Unit – 1 Recombinant DNA technology

INTRODUCTION

Genes are made up of DNA and they produce their phenotypic effects through transcription and translation. Copies of DNA are produced through semi conservative replication. The high fidelity replication of DNA ensures transmission of genes from parents to progeny without change; this is the reason for stability of genetically controlled phenotypes over generations. However, a low frequency (10^{-4} to 10^{-7} per gene per generation) of changes occur in genes naturally ***(spontaneous mutation),*** these mutations are the ultimate source of all the heritable variation observed in living forms. Clearly, *mutations create the variation that is exploited through selection during strain development.*

During evolution, genes appear to have moved across great taxonomic distances; for example, many bacterial genes have been directly integrated into the human genome. One would like to reproduce such and even more distant gene transfers in a controlled manner and at a high enough rates to be of practical application. This is achieved by the technique of recombinant DNA technology, which consists of isolation of multiple copies of a desired gene and then transferring this gene into a suitable organism (called **host).**

The pure multiple copies of desired gene can be obtained in the following three ways.

1. ***Gene cloning*** *is the basic strategy that must be used to obtain the first ever preparation of all the genes.*
2. Once the information on the sequence of a gene (at least of the regions at it's both the ends) becomes available, ***polymerase chain reaction (PCR)*** can be used to prepare copies of the genes much more easily, rapidly and cheaply.
3. When the sequence of whole gene is known, or even the amino acid sequence of its protein product is known, the ***gene could be synthesized*** either chemically or by combining PCR with chemical synthesis. Chemical synthesis of a gene provides an opportunity to effect large scale modifications in the base sequence of the gene.

RECOMBINANT DNA TECHNOLOGY: DEFINITIONS

- ***Rrecombinant DNA molecule***

It is produced by joining together two or more DNA segments usually originating from different organisms.

A vector* (e.g., ***a plasmid, phage or virus) into which the desired DNA fragment has been inserted to enable its cloning in an appropriate host.*** This is achieved by using specific enzymes for cutting the DNA *(restriction enzymes)* into suitable fragments and then for joining together the appropriate fragments *(ligation).*

- ***Chimaeric gene***

A recombinant DNA molecule, which contains a gene from one organism joined to regulatory sequences from another organism; such a gene is called ***chimaeric gene.***

Clearly, the capability to produce recombinant DNA molecules has given man the power and opportunity to create novel gene combinations to suit specific needs.

- ***Vector***

The DNA segments are integrated into an autonomously replicating DNA molecule called ***vector,*** most commonly used vectors are either bacterial plasmids or DNA viruses.

- ***Recombinant DNA technology***

All the steps concerned with piecing together DNA segments of diverse origin and placing them into a suitable vector together constitute ***recombinant DNA technology.***

- ***DNA inserts***

The vectors containing DNA segments to be cloned, called ***DNA inserts.***

- ***Host***

DNA inserts are then introduced into a suitable organism, usually a bacterium; this organism is called ***host.***

- ***Transformation***

The process of introduction of DNA inserts into suitable host is called ***transformation. T***ransformed host cells are selected and cloned.

- ***DNA cloning*** or ***gene cloning***

The step concerned with transformation of a suitable host with recombinant DNA, and cloning of the transformed cells is called ***DNA cloning*** or ***gene cloning.***

- **STEPS IN GENE CLONING**

The entire procedure of gene cloning or recombinant DNA technology may be classified into the following five steps for the convenience in description and on the basis of the chief activity performed.

1. Production and isolation of the DNA fragments to be cloned.
2. Insertion of the isolated gene in a suitable vector to obtain recombinant DNA.
3. Introduction of the recombinant DNA into a suitable organism/cell (usually *E. coli)* called *host (transformation).*
4. Selection of the transformed host cells, and identification of the clone containing the desired gene/DNA fragment.
5. Multiplication/expression of the introduced gene in the host.
6. Where needed, transfer and expression of the gene into another organism.

- **THE TOOLS USED FOR GENE CLONING**

Gene cloning utilizes certain biological products and biological agents for achieving its objectives; these may be termed as *gene cloning tools or genetic engineering tools.*

1. ***Restriction endonucleases*** to cut the DNA at specific sites; these are often called *molecular scissors.*

2. A suitable DNA molecule capable of self-replication in the selected host cell. This DNA molecule is called **vector** and the DNA to be cloned is integrated into this vector.

3. **DNA ligase** to seal the nicks that remain in the recombinant DNA molecule; it is sometimes termed as *molecular glue. In vitro,* DNA ligase joins together two DNA molecules during the production of recombinant DNA.

4. A suitable organism that serves as a ***host*** for propagation of the recombinant DNA, *i.e.*, vector containing the DNA fragment to be cloned, *viz., DNA* ***insert.***

5. ***Reverse transcriptase*** is used to produce cDNA (complementary DNA) copies of, usually, mRNA that are used for creation of cDNA libraries.

6. ***Alkaline phosphatase*** for removing 5'-phosphate from DNA ends.

7. ***T4 Polynucleatide kinase*** for addition of phosphate group to an end having a free 5'-OH.

8. ***SI nuclease*** for removal of single-stranded protrusions from ends; both 3'- and 5'-extensions are removed.

9. **The *Klenow fragment of* E. coli *DNA polymerase I*** to make the protruding ends double-stranded by extending the shorter strand.

10. ***Lambda exonuclease*** for removal of nucleotides from the 5'-ends. An ***exonuclease*** removes one nucleotide at a time from the end of a DNA molecule; it does not produce internal cuts in DNA.

11. ***E. coli exconuclease III*** for removal of nucleotides from 3'-ends.

12. ***Exonuclease Bal31*** for making DNA fragments with blunt ends shorter from both its ends.

13. ***Terminal deoxynucleotidyl transferase*** for addition of single-stranded sequences to 3'-end of blunt-ended fragments.

14. ***Linker*** and ***adapter oligonucleotide sequences*** for modification of the cut ends of DNA fragments.

==========

- **DNA manipulating enzymes**
- **RESTRICTION ENDONUCLEASES**

A **restriction enzyme** (or **restriction endonuclease**) is an enzyme that cuts double-stranded or single stranded DNA at specific recognition nucleotide sequences known as restriction sites. Such enzymes, found in bacteria and archaea, are thought to have evolved to provide a defense mechanism against invading viruses. Inside a bacterial host, the restriction enzymes selectively cut up *foreign* DNA in a process called *restriction*; host DNA is methylated by a modification enzyme (a methylase) to protect it from the restriction enzyme's activity. Collectively, these two processes form the restriction modification system. To cut the DNA, a restriction enzyme makes two incisions, once through each sugar-phosphate backbone (i.e. each strand) of the DNA double helix.

After isolating the first restriction enzyme, HindII, in 1970, and the subsequent discovery and characterization of numerous restriction endonucleases, the 1978 Nobel Prize for Physiology or Medicine was awarded to Daniel Nathans, Werner Arber, and Hamilton O. Smith. Their discovery led to the development of recombinant DNA technology that allowed, for example, the large scale production of human insulin for diabetics using *E. coli* bacteria. Over 3000 restriction enzymes have been studied in detail, and more than 600 of these are available commercially and are routinely used for DNA modification and manipulation in laboratories.

Recognition site

Restriction enzymes recognize a specific sequence of nucleotides and produce a double-stranded cut in the DNA. While recognition sequences vary between 4 and 8 nucleotides, many of them are palindromic, which correspond to nitrogenous base sequences that read the same backwards and forwards. In theory, there are two types of palindromic sequences that can be possible in DNA. The mirror-like palindrome is similar to those found in ordinary text, in which a sequence reads the same forward and backwards on the same DNA strand (i.e., single stranded) as in GTAATG. The inverted repeat palindrome is also a sequence that reads the same forward and backwards, but the forward and backward sequences are found in complementary DNA strands (i.e., double stranded) as in GTATAC (Notice that GTATAC is complementary to

CATATG). The inverted repeat is more common and has greater biological importance than the mirror-like.

