Unit-III : Molecular Biology - I

Molecular biology is the study of biomolecules. OR Molecular biology is the study of biological phenomena at molecular level. In other words, it is the study of the molecular structure of DNA, the information it encondes, & the biochemical basis of gene expression & its regulation.

Nucleic acids :—

Nucleic acids were first discovered by Fredrich Miesches (1869) from nuclei of ~~first~~ pus cells. He named the substance as nuclein. Altman (1889) named it as nucleic acid. The nucleic acids are found in nucleus as well as cytoplasm.

Ther nucleic acids are of two types :—

(a) Deoxyribose nucleic acid (DNA), and

(b) Ribose nucleic acid (RNA).

★ Deoxyribose nucleic acid (DNA) :—

DNA is the genetic material of all the living organisms, except some viruses.

In eukaryotic cells, DNA is linear and occurs in nucleus & also in mitochondria & chloroplasts. In prokaryotic cells, it is circular & occurs in the cytoplasm.

"DNA is a polymer containing thousands of nucleotides linked through phosphodiester bonds".

Chemical composition of DNA :—

DNA is made of three types of chemical components. They are :—

① Deoxyribose sugars, ② phosphate groups and ③ Nitrogenous bases.

① Deoxyribose sugar :—

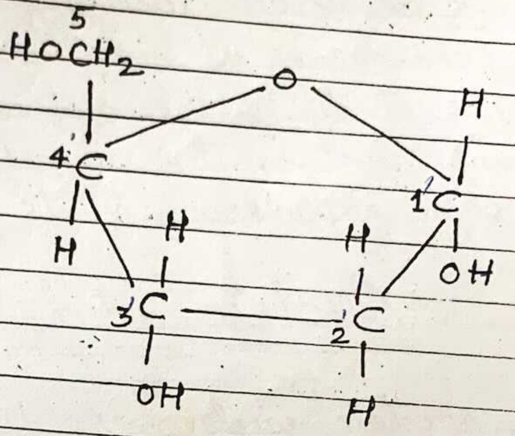
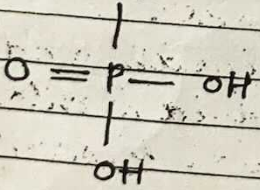


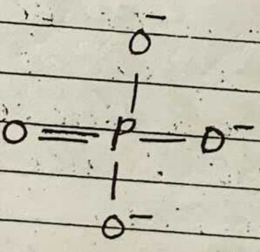
Fig. :- Deoxyribose sugar

It contains 5 carbon atoms & hence called pentose sugar. Four carbon atoms ~~are~~ & one oxygen atom form a ring. The fifth carbon atom is outside the ring and forms a part of $-CH_2$ group. The four carbon atoms of the ring are numbered as 1' (1-prime), 2', 3' and 4'. The carbon atom outside the ring is numbered as 5'. Deoxyribose sugar has 3 hydroxyl ($-OH$) groups in position 1', 3' and 5'. Hydrogen atoms are attached to carbon atoms 1', 2' & 4'. ~~and 5'~~. As the sugar contains one oxygen atom less at 2' carbon atom than ribose sugar, it is called deoxyribose.

② Phosphate groups :—



Phosphoric acid



Phosphate group

Phosphate groups are ionized derivatives of phosphoric acid (H_3PO_4) & ~~containing~~ contain central 'p' with 4 'O' atoms. One phosphate group is attached to 3' carbon of one deoxyribose sugar and 5' carbon of adjacent sugar. ~~It alternates~~ The phosphate groups alternate with deoxyribose sugar molecules. The acidic nature of ~~th~~ nucleic acid is due to phosphoric acid. Out of 3 $-OH$ groups, 2 are involved in formⁿ of sugar-phosphate backbone of each strand of DNA.

③ Nitrogenous bases :-

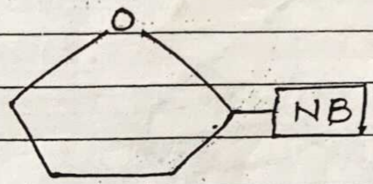
These are nitrogen containing organic heterocyclic compounds. They are of ~~Ay~~ 2 types, as follows :-

(a) Purines :- They are double-ringed (dicyclic) compounds. Two types - Adenine (A) & Guanine (G)

(b) Pyrimidines :- They are single-ringed (mono-cyclic) compounds. - Thymine (T) & cytosine (C).

Nucleosides :-

The pentose sugar combined with a nitrogenous base is called as nucleoside. Nitrogenous base is attached at 1' ~~end~~ carbon atom of pentose sugar.



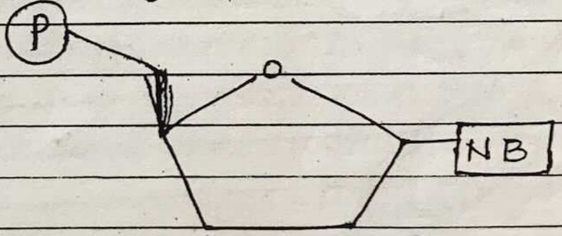
A nucleoside

Nucleoside = A pentose sugar + A nitrogenous base.

eg: Deoxyadenosine, Deoxyguanosine, Deoxythymidine & Deoxycytidine

Nucleotides :-

A nucleoside with a phosphate group is called nucleotide. It is the basic unit of DNA. Phosphate group is attached to 5' carbon of pentose sugar.



A nucleotide

Nucleotide = Nucleoside + phosphate group.

eg: Deoxyadenosine monophosphate
Deoxyguanosine monophosphate
Deoxythymidine monophosphate
Deoxycytidine monophosphate

