

# **Production of recombinant proteins**

# Vectors to express proteins in cells and/or purify recombinant proteins

- In bacteria
- In yeast
- In fly cells
- In vertebrate cells

# Vettori di espressione per la produzione di proteine

## A cosa possono servire le proteine ricombinanti?

- PROTEINE DI INTERESSE TERAPEUTICO.
- PROTEINE DI INTERESSE COMMERCIALE (ENZIMI).
- PROTEINE DA UTILIZZARE COME ANTIGENI PER LA PRODUZIONE DI ANTICORPI POLICLONALI E MONOCLONALI.
- REAGENTI PER LA RICERCA BI BASE E APPLICATA.

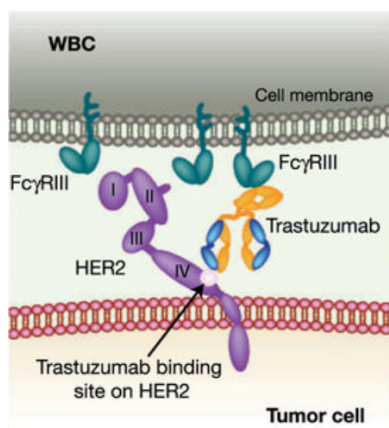
## Produzione di proteine

Brand	Generic	Company	Therapeutic category	Indications
Humulin	Insulin	Eli Lilly	Diabetes	Diabetes
Humatrope	Recombinant Somatropin	Eli Lilly	Hormones	Growth failure
Genotropin	Somatropin	Pfizer	Hormones	Growth failure
Saizen	Somatropin	Serono	Hormones	Growth failure
Nutropin/Protropin	Somatropin/Somatrem	Genentech	Hormones	Growth failure
Intron A &	Interferon alpha-2b/	Schering-Plough	Anti-infective	Viral infections
Avonex	Interferon beta-1a	Biogen Idec	Multiple sclerosis	Chronic inflammatory demyelinating polyneuropathy
Betaseron/Betaferon	Interferon beta-1b	Schering AG	Multiple sclerosis	Multiple sclerosis
Procrit/Eporex	Epoetin alpha	J&J	Blood modifier	Anaemia
Epogen	Epoetin alpha	Amgen	Blood modifier	Anaemia
NeoRecormon	Epoetin beta	Roche	Blood modifier	Anaemia
Kogenate	Factor VIII	Bayer	Blood modifier	Haemophilia
NovoSeven	Factor VIIa	Novo Nordisk	Blood modifier	Haemophilia
Benefix	Factor IX	Wyeth	Blood modifier	Haemophilia
Fabrazyme	Agalsidase beta	Genzyme	Enzymes	Fabry disease
Replagal	Agalsidase alfa	TKT Europe	Enzymes	Fabry disease
Pulmozyme	Dornase alpha	Genentech	Enzymes	Cystic fibrosis
Activase/Actilyse	Alteplase	Genentech	Blood factor	Myocardial infarction

# Produzione di proteine

## Monoclonal Antibodies

Brand	Generic	Company	Therapeutic category	Indications
ReoPro	Abciximab	Eli Lilly	Blood modifier	Acute coronary syndrome
Rituxan	rituxumab	Genentech	Cancer	Non-Hodgkin's lymphoma
Herceptin	Trastuzumab	Genentech	Cancer	Breast cancer
Synagis	Palivizumab	MedImmune	Respiratory	Respiratory syncytial virus
Campath	Alemtuzumab	Schering AG	Cancer	Non-Hodgkin's lymphoma
Humira	Adalimumab	Abbott Labs	Anti-arthritic	Rheumatoid arthritis
Xolair Omalizumab	Omalizumab	Genentech	Respiratory diseases	Paediatric asthma, peanut allergies
Erbitux	Cetuximab	Imclone Systems	Cancer	Colon cancer
Avastin	Bevacizumab	Genentech	Cancer	Colon cancer



Trastuzumab: HER2 positive breast cancer;  
recruits immune effector cells via the antibodies Fc domain

# Vettori di espressione per la produzione di proteine

## QUALI SISTEMI ETEROLOGHI UTILIZZARE PER L'ESPRESSIONE DEI GENI ?

E' virtualmente possibile esprimere geni in sistemi di ogni tipo utilizzando vettori d'espressione appropriati, in funzione di esigenze specifiche.

I più diffusi:

- *Escherichia coli*,
- *Bacillus subtilis*,
- *Lieviti (yeast)*
- *cellule d'insetto/sistemi virali (Insect cells)*
- *cellule vegetali*
- *cellule di mammifero in coltura (mammalian cells)*

**Strategia :**

- **Sovraespressione in sistema autologo or eterologa**
- **Sistema di purificazione**

L'espressione in E.coli è di gran lunga la più semplice e, forse, per questo la più utilizzata come prototipo di espressione genica in **sistemi eterologhi**.

# Codon usage in different organisms

Table 3.2 The genetic code and codon usage in *E. coli* and humans

Codon	Amino acid	Frequency of use in:	
		<i>E. coli</i>	Humans
GAG	Glutamic acid	0.30	0.59
GAA	Glutamic acid	0.70	0.41
CGG	Arginine	0.08	0.19
CGA	Arginine	0.05	0.10
CGU	Arginine	0.42	0.09
CGC	Arginine	0.37	0.19
AGG	Arginine	0.03	0.22
AGA	Arginine	0.04	0.21
CCG	Proline	0.55	0.11
CCA	Proline	0.20	0.27
CCU	Proline	0.16	0.29
CCC	Proline	0.10	0.33
UGA	Stop	0.30	0.61
UAG	Stop	0.09	0.17
UAA	Stop	0.62	0.22

Problem:

**The expression of human cDNAs in bacteria is often inefficient**

**Codoni sinonimi:**

Tre dei 64 codoni sono segnali di terminazione della sintesi proteica (codoni di stop o di terminazione), mentre gli altri 61 codificano per i 20 amminoacidi. **Nei casi in cui uno stesso amminoacido è codificato da più codoni (da 2 a 6), questi si chiamano codoni sinonimi.**

Se alcuni tRNA sono rari in quell'ospite, allora alcuni codoni non verranno riconosciuti e funzioneranno come codoni di stop, causando sintesi premature; oppure potrebbero non corrispondere all'inserzione dell'aminoacido desiderato (codon bias).

**Bisognerà modificare il cDNA ricombinante da clonare in modo che contenga tra i sinonimi quei codoni per i tRNA più frequenti nell'ospite; oppure bisognerà ingegnerizzare l'ospite affinché esprima i tRNA rari (ceppi Rosetta di *E. coli*).**

## Pro and Cons of expressing human proteins in bacteria

### Vantaggi:

- Vasta scelta di vettori di clonaggio
- Vasta scelta di ceppi
- Controllo relativamente semplice dell'espressione a livello genetico
- Buona resa della proteina ricombinante (25% del totale delle proteine)
- La proteina ricombinante può essere espressa come proteina di fusione
- La proteina ricombinante può essere disegnata per essere secreta nel terreno di crescita
- Economico

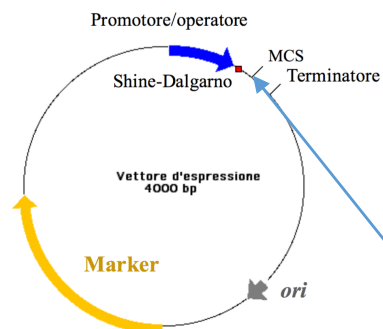
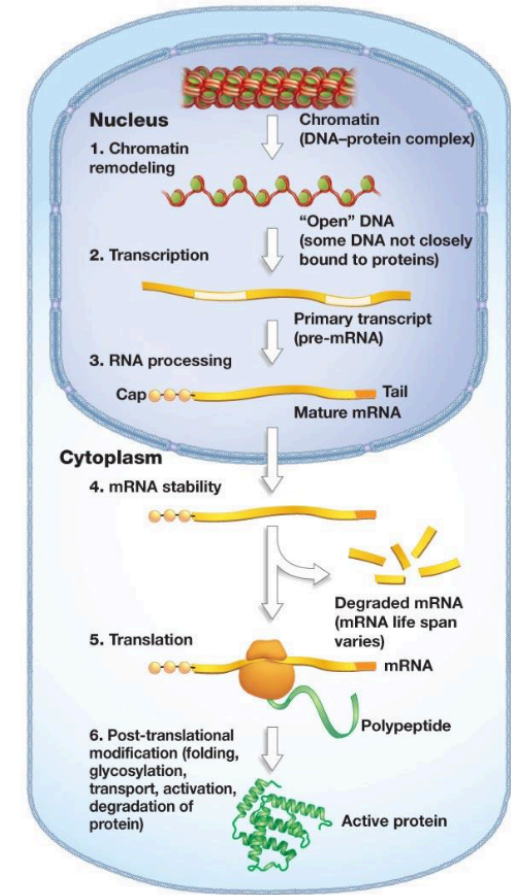
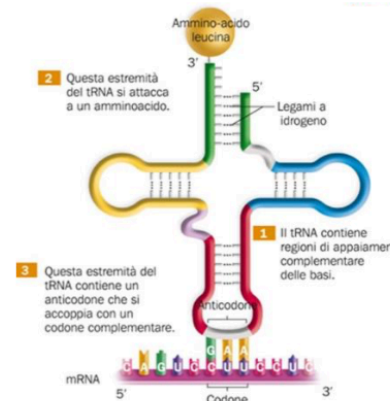
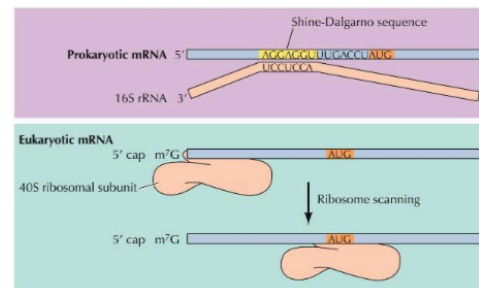
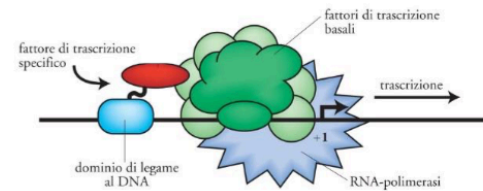
### Svantaggi:

- La proteina ricombinante mancherà di modificazioni post-trasduzionali
- L'attività biologica del ricombinante può essere diversa dalla proteina naturale
- Carico metabolico molto pesante con overespressione di proteine che a volte porta alla formazione di inclusion bodies, aggregati che rendono difficoltosa la purificazione del prodotto e possono ridurre la sua attività biologica

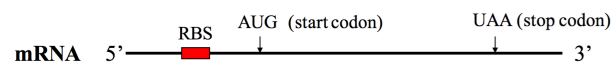


# Fattori influenzanti l'espressione dei geni clonati per ottimizzare l'espressione di proteine di mammifero nei batteri

- Il **promotore** (batterico/eucariotico): la regione legante l'RNA polimerasi e segnali regolanti l'espressione genica (iniziare la sintesi del trascritto di mRNA)
- **cDNA**: più usato, rispetto al clonaggio di DNA genomico contenente introni
- L'inizio della traduzione: **sequenza di Shine-Dalgarno legante il 3' dell'rRNA 16S (nei batteri) e la sua distanza dal codone d'inizio della traduzione**, per inserire un appropriato sito di legame al ribosoma e un codone d'inizio correttamente posizionato. Inoltre, alcuni organismi non usano l'AUG, ma codoni d'inizio alternativi (GUG, UUG, CUG), anche il **sito d'inizio dovrà essere opportunamente modificato base all'ospite**.
- **Codoni sinonimi** (diversi ma codificanti lo stesso aminoacido): la frequenza d'inserimento dell'aminoacido dipenderà dalla **disponibilità nella cellula del tRNA** specifico per ciascun codone.

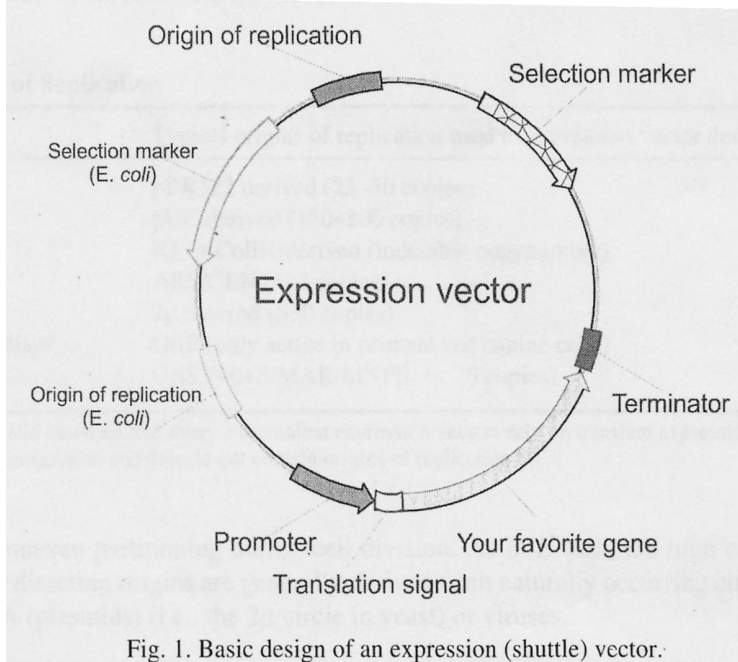


cDNA sequence of insulin (start → stop codon)  
Cloned in MCS (upstream Promoter, downstream transcriptional stop)



# VETTORI D'ESPRESSIONE (Batterie)

## caratteristiche generali

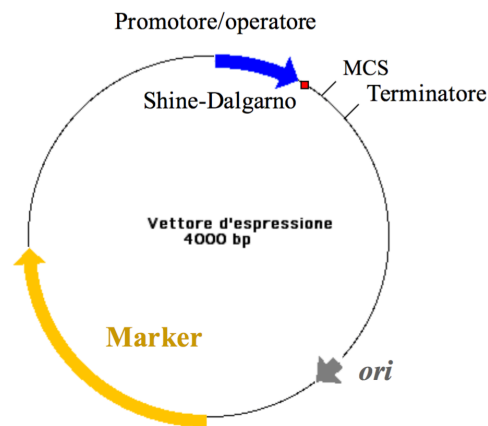


### Elementi standard

- Origine di replicazione
- Marker di selezione
- Promotore (T7, T3, SP6; normalmente inducibile)
- Terminatori di trascrizione
- Sequenze per iniziazione della traduzione
- Codoni di terminazione della traduzione

### Elementi genetici specifici per diverse applicazioni:

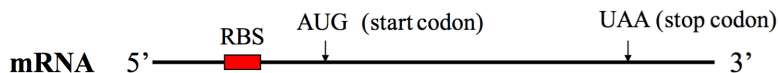
- Sequenze segnale per secrezione
- Peptidi per purificazione (tags)
- Ecc.



