Chapter II:

Host-vector systems and molecular cloning

Host-vector systems and molecular cloning

- Definitions:
  - "Cloning" is a term that can be used to mean two very different process:
    - Cleaving a fragment of DNA from an organism and inserting it into a vector where it can be replicated in a host organism. (**molecular cloning or gene or DNA fragment cloning**)
    - Using the nuclear DNA of one organism to create a second organism with the same nuclear DNA (**cell cloning**)

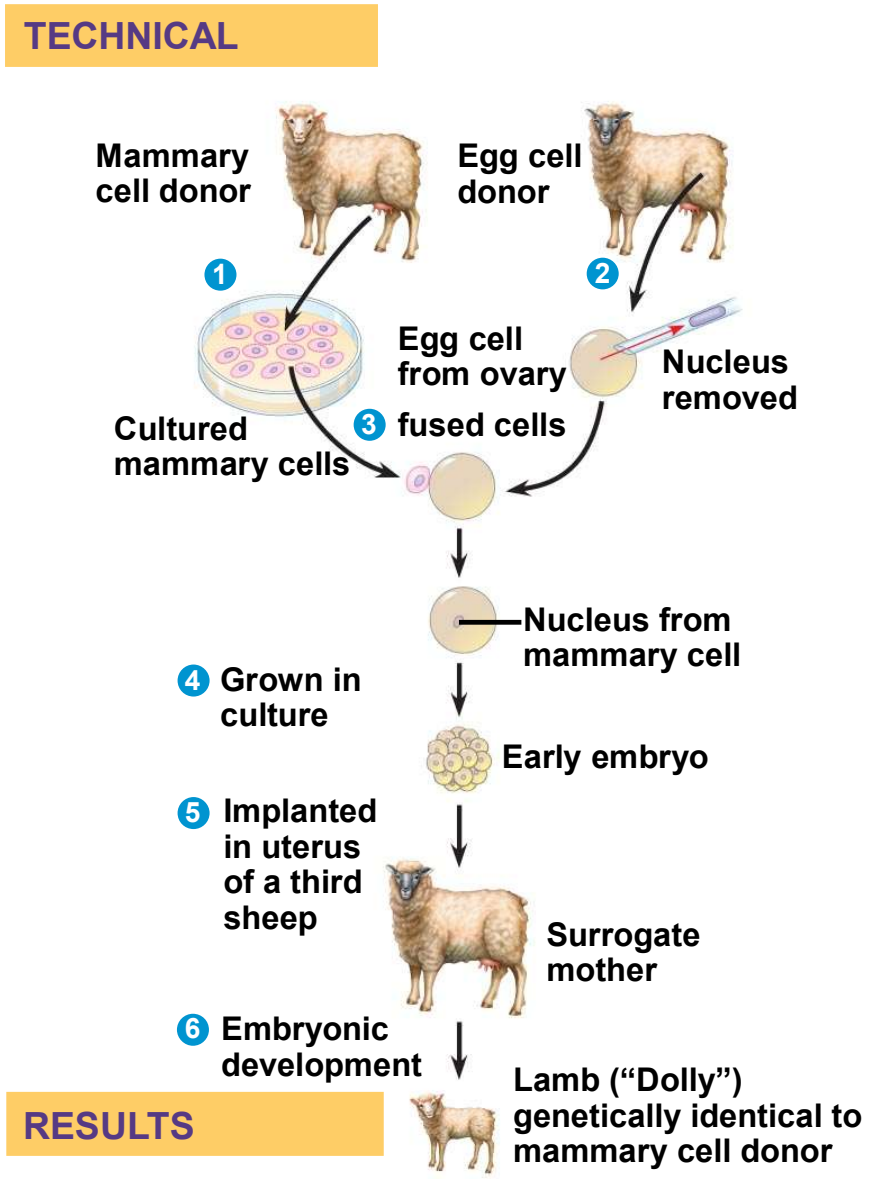
Host-vector systems and molecular cloning

- **Subcloning:**
  - Transfer of a cloned DNA insertion or a portion thereof from one vector to another vector
- **Recombinant plasmid:**
  - Vector into which foreign DNA has been introduced
- **Recombinant organism:**
  - Organism into which a recombinant vector has been introduced

Host-vector systems and molecular cloning

Cell cloning

Cell cloning



Host-vector systems and molecular cloning

Molecular Cloning

Host-vector systems and molecular cloning

- Molecular cloning:
  - Cleaving a fragment of DNA from an organism and inserting it into a vector where it can be replicated in a host organism. (or gene or DNA fragment cloning)
- Why clone?
  - To separate, identify, manipulate, or express a specific fragment or gene of DNA

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 cloning vectors

- Types of Cloning Vectors
  - Plasmids Bacteria
  - Expression vectors
  - Bacteriophage
  - Cosmides
  - Viral vectors
  - Bacterial Artificial Chromosomes (BAC)
  - Yeast Artificial Chromosomes (YAC)
  - Human artificial chromosome

Cloning vectors *vs.* insert size

Vector type	Insert size (kb)
Plasmid	0.1-10
Bacteriophage	10-20
Cosmide	35-45
BAC (bacterial artificial chromosome)	50-300
YAC	100-2,000
Human artificial chromosome	>2,000

Host types (expression systems)

- Bacteria : plasmids , phages
- Yeasts : plasmids , yeast artificial chromosomes (YACs)
- Insects : baculovirus, plasmids
- Mammals :
  - Vectors viral (therapy) gene) :
    - SV40
    - vaccinia virus
    - adenovirus
    - retrovirus

The cloning strategy

- The strategy depends on the initial information and the desired objective.
- The initial information or sources:
  - protein sequence
    - positional cloning information
  - mRNA sequence
  - cDNA libraries
  - genomic DNA banks
  - known or unknown DNA sequence
  - PCR product

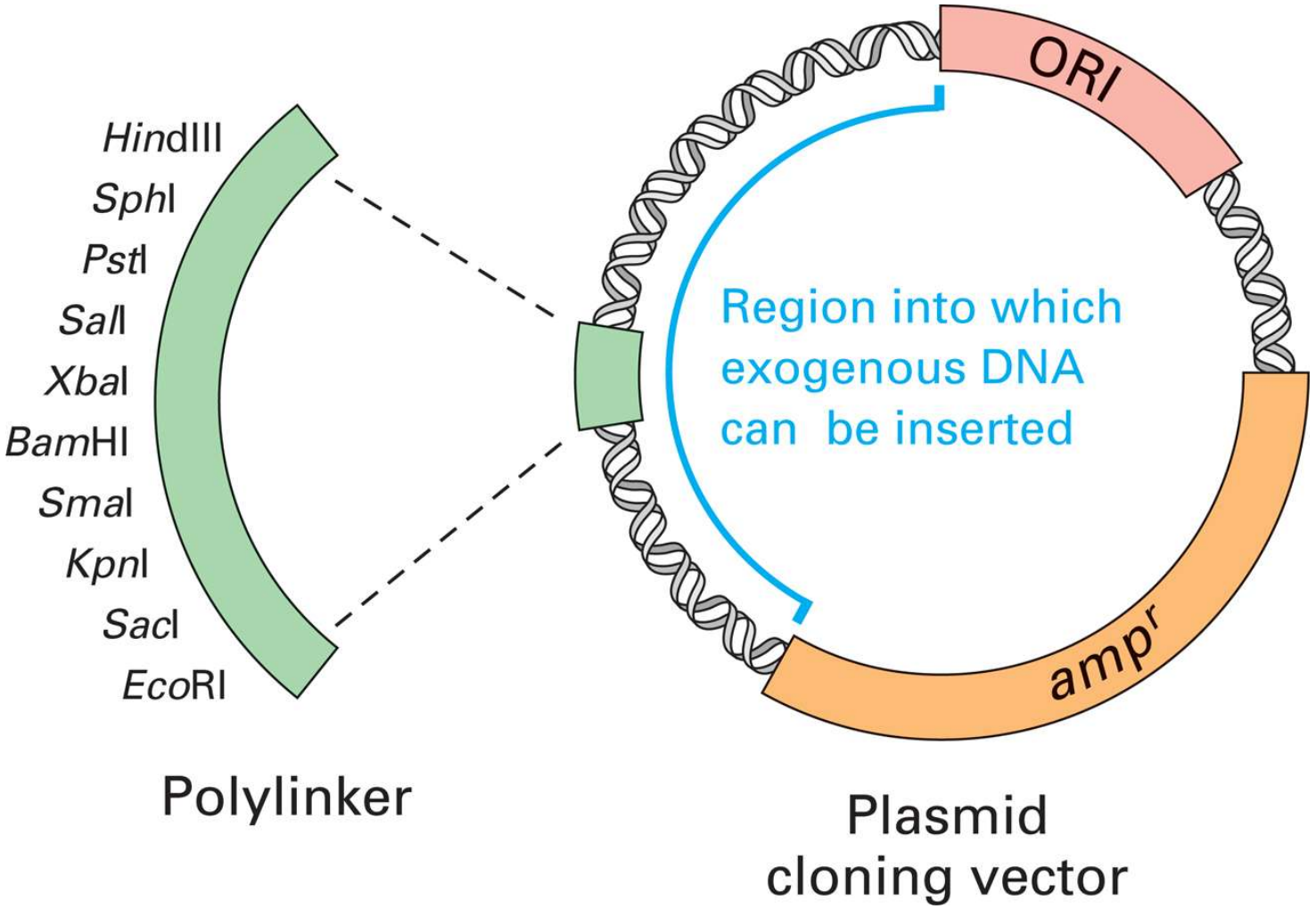
Plasmids as cloning vectors

- Plasmids are circular DNA sequences found naturally in bacteria.
- Small size, and many copies.
- They can carry antibiotic resistance genes, toxins, or other proteins.
- They replicate independently of the organism's genome.
- They can be designed to be useful cloning vectors.

Plasmids as cloning vectors

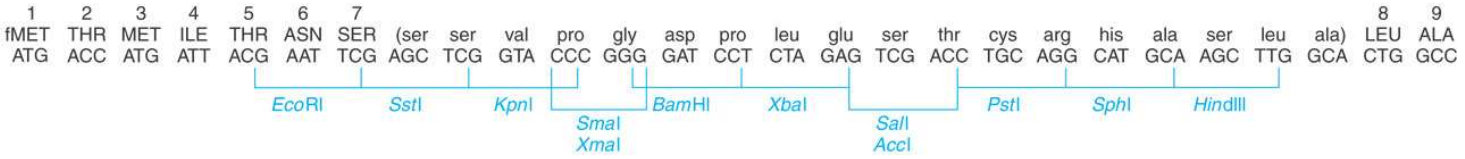
- The plasmid cloning vector:
  - A cloning vector is a generally circular DNA molecule, distinct from chromosomal DNA, that allows genetic information to be integrated into a microorganism.

- The plasmid cloning vector



The advantage of Polylinker

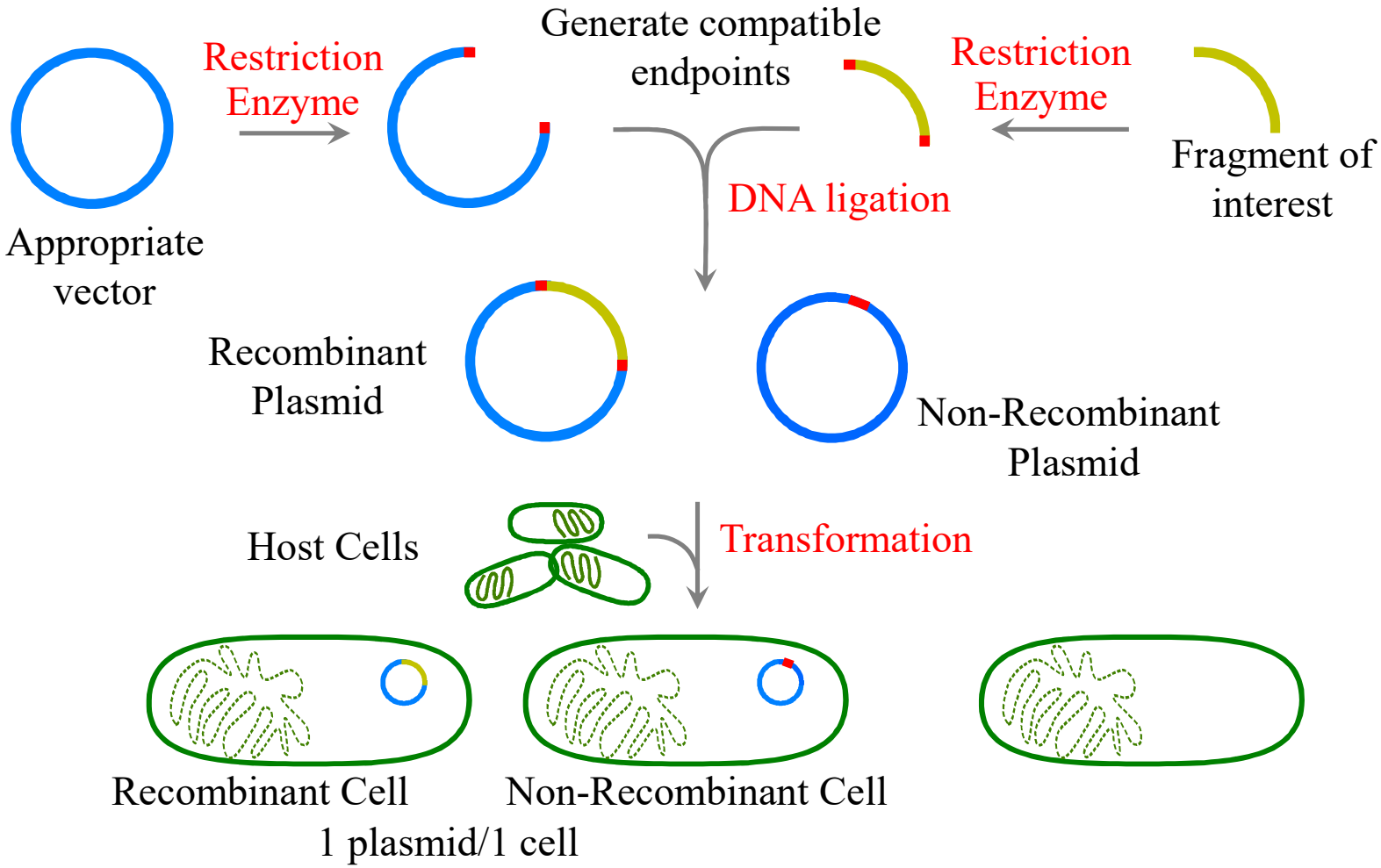
- Unique insertion sites (generally)
- facilitated excision of the insert
- Restriction endonuclease mapping and facilitated subcloning



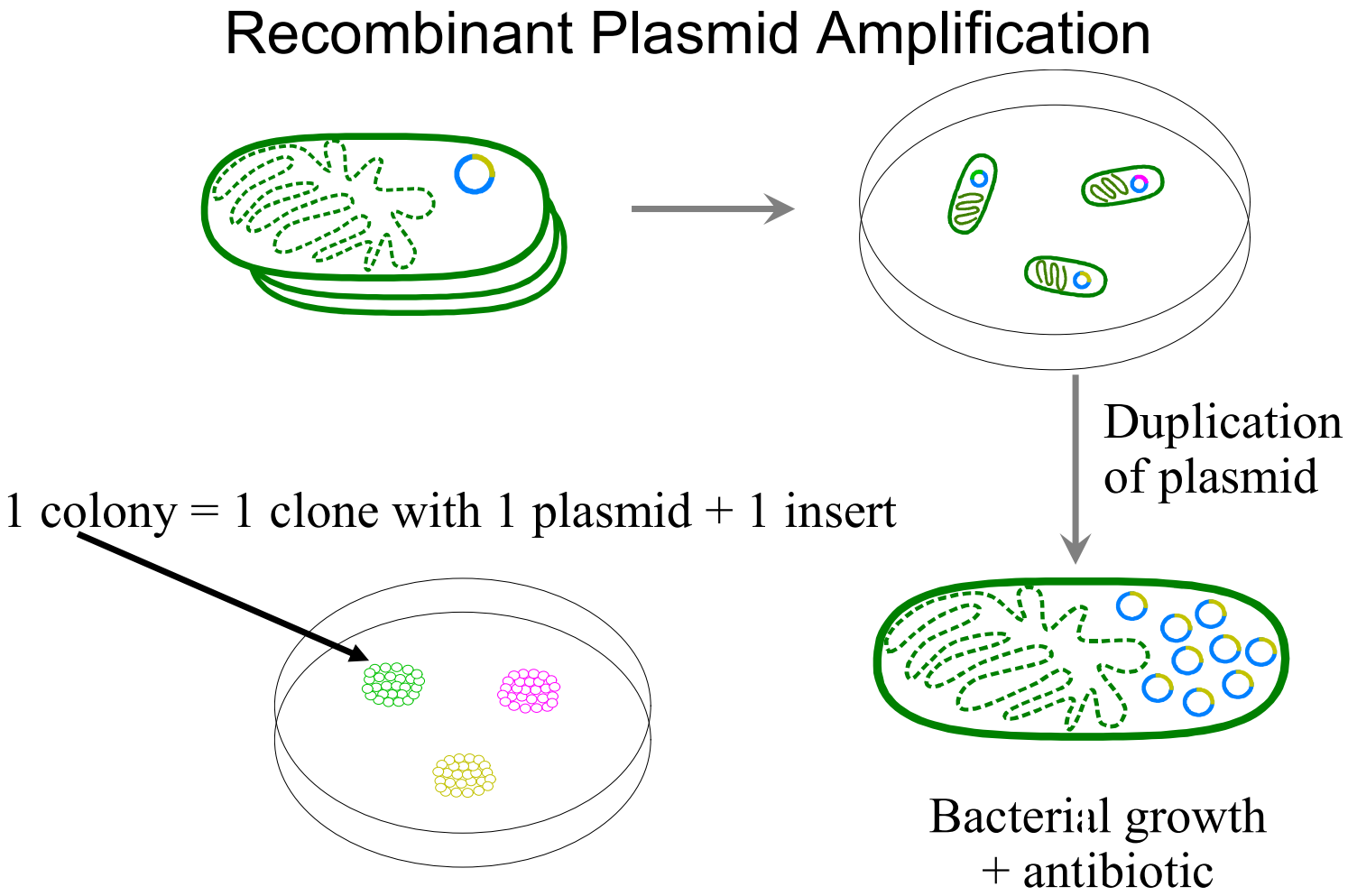
The cloning methodology

- **Cut** the cloning vector with a restriction enzyme of choice (e.g., EcoRI)
- **Cutting** the DNA of interest with the same restriction enzyme yields the same cohesive ends.
- **Mix** the restricted cloning vector and the DNA of interest together.
- **Ligature** the fragments together using DNA ligase
- The Transformation: **Inserting** ligated DNA into a host of choice -
- **Cultivating** host cells under restrictive conditions,
- **Growing** on plates containing an antibiotic (**the (selection**)

The cloning methodology



The cloning methodology



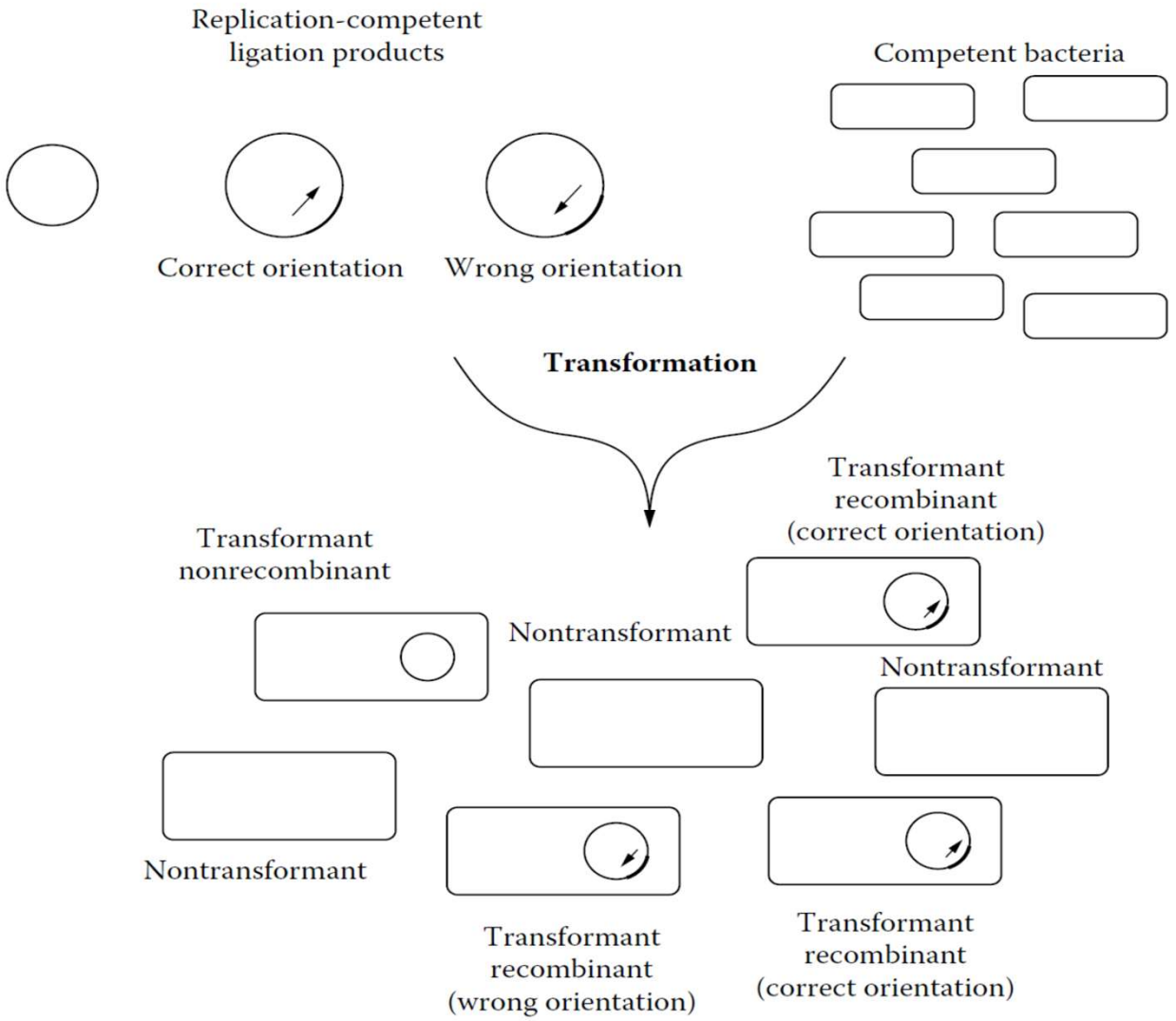
Antibiotics commonly used as selective agents

Antibiotic	Description
Ampicillin (Amp)	Inhibits bacterial cell wall synthesis; inactivated by β-lactamase, which cleaves the β-lactam ring of amplicillin.
Kanamycin (Kan)	It binds to the 30S ribosomal subunit and inhibits protein synthesis; it is inactivated by a phosphotransferase.
Neomycin (Neo)	It binds to the 30S ribosomal subunit and inhibits protein synthesis; it is inactivated by a phosphotransferase.
Tetracycline (Tet)	It binds to the 30S subunit of the ribosome and inhibits protein synthesis;

Bacterial transformation

- The transformation is the act of introduicing the vector into the appropriate host.
- The transformation process results in the insertion of a DNA molecule into the host cell.
- All plasmid and bacteriophage vectors commonly used to clone foreign DNA fragments allow the insertion of a single vector molecule into the host cell.
- This single molecule can be amplified multiple times in the host, but all resulting molecules are identical.

