – provide energy and stabilize ribosome structure.
- **Enzymes & factors** – elongation factors, translocation factors, release factors.

3. Role of DNA and RNA

- **DNA** controls protein synthesis via transcription → produces **mRNA**.
- **Genetic code** specifies which amino acid corresponds to each codon.
- **RNAs** (mRNA, tRNA, rRNA) act as intermediates between DNA and protein.

4. Ribosome Structure and Function

- Ribosomes = **site of translation**.
- Made of **large subunit** and **small subunit**.
- Subunits are **separate in cytoplasm**, but join during protein synthesis (association requires **Mg^{2+} ions**).
- **Binding sites on ribosome for tRNA:**
    1. **A site** = Aminoacyl-tRNA site (entry site for new tRNA-amino acid).
    2. **P site** = Peptidyl-tRNA site (holds growing polypeptide chain).
    3. **E site** = Exit site (tRNA leaves ribosome).
- **Exception:** The first tRNA-amino acid complex enters directly into the **P site**, not the A site.

5. Special Note in Eukaryotes

- A **groove between ribosomal subunits** protects:
    - The **growing polypeptide chain** from cellular enzymes.
    - The **mRNA** from nucleases.

✅ **Exam Tip:**

- Translation requires: *mRNA, tRNA, ribosome, amino acids, ATP, Mg^{2+}*.

- Ribosome = *factory of protein synthesis*.
- Remember: **First tRNA enters P site** (all others enter A site).

Mechanism of Translation (Protein Synthesis)

Steps Involved

Translation = synthesis of **polypeptide chain** from mRNA.
It occurs in **three steps**:

1. **Initiation**
2. **Elongation**
3. **Termination**

1. Initiation of Polypeptide Chain

a. **Activation of amino acids**

- Amino acids are **activated using ATP**.
- Form aminoacyl-tRNA (tRNA bound to specific amino acid).

b. **mRNA binding**

- Small ribosomal subunit binds to **5′ end of mRNA**.
- **Initiator codon (AUG)** is recognized.

c. **Initiator tRNA binding**

- Initiator tRNA with **anticodon UAC** pairs with AUG codon.
- Carries:
    - **Methionine (Met)** in **eukaryotes**.
    - **Formyl methionine (fMet)** in **prokaryotes**.

d. **Assembly of ribosome**

- Large ribosomal subunit joins the small subunit.
- Association requires **Mg^{2+} ions**.
- Result: Initiator tRNA is placed in the **P-site**; **A-site is vacant**.

2. Elongation of Polypeptide Chain

Process of adding amino acids one by one to growing chain.

Cycle has 3 steps:

a. **Codon Recognition**

- New aminoacyl-tRNA enters **A-site** of ribosome.
- Anticodon pairs with codon by **hydrogen bonds**.

b. **Peptide Bond Formation**

- Amino acid at **P-site (methionine)** joins with amino acid at **A-site**.
- Enzyme: **Ribozyme (peptidyl transferase activity of rRNA)** catalyzes bond.
- tRNA at P-site is released.

c. **Translocation**

- Ribosome shifts along mRNA.
- tRNA carrying dipeptide moves from **A-site → P-site**.
- **A-site becomes vacant** → ready for next aminoacyl-tRNA.
- **Uncharged tRNA** is expelled from **E-site**.

This cycle repeats:

- Codon recognition → peptide bond formation → translocation.
- Process is very fast (~0.1 sec per peptide bond).
- Ribosome exposes codons one by one for translation.

