In Vitro Recombinant DNA Technology

Churchward, Gordon; Chandler, Michael

in: Comprehensive Biotechnology
Vol. 1 p. 77 - 111 (1985)
ed. by M. Moo-Young
Pergamon Press

6

In Vitro Recombinant DNA Technology

G. CHURCHWARD and M. CHANDLER
University of Geneva, Switzerland

6.1 INTRODUCTION

The past 10 years have seen a dramatic acceleration in our understanding of the organization and expression of both prokaryotic and eukaryotic genomes. These advances have been due in large part to the availability of a range of techniques known collectively as recombinant DNA technology.

The most important contribution to this rapid growth has been the development of methods for obtaining specific genes in high yield and in a pure form. The magnitude of the problem confronting the molecular biologist wishing to study a specific gene which forms part of a complex genome can best be understood by considering that the human β-globin gene (1.6 kb), for example, represents only 0.00005% of the human genome (2.9×10^6 kb). Previously, detailed study of gene expression at the molecular level was limited to those genes that could be propagated as part of a viral genome or a bacterial plasmid because these autonomously replicating molecules carry relatively few genes and can be purified easily, in quantity, away from the genomic DNA of the host. The discovery of enzymes able to cleave a DNA molecule at specific sites (restriction endonucleases) or to join covalently double-stranded DNA molecules (DNA ligases) *in vitro*, together with a detailed knowledge of the organization of bacterial viruses and plasmids (particularly those of *Escherichia coli*), provided a means of purification and propagation of genes from any source. Bacterial viruses and plasmids can be harnessed as vectors for foreign DNA fragments by covalent linkage of the fragment to the vector at a predetermined position defined by a restriction endonuclease cleavage site. Provided that the insertion of the DNA does not interrupt any essential vector functions, the hybrid molecule can be propagated and easily isolated.

The basic gene enrichment procedure, known as molecular cloning, together with the many complementary methods allowing rapid DNA sequencing and *in vitro* manipulation of the cloned segment, has permitted the purification and study of the structure and the expression of genes obtained from a wide variety of organisms. The subsequent development of host–vector systems for organisms other than *E. coli* and of 'shuttle' vectors (bifunctional replicons able to propagate in several hosts) has introduced a large degree of flexibility into the study of gene structure and function. Not only can a gene be studied in a heterologous system, for example mammalian genes in *E. coli*, but it is now possible to study many genes in an homologous system by reintroduction of a cloned gene into cells of the original organism. This flexibility has profoundly changed the approach of the geneticist. In addition to isolating mutants by screening for a given phenotype and subsequently locating these mutations within a specific gene by employing suitable genetic crosses, recombinant DNA technology has provided the means to introduce predetermined changes in a given gene and to observe the effects of such changes on gene expression. This type of approach has been called 'reversed genetics'.

Although the full potential of these methods has yet to be realized, they have already been applied with success to a wide range of problems of economic, environmental and medical importance.

In this review, we propose to outline briefly the concepts involved in recombinant DNA technology and to present the reader with a comprehensive, although not exhaustive, discussion of many of the techniques available for the cloning and manipulation of genes. Where possible we have cited review articles rather than the original papers. A list of several important articles, some of which include detailed experimental protocols, will be found at the end of this chapter.

We have divided gene cloning into five different stages: (a) generation of DNA fragments for cloning, (b) covalent linkage of the fragments to a vector plasmid or bacteriophage, (c) introduction of this hybrid molecule into a suitable host, (d) detection of the desired clone, and (e) characterization of the purified cloned fragment. Subsequent operations may involve: (f) modification of the cloned piece by removal of inessential DNA regions and mutagenesis, and (g) recloning into a vector to study and optimize expression of the gene in its original host or in other secondary host organisms.

6.2 GENERATION AND CLONING OF DNA FRAGMENTS

6.2.1 Fragmentation of DNA

One of the first problems encountered prior to cloning a given gene is to separate the gene physically from the rest of the DNA molecule of which it forms a part. This is normally accomplished either by fragmentation of the DNA with restriction enzymes or by synthesis of a double-stranded DNA copy of partially purified mRNA (cDNA).

6.2.1.1 Class II restriction enzymes

Fragmentation of a genome for cloning is most conveniently carried out by the use of class II restriction endonucleases of which over 200 are known (Roberts, 1983). Each enzyme recognizes a specific nucleotide sequence and cleaves both strands of the DNA. A specific recognition sequence may consist of six, five or four nucleotides which on average would be expected to occur once every 4096 (1 in 4^6; infrequent cutters), 1024 (1 in 4^5) and 256 (1 in 4^4; frequent cutters) base pairs respectively in DNA with a 50% G + C content.

Many such enzymes recognize palindromic sequences and cleave each strand of the DNA, several base pairs from the axis of symmetry generating protruding 5′ or 3′ complementary, single-stranded extensions (Figure 1, Table 1). This complementarity facilitates the eventual insertion of the fragment into a vector molecule cleaved with the same enzyme, since the extremities of the vector and the fragment are able to hybridize and form metastable structures prior to their covalent linkage (Mertz and Davis, 1972; Section 6.2.2). The role of homology in cloning using this type of enzyme in general precludes the direct insertion of a fragment generated by one enzyme into a vector cleaved with another. However, exceptions exist, for example in the case of the enzymes *Sal*I and *Xho*I, and *Bam*HI and *Bgl*II, where the extensions (but not the entire recognition sequences) are identical (Table 1).

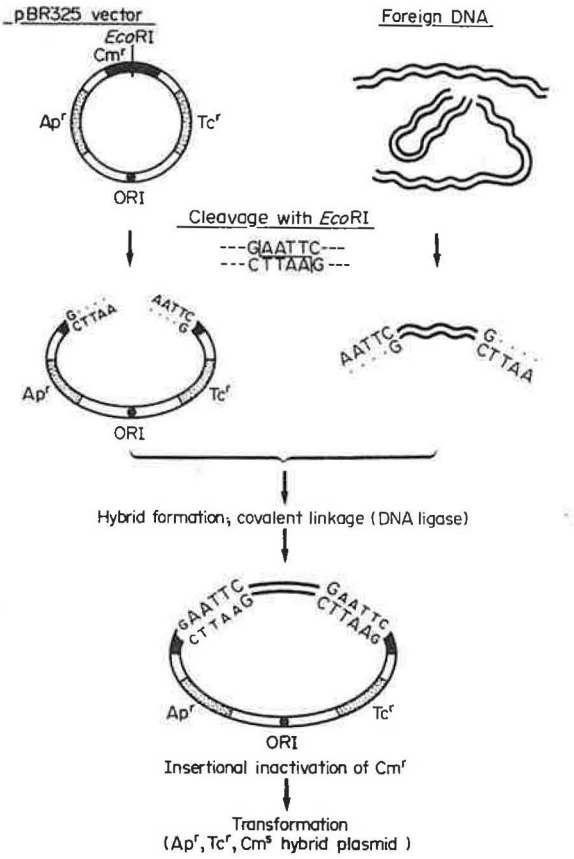

Figure 1 Cloning foreign DNA into an *E. coli* plasmid vector using insertional inactivation. Following cleavage of pBR325 vector and foreign DNA with a restriction endonuclease, *Eco*RI, the DNAs are ligated together. Insertion of DNA into the vector inactivates the gene specifying chloramphenicol resistance (Cm^r). The desired clones may be recovered after transformation as colonies that are resistant to ampicillin (Ap^r) or tetracycline (Tc^r), but sensitive to chloramphenicol

Other class II restriction enzymes cleave both DNA strands at the axis of symmetry of the recognition site and thus generate fragments which carry neither 5′ nor 3′ extensions. These 'flush' or 'blunt' ended fragments can also be cloned directly (Sgaramella *et al.*, 1970). Since no single-

Table 1 Recognition Sites of Some Restriction Endonucleases Used in Cloning

	5′ Extensions		*3′ Extensions*		*Blunt ends*	
(A)	*Bam*HI	-G|G A T C C- -C C T A G|G-	*Kpn*I	-G G T A C|C- -C|C A T G G-	*Pvu*II	-C A G|C T G- -G T C|G A C-
	*Bgl*II	-A|G A T C T- -T C T A G|A-	*Pst*I	-C T G C A|G- -G|A C G T C-	*Sma*I	-C C C|G G G- -G G G|C C C-
	*Eco*RI	-G|A A T T C- -C T T A A|G-				
(B)	*Bam*HI/*Bgl*II hybrid site			-G G A T C A- -G C T A G T-	Cleaved by *Sau*3A	|-G A T C- -C T A G-|

stranded extensions are involved, any 'blunt-ended' fragment can be inserted into any 'blunt-end' site in the vector allowing greater flexibility.

When a fragment produced by digestion of DNA with a restriction enzyme is inserted into a vector digested with the same enzyme, a recognition sequence is regenerated at each junction between vector and inserted DNA. This can facilitate the recovery of fragments for subsequent manipulation. If different enzymes are used to produce the fragment and cleave the vector, hybrid recognition sequences are generated which are usually not susceptible to cleavage by either enzyme. However, they may be cleaved by a third enzyme (Table 1B).

6.2.1.2 Random DNA fragments and the generation of genomic libraries

If the genome of interest is small, for example a virus or plasmid, or if the gene can be enriched in the population, direct cloning using class II restriction enzymes should be the method of choice. For larger genomes, however, it is usual to generate a genomic library (a collection of clones which together include most or all of the genome), and to screen amongst groups of these clones for the presence of the desired gene (Maniatis *et al.*, 1978). For this purpose fragmentation of the DNA should be random. This can be achieved by shearing to generate fragments of a suitable length. However, partial digestion of the genome with a restriction enzyme which cleaves frequently is normally employed (*e.g.* Ish-Horowicz and Burke, 1981). For example, DNA which has been partially digested with *Sau*3A can be inserted into a vector cleaved with *Bam*HI, making use of the homologous 5′ extensions generated by these two enzymes (see Table 1). The number of independent clones required to ensure the appearance of a given gene in the library is determined by the average length of the cloned fragments, the length of the desired sequence and the total length of the genome (Carbon *et al.*, 1977). With fragments of average size 20 kb, about 900 clones are required to ensure that a given *E. coli* gene (genome size 4000 kb) is present with 99% probability. This increases to 3100 in the case of yeast (13 500 kb), 6.7×10^5 for the human genome (2.9×10^6 kb) and 2.4×10^7 for the South American lungfish (1.02×10^8 kb). For a detailed discussion of the theoretical and practical problems in constructing genomic libraries the reader is referred to the paper by Seed *et al.* (1982).

6.2.1.3 Enrichment for specific DNA sequences

In certain circumstances it is possible to obtain an initial enrichment of the desired DNA fragment prior to cloning. Where the fragment can be defined by hybridization with purified or enriched mRNA or other specific probes, an initial enrichment can be obtained by fractionation of the population of fragments according to their size. The most convenient methods of fractionation are agarose or acrylamide gel electrophoresis (see Southern, 1979) followed by extraction of the DNA from the gel matrix (Smith, H. O., 1980; Yang *et al.*, 1979) or sedimentation in sucrose density gradients.

6.2.1.4 Synthesis of cDNA

For the isolation of a gene from large, complex eukaryotic genomes, a second method is frequently employed. This involves the synthesis of a double-stranded cDNA copy from isolated

mRNA molecules. The power of this technique is that only a fraction of the genome (that fraction which is transcribed into mRNA) is copied. The resulting cDNA clones can subsequently be used as probes to identify genomic fragments contained in a genomic library. The coding region of many eukaryotic structural genes is frequently interrupted by introns (intervening sequences) which are removed from the mature RNA by splicing (*e.g.* Breathnach and Chambon, 1981). Since a primary bacterial host will not possess splicing systems, cloning of cDNA copies of mature mRNA is the only way to obtain expression of such genes in bacteria. A comparison of cDNA clones with their genomic counterparts also provides important information concerning RNA processing.

Schemes for purification of bulk eukaryotic mRNA often make use of the ability of polyadenylated 3′ ends to hybridize with, and be reversibly retained by, oligodeoxythymidilate residues immobilized on a solid matrix (Aviv and Leder, 1972). Species such as ribosomal RNA which do not contain large tracts of adenine residues cannot hybridize with the immobilized oligo(dT) and are therefore not retained by the matrix (Figure 2). The polyadenylated mRNA is then eluted from the matrix. If a suitable assay exists for the gene product, an additional purification can be carried out at this stage by fractionating the purified mRNA according to size on a sucrose gradient and testing the capability of each fraction to direct the synthesis of the gene product (*e.g.* Buell *et al.*, 1978; Gray *et al.*, 1982; Section 6.4.1.2). These methods do not generally result in sufficient purification to make cloning cDNA copies of rare mRNA species a straightforward task. If suitable antibodies are available they can be used to precipitate polysomes to purify a specific mRNA. This can greatly enrich for rare mRNA molecules (Korman *et al.*, 1982).

Double-stranded DNA copies are synthesized in several steps from the collection of purified mRNA molecules (Figure 2). DNA synthesis is carried out by use of avian myeloblastosis virus reverse transcriptase (AMVrt) and is primed by an oligo(dT) residue which specifically hybridizes with the 3′ polyadenylated end of the mRNA. After hydrolysis of the mRNA template, conditions are employed which lead to the formation of small 'hairpins' at the 3′ end of the newly synthesized DNA. The small hairpin is used to prime DNA synthesis by AMVrt (Rougeon and Mach, 1976) or by *Escherichia coli* DNA polymerase I (Efstratiadis *et al.*, 1976) using the newly synthesized cDNA copy as template. Both strands of the resulting double-stranded cDNA copy are therefore covalently joined at one end. Prior to cloning, this covalent linkage must be broken. The most common method has been the use of the single strand specific endonuclease S1 (Figure 2). Self-priming employing hairpin structures, followed by S1 treatment, can seriously reduce the probability of obtaining full length cDNA, particularly if the mRNA is large. This problem can be circumvented by extending the 3′ end of the initial single-stranded cDNA copy using dCTP and terminal deoxynucleotidyl transferase (Section 6.2.3), removal of the mRNA template by hydrolysis, and priming of complementary DNA synthesis with oligo(dG) which has been hybridized to the 3′ oligo(dC) extremity (Figure 2, Land *et al.*, 1981). More recently, an extremely efficient system has been developed which uses a specialized vector plasmid (Okayama and Berg, 1982). With this system, the cDNA is synthesized directly into the vector and cDNA copies of RNA molecules capable of encoding proteins in excess of 200 000 D have been cloned.

6.2.1.5 Chemical synthesis of DNA

Although the methods for generating DNA fragments described above are those most commonly used, an increasingly important method for generating specific DNA sequences is that of chemical synthesis. Solid phase synthetic methods can give a chain growth rate of up to one residue per 30 min and generate tetradecamers with a yield in excess of 70% (Alvarado-Urbina *et al.*, 1981). Rapid development in this area has resulted in the production of an enormous variety of DNA sequences. Short, double-stranded oligonucleotide linkers encoding a restriction enzyme recognition sequence can be used to modify the extremities of DNA fragments prior to cloning (Section 6.2.3). Single-stranded oligonucleotides can be used to introduce short deletions or specific base changes into any DNA fragment (Section 6.5.1) and are of increasing importance in DNA sequencing techniques (Anderson *et al.*, 1980).

A very important use of chemically synthesized oligonucleotides is in the screening of cDNA libraries for specific genes (Section 6.4.1). If the partial amino acid sequence of a protein is known, a synthetic oligonucleotide can sometimes be synthesized, permitting detection of the cloned gene by nucleic acid hybridization. Often this is the only feasible approach to cloning a particular gene since other methods of detection are too laborious.

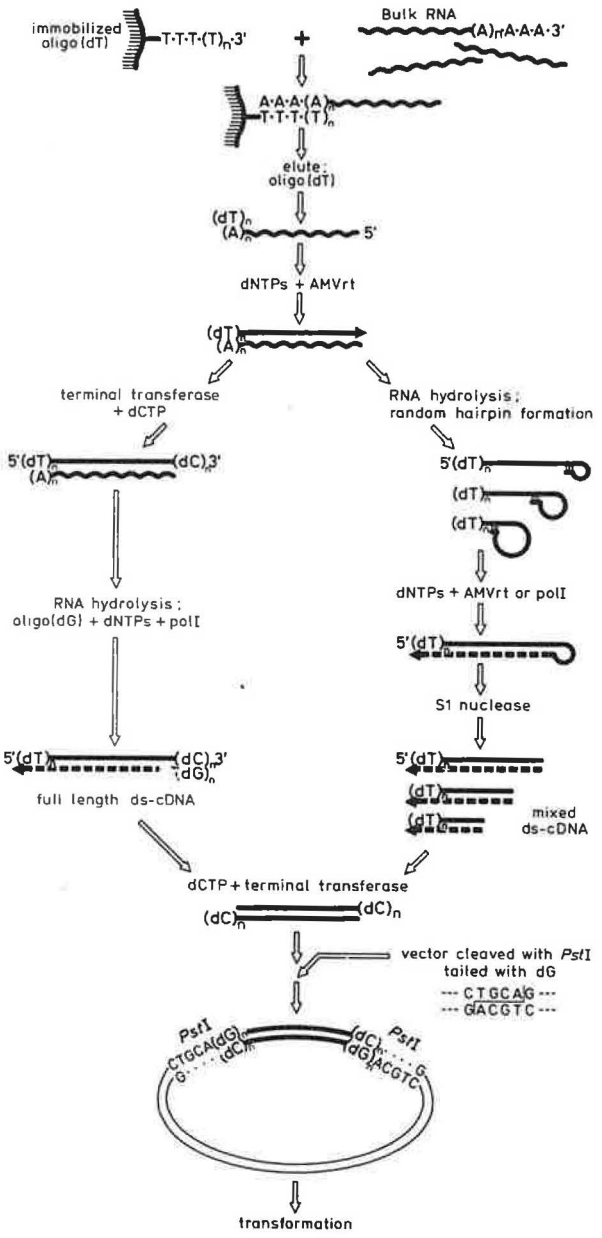

Figure 2 Synthesis and cloning of cDNA. The steps in the synthesis of cDNA copies of a population of polyadenylated mRNA molecules and the subsequent complementary strand synthesis are shown. Two alternatives for complementary strand priming are demonstrated: either by tailing with (dc) and terminal transferase, and using oligo(dG) as primer, or by hairpin formation and subsequent S1 nuclease digestion. The double-stranded cDNA is cloned into a vector by tailing, although addition of synthetic linkers is a frequently used alternative

6.2.2 Covalent Linkage of DNA Fragments to Vector Molecules

In order to purify and amplify a given DNA fragment, it must be inserted and covalently joined to a suitable vector molecule. In a ligation reaction both inter- and intra-molecular ligation may occur and often a large proportion of the products are reconstituted vector molecules that do not contain an inserted DNA fragment. In many cases, it is thus important to ensure that the majority of resulting molecules do indeed carry an inserted segment of foreign DNA since this greatly reduces the effort involved in the subsequent identification of the hybrid molecule.

6.2.2.1 Ligation to vector molecules

Covalent linkage of DNA fragments to a suitable vector molecule is generally accomplished by treatment of a DNA mixture with T4 or, less frequently, *E. coli* DNA ligase. These enzymes use the high energy pyrophosphate linkage of a nucleotide cofactor (ATP for T4; NAD for *E. coli*) to generate a phosphodiester bond between 5′-phosphoryl and 3′-hydroxyl termini of DNA fragments (*e.g.* Higgins and Cozzarelli, 1980). Since nearly all class II restriction enzymes generate 5′-phosphoryl and 3′-hydroxyl ends on cleavage of DNA, fragments generated by these enzymes are excellent substrates for ligation.

