# Lesson 05: Protein synthesis

## 1- Protein synthesis tools: Gene, ribosomes and tRNA

### 1-1-The embarrassed

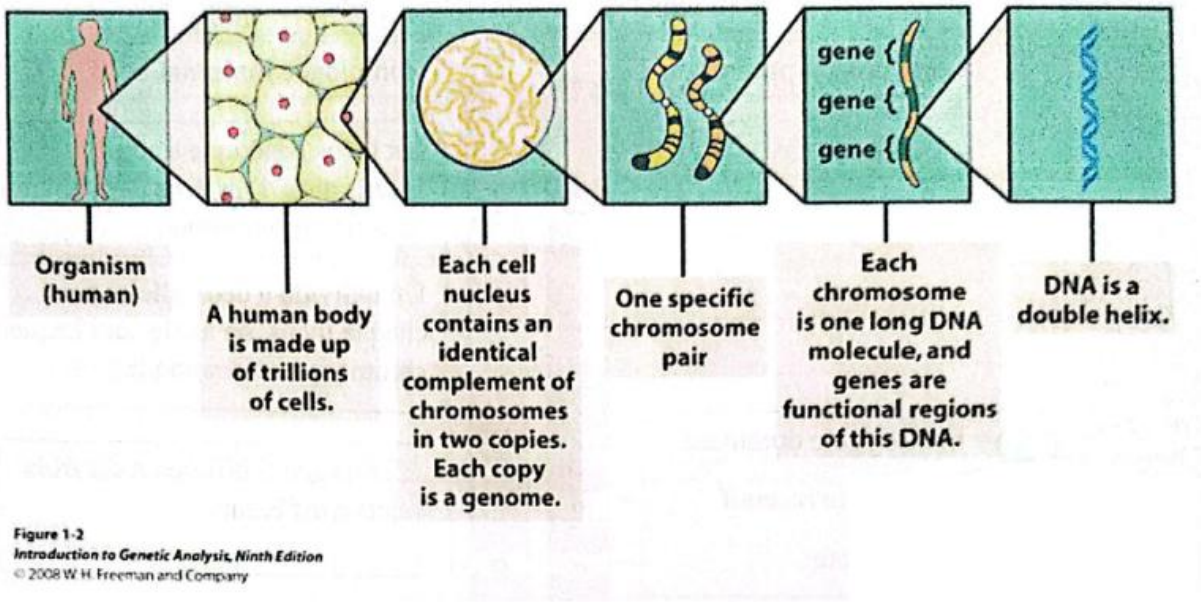



**Figure: From a living organism to a gene**

Genes contain the instructions that allow cells to polymerize amino acids in a specific order and thus synthesize specific proteins.

The precise order in which nucleotides are arranged along the DNA chain gives the specific code for proteins.

Genes are hereditary factors that are responsible for the phenotypic characteristics of the individual (a characteristic can be associated with several genes).

- **Gene structure**

The genes of eukaryotes have a more complex structure than those of prokaryotes, they contain:

- **Promoter** : controls gene activation, sequence located upstream of the gene adjacent to the gene transcription initiation site.
- **Exons** : In eukaryotes the coding information of the gene is often divided into a series of coding sequences called exons, separated from each other by introns.
- **Introns** : These are non-coding sequences within the gene. Introns are generally longer than exons, sometimes making up the majority of the gene. The number and length of introns and exons are highly variable.

- **Terminator** : sequence located at the end of the gene adjacent to the gene's transcription termination site.

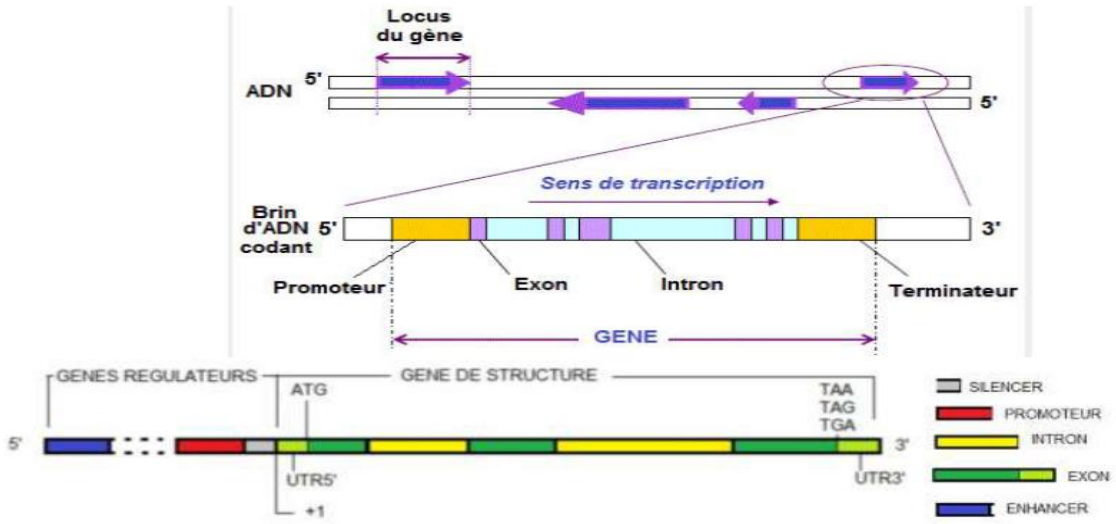
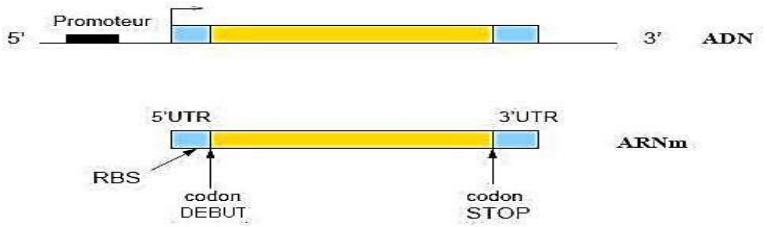



