

Expression Vector

EXPRESSION VECTOR

- A vector used for **expression of a cloned DNA fragment** in a host cell is called as an expression vector.
- These vectors are **frequently engineered** to contain regulatory sequences that act as **promoter and/or enhancer regions** and lead to efficient transcription of the insert gene.
- Expression vectors are used for molecular biology techniques such as **site-directed mutagenesis**.
- The goal of a well-designed expression vector is the production of large amounts of **stable messenger RNA**, and in **extension, proteins**.
- Expression vectors are **basic tools for biotechnology** and the production of proteins.

CHOICE OF EXPRESSION VECTORS

- Strong promoter
- Intact ORF
- Ribosomal binding site
- Termination sequence

EXPRESSION VECTOR COMPONENTS

Goal	Component
Insert cargo into the plasmid and verify the insert sequence accuracy	MCS – restriction sites OR recombination regions 5' and 3' Primer sites for sequence verification
Insert plasmid into cells, enable the plasmid to replicate inside the host, & select for cells carrying the plasmid	<ul style="list-style-type: none">• Backbone compatible with cloning method• Origin of replication• Selection marker and/or screening marker
Transcribe mRNA from the plasmid	<ul style="list-style-type: none">• Promoter (constitutive or inducible)operator, terminator
Translate mRNA into protein	Ribosome Binding Site, start codon, stop codon
Promote proper folding of nascent protein	<ul style="list-style-type: none">• co-expression of chaperones• Solubilization tags• custom-designed synthetic RBS• Codon-optimized ORF
Detect or Purify target protein	<ul style="list-style-type: none">• Epitope tags (His)• reporters (GFP)

EXPRESSION VECTOR TWO TYPE

1. PROKARYOTIC EXPRESSION VECTOR

- Bacterial expression system (e.g. E.coli)

2. EUKARYOTIC EXPRESSION VECTOR

- Yeast expression system (e.g. S.cervisiae)
- Viral expression system (e.g. Baculovirus)
- Mammalian cell expression system

DELIVERY METHODS

Host	Commonly used methods to introduce expression construct (plasmid)	
Prokaryotic cells	Transformation via	<ul style="list-style-type: none"> • CaCl₂ + heatshock • electroporation
Yeast cells	Transformation via	<ul style="list-style-type: none"> • LiAc/PEG/ssDNA • Electroporation • <i>spheroplasts, biolistics, glass beads</i>
Plant cells	Transformation via	<ul style="list-style-type: none"> • <i>Agrobacterium-mediated transformation</i> • Gene gun
Mammalian cell lines	Transfection via	<ul style="list-style-type: none"> • Liposomes • Electroporation • <i>calcium phosphate, nanoparticles</i>
Mammalian (primary) cells	Transduction via	lentivirus
<i>in vivo delivery into live animals</i>	Transduction via	<ul style="list-style-type: none"> • adenovirus • AAV

SELECTION / SCREENING MARKERS

Host	Selection Markers
Prokaryotic cells	Ampicillin, Kanamycin, tetracycline, chloramphenicol
Mammalian cells	Neomycin; Puromycin; Hygromycin; Zeocin
Yeast	HIS4, Zeocin Auxotrophy: URA3, TRP1, LEU2, HIS3

Host	Screening Markers
Prokaryotic cells	Lacz for Blue-white screening
Prokaryotic or Eukaryotic cells	Fluorescent Proteins (GFP, mCherry) Luciferase

TRANSCRIPTIONAL PROMOTERS

Host	Commonly Used Promoters
Bacteria	<ul style="list-style-type: none">•Lac•T7•araBAD
Yeast	<ul style="list-style-type: none">•GAL4•PGK•ADH1•ADE2•TRP1
Mammalian	<ul style="list-style-type: none">•Constitutive: CMV, SV40, EF1a, CAG•Inducible: Tet•<i>Tissue-specific for in vivo work: varied</i>

EPITOPE TAGS / FUSION PROTEINS

Goal	Commonly Use Tag
Detect target protein (common Ab epitope tags)	FLAG (DYKDDDDK) HA (YPYDVPDYA)
Purify target protein (affinity tags)	His6 glutathione-S-transferase (GST)
Improve solubility of target protein	maltose-binding protein (MBP) glutathione-S-transferase (GST)

Tags can be removed at cleavage sites place between target ORF and tag:

Enzyme	Cleavage site
Thrombin	Leu-Val-Pro-Arg-Gly-Ser
Enterokinase	Asp-Asp-Asp-Asp-Lys
Factor Xa	Ile-Glu/Asp-Gly-Arg
TEV	Glu-Asn-Leu-Tyr-Phe-Gln-Gly ⁹

Saccharomyces cerevisiae

- It is **the most common eukaryotic system** and there is a great deal of study about this organism
- It is a **single-celled** and behaves like a bacterial culture and can be grown in **relatively simple media** in both **small and large-scale production**
- Well characterized with **many strong regulatory promoters** with **naturally occurring plasmids**
- Carry out **post-translational modifications**
- Secretes **very few of its own proteins**
- **Recognized as safe** by USDA and FDA

Recombinant proteins produced by *S. cerevisiae* expression systems

Vaccines	Diagnostics	Human therapeutic agents
<ul style="list-style-type: none">•Hepatitis B virus surface antigen•Malaria circumsporozine protein•HIV-1 envelope protein	<ul style="list-style-type: none">•Hepatitis C virus protein•HIV-1 antigens	<ul style="list-style-type: none">•Epidermal growth factor•Insulin•Fibroblast growth factor•Hirudin•Human growth factor¹¹

Saccharomyces cerevisiae

- There are **three main classes** of *S. cerevisiae* expression vectors.
 - Yeast episomal plasmids (YEps)
 - Yeast integrating plasmids (YIps)
 - Yeast artificial chromosomes (YACs)

