CLONING & EXPRESSION OF RECOMBINANT PROTEIN

Dr. Saeideh Sadat Shobeiri

PhD of Immunology

Outline

- **Cloning**

✓ Restriction enzymes

✓ Vectors

✓ Transformation

- **Expression**

✓ Expression Systems

✓ Tags and fusion proteins

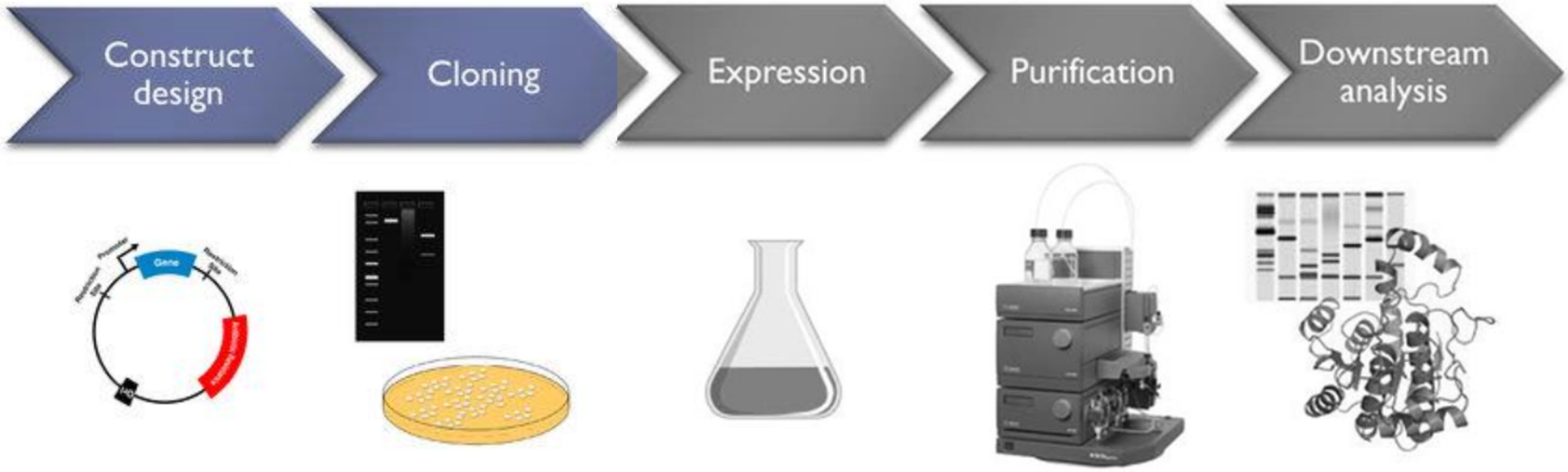
✓ Bacterial host (advantages & dis)

✓ Protein validation

✓ Protein purification

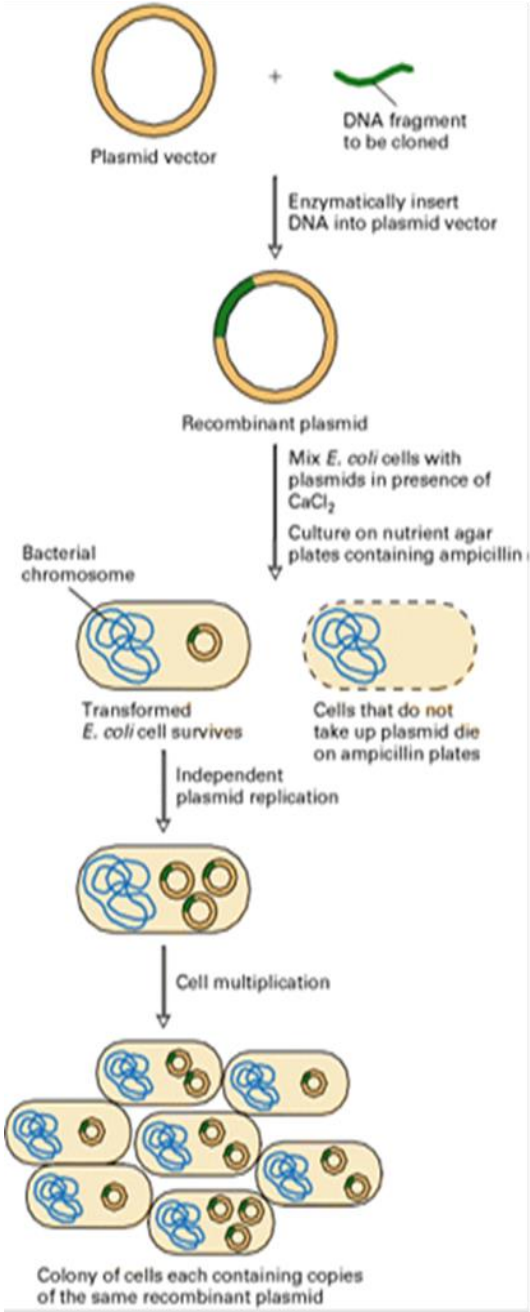
✓ Protein concentration

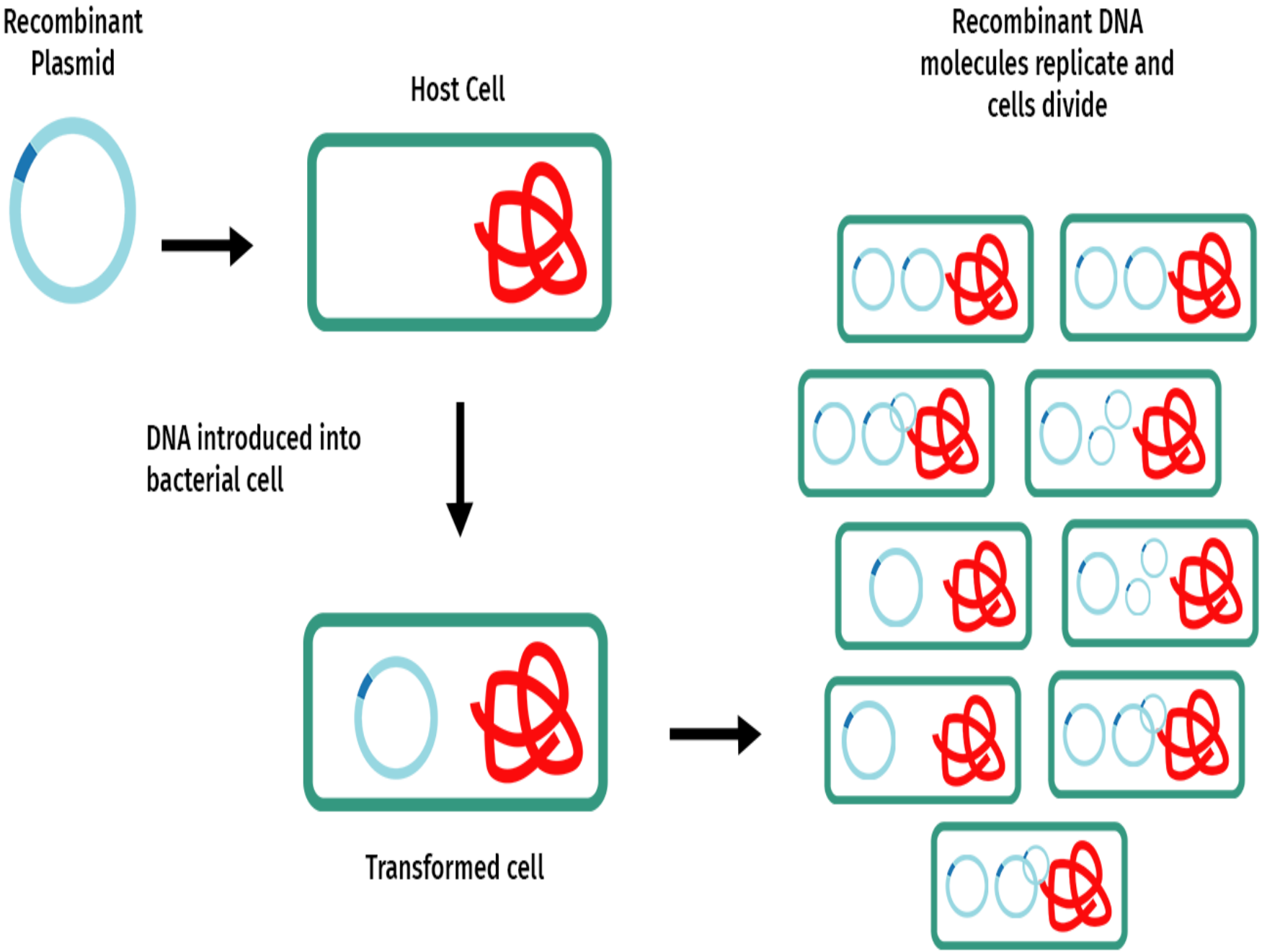
Cloning


Construct design
Cloning
Expression
Purification
Downstream analysis

Defining Cloning

- Cutting a piece of DNA from one organism and inserting it into a vector where it can be replicated by a host organism. (Sometimes called subcloning, because only part of the organism's DNA is being cloned) and later expressing in either prokaryotic or eukaryotic cells.

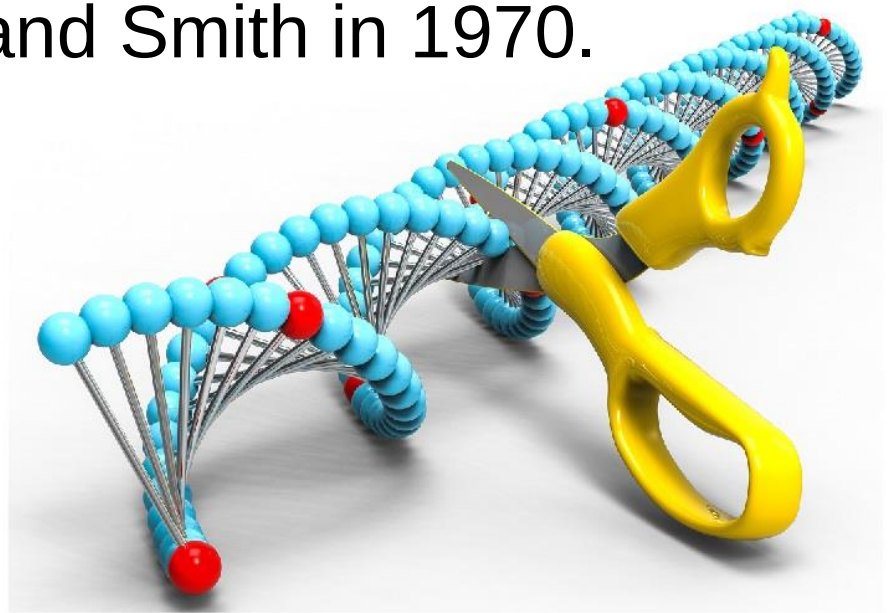



Recombinant Plasmid
Host Cell
Recombinant DNA molecules replicate and cells divide
DNA introduced into bacterial cell
Transformed cell

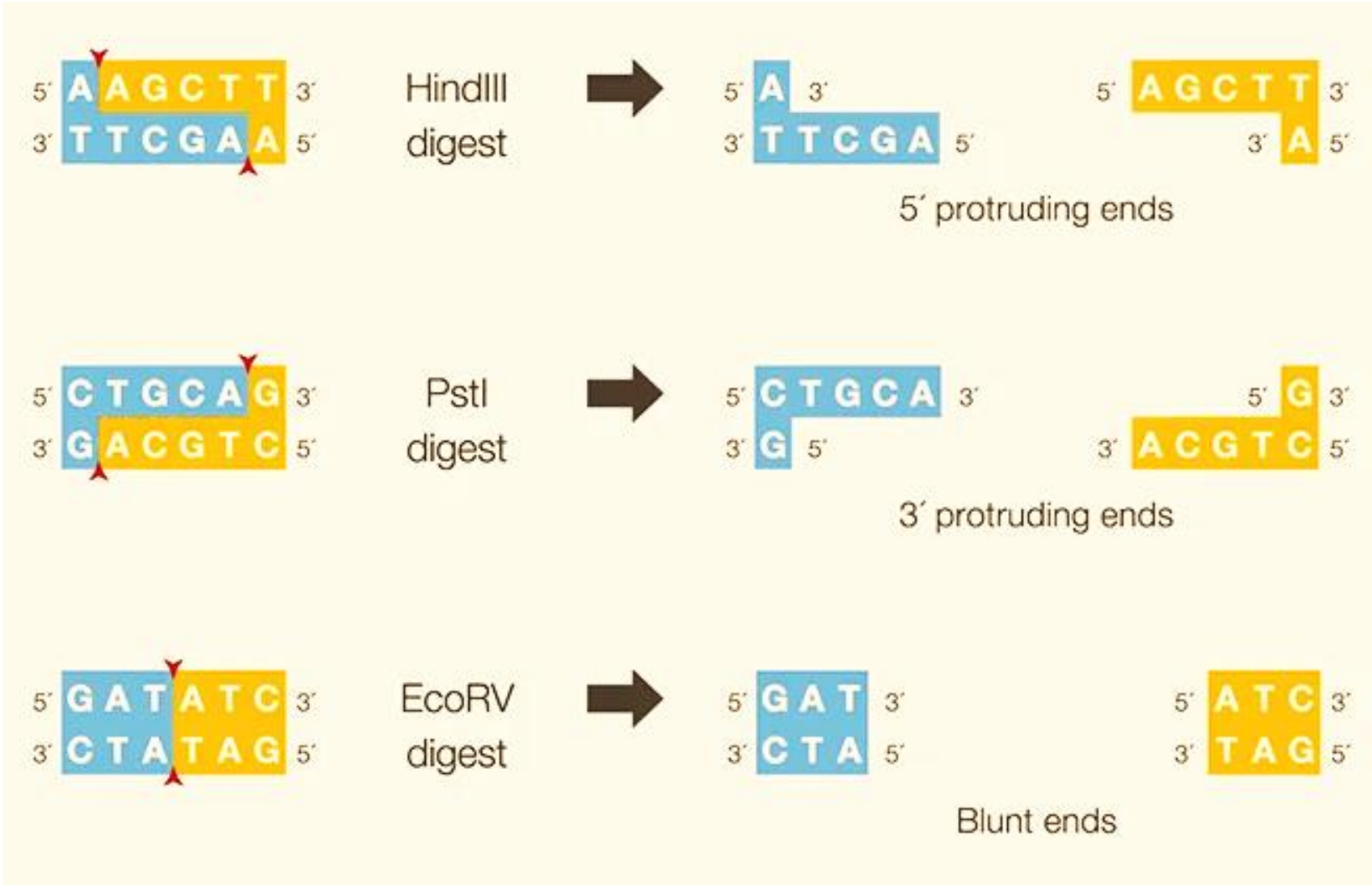
Restriction Enzymes

Restriction Enzymes (also called Restriction Endonucleases) are proteins that cleave DNA molecules at specific sites, producing discrete fragments of DNA.

