STUDY MATERIAL FOR

B.G 5th Semester

SUBJECT: BIOTECHNOLOGY

GENETIC ENGINEERING

GENETIC ENGINEERING

Genetic engineering primarily involves the manipulation of genetic material (DNA) to achieve the desired goal in a pre-determined way. Some other terms are also in common use to describe genetic engineering.

- Gene manipulation
- Recombinant DNA (rDNA) technology
- Gene cloning (molecular cloning)
- Genetic modifications
- New genetics

A *clone* is an identical copy. This term originally applied to cells of a single type, isolated and allowed to reproduce to create a population of identical cells. **DNA cloning** involves separating a specific gene or DNA segment from a larger chromosome, attaching it to a small molecule of carrier DNA, and then replicating this modified DNA thousands or millions of times through both an increase in cell number and the creation of multiple copies of the cloned DNA in each cell. The result is selective amplification of a particular gene or DNA segment.

MOLECULAR CLONING

The basic procedure of molecular cloning involves a series of steps.

1. *Cutting DNA at precise locations.* Sequence-specific endonucleases (restriction endonucleases) provide the necessary molecular scissors.
2. *Selecting a small molecule of DNA capable of self-replication.* These DNAs are called **cloning vectors** (a vector is a delivery agent). They are typically plasmids or viral DNAs.
3. *Joining two DNA fragments covalently.* The enzyme DNA ligase links the cloning vector and DNA to be cloned. Composite DNA molecules comprising covalently linked segments from two or more sources are called **recombinant DNAs.**
4. *Moving recombinant DNA from the test tube to a host cell* that will provide the enzymatic machinery for DNA replication.
5. *Selecting or identifying host cells that contain recombinant DNA.*

The methods used to accomplish these and related tasks are collectively referred to as **recombinant DNA technology** or, more informally, **genetic engineering.**

CLONING VECTORS

Cloning vectors are carrier DNA molecules. Four important features of all cloning vectors are that they: (i) can independently replicate themselves and the foreign DNA segments they carry; (ii) contain a number of unique restriction endonuclease cleavage sites that are present only once in the vector; (iii) carry a selectable marker (usually in the form of antibiotic resistance genes or genes for enzymes missing in the host cell) to distinguish host cells that carry vectors from host cells that do not contain a vector; and (iv) are relatively easy to recover from the host cell. There

are many possible choices of vector depending on the purpose of cloning. The greatest variety of cloning vectors has been developed for use in the bacterial host *E. coli*. Thus, the first practical skill generally required by a molecular biologist is the ability to grow pure cultures of bacteria.

Choice of vector is dependent on insert size and application

The classic cloning vectors are plasmids, phages, and cosmids, which are limited to the size insert they can accommodate, taking up to 10, 20, and 45 kb, respectively (Table below).

A cosmid is a plasmid carrying a phage λ *cos* site, allowing it to be packaged into a phage head. Cosmids infect a host bacterium as do phages, but replicate like plasmids and the host cells are not lysed. Mammalian genes are often greater than 100 kb in size, so originally there were limitations in cloning complete gene sequences. Vectors engineered more recently have circumvented this problem by mimicking the properties of host cell chromosomes. This new generation of artificial chromosome vectors includes bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs), and mammalian artificial chromosomes (MACs).

Table 8.2 Principal features and applications of different cloning vector systems.

Vector	Basis	Size limits of insert	Major application
Plasmid	Naturally occuring multicopy plasmids	≤ 10 kb	Subcloning and downstream manipulation, cDNA cloning and expression assays
Phage	Bacteriophage λ	5–20 kb	Genomic DNA cloning, cDNA cloning, and expression libraries
Cosmid	Plasmid containing a bacteriophage λ *cos* site	35–45 kb	Genomic library construction
BAC (bacterial artificial chromosome)	*Escherichia coli* F factor plasmid	75–300 kb	Analysis of large genomes
YAC (yeast artificial chromosome)	*Saccharomyces cerevisiae* centromere, telomere, and autonomously replicating sequence	100–1000 kb (1 Mb)	Analysis of large genomes, YAC transgenic mice
MAC (mammalian artificial chromosome)	Mammalian centromere, telomere, and origin of replication	100 kb to > 1 Mb	Under development for use in animal biotechnology and human gene therapy

PLASMIDS AS CLONING VECTORS

Plasmids are circular DNA molecules that replicate separately from the host chromosome. Naturally occurring bacterial plasmids range in size from 5,000 to 400,000 bp. They can be introduced into bacterial cells by a process called **transformation.** The cells (generally *E.* For reasons not well understood, some of the cells treated in this way take up the plasmid DNA *coli*) and plasmid DNA are incubated together at 0 °C in a calcium chloride solution, then subjected to a shock by rapidly shifting the temperature to 37 to 43 °C.. Some species of bacteria are naturally

competent for DNA uptake and do not require the calcium chloride treatment. In an alternative method, cells incubated with the plasmid DNA are subjected to a high-voltage pulse. This approach, called **electroporation,** transiently renders the bacterial membrane permeable to large molecules.

Regardless of the approach, few cells actually take up the plasmid DNA, so a method is needed to select those that do. The usual strategy is to use a plasmid that includes a gene that the host cell requires for growth under specific conditions, such as a gene that confers resistance to an antibiotic. Only cells transformed by the recombinant plasmid can grow in the presence of that antibiotic, making any cell that contains the plasmid "selectable" under those growth conditions. Such a gene is called a selectable marker.

Investigators have developed many different plasmid vectors suitable for cloning by modifying naturally occurring plasmids. The *E. coli* plasmid pBR322 offers a good example of the features useful in a cloning vector (Fig. 9-4):

1. pBR322 has an origin of replication, ori, a sequence where replication is initiated by cellular enzymes (Chapter 25). This sequence is required to propagate the plasmid and maintain it at a level of 10 to 20 copies per cell.

2. The plasmid contains two genes that confer resistance to different antibiotics (tet^R, amp^R), allowing the identification of cells that contain the intact plasmid or a recombinant version of the plasmid (Fig. 9–5).

3. Several unique recognition sequences in pBR322 (*Pst*I, *Eco*RI, *Bam*HI, *Sal*I, *Pvu*II) are targets for different restriction endonucleases, providing sites where the plasmid can later be cut to insert foreign DNA.

4. The small size of the plasmid (4,361 bp) facilitates its entry into cells and the biochemical manipulation of the DNA.

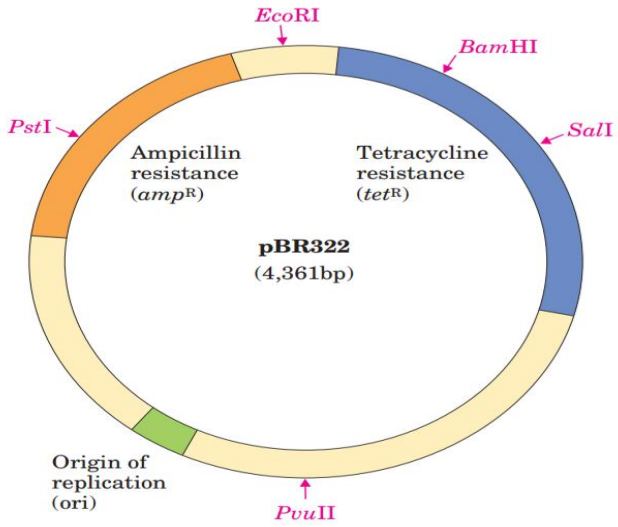


FIGURE 9–4 The constructed *E. coli* plasmid pBR322. Note the location of some important restriction sites—for *Pst*I, *Eco*RI, *Bam*HI, *Sal*I, and *Pvu*II; ampicillin- and tetracycline-resistance genes; and the replication origin (ori). Constructed in 1977, this was one of the early plasmids designed expressly for cloning in *E. coli*.

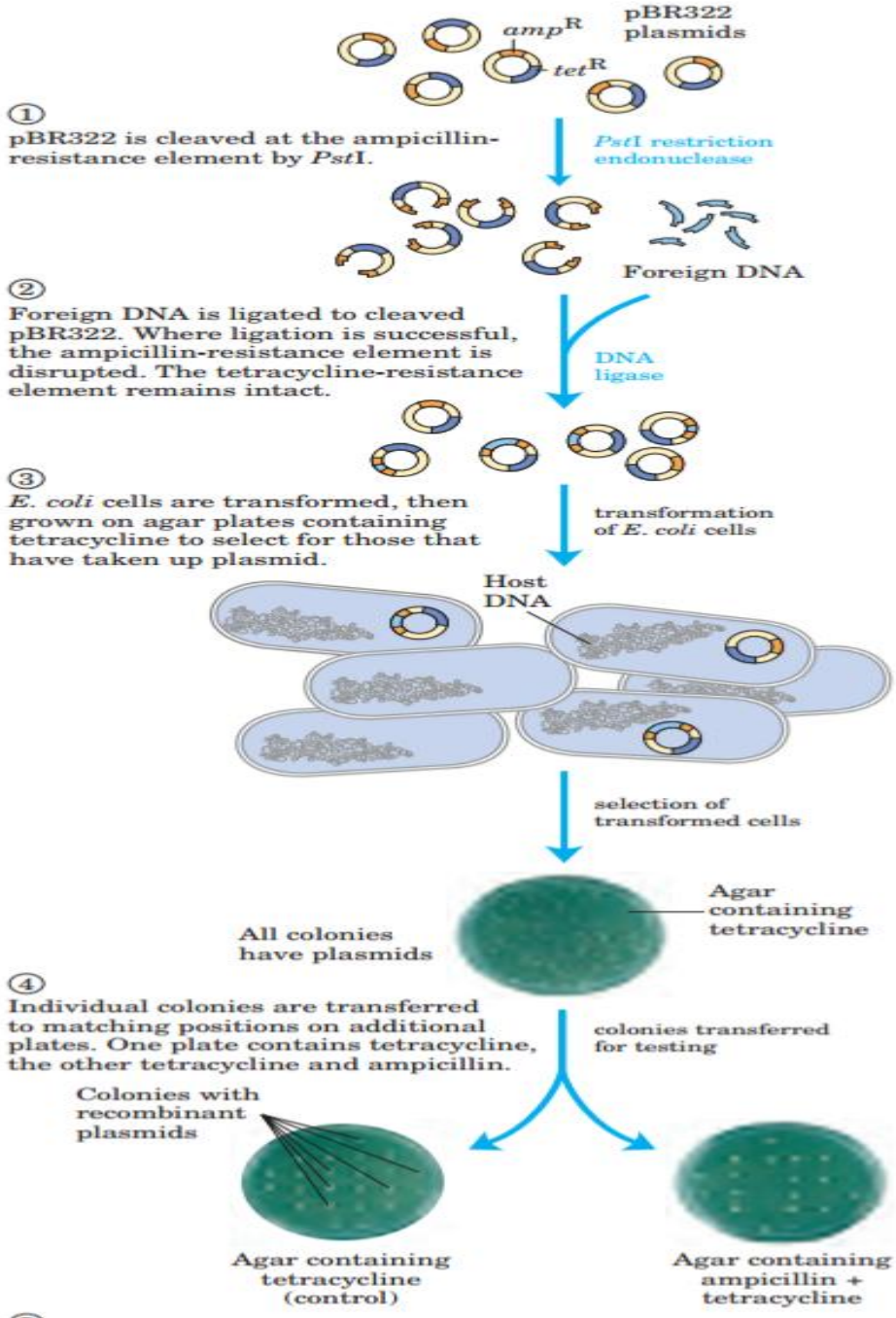


FIGURE 9–5 Use of pBR322 to clone and identify foreign DNA in *E. coli*. Plasmid Cloning

BACTERIOPHAGES AS CLONING VECTORS

Bacteriophages are the viruses that infect bacterial cells after infecting their genetic material (DNA or RNA) and kill them. The viral DNA replicates and expresses inside the bacterial cells and produces a number of phage particles released after bursting the bacterial cells. This is called lytic cycle of bacteriophage. The released phage re-infects the live cells. The ability of transforming the viral DNA from phage capsid specific bacterial cell gave insight to the scientists to exploit bacteriophages and design them as cloning vectors. The two bacteriophages Phage λ and M13 have been modified extensively and commonly used as cloning vectors

Bacteriophage lambda (λ) has been widely used in recombinant DNA since engineering of the first viral cloning vector in 1974. Phage λ vectors are particularly useful for preparing genomic libraries, because they can hold a larger piece of DNA than a plasmid vector. Today many variations of λ vectors exist. Insertion vectors have unique restriction endonuclease sites that allow the cloning of small DNA fragments in addition to the phage λ genome. These are often used for preparing cDNA expression libraries. Replacement vectors have paired cloning sites on either side of a central gene cluster. This central cluster contains genes for lysogeny and recombination, which are not essential for the lytic life cycle (see Fig. 8.1).

