MOLECULAR BIOLOGY

بسم الله الرحمن الرحيم

FINAL – Lecture 18

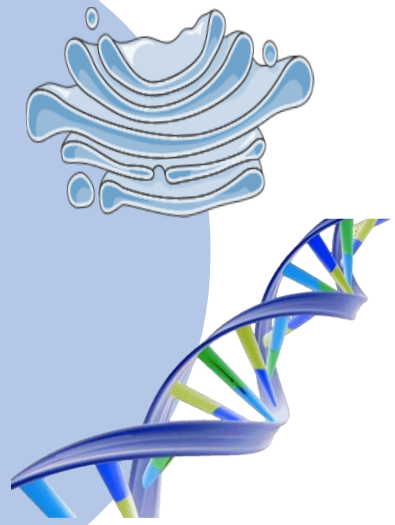
# Recombinant Proteins

﴿ وَإِن تَتَوَلَّوْا يَسْتَبْدِلْ قَوْمًا غَيْرَكُمْ ثُمَّ لَا يَكُونُوا أَمْثَالَكُم ﴾

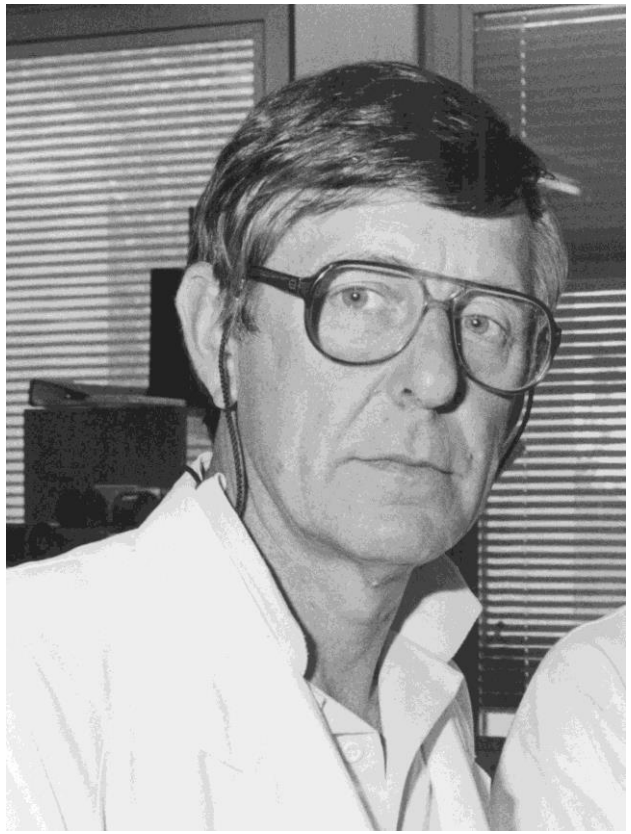
اللهم استعملنا ولا تستبدلنا


Written by :

- Raya Al Weshah
- Aya Altaki

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



# Expression of human proteins in bacteria

**You know the drill :)**

**إذا لسا ما صليت قوم صلي وادعي لإخواننا في غزة**

**بسم الله نبدأ**

# Brief introduction

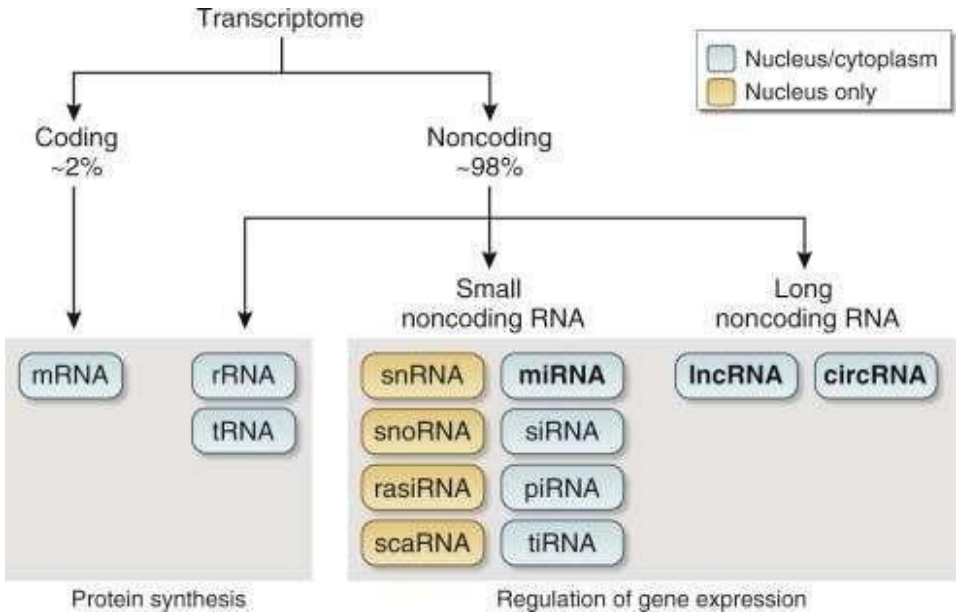
In previous lectures, we dove into cDNA libraries and cloning human DNA in bacteria by way of cloning vectors. In order to express a specific gene in bacteria, the gene's mRNA must be selected. However, there are two problems when it comes to using human mRNA for gene expression in bacteria, both of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

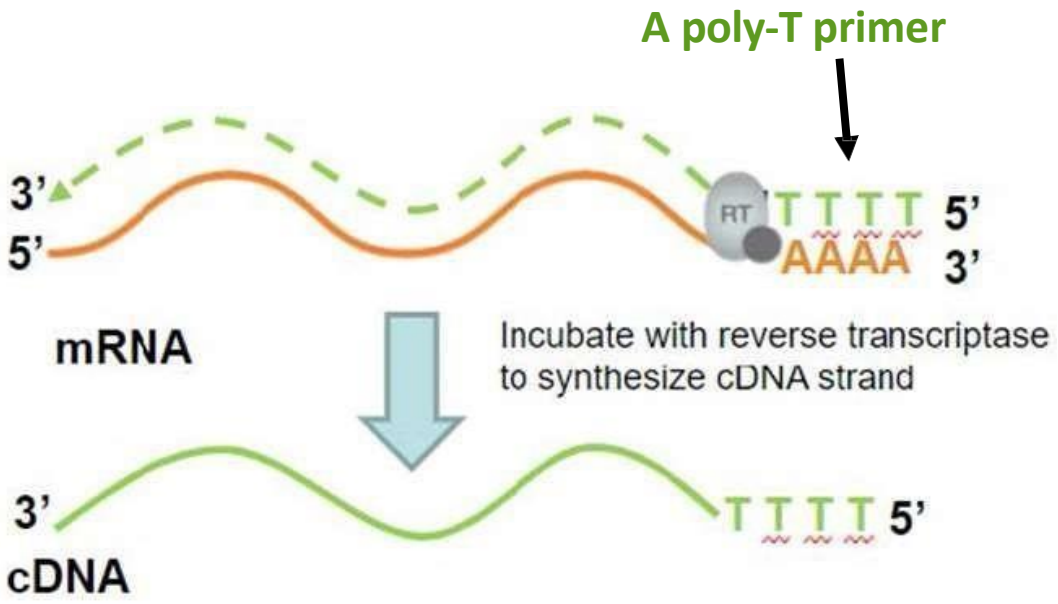


# How do we select for human mRNA?

## 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**

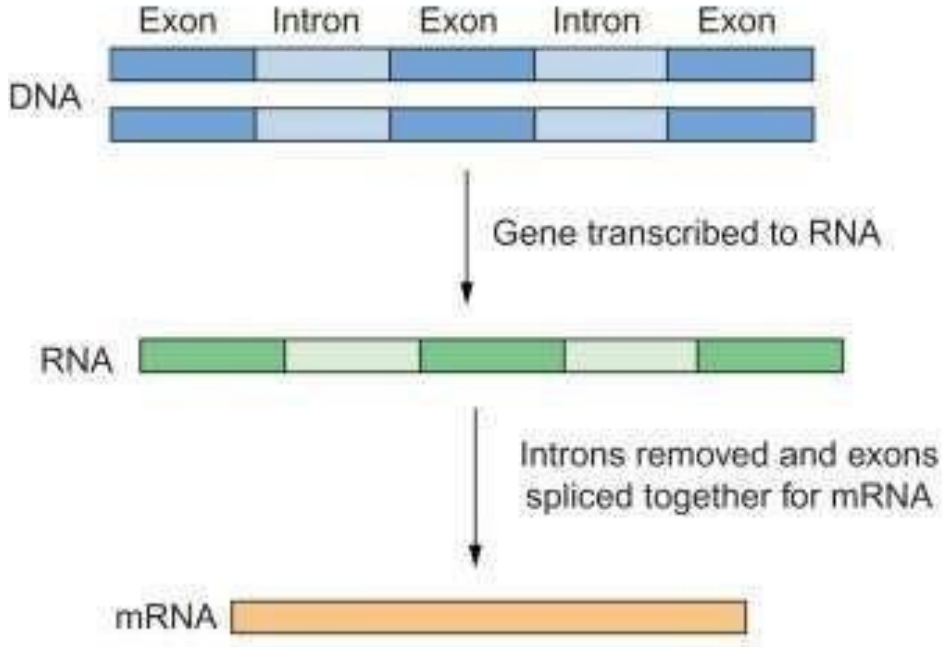


# How do we deselect introns?

## The power of reverse transcriptase (part 2)

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.

