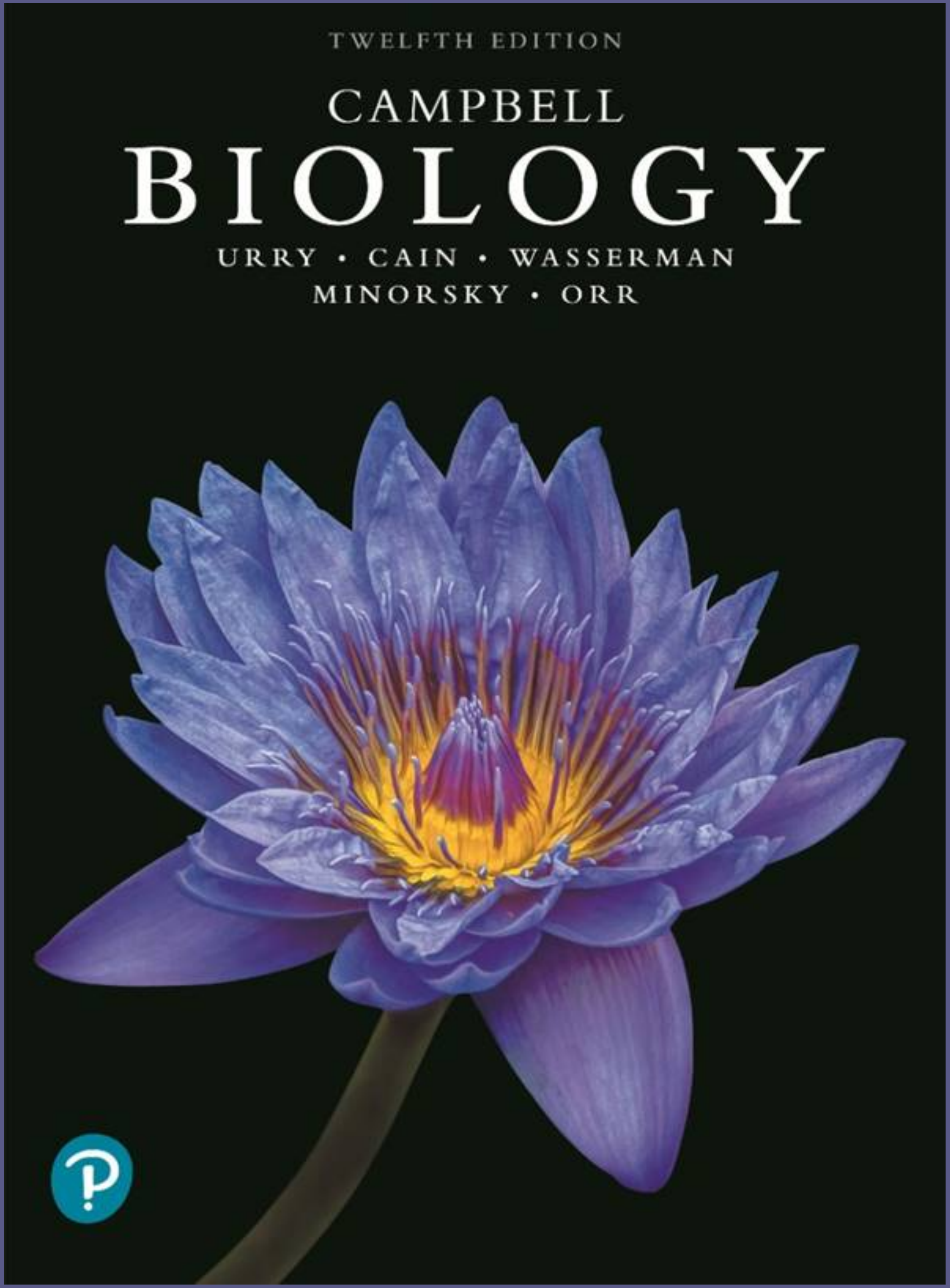


Chapter 5

The Structure and Function of Large Biological Molecules

Lecture Presentations by
Nicole Tunbridge and
Kathleen Fitzpatrick

Figure 5.1

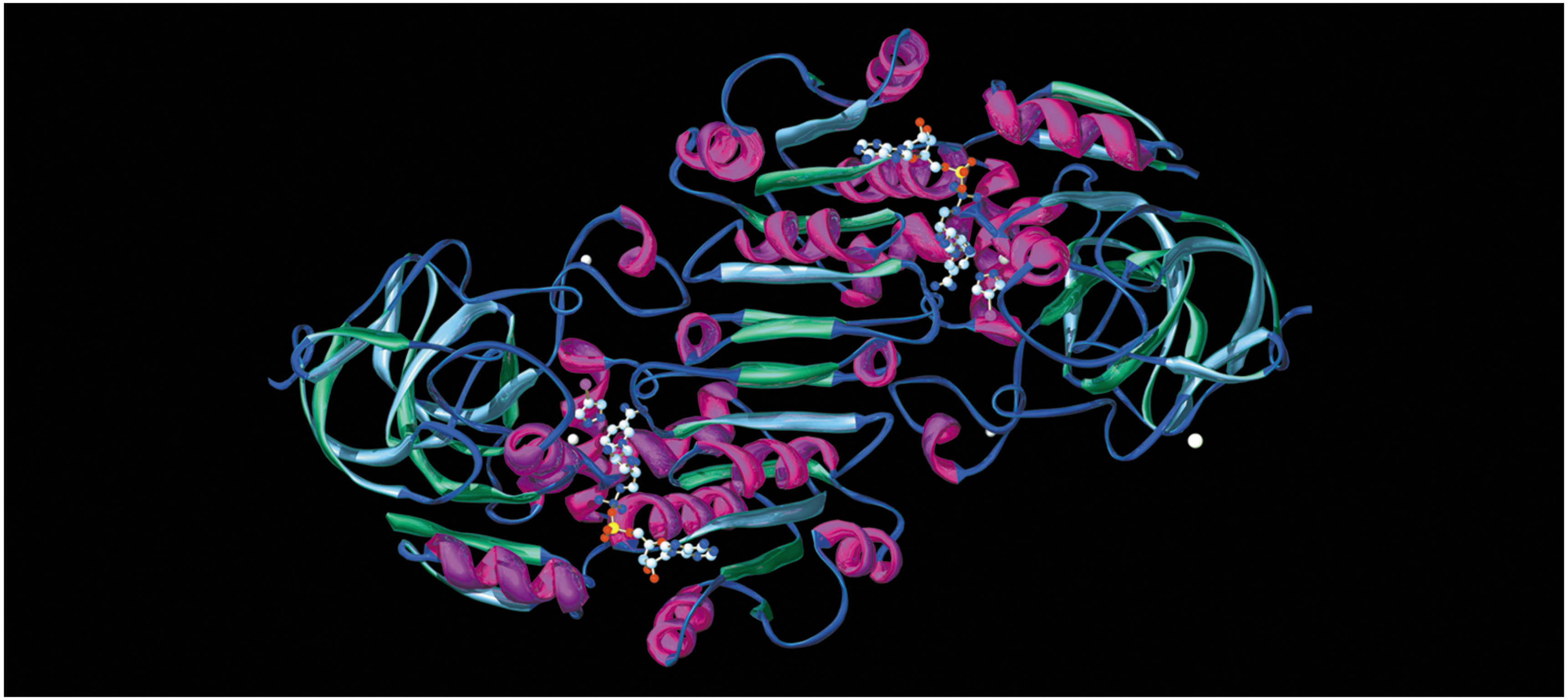
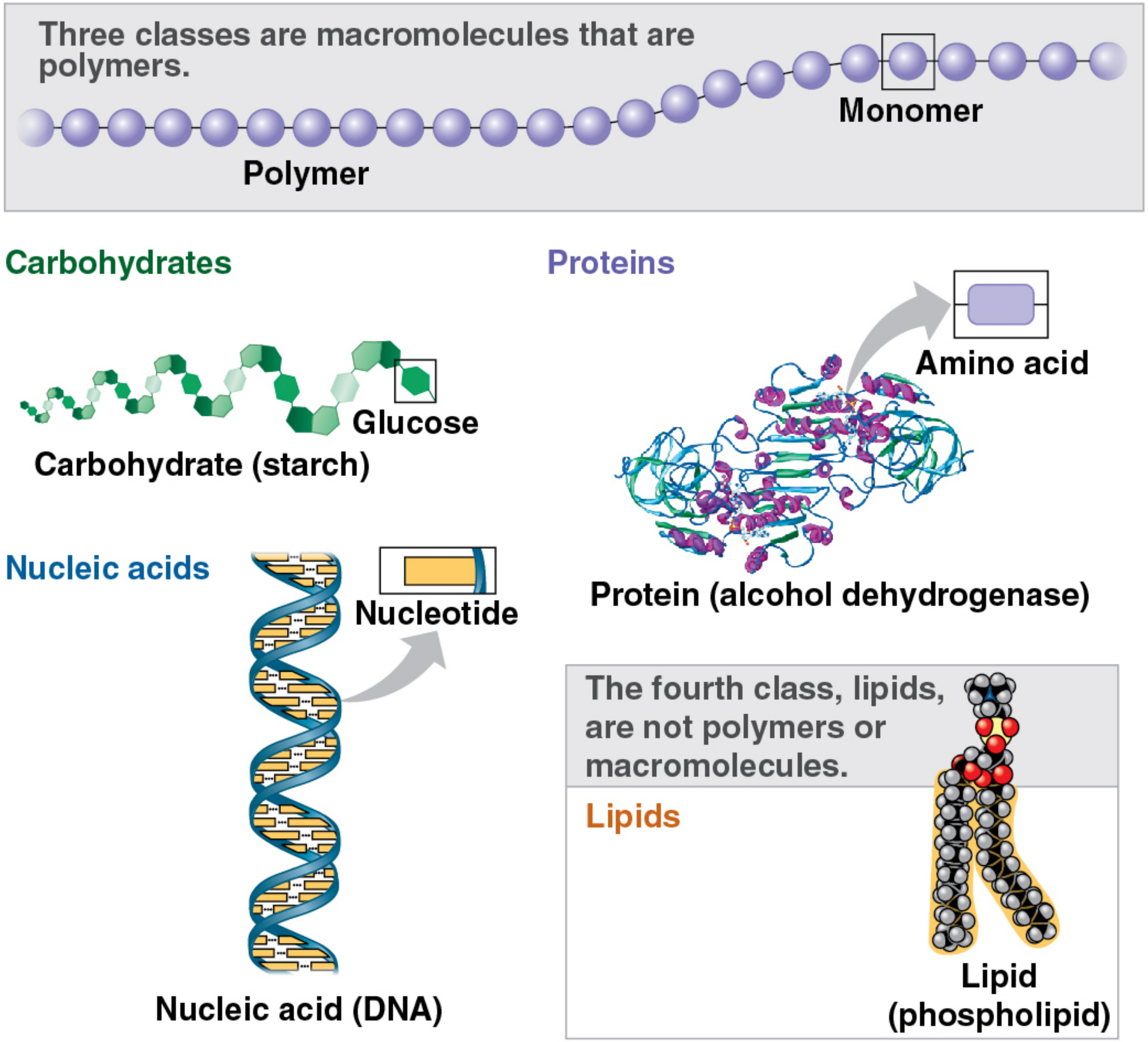


Figure 5.1a

What are the structures and functions of the four important classes of biological molecules?



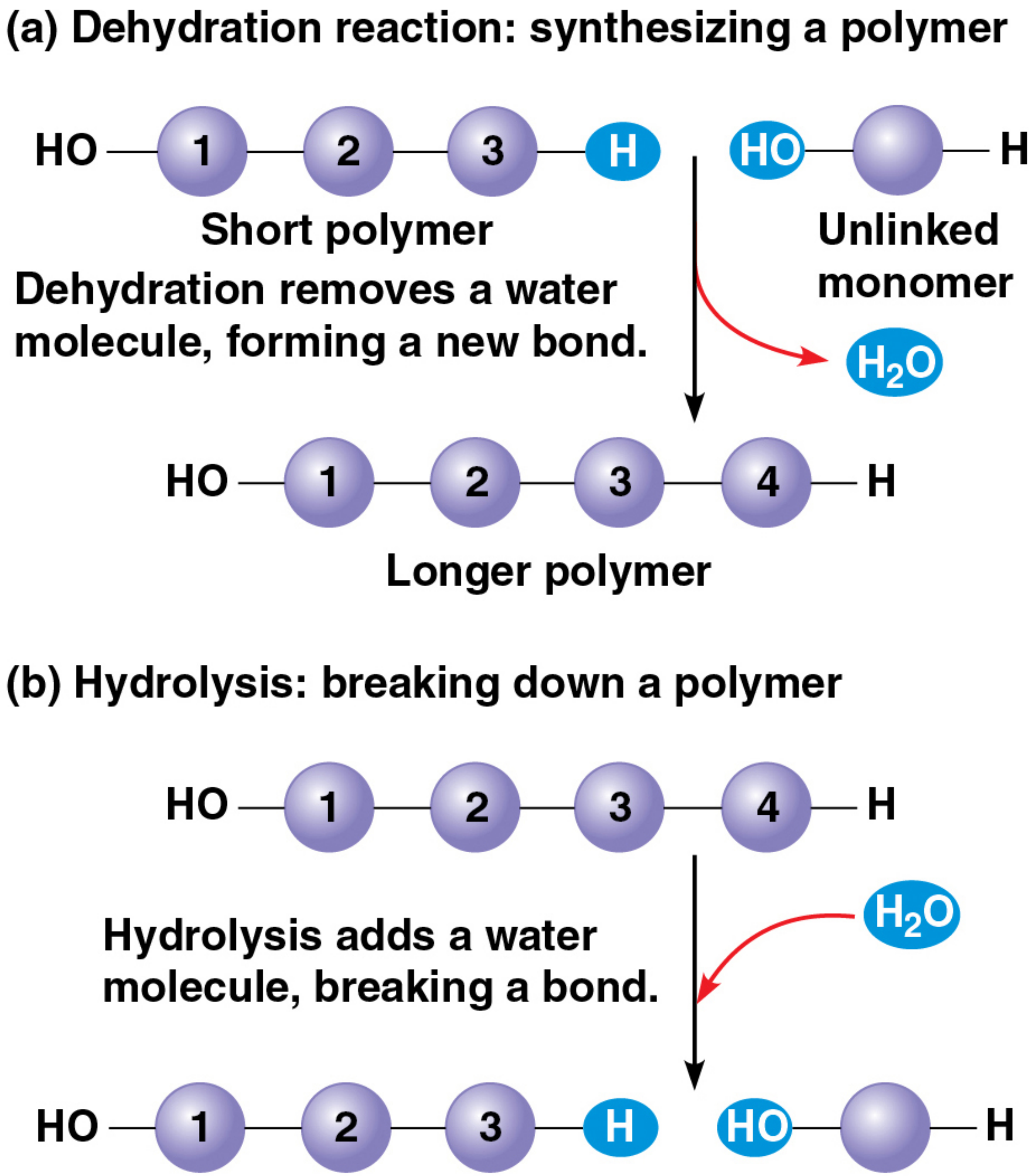
CONCEPT 5.1: Macromolecules are polymers, built from monomers

- Large polymers are known as **macromolecules** for their huge size
- A **polymer** is a long molecule consisting of many similar building blocks
- The repeating units that serve as building blocks are called **monomers**
- Carbohydrates, proteins, and nucleic acids are polymers

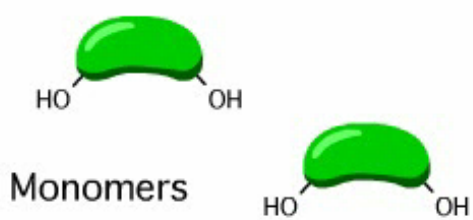
The Synthesis and Breakdown of Polymers

- **Enzymes** are specialized macromolecules that speed up chemical reactions such as those that make or break down polymers
- A **dehydration reaction** occurs when two monomers bond together through the loss of a water molecule
- Polymers are disassembled to monomers by **hydrolysis**, a reaction that is essentially the reverse of the dehydration reaction

Figure 5.2


(a) Dehydration reaction: synthesizing a polymer
HO 1 2 3 H
HO H
Short polymer
Unlinked monomer
Dehydration removes a water molecule, forming a new bond.
H_2O
HO 1 2 3 4 H
Longer polymer
(b) Hydrolysis: breaking down a polymer
HO 1 2 3 4 H
Hydrolysis adds a water molecule, breaking a bond.
H_2O
HO 1 2 3 H
HO H

Video: Polymers



The Diversity of Polymers

- A cell has thousands of different macromolecules
- Macromolecules vary among cells of an organism, vary more within a species, and vary even more between species
- A huge variety of polymers can be built from a small set of monomers

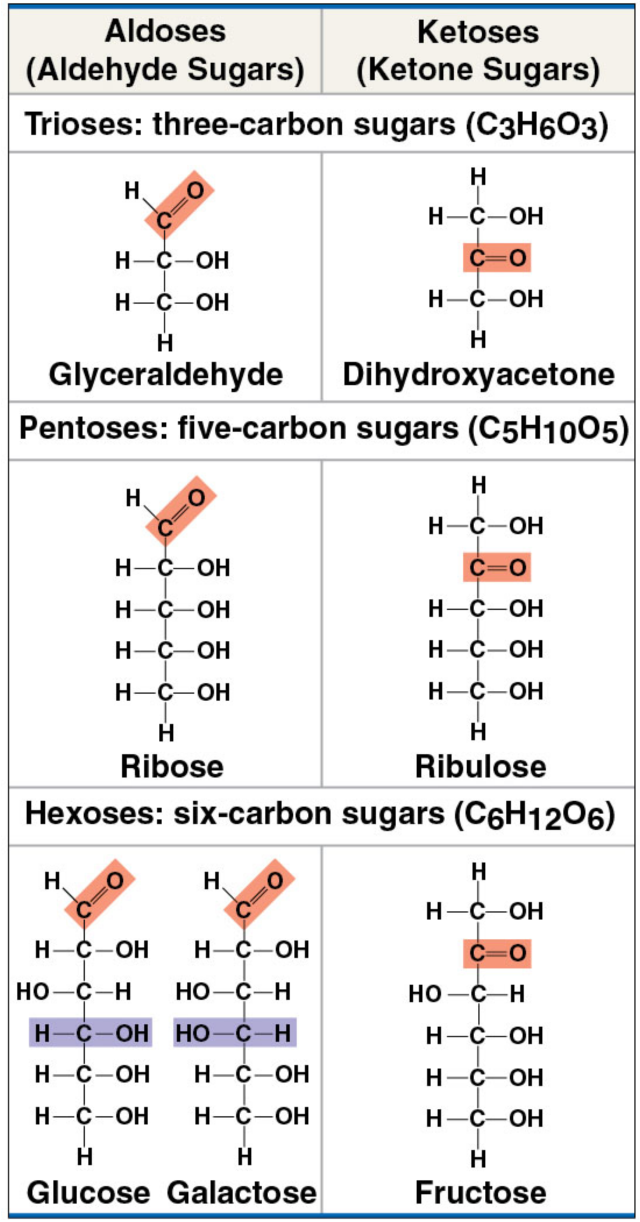
CONCEPT 5.2: Carbohydrates serve as fuel and building material

- **Carbohydrates** include sugars and polymers of sugars
- The simplest carbohydrates are monosaccharides, or simple sugars
- Carbohydrate macromolecules are polysaccharides, polymers composed of many sugar building blocks