The central gene cluster can be removed and foreign DNA inserted between the "arms." All phage vectors used as cloning vectors have been disarmed for safety and can only function in special laboratory conditions.

A typical strategy for the use of a phage λ replacement vector is depicted in Fig. 8.8. The recombinant viral particle infects bacterial host cells, in a process called "transduction." The host cells lyse after phage reproduction, releasing progeny virus particles. The viral particles appear as a clear spot of lysed bacteria or "plaque" on an agar plate containing a lawn of bacteria. Each plaque represents progeny of a single recombinant phage and contains millions of recombinant phage particles. Most contemporary vectors carry a *lacZ′* gene allowing blue-white selection.

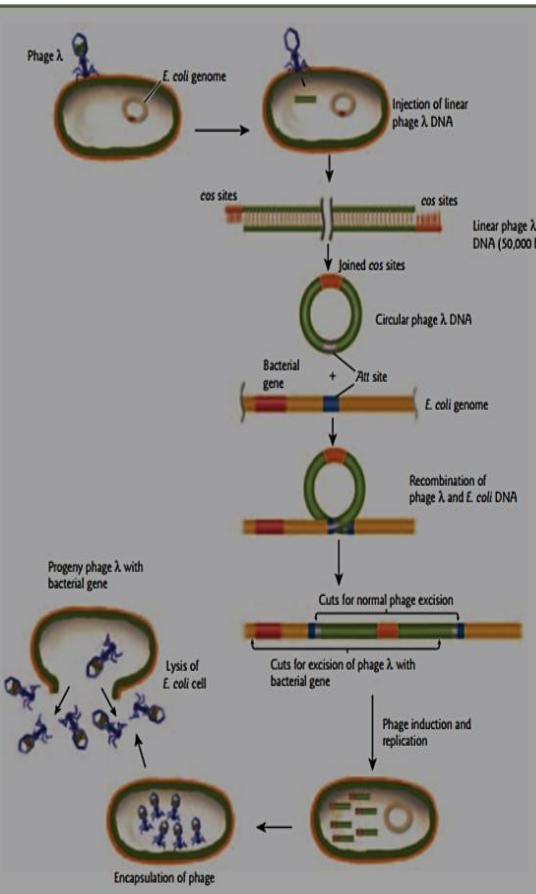


Figure 8.1 Bacteriophage lambda (λ) cohesive sites. Following the injection of a linear phage λ DNA into *E. coli* host cells, the phage λ genome circularizes by joining of the cohesive (*cos*) sites. In the lysogenic mode of replication, phage DNA is incorporated into the host genome by recombination at attachment (*Att*) sites on the phage and bacterial chromosome, and replicated as part of the host DNA. Under certain conditions, such as when the host encounters mutagenic chemicals or UV radiation, reversal of this recombination event leads to excision of the phage DNA. Rare excision events at different places allow phage λ to pick up bacterial genes. In the lytic mode of the phage life cycle, phage λ progeny with bacterial genes incorporated in their genomes are released from the lysed *E. coli*.

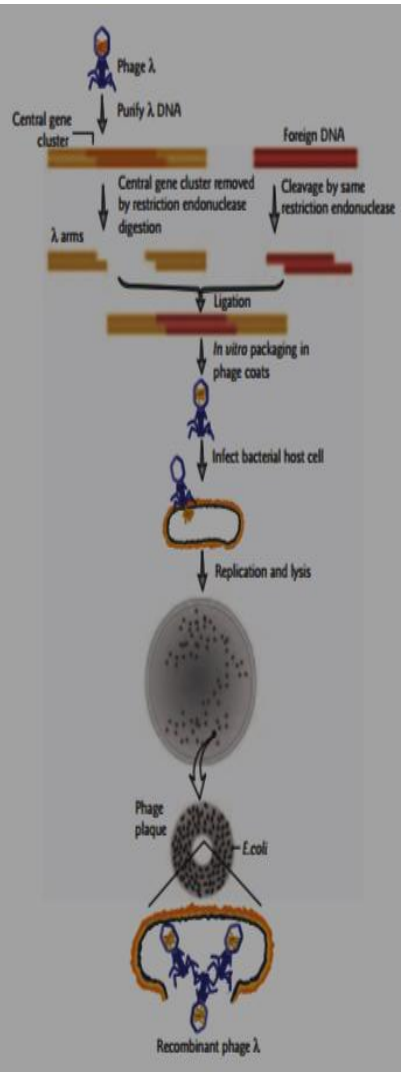


Figure 8.8 Use of bacteriophage lambda (λ) as a cloning vector. DNA is extracted from phage λ and the central gene cluster is removed by restriction endonuclease digestion. The foreign DNA to be cloned is cut with the same enzyme and ligated to the left and right "arms" of the phage λ DNA. The recombinant DNA is then mixed with phage proteins *in vitro*. The DNA is packaged into the phage head and tail fibers are attached via a self-assembly pathway. The recombinant viral particle is then able to infect bacterial cells on an agar plate. The phage replicates its genome, including the foreign DNA insert. Recombinant phage λ DNA directs the cell to make phage particles. The bacteria become filled with new phage particles, break open (lyse), and release millions of recombinant phages. The holes in the lawn of host bacteria, called plaques, are regions where phages have killed the bacteria. Each plaque represents progeny of a single recombinant phage.

Following are the advantages of phage cloning system over the Plasmids:

1. DNA can be packed in vitro into phage particles and transduced into E. coli with high efficiency
2. Foreign DNA upto 25 kb in length can bee inserted into phage vector.
3. Screening and storage of recombinant DNA is easier.

RECOMBINANT DNA TECHNOLOGY TOOLS

The standard use of recombinant DNA technology was to add genetic information to the genome, generating a *gain of function* in the resulting cells and transgenic organisms. Now that such gene-transfer processes are routine, the focus of transgenic technology has shifted to the provision of more control over the behavior of transgenes. New techniques have evolved in parallel in animals and plants. Advances have come about through the development of inducible expression systems that facilitate external regulation of transgenes and the exploitation of site-specific recombination systems to make precise modifications in target genomes. In mice, the combination of gene targeting, site-specific recombination and inducible transgene expression makes it possible to selectively switch on and off both transgenes and endogenous genes in a conditional manner. Other routes to gene silencing have also been explored, such as the expression of antisense RNA and the recently described phenomenon of RNA interference. Transgenic animals and plants are also being used increasingly as tools for the analysis of genomes.

Moreover, the cloned genes are utilized commercially in medicals, agriculture and industry and production of variable products.

An outline of these applications can be given as:

1. **CLONED GENES AFOR PRODUCTION OF GROWTH HORMONES, VACCINES AND COMMERCIAL CHEMICALS**

HUMAN PEPTIDE HORMONE GENES

Insulin:

Itakura et al chemically synthesized DNA sequence for two chains A and B of insulin and were separately inserted into two PBR322 plasmids. The recombinant plasmids were separately transferred into E. coli cells which secreted fused β-galactosidase –B chain. These were isolated and insulin was produced

Somatotropin (growth hormone)

Biosynthesis of Somatotropin was achieved through gene cloning procedures. For that double stranded cDNAs were produced from mRNA precursor of hGH gene which was then incorporated into bacterial cells where it was expressed in non-precursor form. The rDNA plasmids are allowed to transform E. coli

Somatostatin

Two plasmids, PSOM 1 and PSOM 111 were constructed for the synthesis of Somatostatin. The synthetic gene was introduced into E. coli and Somatostatin was produced.

B-endorphin

Shine et al integrated DNA sequences of B-endorphin obtained from mRNA adjacent to β-galactosidase genes on plasmid.

HUMAN INTERFERON GENES

There are three types of human interferon α, β and γ. These are being successfully produced from genetically engineered E. coli cells

By isolation of mRNA from leukocytes and fibroblasts, production of cDNA, its integration into PBR322 and incorporation and cloning into E. coli cells

GENES FOR VACCINES AND IMMUNOGENIC SUBSTANCES

Some of the vaccines synthesized by recombinant DNA technology are:

- Recombinant vaccine for Hepatitis B virus
- Indigenous Hepatitis B vaccine
- Vaccines for rabies virus
- Vaccines for polio virus
- Vaccines for foot and mouth disease virus
- Vaccines for small pox virus
- Malaria vaccines etc.

GENES ASSOCIATED WITH GENETIC DISEASES

Recombinant DNA technology is used to cure certain defective genetic diseases like

- Phenylketonuria genes
- Urokinase genes
- Thalassaemia genes
- Haemophilia genes

PRODUCTION OF COMMERCIAL CHEMICALS

There are several chemicals which are produced by using the recombinant DNA technologies. A few are:

- Vitamins
- Organic Acids
- Alcohols
- Antibiotics

2. PRESERVATION, DIAGNOSIS AND CURE OF DISEASES

PRESERVATION OF DISEASES

For preservation of diseases several immunogenic polypeptides (vaccines) and proteins (antibiotics) have been synthesized

DIAGNOSIS OF DISEASES

Such techniques have been solved the problem of conventional methods for diagnosis of many diseases. DNA probe, monoclonal antibodies and antenatal diagnosis are some of methods used for diagnosis of disease.

GENE THERAPY

A remarkable advancement in recombinant DNA technology has occurred and accumulated such knowledge that has made possible to transfer genes for treatment of human diseases.

The abnormal gene expression can manifest in the form of specific genetic disorders. The process of delivering of genes in humans to correct diseases is called as "Gene Therapy"

The ultimate goal of the gene therapy is the gene replacement therapy. Gene therapy permits physiological regulation of the transgenes and elimination of other cellular genes which occur at the time of random integration of foreign gene.

Gene therapy can be somatic therapy when genetic defects are corrected in somatic cells of the body and germ-line gene therapy, when genetic defects are corrected in germ cell.

3. DNA FINGER PRINTING OR PROFILING

The complicated technology that facilitates the identification of individuals at genetic level is known as DNA fingerprinting or DNA profiling. This genetic analysis is based on identification of tiny segments of the hereditary material which testify the unique molecular signature which cannot be altered. With the help of DNA profiling technology identify of burnt or unidentifiable dead body can be done.

