

## **DNA MODIFYING ENZYMES**

Restriction enzymes and DNA ligases represent the cutting and joining functions in DNA manipulation. All other enzymes involved in genetic engineering fall under the broad category of enzymes known as DNA modifying enzymes. These enzymes are involved in the degradation, synthesis and alteration of the nucleic acids.

**DNA ligase** is an important cellular enzyme, as its function is to repair broken phosphodiester bonds that may occur at random or as a consequence of DNA replication or recombination. In genetic engineering it is used to seal discontinuities in the sugar—phosphate chains that arise when recombinant DNA is made by joining DNA molecules from different sources. It can therefore be thought of as molecular glue, which is used to stick pieces of DNA together. This function is crucial to the success of many experiments, and DNA ligase is therefore a key enzyme in genetic engineering.

The enzyme used most often in experiments is T4 DNA ligase, which is purified from *E. coli* cells infected with bacteriophage T4. Although the enzyme is most efficient when sealing gaps in fragments that are held together by cohesive ends, it will also join blunt-ended DNA molecules together under appropriate conditions. T4 DNA Ligase catalyzes the formation of a phosphodiester bond between juxtaposed 5'-phosphate and 3'-hydroxyl termini in duplex DNA or RNA. The enzyme repairs single-strand nicks in duplex DNA, RNA, or DNA/RNA hybrids. It also joins DNA fragments with either cohesive or blunt termini, but has no activity on single-stranded nucleic acids. The T4 DNA Ligase requires ATP as a cofactor.

## **RESTRICTION ENDONUCLEASES**

- Also called restriction enzymes
- 1962: “molecular scissors” discovered in bacteria
- *E. coli* bacteria have an enzymatic immune system that recognizes and destroys foreign DNA
- 3,000 enzymes have been identified, around 200 have unique properties, many are purified and available commercially

Named for bacterial genus, species, strain, and type

Example: EcoRI

Genus:Escherichia

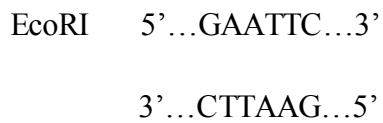
Species:coli

Strain:R

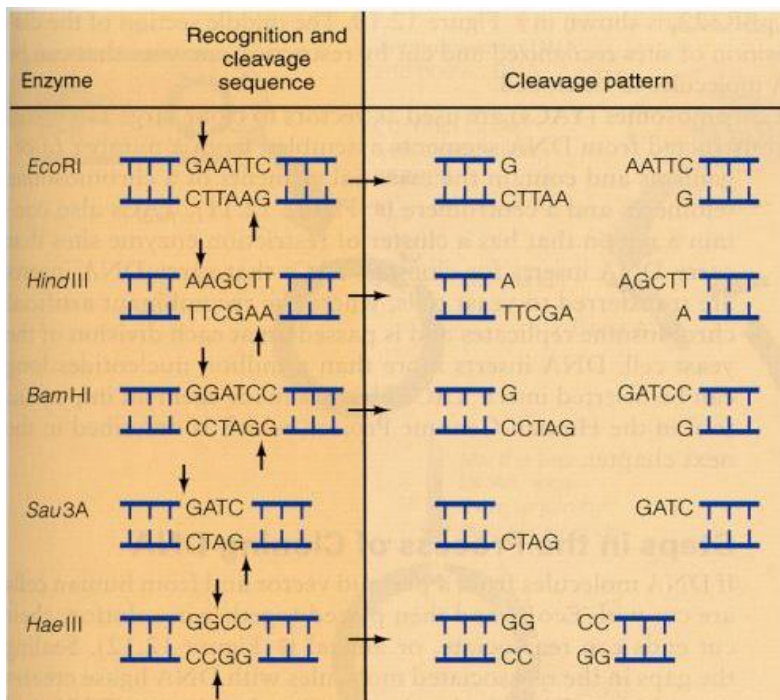
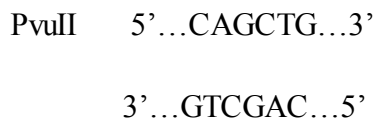
Order discovered: 1

Enzymes recognize specific 4-8 bp sequences which are palindromic.

Some enzymes cut in a staggered fashion - “sticky ends”