Restriction Enzymes (RE) were first isolated by Nathans and Smith in 1970.



Restriction Enzymes



Why Restriction Enzymes?

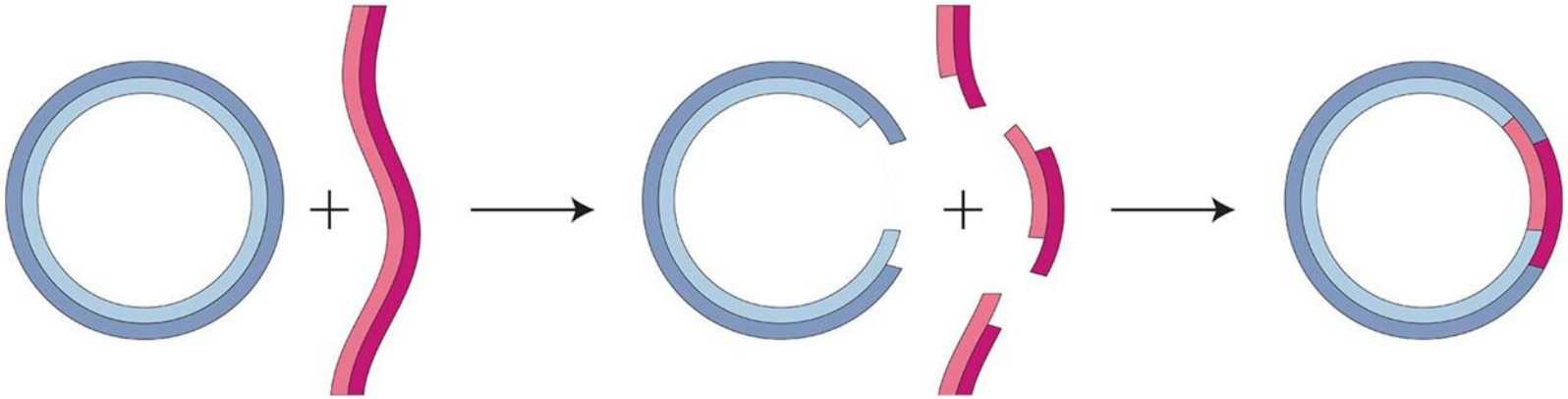
- Why would bacterial cells contain proteins that cleave DNA at specific sequences?
  - Generally restriction enzymes are thought to protect bacterial cells from phage (bacterial virus) infection. Bacterial cells that contain restriction enzymes can “cut up” invasive viral DNA without damaging their own DNA.

Joining DNA Fragments

- In 1972, Paul Berg and colleagues made the first “artificial” recombinant DNA molecule.
- Demonstrated that the DNA of Simian virus 40 could be linearized by EcoR1
- Created a circular DIMER of Simian virus DNA by joining two linearized fragments
- Also inserted pieces of Lambda phage DNA into linearized Simian 40 virus molecule.

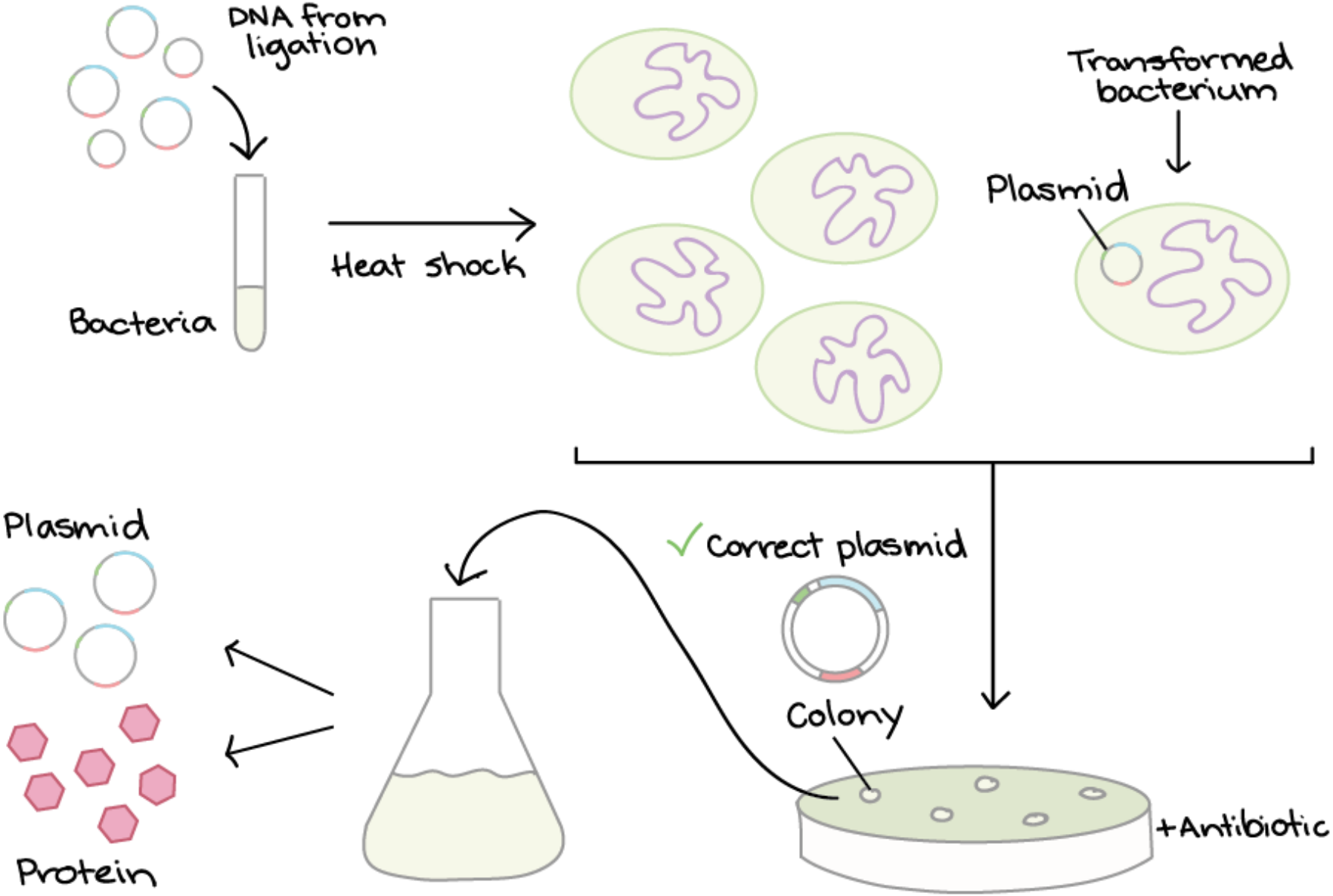
Isolating Genes

- Herbert Boyer and Stanley Cohen built on the work of Berg, Nathans and Smith to use restriction enzymes to isolate a single gene, place it into a plasmid vector.
- Bacterial cells were then transformed with the recombinant plasmid.
- The bacteria host cells replicated the plasmid, producing many copies of the gene, thus amplifying it.



Advantages

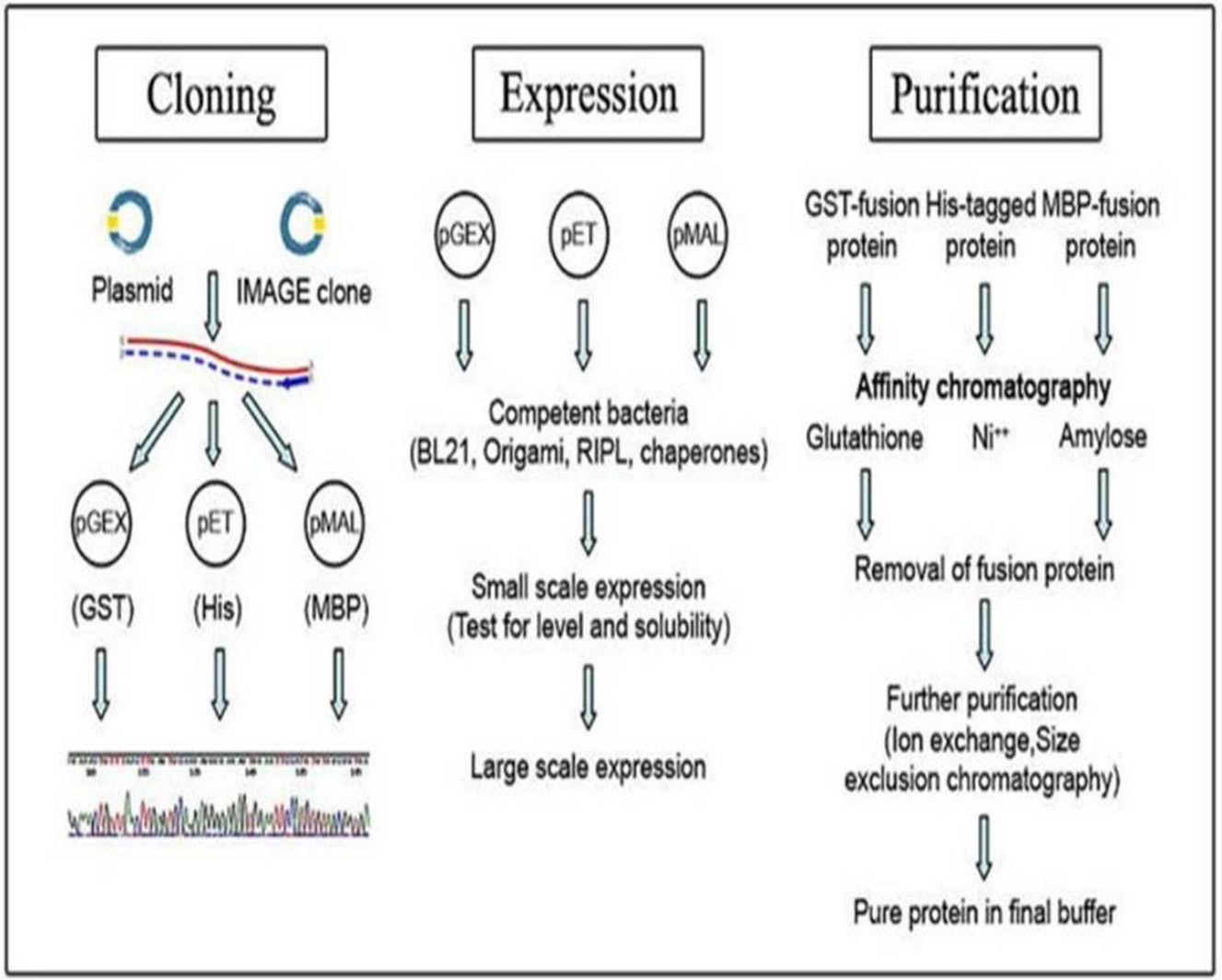
- The practical application was that expensive human protein products, like insulin, which were used to treat disease, could eventually be produced from recombinant molecules in the laboratory using bacteria or another host.
- Human protein products like insulin could be used in very large quantities from the recombinant molecule. Patients no longer had to use insulin isolated from pigs or cows.


DNA from ligation
Bacteria
Heat shock
Transformed bacterium
Plasmid
Correct plasmid
Colony
+Antibiotic
Plasmid
Protein

vectors

Vector	Insert size (kb)
Plasmid	<10 kb
Bacteriophage	9 – 15 kb
Cosmids	23 – 45 kb
BACs	≤ 300 kb
PACs	100 – 300 kb
YACs	100 – 3000 kb

vectors

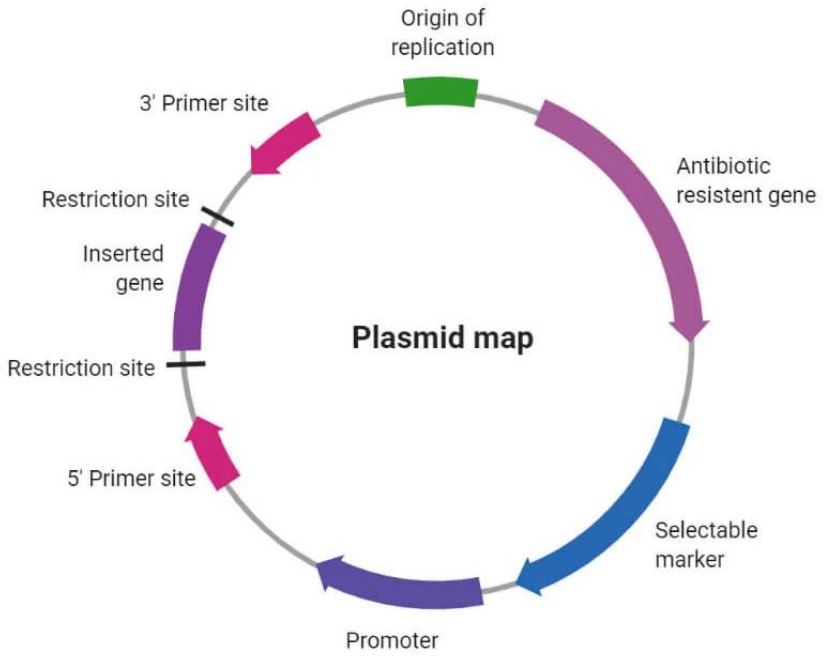

Cloning
Expression
Purification
Plasmid
IMAGE clone
pGEX
pET
pMAL
(GST)
(His)
(MBP)
pGEX
pET
pMAL
Competent bacteria
(BL21, Origami, RIPL, chaperones)
Small scale expression
(Test for level and solubility)
Large scale expression
GST-fusion protein
His-tagged protein
MBP-fusion protein
Affinity chromatography
Glutathione
Ni++
Amylose
Removal of fusion protein
Further purification
(Ion exchange,Size exclusion chromatography)
Pure protein in final buffer

Plasmid Vectors

- Plasmids are circular pieces of DNA found naturally in bacteria.
- Plasmids can carry antibiotic resistance genes, genes for receptors, toxins or other proteins.
- Plasmids replicate separately from the genome of the organism.
- Plasmids can be engineered to be useful cloning vectors.

