Genetics & molecular biology

◉ Sheet ○ Slide

Number:	10
Done by:	Rawan Almujaibel
Corrected by:	Abduallah Alaraj
Doctor:	Ma'moun Ahram

This sheet was written from the record in the website. It is included everything you don't need to go back to the slide.

In previous lecture we discussed:

- Translesion DNA synthesis and how it occurs.
- Recombination repair: non-homologous end joining (NHEJ) and homologous repair (HR), how they linked to breast cancer.
- Then we discussed some controversial issue such as CRISPR Cas9.
- Afterward, we started talking about transcription in eukaryotes including: (RNA polymerase, transcription factors, promoters, tissuespecific transcription factors, enhancers and repressors, mechanisms of transcription)
- Then, we started to talk about eukaryotic genes.
- Ending up our previous lecture talking about alternative splicing

Processing of mRNA in eukaryotes

The RNA polymerase II gets activated by phosphoraylation by **protein TFIIH** (it functions as a kinase and helicase as well). When polymerase gets activated, 2 things will happen:

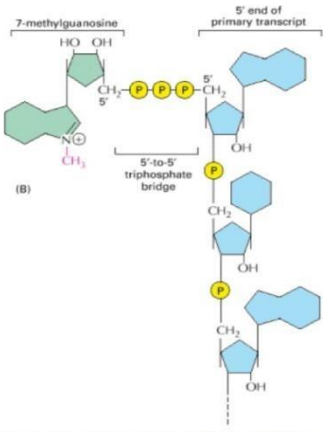
1. It starts transcription.
2. The tail will be a binding site to the proteins (these proteins do the processing to the mRNA – the processing happens only in eukaryotic cells).

The processing is composed of 3 types of processes:
- Capping
- Splicing
- Polyadenylation

When the mRNA is synthesized by the RNA polymerase, the proteins that are bonded to the tail jump on the mRNA to start processing it.

Let's talk about these processes:

1- **Capping** (addition of a cap, like something locks the mRNA) :
As soon as RNA polymerase II has produces about 25 nucleotides of RNA, the 5' end of the new RNA molecule is modified by addition of a "cap" that consists of GTP in reverse orientation.



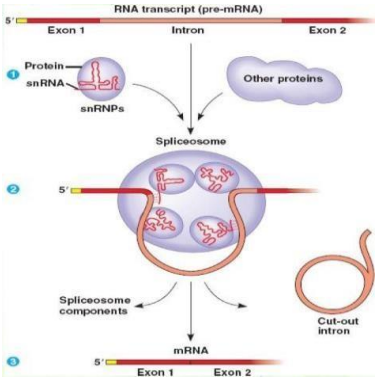
What are the benefits of having Cap?

1- Cap prevents mRNA from degradation and protects it and stabilizes it.

2- It signals the 5’ end of the eukaryotic mRNA which helps the cell distinguish it from other uncapped mRNA molecules
3- It helps transporting mRNA outside nucleus by binding the cap to a protein complex called CBC (cap binding protein which helps the RNA to be exported into the cytoplasm)
4- When the capped mRNA goes to the cytoplasm, the cap helps the cell recognizes it so it can start translation , (it helps in translation)
NOTE: CAP is methylguanine in a reverse orientation (refer to the diagram)

2- **Splicing**:
-Is the action of removing the introns and connecting the exons.

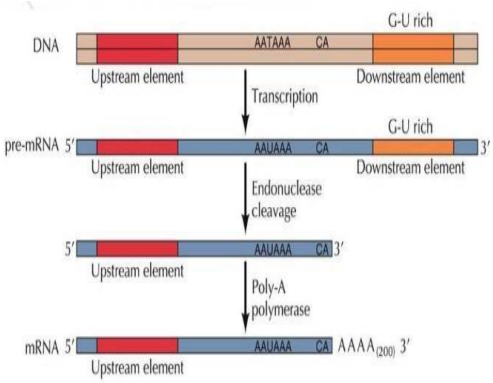
-splicing is very accurate, because it is connecting the end of each exon to the begging of the next exon and any mistake happens during this process it would cause **frame shift mutation** (producing defective and totally different type of protein), that's why it is very accurate.



-RNA splicing is catalyzed by big molecule called **Spliceosomes**.

-Spliceosomes are a complex of proteins and small nuclear RNAs (The catalytic site itself is largely formed by RNA molecules instead of proteins, so they are basically the enzymes are doing the splicing work, not the proteins).

