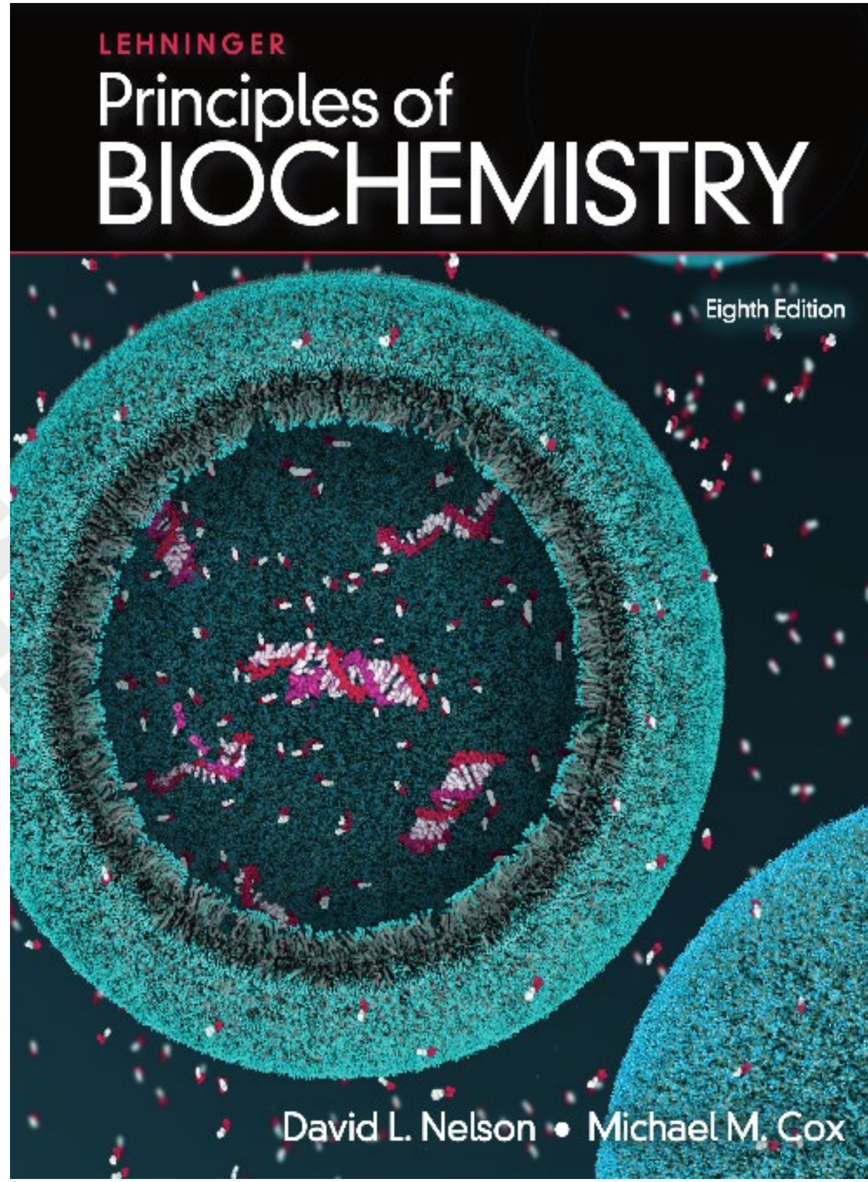
# 8 Nucleotides and Nucleic Acids



# Nucleotides

- have a variety of roles in cellular metabolism:
  - energy currency in metabolic transactions
  - essential chemical links in the response of cells to hormones and other extracellular stimuli
  - structural components of an array of enzyme cofactors and metabolic intermediates
  - constituents of nucleic acids: **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**

# P1 Principle 1 (1 of 6)

**Nucleic acids are both repositories and functional expressions of biological information.** Biological information is one of the required conditions for life, a blueprint for each species transmitted from one generation to the next. RNA can be a functional expression of that information, directing the synthesis of proteins or, in some cases, acting directly as a signal or a reaction catalyst.

# P2 Principle 2 (1 of 4)

**The transmission of biological information relies on molecular complementarity.** Chromosomes are the largest molecules in any cell. They are polymers composed of a small set of common nucleotides, with information embedded in the nucleotide sequence. The common nucleotides in RNA and DNA are organized so that two strands of nucleic acid can maintain a complementary and uniform structure over vast molecular distances. This extended potential for both variable sequence and complementarity, and thus information storage and transmission, is a property shared by no other class of biological molecule.

# P3 Principle 3 (1 of 3)

**Biological information is subject to natural damage and change.** DNA damage is a constant, and it results in occasional mutation — the raw material for evolution.

# P4 Principle 4 (1 of 4)

**Biological information can be accessed, interpreted, and altered in the laboratory.** The information embedded in nucleic acids
is of singular importance to biochemistry and molecular biology. The techniques for sequencing, synthesizing, and altering nucleic acids are continually advancing.

# P5 Principle 5 (1 of 3)

**Nucleoside triphosphates occupy a central role in cellular metabolism, serving as an energy currency and as important regulatory signals.** ATP is the ultimate product of catabolic pathways, providing fuel for anabolic pathways.

# 8.1 Some Basic Definitions and Conventions

# P1 Principle 1 (2 of 6)

**Nucleic acids are both repositories and functional expressions of biological information.** Biological information is one of the required conditions for life, a blueprint for each species transmitted from one generation to the next. RNA can be a functional expression of that information, directing the synthesis of proteins or, in some cases, acting directly as a signal or a reaction catalyst.

# The Functions of DNA

- **gene =** a segment of a DNA molecule that contains the information required for the synthesis of a functional biological product, whether protein or RNA

- only known functions of DNA:
  - storage of biological information
  - transmission of that information to the next generation

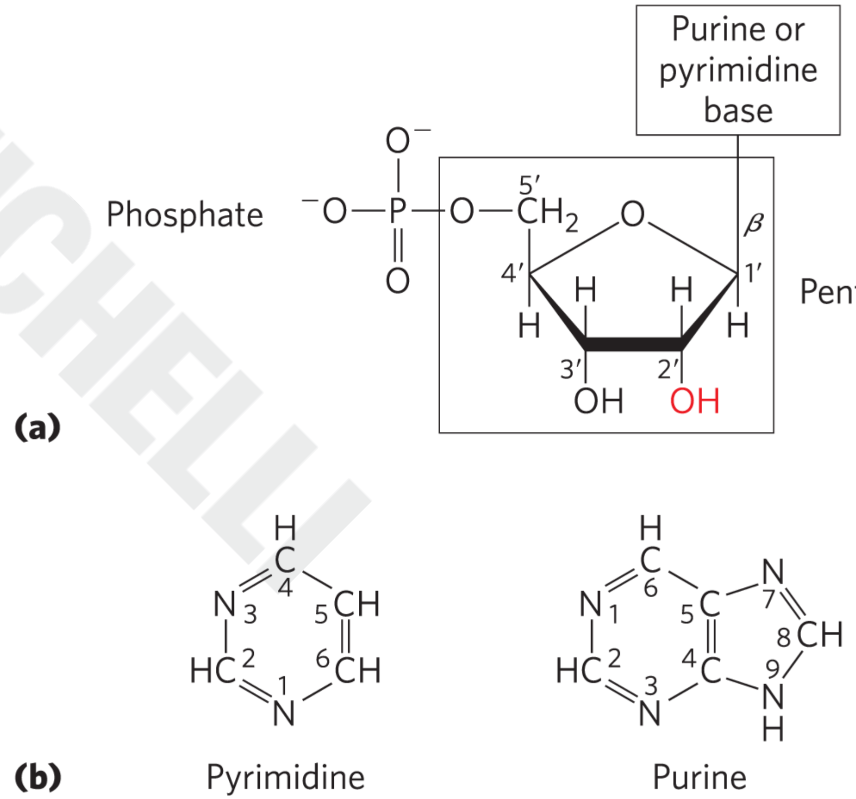
# The Functions of RNA

- several classes of RNA:
  - **ribosomal RNAs (rRNAs)** = components of ribosomes
  - **messenger RNAs (mRNAs)** = intermediates in protein synthesis
  - **transfer RNAs (tRNAs)** = adapter molecules that translate the information in mRNA into a specific amino acid sequence
  - **noncoding RNAs (ncRNAs)** = wide variety of functions

# Nucleotides and Nucleic Acids Have Characteristic Bases and Pentoses

- **nucleotides** have three components:
  - a nitrogenous base (**pyrimidine** or **purine**)
  - a pentose
  - 1+ phosphates

- **nucleoside** = the molecule without a phosphate group



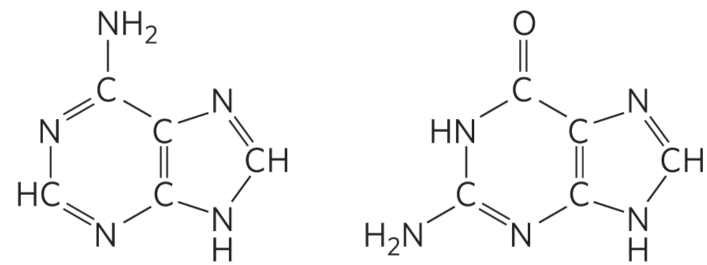
# Nucleotide Bonds

- *N*-β-glycosyl bond = covalently joins the 1′ carbon of the pentose to the base (at N-1 of pyrimidines and N-9 of purines)
  - formed by removal of the elements of water

- the phosphate is esterified to the 5′ carbon

# Nucleotide Nitrogenous Bases

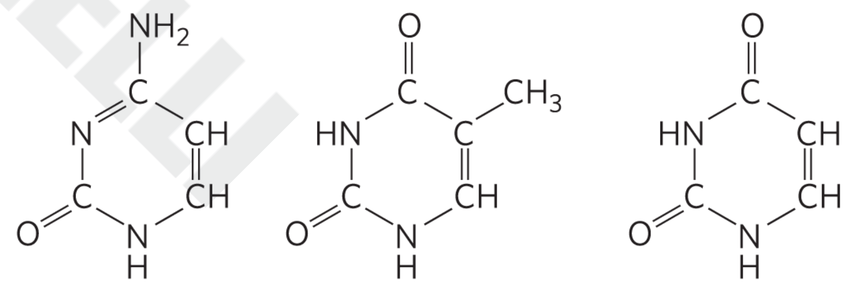
- major purine bases:
  - **adenine (A)** = in DNA and RNA
  - **guanine (G)** = in DNA and RNA



Adenine Guanine

**Purines**

- major pyrimidine bases:
  - **cytosine (C)** = in DNA and RNA
  - **thymine (T)** = only in DNA
  - **uracil (U)** = only in RNA



Cytosine Thymine (DNA) Uracil (RNA)

**Pyrimidines**

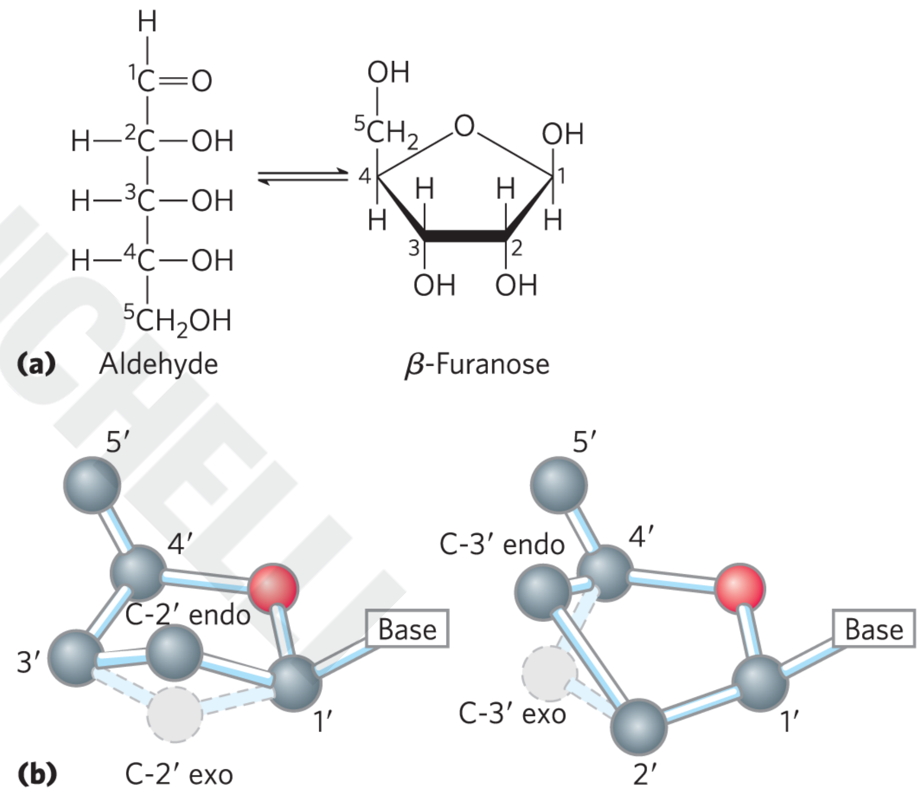
# Nucleotide and Nucleic Acid Nomenclature

| Table 8-1 Nucleotide and Nucleic Acid Nomenclature | | | |
|---|---|---|---|
| **Base** | **Nucleoside** | **Nucleotide** | **Nucleic Acid** |
| Purines: Adenine | Adenosine<br>Deoxyadenosine | Adenylate<br>Deoxyadenylate | RNA<br>DNA |
| Purines: Guanine | Guanosine<br>Deoxyguanosine | Guanylate<br>Deoxyguanylate | RNA<br>DNA |
| Pyrimidines: Cytosine | Cytidine<br>Deoxycytidine | Cytidylate<br>Deoxycytidylate | RNA<br>DNA |
| Pyrimidines: Thymine | Thymidine or deoxythymidine | Thymidylate or deoxythymidylate | DNA |
| Pyrimidines: Uracil | Uridine | Uridylate | RNA |

