MOLEULAR BIOCHEMISTRY

In the past, efforts to understand genes and their expression have been confounded by the immense size and complexity of human DNA. The human genome contains DNA with approximately three billion (10^9) base pairs that encode 30,000 to 40,000 genes located on 23 pairs of chromosomes. It is now possible to determine the nucleotide sequence of long stretches of DNA, and essentially the entire sequence of the human genome is now known. This effort (called the Human Genome Project) was made possible by several techniques that have already contributed to our understanding of many genetic diseases. These include, first, the discovery of restriction endonucleases that permit the dissection of huge DNA molecules into defined fragments. Second, the development of cloning techniques, providing a mechanism for amplification of specific nucleotide sequences. Finally, the ability to synthesize specific probes, which has allowed the identification and manipulation of nucleotide sequences of interest. These and other experimental approaches have permitted the identification of both normal and mutant nucleotide sequences in DNA. This knowledge has led to the development of methods for the prenatal diagnosis of genetic diseases, and initial successes in the treatment of patients by gene therapy.

RESTRICTION ENDONUCLEASES

One of the major obstacles to molecular analysis of genomic DNA is the immense size of the molecules involved. The discovery of a special group of bacterial enzymes, called **restriction endonucleases (restriction enzymes),** which cleave double-stranded DNA into smaller, more manageable fragments, has opened the way for DNA analysis.

Because each enzyme cleaves DNA at a specific nucleotide sequence, restriction enzymes are used experimentally to obtain precisely defined DNA segments called **restriction fragments.**

Specificity of restriction endonucleases

Restriction endonucleases recognize short stretches of DNA (generally four or six base pairs) that contain specific nucleotide sequences. These sequences, which differ for each restriction endonuclease, are **palindromes,** that is, they exhibit two-fold rotational symmetry (Figure 32.2). This means that, within a short region of the double helix, the nucleotide sequence on the "top" strand, read 5'→3'is identical to that of the "bottom" strand, also read in the 5'→3' direction. Therefore, if you turn the page upside down that is, rotate it 180 degrees around its **axis of symmetry the** structure remains the same.

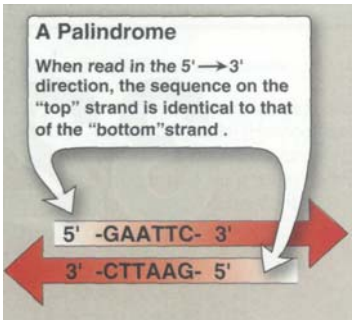

Figure 32.3
Recognition sequence of restriction endonuclease *Eco*RI shows two-fold rotational symmetry

Nomenclature

A restriction enzyme is named according to the organism from which it was isolated. The first letter of the name is from the genus of the bacterium. The next two letters are from the name of the species. An additional subscript letter indicates the type or strain, and a final number is appended to indicate the order in which the enzyme was discovered in that particular organism. For example, *Hae*III is the third restriction endonuclease isolated from the bacterium *Haemophilus aegyptius*.

"Sticky" and "blunt" ends

Restriction enzymes cleave DNA so as to produce a 3'-hydroxyl group on one end and a 5'-phosphate group on the other. Some restriction endonucleases, such as *Taq*\, form staggered cuts that produce "sticky" or cohesive ends—that is the resulting DNA fragments have single-stranded sequences that are complementary to each other (Figure 32.3). Other restriction endonucleases, such as *Hae*III cleave in the middle of their recognition sequence that is, at the axis of symmetry (see Figures 32.2 and 32.3)—and produce fragments that have "blunt" ends that do not form hydrogen bonds with each other. Using the enzyme *DNA ligase*, sticky ends of a DNA fragment of interest can be covalently joined with other DNA fragments that have sticky ends produced by cleavage with the same restriction endonuclease (Figure 32.4). [Note: Another ligase, encoded by bacteriophage T4, can covalently join blunt-ended fragments.] The hybrid combination of two fragments is called a **recombinant DNA molecule.**

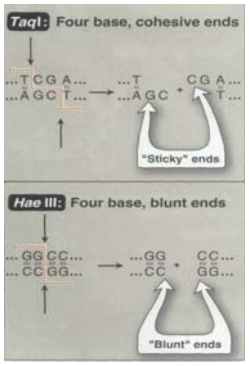

Figure 32.3

Specificity of TaqI and HaeIII restriction endonucleases.

Restriction sites

A DNA sequence that is recognized by a restriction enzyme is called a restriction site. These sites are recognized by restriction endonucleases that cleave DNA into fragments of different sizes. For example, an enzyme that recognizes a specific four base-pair sequence

produces many cuts in the DNA molecule. In contrast, an enzyme requiring a unique sequence of six base pairs produces fewer cuts and, hence, longer pieces. Hundreds of these enzymes, having different cleavage specificities (varying in both nucleotide sequences and length of recognition sites), are commercially available as analytic reagents.