Watson & Cricks model of DNA :—

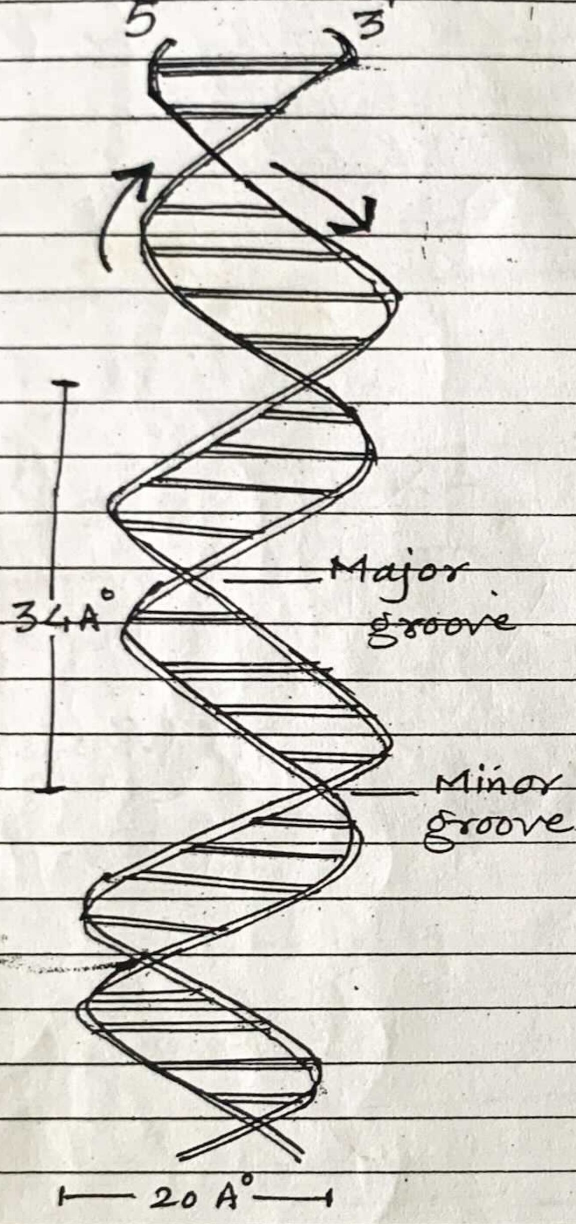


Fig.: DNA double helix

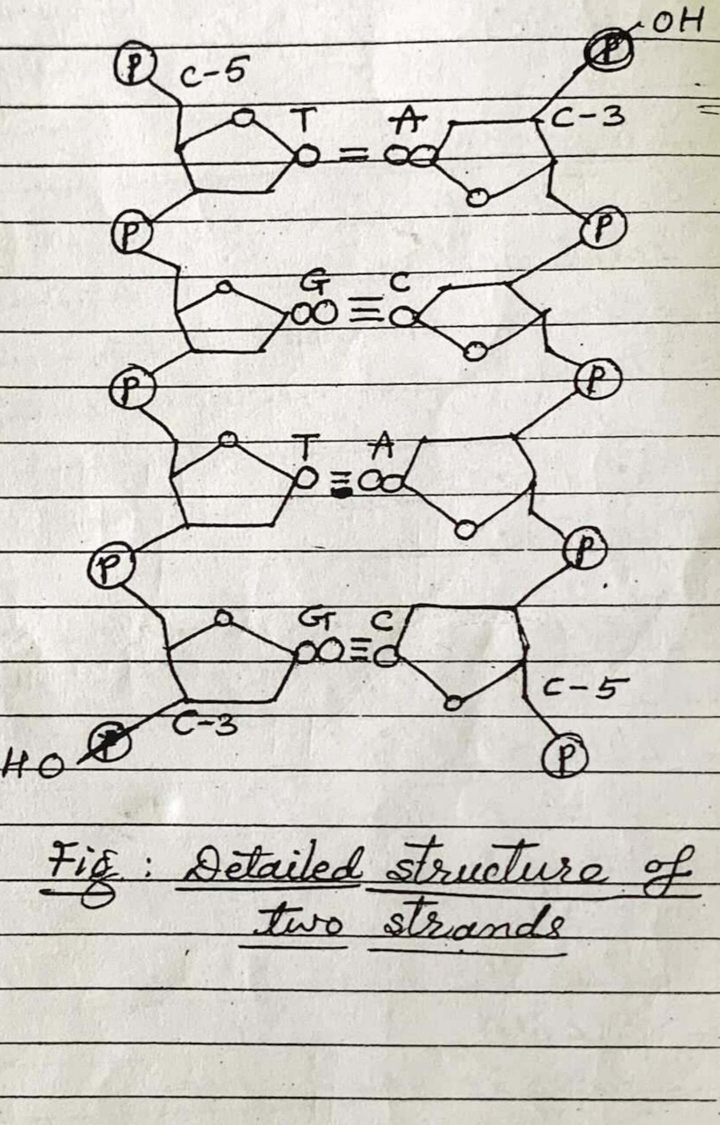


Fig : Detailed structure of two strands

James Watson & Francis Crick (1953) proposed the most accepted model of DNA structure. ~~They~~ Watson, Crick along with Wilkins were awarded Nobel prize in 1962, for Medicine & physiology.

In 1952, Wilkins obtained very fine X-ray diffraction photographs of DNA. The photographs showed that, the DNA molecule is double helical.

7 The helix is right-handed & turns clockwise. DNA molecule has a form of spiral ~~lad~~ ladder 10^{-8} cm = 10^{-10} m. Two strands are compared with vertical bars of the 10^{-9} ~~ladder~~ ladder. The paired nitrogenous bases linked by hydrogen bonds are compared with the steps of the ladder.

① **Double helix :-** (Right-handed)

According to Watson & Crick, the DNA molecule ~~cons~~ is double helical. It consists of two strands, which are spirally coiled around each other & a central imaginary axis to form a double helix. The two strands are joined together at regular intervals. ~~The helix has a diameter of 20 A°, and completes one turn its complete one turn is 34 A° long).~~

③ **3' & 5' ends** ~~[illegible]~~ :- The strands run in opposite directions. Such strands are called antiparallel. One strand runs in 5'–3' direction & other runs in ~~5~~ 3'–5' direction. One end of each strand has free –OH group at 3' carbon. This end is called 3' end. It is not linked to any nucleotide. The other end has one phosphate group at 5' carbon. This end is called 5' end. It is also not linked to any nucleotide.

④ **Grooves:** The two strands coil around each other. ~~& they~~ ~~a~~ The helix also ~~co~~ coils in a right-handed manner around imaginary central axis. This coiling results in formation of major ~~grooves~~ grooves & minor grooves.
Coiling of double helix – major / deep grooves.
Twisting of two strands around one another – minor or shallow groove

⑤ **Dimensions :-**

Diameter of DNA double helix is 20 A°. The length of one complete Turn / Spiral is 34 A°. One turn is distance between two successive major grooves or minor grooves. In one complete turn of DNA, 10 base pairs are present. The angle between successive base pairs is 36°. Hence, the angle of one turn is 360°.

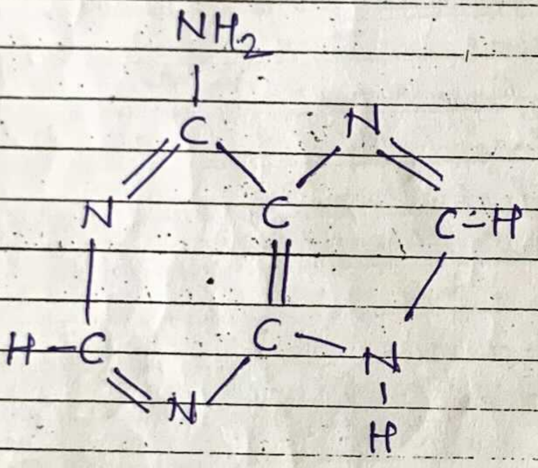
⑥ **Base pairing :-**

The nitrogenous bases of two strands exist in pairs:- One purine & other pyrimidine. Purine base Adenine (A) pairs with Thymine (T) of other strand. Whereas purine base Guanine (G) pairs with Cytosine (C) of other strand. A & T are joined by two hydrogen bonds, while G and C are joined by three hydrogen bonds. – ie. A = T, G ≡ C.

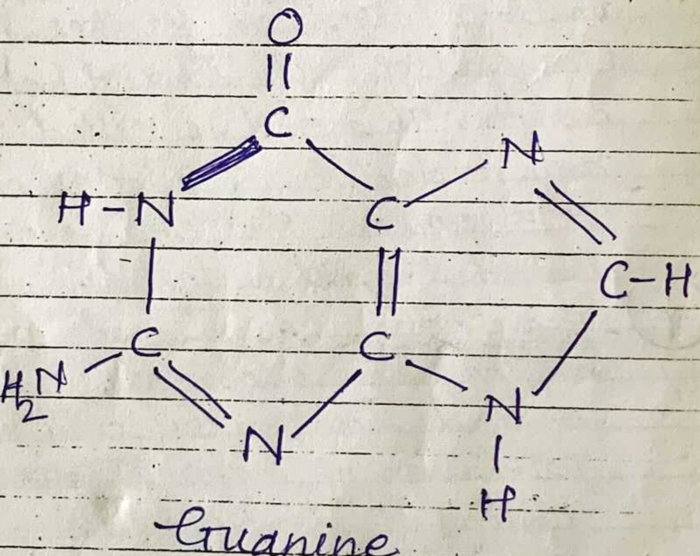
⑦ **Complementary strands** :- As the base sequences present on one strand of DNA decides the base sequence of other strand, ~~are~~ the strands are regarded as complementary strands. The nitrogenous bases along with hydrogen bonds make up the rungs (steps) of the ladder. These are called transverse bars. The distance betⁿ two successive steps is 3.4 Å.

~~hydrogen bonds, while G & C are joined by three hydrogen bonds. ie A = T, G ≡ C.~~

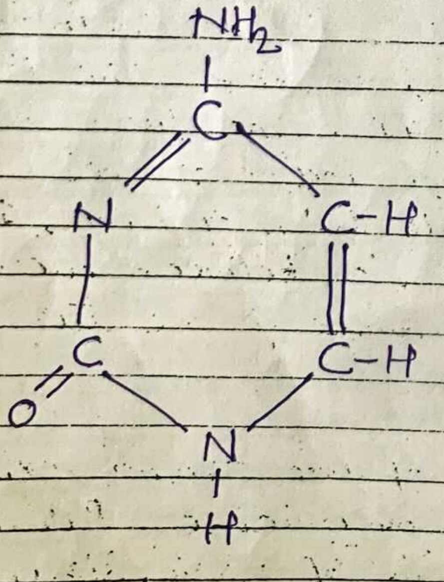
⑧ **Purine-pyrimidine ratio** :- Total No. of Adenine (A) is equal to Thymine(T) & total No. of Guanine (G) is equal to Cytosine (C). ie purine content is equal to pyrimidine content. Thus, purine-pyrimidine ratio is 1 : 1.



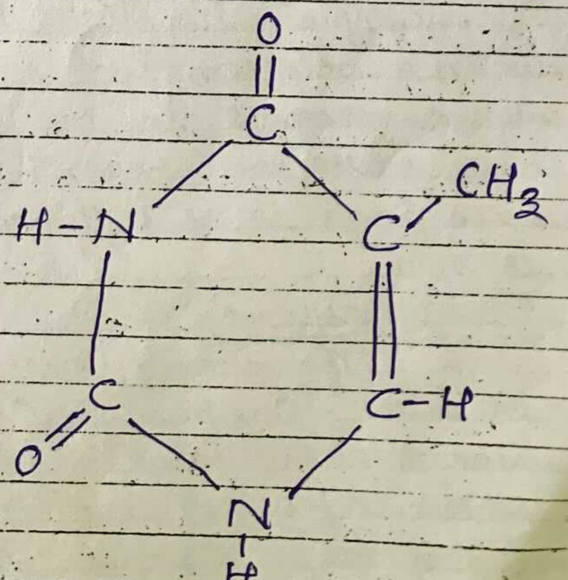
Adenine



Guanine



Cytosine



Thymine

Forms of DNA :-

The more refined X-ray diffraction studies ~~of~~ by Wilkins & his colleagues (1950) & some other techniques of 1970s have shown that, DNA fibres can have three possible structures — A, B and Z. It depends on base composition, salt content & relative humidity.