**Go for mature mRNA**



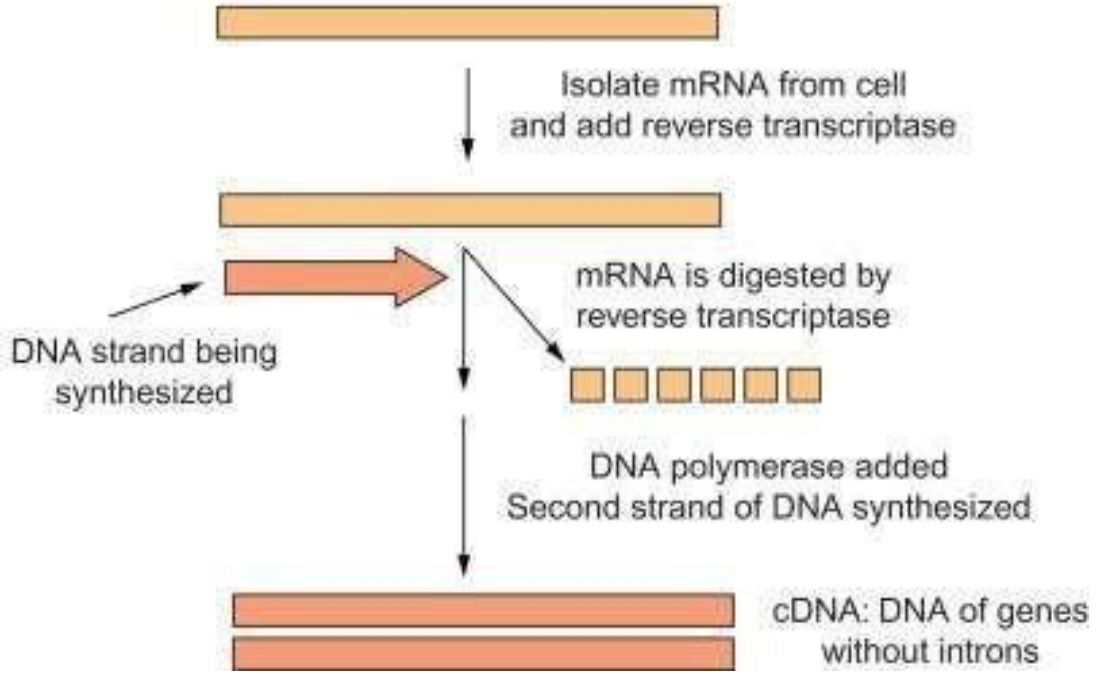

اللهم بارك لي في وقتي وأصلح لي شأني ولا تكلني

# How do we deselect introns?

## The power of reverse transcriptase (part 2)

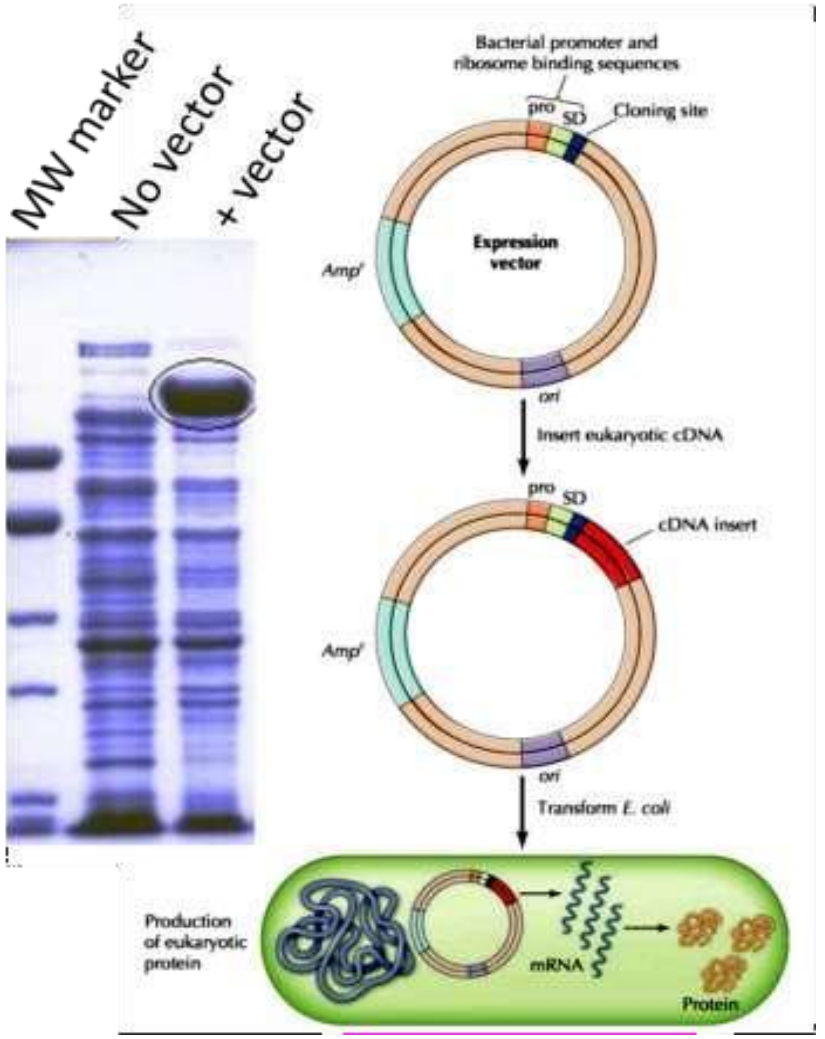
**Reverse transcriptase** is used to synthesize complementary dsDNA from mRNA. This complementary DNA, without introns or a poly-A-tail, is cloned into a plasmid. Remember, bacteria do not have introns, exons, or mRNA molecules with poly-A-tails.

**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

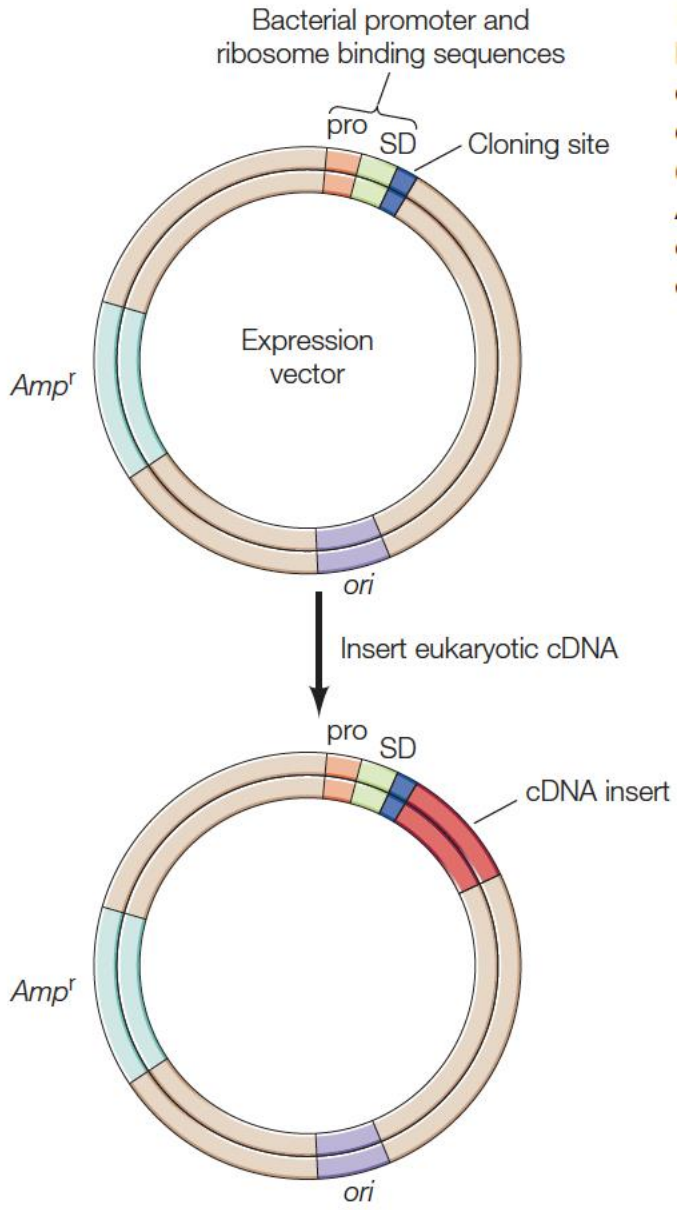



لا إله إلا الله، وحده لا شريك له، له الملك، وله الحمد، وهو

# Expression vectors

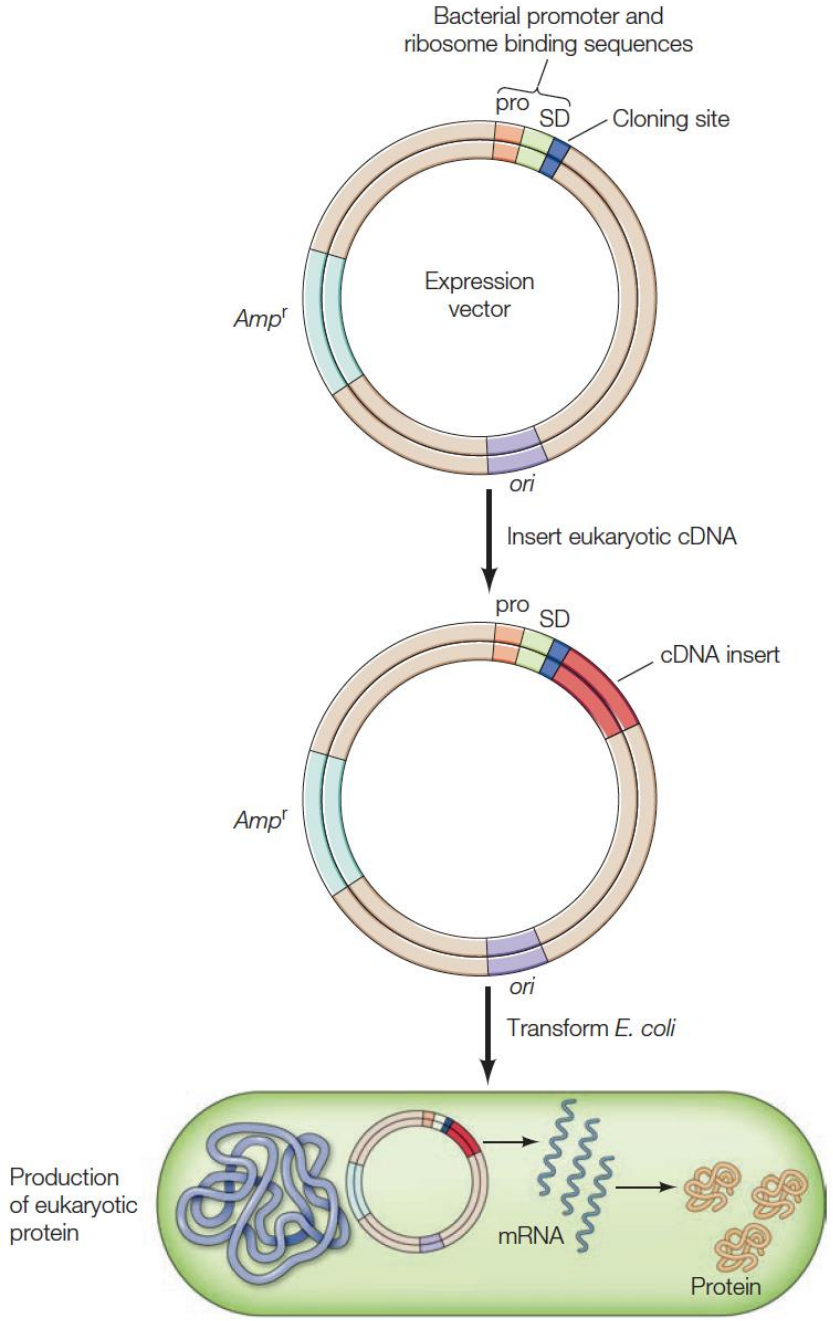
Recall DNA cloning. The vectors used are called **cloning vectors.** The purpose of these cloning vectors is to amplify DNA in bacterial cells, regardless of what the DNA could be (introns, exons, promoters, enhancers, etc.). When discussing **expression vectors,** it means the expression of human proteins, 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:

1) **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.
2) **Ribosomal binding sequences**: once the gene is expressed, the **Shine-Dalgarno sequence** is needed for the ribosome to bind to.
3) **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

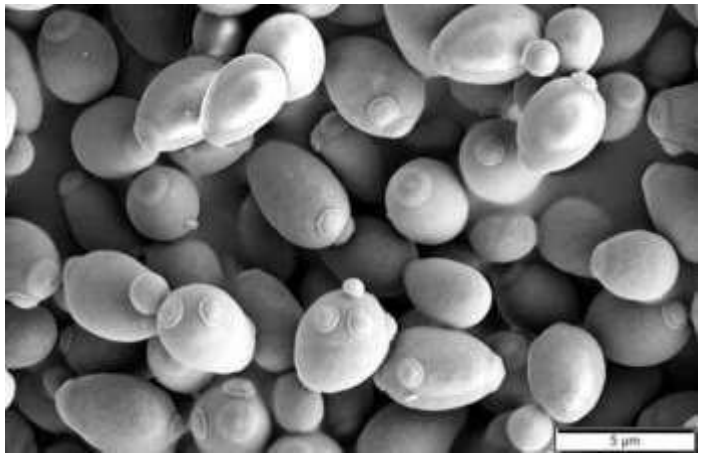
The protein can be extracted from the bacterial cell and put through SDS-PAGE to be filtered through size. Look at the following picture. The “No vector” column represents all bacterial proteins, taken from bacteria that do not contain the plasmid. Comparatively, the ”+ vector” column expresses proteins 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.

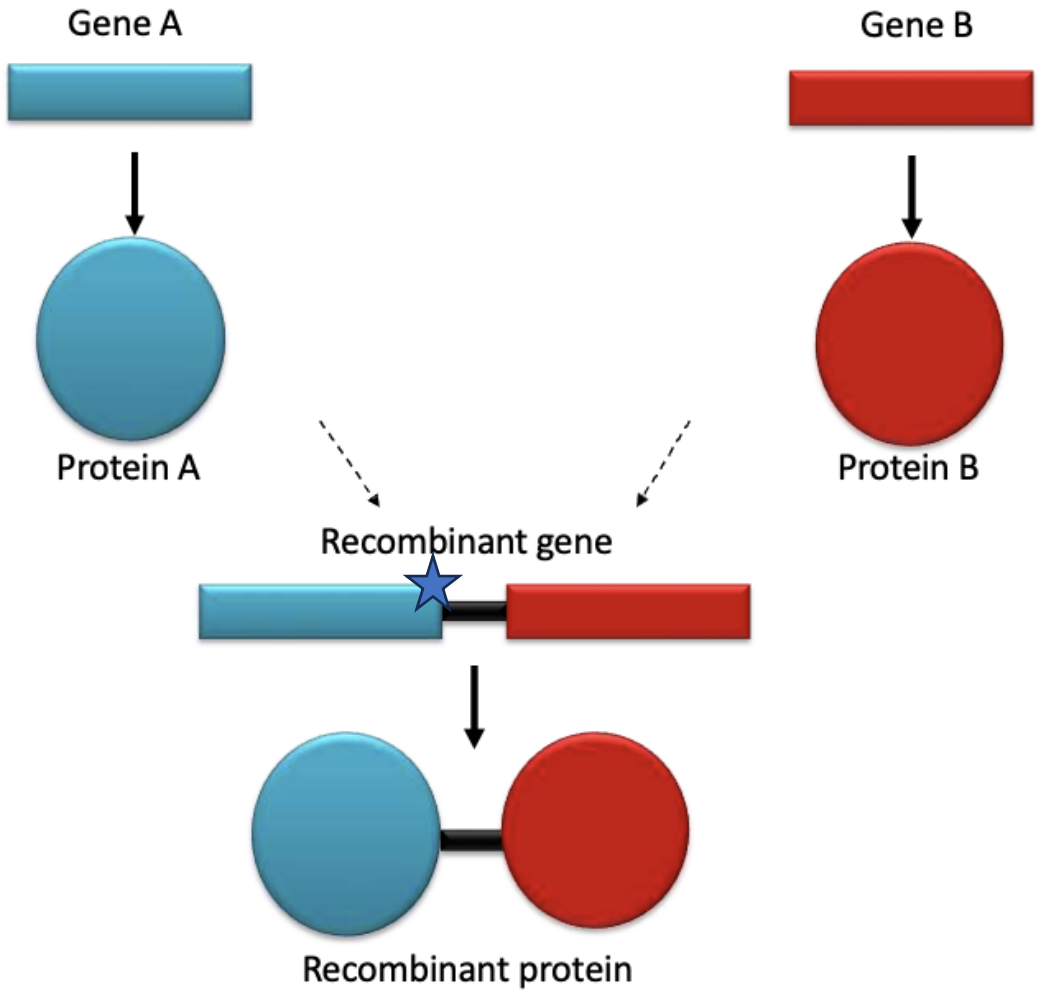
- Solution: use a eukaryotic system such as yeast

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

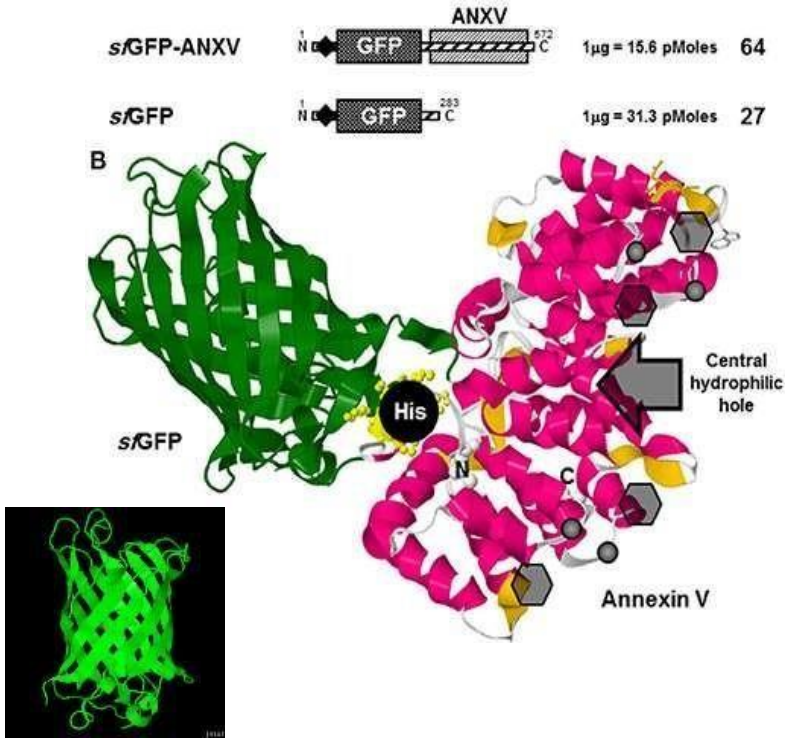
A recombinant protein is a protein that contains two different proteins. Separately, gene A and gene B produce proteins A and B, respectively. However, when these genes are fused together, they from a recombinant gene. The stop codon is deleted from gene A, so that when a protein is eventually translated, 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.