Both T4 and *E. coli* ligase can efficiently join molecules carrying complementary extensions but T4 DNA ligase is also able to join fragments lacking single-stranded terminal extensions. The latter reaction requires high concentrations of DNA extremities and of ligase (Sgaramella and Khorana, 1972) and is preferentially inhibited by high concentrations of ATP (Ferretti and Sgaramella, 1981). *E. coli* ligase is unable to catalyze this type of reaction.

Optimal ligation conditions depend not only on the type of DNA extremity involved but also on the type of vector employed and on the sizes of both vector and fragment. In the case of ligation of DNA fragments to a plasmid vector, reaction conditions should favour an initial joining of vector to fragments and subsequently a circularization of the hybrid molecule (Figure 1). In the case of cloning using a linear vector such as phage λ, three fragments must be joined: the two arms of λ and the foreign DNA fragment. Intramolecular ligation (circularization) or intermolecular ligation can be favoured in a particular reaction by varying the overall concentration of DNA and the relative concentration of fragment and vector (Dugaiczyk *et al.*, 1975).

Two reaction parameters are important: the effective concentration of one end of a DNA fragment in the immediate neighbourhood of the other (j), which depends on the length of the fragment and is independent of the DNA concentration; and the concentration of the DNA termini in the reaction mixture (i) (Dugaiczyk *et al.*, 1975). Circularization is more likely at low DNA concentrations where j for a given molecule is greater than the sum of the concentration of all ligatable ends in the mixture ($j > \Sigma i$). Intermolecular ligation and thus the formation of hybrid molecules are favoured at high DNA concentrations when $j < \Sigma i$. Optimal conditions for insertion of foreign DNA into a λ vector have been determined empirically to require a ratio of vector ends to fragment ends of about 2:1 and a DNA concentration sufficiently high to ensure that i for each DNA species is greater than j (Maniatis *et al.*, 1978). For cloning using a plasmid vector, conditions that result in the insertion of single DNA fragments into the vector molecule are normally chosen empirically.

6.2.2.2 Methods favouring formation of hybrid DNA molecules

Methods have been developed which favour intermolecular ligation and thus increase the probability of obtaining an insert into a vector molecule. In certain cases, where insertion results in the destruction of a unique restriction site in the vector and no site is present in the inserted DNA, the ligation mixture can simply be redigested with the enzyme prior to introduction of the DNA into the host organism. In the second method, the vector and the DNA to be cloned are both cleaved with two restriction endonucleases which each generate single strand extensions of different sequence. Recovery of reconstituted vector molecules is frequently reduced since recircularization requires the formation of a dimer molecule. In *E. coli*, for example, head to head dimer molecules normally cannot be propagated.

The third method exploits the requirement for 5′-phosphoryl and 3′-hydroxyl groups in the ligase-catalyzed reaction. After cleavage with a restriction enzyme the vector molecule is treated with alkaline phosphatase to remove the 5′-phosphoryl groups (Table 2). This precludes both circularization of the vector and the formation of oligomeric vector molecules, and ensures the insertion of 5′-phosphorylated foreign DNA (Ullrich *et al.*, 1977). Ligation results in the covalent linkage of the 5′-phosphorylated extremities of the fragment to the vector. Covalent linkage of the 3′ extremities to the dephosphorylated 5′ extremities of the vector presumably occurs *in vivo* following the introduction of the molecule into the host cell.

A fourth method which ensures covalent linkage of the vector to foreign DNA is that of homopolymer tailing and is described below.

The 'selection' procedures described above operate at the level of the ligation reaction and select against the reconstitution of the vector molecule. Additional or alternative methods which

Table 2 Enzymes Used in the Manipulation of DNA

Class II restriction enzymes (*e.g. Eco*RI)	5′-G.A.A.T.T. C-3′ 3′-C.T.T. A.A.G-5′	$-G^{OH}$ $^{P}A.A.T.T.C-$ $-C.T.T.A.A._{P}$ $_{HO}G-$	Site specific cleavage
T4 DNA ligase	$-(N)^{OH} + {}^{P}(N)-$ $-(N')_{P}$ $_{HO}(N')-$	$-(N)\cdot(N)-$ $-(N')\cdot(N')-$	Covalent linkage
Alkaline phosphatase	$-(N)^{OH \text{ or } P}$ $-(N')_{P}$	$-(N)^{OH \text{ or } P}$ $-(N')_{OH}$	(i) Vector manipulation prior to ligation (ii) End-labelling together with polynucleotide kinase
DNA polymerase I (a) 5′–3′ polymerase (b) 5′–3′ exonuclease (c) 5′–3′ exonuclease	$-N^{OH} + \text{dNTP}$ $-N'.N'.N'.N'_{P \text{ or } OH}$ $^{P}N.N.N.N-$ $N'.N'.N'.N'-$ $_{HO \text{ or } P}$ $-N.N.N.N^{OH}$ $-N'.N'.N'.N'_{P \text{ or } OH}$	$-N.N.N.N.^{OH}$ $-N'.N'.N'.N'._{P \text{ or } OH}$ ^{P}N $_{HO \text{ or } P}N'.N'.N'.N'-$ $-N^{OH}$ $-N'.N'.N'.N'_{P \text{ or } OH}$	(a) + (b) nick translation for *in vitro* labelling
Klenow fragment	Activities (a) and (c) of DNA polymerase I		(i) End-labelling (ii) Removal of 3′ extensions (iii) Conversion of 5′ extensions to blunt ends
Polynucleotide kinase	$-(N)^{P \text{ or } OH}$ $-(N')_{OH}$	$-(N)^{OH}$ $-(N')_{P}$	End-labelling
Terminal transferase	$-(N)^{OH} + \text{dNTP}$ $-(N')_{P \text{ or } OH}$	$-(N)\cdot(N)^{OH}$ $-(N')_{P \text{ or } OH}$	(i) End-labelling (ii) Tailing
Exonuclease III	$-N.N.N.N^{P \text{ or } OH}$ $-N'.N'.N'.N'_{P \text{ or } OH}$	$-N^{OH}$ $-N'.N'.N'.N'_{P \text{ or } OH}$	

select against the propagation of reconstituted vector molecules are also available. These are described in Section 6.3.

6.2.3 Modification of DNA Extremities

In many situations, cloning a given segment of DNA can be greatly facilitated by modification of its extremities. For a variety of reasons it might be desirable to generate cohesive ends, change one type of cohesive end for another or generate blunt ends from cohesive ends.

In cloning random, sheared DNA fragments or double-stranded cDNA into plasmid vectors, modification frequently involves single-strand extension of the polynucleotide chain from the 3′ extremity using terminal nucleotidyl transferase (tailing; *e.g.* Nelson and Brutlag, 1980). This enzyme catalyzes the addition of deoxynucleotide 5′-monophosphates and does not require a template. Tailing of the molecule is carried out using a single deoxynucleotide triphosphate to generate a homopolymer at the 3′ end. Vector DNA carrying 3′ extensions of the complementary homopolymer is synthesized and can then be annealed to the fragment to form non-covalently linked circular molecules. The annealed molecules, while less efficient in establishing themselves than covalently closed circular molecules, can nevertheless be introduced into a suitable host and undergo covalent linkage within the host cell (Wensink *et al.*, 1974). If a suitable choice of homopolymer and vector cleavage site is made, it is possible to regenerate the original recognition sequence at the joints between vector and foreign DNA. This approach is important in the eventual recovery of the cloned fragment. The most common procedure has been to use the restriction enzyme *Pst*I to cleave the vector. Tailing of the vector with dGTP and the fragment with dCTP usually results in regeneration of the *Pst*I recognition sequence at each junction of the hybrid molecule (Figure 2; Villa-Komaroff *et al.*, 1978). If dA:dT homopolymers are employed, conditions of partial denaturation can be found in which the cloned fragment can be isolated by specific cleavage of the polydA:dT duplex regions with endonuclease S1 (Hofstetter *et al.*, 1976). A considerable advantage of the tailing technique is that it prevents recircularization of the vector molecule and ensures the formation of hybrid molecules.

Several additional methods have been developed for modifying DNA extremities. Both 5′ and 3′ cohesive ends can be converted to blunt ends to increase their flexibility in cloning: 5′ extensions can be converted by 'filling in' the duplex strand from its 3′ recessed end using T4 DNA polymerase or *E. coli* DNA polymerase I (Klenow fragment); 5′ and 3′ extensions may be removed by controlled digestion with S1 endonuclease, and 3′ extensions by the 3′–5′ exonuclease activity of the Klenow fragment (Table 2).

Perhaps the most important and most flexible technique available for modifying DNA extremities is the use of oligonucleotide linkers which contain recognition sequences for one or several restriction endonucleases. The most common type of linker is a duplex molecule which can be attached to the extremities of a DNA molecule by blunt end ligation. Once attached, the modified DNA segment is subjected to digestion by the restriction endonuclease for which the linker carries a recognition sequence. This generates specific cohesive ends which enable the fragment, carrying its new extremities, to be ligated into a suitable site in a vector molecule. It is, of course, essential that the inserted piece does not carry a site susceptible to the enzyme or, if present, that the site can be protected (*e.g.* using *Eco*RI methylase together with *Eco*RI). This technique has been used to modify the ends of randomly sheared DNA or double-stranded cDNA in preparation for cloning (Maniatis *et al.*, 1978). It can be used to aid manipulation of cloned DNA fragments and for subsequent mutagenesis (Section 6.5.1). The number of these specific oligonucleotides available at present is large enough to allow a wide choice of extremities and thus allows the investigator to optimize the use of available vector molecules.

6.2.4 Isolation of Recombinant Molecules and Interspecies DNA Transfer

Following ligation, the desired hybrid molecules must be purified from the ligation mixture and propagated. This is accomplished in a first stage by introduction of the mixture into a suitable recipient organism. Since the introduction and maintenance of a given molecule in a host cell requires that the molecule is able to replicate, only those molecules containing the vector will be isolated. If the gene of interest has no selectable character, the resulting clones must be screened further (Section 6.4.1).

Several host–vector systems are at present available. These include various bacterial species, fungi and higher eukaryotes (Section 6.3.4). While certain of these are relatively efficient, and have been used to isolate cloned genes directly, the organism of choice in an initial cloning is generally *Escherichia coli*. Not only have very efficient methods been developed for the introduction of cloned DNA fragments into this host, but the enormous body of information on the genetics of *E. coli* and its bacteriophages and plasmids has resulted in the construction of a wide range of vector molecules. These allow convenient isolation of large quantities of the fragment of interest and subsequent *in vivo* and *in vitro* manipulation of the gene. *E. coli* also has the advantage that problems of restriction, the degradation of unmodified foreign DNA by host-specific enzymes, can be avoided by the use of restriction-deficient mutants as recipient cells. Restriction is often encountered in bacteria and can reduce the probability of introducing a foreign DNA molecule by several orders of magnitude. It should be noted that restriction-deficient mutants of other bacteria, in particular the pseudomonads (Bagdasarian and Timmis, 1982) and *B. subtilis*, are available.

Ligated DNA may be introduced into a cell in two ways: in its native state, in which case the process is referred to as transformation (for plasmid and chromosomal DNA) and transfection (for viral DNA), or, where a viral vector is employed, the DNA may be packaged *in vitro* into mature virus particles and used to infect the host (see also Holt and Saunders, Volume 1, Chapter 5).

6.2.4.1 Transformation and transfection

In *E. coli*, the efficiency of transformation or transfection can be as high as 10^8 transformants per microgram of vector DNA (see Maniatis *et al.*, 1982). The process requires prior treatment of the cells with $CaCl_2$. The treated cells are mixed with the sample and DNA uptake proceeds for about one hour at 4 °C. The final yield of transformants is significantly increased in most cases by a short heat shock (Cohen *et al.*, 1972; but see also Norgard *et al.*, 1978). Following a period of growth, to allow expression of a vector-associated marker, the cells are plated on a selective medium.

Similar protocols have been used with some success for other Gram-negative bacteria, in par-

ticular the pseudomonads (*e.g.* Bagdasarian and Timmis, 1982) and *Klebsiella* (Cannon *et al.*, 1977). In certain cases, addition of RbCl with the $CaCl_2$ has been reported to increase the transformation frequency (Norgard *et al.*, 1978; Bagdasarian and Timmis, 1982; Hanahan, 1983).

Certain Gram-negative bacteria such as *Agrobacterium tumefaciens* seem not to be susceptible to $CaCl_2$-mediated DNA uptake. Since these organisms have a unique relationship with the plants they infect and may prove a powerful tool in introducing cloned DNA into plant cells using their natural plasmid transfer systems (Section 6.3.4), it is important to develop methods for the introduction of cloned DNA into these cells. A treatment which has been used for *A. tumefaciens* is to freeze and thaw the recipient cells. The DNA presumably enters the damaged cells which are then allowed to regenerate (Holsters *et al.*, 1978). This process is inefficient and thus precludes the use of *A. tumefaciens* as a primary host for cloning. Introduction of cloned DNA fragments can be accomplished by cloning using a broad host range plasmid in *E. coli* and subsequent transfer of the plasmid, using natural conjugation and mobilization systems into *A. tumefaciens*. This cross-species transfer has been applied to several Gram-negative bacterial species (Section 6.3.4).

A more general technique which has been applied with success to a variety of organisms including Gram-positive bacteria, fungi and plants, involves the use of polyethylene glycol (PEG) to promote DNA uptake. Here it is necessary to remove all or part of the cell wall to generate protoplasts or spheroplasts prior to PEG treatment. This can, in general, be accomplished by treating the cells with enzymes such as lysozyme, lysostaphin (Gram-positive bacteria), zymolase or snail gut helicase (yeast) which degrade the cell wall (see Hopwood, 1981; Ferenczy, 1981; and Hicks *et al.*, 1978). DNA uptake is promoted by a short exposure to PEG in the presence of $CaCl_2$ and the spheroplasts are then allowed to regenerate. The time of exposure, concentration and molecular weight of the PEG are all critical in optimizing transformation (or transfection) frequencies (Hopwood, 1981). PEG is known to provoke protoplast aggregation and fusion. Indeed, in the case of *Saccharomyces cerevisiae*, it appears that protoplast fusion may be intimately associated with DNA uptake (see Hicks *et al.*, 1978).

With this type of technique, transformation frequencies approaching 10^7 per microgram of vector DNA have been reported for both *B. subtilis* and *Streptomyces* (Chang and Cohen, 1979; see Hopwood, 1981 and Chater *et al.*, 1982) using plasmids of between 5 and 30 kb. For yeast, values of 10^4 per microgram of DNA have been obtained (Kingsman *et al.*, 1979).

A modification of this procedure, which has been reported to increase transformation frequencies in *Streptomyces* (Makins and Holt, 1982), plants (Lurquin, 1979) and animal cells, is to add the transforming DNA to a dried phospholipid film to generate artificial phospholipid vesicles or liposomes. The liposomes are then mixed with cell protoplasts and treated with PEG (see Fraley and Papahadjopoulos, 1982 for an extensive discussion).

Transformation and transfection techniques have also been developed for animal cells. Two methods are currently employed for DNA delivery to large numbers of cells. The first involves the direct addition of a calcium phosphate/DNA precipitate to growing cells (Wigler *et al.*, 1979). In the second, used primarily for transfection with SV40-based vectors (Section 6.3.4), DNA uptake is promoted by DEAE–dextran in the culture medium (*e.g.* Sompayrac and Danna, 1981). A third technique, but one limited to the delivery of DNA to a small number of cells, is microinjection. In some cases microinjection has proved to be an extremely efficient means of introducing cloned DNA into animal cells (see Colbère-Garapin *et al.*, 1982).

An efficient method for introducing cloned DNA into a given host is essential in the primary isolation of a rare DNA segment. Less efficient systems can be tolerated when the cloned DNA is to be introduced in a pure form into a secondary host after initial isolation and propagation in a primary host. In addition to transformation, other methods of interspecies DNA transfer have been developed. Potentially the most general technique involves PEG-induced protoplast fusion directly between bacterial cells carrying the cloned DNA and the recipient cells. This technique has been used to transfer DNA cloned and propagated in *E. coli* into yeast (Kingsman *et al.*, 1979) and animal cells (*e.g.* Rassoulzadegan *et al.*, 1982). Very high efficiency fusion between various cell types has also been obtained by electro-pulse techniques (Zimmermann *et al.*, 1981). These techniques may prove valuable in the future.

6.2.4.2 In vitro *packaging*

One problem encountered in the use of transformation and transfection techniques is that while the frequency per microgram of vector DNA can be quite high, the overall efficiency of

introducing and establishing a viable hybrid molecule is quite low (*e.g.* 10^{-4} transformants per viable cell in the case of *E. coli*). The use of cosmids and vectors based on the bacteriophage λ (Section 6.3.2) allows an alternative procedure to be employed. DNA in the ligation mixture can be packaged directly into phage particles *in vitro*. Once the phage particles are assembled, they may be used to infect a suitable *E. coli* host. Packaging followed by infection can give a frequency two orders of magnitude higher than transfection using native DNA (Enquist and Sternberg, 1979). The packaging reaction employs a mixture of complementing bacterial lysates obtained from two λ lysogens in which each λ prophage carries a mutation in a different gene involved in phage assembly. When provided with a suitable DNA preparation, the mixed lysate is capable of assembling phage particles *in vitro*. In order to optimize the packaging reaction and to reduce the probability that phage present in the bacterial lysates will contribute to the final *in vitro* packaged phage particles, mutant prophages are used and the lysates can be irradiated with UV light prior to mixing with the ligated DNA (Enquist and Sternberg, 1979; Hohn, 1979; Collins, 1979).

6.3 CLONING VECTORS

The choice of vector is an important consideration in designing experimental strategies for cloning. The number and variety of available vectors, however, preclude a comprehensive discussion here. The following is, therefore, simply intended to demonstrate the useful features of the various types of vector available. For more detailed information, the reviews of Bernard and Helinski (1980), Timmis (1981) and Williams and Blattner (1980) should be consulted.

DNA molecules to be used as cloning vectors should be relatively small, autonomous replicons which are physically and genetically well characterized. They must carry an easily selectable character for introduction into a host and should be simple to isolate in large quantities. Furthermore, insertion of a foreign DNA fragment into the vector must not disrupt its replication functions or the expression of the selectable character. Additional properties which are desirable but not essential include: the presence of single recognition sequences for the restriction enzymes to be used in cloning, the availability of a means to identify recombinant molecules, and the presence of suitable vector-associated promoters which will allow transcription of the cloned fragment. The presence of vector-associated promoters is important in cases where the endogenous promoter is not recognized by the primary host or where it has been separated from the gene during cloning. Both bacterial viruses and plasmids have been modified to fulfil some or all of these criteria and, in certain cases, analogous replicons from other organisms have also been successfully harnessed.

Initial cloning is usually carried out using *E. coli* vectors together with an *E. coli* host. These vectors are of three basic types: plasmids, bacteriophages and cosmid vectors which are maintained as plasmids but which can be packaged into phage particles *in vitro* and then used to infect the primary host. Each type of vector has specific properties which lend themselves to different experimental problems.

