Gene manipulation and recombinant DNA technology

The majority of approved biopharmaceuticals are proteins produced in engineered cells by recombinant means. Examples include the production of insulin in recombinant *E. coli*.

Recombinant DNA technology, rDNA technology (Gene manipulation, gene cloning and genetic engineering), describes the process of manipulating genes. It generally involves the isolation, manipulation and subsequent reintroduction of stretches of DNA into cells and is usually undertaken in order to confer on the recipient cell the ability to produce a specific protein, such as a biopharmaceutical.

Recombinant DNA, rDNA, is a piece of DNA artificially created *in vitro* which contains DNA (natural or synthetic) obtained from two or more sources.

Recombinant production of therapeutic protein

Production of any protein via rDNA technology entails the initial identification and isolation of a DNA sequence coding for the target protein. This sequence can be direct genomic DNA, but mRNA coding for the protein of interest can also act as a starting point. Total eukaryotic cellular mRNA can be purified from the cell via an affinity-based mechanism. The mRNA is enzymatically 'reverse transcribed' into cDNA. If the target therapeutic protein is eukaryotic then the genomic DNA will

contain both coding (exon) and non-coding (intron) sequences, whereas the cDNA will be a reflection of the exons only.

The desired gene/cDNA is normally amplified, sequenced and then introduced into an expression vector that facilitates its introduction and expression (transcription and translation) in an appropriate producer cell type. All recombinant therapeutic proteins approved to date are produced in *E. coli*, *S. cerevisiae* or in animal cell lines.

Recombinant DNA technology has had a fourfold positive impact upon the production of pharmaceutically important proteins:

- **It overcomes the problem of source availability**. Many proteins of therapeutic potential are produced naturally in the body in minute quantities. Examples include interferons, interleukins and colony-stimulating factors (CSFs). This rendered impractical their direct extraction from native source material in quantities sufficient to meet likely clinical demand. Recombinant production allows the manufacture of any protein in whatever quantity it is required.

- **It overcomes problems of product safety.** Direct extraction of product from some native biological sources has, in the past, led to the unwitting transmission of disease. Examples include the transmission of blood-borne pathogens such as hepatitis B and C and human immunodeficiency virus (HIV) via infected blood

products and the transmission of Creutzfeldt–Jakob disease to persons receiving human growth hormone (GH) preparations derived from human pituitaries.

❖ It provides an alternative to direct extraction from inappropriate/dangerous source material.

A number of therapeutic proteins have traditionally been extracted from human urine. Folliclestimulating hormone (**FSH**), the fertility hormone, for example, is obtained from the urine of postmenopausal women, and a related hormone, human chorionic gonadotrophin (**hCG**), is extracted from the urine of pregnant women. Urine is not considered a particularly desirable source of pharmaceutical products. Although several products obtained from this source remain on the market, recombinant forms have now also been approved. Other potential biopharmaceuticals are produced naturally in downright dangerous sources. **Ancrod,** for example, is a protein displaying anti-coagulant activity and, hence, is of potential clinical use. It is, however, produced naturally by the Malaysian pit viper. Although retrieval by milking snake venom is possible, and indeed may be quite an exciting procedure, recombinant production in less dangerous organisms, such as *Escherichia coli* or *Saccharomycese cerevisiae*, would be considered preferable by most.

❖ **It facilitates the generation of engineered therapeutic proteins displaying some clinical advantage**

Over the native protein product. Techniques such as **site-directed mutagenesis** facilitate the logical introduction of predefined changes in a protein's amino acid sequence. Such changes can be as minimal as the insertion, deletion or alteration of a single amino acid residue, or can be more substantial (e.g. the alteration/deletion of an entire domain, or the generation of a novel hybrid protein). Such changes can be made for a number of reasons, and several engineered products have now gained marketing approval.

Despite the undoubted advantages of recombinant production, it remains the case that many protein-based products extracted directly from native source material remain on the market. In certain circumstances, direct extraction of native source material can prove equally/more attractive than recombinant production. This may be for an economic reason if, for example, the protein is produced in very large quantities by the native source and is easy to extract/purify, e.g. human serum albumin (HSA). Also, some blood factor preparations purified from donor blood actually contain several different blood factors and, hence, can be used to treat several haemophilia patient types. Recombinant blood factor preparations, on the other hand, contain but a single blood factor and, hence, can be used to treat only one haemophilia type.

