Doctor 022

Biochemistry

Sheet no. 2

RECOMBINANT DNA-BASED MOLECULAR TECHNIQUES (PART II)

Writer: Al-Razi Node

Corrector: Al-Razi Node

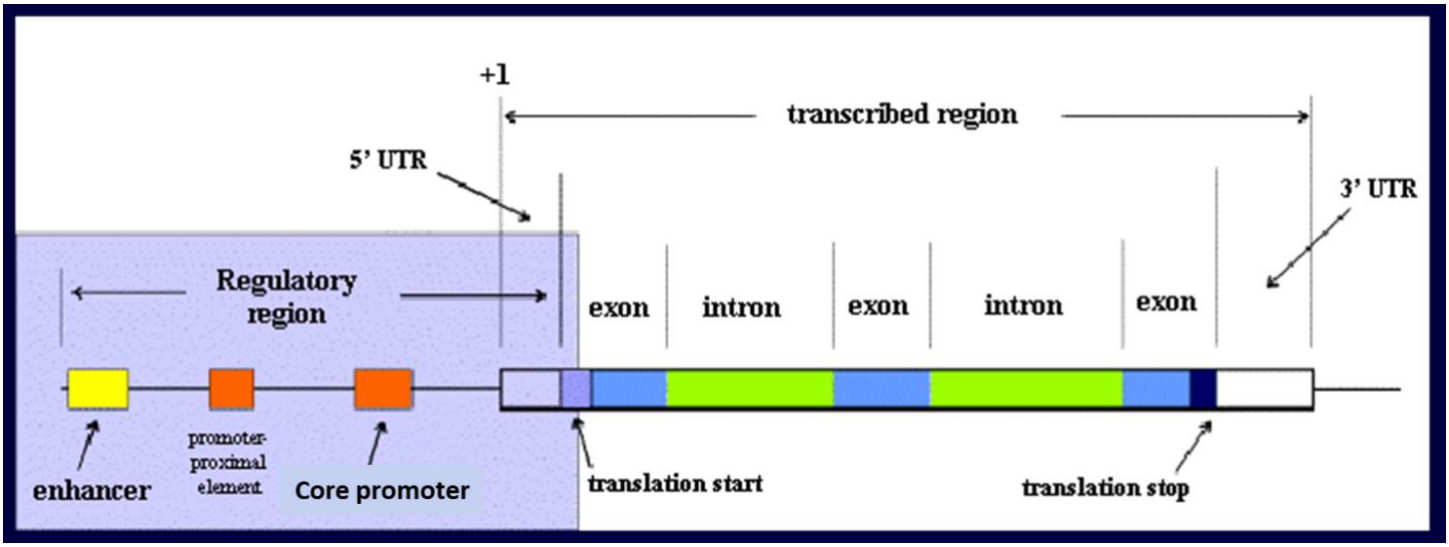
Doctor: Dr Mamoun

RECOMBINANT DNA-BASED MOLECULAR TECHNIQUES (PART II)

In this part, we will talk about techniques that we use them to analyze DNA regulatory sequences, and techniques related to investigating protein-protein interaction.

ANALYSIS OF TRANSCRIPTIONAL REGULATORY SEQUENCES: ROLE OF ENZYMES

What are transcriptional regulatory sequences?



Here we have a eukaryotic gene that contains: transcriptional start site (at +1 site), introns and exons.

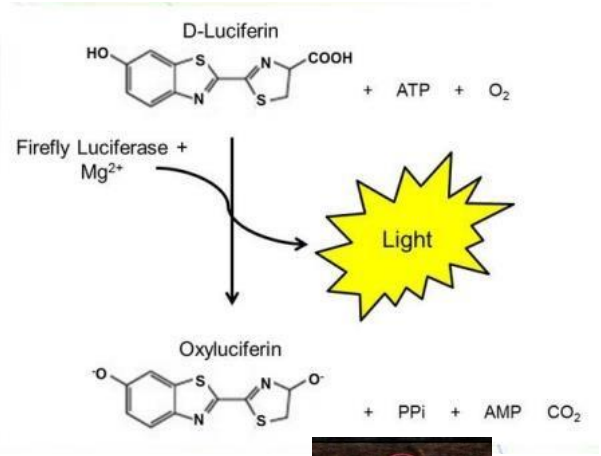
- There are also non-coding sequences which regulate the activity of gene:
  - the basic promoter region (the core promoter) contains the RNA Polymerase binding site and regulatory sequencing, then coming the RNA polymerase and bind in the core promoter and make the mRNA. The mRNA will do on it 3 major processes: -
    A. addition of A cap on the 5' N.
    B. addition of poly A tail on the 3' N.
    C. splicing to remove the introns and connect exons together.
  - promoter proximal element (gene specific) any regulatory sequence in eukaryotic DNA that is located close to (within 200 base pairs) a promoter and binds a specific protein thereby modulating transcription of the associated protein-coding gene. And it should to be near the core promoter and it make the regulatory function.

