Chapter 11

Genes: Structure, Replication, and Mutation

Terms and concepts

- clone
  - population of cells that are genetically identical
- genome
  - all genes present in a cell or virus
    - haploid – one set of genes
    - diploid – two sets of genes
- genotype
  - specific set of genes an organism possesses
- phenotype
  - set of observable characteristics

DNA as Genetic Material

- established by several critical experiments
  - Fred Griffith (1928)
  - Oswald T. Avery, C. M. MacLeod, and M. J. McCarty (1944)
  - Alfred D. Hershey and Martha Chase (1952)

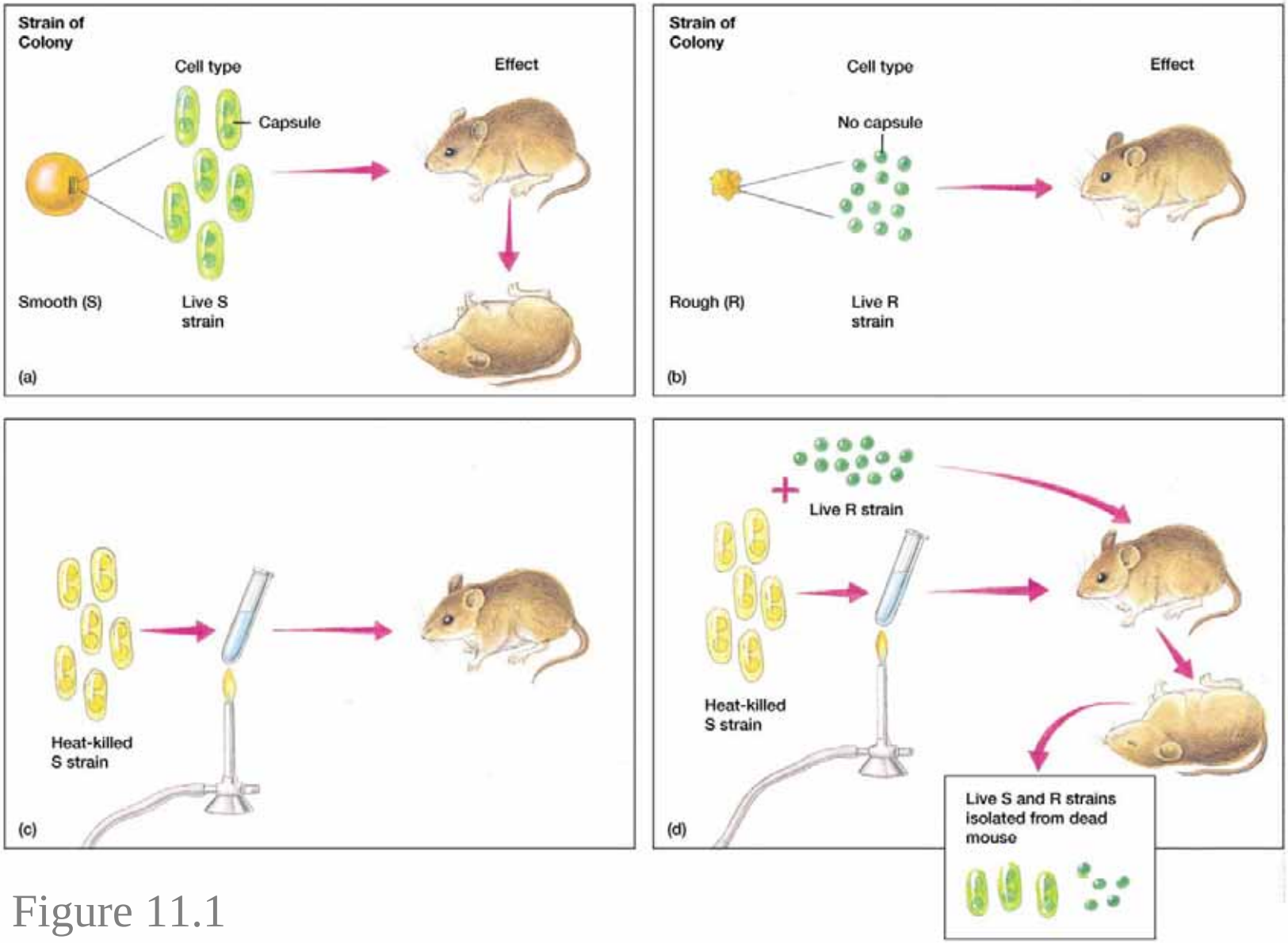

Strain of Colony
Cell type
Effect
Capsule
Smooth (S)
Live S strain
(a)
Strain of Colony
Cell type
Effect
No capsule
Rough (R)
Live R strain
(b)
Heat-killed S strain
(c)
Live R strain
Heat-killed S strain
Live S and R strains isolated from dead mouse
(d)

Figure 11.1

Transforming principle

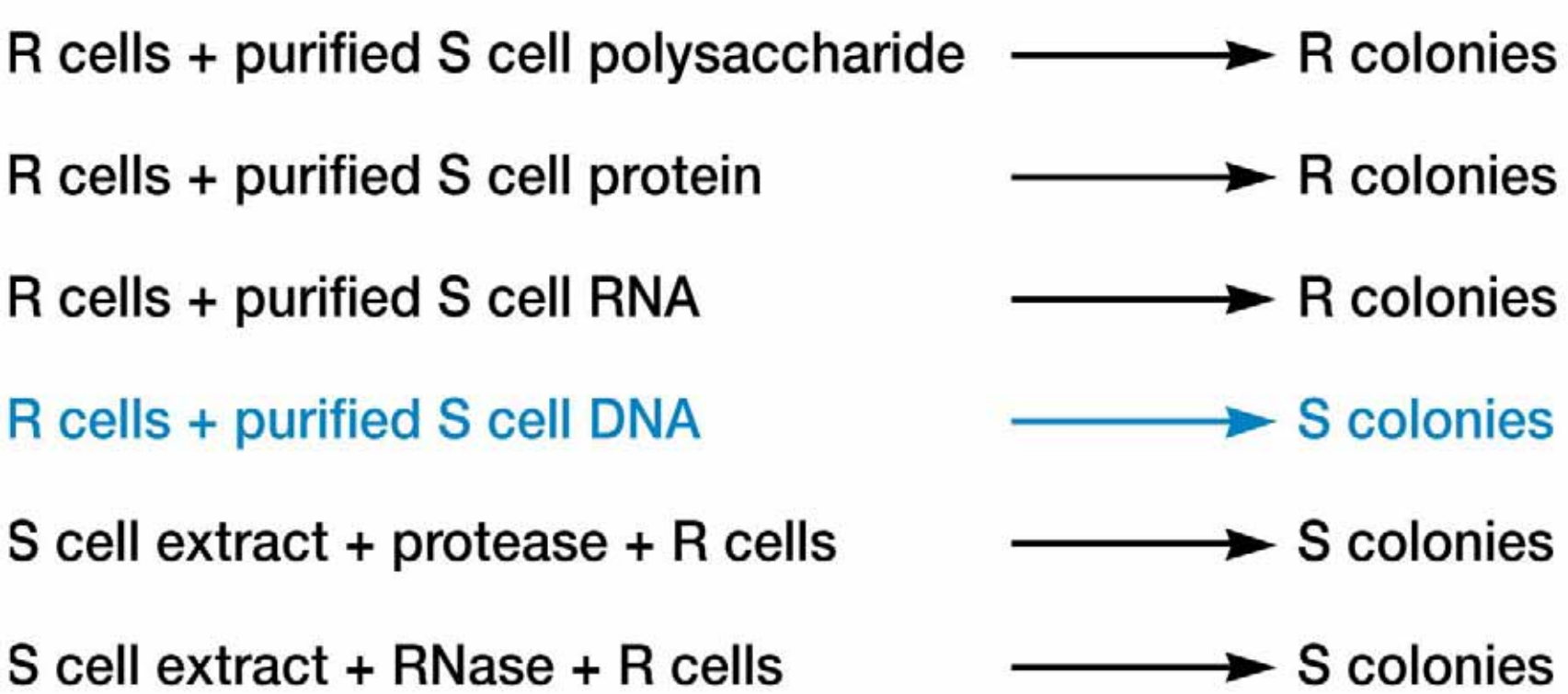


Figure 11.2

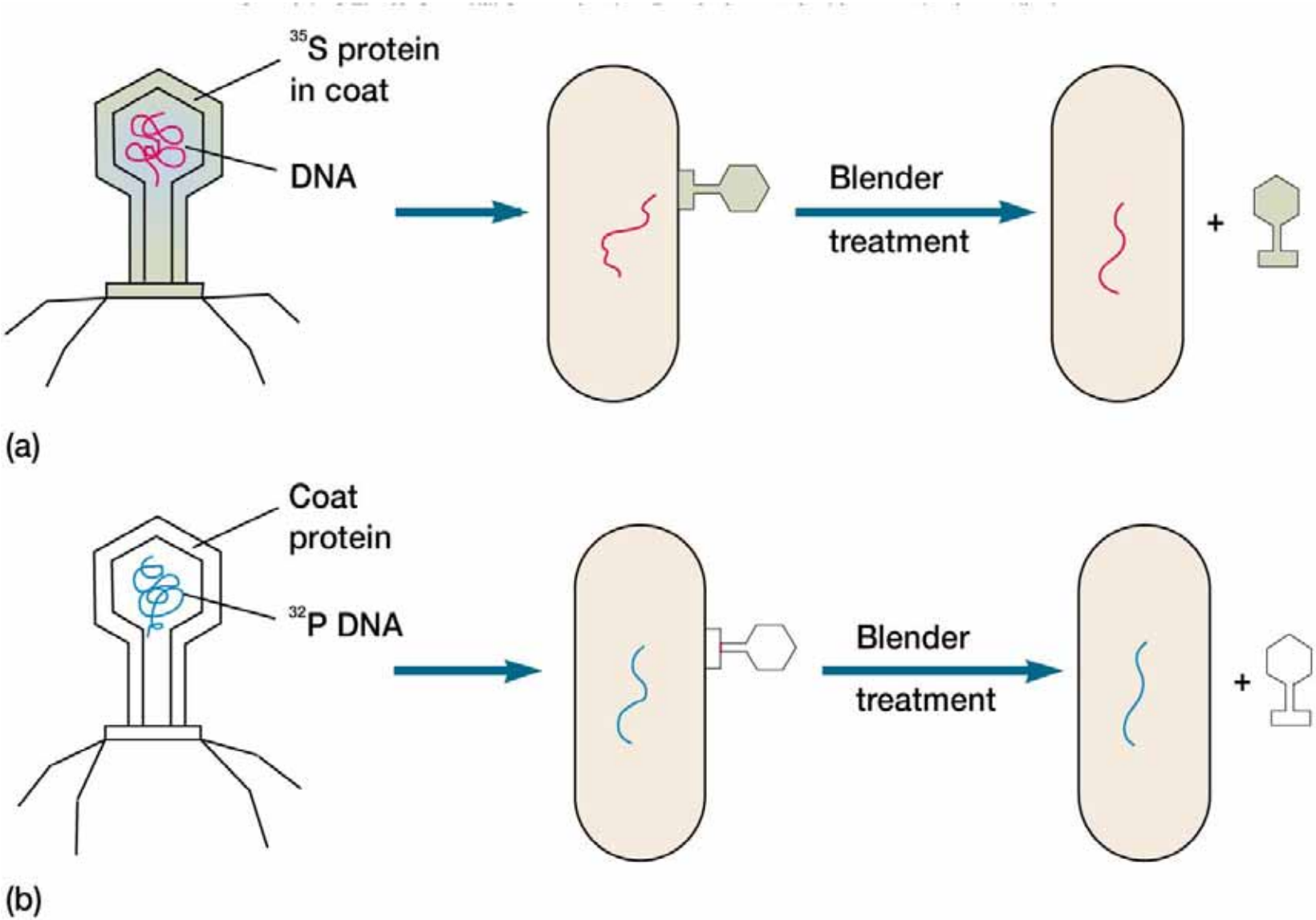

^{35}S protein in coat
DNA
Blender treatment
+
(a)
Coat protein
^{32}P DNA
Blender treatment
+
(b)

