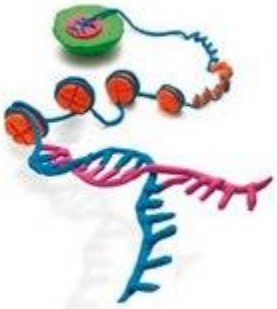
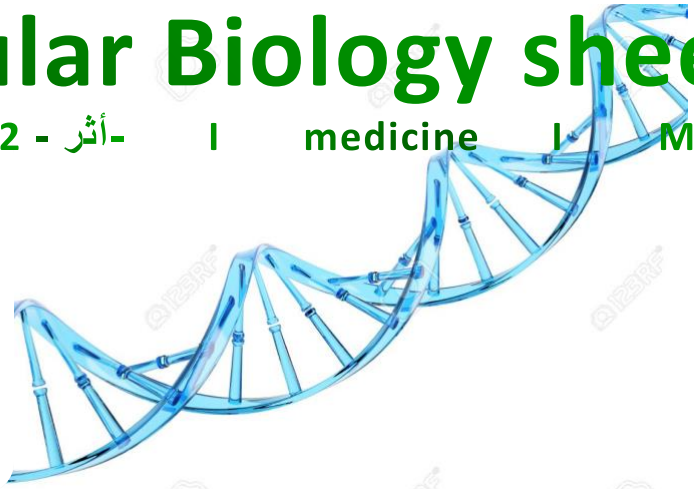
1

Molecular Biology sheet

Doctor 2022 - أثر- I medicine I MU





DOCTOR

Dr. Sameer Mahjoub

DONE BY:

Ibrahem Abumustafa

CORRECTED BY:

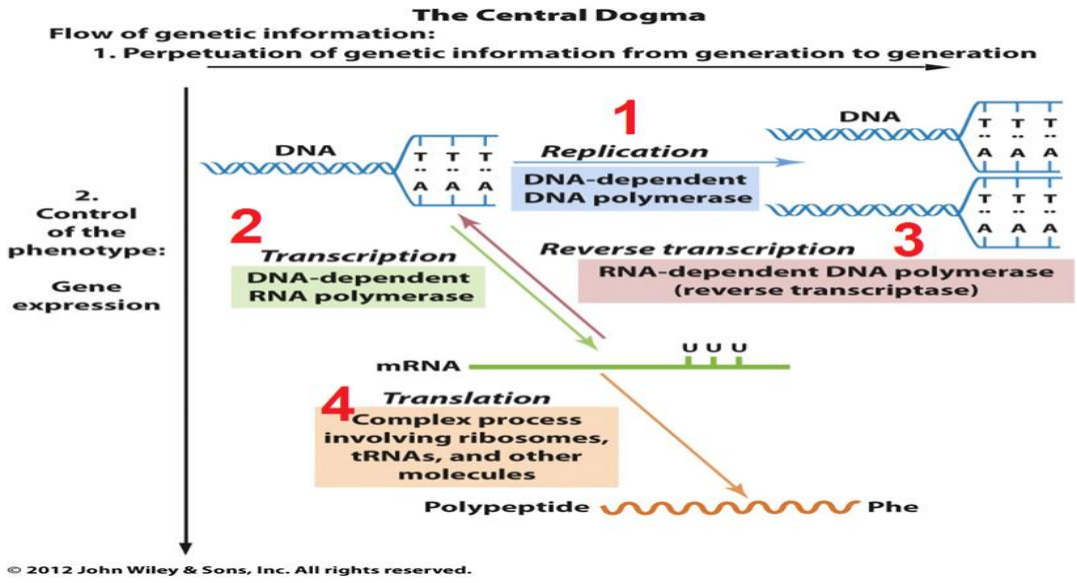
Layan Husam

Transcription
RNA Processing

Kindly note that the sheet is full of repetitions, do not be scared of the pages number 😊

The Central Dogma of life

We are talking about everything that happens in the nucleus.



No synthesis of DNA nor RNA molecules happens without a template to build on. That being said, to build any of them go to the one that it <u>depends</u> on it. So, what processes do we have?

1. **Replication: using DNA- dependent DNA Polymerase (we use DNA template to build DNA via DNA polymerase).**
2. **Transcription: we use DNA template to build RNA using RNA polymerase to produce mRNA (messenger RNA) later on.**
3. **Reverse Transcription: we synthesize DNA from RNA using DNA polymerase, it mostly found in viruses and it is behind the disasters of viral diseases. (we are not going to discuss DNA polymerase in this lecture).**
4. **Translation: it means production of proteins, and it happens <u>in protein synthesizing machinery rather than ribosomes</u>; as ribosomes are not the only ones responsible for proteins production and the proteins it makes are usually not active nor mature yet it requires further modifications that happen in: endoplasmic reticulum (ER), Golgi Apparatus and even sometimes after the proteins are secreted.**

Overview of transcription

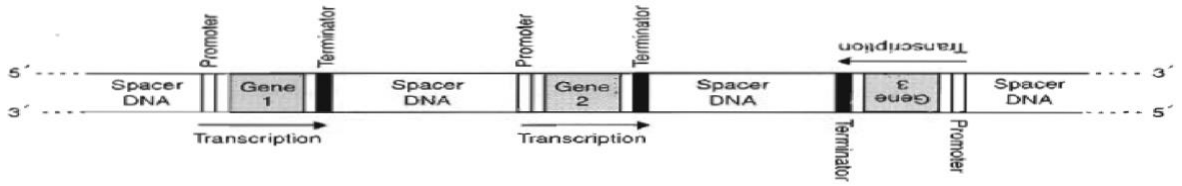
The first stage in the expression of genetic information is transcription of the information in the DNA deoxyribonucleotides sequence into RNA ribonucleotides sequence.

- For any gene, only one strand of the DNA molecule, called the template strand, is transcribed by RNA polymerase. And regardless the polymerase type (RNA or DNA) it always read from 3' (three prime) to 5' direction and builds from 5' to 3' direction.

- Because RNA polymerase moves in the 3' to 5' direction along the template strand of DNA, the RNA product is antiparallel and complementary to the template (1st strand) and identical to the 2nd strand which is called non-template (anti-template) or coding strand.

- RNA polymerase and any enzyme working on RNA are called blind enzymes (can't recognize where to start and where to end), that is why we need marks (signals or signs). RNA polymerase recognizes start signals (promoters) at which the 1st point is called operator and it is considered the transcription initiation site and stop signals (terminators) for each of the thousands of transcription units in the genome of an organism.

To sum up, the enzyme is RNA polymerase, the function is to synthesize RNA from DNA template, it starts from the template strand in the 3' to 5' direction as the one to be synthesized will be from 5' to 3' direction and it is a blind enzyme that does not know where to start and where to end and consequently it needs guiding signals.



Transcription of several genes on a chromosome

Types of RNA

- RNA molecules play a variety of roles in the cell. The 9 types of RNA are:

1- <u>Ribosomal RNA (rRNA),</u> which is the most abundant type of RNA in the cell, due to huge number of ribosomes formed, and rRNA is not related to the status of ribosomes (attached or not). In Eukaryotes we have 4 types of rRNA: 28s, 18s, 5.8s and 5s while in prokaryotes we have 3 types: 23s, 16s and 5s. Where S stands for: 1) Svedberg (the scientist who first discovered it) and 2) Sedimentation rate (so we are not talking about weights but the rates of sedimentation) a possible exam question. We can know the number (#) of rRNA if we knew the # of ribosomes we have (countable).

2- Transfer RNA (tRNA), which is the second most abundant type of RNA. From its name; it transfers (carries) amino acids (AA) from cytoplasm to the site of protein synthesis machinery. If I have a specific number of AA, I know that I at least have 20 types of tRNA (countable) each one of them is responsible for carrying from cytoplasm to the site of protein synthesizing machinary.

3- Messenger RNA (mRNA) the only type of RNA that is translated, which carries the genetic information specifying the amino acid sequence of a protein to the ribosome. The mRNA population in a cell is very(the most one) heterogeneous in size (length) and base sequence (the most important yet the most variable one), as the cell has essentially a different mRNA molecule for each of the thousands of different proteins made by that cell. Can we know the # of mRNA molecules? NO! the # depends on the cell requirements, so we may have endless number of possibilities based on the proteins we want to synthesize (uncountable).

The previous 3 RNA types (rRNA, tRNA and mRNA) are the major functioning ones, the rest of the types help them in a way or another.