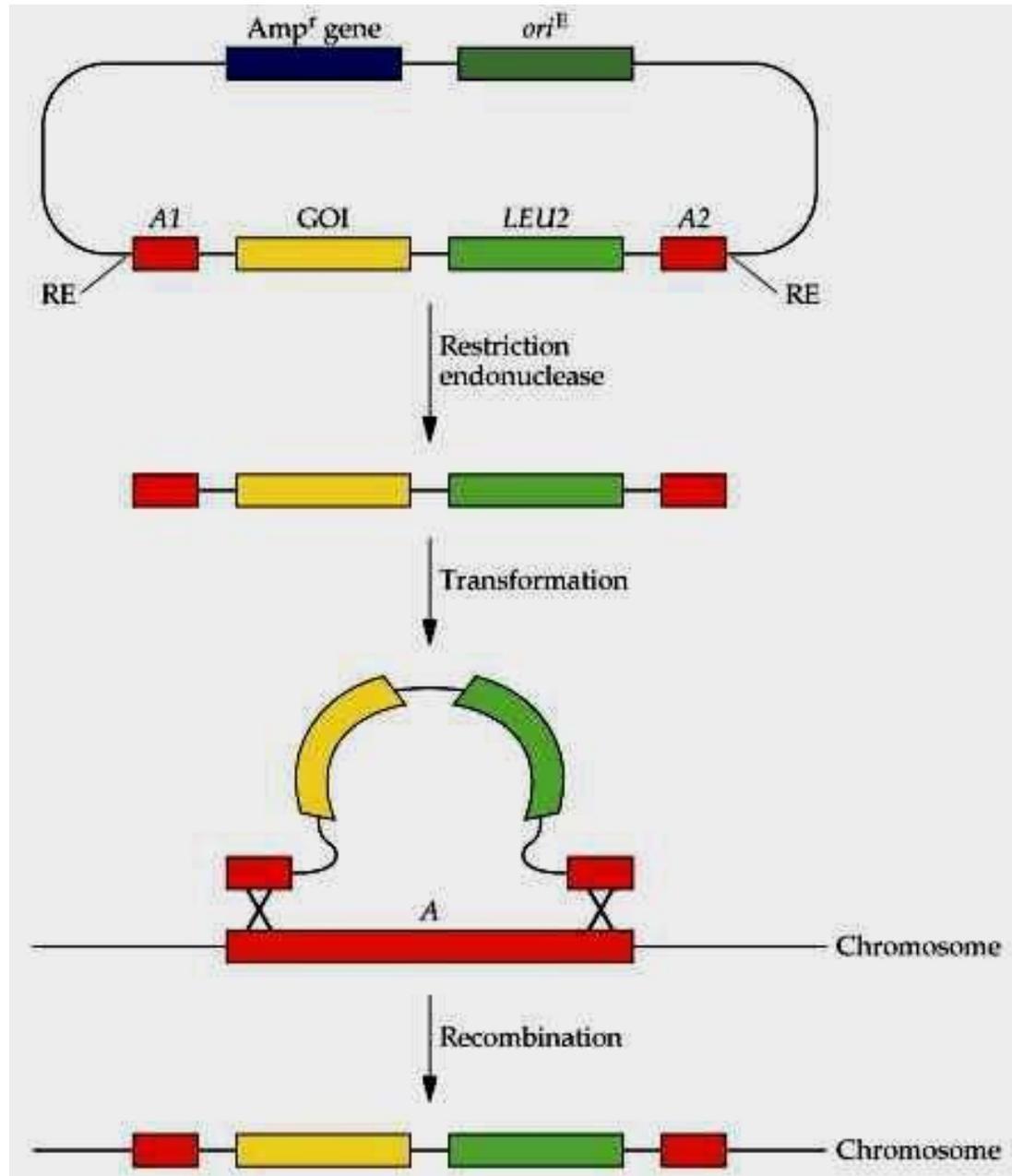
- **Yeast episomal plasmids** have been used extensively for the production of either **intra- or extracellular heterologous proteins**

- Typically, vectors **function in both** *E. coli* and *S.cerevisiae*.

Saccharomyces cerevisiae

- The YEps vectors are based on the **high-copy-number 2 μ m plasmids**
- The vectors replicate independently via a **single origin of replication**.
- There are more than **30 copies per cell**.
- **Selection scheme rely on mutant host strains** that require a particular amino acid (histidine, tryptophan, or leucine) or nucleotide (uracil).
- When a Yep with a wild-type *LEU2 gene* is transformed into a mutant *leu2 host cell*, *only cells* that carry plasmid will grow.
- A **Yip vector** is used to integrate a heterologous gene into the host genome to provide a **more reliable production system**.