Figure 11.3

The Central Dogma

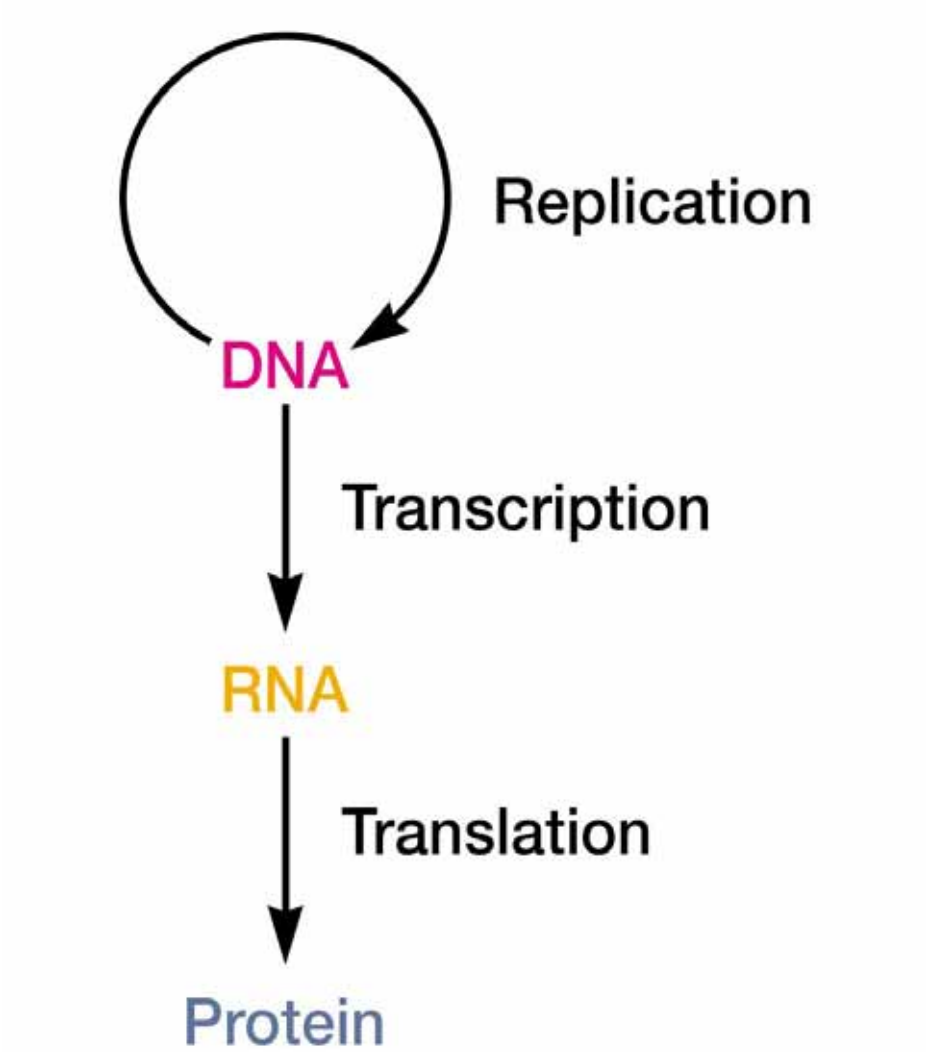


Figure 11.4

Nucleic Acid Structure

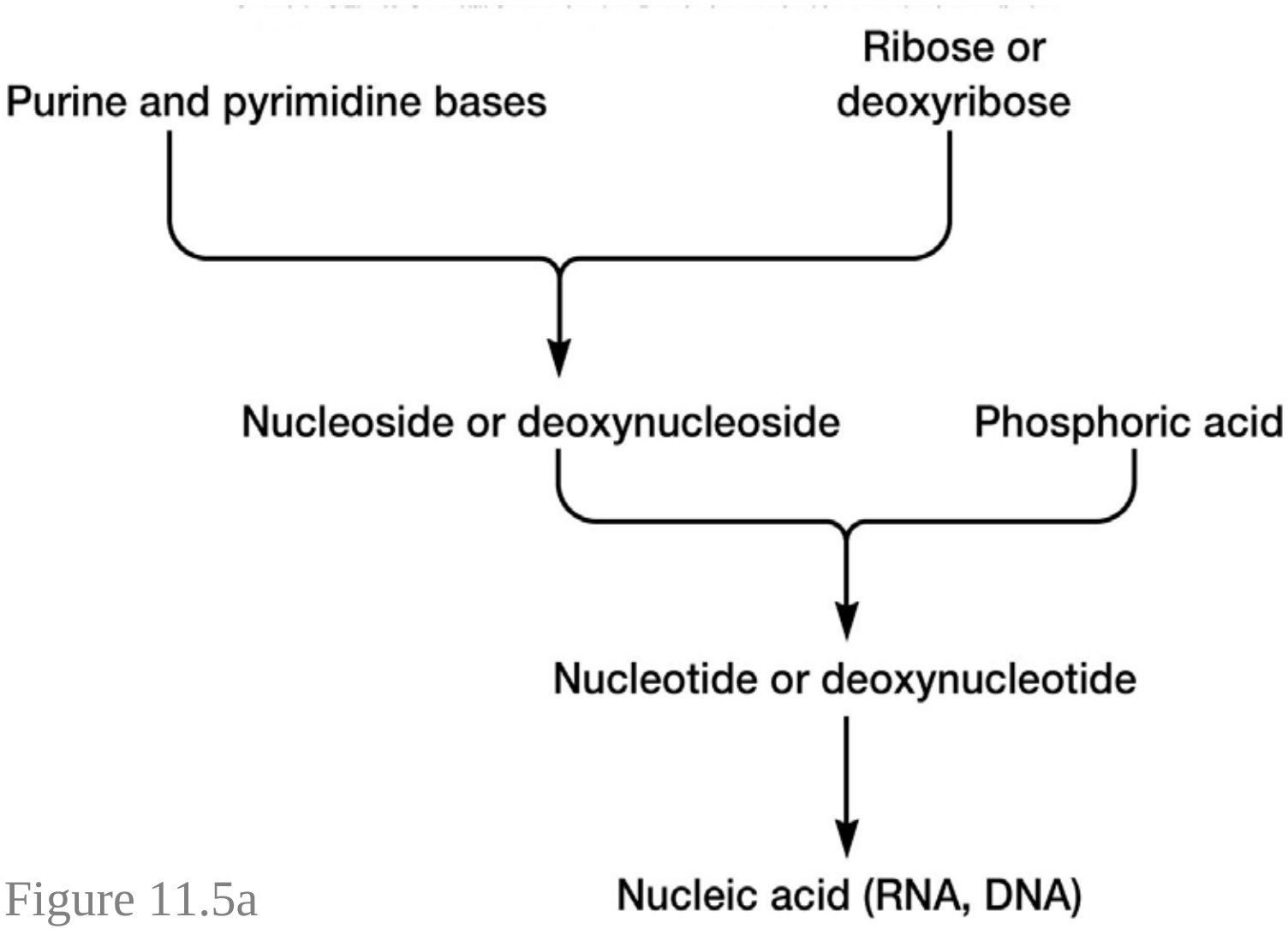


Figure 11.5a

Examples of nucleosides

nucleoside =
nitrogenous base
+ pentose sugar

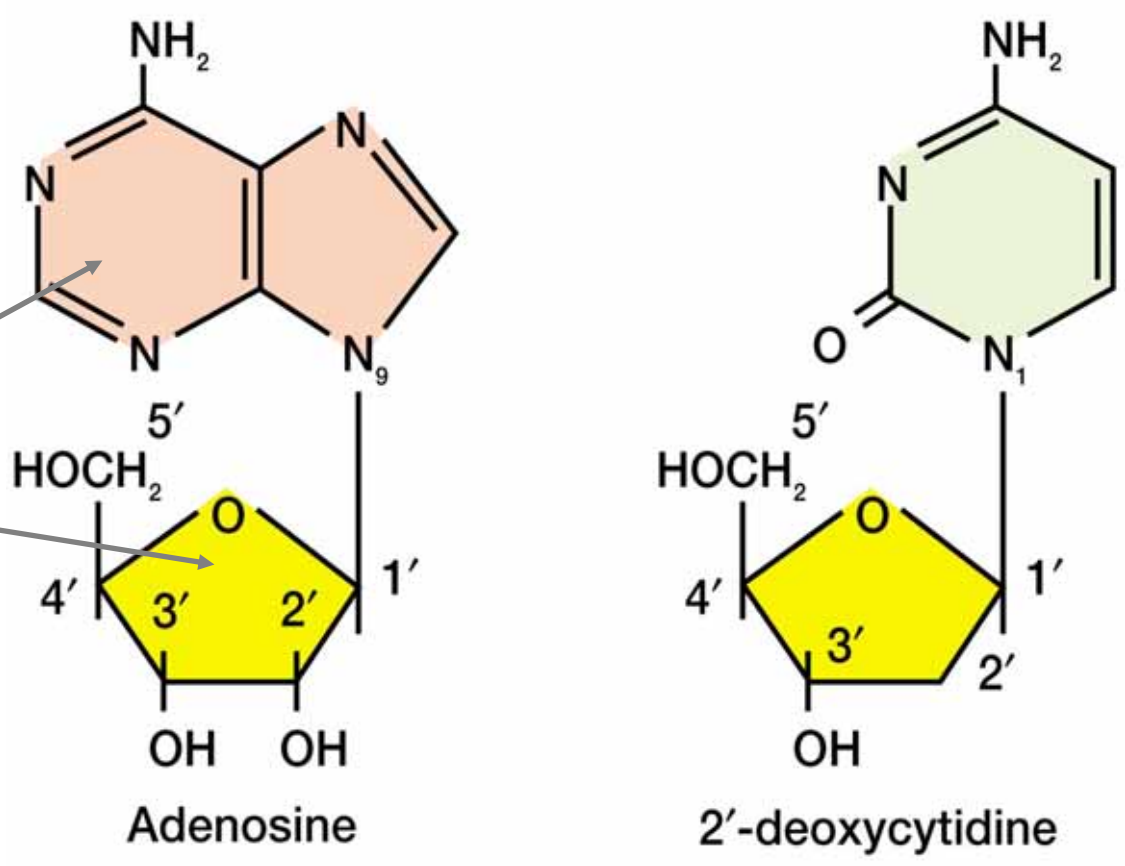
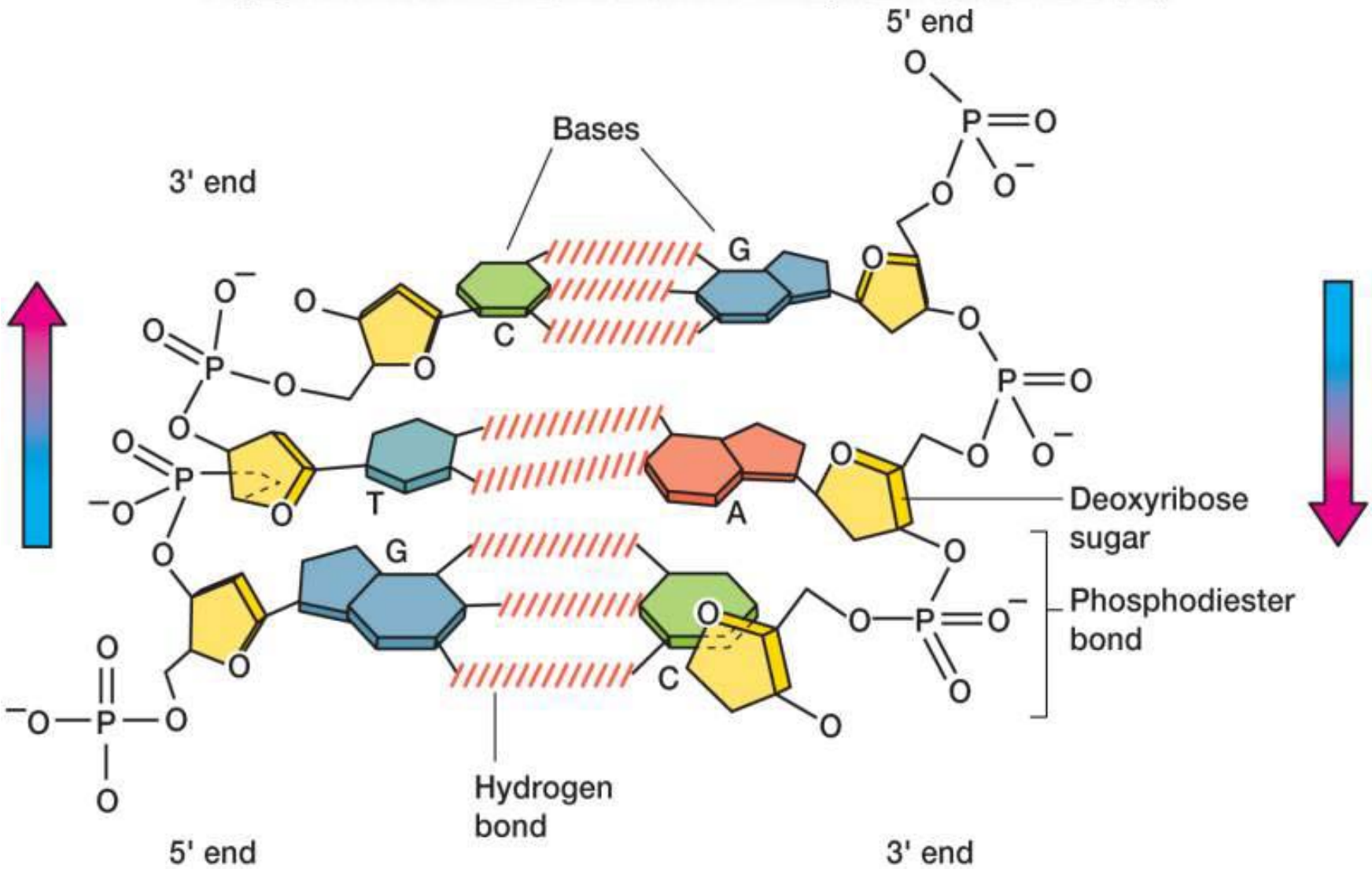


Figure 11.5b

DNA Structure

- nitrogenous bases
  - A, T, G, C
- pentose sugar
  - deoxyribose
- chain of nucleotides linked by phosphodiester bonds
- usually a double helix, composed of two complementary strands
  - base pairing rules
    - A with T
    - G with C



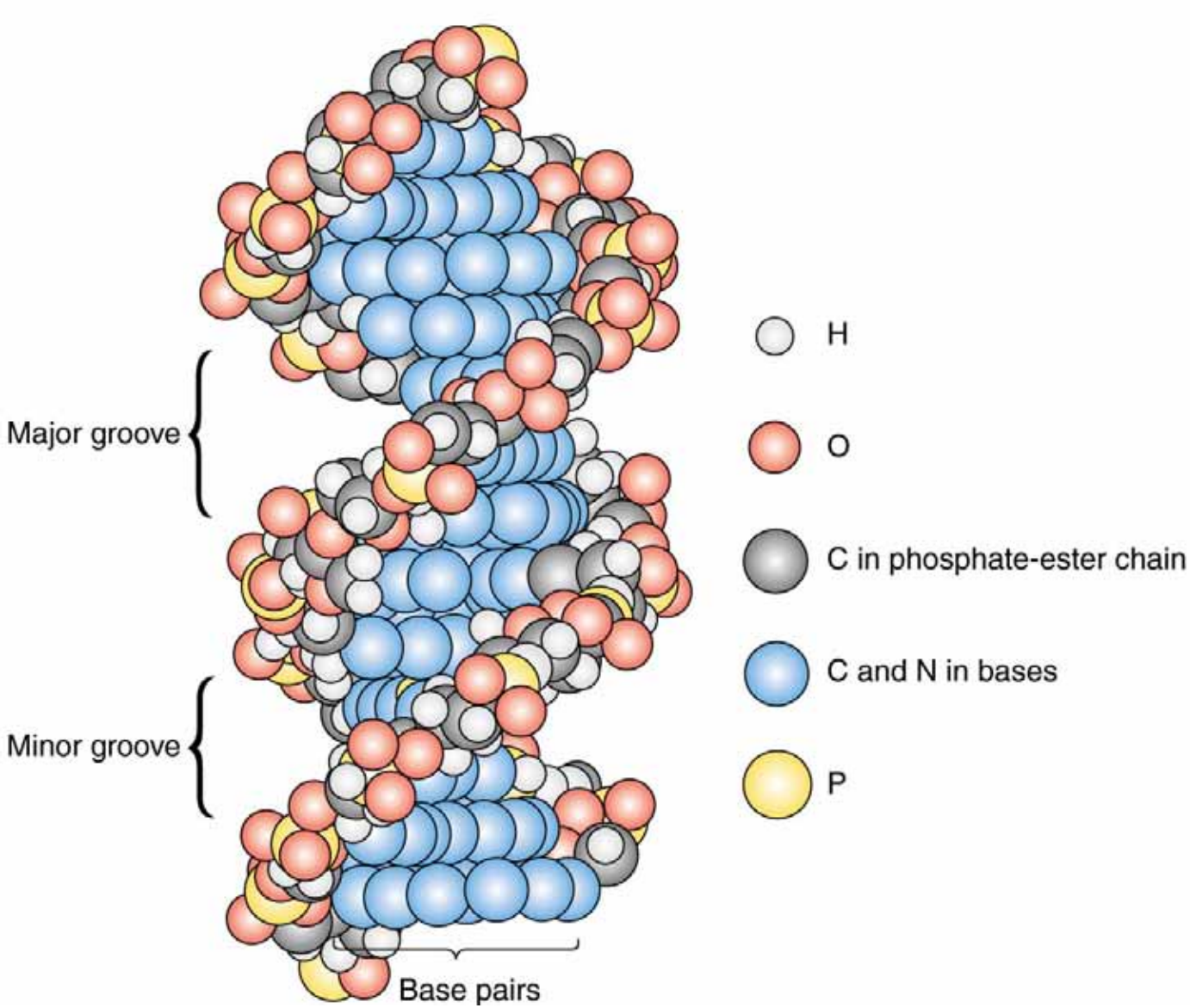

Major groove
Minor groove
Base pairs
H
O
C in phosphate-ester chain
C and N in bases
P

Figure 11.7a

two polynucleotide chains are anti-parallel

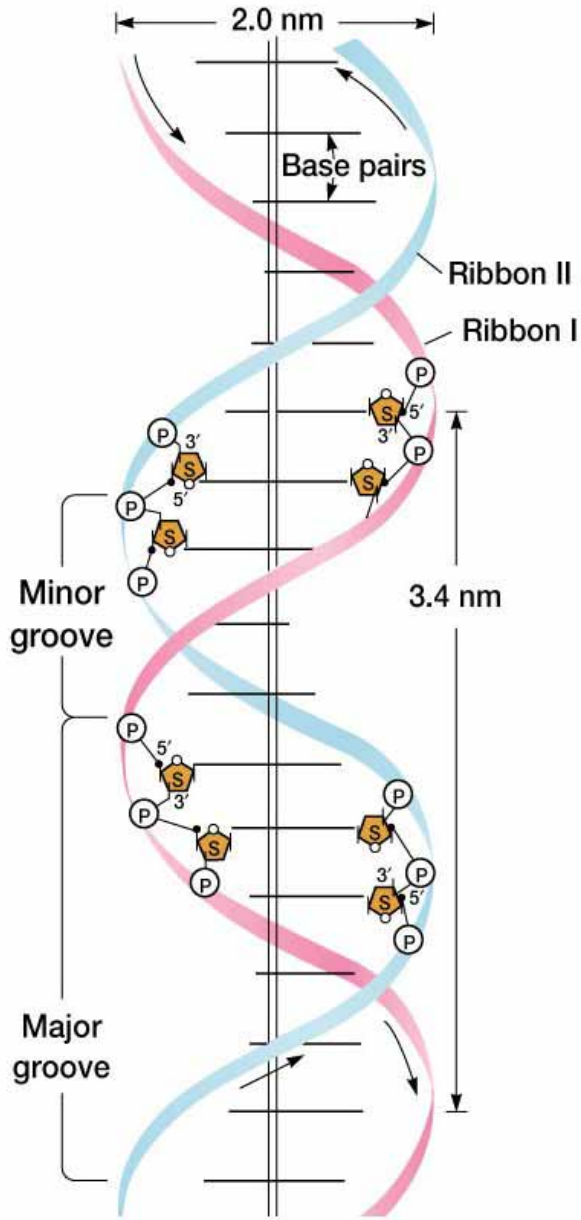


Figure 11.7b

RNA Structure

- nitrogenous bases
  - A, G, C, U (instead of T)
- pentose sugar
  - ribose
- usually consists of single strand of nucleotides linked by phosphodiester bonds
  - can coil back on itself
    - forms hairpin-shaped structures with complementary base pairing and helical organization
    - base pairing rules
      - A with U
      - G with C

Types of RNA

- three types
  - ribosomal RNA (rRNA)
  - transfer RNA (tRNA)
  - messenger RNA (mRNA)
- differ from each other in function, site of synthesis in eucaryotic cells, and structure

The Organization of DNA in Cells

- organization differs in two cell types

Procaryotic DNA

- usually exists as closed circular, supercoiled molecule associated with basic proteins

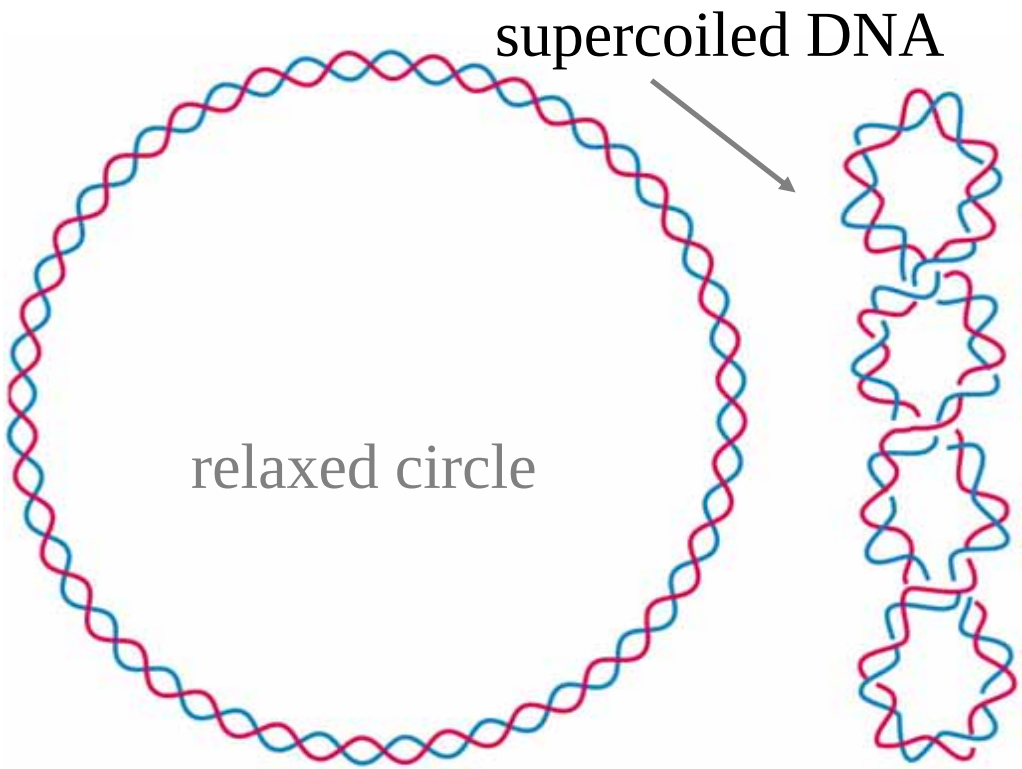


Figure 11.9

Eucaryotic DNA

- linear molecules
- associated with histones
- coiled into repeating units called nucleosomes

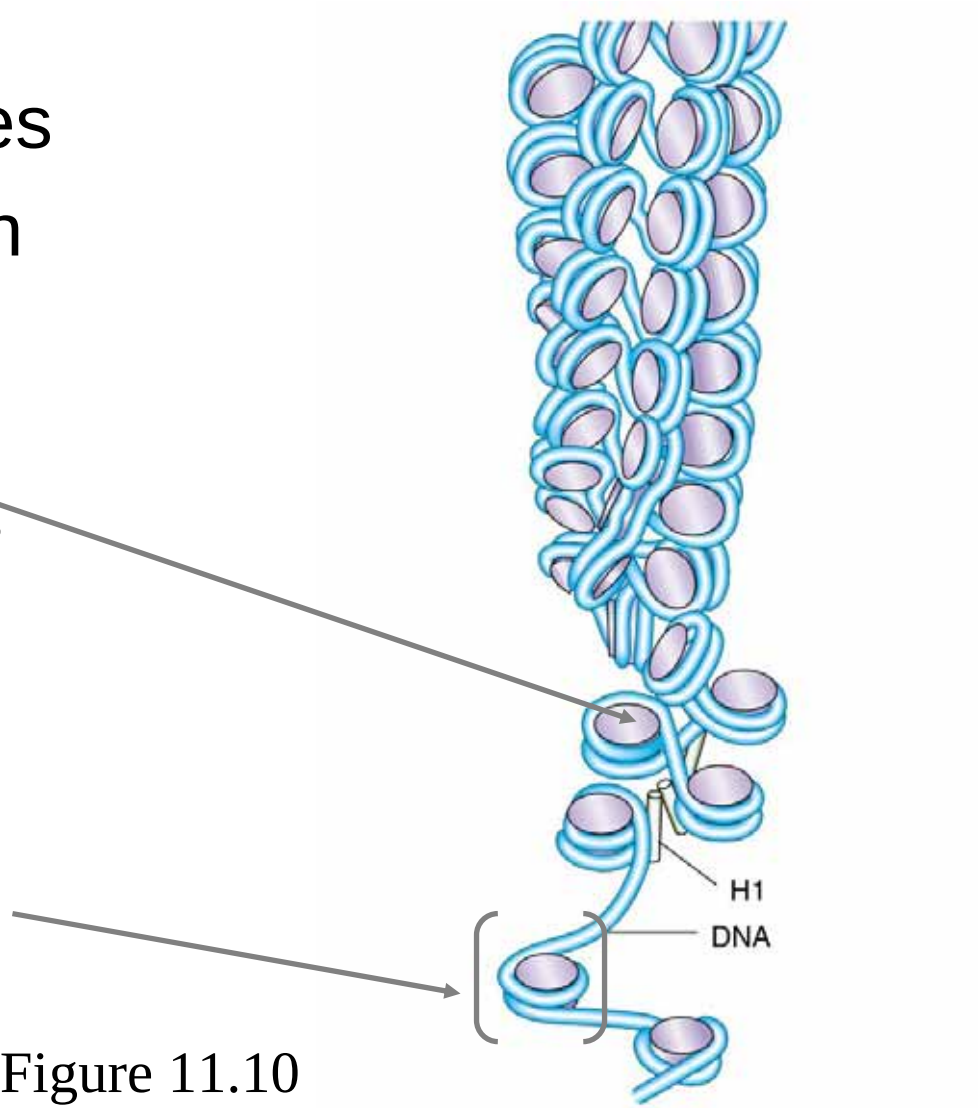


Figure 11.10

DNA Replication

- complex process involving numerous enzymes and proteins
- in general, process is similar in all organisms

