Swayam Course - **Analytical Techniques**

Week 13, Tutorial 33 - **Expression Techniques of Recombinant proteins using Bacteria and yeast**

Content Writer - **Dr. Ruby Dhar, DBT-Biocare Fellow, Department of Biochemistry, All India Institute of Medical Sciences, New Delhi**

OBJECTIVES:

1.Introduction
- **1.1 What are recombinant proteins**
- **1.2 Why are recombinant proteins important**

2. Various protein expression system

3. Cloning vectors
- **3.1 Properties of cloning vector**
- **3.2 Types of cloning vectors**

4. Expression vector
- **4.1 Properties of expression vector**
- **4.2 How do expression vector work**

5. Basic layout of recombinant protein expression

6. Basic steps of protein purification

What are protein tags and Purification of tagged proteins

7. Expression of recombinant proteins in bacteria

8. Expression of recombinant proteins in yeast

Introduction:

Recombinant Protein is a protein encoded by a gene — recombinant DNA — that has been cloned in a system that supports expression of the gene and translation of messenger RNA. The formation of recombinant protein is carried out in specialized vehicles known as vectors. Recombinant technology is the process involved in the formation of recombinant protein.

These proteins are produced to scale up the production of protein, generating mutant variety of the protein and large scale production for commercial use.

The expression and purification of recombinant proteins using bacterial vectors is a mature and preferred system to obtain folded and stable proteins. However, functional post-translational protein modifications, such as glycosylation or phosphorylation, can only be achieved using eukaryotic expression systems. In addition, insolubility is another challenge when using proteins expressed in *Escherichia coli*, such as certain intrinsically disordered proteins, which are more prone to aggregation than folded proteins. Eukaryotic protein expression systems, including human cells, baculovirus/insect cells, and yeast, have become indispensable for the production of functional eukaryotic proteins.

This module discusses the expression of recombinant proteins in bacteria and yeast.

There is a separate module which discusses the recombinant protein expression in mammalian cells and insect cells.

1.1 What are recombinant proteins?

Recombinant Protein is a protein encoded by a gene — recombinant DNA — that has been cloned in a system that supports expression of the gene and translation of messenger RNA. Recombinant protein is a manipulated form of protein, which is generated in various ways to produce large quantities of proteins, modify gene sequences and manufacture useful commercial products. The formation of recombinant protein is carried out in specialized vehicles known as vectors. Recombinant technology is the process involved in the formation of recombinant protein. Recombinant proteins can be used to identify and locate the position of the protein encoded by a specific gene, and to uncover the function of other genes in various cellular activities such as cell signaling, metabolism, growth, replication and death, transcription, translation, and protein modification

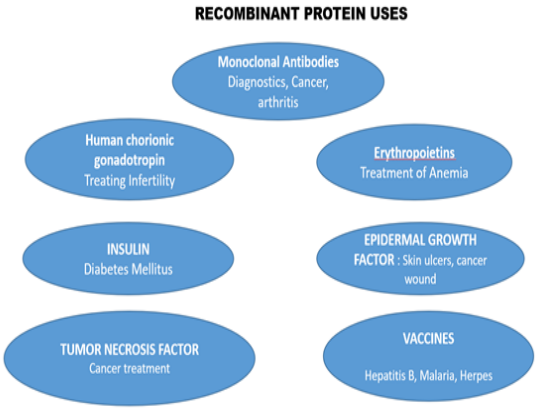
Recombinant protein is encoded by recombinant DNA, which has been cloned in a foreign expression system to supports the expression of the exogenous gene. This recombinant DNA construct can be used to manufacture large quantities of useful protein products. The recombinant DNA, usually the cDNA sequence of the target protein, is designed to be under the control of a well-characterized promoter to express the target protein within the chosen host cell to achieve high-level protein expression. Modification of the gene by recombinant DNA technology can lead to expression of a mutant protein. Proteins coexpressed in bacteria will not possess post-translational modifications, e.g. phosphorylation or glycosylation; eukaryotic expression systems are needed for this kind of modification of expressed proteins.

1. 2 Why are recombinant proteins Important?

Here are few important uses of expressing recombinant proteins

1. To identify the polypeptide coded by a DNA sequence
2. To analyse the biological activity
3. To study the structure -function relationships and interactions
4. To produce vaccines
5. To produce biotechnological enzymes
6. To produce therapeutic proteins
7. To produce target specific drugs
8. To raise specific antibodies
9. To study 3D structure
10. To do protein engineering and design

Examples of recombinant protein expressed using bacteria and yeast.



RECOMBINANT PROTEIN : HUMAN INSULIN

The synthesis of human insulin gene was accomplished by four scientists at City of Hope Medical Center led by Roberto Crea, Ph.D., and Keichi Itakura, Ph.D. Scientists at Genentech, led by David Goeddel, Ph.D. and Dennis Kleid, Ph.D., joined the genes that were made in sections and inserted them, along with the control mechanism into the E. Coli bacterium.

Arthur Riggs, Ph.D. at the City of Hope and Dr. Goeddel of Genentech were responsible for developing the final assays, purification and joining techniques.

2. Various protein expression system

Different protein expression systems have different features and applications.

The various protein expression systems are bacteria, yeast, insect or mammalian systems. The development of genetic engineering and cloning has opened many possibilities of expression and isolation of heterologous proteins for research purposes. Considerable advances in technology have enabled expression and isolation of recombinant proteins in large scale. However, for large scale applications such as enzyme, antibody or vaccine production, the amount of protein required is considerably high. In such cases the system in which the protein is expressed must be easy to culture and maintain, grow rapidly, and produce large amounts of protein. Since mammalian proteins also undergo various post-translational modifications. These requirements led to the discovery of protein expression systems. The various protein expression systems are bacteria, yeast, insect or mammalian systems.