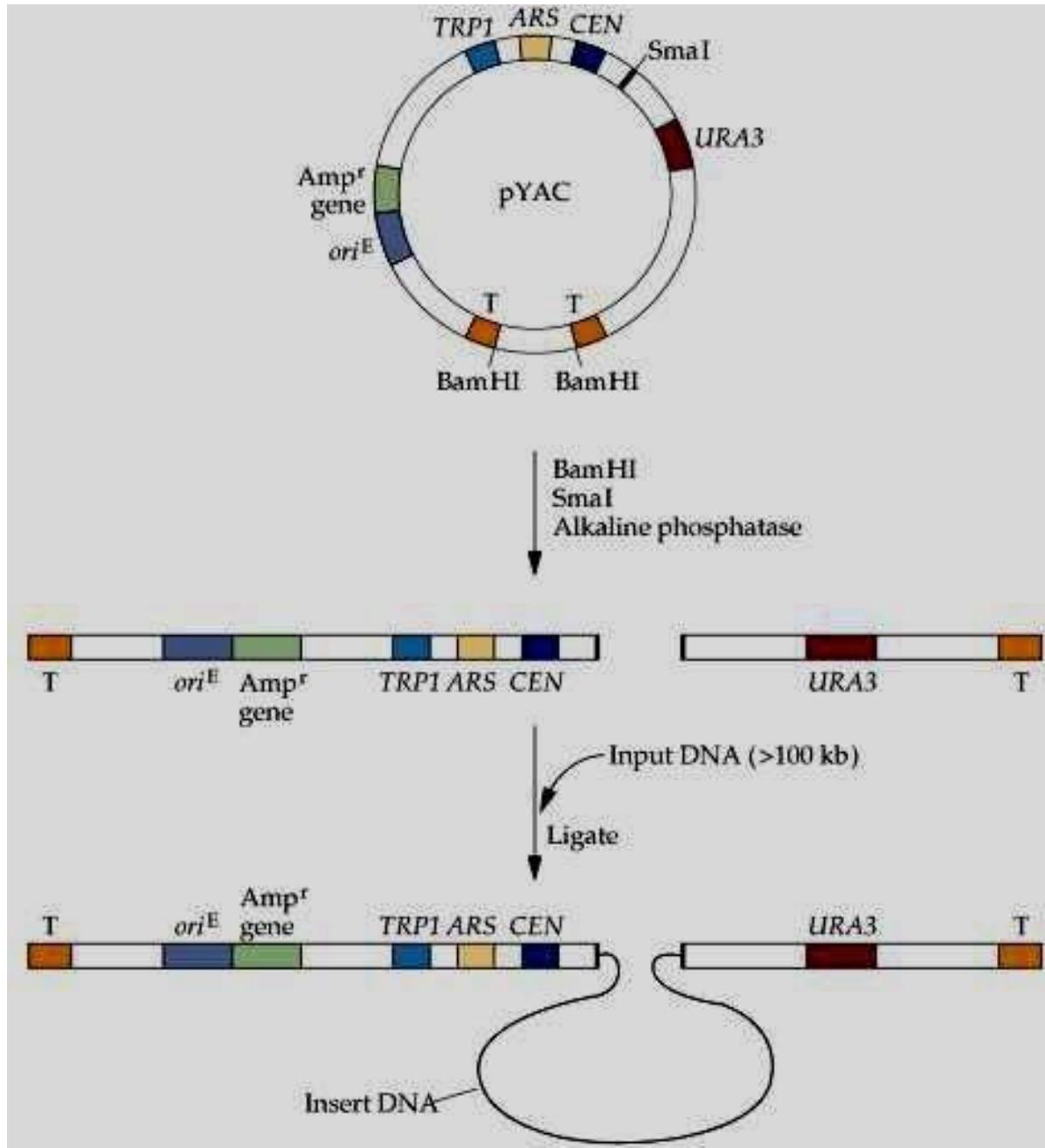
Integration of DNA with a Yip vector



YAC CLONING SYSTEM

- A YAC is designed to clone a **large segment of DNA (100kb)**, which is then maintained as a separate chromosome in the host yeast cell.
- It is **highly stable** and has been used for the physical mapping of **human genomic DNA**, the analysis of transcription units, and **genomic libraries**.
- It has a sequences that act as **ARS for replication**, **centromere** for cell division, and **telomere** for stability.
- To date, they have **not been used** as expression systems for the **commercial production**.

YAC CLONING SYSTEM



Pichia pastoris Expression Systems

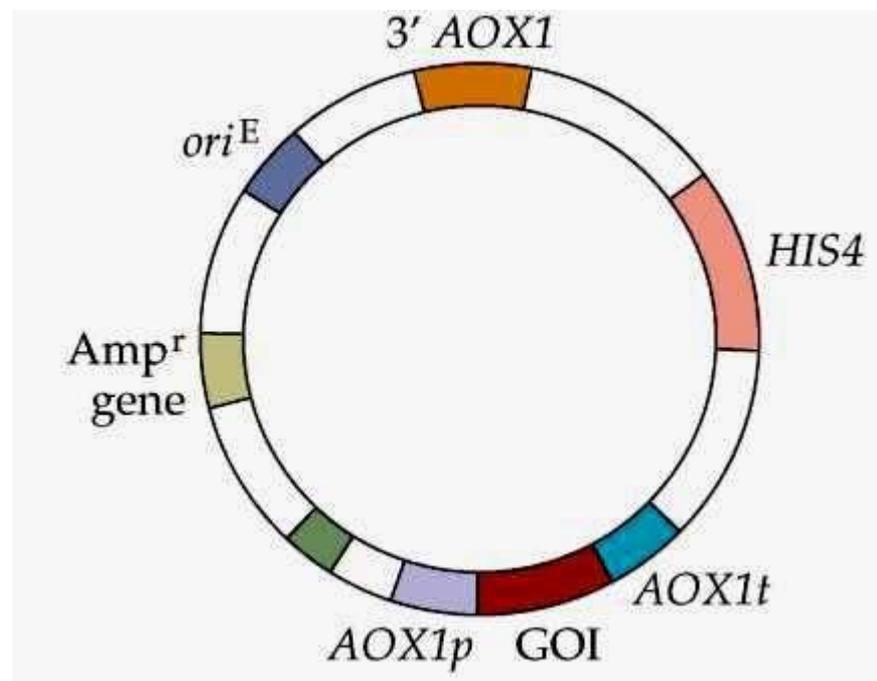
- Though *S. cerevisiae* is **successfully used** to produce recombinant proteins for human, it has major **drawbacks**.
- The level of **protein production is low**.
- There is the tendency for **hyperglycosylation** resulting in change of protein function.
- **Proteins** are often **retained in periplasm**, increasing time and **cost for purification**.
- It **produces ethanol** at high cell densities, which is toxic to cells.

Pichia pastoris Expression Systems

- *P. pastoris* is a **methylophilic yeast** that is able to utilize methanol as a source of carbon and energy.
- **Glycosylation** occurs to a **lesser extent** and the linkages between sugar residues are of the **α -1,2 type**.
- *P. pastoris* strain was extensively engineered with the aim of developing a “**humanized**” strain that glycosylate proteins in a manner identical to that of human cells.
- It does **not produce ethanol**.
- It normally secretes very few proteins, thus **simplifying the purification** of secreted recombinant proteins.

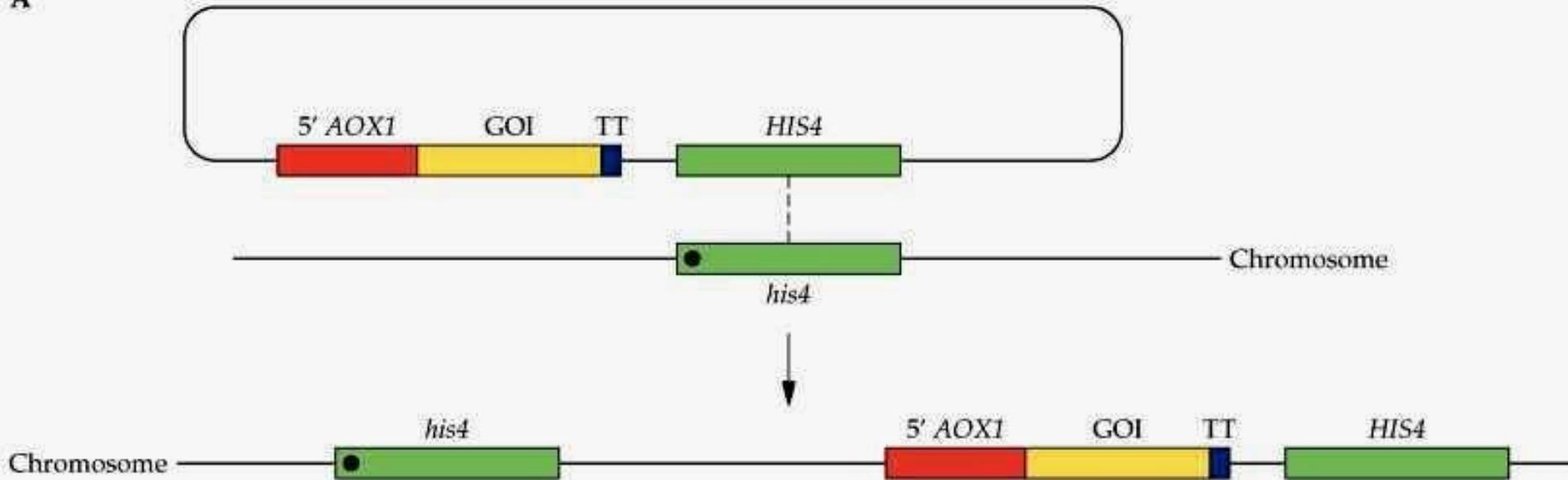
Pichia pastoris Expression Systems

➤ A **double recombination** event between the **AOX1p** and **AOX1** regions of the vector and the **homologous segments** of chromosome DNA results in the **insertion of the DNA** carrying the gene of interest and the **HIS4** gene.

