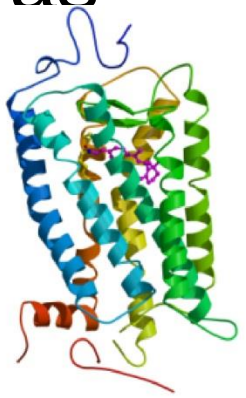
DNA cloning and protein expression

So for years, the only proteins that are well studied are the ones that are abundant in nature

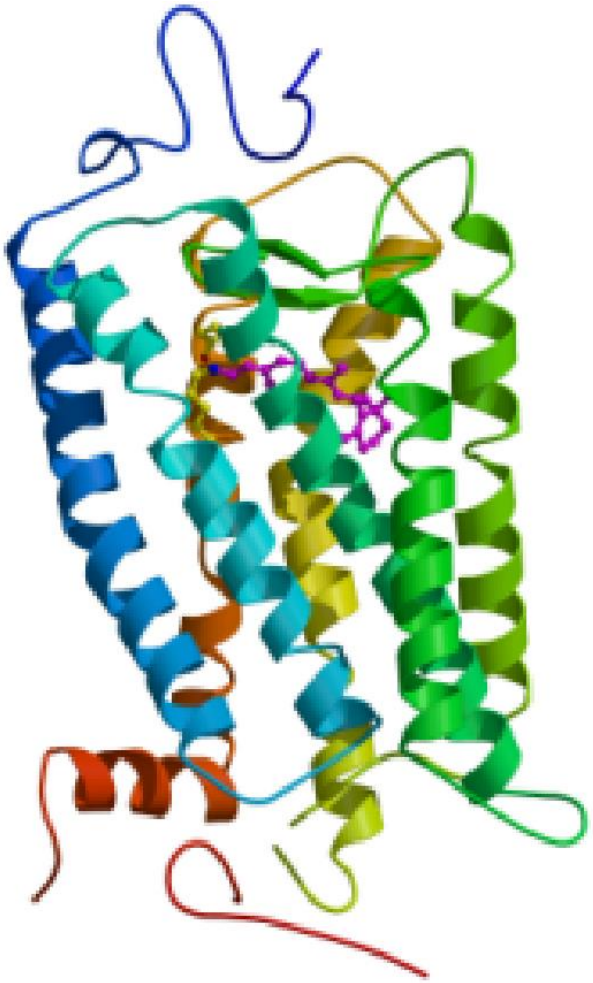
- For example, 70% of the drugs are market are targeting one category of proteins, G-protein coupled receptors (GPCR).
- But most GPCRs can not be purified. The one that was well studied and served as the model for everything, rhodopsin.



Rhodopsin is the most abundant protein in mammalian eyes to sense light.

Grams of proteins can be obtained by grinding eyes of cows.

But other GPCRs were not purified in large amount until recently.
Beta-2 adrenergic receptor (target of beta-blocker and a lot of asthma drugs)



Bovine Rhodopsin

Most proteins don’t exist in large quantity in cells

- Overexpression. Insert target gene into plasmid and use host cells to express the protein. The expression can be turned on once the cell reaches a certain density. At which point, a chemical trigger will be added and cells will produce target protein.

Mostly used organisms for protein overexpression

First, you are a CEO of a biotech startup, you realize that expression system is in high demand, which is why you choose it as the main item on your company portfolio.

So you tell your team to choose a target organism. What the benchmark of this ideal system?

Mostly used organisms for protein overexpression

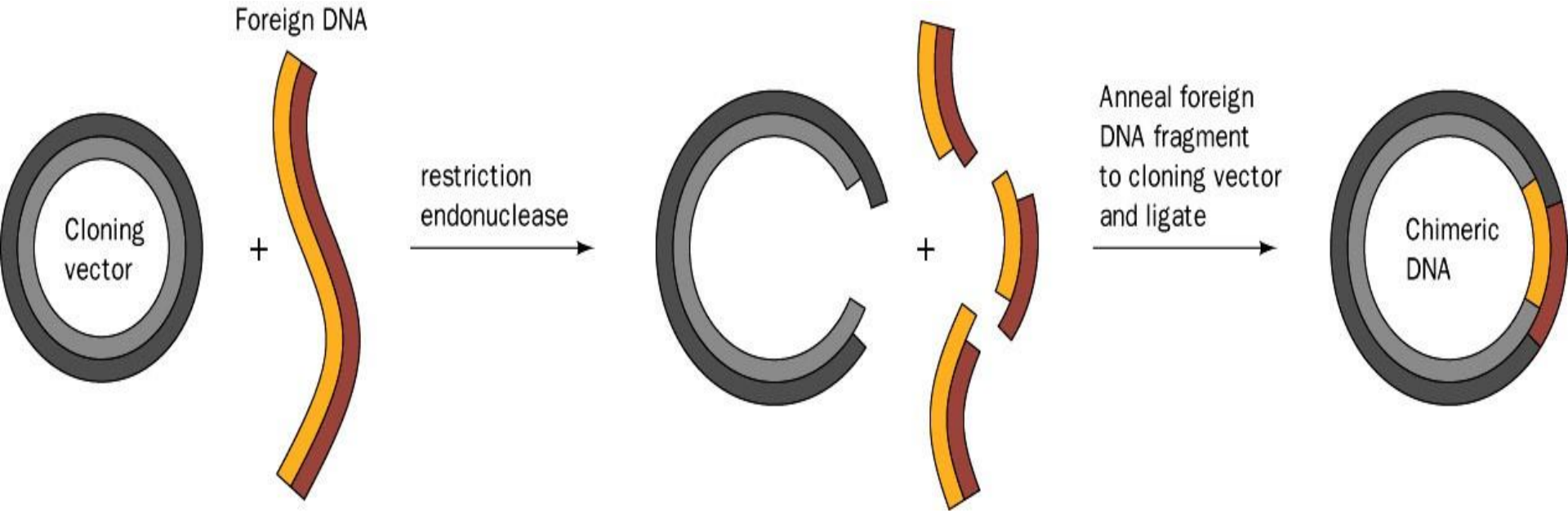
- Bacterial culture
- Yeast
- Drosophila (or silk worm)
- Insect cells
- Mammalian

Mostly used organisms for protein overexpression

- Bacterial culture
  - Easy to manipulate.
  - Don’t get contaminated easily.
  - Grow fast. (one day turn-around).
  - Cheap!

