TRANSCRIPTION

- Formation of RNA over DNA templet is called transcription. Out of two strand of DNA only one strand participates in tr anscription and called ***"Antisense strand".***
- The segment of DNA involved in transcription is ***"Cistron".***
- RNA polymerase enzyme involved in transcription. In eukaryotes there are three types of RNA plymerases.
  - RNA polymerase - I for 28s rRNA, 18s rRNA, 5.8s rRNA.
  - RNA polymerase-II for m-RNA.
  - RNA polymerase-II for t-RNA, 5s rRNA, SnRNA
- In eucaryotes RNA polymerase enzyme composed of 10-15 polypeptide chains.
- Prokaryotes have only one type of RNA polymerase which synthesizes all types of RNAs.
- RNA polymerase of E. Coli has five polypeptide chains β, β', α, α and σ.
- σ polypeptide chain is also known as σ factors (sigma factor).
- Core enzyme + Sigma factor $\Rightarrow$ RNA polymerase

  $(\beta, \beta' \alpha, \alpha)$ (σ)

  Following steps are present in transcription -

(1) **INITIATION :**

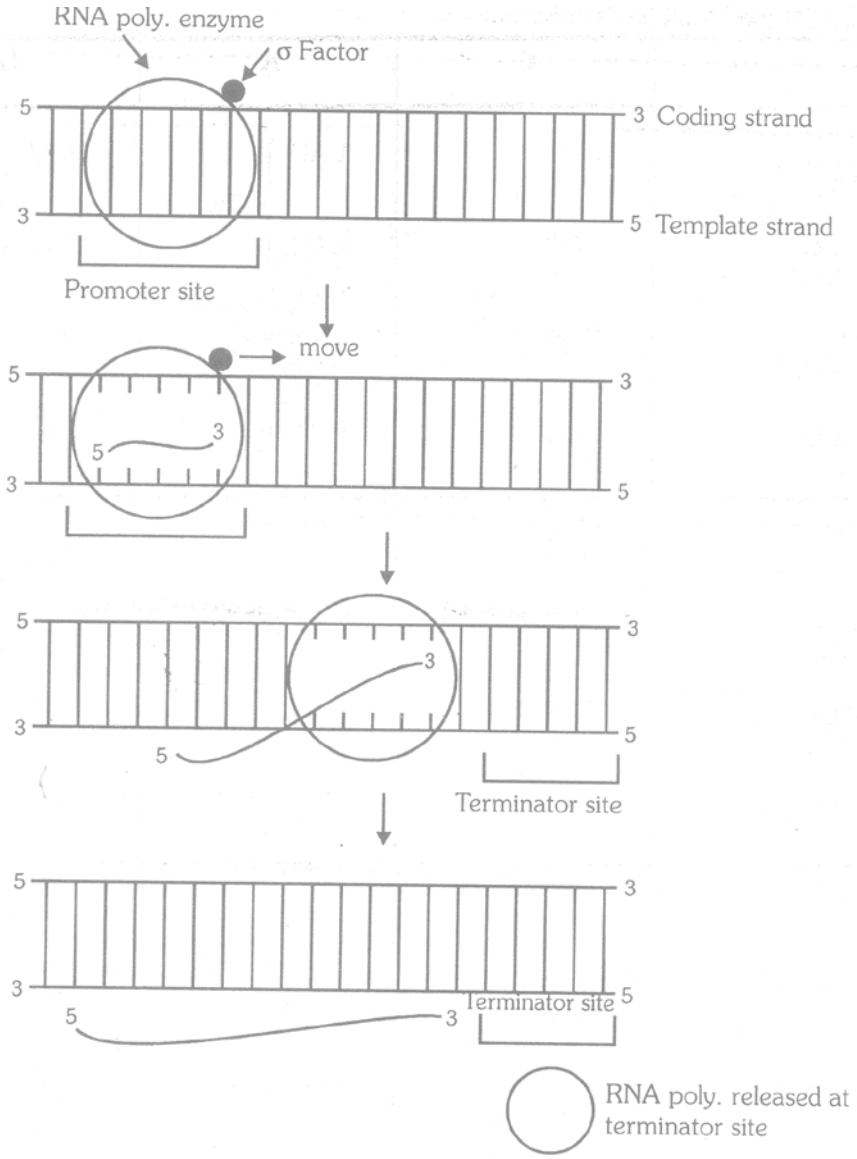
- DNA has a ***"Promoter site or initiation site"*** where transcription begins and a ***"Terminator site"*** where transcription stops.
- Sigma factor (σ) recognises the promoter site of DNA.
- With the help of sigma factor RNA polymerase enzyme attached to a specific site of DNA called ***"Promoter site".***
- In prokaryotes before the 10 N_2 base from ***"Promoter site"*** a sequence of 6 base pairs (TATAAT) is present on DNA, which is called ***"Pribnow box".***
- In eukeryotes before the 20 N_2 base from ***"Promoter site"*** a sequence of 7 base pairs (TATAAAA) or (TATATAT) is present on DNA which is called **"TATA box** or **Hogness box"**
- At promoter site RNA polymerase enzyme breaks H-bonds between two DNA strands and separates them.
- One of the m strand takes part in Transcription. Transcription proceeds in 5' $\rightarrow$ 3" direction.
- Ribonucleotide triphosphate come to lie opposite complementary nitrogen bases of anti sense strand.
- These Ribonucleotides present in the form of triphosphate ATP, GTP, UTP and CTP in nucleoplasm. When they used in transcription, pyrophosphatae hydrolyse two phosphates from each activated nucleotide. This releases energy.
- This energy used in process of transcription.

