KARPAGAM ACADEMY OF HIGHER EDUCATION

(Deemed to be University)
(Established Under Section 3 of UGC Act 1956)
Coimbatore - 641021.

(For the candidates admitted from 2017 onwards)

DEPARTMENT OF BIOCHEMISTRY

STAFF NAME	**:**	**Dr. L. HARIPRASATH**		
SUBJECT	**:**	**GENETIC ENGINEERING AND BIOTECHNOLOGY**		
SUBJECT CODE	**:**	**17BCU601A**		
SEMESTER	**:**	**IV**	**CLASS**	**: III B.Sc. Biochemistry**

Course objective

- This course intended to acquaint the students with simple molecular biology techniques.
- To teach isolation and manipulation of genetic material for achieving the desired goal.

Course outcomes (CO's)

1. The students acquire basic knowledge of engineering and skills to design and conduct experiments, analyze data and interpret the results.
2. The students acquire the ability to demonstrate understanding of modern engineering techniques used in biotechnology with reinforcement.

Unit I: Introduction to recombinant DNA technology

Overview of recombinant DNA technology. Restriction and modification systems, restriction endonucleases and other enzymes used in manipulating DNA molecules, separation of DNA by gel electrophoresis. Extraction and purification of plasmid and bacteriophage DNA.

Joining of DNA fragments : Ligation of DNA molecules. DNA ligase, sticky ends, blunt ends, linkers and adapters. Synthetic oligonucleotides, synthesis and use.

Unit II: Cloning vectors for prokaryotes and eukaryotes

Plasmids and bacteriophages as vectors for gene cloning. Cloning vectors based on *E. coli* plasmids, pBR322, pUC8, pGEM3Z. Cloning vectors based on M13 and λ bacteriophage. Vectors for yeast, higher plants and animals.

DNA sequencing : DNA sequencing by Sanger's method, modifications based on Sanger's method. Automated DNA sequencing. Pyrosequencing.

Unit III: Introduction of DNA into cells and selection for recombinants

Uptake of DNA by cells, preparation of competent cells. Selection for transformed cells. Identification for recombinants - insertional inactivation, blue-white selection. Introduction of phage DNA into bacterial cells. Identification of recombinant phages. Introduction of DNA into animal cells, electroporation.

Methods for clone identification

The problem of selection, direct selection, marker rescue. Gene libraries, identification of a clone from gene library, colony and plaque hybridization probing, methods based on detection of the

translation product of the cloned gene.

Unit IV: Polymerase chain reaction
Fundamentals of polymerase chain reaction, designing primers for PCR. Studying PCR products. Cloning PCR products. Real time PCR.

Expression of cloned genes : Vectors for expression of foreign genes in *E. coli*, cassettes and gene fusions. Challenges in producing recombinant protein in *E. coli*. Production of recombinant protein by eukaryotic cells. Fusion tags and their role in purification of recombinant proteins.

Unit V: Applications of genetic engineering in Biotechnology
Site–directed mutagenesis and protein engineering. Applications in medicine, production of recombinant pharmaceuticals such as insulin, human growth hormone, factor VIII.
Recombinant vaccines. Gene therapy. Applications in agriculture - plant genetic engineering, herbicide resistant crops, problems with genetically modified plants, safety concerns.

SUGGESTED READING

1. Brown, T.A., (2010). Gene Cloning and DNA Analysis 6th ed., Wiley-Blackwell publishing (Oxford, UK), ISBN: 978-1-4051-8173-0.

2. Primrose, S.B., and Twyman, R. M., (2006). Principles of Gene Manipulation and Genomics 7th ed., Blackwell publishing (Oxford, UK) ISBN:13: 978-1-4051-3544-3.

3. Glick, B.R., Pasternak, J.J. and Patten, C.L., (2010). Molecular Biotechnology: Principles and Applications of Recombinant DNA 4th ed., ASM Press (Washington DC), ISBN: 978-1-55581-498-4 (HC).

KARPAGAM ACADEMY OF HIGHER EDUCATION
(Deemed to be University Established Under Section 3 of UGC Act 1956)
Coimbatore – 641 021.

LECTURE PLAN
DEPARTMENT OF BIOCHEMISTRY

STAFF NAME: **Dr. L. Hariprasath**

SUBJECT NAME: **GENETIC ENGINEERING AND BIOTECHNOLOGY** SUB.CODE: **17BCU601A**

SEMESTER: **VI** CLASS: **III B.Sc Biochemistry**

Sl. No	Duration of Period	Topics to be Covered	Books Referred Page No	Web Page referred
		UNIT – I		
1.	1	**Introduction to recombinant DNA technology** : Overview of recombinant DNA technology	R1: 3-4	
2.	1	Restriction and modification systems, restriction endonucleases and other enzymes used in manipulating DNA molecules	R1: 3-4	W1
3.	1	Separation of DNA by gel electrophoresis. Extraction and purification of plasmid and bacteriophage DNA	R1: 16-18, 55-57	
4.	1	**Joining of DNA fragments :**Ligation of DNA molecules. DNA ligase, sticky ends, blunt ends, linkers and adapters	T1: 58-60	
5.	1	Synthetic oligonucleotides, synthesis and use.	T1: 58-60	
6.	1	Revision		
Total number of hours planned for Unit I: 6				
		UNIT – II		
1.	1	**Cloning vectors for prokaryotes and eukaryotes** : Plasmids and bacteriophages as vectors for gene cloning.	R1: 55-57	
2.	1	Cloning vectors based on *E. coli* plasmids, pBR322, pUC8, pGEM3Z.	T1: 66-78	
3.	1	Cloning vectors based on M13 and λ bacteriophage.	T2: 94-102	

4.	1	Vectors for yeast, higher plants and animals.	T2: 109-111	
5.	1	**DNA sequencing :** DNA sequencing by Sanger's method, modifications based on Sanger's method.	T2: 141-145	
6.	1	Automated DNA sequencing, Pyrosequencing.	T2: 141-145	
Total number of hours planned for Unit II: 6				
UNIT – III				
1.	1	**Introduction of DNA into cells and selection for recombinants** : Uptake of DNA by cells, preparation of competent cells. Selection for transformed cells.	T1: 84-89	
2.	1	Identification for recombinants - insertional inactivation, blue-white selection. Introduction of phage DNA into bacterial cells.	R1:101-104 T3:86-89	
3.	1	Identification of recombinant phages. Introduction of DNA into animal cells, electroporation.	T3:86-89	
4.	1	**Methods for clone identification** The problem of selection, direct selection, marker rescue.	T1: 32-38	
5.	1	Gene libraries, identification of a clone from gene library, colony and plaque hybridization probing, methods based on detection of the translation product of the cloned gene.	T1: 37-38	
Total number of hours planned for Unit III: 5				
UNIT – IV				
1.	1	**Polymerase chain reaction:** Fundamentals of polymerase chain reaction, designing primers for PCR. Studying PCR products.	T1: 116-124	
2.	1	Cloning PCR products. Real time PCR.	T1: 124-127	
3.	1	**Expression of cloned genes :** Vectors for expression of foreign genes in *E. coli*, cassettes and gene fusions.	T2: 45-46	
4.	1	Challenges in producing recombinant protein in *E. coli*. Production of recombinant protein by eukaryotic cells.	R1: 189-194	W2
5.	1	Fusion tags and their role in purification of recombinant proteins.	R1: 212-213	
6.	1	Revision		

Total number of hours planned for Unit IV: 6				
UNIT V				
1.	1	**Applications of genetic engineering in Biotechnology:** Site–directed mutagenesis and protein engineering.	R1:141-144	
2.	1	Applications in medicine, production of recombinant pharmaceuticals such as insulin, human growth hormone, factor VIII.	T1:220-225	
3.	1	Recombinant vaccines. Gene therapy	T3:193-196	
4.	1	Applications in agriculture - plant genetic engineering, herbicide resistant crops, problems with genetically modified plants, safety concerns.	T1: 256-262	
5.	1	Revision		
Total number of hours planned for Unit V: 5				
Previous year ESE Question Paper Discussion				
1.	1	Previous year ESE question paper discussion		
2.	1	Previous year ESE question paper discussion		
Total Hours Planned: 30				

Reference books

R1: Principles of Gene Manipulation and Genomics (2006) 7th ed., Primrose, S.B., and Twyman, R. M., Blackwell publishing (Oxford, UK) ISBN:13: 978-1-4051-3544-3.

T1. Desmond S. T. Nicholl, (2008) An Introduction to Genetic Engineering Third Edition, Cambridge University Press

T2. Dubey, R.C. (2009). *A Textbook of Biotechnology* (19th ed.). New Delhi, S. Chand and Compamy Ltd.

T3. Satyanarayana, U. (2005). *Biotechnology*. (th ed.) Kolkata, Books and Allied (P). Ltd

W1: https://www.ndsu.edu/pubweb/~mcclean/plsc731/dna/dna5.html

W2: http://www.biologydiscussion.com/gene/gene-expression/manipulation-of-gene-expression-in-host-cells-with-diagram/9913

UNIT-I
SYLLABUS

Introduction to recombinant DNA technology : Overview of recombinant DNA technology. Restriction and modification systems, restriction endonucleases and other enzymes used in manipulating DNA molecules, separation of DNA by gel electrophoresis. Extraction and purification of plasmid and bacteriophage DNA.

Joining of DNA fragments :Ligation of DNA molecules. DNA ligase, sticky ends, blunt ends, linkers and adapters. Synthetic oligonucleotides, synthesis and use.

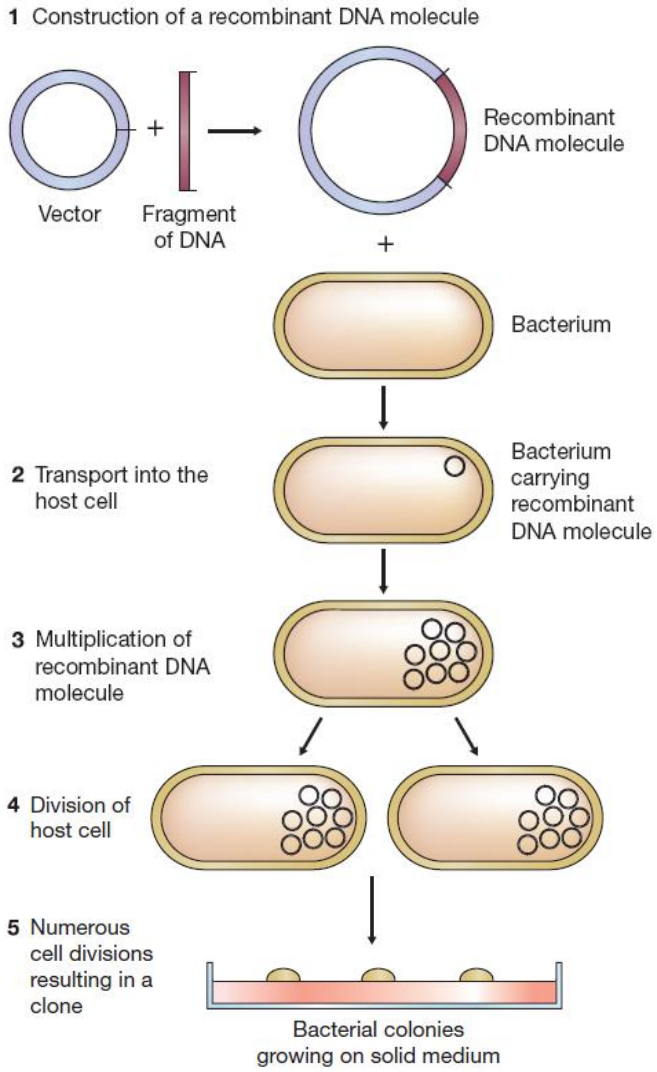
Introduction to recombinant DNA technology :

Overview of recombinant DNA technology.

The goal of rDNA is gene cloning to generate large amounts of pure DNA that can be manipulated and studied. The following are the basic steps involved in the process of the rDNA technique for gene cloning:

The steps in a gene cloning experiment:

- A fragment of DNA, containing the gene to be cloned, is inserted into a circular DNA molecule called a vector, to produce a recombinant DNA molecule.
- The vector transports the gene into a host cell, which is usually a bacterium, although other types of living cell can be used.
- Within the host cell the vector multiplies, producing numerous identical copies, not only of itself but also of the gene that it carries.
- When the host cell divides, copies of the recombinant DNA molecule are passed to the progeny and further vector replication takes place.
- After a large number of cell divisions, a colony, or clone, of identical host cells is produced. Each cell in the clone contains one or more copies of the recombinant DNA molecule; the gene carried by the recombinant molecule is now said to be cloned.



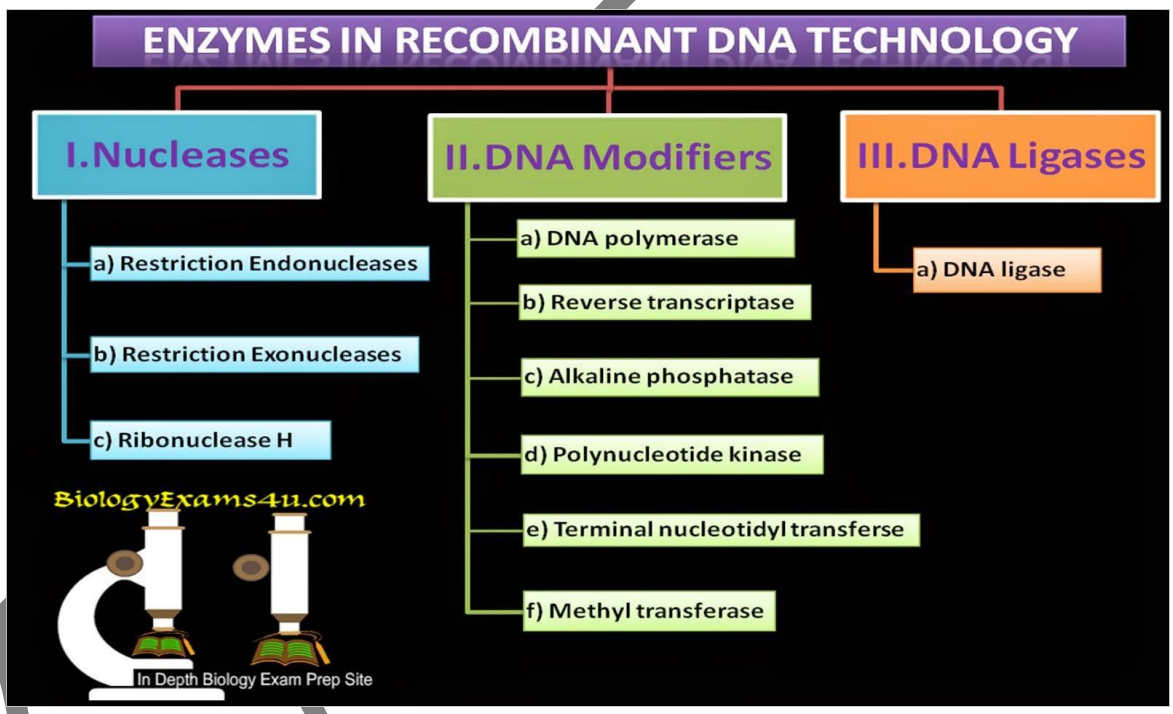
Restriction Modification System

Phage (or viruses) invade all types of cells. Bacteria are one favorite target. Defense mechanisms have been developed by bacteria to defend themselves from these invasions. The system they possess for this defense is the restriction-modificiation system. This system is composed of a restriction endonuclease enzyme and a methylase enzyme and each bacterial species and strain has their own combination of restriction and methylating enzymes

Restriction endonucleases and other enzymes used in manipulating DNA molecules

DNA manipulative enzymes can be grouped into four broad classes, depending on the type of reaction that they catalyze:

- **Nucleases** are enzymes that cut, shorten, or degrade nucleic acid molecules.
- **Ligases** join nucleic acid molecules together.
- **Polymerases** make copies of molecules.
- **Modifying enzymes** remove or add chemical groups.



1. Nucleases

Nucleases degrade DNA molecules by breaking the phosphodiester bonds that link one nucleotide to the next in a DNA strand. There are two different kinds of nuclease (Figure):

- **Exonucleases** remove nucleotides one at a time from the end of a DNA molecule.
- **Endonucleases** are able to break internal phosphodiester bonds within a DNA molecule.

The main distinction between different exonucleases lies in the number of strands that are degraded when a double-stranded molecule is attacked. The enzyme called **Bal31** (purified from the bacterium *Alteromonas espejiana*) is an example of an exonuclease that removes nucleotides from both strands of a double-stranded molecule (Figure). In contrast, enzymes such as *E. coli* exonuclease III degrade just one strand of a double-stranded molecule, leaving single-stranded DNA as the product (Figure).

The same criterion can be used to classify endonucleases. S1 endonuclease (from the fungus *Aspergillus oryzae*) only cleaves single strands (Figure), whereas deoxyribonuclease I (DNase I), which is prepared from cow pancreas, cuts both single and double-stranded molecules (Figure).

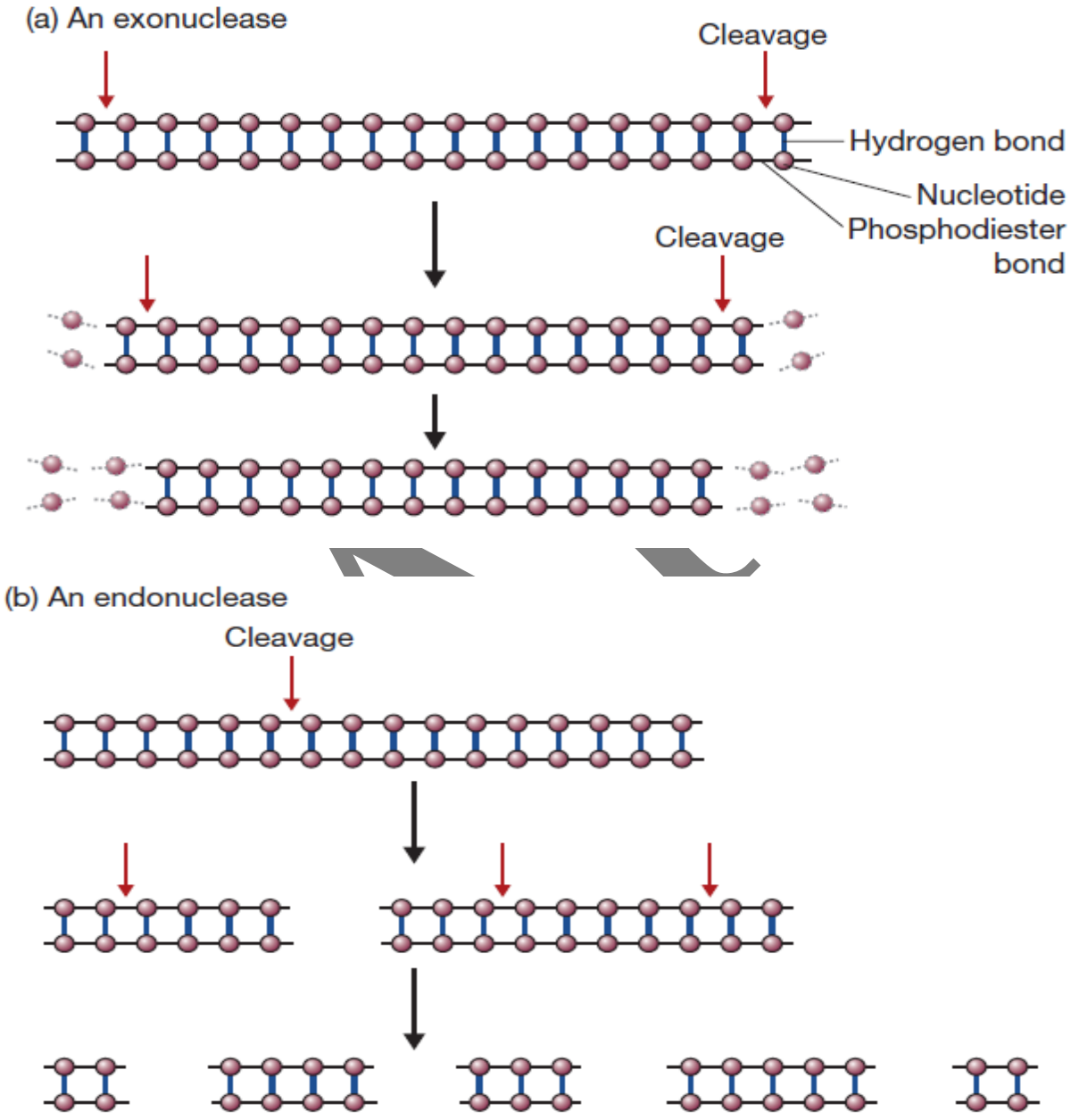


Fig: The reactions catalyzed by the two different kinds of nuclease. (a) An exonuclease, which removes nucleotides from the end of a DNA molecule. (b) An endonuclease, which breaks internal phosphodiester bonds.

2.Ligases

In the cell the function of DNA ligase is to repair single-stranded breaks ("discontinuities") that arise in double-stranded DNA molecules during, for example, DNA replication. DNA ligases from most organisms can also join together two individual fragments of double-stranded DNA

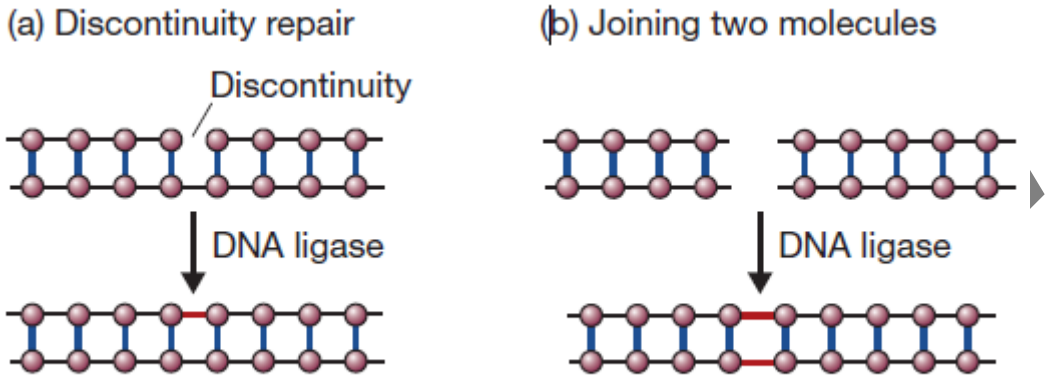


Fig : The two reactions catalyzed by DNA ligase. (a) Repair of a discontinuity—a missing phosphodiester bond in one strand of a doublestranded molecule. (b) Joining two molecules together.

3.Polymerases

DNA polymerases are enzymes that synthesize a new strand of DNA complementary to an existing DNA or RNA template (Figure 4.5a). Most polymerases can function only if the **template** possesses a double-stranded region that acts as a **primer** for initiation of polymerization.

Four types of DNA polymerase are used routinely in genetic engineering.

The first is DNA polymerase I, which is usually prepared from *E. coli*. This enzyme attaches to a short single-stranded region (or **nick**) in a mainly double-stranded DNA molecule, and then synthesizes a completely new strand, degrading the existing strand as it proceeds (Figure 4.5b). DNA polymerase I is therefore an example of an enzyme with a dual activity—DNA polymerization and DNA degradation. The polymerase and nuclease activities of DNA polymerase I are controlled by different parts of the enzyme molecule. The nuclease activity is contained in the first 323 amino acids of the polypeptide, so removal of this segment leaves a modified enzyme that retains the polymerase function but is unable to degrade DNA. This modified enzyme, called the **Klenow fragment**, can still synthesize a complementary DNA strand on a single-stranded template, but as it has no nuclease activity it cannot continue the synthesis once the nick is filled in (Figure).

Several other enzymes—The *Taq* DNA polymerase used in the polymerase chain reaction (PCR) (see Figure) is the DNA polymerase I enzyme of the bacterium *Thermus aquaticus*. This organism lives in hot springs, and many of its enzymes, including the *Taq* DNA polymerase, are thermostable, meaning that they are resistant to denaturation by heat treatment. This is the special feature of *Taq* DNA polymerase that makes it suitable for PCR, because if
it was not thermostable it would be inactivated when the temperature of the reaction israised to 94°C to denature the DNA.

The final type of DNA polymerase that is important in genetic engineering is **reverse transcriptase**, an enzyme involved in the replication of several kinds of virus. Reverse transcriptase is unique in that it uses as a template not DNA but RNA (Figure). The ability of this enzyme to synthesize a DNA strand complementary to an RNA template is central to the technique called complementary DNA (cDNA) cloning.

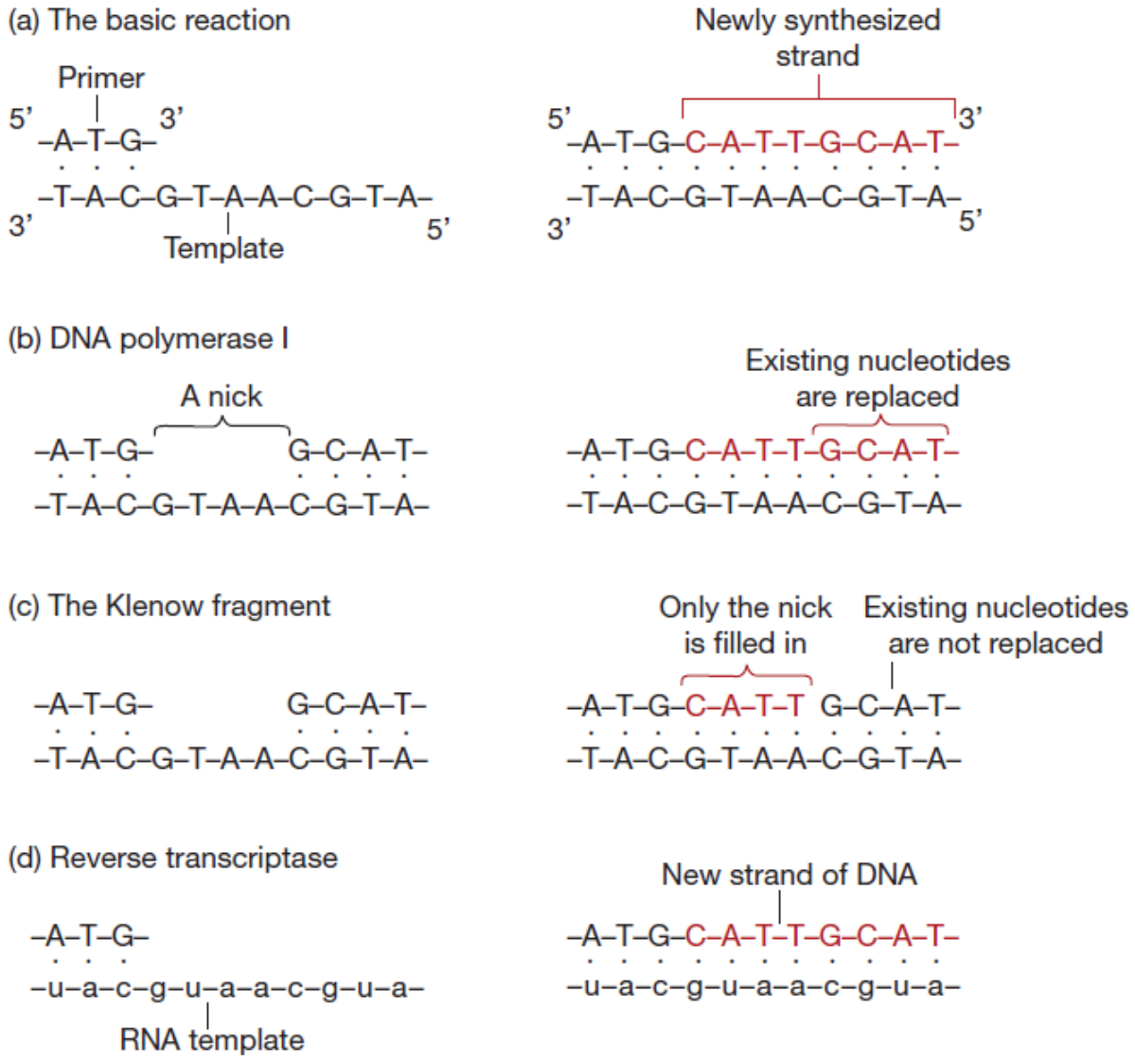


Fig: The reactions catalyzed by DNA polymerases. (a) The basic reaction: a new DNA strand is synthesized in the 5′ to 3′direction. (b) DNA polymerase I, which initially fills in nicks but then continues to synthesize a new strand, degrading the existing one as it proceeds. (c) The Klenow fragment, which only fills in nicks. (d) Reverse transcriptase, which uses a template of RNA.

4. DNA modifying enzymes

There are numerous enzymes that modify DNA molecules by addition or removal of specific chemical groups. The most important are as follows:

l **Alkaline phosphatase** (from *E. coli*, calf intestinal tissue, or arctic shrimp), which removes the phosphate group present at the **5′ terminus** of a DNA molecule(Figure a).

l **Polynucleotide kinase** (from *E. coli* infected with T4 phage), which has the reverseeffect to alkaline phosphatase, adding phosphate groups onto free 5′termini(Figure b).
l **Terminal deoxynucleotidyl transferase** (from calf thymus tissue), which adds one or more deoxyribonucleotides onto the **3′□terminus** of a DNA molecule (Figure c)

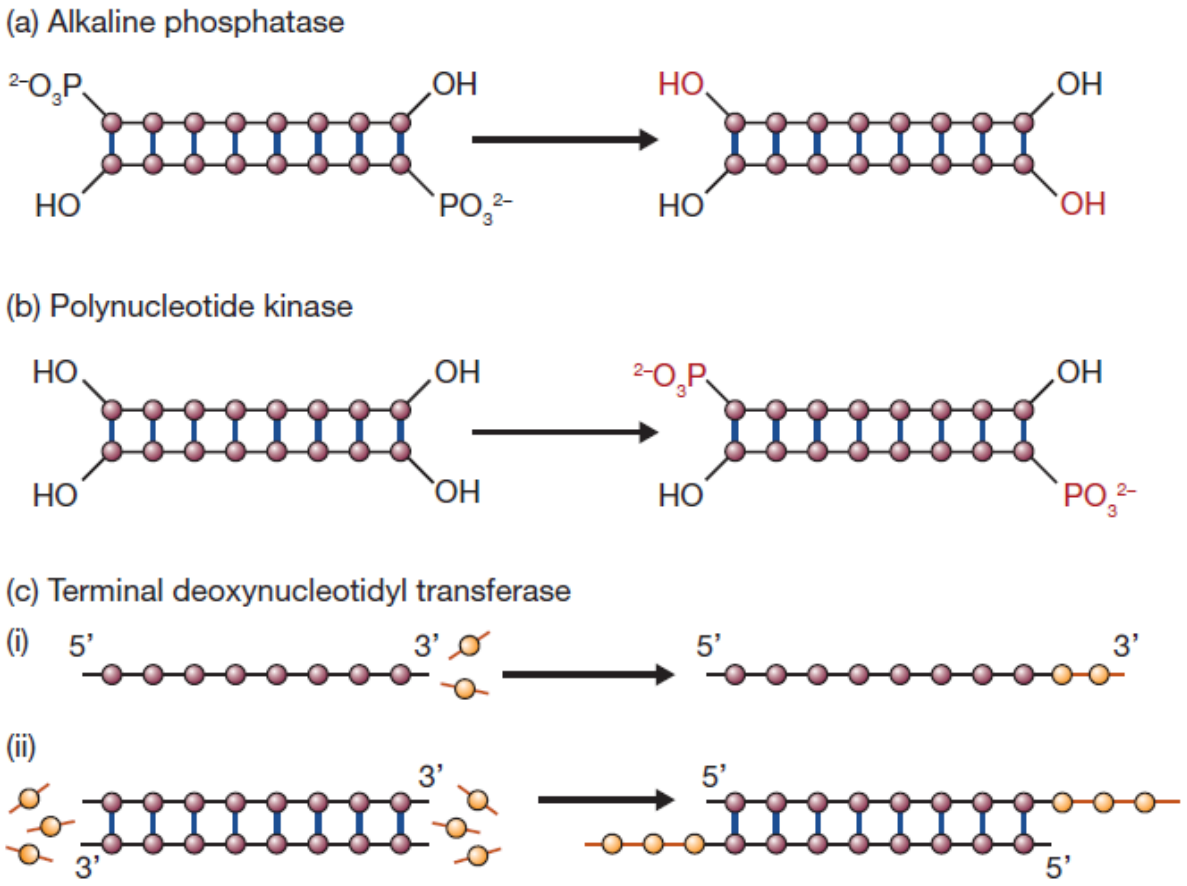


Fig: The reactions catalyzed by DNA modifying enzymes. (a) Alkaline phosphatase, which removes 5′-phosphate groups. (b) Polynucleotide kinase, which attaches 5′-phosphate groups. (c) Terminal deoxynucleotidyl transferase, which attaches deoxyribonucleotides to the 3′ termini of polynucleotides in either (i) single-stranded or (ii) double-stranded molecules.

Enzymes for cutting DNA—RESTRICTION ENDONUCLEASES
Gene cloning requires that DNA molecules be cut in a very precise and reproducible fashion. Purified restriction endonucleases allow the molecular biologist to cut DNA molecules in the precise, reproducible manner required for gene cloning. The discovery of these enzymes, which led to Nobel Prizes for W. Arber, H. Smith, and D. Nathans in 1978, was one of the key breakthroughs in the development of genetic engineering.

The discovery and function of restriction endonucleases

The initial observation that led to the eventual discovery of restriction endonucleases was made in the early 1950s, when it was shown that some strains of bacteria are immune to bacteriophage infection, a phenomenon referred to as **host-controlled restriction.** Restriction occurs because the bacterium produces an enzyme that degrades the phage DNA before it has time to replicate and direct synthesis of new phage particles (Figure a). The bacterium's own DNA, the destruction of which would of course be lethal, is protected from attack because it carries additional methyl groups that block the degradative enzyme action (Figure b). These degradative enzymes are called restriction endonucleases and are synthesized by many, perhaps all, species of bacteria: over 2500 different ones have been isolated and more than 300 are available for use in the laboratory.

The first part of this system is the restriction endonucleases that selectively recognizes specific DNA sequences and fragments the DNA having these sequences, which leads to its complete degradation. Since these enzymes are responsible for the restricted growth and multiplication of bacterial viruses in bacteria they are called **restriction endonucleases.**

The second part is the modification system, which recognizes the restriction sites and modifies those sites by adding methyl groups to one or two bases within those restriction sites. Thus, it modifies the restriction sites by methylation, so that the restriction enzymes are not able to cut those sites. This usually happens for the DNA of the host cells. Thus, these modification systems protect the DNA of the host cell from the action of its own restriction systems.

The discovery of these enzymes marked the beginning of recombinant DNA research and sequence-specific modification of DNA molecules. Since these restriction endonucleases are responsible for converting the DNA molecules into a specific number of fragments by cutting the molecule at certain specific sequences known as **restriction sites**, these enzymes are known as **molecular scissors.**

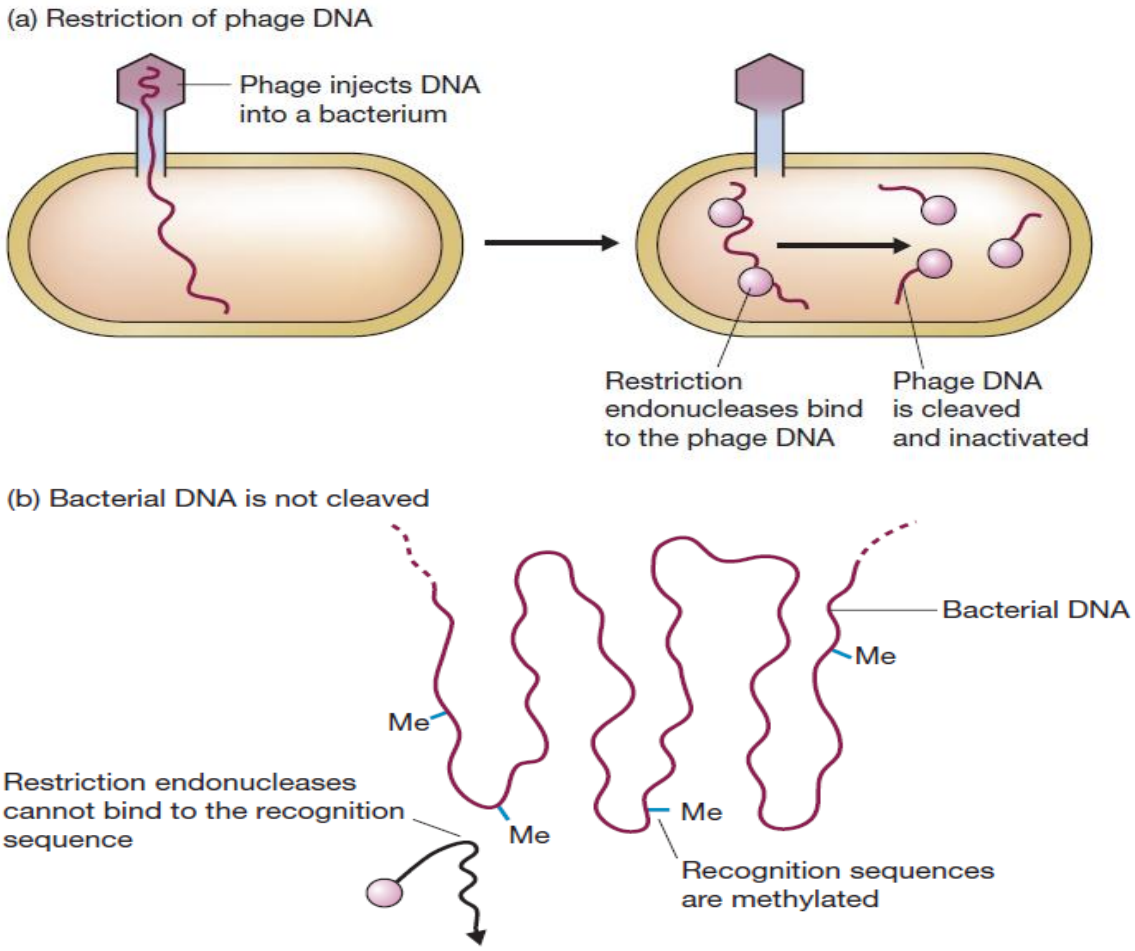


Fig:Host restriction mechanism: Type II restriction endonucleases cut DNA at specific nucleotide sequences

Nomenclature

These restriction enzymes generally have names that reflect their origin—the first letter of the name comes from the genus and the second two letters come from the species of the prokaryotic cell from which they were isolated. For example, *Eco*RI comes from *Escherichia coli* RY13 bacteria, while *Hind*-II comes from hemophilus influenzae strain Rd. Numbers following the nuclease names indicate the order in which the enzymes were isolated from single strains of bacteria.

Types:

Three different classes of restriction endonuclease are recognized, each distinguished by a slightly different mode of action. **Types I and III** are rather complex and have only a limited role in genetic engineering. **Type II restriction endonucleases**, on the other hand, are the cutting enzymes that are so important in gene cloning.

Type I and III restriction endonucleases: They have both endonuclease and methylase activities on a single protein. Type I REs cleave the DNA at a random site located at 1,000 base pairs from the recognition site. Type III RE does the same at 24 to 24 bp away from the recognition sites.

Type II RE: They cleave DNA at specific site within the recognition sequence. This property has made type II RE a most important tool in molecular biology.

Restriction sites

The central feature of type II restriction endonucleases (which will be referred to simply as "restriction endonucleases" from now on) is that each enzyme **has a specific recognition sequence at which it cuts a DNA molecule.** The restriction sites or the recognition sequences are usually four to eight nucleotides in length and are palindromic. The palindromic sequences read the same on both the strands of DNA in a 5′–3′ direction. For example, the restriction site of *Eco*RI is 5′GGATCC3′ and the cleave site is between the G and A on the complementary strands, which is demonstrated in Figure .

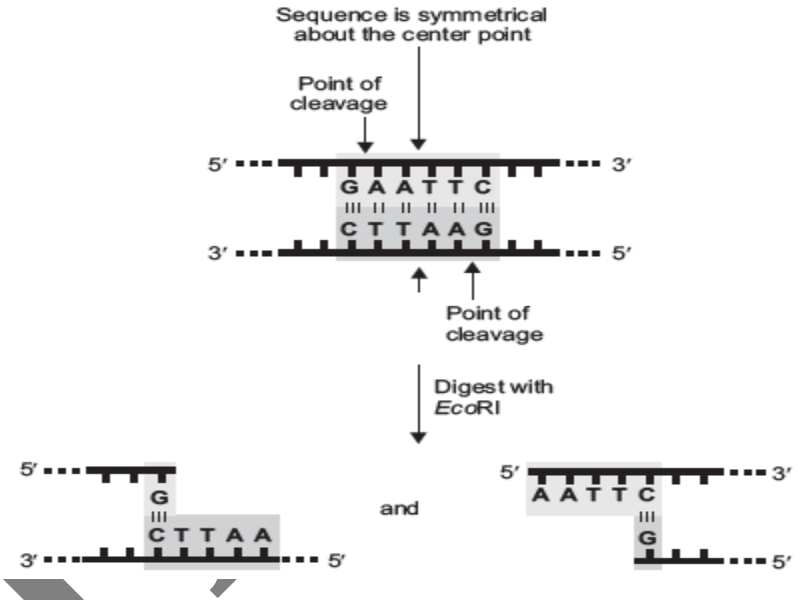


FIGURE The restriction site sequence of EcoRI is a palindrome of a hexamer and it cuts between the bases G and A on both the strands resulting in sticky ends or cohesive ends

The restriction-site recognition sequence for different restriction enzymes is unique and therefore different endonucleases yield different sets of cuts or DNAfragments. The restriction endonuclease called *Pvu*I (isolated from *Proteus vulgaris*) cuts DNA only at the hexanucleotide CGATCG.

Many restriction endonucleases recognize hexanucleotide target sites, but others cut at four, five, eight, or even longer nucleotide sequences. *Sau*3A (from *Staphylococcus aureus* strain 3A recognizes GATC, and *Alu*I (*Arthrobacter luteus*) cuts at AGCT. There are also examples of restriction endonucleases with degenerate recognition sequences, meaning that they cut DNA at any one of a family of related sites. *Hin*fI (*Haemophilus influenzae* strain Rf), for instance, recognizes GA**N**TC, so cuts at GAATC, GATTC, GAGTC, and GACTC. The recognition sequences for some of the most frequently used restriction endonucleases are listed in Table.

TABLE Examples of Type II restriction endonucleases used in recombinant DNA experiments.

Enzyme	Organism from which Derived	Target Sequence (cut at*) 5′ → 3′
Ava I	*Anabaena variabilis*	5′C*C/TCGA/GG 3′ 3′GG/AGCT/C*C5′
Bam HI	*Bacillus amyloliquefaciens*	5′G*GATCC3′ 3′CCTAG*G5′
Bgl II	*Bacillus globigii*	5′A*GATCT3′ 3′TCTAG*A5′
Eco RI	*Escherichia coli* RY 13	5′G*AATTC3′ 3′CTTAA*G
Eco RII	*Escherichia coli* R245	5′*CCA/TGG3′ 3′GGT/ACC*5′
Hae III	*Haemophilus aegyptius*	5′GG*CC3′ 3′CC*GG5′
Hha I	*Haemophilus haemolyticus*	5′GCG*C3′ 3′C*GCG5′
Hind III	*Haemophilus inflenzae* Rd	5′A*AGCT T3′ 3′T TCGA*A5′
Hpa I	*Haemophilus parainflenzae*	5′GTT*AAC3′ 3′CAA*TTG5′
Kpn I	*Klebsiella pneumoniae*	5′G GTAC*C3′

Separation of DNA by gel electrophoresis

Gel Electrophoresis is a procedure used in molecular biology to separate and identify molecules (such as DNA, RNA, protein, complexes) by size. The separation of these molecules is achieved by placing them in a gel made up of small pores and setting an electric field across the gel. The molecules will move based on their inherent electric charge (i.e., negatively charged molecules

move away from the negative pole) and smaller molecules will move faster than larger molecules; thus, a size separation is achieved within the pool of molecules running through the gel. The gel works in a similar manner to a sieve separating particles by size; the electrophoresis works to move the particles, using their inherent electric charge, through the sieve.

Purpose of Gel Electrophoresis

The purpose of gel electrophoresis is to visualize, identify and distinguish molecules that have been processed by a previous method such as PCR, enzymatic digestion or an experimental condition. Often, mixtures of nucleic acids or proteins that are collected from a previous experiment/method are run through gel electrophoresis to determine identity or differentiate molecules.

Types of Gel Electrophoresis

There are two types of gel electrophoresis: native and denaturing. A native gel electrophoresis usually attempts to keep RNA or protein in its native structure while running it through the gel. A denaturing gel electrophoresis attempts to reduce the RNA or protein into its most linear structure before or during gel electrophoresis. The denaturation of the RNA or protein is accomplished by adding a reducing agent to the sample, gel and/or buffer. The reducing agent separates bonds within the RNA or protein molecule and thereby reduces its secondary structure. The secondary structure of a protein or RNA will influence, in a non-linear manner, how fast it migrates through a gel. A denatured, linear form of RNA or protein, however, will migrate proportionally to its linear size (base pairs or kilo Daltons). Denaturing gel electrophoresis is often more accurate for size identification, whereas native gel electrophoresis is usually used to identify protein complexes.

Examples of Gel Electrophoresis

- TAE Agarose Gel Electrophoresis is most commonly used for DNA.
- TBE and Denaturing PAGE (polyacrylamide gel electrophoresis) are common for RNA separation.
- SDS PAGE is a denaturing gel electrophoresis commonly used for protein identification and separation.

Gel Electrophoresis Steps

The broad steps involved in a common DNA gel electrophoresis protocol:

1. Preparing the samples for running

The DNA is isolated and preprocessed (e.g. PCR, enzymatic digestion) and made up in solutionwith some basic blue dye to help visualize the movement of the sample through the gel.

2. An agarose TAE gel solution is prepared

TAE buffer provides a source of ions for setting up the electric field during electrophoresis. A weight to volume concentration of agarose in TAE buffer is used to prepare the solution. For example, if a 1% agarose gel is required, 1g of agarose is added to 100mL of TAE. The agarose percentage used is determined by how big or small the DNA is expected to be. If one is looking at separating a pool of smaller size DNA bands (<500bp), a higher percentage agarose gel (>1%) is prepared. The higher percentage agarose creates a denser sieve to increase separation of small DNA length differences. The agarose-TAE solution is heated to dissolve the agarose.

3. Casting the gel

The agarose TAE solution is poured into a casting tray that, once the gel solution has cooled down and solidified, creates a gel slab with a row of wells at the top.

4. Setting up the electrophoresis chamber

The solid gel is placed into a chamber filled with TAE buffer. The gel is positioned so that the chamber wells are closest to the negative electrode of the chamber.

5. Loading the gel

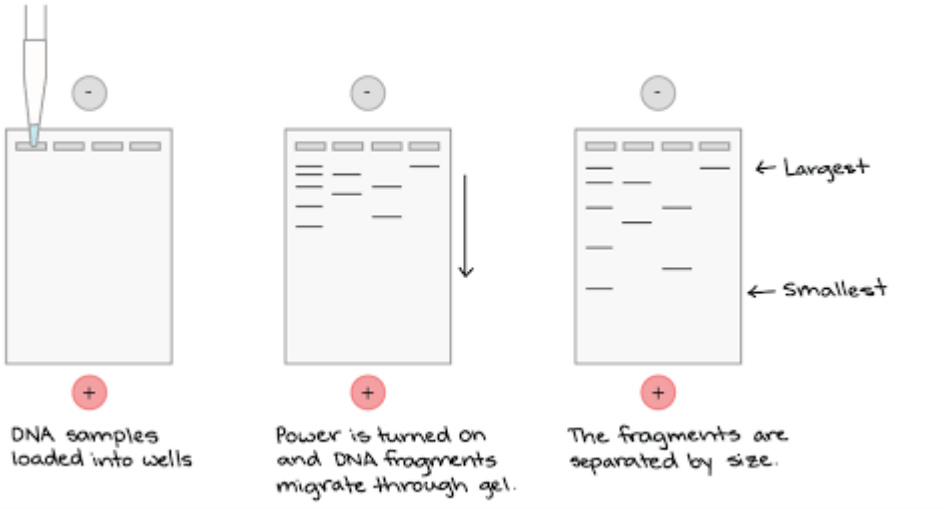
The gel chamber wells are loaded with the DNA samples and usually a DNA ladder is also loaded as reference for sizes.

6. Electrophoresis

The negative and positive leads are connected to the chamber and to a power supply where voltage is set. Turning on the power supply sets up the electric field and the negatively charged DNA samples will start to migrate through the gel and away from the negative electrode towards the positive.

7. Stopping electrophoresis and visualizing the DNA

Once the blue dye in the DNA samples has migrated through the gel far enough, the power supply is turned off and the gel is removed and placed into an ethidium bromide solution. Ethidium bromide intercalates between DNA and is visible in UV light. Sometimes ethidium bromide is added directly to the agarose gel solution in step 2. The ethidium bromide stained gel is then exposed to UV light and a picture is taken. DNA bands are visualized in from each lane corresponding to a chamber well. The DNA ladder that was loaded is also visualized and the length of the DNA bands can be estimated.



Extraction and purification of plasmid and bacteriophage DNA

2.PREPARATION OF PLASMID DNA

Purification of plasmids from a culture of bacteria involves the same general strategy as preparation of total cell DNA. A culture of cells, containing plasmids, is grown in liquid

medium, harvested, and a cell extract prepared. The protein and RNA are removed, and the DNA probably concentrated by ethanol precipitation. However, there is an
important distinction between plasmid purification and preparation of total cell DNA. In a plasmid preparation it is always necessary to separate the plasmid DNA from the large amount of bacterial chromosomal DNA that is also present in the cells. Separating the two types of DNA can be very difficult, but is nonetheless essential if the asmids are to be used as cloning vectors. The presence of the smallest amount of contaminating bacterial DNA in a gene cloning experiment can easily lead to undesirable results. Fortunately several methods are available for removal of bacterial DNA during plasmid purification, and the use of these methods, individually or in combination, can
result in isolation of very pure plasmid DNA.

The methods are based on the several physical differences between plasmid DNA and bacterial DNA, the most obvious of which is size. The largest plasmids are only 8% of the size of the *E. coli* chromosome, and most are much smaller than this. Techniques that can separate small DNA molecules from large ones should therefore effectively
purify plasmid DNA.
In addition to size, plasmids and bacterial DNA differ in **conformation**. When applied to a polymer such as DNA, the term conformation refers to the overall spatial configuration of the molecule, with the two simplest conformations being linear and circular. Plasmids and the bacterial chromosome are circular, but during preparation of the cell extract the chromosome is always broken to give linear fragments. A method for separating circular from linear molecules will therefore result in pure plasmids.

a)Separation on the basis of size
The usual stage at which size fractionation is performed is during preparation of the cell extract. If the cells are lysed under very carefully controlled conditions, only a minimal amount of chromosomal DNA breakage occurs. The resulting DNA fragments are still very large—much larger than the plasmids—and can be removed with the cell debris by centrifugation. This process is aided by the fact that the bacterial chromosome is physically attached to the cell envelope, so fragments of the chromosome sediment with the cell debris if these attachments are not broken.
Cell disruption must therefore be carried out very gently to prevent wholesale breakage of the bacterial DNA. For *E. coli* and related species, controlled lysis is performed as shown in Figure 3.11. Treatment with EDTA and lysozyme is carried out in the presence of sucrose, which prevents the cells from bursting immediately. Instead, **sphaeroplasts** are formed, cells with partially degraded cell walls that retain an intact cytoplasmic membrane. Cell lysis is now induced by adding a non-ionic detergent such as Triton X-100 (ionic detergents, such as SDS, cause chromosomal breakage). This method causes very little breakage of the bacterial DNA, so centrifugation leaves a **cleared lysate**, consisting almost entirely of plasmid DNA. A cleared lysate will, however, invariably retain some chromosomal DNA. Furthermore, if the plasmids themselves are large molecules, they may also sediment with the cell debris. Size fractionation is

therefore rarely sufficient on its own, and we must consider alternative ways of removing the acterial DNA contaminants.

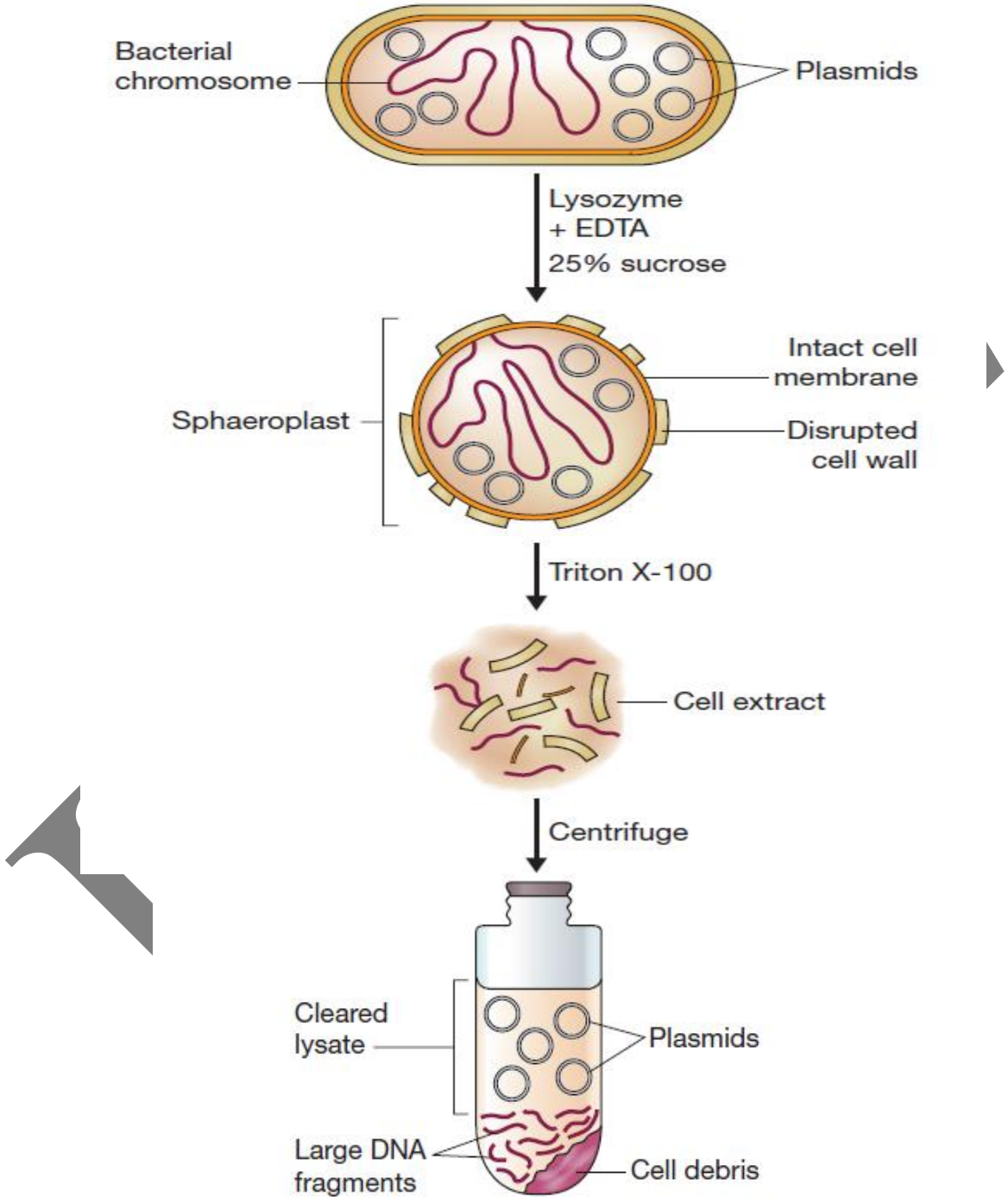


Fig: Preparation of a cleared lysate.

b) Separation on the basis of conformation

Before considering the ways in which conformational differences between plasmids and bacterial DNA can be used to separate the two types of DNA, we must look more closely at the overall structure of plasmid DNA. It is not trictly correct to say that plasmids have a circular conformation, because double-stranded DNA circles can take up one of two quite distinct configurations. Most plasmids exist in the cell as **supercoiled** molecules (Figure). Supercoiling occurs because the double helix of the plasmid DNA is partially unwound during the plasmid replication process by enzymes called topoisomerases (p. 69). The supercoiled conformation can be maintained only if both polynucleotide strands are intact, hence the more technical name of **covalently closedcircular (ccc) DNA**. If one of the polynucleotide strands is broken the double helix reverts to its normal **relaxed** state, and the plasmid takes on the alternative conformation, called **open-circular (oc)** (Figure).

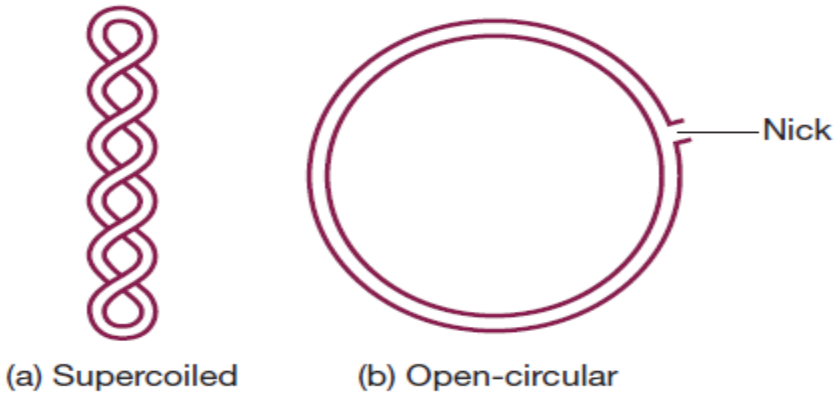


Fig: Two conformations of circular double-stranded DNA: (a) supercoiled—both strands are intact; (b) open-circular
—one or both strands are nicked

Supercoiling is important in plasmid preparation because supercoiled molecules can be fairly easily separated from non-supercoiled DNA. Two different methods are commonly used. Both can purify plasmid DNA from crude cell extracts, although in practice best results are obtained if a cleared lysate is first prepared.

i)Alkaline denaturation

The basis of this technique is that there is a narrow pH range at which non-supercoiled DNA is denatured, whereas supercoiled plasmids are not. If sodium hydroxide is added to a cell extract or cleared lysate, so that the pH is djusted to 12.0–12.5, then the hydrogen bonding in non-supercoiled DNA molecules is broken, causing the double
helix to unwind and the two polynucleotide chains to separate (Figure 3.13). If acid is now added, these denatured bacterial DNA strands reaggregate into a tangled mass. The insoluble network can be pelleted by centrifugation, leaving plasmid DNA in the supernatant. An additional advantage of this procedure is that, under some circumstances (specifically cell lysis

by SDS and neutralization with sodium acetate), most of the protein and RNA also becomes insoluble and can be removed by the centrifugation step. Further purification by organic extraction or column chromatography may therefore not be needed if the alkaline denaturation method is used.

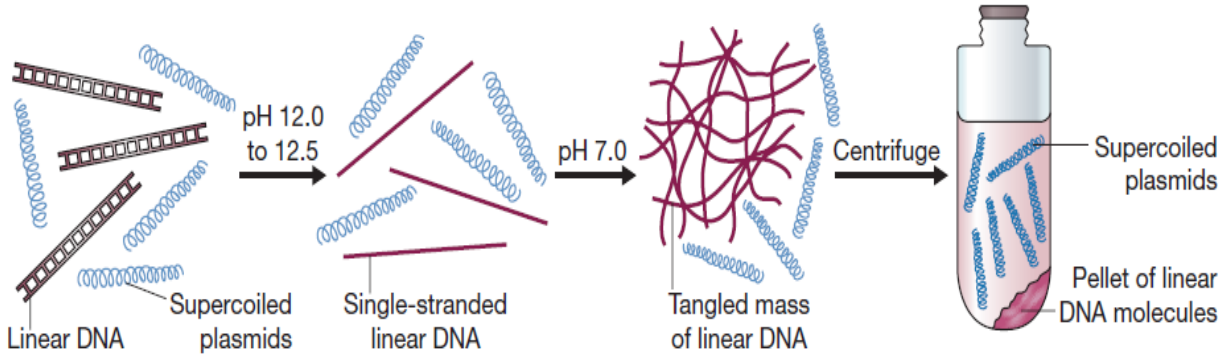


Fig: Plasmid purification by the alkaline denaturation method.

ii)Ethidium bromide–caesium chloride density gradient centrifugation

This is a specialized version of the more general technique of equilibrium or **density gradient centrifugation**. A density gradient is produced by centrifuging a solution of caesium chloride (CsCl) at a very high speed (Figure). Macromolecules present in the CsCl solution when it is centrifuged form bands at distinct points in the gradient
(Figure 3.14b). Exactly where a particular molecule bands depends on its **buoyant density**: DNA has a buoyant density of about 1.70 g/cm3, and therefore migrates to the point in the gradient where the CsCl density is also **1.70** g/cm3. In contrast, protein molecules have much lower buoyant densities, and so float at the top of the tube, whereas RNA forms a pellet at the bottom (Figure). Density gradient centrifugation can therefore separate DNA, RNA, and protein and is an alternative to organic extraction or column chromatography for DNA purification.

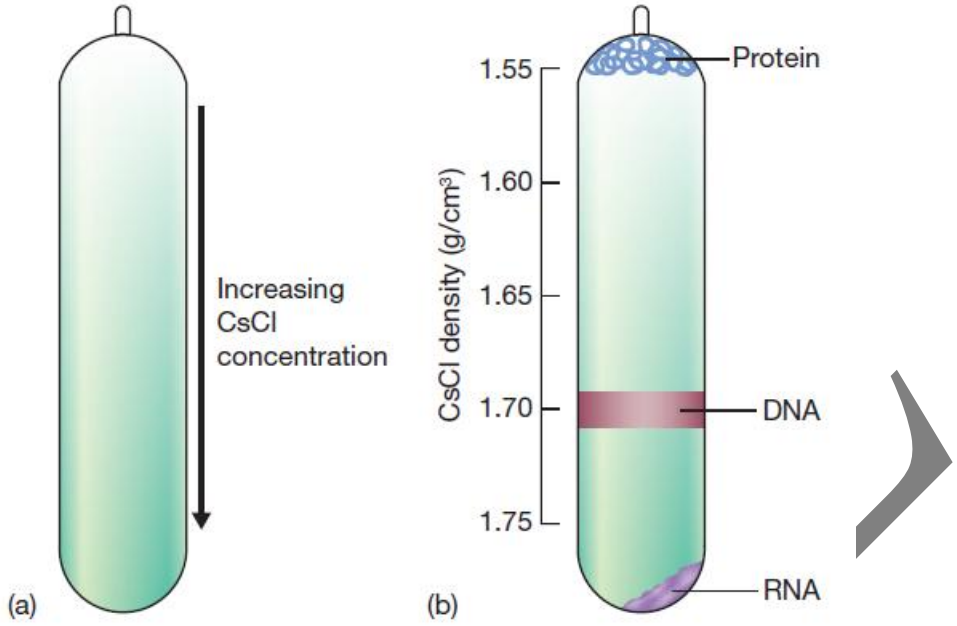
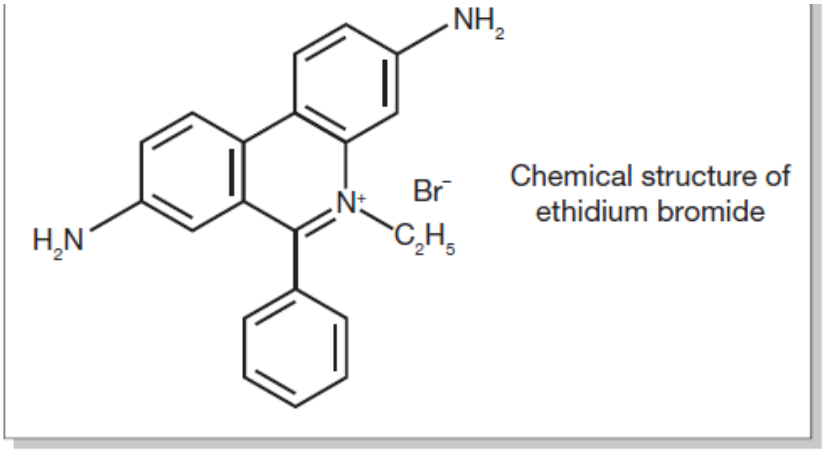


Fig: Caesium chloride density gradient centrifugation. (a) A CsCl density gradient produced by high speed centrifugation. (b) Separation of protein, DNA, and RNA in a density gradient.

More importantly, density gradient centrifugation in the presence of **ethidium bromide (EtBr)** can be used to separate supercoiled DNA from non-supercoiled molecules. Ethidium bromide binds to DNA molecules by intercalating between adjacent base pairs, causing partial unwinding of the double helix (Figure).