4- Heterogeneous nuclear RNA (hnRNA or pre-mRNA or mRNA mother) (the immediate product of gene transcription), which is found only in the nucleus of eukaryotic cells not in prokaryotes and it represents precursors of mRNA, 75% is degraded in the nucleus and 25% only is processed to mature RNA. Processing is required to form mature mRNA. Since processing does not happen in prokaryotes and the synthesized mRNA is directly transcribed, we don't have hnRNA in prokaryotes. Moreover, all the modifications are related to RNA in eukaryotes. 75% of hnRNA is degraded due to possible errors.

5- Small nuclear RNA (snRNA) or Ribozymes (RNA molecules with enzymatic activity), which is also only found in the nucleus of eukaryotes not in prokaryotes (for the same reason in hnRNA), small in size (90- 300 ribonucleotides) and complexed with proteins (forming ribonucleoproteins). One of its major functions is to participate in splicing (removal of introns) mRNA it acts as enzymes for mRNA processing.

6- Small cytoplasmic RNA (scRNA), from its name works in cytoplasm, has catalytic activity in tRNA processing to be converted to mature form and acts as signal recognition particle.

7- Small nucleolar (snoRNA) from its name works in nucleoli, acts in rRNA processing/ maturation/methylation to be converted to mature form.

To sum up, all 3 main types of RNA (tRNA, mRNA and rRNA) have a specific type to convert it from (immature, inactive and non-functioning) form to (mature, active and functioning) one. So, the main type (or its precursor) its processing type are: 1) hnRNA snRNA or ribozymes, 2) tRNA scRNA and 3) rRNA snoRNA.

Now we will talk about the regulatory two types; so that transcription remains controlled by the cell.

8- Micro-RNA newly discovered, short from its name, non-coding, ~ 22 nucleotide long, generated by nucleolytic processing of the products of distinct genes or transcription unites, at least some of which control the expression of other genes during development mRNA will be converted by cutting enzymes to fragments, that will undergo hybridization with non-complimentary sequence of mRNA (for example, G may bind A and U may bind C), (mature micro RNA molecules can hybridize together to form imperfect RNA- RNA duplex within the 3' untranslated regions of specific target mRNA causing unexplained gene expression regulation in at least half 50% of the human genes is regulated by micro RNA). The goal is to do unusual (unexplained) genes to initiate enzymatic systems which will result in degrading the newly synthesized mRNA. Even though it is newly discovered, it has a great importance in diseases diagnoses; many micro-RNA biomarkers are found in some cancers and heart diseases (as in angina pectoris).

9- Small interfering RNA (siRNA) are derived by specific nucleolytic cleavage of larger double stranded RNAs to form small 21-25 long products. They form perfect RNA-RNA hybrids (it has to be complimentary to one another, the complementarity we are talking about is between the fragments of siRNA and mRNA) with their targets anywhere within the length RNA where the complementary sequence exists resulting in reduction of specific protein production (despite its perfectness, it is also considered as an abnormality that will initiate degradation pathways) because siRNA-mRNA complexes are degraded by nucleolytic machinery (interferes with the expression of specific gene by hybridizing to its corresponding RNA sequence in the target mRNA, then activates degradation of mRNA which can not be translated into proteins).

To sum up, anything unusual (hybridization with microRNA or siRNA) within the nucleus -wither its perfect or not- has to be degraded.

Transcription: important concepts and terminology

- RNA polymerase locates genes in DNA by searching for promoter regions in DNA template.

- The promoter is the binding site for transcription factors and RNA polymerase.

- Binding establishes where transcription begins, which strand of DNA is used as the template, and in which direction transcription proceeds.

- RNA polymerase moves along the template strand in the 3' to 5' direction as it synthesizes the RNA product in the 5' to 3' direction using NTPs (ATP: adenosine triphosphate, GTP: Guanosine triphosphate, CTP: Cytidine triphosphate, UTP: uridine triphosphate) as substrates for energy.

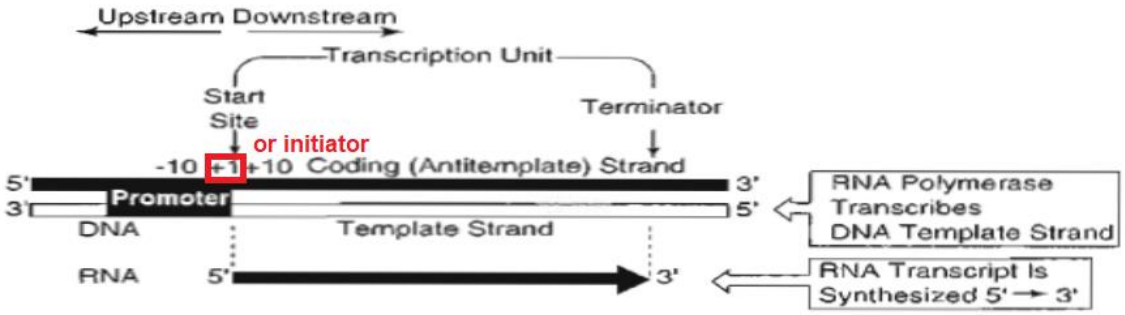
- RNA polymerase does not proofread its work.

- The RNA product is complementary and antiparallel to the template strand.

- The coding (non-template) identical strand is not used during transcription. It is identical in sequence to the RNA molecule, except that RNA contains uracil instead of the thymine found in DNA.

- By convention, the base sequence of a gene is given from the coding strand (5' → 3') while it moves (reads) from 3' 5'

- Transcription ends when RNA polymerase reaches a termination signal.



Transcription of DNA

Prokaryotic RNA Polymerases

- There is a single prokaryotic RNA polymerase that synthesizes all types of RNA in the cell. Only one with 4 subunits (tetramer)

- The core polymerase has the subunit structure (α2ββ`).

- A protein factor called sigma (σ) is required for the initiation of transcription at the promoter. Consider RNA polymerase a blind person and sigma factor is the guide who leads them.

- σ factor is released immediately after transcription initiation, as it is no longer needed.

Functions of the subunits: 1) α (alpha): assembly of the tetrameric core (the 4 subunits) 2) β (Beta): ribonucleoside triphosphate binding site (link ribonucleotides together) 3) β'(Beta Dash): DNA template binding region (recognition site) and 4) σ (sigma factor): initiation of transcription (which is considered the most important one, as it RNA polymerase is a blind enzyme, so it needs this factor to lead it to the transcription starting site.

- Termination of transcription sometimes requires a protein called rho (ρ) factor. It shows RNA polymerase the ending site.

- This enzyme is inhibited by rifampin also known as rifampicin and actinomycin D both are antibiotics they work on prokaryotes like fungi, parasites, bacteria and any other diseases causative agents.

RNA polymerase in prokaryotes is considered an apoenzyme (it needs a cofactor - which is sigma in this case- to be in the active holoenzyme state).

Promoter "Strength" (activity)

- Affects amount of RNA made, so, it affects level of expression for that gene.

- Not all promoters have same "strength" it means the number of initiations to take place in one second. It varies depending on the requirements of the specific cell, accordingly, Promoters differ in DNA sequences and "strength"

- RNA polymerase binds differently to different sequences "Strong promoters" initiate transcription more often than "weak promoters"

- rRNA has strong promoter: ~1 initiation per second while lacZ has a weak promoter: ~1 initiation per minute (60 seconds). LacZ is responsible for producing an enzyme for regulating gene expression.

A pro-tip from Professor. Samir: It has been scientifically proven that studying while fasting is better for memorizing; firstly, due to the distractive effects of sugars on the brain in the fed status and secondly, because the stress of fasting on our bodies elevates noradrenaline (norepinephrine) levels along with adrenaline (epinephrine) one hormone for memorizing and heightened focus. To relate more, it is similar to what happens when you study something at dawn (AlFajr) you remember it better.

Eukaryotic RNA polymerases:

- Three types which can be distinguished by the particular types of RNA they produce:

1- RNA polymerase I (1) is located in the nucleolus and synthesizes 28S, 18S, and 5.8S rRNAs, 3 out of 4 rRNAs. ONLY for rRNA. Remember that s here stands for: 1) Svedberg and 2) Sedimentation rate.

2- RNA polymerase II (2) is located in the nucleoplasm and synthesizes hnRNA to be processed into mRNA and some snRNA (so, snRNA is not one type only!)

3- RNA polymerase III (3) is located in the nucleoplasm and synthesizes tRNA, some snRNA (other than those produced by #2), and 5S rRNA (the last one of the 4 types).