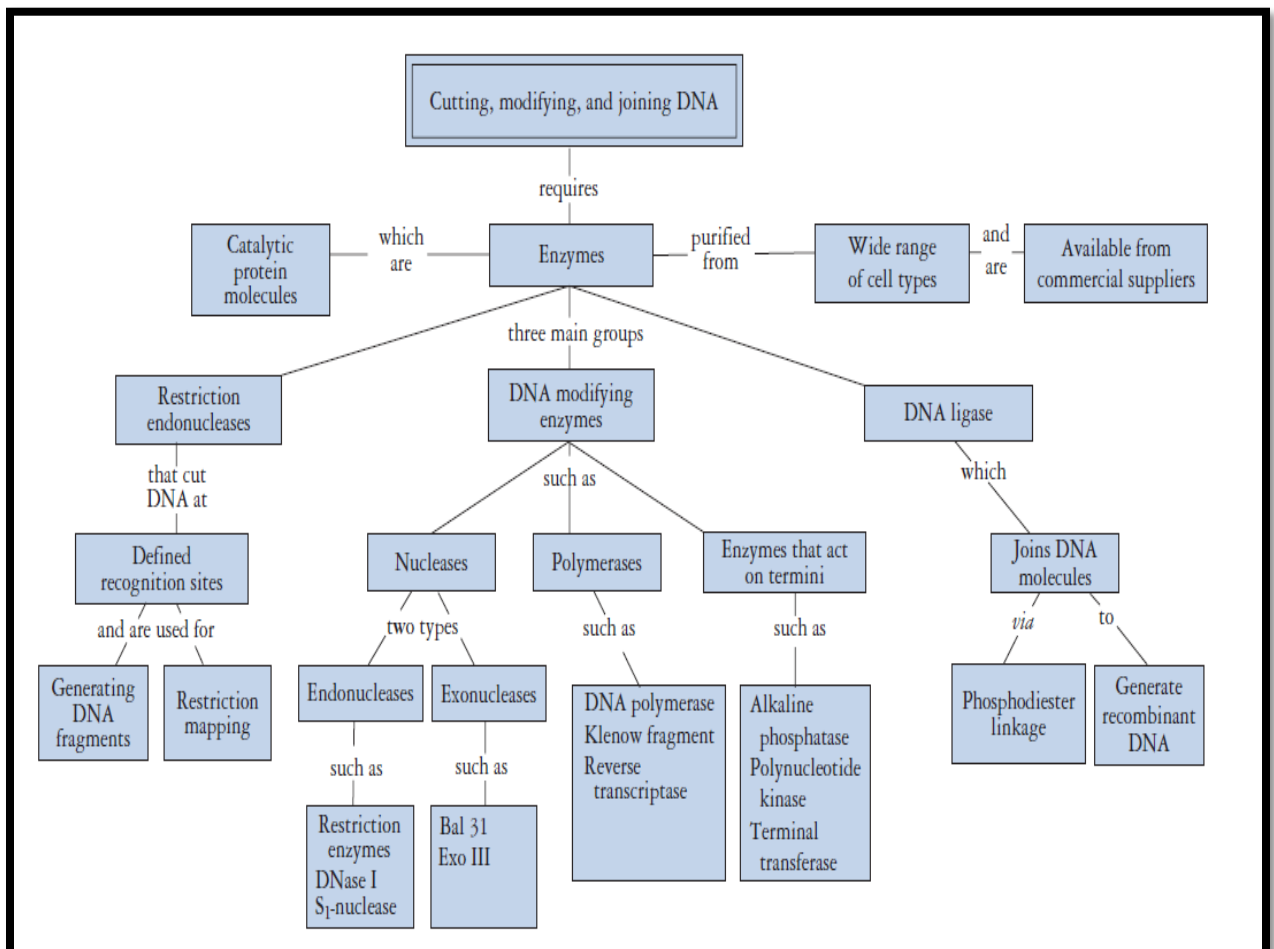


Some enzymes cut in a direct fashion – “blunt ends”



### Uses of restriction enzymes

- RFLP analysis (Restriction Fragment Length Polymorphism)
- DNA sequencing
- DNA storage – libraries
- Transformation
- Large scale analysis – gene chips



### TYPES OF DNA MODIFYING ENZYMES

Nuclease enzymes degrade nucleic acids by breaking the phosphodiester bond that holds the nucleotides together. Restriction enzymes are good examples of endonucleases, which cut within a DNA strand. A second group of nucleases, which degrade DNA from the

termini of the molecule, are known as **exonucleases**. Apart from restriction enzymes, there are four useful nucleases that are often used in genetic engineering. These are

- **Bal 31**
- **exonuclease III** (exonucleases),
- **deoxyribonuclease I** (DNase I)
- **S1-nuclease** (endonucleases).

These enzymes differ in their precise mode of action and provide the genetic engineer with a variety of strategies for attacking DNA.

- Nuclease Bal 31 is a complex enzyme. Its primary activity is a fast-acting 3' exonuclease, which is coupled with a slow-acting endonuclease. When Bal 31 is present at a high concentration these activities effectively shorten DNA molecules from both termini.
- Exonuclease III is a 3' exonuclease that generates molecules with protruding 5' termini.
- DNase I cuts either single-stranded or double-stranded DNA at essentially random sites.
- Nuclease S1 is specific for single-stranded RNA or DNA.

## **POLYMERASES**

Polymerase enzymes synthesise copies of nucleic acid molecules and are used in many genetic engineering procedures. When describing a polymerase enzyme, the terms 'DNA-dependent' or 'RNA-dependent' may be used to indicate the type of nucleic acid template that the enzyme uses.

### **DNA POLYMERASE I**

The enzyme DNA polymerase I has, in addition to its polymerase function, 5'→3' and 3'→5' exonuclease activities. The enzyme catalyses a strand-replacement reaction, where the 5'→3' exonuclease function degrades the non-template strand as the polymerase synthesises the new copy. A major use of this enzyme is in the nick translation procedure for radiolabelling DNA. The 5'→3' exonuclease function of DNA polymerase I can be

removed by cleaving the enzyme to produce what is known as the **Klenow fragment**. This retains the polymerase and 3'→5' exonuclease activities. The Klenow fragment is used where a single-stranded DNA molecule needs to be copied; because the 5'→3' exonuclease function is missing, the enzyme cannot degrade the non-template strand of dsDNA during synthesis of the new DNA. The 3'→5' exonuclease activity is suppressed under the conditions normally used for the reaction. Major uses for the Klenow fragment include radiolabelling by primed synthesis and DNA sequencing by the dideoxy method in addition to the copying of single-stranded DNAs during the production of recombinants. A modified form of DNA polymerase I called the Klenow fragment is a useful polymerase that is used widely in a number of applications

**Reverse transcriptase** (RTase) is an RNA-dependent DNA polymerase, and therefore produces a DNA strand from an RNA template. It has no associated exonuclease activity. The enzyme is used mainly for copying mRNA molecules in the preparation of cDNA (**complementary or copy DNA**) for cloning, although it will also act on DNA templates. Reverse transcriptase is a key enzyme in the generation of cDNA; the enzyme is an RNA-dependent DNA polymerase, which produces a DNA copy of an mRNA molecule.

#### ENZYMES THAT MODIFY THE ENDS OF DNA MOLECULES

The enzymes alkaline phosphatase, polynucleotide kinase, and terminal transferase act on the termini of DNA molecules and provide important functions that are used in a variety of ways. The phosphatase and kinase enzymes, as their names suggest, are involved in the removal or addition of phosphate groups.

