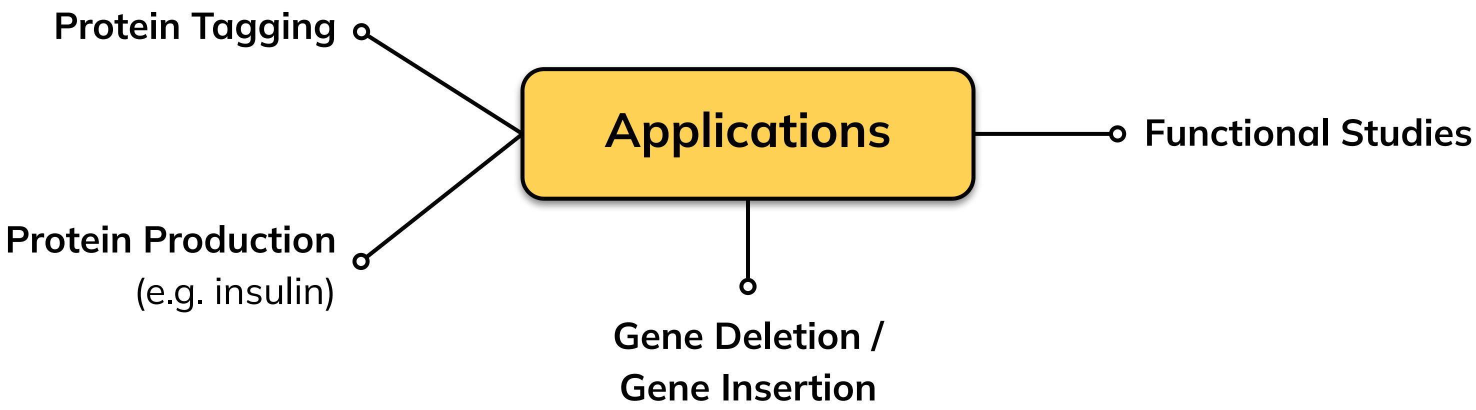
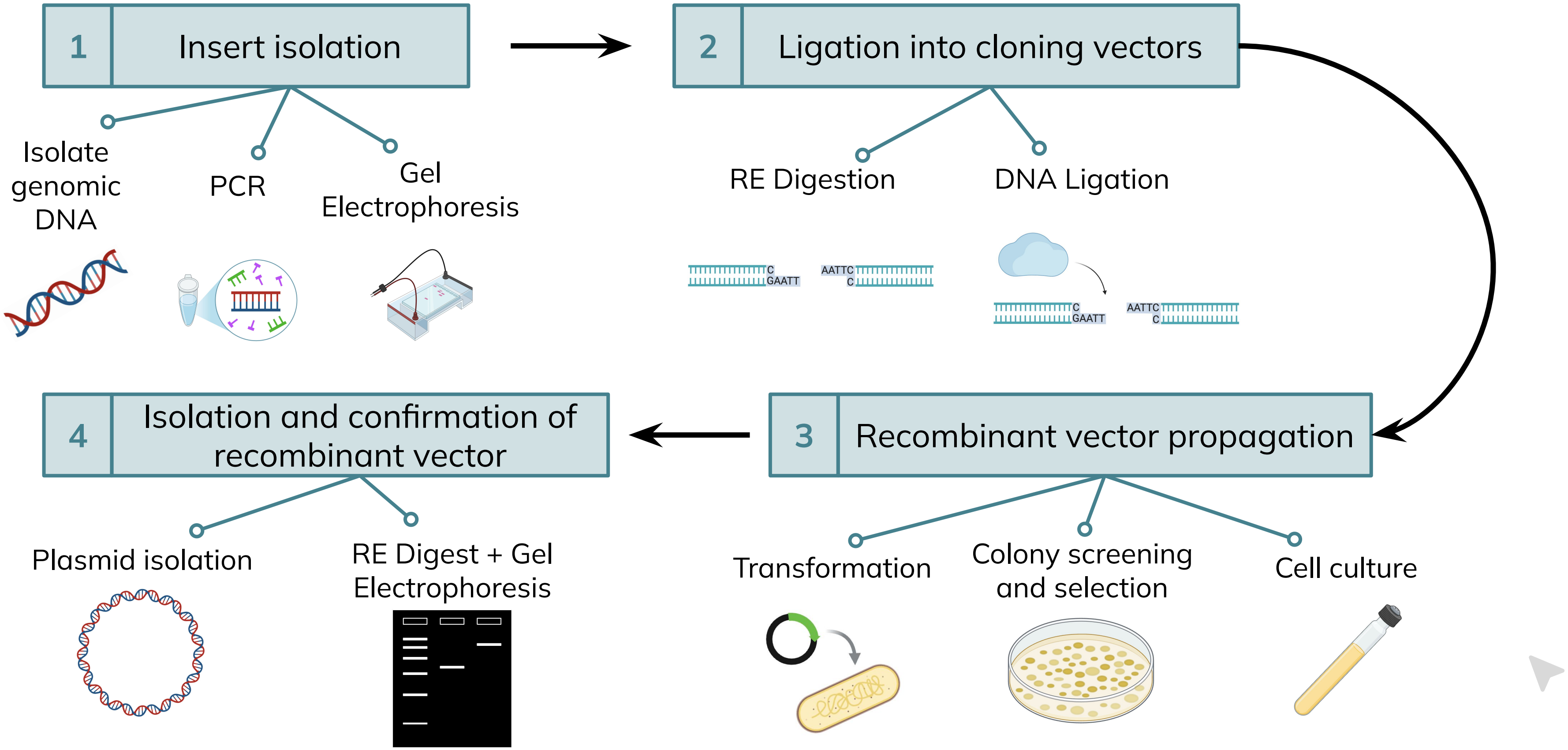
# Molecular Cloning Pipeline

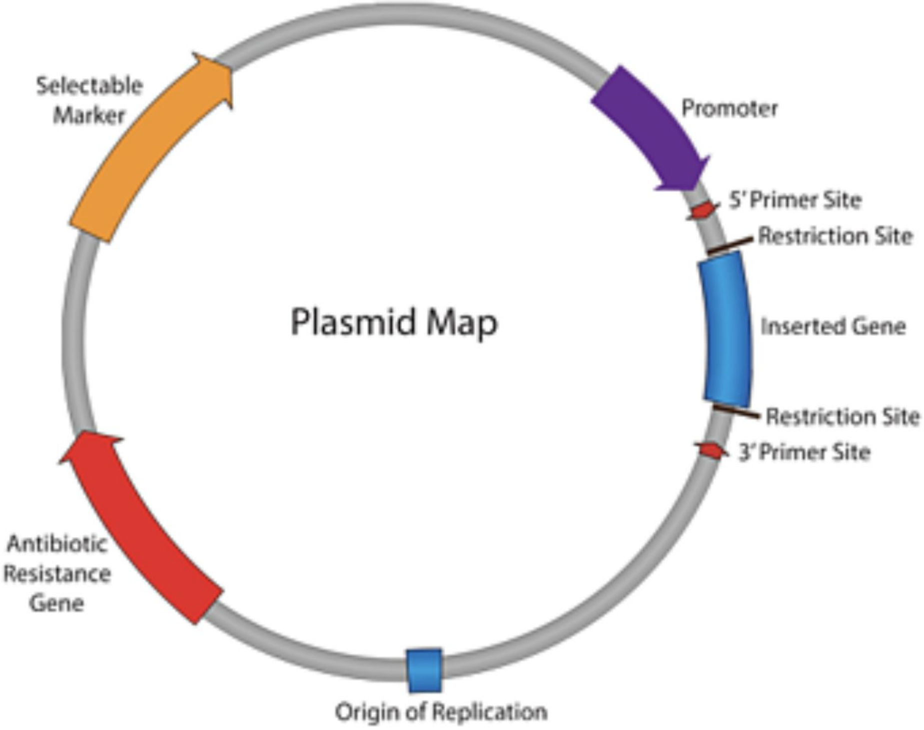
**Objective:** To insert a **gene of interest** into a **cloning vector** to produce a **recombinant plasmid** expressing the *protein of interest*



