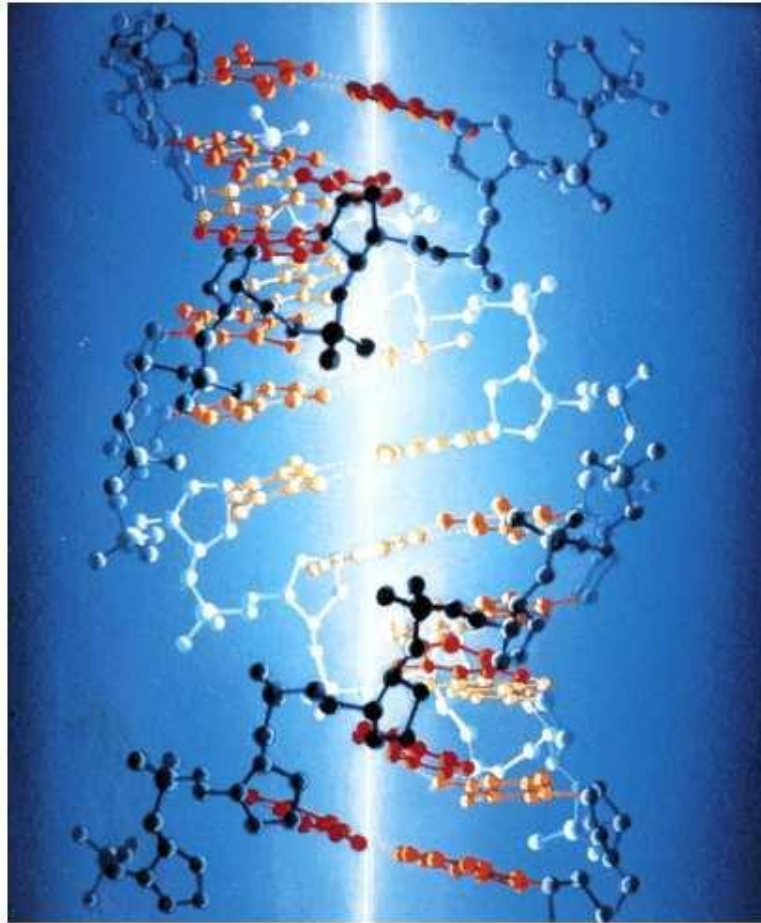


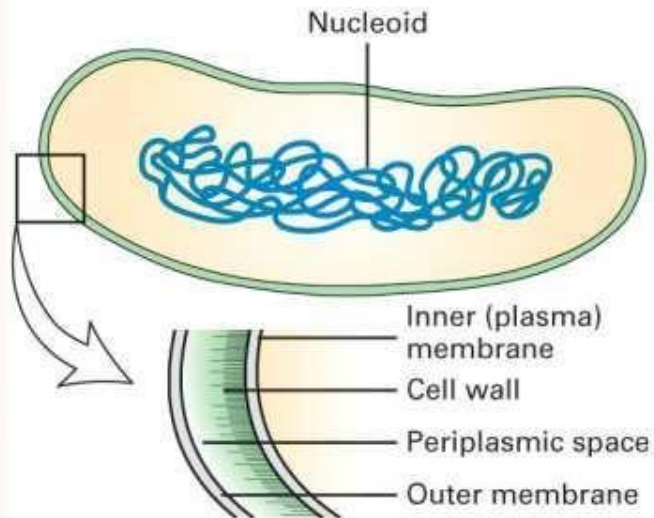
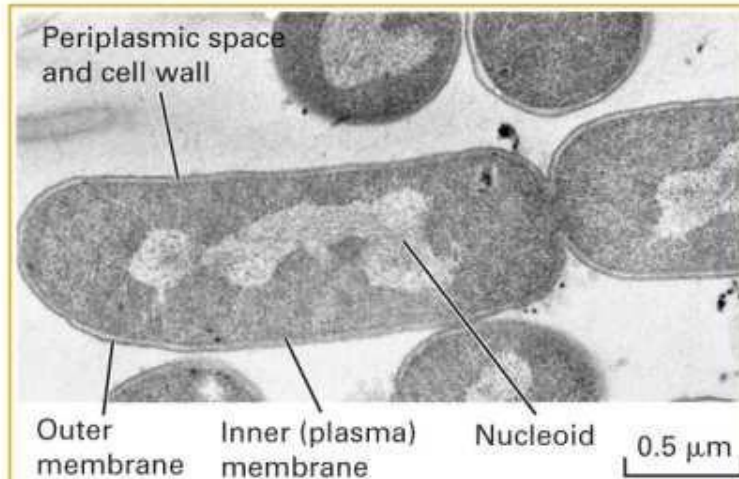
# DNA, RNA, Recombinant DNA Technology



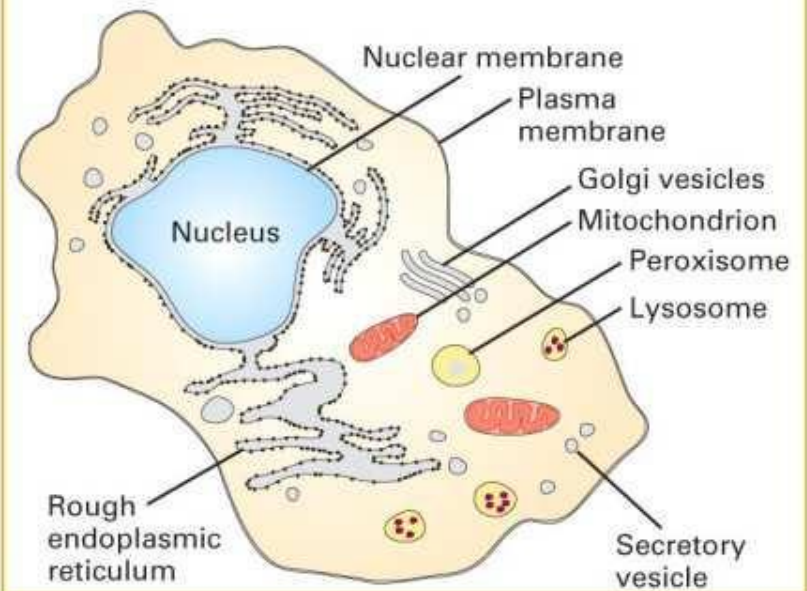
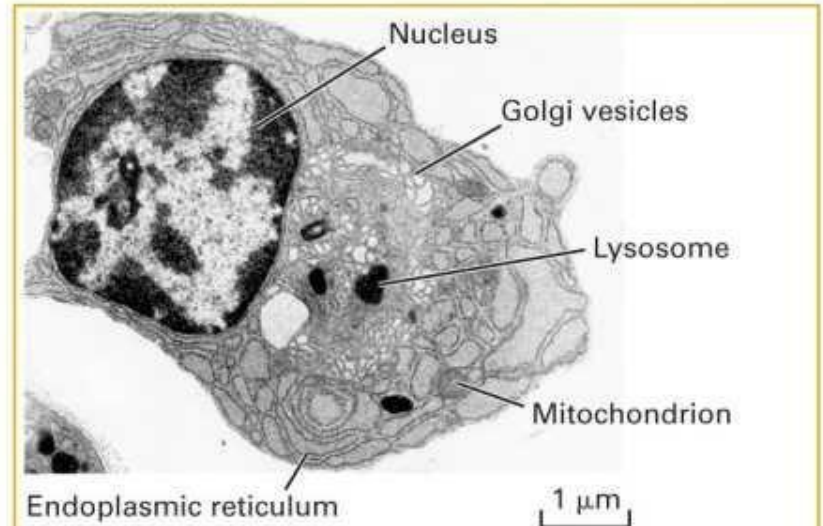
Chapter 3 Opener Fundamentals of Biochemistry, 2/e

# “Metabolic pathways” expanded

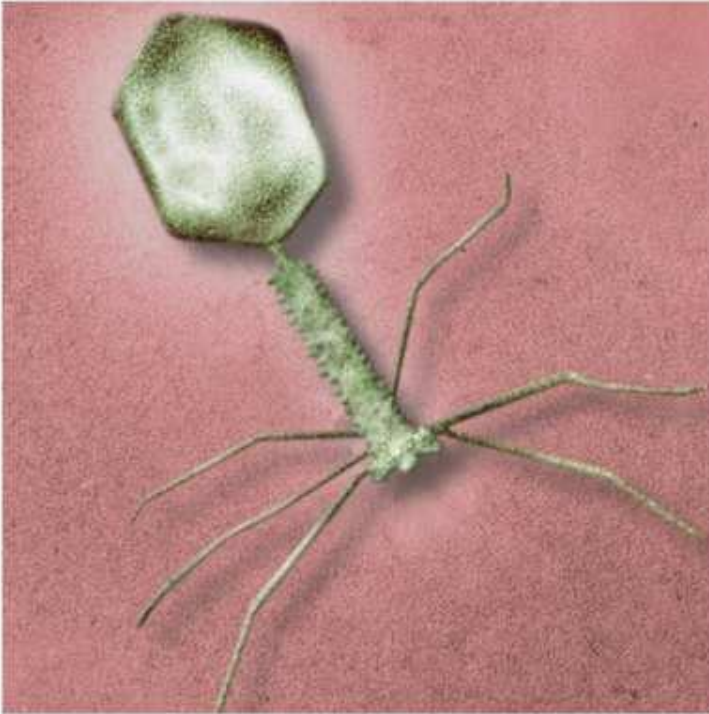
(a) Prokaryotic cell



(b) Eukaryotic cell



Model organisms: Cellular biology, biochemistry...  
molecular biology



## Viruses

Proteins involved in DNA, RNA,  
protein synthesis

Gene regulation

Cancer and control of cell  
proliferation

Transport of proteins and  
organelles inside cells

Infection and immunity

Possible gene therapy approaches



## **Fruit fly (*Drosophila melanogaster*)**

Development of the body plan

Generation of differentiated cell lineages

Formation of the nervous system, heart, and musculature

Programmed cell death

Genetic control of behavior

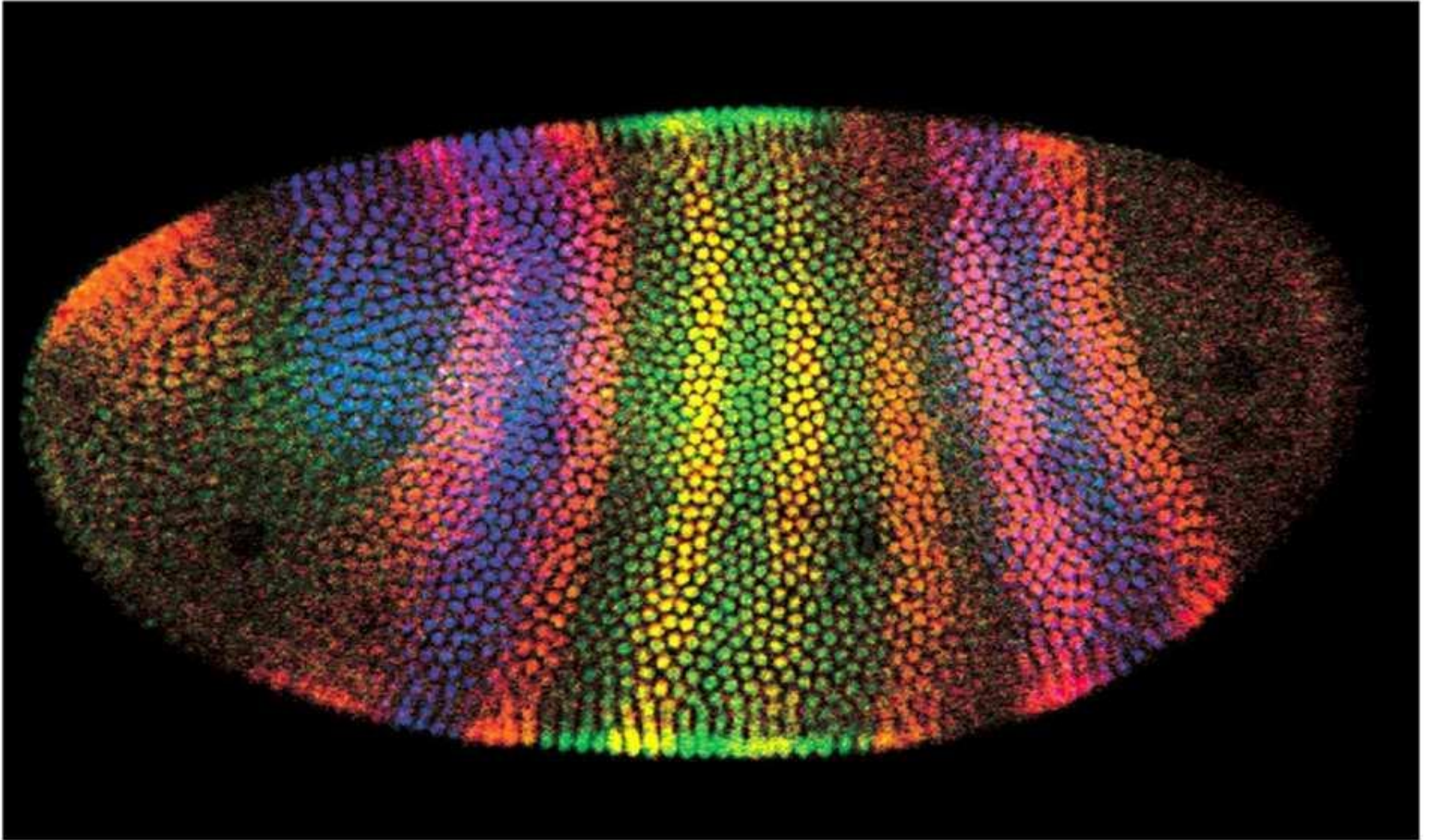
Cancer genes and control of cell proliferation

Control of cell polarization

Effects of drugs, alcohol, pesticides



Developmental biology.....



# Fly mutation “eyeless”

(b)



(c)



‘The fly and you are not much different.’



## **Mice**, including cultured cells

Development of body tissues  
Function of mammalian immune system

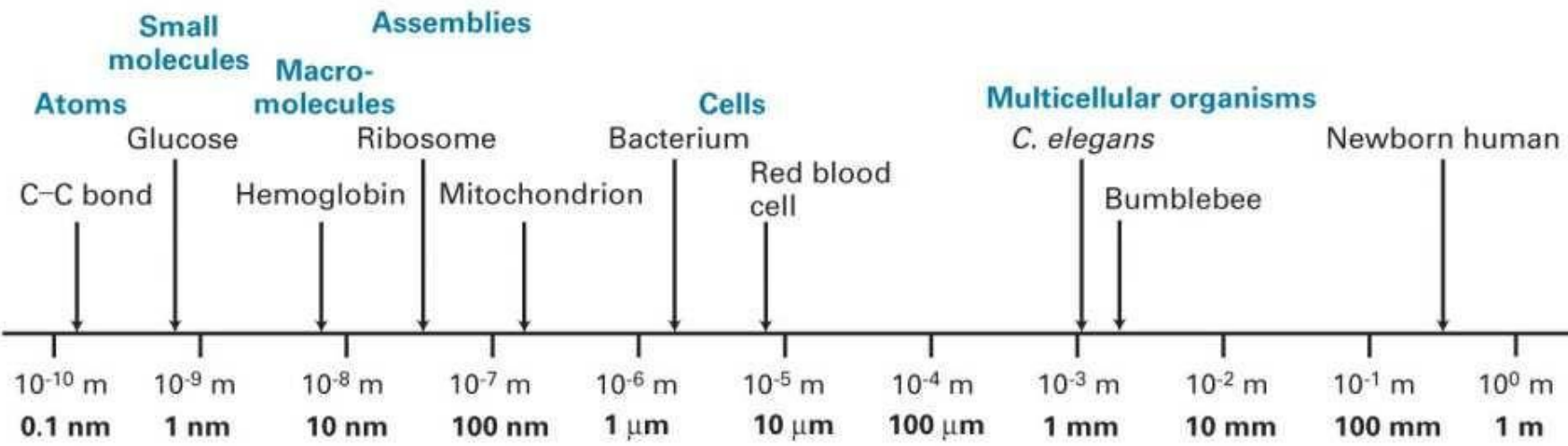
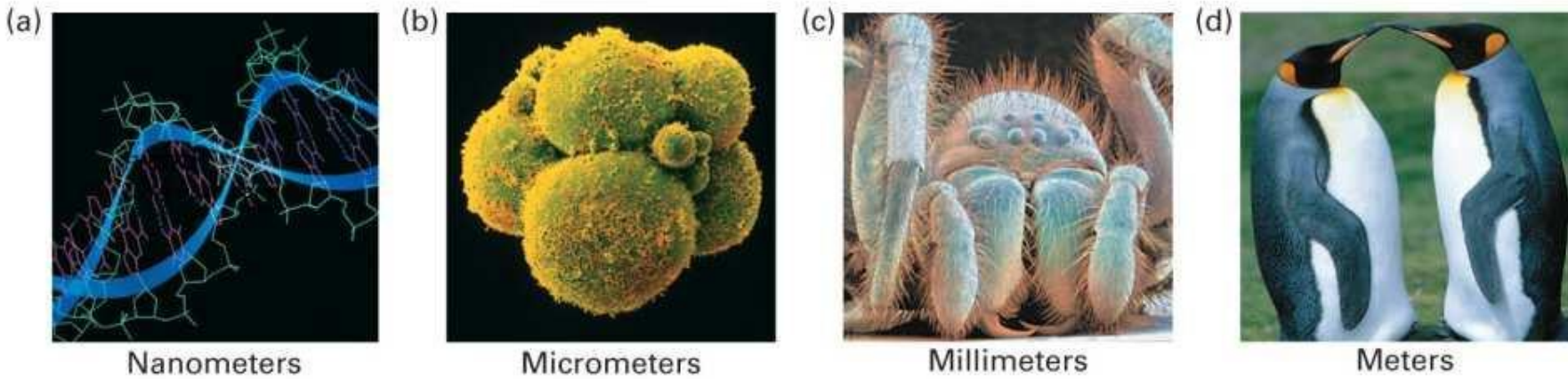
Formation and function of brain and nervous system

Models of cancers and other human diseases

Gene regulation and inheritance  
Infectious disease

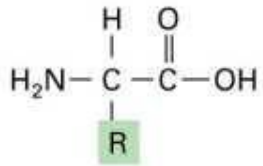


The 'first' science: "technology drives research drives technology dri..."



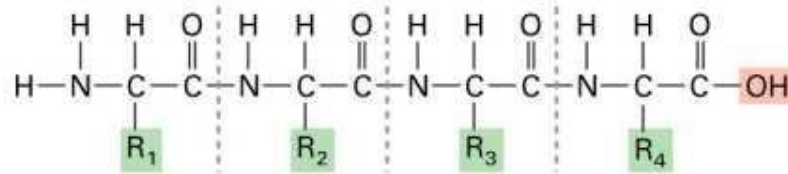


### MONOMERS

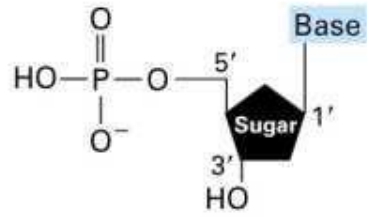
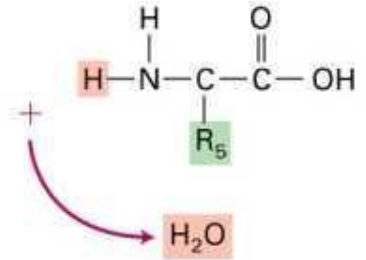


**Amino acid**

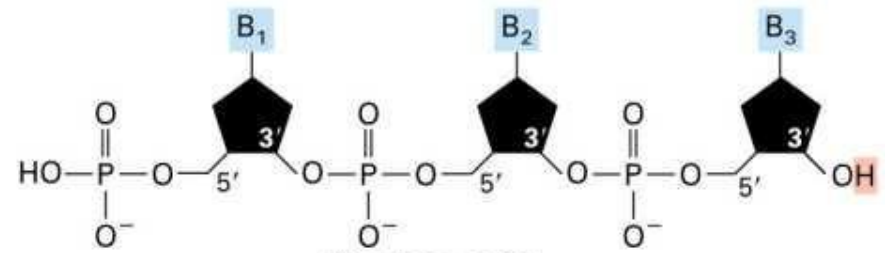
### POLYMERS



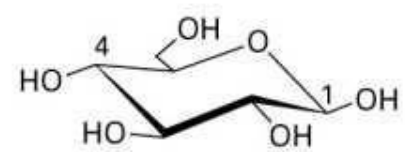
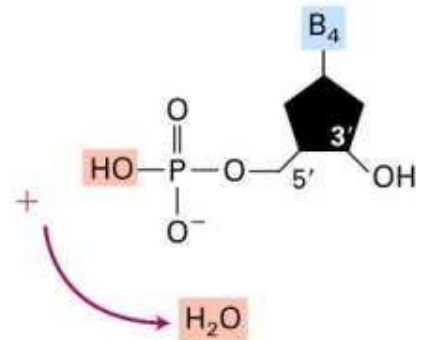
**Polypeptide**



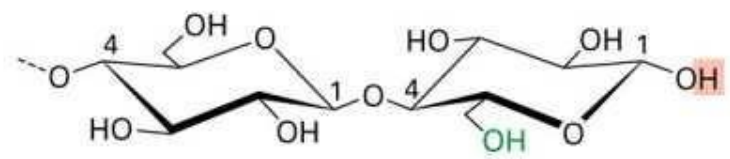
**Nucleotide**



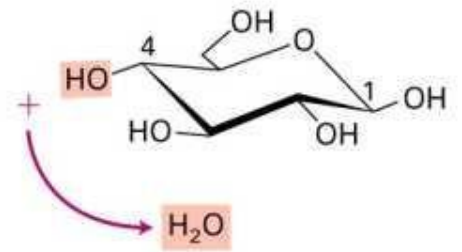
**Nucleic acid**



**Monosaccharide**

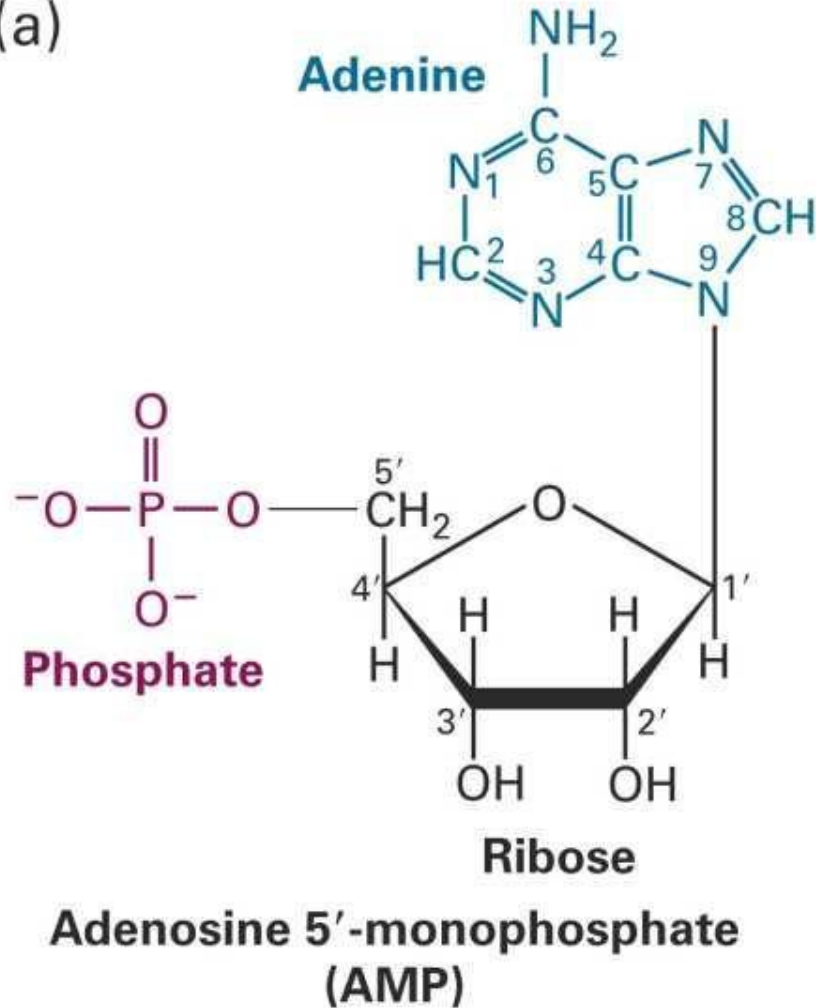


**Polysaccharide**

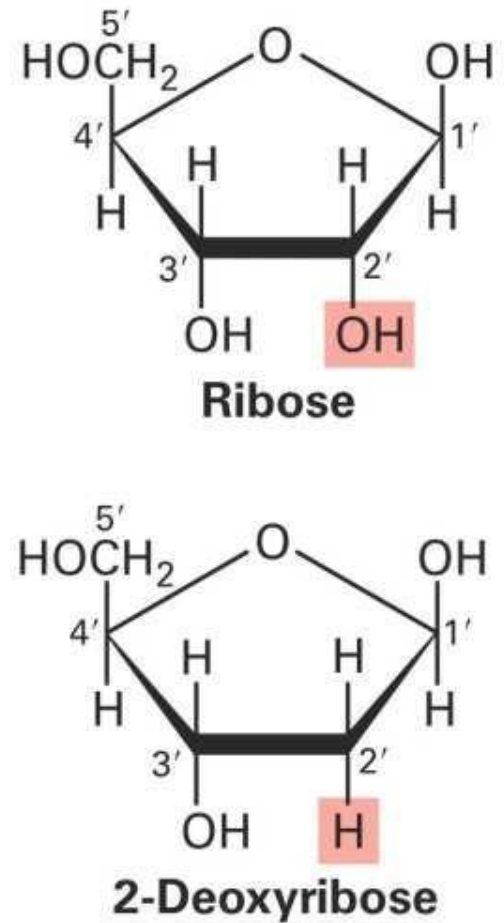


# Nucleic Acids - DNA and RNA

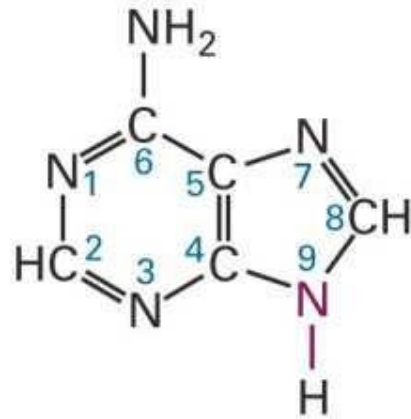
(a)



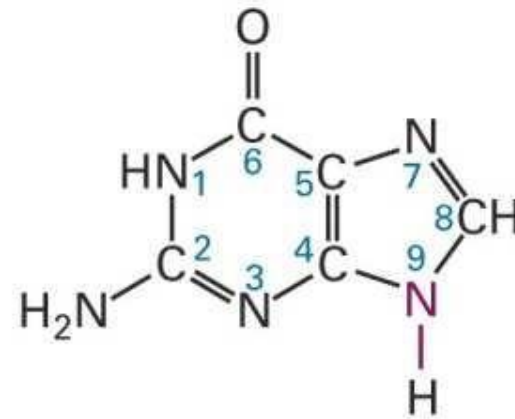
(b)



## PURINES

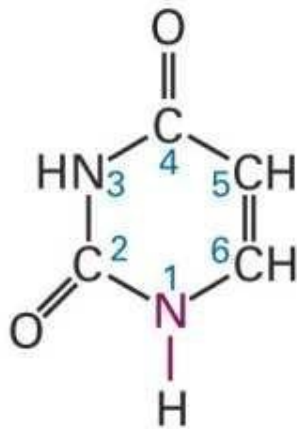


**Adenine (A)**

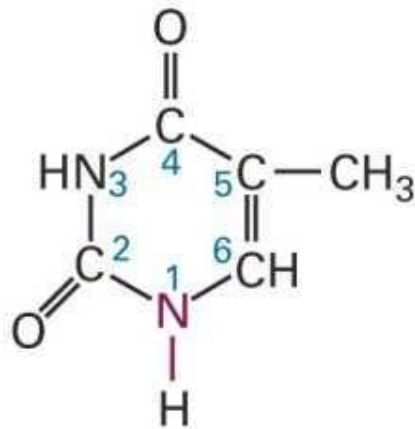


**Guanine (G)**

## PYRIMIDINES



**Uracil (U)**



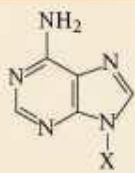
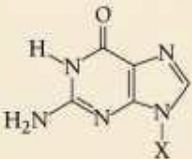
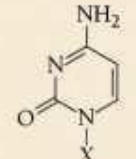
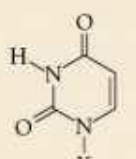
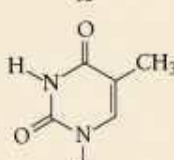
**Thymine (T)**



**Cytosine (C)**



**Table 3-1 Names and Abbreviations of Nucleic Acid Bases, Nucleosides, and Nucleotides**

Base Formula	Base (X = H)	Nucleoside (X = ribose <sup>a</sup> )	Nucleotide <sup>b</sup> (X = ribose phosphate <sup>a</sup> )
	Adenine Adc A	Adenosine Ado A	Adenylic acid Adenosine monophosphate AMP
	Guanine Gua G	Guanosine Guo G	Guanylic acid Guanosine monophosphate GMP
	Cytosine Cyt C	Cytidine Cyd C	Cytidylic acid Cytidine monophosphate CMP
	Uracil Ura U	Uridine Urd U	Uridylic acid Uridine monophosphate UMP
	Thymine Thy T	Deoxythymidine dThd dT	Deoxythymidylic acid Deoxythymidine monophosphate dTMP