A step-by-step introduction of molecular cloning

The goal


Foreign DNA
Cloning vector
+
restriction endonuclease
+
Anneal foreign DNA fragment to cloning vector and ligate
Chimeric DNA

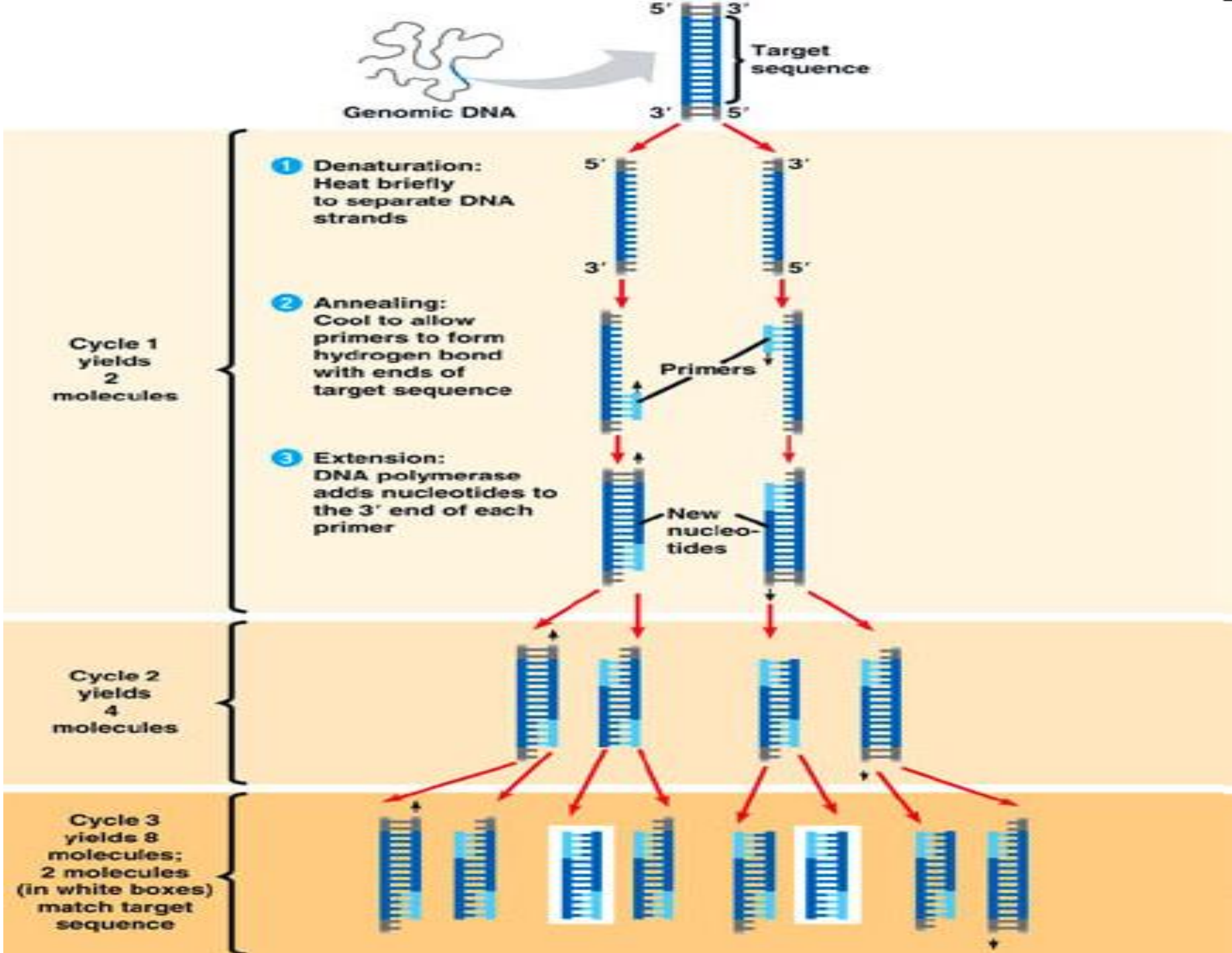
- Step 1: PCR the target gene. Decide which restriction enzyme you want. And design primers with it.
- Step 2. examine PCR product and purify the DNA.
- Step 3: treat PCR product by restriction enzymes.
- Step 4. purify the cut PCR product.

At the same time

- Step 5, treat vector with restriction enzyme.
- Step 6, purify cut vector.
- Finally put your gene (step4) and vector (step 6) together and ligate.
- Transform into bacteria.

- Step 1: PCR the target gene. Decide which restriction enzyme you want. And design primers with it.

What is the component in setting up PCR reaction


Genomic DNA
5′
3′
Target sequence
3′
5′
Cycle 1 yields 2 molecules
1 Denaturation: Heat briefly to separate DNA strands
2 Annealing: Cool to allow primers to form hydrogen bond with ends of target sequence
Primers
3 Extension: DNA polymerase adds nucleotides to the 3′ end of each primer
New nucleo-tides
Cycle 2 yields 4 molecules
Cycle 3 yields 8 molecules; 2 molecules (in white boxes) match target sequence

The components of PCR reaction

- **Template DNA**
- **primers**
- **dNTPs (dATP, dTTP, dCTP & dGTP)**
- **DNA polymerase**
- **$MgCl_2$**
- **PCR buffer (good pH and salt for specific DNA polymerase)**
- **Possible additive (DMSO, glycerol, etc)**