3- **Polyadenylation (it is associated with termination):**
-A certain sequence in the mRNA (AAUAAA) followed by (CA), followed by (G-U rich) as <u>shown in the diagram</u> in the 3' ends of mRNAs is recognized by RNA-binding proteins and RNA-processing enzymes that cleave the RNA.



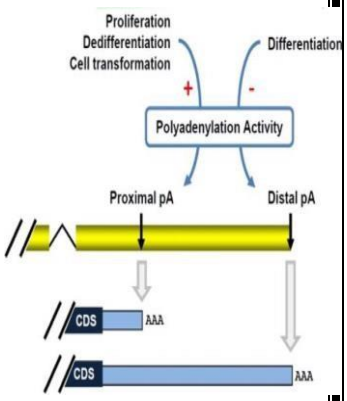
-The mRNA cleavage is occurs in the (CA) subunit, then it followed by the addition of poly-A tail (200 A's are added to the mRNA and <u>the nucleotide precursor for these addition is ATP</u>) at the 3' end.

-The adenylation is catalyzed by an enzyme called **Poly-A polymerase** (the enzyme does not use the DNA to add the A's , it doesn’t require a template and the poly A tail is not coded) <u>Significance of polyadenylation:</u>

- It helps in transporting mRNA from the nucleus to the cytosol.

- It helps in translation (same as capping).
- It stabilizes mRNA (same as capping).

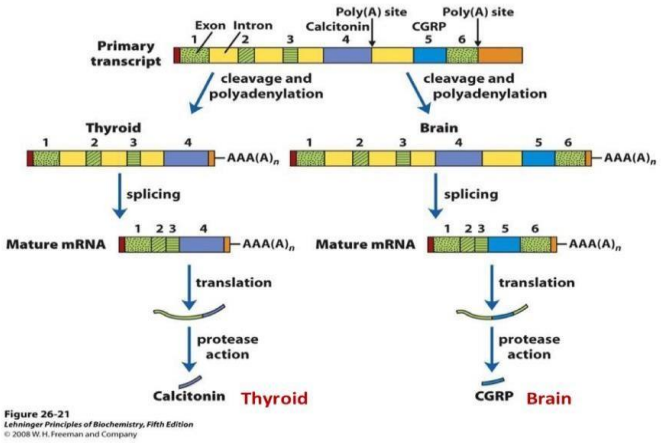
Alternative polyadenylation (it happens in certain genes only):



- In this process in some certain genes we have more than one poly-A signals (it is either in proximal, in middle or distal sites), so the cell itself would choose the suitable site for it. As a result producing different mRNA lengths, leading to produce different proteins.

  For example, as it is shown in the diagram below, there are 2 different proteins produced from the same gene. The proteins are Calcitonin (in thyroids) and Calcitonin geno-related protein [CGRP] (in Brain). <u>How does that happen?</u>

  We have more than one poly-A sites in a gene (one after exon 4 and the other after exon 6). In the thyroid cells we will have a termination on the proximal poly-A site resulting in the production of mRNA that contains exons 1, 2, 3, and 4 after splicing producing Calcitonin. However, in the brain the proximal poly-A site gets **masked** (hidden) and the cell does not see the poly-A site, so the mRNA produced with very long RNA molecule that contains exons 1,2,3,5, and 6 after splicing. The exon number 4 is removed along with introns as it was masked, producing CGRP.



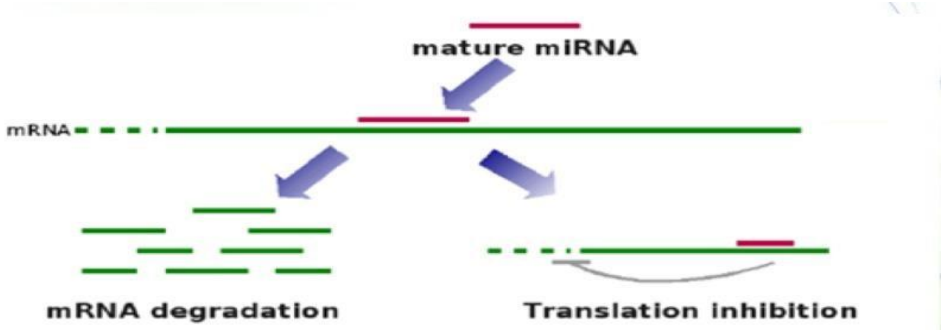
The purpose from all these above is to increase the diversity. We have 20,000 genes in our bodies, these genes make hundreds and thousands of proteins that have different functions, and they are tissue specific as well. And it gives us kind of answers around how one of the proteins would be produced in one cell and couldn't in other cells or tissues.

