Molecular Cloning

A glaring problem in most areas of biochemical research is obtaining sufficient amounts of the substance of interest. For example, a 10 L culture of *E. coli* grown to its maximum titer will only contain about 7 mg of DNA polymerase I, and many other proteins in much lesser amounts. Furthermore, only rarely can as much as half of any protein originally present in an organism be recovered in pure form. Eucaryotic proteins are even more difficult to obtain because tissue samples are usually only available in small quantities. With regards to the amount of DNA present, the 10 L *E. coli* culture would contain about 0.1mg of any 1000 bp length chromosomal DNA but its purification in the presence of the rest of the chromosomal DNA would be a very difficult task. These difficulties have been greatly reduced through the development of **molecular cloning techniques**. These methods, which are also referred to as genetic engineering and **recombinant DNA technology**, deserve much of the credit for the enormous progress in biochemistry and the dramatic rise of the biotechnology industry.

The main idea of molecular cloning is to insert a DNA segment of interest into an autonomously replicating DNA molecule, a so-called cloning **vector**, so that the DNA segment is replicated with the vector. Cloning such a chimeric vector in a suitable host organism such as *E. coli* or yeast results in the production of large amounts of the inserted DNA segment. If a cloned gene is flanked by the properly positioned control sequences for RNA and protein synthesis, the host may also produce large quantities of the mRNA and protein specified by that gene. The techniques of genetic engineering are discussed in detail.

Plasmid-based cloning vectors

Plasmids are circular DNA duplexes of 1 to 200 kb that contain the requisite genetic machinery, such as replication origin, to permit their autonomous propagation in a bacterial host or in yeast. Plasmids may be considered molecular parasites but in many instances they benefit their host by providing functions, such as resistance to antibiotics, that the host lacks. Indeed, the widespread and alarming appearance, since antibiotics came into use, of antibiotic-resistant pathogens is a result of the rapid proliferation among these organisms of plasmids containing genes that confer resistance to antibiotics.

Some types of plasmids, which are present in one or a few copies per cell, replicate once per cell division as does the bacterial

