University of Jordan
Faculty of Medicine

GENETICS & Molecular Biology

Number: 11

Done By: Majd Al Rawashdeh

Corrected By: Alma Jarkas

Subject: **Transcription in eukaryotes & mRNA stability**

Doctor: Mamoun Ahram

Price:

Date:

Transcription in eukaryotes

Things in general are more complicated in eukaryotes than in prokaryotes (i.e: in eukaryotes we need more proteins, more steps, and more regulation than in prokaryotes). In contrast to bacteria, which contain a single type of RNA polymerase, eukaryotic nuclei have three:

* **RNA polymerase I**: transcribes ribosomal RNA (rRNA) genes that are not translated.
* **RNA polymerase II**: transcribes protein-encoding genes, it's the one responsible for the production of the mRNAs that will eventually produce proteins.
* **RNA polymerase III**: transcribes tRNA genes and one of the rRNA genes (we will talk about them in details).

Note : ribosomal proteins are transcribed by RNA polymerase II , so :

- Ribosomal RNA → RNA polymerase I .
- Ribosomal protein → RNA polymerase II .

For the bacterial RNA polymerase (there is one polymerase): it sets on the DNA, unwinds it, and move to transcribe the DNA into RNA then terminate by itself. Some genes require role protein to terminate the transcription, but overall, the RNA polymerase in bacteria does almost all the work.

But the eukaryotic RNA polymerases don't, they require proteins; which are called **General Transcription Factors**, called "general" because they work on all genes .so, we can say that the eukaryotic RNA polymerase is lazy; it allows other protein to participate in the job.

Now, there are general transcription factors that are necessary for RNA polymerase II (called TFII), other factors are necessary for RNA polymerase I (TF I).

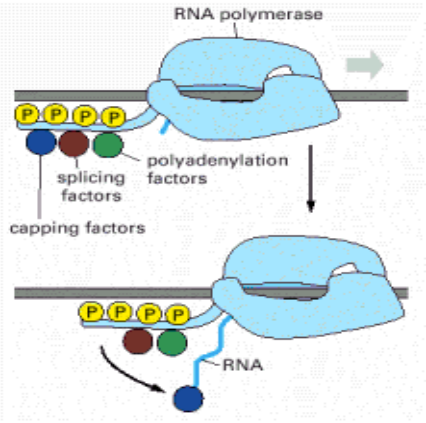
If RNA polymerase II transcribes a gene it needs those proteins (TFII), if it transcribes another gene it needs the same set of proteins; that's why they are called "general".

These transcription factors do a lot of work for the RNA polymerase; they:

- Help position the RNA polymerase on the promoter; they stabilize the interaction between the polymerase and the DNA.
- They act as helicases; they pull apart (separate) the 2 strands from each other.
- Activate the RNA polymerase to start transcription, so they give the RNA polymerase a push.

Again, **TF** = Transcription factor, **TFII** = Transcription factors which are necessary for RNA polymerase II; they are listed as TFIIA, TFIIB, etc.

Mechanism of transcription



❖ Initiation

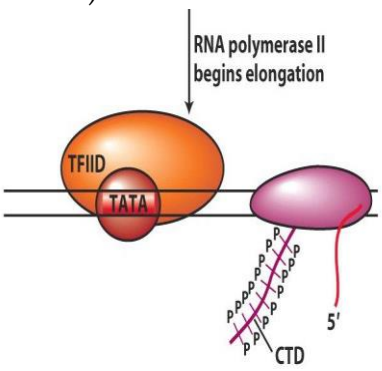
There is a large set of proteins but we will talk about only two of them that are necessary for initiation and for transcription in general:

* **TFIID:** helps the RNA polymerase bind to the promoter region (Consensus sequences), the binding of TFIID causes a bend in the DNA, and this bend in the DNA is the signal for other transcription factors to come along and bind to the RNA polymerase II and for transcription to start.

*__TFIIH__: act as helicase; unwinds the DNA in the promoter region forming open promoter complex.
we talked about it in excision – repair system (**coupling of transcription and repair)**.