.....

microRNA and alternative polyadenylation:

What are **microRNA** molecules?

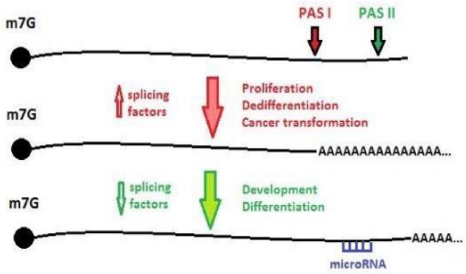
- They are short RNA molecule (about 22 nucleotides).
- They have sequences that can be complementary to certain mRNA, which will bind to the mRNA.
- When microRNA binds to mRNA, it will either **degrading** the mRNA **OR inhibiting** the translation of the mRNA.
  So the cell regulates how much proteins are need by regulating how much mRNA is available for translation.
  Logically, when we have a small amount of the mRNA it will lead to produce a small amount of protein.



How is that associated with the poly-A site?

let say that we have a gene has 2 poly-A sites (PAS) such as PAS1 and PAS2 and if the cell chooses the second site which produces a long mRNA. In this case, the gene doesn't have exons, instead it has microRNA binding site appears which won't be appear when the cell chooses the PAS1 because the mRNA is shorter.

So, if the cell produces the longer mRNA, the amount of the mRNA would be lower or the translation is reduced as a result of having microRNA binding site and it will lead to produce a small amount of proteins. So the cells can harmonies which site they want to regulate or to produce at a certain condition.



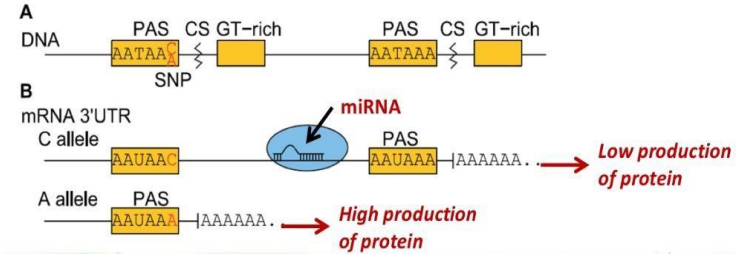
- The choice of poly-A site can be influenced by extracellular stimuli that regulate the expression of the proteins that take part in polyadenylation.
- Having a shorter transcripts would remove regulatory elements in the 3'-UTR and influence the half-lives of mRNA and, hence the amount of generated proteins.
- Example: Longer 3'-UTR would contain binding sites for microRNAs at the 3'-UTR, which tend to repress translation and promote degradation of the mRNAs.

.....

SNPs and alternative polyadenylation:

The presence of SNPs within the polyadenylation signal can also alter the length of the mRNA and, hence, protein amount in cells

In this diagram shows us that poly-A site [PAS] as (AATAAA) which is a termination signal and the gene has 2 PASs one at the proximal site and the other is in the distal site.



In the SNPs condition for example the sequence would be (AATAAC) instead of (AATAAA), so by this we will have only one PAS (2nd site) and that leads to produce long mRNA that includes miRNA binding site. The miRNA will bind to the mRNA either degrades it or reduces the translation which is leading to reduce the production of protein.

The other condition is when the termination site (AATAAA) PAS is normal, so the mRNA will be shorter which is not including the miRNA binding site , so we have high production of protein.

BY SNPs we will get an individual variation which will cause a phenotype

And that explains why some people need to take 2 panadol tablets while the other people one tablet is enough for them, because one of them has more enzyme than the other results to take different dosage.

mRNA transport (regulation)

When the mRNA is capped, spliced, and polyadenaylated and everything is in a good matter. It will be transported from the nucleus to the cytoplasm where it is translated into protein later. Transportation is highly selective and is associated with correct RNA processing.

If we have a gene has only one poly-A site and it is mutated, the transcription of the mRNA won't be terminated and as a result of this, the nucleus will produce very long mRNA and the **cell wont transport it out of the nucleus** which leads to produce very low proteins or even not producing it at all.