DNA + RNA

DNA + RNA

DNA + RNA

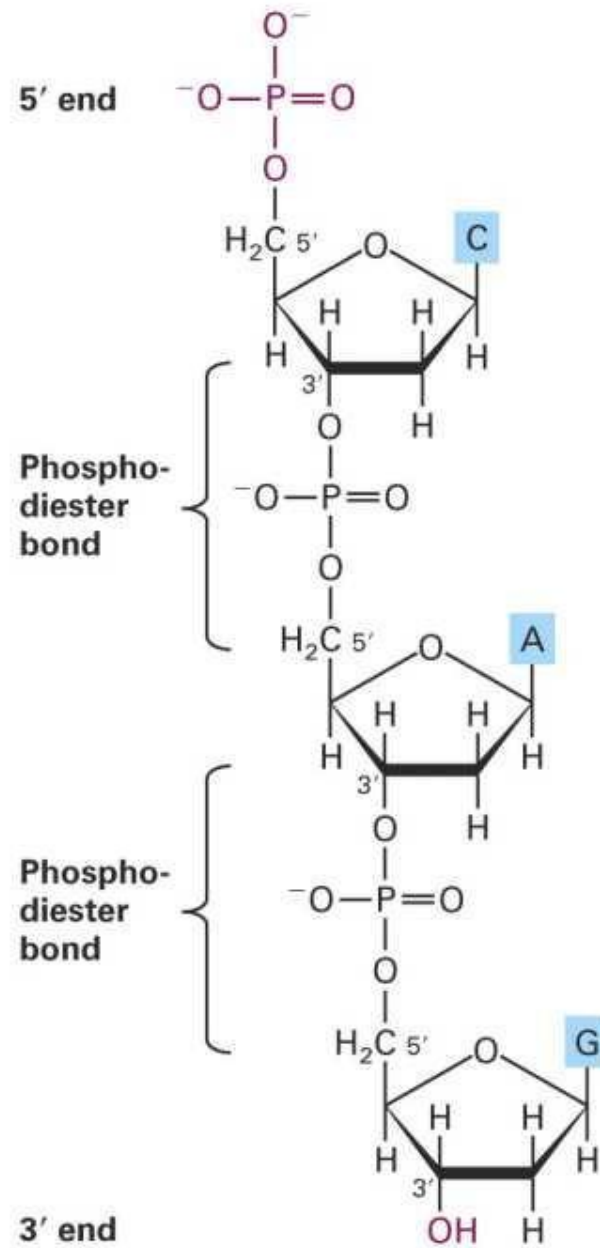
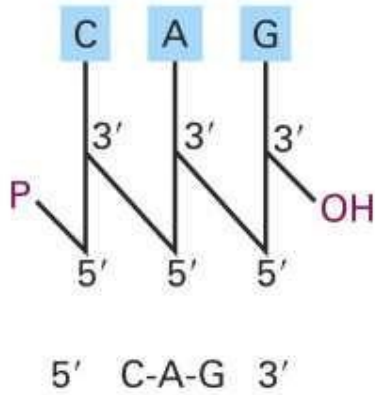
RNA ←

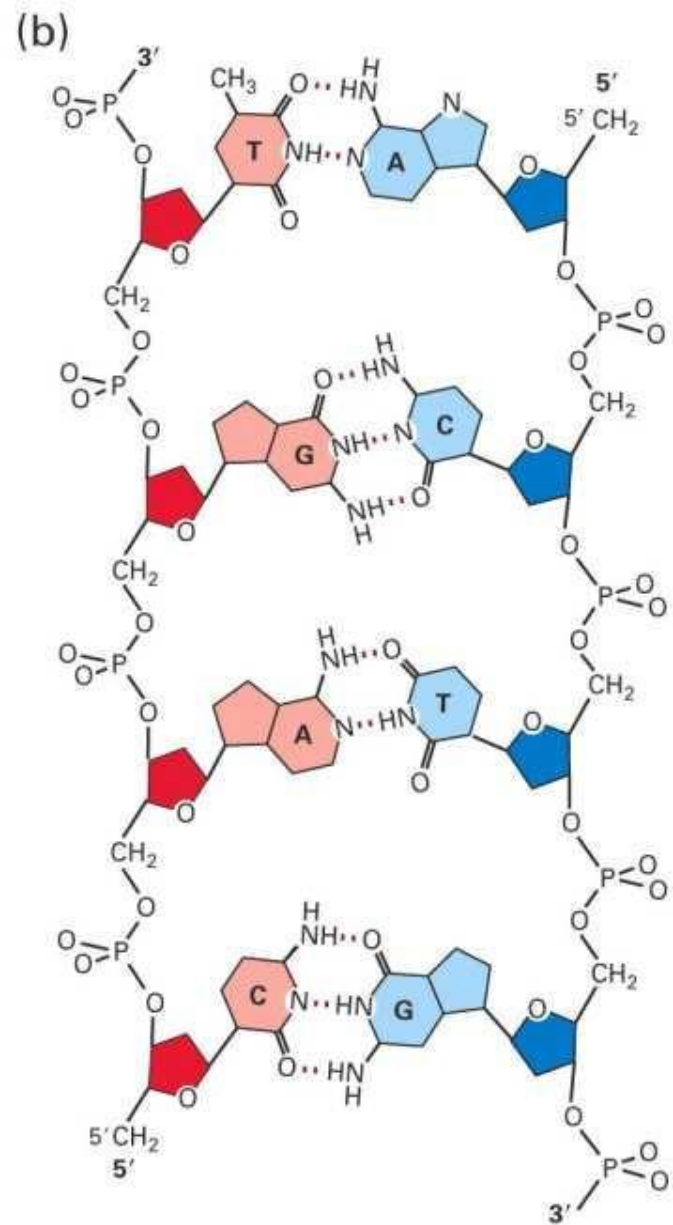
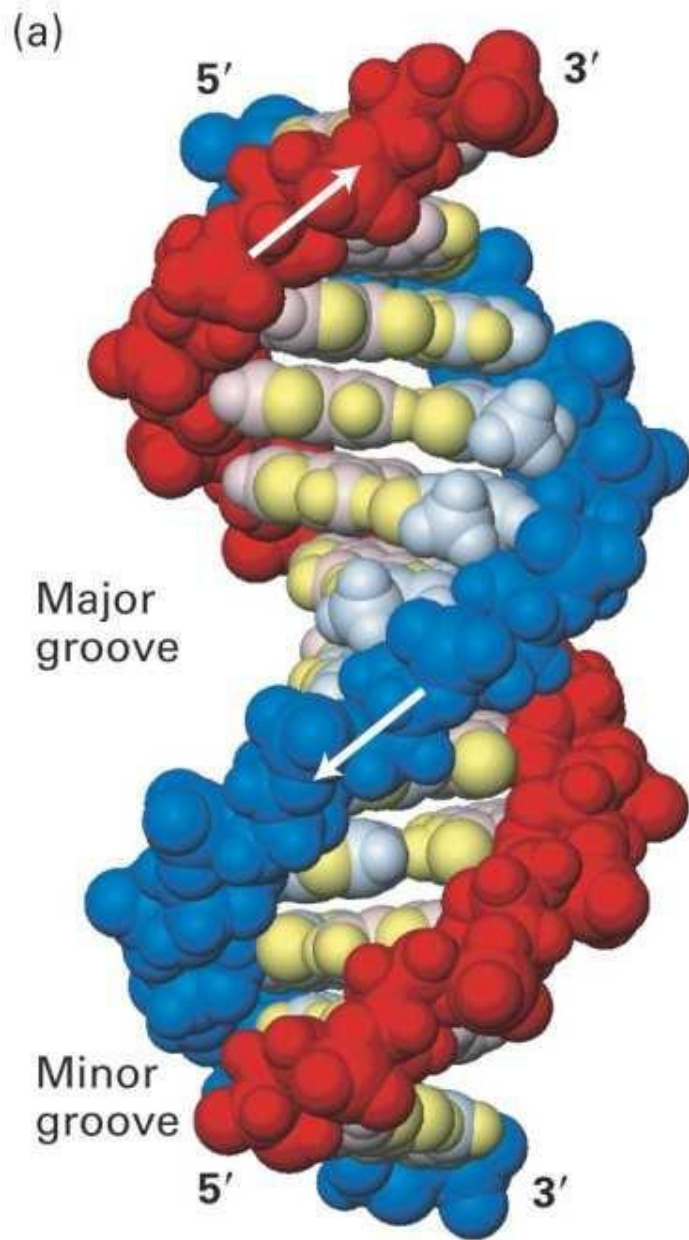
DNA ←

<sup>a</sup>The presence of a 2'-deoxyribose unit in place of ribose, as occurs in DNA, is implied by the prefixes "deoxy" or "d." For example, the deoxy-nucleoside of adenine is deoxyadenosine or dA. However, for thymine-containing residues, which rarely occur in RNA, the prefix is redundant and may be dropped. The presence of a ribose unit may be explicitly implied by the prefix "ribo."

<sup>b</sup>The position of the phosphate group in a nucleotide may be explicitly specified as in, for example, 3'-AMP and 5'-GMP.

# DNA structure -> sequence

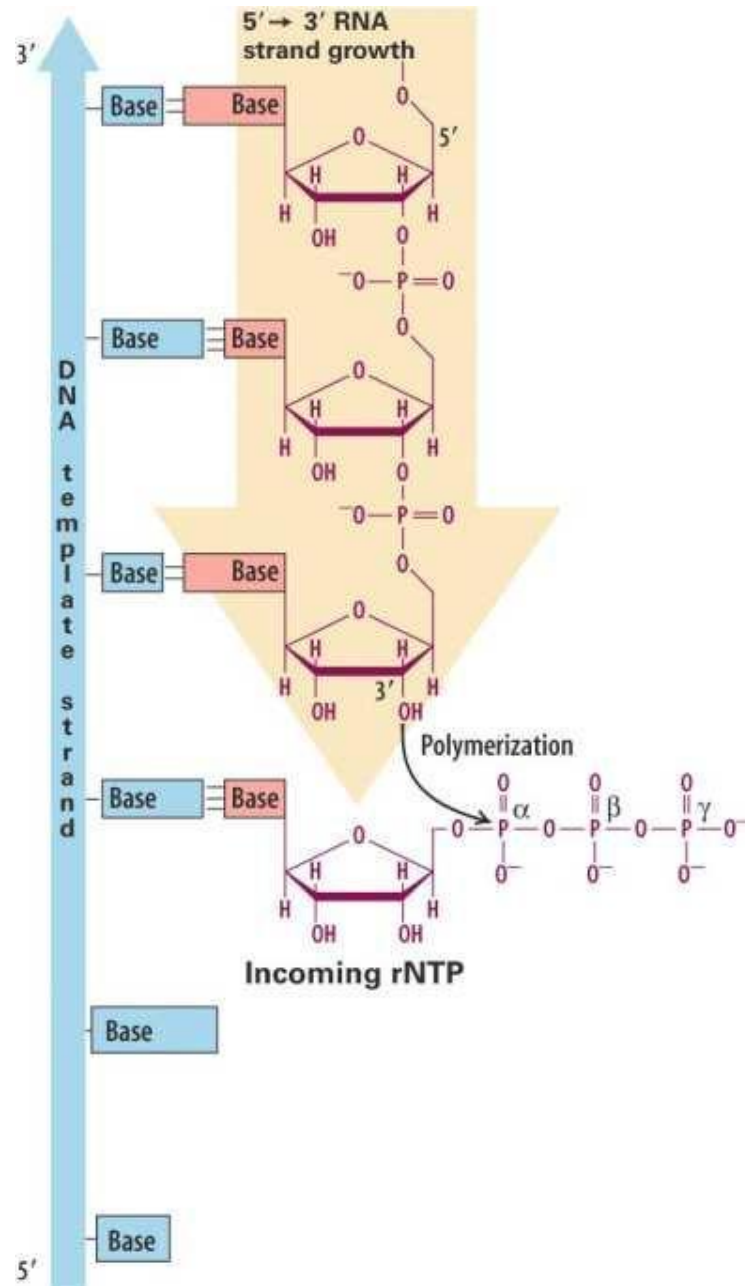




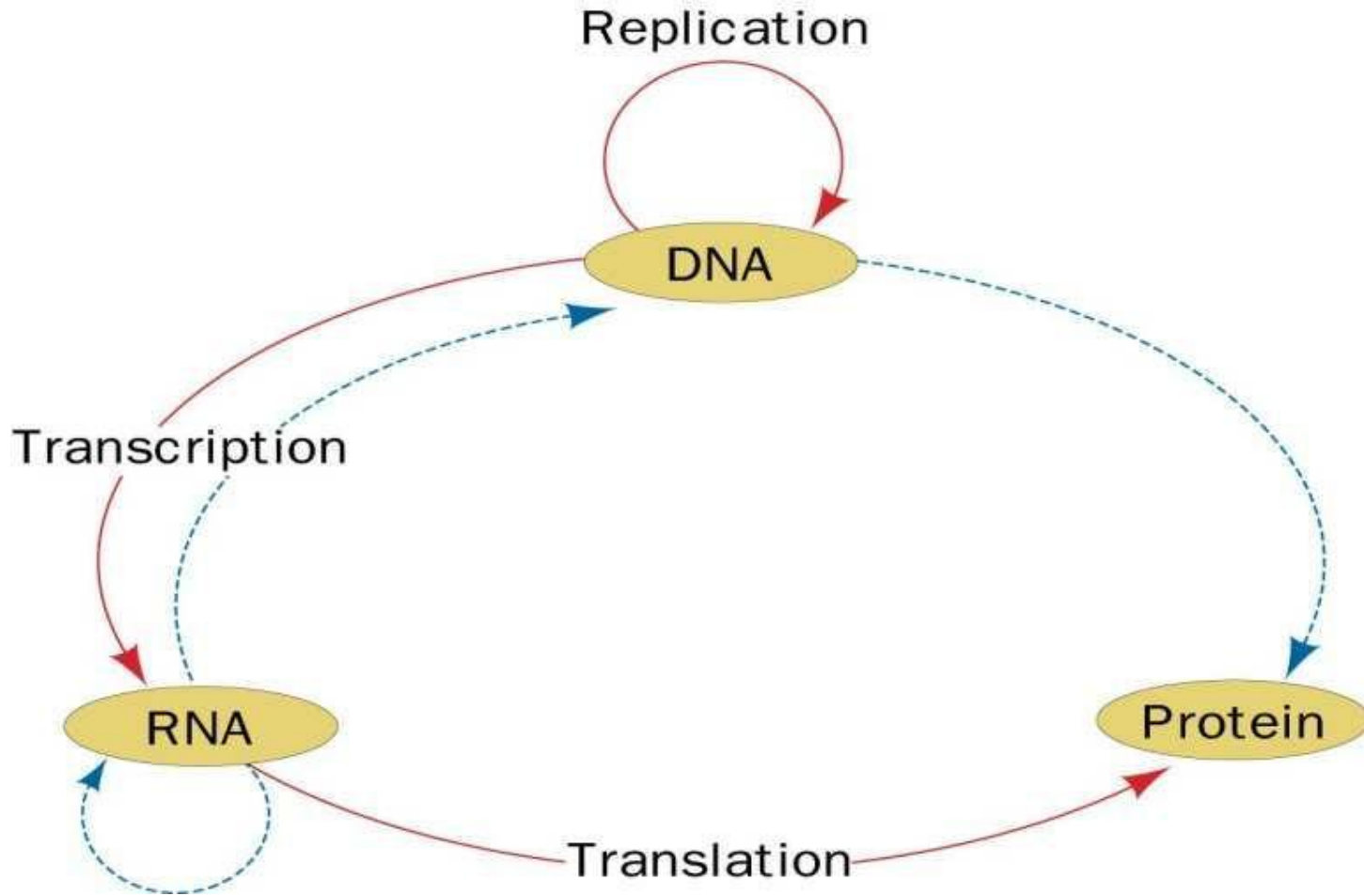


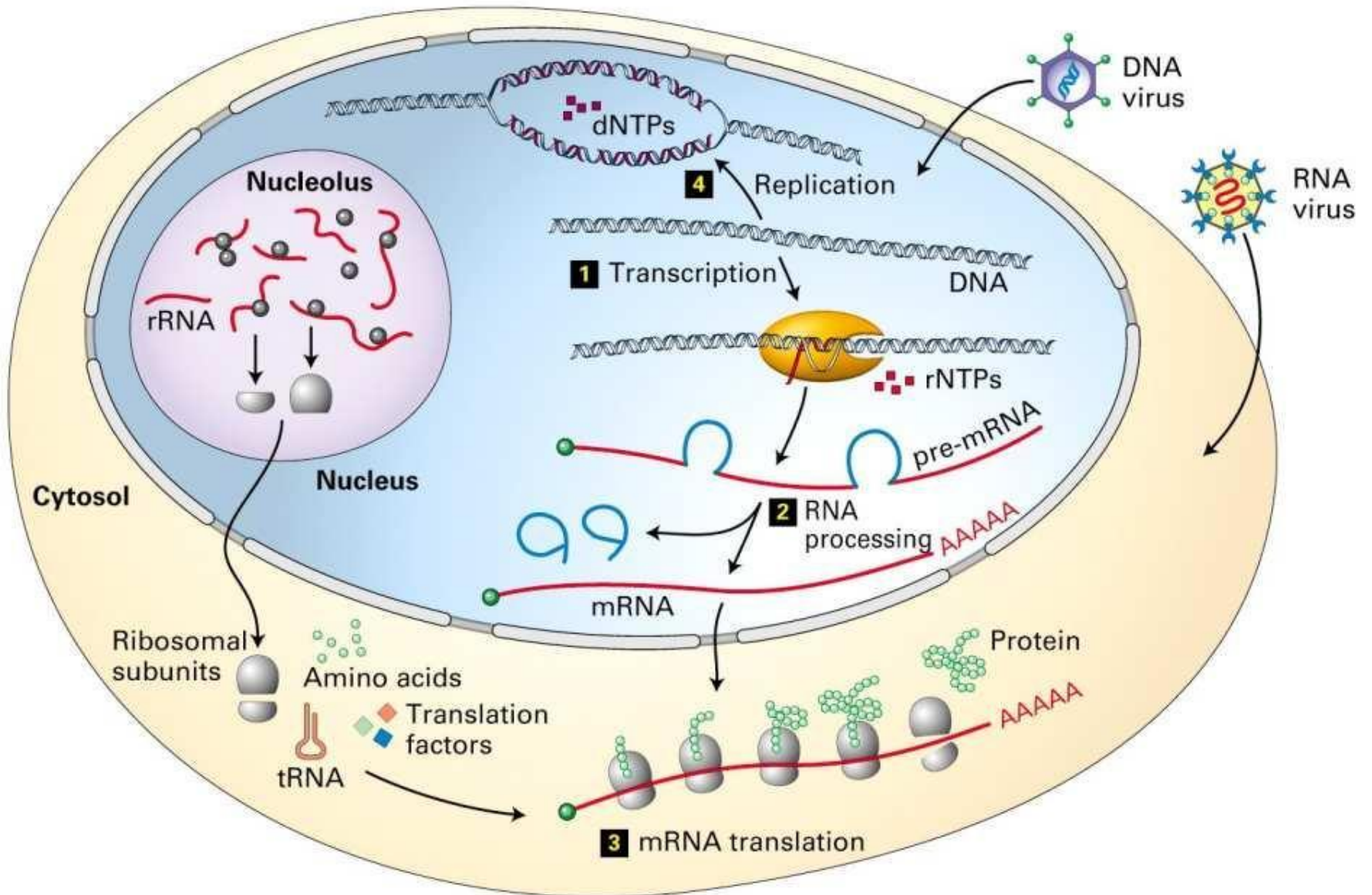
Polymerase reaction:

5' → 3'



# The central dogma







# Gene expression.

DNA 5' — A-G-A-G-G-T-G-C-T — 3'  
3' — T-C-T-C-C-A-C-G-A — 5'

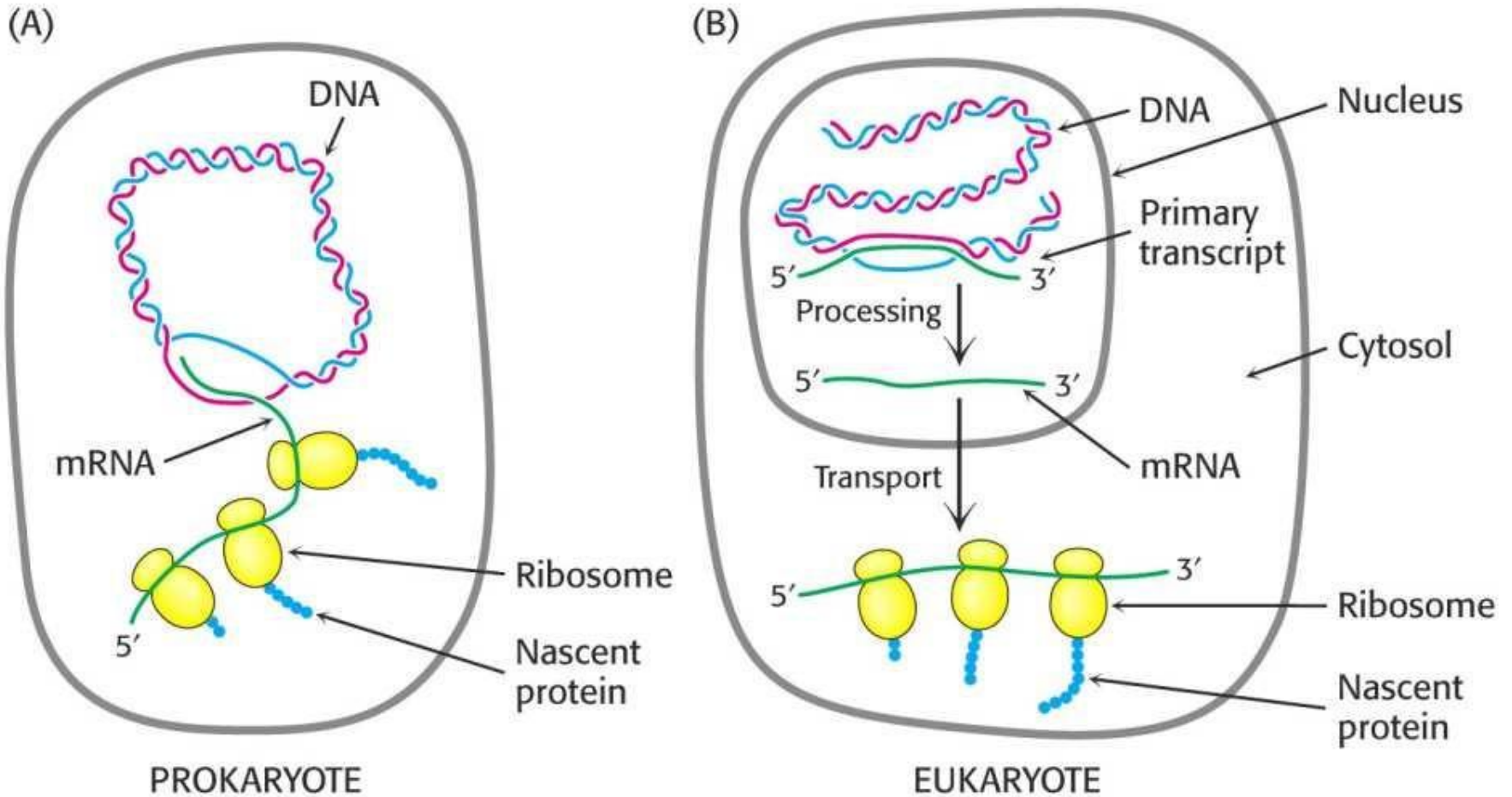


mRNA 5' — A-G-A-G-G-U-G-C-U — 3'  
tRNAs U-C-U C-C-A C-G-A

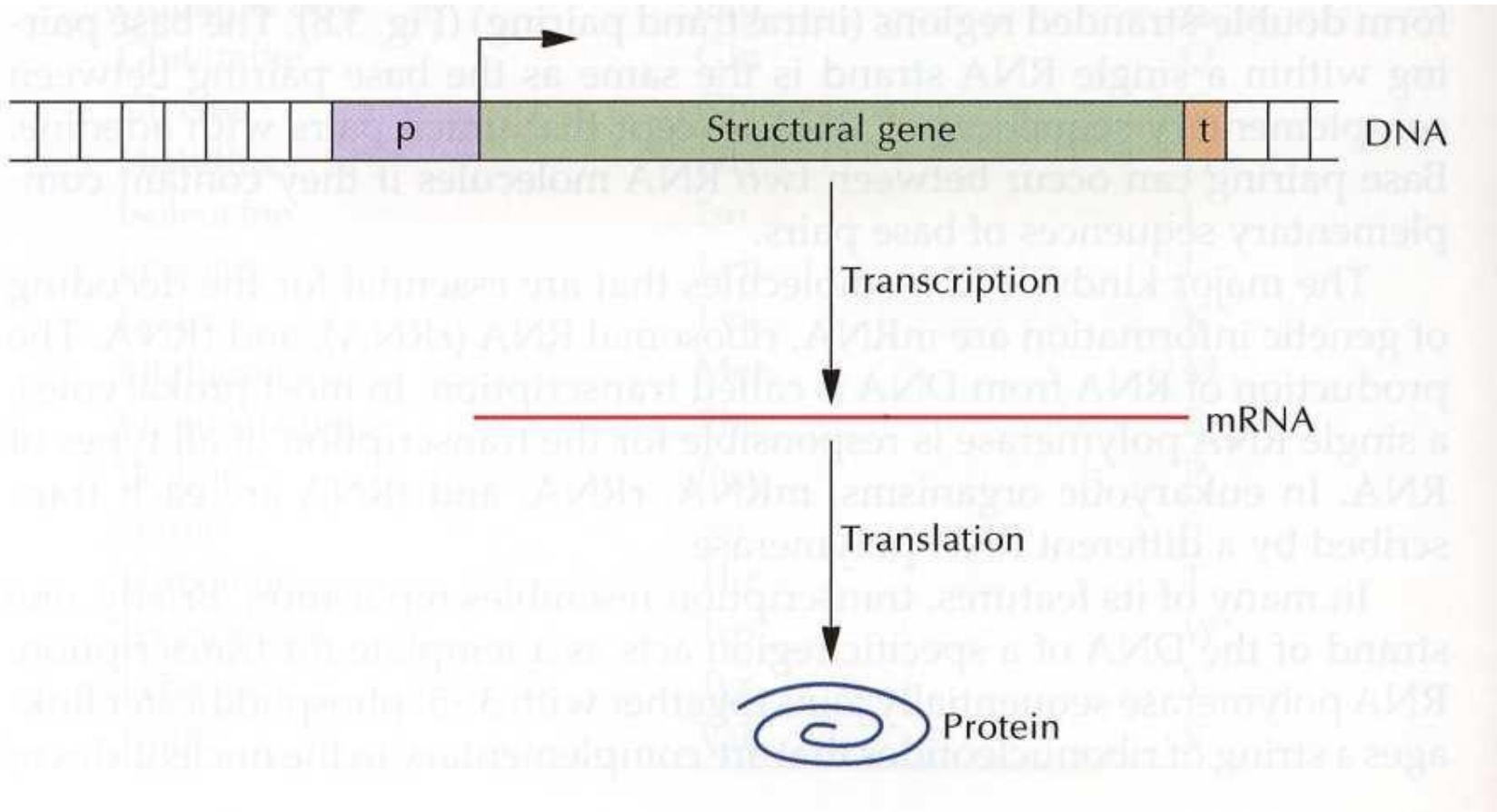
┌───┐ ┌───┐ ┌───┐  
| | | | | |  
└───┘ └───┘ └───┘  
Arginine Glycine Alanine