① **B-form** :- It is the structure proposed by Watson and Crick. It is obtained at high humidity ([illegible] (92%) ~~56%)~~ & in presence of excess salt. It consists of a right-handed antiparallel double helix. Purines & pyrimidines are roughly perpendicular to the axis of helix. One turn of helix consists of ten base pairs. Pitch of the helix is 34A°.

② **A-form** :— It is obtained at 75% relative humidity. It consists of a left-handed antiparallel double helix. One turn of helix consists of 11 base pairs. Pitch of the helix is 28.15A°.

③ **Z-form** :— ~~Z-DNA~~ :- ~~It is~~ The DNA duplex is formed by alternating right & left-handed helices arranged side by side. This structure is called right-left handed helix. One turn consists of 10 base pairs. The helix is arranged in zigzag manner & hence called Z- ~~DNA~~ form of DNA.

→ ~~St~~ Continued from page 6 - - - - -

② **Structure of each strand** :- Each strand is a polynucleotide chain ie formed of many nucleotides. The nucleotide is basic unit, formed by 3 components - deoxyribose sugar, a phosphate group & a nitrogenous base. The nucleotides are of 4 types - adenine nucleotide, guanine nucleotide, cytidine nucleotide & thymidine nucleotide. They are linked to each other by phosphodiester bonds to form polynucleotide strand. The phosphodiester bonds is formed betn 3' carbon of one deoxyribose sugar & 5' carbon of another deoxyribose sugar. The backbone of the strand is formed by alternately arranged deoxyribose & phosphate group. Each deoxyribose sugar has one nitrogenous base at 1' carbon. It may be A, T, G or C

⑨ Polarity :- Both the strands show polarity. O atom ~~of each polynucleotide~~ sugar in 1st strand is directed towards one end, whereas in other ~~sto~~ strand, it is directed to opposite pole

* Semi-conservative Replication of DNA :-

2M/4M. (Messelson and Stahl Experiment)

Messelson and stahl (1958) showed that, the double-stranded DNA suggested by Watson & Crick replicates by a simple method. During replication, the two strands of DNA molecule gets separated from each-other & each one acts as template for the synthesis of a complementary strand. When the replication is completed, two ~~identical~~ DNA molecules identical to each-other & identical to the original are produced. ~~ie~~ One half of each new molecule of DNA is old and one half is new. This mode of replication is called as semi-conservative mode. It means that, one strand comes from original helix & one is newly formed.

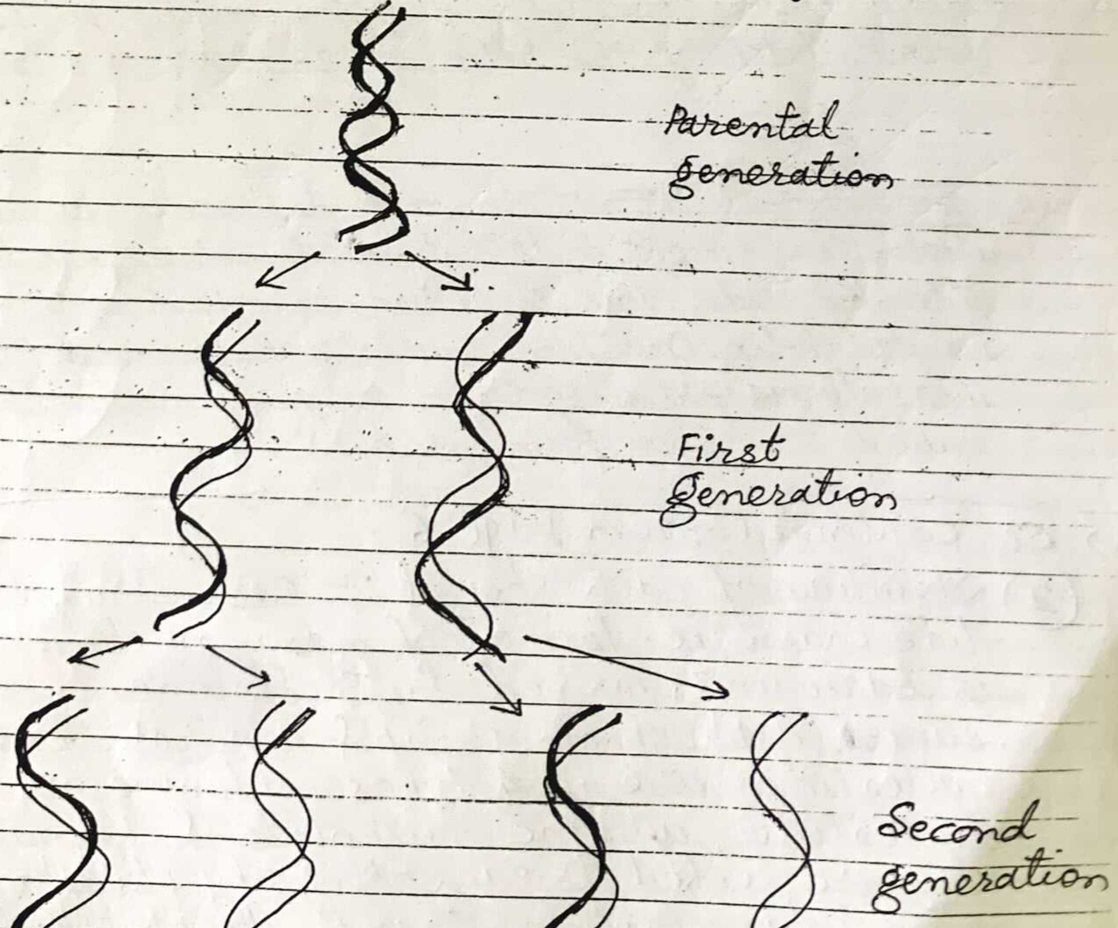


Fig: Messelson & Stahl expt.

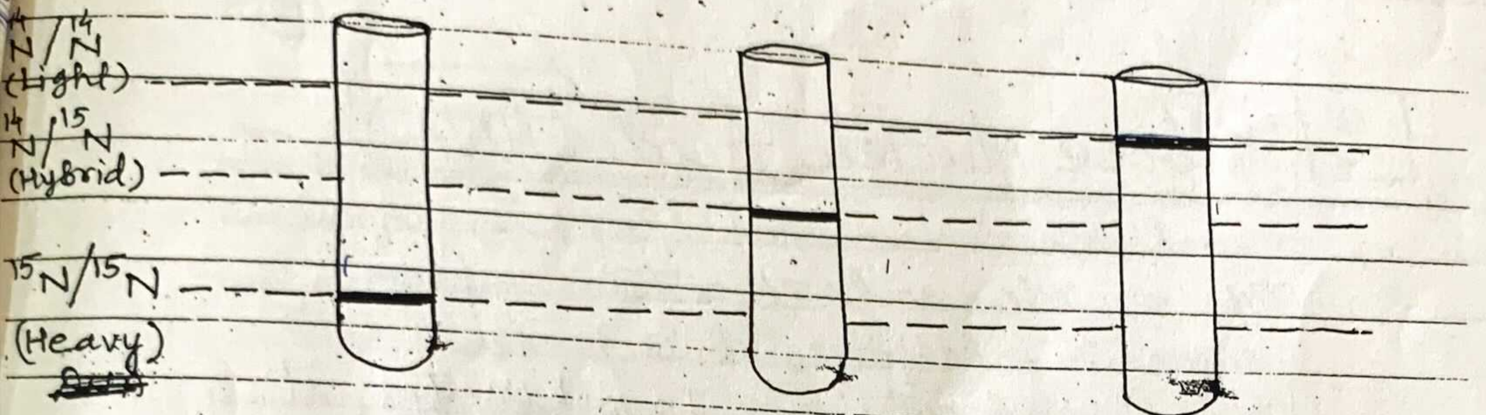


Fig.: Messelson & Stahl Experiment showing hybrid DNA

Nitrogen is the major constituent of DNA. It is found in ^{14}N isotope. *Escherichia coli* was grown ~~in~~ for several generations in a medium with ^{15}N. The DNA of the resulting cells had a higher density (was heavier). After that, the cells ~~were~~ with ^{15}N in their DNA were put back into a ^{14}N medium and allowed to divide. DNA was extracted from a cell & was found to be hybrid ~~to~~ ($^{14}N + {}^{15}N$). This hybrid was not heavy as ^{15}N and not light as ^{14}N, but had an intermediate density.

