COURSE: MSc Part -II

PAPER – XI

TOPIC- Molecular Biology (4)

PREAPARED BY: Ajai Kishore Sharan

Topic- 8

Transcription and Translation in Eukaryotes

Transcription is the process by which the information in a strand of DNA is copied into a new molecule of RNA. It is the first step of gene expression, in which a particular segment of DNA is copied into RNA (especially mRNA) by the enzyme RNA polymerase. It results in a complementary, antiparallel RNA strand called a primary transcript.

Transcription in Eukaryotes

Transcription occurs in eukaryotes in a way that is similar to prokaryotes with reference to the basic steps involved. However, some major differences between them include:

Transcription is carried out by three enzymes (RNA polymerases I, II and III). The regulation of transcription is more extensive than prokaryotes.

Eukaryotic Transcription

Enzyme(s) Involved in Eukaryotic Transcription

Unlike prokaryotes where all RNA is synthesized by a single RNA polymerase, the nucleus of a eukaryotic cell has three RNA polymerases responsible for transcribing different types of RNA.

a) RNA polymerase I (RNA Pol I) is located in the nucleolus and transcribes the 28S, 18S, and 5.8S rRNA genes.

b) RNA polymerase II (RNA Pol II) is located in the nucleoplasm and transcribes protein-coding genes, to yield pre-mRNA, and also the genes encoding small nucleolar RNAs (snRNAs) involved in rRNA processing and small nuclear RNAs (snRNAs) involved in mRNA processing, except for U6 snRNA.

c) RNA polymerase III (RNA Pol III) is also located in the nucleoplasm. It transcribes the genes for tRNA, 5S rRNA, U6 snRNA, and the 7S RNA associated with the signal recognition particle (SRP) involved in the translocation of proteins across the endoplasmic reticulum membrane. Each of the three eukaryotic RNA polymerases contains 12 or more subunits and so these are large complex enzymes.

Features of Eukaryotic Transcription

Transcription in eukaryotes occurs within the nucleus and mRNA moves out of the nucleus into the cytoplasm for translation. The initiation of RNA synthesis by RNA polymerase is directed by the presence of a promoter site on the 5’ side of the transcriptional start site. The RNA polymerase transcribes one strand, the antisense (-) strand, of the DNA template. RNA synthesis does not require a primer. RNA synthesis occurs in the 5’ → 3’ direction with the RNA polymerase catalyzing a nucleophilic attack by the 3-OH of the growing RNA chain on the alpha-phosphorus atom on an incoming ribonucleoside 5-triphosphate.mRNA in eukaryotes is processed from the primary RNA transcript, a process called maturation.

RNA processing

The primary eukaryotic mRNA transcript is much longer and localised into the nucleus, when it is also called heterogenous nuclear RNA (hnRNA) or pre- mRNA.It undergoes various processing steps to change into a mature RNA:

Cleavage

Larger RNA precursors are cleaved to form smaller RNAs.Primary transcript is cleaved by ribonuclease-P (an RNA enzyme) to form 5-7 tRNA precursors.

Capping and Tailing

Initially at the 5′ end a cap (consisting of 7-methyl guanosine or 7 mG) and a tail of poly A at the 3′ end is added. The cap is a chemically modified molecule of guanosine triphosphate (GTP).

Splicing

The eukaryotic primary mRNAs are made up of two types of segments; non-coding introns and the coding exons.The introns are removed by a process called RNA splicing where ATP is used to cut the RNA, releasing the introns and joining two adjacent exons to produce mature mRNA.