Set up your PCR experiment

- Template
- Primer
- Buffer
- Mg ion
- dNTP
- Water

- Last – add polymerase

Set up your PCR experiment

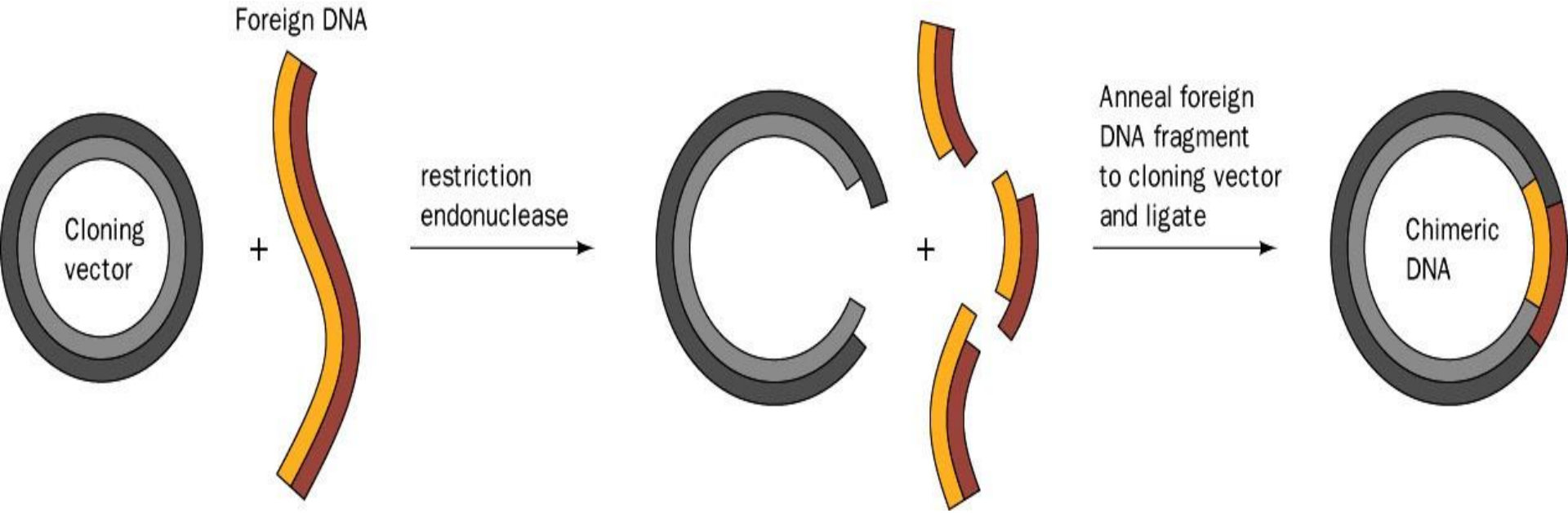
- First cycle of denature. (30s-2 min) 95 or 98°C.
- Denature.
- **Annealing** (annealing temperature)
- Extension (68 or 72°C) but **how long**
- How many circles?

What is an ideal DNA polymerawse

To design primers

- A short preview

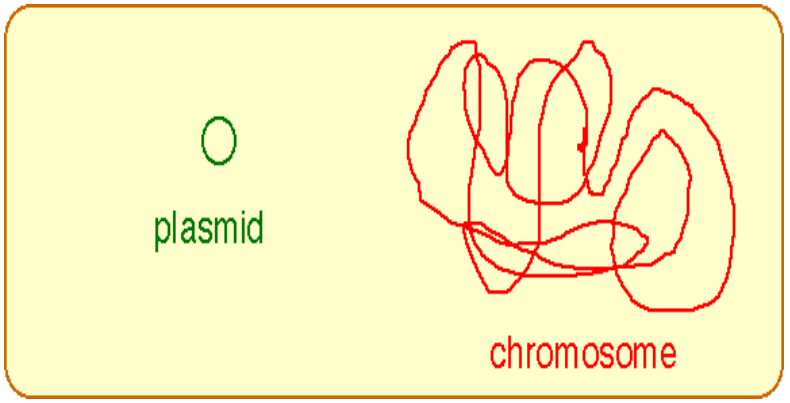
The goal

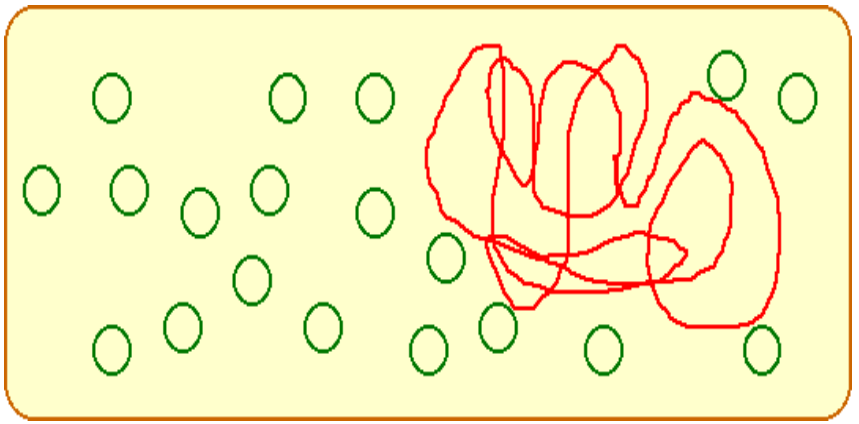

Foreign DNA
Cloning vector
+
restriction endonuclease
+
Anneal foreign DNA fragment to cloning vector and ligate
Chimeric DNA

Step 4. purify the cut PCR product.