4. ANIMAL AND PLANT IMPROVEMENT

TRANSGENIC ANIAMLS

By the help of rDNA technology good yielding transgenic farm animals have been produced by incorporation of different transgenes

CROP IMROVEMENT

Following are some steps towards crop improvement by rDNA technology:

- Transgenic plants
- Nif gene transfer
- Phaseolin gene transfer
- Conversion of C3 into C4 plants
- Herbicide resistant plants
- Insect pest resistant plants
- Plant improvement through genetic transformation
- Agro bacterium mediated gene transfer

RESTRICTION ENZYMES AND DNA LIGASES

Two major categories of enzymes are important tools in the isolation of DNA and the preparation of recombinant DNA: restriction endonucleases and DNA ligases. Restriction endonucleases recognize a specific, rather short, nucleotide sequence on a double-stranded DNA molecule, called a restriction site, and cleave the DNA at this recognition site or elsewhere, depending on the type of enzyme. DNA ligase joins two pieces of DNA by forming phosphodiester bonds.

1. RESTRICTION ENZYMES

Restriction endonucleases are found in a wide range of bacterial species. Werner Arber discovered in the early 1960s that their biological function is to recognize and cleave foreign DNA (the DNA of an infecting virus, for example); such DNA is said to be *restricted.* In the host cell's DNA, the sequence that would be recognized by its own restriction endonuclease is protected from digestion by methylation of the DNA, catalyzed by a specific DNA methylase. The restriction endonuclease and the corresponding methylase are sometimes referred to as a **restriction-modification system.**

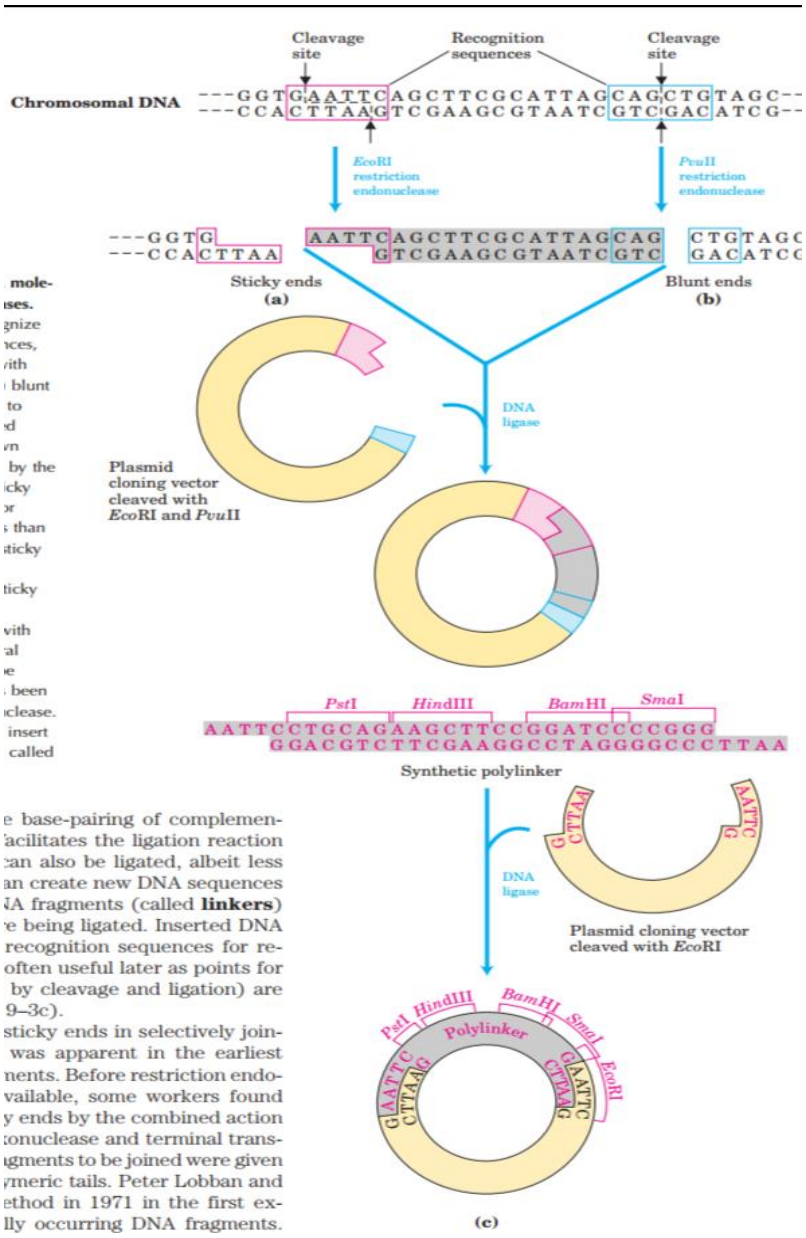
Type I and III restriction endonucleases are not useful for gene cloning because they cleave DNA at sites other than the recognition sites and thus cause random cleavage patterns. In contrast, type II endonucleases are widely used for mapping and reconstructing DNA in vitro because they recognize specific sites and cleave just at these sites (Table 8.1).

Table 8.1 Major classes of restriction endonucleases.

Class	Abundance	Recognition site	Composition	Use in recombinant DNA research
Type I	Less common than type II	Cut both strands at a nonspecific location > 1000 bp away from recognition site	Three-subunit complex: individual recognition, endonuclease, and methylase activities	Not useful
Type II	Most common	Cut both strands at a specific, usually palindromic, recognition site (4–8 bp)	Endonuclease and methylase are separate, single-subunit enzymes	Very useful
Type III	Rare	Cleavage of one strand only, 24–26 bp downstream of the 3′ recognition site	Endonuclease and methylase are separate two-subunit complexes with one subunit in common	Not useful

Some restriction endonucleases make staggered cuts on the two DNA strands, leaving two to four nucleotides of one strand unpaired at each resulting end. These unpaired strands are referred to as **sticky ends** (Fig. 9–3a), because they can base-pair with each other or with complementary sticky ends of other DNA fragments.

Other restriction endonucleases cleave both strands of DNA at the opposing phosphodiester bonds, leaving no unpaired bases on the ends, often called **blunt ends** (Fig. 9–3b). The base-pairing of complementary sticky ends greatly facilitates the ligation reaction (Fig. 9–3a). Blunt ends can also be ligated, albeit less efficiently. Researchers can create new DNA sequences by inserting synthetic DNA fragments (called **linkers**) between the ends that are being ligated. Inserted DNA fragments with multiple recognition sequences for restriction endonucleases (often useful later as points for inserting additional DNA by cleavage and ligation) are called **polylinkers** (Fig. 9–3c).



mole-
ses.
gnize
ces,
with
blunt
to
d
vn
by the
icky
or
than
sticky
ticky
with
al
e
been
clease.
insert
called

e base-pairing of complemen-
facilitates the ligation reaction
can also be ligated, albeit less
an create new DNA sequences
A fragments (called **linkers**)
e being ligated. Inserted DNA
recognition sequences for re-
often useful later as points for
by cleavage and ligation) are
9–3c).
sticky ends in selectively join-
was apparent in the earliest
nents. Before restriction endo-
vailable, some workers found
y ends by the combined action
onuclease and terminal trans-
gments to be joined were given
ymeric tails. Peter Lobban and
ethod in 1971 in the first ex-
lly occurring DNA fragments.

Restriction endonclease nomenclature

Restriction endonucleases are named for the organism in which they were discovered, using a system of letters and numbers. For example, *Hin*dIII (pronounced "hindee-three") was discovered in *Haemophilusinfluenza* (strain d). The *Hin* comes from the first letter of the genus name and the first two letters of the species name; d is for the strain type; and III is for the third enzyme of that type. *Sma*I is from *Serratia marcescens* and is pronounced "smah-one," *Eco*RI (pronounced "echo-r-one") was discovered in *Escherichia coli* (strain R), and *Bam*HI is from *Bacillus amyloliquefaciens* (strain H). Over 3000 type II restriction endonucleases have been isolated and characterized to date. Approximately 240 are available commercially for use by molecular biologists.

2. DNA LIGASES

The study of DNA replication and repair processes led to the discovery of the DNA-joining enzyme called DNA ligase. DNA ligases catalyze formation of a phosphodiester bond between the 5′-phosphate of a nucleotide on one fragment of DNA and the 3′-hydroxyl of another (see Fig. 6.14). This joining of linear DNA fragments together with covalent bonds is called ligation. Unlike the type II restriction endonucleases, DNA ligase requires ATP as a cofactor.

Because it can join two pieces of DNA, DNA ligase became a key enzyme in genetic engineering. If restriction-digested fragments of DNA are placed together under appropriate conditions, the DNA fragments from two sources can anneal to form recombinant molecules by hydrogen bonding between the complementary base pairs of the sticky ends. However, the two strands are not covalently bonded by phosphodiester bonds. DNA ligase is required to seal the gaps, covalently bonding the two strands and regenerating a circular molecule. The DNA ligase most widely used in the lab is derived from the bacteriophage T4. T4 DNA ligase will also ligate fragments with blunt ends, but the reaction is less efficient and higher concentrations of the enzyme are usually required *in vitro*. To increase the efficiency of the reaction, researchers often use the enzyme terminal deoxynucleotidyl transferase to modify the blunt ends. For example, if a single-stranded poly (dA) tail is added to DNA fragments from one source, and a single stranded poly (dT) tail is added to DNA from another source, the complementary tails can hydrogen bond (Fig. 8.5). Recombinant DNA molecules can then be created by ligation.

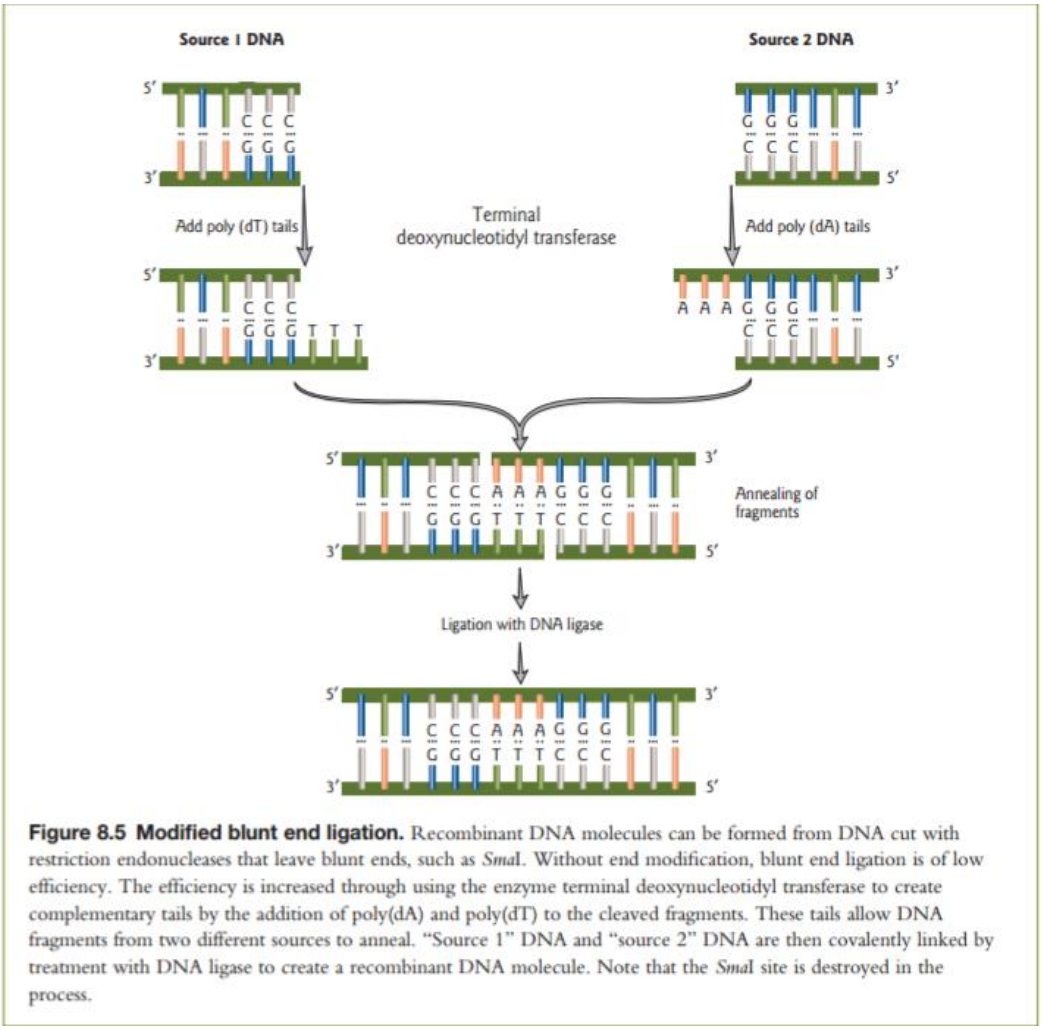


Figure 8.5 Modified blunt end ligation. Recombinant DNA molecules can be formed from DNA cut with restriction endonucleases that leave blunt ends, such as *Sma*I. Without end modification, blunt end ligation is of low efficiency. The efficiency is increased through using the enzyme terminal deoxynucleotidyl transferase to create complementary tails by the addition of poly(dA) and poly(dT) to the cleaved fragments. These tails allow DNA fragments from two different sources to anneal. "Source 1" DNA and "source 2" DNA are then covalently linked by treatment with DNA ligase to create a recombinant DNA molecule. Note that the *Sma*I site is destroyed in the process.