- Bacterial alkaline phosphatase (there is also a similar enzyme, calf intestinal alkaline phosphatase) removes phosphate groups from the 5 ends of DNA, leaving a 5-OH group. The enzyme is used to prevent unwanted ligation of DNA molecules, which can be a problem in certain cloning procedures. It is also used prior to the addition of radioactive phosphate to the 5 ends of DNAs by polynucleotide kinase.
- Terminal transferase (terminal deoxynucleotidyl transferase) repeatedly adds nucleotides to any available 3 terminus. Although it works best on protruding 3 ends, conditions can be adjusted so that blunt-ended or 3-recessed molecules may be utilised. The enzyme is

mainly used to add homopolymer tails to DNA molecules prior to the construction of recombinants

## VECTORS

A cloning **vector** is a DNA molecule used as a vehicle to transfer foreign genetic material

Few of the generally used cloning vectors include:

1. Plasmids
2. Bacteriophages
3. Cosmids
4. Yeast artificial chromosomes (YAC's)
5. Bacterial artificial chromosomes (BAC's)

### Comparison of vectors available for cloning DNA fragments

Vector	Host cell	Vector structure	Insert range (kb)
M13	<i>E. coli</i>	Circular virus	1-4
Plasmid	<i>E. coli</i>	Circular plasmid	1-5
Phage $\lambda$	<i>E. coli</i>	Linear virus	2-25
Cosmids	<i>E. coli</i>	Circular plasmid	35-45
BACs	<i>E. coli</i>	Circular plasmid	50-300
YACs	<i>S. cerevisiae</i>	Linear chromosome	100-2000

### Vectors for yeast and other fungi

The yeast *Saccharomyces cerevisiae* is one of the most important organisms in biotechnology. Development of cloning vectors for yeast was stimulated by the discovery of a plasmid that is present in most strains of *S. cerevisiae*. The 2  $\mu$ m plasmid, as it is called, is one of only a very limited number of plasmids found in eukaryotic cells.

### Why is it popularly used?

Yeast cells are much easier to grow and manipulate than plant and animal cells.

The cellular biochemistry and regulation of yeast are very like those of higher eukaryotes.

There are many yeast homologues of human genes, e.g. those involved in cell division. Thus yeast can be a very good surrogate host for studying the structure and function of eukaryotic gene products.

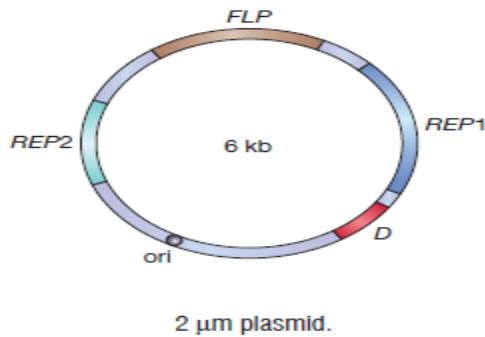
### The different kind of cloning vector derived from yeast are:

1. 2  $\mu$ m plasmid
2. yeast episomal plasmids (YEps)
3. Yeast integrative plasmids (YIps)
4. Yeast replicative plasmids (YRps)
5. yeast artificial chromosome (YAC)

Vector	Transformation frequency	Copy no./cell	Loss in non-selective medium	Disadvantages	Advantages
YIp	$10^2$ transformants per $\mu$ g DNA	1	Much less than 1% per generation	1 Low transformation frequency 2 Can only be recovered from yeast by cutting chromosomal DNA with restriction endonuclease which does not cleave original vector containing cloned gene	1 Of all vectors, this kind give most stable maintenance of cloned genes 2 An integrated YIp plasmid behaves as an ordinary genetic marker, e.g. a diploid heterozygous for an integrated plasmid segregates the plasmid in a Mendelian fashion 3 Most useful for surrogate genetics of yeast, e.g. can be used to introduce deletions, inversions and transpositions (see Botstein & Davis 1982)
YEp	$10^3$ – $10^5$ transformants per $\mu$ g DNA	25–200	1% per generation	Novel recombinants generated <i>in vivo</i> by recombination with endogenous 2 $\mu$ m plasmid	1 Readily recovered from yeast 2 High copy number 3 High transformation frequency 4 Very useful for complementation studies
YRp	$10^4$ transformants per $\mu$ g DNA	1–20	Much greater than 1% per generation but can get chromosomal integration	Instability of transformants	1 Readily recovered from yeast 2 High copy number. Note that the copy number is usually less than that of YEp vectors but this may be useful if cloning gene whose product is deleterious to the cell if produced in excess 3 High transformation frequency 4 Very useful for complementation studies 5 Can integrate into the chromosome
YCp	$10^4$ transformants per $\mu$ g DNA	1–2	Less than 1% per generation	Low copy number makes recovery from yeast more difficult than that with YEp or YRp vectors	1 Low copy number is useful if product of cloned gene is deleterious to cell 2 High transformation frequency 3 Very useful for complementation studies 4 At meiosis generally shows Mendelian segregation
YAC		1–2	Depends on length: the longer the YAC the more stable it is	Difficult to map by standard techniques	1 High-capacity cloning system permitting DNA molecules greater than 40 kb to be cloned 2 Can amplify large DNA molecules in a simple genetic background

## Vectors for yeast and other fungi

### 2 $\mu$ m plasmid



The 2  $\mu$ m plasmid is an excellent basis for a cloning vector. It is 6 kb in size and has a high copy number of between 70 and 200.

Replication makes use of a plasmid origin, several enzymes provided by the host cell, and the proteins coded by the REP1 and REP2 genes carried by the plasmid.

However, all is not perfectly straightforward in using the 2  $\mu$ m plasmid as a cloning vector.

First, there is the question of a selectable marker. For this purpose a normal yeast gene is used, generally one that codes for an enzyme involved in amino acid biosynthesis. An example is the gene LEU2, which codes for  $\beta$ -isopropyl-malate dehydrogenase, one of the enzymes involved in the conversion of pyruvic acid to leucine.

In order to use LEU2 as a selectable marker, a special kind of host organism is needed. The host must be an auxotrophic mutant that has a non-functional LEU2 gene.

Such a leu2<sup>-</sup> yeast is unable to synthesize leucine and can survive only if this amino acid is supplied as a nutrient in the growth medium. Selection is possible because transformants contain a plasmid-borne copy of the LEU2 gene, and so are able to grow in the absence of the amino acid. In a cloning experiment, cells are plated out onto minimal medium, which contains no added amino acids. Only transformed cells are able to survive and form colonies and thus are easily differentiated.