The DNA ~~was in~~ had intermediate density because it had ~~an~~ a ^{15}N DNA strand & a ^{14}N strand. The ^{15}N strand was ~~one of the~~ original strand and ^{14}N strand was newly synthesized strand. This shows that, the replication of DNA is of semi-conservative manner.

[2] Ribose Nucleic Acid (RNA) :—

Ribose nucleic acid (RNA) is another type of nucleic acid. It contains ribose sugar. In eukaryotes, it is present both in nucleus & cytoplasm. In prokaryotes, it is present in cytoplasm.

Chemical composition of RNA

RNA is made up of 3 components — ribose sugars, phosphate groups and nitrogenous bases :-

(I) **Ribose sugar** :— It is a pentose sugar having one oxygen atom ~~two~~ more as compared to deoxyribose sugar.

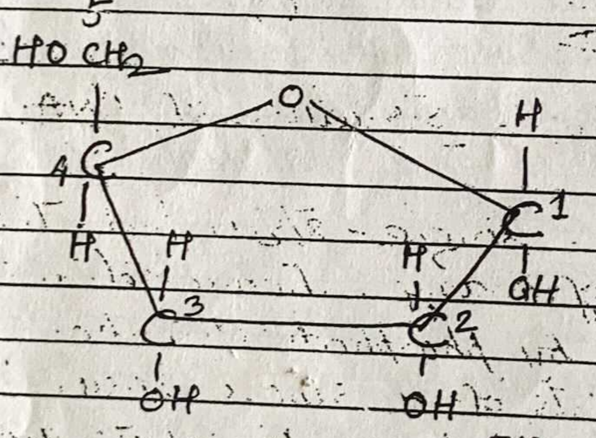


Fig: Ribose sugar

(II) **Phosphate group** :— It is an inorganic compound present in the form of phosphoric acid.

AKU
Fun^n ER
Lysosome mRN
prokaryotic cel
mitosis
STEN of gene
Golgi comple
rRNA
8 M = Rel^n of
Gene & m
& disea

(III) Nitrogenous bases :— The nitrogenous bases are of two types — purines & pyrimidines. The purines are Adenine (A) & Guanine (G), while pyrimidines are Uracil (U) and Cytosine (C).

Structure of RNA :—

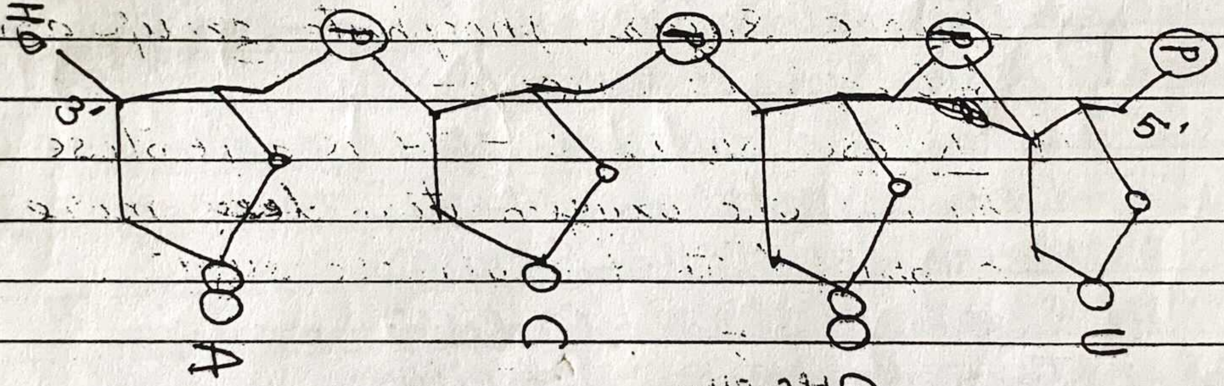


Fig.1- Structure of RNA.

RNA is single-stranded. The strand is made up of many ribonucleotides, linked to each-other by phosphodiester bonds. Each nucleotide consists of a ribose sugar, a phosphate group & a nitrogenous base. Phosphate group is attached to 5' carbon atom & nitrogenous base to 1' carbon atom of the ribose sugar. Nitrogenous bases are Adenine (A), Guanine (G), Uracil (U) and cytosine (C).

The single strand of RNA may be straight or folded at some regions. In folded regions, there is pairing between A & U and G & C. In unfolded regions, nitrogenous bases remain unpaired. Because of this, RNA does not show purine-pyrimidine equity.

Types of RNA :-

RNA is of two main types:-

① **Genetic RNA** :- RNA which acts as genetic material is called genetic RNA. It occurs in certain viruses) mostly plant viruses and many bacteriophages (bacterial viruses).

② **Non-genetic RNA** :- RNA which does not act as genetic material is called non-genetic RNA. It is found in all those organisms, where DNA acts as genetic material. Non-genetic RNA is involved in protein synthesis. It is of 3 types:

i) Messenger RNA (mRNA)
ii) Ribosomal RNA (rRNA), and
iii) Transfer RNA (tRNA)

(i) Messenger RNA (mRNA) :-

- It is called mRNA because it carries information for protein synthesis from DNA to ribosomes.
- Jacob & Monod (1961) proposed the name messanger RNA for RNA carrying information for protein synthesis from DNA.
- Usually each gene transcribes its own mRNA. Therefore there are as many types of mRNA molecules as there are genes. There may be 1000-10,000 different species of mRNA in a cell. These mRNAs differ from each-other in sequence of bases & their length.

The mRNA shows following structural features:-

1) <u>Cap</u> :- At the ~~end of~~ 5' end, cap is present. This is blocked methylated structure. Without cap, mRNA rarely binds to the ribosome.
2) <u>Non-coding region-I (NC-I)</u> : The cap is followed by non-coding region-I. It consists of 10-100 nucleotides. It is rich in A & U and does not translate proteins. → It initiates protein synthesis.
3) <u>Initiation codon</u> : It is AUG in both prokaryotes & eukaryotes.
4) <u>Coding region</u> : It consists of about 1500 nucleotides & translates proteins
5) <u>Termination codon</u> : It terminates protein synthesis. It is UAA, UAG or UGA
6) <u>Non-coding region-II (NC-II)</u> : It consists of 50-150 nucleotides & doesnt Transla ~~protein~~ proteins. This region contains AAU AAA sequence.
7) <u>Poly A sequence</u> : At the 3' end, there is polyadenylate or poly A sequence. It consists of 200-250 nucleotides, but which becomes shorter with age. It is added before mRNA reaches the cytoplasm.

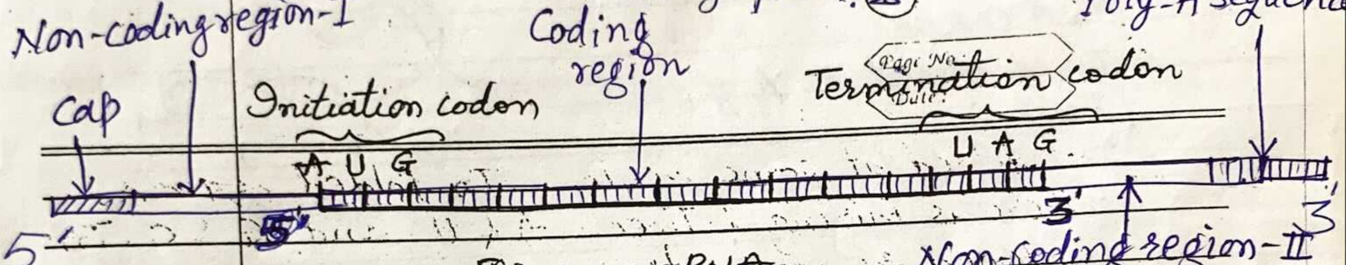


Fig :- mRNA

- ~~The mRNA~~ is about 3-5 % of total RNA content of the cell. - It is always straight & base pairing is absent.
- It is produced on the DNA strand. The process is called <u>transcription</u>.
- The base sequence of mRNA is complementary to that of DNA strand.
- The bases are organized into triplets. The triplets of bases on mRNA are ~~is~~ called <u>codons</u>.
- Each codon specifies one amino acid. Hence, sequence of codons of mRNA indicates the sequence of amino acids for the synthesis of protein.
- The codon by which protein synthesis begins is called <u>initiation codon</u>. It is AUG. (5' end)
- The codon by which protein synthesis ends is called <u>termination codon</u>. It may be UAA, UAG or UGA. (3' end)
- <u>Functions</u> :- mRNA carries genetic information for protein synthesis from DNA to ribosomes ie site of protein synthesis as they have codons on them. The sequence of bases on mRNA decides the sequence of amino acids in proteins. They come to the ribosomes & supply information for protⁿ synthesi

Page No.:
Date :

(ii) Ribosomal RNA (rRNA) :—

- It is present in ~~cyto~~ ribosomes & hence called ribosomal RNA (rRNA).
- It is about 80% of total RNA.
- It is folded upon itself in certain regions.
- In folded regions, complementary base pairing occurs.