❖ Elongation

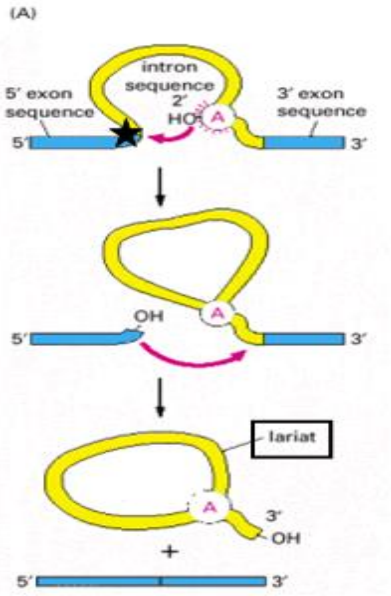
RNA polymerase can't move forward & can't start transcription until it's activated.
It's activated by phosphorylation , it has a tail (polypeptide extension) and this tail is phosphorylated & this is the signal that tells the RNA polymerase to start.



The protein that does this phosphorylation is **TFIIH** (the same one that functions in DNA repair and functions as helicase) here it acts as a kinase , it phosphorylates the RNA polymerase activating it to start transcription .

❖ Termination

Just like bacteria there is a consensus sequence which is (AAUAAA) in the RNA , once this sequence appears that's the signal ; the termination/stop point for transcription , then the RNA polymerase is released from the DNA releasing the RNA , then it can move backward to bind to another promoter region for another round of transcription .
This termination is coupled to another process known as **polyadenylation** .



A huge difference between prokaryotic transcription and eukaryotic transcription is that in bacteria the RNA is ready for the next step (translation) , in eukaryotic cells the RNA is extensively modified before translation ; there are **three major mechanisms of processing of the RNA** :

* **Capping .**
* **Splicing .**
* **Polyadenylation .**

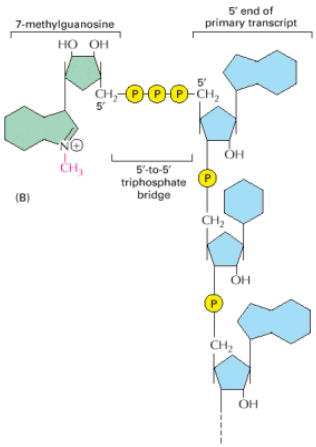
These modifying proteins bind to the tail of the RNA polymerase II , **only when the tail is phosphorylated** . so you can think of RNA polymerase as a train that carries these proteins and moves forward.

As the RNA polymerase is transcribing , these proteins can jump on the messenger RNA and they start processing & modifying it .

❖ **RNA capping**

What is a cap ?

A cap is a reverse guanine nucleotide , meaning that it's a guanine nucleotide but it's bound to the very first nucleotide of the mRNA in a reverse orientation (5′ to 5′ instead of 5′ to 3′) .



What is the importance of capping ?

- It **signals** to the cell that this is a messenger RNA ; so that the cell prepares itself to do translation .
- It **binds** to a protein complex (a set of proteins) known as **Cap Binding Complex (CBC) ,** and transport the mRNA from the nucleus to the cytoplasm . If the mRNA is not capped it will stay in the nucleus and it will not be transported.
- It helps in the **translation** ; in the cytoplasm there is other sets of proteins and ribosomal proteins that recognize the mRNA that need to be translated and bind to the cap , then they start the process of translation starting from the cap not from the middle of the mRNA .

* **Initiation** is by TFIID & TFIIH .

* **Elongation** is activated by TFIIH and phosphorylation .

* **Termination** is signaled by a sequence in the RNA (AAUAAA) and is coupled to polyadenylation .

RNA splicing

Splicing Is connecting exons to each other and it's catalyzed by RNA molecules(connected to a number of protein forming a complex known as the spliceosome) known as **snRNAs** (small nuclear RNAs) and each one of them is complexed with protein subunits to form a **snRNP** (small nuclear ribonucleoprotein) .