- elements or sequences that are regulatory DNA sequences that, when bound by specific proteins, regulate the transcription of an associated gene. They can be located near, within, after, and/or very far away from the gene, and, if lipped or relocated, are still functional. Positively regulate gene expression (enhancers) and other sequences that negatively regulate gene expression (silencers).

  How can we investigate these sequences? how can we pinpoint specific elements that regulate gene activity? how do we know if it is a positive/negative regulatory element?

FIREFLY LUCIFERASE

Scientists investigated why the bottom of these flies fluoresce (produces light) at night, it turns out that thereis a molecule called Luciferin which is converted into Oxyluciferin, that produces light by an enzyme called Luciferase



(this is the basic reaction in producing this light).
Scientists have taken advantage of Luciferase and they have used the recombinant DNA technology as follows:

LUCIFERASE REPORTER ASSAY

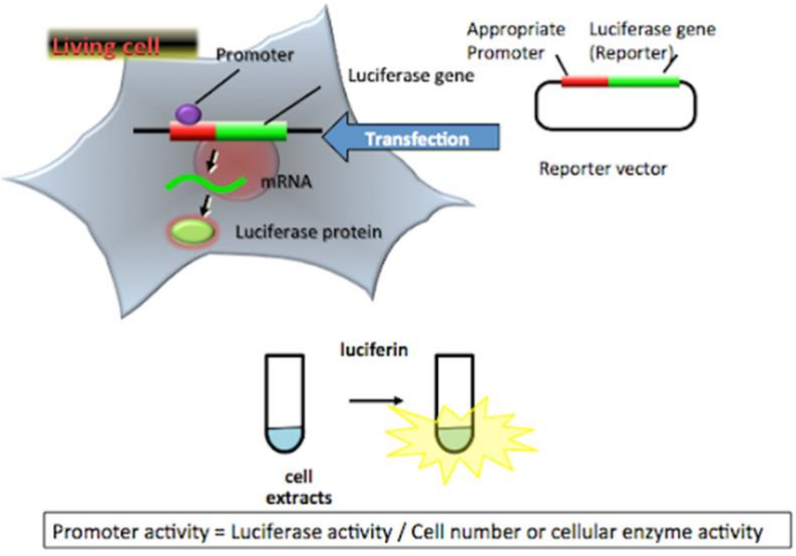
- **Purpose: study the activity of a gene at certain conditions or identify the function of certain regions of the promoter**

A reporter gene: it is a gene that gives us a signal (an idea of something), and here we want to know the activity of a certain region inside the cell under certain conditions or the importance of several regions, in terms of regulating gene expression.

- Only the regulatory region (e.g. promoter, PPE, etc.) of the gene is placed upstream of a “reporter gene” such as the luciferase gene in a plasmid. So, the luciferase gene is under control of the promoter region of the gene that we want to analyze, so we are concerned about the regulatory element that controls the gene expression not the gene itself.
- The plasmid is transfected (inserted) into cells, and the expression level of luciferase (instead of the original gene itself) is measured.

The regulatory proteins that control the expression of the gene of interest would control the expression of the luciferase. If This gene of interest, is at certain condition where it should be expressed, instead of expressing that gene we have the luciferase gene expressed instead.

we add the luciferin molecule to the cells, if the gene of interest is highly expressed it means that the luciferase gene would be expressed instead, so more luciferin is converted, more luciferase is produced, the color is stronger.

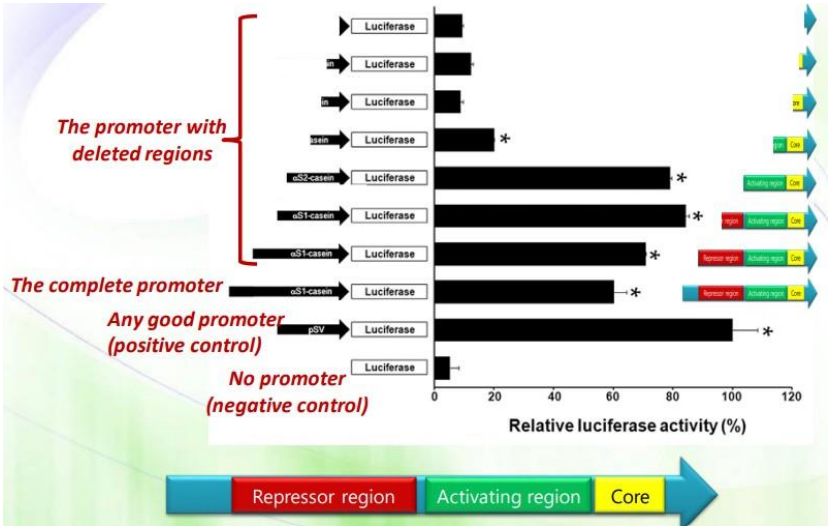


So here we are reporting:
the more the luciferin is converted, the more luciferase produced, the stronger the color, and vice versa.

