Cellular & Molecular Biology
SQBS 1143

Chapter 8:
Techniques in Molecular Biology

Prof. Dr. Noor Aini Abdul Rashid
Dr. Chan Giek Far

Introduction

- Suppose we want to create our very own monster by genetic engineering.

- How do we go about it? Frankenstein made his monster by sewing a brain into a body and then charging up his creation with a lightning bolt.

- Genetic engineers make 'patchwork' organisms not by joining organs but splicing genes together. So let's get started on some basic operations.

What is Genetic Engineering?

- refers to a set of technologies used to **change the genetic make-up** of cells and **move genes across species** boundaries to produce novel organisms

- techniques are highly sophisticated in manipulation of the genetic material

- several other terms that can be used to describe this technology: **gene manipulation**, **gene cloning**, **recombinant DNA technology**, **gene modification** and the **new genetics**.

First, Catch The Mouse

- The first step is **to get hold of some DNA**. Since all living organisms have genes, they all have DNA which you can **extract**.

- If you intend to tackle something large like a tree or a hippopotamus, it is only necessary to hack off a small sample.

- Nowadays, genetic manipulations using laboratory mice or rats are often done by snipping off a piece from the end of the tail.

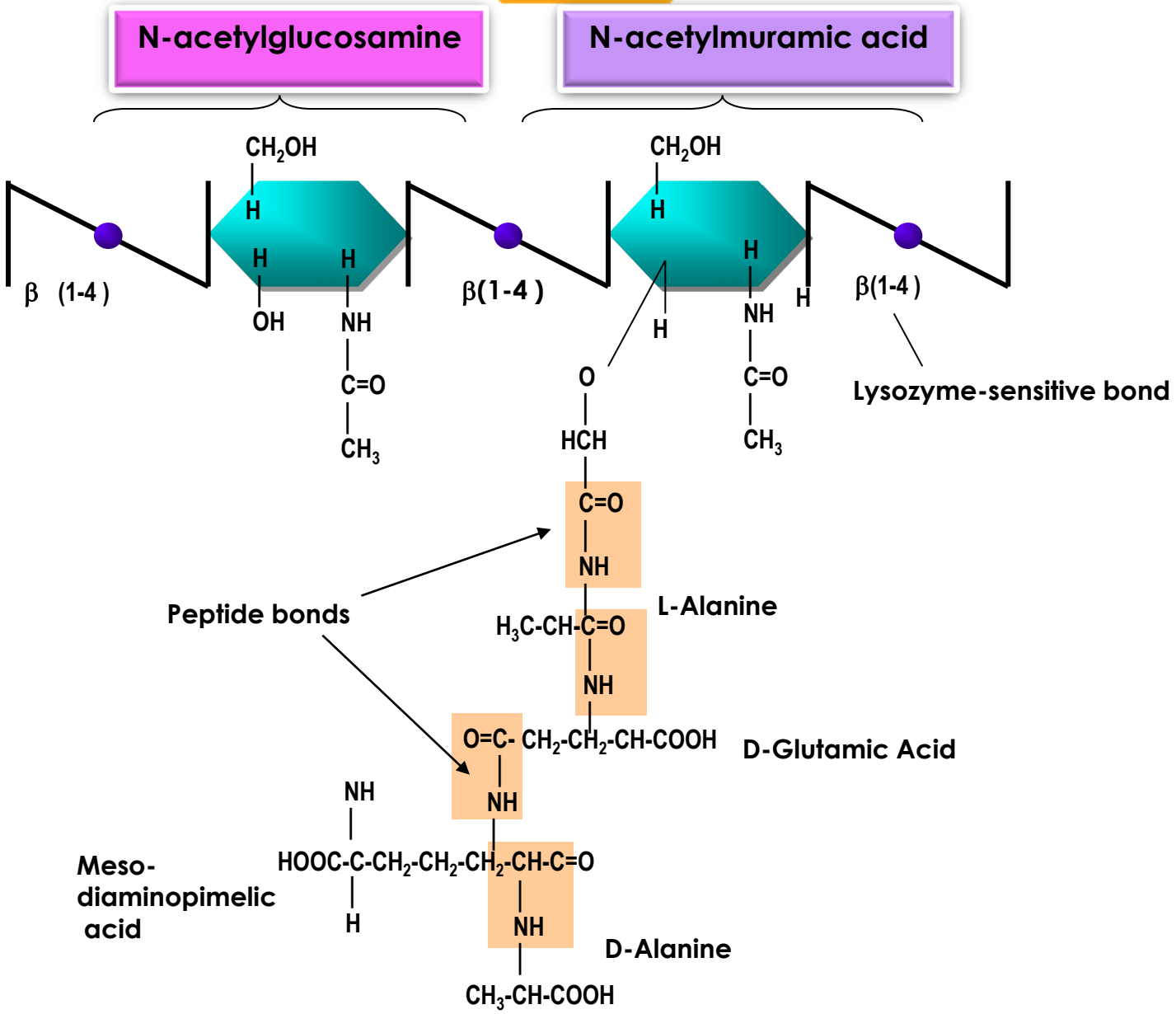
- However, by far the easiest place to get DNA is from **bacteria**. A few drops of a bacterial culture will give plenty of DNA for most purposes and since the bacteria are single cells and contain no bone, fat, etc., the DNA is relatively easy to get out.

- For **animals and plants**, we must grind sample into tiny fragments before proceeding.

- Next, we break open the cells. This may be done **mechanically** in a blender or it may be done using **chemicals** to degrade and dissolve the components of the cell walls.

- For **bacteria**, we usually use a mixture of **lysozyme** that digests the outer layer of the cell wall, followed by a **detergent** that dissolves the greasy cell membrane.

- For a mouse's tail we use **enzymes** that degrade the connective tissue and disperse the cells. The DNA is then liberated into solution and is purified by further series of steps.


N-acetylglucosamine
N-acetylmuramic acid
β (1-4)
β(1-4)
β(1-4)
Lysozyme-sensitive bond
Peptide bonds
L-Alanine
D-Glutamic Acid
Meso-diaminopimelic acid
D-Alanine

Second, Purifying the DNA

- Two general types of procedure are used here, **centrifugation** and **chemical extraction**. First we centrifuge, then we extract.

- Centrifugation is a routine technique in molecular biology labs.

- The basic idea is quite simple: the sample is spun at high speed and the centrifugal force causes the larger or heavier components to sediment to the bottom of the tube.

- For example, after destroying the cell wall of bacteria by lysozyme and detergents, we are left with a solution containing **the fragments of the wall, which are small,** and **the DNA, which is a gigantic molecule**.

- So the next step is to centrifuge. The DNA and some other large components are hurled to the bottom of the tube and the garbage from the destruction of cell wall remains in solution and is discarded.

- The **sedimented DNA** is then redissolved. However, it still has **lots of proteins and RNA mixed in** with it. We extract these by chemical means. For example, one step used in almost all DNA purifications uses the chemical phenol.

- **Phenol**, also known as carbolic acid, is very corrosive and extremely dangerous. The reason for this is that, it dissolves and denatures the proteins which make up 60 to 70 percent of all living matter.