The following factors determine the type of expression system used to produce recombinant proteins:

- time spent in expressing the protein
- ease of handling the expression system
- amount of protein needed
- mass of the protein
- type of post-translational modifications, number of disulfide bonds
- destination of the expressed protein

Characteristic	E.Coli	Yeast	Mammalian	Insect
Proteolytic cleavage	?	?	Y	Y
glycolation	N	?	Y	?
secretion	?	Y	Y	Y
Folding	?	?	Y	Y
phosphorylation	N	?	Y	?
acetylation	N	Y	Y	?
Amidation	N	Y	Y	Y
%Yield	>50%	1%	<1%	>30%

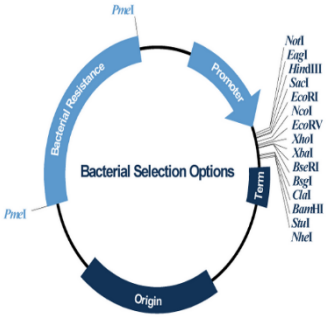
The key problem facing the researcher is that it is still not possible to predict which expression system will work best for a particular protein and a particular end-use. A universally applicable expression system does not yet exist. When selecting an expression system the researcher should bear in mind the fundamental properties of each system, their pros and cons and how any particular

limitations of that system can be overcome. Decisions should be informed by knowledge of the protein expression target/family members and the ultimate use of the recombinant protein. If resources permit, it may be prudent to explore two (or more) expression systems in parallel.

3. Cloning Vectors

What are Cloning vectors?

Cloning vectors are DNA molecules that are used to "transport" cloned sequences between biological hosts and the test tube.
•Most vectors are genetically engineered.
•A vector is used to amplify a single molecule of DNA into many copies.



Cloning vectors share four common properties:
1. Ability to replicate.
2. Contain a genetic marker for selection.
3. Unique restriction sites to facilitate cloning of insert DNA.
4. Minimum amount of nonessential DNA to optimize cloning

DIFFERENCE BETWEEN CLONING VECTORS VS EXPRESSION VECTORS:

CLONING VECTORS

- Expression vector is a plasmid which is used to introduce a specific gene into a target cell and utilizes cells mechanisms to produce the relevant gene.
- Used to obtain RNA or protein of the inserted DNA segment
- A plasmid vector
- It contains enhancers, promoter region, termination codon, transcription initiation sequence and translation initiation sequence in addition to the classical features of a cloning vector

EXPRESSION VECTORS

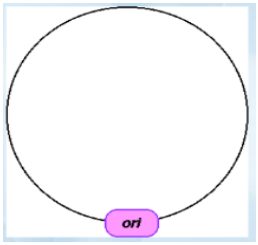
- Cloning vector is a small piece of DNA which can be stably maintained within a host cell
- Used to obtain numerous copies of the Inserted DNA segment
- Can be plasmids, cosmids, phages, BACs, YACs or MACs
- It contains origin of replication, multiple cloning sites and a selection marker

3.1 PROPERTIES OF CLONING VECTOR:

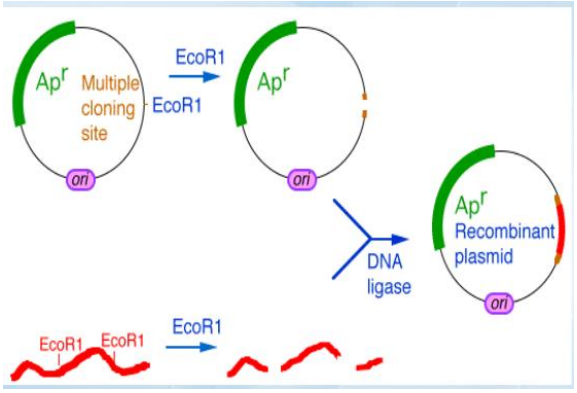
- **Origin of replication**

Origin of replication is a DNA segment recognized by the cellular DNA-replication enzymes.

Without replication origin, DNA cannot be replicated in the cell.

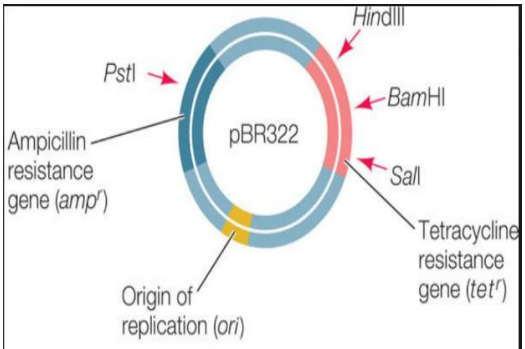


- **Multiple cloning site:**
- Many cloning vectors contain a multiple cloning site or polylinker: a DNA segment with several unique sites for restriction endonucleases located next to each other
- Restriction sites of the polylinker are not present anywhere else in the plasmid.
- Cutting plasmids with one of the restriction enzymes that recognize a site in the Polylinker does not disrupt any of the essential features of the vector
- Gene to be cloned can be introduced into the cloning vector at one of the restriction sites present in the polylinker



- **Selection marker:**

Usually an antibiotic marker is part of the cloning vector which is required for the maintenance of the plasmid in the cell. Under the selective conditions, only cells that contain plasmids with the appropriate selectable marker can survive. Commonly, genes that confer resistance to various antibiotics are used as selective markers in cloning vectors.



Here is a list of common antibiotic markers used in prokaryotes with their mechanism of action:

Antibiotic	Type of Cell	Mechanism of Action	Mechanism of Resistant Gene
Ampicilin	Prokaryote	Binds to and inhibits a number of enzymes in the bacterial membrane that are involved in the synthesis of the cell wall.	The amp^r gene on the plasmid codes for an enzyme that is secreted into the periplasmic space of the bacterium, where it catalyzes hydrolysis of the B-lactam ring of ampicillin, with concomitant detoxification of the drug.
Chloramphenicol	Prokaryote	Binds to the ribosomal 50S subunit and inhibits bacterial protein synthesis.	The cam^r gene encodes for a tetrameric, cytosolic protein (MW of each subunit 23,000) that in the presence of acetyl coenzyme A, catalyzes the formation of hydroxyl acetoxy derivatives of chloraphenicol that are unable to bind to the ribosomes.
Kanamycin	Prokaryote	Binds to ribosomal 70S subunit and inhibits bacterial protein synthesis.	
Tetracyclin	Prokaryote	Binds to a protein of the 30S subunit of the ribosome and inhibits ribosomal translocation	The tet^r gene encodes a membrane associated protein that prevents the antibiotics from entering the cell.
Blasticidin S	Prokaryote & Eukaryote	Inhibits protein synthesis.	The blasticidin^r gene encodes for a blasticidin deaminase, which converts blasticidin S to a non toxic deaminohydroxy.