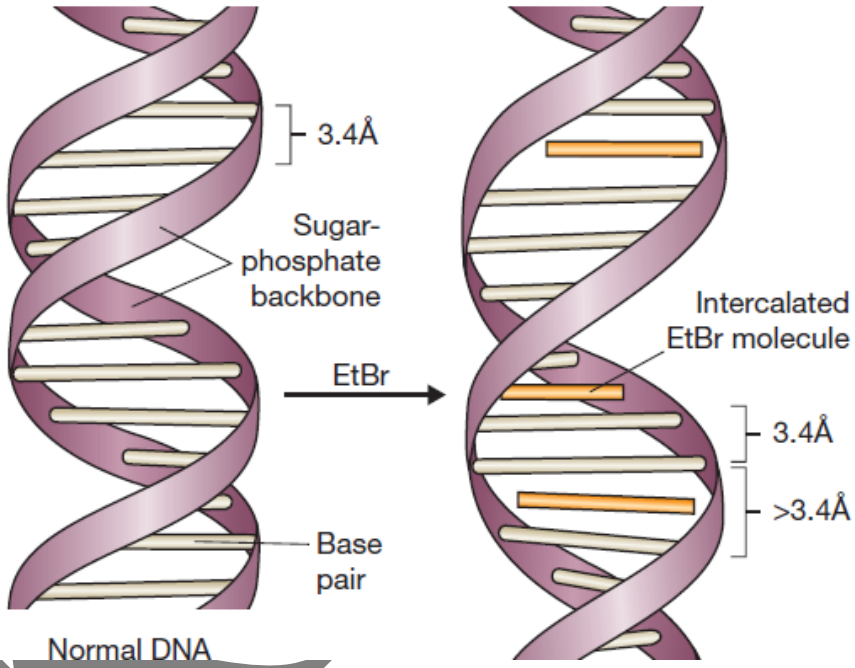


Fig: Partial unwinding of the DNA double helix by EtBr intercalation between adjacent base pairs. The normal DNA molecule shown on the left is partially unwound by taking up four EtBr molecules, resulting in the "stretched" structure on the right.

This unwinding results in a decrease in the buoyant density, by as much as 0.125 g/cm3 for linear DNA. However, supercoiled DNA, with no free ends, has very little freedom to unwind, and can only bind a limited amount of EtBr. The decrease in buoyant density of a supercoiledmolecule is therefore much less, only about 0.085 g/cm3. As a consequence, supercoiledmolecules form a band in an EtBr–CsCl gradient at a different position to linear and

open-circular DNA (Figure)

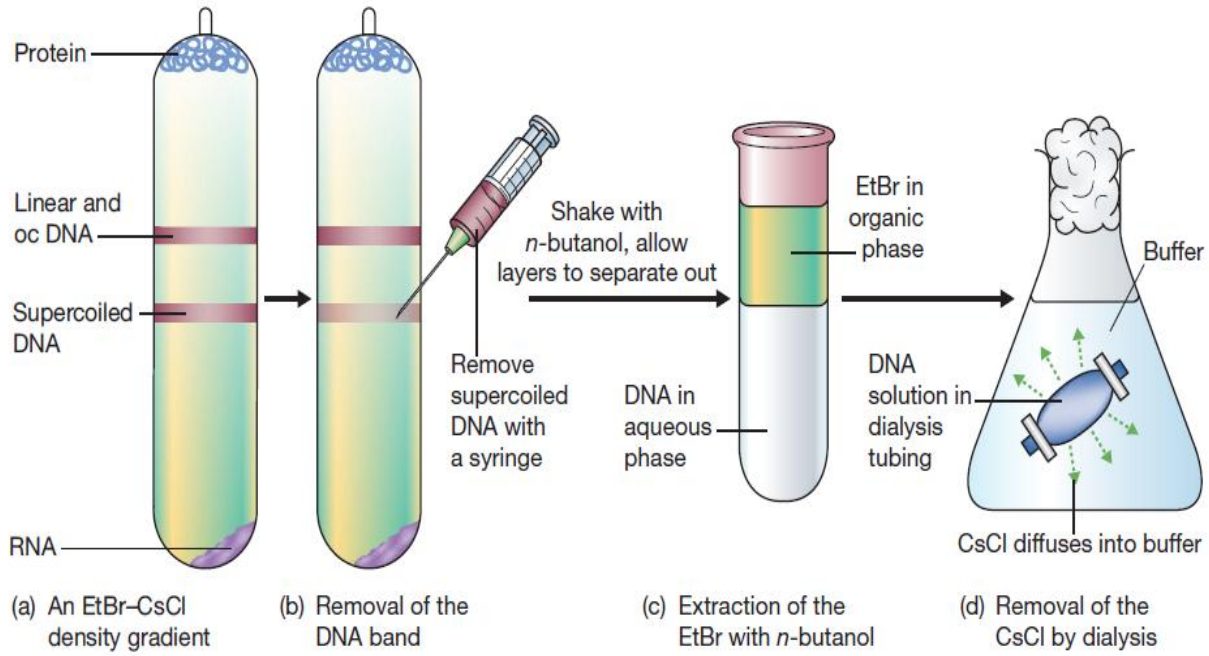


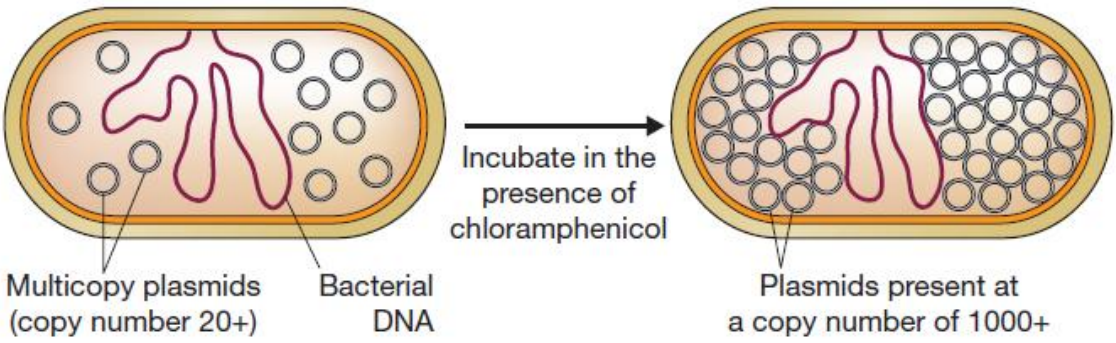
Fig: Purification of plasmid DNA by EtBr–CsCl density gradient centrifugation

Ethidium bromide–caesium chloride density gradient centrifugation is a very efficient method for obtaining pure plasmid DNA. When a cleared lysate is subjected to this procedure, plasmids band at a distinct point, separated from the linear bacterial DNA, with the protein floating on the top of the gradient and RNA pelleted at the bottom. The position of the DNA bands can be seen by shining ultraviolet radiation on the tube, which causes the bound EtBr to fluoresce. The pure plasmid DNA is removed by puncturing the side of the tube and withdrawing a sample with a syringe (Figure). The EtBr bound to the plasmid DNA is extracted with *n*-butanol (Figure) and the CsCl removed by dialysis (Figure). The resulting plasmid preparation is virtually 100% pure and ready for use as a cloning vector.

Plasmid amplification

Preparation of plasmid DNA can be hindered by the fact that plasmids make up only a small proportion of the total DNA in the bacterial cell. The yield of DNA from a bacterial culture may therefore be disappointingly low. **Plasmid amplification** offers a means of increasing this yield. The aim of amplification is to increase the copy number of a plasmid. Some **multicopy plasmids** (those with copy numbers of 20 or more) hav e the useful property of being able to replicate in the absence of protein synthesis. This contrasts with the main bacterial chromosome, which cannot replicate under these conditions. This property can be utilized during the growth of a bacterial culture for plasmid DNA purification. After a satisfactory cell density has been reached, an inhibitor of protein synthesis (e.g., chloramphenicol) is added, and the culture incubated for a

further 12 hours. During this time the plasmid molecules continue to replicate, even though chromosome replication and cell division are blocked (Figure). The result is that plasmid copy numbers of several thousand may be attained. Amplification is therefore a very efficient way of increasing the yield of multicopy plasmids.



2.Preparation of bacteriophage DNA

The key difference between phage DNA purification and the preparation of either total cell DNA or plasmid DNA is that for phages the starting material is not normally a cell extract. This is because bacteriophage particles can be obtained in large numbers from the extracellular medium of an infected bacterial culture. When such a culture is centrifuged, the bacteria are pelleted, leaving the phage particles in suspension (Figure 3.18). The phage particles are then collected from the suspension and their DNA extracted by a single deproteinization step to remove the phage capsid. This overall process is more straightforward than the procedure used to prepare total cell or plasmid DNA. Nevertheless, successful purification of significant quantities of phage DNA is subject to several pitfalls. The main difficulty, especially with e, is growing an infected culture in such a way that the extracellular phage titer (the number of phage particles per ml of culture) is sufficiently high. In practical terms, the maximum titer that can reasonably be expected for e is 1010 per ml; yet 1010 e particles will yield only 500 ng of DNA. Large culture volumes, in the range of 500–1000 ml, are therefore needed if substantial quantities of e DNA are to be obtained.

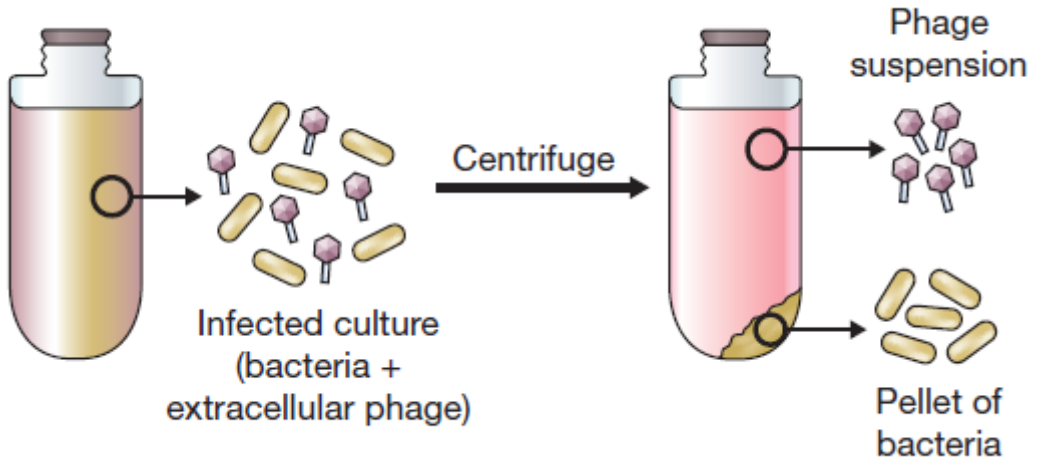


Fig: Preparation of a phage suspension from an infected culture of bacteria.

Collection of phages from an infected culture

The remains of lysed bacterial cells, along with any intact cells that are inadvertently left over, can be removed from an infected culture by centrifugation, leaving the phage particles in suspension (see Figure). The problem now is to reduce the size of the suspension to 5 ml or less, a manageable size for DNA extraction.

Phage particles are so small that they are pelleted only by very high speed centrifugation. Collection of phages is therefore usually achieved by precipitation with **polyethylene glycol (PEG)**. This is a long-chain polymeric compound which, in the presence of salt, absorbs water, thereby causing macromolecular assemblies such as phage particles to precipitate. The precipitate can then be collected by centrifugation, and redissolved in a suitably small volume (Figure).

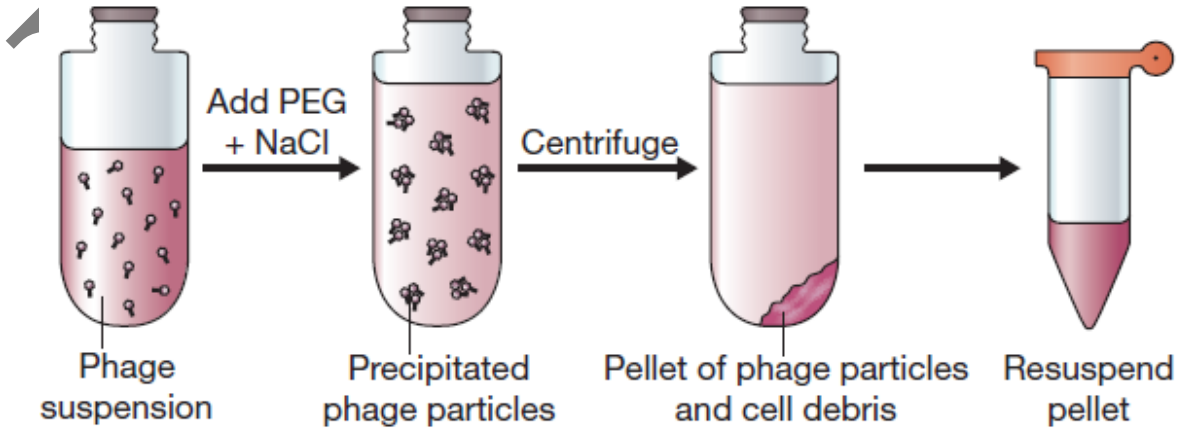


Fig: Collection of phage particles by polyethylene glycol (PEG) precipitation

Purification of DNA from * phage particles

Deproteinization of the redissolved PEG precipitate is sometimes sufficient to extract pure phage DNA, but usually e phages are subjected to an intermediate purification step. This is necessary because the PEG precipitate also ontains a certain amount of bacterial debris, possibly including unwanted cellular DNA. These contaminants can be separated from the e particles by CsCl density gradient centrifugation. The e particles band in a CsCl gradient at **1.45–1.50 g/cm3** (Figure), and can be withdrawn from the gradient as described previously for DNA bands (see Figure). Removal of CsCl by dialysis leaves a pure phage preparation from which the DNA can be extracted by either phenol or protease treatment to digest the phage protein coat.

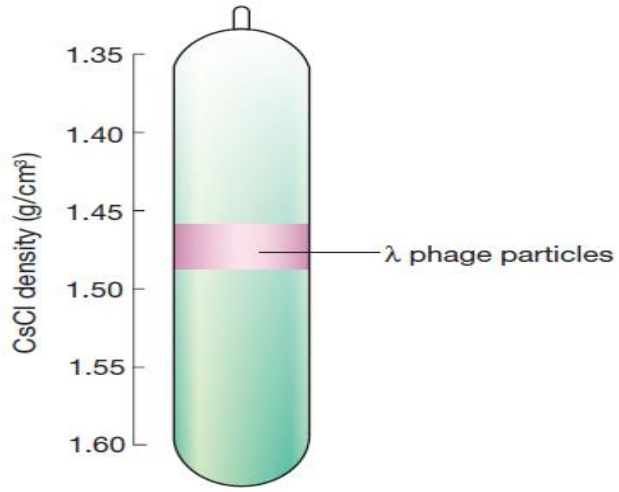


Fig: Purification of Lambda phage particles by CsCl density gradient centrifugation.

Joining of DNA fragments
Ligation of DNA molecules
DNA ligase

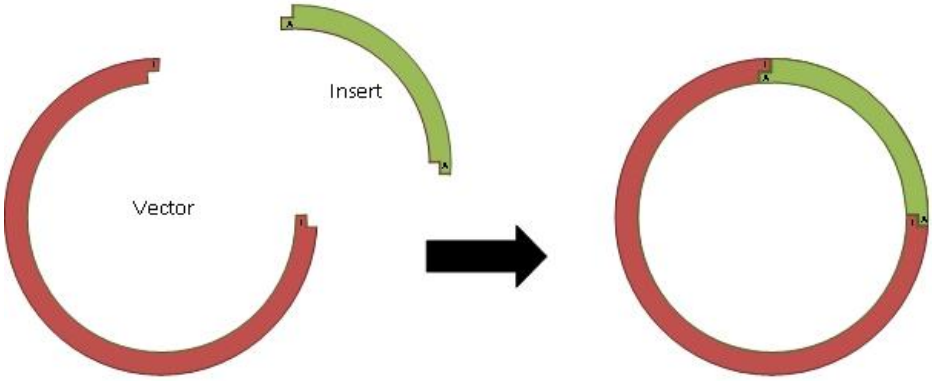
DNA ligase is an enzyme which can connect two strands of DNA together by forming a bond between the phosphate group of one strand and the deoxyribose group on another. It is used in cells to join together the Okazaki fragments which are formed on the lagging strand during DNA replication.

In molecular biology, DNA ligase can be used to insert genes of interest into plasmid vectors, or to create fusion genes by joining one gene onto another. This process is called ligation (literally "tying a knot").

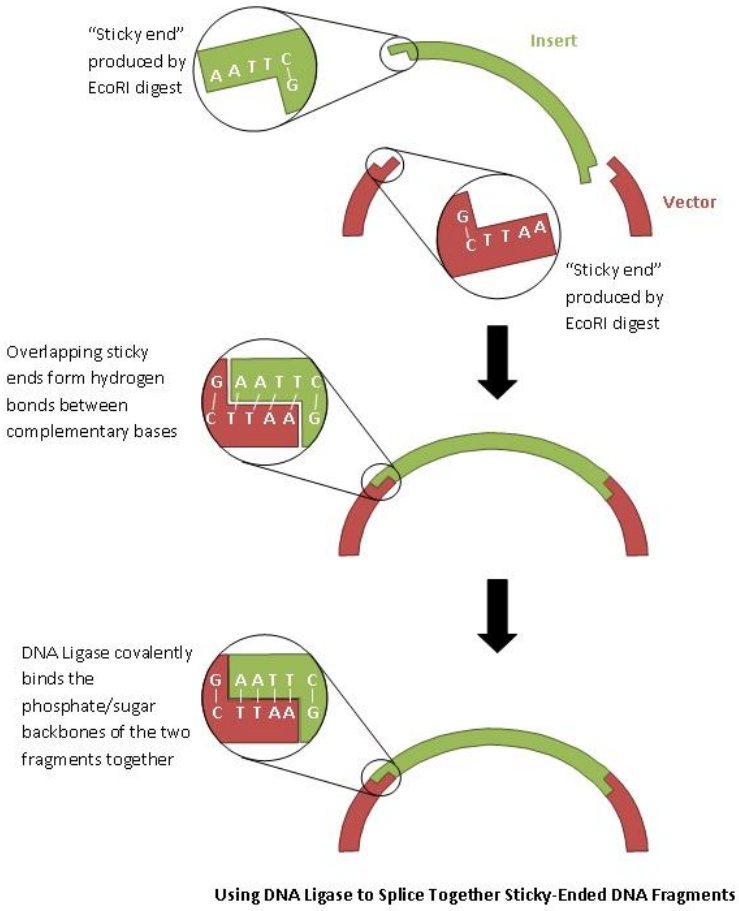
Ligation can be performed on lengths of DNA which have "blunt" or "sticky" ends following restriction digests.

In "blunt end" ligation, the DNA fragments are joined directly together by the DNA ligase. There is less control over the orientation of the resultant insertion, however this can be improved

by creating single base overhangs (eg. "A-Tailing" with Taq Polymerase and dATP – see figure below).



In "sticky end" ligation, overlapping regions of complementary single stranded DNA hydrogen bond to each other, and the DNA ligase enzyme connects the sugar phosphate backbones together. Through careful selection of restriction enzymes to create the sticky ends, a great deal of control can be exercised over the site of ligation (see figure below).



Using DNA Ligase to Splice Together Sticky-Ended DNA Fragments

Sticky ends, blunt ends

Type II restriction enzymes cut DNA and produce two types of fragments. Some restriction enzymes produce fragments with blunt ends, whereas others produce fragments with sticky (overhanging or staggered) ends.

Many restriction endonucleases make a simple double-stranded cut in the middle of the recognition sequence (Figure 4a), resulting in a **blunt end** or **flush end**. *Pvu*II, *Hind*-II, *Sma*-I and *Alu*I are examples of blunt end cutters.

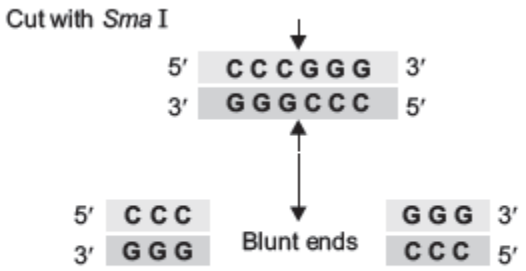


FIGURE 15.3 **Restriction cleavage with *Sma*-I resulting in blunt ends.**

Many endonucleases cleave the DNA backbones in positions that are not directly opposite to each other as in the case of *Bam* HI which results in protruding ends or sticky ends as illustrated in Figure.

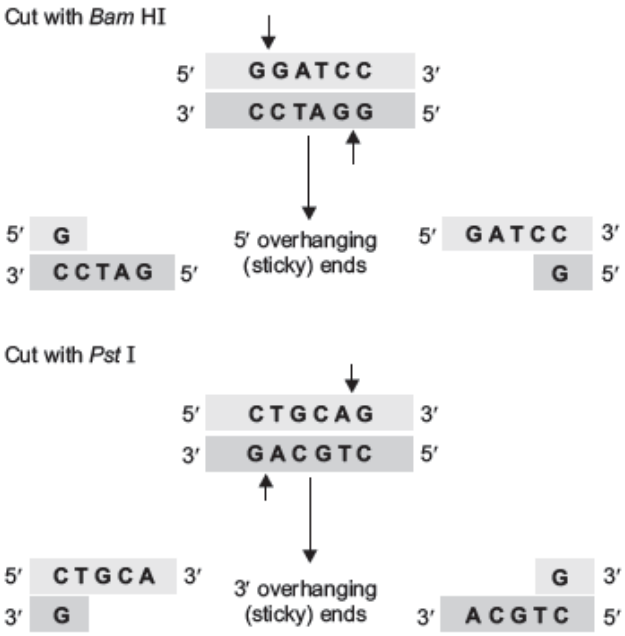


FIGURE 1Restriction digestion with Bam HI resulting in 5′□ overhanging ends and with Pst-I resulting in 3′□ overhanging ends.

Sticky ends are useful for DNA cloning because complementary sequences anneal and can be ligated directly by DNA ligase. If two different DNA samplescleaved with the same type of restriction enzymes are mixed together in the presence of DNA ligase, a recombinant DNA

molecule can be generated. This is possible because of the presence of the same type of sticky ends.

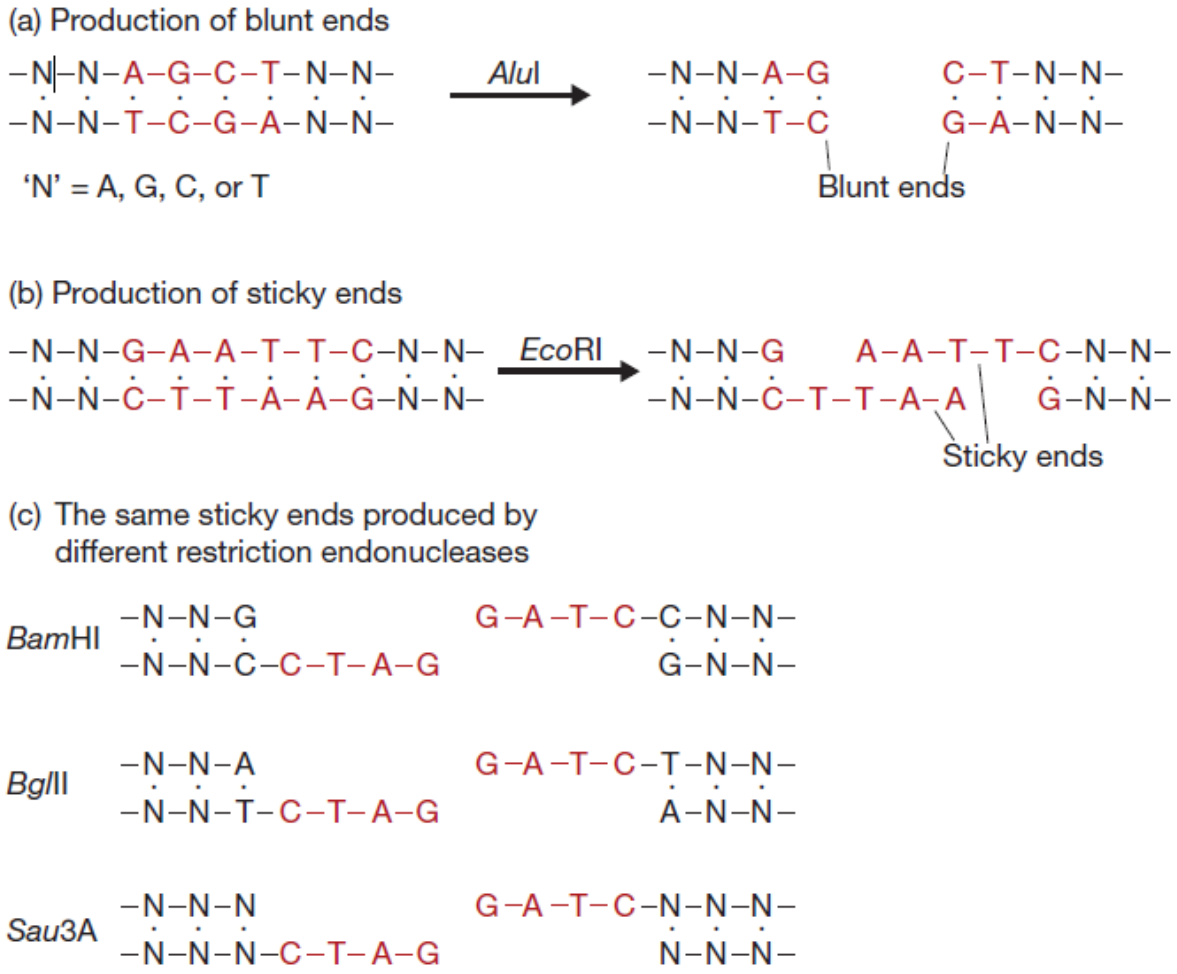


Fig: The ends produced by cleavage of DNA with different restriction endonucleases. (a) A blunt end produced by *Alu*I.(b) A sticky end produced by *Eco*RI. (c) The same sticky ends produced by *Bam*HI, *Bgl*II and *Sau*3A.

Putting sticky ends onto a blunt-ended molecule

For the reasons detailed in the preceding section, compatible sticky ends are desirable on the DNA molecules to be ligated together in a gene cloning experiment. Often these sticky ends can be provided by digesting both the vector and the DNA to be cloned with the same restriction endonuclease, or with different enzymes that produce the same

sticky end, but it is not always possible to do this. A common situation is where the vector molecule has sticky ends, but the DNA fragments to be cloned are blunt-ended. Under these circumstances one of three methods can be used to put the correct sticky ends onto the DNA fragments.

Llinkers and adapters

Linkers

The first of these methods involves the use of **linkers**. These are **short pieces of double stranded DNA,** of known nucleotide sequence, that are synthesized in the test tube. A typical linker is shown in Figure a. It is blunt-ended, but **contains a restriction site**, *Bam*HI in the example shown. DNA ligase can attach linkers to the ends of larger blunt ended DNA molecules. Although a blunt end ligation, this particular reaction can be performed very efficiently because synthetic oligonucleotides, such as linkers, can be made in very large amounts and added into the ligation mixture at a high concentration. More than one linker will attach to each end of the DNA molecule, producing the

chain structure shown in Figure. However, digestion with *Bam*HI cleaves the chains at the recognition sequences, producing a large number of cleaved linkers and the original DNA fragment, now carrying *Bam*HI sticky ends. This modified fragment is ready for ligation into a cloning vector restricted with *Bam*HI.

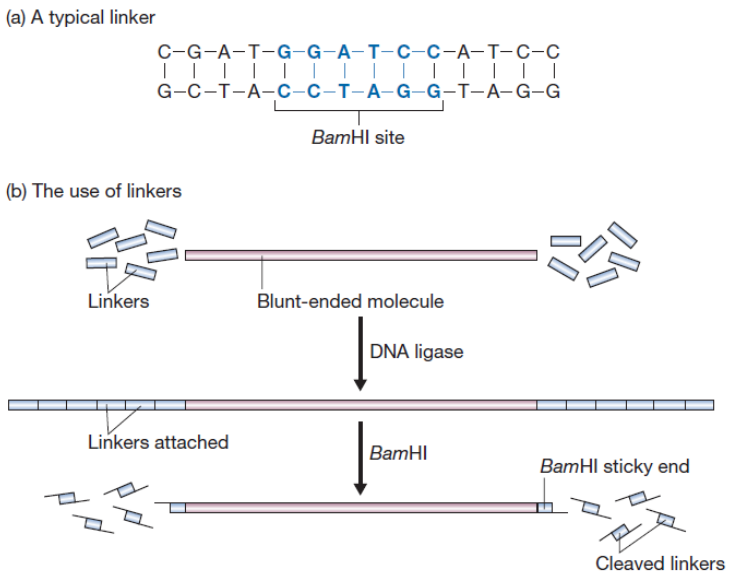


Fig: Linkers and their use: (a) the structure of a typical linker; (b) the attachment of linkers to a blunt-ended molecule.

Adaptors

There is one potential drawback with the use of linkers. Consider what would happen if the blunt-ended DNA olecule shown in Figure contained one or more *Bam*HI recognition

sequences. If this was the case, the restriction step needed to cleave the linkers and produce the sticky ends would also cleave the blunt-ended molecule (Figure). The resulting fragments will have the correct sticky ends, but that is no consolation if the gene contained in the blunt-ended fragment has now been broken into pieces.

The second method of attaching sticky ends to a blunt-ended molecule is designed to avoid this problem. **Adaptors**, like linkers, **are short synthetic oligonucleotides**. But unlike linkers, an adaptor is synthesized so that it **already has one sticky end** (Figure a). The idea is of course to ligate the blunt end of the adaptor to the blunt ends of the DNA fragment, to produce a new molecule with sticky ends. This may appear to be a simple method but in practice a new problem arises. The sticky ends of individual adaptor molecules could base pair with each other to form dimers (Figure b), so that the new DNA molecule is still blunt-ended (Figure c). The sticky ends could be recreated by digestion with a restriction endonuclease, but that would defeat the purpose of using adaptors in the first place.

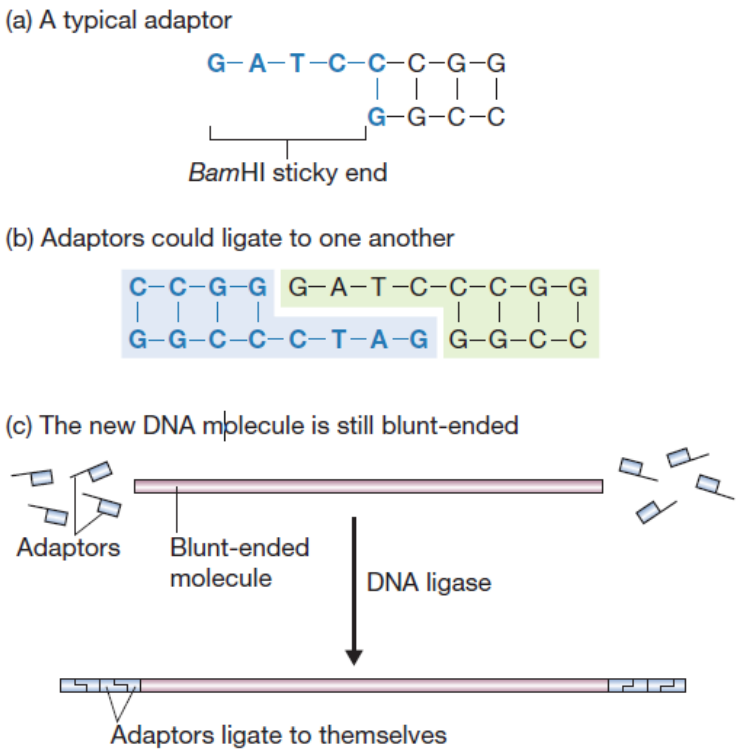


Fig: Adaptors and the potential problem with their use. (a) A typical adaptor. (b) Two adaptors could ligate to one another to produce a molecule similar to a linker, so that (c) after ligation of adaptors a blunt-ended molecule is still blunt-ended and the restriction step is still needed.

Synthetic oligonucleotides: synthesis and use

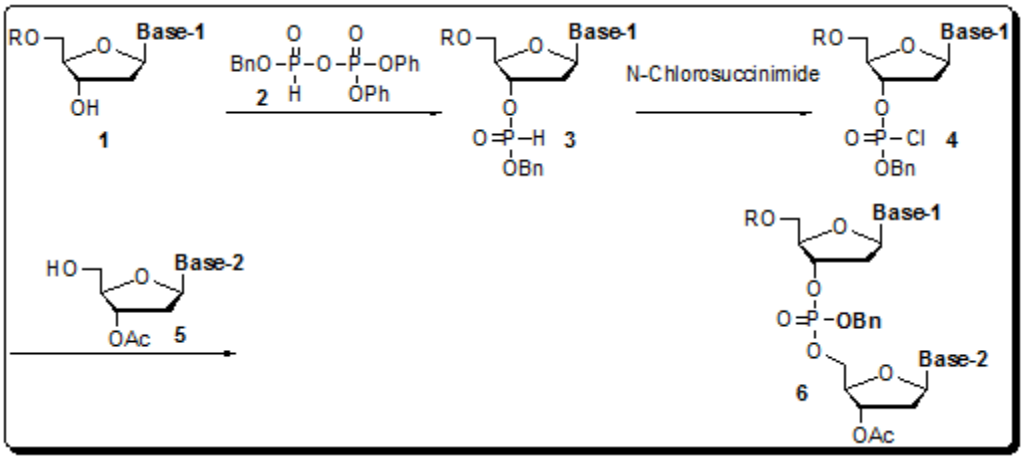
Oligonucleotide synthesis is the chemical synthesis of relatively short fragments of nucleic acids with defined chemical structure/sequence. The technique is extremely useful in current laboratory practice because it provides a rapid and inexpensive access to custom-made oligonucleotides of the desired sequence. Whereas enzymes synthesize DNA and RNA in a 5' to 3' direction, chemical oligonucleotide synthesis is carried out in the opposite, 3' to 5' direction. Currently, the process is implemented as solid-phase synthesis using phosphoramidite method and phosphoramidite building blocks derived from protected 2'-deoxynucleosides (dA, dC, dG, and T), ribonucleosides (A, C, G, and U), or chemically modified nucleosides, *e.g.* LNA. To obtain the desired oligonucleotide, the building blocks are sequentially coupled to the growing oligonucleotide chain in the order required by the sequence of the product. Upon the completion of the chain assembly, the product is released from the solid phase to solution, deprotected, and collected. The occurrence of side reactions sets practical limits for the length of synthetic oligonucleotides (up to about 200 nucleotide residues) because the number of errors accumulates with the length of the oligonucleotide being synthesized. Products are often isolated by HPLC to obtain the desired oligonucleotides in high purity. Typically, synthetic oligonucleotides are single-stranded DNA or RNA molecules around 15–25 bases in length. They are most commonly used as antisense oligonucleotides, small interfering RNA, primers for DNA sequencing and amplification, probes for detecting complementary DNA or RNA via molecular hybridization, tools for the targeted introduction of mutations and restriction sites, and for the synthesis of artificial genes.

4.17.1. Types of Chemical Synthesis

The evolution of oligonucleotide synthesis comprised of four major methods of the formation of inter-nucleosidic linkages.

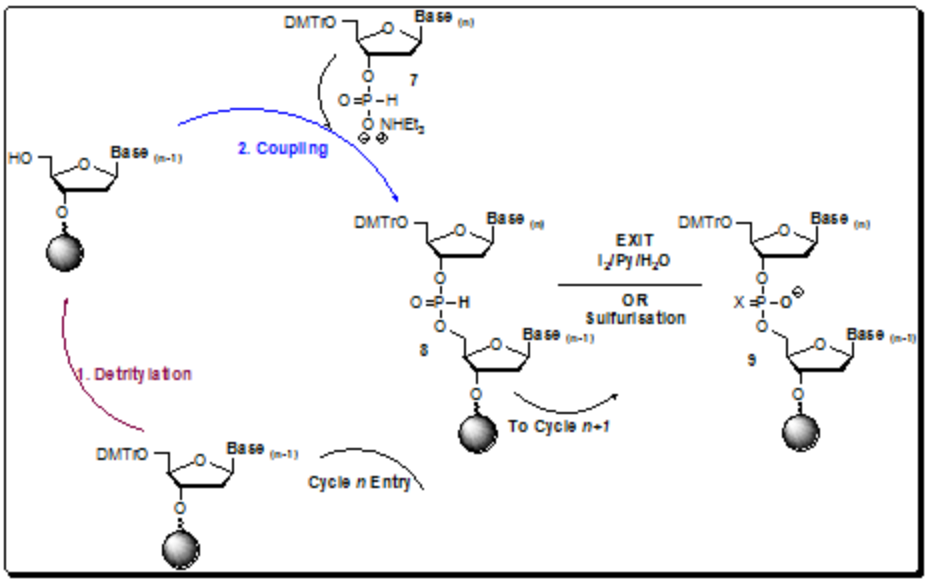
4.17.1.1. Early work and contemporary H-phosphonate synthesis

A. **Todd's synthesis:** In the early 1950s, Alexander Todd's group pioneered H-phosphonate and phosphate triester methods of oligonucleotide synthesis. The reaction of compounds **1**and **2** to form H-phosphonate diester **3** is an H-phosphonate coupling in solution while that of compounds **4** and **5** to give **6** is a phosphotriester coupling.



Scheme 4.1: Todd's DNA synthesis.

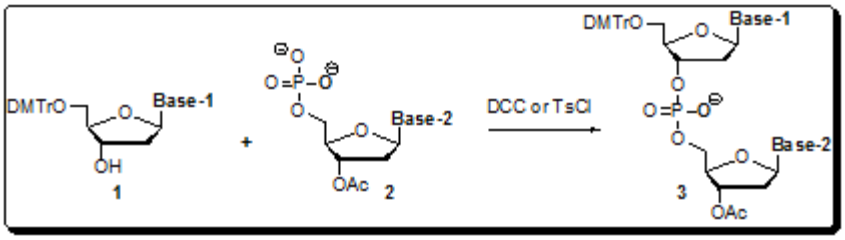
B. **H-Phosphonate Synthesis:** Thirty years later, this work inspired, independently, two research groups to adopt the H-phosphonate chemistry to the solid-phase synthesis using nucleoside H-phosphonate monoesters **7** as building blocks and pivaloyl chloride, 2,4,6-triisopropylbenzenesulfonyl chloride (TPS-Cl), and other compounds as activators. The practical implementation of H-phosphonate method resulted in a very short and simple synthetic cycle consisting of only two steps, detritylation and coupling (**Scheme 4.2**). Oxidation of inter-nucleosidic H-phosphonate diester linkages in **8** to phosphodiester linkages in **9** (X = O) with a solution of iodine in aqueous pyridine is carried out at the end of the chain assembly rather than as a step in the synthetic cycle. Alternatively, **8** can be converted to phosphorothioate **9** (X = S).



Scheme 4.2: Synthetic cycle in H-phosphonate method of oligonucleotide synthesis.

4.17.1.2. Phosphodiester synthesis

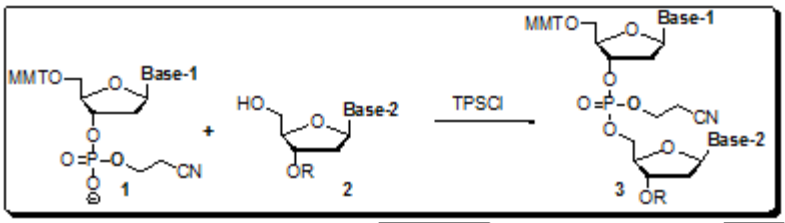
In the 1950s, Khorana and co-workers developed a phosphodiester method where 3'-*O*-acetylnucleoside-5'-*O*-phosphate **2** (**Scheme 4.3**) was activated with *N*,*N*'-dicyclohexylcarbodiimide (DCC) or 4-toluenesulfonylchloride (Ts-Cl) and a 5'-*O*-protected nucleoside **1** was reacted with the activated species to give a protected dinucleoside monophosphate **3**. Upon the removal of 3'-*O*-acetyl group using base-catalyzed hydrolysis, further chain elongation was carried out. Following this methodology, sets of tri- and tetradeoxyribonucleotides were synthesized and enzymatically converted to longer oligonucleotides, which allowed elucidation of the genetic code. The major limitation of the phosphodiester method consisted in the formation of pyrophosphate oligomers and oligonucleotides branched at the internucleosidic phosphate. The method seems to be a step back from the more selective chemistry described earlier; however, at that time, most phosphate-protecting groups available now had not yet been introduced. The lack of the convenient protection strategy necessitated taking a retreat to a slower and less selective chemistry to achieve the ultimate goal of the study.



Scheme 4.3: Oligonucleotide coupling by phosphodiester method.

4.17.1. 3. Phosphotriester Synthesis

In the 1960s, groups led by R. Letsinger and C. Reese developed a phosphotriester approach. The defining difference from the phosphodiester approach was the protection of the phosphate moiety in the building block **1** (**Scheme 4.4**) and in the product **3** with 2-cyanoethyl group. This precluded the formation of oligonucleotides branched at the internucleosidic phosphate. The higher selectivity of the method allowed the use of more efficient coupling agents and catalysts, which dramatically reduced the length of the synthesis. The method, led to the automation of the oligonucleotide chain assembly.



Scheme 4.4: Oligonucleotide coupling by phosphotriester method.

4.17.1.4. Phosphite Triester Synthesis

In the 1970s, substantially more reactive P(III) derivatives of nucleosides, 3'-*O*-chlorophosphites, were successfully used for the formation of inter-nucleosidic linkages. This led to the discovery of the phosphite triester methodology. The group led by M. Caruthers took the advantage of less aggressive and more selective 1*H*-tetrazolidophosphites and implemented the method on solid phase. The use of 2-cyanoethyl phosphite-protecting group in place of a less user-friendly methyl group led to the nucleoside phosphoramidites currently used in oligonucleotide synthesis.

4.17.1. 5. Synthesis by the Phosphoramidite Method

4.17.1. 5.1. Building Blocks: Nucleoside phosphoramidites

The naturally occurring nucleotides (nucleoside-3'- or 5'-phosphates) and their phosphodiester analogs are insufficiently reactive to afford expedite synthetic preparation of oligonucleotides in high yields. The selectivity and the rate of the formation of inter-nucleosidic linkages is dramatically improved by using 3'-*O*-(*N*,*N*-diisopropyl phosphoramidite) derivatives of nucleosides called nucleoside phosphoramidites that serve as building blocks in phosphite triester methodology. To prevent undesired side reactions, all other functional groups present in nucleosides have to be rendered unreactive by attaching protecting groups. Upon the completion of the oligonucleotide chain assembly, all the protecting groups are removed to yield the desired oligonucleotides.

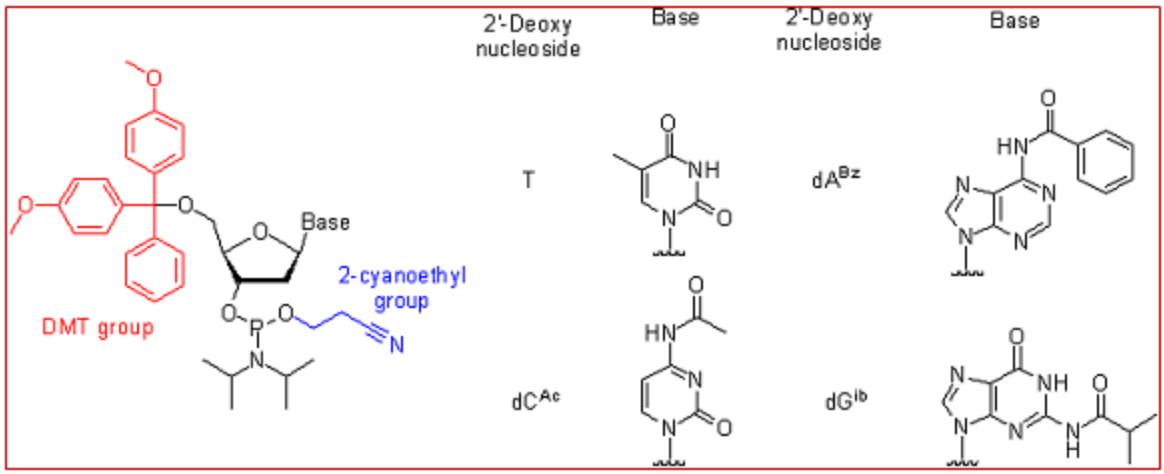


Figure 4.20: Protected 2'-deoxynucleoside phosphoramidites.

Below, the protecting groups currently used in commercially available and most common nucleoside phosphoramidite building blocks are briefly described.

- The 5'-hydroxyl group is protected by an acid-labile DMT (4,4'-dimethoxytrityl) group.'
- Thymine and uracil, nucleic bases of thymidine and uridine, respectively, do not have exocyclic amino groups and hence do not require any protection. Although the nucleic base of guanosine and 2'-deoxyguanosine have an exocyclic amino group, its basicity is low to an extent that it does not react with phosphoramidites under the conditions of the coupling reaction. However, a phosphoramidite derived from the N2-unprotected 5'-*O*-DMT-2'-deoxyguanosine is poorly soluble in acetonitrile, the solvent commonly used in oligonucleotide synthesis. In contrast, the N2-protected versions of the same compound dissolve in acetonitrile well and hence are widely used. Nucleic bases adenine and cytosine bear the exocyclic amino groups reactive with the activated phosphoramidites under the conditions of the coupling reaction. Although, at the expense of additional steps in the synthetic cycle, the oligonucleotide chain assembly may be carried out using dA and dC phosphoramidites with unprotected amino groups, most often these are kept permanently protected over the entire length of the oligonucleotide chain assembly. The protection of the exocyclic amino groups has to be orthogonal to that of the 5'-hydroxy group because the latter is removed at the end of each synthetic cycle. The simplest to implement and hence the most widely accepted is the strategy where the exocyclic amino groups bear a base-labile protection. Most often, two protection schemes are used.
- In the first, the standard and more robust scheme, Bz (benzoyl) protection is used for A, dA, C, dC, G, and dG are protected with isobutyryl group. More recently, acetyl group is often used to protect C and dC.
- In the second, mild protection scheme, A and dA are protected with isobutyryl or phenoxyacetyl groups (PAC). C and dC bear acetyl protection, and G and dG are protected with 4-isopropylphenoxyacetyl (i-Pr-PAC) or dimethylformamidino (dmf) groups. Mild protecting groups are removed more readily than the standard protecting

groups. However, the phosphoramidites bearing these groups are less stable when stored in solution.

- The phosphite group is protected by a base-labile 2-cyanoethyl group. Once a phosphoramidite has been coupled to the solid support-bound oligonucleotide and the phosphite moieties have been converted to the P(V) species, the presence of the phosphate protection is not mandatory for the successful conducting of further coupling reactions.
- In RNA synthesis, the 2'-hydroxy group is protected with TBDMS (*t*-butyldimethylsilyl) group or with TOM (*t*-butyldimethylsilyloxymethyl) group, both being removable by treatment with fluoride ion.

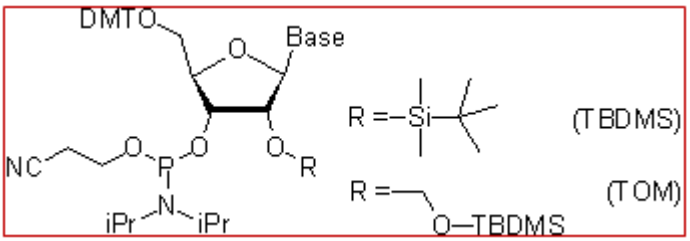


Figure 4.21: 2'-*O*-Protected ribonucleoside phosphoramidites.

- The phosphite moiety also bears a diisopropylamino (iPr2N) group reactive under acidic conditions. Upon activation, the diisopropylamino group leaves to be substituted by the 5'-hydroxy group of the support-bound oligonucleotide.

4.17.1.5.2. Non-nucleoside Phosphoramidites

Non-nucleoside phosphoramidites are the phosphoramidite reagents designed to introduce various functionalities at the termini of synthetic oligonucleotides or between nucleotide residues in the middle of the sequence. In order to be introduced inside the sequence, a non-nucleosidic modifier has to possess at least two hydroxy groups, one of which is often protected with a DMT group while the other bears the reactive phosphoramidite moiety. Non-nucleosidic phosphoramidites are used to introduce desired groups that are not available in natural nucleosides or that can be introduced more readily using simpler chemical designs. A very short selection of commercial phosphoramidite reagents is shown in Scheme for the demonstration of the available structural and functional diversity. These reagents serve for the attachment of 5'-terminal phosphate (**1**), NH2 (**2**), SH (**3**), aldehydo (**4**), and carboxylic groups (**5**), CC triple bonds (**6**), non-radioactive labels and quenchers (exemplified by 6-FAM amidite **7** for the attachment of fluorescein and dabcyl amidite **8**, respectively), hydrophilic and hydrophobic modifiers (exemplified by hexaethyleneglycol amidite **9** and cholesterol amidite **10**, respectively), and biotin amidite **11**.

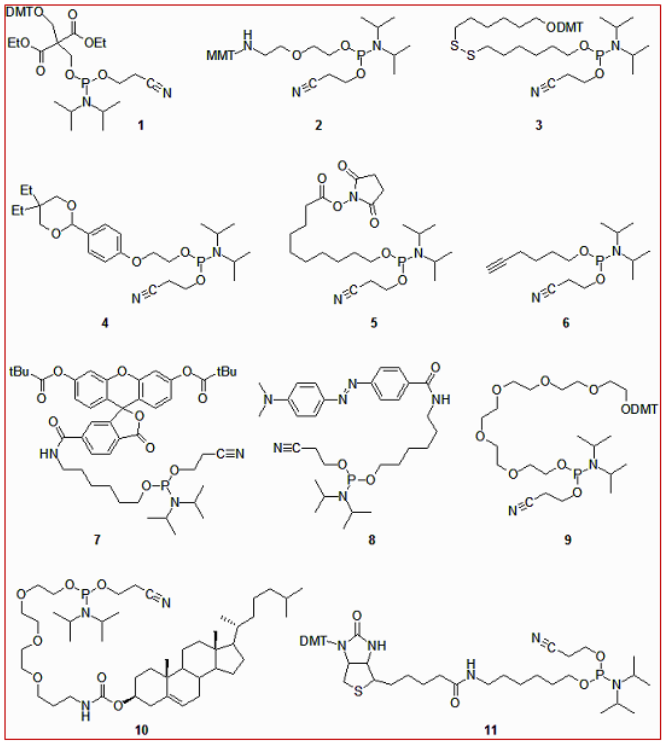
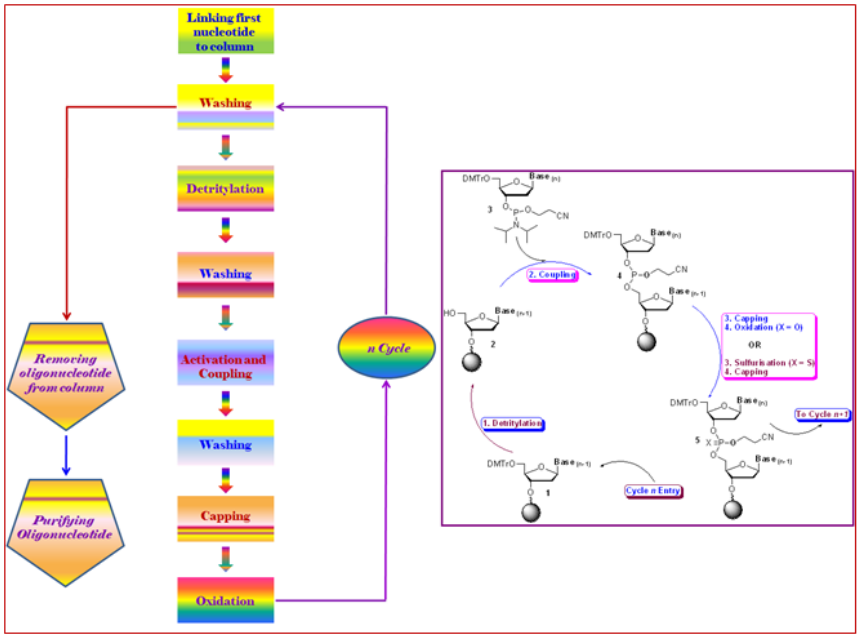


Figure 4.22: Non-nucleoside phosphoramidites for 5'-modification of synthetic oligonucleotides.

4.17.1.5.3. The Synthetic cycle



Scheme 4.6: Synthetic cycle for preparation of oligonucleotides by phosphoramidite method.

Oligonucleotide synthesis is carried out by a stepwise addition of nucleotide residues to the 5'-terminus of the growing chain until the desired sequence is assembled. Each addition is referred to as a synthetic cycle (**Scheme 4.6**) and consists of four chemical reactions:

- ***Step 1 – De-blocking (detritylation):*** The DMT group is removed with a solution of an acid, such as TCA or Dichloroacetic acid (DCA), in an inert solvent (dichloromethane or toluene) and washed out, resulting in a free 5'-terminal hydroxyl group on the solid support-bound oligonucleotide precursor.
- ***Step 2 – Coupling:*** A nucleoside phosphoramidite (or a mixture of several phosphoramidites) is activated by an acidic azole catalyst, 1*H*-tetrazole, 2-ethylthiotetrazole, 2-benzylthiotetrazole, 4,5-dicyanoimidazole, or a number of similar compounds. This mixture is brought in contact with the starting solid support (first coupling) or a support-bound oligonucleotide precursor (following couplings) whose 5'-hydroxy group reacts with the activated phosphoramidite moiety of the incoming nucleoside phosphoramidite to form a phosphite triester linkage. This reaction is very rapid and requires, on small scale, about 20 s for its completion. The phosphoramidite coupling is also highly sensitive to the presence of water and is commonly carried out in anhydrous acetonitrile. Unbound reagents and by-products are removed by washing.
- ***Step 3 – Capping:*** After the completion of the coupling reaction, a small percentage of the solid support-bound 5'-OH groups (0.1 to 1%) remains unreacted and needs to be permanently blocked from further chain elongation to prevent the formation of oligonucleotides with an internal base deletion commonly referred to as (n-1) short oligomers. This is done by acetylation of the unreacted 5'-hydroxy groups using a mixture of acetic anhydride and 1-methylimidazole or, less often, DMAP as catalysts. Excess reagents are removed by washing.
- ***Step 4 – Oxidation:*** The newly formed tricoordinated phosphite triester linkage is not natural and is of limited stability under the conditions of oligonucleotide synthesis. The treatment of the support-bound material with iodine and water in the presence of a weak base (pyridine, lutidine, or collidine) oxidizes the phosphite triester into a tetracoordinated phosphate triester, a protected precursor of the naturally occurring phosphate diester internucleosidic linkage. This step can be substituted with a sulfurization step to obtain oligonucleotide phosphorothioates. In the latter case, the sulfurization step is carried out prior to capping.

Post-synthetic processing

After the completion of the chain assembly, the solid support-bound oligonucleotide is fully protected:

- The 5'-terminal 5'-hydroxy group is protected with DMT group;
- The internucleosidic phosphate or phosphorothioate moieties are protected with 2-cyanoethyl groups;
- The exocyclic amino groups in all nucleoside bases except for T and U are protected with acyl protecting groups.

To furnish a functional oligonucleotide, all the protecting groups have to be removed. The N-acyl base protection and the 2-cyanoethyl phosphate protection may be, and is often removed

simultaneously by treatment with inorganic bases or amines. However, the applicability of this method is limited by the fact that the cleavage of 2-cyanoethyl phosphate protection gives rise to acrylonitrile as a side product. Under the strong basic conditions required for the removal of N-acyl protection, acrylonitrile is capable of alkylation of nucleic bases, primarily, at the N3-position of thymine and uracil residues to give the respective N3-(2-cyanoethyl) adducts *via* Michael reaction. The formation of these side products may be avoided by treating the solid support-boud oligonucleotides with solutions of bases in an organic solvent, for instance, with 50% triethylamine in acetonitrile or 10% diethylamine in acetonitrile. The solid support-bound oligonucleotides are deprotected using one of the two general approaches.

- Most often, 5'-DMT group is removed at the end of the oligonucleotide chain assembly. The oligonucleotides are then released from the solid phase and deprotected by treatment with aqueous ammonium hydroxide. This removes all remaining protection groups from 2'-deoxyoligonucleotides, resulting in a reaction mixture containing the desired product. The fully deprotected product is purified by desalting using ethanol precipitation, or size exclusion chromatography, or reverse-phase HPLC.
- The second approach is only used when the intended method of purification is reverse-phase HPLC. In this case, the 5'-terminal DMT group that serves as a hydrophobic handle for purification is kept on at the end of the synthesis. The oligonucleotide is deprotected under basic conditions as described above and, upon evaporation, is purified by reverse-phase HPLC. The collected material is then detritylated under aqueous acidic conditions. Finally, the product is desalted
- For some applications, additional reporter groups may be attached to an oligonucleotide using a variety of post-synthetic procedures.

Characterization

As with any other organic compound, it is prudent to characterize synthetic oligonucleotides upon their preparation. In more complex cases oligonucleotides are characterized after their deprotection and after purification. In day-by-day practice, it is sufficient to obtain the molecular mass of an oligonucleotide by recording its mass spectrum. Two methods are currently widely used for characterization of oligonucleotides: electrospray mass spectrometry (ES MS) and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF).

POSSIBLE QUESTIONS

8 MARKS

1. What are the basic steps involved in gene cloning? Explain in detail.
2. Discuss and differentiate the roles of restriction endonucleases.
3. Write in detail about the features of synthetic oligo nucleotides.
4. Explain in detail about bacteriophage and plasmid isolation and purification.
5. Give a detailed account on vectors designing.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
III BSc BIOCHEMISTRY
17BCU601A-GENETIC ENGINEERING and BIOTECHNOLOGY
MULTIPLE CHOICE QUESTIONS

Unit-I

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	Which one is specific for ssRNA?	Dnase I	Rnase H	S1 nuclease	Rnase A	RNase A
	One of the following enzyme is used for synthesizing a small stretch of DNA molecules	Reductase	DNA polymerase	Reverse transciptase	Transferase	DNA polymerase
2						
3	Which one is artificially making copy of mRNA?	rDNA	SiRNA	DNA	cDNA	cDNA
	The gene formed by the joining of DNA segments from two different sources are called	Recombinant gene	Joined gene	Both a and b	Chimeric gene	Chimeric gene
4	as					
	Who created the first rDNA molecule	Nathan, Arber and Smith	Watson, Crick and Wilkins	Boyer and Cohen	Paul Berg	Paul Berg
5						
6	One of the following enzyme is used for transcription of a gene?	DNA polymerase	Endonuclease	RNA polymerase	Transcriptase	RNA polymerase
7	The linkers are synthetic_____	Dideoxynucleotides	Oligonucleotides	Ribonucleotides	None of the above	Oligonucleotides
8	Linkers are connected with the following	DNA ligase	T4 DNA polymerase	T4 DNA ligase	T4 DNA reductase	T4 DNA ligase
9	Which one of the following not related with transcription?	Annealing	Initiation	Template recognition	Termination	Annealing
10	First region of the promoter is consisted of	Oxynucleotide	Deoxynucleotide	Ribonucleotides	Hexanucleotides	Hexanucleotides
11	Which one of the following is important regulatorv element?	Receptor	Acceptor	Promoter	Connecter	Promoter
12	Transcription is	Synthesis of DNA	Synthesis of RNA	Synthesis of rDNA	Synthesis of cDNA	Synthesis of RNA
13	Which one is not related with gene expression in regulation	Promoter gene	Integrator gene	Sensor gene	Receptor gene	Promoter gene
14	O-linked sugars or glycosylation is synthesized in	Endoplasmic reticulum	Golgi apparatus	Both the A and B	Ribosomes	Both the A and B
15	Negative regulation is	Chaperons	Hylurons	trp operon	None of the above	trp operon
16	Autoregulation is done in	Tripticase	trpR gene	Repressor gene	DNAse	trpR gene
17	DNA damage is not related to	Deamination	Thymine dimer formation	Dehydration	Depurination	Dehydration
18	Polylinker site is present in _____ vector	pBR322	Lambda Phage	pUC8	Phagemid	pUC8
19	Natural genetic engineering process is carried out by _____ bacterium	*E.coli*	*S. Cocci*	*A.tumifaciens*	*R.oryzae*	*A.tumifaciens*
	Once the two fragments are joined, the final sealing phosphodiester bond is formed by	Polymerase	Exonuclease	Ligase	Alkaline phosphatase	Ligase
20						
	Who discovered restriction enzymes	Nathan, Arber and Smith in 1970	Watson, Crick and Wilkins in 1970	Boyer and Cohen in 1975	Paul Berg in 1975	Nathan, Arber and Smith in 1970
21						
22	Which of the following polymerase enzyme would you select for cDNA synthesis?	DNA polymerase	Reverse transcriptase	Taq polymerase	Vent polymerase	Reverse transcriptase
23	How many types of DNA ligases are available?	One	Two	Three	Four	Four
24	The enzyme that produces RNA is called	Restriction endonucleases	DNA polymerase	RNA polymerase	Exonuclease	RNA polymerase
25	How many types in DNA polymerases are present in prokaryote?	One	Two	Three	Four	Two
	Which of the following statements are true regarding rDNA technology	rDNA technology is used to obtain large number of copies of specific DNA fragments	rDNA technology is used to obtain large quantities of the protein produced by the concerned gene	rDNA technology is used to integrate gene of interest into chromosomes where it expresses itself	All the above	All the above
26						
27	*Escherichia coli* DNA ligase has the molecular weight of	73kDa	72 kDa	50kDa	35kDa	73kDa
28	Taq DNA Polymerase requires ..... Cofactor for activity	Mg_{2+}	Na	Fe	Cu	Mg2+
29	Reverse transcriptase isolated from	Retrovirus	Adenovirus	Baculovirus	Gemini virus	Retrovirus
	"Naked" DNA	Is free of nucleic acids	Is free of cell	Contains just sugar-phosphate backbone	Is free of protein	Is free of cell
30						
31	Ligases are called	Molecular Scissors	Molecular glue	Molecular tool	None of the above	Molecular glue
32	EcoRI is isolated from	*Bacillus*	*Pseudomonas*	*Escherichia coli*	*vibrio*	*Escherichia coli*
33	T4 DNA Polymerase optimum temperature is	37 degree	45 degree	94 degree	50 degree	94 degree
34	Adaptors have a	3' OH and 5' P end	a short nucleotide sequence	Linker end	Homopolymer tails	3' OH and 5' P end
35	Alkaline phosphatase catalyze	Removal of phosphate at 5'terminal	Removal of3" terminal hydr	Removal of 3 'terminal phos	None of the above	Removal of phosphate at 5'terminal
36	Type II restriction enzymes cleaves DNA at	tetra nucleotide sequence	Penta nucleotide sequence	Hexa nucleotide sequence	tri nucleotide sequence	Penta nucleotide sequence
37	Smith and Nathans received nobel prize during	1973	1978	1980	1981	1978
38	Lac Z code for _____ enzyme	Alpha galactosidase	Beta galactosidase	Permease	Lactose transferase	Beta galactosidase
39	Which enzyme used to add or remove chemical groups?	Nucleases	Ligases	Polymerases	Modifying enzymes	Modifying enzymes
40	Recoginition sequence for Eco R1	AATTCCTTAA	AG/CT	TC/GA	GA/TC	GA/TC
41	What is the product produced by reverse transcriptase	single strand DNA	double strand DNA	single strand RNA	double strand RNA	single strand DNA
42	Endonucleases are enzymes that cleaves DNA at	Defined sequence	3' end of nucleotide	Internal position in random	external position	Internal position in random manner
43	Enzymes were classified by	Smith and Nathans	arber	rodriguez	Boliver	Smith and Nathans
44	Type II restriction enzymes cleaves DNA at	Defined recognition site	Random sites	25 bases away	1000 basepairs away	Defined recognition site
45	Suitable host in genetic engineering to introduce DNA fragments of donor is	*Bacillus subtilis*	*Escherechia coli*	Bacteriophages	virus	*Bacillus subtilis*
	What is the natural function of restriction enzymes?	Protecting bacteria by cleaving the DNA of infecting viruses.	Protecting bacteria by cleaving their own DNA.	Protecting bacteria by methylating their own DNA.	Protecting bacteria by methylating the DNA of infecting viruses	Protecting bacteria by cleaving the DNA of infecting viruses.
46						
47	How many types of Restriction endonucleases are present?	One	Two	Three	Four	Three
48	Term endonucleases was coined by _____	Lederberg and Meselson	Lederberg and tatum	Lederberg and yaun	Smith and Nathans	Lederberg and Meselson
49	Co-factor needed for type II restriction enzyme were	ATP and Mg 2+	S- Adenosyl Methionine	Fe	DNP+	ATP and Mg 2+
50	Restriction enzyme from Escherichia coli K 12 was first isolated from	Meselson and yaun	Lederberg and Meselson	yaun	Meselson	Meselson and yaun
	A recombinant DNA molecule is produced by	Joining of two DNA fragments	Joining of two or more DNA fragments	Both a and b	Joining of two or more DNA fragments originating from different organisms	Joining of two or more DNA fragments originating from different organisms
51						
52	Which organism produces Taq DNA polymerase	*Thermus aquaticus*	*E.coli*	*Bacillus* spp	*Pseudomonas* spp	*Thermus aquaticus*
53	*Escherichia coli* ligase seal nick at	3°C	4°C	5°C	8°C	4°C
54	Restriction enzymes cleave DNA at----------	Specific nucleotide sequence	Interior part of nucleotide se	Ends of nucleotide sequenc	middle part of sequence	Specific nucleotide sequence
55	Type I restriction enzyme have	3 subunit	5 subunit	2 subunit	1 subunit	3 subunit
56	Ligases form phosphodiester bond between	3" hydroxyl and 5" phosphate	3" hydroxyl	5" hydroxyl	3" hydroxyl and 5" triphosphate	3" hydroxyl and 5" phosphate
57	How many types of nucleases are present?	One	Two	Three	Four	Two
58	Sources of Ligases are	Bacillus	Bacteriophage	Bacteriophage infected bact	retrovirus	Bacteriophage
59	Bacteriophage ligases have molecular weight of	68 kDa	70 kDa	60 kDa	95 kDa	68 kDa
60	Linkers used for producing sticky end have	Modified blunt end	A short nucleotide sequence	Modified sticky end	Both a and c	A short nucleotide sequence with restriction site

UNIT-II
SYLLABUS

Cloning vectors for prokaryotes and eukaryotes : Plasmids and bacteriophages as vectors for gene cloning. Cloning vectors based on *E. coli* plasmids, pBR322, pUC8, pGEM3Z. Cloning vectors based on M13 and λ bacteriophage. Vectors for yeast, higher plants and animals.

