BIOC4004 - Industrial Biochemistry
Lecture 08 – Mon Jan 26, 04

Topics for the Day:

- heterologous expression systems
  - bacterial
  - Eukaryotic
    - *Pichia pastoris*
    - *BEVS*
- Mammalian expression systems

Why Heterologous Expression of Recombinant Proteins ?

In order to work on a protein, you must isolate it and purify it

Before recombinant DNA

- isolation and purification from original sources
- find a source tissue that is enriched for the protein of interest !!!!

- **Advantages:**
  - working with the real deal !!!!
- **Disadvantages:**
  - need high amounts of starting material
  - low yield in most cases
  - purification of individual protein from a myriad of other proteins!!!

After recombinant DNA:

- cloning of the gene, expression of the gene in heterologous system

- **Advantages:**
  - don't need to have continuous source of original source !!!!
  - higher "starting purity"
    - over-expression: higher proportion of total protein content
    - protein secretion: less contamination from intracellular proteins
  - use protein engineering to facilitate purification
    - purification tags
    - protein secretion
- **Disadvantages:**
  - protein may have incorrect structure
    - folding
    - post-translational modifications

- **Protein chemistry is tough enough**
- **Use rDNA tech to solve some of the problems of protein purification**
  - **ultimately, it's all about protein purification**

Components of a Heterologous Expression System

The Host ie. who's gonna express your gene ?

- bacterial cells
- unicellular eukaryotes
- insect cells
- mammalian cells
- plants
- animals

The Vector ie. how are you gonna ensure gene expression in the host?

- plasmids
- bacteriophages
- viruses
    - strong promoter
    - good transcriptional regulatory sequences
    - all of the signals to ensure good translation
        - Ribosomal Binding Site
        - in-frame cloning of the coding sequence
        - translation termination

The "DNA transfer Method" ie. how are you gonna get the DNA into host cells ?

- chemical transformation
- viral transduction
- protoplast fusion
- liposomes
- electroporation
- microprojectile
- microinjection

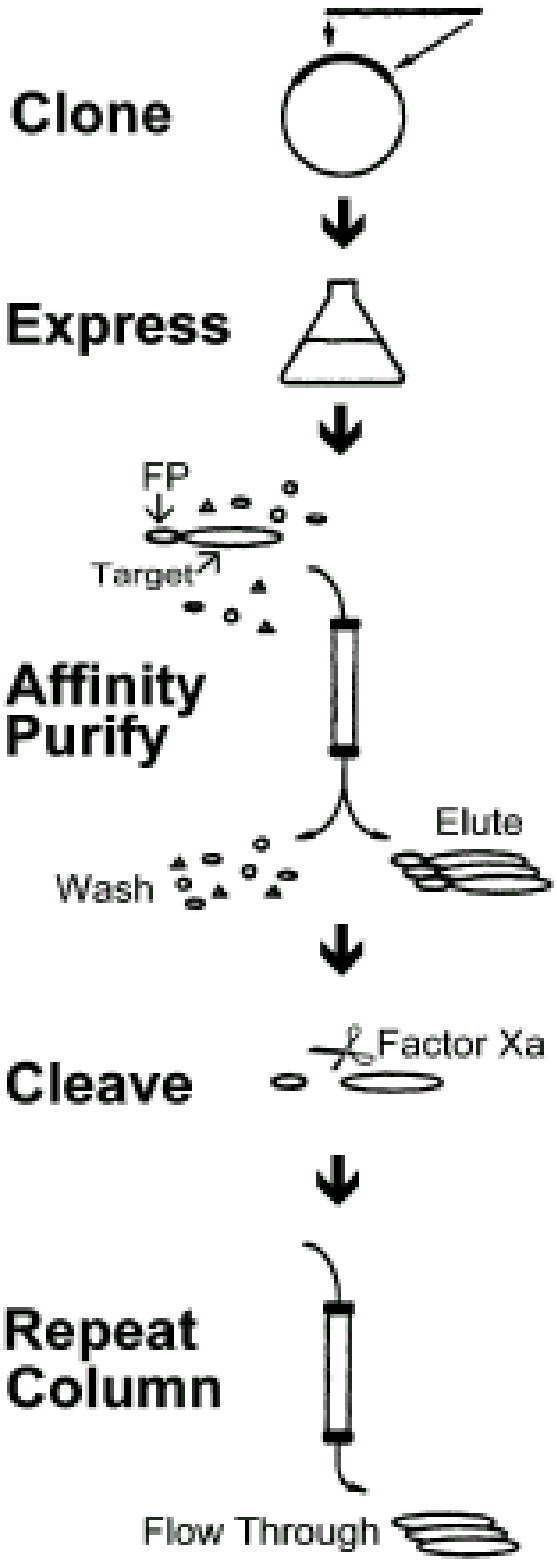
The Production Method ie. how are you gonna mass produce the protein ?

- shake cultures
- fermentors
- bioreactors
- the plant/animal

The Purification System ie. how are you gonna get pure protein ?