To sum up, SnRNA is synthesized by 2 RNA polymerases (#2 and #3), mRNA by only 1 which is #2 and rRNA by 2 (#1 and #3).

- Transcription factors (such as TFIID for RNA polymerase II) help to initiate transcription.

- The requirements for termination of transcription in eukaryotes are not well understood. - In addition, RNA polymerase II is inhibited by (α amanitin) a toxin from certain mushrooms do not eat wild mushrooms. It inactivates RNA pol II completely and can kill a person, while, RNA pol I and III are less affected by toxin all in eukaryotes, in prokaryotes the inhibitors are rifampin (rifampicin) and actinomycin D antibiotics.

Prokaryotic	Eukaryotic
Single RNA polymerase ($\alpha_2\beta\beta'$)	RNAP 1: rRNA (nucleolus), except 5S rRNA RNAP 2: hnRNA/mRNA and some snRNA RNAP 3: tRNA, 5S rRNA
Requires sigma (σ) to initiate at a promoter	No sigma, but transcription factors (TFIID) bind before RNA polymerase
Sometimes requires rho (ρ) to terminate	No rho required
Inhibited by rifampin → Actinomycin D	RNAP 2 inhibited by α-amanitin (mushrooms) Actinomycin D

site of initiation

Comparison of eukaryotic and prokaryotic RNA polymerases

We know that sigma factor helps in recognizing the starting site (the promoter), but how exactly? What are the sings? BOXES!

- The following events occur during the expression of a prokaryotic gene:

- With the help of sigma factor, RNA polymerase recognizes and binds to the promoter, region.

- The bacterial promoter contains two "consensus" sequences, the 1st is called the Pribnow box [TATA (TATTAT) box] which is 10 ribonucleotides away from the transcription initiation site and recognizes 6 of them, what is was missed? We go the

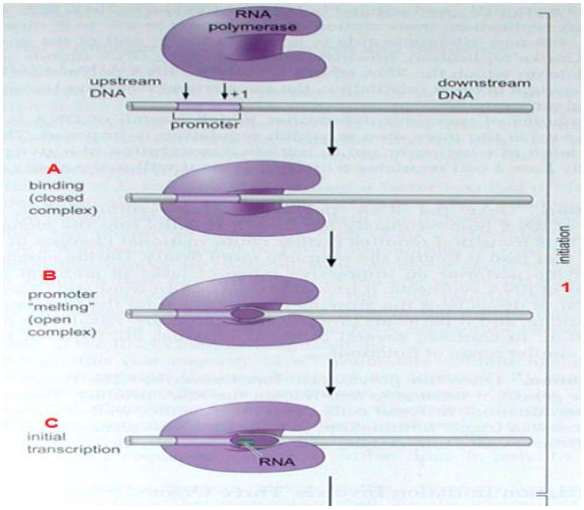
to the 2nd one which is a -35 base pairs away and has the sequence (TGTTGACA) of 8. Note that the following 6 points are connected.

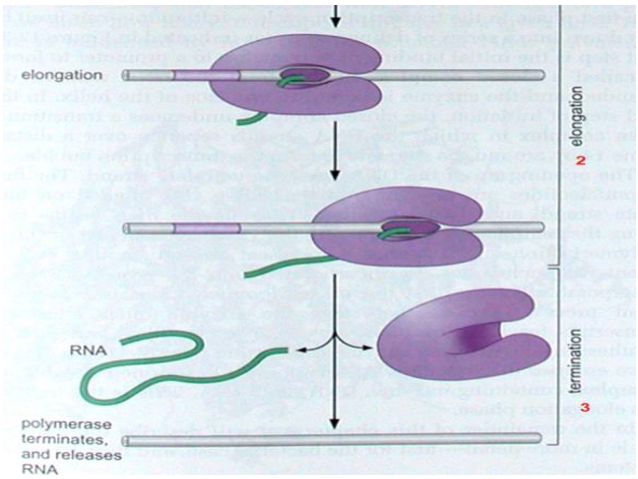
1- The promoter identifies the start site for transcription and orients the enzyme on the template strand.

2- Transcription begins at the + 1 base pair. Sigma factor is released as soon as transcription is initiated.

3- The core polymerase continues moving reading along the template strand in the 3' to 5' direction, synthesizing the mRNA in the 5' to 3' direction.

The 3 processes of transcription are: 1) initiation (with 3 substages↓), 2) elongation or extension and 3) termination. RNA polymerases -just like DNA polymerases in replication- can not act on a double stranded DNA directly. Accordingly, helicase enzyme will also work here for separating the double strands -however, it is only for a limited region in the transcription that is 1 gene- and unwinding the two strands (prevents them from forming a helix structure again) which ensures an access of RNA polymerase to the space in-between the two strands to start transcribing the template strand to be identical to the non-template strand. Finally, the RNA is released.



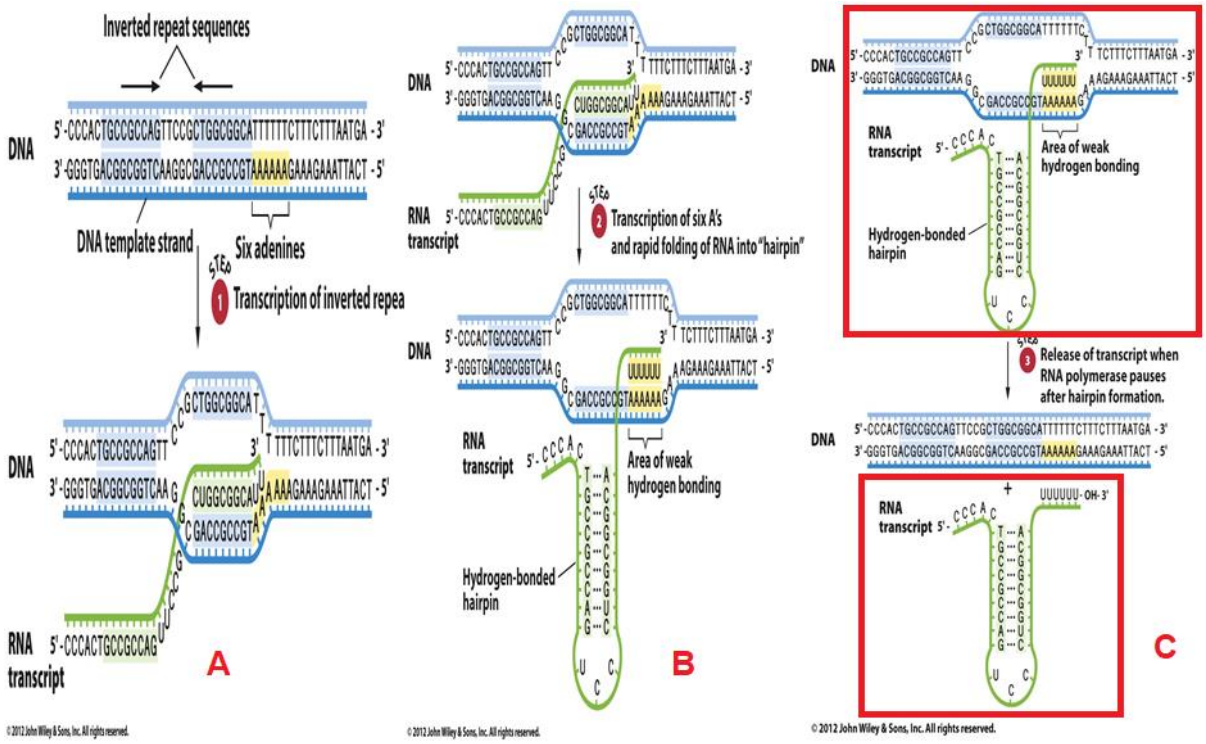


To sum up, we started at the two boxes; TATTAT box and TGTTGACA -35 nucleotides away- box. They were recognized by sigma factor who held the blind RNA polymerase hand and took it to the starting point where it started transcribing in a relatively massive speed until it reaches a specific termination site where it will be stopped.