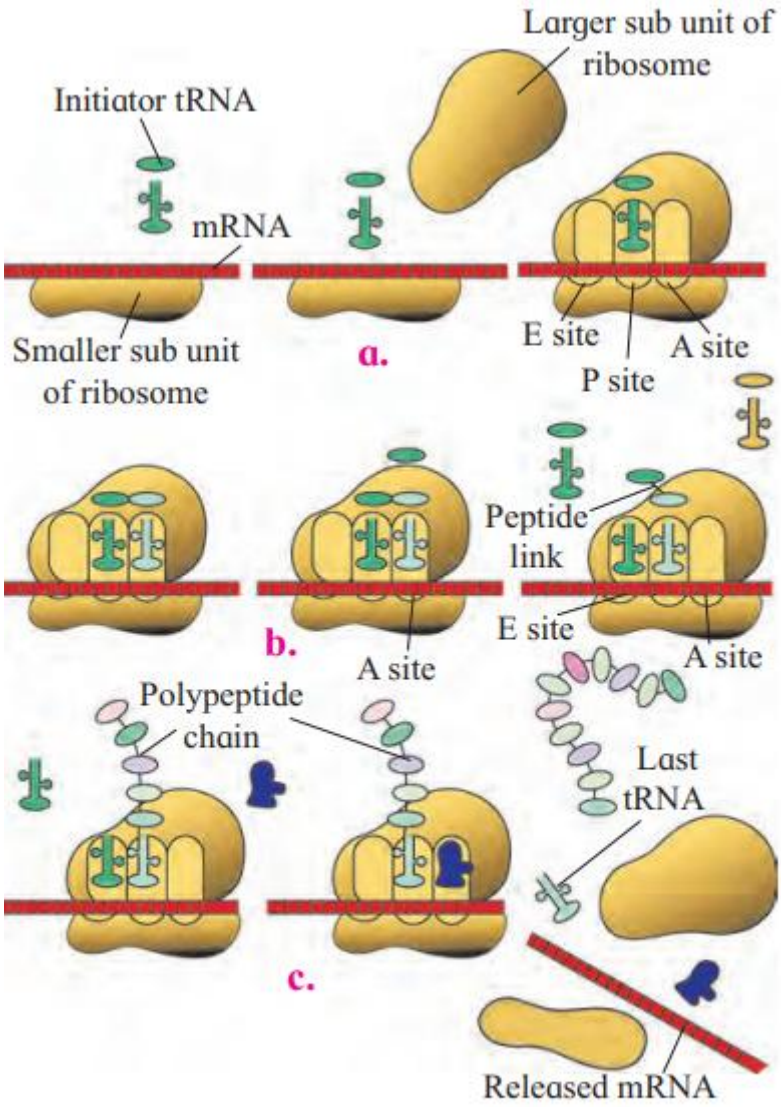


Fig. 4.15 : Translation Protein synthesis
a. Initiation, b. Elongation, c. Termination

✅ **Exam Tip:**

- Initiator codon = **AUG (methionine)**.
- First tRNA enters directly at **P-site**, others enter at **A-site**.
- Enzyme for peptide bond = **Ribozyme (23S rRNA in prokaryotes / 28S rRNA in eukaryotes)**.

3. Termination of Translation and Release of Polypeptide

1. Stop Codons

- At the end of mRNA, a **stop codon** appears at the **A-site**.

- Stop codons = **UAA, UAG, UGA**.
- These codons are **not recognized by any tRNA anticodon**.

2. Role of Release Factors

- Special proteins called **release factors** bind to stop codons.
- This action **terminates translation**.
- The completed **polypeptide chain** is released into the cytoplasm.

3. Dissociation of Complex

- After polypeptide release:
  - **Ribosomal subunits** (large + small) separate.
  - The **last tRNA** is set free into the cytoplasm.
  - The **mRNA** is released.

4. Untranslated Regions (UTRs)

- mRNA contains **non-coding sequences** at both ends:
  - **5′ UTR** (before start codon AUG).
  - **3′ UTR** (after stop codon).
- Functions: Help in **efficient initiation, regulation, and stability** of translation.

5. Fate of mRNA

- Once translation is complete, mRNA is **rapidly degraded by nucleases**.
- Hence, **mRNA is short-lived** in cytoplasm (especially in prokaryotes).