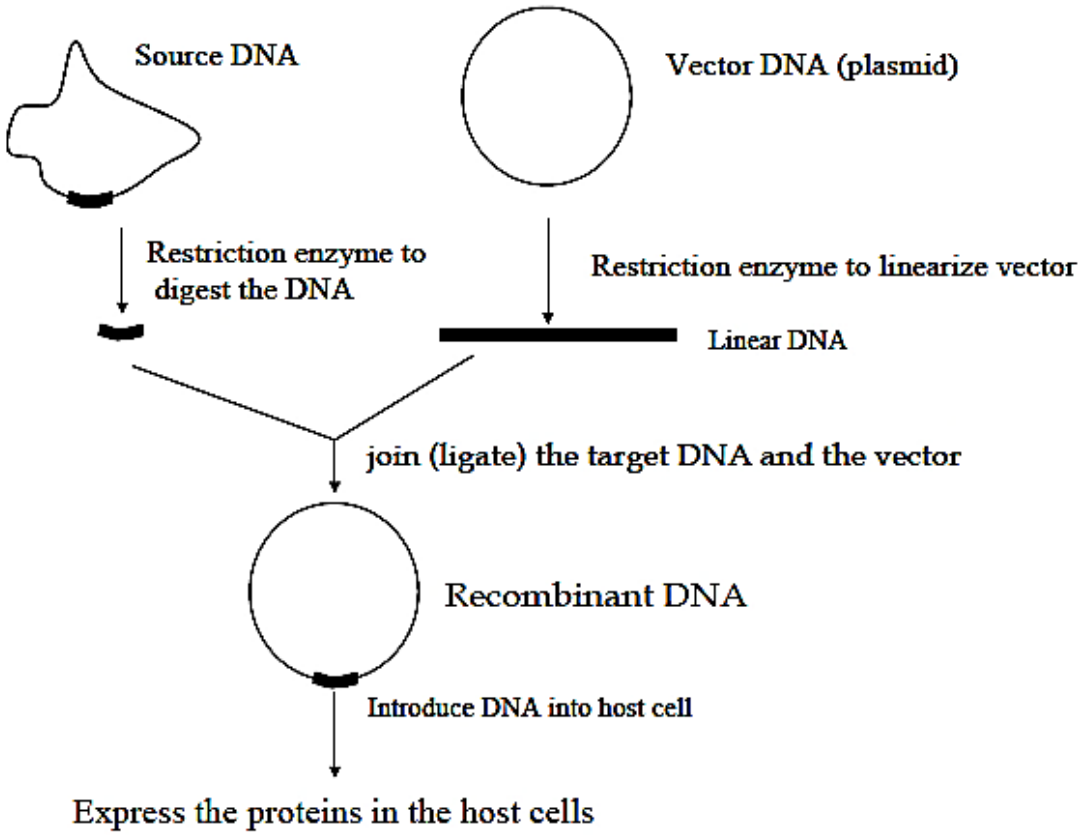
Recombinant DNA technology steps

Recombinant DNA technology relates to the usage of three main tools:

enzymes (restriction enzymes, polymerases, and ligases); (2) vectors;

and (3) host organism.

In generally, a recombinant DNA technology has the following steps: (1) isolating DNA (to be used as a template to amplify the gene). (2) Amplifying the gene copies by PCR (3) cutting the desired DNA by restriction sites, (4) inserting the genes into the vectors, (5) transferring the vectors into host organism, and (6) obtaining the products of recombinant genes.



Polymerase chain reaction

Polymerase chain reaction allows obtaining large amounts of the gene of interest. The process begins by extraction of total genomic DNA from the source of interest. Oligonucleotide primers whose sequences flank the target gene/DNA segment are synthesized and used to amplify that portion of DNA selectively. Recognition sites for REs can be incorporated into the oligonucleotides to allow cloning of the amplified gene.

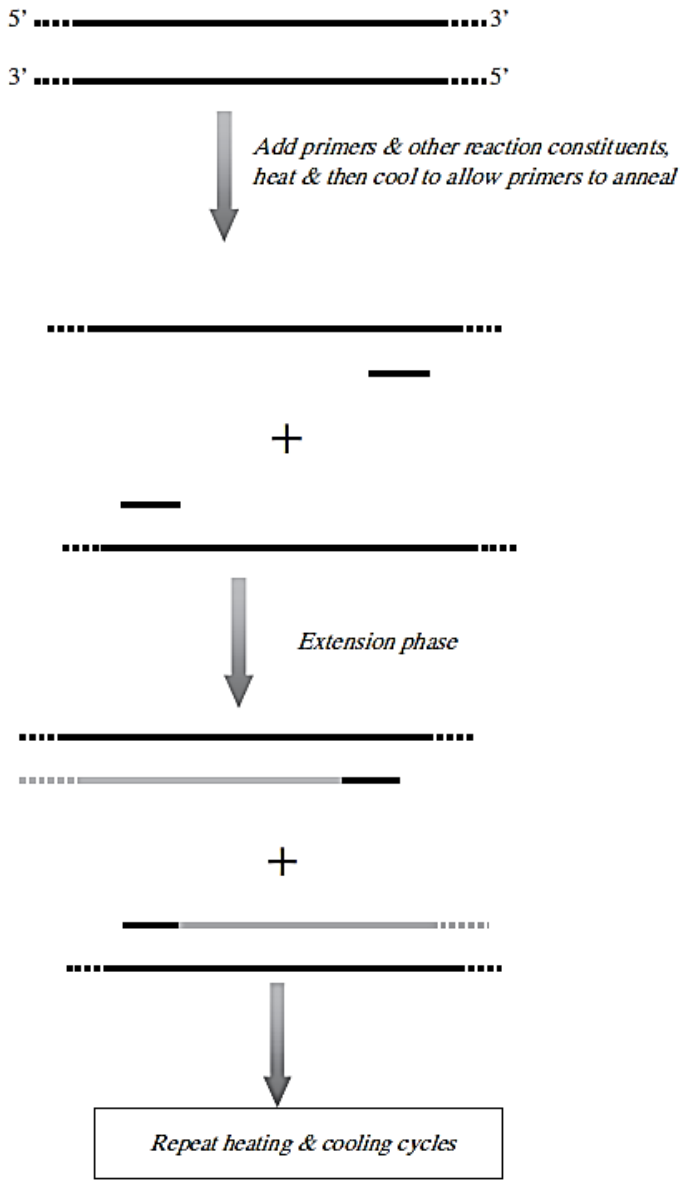


Figure The PCR is initiated by separation of the double-stranded DNA into its two constituent strands. This is achieved by heating the sample (usually to 94 °C). Also present in the reaction mixture are: (a) two chemically synthesized oligonucleotide primers ('oligos') whose sequences are complementary to the sequences flanking the gene of interest; (b) the enzyme DNA polymerase, which can extend the primers to synthesize a new DNA strand of complementary sequence to the single stranded DNA template; (c) all the nucleoside precursors required for synthesis of the growing DNA strand (i.e. the deoxynucleoside triphosphates or dNTPs). Once strand separation has been achieved the reaction temperature is reduced, in order to allow primers to anneal to complementary sequences on each strand and allow the DNA polymerase to extend the primers. This extension phase is normally carried out at 74 °C. The DNA polymerase used is sourced from the thermophilic microorganism *Thermus aquaticus*; therefore, it is heat stable and not inactivated by PCR operational temperatures. This completes the first cycle of the PCR process; the result is a doubling of the amount of target DNA present in the reaction mixture. The cycle is then repeated and with each repeat comes a doubling of the amount of target DNA. After 25–30 cycle repeats, several hundred million copies of the target DNA have been generated

Cutting DNA The breakthrough that made recombinant DNA technology possible was the discovery and characterization of **restriction enzymes.** Restriction enzymes are produced by bacteria as a defense mechanism against phages. The enzymes act like scissors, cutting up the DNA of the phage and thereby inactivating it. Importantly, restriction enzymes do not cut randomly; rather, they cut at specific DNA target sequences, which is one of the key features that make them suitable for DNA manipulation.