PLASMIDS

- Bacterial cells may contain extra-chromosomal DNA called plasmids.
- Plasmids are usually represented by small, circular DNA.
- Some plasmids are present in multiple copies in the cell





VECTORS

- **Cloning vectors are DNA molecules that are used to "transport" cloned sequences between biological hosts and the test tube:**
  - **What is the minimal requirement for the vectors to function as transporters of targeted genes.**

VECTORS

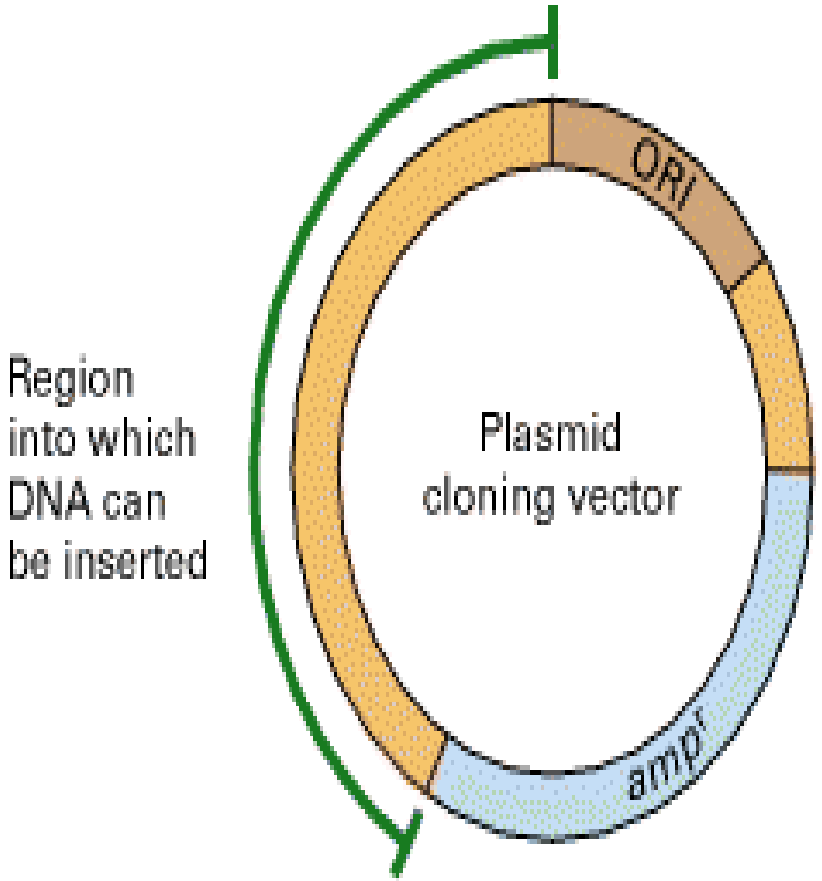
- **Cloning vectors are DNA molecules that are used to "transport" cloned sequences between biological hosts and the test tube:**

  - Ability to promote autonomous replication.

  - Contain a genetic marker (usually dominant) for selection.

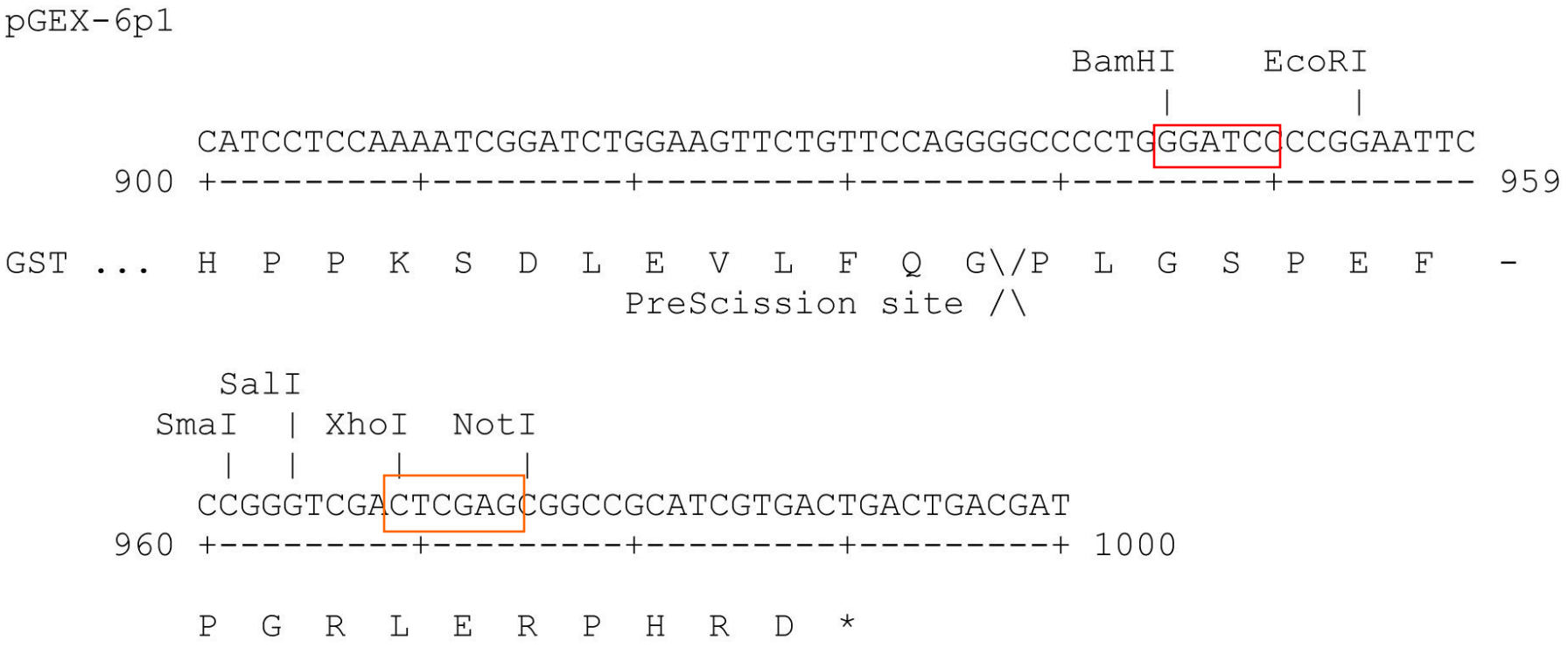
  - Unique DNA sequence to facilitate cloning of insert DNA.