> Note that : **lipoprotein** is a protein that is modified by **lipids** .
>
> **Glycoprotein** is a protein that is modified by **sugar** .
>
> So , **Nucleoprotein** is a protein that is modified by **nucleic acid**

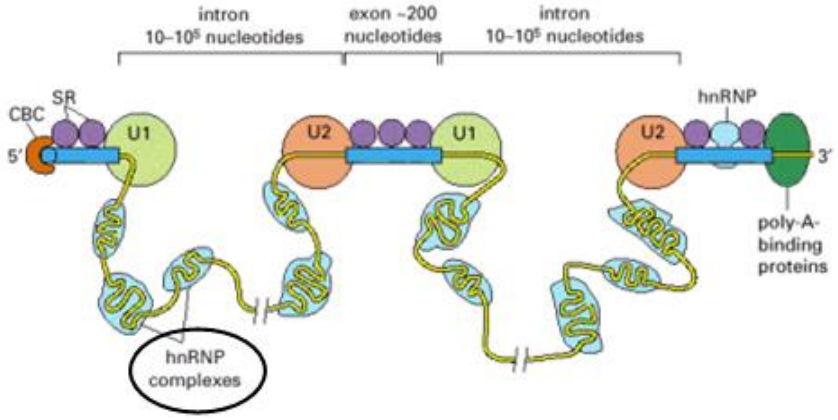
So the snRNP is a protein that is bound to the small nuclear RNA molecules , they form a complex that can do the splicing .

They bind to the borders of exons and introns , they cut the introns and release it in the form of **'lariat'** (as in cowboy movies) .

There are certain sequences at the borders (there is one near the end of intron)* that tells the snRNP that this is an intron that needs to be removed .

hnRNP:

There are other RNA molecules known as heterogeneous nuclear RNA molecules , they bind to proteins forming **heterogeneous nuclear ribonucleoproteins (hnRNP) .**



They modify the introns , as the pre-mRNA comes out of the RNA polymerase , these proteins binds to introns ; marking & covering it .

So it's actually the RNA molecules that do the splicing and recognize the introns ; this enzymatic & functional aspect of RNA molecules is amazing when thinking about evolution .

Accuracy of splicing :

Splicing is really accurate ; mistakes are not allowed to happen .

This accuracy is ensured by number of mechanisms :

- In general , exons are relatively small and introns are larger (variable sizes) so there is consistency when it come to the size of exons (more uniform than introns).
- The assembly of spliceosomes (complex of 5 small nuclear ribonucleoproteins **snRNPs**) occurs as the mRNA is coming out of the RNA polymerase .

As the mRNA is coming out of the RNA polymerase ; its marked immediately ,not after finishing transcription .This increases the accuracy of splicing and marking of introns and exons .

- We also have **SR proteins** that mark the borders as the mRNA is coming out they bind to the borders of exons and introns (introns are covered by **hnRNP**)

* **Marking** of mRNA and assembly of spliceosomes →occurs **during** transcription (**co-transcriptionally**) .

* **Splicing** the mRNA →occurs **after** the transcription is finished (**post-transcriptionally**) .
Just to make sure that all introns and borders are marked , & the splicing is accurate .

❖ **RNA polyadenylation**

Remember :
* The termination sequence (AAUAAA) , tells the RNA polymerase to terminate the transcription .
* Termination is coupled to polyadenylation .

Polyadenylation proteins are associated with the tail of RNA polymerase (which act as a train that carries all of these proteins) so when the proteins see the termination sequence they jump on the mRNA and do two things :

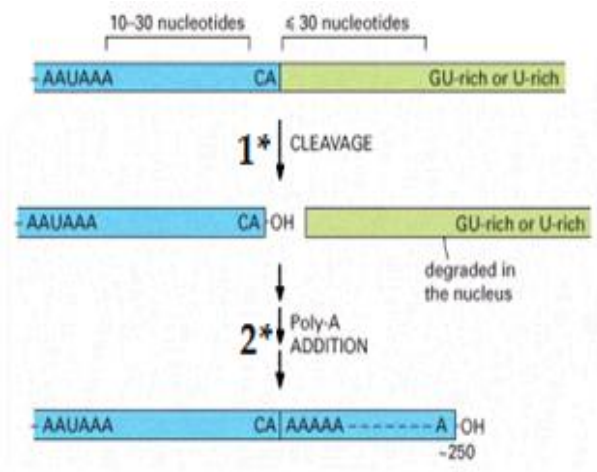
1* **Cleavage** : a cut is made few nucleotides after this sequence .

2* **Addition** of AAAAA… a sequence of about 200 A's at the end of mRNA . so the mRNA has a tail of A's at the end of it .

This is another signal/ sign/ mark that this is a messenger RNA ,the first one is **the cap** , the second one is **the presence of sequence of A's** .