Many of these vectors have been modified to permit more detailed analysis of a DNA fragment or the isolation of DNA with particular functions. Vectors are available which allow optimization of gene expression, facilitate DNA sequencing or permit the isolation of transcription and translation signals, or signal sequences involved in replication.

6.3.1 Plasmid Vectors

The plasmid vectors most widely used in *E. coli* are derived from the small, multicopy plasmids pMB9, p15A and colE1. They possess the important property of continued replication in the absence of protein synthesis, in contrast to the host chromosome. Treatment of the host cells with chloramphenicol, or other antibiotics which inhibit protein synthesis, results in a large increase in the number of copies of the plasmid and a concomitant increase in the proportion of plasmid DNA in the culture (amplification; Clewell and Helinski, 1972). A very significant advantage in the use of pMB9 derivatives such as pBR322 and pBR325 (Bolivar *et al.*, 1977a, 1977b; Bolivar, 1978) is that the entire nucleotide sequence of these plasmids is known (Sutcliffe, 1979; Prentki *et al.*, 1981). This together with the precise location of the plasmid promoters (West and Rodriguez, 1980; Stüber and Bujard, 1981; Queen and Rosenberg, 1981) greatly facilitates physical studies of cloned DNA fragments. Extensive knowledge of the organization of pBR322 and

pBR325 has also led to the generation of deletion derivatives which exhibit even higher copy numbers (Twigg and Sheratt, 1980; Soberon *et al.*, 1980; Covarrubias and Bolivar, 1982).

A second type of plasmid vector is based on low copy number plasmids such as F and pSC101. These differ from those described above in being unable to undergo amplification in the absence of host protein synthesis. This kind of plasmid vector is useful in cases where the product of the cloned gene might be deleterious to the host cell, since the gene will be maintained in fewer copies per cell than when cloned in a multicopy plasmid, and might be expected to produce less product.

A third type of vector has been constructed from 'runaway' mutants of the low copy number plasmid R1 (Uhlin *et al.*, 1979; Bittner and Vapnek, 1981). At low temperature these plasmids exhibit the wild-type copy number while at elevated temperatures replication is no longer controlled and a large amplification of plasmid DNA occurs. This eventually results in cell death. The use of such a vector should also allow the cloning and expression of genes which are normally deleterious to the host.

All these plasmids, however, have relatively restricted host ranges. Other plasmids with broader host ranges allowing eventual propagation in a variety of bacterial species have been developed as vectors and are described below.

The size of DNA fragments that can be cloned into plasmid vectors is limited only by the reduction in transformation frequency associated with increased plasmid size (see Hanahan, 1983).

Many plasmid vectors have been designed to allow easy detection of hybrid molecules. This is important since the methods available for ensuring the formation of hybrid molecules at the ligation step in the cloning procedure (Section 6.2.2.2) are often only partially effective. Commonly these vectors specify resistance to at least two antibiotics. The resistance genes ideally carry unique sites for restriction enzymes commonly used in cloning. One of these characters is employed to select for the introduction of the plasmid into the host cell. The other can be employed to detect hybrid molecules since cloning within an antibiotic resistance gene generally inactivates the gene. Following ligation and transformation, clones are normally selected by resistance to one antibiotic and then screened for sensitivity to the second antibiotic by replica plating. Clones sensitive to the second antibiotic should all carry recombinant DNA molecules in which the resistance gene has been interrupted by insertion of the cloned piece. This procedure is known as insertional inactivation (see Timmis, 1981). In several cases screening can be performed directly on the primary transformants by use of indicator plates. Thus inactivation of the chloramphenicol transacetylase gene (Figure 1) can be scored by including crystal violet in the selective plates (Proctor and Rownd, 1982) and clones in which the β-lactamase gene has been interrupted can be scored using the β-lactam nitrocephin or starch/iodine (O'Callaghan *et al.*, 1972; Boyko and Ganschow, 1982).

An analogous method based on the interruption of a cloned β-galactosidase gene has also been developed, using a derivative of pBR322 (pUR222; Rüther *et al.*, 1981) which carries a cloned fragment of the *lacZ* gene of *E. coli*. The cloned *lacZ* fragment carries several unique restriction sites at its 5′ extremity. Certain *lacZ*$^-$ host cells are complemented by the protein fragment synthesized by the plasmid to produce active β-galactosidase and form blue colonies on selective media containing the chromogenic substrate 5-bromo-4-chloroindoyl-β-D-galactosidase (X-gal). Insertion of DNA into the *lacZ* gene fragment stops synthesis of the protein fragment and, in such a host background, results in the production of white colonies.

Inactivation of an antibiotic resistance gene can also be used to enrich for recombinant clones *prior* to plating the transformation mixture. If the action of the antibiotic is bacteriostatic, *i.e.* prevents the growth of the cells (*e.g.* tetracycline), the transformation mixture can be treated with this antibiotic and subsequently with a second antibiotic whose action is bacteriocidal but which kills only growing cells (*e.g.* ampicillin or cycloserine). Cells carrying clones in which a tetracycline resistance gene has been inactivated will not grow in the presence of the drug. Those clones which have retained tetracycline resistance will continue to grow and will be killed by the action of the second antibiotic, resulting in an overall enrichment for recombinant clones (Bolivar *et al.*, 1977a). Methods also exist for the direct selection of inactivation of both tetracycline (Malloy and Nunn, 1981) and kanamycin resistance genes (Slutsky *et al.*, 1980).

Frequently, complementation between different cloned fragments must be examined. This cannot be easily achieved when the two fragments are cloned in identical plasmid vectors. In the absence of a selection for each plasmid, two plasmids with the same replication system will not be stably maintained but will segregate, giving rise to a mixed population of cells that contain one or the other plasmid. This phenomenon is known as incompatibility. Two derivatives of p15A,

pACYC177 and pACYC184, have been constructed and are compatible with pMB9 derived plasmids (Chang and Cohen, 1978). Other vectors are also available which circumvent this problem.

6.3.2 Vectors Derived from Bacteriophage λ

6.3.2.1 Phage vectors

The attraction of bacteriophage λ as a vector results from the fact that 40% of the phage genome (20 kb) is inessential for phage growth, that reconstituted phage particles can be efficiently prepared *in vitro* and that once formed, recombinant molecules can be propagated as phage particles *in vivo*. This allows the cloning of large fragments of foreign DNA and efficient introduction of the hybrid molecules into the primary *E. coli* host, and facilitates the isolation and handling of large numbers of independent clones. This is particularly important in the construction of genomic libraries of higher eukaryotic cells which may require the propagation of at least 10^7 independent clones. In addition, size constraints imposed on the phage genome in both *in vivo* and *in vitro* packaging to produce infectious phage particles (see below) present a powerful selective technique for ensuring the isolation of recombinant phages. These properties have made λ-based vectors important in the construction of genomic libraries. An additional advantage of λ vectors stems from the fact that most of the inessential DNA is located in a single region of the genome whose transcription is directed by a strong phage promoter P_L. Foreign DNA cloned in the inessential region would therefore be expected to undergo transcription from the phage promoter. Finally, phage vectors are ideal for cloning genes whose continuous expression adversely affects the growth of the host cells because the cloned fragment is not continuously propagated within the host cell.

Two basic types of λ vector have been developed: (i) insertion vectors which carry a deletion in the phage genome and allow the insertion of foreign DNA into a unique restriction site without the concomitant removal of phage DNA; and (ii) replacement vectors in which an inessential region of the phage genome, flanked by suitable restriction sites, can be replaced by the foreign DNA fragment. Useful variants of the replacement vectors have been constructed in which the replaceable fragment carries genes whose inactivation or loss permits the direct selection or screening of the recombinant clones. An extensive list of λ vectors can be found in the article by Williams and Blattner (1980).

For both types of vector, the final recombinant genome must be between 39 and 52 kb (78% and 105% of the wild-type λ genome) if they are to be packaged into infectious particles. Insertion vectors must therefore be at least 39 kb in length to maintain their viability. This places an upper limit of about 12 kb on the size of foreign DNA fragments which can be inserted. Replacement vectors have a larger capacity because the entire inessential region can be replaced, allowing the cloning of fragments up to 22 kb.

A disadvantage in the use of λ, as with other vectors, is that a substantial proportion of vector molecules can be reconstituted during ligation. With λ replacement vectors, this problem can be avoided in two ways. Since λ has a linear genome, the two fragments carrying essential functions can be purified away from the replaceable fragment by physical methods, such as sucrose gradient centrifugation, prior to ligation to the foreign DNA. The purified fragments rather than the entire phage can then be employed in ligation. Purification of the vector fragments can be avoided if the replaceable fragment carries functions that prevent plaque formation (lytic growth) of the reconstituted vector on particular host strains. Only recombinant phages of the correct size and in which the fragment has been replaced will form plaques, making this type of λ vector the most useful for constructing genomic libraries.

The observation that wild-type λ phages are unable to form plaques on bacteria lysogenic for phage P2 ('Spi^+': sensitivity to P2 inhibition) permits a powerful selection of this type. The inhibition of plaque formation is due to the expression of the λ *red* and *gam* genes. These functions are located in the inessential region of λ, and vectors have been constructed where the genes are replaced during the cloning procedure (Loenen and Brammer, 1980). Recombinant phages that have lost the *red* and *gam* genes are selected on a P2 lysogenic strain.

Several types of vector have been developed which allow direct screening for recombinant phages and are useful for cloning specific DNA fragments. Many of these vectors synthesize β-galactosidase, and form blue plaques on X-gal media. Cloning DNA fragments into the vector

prevents the synthesis of β-galactosidase and recombinant phages form colourless plaques on *lacZ*$^-$ bacteria which can be easily detected. Many different screening and selection procedures can be used with λ vectors and are described in detail by Williams and Blattner (1980).

The use of λ vectors generally entails the use of the λ *in vitro* packaging system and the delivery of the hybrid molecules by infection of the primary *E. coli* host. The λ packaging system normally requires concatamers of phage DNA as a substrate. Such concatamers are a normal product of phage replication in the host cell. A specific 12 bp sequence (the *cos* site) is recognized by the packaging machinery and DNA is presumably drawn into the phage head until a second *cos* site is reached. The recognition of the second site results in cleavage of the λ DNA. The smallest DNA molecule which can be packaged will therefore contain two *cos* sites separated by approximately one genome's length of DNA. The *cos* sites must occur in the correct orientation. For this reason, any ligation mixture in which the delivery to the primary *E. coli* host relies on *in vitro* packaging must be sufficiently concentrated to generate concatameric molecules. This constraint applies both to λ phage vectors and to cosmids (see below).

6.3.2.2 Cosmid vectors

Cosmids are plasmids that can be packaged into λ phage particles by virtue of the presence of a cloned λ *cos* site (Collins, 1979; Hohn and Hinnen, 1980). Any plasmid molecule that contains a *cos* site can be packed into a phage particle provided that it is large enough to overcome the size constraints of the λ packaging system. Thus cosmids have advantages of both plasmid and bacteriophage vectors since they can be delivered to the host by the more efficient infection procedures rather than by transformation. Since phage heads will accept about 50 kb of DNA, small cosmid vectors can be used to clone very large fragments of DNA. For example, a 5 kb cosmid (approximately the size of many of the useful plasmid vectors described above) cannot be packaged as a monomeric unit unless it contains between 40 and 45 kb of additional DNA. This is a significant increase in capacity compared to λ phage vectors and reduces the number of clones required to generate a genomic library. Recently, an elegant system of cosmid cloning has been described by Ish-Horowicz and Burke (1981). Normally, ligation produces a random mixture of different polymers of vector and insert DNA. Only a small fraction of these have the correct structure to be packaged. By consecutive use of alkaline phosphatase (Table 2; Section 6.2.2.2) and suitable restriction enzymes for which the vector carries unique sites, the only ligation products generated using this technique are hybrid molecules which carry two *cos* sites and a single inserted foreign DNA fragment.

6.3.3 Special Purpose Cloning Vectors

6.3.3.1 Expression vectors

Once a gene has been cloned and identified, subsequent steps will probably involve its recloning into a secondary vector so that its transcription is directed by a strong vector-associated promoter. The *E. coli lac*, *trp* and the λ P_L promoters have all been employed for this purpose. Various vectors have been developed in which it is possible to insert DNA fragments downstream from these promoters (Mercereau-Puijalon *et al.*, 1978; Hallewell and Emtage, 1980; Remaut *et al.*, 1981). An additional problem in maximizing expression of cloned genes in *E. coli* which is frequently encountered with genes from a heterologous source is that the gene carries no translation start signal which can be efficiently recognized by the *E. coli* translation system. This problem may arise for heterologous genes cloned into any host. Thus even though the gene can be transcribed from a promoter within the *E. coli* vector, the resulting mRNA is poorly translated and little or no protein product will be synthesized. In such cases alternative strategies must be employed. One is to fuse the gene to the amino-terminal portion of a vector gene that is efficiently translated in the host. Both translation and transcription start signals are thus provided by the vector-associated gene. The fusion must be engineered in the correct reading frame and preferably as few amino acids as possible deleted from, or added to, the cloned gene. Suitable vectors have been developed to enable fusion of a cloned gene to β-galactosidase (Charnay *et al.*, 1978), anthranilate synthetase (Tacon *et al.*, 1980) and β-lactamase (Talmadge and Gilbert,

1981). The β-lactamase fusion vectors are also useful in cases where secretion of the product by the *E. coli* host is desirable (Section 6.5.2).

An alternative strategy is to couple the coding region of the cloned gene to a DNA segment carrying both a strong promoter and a ribosomal binding site (Guarente *et al.*, 1980; Panayotatos and Truong, 1981; see Section 6.5.2).

6.3.3.2 Single-stranded phage vectors

An interesting class of *E. coli* vector has been derived from the single-stranded DNA phages fd and M13 (see Barnes, 1980). These phages do not kill the host cell but are continuously secreted into the culture medium allowing simple isolation of DNA from the phage. Cloning into these vectors requires the isolation of double-stranded replication intermediates from the host cell. Various antibiotic resistance genes have been cloned into the vectors to allow simple selection for their introduction into the host cell. The *lac* gene has also been used for detection of recombinant molecules in the same way as with other vectors.

Because isolation of the phage DNA is so simple, use of these vectors is an excellent method for the preparation and purification of single strands of cloned DNA for sequencing studies, mutagenesis (Section 6.5.1) and analysis of replication properties (Ray *et al.*, 1982). The most common use of single-stranded phage vectors is in sequencing large regions of DNA. Short fragments of DNA are cloned at random. Sequencing is carried out by annealing a single-stranded fragment complementary to the phage. This is derived from either a purified phage restriction fragment or a synthetic oligonucleotide. The annealed fragment then serves as a primer in the chain termination sequencing method (Sanger *et al.*, 1977; see Smith, A. J. H., 1980). The overall sequence can be compiled from a series of overlapping clones using established computer scanning techniques. Sequencing of both strands of a given fragment can be simply accomplished by cloning the fragment in both orientations within the vector.

Fragments of DNA cloned in single-stranded vectors should be ideal hybridization probes for colony or plaque screening, or for hybridization to 'Southern' or 'Northern' blots (Sections 6.4.1 and 6.4.2.2), since there is no complementary probe strand to compete in the hybridization reactions. This competition can, for example, make it impossible to detect rare mRNA species in such experiments. It has been difficult to obtain sufficiently high specific activities of radioactive labelling of phage DNA, but techniques are now available that circumvent this problem (Brown *et al.*, 1982).

6.3.3.3 Plasmid vectors for subcloning and sequencing

A major advantage in the use of single-stranded phage vectors in DNA sequencing is that the DNA needs no extensive purification. Recently, two plasmid vectors have also been developed which permit DNA sequencing without extensive purification of radioactively labelled DNA. One, pHP34 (Prentki and Krisch, 1982), derived from pBR322, allows the cloning of any blunt end DNA fragment with the possibility of its easy recovery. Fragments cloned in this vector can be sequenced directly by the chemical method of Maxam and Gilbert (1980). The second plasmid, pUR222 (Rüther *et al.*, 1981; Section 6.3.1) has similar properties. An improved version of pUR222, pUR250, permits sequencing of both strands of the cloned fragment (Rüther, 1982). These plasmids are ideal vectors for the rapid sequencing protocol described by Guo and Wu (1982) which uses the chain terminator method rather than the chemical method.

6.3.3.4 Vectors for the detection of transcription and translation signals

Two series of *E. coli* plasmid vectors have been described that permit the detection of promoters and transcription termination signals (Casabadan *et al.*, 1980; McKenny *et al.*, 1981). Both types of vector contain genes that are not transcribed and therefore not expressed unless a promoter is cloned into a suitable site on the plasmid. One carries the gene for β-galactosidase, and the other a gene encoding galactokinase. The levels of expression of both these genes affect the colour of colonies growing on indicator media facilitating the isolation and subsequent mutational analysis of promoter sequences. A plasmid that expresses β-galactosidase or galactokinase can be used to isolate and analyze sequences involved in the termination of transcription. Apart

from their utility in the analysis of transcription in *E. coli*, these plasmids can also be used to isolate DNA sequences from other organisms that function as promoters in *E. coli* (*e.g.* Casabadan and Cohen, 1980).

The vectors facilitate the analysis of the control of expression of those genes whose products cannot be easily assayed. Not only do the β-galactosidase vectors allow analysis of transcriptional control (by isolating promoters and terminators) but they also permit the fusion of coding sequences in a cloned DNA segment with a DNA sequence encoding β-galactosidase. This results in the production of a hybrid protein that retains β-galactosidase activity and therefore allows investigation of regulation at the translational level.

6.3.4 Vector Systems for Organisms other than *E. coli*

Establishment and maintenance of cloned genes in organisms other than *E. coli* represents an important aspect of recombinant DNA technology. The ability to introduce genes into an homologous host allows the study of their expression in a natural environment. Moreover, the availability of a range of host–vector systems permits the study of a given gene in a variety of species. Such studies should provide essential information on gene expression and aid in the optimization of production of specific gene products.

Techniques which allow the introduction of a gene by integration into the host genome following transformation have been available for several years both for bacteria and higher organisms. While these methods can be useful in strain improvement they have limited utility. In cases where there is no direct selection, the gene of interest must be delivered to the host in a relatively pure form, preferably following an initial cloning in a more suitable host–vector system, together with DNA containing a selectable genetic marker to select transformed clones (*e.g.* Wigler *et al.*, 1979; Colbère-Garapin *et al.*, 1982). A disadvantage of this method is that once integrated, the recovery of the gene is not always simple. In addition, integration is not the method of choice where high yield of a gene product is required since, in general, the integrated gene will be present in only a few copies per cell.

In view of these considerations it is frequently more useful if the cloned DNA can be introduced into the host cell as part of an autonomously replicating molecule. Since initial cloning is usually performed in an organism such as *E. coli* where well-developed vector systems are available, a useful feature is to design a vector so that it can be propagated in more than one host. This is generally obtained by the *in vitro* fusion of two replicons (see below). While not yet as sophisticated as those of *E. coli*, vector systems exist for a variety of organisms and are being rapidly improved.