(2) ELONGTION :

- RNA polymerase enzyme establishes phosphodiester bond between adjacent ribonucleotides.
- Sigma factor separates and core enzyme moves along the anti sense strand till it reaches terminator site

(3) TERMINATION :

- When RNA polymerase enzyme reaches at terminator site, it separates from DNA templet.
- In terminator site on DNA, N_2 bases are present in palindromic sequence.
- In most cases RNA polymerase enzyme can recognise the ***Terminator site'*** and stop the synthesis of RNA chain, but in prokaryotes, it recognises the terminator site with the help of **Rho factor (ρ factor)**
- Rho (ρ) factor is a specific protein which helps RMA polymerase enzyme to recognise the terminator site.



GENETIC CODE

Term Given by **George Gamow.**

- The relationship between the sequence of amino acids in a polypeptide chain and nucleotide sequence of DNA of m-RNA is called ***genetic code.***
- There occur 20 types of amino acids which participate in protein synthesis. DNA contains information for the synthesis of any types of polypeptide chain. In the process of transcription, information transfers from DNA to m-RNA in the form of complementary N_2-base sequence.
- m-RNA contains code for each amino acid and it is called codon. A codon is the nucleotide sequce in m-RNA which codes for particular amino acid ; where the genetic code is the sequence of nucleotides in m-RNa molecule, which contains information for the synthesis of polypeptide chain.

Triplet Code :

- The main problem of genetic code was to determine the exact number of nucleotide in a codon which codes for one amino acid.
- There are found types of N_2-bases in m-RNA (A, U, G, C) for 20 types of amino acids.

	U	**C**	**A**	**G**	
U	UUU, UUC } Phe UUA, UUG } Leu	UCU, UCC, UCA, UCG } Ser	UAU, UAC } Tyr UAA } Ochre (ter min ator) UAG } Amber (ter min ator)	UGU, UGC } Cys UGA } Ter min ator UGG } Try	U C A G
C	CUUU, CUC, CUA, CUG } Leu	CCU, CCC, CCA, CCG } Pr o	CAU, CAC } His CAA, CAG } G In	CGU, CGC, CGA, CGG } Arg	U C A G
A	AUU, AUC, AUA } lie 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	CGU, GGC, GGA, GGG } Gly	U C A C

Fig : Triplet codons of mRNA for amino acids represented in tabular form.

- If genetic code is ***singlet*** i.e. codon is the combination of only one nitrogen base, then only four codons are possible A,C,G and U. These are insufficient to code for 20 types of amino acids.
- Singlet code = 4' = 4 × 1 = 4 codons
- If genetic code is ***doublet*** (i.e. codon is the combination of two nitrogen bases) then 16 codons are formed.
- Doublet code = 4^2 = 4 × 4 = 16 codons.
- 16 codons insufficient for 20 amino acid

A
C
G
U

Codons

Singlet code : 4 × 1 = 4 codons

AA	AC	AG	AU
CC	CA	CG	CU
GG	GA	GC	GU
UU	UA	UG	UC

Doublet code : 4 × 4 = 16 codons

Gamow :

- (1954) Pointed out the possibility of three letter code (Triplet code).
- Genetic code is triplet i.e. one codon consists of three nitrogen bases Triplet code = 4^3 = 4 × 4 × 4 = 64 codons
- In this case there occurs 64 codons in dictionary of genetic code.
- 64 codons are sufficient to code 20 types of amino acids.