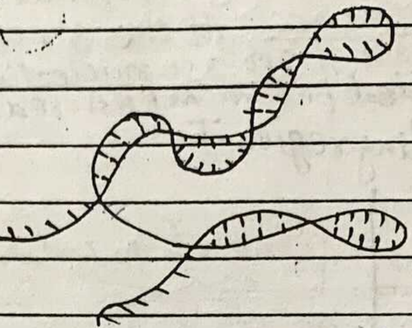


fig: rRNA

→ **Functions** :— rRNA provides proper binding site for mRNA on the ribosome. It orients mRNA is such a way that, codons are properly read. It protects mRNA from the action of enzymes. It releases tRNA ~~I~~ after transfer of amino acid to polypeptide chain.

~~(iii) Transfer RNA (tRNA)~~ :—

- The 70S ribosome of prokaryotes contains 30S subunit & 50S subunit. The 50S subunit contains 23S and 5S rRNAs, while 30S contains 16S rRA
eukaryotes, 80S ribosome consists of 40S and 60S subunits. The 40S subunit contains 16–18S rRNA and the 60S subunit contains 25–28S rRNA, 5.8S rRNA & 5S rRNA.
- In the helical region, base pairs are complementary.
- In the unfolded region, the bases have not complements.
- Hence Ribosomal RNA does not show purine-pyrimidine equality.
- The rRNA strands unfold upon heating & folded upon cooling.

(iii) Transfer RNA :-
(tRNA)

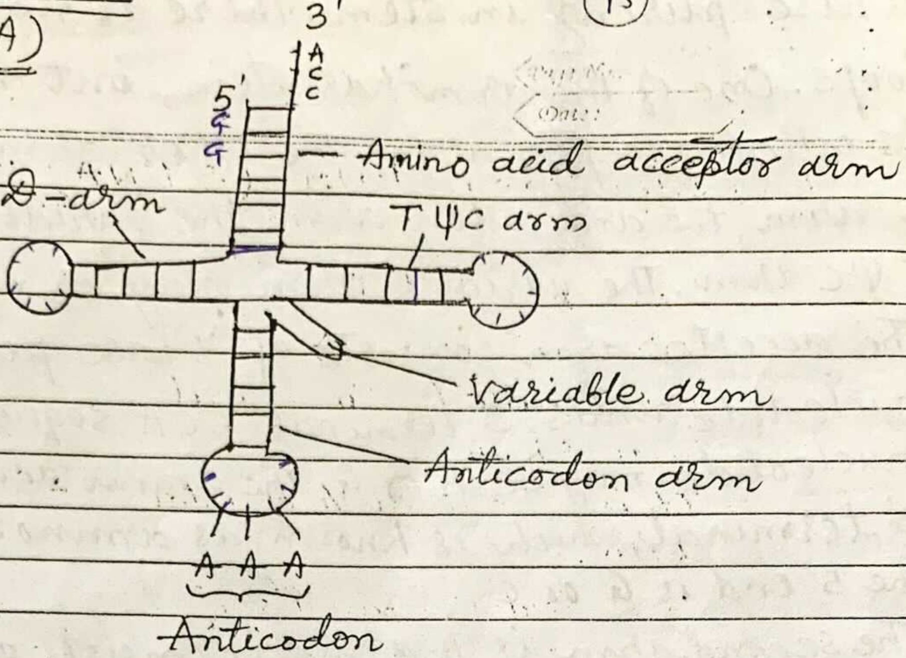


Fig: tRNA (Clover leaf model)

- It transfers amino acids to the site of protein synthesis. Hence, it is called transfer RNA.
- It is also called soluble RNA as it can not be precipitated by centrifugation.
- It is about 10 - 20% total RNA content of the cell.
- It consists of a single strand, folded upon itself. It results in either cloverleaf model or hair pin model.
- ~~On one loop of tRNA, there are anticodon is a triplet of nitrogenous bases called anticodon.~~
- ~~The anticodon pairs with complementary codon of mRNA molecule.~~ ~~Functions:-~~
- ~~The tRNA carries amino acids to mRNA during protein synthesis. Each type of amino acid is carried by a specific type of tRNA.~~
- The tRNA molecule is made up of 73-93 nucleotides.
- Only 0.025% of DNA codes for tRNA.
- The tRNA consists of CCA sequence at 3' end & G or C at 5' end.
- The nucleotidesequence of tRNA was worked out by Holley et al. (1965). Cloverleaf model suggested by Holley is widely accepted.
- According to cloverleaf model, the single poly-nucleotide chain of tRNA is folded upon itself to form 5 arms. As a result of the folding, the

5' and 3' ~~arms~~ ends of the chain come near each -other. An arm consists of a stem and a loop. There is base-pairing in stem. There is no base pairing in loops. One of the arm has stem, but not a loop and is called acceptor arm. The other arms are called the D-arm, the anticodon arm, the variable arm and the TψC arm. The variable arm may or may not have stem.

1)- The acceptor arm consists of 7 base pairs & 4 unpaired nucleotide units: 3' terminal CCA sequence & fourth nucleotide may be A or G. The amino acid attaches to the 3' terminal, which is known as aminoacid binding site. The 5' end is G or C.

2)- The second arm is D-arm. It consists of ~~15-18 nucleotide~~ 3-4 base pairs in stem & 7-11 unpaired bases in loop. The loop is called Loop-I or dihydroxyuridine (DHU) loop or D-loop.

3) The third arm or anticodon arm consists of a stem of 5 base pairs & a loop called Loop-II or anticodon loop. This loop consists of 7 unpaired nucleotides, of which the middle 3 form anticodon. The anticodon recognises the three complementary bases which constitute codon of mRNA.

4)- The variable arm (mini arm / mini loop) has a loop containing 4-5 bases, but no stem. Its function is not known.

5)- The TψC arm consists of a stem having 5 base pairs & loop of 7 nucleotides. The outermost of the 5 pairs of stem is C-G. The TψC loop contains a constant TψC sequence. All the tRNAs have a ribosome recognition site on the TψC loop.

(Modified bases)

ψ-pseudouridine
T-ribothymine

Functions of tRNA :- Transfer RNA carries ~~specific~~ aminoacid from cytoplasm to the site of protein synthesis (ribosomes). There is a specific tRNA for a ~~amino acid~~ particular amino acid. ~~tRNA~~ The tRNA has a site for amino acid attachment at its 3' or CCA end & anticodons on the anticodon loop. Hence, it gives a particular amino acid for protein synthesis.

Genetic code :-

Genetic code is defined as "the sequence of nitrogenous bases (nucleotides) in mRNA molecule, which contains the information for the protein synthesis." ~~of protein molecules~~.

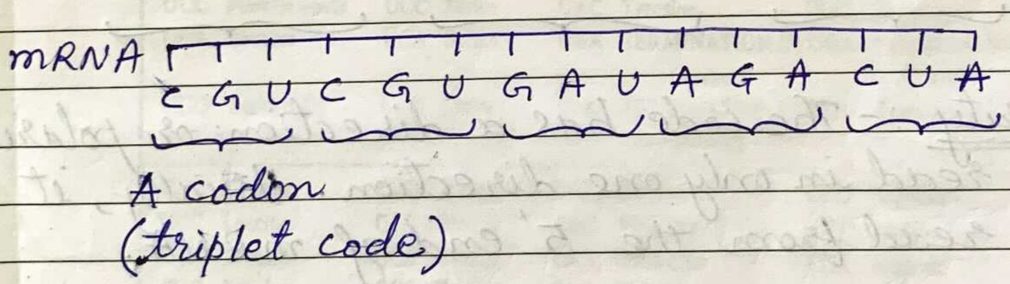
A codon is defined as "the sequence of 3 nitrogenous bases (nucleotides) in mRNA, which codes for a single amino acid."

Nirenberg & Mathaei (1961) experimentally proved that, a single amino acid is determined by a sequence of three nitrogenous bases. The sequence of three nitrogenous bases determining a single amino acid is called ~~a~~ triplet code.