4-RNA polymerase eventually reaches a transcription termination signal, at which point it will stop transcription and release the completed mRNA molecule. There are two kinds of transcription terminators commonly found in prokaryotic genes:

A- <u>Rho-dependent termination</u> **requires participation of rho factor (a protein at the end of the gene near the termination site). This protein binds to the newly formed RNA where it moves toward the site of RNA polymerase that has paused at a termination site. Rho then displaces RNA polymerase from the 3' end of the RNA. It means that RNA polymerase can read DNA sequence of the template strand from a distance.**

B- <u>Rho-independent termination</u> **(figure C below) occurs when the newly formed RNA folds back on itself to form a GC-rich hairpin loop (where 3 strong hydrogen bonds [H bonds] form between them) closely followed by 6-8 U residues forming weak H bonds where a mechanical stress will be exerted on the mRNA -due to the strong vs weak bonds- which is still base-pairing with the template DNA. These two structural features (and the resulting strong/ weak H bonds) of the newly synthesized RNA promote dissociation of the RNA from DNA template (when the weak H bonds break).**



Transcription Termination (Important)

To sum up, termination occurs either by going to the RNA polymerase where it dissociates itself from the template (rho dependent) OR when RNA folds back on itself forming GC rich regions with strong H Bonds to exert a mechanical stress on the H Bonds between the newly synthesized mRNA and DNA template.

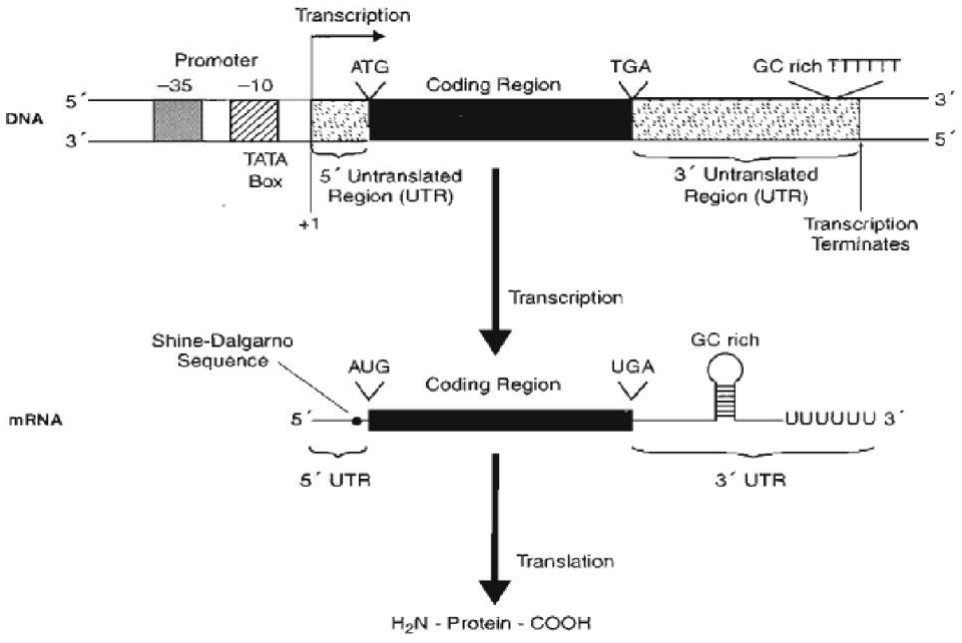
5. Transcription and translation can occur simultaneously in bacteria because there is no processing of prokaryotic mRNA (generally no introns) and since there is no nuclear membrane to separate the two processes from each other (both occur in the cytoplasm), once transcription is completed (mRNA is formed) it will directly bind ribosomes where it can begin translating the message and sometimes even before transcription is complete.

- Ribosomes bind to a sequence called Shine-Dalgarno sequence in the 5' untranslated region (UTR) of the message where the translation process begins and it can be identified by the ribosomes: it is in the 5' end of the mRNA that will make the amino group terminus of the to-be synthesized protein (the 3' terminus is the carboxylic acid group).

- Protein synthesis begins at an AUG codon at the beginning of the coding region and continues until the ribosome reaches a stop codon at the end of the coding region.

6. The ribosome translates the message in the 5' to 3' direction, synthesizing the protein from amino terminus to carboxyl terminus.

A

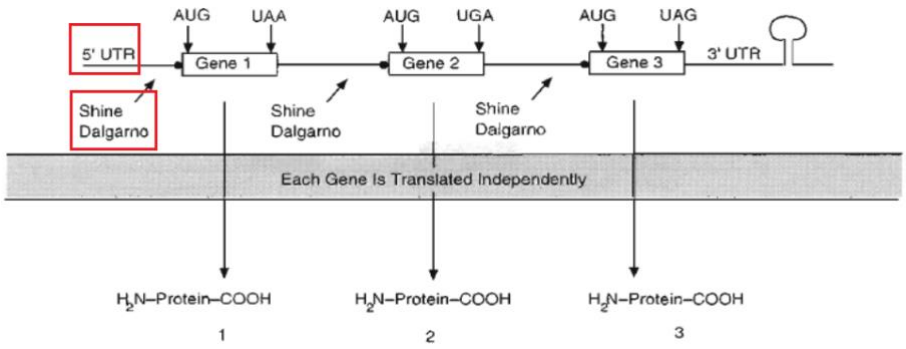


prokaryotic transcription unit.

Important slide:

- The mRNA produced by the gene shown above is a monocistronic message (monocistronic mRNA). That is, it is transcribed from a single gene and codes for only a single protein. This is the case in MOST eukaryotes; as each gene is separated from one another.

- The word cistron is another name for a gene. Some bacterial operons (clusters of co-regulated genes with related functions) produce polycistronic messages (polycistronic mRNAs). In these cases, related genes grouped together in the DNA are transcribed as one unit. Genes in prokaryotes are found in clusters polycistronic mRNA (in prokaryotes and viruses). If a bacterial cell is to metabolize a substance – let it be glucose- and it requires a number of enzymes – let it be 4-, upon transcription, 1 mRNA molecule will be formed carrying the mRNA of the 4 genes related to those enzymes. Simply put, polycistronic mRNA contains mRNA for multiple genes not only one, then by peptidases, it will be cleaved into 4 different mRNAs. The mRNA in this case contains information from several genes and codes for several different proteins



Prokaryotic polycistronic message codes for several different proteins

Here are 3 genes, upon transcription it will form 1 mRNA molecule then by peptidases it will form 3 different enzymes.

Production of eukaryotic mRNA

RNA is not processed (modified) in prokaryotes contrary to eukaryotes. Why?

- In eukaryotes, most genes are composed of coding segments (exons) translated, and it is found in both eukaryotes and prokaryotes interrupted by noncoding segments (introns) not translated, found only in eukaryotes. If we have an mRNA with linked exons and introns, ribosomes that read the genetic codons in the exons won't be able to read the introns intermittent protein (a huge problem) that is why modifications (processing, particularly introns splicing [removal]) occur to eukaryotic RNA (mRNA included); to form a continuous mRNA to be translated to a functioning protein. In prokaryotes no need for the modifications as there are only exons can be read continuously.

- Both exons and introns are transcribed in the nucleus.

- Introns are removed during processing of the RNA molecule in the nucleus.

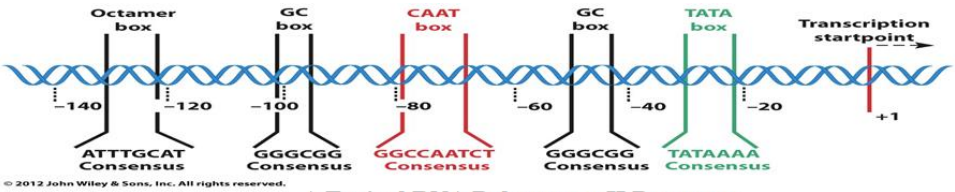
- In eukaryotes, all mRNA is monocistronic.

- The mature mRNA is translated in the cytoplasm.

- Transcription of a typical eukaryotic gene occurs as follows:

1. With the help of proteins called transcription factors (TFs) these factors are alternatives to sigma factor in prokaryotes (no sigma factor in eukaryotes), TF+ RNA polymerase= BRE initiation complex. So, TF are guides that take the blind RNA polymerase hand to the initiation site in eukaryotes. RNA polymerase II recognizes and binds to the promoter region. The basal promoter region of eukaryotic genes usually has two consensus sequences called the 1) TATA box (also called Hogness box) which is 25 base pairs away from transcription initiation site and 2) the CAAT box which is 70- 90 ribonucleotides away a 3rd) GC Box is sometimes found yet it does not have a specific position: before CAAT, after TATA or in-between (position independent). Remember, in prokaryotes the promoter contains 1) TATTAT box and 2) TGTTGACA box. Note that the sequence of any promoter does not have to be alone as long as it is found. For example: TATAAAA, TATAT and TATA are all the same.

