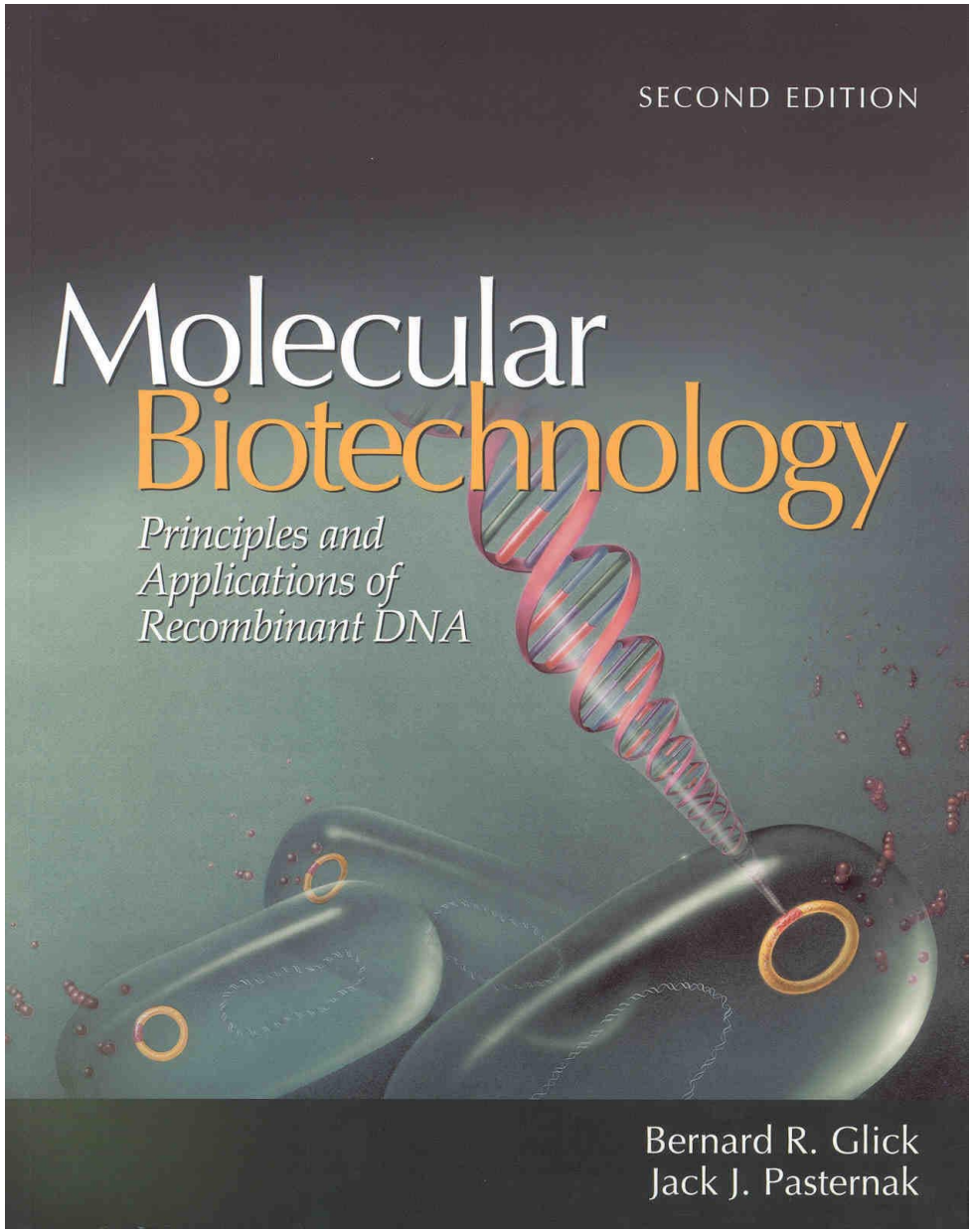
Molecular Biotechnology

Genetic-Engineering of Microorganisms

David R. Shonnard

CM4710 Biochemical Processes

November 28, 2003


SECOND EDITION
Molecular Biotechnology
Principles and Applications of Recombinant DNA
Bernard R. Glick
Jack J. Pasternak

Recombinant DNA Cloning Procedure

1. Identify a cloning vector (Plasmid)
2. Identify a target DNA
3. Open up plasmid and target DNA using Restriction Endonuclease
4. Join target DNA with cloning vector
5. Introduce recombinant DNA construct into host cell
6. Isolate cells with cloned DNA

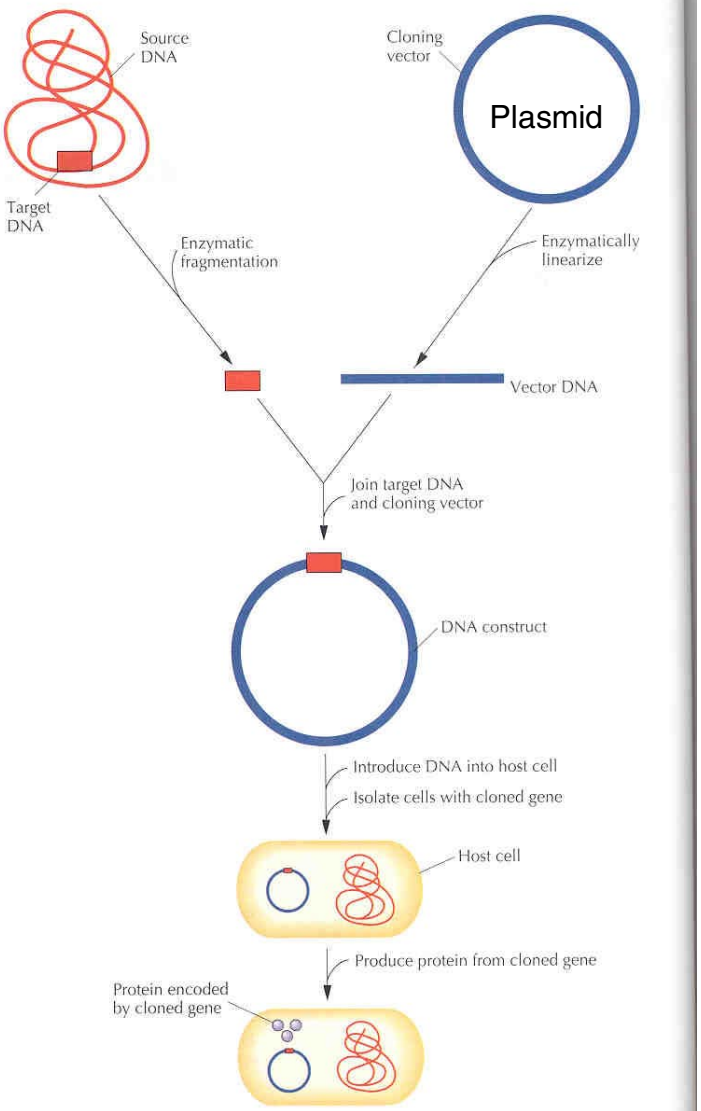
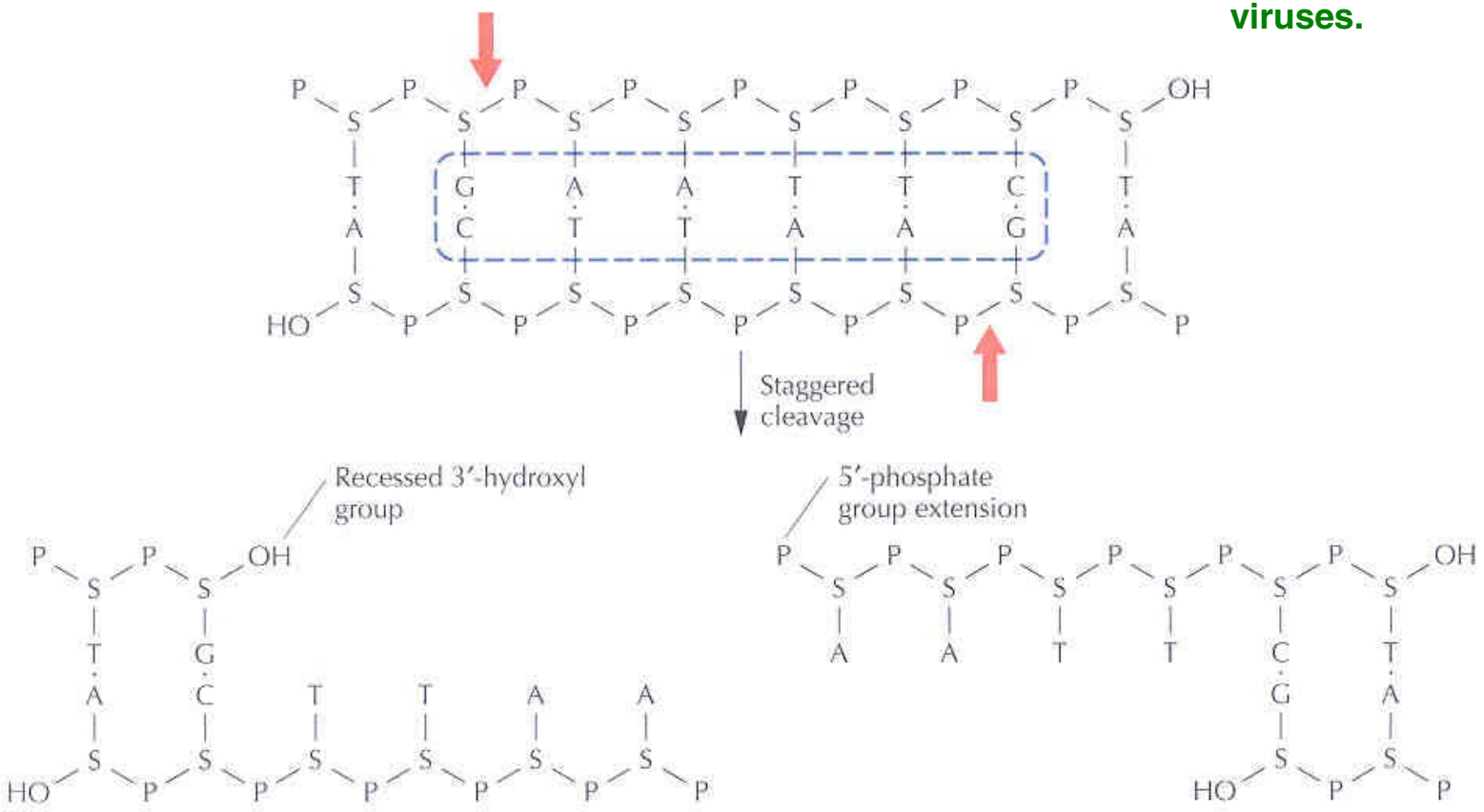


Figure 4.1 Recombinant DNA cloning procedure. DNA from a source organism is cleaved with a restriction endonuclease and inserted into a cloning vector. Then the cloning vector–insert DNA construct is introduced into a target host cell, and those cells that carry the construct are identified and grown. If required, the cloned gene can be expressed in the host cell and its protein can be produced and harvested.