Depending upon the specific RE utilized, DNA cleavage may yield blunt ends (e.g. BsaAI and EcoRV) or staggered ends – the latter are often referred to as sticky ends.

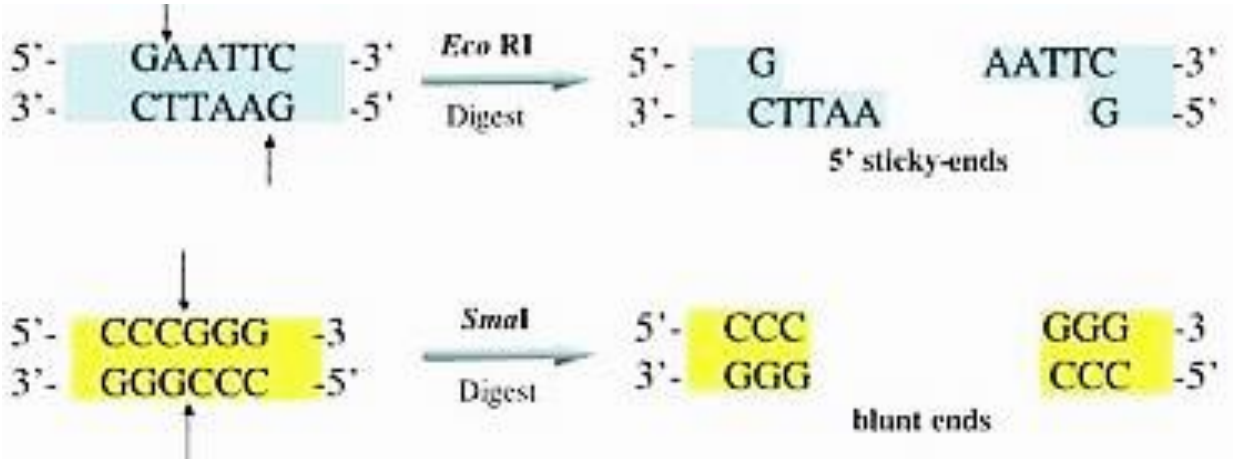


Figure: Sticky and blunts ends

Table Some commercially available REs, their sources, DNA recognition sites and cleavage points

Restriction enzyme	Source	DNA recognition sequence and cleavage site[a]
BclI	*Bacillus caldolyticus*	5′-T↓GATCA-3′ 3′-ACTAG↑T-5′
BglII	Recombinant *E. coli* carrying *BglII* gene from *Bacillus globigii*	5′-A↓GATCT-3′ 3′-TCTAG↑A-5′
BsaAI	Recombinant *E. coli* carrying *BsaAI* gene from *Bacillus stearothermophilus* A	5′-PyAC↓GTPu-3′ 3′-PuTG↑CAPy-5′
BsaJI	*B. stearothermophilus* J	5′-C↓CNNGG-3′ 3′-GGNNC↑C-5′
BsiEI	*B. stearothermophilus*	5′-CGPuPy↓CG-3′ 3′-GC↑PyPuGC-5′
EcoRV	Recombinant *E. coli* carrying *EcoRV* gene from the plasmid J62 pIg 74	5′-GAT↓ATC-3′ 3′-CTA↑TAG-5′
MwoI	Recombinant *E. coli* carrying cloned *MwoI* gene from *Methanobacterium wolfeii*	5′-GCNNNNN↓NNGC-3′ 3′-CGNN↑NNNNNCG-5′
Tsp509I	*Thermus* sp.	5′-↓AATT-3′ 3′-TTAA↑-5′
XbaI	Recombinant *E. coli* carrying *XbaI* gene from *Xanthomonas badvii*	5′-T↓CTAGA-3′ 3′-AGATC↑T-5′
XhoI	Recombinant *E. coli* carrying *XhoI* gene from *X. holcicola*	5′-C↓TCGAG-3′ 3′-GAGCT↑C-5′

[a]G: guanine; C: cytosine; A: adenine; T: thymine; Pu: any purine; Py: any pyrimidine; N: either a purine or pyrimidine. Arrow indicates site of cleavage.

3- Inserting the genes into the vectors.

Once the gene/cDNA coding for a potential target protein has been isolated, the goal usually becomes one of achieving high levels of expression of this target gene. This process entails ligation of the gene into a vector that will support high-level transcription and translation.

Essential features of cloning vectors

1. Origin of replication

2. Multiple unique restriction sites into which the foreign DNA can be inserted (multiple cloning site, MCS).

3. Selectable markers for identifying cells harboring the cloning vector-insert DNA construct, and whether the foreign DNA has been inserted.

- ex: pBR322
- One of the best-studied and often used
- Two antibiotics resistance genes: Ampicillin and tetracycline.

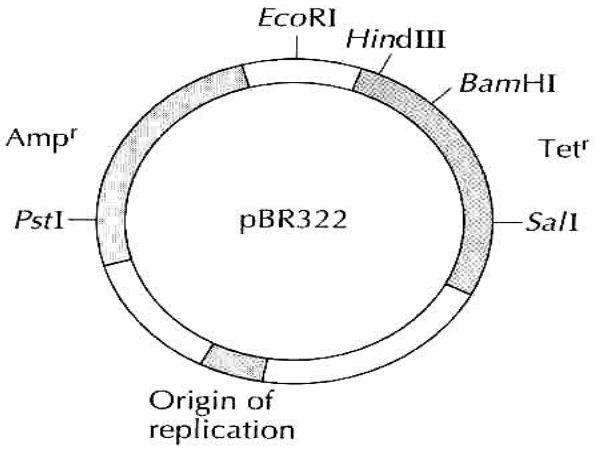


Figure Genetic map of the plasmid cloning vector pBR322. Unique *Hin*dIII, *Sal*I, *Bam*HI, and *Pst*I recognition sites are present the genes for tetracycline resistance (Tet^r) and ampicillin resistance (Amp^r). The unique *Eco*RI site is just outside the tetracycline resistance gene. The origin of replication functions in the bacterium *E. coli*. The complete DNA sequence of pBR322 consists of 4,361 bp.