The genetic code has following properties —

1) Triplet :-

A single amino acid is specified by a set of three nucleotides in the mRNA. This is called as codon. Hence, a codon is a triplet.

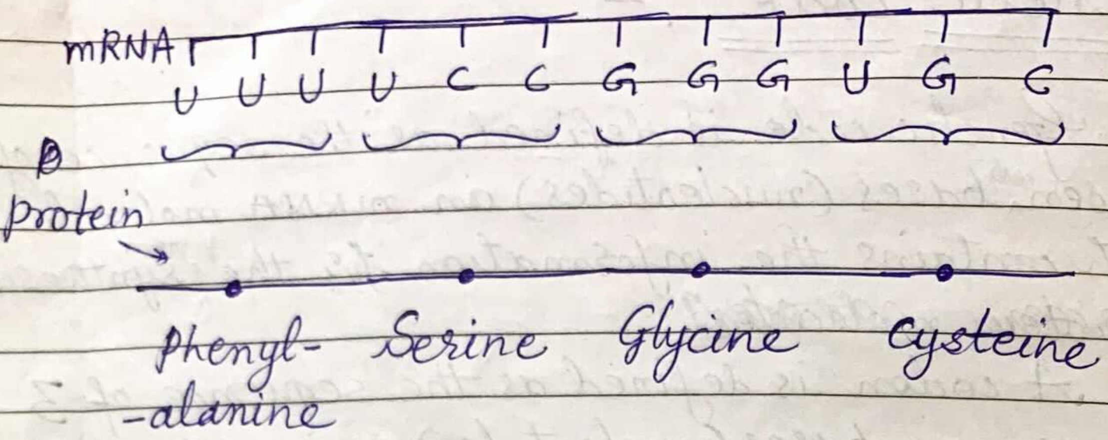


2) Universal :-

Genetic code is universal in nature. One type of code determines the same amino acid in all organisms including viruses, bacteria, plants and animals.

3) Colinearity :-

The codons in mRNA and the amino acids in polypeptide chains have a linear arrangement.

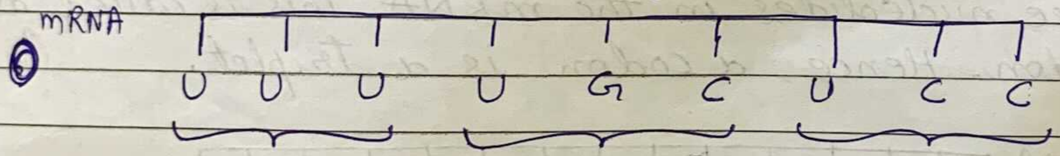


4) Commaless :-

There is no comma or punctuation between the adjacent codons. That is, each codon is immediately followed by the next codon without spaces in between them.

5) Non-overlapping :-

The codons are non-overlapping. ie. each letter is read only once.



6) Polarity :- The code has a direction or polarity. It is read in only one direction. Usually, it is read from the 5' end of mRNA.

7) Initiation codon :-

The synthesis of a polypeptide chain is initiated by a codon called initiation codon. It is located in the begining of a cistron (gene). It is AUG & codes for amino acid methionine.

8) Termination codon :-

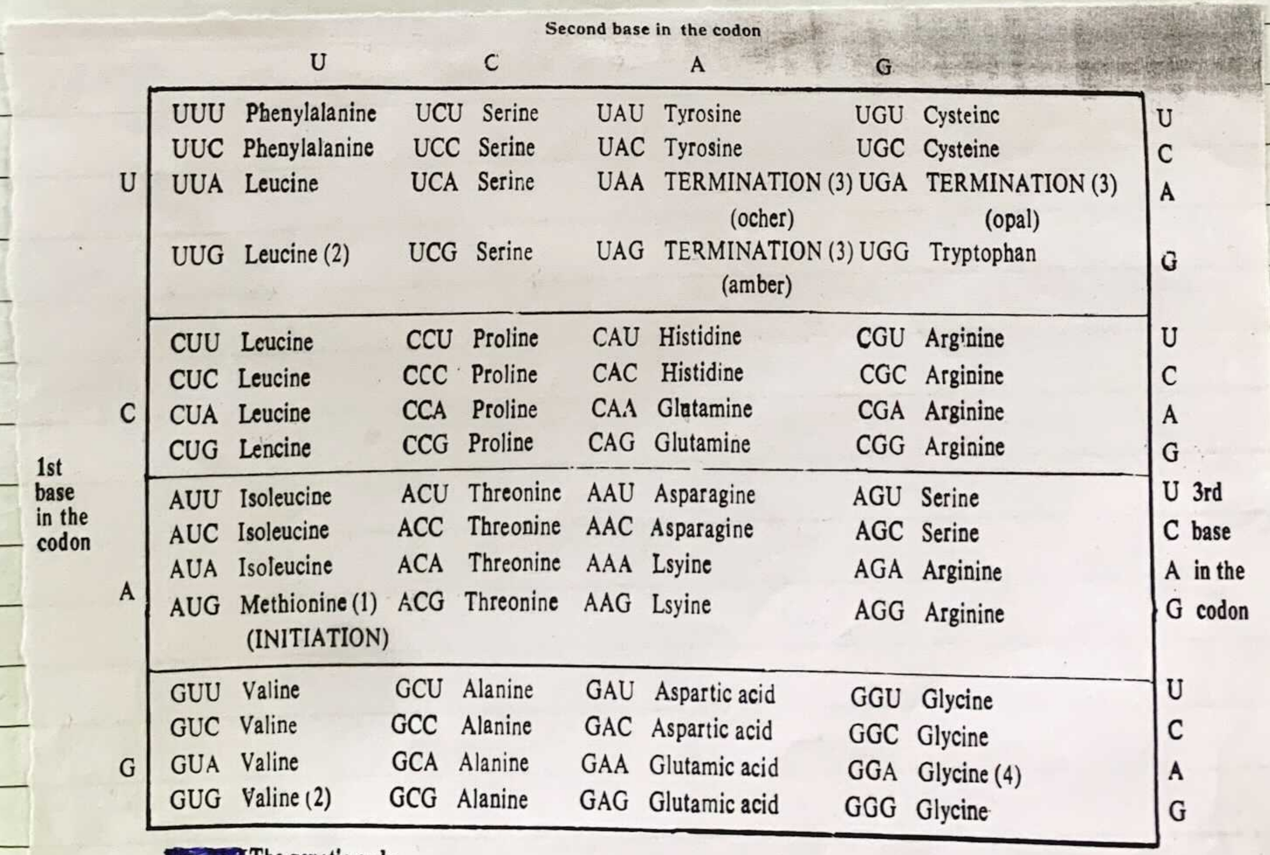
The synthesis of a polypeptide chain is completed by a codon called termination codon. It is located at the end of a cistron (gene). The termination codon may be UAA, UGA or UAG. The termination codon does not code for any amino acid. Hence, it is also called non-sense codon. There is no anticodon for non-sense codon.

9) Codons & anticodons :-

The codons of mRNA can pair with complementary codons of tRNA. The ~~codon~~ triplet of bases on tRNA is called anticodon. In base pairing, A pairs with U and G with C.

10) Degeneracy :-

A single amino acid may be specified by many codons. This is called degeneracy. Such codons are called degenerate codons. eg: Phenylalanine has two codons, such as UUU and UUC.



Second base in the codon

1st base in the codon	U	C	A	G	3rd base in the codon
U	UUU Phenylalanine	UCU Serine	UAU Tyrosine	UGU Cysteinc	U
	UUC Phenylalanine	UCC Serine	UAC Tyrosine	UGC Cysteine	C
	UUA Leucine	UCA Serine	UAA TERMINATION (3) (ocher)	UGA TERMINATION (3) (opal)	A
	UUG Leucine (2)	UCG Serine	UAG TERMINATION (3) (amber)	UGG Tryptophan	G
C	CUU Leucine	CCU Proline	CAU Histidine	CGU Arginine	U
	CUC Leucine	CCC Proline	CAC Histidine	CGC Arginine	C
	CUA Leucine	CCA Proline	CAA Glntamine	CGA Arginine	A
	CUG Lencine	CCG Proline	CAG Glutamine	CGG Arginine	G
A	AUU Isoleucine	ACU Threonine	AAU Asparagine	AGU Serine	U
	AUC Isoleucine	ACC Threonine	AAC Asparagine	AGC Serine	C
	AUA Isoleucine	ACA Threonine	AAA Lsyine	AGA Arginine	A
	AUG Methionine (1) (INITIATION)	ACG Threonine	AAG Lsyine	AGG Arginine	G
G	GUU Valine	GCU Alanine	GAU Aspartic acid	GGU Glycine	U
	GUC Valine	GCC Alanine	GAC Aspartic acid	GGC Glycine	C
	GUA Valine	GCA Alanine	GAA Glutamic acid	GGA Glycine (4)	A
	GUG Valine (2)	GCG Alanine	GAG Glutamic acid	GGG Glycine	G

[redacted] The genetic code.