2. RNA polymerase II separates the strands of the DNA over a short region to initiate transcription and read the DNA sequence. The template strand is read in the 3' to 5' direction as the RNA product (the primary transcript) is synthesized in the 5' to 3' direction.



A Typical RNA Polymerase II Promoter

Class II promoters (most similar to bacterial promoters)

- Common type of promoter (most genes use this)

- Many variations, but "consensus" (generally agreed that it) has a 1)"Core" (minimal portion of the promoter required to properly initiate transcription) + 2) "Upstream Core" ± 3) "Downstream".

- Core (3 elements): 1. "TATA box" (5'-TATA-3') 25 base pairs away 2. TFIIB (Transcription Factor 2 B) recognition element (BRE) (BRE is the element) not TF2B 3. Initiator box (Inr) with an "A" adenine at +1, most common

- Downstream promoter element (DPE, less common) yes, we have sings here too but not that important

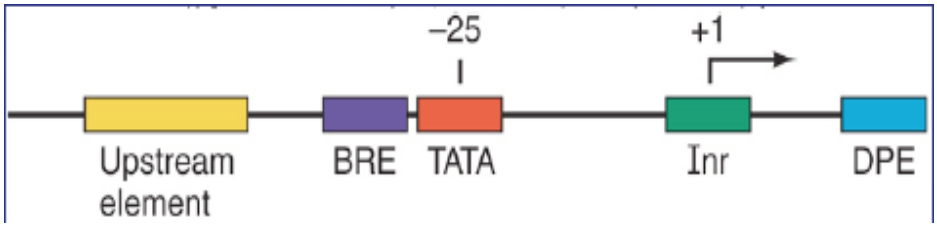
- Core promoter is recognized by general TFs that associate with RNA pol (polymerase) to form a pre-initiation complex at great majority of promoters

- At least one of these elements is missing in most promoters e.g., highly expressed specialized genes (active produce so many mRNAs and should be controlled all 3 elements must be there), however, some of them may even tend to have TATA boxes

(more than 1 box), but promoters for housekeeping genes (a weaker genes with smaller amounts of products) tend to lack them

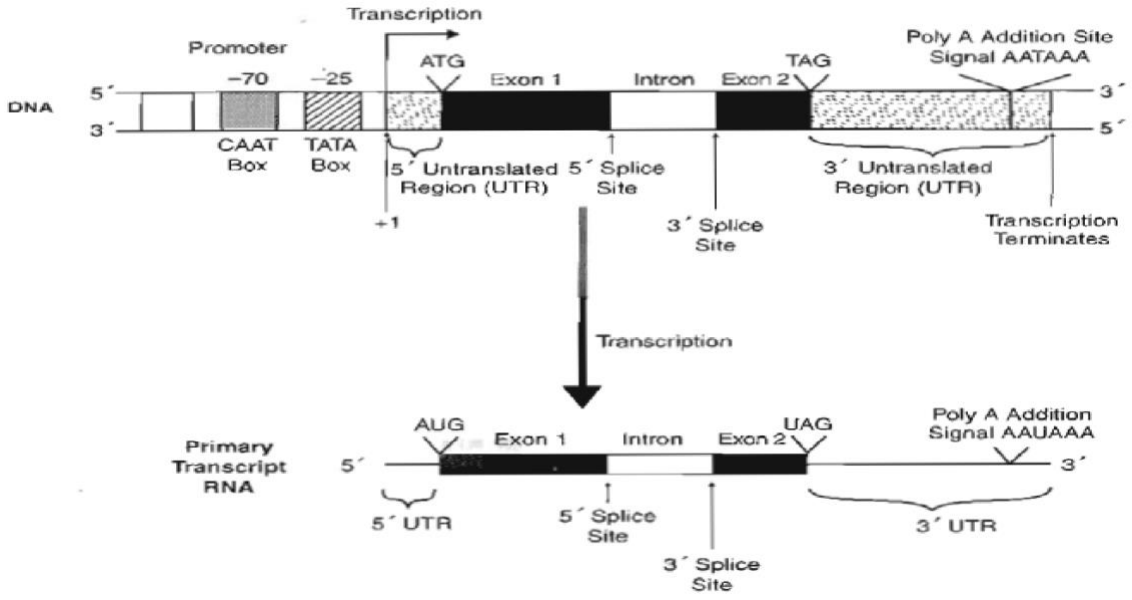
Now comes the most important part:

- Upstream core elements include: TATA box, GC box, CAAT box, TF2B and Inr box. Those upstream elements: quite varied in number and can be 1) orientation-independent (GC box may come after CAAT, before TATA or in-between, but 2) relatively position-dependent (in the upstream region [upstream region means towards 5' while downstream region means towards 3']) & recognized by other TFs 3)relatively gene-specific (highly expressed specialized gene may have more than one of a certain box [TATA or CAAT for example] while house-keeping genes lack many) that participate in initiation at smaller sub-sets of promoters. 1. GC box (GC rich)2. CAAT box (5'-CCAAT-3). RNA polymerase II ends transcription when it reaches a termination signal. These signals are not well understood in eukaryotes. It might be TF (transcripton factors)guided while in prokaryotes is rho dependent vs independent terminations.



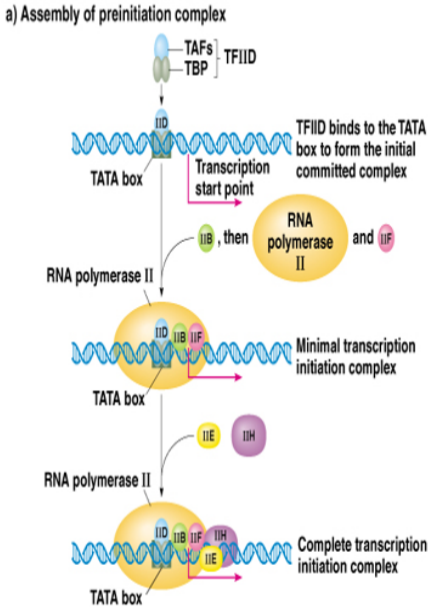
TFIIB recognition element; Inr: initiator box; DPE:

downstream promoter element



A eukaryotic transcription unit

Order of binding is: IID + IIA + IIB + RNA poly. II + IIF +IIE +IIH



1- TBP in TFIID binds to the TATA box
2- TFIIA and TFIIB are recruited with TFIIB binding to the BRE
3- RNA Pol II-TFIIF complex is then recruited
4- TFIIE and TFIIH then bind upstream of Pol II to form the pre-initiation complex
- Promoter melting using energy from ATP hydrolysis by TFIIH)
- Promoter escapes after the phosphorylation of the C-terminal domain tail

The doctor did not give any direct comments regarding this page's figures.

The following table is very important: (numbers are just to link the notes the doctor said/ read)

Table 11.3

General Transcription Initiation Factors

Factor	Subunits ?	Size (kDa) ?	Function
1 TFIID-TBP (TATA Binding Protein)	1	27	TATA box recognition, positioning of TATA box DNA around TFIIB and Pol II
2 TFIID-TAF_{II}s 2	14	15–250	Core promoter recognition (non-TATA elements), positive and negative regulation
3 TFIIA	3	12, 19, 35	Stabilization of TBP binding; stabilization of TAF–DNA binding
4 TFIIB	1	38	Recruitment of Pol II and TFIIF; start-site recognition for Pol II
5 TFIIF	3	156 total	Promoter targeting of Pol II
6 TFIIE	2	92 total	TFIIH recruitment; modulation of TFIIH helicase ATPase, and kinase activities; promoter melting
7 TFIIH	9	525 total	Promoter melting; promoter clearance via phosphorylation of CTD

© Brooks/Cole, Cengage Learning

1. **TFIID- TBP: for TATA box recognition on promoter and it also puts TATA box around both TFIIB & polymerase RNA (positioning) for stabilization**
2. **TFIID-TAF_{II}s: more powerful than #1 as it does 2 important roles: 1st a core promoter recognition of non- TATA elements (CAAT and GC boxes) 2nd regulation of TF: positively (increases it) and negatively (decreases it).**
3. **TFIIA: stabilization of both #1 and #2.**
4. **TFIIB: recruitment of RNA polymerase II and start site recognition for RNA polymerase II.**
5. **TFIIF: promoter targeting of RNA polymerase II it helps in making RNA polymerase II stay inside the promoter (not to contact DNA from outside nor bind any other TFs).**
6. **TFIIE: #7 recruitment, modulation of #7 helicase ATPase and kinase (phosphorylation) activities; promoter melting**
7. **TFIIH: Promoter melting (supports helicase enzyme in separating the two strands to give access to RNA polymerase to start transcription), promoter clearance via phosphorylation of CTD (moving the RNA polymerase from the transcription initiation site along the DNA template by phosphorylation of serine AA @ the carboxylic end of the enzyme. Once it is phosphorylated it moves along the DNA template. It is considered the last factor.**

   Regarding the Subunits and Size (kDa): a question might come about the largest/ smallest, those of the same size/subunits, the one with the least/most # of subunits ..etc (and the doctor did it before). However, "since I stated the question it won't come, inshalla", the doctor said.