**Figure : Structure of a embarrassed at the house of eukaryotes**




UTR : Région non traduite (UnTranslated Region)
RBS : Site de fixation ribosome (Ribosome binding site)

**Figure : Structure of a embarrassed at the house of prokaryotes**

### 1-2- The ribosome

The ribosomes are "organelles" of composition complex. Each is composed of :

- Ribosomal proteins
- rRNA synthesized in a area particular of nucleoplasm : the nucleolus

He constituted of **two below units ribonucleoproteins** generally designated by their sedimentation coefficient in Svedbergs (S):

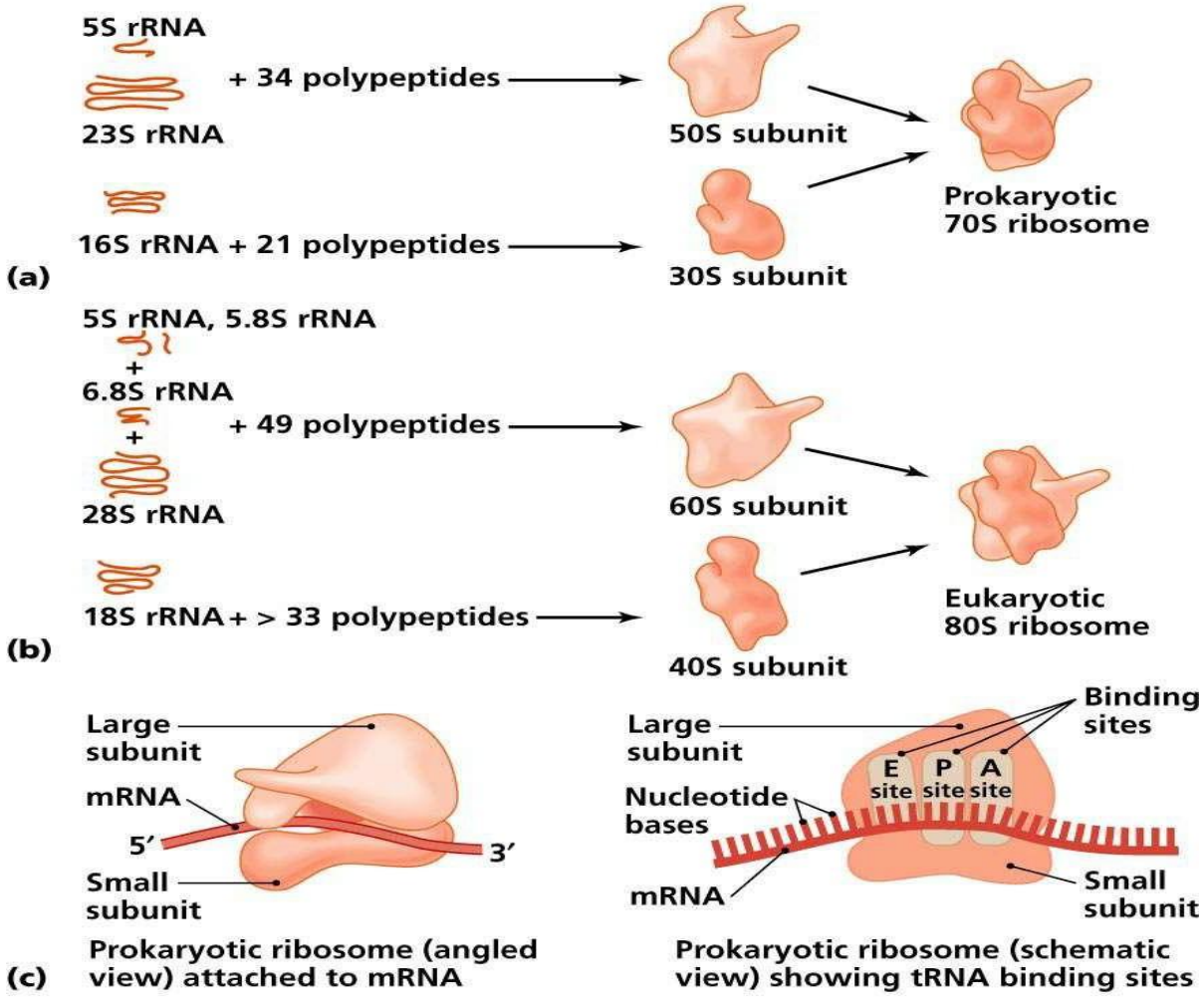
- There big subunit : 50S at the prokaryotes and 60S at the eukaryotes
- There small subunit : 30S at the prokaryotes and 40S at the eukaryotes

There synthesis of different subunits ribosomal is In cell nucleus, The subunits win the cytoplasm through nuclear pores.

Ribosomes are going be **free** in yhe cytosol or **related At reticulum endoplasmic** (REG).

In the cytoplasm the subunits ribosomal are going associate to mRNA For participate has there protein synthesis. The two elements of the ribosome are essential for **translation** and will be put in place at the time of **the initiation phase** of this process.

The small subunit recognizes the mRNA and binds to it fixed, the large subunit completes the ribosome and presents **recognition and processing sites for each tRNA** loaded with amino acids.





- **Figure : Structure and types ribosomes**

### 1-3- tRNA (RNA of transfer)

tRNAs are small molecules (about 70 nucleotides) synthesized in the nucleoplasm, having a structure secondary in "leaf of clover" represented In there figure below. Two Characteristic

regions are distinguished:

- The arm acceptor understand the two ends, It is the end **3'** Who go **fix acid amino acid** , it is invariable for all tRNAs ( CCA **)** .
- A succession of double helix structures interrupted by loops including the **anticodon loop** which contains **the specific triplet** , and constitutes the recognition site of the **codon** of the mRNA linked to the ribosome.