- **Ligation**

Integration of the DNA fragments into the chosen vector is undertaken by ‘opening up’ the circular vector via treatment with the same **RE** as used to generate the DNA fragments for cloning, followed by co-incubation of the cleaved vector and the fragments under conditions that promote the annealing of complementary sticky ends. The plasmids are then incubated with another enzyme, a **DNA ligase** (mainly from bacteriophage T4), which catalyses the formation of phosphodiester bonds in the DNA backbone and thus will seal or ‘ligate’ the plasmid.

Note:

The plasmid DNA can self-ligate after restriction enzyme digestion. To minimize the self-ligation the cleaved plasmid can be treated with alkaline phosphatase to remove the phosphate group.

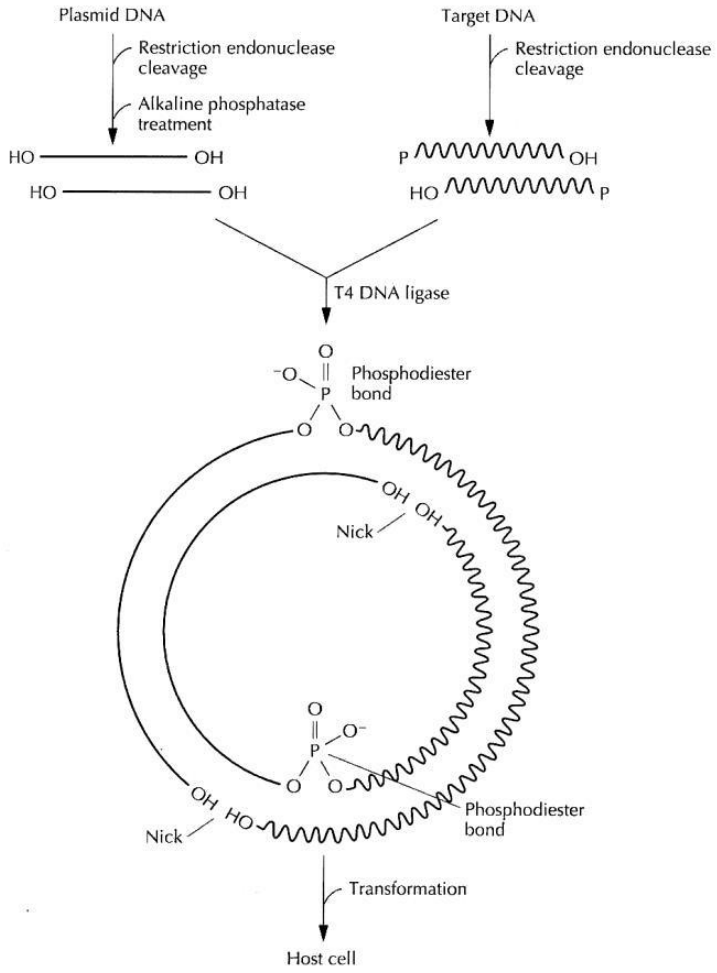


Figure : Ligation of the digested DNA and the digested plasmid. The two phosphodiester bonds are formed by T4 ligase and able to hold both molecules together despite the nicks.

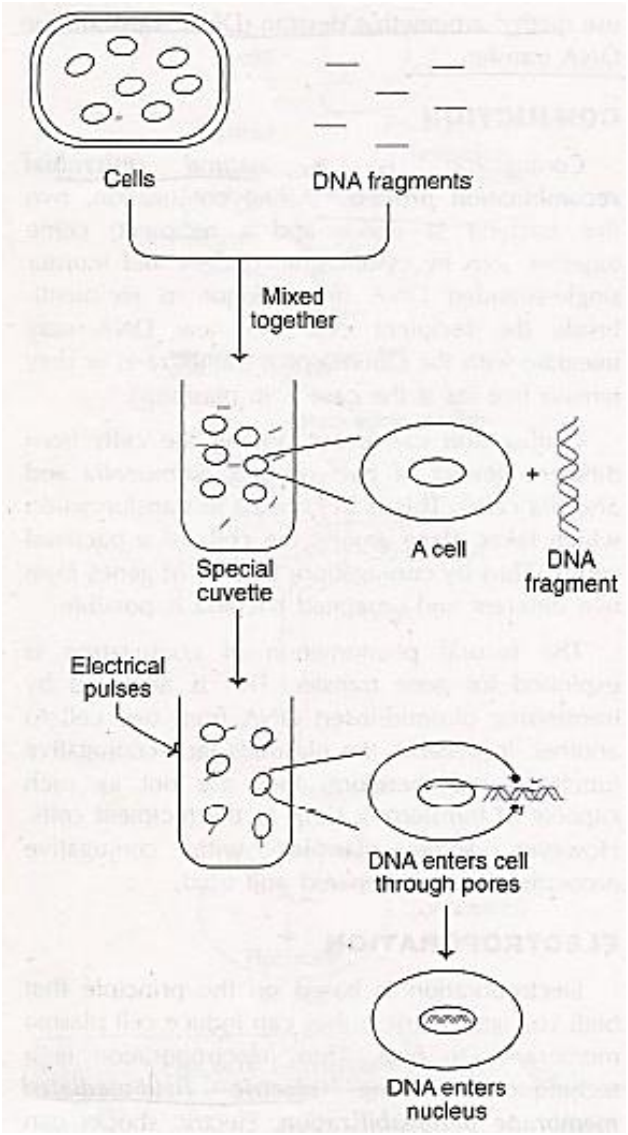
3- Transferring the vectors into host organism.