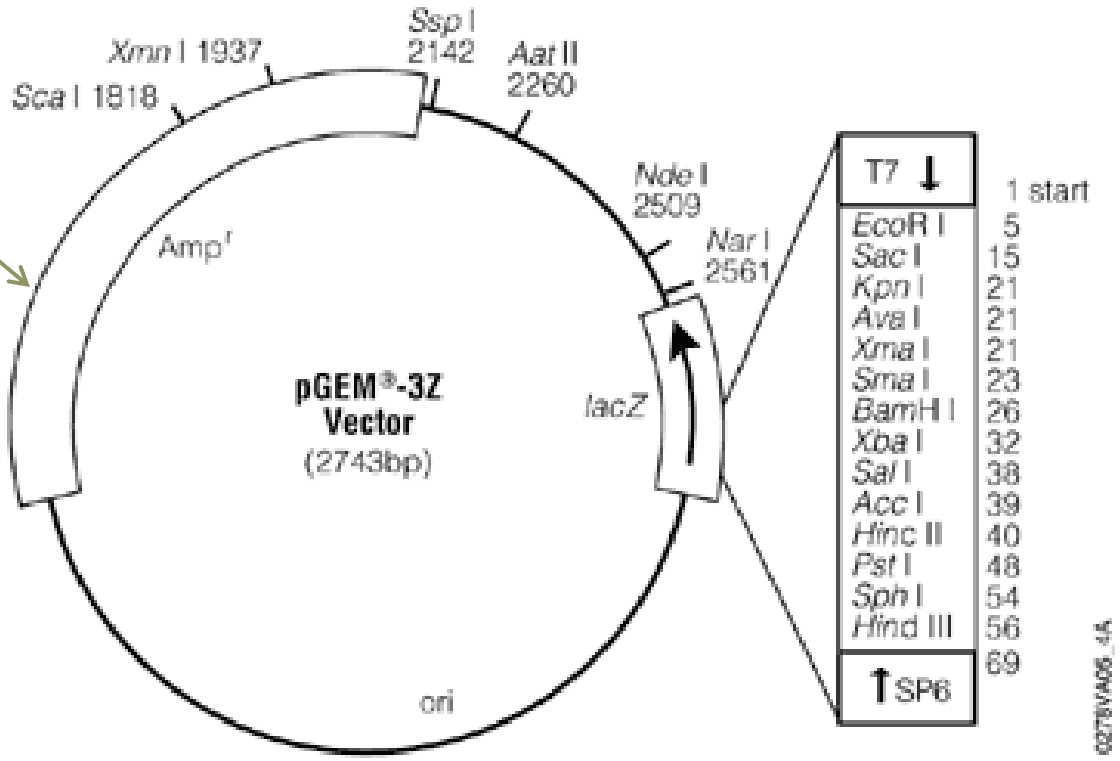
Plasmid Vectors (continued)

- Plasmid vectors can be designed with a variety of features:
  - Antibiotic resistance
  - Colorimetric “markers”
  - Strong or weak promoters for driving expression of a protein

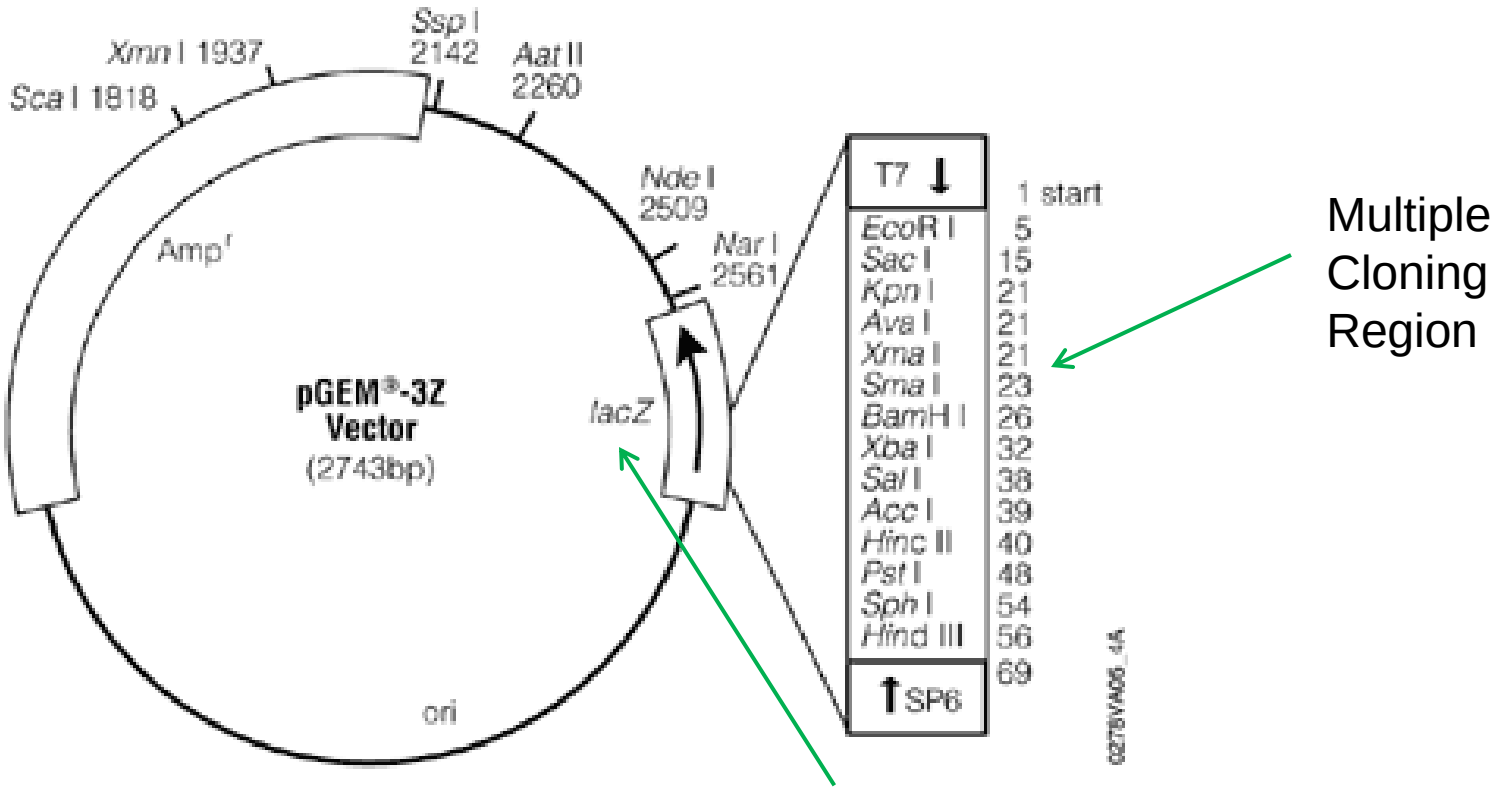


Antibiotic Resistance Markers

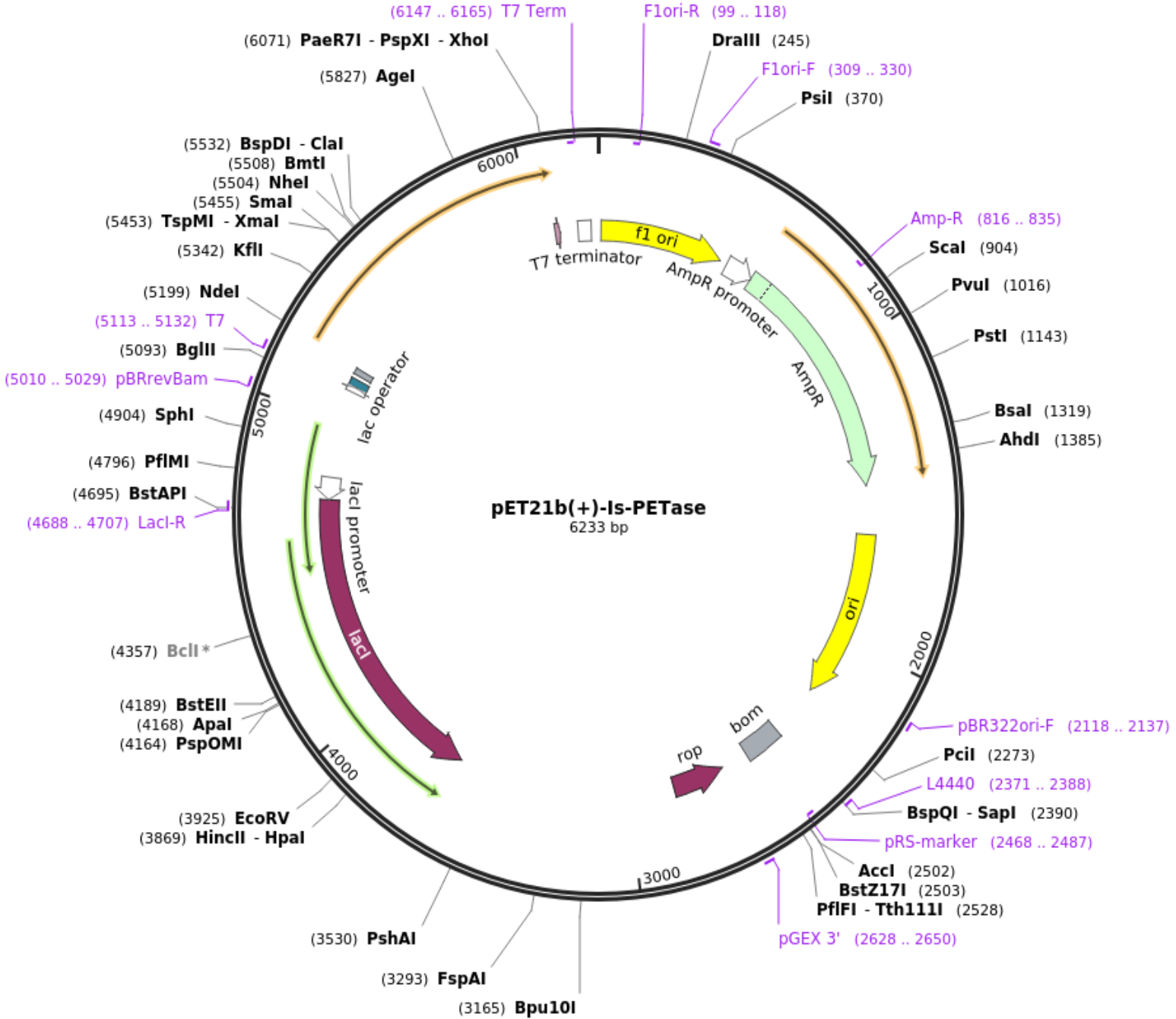
Antibiotic Resistance Gene

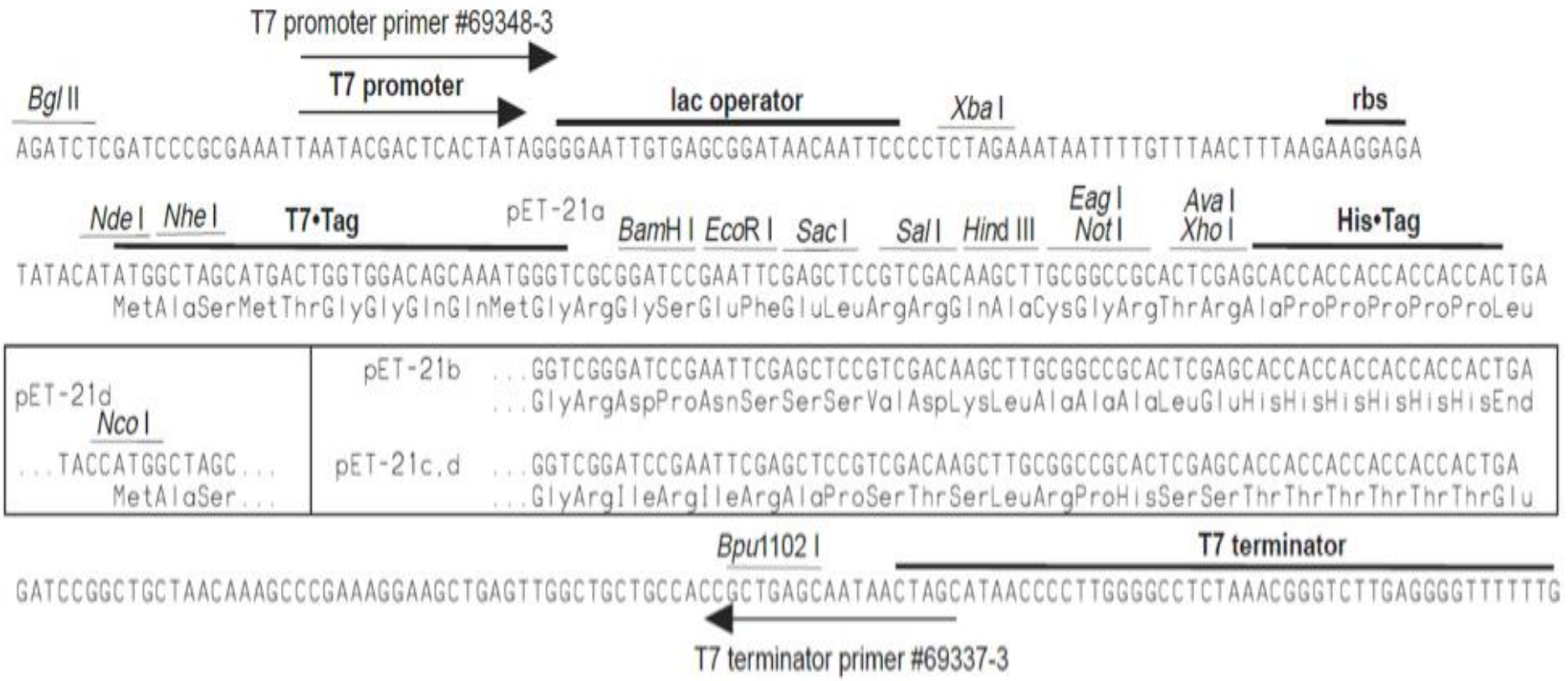


Multiple Cloning Region



The cloning marker for this plasmid is the *lac*Z gene.