Yeast artificial chromosome (YAC) is a human-engineered DNA molecule used to clone DNA sequences in yeast cells. YACs are often used in connection with the mapping and sequencing of genomes. Segments of an organism's DNA, up to one million base pairs in length, can be inserted into YACs. The YACs, with their inserted DNA, are then taken up by yeast cells. As the yeast cells grow and divide, they amplify the YAC DNA, which can be isolated and used for DNA mapping and sequencing. Its abbreviation is YAC.

They carry large amounts of DNA so a long distance map of the region can be obtained in several steps. The repetitive regions may be 10-20 kb in length they are rarely longer than 50 kb. Thus a YAC with 100kb will contain some region that is single copy which can be used for further steps in the walk.

### Features of YACs

1. Large DNA (>100 kb) is ligated between two arms. Each arm ends with a yeast telomere so that the product can be stabilized in the yeast cell. Interestingly, larger YACs are more stable than shorter ones, which favors cloning of large stretches of DNA.
2. One arm contains an autonomous replication sequence (ARS), a centromere (CEN) and a selectable marker (*trp1*). The other arm contains a second selectable marker (*ura3*).
3. Insertion of DNA into the cloning site inactivates a mutant expressed in the vector DNA and red yeast colonies appear.
4. Transformants are identified as those red colonies which grow in a yeast cell that is mutant for *trp1* and *ura3*. This ensures that the cell has received an artificial chromosome with both telomeres (because of complementation of the two mutants) and the artificial chromosome contains insert DNA (because the cell is red).

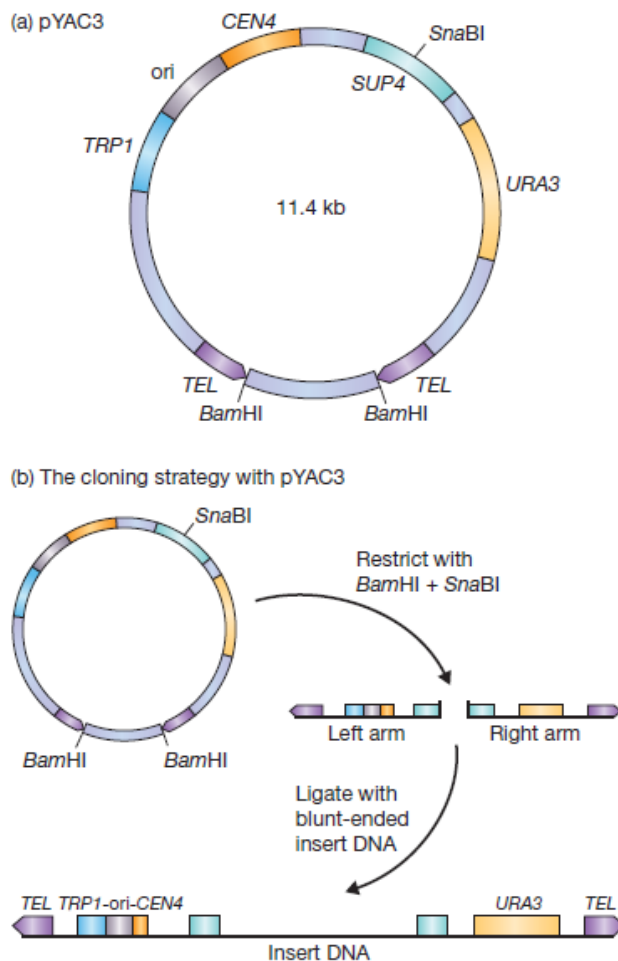
### Example : pYAC3

pYAC3 is essentially a pBR322 plasmid into which a number of yeast genes have been inserted. Two of these genes, URA3 and TRP1, have been encountered already as the selectable markers for YIp5 and YRp7, respectively. As in YRp7, the DNA fragment that carries TRP1 also

contains an origin of replication, but in pYAC3 this fragment is extended even further to include the sequence called CEN4, which is the DNA from the centromere region of chromosome 4. The TRP1–origin–CEN4 fragment therefore contains two of the three components of the artificial chromosome.

The third component, the telomeres, is provided by the two sequences called TEL.

These are not themselves complete telomere sequences, but once inside the yeast nucleus they act as seeding sequences onto which telomeres will be built. This just leaves one other part of pYAC3 that has not been mentioned: SUP4, which is the selectable marker into which new DNA is inserted during the cloning experiment.



## **Bacterial artificial chromosome**

A vector used to clone DNA fragments (100- to 300-kb insert size; average, 150 kb) in *Escherichia coli* cells. Based on naturally occurring F-factor plasmid found in the bacterium *E. coli*.

Bacterial Artificial Chromosomes (BAC) have been developed to hold much larger pieces of DNA than a plasmid can. BAC vectors were originally created F' plasmid present in few bacterial plasmids. It was found through studies that F' plasmids were able to hold up to a million basepairs of DNA from nonself bacteria. In 1992, Hiroaki Shizuya took the parts of F' that were important, cleaned it up, and turned it into a vector.

BAC vectors are able to hold up to 350 kb of DNA and have all of the tools that a vector needs to work properly, like replication origins, antibiotic resistance genes, and convenient places where clone DNA can insert itself.<sup>[8]</sup> By these vectors that can hold larger pieces of DNA, the number of clones required to cover the human genome could be reduced drastically from 1.8 billion to about 50 million.