Characteristics of Genetic Code :

(i) **Triplet in Nature :**

- A codon is composed of three adjacent nitrogen bases which specifies the one amino in polypeptide chain.

  For Ex. :

  - In m-RNA is there are total 90 N_2 - bases.
  - Then this m-RNA determines 30 amino acids in polypeptide chain.
  - In above example, number of Nitrogen bases are 90 so codons $\Rightarrow$ 30 and 30 codons decide 30 amino acid in polypeptide chain.

(ii) **Universality :**

The genetic code is applicable universally. The same genetic code is present in all kinds of living organism including viruses, bacteria, unicellular and multicellular organism.

(iii) **Non - Ambiguous :**

- Genetic code is non ambiguous i.e. one codon specifies only one amino acid and not any other.
- In this case one codon never code two different amino acids. Exception GUG codon which code both valine and methionine amino acid.

(iv) **Non - Overlapping :**

A nitrogen base i a constituent of only one codon.

(v) **Comma less :**

- There is no punctuation (comma) between the adjacent codon i.e. each codon is immediately followed by the next codon.
- If a nucleotide is deleted or added, the whole genetic code read differently.
- A Polypeptide chain having 50 amino acid shall be specialized by a linear sequence of 150 nucleotides. It a nucleotide is added in the middle of this sequence, the first 25 amino acids of polypeptide will be same but next 25 amino acids will be different.

(vi) **Degeneracy of Genetic code :**

- There are 64 codons for 20 types of amino acids, so most of the amino acids (expect two) can be coded by more than one codon. Single amino acid coded by more than one codon is called ***"Degeneracy of genetic code".*** This incident was discovered by ***Baurnfield*** and ***Nirenberg.***
- Only to amino acids ***Tryptophan*** and ***methionine*** are specified by single codon.
  - UGC for Tryptophan
  - AUG for methionine.
- All the other amino acid are specified or coded by 2 to 6 codons
- Leucine, serine and arginine are coded or specified by 6-codons.

  Leucine = CUU, CUC, CUA, CUG, UUA & UUG

  Sperine = UCU, UCC, UCA, UCG, AGU, AGC

  Arginine -= CGU, CGC, CGA, CGG, AGA, AGG
- Degeneracy of genetic code is related to third position (3' - end of triplet codon) of codon. The third base is described as ***"Wobbly base".***

Exception :

Different Codon :

- Normally UA and U GA are chain termination codon but in Paramecium and some other ciliates UAA and UGA code for glutamine amino acid.

Mitochonodial Gene :

- Normally AGG and AGA code for Arginine amino acid but in human mitochondria these function as stop codon.
- UGA, a termination codon corresponds to tryptophan while AUA (Codon for isoleucine) denotes methionine in human mitochondria.

Chain Initiation and Chain Termination Codon :

- Polypeptide chain synthesis is signalled by two initiation codons AUG or GUG.
- AUG codes methionine amino acid in eucaryotes and in prokaryotes AUG codes N-formyl methionine.
- Some times GUG also functions as start codon it codes for valine amino acid normally but when it is present at starting position it code for methionine amino acid.
- Out of 64 codons 3-codons are stopping or nonsense or termination codon.

  Nonsense codon do not specify and amino acid.

  UAA (Ochre)

  UAG (Amber) **Non-Sense Codon**

  UGA (Opal)
- So only 61 codons are sense codons which specify 20 amino acid.

WOBBLE HYPOTHESIS

- It was propounded by CRICK.
- Normally an anticodon recognises only one codon, but sometimes as anticodon recognises more than one codon. This is known as Wobbling. Wobbling normally occurs for third nucleotide of codon.
- For e.g. Anticodon AAG can recognise two anticodons i.e. UUU and UUC, both stand for phenyl alanine.

Type of m-RNA - m-RNA is of 2 types-

(1) **Monocistromic -** The m - RNA in which genetic signal is present for the formation of only one polypeptide chain.

(2) **Polycistronic :** The m-RNA, in which genetic signal is present for the formation of more than one polypeptide chains.

- Non sense codons are found in middle position in polycistronic m-RNA.