# Nucleotide Pentoses

- two kinds of pentoses:
  - 2′-deoxy-d-ribose = in DNA
  - D-ribose = in RNA

- both are in their β-furanose (closed five-membered ring) form

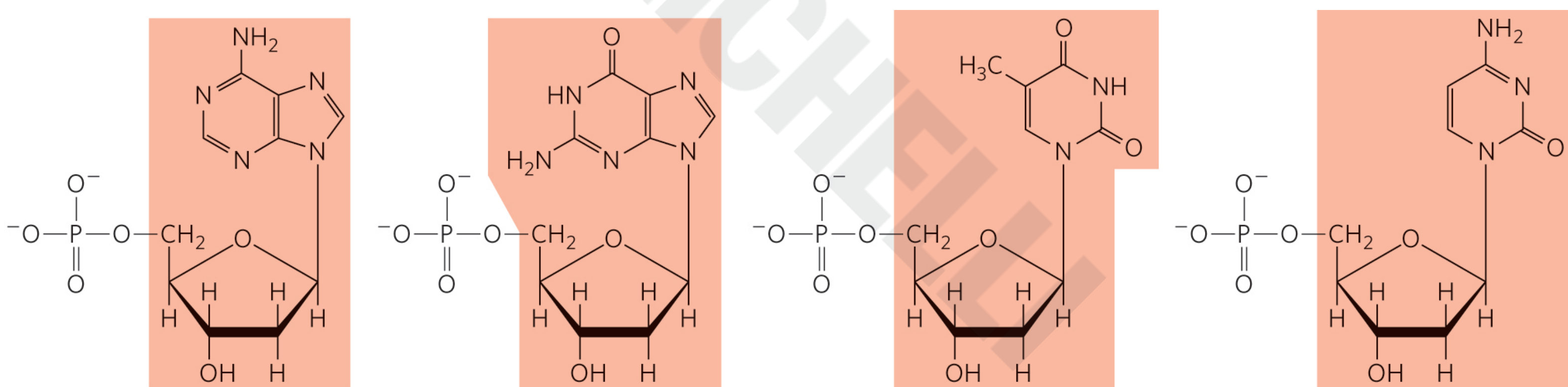



Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

# Structure and Names of The Four Major Deoxyribonucleotides

- **deoxyribonucleotides** = structural units of DNA
  - also called deoxyribonucleoside 5′-monophosphates, deoxynucleotides, and deoxynucleoside triphosphates





| | | | | |
|---|---|---|---|---|
| **Nucleotide:** | Deoxyadenylate (deoxyadenosine 5′-monophosphate) | Deoxyguanylate (deoxyguanosine 5′-monophosphate) | Deoxythymidylate (deoxythymidine 5′-monophosphate) | Deoxycytidylate (deoxycytidine 5′-monophosphate) |
| **Symbols:** | A, dA, dAMP | G, dG, dGMP | T, dT, dTMP | C, dC, dCMP |
| **Nucleoside:** | Deoxyadenosine | Deoxyguanosine | Deoxythymidine | Deoxycytidine |

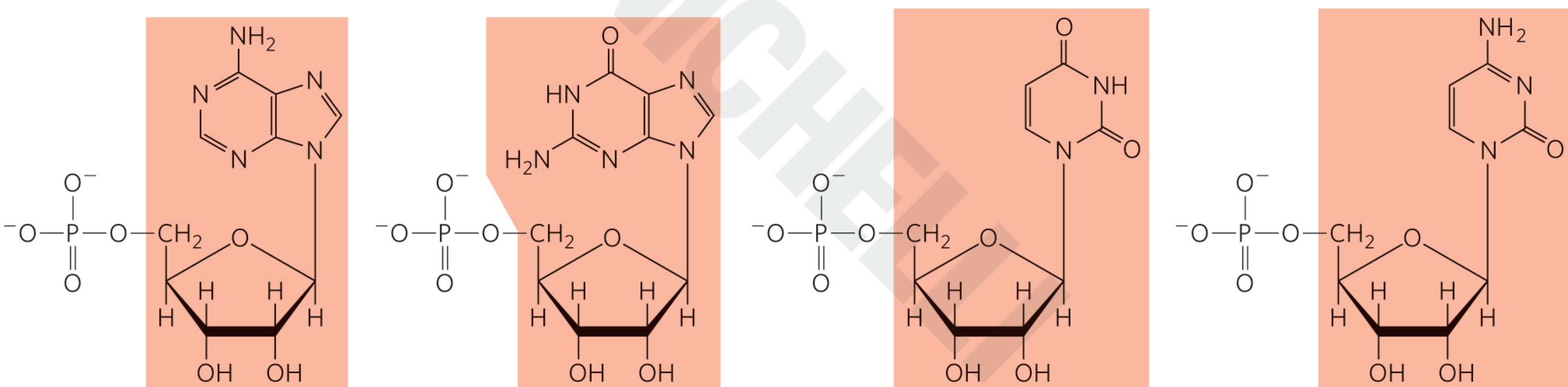
**(a) Deoxyribonucleotides**

Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

# Structure and Names of The Four Major Ribonucleotides

- **ribonucleotides** = structural units of RNA
  - also called ribonucleoside 5′-monophosphates





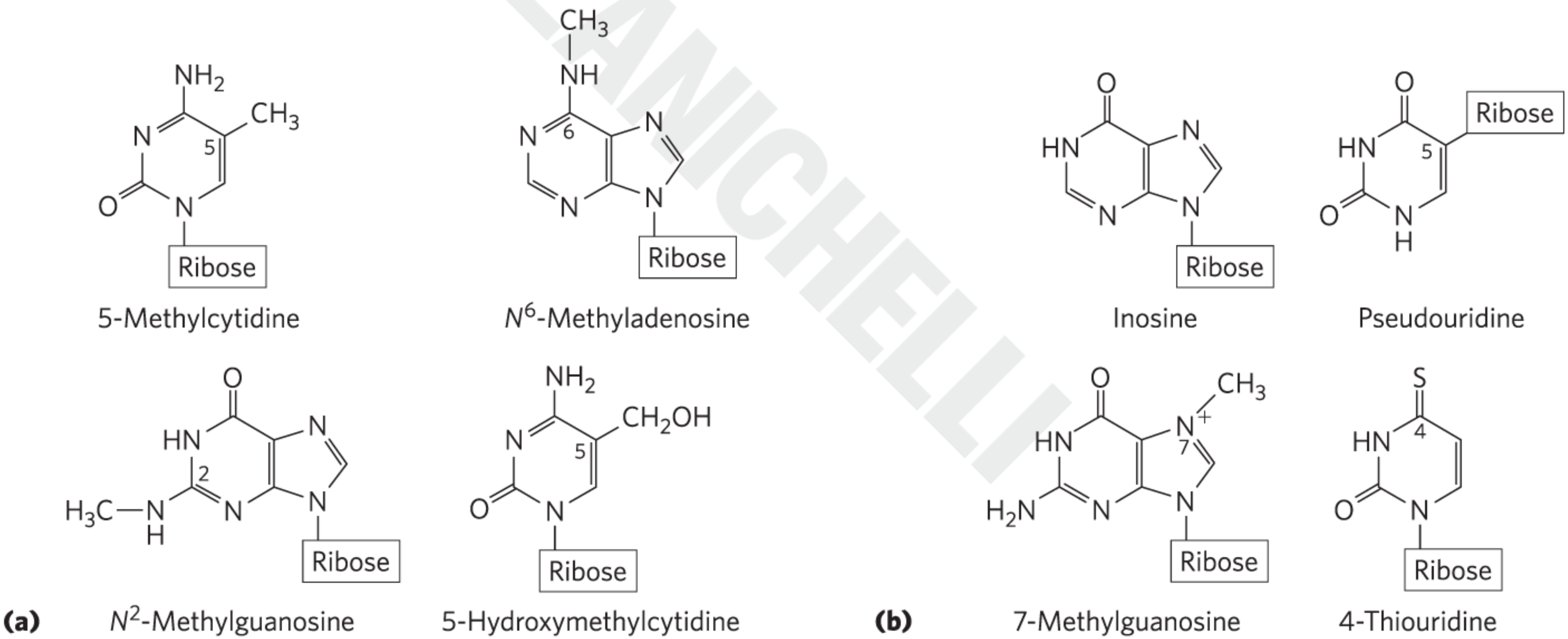
| | | | | |
|---|---|---|---|---|
| **Nucleotide:** | Adenylate (adenosine 5′-monophosphate) | Guanylate (guanosine 5′-monophosphate) | Uridylate (uridine 5′-monophosphate) | Cytidylate (cytidine 5′-monophosphate) |
| **Symbols:** | A, AMP | G, GMP | U, UMP | C, CMP |
| **Nucleoside:** | Adenosine | Guanosine | Uridine | Cytidine |

**(b) Ribonucleotides**

Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

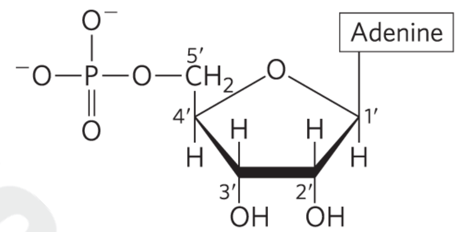
# Minor Purine and Pyrimidine Bases

- methylated forms – most common in DNA

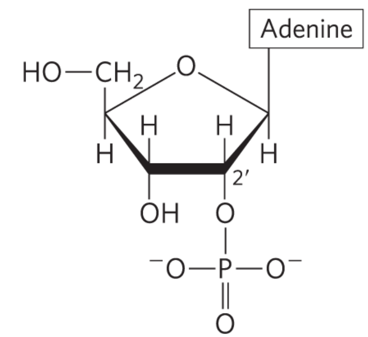


# Nucleotides with Phosphate Groups in Different Positions

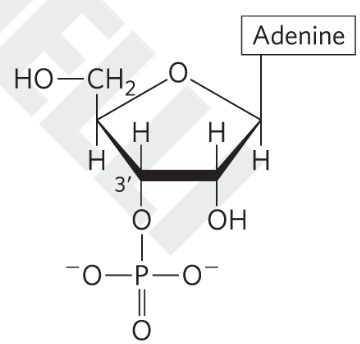
- **ribonucleoside 2′,3′-cyclic monophosphates** = isolatable intermediates
- **ribonucleoside 3′-monophosphates** = end products of RNA hydrolysis
- adenosine 3′,5′-cyclic monophosphate (cAMP)
- guanosine 3′,5′-cyclic monophosphate (cGMP)