Restriction Endonuclease Enzyme - Staggered Cut (sticky-ended)

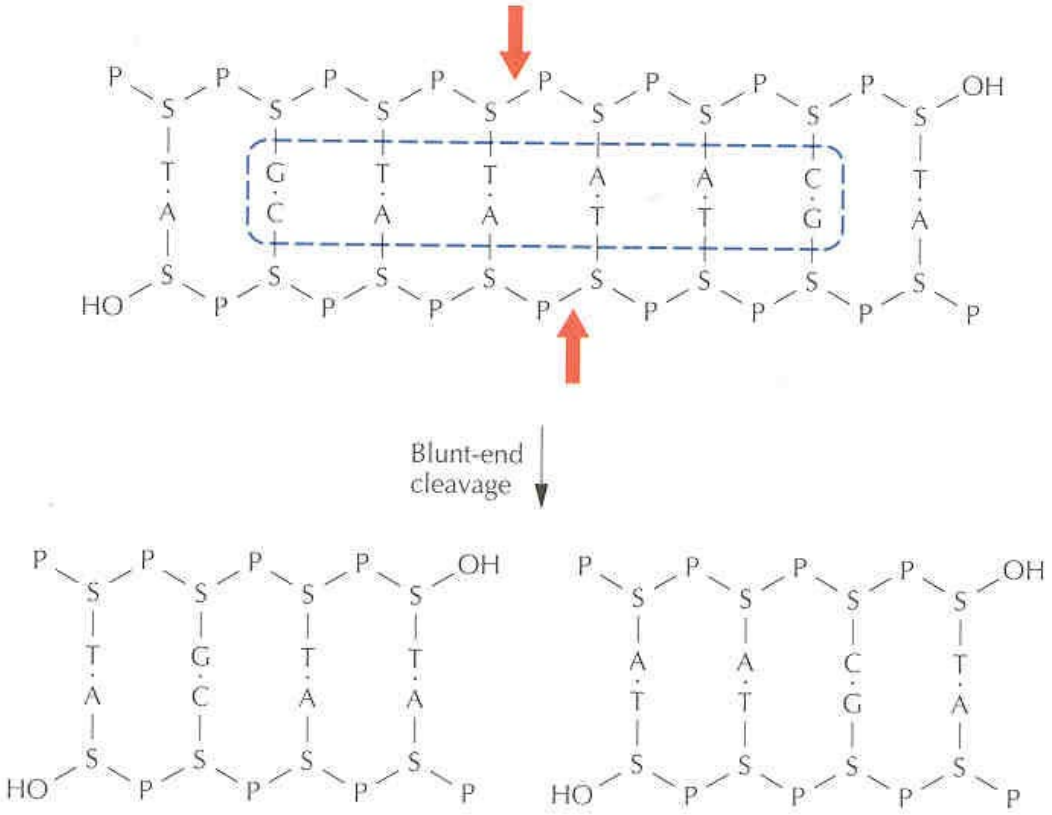
Figure 4.2 Symmetrical, staggered cleavage of a short fragment of DNA by the type II restriction endonuclease *Eco*RI. The bold arrows show the sites of cleavage in the DNA backbone. S, deoxyribose sugar; P, phosphate group; OH, hydroxyl group; A, adenine; T, thymine; C, cytosine; G, guanine. The *Eco*RI recognition sequence is highlighted by the dashed line.

Restriction Endonuclease is present in bacteria to help combat invading DNA from viruses.



Restriction Endonuclease Enzyme- Blunt-end Cut

Figure 4.3 Blunt-end cleavage of a short fragment of DNA by the type II restriction endonuclease *Hin*dII. The bold arrows show the sites of cleavage in the DNA backbone. For abbreviations, see the legend to Fig. 4.2. The *Hin*dII recognition sequence is highlighted.



Restriction Endonuclease Enzymes

Isolated from different bacteria

E. coli

Table 4.1 Recognition sequences of some restriction endonucleases

Enzyme	Recognition site	Type of cut end
*Eco*RI	G↓A—A—T—T—C C—T—T—A—A↑G	5′-phosphate extension
*Bam*HI	G↓G—A—T—C—C C—C—T—A—G↑G	5′-phosphate extension
*Pst*I	C—T—G—C—A↓G G↑A—C—G—T—C	3′-hydroxyl extension
*Sau*3AI	↓G—A—T—C C—T—A—G↑	5′-phosphate extension
*Pvu*II	C—A—G↓C—T—G G—T—C↑G—A—C	Blunt end
*Hpa*I	G—T—T↓A—A—C C—A—A↑T—T—G	Blunt end
*Hae*III	G—G↓C—C C—C↑G—G	Blunt end
*Not*I	G↓C—G—G—C—C—G—C C—G—C—C—G—G—C↑G	5′-phosphate extension

Cleavage site

Plasmids

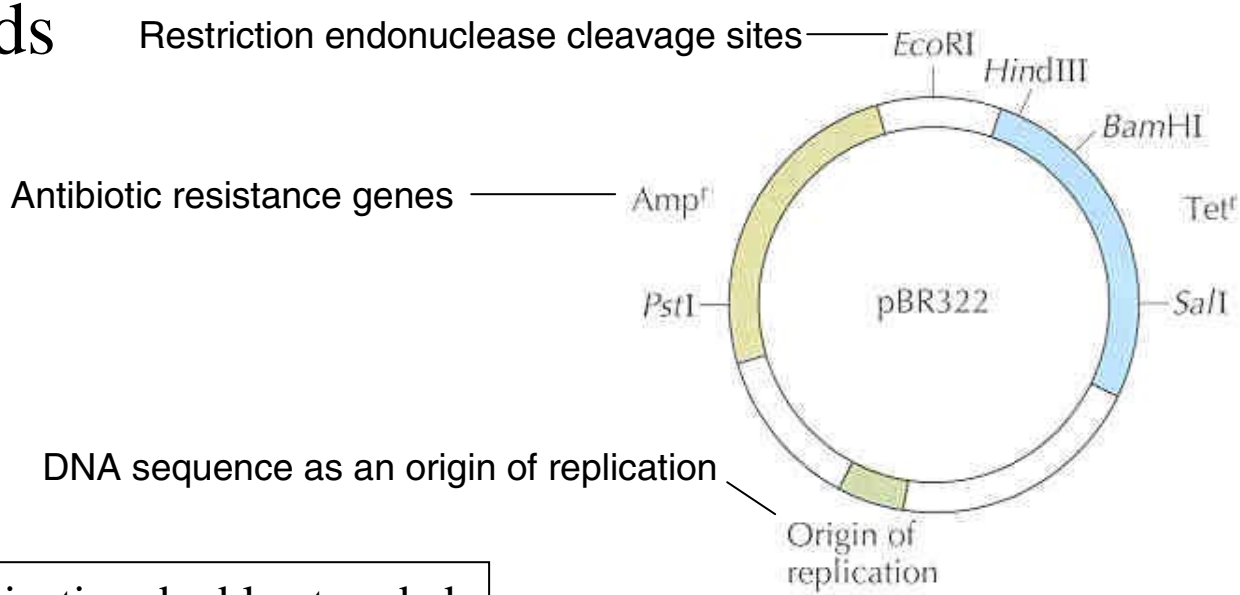


Figure 4.7 Genetic map of the plasmid cloning vector pBR322. Unique *Hin*dIII, *Sal*I, *Bam*HI, and *Pst*I recognition sites are present the genes for tetracycline resistance (Tet^r) and ampicillin resistance (Amp^r). The unique *Eco*RI site is just outside the tetracycline resistance gene. The origin of replication functions in the bacterium *E. coli*. The complete DNA sequence of pBR322 consists of 4,361 bp.

1. Self-replicating double-stranded circular strands of DNA
2. Not part of cell chromosome
3. F plasmid - transfer information from cell to cell
4. R plasmid - antibiotic resistance
5. Degradative plasmid - utilization of unusual metabolites
6. 1 - 500 kilobasepairs long

Creating a Gene Library - Target DNA

Make a Gene Library -

a. Cut DNA from a source microorganism with a 4-cutter restriction endonuclease (many pieces of DNA are created, some large and some small)

b. All resulting fragments of DNA are inserted into the cloning vector (plasmid) and the plasmid is introduced into host cells, yielding many recombinant (transformed) cells - called “transformation”

c. Identify the proper cell line (clone) from the many cell lines created. This is a process called “screening”.

Creating a Gene Library - Partial Digestion of Target DNA with one or more Restriction Endonucleases