سبحان الله العظيم وبحمده عدد خلقه، سبحان الله رضا نفسه، سبحان الله زنة عرشه

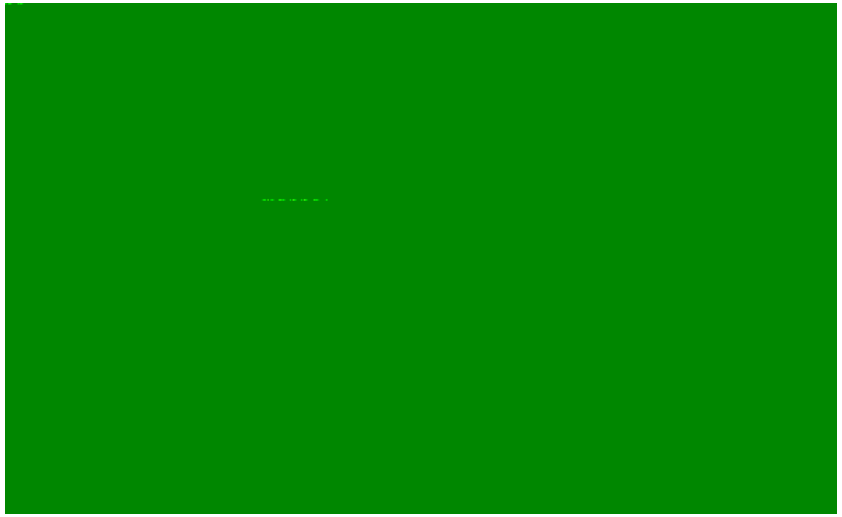
# GFP-tagged proteins

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



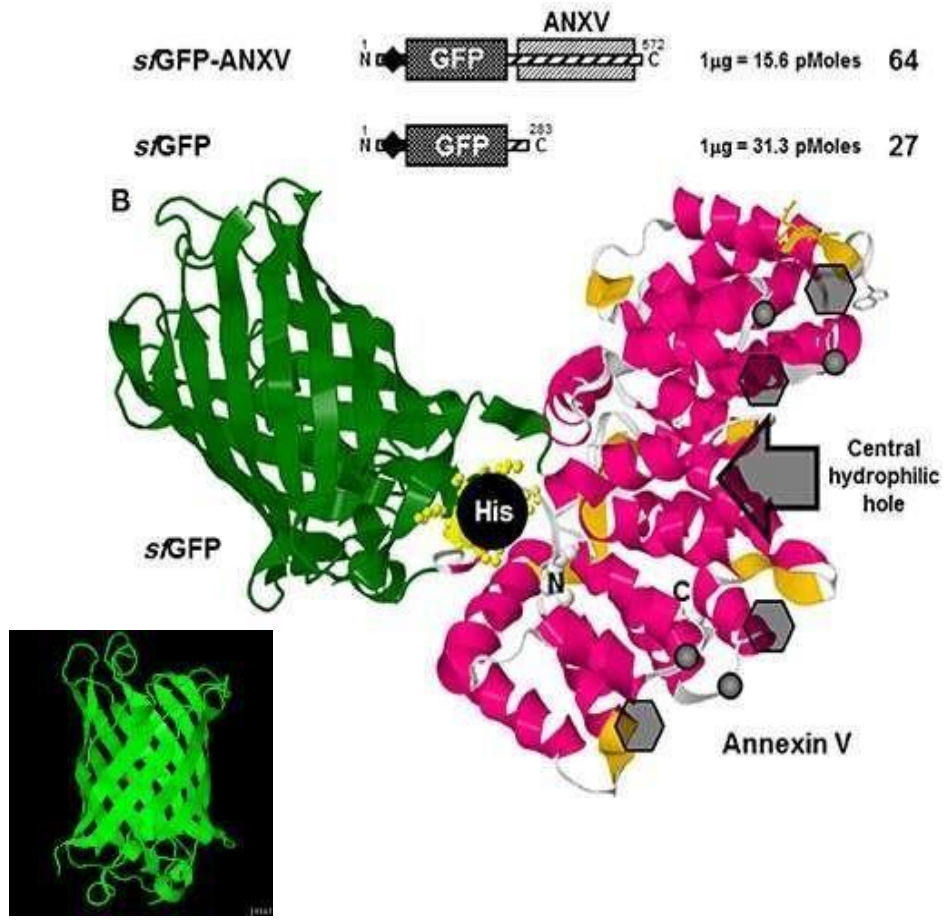


You can view this animation in the recorded lecture (19:46-20:11)



# GFP-Tagged Proteins

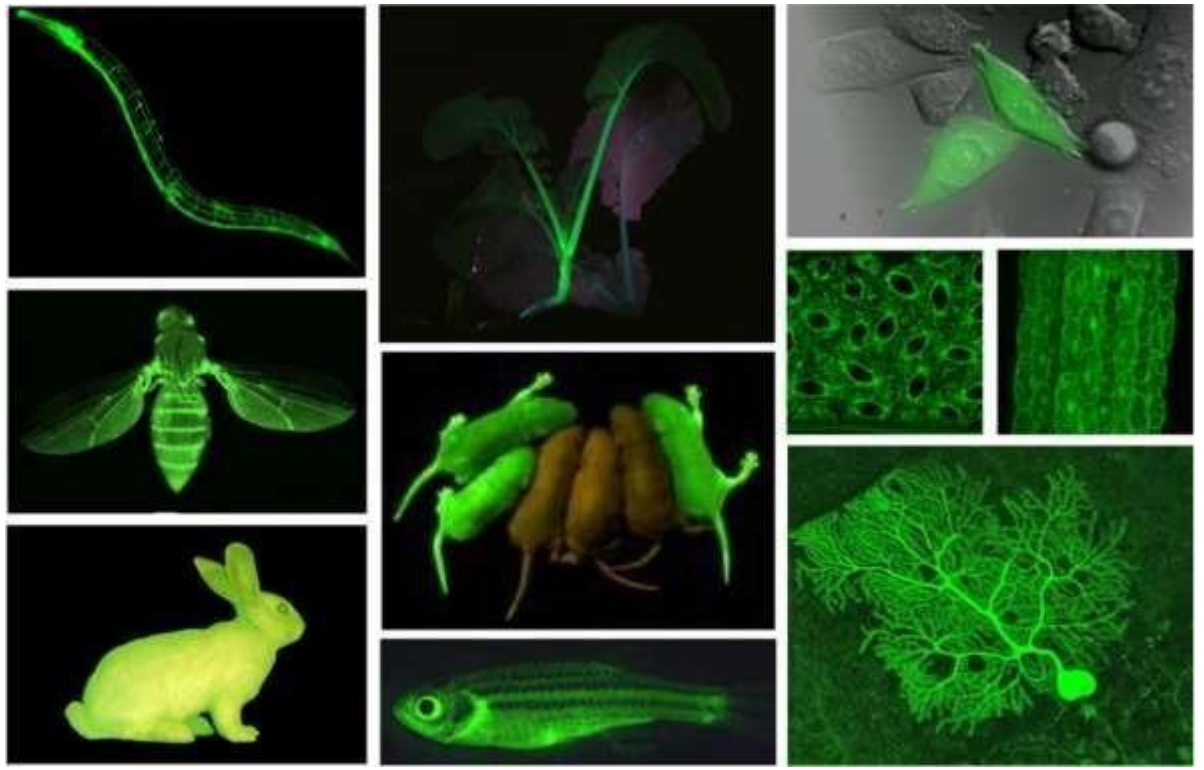
One example of a recombinant protein is the green fluorescent protein (GFP). This protein is isolated from jellyfish and can produce fluorescence. This gene can be cloned upstream or downstream of another gene, the protein of interest, by its insertion in an expression vector, then transcription & translation. 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.

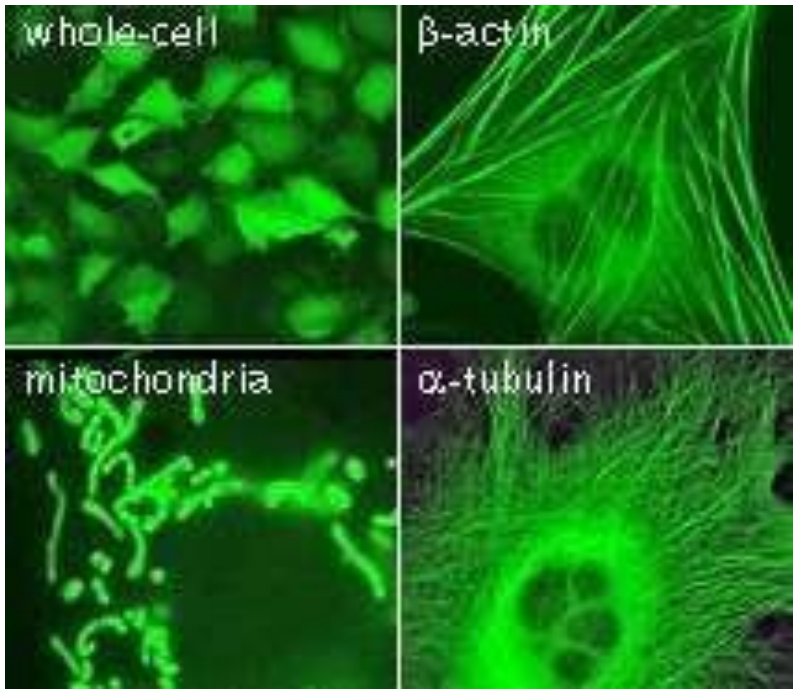


# A world of possibilities

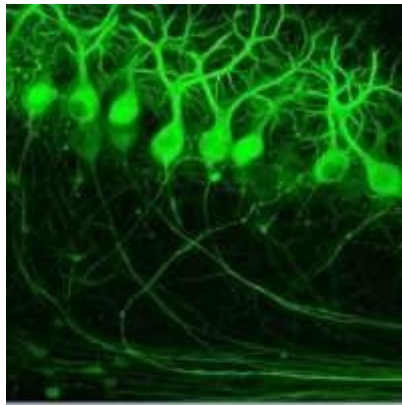
These are some examples of proteins that have been **tagged** or have another protein (GFP) attached to them.

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






Nerve cell fluorescence helps us comprehend neural connections

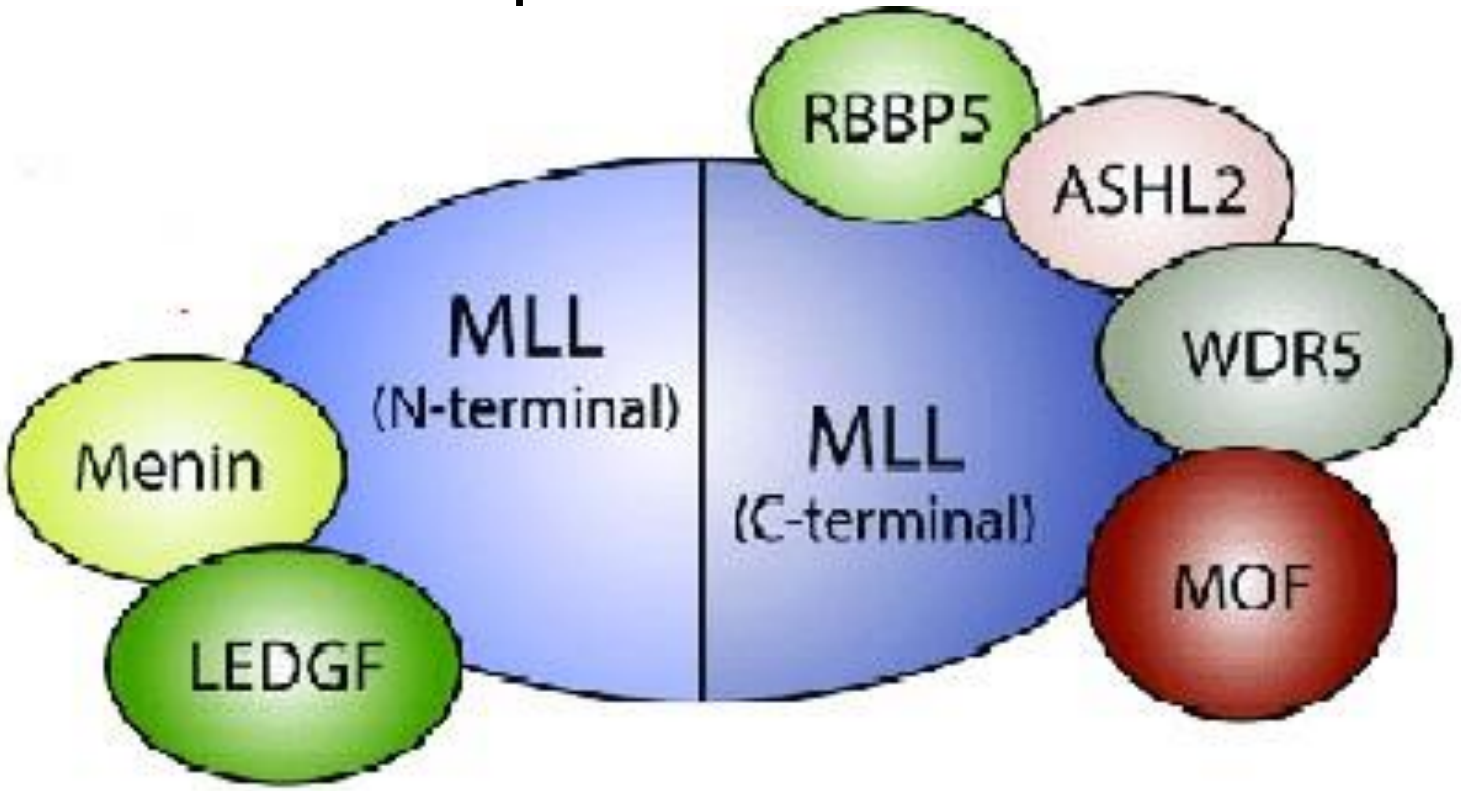


# Protein-protein interaction

## Co-immunoprecipitation

How we can understand protein protein interaction? now we will introduce two techniques :one of them is known as co-immune precipitation, precipitation means: you know it settles down ( ترسيب)
, immune:once you heard the word immune it means I am using antibodies here ,co : means together
So, I am precipitating a protein with another protein that is accompanied by another protein using antibodies

