CHAPTER 3
Basic Tools and Techniques of DNA Science

The comedy movie "Sleeper," which tells of a totalitarian society of the 21st century, contains a cloning scene that is hard to forget. "The Leader" has been assassinated, but his nose has been kept alive through the wonders of futuristic science. The bungling Miles Monroe (Woody Allen) and his sidekick Luna Schlosser (Diane Keaton) masquerade as surgeons who know the procedure to reproduce a new Leader from the saved nose cells. Asked if he minds performing the operation while numerous doctors observe in the surgical theater, Miles replies, "Never clone alone." Setting the nose in its proper position on a person-shaped cloning table, he confirms his mission: "What you want, basically, is a whole entire person connected to that nose, right? Do you want me to leave room for a mustache?"

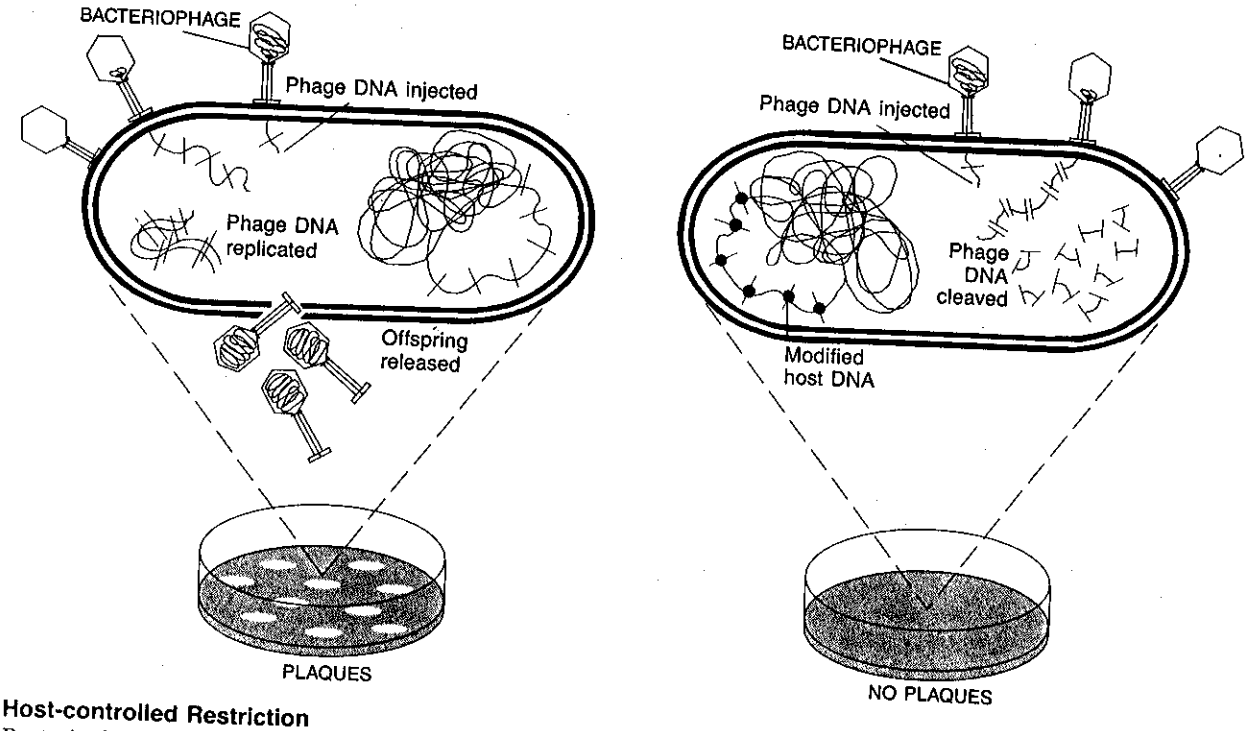
This scene satirizes the hype and misconception of the word cloning. It conjures up images of regenerating an entire human being, or other complex organism, from a small sample of body tissue. In theory, this is possible. Since every cell carries in its chromosomes the same basic set of genetic instructions as in the fertilized ovum, somatic cells that make up body tissues could provide the raw genetic information needed to recreate "from scratch" an entire organism.

Woody Allen and Diane Keaton as Cloners in "Sleeper", 1973
(Courtesy of the Museum of Modern Art Film Stills Archive.)

There are real-life examples of this sort of cloning. The mature cells from certain plants *can* be stimulated to produce a new adult plant. All animals have some ability to regenerate injured cells and, in some cases, even entire body parts. Amphibians and reptiles, for example, can regenerate new tails. In practice, however, fully differentiated cells taken from most adult organisms lack the regenerative ability of plant cells and cannot be stimulated to "clone" a complete organism. In the vast majority of higher organisms, only the fertilized ovum is able to progress through the complex developmental stages that culminate in a fully functional adult. Technology for cloning mammals from single cells does not now exist, nor does it seem very close at hand.

Taken in its pure sense, the term cloning refers to the reproduction of daughter cells by fission or mitotic division. During these processes, DNA from a parent cell is replicated, and identical sets of genetic information are passed on to daughter cells. Successive generations of cells in turn divide, giving rise to a population of genetically identical *clones*, all derived from a single ancestral cell.

Gene cloning uses the natural replicating ability of cells to isolate and duplicate an individual gene. First, the gene of interest is inserted into a carrier DNA molecule, termed a vector, that can self-replicate within a host cell. Then the vector, with its gene insert, is introduced ("transformed") into an appropriate host cell. Subsequent mitosis of the host cell creates a population of clones, each containing the gene of interest.



Host-controlled Restriction
Bacteriophage infects strain at left, producing clear plaques in bacterial lawn. Resistant strain at right possesses a restriction enzyme that cleaves incoming phage DNA, as well as a modifying enzyme that methylates its own DNA to protect it from cleavage. (Art concept developed by Lisa Shoemaker.)

Restriction-Modification

In the early 1950s, Salvador Luria and Giuseppe Bertani (University of Illinois) and Jean Weigle (California Institute of Technology) found evidence for a sort of primitive immune system in bacteria. They observed that certain strains of the bacterium *Escherichia coli* are resistant to infection by various bacteriophages. The phenomenon seemed to be a property of the bacterial cell, which is able to *restrict* the growth and replication of phages. In 1962, Werner Arber, at the University of Geneva, provided the first evidence that the resistant bacterium possesses an enzyme system that selectively recognizes and destroys foreign phage DNA within the bacterial membrane and that also *modifies* the chromosomal DNA of the bacterium to prevent self-destruction.

Several years later, Arber and his associate Stuart Linn, as well as Matthew Meselson and Robert Yuan at Harvard University, isolated *E. coli* extracts that efficiently cleave phage DNA. These extracts contained the first known *restriction endonucleases*, enzymes that attack and digest internal regions of the DNA of an invading bacteriophage but not that of the host. This was shown to occur because, in addition to a restriction (cutting) activity, these enzymes also possess a modification (protecting) activity.

It was later found that the modifying enzyme protects host DNA from digestion by adding methyl groups to a nucleotide within the sequence recognized by the restriction enzyme. This modification of the sequence blocks the restriction enzyme from digesting the DNA. Although typically both strands of host DNA are methylated, bacterial DNA is protected from digestion even when one strand is methylated. Therefore, during DNA replication, the methylated parental strand protects the molecule until the newly synthesized daughter strand can also be methylated.

Because their cutting activity is not precise, the enzymes isolated by Arber, Linn, Meselson, and Yuan were of no practical value as tools for manipulating DNA. Although these enzymes recognize specific nucleotide sequences, they cut the DNA molecule randomly at positions that may be thousands of nucleotides distant from the recognition site.

The phenomenon of restriction-modification thus remained of purely academic interest until 1970, when Hamilton Smith and his student Kent Wilcox, at Johns Hopkins University, isolated a new restriction endonuclease from *Haemophilus influenzae*. The restriction activity of this enzyme, named *Hin*dII, differs from those previously described in two important ways. First, the restriction activity is separate from the modification activity. Second, it cleaves DNA predictably, cutting *within* its recognition sequence.