CENTRAL DOGMA

- Central dogma term was given by Crick.
- The formation (production) of m - RNA from DNA and then synthesis of protein from it, is known as **Central Dogma.**

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

- It means, it includes transcription and translation.
- The central dogma scheme of protein synthesis was presented by **Jacob and monad.**
- The detailed study of central dogma is done by **Nirenberb, Methai and Khorana.**
- **Beedle and Tatum** studies central dogma in a fungus ***Neurospora.***

$$\text{Replication} \underset{\text{Reserve Transcription}}{\overset{\text{Transcription}}{\rightleftharpoons}} \xrightarrow{\text{Translation}} Protein$$

Reverse Transciption :

- The formation of DNA from RNA is known as **Reserve - transcription.** It was discovered by **Temin and Baltimore** is Rous - sarcoma virus. So it also called **Teminism.**
- ss-RNA of Rous-Sarcom visur (Retro virus) produces ds-DNA is hot's cell with the help of enzyme reverse transcriptase (DNA-polymerase). This DNA is called c-DNA (Complimentary DNA). Some times this DNA moves in host genome. Such mobile DNA is called "Retroposon" (Oncogene).

PROTEIN - SYNTHESIS

(1) Activation of Amino acid :

- 20 types of amino acid participate in protein sysnthesis.
- Amino acid reacts with ATP to form ***"Amino acyl AMP enzyme complex"***, which is also known as **"Activated Amino acid***".*

$$\text{Amino acid} + ATP \xrightarrow[t-RNA\ \text{synthetase}]{\text{Amino acyl}} \text{Amino acyl AMP - enzyme complex} + PP$$

- This reaction is catalyzed by a specific **'Amino acyl t-RNA synthetse'** enzyme.
- There is a separate **'Amino acyl t-RNA synthetase'** enzyme for each kind of amino acid.

(2) ***Charging of t-RNA :***

- Specific activated amino acid is recognised by its specific t-RNA.
- No amino acid attached to the **'Amino acid attachment site'** of its specific t-RNA and **AMP** and enzyme are separated from it.

Amino acyl AMP enzyme complex + tRNA $\rightarrow$ Amino acyl t-RNA complex+ AMP + enzyme

- Amino acyl t-RNA complex is also called **"Charged t-RNA"**
- Now Amino acyl t-RNA moves to the ribosome for protein synthesis.

(3) ***Translation :*** *3 steps-*

(A) **Initiation of Polypeptice chain :**

- In this step 30 's' and 50 's' sub units of Ribosome, TPG, Mg^{+2} charged t-RNA, m-RNA and some initiation factors re required.
- In prokaryotes there are three initiation factors present - **IF1. IF2, IF3.**
- In Eukaryotes there are more than 3 initiation factors are present. Ten initiation factors have been identified in red blood cells -

eIF1, eIF2, eIF3, eIF4A, eIF4B, eIF4C, eIF4D, eIF4F, eIF5, eIF6.

- Initiation factors re specific protein.
- CTP and initiation factors promote the initiations process.
- In starting the both sub units of ribosome are separated with the help of IF3 factor.
- In prokaryotes with the help of **"S D sequence" (Shine-Delgarno sequence)** m-RNA recognises the smaller sub unit of ribosome. A sequence of 8 N_2 base is present before the 4-12 N_2 base of initiation Codon on mRNA, called **"SD sequence".** In smaller subunit of ribosome, a complementary sequence of **"SD sequence"** is present of 16 's' rRNA, which is called **"Anti Shine-delgarno sequence" (AD sequence)**

- With the help of 'SD' and 'ASD sequence' mRNA recognises the smaller su unit of ribosome.
- While is Eukaryotes, smaller sub unit of ribosome is recognised by **"7mG cap".**
- In Eukaryotes, 18 's' rRNA of smaller sub unit has a complementary sequence of **"7mG cap"**

$$\text{30 'S' sub unit+ m-RNA} \xrightarrow[Mg^{+2}]{\text{IF3}} \text{30 's' m-RNA - complex}$$

- This **"30 's' m-RNA - complex"** reacts with **'Formyl methionyl t-RNA - complex'** and **"30 's' m-RNA - formly methionyl t-RNA - complex"** form. This t-RNA attached with codon part of m-RNA. A CTP molecule is required.