3. PHOSPHATASES

A Phosphatase is an enzyme that removes a phosphate group from its substrate by hydrolyzing phosphoric acid monoesters into a phosphate ion and a molecule with a free hydroxyl group. This action is directly opposite to that of phosphorylases and kinases, which attach phosphate groups to their substrates by using energetic molecules like ATP.

A common phosphatase in many organisms is alkaline phosphatase. Another large group of proteins present in archaea, bacteria and eukaryotes are deoxyribonucleotide and ribonuceotide phosphatase or pyrophosphatase that catalyse the decomposition of dNTP/NTP in dNDP/NDP and free phosphate ion ot dNMP/NMP and a free pyrophosphate ion. The other group of phosphatase is collectively called as protein phosphatase, which removes a phosphate group from the phosphorylated amino acid residue of the substrate protein. Protein phosphorylation is common post translational modification of protein catalyzed by protein kinase, and protein phosphatases reverse the effect

4. T4 POLYNUCLEOTIDE KINASE

T4 Polynucleotide Kinase catalyzes the transfer and exchange of P_i from the γ position of ATP to the 5′- hydroxyl terminus of double and single stranded DNA and RNA, as well as nucleoside 3′-mono phosphates.

Oligonucleotide labeling

Methods for end-labeling oligonucleotides do not involve DNA synthesis reactions. Instead, they make use of other enzymes. In 5′ end-labeling, a *g*-phosphate from ATP is added to the 5′ end of an oligonucleotide by the enzyme T4 polynucleotide kinase. In 3′ end-labeling, a labeled dNTP is added to the 3′ end by the enzyme terminal transferase

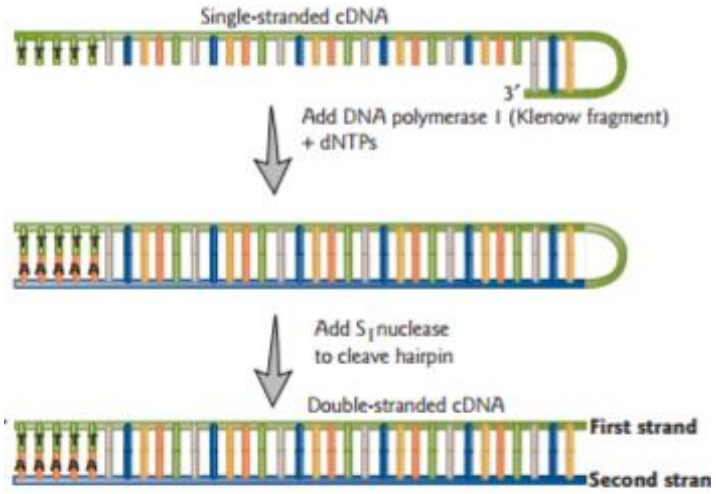
5. KLENOW FRAGMENT OF DNA POLYMERASE I

It is the product of enzymatic breakdown of DNA pol I from E. coli. Its molecular weight is 76 Kd. It has 3′ → 5′ exonuclease activity. It has the ability to do polymerization under in vitro conditions. It lacks 5′ → 3′ exonuclease activity

For most applications, including cloning of cDNAs, double stranded DNA is required. The second DNA strand is generated by the Klenow fragment of DNA polymerase I from *E. coli* (Fig. 1C). The 5′ → 3′ exonuclease activity of DNA polymerase I from *E. coli* makes it unsuitable for many applications. However, this enzymatic activity can be readily removed from the holoenzyme by exposure to a protease. The large or Klenow fragment of DNA polymerase I generated by proteolysis has 5′ → 3′ polymerase and 3′ → 5′ exonuclease (proofreading) activity, and is widely used in molecular biology. Commercially available Klenow fragments are usually produced by expression in bacteria from a truncated form of the DNA polymerase I gene.

Klenow fill-in

A "fill-in" reaction is used to generate blunt ends on fragments created by cleavage with restriction endonucleases that leave 5′ single-stranded overhangs. The Klenow fragment of *E. coli* DNA polymerase I is used to fill in the gaps from 5′ to 3′, in the presence of dNTPs, including one labeled dNTP (Fig.below). The result is a doublestranded DNA with the 3′ ends labeled.



DNA SEQUENCING

Until the late 1970s, determining the sequence of a nucleic acid containing even five or ten nucleotides was difficult and very laborious. The development of two new techniques in 1977, one by Alan Maxam and Walter Gilbert (Chemical Degradation Method) and the other by Frederick Sanger (enzymatic method/ dideoxy method) , has made possible the sequencing of ever larger DNA molecules with an ease unimagined just a few decades ago. The techniques depend on an improved understanding of nucleotide chemistry and DNA metabolism, and on electrophoretic methods for separating DNA strands differing in size by only one nucleotide. Electrophoresis of DNA is similar to that of proteins. Polyacrylamide is often used as the gel matrix and in work with short DNA molecules (up to a few hundred nucleotides), Agarose is generally used for longer pieces of DNA.

MAXAM AND GILBERT'S CHEMICAL DEGREDATIION METHOD

This method is not much popular because it is time consuming and labour intensive. In this method, the DNA molecule can be radiolabelled at either 5′ end by using polynucleotide kinase or 3′ end by terminal transferase. One end of radiolabelled double stranded DNA is removed using endonuclease. A base is modified chemically follwed by cleavage of sugar-phosphate backbone of DNA. Not any specific reaction for 4 bases is carried out, except specific reaction to G only and purine specific reaction which removes A or G. a difference in these reactions indicate the presence of A.

Cleavage of purine: the mixture is separated in four sets, each treated with different reagents which degredate only G or C or A and G or C and T. in one set, DNA is treated with acid followed by dimethyl sulphate. This causes methylation of A (at 3 ′ position) and G (at 7′ position). Subsequently addition of alkali (-OH) and piperdine results in cleavage of DNA and removal of purines (A or G)

Clevage of pyrimidines: similar to the cleavage of purine, pyrimidine (C or T) is also cleaved in the presence of 1-2 M Nacl Soln. It works only with C. Difference b/w these two indicate the presence of T in the DNA sequence cleavage of pyrimidines also takes place through hydrazine hydrolysis.

Partial chemical cleavage of DNA fragments as done above generates the populations of radiolabelled molecules extending from radiolabelled terminus to the site of chemical cleavage. These fragments are of different sizes that represent unique pairs of 5′ and 3′ cleavage products in the random collection. A complete set is formed by these products, the length of each no. is short by one nucleotide. These can be separated by gel electrophoresis. The fragments containing labelled terminus can be observed by autoradiography of the gel. Following the order of fragments obtained from digestions the sequence of nucleotides is deduced and interpreted.

FREDERICK SANGER'S (ENZYMATIC METHOD/ DIDEOXY METHOD)

In this method, single-stranded DNA is mixed with a radioactively labeled primer to provide the 3′-OH required for DNA polymerase to initiate DNA synthesis. The primer is usually complementary to a region of the vector just outside the multiple cloning sites. The sample is then split into four aliquots, each containing DNA polymerase, four dNTPs (at high

concentration), and a low concentration of a replication terminator. The replication terminators are dideoxynucleoside triphosphates (ddNTPs) that are missing the 3′-OH. Because they lack the 3′-OH, they cannot form a phosphodiester bond with another nucleotide. Thus, each reaction proceeds until a replication-terminating nucleotide is added, and each of the four sequencing reactions produces a series of single-stranded DNA molecules, each one base longer than the last.

The polymerase of choice for DNA sequencing is phage T7 DNA polymerase (called "Sequenase"). The sequencing mixtures are loaded into separate lanes of a denaturing polyacrylamide gel and electrophoresis is used to separate the DNA fragments. Autoradiography is used to detect a ladder of radioactive bands. The radioactive label (primer) is at the 5′ end of each newly synthesized DNA molecule. Thus, the smallest fragment at the bottom of the gel represents the 5′ end of the DNA. Reading the sequence of bases from the bottom up (5′ → 3′) gives the sequence of the DNA molecule synthesized in the sequencing reaction. The sequence of the original strand of DNA is complementary to the sequence read from the gel (3′ → 5′).

This method can separate DNA of approximately 500 nt. For longer sequences, overlapping fragments are sequenced. The technique is very laborious and the sequences have to be read by hand.