Daniel Nathans, a colleague of Smith's at Johns Hopkins, then showed the broad applicability of restriction endonucleases. He used the *Hin*dII enzyme to cut the purified DNA of a small virus that infects monkeys (simian virus 40, or SV40) and separated the resulting *restriction fragments* by size in an electrical field. He deduced the order of the

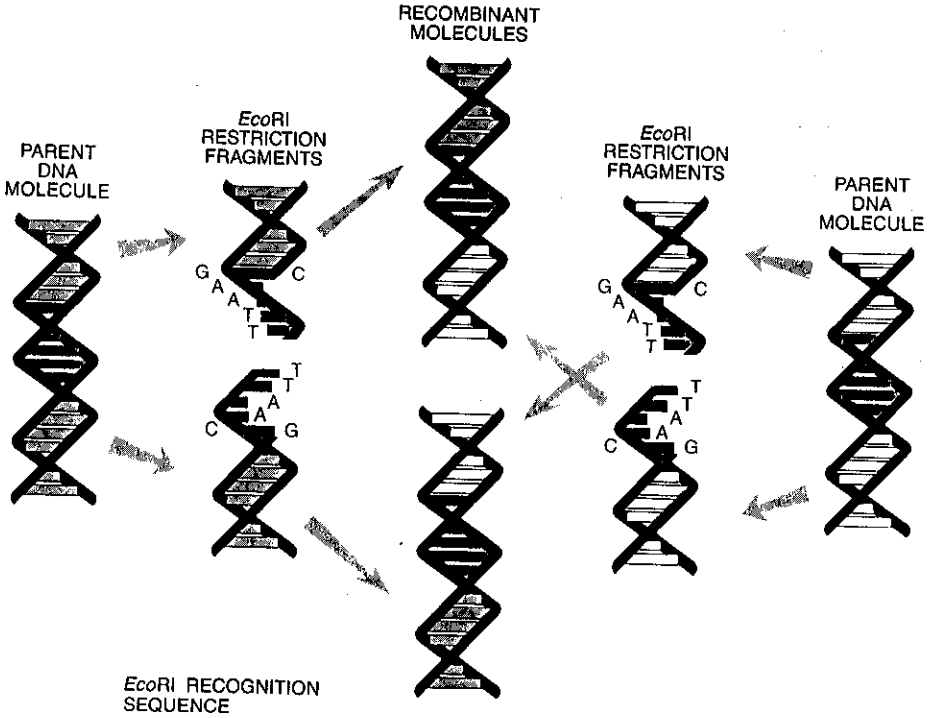
fragments (and corresponding restriction sites) in the 5000-nucleotide circular chromosome, creating a *restriction map* that was then related to the existing genetic map of SV40.

The First Recombinant DNA Molecules

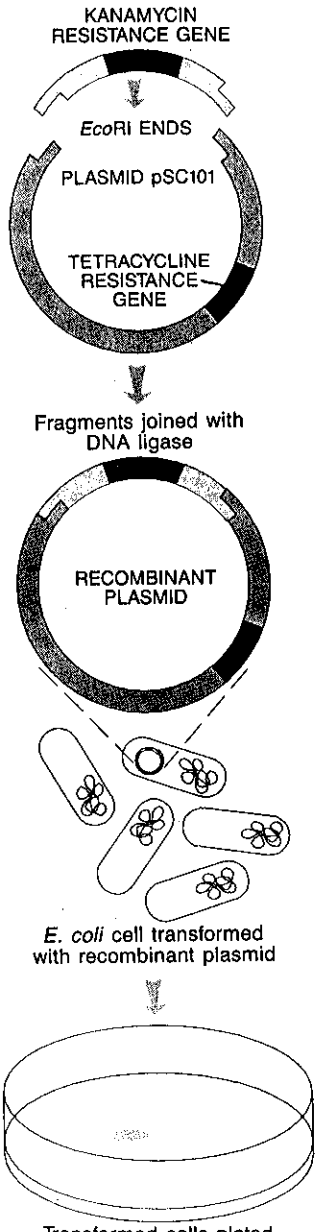
In the early 1970s, the first recombinant molecules were produced independently using two methods. One method was discovered at Stanford University under the direction of Paul Berg; the other was developed jointly by the laboratories of Stanley Cohen, also at Stanford, and Herb Boyer at the University of California, San Francisco.

In 1972, Berg's group worked out the "tailing" method of joining DNA molecules, which was modeled after the "sticky ends" found at the chromosome ends of the bacteriophage λ. These complementary, single-stranded regions of DNA join spontaneously to form a circular DNA molecule. Using the restriction enzyme *Eco*RI that had been recently isolated from *E. coli*, they cut the circular SV40 chromosome and a small, circular chromosome (termed a plasmid) from *E. coli*. The restriction enzyme cuts each molecule in only a single place, opening the DNA loops to produce linear strands.

To produce sticky ends, a single-stranded "tail" of 50–100 adenine residues was added to the ends of the SV40 DNA using the enzyme terminal transferase. A tail of thymine residues was added to the plasmid DNA by the same method. When the two DNAs were mixed together, the complementary adenine and thymine tails base-paired to form a



Joining Complementary "Sticky Ends" of *Eco*RI Restriction Fragments



The Boyer-Cohen-Chang Experiment, 1973

circular, recombinant DNA molecule. Two enzymes completed the job: DNA polymerase filled in any single-stranded gaps, and DNA ligase sealed the junction points between the SV40 and plasmid fragments.

Berg had planned to introduce his recombinant plasmid into animal cells and then look for expression of the bacterial genes in the infected animal cells. We know in hindsight that this would have failed, because expression of a foreign gene requires a detailed understanding of its regulation (see Chapter 4). Regardless, Berg never got the opportunity to test his new DNA molecule. Concern in the scientific community over the potential danger of the experiment halted research on gene transfer into mammalian cells for several years.

While not completed, Berg's experiment provided two key pieces of the recombinant DNA puzzle. He showed that a restriction enzyme can be used to cut DNA in a predictable manner and that DNA fragments from different organisms can be joined together. In 1973, Stanley Cohen and Annie Chang carried the Berg experiment a step further, adding the final piece to the puzzle. They showed that a recombinant DNA molecule can be maintained and replicated within *E. coli*.

Cohen and Chang purified from *E. coli* a plasmid that incorporated several important features: (1) a unique restriction recognition site, to allow the molecule to be cut at a single location by the restriction enzyme *Eco*RI; (2) a nucleotide sequence, the origin of replication, to allow the plasmid to be replicated within a host bacterial cell; and (3) a gene coding for resistance to the antibiotic tetracycline, to allow for selection of bacteria incorporating the molecule. The new plasmid was named pSC101 (SC for Stanley Cohen).

The first step was to devise a method to introduce this plasmid into *E. coli* efficiently. Using a modification of a transformation protocol described 3 years earlier by Morton Mandel and A. Higa (at the University of Hawaii), they inserted pSC101 into *E. coli* cells. The transformed cells were then spread on nutrient agar plates containing tetracycline; the appearance of colonies of bacteria with resistance to the antibiotic showed that the plasmid had been taken up and expressed.

The next step was to construct and insert a recombinant plasmid. For this, a second plasmid, pSC102, was isolated that contained an antibiotic resistance gene for kanamycin. Herb Boyer's laboratory had recently demonstrated that *Eco*RI creates its own sticky ends when it cuts, making Berg's tailing method unnecessary. So, pSC101 and pSC102 were simply cut with *Eco*RI, mixed together, and rejoined with DNA ligase. The resulting recombinant plasmid was transformed into *E. coli* cells, which were plated on media containing both tetracycline and kanamycin. The appearance of colonies with *dual* resistance confirmed that a recombinant DNA molecule had been introduced successfully into living bacterial cells.

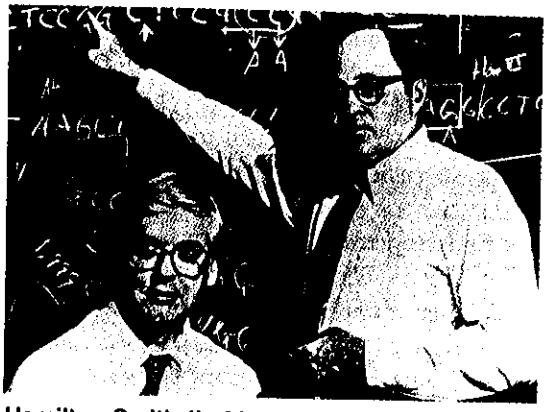
Shortly thereafter, Cohen and Boyer teamed up to produce a recombinant molecule containing DNA from two different species. To accomplish this, they spliced a gene encoding a ribosomal RNA from the toad *Xenopus laevis* into the plasmid pSC101 and transformed the recombined DNA into *E. coli*. Some of the colonies resistant to tetracycline were found to contain the toad ribosomal RNA gene.

The recombinant DNA techniques introduced by Berg, Chang, Boyer, and Cohen enabled biologists to change the genetic constitution of a living thing in a controlled manner and to transcend established species barriers. Each novel DNA combination creates a new biological entity with altered genetic and biochemical characteristics.

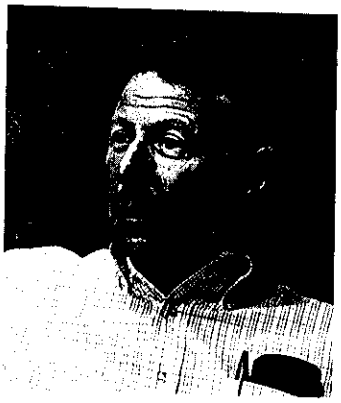
Restriction Endonucleases

Restriction endonucleases, or restriction enzymes, are used as molecular scalpels to cut DNA in a precise and predictable manner. They are members of the class of nucleases, which display the general property of breaking the phosphodiester bonds that link adjacent nucleotides in DNA and RNA molecules. *Endo*nucleases cleave nucleic acids at internal positions, while *exo*nucleases progressively digest from the ends of nucleic acid molecules.