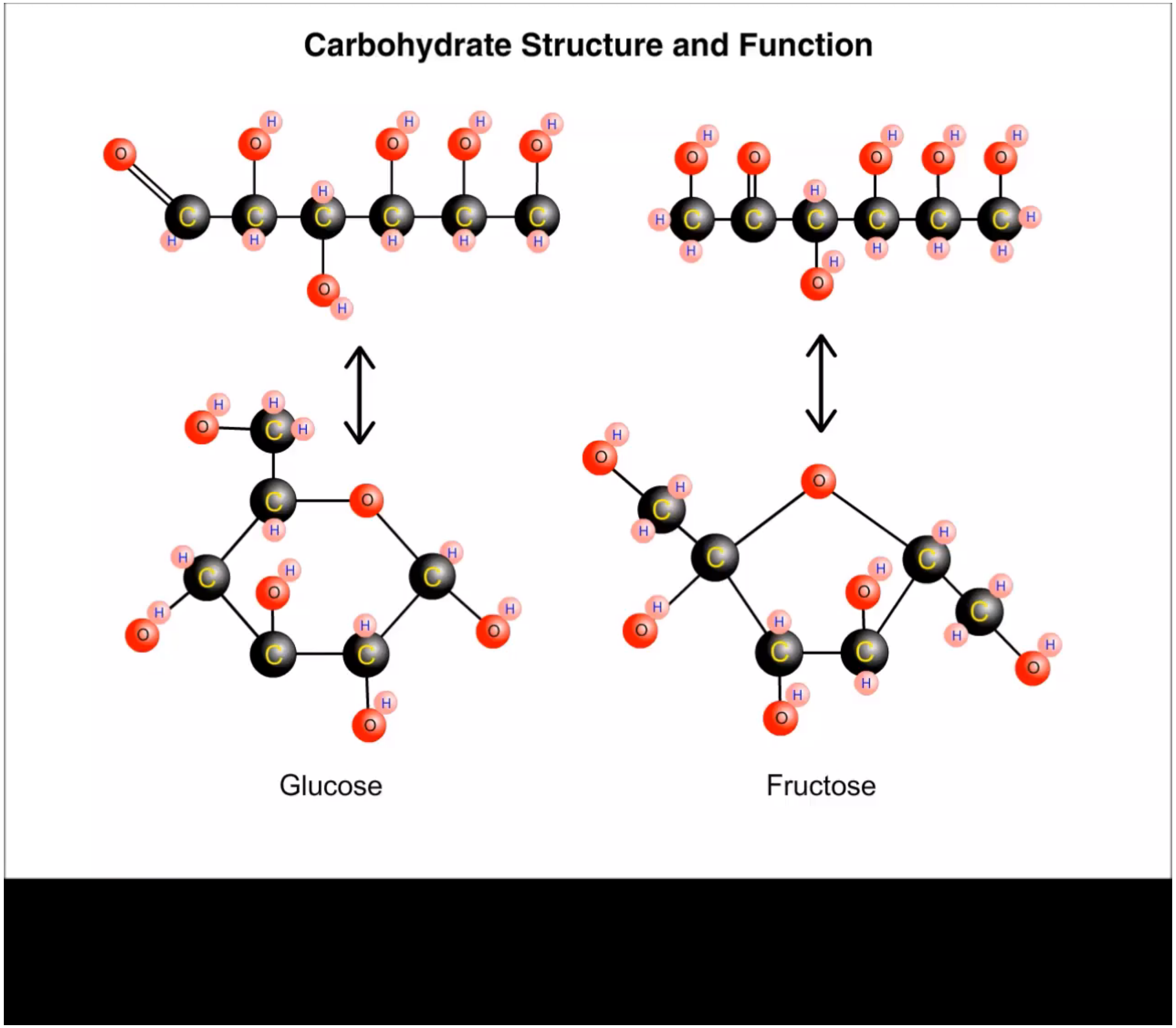
Sugars

- **Monosaccharides** have molecular formulas that are usually multiples of CH_2O
- Glucose ($C_6H_{12}O_6$) is the most common monosaccharide
- Monosaccharides are classified by
  - The location of the carbonyl group (as aldose or ketose)
  - The number of carbons in the carbon skeleton

Figure 5.3

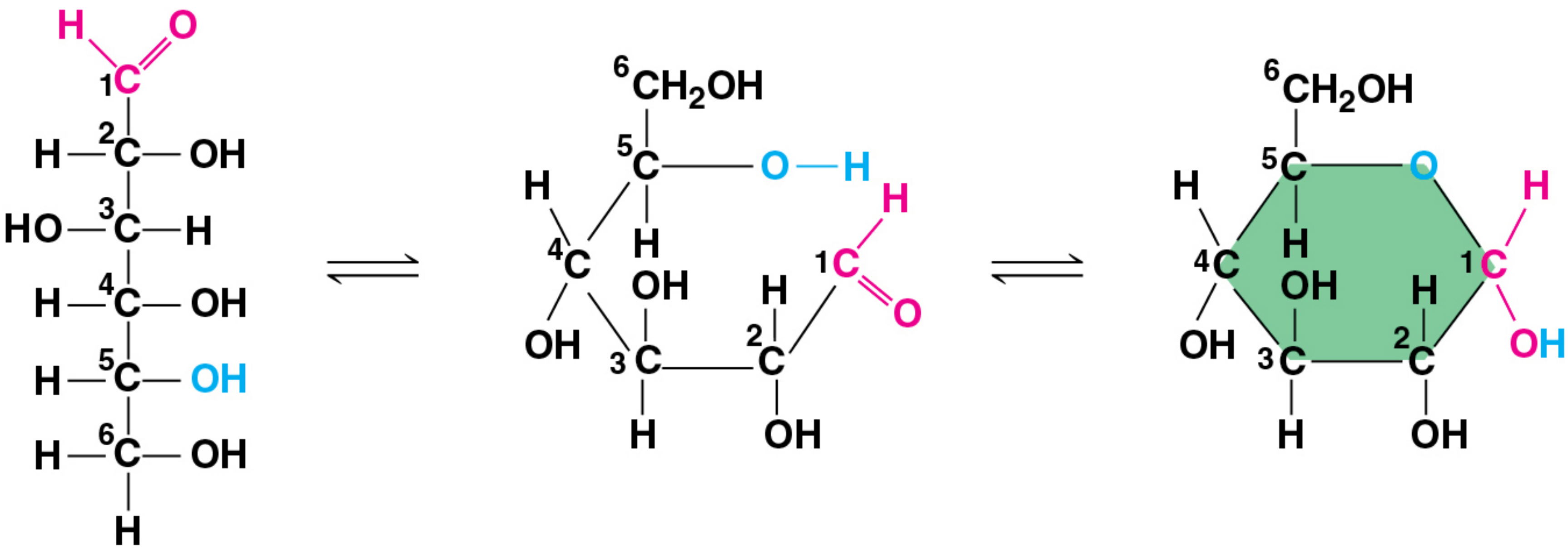

Aldoses (Aldehyde Sugars)
Ketoses (Ketone Sugars)
Trioses: three-carbon sugars ($C_3H_6O_3$)
Glyceraldehyde
Dihydroxyacetone
Pentoses: five-carbon sugars ($C_5H_{10}O_5$)
Ribose
Ribulose
Hexoses: six-carbon sugars ($C_6H_{12}O_6$)
Glucose
Galactose
Fructose

Video: Synthesis of Sucrose

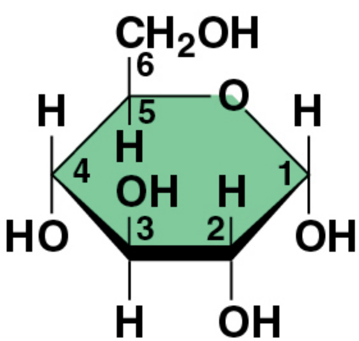


- Though often drawn as linear skeletons, in aqueous solutions many sugars form rings
- Monosaccharides serve as a major fuel for cells and as raw material for building molecules

Figure 5.4



(a) Linear and ring forms

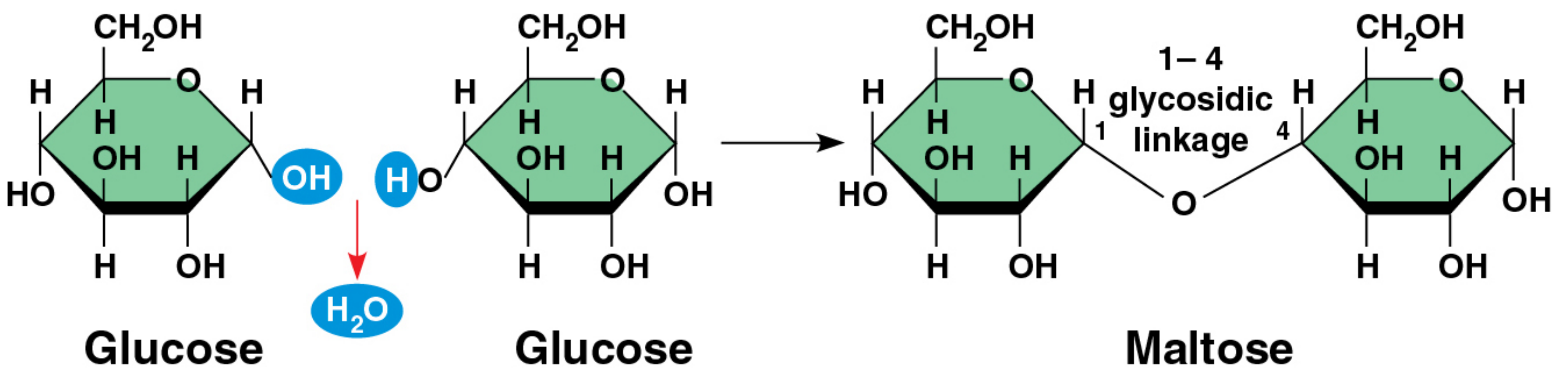


(b) Abbreviated ring structure

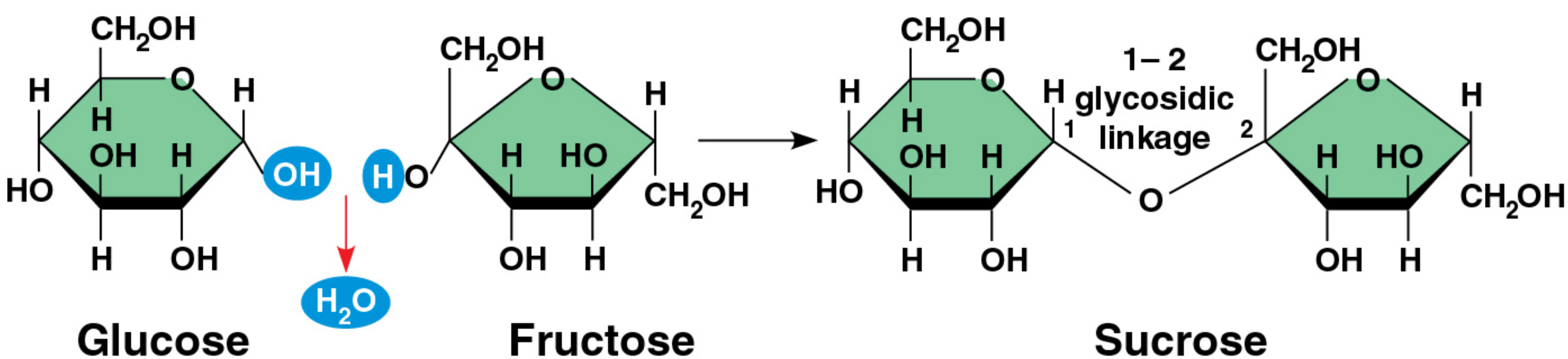
- A **disaccharide** is formed when a dehydration reaction joins two monosaccharides
- This covalent bond between two monosaccharides is called a **glycosidic linkage**

Figure 5.5

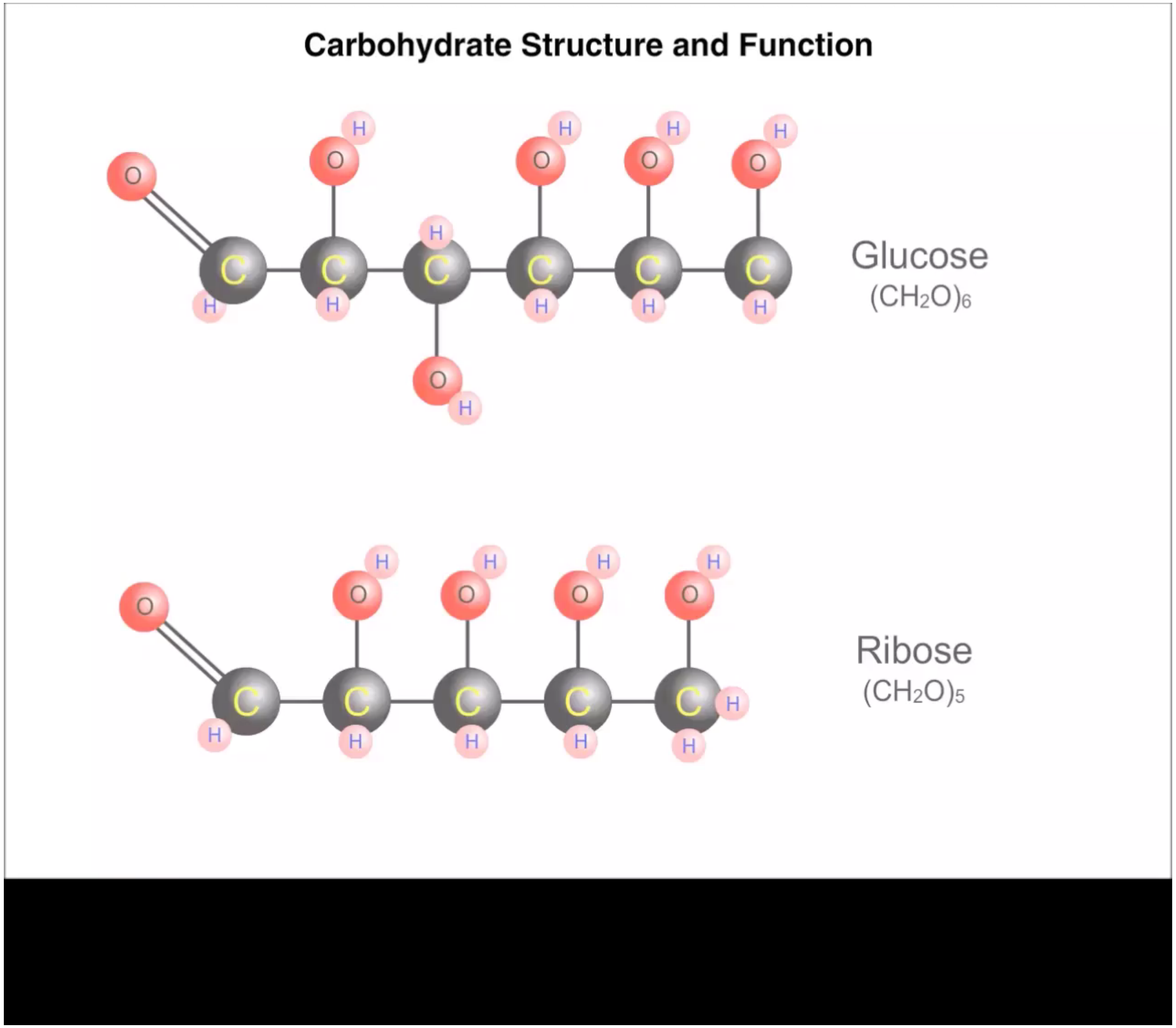
(a) Dehydration reaction in the synthesis of maltose



(b) Dehydration reaction in the synthesis of sucrose



Video: Monosaccharides



Polysaccharides

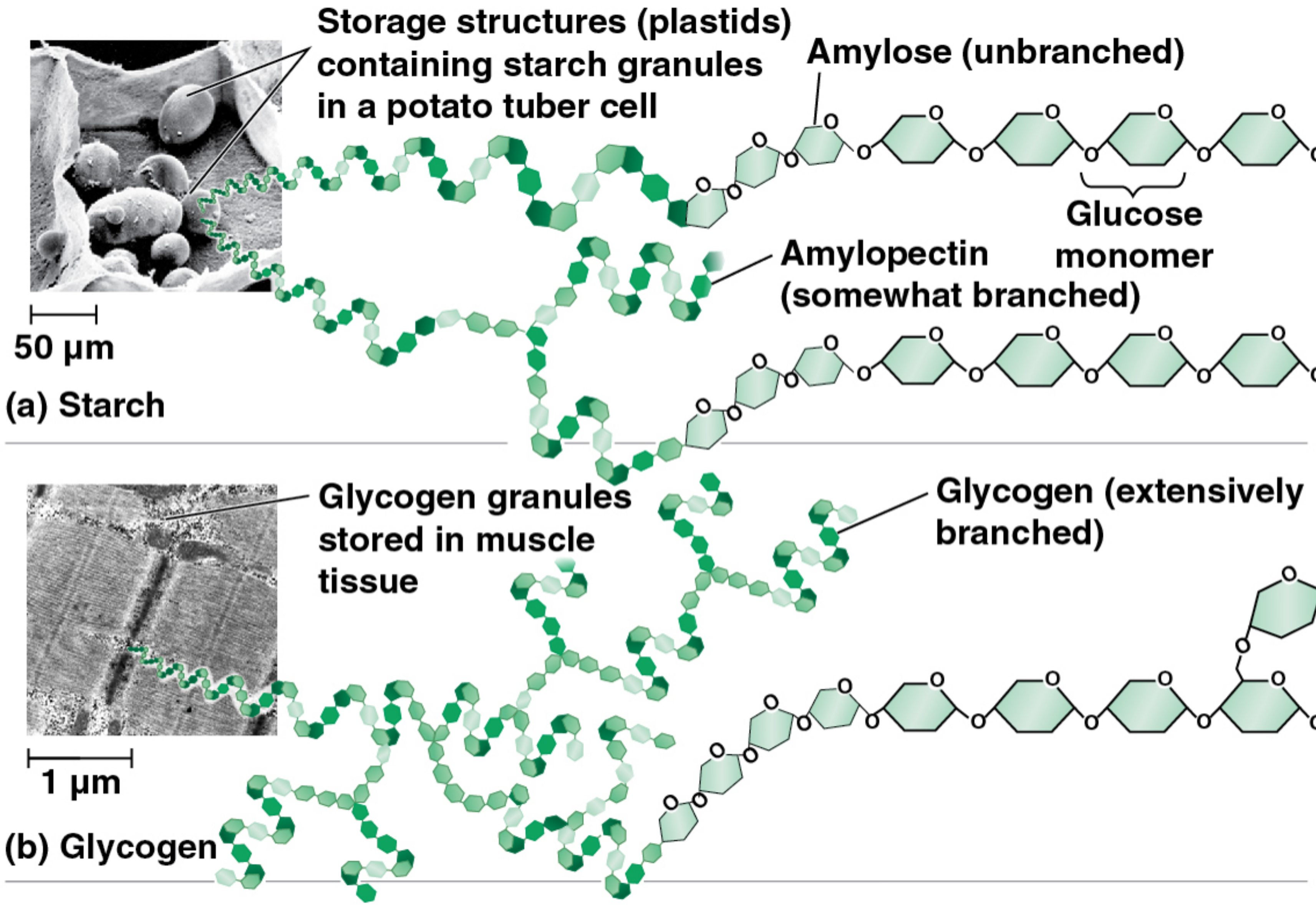
- **Polysaccharides**, the polymers of sugars, have storage and structural roles
- The architecture and function of a polysaccharide are determined by its sugar monomers and the positions of its glycosidic linkages

Storage Polysaccharides

- **Starch**, a storage polysaccharide of plants, consists of glucose monomers
- Plants store surplus starch as granules within chloroplasts and other plastids
- The simplest form of starch is amylose

- **Glycogen** is a storage polysaccharide in animals
- Glycogen is stored mainly in liver and muscle cells
- Hydrolysis of glycogen in these cells releases glucose when the demand for sugar increases