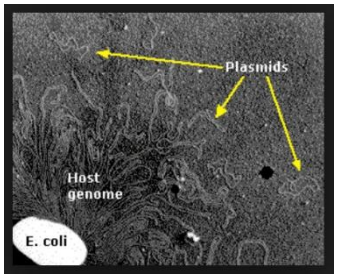
3.2 Different types of cloning vectors:

Different types of cloning vectors are used for different types of cloning experiments.
•The vector is chosen according to the size and type of DNA to be cloned

Types of vectors	Maximum insert size(kb-1000bp)
•Bacterial plasmid	6-12
•bacteriophage	25
•Cosmids	25

Bacterial Plasmid:

Most bacterial DNA is on a single large chromosome, but some DNA is in a small circle called a plasmid.

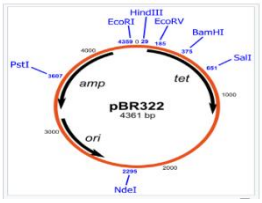


Plasmids occur naturally in bacteria and usually carry genes that are useful but not essential to survival

There can be as many as several hundred copies of a single plasmid in each bacteria.

They can replicate independently of the host cell.

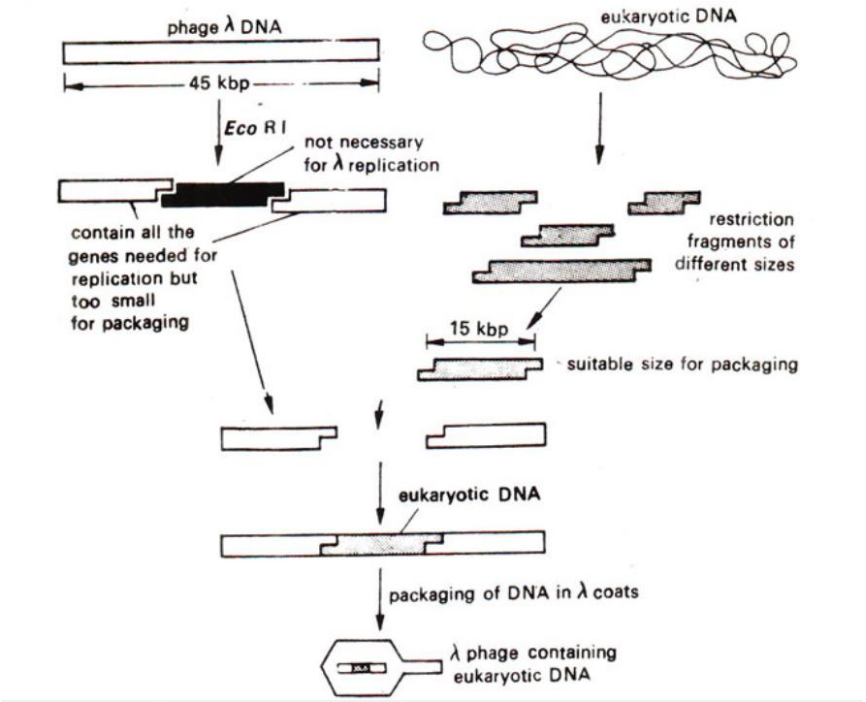
Examples of plasmid vectors are pBR322, pUC 8 ,pUC 18, pUC19



Schematic diagram of pBR322 series plasmid

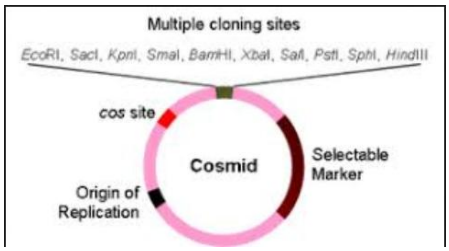
Bacteriophage:

Bacteriophages provide another source of cloning vectors. Since usually, a phage has a linear DNA molecule, a single break will generate two fragments, which are later joined together with foreign DNA to generate a chimeric phage particle . The chimeric phage can be isolated after a lytic cycle. The use of phage particles as vector imposes a limitation on the size of foreign DNA. which can be cloned, because the capacity of phage head is only limited, and if the size of foreign DNA is too long, size of phage DNA may not be accommodated in phage head. In order to overcome this problem, those segments of DNA, which do not contain essential genes may be removed. Such a technique has been followed in phage lambda (λ)to create a smaller vector genome having single restriction site for the enzyme *EcoRl.* Since the reduced size also fails to be adequately packed in phage head (there is also a requirement of a minimum size of DNA), this automatically provides a selection method, in which only the chimeric particles will be obtained in the phage progeny, and vector particles lacking cloned segment will be eliminated due to its reduced size. Some of the Lamda phage and M13 vectors are λgt10, λgt11, EMBL13, EMBL14 and M13 .



Cosmids:

Cosmids are plasmid particles, into which certain specific DNA sequences, namely those for *cos* sites are inserted. Since these *cos* sites enable the DNA to get packed in lambda particle, cosmids allow the packaging of DNA in phage particle *in vitro,* thus permitting their purification. Like plasmids, these cosmids perpetuate in bacteria and do not carry the genes for lytic development.



Cosmids are medium-sized cloning vectors. The cloning capacity of these vectors is 35–45 kbp. The first cosmid vector was described by Collins in 1978. Cosmid vectors are developed by combining the features of the **plasmid** vector and the bacteriophage vector. Origin of replication, *multiple cloning* site and selectable marker are obtained from the plasmid and only the cohesive site or **cos site** region is taken from lambda phage. These are fused together to obtain the cosmid vector. Approximately 200 bp lambda phage sequence is cloned into the cosmid vector. This consists of cosN, cosB and cosQ. A cosmid vector may have one or two cos sites. Cosmid vectors are used in the construction of genomic libraries. The cloning of a foreign DNA in cosmid vector involves the following steps: (1) ligation of the foreign DNA between two cos sites; (2) making a concatemeric DNA; (3) ***in vitro*** **packaging** to introduce the DNA into the phage head to form the matured phage particle; and (4) introduction of the cloned DNA into *E. coli* by transduction. After their entry into the host cell, the cosmids are maintained as plasmids. Usually cosmids are maintained as high copy plasmids.