- If you add phenol to a sample of DNA, it will dissolve and remove most of the proteins in your sample, and in doing so, purify the DNA. When phenol is added to water, the two liquids do not mix to form a single solution; instead **the denser phenol** forms a separate **layer below the water**.

- When shaken, the two layers mix temporarily; when the shaking stops, the layers separate again. The phenol **dissolves almost all of the protein**, but the nucleic acids, DNA and RNA, stay in the water layer.

- To ensure good separation of the layers we centrifuge briefly. Then, we suck off and **keep the water layer** with the DNA and RNA.

Two layers mixed temporarily.

Separation of two layers.

Getting Rid of the RNA

- To get rid of the RNA, we use an enzyme that **degrades RNA**, **ribonuclease**. This converts the RNA into tiny fragments but leaves the DNA unchanged as a giant macromolecule.
- We now add an equal volume of alcohol. The alcohol dissolves in the water so enthusiastically that it occupies all of the water and pushes the larger and less soluble DNA out of solution. However, the small RNA fragments remain dissolved.

- We now centrifuge again and the DNA is sedimented to the bottom of the tube and we can pour off the solution containing the RNA fragments.

- The **tiny pellet of DNA left** at the bottom of the tube is **often scarcely visible**. Nonetheless, it contains billions of DNA molecules, sufficient for most experiments. This DNA is redissolved and is now ready for use in genetic engineering.

Cutting up the DNA

- The DNA isolated from a typical bacterial cell is a giant circular chromosome with about 3,000 genes on it. In genetic engineering, we normally work with one or two genes, or only a fragment of a gene, at a time.

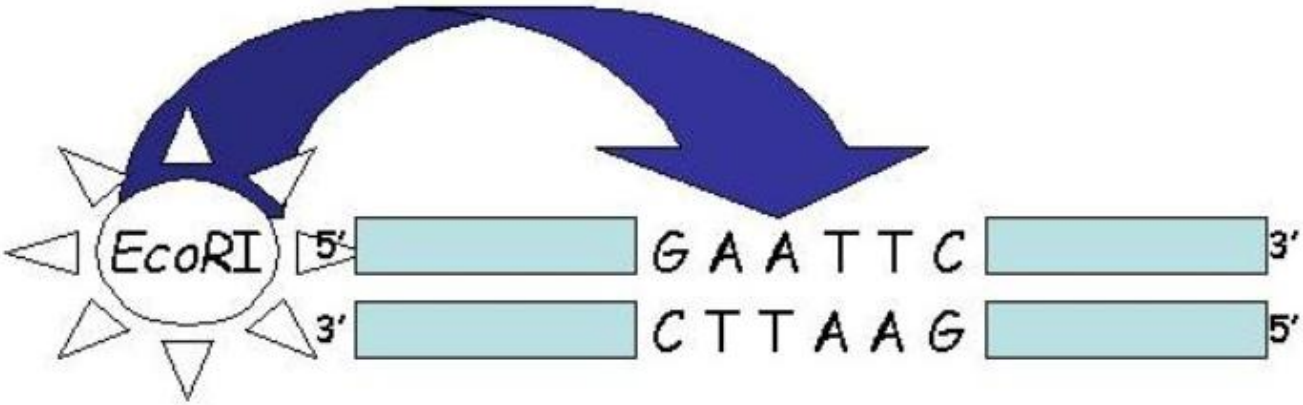
- So we need to cut the DNA into pieces of a manageable size. This is done by a special type of enzyme known as a **restriction enzyme** which acts as molecular scissors.

- Restriction enzymes all bind to DNA at a specific sequence of bases, the **recognition site**. This base sequence is usually four, six or eight bases long and the bases form an **inverted repeat**.

- In an inverted repeat, the sequence on the top strand of the DNA is the same as the sequence of the bottom strand read in the reverse direction.

- Each restriction enzyme has its own **specific recognition site**. Since any random series of four bases will be found quite frequently, enzymes that recognize four bases will cut DNA into lots of very short pieces.

- Conversely, finding a particular eight-base sequence is rare, so eight-base recognizing enzymes cut DNA only at relatively long intervals and generate few large pieces.


EcoRI
5′
G A A T T C
3′
3′
C T T A A G
5′

- The **six-base enzymes** are the most convenient in practice as they give an intermediate result. Two examples of widely used restriction enzymes and their sites are shown below:

5' -G G A T C C - 3' 3' -C C T A A G – 5' Recognition site for *Bam*HI	5' -G A A T T C -3' 3' -C T T A A G -5' Recognition site for *Eco*RI

Where is the DNA Cut?

- You might think that DNA should be cut at the recognition site where the restriction enzyme binds. This is often true, but not always.

There are **two classes** of restriction enzyme:

- **Type I restriction enzymes** cut the DNA a thousand or more base pair away from the recognition site. This is done by looping the DNA so the enzyme can get a grip on it both at the recognition site and the cutting site.

- Since the exact length of the loop is not constant, and since the base sequence at the cut site is not fixed, these enzymes are not much use to molecular biologists.

- Even more bizarre is that these enzymes are suicidal. Most enzyme carry out the same reaction over and over again on a continual stream target molecules. Type I restriction enzymes perform a kamikaze attack and when they cut DNA, the enzymes are destroyed as well!

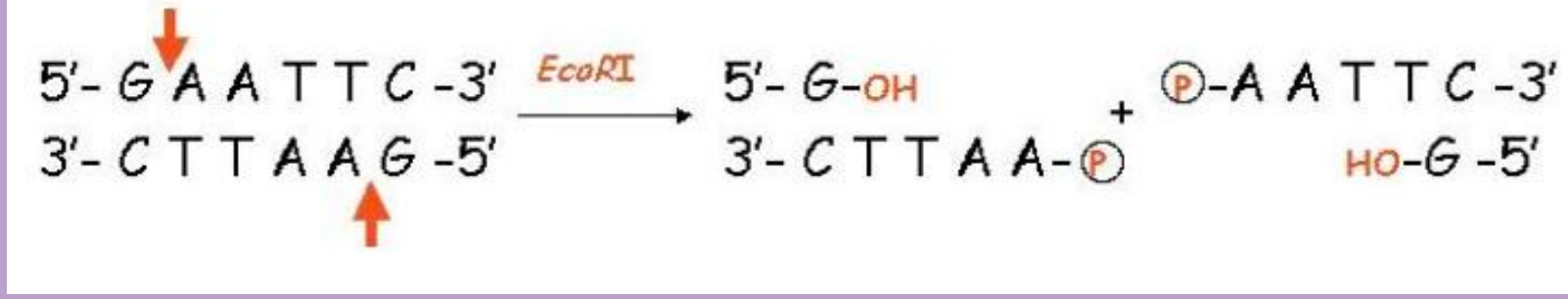
- **Type II restriction enzymes** cut the DNA in the middle of the recognition site. Since the exact position of the cut is known, these are the restriction enzymes which are normally used in genetic engineering.