The tRNA own two functions essential : there possibility, **to bind to a specific amino acid** and on the other hand recognize a specific codon to **an anticodon** complementary to the **codon.**

**Anticodon-codon** recognition is based on base **complementarity** and involves the primary structure of tRNAs ; specific recognition of an amino acid is much more complex and involves **the three-dimensional architecture** of these particular molecules.

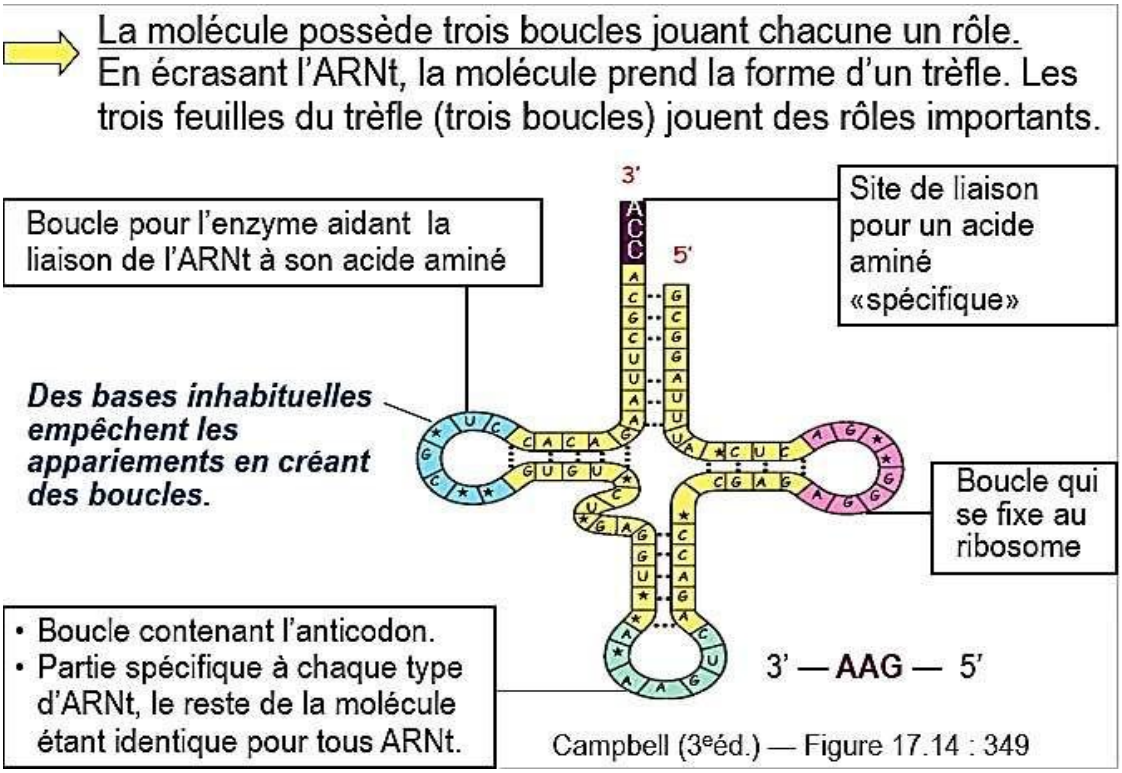



**Figure : Structure of a tRNA**

**2- Transcription**

By definition, transcription corresponds to the synthesis of an RNA molecule from a DNA template. Genes are transcribed only when their products are needed by the cell. These products correspond either to a polypeptide chain or to a functional RNA.

The transcription is a biosynthesis of RNA which rests on there **complementarity of the bases :**

- Alone, of small portions of the genome are transcribed has a era data of the life of there cell
- There transcription begin in a point accurate of DNA for to end in a point also precise, the space between the two constitutes **a transcription unit.**
- **A alone strand of DNA** is transcribed **(The strand 3'- 5')**, It has say serves of model has there polymerization ribonucleotides. The result is an **RNA molecule whose 5'-3' orientation** corresponds to the NH2 - COOH orientation of the protein. The reading of the code (translation) is done in the same direction as transcription.
- The entire transcription product does not match the protein code; on the 5' side (upstream) a guide sequence allows the attachment of the messenger RNA to the ribosome.
- Transcription is carried out by an **RNA polymerase** which uses single-stranded DNA to polymerize ribonucleotides (local denaturation of the DNA molecule is necessary).
- Many **factors proteins** intervene For ensure THE steps of the transcription.
- A gene can be transcribed simultaneously by several polymerases.
- The differences sensitive exists between transcription at prokaryotes the eukaryotes.

- **The mechanism of transcription** (see figures below) **:**

- **Initiation**

RNA polymerase binds to the gene promoter. The two strands of DNA unwind at this point (break the connections hydrogens weak) and there transcription begin in **a alone strand (The strand 3'- 5').**

- **Elongation**

Transcription is carried out by a family of enzymes in the 5'-3' direction: **RNA polymerases** (different according to the kind of RNA transcribed), RNA polymerase in the Prokaryotes is RNA polymerase II and in Eukaryotes have RNA polymerase I and III.

**ARN polymérase I :** transcrit les gènes codant pour les deux molécules d'ARN ribosomial
**ARN polymérase III** : transcrit les gènes de l'ARN transfert, de l'ARN 5S, et quelques autres petites molécules d'ARN
**ARN polymérase II** : transcrit la plupart des gènes des eucaryotes qui codent pour des protéines, leur régulation est la plus complexe

Of manner antiparallel by report has DNA copy, of way

| ADN | ARN |
|---|---|
| A | U |
| T | A |
| C | G |
| G | C |

Complementary polymerase unwinds the coding strand and exposes 10 to 20 bases at a time.