There are three major classes of restriction endonucleases. Type I and type III enzymes have both restriction (cutting) and modification (meth-



Hamilton Smith (Left) and Daniel Nathans, 1978



Paul Berg, 1980

Herbert Boyer, 1977

Stanley Cohen and Annie Chang, 1973

Key Figures in the Construction of the First Recombinant DNA Molecules
(Courtesy of (*Top*) The Baltimore Sun; News and Publication Service, Stanford University; (*Bottom*) Paul Conklin for Academy Forum; News and Publications Service, Stanford University.)

ylating) activity. Both types cut at sites some distance from their recognition sequences; ATP is required to provide energy for movement of the enzyme along the DNA molecule from recognition site to cleavage site. Type I enzymes, such as *Eco*K isolated by Meselson and Yuan, cut at random sites 1000 nucleotides or more away from the recognition sequence. Type III enzymes cut at specific sites quite near the recognition sequence, but these may be difficult to predict.

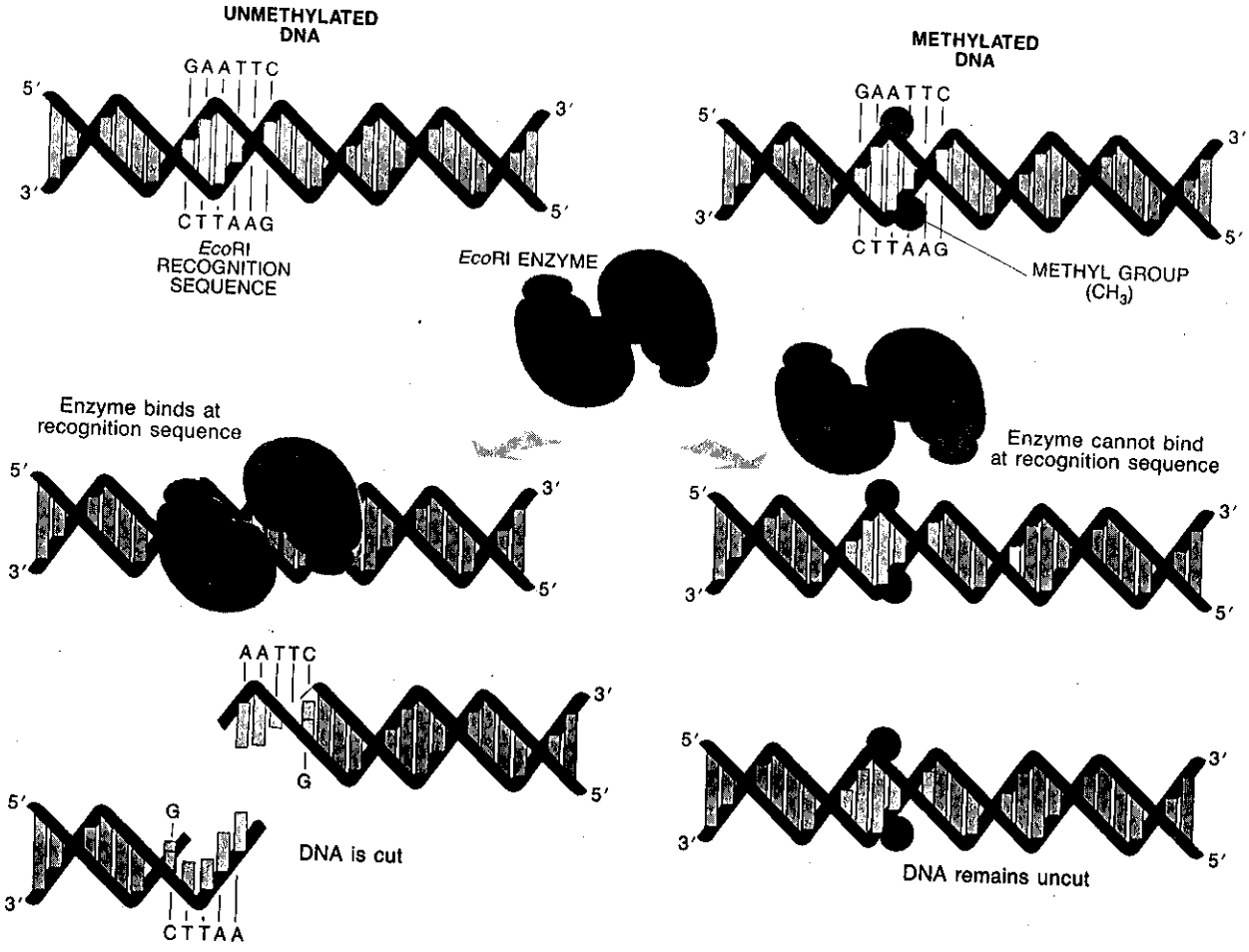
The restriction enzymes used in DNA science are invariably type II. They are most useful for several reasons: (1) Each has only restriction activity; modification activity is carried by a separate enzyme. (2) Each cuts in a predictable and consistent manner, at a site within or adjacent to the recognition sequence. (3) They require only magnesium ion (Mg^{++}) as a cofactor; ATP is not needed.

Today, more than 1200 type II enzymes have been isolated from a variety of prokaryotic organisms. Enzymes have been identified that recognize more than 130 different nucleotide sequences; more than 70 types are commercially available. To avoid confusion, restriction endonucleases are named according to the following nomenclature:

1. The first letter is the initial letter of the genus name of the organism from which the enzyme is isolated.
2. The second and third letters are usually the initial letters of the organism's species name. Since they are derived from scientific names, the first three letters of the endonuclease name are italicized.
3. A fourth letter, if any, indicates a particular strain of organism.
4. Originally, Roman numerals were meant to indicate the order in which enzymes, isolated from the same organism and strain, are eluted from a chromatography column. More often, though, the Roman numerals indicate the order of discovery.

*Eco*RI	*E*	=	genus *Escherichia*
	co	=	species *coli*
	R	=	strain RY13
	I	=	first endonuclease isolated
*Bam*HI	*B*	=	genus *Bacillus*
	am	=	species *amyloliquefaciens*
	H	=	strain H
	I	=	first endonuclease isolated
*Hin*dIII	*H*	=	genus *Haemophilus*
	in	=	species *influenzae*
	d	=	strain Rd
	III	=	third endonuclease isolated

By some unknown mechanism, a type II restriction endonuclease scans a DNA molecule, stopping only when it recognizes a specific sequence of nucleotides. Most restriction enzymes recognize a four- or six-nucleotide sequence. Assuming that the four component nucleotides (A,C,T,G) are distributed randomly within a DNA molecule, then any four nucleotides



Molecular Detail of *Eco*RI Restriction-Modification
(Art concept developed by Lisa Shoemaker.)

will occur, on average, every 256 nucleotides ($4 \times 4 \times 4 \times 4$), and a six-nucleotide recognition site will occur every 4096 nucleotides ($4 \times 4 \times 4 \times 4 \times 4 \times 4$).

Many restriction enzymes have recognition sites that are composed of symmetrical, or palindromic, nucleotide sequences. This means that the recognition sequence read forward on one DNA strand is identical to the sequence read in the opposite direction on the complementary strand. Put another way, the 5′→3′ sequence is identical on each DNA strand.

In a general sense, the terms 5′ and 3′ refer to either end of a single DNA strand. Specifically, they designate carbon atoms on opposite sides of the deoxyribose ring that are joined to form the DNA polymer. The 3′ carbon is linked, through phosphoester bonds with an intervening phosphate, to the 5′ carbon of the adjacent nucleotide. By convention, the nucleotide sequence is "read" from the 5′ end to the 3′ end. In duplex DNA, it is usual to show only one of the strands, since the complementary strand can be deduced from the base-pairing rules. Note that in duplex DNA, the strands are antiparallel; that is, 5′→3′ reads in opposite directions on the two complementary strands.

Within or very near the recognition site, the restriction enzyme catalyzes a hydrolysis reaction that uses water to break a specific phosphodiester linkage on each strand of the DNA helix. Two DNA fragments are produced, each with a phosphate group at the 5′ end and a hydroxyl group at the 3′ end.

For each type II endonuclease, there is a corresponding modifying enzyme that blocks restriction activity by methylating specific nucleotides within the recognition sequence. The protruding methyl group presumably prevents catalysis by interfering with the close molecular interaction between the restriction enzyme and its recognition site. *Eco*RI methylase, for example, adds a methyl group to the second adenine residue within the *Eco*RI recognition site.