Subito dopo la sequenza di Shine-Dalgarno deve essere presente un codone di inizio, quasi **sempre AUG** (in una piccola percentuale di casi può essere presente il codone GUG).

La spaziatura ottimale tra SD e AUG è di 8 bp

E' importante che la sequenza nucleotidica tra la SD e il codone d'inizio non sia disturbata da strutture secondarie (es. hairpin loops) che possono interferire drasticamente con il legame al ribosoma e la conseguente traduzione.



# Caratteristiche del prodotto proteico ricombinante

**Geni in procarioti possono avere:**

- Espressione Costitutiva
- ESPRESSIONE REGOLATA (= inducibile!!! Es. Lac Operon)

**Una produzione continua può provocare:**

- Inibizione Funzioni Cellula
- Perdita Energia
- Perdita Plasmide
- Formazione aggregati (inclusion bodies)

**Nella produzione di proteine eterologhe in batteri vengono utilizzati spesso promotori forti e regolabili**

# Promotori per la espressione di proteine ricombinanti in *E. coli*

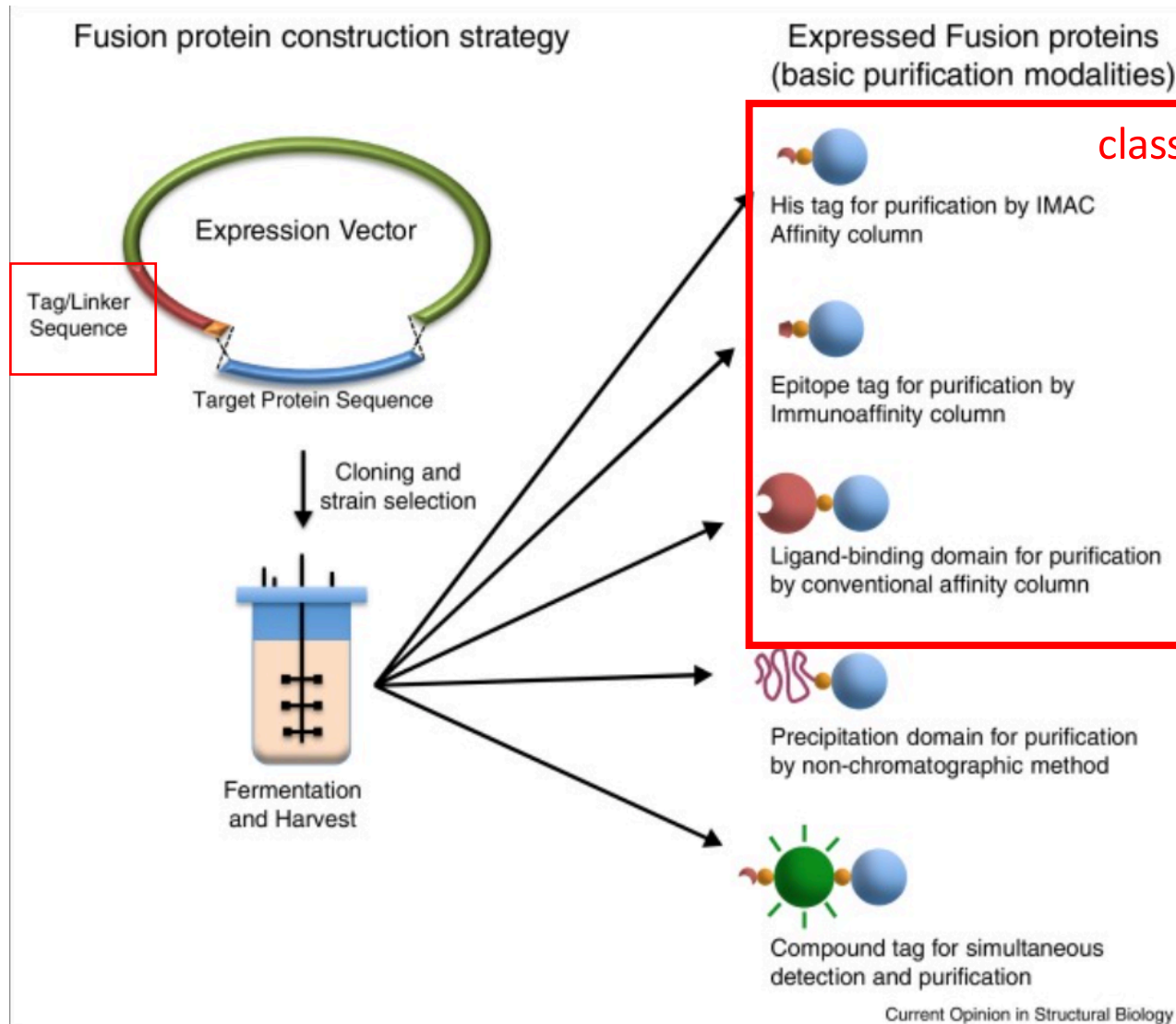
**Table 4** Some *E. coli* promoter systems that are in use for heterologous protein production and their characteristics

Expression system based on	Induction (range of inducer)	Level of expression	Key features	Original reference
<i>lac</i> promoter	Addition of IPTG 0.2 mM (0.05–2.0 mM)	Low level up to middle	Weak, regulated suitable for gene products at very low intracellular level Comparatively expensive induction	Gronenborn (1976)
<i>trc</i> and <i>tac</i> promoter	Addition of IPTG 0.2 mM (0.05–2.0 mM)	Moderately high	High-level, but lower than T7 system Regulated expression still possible Comparatively expensive induction	Brosius et al. (1985)
T7 RNA polymerase	Addition of IPTG 0.2 mM (0.05–2.0 mM)	Very high	High basal level Utilizes T7 RNA polymerase High-level inducible over expression T7 <i>lac</i> system for tight control of induction needed for more toxic clones Relative expensive induction	Studier and Moffatt (1986)
Phage promoter <i>p<sub>L</sub></i>	Shifting the temperature from 30 to 42 °C (45 °C)	Moderately high	Basal level depends on used strain (pLys) Temperature-sensitive host required Less likelihood of “leaky” uninduced expression Basal level, high basal level by temperatures below 30 °C	Elvin et al. (1990)
<i>tetA</i> promoter/operator	Anhydrotetracycline 200 µg/l	Variable from middle to high level	No inducer Tight regulation Independent on metabolic state Independent on <i>E. coli</i> strain Relative inexpensive inducer Low basal level	Skerra (1994)
<i>araBAD</i> promoter ( <i>P<sub>BAD</sub></i> )	Addition of L-arabinose 0.2 % (0.001–1.0 %)	Variable from low to high level	Can fine-tune expression levels in a dose-dependent manner Tight regulation possible Low basal level Inexpensive inducer	Guzman et al. (1995)
<i>rhaP<sub>BAD</sub></i> promoter	L-rhamnose 0.2 %	Variable from low to high level	Tight regulation Low basal activity Relative expensive inducer	Haldimann et al. (1998)

IPTG binds and inactivates lac repressor; lacI encoded by vector)

# Sistemi per la purificazione di proteine ricombinanti in E. coli

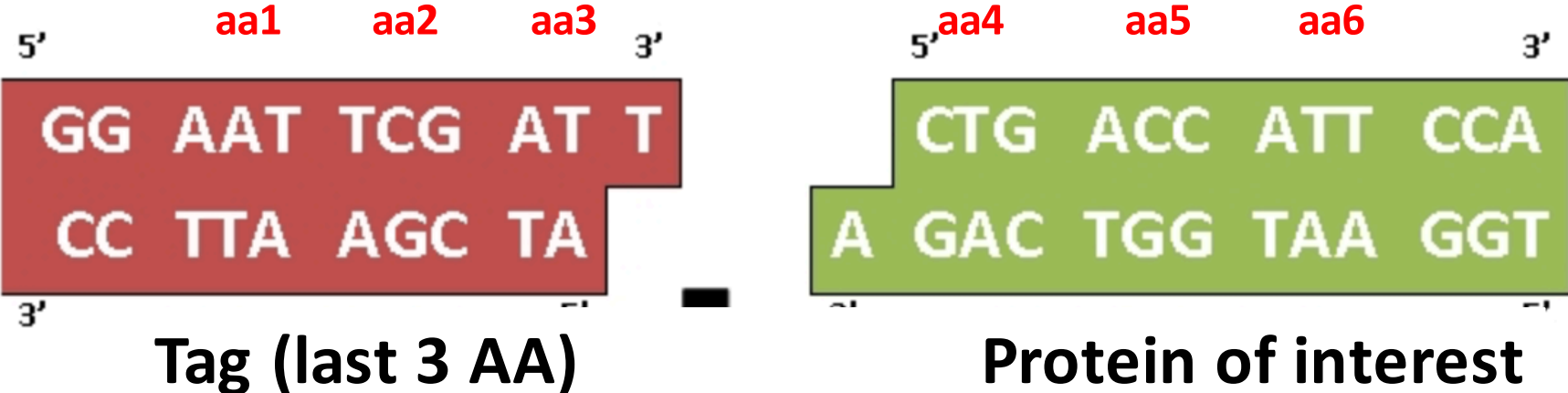
Si preferisce esprimere proteine di fusione (Tag + proteina di interesse), per la loro maggiore stabilità, gli alti livelli di espressione e la relativa facilità con cui si purificano.



- Tag is a short peptide sequence fused to the protein of interest (NH<sub>2</sub> or COOH terminus of protein) that can be used to purify the recombinant protein
- Tag can be captured by a special surface of a solid resin (in suspension) (by IMAC = Immobilized metal affinity chromatography; or Immunoaffinity column)
- Rest of cell lysate is discarded
- Recombinant protein bound to resin
- Washing
- Elution from column

# Sistemi per la purificazione di proteine ricombinanti in E. coli

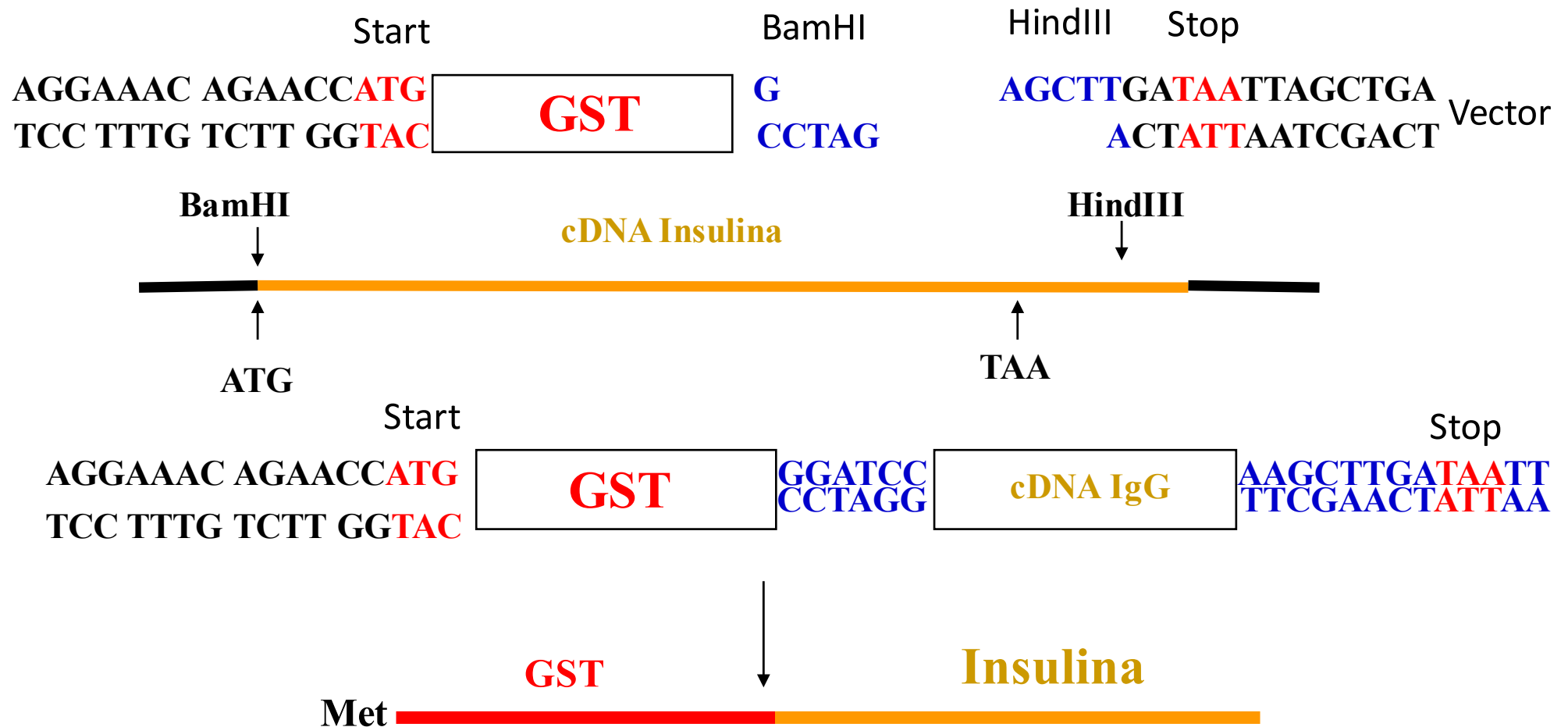
TAG and Protein must be “in frame”)



# Sistemi per la purificazione di proteine ricombinanti in E. coli

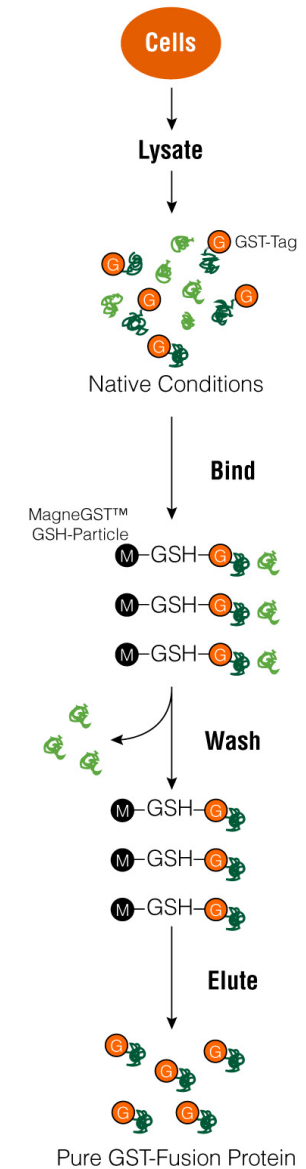
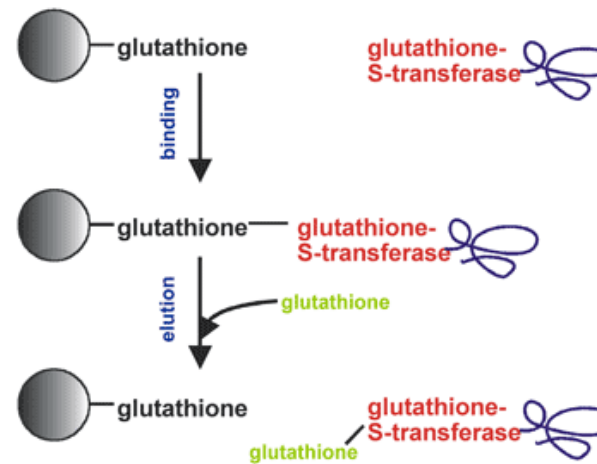
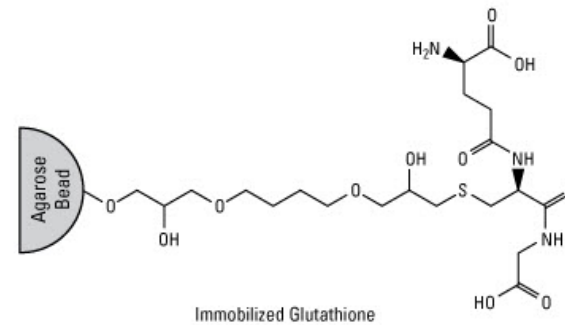
## GST-tagged proteins:

A convenient method of protein expression and subsequent purification is to fuse a protein with a glutathione-S-transferase (GST) domain. The DNA encoding for this 25 kDa protein domain is ligated in-frame with the gene for the desired protein so that, upon expression, your desired protein is fused to the GST domain. This is an incredible help in protein purification, since **GST binds glutathione extremely strongly**. The general purification strategy is thus to bind the GST fusion protein on a column of immobilized glutathione, wash away all the undesired molecules, and then elute the protein.