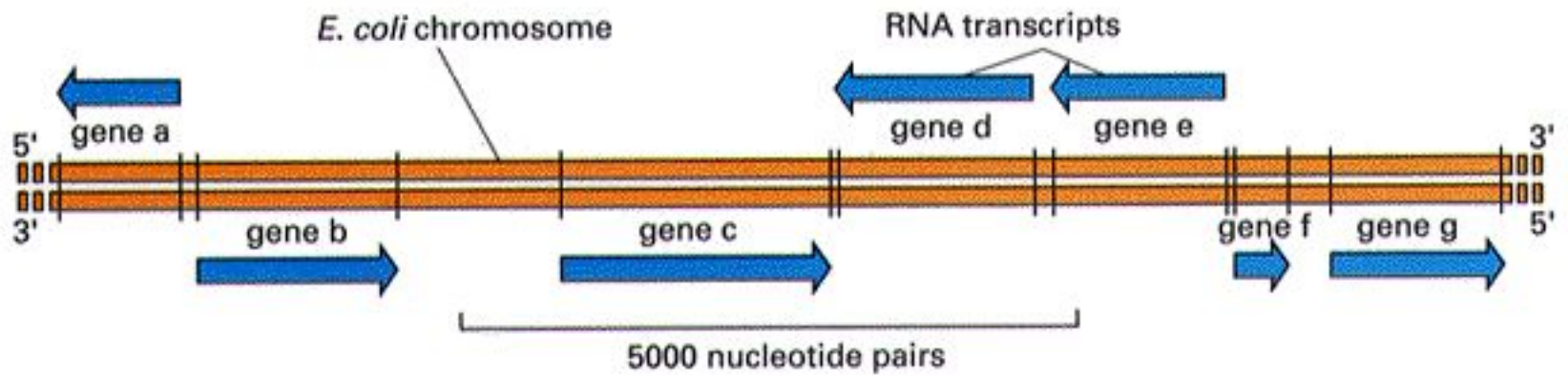


Polypeptide —Arg-Gly-Ala—



# What is a gene?









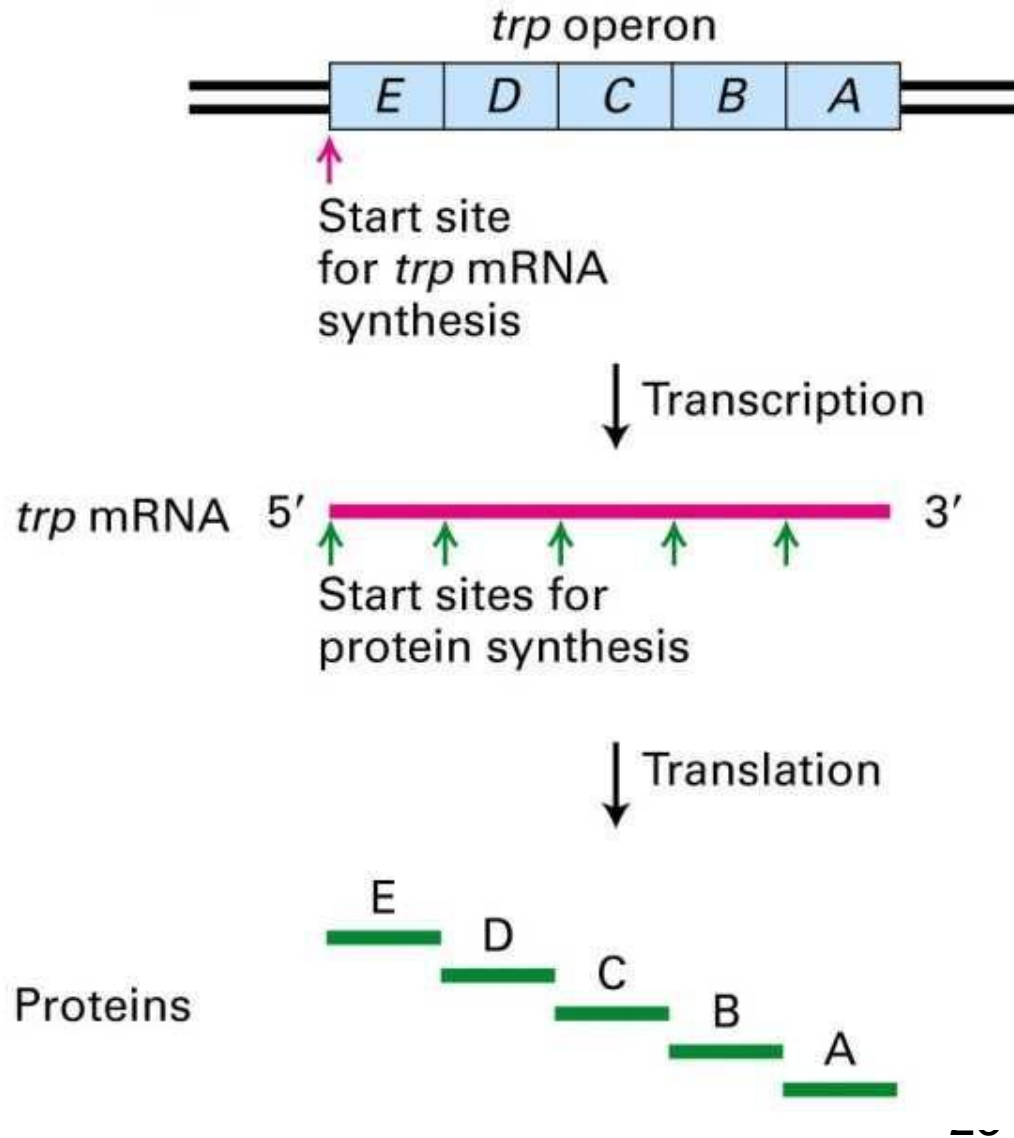
**(a) Monocistronic**



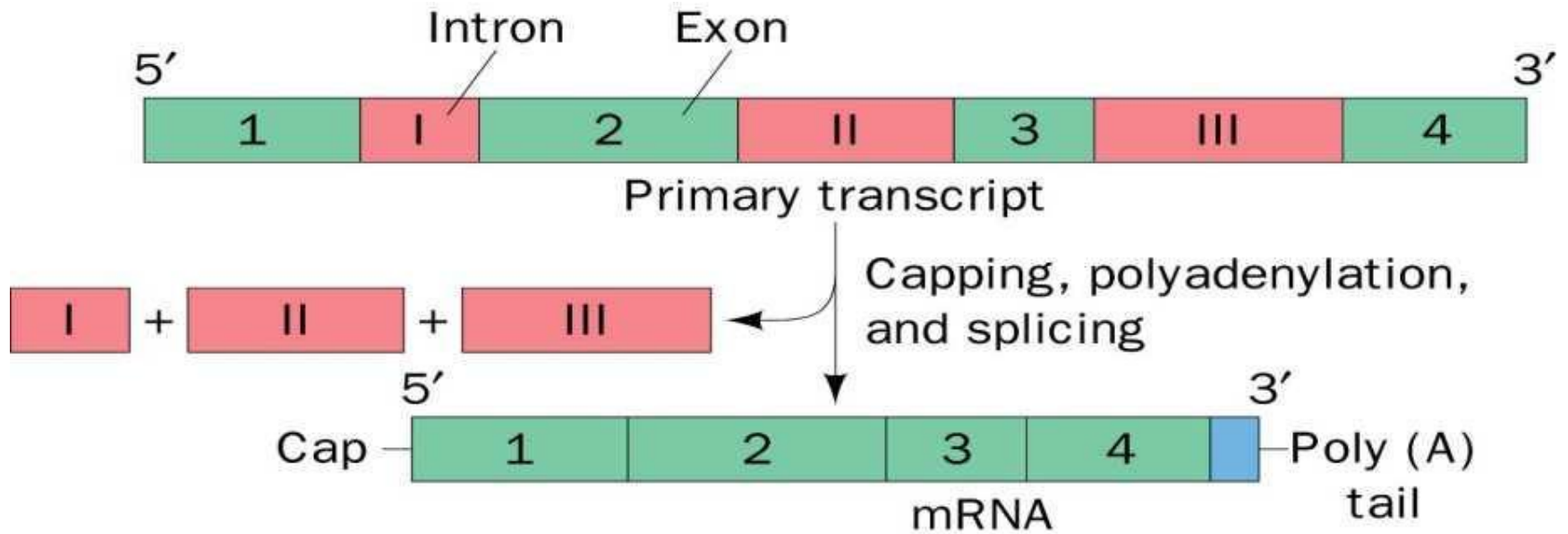
**(b) Polycistronic**

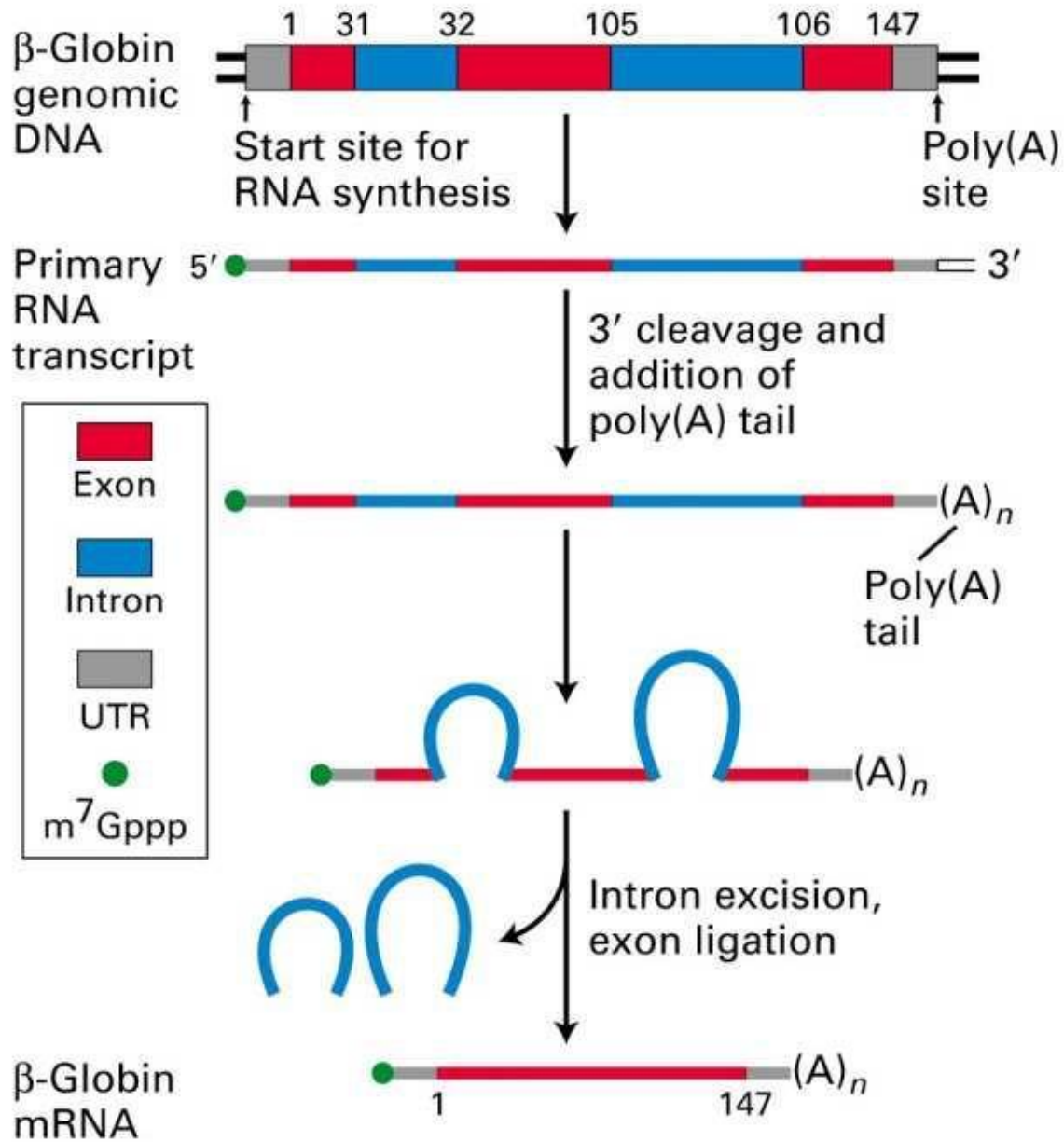
(a) Prokaryotes

*E. coli* genome

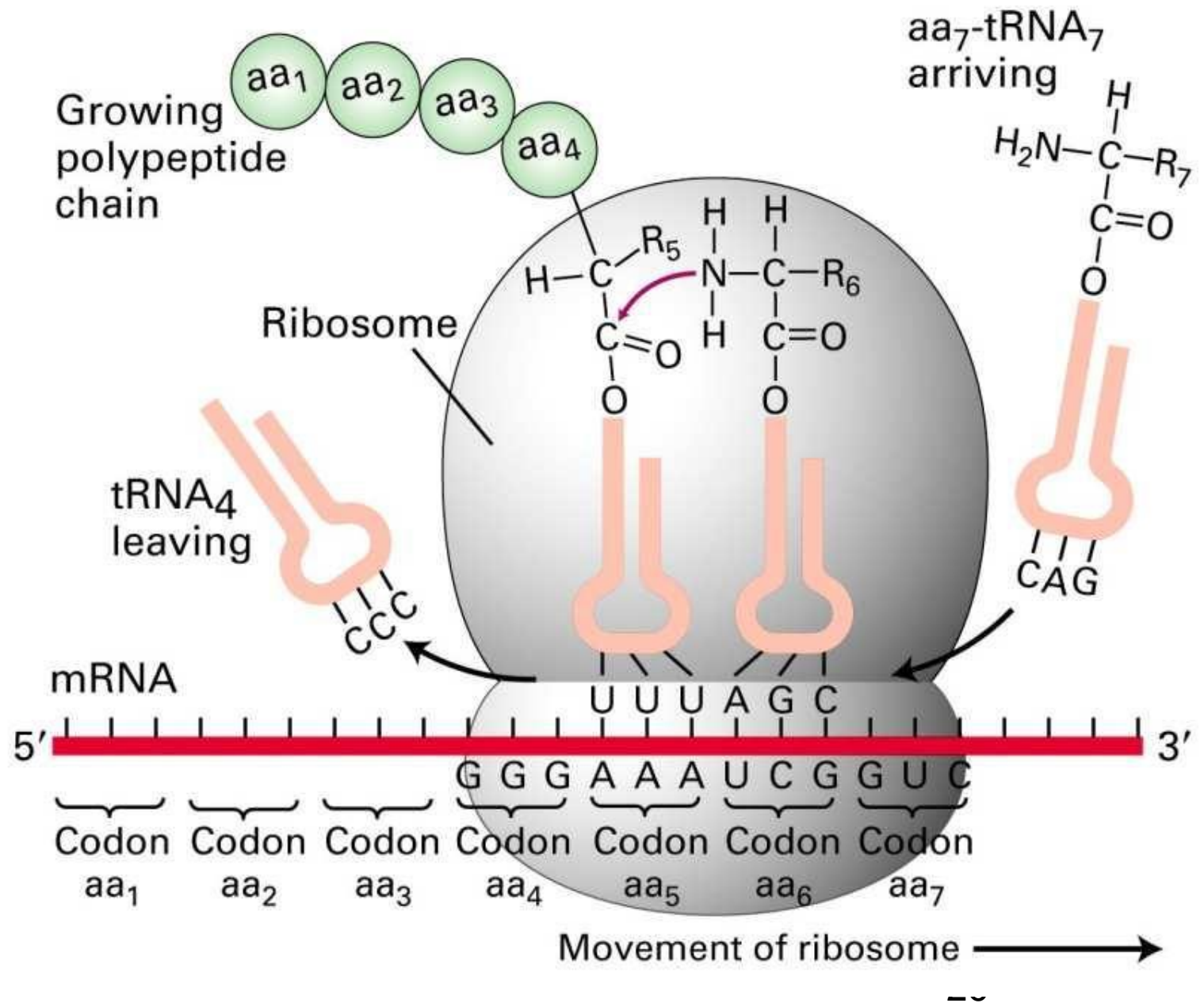


# Eukaryotes - Intron-Exon concept







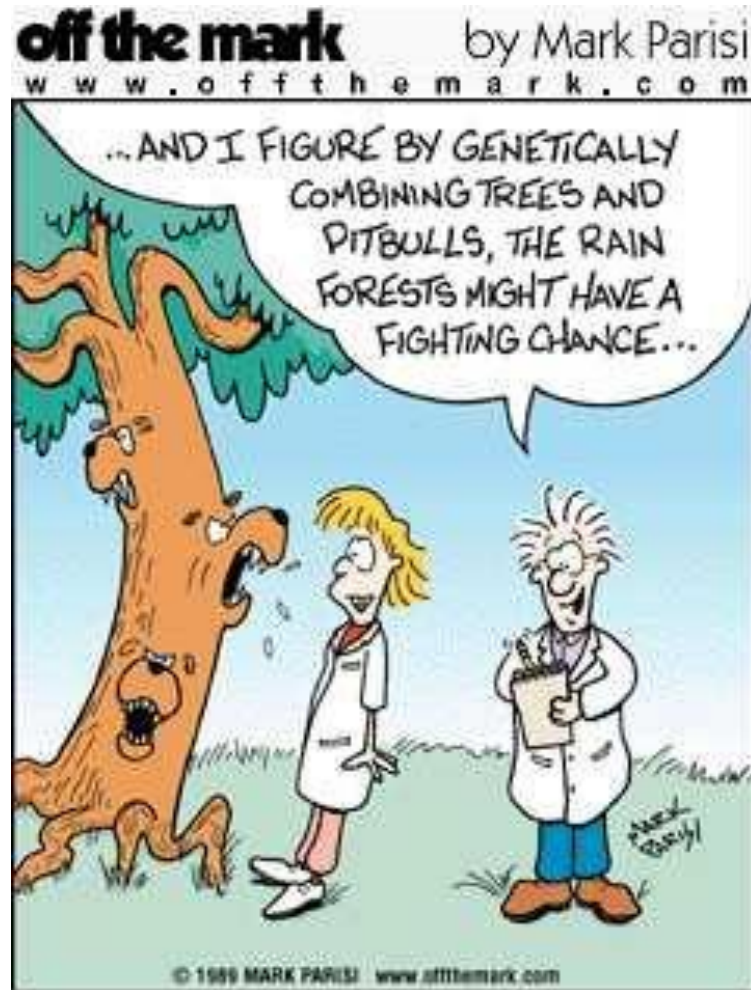


**TABLE 5.4 The genetic code**

First position (5' end)	Second position				Third position (3' end)
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

*Note:* This table identifies the amino acid encoded by each triplet. For example, the codon 5' AUG 3' on mRNA specifies methionine, whereas CAU specifies histidine. UAA, UAG, and UGA are termination signals. AUG is part of the initiation signal, in addition to coding for internal methionine residues.

# Recombinant DNA Technology



# Definitions

**Recombinant DNA**, a DNA construct created by fusing different fragments of DNA

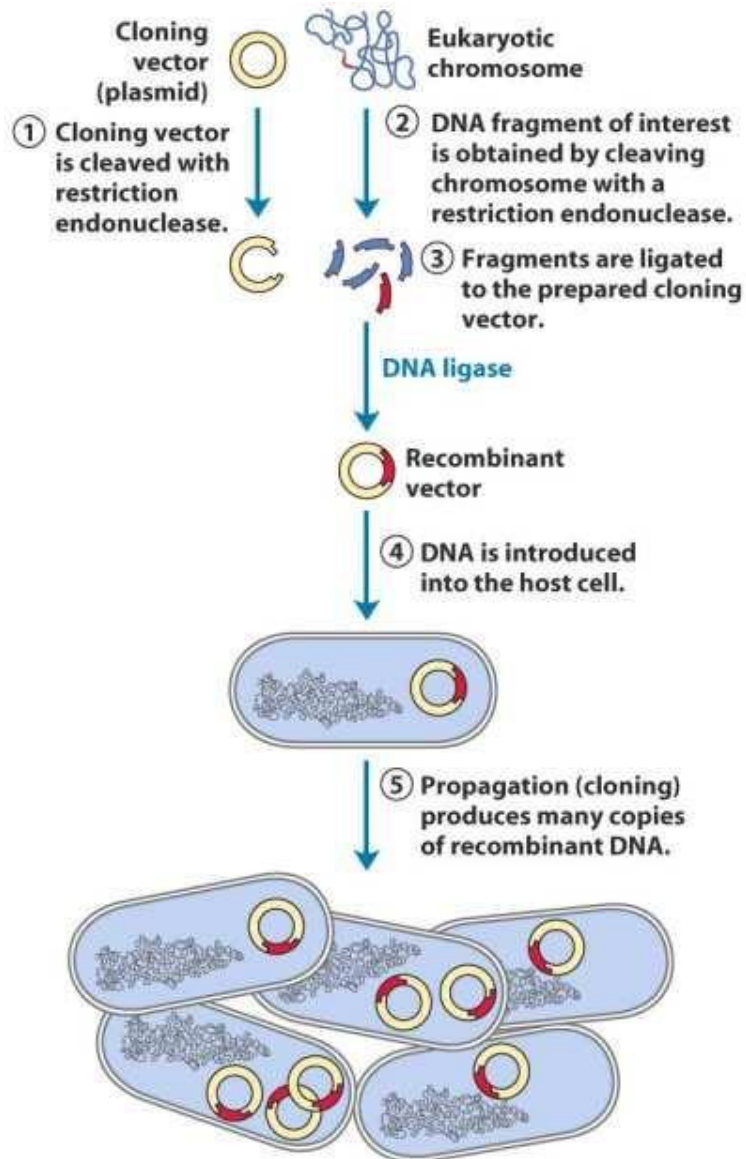
**Genetic Engineering**, the deliberate alteration of DNA through the creation of recombinant DNA

**Genetically Modified Organism**, a living entity modified through genetic engineering

**Transgenic**, a genetically modified organism containing DNA from another source



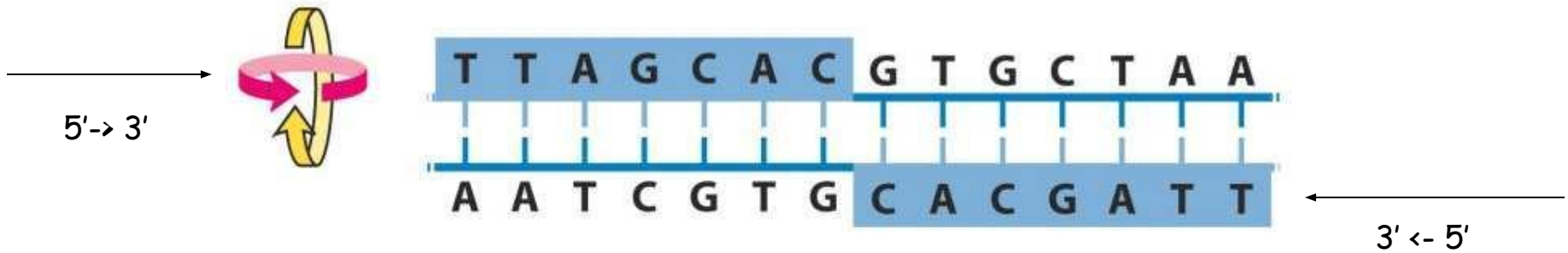
# Recombinant DNA Technology



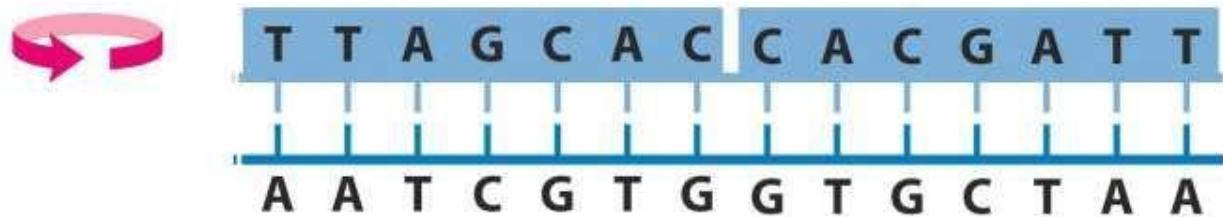
**Clones** -> Cells or organisms with identical DNA

# Restriction endonucleases

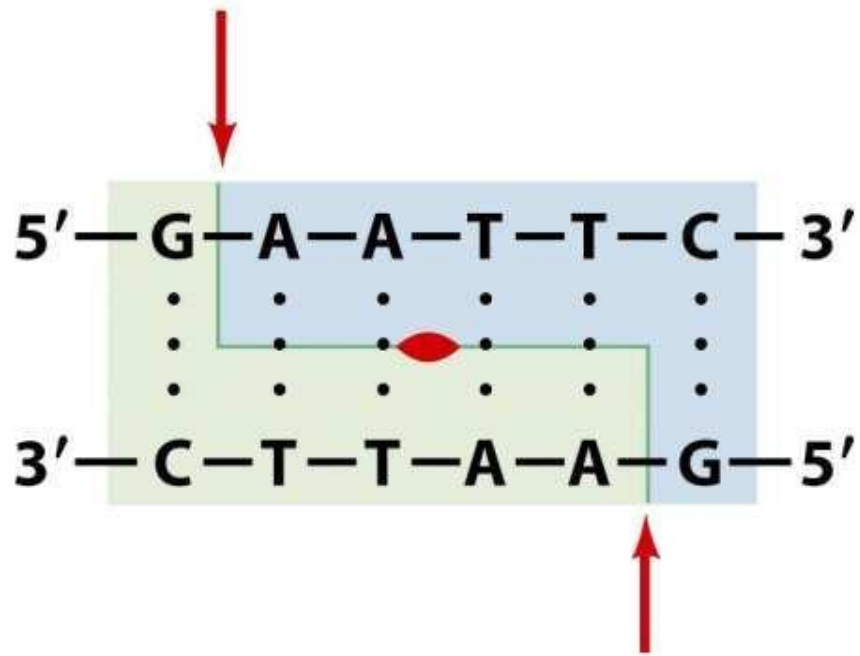
## Palindrome



## Mirror repeat

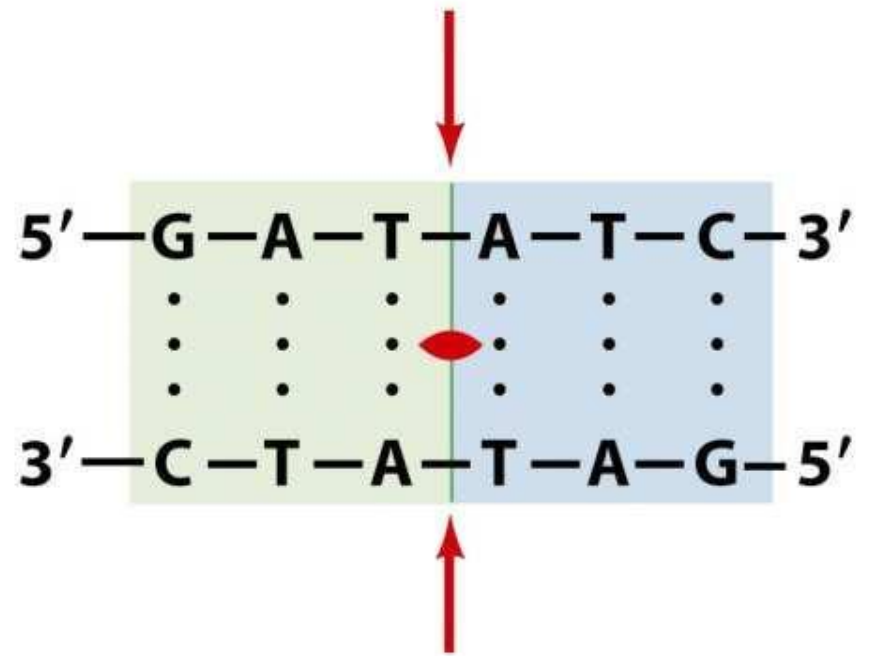


**(a) *EcoRI***



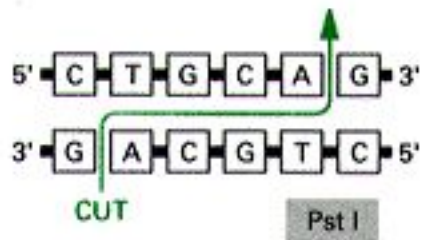
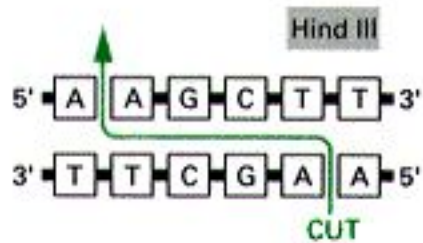
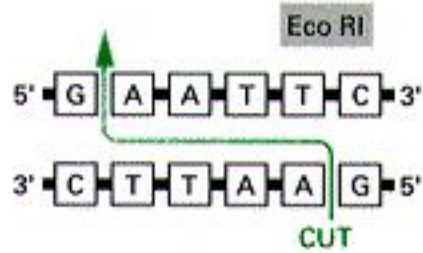
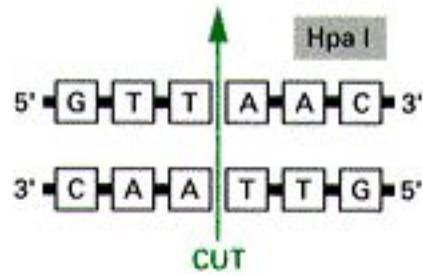
 **Cleavage site**

**(b) *EcoRV***



 **Twofold symmetry axis**

Figure 3-16 Fundamentals of Biochemistry, 2/e  
© 2006 John Wiley & Sons





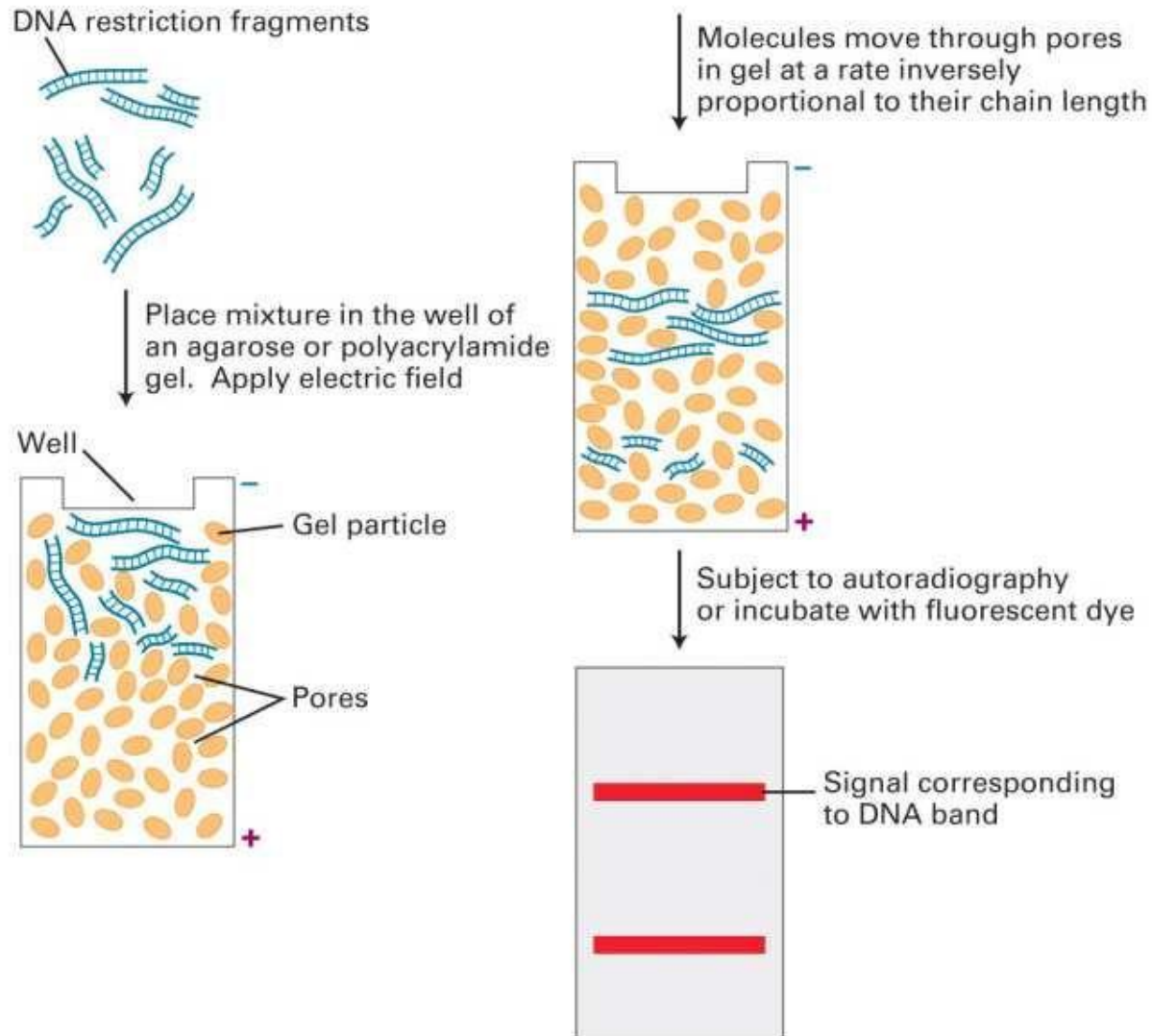
**Table 3-2** Recognition and Cleavage Sites of Some Restriction Enzymes

Enzyme	Recognition Sequence <sup>a</sup>	Microorganism
<i>Alu</i> I	AG↓CT	<i>Arthrobacter luteus</i>
<i>Bam</i> HI	G↓GATCC	<i>Bacillus amyloliquefaciens</i> H
<i>Bgl</i> I	GCCNNNNN↓NGGC	<i>Bacillus globigii</i>
<i>Bgl</i> II	A↓GATCT	<i>Bacillus globigii</i>
<i>Eco</i> RI	G↓AATTC	<i>Escherichia coli</i> RY13
<i>Eco</i> RII	↓CC(↓)GG	<i>Escherichia coli</i> R245
<i>Eco</i> RV	GAT↓ATC	<i>Escherichia coli</i> J62 pLG74
<i>Hae</i> II	RGCGC↓Y	<i>Haemophilus aegyptius</i>
<i>Hae</i> III	GG↓CC	<i>Haemophilus aegyptius</i>
<i>Hind</i> III	A↓AGCTT	<i>Haemophilus influenzae</i> R <sub>d</sub>
<i>Hpa</i> II	C↓CGG	<i>Haemophilus parainfluenzae</i>
<i>Msp</i> I	C↓CGG	<i>Moraxella</i> species
<i>Pst</i> I	CTGCA↓G	<i>Providencia stuartii</i> 164
<i>Pvu</i> II	CAG↓CTG	<i>Proteus vulgaris</i>
<i>Sal</i> I	G↓TCGAC	<i>Streptomyces albus</i> G
<i>Taq</i> I	T↓CGA	<i>Thermus aquaticus</i>
<i>Xho</i> I	C↓TCGAG	<i>Xanthomonas holcicola</i>

<sup>a</sup>The recognition sequence is abbreviated so that only one strand, reading 5' to 3', is given. The cleavage site is represented by an arrow (↓). R, Y, and N represent a purine nucleotide, a pyrimidine nucleotide, and any nucleotide, respectively.

Source: Roberts, R.J. and Macelis, D., REBASE—the restriction enzyme database, <http://rebase.neb.com>.