Patterns of DNA Synthesis

- semiconservative
  - each parental strand is conserved
  - two parental strands separate and serve as templates for synthesis of new strands

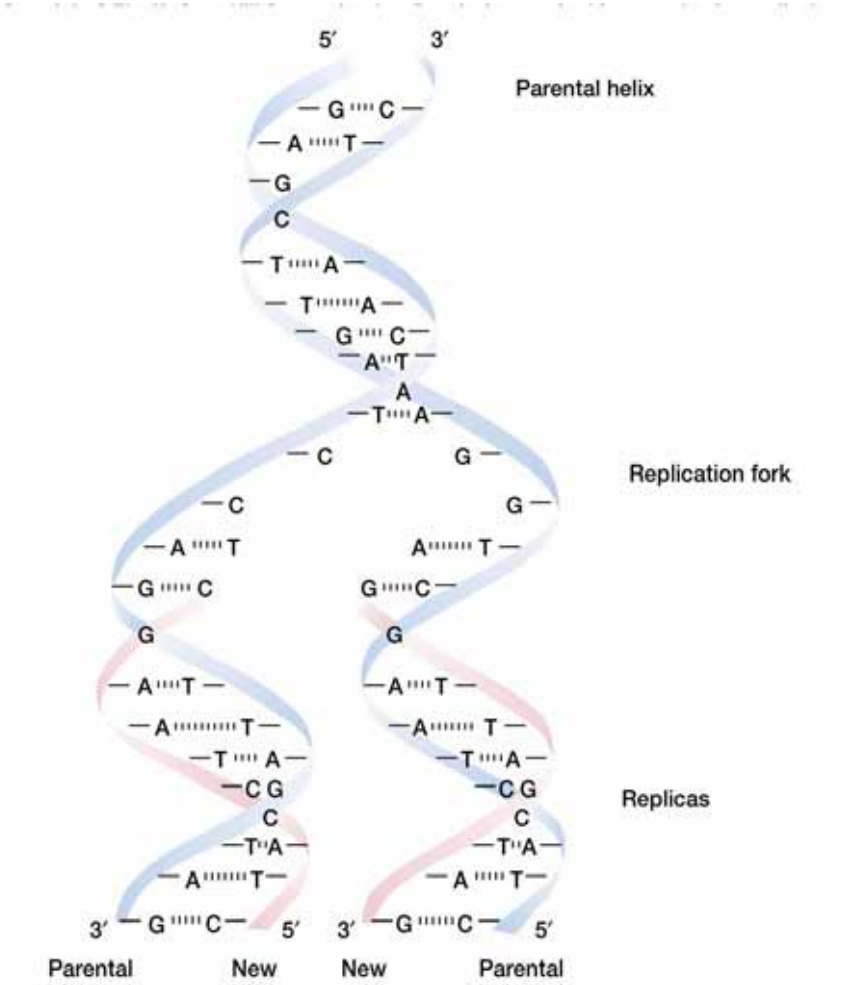


Figure 11.11

Patterns of DNA synthesis…

- in procaryotes
  - bidirectional from a single origin of replication
  - replicon
    - portion of the genome that contains an origin and is replicated as a unit

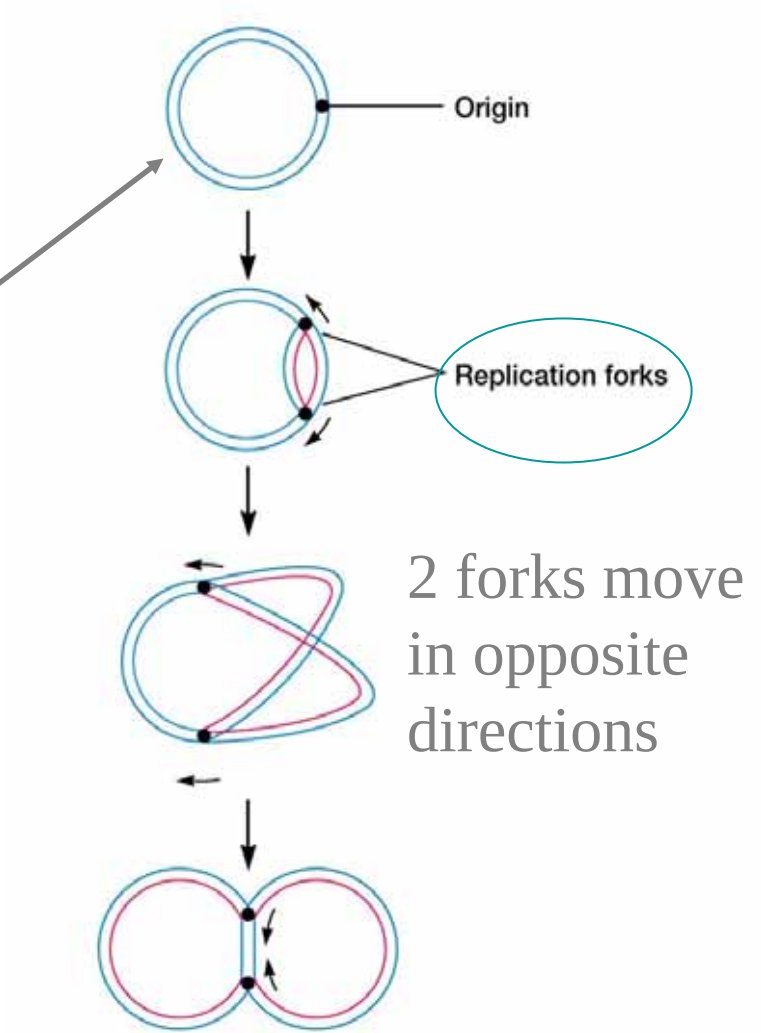


Figure 11.12

Patterns of DNA synthesis…

- in eucaryotes
  - bidirectional
  - multiple origins of replication

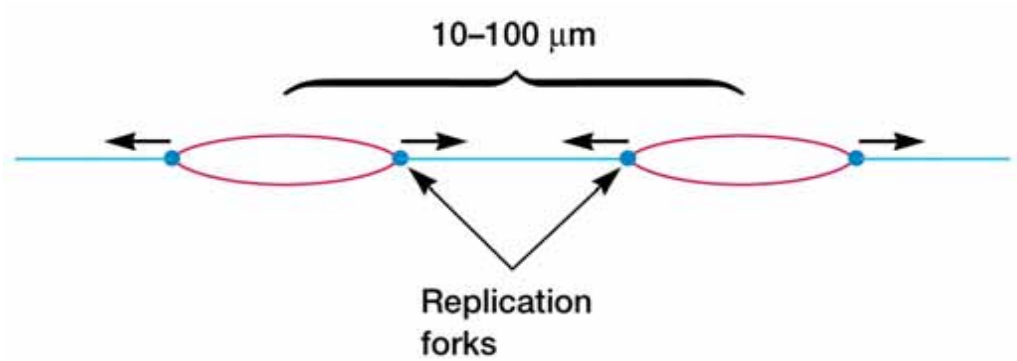


Figure 11.14

Patterns of DNA synthesis…

- some small circular genomes (e.g., viruses and plasmids
  - replicated by rolling-circle mechanism