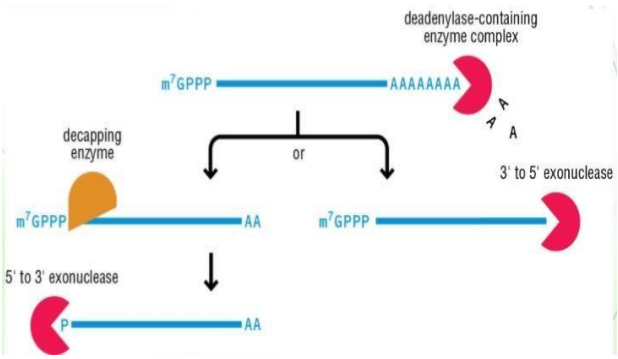
……

Degradation of mRNA

It is type of regulation to balance out the amount of the proteins in a cell.
The first thing happens in degradation is the removal of the poly-A tail, either by:

1- Deadenylation through exonucleases at 3' end only **or**
2- Deadenylation followed by de-capping (removal of cap) followed by exonucleases at the 5' end.

The stability of the mRNA is dependent on **the cell type** and according to the gene. The vast majority of mRNAs in a bacterial cell are very unstable, having a half-life of about 3 minutes. However, the mRNA in eukaryotic cells are more stable (up to 10 hours; average of 30 minutes).



……

Some phenomena in eukaryotes

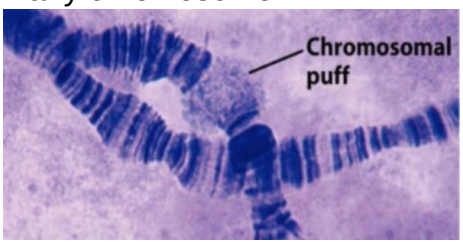
In general what happens in eukaryotic is the cell produces first pre-mRNA (it gets to be processed by capping>splicing> polyadenylation>transporting to the cytoplasm > translation , but exceptions happen to some certain genes. These exceptions can regulate:

1- How much protein in the cell.
2- How stable the proteins are.
3- Increases the diversity of the proteins.

The phenomena:

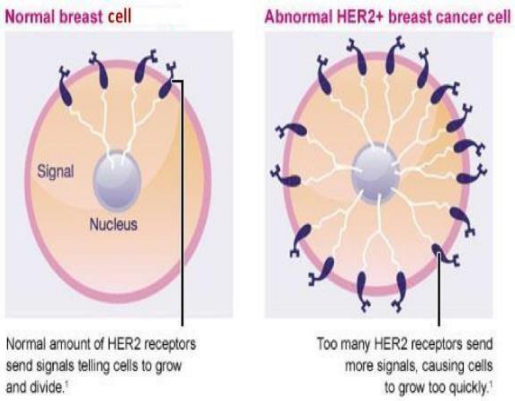
1-Gene amplification:

In any normal cell, we will find a single copy of a gene in any chromosome. Let say if the cell needs to produce more protein of a certain gene because one copy of that gene is not enough to the cell purposes. In this case the cell **amplifies** the gene (means that on the same chromosomes (on the same DNA strand) we will have multiple copies of the same gene and all of these copies are active, therefore they increase the protein production). For example the dihydrofolate reductase that helps in the escape mechanism of the cancer cells.



The cancer cells usually take an advantage of something that naturally and physiology found in the cell. So they take an advantage of gene amplification, **how**?

Let's take a breast cancer cell as an example, as you can see in the pic. The left picture is the normal mammary cell in breast; it has certain number of receptors (HER2 – Human epithelial growth factor receptor). This receptor is bind to a ligand that induces cell growth, so it is supposed to be under control. The picture on the right is the cancerous cell, so it amplifies the HER2 receptor to overproduce the receptors and so the cell will be highly sensitive to her2 ligand therefore will cause overgrowth of the cells by increasing the division.



In diagnosing breast cancer, the physician stains the HER2 receptors. If these receptors high in number, then the physician looking for any amplification in the genes to make sure what is the cause of that over production of that protein. Then the treatment plan is made due to DNA staining by either giving a patient medication that target the HER2 receptors called Hercepten (which is a drug that deactivate the receptors, and that results in decreasing the growth leading to decreasing the tumor size or cell death).