- Protein chemistry approaches
- rDNA approaches

Super-easy Heterologous Expression in *E. coli*



• The gene of interest is cloned into the expression vector

• Transformed *E. coli* are grown in large-scale liquid cultures and after reaching the appropriate density are induced to express the recombinant protein.

• Cells are harvested, lysed and the lysate is passed over an affinity column

• The recombinant fusion protein is retained by the column and other *E. coli* proteins are removed. The fusion protein is eluted with free ligand or by other means.

• Many expression vectors have a specific protease site (eg.,factor Xa) engineered between the fusion partner and the protein of interest. Treatment with the specific protease may result in separation of the two proteins.

• The fusion partner can be separated from the protein of interest by repeating the affinity chromatography since the fusion partner will be retained by the column.

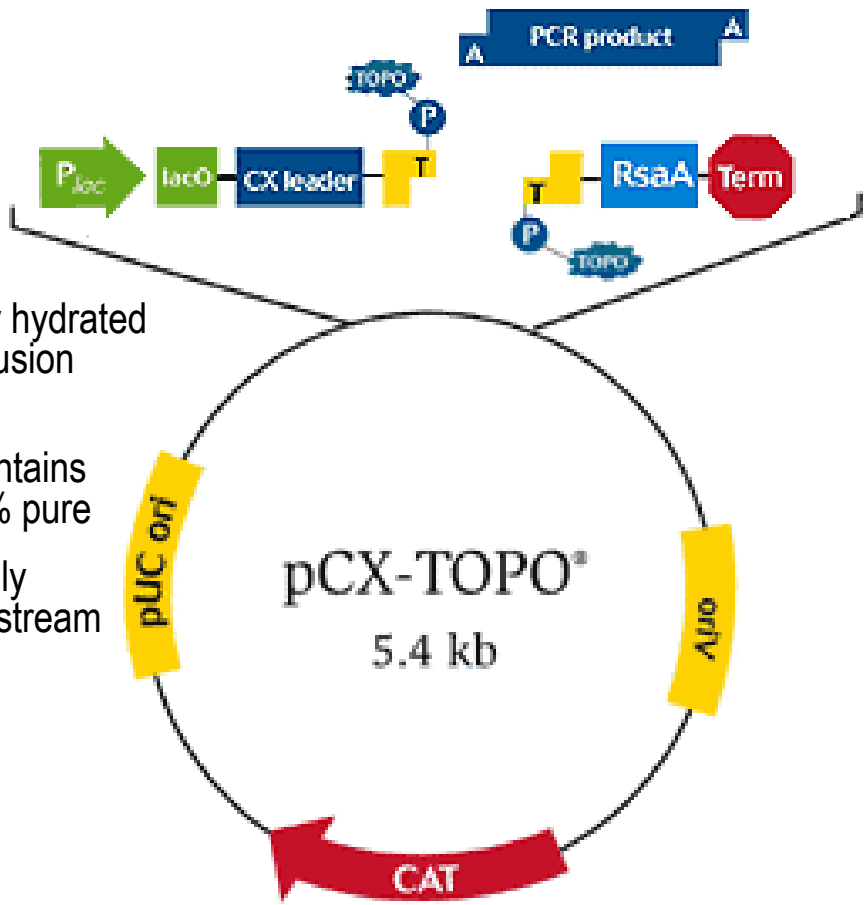
More Specialized Bacterial Systems: *Caulobacter*

The PurePro™ *Caulobacter* Expression System offers the following two advantages over *E. coli* expression systems:

- Proteins are secreted into the medium for simplified downstream processing
- No expensive affinity columns or lengthy separation steps are needed for protein purification

How does it work ?

- *Caulobacter* secretes large amounts of a hydrophilic structural protein, RsaA
- Invitrogen's PurePro™ *Caulobacter* Expression System takes advantage of the efficient RsaA secretion signal to direct the secretion of the recombinant protein into the medium



- RsaA fusions form a highly hydrated aggregates of the secreted fusion protein.
- The aggregate typically contains protein that is more than 90% pure
- The aggregate can be easily solublized and used in downstream applications

E. coli and Caulobacter are great but….they're still prokaryotes

Problems with expression of eukaryotic proteins in prokaryotic expression systems:

- genetic stability of the coding sequence
- protein stability once expressed
- potential problem with correct folding
- post translational modification machinery is very different
  - no Golgi apparatus
  - no Endoplasmic Reticulum
  - ie. no glycosylation and few post-translational mods
- potential toxicity because of endotoxin (LPS)

