University of Jordan
Faculty of Medicine

GENETICS & Molecular Biology

Number: 12

Done By: Reem Akiely

Corrected By: Rand Tarawneh

Subject: Gene Amplification, Diversity of Antibodies, Regulation of Transcription

Doctor: Ma'moun Ahram

Price:

Date:

**** This sheet was mainly written according to the record that belongs to pack 3, but some parts were written according to that of pack 1 in which these parts were a bit more clear.**

❖ Topics of this lecture:

*Gene amplification.

*Diversity among antibodies.

*Regulation of transcription in prokaryotes (slides 281-295).

*Regulation of transcription in eukaryotes (slides 296-304).

❖ Gene Amplification

A phenomenon that happens in cancer cells.

(but can also happen physiologically according to the record of pack 1)

It's usually a reaction to something; for example, some cells develop resistance to certain chemotherapeutic agents -like methotrexate- by amplifying certain genes that would increase the expression of a certain enzyme that would eliminate the effect of methotrexate so that cells become resistant to the chemotherapeutic agent. (First example on gene amplification).

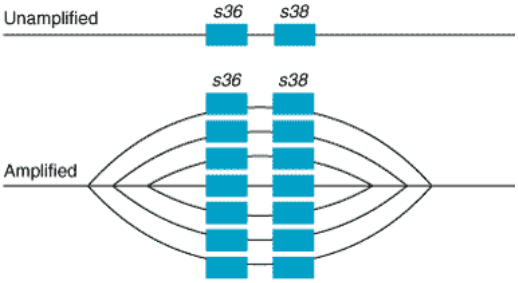
Gene amplification: a mechanism for increasing expression of a certain enzyme or protein by increasing the number of copies of a certain region of a chromosome, increasing the quantity of DNA in this region.

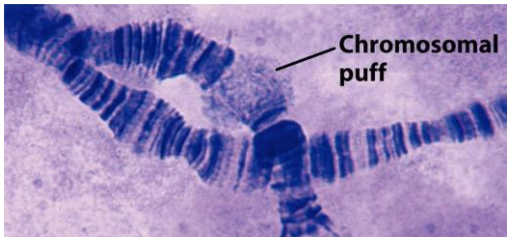
Gene amplification that occurs in breast cancer is also a good example:

- Breast tumor cells can progress and become more aggressive by amplifying the gene of Human Epidermal Growth Factor Receptor 2 (amplifying the **gene** of HER2).
- HER2 (the receptor) normally exists on the cell surface. *The function of HER2 is to stimulate cell growth when it binds to its ligand (a transcription factor)*. But when cancer cells amplify the responsible gene, they increase the number of these receptors on the cell surface leading to continuous growth and division of these cells.

- Look at the figure below to the left (gene amplification in cancer cells),
  - You can notice the location of the HER2 gene in the normal dsDNA. (The upper part of the figure).
  - What happens in cancer cells is that this region in the chromosome that contains this dsDNA becomes amplified meaning that this region would have **several copies** of the same gene, so rather than having 2 copies of the gene (one on each allele), each chromosome would have 10 copies –for example- of the same region with a total number of 20 copies that would all be active producing the protein.
  - From the figure, you can notice that we have several copies of a certain region of the same strand on top of each other, and each one of them contains a promoter, coding regions (exons), introns, and everything needed.

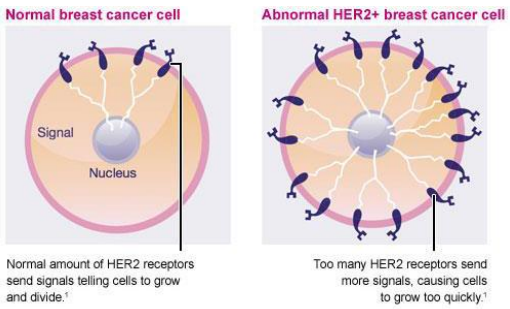
- Gene amplification results in what is called a chromosomal puff (انتفاخ) in the structure of the chromosomes. Chromosomal puffs can be seen when chromosomes are stained. (See the figure below to the right).





- Gene amplification is one way by which cells can increase expression of a certain gene.

- In cancer cells, rather than having a limited number of HER2 receptors, the cells would have large numbers of these receptors (the figure to the right), and these receptors stimulate cell growth when growth factors bind to them causing signal transduction pathway that tells the cells to keep on growing and divide continuously. These cancerous cells become really aggressive (advanced) because they keep on amplifying the gene forming tumor (and then cancer).