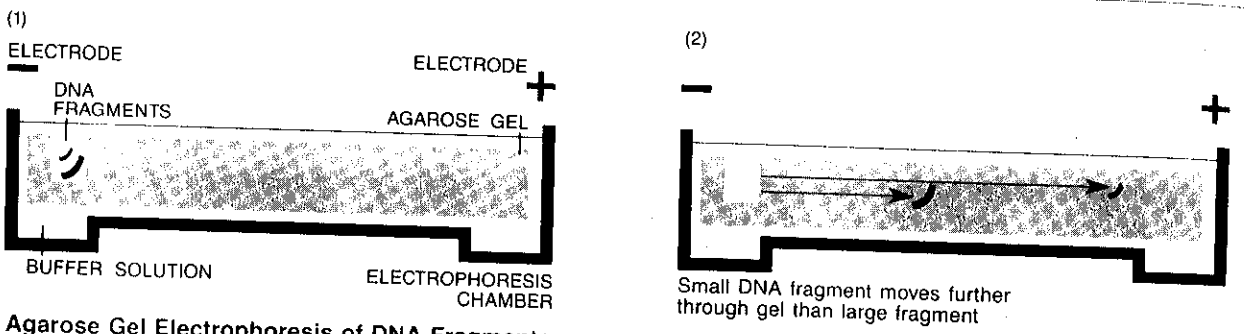
Some endonucleases, such as *Hin*dII, cut cleanly through the DNA helix by cleaving both complementary strands at the same nucleotide position, typically in the center of the recognition site. These enzymes leave flush- or blunt-ended fragments. Other endonucleases cleave each strand off-center in the recognition site, at positions two to four nucleotides apart. This creates fragments with exposed ends of short, single-stranded sequences. Various enzymes leave single-stranded "overhangs" on either the 5′ or 3′ ends of the DNA fragments. *Eco*RI, *Bam*HI, and *Hin*dIII, for example, each leave 5′ overhangs of four nucleotides.

Single-stranded overhangs, also called cohesive or "sticky" ends, are extremely useful in making recombinant DNA molecules. These exposed nucleotides serve as a template for realignment, allowing the complementary nucleotides of two like restriction fragments to hydrogen bond to one another. A given restriction enzyme cuts all DNA in exactly the same fashion, regardless of whether the source is a bacterium, a plant, or a human being. Thus, any sticky-ended fragment can be recombined with any other fragment generated by the same restriction enzyme.

Specificities of Some Typical Restriction Endonucleases

Source	Enzyme	Recognition sequence	Number of cleavage sites		
			λ	Adenovirus-2	SV40
Bacillus amyloliquefaciens H	*Bam*HI	G↓GATCC	5	3	1
Bacillus globigii	*Bgl*II	A↓GATCT	6	12	0
Escherichia coli RY13	*Eco*RI	G↓A*ATTC	5	5	1
Escherichia coli R245	*Eco*RII	↓CCTGG	>35	>35	16
Haemophilus aegyptius	*Hae*III	GG↓CC*	>50	>50	19
Haemophilus influenzae R_d	*Hin*dII	GTPy↓PuAC	34	>20	7
Haemophilus influenzae R_d	*Hin*dIII	A↓AGCTT	6	11	6
Haemophilus parainfluenzae	*Hpa*II	C↓CGG	>50	>50	1
Nocardia otitidis-caviarum	*Not*I	GC↓GGCCGC	0	7	0
Providencia stuartii 164	*Pst*I	CTGCA↓G	18	25	2
Serratia marcescens S_b	*Sma*I	CCC↓GGG	3	12	0

Recognition sequences are written 5′ to 3′. Only one strand is represented. The arrows indicate cleavage sites. Pu (purine) denotes that either A or G will be recognized. Py (pyrimidine) denotes that either C or T will be recognized. Asterisks represent positions where bases can be methylated.



Agarose Gel Electrophoresis of DNA Fragments
(Art concept developed by Lisa Shoemaker.)

Agarose Gel Electrophoresis

Polyacrylamide gel electrophoresis had been used for many years to separate RNA and protein molecules by molecular weight. In 1970, Daniel Nathans used polyacrylamide gel electrophoresis as a simple and rapid means to separate DNA restriction fragments. Prior to this, DNAs of different sizes were separated by velocity-sedimentation ultracentrifugation, a laborious method that determines only the relative sizes of DNA restriction fragments. Whereas centrifugation uses gravitational force to separate molecules, electrophoresis means literally *to carry with electricity*.

Gel electrophoresis takes advantage of the fact that, as an organic acid, DNA is negatively charged. DNA owes its acidity to phosphate groups that alternate with deoxyribose to form the rails of the double helix ladder. In solution, at neutral pH, negatively charged oxygens radiate from phosphates on the outside of the DNA molecule. When placed in an electric field, DNA molecules are attracted toward the positive pole (anode) and repelled from the negative pole (cathode).

During electrophoresis, DNA fragments sort by size in the polyacrylamide gel. The porous gel matrix acts as a molecular sieve through which smaller molecules can move more easily than larger ones; thus, the distance moved by a DNA fragment is inversely proportional to its molecular weight. In a given period of time, smaller restriction fragments migrate relatively far from the origin compared to larger fragments. Because of its small pore size, polyacrylamide efficiently separates small DNA fragments of up to 1000 nucleotides. However, this level of resolution was inappropriate for isolating gene-sized fragments of several thousands of nucleotides.

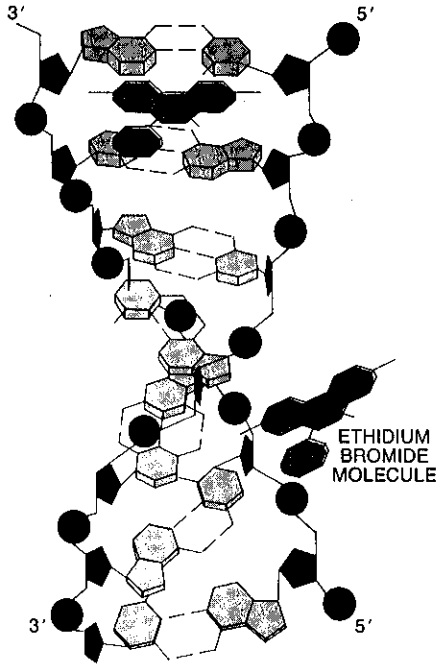
Although less time consuming than centrifugation, early polyacrylamide gel electrophoresis was still labor intensive and required the use of radioactively labeled DNA fragments. Following electrophoresis, the polyacrylamide gel was cut into many bands, and the amount of radioactivity in each slice was determined in a scintillation counter. The pattern of radioactivity was used to reconstruct the pattern of DNA bands in the gel.

A research team at Cold Spring Harbor Laboratory, led by Joseph

Sambrook, introduced two important refinements to DNA electrophoresis that made possible rapid analysis of DNA restriction patterns. First, they replaced polyacrylamide with agarose, a highly purified form of agar. An agarose matrix can efficiently separate larger DNA fragments ranging in size from 100 nucleotides to more than 50,000 nucleotides. DNA fragments in different size ranges can be separated by adjusting the agarose concentration. A low concentration (down to 0.3%) produces a loose gel that separates larger fragments, whereas a high concentration (up to 2%) produces a stiff gel that resolves small fragments.

Second, they used a fluorescent dye, ethidium bromide, to stain DNA bands in agarose gels. Following a brief staining step, the fragment pattern is viewed directly under ultraviolet (UV) light. This technique is extremely sensitive; as little as 5 ng (0.005 μg) of DNA can be detected. Thus, it is not difficult to understand why ethidium bromide staining quickly replaced radioactive labeling for the routine analysis of DNA restriction patterns.

Currently used methods are identical to those published by the Cold Spring Harbor team in 1973. Molten agarose is poured into a casting tray in which a plastic or Plexiglas comb is suspended. As it cools, the agarose hardens to form a Jell-O-like substance consisting of a dense network of cross-linked molecules. The solidified gel slab is immersed in a chamber filled with buffer solution, which contains ions needed to conduct electricity. When the gel solidifies, the comb is removed, leaving a number of wells into which DNA samples are loaded. Just prior to loading, the digested DNA is mixed with a loading solution that consists of sucrose and one or more visible dyes. The dense sucrose solution weights the DNA sample, helping it to sink when loaded into a well.



Intercalation of Ethidium Bromide into DNA Helix

Current supplied through electrodes at either end of the chamber creates an electric field across the gel. The negatively charged DNA fragments move from the wells into the gel, migrating through the pores in the matrix toward the positive pole. The negatively charged dye molecules do not interact with the DNA, but migrate independently toward the positive pole. For example, the commonly used marker bromophenol blue migrates at a rate equivalent to a DNA fragment of approximately 300 nucleotides (in a 1% gel). The visible movement of the dye allows one to monitor the relative migration of the unseen DNA bands.

Following electrophoresis, the gel is soaked in a dilute solution of ethidium bromide. The stain diffuses throughout the gel, becoming highly concentrated in regions where it binds to DNA fragments. (Alternately, ethidium bromide is incorporated into the gel and buffer prior to beginning electrophoresis.) A planar molecule, the ethidium bromide intercalates between the stacked nucleotides of the DNA helix, staining DNA bands in the gel.