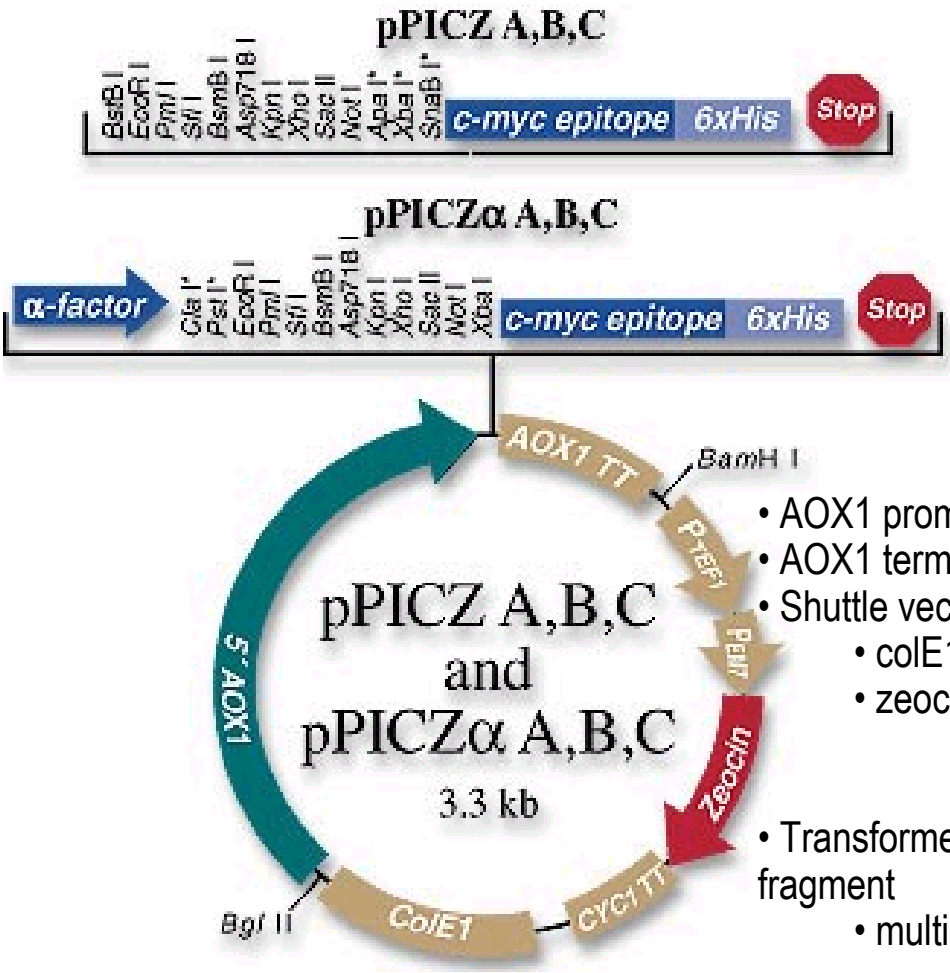
Eukaryotic Expression Systems

- in theory, plasmids can be introduced into any host
- yeast are easy to maintain in lab
  - *Saccharomyces cerevisiae*
  - *Pichia pastoris*
- viruses
  - several mammalian
  - baculovirus (insect)

vaccinia (lytic)
adenovirus (lytic)
papilloma (episomal)
retrovirus (integrated)

Pichia pastoris yeast as a heterologous expression system

- grows on methanol(cheap food !!!)
- fast growth
- no endotoxin
- very strong alcohol oxidase (AOX1) promoter; methanol induced
- typically has good folding and pretty accurate glycosylation
    - Brewer's yeast has atypical glycosylation
- most proteins are secreted into the medium
    - super high levels !!!! (unlike brewer's yeast)
    - Up to 30% of the protein content can be your engineered protein
- examples: human serum albumin, human lysozyme, tetanus toxin fragment