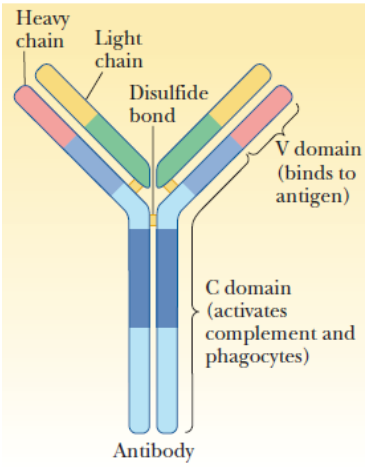
Note: the increased number of receptors –due to gene amplification- increases the sensitivity of the receptors to the presence of the ligand (even a little amount of the ligand can cause a high rate of cell growth).

❖ Formation of diversity of antibodies

(note: this section was written acording to the record of pack 1)

*A phenomenon that occurs in immune cells normally.

*Remember the structure of an antibody:



- two identical heavy chains and 2 identical light chains held together by disulfide bonds.
- Both (light and heavy chains) contain constant and variable regions.
- The variable regions are responsible for recognition of antigens.

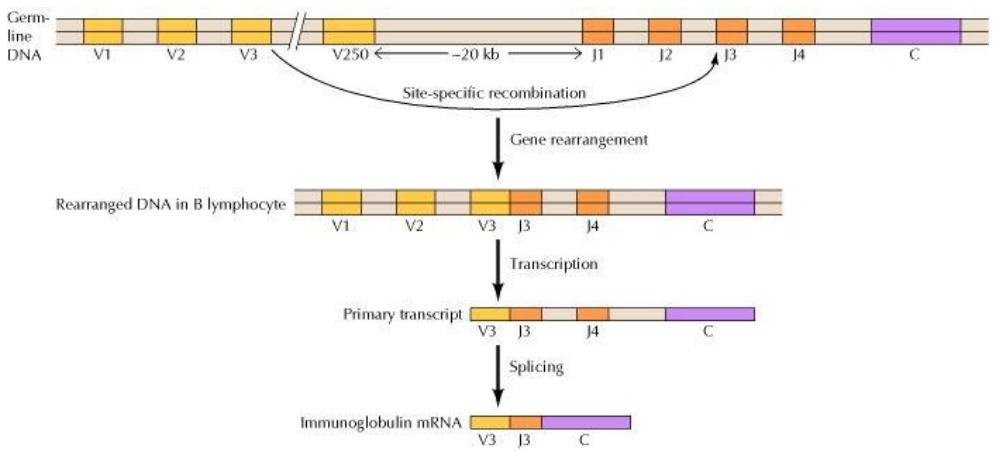
*Our bodies should make large numbers of antibodies because there are so many antigens. We're talking about 10^{11} different types of antibodies.

*How can the B-cell produce all of these different antibodies from only one genome?

By gene rearrangement (or genetic recombination).

1. Gene rearrangement of the light chain (the figure below)
   There are 2 types of immunoglobulin light chains; κ (kappa) and λ (lambda) depending on the constant region.
   Each light chain is a product of at least 3 genes:
   - Variable gene; there are 250 different genes (regions) to choose the variable gene from. (You can think of the 250 regions as 250 exons from which we'll choose one).
     → 250 possibilities in this location
   - Joining region gene; there are 4 different genes to choose from
     → 4 possibilities
   - Constant region gene: one gene; κ or λ → 2 possibilities

   The number of possible combinations we can have :
   250 *4*2 = 2000 (different combinations of light chains that can be produced)



2. Gene rearrangement of the heavy chain (the figure below)

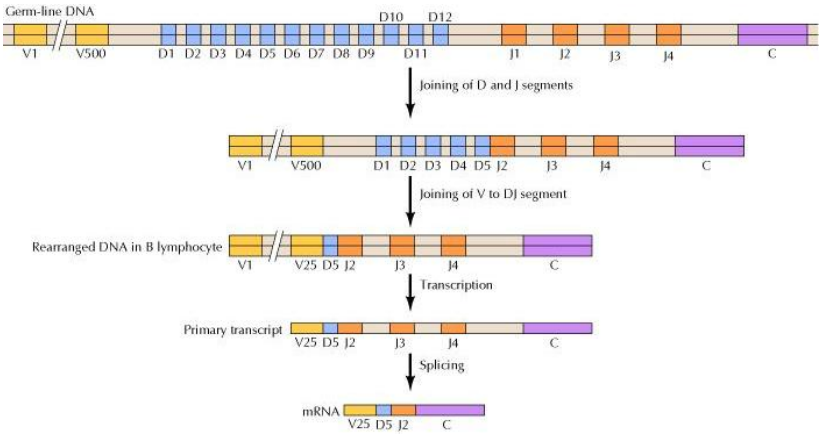
   There are 5 types of heavy chains depending on the type of the constant region that could be α, γ, δ, ε, or μ (mu).