Of the nucleosides tri-P of RNA (synthesized In the cytoplasm Then entered In the nucleus) pair with the bases.

A times matched, the nucleosides are polymerized by of the links phosphodiester .

Behind there wave synthesis, the RNA detaches and the DNA rewinds .

- **Termination**

There transcription to continues until there END of there region terminal. THE transcribed of RNA And there polymerase are released.

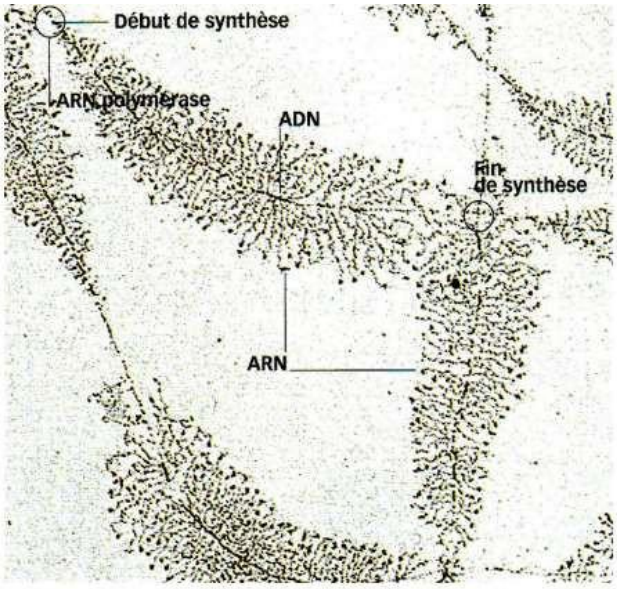



**Figure: Observation At microscope of there transcription of DNA**

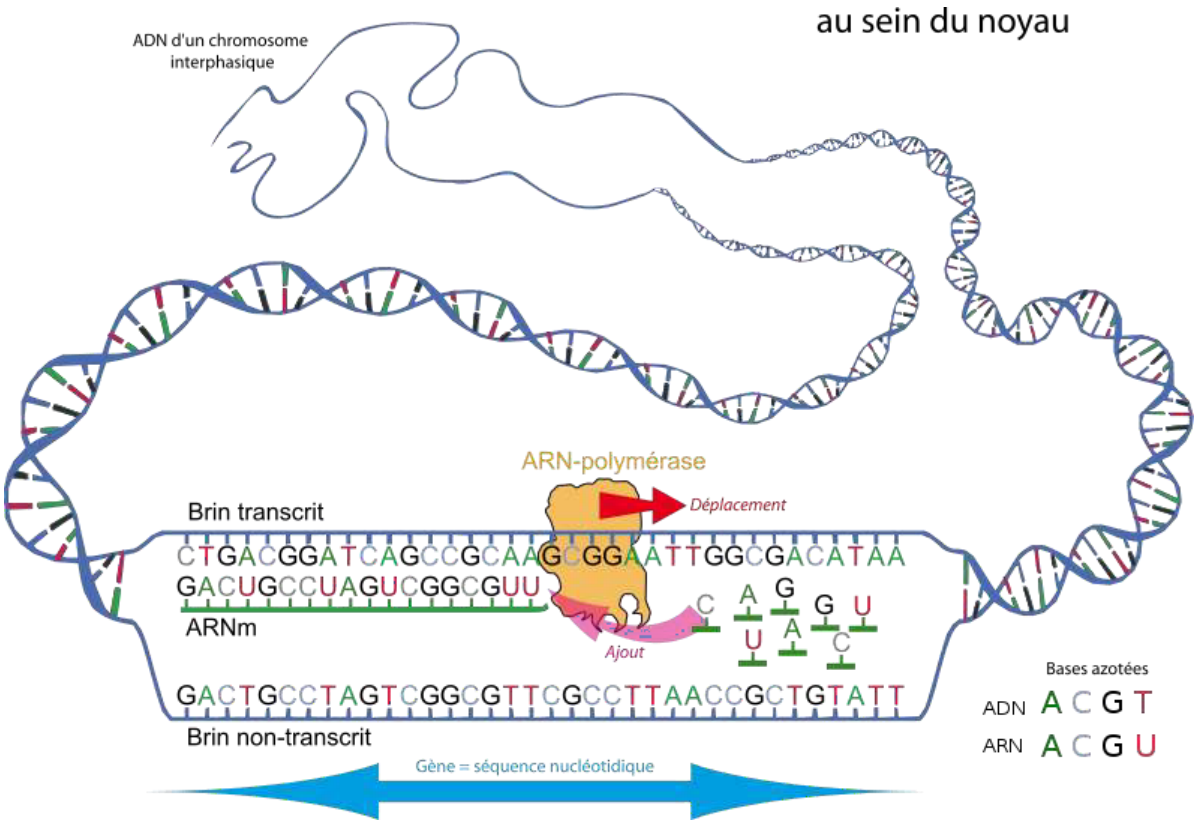



**Figure : Transcription DNA**

- **Prokaryotic transcription**

Transcription occurs in 3 steps:

**1- Initiation**

Prokaryotic RNA polymerase binds directly to the promoter. RNA polymerase begins to unwind the DNA helix.

**2- Elongation**

The region containing the RNA polymerase, template DNA, and RNA is called the transcription bubble. The position of the 3' end of the RNA interacts with an incoming ribonucleotide triphosphate. After passing the transcription bubble, the transcribed DNA coils up again as it leaves the transcription bubble.