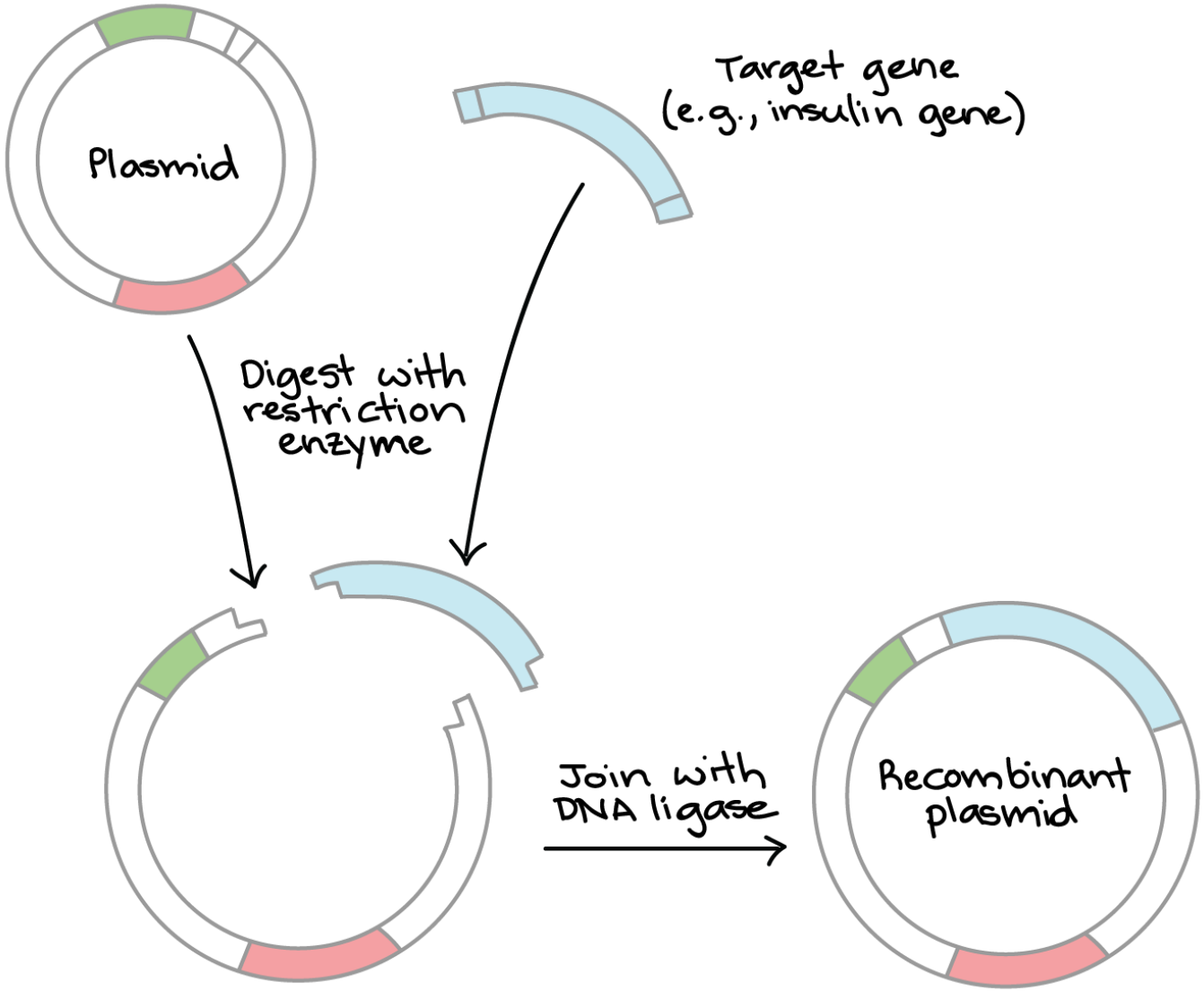


pET-21a-d(+) cloning/expression region

Fusion Tags Available for pET Constructs

Tag	N/C Terminal or Internal	Size (aa)	Basis for Detection and/or Purification	Applications	pET Vector Series
T7•Tag®	N, I	11 or 260	monoclonal antibody	western blot immunoprecipitation purification	3, 5, 9, 11, 17 17x, 21, 23, 24, 28, 33, pSCREEN
S•Tag™	N, I	15	S-protein (104aa) affinity	western blot quantitative assay purification	29, 30, 32, 34–37, 39–42, pSCREEN
His•Tag®	N, C, I	6, 8, or 10	metal chelation chromatography (native or denaturing)	His•Bind, resin purification	14–16, 19–42, pSCREEN

Cloning a Piece of DNA



Performing the Restriction Digests

- You will need to set up a restriction digest of your plasmid vector and your DNA of interest
- Restriction enzymes all have specific conditions under which they work best. Some of the conditions that must be considered when performing restriction digest are:

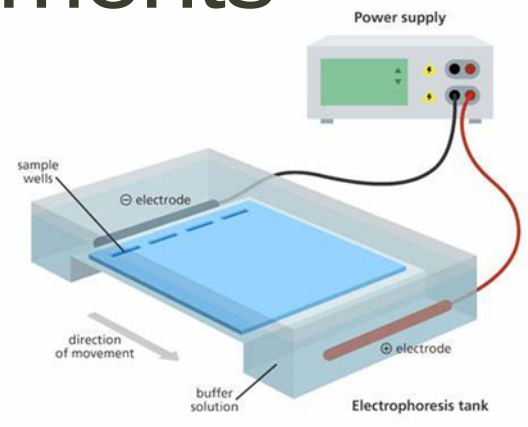
✓temperature

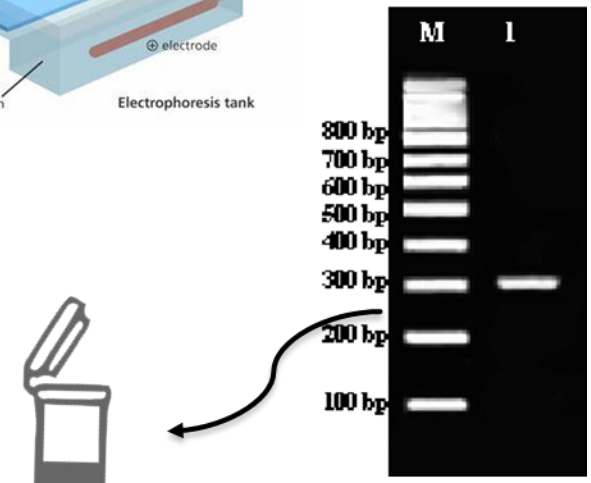
✓salt concentration

✓purity of the DNA

Purify your DNA Fragments

- The insert of interest that you want to clone into your plasmid needs to be separated from the other DNA
- You can separate your fragment using Gel Electrophoresis
- You can purify the DNA from the gel by cutting the band out of the gel and then using a variety of techniques to separate the DNA from the gel matrix

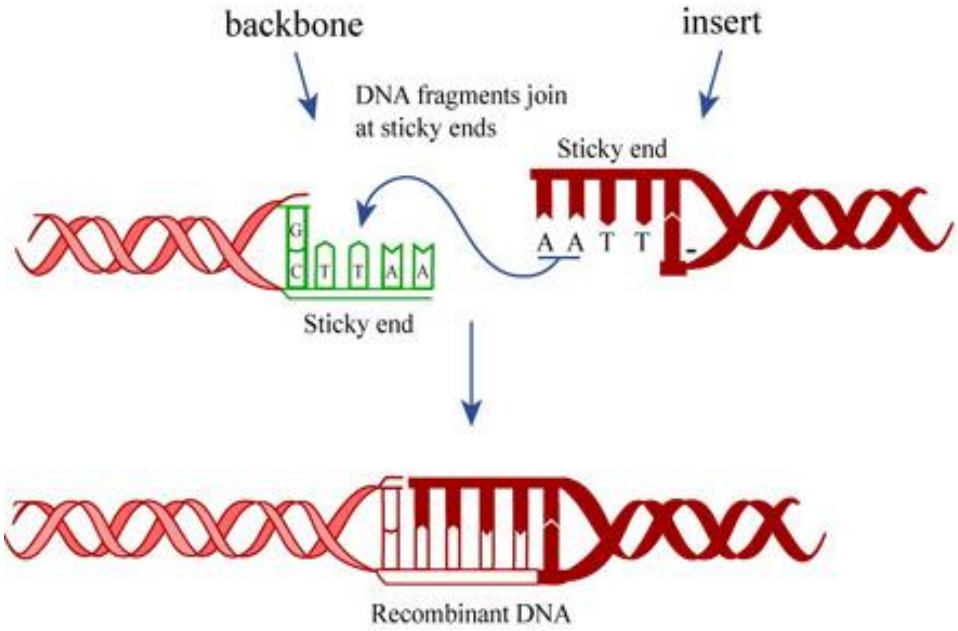




Extraction &purification

Ligation

- Ligation is the process of joining two pieces of DNA from different sources together through the formation of a covalent bond.
- DNA ligase is the enzyme used to catalyze this reaction.
- DNA ligation requires ATP.



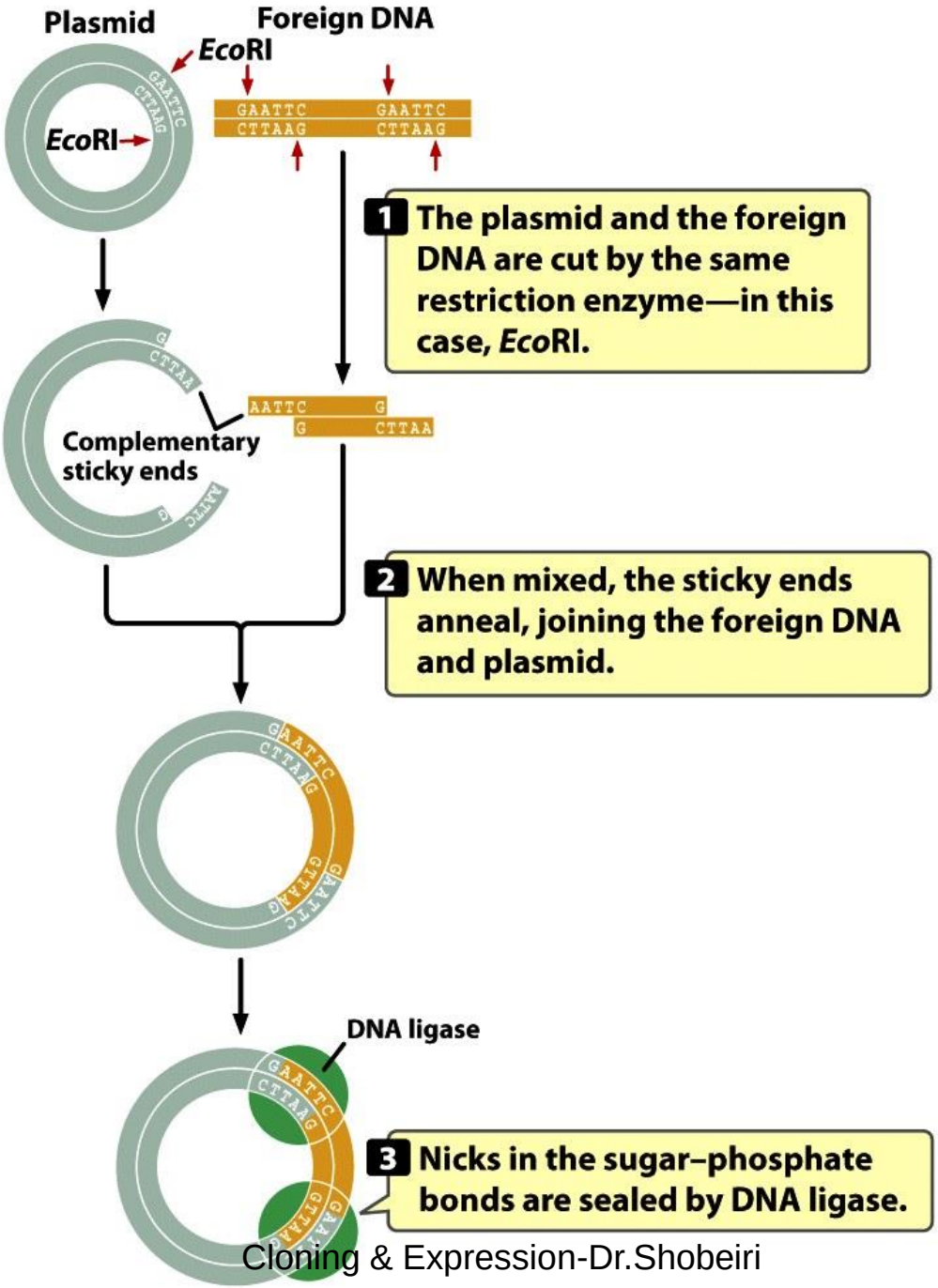
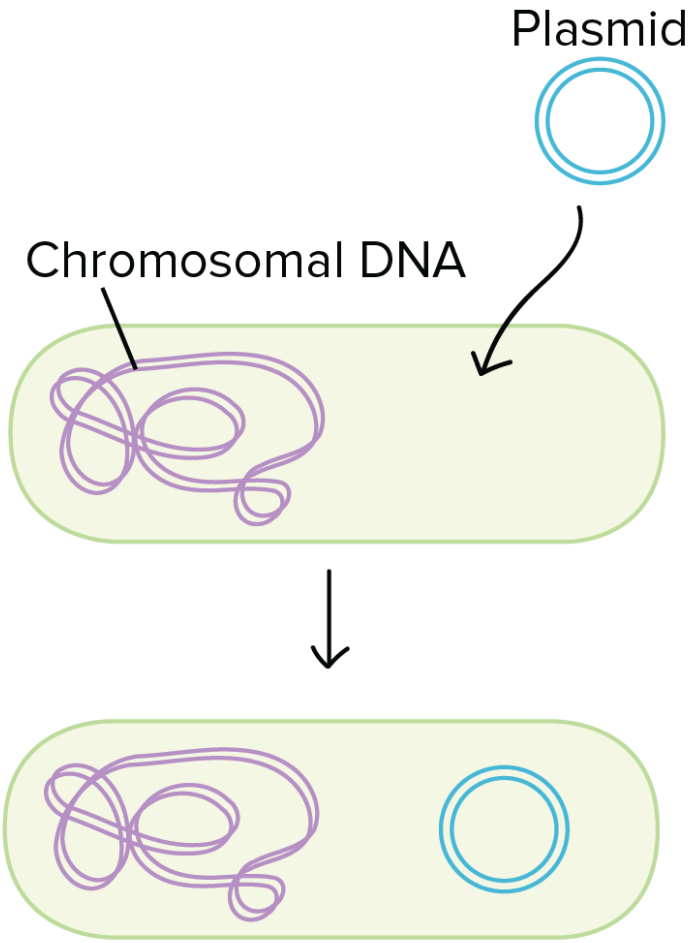


