بسم الله الرحمـٰن الرحيم

(وَفَوْقَ كُلِّ ذِي عِلْمٍ عَلِيمٌ)

Cytology & Molecular Biology | FINAL 19

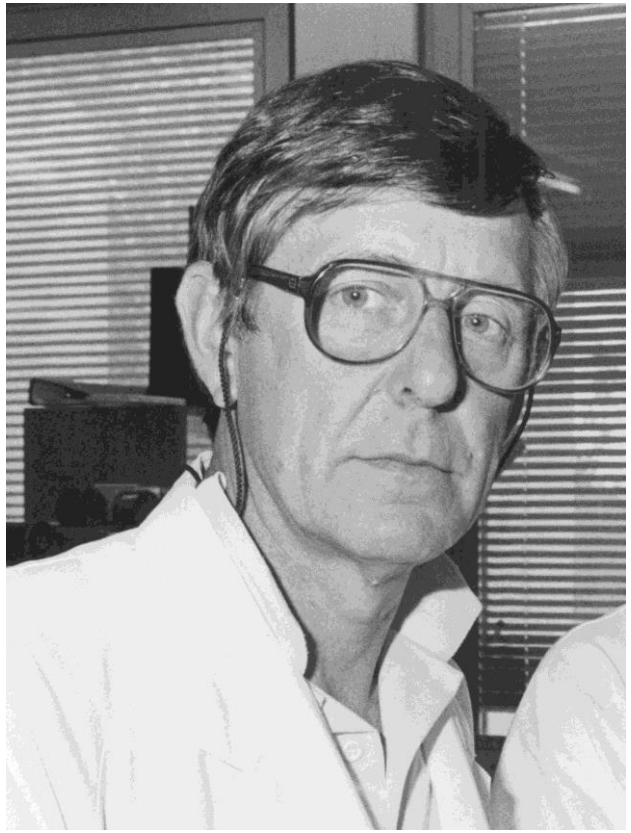
Recombinant Proteins

Written by : DST

Reviewed by : Reem Alfaqeh

Leen Alzoubi

Click on John Shine or Lynn Dalgarno to access last lecture's quiz :)



Expression of human proteins in bacteria

Brief introduction

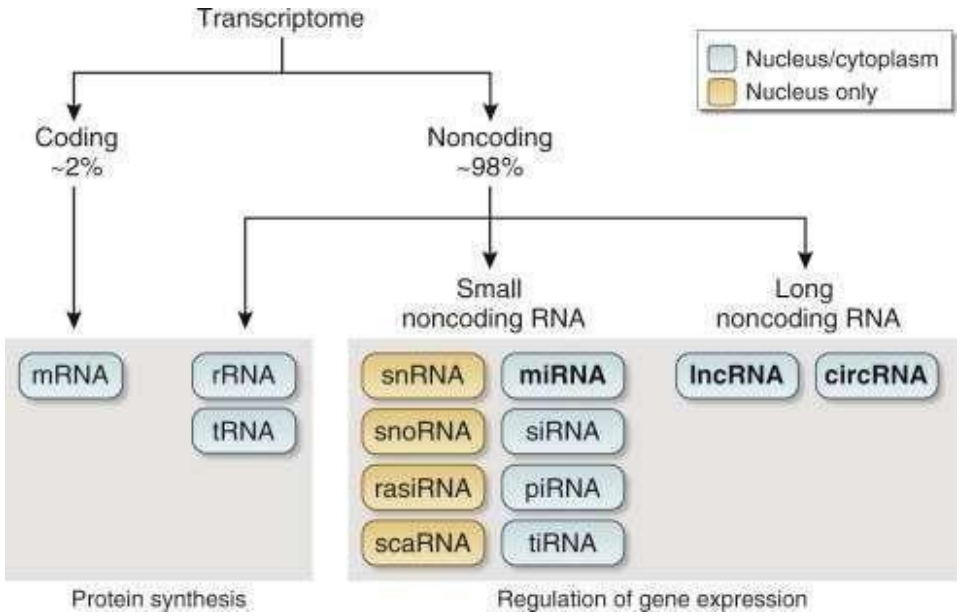
In previous lectures, we dove into cDNA libraries and cloning human DNA in bacteria by way of cloning vectors. In order to express a specific gene in bacteria, the gene's mRNA must be selected. However, there are two problems when it comes to using human mRNA for gene expression in bacteria, both of which will be discussed in the upcoming slides.

How do we select for human mRNA?

The power of reverse transcriptase (part 1)

There exists many human mRNA molecules inside the cell. About 50 years ago, only 3 RNA molecules were known to man; mRNA, tRNA, and rRNA. In present day, there are approximately 15-20 types of mRNA, including noncoding RNA, both long and short, and miRNA. The abundancy of RNA molecules makes it tricky to select for mRNA molecules specifically. How can we address this problem? By observing the characteristics unique to mRNA. For example, mRNA is the only RNA molecule that has a poly-A-tail. Once this identification has been made, reverse transcriptase catalyzes the conversion of mRNA cDNA to be used for gene expression in bacteria.

The "many types of RNA" challenge

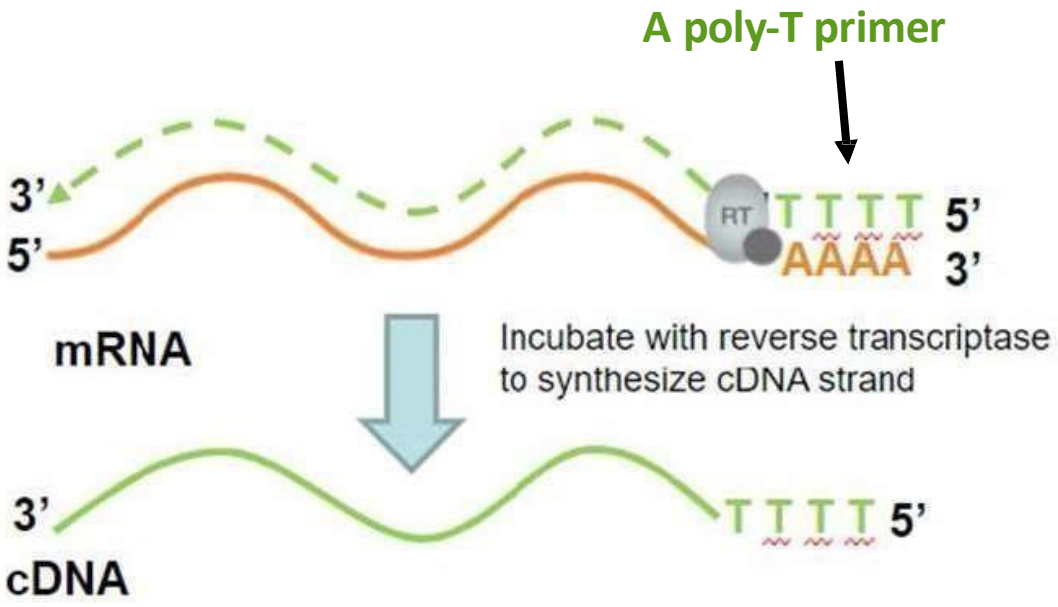


How do we select for human mRNA?

The power of reverse transcriptase (part 1)

Reverse transcriptase, like DNA polymerase, requires a primer-- poly-T primer. Once the poly-T primer is added, reverse transcriptase synthesizes the first strand of DNA. This is the perfect solution for the issue scientists faced in the previous slide. Since the poly-T primer is specific to poly-A-tails, it can be used to pick out mRNA molecules from the other RNA molecules.

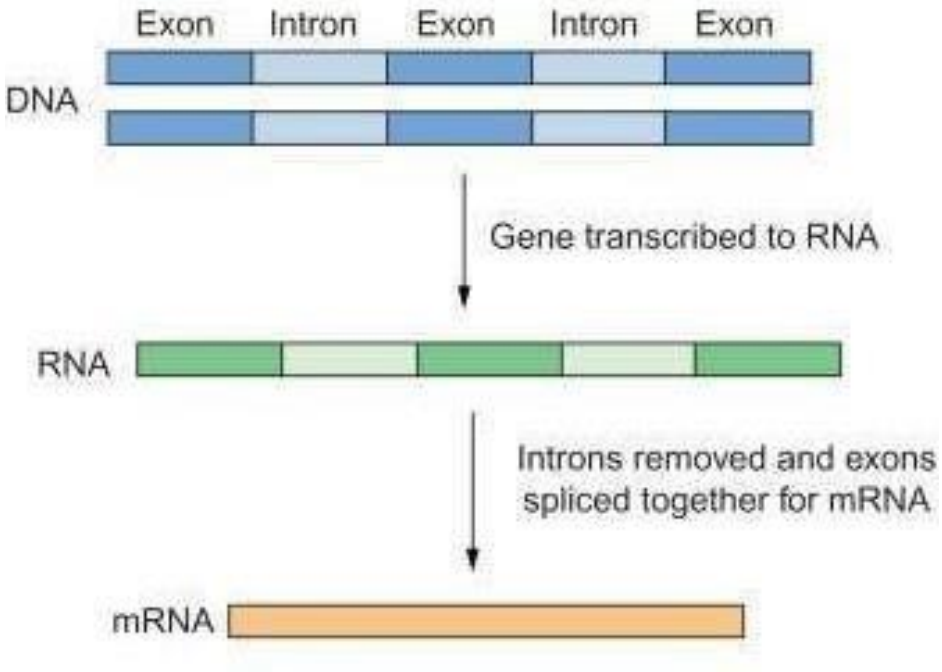
The "poly-T primer" solution



How do we deselect introns?

The power of reverse transcriptase (part (2

Go for mature mRNA
Synthesizing a gene from genomic DNA poses a specific problem that is the transcription of introns. This is why the primary transcript is not used to make copies of a gene. Instead, the mature RNA transcript (mRNA), containing only exons connected to each other, is used.