- **Transformation:**

Heat shock: Transferring genes into prokaryotic cells. Expose bacterial cells to CaCl2 to make cells competent (able to take up exogenous DNA). Mix the cells with the recombinant DNA and apply a heat shock (increase temp to 42º C). The membrane can transiently open to uptake DNA.

Another ways to transfer bacterial DNA is Conjugation and Transduction.

- **Electroporation:** is the process of applying electrical field to a living cell for a brief duration of time in order to create microscopic pores in the plasma membrane called electro-pores. This technique is used for transferring the recombinant DNA molecule into wide range of hosts starting from **bacteria** to **plant** (plant protoplasts) and **animal cells**.



4-Growing of the transformed bacterial cells on agar plates

5-Screening/ identification of the host cell colonies containing the rDNA.

The *E. coli* cells are next spread out on the surface of an agar plate (containing an appropriate antibiotic) and incubated under appropriate conditions.

Three main types of cell will be initially transferred onto these agar plates:

(a) Some cells will have failed to take up any plasmid;

(b) Some transformed cells may have a plasmid in which no foreign DNA had been inserted;

(c) Some cells will house a plasmid that does carry a fragment of the target DNA: These latter cells are the only ones of interest, and **various methods to verify the gene of interest was successfully cloned:**

1. Positive selection vector

An effective method to simplify screening is to use a positive selection vector. Positive selection vectors conditionally express a lethal gene, such as a restriction enzyme that digests the genomic DNA of the bacterial host. The expression of the lethal gene is disrupted by ligation of a DNA insert into the cloning site. As a result, only cells with recombinant plasmids are able to grow.

2. Diagnostic restriction digest

Restriction enzyme digests can also be performed to determine the correct insert. First, isolate plasmid DNA from an overnight bacterial culture using a plasmid miniprep kit. Then, the purified plasmid DNA from recombinant clones is digested using restriction enzymes. The digested plasmid is run on an agarose gel to verify that the vector backbone and insert are of the expected sizes.

3-Colony PCR

Is a method for determining the presence or absence of insert DNA in plasmid constructs. Individual transformants can either be lysed in water with a short heating step or added directly to the PCR reaction and lysed during the initial heating step. This initial heating step causes the release of the plasmid DNA from the cell, so it can serve as template for the amplification reaction. Primers designed to specifically target the insert DNA can be used to determine if the construct contains the DNA fragment of interest.

4. Sequencing

The most accurate way to verify the recombinant colonies is by sequencing. Plasmid DNA is first isolated from an overnight bacterial culture. The insert can be identified using sequencing primers appropriate for the vector. Sequencing across the entire insert is required to verify the exact sequence of insert.

5- Blue-white screening

The most common reporter gene used in plasmid vectors is the bacterial *lacZ* gene encoding beta-galactosidase, an enzyme that naturally degrades lactose but can also degrade a colorless synthetic analog **X-gal**, thereby producing blue colonies on X-gal–containing media. The *lacZ* reporter gene is disabled when the recombinant

DNA is spliced into the plasmid. Because the LacZ protein is not produced when the gene is disabled, X-gal is not degraded and white colonies are produced, which can then be isolated.

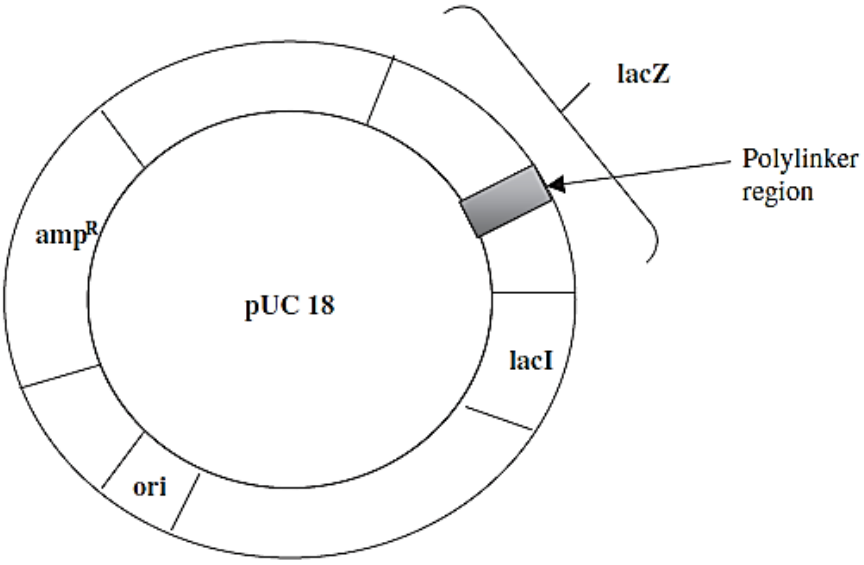


Figure The plasmid pUC18 is often used for cloning purposes. It contains three genes: the ampicillin resistance gene (*amp*^{R}), the *lacZ* gene, which codes for the enzyme β-galactosidase, and the *lacI* gene, which codes for a factor that controls the transcription of *lacZ*. Also present is an origin of replication (ori), essential for plasmid replication within the cell. Note the presence of a short stretch of DNA called the polylinker region located within the *lacZ* gene. The polylinker (also called a multiple cloning site) contains cleavage sites for 13 different REs. This allows genetic engineers great flexibility to insert a DNA fragment for cloning into this area. The polylinker has been designed and positioned within the *lacZ* gene so as not to prevent the expression of functional β-galactosidase. However, if a piece of DNA for cloning is introduced into the polylinker region, then the increased length does block β-galactosidase expression.