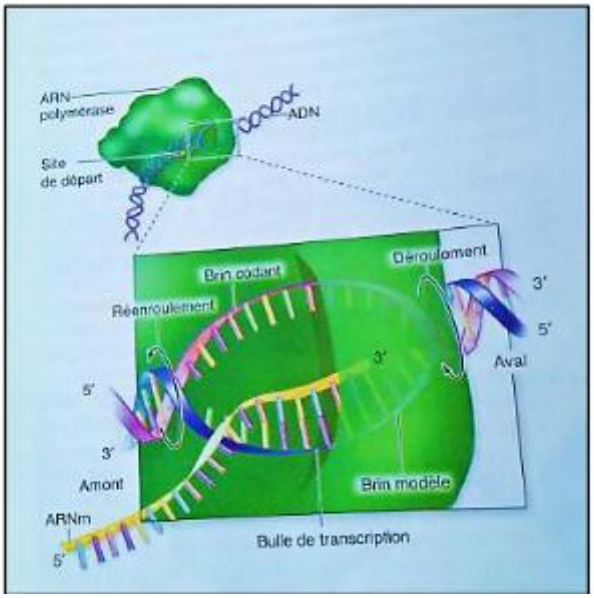



**Figure: Schematic of a transcription bubble**

### 3- Termination

The end of the bacterial transcription unit is marked by termination sequences that signal the polymerase to "stop". At this point, the RNA transcript forms a hairpin, followed by at least 4 uracil ribonucleotides , which stops the RNA polymerase. This causes the RNA and DNA in the transcription bubble to dissociate, releasing the RNA polymerase, and reconstituting the DNA helix.

- **Eukaryotic transcription**

### 1- **Eukaryotic RNA polymerases**

Most eukaryotic cells have three types of RNA polymerases:

- RNA polymerase I transcribes rRNA;

- RNA polymerase II transcribes pre-mRNAs, pnoRNAs , some miRNAs and some pnRNA ;

- RNA polymerase III transcribes small RNA molecules, specifically tRNAs , small rRNAs, some miRNAs , and some pnRNAs .

2- **Structure of the promoter**

RNA polymerase II promoters have control elements located upstream of the initiation site. These DNA sequences are called boxes:

- **The TATA box** : located approximately -25 base pairs from the origin of transcription. It is a sequence of six nucleotides rich in A and T. The so-called consensus sequence (statistically the most encountered) is TATAAA.

- **The GC box** : most often located in the region between -110 and -40. It can be in the form of hexanucleotides : 5'-GGGCGG-3'. The box can be repeated several times.

- **The CCAAT box** : often located in the region between -120 and -80. This box can be located before or after a GC box or even between two GC boxes.

**3- The phases of transcription**

- **Initiation complex**

Eukaryotic RNA polymerase II binds to the promoter via transcription factors including several proteins (TFIIA, TFIIB, etc.). These proteins associated with RNA polymerase II constitute the transcription initiation complex and catalyze the formation of the first phosphodiester bond between the first two nucleotides of the mRNA.

- **Modifications of the primary transcript (pre-mRNA)**

The primary transcript corresponds to a complete copy of the exons and introns of a gene. The final form is designated by the mature mRNA, obtained following the following modifications:

**- Addition of the cap at the 5' end**

The first base of the transcript is usually an Adenine (A) or a Guanine (G), and it is then modified by the addition of GTP to the PO45' group to form the cap. The latter protects the 5' end of the mRNA from attack by degrading enzymes; it is also involved in the initiation of translation.

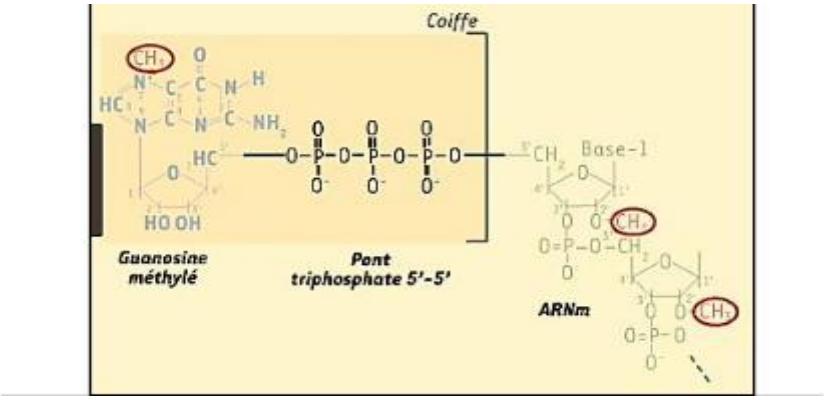



**Figure: Molecular structure of the cap**

**- Addition of the poly(A) tail to the 3' end**

After synthesis, mRNAs are cleaved, by an endonuclease , in their 3' part about twenty bases downstream of a specific sequence: AAUAAA. After this cut, the enzyme poly(A) polymerase in the presence of ATP adds a variable number of A's. The presence of poly(A) would also have a protective function of mRNAs on the 3' end. It also has a role in facilitating the attachment of ribosomes to the mRNA.

**- Excision-splicing**

Excision-splicing allows the maturation of the primary transcript into mRNA. This involves the removal of introns by excision followed by splicing of exons (end-to-end joining of exons). Excision-splicing takes place inside the spliceosome . This process occurs in the nucleus before the export of the mRNA to the cytoplasm.