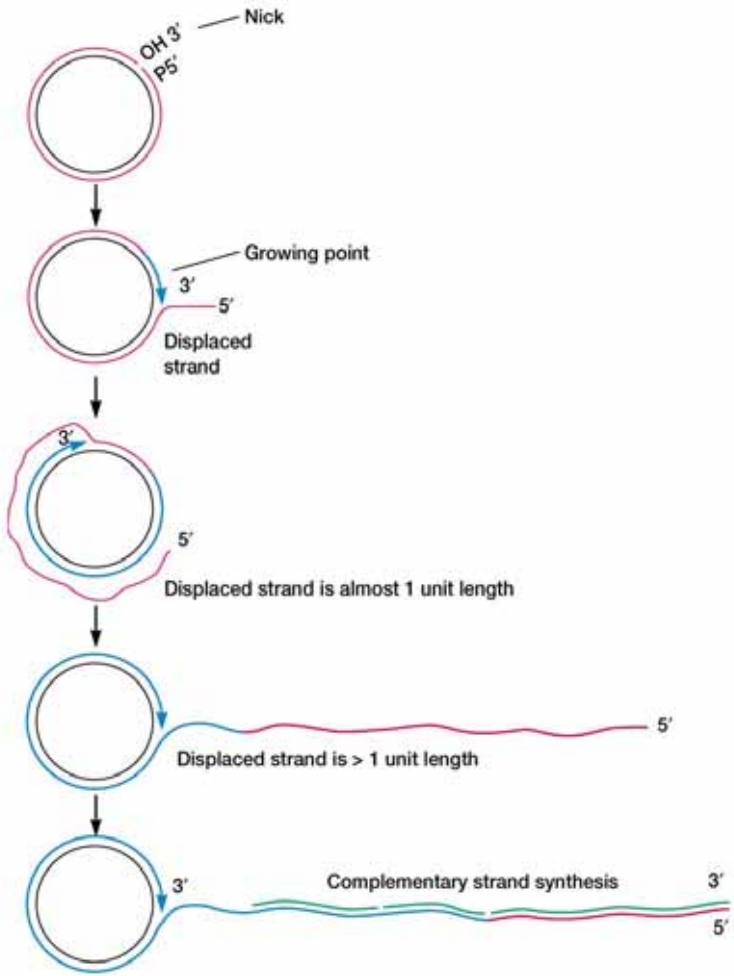


Figure 11.13

Mechanism of DNA Replication

in bacteria

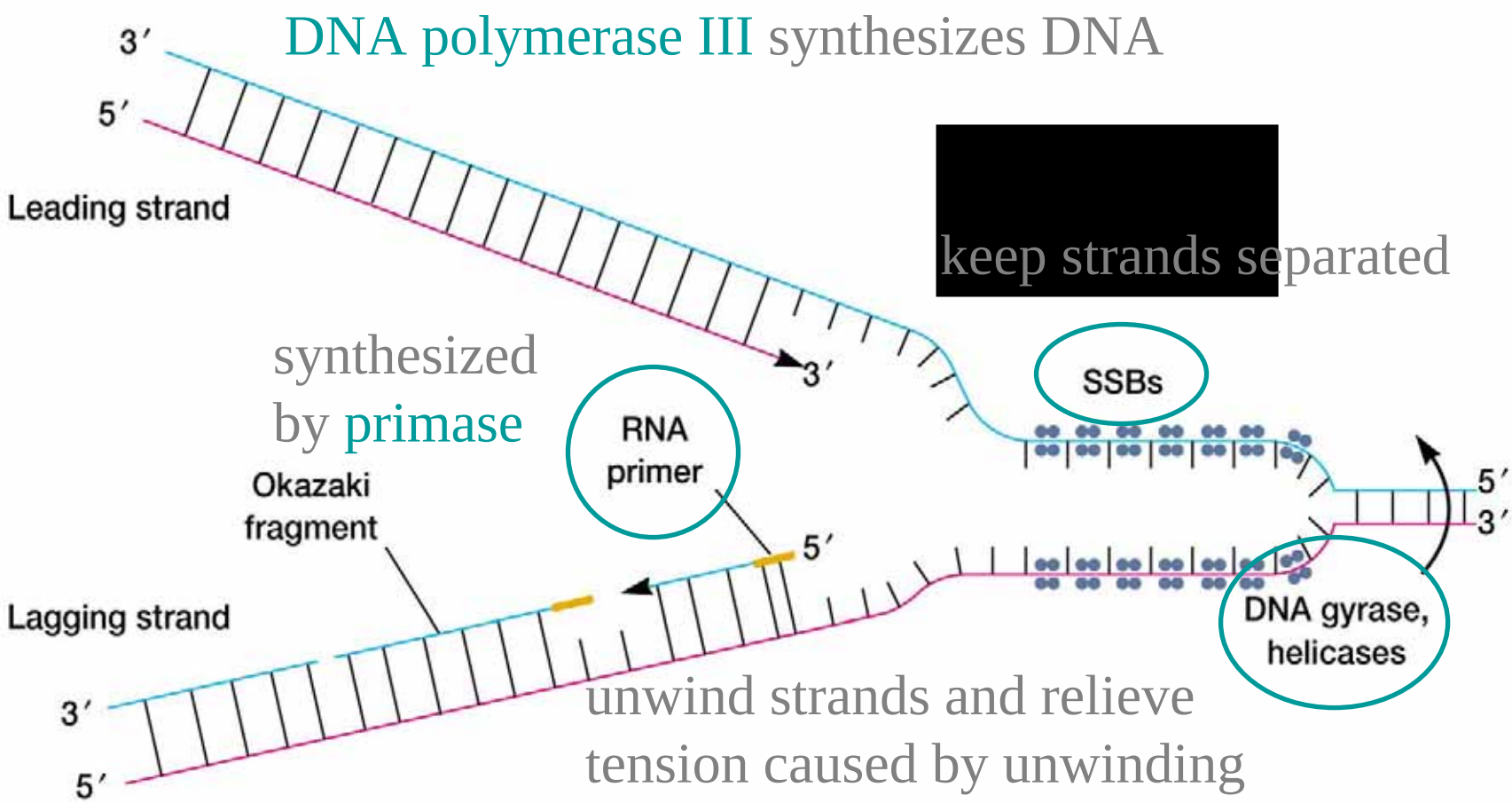


Figure 11.16

DNA polymerase III

- uses each strand as template and synthesizes complementary strands

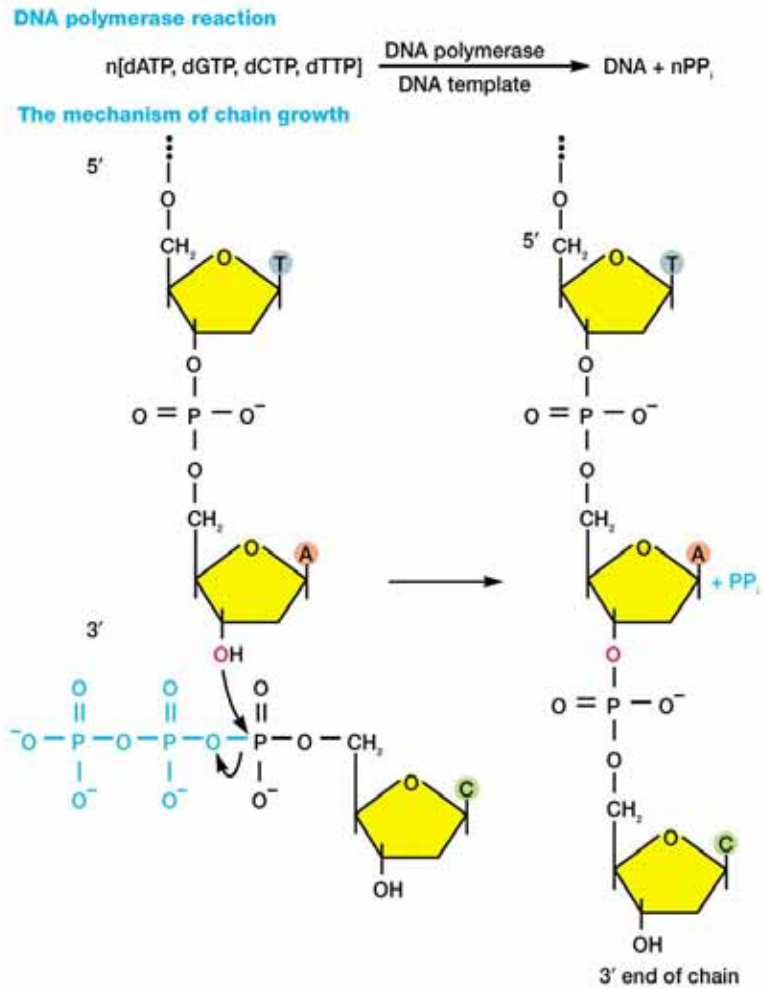


Figure 11.15

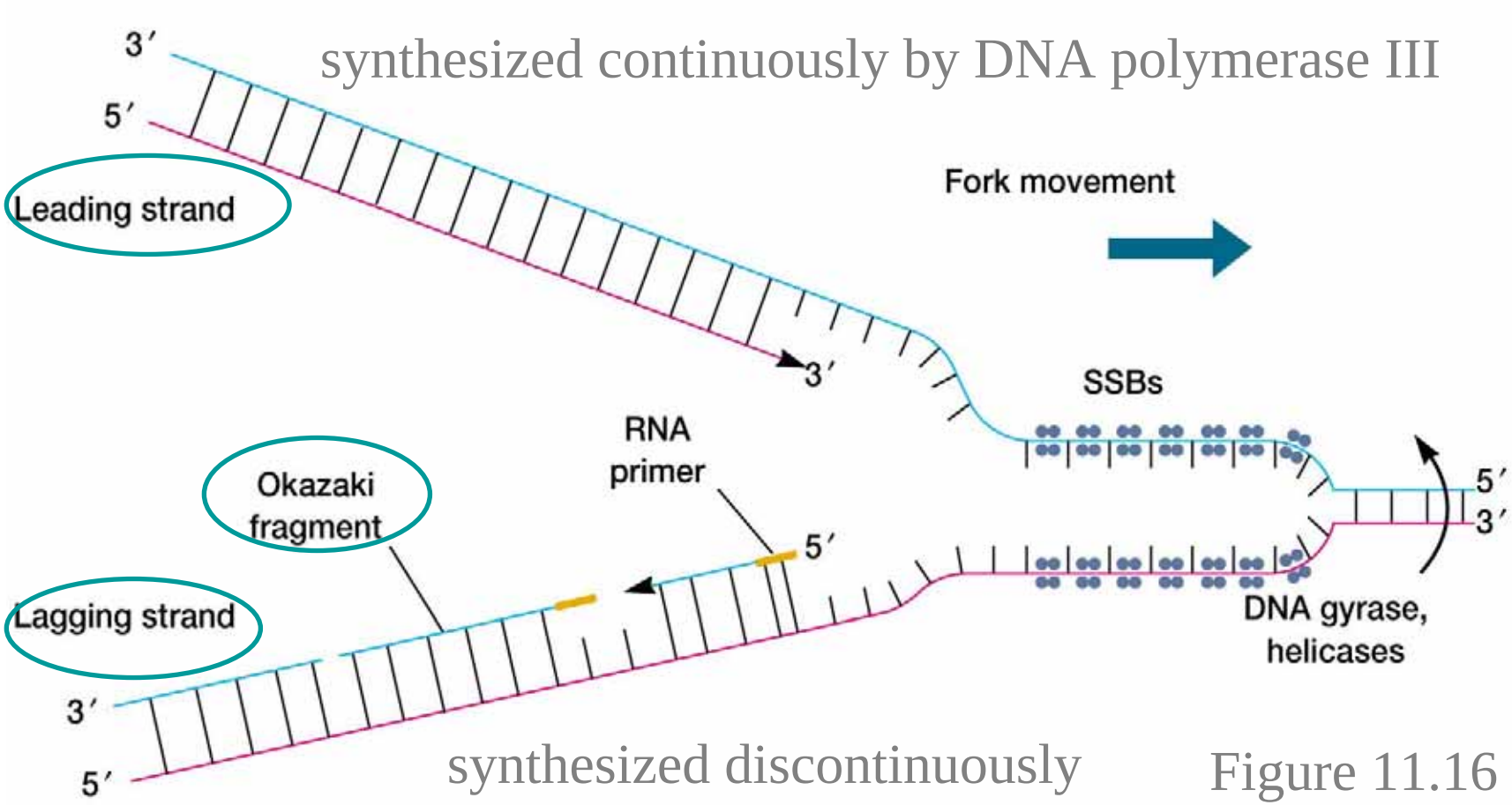


Figure 11.16

DNA polymerase I removes primers and fills gaps

DNA ligase joins fragments to form complete strands of DNA

DNA ligase reaction

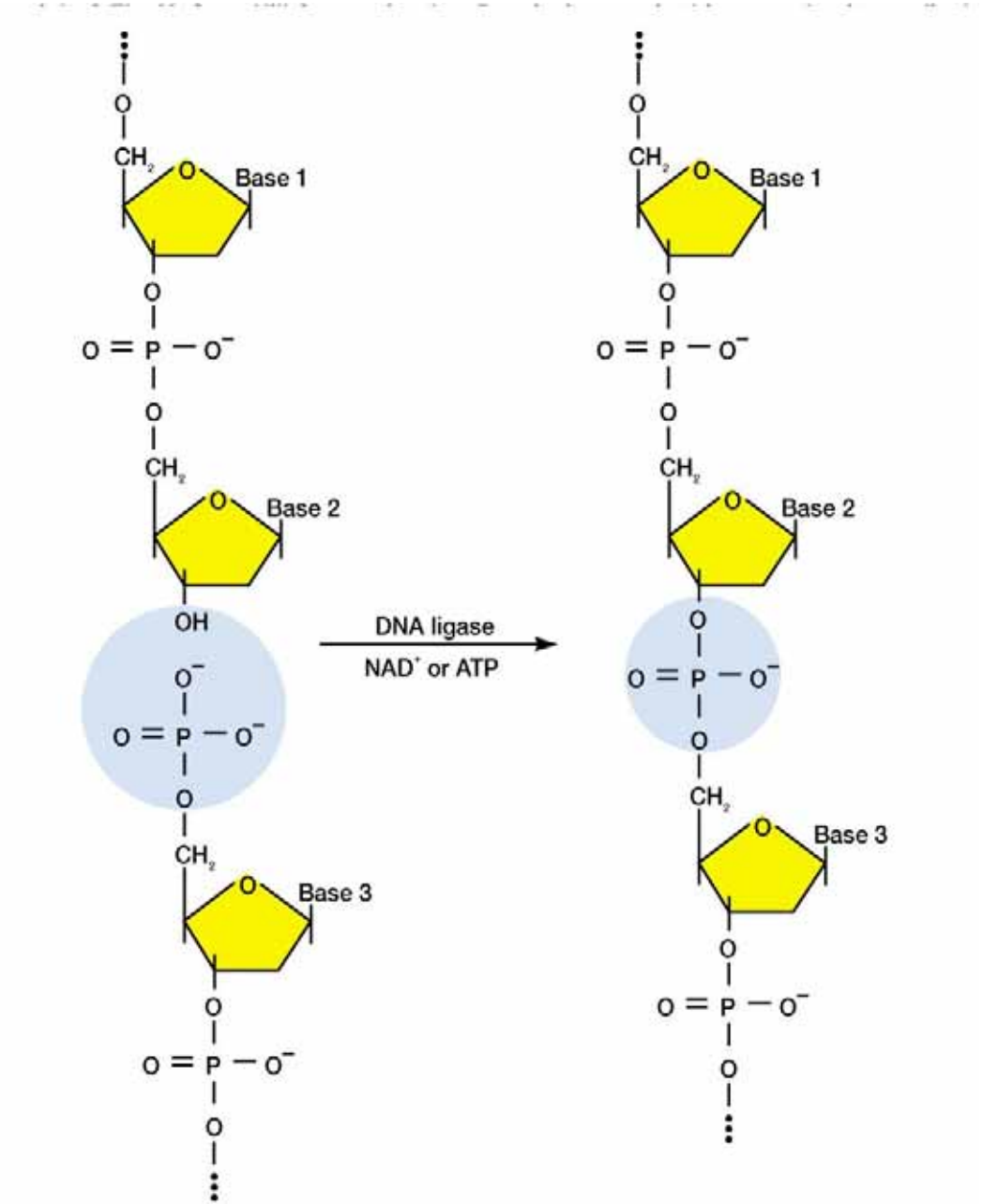


Figure 11.18

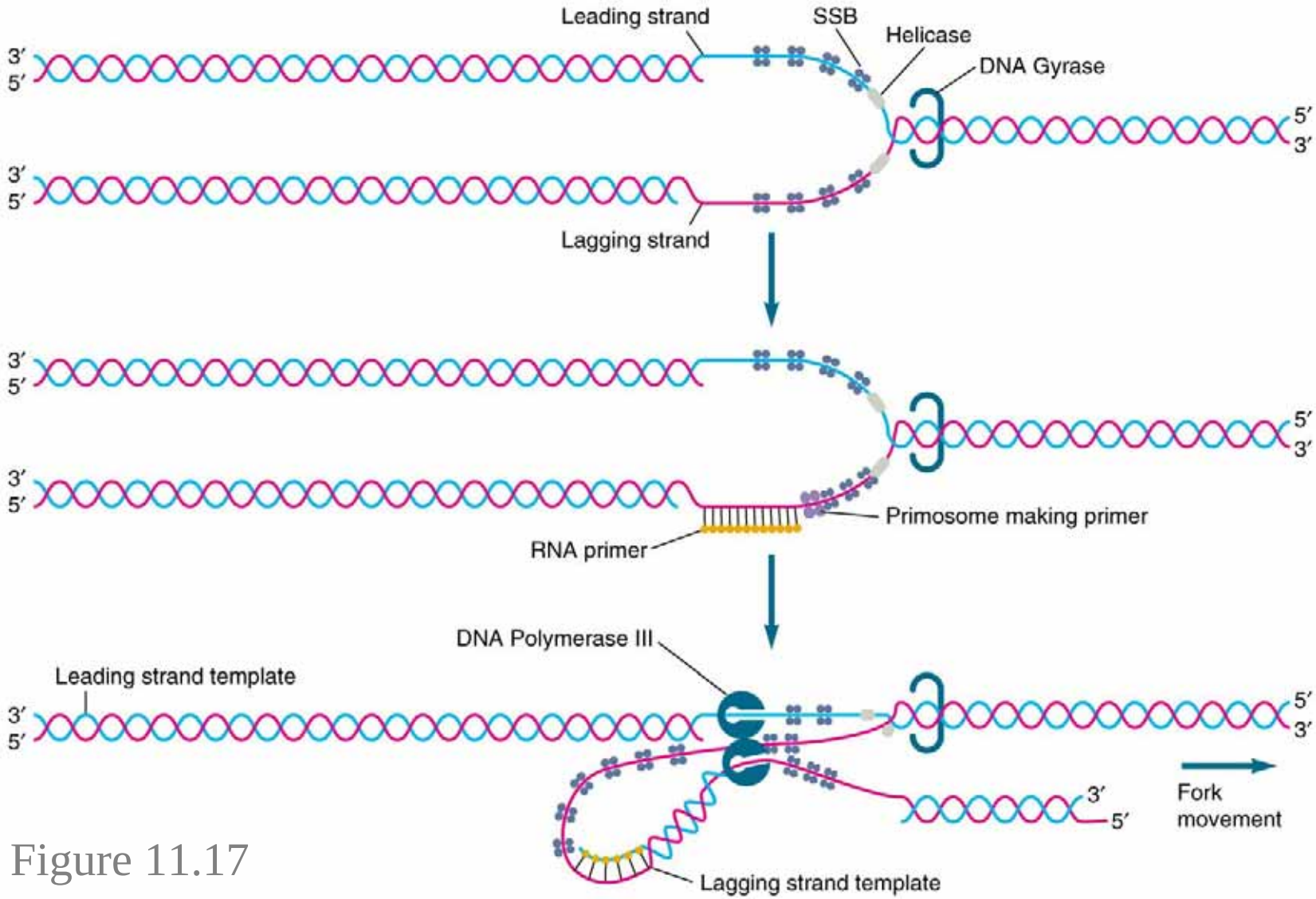

Leading strand
SSB
Helicase
DNA Gyrase
3′
5′
5′
3′
Lagging strand
Primosome making primer
RNA primer
DNA Polymerase III
Leading strand template
Fork
movement
Lagging strand template

Figure 11.17

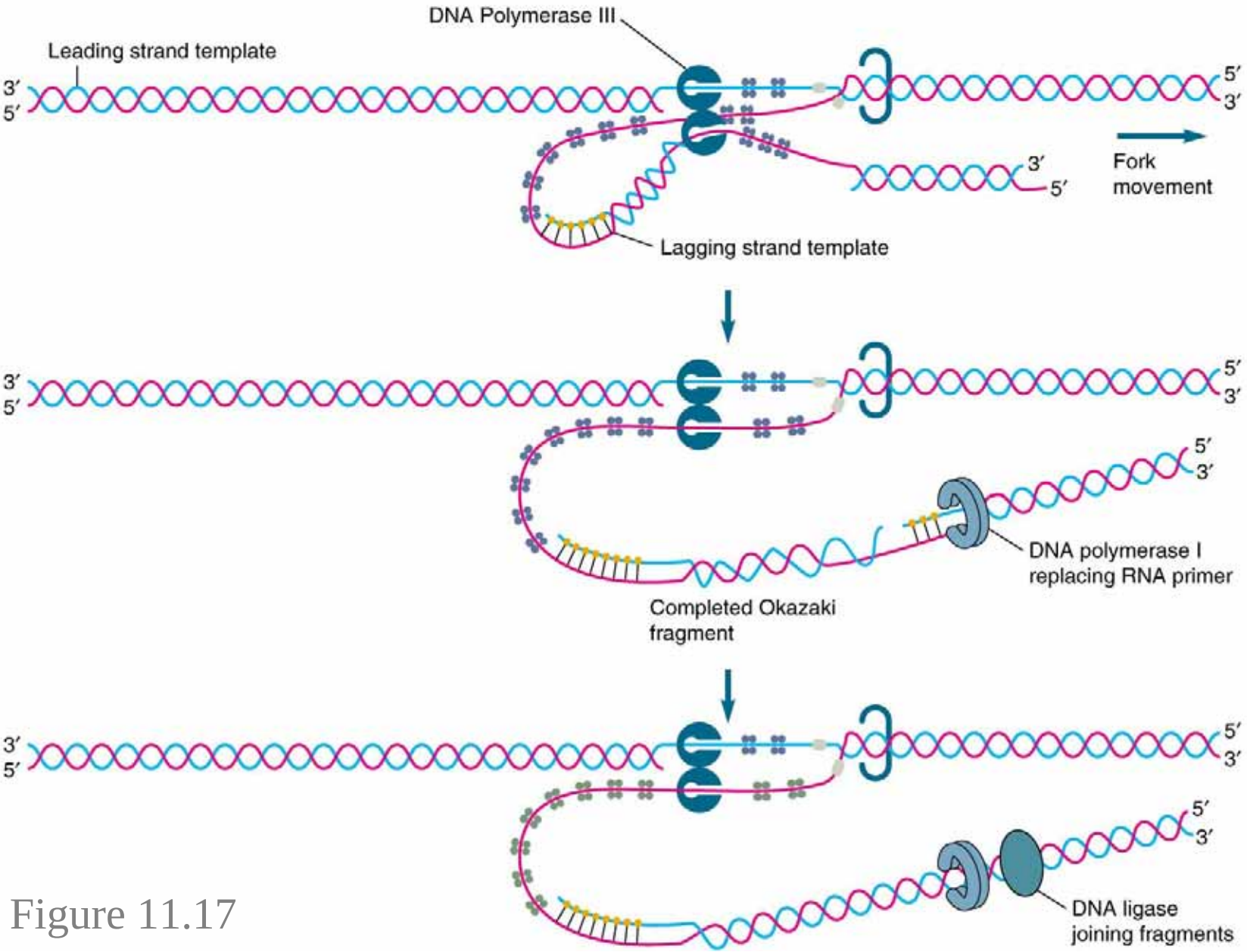

DNA Polymerase III
Leading strand template
3′
5′
5′
3′
3′
5′
Fork
movement
Lagging strand template
3′
5′
5′
3′
5′
3′
DNA polymerase I
replacing RNA primer
Completed Okazaki
fragment
3′
5′
5′
3′
5′
3′
DNA ligase
joining fragments

Figure 11.17

Some amazing facts

- ≥ 30 proteins required to replicate *E. coli* chromosome
- occurs with great fidelity
  - error frequency = 10^{-9} or 10^{-10} per base pair replicated
  - due to proofreading activity of DNA polymerases III and I
- occurs very rapidly
  - 750 to 1,000 base pairs/second in procaryotes
  - 50-100 base pairs/second in eucaryotes

The Genetic Code

- the manner in which genetic instructions for polypeptide synthesis are stored within genome
- colinearity
  - sequence of base pairs in DNA corresponds to the amino acid sequence of polypeptide encoded

Establishment of Genetic Code

- codon
  - genetic code word
  - specifies an amino acid
- codon meanings deciphered by Marshall Nirenberg, et al. in 1960s

Organization of the Code

- code degeneracy
  - up to six different codons can code for a single amino acid
- sense codons
  - the 61 codons that specify amino acids
- stop (nonsense) codons
  - the three codons used as translation termination signals
  - do not encode amino acids

Table 11.1 The Genetic Code

First Position (5′ End)[a]	Second Position: U	Second Position: C	Second Position: A	Second Position: G	Third Position (3′ End)
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA STOP	UGA STOP	A
	UUG Leu	UCG Ser	UAG STOP	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G

[a]The code is presented in the RNA form. Codons run in the 5′ to 3′ direction. See text for details.

Wobble

- loose base pairing
  - 3rd position of codon less important than 1st or 2nd
- eliminates need for unique tRNA for each codon

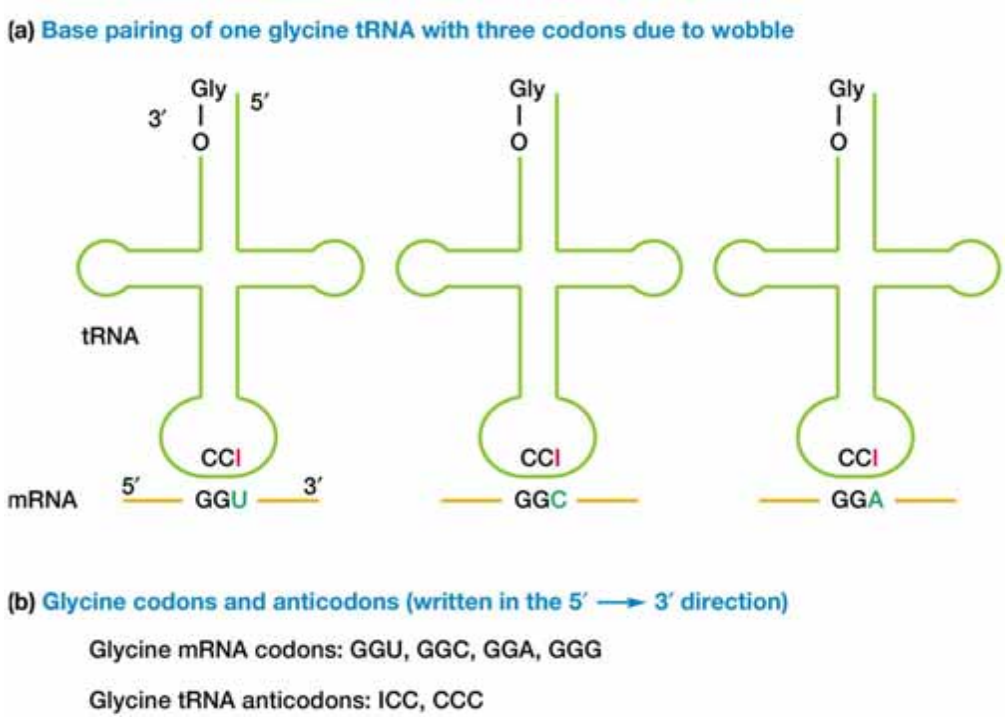


Figure 11.19

Gene Structure

- gene
  - linear sequence of nucleotides with a fixed start point and end point
  - encodes a polypeptide, a tRNA, or an rRNA
    - cistron – gene that encodes a polypeptide
- reading frame
  - organization of codons such that they can be read to give rise to a gene product