DNA CLONING

Introduction of a foreign DNA molecule into a replicating cell permits the amplification (that is, production of many copies) of the DNA. In some cases, a single DNA fragment can be isolated and purified prior to cloning. More commonly, to clone a nucleotide sequence of interest, the total cellular DNA is first cleaved with a specific restriction enzyme, creating hundreds of thousand s of fragments. Therefore, individual fragments cannot be isolated. Instead, each of the resulting DNA fragments is joined to a DNA vector molecule (referred to as a **cloning vector)** to form a hybrid molecule. Each hybrid recombinant DNA molecule conveys its inserted DNA fragment into a single host cell, for example, a bacterium, where it is replicated (or **"amplified").** As the host cell multiplies, it forms a clone in which every bacterium carries copies of the same inserted DNA fragment, hence, the name **"cloning."** The cloned DNA is eventually released from its vector by cleavage (using the appropriate restriction endonuclease) and is isolated. By this mechanism, many identical copies of the DNA of interest can be produced. [Note: An alternative to cloning the **Polymerase Chain** Reaction (PCR)].

Vectors

A vector is a molecule of DNA to which the fragment of DNA to be cloned is joined. Essential properties of a vector include:

1) it must be capable of autonomous replication within a host cell, 2) it must contain at least one specific nucleotide sequence recognized by a restriction endonuclease, and 3) it must carry at least one gene that confers the ability to select for the vector, such as an antibiotic resistance gene. Commonly used vectors include **plasmids** and **bacterial and animal viruses.**

Prokaryotic plasmids:

Prokaryotic organisms contain single, large, circular chromosomes. In addition, most species of bacteria also normally contain small, circular, extrachromosomal DNA molecules called plasmids (Figure 32.5). Plasmid DNA undergoes replication that may or may not be synchronized to chromosomal division. Plasmids may carry genes that convey antibiotic resistance to the host bacterium, and may facilitate the transfer of genetic information from one bacterium to another.

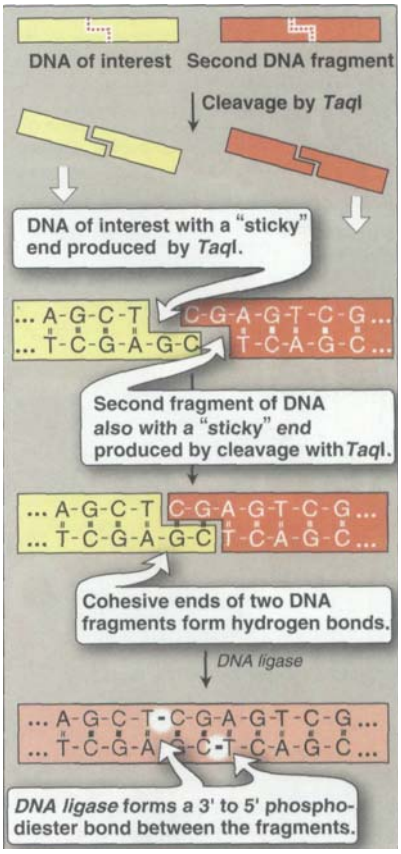

Figure 32.4

Formation of recombinant DNA from restriction fragments with sticky ends.

[Note: Bacteria are grown in the presence of antibiotics, thus selecting for cells containing the hybrid plasmids, which provide antibiotic resistance.] Plasmids can be readily isolated from bacterial cells, their circular DNA cleaved at specific sites by restriction endonucleases, and foreign DNA inserted into the circle. The hybrid plasmid can be reintroduced into a bacterium, and large numbers of copies of the plasmid containing the foreign DNA produced. [Note: The experiment is conducted to favor only one DNA fragment being inserted into each plasmid and only one plasmid being taken up by each bacterium.]

Other vectors

The development of improved vectors that can more efficiently accommodate large DNA segments, or express the passenger genes in different cell types, is an ongoing endeavor of molecular genetics research. In addition to the prokaryotic plasmids described above, bacteriophage lambda (λ) yeast artificial chromosomes (YACs), and mammalian viruses (retroviruses, for example) are currently in wide use as cloning vectors **(Champe, Harvey and Ferrier, 2005)**

.

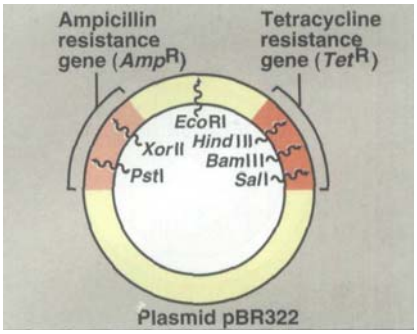

Figure 32.5

A restriction map of plasmid pBR322 indicating the positions of its antibiotic resistance genes and the sites of nucleotide sequences recognized by specific restriction endonucleases.