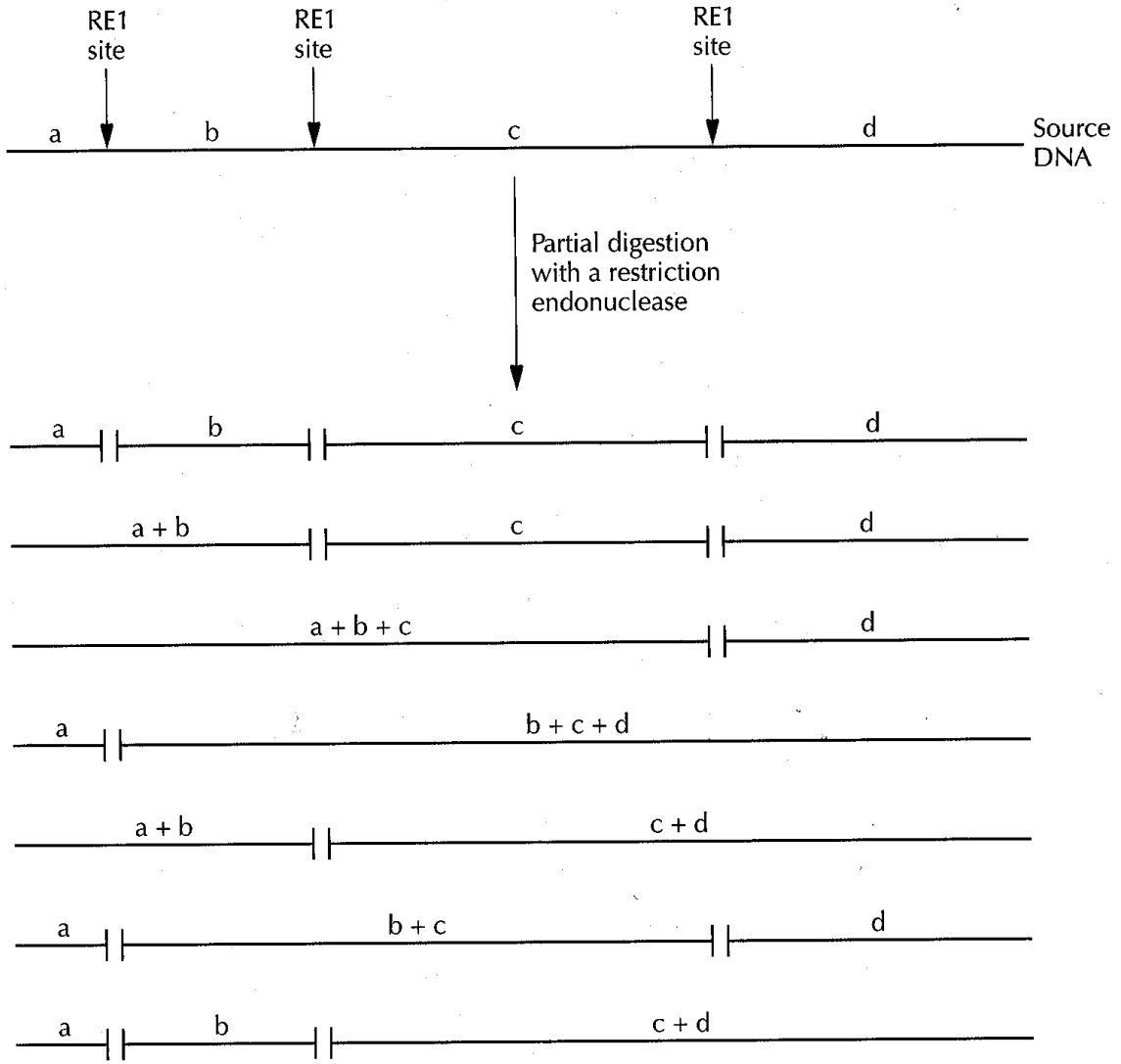
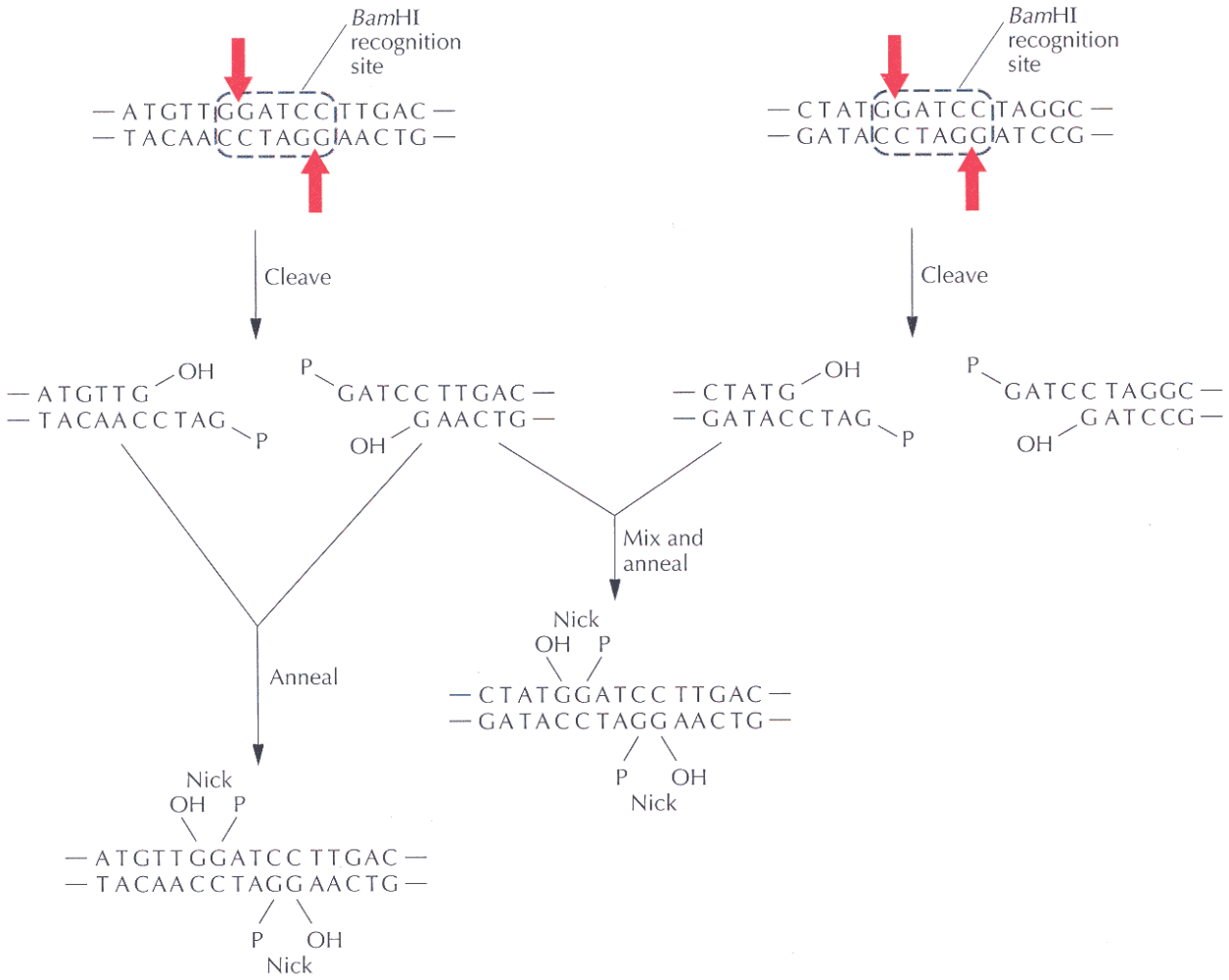


Figure 4.10 Partial digestion of a fragment of DNA with a type II restriction endonuclease. Partial digestions are usually performed by varying either the length of time or the amount of enzyme used for the digestion. In some of the DNA molecules, the restriction endonuclease has cut at all sites (each labeled RE1). In other molecules, fewer cleavages have occurred. The desired outcome is a sample with DNA molecules of all possible lengths.

Creating a Gene Library - Cutting Target DNA

Figure 4.5 Annealing complementary extensions after staggered cleavage with a type II restriction endonuclease. Two different DNA fragments are cut with the restriction endonuclease *Bam*HI, mixed, and annealed. Not all of the possible combinations of annealed DNAs are shown. The four fragments that are generated by the *Bam*HI digestion can anneal to one another to form any of six different DNA molecules. A break in the phosphodiester bond in one strand of duplex DNA is called a nick. The hydrogen bonds of the four base pairs between nicks on opposite strands are not sufficiently strong enough to hold DNA molecules together for long periods in solution. A, C, G, and T represent nucleotides.



Creating a Gene Library - Ligation of Cut DNA

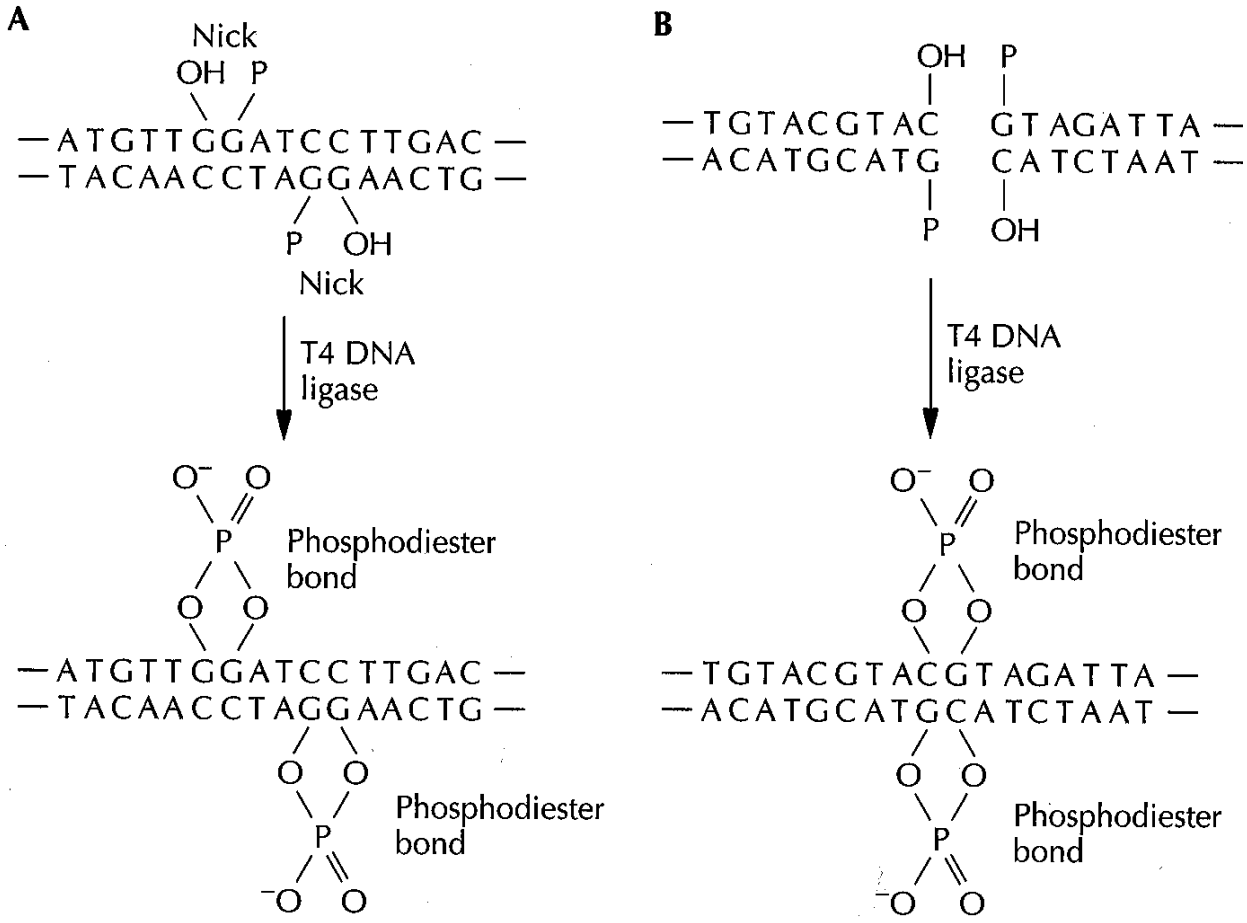


Figure 4.6 Mode of action of T4 DNA ligase. The enzyme T4 DNA ligase forms phosphodiester bonds by joining 5′-phosphate and 3′-hydroxyl groups at nicks in the backbone of double-stranded DNA. **A.** Ligation of sticky-ended DNA. **B.** Ligation of blunt-ended DNA. A, C, G, and T represent nucleotides.