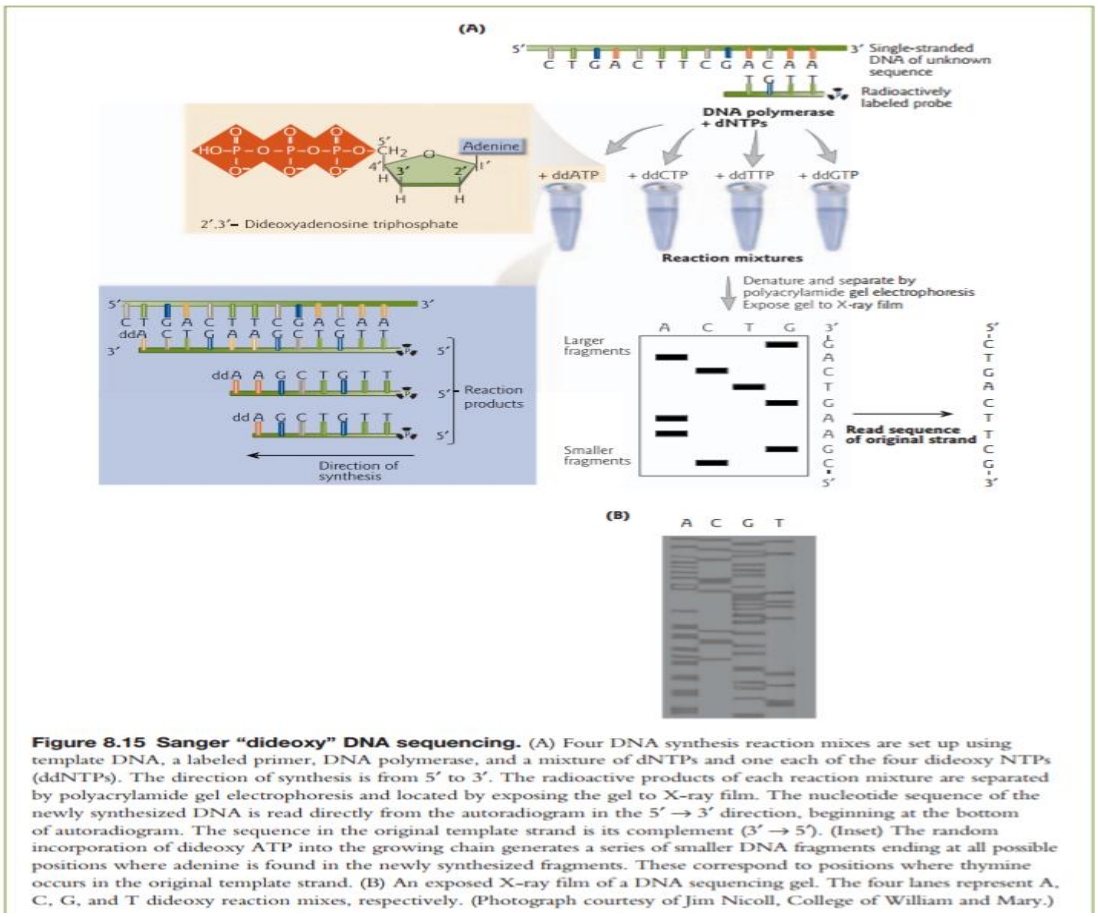


Figure 8.15 Sanger "dideoxy" DNA sequencing. (A) Four DNA synthesis reaction mixes are set up using template DNA, a labeled primer, DNA polymerase, and a mixture of dNTPs and one each of the four dideoxy NTPs (ddNTPs). The direction of synthesis is from 5′ to 3′. The radioactive products of each reaction mixture are separated by polyacrylamide gel electrophoresis and located by exposing the gel to X-ray film. The nucleotide sequence of the newly synthesized DNA is read directly from the autoradiogram in the 5′ → 3′ direction, beginning at the bottom of autoradiogram. The sequence in the original template strand is its complement (3′ → 5′). (Inset) The random incorporation of dideoxy ATP into the growing chain generates a series of smaller DNA fragments ending at all possible positions where adenine is found in the newly synthesized fragments. These correspond to positions where thymine occurs in the original template strand. (B) An exposed X-ray film of a DNA sequencing gel. The four lanes represent A, C, G, and T dideoxy reaction mixes, respectively. (Photograph courtesy of Jim Nicoll, College of William and Mary.)

POLYMERASE CHAIN REACTION (PCR)

The polymerase chain reaction (PCR) is the one of the most powerful techniques that has been developed recently in the area of recombinant DNA research. PCR has had a major impact on many areas of molecular cloning and genetics. With this technique, a target sequence of DNA can be amplified a billion-fold in several hours. Amplification of particular segments of DNA by PCR is distinct from the amplification of DNA during cloning and propagation within a host cell. The procedure is carried out entirely *in vitro*. In addition to its use in many molecular cloning strategies, PCR is also used in the analysis of gene expression, forensic analysis where minute samples of DNA are isolated from a crime scene and diagnostic tests for genetic diseases.

PCR is a DNA polymerase reaction. As with any DNA polymerase reaction it requires a DNA template and a free 3′-OH to get the polymerase started. The template is provided by the DNA sample to be amplified and the free 3′-OH groups are provided by site-specific oligonucleotide primers. The primers are complementary to each of the ends of the sequence that is to be amplified. Note that *in vivo* DNA polymerase would use an RNA primer, but a more stable, more easily synthesized DNA primer is used *in vitro*. The three steps of the reaction are denaturation, annealing of primers, and primer extension (Fig. 1):

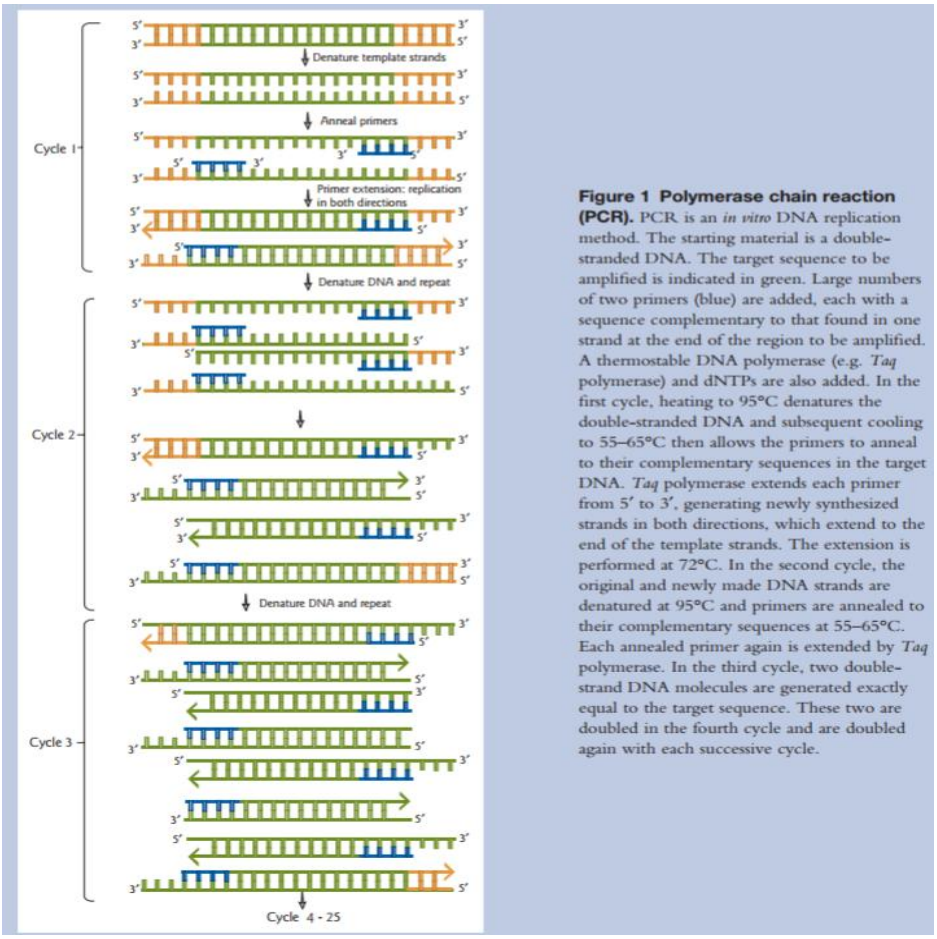


Figure 1 Polymerase chain reaction (PCR). PCR is an *in vitro* DNA replication method. The starting material is a double-stranded DNA. The target sequence to be amplified is indicated in green. Large numbers of two primers (blue) are added, each with a sequence complementary to that found in one strand at the end of the region to be amplified. A thermostable DNA polymerase (e.g. *Taq* polymerase) and dNTPs are also added. In the first cycle, heating to 95°C denatures the double-stranded DNA and subsequent cooling to 55–65°C then allows the primers to anneal to their complementary sequences in the target DNA. *Taq* polymerase extends each primer from 5′ to 3′, generating newly synthesized strands in both directions, which extend to the end of the template strands. The extension is performed at 72°C. In the second cycle, the original and newly made DNA strands are denatured at 95°C and primers are annealed to their complementary sequences at 55–65°C. Each annealed primer again is extended by *Taq* polymerase. In the third cycle, two double-strand DNA molecules are generated exactly equal to the target sequence. These two are doubled in the fourth cycle and are doubled again with each successive cycle.

1. **Denaturation.** In the first step, the target sequence of DNA is heated to denature the template strands and render the DNA single-stranded.
2. **Annealing.** The DNA is then cooled to allow the primers to anneal, that is, to bind the appropriate complementary strand. The temperature for this step varies depending on the size of the primer, the GC content, and its homology to the target DNA. Primers are generally DNA oligonucleotides of approximately 20 bases each.
3. **Primer extension.** In the presence of Mg2+, DNA polymerase extends the primers on both strands from 5′ to 3′ by its polymerase activity. Primer extension is performed at a temperature optimal for the particular polymerase that is used. Currently, the most popular enzyme for this step is *Taq* polymerase, the DNA polymerase from the thermophilic (heat-loving) bacteria *Thermus aquaticus*. This organism lives in hot springs that can be near boiling and thus requires a thermostable polymerase.

These three steps are repeated from 28 to 35 times. With each cycle, more and more fragments are generated with just the region between the primers amplified. These accumulate exponentially. The contribution of strands with extension beyond the target sequence becomes negligible since these accumulate in a linear manner. After 25 cycles in an automated thermocycler machine, there is a 225 amplification of the target sequence. PCR products can be visualized on a gel stained with nucleic acid-specific fluorescent compounds such as ethidium bromide or SYBR green. The error rate of *Taq* is $2 \times 10{-}4$. If an error occurs early on in the cycles, it could become prominent. Other polymerases, such as *Pfu*, have greater fidelity. *Pfu* DNA polymerase is from *Pyrococcus furiosus*. Base misinsertions that may occur infrequently during polymerization are rapidly excised by the $3' \rightarrow 5'$ exonuclease (proofreading) activity of this enzyme.

When Kary Mullis first developed the PCR method in 1985, his experiments used *E. coli* DNA polymerase. Because *E. coli* DNA polymerase is heat-sensitive, its activity was destroyed during the denaturation step at 95°C. Therefore, a new aliquot of the enzyme had to be added in each cycle. The purification, and ultimately the cloning, of the DNA polymerase from *T. aquaticus* made the reaction much simpler. In his first experiments, Mullis had to move the reaction manually between the different temperatures. Fortunately, this procedure has been automated by the development of thermal cyclers. These instruments have the capability of rapidly switching between the different temperatures that are required for the PCR reaction. Thus the reactions can be set up and placed in the thermal cycler, and the researcher can return several hours later (or the next morning) to obtain the products.

APPLICATIONs OF PCR

PCR is used in:

- Diagnosis of pathogens
- Diagnosis of specific mutations (sickle cell anaemia, phenylketonuria and muscular dystrophy)
- Prenatal diagnosis of several genetic diseases
- DNA fingerprinting
- Diagnosis of plant diseases
- Sexing of embryo (sex of human or livestock embryos, Fertilized embryos)
- Molecular archaeology (palaentology)
- molecular cloning strategies,
- Analysis of gene expression,
- forensic analysis

CONSTRUCTING DNA LIBRARIES

Vectors are used to compile a library of DNA fragments that have been isolated from the genomes of a variety of organisms. This collection of fragments can then be used to isolate specific genes and other DNA sequences of interest. DNA fragments are generated by cutting the DNA with a specific restriction endonuclease. These fragments are ligated into vector molecules, and the collection of recombinant molecules is transferred into host cells, one molecule in each cell. The total number of all DNA molecules makes up the library. This library is searched, that is screened, with a molecular probe that specifically identifies the target DNA. Once prepared the library can be perpetuated indefinitely in the host cells and is readily retrieved whenever a new probe is available to seek out a particular fragment. Two main types of libraries can be used to isolate specific DNAs: genomic and cDNA libraries.