The enzyme that adds the A's is known as **poly-A polymerase** , it does **not** require a template ; it adds the A's on its own .

Poly-A binding proteins do something similar to the cap binding proteins ; they :



- Help in transporting the mRNA out of the nucleus .
- Help in protein synthesis .
- Stabilize the mRNA ; so they regulate the stability of mRNA .

Based on a student question :

* There is a reservoir for the free nucleotides .

* The enzymes that produce the nucleotides are regulated by the nucleotides themselves .

Remember the enzyme**" Transcarbamylase"** that is regulated by **CTP** & **ATP** :

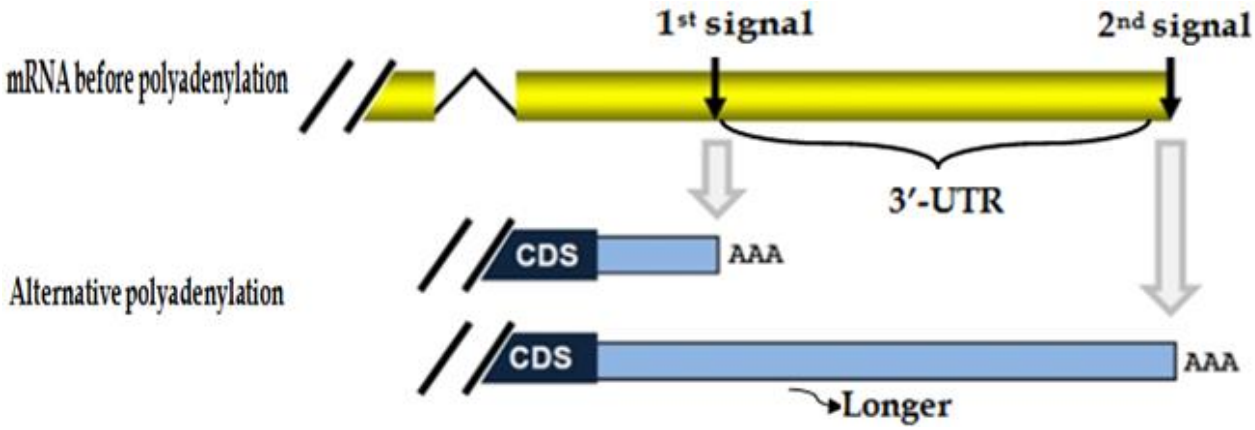
If the levels of **CTP** goes up it **inhibit** the enzyme .

If the levels of **ATP** goes up it **activate** the enzyme .

To make sure that the amount of pyrimidines & purines is balanced .

What differentiate humans from mice or worms is **regulation** ; which is how much of a protein is made , how this protein interacts with the other protein, how this protein and mRNA are regulated .

Alternative polyadenylation



In some genes there are **two polyadenylation signals** resulting in either cleavage at the first signal or the second signal .

3′-UTR : 3′-untranslated region (we will talk about it later) .

So it's the same gene that produces the same exact protein but the mRNAs are a little different (one of them is longer) .

What is the **importance** of alternative polyadenylation ?why do cells utilize it?

It's a mechanism of **regulating** how much **mRNA** there should be in the cell , how much **translation** there is and how much **protein** the cell would have, just regulation .

Example :