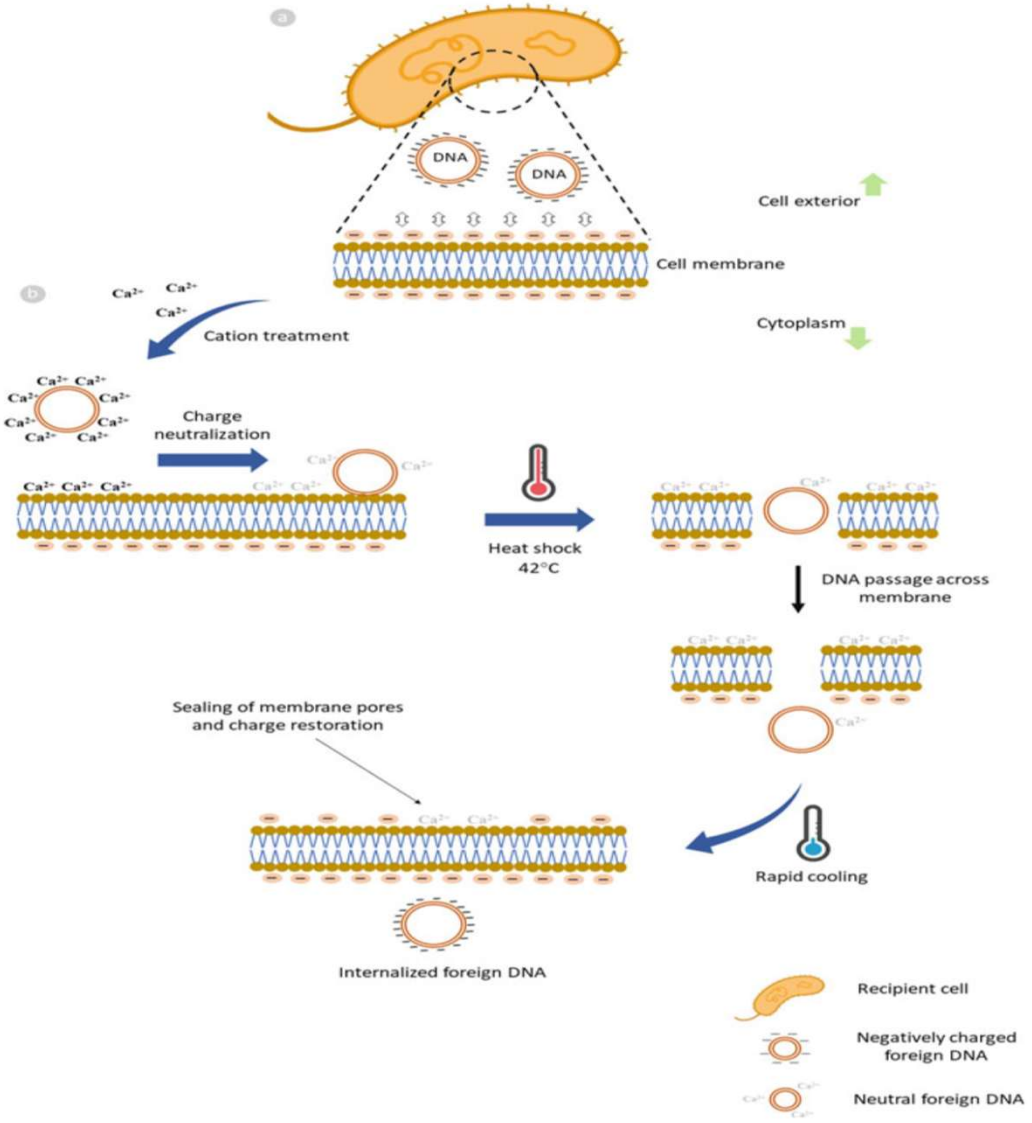
Bacterial transformation



Transformation by thermal shock

- The transformation of competent cells by heat shock is achieved by mixing plasmid DNA with the cells, incubating for 2 min at 42° C is often used to enlarge the nanopores created in the membranes, then the mixture is directly transferred onto ice for 90 seconds and thus stimulate the entry of the plasmid into the cells.
- The transformed cells are generally incubated in a nutrient broth at 37° C for 60 to 90 minutes.

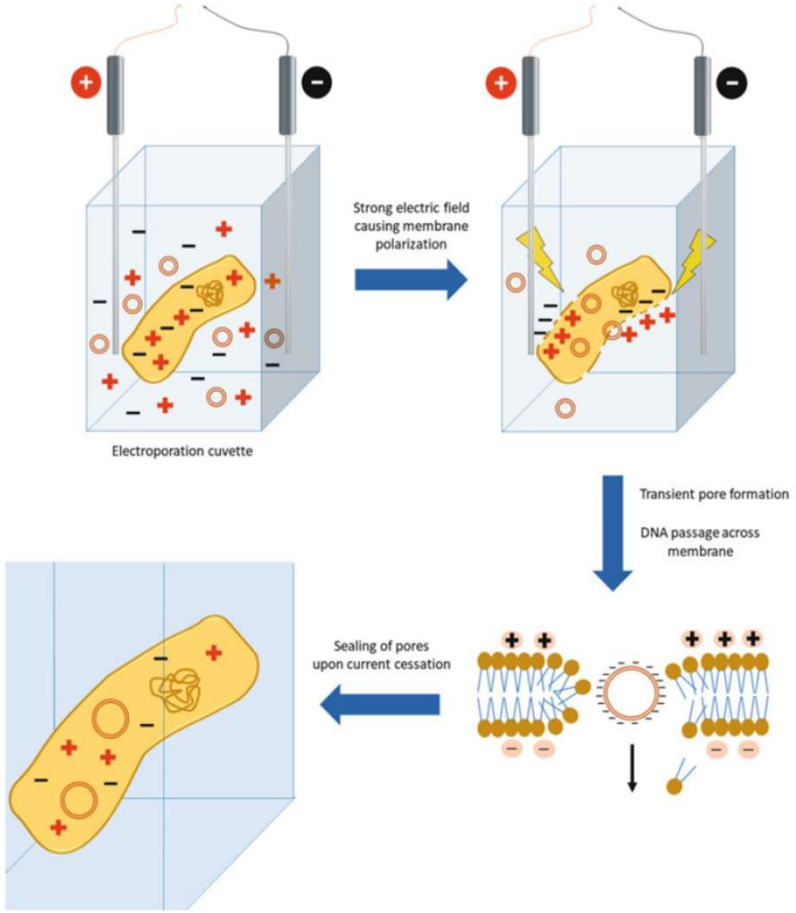
Bacterial transformation



Transformation by electric shock

- Increased transformation efficiencies have been observed using high-voltage electrical pulses in a process called electroporation.
- Electrical shock transformation (or electroporation) is the use of an electric field pulse to induce microscopic pores in a biological membrane

Bacterial transformation



Selection of recombinant clones

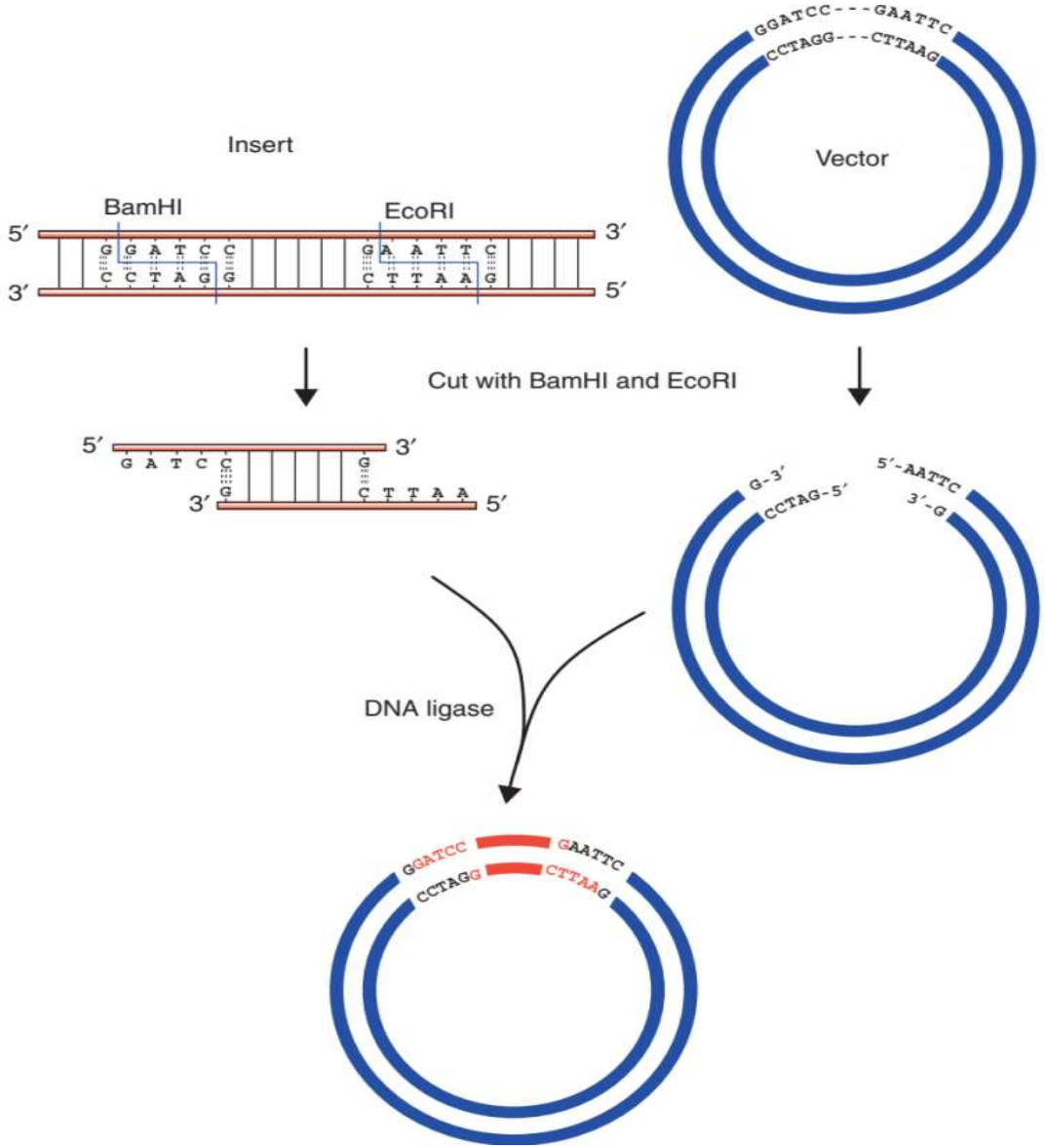
- There are many different ways to screen (select) recombinant clones, depending on the vector used for cloning.
- The most commonly used method for screening transformants, for example, is antibiotic selection, although other methods such as blue-white screening and colony PCR are also available.

Selection of recombinant clones

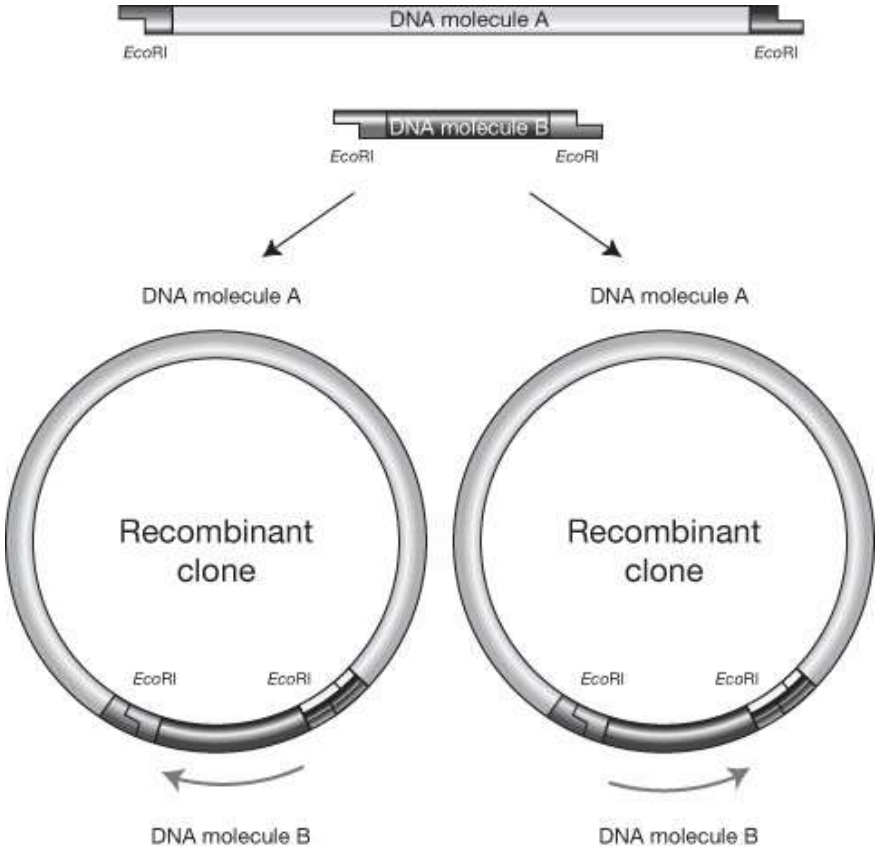
- For plasmid cloning, antibiotic screening is the simplest and fastest way to choose transformants .
- In this method, the transformation mixture is plated onto an agar plate that includes the relevant antibiotic.
- Non- transformers do not carry the plasmid and therefore will not survive on the plate containing the antibiotic, whereas all transformants (recombinant or not) will produce colonies due to the presence of the resistance gene.

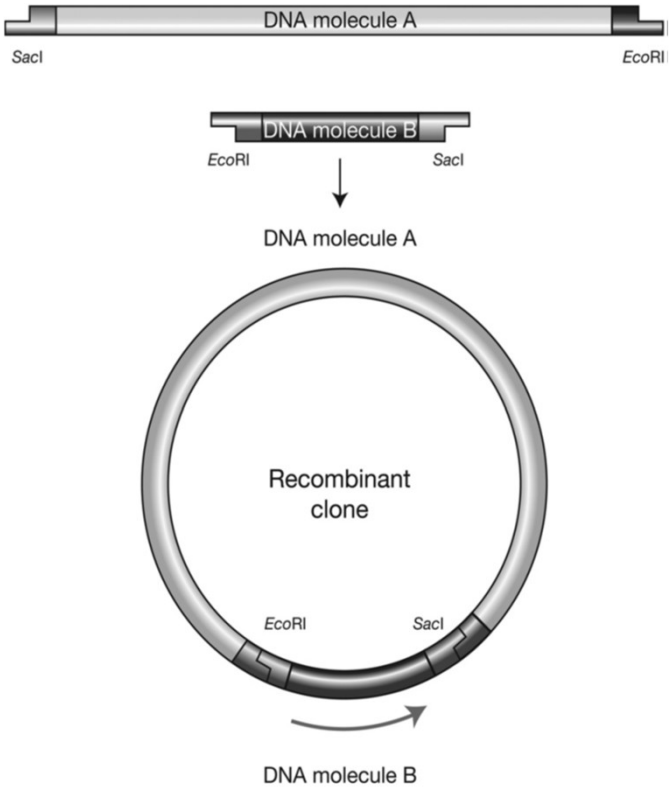
Cloning Direction

- Often, we need to insert foreign DNA in a particular orientation.
- This can be done by performing two cleavages with two different restriction enzymes
- Cutting of foreign DNA by the same two restriction enzymes
- Foreign DNA can only be inserted in one direction

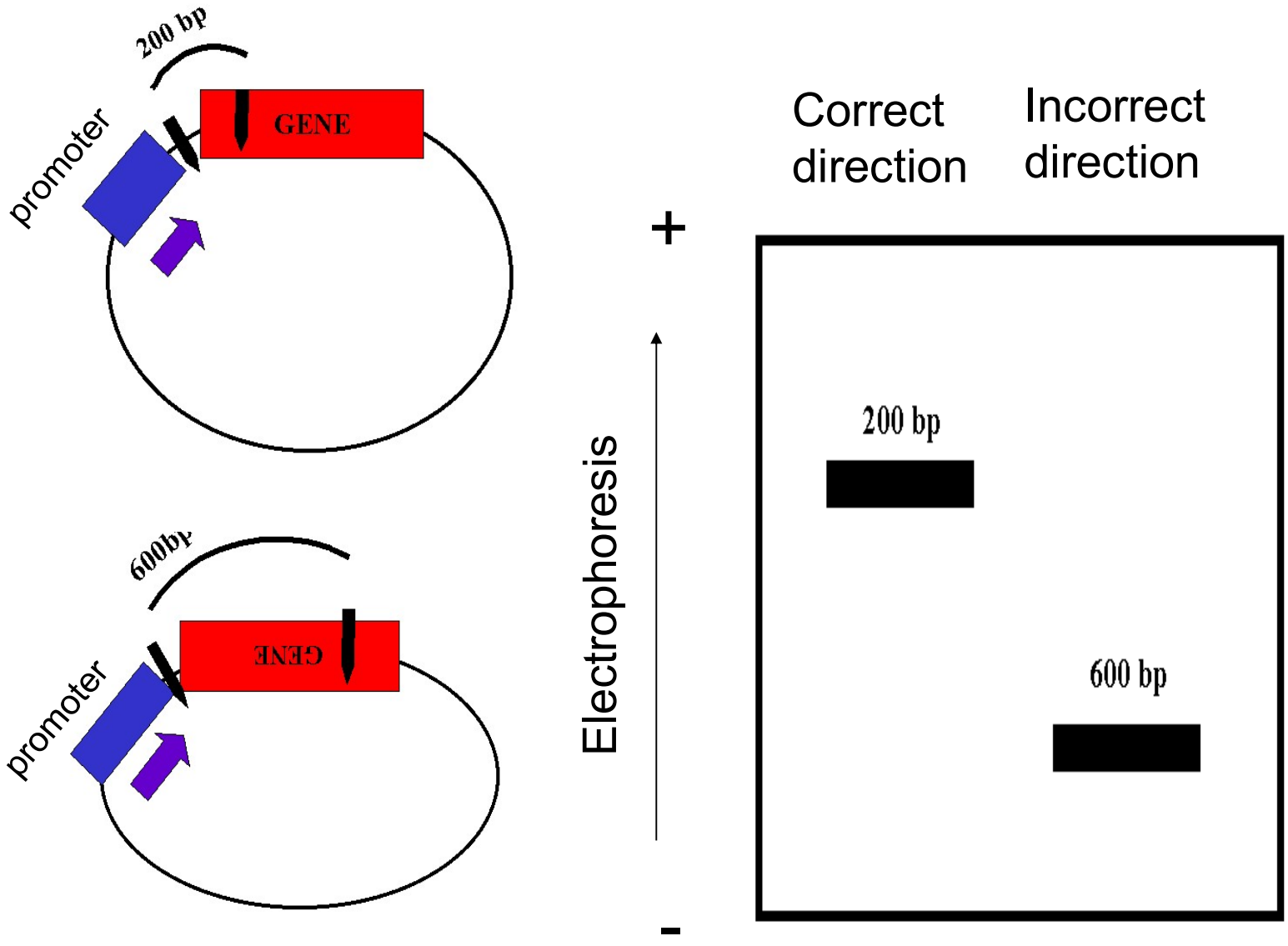

Insert
BamHI
EcoRI
5′
3′
G G A T C C
C C T A G G
G A A T T C
C T T A A G
Vector
GGATCC - - - GAATTC
CCTAGG - - - CTTAAG
Cut with BamHI and EcoRI
G A T C
C T T A A
G-3′
CCTAG-5′
5′-AATTC
3′-G
DNA ligase
GGATCC
GAATTC
CCTAGG
CTTAAG