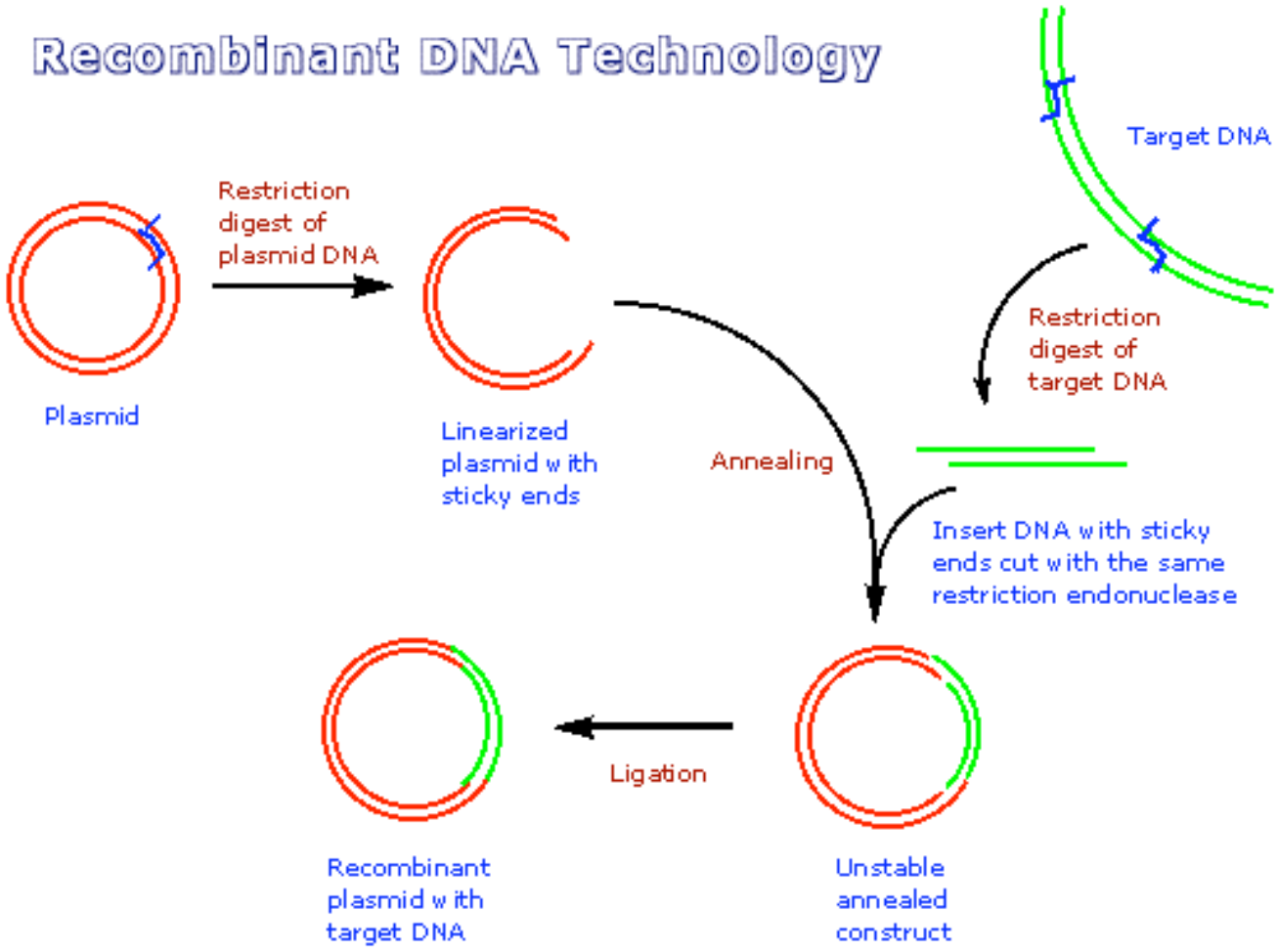
chromosome; their replication is said to be under stringent control. The plasmids used in molecular cloning, however, are under relaxed control; they are normally present in 10 to as many as 700 copies per cell. Moreover, protein synthesis in the bacterial host is inhibited, for example, by the antibiotic chloramphenicol, thereby preventing cell division. These plasmids continue to replicate until two or three thousand copies have accumulated per cell. **The plasmids that have been constructed for molecular cloning are relatively small, replicate under relaxed control, carry genes specifying resistance to one or more antibiotics, and contain a number of conveniently located restriction endonuclease sites into which the DNA to be cloned may be inserted.** Indeed, many plasmid vectors contain a strategically located short segment of DNA known as a polylinker that has been synthesized to contain a variety of restriction sites that are not present elsewhere in the plasmid. You can refer to your restriction map of the PVL1392 vector to visualize this.

The expression of a chimeric plasmid in a bacterial host was first demonstrated in 1973 by Herbert Boyer and Stanley Cohen. The host bacterium takes up a plasmid when the two are mixed together in a process that is greatly enhanced by the presence of divalent cations such as Ca^{2+} (which increase membrane permeability of DNA). An absorbed plasmid vector becomes permanently established in its bacterial host with an efficiency of around 0.1%.

Plasmid vectors cannot be used to clone DNAs of more than 10 kb. This is because the time required for plasmid replication increases with plasmid size. Hence intact plasmids with large unessential inserts are lost through the faster proliferation of plasmids that have eliminated these inserts by random deletions.

Creating of the recombinant plasmids containing the target DNA

A DNA to be cloned is, in many cases, obtained as a defined fragment through the application of **restriction endonucleases**. Recall that most restriction enzymes cleave duplex DNA at specific palindromic sequences so as to yield single-stranded sticky ends that are complementary to each other. Therefore, as Janet Mertz and Ron Davis first demonstrated in 1972, a restriction fragment may be inserted into a cut made in a cloning vector by the same restriction enzyme. The complimentary ends of the two DNAs specifically associate under **annealing** conditions and are covalently joined through action of the enzyme **DNA ligase**. An inherent advantage of this method is the ability to precisely excise the desired fragment through its restriction sites.