- **replication origin (ORI) sequence**
- **a gene that permits selection,**
- **Here the selective gene is *ampr*; it encodes the enzyme b-lactamase, which inactivates ampicillin.**
- **Exogenous DNA can be inserted into the bracketed region .**
- **Under the selective conditions, only cells that contain plasmids with selectable marker can survive**



What restriction enzyme to choose

- First check the vector you want to put your gene in.



So we decided to use BamHI and XhoI

- BamHI

Activity in NEBuffers:
NEBuffer 1: 75%NEBuffer 2: 100%NEBuffer 3: 100%NEBuffer 4: 100%

- XhoI

Activity in NEBuffers:
NEBuffer 1: 75%NEBuffer 2: 100%NEBuffer 3: 100%NEBuffer 4: 100%

- 1. what restriction enzyme site the vector has?
- 2. Does the gene has internal site for this?

http://tools.neb.com/NEBcutter2/

- 3. do you want the stop codon or not?

  - We want restriction enzyme that is specific, active.
  - To do double digestion at the same time, restriction enzyme have good activity at same buffer.

What to consider when design primers- some tips

1. primers should be 17-60 bases in length

2. base composition should be 45-70% (G+C) (why)

3. primers end (3') prefers to be a G or C, (why)

4. T_ms between 55-80^{o}C are preferred

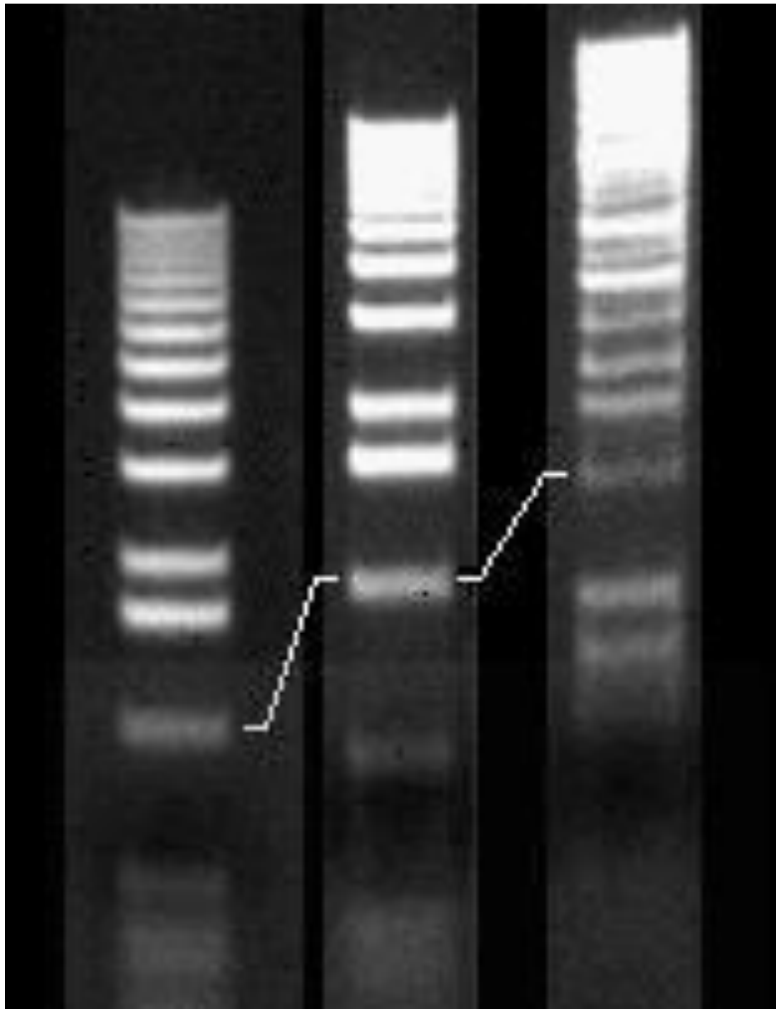
5. 3'-ends of primers should not be complementary (ie. base pair), as otherwise primer dimers will be synthesised preferentially to any other product

6. primer self-complementarity (ability to form secondary structures such as hairpins) should be avoided

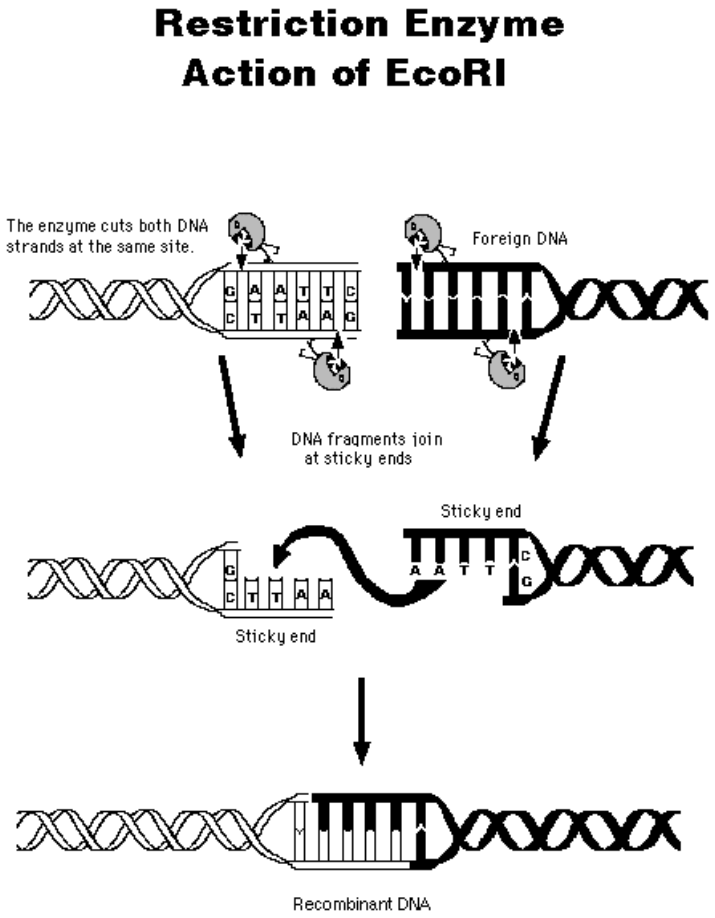
Step 2. examine PCR product and purify the DNA.