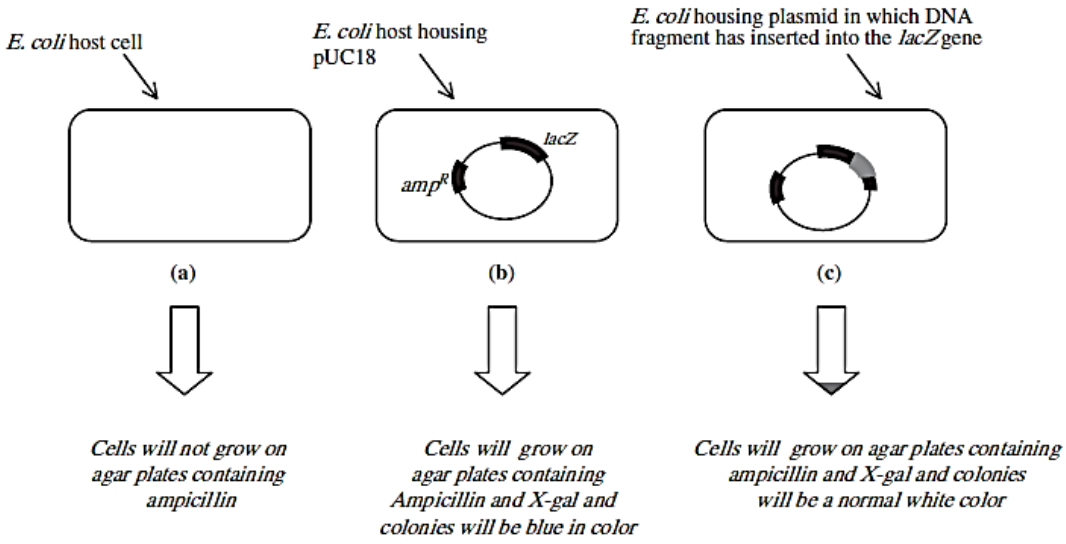
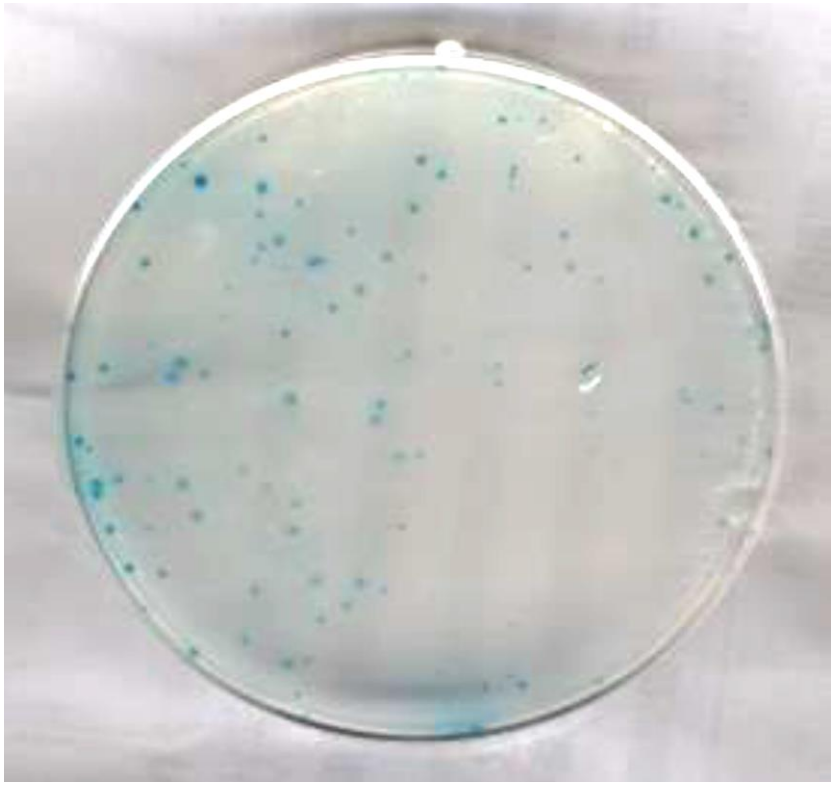


Figure . dentification of *E. coli* host cell clones containing rDNA using pUC18 vectors. After transformation the cells are spread on agar plates containing ampicillin and a chemical called X-gal. Any untransformed cells present will fail to grow on these plates as the host *E. coli* cells contain no ampicillin resistance gene. This step, therefore, identifies cells into which plasmid has successfully been transferred (cell types (b) and (c)). Cells containing plasmid into which no foreign DNA has been inserted (cell type (b)) will grow on ampicillin-containing plates as the plasmid contains the amp^R gene. This cell type also produces functional β-galactosidase . This enzyme will cleave X-gal, liberating a product that is blue in colour.
Colonies, therefore, appear blue. Cells containing a plasmid in which a DNA fragment has been inserted into the *lacZ* gene (cell type (c), i.e. the desired cells) do not produce β-galactosidase; therefore, colonies derived from these cells will be a normal white colour