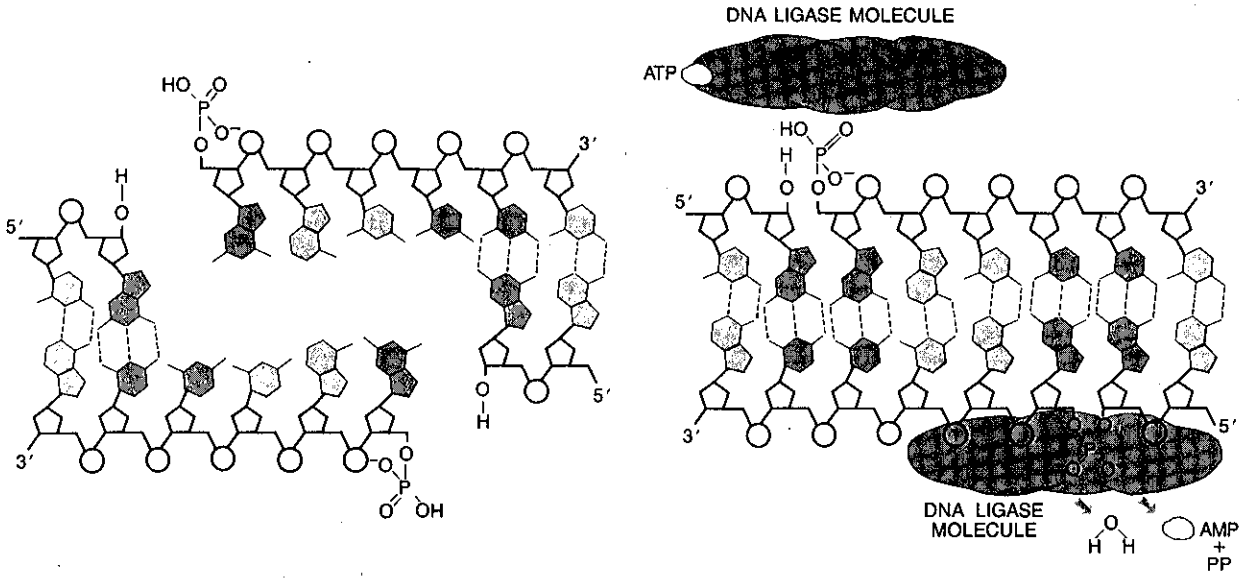
The stained gel is then exposed to medium-wavelength UV light. The DNA/ethidium bromide complex strongly absorbs UV light at 300 nm, retains some of the energy, and reemits visible light in the orange range at 590 nm. Under UV illumination, the stained restriction fragments appear as fluorescent orange bands in the gel. It is important to understand that a band of DNA seen in a gel is not a single DNA molecule. Rather, the band is a collection of millions of identical DNA molecules, all of the same nucleotide length.

DNA Recombination and Ligation

A key step of gene cloning is to recombine a gene of interest into a plasmid vector. Ideally, a gene is isolated on a restriction fragment created by two different endonucleases that cleave on either side of the gene and that generate distinctive single-stranded ends. The sticky ends of the restriction fragment are then rejoined to the complementary ends of a plasmid vector that has been opened up with the same two enzymes. Such "directional cloning" using two different enzymes produces restriction fragments that have noncomplementary overhangs at each end. This prevents any fragment from rejoining its own ends and encourages recombination between different fragments.

The double digestion of the source DNA produces two or more restriction fragments. The entire mixture of fragments can be ligated directly into plasmid vectors. Alternately, if one begins with a small DNA, such as a plasmid or virus, gel electrophoresis can be used to "gel-purify" a particular restriction fragment of interest. The cut DNA is electrophoresed, and the band containing the desired DNA is cut from the agarose gel. The DNA fragment is eluted from the gel and can be ligated into an appropriate vector.

The single-stranded overhang of a sticky end can form hydrogen bonds with the complementary nucleotides in the overhang of another fragment generated by the same restriction enzyme. Hydrogen bonding of several



Molecular Detail of DNA Ligation by T4 Ligase
Hydrogen bonding between complementary nucleotides aligns *Bam*HI fragments while ligase reforms phosphodiester bonds on each side of the DNA molecule.

nucleotides is not sufficient to form a stable molecule, so associations between complementary ends constantly form and break. This transient interaction, however, does hold the two restriction fragments together long enough for DNA ligase to re-form phosphodiester bonds between adjacent nucleotides. This covalently links the deoxyribose-phosphate rails of the two fragments into a stable double helix. During the ligation reaction, an ester linkage is formed between the terminal phosphate of the 5′ overhang of one fragment and the adjacent deoxyribose ring at the 3′ end of the second fragment. This is accompanied by the loss of one molecule of water, making ligation an example of a condensation reaction.

The Host Cell: The Bacterium *Escherichia coli*

The manipulation of DNA and creation of recombinant DNA molecules described so far takes place in the test tube. Ultimately, the propagation of a DNA sequence must take place inside a living cell. Thus, transformation—the cellular uptake and expression of DNA in a bacterium—is crucial to the research process. The following elements are required to make the transformation process efficient and controllable enough to be of general use for introducing foreign genes into living cells: (1) a suitable host organism in which to insert the gene, (2) a self-replicating vector to carry the gene into the host organism, and (3) a means of selecting for host cells that have taken up the gene.

The bacterium *E. coli* has become the most widely used organism in molecular biology because it provides a relatively simple and well-understood genetic environment in which to isolate foreign DNA. Its primary genetic complement is contained on a single chromosome of

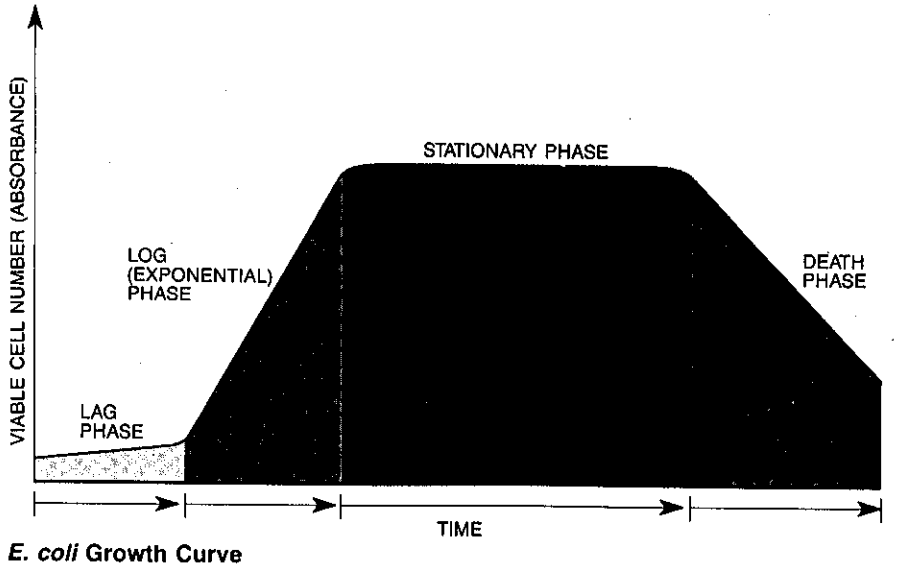
approximately 5 million base pairs, making it only 1/600th the size of a haploid set of human chromosomes (the human genome). The chromosomal locations and sequences of a large number of its genes are known.

Because the genetic code is nearly universal, *E. coli* can accept foreign DNA derived from any organism. The DNA of a bacterium, a human, a corn plant, or a fruit fly is constructed of the same four nucleotides (adenosine, cytosine, guanosine, and thymidine), is assembled in the same structure, and is replicated by the same basic mechanism. Each organism transcribes DNA into messenger RNA, which is in turn translated into proteins according to the genetic code. A foreign gene inside *E. coli* is replicated, and in some cases translated, in exactly the same manner as the native bacterial DNA. *E. coli* "sees" foreign DNA as its own.

Under the best of circumstances, the uptake of a specific foreign gene is a relatively rare occurrence and is thus most easily accomplished in a large population of organisms that are reproducing rapidly. *E. coli* is an ideal genetic organism in this regard.

A recombinant plasmid is biologically amplified when a transformed bacterium replicates by binary fission to create a clone of identical daughter cells. Under favorable conditions, *E. coli* replicates once every 22 minutes, giving rise to 30 generations and more than 1 billion cells in 11 hours. This number of cells can be contained in a single milliliter of culture solution. Moreover, since each bacterium can carry up to several hundred copies of a cloned gene, the foreign DNA sequence is potentially amplified by a factor of several hundred billion.

E. coli is a constituent of the normal bacterial fauna that inhabits the human colon, where it absorbs digested food materials. Thus, it grows best with incubation at 37°C in a culture medium that approximates the nutrients available in the human digestive tract. An example of such a medium is LB broth, which contains carbohydrates, amino acids, nucleotide phosphates, salts, and vitamins derived from yeast extract and milk protein.