After completion of the run, add a DNA staining material and visualize the DNA under UV light.

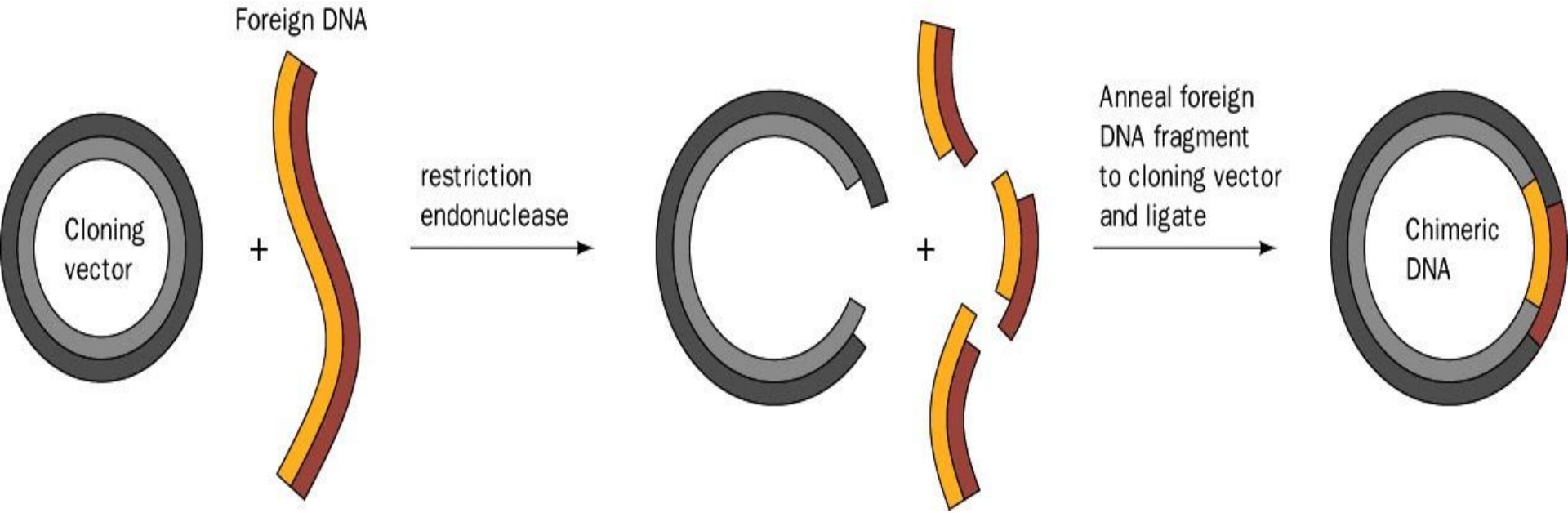
Agarose Concentration



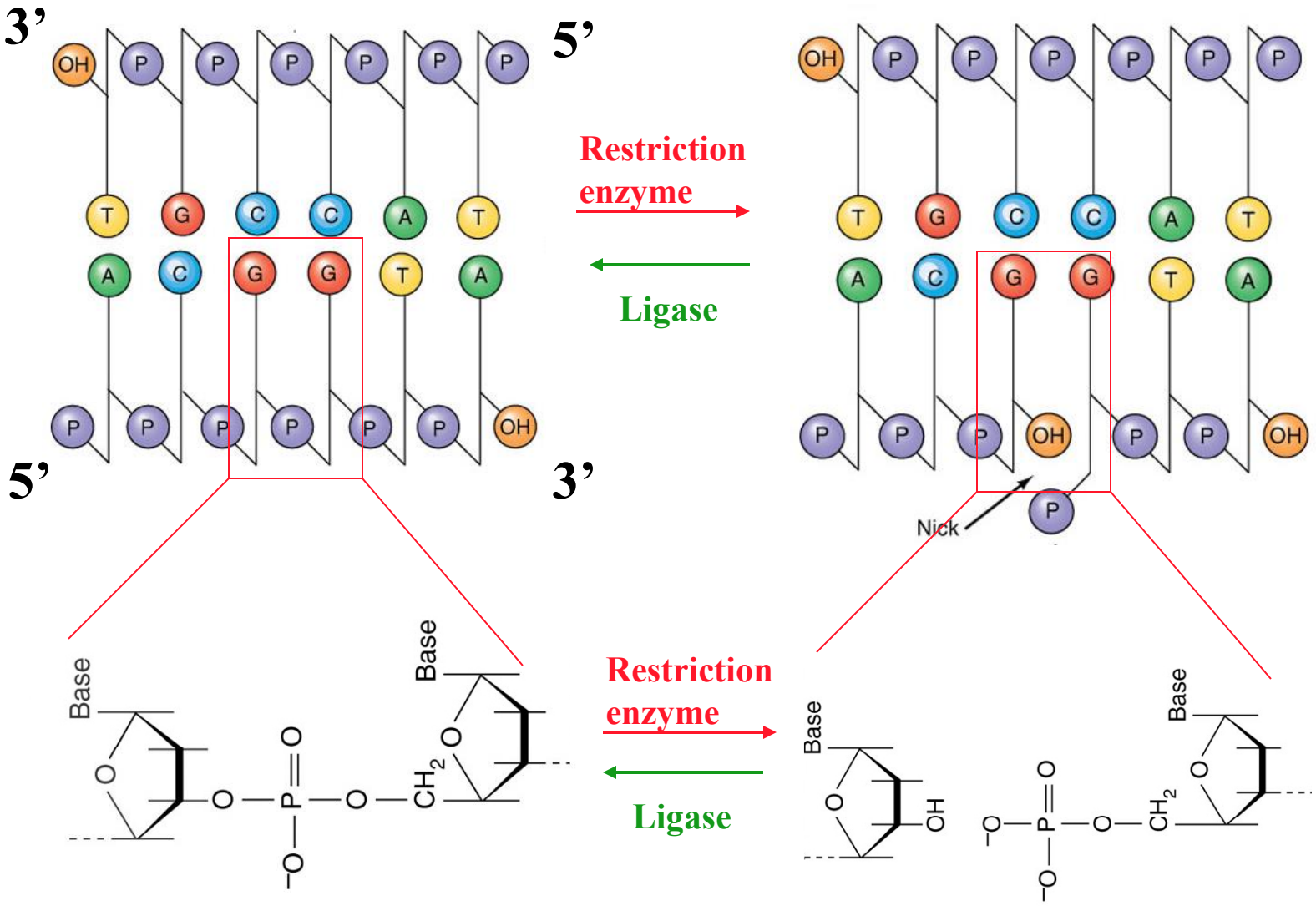
Step 3: treat vector by restriction enzymes and purify by gel purification or spin column