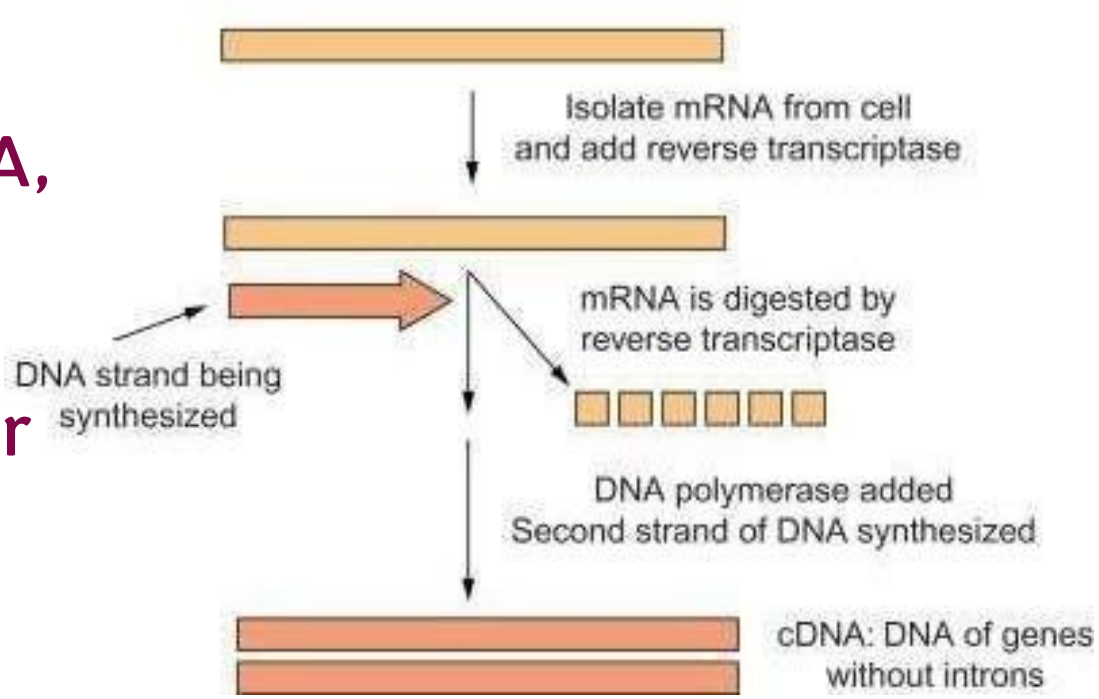


How do we deselect introns?

The power of reverse transcriptase (part 2)

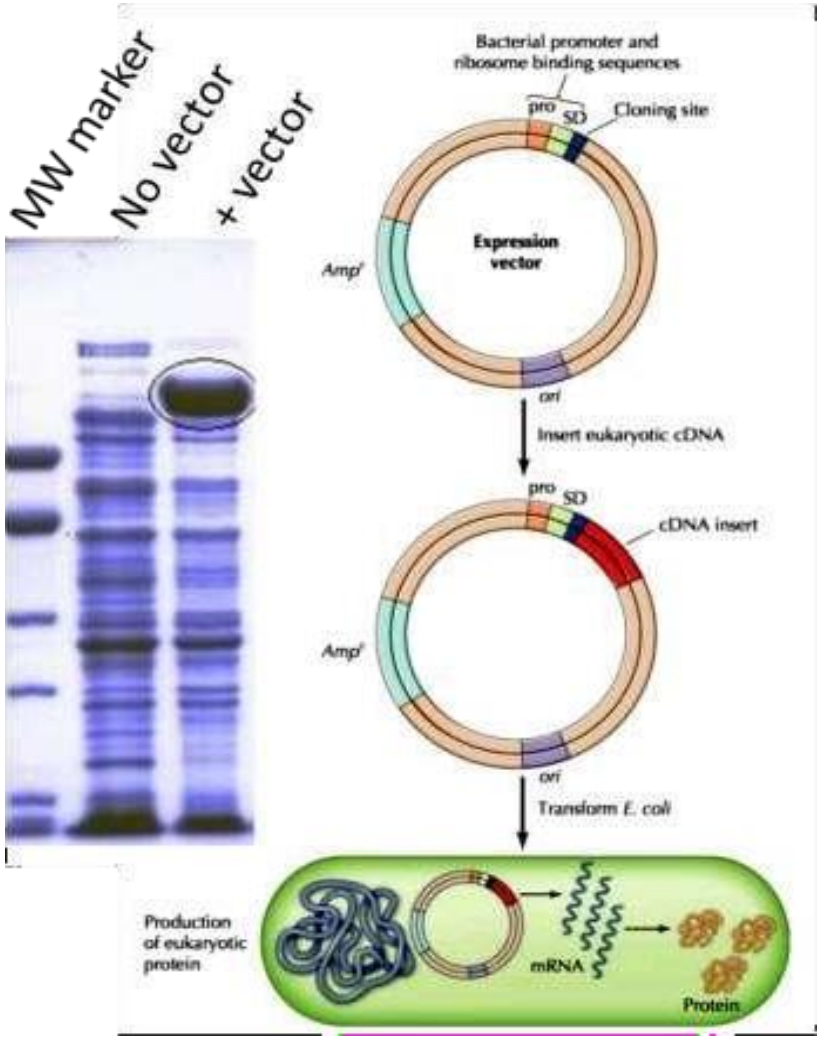
Reverse transcriptase is used to synthesize complementary dsDNA from mRNA. This complementary DNA, without introns or a poly-A-tail, is cloned into a plasmid. Remember, bacteria do not have introns, exons, or mRNA molecules with poly-A-tails.

The "reverse" solution



Expression vectors

- Expression vectors contain additional sequences:
  - Promoter sequences upstream of gene to be inserted,
  - Ribosomal binding sequences (Shine-Dalgarno [SD] sequences),
  - A transcription termination sequence.
- The protein is expressed and purified.
- Examples: insulin, growth hormone, plasminogen activator, erythropoietin



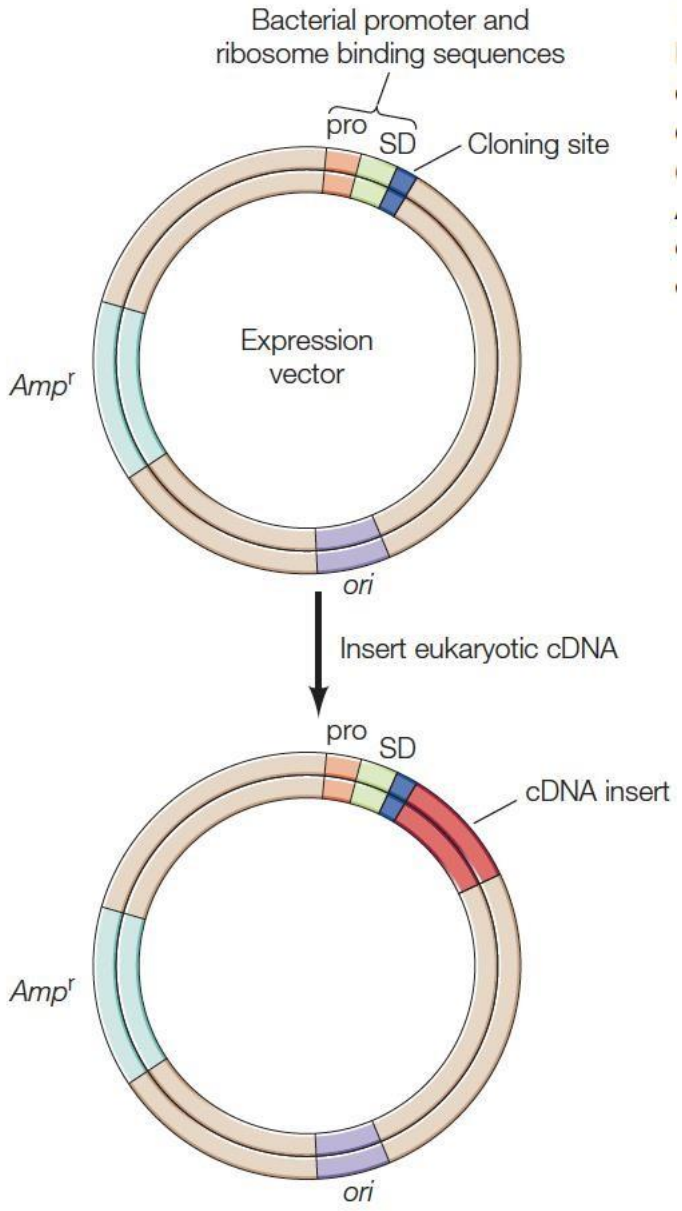
Expression vectors

Recall DNA cloning. The vectors used are called cloning vectors. The purpose of these cloning vectors is to amplify DNA in bacterial cells, regardless of what the DNA could be (introns, exons, promoters, enhancers, etc.). When discussing expression vectors, it means the expression of human proteins, starting with mRNA. mRNA is then converted to cDNA, and then cDNA is cloned into an expression vector. The expression vector shares the same properties of cloning vectors; origin of replication and an antibiotic resistance gene. However, expression vectors have three unique properties, which are:

Promoter sequence: it does not have to be the same as that of the original gene. For example, if the insulin gene is inserted into the plasmid, bacterial promoters can be used in place of the insulin gene promoter. The cloning site, which is where the gene is attached, is downstream of the promoter.

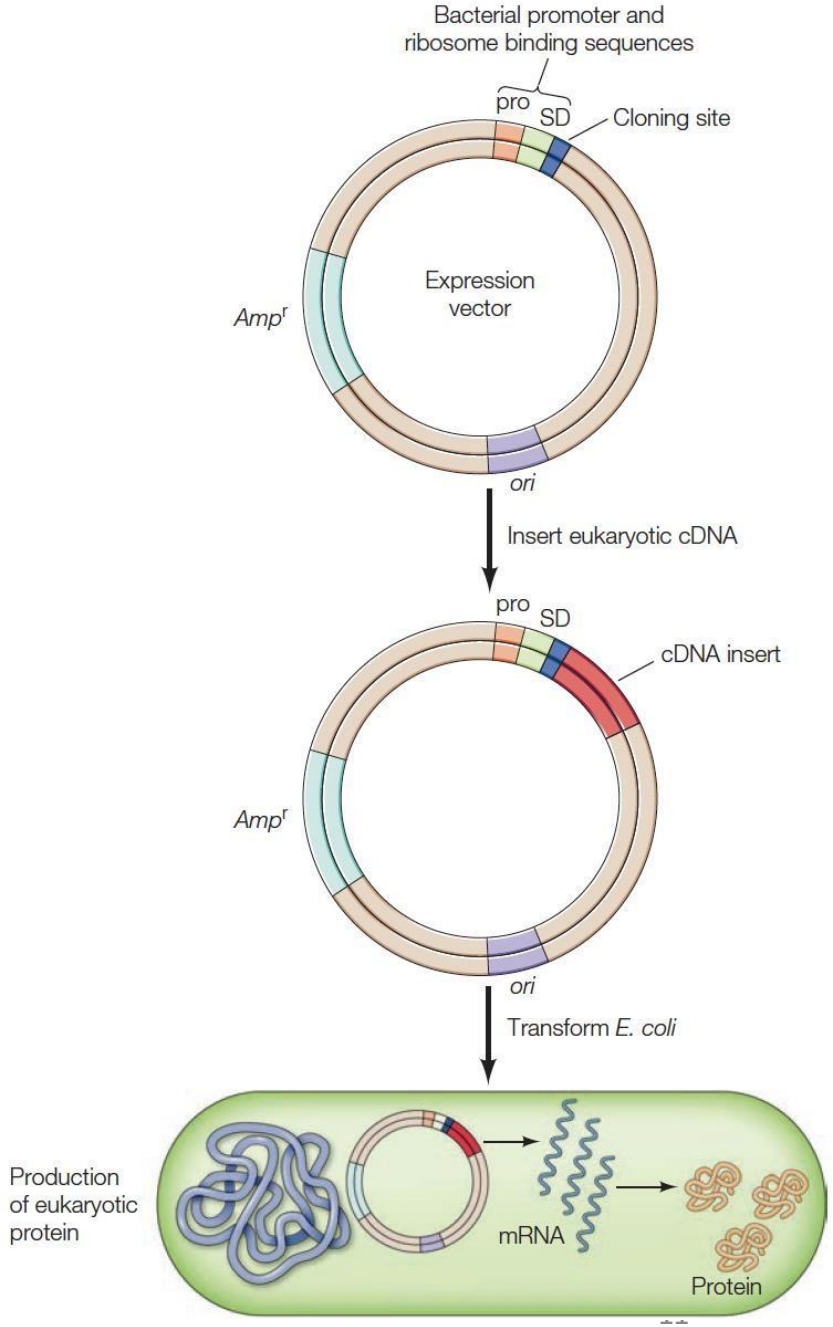
Ribosomal binding sequences: once the gene is expressed, the Shine-Dalgarno sequence is needed for the ribosome to bind to.