***E. coli* Growth Curve**

Bacterial growth falls into several distinct phases. During *lag phase*, cells adjust to the nutrient environment and gear up for rapid proliferation; little or no cellular replication takes place. During *logarithmic (log) phase*, the culture grows exponentially and the cell number doubles every 22 minutes. During *stationary phase*, the cell number remains constant as new cells are produced at the same rate as old cells die. After an extended period, the culture enters *death phase;* the number of viable cells decreases as nutrients deplete and wastes accumulate.

Masses of bacterial cells are grown in a suspension culture; shaking provides aeration and keeps cells suspended in the medium. To isolate individual colonies, cells are spread on the surface of LB agar plates. Although the individual cells are invisible to the naked eye, after plating onto solid medium, each cell divides to form a visible colony of identical daughter cells in 12–24 hours.

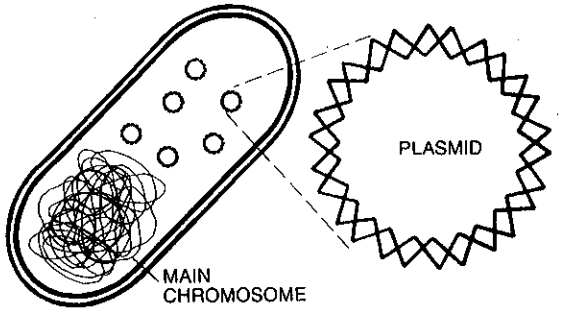
The Plasmid Vector

In medical terminology, a vector is an organism that carries a pathogen from one host organism to another. In molecular biology, a vector is a DNA molecule that is used as a vehicle to carry foreign DNA sequences into *E. coli* or another host cell.

Plasmids are the simplest bacterial vectors. Ranging in length from 1000 to 200,000 base pairs, they are circular DNA molecules that exist separate from the main bacterial chromosome. To be propagated through successive bacterial generations, the plasmid vector must contain specific DNA sequences that allow it to be replicated within the host cell. DNA polymerase and other proteins required to initiate DNA synthesis bind to this region, which is called the origin of replication.

Plasmids can be divided into two broad groups, according to how tightly their replication is regulated. Plasmids that are under *stringent* control only replicate along with the main bacterial chromosome and so exist as a single copy, or at most several copies, within the cell. *Relaxed* plasmids, on the other hand, replicate autonomously of the main chromosome and have copy numbers of 10–500 per cell.

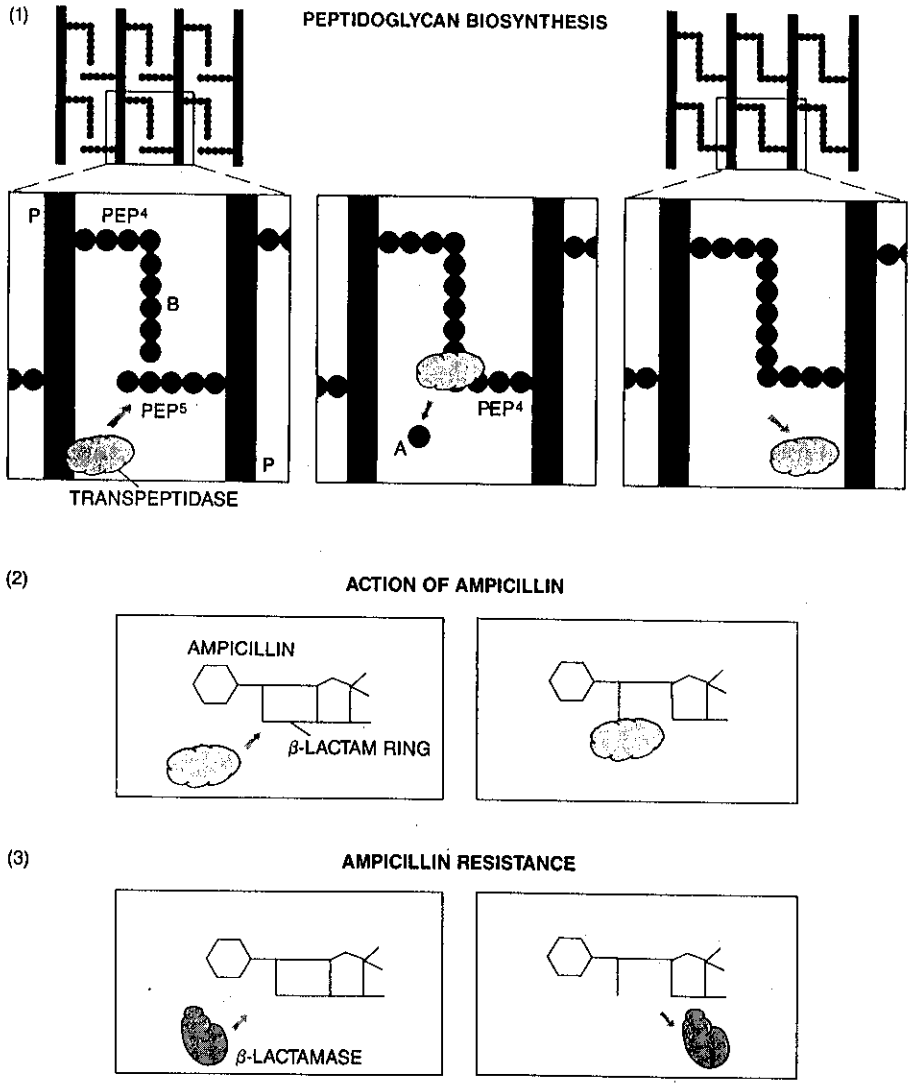
Generally, only those plasmids that confer some selective advantage are maintained in a given bacterial population. A particularly important



E. coli with Plasmids

selective advantage offered by many plasmids is antibiotic resistance genes that code for proteins that disable antibiotics secreted by microorganisms with which bacteria compete. Plasmids, and antibiotic resistance, can be passed from one bacterial strain to another during conjugation.

Antibiotics function by several different mechanisms. For example, members of the penicillin family (including ampicillin) interfere with cell wall biosynthesis. Kanamycin, tetracycline, and chloramphenicol arrest bacterial cell growth by blocking various steps in protein synthesis. Likewise, there are various mechanisms of antibiotic resistance. The



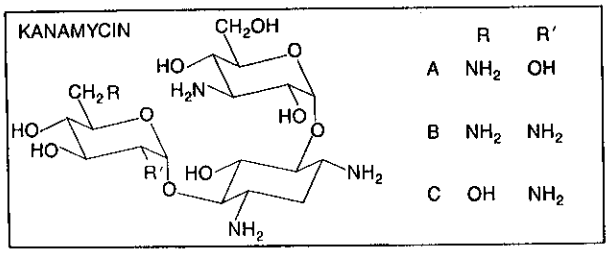
Peptidoglycan Biosynthesis, Ampicillin Action, and Ampicillin Resistance
(*1*) A transpeptidase removes an alanine residue (A) from a pentapeptide (PEP^5). The resulting tetrapeptide (PEP^4) is joined to the peptide bridge (B) to cross-link two adjacent polysaccharide chains (P). (2) The β-lactam ring of ampicillin structurally mimics the peptide bridge and irreversibly binds the transpeptidase, making it unavailable for peptidoglycan synthesis. (*3*) The pAMP resistance protein, β-lactamase, cleaves the β-lactam ring of ampicillin, making it unable to bind the transpeptidase.

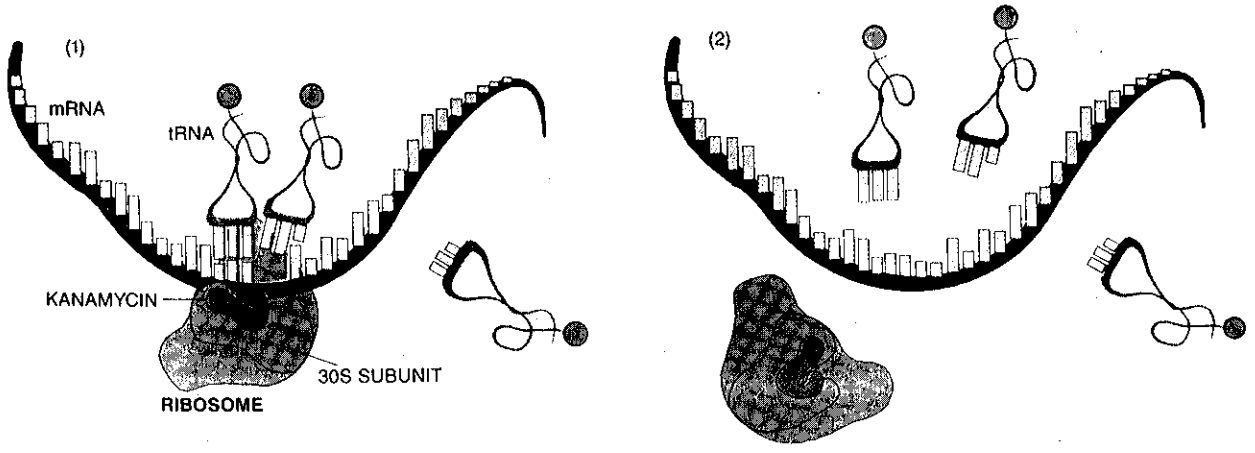
ampicillin, kanamycin, and chloramphenicol resistance genes produce proteins that inactivate their target antibiotics through chemical modification. The tetracycline resistance gene specifies an enzyme that prevents transport of the antibiotic through the cell membrane.