Importance of reading frame

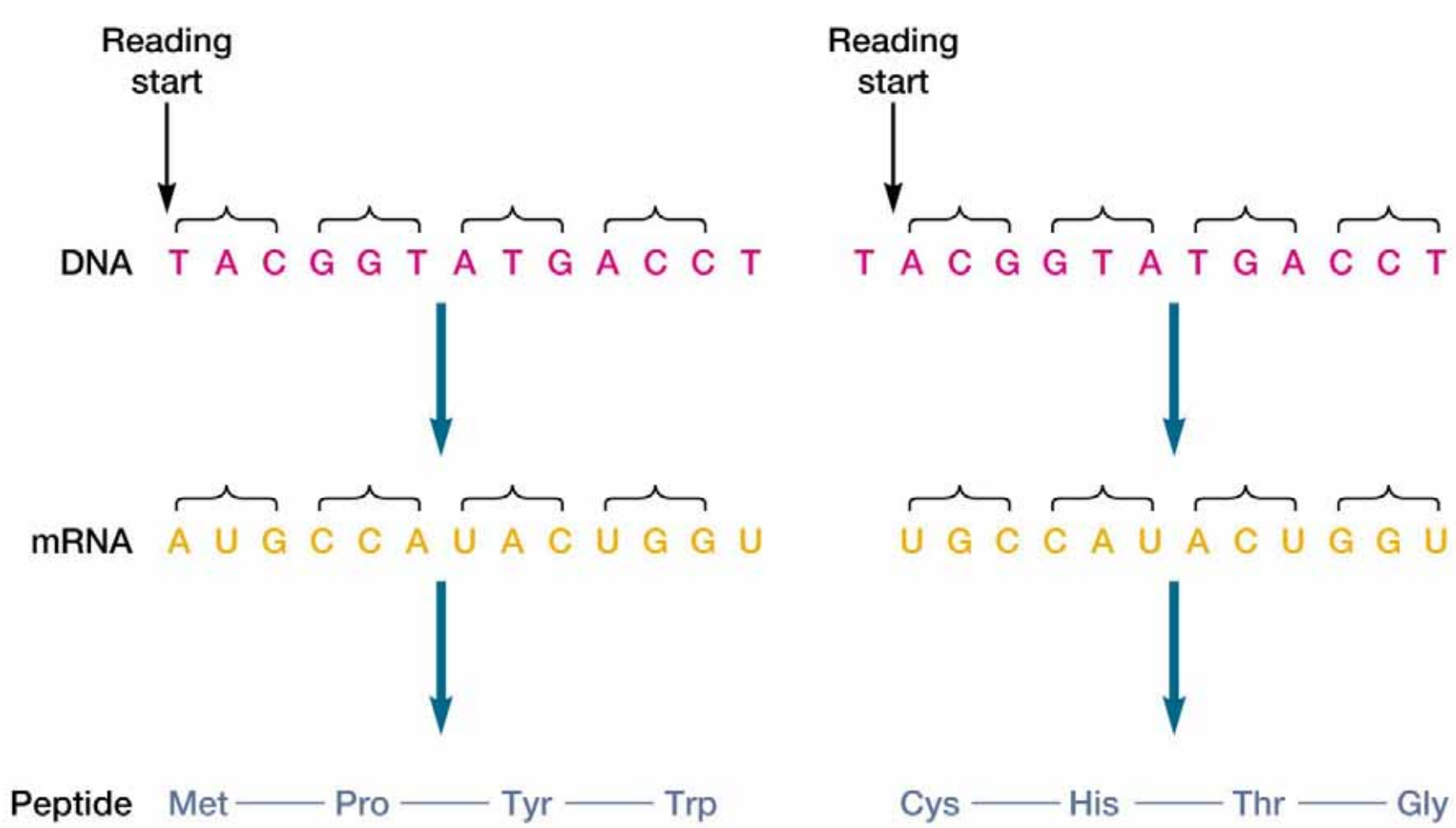


Figure 11.20

Organization of genes on chromosomes

- for most organisms, reading frames do not overlap

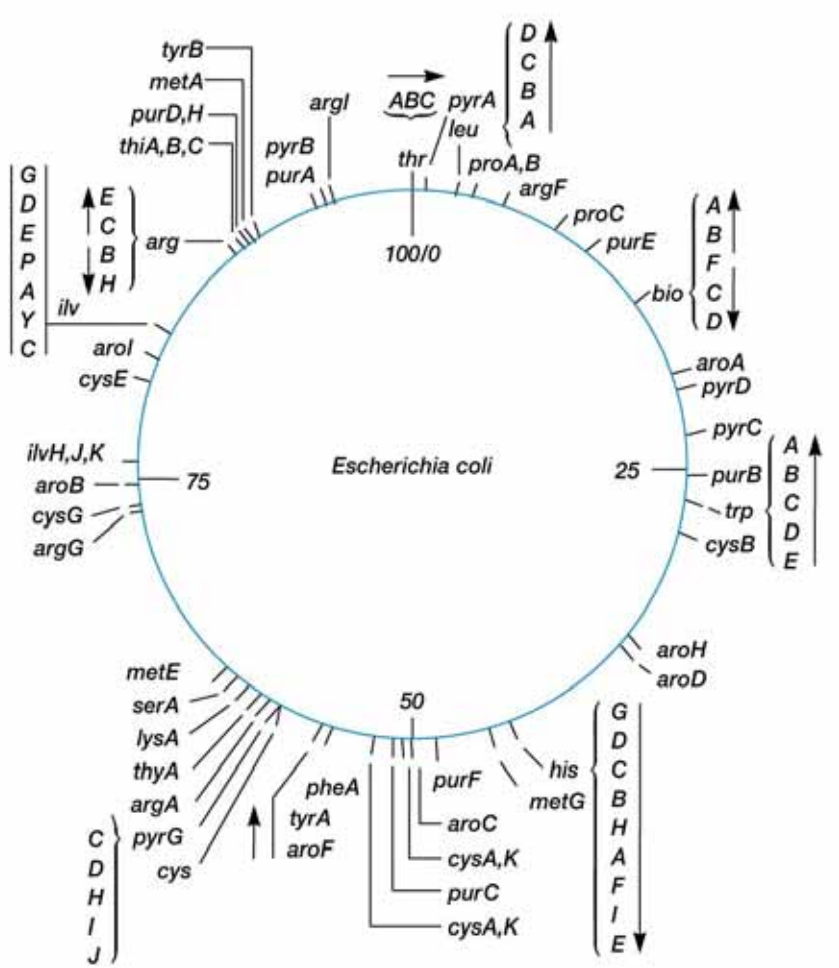


Figure 11.21a

• some bacteria and some viruses have overlapping genes

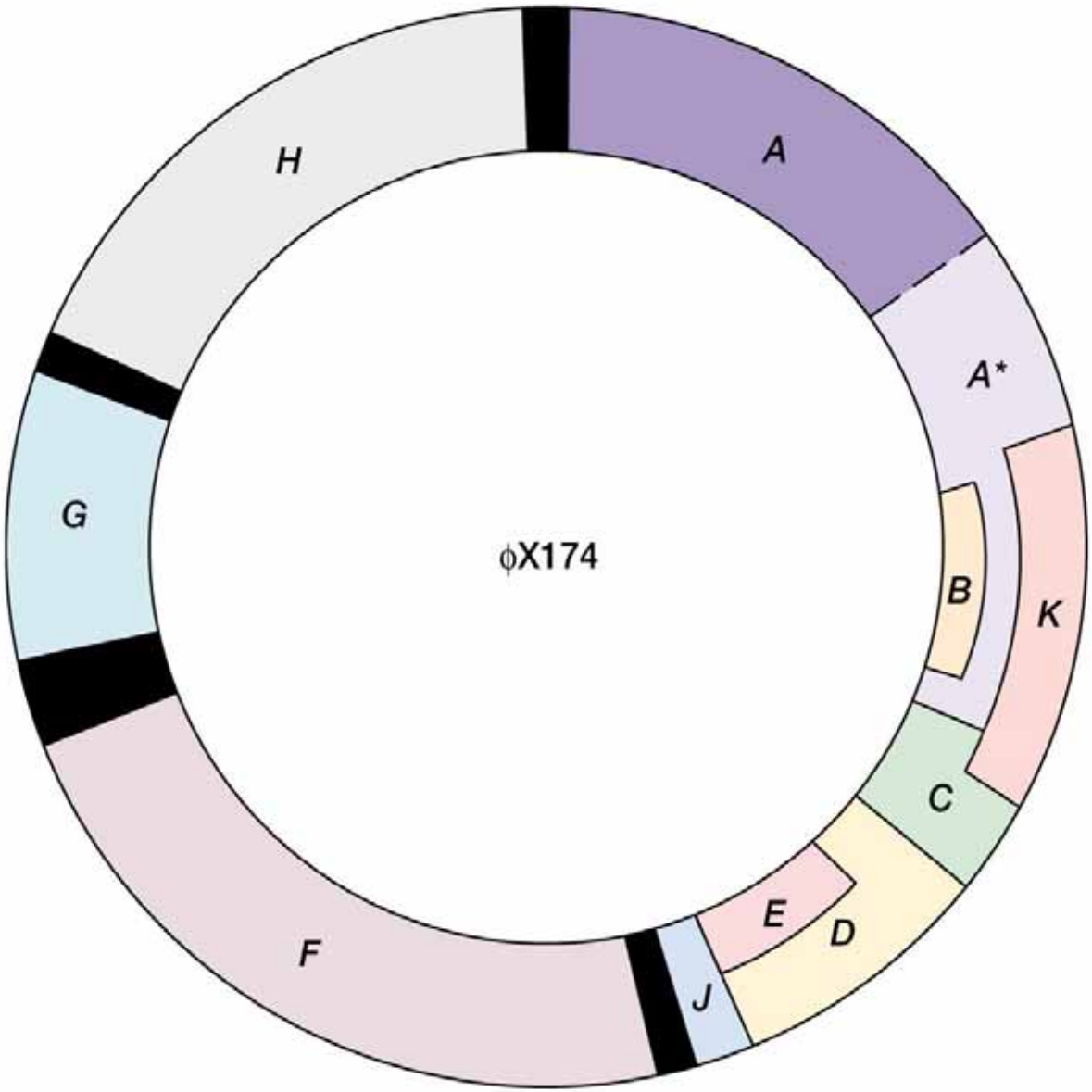


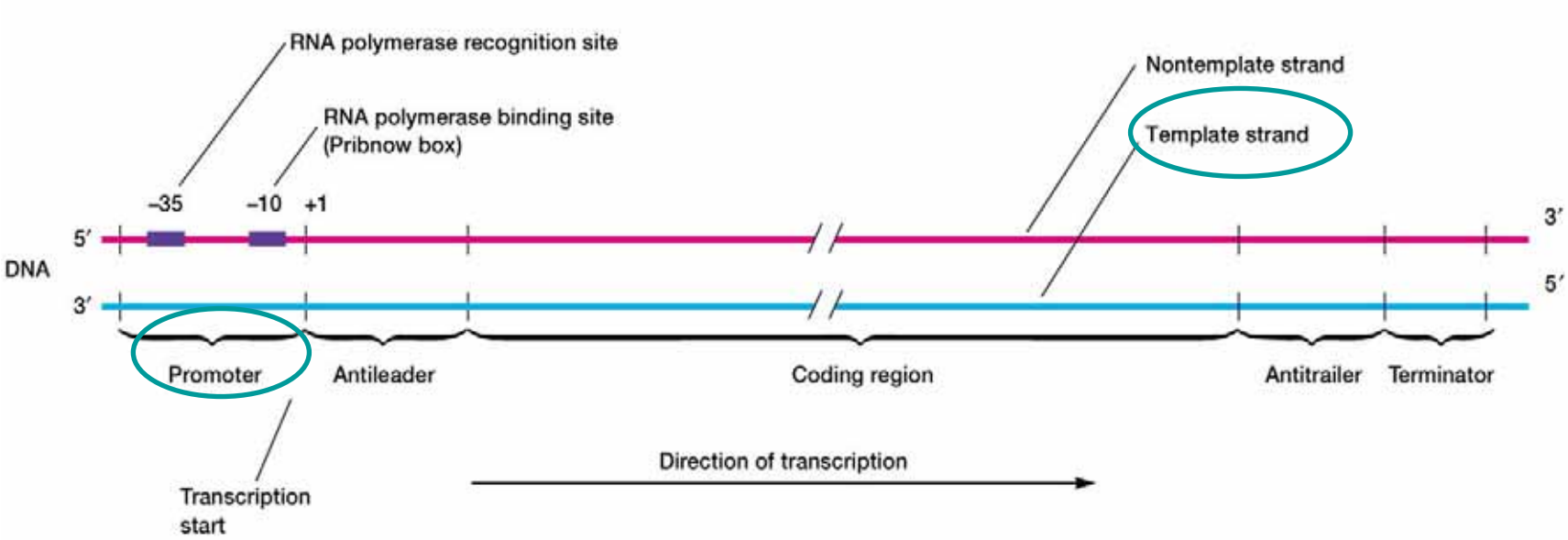
Figure 11.21b

Procaryotic versus eucaryotic genes

- procaryotes (and viruses)
  - coding information is usually continuous
- eucaryotes
  - most genes have coding sequences interrupted by noncoding sequences
    - exons – coding sequences
    - introns – noncoding sequences

Genes That Code for Proteins

strand that contains coding information and directs RNA synthesis



serves as recognition and binding site for RNA polymerase

Figure 11.22

Bacterial promoters

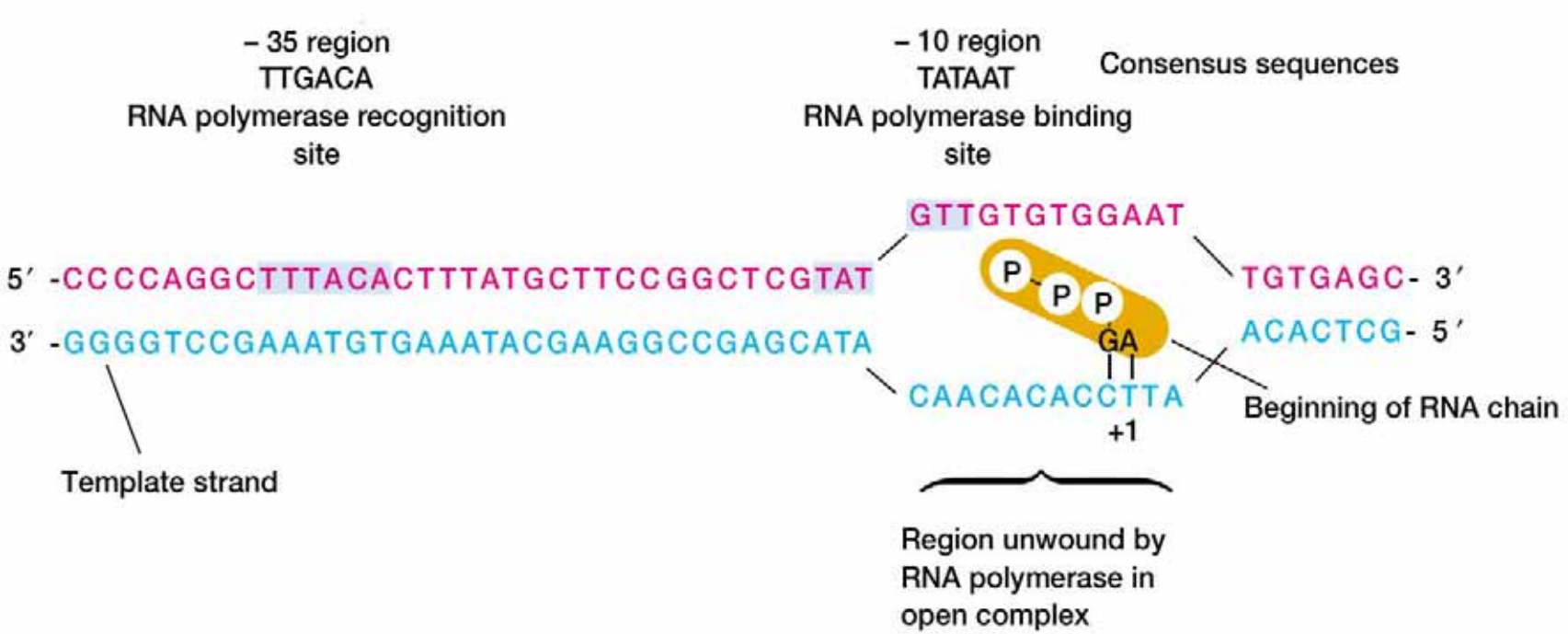


Figure 11.23

also called
Pribnow box

region that specifies sequence of amino acids in a polypeptide

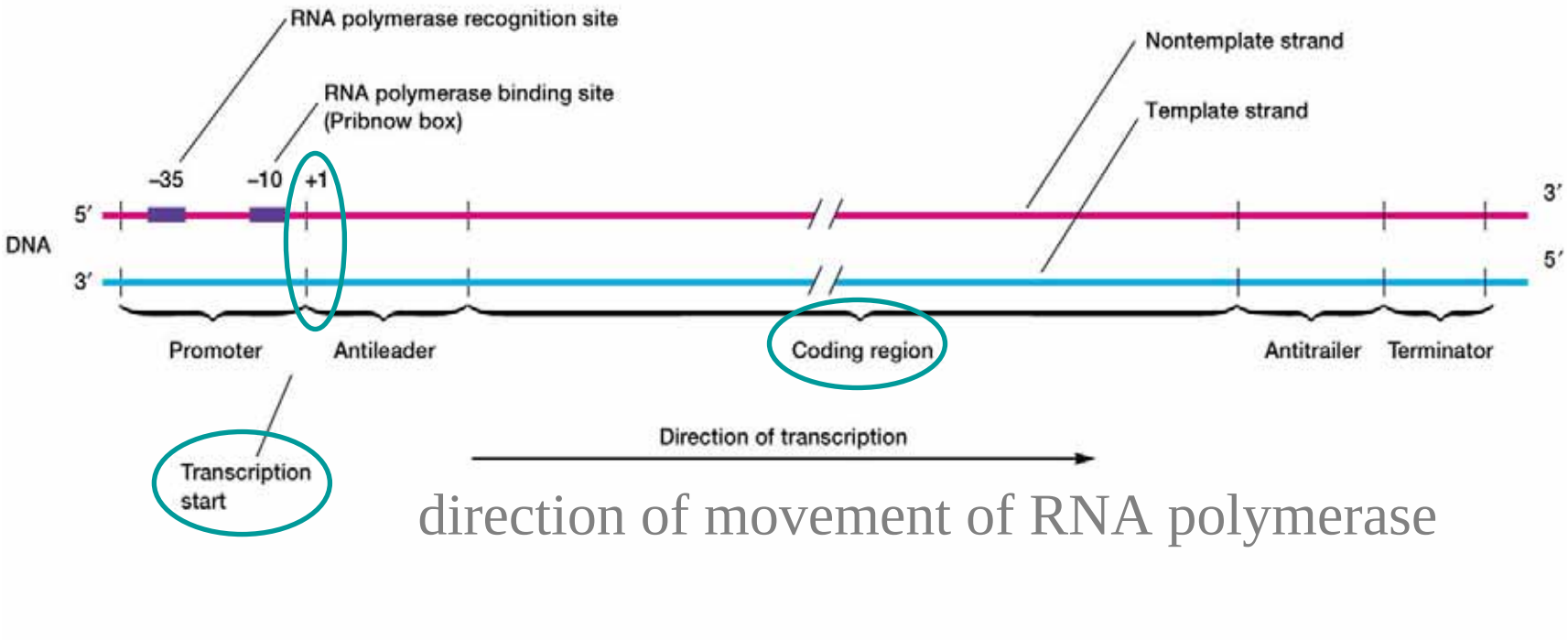
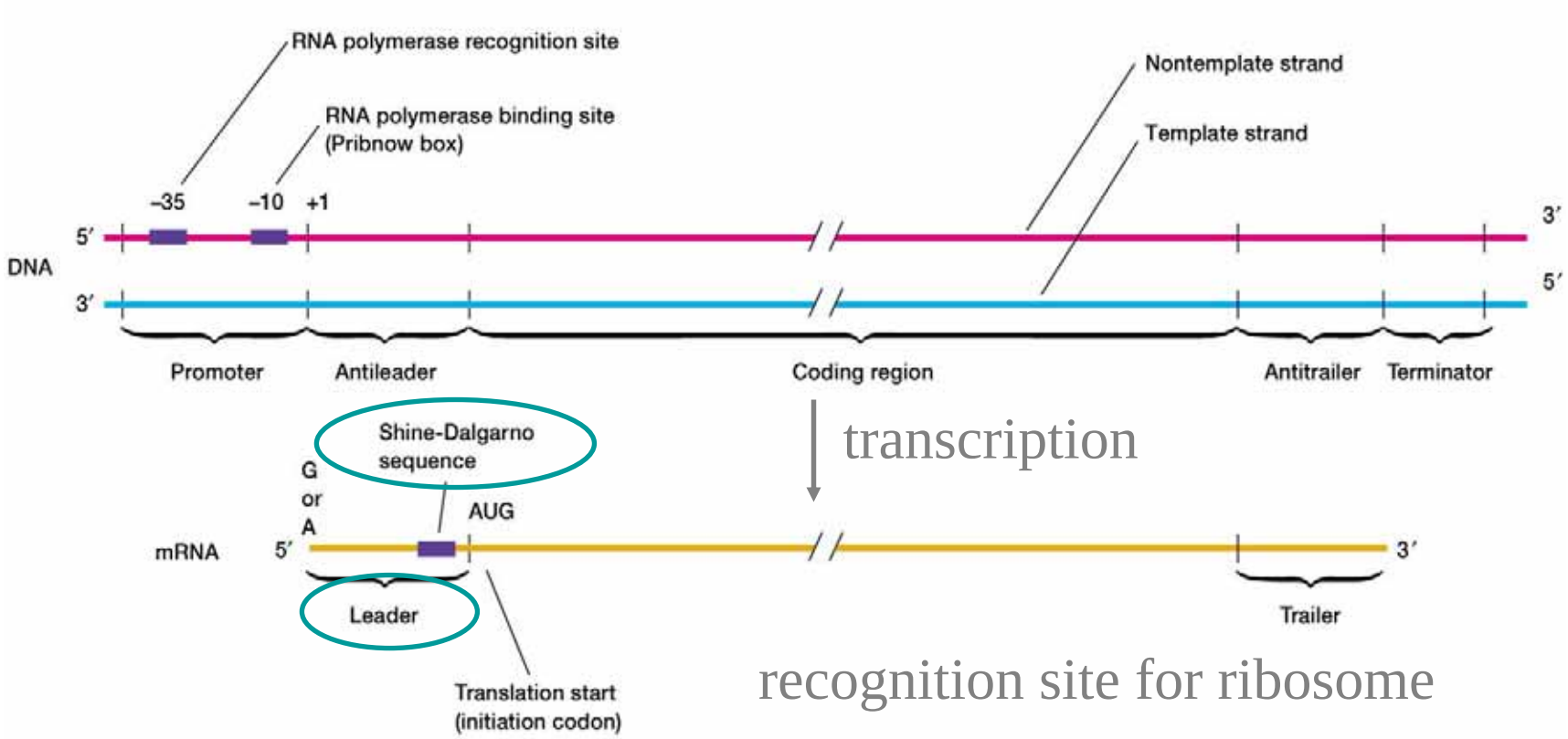