   To sum up,

   One TF recognizes TATA box #1 TFIID-TBP

   One TF recognizes the non-TATA elements #2 TFIID-TAF_{II}s

   One TF recognizes core promoter elements #2 TFIID-TAF_{II}s

   One TF stabilizes TFIID-TBP and TFIID-TAF_{II}s #3 TFIIA

   One TF recruits RNA polymerase II #4 TFIIB

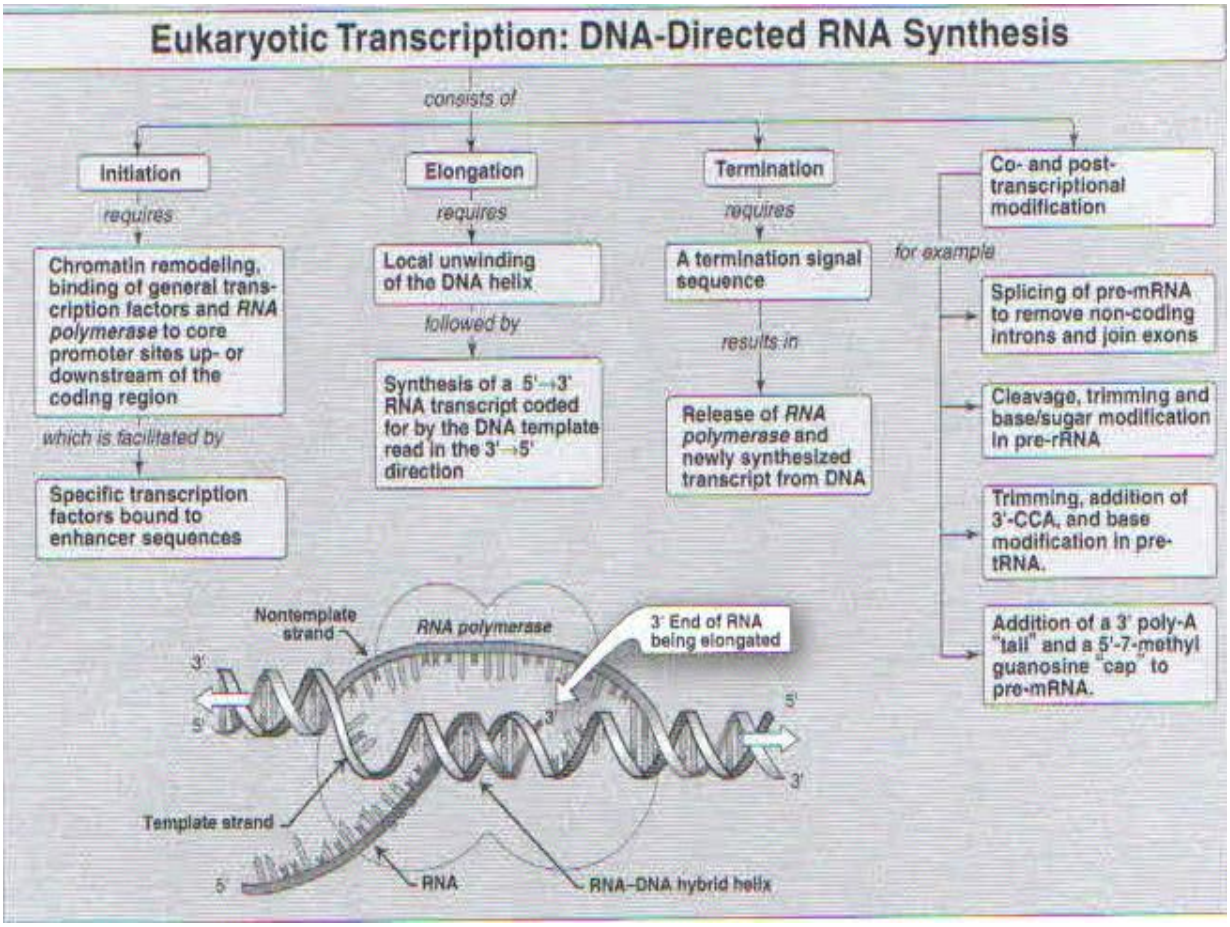
   One TF recognizes the promoter #5TFIIF

   One TF bind RNA polymerase to transcription initiation site #6 TFIIE

   One TF recruits TF responsible for opening the strands to allow clearance#6 TFIIE

   One TF does promoter clearance #7 TFIIH

   One TF moves RNA polymerase II along the template #7 TFIIH



Transcription Enhancers and Silencers: for regulation. Very important slide.

- Both are binding sites for transcription factors (TF's), however, they are <u>non-promoter</u> elements -can be found anywhere- that interact with TFs to either ↑ or ↓ their activities.

- Enhancers: "non-promoter DNA elements that stimulate transcription" They interact with general transcription factors to promote formation of pre-initiation complex to increase the amount of Transcription from a nearby promoter (core + upstream elements)

- Silencers: Decrease amount of Transcription from nearby promoters slows down pre-initiation complex formation.

- Initially Defined as being "Position and orientation independent" (downstream, upstream within the gene itself ..etc).

- Found upstream, within, or downstream of genes, they function in either orientation (not always true)

- Sometimes a DNA element can act as an enhancer or a silencer depending on what is bound to it and what function it is going to do.

Posttranscriptional processing of RNAs: only to eukaryotes.

- Bacterial rRNAs and tRNAs undergoes no additional processing, after being transcribed they are immediately ready for use in translation.

- Translation of bacterial mRNAs can begin even before transcription is completed due to the lack of the nuclear- cytoplasmic separation that exists in eukaryotes and to afford a unique opportunity for regulating the transcription of certain genes.

- An additional feature of bacterial mRNAs is that most are polycistronic which means that multiple polypeptides can be synthesized from a single primary transcript.

- Polycistronic mRNAs are very rare in eukaryotic cells but have been identified.

- In addition, several viruses encode polycistronic RNAs.

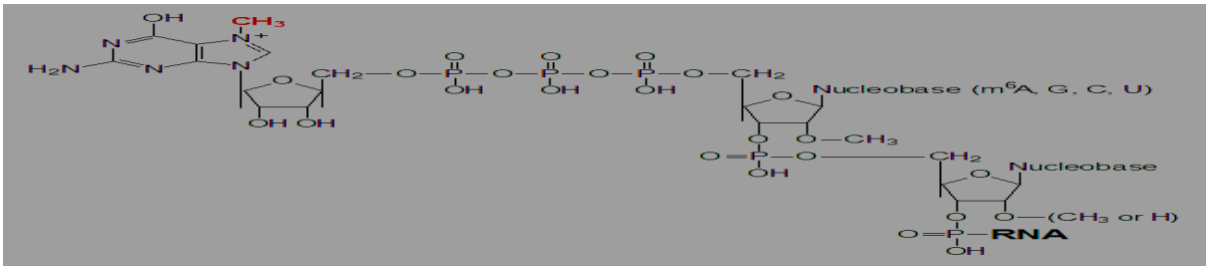
- In contrast to bacterial transcripts, eukaryotic RNAs (all 3 classes: tRNA, mRNA and rRNA) undergo post-transcriptional processing.

- All 3 classes of RNA are transcribed from genes that contain introns.

- The sequences encoded by the intronic DNA must be removed from the primary transcript prior to the RNAs being biologically active, this process of intron removal is called RNA splicing 1)(introns splicing), additional processing occurs to mRNAs, the 5' end of all eukaryotic mRNAs are capped with a unique 5'→5' linkage to a 7-methyl GTP, 2) Capping: @ 5' of mRNA one 7-methyl GTP molecule is added an enzymatic process that needs 3 enzymes. 3) Polyadenylation. Those are the 3 processes that happens to mRNA.