DNA sequencing : DNA sequencing by Sanger's method, modifications based on Sanger's method. Automated DNA sequencing. Pyrosequencing.

Cloning vectors for prokaryotes and eukaryotes :
Plasmids and bacteriophages as vectors for gene cloning

Vectors: The Vehicle for Cloning

One of the major applications of rDNA experiments is the gene cloning or DNA cloning. Vectors are another major component required to make an rDNA molecule for gene cloning. Vectors act as a vehicle for carrying foreign DNA into a host cell for multiplication. Usually small circular DNA molecules of bacterial origin are used as cloning vectors. A DNA molecule should possess the following essential characteristics to act as a cloning vector:

Origin of Replication

It is required for autonomous replication of the plasmid using the host's replication machinery. Almost all commonly used plasmids are based on the ColE1 origin of replication (ori). Naturally occurring origins of replication are negatively regulated to keep the copy number down (typically five to ten copies per cell) to reduce the load on the host's replication machinery. While a high copy number is disadvantageous in a natural system, it is a desirable feature in a cloning vector— since the whole idea of gene cloning is to easily isolate substantial quantities of a particular DNA sequence. Modern plasmid vectors are therefore often called 'runaway replicons' and are present at 100 to 1,000 copies per cell.

Selectable Markers

Selectable markers are essential for the identification of bacteria containing recombinant plasmids. Selection can be divided into two types—**positive selection** and **negative selection**.

- **Positive selection** is used to identify bacteria that contain plasmids. The most common markers used for positive selection are the antibiotic resistance genes carried by the original R factors. While many antibiotics and resistance genes are available, the commonly used ones fall into two general classes: **Antibiotics affecting cell wall synthesis** and **Antibiotics affecting translation.** Ampicillin is a beta-lactam-based antibiotic that acts by inhibiting the synthesis of the bacterial peptidoglycan cell wall. Sensitive bacteria are not actively 'killed,' but on cell division are unable to synthesize the cell wall and suffer from osmotic lysis. The enzyme beta-lactamase is secreted into the periplasmic space where it breaks down the antibiotic, allowing cell wall synthesis to

proceed. The antibiotics tetracycline, kanamycin, and chloramphenicol all act by inhibiting translation. The covalent modification (phosphorylation, acetylation) of these antibiotics blocks their interaction with the translation apparatus. Positive selection is particularly important when introducing plasmids into bacteria by transformation. At best, only about 1 in 10,000 bacteria picks up a plasmid that carries the antibiotic resistance. A strong positive selection system is essential to eliminate the 9,999 bacteria that did not pick up a plasmid from the one that did. By plating the transformation products directly on antibiotic plates, all untransformed bacteria die and only those containing the plasmid (and antibiotic-resistance marker) grow to form colonies.

- A second selection system is necessary to distinguish between plasmids that are merely recircularized from those that carry a foreign DNA insert. Usually, a negative selection method is used. In order to identify those plasmids carrying a foreign DNA fragment, the site of insertion is chosen such that the insertion disrupts a selectable marker—a phenomenon known as **insertional inactivation**. It can be the insertional inactivation of an antibiotic resistance gene or enzymes such as β-galactosidase, a product of *Lac* Z gene of lac operon. The insertion of the foreignDNA into the vector will disrupt the expression of these genes and will produce a color difference for the colony from that of cells with intact vectors.

Multiple Cloning Sites (MCS) or Polylinker

A vector should have a site specific for cloning the foreign DNA fragment provided with one restriction site for most of the commonly used unique restriction endonucleases. All these unique restriction sites are grouped together in a small region of the vector known as the **multiple cloning site (MCS)** or the **polylinker**. The presence of unique restriction sites at the MCS gives flexibility in the choice of restriction enzymes.

Small Size

Relatively small vectors are more desirable because they increase the transformation efficiency and are easy to manipulate. Small size also helps purification procedures to obtain intact plasmids. A large number of vectors with the above characteristics have been developed.

There are six different types of cloning vectors commonly used in recombinant DNA experiments. They are the following:

- Plasmid-cloning vectors
- Bacteriophage cloning vectors
- Cosmid-cloning vectors
- Yeast artificial chromosomes (YACs)
- Bacterial artificial chromosomes (BACs)
- Animal and plant vectors (Shuttle vectors)

Cloning vectors based on *E. coli* plasmids, pBR322, pUC8, pGEM3Z

Plasmids are extra chromosomal circular double stranded DNA. These vectors are derived from bacterial plasmids and are the most widely used, versatile, and easily manipulated ones. Plasmid are used as cloning vectors, include ori site or origin of replication, needed for replication in the bacterial cells. For an example E.coli plasmid vector, should have following features:

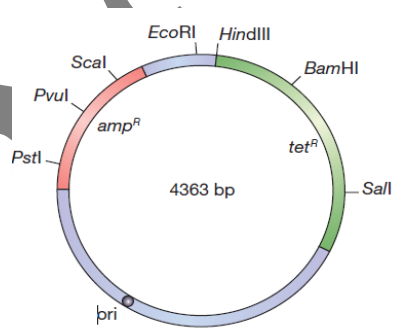
a. Ori site for replication.
b. Selectable marker genes, such as antibiotic resistance gene
c. Unique restriction sites, so that the restriction enzymes can be used to cut the plasmid and DNA of interest can be inserted into the plasmid.

ColE1 of *E. coli* is an example of a naturally occurring plasmid. The *ori* in almost all plasmid vectors is that of ColE1. The following are some of the examples of plasmid-cloning vectors.

a)pBR322 Vectors

The name "pBR322" conforms with the standard rules for vector nomenclature:
"p" indicates that this is indeed a plasmid.
"BR" identifies the laboratory in which the vector was originally constructed
(BR stands for Bolivar and Rodriguez, the two researchers who developed pBR322).
"322" distinguishes this plasmid from others developed in the same laboratory
(there are also plasmids called pBR325, pBR327, pBR328, etc.).

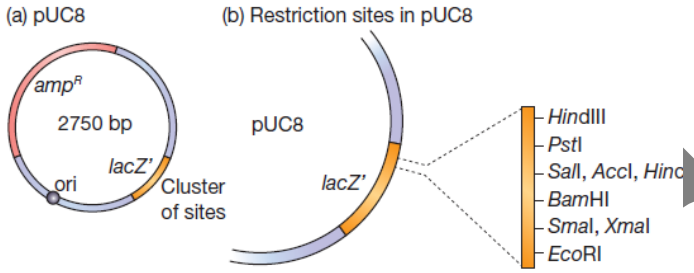
This was the first widely used, purpose built plasmid vector. It has a number of useful features:
Origin of replication. It carries a fragment of the plasmid pMB1 that acts as an origin for DNA replication and thus ensures multiplication of the vector.
Size. It is relatively small at 4,363 bp. This is important because transformation efficiency is inversely proportional to size and above 10 kbp is very low. Thus, there is 'room' in pBR322 for an insert of at least six kbp.
Copy number. Reasonably high copy number (~15 copies per cell), which can be increased 200-fold by treatment with a protein-synthesis inhibitor—chloramphenicol amplification.
Selectable marker. It carries two antibiotic resistance genes—ampicillin and tetracycline.
Cloning sites. It carries a number of unique restriction sites. Some of these are located in one of the antibiotic resistance genes (e.g., sites for *Pst* I, *Pvu* I, and *Sac* I are found in *Ampr* and *Bam*HI and *Hind* III in *Tetr*). Cloning into one of these sites inactivates the gene allowing recombinants to be differentiated from non-recombinants known as **insertional inactivation.**



Gene map of pBR322

b)pUC Series Vectors

Another popular plasmid vector is the pUC series, which is extensively used as cloning as well as expression vector. These vectors have three important additional features compared to pBR322.



pUC8 vector

High copy number. A mutation within the origin of replication produces 500to 600 copies of the plasmid per cell without amplification.

Blue-white screening. This is a special form of insertional activation that can be used during the primary selection of transformants, rather than requiring a second round of screening. It utilizes the N-terminal portion of the *E. coli* beta-galactosidase—the product of lac Z gene (alpha-peptide) encoded by the vector in a form of intermolecular complementation—that restores betagalactosidase activity to a defective enzyme (omega-peptide) encoded by the host. If a chromogenic substrate (X-gal) and a beta-galactosidase inducer (IPTG) are included in the plates on which the primary transformants are selected, non-recombinant molecules will catabolize the colorless substrate to give blue colonies.

A synthetic polylinker. This is a piece of manmade DNA that contains several unique restriction sites. It has been inserted within the portion of the vector encoding the beta-galactosidase alpha-peptide in such a way that it does not affect its expression. However, inserting a foreign DNA fragment into any one of the polylinker restriction sites almost invariably disrupts the enzyme activity. Thus, recombinant colonies remain white but non-recombinants turn blue.

Expression Vectors

In most of the cases the main purpose of the rDNA experiment and cloning is to sufficiently multiply or amplify the inserted DNA fragment. But sometimes the aim of the process will be to produce large quantities of protein encoded by the inserted gene. This can be accomplished by incorporating the necessary regulatory elements along with the gene in the vector. Such vectors should be provided with signals necessary for initiation and termination of transcription such as suitable promoters and terminator sequences; and signals for translation initiation such as a start codon and a ribosome binding site into the vector upstream to the multiple cloning site (MCS). These vectors, which have the regulatory elements and other machinery for the expression of the cloned gene, can be used for the production of recombinant proteins and are known as **expression vectors**. Vectors of the pUC series are example of expression vectors.

pUC19 is an example for an E.coli cloning vector. pUC19 is basically 2,686bp in length. The features of this pUC19 plasmids are as follows:
a. High copy number, as high as nearly hundred copies per bacterial cell. This helps in getting good yield of cloned DNA in short duration.
b. It also has got selectable marker as ampicillin resistance gene.
c. This plasmid also got a cluster of unique restriction sites known as polylinkers or in other words this plasmid has got multiple cloning sites.
d. The polylinkers or multiple cloning sites are also a part of lacZ (galactosidase) gene. That is pUC19 plasmid will complement a β lacZ E.coli allowing it to become lacZ positive. When DNA of interest is cloned into the polylinkers lacZ is disrupted, this prevents the complementation from occurring.
e. A chromogenic analog of lactose known as X-gal, turns blue in presence of galactosidase, and remains white in β absence; therefore blue white colonies can be used as screening method to identify the bacterial colonies which contain recombinant plasmid.

DNA of interest is inserted into a cloning vector using restriction enzyme and then ligating it with enzyme ligase.
a. pUC19 plasmid vector can be cut using restriction enzyme that has got a unique site in the polylinker region.
b. DNA of interest is also cut using the same restriction enzyme.
c. DNA of interest and pUC19 DNA are mixed and allowed for random joining of fragments.
d. Resulting plasmids then transformed into E.coli cells either by chemical treatment method or by electroporation method. Then these bacterial cells are grown on media contain ampicillin and Xgal.
e. Ampicillin resistance is resulted from pUC19 sequence, Blue colonies are resulted from the rejoined plasmids, and white colonies will result from transformed bacterial cells with recombinant plasmids.
Many different types of cloning vectors are designed in the laboratory and they are commercially available for use. These vectors have different array of unique restriction sites in the polylinker site. Many plasmid cloning vectors are designed and available for many prokaryotic organisms as well as for eukaryotic organisms. But the plasmid cloning vectors which carry more than 5 to 10 kb gene of interest are often unstable. Therefore size of the gene of interest is limited in plasmid cloning vectors.
pGEM-3Z♣pGEM-3Z carries a gene for ampicillin resistance (bla)-
so transformed cells can be selected by growth(with gene or not) on an ampicillin plate

Cloning vectors based on M13 and λ bacteriophage

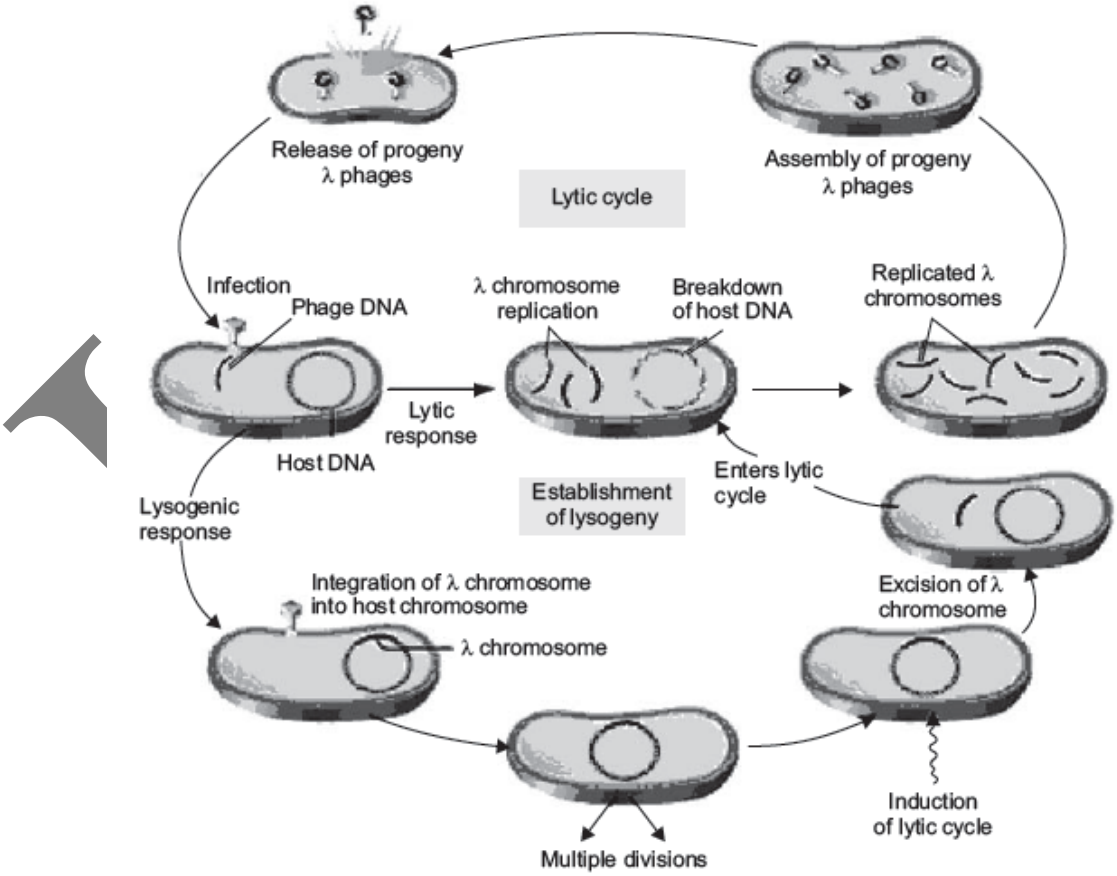
Bacteriophage - basic features, life cycle

These are derived from viruses that specifically infect bacteria. Such viruses usually contain a comparatively small DNA genome surrounded by a protein coat. The viruses that infect bacteria are called **bacteriophages**. Those phages, which most frequently infect *E. coli*, are used

for cloning purposes. Bacteriophages infect bacterial cells by injecting their DNA into the host cytoplasm. This DNA in the host cell selectively replicates and expresses the proteins required for the assembly of new phage particles. These processes result in the production of a large number of phages, which break the cells with a process known as the **lytic cycle**. These new viral particles will re-infect the neighboring cells and the cycle continues. In some cases the phage DNA, after entering into the host cell, is integrated along with the host genome and is separated at a later stage and starts the lytic cycle as explained above. The lytic cycle and phage multiplication are shown in Figure.

The natural way of transferring the viral DNA into a specific bacterial host has attracted scientists and they have modified these viral genome DNA to use as vectors for gene cloning. Bacteriophage lambda (λphage) and M13 are extensively modified for the development of a phage-based cloning vector.

There are many versions or types of bacteriopage λ used as cloning vectors. For example sequence of the gene responsible for the lysogeny is removed, so that only one type that is lytic infection is possible in this bacteriophage λ vector. The central region of the λ chromosome is used for inserting the gene of interest. Gene of interest is inserted using restriction digestion and also DNA ligation method.



The bacteriophage infection and life cycle—lysogeny and lytic cycle.

When this ligated DNA is mixed with phage proteins, this leads to the assembling of phage head and DNA is also packed to form virus particles. Only viral particles with both the phage chromosome arms and also with proper central segment (3752kb) are able to replicate by infecting the E.coli bacterial cell. Progeny phages contain DNA of interest; hence provide a large quantity of DNA of interest. Many types of phage cloning vectors are available with varying features like an expanded array of restriction sites and much more.

a)ƛ Phage-cloning Vectors

These are the engineered version of Lambda bacteriophage that infects *E. coli*. This vector has a double-stranded linear DNA genome having a size of 48,514 bp (48.5 kb). About 12 bases at the ends of this DNA molecule are unpaired and complementary and therefore are sticky or cohesive and known as **cos sites (cohesive end sites).** These cohesive sites are very essential for packaging the DNA into the viral particle during the lytic cycle. But a major portion of the central region of the viral genome DNA is not required for the infection and lytic cycle in *E. coli* cells. The ƛphage vector is designed in such a way that the central region of the ƛchromosome (linear) is cut with a restriction enzyme and is replaced with the foreign DNA digested with the same restriction enzyme. This recombinant DNA is packaged in phage heads to form virus particles (Figure). Phages with both ends of the ƛ chromosome and a 37 to 52 kb insert replicate by infecting *e. coli*. Phages replicate using *E. coli* and the lytic cycle (Figure) and produce large quantities of 37 to 52 kb cloned DNA. A large number of unique restriction sites are also available here.

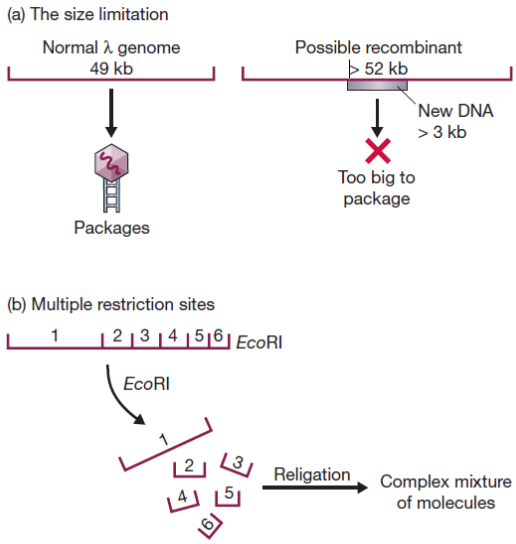


Fig: The two problems that had to be solved before cloning vectors could be developed. (a) The size limitation placed on the ƛ genome by the need to package it into the phage head. (b) ƛ DNA has multiple recognition sites for almost all restriction endonucleases.

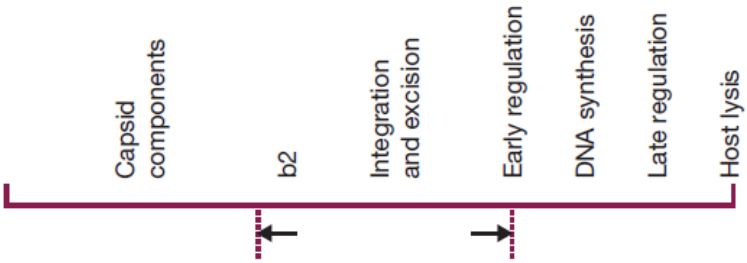


Fig: The λ genetic map, showing the position of the main non-essential region that can be deleted without affecting the ability of the phage to follow the lytic infection cycle. There are other, much shorter non-essential regions in other parts of the genome.

Insertion and replacement vectors for bacteriophages

Once the problems posed by packaging constraints and by the multiple restriction sites had been solved, the way was open for the development of different types of e-based cloning vectors. The first two classes of vector to be produced were λ **insertion** and λ**replacement** (or substitution) vectors.

a)Insertion vectors

With an insertion vector (Figure a), a large segment of the non-essential region has been deleted, and the two arms ligated together. An insertion vector possesses at least one unique restriction site into which new DNA can be inserted. The size of the DNA fragment that an individual vector can carry depends, of course, on the extent to which the non-essential region has been deleted. Two popular insertion vectors are:

λgt10 (Figure 6.12b), which can carry up to 8 kb of new DNA, inserted into a unique *Eco*RI site located in the *c*I gene. Insertional inactivation of this gene means that recombinants are distinguished as clear rather than turbid plaques .

λZAPII (Figure 6.12c), with which insertion of up to 10 kb DNA into any of 6 restriction sites within a polylinker inactivates the *lacZ*2 gene carried by the vector. Recombinants give clear rather than blue plaques on X-gal agar.

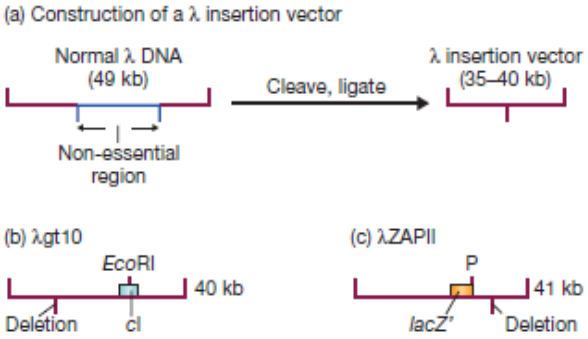


Fig: λ insertion vectors. P = polylinker in the *lacZ*2 gene of ⎣ZAPII, containing unique restriction sites for *Sac*I, *Not*I, *Xba*I, *Spe*I, *Eco*RI, and *Xho*I

b) Replacement vectors

A e replacement vector has two recognition sites for the restriction endonuclease used for cloning. These sites flank a segment of DNA that is replaced by the DNA to be cloned (Figure). Often the replaceable fragment (or **"stuffer fragment"** in cloning jargon) carries additional restriction sites that can be used to cut it up into small pieces, so that its own re-insertion during a cloning experiment is very unlikely. Replacement vectors are generally designed to carry larger pieces of DNA than insertion vectors can handle. Recombinant selection is often on the basis of size, with non-recombinant vectors being too small to be packaged into e phage heads

An example of a replacement vectors is:

ƛEMBL4 (Figure b) can carry up to 20 kb of inserted DNA by replacing a segment flanked by pairs of *Eco*RI, *Bam*HI, and *Sal*I sites. Any of these three restriction endonucleases can be used to remove the stuffer fragment, so DNA fragments with a variety of sticky ends can be cloned. Recombinant selection with eEMBL4 can be on the basis of size, or can utilize the Spi phenotype .

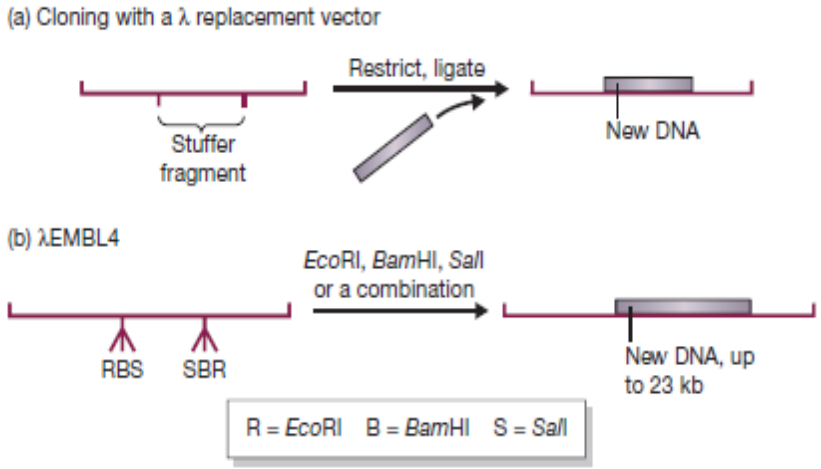


Fig: λ replacement vectors. (a) Cloning with a λ replacement vector. (b) Cloning with λ EMBL4.

b)M13-based Cloning Vectors

The M13 family of vectors is derived from bacteriophage M13. This is a malespecific (infects *e. coli* having *f. pili*), lysogenic filamentous phage with a **circular single-stranded DNA** genome about 6,407 bp (6.4 kb) in length. On infection, this molecule is transferred to *E coli* and converted into the double-stranded replicative form (RF). The replication continues and when there are more than 100 copies of DNA in the cell, the DNA replications become asymmetric and produce copies of the original single-stranded molecule, which are packaged into infective particles and extruded from the cell. The host is never lysed but continues to grow throughout the infection, although at a significantly reduced rate. This difference in growth rate between infected and uninfected cells produces characteristic 'plaques' when M13 phages are plated on a suitable host.

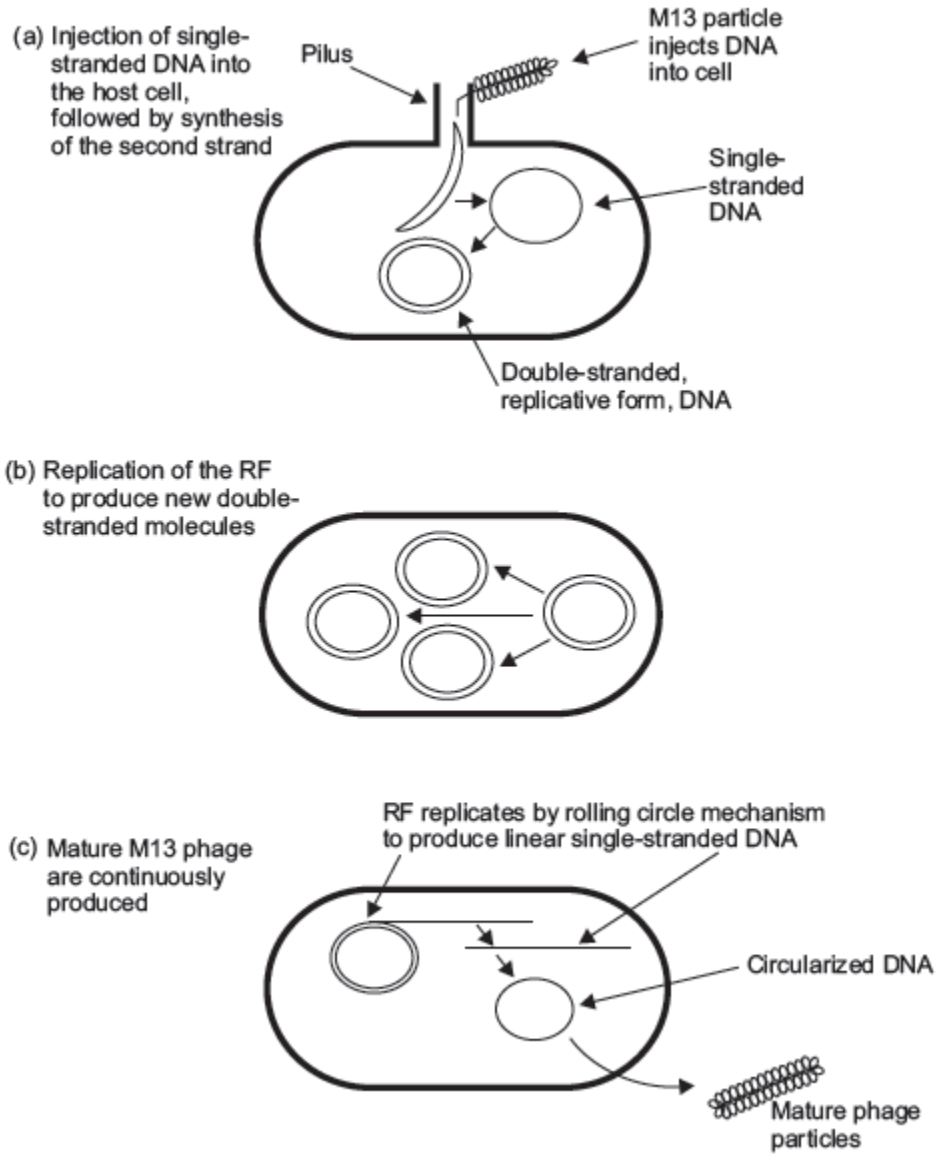


FIGURE Infection and life cycle of M13 phage.

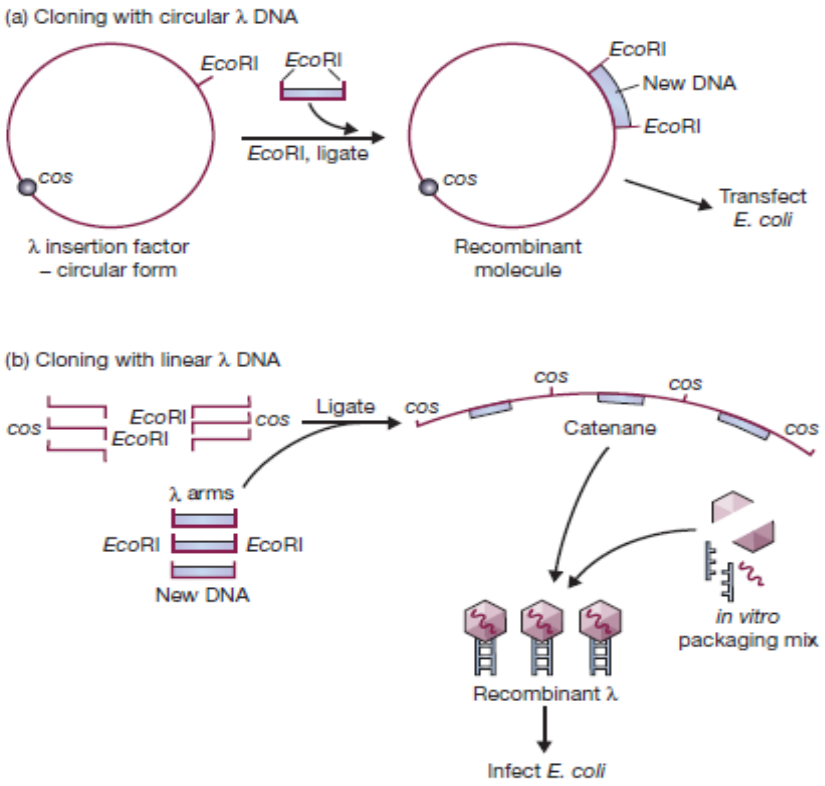
The advantages of M13-based vectors are that they contain the same polylinker and alpha-peptide fragments as the pUC series and recombinants can be selected by the blue → white color test. The RF form of M13 vectors can be isolated by standard plasmid DNA-preparation procedures, and foreign DNA can be inserted into them as if they were conventional plasmids. The size of the genome is below 10 kb and so is easy to handle.

The specific use of M13 is as an aid to DNA sequencing. Once cloned into M13, large amounts of the single-stranded form of any given fragment can be easily isolated from the mature phage that is extruded from infected cells. This is an ideal template for a dideoxy-sequencing reaction.

If an oligonucleotide complementary to the region just downstream of the polylinker is used as the primer for the dideoxy reaction, the cloned DNA fragment can be sequenced. With this method it is possible to determine the sequence of any DNA fragment that has been cloned into M13.

Cloning experiments with insertion or replacement vectors

A cloning experiment with a e vector can proceed along the same lines as with a plasmid vector—the e molecules are restricted, new DNA is added, the mixture is ligated, and the resulting molecules used to transfect a competent *E. coli* host (Figure). This type of experiment requires that the vector be in its circular form, with the *cos* sites hydrogen bonded to each other. Although satisfactory for many purposes, a procedure based on transfection is not particularly efficient. A greater number of recombinants will be obtained if one or two refinements are introduced. The first is to use the linear form of the vector. When the linear form of the vector is digested with the relevant restriction endonuclease, the left and right arms are released as separate fragments. A recombinant molecule can be constructed by mixing together the DNA to be cloned with the vector arms (Figure 6.14b). Ligation results in several molecular arrangements, including catenanes comprising left arm–DNA–right arm repeated many times (Figure 6.14b). If the inserted DNA is the correct size, then the *cos* sites that separate these structures will be the right distance apart for *in vitro* packaging (p. 81). Recombinant phage are therefore produced in the test tube and canbe used to infect an *E. coli* culture. This strategy, in particular the use of *in vitro* packaging, results in a large number of recombinant plaques.



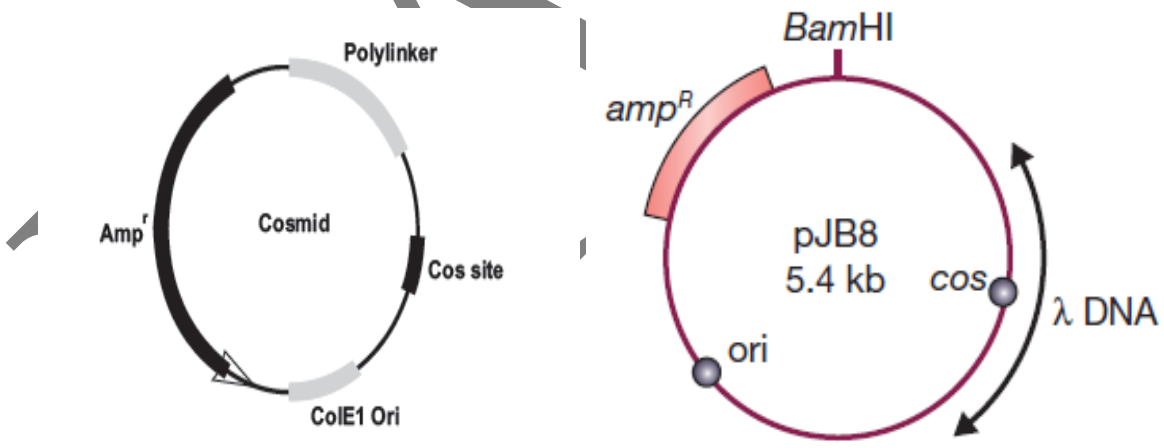
Different strategies for cloning with a λ vector. (a) Using the circular form of λ as a plasmid. (b) Using left and right arms of the λ genome, plus *in vitro* packaging, to achieve a greater number of recombinant plaques

Vectors for yeast, higher plants and animals

Cosmids

Cosmid-cloning vectors were among the first large insert cloning vehicles developed. They were constructed by certain features of plasmid and the *cos* site (which allows the target DNA to be inserted) of λ phage. A cosmid can be defined as a plasmid that contains a *cos* site from the lambda phage genome. The vector replicates as a plasmid (it contains a ColE1 origin of replication), and uses AmpR for positive selection and employs lambda phage packaging to select for recombinant plasmids carrying foreign DNA inserts up to 45 kb in size.

The simplest cosmid vector has a ColE1 origin of replication, selectable markers including the antibiotic-resistance gene and β-galactosidase gene (a part of *lac Z* gene), and suitable polylinker sites and lambda cos site (Figure). Ligation of the cosmid vector and foreign DNA fragments of sizes **upto 45 kb** is similar to ligation into a lambda substitution vector. The desired ligation product is a concatemer of 45 kb foreign DNA fragment and five kb cosmid-vector sequences. This concatemer is then packaged into viral particles (remember, packaging is cos site to cos site) and these are used to infect *E. coli* where the cosmid vector replicates using the ColE1 origin of replication (*ori*). Phage packaging serves only to select for recombinant molecules and to transfer these long DNA molecules (50 kb total) into the bacterial host (50 kb fragments transform very inefficiently while phage infection is very efficient).



A simplest form of cosmid

The following are some advantages of cosmids:

- High transformation efficiency.
- The cosmid vector can carry up to 45 kb whereas plasmid and phage vectors are limited to 25 kb.

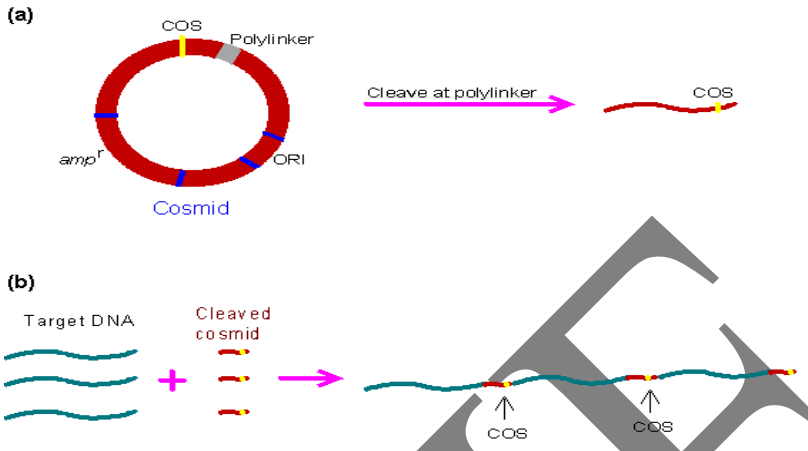
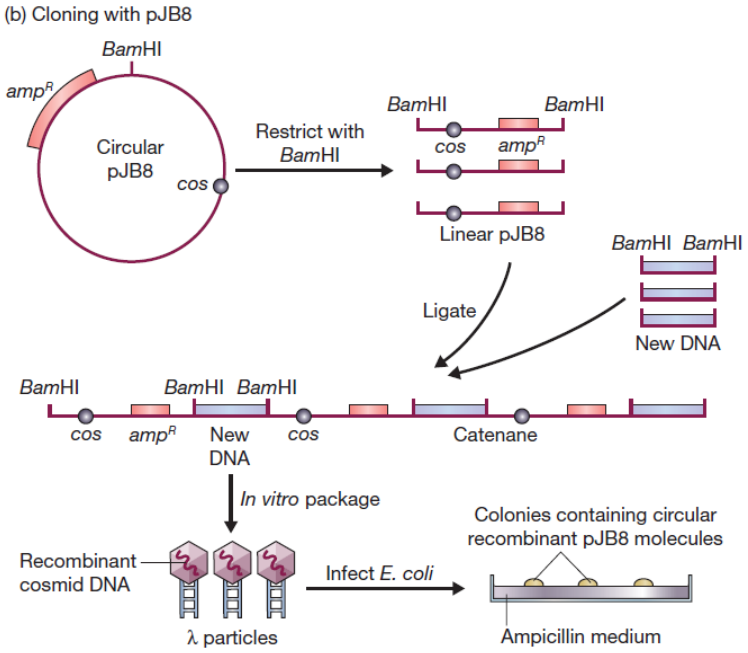


Figure. Cloning by using cosmid vectors. (a) In addition to *amp*^{r}, ORI, and polylinker as in the plasmid vector, the cosmid vector also contains a COS site. (b) After cosmid vectors are cleaved with restriction enzyme, they are ligated with DNA fragments. The subsequent assembly and transformation steps are the same as cloning with phages

A cloning experiment with a cosmid is carried out as follows

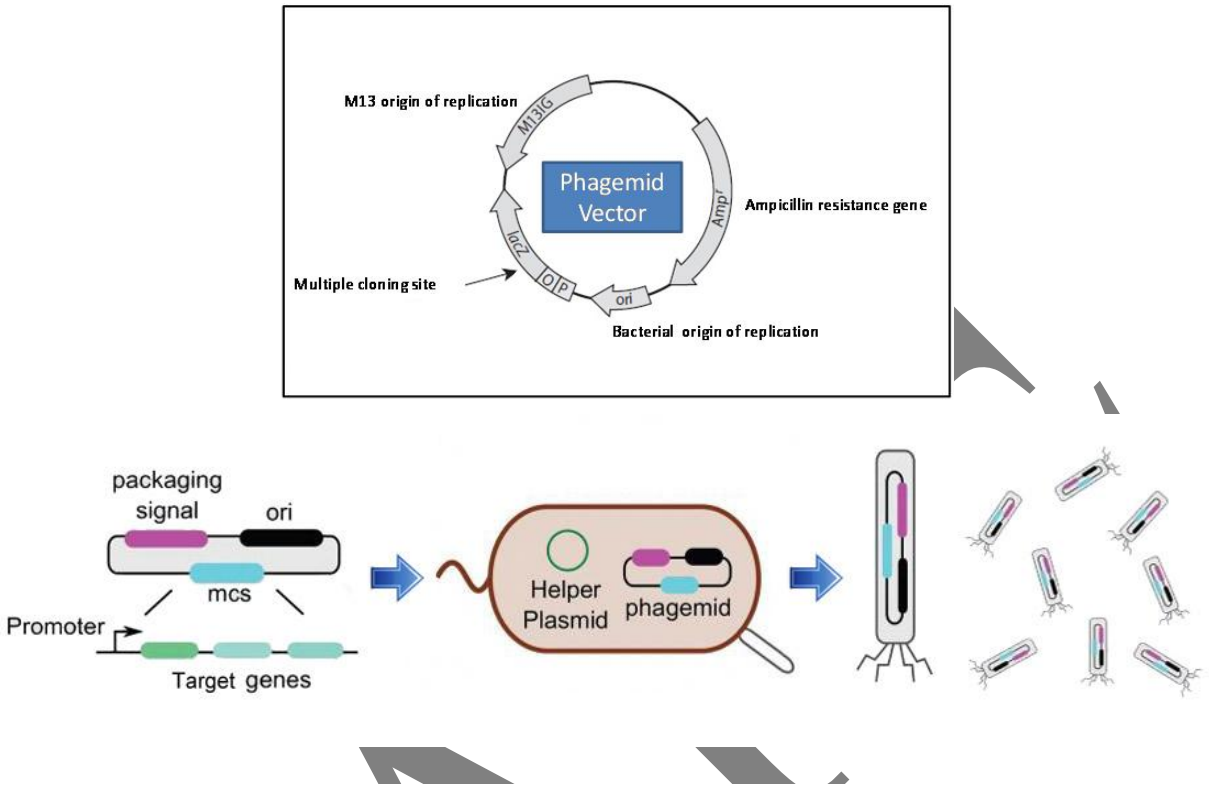
The cosmid is opened at its unique restriction site and new DNA fragments inserted. These fragments are usually produced by partial digestion with a restriction endonuclease, as total digestion almost invariably results in agments that are too small to be cloned with a cosmid. Ligation is carried out so that catenanes are formed. Providing the inserted DNA is the right size, *in vitro* packaging cleaves the *cos* sites and places the recombinant cosmids in mature phage particles. These e phage are then used to infect an *E. coli* culture, though of course plaques are not formed. Instead, infected cells are plated onto a selective medium and antibiotic-resistant colonies are grown. All colonies are recombinants, as non-recombinant linear cosmids are too small to be packaged into e heads.



A typical cosmid and the way it is used to clone long fragments of DNA

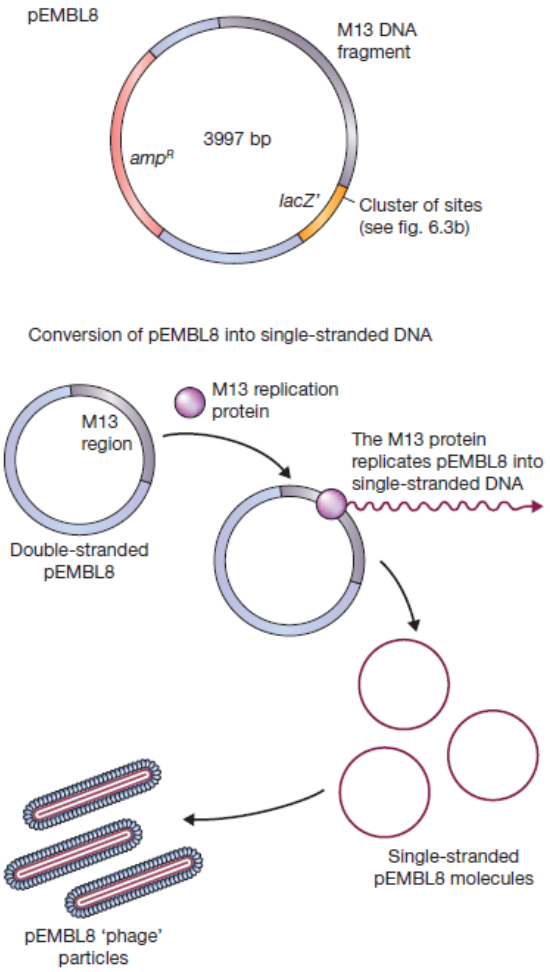
Phagemids

Phagemid vectors are plasmids which have been artificially manipulated so as to contain a small segment of the genome of a filamentous phage, such as M13, fd or f1. The selected phage sequences contain all the cis-acting elements required for DNA replication and assembly into phage particles. They permit successful cloning of inserts several kilobases long (unlike M13 vectors in which such inserts tend to be unstable). They can be replicated in the host cells like plasmids maintaining double stranded replication and high copy number so that large amount of foreign DNA can be recovered. They can also package their single stranded DNA in phage particles because of the presence of M13 origin of replication. Following, transformation or electroporation of a suitable E. coli strain with a recombinant phagemid, the bacterial cells are superinfected with a filamentous helper phage, such as f1, which is required to provide the coat protein. Phage particles secreted from the superinfected cells will be a mixture of helper phage and recombinant phagemids. The mixed single-stranded DNA population can be used directly for DNA sequencing because the primer for initiating DNA strand synthesis is designed to bind specifically to a sequence of the phagemid vector adjacent to the cloning site. A positive selection marker is used to select bacteria containing the phagemid. Fragments of several kilobases of DNA in length can be isolated in single stranded form from the phagemids. They are also utilized for sequencing purposes and for generating templates for site-directed mutagenesis. Commonly used phagemid vectors include the pEMBL series of plasmids and the pBluescript family. λZAP is also an example of phagemid



pEMBL8

pEMBL8 was made by transferring into pUC8 a 1300 bp fragment of the M13 genome. This piece of M13 DNA contains the signal sequence recognized by the enzymes that convert the normal double-stranded M13 molecule into single-stranded DNA before secretion of new phage particles. This signal sequence is still functional even though detached from the rest of the M13 genome, so pEMBL8 molecules are also converted into single-stranded DNA and secreted as defective phage particles. All that is necessary is that the *E. coli* cells used as hosts for a pEMBL8 cloning experiment are subsequently infected with normal M13 to act as a helper phage, providing the necessary replicative enzymes and phage coat proteins. pEMBL8, being derived from pUC8, has the polylinker cloning sites within the *lacZ′* gene, so recombinant plaques can be identified in the standard way on agar containing X-gal. With pEMBL8, single-stranded versions of cloned DNA fragments up to 10 kb in length can be obtained, greatly extending the range of the M13 cloning system.



Viral vectors and plant vectors

- Exploitation of plant viruses as transformation vectors by massive infection may be harmful and even deleterious to the target plants. It is still however able to express and produce foreign proteins.
- Plant viruses must exhibit some of the exemplary features before they are considered as vectors. They should extend their broader host-range, spread of seed transmission and carry additional copies of gene of interest.
- Several viral vectors require suitable modification in order to accommodate extra nucleic acid and also aggressive in infection process. Although several groups of viruses have been identified, some moderate progresses have been made only in two groups. These two groups are Caulimo virus and Gemini virus, which have DNA genome as genetic material.

Cauliflower Mosaic Virus (Caulimovirus):

Cauliflower mosaic virus (CamV) belongs to the group caulimovirus, can be used as potential candidate to deliver foreign gene into the plant. It is perhaps the best studied viruses

among plant virus, which infects several members belonging to Cruciferae family. Cauliflower mosaic virus contains circular double helical DNA as genetic material.

As an infective agent, can cause disease in wide range of commercially important cultivated crops. Cauliflower mosaic DNA has been subjected to a wide range of manipulation. This was the only and first virus to be manipulated and used as a favourable choice for genetic engineering work. Elucidation of 8 kb CamV reveals that, it contains six major and two minor reading frames (Fig. 14.12).

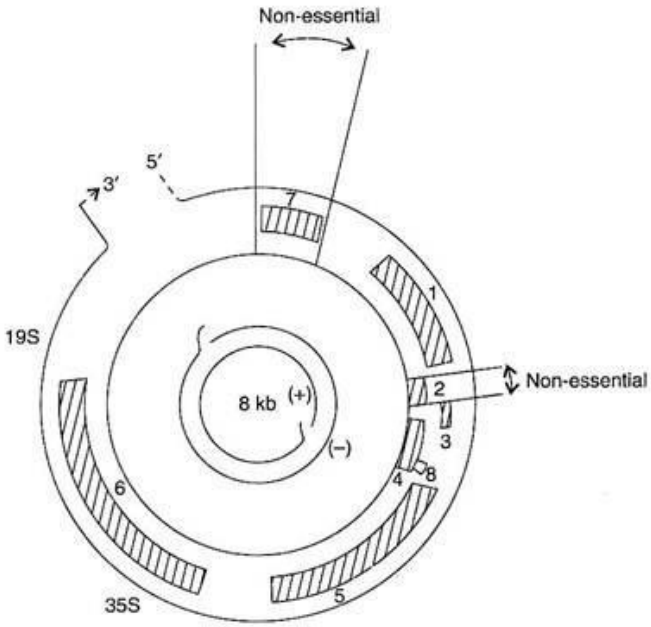


Fig. 14.12. Genetic map of cauliflower mosaic virus.

Presence of ribonucleotide in DNA leads to the conclusion that CamV replication involves the synthesis of negative DNA strand by employing reverse transcriptase and followed by synthesis of positive DNA strand. Once the synthesis of double strand DNA completes, it is then packed into viral particles and continue the cycles of transcription and translation.

CamV Vector:

Cauliflower mosaic virus can be used as a potential vector due to the infective nature of its genetic material. This could be proved by applying viruses on the leaf rubbed with abrasive material. The CamV cannot accommodate foreign DNA, if the size exceeds its normal size. The inserted DNA may destabilize infectious nature of the virus.

Other constraints are the packaging of genome and limitation of the insertion of foreign DNA. Despite the marginal constraints, CamV genome can be packed in nucleosome and is able to undergo transcription by plant RNA polymerase II. The genome of CamV consists of six major and two minor open reading frames (ORF), in tightly packed arrangements.

The two ORF regions, one (ORFII) codes for insect transmission factor and other (PRF VII) with unknown functions can be replaced with gene of interest.

Gemini Virus:

It is a DNA virus, known to infect wide range of economically important and agriculture crops of monocotyledonous and dicotyledonous plants. Several diseases such as maize streak virus and curly top virus are caused by Gemini virus. The genome is single Stranded Circular DNA and its replication takes place by DNA immediately.

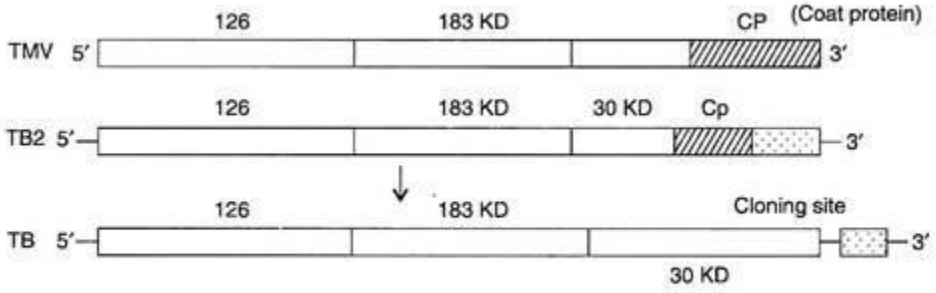
Tobacco Mosaic Virus Based Expression Vector:

Tobacco mosaic virus (TMV) is a RNA virus and shows several advantages by designing expression vector. TMV was the first virus to be purified and sequenced. As far as biohazard is considered TMV could be used as a comparatively safe recombinant virus in the field.

The coat protein of the TMV is one of the most accumulated proteins in plants reachable upto 10% of the dry weight of infected plant. Approximately, under ideal condition, 2000 kg tobacco protein can be produced per acre per year. Moreover, TMV can be purified in crystalline form in substantial quantity by simple methodology.

The single stranded RNA genetic material of TMV encloses 6300 nucleotides with four open reading frames (Fig. 14.13). The filamentous nature of the RNA virus is determined by the length of the viral nucleic acid. Both 183 kD read through protein and 126 kD coat proteins are translated from the 5′ end of the genomic RNA.

These two proteins form replicase complex. In addition to these two proteins, the 30-kD movement protein and 17.5-kD capsid proteins are translated at 3′ region of sub-genomic mRNA during replication.



14.13 Genomic organization of TMV vector (2 and 3)

TMV can be subjected to a wide range of manipulation by replacing the viral coat protein with a foreign protein, for example, replacement of coat protein with reporter gene chloramphenical acetyl transferase (CAT) resulted in a free-RNA virus that generated high CAT activity. Improvisation of TMV vector was achieved by placing CAT gene under the control of a coat protein sub-genomic promoter of TMV into the entire TMV genome.

This recombinant virus produced third sub-genomic mRNA and this was translated into CAT enzyme with high activity. Similarly, another TMV hybrid expression vector TB_2 was designed in which coat protein gene and neomycin phosphotransferase marker gene was placed under the control of sub-genomic promoters of TMV coat protein and ORSV, respectively.

This was referred as extra gene vector. TB_2 effectively produced the foreign protein without any major constraints. Another extra gene in TMV based vector, 4GD-PL, was

developed from tomato green mosaic virus. The 4GD-PL vector was able to express foreign proteins systematically throughout plants.

All these studies demonstrated that proximity of genes to the 3′-untranslated region of the genome increases efficiency of their translation. Another improved TMV vector, 30B, was designed in which the start codon (AUG) of the capsid protein was mutated to AGA, and restriction cloning sites were engineered (40 nucleotide) to provide a full-size sub-genomic RNA promoter.

Possibility of satellite RNA to be used as vector has been considered. They vary in their size between 0.27 and 1.6 kb. They are not indispensable for virus replication. However, their functions can alter pathogenecity of virus.

Cow Pea Mosaic Virus Expression Vector:

Cow pea mosaic virus (CpmV) is also a RNA virus and infects species of legumes. There are two separate positive strand-RNA molecules present in the genetic material of CpmV. The number of nucleotides present in the RNA I and RNA II strand is 5889 and 3480, respectively. Although RNA I alone can replicate on its own but both RNAs are indispensable for infectivity.
The proteins involved in the replication of the virus are encoded by RNA I whereas movement proteins are encoded by RNA II. CpmV capside of both large (L) and small (S) coat protein of 30 copies each are in isohedral symmetry. The two capsid proteins are folded into three antiparallel β-barrel structures.

In the construction of CpmV expression vector, preference was given to the replacement of stable chimeras by insertion of foreign sequences rather than replacement for native residues.
Therefore, in the construction of viable and well refined CpmV vector, precise site of insertion of foreign sequence was given a prime choice by introducing foreign DNA sequence into βB-βC loop of the S protein for most chimeras foreign sequences inserts immediately upstream of proline 23 of the S protein. In view of propagating the chimeras, engineered pCP_2 and pCP_1 are linearised and inoculated on cow pea plants. (Fig. 14.14)

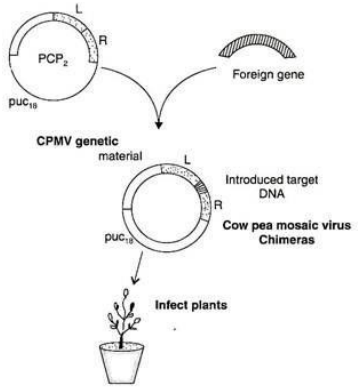


Fig. 14.14 CpmV expression vector construction and infection of plants

Vectors for S. cerevisiae:

There are three types of vectors for S. cerevisiae:

1. Episomal or plasmid vectors.

2. Integrating vectors.
3. Yeast artificial chromosomes (YACs).

1. Plasmid vectors:

Among the vectors, plasmids with single cloned genes are widely used. Manipulation with growth conditions increase the vector stability and expression efficiency. Use of tandem gene arrays has not met with success, since they are unstable.

2. Integrating vectors:

They are basically the integration of cloned genes with chromosomal DNA. These are not frequently used, since the protein production is low.

3. Yeast artificial chromosome (YAC):

Introduced in 1987, YAC is a fragment of yeast DNA that will accept a foreign DNA of about 250-500 kb in length. In fact, the yeast DNA is only about 1% of the total DNA which however, is very important, since it contains three essential genes required for replication. These are the genes for telomere (that protects DNA from nuclease degradation and thus maintains stability), centromere (forms spindles during cell division) and the origin of replication (where DNA polymerase initiates replication). YAC behaves just like a chromosome and replicates.

The construction of the yeast artificial chromosome is depicted in Fig. 11.2. Two opposite ends of a yeast chromosome namely the left telomere and right telomere are chosen. The left telomere is then attached to a centromere. A large segment of the foreign DNA is added and all the three are ligated. Unlike the plasmid vectors, the stability of YAC increases as the size of insert DNA increases.

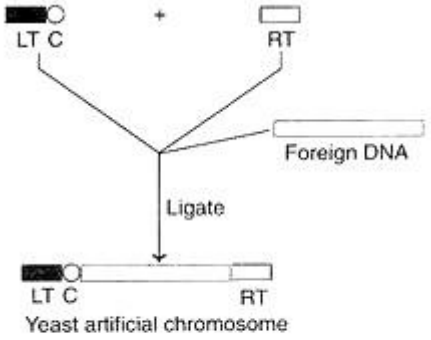


Fig. 11.2 : Construction of yeast artificial chromosome (LT-Left telomere; C-Centromere; RT-Right telomere).

YACs have not been used for commercial production of recombinant proteins. However, they have been employed successfully for physical mapping of genomic DNAs, particularly in human genome project.

Post-translational modifications by S. cerevisiae:

The heterologous proteins synthesized by S. cerevisiae undergo post-translational changes while they are being exported into the extracellular environment. To facilitate protein secretion, a single (leader) peptide is attached to the protein. This peptide is removed by the yeast endoprotease.

Other Yeast Expression Systems:

Despite the very successful use of S. cerevisiae for generating recombinant proteins, there are certain limitations. These include a very low or a limited yield, difficulty in secretion of some proteins and hyper glycosylation. Attempts are being made to explore the utility of other yeasts for the production of hepatitis B virus surface antigen (HBsAg) and bovine lysozyme. The yeast, Hansenula polymorpha, is employed for the synthesis of α- and β-globin chains of human hemoglobin.

Insect Cell Expression Systems:

Cultured insect cells are in use for expressing cloned DNAs. Baculo viruses exclusively infect insect cells. The DNA of these viruses encode for several products and their productivity in cells is very high to the extent of more than 10,000 times compared to mammalian cells. Besides carrying a large number of foreign genes, the baculoviruses can effectively express and process the products formed. Another advantage with these viruses is that they cannot infect humans, other vertebrates or plants. Thus, baculoviruses are safe vectors.

Polyhedrin gene of baculovirus:

The polyhedrin gene is responsible for the synthesis of a matrix protein-polyhedrin. This protein is synthesized in large quantities by baculovirus during the infection cycle. Polyhedrin protects the virus from being inactivated by environmental agents. The promoter for polyhedrin gene is very strong. However, the life-cycle of baculovirus does not depend on the presence of this gene. Polyhedrin gene can be replaced by a cloned gene, and the genetically engineered baculovirus can infect the cultured insect cells.

The cloned gene expresses, and large quantities of recombinant proteins are produced. Because of a close similarity in the post-translational modifications between insects and mammals, biologically active proteins can be produced by this approach. And in fact, by using baculovirus as an expression vector system, a good number of mammalian and viral proteins have been synthesized (Table 11.1).

Table 11.1 Selected examples of recombinant proteins produced by baculovirus expression vector system

Adenosine deaminase
Alkaline phosphatase
Amyloid precursor protein
Anthrax antigen
DNA polymerase α
Erythropoietin
HIV-I envelope protein
Interferons (α, β)
Interleukin-2
Malaria proteins
Pancreatic lipase
Polio virus proteins
Rabies virus proteins
Rhodopsin
Simian rotavirus capsid antigen
Tissue plasminogen activator

Baculovirus expression vector system:

The most commonly used baculovirus is Autograph California multiple nuclear polyhedrosis virus (AcMNPV). It can grow on the insect cell lines (e.g., derived from fall army worm) and produce high levels of polyhedrin or a recombinant protein. The organization of a baculovirus

(AcMNPV) transfer vector is shown in Fig. 11.3A. It consists of an E. coli-based plasmid vector along with the DNA of baculovirus. This in turn has AcMNPV DNA, a polyhedrin promoter region, cloning site for insert DNA and polyhedrin termination region.

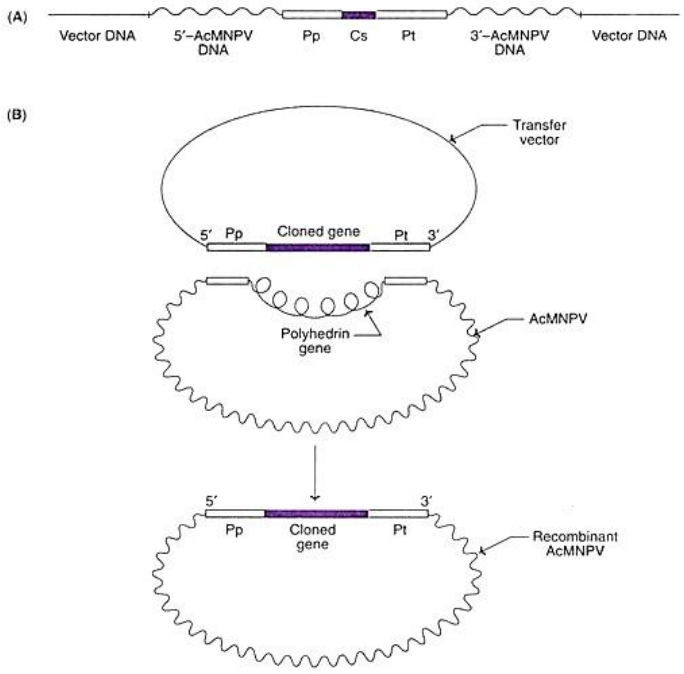


*Fig. 11.3 : Baculovirus expression vector system (A) Organization of baculovirus transfer vector (B) Replacement of the polyhedrin gene of baculovirus with a cloned gene from a transfer vector (AcMNPV–Autographa californica multiple nuclear polyhedrosis virus; Pp–Polyhydrin gene promoter; Cs–Cloning site; Pt–Polyhedrin gene termination; **Note** : The coding region of polyhedrin gene not shown in A).*

When the insect culture cells, transfected with AcMNPV are mixed with transfer vector carrying a cloned gene, a double crossover occurs. The result is that the cloned gene with polyhedrin promoter and termination sequences gets integrated into AcMNPV DNA (Fig. 11.3B). In this process, polyhedrin gene is lost. The recombinant baculovirus containing cloned gene is isolated. The host insect culture cells, on infection with recombinant baculovirus, produce heterologous proteins. A large number and a wide variety of recombinant proteins (around 500) have been synthesized in the laboratory. A majority of them (>95%) have the requisite post-translational modifications. A selected list of recombinant proteins is given in Table 11.1.

Modifications in the production of recombinant baculovirus:

The original method of creating recombinant baculovirus has undergone several changes. Incorporation of a unique Bsu 361 restriction endonuclease site on the polyhedrin gene increases the yield of recombinant baculovirus production to about 30% from the normal 1%.

Bacmid:

This is shuttle vector for E. coli and insect cell baculovirus. Construction of a recombinant bacmid is a novel approach to carry out all the genetic manipulations including the expression of baculovirus vector in E. coli.

Use of yeast cells:

The genetic manipulations of AcMNPV genome can be done in yeast cells with yeast-insect shuttle vector. Then the recombinant baculovirus is introduced into insect cells.

Mammalian Cell Expression Vectors:

Mammalian expression vectors are useful for the production of specific and authentic recombinant proteins (for use as therapeutic agents). In addition, they are also helpful for studying the function and regulation of mammalian genes. In general, the mammalian expression vectors are quite comparable to other eukaryotic expression vectors. . However, large-scale production of recombinant proteins with engineered mammalian cells is costly.

A diagrammatic representation of mammalian vector is shown in Fig. 11.4. It contains a eukaryotic origin of replication from an animal virus such as Simian virus 40 (SV40) and a prokaryotic origin of replication (from E. coli). The mammalian vector has a multiple cloning site and a selectable marker gene. Both of them are under the control of eukaryotic promoter and polyadenylation sequences.

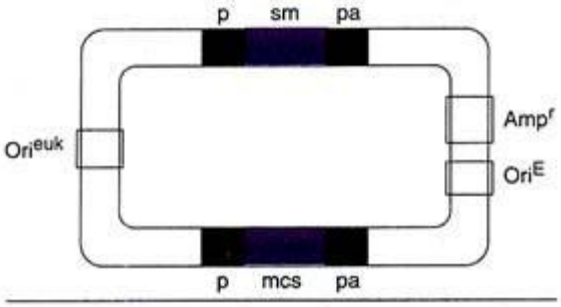


Fig. 11.4 : *A diagrammatic representation of mammalian expression vector (p–Promoter sequence; Pa–polyadenylation sequence; mcs–Multiple cloning site; sm-Selectable marker gene; Ori^{euk}–Origin of eukaryotic replication; Ori^E–Origin of* E. coli *replication; Amp^r–Ampicillin resistant marker gene).*

These sequences are obtained from either animal viruses (SV40, herpes simplex virus) or mammalian genes (growth hormone, metallothionein). The promoter sequences facilitate the transcription of cloned genes (at the multiple cloning site) and the selectable marker genes. On the other hand, the polyadenylation sequences terminate the transcription. Ampicillin resistant marker gene can be used for selecting the transformed E. coli cells.

Markers for mammalian expression vectors:

There are several markers in use for the selection of transformed mammalian cells. The bacterial gene (Neo^r) that encodes for neomycin phosphotransferase is frequently used. The other markers are the genes that encode for the enzyme dihydrofolate reductase (DHFR), and glutamine synthetase (GS).