   Each heavy chain is the product of at least 4 genes:

   - Variable region gene; there are 500 different genes → 500 possibilities
   - Diversity region gene; there are 12 different genes →12 possibilities
   - Joining region gene; there are 4 genes → 4 possibilities
   - Constant region gene; one gene; we can think of it as 5 types (α, γ, δ, ε, and μ).

   The number of possible combinations we can have :

   500 * 12 * 4 = 24000 *(note: when the Dr. was asked why we didn't multiply by 5 (5 possibilities for the constant region), he said that we can multiply by 5, but the numbers here are not as important as the idea, that is the reason behind the diversity)*.



Now, taking into consideration the possible combinations of the light and heavy chains together:

2000(L.chain) * 24000(H.chain) = ~ $5*10^7$ different immunoglobulin molecules.

<u>Is this enough??</u>

No, we need 10^{11} different types of antibodies. We need something more to increase the diversity, and this happens by different molecular mechanisms:

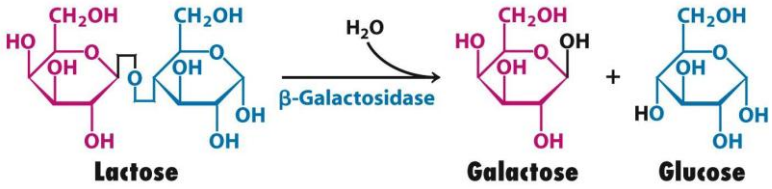
- Genetic recombination or gene rearrangement is not precise. For example, when the variable region combines with the diversity region in a heavy chain, the combination usually will not be very precise so we may lose 1, 2 or several nucleotides. This fact will make a huge difference since there will be addition or deletion of different amino acids that would change the structure of the variable region adding to immunoglobulins diversity.
- During recombination, the proteins of the genetic machinary can by themselves add or delete nucleotides, as a result, different types of variable regions would be produced.

- Further antibody diversity is generated by a process known as hypermutaion;
  We have B-cell carrying the antibody.
  - The antibody binds to the antigen, stimulating the B-cell.
  - The B-cell matures. During this maturation process, there is cell growth, which means also the presence of DNA replication. DNA replication has high rate of errors resulting in somatic hypermutations. (During DNA replication, different types of B-cells are produced –the antibody of each one of them differs by one or two amino acids or so- due to the different and random mutations that occur in certain regions).
  - Some of the formed B-cells will fail to bind to the antigen, so they die. Others that will have higher affinity to the antigen (even higher than the original cells) will get further stimulated and these are the ones that survive , grow, and fight the antigen.

  Note: there are two types of cells in our bodies; germ line cells (eggs/ sperms), and somatic cells (every thing other than germ cells. Somatic cells are diploid)

❖ Regulation of Transcription in Prokaryotes

- Concerning this subject, one example will be discussed to give an idea on how expression is regulated.
- Bacteria usually utilize simple sugars; they can utilize glucose and also galactose. Lactose is a disaccharide composed of glucose and galactose.
- Bacterial cells produce an enzyme known as β-galactosidase that cleaves the glycosidic bond producing glucose and galactose that can be then utilized for production of energy.



- In the 1950s or 60s, two French scientists (Francois Jacob and Jacques Monod) studied what is known as The Lac Operon.

> Remember: an operon is a genetic system that exists in bacteria and is composed of multiple regions of the same genetic unit that produces one mRNA and this one mRNA produces several polypeptides with different functions. One example is the trp operon -was discussed in one of the previous lectures- that is responsible for the **production** of tryptophan. Another example is the lac operon.

The Lac Operon

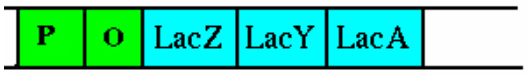
- It is an operon that produces the proteins that are responsible for the **metabolism** of lactose.
- The lac operon can produce 3 proteins:
  1. β-galactosidase: cleaves the glycosidic bond producing galactose and glucose. Glucose mainly can be used for generation of energy.

  2. Lactose permease: a transporter on the cell surface that takes lactose into the cell if it's outside.

  3. A transacetylase: modifies the galactosides (The modification is acetylation). It acetylates β-galactosides.

- The 3 genes responsible for the previous 3 proteins exist in the lac operon.