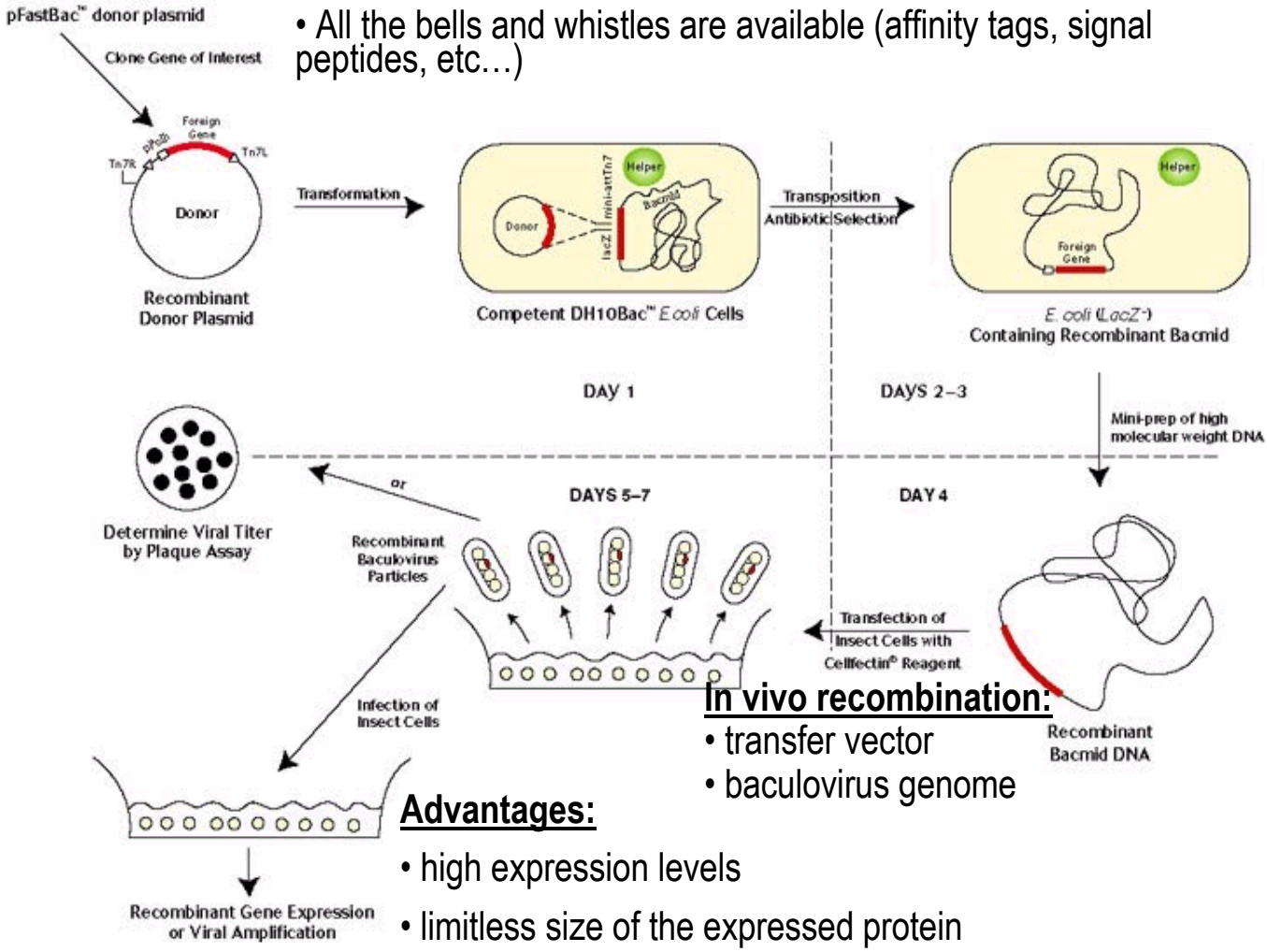


- AOX1 promoter (good expression)
- AOX1 terminator (good processing)
- Shuttle vector:
    - colE1 *E.coli* origin
    - zeocin resistance
        - *E. coli*
        - *Pichia*
- Transformed as linear *Bgl*II *Bam*HI fragment
    - multimer integration possible

Baculovirus Expression Vector System(BEVS)

Based on a family of insect viruses

- *Autographa californica* nuclear polyhedrosis virus (AcMNPV)
- The insect cells are either Sf9 cells or Sf21 cells
  - from *Spodoptera frugiperda* (fall army worm) ovary cells
- Gene expression driven by Polyhedrin promoter (**super** strong !!!!)
- All the bells and whistles are available (affinity tags, signal peptides, etc…)



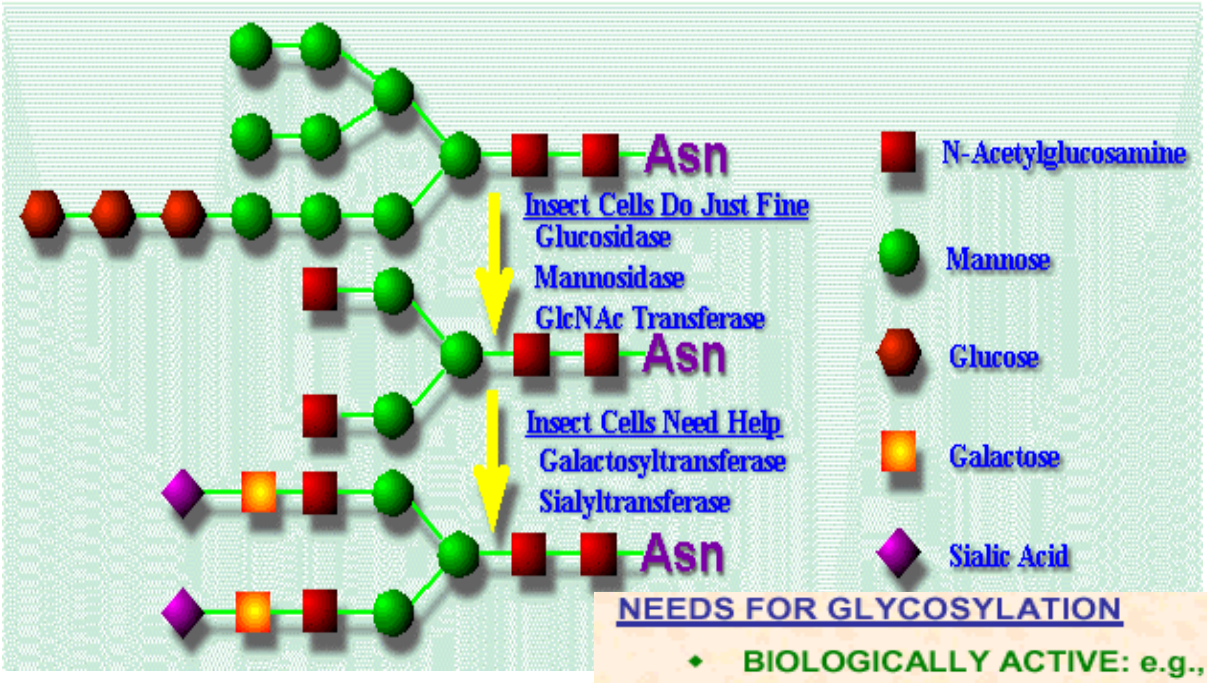
In vivo recombination:

- transfer vector
- baculovirus genome

Advantages:

- high expression levels
- limitless size of the expressed protein
- efficient cleavage of signal peptides and processing
- post-translational modifications
- simultaneous expression of multiple genes
- the cells die during the manufacturing process, the BEVS system is able to produce proteins from genes of unknown function.

Baculovirus and post-translational mods.



NEEDS FOR GLYCOSYLATION

- BIOLOGICALLY ACTIVE: e.g., EPO (NON-GLYCOSYLATED, INACTIVE)
- LONGER HALF-LIFE (NON-GLYCOSYLATED: CLEARED IN BODY MUCH FASTER
- PROTECTION FROM PROTEOLYTIC CLEAVAGE *in vivo*
- HIGHER BIOLOGICAL ACTIVITY
  - LESS TOXICITY (e.g., BETA INTERFERON)
  - *E. coli* DERIVED vs. ANIMAL CELL DERIVED
  - GLYCOSYLATED B-IFN: 1000-FOLD INCREASED ACTIVITY

Other post-translational modifications

Post-translational modification:	***Present in Baculovirus?***
Acylation and phosphorylation	Sometimes
Cleavage of signal sequences removal of hormonal prosequences polyprotein cleavages	Reported
Protein targeting	Conserved
Internal proteolytic cleavages at arginine-rich or lysine rich sequences	Highly inefficient

Heterologous Expression Systems, A Review

	The Good	The Bad
Bacterial • *E. coli* • *B. subtilis* • *Caulobacter*	• Simple fermentation and transformation • High yield • Well studied expression systems • Many proteins are expressed in inclusion bodies	• Many proteins are expressed in inclusion bodies • No post-translational modifications • Improper folding of disulphide linked proteins
Yeasts • *S. cerevisiae* • *P. pastoris*	• Simple fermentation and transformation • High yield • Well studied expression systems • Some limited post-translational modifications • Powerful secretory pathways	• Improper folding of disulphide linked proteins • Hypoglycosylation
Insect Cell Baculovirus	• moderate yield • Most post-translational modifications • Correct folding of product	• More complex fermentation required • Need to produce and maintain recombinant baculovirus

- The choice of system will depend on the protein being worked on.
- Try a simpler system, if it fails, move on !!!!

Sometimes you just need to use the real deal !!!! : Expression in Mammalian Systems

- production of biologically functional proteins for human therapy
    - good folding
    - good post-translational modifications (especially **glycosylation** !!!!)

- **Problems:**
    - difficulty generating **stable** cell lines (need to integrate into genome!!!!)
        - viral vectors
    - slow growth (many days to weeks !!!!)
    - culture contamination(bacterial and viral)
        - antibiotics **can not** be used in industrial production !!!
    - cells **very** sensitive to most growth parameters
        - temperature
        - pH
        - oxygen
        - mechanical stresses
        - toxic byproducts
    - low cell densities
    - problems over-producing protein
    - apoptosis
    - some cell-types require attachment to solid medium

	The Good	The Bad
Mammalian	• All post-translational modifications • Correct folding of product • Correct glycosylation	• More complex fermentation required • low yields • increased production costs

There are **many** problems with mammalian expression, yet sometimes there is no other alternative !!!!