# Overview of the Molecular Cloning Pipeline

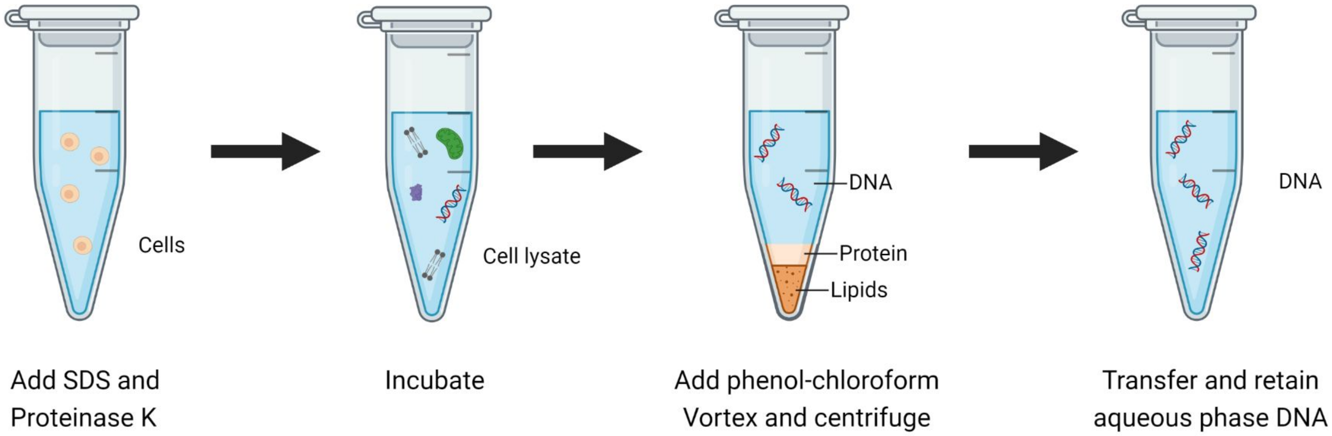


# Prerequisites: Cloning Vectors

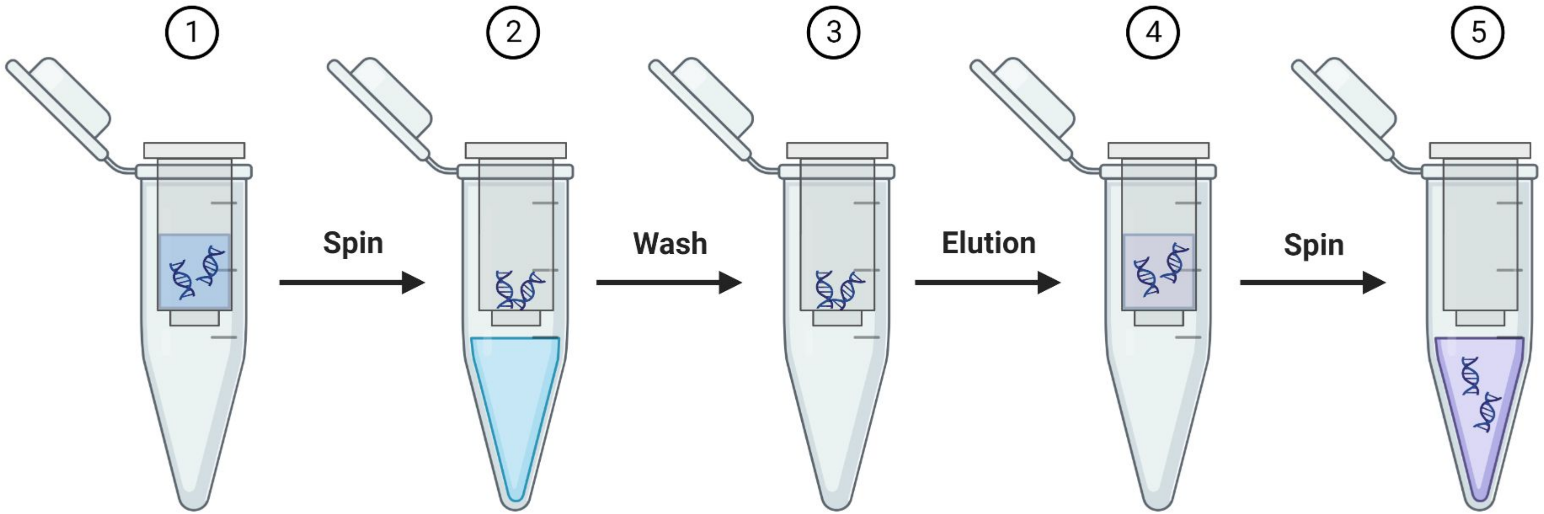


# DNA Isolation

**Purpose:** Obtain **template** for PCR to amplify **insert**.



# DNA Purification



# Polymerase Chain Reaction (PCR)

**Purpose:** Amplify **large amounts** of **insert** for ligation into cloning vector.

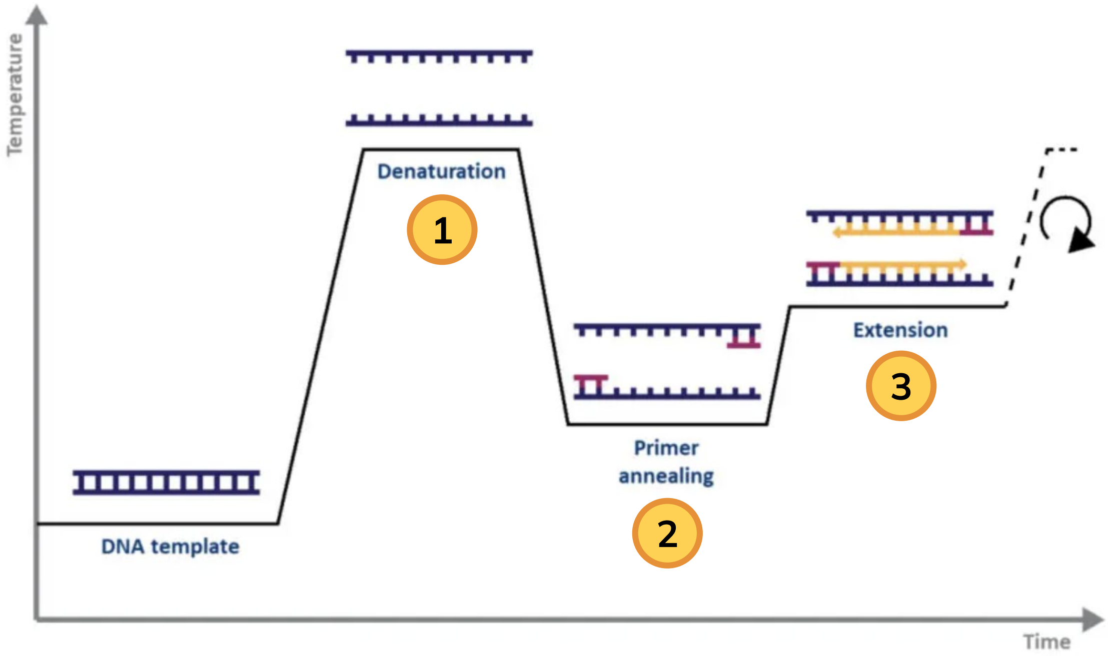



*Image adapted from https://www.enzo.com/note/what-are-the-differences-between-pcr-rt-pcr-qpcr-and-rt-qpcr/*

# PCR Reagents

| Reagent | Purpose |
|---|---|
| Forward and Reverse Primers | **Bind** to sequences flanking insert and provide **3’OH** for DNA Pol to begin extension |
| Thermostable DNA Polymerase (e.g. Taq polymerase) | Catalyses **phosphodiester bond formation** between 3’OH and incoming dNTP |
| Buffer | **Optimal pH** for DNA polymerase and contains **$Mg^{2+}$ ions** for DNA Pol function |
| dNTPs | **Building blocks** for DNA synthesis |

# PCR Considerations

| Consideration | Notes |
| --- | --- |
| Annealing temperature | Depends on **Tm** (hence on GC content and length) of primer pair.<br><br>Set 2–5°C below Tm of the primers. |
| Extension time | Depends on **length** of fragment to be amplified.<br><br>Taq polymerase: ~1kb/min |
| Maximum insert size | For longer fragments (**≥5kb**), mismatch errors likely to accumulate and Taq polymerase begins to ‘fall off’ DNA while extending. |

# Question Walkthrough

One cycle of the PCR reaction doubles the number of DNA fragments. Further, each time one cycle of the PCR reaction progresses, the primer pair, the substrate dNTP, and the DNA polymerase molecule are required double amount, so the amount of these components limits the overall amount of DNA that can be synthesized in PCR.

The length of the DNA fragment to be amplified was 100 base pairs including the primers, and the PCR reaction was started with the primer length of 20 bases. The four types of bases A, C, G, and T are evenly distributed in the sequence to be amplified, and the amplification efficiency of PCR is 100%. As the PCR reaction progresses, the reaction will not be completed due to running out of one of the components in a certain cycle.

**Choose the correct No. of the reaction stop cycle and the limiting component.** 60 (3 points)

Template DNA fragment : 4 copies

Primer : 1,000 sets

dNTPs (dATP, dTTP, dGTP, dCTP) : 48,000 molecules (12,000 molecules each)

DNA polymerase : 1,200 molecules

| No. | Cycles | Limiting component |
|---|---|---|
| (1) | 7 | Primer pairs |
| (2) | 7 | dNTPs |
| (3) | 7 | DNA polymerase |
| (4) | 8 | Primer pairs |
| (5) | 8 | dNTPs |
| (6) | 8 | DNA polymerase |