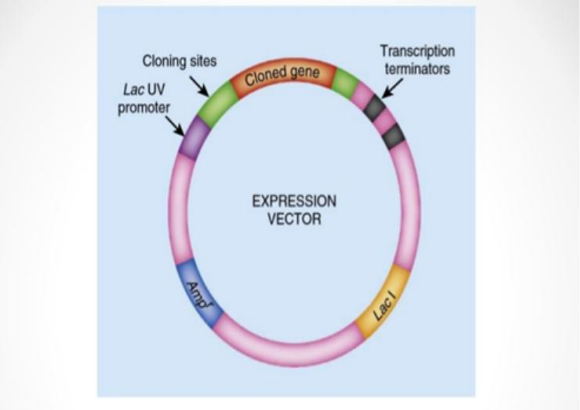
4. Expression Vector

4.1 PROPERTIES OF AN EXPRESSION VECTOR :

Expression vectors are the basic tools in biotechnology for the production of proteins. The expression vectors are vectors which act as vehicles for DNA insert and also allow the DNA insert to be expressed efficiently. These may be plasmids or viruses. The expression vectors are also known as expression constructs.

The simplest expression constructs are also known as transcription vectors; only because they allow transcription of the cloned foreign gene and not its translation. The vectors which facilitate both transcription and translation of the cloned foreign gene are known as protein expression vectors. These protein expression constructs also lead to the production of recombinant protein. The vector is engineered to contain regulatory sequences that act as enhancer and promoter regions and lead to efficient transcription of the gene carried on the expression vector. A common promoter utilized in the expression constructs is the mutant version of the lac promoter, lacUV. The lacUV promoter initiates a high level of transcription under induced conditions. The goal of a well-designed expression vector is the efficient production of protein, and this may be achieved by the production of significant amount of stable messenger RNA, which can then be translated into protein. The expression of a protein may be tightly controlled, and the protein is only produced in significant quantity when

necessary through the use of an inducer, in some systems however the protein may be expressed constitutively.



In order to express a protein an expression vector must have the correct translation initiation sequence such as a ribosomal binding site and start codon, a termination codon, and a transcription termination sequence. There are differences in the machinery for protein synthesis between prokaryotes and eukaryotes, therefore the expression vectors must have the elements for expression that is appropriate for the chosen host. For example, prokaryotes expression vectors would have a Shine-Dalgarno sequence at its translation initiation site for the binding of ribosomes, while eukaryotes expression vectors would contain the Kozak consensus sequence.

The Kozak consensus sequence, Kozak consensus or Kozak sequence is a sequence which occurs on eukaryotic mRNA and has the consensus (gcc)gccRccAUGG. The Kozak consensus sequence plays a major role in the initiation of the translation process. The sequence was named after the scientist who discovered it, Marilyn Kozak.

The sequence is identified by the notation (gcc)gccRccAUGG, which summarizes data analysed by Kozak from a wide variety of sources as follows:

1. a lower-case letter denotes the most common base at a position where the base can nevertheless vary;
2. upper-case letters indicate highly conserved bases, *i.e.* the 'AUGG' sequence is constant or rarely, if ever, changes, with the exception being the IUPAC ambiguity code 'R' which indicates that a purine (adenine or guanine) is always observed at this position (with adenine being claimed by Kozak to be more frequent); and
3. the sequence in parentheses (gcc) is of uncertain significance.

The promoter initiates the transcription and is therefore the point of control for the expression of the cloned gene. The promoters used in expression vector are normally inducible, meaning that protein synthesis is only initiated when required by the introduction of an inducer such as IPTG. Gene expression however may also be constitutive (i.e. protein is constantly expressed) in some expression vectors. Low level of constitutive protein synthesis may occur even in expression vectors with tightly controlled promoters.

Example of a constitutive promoter: GAP promoter of the gene encoding glyceraldehyde-3-phosphate dehydrogenase.

Example of a Inducible promoter: The lac Z gene can be induced by IPTG, isopropylthiogalactosidase in *E.Coli*