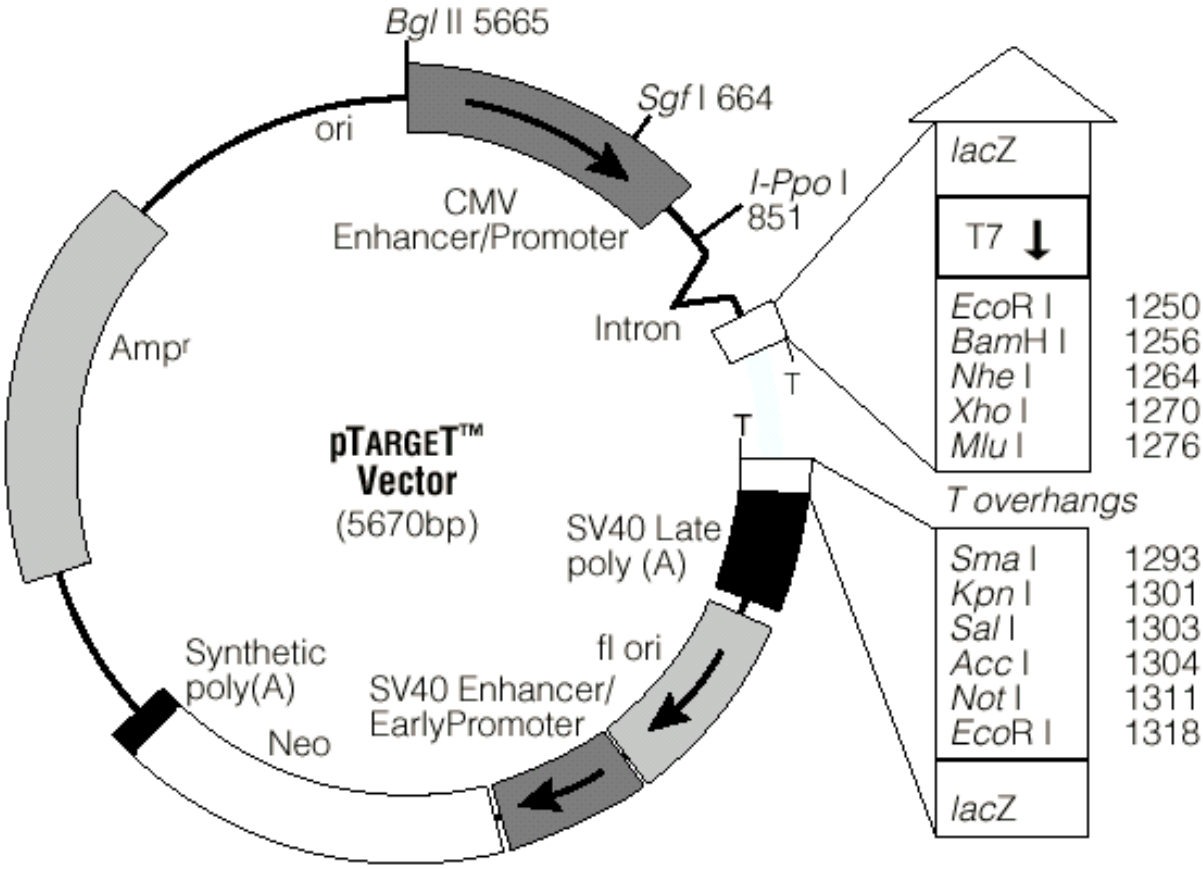
Industrial Recombinant Protein Production

- **e. coli**
  - Lilly rh insulin (1982): <$400/gm
  - bovine growth hormone (1994): <$11.60/gm
- **mammalian cells**
  - 21/33 US-approved products produced in mammalian cells
- **transgenic animals**
  - low speed to product, disease transmission…
- **transgenic plants**
  - glycosylation issues, purification
- **cell-free translation**

Made in Mammalian Cells

– DNASE:
  $500 MILLION,Dissolves DNA (lung); CF treatment
– Erythropoietin (EPO):
  $1.5 BILLION, Enhances RBC formation; Renal Dialysis
– FACTOR VIII:
  $800 MILLION: obviates need for plasma (AIDS/HIV); Hemophilia
– TISSUE PLAMINOGEN ACTIVATOR (TPA)
  $500 MILLION: Dissolves Blood Clots (heart attack)
– BETA INTERFERON:
  $650 MILLION (1996-1997): Multiple Sclerosis
– MONOCLONAL ANTIBODIES (HYBRIDOMA TECHNOLOGY):
  ~ $ 10 BILLION: Diagnostics Industry
  $??? : Untapped therapeutics potential