30 'S' m-RNA - complex + Formyl methionyl t-RNA - complex

$$\text{GTP} \downarrow \begin{matrix}\text{IF2,IF3}\\ Mg^{+2}\end{matrix}$$

30 'S' m-RNA formyl methionyl t-RNA - complex

- Now larger sub unit of ribosome (50 'S' sub unit) joins this complex. The initiation factor released and complete 70 'S' ribosome is formed.
- In larger sub unit or ribosome is formed.
- In larger sub unit of ribosome is formed.

'P' site = Peptidyl site.

'A' site = amino acyl site.

E - site = exit site

- Starting codon of m-RNA is near to **'P' site** of ribosome, so t-RNA with formyl methionine amino acid first attaches to **'P' site** of ribosome and next codon of m-RNA is near to **'A' site** of ribosome. So new t-RNA with new amino acid always attach at **'A' site** of ribosome but in initiation step 'A' site is empty.

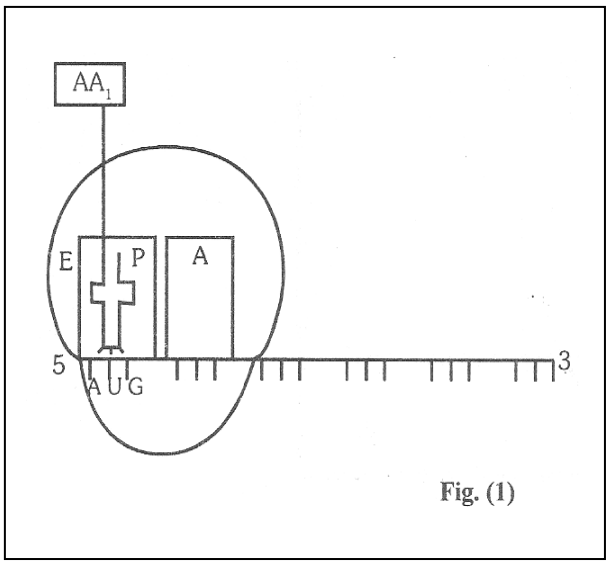


Fig. (1)

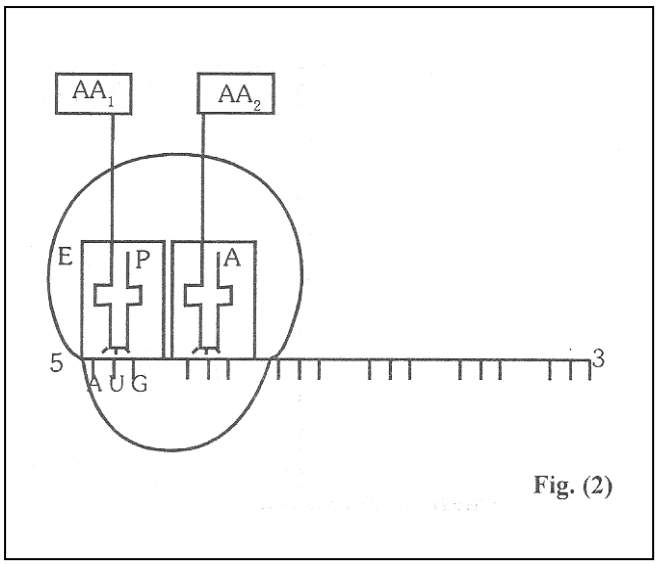


Fig. (2)

(B) **Chain - elongation :**

- New rRNA with new Amino acid attached at 'A' site Ribosome.
- First of all t-RNA of P-site is discharged so -COOH of P-site A. A. becomes free.
- Now peptide bond takes place between **-COOH group** of P site amino acid and **- NH_2 group** of A-site amino acid.
- Peptidyl transferase enzyme includes the formation of peptide bond. In peptide bond formation, a 23 'S' r-RNA is also helpful. This r-RNA acts as an enzyme so it is also called **"Ribozyme".**
- After formation of peptide bond t-RNA of **P site** released from ribosome via E-site and dipeptide attaches with **A site.**
- Now t-RNA of A site is transferred to **P site** and **A site** becomes empty.
- Now ribosome slides over m-RNA strand in **5' → 3"** direction. Due to sliding. Due to sliding of ribosome on m-RNA, new codon of m-RNA continuously available at **A site** of ribosome and according to new codon of m-RNA new amino acid attaches in polypeptide chain.
- **Translocase enzyme** helpful in movement of ribosome. **GTP** provides energy for sliding of ribosome.
- In elongation process some protein factors are also helpful, which known as **'Elongation factors'**
- In prokaryotes three **'Elongation factors'** are present - **EF-Tu, EF - Ts, EF - G.**
- In Eukaryotes two elongation factors are present **- eFF1, eFF2.**