For several Gram-negative bacterial species, a fruitful approach has been the use of broad host range plasmids. Two types of plasmid vector are being employed at present: (i) those based on the small high copy number plasmid RSF1010 (Bagdasarian *et al.*, 1981) and (ii) those based on low copy number plasmids of the P incompatibility group such as RP4 and RK2 (Kahn *et al.*, 1979). A significant advantage of RSF1010-based vectors is their apparent ability to undergo spontaneous amplification as the host cells reach stationary phase (J. Frey and M. Bagdasarian, personal communication). This facilitates isolation of large quantities of cloned DNA. In addition, while RSF1010 is itself non-transmissible, it can be readily mobilized into other cells in the presence of a second, conjugal plasmid such as RP4 (Willetts and Crowther, 1981). These vectors are therefore more versatile than the colE1 related vectors of *E. coli*. They can be used for direct cloning in a variety of species, or, where the transformation system in a particular species is inefficient, cloning can be carried out in *E. coli* and the resulting plasmids susbsequently mobilized into the species of choice (Bagdasarian and Timmis, 1982).

Various derivative plasmids have been constructed, each carrying a selectable antibiotic resistance marker and single sites for several restriction endonucleases. In many cases, recombinant plasmids can be recognized by insertional inactivation, as for *E. coli* vectors (Section 6.3.1). Cosmid vectors allowing for efficient cloning and introduction into *E. coli* using the *in vitro* packaging system of bacteriophage λ have also been constructed (Bagdasarian *et al.*, 1981; J. Frey, personal communication).

Vectors based on the *inc*P group plasmids have provided another useful system. Since these relatively large plasmids specify their own conjugation system, direct cloning into such plasmids allows simple *in vivo* transfer into a secondary bacterial host. In addition, binary vector systems composed of two derivatives of the plasmid have been developed. A mini-plasmid, deleted for

the transfer functions of the parent, is used as a vector. Transfer to an alternative host can be accomplished by the presence of a second plasmid or phage carrying the cloned transfer functions in the same cell (Ditta *et al.*, 1980). Cosmid derivatives of the *inc*P plasmids have been constructed and used with success (Klee *et al.*, 1982).

Progress has also been made in the development of vectors for Gram-positive bacteria. Several low and high copy number plasmids which replicate in both *Staphylococcus aureus* and *Bacillus subtilis* have been modified for use as vectors. As for many of the vectors discussed so far, antibiotic resistance is generally employed as the vector-associated marker. Some vectors carry cloned *Bacillus* genes which can be selected by complementation of a suitable auxotrophic marker in the recipient strain (see Kreft and Hughes, 1982 and Ehrlich *et al.*, 1982). Insertional inactivation can be used in many cases to screen for recombinant plasmids. Certain of these plasmids can be amplified 4–8 fold by treatment with hydroxyurea and in the case of one vector, pUB110, amplification to about 10^3 copies per cell can be achieved by propagation at the non-permissive temperature in strains carrying a temperature-sensitive mutation in chromosomal DNA replication (Ehrlich *et al.*, 1982).

One important problem encountered in the present *B. subtilis* cloning systems is the relatively high instability of the vectors and inserts. This instability derives from two sources: (i) segregation of the plasmids during growth of the cells (a problem which may be overcome by suitable selective pressure), and (ii) deletion or other rearrangements of the cloned DNA. The reasons for the molecular instability are at present unclear. One way in which this phenomenon can be partially overcome, at least operationally, is by the use of shuttle vectors. These are replicons able to propagate in two alternatives hosts. Vectors able to replicate in both *B. subtilis* and *E. coli* have been constructed by *in vitro* fusion of a *B. subtilis* plasmid with the well-characterized *E. coli* plasmid pBR322. They allow the initial cloning and manipulation to be carried out in *E. coli*, where they are stable, and can subsequently be introduced into a suitable *B. subtilis* strain. This type of vector may be superseded, however, by the discovery of small plasmid derivatives which are able to replicate efficiently in *Streptococcus pneumoniae, B. subtilis* and *E. coli* (Baranay *et al.*, 1982). One shuttle vector has been constructed for DNA sequencing. This vector is a hybrid between a *B. subtilis* plasmid and the single strand DNA phage, M13, of *E. coli*. The hybrid replicates as a double-stranded plasmid molecule in *B. subtilis* and can be packaged as a single-stranded phage in *E. coli* (see Kreft and Hughes, 1982). While plasmid vectors for *Bacillus* are relatively well advanced, little use has been made as yet of *Bacillus* phages as vector molecules.

For the industrially important genus *Streptomyces*, both low and high copy number plasmids are available as vectors. Certain of these have a relatively broad host range (Chater *et al.*, 1982). Derivative plasmids have been constructed which allow direct selection for an antibiotic resistance marker and in certain cases allow screening for recombinant molecules by insertional inactivation. While *Streptomyces* phages have not been exploited to the same extent as plasmid vectors, use has been made of a shuttle vector composed of the broad host range *Streptomyces* phage, ϕC31, and pBR322. When grown as a phage in *Streptomyces*, the packaging constraints of ϕC31 can be exploited, in the same way as those of coliphage λ, to generate deletions (Chater *et al.*, 1982).

Amongst the lower eukaryotes, the yeast *S. cerevisiae* has the most developed cloning system. Two classes of autonomous vector have been developed. They are based either on the 2 μm plasmid (see Hollenberg, 1982 and Hinnen and Meyhack, 1982) or on autonomously replicating segments of chromosomal DNA (*ars*). Many of the autonomous vectors also contain an *E. coli* vector plasmid such as pBR322. Cosmid derivatives of such shuttle vectors are available (Hohn and Hinnen, 1980).

Like many *B. subtilis* vectors, yeast vectors based on the 2 μm plasmid are unstable and undergo molecular rearrangements and segregation. Their degree of stability and copy number vary according to the particular vector, the inserted sequence and the presence of endogenous 2 μm plasmids in the recipient cells. The frequency of transformation is, in general, higher than with other types of vector, presumably reflecting relatively high copy numbers (the parent has a copy number of about 50–100 per cell). The other class of high frequency transforming vectors is based on autonomously replicating yeast genomic sequences (*ars*) cloned in an *E. coli* plasmid vector. The plasmids are quite unstable and segregate rapidly in yeast but can be stabilized by the insertion of a yeast centromere. They then behave as a chromosome during meiosis and mitosis (see Hinnen and Meyhack, 1982).

Most yeast vectors employ a yeast biosynthetic marker for selection purposes. Recipients are, therefore, limited to those strains which carry the relevant auxotrophic mutation. More recently,

use has been made of bacterial chloramphenicol and gentamycin G418 resistance as selective markers (Hollenberg, 1982; Jimenez and Davies, 1980).

An expression vector based on the 2 μm plasmid has been developed that carries the control region of the yeast phosphoglycerate kinase gene. This has been shown to promote high level expression of human α-interferon in yeast (Tuite *et al.*, 1982).

Ars sequences from other organisms have been isolated by cloning into yeast and may prove useful as a basis of vector construction in their homologous host although initial results with an *ars* locus derived from *N. crassa*, indicate that the 'plasmid' integrates when reintroduced into its parent organism (Vapnek and Case, 1982).

For mammalian cells, the most widely used autonomously replicating vectors are based on the papova virus SV40. This is perhaps the best characterized of all the animal viruses and has many useful features (see Elder *et al.*, 1981; Hamer, 1980; Gruss and Khoury, 1982). Its genome is composed of a small circular DNA molecule. It can be propagated by lytic infection of permissive cells such as African Green Monkey Kidney (AGMK) cells to yield up to 10^5 copies per cell (Hamer, 1980). In non-permissive cells, SV40 undergoes integration into the host genome thus maintaining a low copy number of any cloned gene (see Elder *et al.*, 1981). The virus can be recovered subsequently by fusion of the 'transformed' non-permissive cells with a permissive cell line (Hamer, 1980). In AGMK cells, the lytic cycle occurs in two phases which involve initial expression of the early functions followed by high level expression of the late functions. In non-permissive cells only early functions are expressed. The regions coding for these functions are located in opposite halves and transcribed in opposite directions on the circular genome. Vectors have been constructed in which the cloned DNA can replace either the early or late regions. As long as the hybrid molecule represents between 70% and 105% of the viral genome in length, it can be isolated and maintained as a viral stock by growth with a complementing helper virus. For production of encapsidated virus, the present vectors are limited to the inclusion of 2.5 kb of cloned DNA. Where it is unnecessary to obtain amplification of a viral stock or reinfection of a secondary host, other derivative vectors are available. These include both defective SV40 molecules which can be propagated in cells engineered to produce constitutively the SV40 T-antigen necessary for replication of the virus, and shuttle vectors composed of part of the SV40 genome combined with an *E. coli* plasmid such as pBR322 (see Elder *et al.*, 1981). Such molecules can also be combined with selectable genetic markers. Examples include the thymidine kinase gene of *Herpes simplex* virus and bacterial neomycin phosphotransferase which specifies resistance to the antibiotic G418 (see Colbère-Garapin *et al.*, 1982).

The use of SV40 vectors to express cloned eukaryotic genes is well advanced. Many of the controlling elements involved in viral gene expression have been precisely located on the SV40 genome and have been successfully harnessed for the expression of foreign genes (Gruss and Khoury, 1982).

Recently an autonomously replicating vector based on bovine papilloma virus has been developed and used to promote expression of cloned human β-globin and β-interferon in mouse cells (Dimaio *et al.*, 1982; Zinn *et al.*, 1982). At the time of writing, other vector systems based on adeno and murine sarcoma viruses are being developed (see Gluzman, 1983).

An alternative to viral vectors for introducing cloned genes into eukaryotic hosts is to use a transposable element. Rubin and Spradling (1982) have shown that it is possible to introduce exogenous DNA into *Drosophila* embryos using the P element as a vector.

Of all the important groups of organisms which have received attention to date, cloning systems in plants are the least well developed. There are, however, several systems under investigation.

Only two classes of the known plant viruses, the gemini viruses and the caulimoviruses, have a DNA genome and are, therefore, potential vectors. The small double-stranded circular DNA genome of cauliflower mosaic virus (CMV) has received the most attention among these (Hohn *et al.*, 1982) although host range is restricted to members of the Cruciferae.

Another potential class of autonomous vectors comprises sequences equivalent to the genomic *ars* sequences of yeast (see above). Such elements, able to replicate in yeast, have been isolated from *Zea* maize (Stinchcomb *et al.*, 1980) but it is as yet unclear whether they are able to replicate autonomously in cells cultured from the parental plant. *Chlamydomonas rheinhardii* genomic segments, able to replicate in the parental cells, have recently been isolated (J.-D. Rochaix, personal communication) and may provide the basis for a useful vector system.

Perhaps the most unusual method for introducing foreign DNA into plants (both dicotyledons and gymnosperms; see Nester and Kosuge, 1981) stems from the observation that segments (T-DNA) of certain plasmids carried by the genus *Agrobacterium* are transferred from the bacterium

into the plant cell where they become integrated into the plant genome. The best studied of these are the Ti (tumour inducing) plasmids of *Agrobacterium tumefaciens*, the causative agent of Crown Gall disease. Similar observations have been made in the case of a plasmid carried by *A. rhizogenes* and which induces 'hairy root' disease. While many problems remain to be solved before this integrating vector can be used as a vehicle for routine genetic manipulation in plants, several successful attempts have been made to introduce foreign DNA cloned into the integrating sequences. Cloning into these plasmids is complicated by their large size. A system has been described in which DNA is first cloned into a suitable shuttle vector, and then subsequently inserted into the Ti plasmid by recombination *in vivo* (Leemans *et al.*, 1982). This has been used to promote the expression of chloramphenicol acyltransferase in plant cells (Herrera-Estrella *et al.*, 1983). Ti plasmid DNA may be introduced into plants cells by the PEG/Ca^{2+} method (Krens *et al.*, 1982), or by the use of liposomes (see Fraley and Papahadjopoulos, 1982), and the T-DNA undergoes integration and is expressed.

6.4 DETECTION AND ANALYSIS OF CLONES

Following their introduction into a host cell, hybrid molecules must be detected and characterized. Detection may be accomplished by selection or screening of transformant clones by physical, chemical or biological methods. Their subsequent characterization requires efficient methods for the isolation of recombinant DNA molecules, for their physical analysis, and in many cases to examine their ability to direct the synthesis of a given gene product.

6.4.1 Screening Recombinant Clones

The simplest procedure for clone identification is direct selection. This is normally only possible when the cloned gene is expressed and complements a defect in the new host. Since each complementation is rarely possible, indirect methods must generally be used.

As described in Section 6.3, many vector systems have been designed to ensure, as far as is practicable, that only recombinant clones are recovered. Once isolated, a set of recombinant clones can sometimes be screened and the desired clone identified simply by size or restriction pattern. More usually, alternative methods are required. Screening procedures based on nucleic acid homology, *in vitro* translation, and immunodetection of a protein product have been developed and successfully employed.

6.4.1.1 Nucleic acid homology

Screening using nucleic acid homology requires that a specific DNA or RNA probe, complementary to the desired DNA fragment, is available. This need not necessarily be highly purified provided that the mixture gives the required specificity of hybridization. The probe may be a radioactive DNA fragment, mRNA labelled *in vitro* (Maniatis *et al.*, 1976) or a radioactive cDNA copy of the mRNA (Rougeon *et al.*, 1975). Labelled DNA probes can be prepared by 'nick translation' (Rigby *et al.*, 1977) using *E. coli* DNA polymerase I or by 'end labelling' (Table 2). Where no such probe can be made an alternative approach is to synthesize chemically a complementary oligonucleotide. This requires that some part of the protein or nucleotide sequence of the gene is already known. Mixtures of oligonucleotides that derive from partial protein sequence data have been used to screen cDNA libraries and show great reliability in the detection of rare clones (*e.g.* Edlund *et al.*, 1983).

Such probes can be used to screen clones by hybridization, either to bacterial colonies (Grunstein and Hogness, 1975), or to phage plaques (Benton and Davis, 1977). The colonies or plaques are replicated from a Petri dish onto nitrocellulose or paper filters and treated with alkali, which in colony hybridization lyses the bacteria, and in both cases denatures the DNA. The DNA is then baked onto the filter. Following hybridization with the probe, the filters are autoradiographed and compared with the original plate to detect those clones exhibiting homology. Under optimal conditions, initial screening can be performed at high density (10^5 colonies per plate; Hanahan and Meselson, 1980). Examples of the methods used to maintain and screen large numbers of clones are given by Maniatis *et al.* (1978) and Gergen *et al.* (1979).

One important use of this technique is in the screening of cDNA clones. While initially this was performed using purified radioactive mRNA or its cDNA copy, in most cases such a purified probe is not available. An alternative method ('plus and minus' screening) is to isolate mRNA from tissues which express the gene of interest and from tissues which do not. Duplicate filters can be prepared from each Petri dish and hybridized with either radioactive mRNA or cDNA synthesized from one or other of the probe preparations. Clones that give an increased response to the 'plus' preparation can then be analyzed further (*e.g.* Gray *et al.*, 1982). The second important use of this type of screening is in the isolation of genomic DNA clones. Using appropriate probes, overlapping cloned segments of genomic DNA can be isolated (walking; W. Bender, P. Spierer and D. Hogness, personal communication) allowing the structural analysis of extended regions of chromosomal DNA.

6.4.1.2 Translation in vitro

For screening cDNA libraries two methods that rely on the *in vitro* translation of RNA have been developed. In one, hybrid arrest translation (Paterson *et al.*, 1977), cloned DNA is hybridized to the bulk mRNA used to generate the cDNA and the preparation is then used to direct the synthesis of proteins in an appropriate *in vitro* system. Cloned DNA will titrate homologous mRNA from the preparation by duplex formation and thus prevent the synthesis of the protein product. The second technique relies on the hybridization of the cloned DNA with its homologous RNA to select the mRNA of interest. The RNA can then be recovered and used to direct an *in vitro* translation reaction (Ricciardi *et al.*, 1979; Parnes *et al.*, 1981). DNA is prepared from pools of recombinant clones, denatured and immobilized on filters (Kafatos *et al.*, 1979). The filters are then hybridized with the RNA which is then eluted and translated. This second technique has the advantage that a positive result is obtained. Both techniques require that the protein product of interest can be detected by size, activity or antigenicity.

6.4.1.3 Immunological screening

Methods analogous to those for plaque and colony hybridization have been developed to allow detection by antibody binding of clones that synthesize a specific protein product. Broome and Gilbert (1978) describe a technique where PVC sheets, to which antibody has been bound, are applied directly to plaques or lysed colonies on a Petri dish. The protein product binds specifically to the immobilized antibody and the sheets are exposed to a second, radioactive preparation of the antibody which binds to the immobilized antigenic protein product. Autoradiography can then be used to locate the positive clones on the original plate. An alternative sandwich technique uses radioactive protein A from *S. aureus* to detect antibody binding (Erlich *et al.*, 1978).

The use of antibodies to screen transformed colonies implies that at least part of the cloned gene is expressed and produces an antigenic protein fragment. To accomplish this, cDNA (or fragmented genomic DNA) may be cloned directly into an expression vector. Transformed colonies can then be replica plated onto nitrocellulose filters and screened for antibody binding (Heitman *et al.*, 1983). An alternative is to clone into a specialized vector that contains the *lac* promoter and β-galactosidase gene with a frameshift mutation that prevents the synthesis of active β-galactosidase. Selection is made for the insertion of DNA fragments that restore the correct reading frame by selecting for Lac^+ colonies. These colonies can then be screened by an antibody sandwich technique (Rüther *et al.*, 1982).

6.4.2 Characterization of Cloned DNA

6.4.2.1 Isolation of cloned DNA

The method of choice for isolating cloned DNA will be determined by the nature of the vector molecule and of the host. It will also depend on whether DNA is to be isolated for analytical or for more extensive studies.

It is often more convenient to screen directly DNA isolated from many individual transfor-

mants than to screen potential clones by one of the procedures discussed above. Plasmid content, size and restriction pattern can be conveniently determined from single *E. coli* transformant colonies or small cultures (Telford *et al.*, 1976; Barnes, 1976; Kado and Liu, 1981; Klein *et al.*, 1980). Similar techniques are available for a variety of Gram-negative bacteria (Kado and Liu, 1981) and yeast (Davis *et al.*, 1980). Isolation of supercoiled plasmid or viral DNA in large quantities requires cell lysis and enrichment by preferential removal of high molecular weight linear chromosomal DNA and cell debris. Lysis methods differ from organism to organism. In the case of *E. coli*, cells can be treated with lysozyme and partially lysed using a non-ionic detergent. Cell debris is removed by centrifugation ('cleared-lysate'; Clewell and Helinski, 1969) and the enriched plasmid DNA in the supernatant is subsequently purified by centrifugation in a CsCl–ethidium bromide density gradient. Modifications of this basic technique have been used with success in the isolation of plasmid DNA from *Pseudomonas putida* (J. Frey, personal communication) and *B. subtilis* (Lovett and Keggins, 1979).

A more general technique which has been adapted for several Gram-negative bacterial species, the Gram-positive *Streptomyces* and *B. subtilis*, as well as yeast and animal cells, involves total lysis with SDS, followed by removal of cell debris and chromosomal DNA. This can be achieved by preferential precipitation of high molecular weight DNA and membrane complexes with NaCl, NaAc or KAc (Hirt, 1967; Birnboim and Doly, 1979; Gryczan *et al.*, 1978; Bibb *et al.*, 1977; Hansen and Olsen, 1978) or by preferential denaturation of chromosomal DNA and its removal by treatment with acidic phenol (*e.g.* Zasloff *et al.*, 1978; Kado and Liu, 1981). Plasmid DNA is then purified by centrifugation on a CsCl–ethidium bromide density gradient. Both methods can give milligram yields of high copy number plasmid DNA from *E. coli*.