5'-GTATAC-3'
::::::
3'-CATATG-5'

A palindromic recognition site reads the same on the reverse strand as it does on the forward strand

EcoRI digestion produces "sticky" ends,

G|AATTC
CTTAA|G

whereas SmaI restriction enzyme cleavage produces "blunt" ends

CCC|GGG
GGG|CCC

Recognition sequences in DNA differ for each restriction enzyme, producing differences in the length, sequence and strand orientation (5' end or the 3' end) of a sticky-end "overhang" of an enzyme restriction.

Different restriction enzymes that recognize the same sequence are known as neoschizomers. These often cleave in a different locales of the sequence; however, different enzymes that recognize and cleave in the same location are known as an isoschizomer.

Types

. There are four classes of restriction endonucleases: types I, II,III and IV based on their composition and enzyme cofactor requirements, the nature of their target sequence, and the position of their DNA cleavage site relative to the target sequence. All types of enzymes recognize specific short DNA sequences and carry out the endonucleolytic cleavage of DNA to give specific double-stranded fragments with terminal 5'-phosphates.

- Type I enzymes cleave at sites remote from recognition site; require both ATP and S-adenosyl-L-methionine to function; multifunctional protein with both restriction and methylase activities.
- Type II enzymes cleave within or at short specific distances from recognition site; most require magnesium; single function (restriction) enzymes independent of methylase.
- Type III enzymes cleave at sites a short distance from recognition site; require ATP (but doesn't hydrolyse it); S-adenosyl-L-methionine stimulates reaction but is not required; exist as part of a complex with a modification methylase.
- Type IV enzymes target methylated DNA

Type I restriction endonucleases

Type I restriction enzymes were the first to be identified and were first identified in two different strains (K-12 and B) of *E. coli*. These enzymes cut at a site that differs, and is a random distance (at least 1000 bp) away, from their recognition site. Cleavage at these random sites follows a process of DNA translocation, which shows that these enzymes are also molecular motors. The recognition site is asymmetrical and is composed of two specific portions—one containing 3–4 nucleotides, and another containing 4–5 nucleotides—separated by a non-specific spacer of about 6–8 nucleotides. These enzymes are multifunctional and are capable of both restriction and modification activities, depending upon the methylation status of the target DNA. The cofactors S-Adenosyl methionine (AdoMet), hydrolyzed adenosine triphosphate (ATP), and magnesium (Mg^{2+}) ions, are required for their full activity.

Type I restriction enzymes possess three subunits called HsdR, HsdM, and HsdS; HsdR is required for restriction; HsdM is necessary for adding methyl groups to host DNA (methyltransferase activity) and HsdS is important for specificity of the recognition (DNA-binding) site in addition to both restriction (DNA cleavage) and modification (DNA methyltransferase) activity.

Type II restriction endonucleases

Typical type II restriction enzymes differ from type I restriction enzymes in several ways. They are a dimer of only one type of subunit; their recognition sites are usually undivided and palindromic and 4–8 nucleotides in length, they

recognize and cleave DNA at the same site, and they do not use ATP or AdoMet for their activity—they usually require only Mg^{2+} as a cofactor.

These are the most commonly available and used restriction enzymes. In the 1990s and early 2000s, new enzymes from this family were discovered that did not follow all the classical criteria of this enzyme class, and new subfamily nomenclature was developed to divide this large family into subcategories based on deviations from typical characteristics of type II enzymes. These subgroups are defined using a letter suffix.

Type IIB restriction enzymes (e.g. BcgI and BplI) are multimers, containing more than one subunit. They cleave DNA on both sides of their recognition to cut out the recognition site. They require both AdoMet and Mg^{2+} cofactors.

Type IIE restriction endonucleases (e.g. NaeI) cleave DNA following interaction with two copies of their recognition sequence. One recognition site acts as the target for cleavage, while the other acts as an allosteric effector that speeds up or improves the efficiency of enzyme cleavage.

Type IIF restriction endonucleases (e.g. NgoMIV) interact with two copies of their recognition sequence but cleave both sequences at the same time.

Type IIG restriction endonucleases (Eco57I) do have a single subunit, like classical Type II restriction enzymes, but require the cofactor AdoMet to be active.

Type IIM restriction endonucleases, such as DpnI, are able to recognize and cut methylated DNA.

Type IIS restriction endonucleases (e.g. FokI) cleave DNA at a defined distance from their non-palindromic asymmetric recognition sites. These enzymes may function as dimers.

Type IIT restriction enzymes (e.g., Bpu10I and BslI) are composed of two different subunits. Some recognize palindromic sequences while others have asymmetric recognition sites.

Type III restriction endonucleases

Type III restriction enzymes (e.g. EcoP15) recognize two separate non-palindromic sequences that are inversely oriented. They cut DNA about 20-30 base pairs after the recognition site. These enzymes contain more than one subunit and require AdoMet and ATP cofactors for their roles in DNA methylation and restriction, respectively.

Artificial Restriction Enzymes

Artificial restriction enzymes can be generated by fusing a natural or engineered DNA binding domain to a nuclease domain (often the cleavage domain of the type IIS restriction enzyme FokI. Such artificial restriction enzymes can target large DNA sites (up to 36 bp) and can be engineered to bind to desired DNA sequences. Zinc finger nucleases are the most commonly used artificial restriction enzymes.

Nomenclature

Each enzyme is named after the bacterium from which it was isolated using a naming system based on bacterial genus, species and strain. For example, the name of the EcoRI restriction enzyme was derived as shown in the box.

Derivation of the EcoRI name		
Abbreviation	**Meaning**	**Description**
E	*Escherichia*	genus
co	*coli*	species
R	RY13	strain
I	First identified	order of identification in the bacterium

Applications of Restriction endonucleases

Isolated restriction enzymes are used to manipulate DNA for different scientific applications.

i. They are used to assist insertion of genes into plasmid vectors during gene cloning and protein expression experiments. To clone a gene fragment into a vector, both plasmid DNA and gene insert are typically cut with the same restriction enzymes, and then glued together with the assistance of an enzyme known as a DNA ligase.
ii. Restriction enzymes can also be used to distinguish gene alleles by specifically recognizing single base changes in DNA known as single nucleotide polymorphisms.
iii. Restriction enzymes are used to digest genomic DNA for gene analysis by Southern blot.

Examples of Restriction enzymes

Enzyme	**Source**	**Recognition Sequence**	**Cut**
EcoRI	*Escherichia coli*	5'GAATTC 3'CTTAAG	5'---G AATTC---3' 3'---CTTAA G---5'
EcoRII	*Escherichia coli*	5'CCWGG 3'GGWCC	5'--- CCWGG---3' 3'---GGWCC ---5'
BamHI	*Bacillus amyloliquefaciens*	5'GGATCC 3'CCTAGG	5'---G GATCC---3' 3'---CCTAG G---5'
HindIII	*Haemophilus influenzae*	5'AAGCTT 3'TTCGAA	5'---A AGCTT---3' 3'---TTCGA A---5'
TaqI	*Thermus aquaticus*	5'TCGA 3'AGCT	5'---T CGA---3' 3'---AGC T---5'
PovII*	*Proteus vulgaris*	5'CAGCTG 3'GTCGAC	5'---CAG CTG---3' 3'---GTC GAC---5'
SmaI*	*Serratia marcescens*	5'CCCGGG 3'GGGCCC	5'---CCC GGG---3' 3'---GGG CCC---5'
HaeIII*	*Haemophilus aegyptius*	5'GGCC 3'CCGG	5'---GG CC---3' 3'---CC GG---5'

- **DNA ligase**

Mertz and Davis (1972) for the first time demonstrated that cohesive termini of cleaved DNA molecules could be covalently sealed with *E. coli* DNA ligase and were able to produce recombinant DNA molecules (Fig. 4.3). DNA ligase seals single strand nicks in DNA which has 5' ⟶ 3' -OH (hydroxyl) termini. There are two enzymes which are extensively used for covalently joining restriction fragments: the ligase from *E. coli* and that encoded by T4 phage. The main source of DNA ligase is T4 phage, hence, the enzyme is known as T4 DNA ligase.