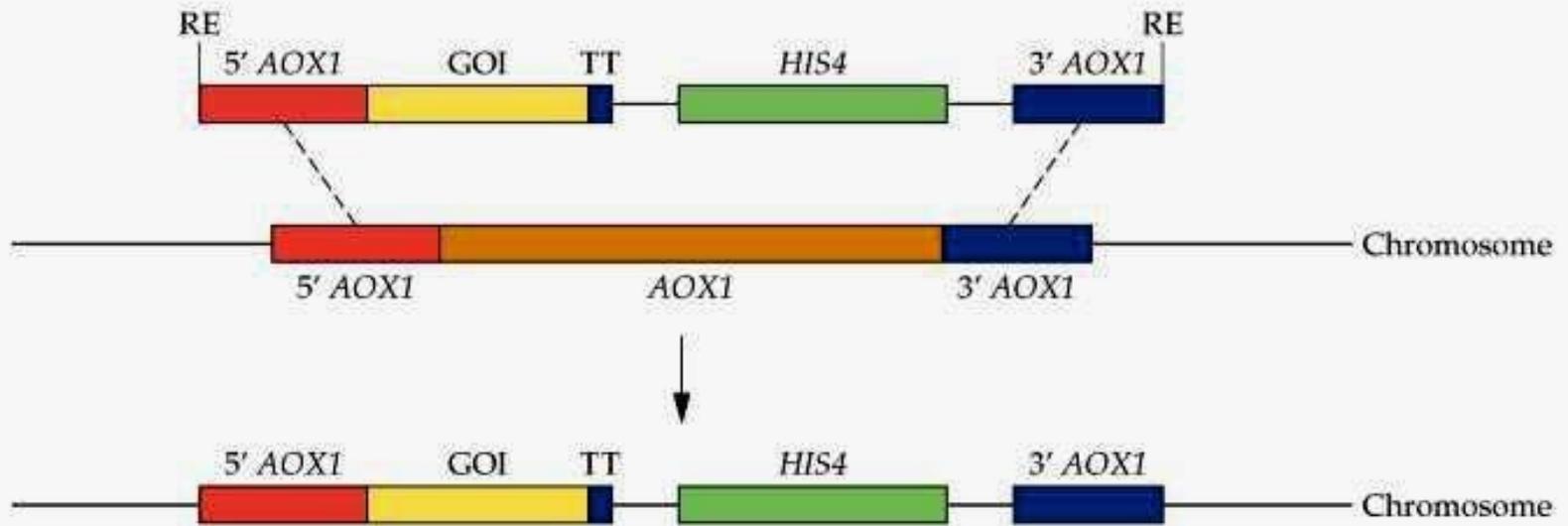


Pichia pastoris Expression Systems

A



B



Baculovirus-Insect Cell Expression

- Baculoviruses are a large, diverse group of viruses that specifically **infect arthropods**, and are not infectious to other animals.
- During the infection cycle, **two forms of baculovirus** are produced.
- A **single nucleocapsid (virus particle)** which can infect more midgut cells.
- **Clusters of nucleocapsids** that are produced outside of the cells (**virions**) in a protein matrix (polyhedrin).

Baculovirus-Insect Cell Expression

➤ The **polyhedrin gene** is replaced with a coding sequence for a heterologous protein, followed by infection of cultured insect cells, resulting in the production of the heterologous protein.

TABLE 7.6 Some of the recombinant proteins that have been produced by the baculovirus expression vector system

α -Interferon	G-protein-coupled receptors	Malaria proteins
Adenosine deaminase	HIV-1 envelope protein	Mouse monoclonal antibodies
Anthrax antigen	HSV capsid proteins	Multidrug transporter protein
β -Amyloid precursor protein	Human alkaline phosphatase	Poliovirus proteins
β -Interferon	Human DNA polymerase α	Pseudorabies virus glycoprotein 50
Bovine rhodopsin	Human pancreatic lipase	Rabies virus glycoprotein
Cystic fibrosis transmembrane conductance regulator	Influenza virus hemagglutinin	Respiratory syncytial virus antigen
Dengue virus type 1 antigen	Interleukin-2	Simian rotavirus capsid antigen
Erythropoietin	Lassa virus protein	Tissue plasminogen activator

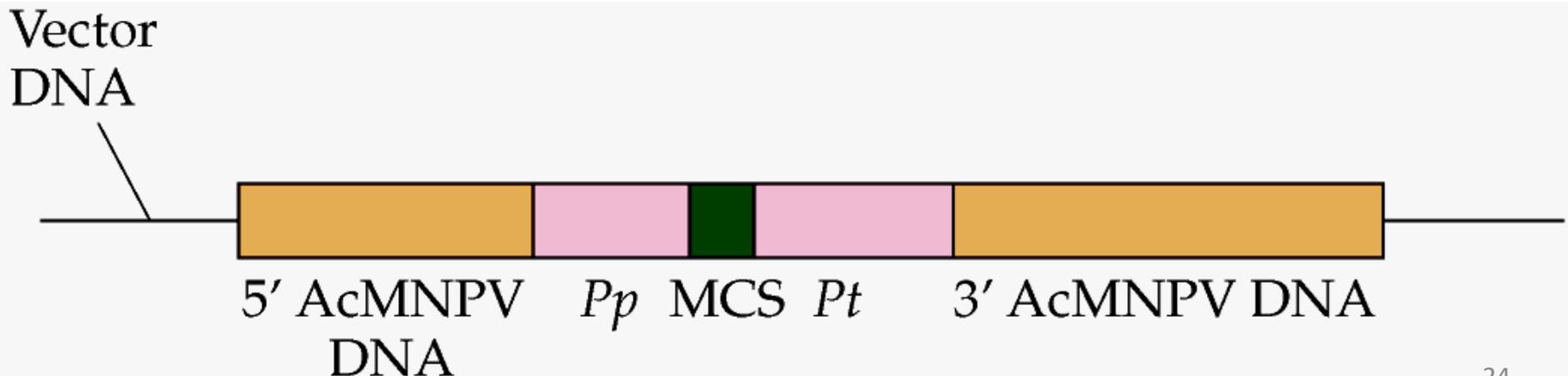
HIV-1, human immunodeficiency virus type 1; HSV, herpes simplex virus.