# Sistemi per la purificazione di proteine ricombinanti in E. coli

1. Construction of GST-tagged cDNA in frame. Cloned into bacterial expression vector.
2. Transformation of bacteria
3. Growth of bacteria
4. Induction of protein expression via the use of an inducible promoter
5. Cell harvest and lysis
6. Binding of recombinant protein to resin with immobilized glutathione
7. Bound protein can be eluted with excess of glutathione (in elution buffer – competition).





# Sistemi per la purificazione di proteine ricombinanti in E. coli

AGGAAAC AGAACCATG  
TCC TTTG TCTT GGTAC

HisHisHisHisHisHis

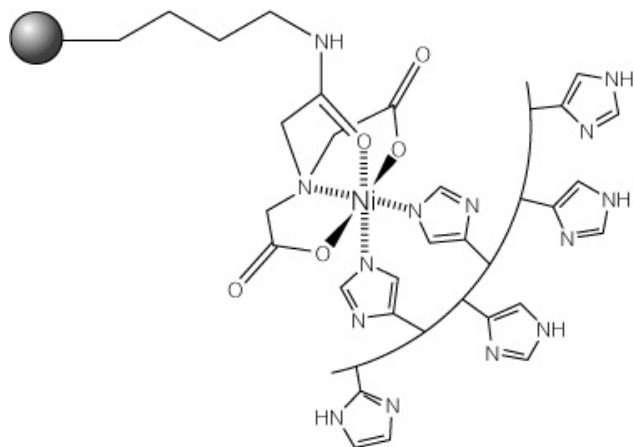
GGATCC  
CCTAGG

cDNA IgG

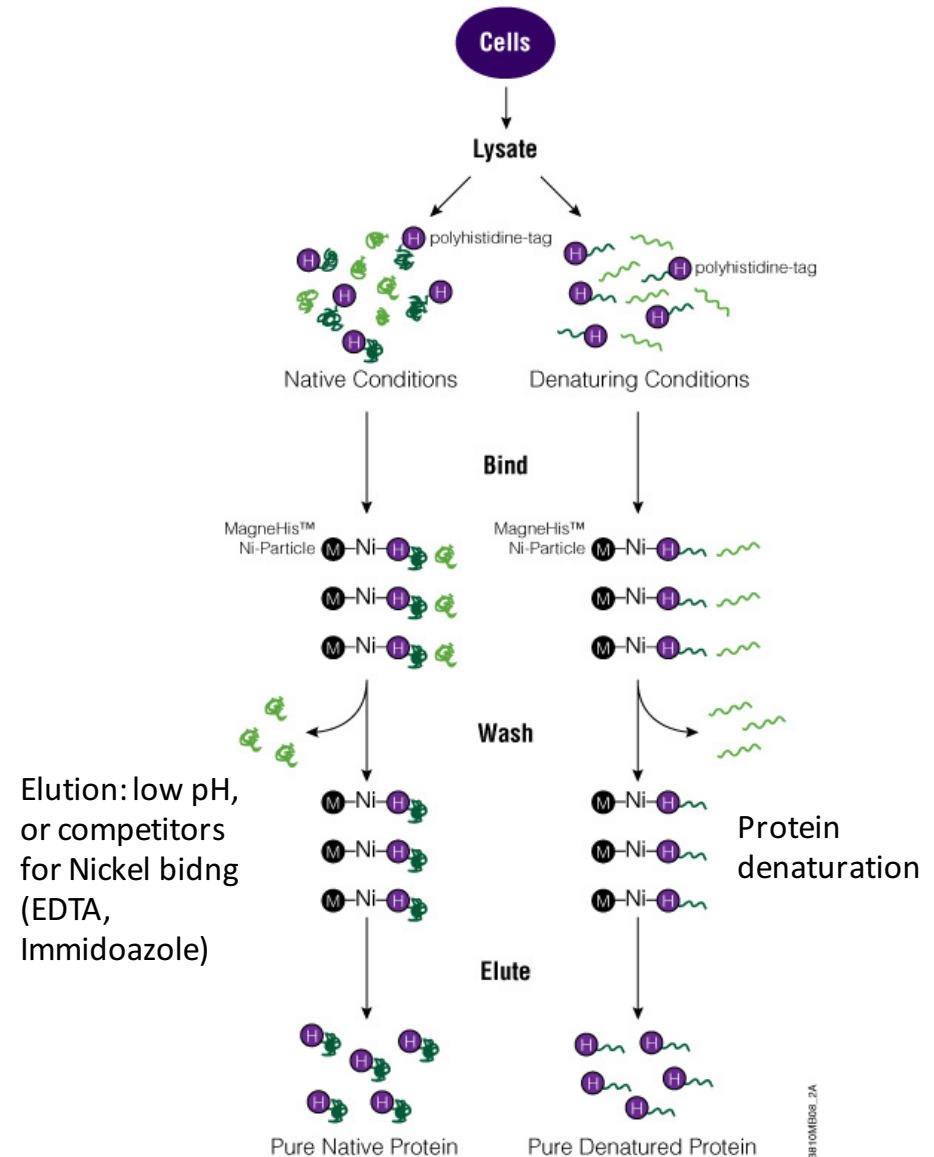
AAGCTT  
TTCGAA

## HISTIDINE (His)-tagged proteins:

In general, proteins possess more or less the ability to coordinate metal ions on their surface, **Histidine is strongly involved in the coordinate bond with metal ions.** Therefore, if a number of histidines are added to the end of the protein by genetic engineering, the affinity of the protein for the metal ion is remarkably increased and the basic idea is that purification can be easily carried out. When a protein having a His tag is brought into contact with a carrier on which a metal ion such as **nickel** is immobilized under the condition of pH 8 or higher, the histidine residue chelates the metal ion and binds to the carrier.



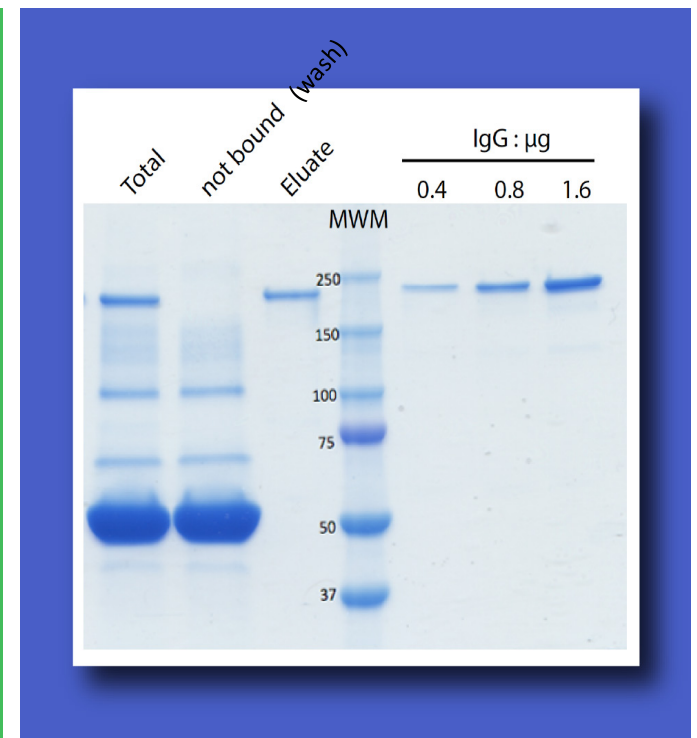
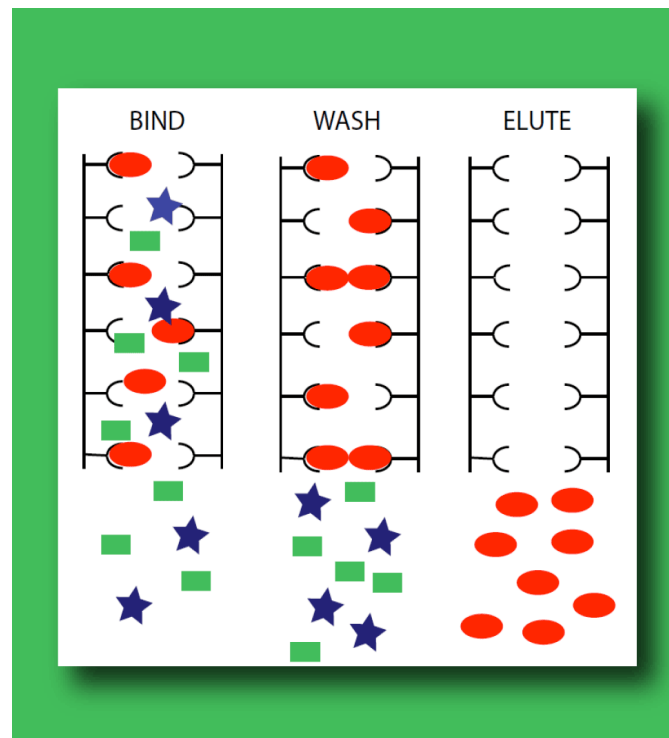
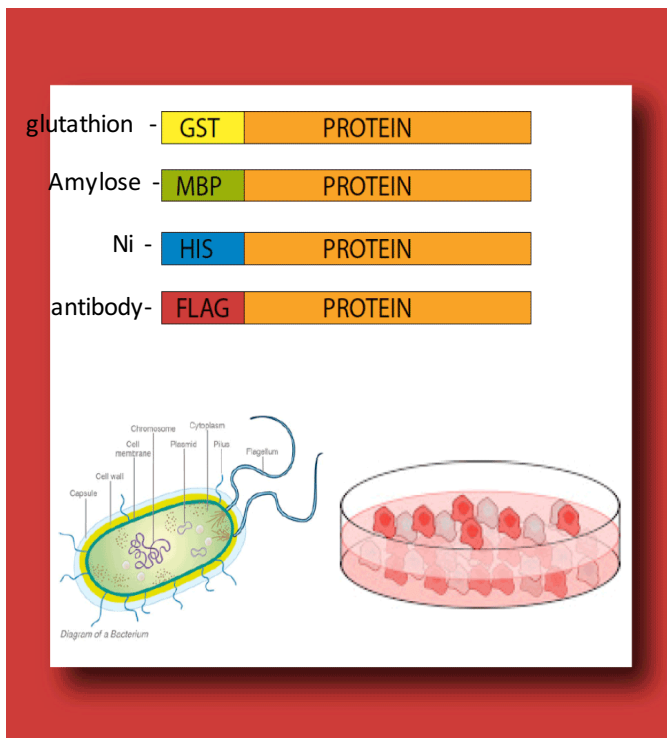
4420MA12\_3A



3810MBoe\_2A

# COMMON TAGs FOR EXPRESSION OF RECOMBINANT PROTEINS IN E. COLI

- Calmodulin binding peptide (CBP)
- Proteina A (IgG binding domain)
- Chitin binding domain (CBD)
- MBP (Maltose binding protein)
- Strep tag (Streptavidin binding tag)



## 1. Vectors express and purify recombinant proteins

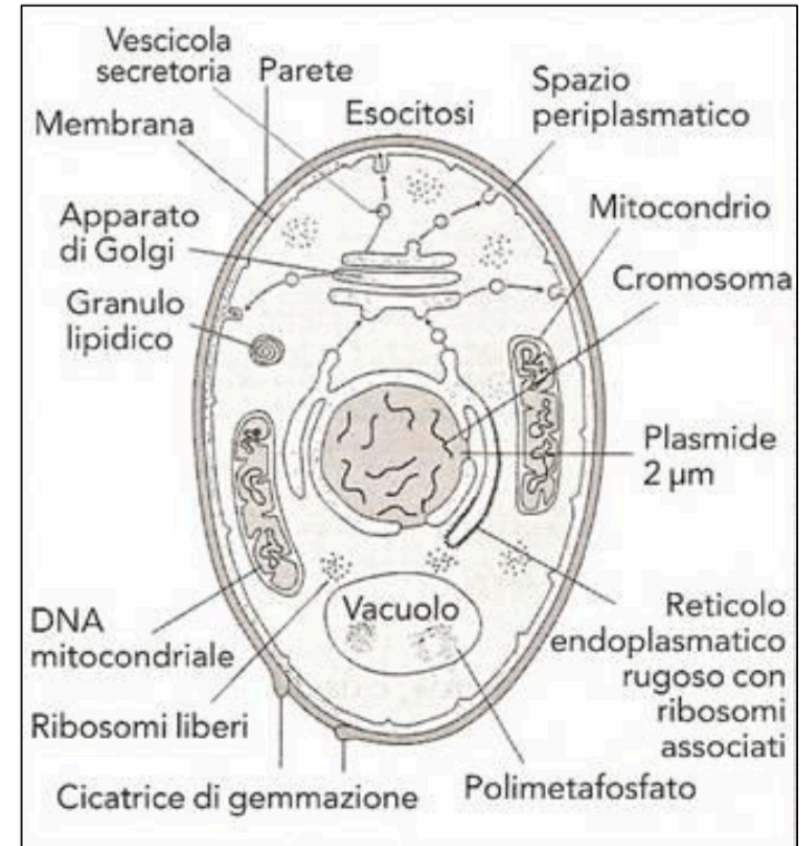
- In bacteria
- **In yeast**
- In fly cells
- In vertebrate cells