For the joining reactions, the *E. coli* DNA ligase uses nicotinamide adenine dinucleotide (NAD^+) as a cofactor, while T4 DNA ligase requires ATP for the same. Both the enzymes contain a $-NH_2$ group on lysine residue. In both cases, cofactor breaks into AMP (adenosine monophosphate) (Fig. 4.3.) which in turn adenylate the enzyme (E) to form enzyme -AMP complex (EAC). EAC binds to nick containing 3' -OH and 5' $-PO_4$ ends on a double stranded DNA molecule. The 5' -phosphoryl terminus of the nick is adenylated by the EAC with 3'-OH terminus resulting in formation of phosphodiester and liberation of AMP (Lehman, 1974). After formation of phosphodiester nick is sealed (Fig. 4.3). T4 enzyme has the ability to join the blunt ends of DNA fragments, whereas *E. coli* DNA ligase joins the cohesive ends produced by restriction enzymes. Additional advantage with T4 enzymes is that it can quickly join and produce the full base pairs but it would be difficult to retrieve the inserted DNA from vector. However, cohesive end ligation proceeds about 100 times faster than the blunt end ligation.

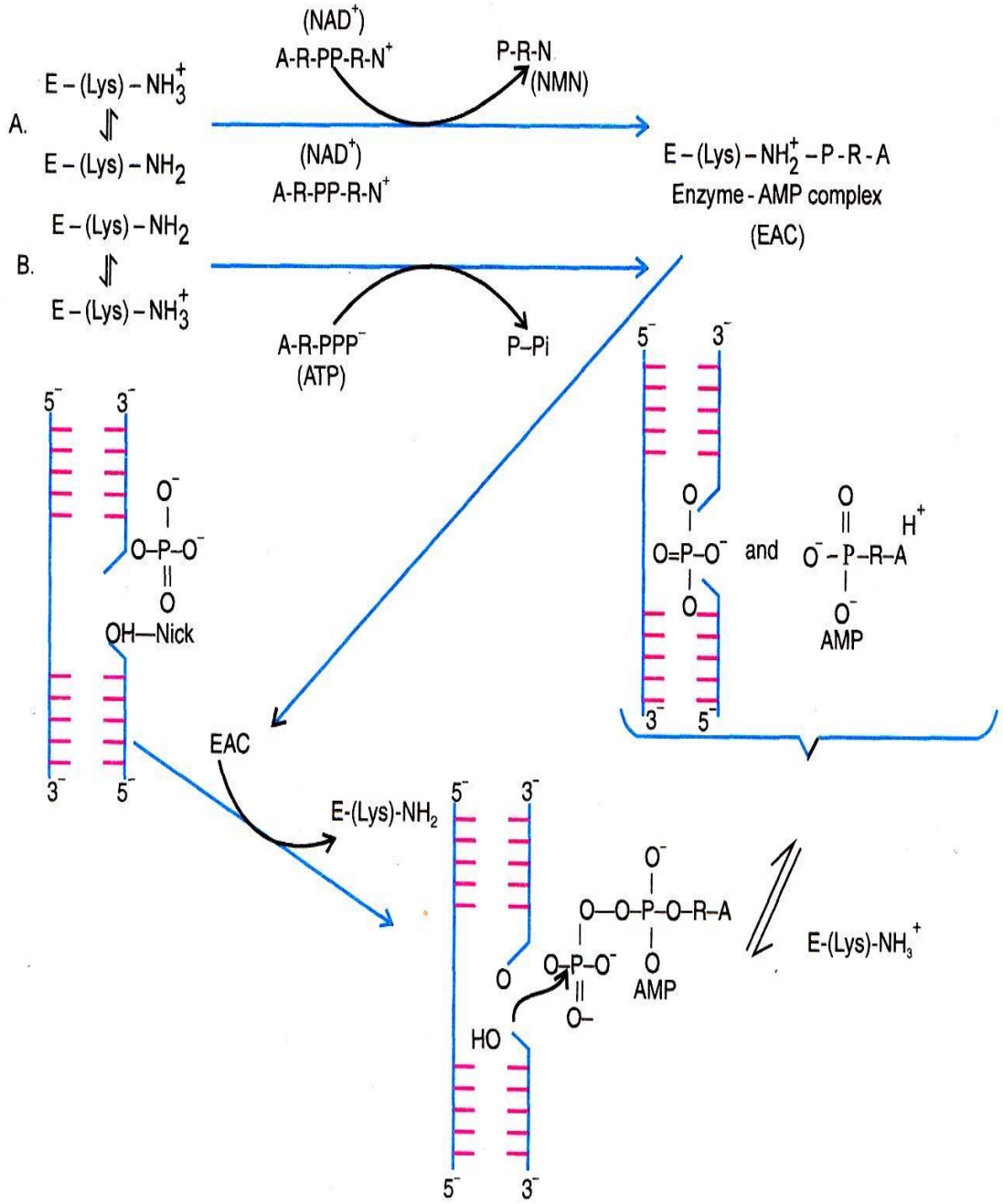

Fig 4.3. Mechanism of DNA ligase enzyme - AMP complex formation and sealing of nick on a double stranded DNA molecule; A, *E. coli* DNA ligase; B, T4 DNA ligase; A, adenine; P, PO_4; R, ribose; N, nicotinamide; NMN, nicotinamide mononucleotide

- **Alkaline phosphatases**

When plasmid vector, for joining a foreign DNA fragment, is treated with restriction enzyme, the major difficulty arises at the same time. Because the cohesive ends of broken plasmids, instead of joining with foreign DNA join the cohesive end of the same DNA molecules and get recircularized. To overcome

this problem, the restricted plasmid *(i.e.* plasmid treated with restriction enzymes) is treated with an enzyme, alkaline phosphatase that digests the terminal 5' phosphoryl group (Fig. 4.4). The restriction fragments of the foreign DNA to be cloned are not treated with alkaline phosphatase. Therefore, the 5' end of foreign DNA fragment can covalently join to 3' end of the plasmid. The hybrid or recombinant DNA obtained has a nick with 3' and 5' hydroxy ends. Ligase will only join 3' and 5' ends of recombinant DNA together if the 5' end is phosphorylated. Thus, alkaline phosphatase and ligase prevent recircularization of the vector and increase the frequency of production of recombinant DNA molecules. The nicks between two 3' ends of DNA fragment and vector DNA are repaired inside the bacterial host cells during the transformation.
The mechanism of action of alkaline phosphatase is as below:

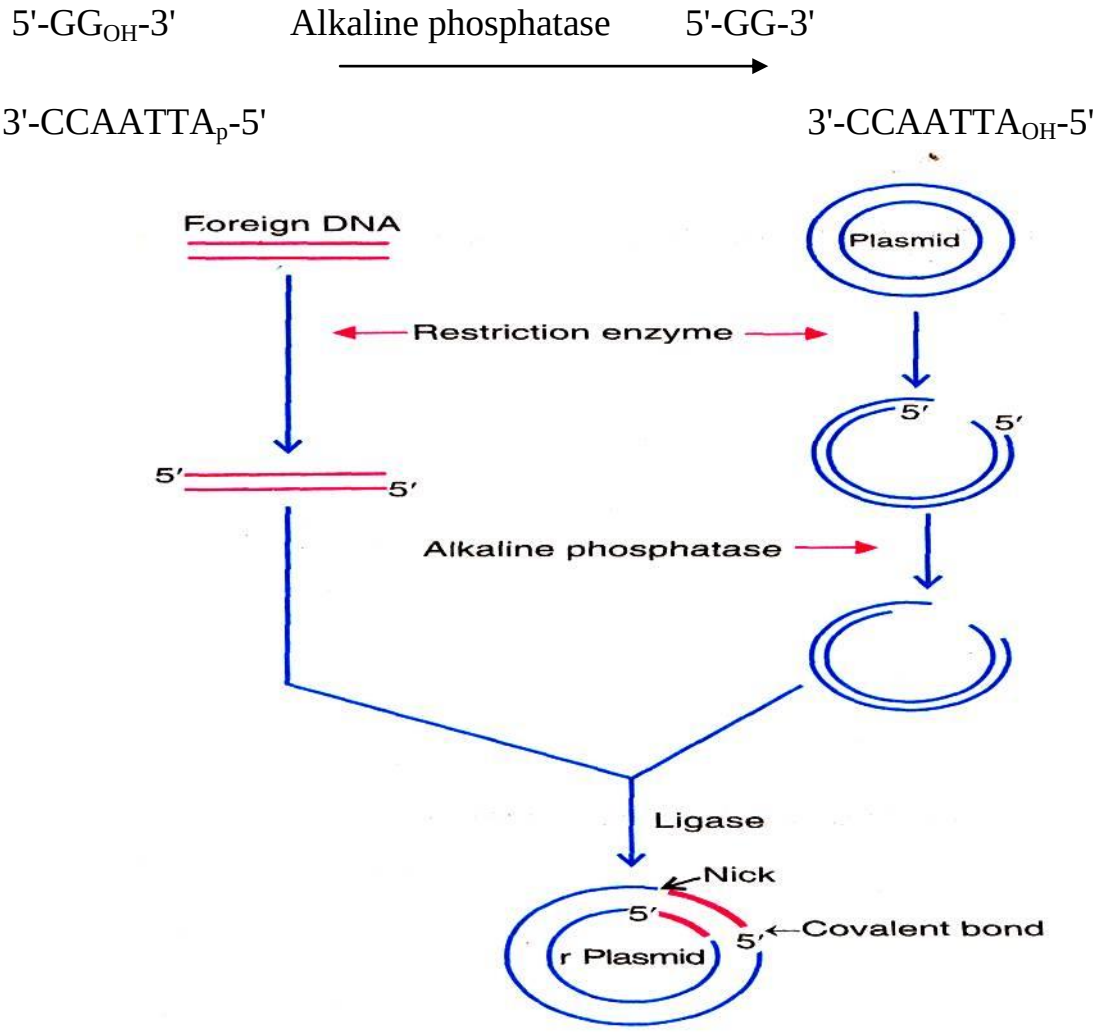

Fig. 4.4. Inhibition of recircularization by alkaline phosphates (to increase recombinant plasmids)

- **Polynucleotide Kinase**

The T4 polynucleotide kinase catalyses the transfer of $\gamma^{32}PO_4$ of ATP to 5' terminus of a dephosphorylated DNA (i.e. DNA lacking -PO_4 residue) or RNA as below:

$$\begin{array}{lcl} 5'\text{-}GG_{OH}\text{-}3' & \text{Polynucleotide kinase} & 5'GG_{OH}\text{-}3' \\ 3'\text{-}CCTTAA_{OH}\text{-}5' & \xrightarrow[\gamma^{32}P\text{-}ATP]{} & 3'\text{-}CCTTAA_{P^*}\text{-}5' \end{array}$$

This enzyme shows both phosphorylation as well as phosphatase activities; therefore, two types of reaction can be used accordingly. In phosphorylation activity the γ-PO_4 is transferred to the 5'end of dephosphorylated DNA. When ADP is in excess, it causes the T4 polynucleotide kinase to transfer the terminal phosphate (attached to a nucleotide) from phosphorylated DNA to ADP. By using radiolabelled gamma-phosphate ($\gamma^{32}PO_4$), it was found that DNA is rephosphorylated by transfer of $\gamma^{32}P$ from $\gamma^{32}P$-ATP. This shows that the enzyme plays a dual role of phosphorylation activity and phosphatase activity.

Due to the presence of both properties, this enzyme is used in radiolabelling of the 5' end of duplex during DNA sequencing (*see* DNA sequencing method), constructing the terminally labelled DNA, phosphorylating the synthetic linkers/oligonucleotide linkers/other DNA fragments lacking terminal 5'phosphates, etc.

- **Reverse Transcriptase**

In addition to these enzymes, reverse transcriptase is used to synthesize the copy DNA or complementary DNA (cDNA) by using mRNA as a template. Reverse transcriptase is very useful in the synthesis of cDNA and construction of cDNA clone bank.

Until recently, it was known that the genetic information of DNA pass to protein through mRNA. During 1960s, Temin and co-workers postulated that in certain cancer causing animal viruses which contain RNA as genetic material, transcription of cancerous genes (on RNA into DNA) takes place most probably by DNA polymerase directed by virus RNA. Then DNA is used as template for synthesis of many copies of viral RNA in a cell. In 1970, S. Mizutani, H.M. Temin and D. Baltimore discovered that information can also pass back from RNA to DNA. They found that retroviruses (possessing RNA) contain RNA dependent DNA polymerase which is also called as reverse transcriptase. This produces single stranded DNA, which in turn functions as template for complementary long chain of DNA.

- **Formation of Recombinant DNA**

Most enzymes produce ***staggered cuts*** in which the two strands of a DNA double helix are cleaved at different locations; this generates ***protruding*** (3′ or 5′) ***ends*** (Fig. 2.3), *i.e.*, one strand of the double helix extends some bases beyond the other. Due to the palindromic (symmetrical) nature of the target sites, the two protruding ends generated by such a cleavage by a given enzyme have complementary base sequence. As a result, they readily pair with each other under annealing conditions; such ends are called ***cohesive*** or ***sticky ends***. An important consequence of this fact is that when fragments generated by a single restriction enzyme from different DNAs are mixed, they join together due to their sticky ends (Fig. 2.4). Therefore, this property of the restriction enzymes is of great value for the construction of recombinant DNAs.

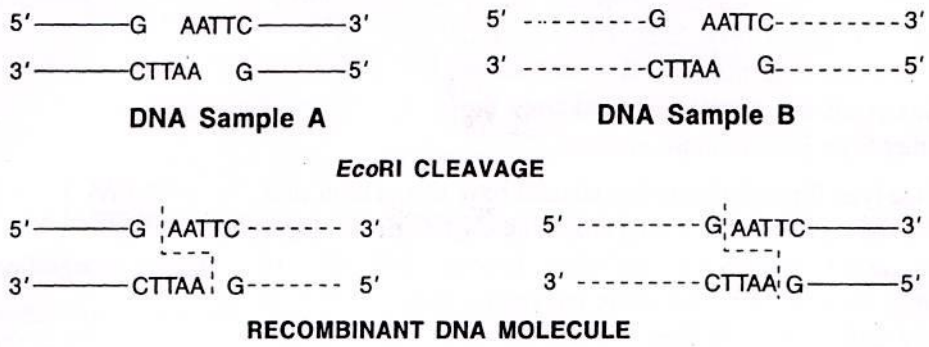

Fig. 2.4. Two distinct samples (A and B) of DNA are cleaved with the same restriction enzyme (*Eco*RI). The fragments from sample A readily join with those from sample B due to their cohesive (or sticky = complementary) protruding ends. The dotted line within the recognition sequences separates the base sequences belonging to the different fragments.