Transcription termination sequence: not a stop codon!



Expression vectors

To put it all together: the plasmid, containing the cDNA, and all the previously mentioned qualities, is inserted into a bacterium, which recognizes the promoter region. From there, it initiates transcription and translation of the targeted protein. The synthesized protein is then folded and assembled into its functional form. This can be applied to insulin, growth hormones, plasminogen activators, and erythropoietin. Through this process, bacterial cells have been essentially turned into protein factories. These proteins are of clinical importance as they can be used to treat patients.



Expression vectors

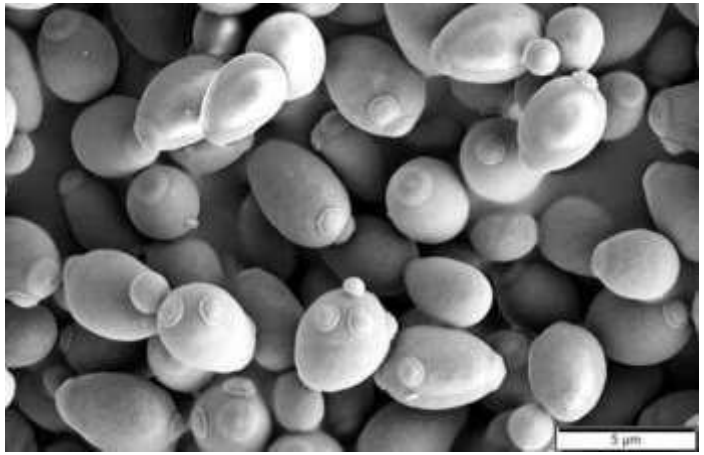
The protein can be extracted from the bacterial cell and put through SDS-PAGE to be filtered through size. Look at the following picture. The “No vector” column represents all bacterial proteins, taken from bacteria that do not contain the plasmid. Comparatively, the ”+ vector” column expresses proteins from a bacterial cell that does contain the vector. The thick band observed represents the human protein, and the thickness of the band signifies the high expression of that protein.

Challenges of protein expression in bacteria

- No internal disulfide bonds
- No post-translational modification (example: glycosylation)
- Protein misfolding — Especially large proteins. Much less for smaller proteins.
- Protein degradation — Applicable to **human** proteins, due to the bacteria not recognizing it.

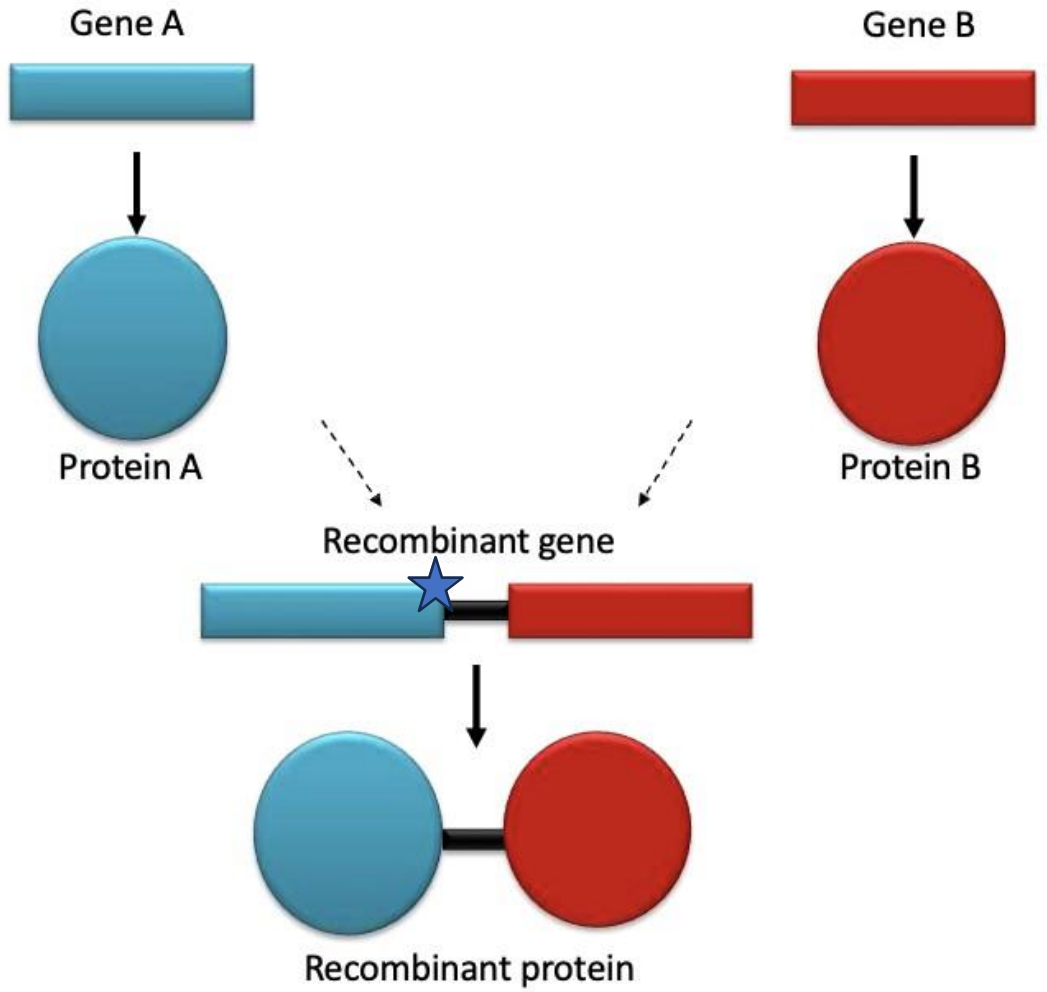
- Solution: use a eukaryotic system such as yeast

Yeast are unicellular, rapidly-growing, eukaryotic organisms. Due to them being eukaryotic, they carry out the same functions that humans do. Plasmids can be inserted in yeast; bacterial cells can fooled, in a sense, because they cannot really differentiate between plasmid and genomic DNA.



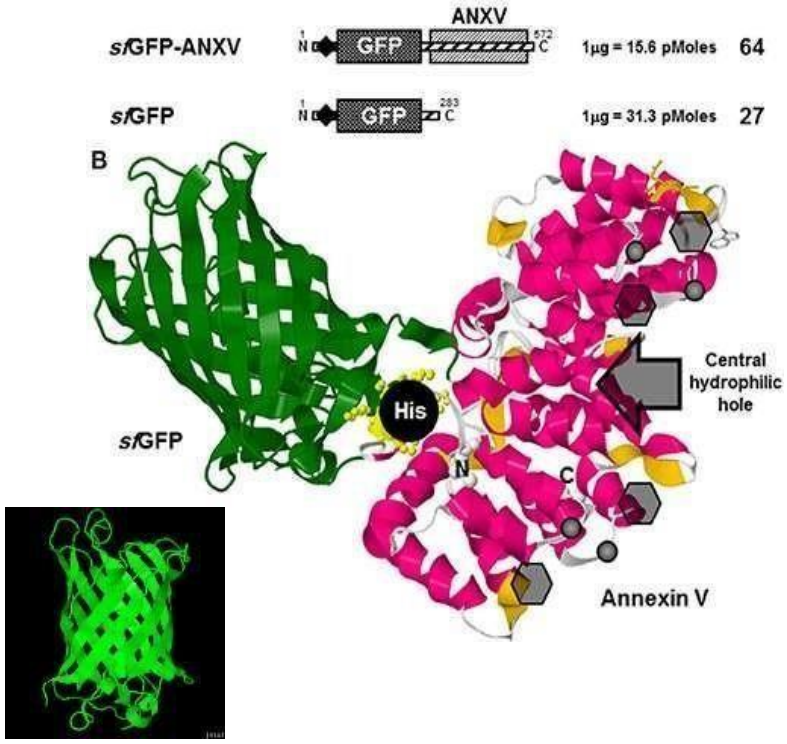
Production of a recombinant protein

A recombinant protein is a protein that contains two different proteins. Separately, gene A and gene B produce proteins A and B, respectively. However, when these genes are fused together, they from a recombinant gene. The stop codon is deleted from gene A, so that when a protein is eventually translated, translation doesn't stop at the end of the gene (★), but continues until the stop codon of gene B is reached. The two synthesized proteins are joined by a linker that does not affect protein folding. The blue and red proteins fold independently of each other.

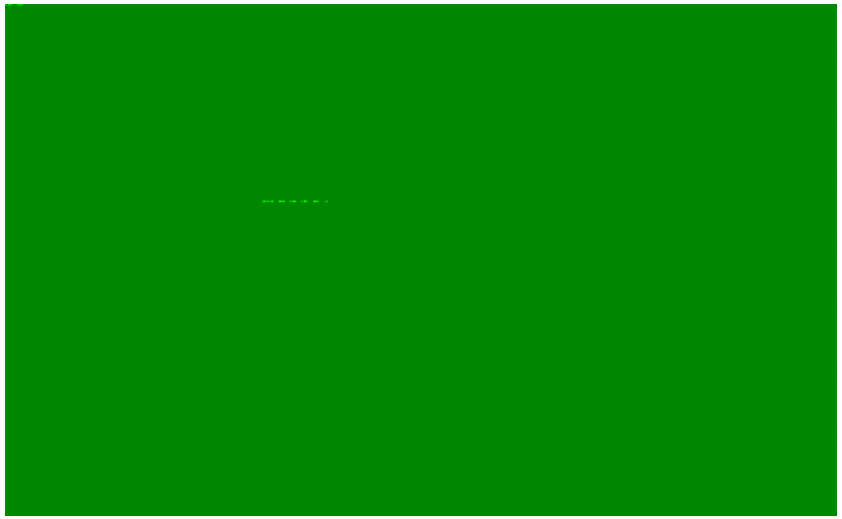


GFP-tagged proteins

- Green Fluorescent Protein (GFP) allows for protein detection rather than for purification purposes.