- **The structure of the lac operon:**



  - Lac**Z**: the part of the operon that produces the galactosida**S**e.
  - Lac**Y**: the part of the operon that produces the perm**ea**se.
  - Lac**A**: the part that produces the trans**A**cetylase.

  → 3 different proteins can be produced from this operon.

Upstream of the Lac Z, Y, A, we find:

- P : the Promoter (the binding site of the RNA polymerase upstream of the starting point of transcription)
- O: the Operator

The operator:

- A regulatory sequence that regulates if the gene should be active or not, and how active it should be (the operator balances the expression of the gene).
- The operator is the binding site of another protein called the Repressor (the lac repressor).

The lac repressor:

- Also called the *i protein* is produced from a different gene that has its own promoter (a gene that is not related to the lac operon) , the gene is called "the lacI gene" (or the *i gene*). (I stands for inhibition)
- The lacI gene is not an operon because this genetic unit would produce one polypeptide only.
- The lac repressor (*the i protein*) binds to the operator (*the o region*) and it represses (blocks) transcription by blocking the movement of RNA polymerase (RNA polymerase will not bind to the promotor).

A student's question: "In the previous figure, why is the Pi (the promoter of the lacI) located to the right of the lacI gene although we said before that a promoter should always be upstream of its gene?"

The promoter should be upstream for sure, and since the promoter is on the right, this means that the strand that is transcribed is the opposite strand compared to the one transcribed in the case of the lac operon. (Remember the figure in slide 234 from a previous lecture).

- There are two types of regulation; positive regulation and negative regulation.

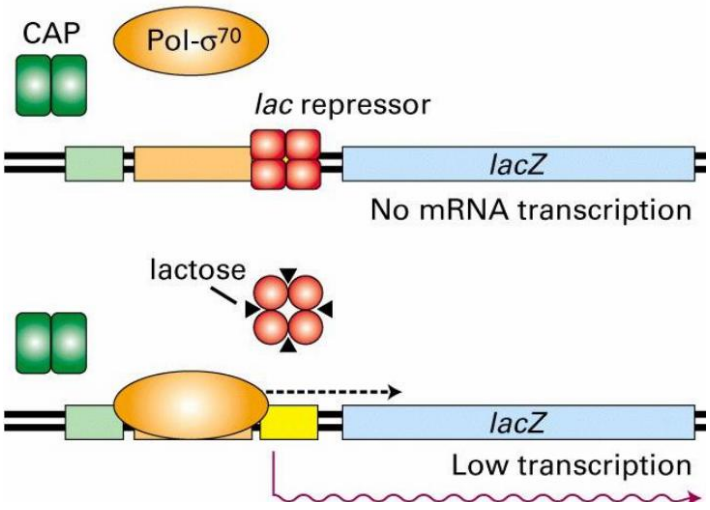
Positive regulation

- Controlled by lactose.
  This actually makes sense. If there is lactose, we should have gene expression (we mean expression of the lac operon).
  If there is no lactose, there should be no expression.

- Regulation by lactose is **positive** because the presence of a certain molecule –here, it is lactose- **stimulates** (increases) gene expression.

- What happens in the presence of lactose? (the lower part of figure below)
  Some lactose molecules bind to the repressor (that usually binds to the operator) and they release the repressor from the operator in the promoter region by changing the structure of the repressor.
  - Now, RNA polymerase can bind to the promoter region and transcribe the gene → "induction of expression". This is positive regulation.

- What happens if there is no lactose? (the upper part of the figure)
  The repressor will stay sitting on the operator preventing RNA polymerase from binding to the promoter and thereby preventing the induction of expression.



Note: the repressor is always there. It's not regulated at the molecular level (at the DNA level), but it's regulated at the protein level.

Before moving to negative regulation, two important concepts must be discussed:

Cis. Vs Trans regulatory elements

- In terms of molecular regulation, we can think of the mentioned regulatory elements as cis and trans.