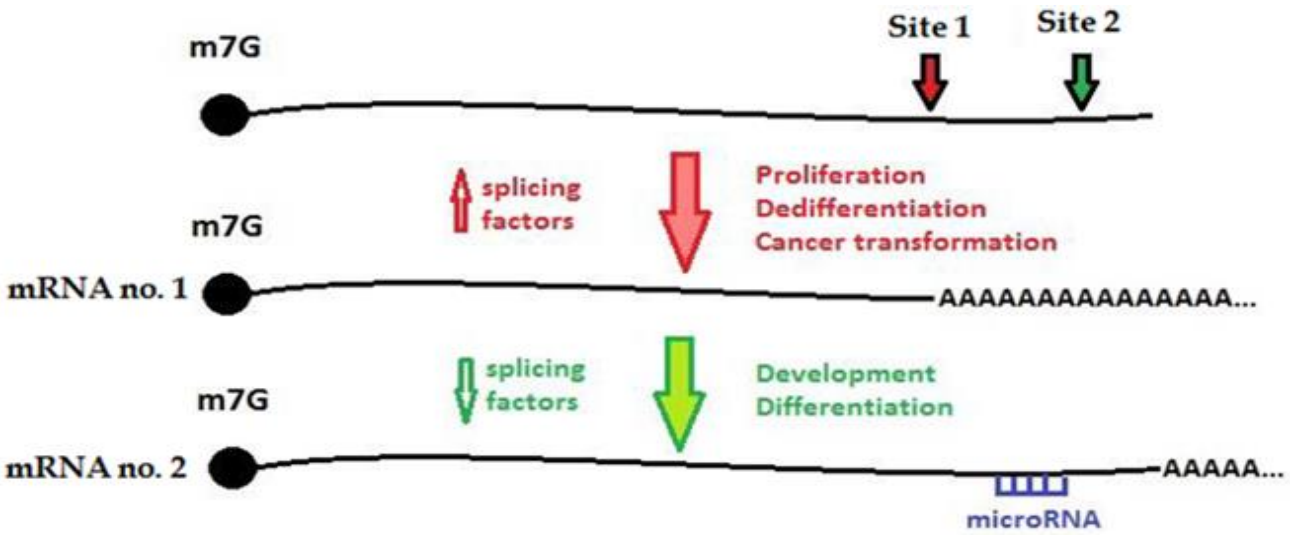
You have one certain mRNA , it can be cleaved at site 1 or site 2 and that will result in two different mRNAs :

mRNA number 2 have extension which contain a regulatory sequence that can bind to a **microRNA** (its complementary to the regulatory sequence of the mRNA and it can block translation or it can signal to the cell to degrade this mRNA).

If mRNA number 2 is degraded and the level of it is reduced ; the protein levels will go down as well .

So if the cell needs **more** amount of the protein , it will cleave the mRNA at **site 1** , **preventing microRNA** from binding to the mRNA (as in mRNA no. 1) .

If the cell wants a **little** amount of the protein , it will cleave the mRNA at **site 2** , having the extension that **binds to the microRNA** →the RNA is degraded or translation is blocked making less of this protein .



How the cleavage is regulated ?

By certain **regulatory protein** that bind to site 1 and **mask** it (hide it) , so the cleavage enzymes that cleave the mRNA cannot recognize it , but they can recognize site 2 and make cleavage there .

If site 1 is not masked (not covered) then the cleavage will happen there (at site 1).

SNPs and alternative polyadenylation

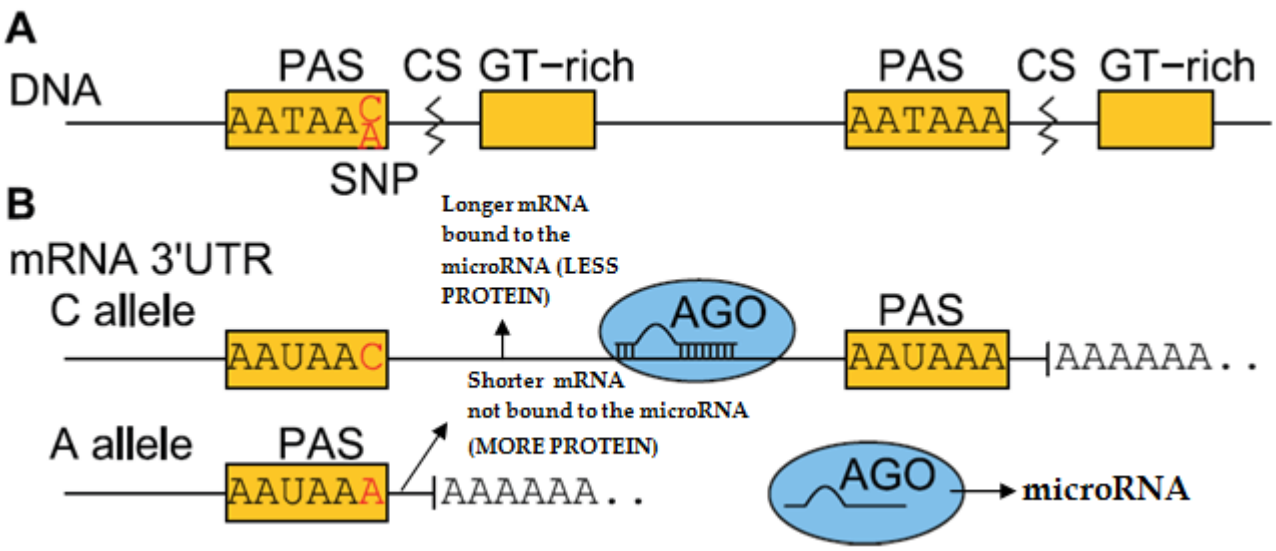
Remember : SNPs can be linked or causative , and if they are within the coding region, they can generate proteins with different amino acids at certain positions (individual variations).