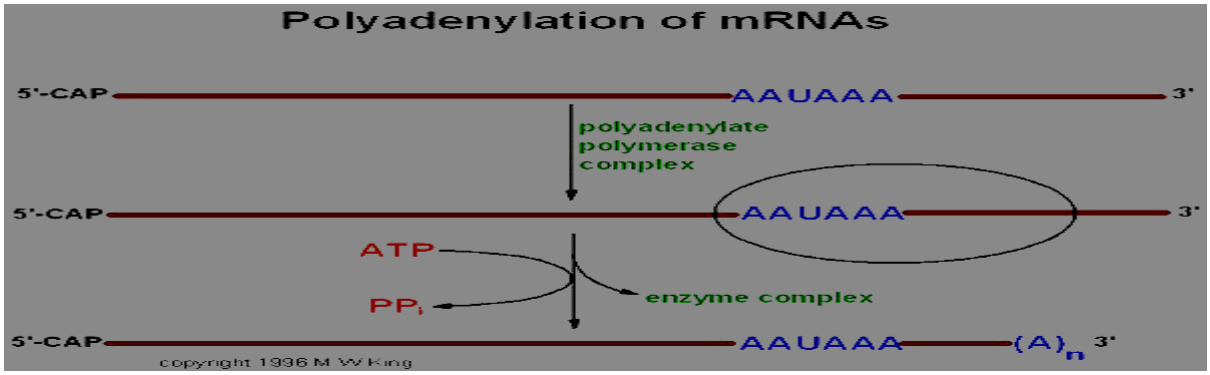
- Capping: The capped end of the mRNA is thus, functions in making it 1) protected from 5' exonucleases (attacks on the 5' end of mRNA) and more importantly is to be 2) recognized by specific proteins of the translational machinery (ribosomes) 5'to 3' if capped with 7- methyl GTP [on 5' end] it will be identified by ribosomes, while if 3'5' is capped with it [on 3' end] it will not be identified (no translation because it has a stop codon), moreover, the 5' end in protein is amino group while the 3' end in mRNA is carboxylic acid so it has to be plugged correctly (to the 5' end). So, its function eventually is to start translation in the correct orientation

- The capping process occurs after the newly synthesizing mRNA is around 20–30 bases long and not until it is fully formed.



Structure of the 5'-cap of eukaryotic mRNAs

- Polyadenylation: Messenger RNAs also are polyadenylated at the 3' end. A specific sequence, AAUAAA, is recognized by the endonuclease activity of by polyadenylate polymerase enzyme which cleaves the primary transcript approximately 11–30 bases 3' of the sequence element. A stretch of 20–250 A residues is then added to the 3' end by the polyadenylate polymerase activity. Polyadenylate polymerase recognizes the AAUAAA sequence at the 3' then it cleaves it and plug a tail consisting of a # of adenine AA to provide a protection from 3' exonucleases.



Splicing of RNAs (introns [non-coding or non-translated proteins no need for it] splicing): it happens in the nucleus (if it was left to be done in the cytoplasm direct translation introns will not be read by ribosomes intermittent translation improper protein.

- There are several different classes of reactions involved in intron removal.

- The 2 most common are the group I and group II introns. Group I introns are found in nuclear, mitochondrial and chloroplast rRNA genes, group II in mitochondrial and chloroplast mRNA genes. In general, there is 3 types of introns.

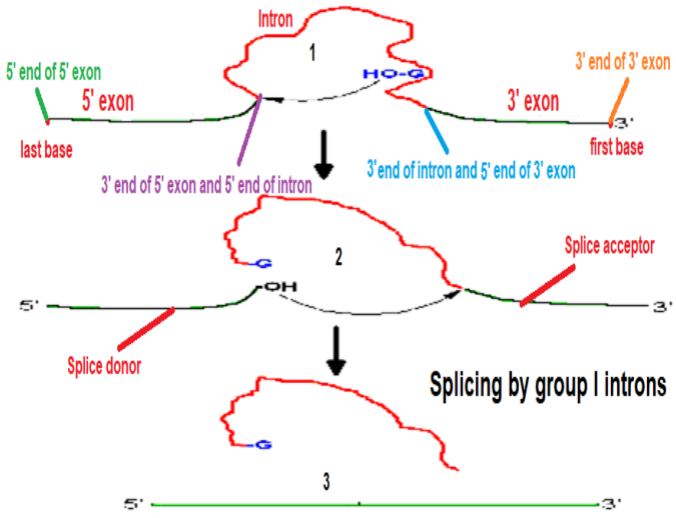
- Many of the group I and group II introns are self-splicing, however, some of group I introns require an external guanosine as a cofactor.

Group I introns:

- The 3'–OH of the guanosine nucleotide acts as a nucleophile to attack the 5'–phosphate of the 5' nucleotide of the intron.

- The resultant 3'–OH at the 3' end of the 5' exon then attacks the 5' nucleotide of the 3' exon releasing the intron and covalently attaching the two exons together.

- The 3' end of the 5' exon is termed the splice donor site and the 5' end of the 3' exon is termed the splice acceptor site.



In other words, guanosine with its OH group at 3' end of intron has a nucleophilic ability (can attack as a nucleophile), it will attack the phosphodiester bond between the last ribonucleotide @ the 3' end of the intron and the first ribonucleotide @ the 5' end of the 3 end intron. The OH now will be transferred to the 3' end of the 5' exon making it a splice donor that will attack itself the phosphodiester bond between the 1st base on 3' end of intron and the last base on 5' end of 3' exon (now called a splice acceptor). The intron is now released and the exons are fused together. That being said, intron splicing entails not only the removal of introns but fusing the exons together to form mRNA later on.

(الفقرة مش متأكدين منها كونها ذُكرت بشكلين مختلفين في مواضع مختلفة)

Group II introns:

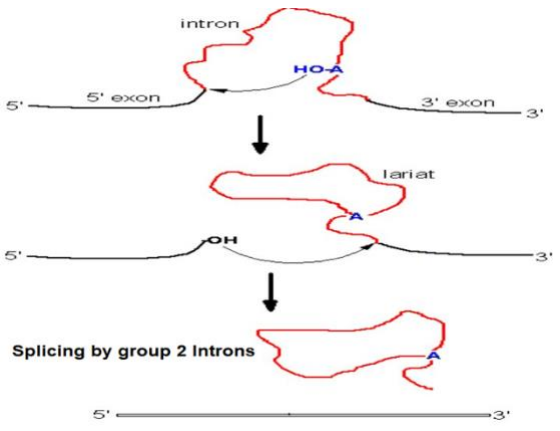
- Group II introns are spliced similarly except that instead of an external nucleophile (guanosine), the 2'–OH of an adenine residue within the intron is the nucleophile.

- This residue attacks the 3' nucleotide of the 5' exon forming an internal loop called a lariat structure (lariat: a lasso, similar to that used to catch bulls while riding horses).

- The 3' end of the 5' exon then attacks the 5' end of the 3' exon as in group I splicing releasing the intron and covalently attaching the two exons together

(الفقرة مش متأكدين منها كونها ذُكرت بشكلين مختلفين في مواضع مختلفة)

In other words, the internal adenosine at with its 2' OH will attack the phosphodiester bond between the last ribonucleotide of the 3' end of the 5' exon and the 1st base of the 5' end of the intron, to form a lariat shape. Then, the OH will be transferred to the 3' end of 5' exon (splice donor) which will attack the phosphodiester bond between the 1st base of 3' end of the intron and last base of the 5' end of the 3' exon. Exons are fused together and intron is released.



Group III introns: the most important type

- The third class of introns is also the largest class found in nuclear mRNAs, that undergoes a splicing reaction similar to group II introns in that an internal lariat structure is formed.

- However, the splicing is catalyzed by specialized RNA– protein complexes called small nuclear ribonucleoprotein particles (snRNPs), snRNP are ribozymes (snRNA) that bound to a protein.

- The RNAs found in snRNPs are identified as U1, U2, U4, U5 and U6.