To summarize:

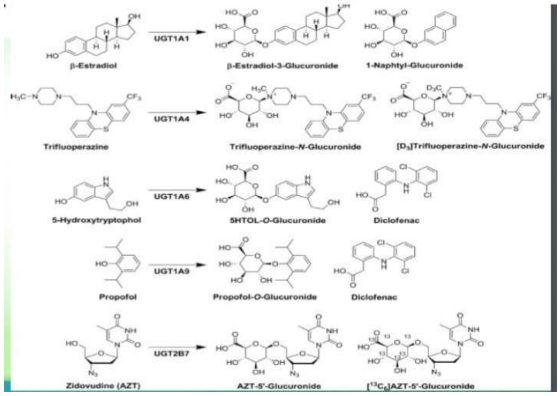
- It is an increase in copy number of a restricted region of a chromosome increasing the quantity of DNA in these regions.
- It is a mechanism that cancer cells use to escape resistance from methotrexate whereby the target gene, dihydrofolate reductase, is amplified.
- It is also a mechanism by which breast tumor cells progress and become more aggressive whereby they amplify the human epidermal growth factor receptor 2 (HER2), which stimulates cell growth.

2-UDP-glucuronosyltransferase (UGT):

When we take any type of a drug or ingest a product (poison, any chemicals, etc...) (a xenobiotic) from outside of our body. All these products will go to the liver and get modified. The modification is addition of a hydrophilic group onto xenobiotics to make elimination process easier. One of those enzymes that can modify xenobiotics is called **UDP-glucuronosyltransferase (UGT).**

It adds glucuronic acid onto xenobiotics making the drug or the chemical more hydrophilic.

Substrates
Etoposide
Genistein
Tamoxifen
PCBs
heterocyclic amines
Benzo[a]phrene
Nicotine
Raloxifene



These 2 pictures are not for memorization, just for your own knowledge

UGT can act on many substrates with different structures, chemicals or drugs. Moreover, the reaction between the enzyme and substrate are catalyzed in different tissues. **How**?

Get this concept first:

Will exemplify it as having a drill with different flutes, the drill do the same work regardless of changing the flutes, however the flutes select the end results of the drill like the hole sizes (bigger or smaller and so on).

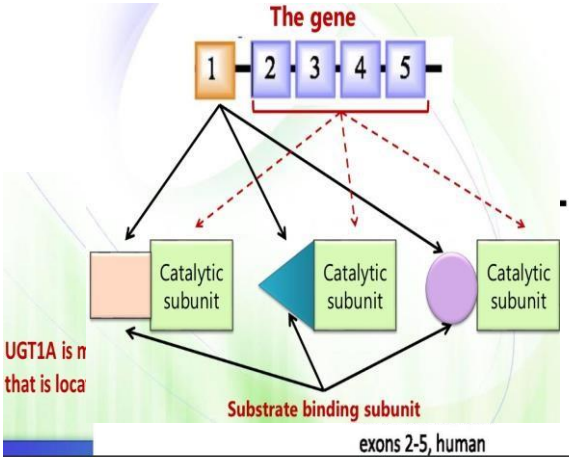
Another example here, these photos for the same person with different hats, when the person changes his hat, will change his look.

The same concept is applied for the UGT enzyme is by adding glucuronic acid on every substrate.

- It is a family of enzymes that is responsible for the glucuronidation of hundreds of compounds, including hormones, flavonoids and environmental mutagens.

Let's dig deep in genetic material and see how this happens:

In the diagram, the proteins that have made from the gene, they have the same catalytic subunit (which adds glucoronic acid), however the substrate binding site is different in different tissues for different substrate. If you see the gene structure itself, it has 5 exons; they are numbered from 1 to 5. The exons number 2,



3, 4, and 5 determine the catalytic subunit. Whereas, exon number 1 determines the substrate binding site.

The 2nd pic shows the exon number 1 closely. Surprisingly**, they found that is not a single exon (multiple exons – A1, A3, A4, A5, A6,**

A7, A9, etc…). Each of these exons has their **own promoter,** so the cell can pick and choose which exon would be suitable for it. For example, some cell chooses to have exon A1 connected with exon 2, 3, 4, 5. The other cell chooses to have exon A9, 2, 3, 4, and 5. In this way we end up producing proteins with **different substrate binding** sites and the catalytic site will be the same. They produce **different proteins isoforms in the tissues**, which act on different substrates or drugs.

It explains why it has many substrates with different structures and the reactions can be catalyzed in different tissues as well.

……

Regulation of mRNA stability

As we said previously, if the mRNA is stable for long period of time that means the proteins would be produced in a high amount. However, if we degrade the mRNA the protein that is produced from that mRNA will very low.

First example of regulation is **Iron metabolism**:

In our body iron is very important, so it gets stored. It should not be lost from our body. Due to that, our body produces enzymes that regulates the iron amount by producing some protein, they are: 1- **Ferritin** (protein found inside liver cells that stores the Iron).