| (7) | 9 | Primer pairs |
| (8) | 9 | dNTPs |
| (9) | 9 | DNA polymerase |
| (0) | Others | |

*Images adapted from IBO 2020 Theory 2 Question 14*

# Primer Design for Molecular Cloning

| Consideration | Notes |
| --- | --- |
| Primer Length | 18bp–30bp |
| Melting Temperature (Tm) | $Tm = 4(G+C) + 2(A+T)$<br>Ideal Tm is between 52–58°C, and the primer pair should be within 2-3°C of each other's Tm. |
| GC Content | Ideally 50-55% |
| Primer Dimers | Primer dimers form through binding between complementary regions of two primers.<br>Ensure primers are designed to minimise complementarity. |

# How to Design a Primer

| Primer Component | Purpose |
| --- | --- |
| Complementary Sequences | **Forward** primer: Sequence flanking **upstream** of target gene<br><br>**Reverse** primer: Sequence flanking **downstream** of target gene |
| Restriction Sites | Overhang sequences with **restriction sites** and three additional nucleotides. |
| 3’ GC clamp | Presence of **G/C bases** at 3’ end of primer to promote binding stability at DNA polymerase extension site.<br><br>Note: This component may not always be present. |

# How to Design a Primer: Forward Primer

5’ATGAAGTTATTGAGCAATAGTCTAATGTTCCTTCCTTCATATGCTTCGTCTTGA3’

3’TACTTCAATAACTCGTTATCAGATTACAAGGAAGGAAGTATACGAAGCAGAACT5’

**EcoRI:** 5’GAATTC3’

**Forward Primer:** 5’CCC**GAATTC**ATGAAGTTATTGAGC3’

# How to Design a Primer: Reverse Primer

5'ATGAAGTTATTGAGCAATAGTCTAATGTTCCTTCCTTCATATGCTTCGTCTTGA3'

3'TACTTCAATAACTCGTTATCAGATTACAAGGAAGGAAGTATACGAAGCAGAACT5'

**SalI:** 5'GTCGAC3'

**Reverse Primer:** 5'CCC**GTCGAC**TCAAGACGAAGCATA3'

# Question Walkthrough

Part of the sequence of vector A, which is for protein expression using *Escherichia coli* as a host, is shown. It was planned to express a plant-derived gene X using vector A. Vector A is a plasmid vector that expresses a protein fused to the N-terminus His-tag, which enables efficient purification of the expressed protein. As shown in Figure 1, translation of the protein occurs from the start codon immediately before the His tag with six consecutive His residues. The DNA sequences of the 5' and 3' regions of gene X are shown in Figure 2. We planned to clone gene X using restriction enzyme sites, EcoRI, SmaI, or SalI in vector A. When the gene X is amplified by PCR, a fragment with a restriction enzyme site at the end can be amplified using the primer with a restriction enzyme site. Since the restriction enzyme site is not recognized if it is located at the end of the DNA fragment, three "Cs" were also attached in addition to the restriction enzyme site.

For example, in order to add the EcoRI site to 5'-XXXXXXXXXX----, the primer is designed as below.