- There are two different ways of cutting the recognition site in half. One way is to cut both strands of the double stranded DNA at the same point. This leaves **blunt ends**. The alternative is to cut the two strands at different places which generates overhangs. Because these overhangs are 'unmarried', they will base pair with each other and they are known as **sticky ends**.

BLUNT (*Hae* III)

Old genetic engineers don't die, they just come to sticky ends!

STICKY OR OVERHANGING (*Eco*RI)



- Enzymes that generate sticky ends are far more useful. This is because if two different pieces of DNA were cut with the same restriction enzyme, the same sticky ends should be generated on both fragments.

- This allows the two different pieces of DNA to be bound together by **matching the sticky ends**.

- Such pairing (called **annealing**) between the ends is temporary since the pieces of DNA are only held together by **hydrogen bonding** between the base pairs, not by permanent covalent bonds. Nonetheless, this gives time for the permanent bonding of the sugar-phosphate backbone.

How are Fragments of DNA Joined Together?

- Another type of enzyme is used to join DNA fragments permanently is **DNA ligase**, (the same enzyme we studied before) where it joined up the fragments of the lagging strand during DNA replication.

- If DNA ligase finds two DNA fragments touching each other, end to end, it will join them together.

- In practice, pieces of DNA with matching sticky ends will tend to stay attached much of the time and consequently DNA ligase will join them permanently without much trouble.

- Since DNA fragments with blunt ends have no way to hold on to each other, they drift apart most of the time. Ligating them takes a very long time and a lot of DNA ligase and even then it is the inefficient. In fact, bacterial ligase called ***E. coli* DNA ligase** cannot join blunt ends at all.

- In practice genetic engineers normally use **T4 DNA ligase**. Originally this came from the T4 bacterial virus and is both easier to use in the test tube and can join of blunt ends if need be.

Where Do Restriction Enzymes Come From?

- Several hundred restriction enzymes different recognition sites are now available commercially. When bacteria are attacked by viruses, the virus coat is left outside and only the virus enters the target cell.

- The virus DNA will take over the victim's cellular machinery and use it to manufacture more virus particles unless bacterial cell fights back.

- The bacteria usually fight by chopping up the DNA of the virus thereby destroying it.

- This "restricts" the entry of the virus and so the enzymes that chopped up the DNA were called **restriction enzymes**.

- In practice, viruses get wasted by the victims they attack. Only a few survive and succeed in conquering the bacteria.

Where Do Specific Restriction Enzymes Get Their Funny Names?

- Restriction enzymes have names made up of the initials of the restriction bacteria they come from.

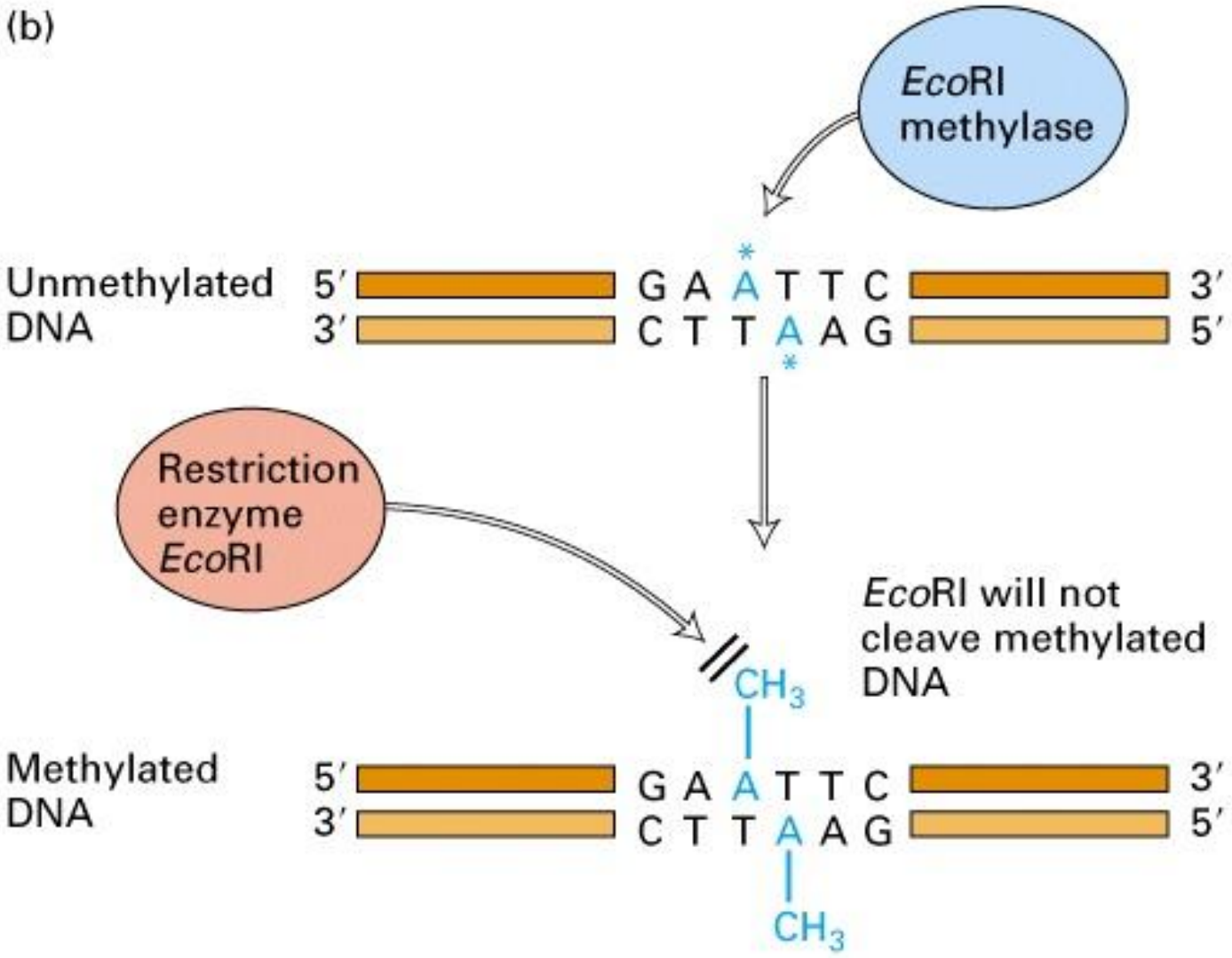
- Let's take ***EcoRI***. *Eco*R means that the enzyme was found in ***E****scherichia* ***c****oli* strain **R**Y13 and the "**I**" ("one" not 'eye' means it was the first restriction enzyme found in this strain).

Protection of the Cell's Own DNA

- Why don't restriction enzymes destroy the DNA of their own cell? In practice, whenever a bacterial cell makes a restriction enzyme it also makes a "kinder and gentler" protein known as a **modification enzymes**.

- This chemically modifies the DNA of the bacterial recognition site.

- Modification enzymes bind to the DNA at the same recognition site as the corresponding restriction enzymes. They then **add a molecular tag** to one of the bases in the recognition sequence.