# Trasformazione di cellule eucariotiche di lievito (e funghi filamentosi)

## ❑ Inserimento di DNA ricombinante in cellule di *Saccaromices cerevisiae*:

- Biochimica cellulare e i meccanismi di regolazione del lievito sono **simili agli eucarioti superiori**
- Trasduzione del segnale e regolazione della trascrizione da parte di **ormoni steroidei simile ai mammiferi**
- Molti geni sono omologhi di geni umani, tra questi molti coinvolti nella divisione cellulare
- **Sistemi modello di studio**
- Possono **esprimere proteine** eterologhe (farmaci proteici)
- Rappresentano **modelli di screening di farmaci**

- ▶ Production of sferoplasts (treatment of call wall with enzymes such as zymolyase, glucanase or lyticase) and treatment with polyethyleneglycol
- ▶ Electroporation



# Vettori di espressione per la produzione di proteine in lievito

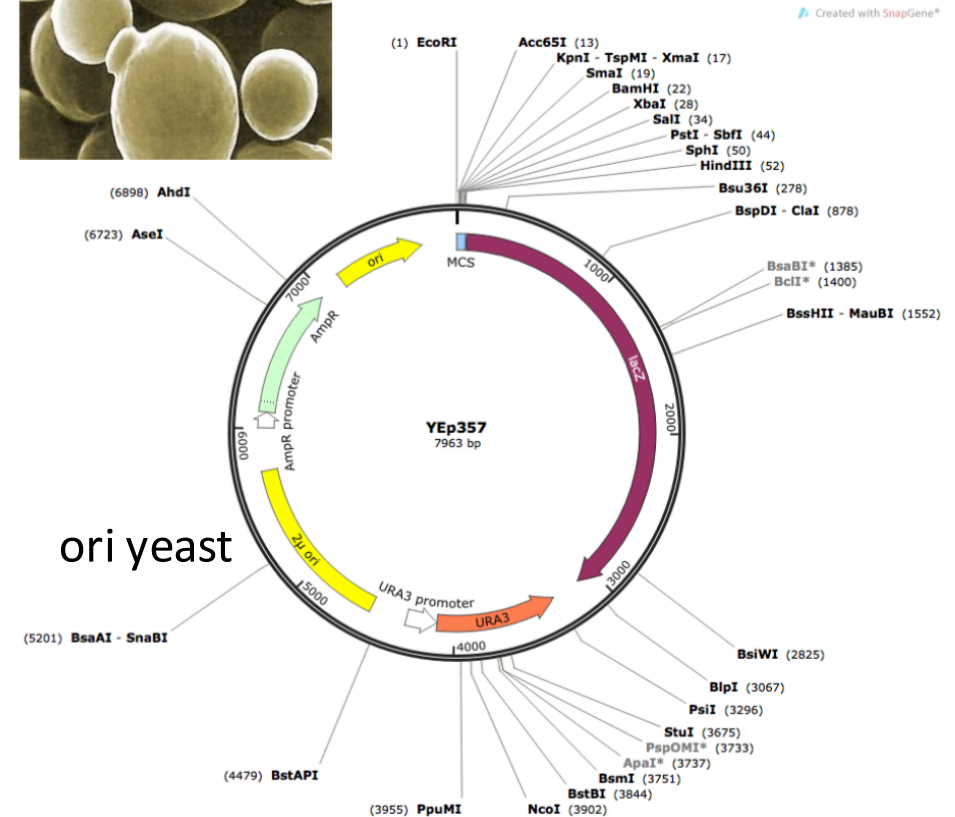
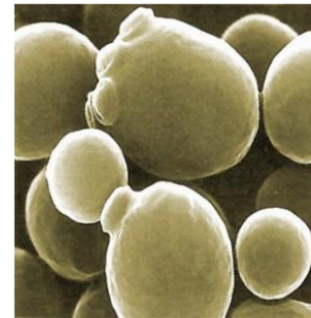
**Sono funghi**, organismi eucariotici unicellulari di forma tonda o ellittica

Dimensioni: 3-50 micron

Più di 1000 specie, ma i più usati (per lievitare il pane e fermentare bevande alcoliche) appartengono agli Ascomiceti: *Saccharomyces cerevisiae*

## Yest Episomal Plasmid (Yep; repliconi episomali):

1. Si replicano autonomamente ad alto numero, **20-100 copie per cellula**, grazie all'aggiunta la **sequenza di replicazione autonoma ars**. Essendo instabili, tendono ad essere persi durante la duplicazione delle cellule di lievito.
2. Hanno un **ori batterica**, che consente la loro crescita e manipolazione anche nei batteri (*vettori shuttle*)
3. Hanno anche un **marcatore per la resistenza agli antibiotici**, per la selezione in *E. coli*
4. Ma è necessario anche un secondo marcatore di selezione rispetto ai batteri, poiché i lieviti sono sensibili a un numero ridotto di antibiotici, sono quindi usati altri marcatori di selezione, come la «**complementazione di mutazioni auxotropiche**», associate alla **produzione di enzimi coinvolti nella biosintesi di aminoacidi o altre molecole**: *trp1* (triptofano), *leu2* (leucina), *his3* (istidina) o *ura3* (uracile).



Da utilizzare con un ceppo *ura3<sup>-</sup>*

# Vettori di espressione per la produzione di proteine in lievito

## Nuovi costrutti hanno anche un centromero **Yest Centromero Plasmid (YCp)**:

Le sequenze centromeriche si associano a proteine specifiche che formano un complesso multiproteico capace di legare i microtubuli del fuso responsabili della segregazione, consentendo una corretta distribuzione delle copie episomali durante la divisione cellulare. Basso numero di copie **1-2 copie per cellula**.

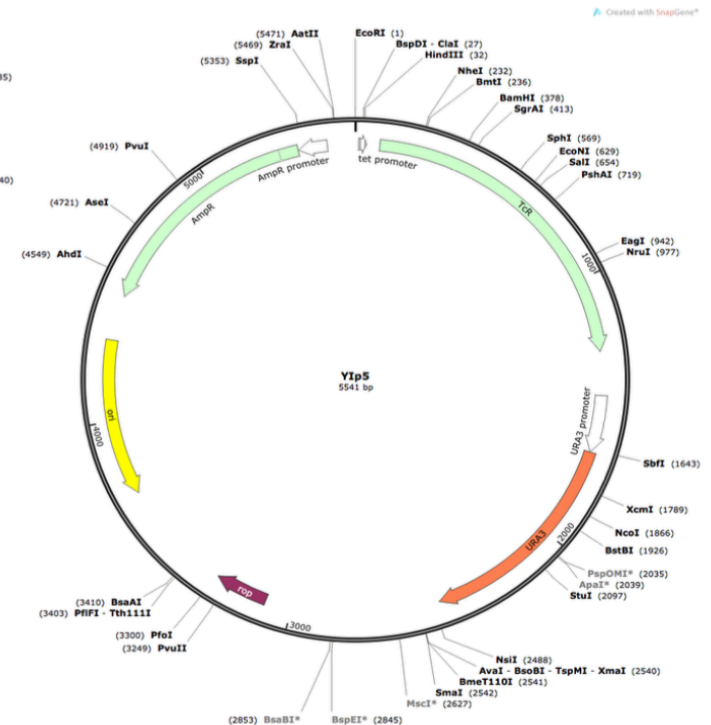
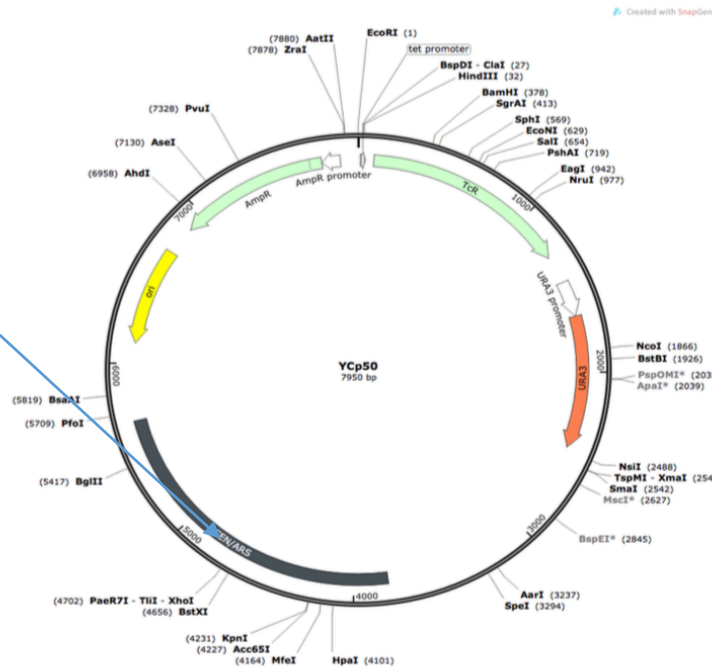
Vantaggiosi **quando il gene clonato è dannoso per l'ospite, o se ne vuole studiare la regolazione (studi d'induzione dell'espressione)**

## **Yest Integrating Plasmid (YIp)**:

Si replicano autonomamente, ma **si integrano anche nel cromosoma di lievito per ricombinazione omologa**, hanno bassa frequenza di trasformazione, ma sono **più stabili**

## **Yest Artificial chromosomes (YAC)**:

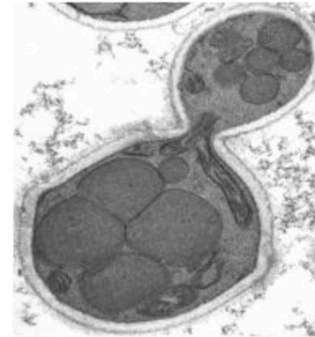
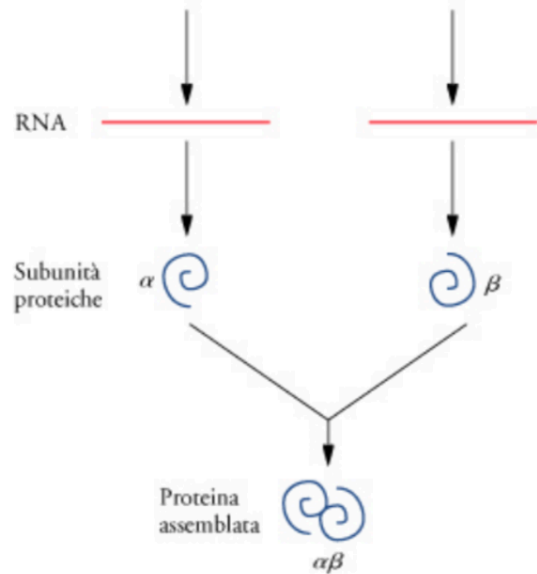
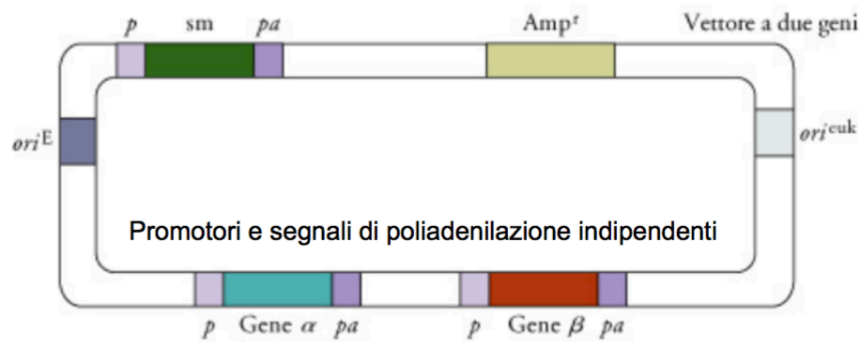
Usati per clonare **frammenti molto grandi**



# Vettori di espressione per la produzione di proteine in lievito

## espressione di 2 geni in una stessa cellula

### vettori a doppia cassetta



*Hansenula polymorpha*

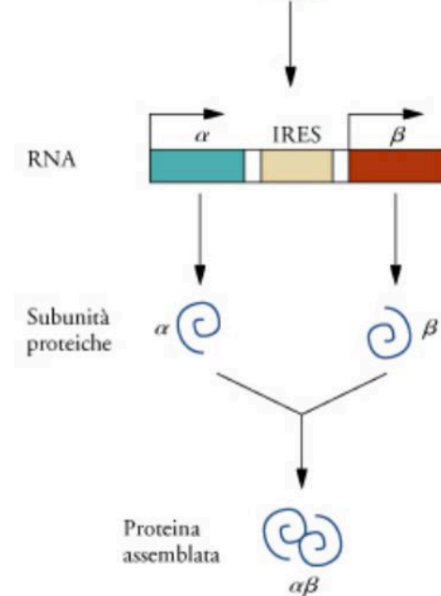
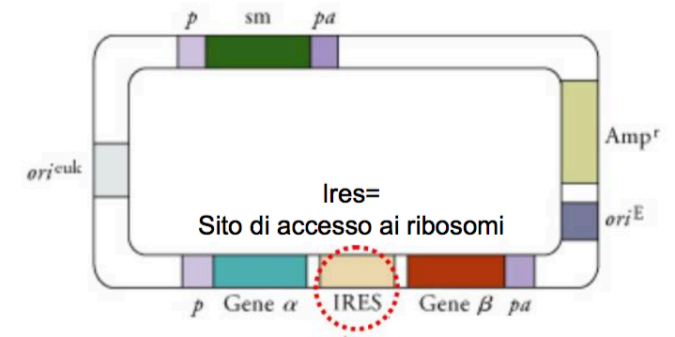
### Geni clonati in tandem:

- Necessaria quando la forma attiva di una proteina è costituita da 2 proteine diverse (eterodimero)
- Preferibile all'assemblaggio *in vitro*
- Necessità di esprimere le subunità contemporaneamente e stechiometricamente
- Promotore MOXp (metanolo ossidasi)
- Sequenza di arresto della trascrizione MOXt
- Integrazione fortuita dopo varie generazioni
- Produzione del tetramero  $\alpha_2\beta_2$  funzionale

### Esempi:

- Ormone stimolatore della tiroide ( $\alpha\beta$ )
- Emoglobina ( $\alpha_2\beta_2$ )

### vettori dicistronici



## Promoters to express recombinant proteins in yeast cells

*Table 7.1* Promoters for *S. cerevisiae* expression vectors

Promoter	Expression conditions	Status
Acid phosphatase ( <i>PH05</i> )	Phosphate-deficient medium	Inducible
Alcohol dehydrogenase I ( <i>ADHI</i> )	2–5% Glucose	Constitutive
Alcohol dehydrogenase II ( <i>ADHII</i> )	0.1–0.2% Glucose	Inducible
Cytochrome <i>c</i> <sub>1</sub> ( <i>CYC1</i> )	Glucose	Repressible
Gal-1-P Glc-1-P uridylyltransferase	Galactose	Inducible
Galactokinase ( <i>GAL1</i> )	Galactose	Inducible
Glyceraldehyde-3-phosphate dehydrogenase ( <i>GAPD, GAPDH</i> )	2–5% Glucose	Constitutive
Metallothionein ( <i>CUP1</i> )	0.03–0.1 mM copper	Inducible
Phosphoglycerate kinase ( <i>PGK</i> )	2–5% Glucose	Constitutive
Triose phosphate isomerase ( <i>TPI</i> )	2–5% Glucose	Constitutive
UDP galactose epimerase ( <i>GAL10</i> )	Galactose	Inducible