DNA sequencing

DNA sequencing is the determination of the precise sequence of nucleotides in a sample of DNA. Before the development of direct DNA sequencing methods, DNA sequencing was difficult and indirect. The DNA had to be converted to RNA, and limited RNA sequencing could

be done by the existing cumbersome methods. Thus, only shorter DNA sequences could be determined by this method. Using this method, Walter Gilbert and Alan Maxam at Havard University determined that the Lac operator is a 27 bp long sequence.

The development of direct DNA sequencing techniques changed the scope of biological research. The evolution of DNA sequencing technology from plus-minus sequencing to pyro-sequencing within about 20 years parallels the progress in biology from molecular biology to genomics.

The development of DNA sequencing techniques with enhanced speed, sensitivity and throughput are of utmost importance for the study of biological systems. Sequence determination is most commonly performed using di-deoxy chain termination technology. Pyro-sequencing, a non-electrophoretic real- time bio-luminometric method for DNA sequencing has emerged as a state of the art sequencing technology.

This technology has the advantage of accuracy, ease of use, and high flexibility for different applications. Pyro-sequencing allows the analysis of genetic variations including SNPs, insertion/deletions and short repeats, as well as assessing RNA allelic imbalance, DNA methylation status and gene copy number.

DNA sequencing by Sanger's method :

The first DNA sequencing method devised by Sanger and Coulson in 1975 was called plus and minus sequencing that utilized E. coli DNA pol I and DNA polymerase from bacteriophage T4 with different limiting triphosphates. This technique had a low efficiency. Sanger and co-worker (1977) eventually invented a new method for DNA sequencing via enzymatic polymerization that basically revolutionized DNA sequencing technology.

The most popular method for doing this is called the dideoxy method or Sanger method (named after its inventor, Frederick Sanger, who was awarded the 1980 Nobel prize in chemistry [his second] for this achievement). Finding a single gene amid the vast stretches of DNA that make up the human genome – three billion base-pairs' worth – requires a set of powerful tools. These tools include genetic maps, physical maps and DNA sequence which is a detailed description of the order of the chemical building blocks, or bases, in a given stretch of DNA.

Scientists need to know the sequence of bases because it tells them the kind of genetic information that is carried in a particular segment of DNA. For example, they can use sequence information to determine which stretches of DNA contain genes, as well as to analyze those genes for changes in sequence, called mutations, that may cause disease.

The first methods for sequencing DNA were developed in the mid-1970s. At that time, scientists could sequence only a few base pairs per year, not nearly enough to sequence a single gene, much less the entire human genome. By the time the HGP began in 1990, only a few laboratories had managed to sequence a mere 100,000 bases, and the cost of sequencing remained very high. Since then, technological improvements and automation have increased speed and lowered cost to the point where individual genes can be sequenced routinely, and some labs can sequence well over 100 million bases per year.

DNA is synthesized from four deoxynucleotide triphosphates. The top formula shows one of them: deoxythymidine triphosphate (dTTP) (Fig. 23.7). Each new nucleotide is added to the 3′ – OH group of the last nucleotide added.

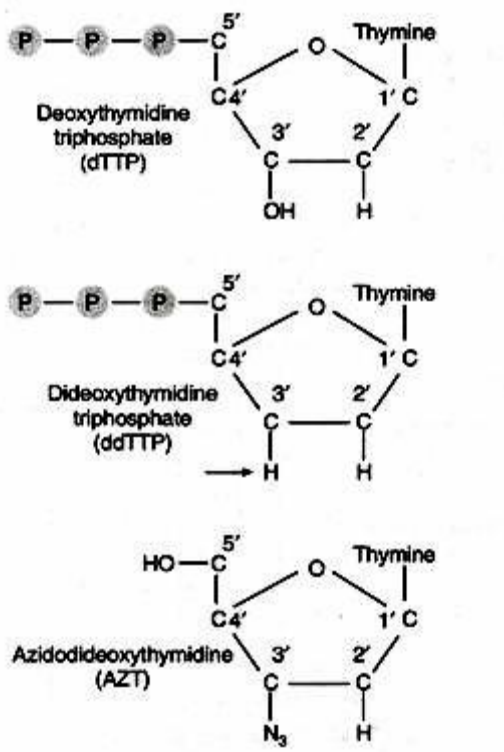


Fig. 23.7. Structure of dideoxynucleotides.

The dideoxy method gets its name from the critical role played by synthetic nucleotides that lack the -OH at the 3′ carbon atom. A dideoxynucleotide (dideoxythymidine triphosphate – ddTTP as shown here) can be added to the growing DNA strand. When it is added it stops chain elongation because there is no 3′ -OH for the next nucleotide to be attached. For this reason, the dideoxy method is also called the chain termination method.

The bottom formula shows the structure of azidothymidine (AZT), a drug used to treat AIDS. AZT (which is also called zidovudine) is taken up by cells where it is converted into the triphosphate. The reverse transcriptase of the human immunodeficiency virus (HIV) prefers AZT triphosphate to the normal nucleotide (dTTP). Because AZT has no 3′ -OH group, DNA synthesis by reverse transcriptase halts when AZT triphosphate is incorporated in the growing DNA strand. Fortunately, the DNA polymerases of the host cell prefer dTTP, so side effects from the drug are not as severe as might have been predicted.

The Procedure:

The DNA to be sequenced is prepared as a single strand (Fig. 23.8).

This template DNA is mixed with the following:

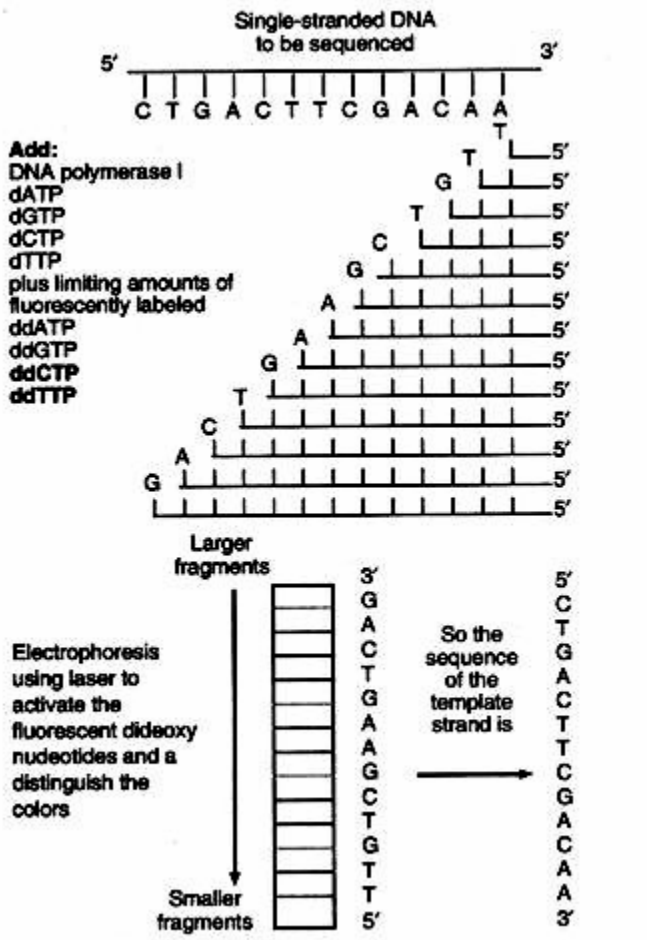


Fig. 23.8. Sanger's method of DNA sequencing.

(a) A mixture of all four normal (deoxy) nucleotides in sample quantities
i. dATP
ii. dGTP
iii. dCTP
iv. dTTP

(b) A mixture of all four dideoxynucleotides, each present in limiting quantities and each labeled with a "tag" that fluoresces a different colour:
i. ddATP
ii. ddGTP
iii. ddCTP
iv. ddTTP

(c) DNA polymerase I:

Because all four normal nucleotides are present, chain elongation proceeds normally until, by chance, DNA polymerase inserts a dideoxy nucleotide instead of the normal deoxynucleotide. If the ratio of normal nucleotide to the dideoxy versions is high enough, some DNA strands will

succeed in adding several hundred nucleotides before insertion of the dideoxy version halts the process.

At the end of the incubation period, the fragments are separated by length from longest to shortest. The resolution is so good that a difference of one nucleotide is enough to separate that strand from the next shorter and next longer strand. Each of the four dideoxynucleotides fluoresces a different colour when illuminated by a laser beam and an automatic scanner provides a printout of the sequence.

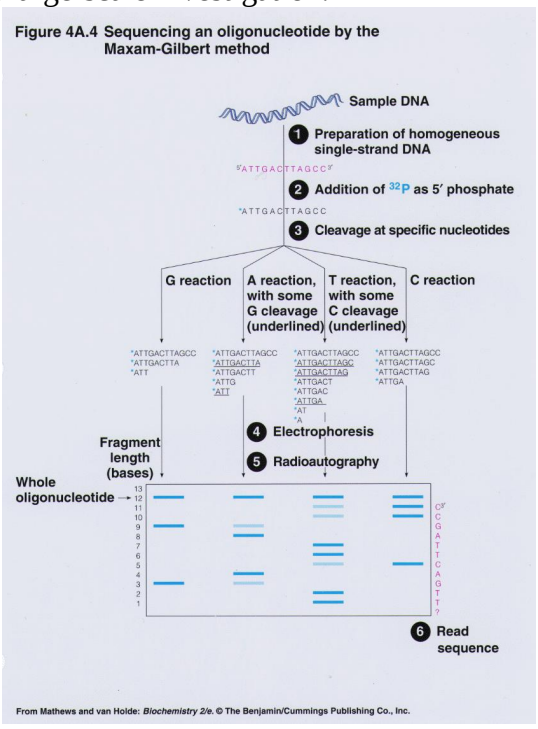
Modifications based on Sanger's method

Maxam and Gilbert Method:

In 1977, Maxam and Gilbert described a sequencing method based on chemical degradation at specific locations of the DNA molecule. The end labeled DNA fragments are subjected to random cleavage at adenine, cytosine, guanine or thymine positions using specific chemical agents and the products of these fours reactions are separated using polyacrylamide gel electrophoresis (PAGE). As in Sanger method, the sequence can be easily read from four parallel lanes in the sequencing gel.

Double stranded or single stranded DNA from chromosomal DNA can be used as template. Originally, end labeling was done with P phosphate or with a nucleotide linked to P and enzymatically incorporated into the end fragment. The read length is up to 500bp. The chemical reactions in the technique are slow and involved hazardous chemicals that require special handling in the DNA cleavage reaction.

As in Sanger's method, additional cautions in Maxam and Gilbert method include purification and separation of DNA fragments and higher analysis time. Therefore, this technology is not suitable for high throughput large-scale investigation.



Hybridization Method:

Ed Southern's (1990) sequencing by hybridization technique relies on detection of specific DNA sequences using hybridization of complementary probes. It utilizes a large number of short nested oligonucleotides immobilized on a solid support to which the labeled sequencing template is hybridized. The target sequence is deduced by computer analysis of the hybridization pattern of the sample DNA.

DNA sequence can also be analyzed by sequencing by synthesis. Sequencing by hybridization makes use of a universal DNA microarray, which harbors all nucleotides of length k (called "k-words", or simply words when k is clear). These oligonucleotides are hybridize to an unknown DNA fragment, whose sequence one would like to determine.

Under ideal conditions, this target molecule will hybridize to all words whose Watson-Crick complements occur somewhere along its sequence. Thus, in principle, one would determine in a single microarray reaction the set of all k-long substrings of the target and try to infer the sequence from those data.

The average length of a uniquely resconstructible sequence using an 8-mer array is <200 bases, far below a single read length on commercial gel-lane machine. The main weakness of sequencing by hybridization is ambiguous solutions-when several sequences have the same spectrum; there is no way to determine the true sequence.

Pal Nyren's Method:

In 1996, Pal Nyren's group reported that natural nucleotide can be used to obtain efficient incorporation during a sequencing-by-synthesis protocol. The detection was based on the pyrophosphate (inorganic biphosphate) released during the DNA polymerase reaction, the quantitative conversion of pyrophosphate to ATP by sulfurylase and the subsequent production of visible light by firefly luciferase.

The first major improvement was inclusion of dATPaS in place of dATP in the polymerization reaction, which enabled the pyrosequencing reaction to be performed in homogeneous phase in real time.

The non-specific signals were attributed to the fact that dATP is a substrate for luciferase. Conversely, dATPaS was found to be inert for luciferase, yet could be incorporated efficiently by all DNA polymerases tested. The second improvement was the introduction apyrase to the reaction to make a four-enzyme system. Apyrase allows nucleotides to be added sequentially without any intermediate washing step.

Pyrosequencing nonelectrophoretic real-time DNA sequencing method is based on sequencing by synthesis based on the pyrophosphate (inorganic biphosphate) released during the DNA polymerase reaction.

In a cascade of enzymatic reaction, visible light is generated that is proportional to the number of incorporated nucleotides. The cascade starts with a nucleic acid polymerization reaction in which inorganic bip-hosphate (PPi) is released as a result of nucleotide incorporation by polymerase.

The released PPi is subsequently converted to ATP by ATP sulfurylase, which provides the energy to luciferase to oxidize luciferin and generate light. The light so generated is captured by a CCD camera and recorded in the form of peaks known as pyrogram (compared with

electropherograms in Sanger's method). Because the added nucleotide is known the sequence of template can be determined.

Standard pyrosequencing uses the Klenow fragment of E. coli DNA pol I, which is relatively slow polymerase. The ATP sulfurylase used in pyrosequencing is a recombinant version from the yeast and the luciferase is from the American firefly. The overall reaction from polymerization to light detection takes place within three to four seconds at real time.

One pmol of DNA in a pyrosequencing reaction yields 6×10^{11} ATP molecules which in turn, generate more than 6×10^{9} photons at a wavelength of 560 nm. This amount of light is easily detected by a photodiode, photomultiplier tube or a CCD camera. Pyrosequencing technology has been further improved into array-based massively parallel microfluidic sequencing platform.

Slab Gel Sequencing Systems:

These systems make use of ultrathin (75 µm) slab gels and involve running of atleast 96 lanes per gel. In these systems, automation in sample loading of sequencing gels has also been achieved, by using a plexiglass block having wells that are same distance apart as the comb teeth cut in a porous membrane that is used as a comb for drawing samples by capillary action.

Each well in plexiglass block is filed with a sample (PCR dideoxy-reaction mixture), so that when the porous membrane comb is lowered onto the sample wells in the pexiglass the samples are drawn up automatically into the comb teeth by capillary action.

Using this approach of employing porous combs, automated loading of up to 192, 384 or 480 samples per gel has been achieved. The porous comb with the samples is placed between the glass plates of the gel apparatus above the flat surface of the polymerized gel and the samples are driven from the comb into the gel by electrophoresis.

Capillary Gel Electrophoresis:

In these systems, slab gel electrophoresis is replaced by capillary gel electrophoresis to analyse DNA samples. In these systems, instead of scanning DNA as it migrates through 96 lanes each in a series of 96 capillary tubes, DNA fragments pass are scanned.

In the original models of the above old slab gel machines, gels must be poured and reagents frequently reloaded, interrupting the sequencing.

In capillary gel sequencing systems, on the other hand, the robot moves the DNA samples and reagents through the tubes continuously, requiring attention only once a day. The system produces a steady flow of data, each signal representing one of the four DNA bases (adenine, cytosine, guanine and thymine).

Automated DNA sequencing

A variant of the above dideoxy-method was developed, which allowed the production of automatic sequencers. In this new approach, different fluorescent dyes are tagged either to the oligonucleotide primer (dye primers) in each of the four reaction tubes (blue for A, red for C, etc), or to each of the four ddNTPs (dye terminators) used in a single reaction tube: when four tubes are used, they are pooled.

After the PCR reaction is over, the reaction mixture is subjected to separation of synthesized fragments through electrophoresis (Fig. 23.9). Depending upon the electrophoretic system used,

whether slab gel electrophoresis or capillary electrophoresis, following two types of automatic sequencing systems have been designed.

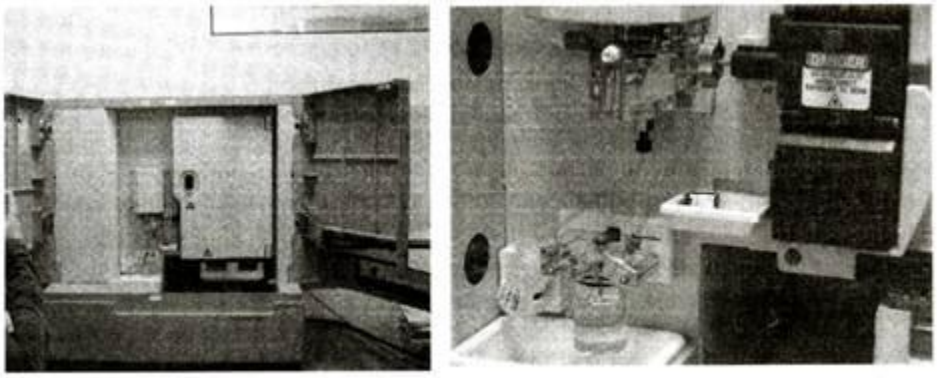


Fig. 23.9. Automated DNA sequencer (left) and details of sample loop (right).

Pyrosequencing

In laboratories around the world there is an intense desire to sequence more genomes.

- those of a wide variety of organisms to aid in establishing evolutionary relationships;
- those of pooled populations of microorganisms in, for examples, sea water, soil, the large intestine;
- other humans to look for
    - genes that predispose to disease;
    - genetic patterns in various ethnic groups.

All of the sequenced genomes listed in Genome Sizes were determined using the dideoxy method invented by Frederick Sanger and described in the page DNA Sequencing.

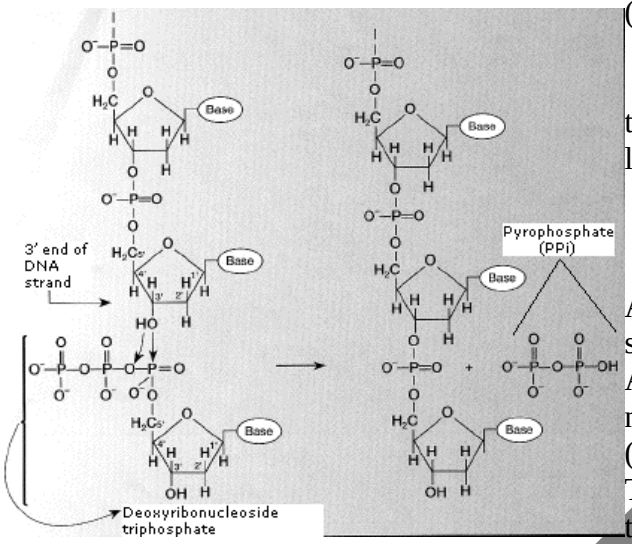
But now a great effort is being expended to find ways to sequence DNA more rapidly (and more cheaply).

Several new methods are being developed and one is already commercially available (the Genome Sequencer 20 System). Its method is called **pyrosequencing** or **sequencing by synthesis**.

It works like this.

- The DNA to be sequenced is broken up into fragments of ~100 base pairs and denatured to form single-stranded DNA (ssDNA).
- Single ssDNA fragments are attached to microscopic beads, which are separated from each other.
- The polymerase chain reaction (PCR) is run on each bead so that each becomes coated with ~ 10 million identical copies of that fragment.
- The beads are placed singly into separate, microscopic wells (~200,000 of them).
- Each well receives a cocktail of reagents:
    - **DNA polymerase** — for adding deoxyribonucleotides to the ssDNA
    - **adenosine phosphosulfate** (**APS**)

- **ATP sulfurylase** — an enzyme that forms ATP from adenosine phosphosulfate (APS) and pyrophosphate (PPi).
- **luciferin**
- **luciferase** — an ATPase that catalyzes the conversion of **luciferin** to **oxyluciferin** with the liberation of **light**.



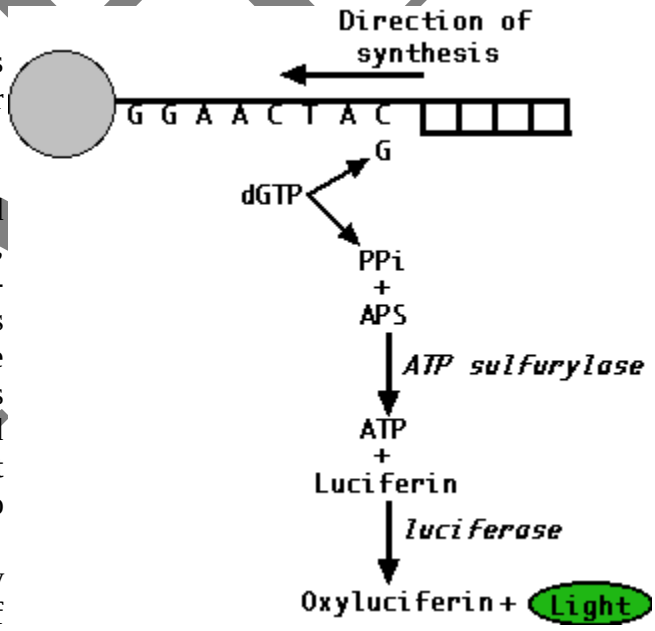
After a primer is annealed to the end of the ssDNA, synthesis is ready to begin.

As is always true of DNA synthesis, incoming nucleotides are added to the 3' end of the growing chain (left).

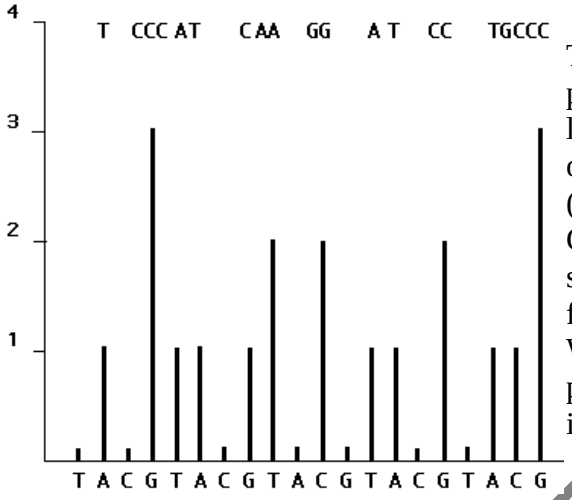
The nucleotides are supplied as four deoxynucleoside triphosphates. As each nucleotide is added, a molecule containing two phosphate groups — called **pyrophosphate** (PPi) is split off.

The sequencing run:

- Each of the thousands of wells is flooded with one four deoxyribonucleotides
  - **dTTP**, **dCTP**, and **dGTP** but
  - instead of dATP (which would trigger the luciferin reaction), deoxyadenosine alpha-thiotriphosphate (dATPαS) is used instead. DNA polymerase ignores the difference and uses it whenever a T is encountered on the ssDNA template, but luciferase doesn't recognize to it.



- In any well where the complementary nucleotide is present at the 3' end of the template, the nucleotide is added and pyrophosphate is liberated.
- The amount of light is proportional to the number of that nucleotide added. So if, for example, the incoming nucleotide is dGTP, and there is a string of 3 Cs on the template, the light emitted will be 3 times brighter than if only one C is present.
- A detector picks up the light (if any) from each well and the data are recorded.
- Then each of the remaining 3 nucleotides are added in sequence.
- Then the sequence of 4 additions is repeated until synthesis is complete.



The diagram on the left shows the type of data produced in a single well. The height of the peak of light production gives the number of additions that occurred when a particular nucleotide was added (bottom).

Computer software then displays the template sequence (top) for each of the thousands of different fragments sequenced.

With this technology, as many as 20 million base pairs of genome sequence can be learned in an instrument run of less than 6 hours.

POSSIBLE QUESTIONS

8 MARKS

1. Explain about vectors of gene cloning in detail.
2. Discuss and differentiate the roles of cosmids and phagemids.
3. Write in detail about the basic features, classification, copy number of plasmids with examples.
4. Explain in detail about bacteriophage and its life cycle.
5. Give a detailed account on plant viral vectors.
6. Explain the method of DNA sequencing in detail.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
III BSc BIOCHEMISTRY
17BCU601A-GENETIC ENGINEERING and BIOTECHNOLOGY
MULTIPLE CHOICE QUESTIONS

Unit-II

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	How many types in DNA polymerases are present in eukaryote?	One	Two	Three	Four	Three
2	The DNA molecule to which the gene of insert is integrated for cloning is called	Carrier	Transformer	Vector	Transducer	Vector
3	λ-Bacteriophages are the double stranded DNA of	47.5 kb	48.6 kb	45.7 kb	46.3 kb	48.6 kb
4	DNA sequencing was done by	Ribonucleotide chain terminators	Dideoxynuleotide chain terminators	cDNA chain terminators	mRNA chain terminators	Dideoxynucleotide chain terminators
5	Which gene is more complex in nature	Prokaryotic gene	Eukaryotic gene	Both	None	Eukaryotic gene
6	Which one is very helpful in the base sequence of a DNA molecule	RFLPs	Restriction maps	Both of the above	Generic maps	Both of the above
7	Gene cloning by the following one of the enzyme	Ectonuclease	Endonuclease	Restriction endonuclease	Ribonuclease	Restriction endonuclease
8	Which one helps to transfer the foreign DNA molecule into host cell?	cDNA	Cloning vectors	mRNA	Vehicle	Cloning vectors
9	One of the following is not connected with cloning vectors?	Bacteria	Plasmids	Bacteriophages	Cosmids	Bacteria
10	Plasmid was coined by	Maccarthy	Avery	Lederberg	Smith	Lederberg
11	_____ vectors provide signals that maximize gene expression	Cloning	Inducible	Expression	Strong	Expression
12	Plasmid carries	PMBI replicon	PI5A replicon	ColEi replicon	PkN402	PMBI replicon
13	Eukaryotes that habour a plasmid	*Fusarium*	*Saccharromyces cerevisiae*	*Pencillium*	*Aspergillus*	*Saccharromyces cerevisiae*
14	Recombinants using pBR322 plasmid vector are	Amp^R Tet^R	Amp^S Tet^S	Amp^R Tet^S	Amp^S Tet^R	Amp^R Tet^S
15	There are two types of naturally occurring plasmids	Col plasmid and F plasmid	E plasmid	*E. coli*	PBR322	Col plasmid and F plasmid
16	Transfer genes have cluster of	12 different genes	11 different genes	10 different genes	20 different genes	12 different genes
17	LacZ- *E.coli* cell synthesis	Beta galatosidase	Alpha galatosidase	Partial Beta galatosidase	Partial Alpha galatosidase	Partial Beta galatosidase
18	Plasmid contain autonomous replication (rep) includes	Origin of replication	chemicals	Copy number	genes	Origin of replication
19	PBR322 composed of	3 section	2 section	1 section	5 section	3 section
20	pUC8 Recombinant are	$Amp^{R.}$ make β galatosidase	$Amp^{R.}$ unable to make β galatosidase	$Amp^{S.}$ make β galatosidase	$Amp^{S.}$ unable to make β galatosidase	$Amp^{R.}$ unable to make β galatosidase
21	Phage infection in *E.coli* is visualized in agar medium through	Colony	Plaque	Blue colour	White colour	Plaque
22	Plaque formation is due to lysis of_____	Bacteria	Bacterio phage	Virus	Bacterio phage and virus	Bacteria
23	λEMBL4 carries up to	10 Kb	20kb	l5kb	25 kb	10 Kb
24	Ti Plasmids are	Transgenic plasmids	Tumour inducing plasmids	Inhibiting plasmids	Haemorhageic plasmids	Tumour inducing plasmids
25	Lambda DNA has a genetic material as	ds DNA	Ss DNA	ds RNA	ssRNA	Ss DNA
26	Vectors, which contains a unique site (x) for the insertion of foreign DNA ha	Expression vector	Replacement vector	Insertional vector	Shuttle vector	Insertional vector
27	Co-factor needed for type II restriction enzyme were	ATP and Mg 2+	S- Adenosyl Methionine	Fe	DNP+	ATP and Mg 2+
28	Alkaline phosphatase contain	Zn2+	Mg2+	Fe	Energy containing enzyme	Zn2+
29	Bacterial Alkaline Phosphatase action is inhibited by	Lysozyme	CTAB	EDTA	MgSO4	EDTA
30	What is the product produced by reverse transcriptase	single strand DNA	double strand DNA	single strand RNA	double strand RNA	Single strand DNA
31	Cassettes produce	Only coded protein	Produce *E.coli* protein	Fusion protein	Partial coded protein	Fusion protein
32	Mob function defined by	3 regions of plasmid DNA	5 regions of plasmid DNA	7 regions of plasmid DNA	2 regions of plasmid DNA	2 regions of plasmid DNA
33	The bacteria associated with the discovery of R plasmids	Bacillus	Pseudomonas	Shigella	Salmonella	Shigella
34	Tn3 from transposon for PBR322 have	Tetracyclin resistant	Streptomycin resistant	Ampicillin resistant	Chloramphenicol resistant	Ampicillin resistant
35	The best-studied Col plasmid is	ColA	ColB	ColE	ColD	ColE
36	One gene important for ampicillin resistant in PBR322 is	D galactose gene	P -galactose gene	β-lactamase gene	β-lactase gene	β-lactamase gene
37	PBR 322 vectors are restricted to	Gram positive bacteria	Gram negative bacteria	Archae bacteria	Eubacteria	Gram negative bacteria
38	Expression vectors signals that maximize gene expression	P, R, T	R, P, T	T, R, P	P, T, R	P, R, T
39	Lambda genome has a length of	48502bp	33402bp	40000bp	50,435bp	48502bp
40	All are problem with *E.coli* host except_____	Codon bias	Premature termination	Intron excision	Codon duplication	Codon duplication
41	The non-essential region of lambda phage is	int, xis, mob	int, xis, att	tra, mob, xis	B2 region	B2 region
42	All are vectors in plant except_____	Gemini virus	CaMV	Adeno virus	Monopartite virus	Adeno virus
43	The expression vector for somatostatin is	Psom II 3	P som II	P BH 20	P som I	Psom II 3
44	Inert materials used in microcarriers are	Gold, Copper , tungsten	Gold, Platinum , tungsten	Gold, platinum, Copper	Platinum, copper, tunsten	Gold, Platinum , tungsten
45	Non-conjugative plasmids have	tra+/ mob+ genes	tra-/ mob- genes	tra-/ mob+ genes	Mob+/ mob+ genes	tra-/ mob+ genes
46	Runaway vectors replicate upto	34°C	40^{o}C	94°C	None	34°C
47	Initiation codon always lies near this site	Ribosome binding site	Promoter	Terminator	Consenses sequence	Ribosome binding site
48	Uptake of naked DNA into the cell is called as _____	Transfection	Transformation	Transduction	Transdution	Transformation
49	The cell is made competent by the treatment of _____	50mM $CaCl_2$ and 42°C	42mM $CaCl_2$ and 50°C	50mM $CaCl_2$ and 50°C	42M $CaCl_2$ and 42°C	50mM $CaCl_2$ and 42°C
50	λgt10 carry up to	8 Kb of new DNA	10kb new DNA	15kb new DNA	20 kb new DNA	10kb new DNA
51	_____ gene expression can be quantified	Scorable	Selectale	Reporter	Readable	Reporter
52	The main difference between transformation and transfection is	Plasmid DNA	Phage DNA	Heat shock	Competent *E.coli* cells	Phage DNA
53	EM vector have	SP6 promoter	T7 Promoter	Both SP6 and T7	T8 Promoter	Both SP6 and T7
54	PH V 33 Multiply both in	*E. coli* and *Bacillus*	*Bacillus* and *Streptomyces*	*Bacillus* and *Yeast*	*Bacillus* and *Pseudomonas*	*E.coli* and *Bacillus*
55	plasmids are extra chromosomal DNA that vary in size from	lKb to More than 200kb	2Kb to More than 100kb	5Kb to More than 200kb	More than 300kb	lKb to More than 200kb
56	Defective *cos* site is seen _____ strain of *E. coli*	SMR 10	MSR 20	BHB2688	BHB2690	SMR 10
57	The term episome was coined by	A. Jacob and Wollman	Lederberg	Watson	Collins	A. Jacob and Wollman
58	The process by which the recombinant DNA is introduced into bacterium through bacterio phage	Transfection	Transformation	Transduction	Transdution	Transfection
59	The process of physical and chemical treatment that enhance the ability of cell to take up DNA is	Transfection	Transformation	Transduction	Competent	Competent
60	The genes for drug resistance is initially coded on	PBR322	Transposons	Yeast plasmid	None of the above	PBR322

UNIT-III
SYLLABUS

Introduction of DNA into cells and selection for recombinants : Uptake of DNA by cells, preparation of competent cells. Selection for transformed cells. Identification for recombinants - insertional inactivation, blue-white selection. Introduction of phage DNA into bacterial cells. Identification of recombinant phages. Introduction of DNA into animal cells, electroporation.

Methods for clone identification
The problem of selection, direct selection, marker rescue. Gene libraries, identification of a clone from gene library, colony and plaque hybridization probing, methods based on detection of the translation product of the cloned gene.

Introduction of DNA into cells and selection for recombinants :

Uptake of DNA by cells, preparation of competent cells.

Conjugation:

In this process, the exchange of genetic material takes place through a conjugation tube between the two cells of bacteria. The process was first postulated by Joshua Lederberg and Edward Tatum (1946) in Escherichia coli. They were awarded the Nobel Prize in 1958 for their work on bacterial genetics. Later on, it has also been demonstrated in Salmonella, Vibrio and Pseudomonas.

There are two mating types of bacteria, one is. male type or F^+ or donor cell, which donates some DNA. The other one is female type or F^- or recipient cell, which receives DNA.

Later, after receiving DNA, the recipient cell may behave as donor cell i.e., F^+ type. The F-factor is the fertility factor, sex-factor or F-plasmid present in the cell of F^+ i.e., donor cell or male type. The plasmid takes part in conjugation is called episome.

In this process, two cells of opposite mating type i.e., F^+ and F^- become temporarily attached with each other by sex pilus (Fig. 2.26). The sex pilus has a hole of 2.5 pm diameter through which DNA can pass from donor to recipient cell. The F-factor or F-plasmid is a double stranded DNA loop, present in the cytoplasm; apart from the nucleoid. The F-factor contains about 20 genes.

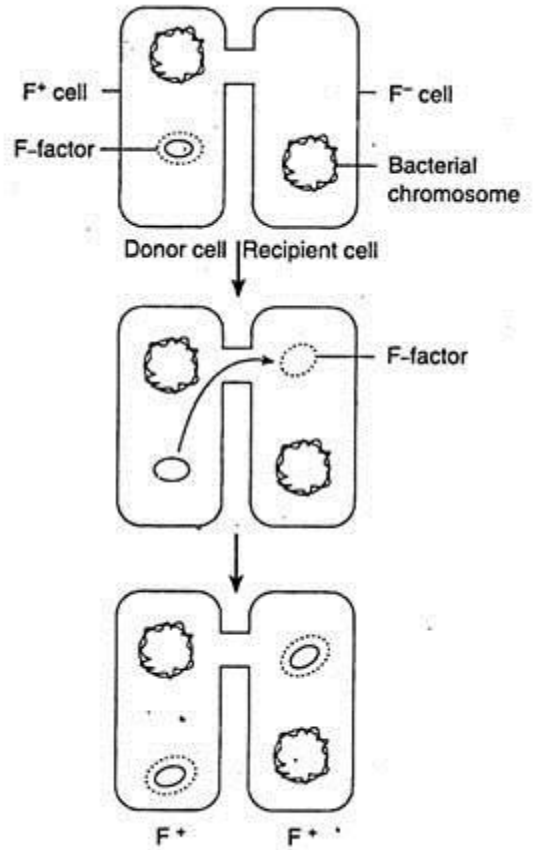


Fig. 2.26 : Recombination by F-factor in conjugation

After the establishment of conjugation tube, the F-factor prepares for replication by the rolling circular mechanism. The two strands of F- factor begin to separate from each other and one of them passes to the recipient i.e., F^- cell.

After reaching in F^- cell, enzymes synthesise a complementary strand that forms a double helix, which bends into a loop. The conversion process is thus completed. In the donor cell i.e., in F^+, a new DNA strand also forms to complement the left over DNA strand of the F-factor.

There is another type of conjugation where passage of nucleoid DNA takes place through conjugation tube. Strains of bacteria are known as Hfr (high frequency of recombination) strain. William Hayes discovered such strains of E. coli in 1950s. The Hfr factor is also called episome. In Hfr strain, the F-factor is attached with the nucleoid DNA i.e., the bacterial chromosome.

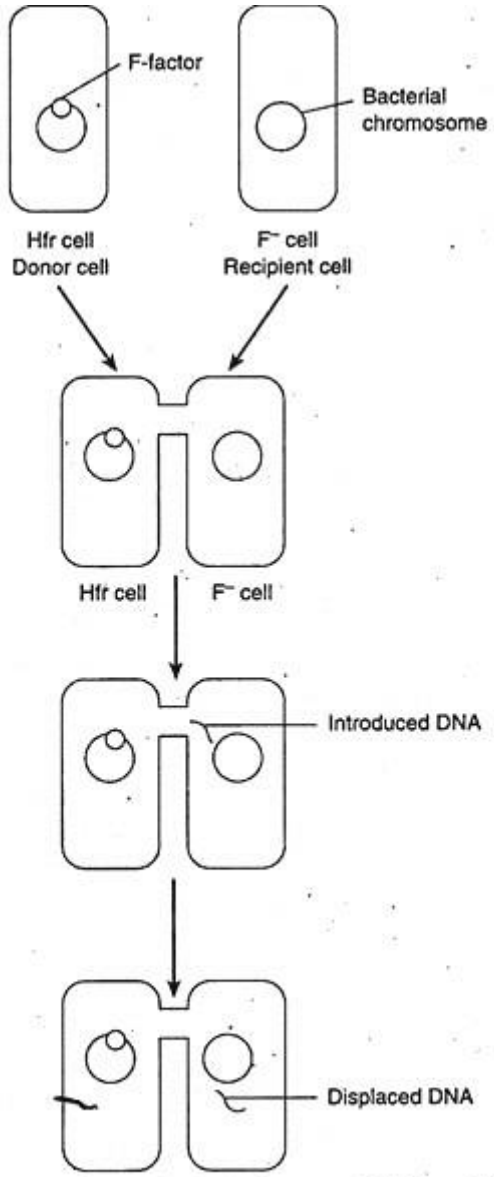


Fig. 2.27 : Recombination by fragment of DNA in conjugation

In this process, Hfr and F⁻ cells become attached with each other by sex pilus (Fig. 2.27). At the point of attachment of F-factor, the bacterial chromosome opens and a copy of one strand is formed by the rolling circular mechanism.

A portion of single stranded DNA then passes into the recipient cell through pilus. Due to agitation in medium, the conjugation tube may not survive for long time because of broken pilus. Thereby, the total length of transfer DNA may not be able to take entry to the recipient cell.

The behaviour of the transferred DNA depends on the presence and absence of F-factor:

If F-factor is indeed transferred, then it usually remains detached from the chromosome of recipient cell and enzymes synthesise a complementary DNA strand. The factor then forms a loop and exists as a plasmid, thereby the recipient cell becomes a donor.

If F-factor remains at the rear end of the transfer DNA during its entry to the recipient cell, the F-factor may not be able to take entry due to broken pilus and only a portion with new genes (Fig. 2.27) takes up the entry. Thereby, the F^- strain remains as recipient one. In F^-strain, genetic recombination takes place between donor fragment and recipient DNA.

(iii) Sometimes, if the F-factor gets free from the Hfr cell and maintains an independent status, then the Hfr cell converts to a F^+ cell. Sometimes during the leaving of F-factor from the bacterial chromosome, it takes a segment of chromosomal DNA. The F-factor with segment of chromosomal DNA is called F'-factor.

Later on, during conjugation, when this F'-factor is transferred, the recipient cell receives some chromosomal DNA from the donor cell. This process is called sexduction. In this process, the recipient cell receives a portion of chromosomal DNA which duplicates with the existing one for a specific function, thereby the recipient cell is a partial diploid.

Transformation:

It is a kind of genetic recombination where only the carrier of genes, i.e., the DNA molecules of donor cell, pass into the recipient cell through the liquid medium:

It was described by Frederick Griffith (1928), an English bacteriologist. He had done his experiment with laboratory mice and two types of Diplococcus pneumoniae, the pneumonia causing organism. One type has rough (R) non-capsulated cells and another one with smooth (S) capsulated cells. The R-type is non-pathogenic, while the S-type is pathogenic.

The process of transformation is mentioned below (Fig. 2.28):

(i) When live non-pathogenic (R-type) cells are injected in mice, the mice remain alive.

(ii) When dead pathogenic (S-type) cells are injected in mice, the mice also remain alive.'

(iii) When pathogenic (S-type) cells are injected in mice, they suffer from pneumonia and died.

(iv) When live non-pathogenic (R-type) cells are mixed with dead pathogenic (S-type) cells and are injected in mice, they also suffered from pneumonia and died. On isolation of dead tissue of mice, the smooth (S) qapsulated cells are found on agar. The above experiment indicates the conversion of R-type to S-type, called transformation.

Later, James L. Alloway (1932), transformed the rough type cells to smooth type, by using the fragments from dead smooth-type cells and confirmed Griffith's work.

Further, Oswald T. Avery, Colin M. MacLeod and Maclyn N. McCarty (1944) also found that DNA isolated from the fragments could induce the transformation. Their experimental result was the first proof of DNA as the genetic material in living organism. The possible mechanism of transformation can be explained (Fig. 2.29).

The transformation takes place in a few cell of the mixed population. It is an important method of genetic recombination. A few donor cells break apart and an explosive release and fragmentation of DNA take place. A fragment of double stranded DNA (10-20 genes) then gets attached with the recipient cell for entry (Fig. 2.29).

During entry one strand of the fragment becomes dissolved by enzyme leaving the second strand, which then passes to the recipient cell through cell wall and cell membrane.

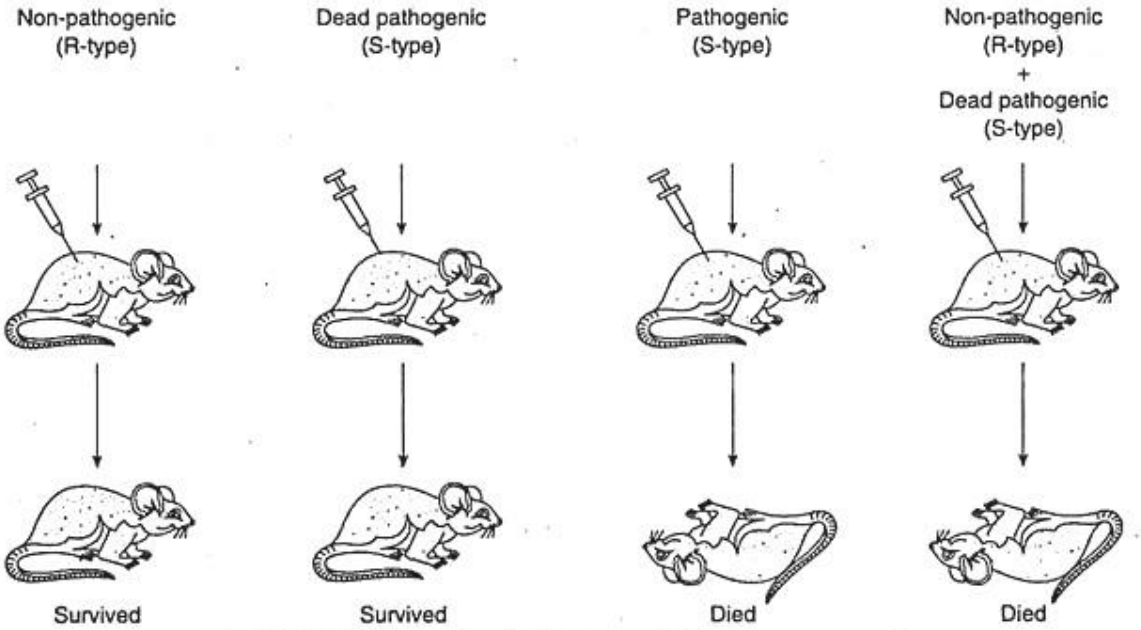


Fig. 2.28 : Griffith's experiment with mice and *Diplococcus pneumoniae*

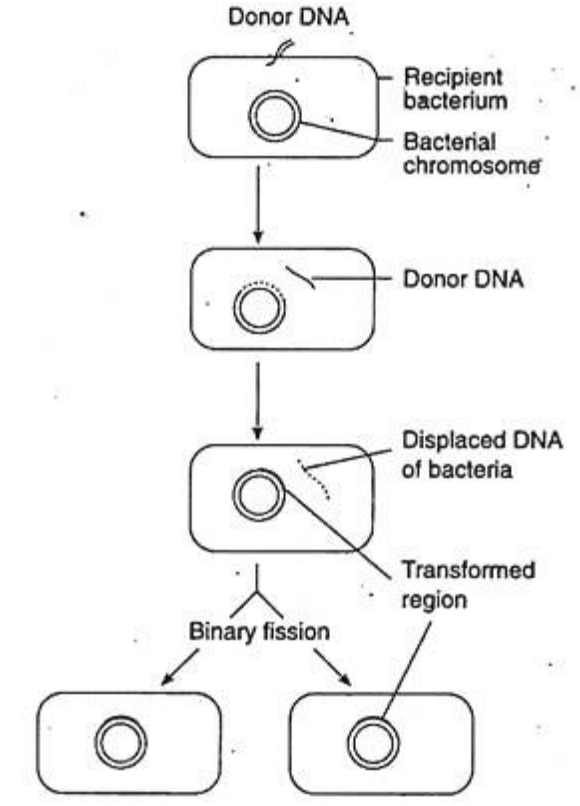


Fig. 2.29 : Diagrammatic representation of Transformation

After entry, a portion of single strand of double stranded DNA of recipient cell gets displaced by enzyme and then replaced by the DNA of donor cell. The displaced DNA is then dissolved by other enzyme. Thus the recipient cell becomes transformed which will display its own as well as the characters of the newly incorporated DNA.

Detailed mechanism of transformation, with especial emphasis on natural and induced competence and DNA uptake:

Thus the transformation takes place by horizontal gene transfer through uptake of free DNA by other bacteria. This transformation takes place either spontaneously by taking DNA from the environment, i.e., Natural, or by forced uptake under laboratory condition i.e., Artificial process.

A. Natural Transformation:

During natural transformation, free naked fragments of double stranded DNA of donor cell become attached to the surface of the recipient cell. The free double stranded ON A molecules may be available in the medium by lysis or natural decay of bacteria (Fig. 2.30).

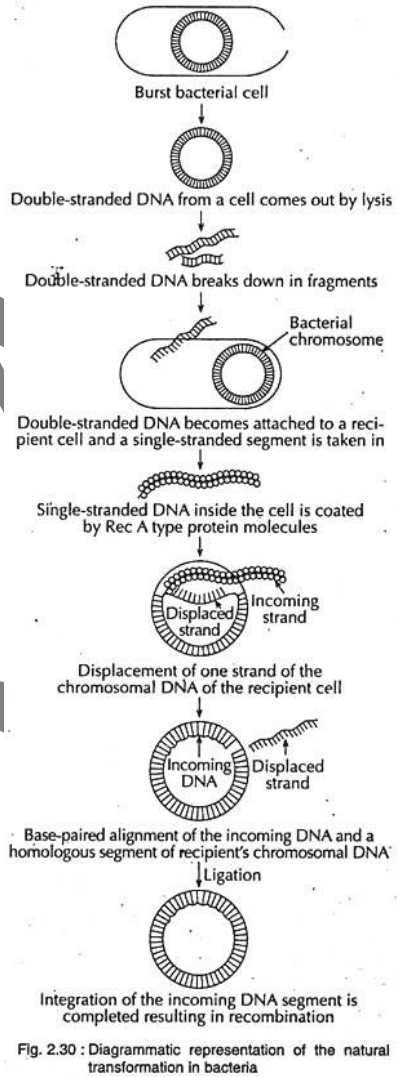


Fig. 2.30 : Diagrammatic representation of the natural transformation in bacteria

After attachment of donor double stranded DNA with the surface of recipient bacterium, one strand is digested by the bacterial nuclease and the remaining one strand is then taken in by an energy-requiring transport system. This uptake of DNA takes place during late logarithmic phase of growth.

During this process, Rec A type of protein plays an important role. The Rec A protein binds with the single stranded DNA and forms a coating around the DNA (Fig. 2.30). The coated single stranded DNA and DNA of recipient cell then move close to each other to get homologous sequence.

After reaching at proper place, the Rec A protein actively displaces one strand of chromosomal DNA of recipient cell. The process requires hydrolysis of ATP to get energy. The incoming DNA strand is then integrated with one strand of bacterial DNA by base pairing and ligation takes place by DNA ligase.

The displaced DNA strand of recipient cell is then digested by cellular DNase activity. Any mis-match between the two strands of new region is corrected by them. Thus the transformation is completed. If the introduced single stranded DNA fails to recombine with the recipient DNA, it is digested by cellular DNase and gets lost.

B. Artificial Transformation:

The E. coli, an ideal material for research is not transformed naturally. Later, it has been discovered that the transformation in E. coli can be done by special physical and chemical treatments. This can be done by exposure of E. coli to high voltage electric field and also by high concentration of $CaCI_2$. Under such condition, the bacterial cells are forced to take up foreign DNA. This type of transformation is called artificial.

During this process, the recipient bacterial cells are able to take up double stranded DNA fragments.

Physical or chemical treatment forces the recipient bacterial cell to receive exogenous DNA. The foreign DNA is then integrated with the chromosome by homologous recombination, mediated by Rec A protein. The Rec A protein catalyses the annealing of two DNA segments and exchange of homologous region.

This involves nick i.e., small cut of DNA strands and rejoining of exchanged parts i.e., breakage and reunion. The generally accepted model of the above phenomenon is given below (Fig. 2.31):

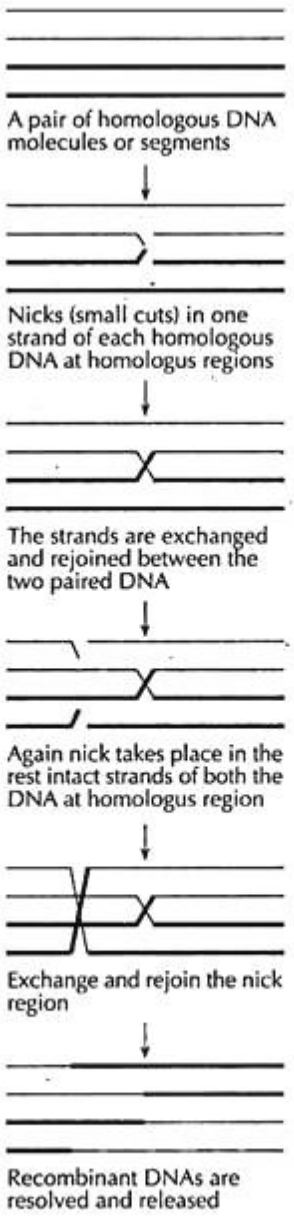


Fig. 2.31 : Diagrammatic representation of a model of homologous recombination by reciprocal crossing over between two DNA molecules or DNA segments

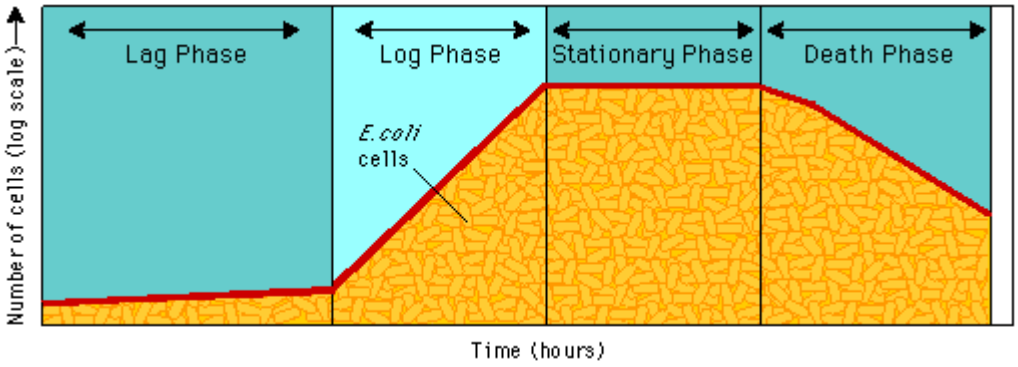
Difference between conjugation and transformation

Conjugation	Transformation
1. It takes place between the members of same species and also between the members of two different species of *Escherichia* and *Salmonella* or *Shigella; Salmonella* and *Serratia*.	It takes place between members of the same species.
2. Genetic recombination is temporary.	Genetic recombination is

Preparation of competent cells

E. coli cells are more likely to incorporate foreign DNA if their cell walls are altered so that DNA can pass through more easily. Such cells are said to be "competent." Cells are made competent by a process that uses calcium chloride and heat shock. Cells that are undergoing very rapid growth are made competent more easily than cells in other stages of growth.

The growth rate of a bacterial culture is not constant. In the early hours (lag phase), growth is very slow because the starting number of dividing cells is small. This is followed by a time of rapid cell division known as the log phase. The actual length of each phase depends on the temperature at which the cells are incubated.



Selection for transformed cells

After transformation, it is necessary to identify the cells that contain plasmid-cloned DNA constructs. In pBR 322 in which target DNA was inserted onto the BamHI site, recombinant bacteria (bacteria with recombinant plasmid) are selected.

All cells are grown successively on media containing antibiotic, ampicillin or tetracycline and cells showing the recombinant DNA (depending upon the restriction enzyme site and loss of particular antibiotic resistance due to disruption of the gene). Selected recombinant bacteria are grown in bioreactor to obtain the gene product (Fig. 14.10).

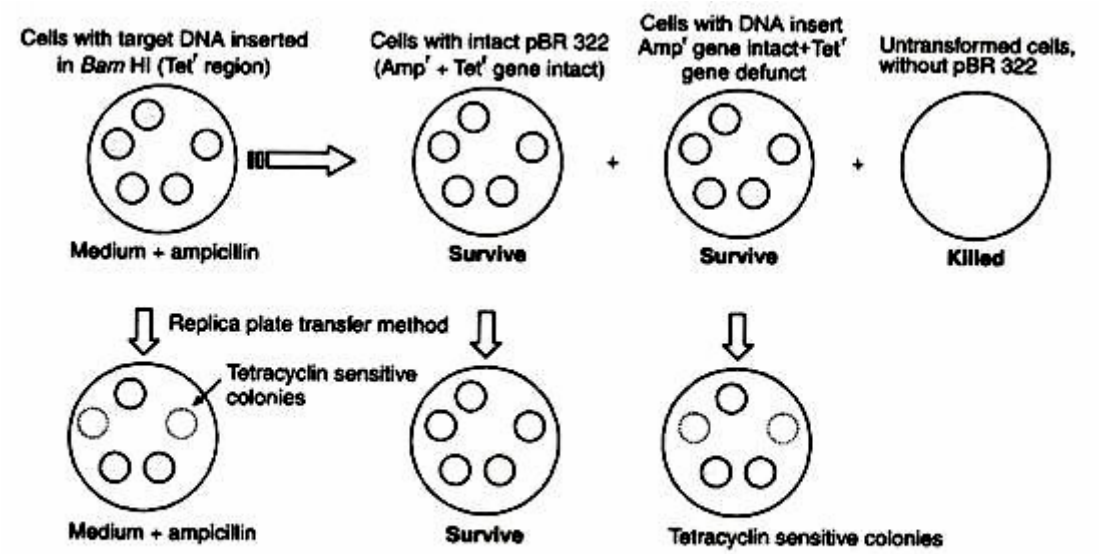


Fig. 14.10. Selection of recombinant bacterial cells containing recombinant plasmid. Colonies corresponding to tetracyclin sensitive colonies (B) are selected from original cultures (A). Growth of three types of cells obtained after transformation is also presented.

Identification for recombinants

Recombinant selection and screening

Mutations are genetic changes or modifications caused by chemical and physical mutagens. Mutations can results from modification of a single base or few bases. However this can result in change or modification of a phenotypic character which can be used to recognize them. This feature is widely used in DNA recombinant technology. Plasmid vectors carry genes for drug resistance, toxin production which can be used to distinguish recombinants. When genes of interest are inserted into the plasmid, the reading frame for the marker genes can be altered. This results in mutants who can be identified using special chemicals/ media.

Insertional Inactivation Method:

It is based on basic principle that cloned DNA fragment disrupts the coding sequence of gene.

To identify recombinants, one of the important approaches is to use DNA probe. In a DNA molecule, the two complementary strands are held together by hydrogen bonds. If two similar DNA pieces are mixed together and hydrogen bonds broken (by heating) the strands will separate.

Upon lowering the temperature, the hydrogen bonds are formed again. Some of the resultant double-stranded DNA will be hybrids i.e., composed of one strand of one type and one strand of the other type. This concept of DNA hybridization has been exploited for utilizing the DNA molecules as probes (Fig. 11.22).

The transformed colonies are replica plated to a nitrocellulose filter and are lysed to release the DNA. This DNA is denatured (by raising the temperature) and fixed to the nitrocellulose so as to produce a DNA print corresponding exactly to the position of the colonies on the original plate.

The DNA print is then hybridized with the probe which has been previously radioactively labelled. After washing off unhybridized DNA, the position of the radioactive spots on the filter is indicated by autoradiography in order to identify the presence of the required DNA.

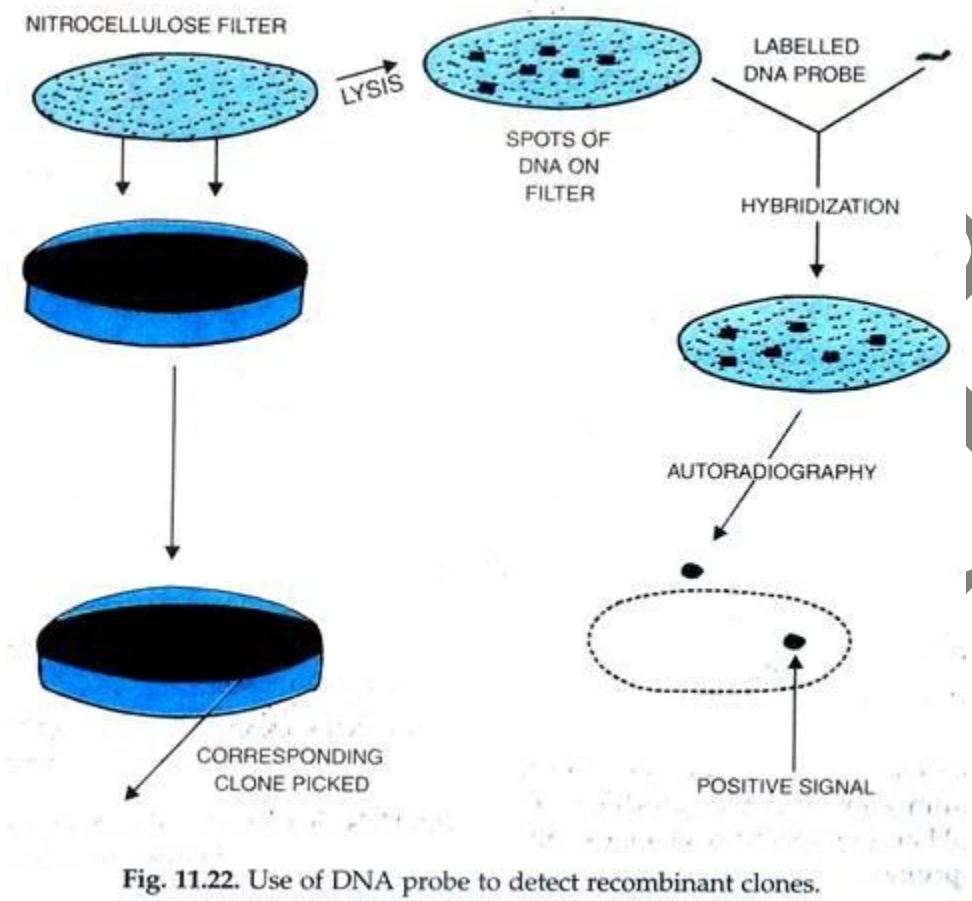


Fig. 11.22. Use of DNA probe to detect recombinant clones.

Blue white selection

It is a widely used method in screening recombinants in cloning. This is based on the gene product of lac z gene. The plasmid vectors contain this gene which produces **β galactosidase** enzyme. When a gene is inserted close to **lac z gene**, the reading frame will be distorted and the gene is inactivated. So the transformed cells will not produce this enzyme and are called competent cells. After the recombination, the bacterial cells are grown in a medium containing X gal (5-bromo-4-chloro-indolyl-β-D-galactopyranoside) and **IPTG (Isopropyl β-D-1-thiogalactopyranoside).** IPTG acts as the inducer for lac z gene and enhance the production of β galactosidase. When it is produced, combines with X gal to form a blue colour complex called **5,5'-dibromo-4,4'-dichloro-indigo** which is insoluble. The transformed colonies will appear white in colour and non- transformed cells will appear blue in colour. This method is also called as insertional inactivation of lac z gene.

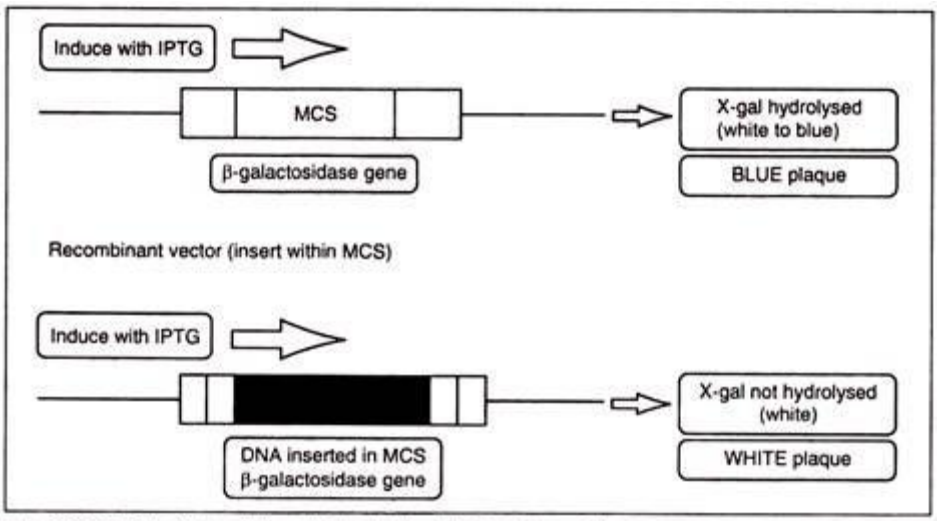


Fig. 4.17: Principle of blue-white selection for the detection of recombinant vectors

Introduction of phage DNA into bacterial cells

Transduction:

It is a special method of genetic recombination where genetic material is transferred from the donor to the recipient cell through a non- replicating bacteriophage — temperate bacteriophage. This was discovered by Joshua Leaderberg and Nortor Zinder (1952) during their research with Salrv onella typhimurium.

In this process, a small fragment of bacterial DNA is incorporated into an attacking bacteriophage (i.e., virus which infect bacteria) and when this bacteriophage infects a new bacterial cell, it transfers the genetic material into it, and thus genetic recombination takes place.

Transduction are of two types:

A. Specialised transduction, and

B. Generalized transduction.

A. Specialised Transduction:

In this process, the bacteriophage gets attached to a bacterial cell wall at the receptor site and the nucleic acid of bacteriophage is transferred into the cytoplasm of the host cell (Fig. 2.32A). The phage does not cause the lysis of the host bacterium. In the bacterial cell, the phage nucleic acid codes for the synthesis of specific proteins, the repressor proteins.

The repressor proteins prevent the virus to produce the material require for its replication. In the bacterial cell, the viral DNA may exist as a fragment in the cytoplasm or it may attach itself to the chromosome, known as prophage (Fig. 2.32B). The bacterial cell which carries the prophage is called lysogenic and the phenomenon where the phage DNA and bacterium exist together is called lysogeny.

The bacterial cell may remain lysogenic for many generations and during this period the viral DNA replicates many times together with the bacterial chromosome.

However, in course of time, the phage stops the synthesis of repressor proteins in the bacterial cell, and then the synthesis of phage components starts. Now the phage DNA separates from the bacterial chromosome and starts the synthesis of phage proteins (Fig. 2.32C).

During this separation, a number of genes of the bacterium get attached to it. These attached genes keep on replicating along with the phage DNA (Fig. 2.32D) and later on it develops into phage particles, those come out from the bacterial cell by bursting (Fig. 2.32E).

When the new phage particle (Fig. 2.32F) infects a new bacterial cell (Fig. 2.32G, H), the attached bacterial genes present along with phage particle enters in the chromosome of the new bacterium and causes recombination (Fig. 2.321).

Thus the new bacterial cell contains its own genes and several genes from the parent bacterial cell. This type of transduction is known as specialised transduction, which is an extremely rare event.

B. Generalised Transduction:

This process of transduction is more common than specialized transduction. Here the prophage particle is present in the cytoplasm of the infected bacterial cell (Fig. 2.32J). In this process, the phage DNA starts synthesising new phages.

During this process chromosome of bacterial cell gets fragmented (Fig. 2.32K) and some of the fragments become attached with the DNA of some new phage particle, while others remain with phase DNA (Fig. 2.32L).

When the newly formed phage with fragment of bacterial chromosome in its DNA (Fig. 2.32M) attacks a new bacterium, the gene of the parent bacterium is transferred to the new bacterium and causes recombination. This type of transduction is called generalised transduction. This type of transduction is also rare.