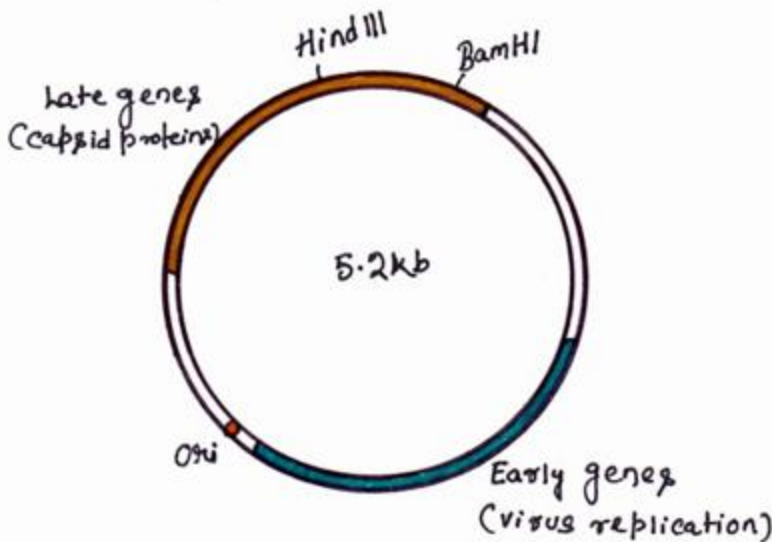
Researchers have modified BAC vectors to become more convenient to use and more useful in specialized situations. In addition to the antibiotic resistance gene that was added to identify transfected bacteria, a gene was added that enabled the bacteria to turn the colourless substance X-gal/IPTG blue. So it is possible to tell not only if a bacteria had been transfected (meaning incorporated into the cell), but also if the bacteria was transfected with the vector containing insert DNA or just the vector alone (remembering that if the vector has properly incorporated the clone DNA, it will have lost its ability to change X-gal/IPTG blue).- Modifications to BAC vectors make them more specialized.

**Few advantages:**

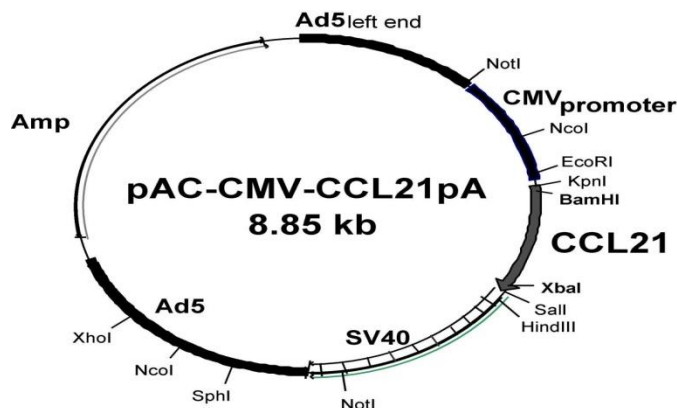
- The creation of BAC vectors has allowed researchers to do many things that they could not do before and do them more quickly and more easily. <sup>[8]</sup>
- BACs have allowed researchers to look at microbial DNA without having to actually grow the organisms<sup>[8]</sup>
- BAC vectors are also useful for studying *pathogens*, and are helpful in the development of vaccines. <sup>[8]</sup> BAC vectors are playing a tremendous role in discovering new and powerful antibiotics in the environment.
- Discovering enzymes that are able to help clean up oil spills or help farmers breed healthier farm animals or even process radioactive waste are just a few examples of what Bacterial Artificial Chromosomes can do.

**Animal viruses as cloning vectors**

The first eukaryotic DNA virus was SV40, for which a complete nucleotide sequence and a detailed understanding of transcription were available. The genome of SV40 contains very little non-essential DNA so it is necessary to insert the foreign gene in place of essential viral genes and to propagate the recombinant genome in the presence of a helper virus. This virus is capable of infecting several mammalian species, following a lytic cycle in some host and a lysogenic cycle in others. The genome is 5.2kb in size and contains two sets of genes, the early genes, expressed early in the infection cycle and coding for proteins involved in viral DNA replication, and the late genes, coding for viral capsid proteins (Figure-18). However, all work using SV40 virions to propagate recombinant DNA molecules is severely constrained by the facts that the viral genome is small, 5.24 kb, and that the packaging limits are strict. Such systems cannot, therefore, be used for the analysis of most eukaryotic genes.

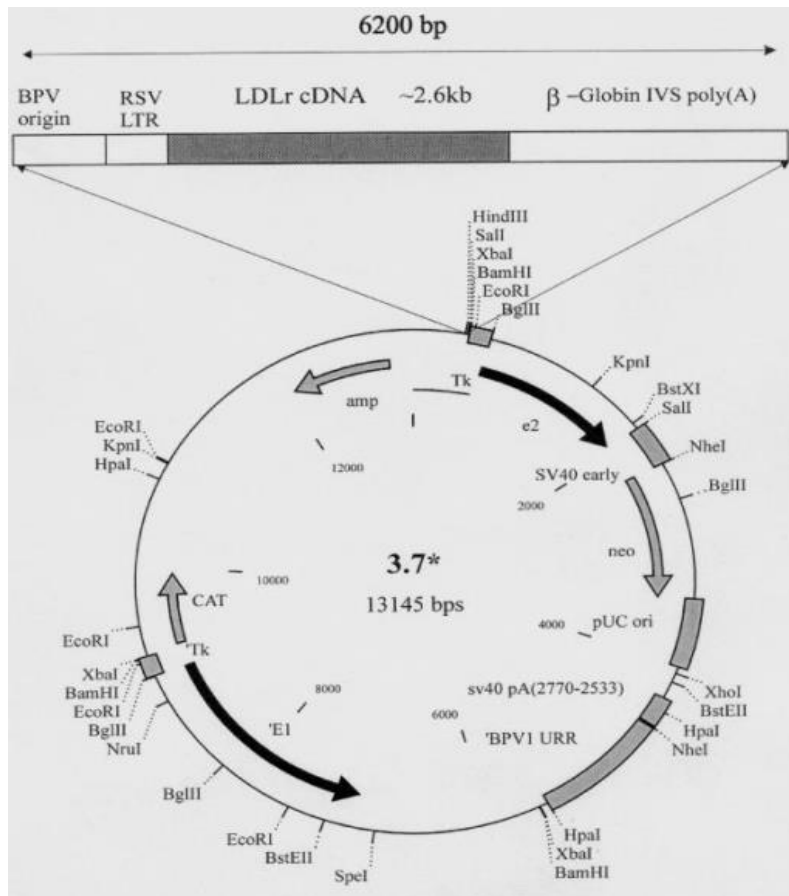


**Adenoviruses**, are a group of viruses which enable larger fragments of DNA to be cloned than it is possible with an SV40 vector. Adenoviruses are more difficult to handle because the genomes are bigger.