Figure 5.6a,b

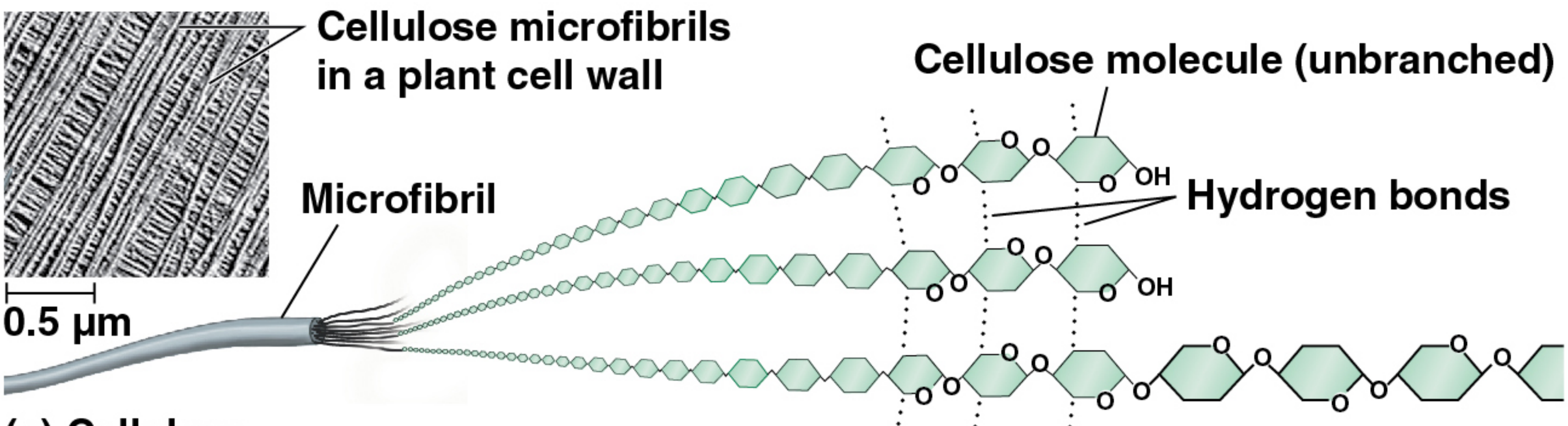


Storage structures (plastids) containing starch granules in a potato tuber cell
Amylose (unbranched)
Glucose monomer
Amylopectin (somewhat branched)
50 µm
(a) Starch
Glycogen granules stored in muscle tissue
Glycogen (extensively branched)
1 µm
(b) Glycogen

Structural Polysaccharides

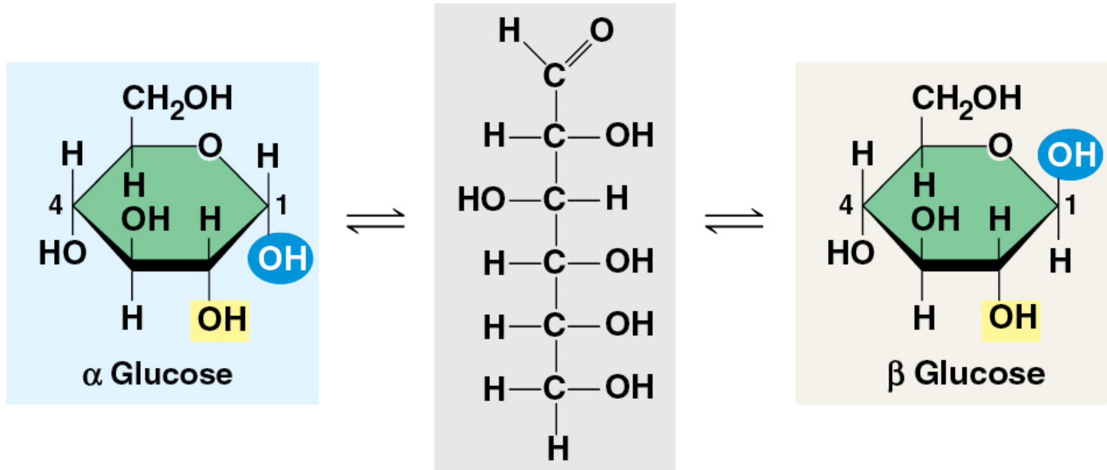
- The polysaccharide **cellulose** is a major component of the tough wall of plant cells
- Like starch, cellulose is a polymer of glucose, but the glycosidic linkages differ
- The difference is based on two ring forms for glucose: alpha (α) and beta (β)

Figure 5.6c

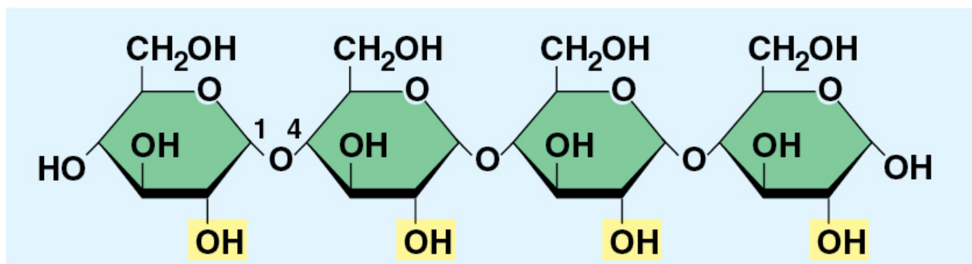


(c) Cellulose

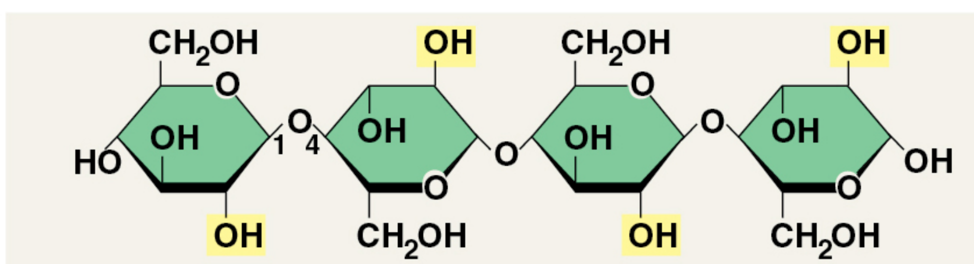
Figure 5.7



(a) α and β glucose ring structures

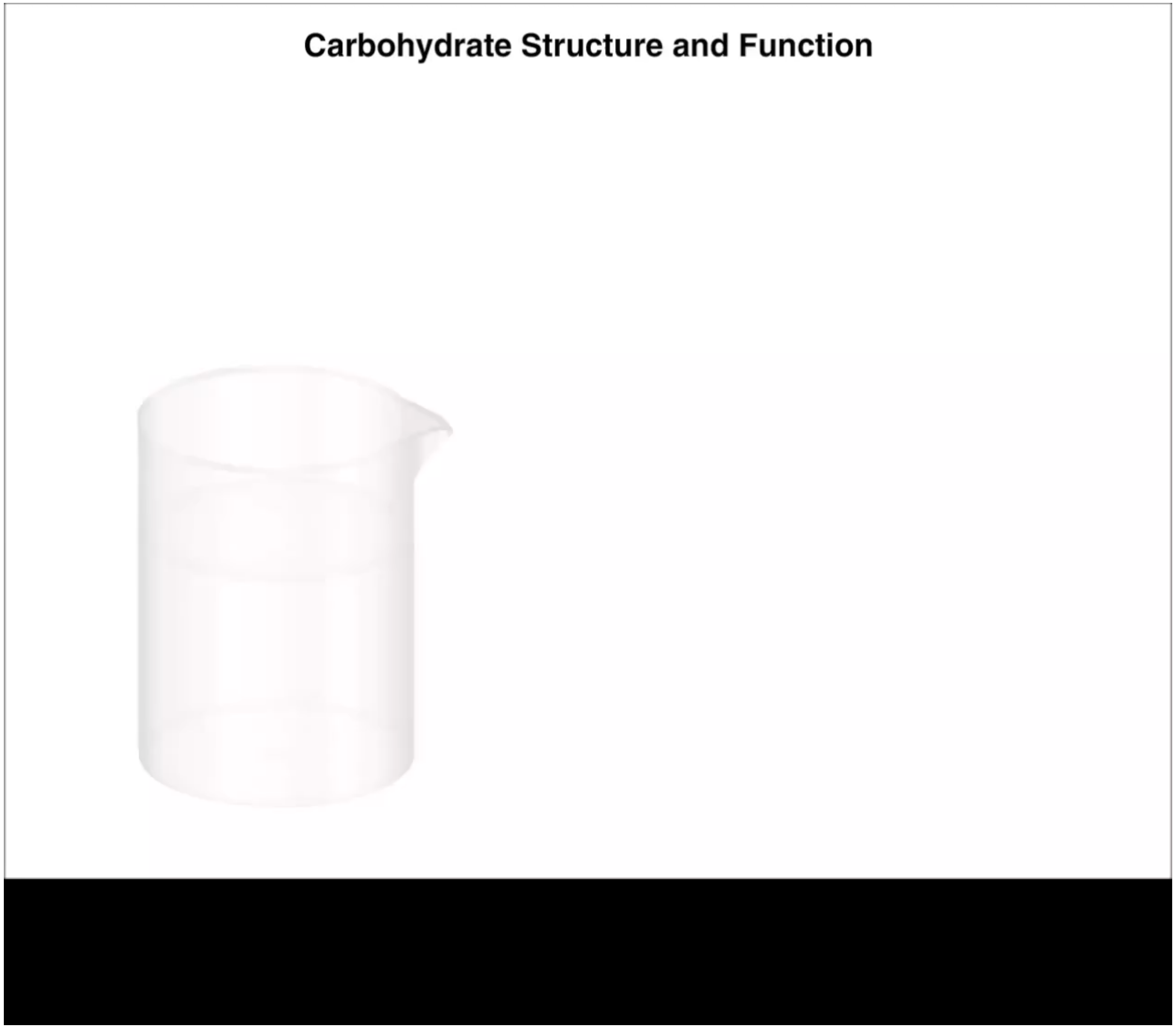


(b) Starch: 1–4 linkage of α glucose monomers



(c) Cellulose: 1–4 linkage of β glucose monomers

Video: Polysaccharides

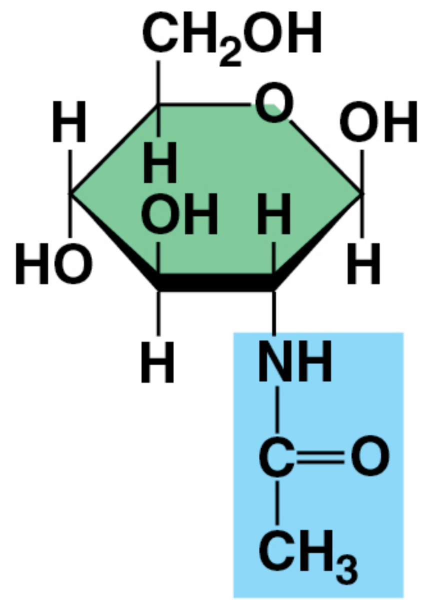


- Starch (α configuration) is largely helical
- Cellulose molecules (β configuration) are straight and unbranched
- Some hydroxyl groups on the monomers of cellulose can hydrogen-bond with hydroxyls of parallel cellulose molecules

- Enzymes that digest starch by hydrolyzing α linkages can’t hydrolyze β linkages in cellulose
- The cellulose in human food passes through the digestive tract as “insoluble fiber”
- Some microbes use enzymes to digest cellulose
- Many herbivores, from cows to termites, have symbiotic relationships with these microbes

- **Chitin**, another structural polysaccharide, is found in the exoskeleton of arthropods
- Chitin also provides structural support for the cell walls of many fungi

Figure 5.8



◀ The structure of the chitin monomer

◀ Chitin, embedded in proteins, forms the exoskeleton of arthropods.

CONCEPT 5.3: Lipids are a diverse group of hydrophobic molecules

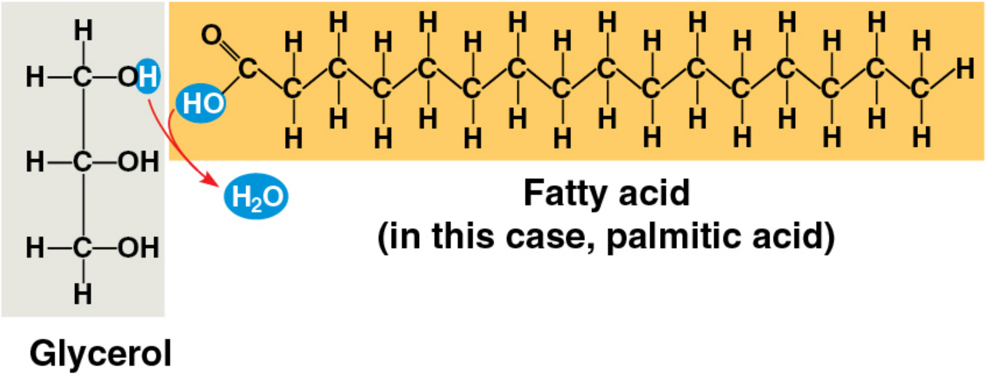
- **Lipids** are the one class of large biological molecules that does not include true polymers
- The unifying feature of lipids is that they mix poorly, if at all, with water
- Lipids consist mostly of hydrocarbon regions
- The most biologically important lipids are fats, phospholipids, and steroids

Fats

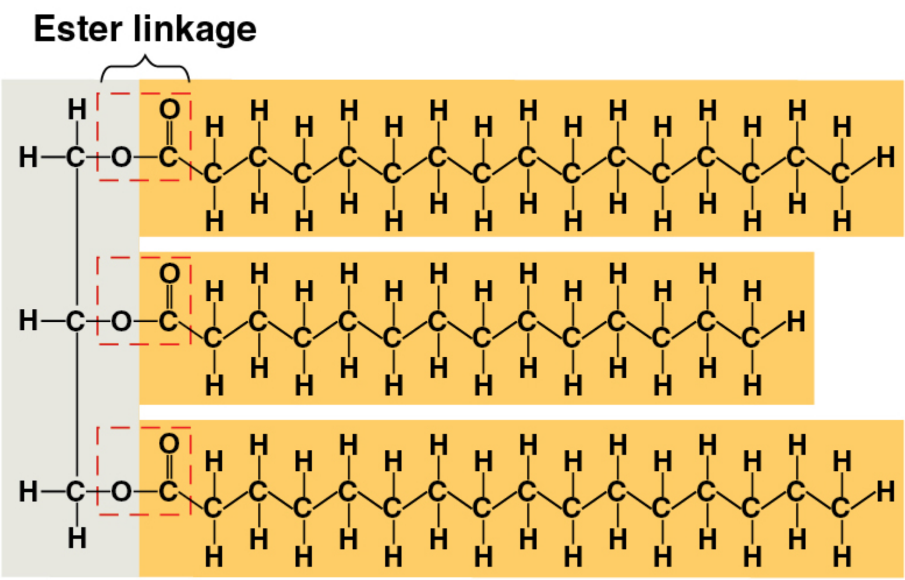
- **Fats** are constructed from two types of smaller molecules: glycerol and fatty acids
- Glycerol is a three-carbon alcohol with a hydroxyl group attached to each carbon
- A **fatty acid** consists of a carboxyl group attached to a long carbon skeleton

- Fats separate from water because water molecules hydrogen-bond to each other and exclude the fats
- In a fat, three fatty acids are joined to glycerol by an ester linkage, creating a **triacylglycerol**, or triglyceride
- The fatty acids in a fat can be all the same or of two or three different kinds

Figure 5.9



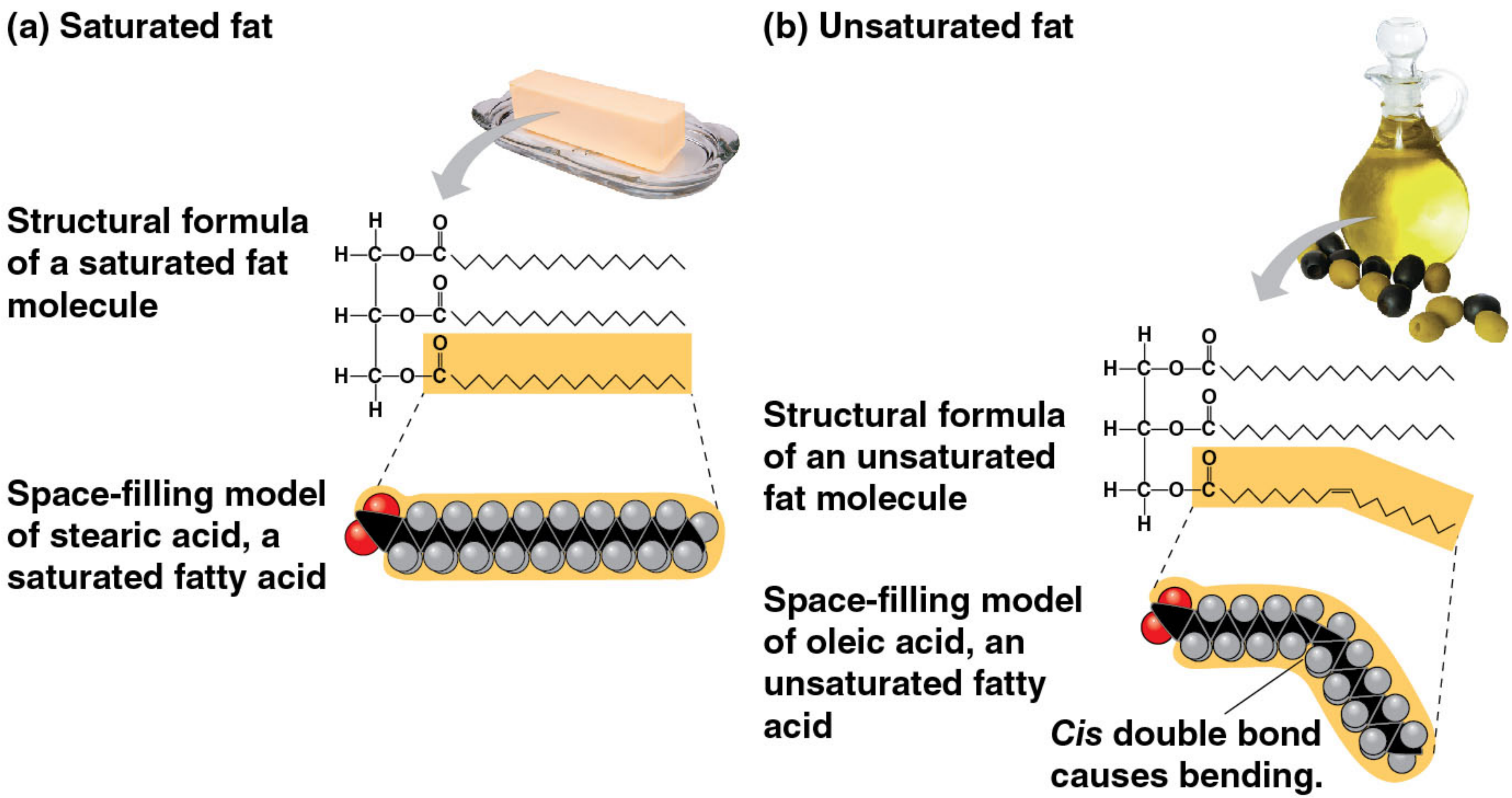
(a) One of three dehydration reactions in the synthesis of a fat. One water molecule is removed for each fatty acid joined to the glycerol.



(b) A fat molecule (triacylglycerol) with three fatty acid units. In this example, two of the fatty acid units are identical.

- Fatty acids vary in length (number of carbons) and in the number and locations of double bonds
- **Saturated fatty acids** have the maximum number of hydrogen atoms possible and no double bonds
- **Unsaturated fatty acids** have one or more double bonds

Figure 5.10


(a) Saturated fat
Structural formula of a saturated fat molecule
Space-filling model of stearic acid, a saturated fatty acid
(b) Unsaturated fat
Structural formula of an unsaturated fat molecule
Space-filling model of oleic acid, an unsaturated fatty acid
Cis double bond causes bending.