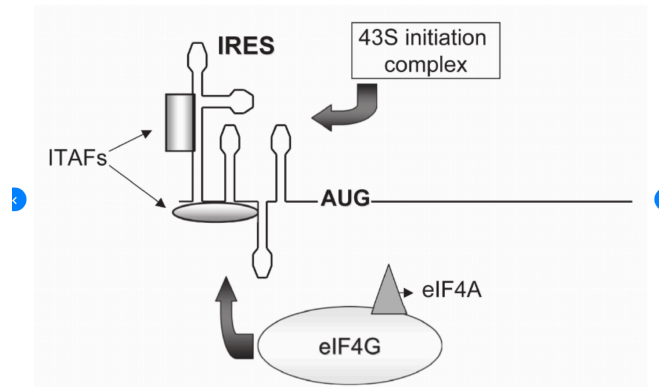
# Pre-requisites for recombinant protein expression in *S. cerevisiae*

## 1. Cap-dependent translation:

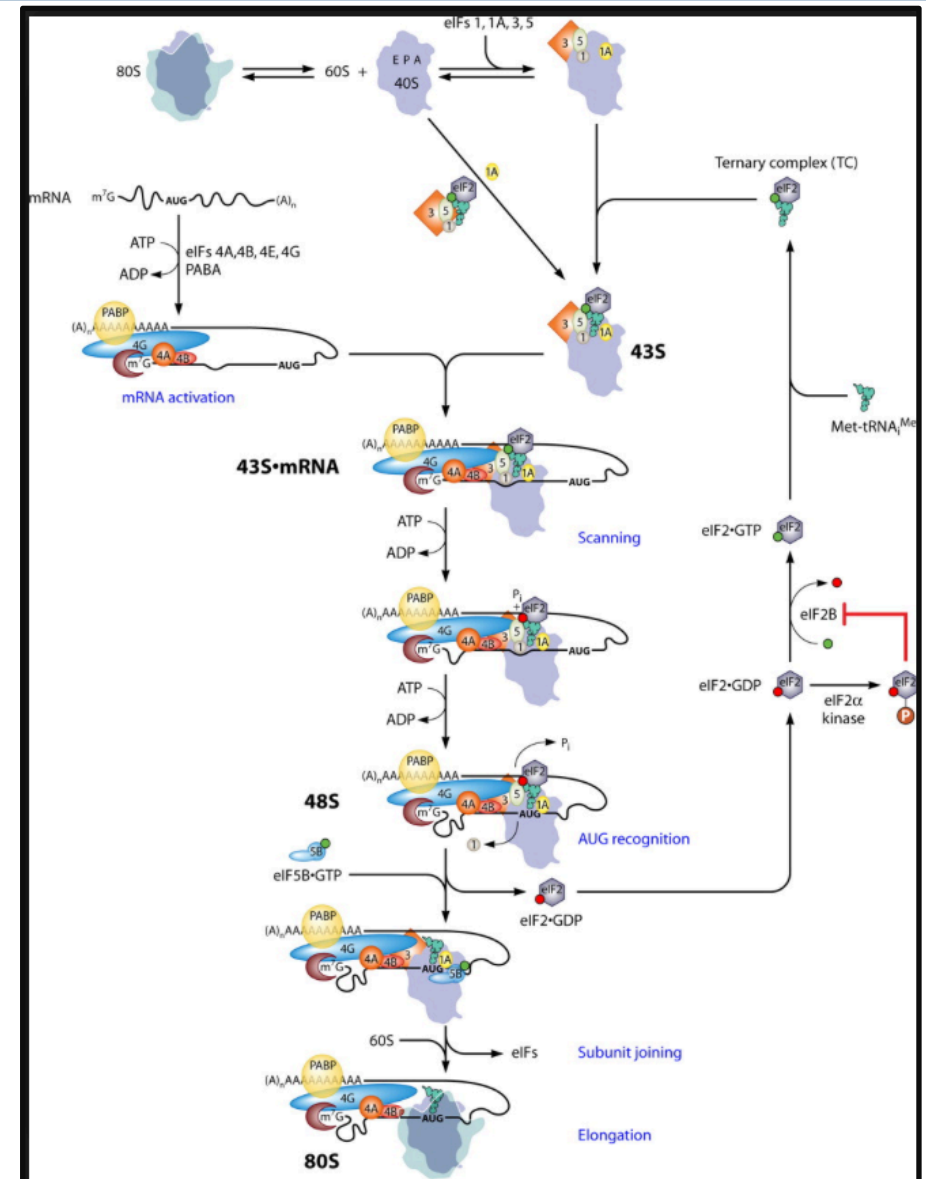
- **Cap structure** and the cap binding proteins are responsible for proper ribosome binding to mRNA and recognition of the correct **initiation codon**.
- The first AUG codon in the 5' end of mRNA functions as the initiation codon.
- **Kozak sequence** may be present around the initiation codon. (ACCAUGG)

## 2. Cap-independent translation:

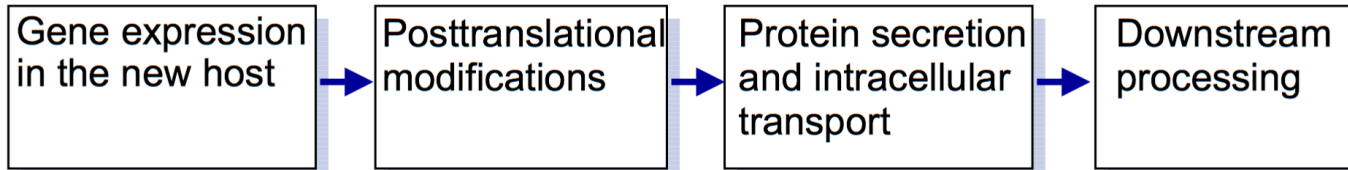
Ribosome binding to mRNA occurs through 'internal ribosome entry site' (IRES) on mRNA. Found in poliovirus (PV) and encephalomyocarditis virus (EMCV) RNA genomes



Schematic representation of internal ribosome entry site (IRES)-mediated translation initiation. Internal initiation is an alternative mechanism to cap-dependent translation initiation which allows loading of the 40S ribosomal subunit on the mRNA in a 5' end- and cap- independent fashion. Among the different IRESs canonical initiation factor requirements are variable. However, most IRESs require specific cellular proteins, IRES trans-acting factors (ITAFs), to be functional. See the text for details. For diagram simplicity, other proteins, as well as a second eIF4A molecule known to interact with eIF4G, have been omitted.



# Secretion of recombinant proteins from yeast cells



- Regulation of gene expression
- Molecular biology methods

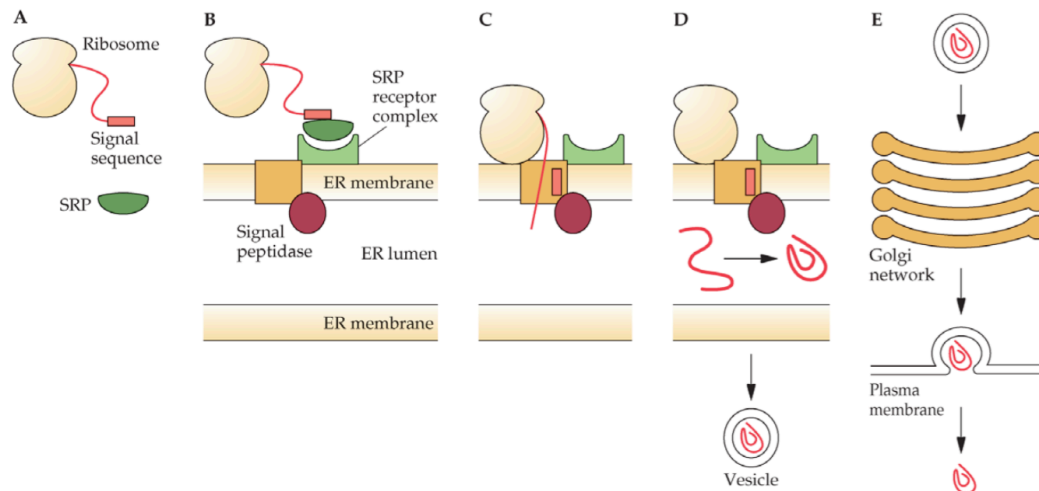
- Translocation to the secretory pathway
- Folding and modifications
- Glycosylation

- The secretory pathway: ER → Golgi → secret. vesicles → PM
- Secretion to the growth medium

- Chemical character of the protein
- Protein engineering
- Recovery & purification

Signal sequence of protein is recognized by signal recognition particle

- SRP (signal recognition particle) directs secretory protein to the translocation machinery on the ER-membrane; SRP cleaved off
- Translation and translocation process are coupled
- After translocation proteins undergo:
  - Folding
  - Post-translational modifications
  - Is transported through secretory pathway
  - Released into the medium



## Secretion

The yeast secretory pathway is very similar to that in higher eukaryotes.

N-terminal signal sequences for co-translational translocation of secreted proteins into the ER are removed by a signal peptidase. Examples of popular signal sequences used for secretion of heterologous proteins -these of Pho5, Suc2 and the  $\alpha$ -factor.

# Recombinant proteins produced in Yeast

## VACCINES

- Hepatitis B virus surface antigen
- Malaria circumsporozoite protein
- HIV-1 envelope protein

## DIAGNOSTICS

- Hepatitis C virus protein
- HIV-1 antigens

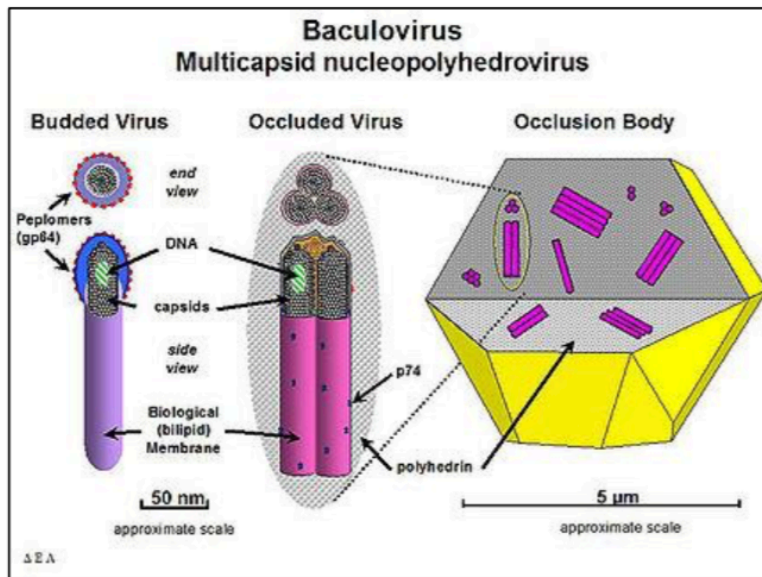
## HUMAN THERAPEUTIC AGENTS

- Epidermal growth factor
- Insulin
- Insulin-like growth factor
- Platelet-derived growth factor
- Proinsulin
- Fibroblast growth factor
- Granulocyte-macrophage colony-stimulating factor
- $\alpha_1$  antitrypsin
- Blood coagulation factor XIIIa

## 2. Vectors express and purify recombinant proteins

- In bacteria
- In yeast
- **In fly cells – insect cells**
- In vertebrate cells

# I baculovirus - *Autographa californica* nucleopolyhedrovirus (*AcMNPV*)



***Autographa californica***

The baculovirus life cycle involves two distinct forms of virus. **Occlusion derived virus (ODV)** is present in a protein matrix (**polyhedrin** or **granulin**) and is responsible for the primary infection of the host while the **budded virus (BV)** is released from the infected host cells later during the secondary infection

Typically, the initial infection occurs when a susceptible host insect feeds on plants that are contaminated with the occluded form of the virus. The protein matrix dissolves in the alkaline environment of the host midgut (stomach), releasing ODV that then fuse to the columnar epithelial cell membrane of the host intestine and are taken into the cell in endosomes. Nucleocapsids escape from the endosomes and are transported to nucleus. This step is possibly mediated by actin filaments. Viral transcription and replication occur in the cell nucleus and new BV particles are budded out from the basolateral side to spread the infection systemically. During budding, BV acquires a loosely fitting host cell membrane with expressed and displayed viral glycoproteins.