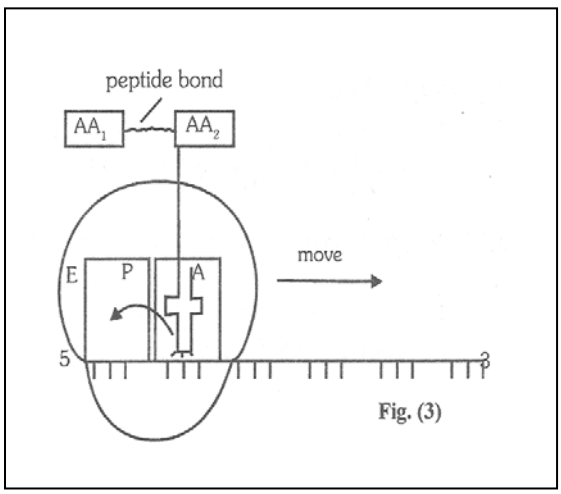


Fig. (3)

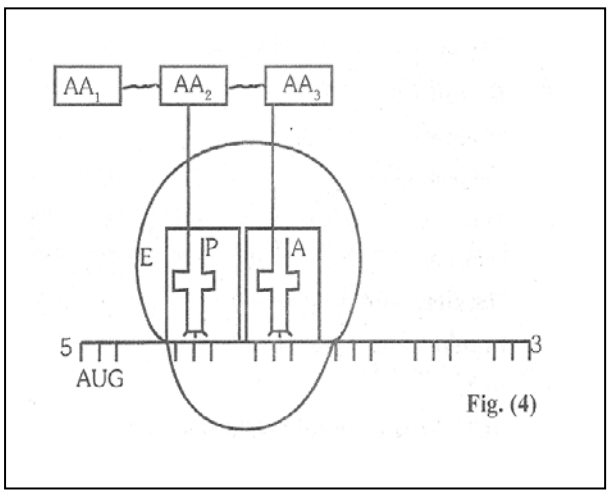


Fig. (4)

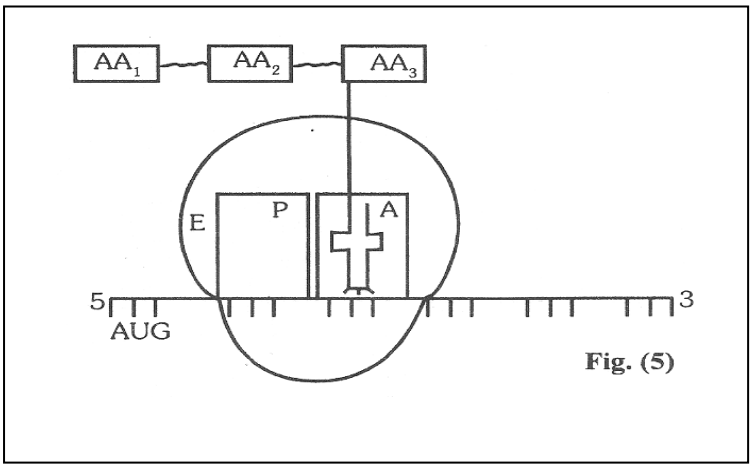


Fig. (5)

(C) **Chain - Termination :**

☛ Due to sliding of ribosome over m-RNA when any **Nonsense codon (UAA, UAG, UGA)** available at **A site** of ribosome, then polypeptide chain terminate.

☛ The linkage between the last t-RNA and the polypeptide chain is broken by three **release factor** called **RF1, RF2, RF3** with the help of **GTP**

☛ Peptidyl transferase enzyme also catalysed the releasing process.