- Fats made from saturated fatty acids are called saturated fats and are solid at room temperature
- Most animal fats are saturated
- Fats made from unsaturated fatty acids are called unsaturated fats or oils and are liquid at room temperature
- Plant fats and fish fats are usually unsaturated

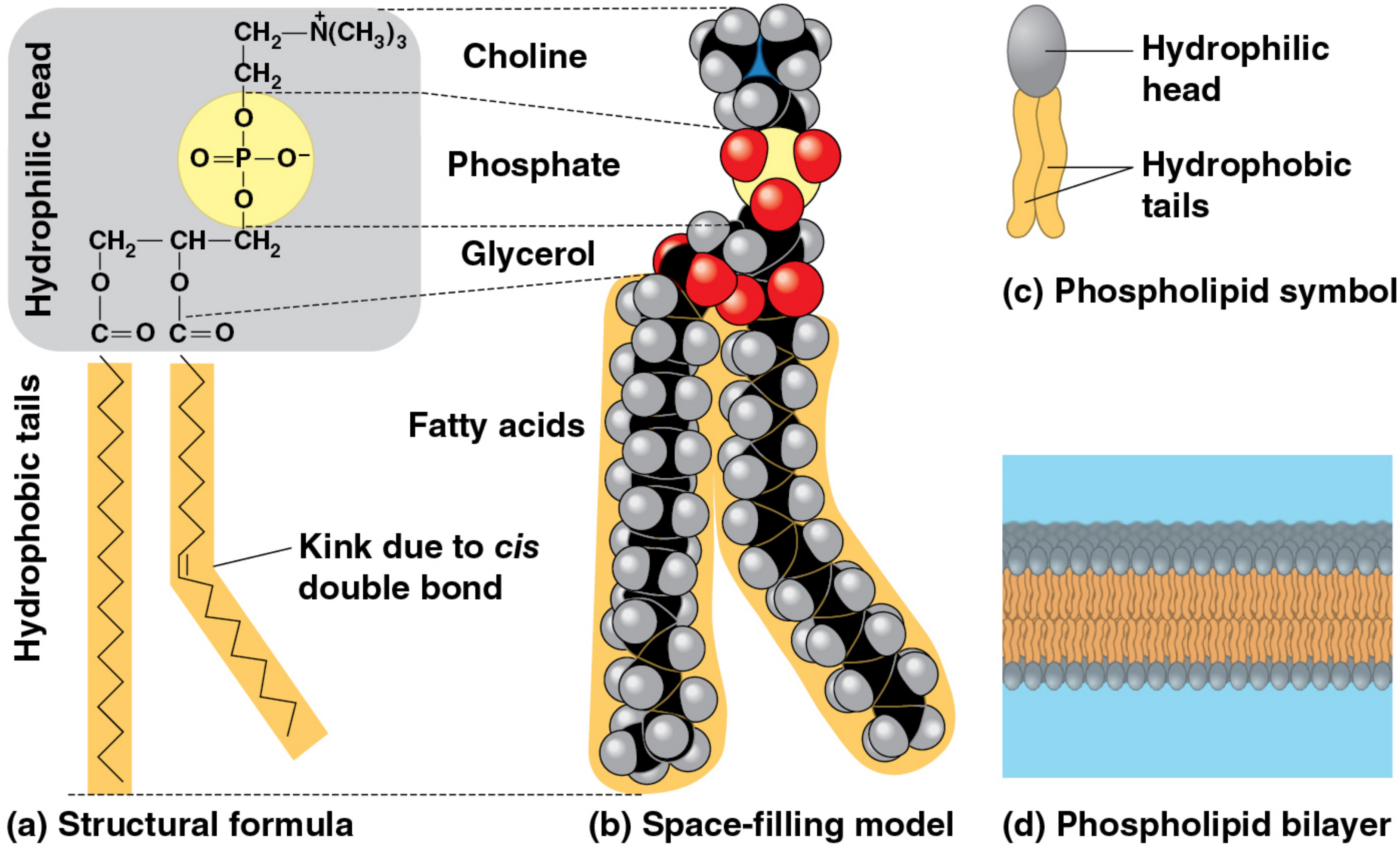
- A diet rich in saturated fats may contribute to cardiovascular disease through plaque deposits
- Hydrogenation is the process of converting unsaturated fats to saturated fats by adding hydrogen
- Hydrogenating vegetable oils also creates unsaturated fats with *trans* double bonds
- These ***trans*** **fats** may contribute more than saturated fats to cardiovascular disease

- The major function of fats is energy storage
- Humans and other mammals store their long-term food reserves in adipose cells
- Adipose tissue also cushions vital organs and insulates the body

Phospholipids

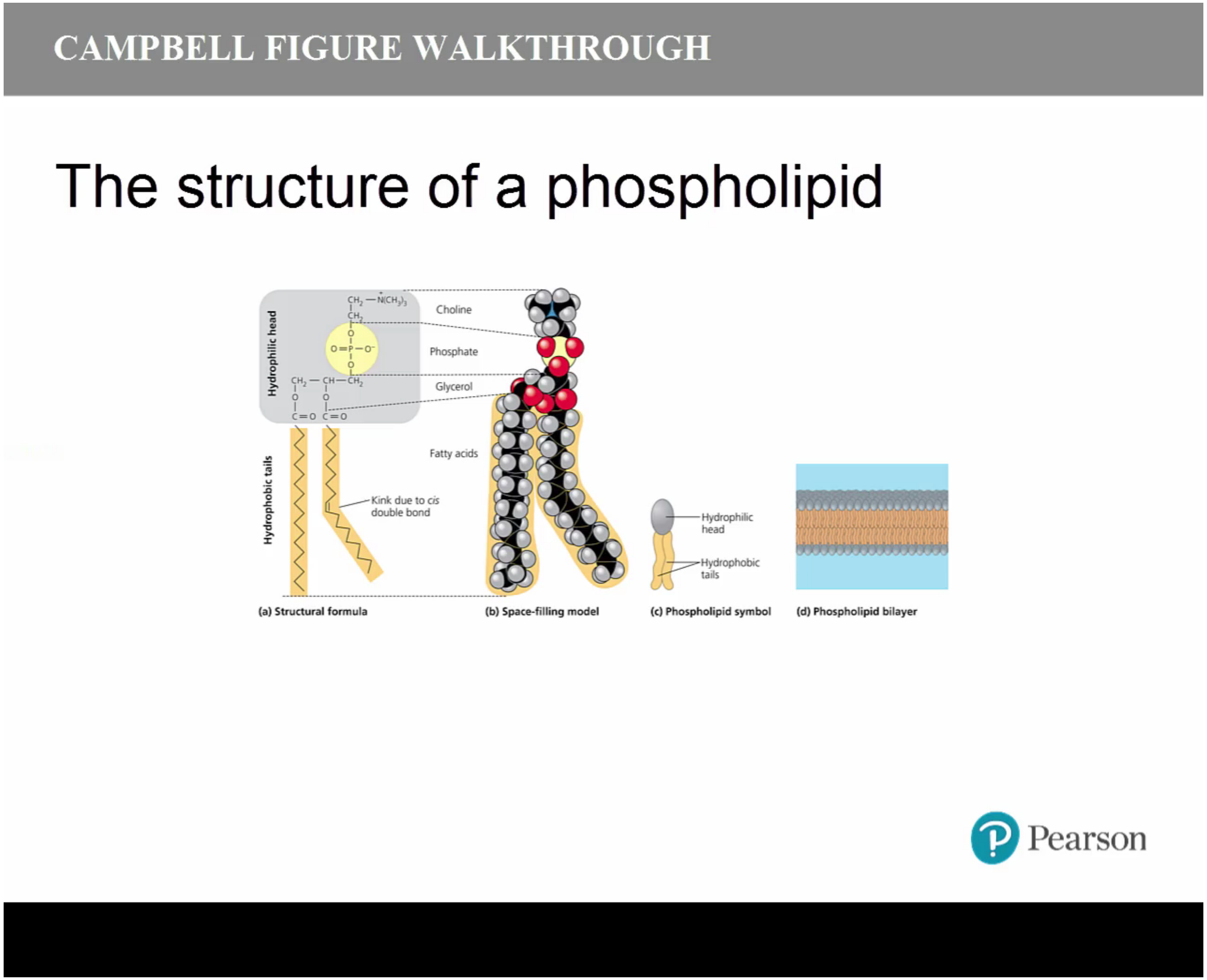
- In a **phospholipid**, two fatty acids and a phosphate group are attached to glycerol
- The two fatty acid tails are hydrophobic, but the phosphate group and its attachments form a hydrophilic head

Figure 5.11

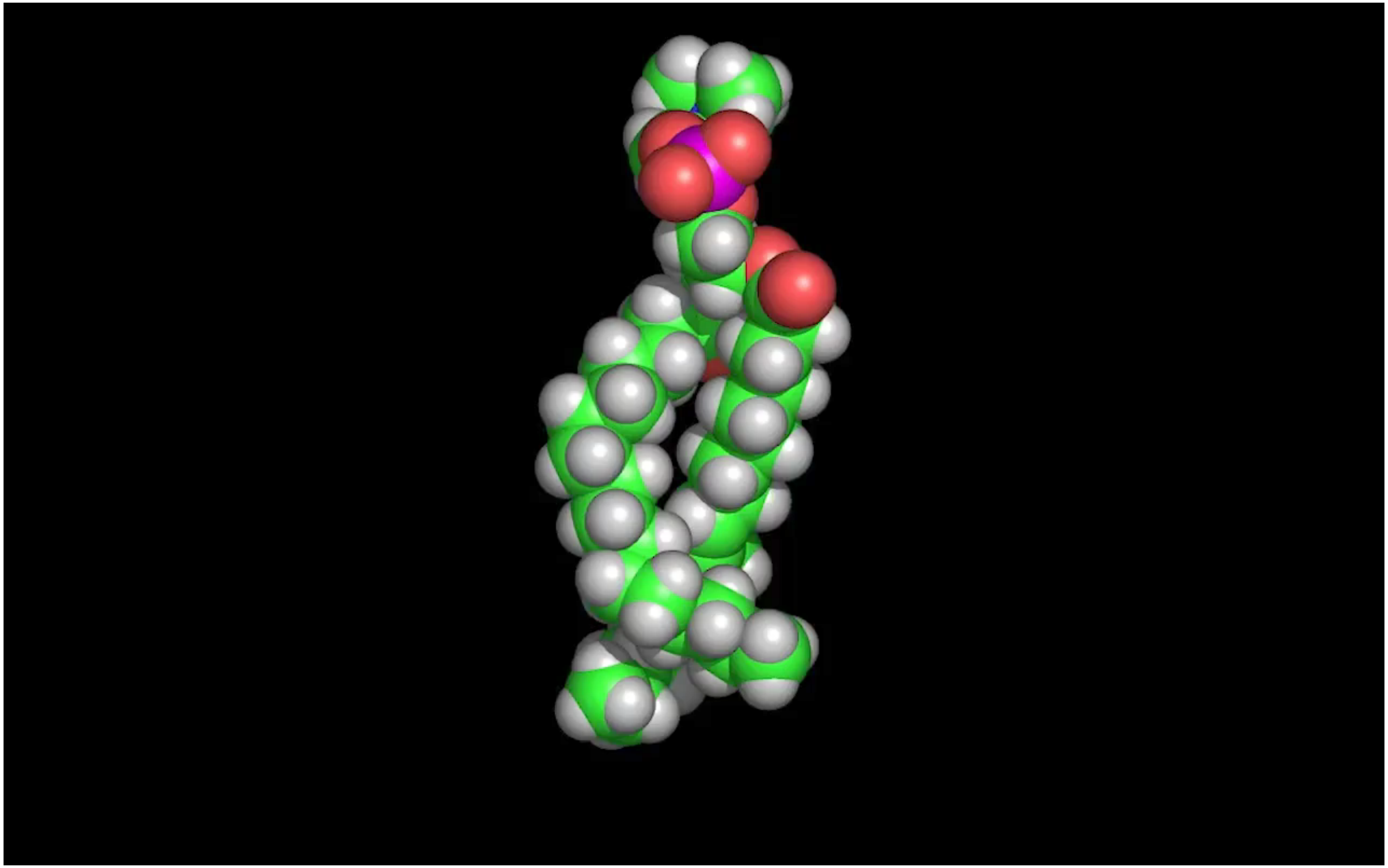


Hydrophilic head
Hydrophobic tails
CH2—N(CH3)3
CH2
O
O=P—O−
O
CH2—CH—CH2
O O
C=O C=O
Choline
Phosphate
Glycerol
Fatty acids
Kink due to cis double bond
(a) Structural formula
(b) Space-filling model
Hydrophilic head
Hydrophobic tails
(c) Phospholipid symbol
(d) Phospholipid bilayer

Video: Phospholipids



Video: Model of Phospholipid

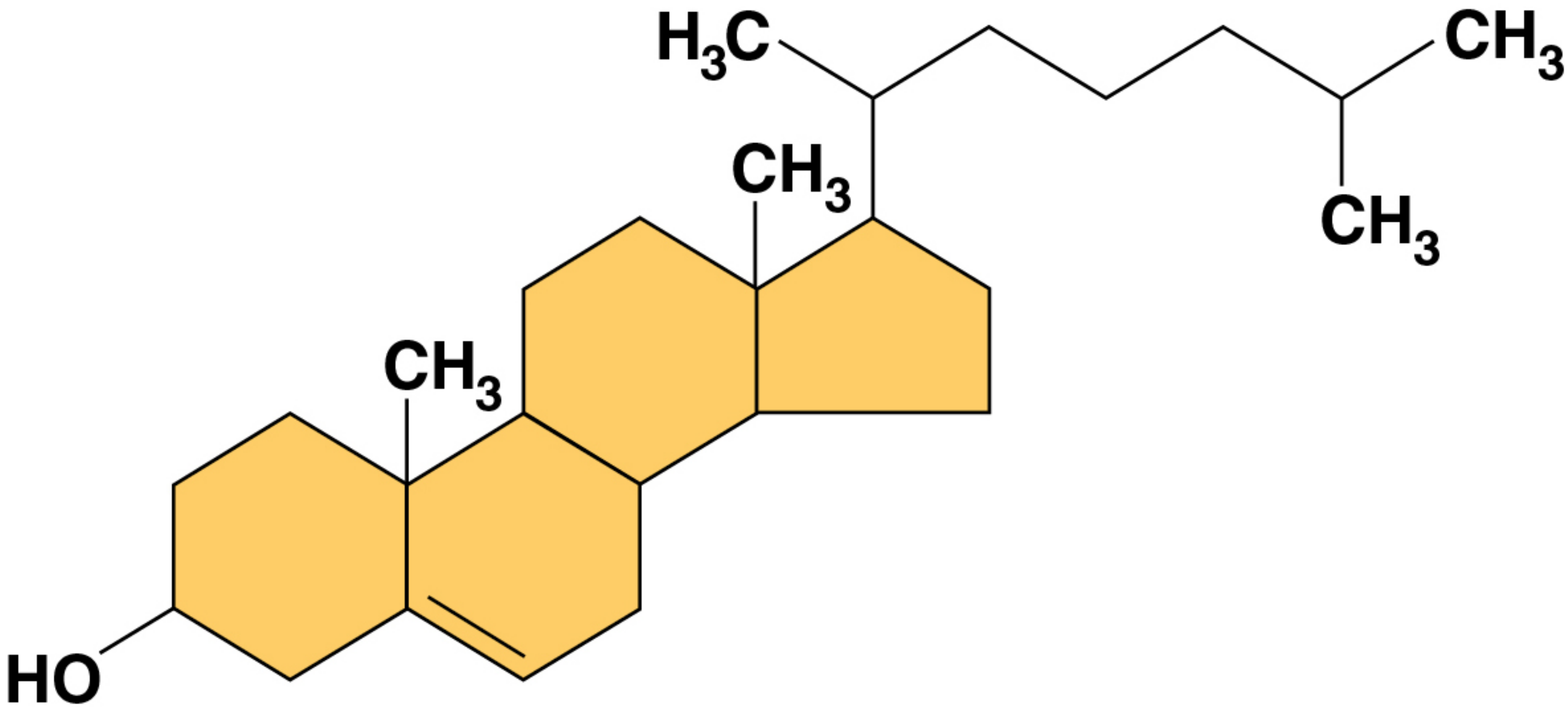


- When phospholipids are added to water, they self-assemble into double-layered sheets called bilayers
- At the surface of a cell, phospholipids are also arranged in a bilayer, with the hydrophobic tails pointing toward the interior
- The phospholipid bilayer forms a boundary between the cell and its external environment

Steroids

- **Steroids** are lipids characterized by a carbon skeleton consisting of four fused rings
- **Cholesterol**, a type of steroid, is a component in animal cell membranes and a precursor from which other steroids are synthesized
- A high level of cholesterol in the blood may contribute to cardiovascular disease

Figure 5.12



CONCEPT 5.4: Proteins include a diversity of structures, resulting in a wide range of functions

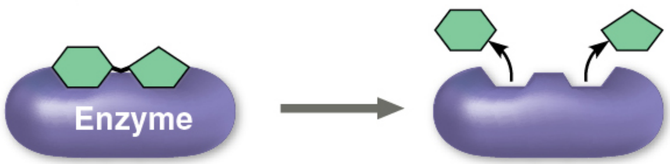
- Proteins account for more than 50% of the dry mass of most cells
- Some proteins speed up chemical reactions
- Other protein functions include defense, storage, transport, cellular communication, movement, and structural support

Figure 5.13a

Enzymatic proteins

Function: Selective acceleration of chemical reactions

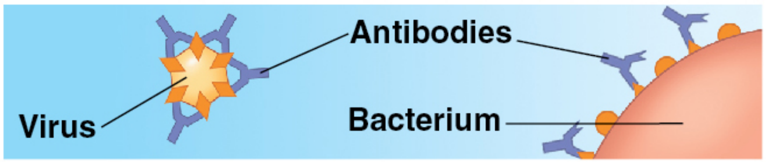
Example: Digestive enzymes catalyze the hydrolysis of bonds in food molecules.



Defensive proteins

Function: Protection against disease

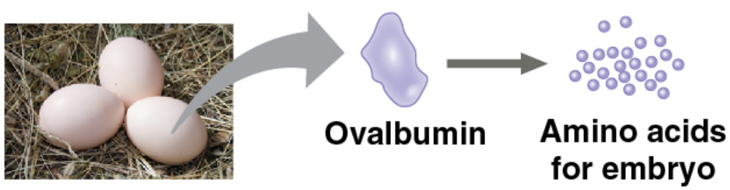
Example: Antibodies inactivate and help destroy viruses and bacteria.



Storage proteins

Function: Storage of amino acids

Examples: Casein, the protein of milk, is the major source of amino acids for baby mammals. Plants have storage proteins in their seeds. Ovalbumin is the protein of egg white, used as an amino acid source for the developing embryo.



Transport proteins

Function: Transport of substances

Examples: Hemoglobin, the iron-containing protein of vertebrate blood, transports oxygen from the lungs to other parts of the body. Other proteins transport molecules across membranes, as shown here.