(b)
EcoRI methylase
Unmethylated DNA
5′ G A A T T C 3′
3′ C T T A A G 5′
Restriction enzyme EcoRI
EcoRI will not cleave methylated DNA
CH3
Methylated DNA
5′ G A A T T C 3′
3′ C T T A A G 5′
CH3

- Once the cut site has been tagged or "modified", the restriction enzymes can no longer cut it. The result is that the DNA in a bacterial cell is **immune to that cell's own restriction enzymes**.

- Only foreign DNA that tries to invade the cell from outside is attacked. Clever, huh?

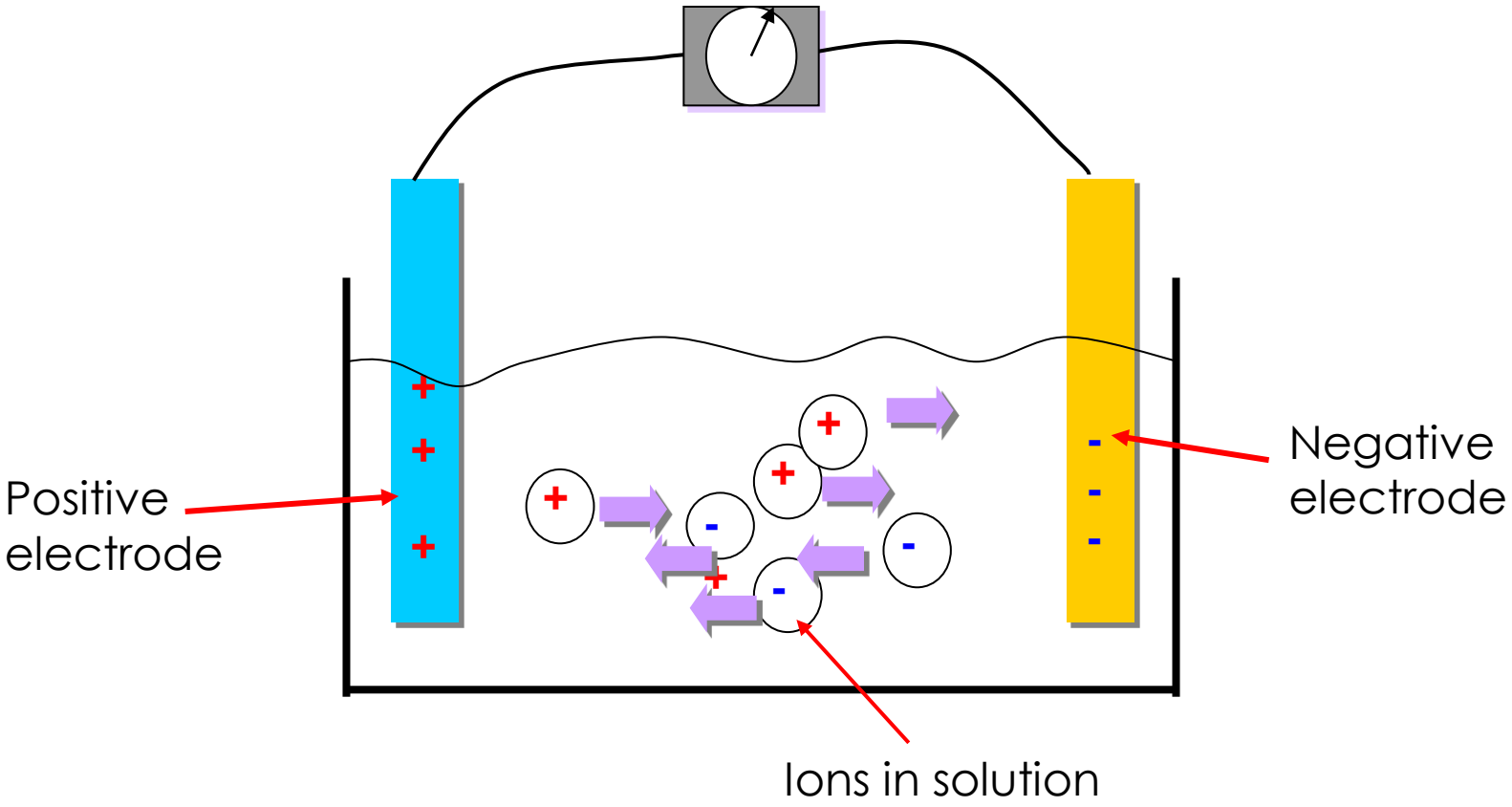
How Can Individual Fragments be Separated out?

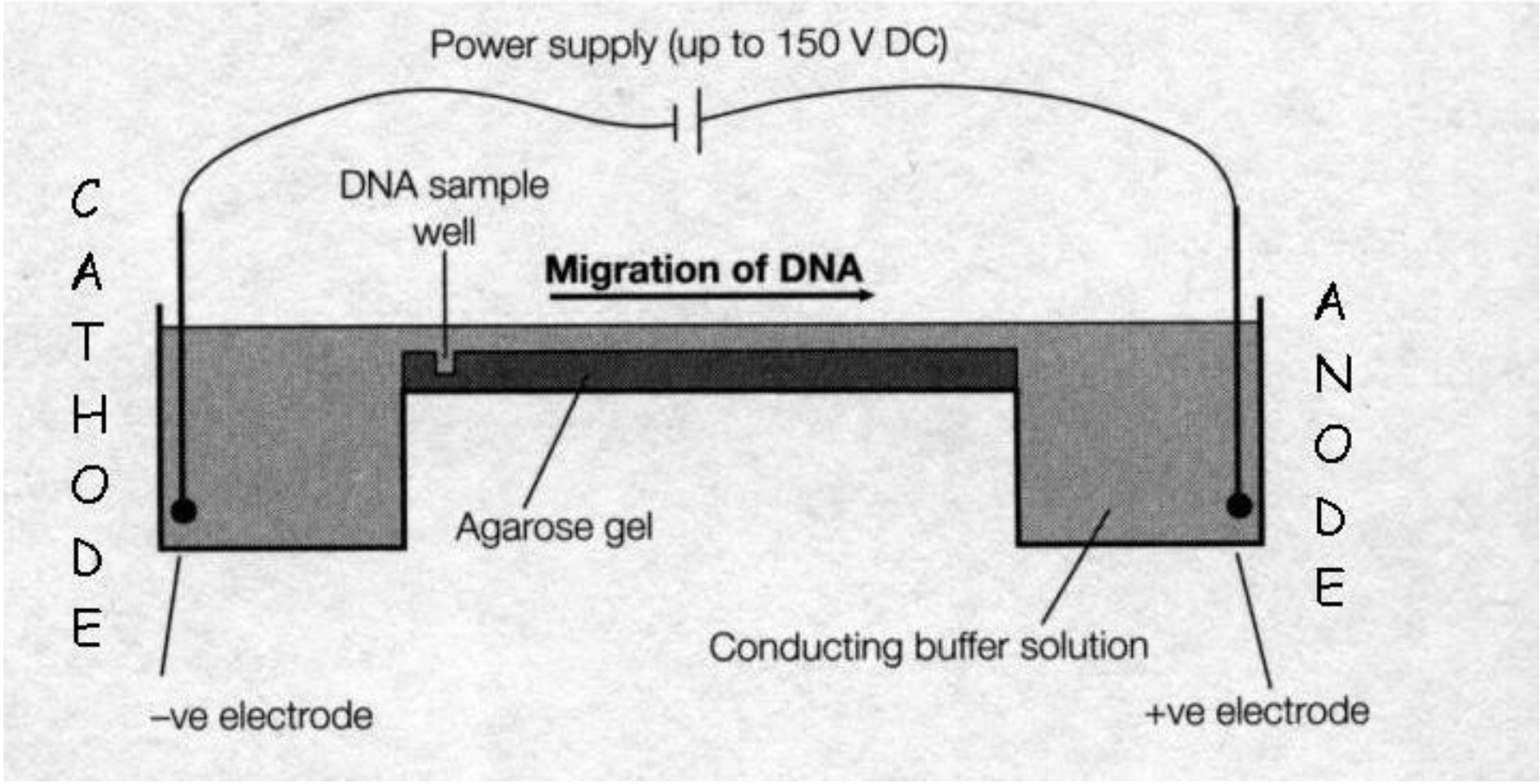
- After cutting up a long piece of DNA we will probably need to separate the pieces from one another. This is usually done by **gel electrophoresis**.
- This technique will separate molecules only if they are electrically charged. Two electrodes, one positive and the other negative are connected up to a high voltage source.
- Positively charged molecules move towards negative electrode and negatively charged molecules move towards positive electrode.