Baculovirus infection can be divided to three distinct phases:

**Early (0–6 h post-infection),**

**Late (6–24 h p.i.)**

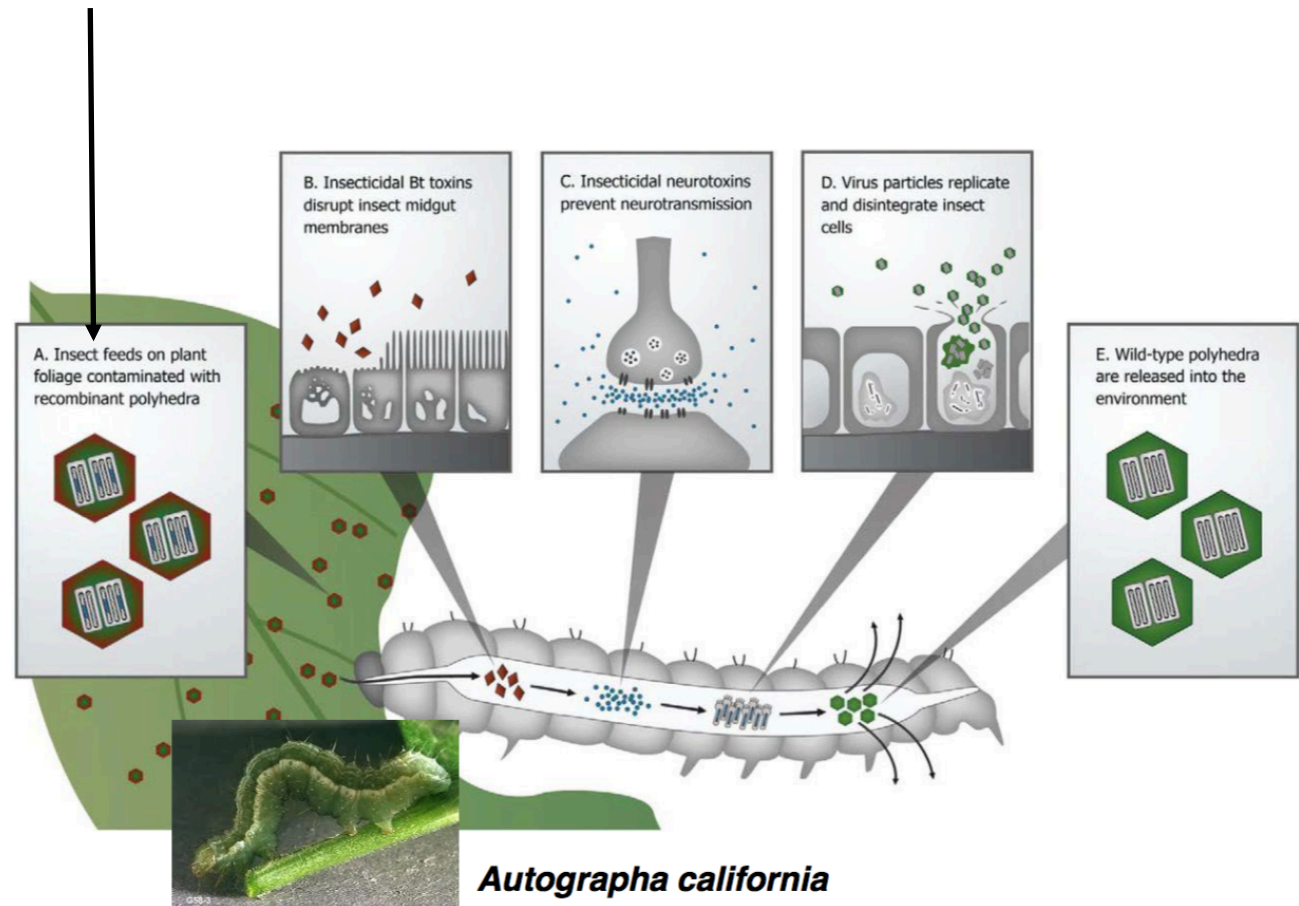
**Very late phase (18–24 to 72 h p.i.)**

While BV is produced in the late phase, the ODV form is produced in the very late phase acquiring the envelope from host cell nucleus and embedded in the matrix of occlusion body protein. These occlusion bodies are released when cells lyse to further spread baculovirus infection to next host. The extensive lysis of cells frequently causes the host insect to literally disintegrate, thus the reason for the historic name "wilting disease." The complete ODV-polyhedrin particles are resistant to heat and light inactivation, whereas the naked BV virion is more sensitive to environment

# I baculovirus - *Autographa californica* nucleopolyhedrovirus (AcMNPV)

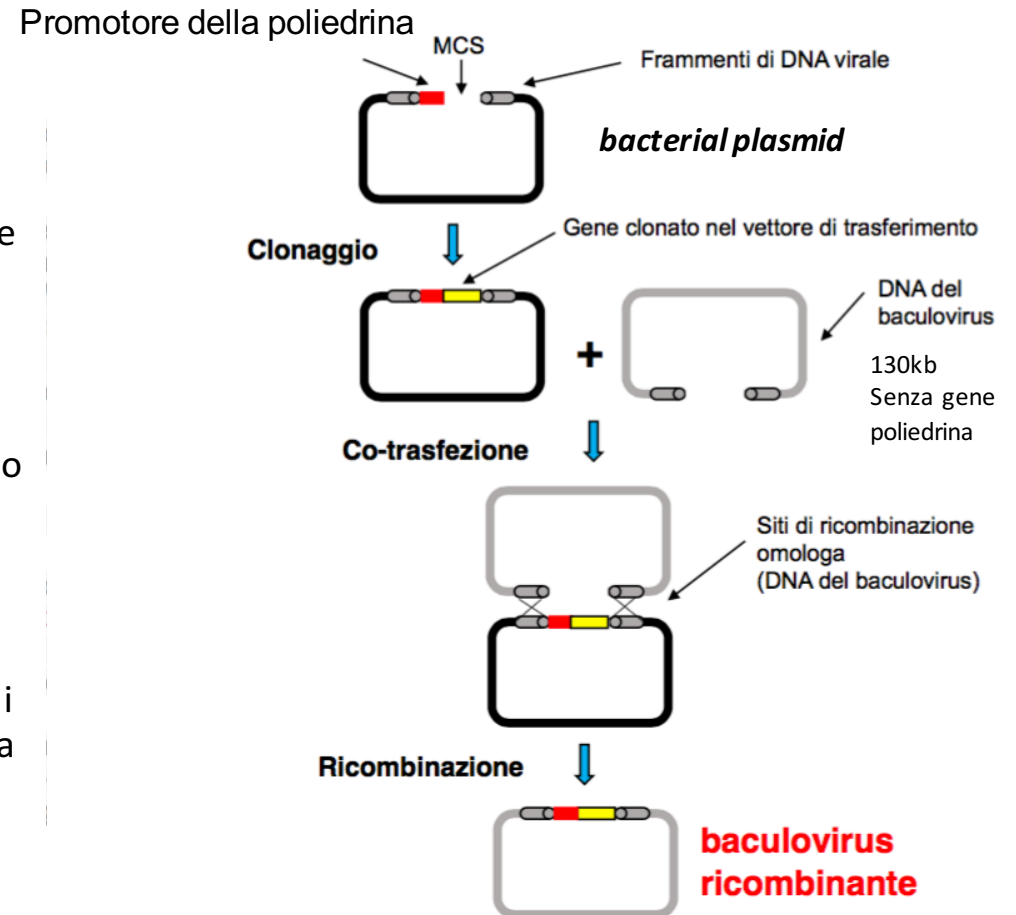
*Autographa californica* multiple nuclear polyhedrosis virus (AcMNPV) has a double-stranded circular DNA genome of approximately 130 kb that contains more than 150 open reading frames. The ability of AcMNPV to infect insect cells has led to its use in multiple protein expression systems and as plant insecticides.

Grosse quantità di **poliedrina**, una proteina della matrice, che il virus produce grazie ad un promotore molto forte, sono ingerite dall'insetto



# Use of baculovirus for expression of recombinant protein

1. I baculovirus (AcMNPV) usati per infettare le cellule d'insetto: Il **promotore della poliedrina** è usato per l'**espressione del gene esogeno**
3. DNA virale è grande (130kb): il gene da esprimere deve essere inserito nel gene della poliedrina usando un vettore di trasferimento più piccolo (contenente anche un ori per la sua manipolazione in E. coli)
4. **Ricombinazione** del vettore ("donor plasmid") di trasferimento col DNA virale modificato. (contiene sequenze BAC = Bacmid). Vettore di trasferimento contiene di un gene marcatore, come la  $\beta$ -galattosidasi nel vettore di trasferimento per selezionare i ricombinanti (placche blu), sostituito spesso con la GFP (Green Fluorescent Protein), che consente di identificare più facilmente i ricombinanti illuminandoli con una lunghezza d'onda appropriata (emissione 509 nm).
5. Dopo ricombinazione il vettore si comporta come un BAC  $\rightarrow$  preparazione di DNA Bacmid da E.coli possibile



Commonly used cell lines are immortalized sf9 & sf21 derived from the pupal ovarian tissue of the fall army worm *spodoptera frugiperda* and high five (BTI-Tn-5B1-4 cells) cells derived from the ovarian cells of the cabbage looper



# Use of baculovirus for expression of recombinant protein

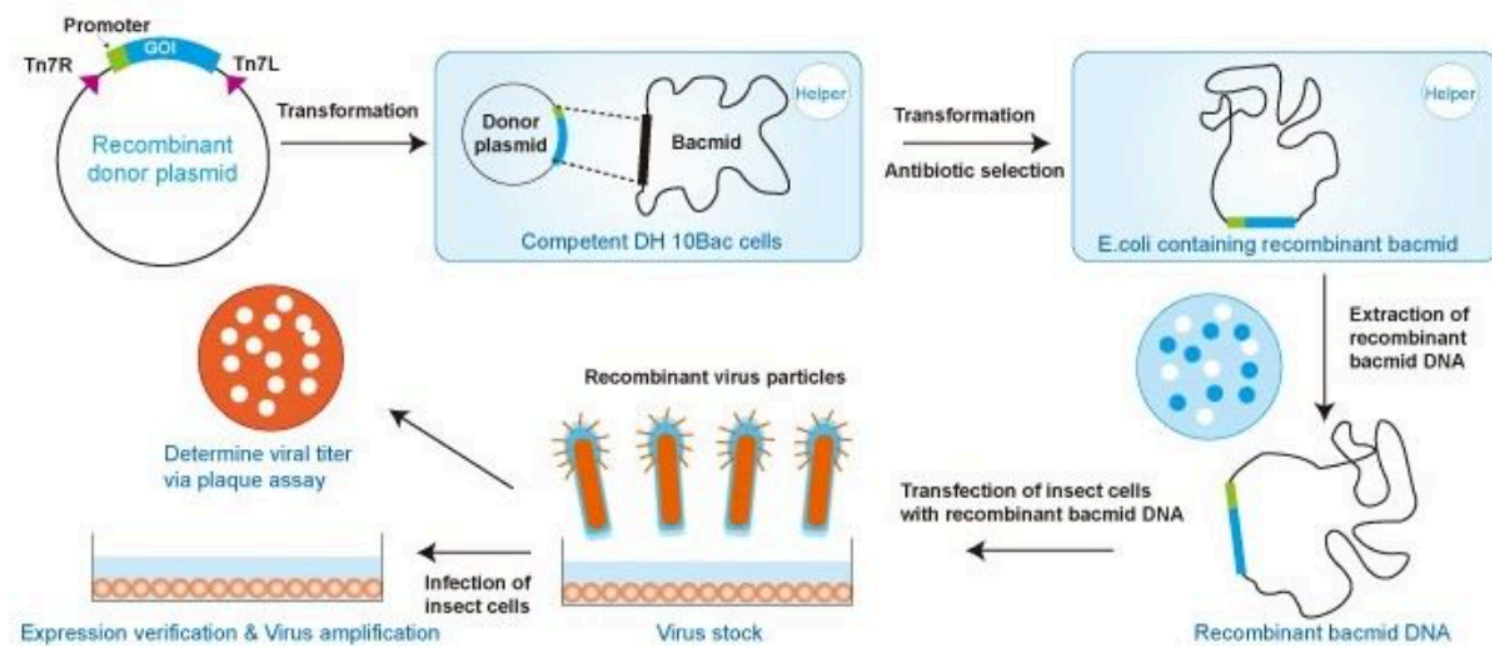


Figure 1. Baculovirus-insect cell expression system

6. I virus ricombinanti non produrranno poliedrina, pero producono il capside normale
  7. Col virus ricombinante viene effettuata l'infezione su larga scala utilizzando cellule sf9 (o altre)
  8. Alti livelli di proteina prima della lisi delle cellule dell'ospite: la proteina eterologa (prodotta all'interno della cellula, oppure secrete nel mezzo di coltura: proteine di fusione (**His6**, **GST**, fusion proteins or other systems) viene raccolta dopo 4-5 giorni dall'infezione
- NOTE: Coltura dell'ospite più costosa rispetto al lievito, ma possibile l'inserimento delle modificazioni post-traduzionali e la co-espressione (nel baculovirus) di geni che codificano per enzimi di glicosilazione necessari per la corretta modificazione della proteina

# Use of baculovirus for expression of recombinant protein

## PROs:

- The polyhedrin gene is **not required** for the continuous production of infectious virus in insect cell culture. Its sequence is replaced with that of the heterologous gene.
- The polyhedrin gene promoter is very **strong**. This determines a very high level of production of recombinant protein.
- Very late expression allows for the production of very toxic proteins.
- **This system is capable of post-translational modifications**

## CONS:

- Expensive.
- Glycosylation in insect cells is different (insect cells unable to produce complex N-linked side chains with penultimate galactose and terminal sialic acid) from that in vertebrate cells, therefore, a problem for therapeutic proteins.
- A large fraction of the RP can be poorly processed and accumulates as aggregates.
- Discontinuous expression: baculovirus infection of insect cells kills the host and hence the need to reinfect fresh cultures for each round of protein synthesis.
- Inefficient for production on a commercial scale

Adenosina deaminasi  
Anticorpi monoclonali del topo  
Antigene del capsido del rotavirus della scimmia  
Antigene del carbonchio  
Antigene del virus sinciziale respiratorio  
Antigene tipo 1 del virus dengue  
Attivatore tissutale del plasminogeno  
Antigene di neutralizzazione del virus della malattia della lingua blu

Emagglutinina del virus influenzale  
Eritropoietina  
Fosfatasi alcalina umana  
Glicoproteina 50 del virus della pseudorabbia  
Glicoproteina del virus della rabbia  
Interferon  $\alpha$   
Interferon  $\beta$   
Lipasi pancreatica umana  
Proteina del Lassavirus

Proteina dell'involucro di HIV-1  
Proteina precursore dell'amiloide  $\beta$   
Proteina trasportatrice multifarmaco  
Proteina del capsido di HSV  
Proteine del poliovirus  
Proteine della malaria  
Recettori accoppiati con la proteina G  
Regolatore della conduttanza  
Transmembrana della fibrosi cistica  
Rodopsina bovina

## 2. Vectors express and purify recombinant proteins

- In bacteria
- In yeast
- In fly cells – insect cells
- **In mammalian cells**

# Introduction of vectors into mammalian cells

## Expression systems

### Transient

High expression over limited time (2-3 days)  
Plasmid is not segregated by mechanism and will get lost in a few days

- Electroporation
- Lipofection
- Calciumphosphate

### Stable

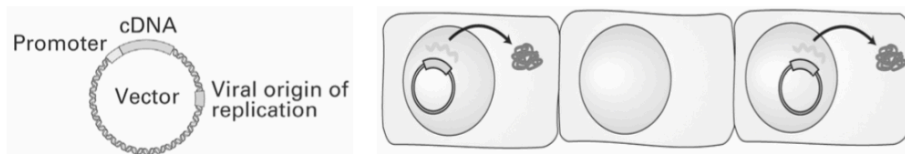
Continuous expression; requires stable integration of vector into genomic DNA; vector must contain a selectable marker

- Electroporation
- Lipofection
- Calciumphosphate
- Viral transduction (stable and episomal)

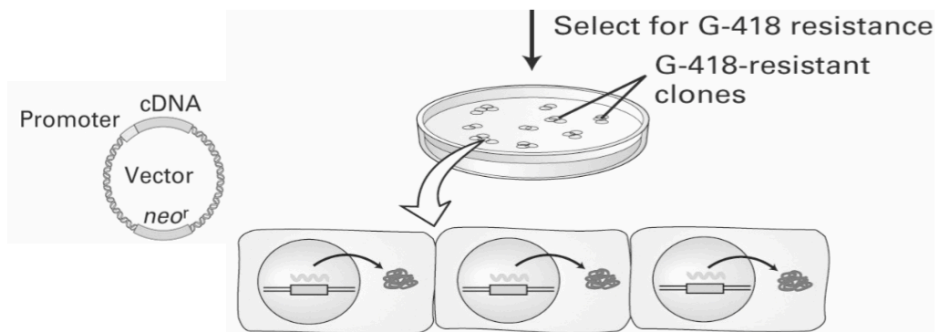
After introduction of vector → apply selection for genome integrants