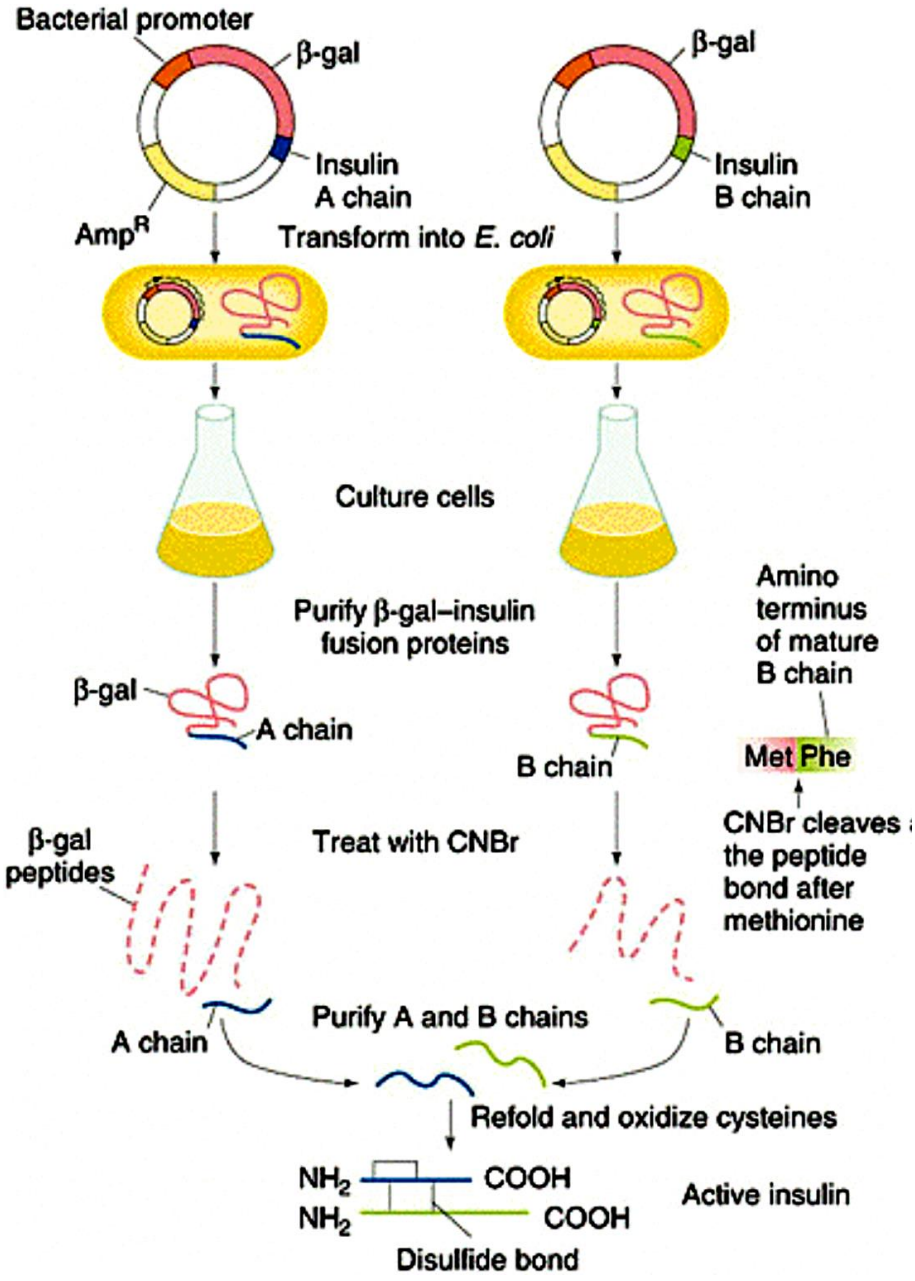


Applications of Recombinant DNA Technology

TABLE Some Recombinant Proteins Were Produced and Commercialized in Medicine

Brand	Generic	Company	Therapeutic category	Indications
THE FIRST GENERATION OF THERAPEUTIC PROTEINS				
Humulin	Insulin	Eli Lilly	Diabetes	Diabetes
Hematrope	Recombinant somatropin	Eli Lilly	Hormones	Growth failure
Genotropin	Somatropin	Pfizer	Hormones	Growth failure
Saizen	Somatropin	Serono	Hormones	Growth failure
Nutropin/ Protropin	Somatropin/ Somatrem	Genetech	Hormones	Growth failure
Intron A	Interferon alpha 2b	Schering-Plough	Anti-infective	Viral infections
Avonex	Interferon beta-1a	Biogen Idec	Multiple sclerosis	Chronic inflammatory demyelinating polyneurophathy
Betaseron/ Betaferon	Interferon beta-1b	Schering AG	Multiple sclerosis	Multiple sclerosis
Procrit/Eprex	Epoetin alpha	J&J	Blood modifier	Anemia
Epogen	Epoetin alpha	Amgen	Blood modifier	Anemia
NeoRecormon	Epoetin beta	Roche	Blood modifier	Anemia
Kogenate	Factor VIII	Bayer	Blood modifier	Hemophilia
NovoSeven	Factor VIIa	Novo Nordisk	Blood modifier	Hemophilia
Benefix	Factor IX	Wyeth	Blood modifier	Hemophilia
Fabrazyme	Agalsidase beta	Genzyme	Enzymes	Fabry disease
Replagal	Agalsidase alfa	TKT Europe	Enzymes	Fabry disease

Insulin – 1st recombinant product to be licensed for therapeutic use in 1982



Transfection: This term is used to for the transfer of DNA into Eukaryotic cells.

- **Liposome mediated gene transfer (Lipofection)**: Encapsulated DNA in liposomes. This technique is found very successful in the transfection of plant protoplasts and animal host cells.

Liposomes are microscopic vesicles developed in a laboratory environment. Each liposome is a spherical ball like structure made up of phospholipid bilayers with a hollow central space, allowing liposomes to interact directly with cells.

A liposome can fuse with the cell membrane of the taken host cell and can deliver its content to it. The recombinant DNA enclosed in the liposome vesicles penetrates into the protoplast of the host cell. In this technique the recombinant DNA, which is negatively charged at a near neutral pH because of its phosphodiester backbone, is mixed with the lipid molecules with positively charged (cationic) head groups. The lipid molecules form a bilayer around the recombinant DNA molecules. This results in the formation of liposomes which are further mixed with the host cells. Most eukaryotic cells are negatively charged at their surface, so the positively charged liposomes interact with the cells. Cells take up the lipid-recombinant DNA complexes, and some of the transfected DNA enters the nucleus.