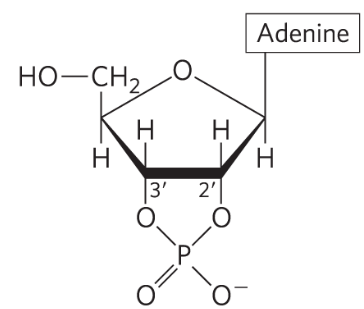
Adenosine 5′-monophosphate



Adenosine 2′-monophosphate



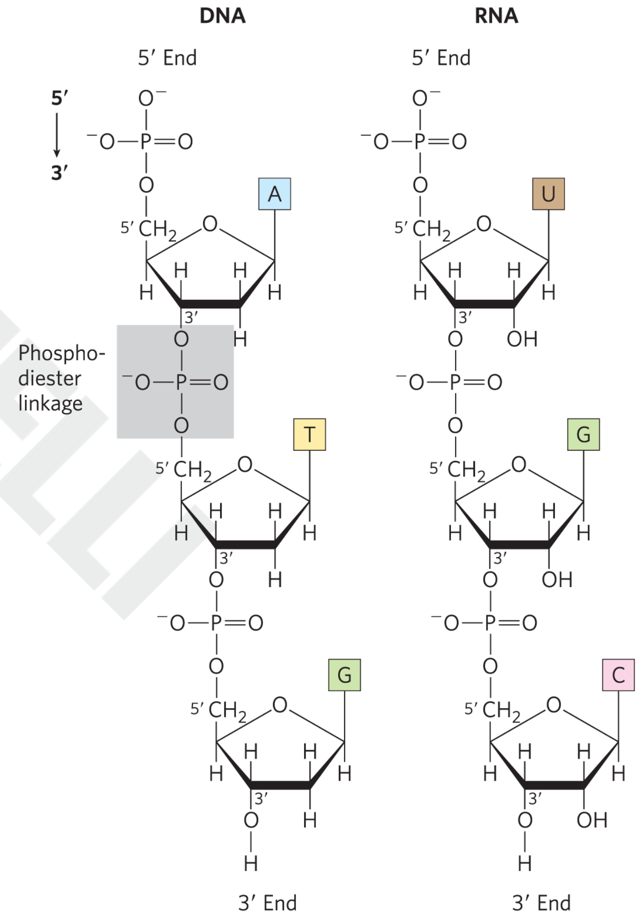
Adenosine 3′-monophosphate



Adenosine 2′,3′-cyclic monophosphate

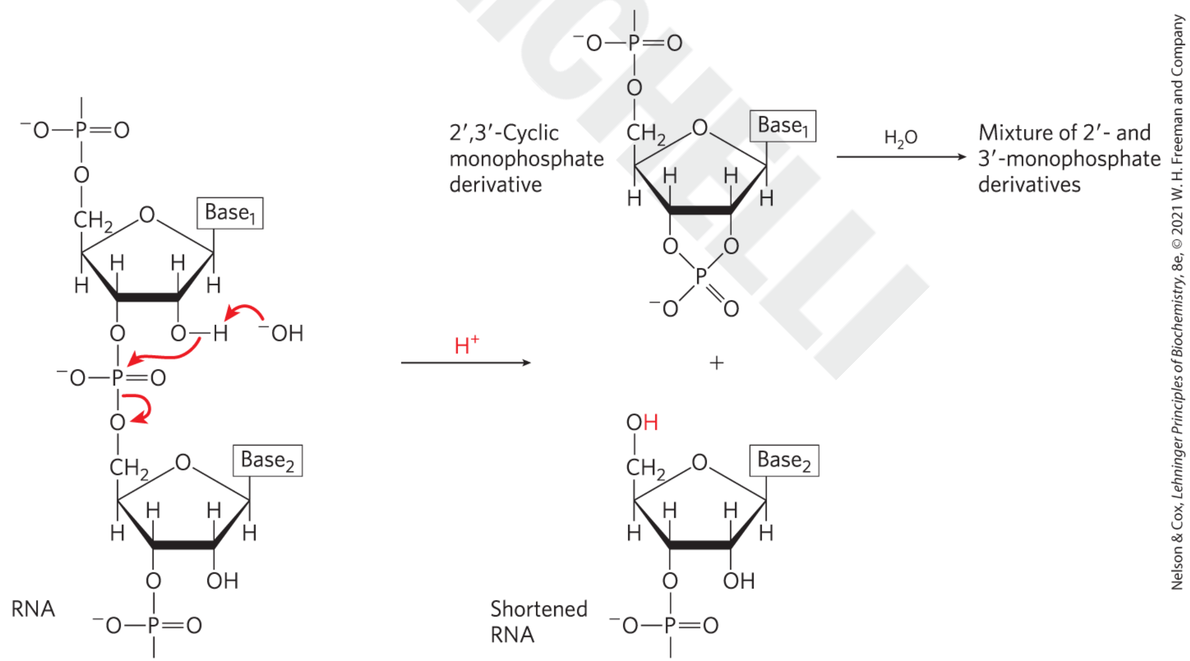
# Phosphodiester Bonds Link Successive Nucleotides in Nucleic Acids

- **phosphodiester linkage** = covalent bond that joins successive nucleotides of both DNA and RNA
  - between the 5′-phosphate group of one nucleotide unit and the 3′-hydroxyl group of the next nucleotide



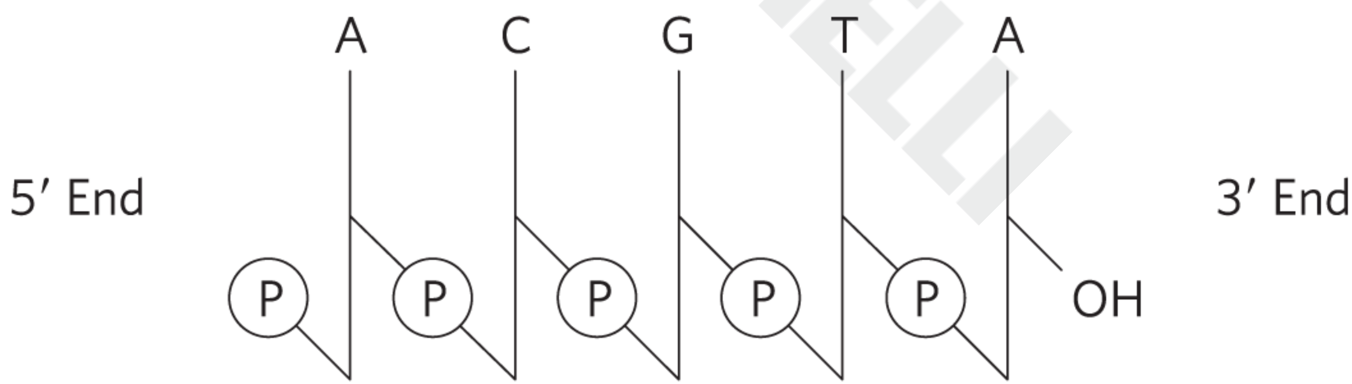
# Hydrolysis of DNA and RNA

- under alkaline conditions:
  - RNA is rapidly hydrolyzed due to the presence of 2′-hydroxyl groups
  - DNA is not rapidly hydrolyzed



# Schematic Representation of Nucleotide Sequences

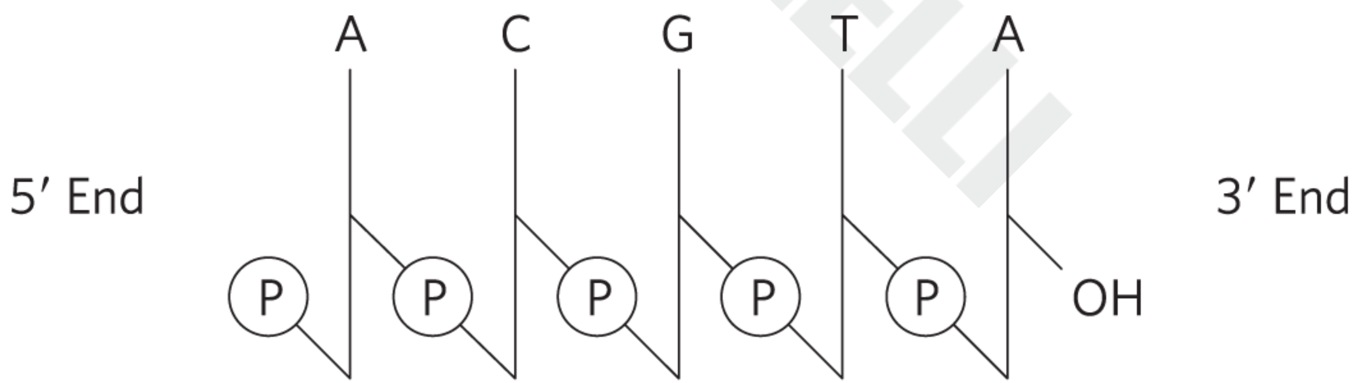
- phosphate groups symbolized by Ⓟ
- deoxyribose symbolized by a vertical line, from C-1′ at the top to C-5′ at the bottom
- lines connecting nucleotides drawn diagonally from C-3′ to C-5′




Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

# Simpler Representations of Nucleotide Sequences

- can also be written as:
  - pA-C-G-T-$A_{OH}$
  - pApCpGpTpA
  - pACGTA




Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

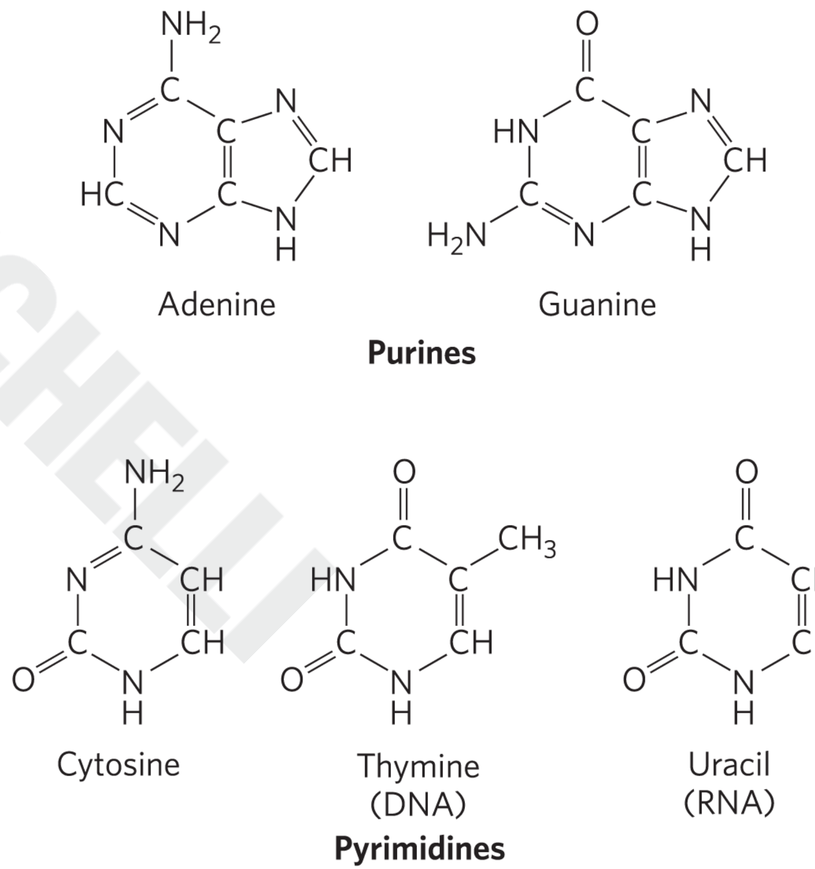
# Oligonucleotides and Polynucleotides

- **oligonucleotide** = short (typically < 50 nucleotides) nucleic acid

- **polynucleotide** = longer nucleic acid

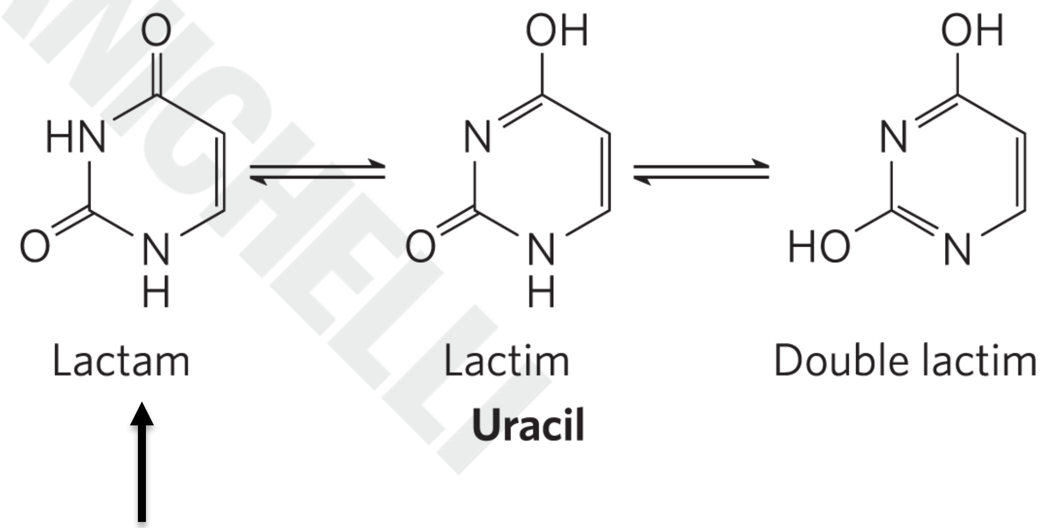
# The Properties of Nucleotide Bases Affect the Three-Dimensional Structure of Nucleic Acids

- weakly basic compounds
- aromatic molecules
- because most bonds in the ring have partial double-bond character:
  - pyrimidines are planar
  - purines have a slight pucker



# Free Pyrimidine and Purine Bases May Exist as Tautomers

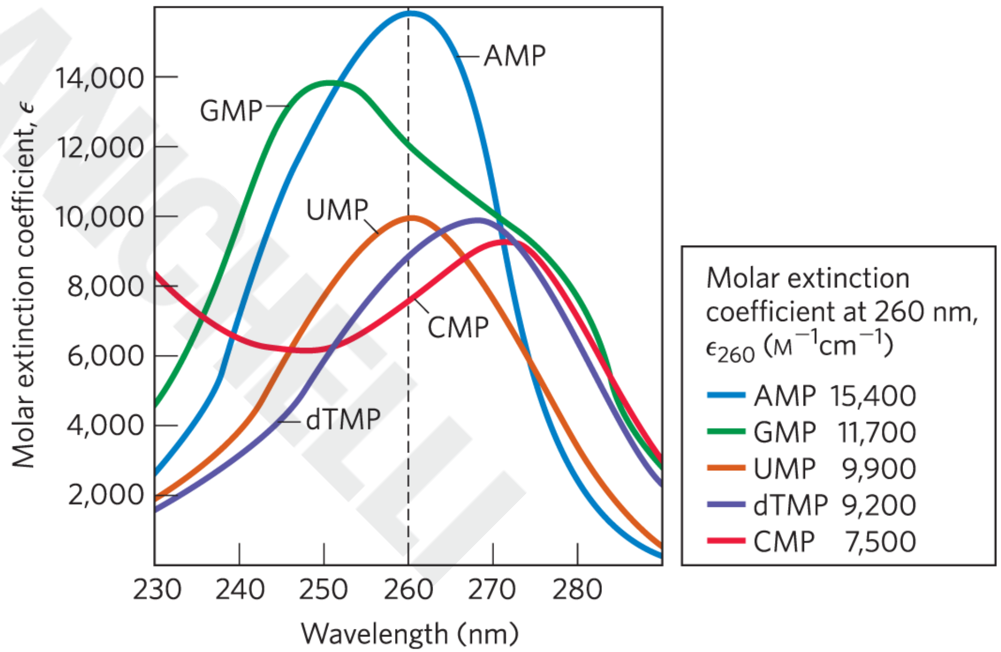
- may exist in readily interconverted forms called **tautomers**
  - lactam
  - lactim
  - double lactim




predominates at pH 7.0

# Absorption Spectra of Common Nucleotides

- all nucleotide bases absorb UV light
- strong absorption near 260 nm



# Solubility of Nucleotides

- hydrophobic and relatively insoluble in pH 7.0 water
  - leads to stacking interactions (van der Waals and dipole-dipole)