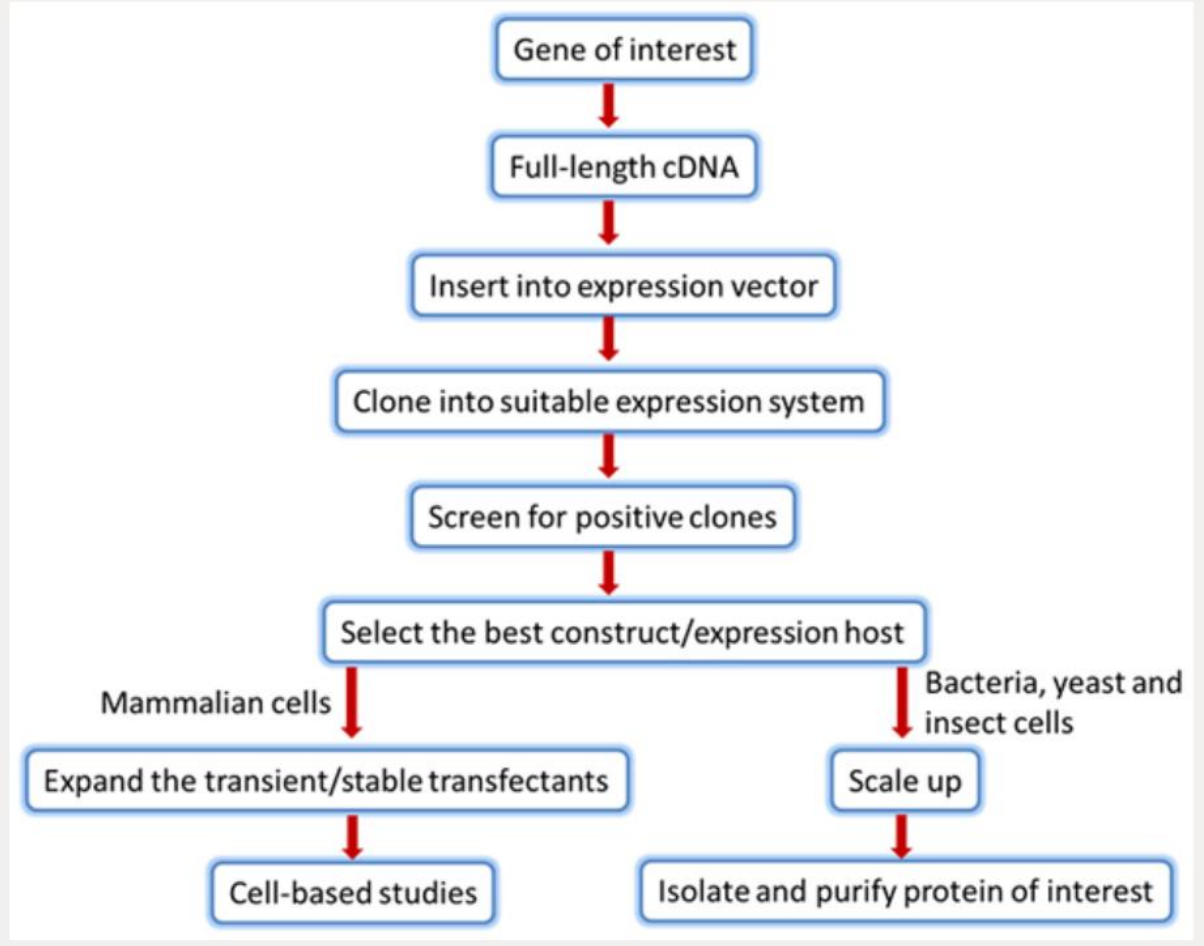
The **GAL1** and **GAL10** promoters are other examples. They are induced by galactose and are suitable for protein expression in *Saccharomyces cerevisiae.*

4.2 How do Expression vectors work ?

- Once the expression construct is inside the host cell, the protein encoded by the gene of interest is produced by the transcription. Thereafter, it utilizes the translation machinery and ribosomal complexes of the host organism.
- Frequently, the plasmid is genetically engineered to harbor regulatory elements like enhancers and promoters. These regulator sequences aid in efficient transcription of the gene of interest.
- Expression vectors are extensively used as tools which help in the production of mRNAs and, in turn, stable proteins. They are of much interest in biotechnology and molecular biology for the production of proteins like insulin. Insulin is the chief ingredient in the treatment of the complex disease, Diabetes.
- When the protein product is expressed, it is to be then purified. The purification of a protein poses a challenge since the protein of interest, whose gene is carried on the expression vector, is to be purified independently of the proteins of the host organism. To make the process of purification simpler, the gene of interest carried on the expression vector should always have a 'tag'. This tag can be any marker peptide or histidine (His tag).
- Expression vectors are considerably exploited in techniques like site-directed mutagenesis. Cloning vectors introduce the gene of interest into a plasmid which in turn replicates in bacteria. These cloning vectors need not necessarily result in the expression of a protein.

5.Basic layout of protein expression using various systems:

- identification of the gene that encodes the protein of interest
- generation of cDNA from the respective mRNA
- selection of suitable expression vector to insert the gene sequence
- selection of suitable system that can express the vector
- appropriate screening
- scaling up methods



- purification

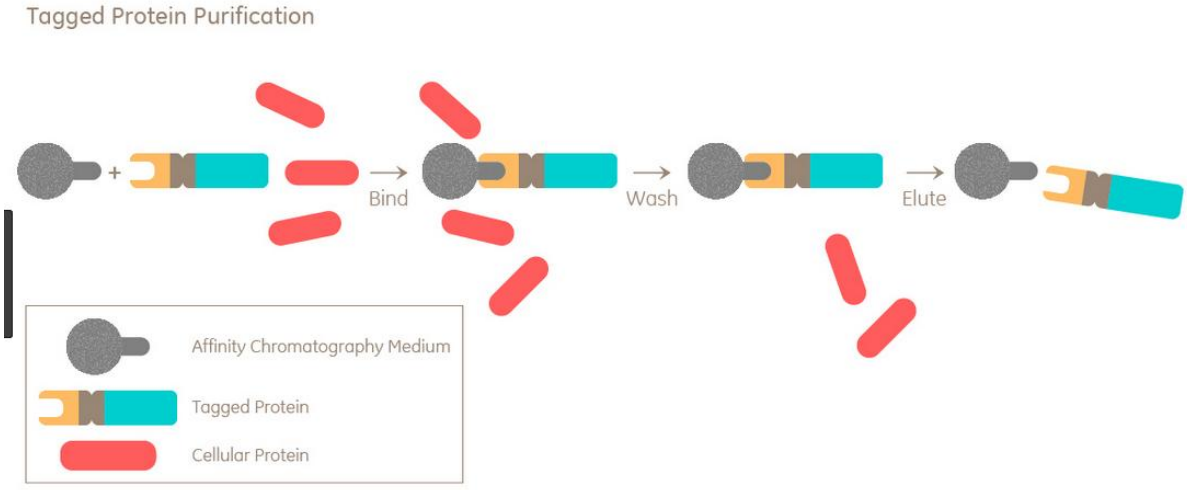
6.Basic steps of protein purification

Once the desired gene of interest is expressed in the appropriate model system, then the protein of interest needs to be purified. Purification of the protein can be done using the following techniques like

1. Ion exchange chromatography
2. Gel Filtration chromatography
3. Affinity chromatography(Immunoaffinity)
   These techniques have been discussed in detail in other modules. Kindly refer to the modules highlighting the respective chromatography techniques.

What are Protein tags and purification of tagged proteins

After the expression of the gene product, it is usually necessary to purify the expressed protein; however, separating the protein of interest from the great majority of proteins of the host cell can be a protracted process. To make this purification process easier, a purification tag may be added to the cloned gene. This tag could be histidine (His) tag, other marker peptides, or a fusion partners such as glutathione S-transferase or maltose-binding protein. Some of these fusion partners may also help to increase the solubility of some expressed proteins. Other fusion proteins such as green fluorescent protein may act as a reporter gene for the identification of successful cloned genes.