✅ **Exam Tip:**

- Stop codons = **UAA, UAG, UGA** (remember: U Are Annoying, U Go Away, U Are Gone).
- Release factors terminate translation.
- mRNA is **short-lived** (unstable).

4.6 Regulation of Gene Expression

1. Definition

- **Regulation of gene expression** = Multistep process controlling whether a gene is turned **ON or OFF**, and how much of its product (polypeptide/protein) is made.
- Ensures that genes are expressed only when needed.

2. Levels of Regulation in Eukaryotes

Gene expression can be regulated at several stages:

1. **Transcriptional level** → control of **primary transcript formation**.
2. **Processing level** → regulation of **splicing** of hnRNA into mature mRNA.
3. **Transport level** → control of mRNA movement from **nucleus → cytoplasm**.
4. **Translational level** → regulation of **protein synthesis on ribosomes**.

3. Example: E. coli and β-Galactosidase

- **β-Galactosidase enzyme** hydrolyses lactose:

$$\text{Lactose} + H_2O \xrightarrow{\beta\text{-galactosidase}} \text{Glucose} + \text{Galactose}$$

- If **lactose is absent** → β-galactosidase is **not synthesized**.
- If **lactose is present**, → enzyme is **synthesized**.
- Shows **environmental regulation** of gene expression.

4. Role in Development

- In higher organisms, **development & differentiation** (embryo → adult) occur through **coordinated regulation** of multiple gene sets.

5. Induction and Inducible Enzymes

- Some enzymes are **adaptive/inducible**, produced only when substrate is present.
- Mechanism:
    - A set of genes is **switched ON** only when required.
    - Small molecule (substrate) that triggers gene expression = **Inducer**.
    - Example: Lactose acts as an **inducer** in *E. coli*.
- This regulation = **Positive control**.

4.7 Operon Concept

1. Definition

- **Operon** = A transcriptional control mechanism in prokaryotes where a cluster of genes is regulated together as a unit.
- Proposed by **Francois Jacob and Jacques Monod (1961)**.
- Example: **Lac Operon** in *E. coli*.

2. Lac Operon in *E. coli*

- When **lactose** is available, *E. coli* induces the synthesis of enzymes needed to digest it.
- Enzymes:
    1. **β-galactosidase (Z gene)** → digests lactose → glucose + galactose.
    2. **Permease (Y gene)** → helps lactose enter the cell.
    3. **Transacetylase (A gene)** → transfers acetyl group from acetyl-CoA to galactoside.

3. Components of the Operon

- **Structural Genes (z, y, a)** → code for lactose-metabolizing enzymes.
- **Operator (O site)** → DNA sequence that controls whether structural genes are transcribed or not.
- **Regulator Gene (i gene)** → produces **repressor protein**.
- **Promoter (P site)** (not mentioned in your text but important for exams) → binding site for RNA polymerase.

4. Mechanism of Regulation

- **Repressor Protein (from i gene):**
    - Can bind to the **operator** and block RNA polymerase.
    - If repressor is bound → **switch OFF** (no transcription, no enzyme synthesis).
    - If repressor is removed → **switch ON** (RNA polymerase transcribes z, y, a genes).
- **Result:**
    - If ON → Structural genes transcribed into a single **polycistronic mRNA**.
    - Each gene functions as a **cistron** (unit coding for one polypeptide).

5. Key Terms

- **Polycistronic mRNA** → single mRNA carrying information for several genes (typical in prokaryotes).
- **Cistron** → segment of DNA coding for one polypeptide.
- **Switching ON/OFF** → determined by repressor binding to operator.

✅ **Exam Tip:**

- Always write: *Operon concept → Jacob & Monod, 1961 → Lac Operon example*.
- Structural genes = z, y, a.
- Regulation by **operator + repressor protein**.