Alternative SV40 system allows episomal maintenance of plasmid

#### Transient transfection

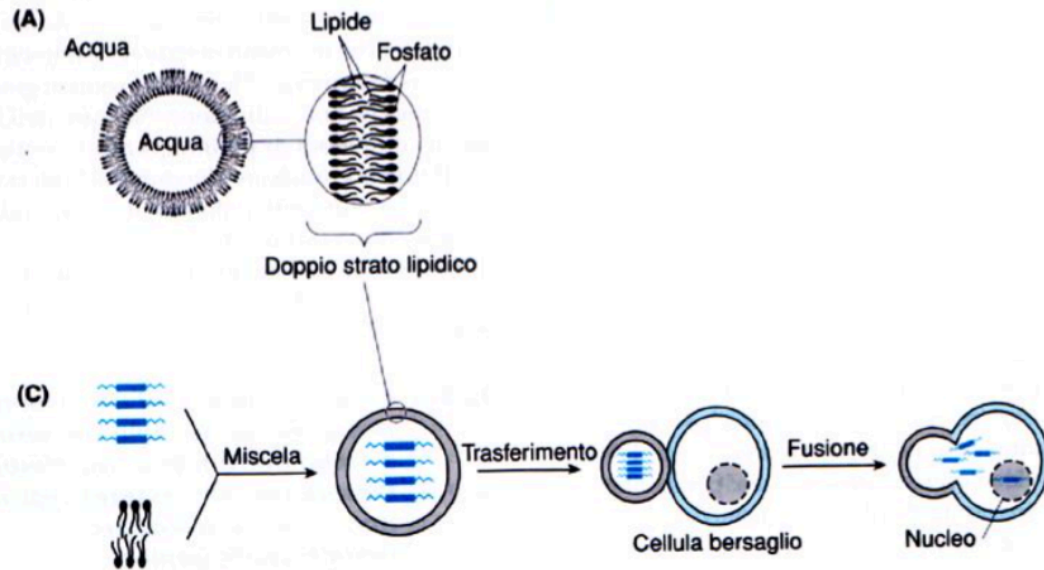


#### Stable transfection

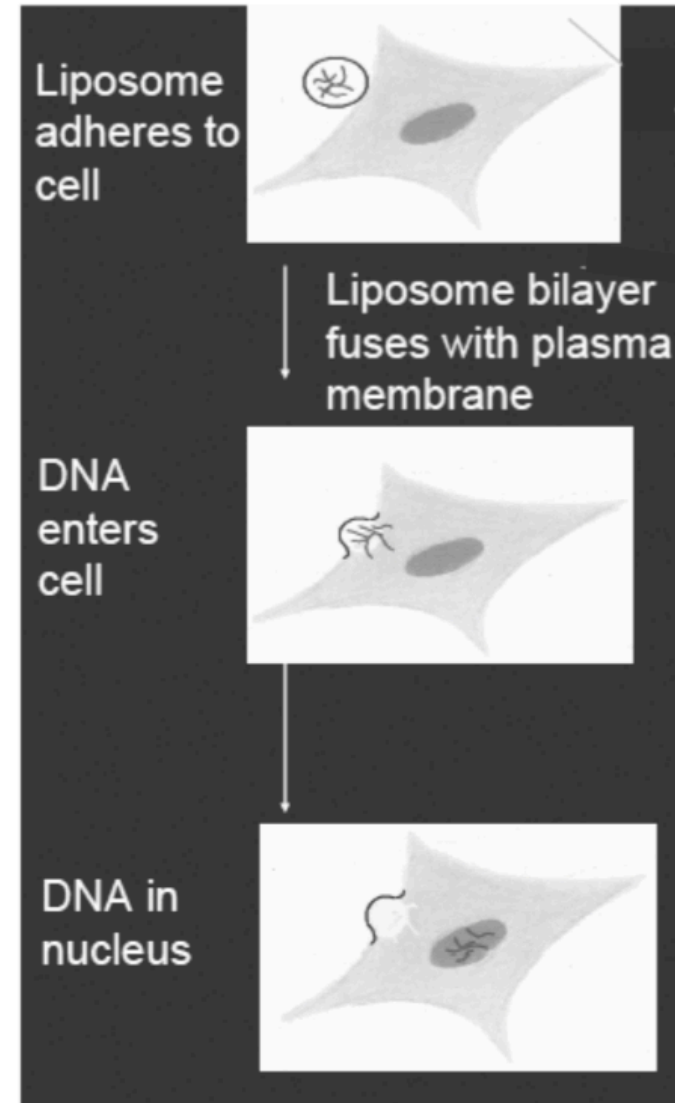


- INCUBAZIONE CON DNA CO-PRECIPITATO CON FOSFATO DI CALCIO O DIETILAMINOETIL-DESTRANO (DEA)
- ELETTROPORAZIONE
- LIPOSOMI
- VEICOLAZIONE DI DNA DA PARTE DI UN VIRUS (Retrovirus, Lentivirus)

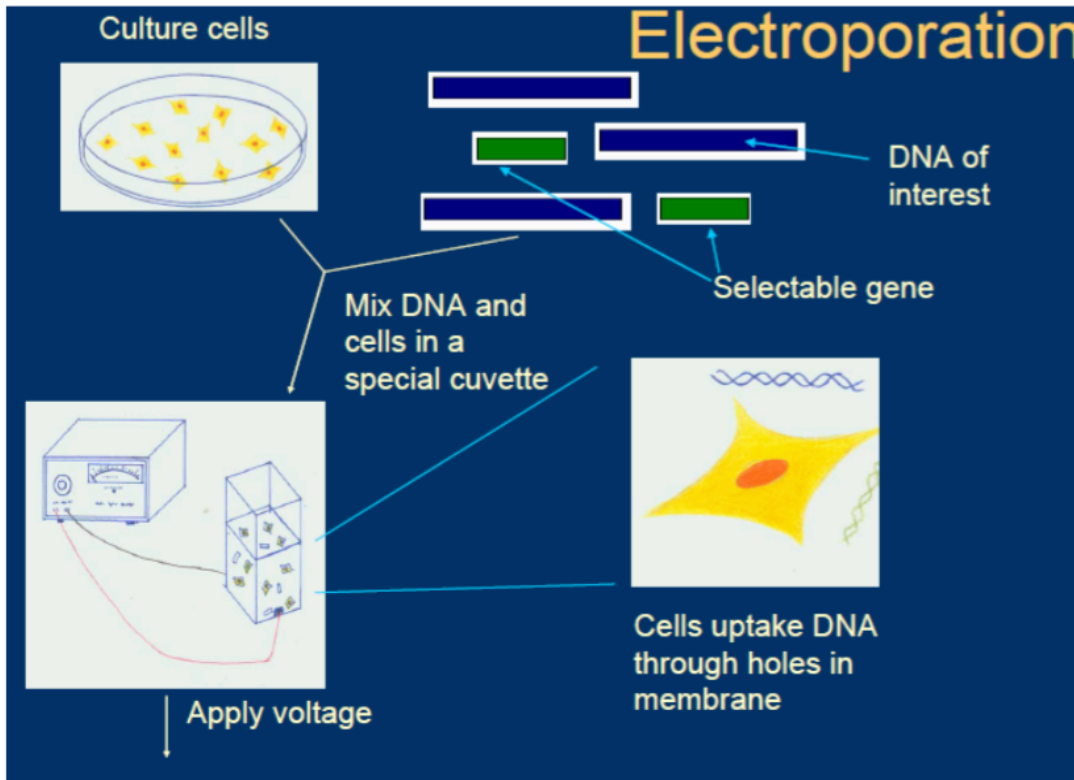
# LIPOFECTION OF MAMMALIAN CELLS



Lipids form spheres and integrate DNA  
Lipids fuse with nuclear membrane, DNA enters cell  
DNA (vector) ready for transcription

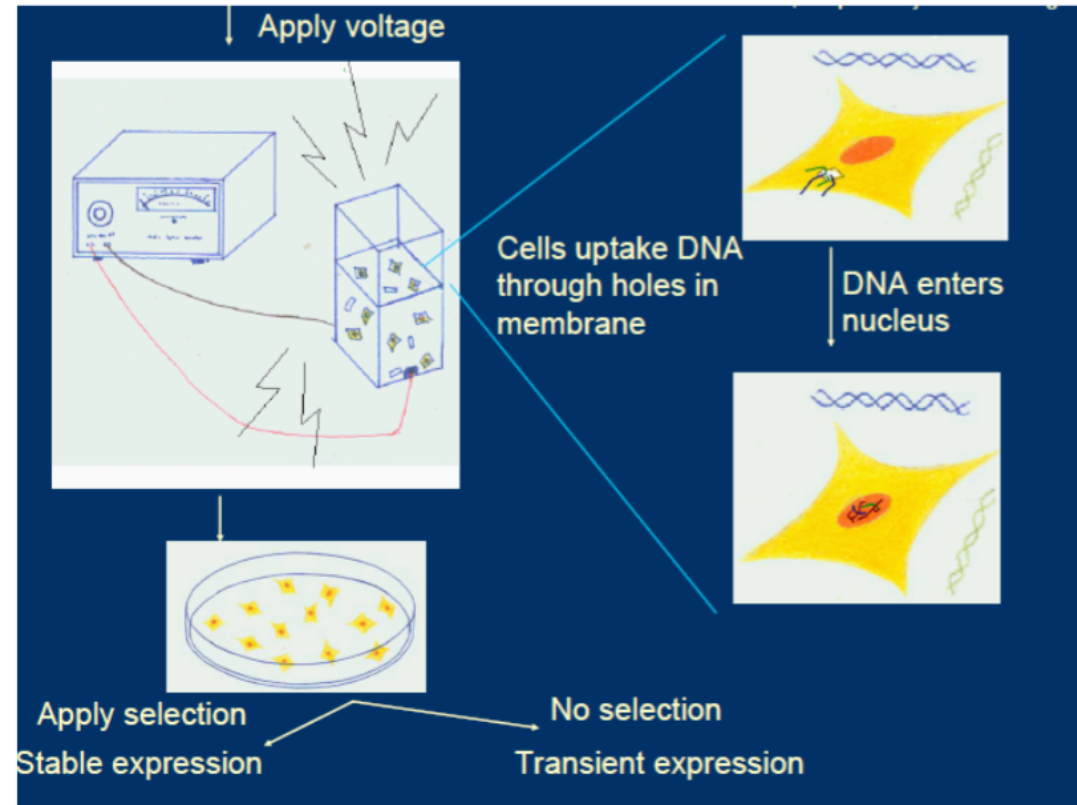


# ELECTROPORATION OF MAMMALIAN CELLS



DNA of interest: linear or circular

Application of current cause transient perforation of cell membrane and DNA can enter cell.

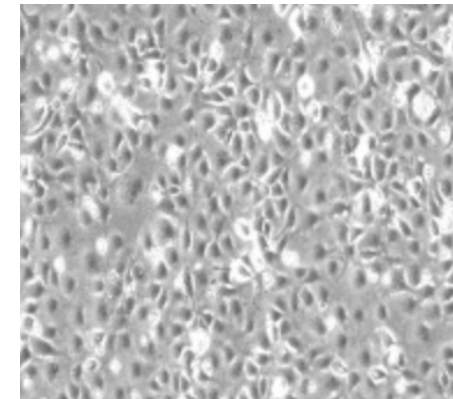
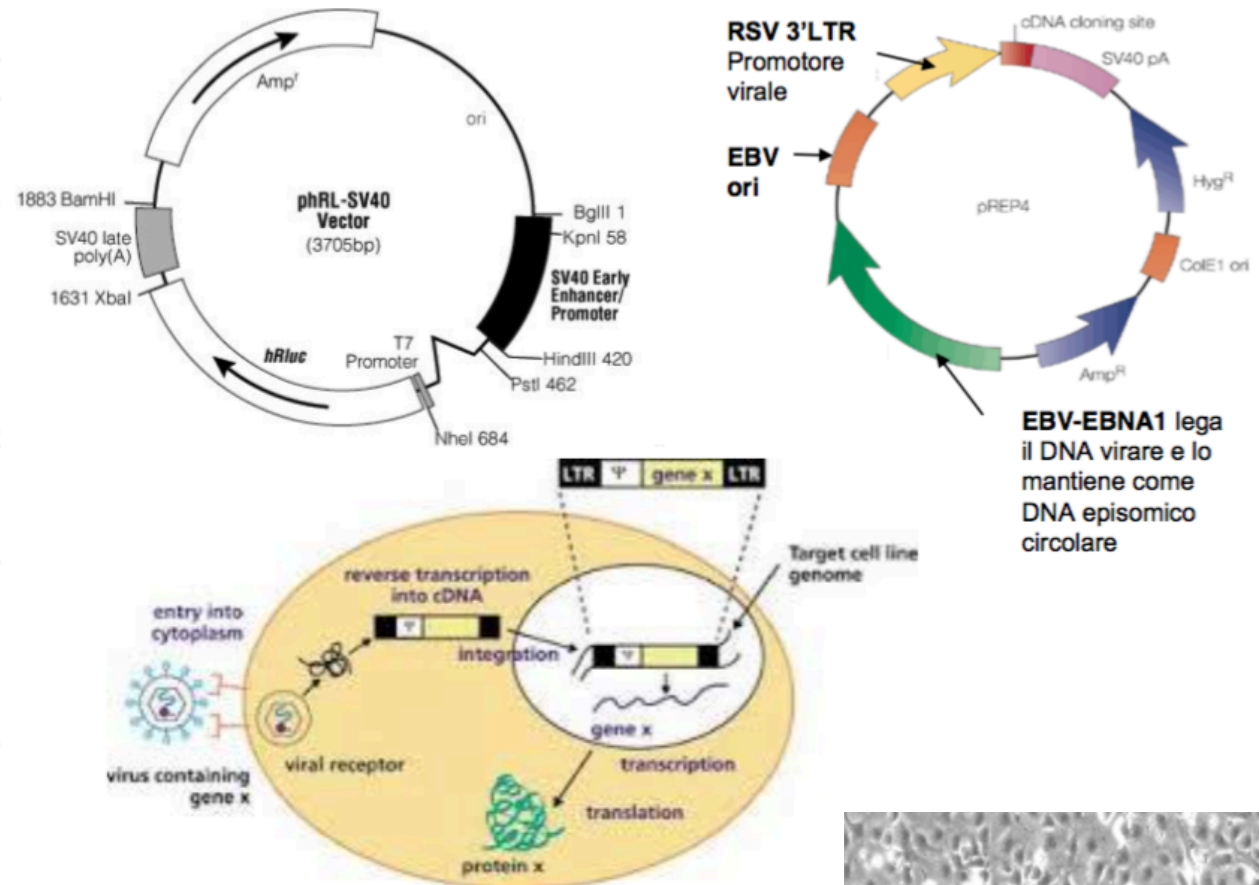