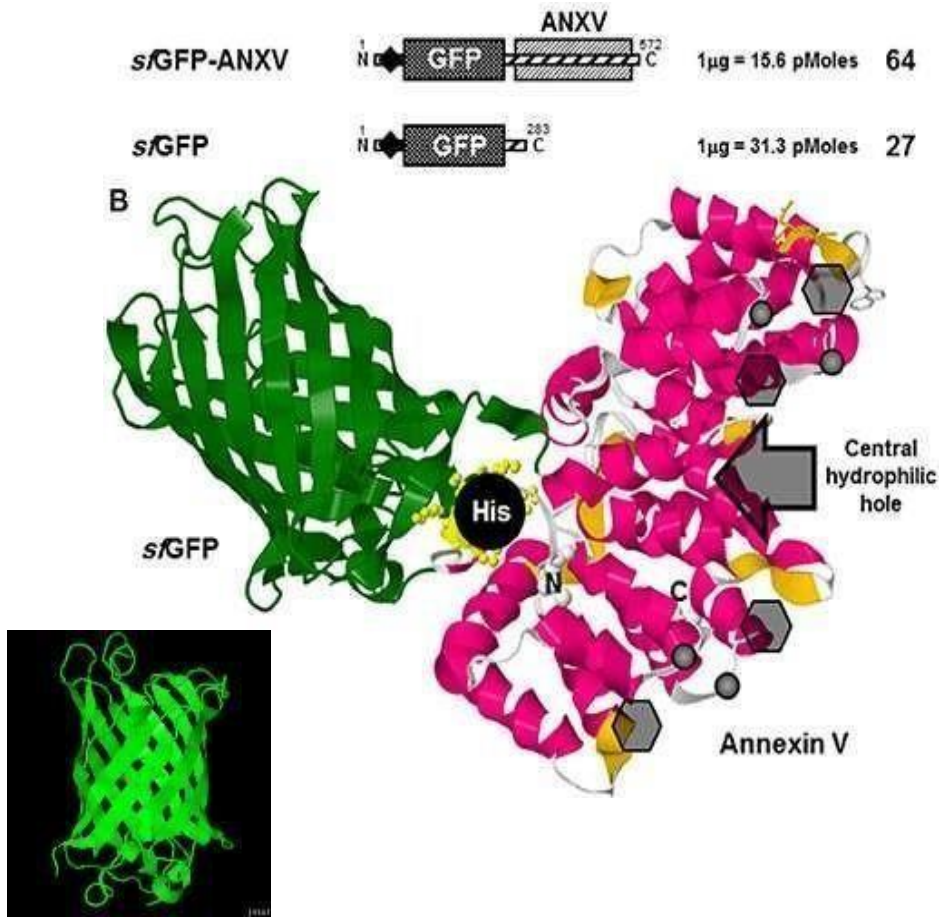


[You can view this animation in the recorded lecture (19:46-20:11)]



GFP-Tagged Proteins

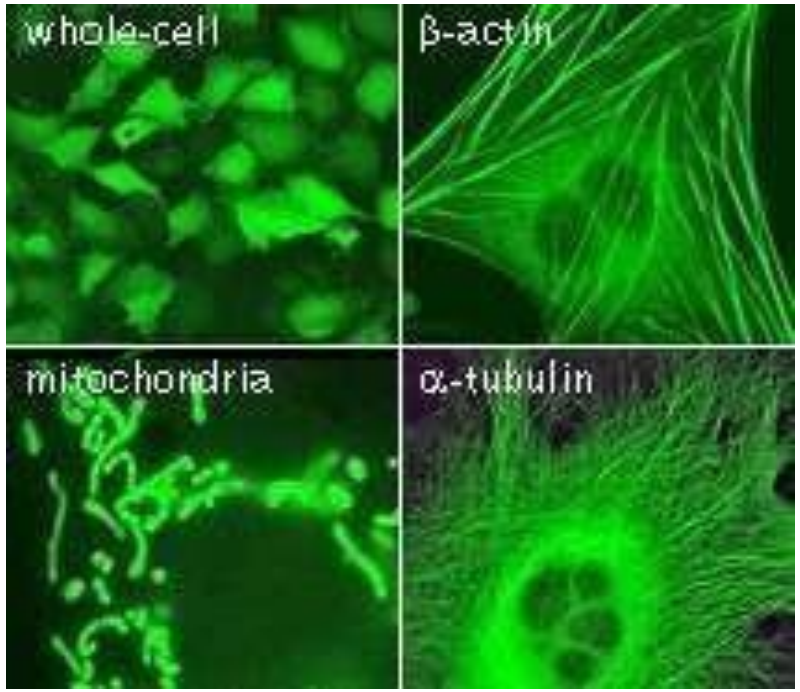
One example of a recombinant protein is the green fluorescent protein (GFP). This protein is isolated from jellyfish and can produce fluorescence. This gene can be cloned upstream or downstream of another gene, the protein of interest, by its insertion in an expression vector, then transcription & translation. The synthesized protein is large; it contains both the green fluorescent protein and the protein of interest. The protein of interest fluoresces because the green protein fluoresces as well. Both proteins fold independently of each other. This goes to show the advantages of domains. Due to the fluorescence of the protein, its movement can be tracked in the live cell.

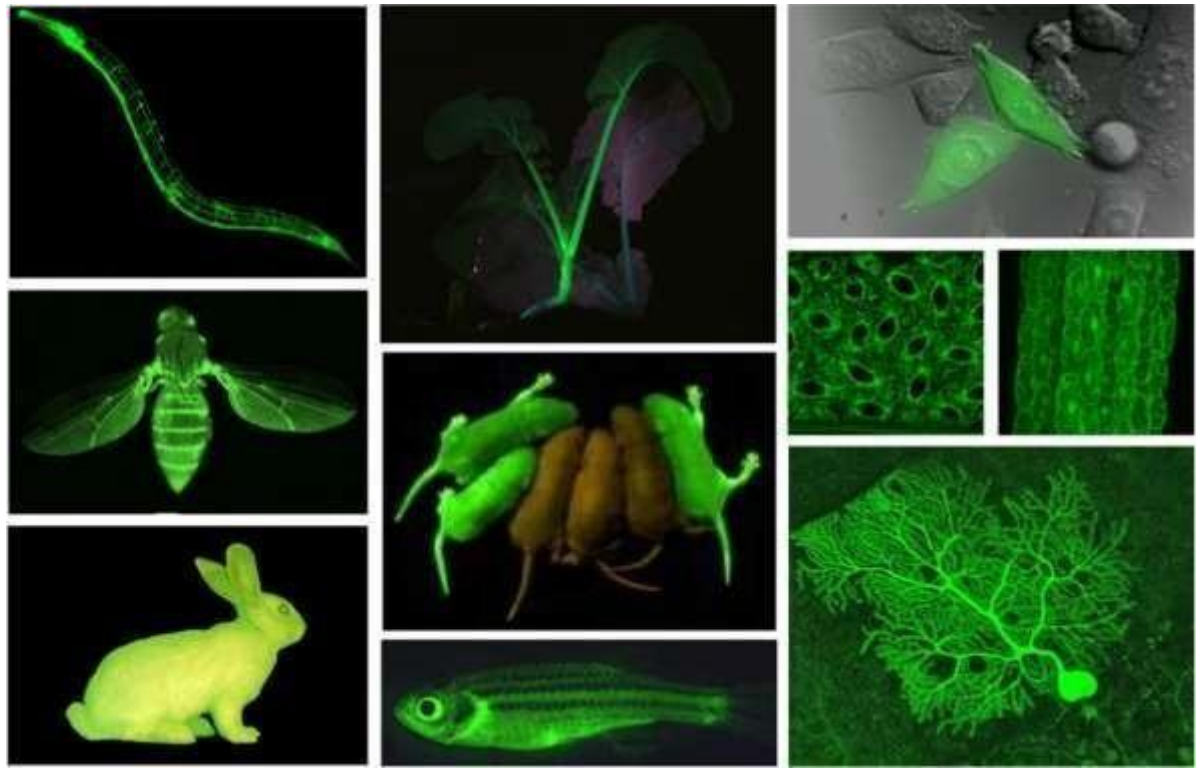


A world of possibilities

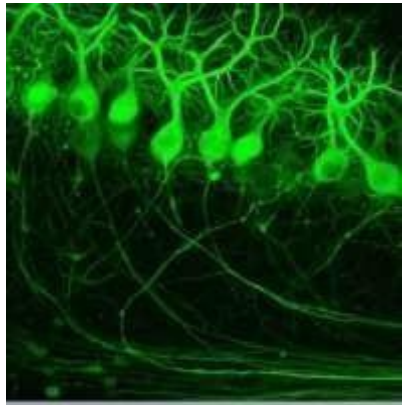
These are some examples of proteins that have been
tagged or have another protein (GFP) attached to them.

Fluorescence helps us understand morphology and organization of tissues and organs.