Genomic library

A genomic library contains DNA fragments that represent the entire genome of an organism. The first step in creating a genomic library is to break the DNA into manageable size pieces (e.g. 15–20 kb for phage λ vectors), usually by partial restriction endonuclease digest. Under limiting conditions, any particular restriction site is cleaved only occasionally, so not all sites are cleaved in any particular DNA molecule. This generates a continuum of overlapping fragments. The second step is to purify fragments of optimal size by gel electrophoresis or centrifugation techniques. The final step is to insert the DNA fragments into a suitable vector. In humans, the genome size is approximately 3×10^9 bp. with an average insert size of 20 kb, the number of random fragments to ensure with high probability (95–99%) that every sequence is represented is approximately 106 clones for humans. The maths actually works out to 1.5×10^5 (i.e. $(3 \times 10^9$ bp$)/(2 \times 10^4$ bp) but more clones are needed in practice, since insertion is random. Bacteriophage λ or cosmid vectors are typically used for genomic libraries. Since a larger insert size can be accommodated by these vectors compared with plasmids, there is a greater chance of cloning a gene sequence with both the coding sequence and the regulatory elements in a single clone.

cDNA library

The principle behind cDNA cloning is that an mRNA population isolated from a specific tissue, cell type, or developmental stage (e.g. embryo mRNA) should contain mRNAs specific for any protein expressed in that cell type or during that stage, along with "housekeeping" mRNAs that encode essential proteins such as the ribosomal proteins, and other mRNAs common to many cell types or stages of development. Thus, if mRNA can be isolated, a small subset of all the genes in a genome can be studied. mRNA cannot be cloned directly, but a cDNA copy of the mRNA can be cloned. Because a cDNA library is derived from mRNA, the library contains the coding region of expressed genes only, with no introns or regulatory regions. This latter point becomes important for applications of recombinant DNA technology to the production of transgenic animals and for human gene therapy

COMPLEMENTARY DNA (cDNA) SYNTHESIS

A cDNA fragment is prepared directely by using mRNA as tempelate. It involves following steps:

Isolation of mRNA

Most eukaryotic mRNAs are poly-adenylated at the 3′ end to form a poly(A) tail. This has an important practical consequence that has been exploited by molecular biologists. The poly(A) region can be used to selectively isolate mRNA from total RNA by affinity chromatography (Fig. 1A). The purified mRNA can then be used as a template for synthesis of a complementary DNA (cDNA).

Purification of mRNA

Total RNA is extracted from a specific cell type that expresses a specific set of genes. Of this total cellular RNA, 80–90% is rRNA, tRNA, and histone mRNA, not all of which have a poly (A) tail. These RNAs can be separated from the poly (A) mRNA by passing the total RNA through an affinity column of oligo(dT) or oligo(U) bound to resin beads. Under conditions of relatively high salt the poly (A) RNA is retained by formation of hydrogen bonds with the complementary bases, and the RNA lacking a poly(A) tail flows through.

The salt conditions for hybridization are similar to the ion concentration in cells (e.g. 0.3–0.6 M NaCl). The poly(A) mRNA is then eluted from the column in low salt elution buffer (e.g. 0.01 M NaCl), which promotes denaturation of the hybrid.

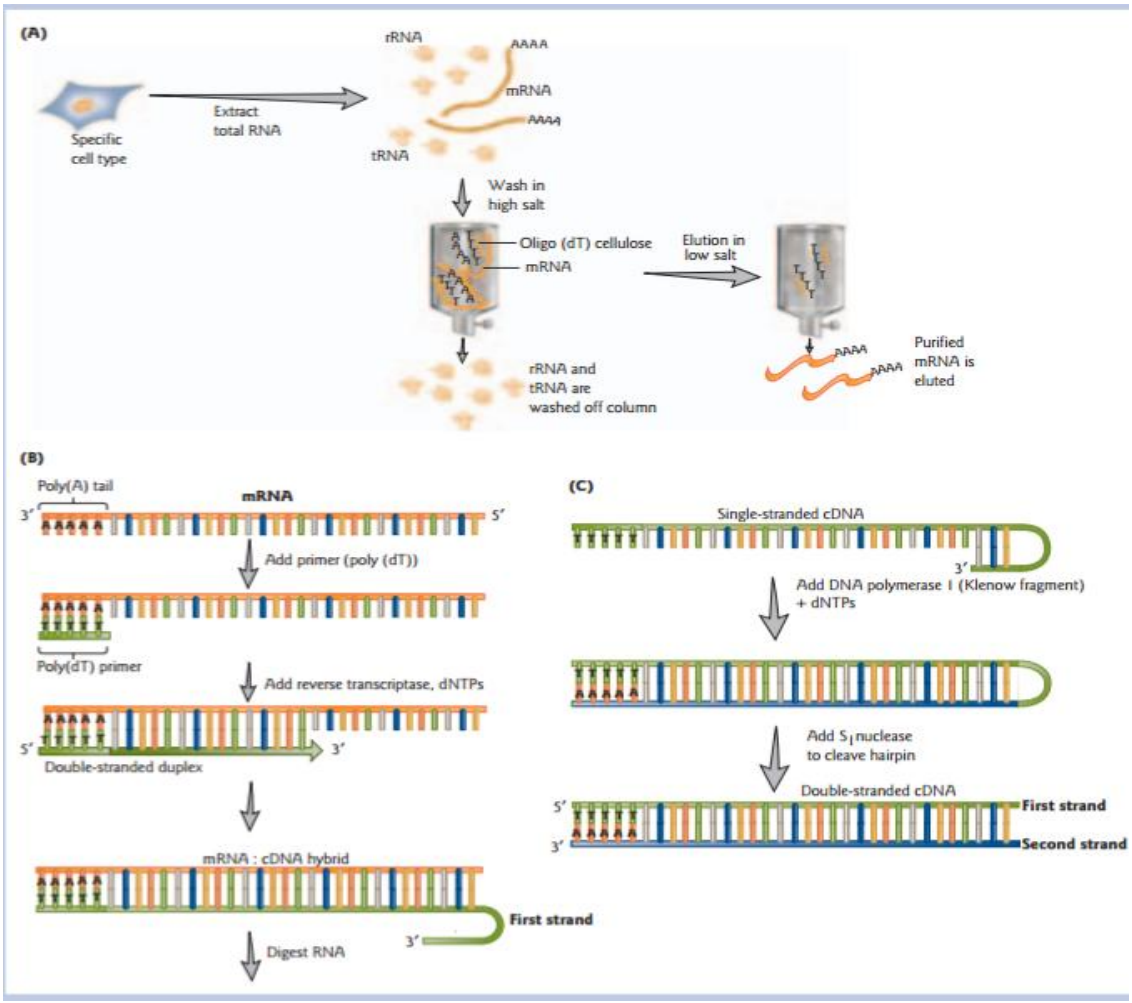
First strand synthesis of cDNA (Reverse transcription)

A number of strategies can be used to synthesize cDNA from purified mRNA. One strategy is as follows. In brief, cDNA is synthesized by the action of reverse transcriptase and DNA polymerase (Fig. 1B). The reverse transcriptase catalyzes the synthesis of a single-stranded DNA from the mRNA template. Like a regular DNA polymerase, reverse transcriptase also needs a primer to get started. A poly(dT) primer is added to provide a free 3′-OH end that can be used for extension by reverse transcriptase in the presence of deoxynucleoside triphosphates (dNTPs). Usually a viral reverse transcriptase is employed such as one from avian myeloblastosis virus (AMV). The reverse transcriptase adds dNTPs from 5′ to 3′ by complementary base pairing. This is called first strand synthesis. The mRNA is then degraded with a ribonuclease or an alkaline solution.

Second strand synthesis of cDNA

For most applications, including cloning of cDNAs, doublestranded DNA is required. The second DNA strand is generated by the Klenow fragment of DNA polymerase I from *E. coli* (Fig. 1C). The 5′ → 3′ exonuclease activity of DNA polymerase I from *E. coli* (see Focus box 6.1) makes it unsuitable for many applications. However, this enzymatic activity can be readily removed from the holoenzyme by exposure to a protease. The large or Klenow fragment of DNA polymerase I generated by proteolysis has 5′ → 3′ polymerase and 3′ → 5′ exonuclease (proofreading) activity, and is widely used in molecular biology. Commercially available Klenow fragments are usually produced by expression in bacteria from a truncated form of the DNA polymerase I gene.

There is a tendency for the reverse transcriptase enzyme used in first strand synthesis to loop back on itself and start to make another complementary strand. This hairpin forms a natural primer for DNA polymerase and a second strand of DNA is generated. S1 nuclease (from *Aspergillus oryzae*) is then added to cleave the single-stranded DNA hairpin. Double-strand DNA linkers with ends that are complementary to an appropriate cloning vector are added to the double-strand DNA molecule before ligation into the cloning vector. The end result is a double-stranded cDNA in which the second strand corresponds to the sequence of the mRNA, thus representing the coding strand of the gene. The sequences that appear in the literature are the 5′ → 3′ sequences of the second strand cDNA (Fig. 1). Sequences corresponding to introns and to promoters and all regions upstream of the transcriptional start site are not represented in cDNAs. The library created from all thecDNAs derived from the mRNAs in the specific cell type forms the cDNA library of cDNA clones.



Real-time PCR

Real time polymerase chain reaction is a technique laboratory technique of Molecular Biology based on PCR. It monitors the amplification of targeted DNA molecule during PCR i.e in real time not in its end, as in conventional PCR.

There are many applications of the PCR where it would be advantageous to be able to quantify the amount of starting material. Theoretically, there is a quantitative relationship between the amount of starting material (target sequence) and the amount of PCR product at any given cycle. In practice, replicate reactions yield different amounts of product, making quantitation unreliable. Higuchi *et al*. (1992, 1993) pioneered the use of ethidium bromide to quantify PCR products as they accumulate. Amplification produces increasing amounts of double stranded DNA, which binds ethidium bromide, resulting in an increase in fluorescence. By plotting the increase in fluorescence versus cycle number it is possible to analyse the PCR kinetics in real time. This is much more satisfactory than analysing product accumulation after a fixed number of cycles.

The principal drawback to the use of ethidium bromide is that both specific and non-specific products generate a signal. This can be overcome by the use of probe-based methods for assaying product accumulation (Livak *et al*. 1995). The probes used are oligonucleotides with a reporter fluorescent dye attached at the 5′ end and a quencher dye at the 3′ end. While the probe is intact, the proximity of the quencher reduces the fluorescence emitted by the reporter dye. If the target sequence is present, the probe anneals downstream from one of the primer sites. As the primer is extended, the probe is cleaved by the 5′ nuclease activity of the *Taq* polymerase (Fig). This cleavage of the probe separates the reporter dye from the quencher dye, thereby increasing the reporter-dye signal. Cleavage removes the probe from the target strand, allowing primer extension to continue to the end of the template strand. Additional reporter-dye molecules are cleaved from their respective probes with each cycle, affecting an increase in fluorescence intensity proportional to the amount of amplicon produced.

Instrumentation has been developed which combines thermal cycling with measurement of fluorescence, thereby enabling the progress of the PCR to be monitored in real time. This revolutionizes the way one approaches PCR-based quantitation of DNA. Reactions are characterized by the point in time during cycling when amplification of a product is first detected, rather than by the amount of PCR product accumulated after a fixed number of cycles.

The higher the starting copy number of the target, the sooner a significant increase in fluorescence is noted. Quantitation of the amount of target in unknown samples is achieved by preparing a standard curve, using different starting copy numbers of the target sequence.