**Papillomaviruses**, which also have a relatively high capacity for inserted DNA, have the important advantage of enabling a stable transformed cell line to be obtained. Papillomavirus transformed cells don't contain integrated viral DNA rather they contain between 50 and 300 copies of unintegrated, circular viral DNA although some proportion of these viral genomes exists as concatamers and/or catenates. Bovine papillomavirus (BPV), which causes warts on cattle, has an unusual infection cycle in mouse cells, taking the form of a multi copy plasmid with about 100 molecules present per cell. It doesn't cause the death of the mouse cell, and BPV molecules are passed to daughter cells on cell division. Shuttle vectors consisting of BPV and

pBR322 sequences, and capable of replication in both mouse and bacterial cells, are therefore of great value in animal cell biotechnology.



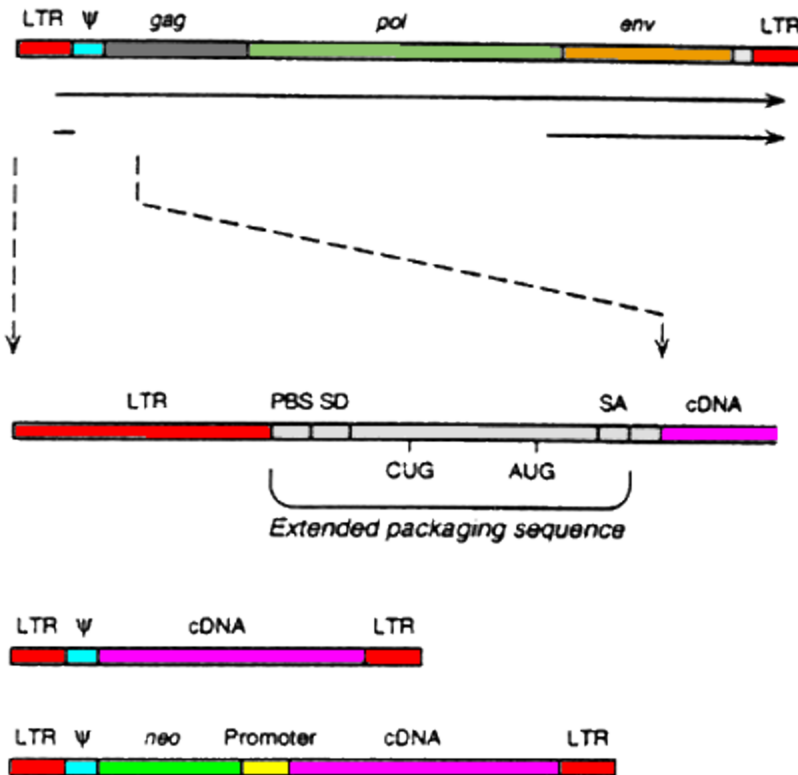
**Retroviruses**, though have single-stranded RNA genomes but provides perhaps the most promising vector system of all. During the process of reverse transcription, sequences from the termini of viral RNA are duplicated to generate long terminal repeats(LTRs). These long terminal repeats contain both the promoter and the polyadenylation signal for the transcription of viral mRNAs. The specificity of proviral DNA integration is also determined by the long terminal repeats. Although retroviruses can integrate at many sites within the cellular genome, integrative recombination always occurs at particular sites at the ends of the LTRs. The sequences appropriately inserted between the two LTRs will be integrated intact which contrasts sharply with the integration of papovavirus or adenovirus DNA, during which extensive rearrangements of the integrated viral sequences are commonplace. A further great advantage of retroviruses is that they are natural transducing viruses.

Retroviral vectors are most frequently based upon the Moloney murine leukaemia virus (Mo-MLV), which is an amphotrophic virus, capable of infecting both mouse cells, enabling vector development in mouse models, & human cells, enabling human treatment. The viral genes (gag, pol & env) are replaced with the transgene of interest & expressed on plasmids in the packaging cell line. Because the non-essential genes lack the packaging sequence ( $\psi$ ) they are not included in the virion particle. To prevent recombination resulting in replication competent retroviruses, all regions of homology with the vector backbone should be removed & the non-essential genes should be expressed by at least two transcriptional units.

The essential regions include the 5' & 3' LTRs & the packaging sequence lying downstream of the 5' LTR. To aid identification of transformed cells selectable markers, such as neomycin & beta galactosidase, can be included & transgenes expression can be improved with the addition of internal ribosome sites. The available carrying capacity for retroviral vectors is approximately 7.5 kb, which is too small for some genes even if the cDNA is used.

A requirement for retroviral integration & expression of viral genes is that the target cells should be dividing. This limits gene therapy to proliferating cells *in vivo* or *ex vivo*, whereby cells are removed from the body, treated to stimulate replication & then transduced with the retroviral vector, before being returned to the patient. When treating cancers *in vivo*, tumour cells are preferentially targeted. However, *ex vivo* cells can be more efficiently transduced, due to exposure to higher virus titres & growth factors. Furthermore *ex vivo* treated tumour cells will associate with the tumour mass & can direct tumouricidal effects.

Though transgene expression is usually adequate *in vitro* & initially *in vivo*, prolonged expression is difficult to attain. Retroviruses are inactivated by c1 complement protein & an anti-alpha galactosyl epitope antibody, both present in human sera. Transgene expression is also reduced by inflammatory interferons, specifically IFN-alpha & IFN-gamma acting on viral LTRs. As the retroviral genome integrates into the host genome it is most likely that the viral LTR promoters are being inactivated, therefore one approach has been to use promoters for host cell genes, such as tyrosine. Clearly this is an area where continued research is needed.