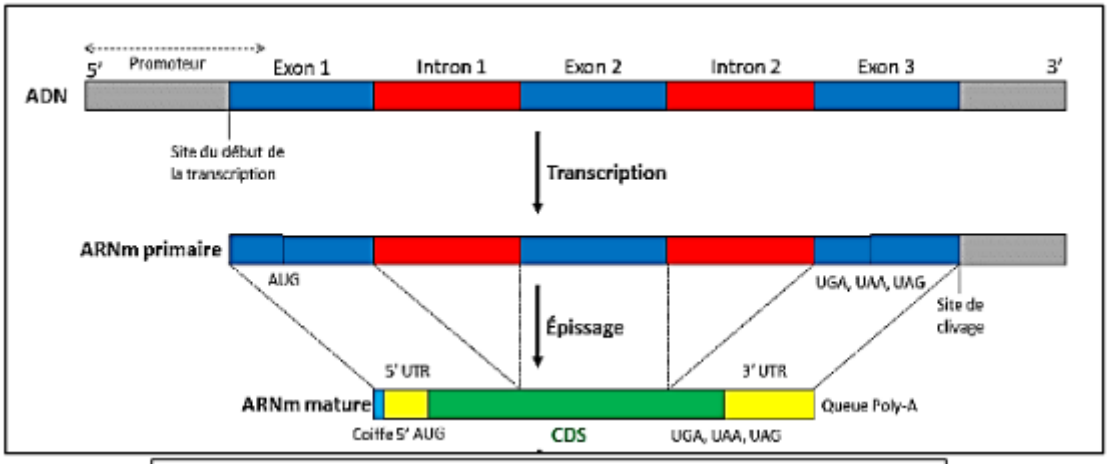



**Figure: Excision-splicing process**

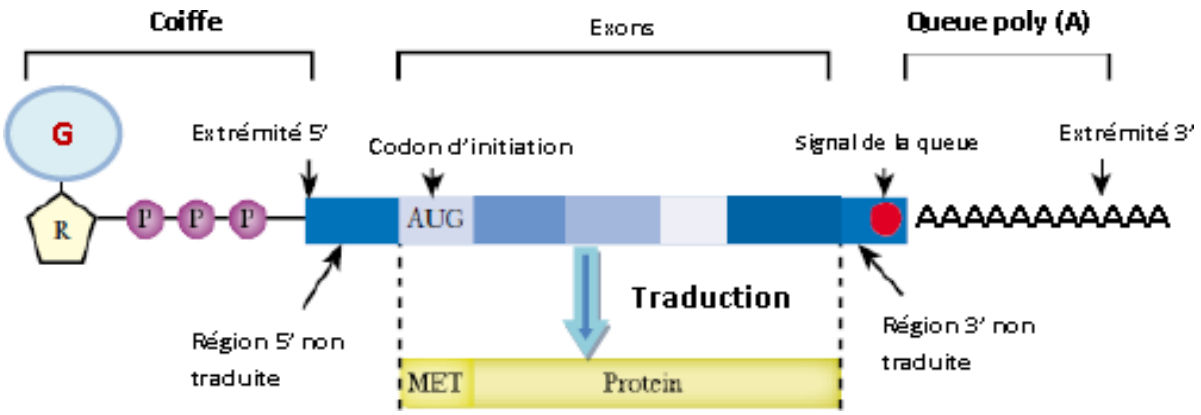



**Figure : Mature mRNA**

### 3- The translation

It is the mechanism by which the flow information go pass of there shape acid nucleic (4-letter alphabet) to protein form (20-letter alphabet) according to a **universal genetic code.**

Translation involves matching the nucleotide sequence of the mRNA to the amino acid sequence of the protein.

**The Genetic Code:** is the system of correspondence between the nucleotide sequences of RNA and the amino acid sequences of the manufactured protein.

The four nitrogenous bases of DNA constitute the genetic alphabet. The correspondence between the genetic information written with these 4 bases and the sequence of the polypeptide chain written with 20 amino acids is ensured by the combination of the four letters giving the genetic code. The basic unit of the genetic code is called a codon (three successive bases).

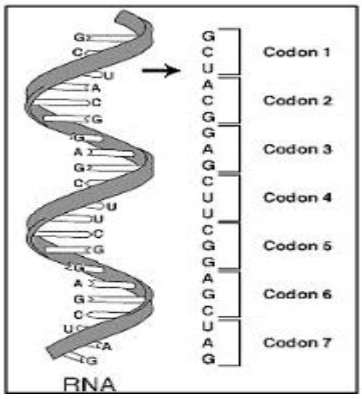
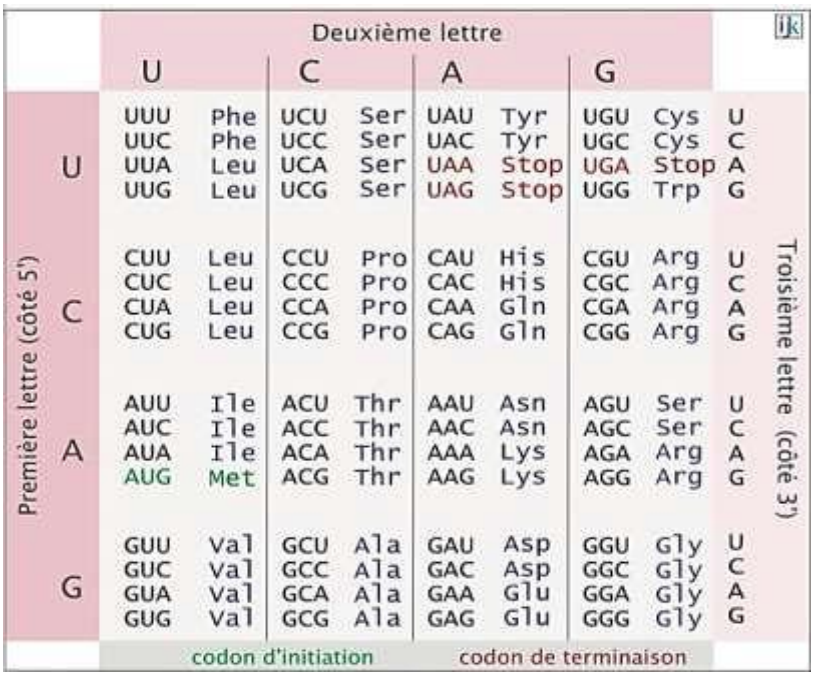