In the case of *E. coli* λ vectors, preparation of cloned DNA can be readily achieved by standard liquid or plate lysate methods (*e.g.* Enquist and Sternberg, 1979), banding of the phage particles on a CsCl step gradient, or precipitation with PEG and extraction of the DNA using phenol.

DNA cloned into the third type of *E. coli* vector, single-stranded DNA phages, is perhaps the easiest to prepare. By centrifugation of the bacterial culture and purification of the phage DNA from the supernatant by phenol extraction, sufficient DNA for sequencing can be obtained from 1 ml of culture (Schreier and Cortese, 1979).

6.4.2.2 Physical characterization of cloned fragments

In characterizing cloned DNA fragments, the most important initial step is to obtain a physical map (restriction map) of the cloned segment. This provides the basis for subsequent analyses. Excision of the cloned piece and purification from the vector DNA may facilitate the analysis. The map is generally obtained in two steps: an initial map is constructed using enzymes which cut infrequently and thus generate rather large subfragments. These enzymes are used both singly and in pairs. A fine-structure map can be generated from the subfragments using enzymes which cleave more frequently. In this step it is advisable to employ isolated subfragments obtained directly from agarose or acrylamide gels (Section 6.2.1.3) or indirectly by recloning into a secondary vector. If the size of the fragments is known accurately, a map can often be generated by simple addition. Normally this procedure can be carried out by hand although computer programs have been developed as an aid (see, for example, *Nucleic Acids Res.*, **10**, 1982).

Unfortunately, a major difficulty experienced in these procedures stems from the error involved in accurately determining the size of each subfragment. Several methods which avoid this problem exist for determining the order of restriction sites on a cloned fragment. One developed by Smith and Birnstiel (1976) involves partial digestion of a fragment after labelling at one extremity (Table 2). After size fractionation on an agarose or acrylamide gel and autoradiography, the size of the resulting 'nested set' (Maxam and Gilbert, 1980) of labelled fragments defines the distances of the cleavage sites from the labelled end. An alternative method involves the sequential deletion of restriction sites using the double-strand exonuclease *Bal*31. The plasmid is cleaved at a unique position in the vector, preferably close to the cloned fragment. The linearized molecule is then treated with *Bal*31 for various times, and then with the appropriate restriction enzyme. The order of disappearance of cleavage sites within the cloned piece is monitored (Legerski *et al.*, 1978). Since deletion occurs in both directions from the initial cleavage site, deletion of cleavage sites in the vector will also occur and thus the vector should be well characterized.

Restriction mapping may also be supplemented by the use of the Southern transfer technique

(Southern, 1979) in which DNA fragments, fractionated on an agarose or acrylamide gel, are denatured, transferred directly to a nitrocellulose filter and hybridized with individual labelled fragments generated by a second restriction enzyme. Several useful modifications of the original technique permit two 'blots' to be obtained from a single gel (Smith and Summers, 1980) and efficient transfer of fragments as small as 100 bp (Thomas, 1980).

Electron microscopy can provide important information complementary to that obtained by restriction mapping. Physical features such as the presence of palindromic sequences, insertions, genomic rearrangements or gross homologies within families of genes can be located with precision by the visualization of homoduplex and heteroduplex molecules formed with the cloned DNAs (see Brack, 1981). In many cases, electron microscopy is faster and simpler than many of the other available techniques. An example of a heteroduplex molecule formed between two related plasmids is shown in Figure 3. The physical structures, inverted repeats and regions of non-homology, represented by single-strand DNA loops in this figure, can be located to within 50 bp under suitable conditions. This technique has also been important in locating introns within eukaryotic genes (see Breathnach and Chambon, 1981).

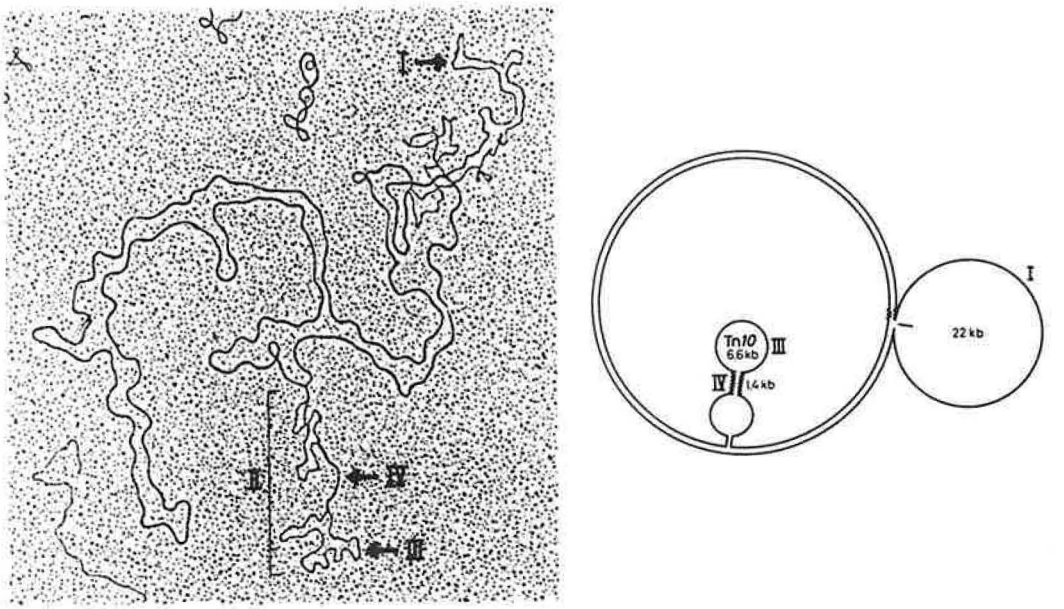

Figure 3 A heteroduplex between two related plasmids. The electron micrograph shows the structures formed after reannealing the single strands of two plasmids following denaturation. An interpretation of the structure is given showing the presence of a 22 kb single-stranded DNA loop due to an insertion into one of the plasmids (I). A second, partially single-stranded loop (II) is visible where DNA has been deleted from one of the plasmids. This deleted DNA includes the transposable element, Tn10, which is a 6.6 kb segment of DNA (III) flanked by sequences repeated in opposite orientations. These sequences can anneal forming a 1.4 kb double-stranded stem (IV)

Frequently, the ultimate goal in the physical characterization of the cloned gene is to obtain its polynucleotide sequence. This can now be performed rapidly, using one of several alternative techniques (Maxam and Gilbert, 1980; Sanger *et al.*, 1977; Guo and Wu, 1982). Primary sequence data provide indispensable information concerning the organization of the gene. They can be compared with concensus sequences, derived from the growing body of sequence data of both prokaryotic and eukaryotic origin, to define possible control regions such as operators and initiation and termination signals for both transcription and translation. Sequence data are essential in designing precise strategies for subsequent modification of the gene.

6.4.2.3 Characterization of products expressed by cloned fragments

Characterization of the expression of a given gene may require analysis of both its transcription and translation. Initially, the most important aspect is to determine whether or not a specific protein product is synthesized. For genes cloned in *E. coli*, four basic methods are available. They all

result in the preferential expression of proteins encoded by the cloned DNA fragment and the vector. Normally, detection is carried out by size fractionation of radioactively labelled proteins on denaturing polyacrylamide gels. This may be coupled with immunological or enzymatic detection of the protein if necessary (Dottin *et al.*, 1979).

The most versatile technique is to use recombinant DNA to direct protein synthesis in a coupled *in vitro* transcription/translation system (Zubay, 1973). Here, exogenous DNA is added to a bacterial cell extract together with radioactive amino acids. The procedure has several advantages. The extract is stable and a single extract can be used with different DNA templates. With suitable modifications (Pratt *et al.*, 1981), linear restriction fragments rather than entire plasmids can be used as templates, facilitating the physical localization of start and stop signals for transcription and translation. The extracts can be prepared from many *E. coli* strains, permitting the use of different genetic backgrounds to analyze the effects of various host functions on expression. In addition, specific negative regulators of transcription and translation should not interfere with the analysis since they are usually diluted in the extract. The major disadvantage is that a variable proportion of the product may be composed of prematurely terminated protein molecules. This can reduce the ability to detect both large proteins, which will have a greater probability of premature termination, and small proteins, which will be obscured by the presence of the protein fragments. The three current *in vivo* methods do not suffer from this problem.

The use of minicells has been the most commonly employed *in vivo* technique. They are small, DNA-less cells that are produced by certain mutants of *E. coli* and other bacterial species (see Frazer and Curtiss, 1975 for a review), and which can readily be separated from the parent cells by centrifugation of sucrose gradients. If the strain carries a plasmid, the plasmid can segregate into the minicells and which support limited synthesis of plasmid proteins when incubated in the presence of exogenous amino acids. Minicells will also support protein synthesis after phage infection (Reeve, 1979). In at least one case, this system has been shown to process correctly *E. coli* membrane proteins (Clement *et al.*, 1982).

An analogous *in vivo* system relies on the preferential degradation of chromosomal DNA following UV irradiation of a *recA uvrA* mutant of *E. coli* (Sancar *et al.*, 1979). Damaged DNA cannot be repaired and is degraded. With the correct dose of UV, plasmid DNA remains undamaged and the resulting 'maxicells' will direct the synthesis of plasmid-encoded proteins. The maxicell system requires fewer manipulations than the minicell system, but for either a separate transformation and preparation must be performed for each plasmid to be analyzed. UV irradiated $uvrA^-$ $recA^+$ cells can be used to study the synthesis of phage-encoded proteins. A much higher UV dose, sufficient to block all transcription in the cells, is given prior to infection (Murialdo and Siminovitch, 1972). With λ vectors many proteins are synthesized and complicate detection of the cloned gene product. It is possible to repress the synthesis of the vector proteins by using lysogenic cells that contain a non-inducible prophage which produces the λ ind^- repressor. Comparison of proteins synthesized in lysogenic and non-lysogenic cells can be used to determine whether or not transcription of a cloned gene can initiate in the cloned DNA segment.

Provided that the cloned gene is strongly expressed, an alternative *in vivo* system can be used. This relies on the selective amplification of colEl-type plasmid vectors in the presence of antibiotics that inhibit protein synthesis. After amplification the antibiotic is removed and the cells are incubated with amino acids. Under these conditions the majority of host proteins are poorly expressed while plasmid products, due to the high copy number of the gene, can be detected (Neidhardt *et al.*, 1980). The advantage of this technique is that no special strains or additional transformation steps are required.

Although no analogous methods exist for examining protein synthesis directly from cloned DNA fragments in eukaryotic organisms, *Xenopus* oocytes will efficiently translate injected mRNA molecules. In addition, *in vitro* translation systems based on extracts of wheat germ or reticulocytes are widely used. The *Xenopus* system will correctly modify proteins and can export heterologous secreted proteins (Colman *et al.*, 1981). A recent report (Contreras *et al.*, 1982) that capped mRNA molecules can be synthesized directly from plasmid DNA in an *in vitro* reaction using *E. coli* RNA polymerase, and that these RNAs are translated efficiently in all three eukaryotic systems, promises to facilitate the screening of cDNA libraries. Protein products may be identified by immunoprecipitation followed by gel electrophoresis. Alternatively, proteins may be transferred to nitrocellulose filters after electrophoresis ('Western blotting'; Towbin *et al.*, 1979; Bowen *et al.*, 1980) and exposed to radioactive antiserum.

Transcription products of cloned DNA can be characterized by methods analogous to those used for the characterization of translation products. Minicells will incorporate exogenous labelled uridine into RNA. The RNA products can be extracted and fractionated by gel electro-

phoresis (see Frazer and Curtiss, 1975). Similar experiments can be performed *in vivo* with maxicells and with phage infected, UV irradiated cells (Lund *et al.*, 1976) or *in vitro* using either purified bacterial RNA polymerase, or eukaryotic cell extracts (Manley *et al.*, 1980; Weil *et al.*, 1979).

For mapping *in vivo* transcription products, a modification of the Southern transfer technique can be employed (Southern blotting; Southern, 1979). Radioactive RNA from the original organism is hybridized to DNA of a clone which has been digested with various restriction enzymes, fractionated by gel electrophoresis and transferred to a nitrocellulose filter. This will locate the transcript relative to the restriction map of the cloned fragment. To identify various RNA products, for example primary and processed transcripts, unlabelled RNA can be fractionated by gel electrophoresis and transferred to chemically treated paper or nitrocellulose ('Northern blotting', Alwine *et al.*, 1979; Thomas, 1980) in an analogous manner. Radioactively labelled subclones of the DNA fragment can be used as hybridization probes.

To obtain more precise information as to the location of the ends of the transcript, a technique known as S1 mapping can be used (Berk and Sharp, 1977). RNA is extracted from cells and hybridized to specific, radioactively labelled, denatured DNA restriction fragments. The hybridization is performed under conditions where only DNA–RNA hybrids would be expected to form and the resulting hybrids are treated with S1 nuclease which degrades single-stranded nucleic acid. Single-stranded radioactive DNA is thus degraded but hybridized DNA is protected. The size of the protected DNA is then determined by electrophoresis on denaturing gels. With appropriate DNA probes and electrophoresis conditions, the method can be used to determine the 5′ and 3′ ends of a transcript, often with an accuracy of one or two bases. An alternative, used to map 5′ ends of transcripts, is to hybridize a labelled restriction fragment to the RNA. The fragment is chosen so that the 5′ end of the RNA remains single stranded. The hybridized DNA fragment is then used to prime DNA synthesis with AMVrt (Proudfoot *et al.*, 1980). Synthesis proceeds to the 5′ end of the RNA molecule and the size of the elongated DNA fragment is then determined by electrophoresis and autoradiography.

Promoter sites, as well as the direction of transcription, can be determined by electron microscopy. Complexes between RNA polymerase and DNA, formed by incubation of DNA with purified enzyme, can be visualized directly (Williams, 1977). The direction of transcription can be detected by measurement of the RNA chain length and the position of the growing point in preparations isolated at different times from an *in vitro* transcription reaction (Brack, 1981; Stuber and Bujard, 1981). For an alternative method using gel electrophoresis rather than electron microscopy see Chelm and Geiduschek (1979).

6.5 MANIPULATION OF CLONED GENES *IN VITRO*

6.5.1 Mutagenesis

The development of *in vitro* methods to make defined alterations in cloned DNA sequences not only provides a greater degree of control than the use of *in vivo* methods, but, more importantly, it enables mutations to be generated and identified in the absence of any selective system. The types of mutation which can be introduced *in vitro* include all those which occur *in vivo*: rearrangements, particularly deletion and insertion of DNA segments, and the alteration of specific base pairs (point mutations). The techniques can be used both in localizing a given gene within a large region of cloned DNA and in generating specific changes within a gene in order to study its expression or the function of its product. In the former case, mutations are generated at random throughout the segment and are generally deletion or insertion mutations whilst, in the latter, they are directed to a specific region or site. Many of these methods have been extensively reviewed recently (Shortle *et al.*, 1981).

6.5.1.1 Generation of deletions and insertions

The deletion and insertion of fragments in cloned DNA is important in defining specific genes, in analyzing their expression and in modifying the fragment for subsequent manipulation by, for example, the introduction or removal of restriction sites. While powerful *in vivo* techniques employing transposable elements are available and are often simpler to use (see Guyer, 1978;

Kleckner *et al.*, 1977), the advantage of *in vitro* manipulation is that the investigator retains a greater degree of control in site specificity.

The simplest method for generating deletions is by the use of restriction enzymes to remove a segment of DNA (Figure 4). This, of course, requires that suitable sites are present in the region of interest and that only a few such sites occur per molecule (Lai and Nathans, 1974). Methods for modification of DNA extremities (Section 6.2.3) can be conveniently employed to enable rejoining of ends generated by two different restriction endonucleases.

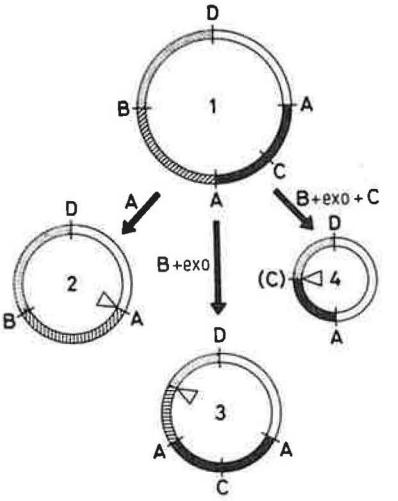

Figure 4 Construction of deletions *in vitro*. Deletion of DNA sequences in a plasmid molecule by restriction enzyme (A) cleavage and religation (i), or by restriction enzyme (B) cleavage followed by exonuclease digestion prior to ligation (ii). The first procedure results in deletion of DNA between the cleavage sites, and the second in deletions extending symmetrically from the original cleavage site. Cleavage of the DNA with a second restriction enzyme (C) prior to ligation, after cleavage with B and digestion with endonuclease (iii), results in deletions extending for a variable distance from a fixed site. Under appropriate conditions a cleavage site (C) is regenerated at the site of the deletion

More versatile *in vitro* methods for deletion formation rely on controlled nuclease digestion from a target site generated in a closed circular plasmid molecule. This target site can be a double-strand break, introduced by use of a restriction enzyme, or a single-strand break, introduced by an endonuclease such as DNaseI. In all cases, it is usually desirable to generate a single target site in each molecule. This can be achieved by partial cleavage with restriction enzymes for which there are two or more sites per molecule or by limited digestion with DNaseI. In the case of DNaseI, digestion in the presence of Mn^{2+} rather than Mg^{2+} results in double-strand cleavage instead of the introduction of single-strand breaks (Shenk *et al.*, 1976). Partial cleavage can often be conveniently controlled by the addition of ethidium bromide to the reaction. This limits cleavage of supercoiled molecules by DNaseI to a single nick and some restriction enzymes to a single, double-strand break (Greenfield *et al.*, 1975; Parker *et al.*, 1977; Shortle and Nathans, 1978).

The nature of the deletion generated depends upon both the exonuclease used and the target site. If a double-strand break is introduced into a plasmid, a deletion extending in both directions from the target site can be generated by the 3′–5′ single-strand activity of exonuclease III, followed by digestion of the resulting single-stranded regions with the nuclease S1 (Heffron *et al.*, 1977; Figure 4). If the target is a single-stranded break, this procedure will result in the formation of a deletion extending only in one direction, since only one DNA strand is digested by exonuclease III. It may be more convenient to use the nuclease *Bal*31 (Legerski *et al.*, 1978) since this nuclease has a double-strand exonuclease activity. It also possesses a single-strand endonuclease activity and is apparently capable of cleaving DNA opposite a single-strand break. Treatment of DNA with *Bal*31 should thus always produce deletions extending in both directions from the target site. If a suitable restriction site exists in the molecule, this can be used to define one end point of a deletion by digestion of exonuclease-treated DNA prior to ligation

(Figure 4). For ligation to occur, the ends of the DNA must usually be treated to generate flush ends (Section 6.2.3) and in some cases the restriction site can be regenerated and used in further manipulation (Panayotatos and Truong, 1981).