Baculovirus-Insect Cell Expression

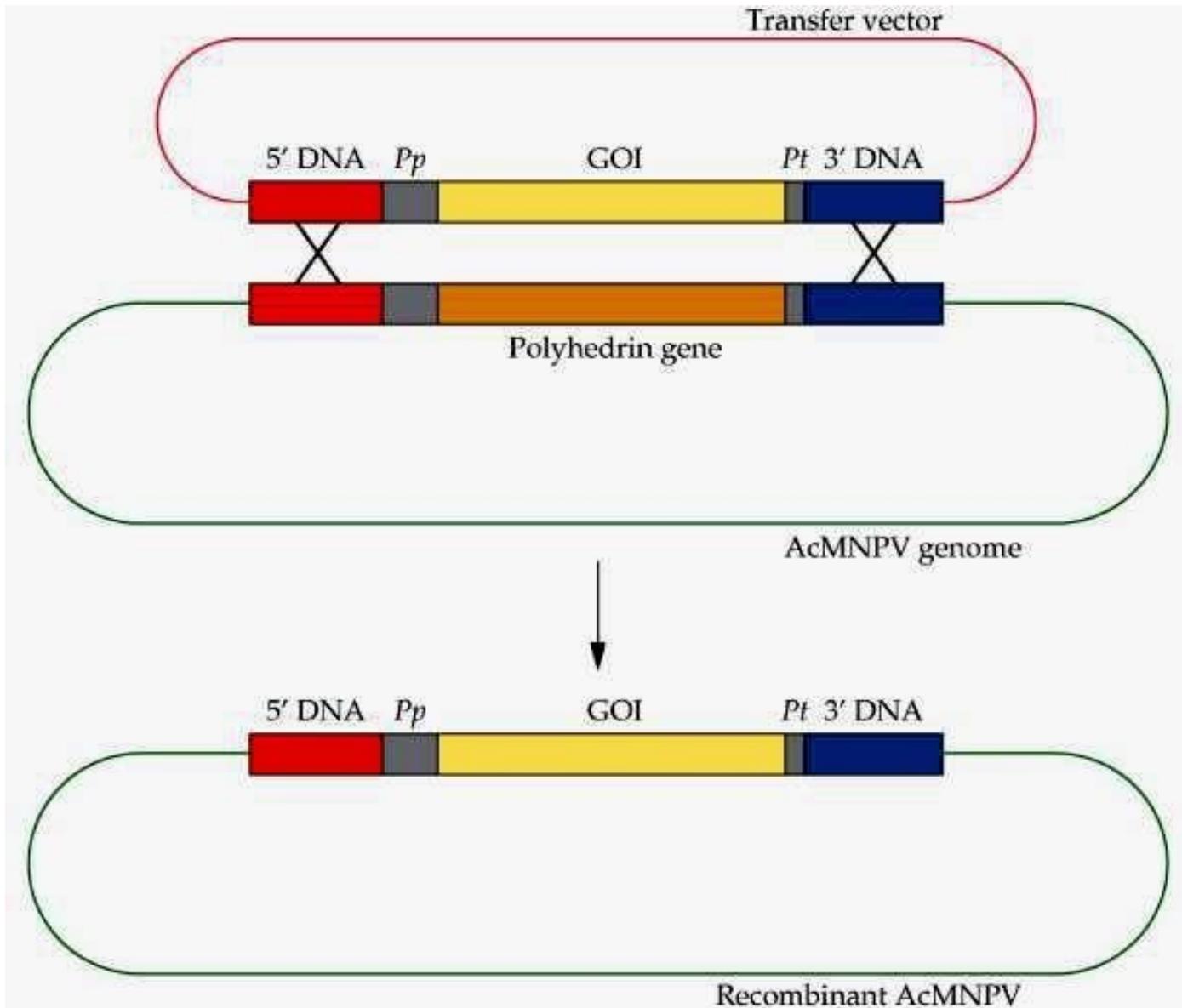
- Constructs have been made using the **polyhedrin promoter** to produce large quantities of extracellular protein.
- Most proteins are **modified and secreted** properly.
- Grows very well in many insect cell lines allowing **easy production**.
- Minor problem that (galactose and sialic acid; N-linked.) **doesn't process certain mammalian glycosylation types correctly**

Baculovirus-Insect Cell Expression

- The specific baculovirus that has been used extensively is *Autographa californica multiple nuclear polyhedrosis virus* (AcMNPV.)
- A gene of interest is inserted into the MCS and the transfer vector is propagated in *E. coli*.
- Next, insect cells in culture are cotransfected with AcMNPV DNA and the transfer vector carrying the cloned gene.



Baculovirus-Insect Cell Expression



Mammalian Cell Expression Systems

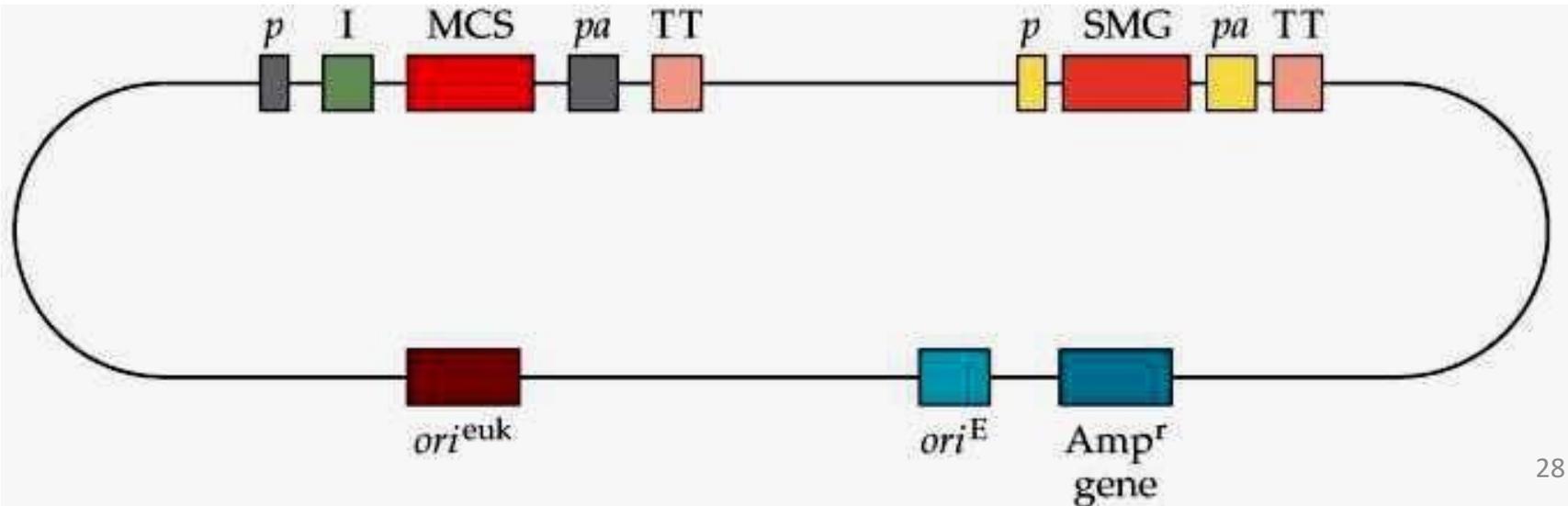
- Important for producing proteins with all **post translational modifications**.
- Many **established cell lines** are useful.
- **Transient expression**: African green monkey, baby hamster & human embryonic (all kidney tissue cell lines.)
- **Long-term expression**: Chinese hamster ovary and mouse myeloma cells.