# Gel Electrophoresis



# Gel Electrophoresis

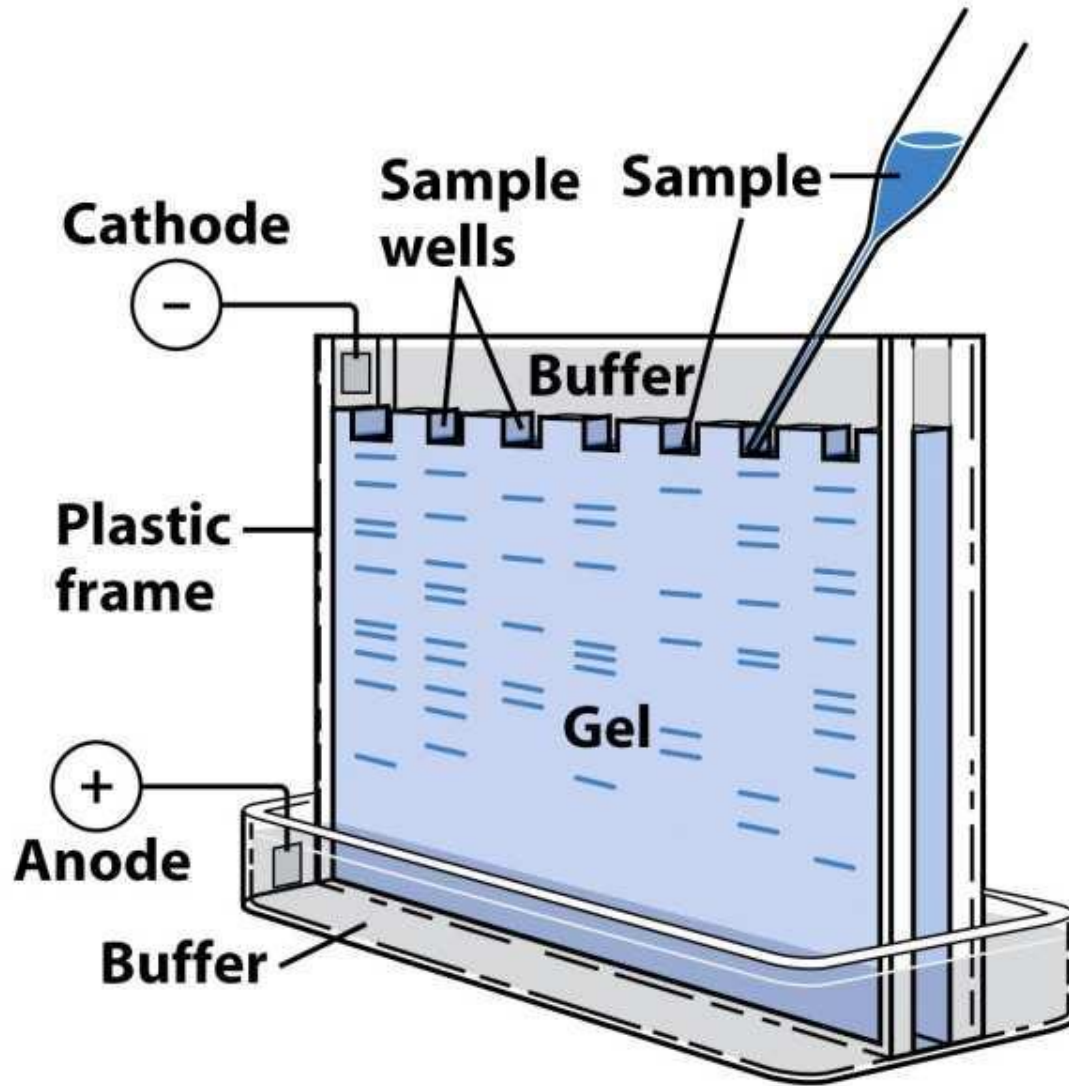
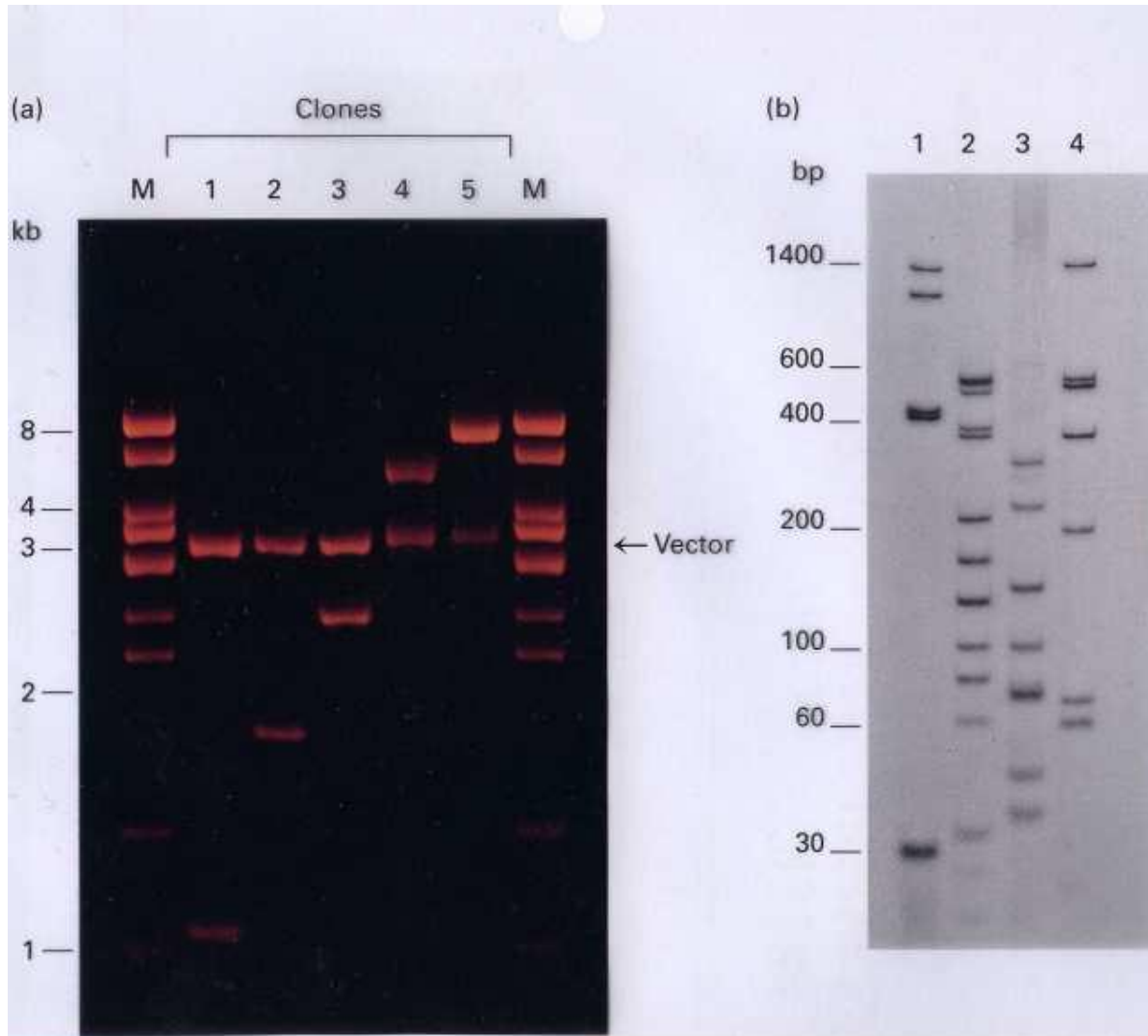
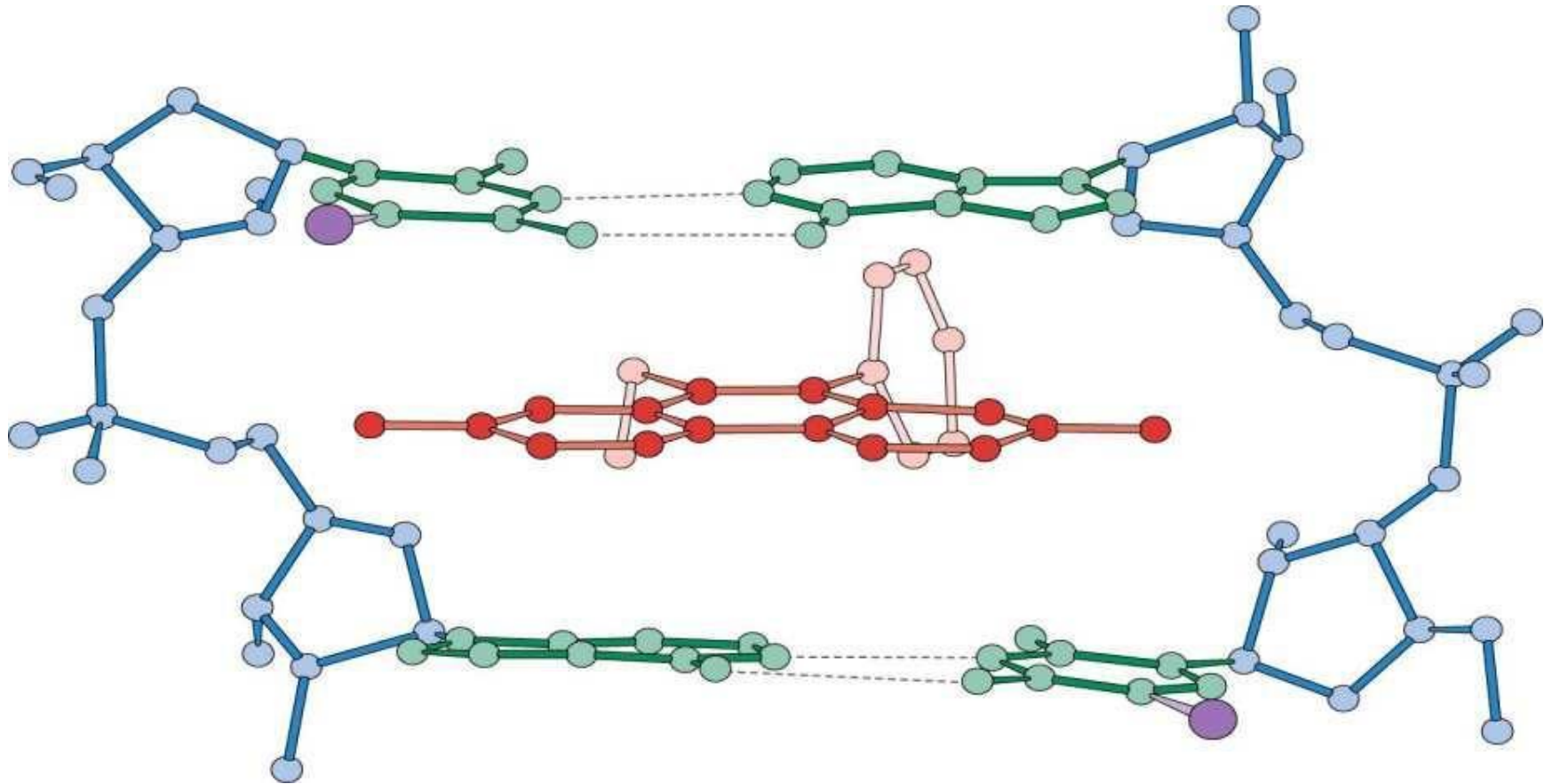


Figure 3-17 Fundamentals of Biochemistry, 2/e  
© 2006 John Wiley & Sons

# Gel Electrophoresis

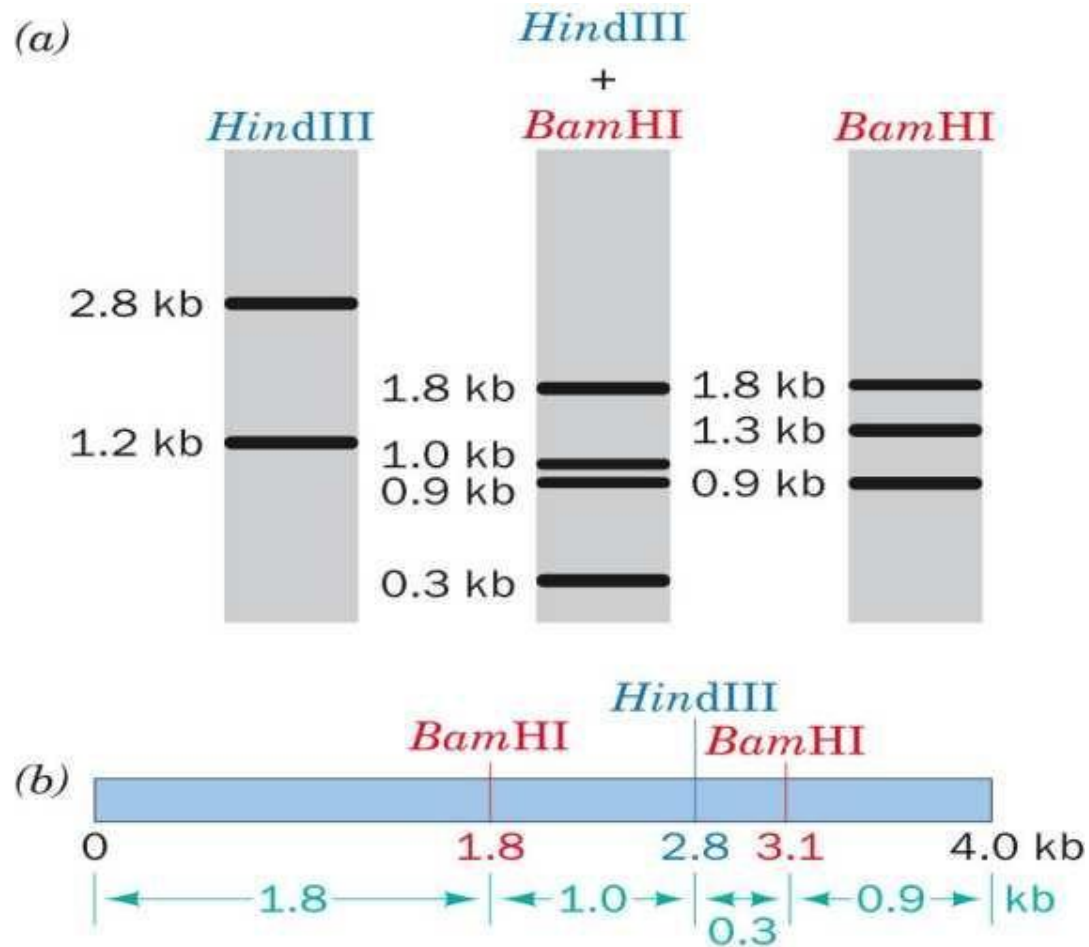


# X-Ray structure of a complex of ethidium bromide with DNA.

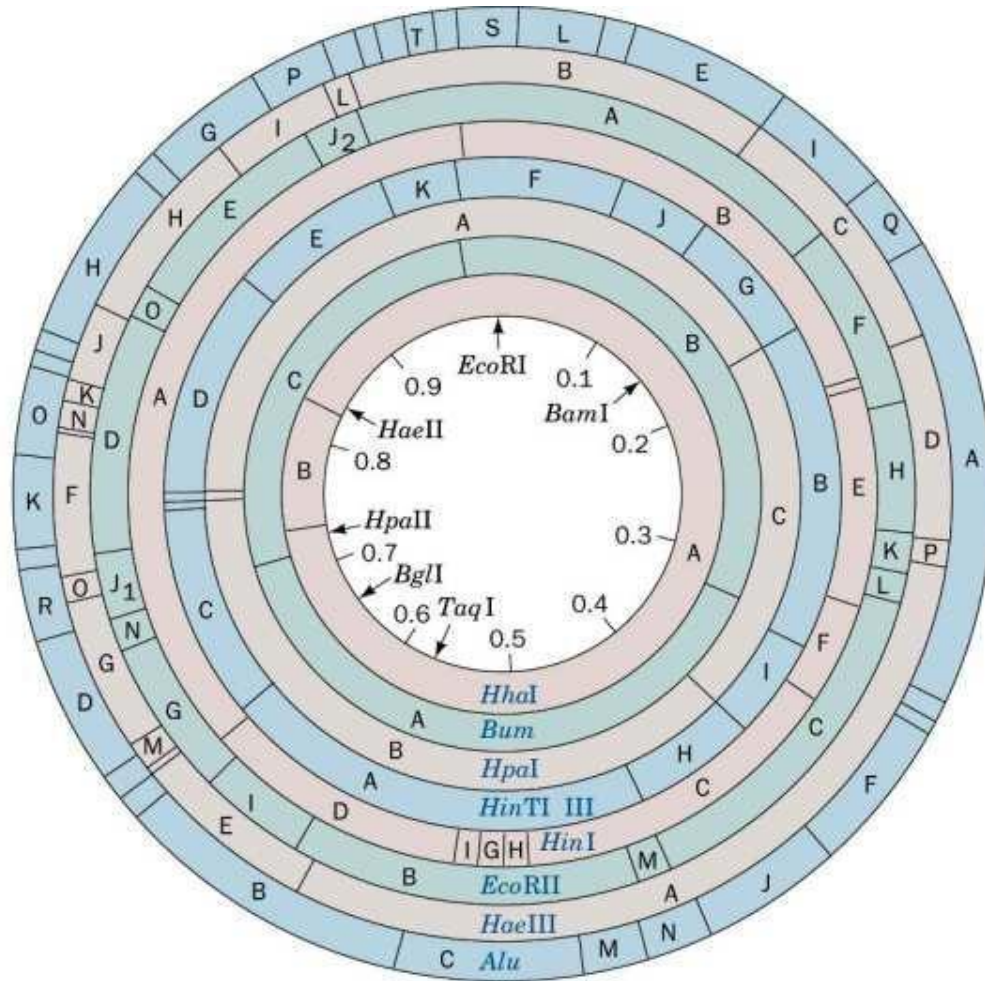


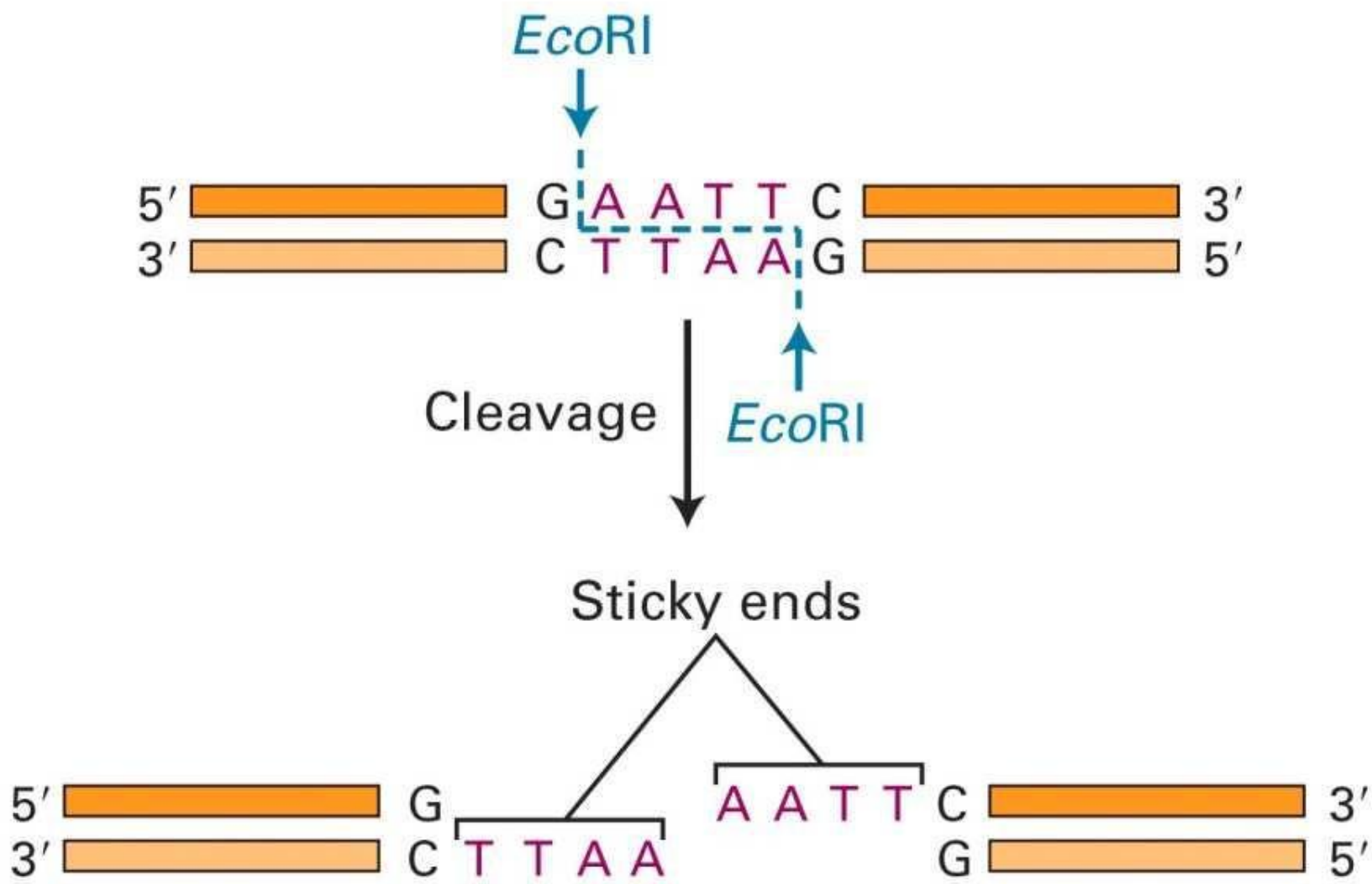


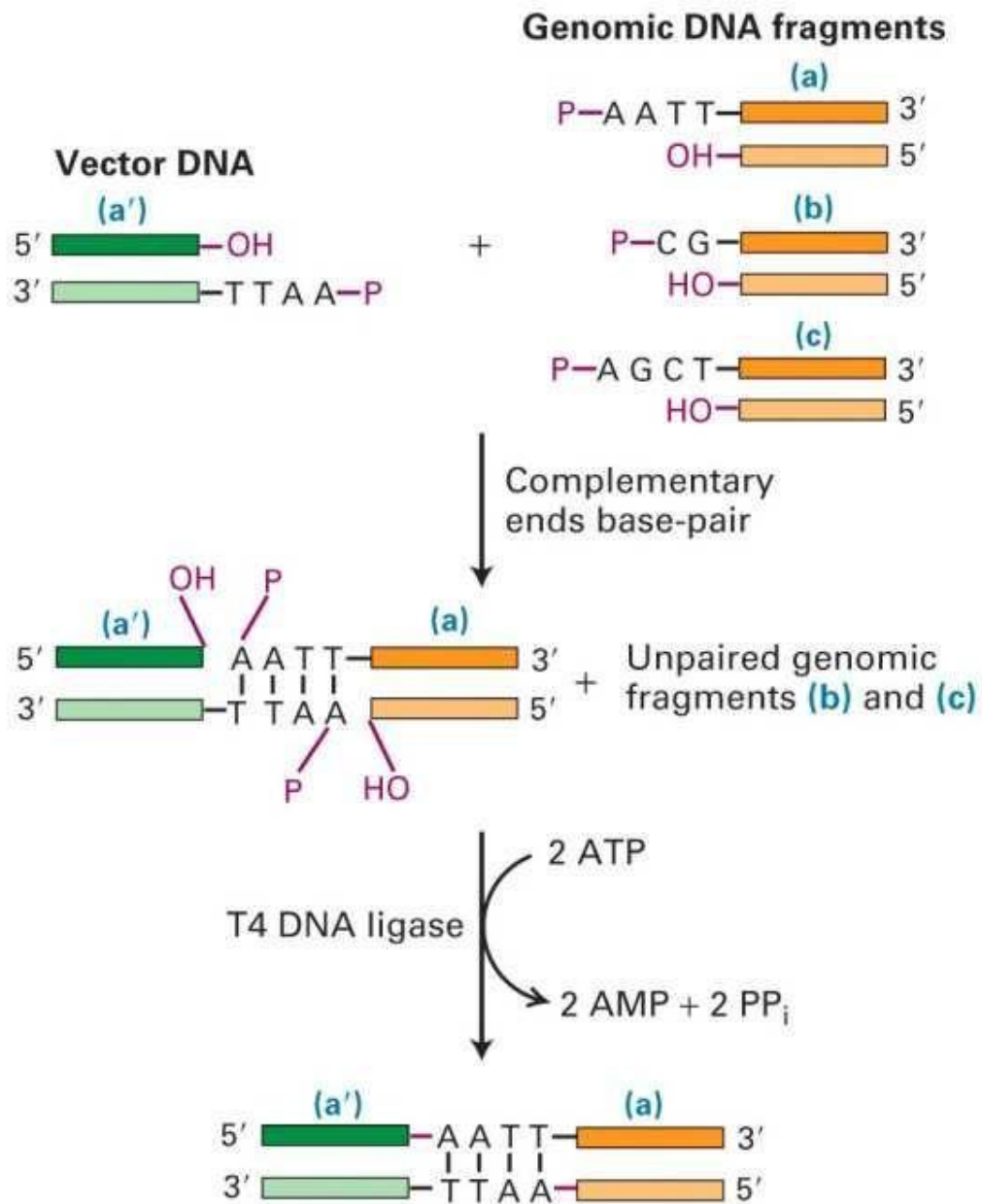
# Construction of a restriction map.



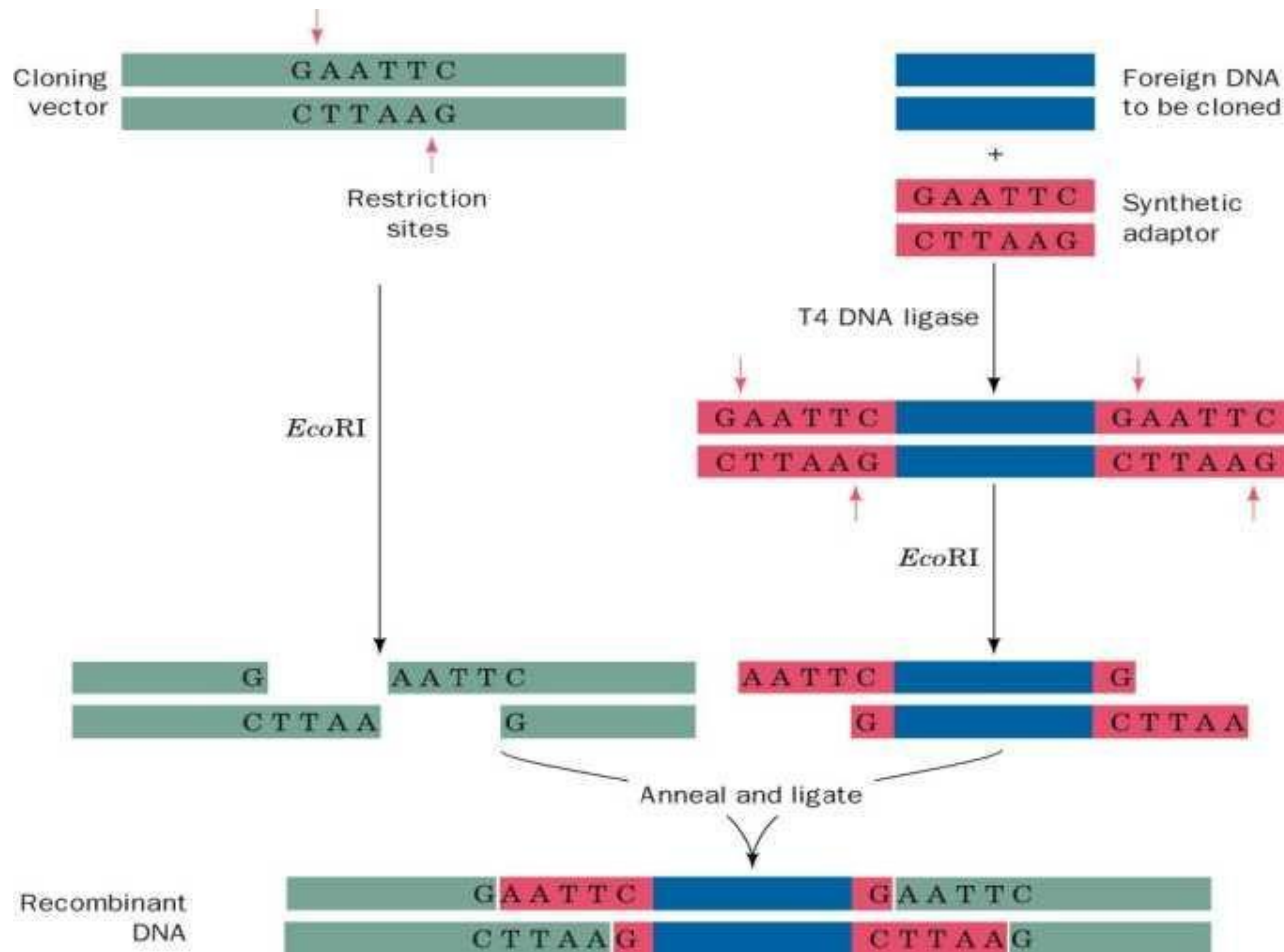
# Restriction map for the 5243-bp circular DNA of SV40.