Following table contains several popular restriction endonucleases and their sequence specificities.

Endonuclease	Sequence cleaved	Bacterial source
BamH I	G|G-A-T-C-C C-C-T-A-G|G	*Bacillus amyloliquefaciens* H
Bgl II	A|G-A-T-C-T T-C-T-A-G|A	*Bacillus globigii*
EcoR I	G|A-A-T-T-C C-T-T-A-A|G	*Escherichia coli* R245
Hind III	A|A-G-C-T-T T-T-C-G-A|A	*Haemophillus influenzae* R_d
Nco I	C|C-A-T-G-G G-G-T-A-C|C	---
Not I	G-C|G-G-C-C-G-C C-G-C-C-G-G|C-G	---
Xho I	C|T-C-G-A-G G-A-G-C-T|C	---

Selection of the recombinant plasmid

How can one select only those host organisms that contain a properly constructed vector? In the case of plasmid transformation, this is usually done through the use of antibiotics and/or chromogenic substrates. For example, *E. coli* transformed with a plasmid containing a foreign DNA insert in its polylinker region that lacks β-galactosidase activity because the insert interrupts the protein-encoding sequence of the lacZ' gene. Thus, when grown in the presence of blue dye X-gal bacterial colonies that have an insert in their polylinker region form colorless colonies, whereas bacteria containing only plasmids that lack an insert form blue colonies. Addition of an antibiotic (such as ampicillin) will eliminate any colonies that did not take up the plasmid. The plasmids intact gene contains a gene that confers antibiotic resistance.

Production of proteins coded by target DNA

One of the greatest advantages of recombinant DNA technology is in the production of large quantities of scarce and/or novel proteins. This is a relatively straightforward procedure for bacterial proteins: A cloned structural gene must be inserted into an expression vector, a plasmid that contains the properly positioned transcriptional and translational control sequences for the protein's expression. With the use of a relaxed control plasmid and an efficient promoter, the production of protein of interest may reach 30% of the host's total cellular protein. Such genetically engineered organisms are called overproducers.

Bacterial cells often sequester such large amounts of useless and possibly toxic protein as insoluble and denatured inclusions result. Protein extracted from these inclusions must therefore be renatured, usually dissolving it in a guanidinium chloride or urea solution and then dialyzing away the denaturant. A strategy for circumventing this difficulty is to precede the protein of interest with the signal sequence of a bacterial protein. Such a protein is secreted into the bacterial periplasm with the concomitant removal of its signal sequence by a bacterial protease. Secreted proteins, which are relatively few in number, can be released into the medium by hypotonic disruption of the bacterial outer membrane, so their purification is greatly simplified relative to that of intracellular proteins.

The ability to synthesize a given protein in large quantities is already having enormous medical, agricultural, and industrial impact. Those that are in routine clinical use include human insulin, human growth hormone, and many other beneficial stimulators of human cells and organs.

REFERENCES

1. Voet, D. and Voet, J. G. Biochemistry John Wiley and Sons, Inc. New York, 1995.

GLOSSARY

- **Clone:** A large number of cells or molecules that are identical with a single parental cell or molecule.
- **Endonuclease:** An enzyme that cleaves internal phosphodiester bonds in DNA or RNA.
- **Ligation:** The enzyme-catalyzed joining in phosphodiester linkage of 2 stretches of DNA or RNA into one; the respective enzymes are DNA and RNA ligases.
- **Palindrome:** A sequence of duplex DNA that is the same when the 2 strands are read in the opposite directions.
- **Plasmid:** A small, extrachromosomal, circular molecule of DNA that replicates independently of the host DNA.
- **Recombinant DNA:** The altered DNA that results from the insertion of a sequence of deoxynucleotides not previously present into an existing molecule of DNA by enzymatic or chemical means.
- **Restriction enzyme:** An endonuclease that causes cleavage of both strands of DNA at highly specific sites dictated by the base sequence.
- **Sticky-ended DNA:** DNA with complementary single strands that protrude from opposite ends of a DNA duplex.
- **Vector:** A plasmid, bacteriophage or virus into which foreign DNA can be introduced for the purposes of cloning.