: Protein synthesis :

" The construction of protein molecules in the cell by sequencially arranging amino acids using peptide bonds is called as protein synthesis". The amino acids are linked together in succession to produce a linear polypeptide chain. The polypeptide chain is a unit of a protein molecule. Protein synthesis occurs in cytoplasm.

Central dogma of protein synthesis :-

The DNA is transcribed into mRNA and mRNA is translated into a protein. This is central dogma of protein synthesis.

$$DNA \xrightarrow{\text{Transcription}} mRNA \xrightarrow{\text{Translation}} Protein$$

The mechanism of protein synthesis involves two steps :-

1. Transcription, &
2. Translation.

1 Transcription :-

The formation of mRNA complementary to DNA strand is called transcription.

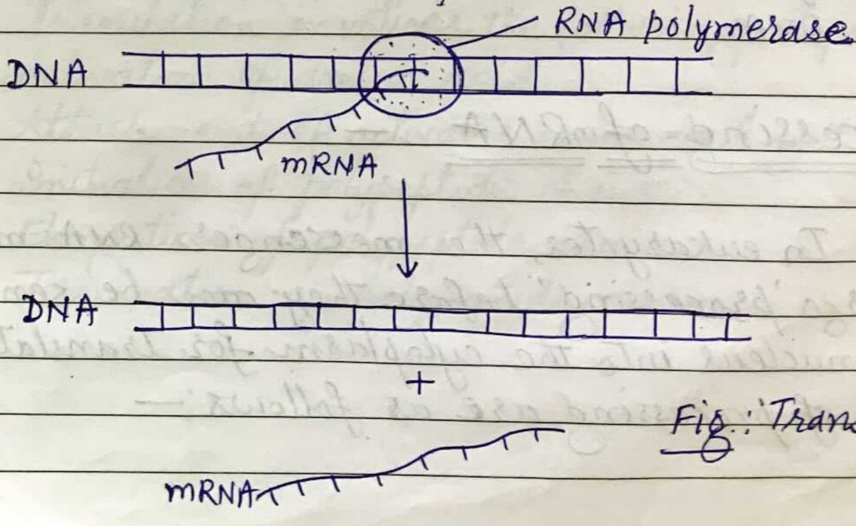


Fig.: Transcription

The mechanism of transcription has three major steps:- Initiation, elongation & termination.

~~1) Elongation~~

1) Initiation :- Transcription begins with initiation. An enzyme called RNA polymerase recognizes & binds to a 'promoter region' which contains the DNA sequence TATA. Proteins called 'transcription factors' may help RNA polymerase bind the promoter. Once bound, RNA polymerase begins to unwind the DNA to allow transcription of template DNA strand, starting at a specific signal called 'initiation site'.

ii) Elongation :- Once initiated, transcription enters the elongation phase. During elongation, RNA polymerase moves along the template DNA strand (3'-5') making the complementary mRNA (5'-3'). If the template DNA displays a G, RNA polymerase adds C; if the template DNA displays an A, RNA polymerase adds an U, and so on...

iii) Termination :- Elongation continues until the RNA polymerase reaches a signal sequence called the 'termination site'. Termination is caused by the protein 'rho factor' which binds to RNA polymerase. RNA polymerase & mRNA transcript are released from DNA.

Processing of mRNA :-

In eukaryotes, the messenger RNAs must undergo 'processing' before they can be sent from nucleus into the cytoplasm for translation. Steps of processing are as follows:-

a) **Splicing** :- Eukaryotic mRNA consists of alternating ~~sequen~~ sequences of introns (non-coding sequences) and exons (coding sequences). During splicing, introns are excised & exons are joined together. The process of excision of introns & rejoining of exons is called splicing.

b) **Capping** :- The 5' end of the mRNA is capped with a guanosine triphosphate (GTP) nucleotide. Capping protects the mRNA from degradation at the 5' end & also helps the mRNA to bind the ribosome. Ribosomes use mRNAs to direct the building of proteins.

c) **Polyadenylation** :- The 3' end of mRNA receives 200-210 adenylate residues. The sequence is AAAAA. This addition is called poly A tail. This protects the mRNA from degradation at 3' end & increases the stability of mRNA.

2 Translation :-

Translation is a process by which the base sequence of mRNA is interpreted into amino acid sequence of a polypeptide chain. In short, Translation is the synthesis of polypeptide chain.

Translation involves the following steps:-

i) Activation of amino acid.
ii) Attachment of activated amino acid with tRNA.
iii) Initiation of polypeptide chain.
iv) Elongation of polypeptide chain.
v) Termination of polypeptide chain.

i) Activation of amino acid :- Amino acids are the building blocks of proteins. They are present in the cytoplasm. They are activated before they are transported ~~to~~ by tRNA. The amino acids are activated by ATP with the help of enzyme amino acyl synthetase. The activated amino acid is called as aminoacyl adenylate or aminoacyl AMP.

$$aa + ATP + E \longrightarrow E\text{-}aa\text{-}AMP + Pi$$

ii) Attachment of activated amino acid with tRNA :- The activated amino acid is attached to the acceptor arm of tRNA. This is called aminoacylation of tRNA. It is catalysed by aminoacyl tRNA synthetase. The product formed is called aminoacyl tRNA complex.

$$E\text{-}aa\text{-}AMP + E + tRNA \longrightarrow aa\text{-}tRNA + AMP + E$$

iv) Elongation ~~Initiation~~ of polypeptide chain :- ~~Protein synthesis is initiated by the selection & transfer of the first amino acid to ribosomes~~. Within the ribosome, tRNAs step through 3 sites - the acceptor site (A site), peptidyl (P site) and exit site (E site). ~~In the A site, the appropriate tRNA carrying its amino acid, binds to the next codon of the mRNA~~. At the begining, ~~of elongation~~, the initiator tRNA sits in the P site. In the A site, an appropriate tRNA carrying its amino acid, binds to the next codon of mRNA. When a correct codon - anticodon match is made in the A site, the ribosome acts as an enzyme, catalyzing the covalent bonding of amino acids attached to the A site and P site tRNAs. As this bonding takes place, the mRNA & tRNAs move through the ribosome exactly one codon. This movement is called translocation. Translocation shifts A site tRNA into P site and shifts the P site tRNA into E site, where it is released from the ribosome. This cycle of recognition, catalysis, translocation & release is repeated for each codon as the mRNA feeds through the ribosome like film through a projector. The product is an elongating chain of covalently linked amino acids. The sequence of amino acids is directed by mRNA.

iii) Initiation of polypeptide chain :- Protein synthesis is initiated by the selection & transfer of first amino acid to ribosome. A mRNA, the small ribosomal unit & a special molecule called initiator tRNA assemble into a complex. Large ribosomal unit binds to this complex. The message on mRNA is in the form of codons, which are read in 5′–3′ direction. Each codon corresponds to specific amino acid. The codon AUG (codes for amino acid methionine) is the site where Translation begins & is the site recognized by the initiator tRNA. Each tRNA is covalently bonded to a unique amino acid & contains an anticodon that recognizes the mRNA codon specific to that amino acid. Codons & anticodons are both made up of RNA, so they can recognize each-other through complementary base-pairing. The enzyme responsible for bonding the correct amino acid to each ~~other~~ kind of tRNA are called aminoacyl-tRNA synthetases.

iv) Elongation of polypeptide chain :-

– see page 23 –

v) <u>Termination of polypeptide chain</u> :- At the end of mRNA, a stop codon signals that translation should stop. There are 3 stop codons - UAA, UAG & UGA. The stop codon is recognized by a 'release factor'. It triggers release of newly synthesized polypeptide chain. The ribosome subunits, tRNAs and mRNA become separated.