Cleavage with two different restriction enzymes

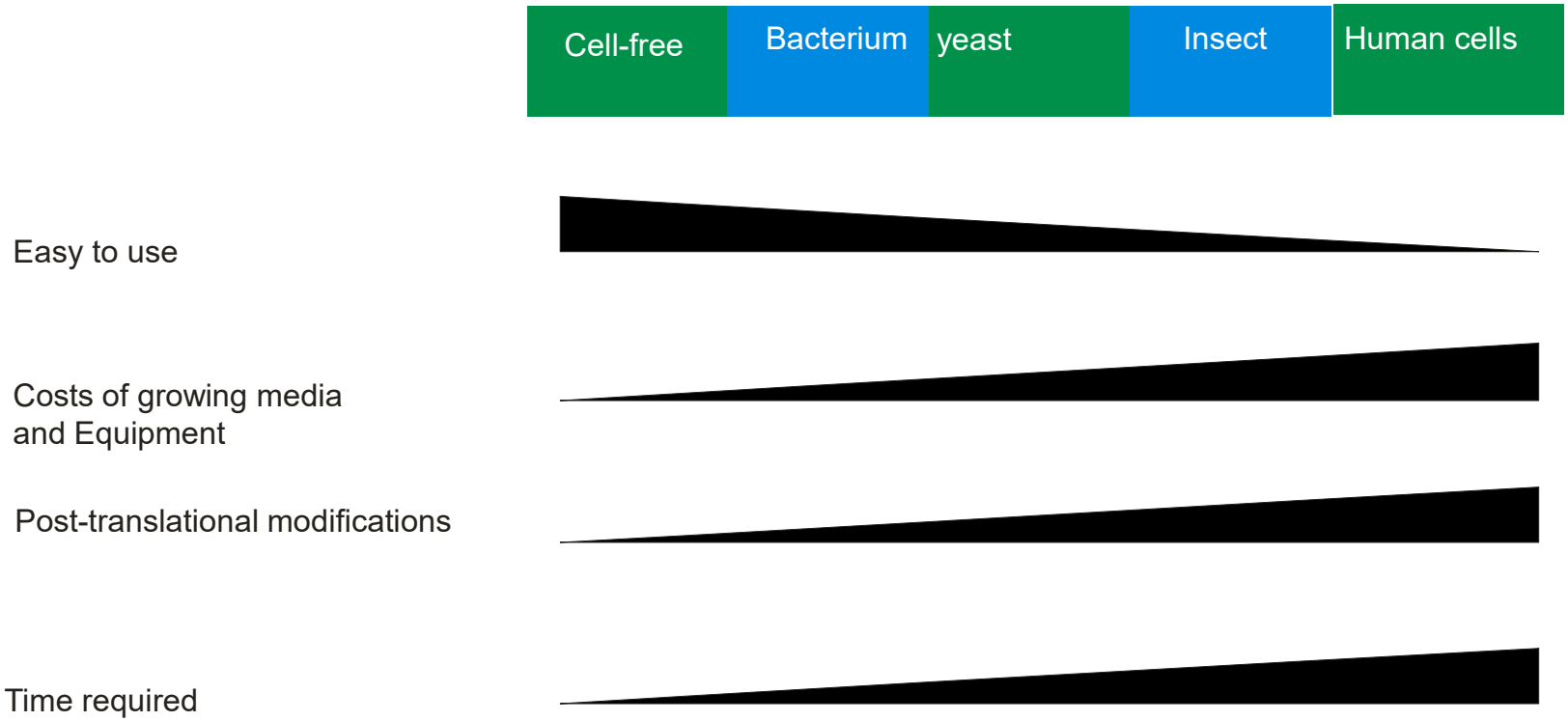




Checking for correct orientation



Host selection



Types of cloning vectors

- Types of Cloning Vectors
  - Plasmids Bacteria
  - Expression vectors
  - Bacteriophage
  - Cosmides
  - Viral vectors
  - Bacterial Artificial Chromosomes (BAC)
  - Yeast Artificial Chromosomes (YAC)
  - Human artificial chromosome

Plasmids as cloning vectors

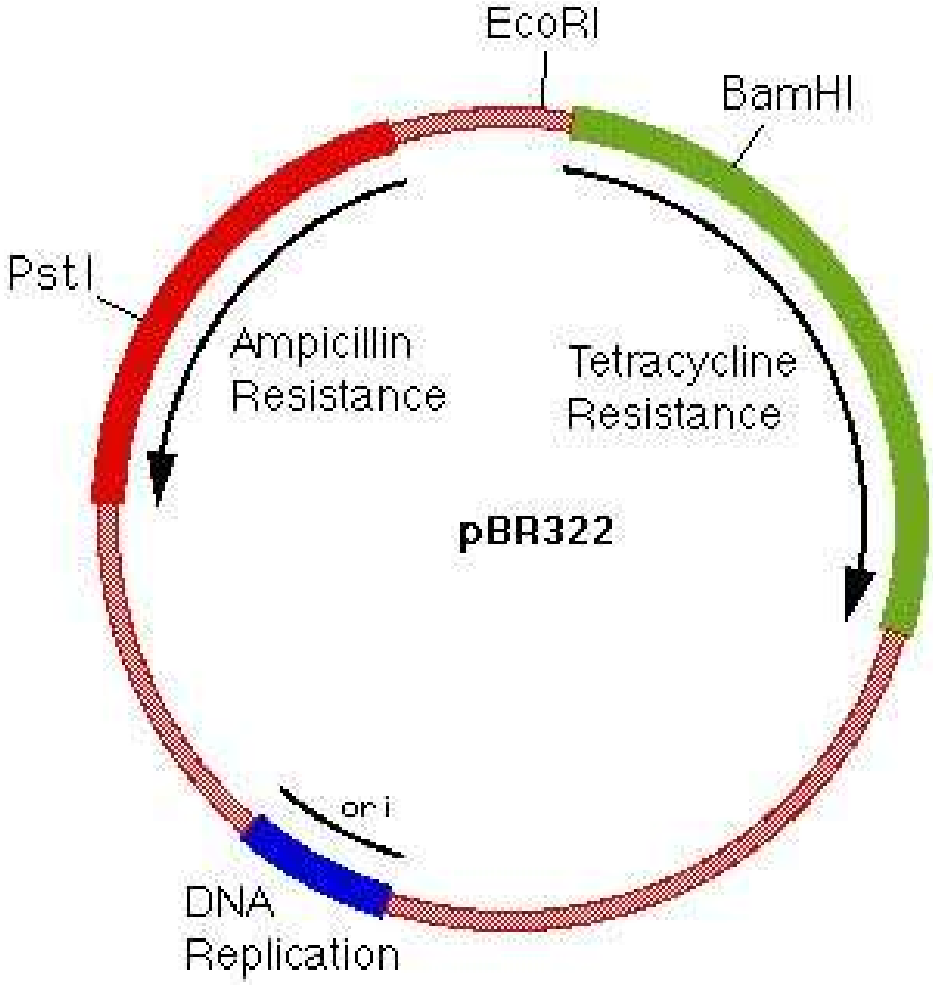
- Plasmid vectors can be designed with a variety of characteristics:
  - Antibiotic resistance (e.g., resistance to ampicillin)
  - Colorimetric markers (e.g., the *lacZ gene)*
  - Strong or weak promoters to direct the expression of a protein

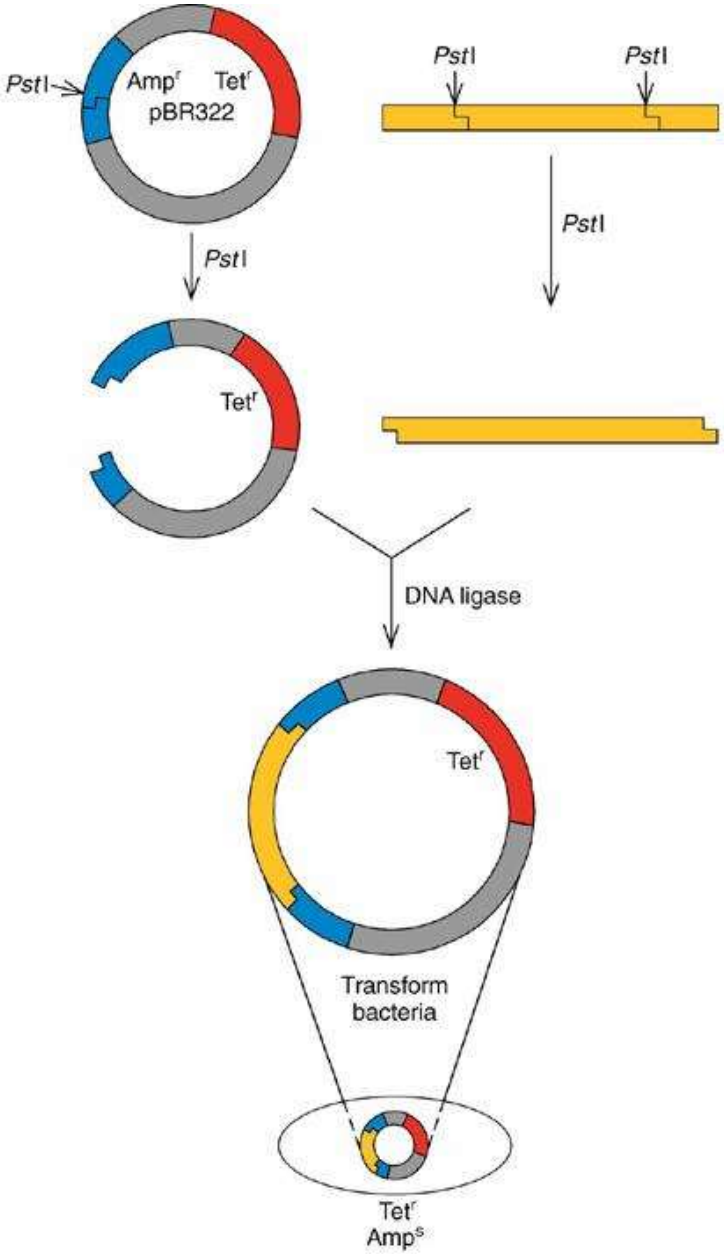
The first plasmid vectors

- **The PBR322 vector**
  - It contains an origin of replication, a tetracycline resistance gene (tetr) and an ampicillin resistance gene (ampr).
  - The ampr gene codes for a protein (β-lactamase) of 286 amino acids capable of catabolizing this antibiotic (ampicillin).
  - Contains numerous restriction sites distributed throughout the sequence. Many of these sites are unique, allowing the circular DNA to be converted into linear DNA. Digestion with BamHI allows the vector to be recombined with a BamHI-digested DNA fragment to be cloned.

The first plasmid vectors

PBR322

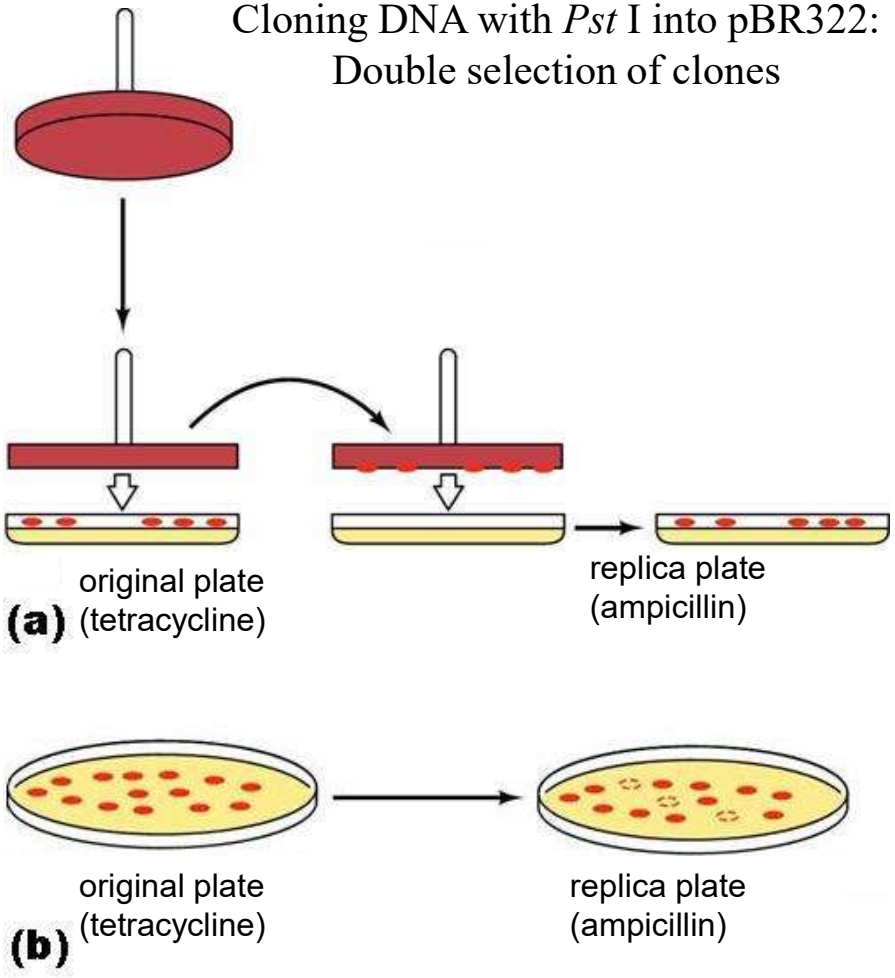




Example:

Cloning of DNA cut by *Pst* I into pBR322.

Cloning DNA with *Pst* I into pBR322:
Double selection of clones



(a) A circular velvet instrument is used to touch the surface of the first dish containing bacterial colonies. Cells from each of these colonies adhere to the velvet and can be transferred to the replica plate in the same positions relative to the other.

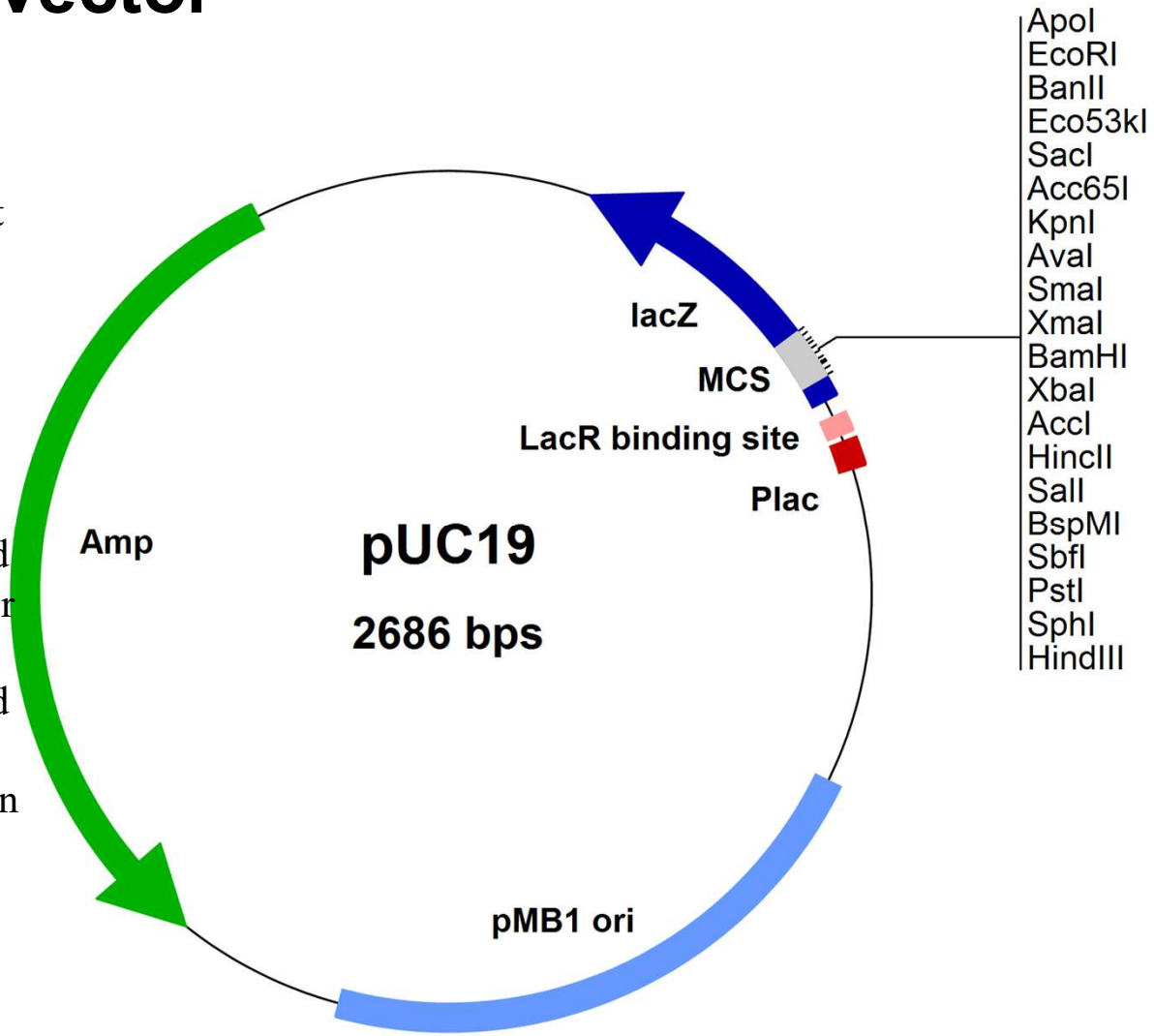
(b) Screening for pBR322 ampicillin resistance gene insertions by replica plating. The original plate contains tetracycline, so all colonies containing pBR322 will grow. The replica plate contains ampicillin, so colonies carrying pBR322 with ampicillin resistance gene insertions will not grow (these colonies are represented by dashed circles). The corresponding colonies from the original plate can then be recovered.