So, we put cells under certain conditions, and we measure the amount oflight that these cells produce. So, for example, this condition actively regulates the expression of the gene of interest according to how much luciferase is expressed, or that other condition negatively regulates the expression of the gene of interest.

❖ In this example below, it shows the results of specific experiment.

We have the transcription start site of the luciferase gene, then there are the regulatory sequences of the gene of interest:

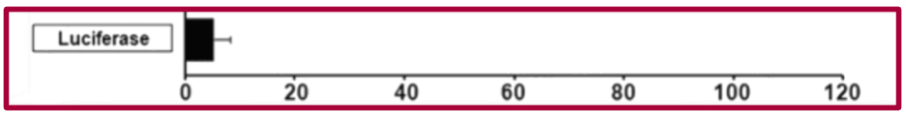


- The core promoter region: the RNA polymerase binding site that produces the basal (minimum and standard)expression of a certain gene.
- The activating region: it is a positive regulatory element.

- Repressor region: it is the negative regulatory element.

Again, we take cells, and the plasmid is transfected into them, and it contains the luciferase gene but without any promoter, so there is no expression of the luciferase gene, and this is called negative control where there should be zero or minimal expression of luciferase gene.

➢ In the picture, we see little signal, and this could be leakage of expressionor just a background color.



➢ We also have <u>a positive control</u>; this is a sample where a luciferase gene is under control of a good promoter (not the promoter of gene of interest) so we get the maximal expression of the luciferase (the color is about 100%.)

- We have <u>the complete control</u> when the luciferase gene is under control of promoter of the gene of interest, and we get a specific signal (60%).

- Then, we start to remove more of the repressor region, so we notice an increase in the sample compared to the complete promoter. This increase tells us that the region we just removed contains a repressing element.

- After that, we remove the activating region so we notice a huge drop in the expression of the luciferase gene. This tells us that there is an activating region; an element that positively regulates the expression of gene of interest.

- Finally, when we remove the core promoter as well, there is no expression, and we can compare this condition to the negative control.

- This indicates what certain regions in the prompter represent. If they are activating or repressing regions.

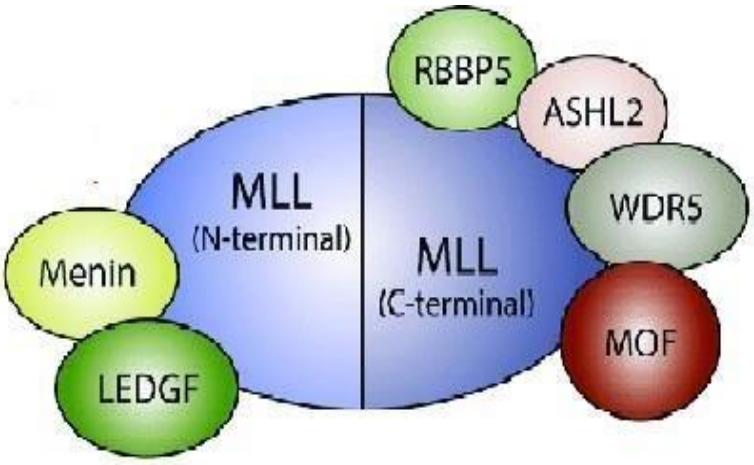
PROTEIN-PROTEIN INTERACTION,

(CO-IMMUNOPRECIPITATION), (YEAST TWO-HYBRID SYSTEM)

STARTING FROM A DNA LIBRARY:

proteins don't act by themselves; they mainly interact with other proteins in order to produce an effect on cells.

Here in the figure, we have a protein contains two domains, one exists at the C - terminus, the second exist at the N - terminus, each of which interact with different proteins, (notice that MOF doesn't react with RBBP5), so these protein-protein interactions are complex, and here comes the great benefit of the techniques:

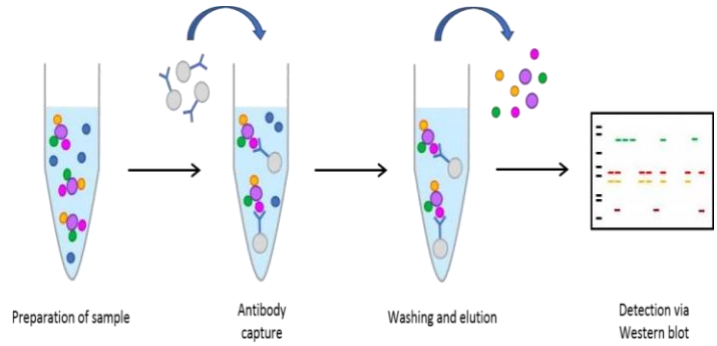


(co -immunoprecipitation), and (yeast two-hybrid system).