Where deletions must be directed to a small region that carries no appropriate restriction site, two elegant methods involving the use of small DNA fragments have been developed. (i) A synthetic oligonucleotide carrying the flanking sequences of the desired deletion is hybridized to a preparation of single-stranded circular plasmid DNA and is used as a primer for synthesis of the complementary strand. The resulting double-stranded plasmid DNA, now carrying a deletion in the newly synthesized strand, is treated with DNA ligase, introduced into a suitable bacterial host by transformation and the transformant colonies are screened by hybridization using the original oligonucleotide as a probe. With stringent hybridization conditions, the probe will only form stable hybrids with mutant plasmid DNA (Wallace *et al.*, 1980). The procedure can be simplified by use of a single-stranded phage vector, thus avoiding the necessity for preparing single-stranded plasmid DNA. (ii) A small restriction fragment, homologous to the region to be mutagenized, is hybridized to double-stranded plasmid DNA to form a displacement loop. The single-stranded region of the circular plasmid DNA, displaced by the fragment, is sensitive to S1 nuclease and digestion finally results in linearization of the plasmid DNA with the loss of several base pairs. Following purification of the linear DNA by gel electrophoresis, it is recircularized with DNA ligase and transformed into the host cells. The technique results in the formation of small deletions (on average 10 bp) throughout the region covered by the single-stranded fragment (Green and Tibbets, 1980).

A powerful technique for mutagenesis by insertion has been developed by Heffron and collaborators (Heffron *et al.*, 1978). It makes use of a synthetic linker, in this case carrying an *Eco*RI cleavage site, although any DNA linker may be used. Normally, a linker is chosen that would generate a frameshift mutation if inserted into a segment of DNA encoding a protein. The linker is ligated to linear plasmid DNA generated by double-strand cleavage with DNaseI. This method can be made more efficient if the linker carries a marker allowing direct selection for plasmids carrying the insertion. Recently, a plasmid has been described that carries a DNA fragment specifying resistance to streptomycin and spectinomycin that is flanked at both ends by cleavage sites for *Eco*RI and *Sma*I (Prentki and Krisch, 1982). The *Sma*I fragment, which has blunt ends, can be used directly for insertion mutagenesis, but results in a large (3.5 kb) insertion. Alternatively, the fragment can be generated by *Eco*RI digestion, and the extremities converted to blunt ends (Section 6.2.3). Mutant plasmids can then be redigested with *Sma*I to delete the drug resistance gene and recircularized leaving an insertion of 14 bp containing a *Sma*I cleavage site. A similar fragment, carrying transcription termination sequences at each end, has recently been constructed (P. Prentki and H. Krisch, personal communication).

6.5.1.2 Point mutations

Single base pair changes can be generated in cloned fragments, either by treating the DNA *in vitro* with chemical mutagens, or by *in vitro* repair synthesis using base analogues or an error-prone DNA polymerase. With these methods any base pair can be replaced by any other, often at high frequency. This is an important consideration if labour intensive methods such as DNA sequencing are to be used to identify clones containing the desired mutations.

Exposure of DNA *in vitro* to either hydroxylamine or bisulfite results in transitions that replace G:C by A:T base pairs. Since hydroxylamine reacts equally well with single- or double-stranded DNA it can only be used for generalized mutagenesis unless a small restriction fragment is purified, mutagenized and recloned. Bisulfite, in contrast, reacts much more rapidly with single-stranded DNA than with double-stranded DNA (see Shortle *et al.*, 1981). Because of this specificity mutations can be limited to a small region of a DNA molecule. Short single-stranded stretches of DNA suitable for bisulfite mutagenesis can be produced by use of a restriction enzyme to introduce a nick into a plasmid from which a small gap of a few nucleotides is then generated by exonuclease digestion. After reaction with bisulfite, the gap is repaired (Shortle *et al.*, 1981). The advantage of this strategy is that frequently the restriction site, if it contains a C residue, is destroyed. This greatly facilitates the identification of mutant clones if no direct selection is available. There are also techniques which expose a number of sites within a larger defined region of the plasmid to the mutagen. DNA containing large single-stranded regions can be prepared: by reannealing purified single plasmid DNA strands after cleavage with different combinations of restriction enzymes and strand separation (Giza *et al.*, 1981); by use of

a single-stranded phage vector (Weiher and Schaller, 1982); or by formation of a displacement loop (Everett and Chambon, 1982). These procedures have the disadvantage that multiple mutations can occur within the single-stranded region. A modification of the displacement loop technique that avoids this problem has been reported (Shortle *et al.*, 1981).

Transitions that result in the replacement of A:T by G:C base pairs can be induced by the use of N^4-hydroxy dCTP in place of dTTP during nick translation with *E. coli* DNA polymerase I (Müller *et al.*, 1978). Since the analogue will pair either with G or A residues, it should also be possible to replace G:C by A:T base pairs if dCTP rather than dTTP is replaced in the reaction.

Certain DNA polymerases, for example AMVrt, will misincorporate nucleotides at reasonably high frequencies *in vitro* and can be used to generate any base substitution. DNA with a single-stranded region produced by any of the methods described above can be repaired in the presence of only three dNTPs. This results in the misincorporation of one of them in place of the missing nucleotide (Shortle *et al.*, 1982). A similar method has been described that involves copying of a single-strand phage template with AMVrt, using an annealed restriction fragment as primer (Zakour and Loeb, 1982). This method has been adapted and successfully used to generate single mutations in defined regions of plasmid DNA (G. Cesareni, personal communication).

Defined point mutations can be introduced into DNA segments cloned in single-stranded phages, or plasmids, using synthetic oligonucleotides (Zoller and Smith, 1982; Dalbadie-McFarland *et al.*, 1982). The advantage of these techniques is that any particular base in a DNA sequence can be altered, since there is no need for an appropriate restriction site, which is normally required to expose a specific site for chemical mutagenesis, or to generate defined primers for repair synthesis.

Although it is not always possible to select directly a mutation in a specific segment of DNA, the choice of a particular vector may permit an indirect selection procedure to be used (Traboni *et al.*, 1982). A short DNA segment can be cloned into a vector at a site where insertion disrupts the synthesis of active β-galactosidase, due either to the introduction of a frameshift, or to the presence of translation termination codons in the cloned fragment. Any of the standard procedures for generating mutations in *E. coli* can be used to produce mutations in the cloned fragment that restore the synthesis of active β-galactosidase. These are then easily detected on media containing X-gal.

The methods for directed mutagenesis described here are often most convenient if applied to short segments of cloned DNA coding for only part of a gene. This means that the mutant DNA must be substituted into the wild-type gene for further analysis. This can be very difficult to do using *in vitro* techniques, and methods that rely on *in vivo* recombination to introduce a mutation generated *in vitro* into a gene are being developed. For example, mutations induced in a cloned fragment of T4 DNA have been efficiently introduced into the phage genome by growing phages on a pool of bacteria transformed with mutagenized plasmid DNA (Volker and Showe, 1980; Figure 5). A more general technique would be to mutagenize a small DNA segment which has been inserted in a plasmid vector. The mutation can then be introduced into the entire gene, cloned in a λ phage vector, by *in vivo* recombination. As shown in Figure 5, *in vivo* recombination results in a mixture of recombinant phages carrying mutant and wild-type alleles of the cloned gene. For this approach to be useful, it must be possible to distinguish readily between the two alleles, for example by screening plaques by hybridization with an oligonucleotide used to generate the mutant allele.

6.5.2 Efficient Expression of Cloned Genes

Since the ultimate goal of many gene cloning experiments is to obtain the maximum yield of the cloned product, the actual cloning and characterization of the gene may represent only the first of many steps. The problems involved in obtaining large yields of an active product are so many and varied as to be beyond the scope of this review. However, the first manipulations, involving *in vitro* recombinant DNA technology, are normally to ensure that the gene is expressed at a high level.

The transcription and translation signals that must be provided will obviously depend upon the host organism that is chosen. These constructions are not necessarily straightforward. For translation in a prokaryotic host, for example, a site that allows ribosome binding must be present, correctly positioned with respect to the initiation codon of the cloned gene. This positioning can

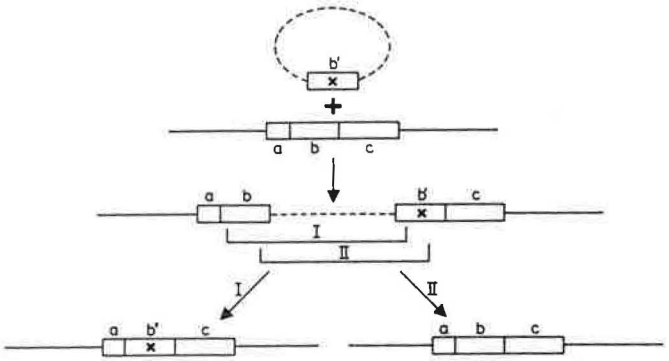

Figure 5 Reconstitution of mutagenized genes by *in vivo* recombination. Transfer of a mutation from a cloned segment to an entire gene. A segment of DNA is cloned into a plasmid and mutagenized (b'). The resulting mutation may be placed in the original gene by *in vivo* recombination. To reconstitute the gene a second recombination event is required, which results in the formation of either a mutant gene, or reconstitution of the original wild-type genes

be crucial in determining the level of translation of a protein (Gold *et al.*, 1981). Similar considerations apply to the expression of genes cloned in vectors that are propagated in eukaryotic hosts. Efficient expression of genes cloned in SV40 vectors, for example, requires that the mRNA transcript of the cloned segment is correctly spliced, capped and polyadenylated (see Hamer, 1980; Elder *et al.*, 1981) and therefore the appropriate signals for all these events must be correctly positioned in the final construction.

6.5.2.1 Constructions that maximize expression

The route followed in the construction of a recombinant DNA molecule that directs the high level expression of a cloned gene product will be different in each particular case. Usually, high level expression is first attempted in a prokaryotic host. The simplest method is to use an expression vector (Section 6.3.3.1) but cloning in many of those that are available frequently results in the addition of extraneous amino acids to the protein product, which might affect its function. In order to avoid this, Guarente *et al.* (1980) have made use of a three-step process that is applicable to any gene cloned in an *E. coli* plasmid vector.

First a hybrid gene is constructed which encodes a protein consisting of the amino-terminal end of the cloned gene product, fused to a portion of β-galactosidase, which retains its enzymatic activity. This is then linked to DNA containing the *lac* promoter and ribosome binding site, although other promoters could be used. This step results in the generation of a series of plasmids with different spacing between the ribosome binding site and the initiation codon of the cloned gene. Clones are examined for the level of β-galactosidase activity by simply comparing the intensity of colouration of colonies on X-gal medium. Those that produce the maximum level of β-galactosidase are picked, DNA is prepared, the β-galactosidase DNA segment is excised from the plasmid and the cloned gene is reconstituted. The advantage of this method is the easy selection of clones that direct the synthesis of high levels of the cloned gene product.

Another method, that also avoids the addition of extraneous amino acids, is to construct expression vectors that permit the insertion of DNA immediately adjacent to a promoter, ribosome binding site and initiation codon (Panayatatos and Truong, 1981).

Alternatively, a synthetic oligonucleotide containing appropriate translation initiation signals may be employed to modify the 5′ end of the cloned gene (see Yelverton *et al.*, 1983).

6.5.2.2 Secretion of cloned products

The manipulations described in the preceding section can be carried one step further to provide the signals necessary for secretion of the cloned gene product from the host organism. This may facilitate recovery of active protein made from the cloned gene. Nearly all secreted proteins are synthesized as preproteins and contain a signal sequence of additional amino acids at the amino-terminal end. This signal sequence is cleaved off during secretion through the cell membrane

(Blobel and Dobberstein, 1975). Some eukaryotic signal sequences, such as that of rat proinsulin, are recognized by *E. coli*, and the product of the cloned eukaryotic gene is secreted into the periplasmic space between the inner and outer cell membranes (Talmadge *et al.*, 1980b). The preprotein is correctly processed to yield a protein with no extraneous amino acids (Talmadge *et al.*, 1980a). After selective release of periplasmic proteins by osmotic shock, the proinsulin has already been separated from the bulk of the host proteins. At the time of writing, secretion vectors are available for both *E. coli* and *B. subtilis* (Talmadge and Gilbert, 1980; Palva *et al.*, 1982).

Secretion may also have another advantage in that the secreted proteins are protected from intracellular proteases by the cell membrane (Talmadge and Gilbert, 1982). In the case of *B. subtilis*, which has no outer membrane, proteins are secreted directly into the culture medium (Palva *et al.*, 1981). Since many bacterial strains produce extracellular proteases, the recovery of large yields of exogenous proteins is not yet a simple task.

Human interferons have been purified from culture media following expression in yeast (Hitzeman *et al.*, 1983). However, two major forms of the interferons were found, one identical to that produced by human cells and the other with three additional amino acids at the *N*-terminal end. Whether this reflects a difference in the mechanism of secretion between yeast and human cells is not yet clear. A considerable proportion of the interferon is not secreted. The cellular location of the non-secreted interferon has not been determined, but it is processed, raising the possibility that the non-secreted interferon is entering various organelles.

If the product of a cloned gene is an integral membrane protein, purification of the protein can present formidable problems. A recent report using influenza virus haemagglutinin as a model suggests that by suitable manipulation of clones such proteins can be secreted into the culture medium rather than being sequestered in the cell membranes of the host organism (Gething and Sambrook, 1982). Many such membrane proteins contain a very hydrophobic segment that serves to anchor the protein in the membrane. In the case of the influenza virus haemagglutinin this segment can be deleted and following expression from an SV40 vector in CV-1 cells, a truncated protein product that is glycosylated and retains normal haemagglutinin antigenic properties is secreted into the culture medium.

6.6 USEFUL SOURCES OF MORE DETAILED INFORMATION

Several collections of review articles contain much more detailed information on both the principles and practice of recombinant DNA technology. These include the series 'Genetic Engineering Principles and Practice' edited by J. K. Setlow and A. Hollaender, Plenum Press, New York and volumes 65 and 68 of 'Methods in Enzymology'. For an extensive review of gene cloning in organisms other than *E. coli* the reader is referred to 'Current Topics in Microbiology and Immunology', volume 96. Two manuals published by Cold Spring Harbor Laboratory, 'Advanced Bacterial Genetics' and especially 'Molecular Cloning' provide detailed experimental protocols.

ACKNOWLEDGEMENTS

During the writing of this review we were supported by grant No. 3.169.81 from the Swiss National Fund to Lucien Caro. We thank Marc Ballivet, Julian Davies and Gerald Selzer for actually reading the manuscript and suggesting many improvements; Otto Jenni for drawing the figures and Brigitte Brun, Jacqueline Snowdon and especially Monique Visini for endlessly retyping the manuscript.

6.7 REFERENCES

Alvarado-Urbina, G., G. M. Sathe, W.-C. Liu, M. F. Gillen, P. D. Duck, R. Bender and K. K. Ogilvie (1981). Automated synthesis of gene fragments. *Science*, **214**, 270–274.

Alwine, J. C., D. J. Kemp, B. A. Parker, J. Reiser, J. Renart, G. R. Stark and G. M. Wahl (1979). Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl paper. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 220–244. Academic, New York.

Anderson, S., M. J. Gait, L. Mayol and I. G. Young (1980). A short primer for sequencing DNA cloned in the single-stranded phage vector M13mp2. *Nucleic Acids Res.*, **8**, 1731–1743.

Aviv, H. and P. Leder (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid cellulose. *Proc. Natl. Acad. Sci. USA*, **69**, 1408–1412.

Bagdasarian, M., R. Lurz, B. Rückert, F. C. H. Franklin, M. M. Bagdasarian, J. Frey and K. Timmis (1981). Specific broad host range, high copy number, RSF1010-derived vectors, and a host–vector system for gene cloning in *Pseudomonas*. *Gene*, **16**, 237–247.

Bagdasarian, M. and K. N. Timmis (1982). Host vector systems for gene cloning in *Pseudomonas*. *Curr. Top. Microbiol. Immunol.*, **96**, 47–67.

Barany, F., J. D. Boeke and A. Tomasz (1982). Staphylococcal plasmids that replicate and express erythromycin resistance in both *Streptococcus pneumoniae* and *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **79**, 2991–2995.

Barnes, W. M. (1976). Plasmid detection and sizing in single colony lysates. *Science*, **195**, 393–394.

Barnes, W. M. (1980). DNA cloning with single-stranded phage vector. In *Genetic Engineering*: *Principles and Methods*, ed. J. K. Setlow and A. Hollaender, vol. 2. Plenum, New York.

Benton, W. D. and R. W. Davis (1977). Screening λgt recombinant clones by hybridisation to single plaques *in situ*. *Science*, **196**, 180–182.

Berk, A. J. and P. A. Sharp (1977). Sizing and mapping of early adenovirus mRNA's by gel electrophoresis of S1 endonuclease-digested hybrids. *Cell*, **12**, 721–732.

Bernard, H. U. and D. R. Helinski (1980). Bacterial plasmid cloning vehicles. In *Genetic Engineering*: *Principles and Methods*, ed. J. K. Setlow and A. Hollaender, vol. 2, pp. 133–168. Plenum, New York.

Bibb, M. J., R. F. Freeman and D. A. Hopwood (1977). Physical and genetical characterization of a second sex factor, ScP2, for *Streptomyces coelicolor* A3 (2). *Mol. Gen. Genet.*, **154**, 155–166.

Birnboim, H. C. and J. Doly (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.*, **7**, 1513–1523.

Bittner, M. and D. Vapnek (1981). Versatile cloning vectors derived from the run-away replication plasmid pKN402. *Gene*, **15**, 319–329.

Blobel, G. and D. Dobberstein (1975). Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes or murine myeloma. *J. Cell Biol.*, **67**, 835–851.

Bolivar, F. (1978). Construction and characterisation of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique EcoRI sites for selection of EcoRI generated recombinant DNA molecules. *Gene*, **4**, 121–136.

Bolivar, F., R. L. Rodriguez, M. C. Betlach and H. W. Boyer (1977a). Construction and characterisation of new cloning vehicles. I. Ampicillin resistant derivatives of the plasmid pMB1. *Gene*, **2**, 75–93.

Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa and J. Falkow (1977b). Construction and characterisation of new cloning vehicles. II. A multipurpose cloning system. *Gene*, **2**, 95–113.

Bowen, B., J. Steinberg, U. K. Laemmli and H. Weintraub (1980). The detection of DNA-binding proteins by protein blotting. *Nucleic Acids Res.* **8**, 1–20.

Boyko, W. L. and R. E. Ganschow (1982). Rapid identification of *Escherichia coli* transformed by pBR322 carrying inserts at the *Pst* I site. *Anal. Biochem.*, **122**, 85–88.

Brack, C. (1981). DNA electron microscopy. *CRC Crit. Rev. Biochem.*, **10**, 113–169.

Breathnach, R. and P. Chambon (1981). Organization and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.*, **50**, 349–383.

Broome, S. and W. Gilbert (1978). Immunological screening method to detect specific translation products. *Proc. Natl. Acad. Sci. USA*, **75**, 2746–2749.

Brown, D. M., J. Frampton, P. Goelet and J. Karn (1982). Sensitive detection of RNA using strand-specific M13 probes. *Gene*, **20**, 139–144.

Buell, G. N., M. P. Wickens, F. Payvar and R. T. Schimke (1978). Synthesis of full length cDNAs from four partially purified oviduct mRNAs. *J. Biol. Chem.*, **253**, 2471–2482.