Cis-acting elements:

- DNA elements (sequences) that control expression of the gene located on the same strand. Cis means "same" → same level
  Example; the promoter and the operator control the expression of the gene **on the same strand**. Another example is something called "the enhancer" (will be discussed in the next section in this sheet).
- If we remove the cis-acting element, for example, if we remove the promoter region of the lac operon and place it somewhere else on a different piece of DNA, nothing will happen. There will be no regulation because it should be on the same piece of DNA with its operon.
  - The promoter is known as a Cis-acting element

Transacting regulatory elements:

- These include the I repressor because it's a protein that can bind to the operator region regulating the expression of the gene.
- Let's say that we removed the gene of the lac repressor and we placed it on another piece of DNA (a plasmid for example), so now it's not part of the same DNA that contains the lac operon, can the repressor still control expression of the lac operon although it exists on a different piece of DNA?
- Yes, because wherever the lac repressor gene is (even if it's on a different piece of DNA), it will produce the repressor protein that will bind to the operator of the lac operon.
- The repressor gene is known as a transacting element because it's basically a DNA sequence that can control expression of a gene on the same piece of DNA or different fragments of DNA.

Effect of mutations

The 2 scientists introduced mutations to the DNA to see the effects of the mutations and found out that there are 2 types of mutations.

Mutations that result in Constitutive expression.

- Constitutive expression means that it's always on.
- Concerning our example (the lac operon), what kind of mutations in the DNA can be introduced to make the gene always on?

  - Deletion of the lac repressor gene.
  - Mutate the operator so that it can no longer bind to the repressor.
  - Mutate the promoter of the repressor gene (the i gene) in a way that would prevent the production of the repressor protein.
  - Mutate the i gene itself so that it produces a defective lac repressor that cannot bind to the operator.

Mutations that result in Noninducible gene

- The gene is always off no matter what.
- Examples on such mutations:

  - Mutate the promoter of the lac operon.
  - Mutate the repressor in a way that would make it always bind to the operator.
  - Mutate the lactose binding site of the repressor so that even if lactose is added, it won't bind to the repressor. (the repressor stays bound to the operator inhibiting transcription).

Negative Regulation

- If we have bacteria in the presence of lactose and glucose, which of them do you think bacteria will prefer?
  Glucose, because it's easier to utilize (in the case of lactose, bacterial cells would need to cleave it to produce glucose, why to do so if glucose is already there?)
  ➔ This means that if glucose is present, bacterial cells will not need to express the gene of galactosidase.

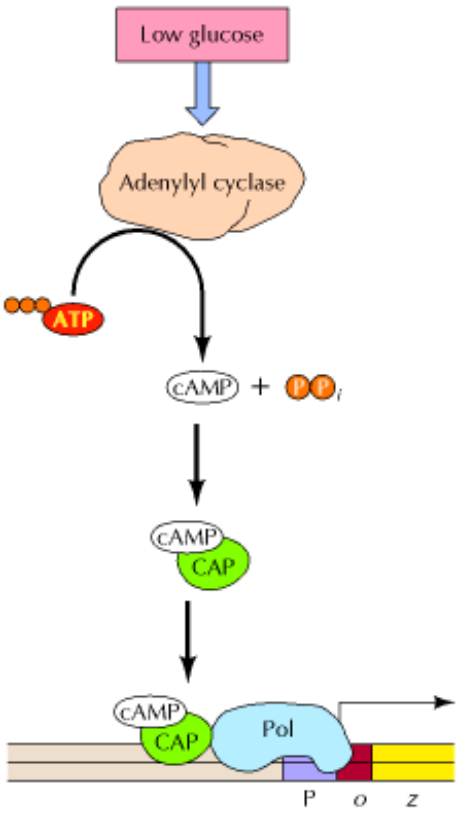
- Glucose can regulate expression of the lac operon negatively (when glucose is present → no expression. This is negative regulation).
- When glucose level goes down, the lac operon should be expressed (we should have stimulation, "positive regulation").

- How does glucose regulate the expression of the lac operon? (Follow the figure from bottom to top).

  ➢ The RNA polymerase is active if there is no repressor and there is lactose but it still needs a push forward to move (need something to make it more active). What it needs is a protein known as CAP (catabolite activating protein).
  CAP binds to the polymerase upstream of the promoter and pushes the polymerase forward inducing (stimulating) expression by making the polymerase more active.

  ➢ Catabolite activating protein is regulated by cAMP (small molecule, a second messenger). When cyclic AMP binds to CAP, CAP then binds to RNA polymerase and activates transcription.

  ➢ Cyclic AMP comes from an enzyme known as Adenylyl Cyclase (or Adenylate cyclase). Adenylate cyclase converts ATP into cAMP. When cAMP is present in high concentration, it binds to CAP...



- Adenylate cyclase is regulated by the level of glucose. Low glucose level means that the enzyme is active and vice versa.