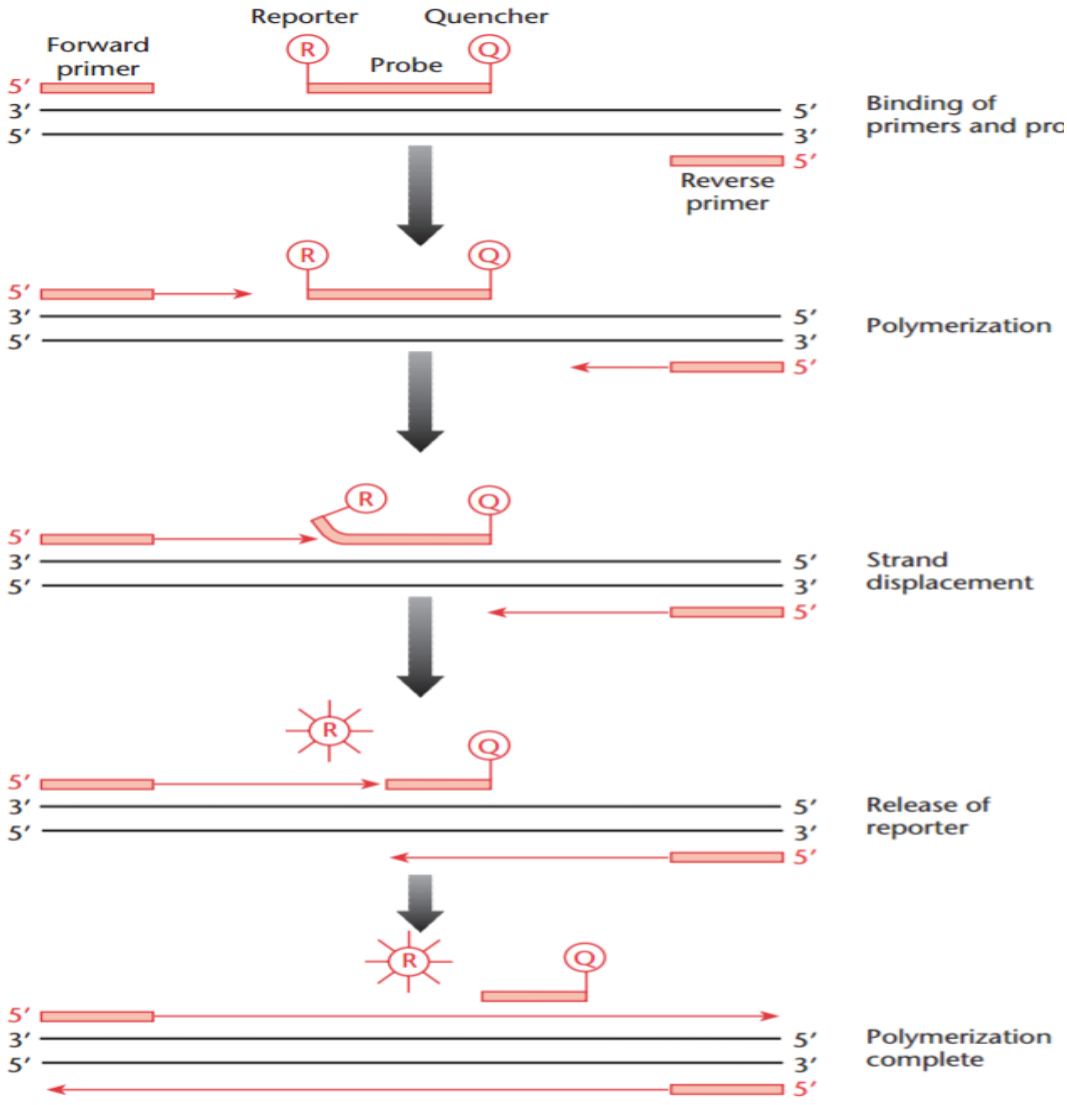


FIG. REAL TIME PCR

MOLECULAR MARKERS

Molecular Marker also called as genetic Marker is a gene or DNA sequence with a known location on chromosome that can be used to identify individuals or species. It can be described as a variation (which may arise due to mutation or alteration in genome loci) that can be observed, a Genetic Marker may be short DNA sequence such as a sequence surrounding a single bp change (single nucleotide polymorphism-SNP) or a long one like mini-satellites.

Some commonly used types of Genetic Markers are:

- RFLP (restriction fragment length polymorphism)
- AFLP (Amplified fragment length polymorphism)
- SSLP (Simple Sequence length polymorphism)
- RAPD (Random Amplification Polymorphic DNA)
- VNTR (Variable Number Tandom Repeats)

- SSR (Simple Sequence Repeats)
- SNP (Single Nucleotide Polymorphism)
- STR (Short Tandom Repeats)
- SFP (Single Feature Nucleotide)
- DAT (Diversity Arrays Technology)
- RAD (Restriction site associated DNA markers)

Genetic Markers are used to know the relationship b/w an inherited disease and its genetic cause.

RFLP – AS GENETIC MARKER

In 1980, Mark Skolnick, Ray White, David Botstein, and Ronald Davis created a restriction fragment length polymorphism (RFLP, pronounced "rif-lip") marker map of the human genome. A RFLP is defined by the existence of alternative alleles associated with restriction fragments that differ in size from each other. RFLPs are visualized by digesting DNA from different individuals with restriction endonucleases, followed by gel electrophoresis to separate fragments according to size, then Southern blotting and hybridization to a labeled probe that identifies the locus under investigation. A RFLP is demonstrated whenever the Southern blot pattern obtained with one individual is different from the one obtained with another individual. These variable regions do not necessarily occur in genes, and the function of most of those in the human genome is unknown. An exception is a RFLP that can be used to diagnose sickle cell anemia (Fig. 8.14). In individuals with sickle cell anemia, a point mutation in the β-globin gene has destroyed the recognition site for the restriction endonuclease *Mst*II. This mutation can be distinguished by the presence of a larger restriction fragment on a Southern blot in an affected individual, compared with a shorter fragment in a normal individual.

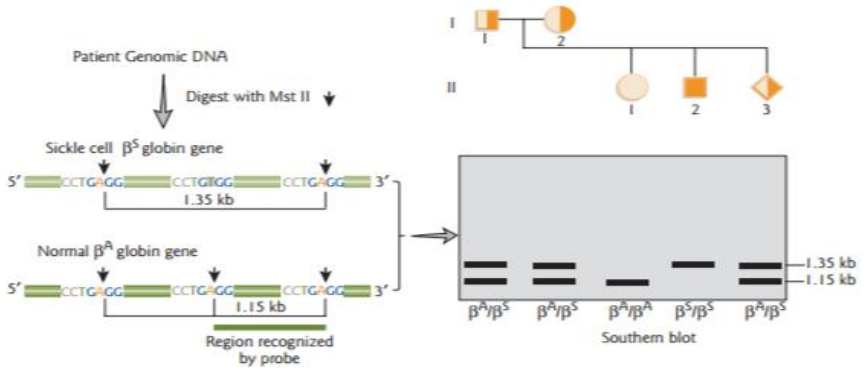


Figure 8.14 Diagnosis of sickle cell anemia by restriction fragment length polymorphism (RFLP) and Southern blot. Black arrows represent the location of recognition sites for the restriction endonuclease *Mst*II in the β-globin gene. In the mutant β-globin gene (β^S), a point mutation (GAG → GTG) has destroyed one *Mst*II recognition site. Digestion of patient genomic DNA with *Mst*II results in a 1.35 kb DNA fragment, compared with a 1.15 kb DNA fragment in normal individuals. For diagnosis, the restriction-digested DNA fragments are separated by gel electrophoresis and transferred to a nylon or nitrocellulose membrane (see Tool box 8.7). The fragments are visualized by hybridization using a probe that spans a portion of the β-globin gene where the 1.15 kb *Mst*II restriction fragment resides. In the pedigree, the family has one unaffected homozygous normal daughter (II-1), an affected homozygous son (II-2), and an unaffected heterozygous fetus (II-3). The genotypes of each family member can be read directly from the Southern blot. (Wilson, J.T., Milner, P.F., Summer, M.E., Nallaseth, F.S., Fadel, H.E., Reindollar, R.H., McDonough, P.G., Wilson, L.B. 1982. Use of restriction endonucleases for mapping the allele for β^S-globin. *Proceedings of the National Academy of Sciences USA* 79:3628–3631; and Chang, J.C., Alberti, A., Kan, Y.W. 1983. A β-thalassemia lesion abolishes the same *Mst*II site as the sickle mutation. *Nucleic Acids Research* 11:7789–7794.)

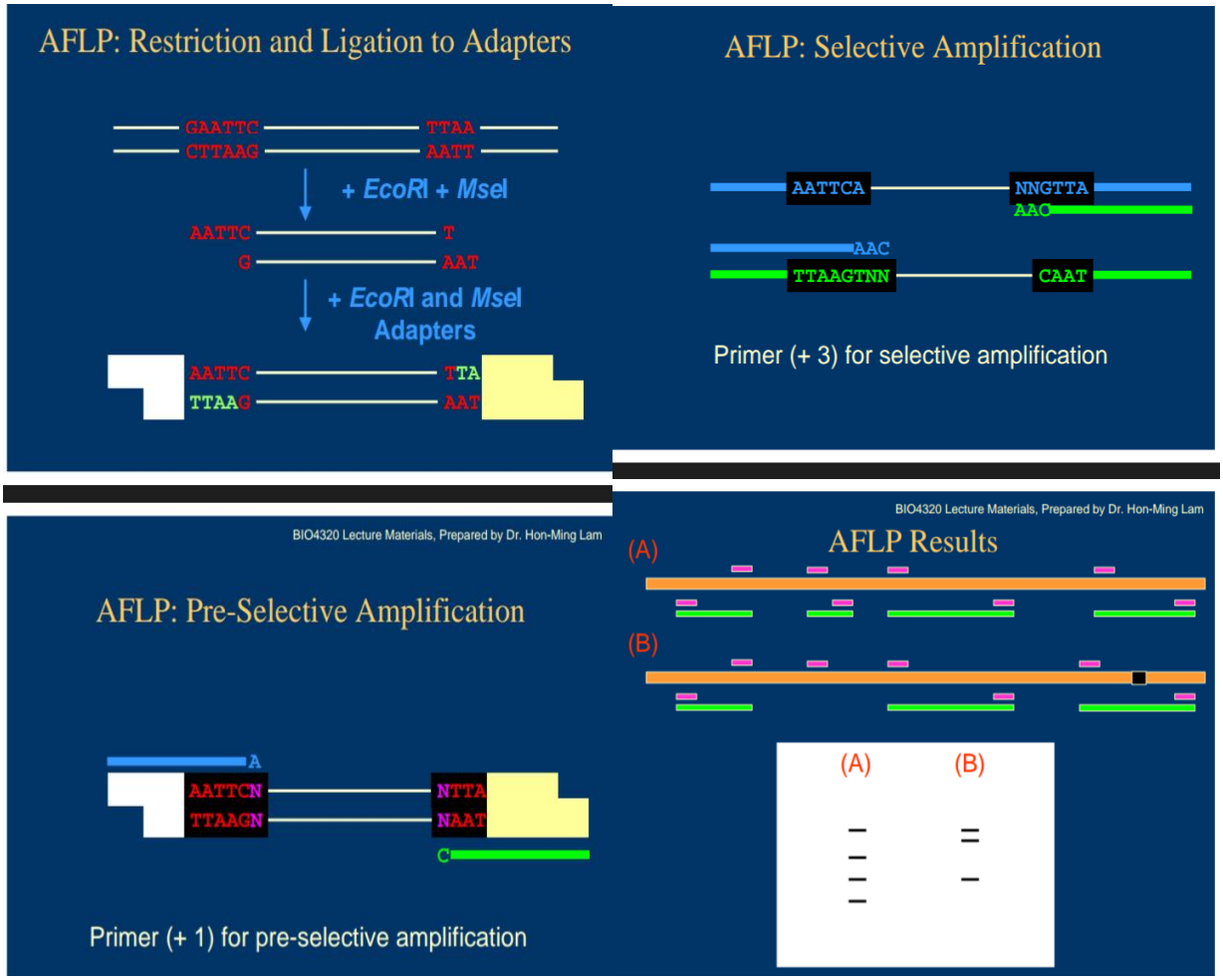
AFLP – AS GENETIC MARKER

AFLP is a PCR based tool used in genetic research, DNA fingerprinting and in practice of genetic engineering. It has been developed in 1990s by Keygene. It involve:

- Restriction endonuclease digestion of genomic DNA
- Ligation of specific adapters
- Amplification of the restriction fragments by PCR using primer pairs containing common sequences of the adapter and two or three arbitrary nucleotides and finally
- Analysis of the amplified fragments using gel electrophoresis

The amplified fragments are separated and visualized on denaturing polyacrylamide gels through autoradiography or via fluorescence methodologies or via automated capillary sequencing instruments.