RNA processing

The primary eukaryotic mRNA transcript is much longer and localised into the nucleus, when it is also called heterogenous nuclear RNA (hnRNA) or pre- mRNA.It undergoes various processing steps to change into a mature RNA:

Cleavage

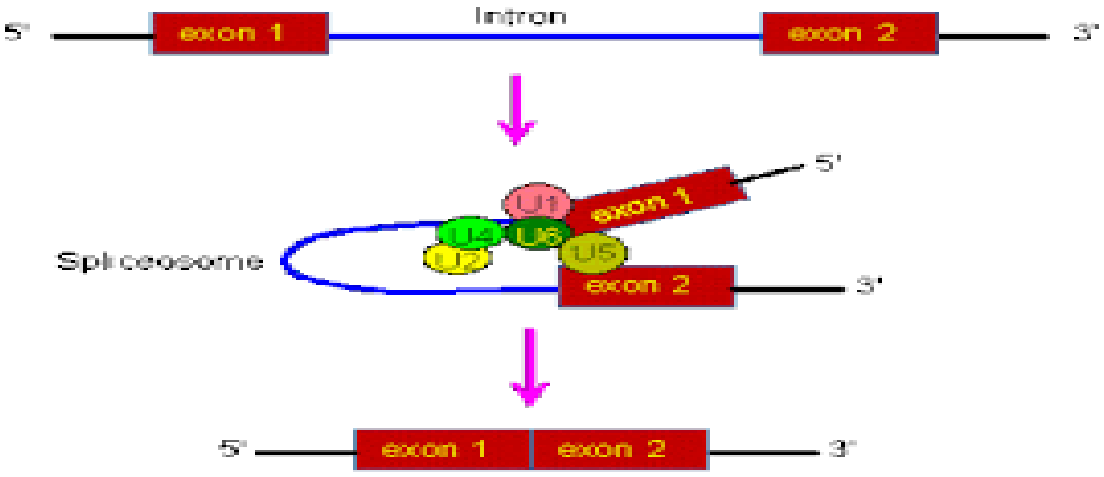
Larger RNA precursors are cleaved to form smaller RNAs.Primary transcript is cleaved by ribonuclease-P (an RNA enzyme) to form 5-7 tRNA precursors.

Capping and Tailing

Initially at the 5′ end a cap (consisting of 7-methyl guanosine or 7 mG) and a tail of poly A at the 3′ end are added. The cap is a chemically modified molecule of guanosine triphosphate (GTP).

Splicing

The eukaryotic primary mRNAs are made up of two types of segments; non-coding introns and the coding exons.The introns are removed by a process called RNA splicing where ATP is used to cut the RNA, releasing the introns and joining two adjacent exons to produce mature mRNA.



Process of Eukaryotic Transcription

The basic mechanism of RNA synthesis by these eukaryotic RNA polymerases can be divided into the following phases:

Initiation Phase

Initiation Phase of Eukaryotic Transcription

During initiation, RNA polymerase recognizes a specific site on the DNA, upstream from the gene that will be transcribed, called a promoter site and then unwinds the DNA locally.Most promoter sites for RNA polymerase II include a highly conserved sequence located about 25–35 bp upstream (i.e. to the 5 side) of the start site which has the consensus TATA(A/T)A(A/T) and is called the TATA box.Since the start site is denoted as position +1, the TATA box position is said to be located at about position -25.The TATA box sequence resembles the -10 sequence in prokaryotes (TATAAT) except that it is located further upstream.Both elements have essentially the same function, namely recognition by the RNA polymerase in order to position the enzyme at the correct location to initiate transcription.The sequence around the TATA box is also important in that it influences the efficiency of initiation. Transcription is also regulated by upstream control elements that lie 5′ to the TATA box.

Some eukaryotic protein-coding genes lack a TATA box and have an initiator element instead, centered around the transcriptional initiation site. In order to initiate transcription, RNA polymerase II requires the assistance of several other proteins or protein complexes, called general (or basal) transcription factors, which must assemble into a complex on the promoter in order for RNA polymerase to bind and start transcription.

These all have the generic name of TFII (for Transcription Factor for RNA polymerase II).

The first event in initiation is the binding of the transcription factor IID (TFIID) protein complex to the TATA box via one its subunits called TBP (TATA box binding protein).As

soon as the TFIID complex has bound, TFIIA binds and stabilizes the TFIID-TATA box interaction. Next, TFIIB binds to TFIID.