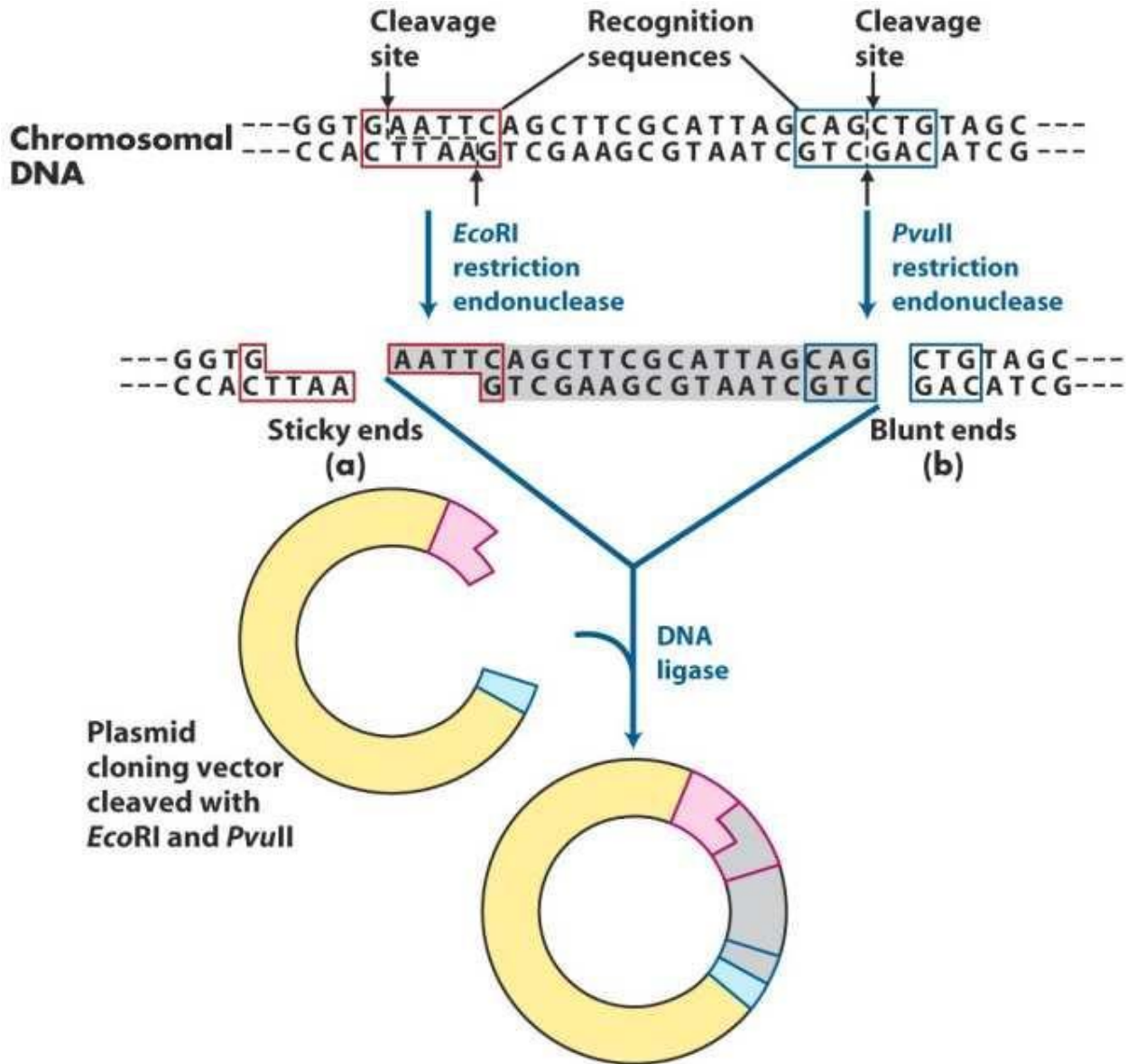




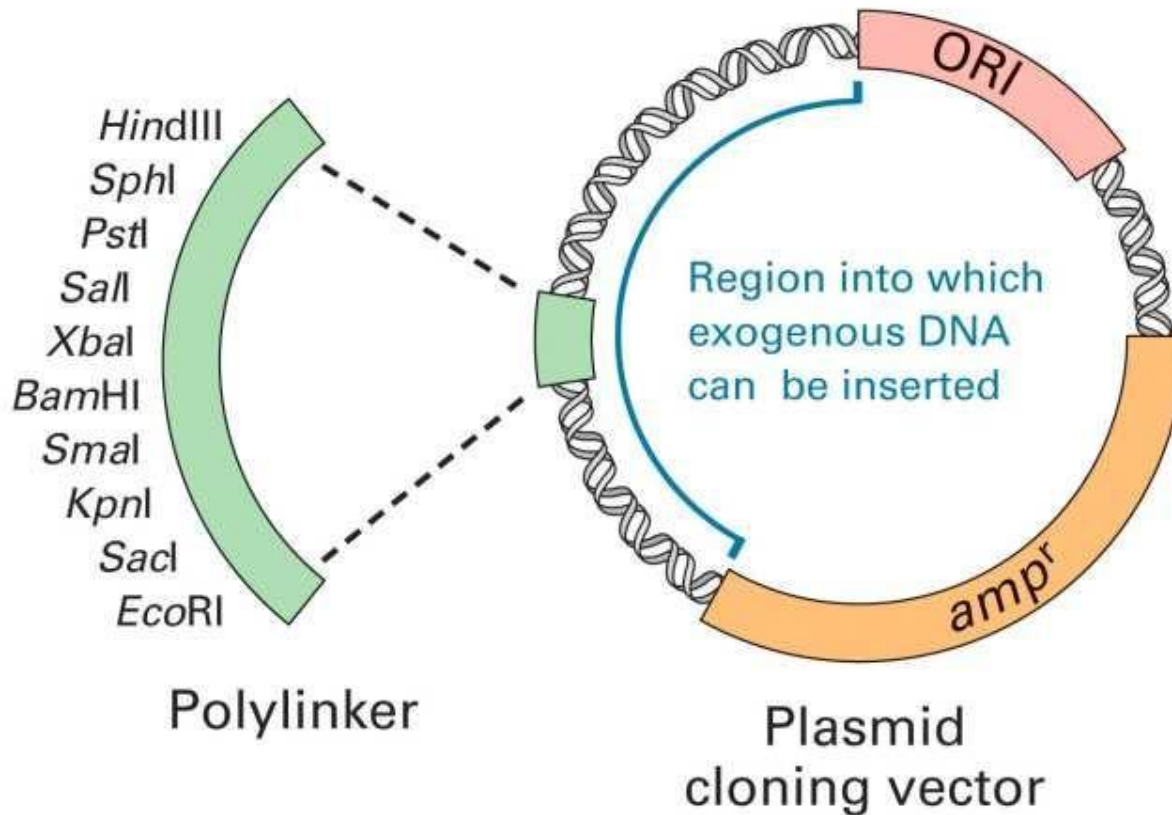
# Construction of a recombinant DNA molecule through the use of synthetic oligonucleotide adaptors







# Plasmid Cloning Vectors



# Plasmid Cloning Vectors

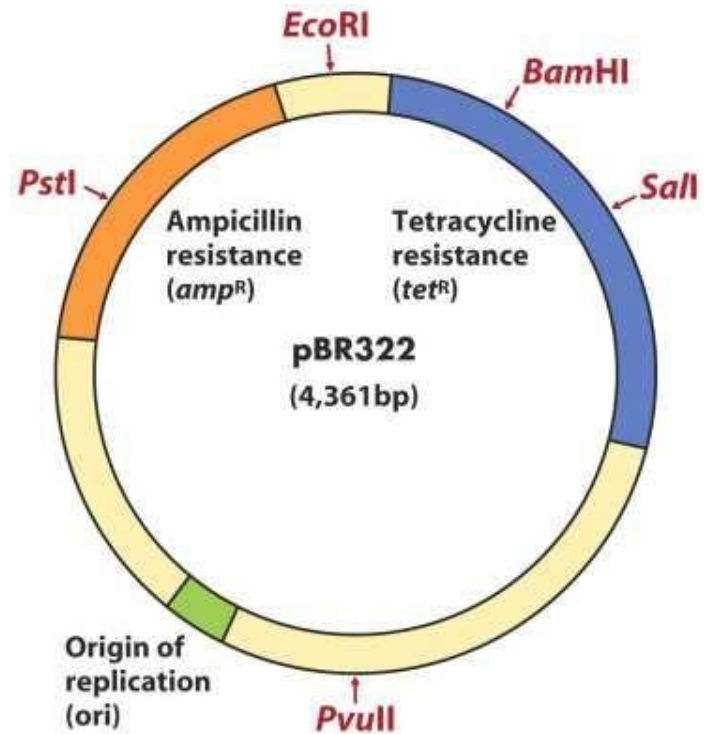
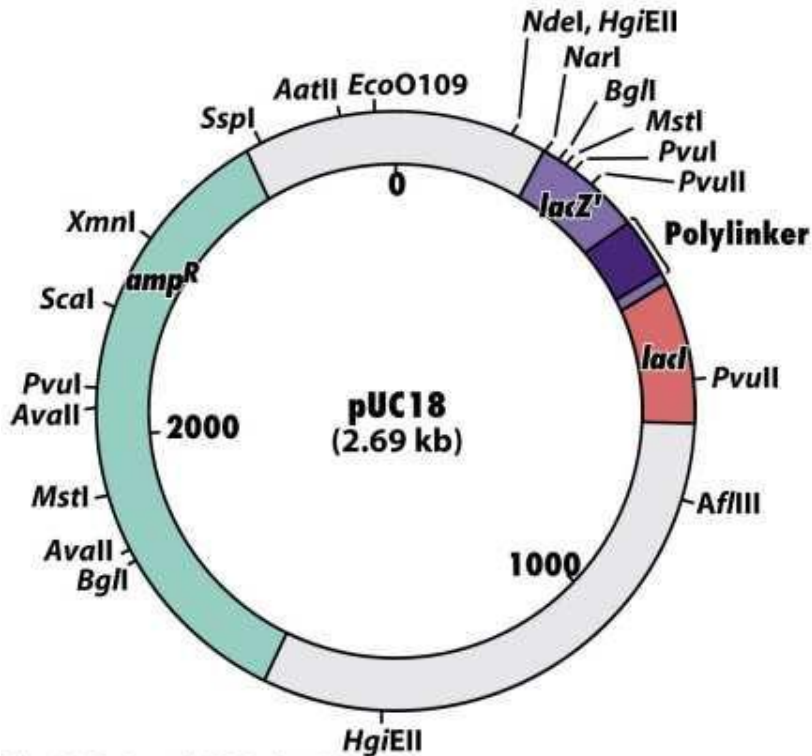
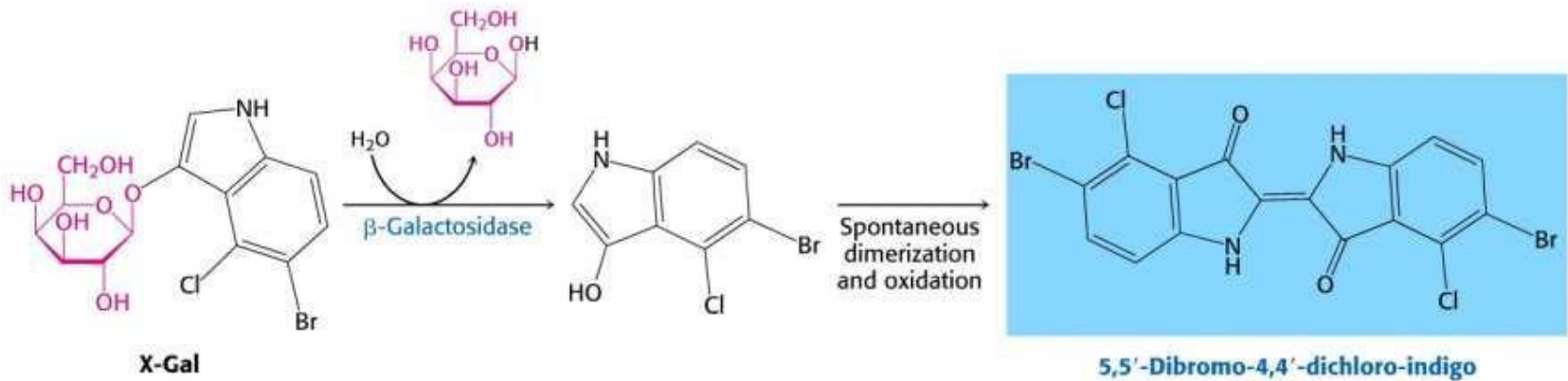
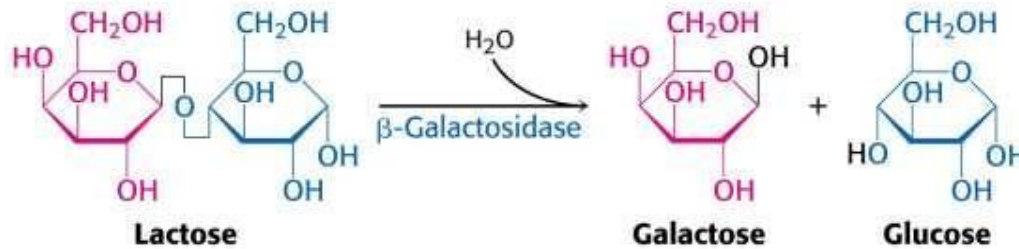


Figure 3-25 Fundamentals of Biochemistry, 2/e  
© 2006 John Wiley & Sons

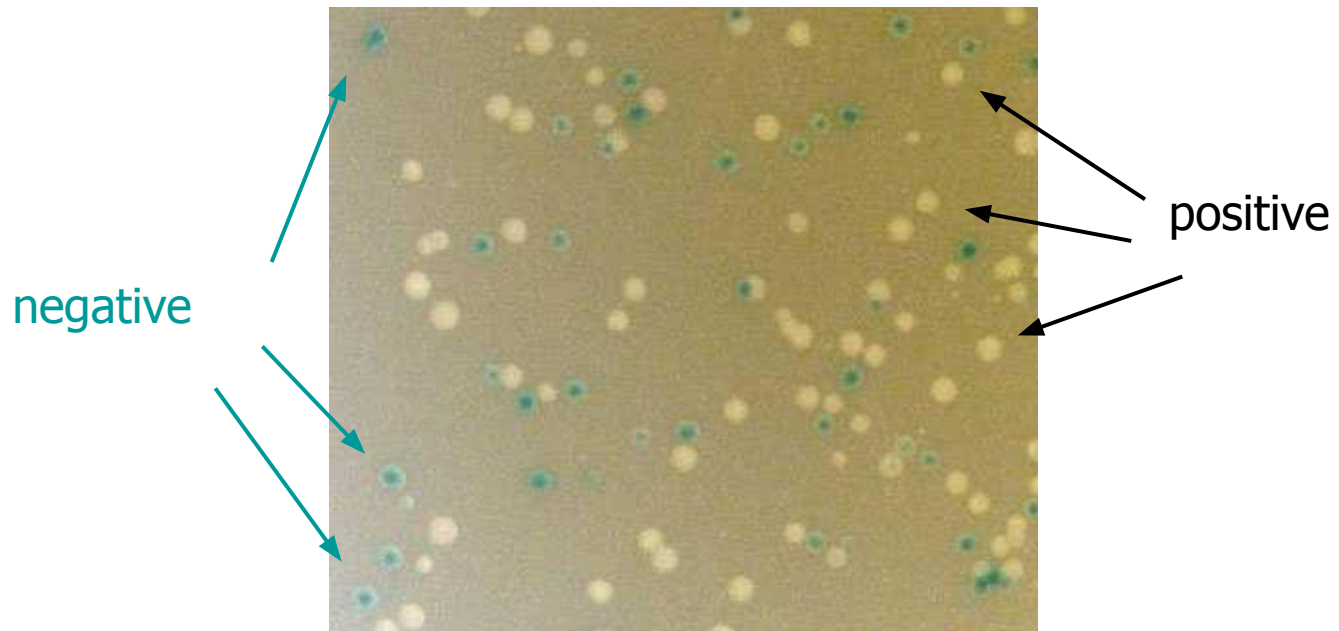
# Insertional inactivation

## Gene in cloning site:

- *LacZ* → pUC18 (*lacZ* complements the host defect in *lacZ*)
  - pUC18 into host organism → active *lacZ* ( $\beta$ -galactosidase) from plasmid → cleavage of X-gal (blue colonies)
  - gene cloned into polylinker → *lacZ* gene disrupted → no cleavage of X-gal (white colonies)



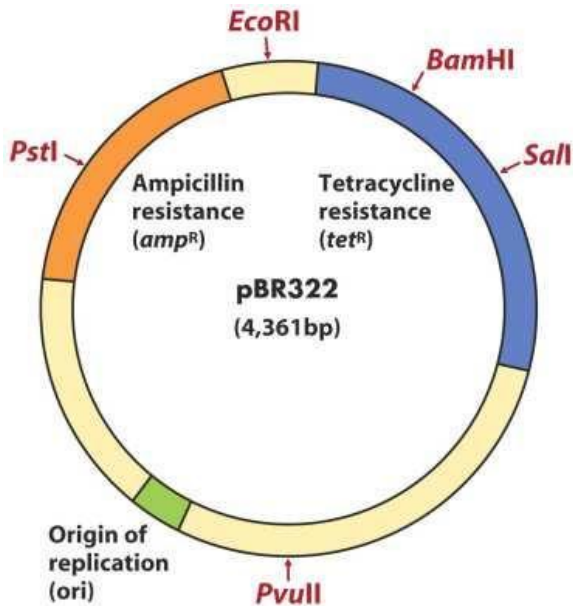
# Blue/White Selection



# Insertional inactivation

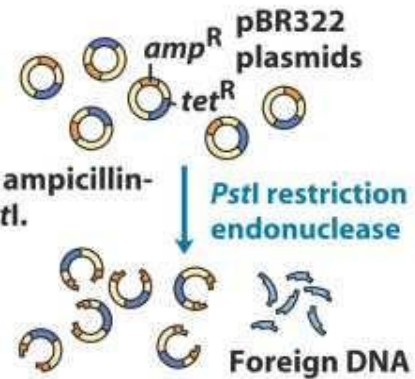
## Gene in cloning site:

- Resistance marker -> pBR322 (cloning sites within antibiotic resistance marker)
  - > plasmid into host -> resistance against 2 antibiotics
  - > gene cloned within one resistance marker -> gene for antibiotic resistance marker disrupted -> sensitive against one antibiotic



①

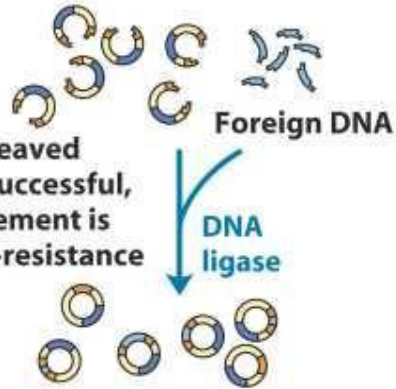
pBR322 is cleaved at the ampicillin-resistance element by *PstI*.





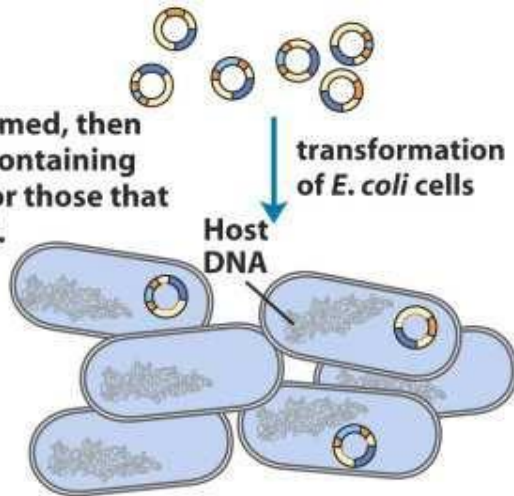
②

Foreign DNA is ligated to cleaved pBR322. Where ligation is successful, the ampicillin-resistance element is disrupted. The tetracycline-resistance element remains intact.



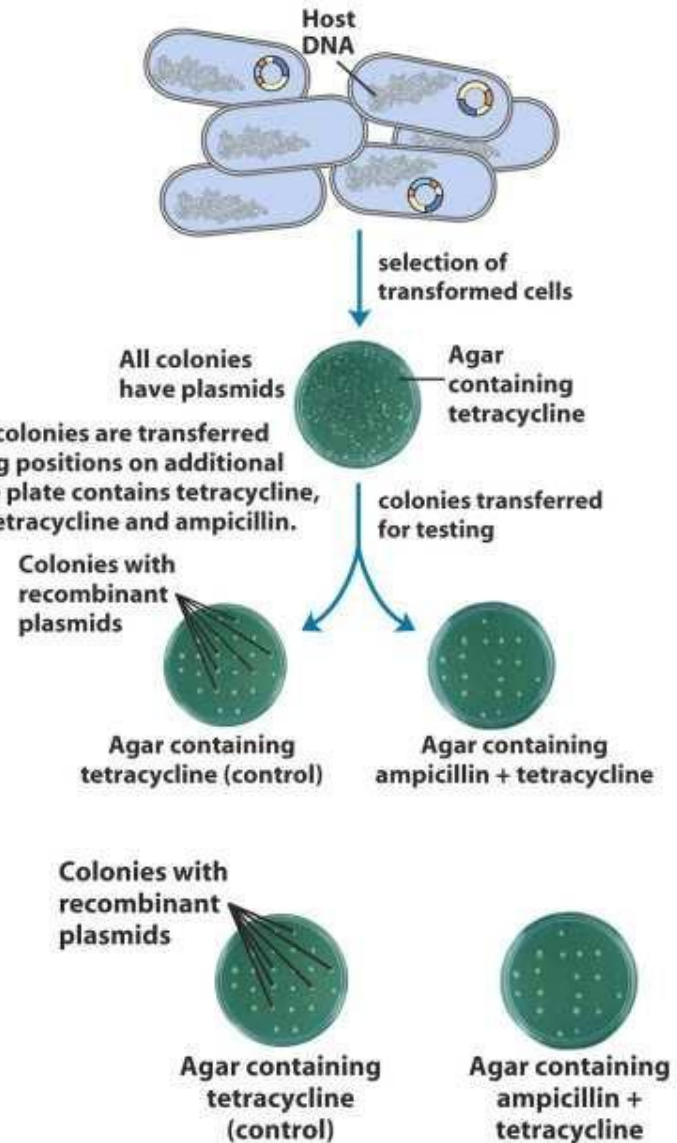
③

*E. coli* cells are transformed, then grown on agar plates containing tetracycline to select for those that have taken up plasmid.



④

Individual colonies are transferred to matching positions on additional plates. One plate contains tetracycline, the other tetracycline and ampicillin.

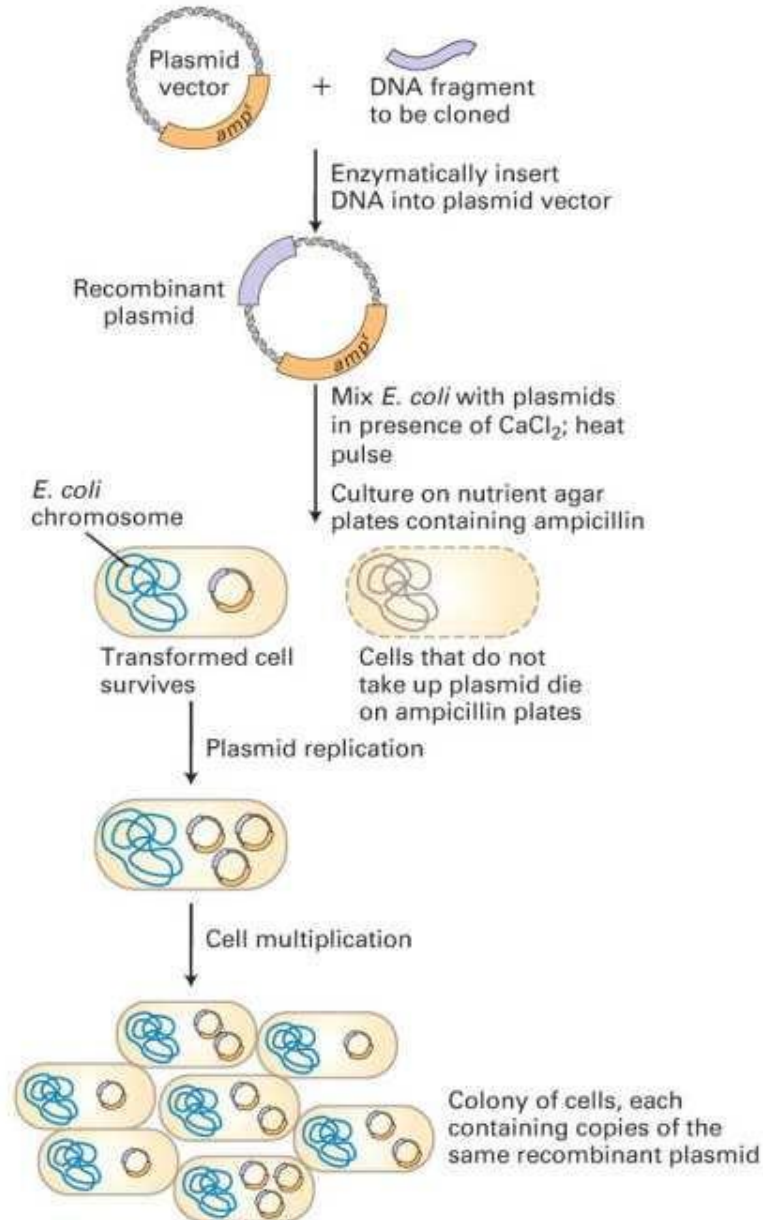


⑤

Cells that grow on tetracycline but not on tetracycline + ampicillin contain recombinant plasmids with disrupted ampicillin resistance, hence the foreign DNA. Cells with pBR322 without foreign DNA retain ampicillin resistance and grow on both plates.



# Transformation and Selection



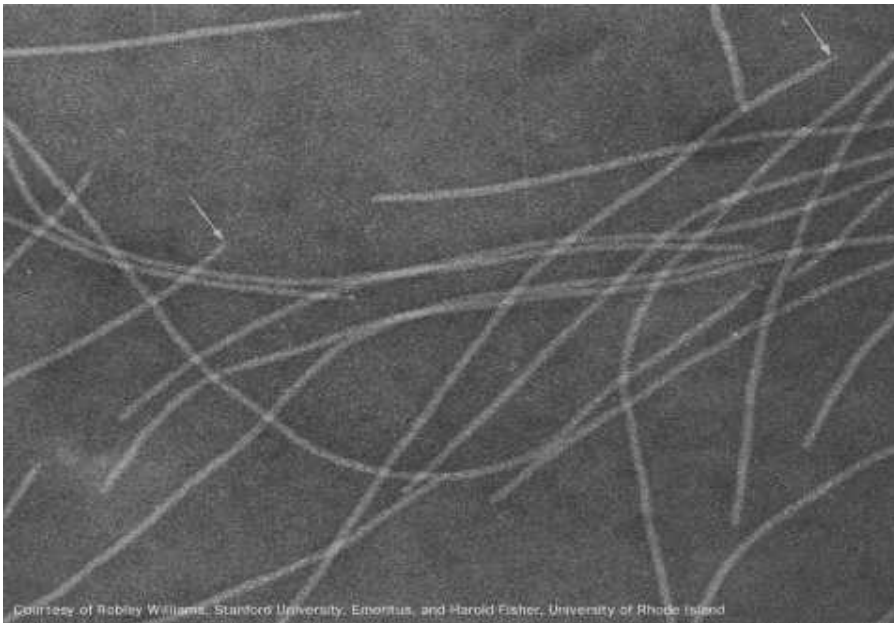
# Horizontal gene transfer

- Transformation -> uptake of naked DNA (chemical transformation, electroporation)
- Conjugation -> DNA transfer by cell - cell contact
- Transduction -> DNA transfer by bacteriophage infection

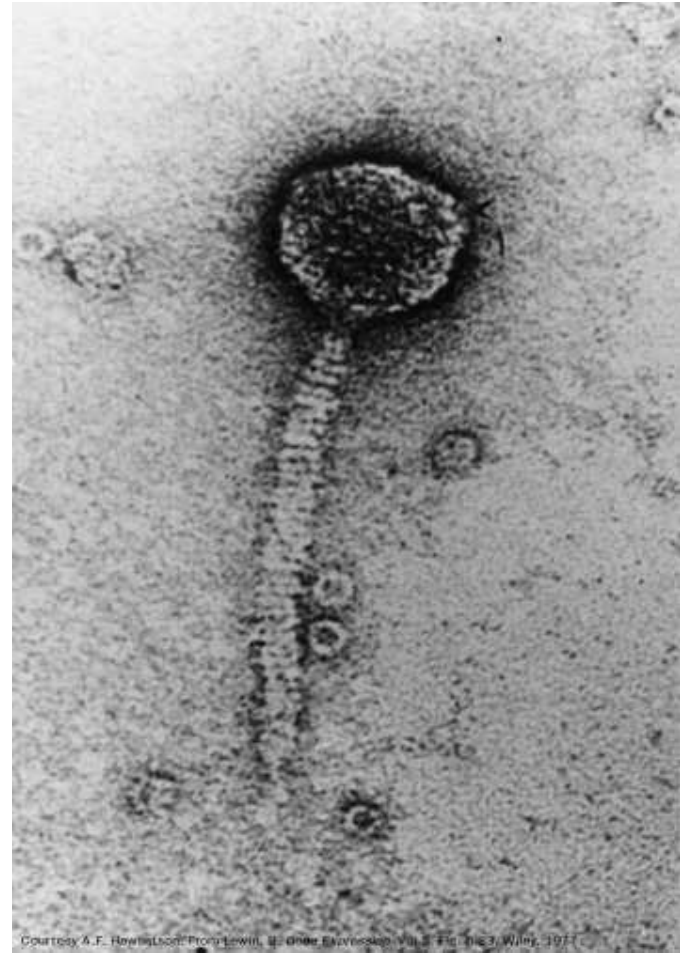
Other methods of Gene transfer -> used with fungi, animal and plant cells:

- Microinjection
- protoplasts

# Bacteriophages

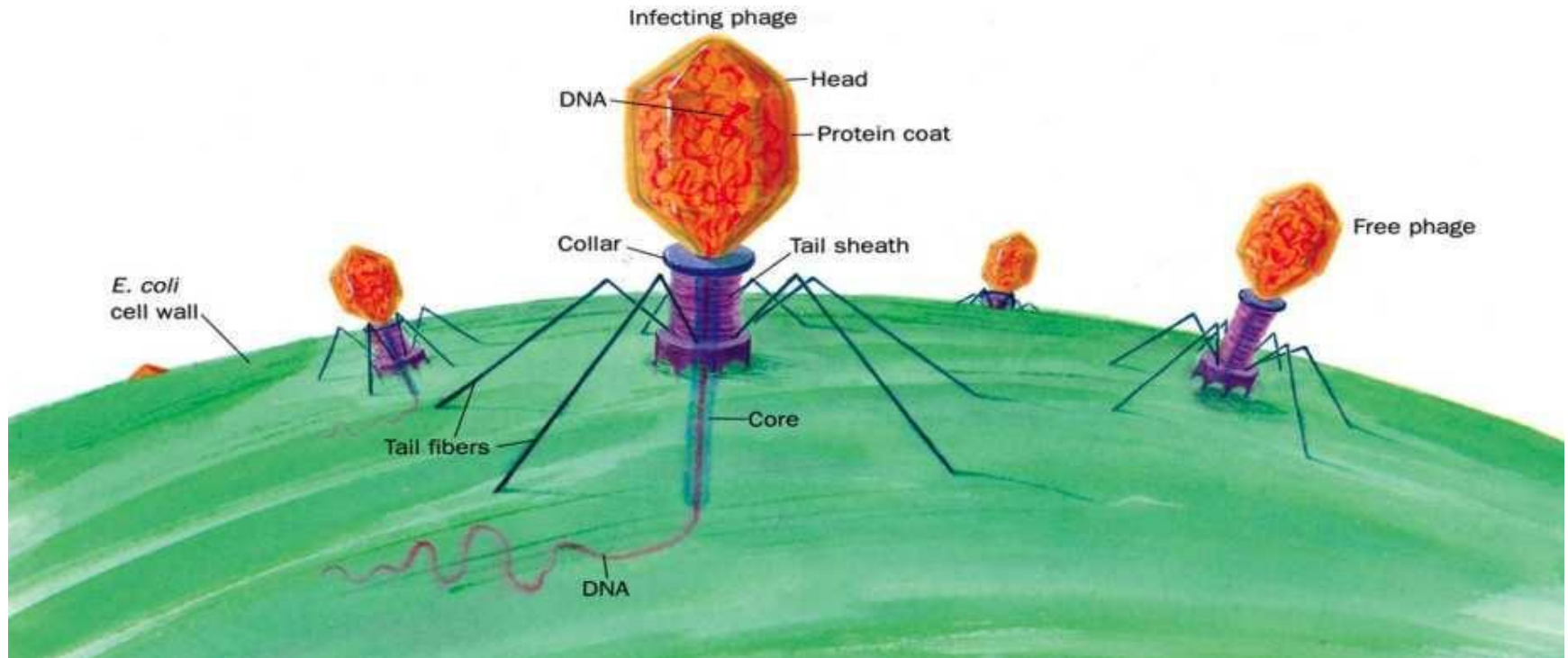


Electron micrograph of the filamentous bacteriophage M13.