In addition, two or more restriction enzymes having different recognition sites generate the same sticky end as is depicted in Fig 2.5 for *Bam*HI, *Bgl*II and *Sau*3A; they all produce GATC sticky ends.

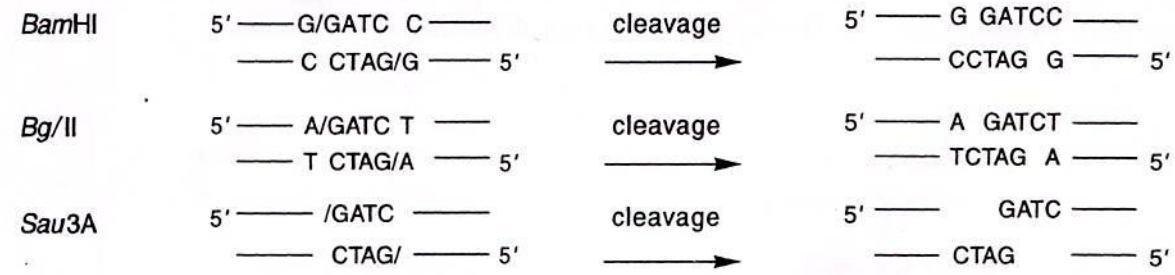

Fig. 2.5. Different restriction enzymes may generate the same sticky end, *e.g.*, 5′GATC3′ in this case.

Therefore, the sticky ends produced by these three enzymes will be complementary with each other, and can be regarded as having been produced by a single enzyme for the purposes of joining together of the different DNA fragments. The same will be true for any other group of such enzymes that generate the same sticky ends.

There are cases where two different restriction enzymes recognize the same target sequence, but one of them is able to recognize both methylated as well non-methylated target sequences, while the other enzyme can recognize only the non-methylated target sequence; such enzymes are known as ***isoschizomers***. For example, restriction enzymes *Hpa*II and *Msp*I are isoschizomers; they both recognize the sequence 5′CCGG3′ when it is unmethylated. But when the second C of the sequence is methylated, *Hpa*II can no longer recognize it, while *Msp*I recognizes it just as well as it does the unmethylated sequence. *Isoschizomers are very useful in determining the state of methylation of a DNA molecule.*

There are Several Methods for Joining DNA Molecules

It is not difficult to see that two DNA molecules cut by the same restriction enzyme producing cohesive ends will stick to each other by virtue of complementarity. Once this has happened, they can be sealed using T4 DNA ligase. However, more than one product can arise when we mix two different DNA molecules cut by the same restriction enzyme. The vector can circularize again by virtue of cohesive ends, more than one fragment of foreign DNA can stick together, and, of course, the desired recombinant DNA can also form (Figure 17.2). Out of all these reactions, the one leading to circularization of the vector is kinetically favored and, as a result, many of the vectors reseal without any insert DNA.

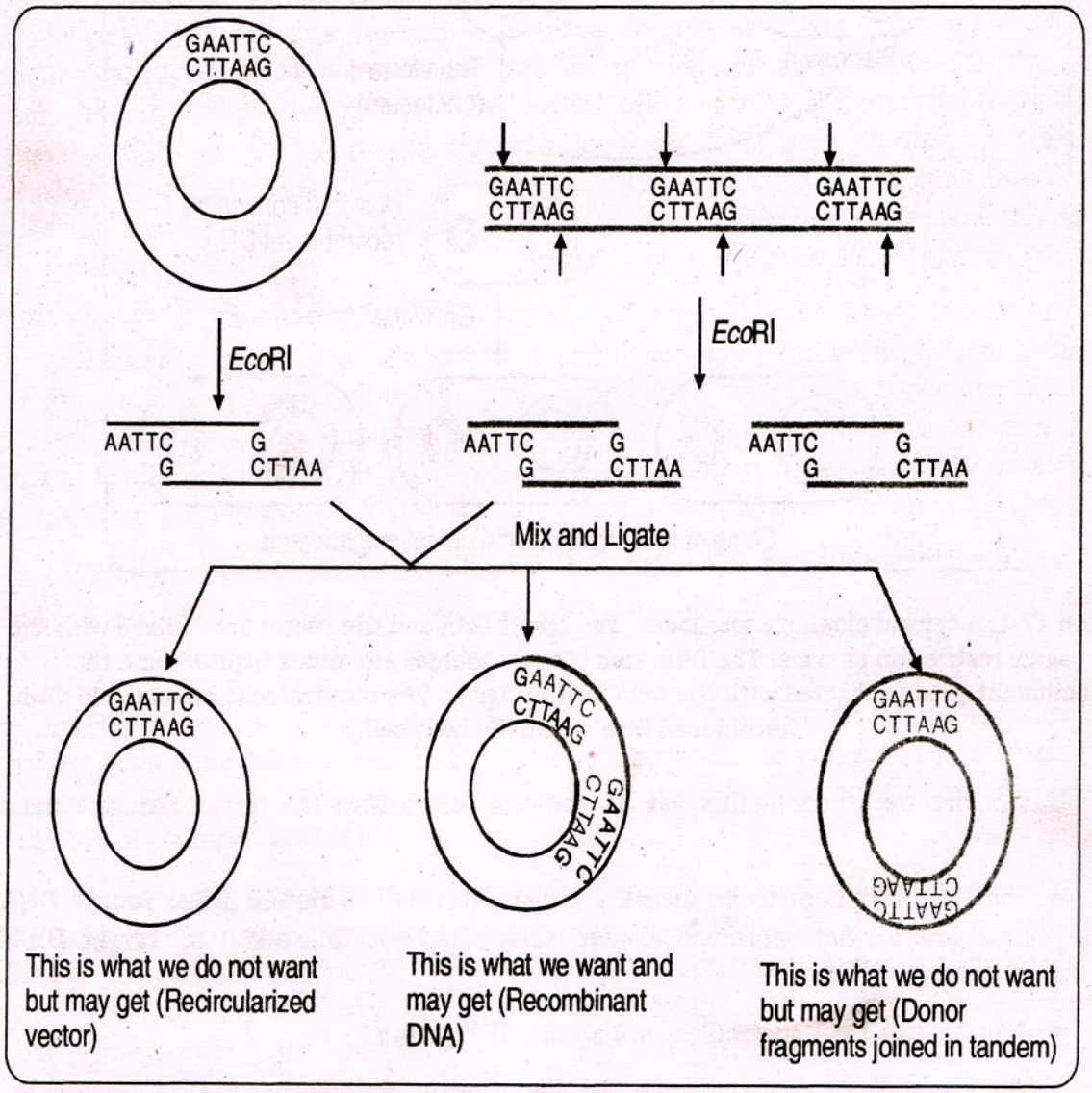

Figure 17.2. When you mix the vector and foreign DNA cut with the same restriction enzyme, recombinant DNA is not the only product formed. The vector can circularize without any insert DNA and several fragments of the foreign DNA can associate end to end.