# Vettori d'espressione di proteine in cellule di mammifero

La produzione di proteine animali in ospiti artificiali, come batteri, lieviti, insetti presenta sempre il **problema potenziale del non corretto processamento dell'mRNA e la produzione di una proteina inattiva o parzialmente attiva**. La resa della proteina ricombinante può essere migliorata inserendo un introne del gene clonato, perché il cDNA può contenere siti di splicing criptici, mentre la presenza dell'introne rende meno probabile lo splicing alternativo sui siti criptici

1. **Nelle cellule di mammifero un plasmide non si replica autonomamente: è possibile con vettori aventi l'origine di replicazione del virus SV40** (simian virus 40, causante negli animali immunodeficienza cronica, lesioni a livello renale e cerebrale) **che si possono replicare in maniera episomale nelle cellule COS** (linea cellulare di fibroblasti renali di scimmia). Avrà vita limitata, transiente, si diluirà man mano che le cellule si dividono, ma è tecnologicamente più facile (più rapida la produzione del clone)
2. **Cloni stabili in forma extracromosomica: contengono l'origine di replicazione del virus Eptein-Barr (EBV)**
3. Inserimento del DNA esogeno in un cromosoma delle cellule di mammifero mediante **Retrovirus**. **Integrazione nel DNA nucleare, si duplica quando la cellula si divide, integrazione stabile**

Per esprimere il gene clonato sono necessari **promotori forti**: derivati da virus animali o da geni di mammifero altamente espressi **SV40**, **citomegalovirus (CMV)**, **herpes simplex virus (HSV)**; oppure **actina  $\beta$** , **timidina chinasi**, **ormone della crescita bovino**



# Expression vectors based on Simian virus 40 (SV40 virus)

1. Nelle cellule di mammifero un plasmide non si replica autonomamente: è possibile con vettori aventi l'origine di replicazione del virus SV40 (simian virus 40, causante negli animali immunodeficienza cronica, lesioni a livello renale e cerebrale) **che si possono replicare in maniera episomale nelle cellule COS** (linea cellulare di fibroblasti renali di scimmia). Avrà vita limitata, transiente, si diluirà man mano che le cellule si dividono, ma è tecnologicamente più facile (più rapida la produzione del clone)

COS cells are fibroblast-like cell lines derived from monkey kidney tissue. COS cells are obtained by immortalizing CV-1 cells with a version of the SV40 virus that can produce large T antigen but has a defect in virus genome replication (no virus particle production possible). SV40 recruits DNA pol alpha to SV40 origin → initiation of replication

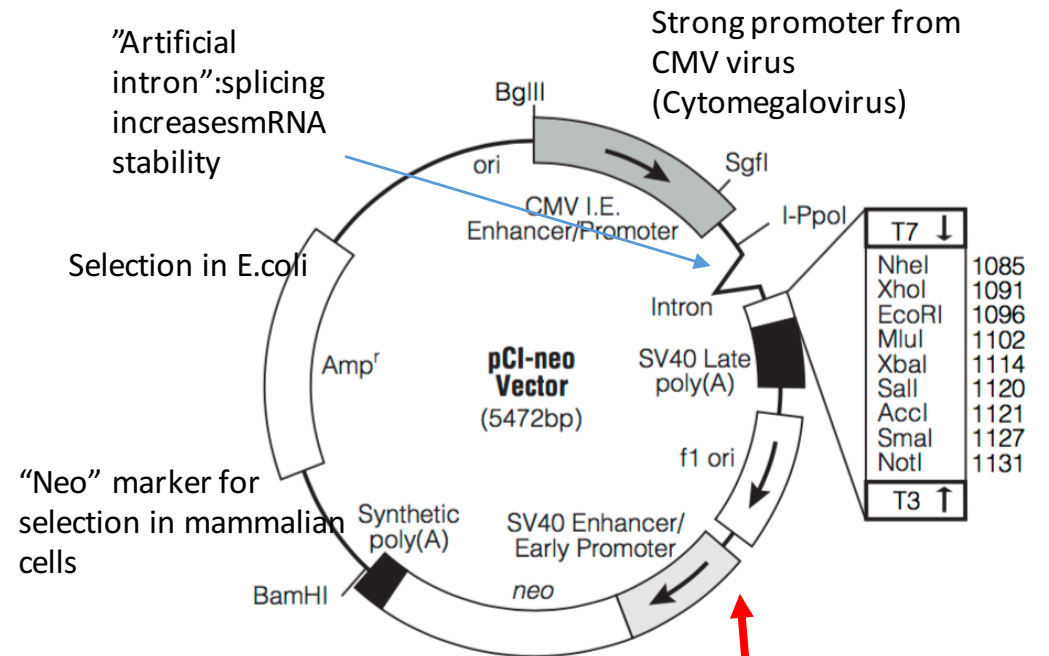
When an expression construct with an SV40 enhancer and early promoter region (contains SV40 origin) is introduced into COS cells, the vector can be replicated substantially by the large T antigen.

## Resistance markers in mammalian cell lines:

-**G418** blocks polypeptide synthesis by inhibiting the elongation step in both prokaryotic and eukaryotic cells

-**Puromycin** is an aminonucleoside antibiotic, derived from the *Streptomyces alboniger* bacterium, that causes premature chain termination during translation taking place in the ribosome.

-**Hygromycin** is active against both prokaryotic and eukaryotic cells. It acts by inhibiting polypeptide synthesis. It stabilizes the tRNA-ribosomal acceptor site, thereby inhibiting translation.

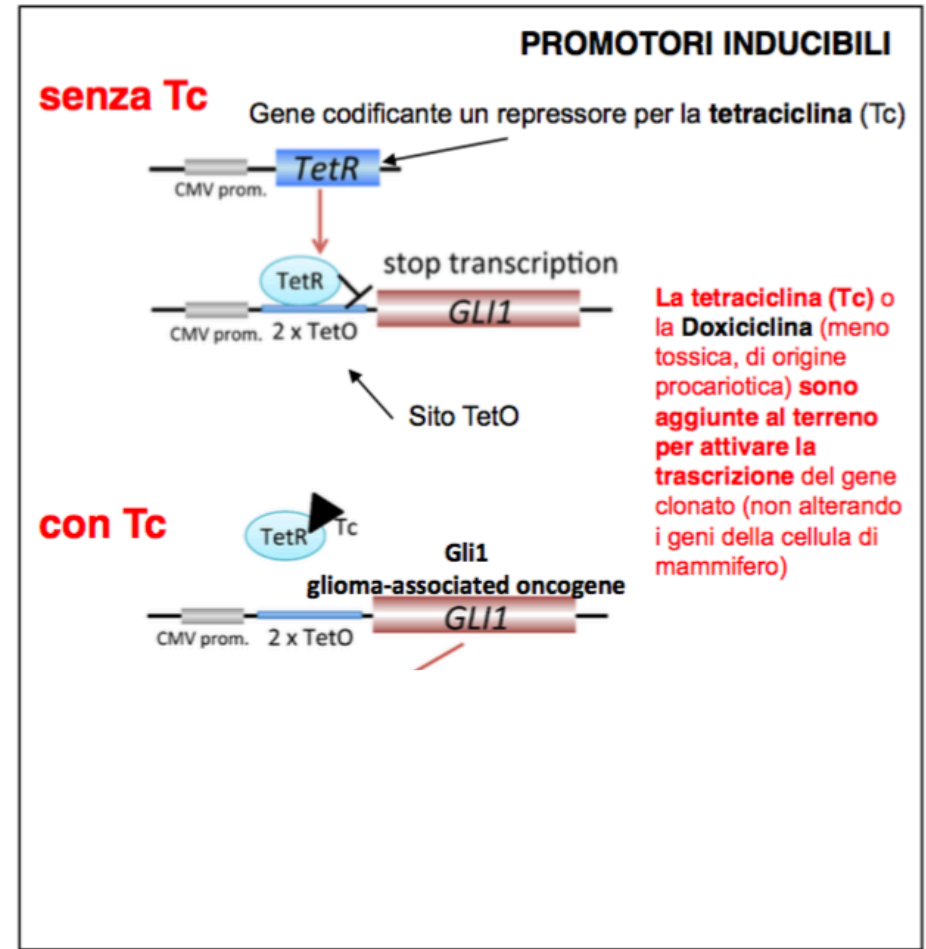
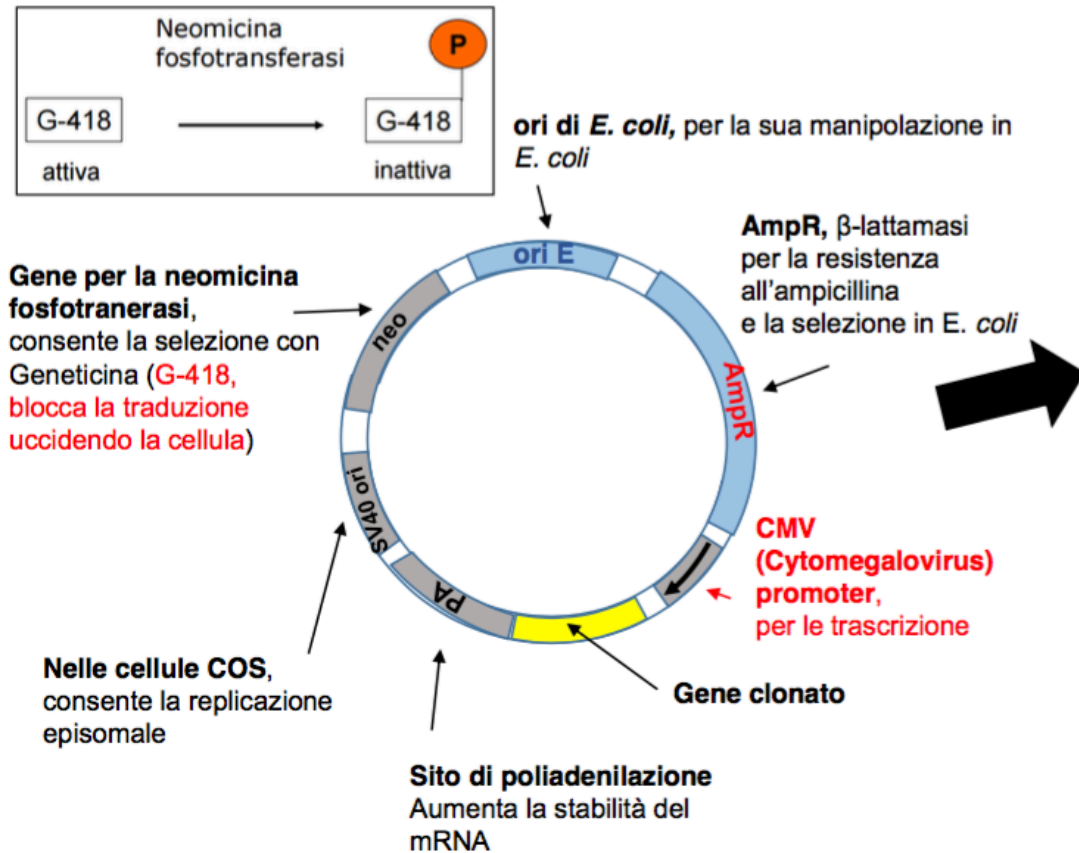


## pCI-neo Sequence Reference Points:

CMV immediate-early enhancer/promoter region	1-750
Chimeric intron	890-1022
T7-EEV sequencing primer binding region	1053-1074
T7 RNA polymerase promoter (-17 to +2)	1067-1085
Multiple cloning region	1085-1137
T3 RNA polymerase promoter (-17 to +3)	1140-1158
SV40 late polyadenylation signal	1167-1388
Phage f1 region	1483-1938
SV40 enhancer and early promoter	2000-2418
SV40 minimum origin of replication	2316-2381
Coding region of neomycin phosphotransferase	2463-3257
Synthetic polyadenylation signal	3321-3369
β-lactamase (Amp <sup>r</sup> ) coding region	3780-4640



# Vectors based on simian virus 40 (SV40 virus) with inducible gene expression system (Tet repressor system)



## Note:

Tetracycline (tc) is a broad family of antibiotics to which bacteria have evolved resistance. Tc normally kills bacteria by binding to the bacterial ribosome and halting protein synthesis. The expression of tc resistance genes is regulated by Tet Repressor Protein. More specifically, TetR represses the expression of TetA, a membrane protein that pumps out substances toxic to the bacteria like tc

# Vettori d'espressione in cellule eucariotiche: aggiunta di marcatori e segnali

1. **Addizione di sequenze per la localizzazione** (nucleo, citoplasma, mitocondri, reticolo endoplasmatico)

2. **Aggiunta di segnali per la secrezione, come un peptide segnale**

- riconosciuto dal meccanismo di secrezione, trasportato attraverso la membrana. Utile se la proteina clonata è anch'essa secreta, poco efficace se è citoplasmatica
- Stabilizza il prodotto del gene clonato proteggendolo dalle proteasi
- Evita l'accumulo della proteina nel citoplasma e la formazione dei corpi di inclusione insolubili.
- Facilitano la purificazione
- Possono essere utilizzate come antigene per produrre anticorpi
- Consentono lo *screening* di genoteche di cDNA

3. **Aggiunta di aminoacidi al N-terminale o C-terminale, come marcatore (*tag*, utile per la purificazione):**

- **Epitopo**, peptide amminoacidico riconosciuto da un anticorpo monoclonale (purificazione per affinità)
- **Coda amminoacidica**, es. con residui di istidina

4. **Per eliminare il peptide tag dalla proteina clonata: si inseriscono tra le due sequenze (peptide e proteina) siti di taglio per proteasi non batteriche specifiche**

- Endopeptidasi
- Trombina
- Fattore Xa



# Comparison of expression systems

Characteristics	<i>E. coli</i>	Yeast	Insect cells	Mammalian cells
Cell growth	rapid (30 min)	rapid (90 min)	slow (18-24 h)	slow (24 h)
Complexity of growth medium	minimum	minimum	complex	complex
Cost of growth medium	low	low	high	high
Expression level	high	low-high	low-high	low-moderate
Extracellular expression	secretion to periplasm	secretion to medium	secretion to medium	secretion to medium
<i>Posttranslational modifications</i>				
Protein folding	refolding usually required	refolding may be required	proper folding	proper folding
N-linked glycosylation	none	high mannose	simple, no sialic acid	complex
O-linked glycosylation	no	yes	yes	yes
Phosphorylation	no	yes	yes	yes
Acetylation	no	yes	yes	yes
Acylation	no	yes	yes	yes
gamma-Carboxylation	no	no	no	yes