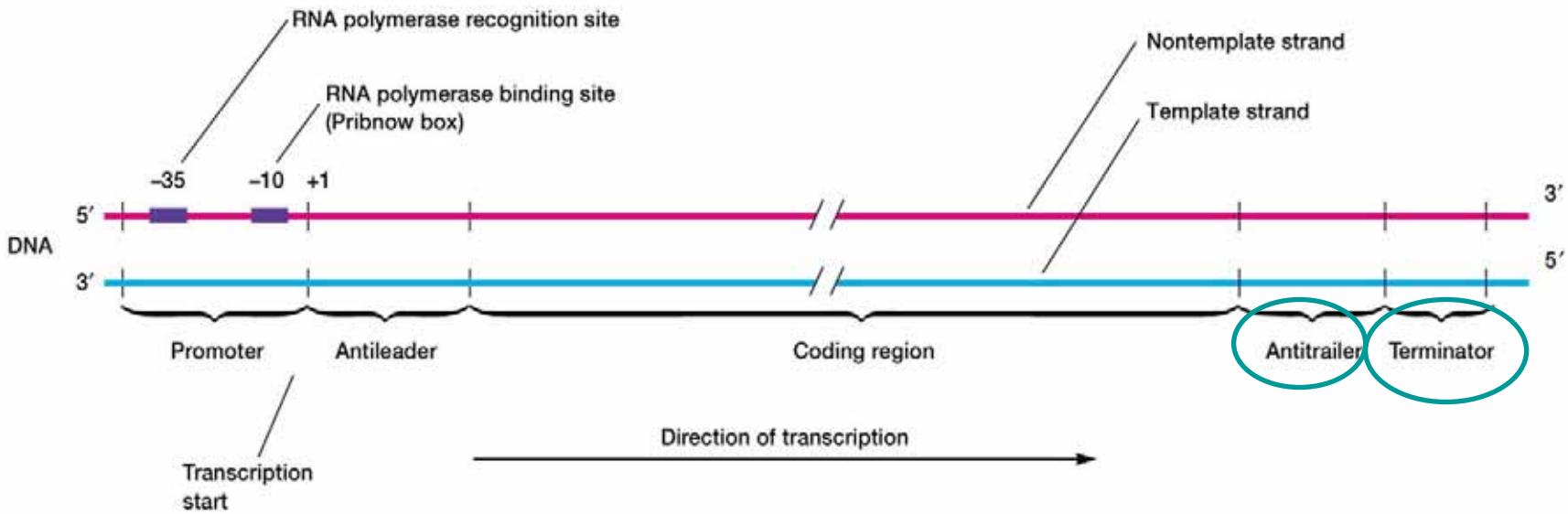
Figure 11.22



sequence that is not translated

Figure 11.22

signal for termination of transcription



transcribed but mRNA sequence is not translated

Figure 11.22

Genes That Code for tRNA and rRNA

tRNA genes have promoter, leader, coding, spacer, and trailer regions

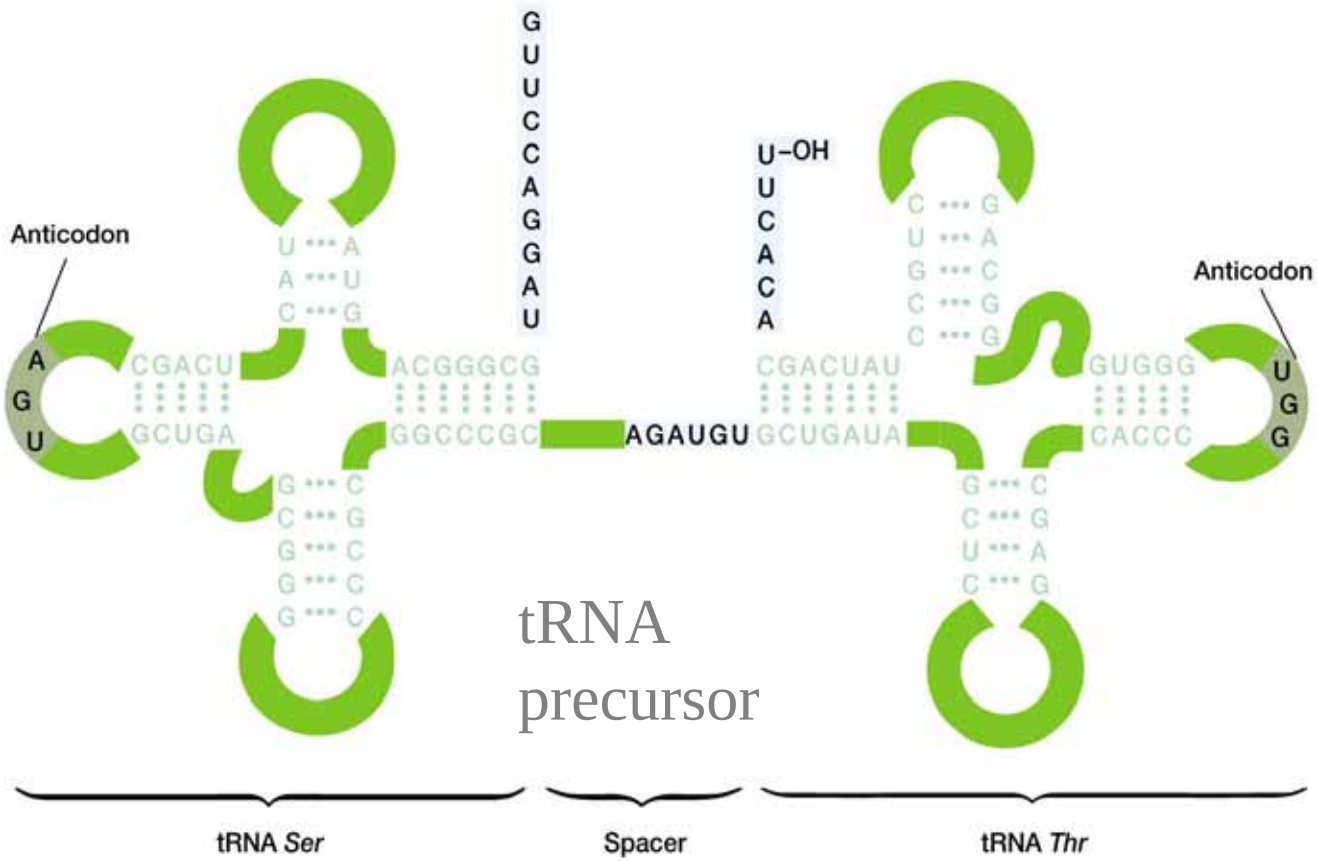


Figure 11.24a

leader, spacer, and trailer removed during maturation process

rRNA genes have promoter, leader, coding, spacer, and trailer regions

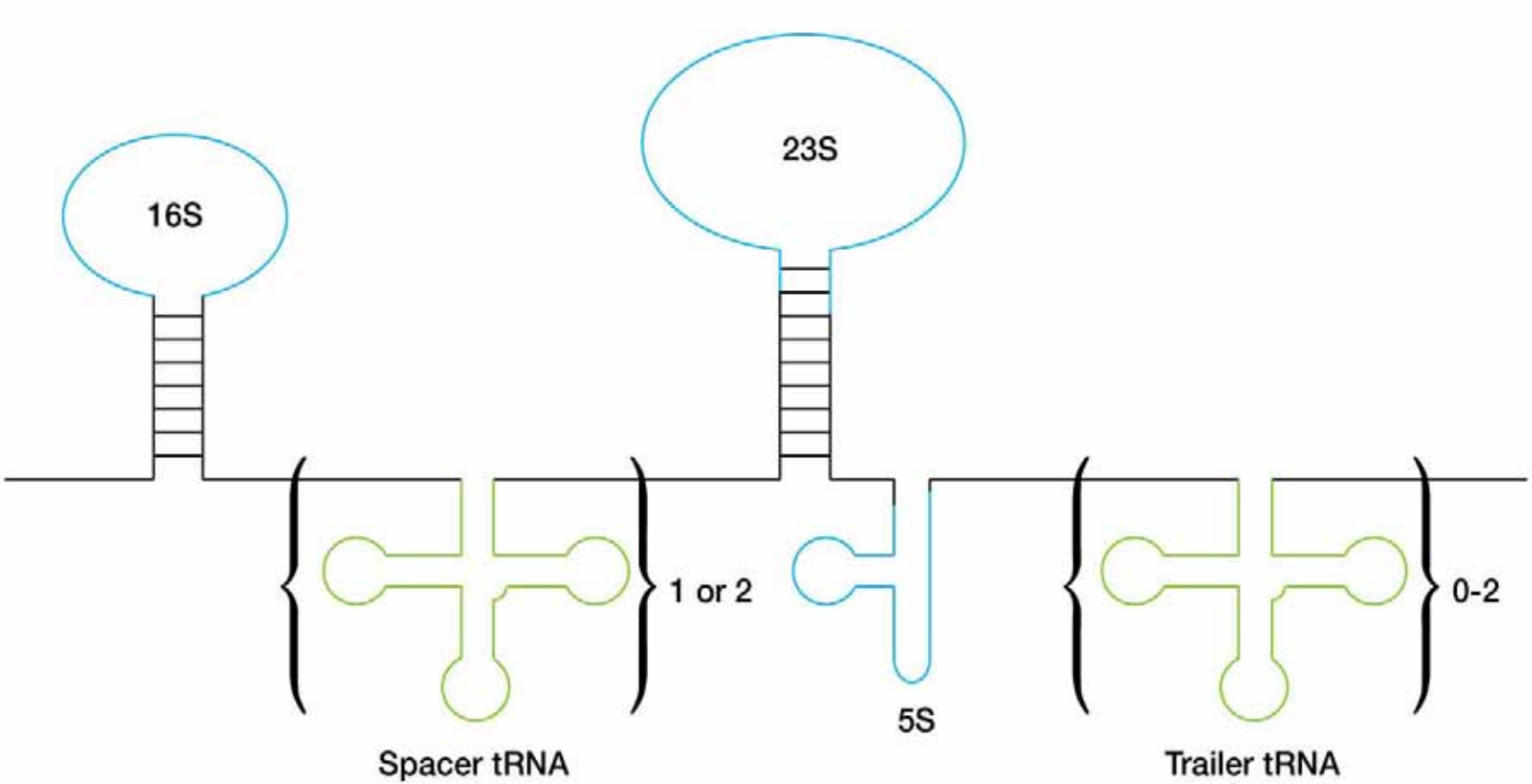


Figure 12.24b

spacer and trailer regions may encode tRNA molecules

Mutations and Their Chemical Basis

- mutations
  - stable, heritable change in nucleotide sequence
  - may or may not have an effect on the phenotype of an organism

Mutations and Mutagenesis

- mutations can be classified in terms of their effect on phenotype
  - morphological mutations
    - change colonial or cellular morphology
  - lethal mutations
    - kill the organism
  - conditional mutations
    - expressed only under certain conditions (e.g., high temperature)

Other types of mutations

- biochemical mutations
  - changes in metabolic capabilities
  - auxotrophs
    - have mutations in biosynthetic pathways
    - cannot synthesize product of pathway
    - require product of pathway as nutrient in minimal growth media
  - prototrophs
    - grow in minimal media without supplements
- resistance mutations
  - resistance to pathogen, chemical, or antibiotic

How mutations arise

- spontaneously
  - develop in absence of any added agent
  - usually thought to arise randomly
    - directed (adaptive) mutation
      - mutations that may result from hypermutation followed by selection
- induced
  - develop after exposure to a mutagen

Spontaneous Mutations

- result of:
  - errors in DNA replication
  - damage to DNA
  - insertion of transposons