However, TFIIB can also bind to RNA polymerase II and so acts as a bridging protein. Thus,RNA polymerase II, which has already complexed with TFIIF, now binds.This is followed by the binding of TFIIE and H. This final protein complex contains at least 40 polypeptides and is called the transcription initiation complex.

Those protein-coding genes that have an initiator element instead of a TATA box appear to need another protein(s) that binds to the initiator element.The other transcription factors then bind to form the transcription initiation complex in a similar manner to that described above for genes possessing a TATA box promoter.

Elongation Phase

TFIIH has two functions:

It is a helicase, which means that it can use ATP to unwind the DNA helix, allowing transcription to begin.In addition, it phosphorylates RNA polymerase II which causes this enzyme to change its conformation and dissociate from other proteins in the initiation complex.The key phosphorylation occurs on a long C-terminal tail called the C-terminal domain (CTD) of the RNA polymerase II molecule.Interestingly, only RNA polymerase II that has a non-phosphorylated CTD can initiate transcription but only an RNA polymerase II with a phosphorylated CTD can elongate RNA.RNA polymerase II now starts moving along the DNA template, synthesizing RNA, that is, the process enters the elongation phase.RNA synthesis occurs in the 5’ → 3’ direction with the RNA polymerase catalyzing a nucleophilic attack by the 3-OH of the growing RNA chain on the alpha-phosphorus atom on an incoming ribonucleoside 5-triphosphate.

The RNA molecule made from a protein-coding gene by RNA polymerase II is called a primary transcript.

Termination Phase of Eukaryotic Transcription

Elongation of the RNA chain continues until termination occurs.

Unlike RNA polymerase in prokaryotes, RNA polymerase II does not terminate transcription at a specific site but rather transcription can stop at varying distances downstream of the gene.

RNA genes transcribed by RNA Polymerse II lack any specific signals or sequences that direct RNA Polymerase II to terminate at specific locations.

RNA Polymerase II can continue to transcribe RNA anywhere from a few bp to thousands of bp past the actual end of the gene.

The transcript is cleaved at an internal site before RNA Polymerase II finishes transcribing. This releases the upstream portion of the transcript, which will serve as the initial RNA prior to further processing (the pre-mRNA in the case of protein-encoding genes.)

This cleavage site is considered the “end” of the gene. The remainder of the transcript is digested by a 5′-exonuclease (called Xrn2 in humans) while it is still being transcribed by the RNA Polymerase II.

When the 5′-exonulease “catches up” to RNA Polymerase II by digesting away all the overhanging RNA, it helps disengage the polymerase from its DNA template strand, finally terminating that round of transcription.

Topic-9

Operon concept in Prokaryotes (Gene expression in Prokaryotes)

When the information stored in our DNA is converted into instructions for making proteins or other molecules, it is called gene expression. Gene expression is a tightly regulated process that allows a cell to respond to its changing environment. It acts as both an on/off switch to control when proteins are made and also a volume control that increases or decreases the amount of proteins made.

There are two types of gene action – constitutive and regulated.

The constitutive gene action is necessary for carrying out vital function of the cell e.g., glycolysis. It does not require repression because it has to continue every moment as long as cell has to survive. Hence, regulator and operator genes are not associated with it. The regulated gene however have a regulatory function and hence requires cooperation of regulator and operator genes. The mechanism of regulation of such genes was first explained by Jacob and Monad and was called as Operon concept.

The Operon Model:

Francois Jacob and Jacques Monod (1961), two French geneticists, discovered while studying bacteria, that the enzymes synthesized by them can be placed in two categories:

(i) Those that are synthesized all the time and occur in relatively constant concentrations, such as enzymes of glycolysis, and

(ii) Those that are synthesized only after a specific stimulation. The first type was named constitutively synthesized and the latter the inducible enzymes.

Operons are regions of DNA that contain clusters of related genes. They are made up of a promoter region, an operator, and multiple related genes. The operator can be located either within the promoter or between the promoter and the genes. RNA polymerase initiates transcription by binding to the promoter region. The location of the operator is important as its regulation either allows or prevents transcription of the genes into mRNA.