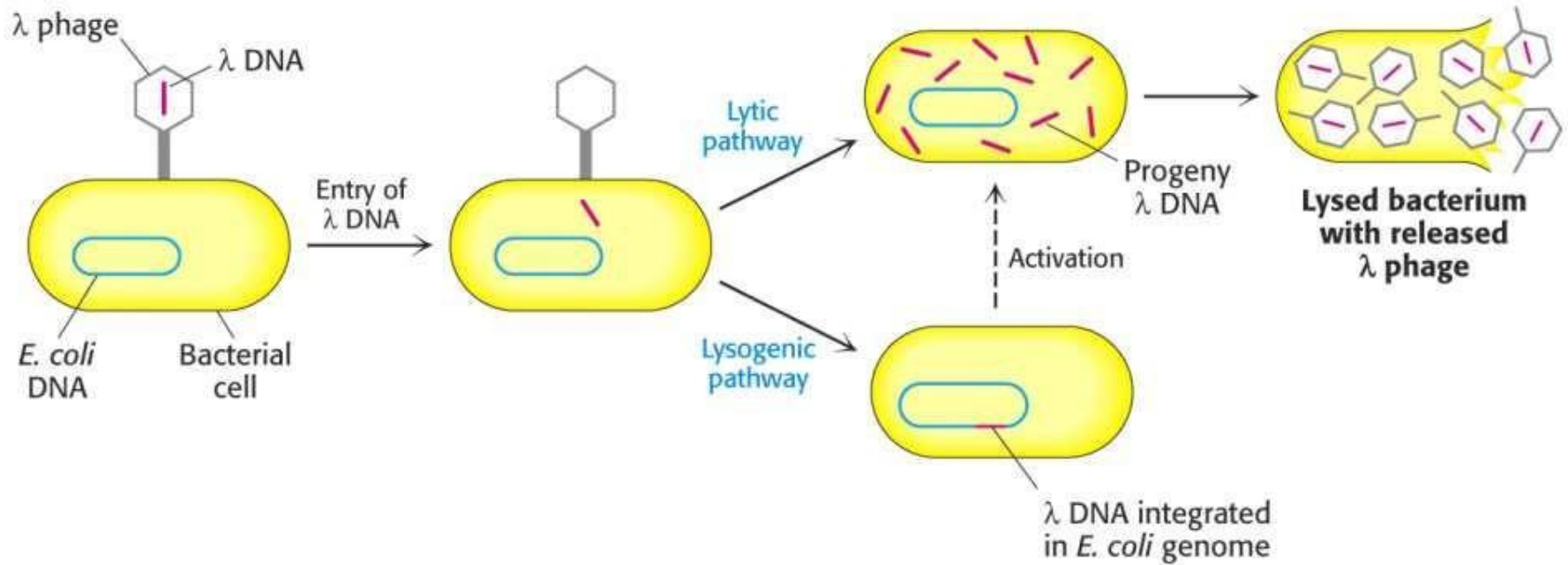


Electron micrograph of bacteriophage  $\lambda$ .

# Bacteriophage T2 injecting its DNA into an *E. coli*

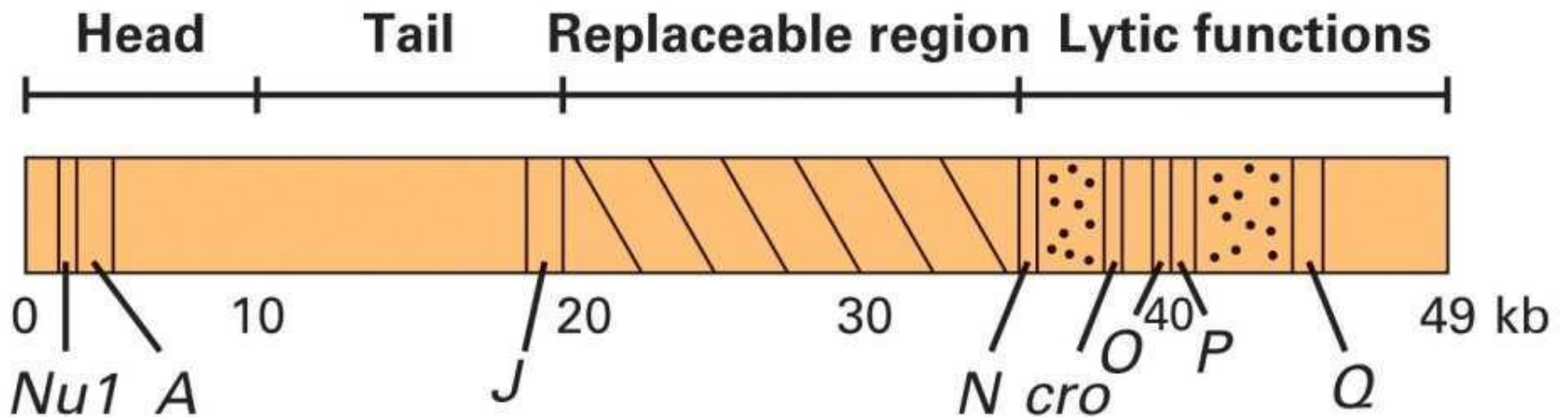


# Life Cycle of Bacteriophage

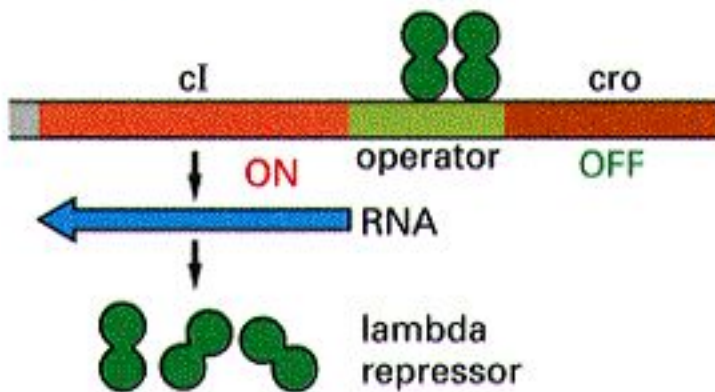




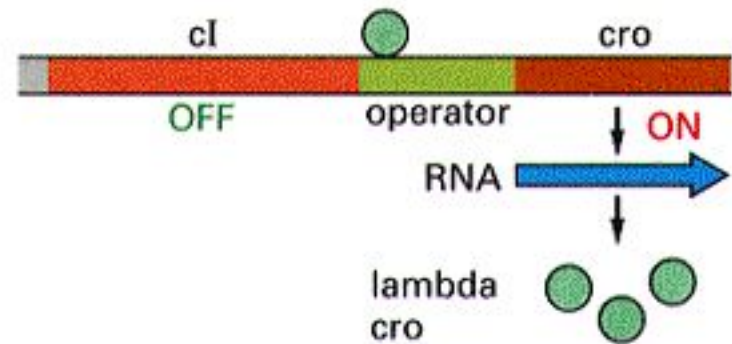
(a)  $\lambda$  Phage genome



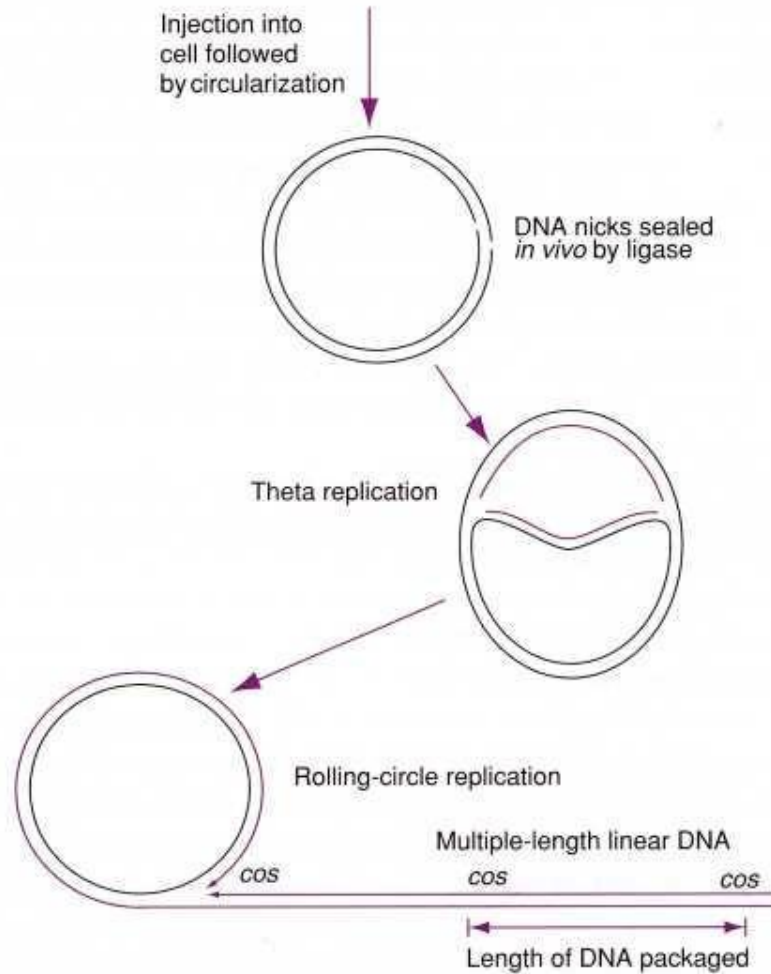
stable state 1: the prophage state  
lambda repressor protein is made



stable state 2: the lytic state  
lambda cro protein is made

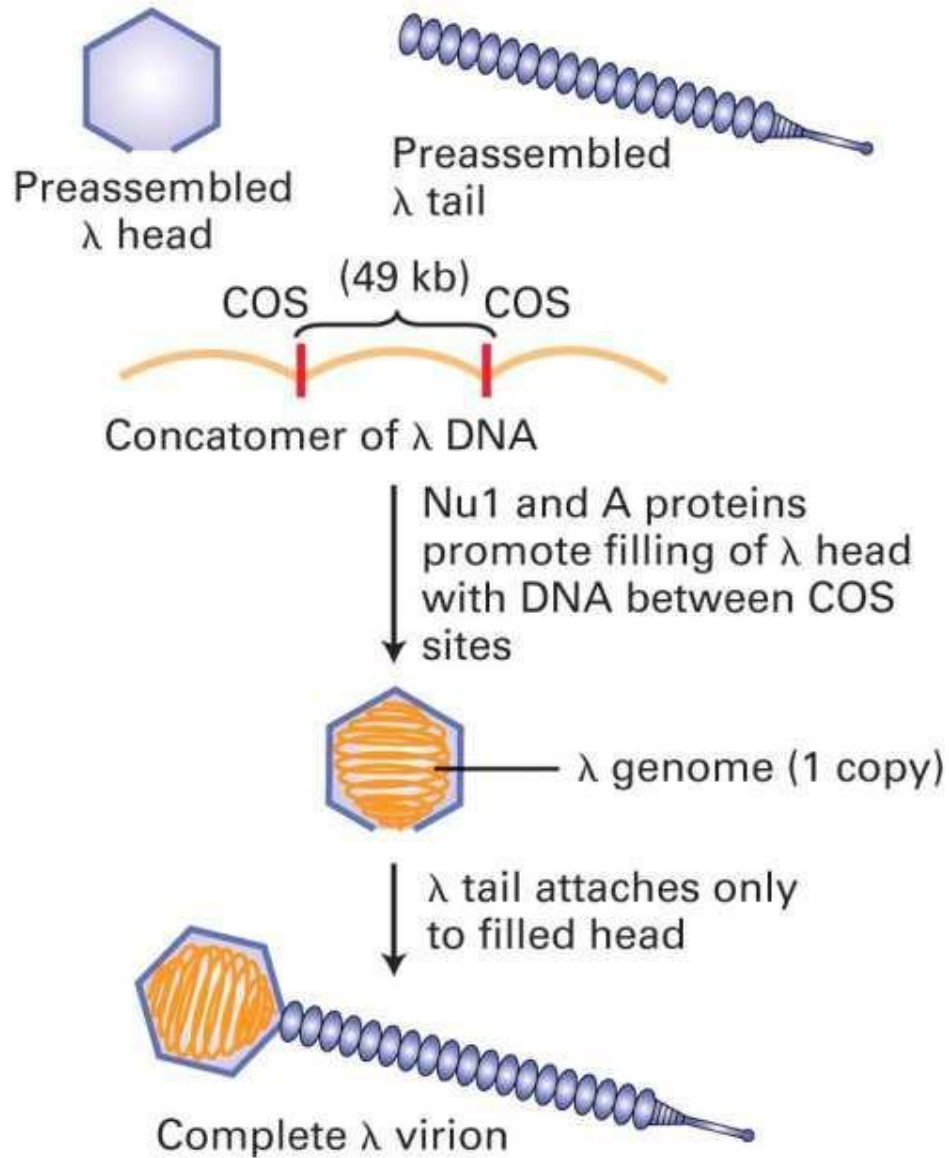


# Replication of bacteriophage upon infection of a cell



**Figure 6.6** Replication of bacteriophage lambda DNA

(b)  $\lambda$  Phage assembly



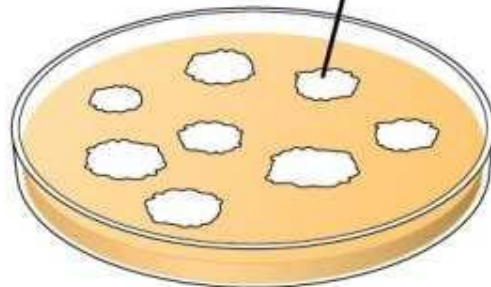
Confluent layer of susceptible host cells  
growing on surface of a plate



Add dilute suspension containing virus;  
after infection, cover layer of cells  
with agar; incubate



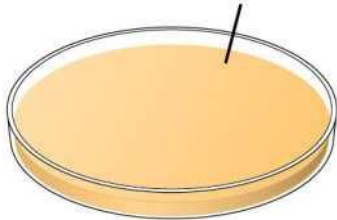
Plaque



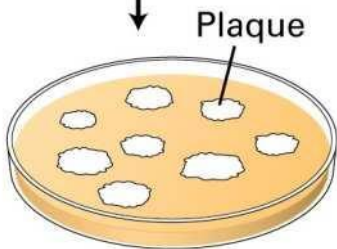
Each plaque represents cell lysis initiated by one viral  
particle (agar restricts movement so that virus can  
infect only contiguous cells)

# Molecular genetics and bacteriophage

Confluent layer of susceptible host cells growing on surface of a plate

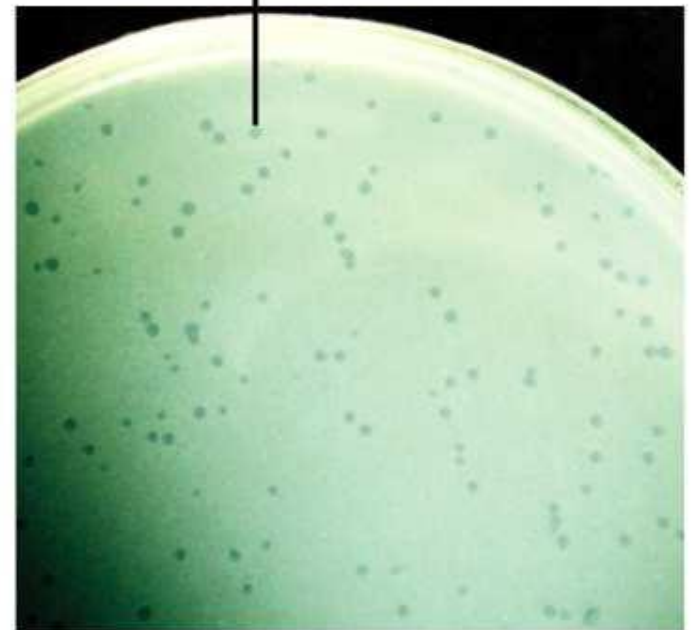


Add dilute suspension containing virus; after infection, cover layer of cells with agar; incubate



Each plaque represents cell lysis initiated by one viral particle (agar restricts movement so that virus can infect only contiguous cells)

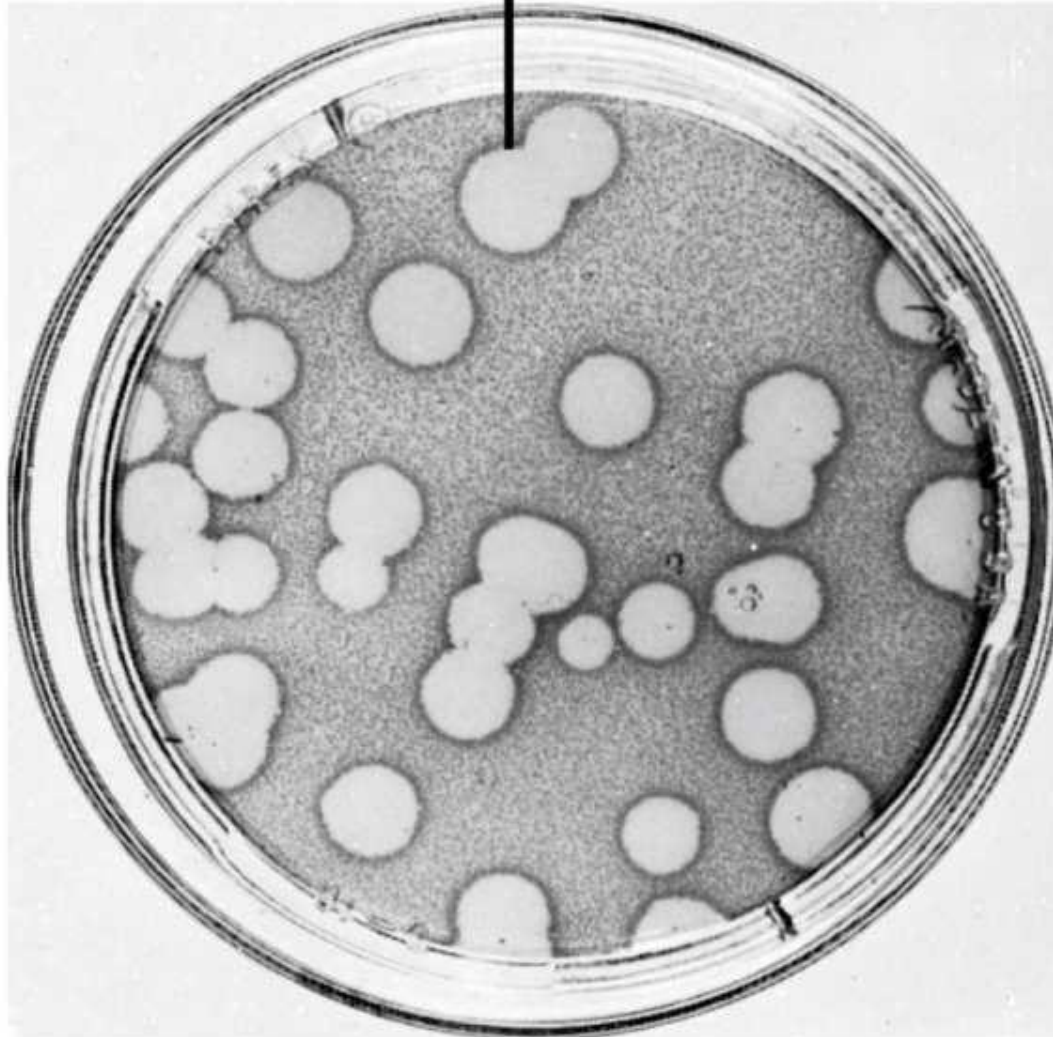
(b) Plaque



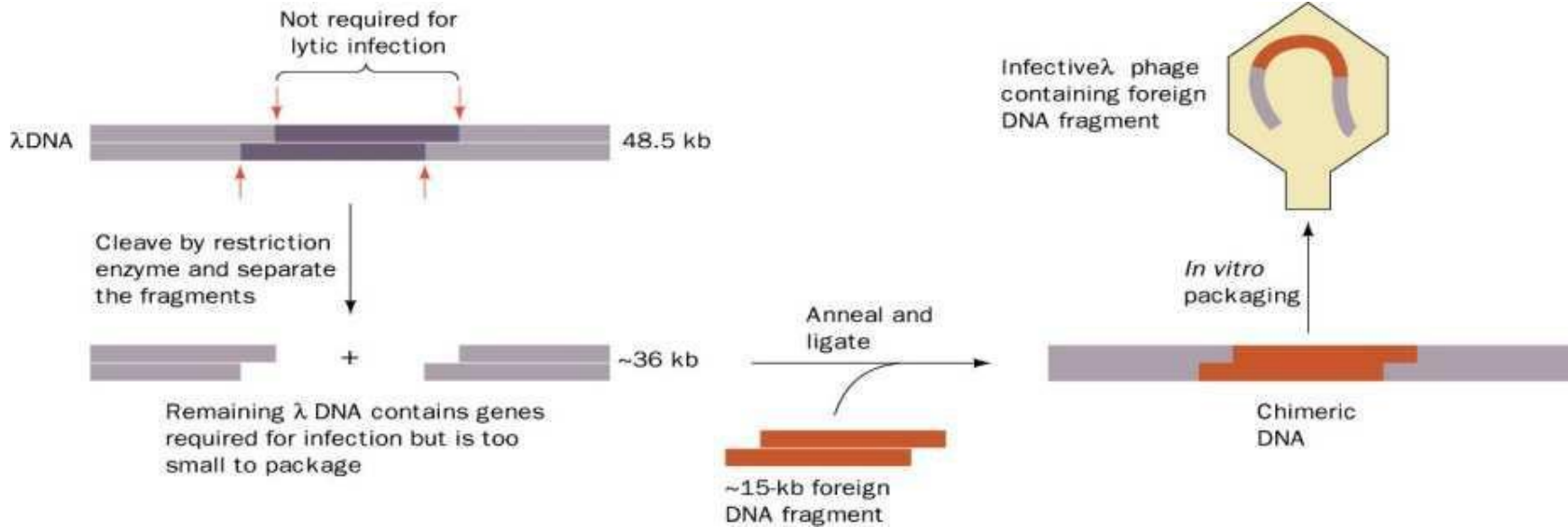


(c)

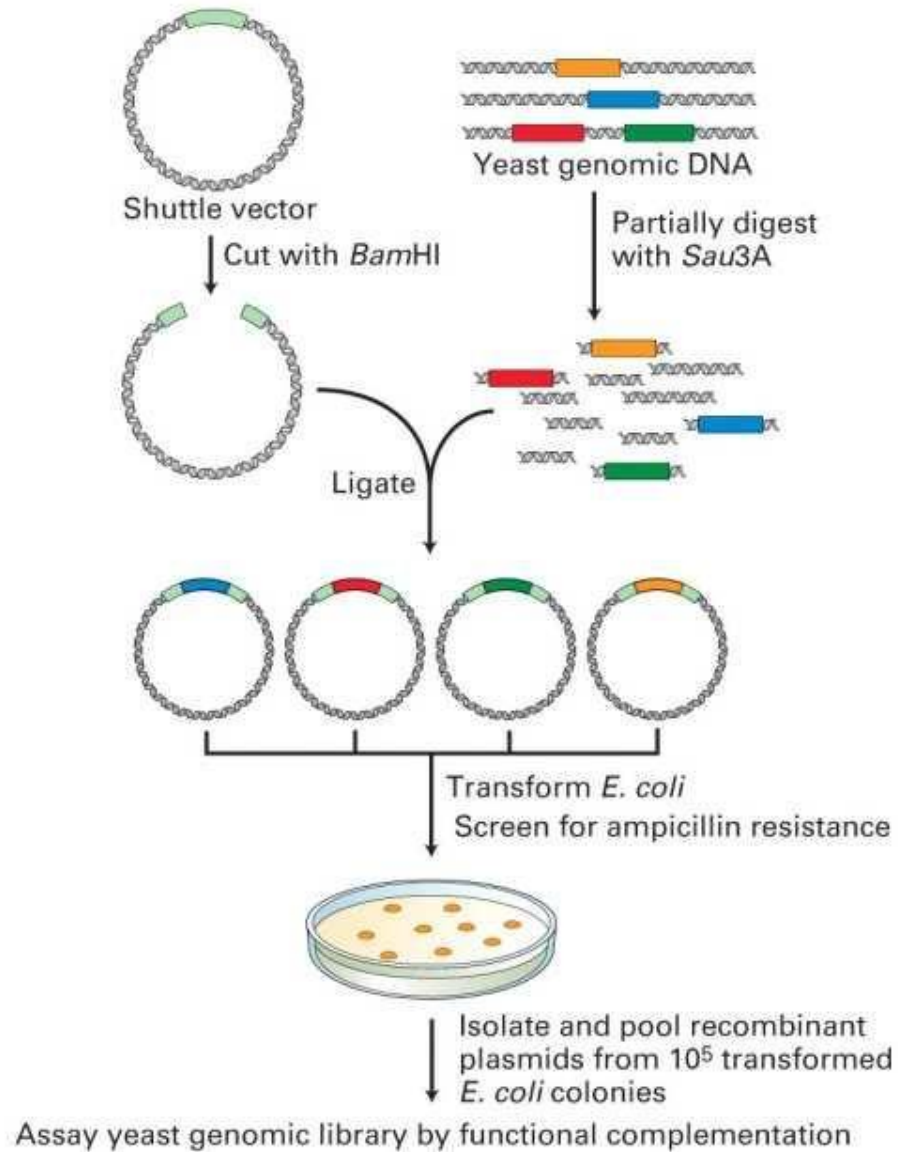
Plaque



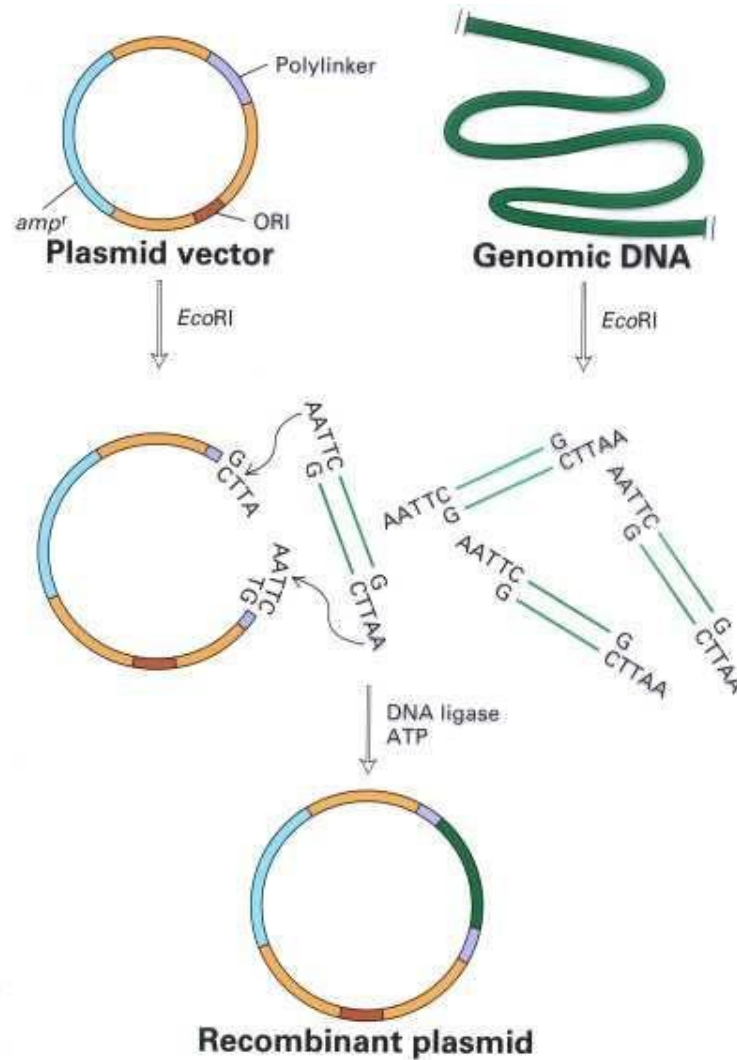
# Cloning of foreign DNA in $\lambda$ phages.



# What is a gene library?



# Creation of Libraries



# Creation of Libraries

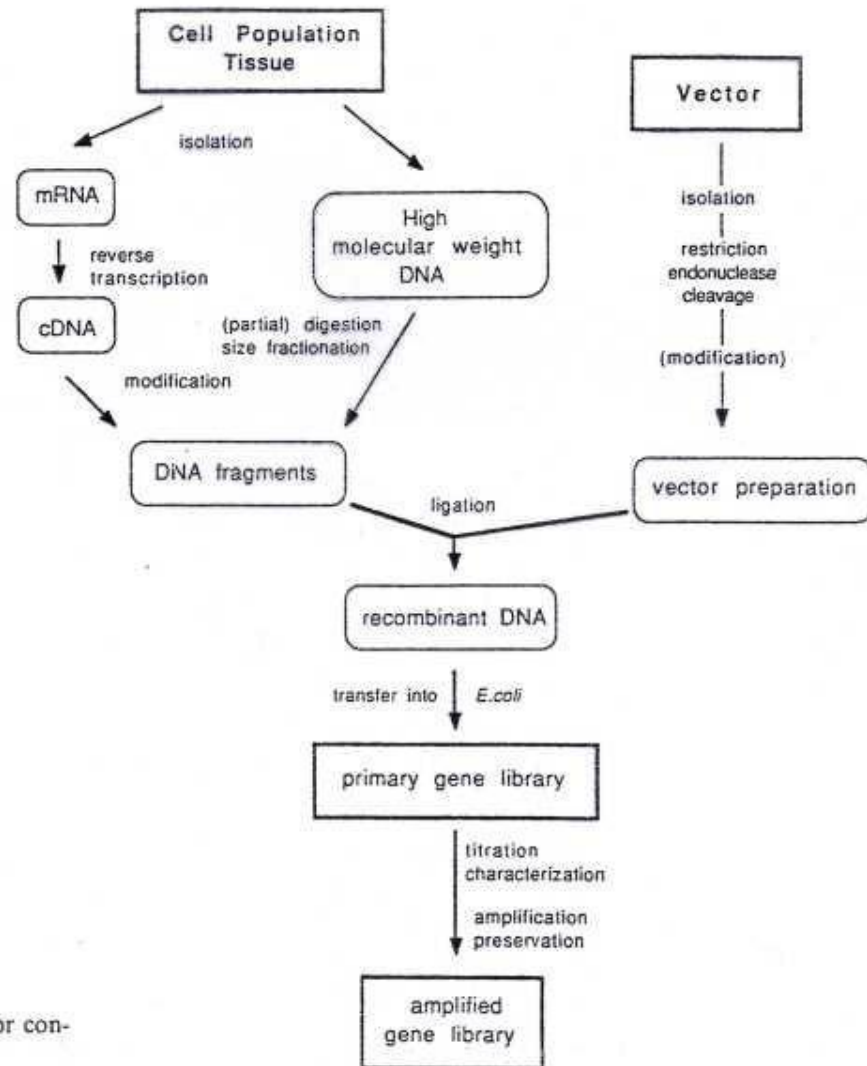


Fig. 6. General scheme for constructing gene libraries.