Lac Operon – Inducible Operon

1. Definition

- **Lac operon** in *E. coli* is an **inducible operon**.
- Switched **ON** only when **lactose (inducer)** is present in the medium.
- Demonstrates **positive control of gene expression**.

2. Components of Lac Operon

1. **Regulator Gene (i gene)**
    - Controls operator gene.
    - Produces **repressor protein**.
    - Repressor binds to operator → **blocks transcription**.
    - In presence of lactose (inducer), repressor is inactivated → operator is turned **ON**.
2. **Promoter Gene (P)**
    - Located **next to operator gene**.
    - Site where **RNA polymerase** binds.
    - Base sequence of promoter decides which DNA strand acts as **template**.
3. **Operator Gene (O)**

- Lies **adjacent to structural genes**.
- Acts as **ON/OFF switch** for transcription.
- **ON** when repressor is inactive (inducer present).
- **OFF** when repressor is bound (inducer absent).

4. **Structural Genes** (z, y, a)
   - Code for enzymes required to metabolize lactose:
     - **lac-z** → β-galactosidase → hydrolyzes lactose → glucose + galactose.
     - **lac-y** → Permease → allows lactose to enter cell.
     - **lac-a** → Transacetylase → transfers acetyl group from acetyl-CoA to lactose derivatives.
   - Together, they produce a single **polycistronic mRNA**.
5. **Inducer (Lactose)**
   - Not part of operon.
   - Inactivates the **repressor protein** by binding to it.
   - Switches operon **ON**.

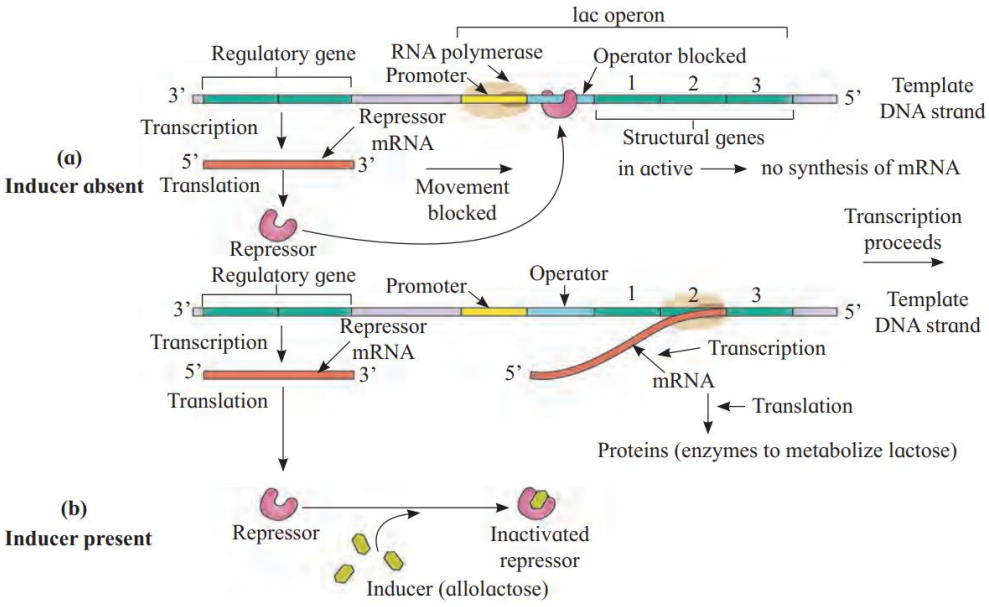


Fig. 4.16 : Working of Lac Operon

3. Mechanism (Induction)

- **No lactose present** → Repressor binds operator → **Operon OFF** (no enzyme synthesis).
- **Lactose present** → Lactose binds repressor → inactivates it → **Operon ON** → transcription of z, y, a genes → enzymes produced → lactose metabolized.

✅ **Exam Tip:**

- Always mention: *Jacob & Monod, 1961* → Lac Operon = **inducible system**.
- Components: Regulator (i), Promoter (p), Operator (o), Structural genes (z, y, a).
- Inducer = Lactose (switches ON the operon).