5'-CCCGAATTCXXXXXXXXXX----,

**Choose true if the primer is a correct one to use, if not, choose false.**

**A.** Forward primer ( $\boxed{\phantom{ATG}}$ : the start codon)

5'-CCC GAA TTC $\boxed{\text{ATG}}$ AAG TTA TTG AGC AAT A-3' $\boxed{60}$

EcoRI site

**B.** Forward primer ( $\boxed{\phantom{ATG}}$ : the start codon)

5'- CCC CCC GGG $\boxed{\text{ATG}}$ AAG TTA TTG AGC AAT A-3' $\boxed{61}$

SmaI site

**C.** Reverse primer ( $\boxed{\phantom{ATG}}$ : the stop codon)

5'-CCC GTC GAC $\boxed{\text{TCA}}$ AGA CGA AGC ATA TGA T-3' $\boxed{62}$

SalI site

**D.** Reverse primer ( $\boxed{\phantom{ATG}}$ : the stop codon)

5'-CCC AAG CTT GTA GGT AGT ATA CGA AGC AGA $\boxed{\text{ACT}}$ -3' $\boxed{63}$

HindIII site

```
        Start codon          His tag
--- ATA CAT [ATG] GCA [CAT CAC CAC CAC CAT CAC] TCC GCG GCT CTT GAA GTC CTC TTT CAG GGA
--- TAT GTA TAC GCA GTA GTG GTG GTG GTA GTG AGG CGC CGA GAA CTT CAG GAG AAA GTC CCT

CCC GGG TAC CAG GAT CCG AAT TCT GTA CAG GCC TTG GCG CGC CCG ACG TCC GTC GAC AAG CTT---
GGG CCC ATG GTC CTA GGC TTA AGA CAT GTC CGG AAC CGC GCG GGC TGC AGG CAG CTG TTC GAA---
SmaI                EcoRI                                       SalI    HindIII
```

Figure 1. DNA sequence of the cloning region of vector A (double strands).

```
Start codon
[ATG] AAG TTA TTG AGC AAT AGT CTA ATG TTC CTT CCT CTG CTG GCT TTG GCT ---
TAC TTC AAT AAC TCG TTA TCA GAT TAC AAG GAA GGA GAG GAC CGA AAC CGA ---

--- TCT TCC TTC CTC AAG GGA ACA CTG CAC CAT CCA TCA TAT GCT TCG TCT [TGA]
--- AGA AGG AAG GAG TTC CCT TGT GAC GTG GTA GGT AGT ATA CGA AGC AGA ACT
                                                          Stop codon
```

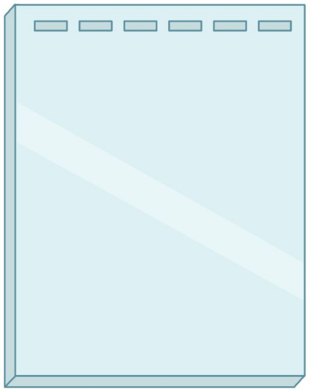
Figure 2. DNA sequence of the gene X showing 5' region and 3' region: 1566 base pair

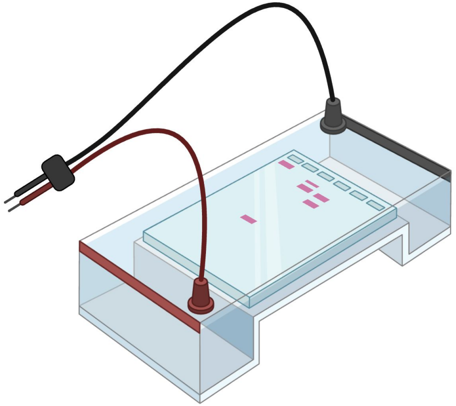
*Images adapted from IBO 2020 Theory 1 Question 16*

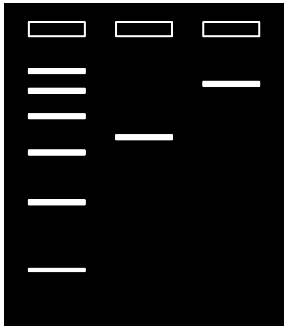
# Agarose Gel Electrophoresis

**Purpose: Separate** DNA fragments based on molecular weight to check whether insert has been **correctly amplified**.

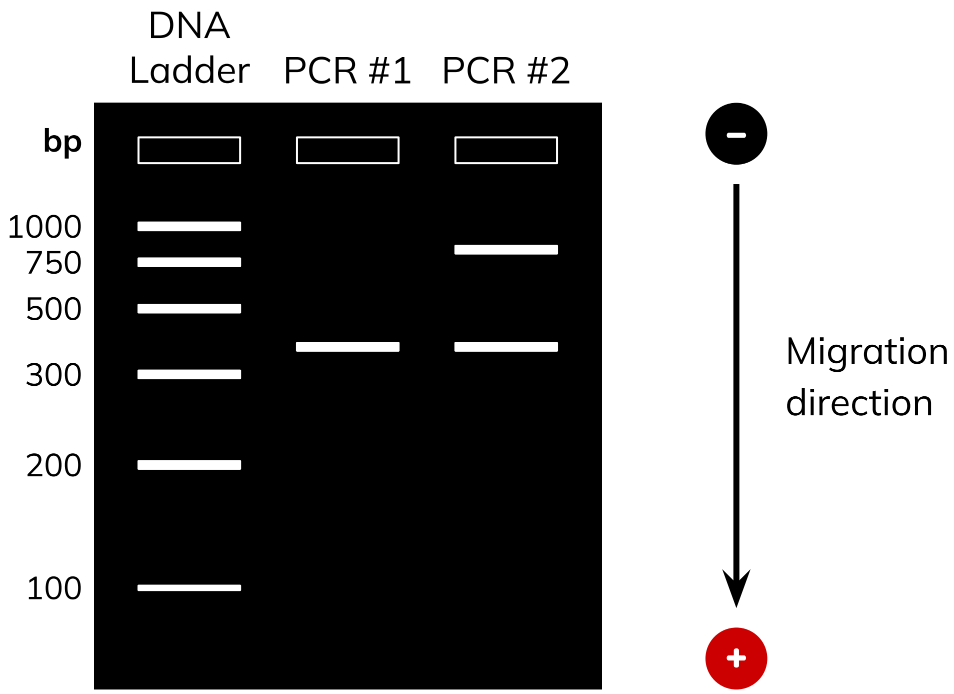
Prepare and stain gel with Ethidium Bromide → Prepare and load samples, then run gel → Image gel under UV







# Agarose Gel Electrophoresis

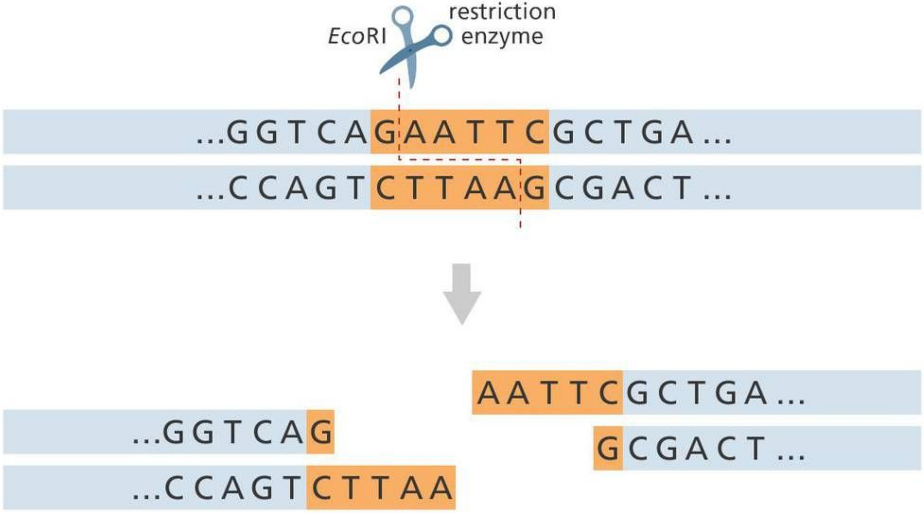


# Agarose Gel Electrophoresis Considerations

| Consideration | Notes |
| --- | --- |
| Agarose percentage | **Higher** percentage for resolving **smaller** fragments.<br>**Lower** percentage for resolving **larger** fragments. |
| Voltage | **Higher** voltage results in **faster** migration and vice versa.<br><br>**Excessive** voltage may result in gel melting and affect DNA migration. |
| Running duration | **Longer** time to resolve **larger** fragments. |