2- **Transferrin receptor** (it is a sensor, senses the amount of iron in our cells and estimates how much of iron could enter the cell).
3- **Ferroportin** and **Diferrin metal transporter** (DMT1) (they are proteins found in the intestinal cells; they absorb iron).

We are going to talk about Ferritin and transferrin receptor.

When the iron amount **is high** in our body, the cell tends to store the iron and the cells stop taking up iron. The cell does this by **increasing the amount of ferritin** and decreasing transferrin receptors on the cells.

However, when the iron amount inside the cell **is low**, the cell decreases the amount of the ferritin protein and increases the transferrin receptor to get more iron inside the cell.

The scientists look at their genes (mRNA) of these proteins; they found that both of them have consensus sequences called Iron response elements (IREs). IREs in the mRNA molecule are like sensors which sense the Iron rate in our body. They are binding site for a protein called Iron response element binding protein (IRE-BP), which influencing protein expression of ferritin, transferring receptor, ferroportin, and DMT1.

Referring to the diagram below:

In transferrin receptor the IREs is in the 3' end of the mRNA. Also, it's found in the 5' end of the mRNA of ferritin. A) Iron deficiency: the IREs-BP can bind to the element of both mRNA but the effect that can cause is different.

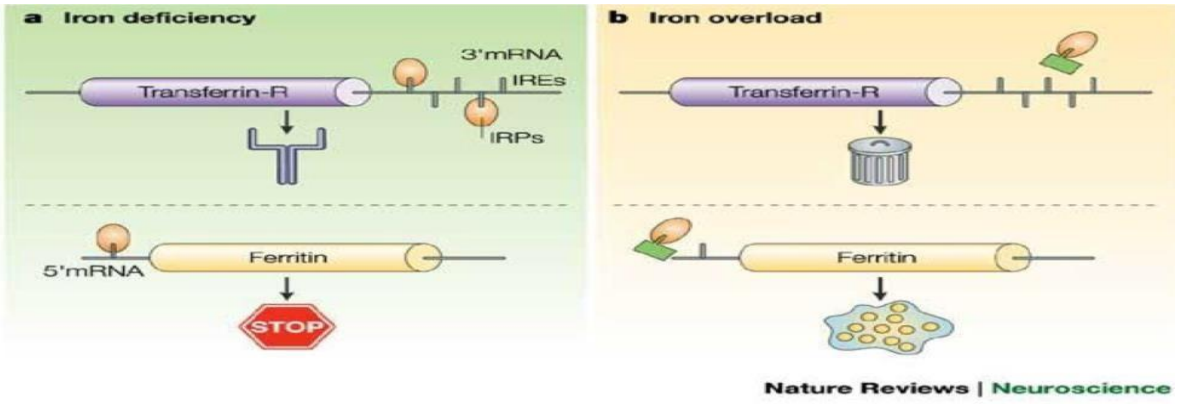
When it binds to the 5' end element of the **ferritin** mRNA, the translation from this mRNA is **blocked** so the production of the ferritin protein is stopped and the amount of the protein became low.

When it binds to the 3' end element of the transferring receptor of mRNA, the mRNA becomes more stable, lasts for a long time and producing more transferrin receptor that we need in our state.

So we are decreasing ferritin because we don't need to store iron at the moment, and increasing transferrin receptor to get more iron inside our cells.

B) Iron overload:

When we have big amount of Iron in our cells, the Iron can bind to the protein therefore it gets released from the 5' region in the ferritin mRNA, the translation resumes producing a lot of ferritin.
At the same time, it gets released from the transferrin receptor mRNA, the mRNA becomes unstable, therefore it gets degraded and the level or the amount transferrin receptor is lower than normal. Which makes sense when we have high level of iron in our cells; we do not need iron at the moment so we decrease the transferrin receptors to reduce iron that gets inside the cell. So, we store iron by ferritin. So we achieve harmony at the end of one of these processes.

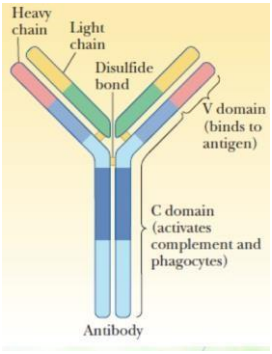


Second example of regulation is **antibodies**:

It is phenomena happen in B cells. They are immune cells that are responsible of producing antibodies or immune-globulins. Immunoglobulins are found on the surface of the cell and they recognize antigens. We have thousands of antigens in nature, does that mean we have to produce same number of genes to produce these immuneglobulins which is exceeding the capacity of our body to come

with this big number. Instead, our body increases the diversity of immuneglobulins that can recognize all different types of antigens.