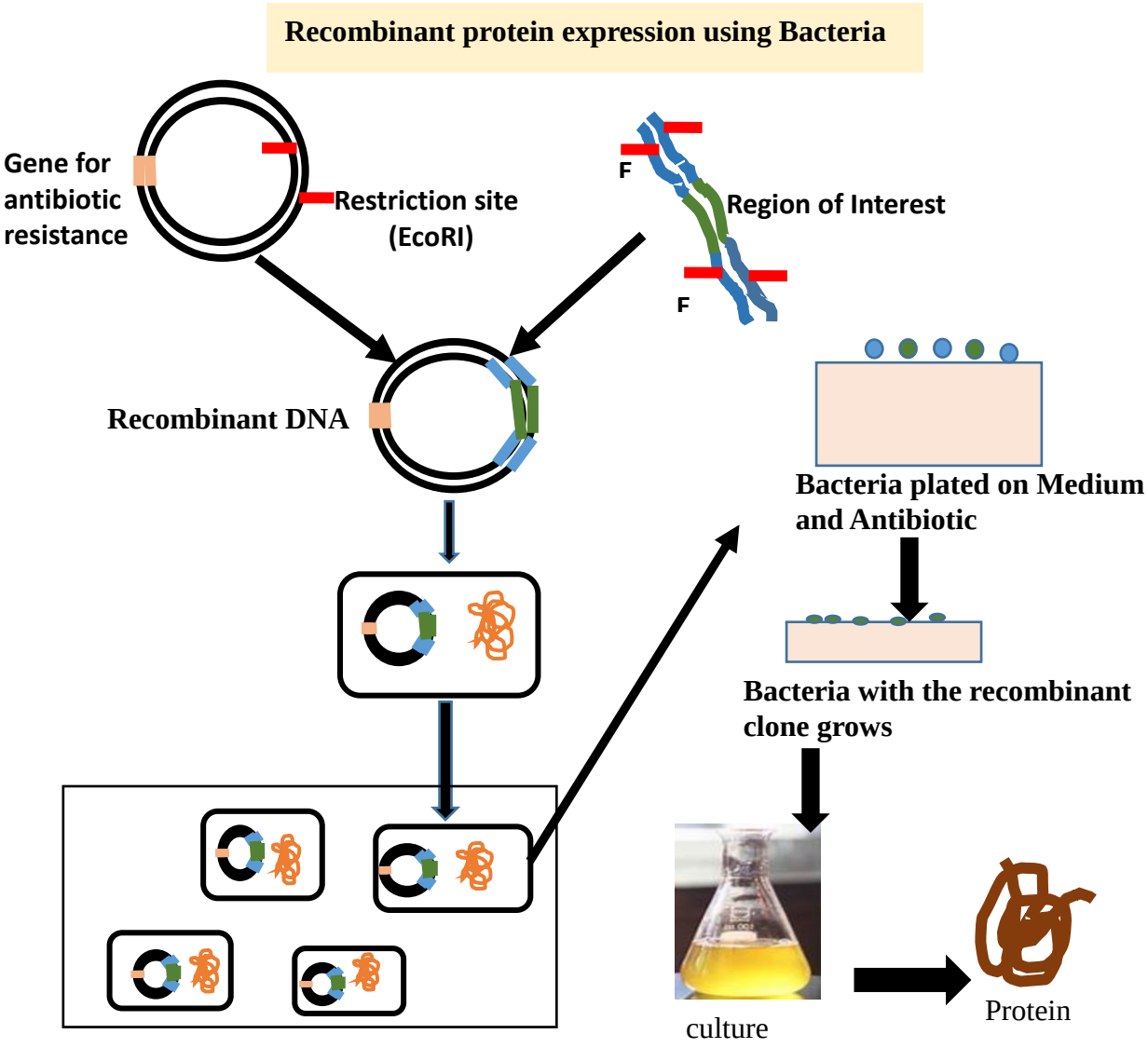


7. Expression of recombinant proteins in bacteria

Bacterial Protein expression system

Bacteria act as rapid and simple systems of expressing recombinant proteins due to the short doubling time. The media required to culture them are not expensive and the methods adapted to scale-up bioproduction are straightforward. The most widely used host system is *E. coli* since there is ample knowledge about its genetics, genome sequence and physiology. The genetic manipulation is easy and it also grows to high densities and is suitable for large-

scale fermentations. However, the cell wall of *E. coli* contains toxic pyrogens and the expressed proteins may have to be extensively tested before use. However, not all proteins can be successfully expressed in *E. coli*, or be expressed with the correct form of post-translational modifications such as glycosylations, and other systems may therefore be used.



8. Expression of recombinant proteins using Yeast :

Yeast systems have been a staple for producing large amounts of proteins for industrial and biopharmaceutical use for many years.

Yeast cell factories combine the advantages of being single cells, fast growing and with easy genetic manipulation, as well as harbours eukaryotic features including a secretory pathway leading to correct protein processing and post-translational modifications.

Engineering of yeast glycosylation to produce glycoproteins of human-like glycan structures is of great interest.

Yeast can be grown to very high cell mass densities in well-defined medium.

Recombinant proteins in yeast can be over-expressed so the product is secreted from the cell and available for recovery in the fermentation solution. Proteins secreted by yeasts are heavily glycosylated at consensus glycosylation sites.

Thus, expression of recombinant proteins in yeast systems historically has been confined to proteins where post-translations glycosylation patterns do not affect the function of proteins. Several yeast expression systems are used for recombinant protein expression, including *Sacharomyces, Scizosacchromyces pombe*, *Pichia pastoris* and *Hansanuela polymorpha*. With the advent of cheaper next-generation sequencing techniques, systems biotechnology approaches focusing on genome scale analyses will advance and accelerate yeast cell factories and thus recombinant protein production processes in the near future

PROPERTIES OF YEAST EXPRESSION VECTORS:

As in any prokaryote, yeast expression systems also require the following key features:

- origin of replication or integration
- a strong promoter
- selection marker

Expression of recombinant proteins in *S. cerevisiae* can be done using three types of vectors: integration vectors (YIp), episomal plasmids (YEp), and centromeric plasmids (YCp).