- Since **DNA carries a negative charge** on each of the many phosphate groups making up its backbone, it will **move towards the positive electrode** during electrophoresis.

- The bigger a molecule the more force required to move it. However, the bigger a DNA molecule, the more negative charges it has!

- So these two factors cancel out because all our fragments of DNA have the same number of charges per unit length. And so they all cruise along at the same speed toward the positive electrode.


Positive electrode
Negative electrode
Ions in solution


Power supply (up to 150 V DC)
DNA sample well
Migration of DNA
CATHODE
ANODE
Agarose gel
Conducting buffer solution
–ve electrode
+ve electrode

- What we want is a way to separate the big fragments from the small ones. To do this, we run them through a **gel**.

- This is a high-tech version of the familiar gelatin dessert without the special flavour and colour!

- Gelatin sets due to a microscopic mesh formed by its own protein fibres.

- Gels for DNA work are made of **agarose**. This is a polysacharide extracted from seaweed. When a hot solution of agarose cools, it congeals to form a meshwork rather like gelatin.

- The sample of DNA is put into a hole made at one end of a slab of agarose and electrodes are connected up in contact with each end of the slab.

- The power is turned on and the DNA starts to move towards the positive electrode. As the DNA molecule move through the gel, they are hindered by **the meshwork** of the agarose fibre.

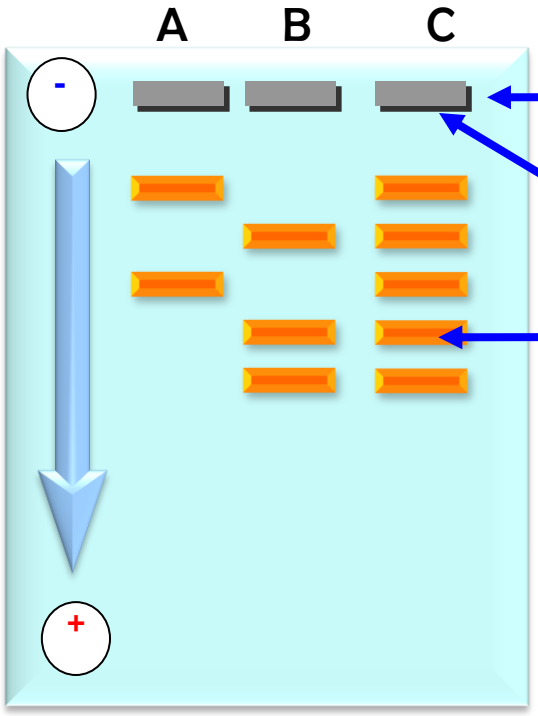
- The **larger molecules** find it more **difficult to squeeze through the gap** but the **smaller ones** could push through faster. The result is that DNA fragment separate **in order of size**.

- Both DNA and agarose are colourless so we cannot see when the DNA has ended up. To locate DNA fragments, a type of stain **ethidium bromide** (EtBr) which binds specifically to RNA is used. EtBr stains DNA orange if **viewed under ultraviolet**.

- After the orange bands are located, they are **cut out of the gel** and the DNA is extracted to yield a pure fragment.

- Agarose gel electrophoresis can be used to **purify DNA** for use in genetic engineering or it can be used to **measure the size** of an unknown piece of DNA, we run a set of standard DNA fragments of known sizes alongside, on the same gel.

Movement of DNA



Sample wells

Sample C is a mix of A plus B

Fluorescent orange bands

Choosing a Vector for our Own Special Purposes

- Now that we have some pieces of DNA containing genes we are interested in, we may need to move them into other cells or manufacture them in bulk quantities for further genetic engineering.

- For either purpose we mount our cloned genes on what are called **vectors**.

In principle any molecule of DNA which can replicate itself inside a cell could work as a vector. To make life as easy as possible we considered the following factors:

1. The vector should be a reasonably small and manageable DNA molecule

2. Moving the vector from cell to cell should be easy

3. Growing and purifying large amounts of vector DNA should be straightforward

- In practice, bacterial **plasmids** that we learned about in the last chapter come closest to these requirements and are the most widely used vectors.

- We will discuss the requirements for vectors using plasmids as examples.

Multicopy Plasmid Vectors

We have discussed plasmids in the former chapters, so let's pick one and get going.

- The **ColE1 plasmid** of *Escherichia coli* is a small circular DNA.
- Molecule that forms the basis of many vectors widely used in molecular biology.
- It exists in up to 40 copies per cell so obtaining plenty of plasmids.
- DNA is easy and it can be moved from cell to cell by **transformation** as described in before.