- To make things more clear, start from the level of glucose:

  a) **Low** level of glucose;
  → Adenylate cyclase is active
  → cAMP is produced
  → cAMP binds to CAP
  →CAP binds to RNA polymerase
  →RNA polymerase activates transcription.

  b) **High** level of glucose
  → Adenylate cyclase is not active
  → the level of cAMP is low
  → cAMP can't bind to CAP
  →CAP is lonely and can't bind to RNA polymerase
  →RNA polymerase is active but not that much.

- The possible 3 cases concerning the presence of lactose and glucose (the figure in the next page):

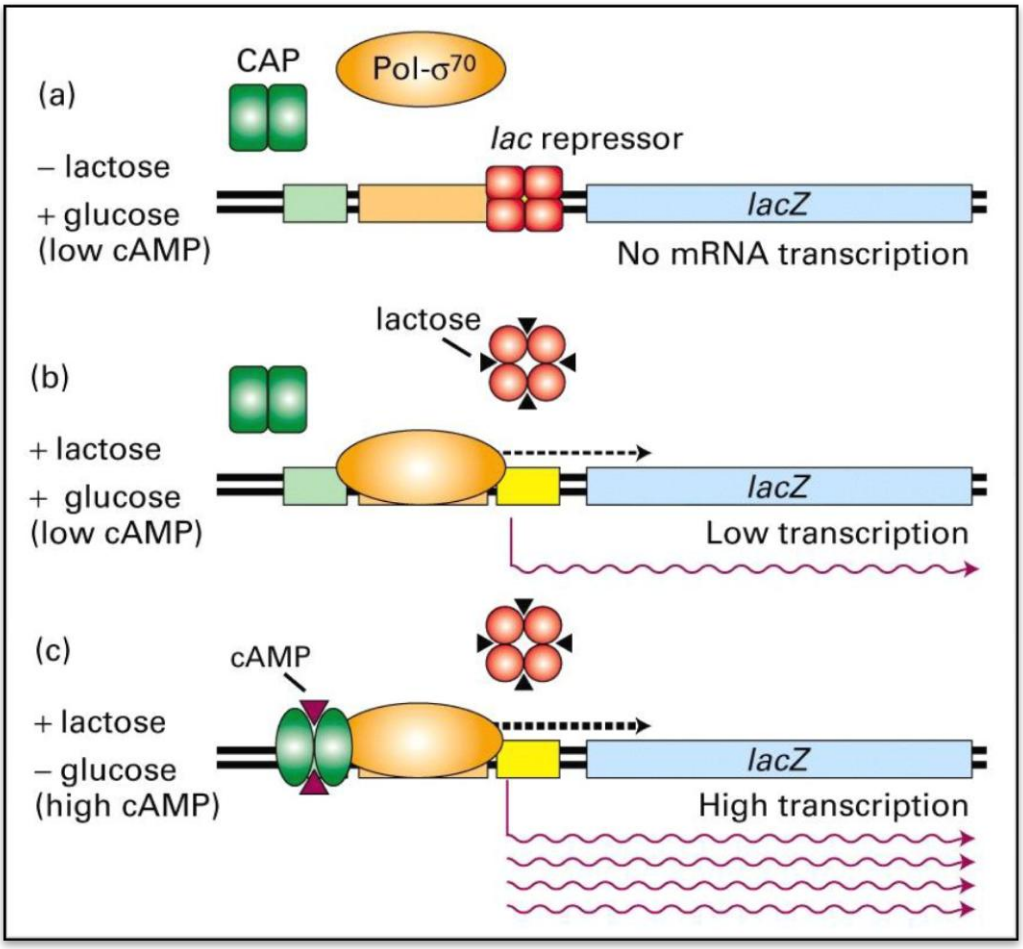
  (a). Glucose is present, lactose is not:
  No need to produce the galactosidase.
  CAP is not active because there is no high level of cAMP.
  The repressor is bound to the operator, RNA polymerase is not active.

  (b). Both, lactose and glucose, are present:
  The lac repressor is not bound, but CAP is not active.
  The polymerase can produce some transcription, but not a lot (low transcription). →low level of expression of the lac operon.

  (c). Lactose is present, glucose is not:
  The repressor is off.
  The CAP is really active now because of the high level of cAMP.
  The polymerase is highly active (high transcription).

- There is an animation link in slide 294 if you want to watch it.
- There is also a figure that you might find useful in slide 295.

*Note: in the exam, some of the questions about the lac operon will be direct, but others will not, and will be based on the concept of regulation of the lac operon. A piece of advice from Dr. Ma'moun: there will be this **indirect simple** question about the lac operon, but it will be a little bit **long**. If you are a fast reader, answer it directly. If you are a slow reader, leave it till the end.*



❖ Regulation of Transcription in Eukaryotes

- More complex than that in prokaryotes.