Replication errors – tautomeric shifts

- keto form→ imino or enol form

- alters hydrogen-bonding characteristics of base

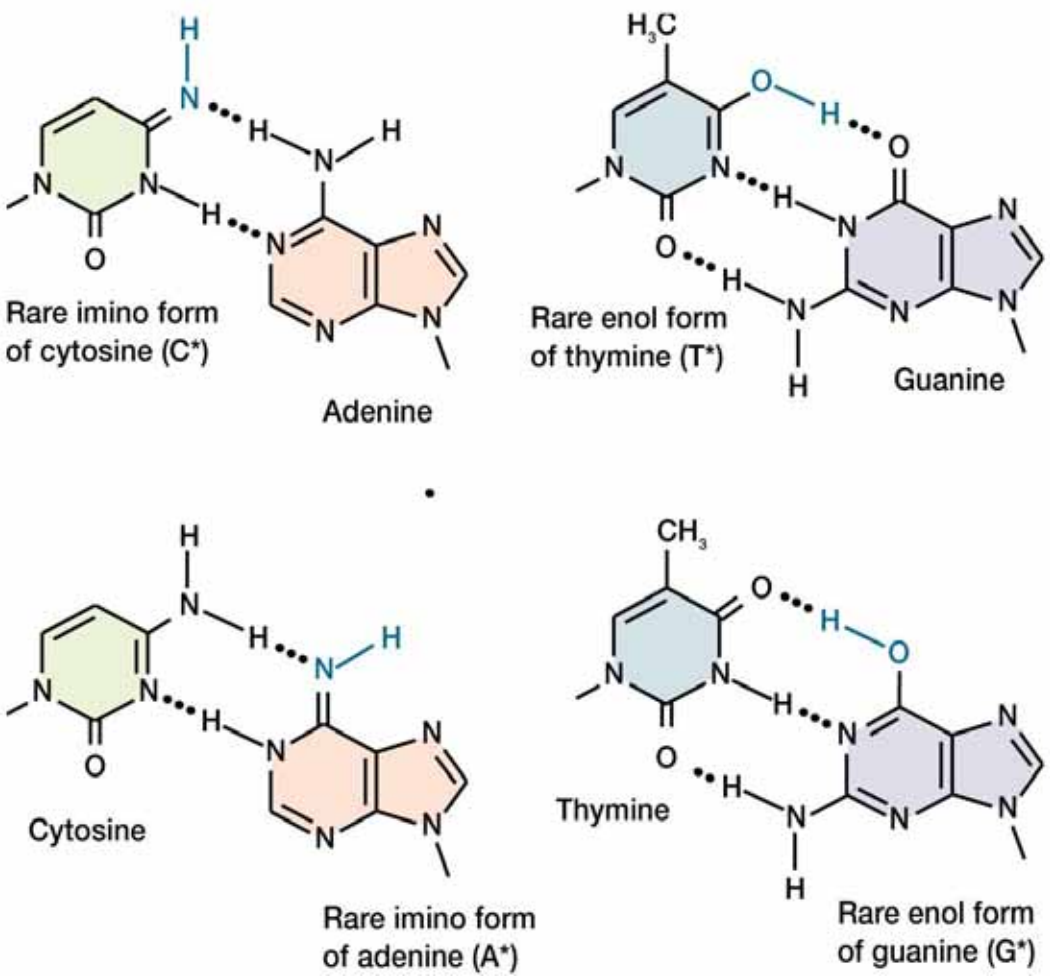


Figure 11.25a

Outcome of tautomeric shift

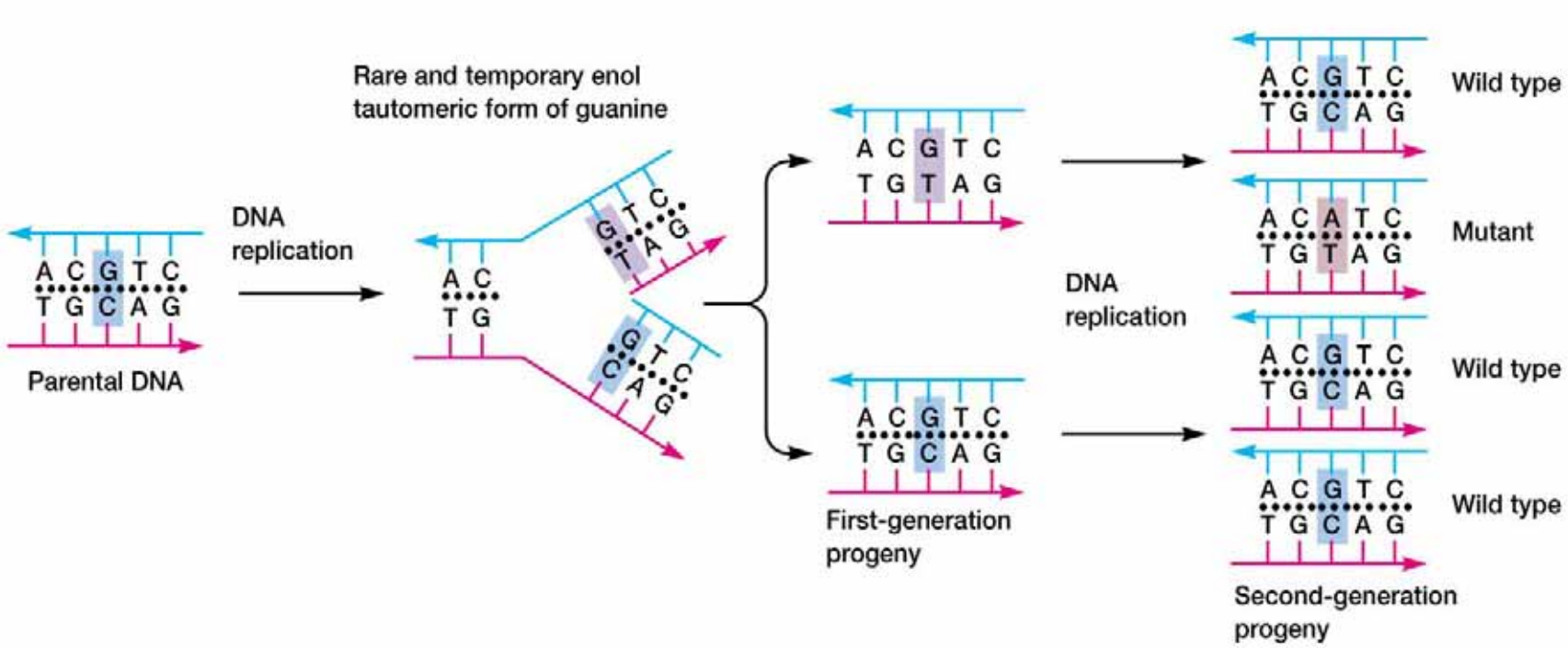


Figure 11.25b

Replication errors - frameshifts

- deletion or addition of base pairs
- alters reading frame

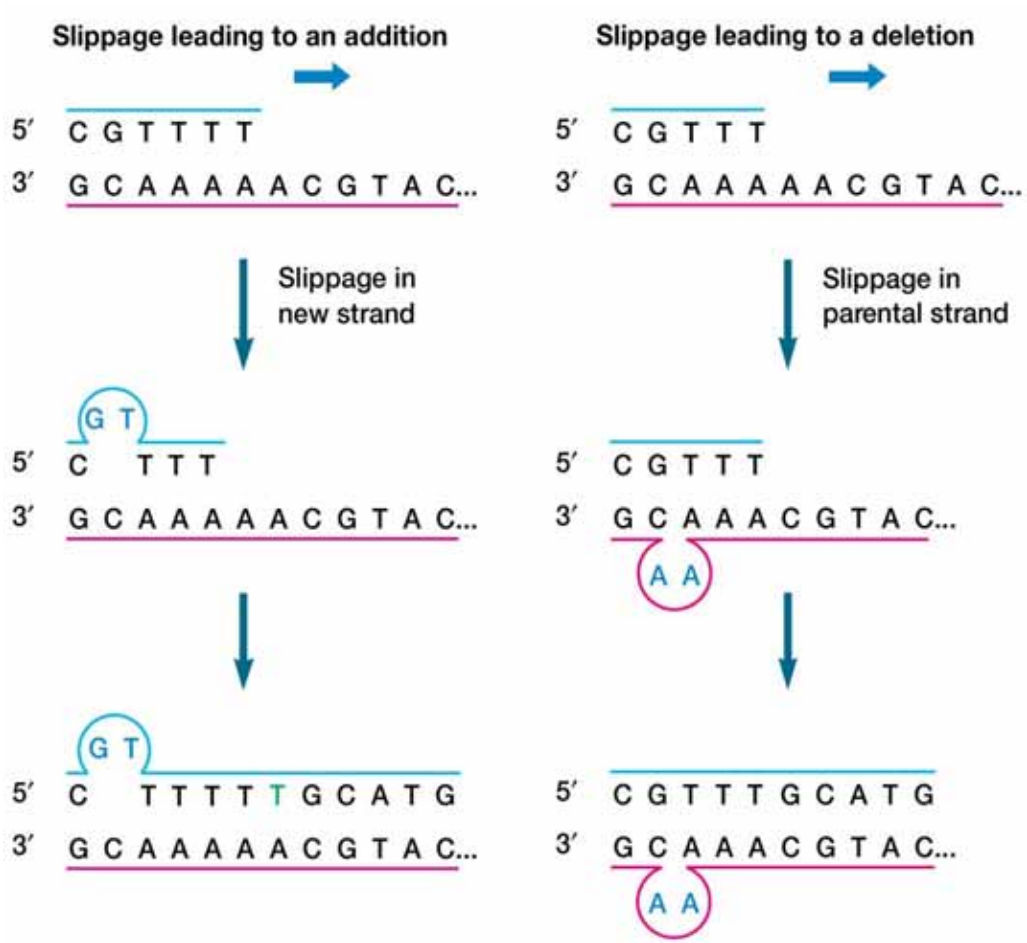


Figure 11.26

Induced Mutations

- caused by chemical or physical agents that damage or alter the chemistry of DNA, or that interfere with DNA repair mechanisms

Base analogs

- similar to nitrogenous bases
- incorporated into DNA during replication
- have different base-pairing characteristics

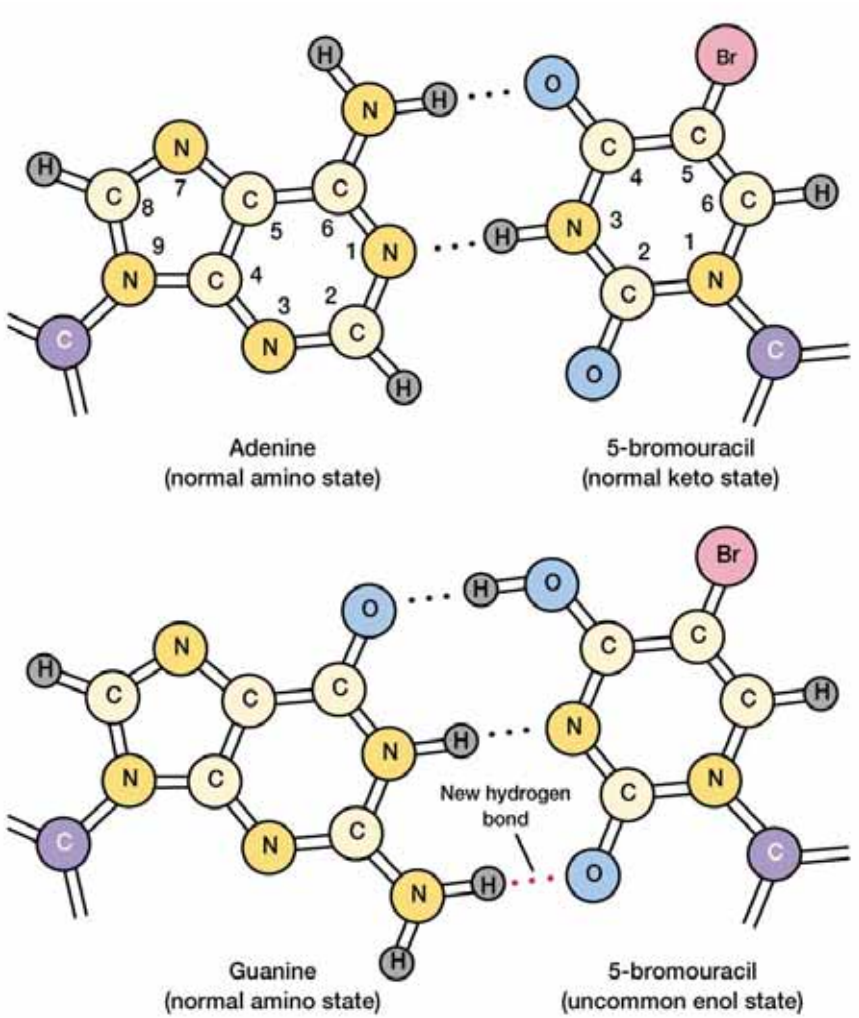


Figure 11.27a

Mutagenesis by the base analog 5-bromouracil (Bu)

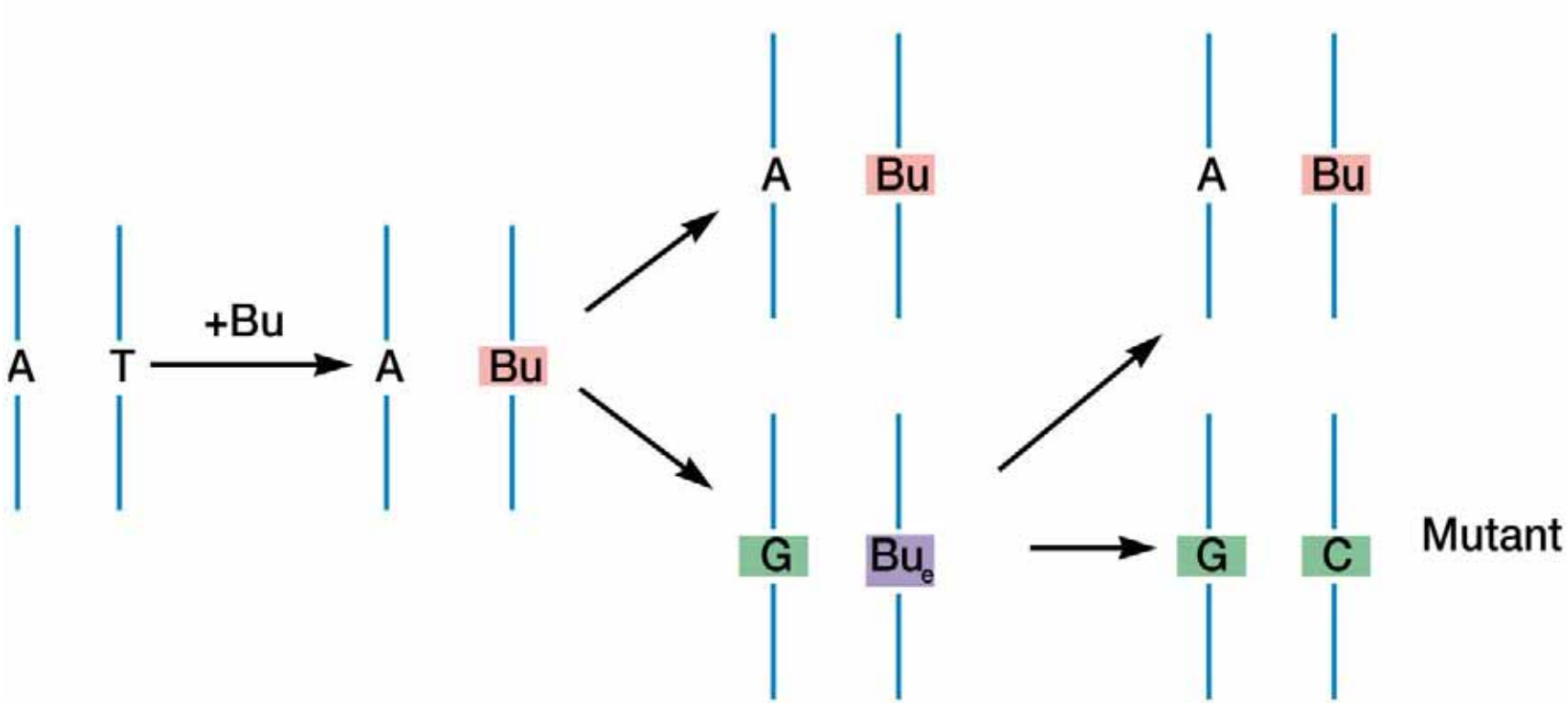


Figure 11.27b

Specific mispairings

- occur when mutagen changes base's structure and pairing characteristics
  - e.g., alkylating agents

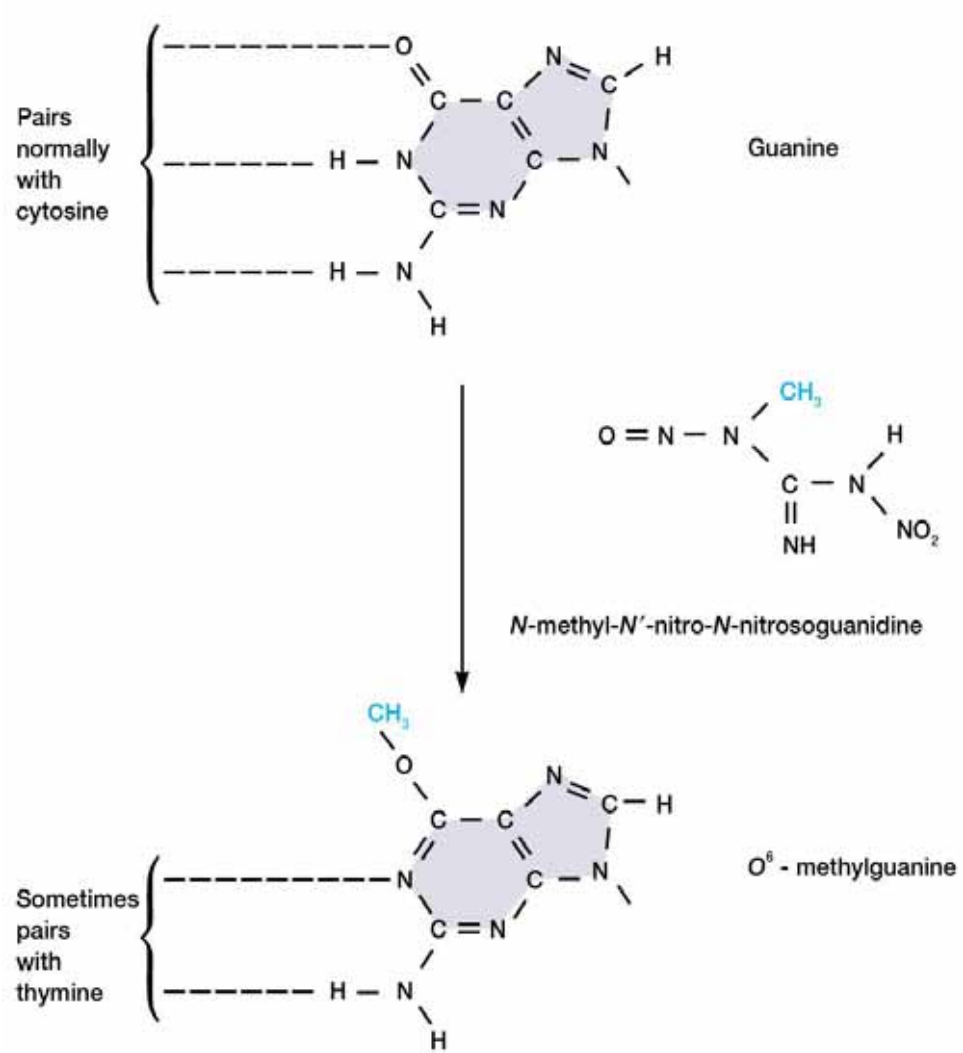


Figure 11.28

Intercalating agents

- planar molecules
- become inserted between stacked bases of helix, distorting DNA
- cause single base pair additions and deletions
- e.g., proflavin and acridine orange

DNA-damaging agents

- severely damage DNA so that it can't serve as template for replication
- repair mechanisms allow survival, but also cause mutations

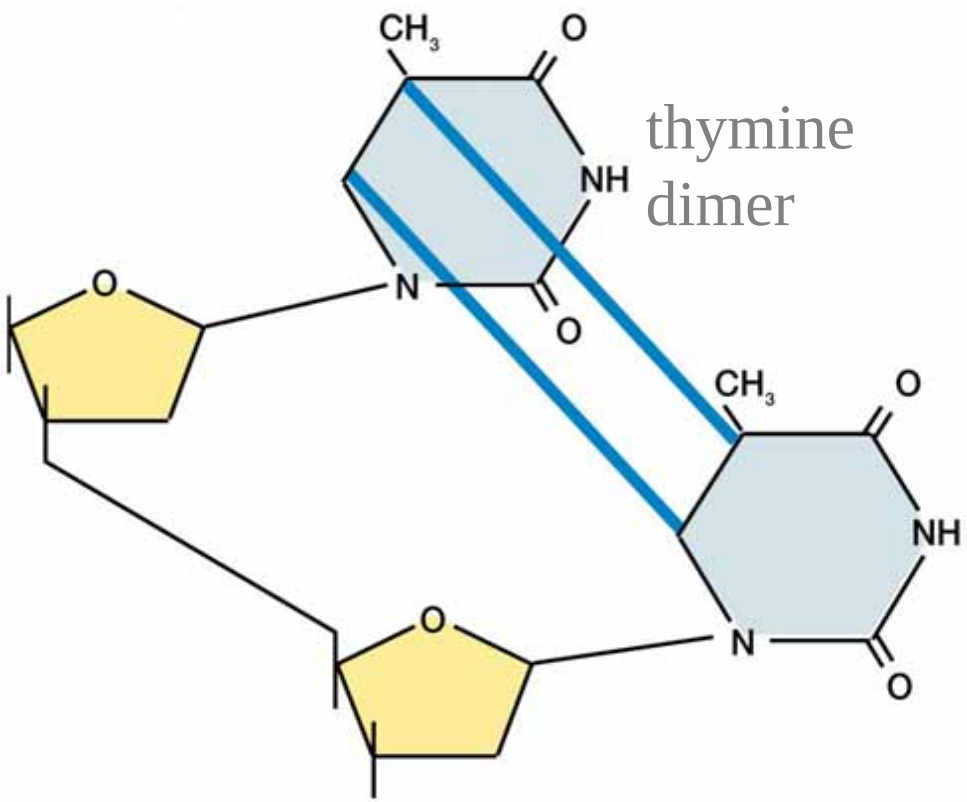


Figure 11.29

The Expression of Mutations

- wild type
  - most prevalent form of gene
- forward mutations
  - wild type $\rightarrow$ mutant form
- reverse mutations
  - mutant phenotype $\rightarrow$ wild type phenotype

Forward mutations

Table 11.2 Summary of Some Molecular Changes from Gene Mutations

Type of Mutation	Result and Example
Forward Mutations	
Single Nucleotide-Pair (Base-Pair) Substitutions	
At DNA Level	
Transition	Purine replaced by a different purine, or pyrimidine replaced by a different pyrimidine (e.g., AT ⟶ GC).
Transversion	Purine replaced by a pyrimidine, or pyrimidine replaced by a purine (e.g., AT ⟶ CG).
At Protein Level	
Silent mutation	Triplet codes for same amino acid: AGG ⟶ **C**GG both code for Arg
Neutral mutation	Triplet codes for different but functionally equivalent amino acid: AAA (Lys) ⟶ A**G**A (Arg)
Missense mutation	Triplet codes for a different amino acid.
Nonsense mutation	Triplet codes for chain termination: CAG (Gln) ⟶ **U**AG (stop)
Single Nucleotide-Pair Addition or Deletion: Frameshift Mutation	Any addition or deletion of base pairs that is not a multiple of three results in a frameshift in reading the DNA segments that code for proteins.
Intragenic Addition or Deletion of Several to Many Nucleotide Pairs	

Reverse mutations

Table 11.2 Summary of Some Molecular Changes from Gene Mutations

Type of Mutation	Result and Example
Reverse Mutations	
True Reversion	AAA (Lys) wild type —forward→ GAA (Glu) mutant —reverse→ AAA (Lys) wild type
Equivalent Reversion	UCC (Ser) wild type —forward→ UGC (Cys) mutant —reverse→ AGC (Ser) wild type
	CGC (Arg, basic) wild type —forward→ CCC (Pro, not basic) mutant —reverse→ CAC (His, basic) pseudo-wild type
Suppressor Mutations	
Intragenic Suppressor Mutations Frameshift of opposite sign at site within gene. Addition of X to the base sequence shifts the reading frame from the CAT codon to XCA followed by TCA codons. The subsequent deletion of a C base shifts the reading frame back to CAT.	CATCATCATCATCATCAT (+) (–) CATXCATATCATCATCAT y x z y y y
Extragenic Suppressor Mutations	
Nonsense suppressors	Gene (e.g., for tyrosine tRNA) undergoes mutational event in its anticodon region that enables it to recognize and align with a mutant nonsense codon (e.g., UAG) to insert an amino acid (tyrosine) and permit completion of the translation.
Physiological suppressors	A defect in one chemical pathway is circumvented by another mutation—for example, one that opens up another chemical pathway to the same product, or one that permits more efficient uptake of a compound produced in small quantities because of the original mutation.