Second-generation vectors

- **The pUC19 vector**

Cloning into a plasmid (pUC):

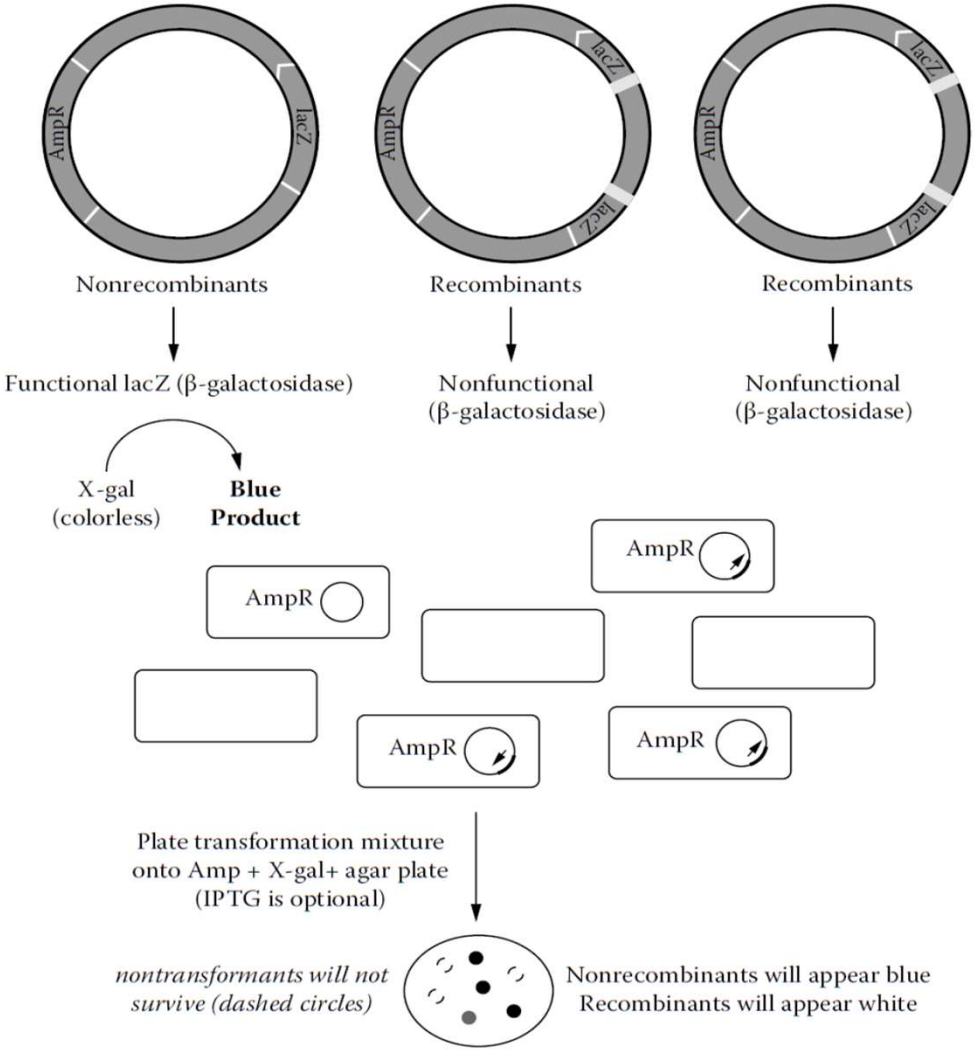
- Replication independent of that of the bacterial chromosome (ori)
- Cloning polylinker at a portion of the E. coli β-galacto-sidase gene (LacZ').
- Selection of transformed bacteria (plasmid with or without insert): ampR
- Selection of transformed bacteria (plasmid with insert): white/blue screen in the presence of IPTG and X-gal

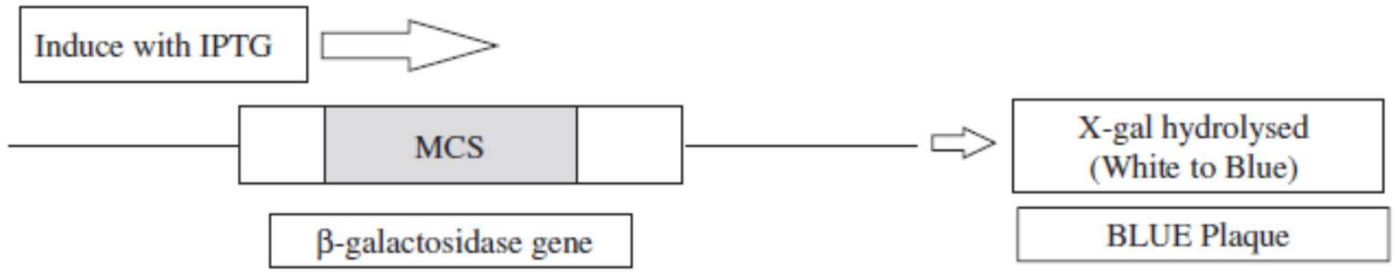


Cloning of the gene of interest

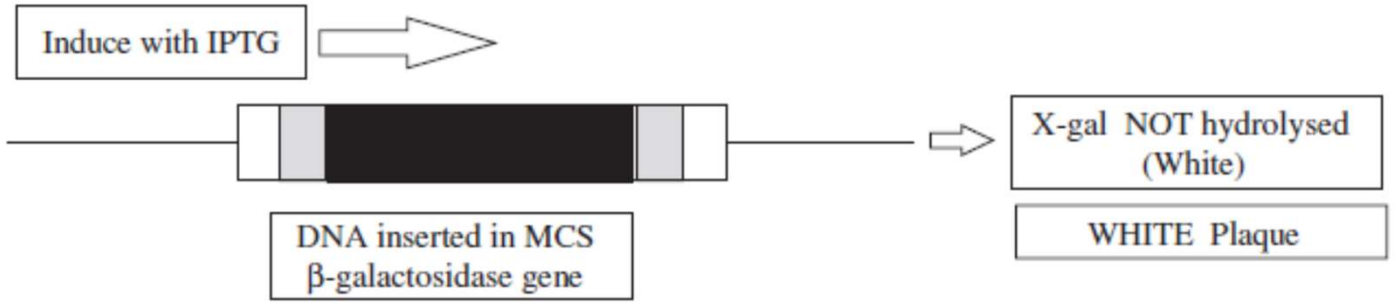
- **Selection**
- There are many different ways to screen (select) recombinant clones, depending on the cloning vector used. The most common method for screening transformants, for example, is antibiotic selection, although other methods such as blue-white screening exist.

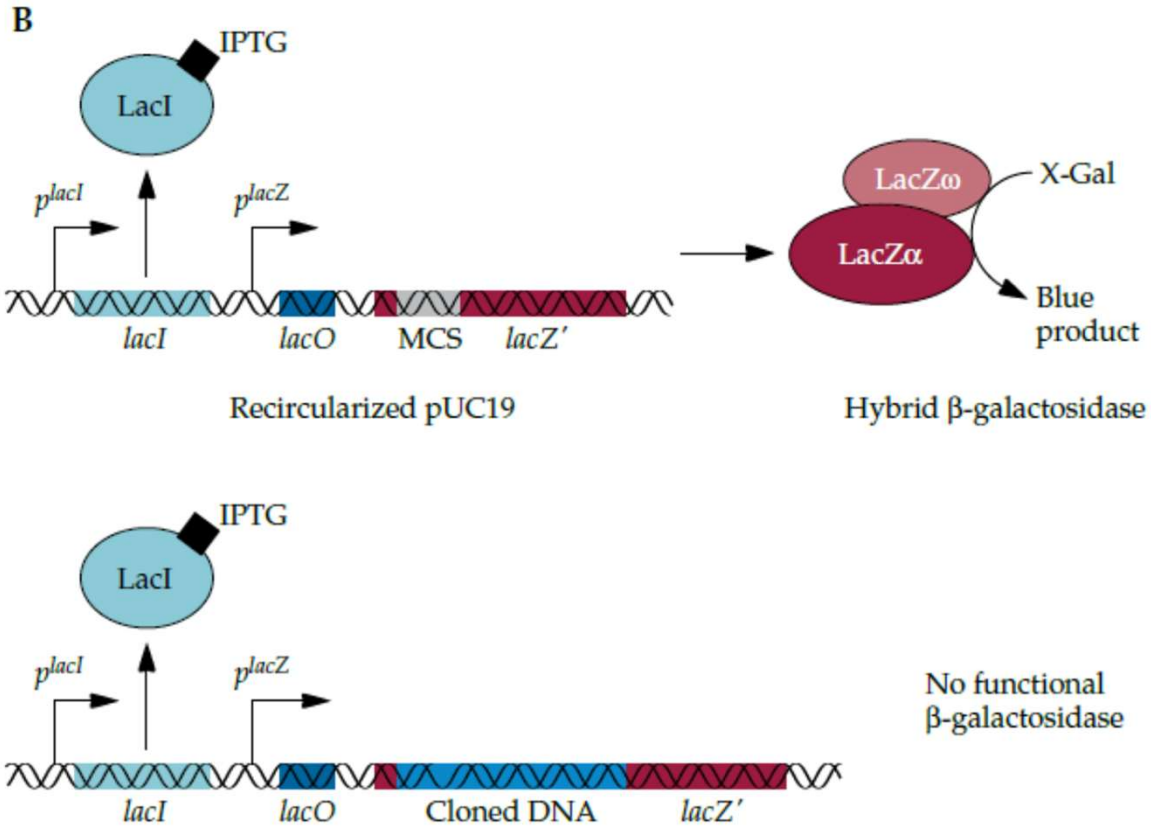
Selection of recombinant clones

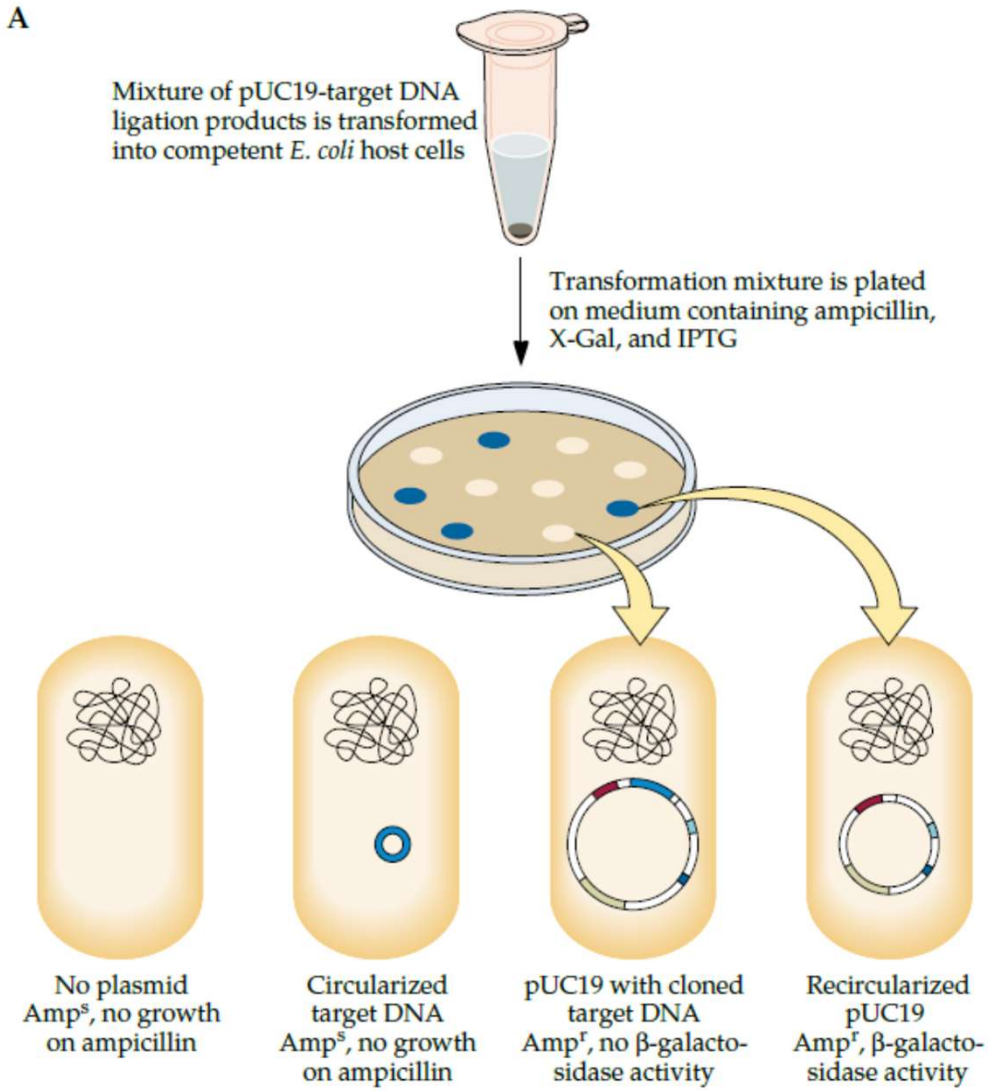


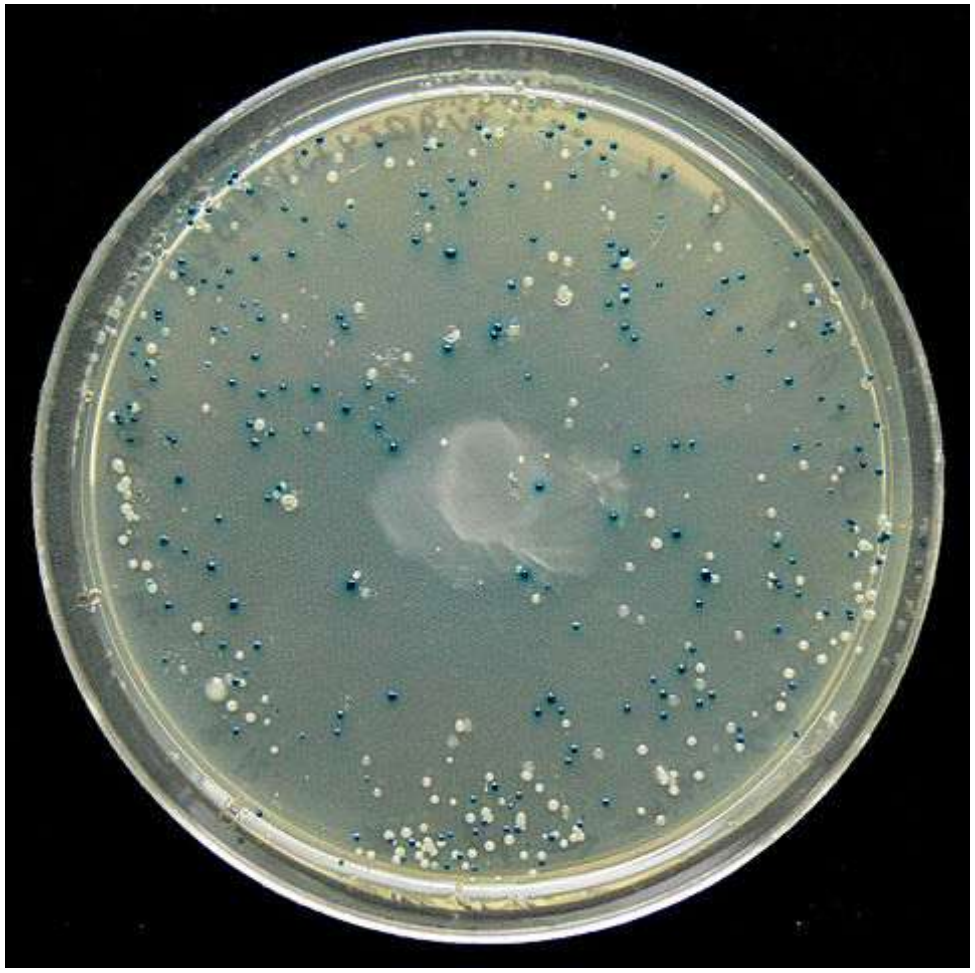


Recombinant vector (insert within MCS)



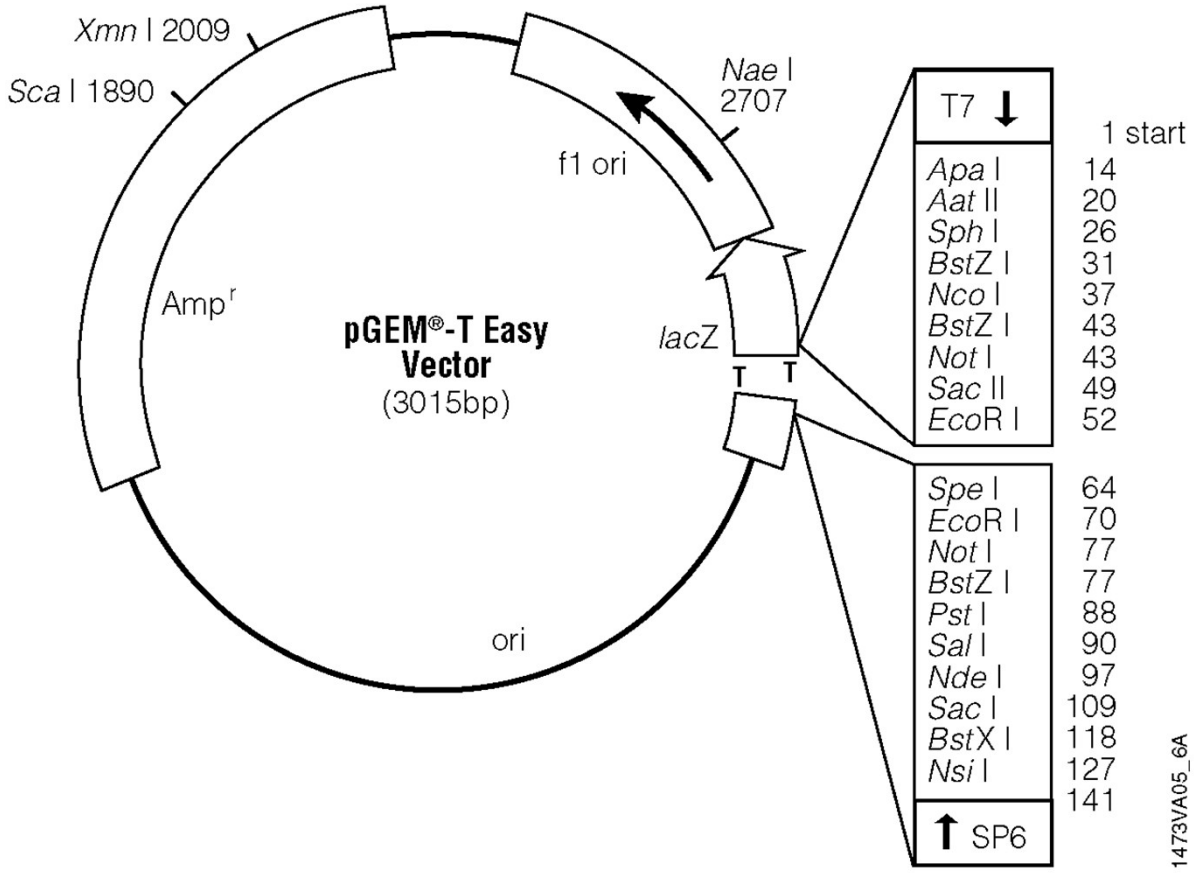

B
IPTG
LacI
p^{lacI}
p^{lacZ}
lacI
lacO
MCS
lacZ′
LacZω
X-Gal
LacZα
Blue
product
Recircularized pUC19
Hybrid β-galactosidase
IPTG
LacI
p^{lacI}
p^{lacZ}
lacI
lacO
Cloned DNA
lacZ′
No functional
β-galactosidase


A
Mixture of pUC19-target DNA ligation products is transformed into competent E. coli host cells
Transformation mixture is plated on medium containing ampicillin, X-Gal, and IPTG
No plasmid
Amp^s, no growth on ampicillin
Circularized target DNA
Amp^s, no growth on ampicillin
pUC19 with cloned target DNA
Amp^r, no β-galacto-sidase activity
Recircularized pUC19
Amp^r, β-galacto-sidase activity



Growth on the plate: blue = uninserted DNA and white = inserted DNA. A certain percentage of digested vectors will recircularize without an insert. Remove 5' phosphates with alkaline phosphatase to prevent recircularization (which also eliminates blue plaques).

Next generation vectors



T-A cloning plasmid: Allows cloning of PCR products,

Expression Vectors

- Expression vectors: allow experimenters to control the expression of cloned genes based on transcriptional control.
- They predict high levels of protein expression
- Strong promoters
- Effective transcription terminators are used to prevent the expression of other genes on the plasmid.