CO-IMMUNOPRECIPITATION:

It differs from immunoprecipitation that here we are not putting down only the protein of interest, rather we are putting down all the other proteins that interact with our protein of interest, and what is done is as follows:

- **Antibody molecules that target a specific protein are conjugated to special beads.**
- **A mixture of cell proteins is added to the beads.**
- **Only the protein of interest is precipitated as well as other proteins bound to it (co-precipitated).**



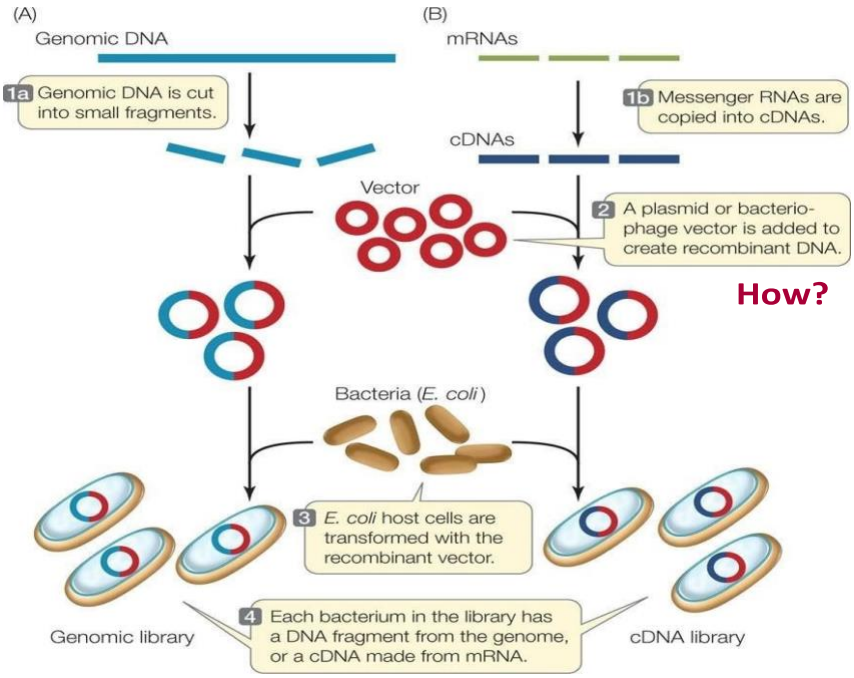
- Side note: after we finish, we can study our sample by immunoblotting or SDS-page, (immunoblotting by the way is also known as western blot), so we have southern blot for DNA, northern blot for RNA, and western blot for proteins.

WHAT IS A DNA LIBRARY?

- **A collection of DNA fragments found in plasmids and ever plasmid we put them in the tube and then we put them in the freezer.**
- **A library can be created for DNA fragments just like book libraries.**
- **You can have clones of bacteria each containing a specific piece of DNA** (like a catalogue).
- **You can save these clones in the freezer and take whichever clone you want to study.**

You can watch the video:
http://www.sumanasinc.com/webcontent/animations/content/dnalibrary.html

We have two types of libraries, a **genomic DNA library** and a **complementary DNA library**, the figure below concludes them all:



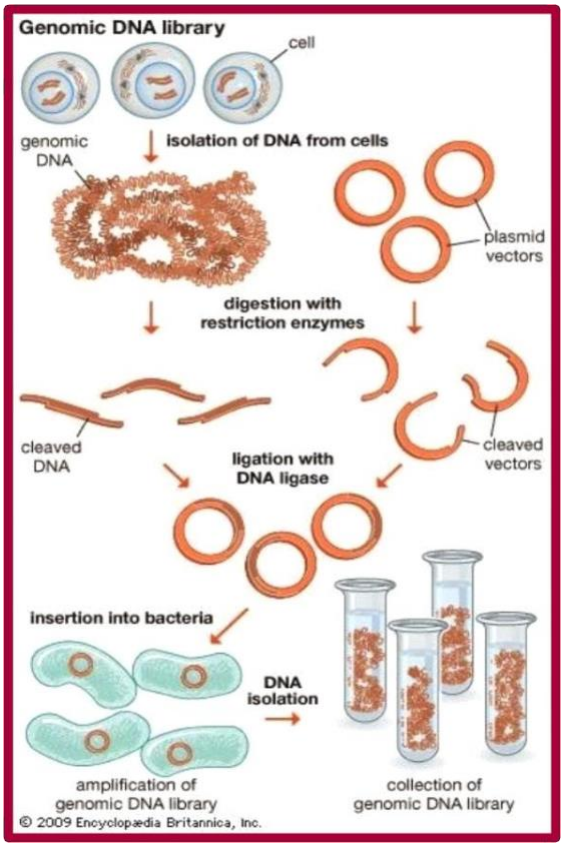
Side note: inside this genomic DNA library we have every single thing, we have introns, exons, motor regions, enhancers, coding regions and non-coding regions.