The structure of immuno-globulin:



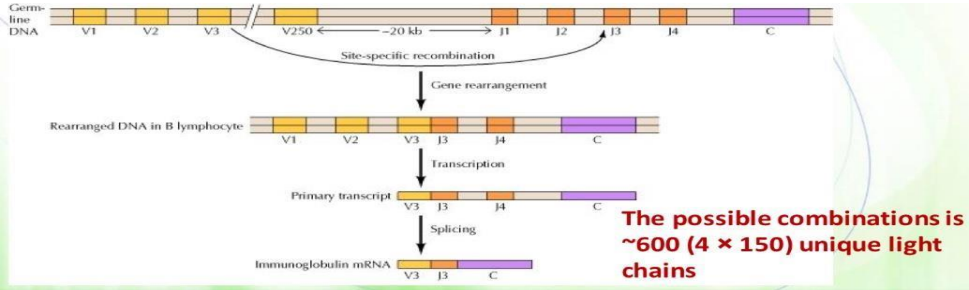
- It is a 4 polypeptide chain protein
- It is made of 2 identical heavy chains and 2 identical light chains.
- Both contain constant and variable regions.
- The variable (VL) regions (both heavy and light chains) are responsible for recognition of antigens. It has the ability to bind to many different antigens.

How is diversity generated by immuno-globulin?

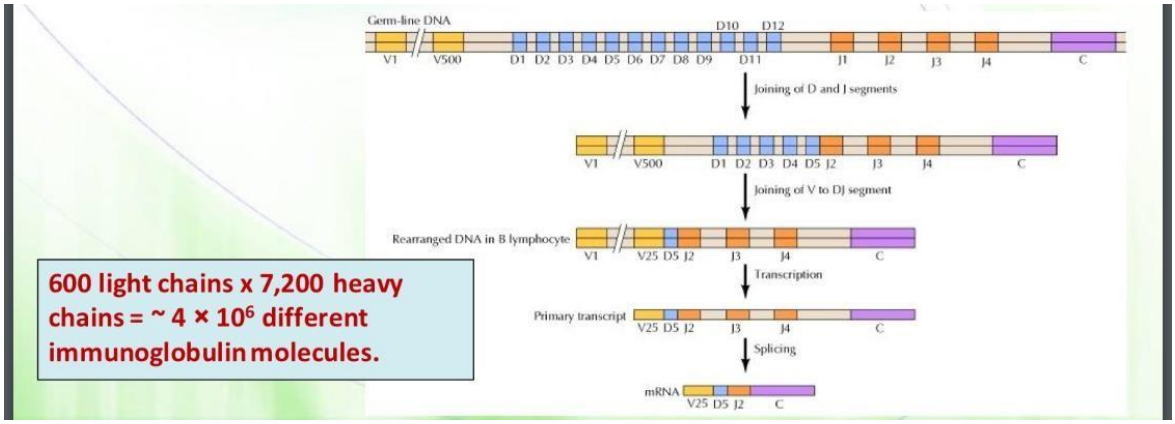
The phenomena called "gene rearrangement" or "gene recombination" of immune globulin genes specifically genes of light and heavy chains.

How?

The gene structure of light chain has 150 exons of V region and 4 exons of joining region. The V region (150 exons) can combine with any of 4 exons in joining region. So the possible combinations we can get is about 600 unique light chains.



The gene structure of heavy chain has 500 exons in V region, 12 different diversity exons in Diversity region (D), and 4 different joining regions. So, any of V region exons can combine with one of diversity region genes and joining region as well. The possible combinations we will get is about 7200 (500*12*4) unique heavy chains.



The antibodies in our body are designed for antigens. They are waiting antigens to enter our body and if there is any associations even a light association between the antibody with an antigen, the B cell will get stimulated (stimulation for cell division>DNA replication)

Having 7 millions of antibodies is not enough for the big number of the antigens, so our cells have additional mechanisms to increase the variation.