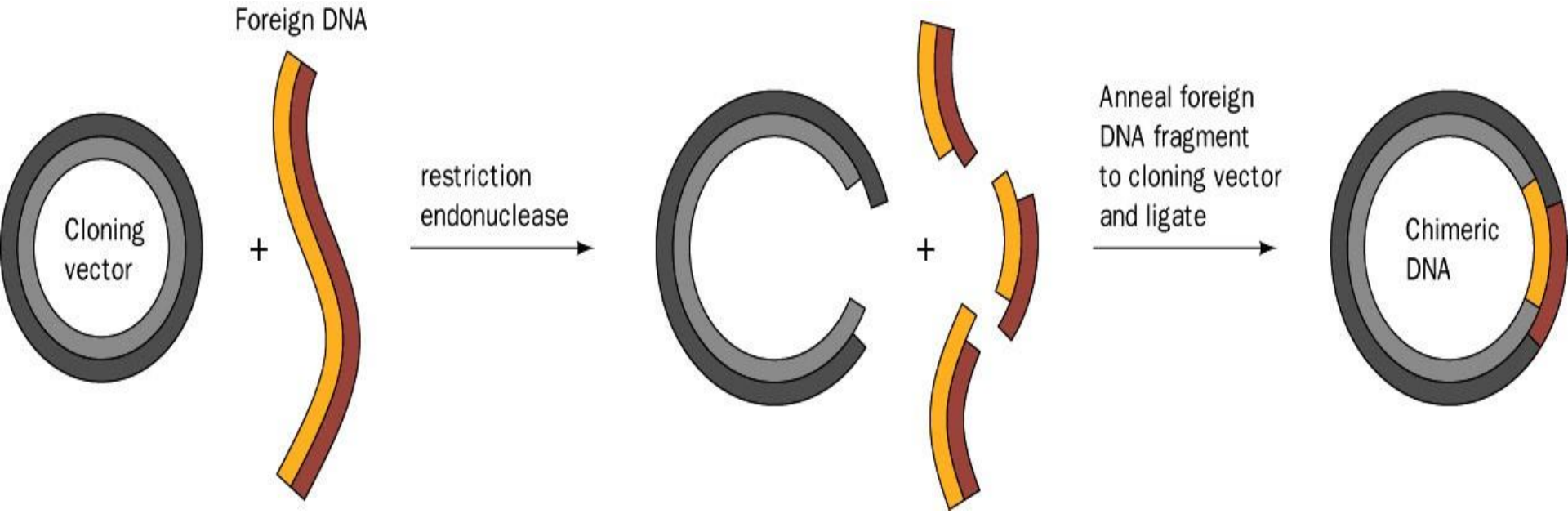
Finally


Foreign DNA
Cloning vector
+
restriction endonuclease
+
Anneal foreign DNA fragment to cloning vector and ligate
Chimeric DNA

DNA ligase covalently links two DNA strands



The rate of ligation is not high



To optimize ligation, different ratio of vector and insert should be tested

1 vector : 3 insert

1 vector : 10 insert

1 vector: 50 insert

Assumption: all cuts are 100%

Transformation

- Upon heat shock or electric pulses, bacterial cell walls can temporarily open up and allow DNA to get in.
- The mechanism for such take-up is still unknown.
- Overall take-up efficiency is low. But once they do, the ligated ones will be duplicated in bacteria along with bacteria growth.

- Antibiotics can be used to select the bacterial cells that take-up the foreign DNA.
- So the specific DNA you want are duplicated many times in bacterial cells.

- When you spread them on agar plate and put it to selection, the ones that have plasmid will grow as colonies.

- Now, purify the DNA and send it for sequencing.

- BINGO!

Step by step bacterial expression

- Step 1: transform vectoring carrying the target gene into bacterial strain.
- Step 2: Grow a small culture overnight. (20-50ml)
- Step 3: Grow a large culture using a small amount from your starter culture.
- Step 4: Induce the expression.
- Step 5: Grow for a few more hours and harvest.

Bacterial expression vectors

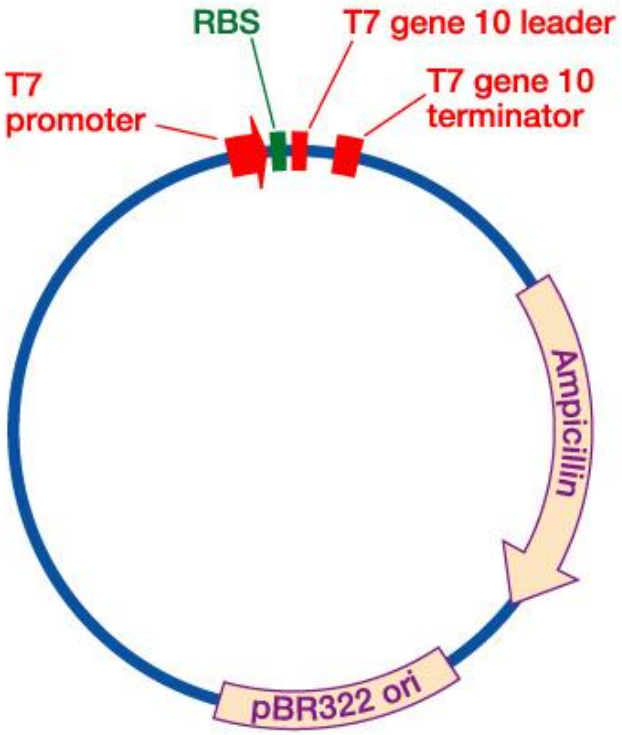


Figure 5-21 part 1 of 2 Human Molecular Genetics, 3/e. (© Garland Science 2004)