Mammalian Cell Expression Systems

- Expression vectors in these systems are usually **derived from an animal virus** such as SV40 (simian virus 40).
- Can be used for expression of **single polypeptides, homooligomers, and heterooligomers**.
- The latter is made possible by transforming with **two or more separate cloned genes**.
- **Industrial production** is however **costly**.

VECTOR DESIGN

- Generalized mammalian expression vector.
- The **MCS** and **SMG** (Selectable Marker Gene) are under the control of **eukaryotic promoter (p)**, **polyadenylation(pa)**, and **terminal sequence(TT)**.
- An **intron(I)** enhances the production of **heterologous protein**.



➤ For the best results, a gene of interest must be equipped with **translation control sequences**.

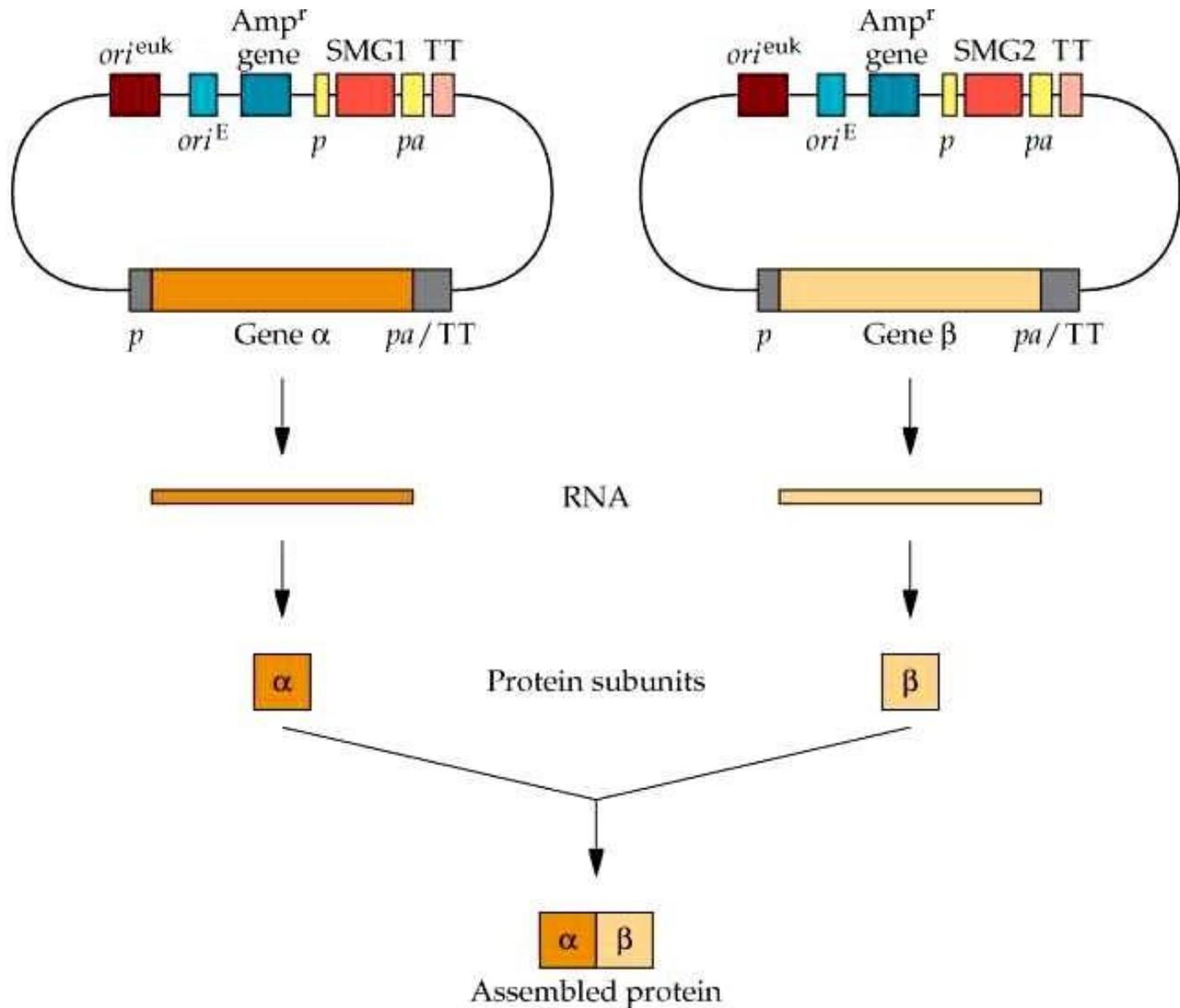
➤ A gene of interest can be fitted with various sequences that enhance translation and facilitate both secretion and purification.

➤ A **Kozak sequence(K)**, specific sequence surrounding the AUG start codon, **signal sequence(S)**, **protein affinity tag(T)** for purification, **proteolytic cleavage site(P)**, and **stop codon(SC)**.