- charged and more soluble at acidic or alkaline pH values

# P1 Principle 1 (3 of 6)

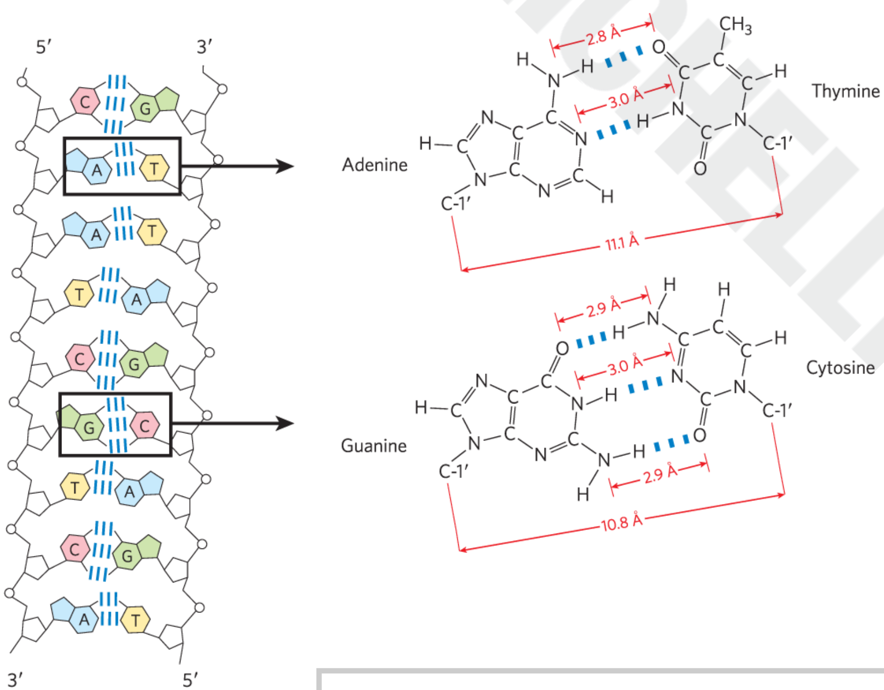
**Nucleic acids are both repositories and functional expressions of biological information.** Biological information is one of the required conditions for life, a blueprint for each species transmitted from one generation to the next. RNA can be a functional expression of that information, directing the synthesis of proteins or, in some cases, acting directly as a signal or a reaction catalyst.

# Principle 2 (2 of 4)

**The transmission of biological information relies on molecular complementarity.** Chromosomes are the largest molecules in any cell. They are polymers composed of a small set of common nucleotides, with information embedded in the nucleotide sequence. The common nucleotides in RNA and DNA are organized so that two strands of nucleic acid can maintain a complementary and uniform structure over vast molecular distances. This extended potential for both variable sequence and complementarity, and thus information storage and transmission, is a property shared by no other class of biological molecule.

# Base Pairing Permits the Duplication of Genetic Information

- **base pairs** = hydrogen-bonding patterns between complementary strands of nucleic acids
  - A bonds specifically to T (or U)
  - G bonds specifically to C




James D. Watson
[Bettmann/Getty Images.]

Francis Crick, 1916–2004
[Bettmann/Getty Images.]

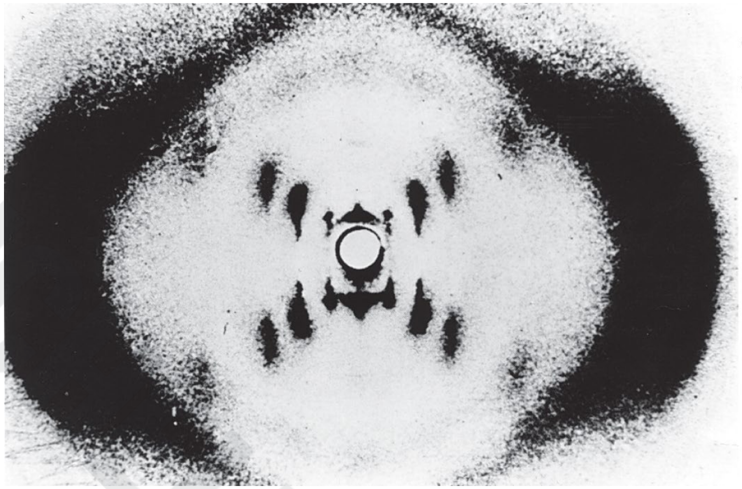
# 8.2 Nucleic Acid Structure

# Hierarchical Levels of Nucleic Acid Structure

- primary structure = covalent structure and nucleotide sequence
- secondary structure = regular, stable structure taken up by some or all the nucleotides
- tertiary structure = complex folding of large chromosomes or the elaborate folding of tRNA or rRNA structures

# DNA Is a Double Helix that Stores Genetic Information

- x-ray diffraction pattern revealed DNA molecules are helical



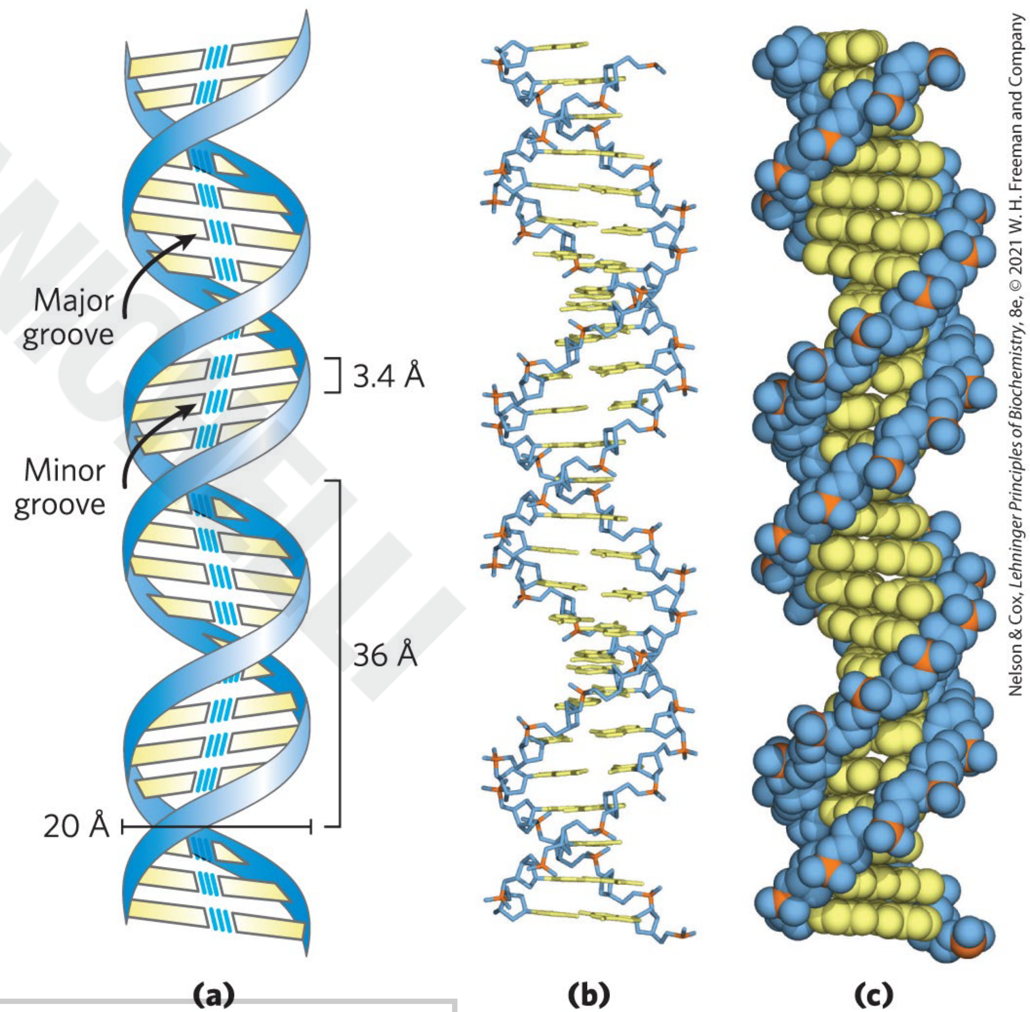
Science Source.

National Library of Medicine/Science Source.

Bettmann/Getty Images.

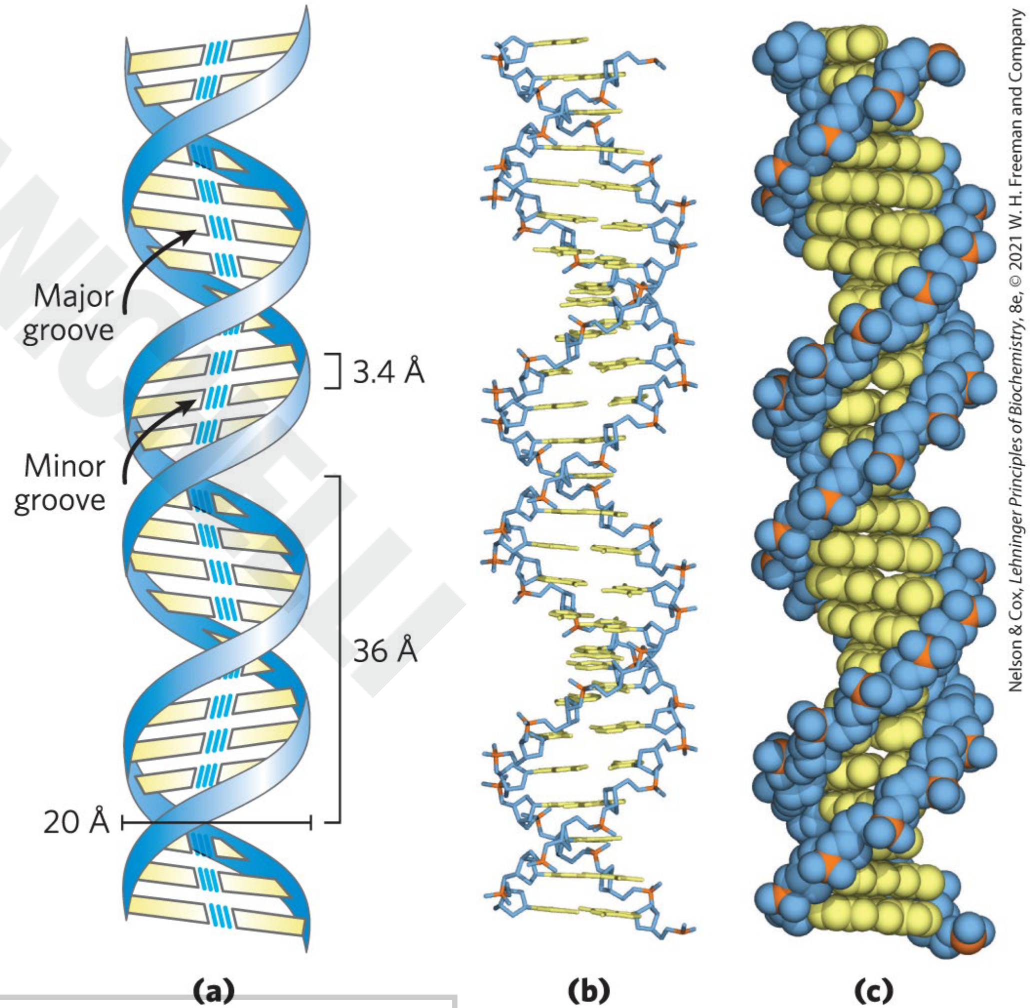
# Watson-Crick Model for the Structure of DNA

- offset pairing of the two strands creates a **major groove** and a **minor groove**
- 3 hydrogen bonds form between G and C
- 2 hydrogen bonds form between A and T