Difference between transformation and transduction

	Transformation	Transduction
1.	Transfer of genetic material takes place from donor to recipient bacterium through the liquid medium.	Transfer of genetic material takes place from donor to recipient bacterium through a Bacteriophage.
2.	The enzyme deoxyribonuclease can completely check the process.	The deoxyribonuclease has no effect on trans-

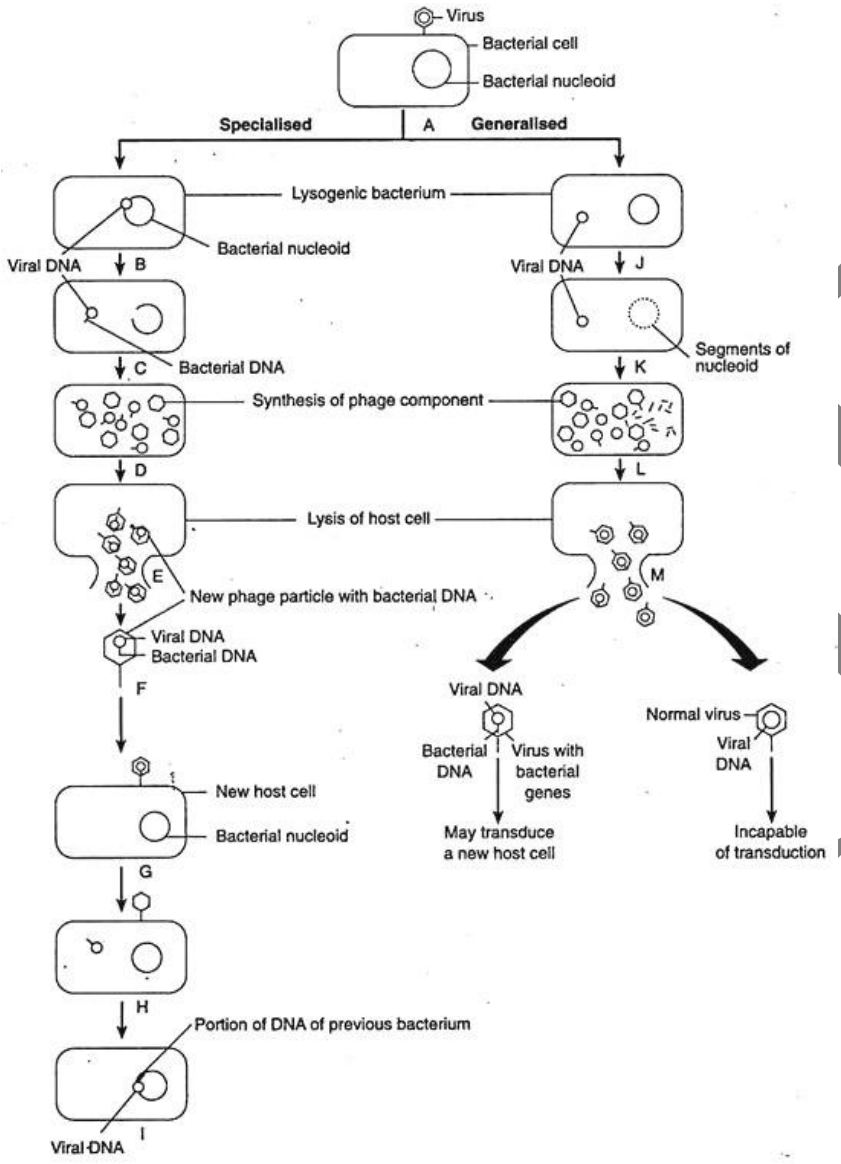


Fig. 2.32 : Diagrammatic representation of Transduction

Identification of recombinant phages

Phage introduction is the process of transfection which is equivalent to transformation excepting that a phage is used instead of bacterial plasmid. In vitro packaging, a vector uses lambda or M 13 phages to produce phage plaques which contain recombinant DNA. The recombinant DNA can be identified using various selection methods. For the first time bacteriophage was used to transfer the foreign DNA into E. coli cells.

If the vector is bacteriophage, its replication in bacterial host would result in phage particles, each carrying an identical copy of target gene. When phage vectors are used, a population of cells is infected with the viruses and virus replication proceeds spontaneously.

Eventually, the phage DNA containing the target gene is inserted into the bacterial chromosome where it will replicate as though it was a part of normal chromosome. Thus these vectors along with target genes are introduced into a bacterial host. This is usually achieved with the enzyme DNA ligase. It is important for ligation that vector DNA and the target DNA have been cut by the same restriction endonuclease.

Introduction of DNA into animal cells

Microinjection

Delivery of nucleic acids to protoplasts or intact cells via microinjection is a labour intensive procedure that requires special capillary needles, pumps, micromanipulators, inverted microscope and other equipment. However, injection into the nucleus or cytoplasm is possible and cells can be cultured individually to produce callus or plants.

In this way selection of transformants by drug resistance or marker genes may be avoided. This method involves skill of the worker to insert needle into the cytoplasm or in the nucleus. The basic technique is similar to that used for animal cell microinjection. In order to microinject protoplasts or other plant cells, the cells need to be immobilized (Fig. 16.3).

The cells are immobilized by:

1. The use of a holding pipette which holds the cells by vacuum.
2. Attachment of cells to poly-L-lysin coated cover slips.
3. Embedding the cells in agarose, agar or sodium alginate.

Glass micropipette are prepared to have openings of about 0.3 µM in diameter and are inserted into plant cell cytoplasm and nuclei with the aid of a micromanipulators device. A syringe like device is used for the controlled delivery of volume (10^{-11} – 10^{-4} 1) into the plant cell.

Most plant cells are injected while keeping inside micro-droplets (2-50 µl) of medium using a chamber which is sterile, vibration free and permits temperature and humidity regulation. A maximum of 100-200 cells per hour can be microinjected by this method.

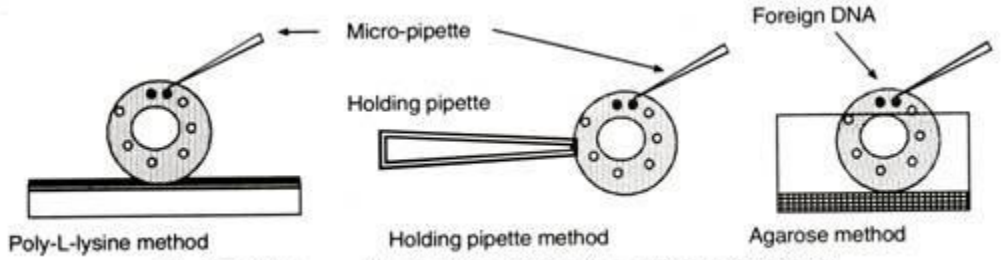


Fig. 16.3. Various methods of immobilizing the cell and microinjection.

The recovery of trans-formants is dependent upon the regeneration ability of the microinjected cells. Different methods have been used to grow injured (microinjected) single cells or protoplasts. Hanging droplets, covered under thin layer of agar or agarose, and micro-culture have been used (Fig. 16.4). Attempts have been made to inject linear or super-coiled DNA, in cytoplasm or in nucleus. Nuclear injections are found better for transformations.

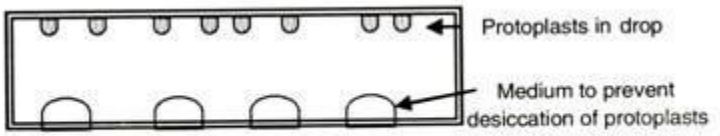


Fig. 16.4. Hanging drop culture method.

Shot-Gun methods

In recent years, it has been shown that DNA delivery to plant cells is also possible, when heavy metallic pellets (tungsten or gold) coated with the DNA of interest are accelerated to a very high initial velocity (1,400 ft/sec.)

These microprojectiles, normally 1-3 nmm diameter, are carried by a macroprojectile' or the **"bullet"** and accelerated into living plant cells (target cells can be pollen, cultured cells, cells in differentiated tissue and meristems) so that they can penetrate cells walls of intact tissue.

The acceleration is achieved either by an explosive charge (cordite explosion) or by using shock waves initiated by a high-voltage electric discharge. The design of two particle guns used for acceleration of microprojectiles are shown in Fig. 11.17.

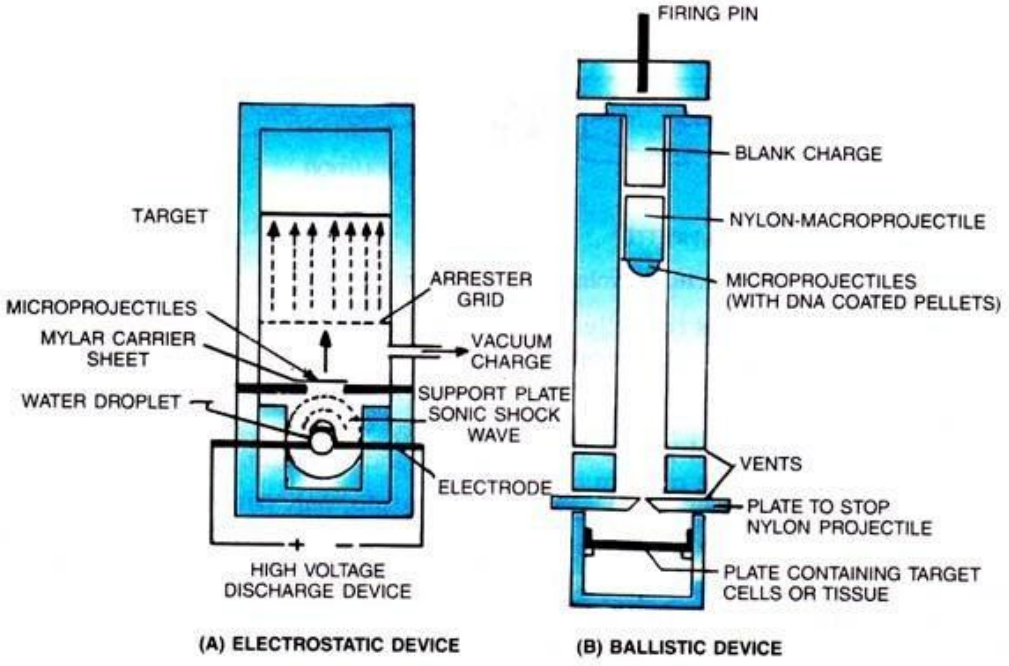


Fig. 11.17. Microprojectile acceleration devices.

The advantages of this method over microinjection include:

(i) Thousands of particles are accelerated at the same time, causing multiple hits resulting transfer of genes into many cells simultaneously.

(ii) Since inact cells can be used, some of the difficulties encounter with the use of protoplasts are automatically circumvented.

(iii) The method is universal in this application, so that cell type, size and shape or the presence/absence of cell walls do not significantly alter its effectiveness. In view of this, particle

bombardment method using microprojectiles has a great promise in a variety of plant species, particularly the cereals.

The phenomenon was first discovered in 1928 from Griffiths experiment with Diplococcus pneumoniae. As this historical experiment has been instrumental for identification of DNA as a hereditary material. Briefly, when a suspension containing a mixture of heat-killed virulent, encapsulated cells and live, non-virulent, non-encapsulated cells was injected into a mouse, a small fraction of the live bacteria became transformed into the virulent encapsulated type.

The transforming ability was inherited by the descendants of the newly transformed live strain. It was inferred that when cells of the virulent strain are killed by heat, their chromosomal material which is somehow liberated from heat-killed cells can pass through the cell wall of the living cells and become incorporated in the host chromosome. Although this experiment involved genes that control presence/absence of the capsule, later on genes controlling other characters could also be transformed by addition of chromosome fragments.

Transformation has proved useful in locating genes in bacteria like Bacillus subtilis, E. coli, Hemophilus influenzae, Rhizobium, Neisseria and others. Attempts have been made to find out if transformation occurs in higher organisms including mammals and man.

All results have been so far negative, except for some special cases where viral chromosome, for example of SV40 (simian virus), which can transform normal human cells in culture into cancerous ones. The mechanism of transformation is not fully known.

Electroporation

Electroporation is a physical transfection method that uses an electrical pulse to create temporary pores in cell membranes through which substances like nucleic acids can pass into cells. It is a highly efficient strategy for the introduction of foreign nucleic acids into many cell types, including bacteria and mammalian cells.

This method is based on the use of the short electrical pulses of high field strength. Electroporation causes the uptake of DNA into protoplasts by temporary permeabilization of the plasma membrane to macromolecules. Protoplasts and foreign DNA are placed in a buffer between two electrodes and a high intensity electric current is passed, the alternating current of about 1 MHz is applied to align the protoplast by di-electrophoresis.

Once aligned, fusion is induced by applying one or more direct current pulses (1-3kV /cm, 10-100 (is), then the alternating field is reapplied briefly to maintain close membrane contact for fusion (Fig. 16.1). Electric field damages membranes and creates pores in membranes. DNA diffuses through these pores immediately after the electric field is applied, until the pores are resealed. Technique is optimized by using appropriate electric field strength (defined as the applied voltage divides by the distance between two electrodes).

The optimum field strength is dependent on the followings:

1. The pulse length of electric current
2. Composition and temperature of the buffer solution
3. Concentration of foreign DNA in the suspension
4. Protoplasts density, and
5. Size of the protoplasts.

It has been demonstrated that the removal of pectin from the plant wall increases the amount of DNA which can be introduced by electroporation. Tobacco mosaic virus was introduced in tobacco protoplasts by this method. Electroporation has been used successfully for transient (when foreign gene which is present in cell but not integrated in the chromosome, shows expression in cytoplasm) and stable transformation (foreign gene integration in host chromosome and is expressed) of protoplasts from a wide range of species.

Plating efficiency (i.e., number of colonies recovered out of number of cells transferred on plates) of electroporated protoplasts grown on selection medium (containing selective marker) can be as high as 0.5%. The highest plant transformation efficiencies have been reported for tobacco, with 0.2% of electroporated leaf mesophyll protoplasts giving rise to transgenic calli. Low transformation efficiency is common in cereals, e.g., in rice 0.002% efficiency was recorded.

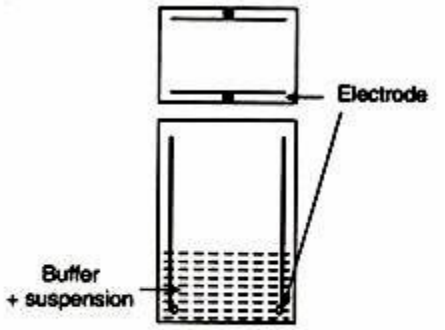


Fig. 16.1. Top (above) and side view (below) of glass cell with electrodes used for electroporation.

Working Protocol

Electroporation is based on a simple process. Host cells and selected molecules are suspended in a conductive solution, and an electrical circuit is closed around the mixture. An electrical pulse at an optimized voltage and only lasting a few microseconds to a millisecond is discharged through the cell suspension. This disturbs the phospholipid bilayer of the membrane and results in the formation of temporary pores. The electric potential across the cell membrane simultaneously rises to allow charged molecules like DNA to be driven across the membrane through the pores in a manner similar to electrophoresis.

The main advantage of electroporation is its applicability for transient and stable transfection of all cell types. Furthermore, because electroporation is easy and rapid, it is able to transfect a large number of cells in a short time once optimum electroporation conditions are determined. The major drawback of electroporation is substantial cell death caused by high voltage pulses and only partially successful membrane repair, requiring the use of greater quantities of cells compared to chemical transfection methods. While more modern instrumentation, such our Neon® Transfection System, overcome high cell mortality by distributing the electrical pulse equally among the cells and maintaining a stable pH throughout the electroporation chamber, optimization of pulse and field strength parameters is still required to balance the electroporation efficiency and cell viability.

Electroporation is based on the principle that high voltage electric pulses can induce cell plasma membranes to fuse. Thus, electroporation is a technique involving electric field-mediated membrane permeabilization. Electric shocks can also induce cellular uptake of exogenous DNA (believed to be via the pores formed by electric pulses) from the suspending solution.

Electroporation is a simple and rapid technique for introducing genes into the cells from various organisms (microorganisms, plants and animals).

The basic technique of electroporation for transferring genes into mammalian cells is depicted in Fig. 6.11. The cells are placed in a solution containing DNA and subjected to electrical shocks to cause holes in the membranes. The foreign DNA fragments enter through the holes into the cytoplasm and then to nucleus.

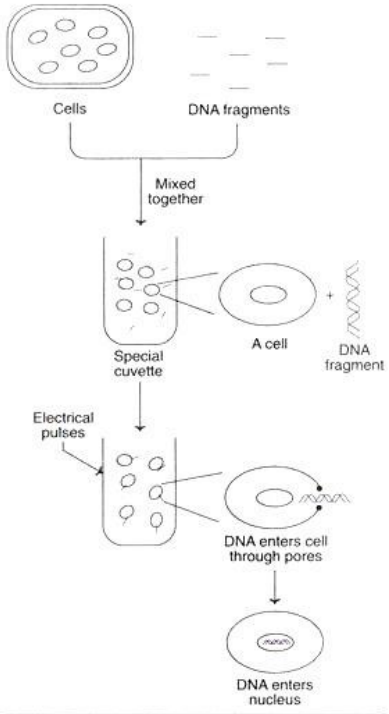


Fig. 6.11 : Gene transfer by electroporation.
(Note : Magnification depicted on right side)

Electroporation is an effective way to transform E.coli cells containing plasmids with insert DNAs longer than 100 kb. The transformation efficiency is around 10^9 transformants per microgram of DNA for small plasmids (about 3kb) and about 10^6 for large plasmids (about 130 kb).

Methods for clone identification

The restriction digest always contains an assortment of donor DNA fragments depending on the number of restriction sites of the enzyme used. A specific gene that is desired to be cloned may be present on a single fragment, or more often on several fragments. When the assortments of fragments are cloned into the vector DNA, only a few of the recombinants contain the desired gene or part of the gene.

The problem of selection

Detection of such recombinants presents a problem. Several techniques have been developed to tackle this problem. One solution is to identify the desired recombinants directly with the help of the marker genes, or indirectly by the technique of colony-hybridization using a radioactive probe of the specific DNA or RNA. Another approach is to use a pre-selected DNA segment containing the desired gene for cloning.

This can be prepared from an m-RNA by reverse transcription (RT). The DNA so produced can be multiplied by the polymerase chain reaction (PCR). The combination of the two is called RT-PCR technique. A desired gene can also be isolated by the Southern blotting technique.

Direct Selection:

A recombinant clone can be selected with the help of marker genes present in the vector as well as in the donor DNA. An example of direct selection can be cloning of threonine synthesis gene (thr^+) in a thr^+-amp^s host through a vector containing an ampicillin-resistance gene (amp^r). In this case, the objective would be to select a clone containing the thr^+ gene.

Transformation with recombinant DNA will form some host cells which have the thr^+-gene. The non-trans-formant host cells fail to grow in a medium which contains ampicillin but no threonine. Also, those host cells which take up the vector DNA other than thr^+ gene fail to grow in such a medium, because threonine is absent, though such host cells possess the amp' gene carried by the vector.

Only those trans-formants are able to form colonies on such a selective medium which have taken up the thr^+ gene from the donor DNA fragment and amp' gene of the vector. The presence of these genes in a recombinant host enables it to synthesise threonine and to destroy ampicillin. As a result, such a trans-formant can form a clone of cells.

This is diagrammatically represented in Fig. 9.136:

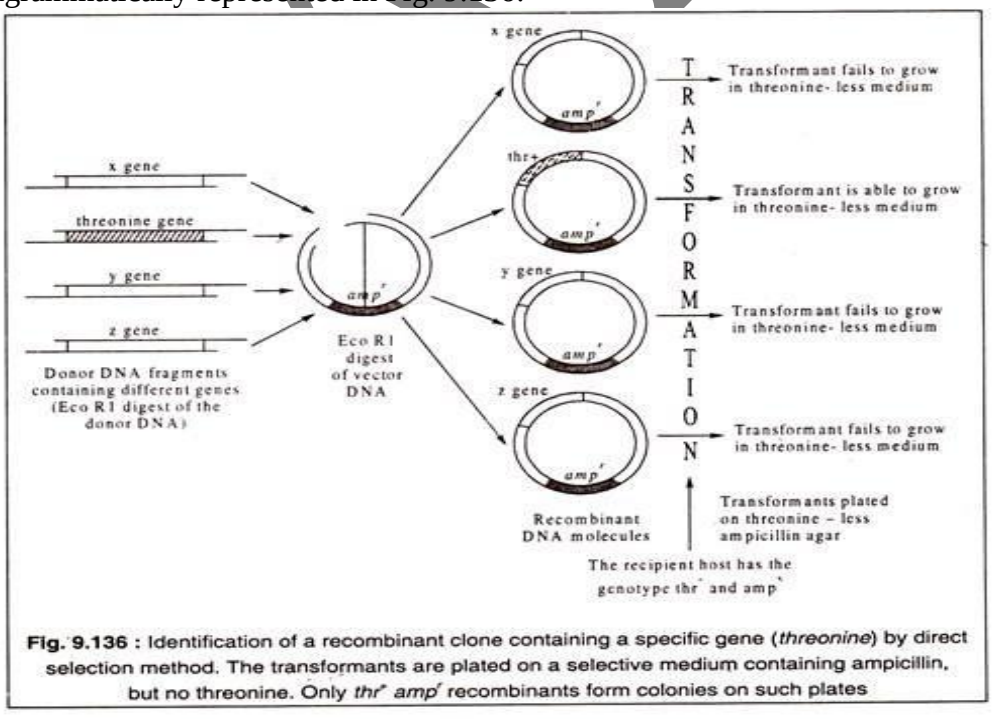


Fig. 9.136 : Identification of a recombinant clone containing a specific gene (*threonine*) by direct selection method. The transformants are plated on a selective medium containing ampicillin, but no threonine. Only *thr*$^+$ *amp*^{r} recombinants form colonies on such plates

Marker rescue

A technique to localize the site of a mutation within a gene. A polynucleotide, the 'marker', from the known wild-type sequence is allowed to recombine with the mutant genome such that, if it includes the site of the mutation, it 'rescues' the mutant by causing it to revert.

Gene Libraries

A genomic library is a collection of independently isolated vector linked DNA fragments derived from a single organism. It contains at least one copy of every DNA sequence in the genome. An ideal library is one that represents all of the sequences with smallest possible number of clones.

The genomic DNA libraries can be prepared by the complete digestion of the total genomic DNA with a restriction enzyme and the fragments are inserted into a suitable vector like X phages (Fig. 22.11). The drawback of this method is that sometimes the sequence of interest may contain multiple restriction sites, so digestion with RE results into two or more pieces.

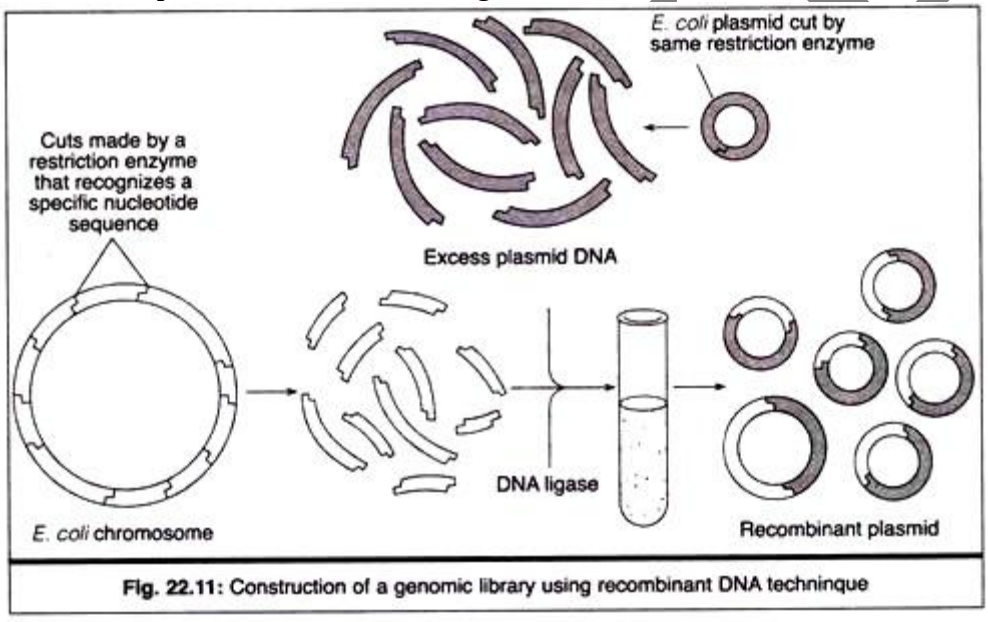


Fig. 22.11: Construction of a genomic library using recombinant DNA techninque

Identification of a clone from gene library

In this method the eukaryotic DNA is broken up into smaller fragments, thus an entire library would necessarily contain a large number of phages, and screening of which is very laborious. The problems of this method can be avoided by random shearing of total DNA and cloning of large fragments.

This method ensures that sequences are not excluded from the cloned library simply because of the distribution of restriction sites. In this procedure the randomly fragmented DNA is partially digested with RE which has short recognition sites. The fragments of desired size are collected through agarose gel electrophoresis, so the population of overlapping fragments that are close to random can be cloned directly.

Colony and plaque hybridization probing

Colony Hybridization

By this procedure, it is possible to detect the colonies of recombinants which have taken up a desired DNA segment containing a gene of choice. For detection of such colonies a sample of

the DNA containing the particular gene which has been labelled with radioactivity must be available. Alternatively, a sample of radioactive m-RNA of that gene may also be used.

After the host cells are transformed by the recombinant vector DNA, they are plated on agar to allow formation of individual discrete colonies. Among these colonies, the majorities are produced by trans-formants without the desired gene, and only very few might contain it. An impression of the colonies is next transferred to a nitrocellulose filter.

They are then treated with alkali (sodium hydroxide) which causes lysis of the cells. At the same time, the DNA molecules liberated from the cells are denatured in situ to produce single-strands. At the next step, the radioactive DNA probe, denatured to produce single-strands or the radioactive m-RNA is added to the filter.

The single-stranded probe DNA or RNA is allowed to hybridize with the single stranded DNA of the colonies. After removing the unbound radioactive nucleic acid by washing, the filter is kept in contact with an X-ray film for a few weeks to detect the presence of colonies which have taken up the desired gene (auto-radiographic technique).

The procedure is schematically shown in Fig. 9.137:

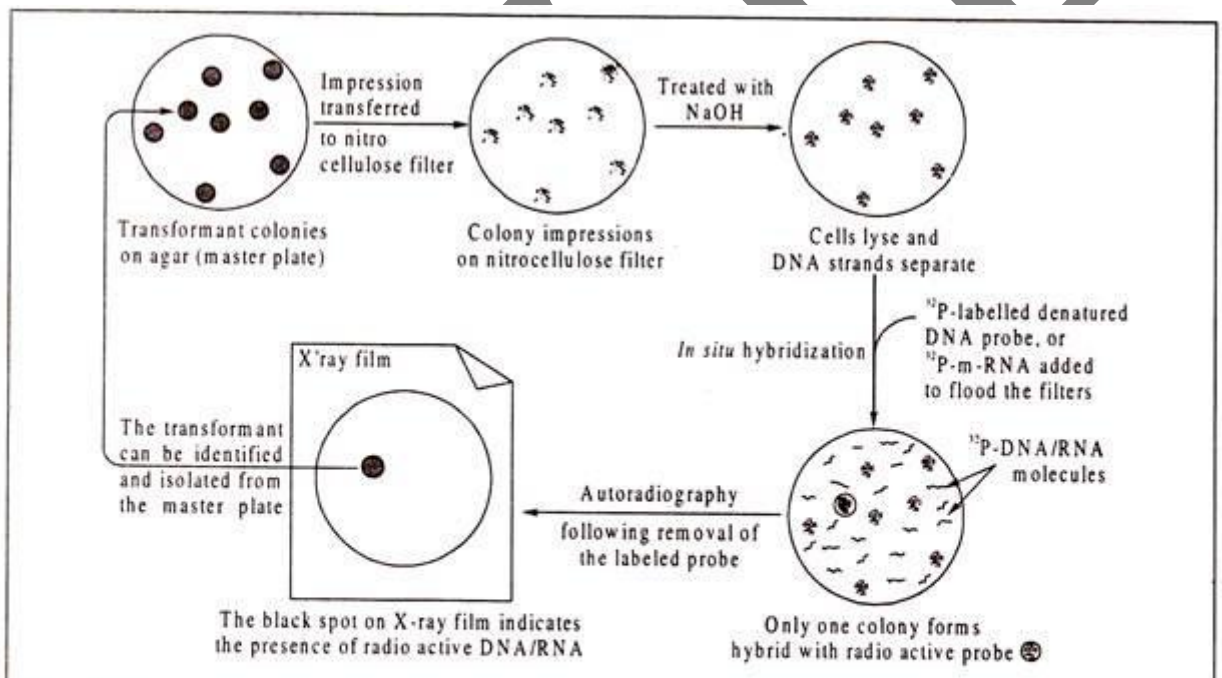


Fig. 9.137 : The procedure of colony hybridization. By this method it is possible to identify a recombinant colony containing a specific gene amongst a large number of colonies with the help of a radioactive probe which contains the DNA of the specific gene. The single-stranded probe DNA or RNA hybridizes with the denatured DNA and the radioactive colony can be detected by autoradiography. The bacterial colony or colonies containing the desired gene can be recovered from the master plate

Plaque hybridization probing

Plaque hybridization, also known as *Plaque lift*, was developed by Benton and Davis in 1977 and employs a filter lift method applied to phage plaques. This procedure is successfully applied to the isolation of recombinant phage by nucleic acid hybridization and probably is the most widely applied method of library screening. The method of screening library by plaque hybridization is described below-

• The nitrocellulose filter is applied to the upper surface of agar plates, making a direct contact between plaques and filter.

• The plaques contain phage particles, as well as a considerable amount of unpackaged recombinant DNA which bind to the filter.

• The DNA is denatured, fixed to the filter, hybridized with radioactive probes and assayed by autoradiography.

Advantages

• This method results in a 'cleaner' background and distinct signal (less background probe hybridization) for λ plaque screening due to less DNA transfer from the bacterial host to the nitrocellulose membrane while lifting plaques rather than bacterial colonies.

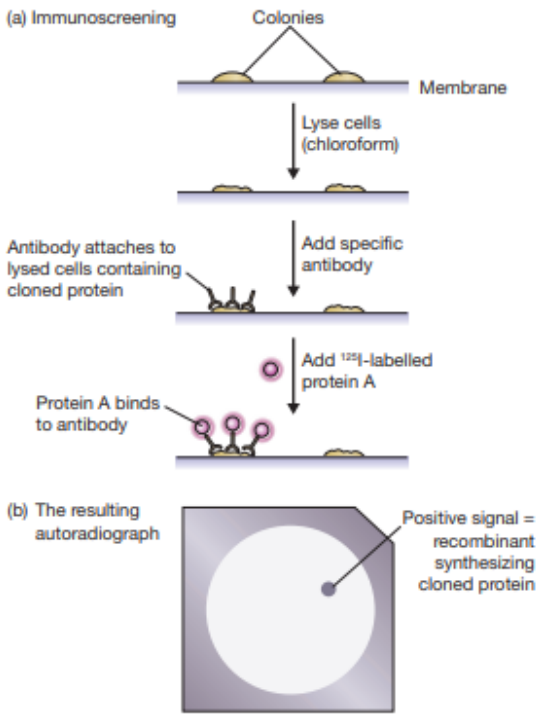
• Multiple screens can be performed from the same plate as plaques can be lifted several times.

• Screening can be performed at very high density by screening small plaques. High-density screening has the advantage that a large number of recombinant clones can be screened for the presence of sequences homologous to the probe in a single experiment.

Methods based on detection of the translation product of the cloned gene

Hybridization probing is usually the preferred method for identification of a particular recombinant from a clone library. The technique is easy to perform and, with modifications introduced in recent years, can be used to check up to 10,000 recombinants per experiment, allowing large genomic libraries to be screened in a reasonably short time. Nevertheless, the requirement for a probe that is at least partly complementary to the desired gene sometimes makes it impossible to use hybridization in clone identification. On these occasions a different strategy is needed. The main alternative to hybridization probing is immunological screening. The distinction is that, whereas with hybridization probing the cloned DNA fragment is itself directly identified, an immunological method detects the protein coded by the cloned gene. Immunological techniques therefore presuppose that the cloned gene is being expressed, so that the protein is being made, and that this protein is not normally present in the host cells. Antibodies are required for immunological detection methods If a purified sample of a protein is injected into the bloodstream of a rabbit, the immune system of the animal responds by synthesizing antibodies that bind to and help degrade the foreign molecule (Figure 8.20a). This is a version of the natural defense mechanism that the animal uses to deal with invasion by bacteria, viruses, and other infective agents. Once a rabbit is challenged with a protein, the levels of antibody present in its bloodstream remain high enough over the next few days for substantial quantities to be purified. It is not necessary to kill the rabbit, because as little as 10 ml of blood provides a considerable amount of antibody (Figure 8.20b). This purified antibody binds only to the protein with which the animal was originally challenged. Using a purified antibody to detect protein in recombinant colonies There are several versions of immunological screening, but the

most useful method is a direct counterpart of colony hybridization probing. Recombinant colonies are transferred to a polyvinyl or nitrocellulose membrane, the cells are lysed, and a solution containing the specific antibody is added (Figure 8.21a). In the original methods, either the antibody itself was labeled, or the membrane was subsequently washed with a solution of labeled protein A, a bacterial protein that specifically binds to the immunoglobulins that antibodies are made of (as shown in Figure 8.21a). In the more modern methods, the bound antibody—the primary antibody—is detected by washing the membrane with a labeled secondary antibody, which binds specifically to the primary antibody. Several secondary antibody molecules can bind to a single primary antibody molecule, increasing the amount of signal that is produced and enabling a clearer detection of each positive colony. In all three methods, the label can be a radioactive one, in which case the colonies that bind the label are detected by autoradiography (Figure 8.21b), or nonradioactive labels resulting in a fluorescent or chemiluminescent signal can be used.



POSSIBLE QUESTIONS

8 MARKS

1. Explain the methods of electroporation in detail.
2. Explain in detail about maximizing the expression of cloned genes in *E.coli.*
3. Explain the methods for to introduce DNA into the living cells in detail.
4. Write in detail about transformation and transfection in *E.coli*
5. Explain the methods for screening recombinants.
6. Write a detailed note on phage DNA introduction into prokaryotes.
7. Give a detailed account on gene libraries.
8. Write a detailed note on colony and plaque hybridization probing.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
III BSc BIOCHEMISTRY
17BCU601A-GENETIC ENGINEERING and BIOTECHNOLOGY
MULTIPLE CHOICE QUESTIONS

Unit-III

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	Which one of the following main constituent in the cell mass?	DNA	Ribosomes	Endoplasmic reticulum	Golgi apparatus	Ribosomes
2	Which virus without synthesizing the DNA intermediate?	Retrovirus	Tobacco viruses	Reoviruses	Polioviruses	Reoviruses
3	One of the final step in the transfer of genetic information?	Transcription	Transformation	Translation	Translocation	Translation
4	Microinjection in mice was performed at	Male pronuclei	Female pronuclei	Blastocyst	Pronucleus	Male pronuclei
5	DNA sequencing was developed by	Watson and Crick	Maxim and Gilbert	Jacob and Monod	Sabin and Salk	Maxim and Gilbert
6	DNA micro array used to determine	Genes present in cell	Genes to be transcribed	Genes that are silenced	Genes that are inactivated	Genes to be transcribed
7	Enzyme used in DNA foot printing is	DNAse I	Reverse transcriptase	RNAse I	DNA polymerase	DNAse I
8	In this technique the interaction with regulatory protein protects the DNA in the region of control sequence from digestive enzyme	DNA foot printing	DNA finger printing	Protein profiling	RNA profiling	DNA foot printing
9	Micro satellites are also known as	RFLP	VNTR	SNP	STR	STR
10	cDNA can be obtained from	DNA	mRNA	tRNA	rRNA	mRNA
11	DNA probe is usually labeled with _____	P^{32}	S^{16}	H^{1}	C^{14}	P^{32}
12	SNP refers to	Single nucleotide position	Single nucleotide polymorphism	Synthetic nucleotide probe	Serine natural protease	Single nucleotide polymorphism
13	Dideoxynucleoside triphosphates (ddNTPs) are used in sequencing DNA because:	ddNTPs are fluorescent.	ddNTPs are incorporated very efficiently into DNA by DNA polymerase	ddNTPs cannot be incorporated into DNA by DNA polymerase.	ddNTPs prevent further DNA synthesis once they are incorporated into the DNA sequence.	ddNTPs cannot be incorporated into DNA by DNA polymerase
14	Antisense refers to short	DNA sequences	RNA sequences	Both DNA and RNA se	Any chemical substances	Both DNA and RNA sequences
15	_____ are designed to be complementary to a s	Oligonucleotides	Gene fragments	Similar sequences	Antisense oligonucleotide	Antisense oligonucleotide
16	Nucleic acid hybridization is the most commonly us	Rapid	screen cDNA libraries	identifies clones	restriction enzymes	Rapid
17	A cDNA version of a gene includes	Codons for a mature mRNA	Sequences corresponding to promoters	Sequences corresponding to introns	both b and c.	Codons for a mature mRNA
18	Inducer in Blue white selection is	IPTG	ITPG	IGPT	IPGT	IPTG
19	Lactose analogue in Blue white selection is	G-gal	X-gal	I-gal	P-gal	X-gal
20	In Blue white selection is recombinants produce _____ colonies	Blue	White	Blue-White	Blue-Blue	White
21	The southern hybridization results was read by	UV transilluminator	Autoradiography	Stainiing	Visible light	Autoradiography
22	A southern blot is a technique used to detect	proteins	DNA	RNA	None	DNA
23	Technique used in DNA fingerprinting	Western blotting	Southern blotting	Variable blotting	Eastern blotting	Southern blotting
24	The technique doesn't involve electrophoresis for th	Dot blotting	Southern blotting	Northern blotting	Western blotting	Dot blotting
25	In blotting Labeled antibody are used to detect	DNA	RNA	protein	Lipid	Protein
26	Northern blots probes are	DNA and RNA	DNA	RNA and protein	RNA	RNA
27	Probe is a	Protein	Short piece of labeled DN	Short piece of labeled pr	Enzyme	Short piece of labeled DNA
28	The probes can be Non-radiolabelled by	Biotin	Toluene	Epinephrin	Globulins	Biotin
29	Probes are short nucleotide of ---------- in len	15 - 20 base pair	20 - 40 base pair	80 base pair	100 base pair	15 - 20 base pair
30	The molecular target for recognition in immunolog	Paratope	Epitope	ab	antigen	Paratope
31	Nitrocellulose membrane is used for	Eastern blotting	Southern blotting	Northern blotting	Both B and C	Southern blotting
32	Method by which recombinants phage DNA introduced in bacterial cell	*In vitro* packaging	Tranduction	Transformation	Transdution	*In vitro* packaging
33	Identification of recombinant phage is achieved by all except_____	LacZ	λ cI gene	Poly linker site	Spi phenotype	Poly linker site
34	_____ is gene that helps in the identification of recombinants	Molecular marker	Selectable marker	Identification Marker	Suitable marker	Selectable marker
35	During high levels of expression recombinant protei	Inclusion bodies	perfussion bodies	forign bodies	macro bodies	Inclusion bodies
36	Southern blotting techniques was developed by	Southern , 1975	Southern, 1971	Southern, 1961	Southern, 1969	Southern , 1975
37	In Southern blotting, the denatured DNA molecules are transferred to _____ membrane	Nitro acetate	Nitro Cellulose	Nitro cellulase	Nitro glycerine	Nitro Cellulose
38	Western blotting is used to analyse the expression	DNA	RNA	Protein	Aminoacid	Protein
39	Techniques useful in understanding the function of gene is	Southern blotting	Northern blotting	Western blotting	DNA micro array	Western blotting
40	Important method for the identification of DNA in any sample is	Southern blotting	Northern blotting	Western blotting	Dot-Blotting	Southern blotting
41	What is the correct sequence of events in Southern blotting?	Hybridization of DNA fragments with a labelled probe sequence followed by separation by electrophoresis and then transfer to a membrane	Separation of DNA fragments by electrophoresis followed by hybridization with a labelled probe sequence and then transfer to a membrane	Separation of DNA fragments by electrophoresis followed by transfer to a membrane and then hybridization with a labelled probe sequence	Transfer of DNA fragments to a membrane followed by separation by electrophoresis and then hybridization with a labelled probe sequence	Separation of DNA fragments by electrophoresis followed by transfer to a membrane and then hybridization with a labelled probe sequence
42	Useful method for mapping restriction sites around a single copy gene sequence	Western blotting	ELISA	Southern blotting	Northern blotting	Southern blotting
43	Recombinant insulin was expressed in	Bacterial expression systen	Plant expression system	Insect expression system	Mammalian expression syste	Mammalian expression system
44	For nucleic acid hybridization nucleic acid is digeste	One	Two	More than two	Other than restriction enzym	More than two
45	_____ refers to blotting of electrophoresed p	Southern blotting	Northern blotting	Western blotting	Dot blotting	Western blotting
46	In Northern blotting, the RNA molecules are transferred to	Nitro cellulose	Nylon Membrane	DBM	PEG	DBM
47	The commonly used membrane in northern blotting	Polyethylene and nylon	Nylon and nitrocellulose	Terylene and nitrocellulo	Nylon and terylene	Nylon and nitrocellulose
48	The original name of nylon is	Nylon	Polystyrene	Artificial wool	Filter 66	Polystyrene
49	The crucial step in northern blotting is	Isolation of m RNA	Transfer of denatured RN	Denaturation of mRNA	Seperation through AGE	Isolation of m RNA
50	DNA blotting technique is used to study about	DNA	RNA	Proteins	Lipids	DNA
51	The upward and backward transfer of DNA into me	Upward capillary transfer	Downward capillary transf	Bi-directional transfer	Vacuum transfer	Bi-directional transfer
52	The method used when the target DNA fragment wi	Capillary method	Vacuum blotting	Bi-directional blotting	Transfer in Alkaline buffer	Vacuum blotting
53	The rate of the movement of DNA is ----------	Slow	Fast	No change	Scattering	Slow
54	The mostly preferred buffer in AGE	TAE	TBE	TPE	TE	TAE
55	The DNA should move from	Negative to Positive	Positive to negative	Negative to neutral	Positive to neutral	Negative to Positive
56	Selection by complementation is the existence of	Auxotroph	Prototroph	Heterotroph	Lethal genes	Auxotroph
57	Which one of the following is not used in Non radio	Biotin	Horse radish peroxidase	Digitoxigenin	Cyanocobalamine	Cyanocobalamine

58	Joining of donor DNA fragments and vector DNA f	Gene cloning	Splicing	Gene manipulation	Moleccular cloning	Gene cloning
59	Which is the most specific recombinant library?	Genomic	protein	cDNA	cRNA	cDNA
60	Pores are formed when the voltage across a plasma	Exceeds its dielectric stren	Exceds its trielectric value	Falls below the dielectric	Maintains at the dielectric va	Exceeds its dielectric strength.

UNIT-IV
SYLLABUS

Polymerase chain reaction: Fundamentals of polymerase chain reaction, designing primers for PCR. Studying PCR products. Cloning PCR products. Real time PCR.

Expression of cloned genes : Vectors for expression of foreign genes in *E. coli*, cassettes and gene fusions. Challenges in producing recombinant protein in *E. coli*. Production of recombinant protein by eukaryotic cells. Fusion tags and their role in purification of recombinant proteins.

Polymerase Chain Reaction

The polymerase chain reaction (PCR) is a laboratory (in vitro) technique for generating large quantities of a specified DNA.

Obviously, PCR is a cell-free amplification technique for synthesizing multiple identical copies (billions) of any DMA of interest. Developed in 1984 by Karry Mullis PCR is now considered as a basic tool for the molecular biologist. As is a photocopier a basic requirement in an office, so is the PCR machine in a molecular biology laboratory!

Fundamentals of polymerase chain reaction

Principle of PCR:

The double-stranded DNA of interest is denatured to separate into two individual strands. Each strand is then allowed to hybridize with a primer (renaturation). The primer-template duplex is used for DNA synthesis (the enzyme- DNA polymerase). These three steps— denaturation, renaturation and synthesis are repeated again and again to generate multiple forms of target DNA.

Technique of PCR:

The essential requirements for PCR are listed below:

1. A target DNA (100-35,000 bp in length).
2. Two primers (synthetic oligonucleotides of 17-30 nucleotides length) that are complementary to regions flanking the target DNA.
3. Four deoxyribonucleotides (dATP, dCTP, dCTP, dTTP).
4. A DNA polymerase that can withstand at a temperature upto 95° C (i.e., thermo-stable).

The reaction mixture contains the target DNA, two primers (in excess), a thermo-stable DNA polymerase (isolated from the bacterium Thermus aquaticus (i.e., Taq DNA polymerase) and four deoxyribonucleoties. The actual technique of PCR involves repeated cycles for amplification of target DNA.

Each cycle has three stages:

1. Denaturation:

On raising the temperature to about 95° C for about one minute, the DNA gets denatured and the two strands separate.

2. Renaturation or annealing:

As the temperature of the mixture is slowly cooled to about 55° C, the primers base pair with the complementary regions flanking target DNA strands. This process is called renaturation or

annealing. High concentration of primer ensures annealing between each DNA strand and the primer rather than the two strands of DNA.

3. Synthesis:

The initiation of DNA synthesis occurs at 3′-hydroxyl end of each primer. The primers are extended by joining the bases complementary to DNA strands. The synthetic process in PCR is quite comparable to the DNA replication of the leading strand.

However, the temperature has to be kept optimal as required by the enzyme DNA polymerase. For Taq DNA polymerase, the optimum temperature is around 75° C (for E. coli DNA polymerase, it is around 37° C). The reaction can be stopped by raising the temperature (to about 95° C).

The 3 stages of PCR in relation to temperature and time are depicted in Fig. 8.1. Each cycle of PCR takes about 3-5 minutes. In the normal practice, the PCR is carried out in an automated machine.

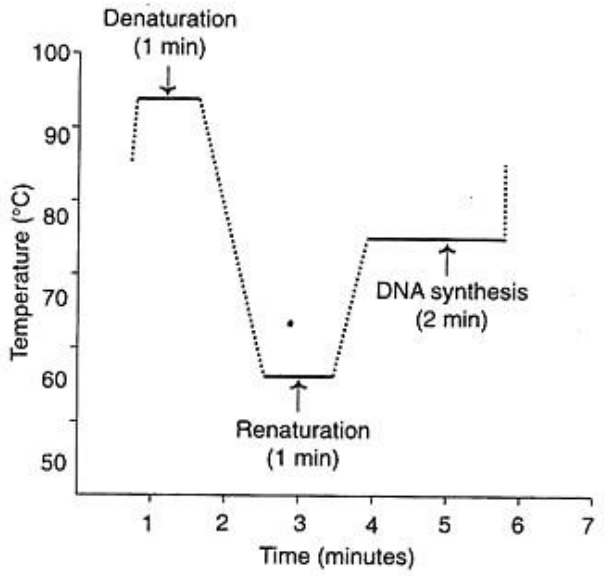


Fig. 8.1 : The three stages in each cycle of PCR in relation to temperature and time (Each cycle takes approximately 3-5 minutes).

As is evident from the Fig. 8.2 (cycle I), the new DNA strand joined to each primer is beyond the sequence that is complementary to the second primer. These new strands are referred to as long templates and they will be used in the second cycle.

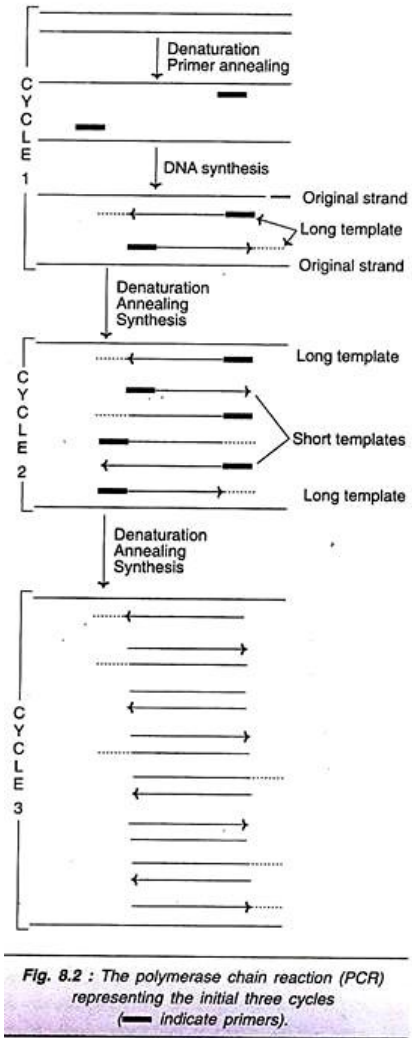


Fig. 8.2 : The polymerase chain reaction (PCR) representing the initial three cycles (— indicate primers).

For the second cycle of PCR, the DNA strands (original + newly synthesized long template) are denatured, annealed with primers and subjected to DNA synthesis. At the end of second round, long templates, and short templates (DNA strands with primer sequence at one end, and sequence complementary to the other end primer) are formed.

In the third cycle of PCR, the original DNA strands along with long and short templates are the starting materials. The technique of denaturation, renaturation and synthesis are repeated. This procedure is repeated again and again for each cycle. It is estimated that at the end of 32nd cycle of PCR, about a million-fold target DNA is synthesized (Table 8.1). The short templates possessing precisely the target DNA as double- stranded molecules accumulate.

Sources of DNA Polymerase:

In the original technique of PCR, Klenow fragment of E. coli DNA polymerase was used. This enzyme, gets denatured at higher temperature, therefore, fresh enzyme had to be added for each cycle. A breakthrough occurred with the introduction of Taq DNA polymerase from thermophilic bacterium, Thermus aquaticus. The Taq DNA polymerase is heat resistant; hence it is not necessary to freshly add this enzyme for each cycle of PCR.

Key Factors for Optimal PCR:

Primers:

Primers play a significant role in determining PCR. The primers (17-30 nucleotides) without secondary structure and without complementarity among themselves are ideal. The complementary primers can hybridize to form primer dimer and get amplified in PCR. This prevents the multiplication of target DNA.

DNA polymerase:

As already described, Taq DNA polymerase is preferred as it can withstand high temperature. In the hot-start protocol, DNA polymerase is added after the heat denaturation step of the first cycle. This avoids the extension of the mismatched primers that usually occur at low temperature.

Taq polymerase lacks proof reading exonuclease (3′-5′) activity which might contribute to errors in the products of PCR. Some other thermo-stable DNA polymerases with proof-reading activity have been identified e.g., Tma DNA polymerase from Thermotoga maritama; Pfu DNA polymerase from Pyrococcus furiosus.

Target DNA:

In general, the shorter the sequence of target DNA, the better is the efficiency of PCR. However, in recent years, amplification of DNA fragments up to 10 kb has been reported. The sequence of target DNA is also important in PCR. Thus, CC-rich regions of DNA strand hinder PCR.

Promoters and inhibitors:

Addition of proteins such as bovine serum albumin (BSA) enhances PCR by protecting the enzyme DNA polymerase. Humic acids, frequently found in archeological samples of target DNA inhibit PCR.

Variations of PCR:

The basic technique of the PCR has been described. Being a versatile technique, PCR is modified as per the specific demands of the situation. Thus, there are many variations in the original PCR; some of them are discussed, hereunder.

Nested PCR:

Sequence similarities between the target DNA and related DNA are very frequently seen. As a result of this, the primers may bind to both the DNAs and therefore even the undesired DNA also gets amplified in PCR. Use of nested primers increases the specificity of PCR, and selectively amplifies target DNA. Nested PCR is illustrated in Fig. 8.3. In the first cycle of PCR, the products are both from target DNA and undesired DNA. A second set of internal primers is now used. They will selectively bind to target DNA and amplification proceeds.

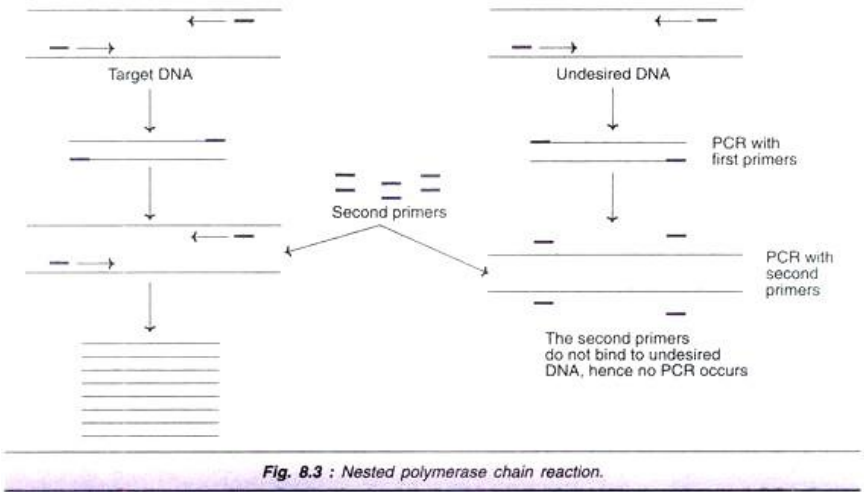


Fig. 8.3 : Nested polymerase chain reaction.

Inverse PCR:

In the inverse PCR, amplification of DNA of the unknown sequences is carried out from the known sequence (Fig. 8.4). The target DNA is cleaved with a restriction endonuclease which does not cut the known sequence but cuts the unknown sequence on either side. The DNA fragments so formed are inverted and get circularized (DNA ligase is employed as a sealing agent).

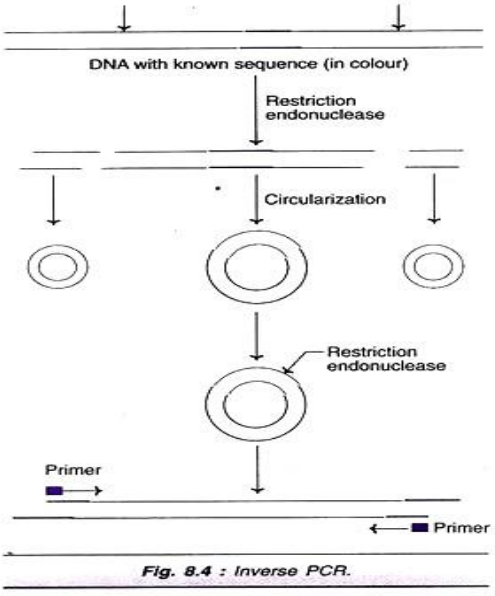


Fig. 8.4 : Inverse PCR.

The circle containing the known sequences is now cut with another restriction enzyme. This cleaves only the known sequence. The target DNA so formed contains the known sequence at both the ends with target DNA at the middle. The PCR amplification can now be carried out. It may be noted that the primers are generated in the opposite direction to the normal, since the original sequence is inverted during circularization.

Anchored PCR:

In the anchored PCR, a small sequence of nucleotides can be attached (tagged) to the target DNA i.e., the DNA is anchored. This is particularly useful when the sequence surrounding the target DNA is not known. The anchor is frequently a poly G tail to which a poly C primer is used. The anchoring can also be done by the use of adaptors. As the adaptors possess a known sequence, the primer can be chosen.

Reverse Transcription PCR:

The PCR technique can also be employed for the amplification of RNA molecules in which case it is referred to as reverse transcription — PCR (RT-PCR). For this purpose, the RNA molecule (mRNA) must be first converted to complementary DNA (cDNA) by the enzyme reverse transcriptase. The cDNA then serves as the template for PCR. Different primers can be employed for the synthesis of first strand of cDNA. These include the use of random primers, oligo dT primer and a sequence specific primer (Fig. 8.5).

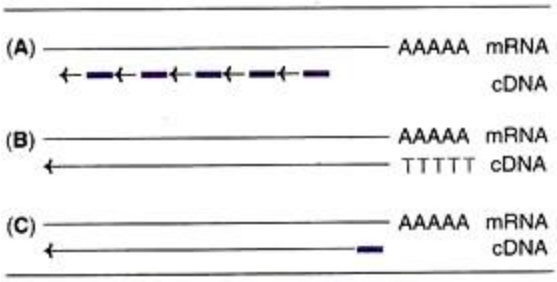


Fig. 8.5 : Synthesis of first strand of cDNA in reverse transcription-PCR with different primers (A) Random primers (B) oligo dT primer (C) Sequence specific primer (Note : The primers are shown in colour).

Asymmetric PCR:

PCR technique can also be used for the synthesis of single-stranded DNA molecules, particularly useful for DNA sequencing. In the asymmetric PCR, two primers in a ratio of 100: 1 are used. After 20-25 cycles of PCR, one primer is exhausted. The result is that in the next 5-10 PCR cycles, only single-stranded DNAs are generated.

Real-Time Quantitative PCR:

The quantification of PCR products in different cycles is not as simple as projected by theoretical considerations (Table 8.1). In practice, large variations occur. The most commonly used technique for measuring the quantity of PCR is by employing a fluorescence compound like eithidium bromide.

The principle is that the double-stranded DNA molecules bind to ethidium bromide which emit fluorescence that can be detected, and DNA quantified. The synthesis of genes by PCR and the role of PCR in site-directed mutagenesis are described elsewhere.

Random Amplified Polymorphic DNA (RAPD):

Normally, the objective of PCR is to generate defined fragments of DNA from highly specific primers. In the case of RAPD (pronounced as rapid), short oligonucleotide primers are arbitrarily selected to amplify a set of DNA fragments randomly distributed throughout the genome. This

technique, random amplified polymorphic DNA is also known as arbitrarily primed PCR (AP-PCR).

The procedure of RAPD is comparable to the general technique of PCR. This method basically involves the use of a single primer at low stringency. A single short oligonucleotide (usually a 9-10 base primer) binds to many sites in the genome and the DNA fragments are amplified from them. The stringency of primer binding can be increased after a few PCR cycles. This allows the amplification of best mismatches.

RAPD can be carefully designed so that it finally yields genome- specific band patterns that are useful for comparative analysis. This is possible since genomic DNA from two different individuals often produces different amplified patterns by RAPD. Thus, a particular DNA fragment may be generated for one individual and not for the other, and this represents DNA polymorphism which can be used as a genetic marker.

RAPD is widely used by plant molecular biologists for the genetic identification of plant species. For this purpose, different combinations of nucleotides, most of them random oligonucleotide primers have been designed and are commercially available. As each random primer anneals to a different region of DNA, many different regions of loci on the DNA can be identified. RAPD is thus useful for the construction of genetic maps and as a method for genomic fingerprinting.

Limitations of RAPD:

The main problem of RAPD is associated with reproducibility. It is often difficult to obtain similar levels of primer binding in different experiments. It is therefore difficult to correlate results obtained by different research groups on RAPD.

Amplified Fragment Length Polymorphism (AFLP):

AFLP is a very Sensitive method for detecting polymorphism in the genome. It is based on the principle of restriction fragment length polymorphism and RAPD. AFLP may be appropriately regarded as a diagnostic fingerprinting technique that detects genomic restriction fragments.

In the AFLP, PCR amplification rather than Southern blotting (mostly used in RFLP) is used for the detection of restriction fragments. It may be noted that AFLP is employed to detect the presence or absence of restriction fragments, and not the lengths of these fragments. This is the major difference between AFLP and RFLP. AFLP is very widely used in plant genetics.

It has not proved useful in the mapping of animal genomes, since this technique is mainly based on the presence of high rates of substitutional variations which are not found in animals. On the other hand, substitutional variations resulting in RFLPs are more common in plants. The basic principle of AFLP involves the amplification of subsets of RFLPs using PCR (Fig. 8.6).

A genomic DNA is isolated and digested simultaneously with two different restriction endonucleases — EcoRI with a 6 base pair recognition site and Msel with a 4 base pair recognition site. These two enzymes can cleave the DNA and result in small fragments (< 1 kb) which can be amplified by PCR. For this purpose the DNA fragments are ligated with EcoRI and Msel adaptors.

These common adaptor sequences (flanking genomic sequences) serve as primer binding sites on the restriction fragments. The DNA fragments can be amplified with AFLP primers each having only one selective nucleotide. These PCR products are diluted and used as templates for the selective amplification employing two new AFLP primers that have 2 or 3 selective nucleotides.

After the selective amplification by PCR, the DNA products are separated on a gel.

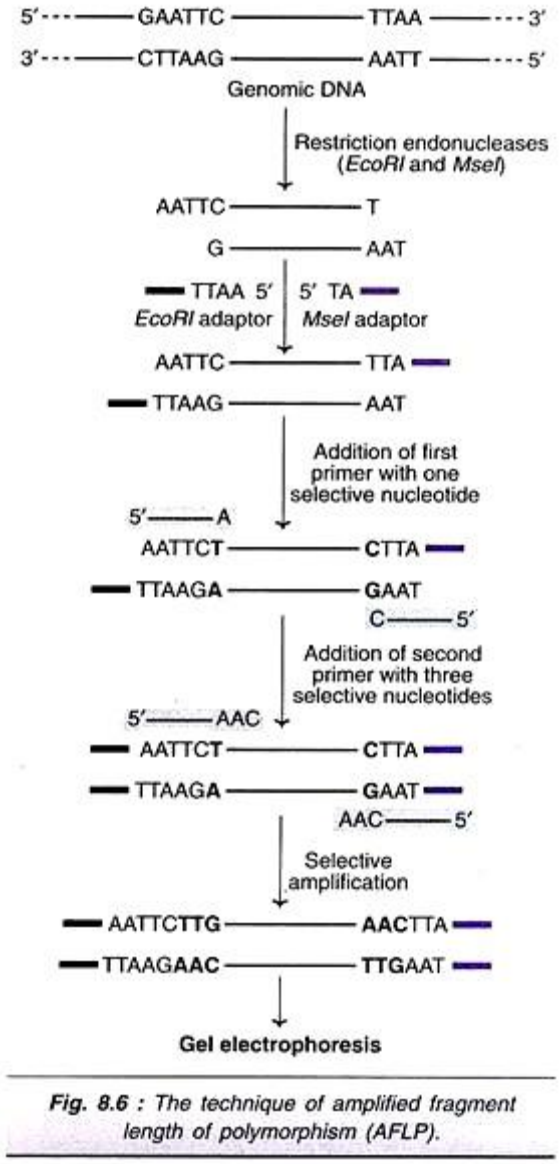


Fig. 8.6 : The technique of amplified fragment length of polymorphism (AFLP).

The resultant DNA fingerprint is identified by autoradiography. AFLP fragments represent unique positions in the genomes, and hence can be used as landmarks to bridge the gaps between genetic and physical maps of genomes. In plants, AFLP is useful to generate high density maps, and to detect genomic clones.

Rapid Amplification of cDNA Ends (RACE):

As already described (See p. 115), reverse transcription, followed by PCR (RT-PCR) results in the amplification of RNA sequences in cDNA form. But the major limitation of RT-PCR is related to incomplete DNA sequences in cDNA. This problem is solved by using the technique rapid amplification of cDNA ends. RACE is depicted in Fig. 8.7, and briefly described below.

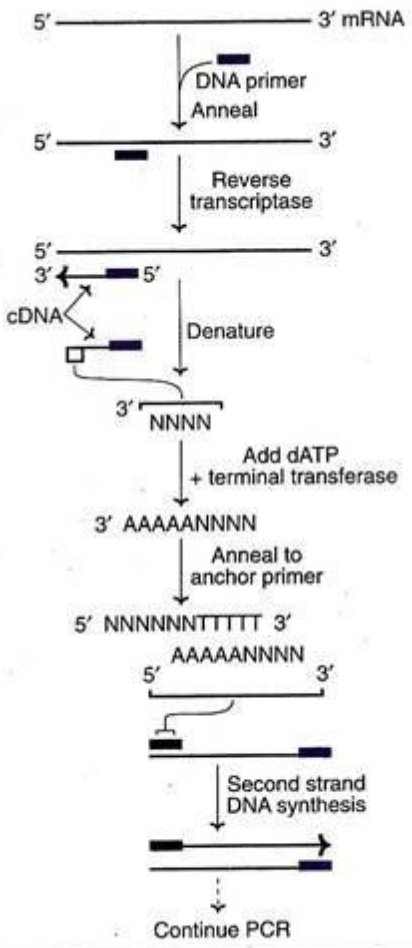


Fig. 8.7 : Rapid amplification of cDNA ends (RACE).

The target RNA is converted into a partial cDNA by extension of a DNA primer. This DNA primer was first annealed at an interval position of RNA, not too far from the 5′-end of the molecule. Now addition dATP (As) and terminal deoxynucleotidyl transferase extends the 3′-end of the cDNA.

This happens due to the addition of a series of as to the cDNA. These as series now act as the primer to anneal to the anchor primer. A second strand of DNA can be formed by extending the anchor primer. The double-stranded DNA is now ready for amplification by PCR. The above procedure described is called 5′- RACE, since it is carried out by amplification of the 5′-end of the starting RNA. Similar protocol can be used to carry out 3′-RACE when the 3′-end RNA sequence is desired.

Limitations of RACE:

Since a specific primer is used, the specificity of amplification of RACE may not be very high. Another disadvantage is that the reverse transcriptase may not fully reach the 5′-ends of RNA,

and this limits the utility of RACE. In recent years, some modifications have been done to improve RACE.

Applications

The top six applications are: (1) PCR in Clinical Diagnosis (2) PCR in DNA Sequencing (3) PCR in Gene Manipulation and Expression Studies (4) PCR in Comparative Studies of Genomes (5) PCR in Forensic Medicine and (6) PCR in Comparison with Gene Cloning.

1. PCR in Clinical Diagnosis:

The specificity and sensitivity of PCR is highly useful for the diagnosis of various diseases in humans. These include diagnosis of inherited disorders (genetic diseases), viral diseases, bacterial diseases etc. The occurrence of genetic diseases frequently identified by restriction fragment length polymorphism (RFLP) can be employed only when there is a mutation resulting in a detectable change in the length of restriction fragment. Many genetic diseases occur without the involvement of RFLP. For all such disorders, PCR technique is a real boon, as it provides direct information of DNA. This is done by amplification of DNA of the relevant region, followed by the direct analysis of PCR products.