**Figure: Succession of codons on an RNA molecule**

With three nucleotides per codon, there are 43 = 64 different possible codons




| Acide aminé | | | Codons |
|---|---|---|---|
| Alanine | A | Ala | GCU, GCC, GCA, GCG. |
| Arginine | R | Arg | CGU, CGC, CGA, CGG ; AGA, AGG. |
| Asparagine | N | Asn | AAU, AAC. |
| Acide aspartique | D | Asp | GAU, GAC. |
| Cystéine | C | Cys | UGU, UGC. |
| Glutamine | Q | Gln | CAA, CAG. |
| Acide glutamique | E | Glu | GAA, GAG. |
| Glycine | G | Gly | GGU, GGC, GGA, GGG. |
| Histidine | H | His | CAU, CAC. |
| Isoleucine | I | Ile | AUU, AUC, AUA. |
| Leucine | L | Leu | UUA, UUG ; CUU, CUC, CUA, CUG. |
| Lysine | K | Lys | AAA, AAG. |
| Méthionine | M | Met | AUG. |
| Phénylalanine | F | Phe | UUU, UUC. |
| Proline | P | Pro | CCU, CCC, CCA, CCG. |
| Pyrrolysine | O | Pyl | UAG, avant élément PYLIS. |
| Sélénocystéine | U | Sec | UGA, avec séquence SECIS. |
| Sérine | S | Ser | UCU, UCC, UCA, UCG ; AGU, AGC. |
| Thréonine | T | Thr | ACU, ACC, ACA, ACG. |
| Tryptophane | W | Trp | UGG. (UGA) |
| Tyrosine | Y | Tyr | UAU, UAC. |
| Valine | V | Val | GUU, GUC, GUA, GUG. |
| Initiation | | | AUG. (UUG, CUG) |
| Terminaison | * | | UAG, UAA ; UGA. |

The genetic code has the following characteristics:

**- Degeneration**

Of the 64 codons, three are stop codons, which signal the end of translation. The remaining 61 codons, called sense codons, encode the 20 amino acids found in proteins.

The code therefore contains more information than is needed to specify these 20 amino acids, and is said to be a degenerate code. Only Trp and Met are specified by a single codon. For the other amino acids, the number of codons specific to each acid is variable (2, 3, 4 or 6). Codons that specify the same amino acid are said to be synonymous.

**- The reading frame and initiation codons**

The code is usually non-overlapping. Each nucleotide normally participates in only one codon. Each nucleotide sequence can be read in three different ways, depending on the reading frame that is applied. The translation system must use the correct reading frame determined by the initiation codon, which is usually AUG which specifies a Methionine (Met).

**- Termination codons**

Three codons – UAA, UAG, and UGA – do not specify an amino acid. These codons signal the end of a protein. They are called stop codons, termination codons, or nonsense codons.

There is no tRNA whose anticodon pairs with a termination codon.

**- The universality of the code**

All living things (with a few exceptions) have the same genetic code. The genetic code is said to be universal.

- On the plan genetics, the elements essentials of the translation are :
  - **messenger RNA (mRNA)** it provides the sequence of codons specifying each amino acid of the protein
  - **The ribosomes** they serve of support For ensure the connection successive of the acids amino acids
  - **The RNA of transfer ( tRNA )** capable to ensure there acknowledgement And there connection between A codon and a specific amino acid
  - **Aminoacyl -tRNA synthetases** are enzymes that ensure the specificity of the bond between a specific tRNA and the corresponding amino acid.
  - **Initiation, elongation and termination factors,** each step requires specific

molecules: guanosine triphosphate (GTP) providing energy, proteins (eIF2, eIF3, eIF4 [A, B, F], eIF5), and release factor RF.

- **Steps of translation**

**1. Initiation**

Initiation involves 3 main stages:

- The small ribosome subunit, assisted by initiation factors, recognizes the 5' cap and binds to it; the small subunit then scans the mRNA like a scanner, until it locates the first AUG codon.
- The initiator tRNA binds to the mRNA at the initiation codon. There is antiparallel base pairing between the mRNA and the small subunit of the ribosome,
- The large subunit of the ribosome binds to the initiation complex.

In eukaryotes, the first amino acid is methionine. In bacteria, it is a methionine formylated on the NH2 terminus ( f-Met).

The progression of each tRNA inside the ribosome can be summarized as follows: site A → site P → site E

The initiator tRNA directly occupies the P site without passing through the A site but all other tRNAs first occupy the A site. Immediately after initiation, the ribosome is attached to the mRNA, the Met- MetRNA is positioned at the start codon in the P site, and the adjacent A site is unoccupied.

**2. Elongation**

Elongation corresponds to protein synthesis by the addition of amino acids to the C-terminal end of the nascent peptide chain. It is assisted by elongation factors (EF-Tu, EF- Ts and EF-G). It proceeds according to a cyclic mechanism in 3 steps:

- In the 1st step, a charged tRNA , associated in complex with the elongation factor EF-Tu and a GTP molecule, enters the A site of the ribosome where the tRNA anticodon pairs with the next codon in the mRNA. Once the charged tRNA is in the A site, GTP is hydrolyzed to GDP and the EF-Tu–GDP complex is released.