Figure 19-7
Genetics: A Conceptual Approach, Third Edition

Transforming Bacteria

- After you create your new plasmid construct that contains your insert of interest , you will need to insert it into a bacterial host cell so that it can be replicated.

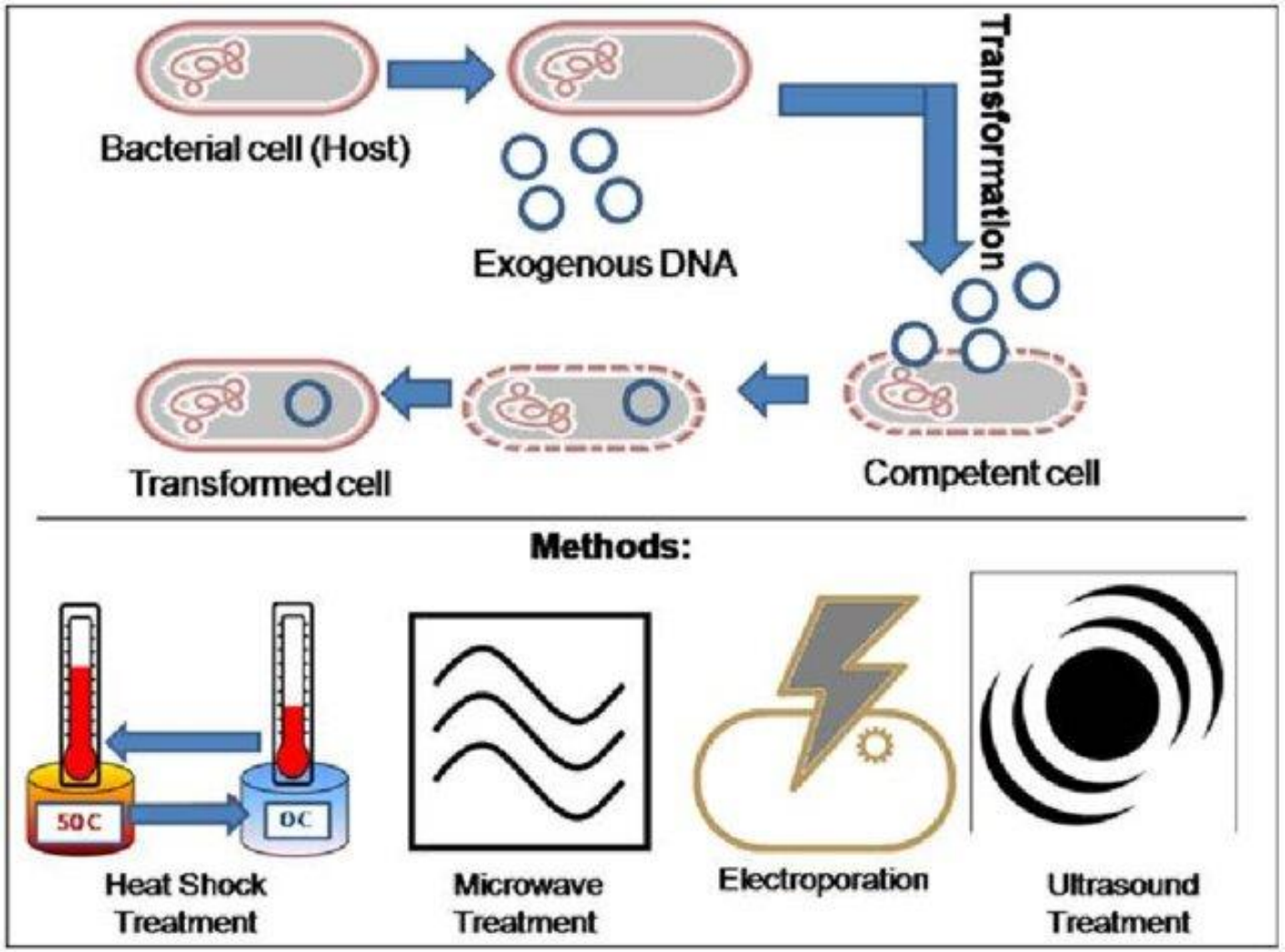
- The process of introducing the foreign DNA into the bacterial cell is called transformation.



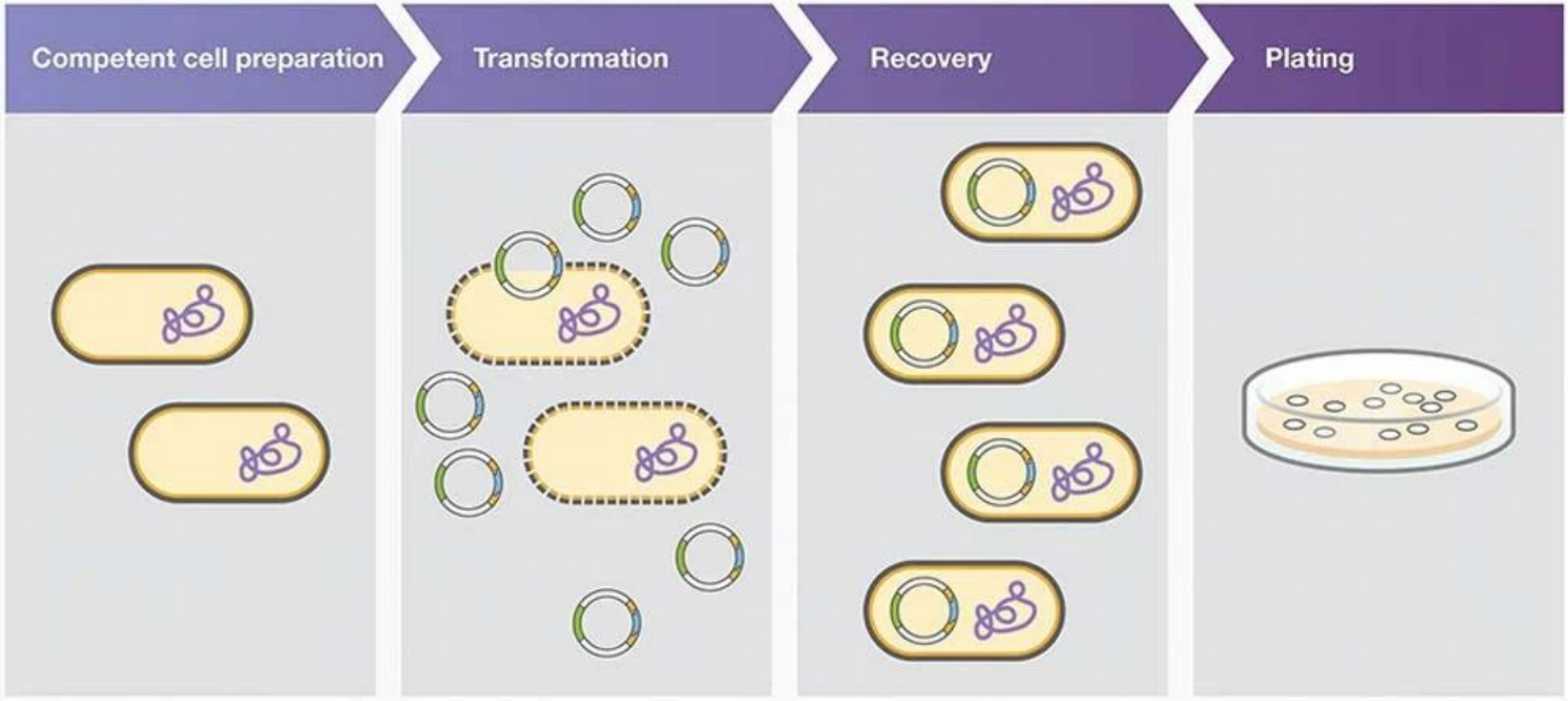
Competent Host Cells

- Not every bacterial cell is able to take up plasmid DNA.
- Bacterial cells that can take up DNA from the environment are said to be competent.
- Can treat cells (electrical current/divalent cations) to increase the likelihood that DNA will be taken up
- Two methods for transforming: heat shock and electroporation

Competent Host Cells

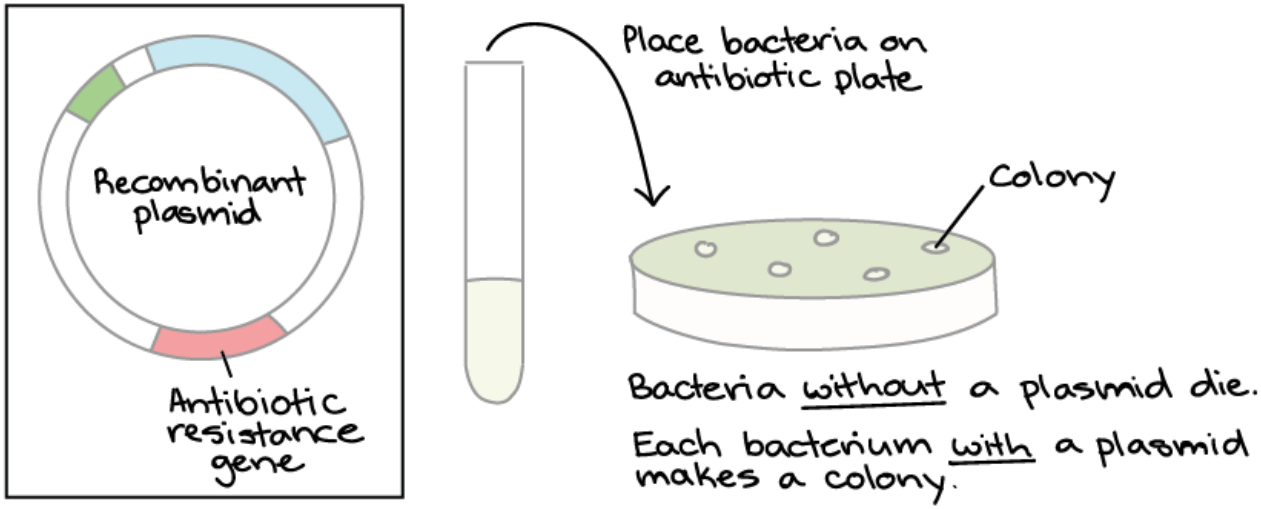


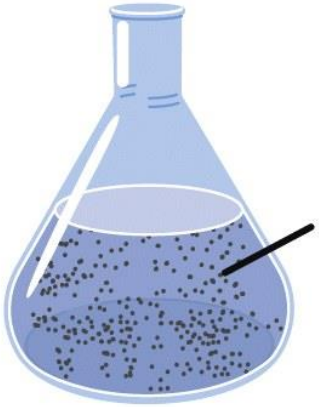
Bacterial Transformation



Selecting for Transformants

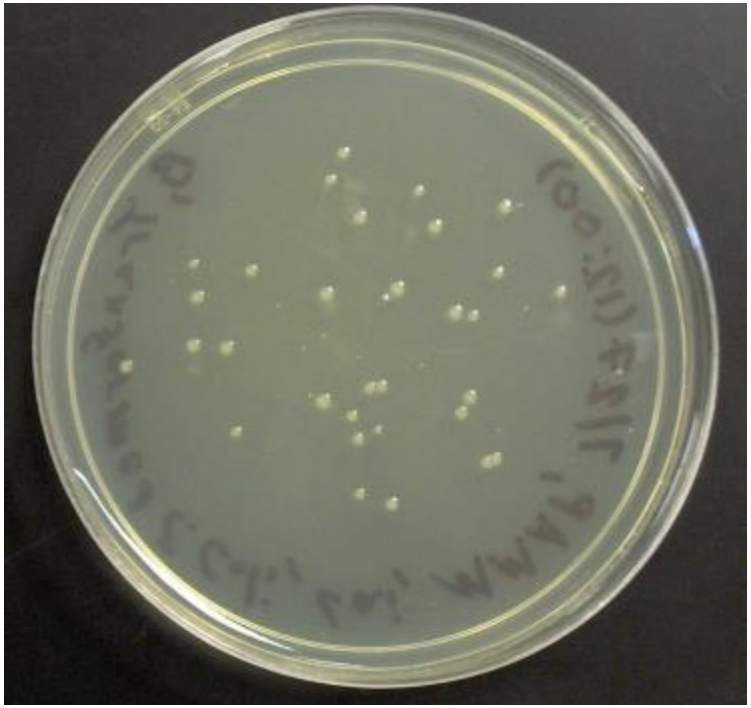
- The transformed bacteria cells are grown on selective media (containing antibiotic) to select for cells that took up plasmid.
- For blue/white selection to determine if the plasmid contains an insert, the transformants are grown on plates containing X-Gal and IPTG.