- Transcription in eukaryotic cells is controlled by:

1) Cis- acting DNA sequences (similar to the operator and promoter in bacteria).

2) Transcriptional regulatory proteins.
Here, we are not talking about the general transcription factors, we are talking about specific regulatory proteins.

3) Repressor proteins that also regulate transcription.

4) Epigenetic (not genetic) modification or control.
Epi means higher or upper. Epigenetic modificattion is a higher level of modification that does not involve changing the DNA sequence. (e.g. methylation).

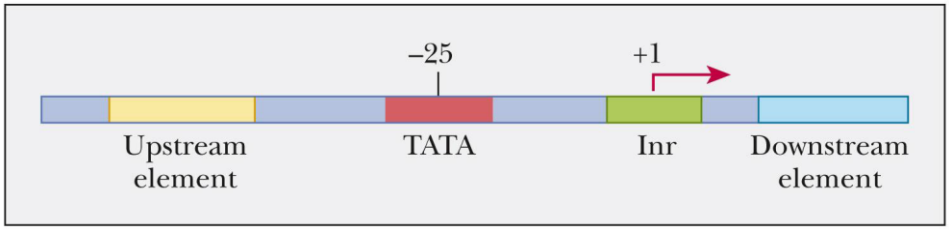
(1) Cis-acting Elements

- Two types:
  1) Promoters that have their own consensus sequences similar to TATA box.
  2) Inhancers

- The promoter region in eukaryotes is a little bit different from that in bacteria, but there is similarity as well.

- Components of the promoter region in eukaryotes:

  - TATA box: the binding site of RNA polymerase.
  - Upstream elements: these are called enhancers (if the induce expression) or repressors (if they repress expression).
  - The initiator element (Inr) that surrounds the +1 site.
  - Some (not all) genes have a downstream element.



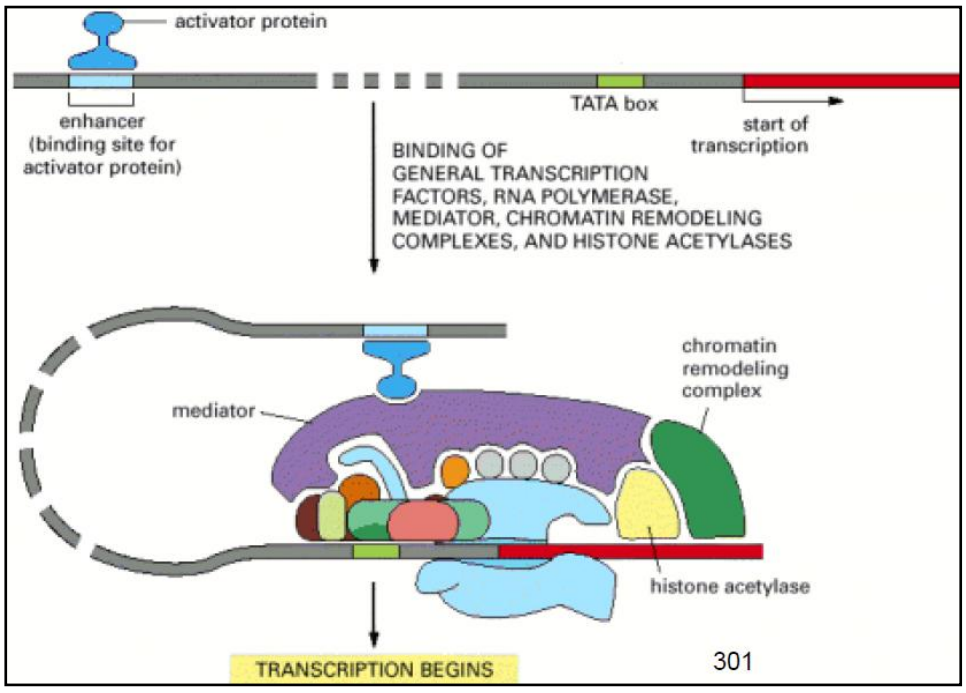
What are the enhancers?