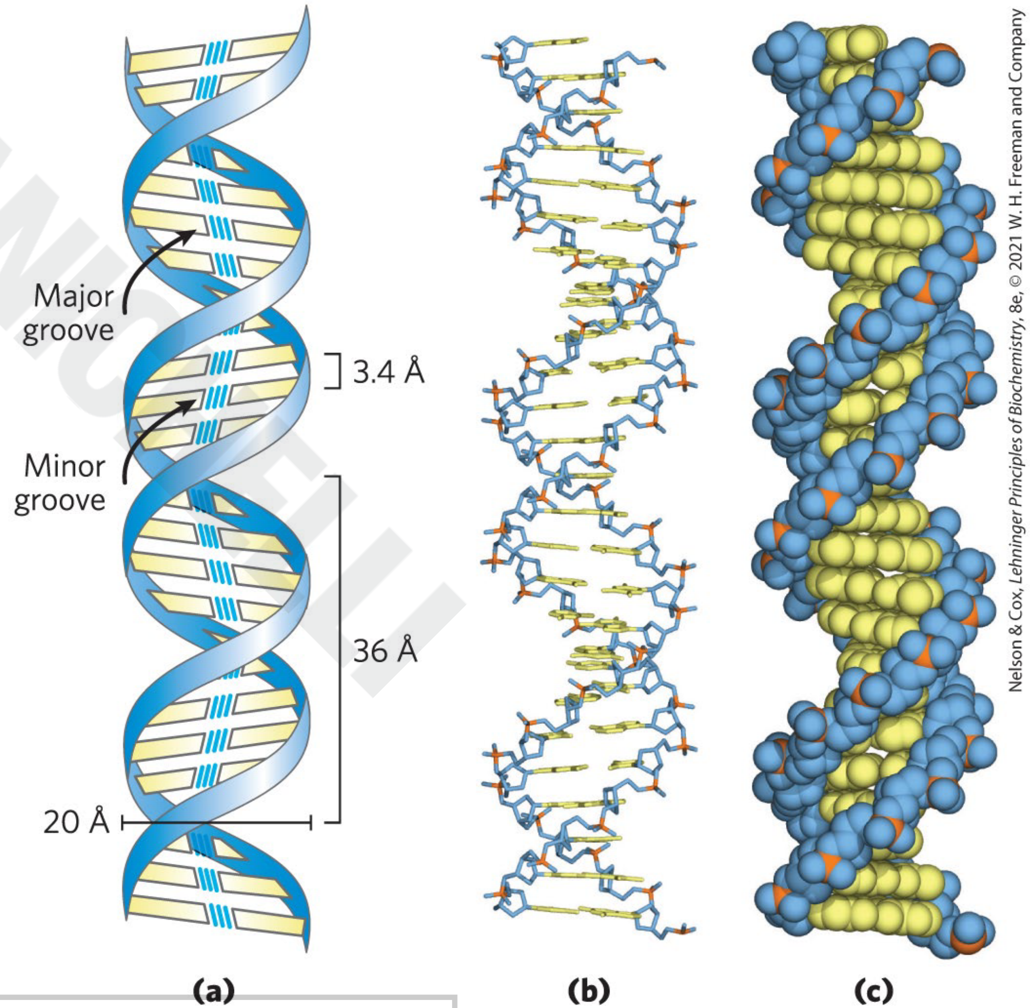
# DNA Strands Are Antiparallel

- **parallel** = 3′ ,5′ - phosphodiester bonds run in the same direction
- **antiparallel** = 3′ ,5′ - phosphodiester bonds run in opposite directions
  – ultimately confirmed by x-ray analysis



# The Double Helix Has 10.5 Base Pairs Per Helical Turn

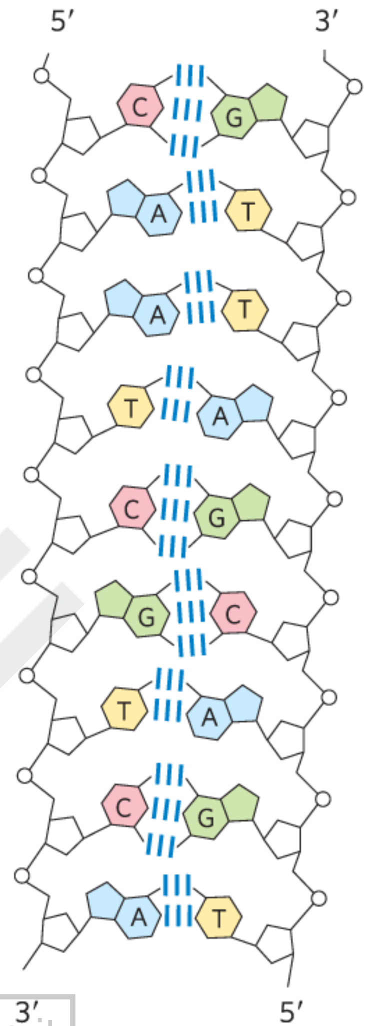
- per helical turn:
  - 10.5 bp
  - 36 Å (3.6 nm)



# Antiparallel Polynucleotides Chains Are Complementary

- double-helical DNA strands are **complementary:**
  - when A occurs in one chain, T is found in the other
  - when G occurs in one chain, C is found in the other

- hydrogen bonding does not contribute significantly to stability of the structure



# Stabilization of the DNA Double Helix

- the double helix is stabilized by:
  - metal cations that shield the negative charges of backbone phosphates
  - base stacking interactions between successive base pairs
    - successive G≡C or C≡G are stronger than successive A=T or T=A
    - duplexes with higher G≡C context are more stable

# P1 Principle 1 (4 of 6)

**Nucleic acids are both repositories and functional expressions of biological information.** Biological information is one of the required conditions for life, a blueprint for each species transmitted from one generation to the next. RNA can be a functional expression of that information, directing the synthesis of proteins or, in some cases, acting directly as a signal or a reaction catalyst.

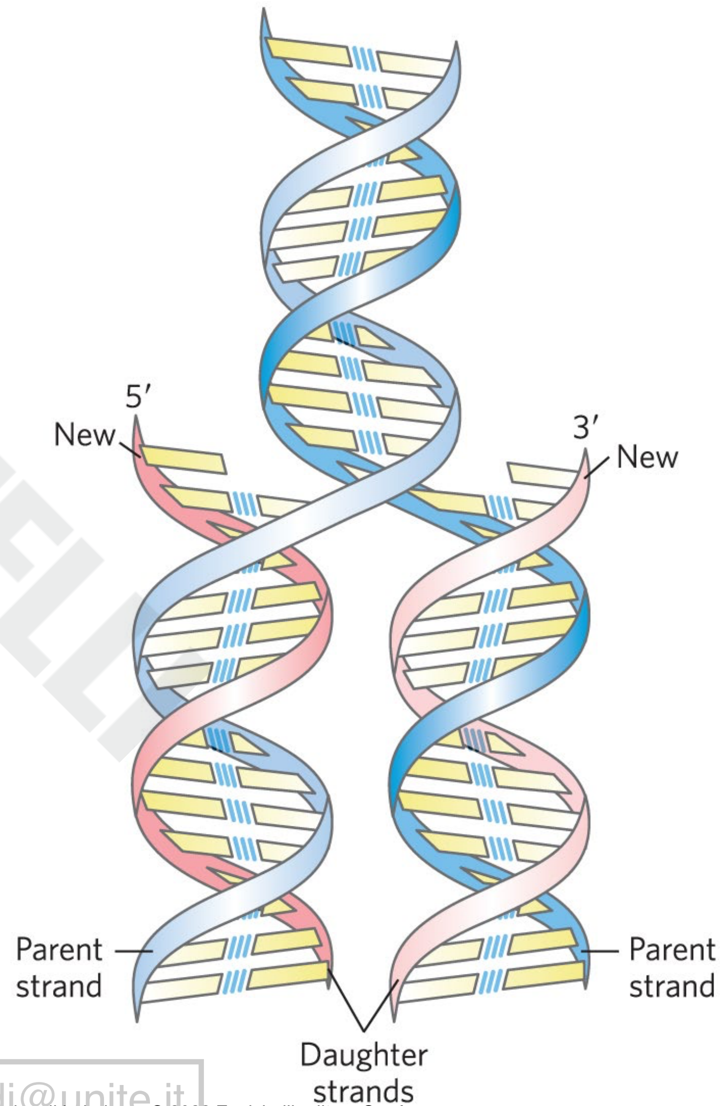
# P2 Principle 2 (3 of 4)

**The transmission of biological information relies on molecular complementarity.** Chromosomes are the largest molecules in any cell. They are polymers composed of a small set of common nucleotides, with information embedded in the nucleotide sequence. The common nucleotides in RNA and DNA are organized so that two strands of nucleic acid can maintain a complementary and uniform structure over vast molecular distances. This extended potential for both variable sequence and complementarity, and thus information storage and transmission, is a property shared by no other class of biological molecule.

# Replication of DNA

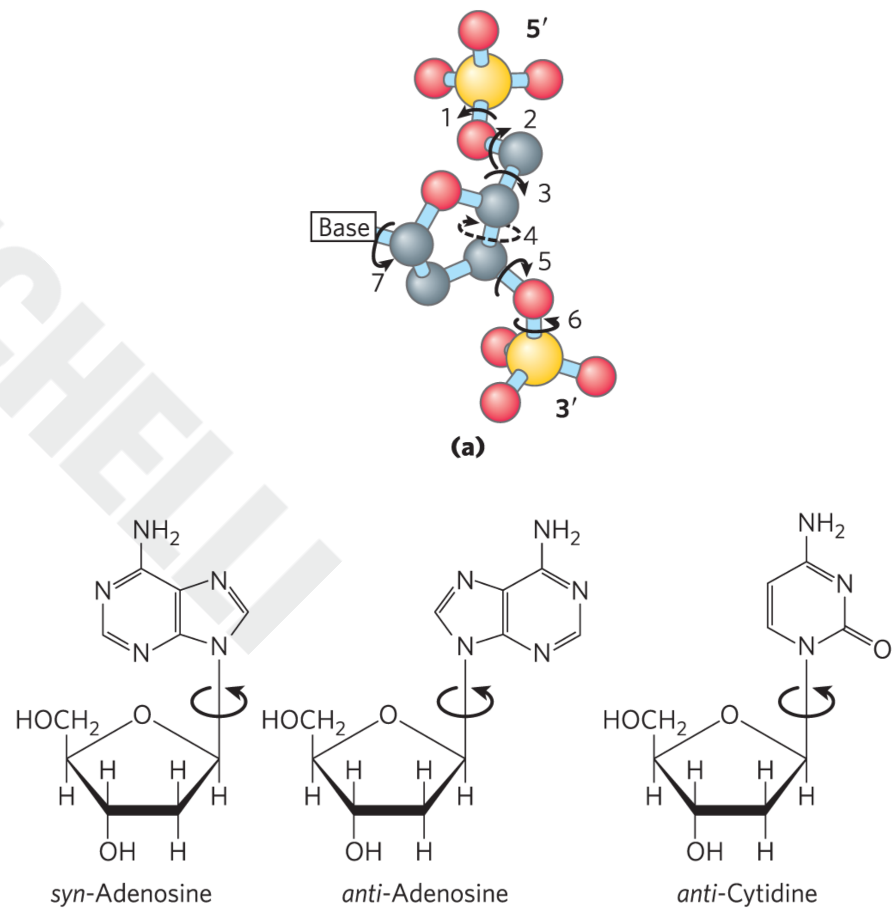
- Step 1: preexisting or "parent" strands become separated

- Step 2: each "parent" strand serves as a template for the biosynthesis of a complementary "daughter strand"



# DNA Can Occur in Different Three-Dimensional Forms

- structural variation in DNA reflects:
  - different possible conformations of the deoxyribose
  - rotation about the contiguous bonds making up the phosphodeoxyribose backbone
  - free rotation about the C-1′–*N*-glycosyl bond

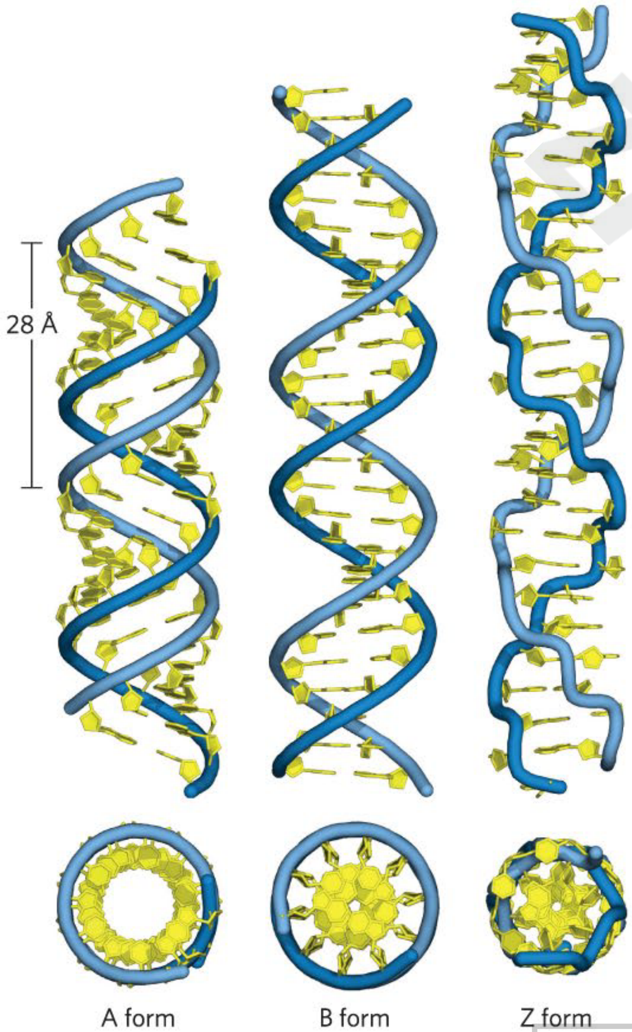



(a) (b)