**Baculoviruses**, enable large amounts of proteins to be obtained from genes cloned in insect cells. One of the major proteins encoded by the virus genome is polyhedrin, which accumulates in very large quantities in the nuclei of infected cells, since the gene has an extremely active promoter. The same promoter can be used to drive the over expression of a foreign gene engineered into the baculovirus genome, and large quantities of protein can be produced in infected insect cells in culture. This method is being used increasingly for large-scale culture of proteins of animal origin, since the insect cells can produce many of the post-translational modifications of animal proteins which a bacterial expression system.

### Expression vector

An **expression vector**, otherwise known as an **expression construct**, is usually a [plasmid](#) or virus designed for [protein expression](#) in cells. The [vector](#) is used to introduce a specific [gene](#) into a target cell, and can commandeer the cell's mechanism for [protein synthesis](#) to produce the [protein encoded](#) by the gene. The plasmid is engineered to contain regulatory sequences that act as [enhancer](#) and [promoter](#) regions and lead to efficient transcription of the gene carried on the



expression vector.<sup>[1]</sup> The goal of a well-designed expression vector is the production of significant amount of stable messenger RNA, and therefore proteins. Expression vectors are basic tools for biotechnology and the production of proteins. An example is insulin which is used for medical treatments of diabetes.

### **Elements for expression**

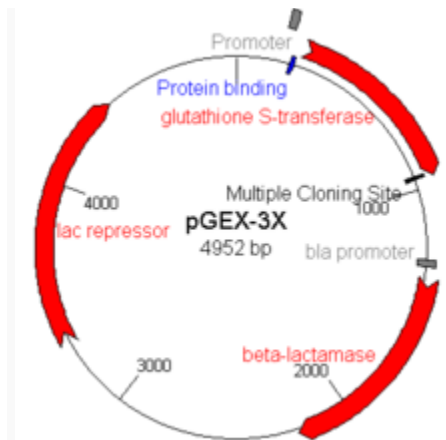
An expression vector must have elements necessary for protein expression. These may include a strong promoter, the correct translation initiation sequence such as a ribosomal binding site and start codon, a strong termination codon, and a transcription termination sequence.<sup>[2]</sup> There are differences in the machinery for protein synthesis between prokaryotes and eukaryotes, therefore the expression vectors must have the elements for expression that is appropriate for the chosen host. For example, prokaryotes expression vectors would have a Shine-Dalgarno sequence at its translation initiation site for the binding of ribosomes, while eukaryotes expression vectors would contain the Kozak consensus sequence.

The promoter initiates the transcription and is therefore the point of control for the expression of the cloned gene. The promoters used in expression vector are normally inducible, meaning that protein synthesis is only initiated when required by the introduction of an inducer such as IPTG. Protein expression however may also be constitutive (i.e. protein is constantly expressed) in some expression vectors. Low level of constitutive protein synthesis may occur even in expression vectors with tightly controlled promoters.

### **Expression systems**

Different organisms may be used to express a target protein, the expression vector used therefore will have elements specific for use in the particular organism. The most commonly used organism for protein expression is the bacterium Escherichia coli. However not all proteins can be successfully expressed in *E. coli*, and other systems may therefore be used.

## Bacterial



An example of a bacterial expression vector is the pGEX-3x plasmid

The expression host of choice for the expression of many proteins is *Escherichia coli* as the production of heterologous protein in *E. coli* is relatively simple and convenient, as well as being rapid and cheap. A large number of *E. coli* expression plasmids are also available suitable for a wide variety of needs. Other bacteria used for protein expression include *Bacillus subtilis*.

Most heterologous proteins are expressed in the cytoplasm of *E. coli*. However, not all proteins formed may be soluble in the cytoplasm, and incorrectly folded proteins formed in cytoplasm can form insoluble aggregates called inclusion body. Such insoluble proteins will require refolding which can be an involved process and may not produce high yield. Where necessarily, for example when the protein can only fold correctly in an oxidizing environment due to the presence of disulphide bonds, the protein may be targeted to the periplasmic space by the use of an N-terminal signal sequence. Other more sophisticated systems are being developed; such systems may allow for the expression of proteins previously thought impossible in *E. coli*, such as glycosylated proteins.

The promoters used for these vector are usually based on the promoter of the lac operon or the T7 promoter, and they are normally regulated by the lac operator. These promoters may also be hybrids of different promoters, for example, the tac promoter is a hybrid of trp and lac promoters. Note that most commonly used lac or lac-derived promoters are based on the lacUV5 mutant which is insensitive to catabolite repression. This mutant allows for expression of protein under the control of the lac promoter when the growth medium contains

glucose since glucose would inhibit protein expression if wild-type *lac* promoter is used. Presence of glucose nevertheless may still be used to reduce background expression through residual inhibition in some systems.

Examples of *E. coli* expression vectors are the pGEX series of vectors where glutathione-S-transferase is used as a fusion partner and protein expression is under the control of the *tac* promoter, and the pET series of vectors which uses a T7 promoter.

It is possible to simultaneously express two or more different proteins in *E. coli* using different plasmids. However, when 2 or more plasmids are used, each plasmid needs to use a different antibiotic selection as well as a different origin of replication, otherwise the plasmids may not be stably maintained. Many commonly-used plasmids are based on the ColE1 replicon and are therefore incompatible with each other; in order for a ColE1-based plasmid to coexist with another in the same cell, the other would need to be of a different replicon, e.g. a p15A replicon-based plasmid such as the pACYC series of plasmids. Another approach would be to use a single two-cistron vector or design the coding sequences in tandem as a bi- or poly-cistronic construct.

### **Mammalian**

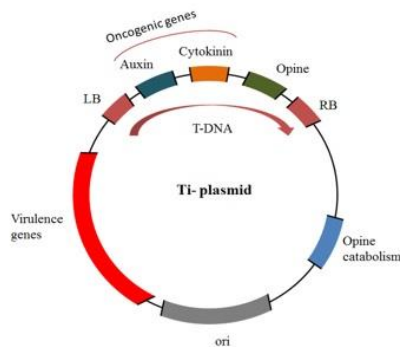
Mammalian expression vectors offer considerable advantages for the expression of mammalian proteins over bacterial expression systems - proper folding, post-translational modifications, and relevant enzymatic activity. It may also be more desirable than other eukaryotic non-mammalian systems whereby the proteins expressed may not contain the correct glycosylations. It is of particular use in producing membrane-associating proteins that require chaperones for proper folding and stability as well as containing numerous post-translational modifications. The downside, however, is the low yield of product in comparison to prokaryotic vectors as well as the costly nature of the techniques involved. Its complicated technology, and potential contamination with animal viruses of mammalian cell expression have also placed a constraint its use in large-scale industrial production.