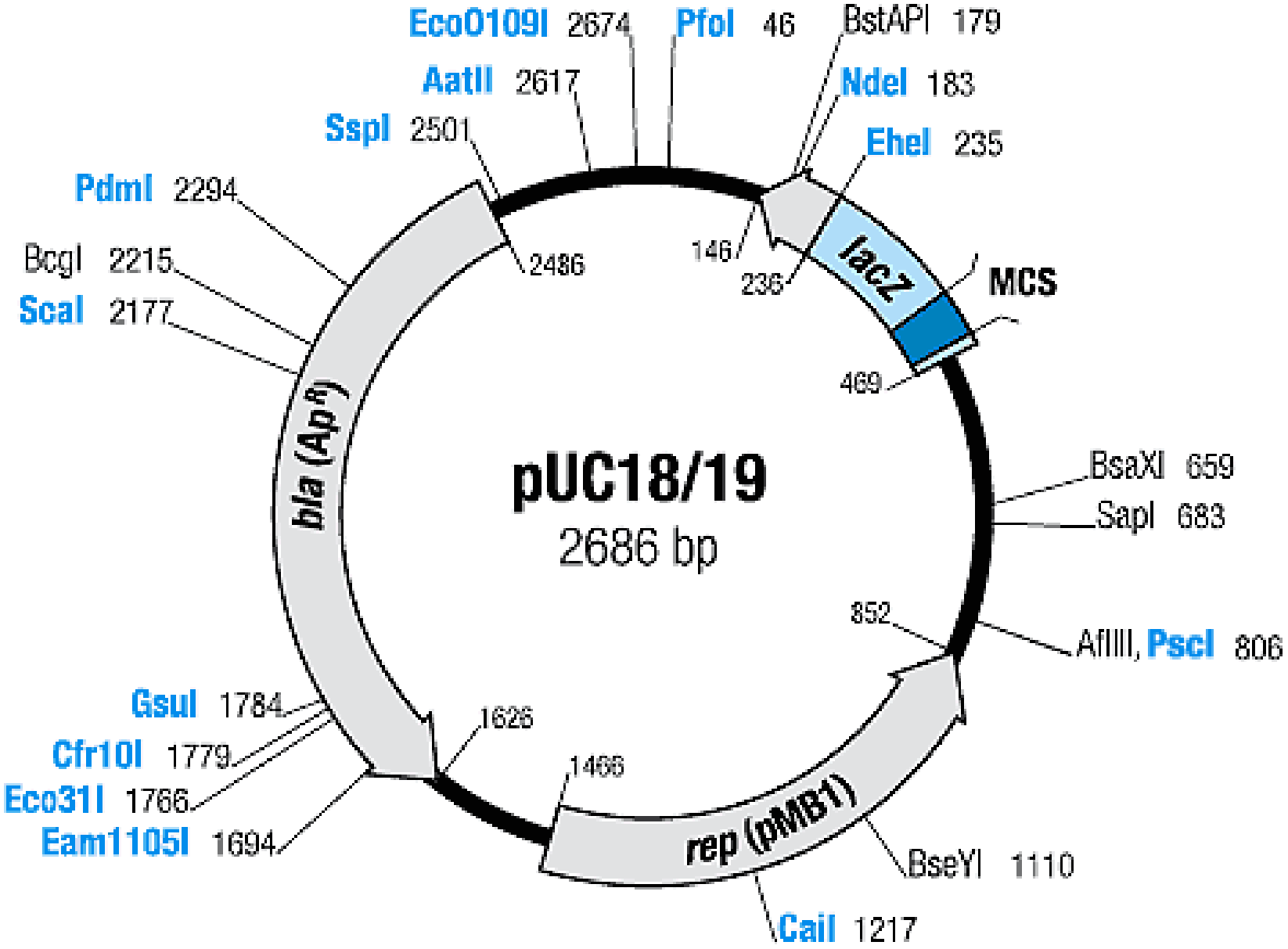
Since our basic requirements are satisfied, let's be greedy. Wouldn't be nice if there was a simple way to:

1. Detect the presence of the vector

2. Directly select for cells that contain the vector (**transformants**)

3. Insert genes into the vector

4. Detect the presence of insert DNA on the vector (**recombinants**)


EcoO109I 2674
PfoI 46
BstAPI 179
AatII 2617
NdeI 183
SspI 2501
EheI 235
PdmI 2294
BcgI 2215
ScaI 2177
2486
146
236
lacZ
MCS
469
bla (Ap^R)
pUC18/19
2686 bp
BsaXI 659
SapI 683
852
AflIII, PscI 806
GsuI 1784
Cfr10I 1779
Eco31I 1766
Eam1105I 1694
1626
1466
rep (pMB1)
BseYI 1110
CaiI 1217

Detecting and Selecting Vectors

- To satisfy the cravings of the newer generation of pampered and spoiled molecular biologists, it has been necessary to improve the original ColE1 plasmid.

- This was done in several stages. First, we get rid of the genes for colicin E1, a toxic protein for killing bacteria as these are obviously not needed. We next add a gene for resistance to the antibiotic ampicillin, a widely used penicillin derivative.

- The ***ampR*** gene is also known as ***bla*** which refers to β-lactamase, the enzyme encoded by this gene, which degrades penicillins and its derivatives.

- When we transform this new vector into bacterial cells, we can directly select those which get the plasmid by incubating in a growth medium containing **ampicillin**. Those cells containing a plasmid survive, while those which did not get a plasmid are killed.

- Any time we wish to check that our vector is still present we merely check to see that the cells are **still ampicillin resistant**. Thus, the antibiotic resistance gene satisfies both requirements #1 and #2 of the preceeding section.

Inserting Genes into Vectors

- So how about inserting genes into the vector? This is done by cutting out a segment of DNA carrying our gene with a suitable restriction enzyme as described above.

- We must then cut the plasmid open with the same restriction enzyme so that the **vector and the insert have matching ends**.

- When we mix the two together and **add DNA ligase**, the enzyme which links together DNA strands, we will end up with our gene inserted into our vector.

- But wait a moment! What if there were more than one cut site in the vector for the restriction enzyme we chose?

- Then our vector would be chopped into pieces, not merely opened conveniently. This would be bad!

- Again, we must avoid inserting the cloned gene into any of the genes needed by the plasmid for its own replication and survival within the cell.

- And, while we are about it, we need flexibility; there are many different restriction enzymes and it would be nice to have a wide choice.

- So let's just fix all these questions at one shot. We put into our plasmid a stretch of artificially synthesized DNA about 50 base pairs long which contains cut sites for seven or eight of our favorite restriction enzymes.