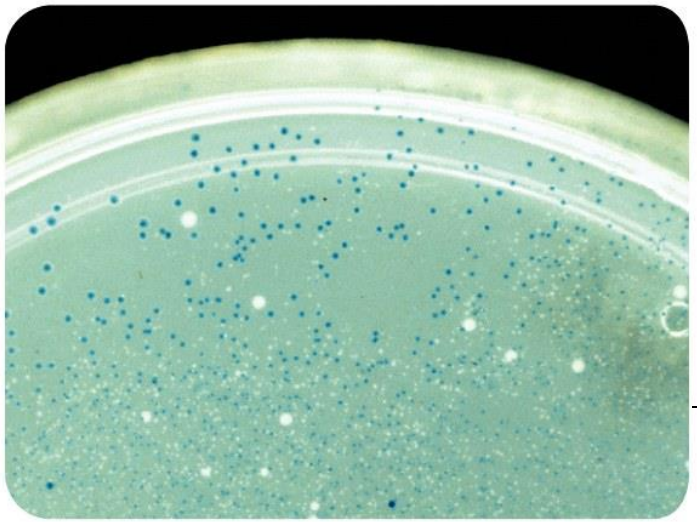




Bacteria

Plate on medium with ampicillin and X-gal



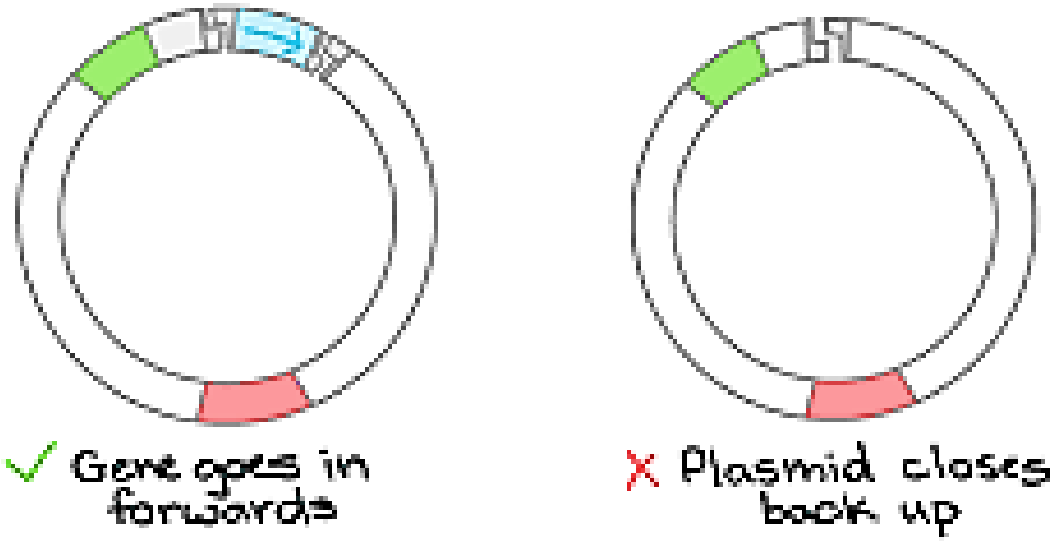


Conclusion: A white colony consists of bacteria carrying a recombinant plasmid.

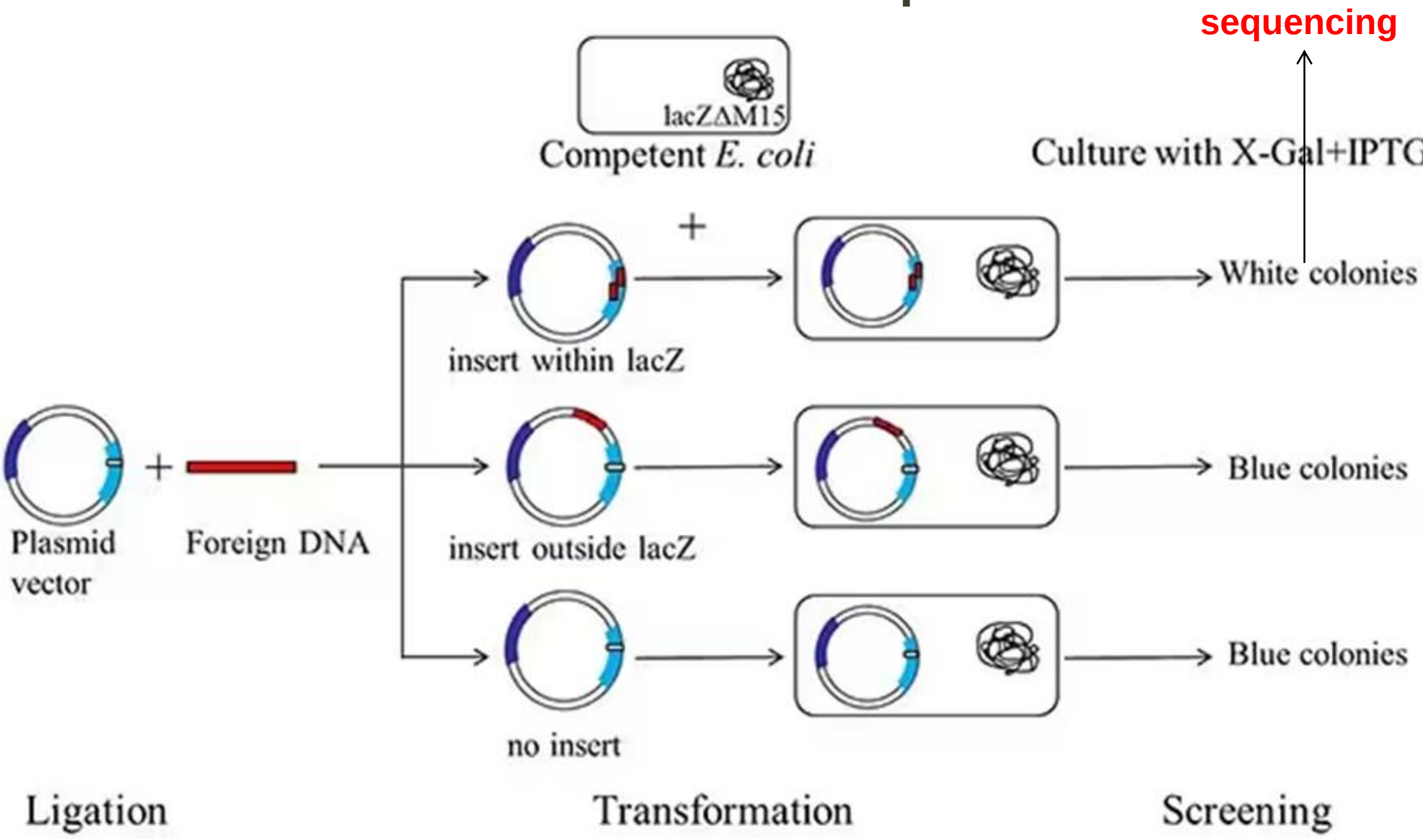
Figure 19-8 part 2
Genetics: A Conceptual Approach, Third Edition

What did the cells take up?

- Plasmid only
- Plasmid with insert cloned
- Foreign DNA from the environment
- Nothing

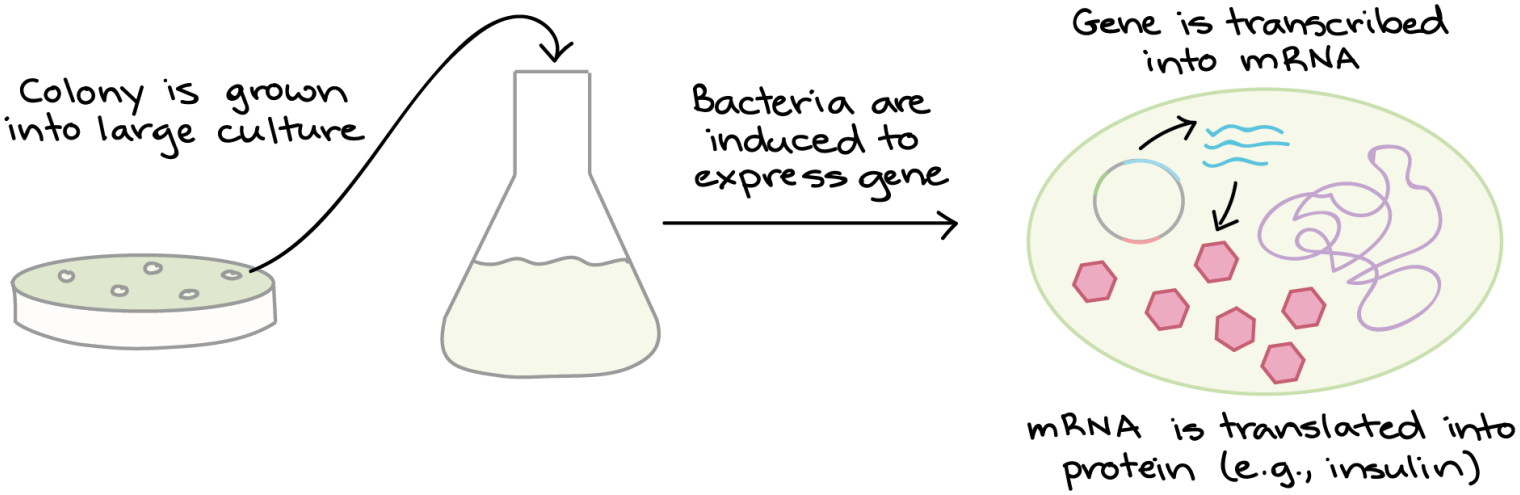


What did the cells take up?

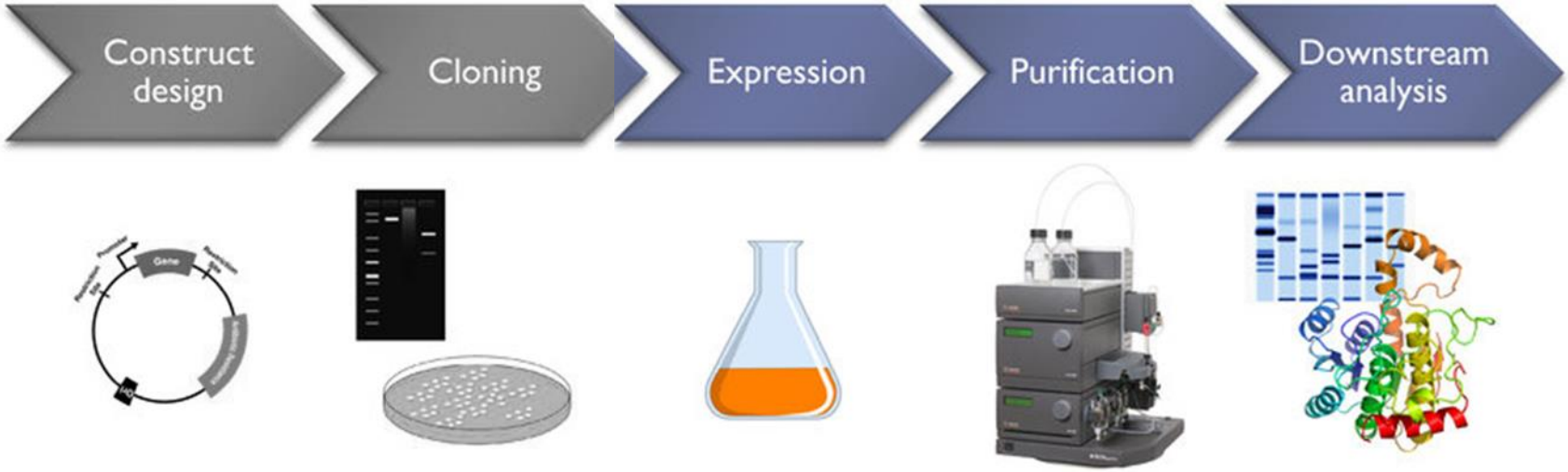


Expressing your cloned gene

- Even if your plasmid contains insert, it may not be able to generate functional protein from your cloned DNA.
  - The gene may not be intact, or mutations could have been introduced that disrupt it.
  - The protein encoded by the gene may require post-translational modifications (i.e., glycosylation or cleavage) to function.
  - Also, some enzymes are a complex of peptides expressed from separate genes.



Recombinant protein expression



Protein Expression Systems

- Expression of a cloned gene can be accomplished by:
  - The *bacteria or yeast* host
  - *Insect or Mammalian cells* (if the plasmid used is designed for expression in mammalian cells)
  - Using an in vitro *cell-free system*.