YIp Vectors. The YIp integrative vectors are vectors that do not replicate autonomously, but integrate into the genome at low frequencies by homologous recombination. Integration of circular plasmid DNA by homologous recombination leads to a copy of the vector sequence flanked by two direct copies of the yeast sequence. Typically, YIp vectors integrate as a single copy. However, methods to integrate multiple copies and stable cell lines with up to 15-20 copies of recombinant gene integrations have been developed for over-expressing specific genes.YIp plasmids with two yeast segments, such as YFG1 and the URA3 marker, have the potential to integrate at either of the genomic loci, whereas vectors containing repetitive DNA sequences, such as Ty elements or rDNA, can integrate at any of the multiple sites within genome.

YEp Vectors. The YEp yeast episomal plasmid vectors replicate autonomously because of the presence of a segment of the yeast 2 μm plasmid that serves as an origin of replication (2 μm ori). The 2 μm ori is responsible for the high copy-number and high frequency of transformation of YEp vectors. Most YEp plasmids are relatively unstable and even under conditions of selective growth, only 60 to 95 percent of the cells retain the YEp plasmid. The copy number of most YEp plasmids ranges from 10 to 40 copies per cell. Although this

system is used for small scale expression studies, the use of YEp vectors in large-scale manufacturing is not advisable.

YCp Vectors. YCp yeast centromere plasmid vectors are autonomously replicating vectors containing centromere sequences (CEN), and autonomously replicating sequences (ARS). The YCp vectors are typically present at very low copy numbers from 1 to 3 per cell. These vectors are also relatively unstable and not very useful in high level expression but are used as regular cloning vectors (e.g., pYC2, pBM272).

Yeast selection markers can be classified into two types: complementation markers and dominant selection markers. Dominant selection markers are antibiotic markers that can be used in yeast such as G418 and cyclohexamide. Complementation markers are marker genes that complement an auxotrophic mutation in the genome like URA3, TRPI, HIS3, and LEU2. These auxotrophic markers are used in selection of recombinants with all three types of expression systems (integration, episomal, and centromeric plasmids). A summary of the various plasmids and their selection systems is shown in the table

Expression of Recombinant Proteins in Yeast

Plasmid	Selection marker	Type
YRp7	TRP1	no
pSZ62	HIS3	Integration
YEp13	LEU2	2 µm
YEp24	URA3	2 µm
YCp50	URA3	CEN4
pRS303	HIS3	Integration
pRS304	TRP1	Integration
pRS314	TRP1	CEN6
pRS424	TRP1	2 µm
pRS425	LEU2	2 µm

Promoter

A range of yeast promoters is available for protein expression. Some like ADH2,SUC2 are inducible and others like GAPDH are constitutive in expression. Similar to the selection markers, a wide variety of combinations of promoters, markers, and expression systems are commercially available

Secretion Signals

A variety of secretion signals are also available for expression in *S. cerevisiae*. These include:

- Prepro alpha factor
- HSp150
- PHO1
- SUC2
- KILM1 (killer toxin type 1)
- GGP1

These systems and combinations of promoter, vectors, and signal sequence can be used for high-level expression of recombinant proteins in yeast.

Transformation methods in yeast :

Yeast Transformation

The introduction of DNA into yeasts:

- **Electroporation:**

Electroporation leads to generation of electric field delivered as an exponentially decaying pulse induces transient pore formations on the cell membrane through which the DNA molecules are able to pass into the cells. The transformation efficiency is affected by the electric field strength, yeast strain, cell number and condition, electroporation volume and buffer. To increase the efficiency of electroporation, many protocols suggest to perform a pretreatment with Lithium acetate or Dithiothreitol.