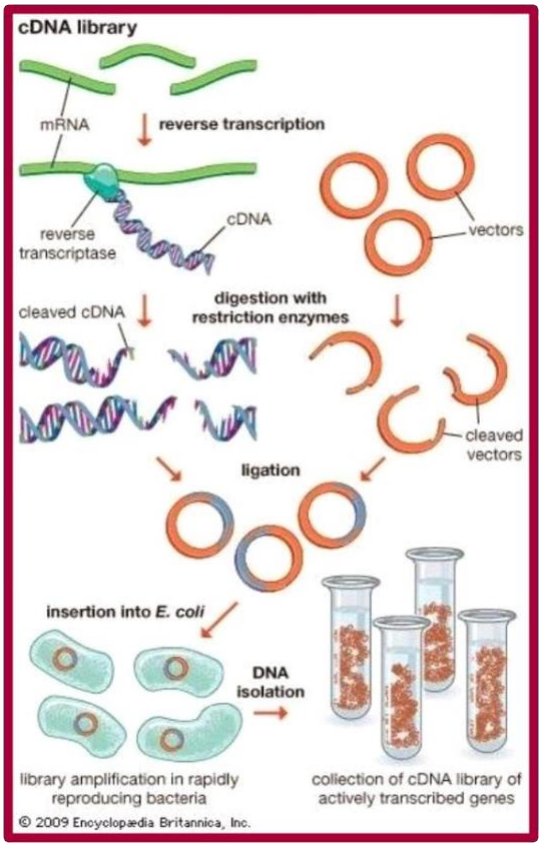
Now regarding cDNA library, it's cleaner, simpler. So, we take mRNAs from cells, we convert them into cDNA by reverse transcriptase, and we insert these DNA molecules inside plasmids and continue it by yourself from the figure, and for god's sake 😂, keep in mind that in cDNA library

there are no introns, nor promoter regions, or enhancers, it contains just exons, poly A tail, stop codon, 5' cap (remember: these exons may contain translated regions and untranslated regions, DON'T BE RUSH!).

So, in genomic DNA libraries we have everything, but just genes in the cDNA libraries, also we have coding and non-coding regions in the genomic DNA libraries, but we have just coding regions in terms of the cDNA libraries, due to that, if we have let's say a skin cell or a neuron or a muscle cell and we create genomic DNA libraries, then these libraries for sure are identical. On the other hand, if we rather create cDNA libraries, these libraries will be different, because genes that are expressed in skin cells are completely different from genes expressed in nerve cells.

The two figures below explain the story:





YEAST TWO-HYBRID SYSTEM

TAKING ADVANTAGE OF DOMAINS

In yeast, an upstream activating sequence (UAS), (a regulatory element), exists.

UAS is the binding site of a transcription factor called Gal4, so clearly, when Gal4 binds to UAS it induces expression of a certain gene (LacZ in the picture).

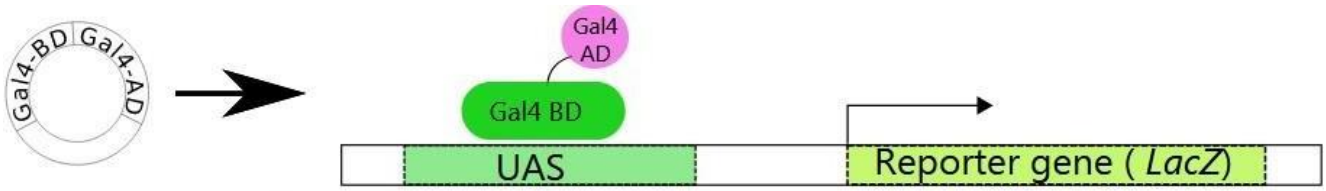
Using genetic engineering and recombinant DNA technology, we have special yeast cells that have a reporter gene (LacZ) which is under the control of UAS and Gal4.