Can we avoid this problem? Can we disallow the reaction leading to circularization of the vector? This is possible if we cut the vector with restriction enzyme and then, before mixing it with foreign DNA, treat the vector with alkaline phosphatase (Figure 21.3). When we do this, we remove the 5′-phosphates from the ends of the vector. If you recall, DNA ligases can seal only if they have a 3′-OH and a 5′-phosphate available. Since we have removed the 5′-phosphate from the vector, the vector cannot circularize. Now if we mix this vector with the foreign DNA, only the foreign DNA can supply the 5′-phosphate and thus the step leading to insertion of this DNA into the vector becomes kinetically favored (Figure 17.3). Thus, by simply removing the 5′-phosphates from the vector, we create a situation favorable for formation of recombinant DNA. One must note here that not all 5′-phosphate groups will be removed from the vector - some will escape enzyme action. Therefore, by this treatment we cannot completely avoid circularization of the vector.

Another question can bug you here. You may have visualized the situation that the vector that circularizes after taking in a target DNA fragment will have only two ligations and two nicks will still remain (Figure 17.3). But rest assured; this molecule is stable enough and can be introduced into a host cell where the host ligase system corrects the defect and seals the nicks.

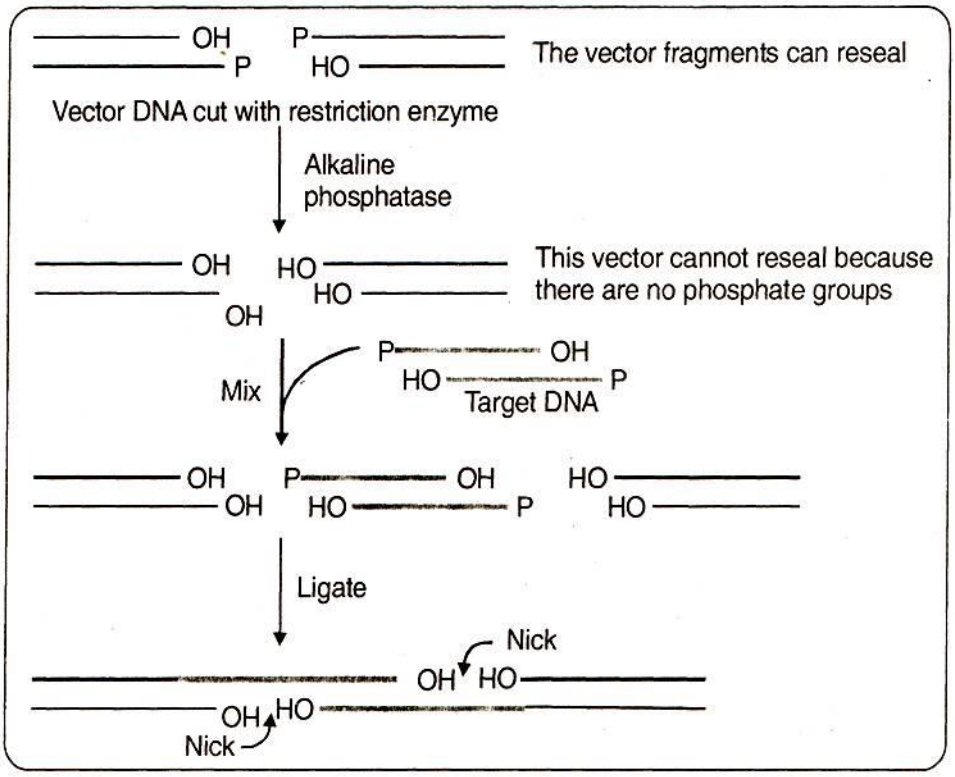

Figure 17.3. Treating the vector with phosphatase before mixing it with foreign DNA promotes formation of recombinant DNA. Mark that the recombinant DNA formed after phosphatase treatment still has two nicks.

Joining of vector and insert DNA generated by cohesive end producing restriction enzymes is self-suggestive. But we know of restriction enzymes which cut in a manner that blunt ends are produced. These enzymes are also used for generating recombinant DNA. How do we join two DNA molecules which have blunt ends?

How do You Ligate Blunt Ends? 1. Doing it Straight

The simplest way to join two blunt-ended DNA molecules is to mix them together in the presence of T4 DNA ligase. This enzyme normally seals nicks in double-stranded DNA molecules having 3'-OH and 5'-P termini. But it can also join two DNA molecules having completely base-paired ends (Figure 17.4). Exactly how T4 DNA ligase achieves this is not known.

The ligation efficiency of this reaction is low because there are no sticky ends to hold the two DNA molecules temporarily till the nicks are sealed by ligase. And if the DNA concentration is low, the yield of recombinant DNA is even lower because ligation of the two ends of the same DNA molecule is the kinetically favored reaction (Figure 17.4). But if DNA concentration is made high, this problem can be minimized and formation of recombinant DNA promoted (Figure 17.4).

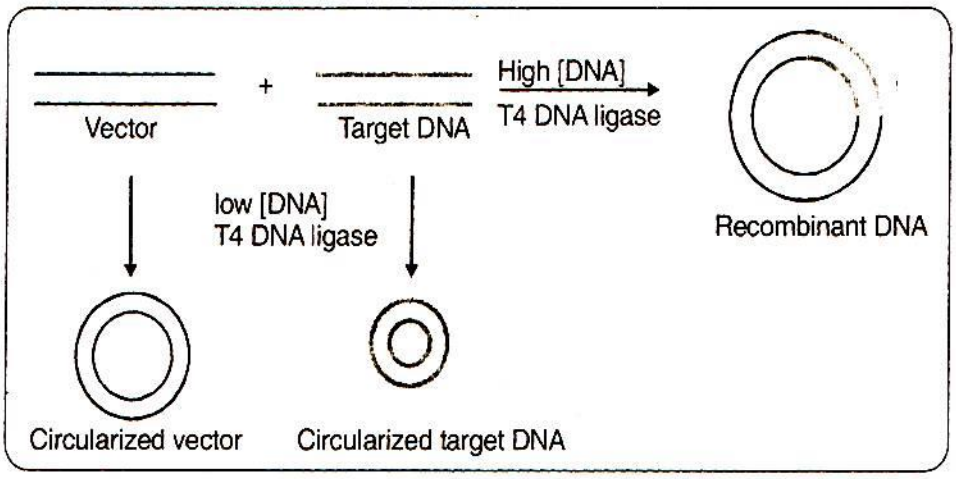

Figure 17.4. Ligation of two blunt ended DNA molecules. Low concentration of DNA promotes intramolecular circularization. High DNA concentration promotes formation of recombinant DNA molecule.

When the concentration of DNA is increased, it should be done in a manner that both vector and target DNA are in equimolar concentration. This minimizes the chances of formation of recombinants with tandem copies of the insert. On the other hand, when the concentration of the target DNA is limiting, the ratio of vector to foreign DNA can be increased to maximize its incorporation into the vector. Ligation reactions are generally carried out at lower temperatures (8-15° C) for extended time periods (2-24 hrs).

How do You Ligate Blunt Ends? 2. Homopolymer Tail Joining

Remember terminal deoxynucleotide transferase? This is an unusual enzyme in that it adds nucleotides to the 3'-OH group of an extended single-stranded segment of DNA. This enzyme has no need of a template and will add nucleotides depending on the concentration of dNTPs in the medium. And if you provide only one dNTP in the medium, it will go on adding that nucleotide till it creates a homopolymer tail at the 3'-end. Taking advantage of this property, we can see to it that (1) the problems associated with blunt end joining are completely removed and that (2) intramolecular ligation cannot take place and therefore only recombinants can form.