Mammalian Expression Vectors…



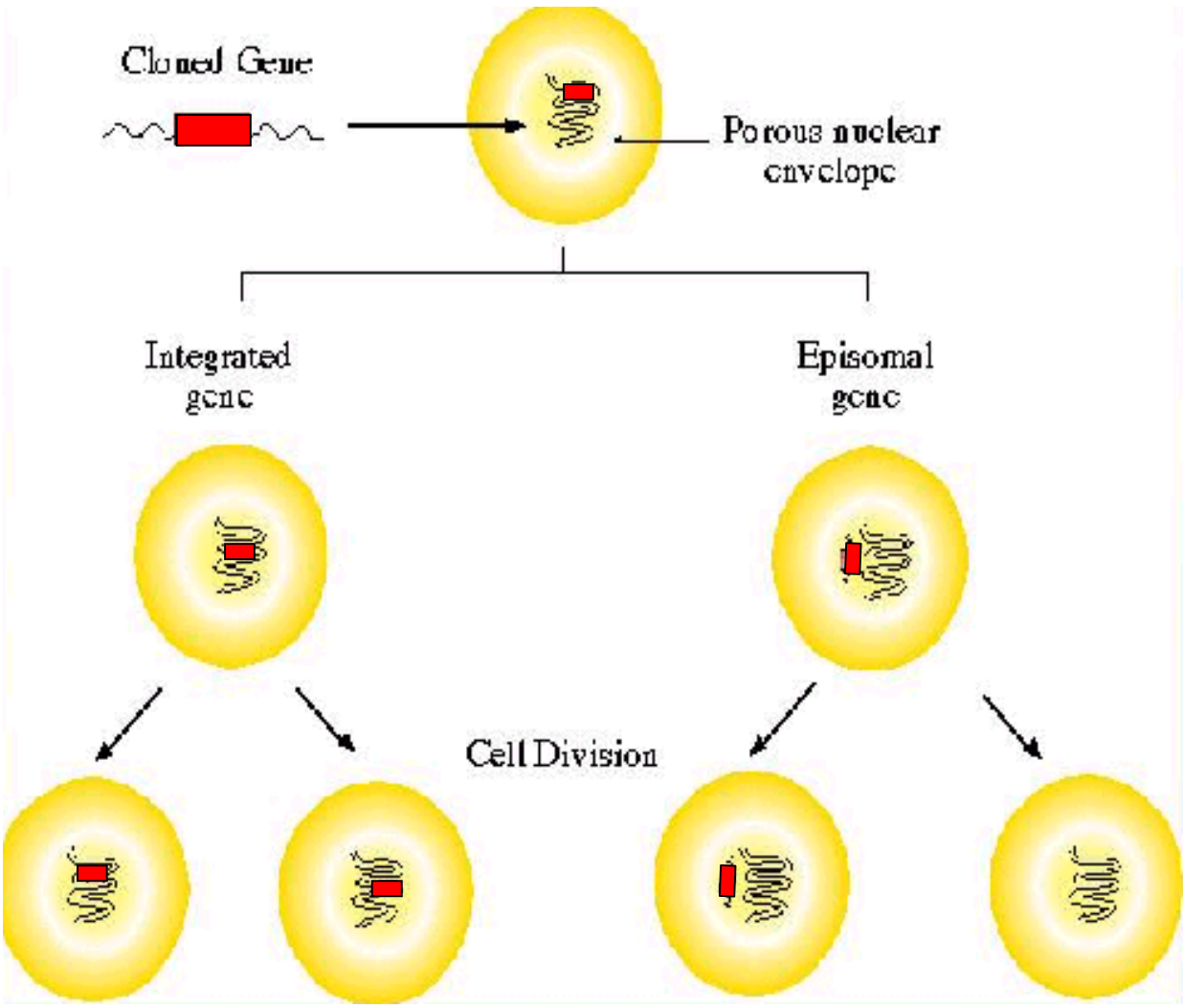
an example: pTarget (invitrogen)

- strong promoter (CMV in this case) with enhancer
- Amp^r for *E.coli*
- Neo for mammalian system
- Intron prevents splicing using cryptic splice sites in coding sequence
  - increases expression

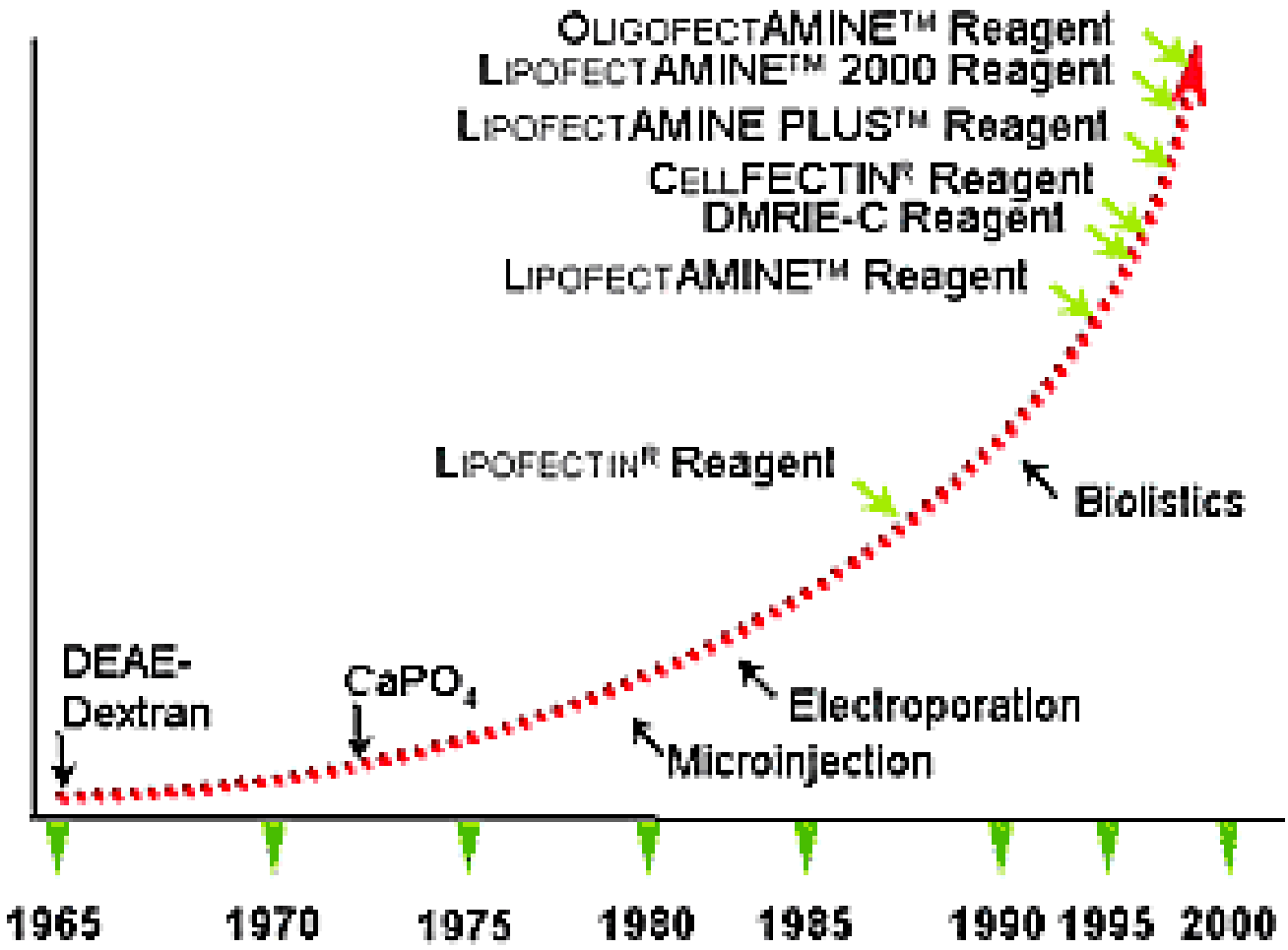
There are a lot of fancy new variations on the theme, but there is **nothing** inherently special about expression vectors for use with mammalian cells.