4.8 Genomics

1. Definition of Genome

- Term **Genome** introduced by **H. Winkler (1920)**.
- Genome = total genetic constitution of an organism.
- It represents a **complete set of DNA (haploid set of chromosomes)** of an organism.

2. Definition of Genomics

- Term **Genomics** coined by **T.H. Roderick (1986)**.
- **Genomics** = the study of genomes by:
    - **Mapping**, **sequencing**, and **analysis of genes**, and
    - Understanding their **functions and expression**.

3. Genome Sequencing Studies

- Sequencing projects started with model organisms:
    - **Yeast, Drosophila (fruit fly), Mouse**.
- Purpose: to facilitate **comparative studies** with humans.
- Many other genomes have been or are being sequenced:

- **Microbes, bee, tomato, and other crops**.

4. Types of Genomics

a. **Structural Genomics**

- Concerned with **mapping, sequencing, and analysis** of genomes.

b. **Functional Genomics**

- Studies the **functions of gene sequences**.
- Includes **gene expression** and regulation in organisms.

5. Applications of Genomics

- **Crop Improvement**
  - Development of **transgenic crops** with desirable traits (e.g., pest resistance, higher yield).
- **Human Health**
  - Treatment of **genetic disorders** using **gene therapy**.
- **Forensic Science**
  - Use of **genetic markers** for identification and forensic analysis.
- **Biotechnology**
  - Insertion of new genes in microbes to produce:
    - **Enzymes**
    - **Therapeutic proteins**
    - **Biofuels**
- **Social Sciences & Medicine**
  - Genomics knowledge applied to **disease prevention, diagnosis, and personalized medicine**.

4.9 Human Genome Project (HGP)

1. Introduction

- **Initiated:** 1990
- **Completed:** 2003
- **Administered by:** International body **HUGO (Human Genome Organization)**.
- **Coordinated by:** US Department of Energy (DOE) & National Institute of Health (NIH).
- **International Partners:** United Kingdom, France, Germany, Japan, India, China + US universities.
- **Nature:** Multinational project → to determine complete **genomic structure of humans**.

2. Main Aims of HGP

1. **Mapping the entire human genome** at nucleotide sequence level.
2. **Storing information** in databases.
3. Developing **tools and techniques** for data analysis.
4. **Technology transfer** to industries and private sectors.
5. Addressing **legal, ethical, and social issues (ELSI)** arising from the project.

3. Association with Bioinformatics

- HGP led to rapid growth of **Bioinformatics** → storage, analysis, and interpretation of genomic data.
- Helped scientists begin to understand the **blueprint of human genome**.

4. Achievements of HGP

- Produced a **complete and accurate sequence** of the **3 billion DNA base pairs** in humans.
- Estimated **~33,000 genes** in humans.
- Provided powerful data for research in **medicine, biotechnology, and life sciences**.

5. Comparative Genomics

- HGP also sequenced genomes of **model organisms** to study gene functions comparatively:
    - **E. coli** (bacterium)
    - **Caenorhabditis elegans** (nematode worm)
    - **Saccharomyces cerevisiae** (yeast)
    - **Drosophila melanogaster** (fruit fly)
    - **Arabidopsis thaliana** (model plant)
    - **Rice** (crop plant)
    - **Mus musculus** (mouse)

6. Importance of HGP

- Helped understand **gene functions and proteins**.
- Major applications in:
    - **Medicine** (diagnosis, personalized treatment, gene therapy).
    - **Biotechnology** (GM crops, microbes producing proteins).
    - **Life sciences** (comparative genomics, evolutionary studies).
- Set foundation for **genomic medicine and functional genomics**.

Table 4.17 : Comparative genome sizes of humans and other models organisms.