Creating a Gene Library - Cleaving of DNA and Ligation of Cut DNA

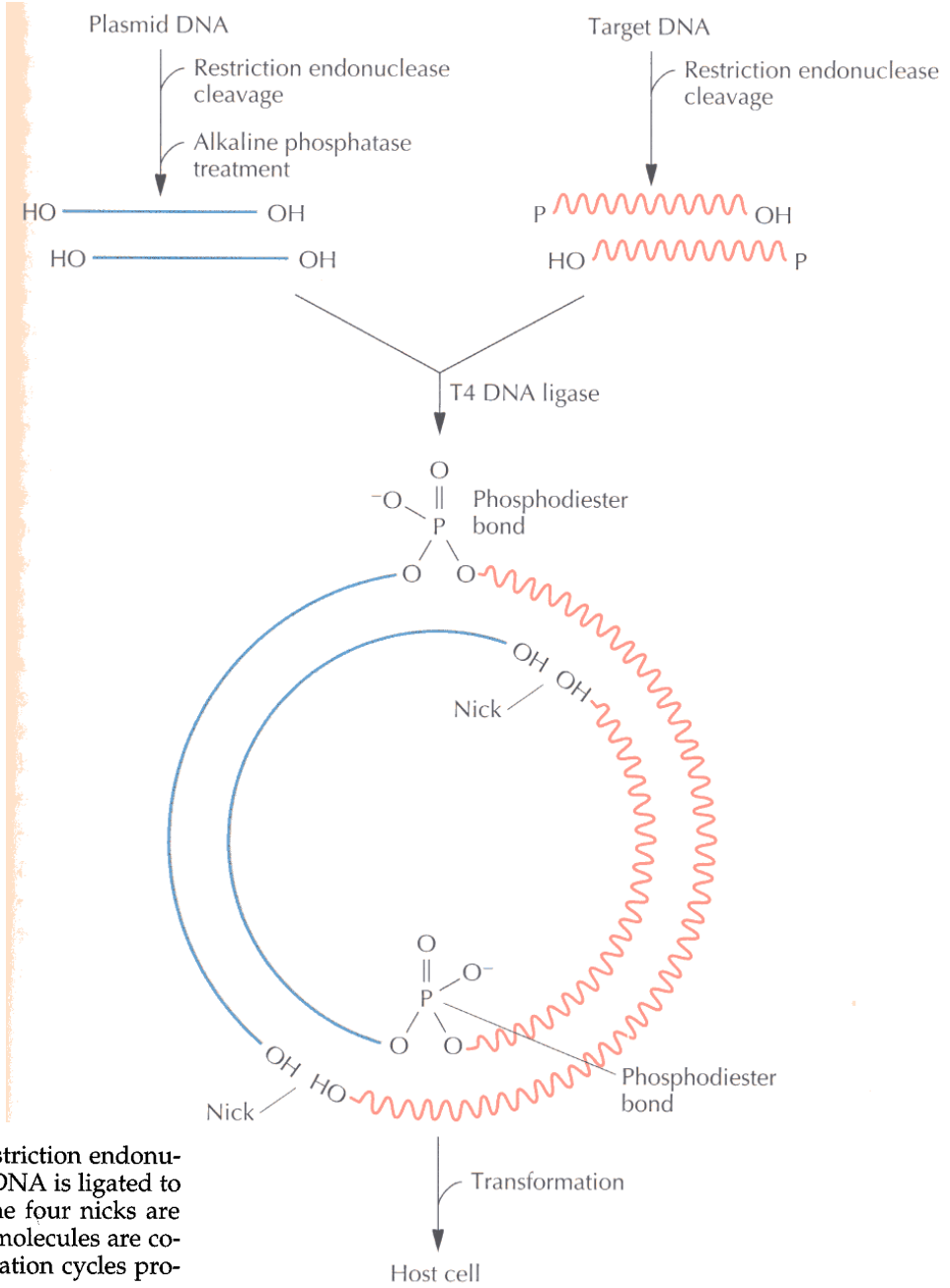


Figure 4.8 Cloning foreign DNA into a plasmid vector. After restriction endonuclease cleavage and alkaline phosphatase treatment, the plasmid DNA is ligated to the restriction endonuclease-digested target DNA, and two of the four nicks are sealed. This molecular configuration is stable, and the two DNA molecules are covalently joined. After introduction into a host cell, ensuing replication cycles produce new complete circular DNA molecules with no nicks.

Creating a Gene Library - Transformation

Chemical Treatment -

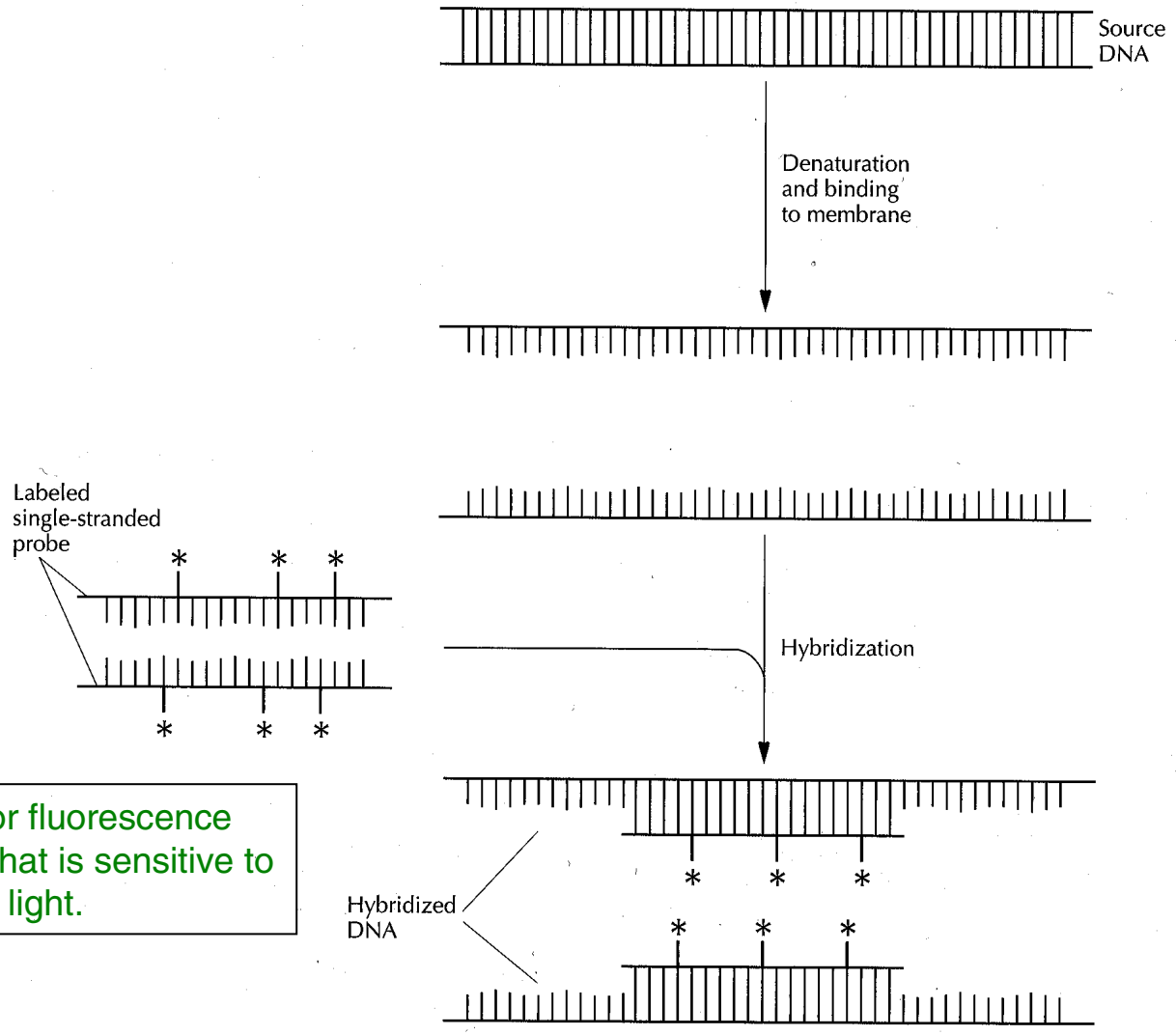
a. Treating cells (*E. coli* often) plus recombinant plasmid DNA with ice-cold $CaCl_2$ and then exposing the culture to higher temperature (42 ˚C) for 90 s. About 1 in very 1,000 cells is transformed. Transformation frequency (about 10^7 - 10^8 transformants per μg plasmid DNA) decreases with increase in plasmid size (small plasmids - < 10 kb).

Electroporation - "electric field-mediated membrane permeabilization"

a. Treating cells (50 μL) plus recombinant plasmid DNA with a single pulse of 25 microfarads, 2.5 kilovolts, and 200 Ohms in 4.6 μs. Transformation efficiency is about 10^9 transformants per microgram of DNA for small plasmids (3 kb) and 10^6 for large plasmids (136 kb).

Creating a Gene Library - Screening of a Library of Recombinant Cells using Hybridization to DNA

Figure 4.11 DNA hybridization. The source DNA is denatured, and the two strands are kept apart, usually by binding them to a solid matrix such as a nitrocellulose or nylon membrane. Labeled probe DNA (often approximately 100 to 1,000 bp) is denatured and added to the denatured source DNA. Hybridization (base pairing) between the probe and source DNA may occur under these conditions. The membrane is then washed to remove unhybridized probe DNA, and the membranes are assayed. If the probe hybridizes with the source DNA, then it can be detected with an assay that identifies its labeled tag. If the probe does not hybridize, then no label is detected. The asterisks (*) denote the labeled tag (signal) of the probe DNA.



Detection of radiation or fluorescence is achieved using film that is sensitive to radioactive decay or to light.

Isolation of desired transformed cell

Target DNA inserted in the *Bam*HI site

1. Transformed cells are resistant to the antibiotic Ampicillin but not to Tetracycline
2. Untransformed cells are not resistant to either
3. Plate cells from transformation mixture onto a medium containing Ampicillin
4. Plate cells surviving step 3 onto a medium containing Tetracycline
5. Cells surviving 3 but not 4 are the transformed cells
6. Further screen surviving cells to find the proper metabolic function

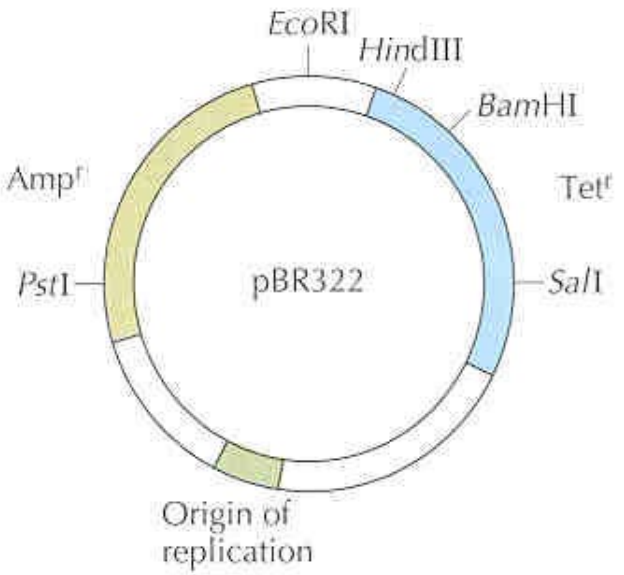


Figure 4.7 Genetic map of the plasmid cloning vector pBR322. Unique *Hin*dIII, *Sal*I, *Bam*HI, and *Pst*I recognition sites are present the genes for tetracycline resistance (Tet^r) and ampicillin resistance (Amp^r). The unique *Eco*RI site is just outside the tetracycline resistance gene. The origin of replication functions in the bacterium *E. coli*. The complete DNA sequence of pBR322 consists of 4,361 bp.