Cultured mammalian cell lines such as the Chinese hamster ovary (CHO), HEK, HeLa, and COS cell lines may be used to produce protein. Vectors are transfected into the cells and the DNA may be integrated into the genome by homologous recombination in the case of stable transfection, or the cells may be transiently transfected. Examples of mammalian expression vectors include the adenoviral vectors, the pSV and the pCMV series of plasmid

vectors, vaccinia and retroviral vectors, as well as baculovirus. The promoters for cytomegalovirus (CMV) and SV40 are commonly used in mammalian expression vectors to drive protein expression. Non-viral promoter, such as the elongation factor (EF)-1 promoter, is also known.

### Ti-plasmid

The virulent strains of *A. tumefaciens* harbor large plasmids (140–235 kbp) known as tumor-inducing (Ti) plasmid involving elements like T-DNA, vir region, origin of replication, region enabling conjugative transfer and o-cat region (required for catabolism of opines).



Ti-plasmid of *Agrobacterium*

### T-DNA

It is a small, specific segment of the plasmid, about 24kb in size and found integrated in the plant nuclear DNA at random site. This DNA segment is flanked by right and left borders.

The functions of T-DNA genes are listed

Gene	Product	Function
<i>ocs</i>	Octopine synthase	Opine synthesis
<i>nos</i>	Nopaline synthase	Opine synthesis
<i>trns1 (iaaH, auxA)</i>	Tryptophan-2-mono-oxygenase	Auxin synthesis
<i>trns2 (iaaM, auxB)</i>	Indoleacetamide hydrolase	Auxin synthesis
<i>trnr (ipt, cyt)</i>	Isopentyltransferase	Cytokinin synthesis

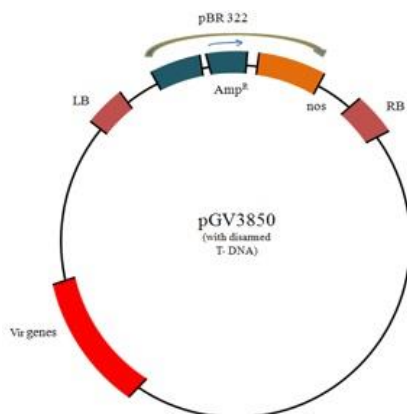
<i>frs</i>	Fructose 1,6-bisphosphate synthase	Opine synthesis
<i>mas</i>	Mannose 6-phosphate synthase	Opine synthesis
<i>ags</i>	Agrobacterium galactosyltransferase	Opine synthesis

### T- DNA:Border Sequences

- T-regions are defined by direct repeats known as T-DNA border sequences (Right and Left Border i.e. RB and LB of 25 bp each).
- These are not transferred intact to the plant genome, but are involved in the transfer process.
- The RB is rather precise, but the LB can vary by about 100 nucleotides.
- Deletion of the RB repeat abolishes T-DNA transfer, but the LB seems to be non-essential. The LB repeat has little transfer activity alone.

### Disarmed Ti-plasmid derivatives as plant vectors

Ti plasmid is a natural vector for genetically engineering plant cells due to its ability to transfer T-DNA from the bacterium to the plant genome. But wild-type Ti plasmids are not suitable as vectors due to the presence of oncogenes in T-DNA that cause tumor growth in the recipient plant cells. For efficient plant regeneration, vectors with disarmed T-DNA are used by making it non-oncogenic by deleting all of its oncogenes. The foreign DNA is inserted between the RB and LB and then integrated into the plant genome without causing tumors.



Structure of the Ti-plasmid pGV3850 with disarmed T-DNA.

- The creation of disarmed T-DNA is an important step forward, but the absence of tumor formation makes it necessary to use an alternative method for the identification of transformed plant cells. Opine production using *pGV3850* was exploited as a screenable phenotype, and the *ocs* and *nos* genes are now widely used as screenable markers.

### Drawbacks

Several drawbacks are associated with disarmed Ti- vector systems as discussed below;

- Necessity to carry out enzymatic assays on all potential transformants.
- Not convenient as experimental gene vectors due to large size.
- Difficulty in *in vitro* manipulation and
- Absence of unique restriction sites in the T-DNA.

### Co- integrate vectors

Co-integrate vectors are the deletion derivatives of Ti-plasmids. The DNA to be introduced into the plant transformation vector is sub cloned in a conventional *Escherichia coli* plasmid vector for easy manipulation, producing a so-called *intermediate vector*. These vectors are incapable of replication in *A. tumefaciens* and also lack conjugation functions. Transfer is achieved using a 'triparental mating' in which three bacterial strains are mixed together:

- (i) An *E. coli* strain carrying a helper plasmid able to mobilize the intermediate vector in *trans*;
- (ii) The *E. coli* strain carrying the recombinant intermediate vector;
- (iii) *A. tumefaciens* carrying the Ti plasmid.

Conjugation between the two *E. coli* strains transfers the helper plasmid to the carrier of the intermediate vector, which in turn is mobilized and transferred to the recipient *Agrobacterium*. Homologous recombination between the T-DNA sequences of the Ti plasmid and intermediate vector forms a large co- integrate plasmid resulting in the transfer of recombinant T-DNA to the plant genome.

**Binary vector**

- Binary vector was developed by Hoekma *et al* (1983) and Bevan in (1984).
- It utilizes the trans- acting functions of the vir genes of the Ti-plasmid and can act on any T-DNA sequence present in the same cell.
- Binary vector contains transfer apparatus(the vir genes) and the disarmed T-DNA containing the transgene on separate plasmids.

**Advantages of Binary vector**

- Small size due to the absence of border sequences needed to define T-DNA region and vir region.
- Ease of manipulation