(note: this part about enhancers was written according to the record of pack 1)

- Enhancers are sequences of DNA that are usually located upstream of the promoter region.
- These enhancers bind to gene-specific transcription factors. (we are not talking here about the general transcription factors like TFIID, TFIIF,.. We are talking about gene-specific transcription factors that target certain genes because these genes contain certain sequences).
  Note: *these gene-specific transcription factors are also called transactivators or coactivators.*
  Note 2*: you can think of CAP that we discussed before as a coactivator. CAP by itself does nothing but it activates the RNA polymerase so it's a coactivator.*
- Enhancers are cis-acting regulatory elements.
- One of the main characteristics of these enhancers is that they can cause looping of the DNA so that the enhancer can bind to the promoter complex. *(DNA is flexible so it can loop)*

- Look at the figure below and pay attention to the following points:

  - You can see the promoter (which has the TATA box)
  - RNA polymerase and all other transcription factors can bind to the promoter region except that they are not activated (they need a push).
  - This push comes from an activator protein that binds to the enhancer region. *Note that the enhancer can be far away from the complex.*
  - Because of this feature of DNA known as DNA looping, DNA can loop and the part with the activator protein can touch the complex activating it so that now it can move forward.
  - One main feature of enhancers is that if we take away the enhancer and change its location putting it downstream of the gene (in other words, if we invert it), it can still be functional. The same activator protein can bind to the DNA, the DNA can loop and the complex can be activated.



Note: *if it was a repressor instead of an enhancer, it would catch the complex keeping it in its place, and no transcription will take place.*

Now, let's talk about the activators (the gene specific regulatory proteins)

(2) Gene-specific Regulatory Proteins

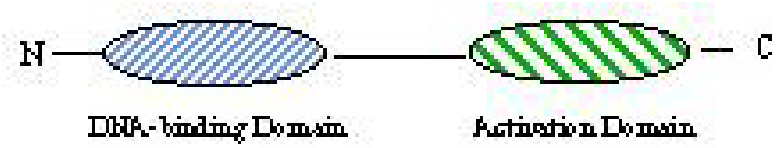
(the activators or the transcriptional regulatory proteins)

- They are composed of -at least- two different domains that are independent of each other:
  1) DNA binding domain
     Binds to specific DNA sequences (binds specifically to the enhancer region or to the regulatory sequence).

2) The activation domain

The domain responsible for the protein-protein interaction as it binds -also specifically- to other proteins as discussed earlier.

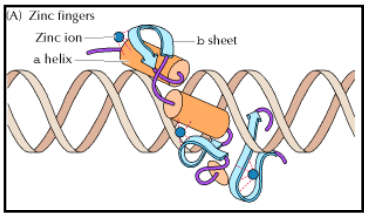
- Both domains (activities) are independent and can be separated from each other.



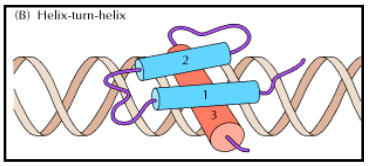
Remember:
A domain is a compact three-dimensional region of a protein that can fold by itself independently of the rest of the protein.

A) There are different structures of DNA binding domains (the Dr. said not to worry about the details)

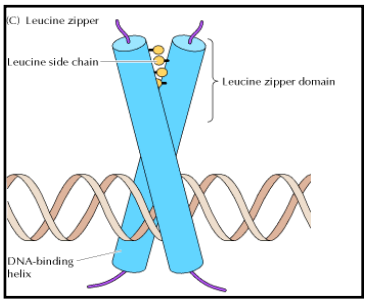
- Zinc finger domains
  These are present in steroid receptors like estrogen and androgen receptors.



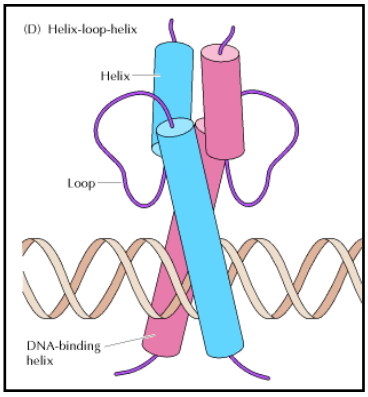
- Helix-turn-helix motif



- Leucine zipper
  Which is present in (specific for) a protein known as CREB , a protein the Dr. will talk about in the upcoming lecture.



- Helix-loop-helix



B) There are also different types of activation domains

(Note: Dr. Ma'moun did not discuss the following but told us to study this slide)

- Acidic domains
- Glutamine-rich domains
- Proline-rich domains

Activation domains are thought to stimulate transcription by interacting with general transcription factors, such as TFIIB or TFIID, thereby facilitating the assembly of a transcription complex on the promoter.

I apologize for any mistake I may have made.

Wish you all best of luck **:D**