The AFLP technology has the capacity to detect various polymorphisms in different genomic regions simultaneously. It is also highly sensitive and reproducible, as a result, AFLP has become widely used for the identification of genetic variation in strains or closely related species of plants, feungi, animals and bacteria



RNA SILENCING, ANTISENSE RNA AND RNA INTERFERENCE

RNA silencing and RNA interference

In higher eukaryotes, including nematodes, fruit flies, plants, and mammals, a class of small RNAs has been discovered that mediates the silencing of particular genes. The RNAs function by interacting with mRNAs, often in the 3′ UTR, results in either mRNA degradation or translation inhibition. In case, the mRNA, and thus the gene that produces it, is silenced. This form of gene regulation controls developmental timing in at least some organisms. It is also used as a mechanism to protect against invading RNA viruses (particularly important in plants, which lack an immune system) and to control the activity of transposons. In addition, small RNA molecules may play a critical (but still undefined) role in the formation of heterochromatin.

The small RNAs are sometimes called micro-RNAs (miRNAs). Many are present only transiently during development, and these are sometimes referred to as small temporal RNAs (stRNAs). Hundreds of different miRNAs have been identified in higher eukaryotes. They are transcribed as precursor RNAs about 70 nucleotides long, with internally complementary sequences that form hairpinlike structures (Fig below). The precursors are cleaved by endonucleases to form short duplexes about 20 to 25 nucleotides long. The best-characterized nuclease goes by the delightfully suggestive name Dicer; endonucleases in the Dicer family are widely distributed in higher eukaryotes. One strand of the processed miRNA is transferred to the target mRNA (or to a viral or transposon RNA), leading to inhibition of translation or degradation of the RNA (Fig.below).

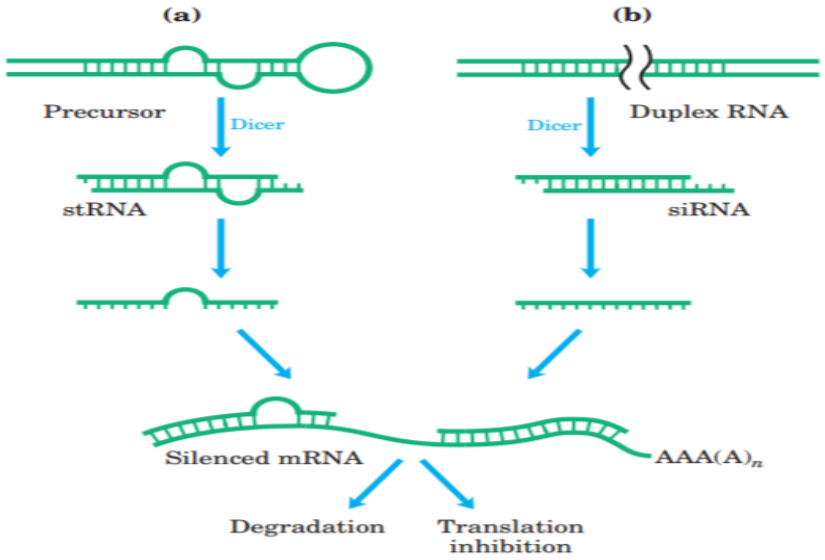


FIGURE 28–33 Gene silencing by RNA interference. **(a)** Small temporal RNAs (stRNAs) are generated by Dicer-mediated cleavage of longer precursors that fold to create duplex regions. The stRNAs then bind to mRNAs, leading to degradation of mRNA or inhibition of translation. **(b)** Double-stranded RNAs can be constructed and introduced into a cell. Dicer processes the duplex RNAs into small interfering RNAs (siRNAs), which interact with the target mRNA. Again, the mRNA is either degraded or its translation inhibited.

This gene regulation mechanism has an interesting and very useful practical side. If an investigator introduces into an organism a duplex RNA molecule corresponding in sequence to virtually any mRNA, the Dicer endonuclease cleaves the duplex into short segments, called small interfering RNAs (siRNAs). These bind to the mRNA and silence it. The process is known as **RNA interference** (**RNAi).** In plants, virtually any gene can be effectively shut down in this way. In nematodes, simply introducing the duplex RNA into the worm's diet produces very effective suppression of the target gene. The technique has rapidly become an important tool in the ongoing efforts to study gene function, because it can disrupt gene function without creating a mutant organism. The procedure can be applied to humans as well. Laboratory-produced siRNAs have already been used to block HIV and poliovirus infections in cultured human cells for a week or so at a time. Although this work is in its infancy, the rapid progress makes RNA interference a field to watch for future medical advances.

RNA interference (RNAi) is a novel phenomenon that has the potential to become an extremely powerful tool for gene silencing in any organism. The process was discovered by Fire *et al.* (1998) while investigating the use of antisense and sense RNA for gene inhibition in the nematode worm *Caenorhabditis elegans*. In one experiment, they introduced both sense and antisense RNA into worms simultaneously and observed a striking and specific inhibitory effect, which was approximately 10-fold more efficient than either single RNA strand alone.

The phenomenon of RNAi appears to be quite general, and this strategy has been used more recently for gene silencing in many other organisms, including *Drosophila*, mice and plants. In *C. elegans*, RNAi is now the standard procedure for gene inactivation, and it is becoming increasingly favoured in Drosophila and plants, due to its potency and specificity.

Like post-transcriptional gene silencing (PTGS) in plants, RNAi is also systemic. RNAi-mediated silencing can be achieved in C. elegans by injecting dsRNA into any part of the body, but, more remarkably, simply placing worms in a solution containing dsRNA or feeding them on bacteria that synthesize dsRNA is sufficient to trigger the effect. The similarities between RNAi and PTGS in plants suggest a common molecular basis.

Antisense RNA

Antisense RNA has the opposite sense to mRNA. The presence of complementary sense and antisense RNA molecules in the same cell can lead to the formation of a stable duplex, which may interfere with gene expression at the level of transcription, RNA processing or possibly translation. Antisense RNA is used as a natural mechanism to regulate gene expression in a number of prokaryote systems and, to a lesser extent, in eukaryotes. Transient inhibition of particular genes can be achieved by directly introducing antisense RNA or antisense oligonucleotides into cells. However, the transformation of cells with antisense transgenes (in which the transgene is inverted with respect to the promoter) allows the stable production of antisense RNA and thus the long-term inhibition of gene expression. This principle was established in transgenic animals and plants at about the same time. Katsuki et al. (1988) constructed an expression cassette in which the mouse myelin basic protein (MBP) cDNA was

inverted with respect to the promoter, thus producing antisense RNA directed against the endogenous gene. In some of the transgenic mice, there was up to an 80% reduction in the levels of MBP, resulting in the absence of myelin from many axons and generating a phenocopy of the myelin-depleted 'shiverer' mutation..

Antisense constructs have been widely used in transgenic animals and plants for gene inhibition. However, the efficiency of the technique varies widely and the effects can, in some cases, be non-specific. In some experiments, it has been possible to shutdown endogenous gene activity almost completely,

ETHICS AND BIOSAFETY OF DNA RECOMBINANT TECHNOLOGY

Paul Berg shared the 1980 Nobel Prize in chemistry for creating the first cloned DNA molecule, a hybrid λ phage that contained the genome of the simian tumor virus, SV40. The fact that he could do this work was worrisome to many people, included him. The recombinant DNA dispute was underway. Berg voluntarily stopped inserting tumor virus genes into phages that attack the common intestinal bacteria *E. coli.*

People continue to worry about the dangers of working with recombinant DNA. One immediate and obvious concern is that cancer or toxin genes will "escape" from the laboratory. In other words, recombinant DNA technology could create a bacterium or plasmid that contained toxin or tumor genes. The modified bacterium or plasmid could then accidentally infect people. A 1974 report by the National Academy of Sciences led to a February 1975 meeting, which took place at the Asilomar Conference Center south of San Francisco. Berg convened this meeting, which over one hundred molecular biologists attended. The recommendations of the Asilomar Committee later formed the basis for official guidelines developed by the National Institutes of Health (NIH). In essence, NIH established guidelines of containment.

Containment means erecting physical and biological barriers to the escape of dangerous organisms. The NIH guidelines defined four levels of risk, from minimal to high, and four levels of physical containment for them (called P1 through P4). The most hazardous experiments, dealing with the manipulation of tumor viruses and toxin genes, require extreme care, which included negative pressure air locks to the laboratory and experiments done in laminar-flow hoods, with filtered or incinerated exhaust air.

Current concern is focused on the acceptability of genetically modified crops (GM crops). As we will discuss later, one fourth of American cropland is planted with genetically modified crops, modified mainly for insect resistance. These modifications have curtailed our use of insecticides. (For cotton and corn, for example, liquid insecticide use dropped by 3.6 million liters and powdered insecticide by 300,000 kilograms in 1999.) However, people are concerned with the effects these modifications might have on natural ecosystems: How many valuable insects will be killed by mistake? Although Third World countries are desperate for these technologies, the United States, European and Asian trading partners are demanding that the crops we export be genetically unmodified. Farmers are also concerned that genetically modified crops have been

modified to be sterile (so called “terminator technology”) so that farmers would need to buy new seeds each year.

More recently, the recombinant DNA dispute has taken a whole new twist. It now has surfaced as a conflict between academic freedom and industrial secrecy. It seems that recombinant DNA technology is very lucrative. Numerous academic scientists have either begun genetic engineering companies or become affiliated with pharmaceutical companies. However, the philosophies of private enterprise and academia are often in conflict. Academic endeavors are presumably open, with free exchange of information among colleagues, whereas private enterprise entails some degree of secrecy, at least until patents are obtained to protect the investments of the companies. Thus, a basic conflict can arise for scientists trained in gene cloning. The conflict has been prevalent since late 1980, when the first patent for recombinant DNA techniques was awarded to Stanford University and the University of California. When, in April 2000, United States President Bill Clinton and British Prime Minister Tony Blair issued a joint statement asking that human genome data not be patented, the American stock market took a major downturn. This is a tumultuous time for biotechnology.

Genetically modified food controversies

There is controversy over GMOs, especially with regard to their use in producing food. The dispute involves buyers, biotechnology companies, governmental regulators, non-governmental organizations, and scientists. The key areas of controversy related to GMO food are whether GM food should be labeled, the role of government regulators, the effect of GM crops on health and the environment, the effect on pesticide resistance, the impact of GM crops for farmers, and the role of GM crops in feeding the world population. In 2014, sales of products which had been labeled as non-GMO grew 30 percent to $1.1 billion.

There is a general scientific agreement that food from genetically modified crops is not inherently riskier to human health than conventional food. No reports of ill effects have been proven in the human population from ingesting GM food.

Although labeling of GMO products in the marketplace is required in many countries, it is not required in the United States and no distinction between marketed GMO and non-GMO foods is recognized by the US FDA. In a May 2014 article in The Economist it was argued that, while GM foods could potentially help feed 842 million malnourished people globally, laws such as those being considered by Vermont's governor, Peter Shumlin, to require labeling of foods containing genetically modified ingredients, could have the unintended consequence of interrupting the benign process of spreading GM technologies to impoverished countries that suffer with food security problems.