Induced Transformation

The phenomenon of transformation, which provided a key clue to understanding the molecular basis of the gene, also provided a tool for manipulating the genetic makeup of living things. The natural transformation described by Griffith and Avery is an exceedingly rare event. However, in 1970, Mandel and Higa found that *E. coli* becomes markedly *competent* for transformation by foreign DNA when cells are suspended in cold calcium chloride solution and subjected to a brief heat shock at 42°C. They also found that cells arrested in early- to mid-log growth can be rendered more competent than can cells in other stages of growth.

Their calcium chloride procedure, which is still in wide use, yields transformation efficiencies of 10^5 to 10^7 transformants per microgram of plasmid DNA. (Transformation efficiency is generally expressed as the number of transformed cells that would be obtained from a microgram of intact plasmid DNA.) However, there is nothing particularly magical about calcium (Ca^{++}) ions.





Kanamycin Action

Kanamycin poisons protein synthesis by irreversibly binding the 30S subunit of the ribosome. (*1*) The kanamycin/ribosome complex initiates protein synthesis by binding mRNA and the first tRNA. (2) However, the second tRNA is not bound, and the ribosome-mRNA complex dissociates. The pKAN resistance protein, from the aminoglycoside family, is a phosphotransferase that adds phosphate groups to the kanamycin molecule, thus blocking its ability to bind the ribosome. Other resistance proteins from this family can be acetyltransferases, which add acetyl groups to the antibiotic. Three forms of the phosphotransferase (A, B, or C) vary by the group present at R and R′, as indicated in the figure.

Subsequent research showed that treatment with other divalent cations, such as magnesium (Mg^{++}), manganese (Mn^{++}), and barium (Ba^{++}), produces comparable or even greater transformation efficiencies. A protocol using a reducing agent and a complex transformation buffer composed of a mixture of positive ions (Ca^{++}, Mn^{++}, K^{+}, and Co^{+++}) achieves efficiencies of up to 10^9 transformants per microgram of plasmid DNA.

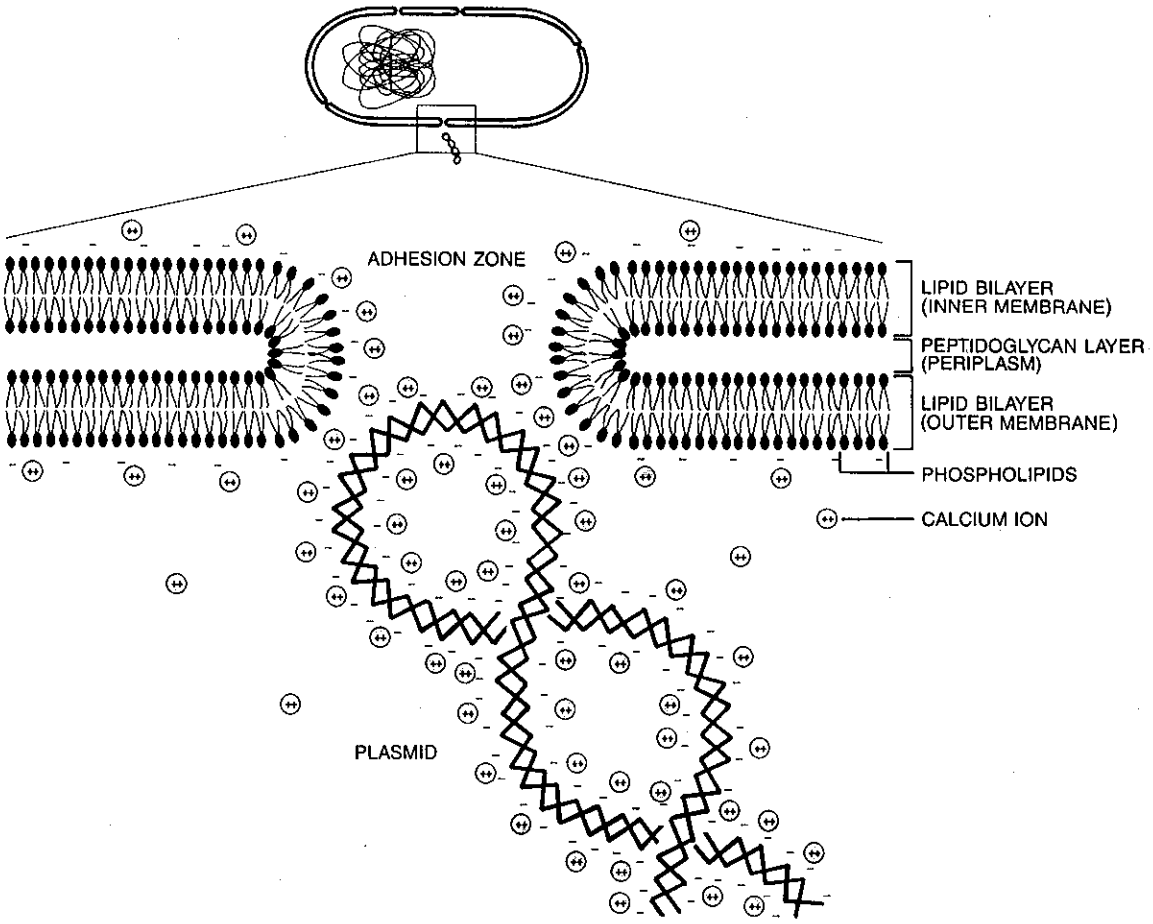
Under conditions that yield high-efficiency transformation (10^9 transformants per microgram of plasmid), approximately 1/10th of all viable cells are rendered competent. At an efficiency of 10^5 transformants per microgram of plasmid, only 1/100,000th of cells become competent. Both size and conformation of the DNA molecule affect transformation efficiency. Small plasmids are more readily taken up than larger ones, although no preferred size cutoff is evident. Linear DNA fragments transform at a negligible rate, in part because they are susceptible to degradation by exonucleases present in the periplasm between the inner and outer cell membrane of *E. coli.*

Because transformation is limited to a subset of cells that are competent, increasing the number of available plasmids does not increase the probability that a cell will be transformed. A suspension of competent cells becomes saturated at a very low concentration of DNA, roughly 0.1–0.2 μg per milliliter of mid-log cells (10^8 cells). Increasing plasmid concentration beyond this point decreases transformation efficiency, because the excess DNA does not transform additional cells. Competent cells will, however, readily take up more than one plasmid under saturating conditions. Experiments have shown that when equal amounts of two different plasmids are added under saturating conditions, 70–90% of all transformed cells contain both plasmids.

Mechanism of DNA Uptake

The mechanism of plasmid DNA uptake by competent *E. coli* cells is unknown. Unlike salts and small organic molecules such as glucose, DNA molecules are too large to diffuse or be readily transported through the cell membrane. Some bacteria possess membrane proteins that recognize DNA and facilitate the absorption of short DNA sequences derived from related species. However, *E. coli* appears not to have evolved such an uptake mechanism.

One hypothesis is that DNA molecules pass through any of several hundred channels formed at *adhesion zones*, where the outer and inner cell membranes are fused to pores in the bacterial cell wall. The fact that adhesion zones are only present in growing cells is consistent with the observation that cells in logarithmic growth can be rendered most competent for plasmid uptake. However, acidic phosphates of the DNA helix are negatively charged, as are a proportion of the phospholipids composing the cell membranes and lining the membrane pore. Thus, electrostatic repulsion between anions may effectively block the movement of DNA through the adhesion zones.