Side note: we have already discussed lacZ gene back in the molecular biology course, but we discussed it as an example of translation regulation in bacteria, not in yeast.

UAS is controlled by a transcription factor (Gal4) that is made of two domains:

- **A DNA-binding domain (BD), which simply binds to the UAS.**
- **An activating domain (AD) that is responsible for the activation of transcription, by stimulating RNA polymerase to transcribe the gene.**

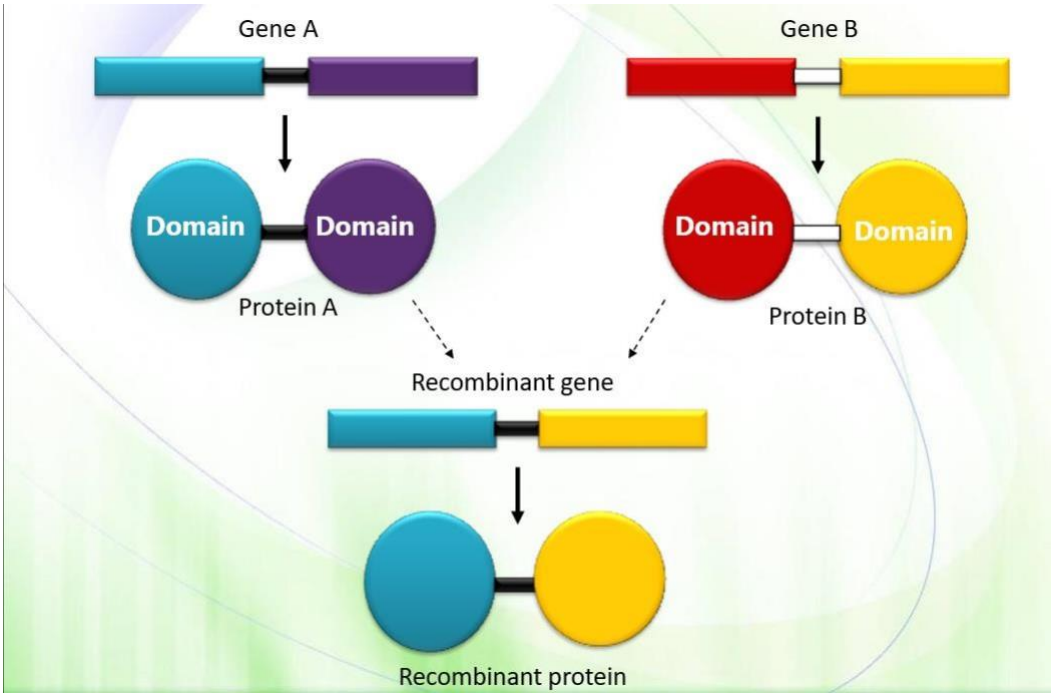
Both must be close to each other in order to transcribe a reporter gene such the LacZ gene.



A. Regular transcription of the reporter gene

https://www.youtube.com/watch?v=okxle_hTaZ0
https://www.youtube.com/watch?v=NxNfibcNk_Y

PRODUCTION OF A RECOMBINANT PROTEIN



As simple as it sounds, recombinant DNA technology is concerned with making recombinant proteins, and this happens by taking advantage of "domains" concept.

A reminder, what are domains?

A protein domain is a compact region (or part) of the protein's polypeptide chain that:

1- Has a defined three-dimensional structure,

2- IS self-stabilizing,

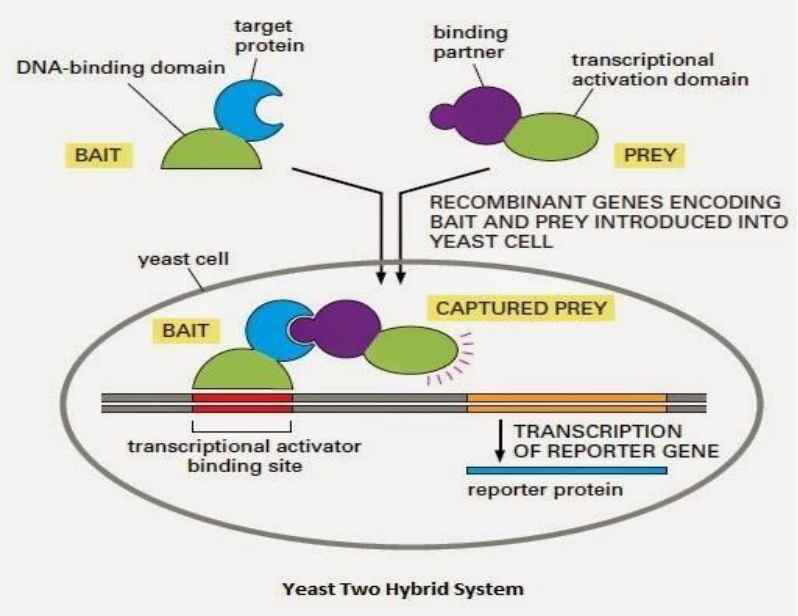
3- Folds independently from the rest.