The same thing happens with polyadenylation , some individuals can have a SNP in the DNA sequence that signals the polyadenylation .

If there is an **A** → there will be **cleavage .**

If there is a **C** → there is **no cleavage** and the other polyadenylation signal will be recognized producing a longer mRNA , in this case again , we have microRNA that can bind to the region that doesn't exist in the shorter mRNA .

So I will have more mRNA with A allele (the shorter mRNA) → more proteins will be produced . These are individual differences .



Summary

The cell can hide a polyadenylation signal or uncover it , regulating how much of the mRNA there is inside the cell , reflecting the amount of the protein (this is within a single cell in an individual) .

Sometimes there is differences among individuals as result of SNPs (single nucleotide polymorphism) , the sequence for polyadenylation is AATAA with A **or** C (some individuals can have A other individuals can have C).

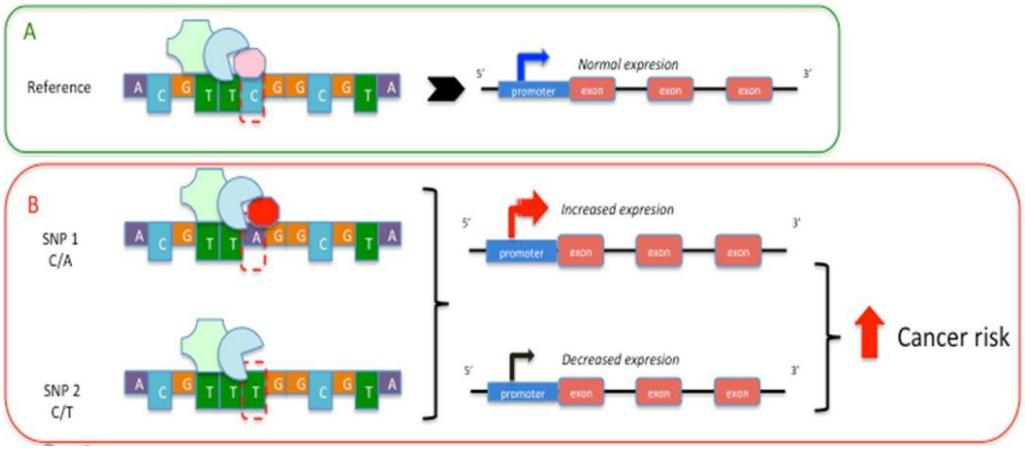
- If there is a **C** →the sequence will not be recognized by the cleavage enzymes , and synthesis/ transcription will continue until it reaches the second **p**oly**a**denylation **s**ignal (PAS) making a cut there .**

Here we have a **longer** mRNA and we can have a binding site for **microRNA** , when it binds, **the amount of protein** that is made would be **less** .

- Individuals that have an **A** →will have shorter mRNA .

SNPs in promoter

We can have **SNPs within the promoter region** (consensus sequence) remember that the RNA polymerase bind to the promoter region , if there is a certain **SNPs** this can either **increase** binding of the RNA polymerase and other transcription factor to the promoter region , or **decrease** the binding ; so certain individuals will have increase expression or decrease expression depending on the presence of a SNP .



Normally , we have <u>C</u> that result in normal expression

If we have an <u>A</u> there is increased production because the polymerase and the transcription factors can bind to the promoter .

If we have a <u>T</u> there is reduced binding resulting in less transcription .

That can really influence disease susceptibility in certain cases .

mRNA transport

It's another level of regulation , some mRNA molecules are transported , others are not ; and that influences how much protein is made in a certain cell .So,mRNA transport is highly selective and is associated to correct RNA processing.

mRNA degradation

When we have the mRNA outside we can regulate the amount of it by controlling how it's degraded .