# Sizes of Some DNA Molecules.

Organism	Number of base pairs (kb) <sup>a</sup>	Contour length (μm)
<b>Viruses</b>		
Polyoma, SV40	5.2	1.7
λ Bacteriophage	48.6	17
T2, T4, T6 bacteriophage	166	55
Fowlpox	280	193
<b>Bacteria</b>		
<i>Mycoplasma hominis</i>	760	260
<i>Escherichia coli</i>	4,600	1,600
<b>Eukaryotes</b>		
Yeast (in 17 haploid chromosomes)	12,000	4,100
<i>Drosophila</i> (in 4 haploid chromosomes)	180,000	61,000
Human (in 23 haploid chromosomes)	3,200,000	1,100,000
Lungfish (in 19 haploid chromosomes)	102,000,000	35,000,000

<sup>a</sup>kb = kilobase pair = 1000 base pairs (bp).

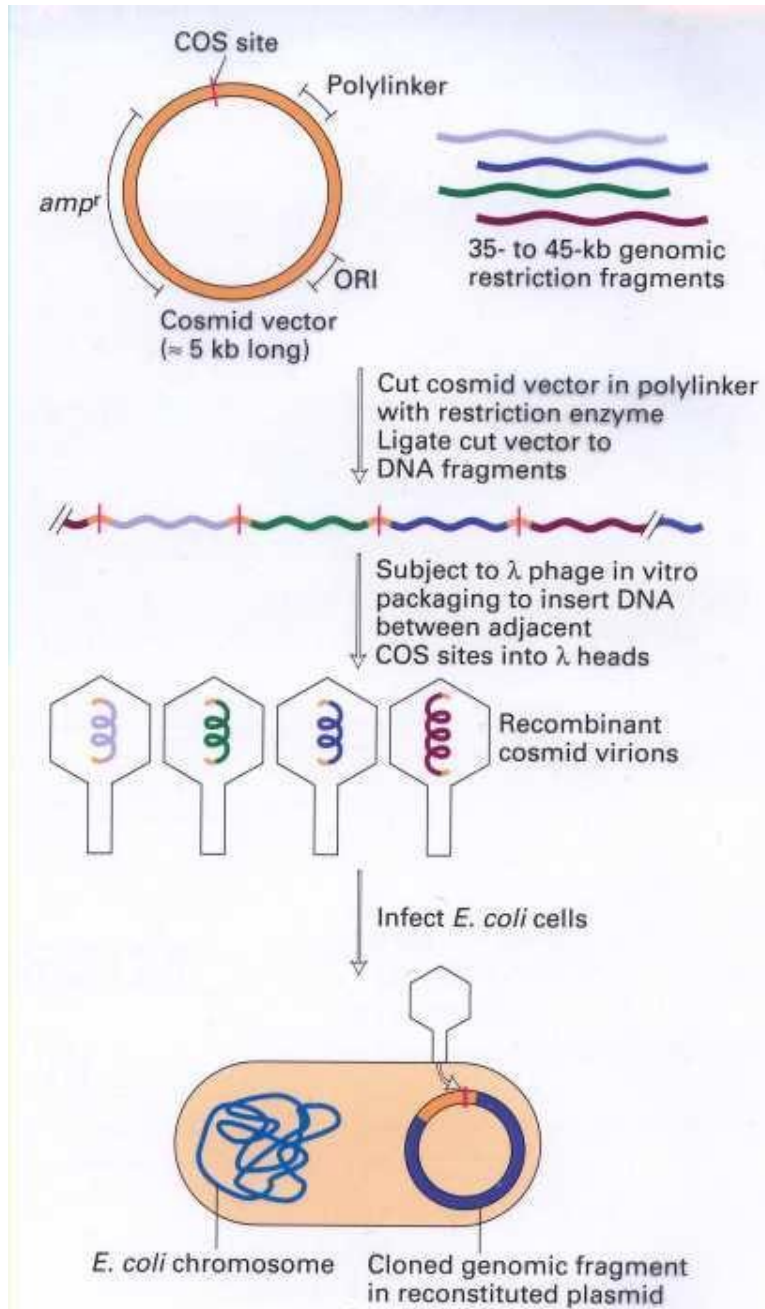
Source: Mainly Kornberg, A. and Baker, T.A., *DNA Replication* (2nd ed.), p. 20, Freeman (1992).

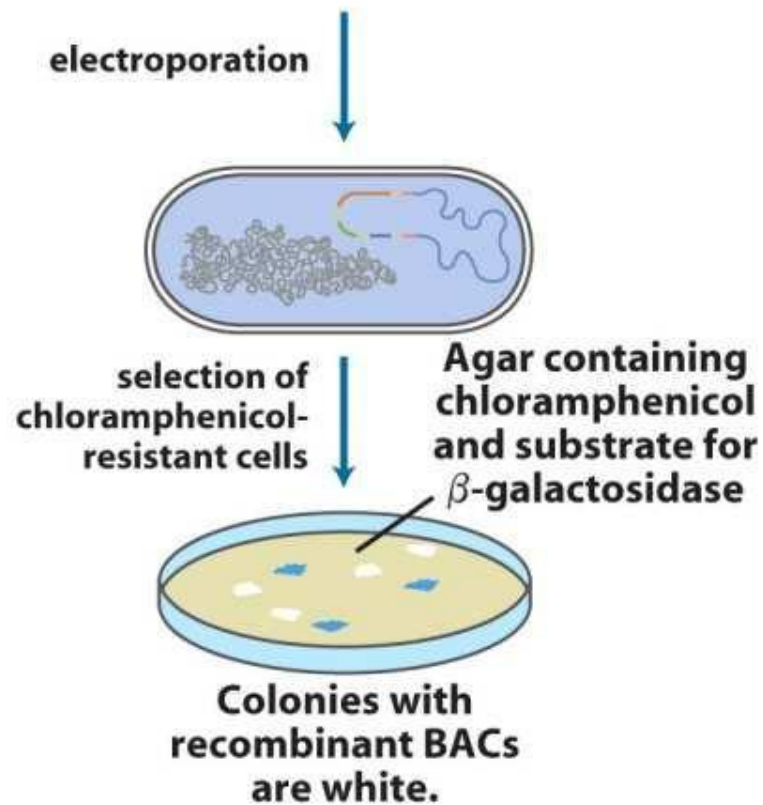
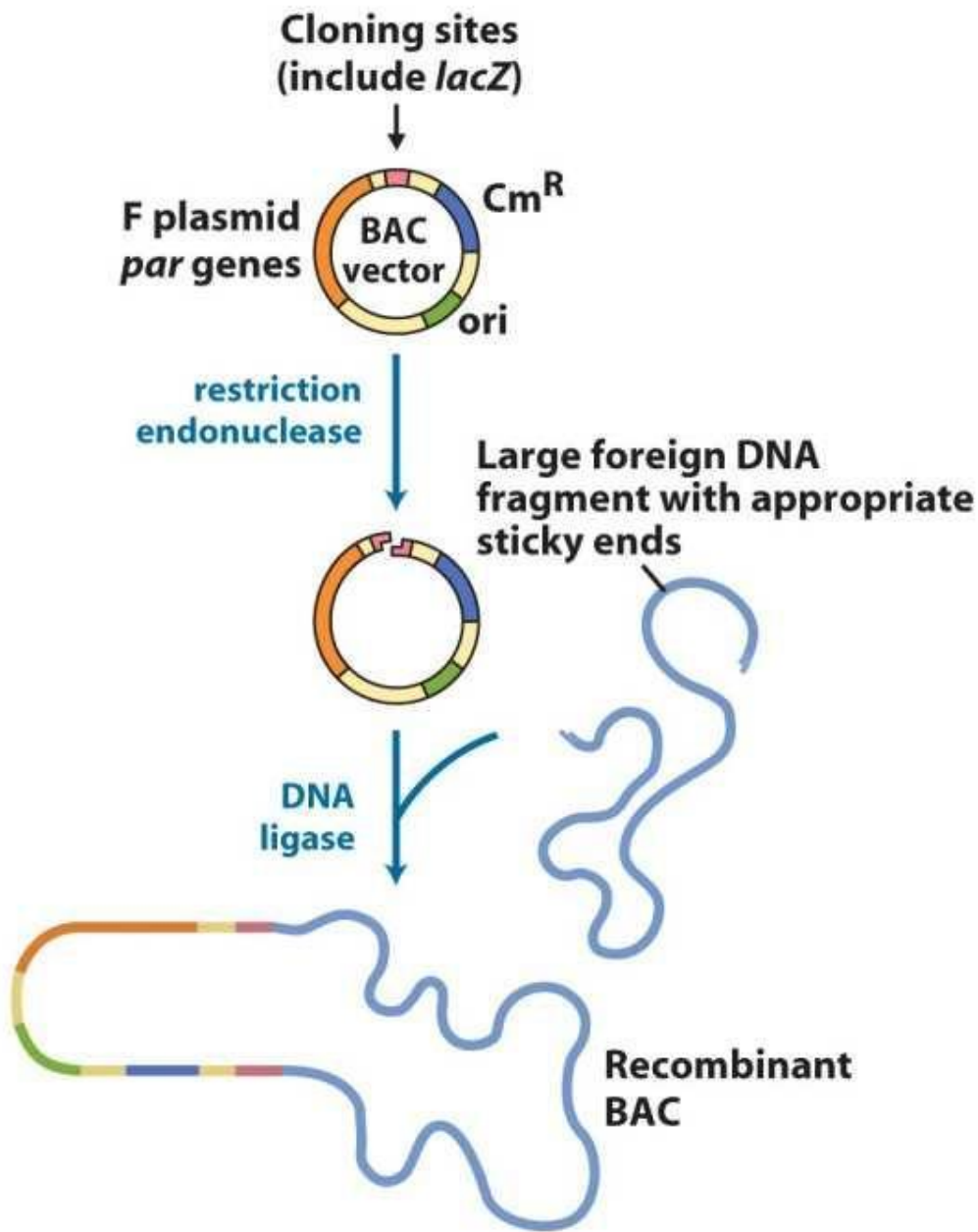


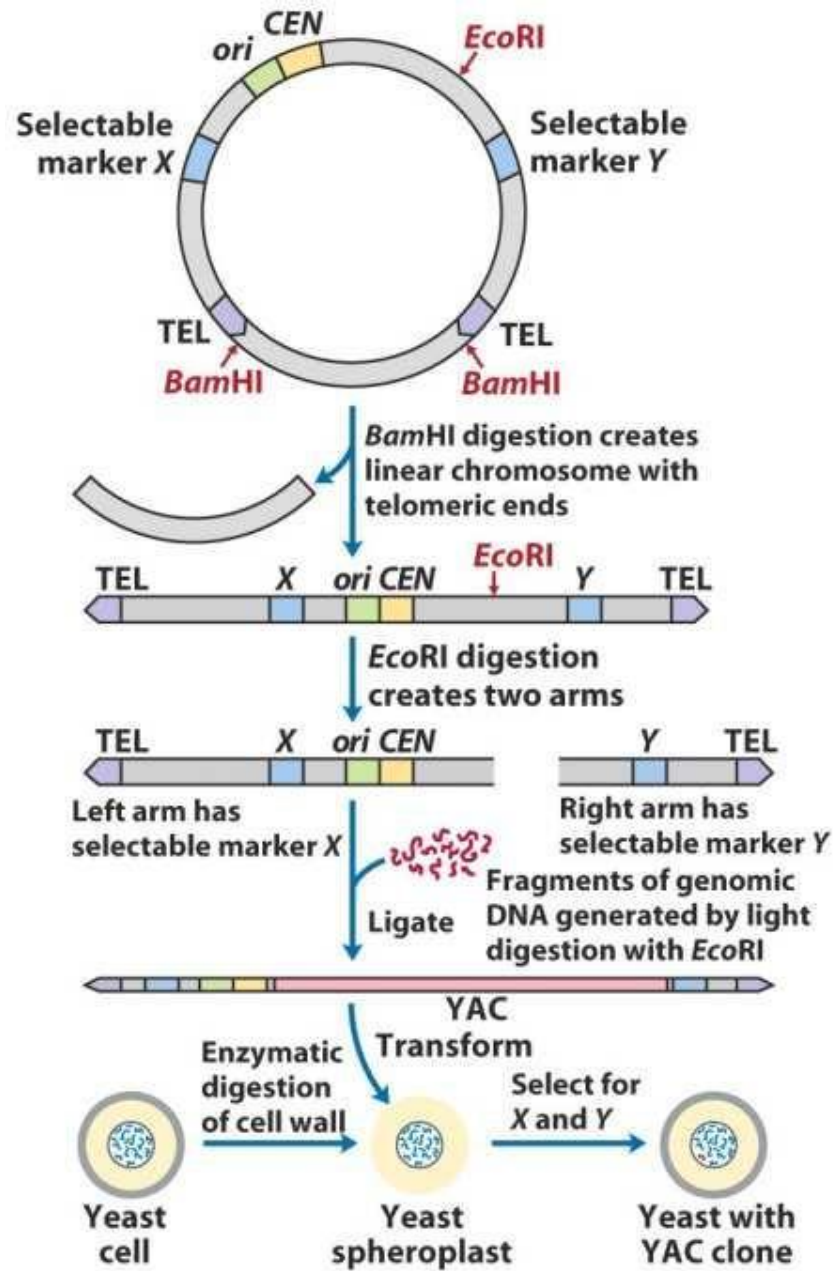
**Table 4.6** Insert capacities of some commonly used vector systems

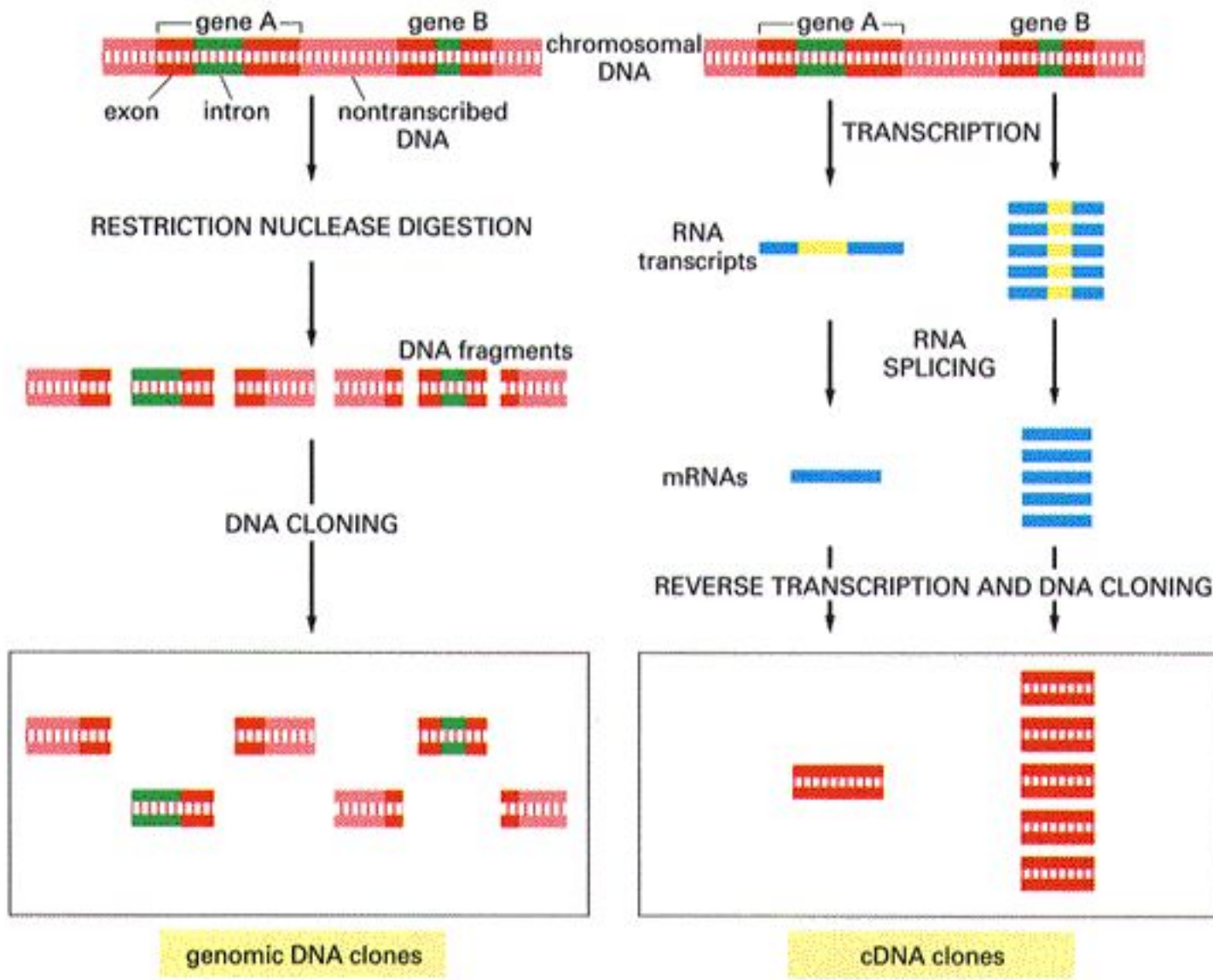
<b>Vector system</b>	<b>Host cell</b>	<b>Insert capacity (kb)</b>
Plasmid	<i>E. coli</i>	0.1–10
Bacteriophage $\lambda$	$\lambda$ / <i>E. coli</i>	10–20
Cosmid	<i>E. coli</i>	35–45
Bacteriophage P1	<i>E. coli</i>	80–100
BAC	<i>E. coli</i>	50–300
P1 bacteriophage-derived artificial chromosome	<i>E. coli</i>	100–300
Yeast artificial chromosome	Yeast	100–2,000
Human artificial chromosome	Cultured human cells	>2,000

# Cosmid = Cos - Plasmid

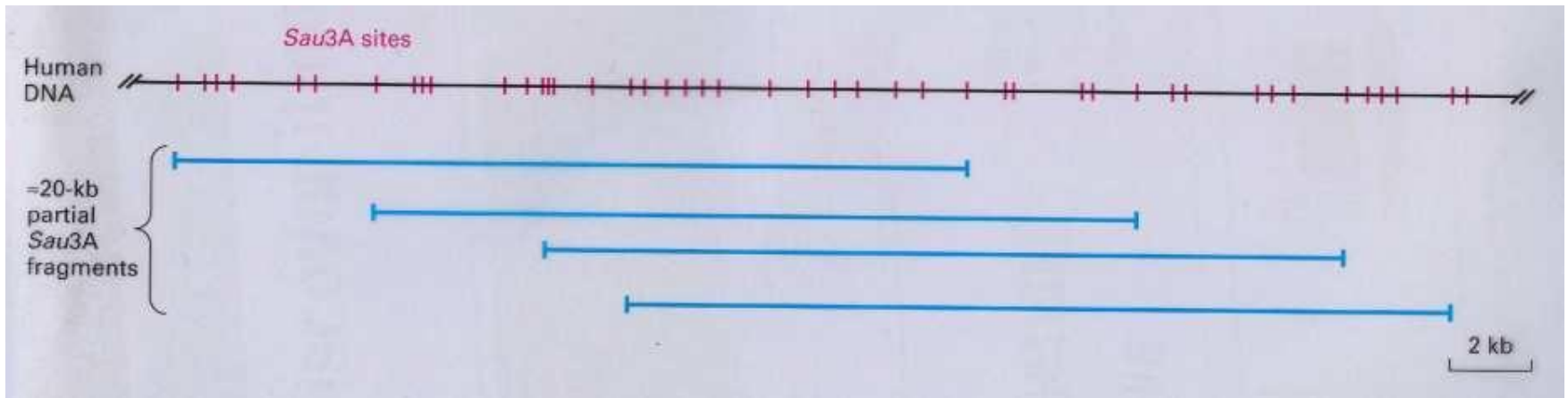




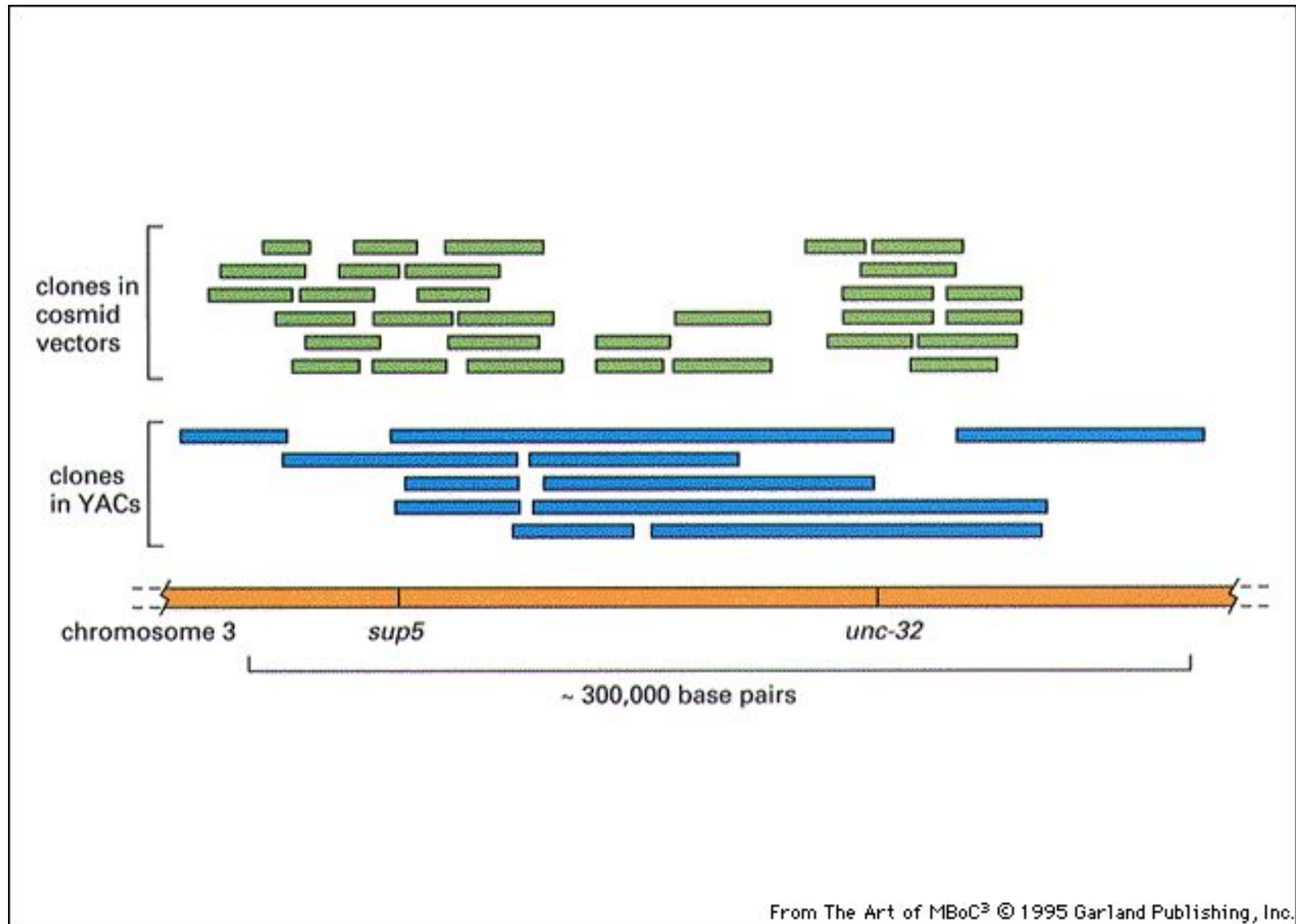




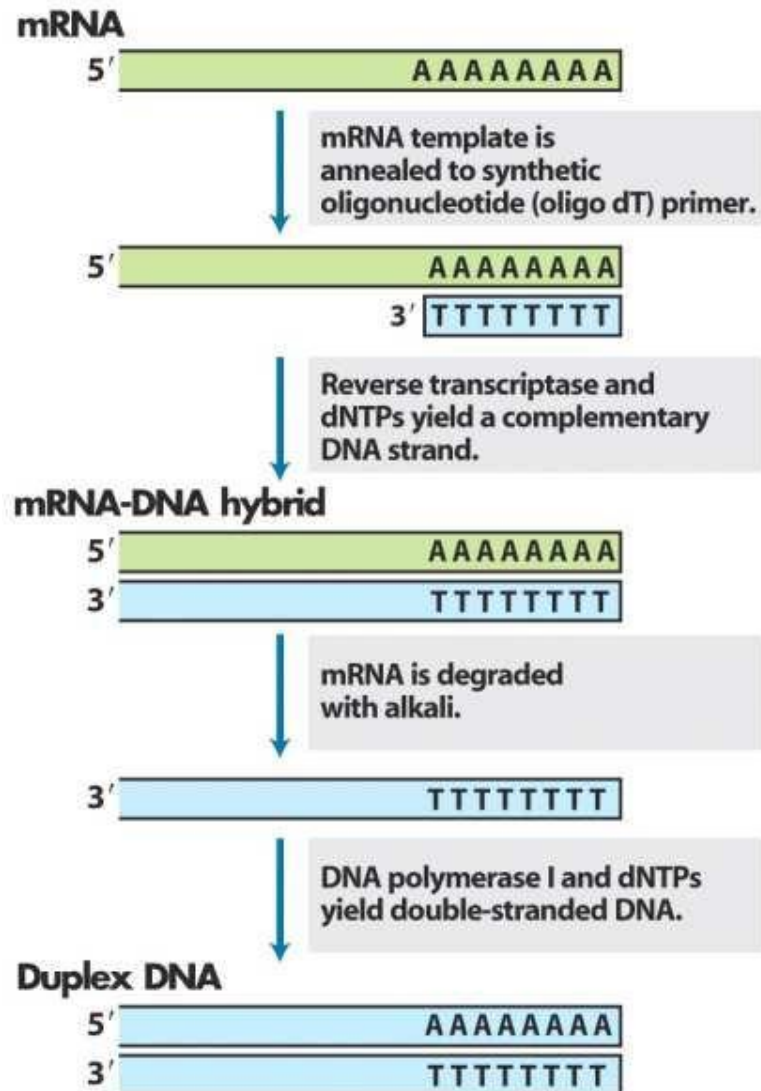
# Fragmentation of genomic DNA







# cDNA synthesis



# DNA Library



Plasmid vectors

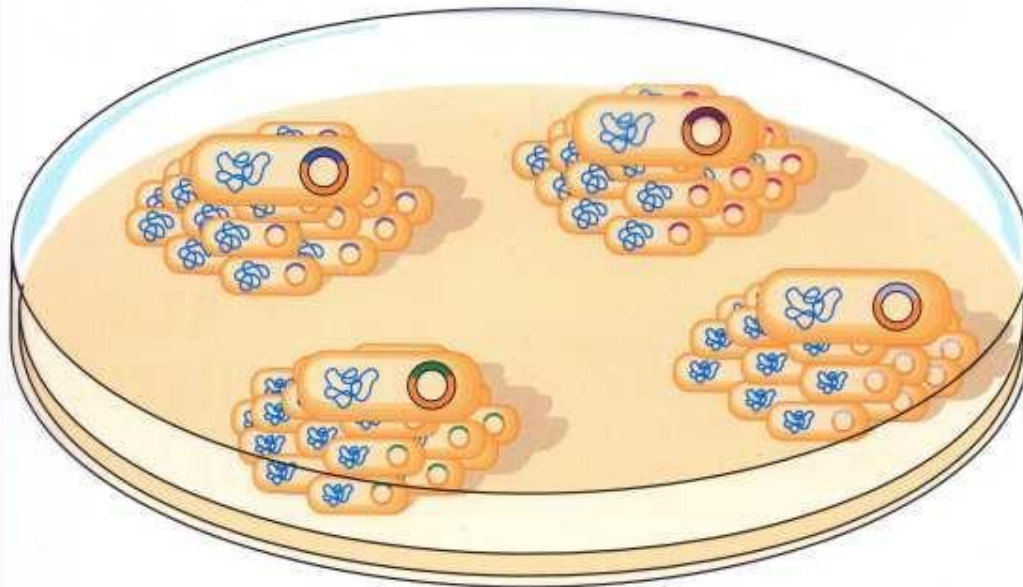


DNA fragments to be cloned

Enzymatically insert DNA fragments into plasmid vectors

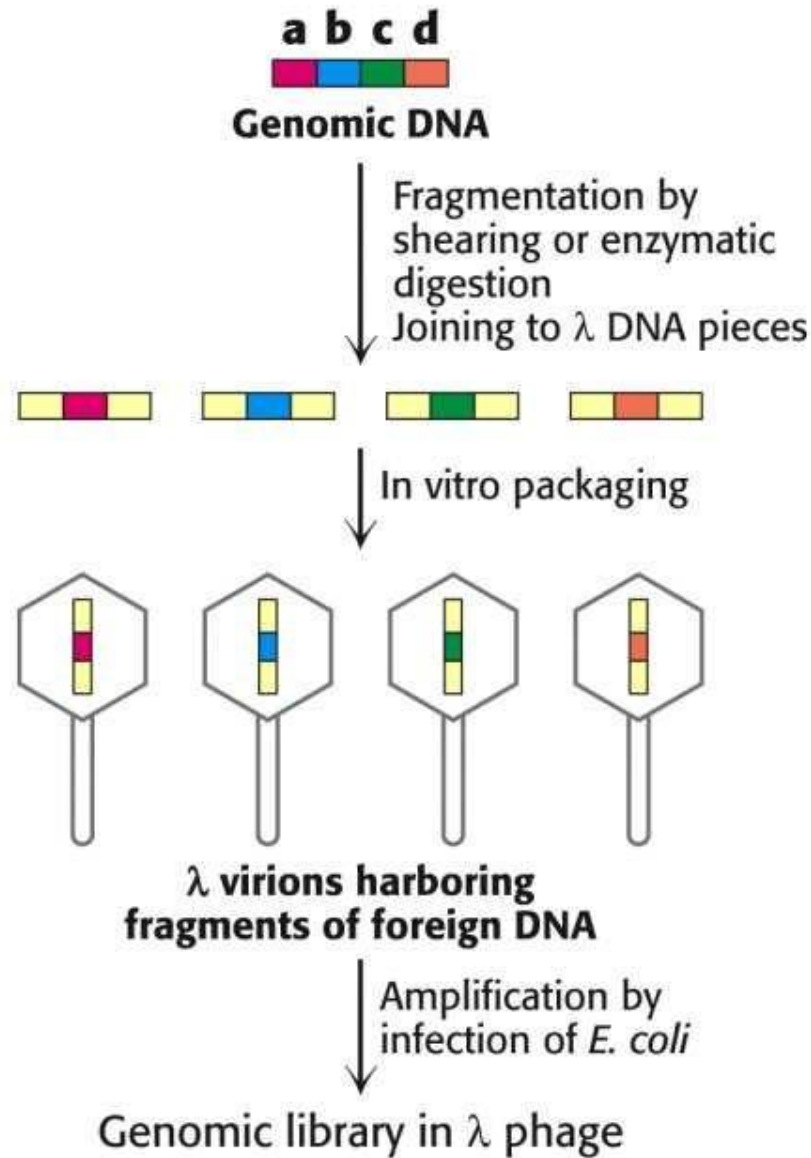


Transform *E. coli* cells and select for ampicillin-resistant colonies



Clones -> genetically identical

# Genomic phage library



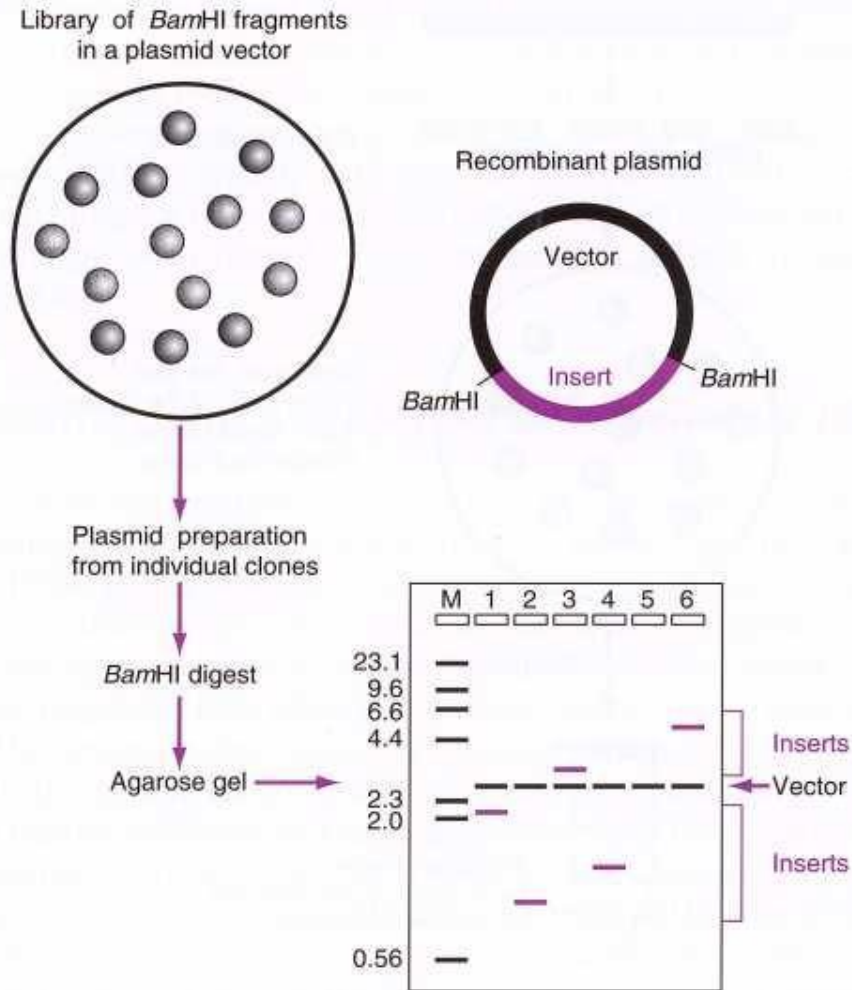
# Evaluation of library

## **Box 7.1** *Estimates of the required size of genomic libraries*

Organism	Genome size	Vector type	Insert size	<i>P</i>	Library size
Bacterium	$4 \times 10^6$ bases	plasmid	4 kb	0.99	$4.6 \times 10^3$
		lambda replacement	18 kb	0.99	$1.0 \times 10^3$
		cosmid	40 kb	0.99	458
		BAC	300 kb	0.99	59
Mammal	$3 \times 10^9$ bases	plasmid	4 kb	0.99	$3.5 \times 10^6$
		lambda replacement	18 kb	0.99	$7.7 \times 10^5$
		cosmid	40 kb	0.99	$3.5 \times 10^5$
		BAC	300 kb	0.99	$4.6 \times 10^4$

The values shown for the genome sizes of bacteria and mammals are examples for the purpose of this calculation. The actual genome sizes vary quite widely from one organism to another. The insert sizes for specific vectors will also vary.