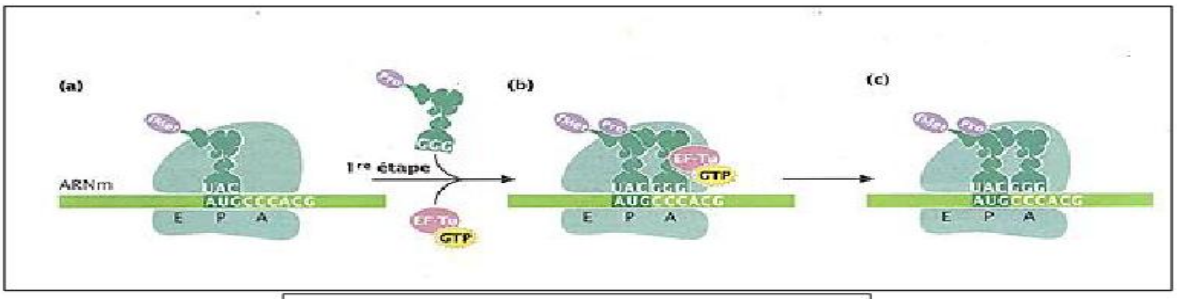



**Figure: First stage of elongation**

- The 2nd step is the formation of a peptide bond between the amino acids attached to the tRNAs occupying the P and A sites. This bond releases the amino acid from its tRNA at the P site.

- The 3rd step of the mechanism is translocation, which is the movement of the ribosome one codon along the mRNA. This step requires the elongation factor (EF-G) and the hydrolysis of GTP to GDP. Since the tRNAs in the P and A sites are still attached to the mRNA by codon-anticodon pairing, they do not follow the movement of the ribosome. Consequently, the tRNA that occupied the P site is now in the E site, which it leaves for the cytoplasm where it can be reloaded. Translocation also causes the tRNA that occupied the A site (which is attached to the growing polypeptide chain) to move to the P site, which leaves the A site open.

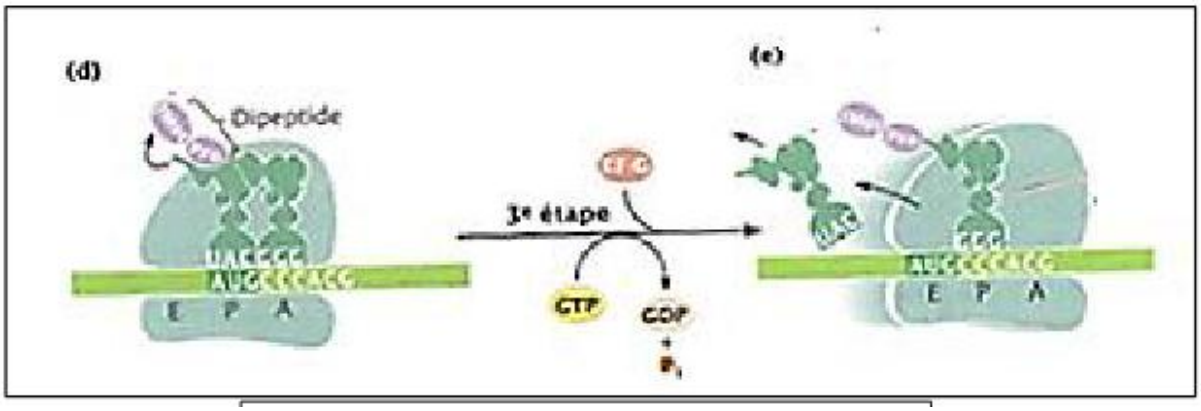



**Figure: Ribosome translocation**

After translocation, the A site is empty and ready to receive the tRNA specified by the next codon.

The cycle repeats: a tRNA and its amino acid occupy the A site, a peptide bond is formed between the amino acids present at the A and P sites, and the ribosome moves on to the next codon.

Throughout the elongation, the polypeptide chain remains attached to the tRNA that occupies the site P. Elongation in eukaryotes proceeds in a similar way.

**3. Termination**

Translation termination occurs at the stop codons UAA, UAG, and UGA, which do not code for any amino acids. These stop codons are recognized by the RF (Releasing Factor) termination factors. Since there are no tRNAs with complementary anticodons to the termination codons, no tRNAs enter the A site. The tRNA is released from the P site, the ribosome detaches from the mRNA and dissociates.

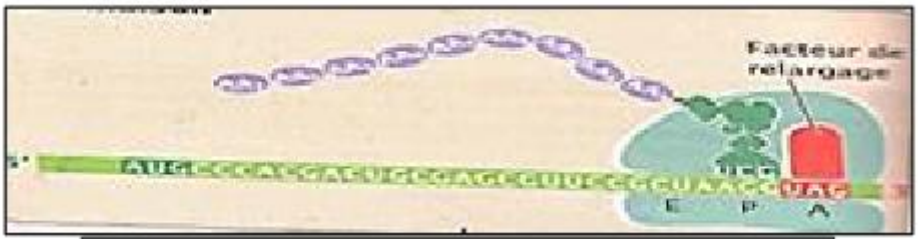



**Figure: Termination of the translation**

- **Post-translational modifications of proteins**

Polypeptide chains undergo post-translational modifications:

- Some proteins are synthesized as larger precursor molecules that must be cleaved and adapted by enzymes to acquire their function.

- For others, glycosylation (the addition of carbohydrate chains) may be necessary for their activation.

- The function of many proteins depends on their correct folding. Some fold spontaneously to acquire their correct shape, but the folding of some others must be assisted by other molecules called molecular chaperones.

**2- Post-translation**

modifications occur after the translation stage ( polymerization of the acids amino acids by rhe ribosome), For there production of the proteins by there cell. We quote:

- Elimination of the amino acid (Methionine): 50% of the proteins of prokaryotic cells And eukaryotes have there Methionine1 removed by the enzyme Met- aminopeptidase .
- Cleavage of the connections peptides : exp . Insulin
- Modification of the residues : glycosylation
- Association has others chains polypeptides identical Or No Or has of the ligands : exp . Hemoglobin.