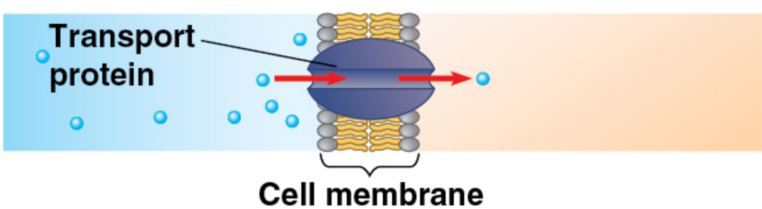
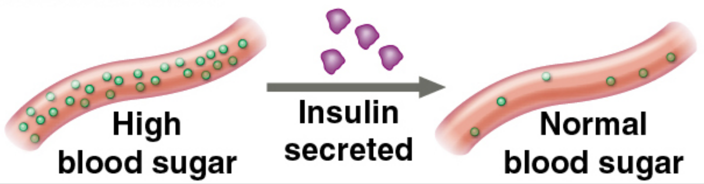


Figure 5.13b

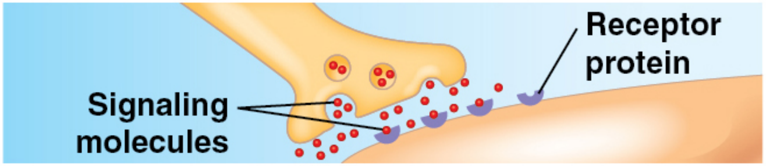
Hormonal proteins

Function: Coordination of an organism's activities
Example: Insulin, a hormone secreted by the pancreas, causes other tissues to take up glucose, thus regulating blood sugar concentration.



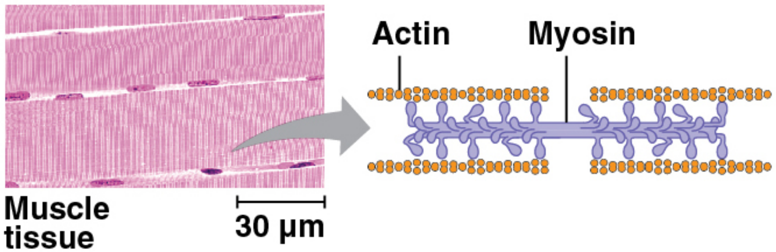
Receptor proteins

Function: Response of cell to chemical stimuli
Example: Receptors built into the membrane of a nerve cell detect signaling molecules released by other nerve cells.



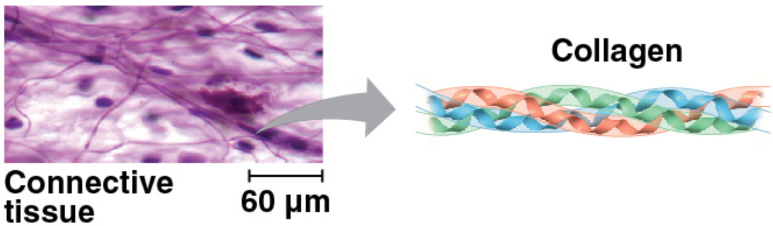
Contractile and motor proteins

Function: Movement
Examples: Motor proteins are responsible for the undulations of cilia and flagella. Actin and myosin proteins are responsible for the contraction of muscles.



Structural proteins

Function: Support
Examples: Keratin is the protein of hair, horns, feathers, and other skin appendages. Insects and spiders use silk fibers to make their cocoons and webs, respectively. Collagen and elastin proteins provide a fibrous framework in animal connective tissues.



Video: Protein Functions

Protein Functions

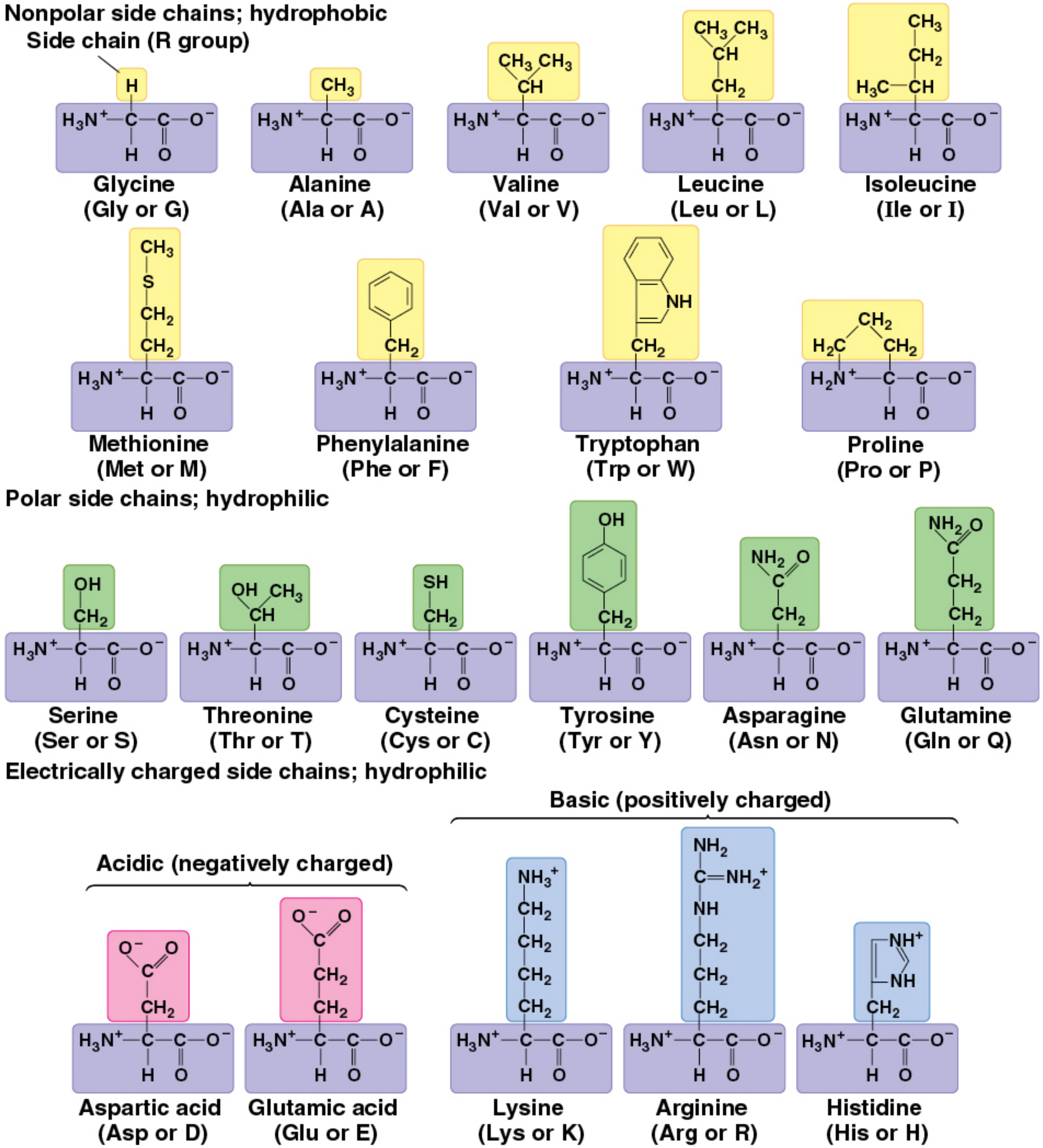
- Enzymes are proteins that act as **catalysts** to speed up chemical reactions
- Enzymes can perform their functions repeatedly, functioning as workhorses that carry out the processes of life

- Proteins are all constructed from the same set of 20 amino acids
- **Polypeptides** are unbranched polymers built from these amino acids
- The bond between amino acids is a peptide bond
- A **protein** is a biologically functional molecule that consists of one or more polypeptides

Amino Acids (Monomers)

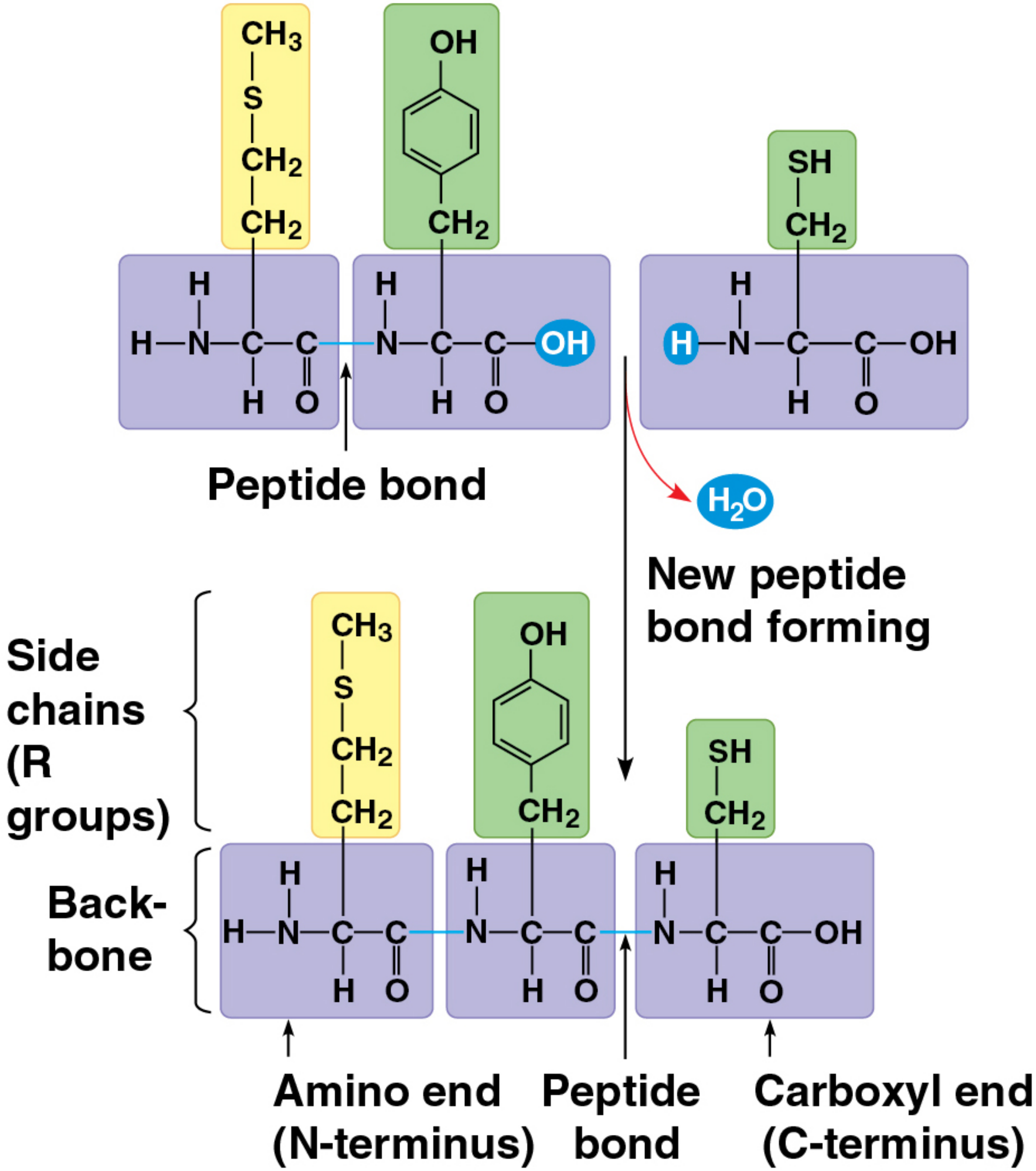
- **Amino acids** are organic molecules with amino and carboxyl groups
- Amino acids differ in their properties due to differing side chains, called R groups

Figure 5.14


Nonpolar side chains; hydrophobic
Side chain (R group)
Glycine
(Gly or G)
Alanine
(Ala or A)
Valine
(Val or V)
Leucine
(Leu or L)
Isoleucine
(Ile or I)
Methionine
(Met or M)
Phenylalanine
(Phe or F)
Tryptophan
(Trp or W)
Proline
(Pro or P)
Polar side chains; hydrophilic
Serine
(Ser or S)
Threonine
(Thr or T)
Cysteine
(Cys or C)
Tyrosine
(Tyr or Y)
Asparagine
(Asn or N)
Glutamine
(Gln or Q)
Electrically charged side chains; hydrophilic
Basic (positively charged)
Acidic (negatively charged)
Aspartic acid
(Asp or D)
Glutamic acid
(Glu or E)
Lysine
(Lys or K)
Arginine
(Arg or R)
Histidine
(His or H)

Polypeptides (Amino Acid Polymers)

- Amino acids are linked by covalent bonds called **peptide bonds**
- A polypeptide is a polymer of amino acids
- Polypeptides range in length from a few to more than 1,000 monomers
- Each polypeptide has a unique linear sequence of amino acids, with a carboxyl end (C-terminus) and an amino end (N-terminus)


CH3
S
CH2
CH2
OH
CH2
SH
CH2
H
H—N—C—C
H O
H
N—C—C—OH
H O
H
H—N—C—C—OH
H O
Peptide bond
H2O
New peptide bond forming
Side chains (R groups)
Back-bone
Amino end (N-terminus)
Peptide bond
Carboxyl end (C-terminus)

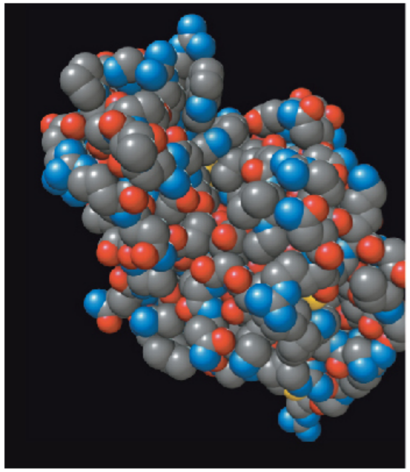
Animation: Proteins

Protein Structure and Function

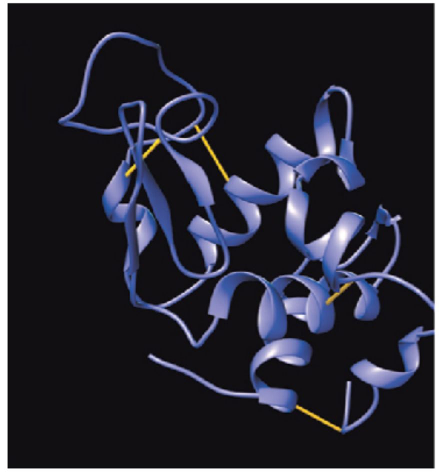
- The specific activities of proteins result from their intricate three-dimensional architecture
- A functional protein consists of one or more polypeptides precisely twisted, folded, and coiled into a unique shape

Figure 5.16

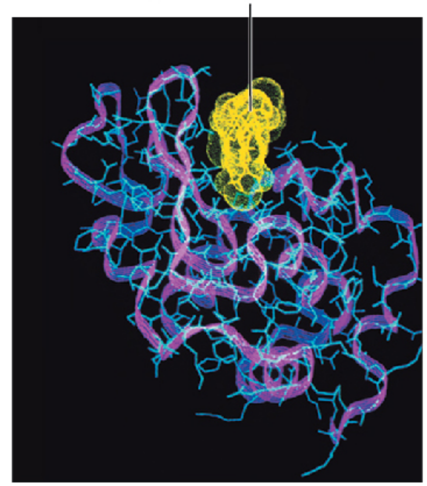
Structural Models



Space-filling model

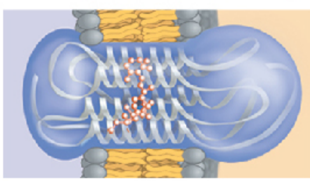


Ribbon model

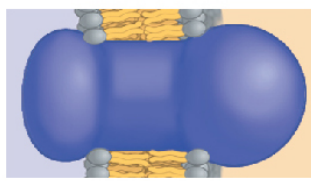


Wireframe model (blue)

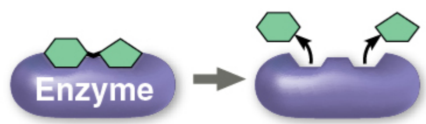
Simplified Diagrams



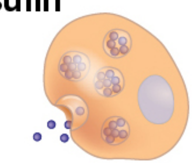
A transparent shape is drawn around the contours of a ribbon model of the protein rhodopsin, showing the shape of the molecule as well as some internal details.



When structural details are not needed, a solid shape can be used.



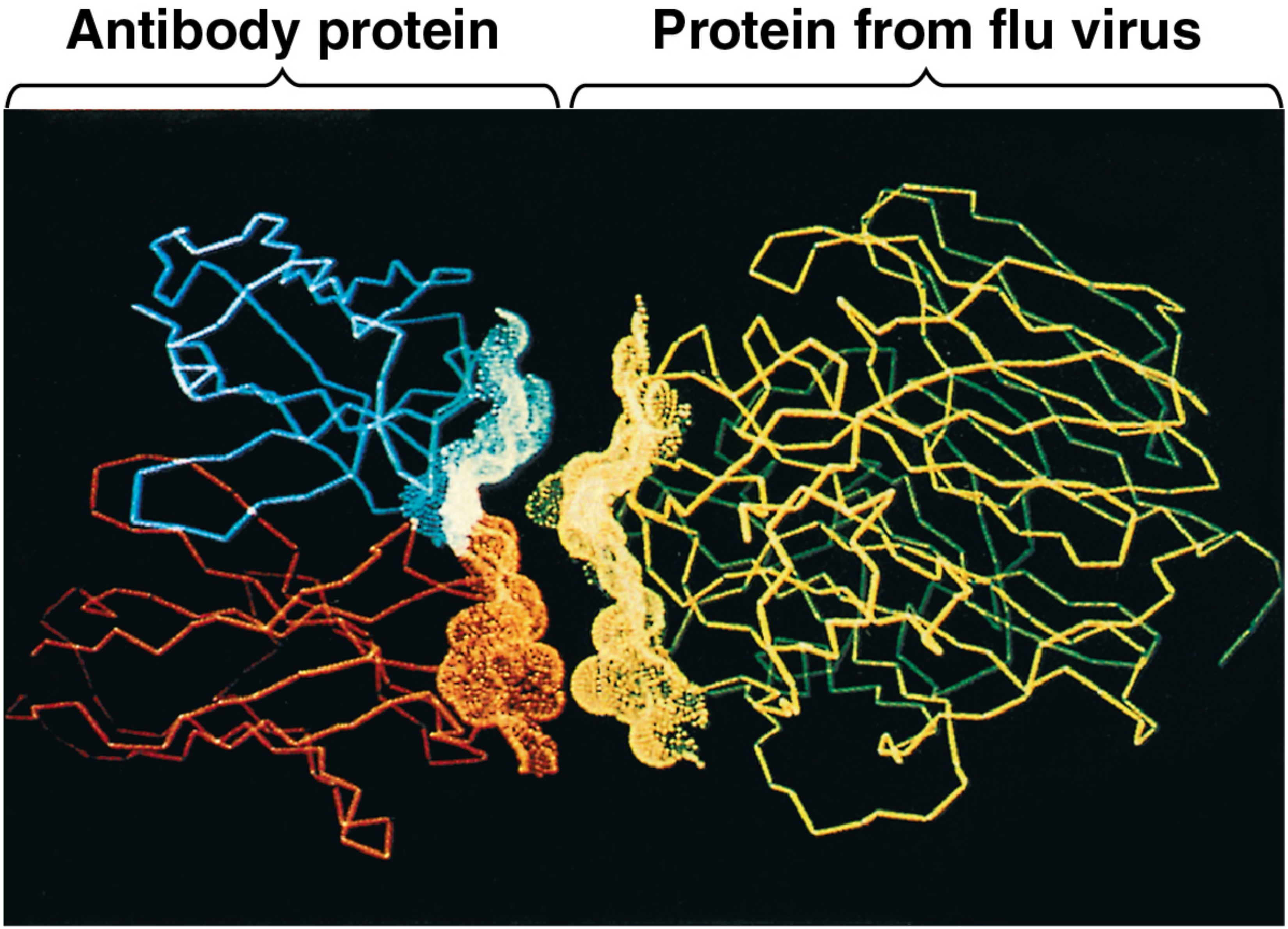
A simple shape is used here to represent a generic enzyme because the diagram focuses on enzyme action in general.



Sometimes a protein is represented simply as a dot, as shown here for insulin.