Plasmid expression vectors

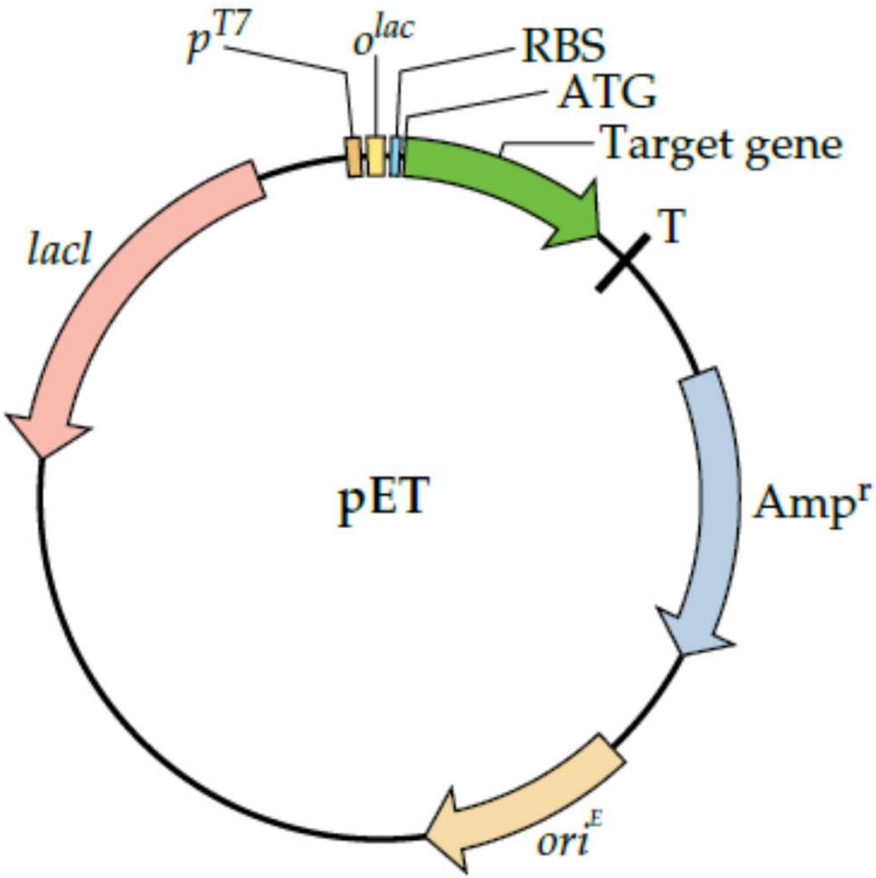
- Unlike cloning plasmids, these expression vectors contain regulatory sequences appropriate for the transcription and translation of the gene of interest.
- A common expression vector must include the following elements:
  - 1. Regulatory sequences to transcribe and translate the gene of interest (e.g. , Promoter…).
  - 2. Origin of replication.
  - 3. Multiple cloning site (MCS).
  - 4. Selectable marker for selection.
  - 5. Tags for the separation and purification of the expressed protein by affinity chromatography.

pET expression vector

- One of the most powerful systems developed for the purpose of cloning genes of interest and expressing recombinant proteins in *E. coli* is the pET vector system.
- This vector system is fundamentally a derivative of the pBR322 plasmid series.
- It was designed to make the most of the T7 bacteriophage and its characteristics.
- The target gene cloned in this vector system is under transcriptional and translational control of the expression signals of bacteriophage T7 RNA polymerase.

pET expression vectors

- The biggest advantage of using T7 RNA polymerase is its extremely high specificity for the T7 promoter.
- Furthermore, the polymerase possesses a high level of activity and translation efficiency due to translation initiation signals.
- Since the T7 promoter is not recognized by the host RNA polymerase, basal expression of the inserted gene can be avoided.
- Expression is further controlled by the practical addition of an inducer (IPTG) when the bacterial culture reaches its logarithmic growth phase.


p^{T7}
o^{lac}
RBS
ATG
Target gene
T
lacI
pET
Amp^r
ori^E

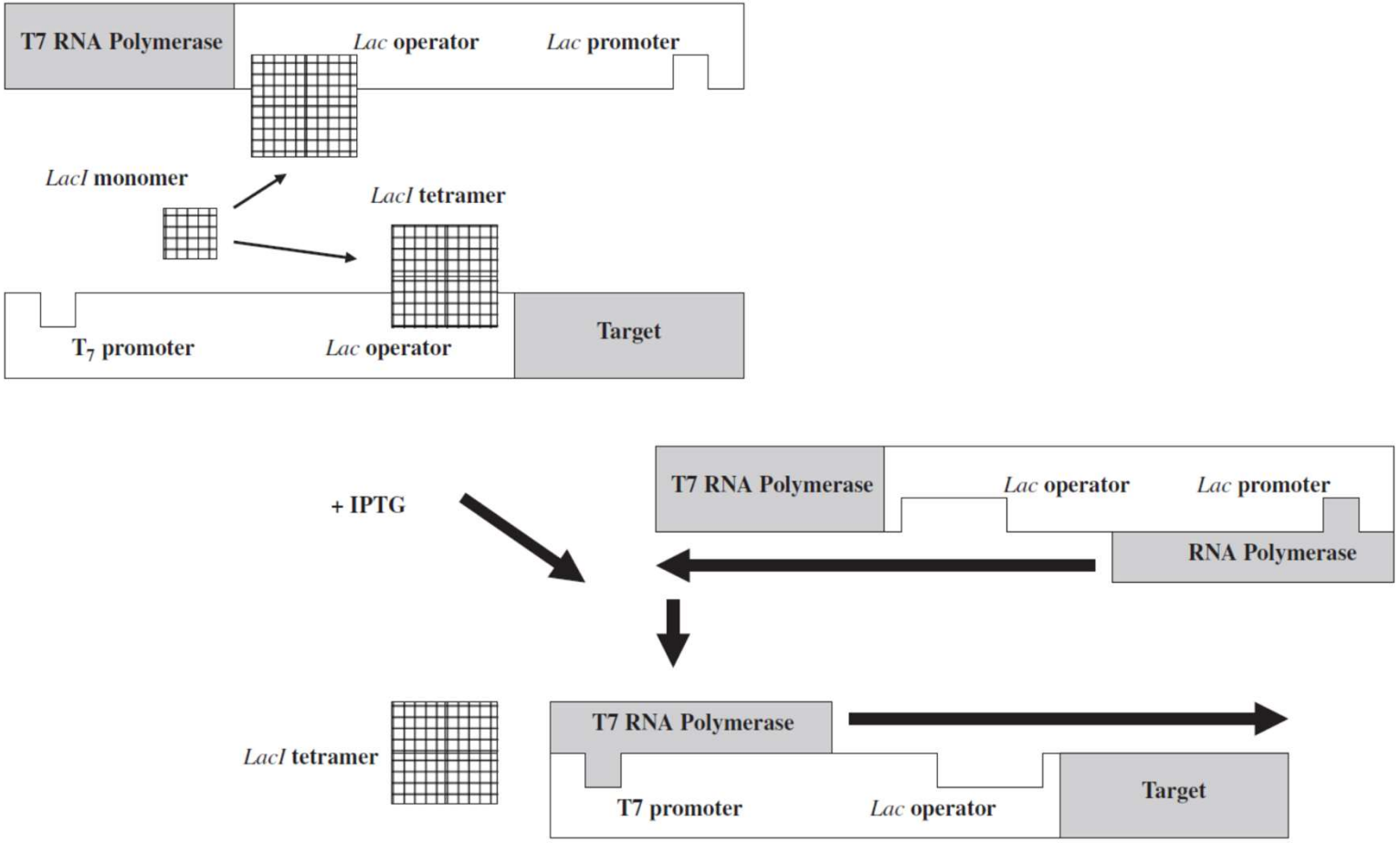
pET expression vector

- LacO : binding site for LacI . This element inhibits the activity of the T7 promoter when the LacI protein is present, preventing the leaky expression of the gene of interest.
- RBS: ribosome binding site is the translation initiation element of bacteriophage T7. This allows for efficient production of the protein of interest.
- ORF: The open reading frame for your gene of interest is placed here.
- T7 Terminator: signal sequence to terminate transcription from the gene of interest, preventing continued transcription.
- AmpicillinR: ampicillin resistance gene. It allows the plasmid to be maintained by ampicillin selection in *E. coli* .
- pBR322 ori : pBR322 replication origin. Plasmids carrying this origin and the Rop gene exist in small copy numbers in *E. coli*

pET expression vectors

- Rop : Primer repressor. It encodes a small protein that regulates the number of plasmid copies. The presence of the Rop protein , in combination with the pBR322 origin of replication on the plasmid, results in a low number of plasmid copies.
- LacI : The natural promoter *of E. coli* and the coding sequence of the lac repressor. In the absence of system induction (i.e., without IPTG), the LacI protein represses the transcription of the gene of interest from the T7lac promoter, as well as the transcription of RNA polymerase T7 from the LacUV5 promoter in host strains used for the production of recombinant proteins.

pET system control

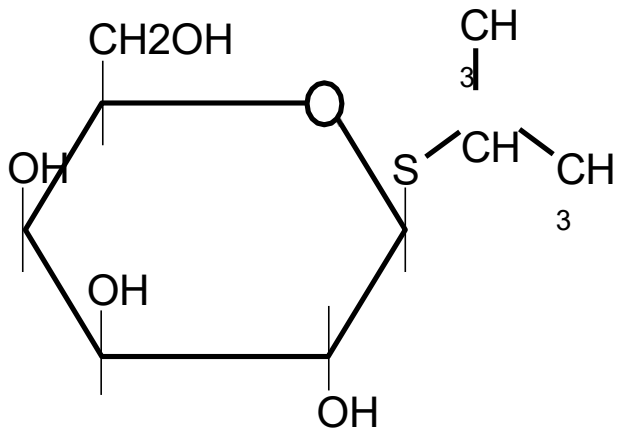


IPTG : Structure - Use

The presence of IPTG allows the induction of the activity of the *LacZ gene, encoding the* ® -galactosidase which allows the hydrolysis of lactose (Gal + Glu) and
which promotes its use, by binding to and inhibiting the *lac repressor.*

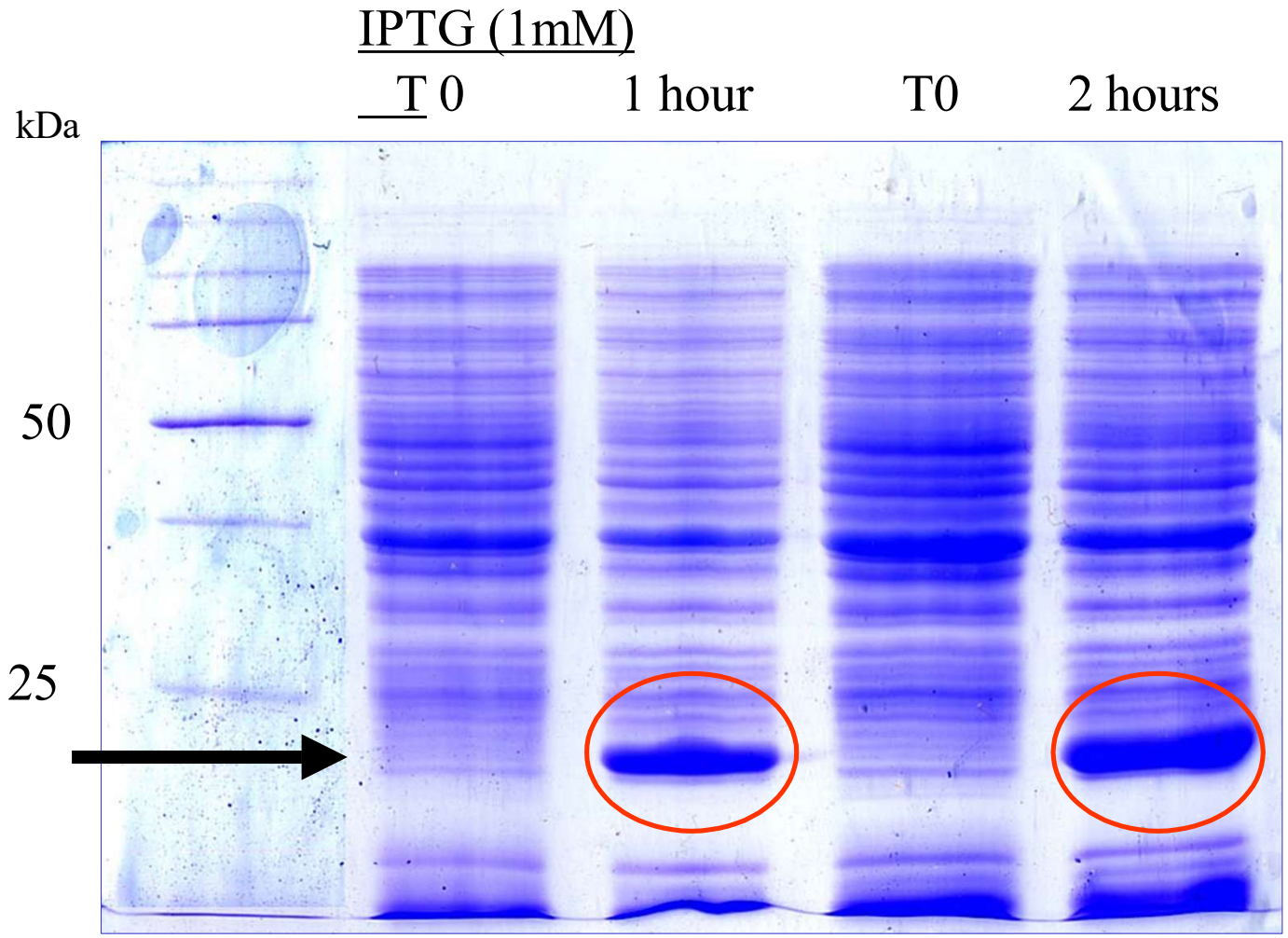
lacZ gene promoter , the presence of IPTG allows that gene to be expressed.

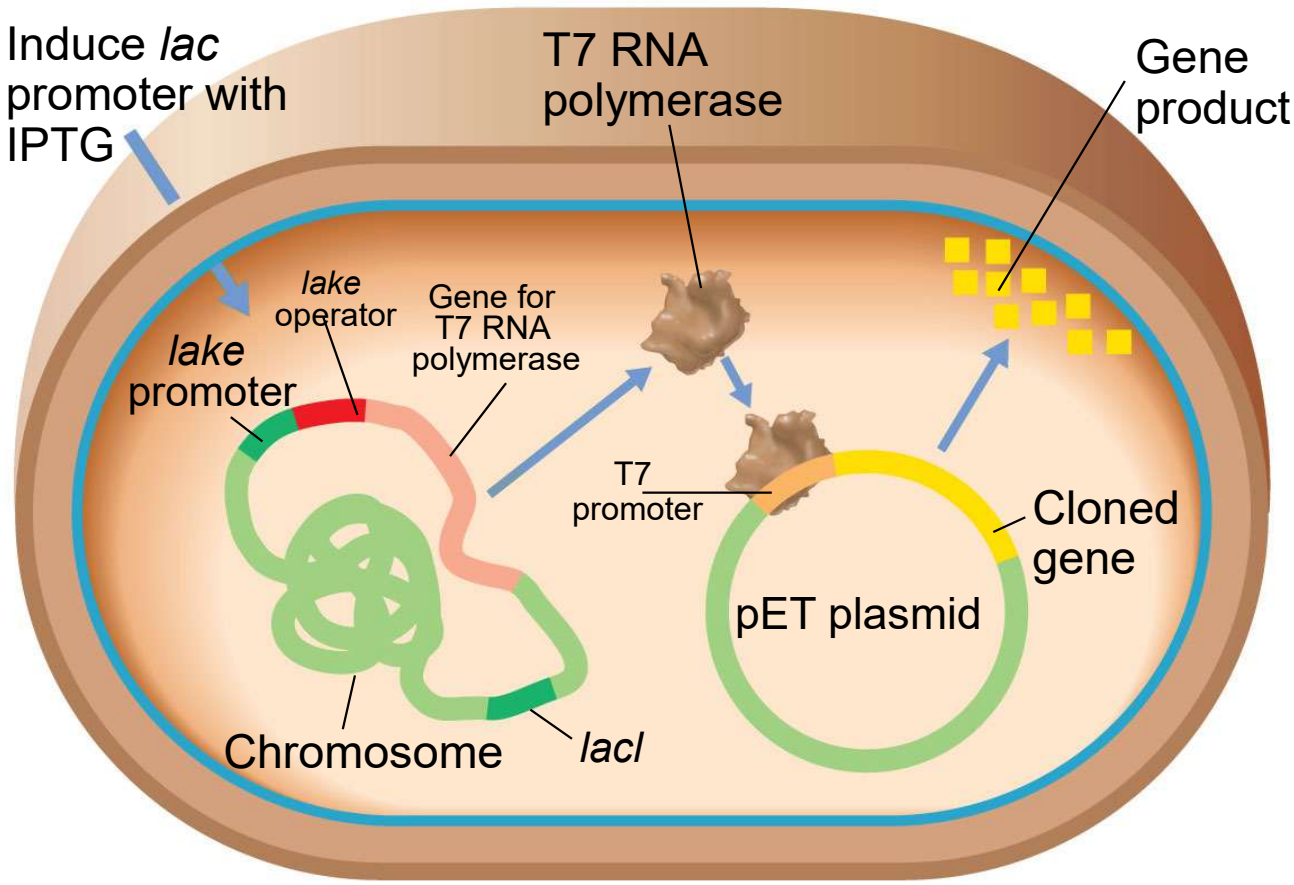
(*) IPTG = Isopropyl- β -D-1-thiogalactopyranoside
(structural analogue of allolactose)



IPTG

Induction by IPTG of there synthesis of a protein recombinant




Induce *lac* promoter with IPTG
T7 RNA polymerase
Gene product
lake operator
Gene for T7 RNA polymerase
lake promoter
T7 promoter
Cloned gene
pET plasmid
Chromosome
lacI

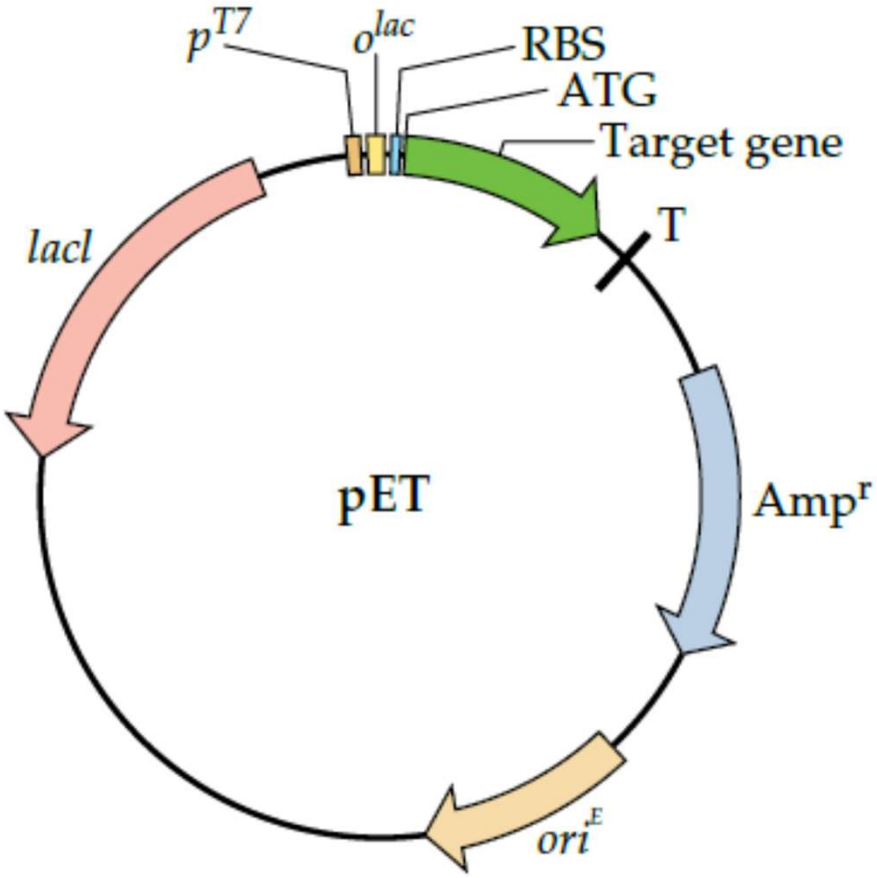
pET expression vectors

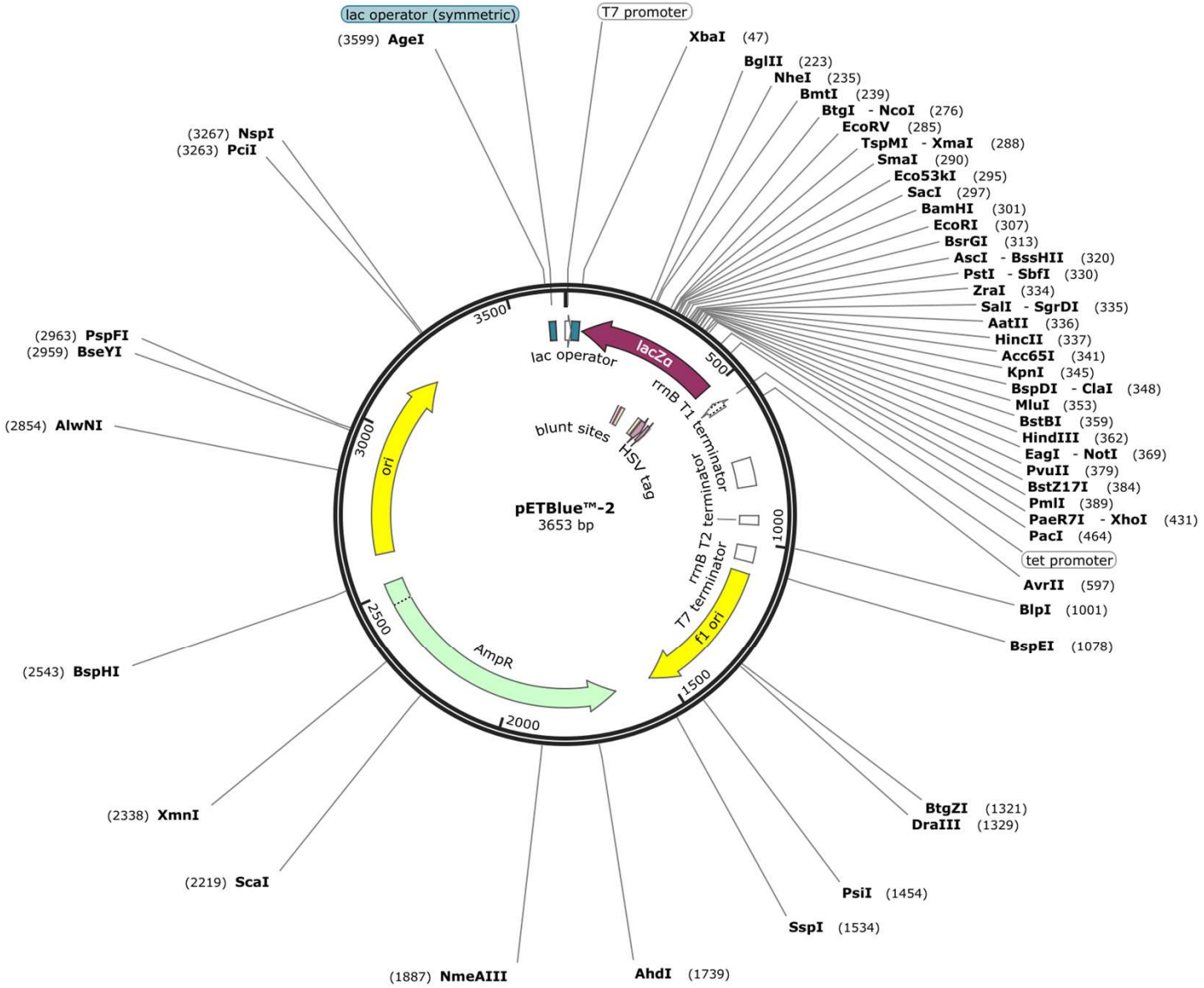
of 2 types:

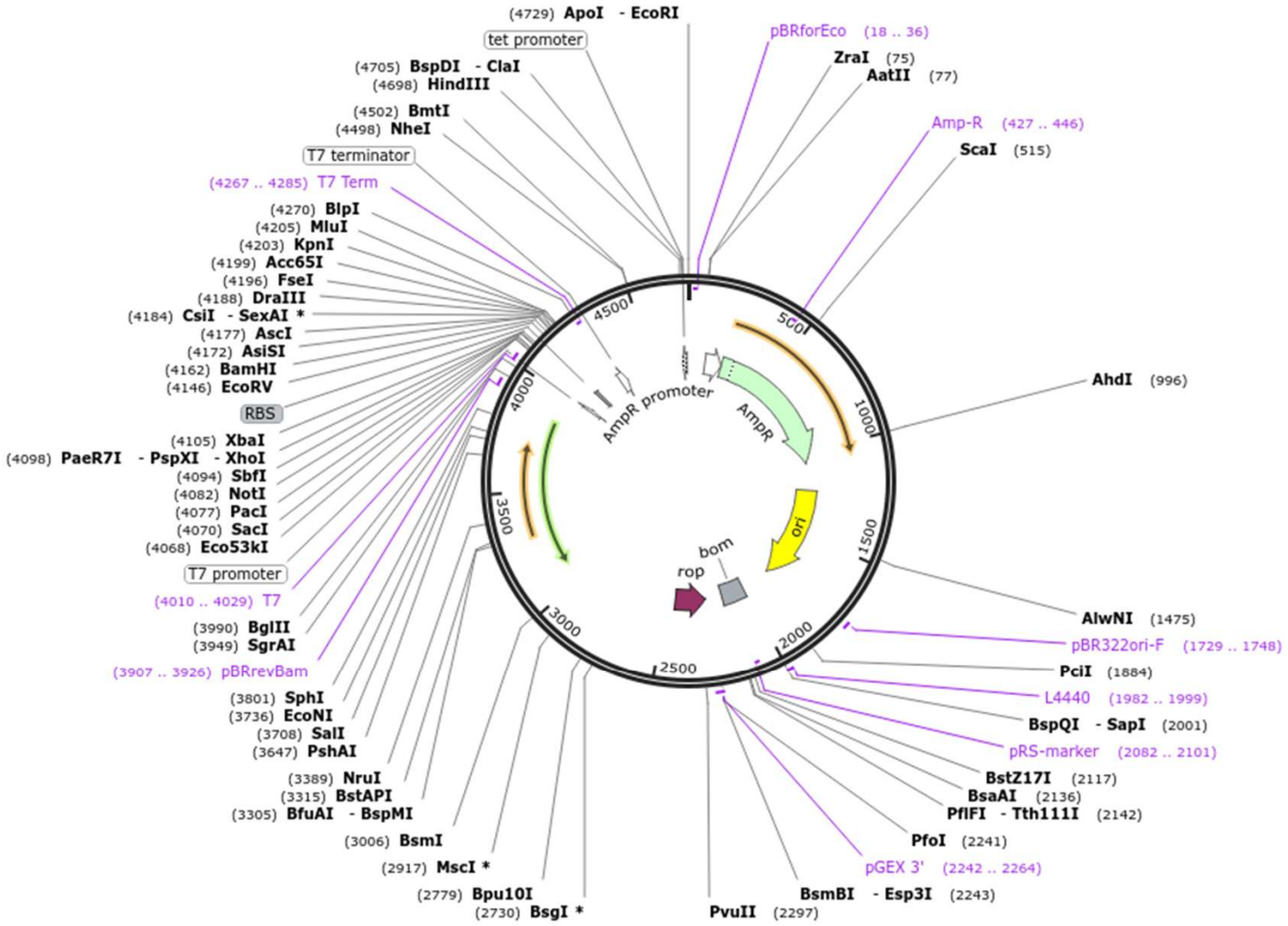
- Two different variants of the pET system are available: **transcriptional vectors** and **translational vectors** .
- Genes that carry their own ribosome binding site (RBS) and the ATG start codon are often cloned into transcriptional vectors . These types of vectors are generally used for cloning and expressing prokaryotic genes.
- Vectors carrying a ribosome binding site (RBS) from phage T7 are called translation vectors. These vectors are generally used for cloning and expressing eukaryotic genes.

pET expression vectors

of 2 types:



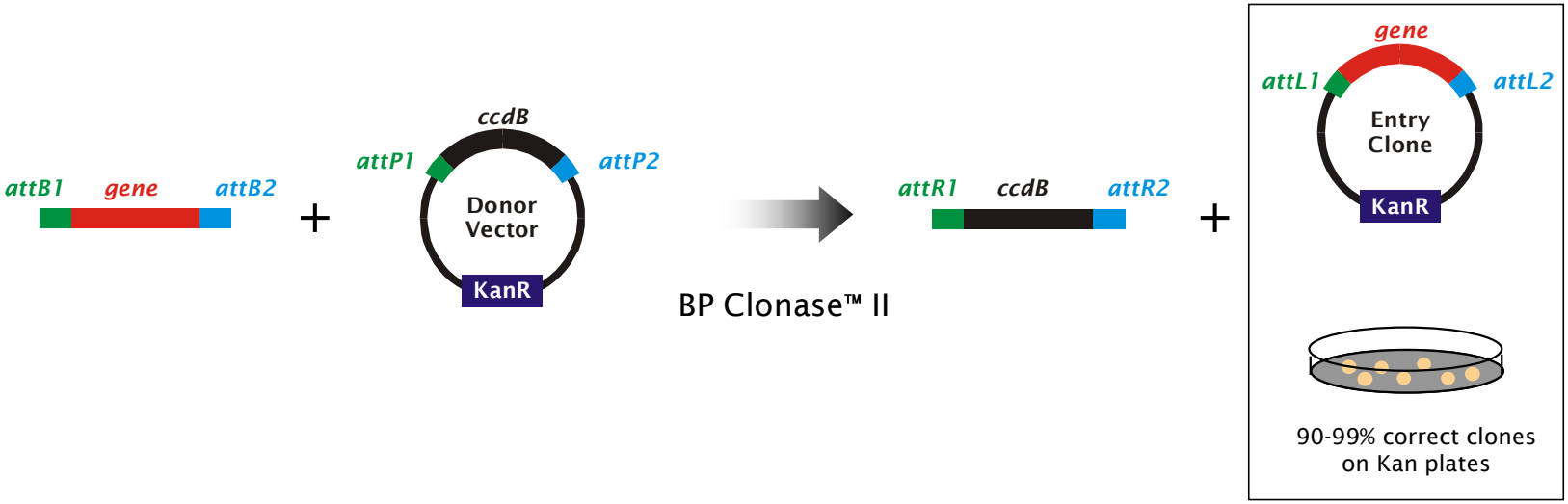




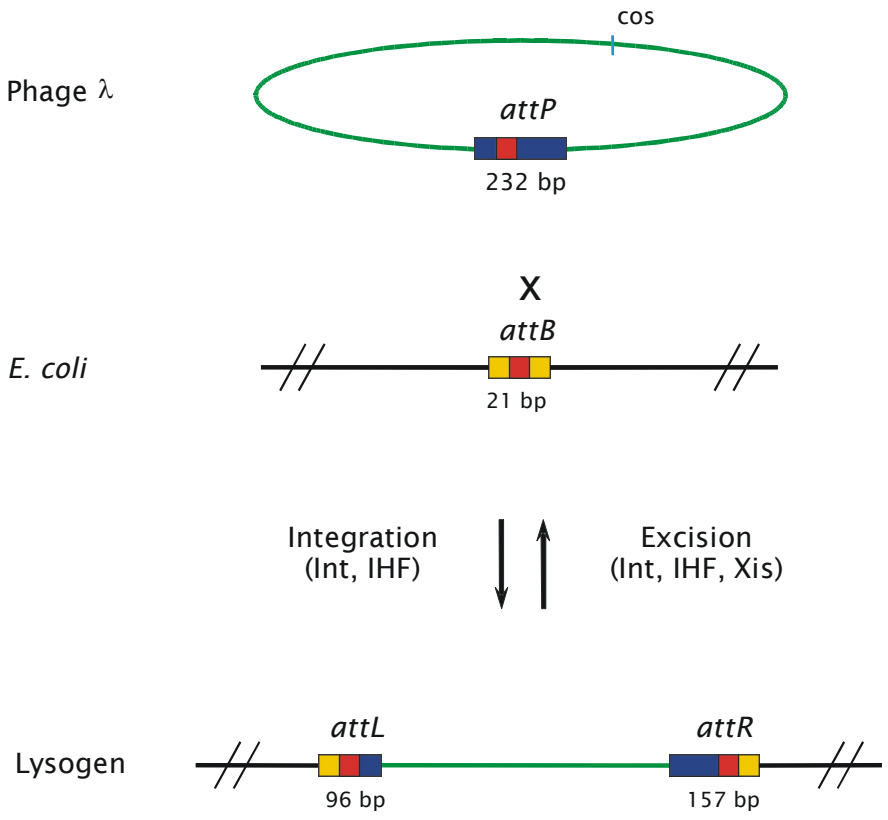
pET LIC cloning vector (2A-T)
4731 bp

Next generation vectors

Gateway® Systm



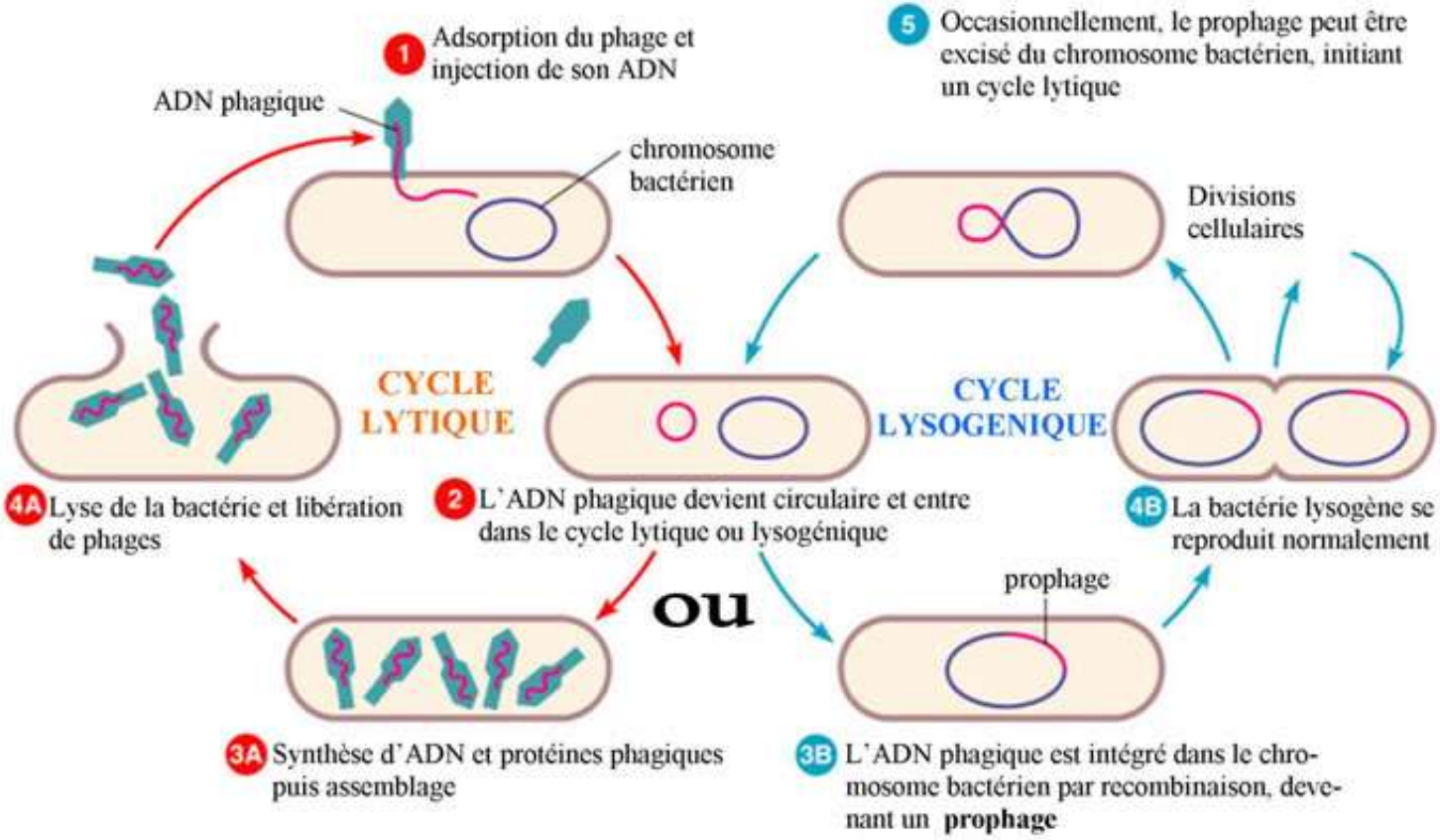
Next generation vectors



The Gateway® System is based on the insertion of a DNA fragment into a plasmid with 2 flanking sequences called "attL1" and "attL2".

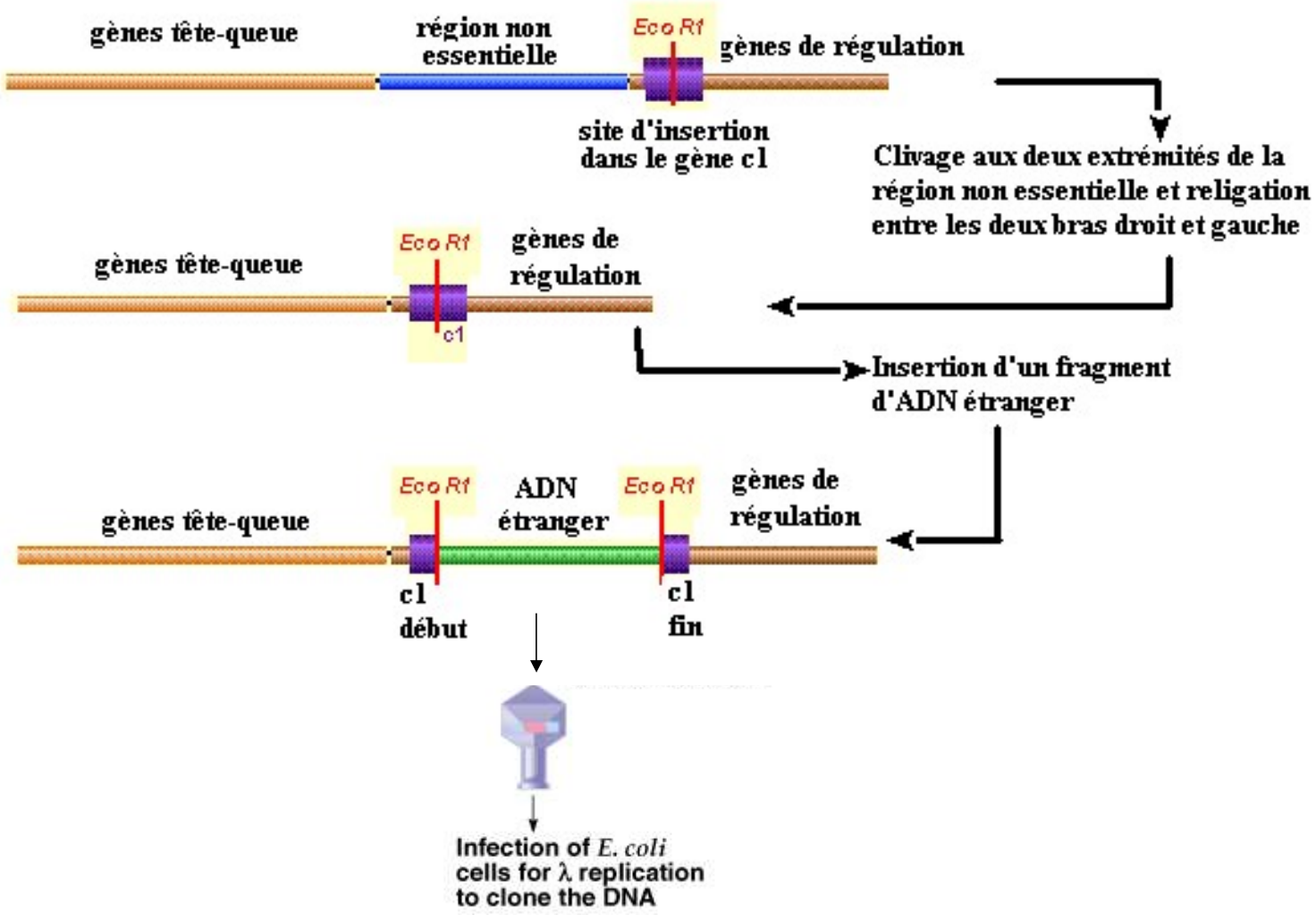
Bacteriophage ƛ

– The lytic cycle and the lysogenic cycle



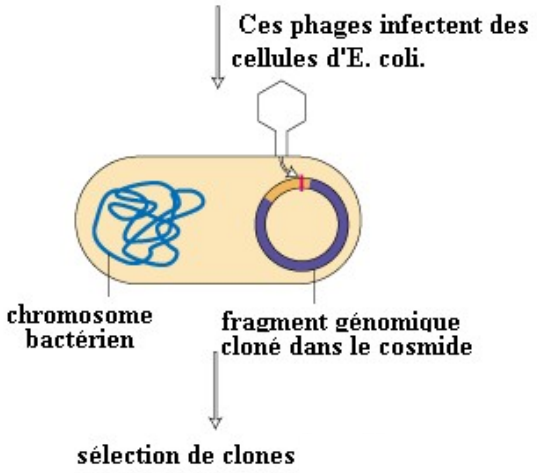
Vecteur d'insertion issu de lambda

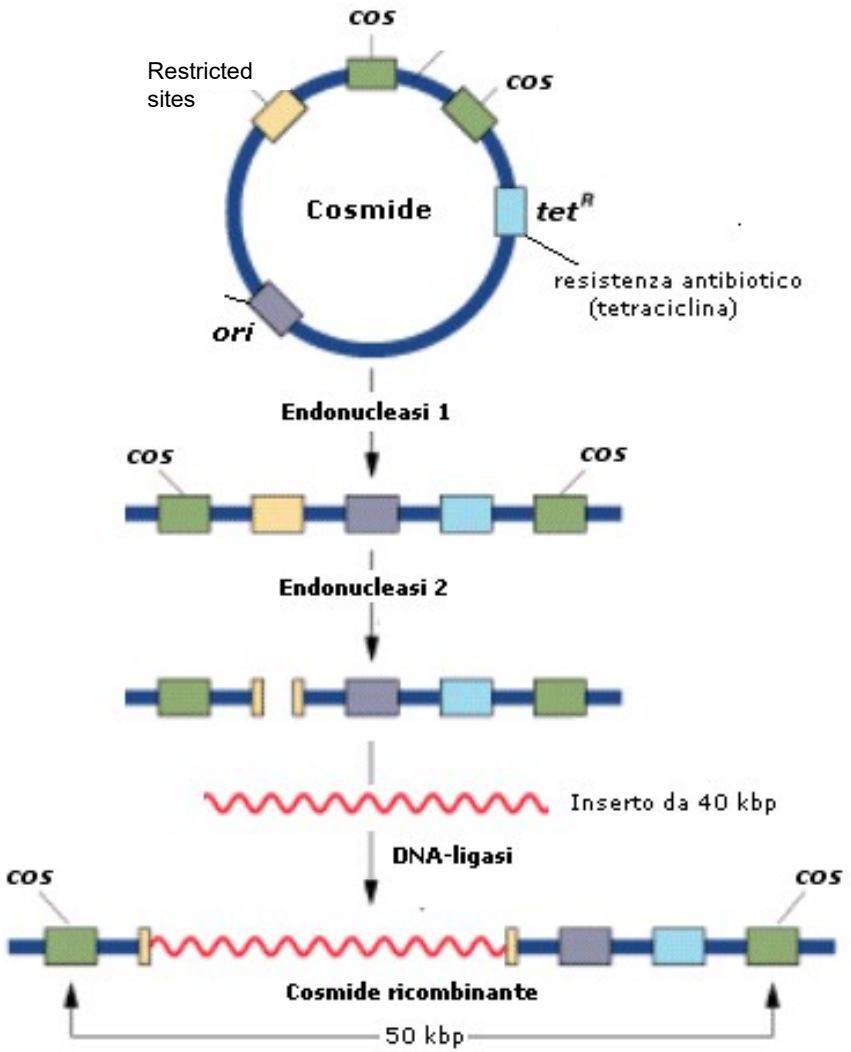
λ gt10



The cosmics

A cosmid is an artificial vector consisting of a hybrid plasmid containing the cos sequence of phage lambda.
Unlike plasmids, cosmids can be encapsidated into bacteriophage capsids via cos sites