Screening of a Library of Recombinant Cells using Hybridization to DNA

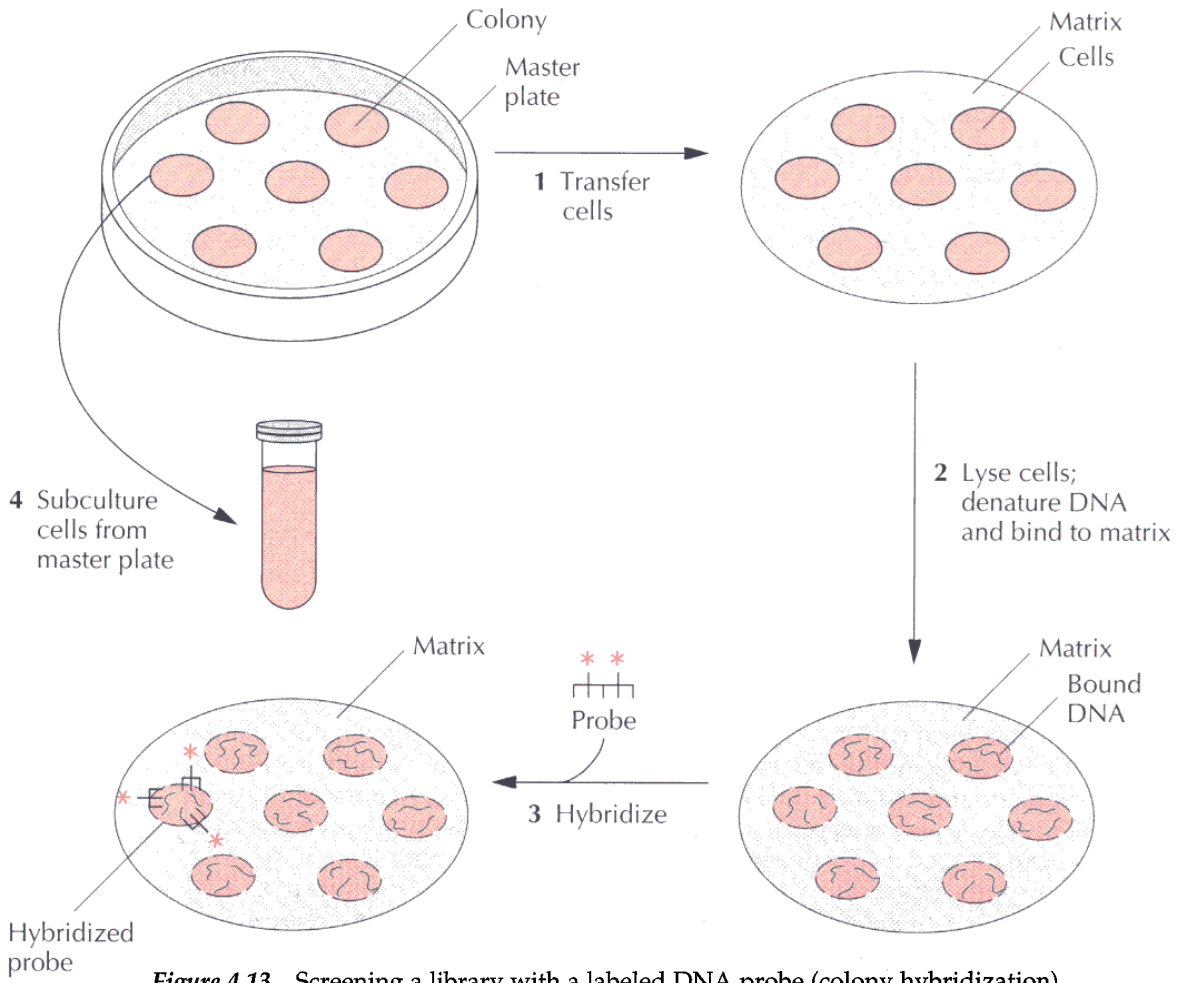


Figure 4.13 Screening a library with a labeled DNA probe (colony hybridization). Cells from the transformation reaction are plated onto solid medium under conditions that permit transformed but not nontransformed cells to grow. (1) From each discrete colony formed on this master plate, a sample is transferred to a solid matrix such as a nitrocellulose or nylon membrane. The pattern of the colonies on the master plate is retained on the matrix. (2) The cells on the matrix are lysed, and the released DNA is denatured, deproteinized, and irreversibly bound to the matrix. (3) A labeled DNA probe is added to the matrix under hybridization conditions. After the nonhybridized probe molecules are washed away, the matrix is processed by autoradiography to determine which cells have bound labeled DNA. (4) A colony on the master plate that corresponds to the region of a positive response on the X-ray film is identified. Cells from the positive colony on the master plate are subcultured because they may carry the desired plasmid–cloned DNA construct.

Screening of a Library of Recombinant Cells using Hybridization to Protein Product with Antibodies

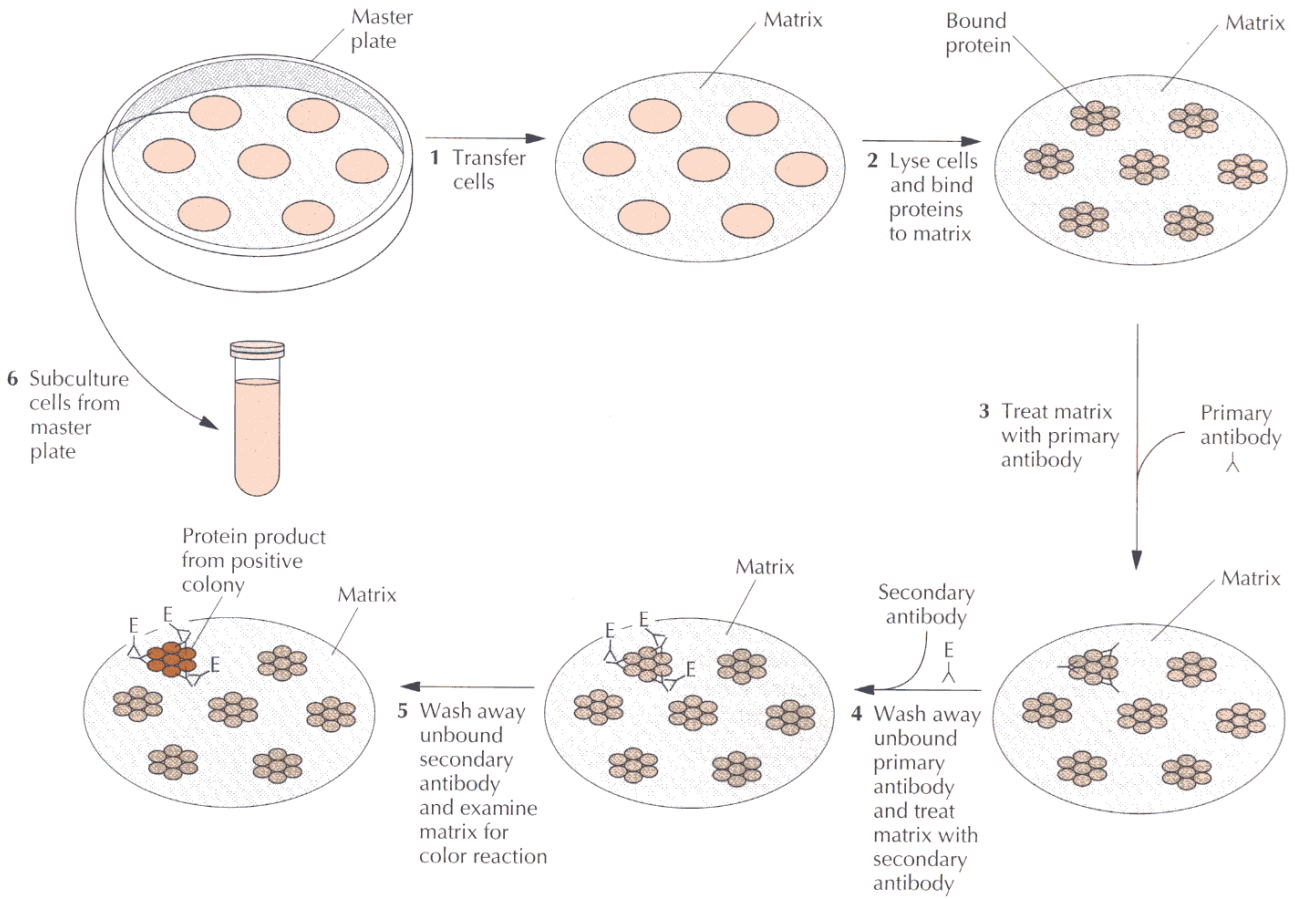


Figure 4.14 Immunological screening of a gene library (colony immunoassay). Cells from the transformation reaction are plated onto solid medium under conditions that permit transformed but not nontransformed cells to grow. (1) From the discrete colonies formed on this master plate, a sample from each colony is transferred to a solid matrix such as a nitrocellulose or nylon membrane. (2) The cells on the matrix are lysed, and their proteins are bound to the matrix. (3) The matrix is treated with a primary antibody that binds only to the targeted protein. (4) Unbound primary antibody is washed away, and the matrix is treated with a secondary antibody that binds only to the primary antibody. (5) Any unbound secondary antibody is washed away, and a colorimetric reaction, which can occur only if the secondary antibody is present, is carried out. (6) A colony on the master plate that corresponds to a positive response on the matrix is identified. Cells from the positive colony on the master plate are subcultured because they may carry the plasmid–insert DNA construct that encodes the protein that binds the primary antibody.

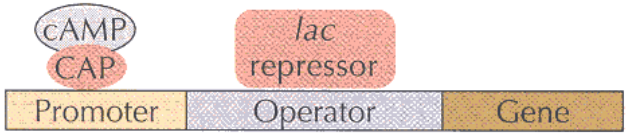
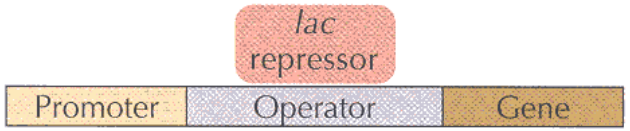
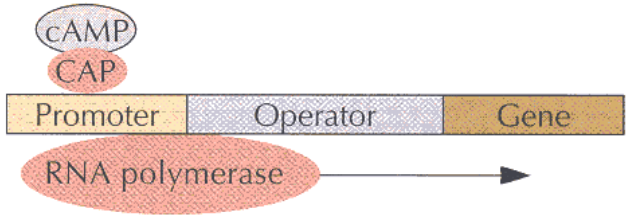
Manipulation of Gene Expression in Prokaryotes

1. Strong and Regulatable Promoters on Plasmid
2. High Copy Number of Plasmids in Host Cell
3. Increasing Rate of Translation – Ribosome Binding
4. Increasing Protein Stability – Fusion Proteins
5. Increasing Protein Stability – N-Terminal Amino Acid
6. Increasing Protein Stability – Overcome O_2 Limitations
7. Increasing Protein Secretion from Host Cell
8. Effects of Plasmid Copy Number on Cell Growth Rate

Regulatable Promoter – *E. coli* Lac Promoter

or

IPTG, isopropyl-β-D-thiogalactopyranoside, a gratuitous inducer molecule.