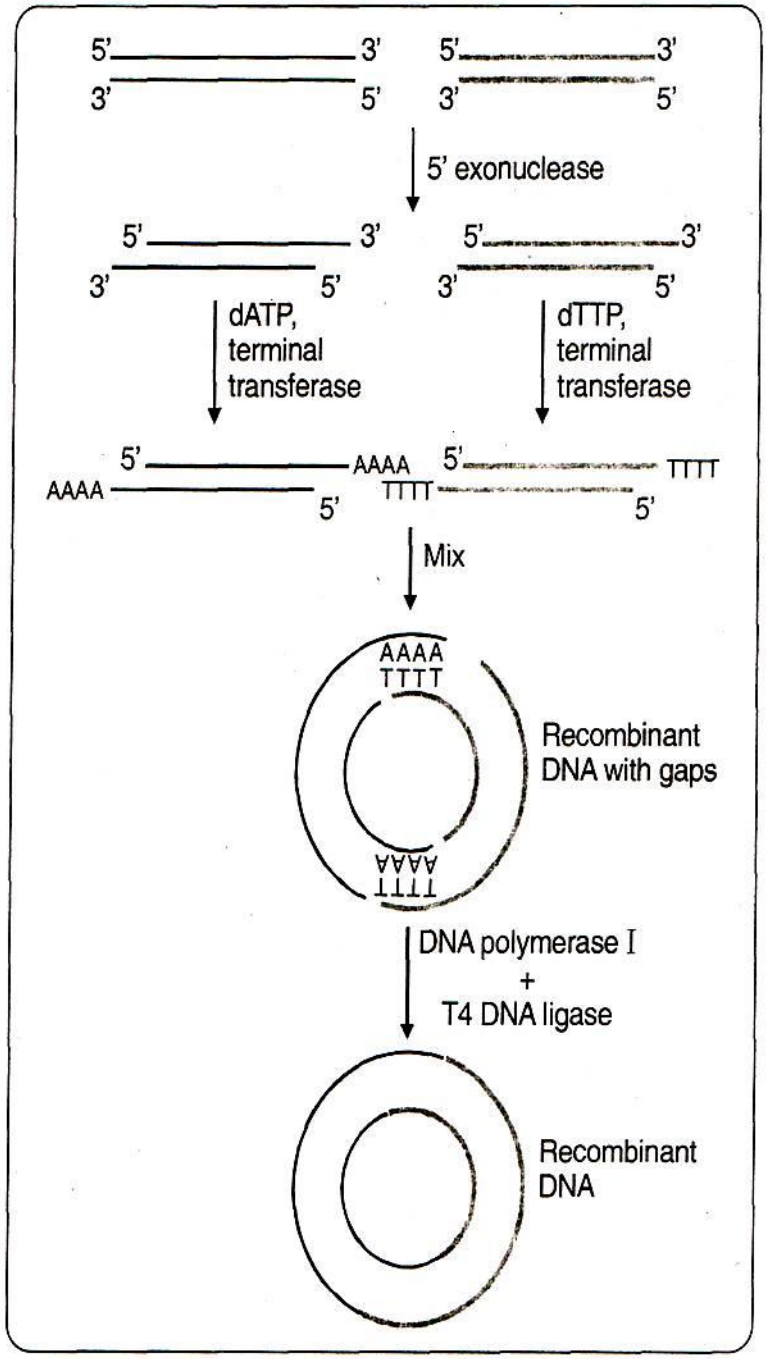

Figure 17.5. Homopolymer tail joining. Note that the vector has a poly(A) tail at the 3′-end and the target DNA has a poly(T) tail. This complementarity promotes hydrogen bond formation between the two DNA molecules. Subsequent treatment with DNA polymerase I fills the gap that exists between the two DNA molecules and ligase then joins the DNA molecules covalently.

The first thing is that since terminal transferase can work with only extended single-stranded 3′-OH termini, we have to create these termini in the blunt-ended DNA molecules. This is easy - we can treat the DNA molecules with a 5′-specific exonuclease to remove a few terminal nucleotides. This will generate 3′-OH overhangs. Now we can mix these DNA molecules with dATP and terminal transferase. This will give rise to a poly(A) tail at the 3′ ends. Supposing this is the treatment we gave to the vector DNA. Now we can take the target DNA, give it 5′-exonuclease treatment and then provide

dTTP and terminal transferase in the medium. This will result in the generation of a poly(T) tail at the 3′ ends. Now if we mix the vector and the target DNA together, they will stick to each other by virtue of complementarity. Note that there will be a gap between the two DNA molecules. This gap is because of your first action of 5′-exonuclease digestion. This gap is now filled by DNA polymerase I. Subsequent reaction with DNA ligase can then seal the two DNA molecules covalently. *What we have achieved here is an artificial creation of sticky ends between the vector and the target DNA.* But these sticky ends are better than the ones created with restriction enzymes - note that the vector cannot circularize all by itself because its termini are not complementary to each other; for the same reason tandem copies of the target DNA cannot form. In one stroke, the problem presented by blunt ends has been turned to an advantageous situation! The entire procedure of homopolymer tail joining is presented in Figure 17.5.

There is one disadvantage of such joining as compared to joining of restriction digested DNA molecules. It is easy to remove the insert from a construct created through joining restriction digested DNAs. All you have to do is to treat the construct with the same restriction enzyme again and, bingo, you get the insert and vector separated. With homopolymer tail mediated recombinant DNA, such a scheme is not possible.

Homopolymer tail joining can not only be used for joining blunt ends created by restriction digestion. We can also use the technique for joining molecules that have been created by mechanical shearing or by random endonuclease treatment. A very important use of this technique is in cloning cDNA. We have already seen what cDNA is and how it is produced starting from RNA (see DNA Technology). Shortly, we will also consider cDNA cloning and see just how important it is. cDNA molecules also have blunt ends and homopolymer tail joining is one of the most important techniques used to clone them.

Homopolymer tail joining suffers from the disadvantage that the recovery of the cloned fragment is not very straightforward. Is there any method of cloning blunt-ended DNA molecules where the recovery of the target DNA is easy? Let's see.

How do you Ligate Blunt Ends? 3. Use of Linkers and Adaptors

Oligonucleotide synthesis is now very easy. With the development of this technology, the use of oligonucleotides has become quite widespread. We can deliberately synthesize oligonucleotides which contain a restriction site. Supposing we synthesize a 6-mer GAATTC. Now this is the *Eco*RI recognition sequence. What if we mix this oligonucleotide with blunt ended fragments of the target DNA? What may happen is that these target DNA fragments will be ligated to these synthetic oligonucleotides at both their ends (Figure 17.6). Once this has happened, you have an *Eco*RI sequence at both the ends of the blunt-ended DNA fragment. Now we cut the vector and this modified target DNA with *Eco*RI, we will have sticky ends protruding at the ends of the previously blunt-ended target DNA and at the ends of the vector. These two DNA molecules can stick to each other by virtue of complementarity. You have your recombinant DNA molecule and, most importantly, if you want to excise your target DNA from the construct, all you have to do is treat the construct with the same restriction enzyme whose site was inbuilt into the oligo. Such oligonucleotides are called **linkers**. One thing you may have thought about - the target DNA should not have a restriction endonuclease site which is the same as the linker sequence. If this is not so, the procedure will not be effective for obvious reasons.