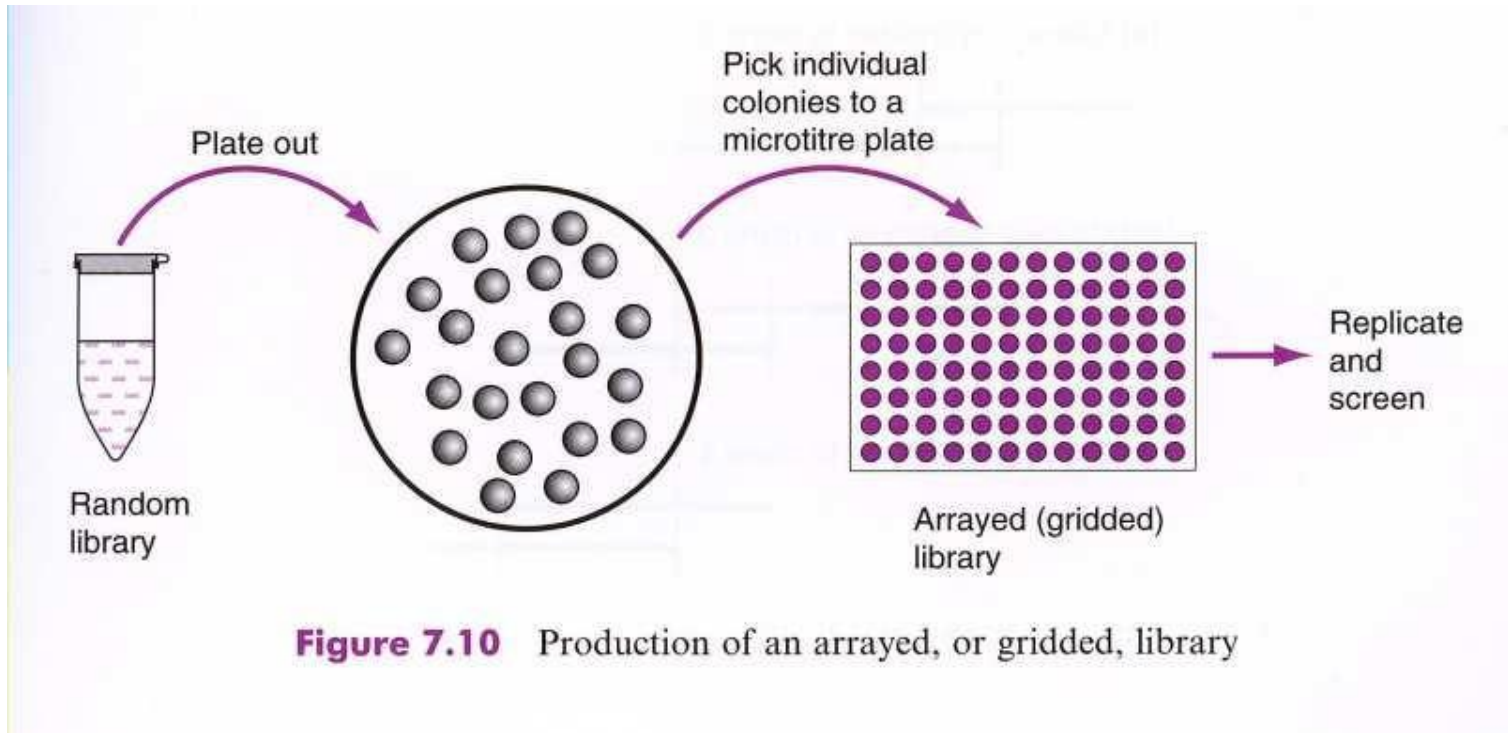
# Evaluation of library



**Figure 7.4** Assessing the quality of a gene library



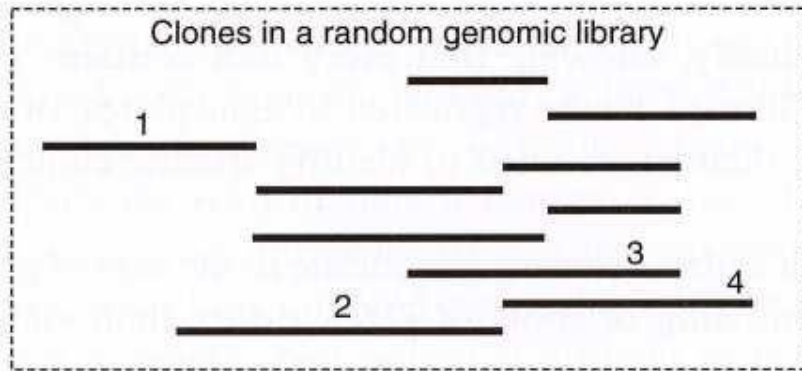
# Ordered library



**Figure 7.10** Production of an arrayed, or gridded, library

—————> **Microarrays**

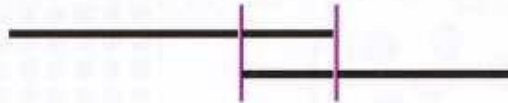
# Ordered library



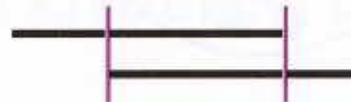
(a) Clone 1 hybridizes to clone 2



(b) Clone 2 hybridizes to clone 3



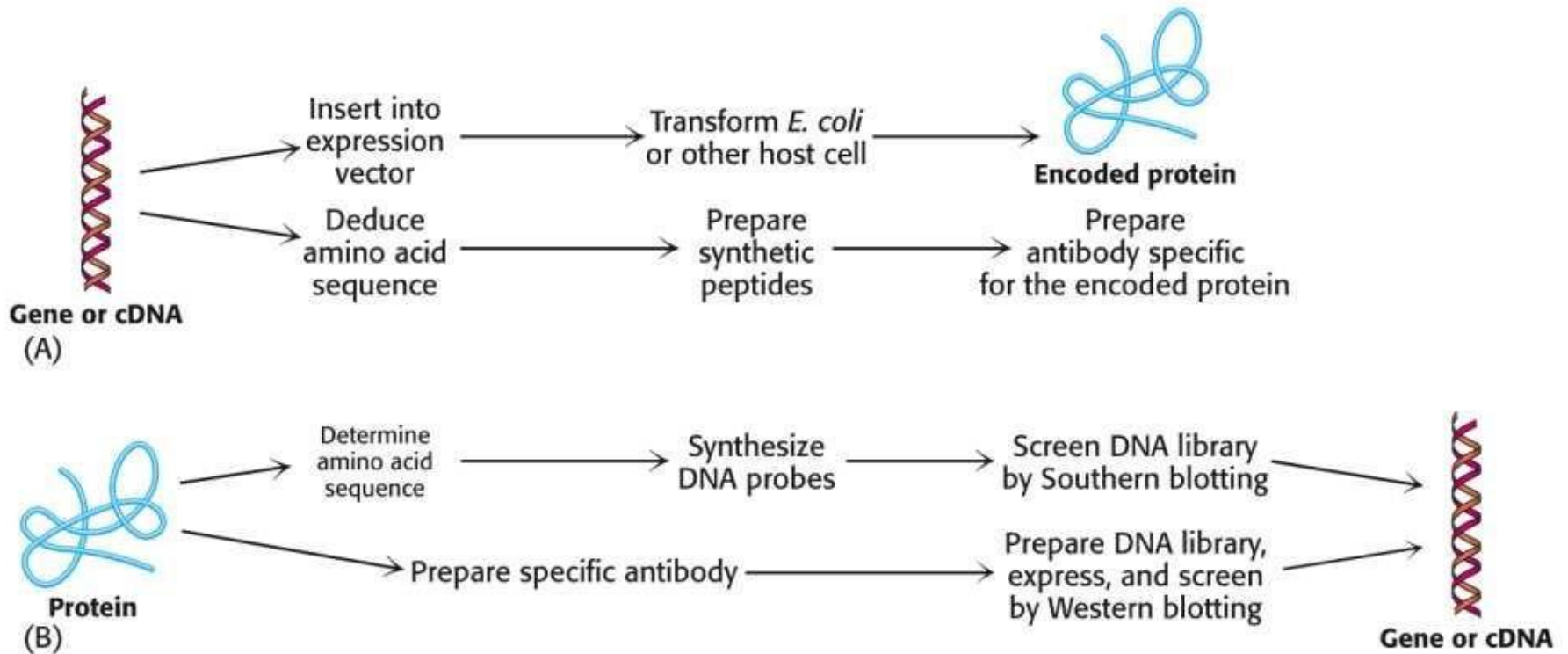
(c) Clone 3 hybridizes to clone 4



→ "Chromosome Walking"  
→ also used in "Human Genome Project"

**Figure 7.11** Production of an ordered library

# Different ways to clone a gene



# Bacterial host engineering



*Escherichia coli* (E. coli) is a type of bacteria normally found in the intestines of people and animals.

Although most strains of E. coli are harmless, some can cause illness or even death.

The most serious form is E. coli **0157:H7**.

E. coli leads to about **73,000 cases of infection** and **61 deaths** each year in the United States.

# Genetic and physical maps of the *E. coli* chromosome

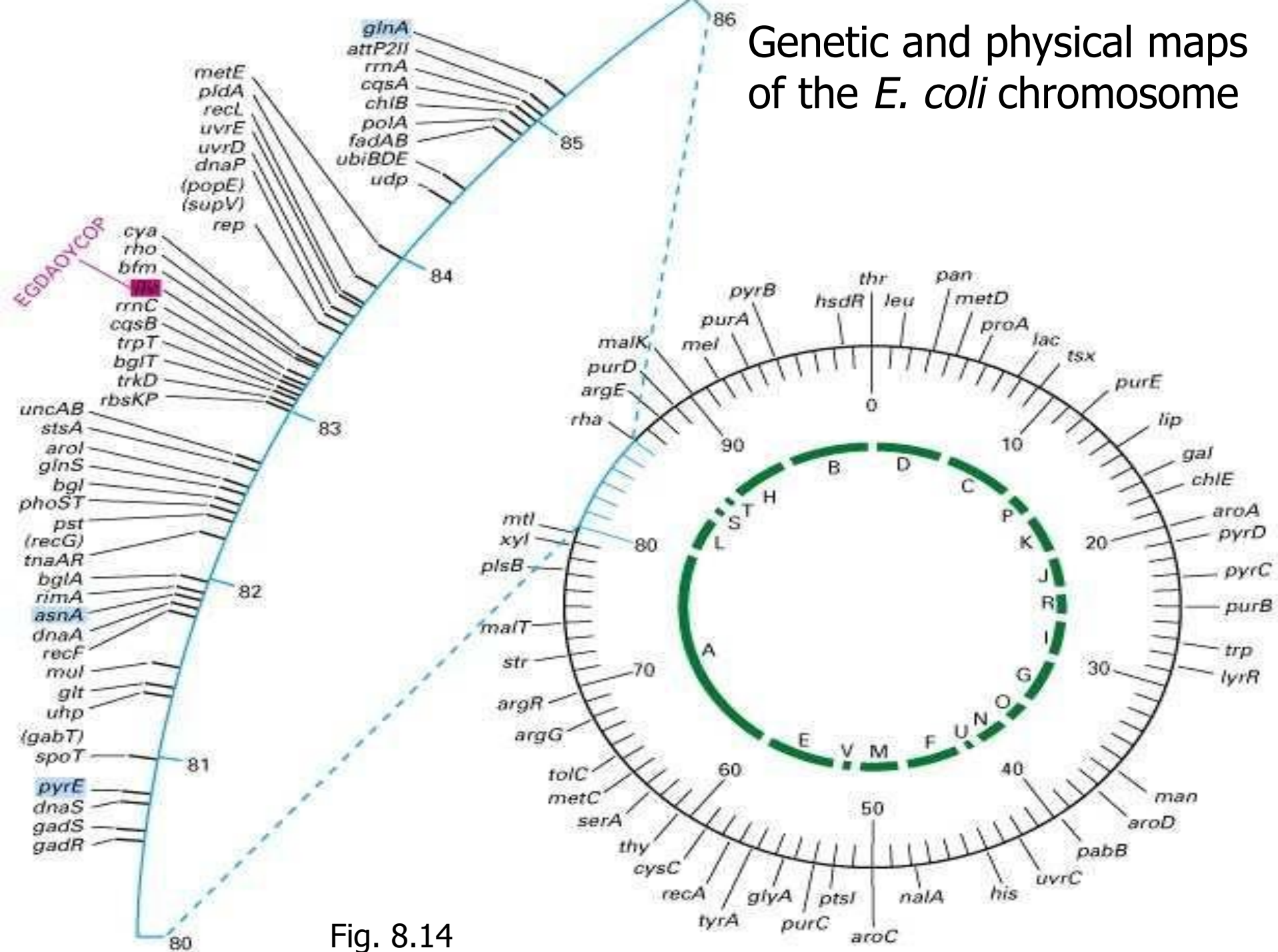


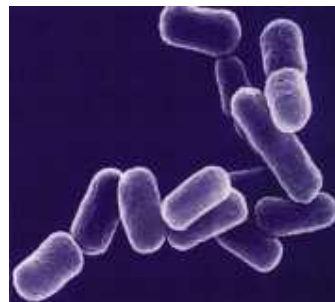
Fig. 8.14

# E.Coli K12 strain has been used for further engineering

The K12 strain was first isolated in 1921 from the stool of a malaria patient and it has been maintained in laboratory stocks as a pure strain for the last 75 years.

NIH Recombinant Advisory Committee (RAC) (1973)

Asilomar Conference on Recombinant DNA (February 1975)



Every strain comes with description of its genotype:  
DH5alpha (recA-; hsdR-; LacIq; uvrA-; mcrA-.....)

Most strain in molecular biology are *recA*<sup>85</sup>*endA*- *hsdR*-



# Additional changes in K12 E.coli for ease of the laboratory practice

## 1. Bacterial restriction modification systems have been removed.

(To prevent its interference with the replication of foreign DNA in bacteria).



### **hsdR/hsdM/hsdS (EcoK) restriction system**

Degrades DNA not methylated  
at the sequence 5'-AAC-(N)5-GTGC-3'

- hsdM** recognises unmethylated DNA
- hsdM** is also involved in methylation of DNA
- hsdR** encodes an endonuclease
- hsdS** encodes DNA sequence specific protein

**hsdR- or hsdS- mutants** facilitate propagation of any foreign DNA

### **mcrA/mcrB/mrr complex**

E.coli DNA is methylated  
by **dcm**, **dam** and **hsdM**

**mcrA/mcrB/mrr** cleaves DNA  
methylated by other systems

**-mcrA-/mcrB-**  
strains are good for cloning  
eukaryotic DNA

# Additional changes in K12 E.coli for ease of the laboratory practice

## 2. DNA recombination systems are modified to prevent rearrangements (RecA-)

(to prevent deletions and rearrangements)

**recA** is a core recombination protein

**recA- strains** allow cloning of repetitive sequences

**recA-/recB-/recC-** are enhanced strains with very low recombination efficiency

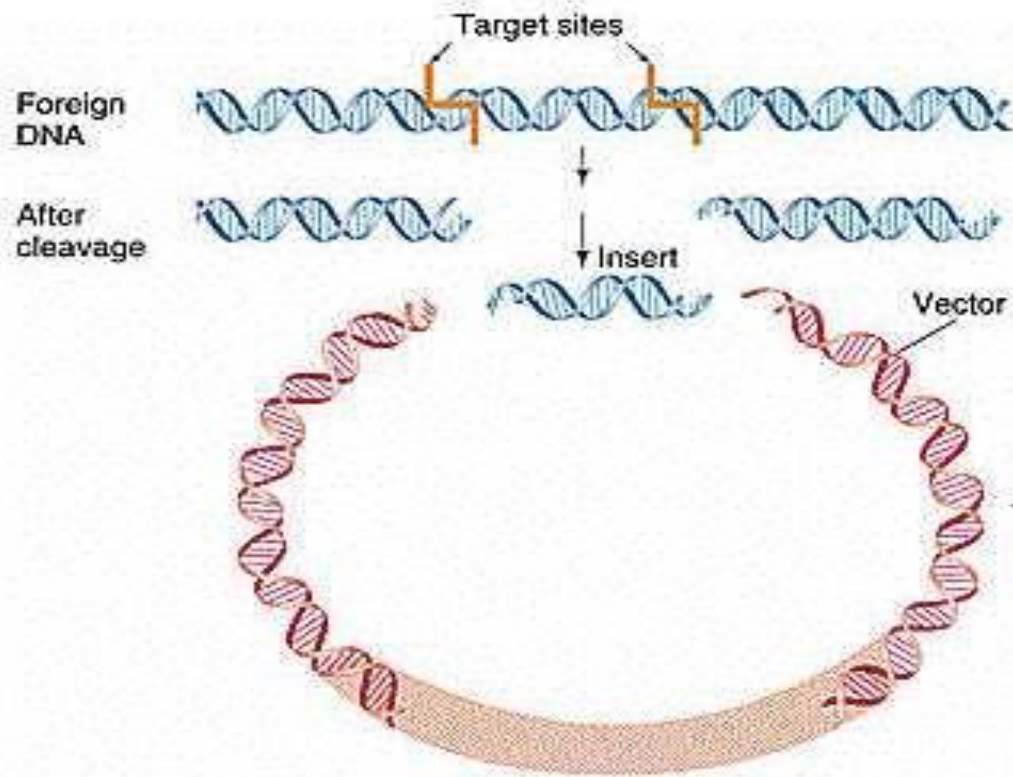
**uvrC/umuC** are involved in DNA repair

**uvrC-/umuC-** are good for cloning of inverted repeats

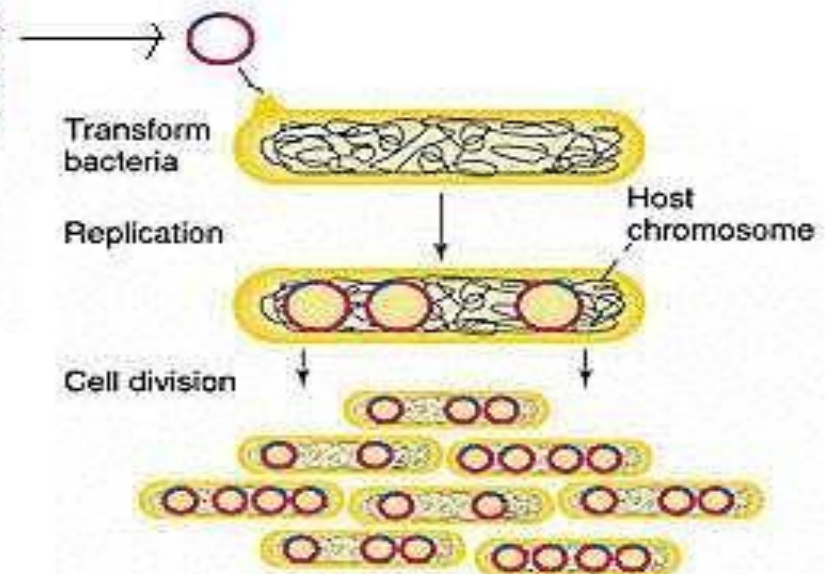
## 3. Endonuclease activity has been mutated (EndA-)

(to increase plasmid yields and improve the quality of DNA – no nicks)

# Transformation of plasmid DNA in competent *E. coli* cells



**Competent (here)  
= able to uptake DNA**

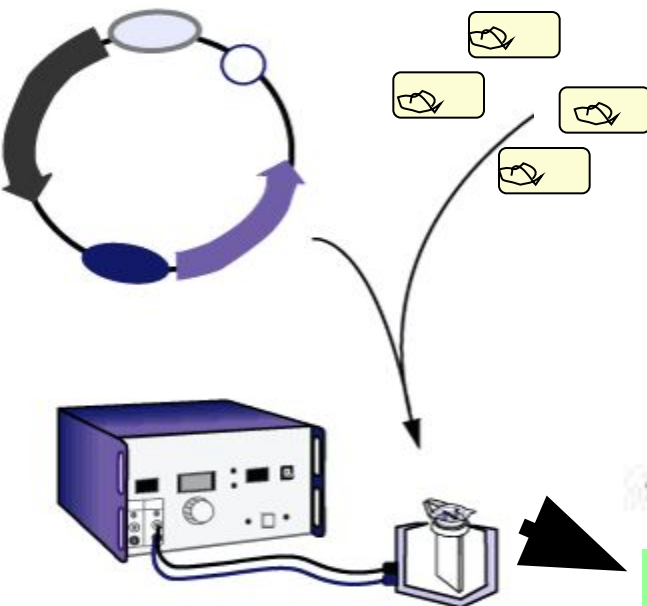


# Transformation of plasmid DNA to competent *E. coli* cells

-- Electroporation and electroporation-competent cells

-- Heat shock transformation and chemically competent cells

## Electroporation



put DNA and cells together

(do something special)



then something happens...

Recovery in rich growing media

## Chemical transformation

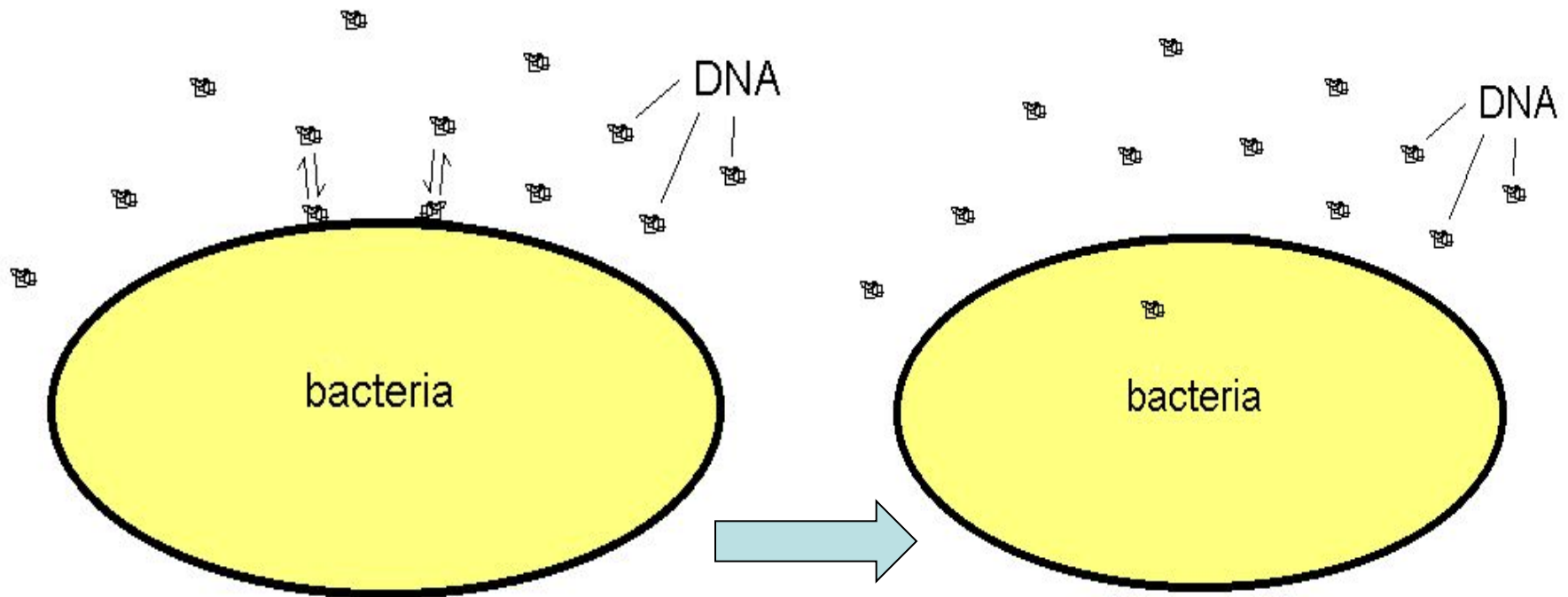
treating *E. coli* with  $\text{CaCl}_2$  will batter the membranes and essentially make the bacteria very unhappy.

$\text{CaCl}_2$  is gaping holes in the membrane

BRIEF HEAT SHOCK

89

# Calcium/phosphate (heat shock) method



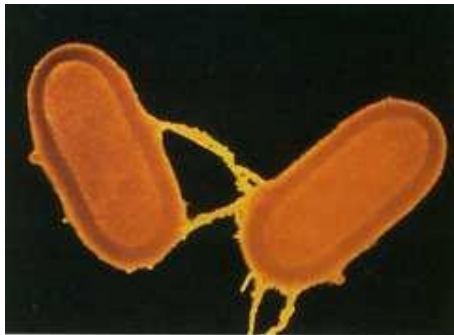
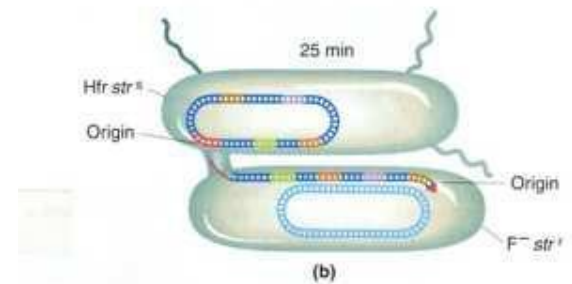
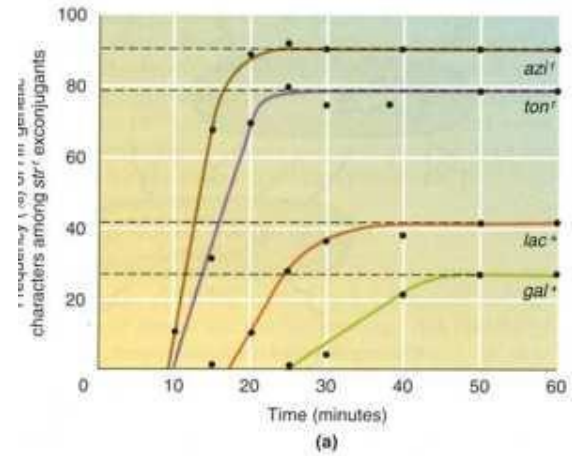
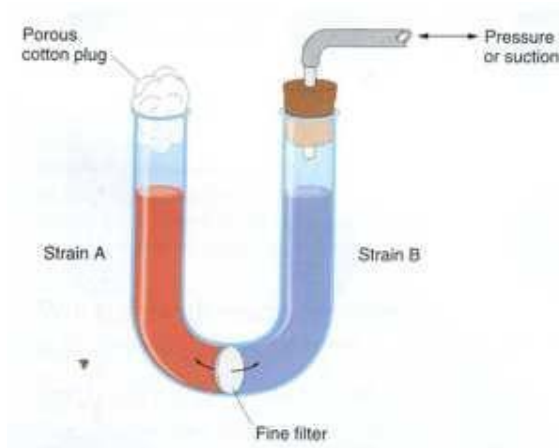
During the incubation on ice, DNA binds to the surface of the bacterium as a calcium-phosphate-DNA complex

Following a sudden increase in temperature, one or more DNA molecules bound to the surface of the cell is taken up by the competent cell.

# Conjugation

## Lederberg

## Monod



- F<sup>-</sup> to F<sup>+</sup>
- 100 minutes
- 4000 genes