# The Three Forms of DNA

- **B-form DNA** = the Watson-Crick structure
  - most stable for a random-sequence DNA molecule under physiological conditions
- **A-form DNA** = right-handed double helix with a wider helix, 11 bp/turn, and a tilted plane
  - favored in solutions devoid of water
- **Z-form DNA** = left-handed helix with 12 bp/turn and a backbone with a zig-zag appearance
  - appears more slender and elongated

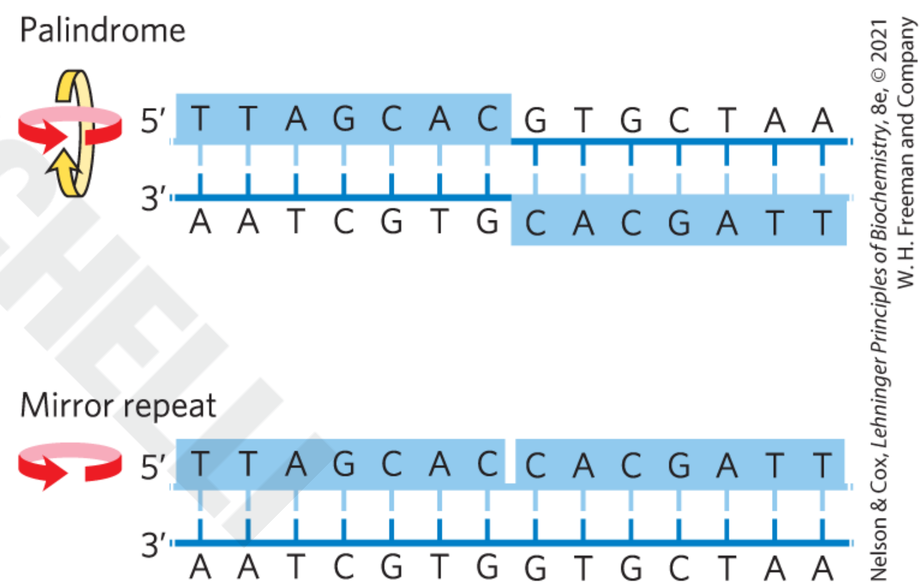
# A, B, and Z Forms of DNA




| | A form | B form | Z form |
|---|---|---|---|
| Helical sense | Right-handed | Right-handed | Left-handed |
| Diameter | ~26 Å | ~20 Å | ~18 Å |
| Base pairs per helical turn | 11 | 10.5 | 12 |
| Helix rise per base pair | 2.6 Å | 3.4 Å | 3.7 Å |
| Base tilt normal to the helix axis | 20° | 6° | 7° |
| Sugar pucker conformation | C-3′ endo | C-2′ endo | C-2′ endo for pyrimidines; C-3′ endo for purines |
| Glycosyl bond conformation | Anti | Anti | Anti for pyrimidines; syn for purines |

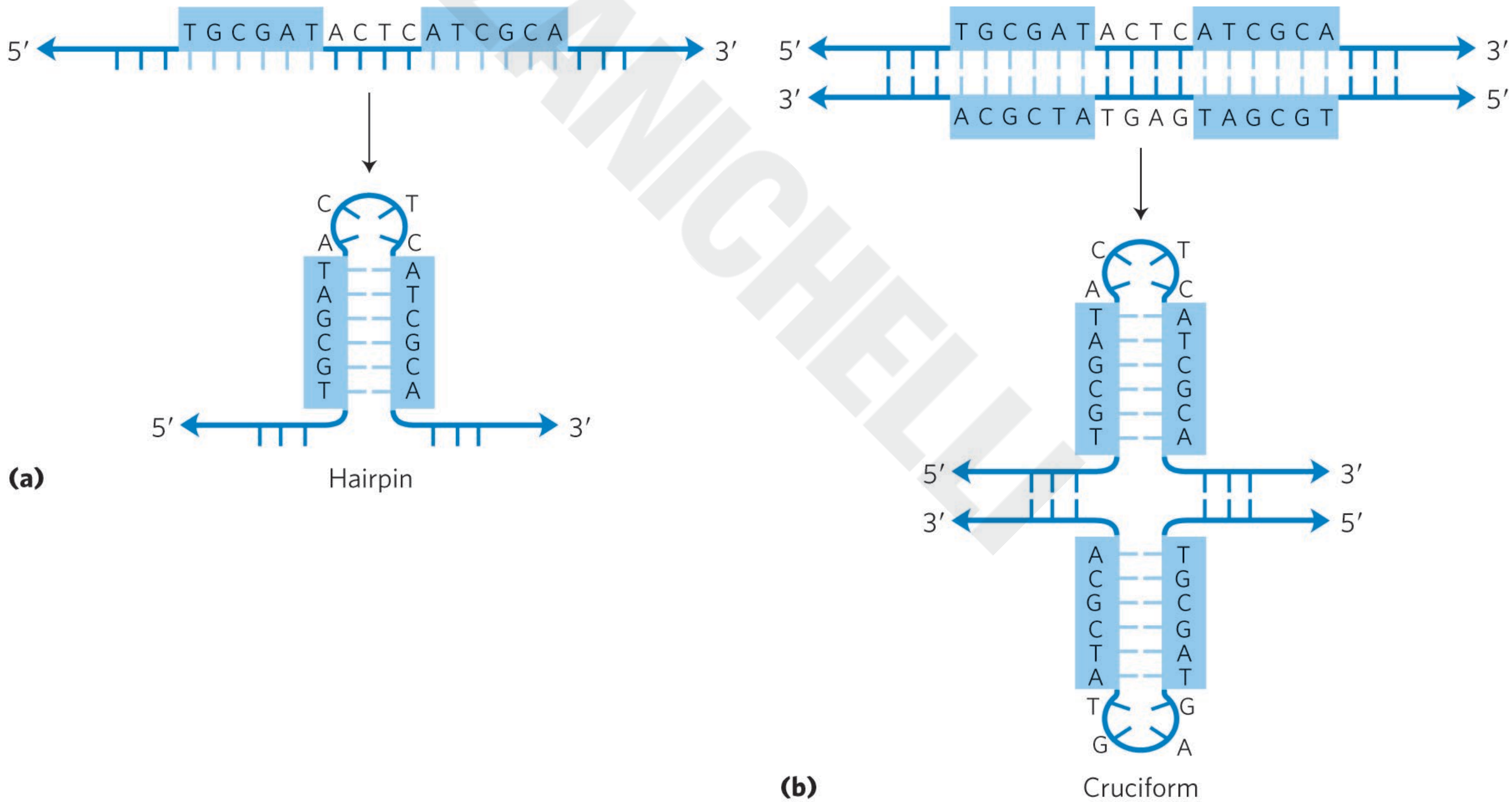
# Certain DNA Sequences Adopt Unusual Structures

- **palindrome** = region of DNA that is identical when read either forward or backward
  - applied to regions of DNA with **inverted repeats**
- **mirror repeat** = sequence when the inverted repeat occurs within each individual strand



# Hairpin and Cruciform Structures

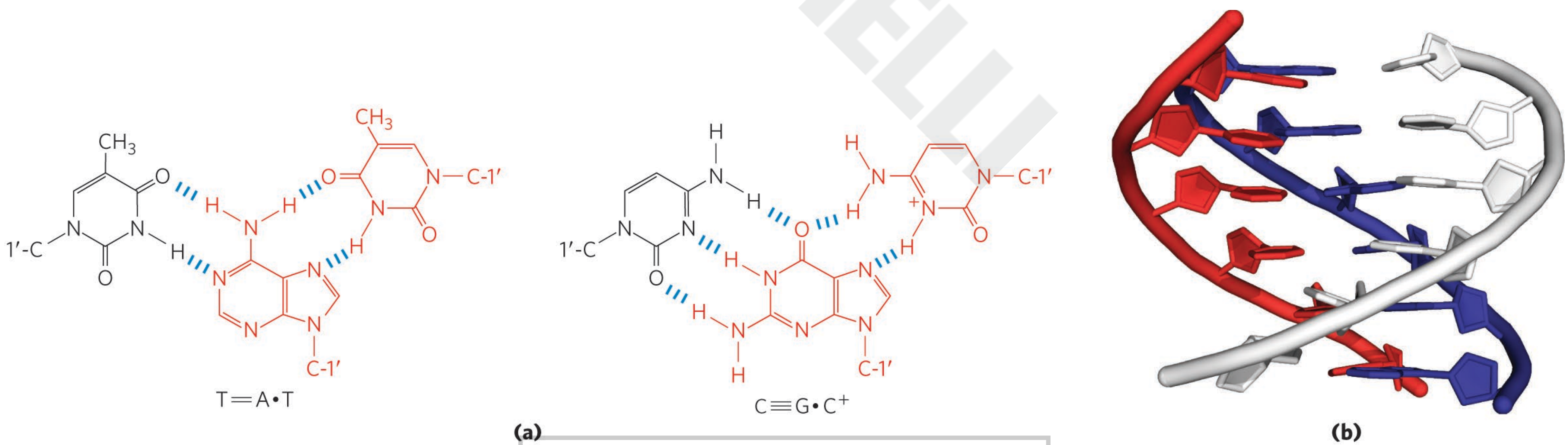
- **hairpin** and **cruciform** structures = form from the self-complementarity within each strand




Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

# DNA Structures Containing Three DNA Strands

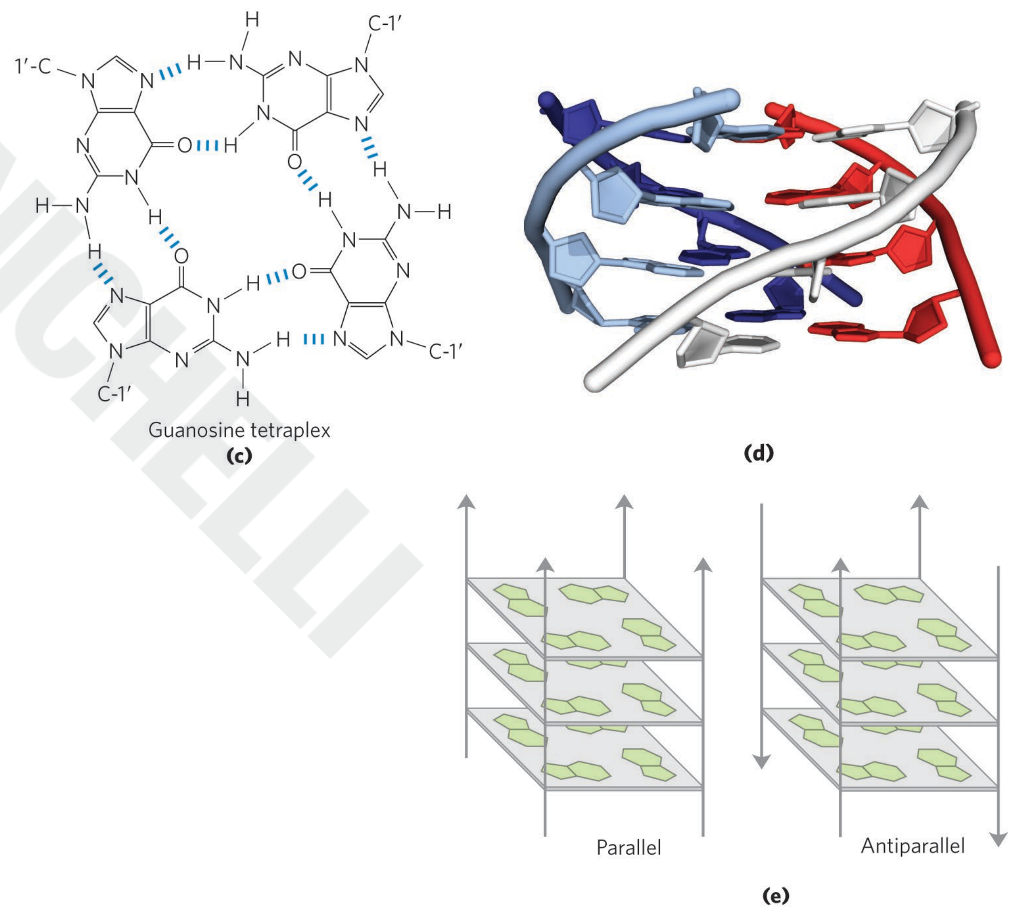
- **Hoogsteen positions** = N-7, $O^6$, and $N^6$ of purines
  - participate in the hydrogen bonding with a third DNA strand
- **Hoogsteen pairing** = the non-Watson-Crick pairing
- **triplex DNAs** = form from Hoogsteen pairing




Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

# DNA Structures Containing Four DNA Strands

- tetraplex DNAs = occur when four DNA strands pair
  - occur readily only for DNA sequences with a very high proportion of G residues
  - **G tetraplex** = very stable




Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

# P1 Principle 1 (5 of 6)

**Nucleic acids are both repositories and functional expressions of biological information.** Biological information is one of the required conditions for life, a blueprint for each species transmitted from one generation to the next. RNA can be a functional expression of that information, directing the synthesis of proteins or, in some cases, acting directly as a signal or a reaction catalyst.