Nerve cell fluorescence helps us comprehend neural connections

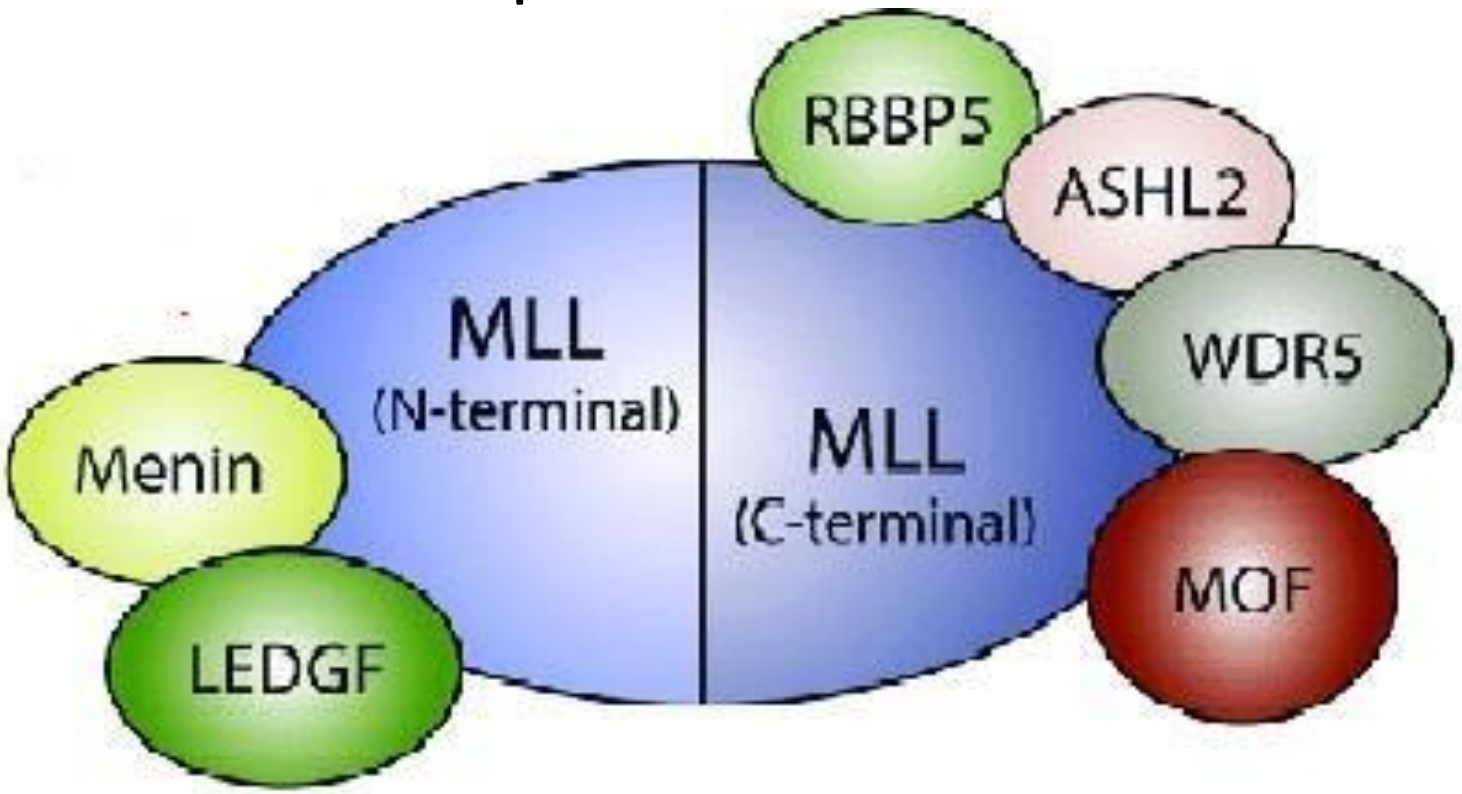


Protein-protein interaction

Co-immunoprecipitation

How can we understand protein-protein interaction? now we will introduce two techniques one of them is known as co-immune precipitation, precipitation means: it settles down, immune: once you hear the word immune you should think of antibodies, co : together
So, I am precipitating a protein with another protein using antibodies

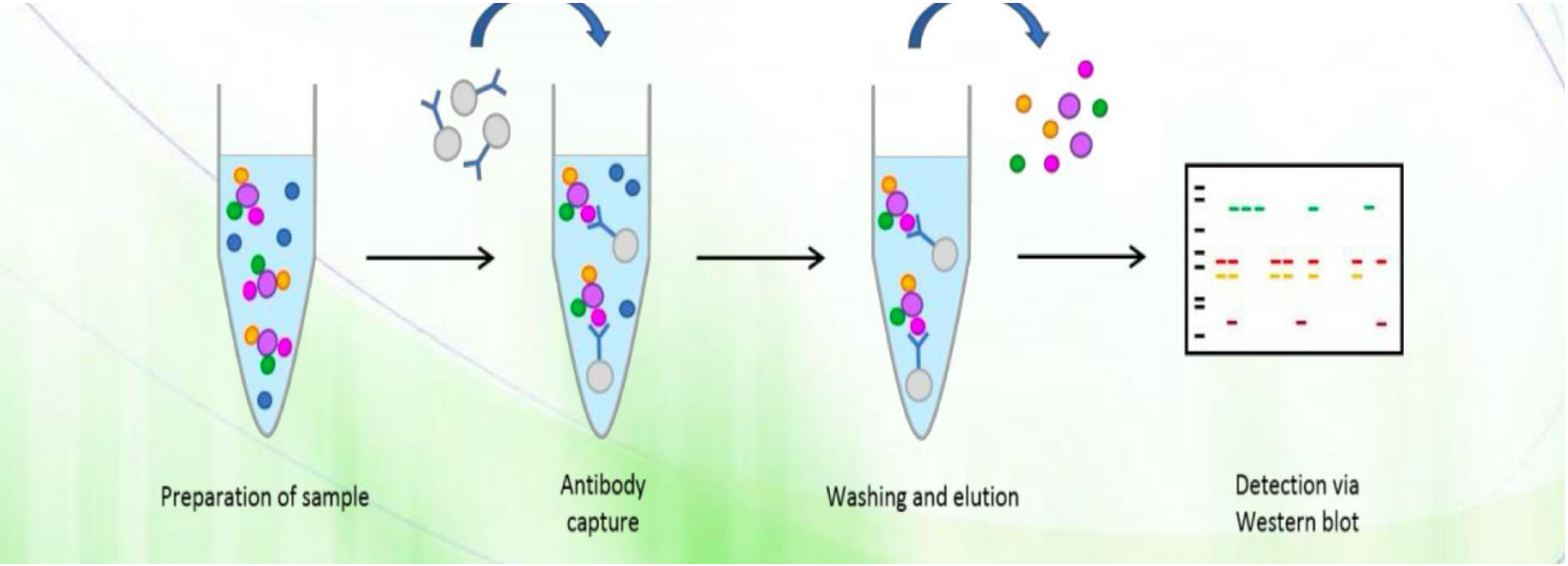
Proteins form complexes

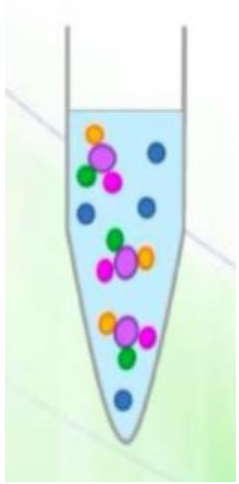


Remember that we can have a large protein let's say this protein MLL has two domains : an N -terminal domain And c-terminal domain and the N -terminal domain can interact with these different proteins, and these proteins can also interact with each other. Now the c-terminal domain can interact with four proteins some of them can interact with each other and others can't like (MoF) protein can't interact directly with the protein(RBBP 5) . How can I identify protein- protein interaction? with the immunoprecipation that's what we do

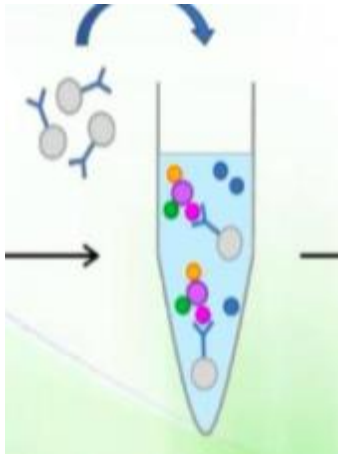
(Co)-Immunoprecipitation

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

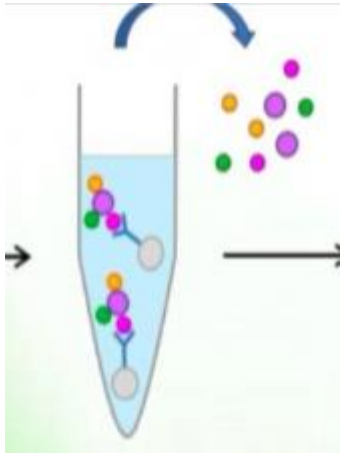




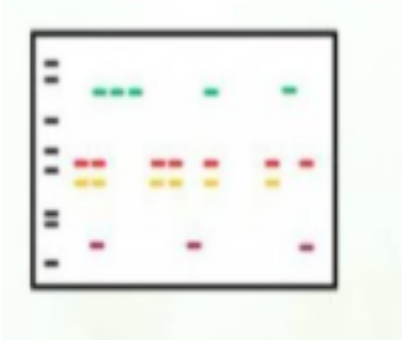
1.we have these proteins complexes . A protein complex is multiple proteins interacting with each other



2. we add an antibody which is specific to one single protein. This antibody then targets the protein of interest (the pink protein) which Interacts with other proteins, so what Will happen is that the antibody is quite heavy and will precipitate along with the protein complexes



3. We remove everything else that doesn't interact with the antibody and then we have these protein complexes that interact with the pink proteins



4. We can take these complexes and study them using for example SDS page (gel electrophoresis) or using western blot (immune blots) in which we separate proteins based on size , then transfer them to a membrane and we add a primary antibody and a secondary antibody to detect the presence of proteins, and we can also determine their sizes

Protein-protein interaction

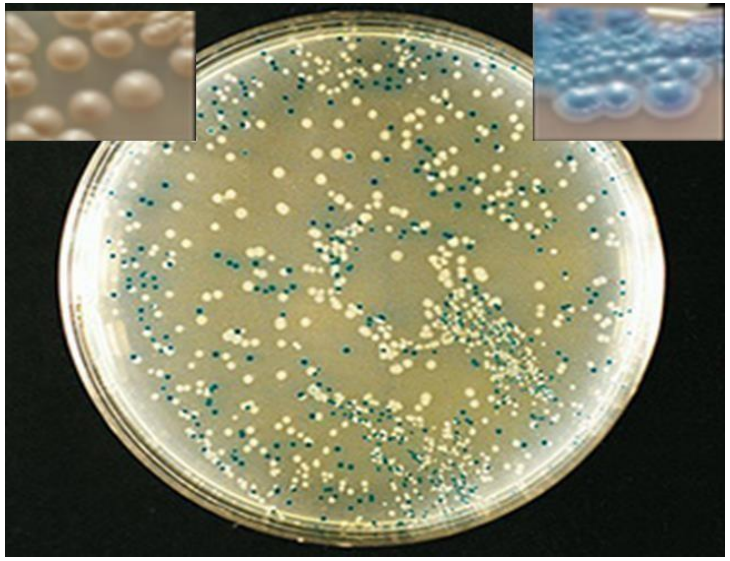
Yeast two-hybrid system

starting from a cDNA library

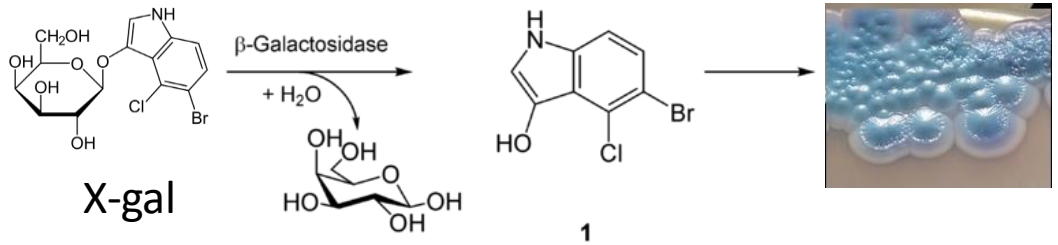
Let's go into something a bit more sophisticated (yeast two-hybrid system) it's a : genetic system where we express hybrids of two proteins or two genes .