Organism	Chromosome number	Estimated gene number	Estimated size (base pairs)
Human (*Homo sapiens*)	46	33,000	3 billion
Mouse (*Mus musculus*)	40	25,000	2.9 billion
Fruit fly (*Drosophila melanogaster*)	8	13,000	165 million
Plant (*Arabidopsis thaliana*)	10	25,000	157 million
Roundworm (*Caenorhabditis elegans*)	12	19,000	97 million
Yeast *(Saccharomyces cerevisiae)*	32	6000	12 million
Bacteria (*Escherichia coli*)	1*	4400	4.6 million

Complexity of Human Genome

1. Key Idea

- Human complexity does **not depend only on the number of genes** (~33,000 genes).
- It depends on **how genes are regulated, expressed, and interact** with each other.

2. Insights from Comparative Genomics

- Many human genes are **similar to those in flies, roundworms, and mice**.
- This shows **evolutionary conservation of genes** across species.
- Studying gene structure and function in these **model organisms** helps in:
    - Understanding **human gene functions**.
    - Exploring **human evolution**.
    -

✅ **Exam Tip:**

- Always write **1990–2003, 3 billion base pairs, ~33,000 genes**.
- HUGO + NIH + DOE + International partners.
- Applications: Medicine, Biotechnology, Agriculture, Bioinformatics.

4.10 DNA Fingerprinting

1. Introduction

- Genes on chromosomes control characters & inheritance.
- Due to **recombination** (paternal + maternal genes) and **mutations**, every individual has a **unique genetic make-up**.
- This unique pattern = **DNA Fingerprint**.
- Technique to identify individuals using DNA restriction analysis is called **DNA Profiling/Fingerprinting**.
- Discovered by **Dr. Alec Jeffreys (1984, UK geneticist)**.

2. Principle

- Based on identification of **nucleotide sequence variations** in DNA.
- **99.9% DNA sequence is identical in all humans**.
- Differences lie in short repeating sequences of 20–100 base pairs called:
    - **Variable Number Tandem Repeats (VNTRs)**.
- Length and number of VNTRs differ in each individual → basis of DNA fingerprinting.

3. Steps in DNA Fingerprinting

1. **Isolation of DNA**
    - DNA is extracted from small samples (blood, hair roots, skin, semen, etc.).
2. **Restriction Digestion**
    - DNA cut into fragments using **restriction enzymes**.

- Creates different fragment lengths → **RFLP (Restriction Fragment Length Polymorphism)**.

3. **Gel Electrophoresis**
   - DNA fragments separated on **agarose gel** under electric field.
   - Fragments move according to **size** (shorter move faster).
   - dsDNA denatured into ssDNA by alkali treatment.

4. **Southern Blotting**
   - Developed by **E. Southern**.
   - DNA fragments transferred from gel → nylon/nitrocellulose membrane.

5. **DNA Probe Selection**
   - A known ssDNA sequence prepared = **DNA Probe**.
   - Labeled with **radioactive isotopes** for detection.

6. **Hybridization**
   - DNA probe binds (base pairs) to complementary sequences on host DNA.
   - DNA-DNA hybrids are formed; unbound probes washed away.

7. **Autoradiography/Photography**
   - Membrane exposed to **X-ray film**.
   - Bands appear → represent unique DNA pattern (DNA fingerprint).

Flowchart: DNA Fingerprinting Process

1️⃣ **Isolation of DNA**

⬇

2️⃣ **Restriction Digestion** → DNA cut into fragments by restriction enzymes (**RFLP**)

⬇

3️⃣ **Gel Electrophoresis** → Separation of DNA fragments by size

⬇

4️⃣ **Southern Blotting** → Transfer of fragments to nylon/nitrocellulose membrane

⬇

5️⃣ **DNA Probe Selection** → ssDNA probe (radioactive/fluorescent labeled) prepared

⬇

6️⃣ **Hybridization** → Probe binds to complementary DNA sequences on membrane

⬇

7️⃣ **Autoradiography (Photography)** → X-ray film developed → unique banding pattern = **DNA Fingerprint**

4. Applications of DNA Fingerprinting

1. **Forensic Science**
    - Solve criminal cases (murder, rape, terrorism).
2. **Parentage Disputes**
    - Identify biological father/mother of a child.
3. **Pedigree Analysis**
    - Used in animals (cats, dogs, horses) and humans.