Transient vs. Stable Transfection

- Hard to maintain "episomal" DNA in mammalian cells
  - they're not used to it !!!!
  - Plasmid DNA will tend to get degraded quickly unless strong selection
    - eventually lost
- Stable expression can be obtained by random integration into genome
  - need integration into active chromatin !!!
  - Low probability event
- Viral vectors can be used but if cells die, you have to re-transfect every time you need more protein !!!



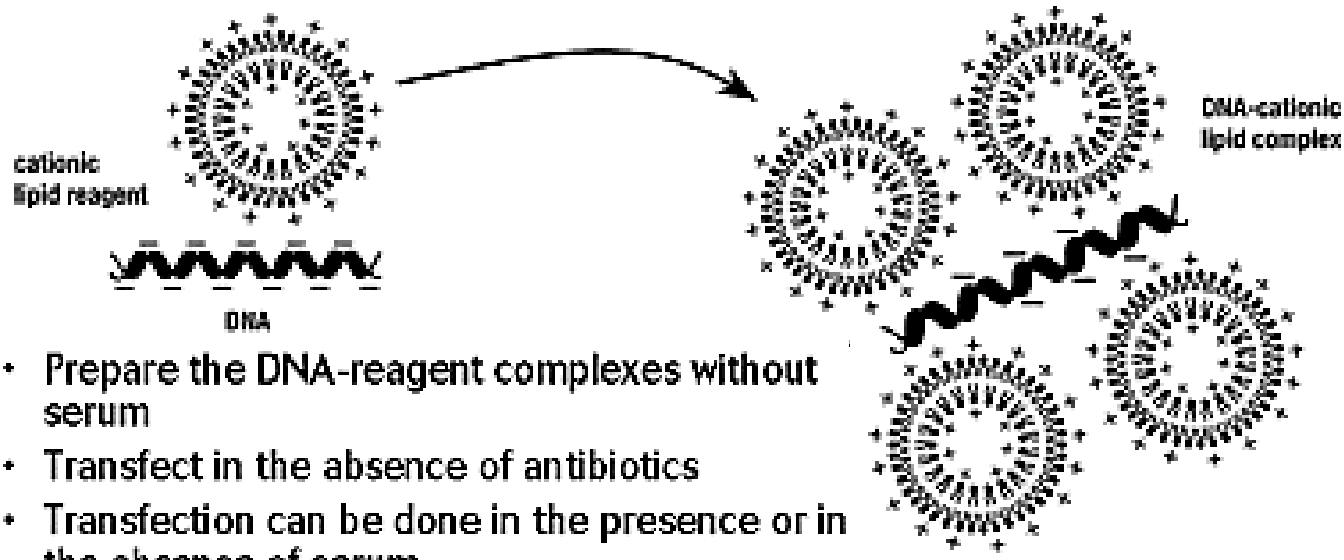
Transfecting Mammalian Cells



Advantages of Cationic Lipid Reagents

- High efficiency
- Work in many cell types
- Easy to use
- Reproducible

Cationic Lipid Reagent-Mediated Transfection



- Prepare the DNA-reagent complexes without serum
- Transfect in the absence of antibiotics
- Transfection can be done in the presence or in the absence of serum
  - Some cells transfect better in the presence of serum, others in the absence of serum
  - Transfection in serum-containing medium is convenient
  - Optimize parameters accordingly
- Transfection can be done in some serum-free media