# P2 Principle 2 (4 of 4)

**The transmission of biological information relies on molecular complementarity.** Chromosomes are the largest molecules in any cell. They are polymers composed of a small set of common nucleotides, with information embedded in the nucleotide sequence. The common nucleotides in RNA and DNA are organized so that two strands of nucleic acid can maintain a complementary and uniform structure over vast molecular distances. This extended potential for both variable sequence and complementarity, and thus information storage and transmission, is a property shared by no other class of biological molecule.

# Messenger RNAs Code for Polypeptide Chains

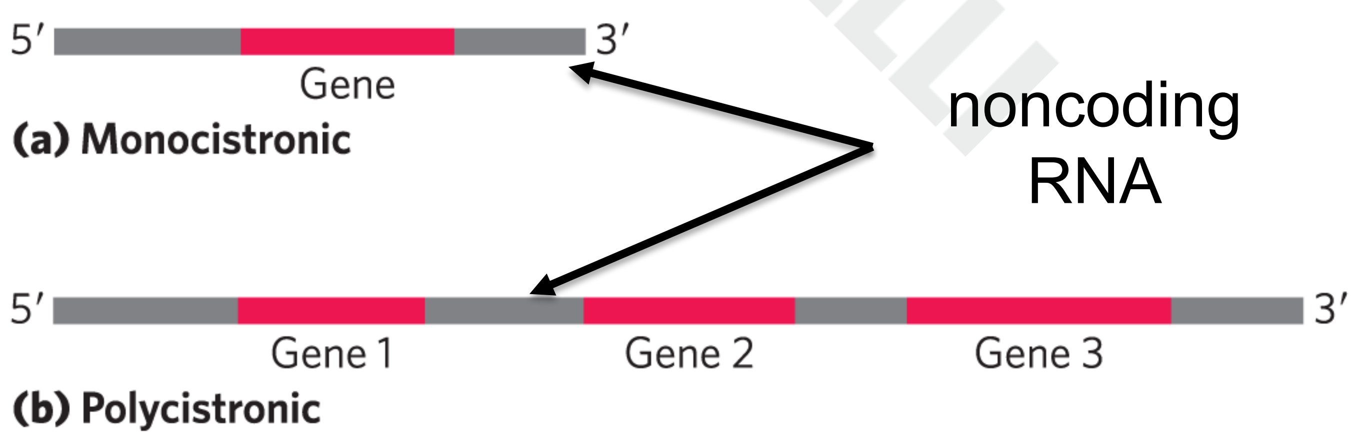
- “messenger RNA” (mRNA) = portion of cellular RNA carrying the genetic information from DNA to the ribosome

- **transcription** = process by which mRNAs are formed on a DNA template

# mRNA Can Be Monocistronic or Polycistronic

- **monocistronic** = codes for only one polypeptide
  - most mRNAs in eukaryotes

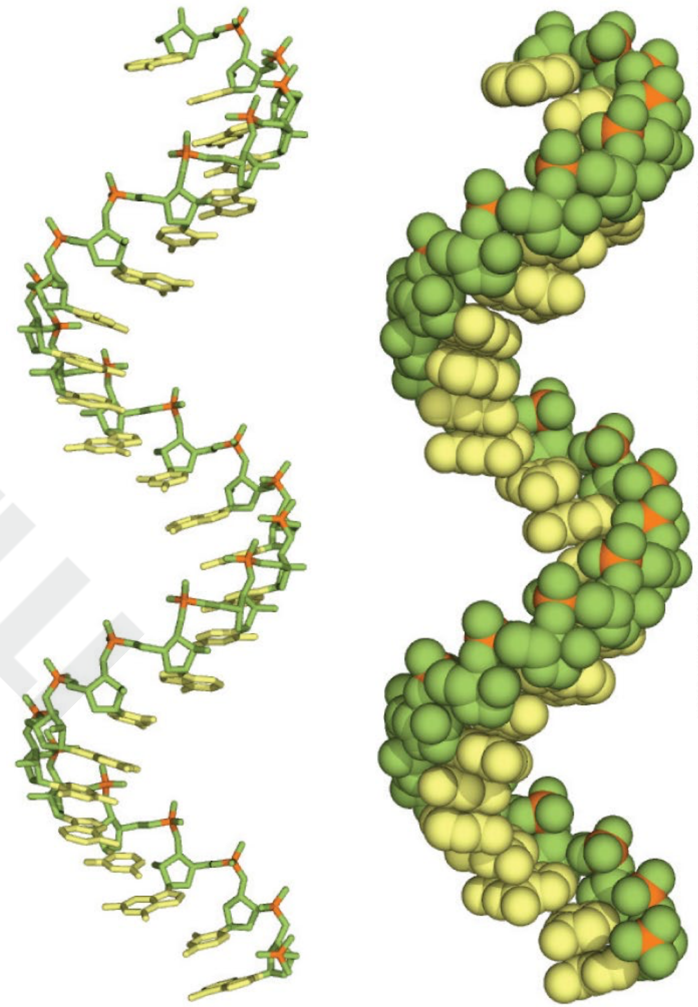
- **polycistronic** = codes for 2+ different polypeptides
  - occurs in bacteria and archaea




Nelson & Cox, *Lehninger Principles of Biochemistry*, 8e, © 2021 W. H. Freeman and Company

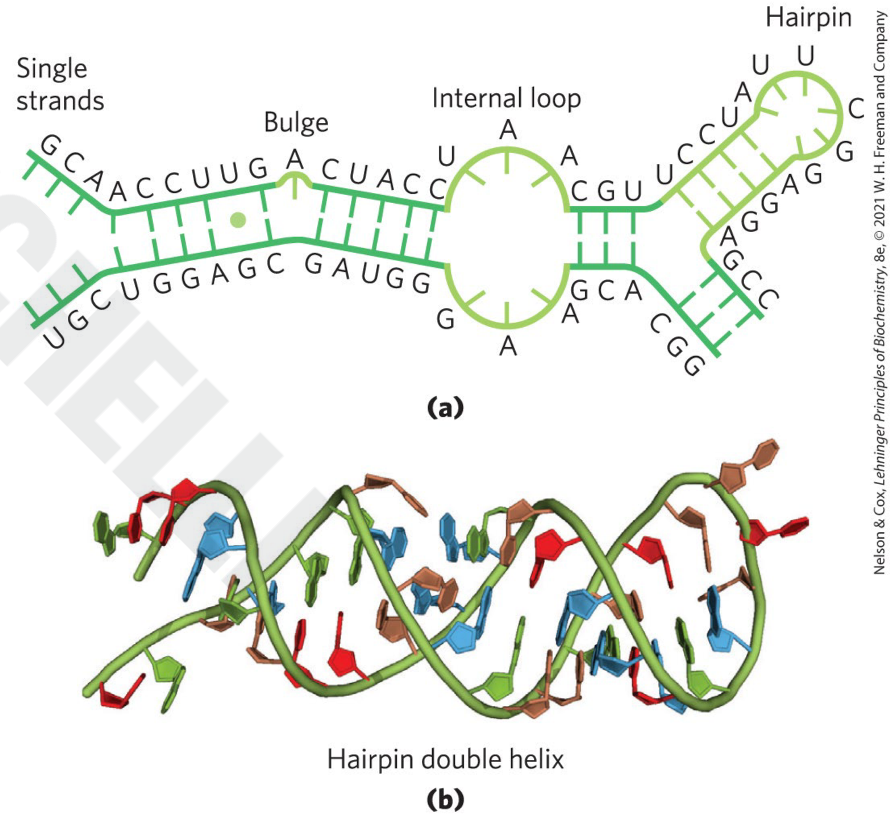
# Many RNAs Have More Complex Three-Dimensional Structures

- mRNA is always single-stranded
  - right-handed helical conformation
  - dominated by base-stacking interactions
    - strongest between two purines
  - can base pair with complementary regions of DNA or RNA
    - paired strands are antiparallel



# Secondary Structure of RNAs

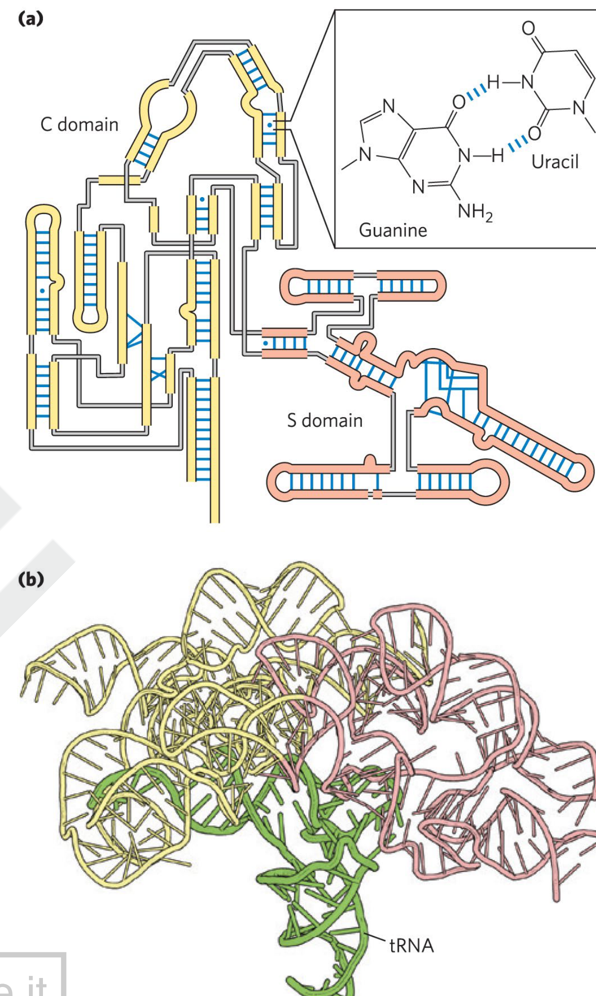
- structure of complementary RNA strands is an A-form right-handed double helix
- breaks caused by mismatched or unmatched bases result in bulges or internal loops
- internal loops form between palindromic sequences



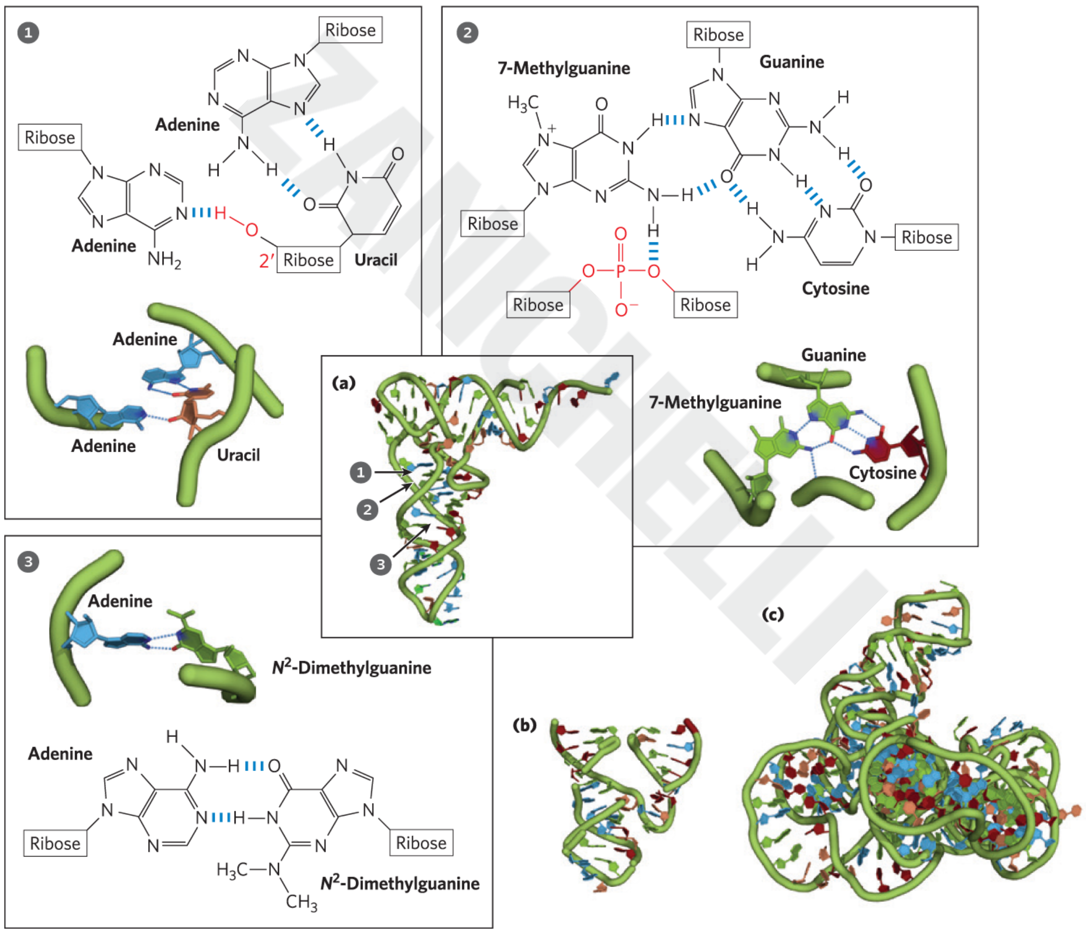
# Base-Paired Helical Structures In RNA

- extensive base-paired helical segments form in RNAs

- hairpins are the most common type of secondary structure



# Three-Dimensional Structure in RNA



# 8.3 Nucleic Acid Chemistry

# P1 Principle 1 (6 of 6)

**Nucleic acids are both repositories and functional expressions of biological information.** Biological information is one of the required conditions for life, a blueprint for each species transmitted from one generation to the next. RNA can be a functional expression of that information, directing the synthesis of proteins or, in some cases, acting directly as a signal or a reaction catalyst.