Prenatal diagnosis of inherited diseases:

PCR is employed in the prenatal diagnosis of inherited diseases by using chorionic villus samples or cells from amniocentesis. Thus, diseases like sickle-cell anemia, p-thalassemia and phenylketonuria can be detected by PCR in these samples.

Diagnosis of retroviral infections:

PCR from cDNA is a valuable tool for diagnosis and monitoring of retroviral infections, e.g., HIV infection.

Diagnosis of bacterial infections:

PCR is used for the detection of bacterial infection e.g., tuberculosis by Mycobacterium tuberculosis.

Diagnosis of cancers:

Several virally-induced cancers (e.g., cervical cancer caused by human papilloma virus) can be detected by PCR. Further, some cancers which occur due to chromosomal translocation (chromosome 14 and 18 in follicular lymphoma) involving known genes are identified by PCR.

PCR in sex determination of embryos:

Sex of human and livestock embryos fertilized in vitro, can be determined by PCR, by using primers and DNA probes specific for sex chromosomes. Further, this technique is also useful to detect sex — linked disorders in fertilized embryos.

2. PCR in DNA Sequencing:

As the PCR technique is much simpler and quicker to amplify the DNA, it is conveniently used for sequencing. For this purpose, single-strands of DNA are required. In asymmetric PCR, preferential amplification of a single-strand is carried out. In another method, strand removal can be achieved by digesting one strand (usually done by exonuclease by its action on 5′-phosphorylated strand).

3. PCR in Gene Manipulation and Expression Studies:

The advantage with PCR is that the primers need not have complementary sequences for the target DNA. Therefore, the sequence of nucleotides in a piece of the gene (target DNA) can be manipulated and amplified by PCR.

By using this method, coding sequence can be altered (thereby changing amino acids) to synthesize protein of interest. Further, gene manipulations are important in understanding the effects of promoters, initiators etc., in gene expression.

PCR is important in the study of mRNAs, the products of gene expression. This is carried out by reverse transcription — PCR.

4. PCR in Comparative Studies of Genomes:

The differences in the genomes of two organisms can be measured by PCR with random primers. The products are separated by electrophoresis for comparative identification. Two genomes from closely related organisms are expected to yield more similar bands. For more details, refer the technique random amplified polymorphic DNA.

PCR is very important in the study evolutionary biology, more specifically referred to as phylogenetic. As a technique which can amplify even minute quantities of DNA from any source (hair, mummified tissues, bone, or any fossilized material), PCR has revolutionized the studies in palaentology and archaelogy. The movie 'Jurassic Park' has created public awareness of the potential applications of PCR!

5. PCR in Forensic Medicine:

A single molecule of DNA from any source (blood strains, hair, semen etc.) of an individual is adequate for amplification by PCR. Thus, PCR is very important for identification of criminals. The reader may refer DNA finger printing technique described elsewhere.

6. PCR in Comparison with Gene Cloning:

PCR has several advantages over the traditional gene cloning techniques .These include better efficiency, minute quantities of starting material (DNA), cost-effectiveness, minimal technical skill, time factor etc. In due course of time, PCR may take over most of the applications of gene cloning.

Designing primers for PCR

Good primer design is essential for successful reactions. The important design considerations described below are a key to specific amplification with high yield. The preferred values indicated are built into all our products by default.

1. Primer Length: It is generally accepted that the optimal length of PCR primers is 18-22 bp. This length is long enough for adequate specificity and short enough for primers to bind easily to the template at the annealing temperature.

2. Primer Melting Temperature: Primer Melting Temperature (T_m) by definition is the temperature at which one half of the DNA duplex will dissociate to become single stranded and indicates the duplex stability. Primers with melting temperatures in the range of 52-58 ^oC generally produce the best results. Primers with melting temperatures above 65^oC have a tendency for secondary annealing. The GC content of the sequence gives a fair indication of the primer T_m. All our products calculate it using the nearest neighbor thermodynamic theory, accepted as a much superior method for estimating it, which is considered the most recent and best available.

Formula for primer T_m calculation:

Melting Temperature $T_m(K)=\{\Delta H/ \Delta S + R \ln(C)\}$, Or Melting Temperature $T_m(^oC) = \{\Delta H/ \Delta S + R \ln(C)\} - 273.15$ where

ΔH (kcal/mole) : H is the Enthalpy. Enthalpy is the amount of heat energy possessed by substances. ΔH is the change in Enthalpy. In the above formula the ΔH is obtained by adding up all the di-nucleotide pairs enthalpy values of each nearest neighbor base pair.

ΔS (kcal/mole) : S is the amount of disorder a system exhibits is called entropy. ΔS is change in Entropy. Here it is obtained by adding up all the di-nucleotide pairs entropy values of each nearest neighbor base pair. An additional salt correction is added as the Nearest Neighbor parameters were obtained from DNA melting studies conducted in 1M Na+ buffer and this is the default condition used for all calculations.

ΔS (salt correction) = ΔS (1M NaCl)+ 0.368 x N x ln([Na+])

Where

N is the number of nucleotide pairs in the primer (primer length -1). [Na+] is salt equivalent in mM.

[Na+] calculation:

[Na+] = Monovalent ion concentration +4 x free Mg2+.

3. Primer Annealing Temperature: The primer melting temperature is the estimate of the DNA-DNA hybrid stability and critical in determining the annealing temperature. Too high T_a will produce insufficient primer-template hybridization resulting in low PCR product yield. Too low T_a may possibly lead to non-specific products caused by a high number of base pair mismatches,. Mismatch tolerance is found to have the strongest influence on PCR specificity.

$T_a = 0.3 \times T_m(\text{primer}) + 0.7\ T_m\ (\text{product}) - 14.9$

where,

T_m(primer) = Melting Temperature of the primers

T_m(product) = Melting temperature of the product

4. GC Content: The GC content (the number of G's and C's in the primer as a percentage of the total bases) of primer should be 40-60%.

5. GC Clamp: The presence of G or C bases within the last five bases from the 3' end of primers (GC clamp) helps promote specific binding at the 3' end due to the stronger bonding of G and C bases. More than 3 G's or C's should be avoided in the last 5 bases at the 3' end of the primer.

6. Primer Secondary Structures: Presence of the primer secondary structures produced by intermolecular or intramolecular interactions can lead to poor or no yield of the product. They adversely affect primer template annealing and thus the amplification. They greatly reduce the availability of primers to the reaction.

i) Hairpins: It is formed by intramolecular interaction within the primer and should be avoided. Optimally a 3' end hairpin with a ΔG of -2 kcal/mol and an internal hairpin with a ΔG of -3 kcal/mol is tolerated generally.

ΔG definition: The Gibbs Free Energy G is the measure of the amount of work that can be extracted from a process operating at a constant pressure. It is the measure of the spontaneity of the reaction. The stability of hairpin is commonly represented by its ΔG value, the energy

required to break the secondary structure. Larger negative value for ΔG indicates stable, undesirable hairpins. Presence of hairpins at the 3' end most adversely affects the reaction.
$\Delta G = \Delta H - T\Delta S$
ii) Self Dimer: A primer self-dimer is formed by intermolecular interactions between the two (same sense) primers, where the primer is homologous to itself. Generally a large amount of primers are used in PCR compared to the amount of target gene. When primers form intermolecular dimers much more readily than hybridizing to target DNA, they reduce the product yield. Optimally a 3' end self dimer with a ΔG of -5 kcal/mol and an internal self dimer with a ΔG of -6 kcal/mol is tolerated generally.

iii) Cross Dimer: Primer cross dimers are formed by intermolecular interaction between sense and antisense primers, where they are homologous. Optimally a 3' end cross dimer with a ΔG of -5 kcal/mol and an internal cross dimer with a ΔG of -6 kcal/mol is tolerated generally.

7. Repeats: A repeat is a di-nucleotide occurring many times consecutively and should be avoided because they can misprime. For example: ATATATAT. A maximum number of di-nucleotide repeats acceptable in an oligo is 4 di-nucleotides.
8. Runs: Primers with long runs of a single base should generally be avoided as they can misprime. For example, AGCGGGGGATGGGG has runs of base 'G' of value 5 and 4. A maximum number of runs accepted is 4bp.
9. 3' End Stability: It is the maximum ΔG value of the five bases from the 3' end. An unstable 3' end (less negative ΔG) will result in less false priming.
10. Avoid Template Secondary Structure: A single stranded Nucleic acid sequences is highly unstable and fold into conformations (secondary structures). The stability of these template secondary structures depends largely on their free energy and melting temperatures(T_m). Consideration of template secondary structures is important in designing primers, especially in qPCR. If primers are designed on a secondary structures which is stable even above the annealing temperatures, the primers are unable to bind to the template and the yield of PCR product is significantly affected. Hence, it is important to design primers in the regions of the templates that do not form stable secondary structures during the PCR reaction. Our products determine the secondary structures of the template and design primers avoiding them.
11. Avoid Cross Homology: To improve specificity of the primers it is necessary to avoid regions of homology. Primers designed for a sequence must not amplify other genes in the mixture. Commonly, primers are designed and then BLASTed to test the specificity. Our products offer a better alternative. You can avoid regions of cross homology while designing primers. You can BLAST the templates against the appropriate non-redundant database and the software will interpret the results. It will identify regions significant cross homologies in each template and avoid them during primer search.
Parameters for Primer Pair Design

1. Amplicon Length: The amplicon length is dictated by the experimental goals. For qPCR, the target length is closer to 100 bp and for standard PCR, it is near 500 bp. If you know the positions of each primer with respect to the template, the product is calculated as: Product length = (Position of antisense primer-Position of sense primer) + 1.
2. Product Position: Primer can be located near the 5' end, the 3' end or any where within specified length. Generally, the sequence close to the 3' end is known with greater confidence and hence preferred most frequently.
3. Tm of Product: Melting Temperature (T_m) is the temperature at which one half of the DNA duplex will dissociate and become single stranded. The stability of the primer-template DNA duplex can be measured by the melting temperature (T_m).
4. Optimum Annealing Temperature (T_a Opt): The formula of Rychlik is most respected. Our products use this formula to calculate it and thousands of our customers have reported good results using it for the annealing step of the PCR cycle. It usually results in good PCR product yield with minimum false product production.
T_a Opt = 0.3 x(T_m of primer) + 0.7 x(T_m of product) - 14.9
where
T_m of primer is the melting temperature of the less stable primer-template pair
T_m of product is the melting temperature of the PCR product.
5. Primer Pair Tm Mismatch Calculation: The two primers of a primer pair should have closely matched melting temperatures for maximizing PCR product yield. The difference of 5^oC or more can lead no amplification.

Studying PCR products

PCR is often the starting point for a longer series of experiments inwhich the amplification product is studied in various ways in order to gain information about the DNA molecule that acted as the original template.
Although a wide range of procedures have been devised for studying PCR products, three techni ques are particularly important:
•Gel electrophoresis of PCR products
•Cloning of PCR products
•Sequencing of PCR products.

Cloning of PCR products

Some applications require that after a PCR the resulting products are ligatedinto a vector and examined by any of the standard methods.
•This may sound easy, but there are complications.
•The first problem arises when the resulting PCR products are blunt-ended.
•Could be inserted into a cloning vector by blunt-end ligation, or alternatively the PCR products could be provided with sticky ends by the attachment of linkers or adaptors.
• Unfortunately, the situation is not so straightforward.
• Taq polymerase tends to add an additional nucleotide, usually an adenosine, to the end of each strand that it synthesizes.

• This means that a double stranded PCR product is not blunt-ended, and instead most 3′ terminal have a single nucleotide overhang

Real Time PCR

RT-PCR also known as quantitative PCR is used to amplify and simultaneously quantify a target DNA. It differs from standard PCR in a way that it can detect the amplified product as the reaction progresses with time but in standard PCR the amplified product is detected at the end of the reaction by agarose gel electrophoresis. RT-PCR is widely used in expression profiling, to determine the expression of a gene or to identify the sequence of an RNA transcript, including transcription start and termination sites. Many probes are used for detection of amplified product in RT-PCR such as TaqMan probe, Molecular beacons, ds DNA binding dyes (eg.SYBR green) of which TaqMan probes are most widely used. TaqMan probes are oligonucleotide probes which have a fluorophore at its 5' end and a quencher at its 3' end. The quencher molecule quenches the fluorescence emitted by the fluorophore when excited by the light source of the cycler via FRET (Fluorescence Resonance Energy Transfer). When DNA polymerase starts synthesizing the new strand, it degrades the probe (due to its 5'-3' exonuclease activity) which releases the fluorophore. This inhibits quenching effect of the quencher and allows fluorescence of the fluorophore. Hence fluorescence detected in RT-PCR is directly proportional to the amount of DNA template.

Expression of cloned genes :

Vectors for expression of foreign genes in *E. coli*

Cloning Vector:

By cloning, one can produce unlimited amounts of any particular fragment of DNA. In principle, the DNA isolated and cut pieces are introduced into a suitable host cell, usually a bacterium such as Escherichia coli, where it is replicated, as the cell grows and divides.

However, replication will only occur if the DNA contains a sequence which is recognized by the cell as an origin of replication. Since such sequences are infrequent, this will rarely be so, and therefore, the DNA to be cloned, has to be attached to a carrier, or vector DNA which does contain an origin of replication.

Criteria of an Ideal Vector:

Vectors are those DNA molecules that can carry a foreign DNA fragment when inserted into it. A vector must possess certain minimum qualifications to be an efficient agent for the transfer, maintenance and amplification of the passenger DNA.

1. The vector should be small and easy to isolate.
2. They must have one or more origins of replication so that they will stably maintain themselves within host cell.
3. Vector should have one or more unique restriction sites into which the recombinant DNA can be inserted.
4. They should have a selectable marker (antibiotic resistance gene) which allows recognition of transformants.
5. Vector DNA can be introduced into a cell.
6. The vector should not be toxic to host cell.

Types of Vector:

Based on the nature and sources, the vectors are grouped into bacterial plasmids, bacteriophages, cosmids and phagemids (Fig. 22.3).

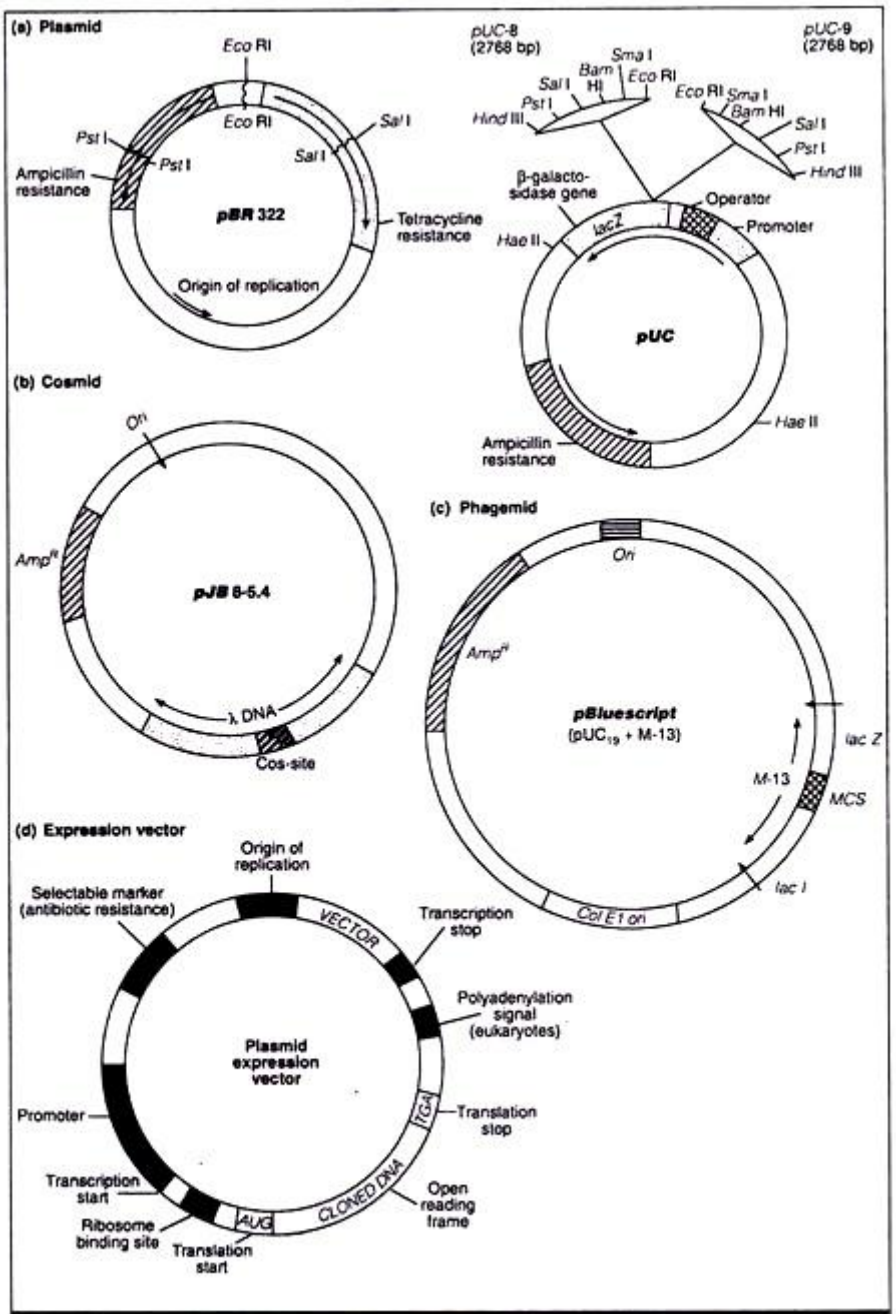


Fig. 22.3: Vectors (a–plasmid, b–cosmid, c–phagemid, d–expression vector)

(a) Plasmid:

Plasmids are the extra-chromosomal, self-replicating, and double stranded closed and circular DNA molecules present in the bacterial cell. A number of properties are specified by plasmids such as antibiotic and heavy metal resistance, nitrogen fixation, pollutant degradation, bacteriocin and toxin production, colicin factors, etc.

Plasmids have following advantages as cloning vehicle (Cohen et a. 1973):

1. It can be readily isolated from the cells.
2. It possesses a single restriction site for one or more restriction enzymes.
3. Insertion of foreign DNA does not alter the replication properties.
4. It can be reintroduced into cell.
5. Selective marker is present.
6. Transformants can be selected easily by using selective medium.
7. Multiple copy numbers are present in a cell.

Some plasmid vectors are pBR 322, pBR 327, pUC vectors, yeast plasmid vector and Ti, Ri plasmids. Ti and Ri Plasmids are widely used in plant system for genetic transformation.

Among higher plants, Ti plasmid of Agrobacterium tumefaciens or Ri plasmid of A. rhizogenes are the best known vectors. T-DNA, from Ti or Ri plasmid of Agrobacterium, is considered to be very potential for foreign gene transfer in cloning experiments with higher plants.

pBR 322 and pUC Vectors:

pBR322 is a derived plasmid from a naturally occurring plasmid col El, composed of 4362 bp DNA and its replication may be more faster. The plasmid has a point of origin of replication (ori), two selectable marker genes conferring resistance to antibiotics, e.g., ampicillin (amp^r), tetracycline (tet^r) and unique recognition sites for 20 restriction endonucleases.

Tetracycline resistance gene has a cloning site and insertion of foreign segment of DNA will inactivate the tet^r gene. The recombinant plasmid will allow the cells to grow only in presence of ampicillin but will not protect them against tetracycline .

Another plasmid used in gene cloning is pUC vector available in pairs with reverse orders of restriction sites relative to lac^z promoter. This is a synthesized plasmid possessing ampicillin resistance gene (amp^r), origin of replication from pBR322(on) and lac^z J gene from E. coli. pUC 8 and pUC 9 make one such pair.

(b) Bacteriophage:

The bacteriophage has linear DNA molecule, a single break will generate two fragments, foreign DNA can be inserted to generate chimeric phage particle. But as the capacity of phage head is limited, some segments of phage DNA, not having essential genes, may be removed. This technique has been followed in λ (Lambda) phage vectors to clone large foreign particle.

Plasmid can clone up to 20 to 25 kb long fragments of eukaryotic genome. The examples of different Lambda phage vectors are λ gt 10, λ gt 11, EMBL 3, etc. M-13 is a filamentous bacteriophage of E. coli whose single stranded circular DNA has been modified variously to give rise M-13 series of cloning vectors.

(c) Cosmid:

Cosmids are plasmid particles, into which certain specific DNA sequences, namely those for cos sites are inserted which enable the DNA to get packed in X particle. Like plasmids, the cosmids

perpetuate in bacteria without lytic development. The cosmids have high efficiency to produce a complete genomic library

(d) Phagemid:

These are prepared artificially by combining features of phages with plasmids. One commonly used phagemid is pBluescript IIKs derived from pUC-19.

(e) Plant and Animal Viruses:

A number of plant and animal viruses have also been used as vectors both for introducing foreign genes into cells and for gene amplification. Cauliflower Mosaic Viruses (CaMV), Tobacco Mosaic Viruses (TMV) and Gemini Virus are three groups of viruses that have been used as vectors for cloning of DNA segments in plant system. SV 40 (Simian Virus 40), human adenoviruses and retroviruses are potential as vectors for gene transfer into animal cells.

(f) Artificial Chromosomes:

Yeast Artificial Chromosome (YAC) or Bacterial Artificial Chromosome (BAC) vectors allow cloning of several hundred kb pairs which may represent the whole chromosome. It can be cloned in yeast or bacteria by ligating them to vector sequences that allow their propagation as linear artificial chromosome.

(g) Transposons:

Transposable elements like Ac-Ds or Mu-1 of Maize, P-element of Drosophila may also be used for cloning vector and transfer of gene among eukaryotes.

Expression Vector:

A vector that has been constructed in such a way that inserted DNA molecule is put under appropriate promoter and terminator sequences for high level expression through efficient transcription and translation. Example: Use of promoters ('nos' from T-DNA) or expression cassettes (pRT plasmids) (Fig. 22.3d).

Shuttle Vector:

There are plasmids capable of propagating and transferring genes between two organisms (e.g., E. coli and A. tumefciciens). It has unique origins of replication for each cell type as well as different selectable markers. It can, therefore, be used to shuttle gene from prokaryotes to eukaryotes. Example: pBin19.

The prime objective of gene cloning is to finally result in the large scale production of proteins for a variety of purposes (industrial, commercial, human health and welfare). This is achieved by the maximal expression of cloned genes through manipulations.

The following are the important features of gene expression that can be considered for manipulation:

i. The presence of regulatable promoters.
ii. The number of copies of cloned genes.
iii. The location of the cloned genes whether inserted into a plasmid or integrated into host genome.
iv. The translation efficiency of the host.
v. The cellular location of the foreign protein and its stability in the host cell.

Some of the strategies that are employed for the manipulation of gene expression in E. coli are discussed hereunder.

Regulatable promoters:

The presence of a strong regulatable promoter sequence is essential for an effective expression of a cloned gene. This is achieved by employing the promoters of E. coli lac (lactose) operon or trp (tryptophan) operon. These promoters have strong affinity for RNA polymerase, and consequently the downstream region (of cloned gene) is transcribed. The promoters thus provide a switch for turning on or turning off the transcription of a cloned gene.

Fusion proteins:

The combination of a foreign protein (encoded by a cloned gene) with the host protein is referred to as a fusion protein. In general, the foreign proteins synthesized are rapidly degraded. This can be reduced by covalently linking a stable host protein to the foreign proteins (i.e., fusion proteins).

The fusion proteins in fact protect the proteolytic degradation of cloned gene product. The synthesis of fusion proteins is achieved by ligating the coding sequences of two genes (cloned gene and host gene). However, it is absolutely essential to ensure that cloned gene contains the correct sequence for the synthesis of the target protein.

Cleavage of fusion proteins:

The fusion proteins, as such, interfere with the biological activity of the target protein. Therefore, these proteins should be cleaved to release the specific desired functional proteins.

Uses of fusion proteins:

The purification of recombinant proteins is much easier in the form of fusion proteins. Fusion proteins are also useful for generating antibodies against target proteins.

Tandem gene arrays:

In general, increase in the number of plasmids (containing cloned gene) proportionately increases the production of recombinant protein. This has a drawback. As the plasmid number increases, the genes coding for antibiotic resistance also increase. The overall effect is that the regular metabolic activates of the host cell are disturbed for the synthesis of plasmid proteins. Consequently, the yield of cloned gene product is not optimum.

An alternative approach is to clone multiple copies of the target gene on a single plasmid (instead of a single gene on a plasmid). In this manner, tandem arrays of a gene can be created. However, each sequence of the genes should be in correct orientation for transcription and translation.

Efficiency of translation:

The quantity of the cloned gene product produced depends on the efficiency of translation. In general, the binding ability of mRNA with the ribosomal RNA, at translational initiation signal called ribosome binding site determines translation. Thus, the efficiency of translation is better if the binding of mRNA to rRNA is stronger. The actual binding between mRNA and rRNA occurs by complementary base pairing of a sequence of 6-8 nucleotides. To achieve maximum translation, the E. coli expression vectors are designed to possess a strong ribosome binding site.

Stability of proteins:

The half-lives of recombinant proteins are highly variable, ranging from minutes to hours. The stability of proteins can be increased by adding amino acids at the N-terminal end of the proteins. Thus, by attaching methionine, serine and alanine to the N-terminal end, the half-life of β-galactosidase can be increased from 2 minutes to 20 hours! Frequently, a single amino acid addition at N-terminal end stabilizes the protein. The yield of recombinant DNA proteins can be enhanced by increasing half-lives.

Secretion of proteins:

The stability of a protein and its secretion are interrelated. An amino acid sequence (signal peptide) may be attached to a protein to facilitate its secretion through cell membrane. Recombinant proteins secreted into the growth medium can be easily purified.

Integration of cloned DNA into the host chromosome:

The use of plasmids for transcription and translation of cloned DNA imposes a metabolic load on the host. In addition, there is often a chance of losing plasmids during cell multiplication. These problems can be overcome by integrating the cloned DNA directly into host chromosomal DNA. Once the cloned DNA becomes a part of genome, it can be maintained for several generations.

Cloned DNA integration into the host DNA is possible only when there is a complementary sequence of about 50 nucleotides between them. The exchange of DNAs occurs by a recombination process (Fig. 11.1). The cloned DNA lies in the middle of plasmid DNA. On physical contact with chromosomal DNA, base pairing occurs between plasmid DNA (x and y) and chromosomal DNA (x' and y'). And the cloned DNA is transferred to host chromosomal DNA by a physical exchange i.e., recombination.

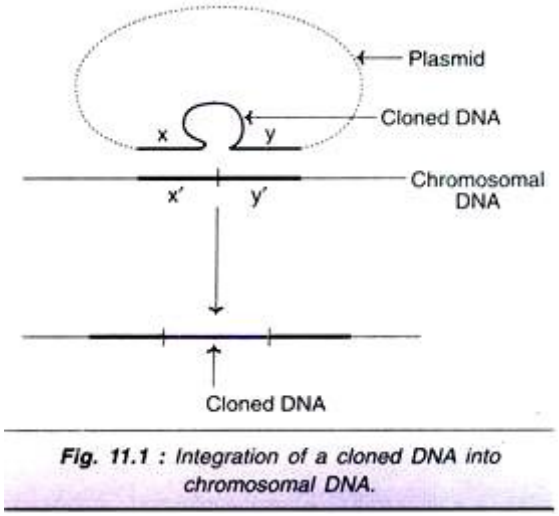


Fig. 11.1 : Integration of a cloned DNA into chromosomal DNA.

Metabolic load:

The presence of cloned DNA alters the metabolism and cellular functions of the host organism. Such metabolic changes are collectively referred to as metabolic load, metabolic drain or metabolic burden. There are several causes for the metabolic load. These include increased utilization of energy for replication and maintenance of plasmids, overproduction of proteins (also drains amino acids, tRNAs), and interference of foreign proteins on the host cell function.

Cassettes and gene fusions

A gene cassette is broadly a modular DNA sequence encoding one or more genes for a single biochemical function. In genetic engineering, a gene cassette refers to a manipulable fragment of DNA carrying, and capable of expressing, one or more genes of interest between one or more sets of restriction sites. It can be transferred from one DNA sequence to another by 'cutting' the fragment out using restriction enzymes and 'pasting' it back into the new context. Integrons are genetic structures in bacteria which express and are capable of acquiring and exchanging 'gene

cassettes'. These cassettes typically carry a single gene without a promoter. The entire series of cassettes is transcribed from an adjacent promoter. The gene cassettes are speculated to be inserted and excised via a circular intermediate. This would involve recombination between short sequences found at their termini and known as 59 base elements - which may not be 59 bases long. The 59-be are a diverse family of sequences that function as recognition sites for the site-specific integrase. Gene cassettes often carry antibiotic resistance genes. An example would be the kanMX cassette which confers kanamycin resistance upon bacteria or fungi. How these cassettes are initially created is not clear.

A gene made by joining parts of two different genes. Fusion genes, and the fusion proteins that come from them, may be made in the laboratory, or made naturally in the body when part of the DNA from one chromosome moves to another chromosome. Fusion proteins produced by this change may lead to the development of some types of cancer. For example, the BCR-ABL fusion gene and protein are found in some types of leukemia. Fusion genes and proteins may also be found in several other types of cancer, including soft tissue sarcoma, cancers of the prostate, breast, lung, bladder, colon, and rectum, and CNS tumors. Fusion genes and proteins are being studied in the diagnosis and treatment of cancer.

Challenges in producing recombinant protein in *E. coli*

1) No/Low expression

When the protein of interest cannot be detected through a sensitive technique (e.g., Westernblot) or it is detected but at very low levels (less than micrograms per liter of culture), the problem often lies in a harmful effect that the heterologous protein exerts on the cell.

Reasons	**Solutions**		
	Vector	**Host strain**	**Growth conditions**
Incorrect vector construction	**Confirm vector by sequencing**		
Rare codons	**Codon optimization**	**Use strains supplementing rare codons (Rosetta, Codon Plus)**	▪ **Lower induction temperature** ▪ **Grow in poor media**

Protein toxicity	▪ **Use promoters with tighter regulation** ▪ **Lower plasmid copy number**	▪ **Use pLysS/pLysE bearing strains in T7-based systems** ▪ **Use strains that are better for the expression of toxic proteins (C41 or C43)**	▪ **Start induction at high OD** ▪ **Shorten induction time** ▪ **Add glucose when using expression vectors containing lac-based promoters** ▪ **Use defined media with glucose as source of carbon**

2) Protein aggregation

The buildups of protein aggregates are known as inclusion bodies (IBs). IB formation results from an unbalanced equilibrium between protein aggregation and solubilization. So, it is possible to obtain a soluble recombinant protein by strategies that ameliorate the factors leading to IB formation**.**

Reasons	**Solutions**		
	Vector	**Host strain**	**Growth conditions**
Incorrect disulfide bond formation	▪ **Add fusion partners, including thioredoxin, DsbA, DsbC** ▪ **Clone in a vector containing secretion signal to cell periplasm**	**Use E. coli strains with oxidative cytoplasmic environment**	▪ **Lower inducer concentration** ▪ **Lower induction temperature**

Incorrect folding	▪ **Use a solubilizing partner** ▪ **Co-express with molecular chaperones**	**Use strains with cold-adapted chaperones**	▪ **Supplement media with chemical chaperones and cofactors** ▪ **Remove inducer and add fresh media** ▪ **Lower inducer concentration** ▪ **Lower temperature**
Proteins with high hydrophobicity or transmembrane domains	▪ **Add fusion tags, including GST, MBP, SUMO, etc.** ▪ **Generate truncated forms of protein**	**Use membrane rich strains (C41/C43)**	▪ **Lower induction temperature** ▪ **Shorten induction time** ▪ **Grow in poor medium** ▪ **Add heat shock chaperones**

3) Truncated protein

Sometimes a truncated form of protein is expressed rather than a complete wild protein. Reasons of the phenomenon and possible solutions are given below.

Reasons	**Solutions**		
	Vector	**Host strain**	**Growth conditions**
Rare codon	**Codon optimization**	**Use strains supplementing rare codons (Rosetta, Codon Plus)**	▪ **Lower induction temperature** ▪ **Grow in poor media**
Protein degradation	**Replace specific protease sites**	**Use low protease strains**	▪ **Induce at high OD** ▪ **Induce at low**

			temperature ▪ Shorten induction time ▪ Use protease inhibitors when breaking cells
Imbalanced translation process of fusion protein	▪ Change another fusion protein ▪ Move fusion protein to C-terminal		▪ Induce at low temperature ▪ Shorten induction time ▪ Change to poor media

4) Protein inactivity

Obtaining a nice amount of soluble protein is not the end of the road. The protein may still be of bad quality, i.e., it does not have the activity it should.

Reasons	Solutions		
	Vector	Host strain	Growth conditions
Low solubility of the protein	Fuse desired protein to a solubility enhancer (fusion partners)		Lower temperature
Lack of essential post translational modification	Change expression system		
Incomplete folding	▪ Use a solubilizing partner	Use strains with cold-adapted chaperones	▪ Monitor disulfide bond formation and allow further

	▪ Co-express with molecular chaperones		folding in vitro ▪ Lower temperature
Mutations in cDNA	**Sequence plasmid before and after induction**	**Use a recA− strain to ensure plasmid stability**	**Transform E. coli before each expression round**

Production of recombinant protein by eukaryotic cells

Expression of cloned genes in eukaryotes has certain advantages. The most important being the ability of eukaryotic organisms to bring about post-translational modifications—glycosylation, phosphorylation, correct disulfide bond formation, proteolytic cleavage etc. Eukaryotic expression systems produce stable and biologically active proteins. This is in contrast to the prokaryotic expression of cloned genes.

In general, the eukaryotic expression of cloned genes is quite comparable to that occurs in the prokaryotes. However, from the technical perspective, it is more difficult to conduct experiments with eukaryotic cells. Many a times, vectors with two distinct origins of replication are used. They serve as shuttle vectors and function in prokaryotic as well as eukaryotic hosts.

The insertion of a foreign DNA into bacterial and yeast cells is referred to as transformation The term transfection is used for the introduction of a foreign DNA into animal cells. The insert DNA in the eukaryotic cells may be associated with vector or integrated into the host chromosomal DNA.

Saccharomyces Cerevisiae— The Yeast in Expressing Cloned Genes:

The common yeast Saccharomyces cerevisiae is widely used as a host for the expression of cloned genes. There are many justifiable reasons for its extensive use.

i. S. cervisiae is single-celled that can be easily grown. Its biochemistry, genetics and physiology are quite known.

ii. It has a naturally occurring plasmid and strong promoters for efficient expression.

iii. S. cerevisiae can bring about many postradiational changes in proteins.

iv. The secreted recombinant proteins can be easily isolated, since very few host proteins are secreted.

v. The U.S. Food and Drug Administration has certified S. cerevisiae as a generally recognized as a safe (CRAS) organism.

As such, S. cerevisiae has been in use for several decades in baking and brewing industries. Biotechnologists work quite comfortably with this yeast to produce a large number of recombinant proteins. These include insulin, a_1 -antitrypsin, hepatitis B virus surface antigen, platelet derived growth factor, fibroblast growth factor and HIV-I antigens. These products are in use as diagnostic agents, vaccines, and therapeutic agents.

Vectors for S. cerevisiae:

There are three types of vectors for S. cerevisiae:

1. Episomal or plasmid vectors.
2. Integrating vectors.
3. Yeast artificial chromosomes (YACs).

1. Plasmid vectors:

Among the vectors, plasmids with single cloned genes are widely used. Manipulation with growth conditions increase the vector stability and expression efficiency. Use of tandem gene arrays has not met with success, since they are unstable.

2. Integrating vectors:

They are basically the integration of cloned genes with chromosomal DNA. These are not frequently used, since the protein production is low.

3. Yeast artificial chromosome (YAC):

Introduced in 1987, YAC is a fragment of yeast DNA that will accept a foreign DNA of about 250-500 kb in length. In fact, the yeast DNA is only about 1% of the total DNA which however, is very important, since it contains three essential genes required for replication. These are the genes for telomere (that protects DNA from nuclease degradation and thus maintains stability), centromere (forms spindles during cell division) and the origin of replication (where DNA polymerase initiates replication). YAC behaves just like a chromosome and replicates.

The construction of the yeast artificial chromosome is depicted in Fig. 11.2. Two opposite ends of a yeast chromosome namely the left telomere and right telomere are chosen. The left telomere is then attached to a centromere. A large segment of the foreign DNA is added and all the three are ligated. Unlike the plasmid vectors, the stability of YAC increases as the size of insert DNA increases.

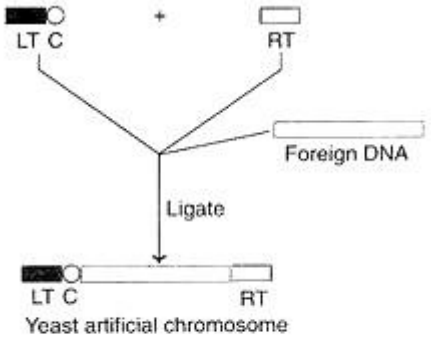


Fig. 11.2 : Construction of yeast artificial chromosome (LT-Left telomere; C-Centromere; RT-Right telomere).

YACs have not been used for commercial production of recombinant proteins. However, they have been employed successfully for physical mapping of genomic DNAs, particularly in human genome project.

Post-translational modifications by S. cerevisiae:

The heterologous proteins synthesized by S. cerevisiae undergo post-translational changes while they are being exported into the extracellular environment. To facilitate protein secretion, a

single (leader) peptide is attached to the protein. This peptide is removed by the yeast endoprotease.

Other Yeast Expression Systems:

Despite the very successful use of S. cerevisiae for generating recombinant proteins, there are certain limitations. These include a very low or a limited yield, difficulty in secretion of some proteins and hyper glycosylation. Attempts are being made to explore the utility of other yeasts for the production of hepatitis B virus surface antigen (HBsAg) and bovine lysozyme. The yeast, Hansenula polymorpha, is employed for the synthesis of α- and β-globin chains of human hemoglobin.

Insect Cell Expression Systems:

Cultured insect cells are in use for expressing cloned DNAs. Baculo viruses exclusively infect insect cells. The DNA of these viruses encode for several products and their productivity in cells is very high to the extent of more than 10,000 times compared to mammalian cells. Besides carrying a large number of foreign genes, the baculoviruses can effectively express and process the products formed. Another advantage with these viruses is that they cannot infect humans, other vertebrates or plants. Thus, baculoviruses are safe vectors.

Polyhedrin gene of baculovirus:

The polyhedrin gene is responsible for the synthesis of a matrix protein-polyhedrin. This protein is synthesized in large quantities by baculovirus during the infection cycle. Polyhedrin protects the virus from being inactivated by environmental agents. The promoter for polyhedrin gene is very strong. However, the life-cycle of baculovirus does not depend on the presence of this gene. Polyhedrin gene can be replaced by a cloned gene, and the genetically engineered baculovirus can infect the cultured insect cells.

The cloned gene expresses, and large quantities of recombinant proteins are produced. Because of a close similarity in the post-translational modifications between insects and mammals, biologically active proteins can be produced by this approach. And in fact, by using baculovirus as an expression vector system, a good number of mammalian and viral proteins have been synthesized (Table 11.1).

Table 11.1 Selected examples of recombinant proteins produced by baculovirus expression vector system

Adenosine deaminase
Alkaline phosphatase
Amyloid precursor protein
Anthrax antigen
DNA polymerase α
Erythropoietin
HIV-I envelope protein
Interferons (α, β)
Interleukin-2
Malaria proteins
Pancreatic lipase
Polio virus proteins
Rabies virus proteins
Rhodopsin
Simian rotavirus capsid antigen
Tissue plasminogen activator

Baculovirus expression vector system:

The most commonly used baculovirus is Autograph California multiple nuclear polyhedrosis virus (AcMNPV). It can grow on the insect cell lines (e.g., derived from fall army worm) and produce high levels of polyhedrin or a recombinant protein. The organization of a baculovirus (AcMNPV) transfer vector is shown in Fig. 11.3A. It consists of an E. coli-based plasmid vector along with the DNA of baculovirus. This in turn has AcMNPV DNA, a polyhedrin promoter region, cloning site for insert DNA and polyhedrin termination region.

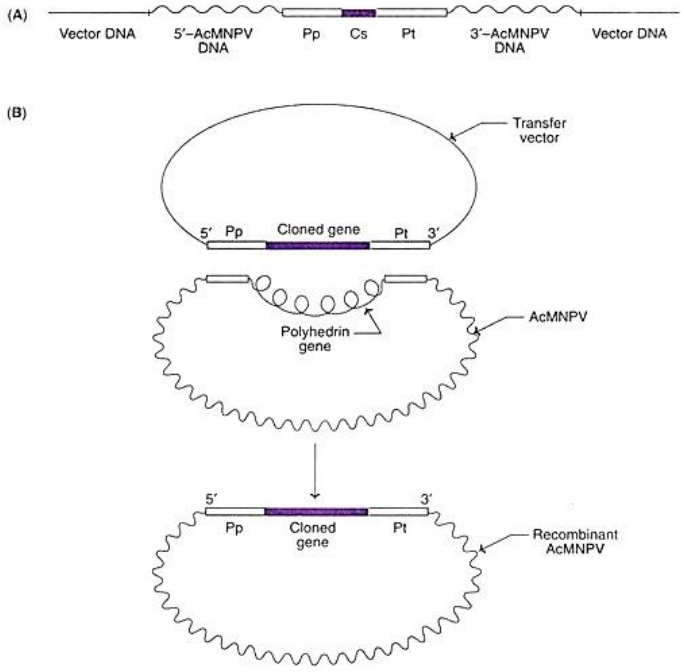


*Fig. 11.3 : Baculovirus expression vector system (A) Organization of baculovirus transfer vector (B) Replacement of the polyhedrin gene of baculovirus with a cloned gene from a transfer vector (AcMNPV–Autographa californica multiple nuclear polyhedrosis virus; Pp–Polyhydrin gene promoter; Cs–Cloning site; Pt–Polyhedrin gene termination; **Note** : The coding region of polyhedrin gene not shown in A).*

When the insect culture cells, transfected with AcMNPV are mixed with transfer vector carrying a cloned gene, a double crossover occurs. The result is that the cloned gene with polyhedrin promoter and termination sequences gets integrated into AcMNPV DNA (Fig. 11.3B). In this process, polyhedrin gene is lost. The recombinant baculovirus containing cloned gene is isolated. The host insect culture cells, on infection with recombinant baculovirus, produce heterologous proteins. A large number and a wide variety of recombinant proteins (around 500) have been synthesized in the laboratory. A majority of them (>95%) have the requisite post-translational modifications. A selected list of recombinant proteins is given in Table 11.1.

Modifications in the production of recombinant baculovirus:

The original method of creating recombinant baculovirus has undergone several changes. Incorporation of a unique Bsu 361 restriction endonuclease site on the polyhedrin gene increases the yield of recombinant baculovirus production to about 30% from the normal 1%.

Bacmid:

This is shuttle vector for E. coli and insect cell baculovirus. Construction of a recombinant bacmid is a novel approach to carry out all the genetic manipulations including the expression of baculovirus vector in E. coli.

Use of yeast cells:

The genetic manipulations of AcMNPV genome can be done in yeast cells with yeast-insect shuttle vector. Then the recombinant baculovirus is introduced into insect cells.

Mammalian Cell Expression Vectors:

Mammalian expression vectors are useful for the production of specific and authentic recombinant proteins (for use as therapeutic agents). In addition, they are also helpful for studying the function and regulation of mammalian genes. In general, the mammalian expression vectors are quite comparable to other eukaryotic expression vectors. . However, large-scale production of recombinant proteins with engineered mammalian cells is costly.

A diagrammatic representation of mammalian vector is shown in Fig. 11.4. It contains a eukaryotic origin of replication from an animal virus such as Simian virus 40 (SV40) and a prokaryotic origin of replication (from E. coli). The mammalian vector has a multiple cloning site and a selectable marker gene. Both of them are under the control of eukaryotic promoter and polyadenylation sequences.

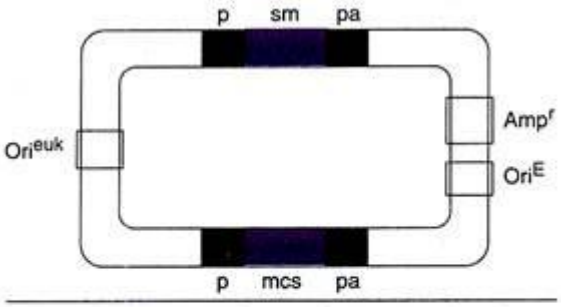


Fig. 11.4 : A diagrammatic representation of mammalian expression vector (p–Promoter sequence; Pa–polyadenylation sequence; mcs–Multiple cloning site; sm-Selectable marker gene; Ori^{euk}–Origin of eukaryotic replication; Ori^E–Origin of E. coli *replication; Amp^r–Ampicilin resistant marker gene).*

These sequences are obtained from either animal viruses (SV40, herpes simplex virus) or mammalian genes (growth hormone, metallothionein). The promoter sequences facilitate the transcription of cloned genes (at the multiple cloning site) and the selectable marker genes. On the other hand, the polyadenylation sequences terminate the transcription. Ampicillin resistant marker gene can be used for selecting the transformed E. coli cells.

Markers for mammalian expression vectors:

There are several markers in use for the selection of transformed mammalian cells. The bacterial gene (Neo^r) that encodes for neomycin phosphotransferase is frequently used. The other markers are the genes that encode for the enzyme dihydrofolate reductase (DHFR), and glutamine synthetase (GS).

Fusion tags and their role in purification of recombinant proteins

The fusion of a small protein or peptide (tag) to the protein of interest is a commonly used method to aid purification of recombinant proteins. Fusion tags can improve protein expression, stability, resistance to proteolytic degradation and solubility. A wide range of fusion tags are available from small peptides to relatively large proteins, each with its own unique characteristics. Many solubility tags are engineered for use in bacterial expression systems to overcome poor protein solubility.

Fusion Tag	Function	Size (kDa)	Description
Polyhistidine (e.g. 6xHis, 10xHis)	Affinity	1-2	The most commonly used affinity tag, binds to metal ions
Strep-tag II	Affinity	1	High affinity for engineered streptavidin
Thioredoxin (Trx)	Solubility	12	Aids in refolding proteins that require a reducing environment
Small Ubiquitin-like Modifier (SUMO)	Solubility	12	Contains a native cleavage sequence enabling tag removal with SUMO protease
Glutathione S-transferase (GST)	Solubility, affinity	26	High affinity for glutathione, often needs to be removed due to large size
Maltose Binding Protein (MBP)	Solubility, affinity	41	Binds to maltose, often needs to be removed due to large size

Fusion Tag Orientation

In any fusion tag system, the sequence encoding the tag is placed directly upstream or downstream of the recombinant gene sequence. A fusion tag that increases solubility or expression is most beneficial if placed at the N-terminus, acting as a positive translation initiator. Fusion tags primarily used for affinity purposes can be useful on the N- or C-terminus. If tag removal is desired, the sequence is often placed at the N-terminus to minimise the number of leftover residues following cleavage.

Combinatorial Fusion Tags

A combination of fusion tags can be used to maximise their functionality. A popular method is to utilise a solubility tag (e.g. GST or MBP) and an affinity tag (e.g. polyhistidine). This combination promotes soluble protein production and provides multiple options for affinity purification.

Fluorescent Proteins

Bioluminescent proteins are used in a wide range of biological applications from ultra-sensitive assay detection to *in vivo* imaging of cellular processes. The current range of bioluminescent proteins are variations of the wild-type Green Fluorescent Protein (GFP) derived from the jellyfish *Aequorea victoria*. Genetic variants featuring fluorescence emission spectral profiles spanning the blue, cyan, red, and yellow regions of the visible spectrum were developed by engineering specific mutations into GFP. These probes are used in a wide range of applications such as:

- Fluorescent Resonance Energy Transfer (FRET)
- Fluorescent Activated Cell Sorting (FACS)
- Photoactivated Localisation Microscopy (PALM)
- Fluorescence Recovery After Photobleaching (FRAP)
- Confocal Microscopy

Tag removal

In many cases it is desirable to remove fusion tags during purification to restore native protein structure. Removal of the tag is achieved by including a cleavage site between the fusion tag and the gene sequence. Commonly used cleavage proteases include:

Protease	Type	Recognition Site
Human Rhinovirus 3C (HRV3C)	Cysteine	LEVLFQ/GP
Tobacco Etch Virus Protease (TEVp)	Cysteine	ENLYFQ/G
Ubiquitin-like Specific Protease (SUMOp)	Cysteine	Recognises tertiary structure
Thrombin	Serine	LVPA/GS

A number of challenges can be encountered during tag removal, including:

- Incomplete cleavage
- Difficulty in separating the protease and tag from the native protein
- Loss of protein from the cleavage process
- Loss of solubility following cleavage

POSSIBLE QUESTIONS

8 MARKS

1. Give a detailed account on PCR techniques and its applications.
2. Write a detailed note on Real Time PCR
3. Explain how you will design primers for PCR in detail.
4. Discuss in detail about gene fusions.
5. Discuss the topic production of recombinant protein by eukaryotic cells in detail.
6. What are fusion tags? How it is employed in purification of recombinant proteins?

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
III BSc BIOCHEMISTRY
17BCU601A-GENETIC ENGINEERING and BIOTECHNOLOGY
MULTIPLE CHOICE QUESTIONS

Unit-IV

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	How many steps in the polymerase chain reaction (PCR)?	3-5 steps	3 steps	5 steps cyclic process	3 steps cyclic process	3 steps cyclic process
2	Which one is not related with the PCR reaction?	Denaturation	Generation	Renaturation	Synthesis	Generation
3	_____ is the template for the protein synthsis?	rRNA	tRNA	mRNA	SiRNA	mRNA
4	The order of three steps in PCR is	Denaturation, Renaturation and synthesis	Denaturation, Synthesis and renaturation	Renaturation, denaturation and synthesis	Synthesis, denaturation and renaturation	Denaturation, Renaturation and synthesis
5	The other name for annealing is	Renaturation	Denaturation	Polymerisation	Extension	Renaturation
6	A PCR reaction that continues for 30 cycles will produce approximately how many PCR products from a single template DNA molecule?	64	128,000	Approximately 1 million	Approximately 1 billion	Approximately 1 billion
7	Enzyme used in PCR is	DNApolymerase	RNA polymerase	Taq polymerase	Reverse transcriptase	Taq polymerase
8	Denaturation in PCR is achieved by	95◦C for 2 min	95◦C for 1 min	90◦C for 2 min	90◦C for 1 min	95◦C for 1 min
9	Annealing temperature in PCR is	95◦C	72◦C	55◦C	80◦C	55◦C
10	PCR was first proposed by	Kary Mullis	Hargobind khurrana	James Watson	Barbara mcclintock	Kary Mullis
11	Asymmetric PCR is to generate	Single stranded copies	Double stranded copies	Multiple copies	half copy	Single stranded copies
12	Enzyme used in PCR is tolerant to	Salt	Heat	Pressure	Ions	Heat
13	Reverse transcriptase PCR uses	mRNA as template to form	protein as template to form I	DNA as tempelate to form	Restriction enzymes	mRNA as template to form cDNA
14	First step of PCR is	Annealing	Renaturation	Denaturation	Ligation	Denaturation
15	Reverese transcription PCR is used for analysis in	Forensic analysis of DNA	Genomic DNA	mRNA	Binding	mRNA
16	PCR is a technique of	DNA degradation	DNA amplification	DNA sequencing	DNA ligation	DNA amplification
17	The basic requirements of PCR reaction excludes	DNA	Primers	Heat stable DNA polymera	Restriction enzymes	Restriction enzymes
18	Which of the following is used yeast as host	*Saccharomyces cerivisiae*	*E. Coli*	*Bacillus subtilis*	*Aspergillus niger*	*Saccharomyces cerivisiae*
19	Ti plasmids are isolated from	*Staphylococcus aureus*	*Agrobacterium tumefaciens*	*Sacchromyces cerivisiae*	*Escherichia coli*	*Agrobacterium tumefaciens*
20	Among the following which is not used for cloning purpose	Bacterio phage	Plasmid	Cosmid	*E.coli*	*E.coli*
21	Which organism produces Taq DNA polymerase	Thermus aquaticus	E.coli	Bacillus spp	Pseudomonas spp	Thermus aquaticus
22	Recoginition sequence for Eco R1	AATTCCTTAA	AG/CT	TC/GA	GA/TC	GA/TC
23	Restriction enzyme from Escherichia coli K 12 was first isol	Meselson and yaun	Lederberg and Meselson	yaun	Meselson	Meselson and yaun
24	The protoplast fusion is done by	PEG method	Electrofusion	Sodium nitrate method	agrobactrium method	PEG method
25	A carrier called a ---------- must be used to deliver the the	Fusion agent	Trancription initiater	Vector	Illucitor	Vector
26	*ex vivo, which means*	Cells are modified outside	genes are changed in cells s	recombination with a very	Recombination approac	Cells are modified outside the body and then transplanted back in again
27	*in vivo, which means*	Cells are modified outside	genes are changed in cells s	recombinationr with a very	Recombination approac	genes are changed in cells still in the body
28	An example of binary vector	pBR322	pUC8	pAL4404	pGV2260	pAL4404
29	DMSO (Dimethyl sulfoxide) is used as	Gelling agent	Alkaylating agent	Chelating agent	Cryoprotectant	Cryoprotectant
30	What is the product produced by reverse transcriptase	single strand DNA	double strand DNA	single strand RNA	double strand RNA	single strand DNA
31	Restriction enzymes cleave DNA at----------	Specific nucleotide sequen	Interior part of nucleotide se	Ends of nucleotide sequen	middle part of sequence	Specific nucleotide sequence
32	Endonucleases are enzymes that cleaves DNA at	Defined sequence	3' end of nucleotide	Internal position in randor	external position	Internal position in random manner
33	Type I restriction enzymes are isolated from the organisms	*Escherichia coli* B	*Virus*	Pig	*Escherichia coli D* am	*Escherichia coli* B
34	EcoRI is isolated from	Bacillus	Pseudomonas	*Escherichia coli*	vibrio	*Escherichia coli*
35	The polymerase chain reaction is a technique that selectiveh	RNA	DNA	Protein	None	DNA
36	The process of binding of primer to the denatured DNA stra	Denaturation	Annealing	Renaturation	None of these	Annealing
37	The technique used to identify specific DNA sequence in ba	Colony hybridization	In situ hybridization	Dot blot	Western blotting	Colony hybridization
38	Reverse transcriptase enzyme is	DNA dependent DNA pol	RNA dependent DNA poly	RNA dependent RNA po	DNA dependent RNA	RNA dependent DNA polymerase
39	Which are responsible for over production of recombinant p	Strong promoter	Gene dosage	week promoter	small promoter	Strong promoter
40	Promoters used in bacterial expression vector are	lac	BAX	TIC	LOCK	lac
41	Recombinant proteins expressed at high levels will sometim	Inclusion bodies	Exclusion bodies	partision bodies	microbodies	Inclusion bodies
42	The tac promoter is a fusion of	Trp and Hfr	Trp and lac	T7 and lac	Lac and ýpl	T7 and lac
43	Shuttle vectors contain	Origin of replication of ba	Origin of replication of euka	marker	operon	Origin of replication of bacteria and eukaryotes
44	Recombinant insulin was expressed in	Bacterial expression syster	Plant expression system	Insect expression system	Mammalian expression	Mammalian expression system
45	Synthetic vesicles composed of lipid bilayer used for transfe	Lipoproteins	Liposomes	Lysosomes	Lysozymes	Liposomes
46	The insoluble recombinant protein aggregates are known as	Inclusion bodies	Fusion proteins	Affinity proteins	Protein clot	Inclusion bodies
47	Vir E codes for which protein that binds to T-DNA during t	Competence factors	SSB proteins	DSB proteins	Tyrosine	SSB proteins
48	In retrovirus gag codes for	Reverse transcriptase	Polymerase	Synthase	Envelop proteins	Reverse transcriptase
49	In retrovirus pol codes for	Reverse transcriptase	Polymerase	Synthase	Envelop proteins	Polymerase
50	In retrovirus env codes for	Reverse transcriptase	Polymerase	Synthase	Envelope proteins	Envelope proteins
51	In second form the protein of matrix is called as ----------	Monohedrin	Polyhedrin	Dihedrin	Trihedrin	Polyhedrin
52	The RNAi pathway is initiated by the	Protein P	REC·	Dicer enzyme	None ofthe above	Dicer enzyme
53	Labeled antibody are used to detect	DNA	RNA	protein	Lipid	protein
54	Source of methionine for methylation obtained from	S- Adenosyl Methionine	Methionine	Protein synthesis	None	S- Adenosyl Methionine
55	Vir E codes for	Single strand binding proteins	Topoisimerase	Polymerase	Helicase	Single strand binding proteins
56	Micro carriers in ATC are prepared from	Collagen , Gelatin, Polystyrene	Collagen , Gelatin, palladium	Collagen , palladium, Polystyrene	Gelatin, Polystyrene, palladium	Collagen , Gelatin, Polystyrene
57	Which of the following gene cannot be used for making resistance against viral infection?	Gene for capsid protein	Gene for nucleocapsid protein	Satellite DNA	Gene for reversve transcriptase	Gene for reversve transcriptase
58	Which cell based plant technology involve the combining of two cells without cell wall from different species?	Clonal propagation	Cybradisation	Protoplast fusion	Mutant selection	Protoplast fusion
59	Which of the following would be useful selectable marker	Antibiotic	DNA	protein	RNA	DNA
60	The tac promoter is a fusion of	Trp and Hfr	Trp and lac	T7 and lac	Lac and ýpl	T7 and lac

UNIT –V
SYLLABUS

Applications of genetic engineering in Biotechnology: Site–directed mutagenesis and protein engineering. Applications in medicine, production of recombinant pharmaceuticals such as insulin, human growth hormone, factor VIII.

Recombinant vaccines. Gene therapy. Applications in agriculture – plant genetic engineering, herbicide resistant crops, problems with genetically modified plants, safety concerns.

Applications of genetic engineering in Biotechnology:
Site-directed mutagenesis and protein engineering

Any heritable change in the genome is commonly called a **mutation**. Biochemically, it is a chemical change or alteration in a nitrogen base of a DNA sequence resulting in the production of a defective protein or a truncated protein, which is not functional. These altered proteins can cause serious problems in metabolism leading to changes in the morphology and physiology of the organism. Mutations, in most cases, are spontaneous and may not be dangerous. Even though the natural mutations are spontaneous and rare, biologists can induce mutations using different methods, which in most cases are not desirable and precise. But, now molecular biologists can alter any amino acid of a protein by changing the corresponding bases in its gene very precisely and accurately resulting in desirable mutations. It is possible to alter properties such as increased stability, temperature resistance, product inhibition, substrate specificity, etc. of any enzyme.

The accurate induction of one or more point mutations on selected regions of a gene resulting in amino acid substitutions or deletion or addition is known as **site directed mutagenesis**. It can be defined as the controlled alteration of selected regions of a DNA molecule.

The **principle of site-directed mutagenesis** is that a mismatched oligonucleotide primer is extended, incorporating the 'mutation' into a strand of DNA that can be cloned. This technique of creating desired molecular mutations in a gene has contributed greatly to the basic understanding of functions of genes, DNA-protein interactions, gene regulations, the role of amino acids in the structure and functions of proteins, role of active centres in the enzyme-substrate interactions, etc. A single base change in a gene permits the evaluation of the role of specific amino acids in the function and structure of a protein. This technique also allows one to create or destroy a restriction site at specific locations within a DNA sequence or gene.

Site-directed mutagenesis is actually one of the applications of PCR. The gene, which has to be mutated, should be made into a single-stranded DNA by cloning into a M13 plasmid vector. By following modern PCR methods, it is possible to carry out the site directed mutation without the participation of M13 vector. The designing and chemical synthesis of the primer is the key factor in this technique. The part of the DNA where the mutation has to be introduced should be synthesized as an oligonucleotide primer, which is complementary to the respective region of the DNA except for the nucleotide that has to be changed. In short, the mutation is introduced to the gene in the form of a primer and the primer is extended with a polymerase reaction.

The site-directed mutagenesis is a multistep process that begins with the cloning of the gene in a bacteriophage like M13 to generate single stranded DNA. M13 is a filamentous bacteriophage

that specifically infects *e. coli* that expresses sex pili encoded by a plasmid F factor. M13 bacteriophage contains DNA in a single-stranded or replicative form, which is replicated to double-stranded DNA within a bacterial cell. The primer is designed and synthesized, which is an oligonucleotide complementary to the region of the DNA to be mutated except for the nucleotide to be changed. This oligonucleotide with the mismatched base or bases hybridize to the single-stranded DNA and serve as the primer to start synthesizing the complementary DNA strand with the help of a suitable DNA polymerase such as T4 DNA polymerase or Taq polymerase. The resulting doublestranded DNA will be a hybrid of the wild type parent strand and the mutated newly synthesized DNA strand. This DNA molecule can be transformed into an E.*coli* cell, where the mutated DNA strand serves as a template to replicate new strands carrying the mutation along with the wild strands. The bacteriophage plaques containing the mutated DNA can be screened by hybridizing with the labeled probe of the original mutated oligonucleotide. By adjusting the hybridization time and wash temperature of the hybridized probe, only the perfectly matched hybrid will remain and all other mismatched hybrids will dissociate. The presence of the desired mutation in the gene can be checked and confirmed after isolating the plasmid DNA from the single positive plaques and sequencing it.

A single base can be mutated in recombinant DNA plasmids with a process called **inverse PCR.** Two primers are synthesized with their antiparallel 5' ends complementary to the adjacent bases on the two strands of DNA. One of the two primers carries a specific mismatched base that is faithfully copied during the PCR amplification resulting in a recombinant plasmid with a single mutated base.

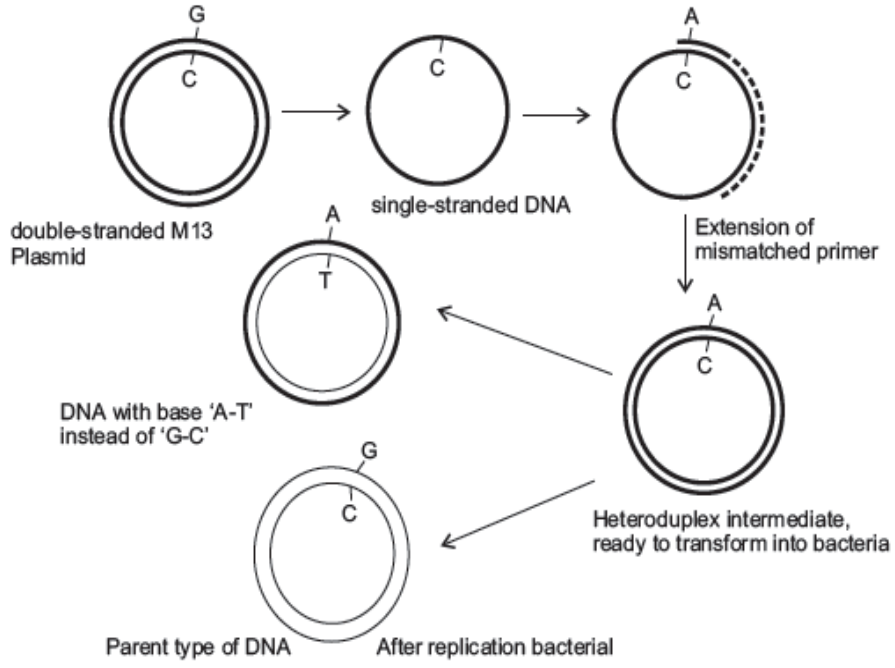


FIGURE: Site-directed mutagenesis of a gene mediated through M13 plasmid

Applications in medicine, production of recombinant pharmaceuticals such as insulin, human growth hormone, factor VIII.

The pharmaceutical products of recombinant DNA technology are broadly divided into the following three categories and briefly discussed with important examples:

1. Therapeutic agents for human diseases.
2. Human protein replacements.
3. Vaccines.

Some authors do not make such categorization and consider all of them together as pharmaceutically important products of biotechnology.

1. Therapeutic agents for human diseases.

Biotechnology is very useful for the production of several therapeutic products for treating human diseases. A selected list of rDNA-derived therapeutic agents along with trade names and their uses in humans are given in Table 15.3.

Some of these are described above (under human protein replacements) while the remaining are discussed below:

TABLE 15.3 A selected list of rDNA-derived therapeutic agents (approved by FDA) with trade names and their applications in humans

rDNA product	*Trade name(s)*	*Applications/uses*
Insulin	Humulin	Diabetes
Growth hormone	Protropin/Humatrope	Pituitary dwarfism
α-Interferon	Intron A	Hairy cell leukemia
Hepatitis B vaccine	Recombinax HB/Engerix B	Hepatitis B
Tissue plasminogen activator	Activase	Myocardial infarction
Factor VIII	Kogenate/Recombinate	Hemophilia
DNase	Pulmozyme	Cystic fibrosis
Erythropoietin	Epogen/Procrit	Severe anemia with kidney damage

Tissue Plasminogen Activator:

Tissue plasminogen activator (tPA) is a naturally occurring protease enzyme that helps to dissolve blood clots. tPA is a boon for patients suffering from thrombosis. The majority of natural deaths worldwide are due to a blockade of cerebral or coronary artery by a blood clot, technically called as thrombus. The phenomenon of thrombus blockage of blood vessels is referred to as thrombosis.

Chemically, thrombus consists of a network of fibrin, formed from the fibrinogen. In the normal circumstances, plasmin degrades fibrin and dissolves blood clots. This plasmin is actually produced by activation of plasminogen by tissue plasminogen activator (Fig. 15.3).