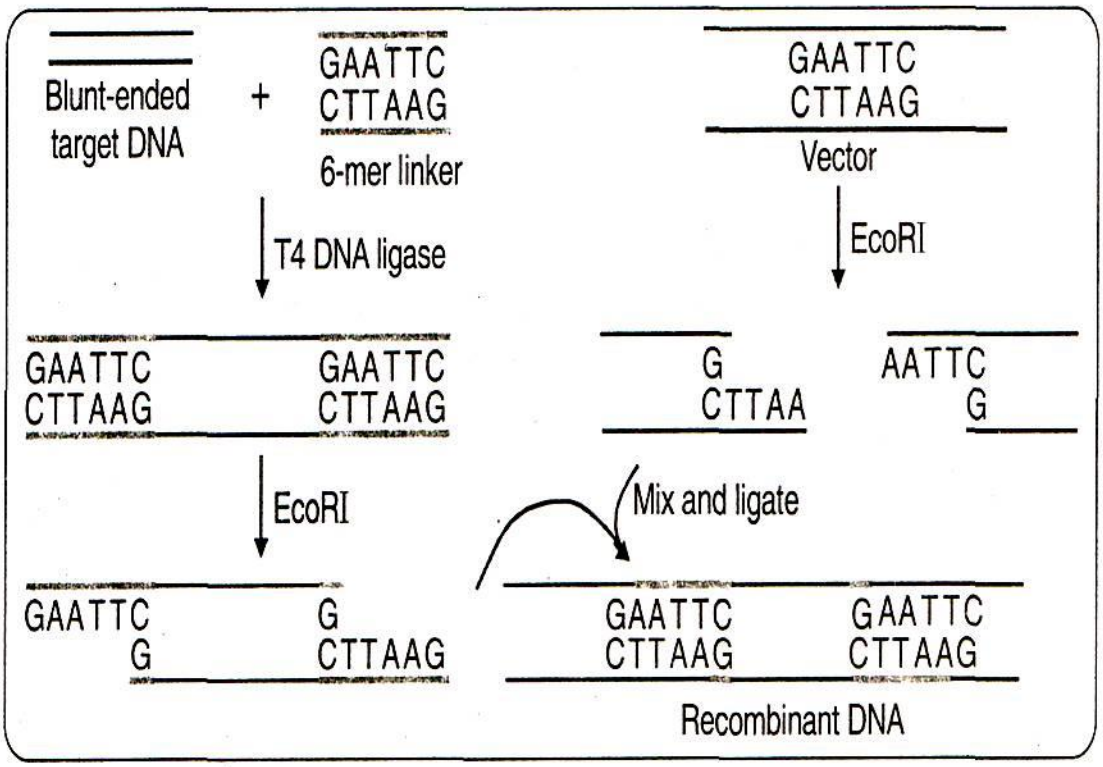

Figure 17.6. Using a linker to generate recombinant DNA.

We can design the linkers slightly differently. Figure 17.7 depicts an oligonucleotide with two restriction sites built into it - *Eco*RI and *Sma*I. Let's see what can be done now. Digest both the vector and the short DNA segment with *Eco*RI. Once this is done, we can ligate the oligonucleotide to the vector. Once this is done, we can take any blunt-ended DNA molecule and recombine it with the vector using now the *Sma*I site (this enzyme gives rise to blunt ends). Later, if we want to excise the cloned fragment, we can easily do it using *Eco*RI. Figure 17.7 depicts the whole scheme. The requirement that is absolute here is that neither the vector nor the target DNA should have the *Eco*RI site. Additionally, the vector should not have *Sma*I site.

Again, some of you may have run ahead of the text and thought of short DNA stretches into which several restriction sites could be built so that it functions as a polylinker and could be used for several occasions. If you have thought so, it is correct. Such polylinkers not only are available, but are usually built into the commercially available vectors. Shortly, we will be discussing such vectors.

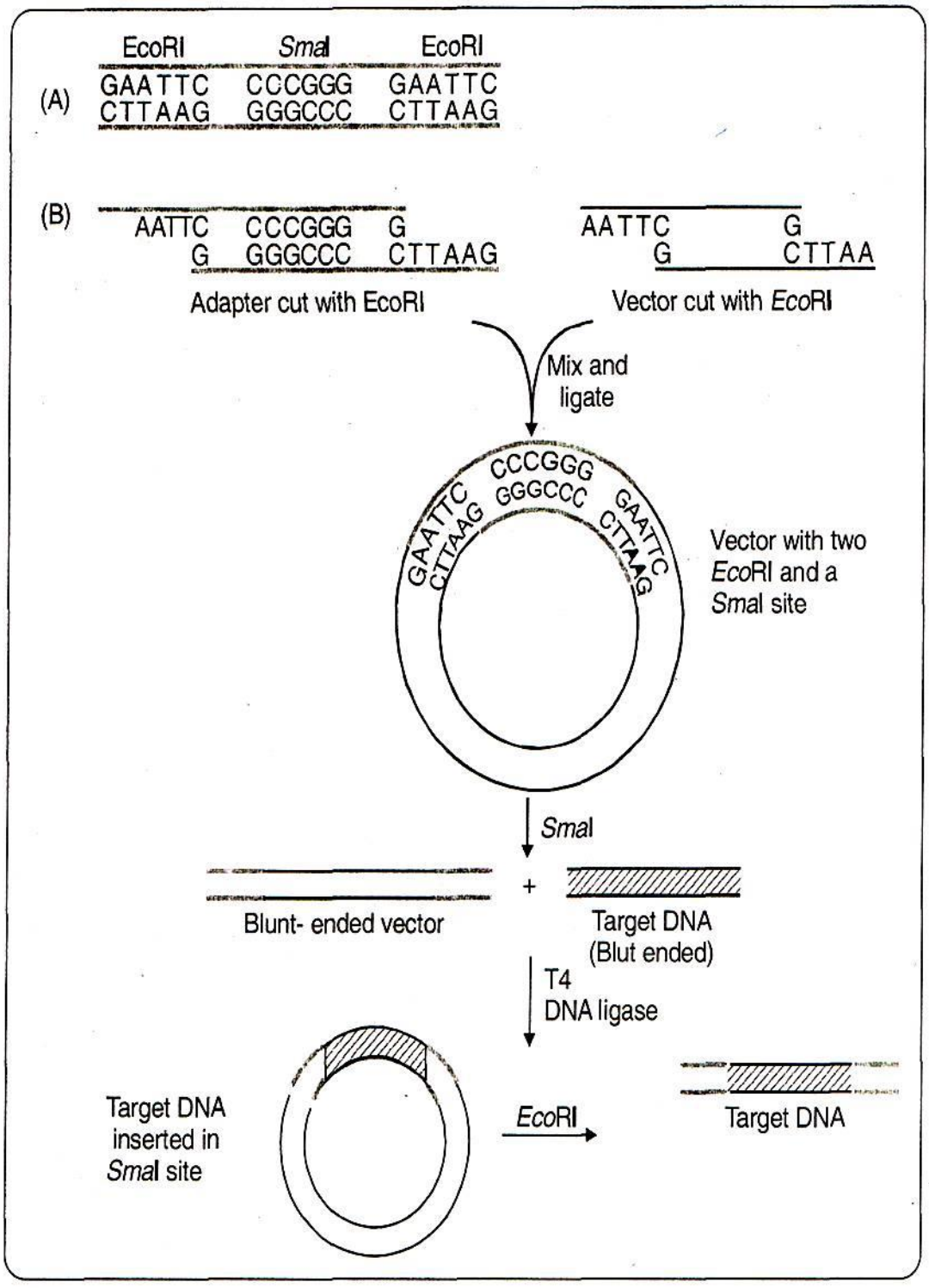

Figure 17.7. Adaptors have more than one restriction sites built into them. (A) The adaptor in the diagram has an *Eco*RI and a SmaI site. (B) One example of the use of an adaptor. Cut vector DNA and the adaptor using *Eco*RI. This gives rise to cohesive ends on both DNA molecules. Mix and ligate the two. Now cut the vector again using *Sma*I. This linearizes the vector and the two ends are blunt. Take a blunt-ended DNA fragment which you want to clone and ligate it with the vector using T4 DNA ligase. Later, when you want to excise the cloned fragment, you can do so by using *Eco*RI.