Why is the LacZ gene used? What is X-gal?

- To test if a protein interacts with another protein, a genetic system is used on yeast cells.
- The yeast cells are allowed to express the bacterial β-galactosidase , whose gene is under control of the gal4 transcription factor.
- The gal4 protein has two domains, a DNA-binding (DB) domain and an activation domain (AD).
- Yeast cells are grown in the presence of a lactose analog called X-gal, which generates a blue product when cleaved.
- When the β-galactosidase gene is activated, beta-galactosidase is produced, which cleaves X-gal generating blue colonies.



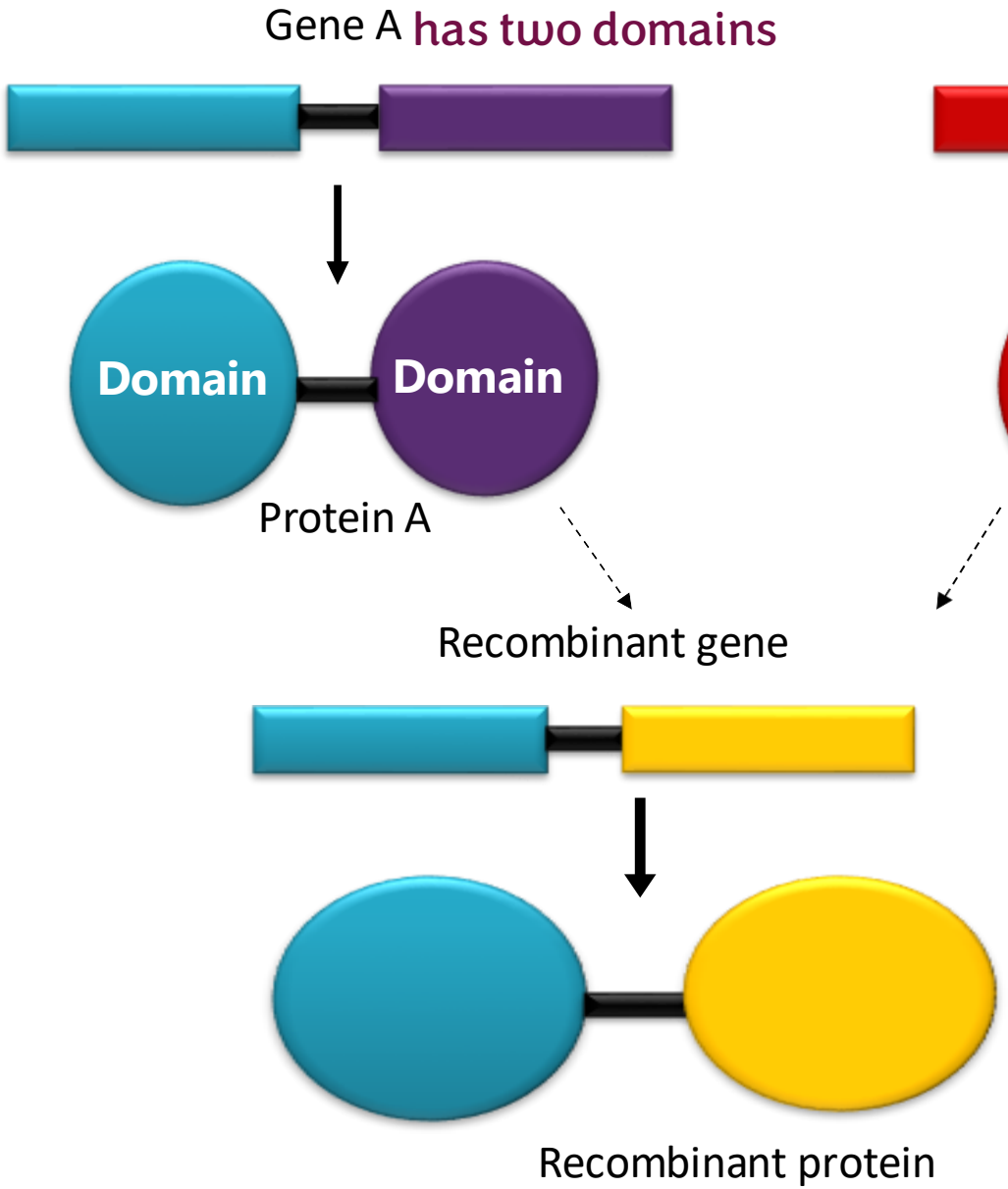
A. Regular transcription of the reporter gene



- There is a genetic system that consists of a regulatory sequence called UAS (this like a promoter, or a promoter proximal element). It's the binding site of the transcription factor (Gal4). Gal4 is a protein that has two domains: a DNA binding domain and an activation domain which can activate the expression of a gene except the gene that we use is a reporter gene just like Lucifer's (it's a gene that gives us a piece of information, it reports something). The gene that we use for the yeast two hybrid is the lac Z Gene . It's the beta galactosidase gene from bacteria it's the gene that produces galactosidase which is an enzyme that cleaves lactose . So, in this genetic system we use a regulatory element from yeast and a bacterial reporter gene which translates into Beta galactosidase that cleaves lactose . If lac z is expressed (if beta galactosidase protein is produced) then it can cleave lactose except that we use a substrate that is similar to lactose, and it's called (x-gal). Beta galactosidase cleaves x gal producing a molecule which gives a blue color, so the yeast cell become bluish .

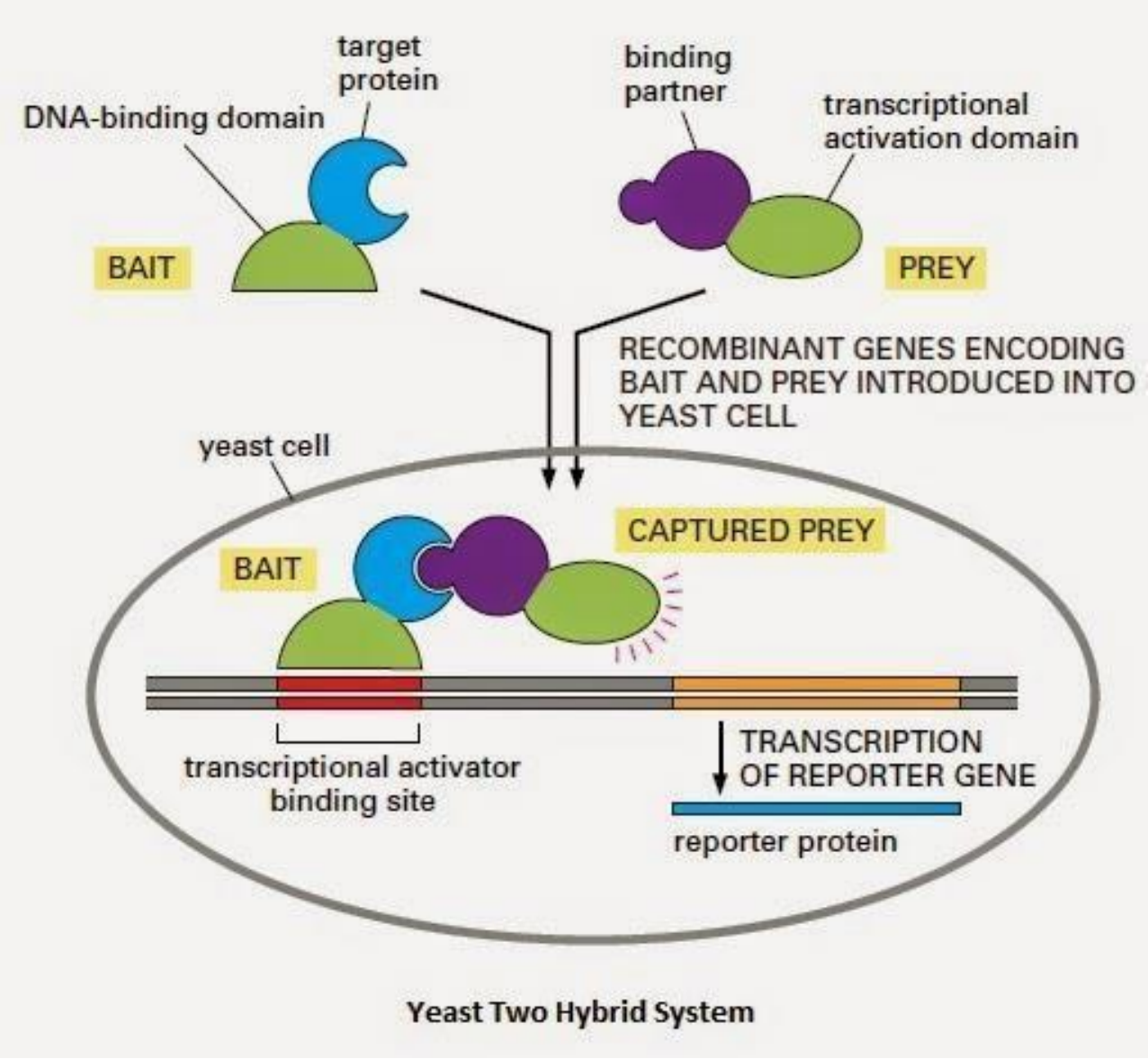
If we grow yeast on a plate each cell can form a colony .in previous figure we have multiple colonies some of them are white , others are blue .the bluish ones are the yeast cells that express the reporter gene ,the white ones originate from yeast cells that don't express beta galactosidase