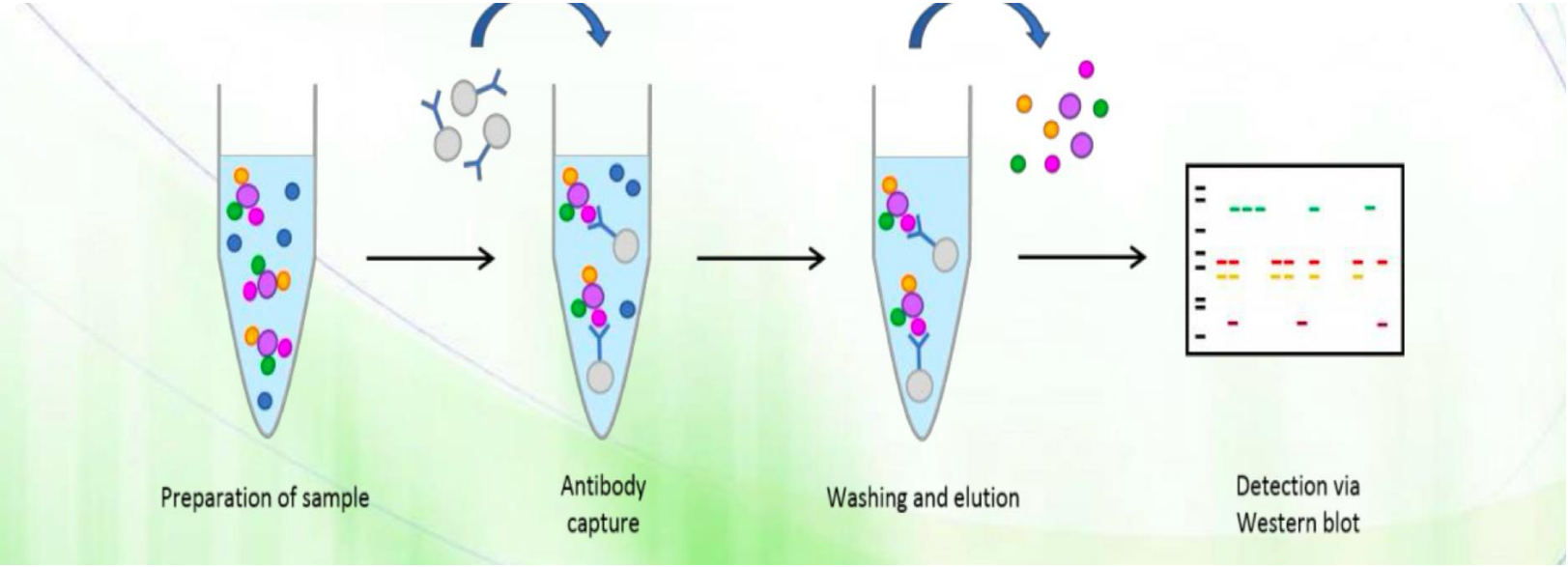
# Proteins form complexes

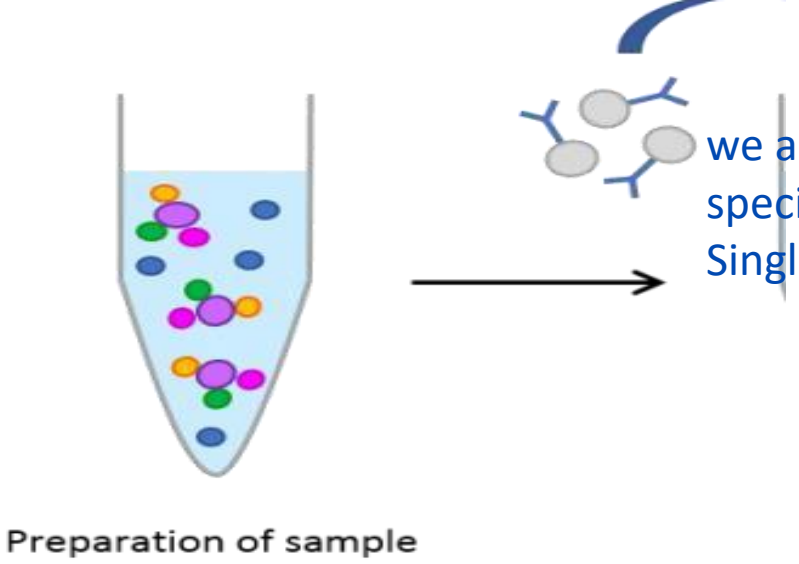



Remember that we can have a large protein let's say this protein MLL has two domains : an N –terminal domain
And c-terminal domain and the N –terminal domain can interact with these different proteins and these proteins can also interact with each other. Now the c-terminal domain can interact with four 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?so with the immunoprecipation that's what we do

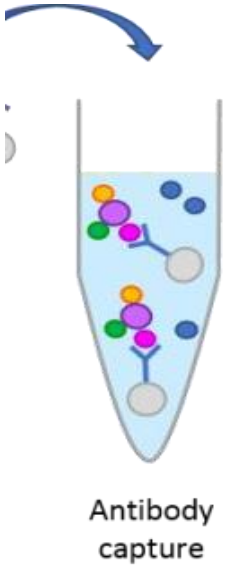
# (Co)-Immunoprecipitation

- 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).



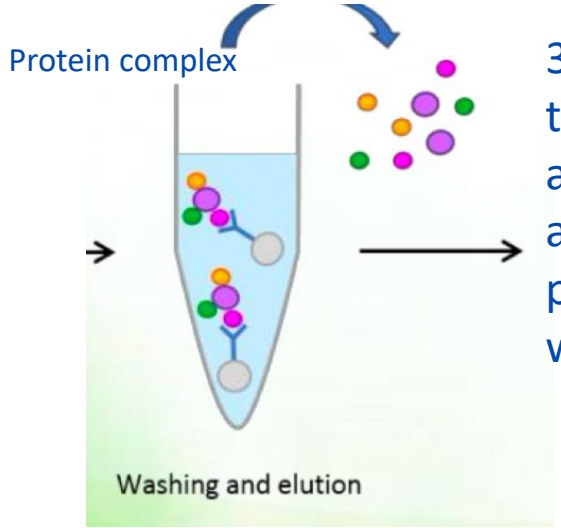



we add an antibody which is
specific to one
Single protein

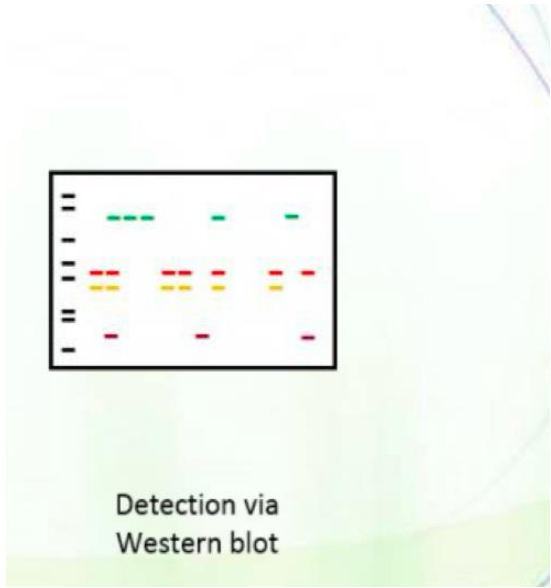



2. This antibody then targets this protein(
the pink protein ) which
Interacts with other proteins,so what
Will happen is that this ( antibody )
Is really quite heavy and it will precipitate
( going down )

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




3. We remove every thing else
that doesn't interact with the
antibody would be removed
and then we have these
protein complexes that interact
with the pink protein




4. We can take complex
, study it using for example
SDs page gel electrophoresis
Or using western blot they
Are immune blots , we separate
Protein on size , then transfer
To a membrane and we add a
Primary antibody and secondary
Antibody to detect the presence of
A protein and we can also determine
The size

# Protein-protein interaction

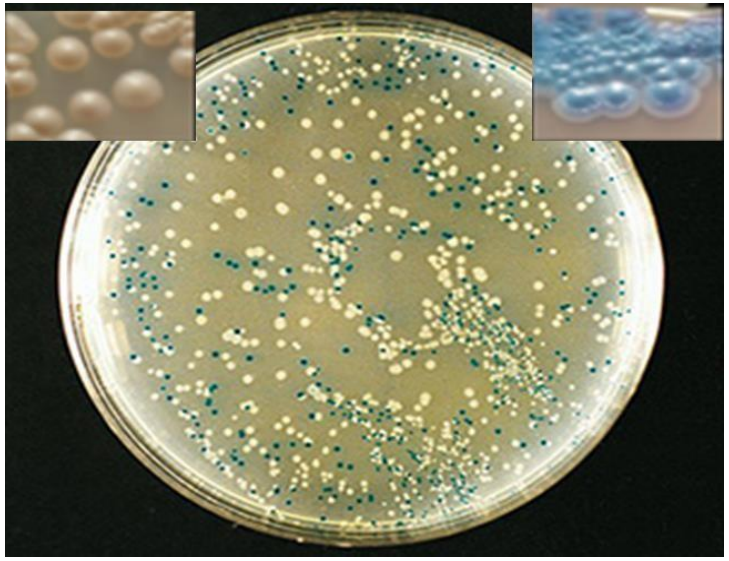
## Yeast two-hybrid system

### starting from a cDNA library

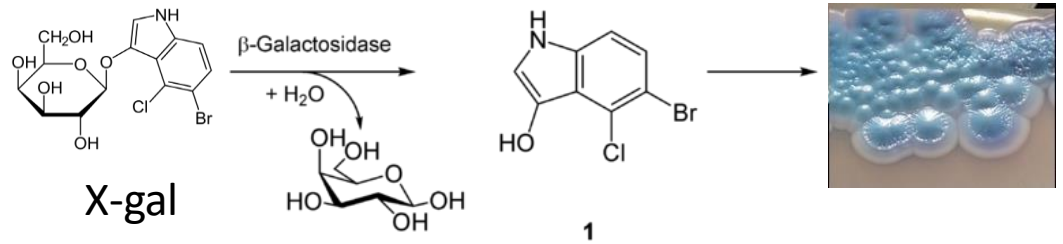
Let’s go into something a bit more sophisticated  (yeast two-hybrid system) it’s a : genetic system where we express hybrids of two proteins or two genes .