Such as：

1- During gene rearrangement, when an exon from V region combines with an exon from joining region is imprecise (the connection between them is not perfect), it could skip 1 or 2 nucleotides causing frameshift mutation. Having mutation will increase the diversity of antibodies (a different immune globulin)
2- Also, during recombination sometimes nucleotides could be **added** or **deleted.** From example: when it adds an exon from the end of the region to the beginning of another region, the addition or deletion will happen. (this is another type of frameshift mutation happens)
3- Somatic hypermutation: when the B cell gets stimulated for division and DNA replication. During DNA replication, the DNA polymerase is not very accurate and the process has to be very fast which results in the introduction of frequent mutations into the variable regions of both heavy-chain and light-chain genes. The resultant is some of B cells antibodies bind to antigens strongly (these cells would survive) and some other will have weak binding (these cells eventually will die). It is type of evolution as well.

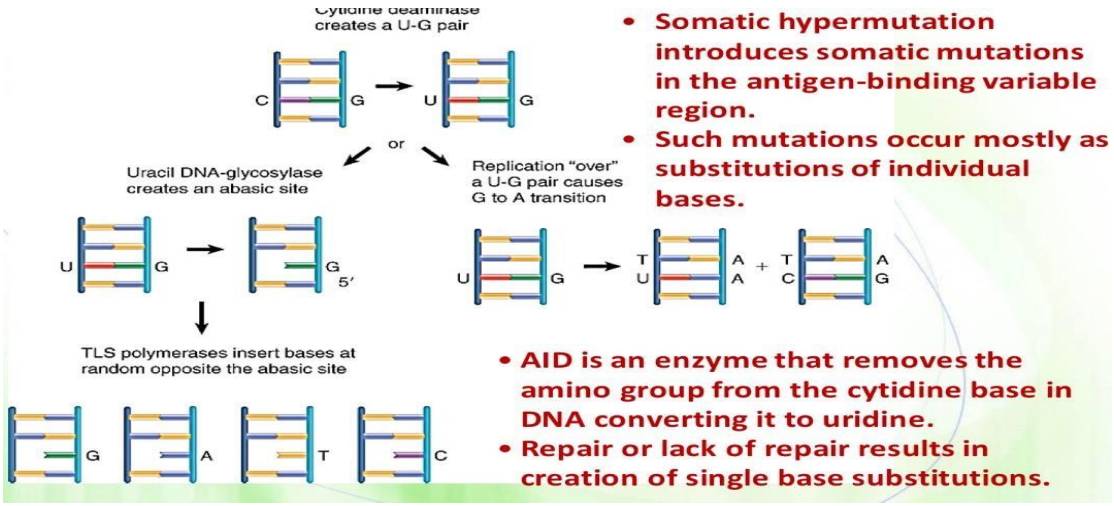
The last mechanism is

Activation-induced deaminase:

There is an enzyme called deaminase, this enzyme is induced when the B cells get activated.

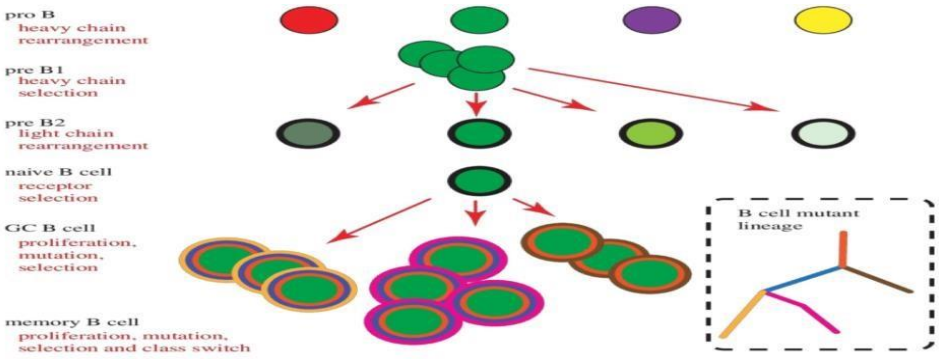
The enzyme starts to deaminate the cytosine, converting it to Uracil. Due to the process happens so fast there is no time for error correction. The Uracil can

stay in the DNA and it will base-pair with Adenine causing a mutation, this also resulting in diversity as well.



TO SUM UP:

1- B cells have antibodies, binding to them antigens (the green one below bound to one)
2- Antigens activate B cells to divide and DNA replication
3- When B cell activated more mutations in B cells (in the gene that produces the immune-globulins)
4- The mutation can cause a stronger affinity makes B cell survive and other mutation can reduce affinity makes B cell to die, that helps in evolution of antibodies.
5- The survived cells produce different types of B cell that is really efficiently recognize the antigens, these cells will be memory cells (stored in our body) afterward, when they are attacked with the same antigens will be activated again.



As you can see this sheet has lots of information, please accept my apologies if there is any mistake….

BEST OF LUCK...