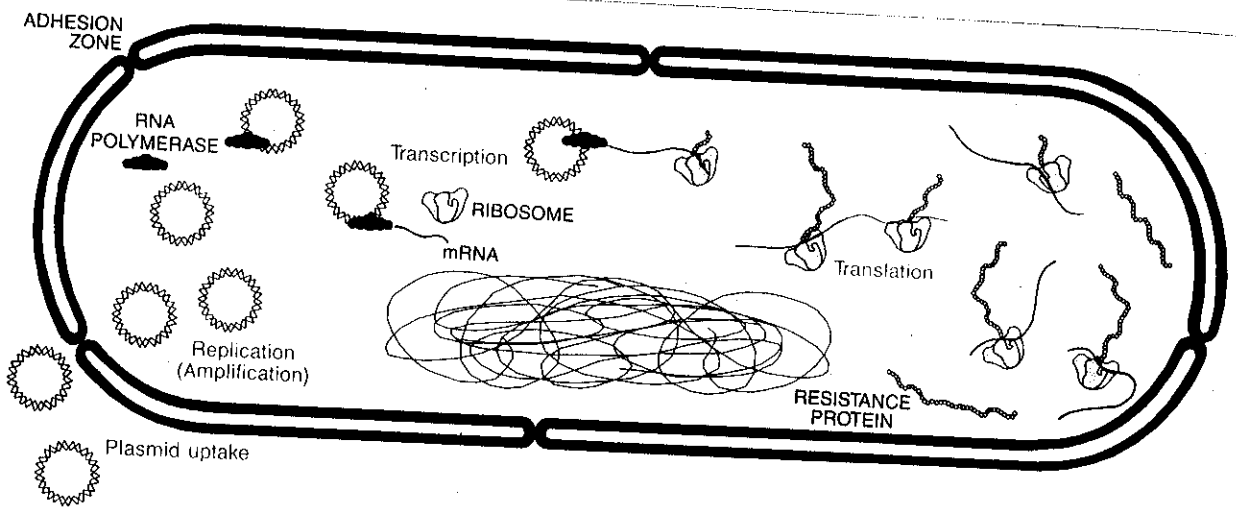


Proposed Molecular Mechanism of DNA Transformation of *E. coli*
Calcium ions (+ +) complex with negatively charged oxygens (−) to shield DNA phosphates from phospholipids at the adhesion zone.

Analysis of this ionic interaction produces a plausible hypothesis for DNA uptake in bacteria. Treatment of the cells at 0°C crystallizes the fluid cell membrane, stabilizing the distribution of charged phosphates. The cations in a transformation solution (Ca^{++}, Mn^{++}, K^{+}, Co^{+++}) form complexes with exposed phosphate groups, shielding the negative charges. With this ionic shield in place, a plasmid molecule can then move through the adhesion zone. Heat shock complements this chemical process, perhaps by creating a thermal imbalance on either side of the *E. coli* membrane that physically helps to pump DNA through the adhesion zone.

Selection of Transformants

Following uptake, the plasmid must become stably established inside the *E. coli* cell. During this period, the plasmid replicates and expresses antibiotic resistance genes that allow selection of a transformed phenotype. Transformed *E. coli* are plated onto culture medium containing one or more antibiotics. Only cells that take up a plasmid and express its



Expression of Antibiotic Resistance by Transformed *E. coli*

antibiotic resistance proteins can thrive on the antibiotic medium; untransformed cells fail to grow. Thus, antibiotic resistance is a selectable marker that allows one to positively identify cells that have been induced to take up plasmid DNA— including recombinant plasmids containing a gene of interest.

This type of culturing strategy, which selects for metabolic phenotypes that cannot be observed directly, was introduced by Beadle and Tatum in 1941. They screened for a *loss of gene function* by supplementing a biochemical mutant with a needed nutrient. Antibiotic selection, on the other hand, screens for a *gain in gene function* by challenging a transformant with an antibiotic.

Isolation and Analysis of Recombinant Plasmids

Growth of colonies on antibiotic medium gives *phenotypic* evidence that the cells have been transformed with a recombinant plasmid. To confirm this at the *genotypic* level, plasmid DNA is isolated from transformants. Restriction analysis of the purified plasmid DNA, along with the parental DNAs used to form the recombinant, provides a "fingerprint" of its genetic identity. Their small size and ring structure make plasmids relatively easy to separate from chromosomal DNA. A rapid method for making a small preparation of purified plasmid DNA, known as a miniprep, is outlined below.

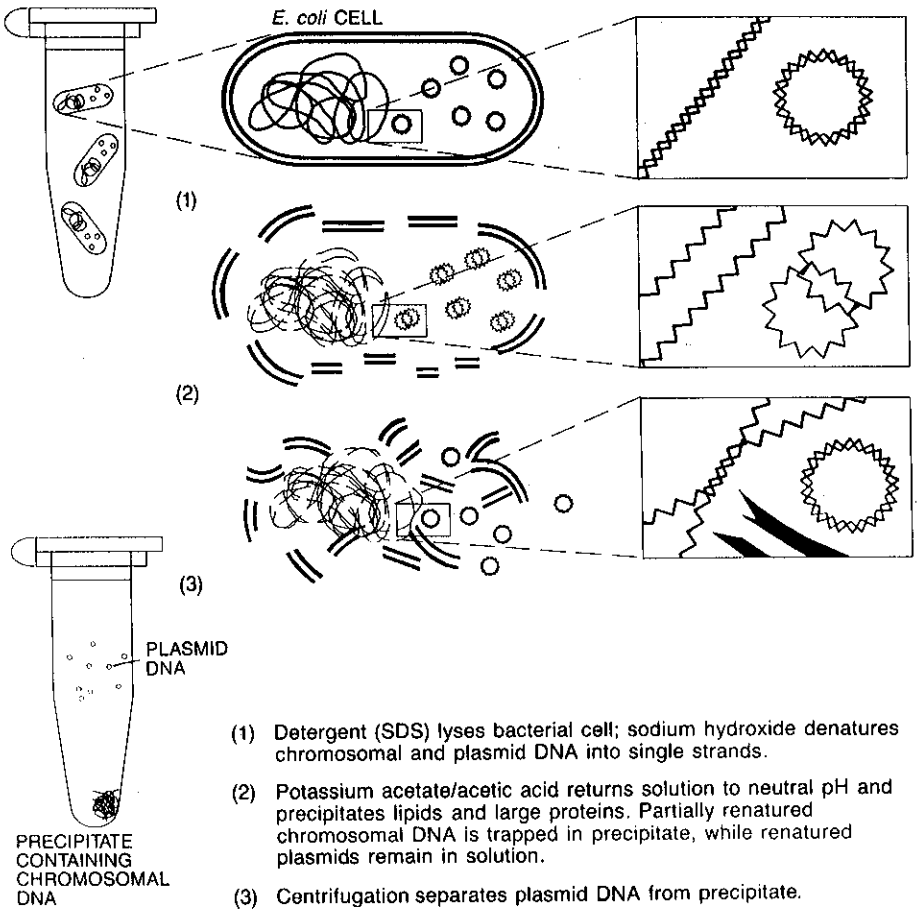
Transformed cells taken from an antibiotic-resistant colony are grown to stationary phase in an overnight suspension culture. The cells are collected by centrifugation and resuspended in a buffered solution of glucose and ethylenediaminetetraacetic acid (EDTA), which binds divalent cations (such as $Mg^{++}$and Ca^{++}) necessary for cell membrane stability.

The resuspended cells are then treated with a mixture of sodium dodecyl sulfate (SDS) and sodium hydroxide. SDS, an ionic detergent,

dissolves the phospholipid and protein components of the cell membrane. This lyses the membrane, releasing the cell contents. Sodium hydroxide denatures both plasmid and chromosomal DNAs into single strands. The chromosomal DNA separates completely into individual strands; however, the single-stranded plasmid loops remain linked together like interlocked rings.

Subsequent treatment with potassium acetate and acetic acid forms an insoluble precipitate of SDS/lipid/protein and neutralizes the sodium hydroxide from the previous step. At neutral pH, the DNAs renature. However, the long strands of chromosomal DNA only partially renature and become trapped in the SDS/lipid/protein precipitate. The plasmid DNA completely renatures into double-stranded molecules that remain in solution and largely escape entrapment in the precipitate.

The precipitate is pelleted by centrifugation and discarded, leaving the plasmid DNA (as well as small RNA molecules) in the supernatant. The supernatant is collected and isopropanol is added to precipitate the plasmid DNA, which is pelleted by centrifugation. The pellet is washed with ethanol and resuspended in a small volume of buffer. Treatment with RNase destroys RNA, leaving relatively clean plasmid DNA.



Plasmid Isolation

Samples of the miniprep plasmid and the parental plasmids are cut with the same restriction enzymes used to make the recombinant constructions. The digested DNAs are run on an agarose gel and the fragment patterns are compared. The bands containing the cloned gene and the vector backbone should line up perfectly with the corresponding fragments in the parental DNA. Each fragment of the same size will have migrated precisely the same distance.

For Further Reading

Cherfas, Jeremy. 1982. *Man-made life*. Pantheon Books, New York.

Sylvester, Edward J. and Lynn C. Klotz. 1987. *The gene age*. Charles Scribner's Sons, New York.

Kornberg, Arthur. 1989. *For the love of enzymes*. Harvard University Press, Cambridge, Massachusetts.

Lewin, Benjamin. 1987. *Genes III*. John Wiley and Sons, New York.

Neidhardt, Frederick C., ed. 1987. Escherichia coli *and* Salmonella typhimurium*: Cellular and molecular biology,* volumes I and II. American Society for Microbiology, Washington, D.C.

Novick, Richard P. 1980. Plasmids. *Sci. Am. 243/6:* 103–127.

Shapiro, James A. 1988. Bacteria as multicellular organisms. *Sci. Am. 258/6:* 82–89.

Watson, James D., John Tooze, and David T. Kurtz. 1983. *Recombinant DNA: A short course*. Scientific American Books, New York.