# 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 there is sort of like a regulatory sequence called: UAS , this like a promoter , promoter proximal element . It's the binding site of the transcription factor ( the transcription factor is the pink one in the previous slide) this transcription factor is known as the Gal4 . It 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 see that's the

Power of genetic engineering. The gene that's 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 we use a regulatory element from yeast and we put it in a DNA upstream of a reporter gene and the

Reporter gene is the bacterial 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) . So it's look like lactose except that

. Beta galactosidase would cleave x gal producing this molecule (look at the previous slide in the bottom number 2) which gives a blue colour so the yeast cell become bluish .

Let's remember the definition of domain. : is a three dimensional structure of a protein and it can fold independently of the rest of the protein and it maintains its function

- To sum up, we have this genetic system, the reporter gene is now under control the uas genetic system. The uas system is the binding site of gal 4

For transcription factor which has a DNA binding domain and an activation domain . If gal 4 binds to the uas then we have expression of the reporter gene

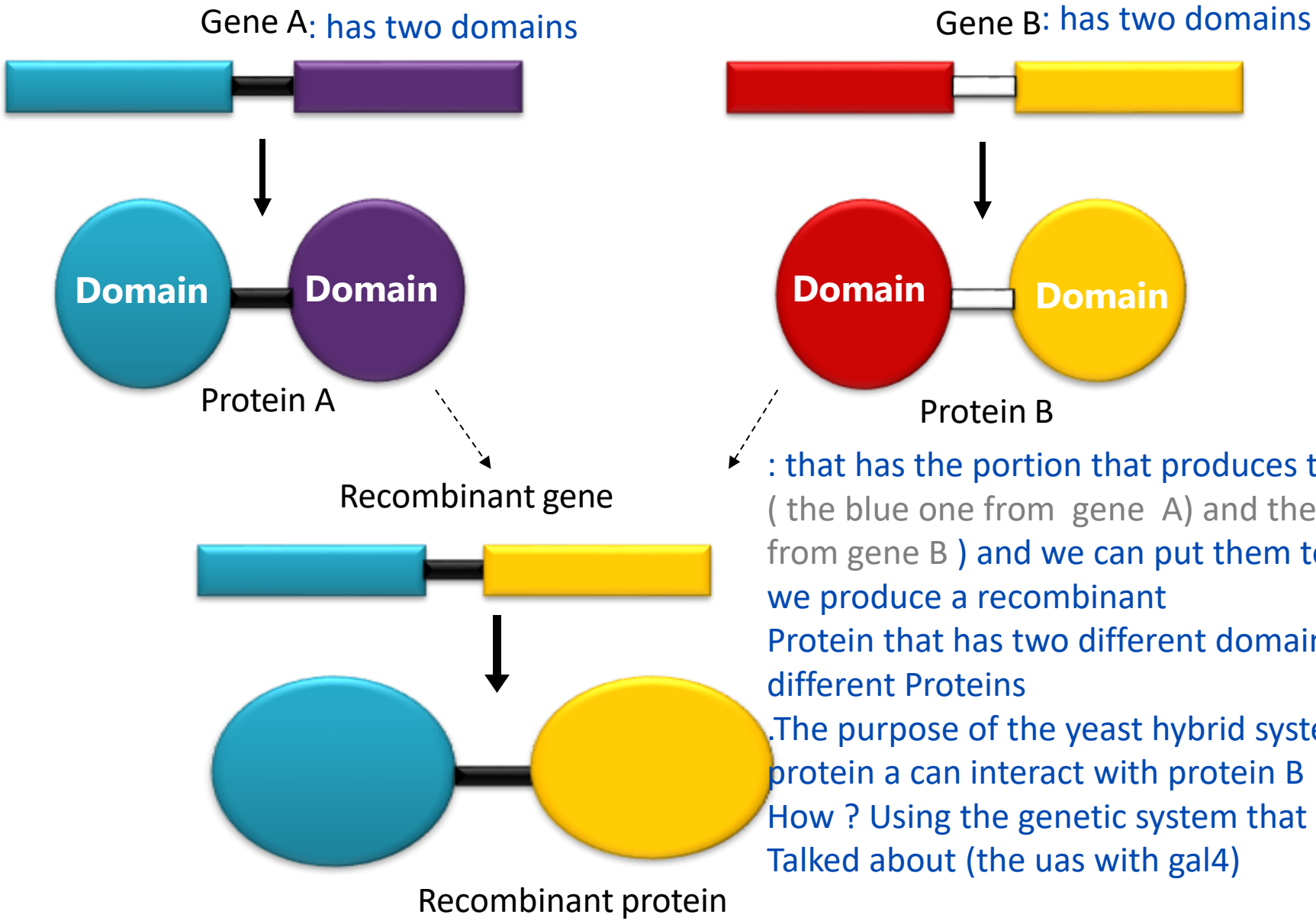
By the RNA polymerase , then translation of the mRNA into beta galactosidase

.if we add x gal to the yeast cells , they can take up x gal ,cleave it and they can

Produce this substrate .it gives us a blue colour . Remember, that 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 doesn’t express beta galactosidase

# Production of a recombinant protein



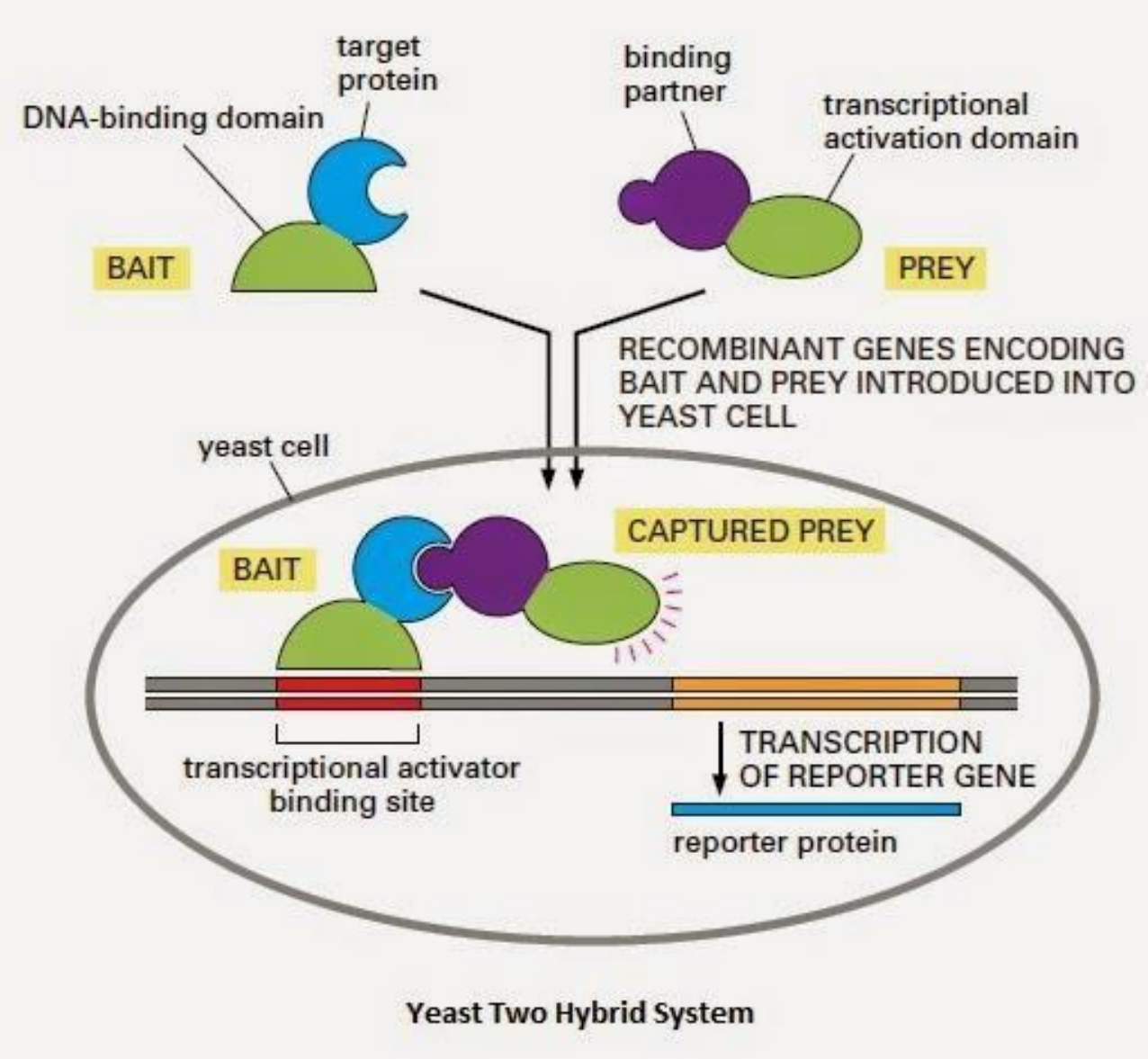

: that has the portion that produces this domain ( the blue one from gene A) and the yellow one from gene B ) and we can put them together so we produce a recombinant