☛ In eukaryotes only one **Release factor** is known - **eRF1.**

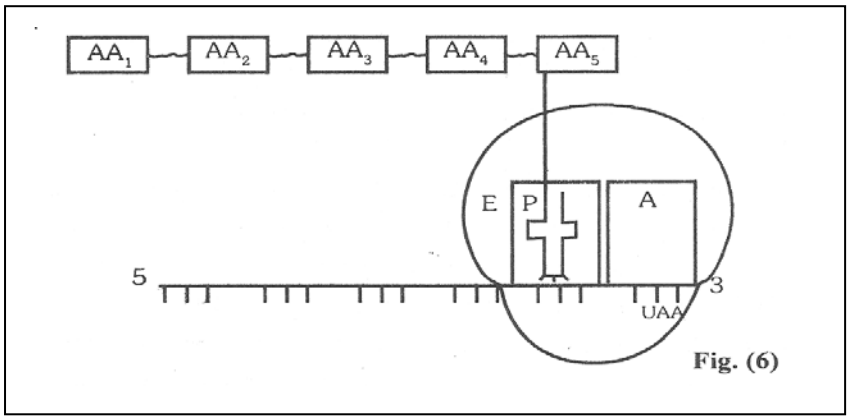


Fig. (6)

ESPECIAL POINTS

(1) The chargaff's fule is not (true) for RNA. It is valid only for double helical DNA i.e. for RNA it is A $\neq$ U and G $\neq$ C.

(2) The duplication of DNA was first of all proved in ***E. coli*** bacterium.

(3) ***E.coli*** Bacterium is mostly used for the study of DNA duplication.

(4) Hargovind singh Khurana first of all recognised the triplet codon for **Cysteine and Valine amino acids.**

(5) Cytoplamic DNA is 1- 5% of total cells DNA.

(6) Three lady scientists named **Avery, Mc - Leod and Mc Carty** (by their transformation experiments on bacteria) Proved that **DNA is a genetic material.**

(7) **Hershey and Chase** first of all proved that DNA is genetic material in bacteriophages.

(8) **Frankel and Conret** proved, RNA as a genetic material is viruses (g-RNA).

(9) AUC
ACU These anticodons do not exist.
AUU

(10) The structure formed by the combination of m - RNA and Ribosomes is knowsn as polyribosomes/Polysomes/Egosomes

(11)The formation of t - RNA takes place from the heterochromatin part of DNA.

(12)The formation of m - RNA takes place from the Euchromatin part of DNA.

(13)m - RNA is least stable. It is continuously formed and finished.

(14)In cytoplasm, t - RNA is present in the form of soluble colloid.

(15)**Nucleases :** Nucleases are the breaking enzymes of nucleic acids. These are of two types :

(1) Endo - Nucleases : These break down the nucleic acids from the inside.

(2) Exo-nucleases : These break done the nucleic acids from the ends (terminal ends). These separate each nucleotide.

(16)**Tay - Sachs - disease -**

This disease takes place due to excess storage of glycolipids.

(17)Excess storage of cerebrosides leads to **Guacher's disease.**

(18)**Some Inhibitors of Bacterial Proteins Synthesis :**

Antibiotic	**Effect**
Tetracyclin	Inhibits binding of amino-acyl tRNA to ribosome
Streptomycin	Inhibits initiation of translation and causes misreading
Chlormphernicol	Inhibits peptidyl transferase and so formation of peptide bonds
Erythromycin	Inhibits translocation of ribosome along mRNA
Neomycin	Inhibits interaction between tRNA and mRNA

(19) **Split gene :**

Discovered by sharp and Roberts. They awarded by Nobel Prize in 1993. Gene which contains non functional part along with functional part is known as split gene. Non functional part is called **intron** and functional part is called **exon.** By transcription split gene produces a RNA which contains conding and non coding sequence and called hn RNA (Hetero genous nuclear RNA). This hn RNA is unstable. Now 200 nucleotides of adenylic acid are added to its 3' and, which is called **poly 'A' tail,** Now it becomes stable. 7 methyl guanonsine is also added to its 5' end, and a cap like structure is formed. It is called **capping.** Bu the process of **RNA splicing** hn-RNA produces functional RNA that is exonic RNA. In RNA splicing non coding parts removed with the help of ribonuclease enzyme and coding part join together with the help of RNA ligase. Some specific proteins are also helpful in RNA - splicing called **"Small nuclear ribonucleoprotein " or 'SnRNP' or 'Snurps'.** These snRNP proteins combine with some other proteins and SnRNA and form **spliceosome complex** This spliceosome complex uses energy of ATP to cut the RNA, releases the non-coding part and joins the coding part to produce functional RNA.

Non coding apart of hn RNA remained inside the nucleus and not translated in to protein. Only coding part moves from nucleus to cytoplasm and translated into protein.

Mostly Eukaryotic genes are example of split gene, but gene which forms histone and interferon protein are non split gene. It contains only and only exonic part.