# What are Restriction Enzymes (REs)?

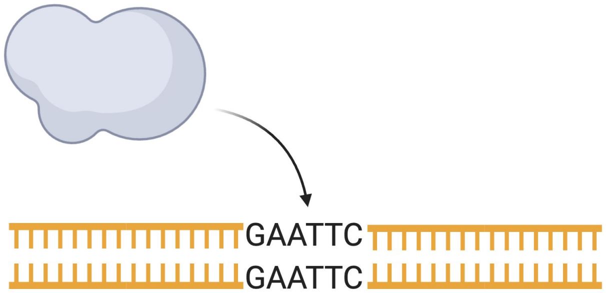
Enzymes that cut DNA at **specific recognition sequences**. *Type II restriction endonucleases* recognise **palindromic** sequences and create **sticky ends**.



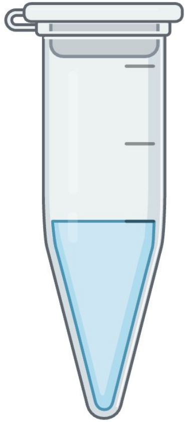
# RE Digestion Workflow

**Purpose:** Generate **sticky ends** for **ligation** of insert into vector.

Choose appropriate *restriction enzymes* and *buffer*




Mix reactants and incubate at 37°C

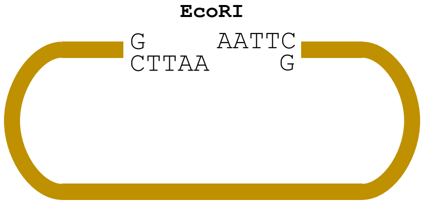


# How to Design RE Digests

## 1-enzyme digests

**Insert**

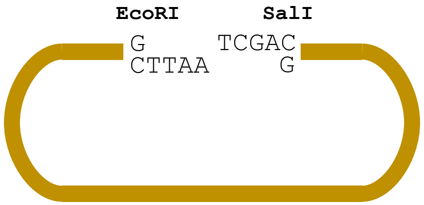

**Plasmid**



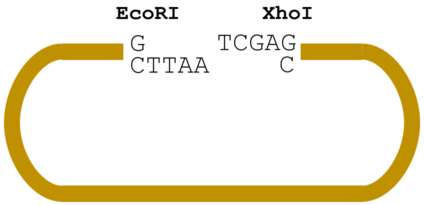

## 2-enzyme digests

**Insert**


**Plasmid**




OR



# DNA Ligation

**Purpose: Ligate** and **seal** the insert into the cloning vector.

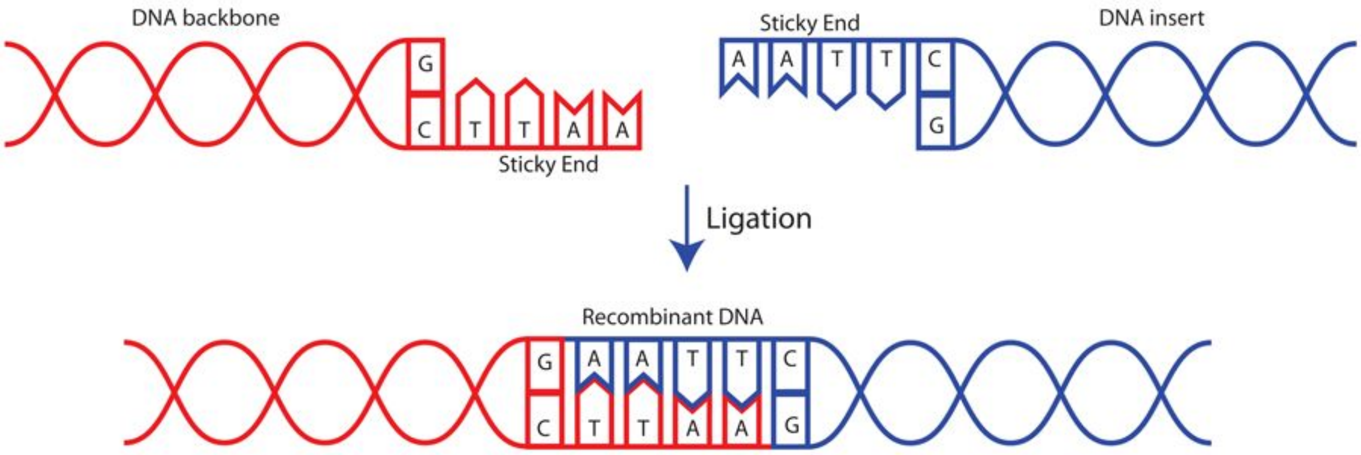
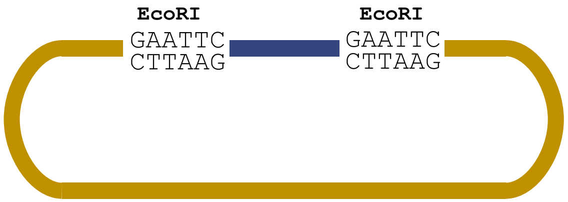



*Image adapted from https://www.addgene.org/protocols/dna-ligation/*

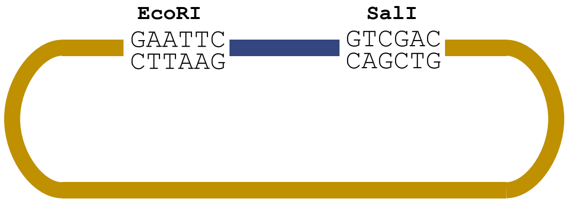
# Outcomes of DNA Ligation

## 1-enzyme digests

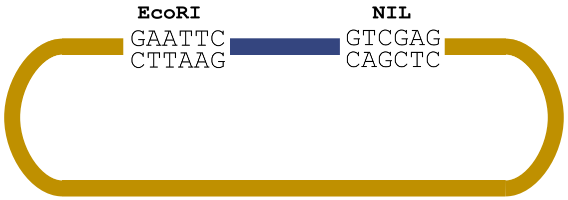



Secondary products are formed

## 2-enzyme digests

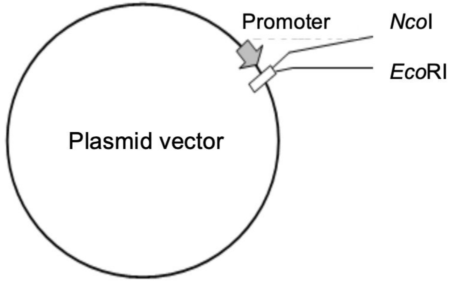



Restriction site destroyed



# Question Walkthrough

You plan to insert the gene *PhoQ* from *Tobibacterium* sp into a plasmid vector containing an artificial promotor followed by a restriction site for *Nco*I (CCATGG) and a restriction site for *Eco*RI (GAATTC).