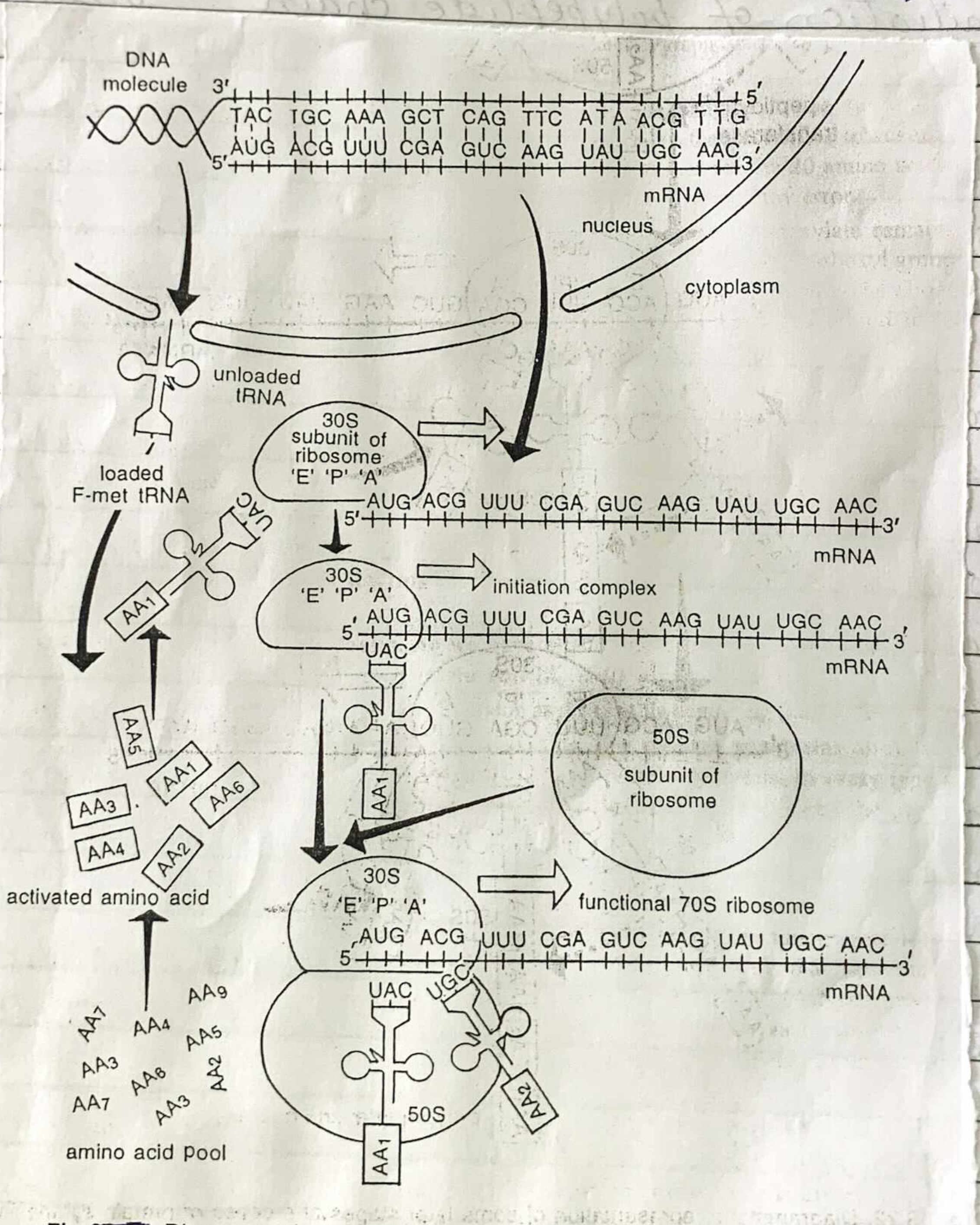


Fig [redacted]. Diagrammatic representation of the some early steps of the process of protein synthesis of *E. coli*.

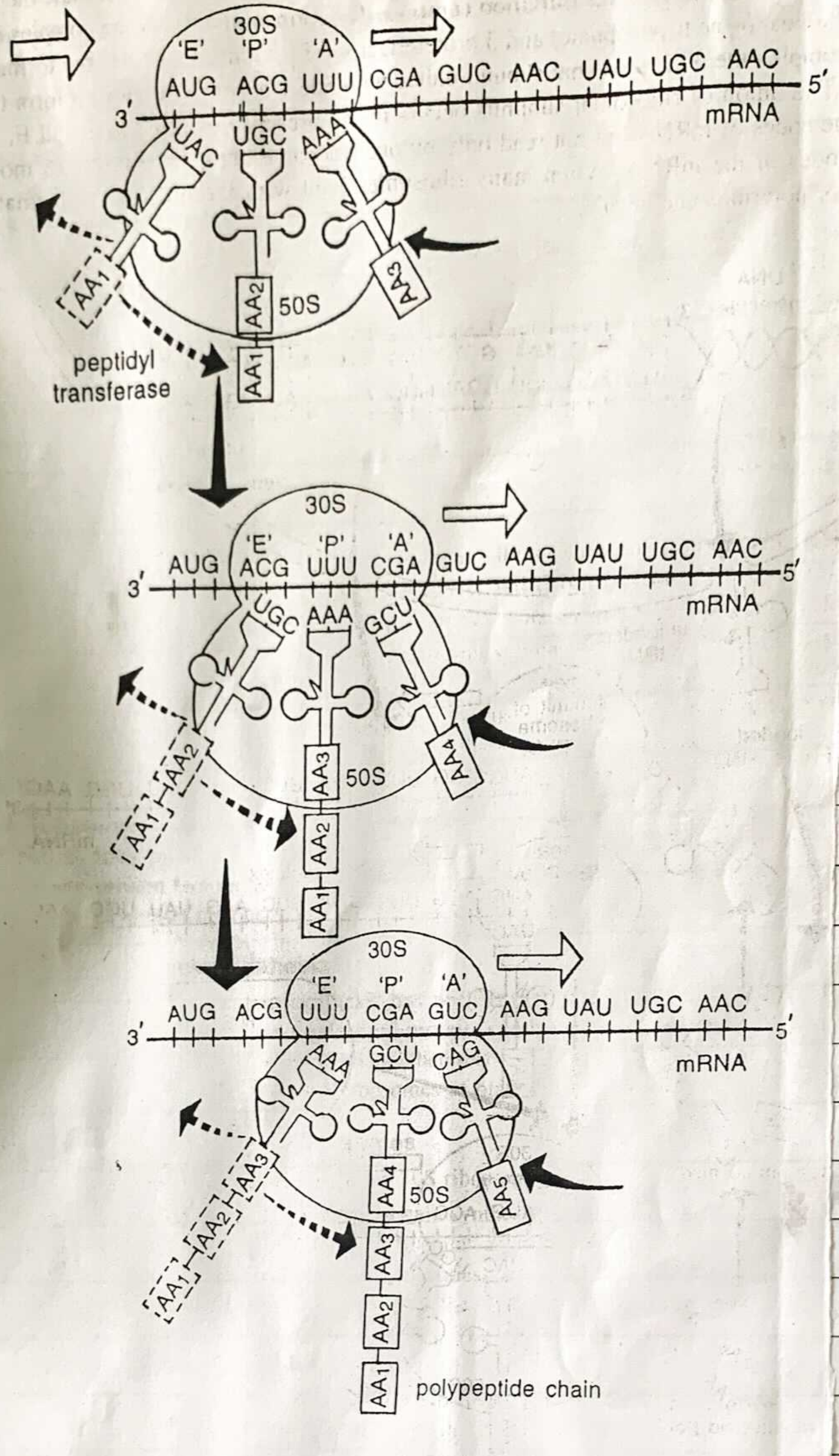


Fig. [redacted] Diagrammatic representation of some later stages of process of protein synt
in *E. coli.* AA_3 = phenylalanine, AA_4 = arginine, AA_5 = valine, AA_6 = lysine.

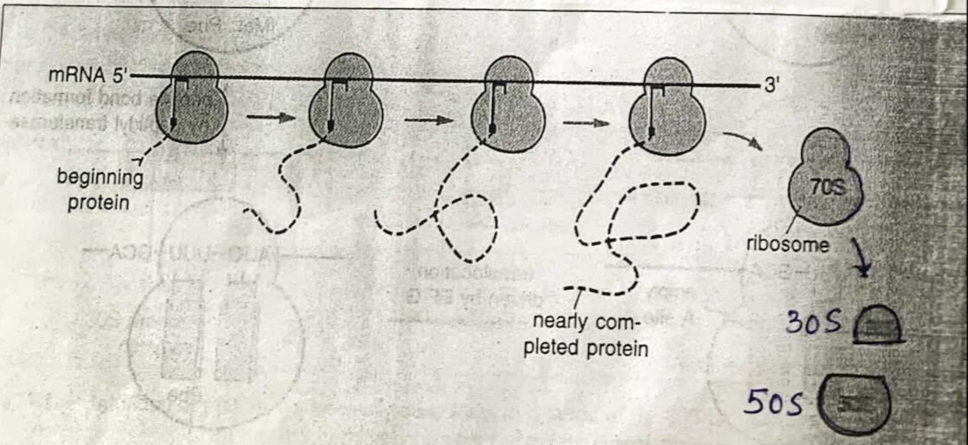


Fig. [redacted] Polysome formation : diagram shows the relative movement of the 70S ribosome & mRNA and growth of the protein chain (after Freifelder, 1985).