Cannon, F. C., E. E. Riedel and F. M. Ausubel (1977). Recombinant plasmid that carries part of the nitrogen fixation gene cluster of *Klebsiella pneumoniae*. *Proc. Natl. Acad. Sci. USA*, **74**, 2963–2967.

Carbon, J., L. Clarke, C. Ilgen and B. Ratzkin (1977). The construction and use of hybrid plasmid gene banks in *Escherichia coli*. In *Recombinant Molecules*: *Impact on Science and Society*, ed. R. F. Beers, Jr. and E. G. Bassett, pp. 355–378. Raven, New York.

Casadaban, M. J. and S. N. Cohen (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. *J. Mol. Biol.*, **138**, 179–207.

Casadaban, M. J., J. Chou and S. N. Cohen (1980). *In vitro* gene fusions that join an enzymatically active β-galactosidase segment to amino terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. *J. Bacteriol.*, **143**, 971–980.

Chang, A. C. Y. and S. Cohen (1978). Construction and characterisation of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.*, **134**, 1141–1156.

Chang, S. and S. N. Cohen (1979). High frequency transformation of *Bacillus subtilis* protoplasts by plasmid DNA. *Mol. Gen. Genet.*, **168**, 111–115.

Charnay, P., M. Perricaudet, F. Galibert and P. Tiollais (1978) Bacteriophage lambda and plasmid vectors, allowing fusion of cloned genes in each of the three transitional phases. *Nucleic Acids Res.*, **12**, 4479–4494.

Chater, K. F., D. A. Hopwood, T. Kieser and C. J. Thompson (1982). Gene cloning in Streptomyces. *Curr. Top. Microbiol. Immunol.*, **96**, 69–95.

Chelm, B. K. and E. P. Geiduschek (1979). Gel electrophoretic separation of transcription complexes: an assay for RNA polymerase selectivity and a method for promoter mapping. *Nucleic Acids Res.*, **7**, 1851–1865.

Clement, J.-M., S. Perrin and J. Hedgpeth (1982). Analysis of λ receptor and β-lactamase synthesis and export using cloned genes in a minicell system. *Mol. Gen. Genet.*, **185**, 302–310.

Clewell, D. B. and D. R. Helinski (1969). Supercoiled circular DNA–protein complex in *E. coli*: purification and induced conversion to an open circular form. *Proc. Natl. Acad. Sci. USA*, **62**, 1159–1166.

Clewell, D. B. and D. R. Helinski (1972). Effect of growth conditions on the formation of the relaxation complex supercoiled colEl deoxyribonucleic acid and protein in *Escherichia coli*. *J. Bacteriol.*, **110**, 1135–1146.

Cohen, S. N., A. C. Y. Chang and L. Hsu (1972). Non-chromosomal antibiotic resistance in bacteria: genetic transformation of *E. coli* by R factor DNA. *Proc. Natl. Acad. Sci. USA*, **69**, 2110–2114.

Colbère-Garapin, F., A. Garapin and P. Kourilsky (1982). Selectable markers for the transfer of genes into mammalian cells. *Curr. Top. Microbiol. Immunol.*, **96**, 145–157.

Collins, J. (1979). *Escherichia coli* plasmids packageable *in vitro* in λ bacteriophage particles. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 309–325. Academic, New York.

Colman, A., C. D. Lane, R. Craig, A. Boolton, T. Mohun and J. Morser (1981). The influence of topology and glycosylation on the fate of heterologous secretory proteins made in *Xenopus* oocytes. *Eur. J. Biochem.*, **113**, 339.

Contreras, R., H. Cheroutre, W. Degrave and W. Fiers (1982). Simple efficient *in vitro* synthesis of capped RNA useful for direct expression of cloned eukaryotic clones. *Nucleic Acids Res.*, **10**, 6353–6362.

Covarrubias, L. and F. Bolivar (1982). Construction and characterisation of new cloning vehicles V1. Plasmid pBR329, a new derivative of pBR328 lacking the 482-base-pair inverted duplication. *Gene*, **17**, 79–89.

Dalbadie-McFarland, G., L. W. Cohen, A. D. Riggs, C. Morin, K. Itakura and J. H. Richards (1982). Oligonucleotide-directed mutagenesis as a general and powerful method for studies of protein function. *Proc. Natl. Acad. Sci. USA*, **79**, 6409–6413.

Davis, R. W., M. Thomas, J. Cameron, P. Thomas, S. Scherer and R. A. Padgett (1980). Rapid DNA isolation for enzymatic and hybridization analysis. In *Methods in Enzymology*, ed. L. Grossman and K. Moldave, vol. 65, pp. 404–411. Academic, New York.

Dimaio, D., R. Treisman and T. Maniatis (1982). Bovine papilloma virus vector that propagates as a plasmid in both mouse and bacterial cells. *Proc. Natl. Acad. Sci. USA*, **79**, 4030–4034.

Ditta, G., S. Stanfield, D. Corbin and D. R. Helinski (1980). Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. USA*, **77**, 7347–7351.

Dottin, R. P., R. E. Manrow, B. R. Fishel, S. L. Aukerman and J. L. Culleton (1979). Localization of enzymes in denaturing polyacrylamide gels. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 513–526. Academic, New York.

Dugaiczyk, A., H. W. Boyer and H. M. Goodman (1975). Ligation of EcoRI endonuclease-generated DNA fragments into linear and circular structures. *J. Mol. Biol.*, **96**, 171–184.

Edlund, T., T. N., M. Ranby, L. O. Heden, G. Palm, E. Holmgren and S. Joselmson (1983). Isolation of cDNA sequences coding for a part of human tissue plasminogen activator. *Proc. Natl. Acad. Sci. USA*, **80**, 349–352.

Efstratiadis, A., F. C. Kafatos, A. M. Maxam and T. Maniatis (1976). Enzymatic *in vitro* synthesis of globin genes. *Cell*, **7**. 279–288.

Ehrlich, S. D., B. Niaudet and B. Michel (1982). Use of plasmids from *Staphylococcus aureus* for cloning of DNA in *Bacillus subtilis*. *Curr. Top. Microbiol. Immunol.*, **96**, 19–29.

Elder, J. T., R. A. Spritz and S. M. Weissman (1981). Simian virus 40 as a eukaryotic cloning vehicle. *Annu. Rev. Genet.*, **15**, 295–340.

Enquist, L. and N. Sternberg (1979). *In vitro* packaging of *λDam* vectors and their use in cloning DNA fragments. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 281–298. Academic, New York.

Erlich, H. A., S. N. Cohen and H. O. McDevitt (1978). Immunological detection and characterization of products, translated from cloned DNA fragments. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 443–453. Academic, New York.

Everett, R. D. and P. Chambon (1982). A rapid and efficient method for the region- and strand-specific mutagenesis of cloned DNA. *EMBO J.*, **1**, 433–437.

Ferenczy, L. (1981). Microbial protoplast fusion. In *Genetics as a Tool in Microbiology. Society for General Microbiology Symposium 31*, ed. S. W. Glover and D. A. Hopwood, pp 1–34. Cambridge University Press, Cambridge.

Ferretti, L. and V. Sgaramella (1981). Specific and reversible inhibition of the blunt endjoining activity of the T4 DNA ligase. *Nucleic Acids Res.*, **9**, 3695–3705.

Fraley, R. and D. Papahadjopoulos (1982). Liposomes: the development of a new carrier system for introducing nucleic acids into plant and animal cells. *Curr. Top. Microbiol. Immunol.*, **96**, 171–191.

Frazer, A. C. and R. Curtiss (1975). Production, properties and utility of bacterial minicells. *Curr. Top. Microbiol. Immunol.*, **69**, 1–84.

Gergen, J. P., R. H. Stern and P. Wensink (1979). Filter replicas and permanent collections of recombinant DNA plasmids. *Nucleic Acids Res.*, **7**, 2115–2136.

Gething, M.-J. and J. Sambrook (1982). Construction of influenza haemagglutinin genes that code for intracellular and secreted forms of the protein. *Nature (London)*, **300**, 598–603.

Giza, P. E., D. M. Schmit and B. L. Murr (1981). Region and strand-specific mutagenesis of a recombinant plasmid. *Gene*, **15**, 331–342.

Gluzman, Y. (ed.) (1983). *Eukaryotic Viral Vectors*. Cold Spring Harbor Laboratory, NY.

Gold, L., D. Pribnow, T. Schneider, S. Shinedling, B. S. Singer and G. Stormo (1981). Translational initiation in prokaryotes. *Annu. Rev. Microbiol.*, **35**, 365–403.

Gray, P. W., D. W. Leung, D. Pennica, E. Yelverton, R. Najarian, C. C. Simonsen, R. Derynck, P. J. Sherwood, D. M. Wallace, S. L. B. Eber, A. D. Levinson and D. V. Goeddel (1982). Expression of human immune interferon cDNA in *E. coli* and monkey cells. *Nature (London)*, **295**, 503–508.

Green, C. and C. Tibbets (1980). Targeted deletion of sequences from cloned circular DNA. *Proc. Natl. Acad. Sci. USA*, **77**, 2455–2459.

Greenfield, L., L. Simpson and D. Kaplan (1975). Conversion of closed circular DNA molecules to single molecules by digestion with DNAseI in the presence of ethidium bromide. *Biochim. Biophys. Acta*, **407**, 365–375.

Grosjean, H. and W. Fiers (1982). Preferential codon usage in prokaryotic genes: the optimal codon–anticodon interaction energy and the selective codon usage in efficiently expressed genes. *Gene*, **18**, 199–209.

Grunstein, M. and D. Hogness (1975). Colony hybridisation: a method for the isolation of cloned DNAs that contain a specific gene. *Proc. Natl. Acad. Sci. USA*, **72**, 3961–3965.

Gruss, P. and G. Khoury (1982). Gene transfer into mammalian cells: use of viral vectors to investigate regulatory signals for the expression of eukaryotic genes. *Curr. Top. Microbiol. Immunol.*, **96**, 159–170.

Gryczan, T. J., S. Contente and D. Dubnau (1978). Characterization of *Staphylococcus aureus* plasmids introduced by transformation into *Bacillus subtilis*. *J. Bacteriol.*, **134**, 318–329.

Guarente, L., T. M. Roberts and M. Ptashne (1980). A technique for expressing eukaryotic genes in bacteria. *Science*, **209**, 1428–1430.
Guo, L. H. and R. Wu (1982). New rapid methods for DNA sequencing based on endonuclease III digestion followed by repair synthesis. *Nucleic Acids Res.*, **10**, 2065–2084.
Guyer, M. S. (1978). The sequence of F is an insertion sequence. *J. Mol. Biol.*, **126**, 347–365.
Hallewell, R. A. and S. Emtage (1980). Plasmid vectors containing the tryptophan operon promoter suitable for efficient regulated expression of foreign genes. *Gene*, **9**, 27–47.
Hamer, D. H. (1980). DNA cloning in mammalian cells with SV40 vectors. In *Genetic Engineering: Principles and Methods*, ed. J. K. Setlow and A. Hollaender, vol. 2, pp. 83–101. Plenum, New York.
Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.*, **166**, 557–580.
Hanahan, D. and M. Meselson (1980). Plasmid screening at high colony density. *Gene*, **10**, 63–67.
Hansen, J. B. and R. H. Olsen (1978). Isolation of large bacterial plasmids and characterization of the P2 incompatibility group plasmids pMG1 and pMG5. *J. Bacteriol.*, **135**, 227–238.
Heffron, F., P. Bedinger, J. J. Champoux and S. Falkow (1977). Deletions affecting the transposition of an antibiotic resistance gene. *Proc. Natl. Acad. Sci. USA*, **74**, 702–706.
Heffron, F., M. So and B. J. McCarthy (1978). *In vitro* mutagenesis of circular DNA molecules using synthetic restriction sites. *Proc. Natl. Acad. Sci. USA*, **75**, 6012–6016.
Helfman, D. M., J. R. Feramisco, J. C. Fiddes, G. P. Thomas and S. H. Hughes (1982). Identification of clones that encode chicken tropomyosin by direct immunogical screening of a cDNA expression library. *Proc. Natl. Acad. Sci. USA*, **80**, 31–35.
Herrera-Estrella, L., A. Depicker, M. van Montagu and J. Schell (1983). Expression of chimaeric genes transferred into plant cells using a Ti-plasmid-derived vector. *Nature (London)*, **303**, 209–213.
Hicks, J. B., A. Hinnen and G. R. Fink (1978). Properties of yeast transformation. *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 1305–1313.
Higgins, N. B. and N. R. Cozzarelli (1980). DNA joining enzymes: a review. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 50–71. Academic, New York.
Hinnen, A. and B. Meyhack (1982). Vectors for cloning in yeast. *Curr. Top. Microbiol. Immunol.*, **96**, 101–117.
Hirt, B. (1967). Selective extraction of polyoma DNA from infective mouse cell cultures. *J. Mol. Biol.*, **26**, 365–369.
Hitzeman, R. A., D. W. Leung, L. J. Perry, W. T. Kohr, H. L. Levine and D. V. Goeddel (1983). Secretion of human interferon by yeast. *Science*, **219**, 620–625.
Hofstetter, H., A. Schambock, J. van den Berg and C. Weissmann (1976). Specific excision of the inserted DNA segment from hybrid plasmids constructed by the poly(dA)-poly(dT) method. *Biochim. Biophys. Acta*, **454**, 587–591.
Hohn, B. (1979). In vitro packaging of λ and cosmid DNA. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 299–309. Academic, New York.
Hohn, B. and A. Hinnen (1980). Cloning with cosmids in *E. coli* and yeast. In *Genetic Engineering: Principles and Methods*, ed. J. K. Setlow and A. Hollaender, vol. 2. Plenum, New York.
Hohn, T., K. Richards and G. Lebeurier (1982). Cauliflower mosaic virus on its way to becoming a useful plant vector. *Curr. Top. Microbiol. Immunol.*, **96**, 193–236.
Hollenberg, C. P. (1982). Cloning with 2 μm DNA vectors and the expression of foreign genes in *Saccharomyces cerevisiae*. *Curr. Top. Microbiol. Immunol.*, **96**, 120–144.
Holsters, M., B. Silva, F. Van Vliet, J. P. Hernalsteens, G. Genetello, M. Van Montagu and J. Schell (1978). *In vivo* transfer of the Ti plasmid of *Agrobacterium tumefaciens* to *Escherichia coli*. *Mol. Gen. Genet.*, **163**, 335–338.
Hopwood, D. A. (1981). Genetic studies with bacterial protoplasts. *Annu. Rev. Microbiol.*, **35**, 237–272.
Ish-Horowicz, D. and J. F. Burke (1981). Rapid and efficient cosmid cloning. *Nucleic Acids Res.*, **9**, 2989–2998.
Jimenez, A. and J. Davies (1980). Expression of a transposable antibiotic element in *Saccharomyces cerevisiae*: a potential selection for eukaryotic cloning vectors. *Nature (London)*, **287**, 869–871.
Kado, C. I. and S. T. Liu (1981). Rapid procedure for detection and isolation of large and small plasmids. *J. Bacteriol.*, **145**, 1365–1373.
Kafatos, F. C., C. W. Jones and A. Efstratiadis (1979). Determination of nucleic acid sequence homologies and relative concentrations by a dot hybridization procedure. *Nucleic Acids Res.*, **7**, 1541–1552.
Kahn, M., R. Kolter, C. Thomas, D. Figurski, R. Meyer, E. Remaut and D. Helinski (1979). Plasmid cloning vehicles derived from plasmids ColE1, F, R6K and RK2. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 268–280. Academic, New York.
Kingsman, A. J., L. Clarke, R. K. Mortimer and J. Carbon (1979). Replication in *Saccharomyces cerevisiae* of plasmid pBR313 carrying DNA from the yeast trpl region. *Gene*, **7**, 141–152.
Kleckner, N., J. Roth and D. Botstein (1977). Genetic engineering *in vivo* using translocatable drug-resistance elements. New methods in bacterial genetics. *J. Mol. Biol.*, **116**, 125–159.
Klee, H. J., M. P. Gordon and E. W. Nester (1982). Complementation analysis of *Agrobacterium tumefaciens* Ti plasmid mutations affecting oncogenicity. *J. Bacteriol.*, **150**, 327–331.
Klein, R. D., E. Selsing and R. D. Wells (1980). A rapid microscale technique for isolation of recombinant plasmid DNA suitable for restriction enzyme analysis. *Plasmid*, **3**, 88–91.
Korman, A. J., P. J. Knudsen, J. F. Kaufman and J. L. Strominger (1982). cDNA clones for the heavy chain of HLA-DR antigens obtained after immunoprecipitation of polysomes by monoclonal antibody. *Proc. Natl. Acad. Sci. USA*, **79**, 1844–1848.
Kreft, J. and C. Hughes (1982). Cloning vectors derived from plasmids and phage of *Bacillus*. *Curr. Top. Microbiol. Immunol.*, **96**, 1–17.
Krens, F. A., L. Molendijk, G. J. Wullems and R. A. Schilperoot (1982). *In vitro* transformation of plant protoplasts with Ti-plasmid DNA. *Nature (London)*, **296**, 72–74.
Lai, C.-J. and D. Nathans (1974). Deletion mutants of SV40 generated by enzymatic excision of DNA segments from the viral genome. *J. Mol. Biol.*, **89**, 179–193.
Land, H., M. Grez, H. Hauser, W. Lindenmaier and G. Schütz (1981). 5'-Terminal sequences of eukaryotic mRNA can be cloned with high efficiency. *Nucleic Acids Res.*, **9**, 2251–2266.
Leemans, J., C. Shaw, R. Deblaere, H. de Greve, J. P. Hernalsteens, M. Maes, M. van Montague and J. Schell (1982).

Site specific mutagenesis of *Agrobacterium* Ti plasmids and transfer of genes to plant cells. *J. Mol. Appl. Genet.*, **1**, 149–162.

Legerski, R. J., J. L. Hodnett and H. B. Gray (1978). Extracellular nucleases of *Pseudomonas* Bal31. III. Use of the double-strand deoxyriboexonuclease activity as the basis of a convenient method of the mapping of fragments of DNA produced by cleavage with restriction enzymes. *Nucleic Acids Res.*, **5**, 1445–1464.

Loenen, W. A. M. and W. J. Brammer (1980). A bacteriophage lambda vector for cloning large DNA fragments made with several restriction enzymes. *Gene*, **10**, 249–259.

Lovett, P. S. and K. M. Keggins (1979). *Bacillus subtilis* as a host for molecular cloning. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 342–357. Academic, New York.

Lund, E., J. E. Dahlberg, L. Lindahl, S. R. Jaskunas, P. P. Dennis and M. Nomura (1976). Transfer RNA genes between 16 and 23S rRNA transcription units of *E. coli*. *Cell*, **7**, 165–177.

Lurquin, P. F. (1979). Entrapment of plasmid DNA by liposomes and their interactions with plant protoplasts. *Nucleic Acids Res.*, **6**, 2773–2784.

Makins, J. F. and G. Holt (1982). Liposome-mediated transformation of streptomyces by chromosomal DNA. *Nature (London)*, **293**, 671–673.

Malloy, J. R. and W. D. Nunn (1981). Selection for loss of tetracycline resistance by *Escherichia coli*. *J. Bacteriol.*, **145**, 1110–1112.

Maniatis, T., S. G. Kee, A. Efstratiadis and F. C. Kaftos (1976). Amplification and characterization of a β-globin gene synthesized *in vitro*. *Cell*, **8**, 163–182.