To conduct this experiment you are required to design forward (the sense strand) and reverse (antisense strand) primers. Part of the 561 nucleotide long coding sequence is shown below.

5'-ATGCGACAGTTCATCACCGA... ______....GCGGGACCGGACTGGGGTAA-3'

**Indicate if each of the following statements is true or false.**

A. The use of two different restriction sites avoids wrong orientation of the inserted fragment

B. A possible forward primer for amplification and insertion of *PhoQ* gene will have the following sequence: 5' – GATCCCATGGATGCGACAGTTC – 3'

C. A possible reverse primer for amplification and insertion of *PhoQ* gene will have the following sequence: 5' – GATCGAATTCAATGGGGTCAGGCC – 3'

D. The final gene product will consist of at least 189 amino acids.

*Image adapted from IBO 2014 Theory 1 Question 3*

# *E. coli* Transformation

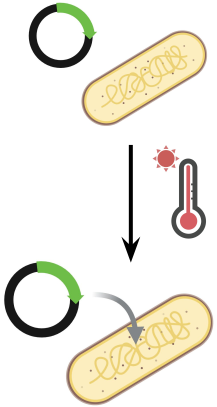
**Purpose:** Introduce recombinant plasmid into bacterial cells to **replicate large amounts of plasmid**.

## Cell Preparation

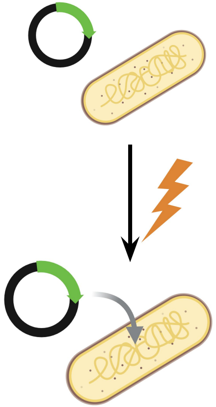
Suspend bacterial cells in **$CaCl_2$** to make them competent → increase *transformation efficiency*

## Transformation Methods

### Heat Shock



### Electroporation

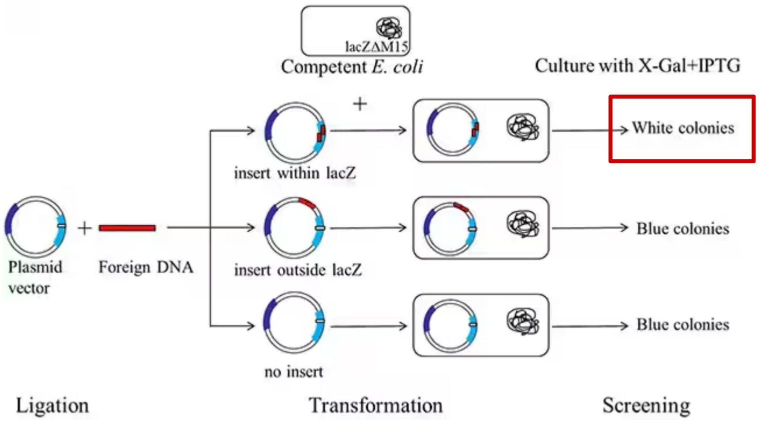


# Screening and Selection of Recombinant Plasmids

## + Selection Systems

Grow on agar plates containing **antibiotics** (antibiotic selection) or **lacking select nutrients** (auxotrophy) to select for transformed cells.

## - Selection Systems



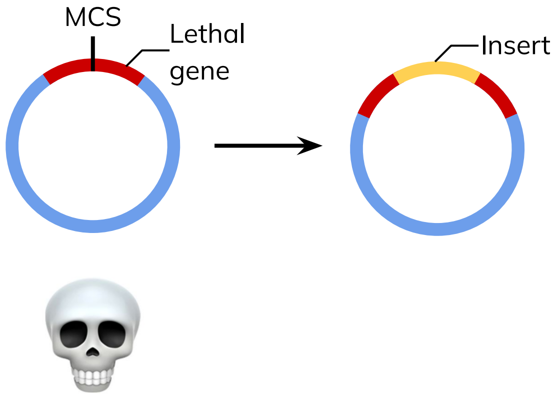
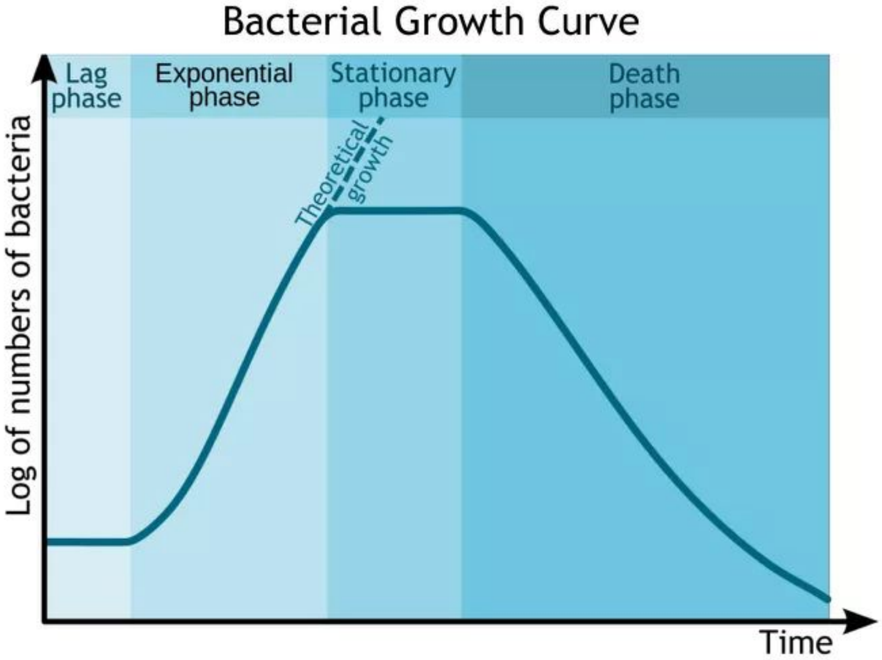




*Image adapted from https://www.sigmaaldrich.com/SG/en/technical-documents/technical-article/genomics/cloning-and-expression/blue-white-screening*

# *E. coli* Cell Culture

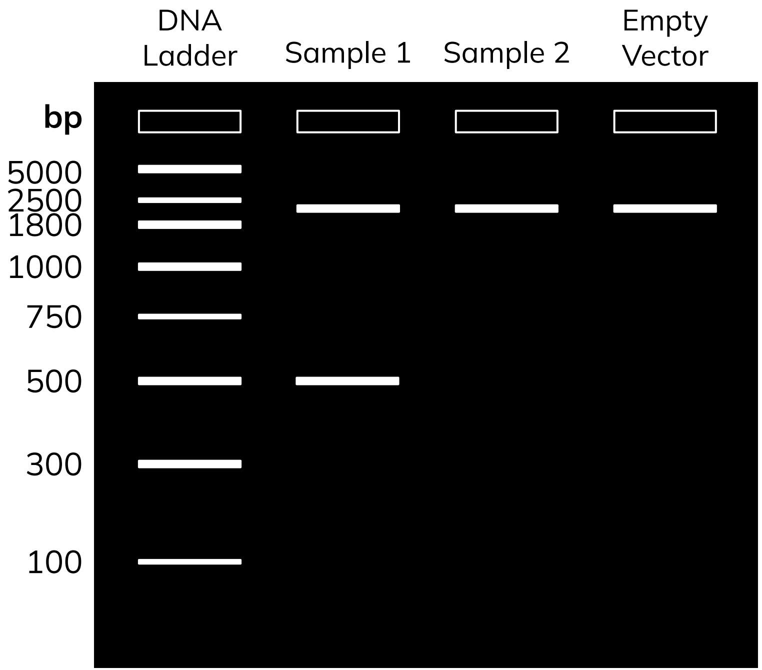
Inoculate **single colonies** into LB media, grow at 37°C until **mid-log phase**.