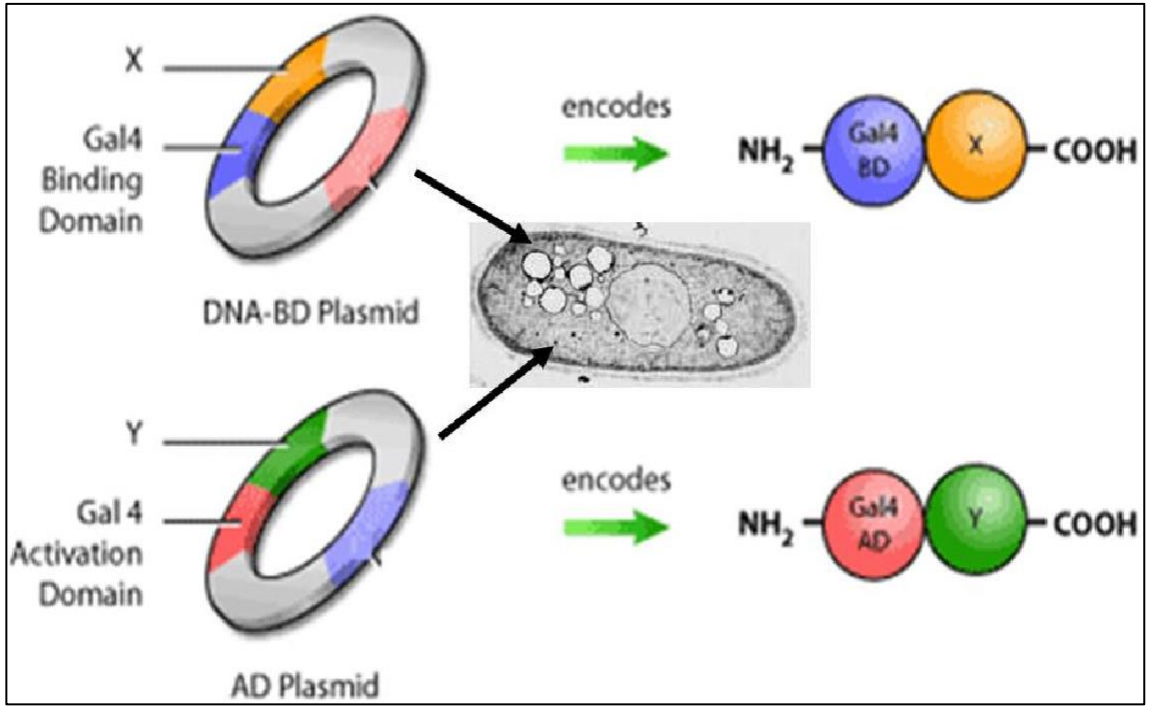
Protein that has two different domains from two different Proteins

.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 Talked about (the uas with gal4)

# Quick illustration



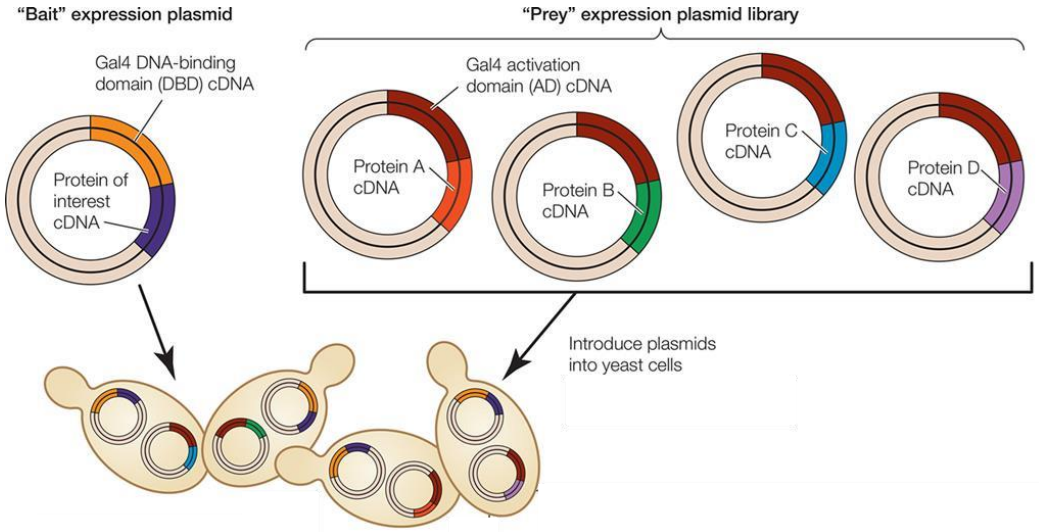



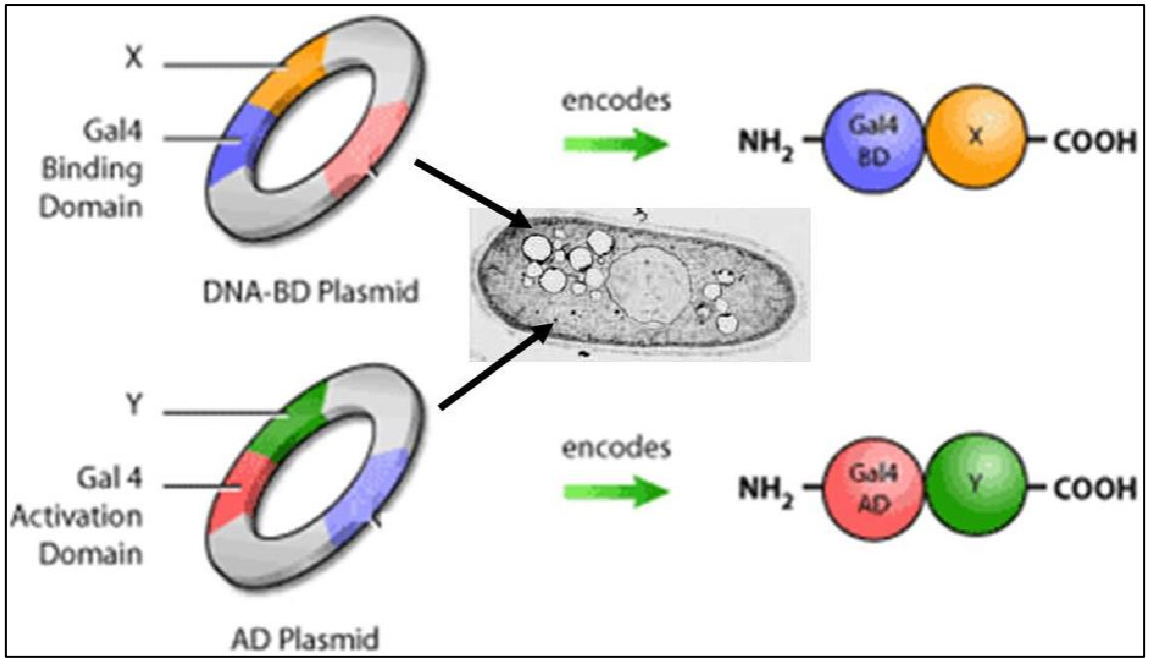
We want to test if protein x interacts with protein y . We are going to use the same genetic system that we just talked About the uas , the regulatory element, gal 4 and the lac z gene that produces beta galactosidase and that's what we Do we create two plasmids ( look at the previous slide in the right) and these two plasmids are 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 ( look at the previous slide in the right , the plasmid that is on the top) . So when x Is produced , it's produced with gal 4 DNA binding domain .
And we have plasmid two ( look at the bottom in the right in the previous slide) this plasmid contains the gene for Protein Y and it's produced as recombinant protein along with Gal 4 activation domain. So we are going to have

.Protein Y with the activation domain of gal 4 . The idea is that if x interacts with Y then these two ( look at the Previous slide (1) and (2) in the right) would be close to each other. So, we're going to have again this protein of Interest with the DNA binding domain ( (3) in the previous slide in the left) . And another protein of interest Protein y with activation domain ( (4) in the previous slide). if these two interact to each other then we're going to have the DNA binding domain and the activation domain very close to each other and we will have transcription of Lac z . If lac z is expressed we will have beta galactosidase, which if it's expressed, it will cleave x gal , if x gal Is cleaved , it will generate a blue product and the yeast cell , and colony become bluish . The colour blue indicates That x interacts with 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 DNA Galactosidase, no cleavage of x- gal , no blue product and the colony would look whitish not bluish

# 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 harder,I want to identify all of the proteins that can interact with x not a single protein .so, we create a CDNA library. It’s a collection of plasmids and each plasmid contains a CDNA representing a specific gene . In the figure, this is a CDNA library,. Different plasmids that

Can produce proteins, recombinant proteins.each protein is produced with

Activation domain of gal 4 . I have one single plasmid that can produce

Protein x and this protein x is produced a recombinant protein with the DNA

Binding domain of gal4. we take this (number 1 in the figure) and we introduce it into yeast cells so this plasmid ( number 1 in the figure) that produces recombinant x would exist in every yeast cell and I have

Let’s say 1 million yeast cells each one of them has the plasmid for the recombinant x gene or ( that can produce recombinant x protein with the DNA binding domain and then I take these one million yeast cells

And I insert these plasmids( look at the figure) into the yeast cell. So, each yeast cell

Would take up one of these plasmids . So each yeast cell would have two plasmids

The plasmid that produces 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 x interacts with a then this yeast cell would have the

Production of beta galactosidase ( can cleave x gal. ) into a bluish

Product and the colony would look blue, if it isn’t interact with protein b the

Binding and activation domain will be far away from each other (no cleavage of x

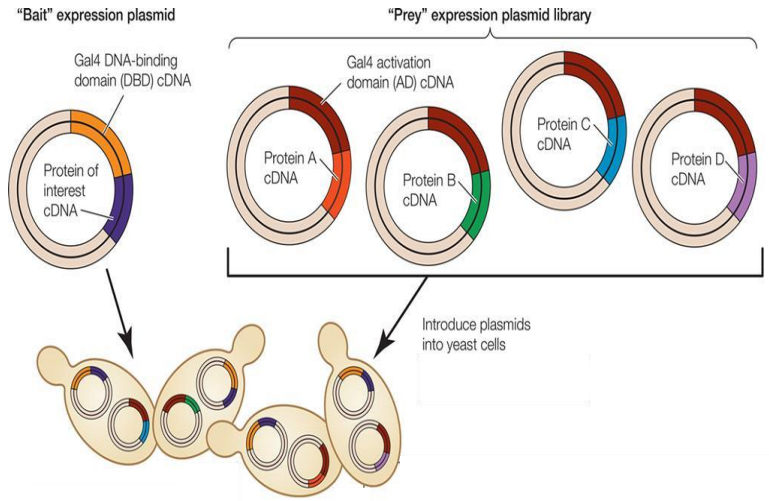
Gal and the colony will look white)