Glucose concentration	Lactose concentration	cAMP concentration	*lac* promoter-operator region	Level of transcription
Low	Low	High		Low
High	Low	Low		Low
High	High	Low		Low
Low	High	High		High

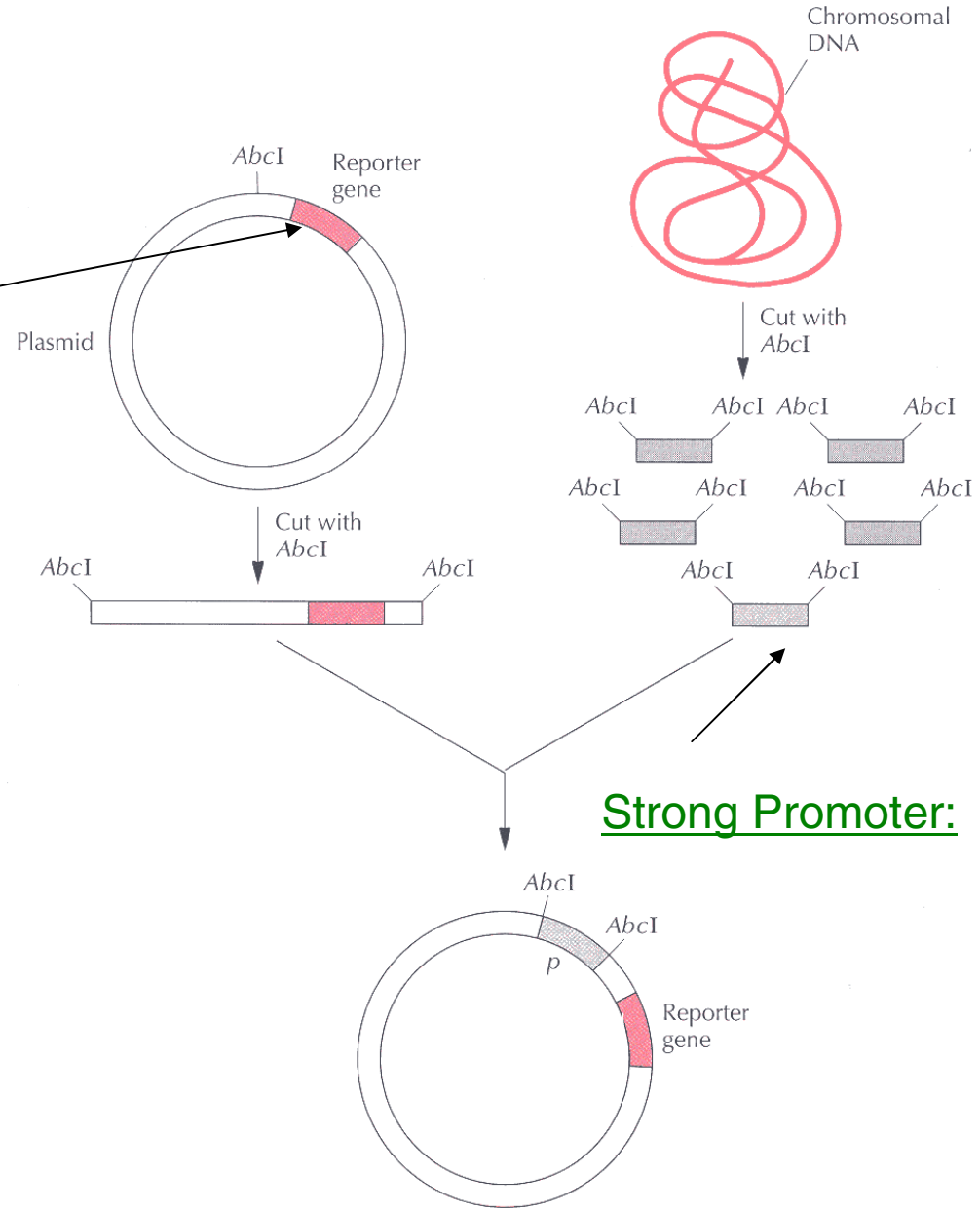
Although the Lac Promoter is regulatable, it is not strong, requiring very specific conditions for Transcription. A strong promoter is needed, replacing the Lac Promoter shown here with one that always has a high affinity for RNA polymerase.

Strong Promoter – Modified *E. coli* Lac Promoter

lacUV5, a variant of the *lac* promoter, contains an altered nucleotide sequence in the -10 region, and is a strong promoter. The -10 region is the region 10 nucleotide pairs upstream from the site of transcription. When the plasmid contains the *lacUV5* promoter, induction of transcription by addition of IPTG or Lactose will cause transcription to occur (m-RNA to be synthesized) without the need for a very low Glucose concentration.

Strong Promoter – Isolating Plasmids

<u>Reporter Gene:</u>
A gene coding for a protein that gives a color change during an assay. A positive assay result proves that the promoter region was cloned in



A Temperature Sensitive Strong Promoter – the p^L Promoter

p^L is a promoter from the bacteriophage, λ. The promoter is controlled by the *cI* repressor protein of bacteriophage, λ, although a temperature -sensitive mutant of *cI*, cI_{857}, is used and is normally integrated into the host cell's chromosome (not into the plasmid). Cells carrying cI_{857} are first grown at 28-30°C, a temperature at which the repressor prevents transcription directed by the p^L promoter. After the cell culture reaches the desired growth stage, the temperature is increased to 42°C, where the repressor is inactive, and transcription is started.

High Copy Number Plasmids

High copy number plasmids require a strong Origin of Replication, a sequence of nucleotides that has a very high affinity for DNA Polymerase. This will assure that the plasmid will be replicated many times during each cell division.

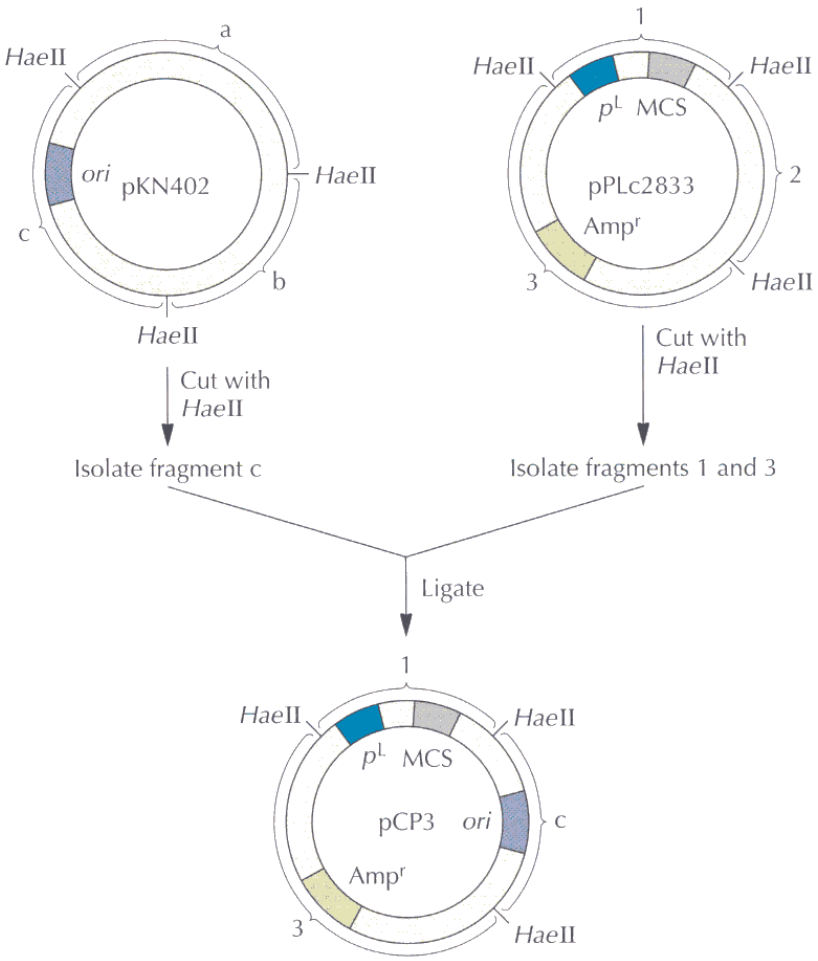


Table 6.1 Effect of temperature on the plasmid copy number of three expression vectors

Plasmid	Plasmids/cell at:		p^L present
	28°C	42°C	
pKN402	82	521	No
pPLc2833	38	42	Yes
pCP3	60	713	Yes

Adapted from Remaut et al., 1983, *Gene* **22**:103–113.

Table 6.6 Effect of plasmid copy number on host cell growth rate

E. coli HB101 with plasmid:	Plasmid copy number	Relative specific growth rate
None	0	1.00
A	12	0.92
B	24	0.91
C	60	0.87
D	122	0.82
E	408	0.77

Adapted from Seo and Bailey, 1985, *Biotechnol. Bioeng.* **27**:1668–1674.

The different plasmids, designated A, B, C, D and E, encode only β-lactamase and are all the same size. The growth rates were normalized to the growth rate value for *E. coli* HB101 without a plasmid.

Increasing Translation – Ribosome Binding

Other factors, such as efficiency of translation, may also affect the amount of protein product from a recombinant cell. Not all procaryote m-RNA are translated with the same efficiency. Increased strength of the Ribosome binding site, a sequence of 6 – 8 nucleotides (e.g., UAAGGAGG), corresponds to a higher efficiency of translation. This sequence of RNA nucleotides is complimentary to the sequence on the RNA component of the small ribosomal subunit (e.g., AUUCCUCC).

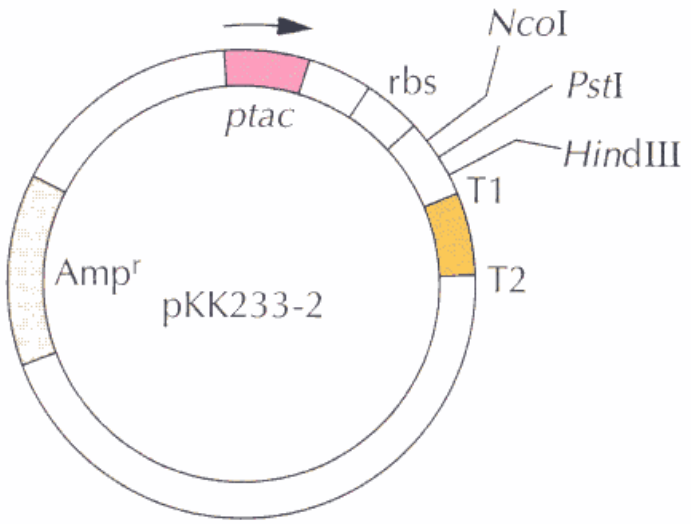


Figure 6.14 The expression vector pKK233-2. The plasmid pKK233-2 codes for the ampicillin resistance gene (Amp^r) as a selectable marker gene, the *tac* promoter (*ptac*), the *lacZ* ribosome binding site (rbs), three restriction endonuclease cloning sites (*Nco*I, *Pst*I, and *Hin*dIII), and two transcription termination sequences (T1 and T2). The arrow indicates the direction of transcription. The plasmid is not drawn to scale.