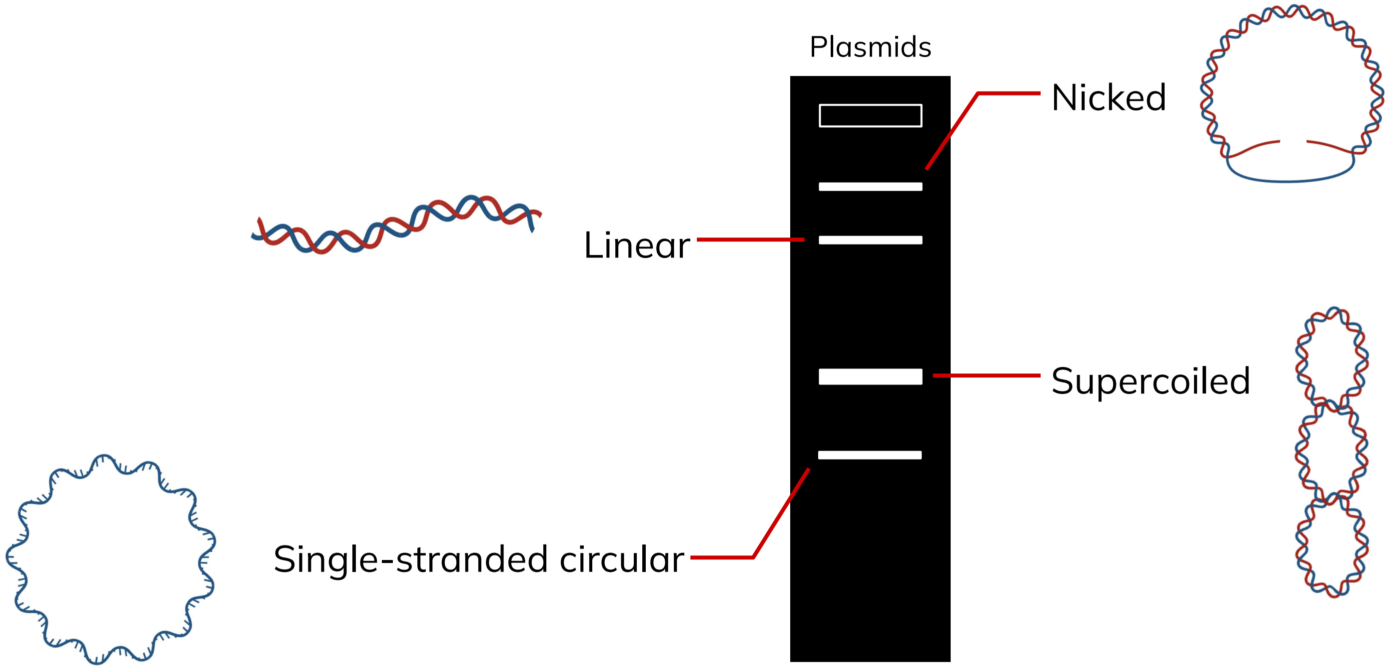
# Confirmation of Recombinant Plasmid

**RE Digest** → Use **same REs** that you did for cloning (assuming the restriction site wasn't destroyed).

**Gel Electrophoresis** → Run alongside cut **empty vector control**.

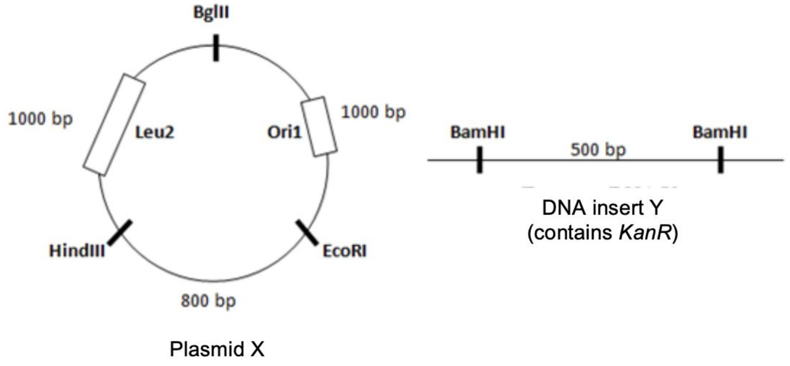


# A Note on Electrophoresis of Plasmids



# Question Walkthrough

An experiment was performed to create a recombinant DNA between plasmid X and DNA insert Y. Plasmid X contains *leu2* gene for leucine biosynthesis while DNA fragment Y contains kanamycin-resistance gene *KanR*. The diagram for X and Y are shown below.




Plasmid X and DNA insert Y were added to a reaction mix containing restriction enzymes *Bgl*II (5'-A*GATCT-3'), *Bam*HI (5'-G*GATCC-3'), and the resulting fragments then transferred to a new reaction mix containing ligase. The resulting DNA was transformed into bacterial culture Z which is sensitive to kanamycin and unable to survive in leucine-deficient medium. Selection for transformed Z cells containing the recombinant plasmid (plasmid X with DNA insert Y) was performed by growing the culture on a selective medium containing kanamycin and no leucine. The recombinant plasmid was then isolated from the culture. Assume that all isolated plasmids are in circular conformation and there is no partial restriction reaction, all plasmids are cut completely by the restriction enzyme. (Note: (*) indicates the location of bond hydrolysis)

**Indicate if each of the following statements is true or false.**

A. Cutting the plasmid after insertion of Y using *Eco*RI will result in single 2800 bp DNA fragment on the electrophoresis gel.

B. If in the reaction mix HindIII (5'-A*AGCTT-3') was used instead of BglII, the transformed bacteria are capable of growing on a medium containing Kanamycin.