# P3 Principle 3 (2 of 3)

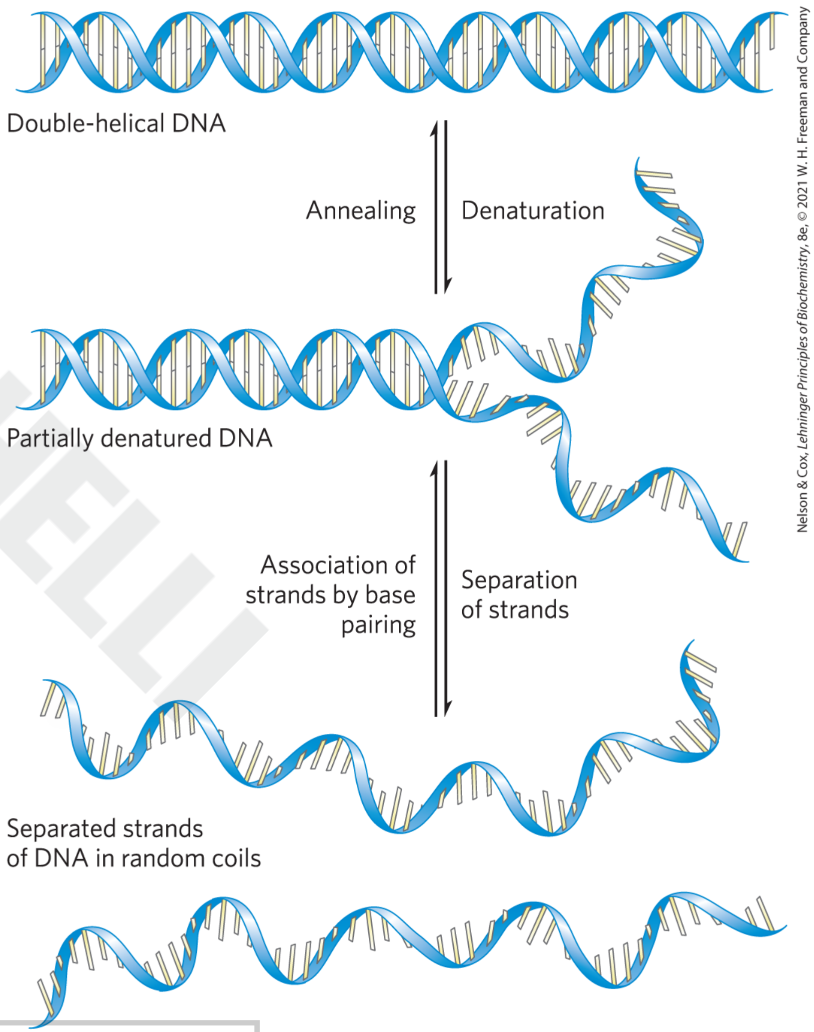
**Biological information is subject to natural damage and change.** DNA damage is a constant, and it results in occasional mutation — the raw material for evolution.

# The Role of DNA as a Repository of Genetic Information

- depends in part on its inherent stability
- chemical transformations are generally very slow in the absence of enzyme catalysts
- carcinogenesis and aging may be linked to slowly accumulating, irreversible DNA alternations

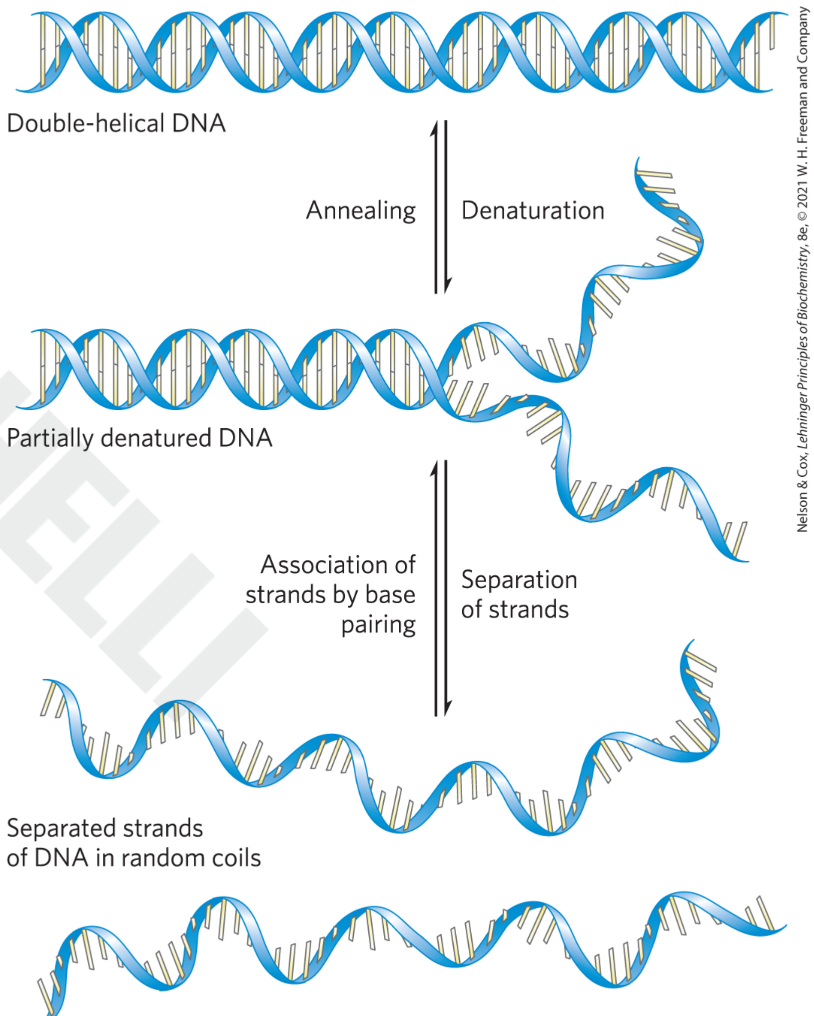
# Double-Helical DNA and RNA Can Be Denatured

- denaturation, or melting, of the double helix:
  - due to pH extremes or high temperatures
  - disrupts hydrogen bonds and base-stacking interactions



# Double-Helical DNA and RNA Can Anneal

- **anneal** = process by which two strands spontaneously rewind when temperature or pH is returned to its normal range
  - two-step process



# The Hypochromic and Hyperchromic Effects

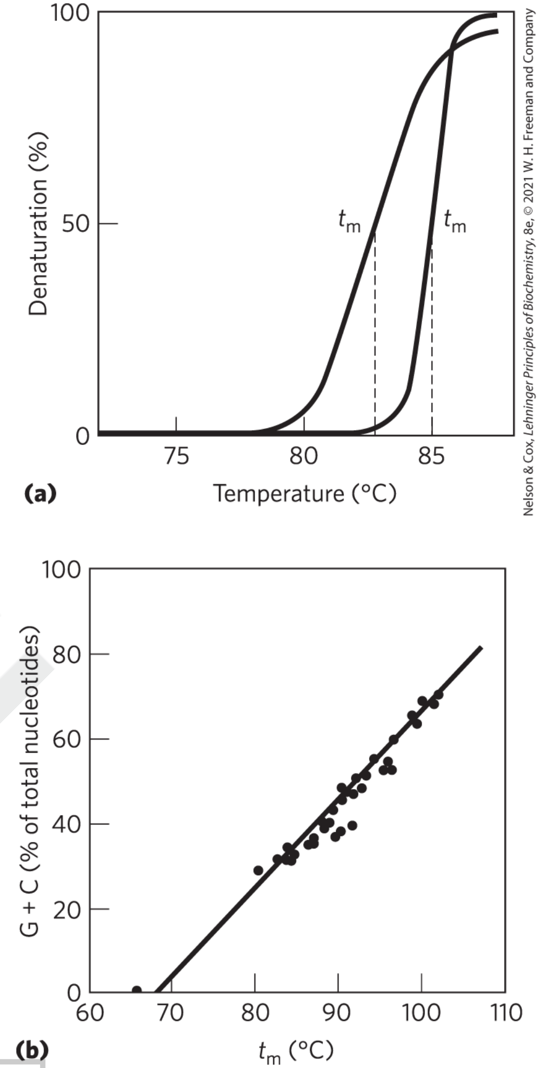
- **hypochromic effect** = the observed decrease in the absorption of UV light when complementary strands are paired

- **hyperchromic effect** = the observed increase in the absorption of UV light when a double-stranded nucleic acid is denatured

- monitoring UV absorption at 260 nm can detect the transition from double-stranded to single-stranded DNA

# Heat Denaturation of DNA

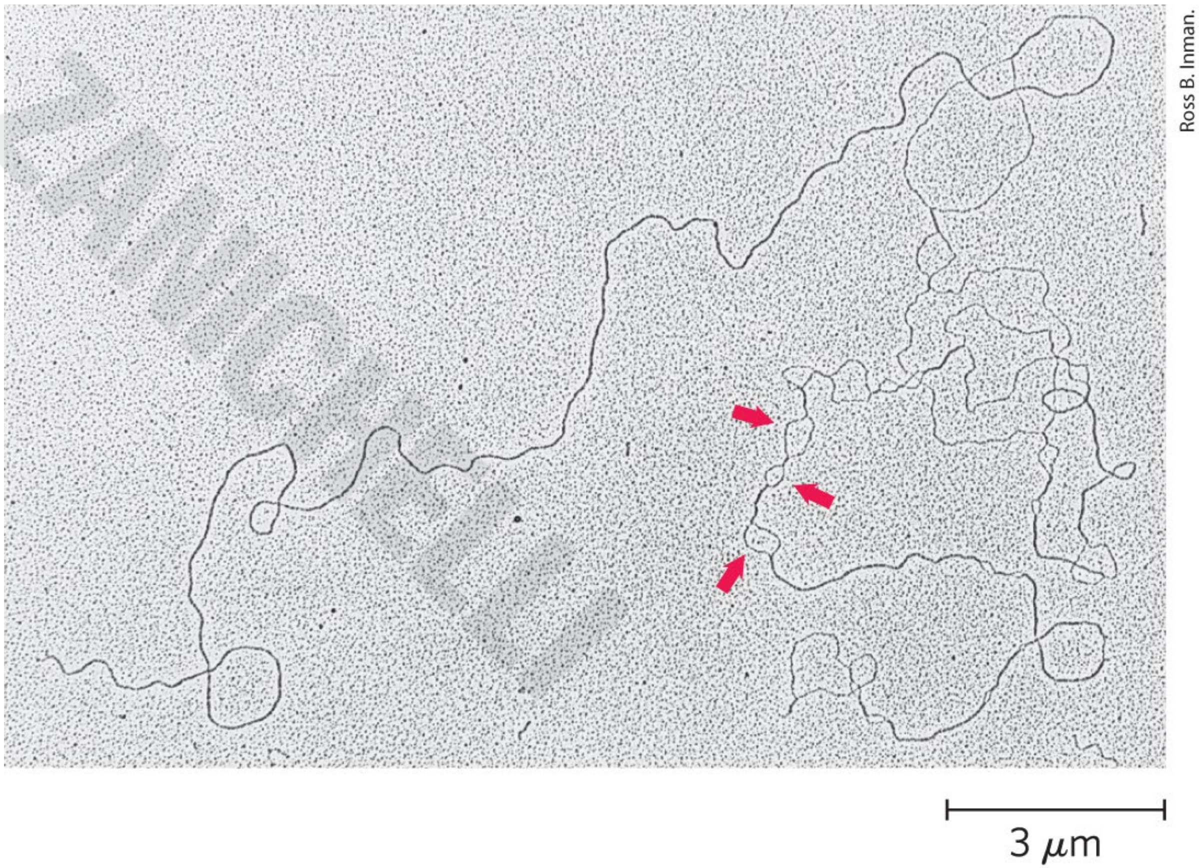
- denaturation temperature, $t_m$ = temperature at which ½ of DNA is present as separated single strands
  - increases with content of G≡C base pairs



# Partially Denatured DNA

- denatured regions form bubbles

- often rich in A=T base pairs



# Denaturation of Double-Stranded RNA and RNA-DNA Hybrids

- RNA duplexes are more stable to heat denaturation than DNA duplexes
- RNA-DNA hybrid stability is generally intermediate

# Nucleotides and Nucleic Acids Undergo Nonenzymatic Transformations

- **mutations** = alterations in DNA structure that produce permanent changes in the genetic information encoded
  - linked to aging and carcinogenesis