Increasing Protein Stability – Fusion Proteins

Foreign proteins expressed in host cells are often degraded by proteolytic enzymes. A way to overcome this is to link a native protein or part of a native protein to the recombinant protein, creating a fusion protein. The cell is tricked into thinking that the foreign protein is native to the cell. This is accomplished at the DNA level by ligating together the coding regions of the two genes. After purifying the fusion protein from the culture, the native protein must be removed.

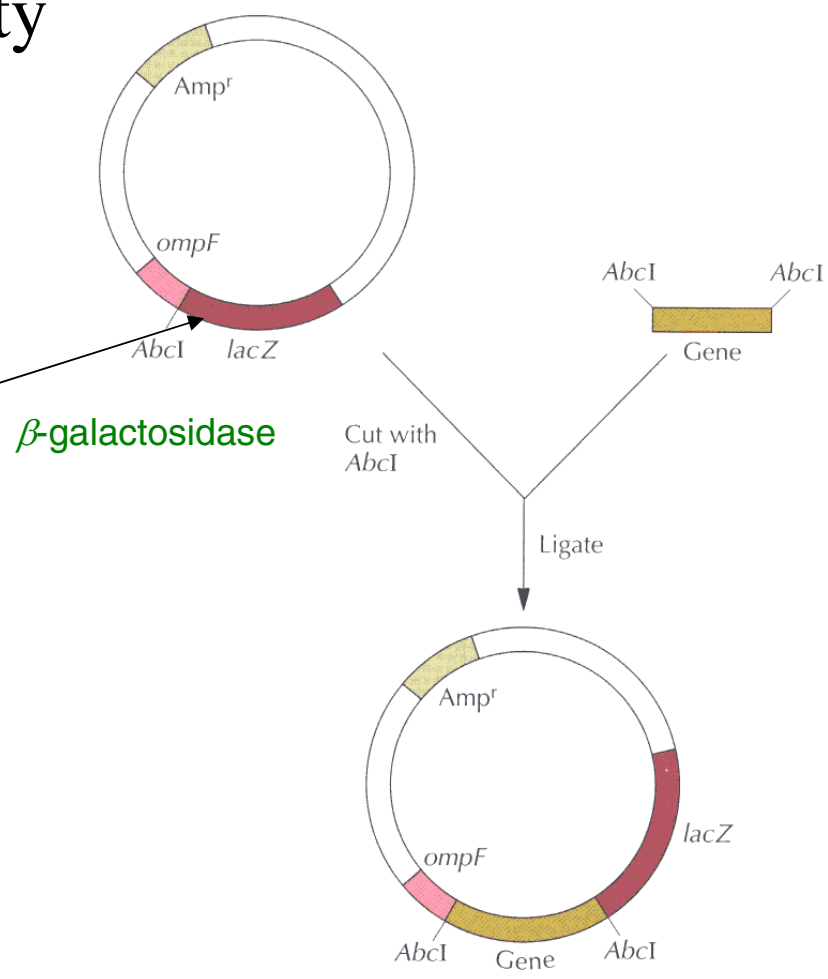


Figure 6.7 Fusion protein cloning vector. The plasmid contains an ampicillin resistance (Amp^r) gene as the selectable marker, a DNA sequence encoding the N-terminal segment of the *E. coli* outer membrane protein (*ompF*), a restriction endonuclease site (*Abc*I) for cloning, and a truncated β-galactosidase gene (*lacZ*). The cloned gene (Gene) is inserted into the *Abc*I site. After transcription and translation, a tribrid protein is produced.

Increasing Protein Stability – N-Terminal Amino Acid

The amino acid located at the N-terminal position has a large effect on the half-life of different proteins within the cell. The table shows the effects of different amino acids at the N-terminal position on the survival time of β-galactosidase.

Table 6.4 Stability of β-galactosidase with certain amino acids added to its N terminus

Amino acid added	**Half-life**
Met, Ser, Ala	>20 h
Thr, Val, Gly	>20 h
Ile, Glu	>30 min
Tyr, Gln	~10 min
Pro	~7 min
Phe, Leu, Asp, Lys	~3 min
Arg	~2 min

Adapted from Bachmair et al., 1986, *Science* **234:** 179–186.

Increasing Protein Stability – Overcome O_2 Limitations

During culture of recombinant cells, oxygen in the medium will become limited during high growth rates at high cell concentration. When oxygen is depleted and cells enter stationary phase, proteases are produced by the cell which may degrade the target protein. A possible solution is to use pure oxygen rather than air, use high aeration rates, and high agitation. Another approach that has been studied is to use smaller and smaller bubble sizes, thus for the same aeration rate, much more surface area between the bubbles and the liquid is created. Some attempts have been made to create mutant cells deficient in protease activity. But 25 proteases are commonly found in cells and eliminating these proteases will hamper the cells normal ability to degrade faulty proteins and may decrease the growth rate of the cells. Another approach is to clone in a gene from a bacterium called *Vitreoscilla*. This gene codes for a hemoglobin-like molecule that binds oxygen and increases the level of oxygen in the cell, even under nearly anaerobic conditions. This strategy has met with some success when this gene is cloned and expressed in *E. coli*.

Increasing Protein Secretion from Host Cell

The stability of protein is dependent on its location in the cell. For example, proteins in the periplasmic space in gram negative bacteria are 10 times more stable than in the cytoplasm. Proteins in the pariplasmic space or even better secreted outside the cell to the medium are much easier to recover. An amino acid sequence at the N-terminus of the protein, termed the signal sequence, facilitates the export of the protein by allowing the protein to pass through the cell membrane. Including the nucleotides for a strong signal sequence in the plasmid next to the target gene will allow for secretion to the pariplasmic space, though secretion through the outer cell wall does not occur. One strategy to allow secretion all the way to the culture medium is to use gram-positive cells or eukaryotic cells, both which lack an outer membrane. Another approach is to genetically engineer a gram negative cell, like *E. coli* to secrete through the outer wall. Some gram-negative cells are able to secrete a bacteriocidal protein called bacteriocin through the cell wall. By putting the bacteriocin release protein gene onto a plasmid under the control of a strong regulatable promoter, *E. coli* cells can be permeabilized at will.

Current chemical/biochemical synthesis of L-Ascorbic Acid (Vitamin C)

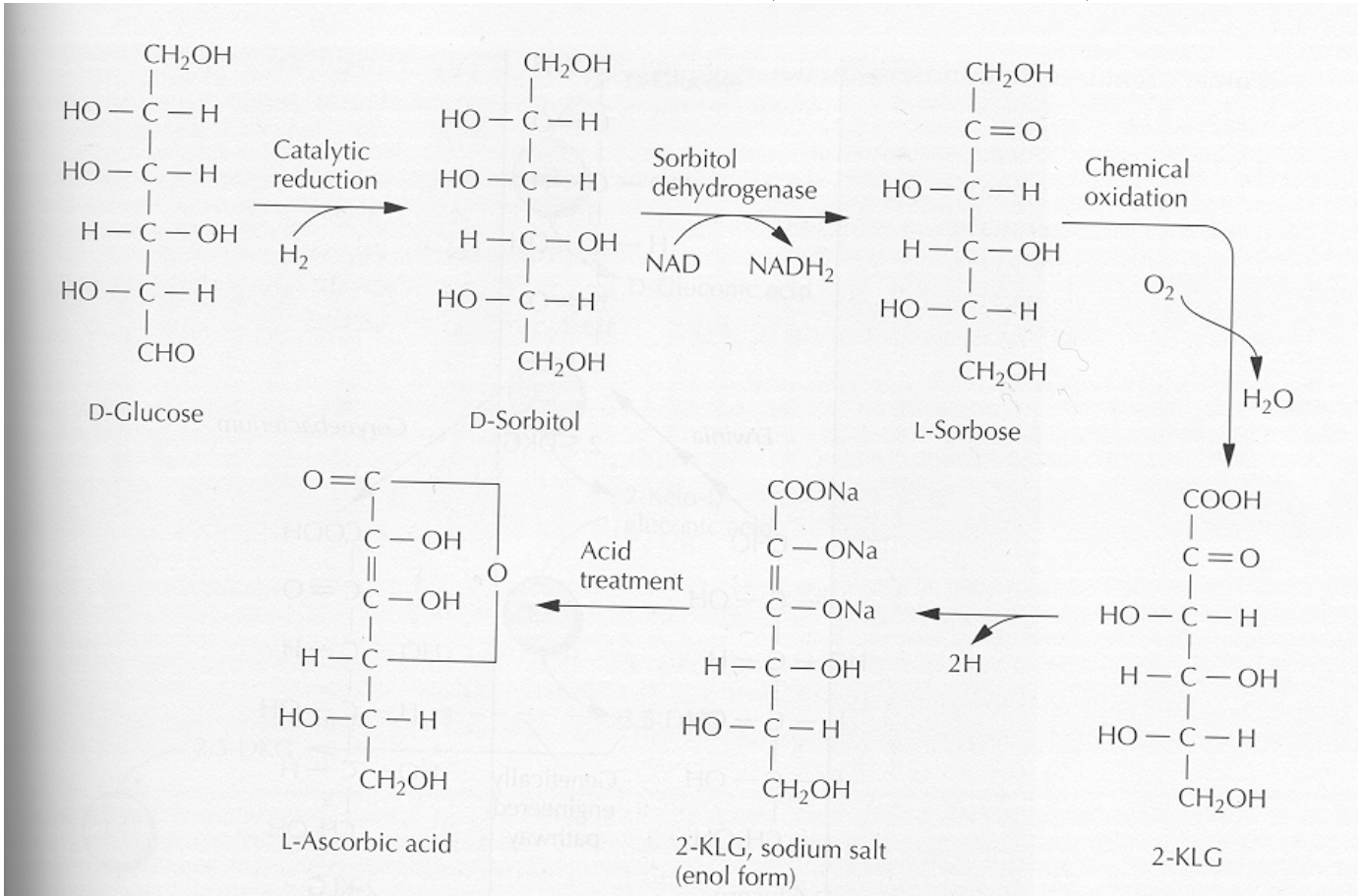


Figure 12.2 Commercial synthesis of L-ascorbic acid. Except for the microbial conversion of D-sorbitol to L-sorbose, all the steps are chemical reactions. The microbial conversion is carried out by *Acetobacter suboxydans*, which produces the enzyme sorbitol dehydrogenase.