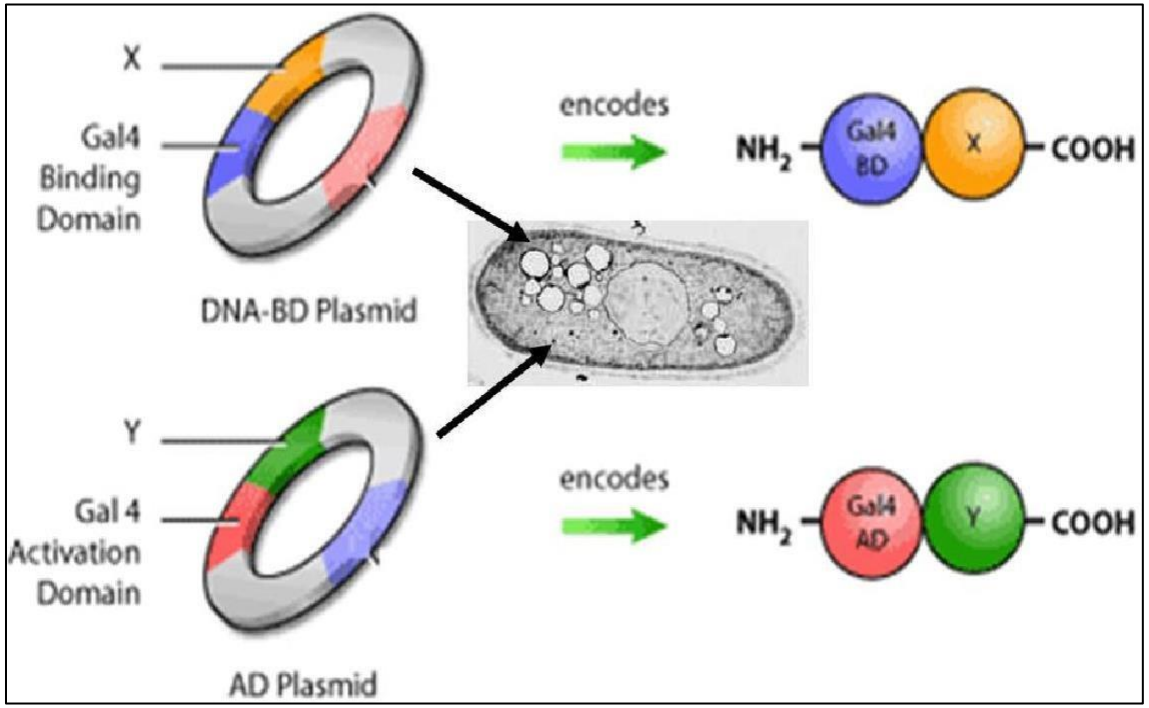
Production of a recombinant protein



The blue region from gene A and the yellow region from gene B, we can put them together to produce a recombinant protein that has two different domains from two different Proteins. The purpose of the yeast hybrid system to test if protein A can interact with protein B. How ? Using the genetic system that we just discussed.

Quick illustration

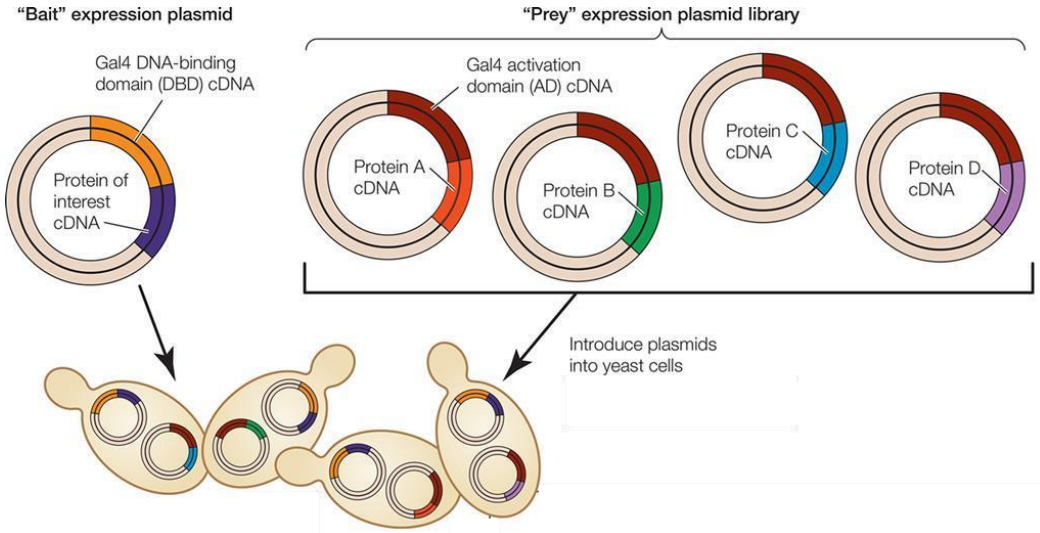


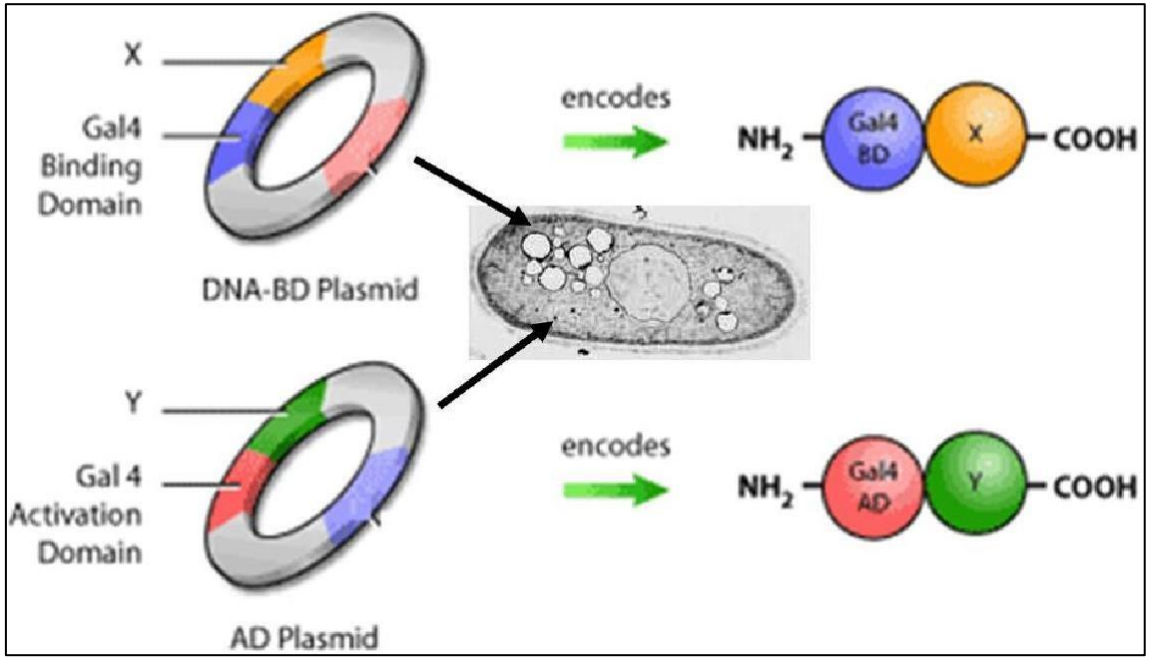


We want to test if protein x interacts with protein y . We are going to use the same genetic system that we just talked about (the regulatory element UAS, gal 4 and the lac z gene that produces beta galactosidase). We start by creating two plasmids (expression plasmids). We are going to put on one of them the gene for protein x and it would be produced as a recombinant protein with the gal 4 DNA binding domain. So, when protein X is translated, it's produced with gal 4 DNA binding domain. Plasmid two contains the gene for protein Y and it's produced as recombinant protein along with Gal 4 activation domain. The idea is that if protein X interacts with protein Y, then these two would be close to each other bringing the DNA binding domain and the activation domain to a very close proximity to each other and we will have transcription of the lac z gene. If lac z is expressed, we will have beta galactosidase which will cleave x gal , if x gal Is cleaved, it will generate a bluish product thus the yeast cell and colony become bluish . Blue indicates that protein X interacts with protein Y. If the reaction doesn't happen then the DNA binding domain would not be close to the activation domain and as a result there will be no transcription of the lac z gene , no production of beta Galactosidase, no cleavage of x- gal , no blue product and the colony would remain whitish.

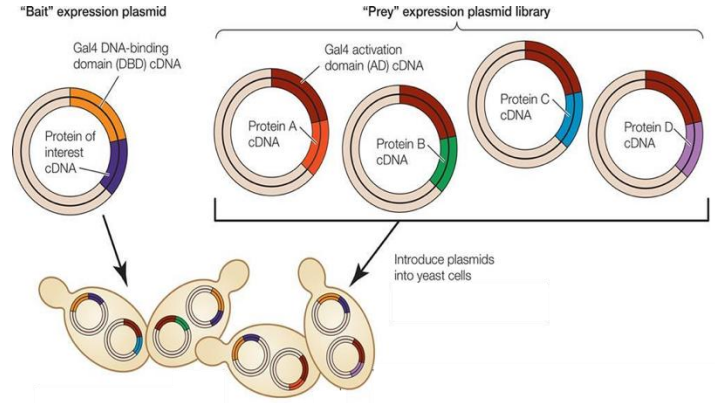
Cloning of hybrid proteins

- 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 DNA binding (DB) domain and the unknown Y gene (or genes) are separately cloned so that they are produced recombined with the activation domain (AD).
- Both recombinant plasmids are transferred into yeast cells so <u>all</u> of them express the known X gene-BD hybrid, but <u>each one</u> expresses a different unknown Y gene-AD hybrid.



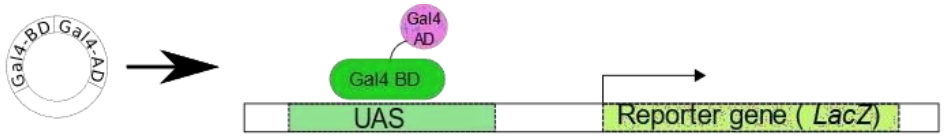


- Let's make things a bit more complicated, we want to identify all the proteins that can interact with protein X. So, we create a cDNA library. It's a collection of plasmids and each plasmid contains a cDNA representing a specific gene. Each one of these proteins is recombinant with the activation domain of gal 4, while protein X is recombinant with the binding domain of gal 4. we take the plasmid that contains the genes for protein X and the binding domain of gal 4 and we introduce it into all the yeast cells and then we insert the other plasmids (containing the genes of the other proteins with the activation domain gene) into the yeast cells. So, each yeast cell would take up one of these plasmids. So, each yeast cell would have two plasmids the plasmid that produces recombinant protein X and one plasmid that would produce either recombinant protein A or B or C or D . We grow them on a plate if protein X interacts with protein A, then this yeast cell produce beta galactosidase (which can cleave x gal) into a bluish product and the colony would look blue, if it doesn't interact with protein B for example the binding and activation domains of gal 4 will be far away from each other (no cleavage of X gal and the colony will look white). After that we take the blue colonies and isolate the plasmids , and we see the identity of the gene that is contained within these plasmids, we can identify what the gene is, and conclude that this gene produces a protein that interacts with protein X.