- The sequence of amino acids determines a protein’s three-dimensional structure
- A protein’s structure determines how it works
- The function of a protein usually depends on its ability to recognize and bind to some other molecule

Figure 5.17



Four Levels of Protein Structure

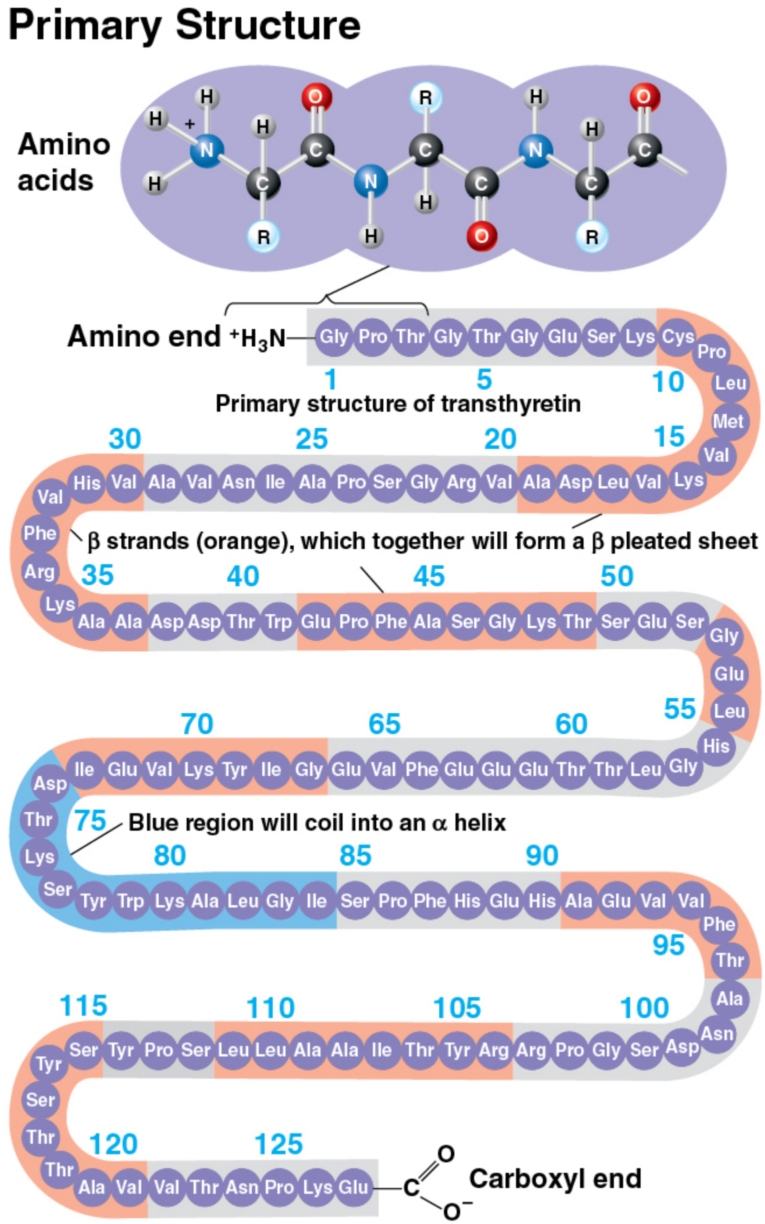
- The primary structure of a protein is its unique sequence of amino acids
- Secondary structure, found in most proteins, consists of coils and folds in the polypeptide chain
- Tertiary structure is determined by interactions among various side chains (R groups)
- Quaternary structure results when a protein consists of multiple polypeptide chains

Video: Protein Structure

Protein Structure

- The **primary structure** of a protein is its sequence of amino acids
- Primary structure is like the order of letters in a long word
- Primary structure is determined by inherited genetic information

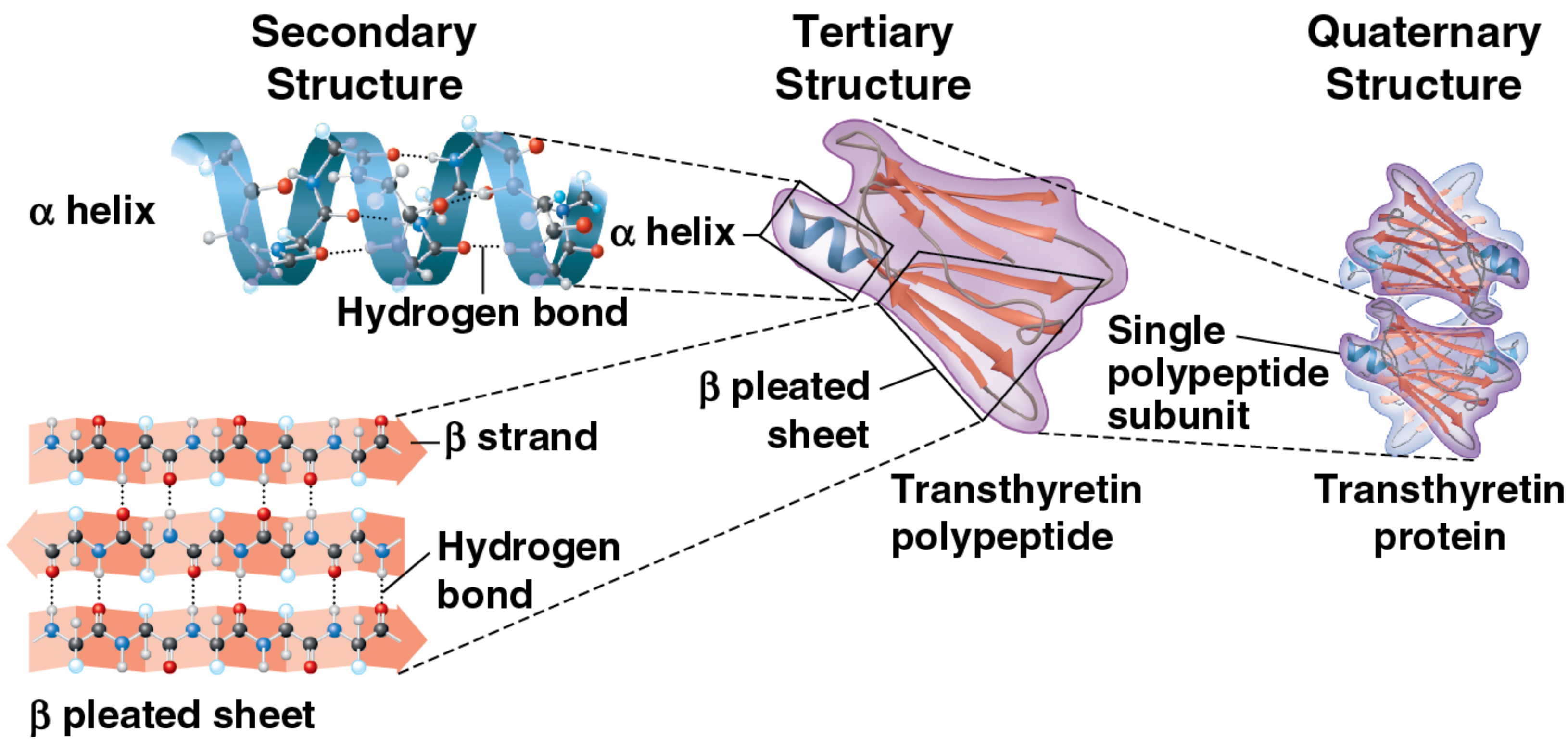
Figure 5.18a



Primary Structure
Amino acids
H H H O R H H O H H C C N C C N C C N H H O R R R
Amino end ^+H_3N
Gly Pro Thr Gly Thr Gly Glu Ser Lys Cys Pro Leu Met Val Lys
1 5 10 15
Primary structure of transthyretin
30 25 20
Val Leu Asp Ala Val Arg Gly Ser Pro Ala Ile Asn Val Ala Val His Val Phe Arg Lys
β strands (orange), which together will form a β pleated sheet
35 40 45 50
Ala Ala Asp Asp Thr Trp Glu Pro Phe Ala Ser Gly Lys Thr Ser Glu Ser Gly Glu Leu His Gly
55
70 65 60
Leu Thr Thr Glu Glu Glu Phe Val Glu Gly Ile Tyr Lys Val Glu Ile Asp Thr Lys Ser
75 Blue region will coil into an α helix
80 85 90
Tyr Trp Lys Ala Leu Gly Ile Ser Pro Phe His Glu His Ala Glu Val Val Phe Thr Ala Asn
95
115 110 105 100
Asp Ser Gly Pro Arg Arg Tyr Thr Ile Ala Ala Leu Leu Ser Pro Tyr Ser Tyr Ser Thr Thr
120 125
Ala Val Val Thr Asn Pro Lys Glu
C O O^-
Carboxyl end

- The coils and folds of **secondary structure** result from hydrogen bonds between repeating constituents of the polypeptide backbone
- Typical secondary structures are a coil called an **α helix** and a folded structure called a **β pleated sheet**

Figure 5.18b

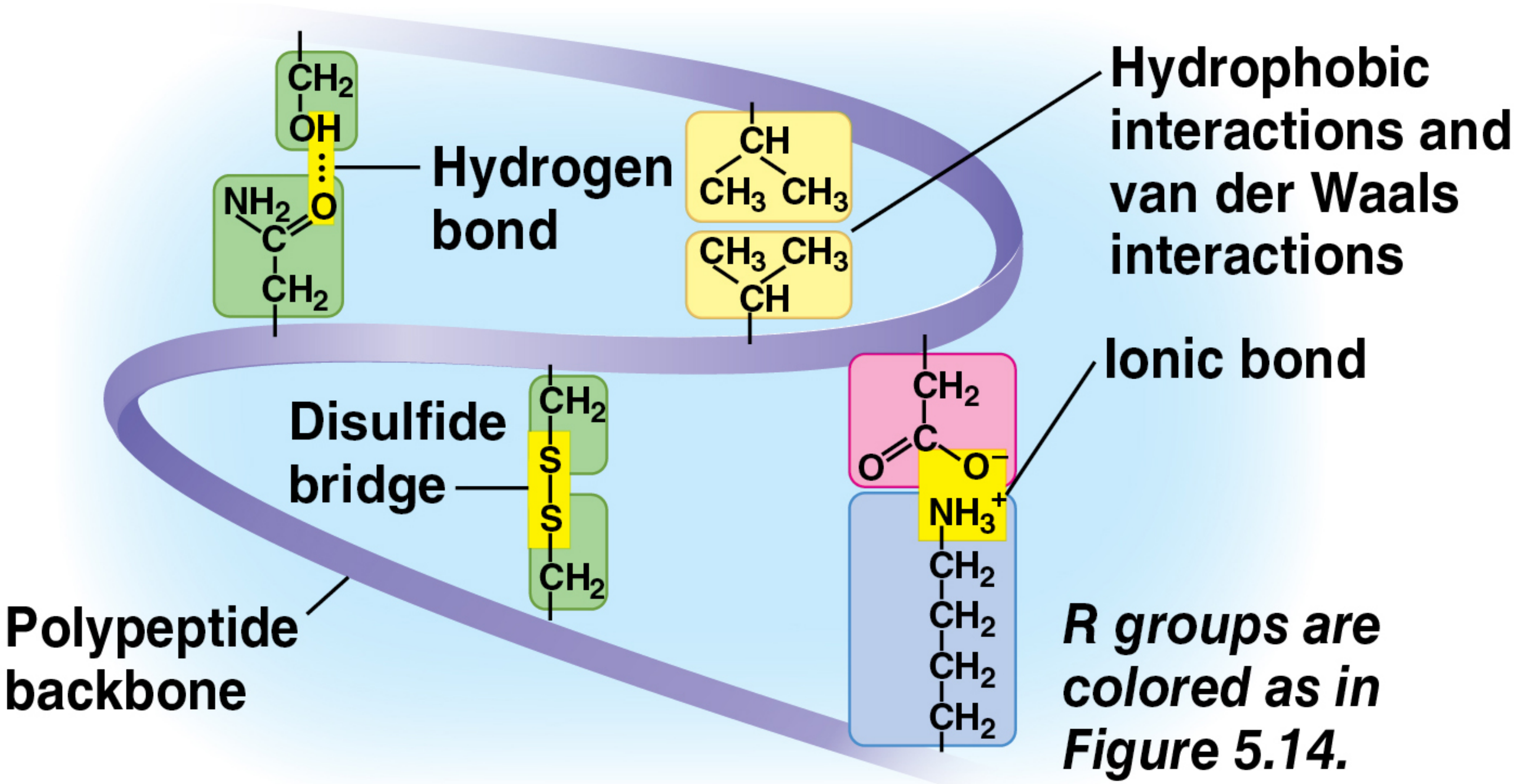


Secondary Structure
Tertiary Structure
Quaternary Structure
α helix
α helix
Hydrogen bond
β pleated sheet
β strand
Hydrogen bond
β pleated sheet
Transthyretin polypeptide
Single polypeptide subunit
Transthyretin protein

Figure 5.18c

- **Tertiary structure**, the overall shape of a polypeptide, results from interactions between R groups, rather than interactions between backbone constituents
- These interactions include hydrogen bonds, ionic bonds, **hydrophobic interactions**, and van der Waals interactions
- Strong covalent bonds called **disulfide bridges** may reinforce the protein's structure

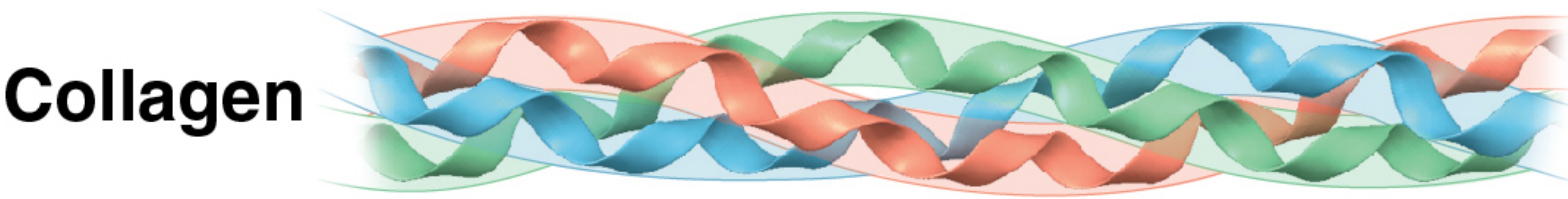
Figure 5.18d

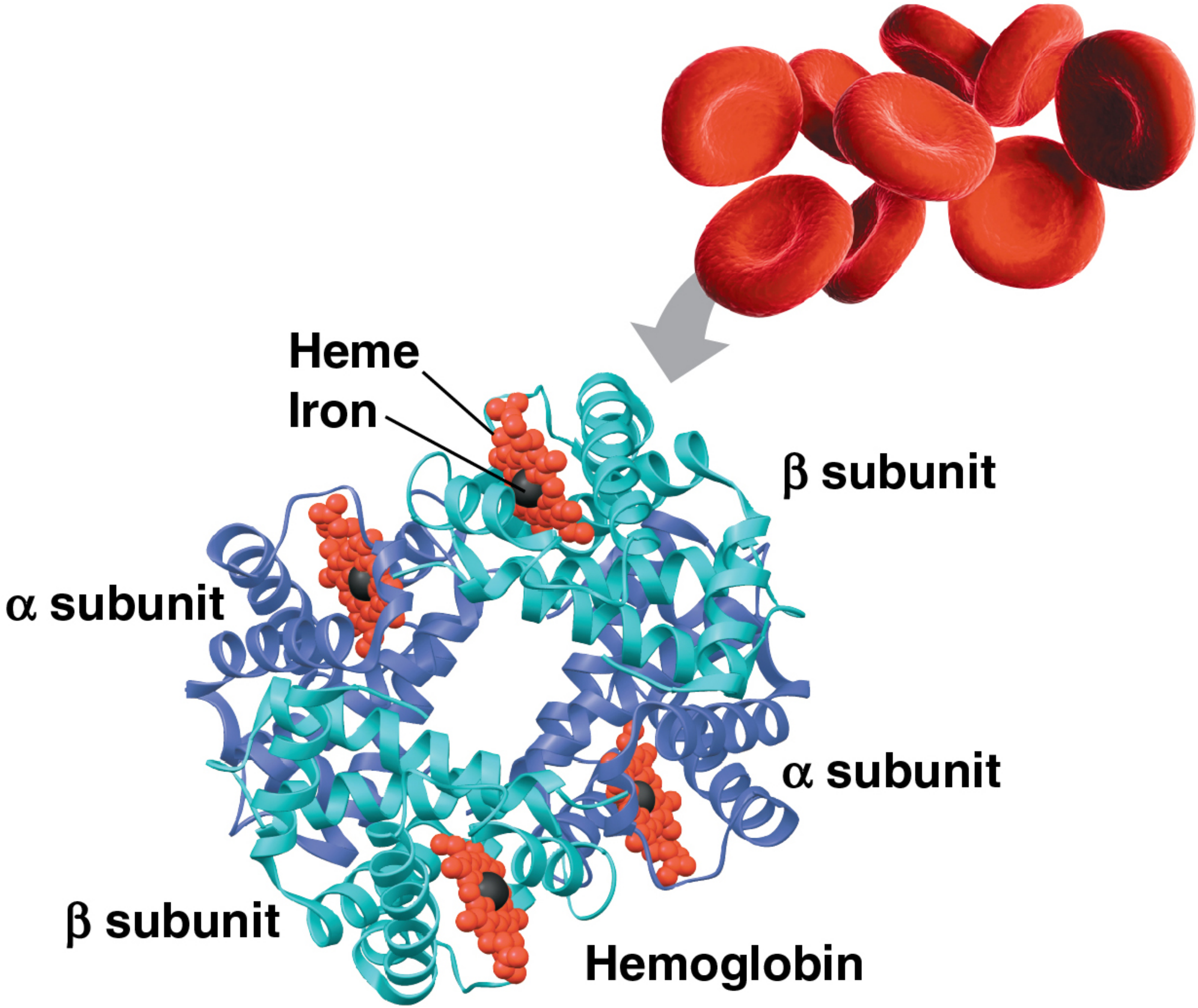


CH2
OH
NH2
O
C
CH2
Hydrogen bond
CH
CH3 CH3
CH3 CH3
CH
Hydrophobic interactions and van der Waals interactions
Disulfide bridge
CH2
S
S
CH2
CH2
C
O
O–
NH3+
CH2
CH2
CH2
CH2
Ionic bond
Polypeptide backbone
R groups are colored as in Figure 5.14.