Bacterial expression vectors

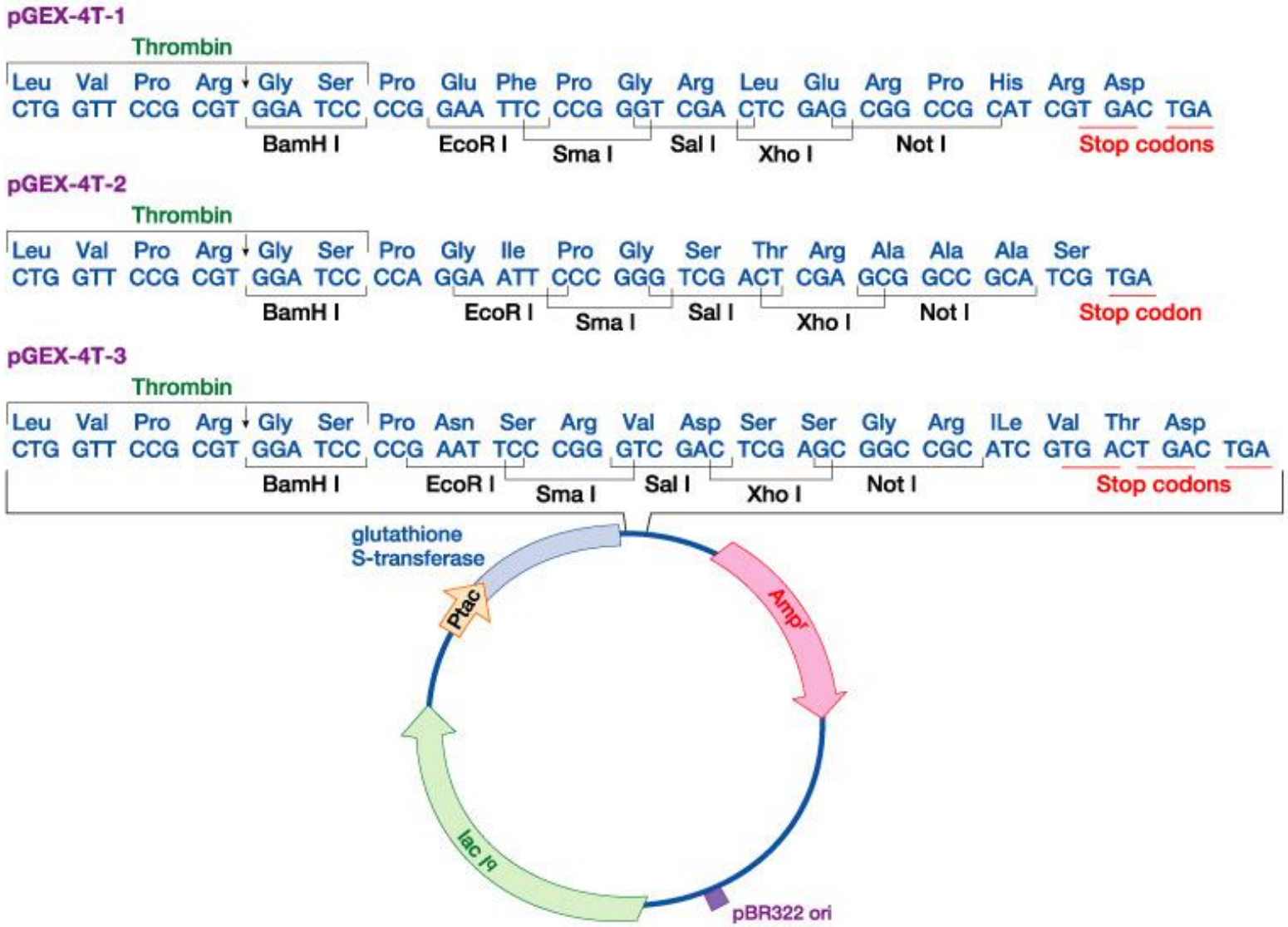


Figure 5-21 part 2 of 2 Human Molecular Genetics, 3/e. (© Garland Science 2004)

So do you use bacterial expression system every single time?

- Codon usage.
- Toxicity
- Sugar linkage or other modifications.
- Folding issue and disulfide bond formation.
- Unstable mRNA

Companies have developed a lot of bacterial strains to improve.

Troubleshoot: Strategies to overcome the problems

- If the protein is from eukaryotes, use codon plus strain if there are a lot of rare codons.
- If the protein is toxic, use (DE3) strain to prevent expression before induction. Additionally, pLysS or pLysE strain can make sure no protein expression until addition of IPTG.
- If protein needs sugar or disulfide bond, co-express proteins that conduct the post-translational modification.
- If the protein has problem in folding, you can either refold from inclusion body or use GroEL or DnaK supplemented strain. Sometimes using special vector with thioredoxin can help.
- Strain with defective mRNA degradation pathway can help stabilize mRNA.

But after you try everything, what to do if it is still not working

- Eukaryotes system:
  - Yeast
    - Cheap and fast turn-around
    - Sometimes good yield.
    - Translational modification still very different from mammalian system.
  - Baculovirus and insect cells.
    - Stable expression.
    - Great yield.
    - Overcome problem like folding or glycosylation.
    - Less non-specific band during purification.
    - But long turn-around.
    - $$$
  - Mammalian
    - Very pricey.
    - Yield low.
    - Expression unstable
    - Only used when small amount is required.

Insect cell expression system

- Step one: generate your expression vector using bacteria.
- Step two: grow virus that contains the vector.
- Step three: transfect virus to cell.
- Step four: harvest cell (many days later)