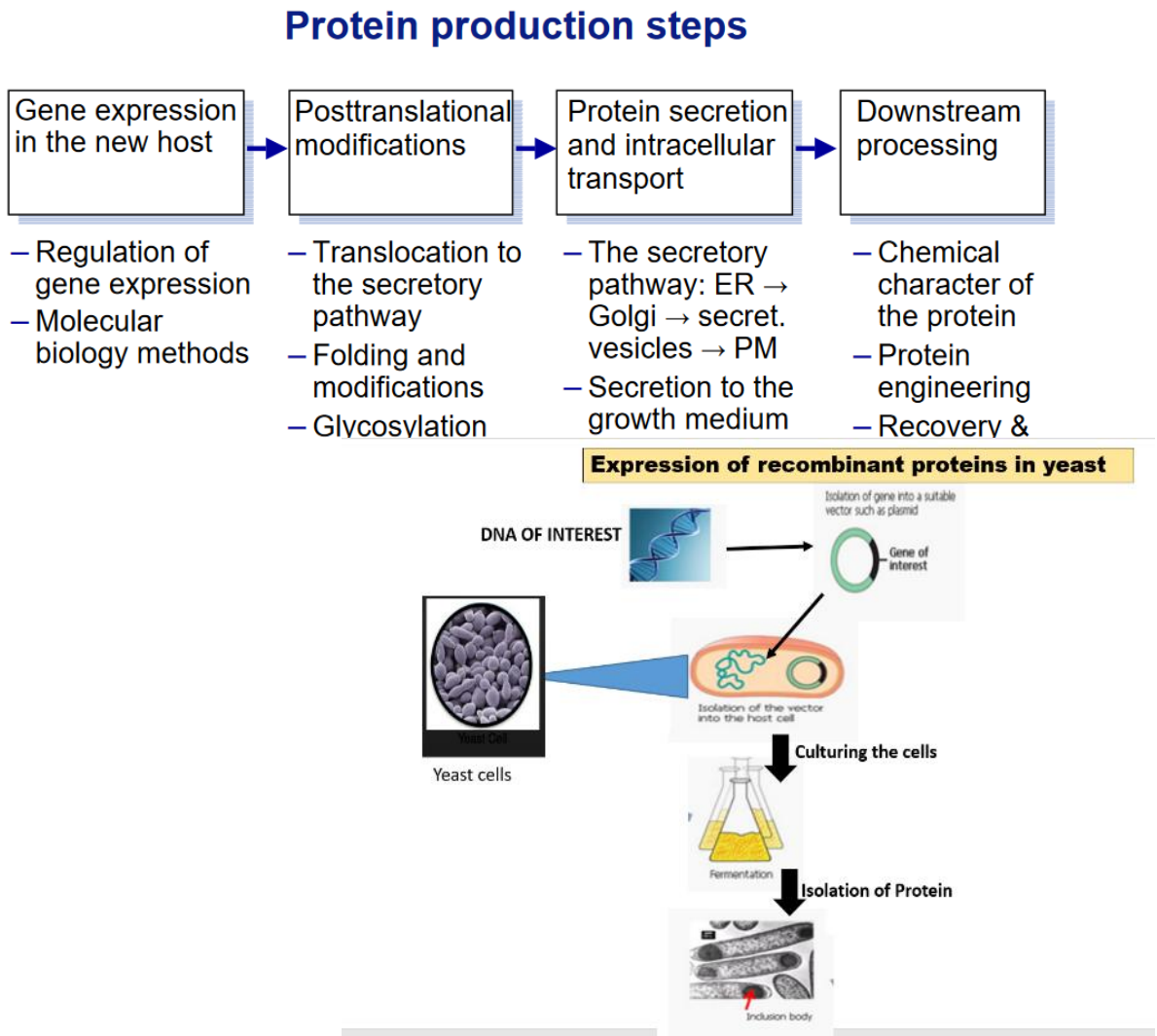
- **Lithium acetate treatment:**

In this method, yeast cells are incubated with a transformation mixture of lithium acetate, PEG 3500, single stranded carrier DNA and foreign plasmid at 42^0C for 40mins.

- **Spheroplast transformation:**

In the spheroplast method, yeast cells are incubated with zymolyase to partially remove cell wall to produce spheroplast.

Spheroplasts are very fragile for osmotic shock but are competent to takes up free DNA at high rate. In addition, polyethyl glycol (PEG) is used to facilitate deposition of plasmid and carrier DNA on cell wall for easier uptake.



Other Yeast Systems

A few other fungal systems have also been successfully used for recombinant protein production. Some of these systems are based on non-fermentation yeast species like *Pichia methanolica* and *Hansenula polymorpha*. These are similar to *Pichia pastoris* and can be grown to very high cell densities and high-level protein expression has been achieved. There is also a wide range of expression systems similar to the *Pichia* and *Saccharomyces* system available for recombinant protein expression.

SUMMARY:

1. Recombinant protein is a manipulated form of protein, which is generated in various ways

to produce large quantities of proteins, modify gene sequences and manufacture useful commercial products.

2. The recombinant DNA, usually the cDNA sequence of the target protein, is designed to be under the control of a

well-characterized promoter to express the target protein within the chosen host cell to achieve high-level protein expression.

3. Modification of the gene by recombinant DNA technology can lead to expression of a mutant protein. Proteins coexpressed in bacteria will not possess post-translational modifications, e.g. phosphorylation or glycosylation;eukaryotic expression systems are needed for this kind of modification of expressed proteins.

4. There are various protein expression system like bacteria, yeast, mammalian and insects.

5. The host selection for recombinant protein expression is dependent on the target protein, downstream use of the recombinant

protein and the properties of the host expression system.

6. Expression vector is a plasmid which is used to introduce a specific gene into a target cell and utilizes cells mechanisms

to produce the relevant gene. It contains enhancers, promoter region, termination codon, transcription initiation sequence

and translation initiation sequence in addition to the classical features of a cloning vector

7. Cloning vector is a small piece of DNA which can be stably maintained within a host cell. It contains origin of replication, multiple cloning sites and a selection marker.

8. Depending on the size and type of DNA to be cloned, cloning vectors are of three types in order of their increasing capacity to hold DNA like bacterial plasmid, bacteriophage and cosmid.

9. Expression vector must have the correct translation initiation sequence such as a ribosomal binding site and start codon, a termination codon, and a transcription termination sequence in order to produce a recombinant protein.

For example, prokaryotes expression vectors would have a Shine-Dalgarno sequence at its translation initiation site for the binding of ribosomes, while eukaryotes expression vectors would contain the Kozak consensus sequence.

10. When the protein product is expressed, it is to be then purified. The purification of a protein poses a challenge since the protein of interest, whose gene is carried on the expression vector, is to be purified independently of the proteins of the host organism. To make the process of purification simpler, the gene of interest carried on the expression vector should always have a 'tag'. This tag can be any marker peptide or histidine (His tag) or a fusion partners such as glutathione S-transferase or maltose-binding protein or a fluorescent tag like GFP.

11. Bacteria act as rapid and simple systems of expressing recombinant proteins due to the short doubling time, less expensive, easy manipulation , but not all proteins can be successfully expressed in *E. coli*, or be expressed with the correct form of post-translational modifications such as glycosylations, and other systems may therefore be used.

12. The cell wall of *E. coli* contains toxic pyrogens and the expressed proteins may have to be extensively tested before use.

13. Most common yeast expression systems used for recombinant protein expression, including *Sacharomyces*, *Scizosacchromyces pombe.*

14. As in any prokaryote, yeast expression systems also require the following key features like a origin of replication or integration, a strong promoter and selection marker.

15. Expression of recombinant proteins in *S. cerevisiae* can be done using three types of vectors: integration vectors **(YIp),** episomal plasmids **(YEp),** and centromeric plasmids **(YCp).**

16.Yeast expression vectors can have either complementation or dominant marker.

17.Yeast systems have been a staple for producing large amounts of proteins for industrial and biopharmaceutical use for many years.

18.Recombinant proteins in yeast can be over-expressed so the product is secreted from the cell and available for recovery in the fermentation solution.

19.Proteins secreted by yeasts are heavily glycosylated at consensus glycosylation sites. Thus, expression of recombinant proteins in yeast systems historically has been confined to proteins where post-translations glycosylation patterns do not affect the function of proteins.

THANKS

End of module