Inducible Operon System (Induction of Operon):

1. Inducible operon system is (a) regulated operon system in which the structural genes remain switched off unless and until an inducer is present in the medium.

2. It occurs in catabolic pathways.

3. Lac operon of *Escherichia coli* is an inducible operon system which was discovered by Jacob and Monod (1961).

4. Lac operon of *Escherichia coli* has three structural genes, z, y, and a.

5. In the induced operon the structural genes transcribe a polycistronic mRNA which produces three enzymes. These are β-galactosidase, galactoside permease and galactoside acetylase.

6. β-galactoside brings about hydrolysis of lactose or galactoside to form glucose and galactose.

7. Galactoside permease is required for entry of lactose or galactoside into the bacterium.

8. Galactoside acetylase is a transacetylase which can transfer acetyle group to β-galactoside.

9. The initiation codon of structural gene z is TAG (corresponding to AUG of mRNA) and is located 10 base pairs away from the end of the operator gene.

10. The substance whose addition induces the synthesis of enzyme is called inducer.

11. Inducer is a chemical which attaches to repressor and changes the shape of operator binding site so that repressor no more remains attached to operator.

12. In the lac operon allolactose is the actual inducer while lactose is the apparent (visible) inducer.

13. Inducers which induce enzyme synthesis without getting metabolized are called gratuitous inducers, e.g. IPTG (Isopropyl thiogalactoside).

14. Regulator gene (gene) produces mRNA that synthesises a biochemical repressor.

15. Repressor is a small protein formed by regulator gene which binds to operator gene and blocks structural enzyme thus checking mRNA synthesis.

16. The represseor of lac operon is a tetrameric protein having a molecular weight of 1, 60,000. It is made up of 4 subunits each having molecular weight of 40,000.

17. The repressor protein has two sites, a head for attaching to operator gene and a groove for attachment of inducer.

18. Promoter gene functions as a recognition point for RNA polymerase. RNA polymerase initially binds to this gene. It becomes functional only when it is able to pass over the operator gene and reach structural genes.

19. Operator gene controls the expressibility of the operon. It is normally switched off due to binding of repressor over it.

20. However, if the repressor is withdrawn by the inducer, the gene allows RNA polymerase to pass from promoter gene to structural gene.

21. In lac operon the operator gene is small, 27 base pairs long. The gene is made of palindromic or self-complementary sequences.

22. If lactose is added, the repressor is rendered inactive so that it cannot attach on operator gene and synthesis of mRNA takes place.