- **Quaternary structure** results when two or more polypeptide chains form one macromolecule
- Collagen is a fibrous protein consisting of three polypeptides coiled like a rope
- Hemoglobin is a globular protein consisting of four polypeptides: two α and two β subunits

Figure 5.18e

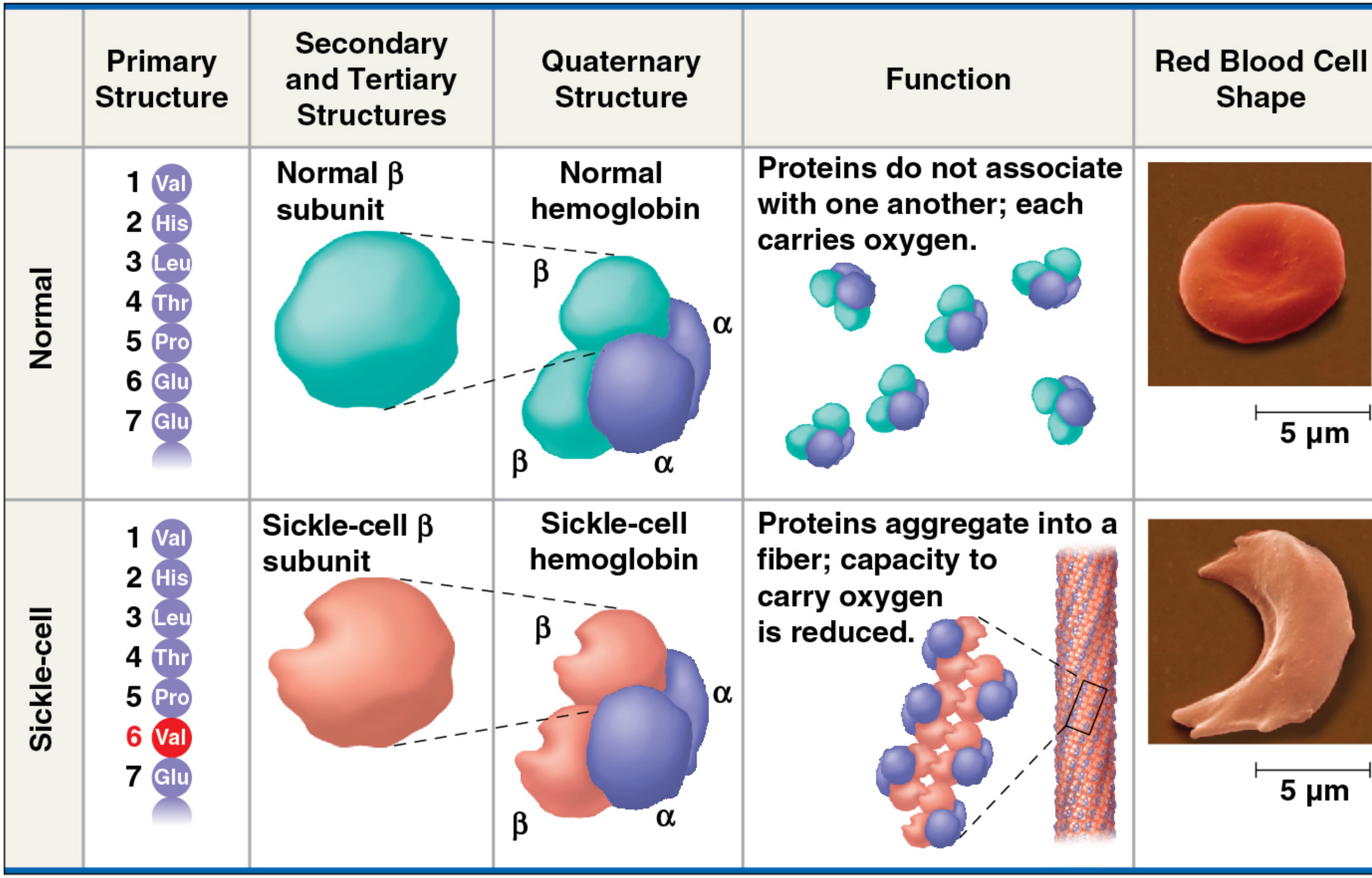



Heme
Iron
β subunit
α subunit
α subunit
β subunit
Hemoglobin

Sickle-Cell Disease: A Change in Primary Structure

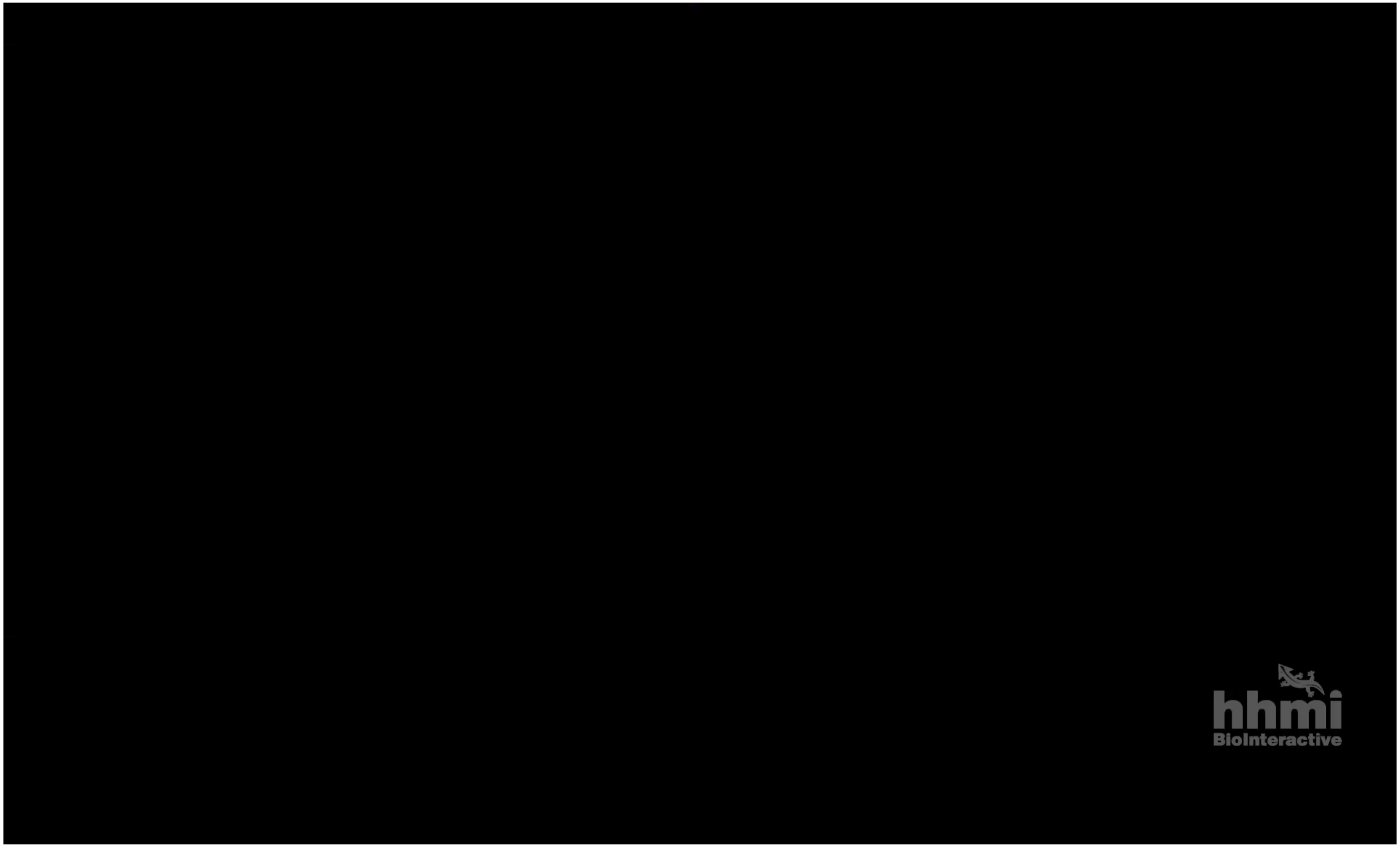
- A slight change in primary structure can affect a protein’s structure and ability to function
- **Sickle-cell disease**, an inherited blood disorder, results from a single amino acid substitution in the protein hemoglobin
- The abnormal hemoglobin molecules cause the red blood cells to aggregate into chains and to deform into a sickle shape

Figure 5.19



Primary Structure
Secondary and Tertiary Structures
Quaternary Structure
Function
Red Blood Cell Shape
Normal
1 Val
2 His
3 Leu
4 Thr
5 Pro
6 Glu
7 Glu
Normal β subunit
Normal hemoglobin
β
α
β
α
Proteins do not associate with one another; each carries oxygen.
5 µm
Sickle-cell
1 Val
2 His
3 Leu
4 Thr
5 Pro
6 Val
7 Glu
Sickle-cell β subunit
Sickle-cell hemoglobin
β
α
β
α
Proteins aggregate into a fiber; capacity to carry oxygen is reduced.
5 µm

Video: Sickle Cell Disease

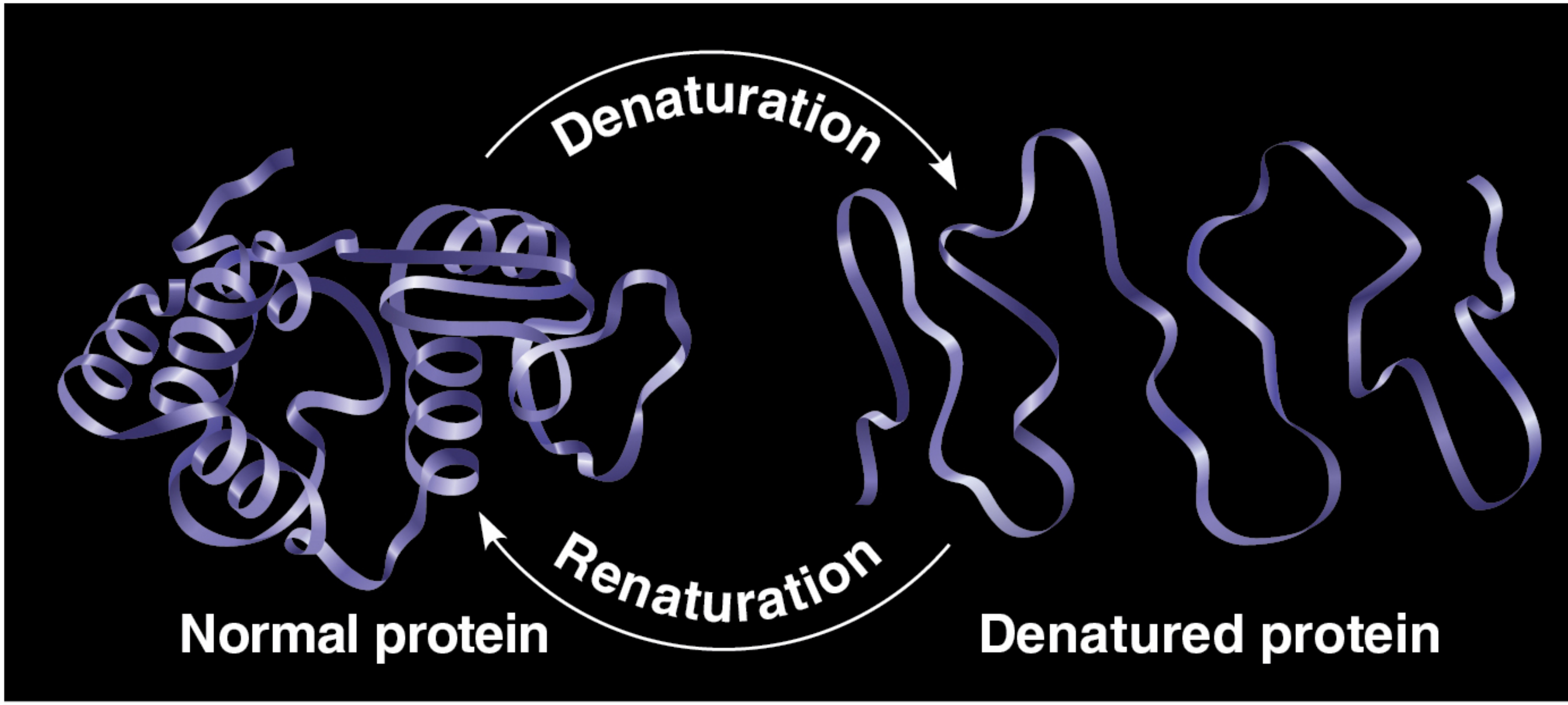


What Determines Protein Structure?

- In addition to primary structure, physical and chemical conditions can affect structure
- Alterations in pH, salt concentration, temperature, or other environmental factors can cause a protein to unravel
- This loss of a protein's native structure is called **denaturation**

- A denatured protein is biologically inactive
- Extremely high fevers can be fatal: proteins in the blood tend to denature at very high body temperatures
- Denaturation can sometimes be reversed when the denaturing agent is removed
- This is not always possible, however

Figure 5.20


Denaturation
Renaturation
Normal protein
Denatured protein

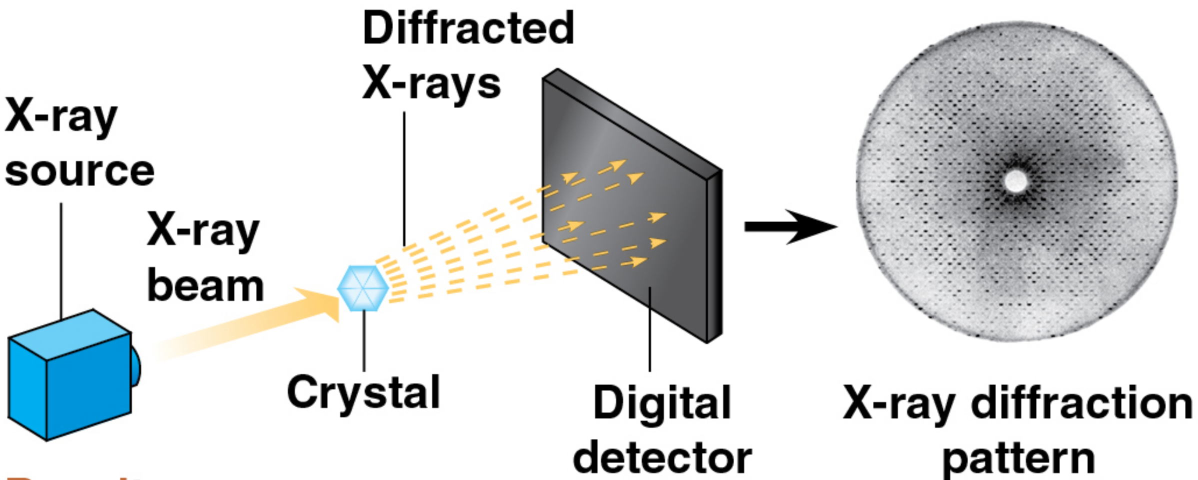
Protein Folding in the Cell

- It is hard to predict a protein's structure from its primary structure
- Most proteins probably go through several stages on their way to a stable structure
- Diseases such as Alzheimer's, Parkinson's, and mad cow disease are associated with misfolded proteins

- Scientists use **X-ray crystallography** to determine a protein's structure
- Another method is nuclear magnetic resonance (NMR) spectroscopy, which does not require protein crystallization
- Bioinformatics is another approach to prediction of protein structure from amino acid sequences

Figure 5.21

Technique



Results

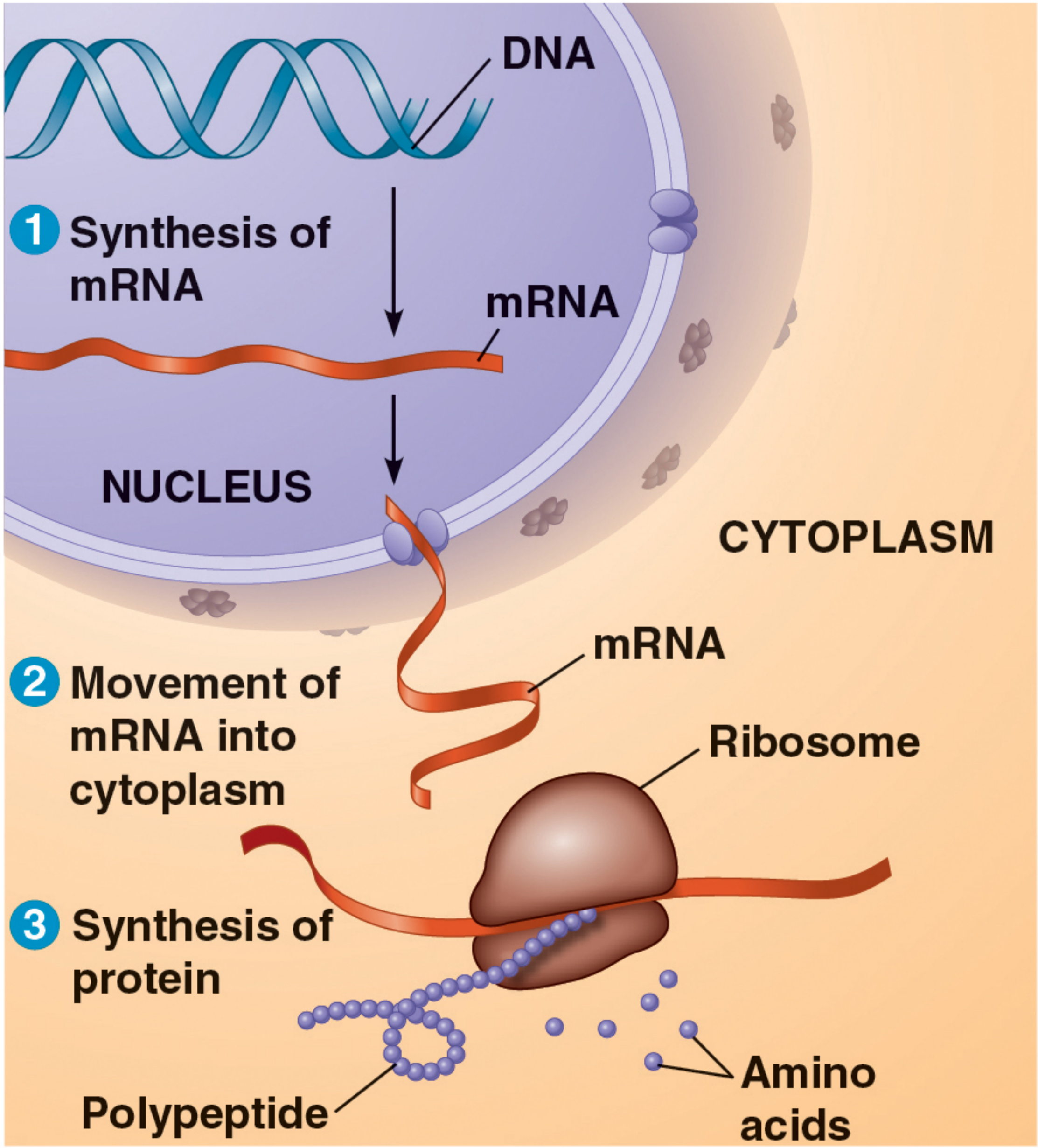
CONCEPT 5.5: Nucleic acids store, transmit, and help express hereditary information

- The amino acid sequence of a polypeptide is programmed by a unit of inheritance called a **gene**
- Genes consist of DNA, a **nucleic acid** made of monomers called nucleotides

The Roles of Nucleic Acids

- There are two types of nucleic acids
  - **Deoxyribonucleic acid (DNA)**
  - **Ribonucleic acid (RNA)**
- DNA provides directions for its own replication
- DNA directs synthesis of messenger RNA (mRNA) and, through mRNA, controls protein synthesis
- This process is called **gene expression**