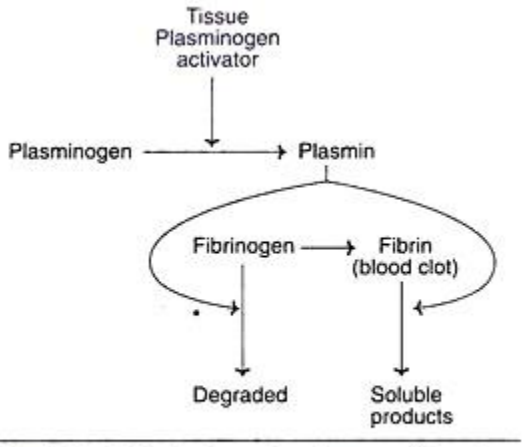


Fig. 15.3 : Role of tissue plasminogen activator in dissolving blood clots.

The natural biological systems is however, not that efficient to remove the blood clots through this machinery. Tissue plasminogen activator is very useful as a therapeutic agent in dissolving blood clots (thrombi) by activating plasminogen. By removing the arterial, thrombi, the possible damage caused by them on heart and brain could be reduced.

Production of recombinant tPA:

DNA technologists synthesized the complementary DNA (cDNA) molecule for tissue plasminogen activator. This cDNA was then attached to a synthetic plasmid and introduced into mammalian cells (Fig. 15.4). They were cultured and tPA-producing cells were selected by using methotrexate to the medium.

tPA-producing cells were transferred to an industrial tank (fermenter). tPA, secreted into the culture medium, is isolated for therapeutic purpose. It may be noted here that tPA was the first pharmaceutical product to be produced by mammalian cell culture.

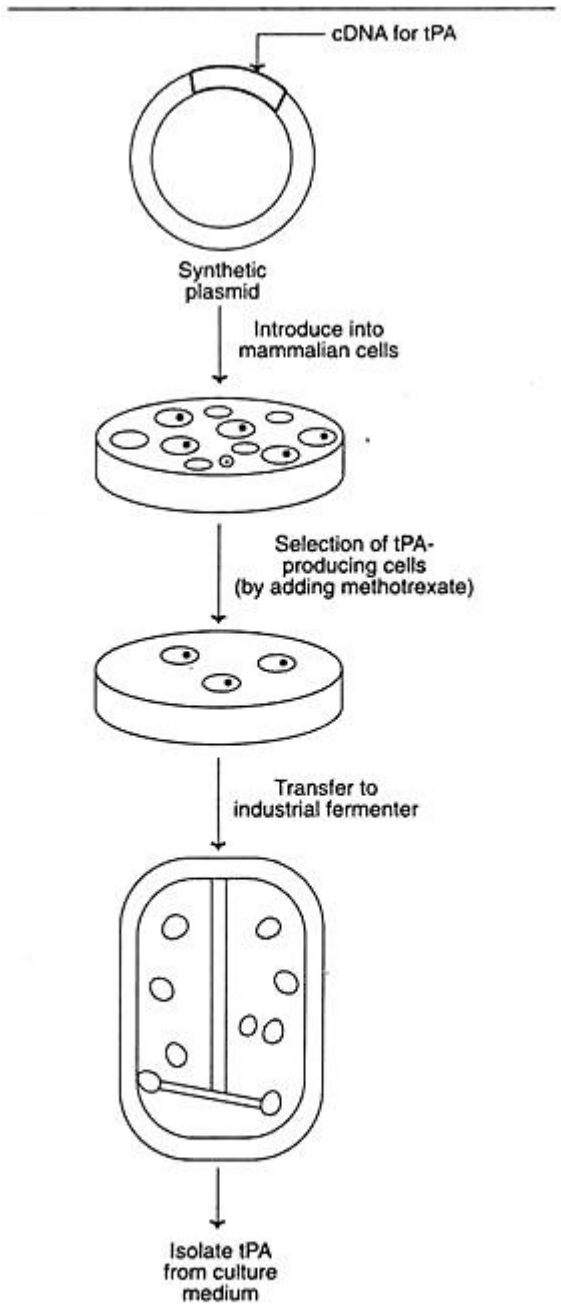


Fig. 15.4 : The production of recombinant tissue plasminogen activator (tPA).

Recombinant tPA has been in use since 1987 for treatment of patients with acute myocardial infarction or stroke. Gene-tech was the first to market tPA with a trade name Activase.

Alteplase and Reteplase:

These are the second generation recombinant tPAs. They have increased in vivo half-lives and are functionally more efficient. The general aspects of second generation recombinant proteins are given elsewhere.

Antibody-plasminogen activator conjugates:

An antibody against fibrin (anti-fibrin antibody) can be conjugated with tissue plasminogen activator. This conjugate is appropriately regarded as immunotherapeutic thrombolytic agent. It quickly and specifically binds to fibrin clots and locally increases the conversion of plasminogen to plasmin to dissolve fibrin (Fig. 15.5). In fact, anti-fibrin monoclonal antibodies have been synthesized, conjugated with tPA and tried for solubilizing blood clots.

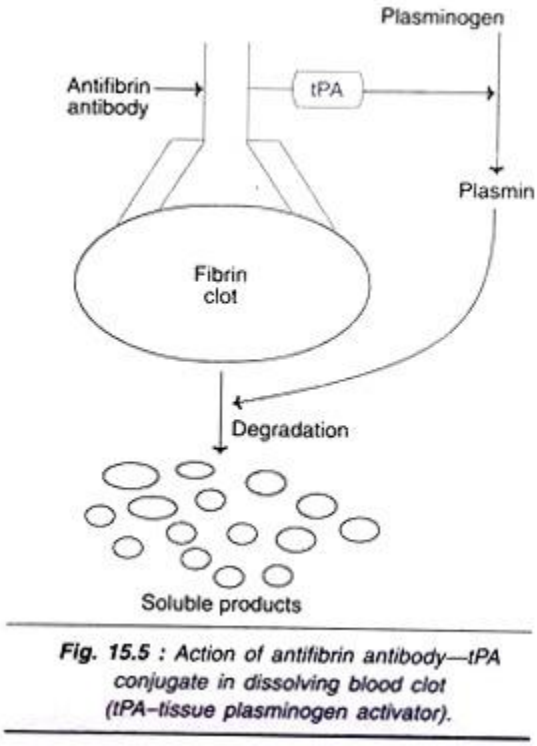


Fig. 15.5 : Action of antifibrin antibody—tPA conjugate in dissolving blood clot (tPA–tissue plasminogen activator).

Advantages of tPA as thrombolytic agent:

Tissue plasminogen activator acts on blood clots (solubilizes by degradation) without reducing the blood clotting capability elsewhere. This is in contrast to the action of urokinase or streptokinase which are more generalized in their action.

Further, tPA can be administered intravenously while urokinase and streptokinase have to be administered directly to the blocked blood vessel. Another merit of using tPA is that its action is much faster than other thrombolytic agents with much reduced side effects.

Interferons:

Interferon is an antiviral substance, and is the first line of defense against viral attacks. The term interferon has originated from the interference of this molecule on virus replication. It was originally discovered in 1957 by Alick Isaacs and Jean Lindemann and was considered to be a single substance.

It is now known that interferon actually consists of a group of more than twenty substances with molecular weights between 20,000-30,000 daltons. All the interferons are proteins in nature and many of them are glycoproteins. They are broadly categorized into three groups based on their structure and function

Interferon-α (IFN-α)
Interferon-β (IFN-β)
Interferon-ɣ (IFN-ɣ)

Mechanism of action of interferons:

Interferons are produced by mammalian cells when infected by viruses. As the virus releases its nucleic acid into cellular cytoplasm, it stimulates the host DNA to produce interferons. These interferons, secreted by the cells, bind to the adjacent cells. Here, they stimulate the cellular DNA to produce a series of antiviral enzymes.

The so formed proteins inhibit viral replication and protect the cells (Fig. 15.6). It is believed that the protective (enzymes) proteins bind to mRNA of viruses and block their protein synthesis.

The action of interferons appears to be species specific. Thus, human interferons operate in humans. Other animal (dog, mouse) interferons are ineffective in man.

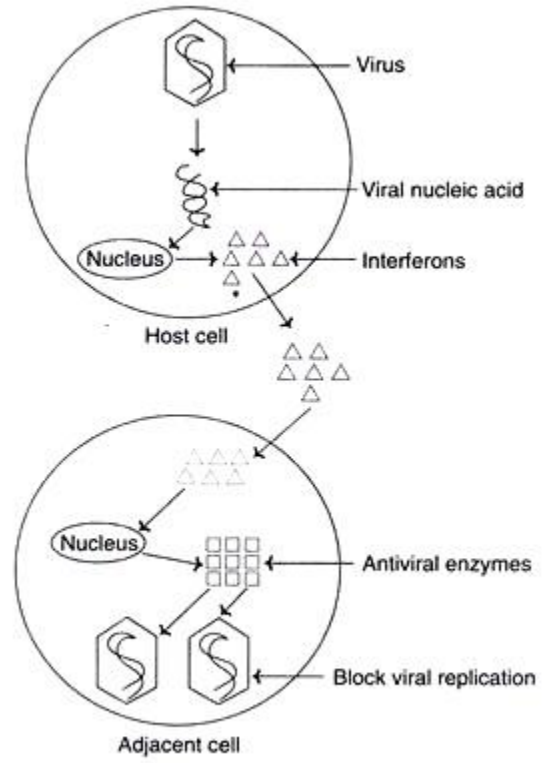


Fig. 15.6 : The mechanism of action of interferons.

Isolation of interferons in the early years:

Blood was the only source of interferons earlier. The procedure was very tedious and the quantity of interferons isolated was very little. Thus, as much as 50,000 litres of human blood

was required to get just 100 mg of interferons. Therefore, it was very difficult to conduct research or use interferons for therapeutic purposes.

Production of recombinant interferons:

The complementary DNA (cDNA) was synthesized from the mRNA of a specific interferon. This is inserted to a vector (say plasmid) which is introduced into E. coli or other cells. The interferon can be isolated from the culture medium. This is the basic mechanism of producing recombinant interferons.

The production of interferons is relatively less in bacterial hosts, although E. coli was the first to be used. This is mainly because most interferons are glycoproteins in nature and bacteria do not possess the machinery for glycosylation of proteins.

Production interferons by yeasts:

The yeast Saccharomyces cerevisiae is more suitable for the production of recombinant interferons. This is mainly because the yeast possess the mechanism to carry out glycosylation of proteins, similar to that occurs in mammalian cells. The DNA sequence coding for specific human interferon can be attached to the yeast alcohol dehydrogenase gene in a plasmid and introduced into 4 yeast cells. The yield of interferons is several fold higher compared to E. coli.

Production of hybrid interferons:

Several attempts have been made to produce hybrid interferons. This is advantageous since different interferons with different antiviral activities can be combined to produce a more efficient interferon. Further, the glycosylation step can be bypassed, and bacteria can be used to produce hybrid interferons. The hybrid interferons are more reactive in performing their function.

The creation of hybrid genes from the genes of IFN-α_2 and IFN-α_3 is illustrated in Fig. 15.7. These genes are digested by restriction endonucleases. The resulting fragments are ligated to generate hybrid genes. The appropriate hybrid genes can be selected and used for producing hybrid interferons. E. coli can be employed for this purpose.

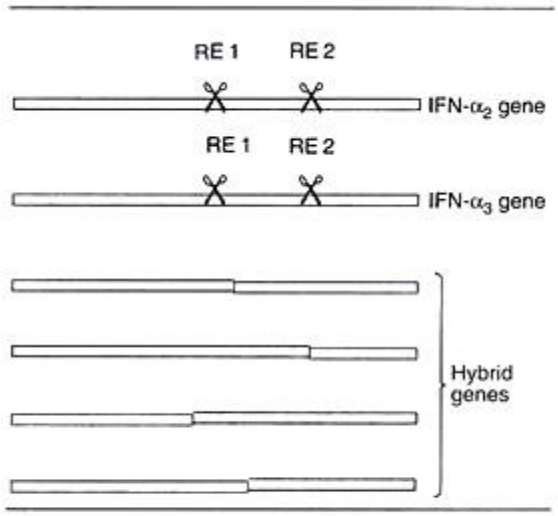


Fig. 15.7 : Creation of hybrid interferon genes (RE-Restriction endonuclease, IFN-Interferon).

Therapeutic applications of interferons:

Interferons-α, -β and -ɣ were respectively approved for therapeutic use in humans in the years 1986, 1993 and 1990. A Swiss biotechnology firm was the first to market interferon-α with a trade name Intron. Interferons are used for the treatment of a large number of viral diseases and cancers.

The cancers include leukemia, kaposis sarcoma, bladder cancer, head and neck cancer, renal cell carcinoma, skin cancer and multiple myeloma. The other diseases employing interferon therapy are AIDS, multiple sclerosis, genital warts, hepatitis C, herpes zoster etc.

Interferons are also employed in the treatment of common colds and influenza. For this purpose, interferons can be used as nasal sprays. The basic mechanism of action of interferons against viruses has already been described.

Interferons are found to cause the death of cancerous cells. This is brought out by stimulating the action of natural killer (NK) cells, a specialised form of lymphocytes that can destroy cancerous cells. Despite the widespread therapeutic applications of interferons, they are not within the reach of a large number of common people due to the cost factor (the cost of production being very high).

Erythropoietin:

Erythropoietin is a hormone synthesized by the kidneys. It stimulates the stem cells of bone marrow to produce mature erythrocytes. Biotechnologists were successful in producing recombinant erythropoietin. An approval for its therapeutic use in humans was obtained in the year 1989. Amgen Inc. first marked erythropoietin with a trade name Epogen. It is useful in treating the patients with severe anemia that accompanies kidney disease.

Another firm Ortho-Biotech company produced Procrit, a genetically engineered erythropoietin in 1997. Procrit acts like the natural hormone and stimulates the production of erythrocytes. It is used in anemic patients undergoing non-cardiac, nonvascular surgery. Procrit administration

before surgery serves as an alternative to blood transfusion. However, therapeutic use of procrit is quite expensive, hence not widely used.

Deoxyribonuclease I (DNase I):

The enzyme DNase I hydrolysis long DNA chains into shorter oligonucleotides. The biotechnology firm Genentech isolated and expressed the gene to produce recombinant DNase I. This enzyme is very useful in the treatment of common hereditary disease cystic fibrosis, as explained hereunder.

Cystic fibrosis (CF) is one of the most common (frequency 1: 25,000) genetic diseases. Patients of CF are highly susceptible to lung infections by bacteria. The presence of live or dead bacteria leads to the accumulation of thick mucus in the lungs making the breathing very difficult.

The major constituent of this mucus is the bacterial DNA (released on bacterial lysis).

Administration of the enzyme DNase I to the lungs of CF patients decreases the viscosity of the mucus, and the breathing is made easier. It must be remembered that DNase I cannot cure cystic fibrosis. It can only relieve the severe symptoms of the disease in most patients.

Alginate Lyase:

Alginate lyase acts on a polysaccharide polymer namely alginate. Alginate is found in soil and marine bacteria. The occurrence of mucus in the lungs of cystic fibrosis patients is partly due to alginate, produced by the bacterium Pseudomonas aeruginosa. Therefore, administration of alginate lyase instead of or in addition to DNase I helps to clear lungs of CF patients.

Alginate lyase gene has been isolated from a Gram-negative soil bacterium, Flavobacterium sp. This gene was used to produce recombinant alginate lyase in E. coli. Trails are being conducted for therapeutic use of this enzyme in CF patients.

Second Generation Therapeutic Proteins (Muteins):

By employing site-directed mutagenesis, the amino acid sequence of a recombinant protein can be suitably modified as desired, by a technique referred to as protein engineering. The mutated proteins are collectively referred to as muteins.

Protein engineering is a rational approach to modify a protein with regard to its stability, solubility, specificity, substrate affinity, pharmacokinetics etc. The muteins obtained by protein engineering technique are considered as Second generation of therapeutic proteins. Selected examples of such proteins (e.g. insulin lispro, Alteplase) are already described.

2. Human Protein Replacements:

The synthesis of the cellular proteins is ultimately under the control of genes. Any defect in a gene produces an incorrect protein or no protein at all. Sometimes, the occurrence of a defective (i.e. functionally ineffective) or deficient protein may cause a disease. Thus, gene defects will result in inherited or genetically linked diseases.

Identification of defective or deficient proteins in the causation of inherited diseases is very important. The recombinant DNA technology can be fruitfully employed to produce human proteins that can be used for the treatment of genetically linked diseases. This is referred to as human protein replacement strategy in biotechnology.

Insulin:

The hormone insulin is produced by the (J-cells of islets of Langerhans of pancreas. Human insulin contains 51 amino acids, arranged in two polypeptide chains. The chain A has 21 amino acids while B has 30 amino acids. Both are held together by disulfide bonds.

Diabetes mellitus:

Diabetes mellitus affects about 2-3% of the general population. It is a genetically linked disease characterized by increased blood glucose concentration (hyperglycemia). The occurrence of diabetes is due to insufficient or inefficient (incompetent) insulin. Insulin facilitates the cellular uptake and utilization of glucose for the release of energy.

In the absence of insulin, glucose accumulates in the blood stream at higher concentration, usually when the blood glucose concentration exceeds about 180 mg/dl, glucose is excreted into urine. The patients of diabetes are weak and tired since the production of energy (i.e. ATP) is very much depressed.

The more serious complications of uncontrolled diabetes include kidney damage (nephropathy), eye damage (retinopathy), nerve diseases (neuropathy) and circulatory diseases (atherosclerosis, stroke). In fact, diabetes is the third leading cause of death (after heart disease and cancer) in many developed countries.

In the early years, insulin isolated and purified from the pancreases of pigs and cows was used for the treatment of diabetics. There is a slight difference (by one to three amino acids) in the structure of animal insulin compared to human insulin. This resulted in allergy in some of the diabetics when animal insulin was administered.

Another problem with animal insulin is that large number of animals have to be sacrificed for extracting insulin from their pancreases. For instance, about 70 pigs (giving about 5 kg pancreatic tissue) have to be killed to get insulin for treating a single diabetic patient just for one year!

Production of recombinant insulin:

Attempts to produce insulin by recombinant DNA technology started in late 1970s. The basic technique consisted of inserting human insulin gene and the promoter gene of lac operon on to the plasmids of E. coli. By this method human insulin was produced. It was in July 1980, seventeen human volunteers were, for the first time, administered recombinant insulin for treatment of diabetes at Guy's Hospital, London.

And in fact, insulin was the first ever pharmaceutical product of recombinant DNA technology administered to humans. Recombinant insulin worked well, and this gave hope to scientists that DNA technology could be successfully employed to produce substances of medical and commercial importance. An approval, by the concerned authorities, for using recombinant insulin for the treatment of diabetes mellitus was given in 1982. And in 1986, Eli Lilly Company received approval to market human insulin under the trade name Humulin.

Technique for recombinant insulin production:

The original technique of insulin synthesis in E. coli has undergone several changes, for improving the yield, e.g. addition of signal peptide, synthesis of A and B chains separately etc. The procedure employed for the synthesis of two insulin chains A and B is illustrated in Fig. 15.1. The genes for insulin A chain and B chain are separately inserted to the plasmids of two different E. coli cultures.

The lac operon system (consisting of inducer gene, promoter gene, operator gene and structural gene Z for β-galactosidase) is used for expression of both the genes. The presence of lactose in the culture medium induces the synthesis of insulin A and B chains in separate cultures. The so

formed insulin chains can be isolated, purified and joined together to give a full-pledged human insulin.

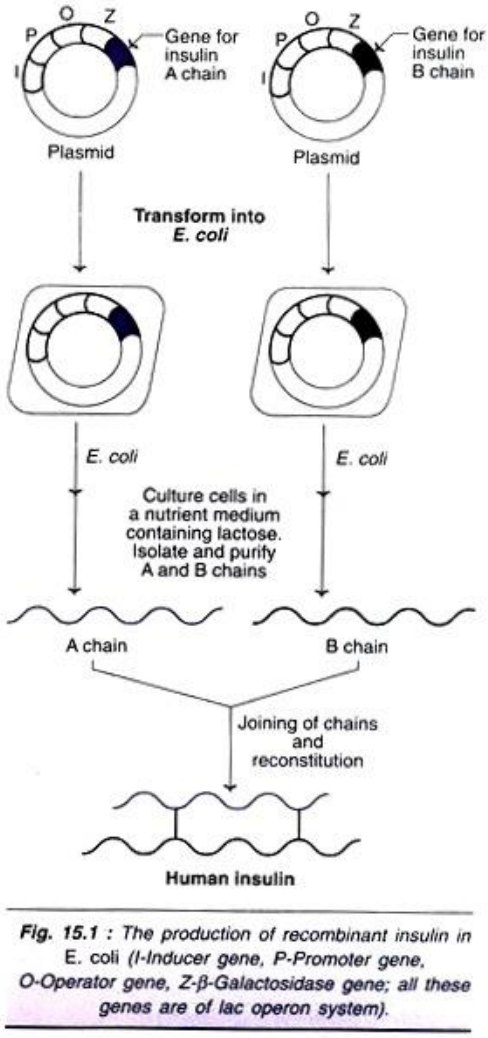


Fig. 15.1 : The production of recombinant insulin in E. coli (I-Inducer gene, P-Promoter gene, O-Operator gene, Z-β-Galactosidase gene; all these genes are of lac operon system).

Second generation recombinant insulin's:

After injecting the insulin, the plasma concentration of insulin rises slowly. And for this reason, insulin injection has to be done at last 15 minutes before a meal. Further, decrease in the insulin level is also slow, exposing the patients to a danger of hyper-insulinemia. All this is due to the existence of therapeutic insulin as a hexamer (six molecules associated), which dissociates slowly to the biologically active dimer or monomer.

Attempts have been made in recent years to produce second generation insulin by site- directed mutagenesis and protein engineering. The second generation recombinant proteins are termed as muteins. A large number of insulin muteins have been constructed with an objective of faster dissociation of hexamers to biologically active forms. Among these is insulin lispro, with modified amino acid residues of the B-chain of insulin. Insulin lispro can be injected immediately before a meal as it attains the pharmacologically efficient levels very fast.

Chemically altered porcin insulin:

As already stated, porcin (pig) insulin differs from human insulin just by one amino acid-alanine in place of threonine at the C-terminal and of B-chain of human insulin. Biotechnologists have developed methods to alter the chemical structure of porcin insulin to make it identical to human insulin. And this chemically modified porcin insulin can also be employed for the treatment diabetes mellitus.

Human Growth Hormone:

Growth hormone is produced by the pituitary gland. It regulates the growth and development. Growth hormone stimulates overall body growth by increasing the cellular uptake of amino acids, and protein synthesis, and promoting the use of fat as body fuel.

Insufficient human growth hormone (hGH) in young children results in retarded growth, clinically referred to as pituitary dwarfism. The child usually is less than four feet in height, and has chubby face and abundant fat around the waist.

Traditional treatment for dwarfism:

The children of pituitary dwarfism were treated with regular injections of growth hormone extracted from the brains of deceased humans. It may be noted that only human growth hormone is effective for treatment of dwarfism. (This is in contrast to diabetes where animal insulin's are employed).

At least eight pituitary glands from cadavers must be extracted to get hGH adequate for treating a dwarf child just for one year! And such treatment has to be continued for 8-10 years!! Further, administration hGH isolated from human brains exposes the children to a great risk of transmitting the cadaver brain diseases (through virus or viral-like agents) e.g. Creuzfeldt- Jacob (CJ) syndrome characterized by convulsions, wasting of muscle etc.

Production of recombinant hGH:

Biotechnologists can now produce hGH by genetic engineering. The technique adopted is quite comparable with that of insulin production. The procedure essentially consists of inserting hGH gene into E. coli plasmid, culturing the cells and isolation of the hGH from the extracellular medium.

Limitation in hGH production:

The hGH is a protein comprised of 191 amino acids. During the course of its natural synthesis in the body, hGH is tagged with a single peptide (with 26 amino acids). The signal peptide is removed during secretion to release the active hGH for biological functions. The entire process of hGH synthesis goes on in an orderly fashion in the body.

However, signal peptide interrupts hGH production by recombinant technology. The complementary DNA (cDNA) synthesized from the mRNA encoding hGH is inserted into the plasmid. The plasmid containing E. coli when cultured, produces full length hGH along with signal peptide. But E. coli cannot remove the signal peptide.

Further, it is also quite difficult to get rid of signal peptide by various other means. Theoretically, cDNA encoding signal peptide can be cut to solve these problems. Unfortunately, there is no restriction endonuclease to do this job, hence this is not possible.

A novel approach for hGH production:

Biotechnologists have resolved the problem of signal peptide interruption by a novel approach (Fig. 15.2). The base sequence in cDNA encoding signal peptide (26 amino acids) plus the

neighbouring 24 amino acids (i.e a. total 50 amino acids) is cut by restriction endonuclease ECoRI.

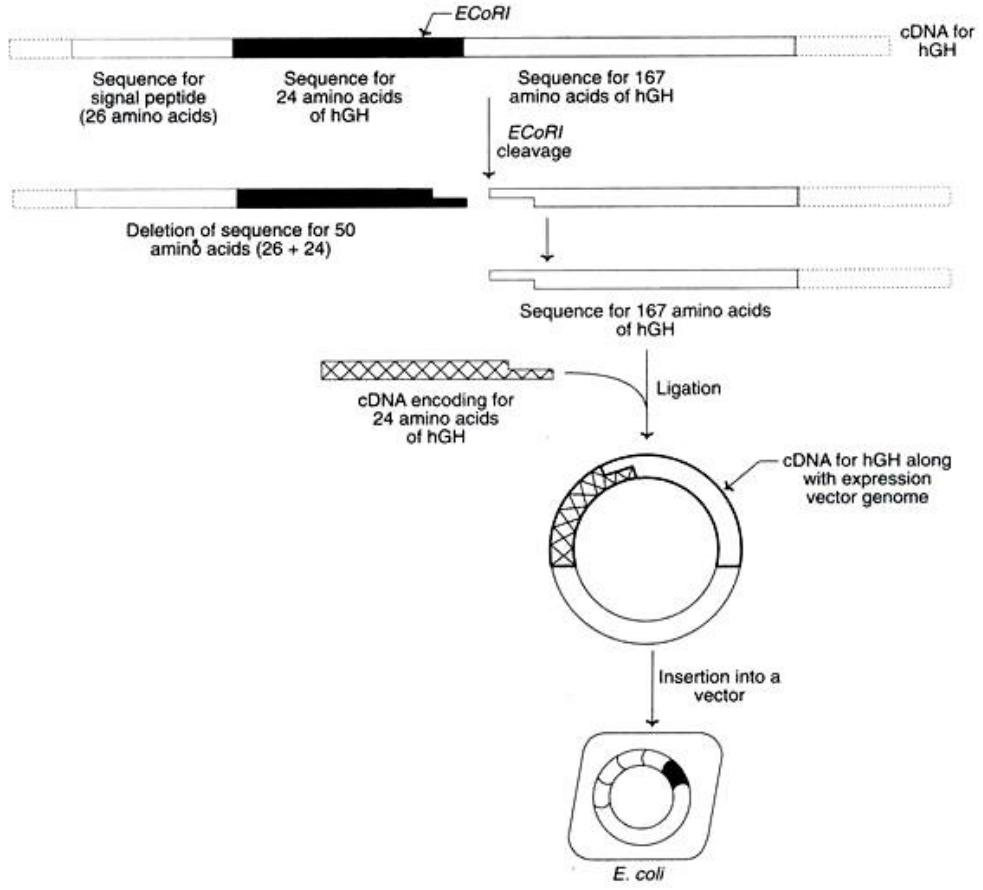


Fig. 15.2 : The production of recombinant human growth hormone (cDNA–Complementary DNA, hGH–Human growth hormone)

Now a gene (cDNA) for 24 amino acid sequence of hGH (that has been deleted) is freshly synthesized and ligated to the remaining hGH cDNA. The so constituted cDNA, attached to a vector, is inserted into a bacterium such as E. coli for culture and production of hGH. In this manner, the biologically functional hGH can be produced by DNA technology. Recombinant hGH was approved for human use in 1985. It is marketed as Protropin by Gene-tech Company and Humatrope by Eri Lilly Company.

Controversy over the use of hGH:

Recombinant hGH can be administered to children of very short stature. It has to be given daily for many years with an annual cost of about $ 20,000. Some workers have reported substantial increase in the height of growth retarded children.

One group of workers observed that the normal growth pattern in children was not restored on hGH administration, although there was an initial spurt. Another big question raised by the

opponents of hGH therapy is whether it is necessary to consider short stature as a disorder at all for treatment!

Use of recombinant growth hormone for farm animals:

Recombinant GH is now available for administration to farm animals to promote early growth and development. Such farm animals yield linear meat, besides increased milk production. However, use of GH in farm animals is a controversial issue.

Clotting Factor VIII:

The clotting factor VIII is required for proper blood clotting process. A genetic defect in the synthesis of factor VIII results in the disorder hemophilia A. This is a sex-linked disease (incidence 1 in 10,000 males) transmitted by females affecting males. The victims have prolonged clotting time and suffer from internal bleeding.

Traditional treatment for hemophilia A:

Clotting factor VIII was isolated from the whole blood and administered to the patients of hemophilia A. This approach requires large quantities of blood. Another problem is the risk of transmission of certain diseases like AIDS to the hemophiliacs.

Production of recombinant factor VIII:

The gene for the formation of factor VIII is located on X chromosome. It is a complex gene of 186 kb (i.e., 186,000 base pairs) in size, organized into 26 exons of varying length. In between the exons, many introns are present. The introns vary in their size, starting from 200 base pairs to as high as 32,000 base pairs.

Biotechnologists were able to isolate mature mRNA (containing only exons and no introns) that is responsible for the synthesis of factor VIII. This mRNA contains 9,000 bases and synthesizes the protein, factor VIII. Factor VIII contains 2332 amino acids, with carbohydrate molecules attached at least at 25 sites.

DNA technologists synthesized the complementary DNA (cDNA) for mature mRNA of factor VIII. This cDNA can be inserted into mammalian cells or hamster kidney cells for the production of recombinant factor VIII. Since 1992, factor VIII is available in the market. It is produced by Genetics Institute in Cambridge, Massachusetts and sold as Recombinant while Miles Laboratories sell under the trade name Kogenate.

Recombinant Vaccines

The three types are: (1) Subunit Vaccines (2) Attenuated Recombinant Vaccines and (3) Vector Recombinant Vaccines.

Recombinant Vaccines—General:

Recombinant DNA technology in recent years has become a boon to produce new generation vaccines. By this approach, some of the limitations (listed above) of traditional vaccine production could be overcome. In addition, several new strategies, involving gene manipulation are being tried to create novel recombinant vaccines.

Types of Recombinant Vaccines:

The recombinant vaccines may be broadly categorized into three groups:

1. Subunit recombinant vaccines:

These are the components of the pathogenic organisms. Subunit vaccines include proteins, peptides and DNA.

2. Attenuated recombinant vaccines:

These are the genetically modified pathogenic organisms (bacteria or viruses) that are made non-pathogenic and used as vaccines.

3. Vector recombinant vaccines:

These are the genetically modified viral vectors that can be used as vaccines against certain pathogens. Some of the developments made in the production of recombinant vaccines against certain diseases are briefly described.

Type # 1. Subunit Vaccines:

As already stated, subunit recombinant vaccines are the components (proteins, peptides, DNAs) of the pathogenic organisms. The advantages of these vaccines include their purity in preparation, stability and safe use. The disadvantages are — high cost factor and possible alteration in native conformation. Scientists carefully evaluate the pros and cons of subunit vaccines for each disease, and proceed on the considered merits.

Hepatitis B:

Hepatitis B is a widespread disease in man. It primarily affects liver causing chronic hepatitis, cirrhosis and liver cancer. Hepatitis B virus is a 42 nm particle, called Dane particle. It consists of a core containing a viral genome (DNA) surrounded by a phospholipid envelope carrying surface antigens (Fig. 16.1 A).

Infection with hepatitis B virus produced Dane particles and 22 nm sized particles. The latter contain surface antigens which are more immunogenic. It is however, very difficult to grow hepatitis B virus in mammalian cell culture and produce surface antigens.

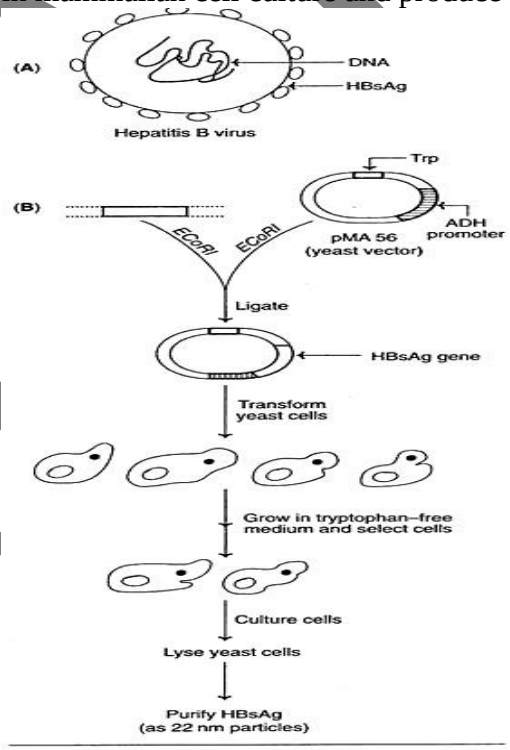


Fig. 16.1 : (A) Hepatitis B virus–Dane particle (42 nm particle); (B) Production of hepatitis B surface antigen (HBsAg) in yeast cells (Trp–Tryptophan, ADH–Alcohol dehydrogenase).

The gene encoding for hepatitis B surface antigen (HBsAg) has been identified. Recombinant hepatitis B vaccine as a subunit vaccine, is produced by cloning HbsAg gene in yeast cells. Saccharomyces cerevisiae, a harmless baking and brewing yeast, is used for this purpose (Fig. 16.1B). The gene for HBsAg is inserted (pMA 56) which is linked to the alcohol dehydrogenase promoter. These plasmids are then transferred and cultured.

The cells grown in tryptophan, free medium are selected and cloned. The yeast cells are cultured. The HBsAg gene is expressed to produce 2nm sized particles similar to those found in patients infected with hepatitis B. (These particles are immunoreactive with anti-HBsAg antibodies). The subunit HBsAg as 22 nm particles can be isolated and used to immunize individuals against hepatitis B.

Hepatitis B vaccine-the first synthetic vaccine:

In 1987, the recombinant vaccine for hepatitis B (i.e. HBsAg) became the first synthetic vaccine for public use. It was marketed by trade names Recombivax and Engerix-B. Hepatitis B vaccine is safe to use, very effective and produces no allergic reactions. For these reasons, this recombinant vaccine has been in use since 1987.

The individuals must be administered three doses over a period of six months. Immunization against hepatitis B is strongly recommended to anyone coming in contact with blood or body secretions. All the health professionals—physicians, surgeons, medical laboratory technicians, nurses, dentists, besides police officers, firefighters etc., must get vaccinated against hepatitis B.

Hepatitis B vaccine in India:

India is the fourth country (after USA, France and Belgium) in the world to develop an indigenous hepatitis B vaccine. It was launched in 1997, and is now being used.

Hepatitis B vaccine tomato?

Biotechnologists have been successful in inserting hepatitis B gene into the cells of the tomato plant. These genetically engineered plants produce hepatitis B antigens. The day may not be far off to get immunized against hepatitis B by having a tomato with lunch!

Foot and Mouth Disease:

Foot and mouth disease (FMD) is a highly contagious disease affecting cattle and pigs. A formalin killed foot and mouth disease virus (FMDV) was previously used to vaccinate against this disease. The genome of FMDV is composed of a single— stranded RNA, covered by four viral proteins (VP1, VP2, VP3 and VP4). Among these, VPI is immunogenic. The nucleotide sequence encoding VPI was identified in the FMDV genome. A double- stranded complementary DNA (cDNA) from the single-stranded viral RNA (genome) was synthesized. This cDNA was then digested with restriction enzymes and the fragments were cloned by using plasmid pBR322 in E. coli. The recombinant vaccine for FMDV in the form of viral protein 1 was used to vaccinate animals. However, VPI vaccination was found to be less effective than that of the whole virus in protecting FMD. Further, studies are being pursued to improve the efficiency of subunit vaccine.

The concept of peptide vaccines:

Theoretically, it is expected that only small portions of a given protein (i.e., domains) are immunogenic and bind to antibodies. Logically, it is possible to use short peptides that are immunogenic as vaccines. These are referred to as peptide vaccines.

Peptide vaccines for foot and mouth disease:

Some details on the FMD are described above. The domains of viral protein I (VPI) of FMDV were chemically synthesized. From the C-terminal end of VPI, amino acids 141 to 160, 151 to 160 and 200 to 213, and from N-terminal end, amino acids 9 to 24, 17 to 32 and 25 to 41 were synthesized.

Each one of these short peptides (domains) was bound to the surface of a carrier protein (Fig. 16.2) and used as a vaccine. Among the peptides used, the one corresponding to amino acids 141 to 160 was found to be effective in immunizing guinea pigs against FMD. In addition, when two peptides were joined together (amino acids 141 to 158, and amino acids 200 to 213), they served as more efficient recombinant vaccines.

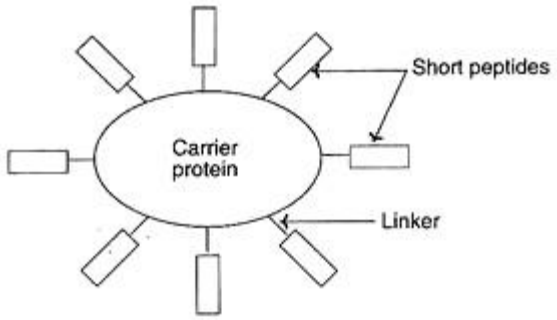


Fig. 16.2 : A general structure of a synthetic peptide vaccine.

The success so far to use recombinant peptides as vaccines has been very limited. This is mainly because a short peptide usually is not enough to be sufficiently immunogenic, since it may not have the same conformation as that of the original viral particle. However, scientists continue their search for specific, inexpensive and safe synthetic peptide vaccines for various diseases.

Herpes Simplex Virus:

Herpes simplex virus (HSV) is an oncogenic (cancer-causing) virus. In addition, it also causes sexually transmitted diseases, encephalitis and severe eye infections. Attempts have been made to produce subunit vaccines against HSV. An envelope glycoprotein D (gD) of HSV that can elicit antibody production has been identified. This is a membrane bound protein, and difficult to isolate and purify. The glycoprotein D was modified by deleting the trans membrane portion of the protein (Fig. 16.3) and the gene was modified.

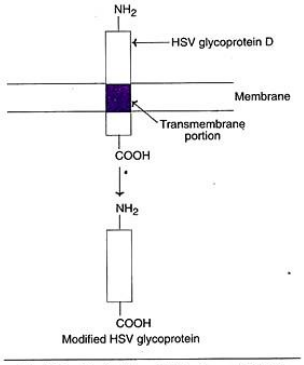


Fig. 16.3 : Modification of HSV glycoprotein D by deleting transmembrane portion. (HSV–Herpes simplex virus)

This gene for gD was cloned in a mammalian vector and expressed in Chinese hamster ovary (CHO). The advantage here is that the protein can get glycosylated (unlike in E. coli system). In the experimental trials, the modified form of gD was found to be effective against HSV.

Tuberculosis:

Tuberculosis is caused by the bacterium Mycobacterium tuberculosis. It is often fatal, and as per some estimates nearly 3 million deaths occur every year due to this highly infectious disease. Antibiotics are used to treat tuberculosis. However, drug-resistant M. tuberculosis strains have been developed making the drug therapy sometimes ineffective. Vaccination for tuberculosis is therefore, advocated.

Bacillus Calmette-Guerin (BCG) vaccine:

In some countries, particularly the developing ones, BCG vaccine is widely used to protect against tuberculosis. However, countries like United States have not approved BCG vaccination for various reasons. BCG vaccine itself causes tuberculosis in some individuals (AIDS victims) and the vaccinated people respond positively for laboratory diagnosis of tuberculosis.

Subunit vaccines:

The secretory (extracellular proteins of M. tuberculosis have been purified and used for immuno-protection against tuberculosis. Of about 100 such proteins, six were found to be useful (either individually or in combination) to immunize guinea pigs. Attempts are underway to develop recombinant subunit vaccine against tuberculosis.

Meningitis:

Group B strain of meningococci, namely Neisseria meningitidis causes meningitis in adolescents and young adults. Meningitis is characterized by inflammation of the membranes covering brain and spinal cord. The symptoms include headache, photophobia, irritability, and neck stiffness. Pizza et al (2000) made a novel approach to develop a vaccine against meningitis. They identified 350 proteins (potential protective antigens) and the entire sequence of genome coding for these proteins in N. meningitidis. All the 350 candidate antigens were expressed in E. coli, purified and used to immunize mice. A good bactericidal antibody response was observed in these mice.

AIDS (Acquired Immunodeficiency Syndrome):

AIDS is a retroviral disease caused by human immunodeficiency virus (HIV). This disease is characterized by immunosuppression, neo-plasma and neurological manifestations. AIDS is invariably fatal, since as of now there is no cure. Development of a vaccine against AIDS is a top priority by DNA technologists world over. In fact, vaccines are being continuously developed and field tested, although there has been no success so far.

Subunit vaccines:

The development of two subunit vaccines, specifically the glycoproteins of HIV envelope is described here. The functions of gp120 and gp41 of HIV are illustrated in Fig. 16.4A. The glycoprotein gp120 projects out of the HIV envelope while the other glycoprotein gp41 lies beneath gp120.

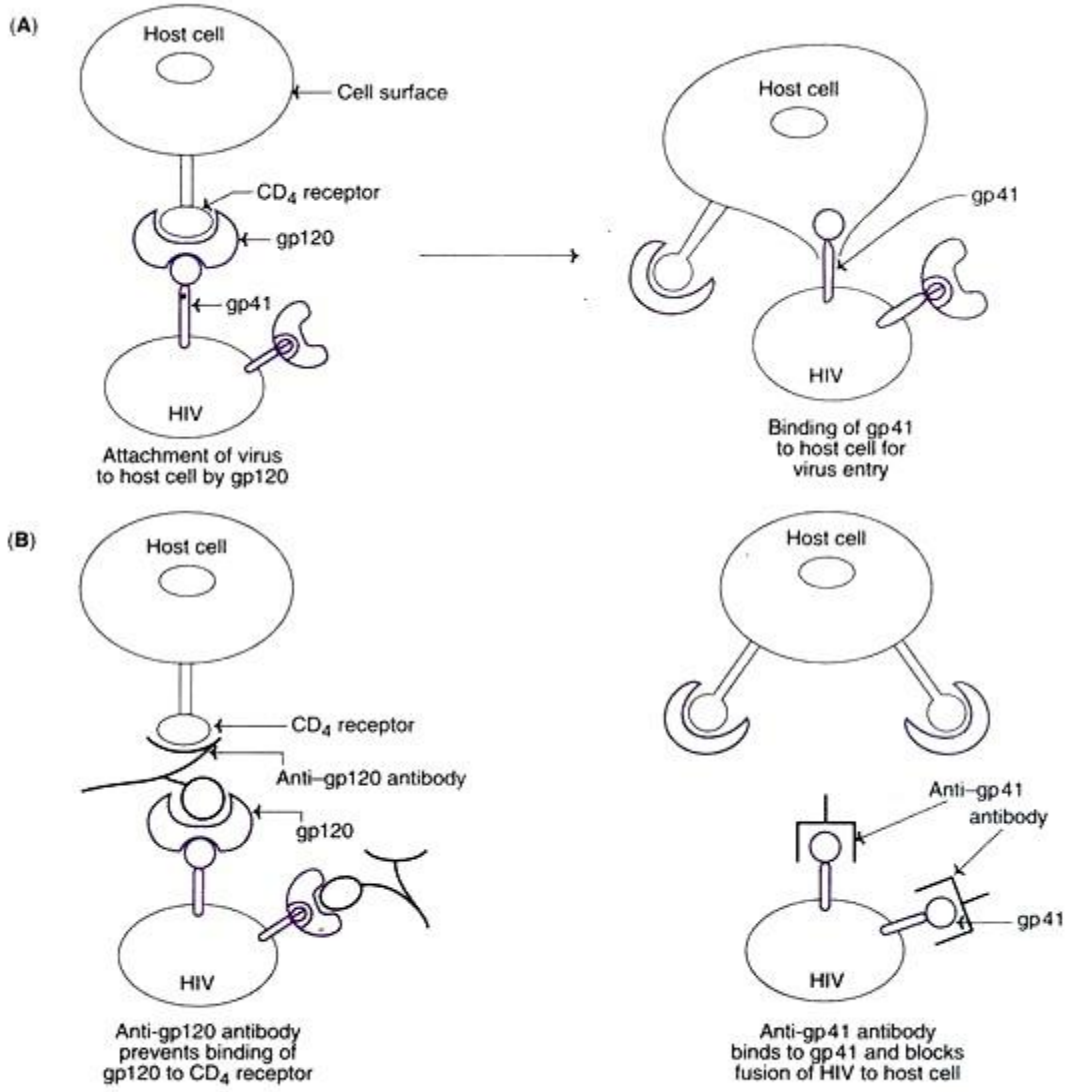


Fig. 16.4 : Subunit recombinant vaccines against AIDS. (A) The functions of gp120 and gp41 of HIV; (B) The action of anti–gp120 and anti–gp41 antibodies to prevent the entry of HIV into host cell.

On entering the body, the HIV binds to the host cells (T-lymphocytes) by attaching gp120 to the CD_4 receptor sites on the cell surface. This attachment uncovers gp41 molecules and the viral envelope. Now gp41 binds to the host cell surface and opens a passage for the entry of the virus into the cell.

Biotechnologists have isolated the genes for gp120 and gp41 and inserted them into the bacterium E. coli. These bacterial cells produce gp120 and gp41 that can be used as recombinant vaccines against AIDS. The action of gp120 and gp41 in immunizing host T-lymphocytes is depicted in Fig. 16.4B.

The gp120 molecules stimulate the host immune system to produce anti- gp120 antibodies. These antibodies bind to gp120 and prevent its attachment to CD_4. In a comparable manner,

gp41 molecules also result in the production of anti-gp41 antibodies. These antibodies also bind to gp41 and block the virus- host cell union. The net result of using gp120 and gp41 vaccines is that the entry of HIV into the host cells is prevented.

Vaccine against AIDS— not yet a reality:

The description of vaccine development against AIDS (given above), which appears attractive is not so simple. The most important limitations are that the HIV has high frequency of mutations. Therefore the vaccines developed cannot bind to the new virus (i.e., mutated one). In addition, gp120 and gp41 are very poor stimulators of immune system. Despite these limitations, scientists have not lost hope, and continue their research to develop vaccines against AIDS.

DNA Vaccines (Genetic Immunization):

Genetic immunization by using DNA vaccines is a novel approach that came into being in 1990. The immune response of the body is stimulated by a DNA molecule. A DNA vaccine consists of a gene encoding an antigenic protein, inserted onto a plasmid, and then incorporated into the cells in a target animal.

The plasmid carrying DNA vaccine normally contains a promoter site, cloning site for the DNA vaccine gene, origin of replication, a selectable marker sequence (e.g. a gene for ampicillin resistance) and a terminator sequence (a poly—A tail).

DNA vaccine—plasmids can be administered to the animals by one of the following delivery methods.

i. Nasal spray
ii. Intramuscular injection
iii. Intravenous injection
iv. Intradermal injection
v. Gene gun or biolistic delivery (involves pressure delivery of DNA-coated gold beads).

DNA Vaccine and Immunity:

An illustration of a DNA vaccine and the mechanism of its action in developing immunity is given in Fig. 16.5. The plasmid vaccine carrying the DNA (gene) for antigenic protein enters the nucleus of the inoculated target cell of the host. This DNA produces RNA, and in turn the specific antigenic protein. The antigen can act directly for developing humoral immunity or as fragments in association with major histocompatibility class (MHC) molecules for developing cellular immunity.

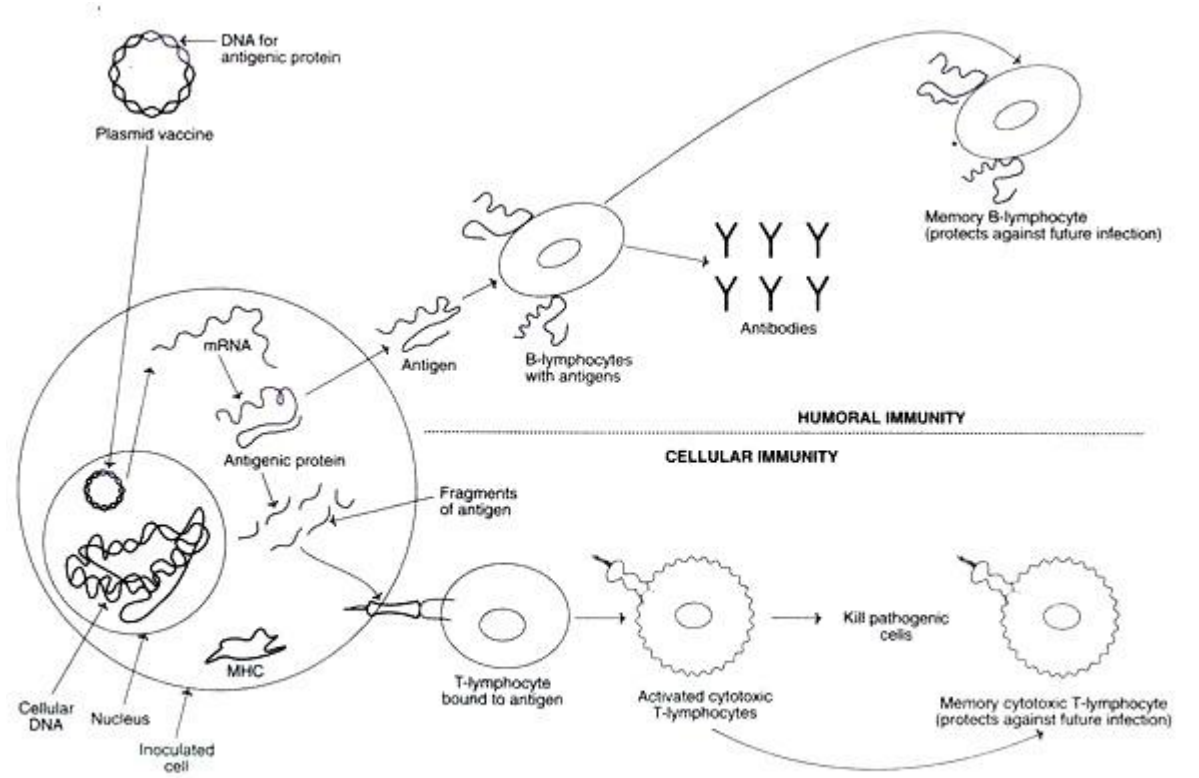


Fig. 16.5 : DNA vaccine and mechanism of its action in developing immunity (MHC–Major histocompatability complex molecule)

Humoral immunity:

As the antigen molecules bind to B- lymphocytes, they trigger the production of antibodies which can destroy the pathogens. Some of the B-lymphocytes become memory cells that can protect the host against future infections.

Cellular immunity:

The protein fragments of the antigen bound to MHC molecules can activate the cytotoxic T-lymphocytes. They are capable of destroying the infected pathogenic cells. Some of the activated T-lymphocytes become memory cells which can kill the future infecting pathogens.

Complementary DNA vaccines:

For genetic immunization, complementary DNA (cDNA) vaccines can also be used. Some workers have successfully used cDNA as vaccines e.g. immunization of mice against influenza.

DNA vaccines for production of antigens and antibodies:

A novel approach for the production of antigens as well as antibodies by DNA vaccine was developed in 1997. In the experiments conducted in mice, researchers injected plasmids containing the genes for malarial parasite and also the genes for the antibodies against malarial parasite. The B-lymphocytes 4 of these mice performed a double duty, and produced antigens for and antibodies against malarial parasite.

The antigens stimulate to produce more and more antibodies. The antibodies so produced react with malarial parasite. The generation of antigens and antibodies by using a DNA vaccine is a

recent development in immunology, and is referred to as antigenic antibody approach of DNA vaccine.

Screening of pathogenic genome for selecting DNA vaccines:

The ultimate goal of scientists is to choose the right DNA fragment from the pathogen to serve as a vaccine for the strongest immune response against the invading pathogen. For this purpose, the pathogen's DNA can be broken into fragments and a large number of vaccines DNA—plasmids can be prepared.

The immune response for each one of the DNA vaccines can be studied by injecting the pathogen. By screening the DNA fragments of the pathogenic genome, it is possible to choose one or few DNA vaccines that can offer maximal immune protection.

Advantages of DNA vaccines:

There are several advantages of using DNA vaccines in immunization:

1. The tedious and costly procedures of purifying antigens or creating recombinant vaccines are not necessary.
2. DNA vaccines are very specific in producing the target proteins (antigens or antibodies). Thus, they trigger immune response only against the specific pathogen.
3. In general, DNA vaccines elicit much higher immune response compared to other kinds of vaccines.
4. DNA vaccines are more stable for temperature variations (low or high) than the conventional vaccines. Thus, the storage and transport problems associated with vaccines are minimal.
5. The delivery methods to the host are simpler for DNA vaccines.

Disadvantages of DNA vaccines:

1. The fate of the DNA vaccine in the host cells is not yet clear. There is a possibility of this DNA getting integrated into the host genome and this may interrupt the normal functions.
2. There also exists a danger of cancer due to DNA vaccines.
3. The post-translational modification of the gene (DNA vaccine) product in host cells may not be the same as that found in the native antigen.

Present status of DNA vaccines:

Since 1990, several groups of workers world-over have been trying to develop DNA vaccines against various diseases in experimental animals. Genetic immunization has been done against a number of pathogenic organisms. These include influenza A virus, rabies virus, hepatitis B virus, bovine herpes virus, HIV type I, and Plasmodium species (malarial parasite). It must be noted that DNA vaccines have not been tried in humans for obvious reasons. The most important being the unknown risks of these foreign DNAs in human subjects.

RNA Vaccines:

Several workers are trying to use RNA molecules as vaccines. These RNAs can readily synthesize the antigenic proteins and offer immunity. But unfortunately, RNAs are less stable than DNAs. This poses a big problem for RNA vaccine manufacture and distribution. Therefore, the progress in the development of RNA vaccines has been rather slow compared to DNA vaccines.

Plants as Edible Subunit Vaccines:

Plants serve as a cheap and safe production systems for subunit vaccines. The edible vaccines can be easily ingested by eating plants. This eliminates the processing and purification

procedures that are otherwise needed. Transgenic plants (tomato, potato) have been developed for expressing antigens derived from animal viruses (rabies virus, herpes virus). A selected list of recombinant vaccines against animal viruses produced in plants is given in Table 16.2.

TABLE 16.2 A selected list of plant edible subunit vaccines

Antigen	*Host plant*
Rabies glycoprotein	Tomato
Foot and mouth virus (VPI)	*Arabidopsis*
Herpes virus B surface antigen	Tobacco
Cholera toxin B subunit	Potato
Human cytomegalovirus glycoprotein B	Tobacco

Edible vaccine production and use:

The production of vaccine potatoes is illustrated in Fig. 16.6. The bacterium, Agrobacterium tumefaciens is commonly used to deliver the DNA (genetic material) for bacterial or viral antigens. A plasmid carrying the antigen gene and an antibiotic resistance gene are incorporated into the bacterial cells (A. tumefaciens).

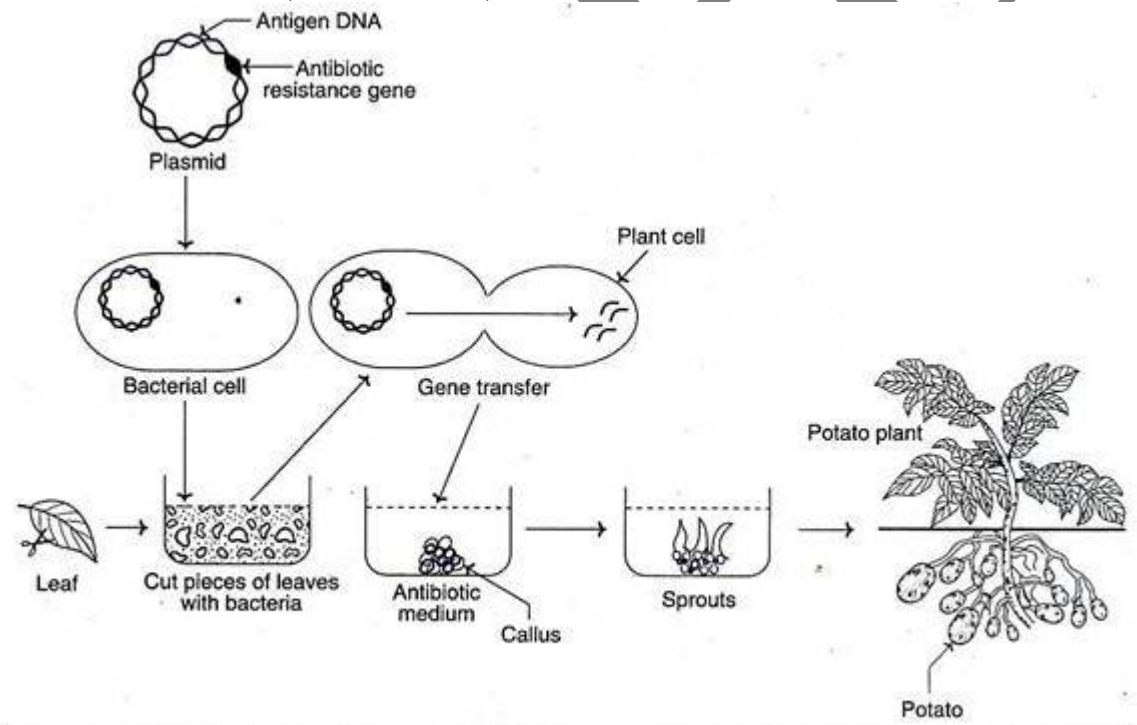


Fig. 16.6 : An illustration for edible vaccine production.

The cut pieces of potato leaves are exposed to an antibiotic which can kill the cells that lack the new genes. The surviving cells (i.e., gene altered ones) can multiply and form a callus (clump of cells). This callus is allowed to sprout shoots and roots, which are grown in soil to form plants.

In about three weeks, the plants bear potatoes with antigen vaccines. The first clinical trials in humans, using a plant- derived vaccine were conducted in 1997. This involved the ingestion of transgenic potatoes with a toxin of E. coli causing diarrhea.

Type # 2. Attenuated Recombinant Vaccines:

In the early years of vaccine research, attenuated strains of some pathogenic organisms were prepared by prolonged cultivation — weeks, months or even years. Although the reasons are not known, the infectious organism would lose its ability to cause disease but retains its capability to act as an immunizing agent. This type of approach is almost outdated now.

It is now possible to genetically engineer the organisms (bacteria or viruses) and use them as live vaccines, and such vaccines are referred to as attenuated recombinant vaccines. The genetic manipulations for the production of these vaccines are broadly of two types:

1. Deletion or modification of virulence genes of pathogenic organisms.
2. Genetic manipulation of non-pathogenic organisms to carry and express antigen determinants from pathogenic organisms.

The advantage with attenuated vaccines is that the native conformation of the immunogenic determinants is preserved; hence the immune response is substantially high. This is in contrast to purified antigens which often elicit poor immunological response.

Some of the important attenuated vaccines developed by genetic manipulations are briefly described.

Cholera:

Cholera is an intestinal disease characterized by diarrhea, dehydration, abdominal pain and fever. It is caused by the bacterium, Vibrio cholera. This pathogenic organism is transmitted by drinking water contaminated with fecal matter. Cholera epidemics are frequently seen in developing countries where the water purification and sewage disposal systems are not well developed.

On entering the small intestine, V. cholera colonizes and starts producing large amounts of a toxic protein, a hexameric enterotoxin. This enterotoxin stimulates the cells lining intestinal walls to release sodium, bicarbonate and other ions. Water accompanies these ions leading to severe diarrhea, dehydration, and even death.

The currently used cholera vaccine is composed of phenol-killed V. cholera. The immuno-protection, lasting for 3-6 months is just moderate. Attempts are being made to develop better vaccines. The DNA technologists have identified the gene encoding enterotoxin (toxic protein). Enterotoxin, an hexamer, consists of one A subunit and five identical B subunits. The A subunit has two functional domains-the A_1 peptide which possesses the toxic activity and A_2 peptide that joins A subunit to B subunits.

By genetic engineering, it was possible to delete the DNA sequence encoding A_1 peptide and create a new strain of V. cholera. This strain is non-pathogenic, since it cannot produce enterotoxin. The genetically engineered V. cholera is a good candidate to serve as an attenuated vaccine.

Creating a new strain of V. cholera:

The development of a new strain of Vibrio cholera that can effectively serve as an attenuated recombinant vaccine is depicted in Fig. 16.7, and briefly described below.

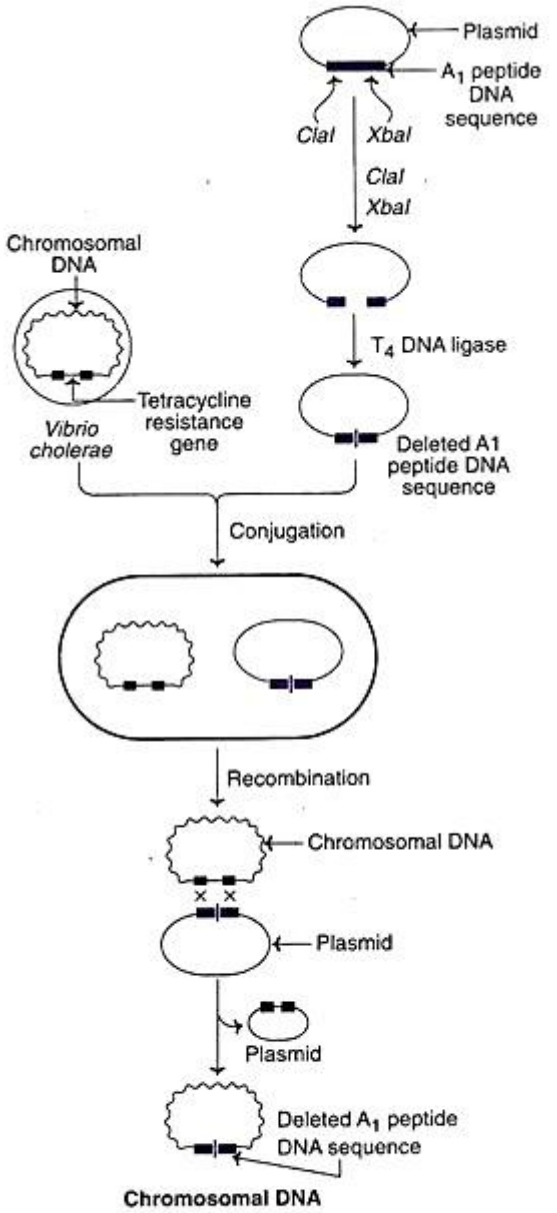


Fig. 16.7 : Development of a new strain of V. cholerae *as an attenuated recombinant vaccine (*ClaI *and* XbaI *are restriction endonucleases).*

1. A tetracycline resistance gene was inserted into the A_1 peptide sequence of V. cholera chromosome. This destroys the DNA sequence encoding for A_1 peptide, besides making the strain resistant to tetracycline. Unfortunately, the tetracycline resistant gene is easily lost and the enterotoxin activity is restored. Because of this, the new strain of V. cholera as such cannot be used as a vaccine.

2. The DNA sequence of A_1 peptide is incorporated into a plasmid, cloned and digested with restriction enzymes (Clal and Xbal). In this manner, the A_1 peptide coding sequence is deleted (the DNA encoding for 183 of the 194 amino acids of the A_1 peptide is actually removed). By using T_4 DNA ligase, the plasmid is re-circularized. This plasmid contains a small portion of A_1 peptide coding sequence.
3. The plasmid, containing the deleted A_1 peptide sequence is transferred by conjugation into the V. cholera strain carrying a tetracycline resistance gene.
4. Recombination can occur between the plasmid (containing a small portion of peptide A_1coding sequence) and the chromosome of V. cholera (carrying tetracycline resistance gene). The result of this double crossover is the formation of V. cholera containing a chromosomal DNA lacking A_1 peptide DNA sequence. As the bacterium, V. cholera multiplies, the plasmids are lost in the next few generations.
5. The V. cholera cells defective in A_1 peptide are selected, based on tetracycline sensitivity. It may be noted that this new strain lacks tetracycline resistance gene.
The genetically engineered V. cholera cells with deleted A_1 peptide DNA sequence are quite stable. They cannot produce active enterotoxin but possess all other biochemical functions of the pathogen. This new strain of V. cholera is undergoing trials for its efficiency as a vaccine. Preliminary results indicate that this attenuated vaccine can protect about 90% of the volunteers against cholera. But there are some side effects. Scientists continue their work to develop a better vaccine against cholera.

Potato as a vehicle for cholera vaccine:

A group workers have developed a gene altered potato containing attenuated cholera vaccine. These potatoes when fed to mice induced immunity against cholera.

Salmonella Species:

The different strains of Salmonella genus are responsible for causing typhoid, enteric fever, food poisoning and infant death. Immunoprotection against Salmonella pathogens is really required. Some workers have been successful in deleting aro genes and pur genes in Salmonella.
Aro genes encode for the enzymes responsible for the biosynthesis of aromatic compounds, while pur genes encode for enzymes of purine metabolism. The new strains of Salmonella can be grown in vitro on a complete medium.
The doubly deleted strains have a very restricted growth in vivo, while they can stimulate immunological response. The genetically engineered attenuated vaccines of Salmonella have been shown to be effective as oral vaccines in experimental animals (mice, cattle, sheep, and chickens). Some workers claim that the new strain of Salmonella offers immunoprotection in humans also.