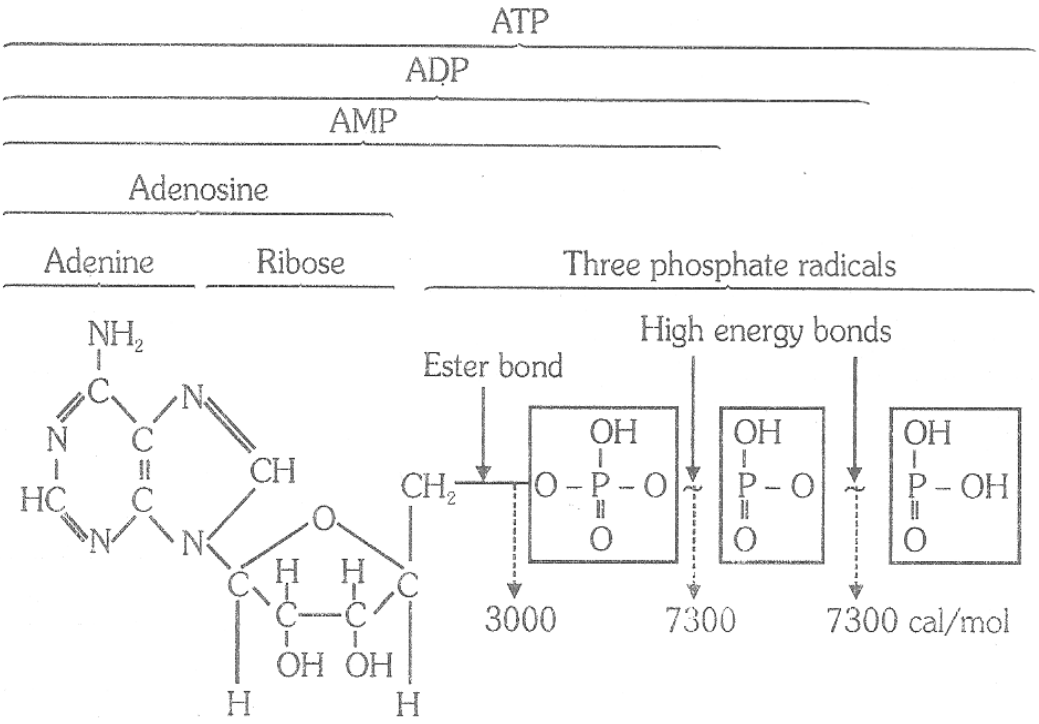
Mostly prokaryotic genes are example of non split gene.

Especial Points :

- In human mitochondria, 4 initiation codons present : AUG, AUA, AUU, AUC.
- **Mic RNA :** It is synthesized sometimes on the sense strand of DNA which is complementary of Antisense strand which is used for mRNA synthesis. Such RNA is used for regulation of gene expression at the level of translation.

- **Higher Nucleotide :** Nucleotides which contain more than one phosphate i.e. ATP, ADP.

  ATP : Discover, Karl Lohmann It is made up by Adenine, D-Ribose and three phosphate. It is high energy compound that release energy when the bond between the phosphate is broken. In ATP two high energy bonds are present. ATP is also called energy currency of cell.



- **Iodine number :** It is the amount of iodine in gram absorbed by 100 gram fat. It is used to determine the degree of unsaturation of fat.
- **Second genetic code :** Interaction between specific t-RNA and amino acyl synthetase enzyme is known as second genetic code.

  GLUT-4 (Glucose transport 4) Proteins : It is a transport protein that allows glucose to enter a cell. GLUT-4 moves in to the plasma membrane in response to insulin or muscle contraction. When amount of glucose increased in blood, more insulin in secreted GLUT-4 then responds to the presence of the insulin and enters the plasma membrane, allowing glucose to enter the cell.
- **DNA-quenching :** Rapid cooling of denatured DNA, fix it in permanently denatured form, it is called DNA quenching.
- ϕ × 174 bacteriophage has 5386 nucleotides λ-bacteriophage has 48502 base pairs, Escherichia coli has 4.6×10^6 base pairs and 6.6×10^9 base pairs in human (2n).
- 2-OH groups present at every nucleotide in RNA as reactive group and makes RNA labile and easily degradable and RNA also has catalytic function so it is more reactive so DNA is chemically less reactive and structurally more stable as compared to RNA.
- An mRNA also have some additional sequences that are not translated and are referred as untranslated regions (UTR). The UTRs are present at both 5' end (before start codon) and at 3'end (after stop codon).