In **bacteria** for example the half life of the mRNA is about **3 min ,** in **our cells** it can be **30 min** maybe **hours** , it depends.

So the amount of mRNA is regulated by how much of it is degraded (how stable it is).

Note : - some sequences influence the half life (increasing or decreasing it) .

- exonucleases are responsible for degradation .

Regulation of mRNA stability

The amount of iron in our body should remain stable, it shouldn't go down and shouldn't go high.
There are certain proteins that **regulate** the amount of iron in the body :

- **Transporters** in intestinal cells that absorb iron into the blood , like DMT1, Ferroportin .
- **Storage proteins** like ferritin (the storage protein for iron) , a single ferritin molecule can bind to 4 thousand iron molecules .

Ferritin :

If the body has high level of iron, it means we should high level of ferritin to store this iron,

If our cells need iron , it means that the level of ferritin should go down , to release this iron into the blood so it goes to our cells.

Transferrin receptor :

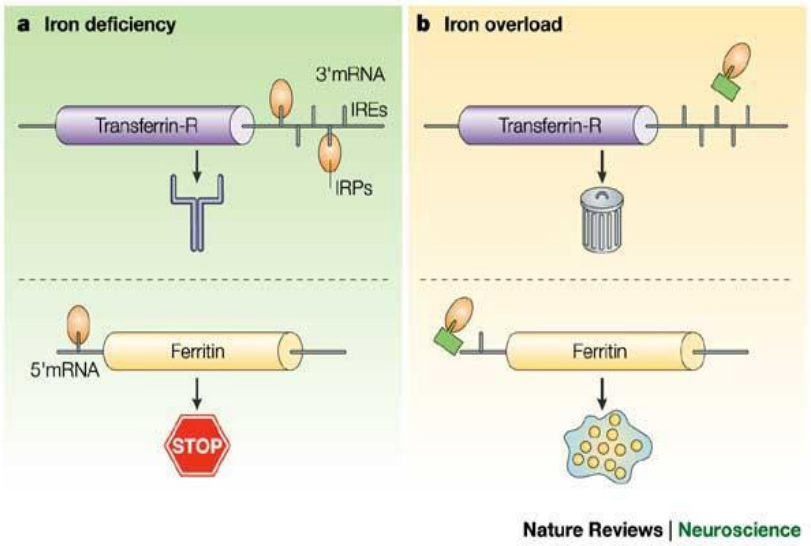
It's a protein that exists on the cell membrane of cells .

If cells don't need iron , transferrin receptor would not show up on the cell surface , because there is no need for iron . (They will stay in , and the production of transferrin receptor should go down).

If the cells need iron , they will have this transferrin receptor on their surface so it takes the iron inside the cell.

So again, if the body has high level of iron, it means that we should store this iron (increase production of ferritin).

If cells have sufficient amount of iron, transferrin receptor should go down .



On the other hand , if cells need iron , they will increase the production of transferrin receptor , and the liver should reduce the expression of ferritin to release this iron to the blood .

⇧ Iron ⇧ ferritin ⇩ transferrin receptor

⇩ Iron ⇩ ferritin ⇧ transferrin receptor

Iron Responsive Element

Ferritin mRNA & transferrin receptor mRNA , contain a certain sequence known as **Iron Responsive Element** , which is a sequence that binds to proteins known as **IRE-BP** (Iron responsive element binding protein) .

These proteins when they bind to the mRNA , they either stabilize it or reduce the stability of it , You should know how these 2 proteins are regulated

Iron deficiency

Here we should reduce the expression of ferritin and increase the expression of transferrin receptor.

The IRE-BPs don't bind to iron (there is no iron) , so these proteins will bind to transferrin receptor mRNA **stabilizing it** , so the mRNA stays inside the cells for a longer time , allowing for translation to happen and for the level of transferrin receptor to increase (Cells need iron).

But for ferritin we don't need it , because there is no iron to store, so the IRE-BP bind to ferritin mRNA and they **block translation** , so ferritin is not made; reducing the amount of it inside cells .

Iron overload

If the amount of iron is high, it will bind to IRE-BPs **releasing** them from the transferrin receptor mRNA, which will be degraded.
The translation of ferritin mRNA is **activated**, increasing the amount of ferritin.
That's how we can control the stability of the mRNA.
Another mechanism is **microRNA;** we will take it later on.

Good luck ^_^