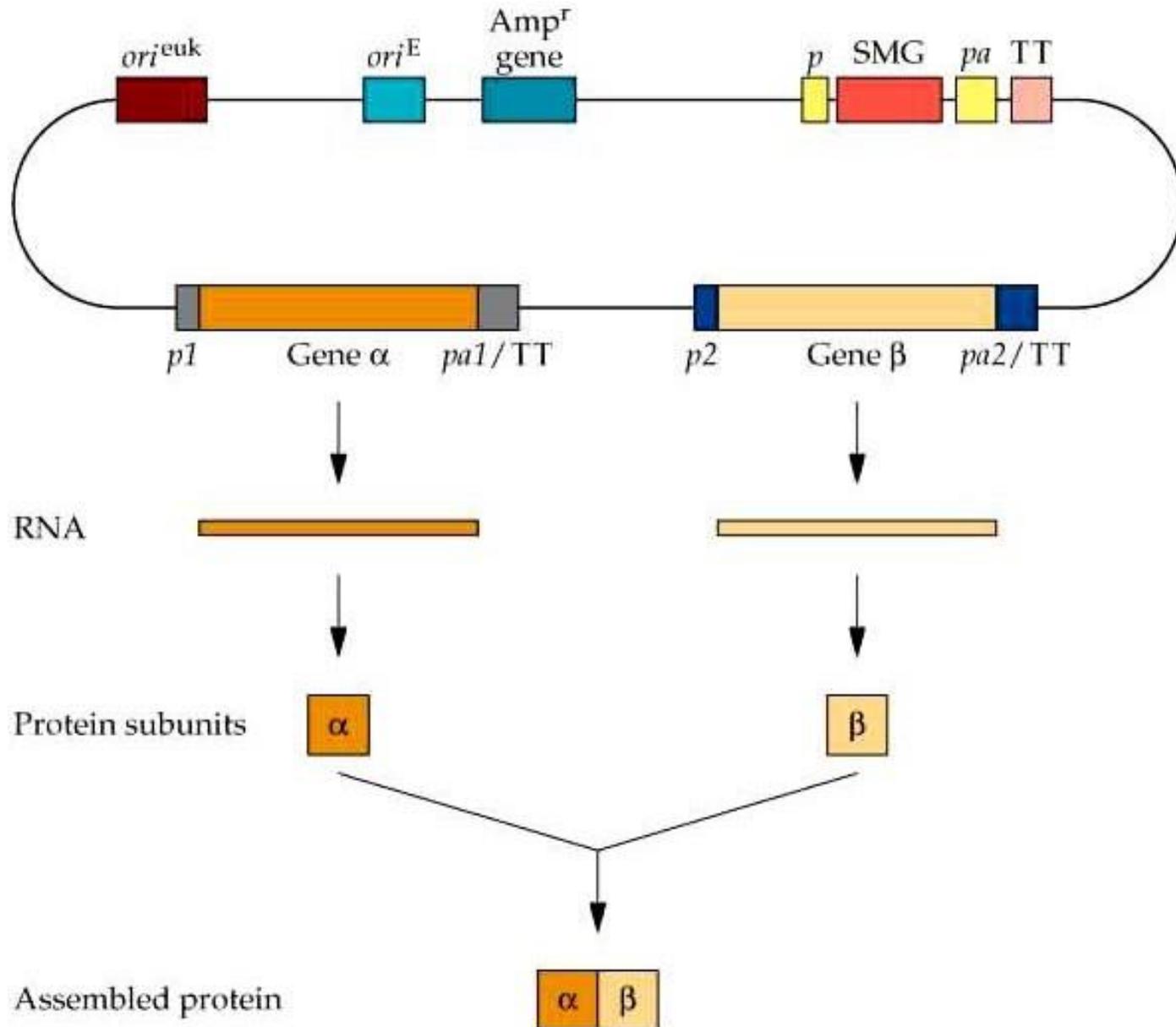
➤ The **5' and 3'UTR** increase the efficiency of translation and contribute to mRNA stability.



Two-Vector Expression System

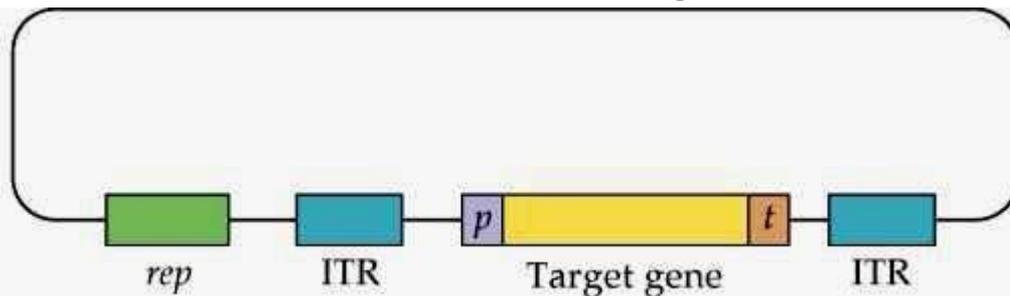


Two-Vector Expression System



Baculovirus Vector in Mammalian Cells

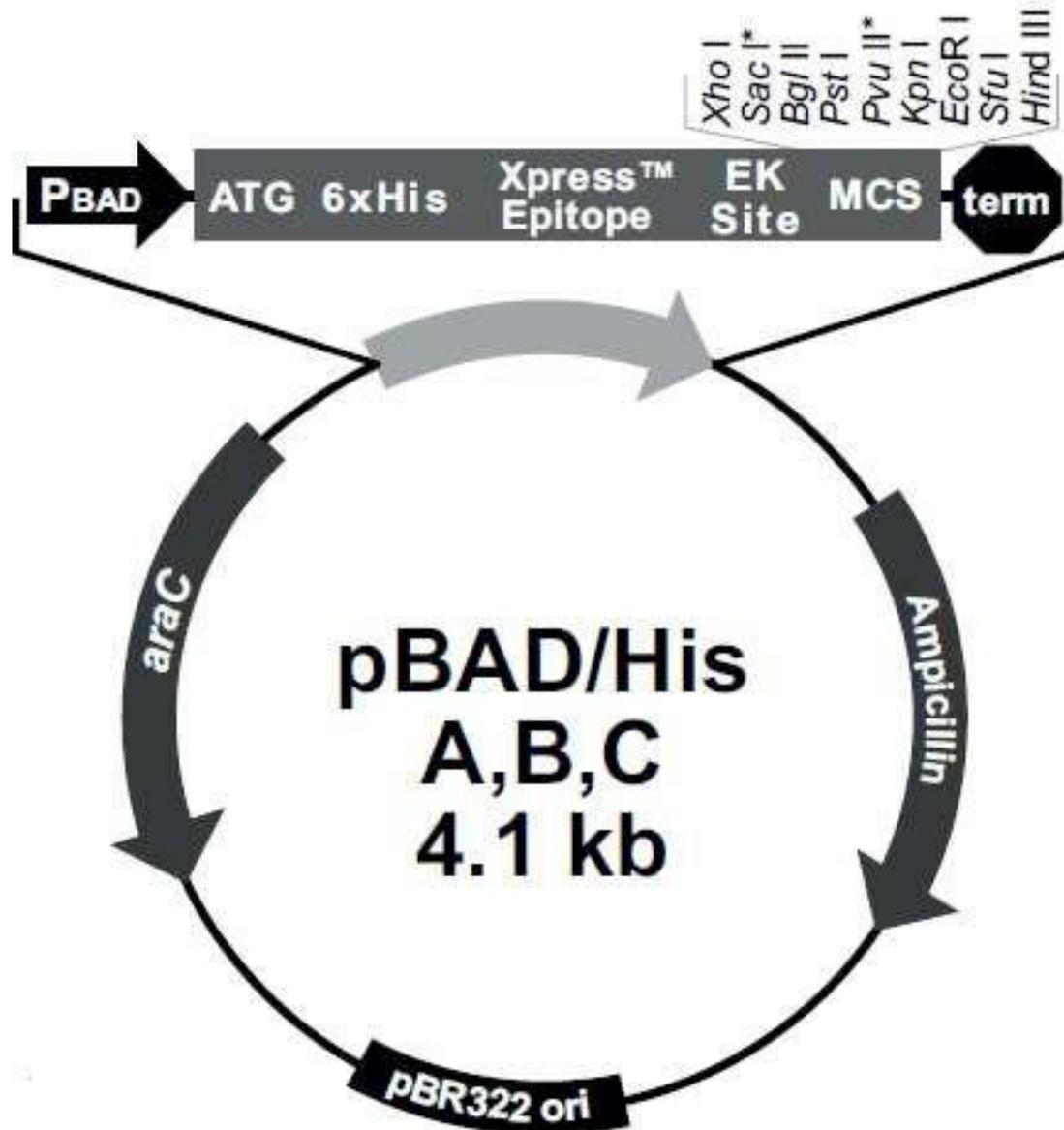
- It is possible to use some of the baculovirus vector to express target proteins in mammalian cells.
- Because **baculovirus cannot replicate in mammalian cells** and the **polyhedron-deficient strains** employed as vectors **cannot infect insects**.
- It is a **safe system**.
- For stable long-term expression, the target gene is inserted between sequences for **adeno-associated virus inverted terminal repeat** to facilitate the integration into the host cells.