Other mutations

- regulatory mutations
  - changes in regulatory sequences
  - alter control of gene expression
- rRNA and tRNA mutations
  - can disrupt protein synthesis
  - some tRNA mutations are suppressor mutations

Detection and Isolation of Mutants

- mutations are generally rare
  - one per 10^7 to 10^{11} cells
- finding mutants requires sensitive detection methods and/or methods to increase frequency of mutations

Mutant Detection

- observation of changes in phenotype
- replica plating technique
  - used to detect auxotrophic mutants

Replica plating

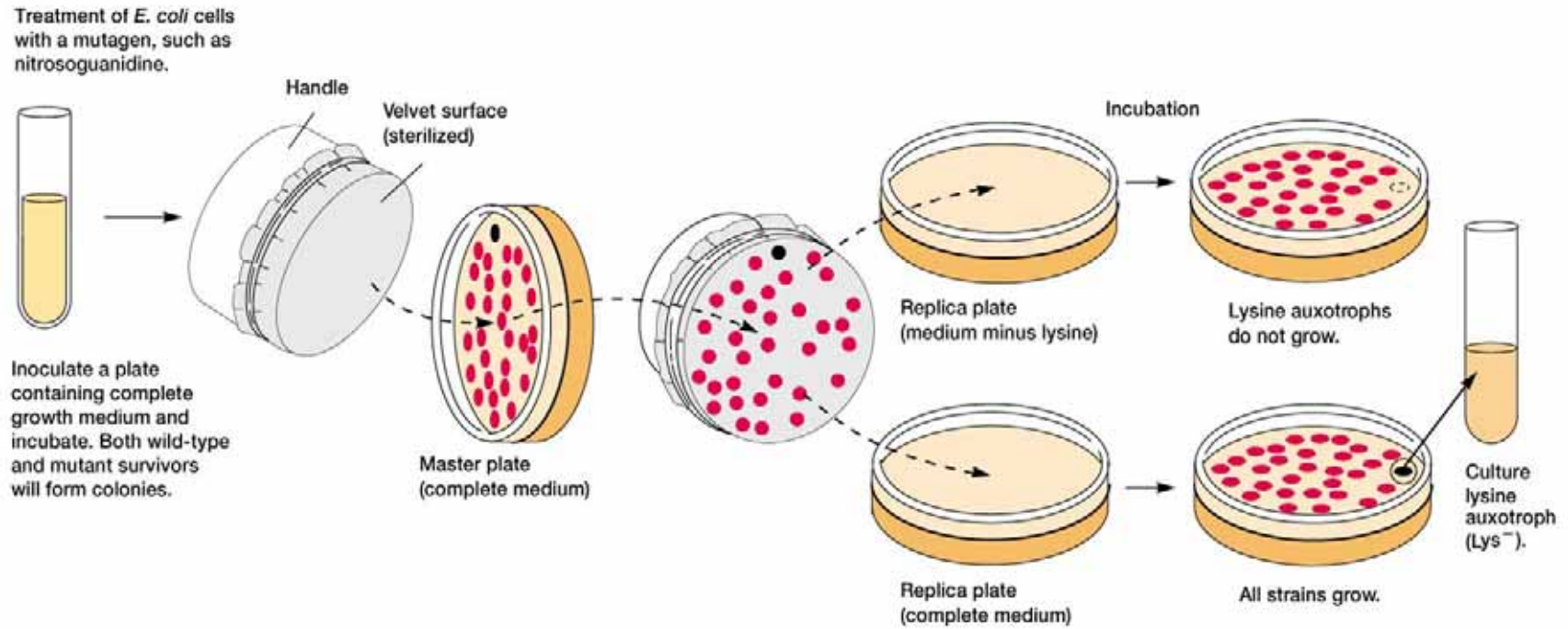


Figure 11.31

Mutant Selection

- use of environmental condition in which only desired mutant will grow
  - e.g., selection for revertants from auxotrophy to prototrophy

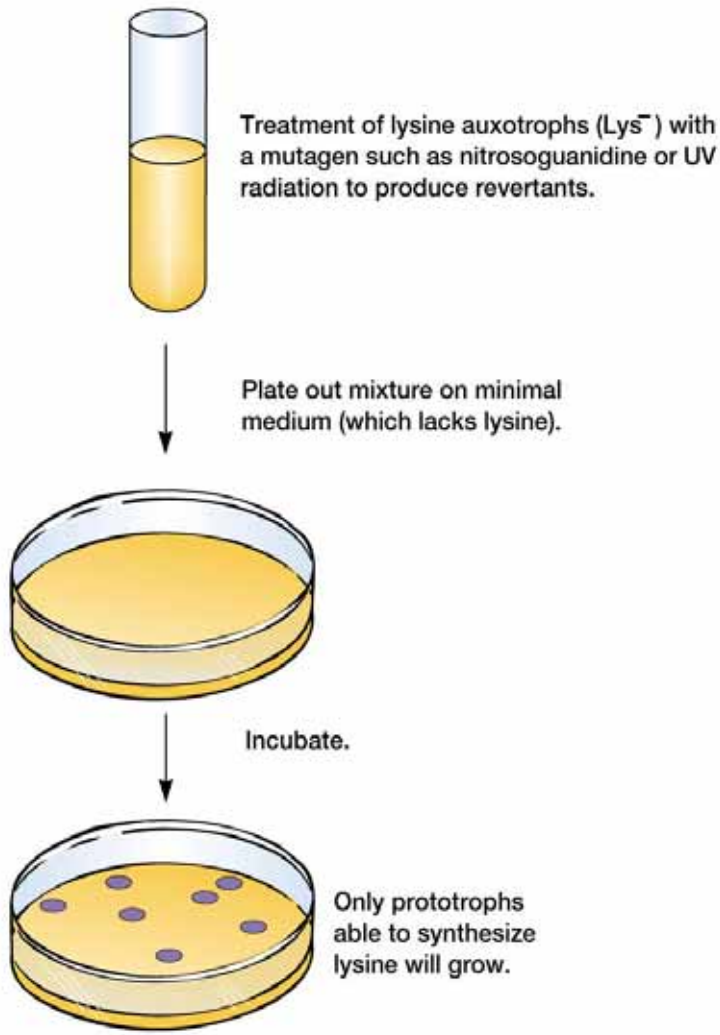


Figure 11.32

Carcinogenicity Testing

- based on observation that most carcinogens are also mutagens
- tests for mutagenicity are used as screen for carcinogenic potential
- e.g., Ames test

Ames Test

- reversion rate in presence of suspected carcinogen > reversion rate in absence of suspected carcinogen

- then, agent is a mutagen, and may be carcinogen

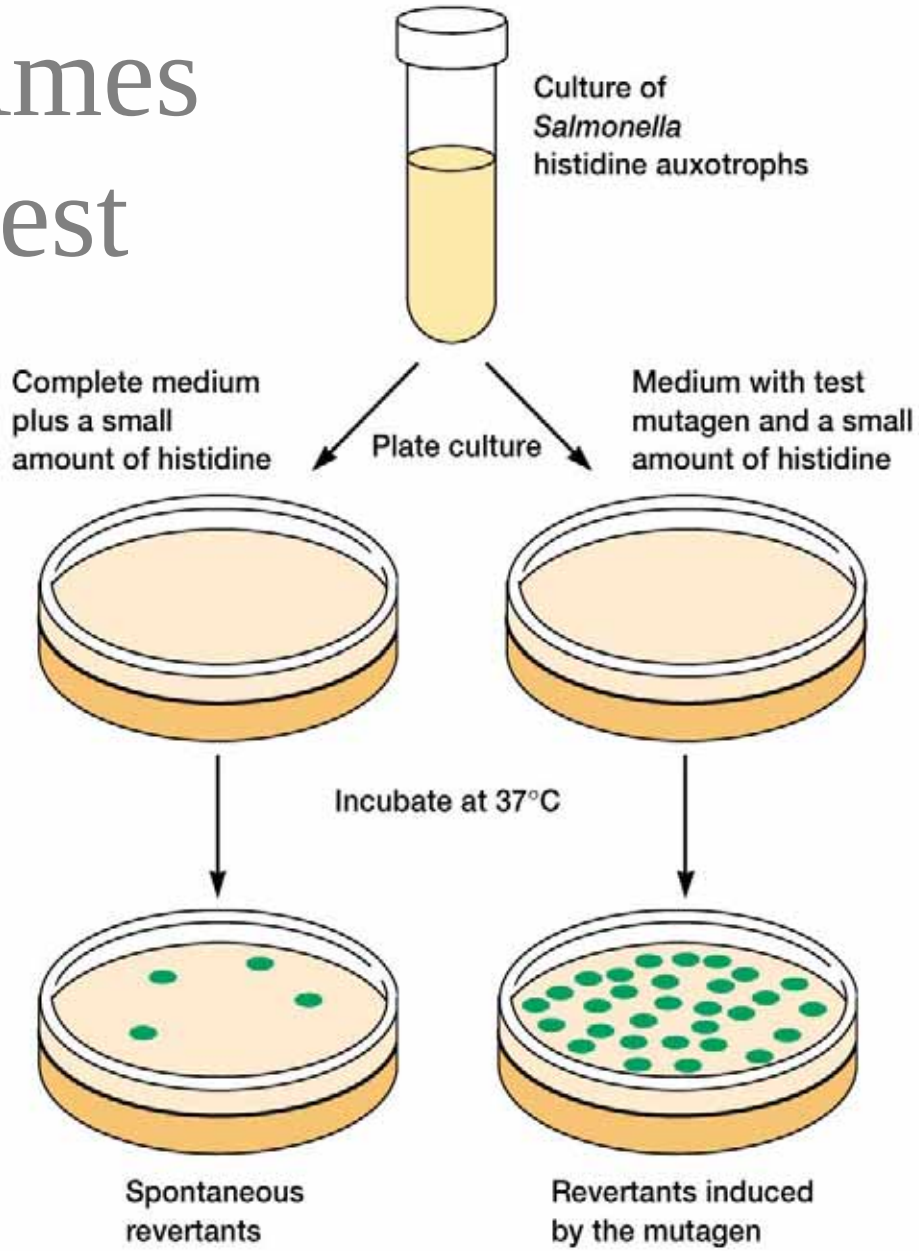


Figure 11.33

DNA Repair

- proofreading
  - correction of errors in base pairing made during replication
  - errors corrected by DNA polymerase
- other repair mechanisms repair incorrect pairings and DNA damage

Excision Repair

- corrects damage that causes distortions in double helix
  - e.g., thymine dimers
  - e.g., apurinic and apyrimidinic sites
  - e.g., damaged bases

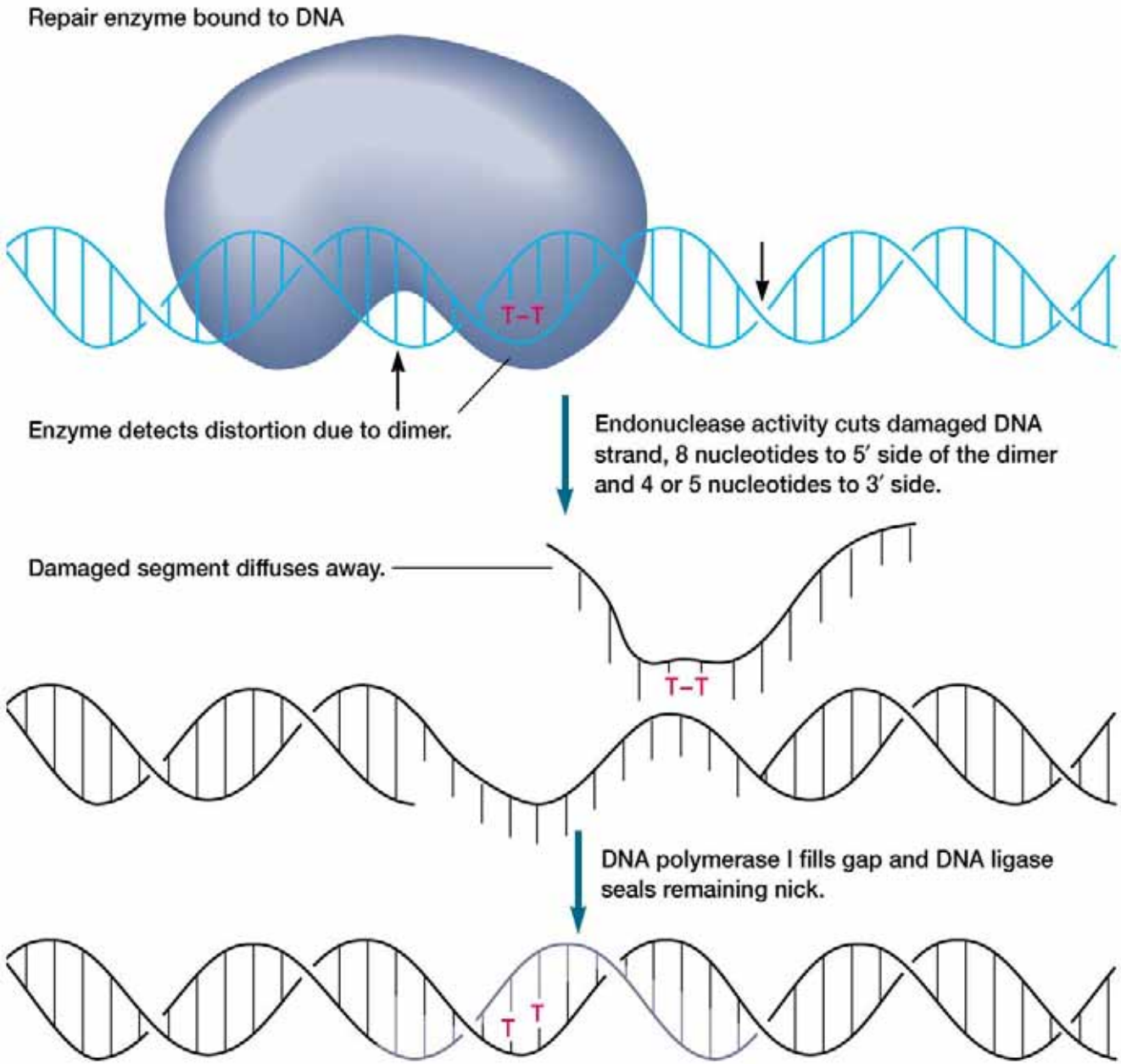

Repair enzyme bound to DNA
T–T
Enzyme detects distortion due to dimer.
Endonuclease activity cuts damaged DNA strand, 8 nucleotides to 5′ side of the dimer and 4 or 5 nucleotides to 3′ side.
Damaged segment diffuses away.
T–T
DNA polymerase I fills gap and DNA ligase seals remaining nick.
T T

Figure 11.34

Removal of Lesions

- photoreactivation
  - used to directly repair thymine dimers
  - thymines separated by photochemical reaction catalyzed by photolyase
- direct repair of alkylated bases
  - catalyzed by alkyltransferase or methylguanine methyltransferase

Postreplication Repair

- type of excision repair
- e.g., mismatch repair system in *E. coli*
  - mismatch correction enzyme scans newly synthesized DNA for mismatched pairs
  - mismatched pairs removed and replaced by DNA polymerase and DNA ligase

DNA methylation

- used by *E. coli* postreplication repair system to distinguish old DNA strands from new DNA strands
  - old DNA methylated; new DNA not methylated
- catalyzed by DNA methyltransferases

Recombination Repair

- repairs DNA with damage in both strands
- involves recombination with an undamaged molecule
  - in rapidly dividing cells, another copy of chromosome is often available
- recA protein catalyzes recombination events

SOS repair

- inducible repair system
- used to repair excessive damage that halts replication, leaving many gaps
  - recA protein initiates recombination repair
  - recA protein also acts as protease, destroying a repressor protein and thereby increasing production of excision repair enzymes

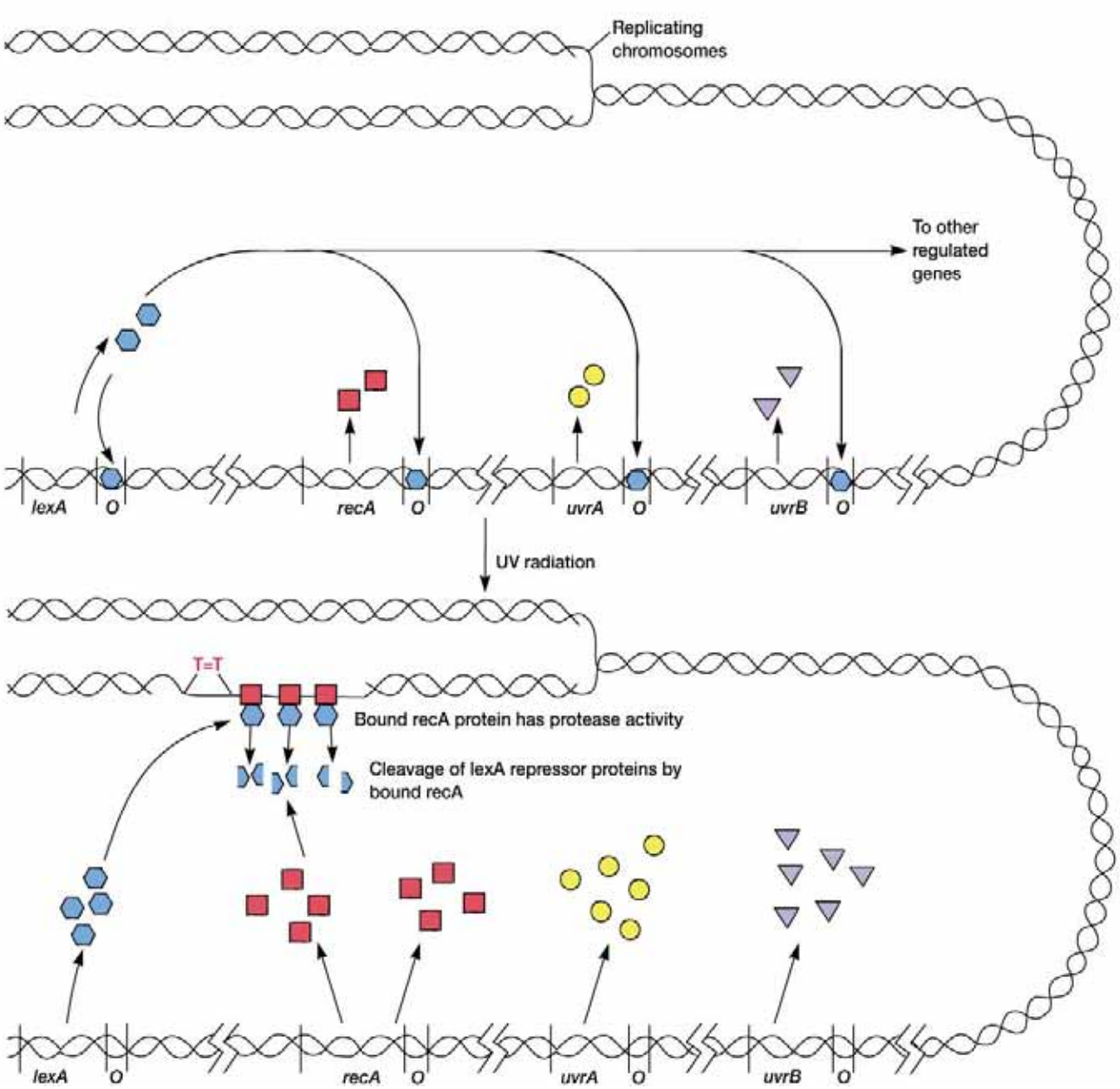

Replicating chromosomes
To other regulated genes
lexA
O
recA
O
uvrA
O
uvrB
O
UV radiation
T=T
Bound recA protein has protease activity
Cleavage of lexA repressor proteins by bound recA
lexA
O
recA
O
uvrA
O
uvrB
O

Figure 11.35