4- Can be disconnected from the protein and, yet maintains its structure and function.

So taking a domain from a protein, fusing it with a domain from another protein (keeping in mind that each domain is still doing its function separately) is our secret magical spell to have a recombinant protein!

QUICK ILLUSTRATION

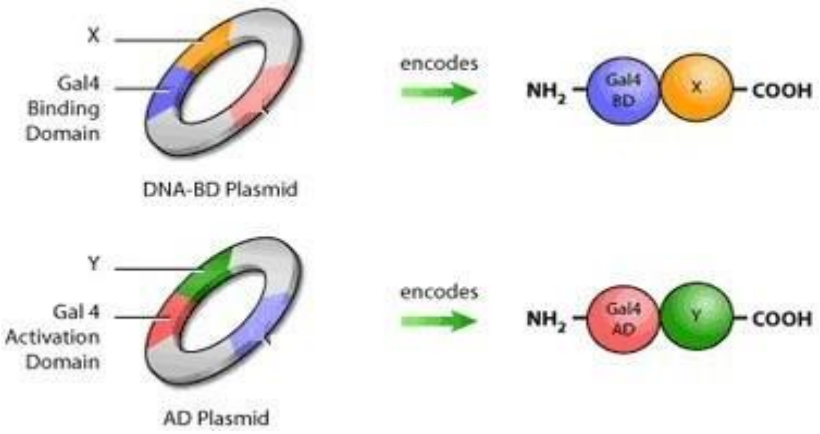
So, the purpose of the yeast two-hybrid system is to investigate if 2 proteins interact with each other, and it's done in the way elaborated in the figure here.



Simply, we take a recombinant protein which contains the DNA binding domain of Gal4 and another domain from another protein, and another recombinant protein that contains the activation domain of Gal4 and another domain from another protein, so the idea that if the other 2 domains (BLUE AND PURPLE IN THE PIC) bind to eachother the two domains of Gal4 will be close to each other, so transcription of reporter gene occurs.

CLONING OF HYBRID PROTEINS

- **In order to discover unknown proteins (Y's) that interact with a known protein (X), the X gene is cloned so it is produced recombined with the DB domain and the unknown Y gene (or genes) are <u>separately</u> cloned, so they are produced recombined with AD.**



- **Both recombinant plasmids are transferred into yeast cells so all of them express the known X gene-BD hybrid, but each one expresses a different unknown Y gene-AD hybrid.**

So again, we insert the known gene (x) with Gal4 DNA binding domain, and the unknown gene (y) with the activation domain of Gal4.

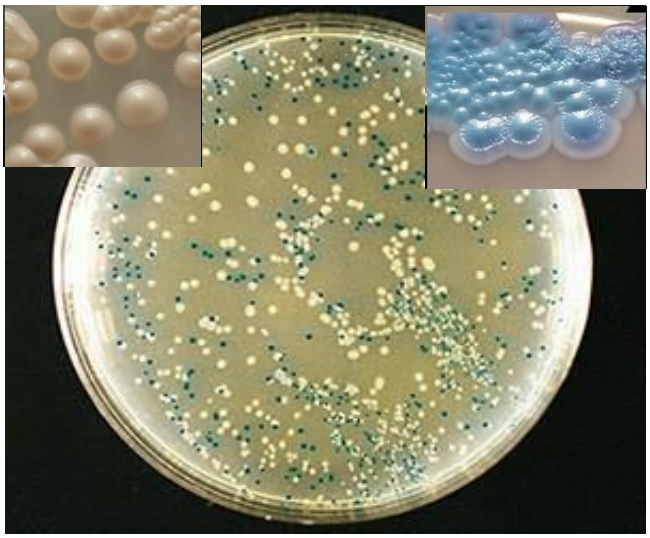
All yeast (**having 2 plasmids**) will produce gene (X) fused to binding domain of Gal4 but each one of them will produce a **different Y** unknown gene fused with the activation domain of Gal4.