Biological production of 2-KLG

1. Co-cultivation of *Erwinia* and *Corynebacterium*
   - advantages/disadvantages?
2. Sequential cultivation of *Erwinia* and *Coryn.*
   - advantages/disadvantages?
3. Genetically-engineered approach
   - which cell to transform?

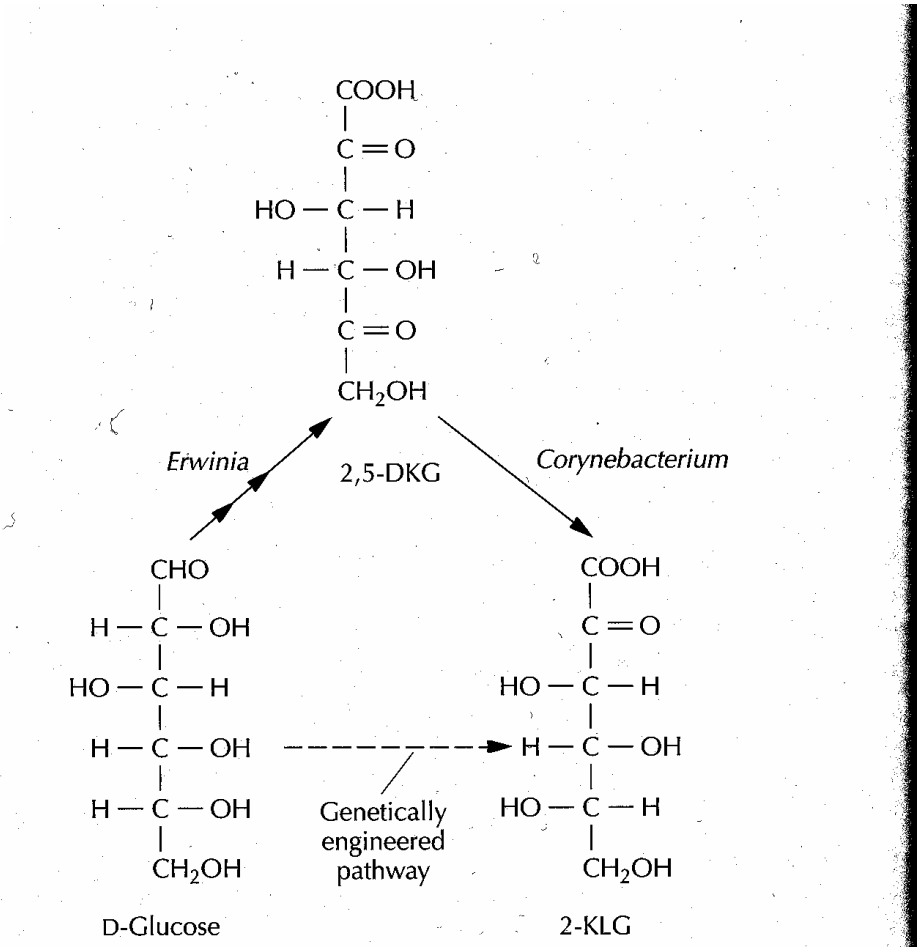


Figure 12.3 Biological production of 2-KLG. *Erwinia* has a set of three enzymes that synthesize 2,5-DKG from D-glucose. *Corynebacterium* has an enzyme that converts 2,5-DKG to 2-KLG. Thus, 2-KLG, which is the immediate precursor of L-ascorbic acid, can be produced from D-glucose either by cocultivating these two microorganisms or by genetically engineering *Erwinia* to express the enzyme from *Corynebacterium*, which converts 2,5-DKG to 2-KLG.

Recombinant *Erwinia* pathway

1. 2,5-DKG reductase from *Corynebacterium* purified
2. 40 amino acids from the N terminal end sequenced
3. Two 43 base pair DNA hybridization probes synthesized
4. *Corynebacterium* clone bank screened using probes
5. Target DNA responsible for 2,5-DKG reductase inserted into a cloning vector (plasmid)
6. *Erwinia* transformed using cloning vector-DNA construct

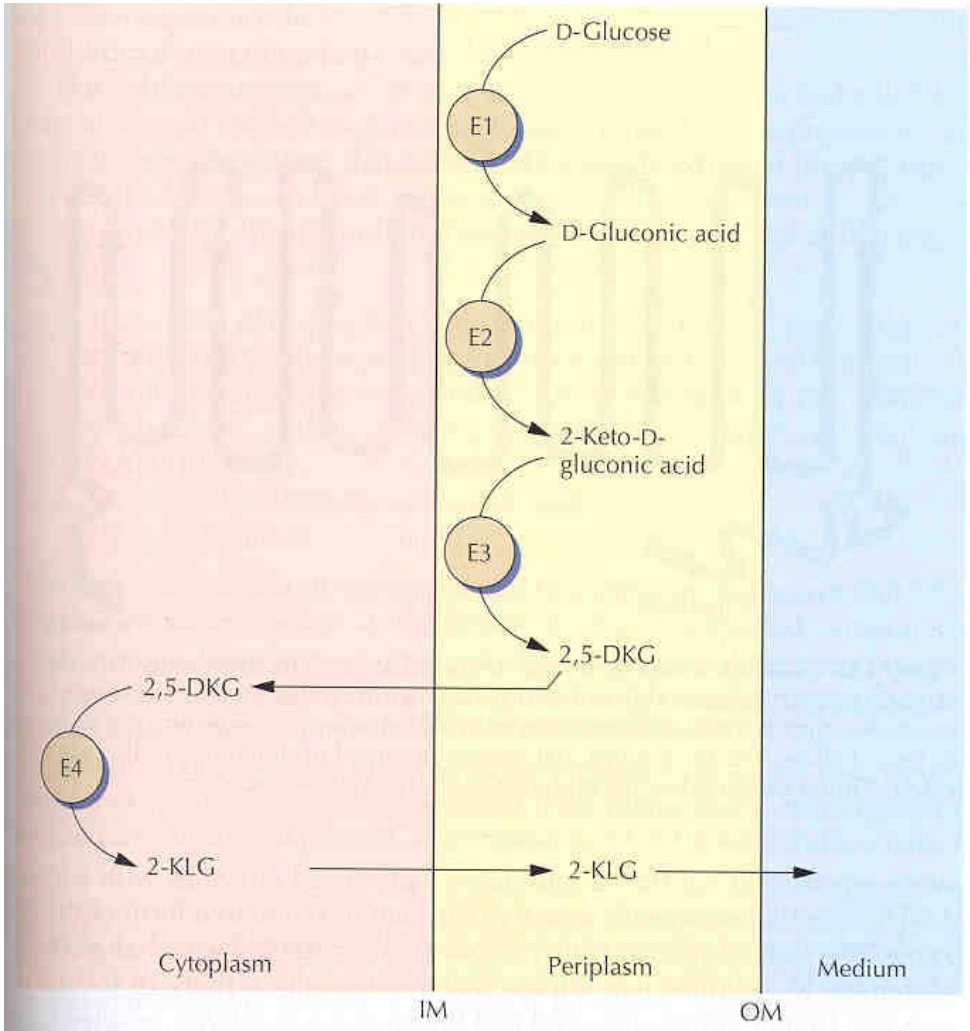


Figure 12.4 Conversion of D-glucose to 2-KLG by recombinant *Erwinia herbicola*. The cellular locations of all of the component enzymes are indicated by the letter E and are numbered consecutively. The major intermediates in the pathway are named. IM and OM denote the inner and outer membranes, respectively.

Increase 2,5-DKG reductase activity

Replace certain amino acids on the protein (enzyme).

How would this task be accomplished?

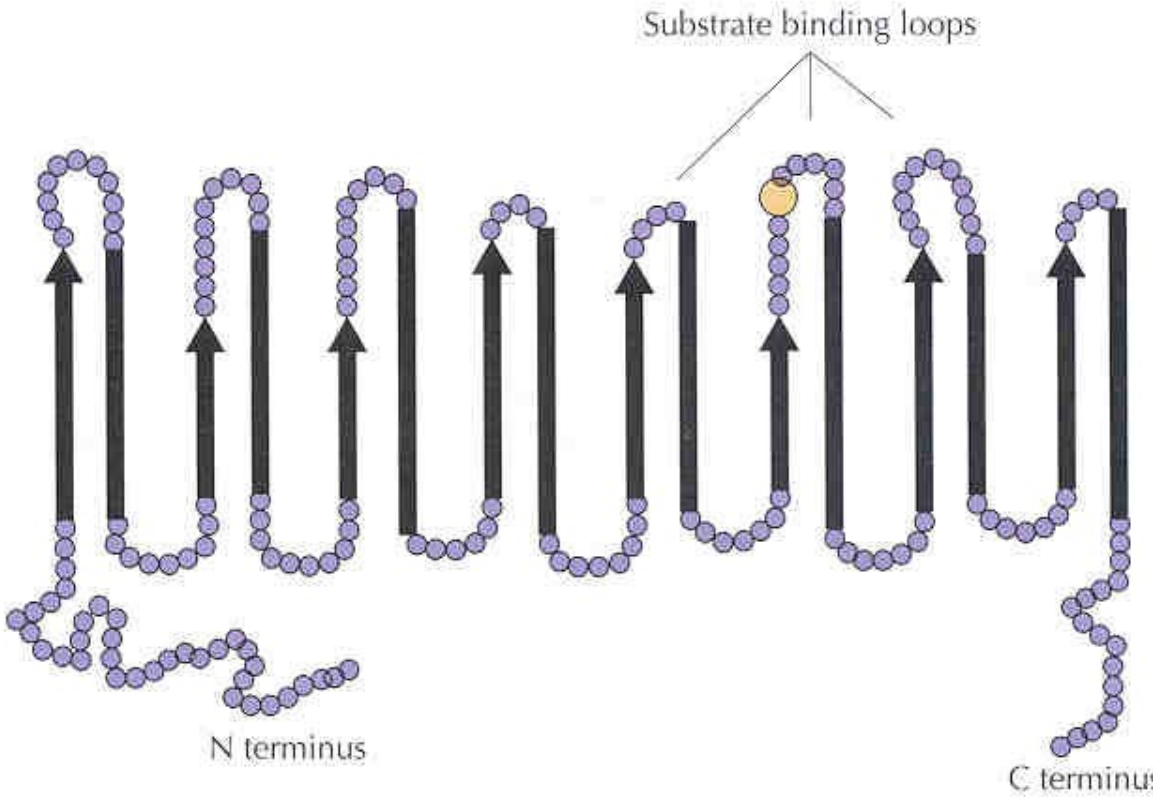


Figure 12.5 Predicted structure of 2,5-DKG reductase. The solid arrows indicate β-stranded regions; the solid bars are α-helical regions; and the circles are amino acid residues either at the N or C terminus or involved in loops connecting the β-strands to the α-helices. The three loops that may be involved in substrate binding are indicated. Amino acid residue 192 is shown as a large yellow circle.