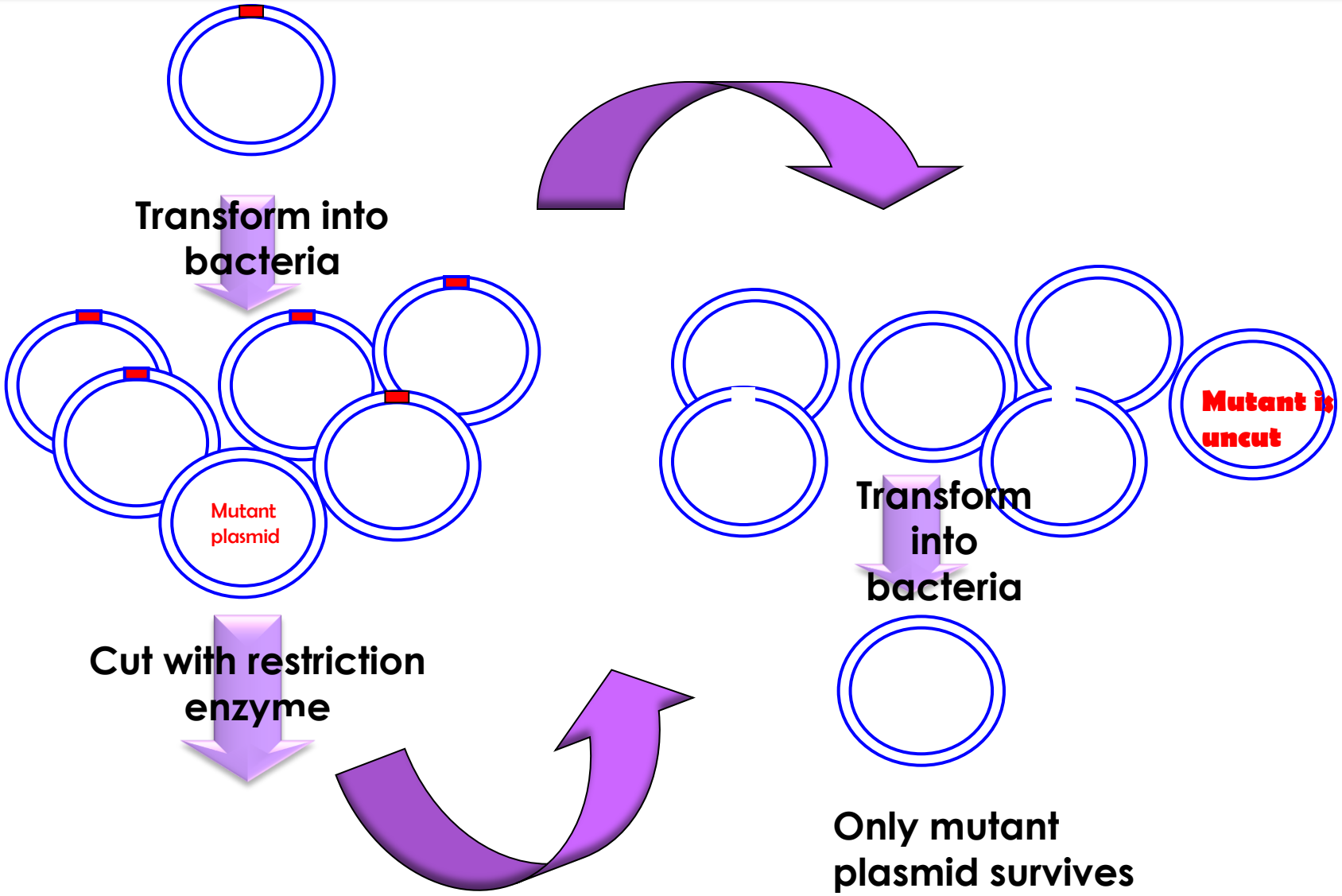
- This is known as **multiple cloning site** or, more often, just as **polylinker**, and not only allows us a wide choice of restriction enzymes, but ensures that the insert does not damage the plasmid and goes into more or less the same location each time.

- One other thing to do is to make sure that there are **no extra cut sites**, on the plasmid for any of the enzymes represented in the polylinker.

- So the easy way to do this is to choose for the polylinker only, be enzymes with zero cut sites in the plasmid we started with.

- Alternatively, we can get rid of unwanted cut sites by the approach shown below. First, we grow cells containing the original plasmid and we make some plasmid DNA.

Removal of unwanted cut sites



- Due to spontaneous mutation, an occasional plasmid will suffer a base change within the cut site about which we are concerned.

- This will abolish recognition of the site by the restriction enzyme. But how do we find this one rare mutant plasmid? Simple, once someone else has thought of it!

- We treat the plasmid sample with the restriction enzyme in question and then transform the plasmid DNA into fresh bacterial cells without religating the break.

- Most plasmids will be destroyed by this procedure, but those few that have lost the cut site by mutation will remain uncut and survive.

Detecting Insertions into Vectors

- Let's suppose we have inserted our cloned gene into a plasmid vector and transformed it back into a bacterial cell. How do we know if our cloned gene is actually there?

- If the cloned gene itself makes something easy to detect, there is no problem. But what if there is no easy test and we have to detect directly whether our cloned piece of DNA was successfully inserted in the vector?

There are basically three approaches:

A. antibiotics,

B. bruteforce and ignorance, and

C. colour.

Let's start with the second approach : B.

- We pick a large number of separate bacterial colonies that have received the plasmid vector, hopefully with DNA inserted.

- We extract plasmid DNA from each of these. We then cut the plasmid DNA with the restriction enzyme used during our original cloning experiment.

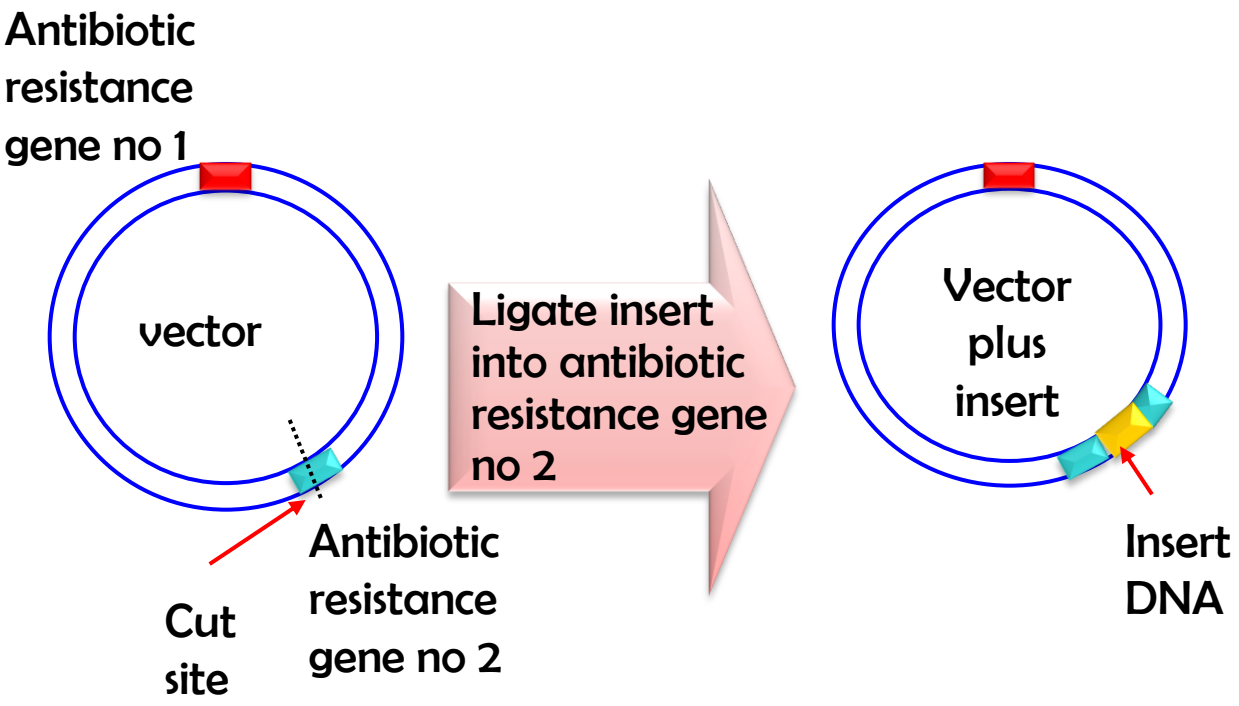
- If there is no insert in the plasmid, this merely converts the plasmid from a circular to a linear molecule of DNA. If the vector **contains inserted DNA**, we will **get two pieces of DNA**, one being the original plasmid and the other the inserted DNA fragment.