We taking the blue colonies ,

Isolating the plasmids , and we see the identity of gene that is contained in

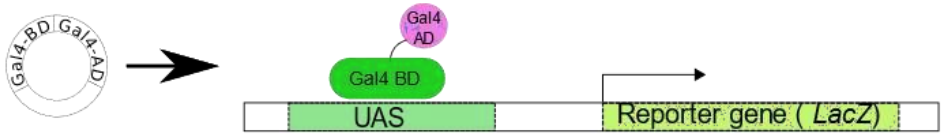
These plasmids, we can identify what the gene , this gene produces protein

That interacts with x

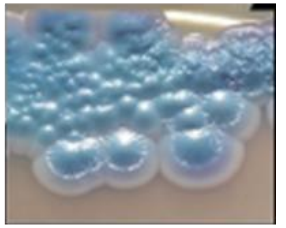


# The possibilities and outcomes

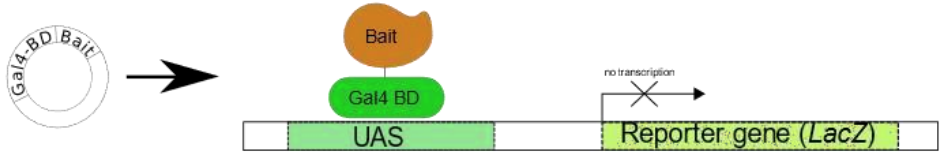
If we have a plasmid that produces the full gal4
Protein. The binding domain (the green one)and the activation domain (the pink one ), then lac z expressed , beta galactosidase would be produced and if we add x gal to this yeast cell we will have a blue colony



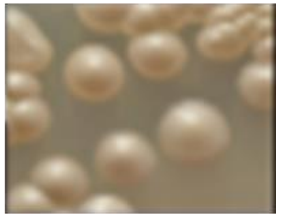
A. Regular transcription of the reporter gene



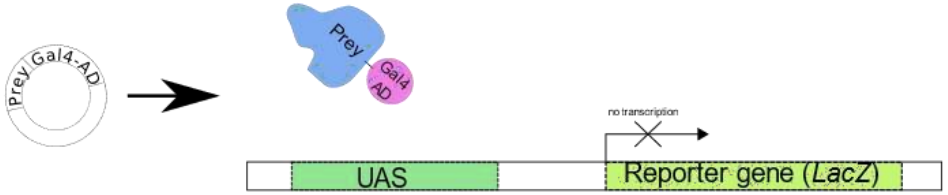
B. If we have plasmid that produces DNA
Binding domain only there would be no production



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

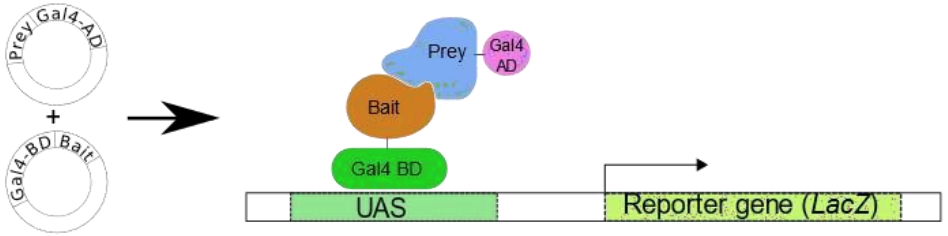


C. If we have plasmid that produces the activation domain only we will have no production of beta galactosidase

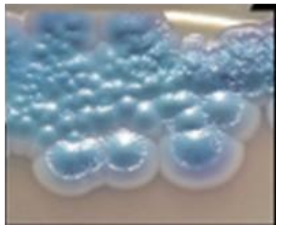


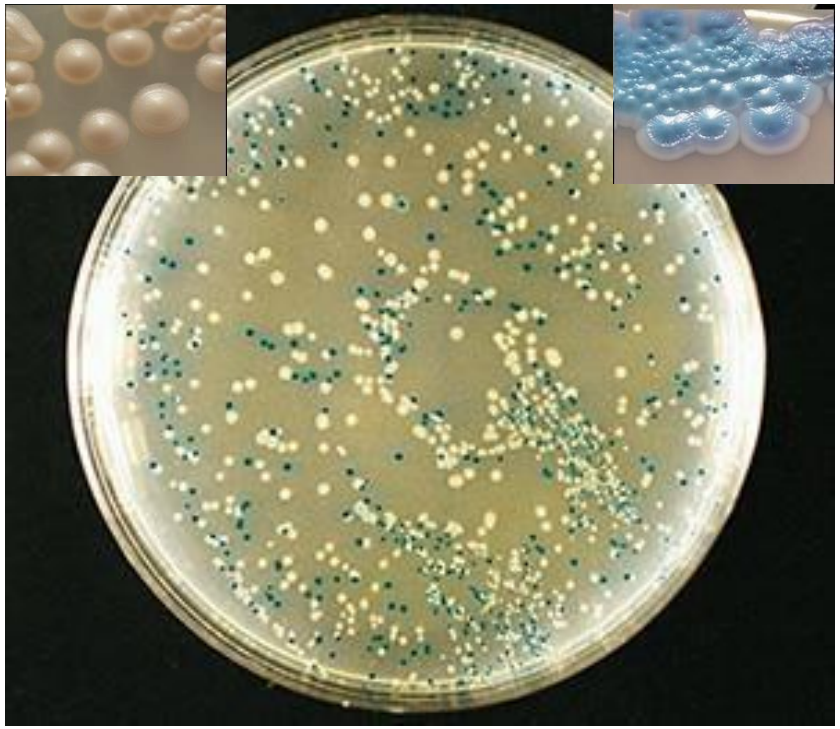
C. One fusion protein only (Gal4-AD + Prey) - no transcription

D. If we produce protein x , protein y and these two proteins
Can interact with each other with the DNA binding domain
Being recombine, 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 that contains this gene bcz I know
Protein x , we will identify this gene that can interact with protein x
.we can do sequencing , immune blots to identify the gene



D. Two fusion proteins with interacting Bait and Prey



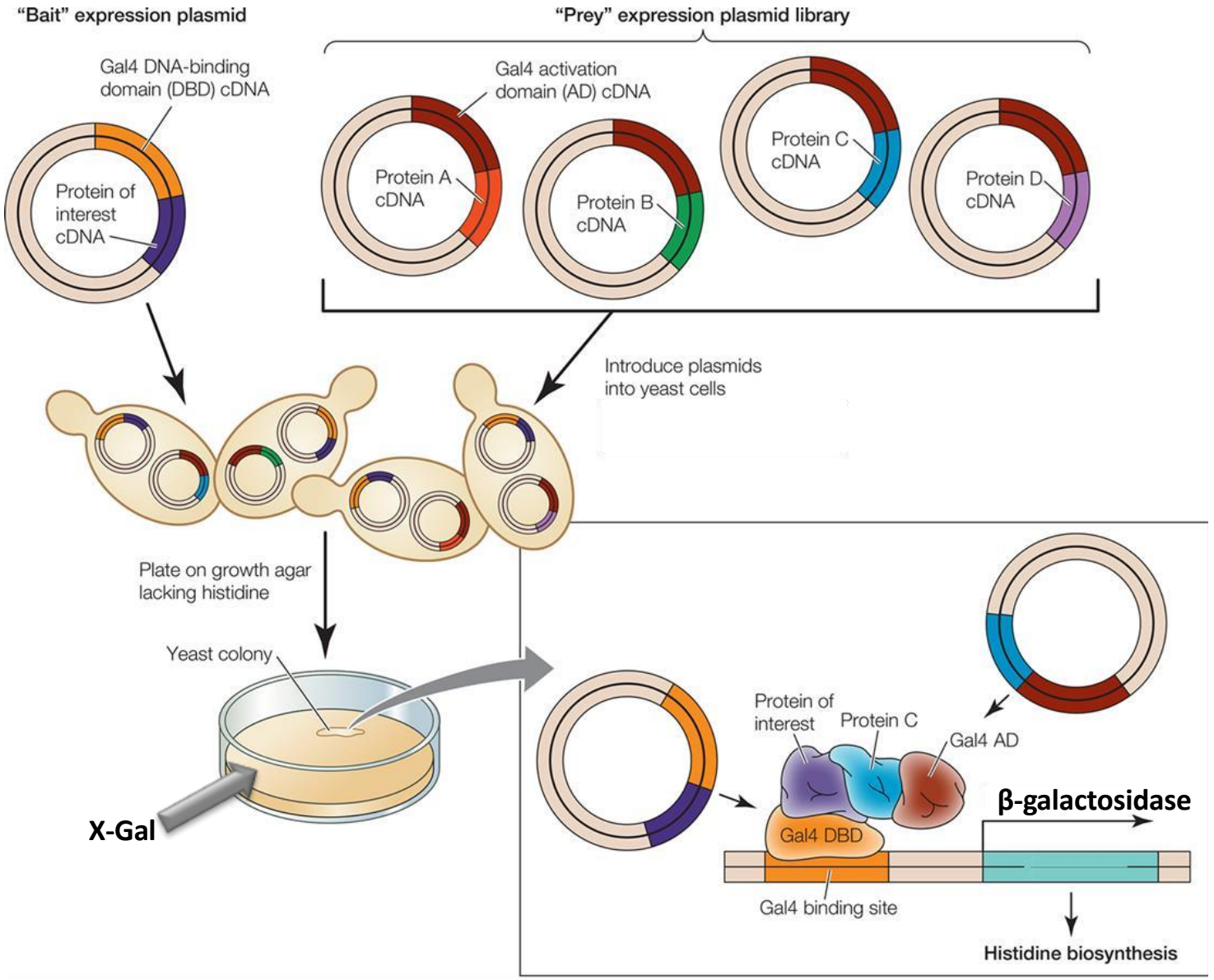


- Blue yeast colonies are picked and plasmids are isolated to identify the unknown genes/proteins that interact with the known gene/protein.

# The procedure

The doctor said 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.



# 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 | | | |

# Additional Resources:

Reference Used:
(numbered in order as cited in the text)

1. First reference

Extra References for the Reader to Use:

1. GFP

# رسالة من الفريق العلمي:

قال مالك بن أنس رحمه الله :
«السنَّةُ سفينةُ نوحٍ مَن ركبها نجا
ومن تخلَّف عنها غرق»

من سنن الرسول صل الله عليه
وسلم:
إزالة الأذى عن الطريق
تجنب الغضب
مساعدة الأهل في أعمال المنزل
التبسم :)