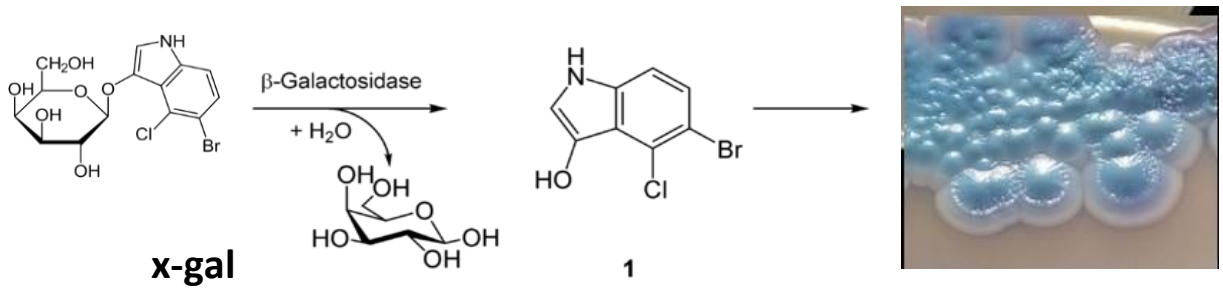
WHY IS THE LACZ GENE USED? WHAT IS X-GAL?

- **Yeast cells are grown in the presence of a lactose analog called X-gal, which generates a blue product when cleaved.**
- **When the LacZ gene is activated (after the binding of x and y genes), beta-galactosidase is produced, which cleaves X- gal generating blue colonies.**

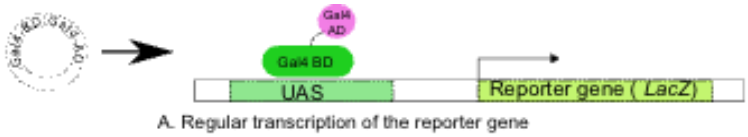
Side note: x-gal is so close to lactose in structure, but it produces a product that gives a blue color upon cleavage.

Side note: if x known gene and y unknown gene don't bind to each other, no blue color will appear, and the colonies will stay white.

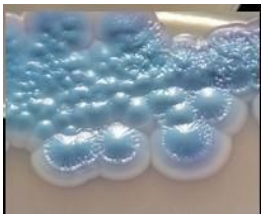


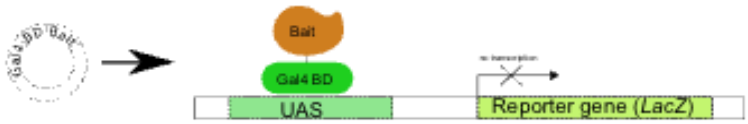


THE POSSIBILITEIS AND OUTCOMES

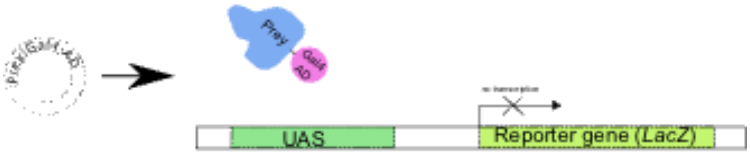


A. Regular transcription of the reporter gene

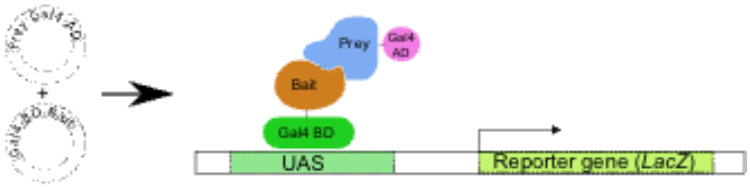




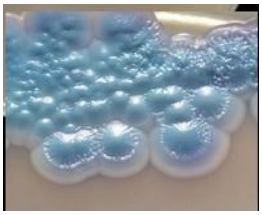
B. One fusion protein only (Gal4-BD + Bait) - no transcription



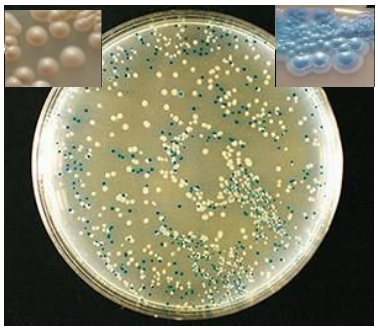
C. One fusion protein only (Gal4-AD + Prey) - no transcription



D. Two fusion proteins with interacting Bait and Prey



1- The normal Gal4 gene binding → transcription → blue color.
2- Only the DNA binding domain (fused with gene x) → no transcription → white color.
3- Only the activation domain (fused with gene y) → no transcription → white color.
4- Gene x and gene y are bound to each other → activation domain and binding domain of Gal4 are close to each other → transcription → blue color.



Blue yeast colonies are picked, and plasmids are isolated to identify the unknown genes (Y gene)/proteins that interact with the known gene (X gene)/protein (using sequencing, PCR, immunoblotting and many other techniques).

White colonies also produce x and y genes, but they are not bound to each other, blue colonies produce them bound.