- To see how many pieces of DNA we actually have, we separate the **cut DNA** by **gel electrophoresis**. If we test enough transformed colonies, sooner or later, perhaps very much later! -we will find one with the DNA fragment inserted into its plasmid.

- Somewhat less laborious is to use an **antibiotic resistance gene**. Actually, we need two antibiotic resistance genes, one to select for cells which have received a plasmid as described above and a second to fool around with to detect inserts.

- Suppose that the cut site for the restriction enzyme we use is within this second antibiotic resistance gene.

- Then, when we insert our cloned fragment of DNA, we will disrupt this antibiotic resistance gene. So, cells that receive a plasmid without an insert will be resistant to both antibiotics. Those receiving a plasmid with an insert will be resistant to only the first antibiotic. This technique is referred to as **insertional inactivation**.



Cells carrying vector are resistant to both antibiotics

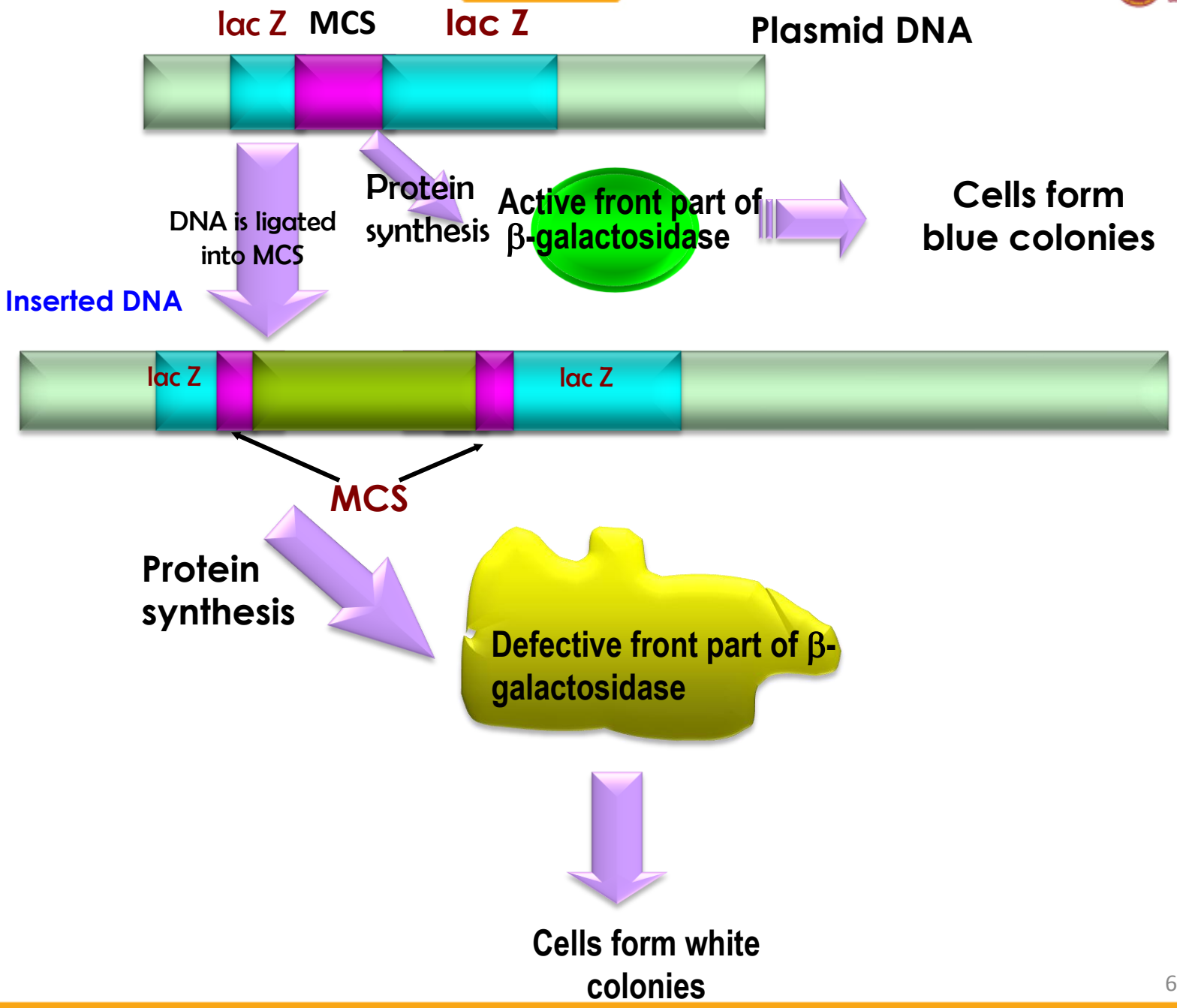
Cells carrying vector and insert are resistant to antibiotic no 1 only

- Finally, for the truly sophisticated, there is colour screening. The best known version is the **blue/whlte screening** done using β-galactosidase and **X-gal**.

- We start with a plasmid carrying the ***lacZ*** gene for **β-galactosidase** and insert a polylinker into the *lacZ* coding sequence, very close to the front of the gene. Luckily, the front most part of the β-galactosidase protein is unusual in being inessential for enzyme activity.

- As long as the polylinker is inserted without disrupting the reading frame we obtain active enzyme. However, if we **insert a foreign segment** of DNA into the polylinker, the *lacZ* gene is **inactivated** and no β–galactosidase will be made.

- This is also termed as **insertional inactivation** as the *lac Z* gene is inactivated due to insertion of foreign DNA.


lac Z
MCS
lac Z
Plasmid DNA
DNA is ligated into MCS
Protein synthesis
Active front part of β-galactosidase
Cells form blue colonies
Inserted DNA
lac Z
lac Z
MCS
Protein synthesis
Defective front part of β-galactosidase
Cells form white colonies

- Cells that contain plasmids are known as **transformants** (whether they have inserts or not has yet to be determined)

- We can detect β-galactosidase because it turns X-gal **blue**. Plasmids without a DNA insert will produce b-galactosidase and the cells carrying them will turn blue. Plasmids with an insert cannot make β-galactosidase and the cells will stay **white**.

Cells that contain plasmids plus insert DNA are known as **recombinants** (confirmed to have inserts)

Cells that contain plasmids are known as **transformants** (whether they have inserts or not has yet to be determined)

References:

- Madigan, M.T., Martinko, J.M., Dunlap, P.V. and Clark, D.P. (2009). Brock Biology of Microorganisms: Pearson Education, USA.

- Clark, D.P. and Russel, L.D. (2000). Molecular Biology Made Simple and Fun: Cache River Press, USA.