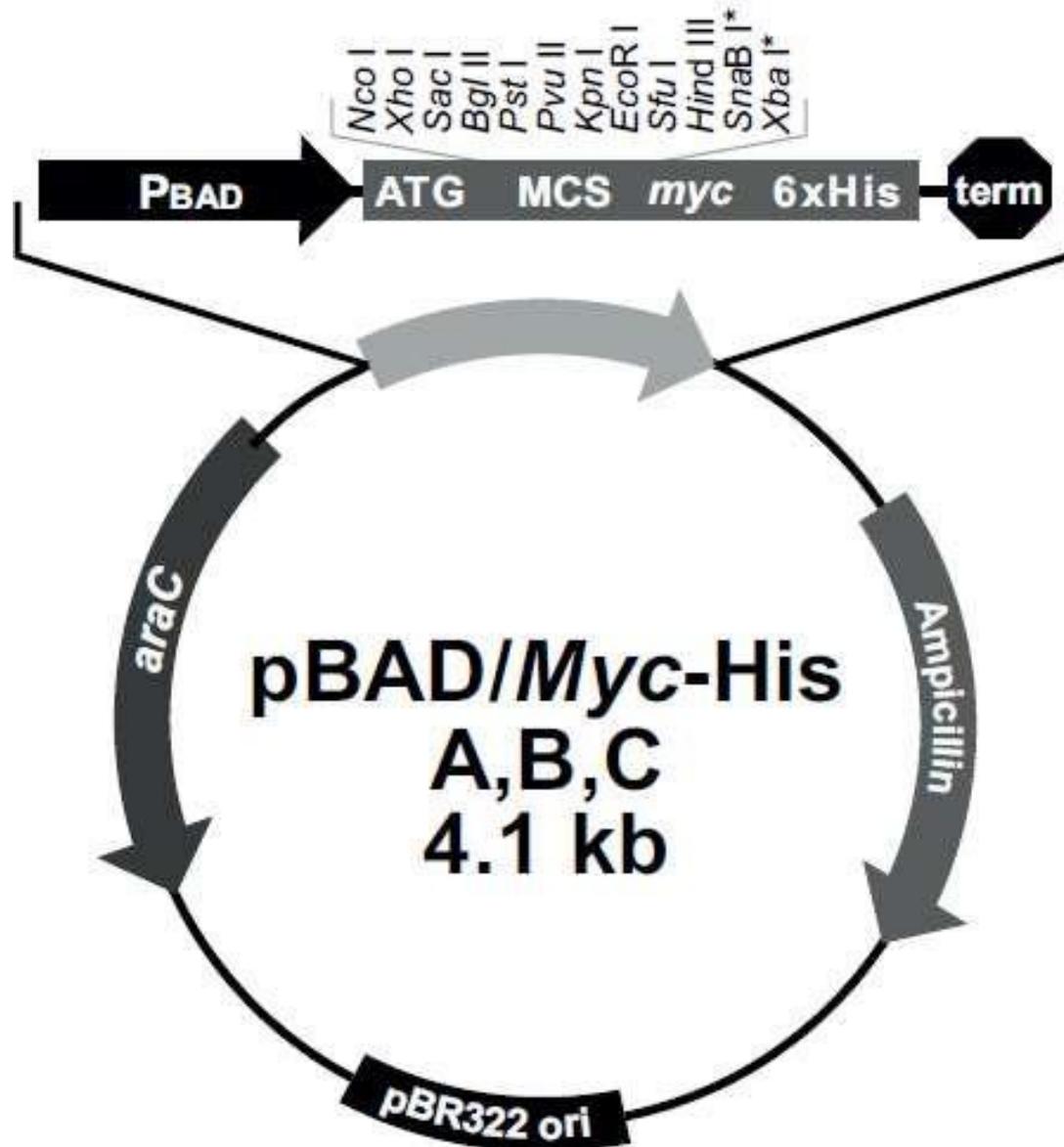
pBAD

- The **pBAD/His** and **pBAD/Myc-His** plasmids are **pBR322-derived expression vectors** designed for regulated, **dose-dependent recombinant protein expression** and purification in *E. coli*.
- Optimum levels of soluble, recombinant protein are possible using the ***araBAD* promoter (PBAD)** from *E. coli*.
- **The regulatory protein, AraC**, is provided on the pBAD/His and pBAD/Myc-His vectors allowing regulation of PBAD.
- In the presence of **L-arabinose**, expression from PBAD is turned on while the absence of L-arabinose produces very low levels of transcription from PBAD (Lee, 1980; Lee *et al.*, 1987).
- By **varying the concentration of L-arabinose**, protein expression levels can be optimized to **ensure maximum expression of soluble protein**.
- In addition, the **tight regulation of PBAD by AraC** is useful for **expression of potentially toxic or essential genes** (Carson *et al.*, 1991; Dalbey and Wickner, 1985; Guzman *et al.*, 1992; Kuhn and Wickner, 1985; Russell *et al.*, 1989; San Millan *et al.*³³,

pBAD



pBAD



APPLICATION OF EXPRESSION VECTOR

- The construction of **expression libraries**.
- The **analysis of gene function** at protein level.
- The commercial **production of proteins**.
- The **production of antibodies**.
- For **in vivo studies** of the protein.

REFERENCES

- Expression vectors: how to choose or customize vectors for gene & protein expression - Rachel Speer, Ph.D. Molecular Biology Specialist
- Expression system BY C.SWORNAKUMARI
M.Sc.,M.PHIL BIOTECHNOLOGY
- Invitrogen Life Technologies
pBAD/His A, B, and C
pBAD/*Myc*-His A, B, and C
Vectors for Dose-Dependent Expression of Recombinant Proteins Containing N- or C-Terminal 6xHis Tags in *E. coli*
Catalog nos. V430-01, V440-01

THANK
YOU