Characteristic	Bacteria	Yeast	Baculovirus	Mammalian
Cell growth	Rapid (30min)	Rapid (90min)	Slow (18-24h)	Slow (24h)
Med. Complexity	Minimum	Minimum	Complex	Complex
Cost	Low	Low	High	High
Expression	High	Low-High	Low-High	Low-Medium

Post-translational modifications (PTMs)

Protein folding	Not reliable	Usually reliable	Very reliable	Very reliable
N-linked glycosylation	None	High mannose	Simple, no sialic acid	Complex
O-linked glycosylation	N	Y	Y	Y
Phosphorylation	N	Y	Y	Y
Acetylation	N	Y	Y	Y
Acylation	N	Y	Y	Y
γ-carboxylation	N	N	N	Y

Protein Expression in Bacteria advantages

Simple, well-understood genetics

Easy to manipulate genetically

Minimal culturing cost

Fast expression with short doubling time

Well-established labeling protocols for stability studies

Established regulatory track record

Easy to scale up by fermentation

Easy to purify from inclusion bodies

Protein Expression in Bacteria Disadvantages

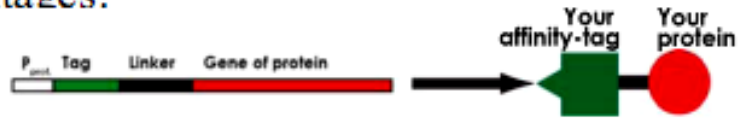
- It does not have the function of eukaryotic post-transcriptional processing
- the protein expression cannot be modified by glycosylation, phosphorylation, etc.,
- it is difficult to form correct disulfide bond pairing and spatial conformation folding.
- The expressed protein is often insoluble and will aggregate into inclusion bodies in the bacteria
- Some pyrogenic sources (endotoxins) may be produced, and E. coli itself contains endotoxins and toxic proteins. The purification of the target protein needs to remove **endotoxins**, which also limits its application.

Features of the BL21-Derived Expression Strains[5]

Expression Strain	Induction	Advantages	Disadvantages
BL21-Gold (DE3) competent cells	IPTG induction of T7 polymerase from *lacUV5* promoter	High-level expression; Ease of induction, direct cloning in expression strain	Leaky expression of T7 polymerase can lead to uninduced expression of potentially toxic proteins.
BL21-Gold (DE3)pLysS competent cells	IPTG induction of T7	Ease of induction, reduced uninduced expression, direct cloning in expression strain	Slight inhibition of induced expression when compared with BL21-Gold (DE3).
BL21 competent cells	Infection with lambda bacteriophage CE6	Tightest control of uninduced expression	Induction not as efficient as DE3 derivatives. Induction (infection) process is more cumbersome.

Protein expression: usage of Tags and fusion proteins

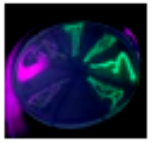
N- or C-terminal fusions of proteins to short peptides (tags) or to other proteins (fusion partners) offer several potential advantages:



- **Improved expression**

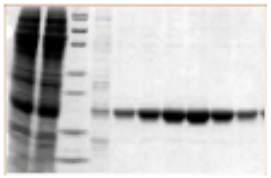
- **Enhanced solubility/folding**
  - E.g. Trx, MBP.
- **Easy detection**

  e.g. Western blot analysis, GFP by fluorescence –



- **Simple purification.**

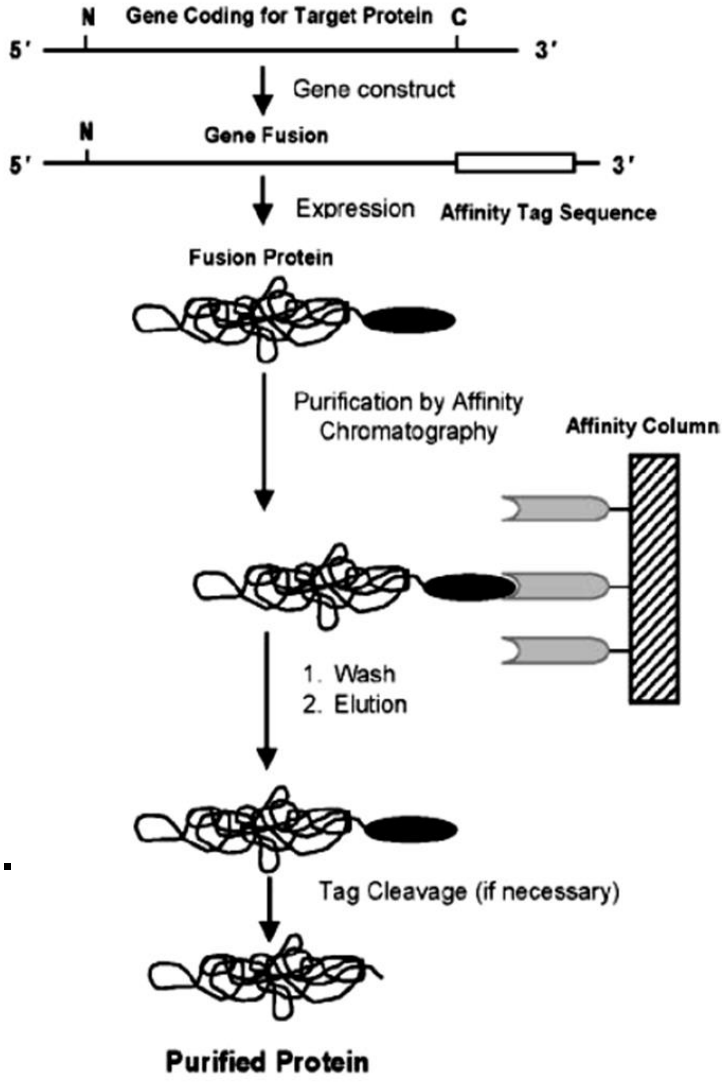
  - fusion partner/tag bind specifically to affinity resins



- Protect from degradation

Increase selectivity of protein purification (Gene fusion strategies)

- Most target protein lack a suitable Affinity ligand usable for capture on a solid matrix.
- A way to circumvent this obstacle is to genetically fuse the gene encoding the target protein with a gene encoding a purification tag.
- When the chimeric protein is expressed, the tag allows for specific capture of the fusion protein.
- Example: 6x his-tag



Tags and Fusions

Fusion	Size	Purification
GST	26 kDa	Glutathione Sepharose
MBP	40 kDa	Amylose resin
Trx	12 kDa	Thiobond

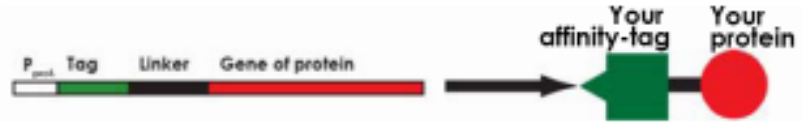
Hunt I, (2005) Protein Exp & Purif 40, 1-22 From gene to protein: a review of new and enabling technologies for multi-parallel protein expression.

Commonly used affinity tag system in recombinant protein expression

- maltose-binding protein (MBP) fusions (provides a factor Xa cleavage site).
- Glutathione-S-transferase (GST) fusion proteins. (contains either a thrombin cleavage site, a factor Xa cleavage site, or an Asp-Pro acid cleavage site).
- Thioredoxin (Trx) fusion proteins (provides an enterokinase cleavage site).
- 6X His-tagged proteins.

Construct design – fusion cleavage sites

- often included to allow removal of tags/fusions



Enzyme	Recognition site	Comments
Thrombin	LVPR/GS	Less specific
Factor Xa	IEGR/	Less specific
Enterokinase	DDDDK/	Very specific

Waugh DS (2005) Trends Biotechnol.23(6):316-20 Making th

A	Alanine	M	Methionine
C	Cisteine	N	Asparagine
D	Aspartic Acid	P	Proline
E	Glutamic Acid	Q	Glutamine
F	Phenylalanine	R	Arginine
G	Glycine	S	Serine
H	Histidine	T	Threonine
I	Isoleucine	V	Valine
K	Lysine	W	Tryptophan
L	Leucine	Y	Tyrosine

Insolubility of heterologous proteins produced in E.coli

Inclusion bodies:

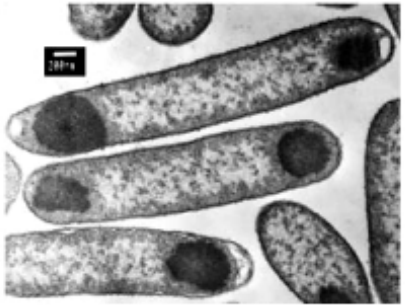
Dense particles, containing precipitated proteins. Their formation depends on protein synthesis rate, growth conditions.

Advantages: proteolysis resistant, big yield, relatively pure, easy to separate.

Disadvantages: inactive product requires in vitro refolding and renaturation

Expression problems

- Protein is insoluble
  - Collect inclusion bodies and refold protein
  - Reduce growth temperature
  - Use heat shock to induce chaperones
  - Use a low/moderate copy number plasmid vector
  - Fuse a periplasmic targetting sequence to N-terminus
  - Co-express chaperones/foldases (PDI…)
  - Try a different fusion partner (e.g. MBP)



Refolding of recombinant proteins

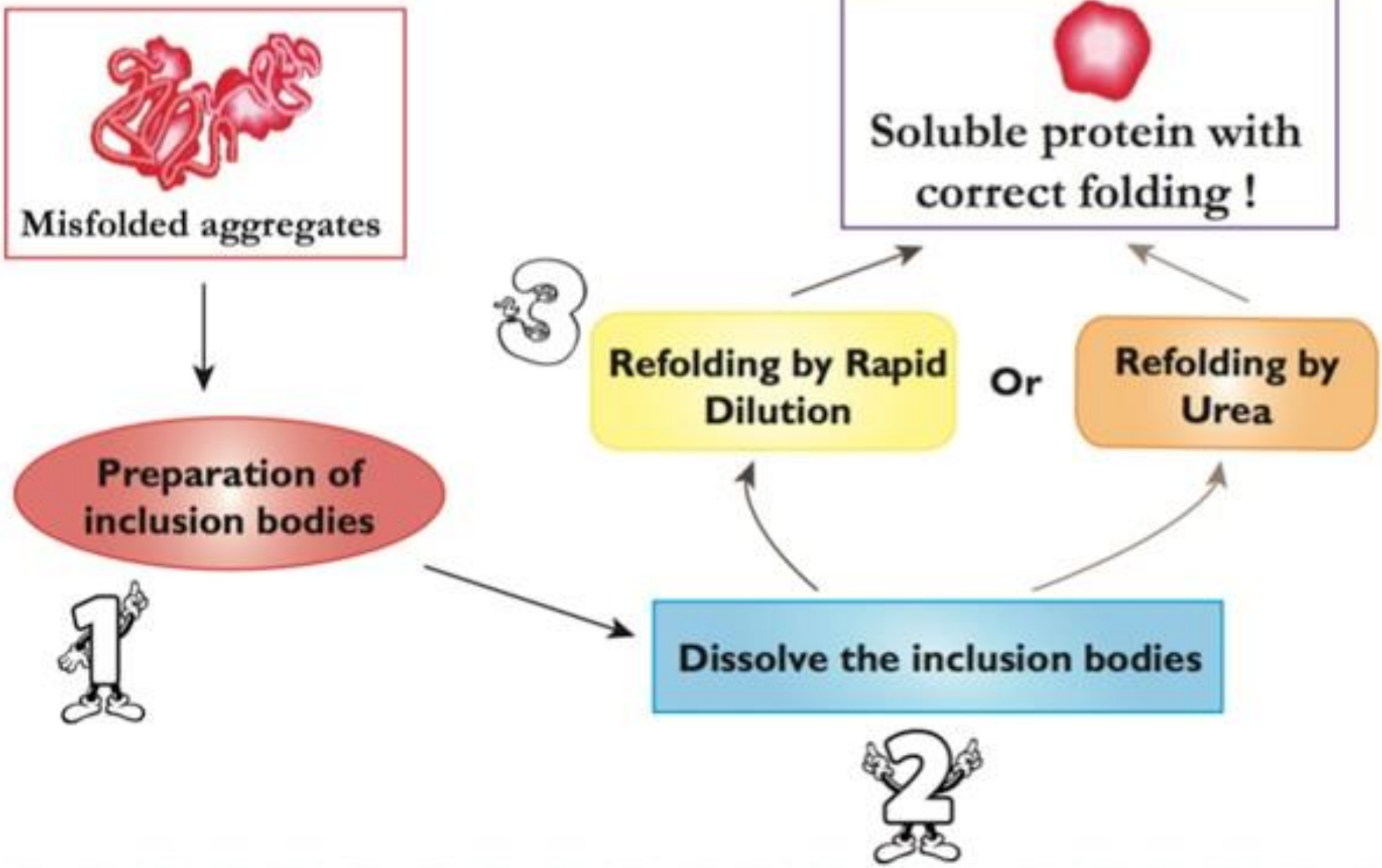
Solubilization:

detergents, high concentration of inorganic salts or organic solvents all used. The most commonly used organic solutes such as urea or guanidine-HCl often used in the presence of reducing agents (mercaptoethanol or DTT). Solubilized proteins can be purified by ion-exchange chromatography or other conventional methods, prior to refolding.