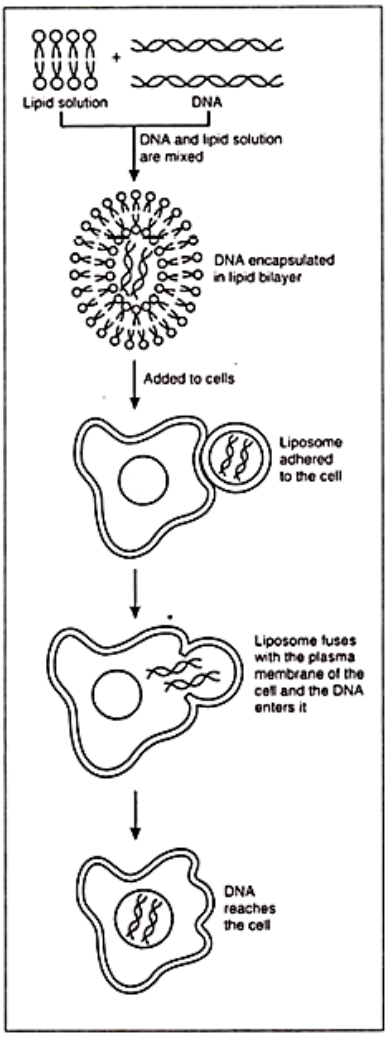
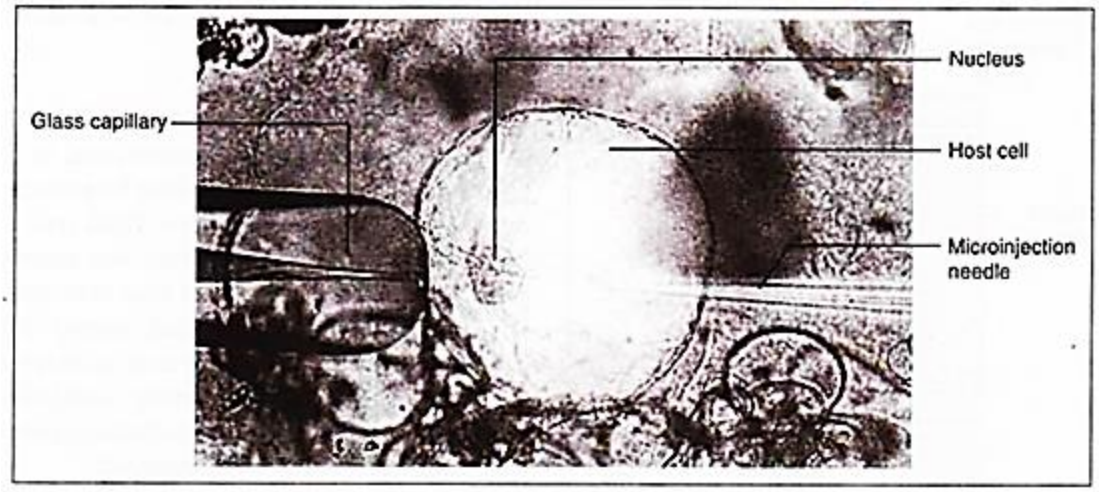


Figure: Liposome mediated gene transfer (Lipofection)

Microinjection: used to introduce genes into multicellular animal (transgenic animals). An extremely fine pipette is used to directly inject DNA into the nucleus of animal cells (e.g. fertilized egg) so DNA is integrated into the chromosome. The transfected egg is then implanted into an animal.



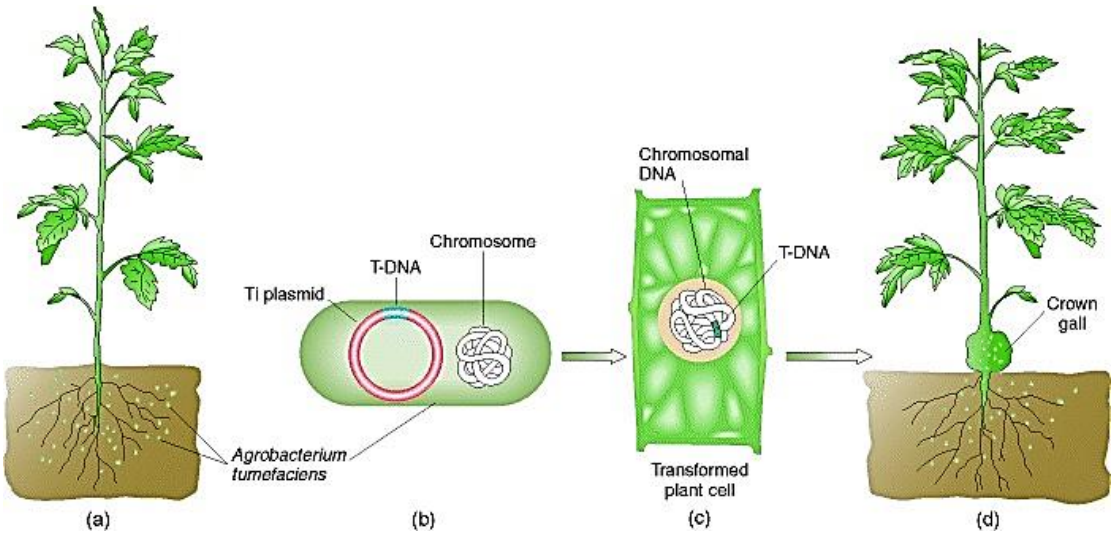
Genetic Engineering in Plants

Because of their economic significance, plants have long been the subject of genetic analysis aimed at developing improved varieties. Recombinant DNA technology has introduced a new dimension to this effort because the genome modifications made possible by this technology are almost limitless. No longer is breeding confined to selecting variants within a given species. DNA can now be introduced from other species of plants, animals, or even bacteria.

The Ti plasmid

The only vectors routinely used to produce transgenic plants are derived from a soil bacterium called *Agrobacterium tumefaciens.* This bacterium causes what is known as *crown gall disease,* in which the infected plant produces uncontrolled growths (tumors, or galls), normally at the base (crown) of the plant. The key to tumor production is a large (200-kb) circular DNA plasmid—the Ti (tumor-inducing) plasmid. When the bacterium infects a plant cell, a part of the Ti plasmid—a region

called *T-DNA*—is transferred and inserted, apparently more or less at random, into the genome of the host plant. The functions required for this transfer are outside the T-DNA on the Ti plasmid. The T-DNA itself carries several interesting functions, including the production of the tumor and the synthesis of compounds called *opines.* Opines are actually synthesized by the host plant under the direction of the T-DNA. The bacterium then uses the opines for its own purposes, calling on opine-utilizing genes on the Ti plasmid.



The natural behavior of the Ti plasmid makes it well-suited for the role of a plant vector. If the DNA of interest could be spliced into the T-DNA, then the whole

package would be inserted in a stable state into a plant chromosome. This system has indeed been made to work essentially in this way, but with some necessary modifications.