23. Transcription is under negative control when lac repressor is inactivated by inducer.

24. Transcription in lac operon is under positive control through cyclic AMP receptor protein (CAP).

25. The catabolite gene activator protein (Cga protein) or cyclic AMP receptor protein (CAP) binds to the Cga site.

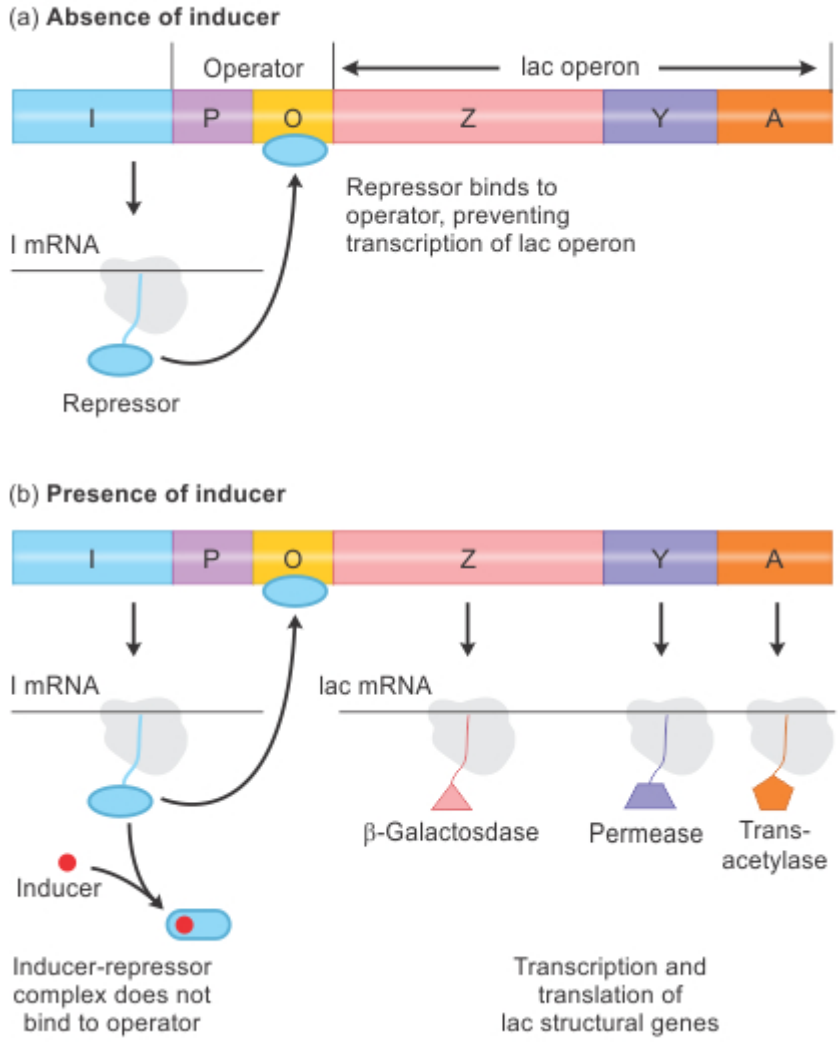
26. When CAP is attached to the binding site the promoter becomes a stronger one.

27. CAP only attaches to the binding site when bound with cAMP.

28. When glucose level is high cAMP does not occur and so CAP does not bind and hence RNA polymerase does not bind, resulting in low transcription.

29. Lac operon will not however remain operative indefinitely despite presence of lactose in the external environment.

30. It will stop its activity with the accumulation of glucose & galactose in the cell beyond the capacity of the bacterium for their metabolism.



Repressible Operon System (Repression of Operon e.g. Tryptophan Operon of *E.coli*):

1. A repressible operon system is a regulated segment of genetic material which normally remains operational but can be switched off when its product is either not required or crosses a threshold value.

2. This system is commonly found in anabolic pathways.

3. Tryptophan operon of *Escherichia coli* is one such repressible operon system.

4. Tryptophan operon has 5 structural gene – E, D, C, and B A.

5. The gene E and D encodes for enzyme anthranilate synthetase, gene C for glycerol phosphate synthetase, gene B for β subunit of tryptophan synthetize and A for α subunit of tryptophan synthetize.

6. Regulator gene (trp-R) produces a biochemical, generally a proteinaceous substance, called aporepressor.

7. Aporepressor alone is unable to block the operator gene because of the absence of the binding head. Therefore, the operon system remains switched on.

8. A complete repressor is formed only when a non-proteinaceous corepressor joins the aporepressor,

9. Corepressor is a non-proteinaceous component or repressor which is also an end product of reaction catalysed by enzymes produced through the activity of structural genes.

10. It (corepressor) combines with aporepressor and forms repressor which then blocks the operator gene to switch off the operon.

11. The structural genes stop transcription and the phenomenon is known as feed-back repression.

12. Corepressor of tryptophan operon is amino acid tryptophan.

13. In tryptophan the repressor gene is not adjacent to promoter but located in another part of E. coli genome.

16. Promoter gene (trp-P) is the recognition as well as initiation point for RNA polymerase. RNA polymerase attaches to promoter gene. It can pass to structural genes provided the operator gene is in the functional state.

17. Operator gene (trp-O) lies in the passage-way between promoter and structural genes. Normally it remains switched on so that RNA polymerase can pass over from promoter gene to structural gene and bring about transcription.