- Helps in **breeding programs** and verifying lineage.

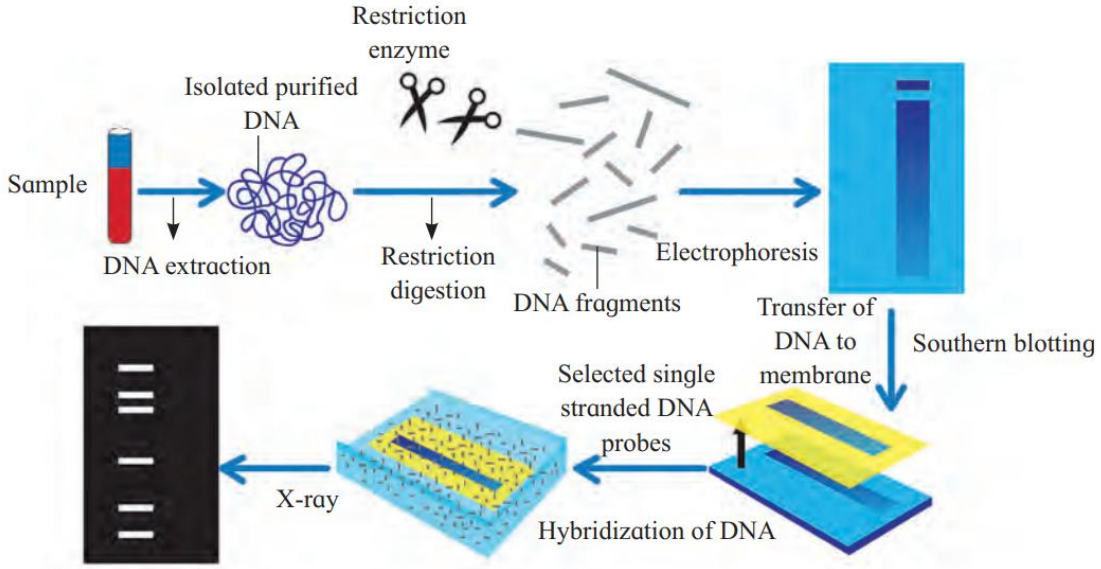


Fig. 4.18 : DNA Fingerprinting

Exam Tip:

- Inventor: **Alec Jeffreys (1984)**.
- Basis: **VNTRs** (unique repeating sequences).
- Remember 3 key applications: **Forensics, Parentage testing, Pedigree analysis**.

Know the scientists

Dr. Lalji Singh
(1947 - 2017)

Father of DNA Fingerprinting in India. He was instrumental in making DNA fingerprinting mainstream in India, for research and its forensic applications. He obtained DNA probe from Y chromosome of female banded krait snake (in this snake female has XY and male has YY chromosome). The unique segment obtained from this chromosome is, banded krait minor (BKM - DNA). It was used to developed probe for the Indigenous DNA fingerprinting technique.

Contributions of Dr. Lalji Singh: i. He installed several dedicated laboratories on aspects of genetics such as population biology, structural biology and transgenic research. ii. His work in the field of DNA fingerprinting technology, contributed for, wildlife conservation, forensics, evolution and phylogeny. iii. Established Centre for DNA Fingerprinting and Diagnostics (CDFD) in late 1990s- making it nodal centre for DNA fingerprinting and diagnostics for all species and several diseases. iv. Founded Laboratory for Conservation of Endangered Species (LaCONES).

********** **END OF CHAPTER** **********