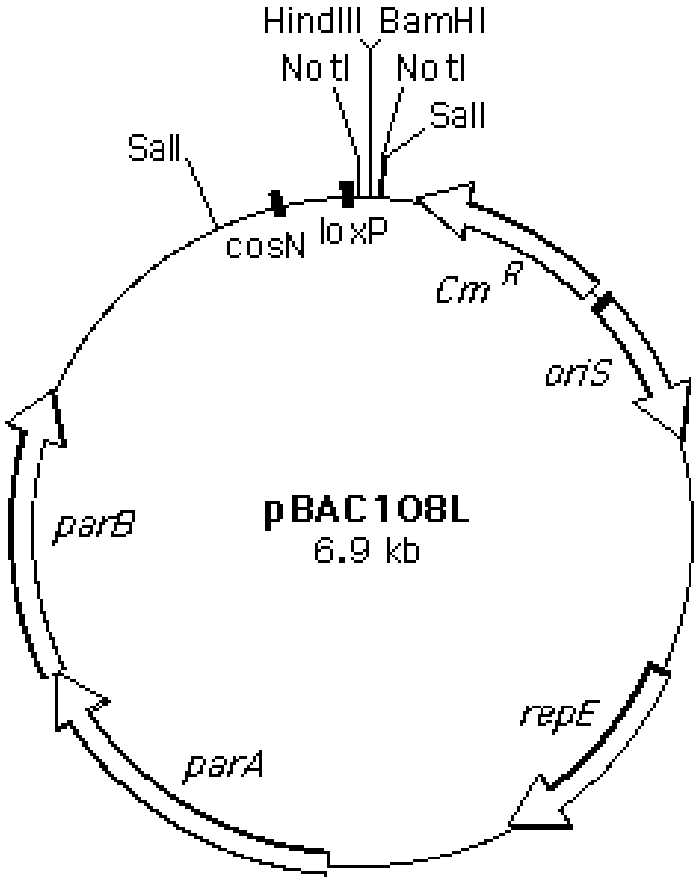




BAC Vectors

BACs are circular DNAs, like a bacterial chromosome, whose origin of replication comes from the F sex factor of *E. coli.*

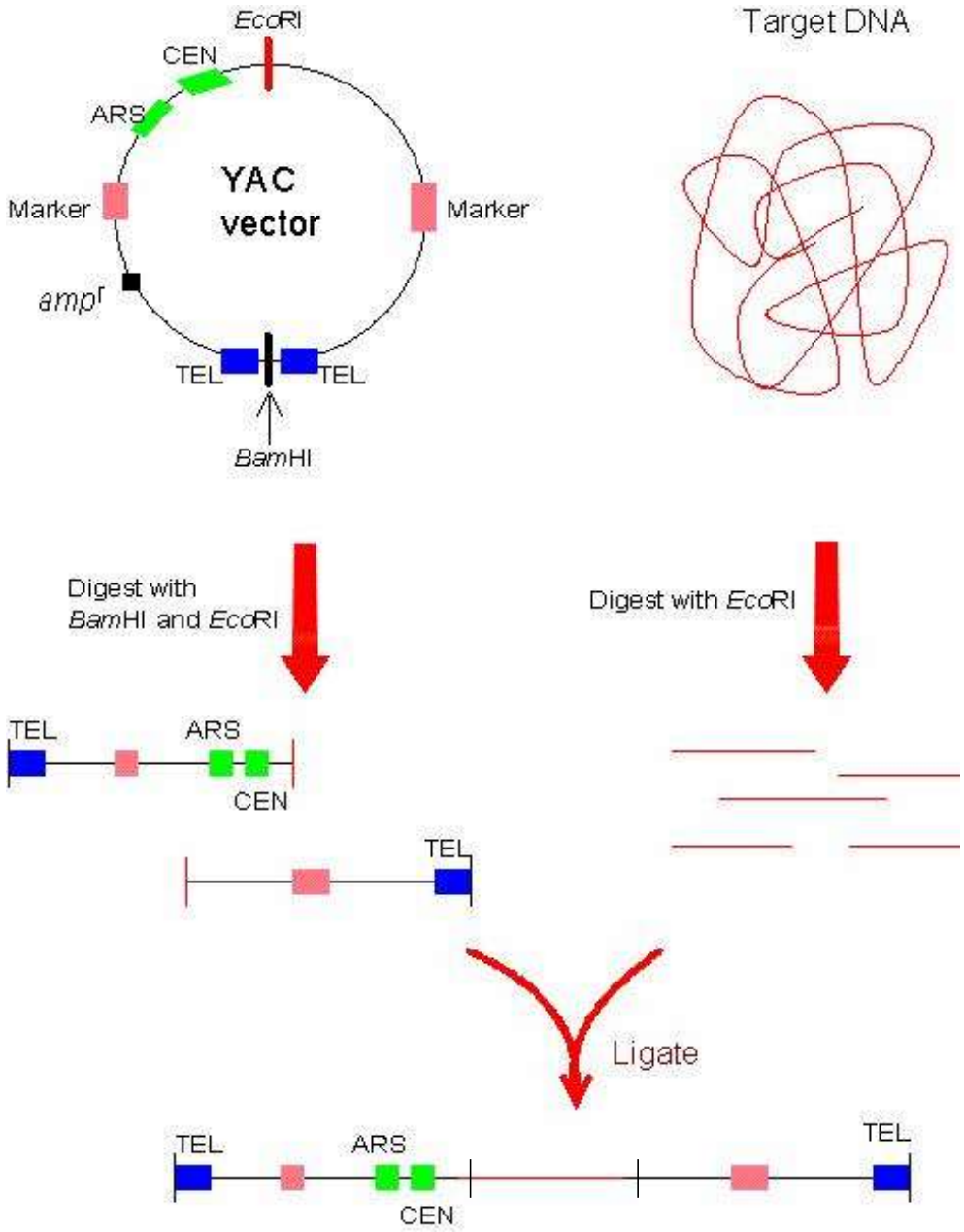
- OriS is the origin of replication of the F factor .
- The parB and parA genes, also of the F factor, are used for the correct distribution of the plasmid into daughter cells.
- The repE gene codes for the factor-specific replication enzyme.
- CosN comes from the lambda bacteriophage.
- Plasmid-derived CmR codes for resistance to chloramphenicol.



YAC Vectors

The essential components of YAC vectors

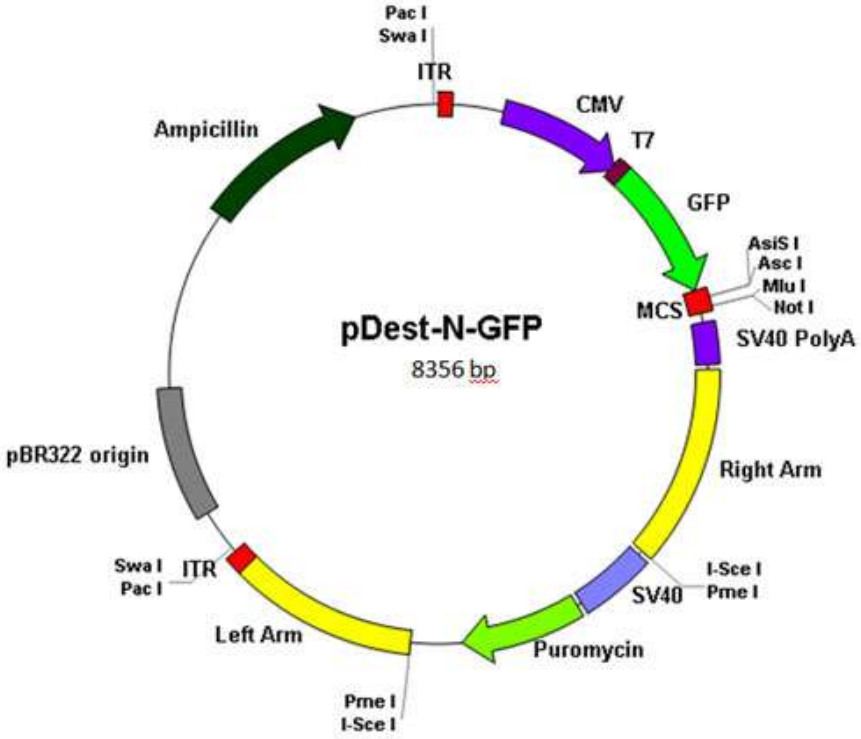
- Centromeres (CEN), telomeres (TEL) and an autonomous replication sequence (ARS) allow proliferation in the host cell.
- AmpR , along with the other two markers TRP1 and URA3, allows for the selective identification of cells containing the YAC vector.
- Restriction sites (e.g., EcoRI and BamHI)


EcoRI
CEN
ARS
YAC
vector
Marker
Marker
ampr
TEL
TEL
BamHI
Target DNA
Digest with
BamHI and EcoRI
Digest with EcoRI
TEL
ARS
CEN
TEL
Ligate
TEL
ARS
TEL
CEN

Viral Vectors

- Viral vectors are used to introduce new or modified genes into the genome of human and animal cells.
- Retroviruses are RNA viruses and are often used as cloning vectors.
- The viral RNA is converted into DNA by viral reverse transcriptase, and then efficiently integrated into the host genome.
- All genes introduced into the viral genome will be integrated into the host chromosome and can remain there virtually indefinitely.
- Retroviral vectors are widely used to study oncogenes of other and human genes.

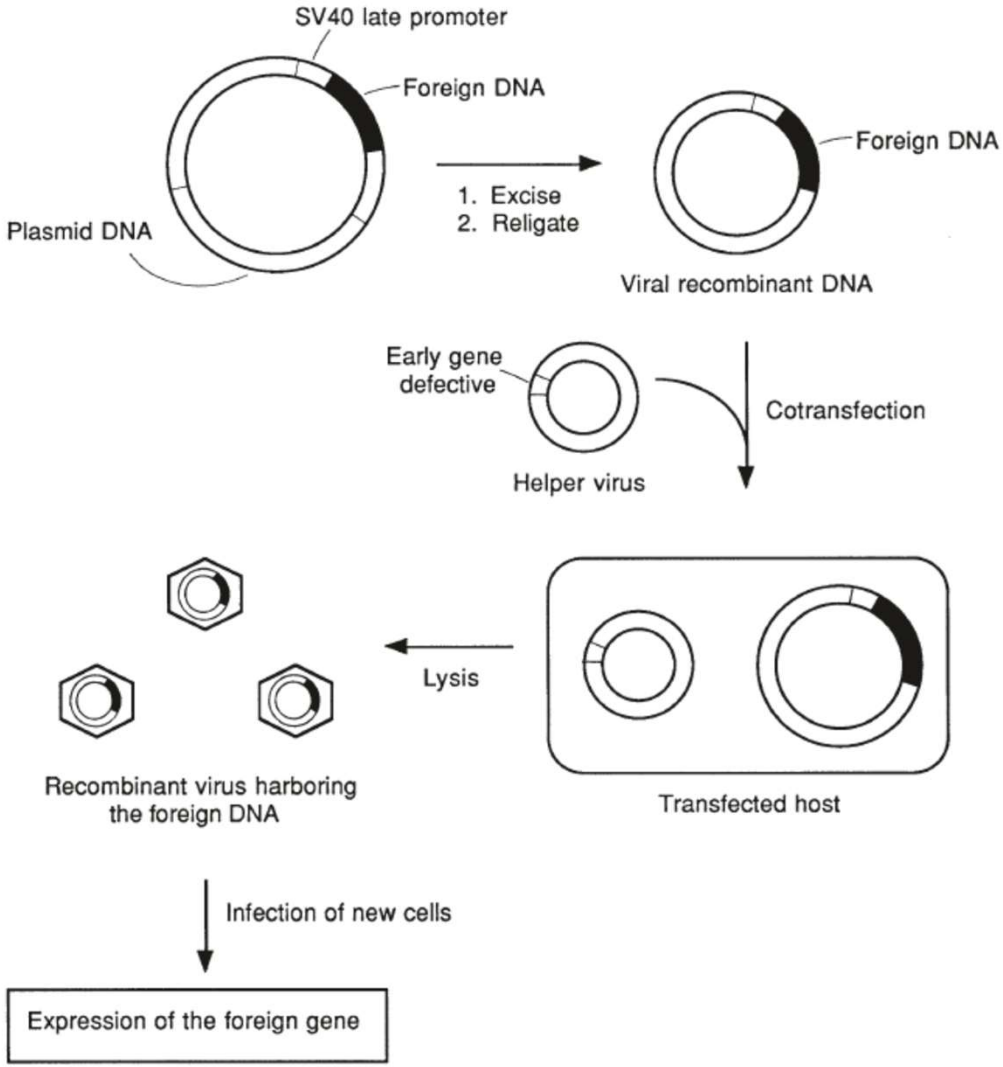
Viral vectors



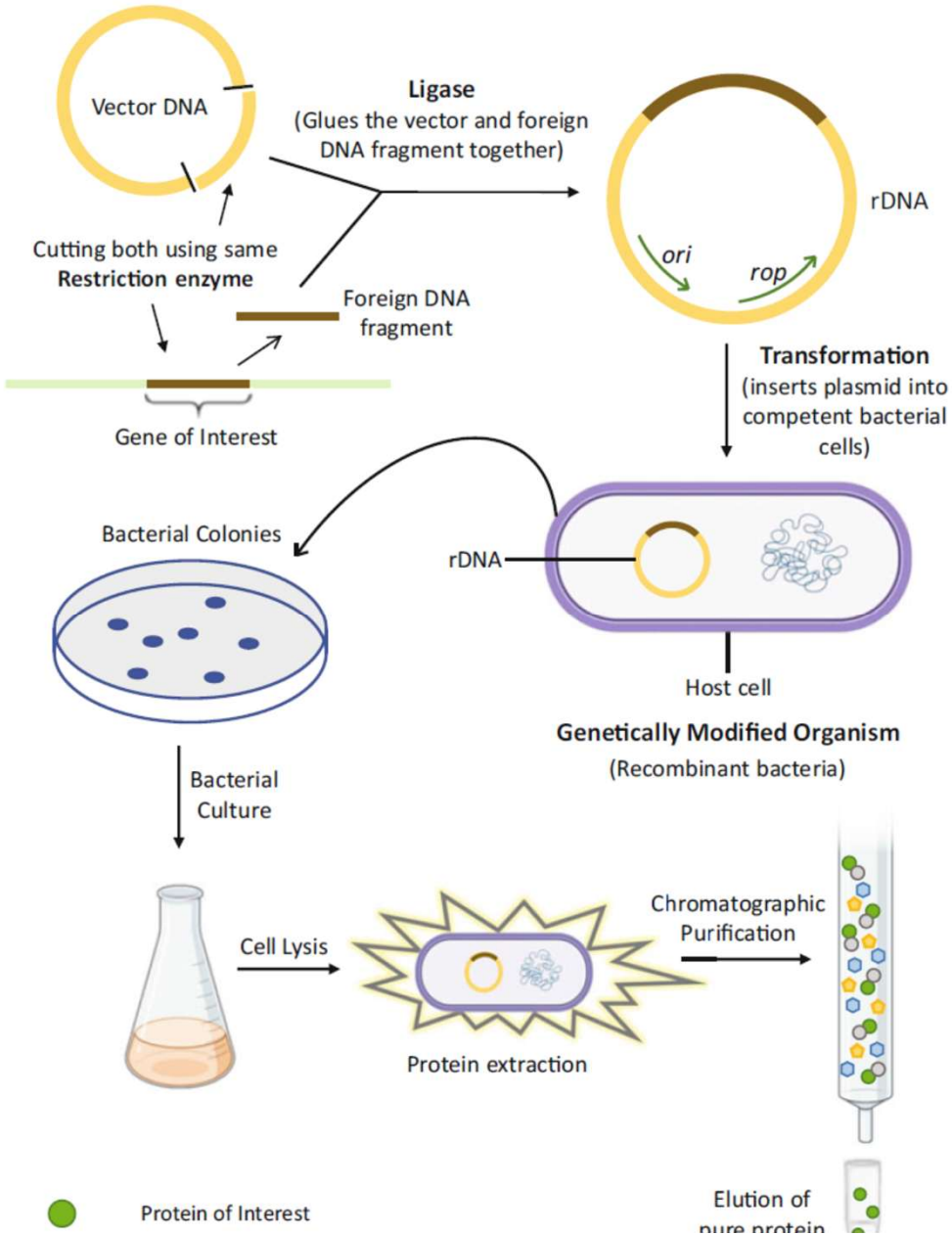
CMV: Cytomegalovirus

GFP: Green fluorescent protein

SV40 cloning strategy



The production of recombinant proteins of therapeutic interest


Vector DNA
Ligase
(Glues the vector and foreign DNA fragment together)
rDNA
ori
rop
Cutting both using same
Restriction enzyme
Foreign DNA fragment
Gene of Interest
Transformation
(inserts plasmid into competent bacterial cells)
Bacterial Colonies
rDNA
Host cell
Genetically Modified Organism
(Recombinant bacteria)
Bacterial Culture
Cell Lysis
Protein extraction
Chromatographic Purification
Protein of Interest
Elution of

INTRODUCTION

- Many proteins that can be used for medical treatment or research are normally expressed at very low concentrations .

- All these therapeutic agents can be produced in large quantities and also more economically using recombinant DNA technology .

- Thanks to this technology, a large quantity of proteins can be produced. This involves inserting the gene for the desired protein into an "expression vector" which must contain a promoter so that the protein can be expressed .

INTRODUCTION

- Recombinant DNA technology is widely used in the production of therapeutic agents such as ;
  - hormones, cytokines, growth factors, antibiotics, vaccines, blood products such as albumin, thrombolytics, fibrinolytic agents, coagulation factors such as factor VII, factor IX, tissue plasminogen activator and many others.

Producing recombinant proteins : The different expression systems

The production of recombinant proteins in:

- Microbial bioreactors
  - *E. coli* expression system
  - *Saccharomyces cerevisiae*
- Mammalian-derived cellular bioreactors
  - Eg Chinese Hamster Ovary cell (CHO).
- Animal bioreactors
  - Production of therapeutic recombinant proteins in the milk of transgenic animals. For example, cows, sheep, etc.

Characteristics of the different expression systems

Characteristic	*E. coli*	Yeasts	Mammals	Insects
Proteolytic cleavage	?	?	Yes	Yes
Glycosylation	No	?	Yes	?
Section	?	Yes	Yes	Yes
Folding	?	?	Yes	Yes
Phosphorylation	No	?	Yes	?
Acetylation	No	Yes	Yes	?
Amidation	No	Yes	Yes	Yes
% yield	>50%	1%	<1%	>30%

Microbial bioreactors

- The first microbial bioreactors, in particular *Escherichia coli* (bacteria) and *Saccharomyces cerevisiae* and *Pichia pastoris* (yeasts), were found to be satisfactory for the production of simple polypeptides such as insulin and human growth hormone.

- However, microbial bioreactors are not suitable for proteins with complex post-translational modifications or complex folding requirements, such as coagulation factors, or monoclonal antibodies.

A fermentor used to grow recombinant bacteria.