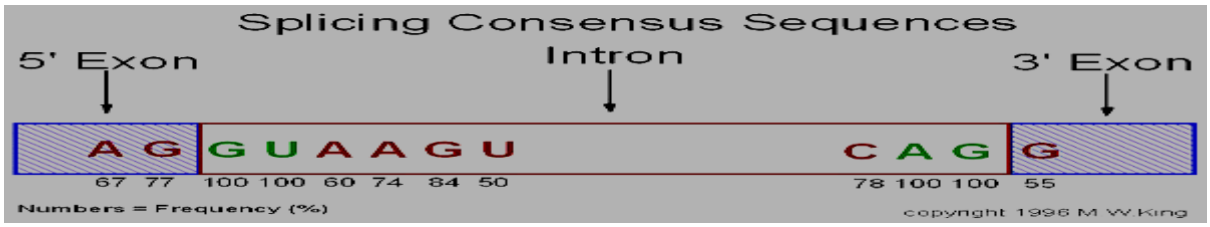
- Analysis of a large number of mRNA genes has led to the identification of highly conserved consensus sequences at the 5' and 3' ends of essentially all mRNA introns.

- The U1 RNA has sequences that are complimentary to sequences near the 5' end of the intron, its binding allows distinguishing the GU at the 5' end of the intron from other randomly placed GU sequences in mRNAs.

- The U2 RNA also recognizes sequences in the intron, in this case near the 3' end. Now, they break the phosphodiester bonds between: 1) the 1st base of 3' end of intron and the last base of 5' end of 3' exon and 2) the last base of 5' end of intron and the 1st base of the 3' end of 5' exon. Then the exons fuse together.

- The addition of U4, U5 and U6 RNAs forms a complex identified as the spliceosome (snRNA plus ~40 proteins) that removes the intron and joins the two exons together.

- U7 is involved in the production of the correct 3'ends of histone mRNA which lacks poly (A) tail.



- An additional mechanism of intron removal is the process of tRNA splicing.

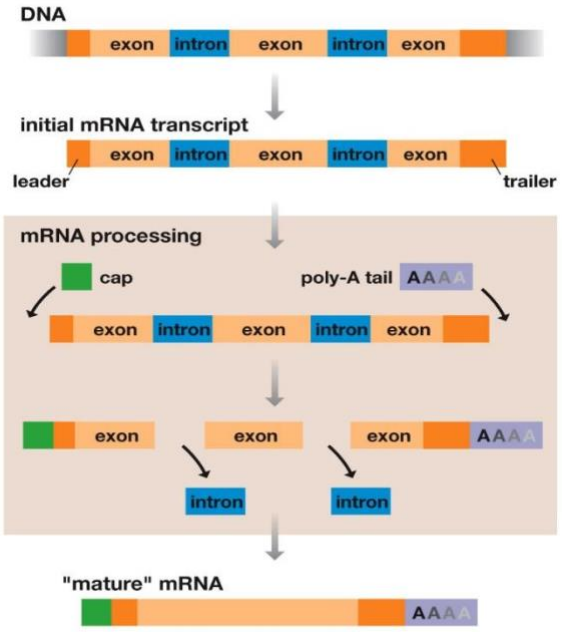
- These introns are spliced by a specific splicing endonuclease that involves a cut-and-paste mechanism.

- In order for tRNA intron removal to occur the tRNA must first be properly folded into its characteristic cloverleaf shape. ☘ if not folded no post-transcriptional processing.

- Misfolded precursor tRNAs are not processed which allows the splicing reaction to serve as a control step in the generation of mature tRNAs.

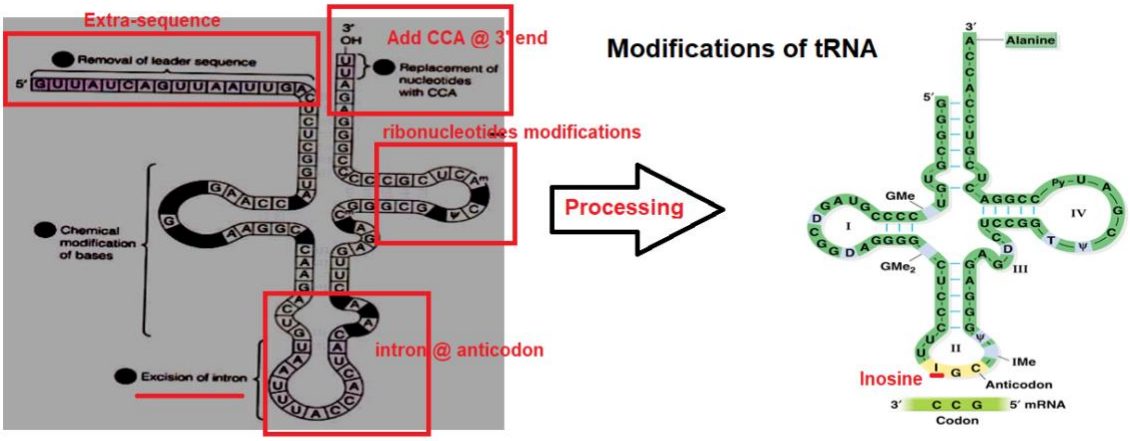
RNA processing: pre-mRNA mRNA

(mRNA is mature: does not have introns)



Modifications of tRNA: function of tRNA: carrying AA from cytoplasm to site of protein synthesis

1- Removal of 5' extra-sequence
2- Addition of: - CCA at 3' end - Anticodon loop
3- 3- Methylation of some bases: to give some unusual ribonucleotides. tRNA is the only RNA type that contain thiamine (a result of uracil methylation). Other unusual bases include pseudouridine and dihydrouridine. You may even find Inosine in the anticodon.



rRNA is used to construct ribosomes

- Eukaryotic ribosomal RNA is transcribed in the nucleolus by RNA polymerase I as a single piece of 45S RNA, which is subsequently cleaved by endonucleases enzymes group to yield 28S rRNA, 18S rRNA, and 5.8S rRNA.

- RNA polymerase III transcribes the 5S rRNA unit from a separate gene. The ribosomal subunits assemble in the nucleolus as the rRNA pieces combine with ribosomal proteins.

- Eukaryotic ribosomal subunits are 60S and 40S. They join during protein synthesis to form the whole 80S ribosome.

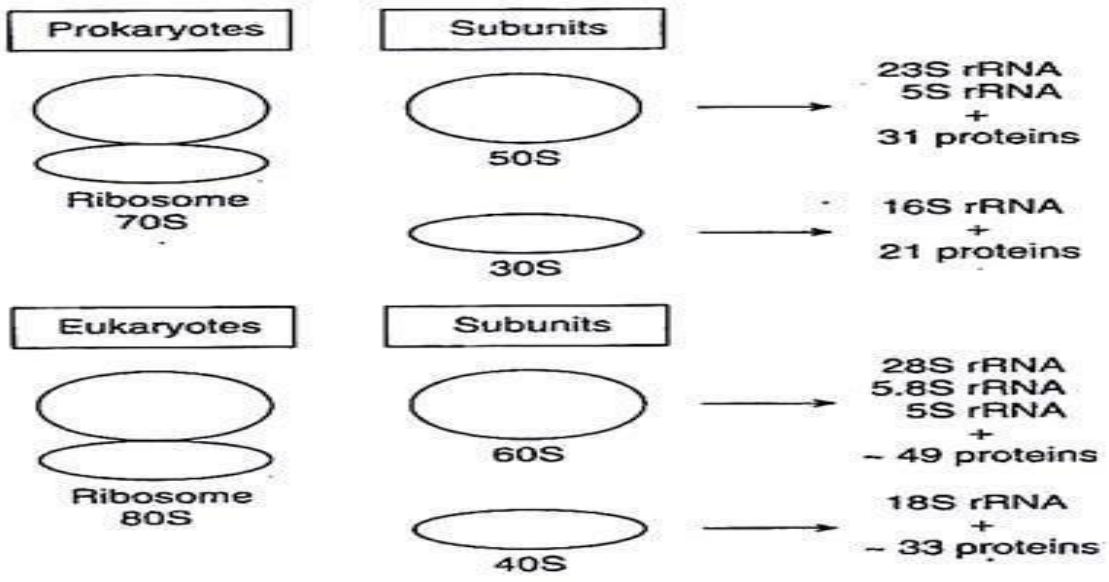


Fig. 2.47: Composition of typical prokaryotic and eukaryotic ribosomes

Subunits are either large (50s and 60s) or small (30s and 40s).
For the large subunits: # of rRNA used are 2 in prokaryotes and 3 in eukaryotes but for small ones it is always 1 subunit.
A # of proteins is also used.
(the doctor read the numbers, sorry)

"إنّ هذا الطّريق صَعب، لا يغُرَّنّك ظلامُ السَّيرِ فيه، إن أبصَرْتَ بقلبك، كَفاكَ ضوء القمَر حتّى تستمرّ".

ويخطو الإنسان بقلبه ما لا تُطيقُه القدم

- أ.قصى العسيلي.