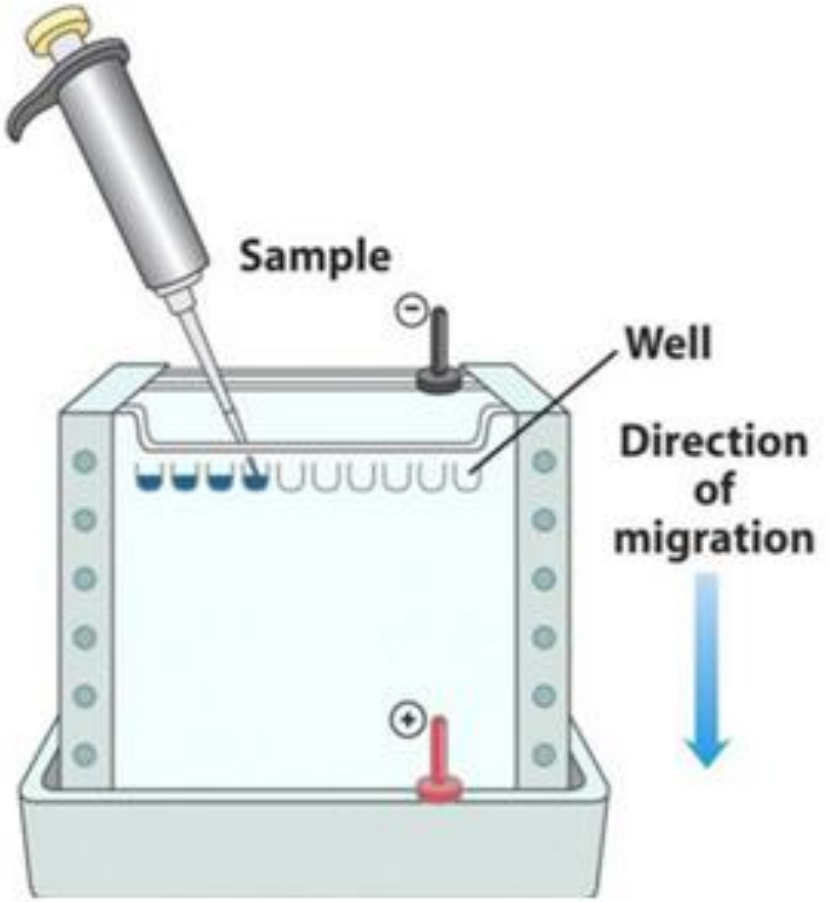
Refolding:

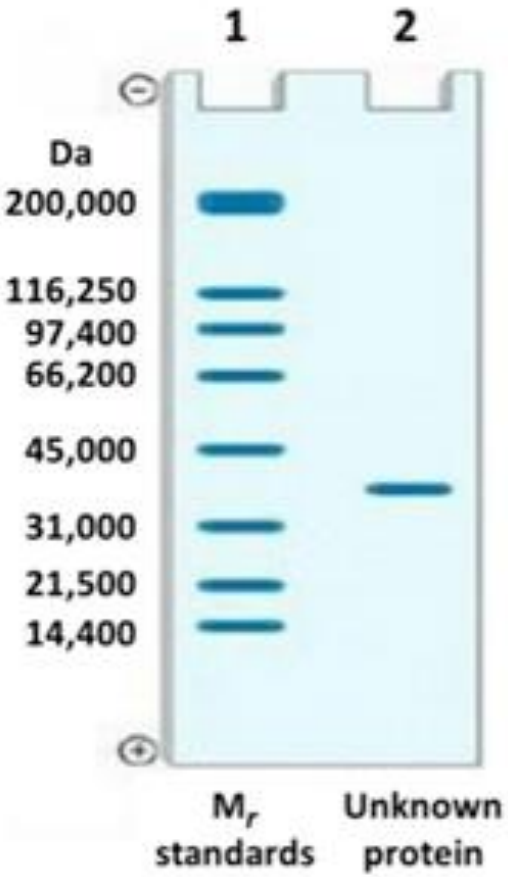
If no S-S bonds present - remove denaturing agent to allow protein to fold correctly.

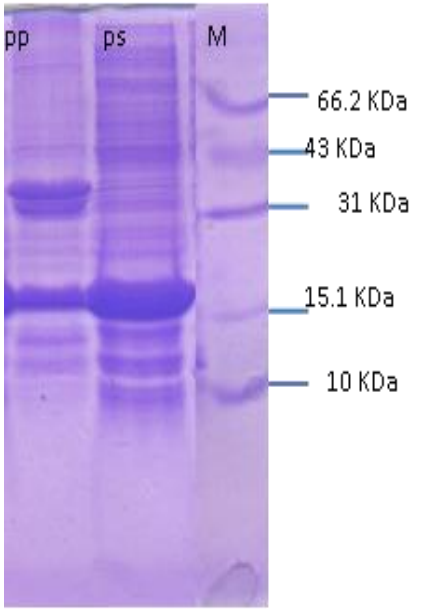
If S-S bonds present - their formation can be accomplished: by air oxidation, catalysed by **trace metal ions**; by a **mixture of reduced and oxidized thiol compounds** - oxidized DTT, reduced DTT; GSSG/GSH; cystine and cysteine, cystamine and cysteamine.


Misfolded aggregates
Soluble protein with correct folding !
3
Refolding by Rapid Dilution
Or
Refolding by Urea
Preparation of inclusion bodies
1
Dissolve the inclusion bodies
2

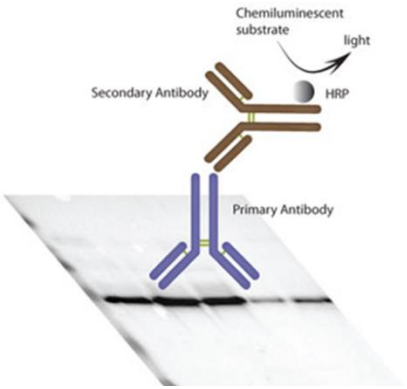
Protein validation SDS-PAGE

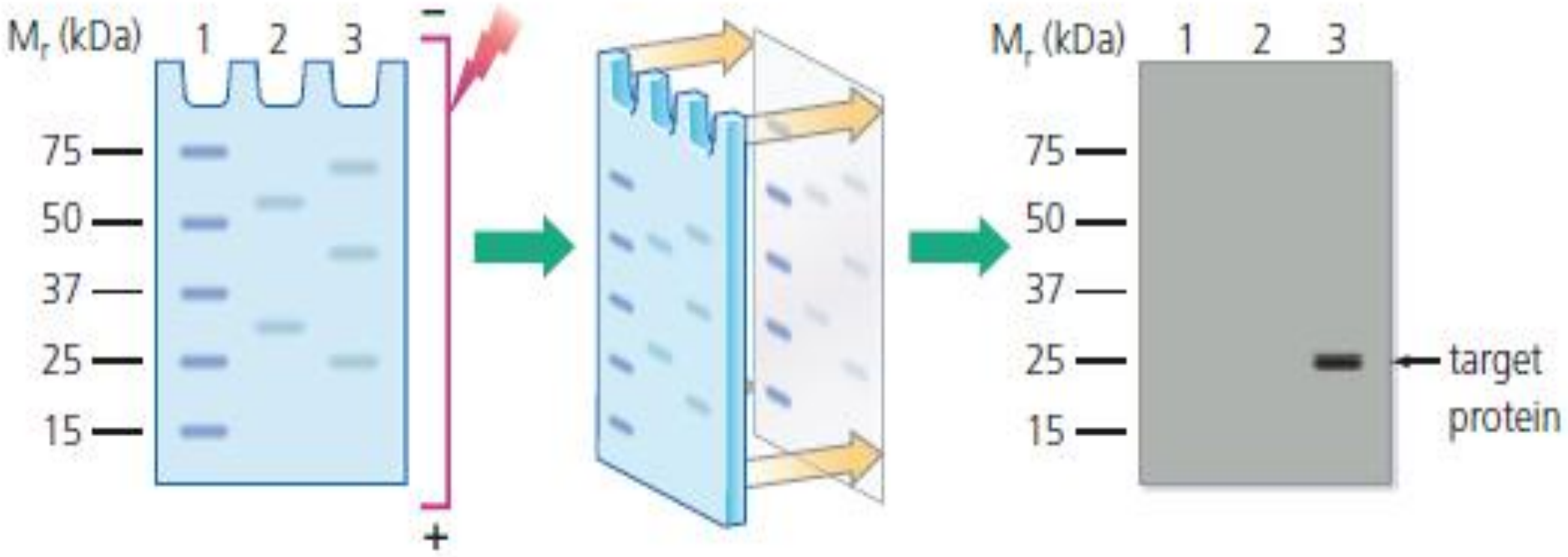




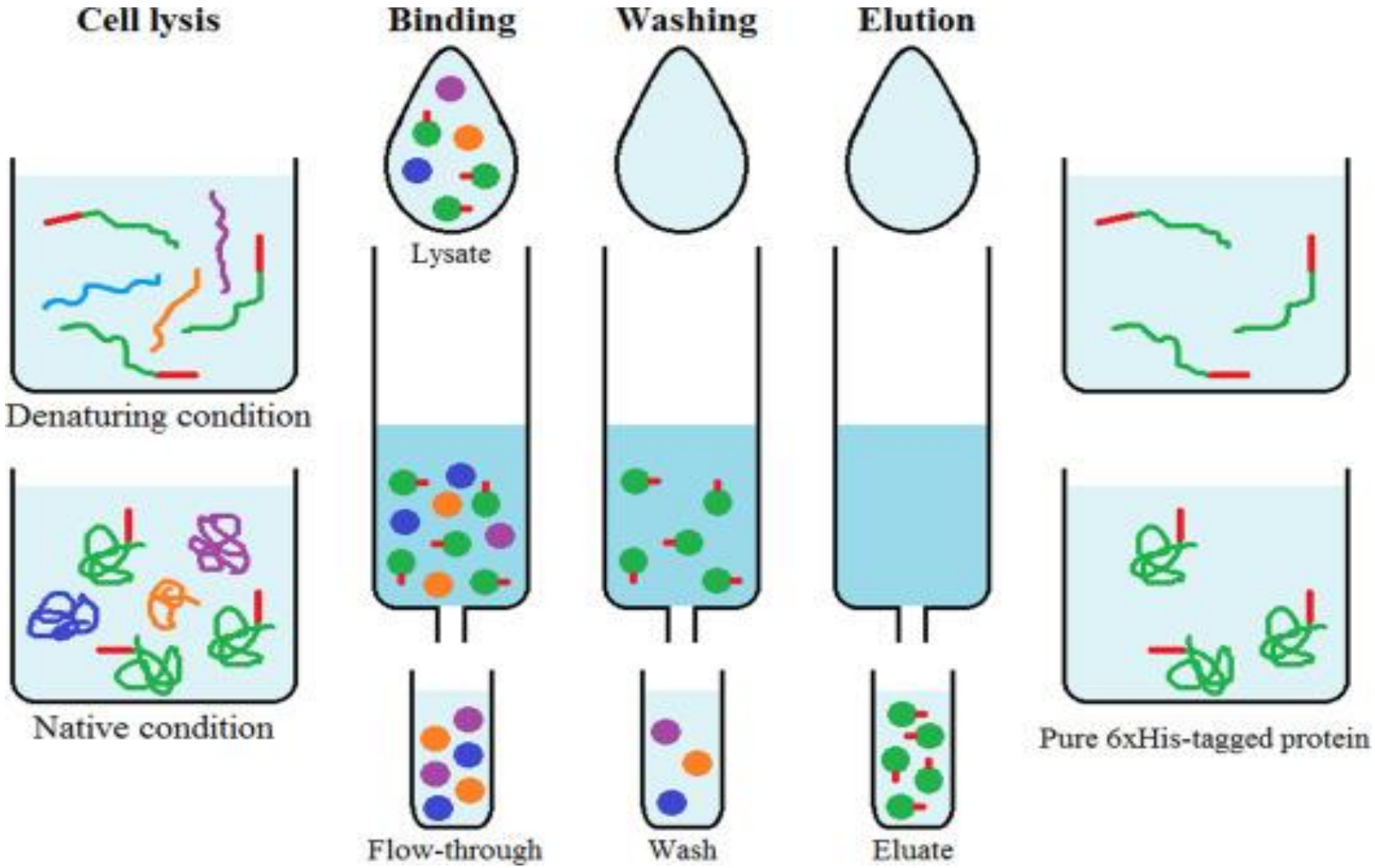


Protein validation Western blotting



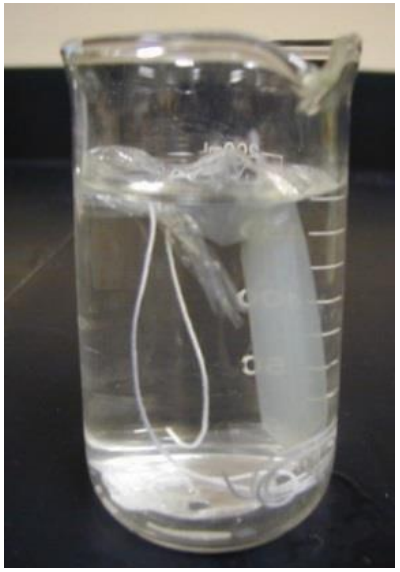


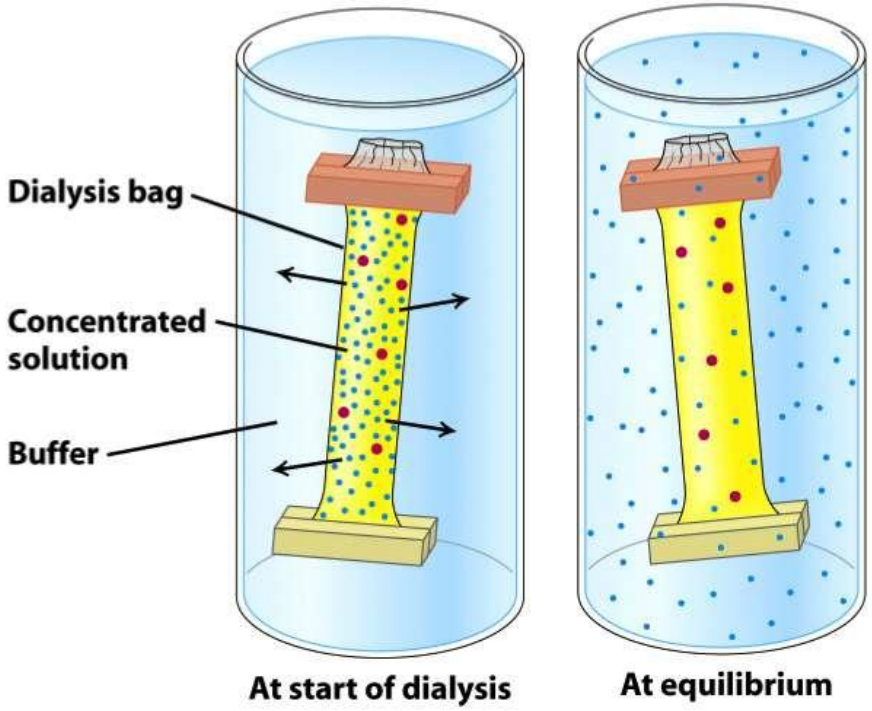
Protein purification



Protein concentration

- Dialysis is usually used **to change** the salt (small-molecule) composition of a macromolecule-containing solution.
- The solution to be dialyzed is placed in a sealed dialysis membrane and immersed in a selected buffer; small solute molecules then equilibrate between the sample and the dialysate.
- dialysis bag Cut off point: 8 kD<






Thank
you!