Maniatis T., R. C. Hardison, E. Lacy, J. Lauer, C. O'Connell and D. Quon (1978). The isolation of structural genes from libraries of eukaryotic DNA. *Cell*, **15**, 687–701.

Maniatis, T., E. F. Fritsch and J. Sambrook (1982). *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, N.Y.

Manley, J. E., A. Fire, A. Cano, P. A. Sharp and M. L. Gefter (1980). DNA-dependent transcription of adenovirus genes in a soluble system. *Proc. Natl. Acad. Sci. USA*, **77**, 3855–3859.

Maxam, A. M. and W. Gilbert (1980). Sequencing end labelled DNA with base-specific chemical cleavages. In *Methods in Enzymology*, ed. L. Grossman and K. Moldave, vol. 65. Academic, New York.

McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady and M. Rosenberg (1981). A system to study promotor and terminator signals recognised by *Escherichia coli* RNA polymerase. In *Gene Amplification and Analysis*, ed. J. C. Chirikjian and T. S. Papos, vol. II. Elsevier/North Holland Biomedical Press, Amsterdam.

Mercereau-Puijalon, O., A. Royal, B. Cami, A. Garapin, A. Krust, F. Gannon and P. Kourilsky (1978). Synthesis of an ovalbumin-like protein by *Escherichia coli* K12 harbouring a recombinant plasmid. *Nature (London)*, **275**, 505–510.

Mertz, J. E. and R. W. Davis (1972). Cleavage of DNA by RI restriction endonuclease generates cohesive ends. *Proc. Natl. Acad. Sci. USA*, **68**, 3370–3374.

Muller, W., H. Weber, F. Meyer and C. Weissman (1978). Site directed mutagenesis of DNA: generation of point mutations in cloned β-globin complementary DNA at positions corresponding to amino acids 121 and 123. *J. Mol. Biol.*, **124**, 343–358.

Murialdo, O. H. and L. Siminovitch (1972). The morphogenesis of bacteriophage lambda. IV. Identification of gene products and control of the expression of morphogenetic information. *Virology*, **48**, 785–823.

Neidhardt, F. C., R. Wirth, M. W. Smith and R. V. Bogelen (1980). Selective synthesis of plasmid-coded proteins by *Escherichia coli* during recovery from chloramphenicol treatment. *J. Bacteriol.*, **143**, 535.

Nelson, T. and D. Brutlag (1980). Addition of homopolymers to the 3'-ends of duplex DNA with terminal transferase. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 41–50. Academic, New York.

Nester, E. W. and T. Kosuge (1981). Plasmids specifying plant hyperplasias. *Annu. Rev. Microbiol.*, **35**, 531–565.

Norgard, M. V., K. Keem and J. J. Monahan (1978). Factors affecting the transformation of *Escherichia coli* strain 1776 by pBR322 plasmid DNA. *Gene*, **3**, 279–292.

O'Callaghan, C. H., A. Morris, S. M. Kirby and A. H. Shinger (1972). Novel method for detection of β-lactamases by using a chromogenic cephalosporin substrate. *Antimicrob. Agents Chemother.*, **1**, 283–288.

Okayama, H. and P. Berg (1982). High efficiency cloning of full length cDNA. *Mol. Cell. Biol.*, **2**.

Palva, I., R. F. Pettersson, N. Kalkkinen, P. Lehtovarra, M. Sarvas, H. Söderlund, P. Takinen and L. Kääriäinen (1981). Nucleotide sequence of the promoter and NH_2-terminal signal peptide region of the α-amylase gene from *Bacillus amyloliquefaciens*. *Gene*, **15**, 43–51.

Palva, I., M. Sarvas, P. Lehtovaara, M. Sibakov and L. Kääriäinen (1982). Secretion of *Escherichia coli* β-lactamase from *Bacillus subtilis* by the aid of α-amylase signal sequence. *Proc. Natl. Acad. Sci. USA*, **79**, 5582–5586.

Panayotatos, N. and K. Truong (1981). Specific deletions of DNA sequences between preselected bases. *Nucleic Acids Res.*, **1**, 5679–5688.

Parker, R. C., R. M. Watson and J. Vinograd (1977). Mapping of closed circular DNAs by cleavage with restriction endonucleases and calibration by agarose gel electrophoresis. *Proc. Natl. Acad. Sci. USA*, **74**, 851–855.

Parnes, J. E., B. Velan, A. Felsenfeld, L. Ramanathan, U. Ferrini, E. Appell and J. G. Seidman (1981). Mouse β-microglobulin cDNA clones: a screening procedure for cDNA clones corresponding to rare mRNAs. *Proc. Natl. Acad. Sci. USA*, **78**, 2253–2257.

Paterson B. M., B. E. Roberts and E. L. Kuff (1977). Structural gene identification and mapping by DNA:mRNA hybrid arrested cell-free translation. *Proc. Natl. Acad. Sci. USA*, **74**, 4370–4374.

Pratt, J. M., G. M. Boulnois, V. Darby, E. Orr, E. Wahle and I. B. Holland (1981). Identification of gene products programmed by restriction endonuclease DNA fragments using an *E. coli in vitro* system. *Nucleic Acids Res.*, **9**, 4459–4474.

Prentki, P., F. Karch, S. Ilda and J. Meyer (1981). The plasmid cloning vector pBR325 contains a 482 base pair-long inverted duplication. *Gene*, **14**, 289–299.

Prentki, P. and H. M. Krisch (1982). A modified pBR322 vector with improved properties for the cloning, recovery and sequencing of blunt ended DNA fragments. *Gene*, **17**, 189–196.

Procter, G. N. and R. H. Rownd (1982). Rosanilins: indicator dyes for chloramphenicol resistant enterobacteria containing chloramphenicol acetyltransferase. *J. Bacteriol.*, **150**, 1375–1382.

Proudfoot, N. J., M. H. H. Shander, J. L. Manley, M. L. Gefter and T. Maniatis (1980). Structure and *in vitro* transcription of human globin genes. *Science*, **209**, 1329–1336.

Queen, C. and M. Rosenberg (1981). A promotor of pBR322 activated by cAMP receptor protein. *Nucleic Acids Res.*, **9**, 3365–3377.

Rassoulzadegan, M., B. Bineruy and F. Cuzin (1982). High frequency of gene transfer after fusion between bacteria and eukaryotic cells. *Nature* (*London*), **285**, 257–259.

Ray, D. S., J. C. Hines, M. H. Kim, R. Imber and N. Nomura (1982). M13 vectors for selective cloning of sequences specifying initiation of DNA synthesis of single-stranded templates. *Gene*, **18**, 231–238.

Reeve, J. N. (1979). Use of minicells for bacteriophage-directed polypeptide synthesis. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 493–502. Academic, New York.

Remaut, E., P. Stanssens and W. Fiers (1981). Plasmid vectors for high efficiency expression controlled by the P_L promotor of coliphage lambda. *Gene*, **15**, 81–93.

Ricciardi, R. P., T. S. Miller and B. E. Roberts (1979). Purification and mapping of specific mRNAs by hybridization selection and cell-free translation. *Proc. Natl. Acad. Sci. USA*, **76**, 4927–4931.

Rigby, P. W. J., M. Dieckman, C. Rhodes and P. Berg (1977). Labelling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. *J. Mol. Biol.*, **113**, 237–251.

Roberts, R. J. (1983). Restriction and modification enzymes and their recognition sequences. *Nucleic Acids Res.*, **11**, 135–167.

Rougeon, F., P. Kourilsky and B. Mach (1975). Insertion of a rabbit β-globin gene sequence into an *E. coli* plasmid. *Nucleic Acids Res.*, **2**, 2365–2378.

Rougeon, F. and B. Mach (1976). Stepwise biosynthesis *in vitro* of globin genes from globin mRNA by DNA polymerase of avian myeloblastosis virus. *Proc. Natl. Acad. Sci. USA*, **73**, 3418–3422.

Rubin, G. M. and A. C. Spradling (1982). Genetic transformation of *Drosophila* with transposable element vectors. *Science*, **218**, 348–353.

Rüther, U. (1982). pUR250 allows rapid chemical sequencing of both DNA strands of its inserts. *Nucleic Acids Res.*, **10**, 5765–5772.

Rüther, U., M. Koenen, K. Otto and B. Müller-Hill (1981). pUR222, a vector for cloning and rapid chemical sequencing of DNA. *Nucleic Acids Res.*, **9**, 4087–4098.

Rüther, V., M. Koenen, A. E. Sippel and B. Müller-Hill (1982). Exon cloning: immunoenzymatic detection of exons of the chicken lysozyme gene. *Proc. Natl. Acad. Sci. USA*, **79**, 6852–6855.

Sancar, A., A. M. Hack and W. D. Rupp (1979). Simple method for identification of plasmid coded protein. *J. Bacteriol.*, **137**, 692–693.

Sanger, F., S. Nicklen and A. R. Coulson (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*, **12**, 5463–5467.

Schreier, P. H. and R. Cortese (1979). A fast and simple method for sequencing DNA cloned in the single-stranded bacteriophage M13. *J. Mol. Biol.*, **129**, 169–172.

Seed, B., R. C. Parker and N. Davidson (1982). Representation of DNA sequences in recombinant DNA libraries prepared by restriction enzyme partial digestion. *Gene*, **19**, 201–209.

Sgaramella, V., J. H. van de Sande and H. G. Khorana (1970). Studies on polynucleotides: C. A novel joining reaction catalyzed by the T4-polynucleotide ligase. *Proc. Natl. Acad. Sci. USA*, **67**, 1468–1475.

Sgaramella, V. and H. G. Khorana (1972). Studies on polynucleotides: CXVI. A further study of the T4 ligase catalysed joining of DNA at base-paired ends. *J. Mol. Biol.*, **72**, 493–502.

Shenk, T., J. Carbon and P. Berg (1976). Construction and analysis of viable deletion mutants of SV40. *J. Virol.*, **18**, 664–671.

Shortle, D. and D. Nathans (1978). Local mutagenesis: a method for generating viral mutants with base substitutions in preselected regions of the viral genome. *Proc. Natl. Acad. Sci. USA*, **75**, 2170–2174.

Shortle, D., D. Dimaio and D. Nathans (1981). Directed mutagenesis. *Annu. Rev. Biochem.*, **15**, 265–297.

Shortle, D., P. Grisafi, J. J. Bencovic and D. Botstein (1982). Gap misrepair mutagenesis: efficient site directed induction of transition, transversion and frameshift mutations *in vitro*. *Proc. Natl. Acad. Sci. USA*, **79**, 1588–1592.

Slutsky, A. M., P. M. Rabinovich, L. Z. Yakubov, I. V. Sineokaya, A. I. Stepanov and V. K. Gordeyev (1980). Direct selection of DNA inserts in plasmid gene of kanamycin resistance. *Mol. Gen. Genet.*, **180**, 487–488.

Smith, A. J. H. (1980). DNA sequence analysis by primed synthesis. In *Methods in Enzymology*, ed. L. Grossman and K. Moldave, vol. 65, pp. 560–579. Academic, New York.

Smith, G. E. and M. D. Summers (1980). The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl-paper. *Anal. Biochem.*, **109**, 123–129.

Smith, H. O. (1980). Recovery of DNA from gels. In *Methods in Enzymology*, ed. L. Grossman and K. Moldave, vol. 65, pp. 371–380. Academic, New York.

Smith, H. O. and M. L. Birnsteil (1976). A simple method for DNA restriction site mapping. *Nucleic Acids Res.*, **3**, 2387–2398.

Soberon, X., L. Covarrubias and F. Bolivar (1980). Construction and characterization of new cloning vehicles. IV. Deletion derivatives of pBR322 and pBR325. *Gene*, **9**, 282–305.

Sompayrac, L. M. and K. J. Danna (1981). Efficient infection of monkey cells with DNA of simian virus 40. *Proc. Natl. Acad. Sci. USA*, **78**, 7575–7578.

Southern, E. (1979). Gel electrophoresis of restriction fragments. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 152–176. Academic, New York.

Stinchcomb, D. T., M. Thomas, J. Kelly, E. Selker and R. W. Davies (1980). Eukaryotic DNA segments capable of autonomous replication in yeast. *Proc. Natl. Acad. Sci. USA*, **77**, 4559–4563.

Stüber, D. and H. Bujard (1981). Organisation of transcriptional signals in plasmids pBR322 and pACYC184. *Proc. Natl. Acad. Sci. USA*, **78**, 167–171.

Sutcliffe, J. G. (1979). Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 77–90.

Tacon, W., N. Carey and S. Emtage (1980). The construction and characterisation of plasmid vectors suitable for the expression of all DNA phases under the control of the *E. coli* tryptophan promotor. *Mol. Gen. Genet.*, **177**, 427–438.

Talmadge, K. and W. Gilbert (1981). Construction of plasmid vectors with unique PstI cloning sites in a signal sequence coding region. *Gene*, **12**, 235–241.
Talmadge, K. and W. Gilbert (1982). Cellular location affects protein stability in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **79**, 1830–1833.
Talmadge, K., J. Kaufman and W. Gilbert (1980a). Bacteria mature preproinsulin to proinsulin. *Proc. Natl. Acad. Sci. USA*, **77**, 3988–3992.
Talmadge, K., S. Stahl and W. Gilbert (1980b). Eukaryotic signal sequence transports insulin antigen in *Escherichia coli*. *Proc. Natl. Acad. Sci. USA*, **77**, 3369–3373.
Telford, J., P. Boseley, W. Schaffner and M. Birnstiel (1976). Novel screening procedure for recombinant plasmids. *Science*, **195**, 391–392.
Thomas, P. S. (1980). Hybridisation of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA*, **77**, 5201–5205.
Timmis, K. N. (1981). Gene manipulation *in vitro*. In *Genetics as a Tool in Microbiology*, *Society for General Microbiology Symposium 31*, ed. S. W. Glover and D. A. Hopwood. Cambridge University Press, Cambridge.
Towbin, H., T. Staehelin and J. Gordon (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc. Natl. Acad. Sci. USA*, **76**, 4350–4354.
Traboni, G., G. Ciliberto and R. Cortese (1982). A novel method for site directed mutagenesis: its application to an eukaryotic $tRNA^{PRO}$ gene promotor. *EMBO J.*, **1**, 415–420.
Tuite, M. F., M. J. Dobson, N. A. Roberts, R. M. King, D. C. Burke, S. M. Kingsman and A. J. Kingsman (1982). Regulated high efficiency expression of human interferon-alpha in *Saccharomyces cerevisiae*. *EMBO J.*, **1**, 603–608.
Twigg, A. J. and D. Sherratt (1980). Trans-complementable copy-number mutants of plasmid colE1. *Nature* (*London*), **283**, 216–218.
Uhlin, B. E., S. Molin, P. Gustafsson and K. Nordström (1979). Plasmids with temperature dependent copy number for amplification of cloned genes and their products. *Gene*, **6**, 91–106.
Ullrich, A., J. Shine, J. Chirgwin, R. Pictet, E. Tischer, W. J. Rutter and H. M. Goodman (1977). Rat insulin genes: construction of plasmids containing the coding sequences. *Science*, **196**, 1313–1319.
van Montagu, M. and J. Schell (1982). The Ti plasmids of *Agrobacterium*. *Curr. Top. Microbiol. Immunol.*, **96**, 237–254.
Vapnek, D. and M. Case (1982). Gene cloning in *Neurospora crassa*. *Curr. Top. Microbiol. Immunol.*, **96**, 98–100.
Villa-Komaroff, L., A. Efstratiadis, S. Broome, P. Lomdico, R. Tizard, S. P. Naber, W. L. Chick and W. Gilbert (1978). A bacterial clone synthesizing proinsulin. *Proc. Natl. Acad. Sci. USA*, **75**, 3727–3731.
Völker, T. A. and M. K. Showe (1980). Induction of mutations in specific genes of bacteriophage T4 using cloned restriction fragments and marker rescue. *Mol. Gen. Genet.*, **177**, 447–452.
Wallace, R. B., P. F. Johnson, S. Tanaky, M. Schöld, K. Itakura and J. Abelson (1980). Directed deletion of a yeast transfer RNA intervening sequence. *Science*, **209**, 1396–1400.
Weiher, H. and H. Schaller (1982). Segment specific mutagenesis and extensive mutagenesis of a *lac* promotor/operator element. *Proc. Natl. Acad. Sci. USA*, **79**, 1408–1412.
Weil, P. A., D. S. Lose, J. Segall and R. G. Roeder (1979). Selective and accurate initiation of transcription at the Ad2 major late promoter in a soluble system dependent on purified RNA polymerase II and DNA. *Cell*, **18**, 469–484.
Wensink, P. C., D. J. Finegan, J. E. Donelson and D. S. Hogness (1974). A system for mapping DNA sequences in the chromosomes of *Drosophila melanogaster*. *Cell*, **3**, 315–325.
West, R. W. and R. L. Rodriguez (1980). Construction and characterization of *E. coli* promotor probe plasmid vectors. II. RNA polymerase binding studies on antibiotic resistance promotors. *Gene*, **9**, 175–193.
Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlamb and L. Chasin (1979). DNA-mediated transfer of the adenine phosphoribosyl-transferase locus into mammalian cells. *Proc. Natl. Acad. Sci. USA*, **76**, 1373–1376.
Willets, N. and C. Crowther (1981). Mobilization of the non-conjugative Inc and plasmid RSF1010. *Genet. Res.*, **37**, 311–316.
Williams, R. C. (1977). Use of polylysine for adsorption of nucleic acids and enzymes to electron microscope specimen films. *Proc. Natl. Acad. Sci. USA*, **74**, 2311–2316.
Williams, W. and F. R. Blattner (1980). Bacteriophage lambda vector for DNA cloning. In *Genetic Engineering*: *Principals and Methods*, ed. J. K. Setlow and A. Hollaender, vol. 2, pp. 133–168. Plenum, New York.
Yank. R. C., J. Lis and R. Wu. (1979). Elution of DNA from agarose gels after electrophoresis. In *Methods in Enzymology*, ed. R. Wu, vol. 68, pp. 176–182. Academic, New York.
Yelverton, E., S. Norton, J. F. Obijeski and D. V. Goeddel (1983). Rabies virus glycoprotein analogs: biosynthesis in *Escherichia coli*. *Science*, **219**, 614–620.
Zakour, P. A. and L. A. Loeb (1982). Site-specific mutagenesis by error-directed DNA synthesis. *Nature* (*London*), **295**, 708–710.
Zasloff, M., G. D. Ginder and G. Felsenfeld (1978). A new method for the purification and identification of covalently closed circular DNA molecules. *Nucleic Acids Res.*, **5**, 1139–1152.
Zimmerman, U., P. Schreurich, G. Pilwat and R. Benz (1981). Cells with manipulated functions: new perspectives for cell biology, medicine and technology. *Angew. Chem.*, **20**, 325–344.
Zinn, K., P. Mellon, M. Ptashne and T. Maniatis (1982). Regulated expression of an extrachromosomal human β-interferon gene in mouse cells. *Proc. Natl. Acad. Sci. USA*, **79**, 4897–4901.
Zoller, J. J. and M. Smith (1982). Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. *Nucleic Acids Res.*, **10**, 6487–6500.
Zubay, G. (1973). *In vitro* synthesis of proteins in microbial systems. *Annu. Rev. Genet.*, **7**, 267–287.