Leishmania Species:

Leishmania species are flagellated protozoan parasites and are responsible for the disease leishmaniasis. This disease is characterized by cutaneous, visceral and mucosal leisons. Leishmaniasis is transmitted by sand flies.
An attenuated strain of leishmania has been created and successfully used in mice to offer immunoprotection against leishmaniasis. In Leishmania major, the genes encoding dihydrofolate reductase-thymidylate synthase can be replaced by the genes encoding resistance to antibiotics G-418 and hygromycin.

This new strain of L. major invariably requires thymidine in the medium for its growth and multiplication. The attenuated strain of L. major can survive only a few days when administered to mice. This short period is enough to induce immunity in mice against the lesions of leishmania. However, more experiments on animals have to be carried out before the leishmania attenuated vaccine goes for human trials.

Type # 3. Vector Recombinant Vaccines:

Some of vectors can be genetically modified and employed as vaccines against pathogens.

Vaccines against Viruses- Vaccinia Virus:

Vaccinia viruses is basically the vaccine that was originally used by Jenner for the eradication of smallpox. The molecular biology of this virus has been clearly worked out. Vaccinia virus contains a double-stranded DNA (187 kb) that encodes about 200 different proteins. The genome of this virus can accommodate stretches of foreign DNA which can be expressed along with the viral genes.

The vaccinia virus can replicate in the host cell cytoplasm (of the infected cells) rather than the nucleus. This is possible since the vaccinia virus possesses the machinery for DNA replication, transcription-DNA polymerase, RNA polymerase etc. The foreign genes inserted into the vaccinia virus can also be expressed along with the viral genome. Thus, the foreign DNA is under the control of the virus, and is expressed independently from the host cell genome.

The vaccinia viruses are generally harmless, relatively easy to cultivate and stable for years after lyophilization (freeze-drying). All these features make the vaccinia virus strong candidates for vector vaccine. The cloned foreign genes (from a pathogenic organism) can be inserted into vaccinia virus genome for encoding antigens which in turn produces antibodies against the specific disease- causing agent.

The advantage with vector vaccine is that it stimulates B-lymphocytes (to produce antibodies) and T-lymphocytes (to kill virus infected cells). This is in contrast to a subunit vaccine which can stimulate only B-lymphocytes. Thus, vaccinia virus can provide a high level of immunoprotection against pathogenic organisms. Another advantage of vaccinia virus is the possibility of vaccinating individuals against different diseases simultaneously. This can be done by a recombinant vaccinia viruses which carries genes encoding different antigens.

Antigen genes for certain diseases have been successfully incorporated into vaccinia virus genome and expressed. Thus, vector vaccines have been developed against hepatitis, influenza, herpes simplex virus, rabies, angular stomatitis virus and malaria. However, none of these vaccines has been licensed for human use due to fear of safety. It is argued that recombinant vaccinia virus might create life threatening complications in humans.

Production of recombinant vaccinia viruses:

The development of recombinant vaccinia virus is carried out by a two-step procedure (Fig. 16.8).

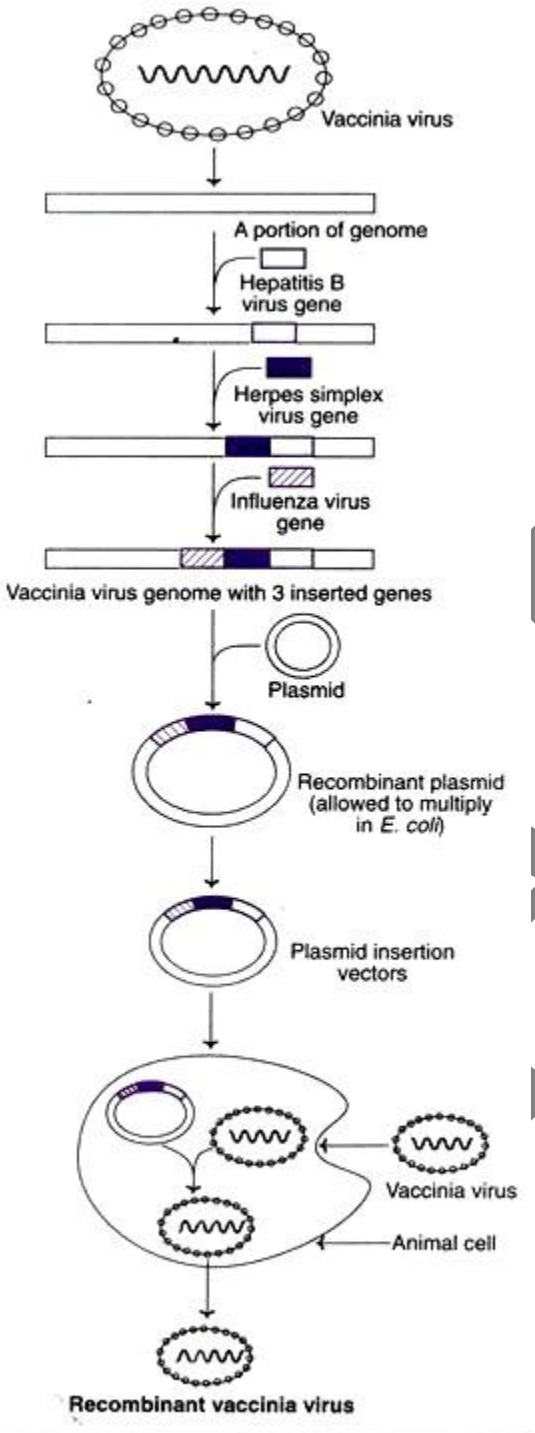


Fig. 16.8 : Production of recombinant vaccinia virus.

1. Assembly of plasmid insertion vector:

Fresh vaccinia (cow pox) viruses are processed to release their DNAs. Now genes from hepatitis B virus, herpes simplex virus and influenza virus are added one after another and inserted into vaccinia virus genome. These DNA clusters are cloned in E. coli for increasing their number and to produce plasmid insertion vectors. The plasmid contains the foreign subunit genes, the natural vaccinia genes, including the promoter genes. The recombinant plasmids are isolated and purified and serve as plasmid insertion vectors.

2. Production of recombinant vaccinia viruses:

The animal cells are infected with plasmid insertion vectors and normal vaccinia viruses. As the viral replication occurs, the plasmids are taken up to produce recombinant vaccinia viruses. The plasmid insertion vector incorporates its genes into vaccinia virus genome at a place that encodes for the enzyme thymidine kinase (TK).

Thus the recombinant viruses have lost their ability to produce TK. There are two advantages of loss of TK gene. One is that it is easy to select recombined vaccinia viruses that lack TK gene and the second is that these viruses are less infectious than the normal viruses. The recombinant vaccinia viruses, released from the cultured animal cells, can be successfully used as vaccines. These live viral vaccines have some advantages over the killed or subunit vaccines.

Advantages:

1. Authenticated antigens that closely resemble natural antigens can be produced.
2. The virus can replicate in the host cells. This enables the amplification of the antigens for their action on B-lymphocytes and T-lymphocytes.
3. There is a possibility of vaccinating several diseases with one recombinant vaccinia virus.

Disadvantages:

1. The most important limitation is the yet unknown risks of using these vaccines in humans.
2. There may be serious complications of using recombinant viral vaccines in immunosuppressed individuals such as AIDS patients.

Other viral recombinant vaccines:

Most of the work on the development of live viral vaccines has been carried out on vaccinia virus. Other viruses such as adenovirus, poliovirus and varicella-zoster virus are also being tried as recombinant vaccines. Scientists are attracted to develop a recombinant poliovirus as it can be orally administered. It might take many more years for the recombinant viral vaccines to become a reality for human use.

Delivery of Antigens by Bacteria:

It is known that the antigens located on the surface of a bacterial cell are more immunogenic than the antigens in the cytoplasm. Based on this observation, scientists have developed strategies to coat the surfaces of non-pathogenic organisms with antigens of pathogenic bacteria.

Flagellin is a protein present in the fragella (thread like filaments) of Salmonella. A synthetic oligonucleotide encoding the epitope of cholera toxin B subunit was inserted into Salmonella flagellin gene. This epitope was in fact found on the flagellum surface. These flagella-engineered bacteria, when administered to mice, raised antibodies against the cholera toxin B subunit peptide. It may be possible in future to incorporate multiple epitopes (2 or 3) into the flagellin gene to create multivalent bacterial vaccines.

Gene therapy

Advances in biochemistry and molecular biology have helped to understand the genetic basis of inherited diseases. It was a dream of the researchers to replace the defective genes with good ones, and cure the genetic disorders.

Gene therapy is the process of inserting genes into cells to treat diseases. The newly introduced genes will encode proteins and correct the deficiencies that occur in genetic diseases. Thus, gene therapy primarily involves genetic manipulations in animals or humans to correct a disease, and keep the organism in good health. The initial experiments on gene therapy are carried out in animals, and then in humans. Obviously, the goal of the researchers is to benefit the mankind and improve their health.

An overview of gene therapy strategies is depicted in Fig. 13.1. In gene augmentation therapy, a DNA is inserted into the genome to replace the missing gene product. In case of gene inhibition therapy, the antisense gene inhibits the expression of the dominant gene.

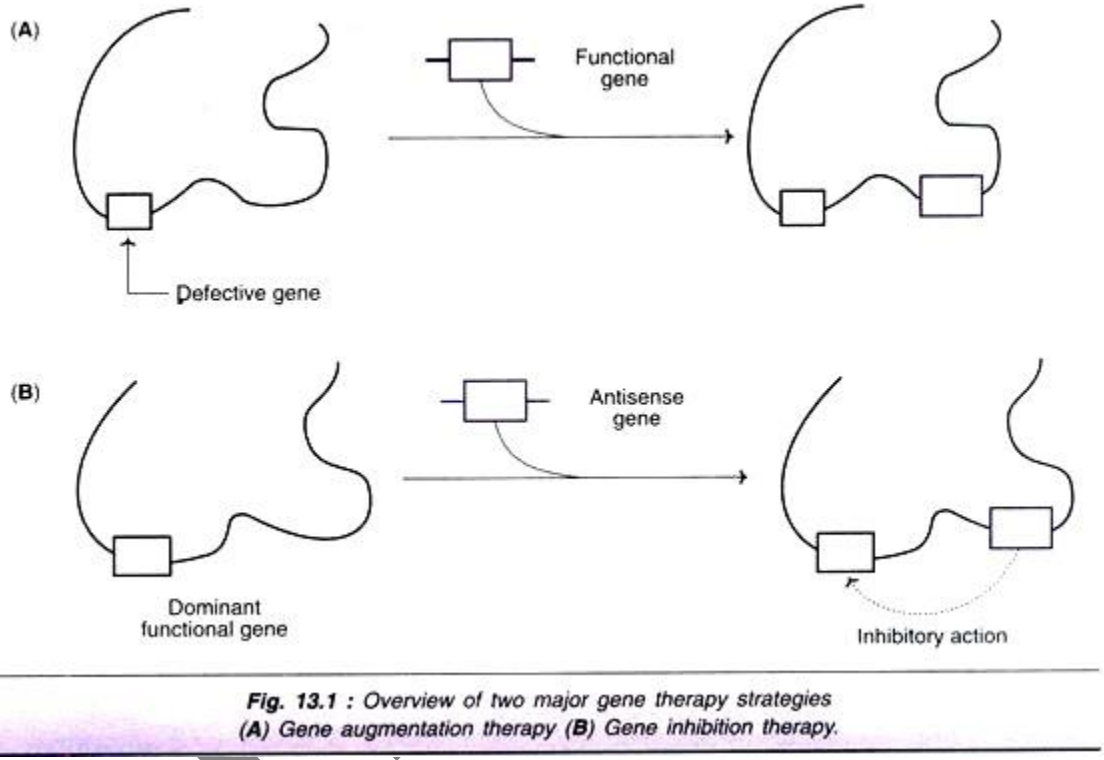


Fig. 13.1 : Overview of two major gene therapy strategies (A) Gene augmentation therapy (B) Gene inhibition therapy.

Approaches for Gene Therapy:

There are two approaches to achieve gene therapy.

1. Somatic Cell Gene Therapy:

The non- reproductive (non-sex) cells of an organism are referred to as somatic cells. These are the cells of an organism other than sperm or eggs cells, e.g., bone marrow cells, blood cells, skin cells, intestinal cells. At present, all the research on gene therapy is directed to correct the genetic defects in somatic cells. In essence, somatic cell gene therapy involves the insertion of a fully functional and expressible gene into a target somatic cell to correct a genetic disease permanently.

2. Germ Cell Gene Therapy:

The reproductive (sex) cells of an organism constitute germ cell line. Gene therapy involving the introduction of DNA into germ cells is passed on to the successive generations. For safety, ethical and technical reasons, germ cell gene therapy is not being attempted at present.

The genetic alterations in somatic cells are not carried to the next generations. Therefore, somatic cell gene therapy is preferred and extensively studied with an ultimate objective of correcting human diseases. Development of gene therapy in humans for any specific disease involves the following steps. In fact, this is a general format for introducing any therapeutic agent for human use.

1. In vitro experiments and research on laboratory animals (pre-clinical trials).
2. Phase I trials with a small number (5-10) of human subjects to test safety of the product.
3. Phase II trials with more human subjects to assess whether the product is helpful.
4. Phase III trials in large human samples for a final and comprehensive analysis of the safety and efficacy of the product.

As such, gene therapy involves a great risk. There are several regulatory agencies whose permission must be sought before undertaking any work related to gene therapy. Recombinant DNA Advisory Committee (RAC) is the supervisory body of the National Institute of Health, U.S.A., that clears proposals on experiments involving gene therapy. A large number of genetic disorders and other diseases are currently at various stages of gene therapy trials. A selected list of some important ones is given in Table 13.1.

TABLE 13.1 Human gene therapy trials

Disease	*Gene therapy*
Severe combined immunodeficiency (SCID)	Adenosine deaminase (ADA).
Cystic fibrosis	Cystic fibrosis transmembrane regulator (CFTR).
Familial hypercholesterolemia	Low density lipoprotein (LDL) receptor.
Emphysema	α_1-Antitrypsin
Hemophilia B	Factor IX
Thalassemia	α- or β-Globin
Sickle-cell anemia	β-Globin
Lesch-Nyhan syndrome	Hypoxanthine-guanine phosphoribosyltransferase (HGPRT).
Gaucher's disease	Glucocerebrosidase
Peripheral artery disease	Vascular endothelial growth factor (VEGF)
Fanconi anemia	Fanconi anemia C
Melanoma	Tumor necrosis factor (TNF)
Melanoma, renal cancer	Interleukin-2 (IL-2)
Glioblastoma (brain tumor), AIDS, ovarian cancer	Thymidine kinase (herpes simplex virus)
Head and neck cancer	p^{53}
Breast cancer	Multidrug resistance I
AIDS	*rev* and *env*
Colorectal cancer, melanoma, renal cancer	Histocompatability locus antigen-B_7 (HLA-B_7)
Duchenne muscular dystrophy	Dystrophin
Short stature*	Growth hormone
Diabetes*	Glucose transporter-2, (GLUT-2), glucokinase
Phenylketonuria*	Phenylalanine hydroxylase
Citrullinemia*	Arginosuccinate synthetase

The Future of Gene Therapy:

Theoretically, gene therapy is the permanent solution for genetic diseases. But it is not as simple as it appears since gene therapy has several inbuilt complexities. Gene therapy broadly involves isolation of a specific gene, making its copies, inserting them into target tissue cells to make the desired protein. The story does not end here.

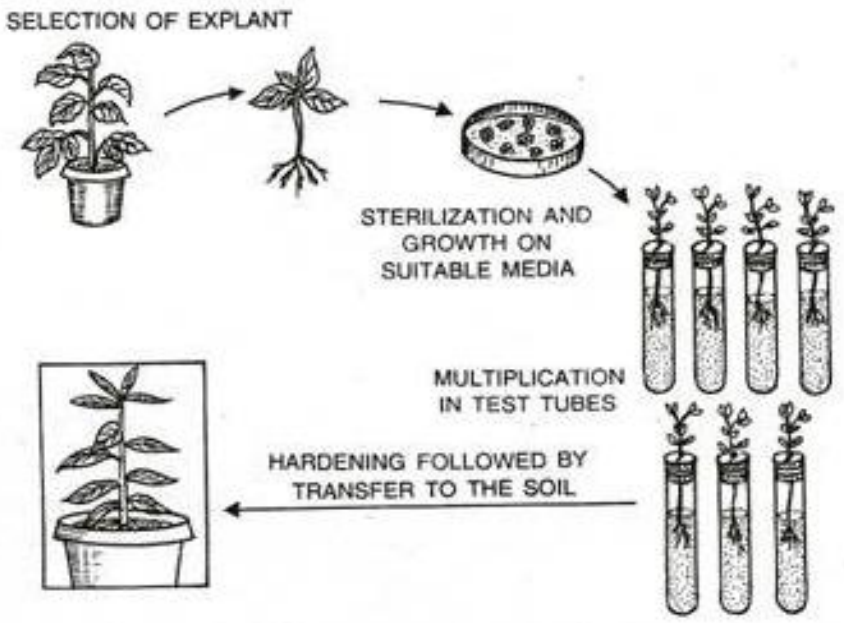
It is absolutely essential to ensure that the gene is harmless to the patient and it is appropriately expressed (too much or too little will be no good). Another concern in gene therapy is the body's immune system which reacts to the foreign proteins produced by the new genes.

The public, in general, have exaggerated expectations on gene therapy. The researchers, at least for the present, are unable to satisfy them. As per the records, by 1999 about 1000 Americans had undergone clinical trials involving various gene therapies.

Unfortunately, the gene therapists are unable to categorically claim that gene therapy has permanently cured any one of these patients! Some people in the media (leading newspapers and magazines) have openly questioned whether it is worth to continue research on gene therapy!

It may be true that as of now, gene therapy due to several limitations, has not progressed the way it should, despite intensive research. But a breakthrough may come anytime, and of course, this is only possible with persistent research. And a day may come (it might take some years) when almost every disease will have a gene therapy, as one of the treatment modalities. And gene therapy will revolutionize the practice of medicine!

Applications in agriculture



Regeneration of whole plants using tissue culture technique.

The use of plant cells to generate useful products and/or services constitutes plant biotechnology. In plant biotechnology, the useful product is a plantlet. The plantlets are used for the following purposes.

1. Rapid Clonal Propagation:

A clone is a group of individuals or cells derived from a single parent individual or cell through asexual reproduction. All the cells in callus or suspension culture are derived from a single explant by mitotic division. Therefore, all plantlets regenerated from a callus/suspension culture generally have the same genotype and constitute a clone. These plantlets are used for rapid clonal propagation. This is done in oil palm.

2. Somaclonal Variation:

Genetic variation present among plant cells of a culture is called somaclonal variation. The term somaclonal variation is also used for the genetic variation present in plants regenerated from a single culture. This variation has been used to develop several useful varieties.

3. Transgenic Plants:

A gene that is transferred into an organism by genetic engineering is known as transgene. An organism that contains and expresses a transgene is called transgenic organism. The transgenes can be introduced into individual plant cells. The plantlets can be regenerated from these cells. These plantlets give rise to the highly valuable transgenic plants.

4. Induction and Selection of Mutations:

Mutagens are added to single cell liquid cultures for induction of mutations. The cells are washed and transferred to solid culture for raising mutant plants. Useful mutants are selected for further breeding. Tolerance to stress like pollutants, toxins, salts, drought, flooding, etc. can also be obtained by providing them in culture medium in increasing dosage. The surviving healthy cells are taken to solid medium for raising resistant plants.

5. Resistance to Weedicides:

It is similar to induction of mutations. Weedicides are added to culture initially in very small concentrations. Dosage is increased in subsequent cultures till the desired level of resistance is obtained. The resistant cells are then regenerated to form plantlets and plants.

Plants Genetic engineering

Genetically modified crops (GMCs, GM crops, or biotech crops) are plants used in agriculture, the DNA of which has been modified using genetic engineering techniques. In most cases, the aim is to introduce a new trait to the plant which does not occur naturally in the species.

Organization of Ti plasmid:

The Ti plasmids (approximate size 200 kb each) exist as independent replicating circular DNA molecules within the Agrobacterium cells. The T-DNA (transferred DNA) is variable in length in the range of 12 to 24 kb, which depends on the bacterial strain from which Ti plasmids come. Nopaline strains of Ti plasmid have one T-DNA with length of 20 kb while octopine strains have two T-DNA regions referred to as T_L and T_R that are respectively 14 kb and 7 kb in length.

A diagrammatic representation of a Ti plasmid is depicted in Fig. The Ti plasmid has three important regions.

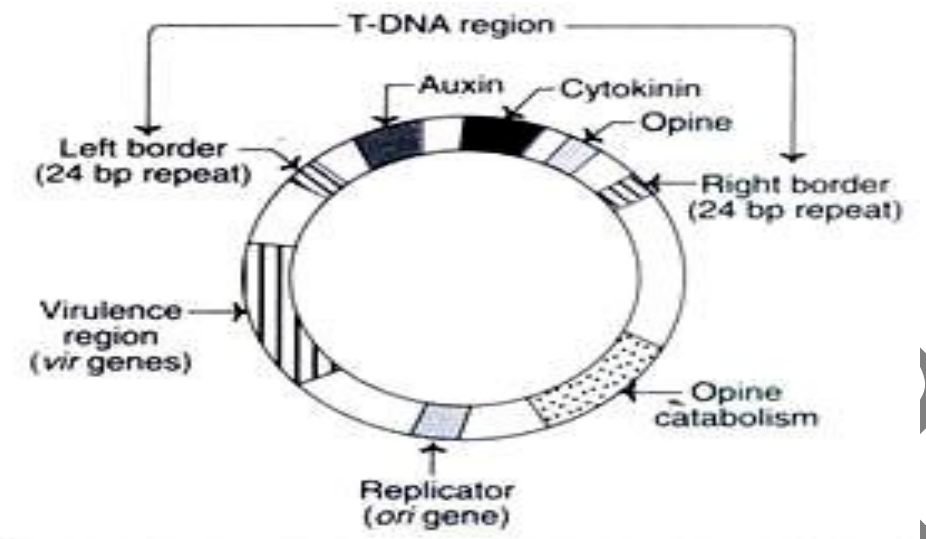


Fig. 49.3 : A dagrammatic representation of a Ti plasmid.

1. T-DNA region:

This region has the genes for the biosynthesis of auxin (aux), cytokinin (cyt) and opine (ocs), and is flanked by left and right borders. These three genes-aux, cyto and ocs are referred to as oncogenes, as they are the determinants of the tumor phenotype.

T-DNA borders — A set of 24 kb sequences present on either side (right and left) of T-DNA are also transferred to the plant cells. It is now clearly established that the right border is more critical for T-DNA transfer and tumori-genesis.

2. Virulence region:

The genes responsible for the transfer of T-DNA into the host plant are located outside T-DNA and the region is referred to as vir or virulence region. Vir region codes for proteins involved in T-DNA transfer. At least nine vir-gene operons have been identified. These include vir A, vir G, vir B_1, vir C_1, vir D_1, D_2 and D_4, and vir E_1, and E_2.

3. Opine catabolism region:

This region codes for proteins involved in the uptake and metabolisms of opines. Besides the above three, there is ori region that is responsible for the origin of DNA replication which permits the Ti plasmid to be stably maintained in A. tumefaciens.

T-DNA transfer and integration:

The process of T-DNA transfer and it integration into the host plant genome is depicted in Fig. 49.4, and is briefly described.

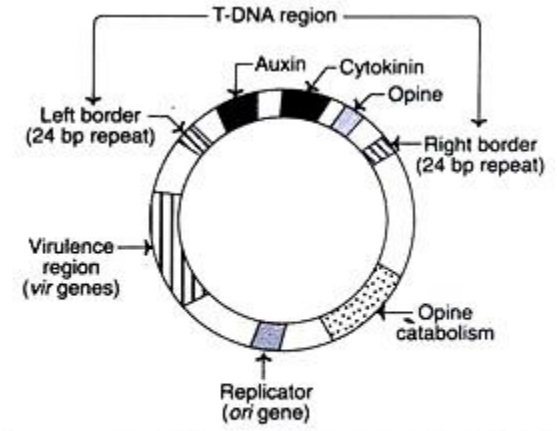


Fig. 49.3 : A dagrammatic representation of a Ti plasmid.

1. Signal induction to Agrobacterium:
The wounded plant cells release certain chemicals- phenolic compounds and sugars which are recognized as signals by Agrobacterium. The signals induced result in a sequence of biochemical events in Agrobacterium that ultimately helps in the transfer of T-DNA of T-plasmid.

2. Attachment of Agrobacterium to plant cells:
The Agrobacterium attaches to plant cells through polysaccharides, particularly cellulose fibres produced by the bacterium. Several chromosomal virulence (chv) genes responsible for the attachment of bacterial cells to plant cells have been identified.

3. Production of virulence proteins:
As the signal induction occurs in the Agrobacterium cells attached to plant cells, a series of events take place that result in the production of virulence proteins. To start with, signal induction by phenolics stimulates vir A which in turn activates (by phosphorylation) vir C. This induces expression of virulence genes of Ti plasmid to produce the corresponding virulence proteins (D1, D2, E_2, B etc.). Certain sugars (e.g. glucose, galactose, xylose) that induce virulence genes have been identified.

4. Production of T-DNA strand:
The right and left borders of T-DNA are recognized by vir D_1/vir D_2proteins. These proteins are involved in the production single-stranded T-DNA (ss DNA), its protection and export to plant cells. The ss T-DNA gets attached to vir D_2.

5. Transfer of T-DNA out of Agrobacterium:
The ss T-DNA — vir D_2 complex in association with vir G is exported from the bacterial cell. Vir B products form the transport apparatus.

6. Transfer of T-DNA into plant cells and integration:
The T-DNA-vir D_2 complex crosses the plant plasma membrane. In the plant cells, T-DNA gets covered with vir E_2. This covering protects the T-DNA from degradation by nucleases; vir D_2 and vir E_2 interact with a variety of plant proteins which influences T-DNA transport and integration.

The T-DNA-vir D_2-vir E_2 — plant protein complex enters the nucleus through nuclear pore complex. Within the nucleus, the T-DNA gets integrated into the plant chromosome through a process referred to illegitimate recombination. This is different from the homologous recombination, as it does not depend on the sequence similarity.

Hairy Root Disease of A. Rhizogenes — R_1 Plasmids:

Agrobacterium rhizogenes can also infect plants. But this results in hairy roots and not crown galls as is the case with A. tumefaciens. The plasmids, of A. rhizogenes have been isolated and characterized. These plasmids, referred to as Ri plasmids, (Ri stands for Root inducing) are of different types. Some of the Ri plasmid strains possess genes that are homologous to Ti plasmid e.g. auxin biosynthetic genes.

Instead of virulence genes, Ri plasmids contain a series of open reading frames on the T-DNA. The products of these genes are involved in the metabolism of plant growth regulators which gets sensitized to auxin and leads to root formation.

Vectors of A. rhizogenes:

As it is done with A tumefaciens, vectors can be constructed by using A. rhizogenes. These vectors are alternate strategies for gene transfer. However, employment of A. rhizogene-based vectors for plant transformation is not common since more efficient systems of A. tumefaciens have been developed.

Importance of hairy roots:

Hairy roots can be cultured in vitro, and thus are important in plant biotechnology. Hairy root systems are useful for the production of secondary metabolites, particularly pharmaceutical proteins.

Ti Plasmid-Derived Vector Systems:

The ability of Ti plasmid of Agrobacterium to genetically transform plants has been described. It is possible to insert a desired DNA sequence (gene) into the T-DNA region (of Ti plasmid), and then use A. tumefaciens to deliver this gene(s) into the genome of plant cell.

In this process, Ti plasmids serve as natural vectors. However, there are several limitations to use Ti plasmids directly as cloning vectors:

i. Ti plasmids are large in size (200-800 kb). Smaller vectors are preferred for recombinant experiments. For this reason, large segments of DNA of Ti plasmid, not essential for cloning, must be removed.

ii. Absence of unique restriction enzyme sites on Ti plasmids.

iii. The phytohormones (auxin, cytokinin) produced prevent the plant cells being regenerated into plants. Therefore auxin and cytokinin genes must be removed.

iv. Opine production in transformed plant cells lowers the plant yield. Therefore opine synthesizing genes which are of no use to plants should be removed.

v. Ti plasmids cannot replicate in E. coli. This limits their utility as E. coli is widely used in recombinant experiments. An alternate arrangement is to add an origin of replication to Ti plasmid that allows the plasmid to replicate in E. coli.

Considering the above limitations, Ti plasmid- based vectors with suitable modifications have been constructed.

These vectors are mainly composed of the following components:

1. The right border sequence of T-DNA which is absolutely required for T-DNA integration into plant cell DNA.
2. A multiple cloning site (poly-linker DNA) that promotes the insertion of cloned gene into the region between T-DNA borders.
3. An origin of DNA replication that allows the plasmids to multiply in E. coli.
4. A selectable marker gene (e.g. neomycin phosphotransferase) for appropriate selection of the transformed cells.

Two types of Ti plasmid-derived vectors are used for genetic transformation of plants— co-integrate vectors and binary vectors.

Co-integrate vector:

In the co-integrate vector system, the disarmed and modified Ti plasmid combines with an intermediate cloning vector to produce a recombinant Ti plasmid (Fig. 49.5).

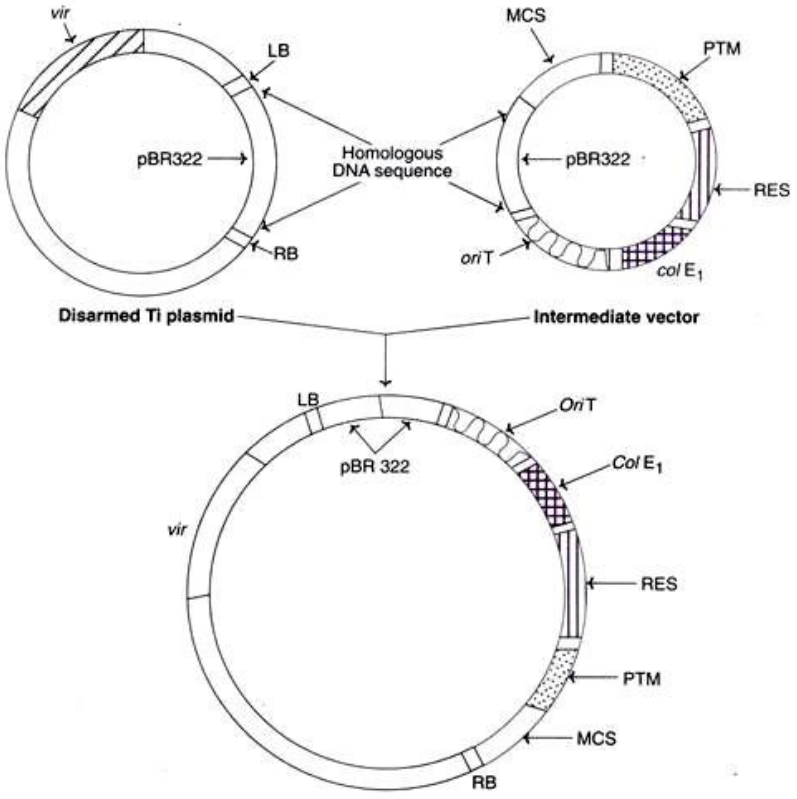


Fig. 49.5 : Cointegrate vector system (vir-Ti plasmid virulence region; pBR322-Bacterial plasmid 322; LB-Left border; RB-Right border; MCS-Multiple cloning site; PTM-Plant transformation marker; RES-Bacterial resistance marker; col E_1-Origin of a replication from col E_1 plasmid; ori T-Origin of transfer site for conjugative plasmid mobilization).

Production of disarmed Ti plasmid:

The T-DNA genes for hormone biosynthesis are removed (disarmed). In place of the deleted DNA, a bacterial plasmid (pBR322) DNA sequence is incorporated. This disarmed plasmid, also

referred to as receptor plasmid, has the basic structure of T-DNA (right and left borders, virulence genes etc.) necessary to transfer the plant cells.

Construction of intermediate vector:

The intermediate vector is constructed with the following components:

i. A pBR322 sequence DNA homologous to that found in the receptor Ti plasmid.

ii. A plant transformation marker (PTM) e.g. a gene coding for neomycin phosphotransferase II (npt II). This gene confers resistance to kanamycin in the plant cells and thus permits their isolation.

iii. A bacterial resistance marker e.g. a gene coding for spectinomycin resistance. This gene confers spectinomycin resistance to recipient bacterial cells and thus permits their selective isolation.

iv. A multiple cloning site (MCS) where foreign genes can be inserted.

v. A Co/E_1 origin of replication which allows the replication of plasmid in E. coli but not in Agrobacterium.

vi. An oriT sequence with basis of mobilization (bom) site for the transfer of intermediate vector from E. coli to Agrobacterium.

Production and use of co-integrate vectors:

The desired foreign gene (target-gene) is first cloned in the multiple cloning site of the intermediate vector. The cloning process is carried out in E. coli, the bacterium where the cloning is most efficient. The intermediate vector is mated with Agrobacterium so that the foreign gene is mobilised into the latter.

The transformed Agrobacterium cells with receptor Ti plasmid and intermediate vector are selectively isolated when grown on a minimal medium containing spectinomycin. The selection process becomes easy since E. coli does not grow on a minimal medium in which Agrobacterium grows.

Within the Agrobacterium cells, intermediate plasmid gets integrated into the receptor Ti plasmid to produce co-integrate plasmid. This plasmid containing plant transformation marker (e.g. npt II) gene and cloned target gene between T-DNA borders is transferred to plant cells. The transformed plant cells can be selected on a medium containing kanamycin when the plant and Agrobacterium cells are incubated together.

Advantages of co-integrate vector:

i. Target genes can be easily cloned

ii. The plasmid is relatively small with a number of restriction sites.

iii. Intermediate plasmid is conveniently cloned in E. coli and transferred to Agrobacterium.

Binary vector:

The binary vector system consists of an Agrobacterium strain along with a disarmed Ti plasmid called vir helper plasmid (the entire T-DNA region including borders deleted while vir gene is retained). It may be noted that both of them are not physically linked (or integrated). A binary vector with T-DNA can replicate in E. coli and Agrobacterium.

A diagrammatic representation of a typical binary vector system is depicted in Fig. 49.6. The binary vector has the following components.

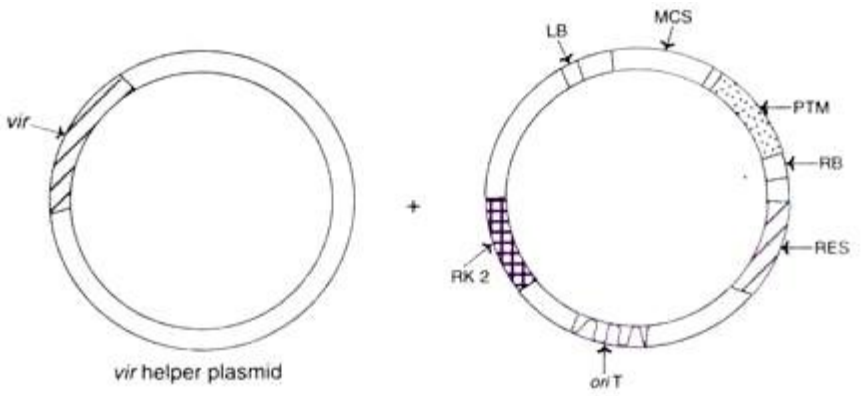


Fig. 49.6 : Binary vector system (vir-Ti plasmid virulence region; LB-Left border; RB-Right border; MCS-Multiple cloning site; PTM-Plant transformation marker; RES-Bacterial resistance marker; oriT-Origin of transfer site for conjugative plasmid mobilization; RK_2-Origin of replication from plasmid).

1. Left and right borders that delimit the T-DNA region.
2. A plant transformation marker (PTM) e.g. npt II that confers kanamycin resistance in plant transformed cells.
3. A multiple cloning site (MCS) for introducing target/foreign genes.
4. A bacterial resistance marker e.g. tetracycline resistance gene for selecting binary vector colonies in E. coli and Agrobacterium.
5. oriT sequence for conjugal mobilization of the binary vector from E. coli to Agrobacterium.
6. A broad host-range origin of replication such as RK_2 that allows the replication of binary vector in Agrobacterium.

Production and use of binary vector:

The target (foreign) gene of interest is inserted into the multiple cloning site of the binary vector. In this way, the- target gene is placed between the right and left border repeats and cloned in E. coli. By a mating process, the binary vector is mobilised from E. coli to Agrobacterium. Now, the virulence gene proteins of T-DNA facilitate the transfer of T-DNA of the vector into plant cells.

Advantages of binary vectors:

i. The binary vector system involves only the transfer of a binary plasmid to Agrobacterium without any integration. This is in contrast to co-integrate vector system wherein the intermediate vector is transferred and integrated with disarmed Ti plasmid.
ii. Due to convenience, binary vectors are more frequently used than co-integrate vectors.

Plant Transformation Technique Using Agrobacterium

Agrobacterium-mediated technique is the most widely used for the transformation of plants and generation of transgenic plants. The important requirements for gene transfer in higher plants through Agrobacterium mediation are listed.

i. The explants of the plant must produce phenolic compounds (e.g. autosyringone) for activation of virulence genes.
ii. Transformed cells/tissues should be capable to regenerate into whole plants.
In general, most of the Agrobacterium-mediated plant transformations have the following basic protocol (Fig. 49.7)

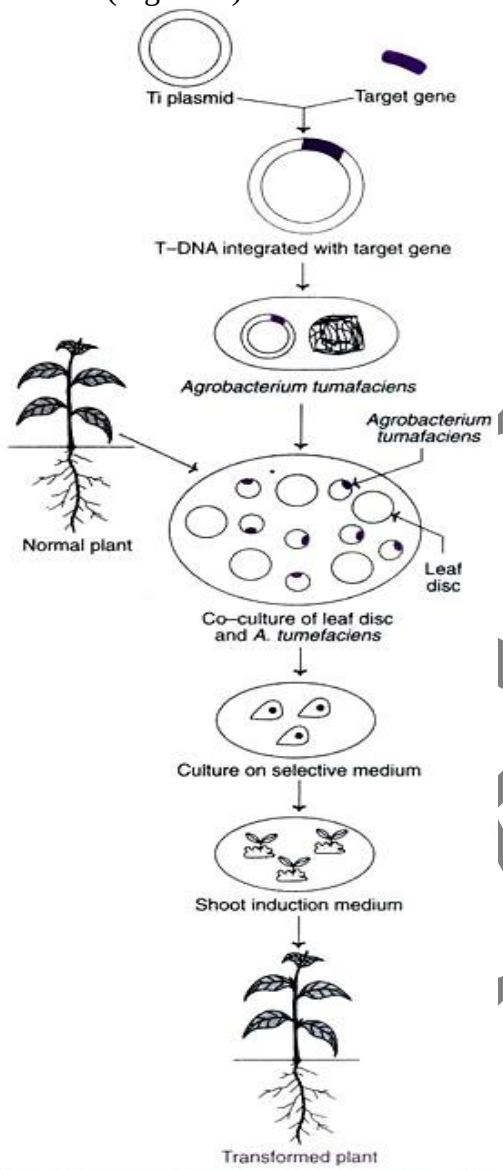


Fig. 49.7 : Transformation technique using Agrobacterium-mediated gene transfer.

1. Development of Agrobacterium carrying the co-integrate or binary vector with the desired gene.
2. Identification of a suitable explant e.g. cells, protoplasts, tissues, calluses, organs.
3. Co-culture of explants with Agrobacterium.
4. Killing of Agrobacterium with a suitable antibiotic without harming the plant tissue.

5. Selection of transformed plant cells.
6. Regeneration of whole plants.

Advantages of Agrobacterium- mediated transformation:

i. This is a natural method of gene transfer.
ii. Agrobacterium can conveniently infect any explant (cells/tissues/organs).
iii. Even large fragments of DNA can be efficiently transferred.
iv. Stability of transferred DNA is reasonably good.
v. Transformed plants can be regenerated effectively.

Limitations of Agrobacterium- mediated transformation:

i. There is a limitation of host plants for Agrobacterium, since many crop plants (monocotyledons e.g. cereals) are not infected by it. In recent years, virulent strains of Agrobacterium that can infect a wide range of plants have been developed.
ii. The cells that regenerate more efficiently are often difficult to transform, e.g. embryonic cells lie in deep layers which are not easy targets for Agrobacterium.

Herbicide Tolerance Technology

Weeds are constant problem in farmers' fields. Weeds not only compete with crops for water, nutrients, sunlight, and space but also harbor insect and disease pests; clog irrigation and drainage systems; undermine crop quality; and deposit weed seeds into crop harvests. If left uncontrolled, weeds can reduce crop yields significantly.

Farmers can fight weeds with tillage, hand weeding, herbicides, or typically a combination of all techniques. Unfortunately, tillage leaves valuable topsoil exposed to wind and water erosion, a serious long-term consequence for the environment. For this reason, more and more farmers prefer reduced or no-till methods of farming.

Similarly, many have argued that the heavy use of herbicides has led to groundwater contaminations, the death of several wildlife species and has also been attributed to various human and animal illnesses.

Advantages of Herbicide Tolerant Crops

- Excellent weed control and hence higher crop yields;
- Flexibility – possible to control weeds later in the plant's growth
- Reduced numbers of sprays in a season;
- Reduced fuel use (because of less spraying);
- Reduced soil compaction (because of less need to go on the land to spray);
- Use of low toxicity compounds which do not remain active in the soil; and
- The ability to use no-till or conservation-till systems, with consequent benefits to soil structure and organisms

Herbicide resistance crops

Weeds (wild herbs) are unwanted and useless plants that grow along with the crop plants. Weeds compete with traps for light and nutrients, besides harbouring various pathogens. It is estimated that the world's crop yield is reduced by 10-15% due to the presence of weeds.

To tackle the problem of weeds, modern agriculture has developed a wide range of weed killers which are collectively referred to as herbicides. In general, majority of the herbicides are broad-spectrum as they can kill a wide range of weeds.

A good or an ideal herbicide is expected to possess the following characteristics:

i. Capable of killing weeds without affecting crop plants.
ii. Not toxic to animals and microorganisms.
iii. Rapidly trans-located within the target plant.
iv. Rapidly degraded in the soil.

None of the commercially available herbicides fulfills all the above criteria. The major limitation of the herbicides is that they cannot discriminate weeds from crop plants. For this reason, the crops are also affected by herbicides, hence the need to develop herbicide-resistant plants. Thus, these plants provide an opportunity to effectively kill the weeds (by herbicides) without damaging the crop plants.

Strategies for engineering herbicide resistance:

A number of biological manipulations particularly involving genetic engineering are in use to develop herbicide-resistant plants.

1. Overexpression of the target protein:

The target protein, being acted by the herbicide can be produced in large quantities so that the affect of the herbicide becomes insignificant. Overexpression can be achieved by integrating multiple copies of the genes and/or by using a strong promoter.

2. Improved plant detoxification:

The plants do possess natural defense systems against toxic compounds (herbicides). Detoxification involves the conversion of toxic herbicide to non-toxic or less toxic compound. By enhancing the plant detoxification system, the impact of the herbicide can be reduced.

3. Detoxification of herbicide by using a foreign gene:

By introducing a foreign gene into the crop plant, the herbicide can be effectively detoxified.

4. Mutation of the target protein:

The target protein which is being affected by the herbicide can be suitably modified. The changed protein should be capable of discharging the functions of the native protein but is resistant to inhibition by the herbicide.

Once the resistant target protein gene is identified, it can be introduced into the plant genomes, and thus herbicide-resistant plants can be developed. For success in the development of herbicide resistant plants, good knowledge of the target protein and the action of herbicides is required.

Some of the developments made in the herbicide resistance of plant are briefly described:

Glyphosate Resistance:

Glyphosate, is a glycine derivative. It acts as a broad-spectrum herbicide and is effective against 76 of the world's worst 78 weeds. Glyphosate is less toxic to animals and is rapidly degraded by microorganisms. In addition, it has a short half-life. The American chemical company Monsanto markets glyphosate as Round up.

Mechanism of action of glyphosate:

Glyphosate is rapidly transported to the growing points of plants. It is capable of killing the plants even at a low concentration. Glyphosate acts as a competitive inhibitor of the enzyme 5-enoylpyruvylshikimate 3-phosphate synthase (EPSPS). This is a key enzyme in shikimic acid pathway that results in the formation of aromatic amino acids (tryptophan, phenylalanine and tyrosine), phenols and certain secondary metabolites (Fig. 50.5).

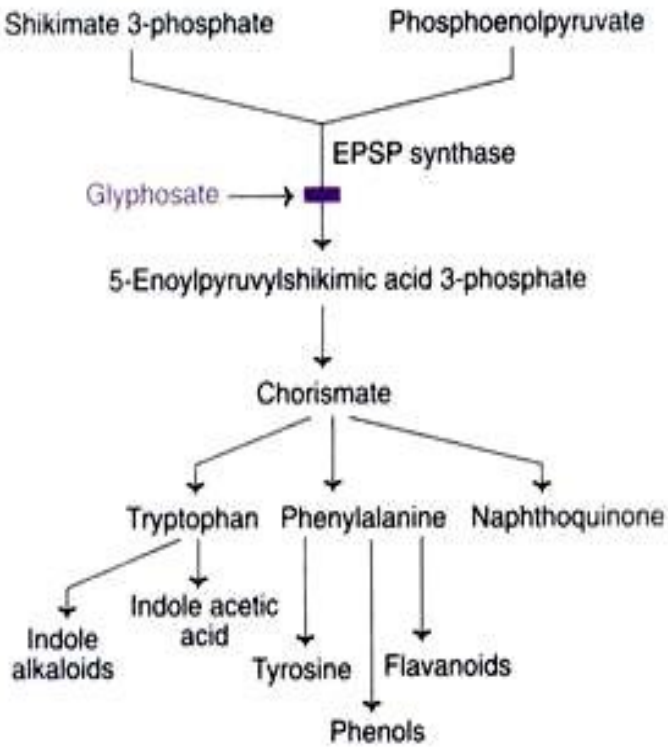


***Fig. 50.5** : Shikimate pathway indicating the action of the herbicide glyphosate (EPSP synthase-5-Enolylpyruvylshikimate 3-phosphate synthase)*

The enzyme EPSPS catalyses the synthesis of 5-enoylpyruvylshikimate 3-phosphate from shikimate 3-phosphate and phosphoenoylpyruvate. Glyphosate has some structural similarly with the substrate phosphoenol pyruvate (Fig 50.6). Consequently, glyphosate binds more tightly with EPSPS and blocks the normal shikimic acid pathway. Thus, the herbicide glyphosate inhibits the biosynthesis of aromatic amino acids and other important products.

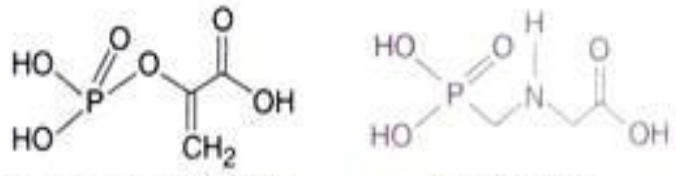


***Fig. 50.6** : Structures of phosphoenolpyruvate (the substrate) and the herbicide glyphosate (the competitive inhibitor).*

This results in inhibition of protein biosynthesis (due to lack of aromatic amino acids). As a consequence, cell division and plant growth are blocked. Further, the plant growth regulator indole acetic acid (an auxin) is also produced from tryptophan. The net result of glyphosate is the death of the plants. Glyphosate is toxic to microorganisms as they also possess shikimate pathway.

Glyphosate is non-toxic to animals (including humans), since they do not possess shikimate pathway. Of the three aromatic amino acids (synthesized in this pathway), tryptophan and phenylalanine are essential and they have to be supplied in the diet, while tyrosine can be formed from phenylalanine.

Strategies for glyphosate resistance:

There are three distinct strategies to provide glyphosphate resistance to plants:

1. Overexpression of crop plant EPSPS gene:

An overexpressing gene of EPSPS was detected in Petunia. This expression was found to be due to gene amplification rather than an increased expression of the gene. EPSPS gene from Petunia was isolated and introduced into other plants. The increased synthesis of EPSPS (by about 40 fold) in transgenic plants provides resistance to glyphosate. These plants can tolerate glyphosate at a dose of 2-4 times higher than that required to kill wild-type plants.

2. Use of mutant EPSPS genes:

An EPSPS mutant gene that conferred resistance to glyphosate was first detected in the bacterium Salmonella typhimurium. It was found that a single base substitution (C to 7) resulted in the change of an amino acid from proline to serine in EPSPS. This modified enzyme cannot bind to glyphosate, and thus provides resistance.

The mutant EPSPS gene was introduced into tobacco plants using Agrobacterium Ti plasmid vectors. The transgene produced high quantities of the enzyme EPSPS. However, the transformed tobacco plants provided only marginal resistance to glyphosate. The reason for this was not immediately identified.

It was later known that the shikimate pathway occurs in the chloroplasts while the glyphosate resistant EPSPS was produced only in the cytoplasm. This enzyme was not transported to the chloroplasts, hence the problem to provide resistance. This episode made scientists to realize the importance of chloroplasts in genetic engineering.

In later years, the mutant EPSPS gene was tagged with a chloroplast-specific transit peptide sequence. By this approach, the glyphosate-resistant EPSPS enzyme was directed to freely enter chloroplast and confer resistance against the herbicide.

3. Detoxification of glyphosate:

The soil microorganisms possess the enzyme glyphosate oxidase that converts glyphosate to glyoxylate and aminomethylphosponic acid. The gene encoding glyphosate oxidase has been isolated from a soil organism Ochrobactrum anthropi. With suitable modifications, this gene was introduced into crop plants e.g. oilseed rape. The transgenic plants were found to exhibit very good glyphosate resistance in the field.

Use of a combined strategy:

More efficient resistance of plants against glyphosate can be provided by employing a combined strategy. Thus, resistant (i.e. mutant) EPSPS gene in combination with glyphosate oxidase gene

are used. By this approach, there occurs glyphosate resistance (due to mutant EPSPS gene) as well as its detoxification (due to glyphosate oxidase gene).

Resistance to other herbicides:

Besides the above, some other herbicide resistant plants have also been developed e.g. bromoxynil, atrazine, phenocarboxylic acids, cyanamide. A list of selected examples of gene transferred herbicide resistant plants is given in Table 50.7.

TABLE 50.7 Selected examples of gene transferred herbicide resistant plants

Herbicide	*Gene transfer/mechanism of resistance*	*Transgenic crop(s)*
Glyphosate	Inhibition of EPSPS	Soybean, tomato
Glyphosate	Detoxification by glyphosate oxidase	Maize, soybean
Phosphinothricin	*bar* gene coding phosphinothricin acetyltransferase	Maize, rice, wheat, cotton, potato, tomato, sugarbeet
Sulfonylureas/imidazolinones	Mutant plant with acetolactate synthase	Rice, tomato, maize, sugarbeet
Bromoxynil	Nitrilase detoxification	Cotton, potato, tomato
Atrazine	Mutant plant with chloroplast *psb A* gene	Soybean
Phenocarboxylic acids	Monooxygenase detoxification (e.g. 2,4-D and 2,4,5-T)	Maize cotton
Cyanamide	Cyanamide hydratase gene	Tobacco

It may however, be noted that some of the herbicide-resistant transgenic plants are at field-trial stage. Due to environmental concern, a few of these plants are withdrawn e.g. atrazine- resistant crops.

Problems with genetically modified plants, safety concerns

Environmental Impact of Herbicide-Resistant Crops:

The development genetically modified (GM) herbicide-resistant crops has undoubtedly contributed to increase in the yield of crops. For this reason, farmers particularly in the developed countries (e.g. USA) have started using these GM crops. Thus, the proportion of herbicide resistant soybean plants grown in USA increased from 17% in 1997 to 68% in 2001.

The farmer is immensely benefited as there is a reduction in the cost of herbicide usage. It is believed that the impact of herbicide-resistant plants on the environment is much lower than the direct use of the herbicides in huge quantities.

There are however, other environmental concerns:

i. Disturbance in biodiversity due to elimination of weeds.

ii. Rapid development of herbicide-resistance weeds that may finally lead to the production of super weeds.

Goals of biotechnological improvements in crops:

There are about 30-40 crops that have been genetically modified, and many more are being added. However, very few of them have got the clearance for commercial use. A selected list is already given in Table 50.9.

The ultimate goals of genetically modified (CM) crop plants are listed below:

i. Resistance to diseases (insect, microorganisms).

ii. Improved nitrogen fixing ability.

iii. Higher yielding capacity.
iv. Resistance to drought and soil salinity.
v. Better nutritional properties.
vi. Improved storage qualities.
vii. Production of pharmaceutically important compounds.
viii. Absence of allergens.
ix. Modified sensory attributes e.g. increased sweetness as in thaumatin.

Concerns about transgenic plants:

The fears about the harmful environmental and hazardous health effects of transgenic plants still exist, despite the fact that there have been no reports so far in this regard. The transfer of almost all the transgenic plants from the laboratory to the crop fields is invariably associated with legal and regulatory hurdles, besides the social and economic concerns.

The major concern expressed by public (also acknowledged by biotechnologists) is the development of resistance genes in insects, generation of super weeds etc. Several remedial measures are advocated to overcome these problems.

The farmers in developing countries are much worried about the seed terminator technology which forces them to buy seeds for every new crop. These farmers are traditionally habituated to use the seeds from the previous crop which is now not possible due to seed terminator technology.

6. Transgenic Plants as Bioreactors:

Another important application of genetically transformed plants is their utility as bioreactors to produce a wide range of metabolic and industrial products.

Applications of Transgenic Plants

The genetic manipulations carried out in plants for the production of transgenic plants have been described .The ultimate goal of transgenic (involving introduction, integration, and expression of foreign genes) is to improve the crops, with the desired traits.

Some of the important ones are listed:

i. Resistance to biotic stresses i.e. resistance to diseases caused by insects, viruses, fungi and bacteria.
ii. Resistance to abiotic stresses-herbicides, temperature (heat, chilling, freezing), drought, salinity, ozone, intense light.
iii. Improvement of crop yield, and quality e.g. storage, longer shelf life of fruits and flowers.
iv. Transgenic plants with improved nutrition.
v. Transgenic plants as bioreactors for the manufacture of commercial products e.g. proteins, vaccines, and biodegradable plastics.

Environmental stresses to plants:

The different types of external stresses that influence the plant growth and development are depicted in Fig. 50.1, These stresses are grouped based on their characters—biotic and abiotic stresses. The biotic stresses are caused by insects, pathogens (viruses, fungi, bacteria), and wounds. The abiotic stresses are due to herbicides, water deficiency (caused by drought, temperature, and salinity), ozone and intense light.

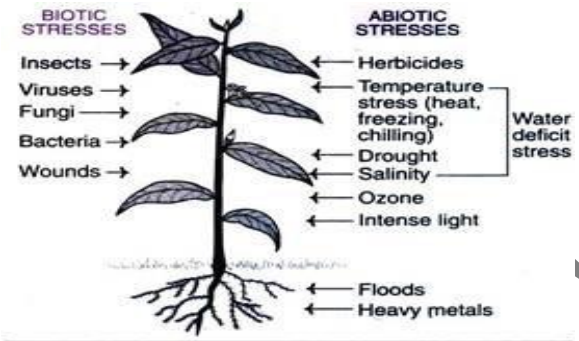


Fig. 50.1 : Biotic and abiotic stresses that affect plant growth, development and yield.

Almost all the stresses, either directly or indirectly, lead to the production of reactive oxygen species (ROS) that create oxidative stress to plants. This damages the cellular constituents of plants which is associated with a reduction in plant yield.

POSSIBLE QUESTIONS

8 MARKS

1. Explain in detail on site directed mutagenesis.
2. Discuss in detail on protein engineering.
3. Explain in detail on applications of genetic engineering in medicine.
4. Discuss on production of recombinant pharmaceuticals.
5. What you know about recombinant vaccines? Explain in detail.
6. Elaborate in detail about gene therapy.
7. Discuss about herbicide resistant crops and its applications.
8. Explain about problems with genetically modified plants. What are the safety measurements to overcome the problems?

.

KARPAGAM ACADEMY OF HIGHER EDUCATION
DEPARTMENT OF BIOCHEMISTRY
III BSc BIOCHEMISTRY
17BCU601A-GENETIC ENGINEERING and BIOTECHNOLOGY
MULTIPLE CHOICE QUESTIONS

Unit-V

S.No	Questions	Option A	Option B	Option C	Option D	Answer
1	Which one is not related with mutation?	Insertion	Regulation	Deletion	Addition	Regulation
2	Oncognes are related with	Ulcer	Diabetes	Cancer	Arthritis	Cancer
3	Which one of the following is not an application of genetic engineering?	Plant gene engineering	Recombinant vaccines	Gene therapy	Reverse omosis	Reverse omosis
4	Asparaginase is used in the treatment of	Diabetes	Leukemia	Blood clotting	Platelet aggregation	Leukemia
5	If the fragment of gene to be mutated lies between two restriction enzyme cleavage site_____ mutagenesis is possible	Oligo nucleotide	Casette	PCR	Direct	Casette
6	Resistance to glyphoshate in transgenic petunia plant has been developed by the transfer of	Gene for EPSPS	Gene for ALS	Gene for GS	Gene for GST	Gene for EPSPS
7	The herbicide bromoxynil is inactivated by the enzyme	ALA synthetase	Nitrilase	EPSP synthetase	PA transferase	Nitrilase
8	*Agrobacterium tumefaciens* infect	Monocotyledons	Dicotyledons	Angiosperms	Gymnosperm	Dicotyledons
9	_____ used in the treatment of myocardial infarction	Lysozyme	Asparaginase	Tissue plasminogen activator	Subtilisin	Tissue plasminogen activator
10	Genes which code for nitrilase confers bromoxynil resistance is	Bar genes	Bxn genes	Cry genes	Trp genes	Bxn genes
11	Tissue plasminogen activator is used in the treatment of	Diabetes	Leukemia	Blood clotting	Lung disease	Blood clotting
12	Bromoxynil, a herbicide acts by inhibiting	Acetolactase synthetase	Photosynthesis	EPSP synthetase	Auxin production	Photosynthesis
13	Genes coding for …… are integrated into plants for preventing fruit ripening and softening during transport	Poly galacturonase	Pectinase	Ethylene	Auxin	Poly galacturonase
14	Somaclonal variations are the ones	Caused by mutagens	Produce during tissue culture	Caused by gamma rays	Induced during sexual embryogeny	Produce during tissue culture
15	Glyophosate acts as herbicide by inhibiting	5-end pyruvyl shikimate 3 phosphate synthetase	Acetolactate synthetase	Photosynthesis	Photothricin acetyl transferase	5-end pyruvyl shikimate 3 phosphate synthetase
16	The gene coding for…… is used to produce phosphinothricin tolerant plants	Phosphinothricin	Phosphinothricin acetyl transferase	Phosphinothricin butyl transferase	Phosphino acetyl transferase	Phosphinothricin acetyl transferase
17	An enzyme used in detergent is	Lysozyme	Asparaginase	Tissue plasminogen activator	Subtilisin	Subtilisin
18	To obtain haploid plant, we culture	Entire anther	Nucleus	Embryo	Apical bud	Entire anther
19	Part of plant used for culturing is called	Stem	Explant	Stock	Callus	Explant
20	Treatment of genetic diseases by introducing proper genes into	Gene therapy	Enzyme therapy	Genetic engineering	Protein engineering	Gene therapy
21	The first transgenic plant was	Tobacco	Pea	Flax	Cotton	Tobacco
22	Callus is	Tissue that forms embryo	An insoluble carbohydrate	Tissue that grows to form embryoid	Un organised actively dividing mass of cells maintained in cultured	Un organised actively dividing mass of cells maintained in cultured
23	Who is the father of tissue culture?	Bonner	Haberlandt	Laibach	Gautheret	Haberlandt
24	The first successful transformation of rDNA molecule into a ba	Nathan, Arber and Smith	Watson, Crick and Wilkins	Boyer and Cohen	Paul Berg	Boyer and Cohen
25	DMSO (Dimethyl sulfoxide) is used as	Gelling agent	Alkaylating agent	Chelating agent	Cryoprotectant	Cryoprotectant
26	Growth hormone producing apical dominance is	Auxin	Gibberellin	Ethylene	Cytokinin	Auxin
27	To develop herbicide resistance in transgenic plants which of the following approach is used	Target molecule is made insensitive to herbicide	Target protein is over produced	A pathway introduce to detoxify the herbicide	a,b and c	a,b and c
28	Cybrids are produced by	Fusion of two different nuclei from two different species	Fusion of two same nuclei from same species	Nucleus of one species but cytoplasm from both the parent species	Nucleus of one species and cytoplasm from single species	Nucleus of one species but cytoplasm from both the parent species
29	Which vector is mostly used in crop improvement?	Plasmid	Cosmid	Phagemid	Agrobacterium	Agrobacterium
30	Substances released by plants in response to wounding and can induce vir genes of Ti plasmid is	Opines	Acetosyringone	Nopaline	Auxin	Acetosyringone
31	A medium which is composed of chemically defined compound is called	Natural media	Synthetic media	Artificial media	Serum free media	Synthetic media
32	Antisence technology	Selectively block expression of gene	Combine genetic material from different species	Combine organelle in a cell	Transfer a cell	Selectively block expression of gene
33	The production of secondary metabolites require the use of	Protoplast	Cell suspension	Meristem	Auxillary buds	Cell suspension
34						
35	Transfer of cloned genes into cells grown in culture and then in	*In vivo* gene therapy	Embryo therapy	Somatic cell gene therapy	*Ex vivo* therapy	*Ex vivo* therapy
36	Amino acid coding changes in DNA can be achieved by	PCR	Site directed mutagenesis	FISH	FACS	Site directed mutagenesis
37	All are applications of site directed mutagenesis except	Increase the substrate affinity of an enzyme	Making the enzyme thermo tolerant	Improving the stability of an enzyme	Decreasing the activity of an enzyme	Decreasing the activity of an enzyme
38	Which tropical fruit crop has been successfully engineered to be protected against lethal virus?	Papaya	Passion fruit	Mango	Lemon	Papaya
39	Hormone pair required for a callus to differentiate are	Auxin and cytokinin	Auxin and ethylene	Auxin and absiccic acid	Cytokinins and gibberllin	Auxin and cytokinin
40	Synthetic seed is produced by encapsulating somatic embryo with	Sodium chloride	Sodium alginate	Sodium acetate	Sodium nitrate	Sodium alginate
41	Agrobacterium tumefaciens infect_____ family	Roseaceae	Solanaceae	Liliaceae	Malvaeceae	Roseaceae
42	Hairy root cultures for secondary metabolite production are induced by transforming plant cells with	Virus	*Agrobacterium tumefaciens*	*Bacillus thuringiensis*	*Agrobacterium rhizogenes*	*Agrobacterium rhizogenes*
43	Which of the following statements are true regarding rDNA tec	Obtain large number copie	Physics technology	Computer technology	Chemical technology	Obtain large number copies of specific DNA fragments
44	Ti plasmid is classified based on	Opine synthesis	Auxin synthesis	Cytokinin synthesis	Acetosyringone synthesis	Opine synthesis
45	The genes which produces toxic crystals in *Bacillus thuringenesis* is	Vir genes	Cry genes	Nif genes	Tra genes	Cry genes
46	The tonic crystals in *B.thuringenesis* kills the insect by	Causing pores in epithelial cells	Auxin synthesis	By inhibition of metabolic enzymes	By inhibiting production of eggs	Causing pores in epithelial cells
47	To obtain haploid plant, we culture	Entire anther	Nucleus	Embryo	Apical bud	Entire anther
48	Artificial seeds are	Seeds produced in laboratory condition	Seeds encapsulated in a gel	Somatic embryos encapsulated in a gel	Zygotic embryos encapsulated in a gel	Somatic embryos encapsulated in a gel
49	The variation in *in vitro* culture is called as	*In vitro* variation	Mutation	Somoclonal variation	Protoplast fusion	Somoclonal variation
50	Genes of Ti plasmid that help in auxin biosynthesis is	vir D	vir E	vir B	vir G	vir D
51	Elicitors are molecules that	Induce cell divison	Stimulate production secondary metabolites	Stimulate hairy root formation that accumulate secondary metabolites	Stimulate senescence	Stimulate production secondary metabolites
52	Genes of Ti plasmid that help in auxin biosynthesis is	iaaH	iaaM	ipt	iaaH and iaaM	iaaH and iaaM
53	Somatic cell gene therapy involves	Introduction of remedial g	Introduction of remedial gen	Introduction of remedial g	Introduction of remedial gene into	Introduction of remedial gene into blood cells
54	Fibroblast are grown well on	Stainless stell	Palladium	Iron	Copper	Palladium
55	In transgenic mice production the gene is introduced into___	Sperm	Egg	Fertilised egg	Ovum	Fertilised egg
56	First cloned animal	Dolly sheep	Dog	Mule	Cat	Dolly sheep
57	The chromosome used for embryo sexing is	X-chromosome	Y-chromosome	XY chromosome	Polytene chromosome	Y-chromosome
58	In ES method of transgenic mice production the gene in introduced through	Microinjection	Electroporation	a & b	Lipofection	a & b
59	First mammalian clone was developed by	Maxim and Gilber, 1971	Wilmut and Camphel, 1997	Salk and Sabin, 1968	Jacob and Mond, 1973	Wilmut and Camphel, 1997
60	First synthetic vaccine is against	Human simplex virus	Hepatitis B	Tuberculosis	Measles	Hepatitis B