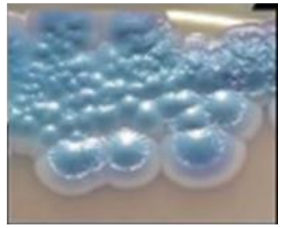


The possibilities and outcomes

A. If we have a plasmid that produces the full gal4 Protein (the binding domain (the green one) and the activation domain (the pink one)), then lac z will be expressed , beta galactosidase would be produced and if we add x gal to this yeast cell, we will have a blue colony

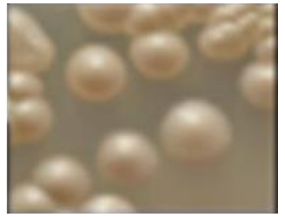


A. Regular transcription of the reporter gene

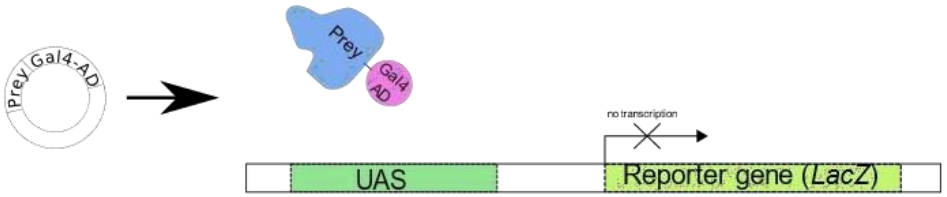


B. If we have plasmid that produces DNA Binding domain only there will be no transcription of lac z gene

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

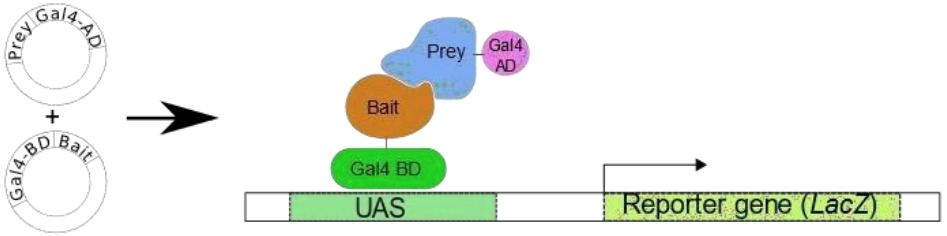


C. If we have plasmid that produces the activation domain only, we will have no production of beta galactosidase

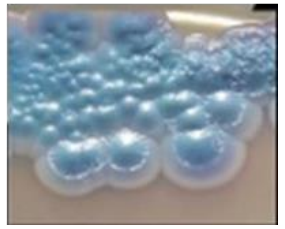


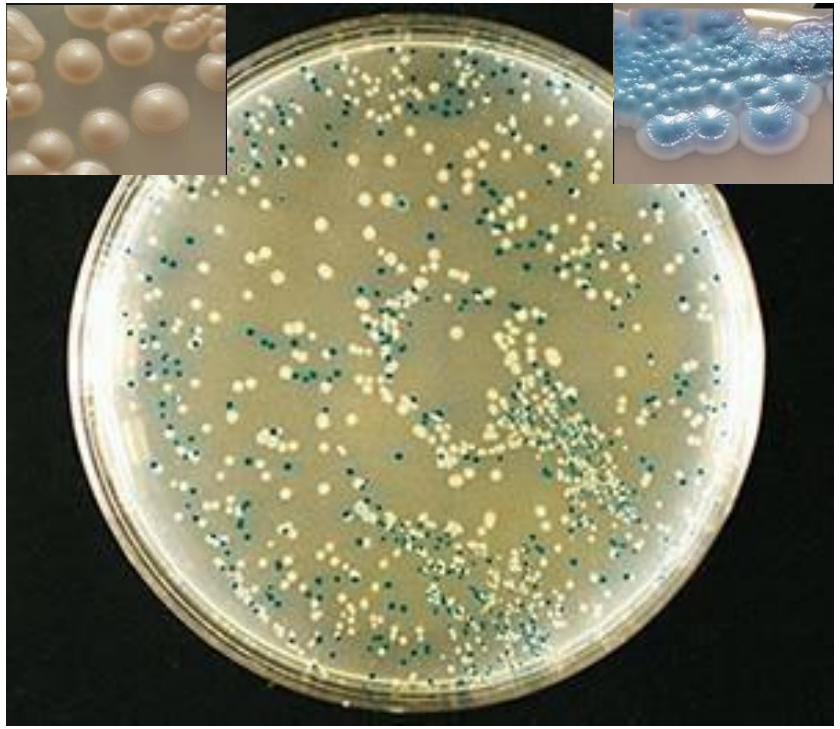
C. One fusion protein only (Gal4-AD + Prey) - no transcription

D. If we produce protein X, protein y and these two proteins can interact with each other with the DNA binding domain being part of protein X and the activation Domain being part of protein Y then these two would be close to each other, we will have expression of lac z , Production of galactosidase, conversion of x gal into a blue product and the colony would be blue. We take this Colony ,isolate the plasmid to identify the gene that can interact with protein X by sequencing or immune blots.



D. Two fusion proteins with interacting Bait and Prey



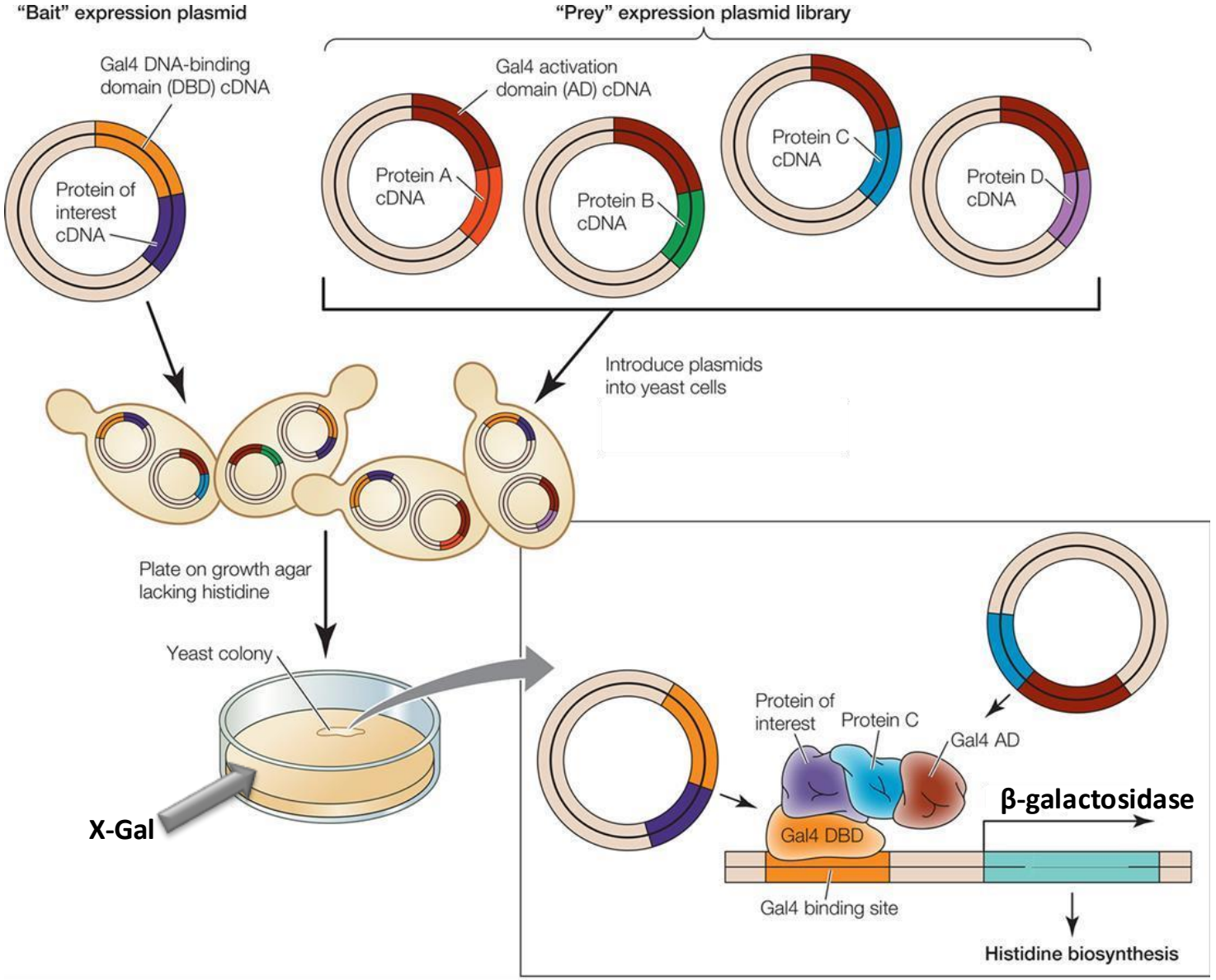


- Blue yeast colonies are picked, and plasmids are isolated to identify the unknown genes/proteins that interact with the known gene/protein.

The procedure

The doctor said to read this slide alone

- A cDNA encoding a protein-of-interest is cloned into an expression plasmid adjacent to a cDNA encoding a DNA-binding domain (DBD) of a transcription factor (e.g., Gal4), yielding a DBD-fusion protein when expressed in cells.
- This plasmid is introduced into all yeast cells.
- A library of cDNAs is cloned into expression plasmids adjacent to a cDNA encoding a transcription factor activation domain (AD), yielding AD-fusion proteins when expressed in cells.
- The plasmids are introduced into the same yeast cells so that each one will have one.
- Protein–protein interactions between DBD- and AD-fusion proteins.
- The cells are grown on plates containing X-gal and each cell form a colony.
- If colonies turn blue, there is interaction.
- If colonies stay white, there is no interaction.



رسالة من الفريق العلمي:

عن النبيّ ﷺ قال: "مثلُ المؤمنين في تَوادِّهم ، وتَرَاحُمِهِم ، وتعاطُفِهِمْ . مثلُ الجسَدِ إذا اشتكَى منْهُ عضوٌ تدَاعَى لَهُ سائِرُ الجسَدِ بالسَّهَرِ والْحُمَّى"

والمعنى: أن المسلمين يستشعرون آلام بعضهم ومصائبهم بالعون وتقديم مساعدة بعضهم بعضًا، كمثل الجسد الواحد، إذا مرض منه عضو انهار له سائر جسده، وهذا تنبيه للمسلمين بأن يكونوا كذلك في جميع شؤونهم.

For any feedback, scan the code or click on it.

Corrections from previous versions:

Versions	Slide # and Place of Error	Before Correction	After Correction
V0 → V1			
V1 → V2			