18. The operator gene can be switched off when both aporepressor and corepressor join together to form repressor. The repressor binds to operator gene to interrupt movement of RNA polymerase.

19. In absence of tryptophan, the RNA polymerase binds to the operator site and thus structural genes are transcribed.

20. The transcription of structural gene leads to the production of enzyme (tryptophan synthetize) that synthesizes tryptophan.

21. When tryptophan becomes available, the enzymes for synthesizing tryptophan are not needed, co-repressor (tryptophan) – repressor complex blocks transcription.

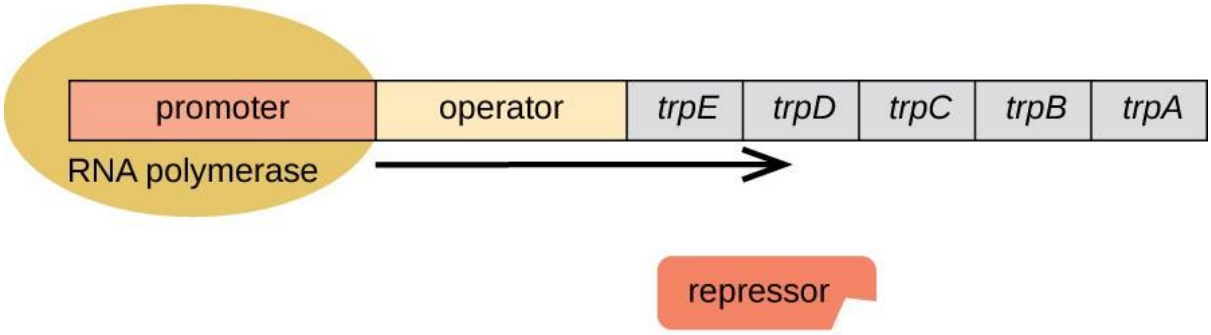
22. One element of tryptophan operon is the leader sequence 'L' that is immediately 5′ end of trp. E gene.

23. This 'L' sequence controls expression of the operon through a process called attenuation.

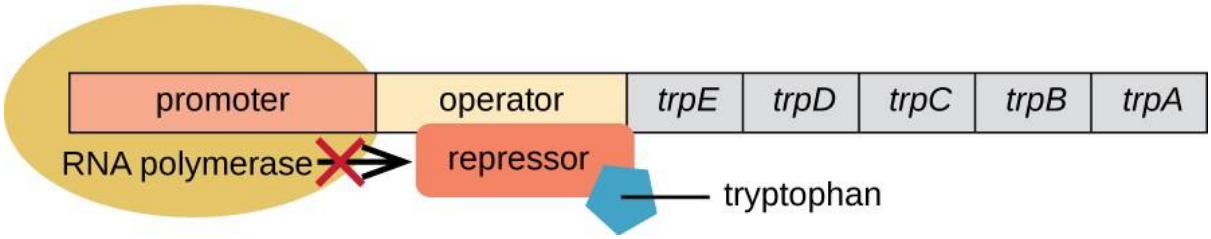
24. Attenuation is the termination of the transcription prematurity at the leader region.

25. The tryptophan operon is a negative control.

In the absence of tryptophan, the *trp* repressor dissociates from the operator, and RNA synthesis proceeds.



When tryptophan is present, the *trp* repressor binds the operator, and RNA synthesis is blocked.



The two operon models described above can be summarized as given below:

(i) Inducible System:

Active Repressor + Operator → System OFF

Active Repressor + Inducer = Inactive Repressor → System ON

(ii) Repressible System:

Apo-repressor and co-repressor complex = Active repressor → System OFF

Apo-repressor = Inactive Repressor → System ON

3. In regulated gene action all the genes required for a multistep reaction can be switched on or off simultaneously.

4. The genes are switched on or off in response to particular chemicals whether required for metabolism or are formed at the end of a metabolic pathway.

5. Gene regulation is required for growth, division and differentiation of cells. It brings about morphogenesis.