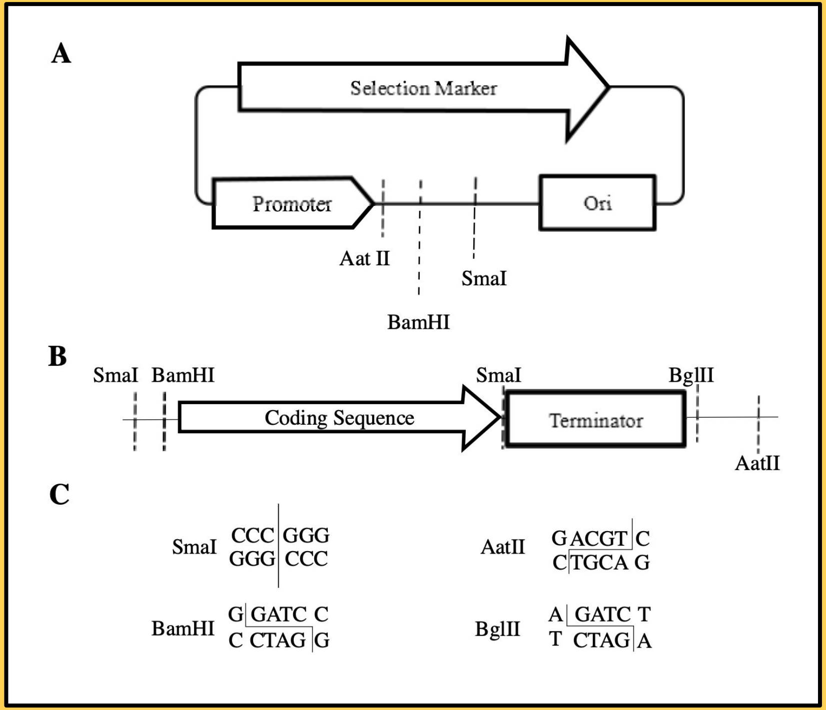
C. The 500 bp insert DNA can be removed from the recombinant plasmid by using restriction enzyme *Bgl*II.

D. The migration pattern of the recombinant plasmid on an electrophoresis gel is different when it is treated with EcoRI or BamHI.

*Images adapted from IBO 2014 Theory 1 Question 5*

# Question Walkthrough

A gene (coding sequence) can be expressed by cloning it into an expression plasmid using restriction enzymes and DNA-ligase. A plasmid (A), a gene of interest (B), and the recognition sequences for four restriction enzymes (C) are shown in the figure. Different cloning strategies, expressed in the statements below, could be used to insert the"Coding sequence and Terminator" of this gene into the plasmid to produce a recombinant plasmid that expresses the gene.




**Indicate if each of the following statements is true or false.**

| | TRUE | FALSE |
|---|---|---|
| Digestion with SmaI followed by ligation can produce the desired recombinant plasmid. | | |
| Digestion with AatII and BamHI followed by ligation can produce the desired recombinant plasmid. | | |
| Digestion with BamHI + BglII followed by ligation can produce the desired recombinant plasmid. | | |
| The 'coding sequence' needs to be in-frame with the promoter | | |

*Images adapted from IBO 2015 Theory 2 Question 16*

Next up... Sequencing!
Covered in next submodule.

# Golden Gate Cloning

**Innovation:** Allows for *simultaneous* and *directional* insertion of **multiple DNA fragments** into a plasmid vector.

Uses **Type IIS REs** – cut outside restriction site and generate non-palindromic overhangs.

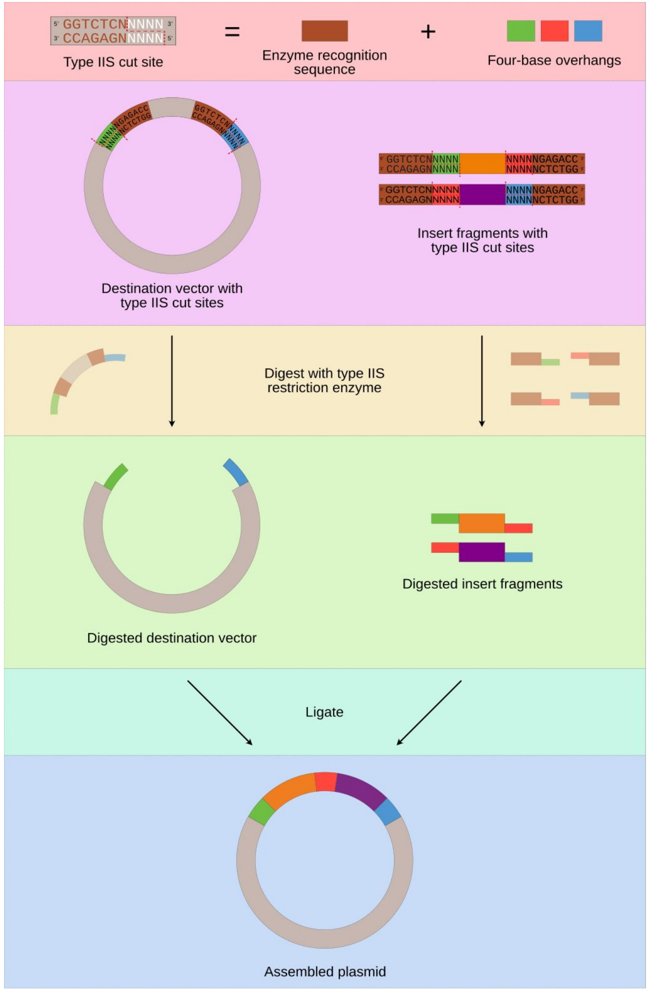



*Image adapted from Rob Hurt - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=88975448*

# Gibson Assembly

**Innovation:** Does not rely on restriction sites for DNA fragment insertion. Instead, it allows for **insertion anywhere in the vector**.

Requires **5' to 3' exonuclease**.

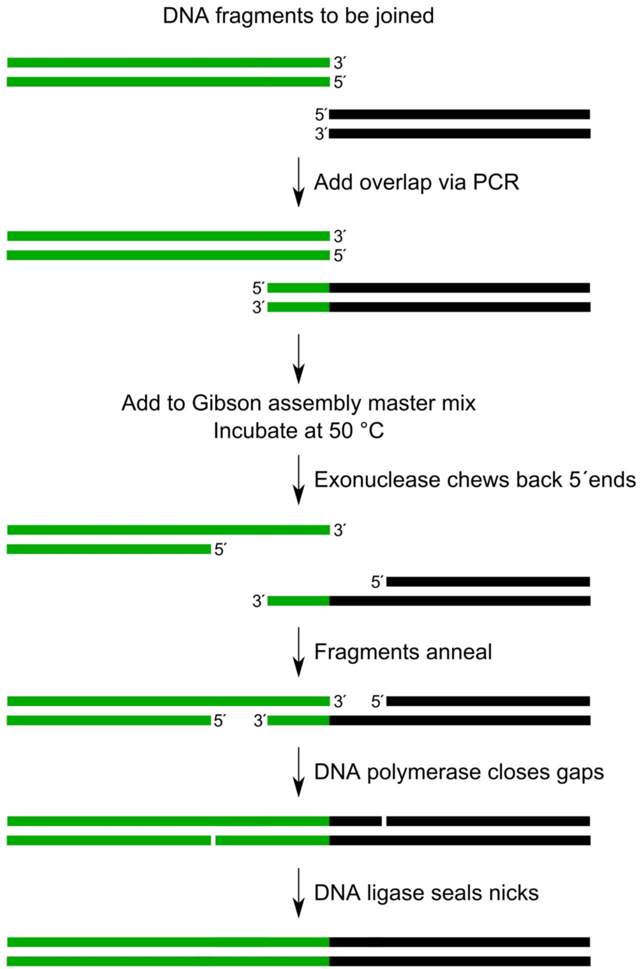



*Image adapted from Tobias Vornholt - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=39652308*

Thank you!