Transcription :- (DNA → m-RNA)

- The process of copying genetic information from one strand of the DNA into RNA is termed as Transcription.
- Here, also, the principle of complementarity governs the process of transcription, except the adenine complements now forms base pair with uracil instead of thymine.

Note In process of replication, the sequence of nitrogenous base gets duplicated when as in transcription only a segment of DNA and only one strand of DNA (template) is copied into RNA.

Why only one strand are not copied during transcription

Ans Either, if both the strands act as template, they would code for two molecules of different sequences, and to form, 'ds' RNA. But, possibly, the sequence of codons code in the transfer would be different. Hence, one segment of the DNA would be coding for two different proteins and this would complicate the genetic information transmission mechanism.

Second, the two RNA molecules if produced would necessarily would be complementary to each other, hence would form a duplex. This would prevent them from being translated into proteins and the matrix of transcription would become a dead end.

Transcription unit :-

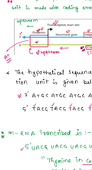
A transcription unit in DNA is defined primarily by the three regions in the DNA

(i) A Promoter
(ii) The structural gene
(iii) A Terminator

- In DNA two strands have opposite polarity and the two operons DNA polymerase also follow the polarities to only one direction, that is 5'→3'. The strand that has polarity 5'→3' act as template, and is also referred to as template strand or antisense strand.

The other strand is known as coding strand and the sequence same as RNA (except uracil in place of thymine) is displayed during transcription. The strand which is not also be coding is referred as coding strand or sense strand.

Note The regions present within defining a transcription unit is made also coding strand.







- The hypothetical sequence from a transcription unit is given below

5' AAGCATGCATGCATGCATGCATGCA' 3' (Template strand)
3' TTCGTACGTACGTACGTACGTACGT' 5' (coding strand)

- m-RNA transcribed is :-

5' UUCG UACG UACG UACG UACG UACG' 3'
"thymine in coding strand is replaced by uracil"

- Promoter :- It is the region on DNA molecule to which DNA Polymerase binds and initiates transcription.

→ Promoter located towards 5' end of coding sequence that is upstream of the structural gene (towards left).

→ In most eukaryotic the promoter has AT rich region called TATA Box (Goldberg-Hogness box).

→ TATA Box is also cis Pribnow Box in Prokaryotes and Hogness box in Eukaryotes.

Besides promoter, eukaryotes also require an enhancer.

→ Promoter is the DNA sequence also defines the Template and coding strand.

→ By indicating the position of the terminator, the direction of coding and template strands could be revealed.

Terminator :- This region of DNA molecule is not transcribed (to the right) of the structural gene. That is towards 3' end of coding strand.

Structural Gene :- It is the area of template strand that is involved in transcription or formation of m-RNA. It is non-sense "transcribable" and "polycistronic".

Carry information for synthesis of one polypeptide chain (mostly in Eukaryotes)

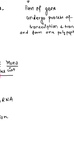
Carry synthesis of more than one polypeptide chain (mostly in Prokaryotes)




- In eukaryotes, the transcribable structural genes have heterogeneous coding sequences - the genes in eukaryotes are split, the coding sequences are expressed separated and distinct as exons. Exons are said to be non-coding sequences that appear in mature or processed RNA. The exons are interrupted by introns. Introns or Intervening sequences do not appear in mature or processed RNA (m-RNA).

The split-gene arrangement further complicates the definition of a gene. In terms of a DNA segment

Substitutions of exons are in also identified by promoter and regulator sequences are loosely defined as regulatory genes even though these sequences do not code for any RNA or protein.




Transcription in Prokaryotes :-

It occurs in cytoplasm with the help of DNA-dependent RNA Polymerase that catalyses the transcription of all the three types of RNA, (eg mRNA, tRNA and rRNA) in bacteria. All three RNA are needed to synthesize a protein in a cell.

(i) Initiation :- RNA polymerase recognizes the promoter region and binds to it. RNA polymerase associates transiently to sigma (σ) factor (Initiation factor). The enzyme recognizes the promoter by the σ-factor. RNA polymerase initiates transcription. It uses nucleotide triphosphates as substrate & polymerizes it in a template dependent fashion following the rule of complementarity. It also facilitates the opening of duplex.

(ii) Elongation :- The RNA polymerase, a core enzyme is only capable of process of elongation in presence of energy and $Mg^{2+}$ ion. Enzyme moves along DNA template, RNA chain becomes longer. Synthesis of RNA continues till the enzyme reaches the terminator.

(iii) Termination :- when the RNA polymerase reaches the terminator region, RNA polymerase transiently bind with Rho (ρ)-factor (Termination factor) which stops the synthesis of RNA chain. It separates RNA polymerase as well as newly formed RNA. The RNA & RNA Polymerase fall off and it results in the termination of transcription.

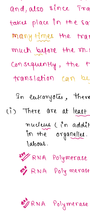
Initiation
Promoter
σ
DNA binding
DNA polymerase
Sigma factor (initiation factor)
Elongation
Elongation
Termination
Termination
RNA
RNA Polymerase
Rho factor
(Termination factor)

Note In Bacteria, since m-RNA does not require any processing to become active, and also since transcription & translation take place in the same compartment, many times the translation can begin much before the m-RNA is fully transcribed consequently, the transcription and translation can be coupled in bacteria.

In eukaryotes, there are two additional complexities.

(i) There are at least three RNA Polymerases in the nucleus (in addition to the RNA polymerase found in the organelles). These to clear cut division of labour.

RNA Polymerase I → rRNA (28S, 18S, 5.8S)
RNA Poly merase II → precursor of m-RNA, the heterogeneous nuclear RNA (hn-RNA)
RNA Polymerase III → t-RNA, 5S rRNA & SnRNA (Small nuclear RNA)




(ii) In eukaryotes, the nascent (new formed) RNA transcribed by RNA polymerase II is called hnRNA or Primary transcript. The primary transcript contains both the exons and the introns and are non-functional. Hence it is subjected to a process of splicing where the introns are removed and exons are joined to a defined order.

The primary transcript is converted into functional m-RNA after post-transcriptional processing which involves three steps :-

(i) Capping at 5' end :- Capping at 5' end of hnRNA occurs shortly after the start of transcription. In capping unusual nucleotide (Modified Guanine Nucleotide - 7-Methylguanosine triphosphate) is methylated to produce m-RNA from degradation by nucleases. It also provides efficient binding site.

(ii) Tailing at 3' end (Polyadenylation) :- This is addition of a poly A tail at 3' end of hnRNA. It contains one-two adenylate residues. It prevents 3' end from degradation by nucleases.

(iii) Splicing :- Introns are removed with the help of Sn-RNA (Pronounced as snurps). "Small nuclear ribonucleoprotein" The ends of the adjacent exons are joined together by ligase enzyme. It is the final step of processing.

Third hn-RNA, now called m-RNA is transported out to the cytoplasm for translation.

Note Split-gene arrangements represent probably an ancient feature of the genome. The process of introns is vestigial of evolutionary and the process of splicing represents the dominance of RNA world.

Splicing you of knowing