Animal bioreactors

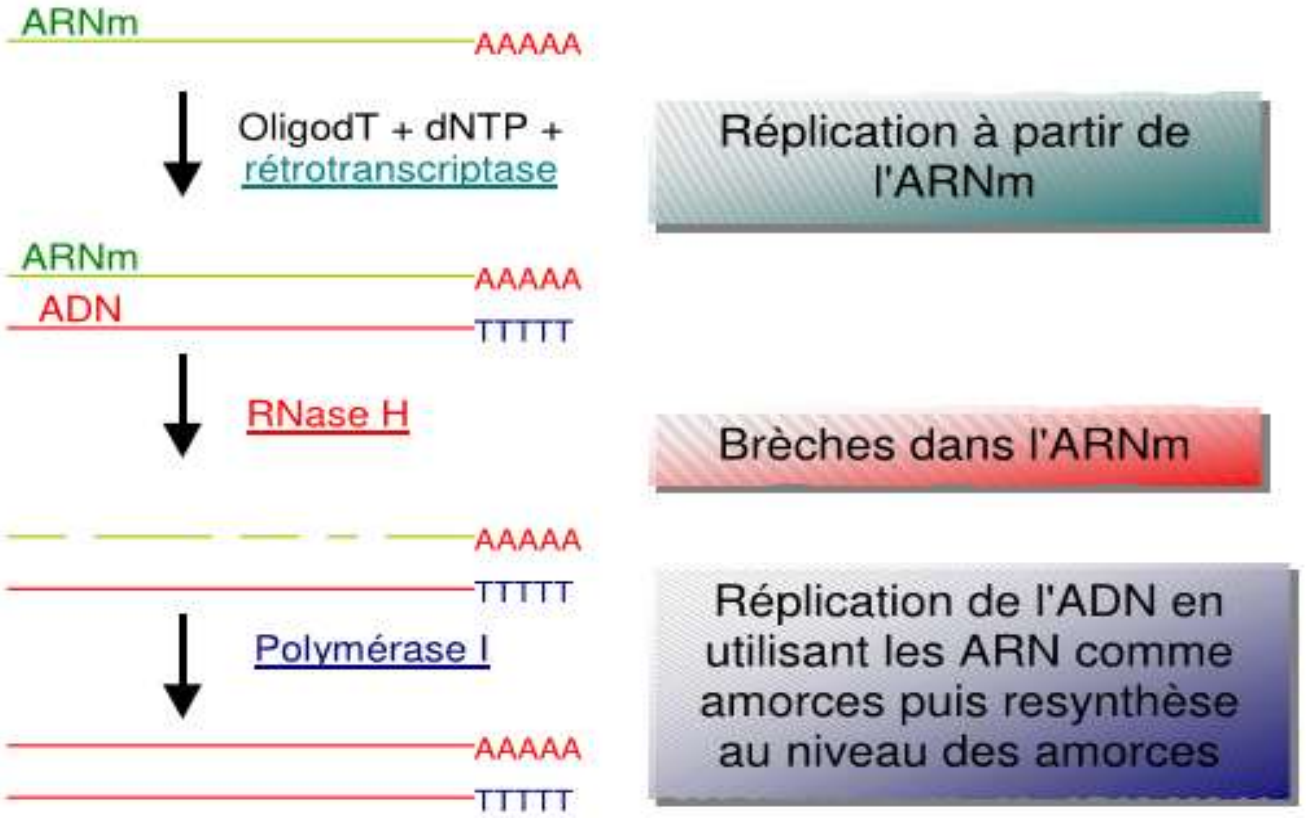
Expression of mammalian genes in bacteria

• It is possible to achieve very high levels of mammalian gene expression in prokaryotes.

•However, the expressed gene must be intron-free.

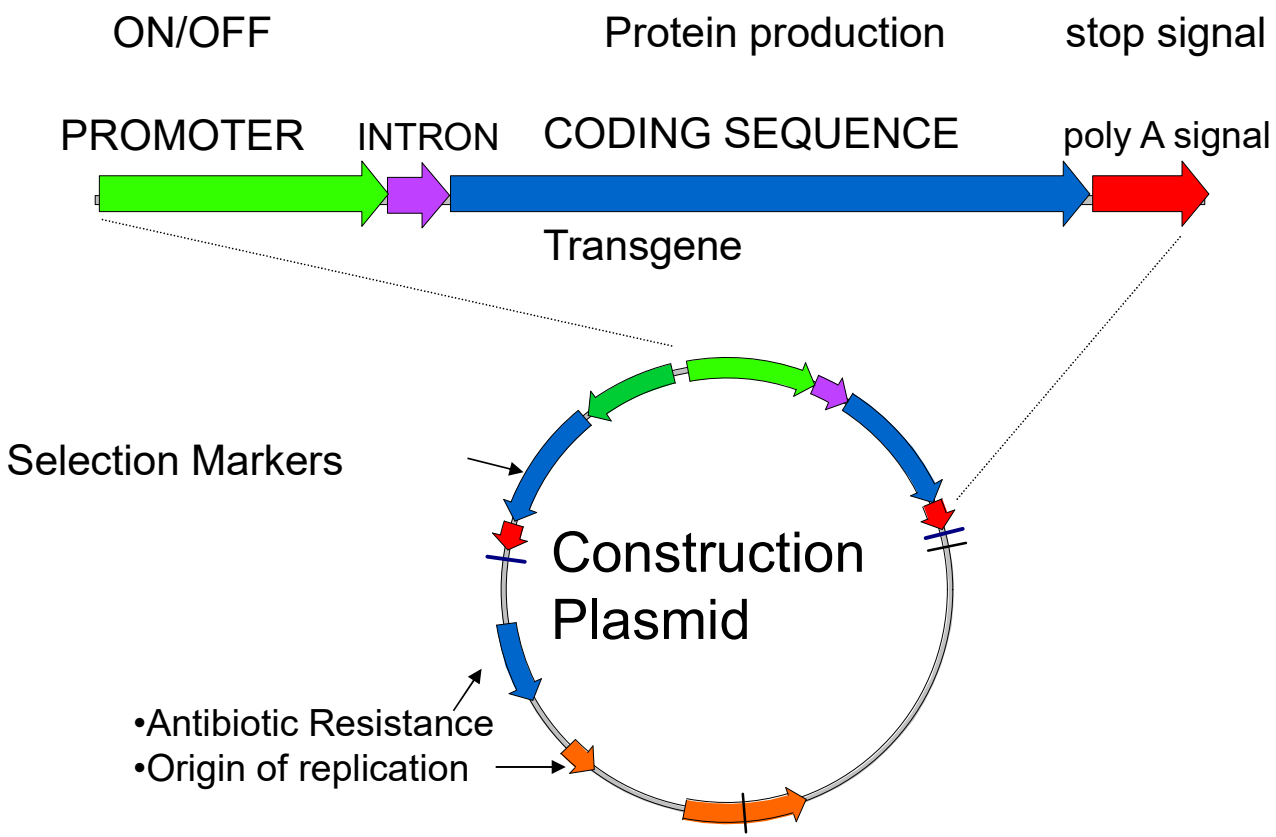
Expression of mammalian genes in bacteria

• This can be accomplished by using a reverse transcriptase to synthesize cDNAs from the mature mRNA encoding the protein of interest.



Expression of mammalian genes in bacteria

- The cDNA is then cloned into a plasmid.



Expression of mammalian genes in bacteria

•The amino acid sequence of a protein can also be used to design and synthesize an oligonucleotide sequence that codes for it. This process is called 'reverse translation'.

Expression of mammalian genes in bacteria

- Example: human insulin,

The first human protein produced through genetic engineering using bacteria was human insulin.

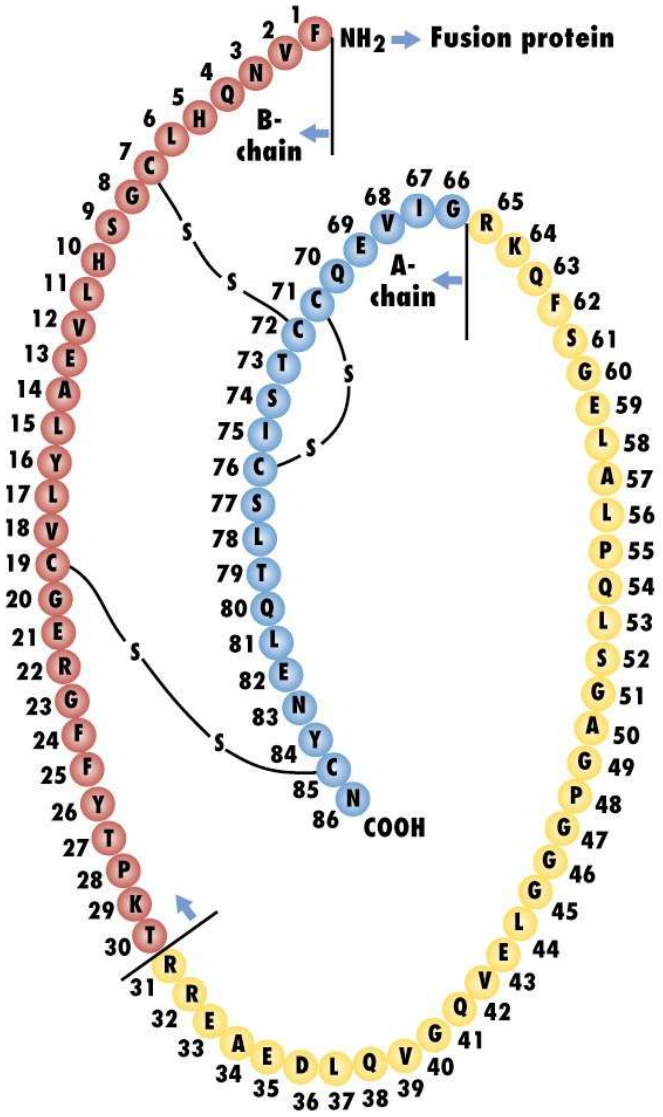
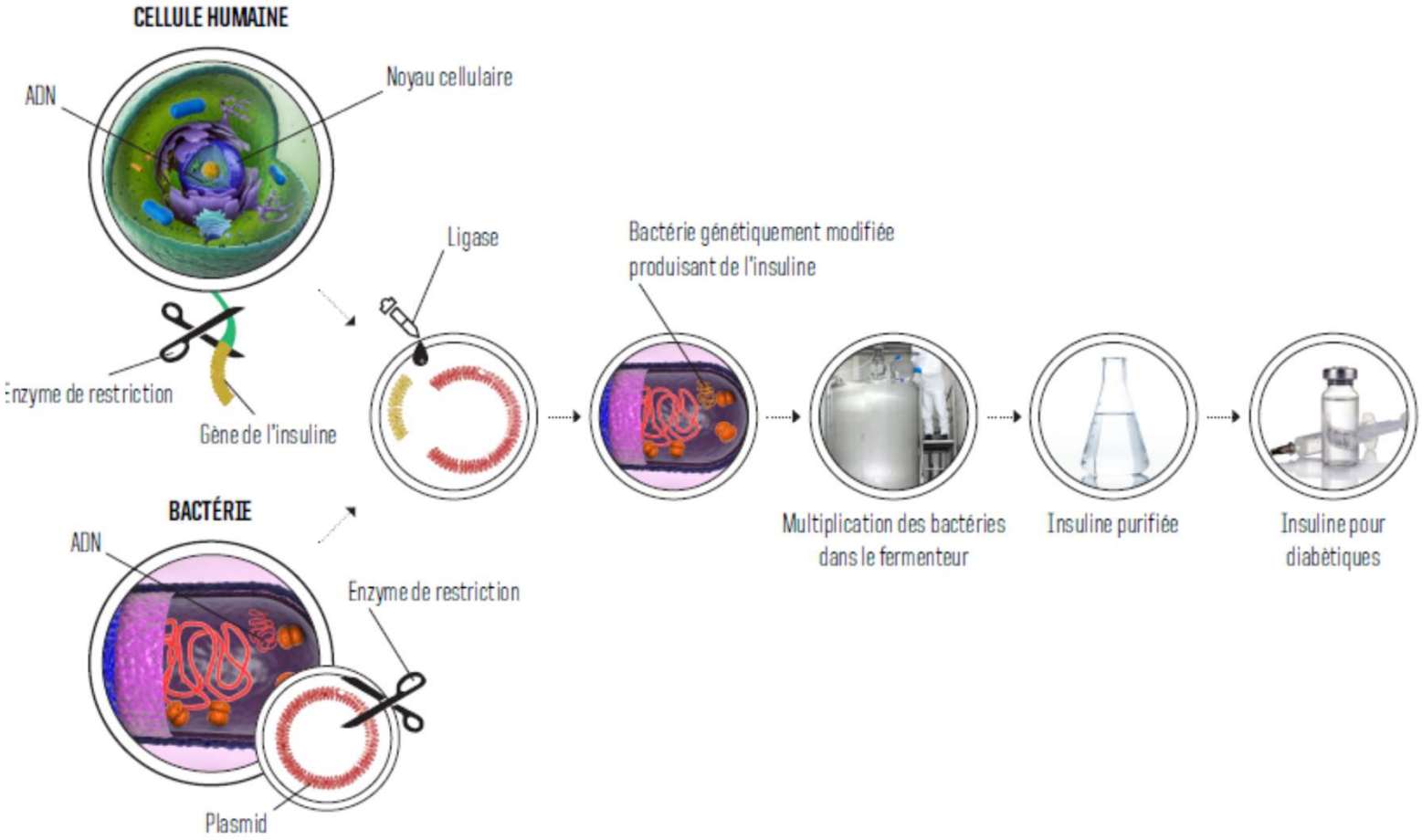


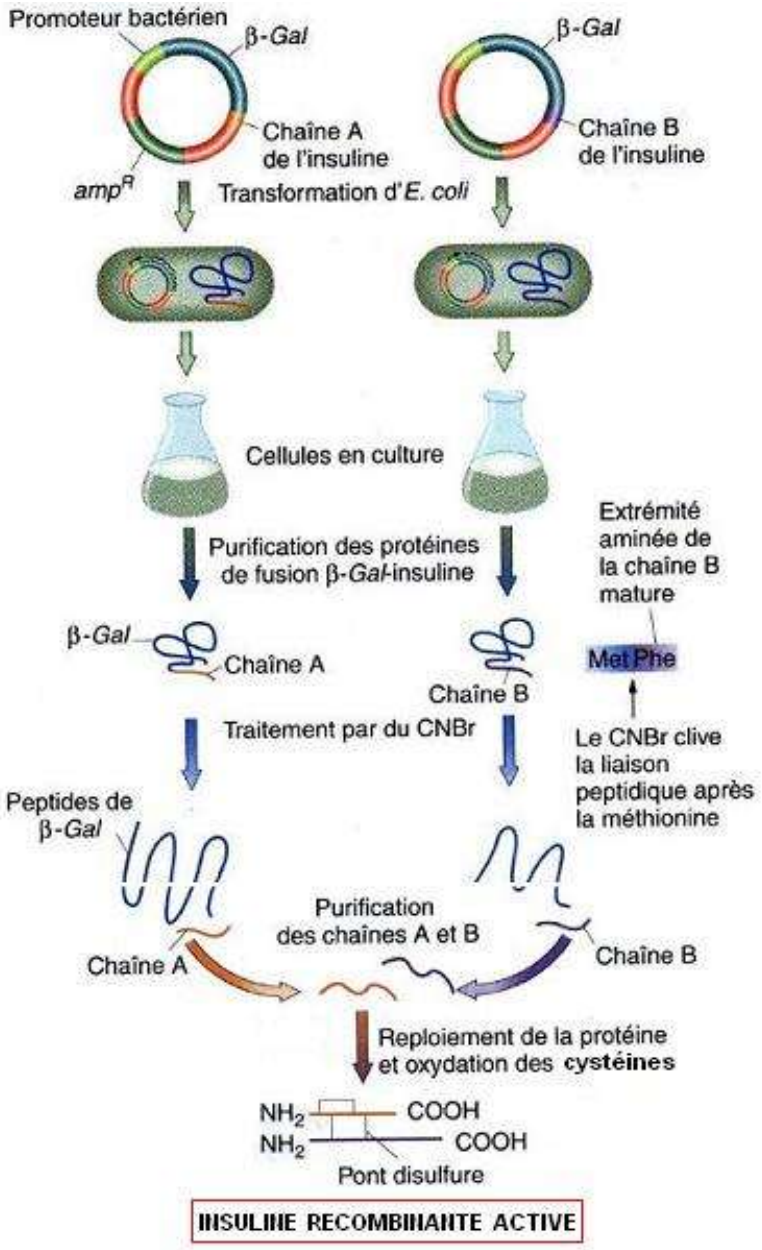
Figure 31-11a Brock Biology of Microorganisms 11/e

We can start with mRNA isolated from tissues that produce the insulin protein. We then use the reverse transcriptase enzyme to reverse transcribe the mRNA into cDNA, which would be intron-free. We can ligate the sticky ends of the cDNA with a plasmid vector.



Production of insulin

1) Diagram of the recombinant insulin production strategy used in 1982 by Eli Lilly



Insulin production

2) Production of Proinsulin:

Production in E. coli.

Synthesis of oligonucleotides encoding the 86 amino acids of proinsulin.

Proinsulin folds spontaneously with the formation of disulfide bonds

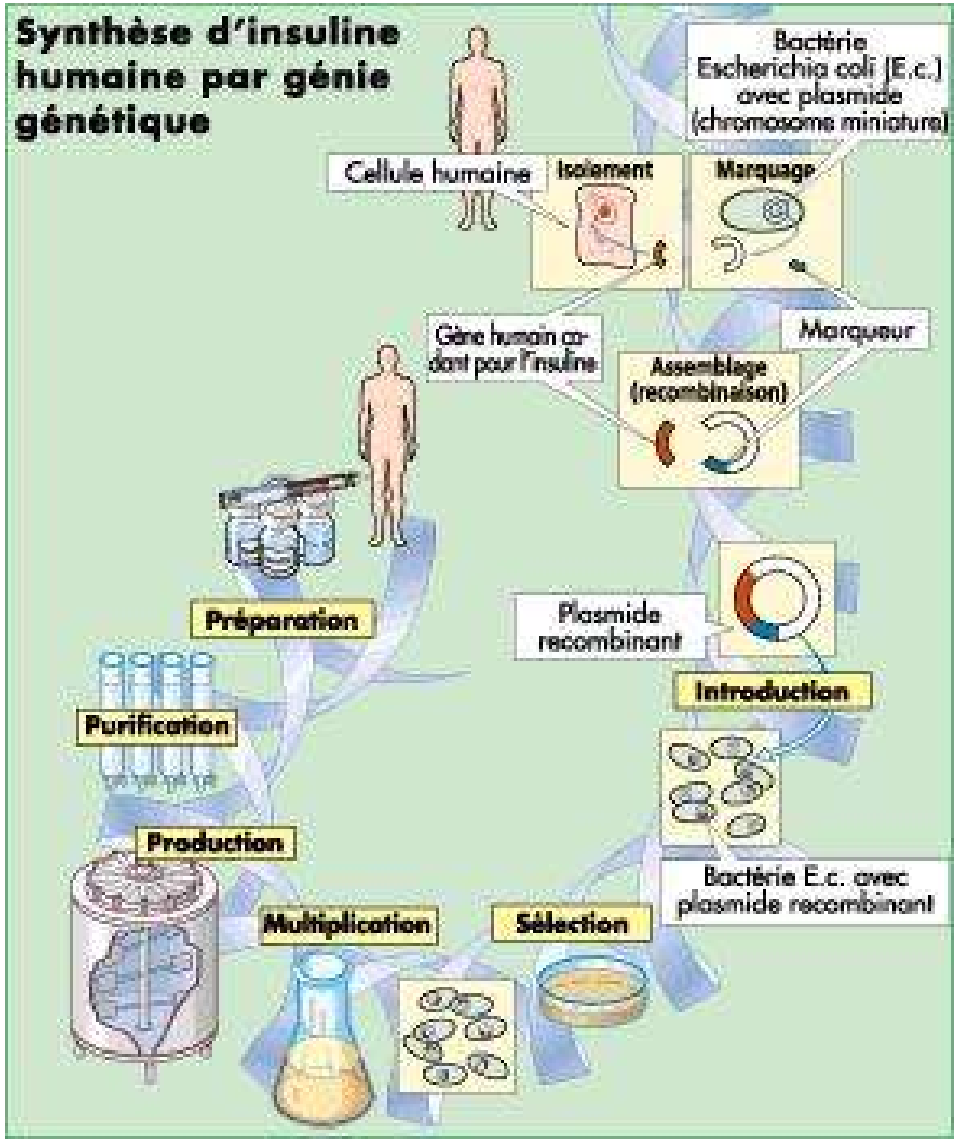
A protease cleaves the C peptide

Final insulin 61aa

Benefits :

- a single production line

- Simplified procedures, and no fusion protein



Production of human growth hormone