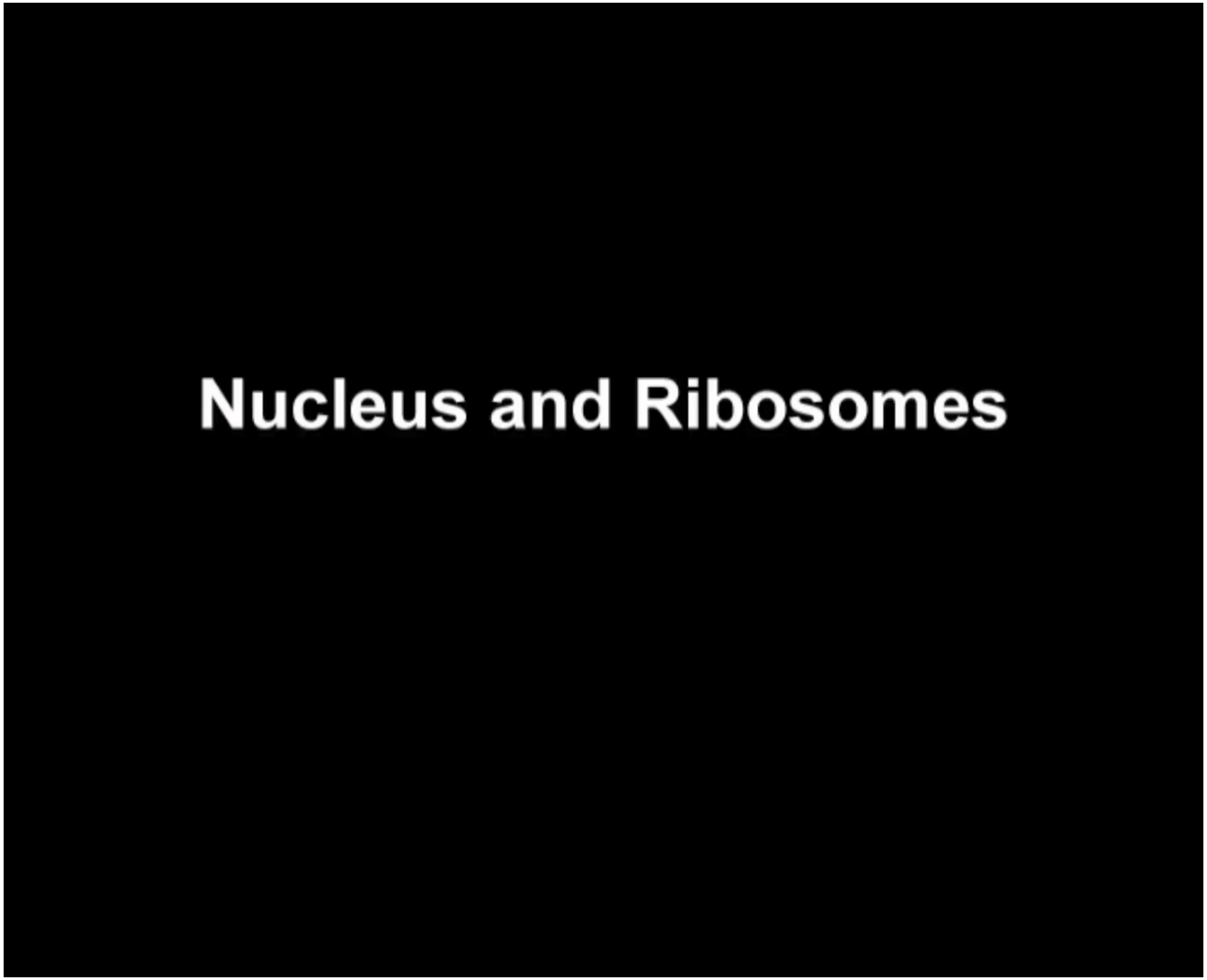
Figure 5.22



DNA
1 Synthesis of mRNA
mRNA
NUCLEUS
CYTOPLASM
mRNA
2 Movement of mRNA into cytoplasm
Ribosome
3 Synthesis of protein
Polypeptide
Amino acids

- Each gene along a DNA molecule directs synthesis of a messenger RNA (mRNA)
- The mRNA molecule interacts with the cell’s protein-synthesizing machinery to direct production of a polypeptide
- The flow of genetic information can be summarized as DNA → RNA → protein

Video: Gene Expression

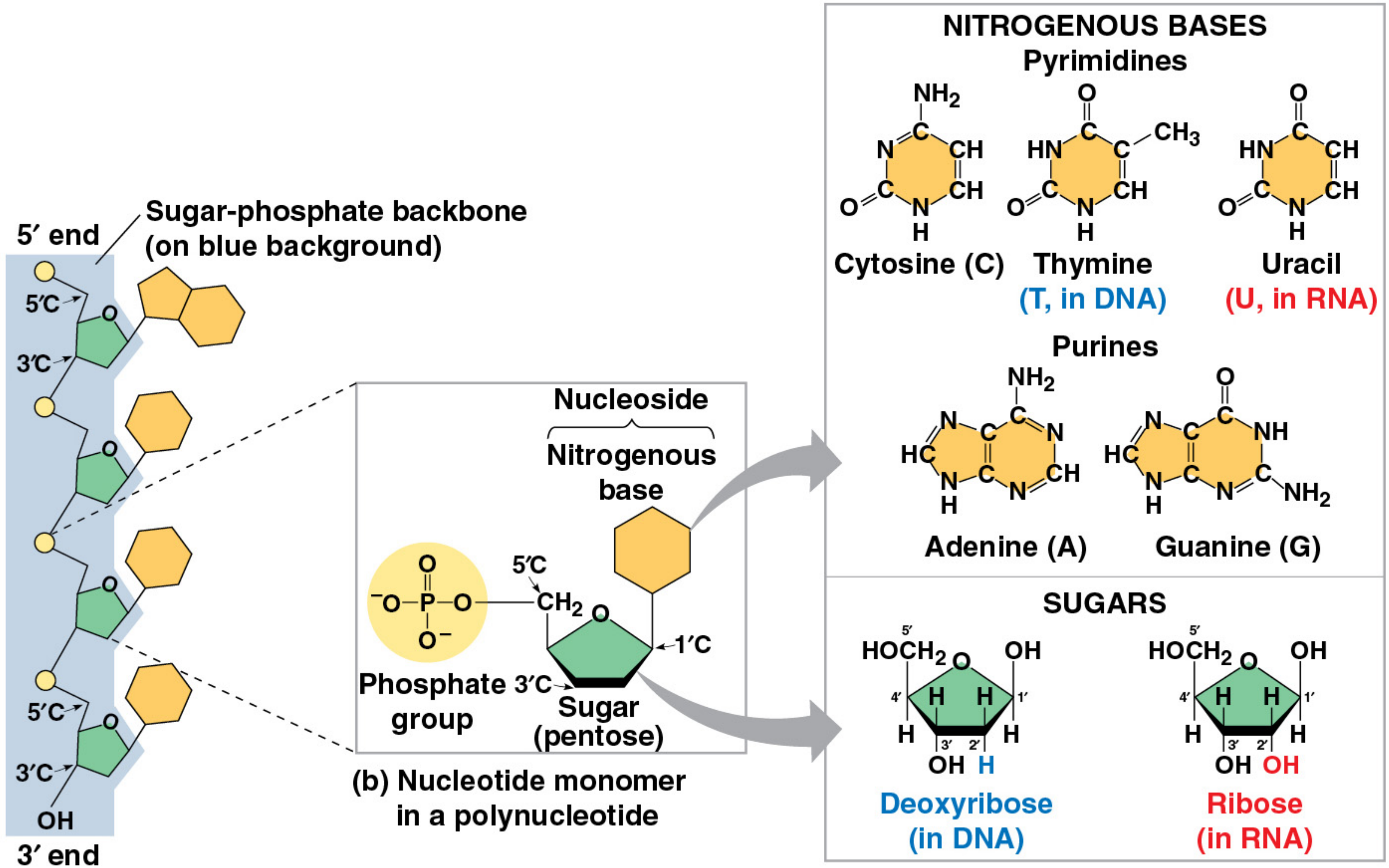


The Components of Nucleic Acids

- Nucleic acids are polymers called **polynucleotides**
- Each polynucleotide is made of monomers called **nucleotides**
- Each nucleotide consists of a nitrogenous base, a pentose sugar, and one or more phosphate groups
- The portion of a nucleotide without the phosphate group is called a nucleoside

- Nucleoside = nitrogenous base + sugar
- There are two families of nitrogenous bases
  - **Pyrimidines** (cytosine, thymine, and uracil) have a single six-membered ring
  - **Purines** (adenine and guanine) have a six-membered ring fused to a five-membered ring
- In DNA, the sugar is **deoxyribose**; in RNA, the sugar is **ribose**
- Nucleotide = nucleoside + phosphate group

Figure 5.23

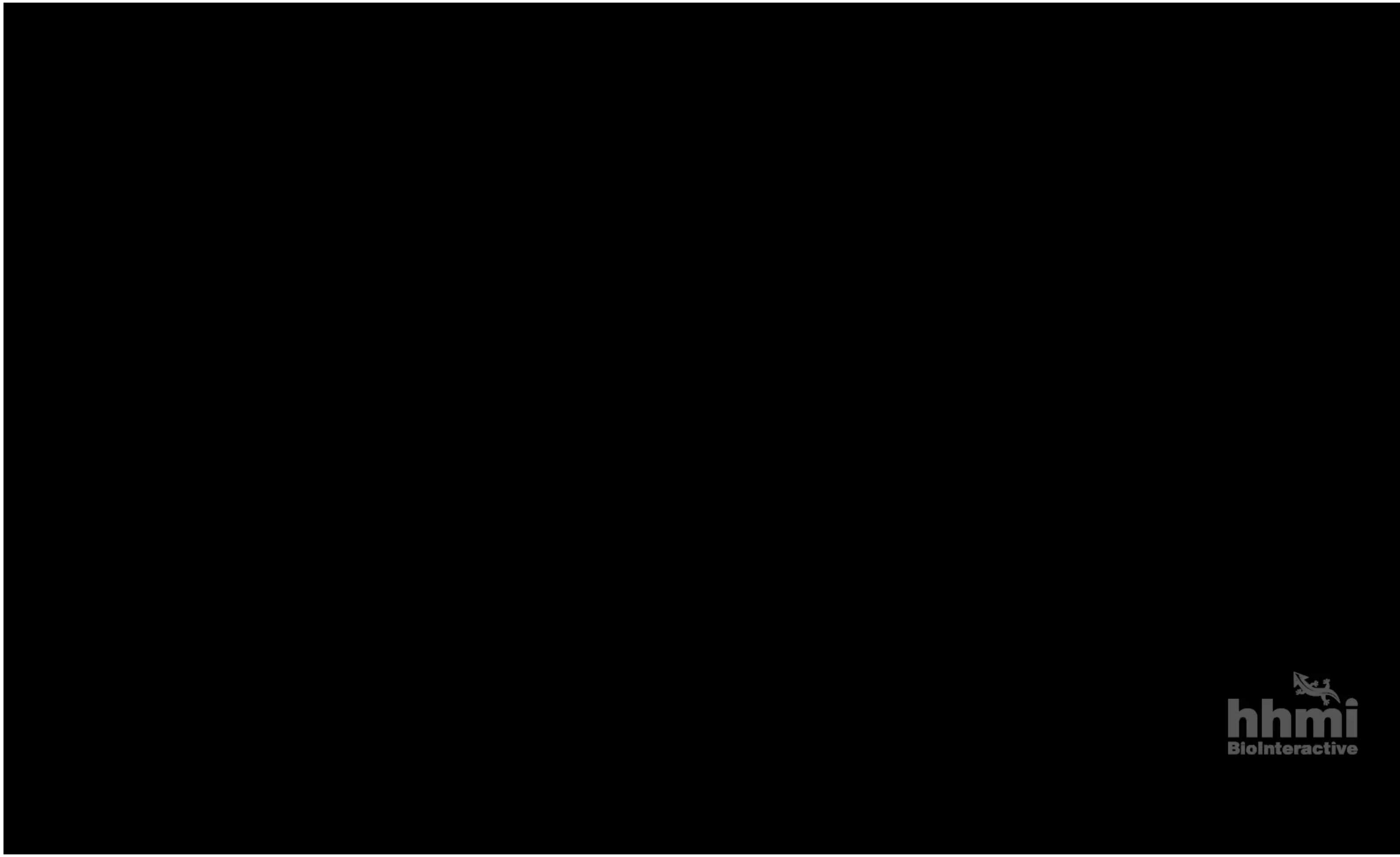


(a) Polynucleotide, or nucleic acid

(b) Nucleotide monomer in a polynucleotide

(c) Nucleoside components

Video: DNA Structure



Nucleotide Polymers

- Nucleotides are linked together by a phosphodiester linkage to build a polynucleotide
- A phosphodiester linkage consists of a phosphate group that links the sugars of two nucleotides
- These links create a backbone of sugar-phosphate units with nitrogenous bases as appendages
- The sequence of bases along a DNA or mRNA polymer is unique for each gene

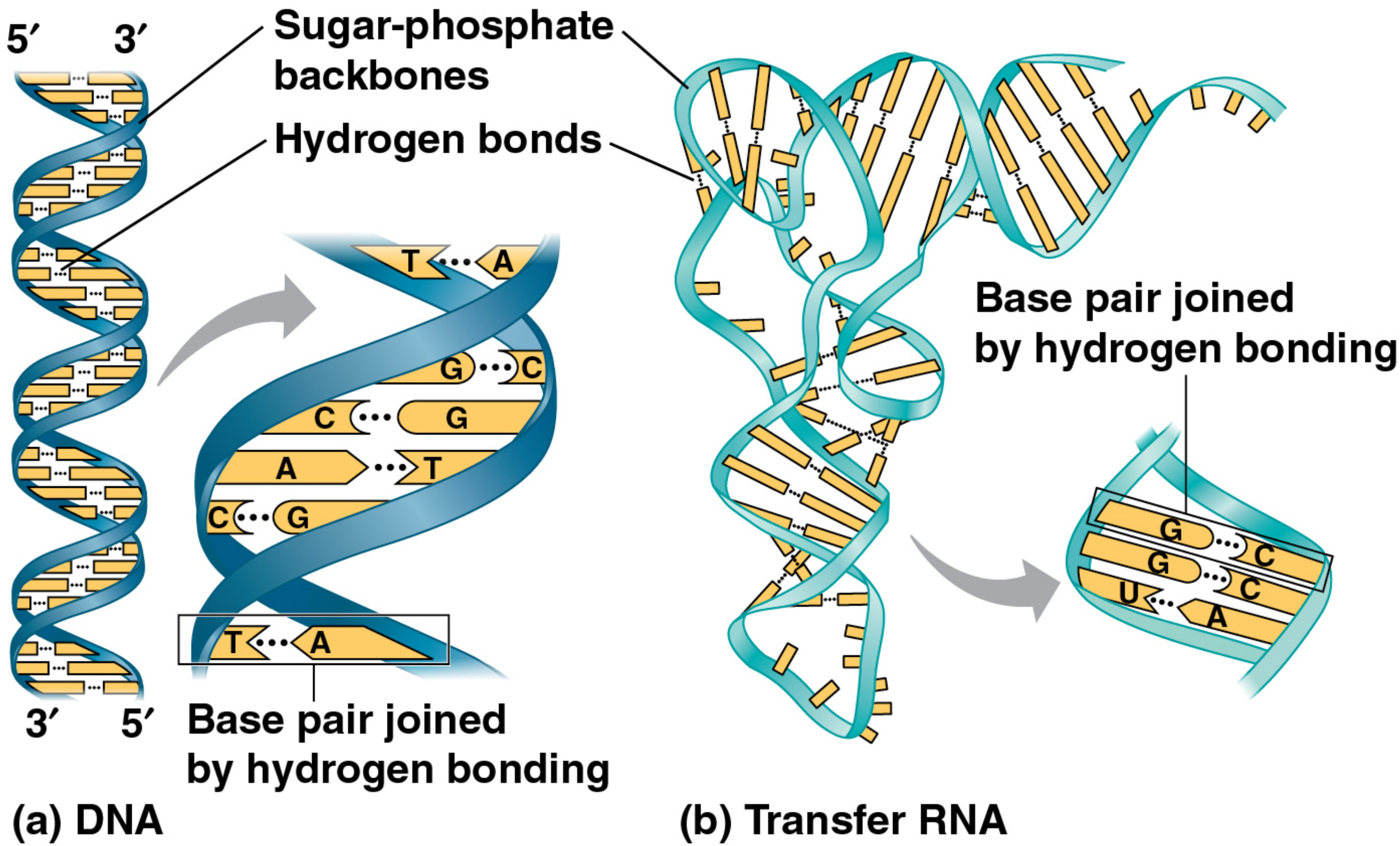
The Structures of DNA and RNA Molecules

- DNA molecules have two polynucleotides spiraling around an imaginary axis, forming a **double helix**
- The backbones run in opposite $5' \rightarrow 3'$ directions from each other, an arrangement referred to as **antiparallel**
- One DNA molecule includes many genes

- Only certain bases in DNA pair up and form hydrogen bonds: adenine (A) always with thymine (T), and guanine (G) always with cytosine (C)
- This is called complementary base pairing
- This feature of DNA structure makes it possible to generate two identical copies of each DNA molecule in a cell preparing to divide

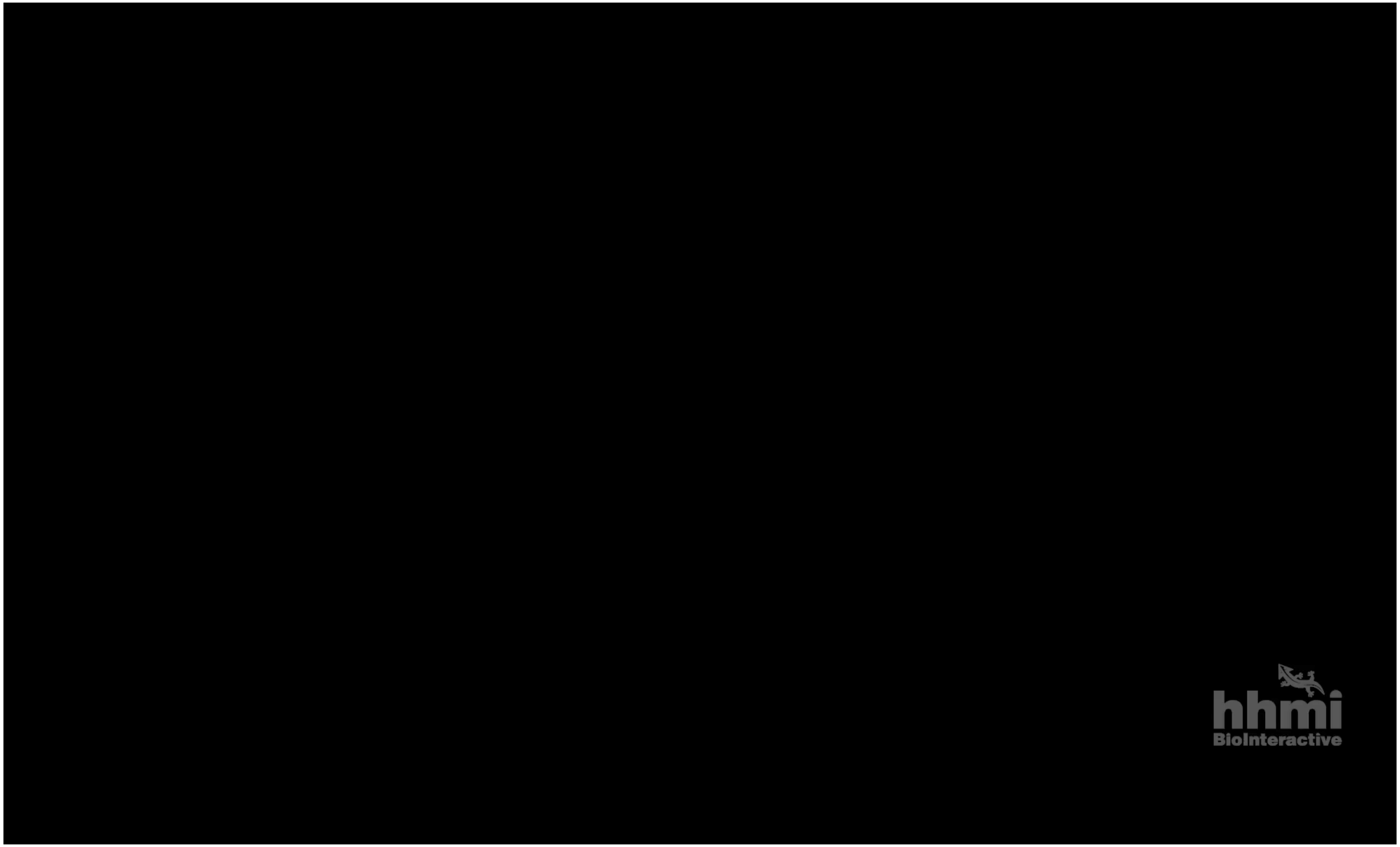
- RNA, in contrast to DNA, is single-stranded
- Complementary pairing can also occur between two RNA molecules or between parts of the same molecule
- In RNA, thymine is replaced by uracil (U), so A and U pair
- While DNA always exists as a double helix, RNA molecules are more variable in form

Figure 5.24



5′
3′
Sugar-phosphate backbones
Hydrogen bonds
T A
G C
C G
A T
C G
T A
3′
5′
Base pair joined by hydrogen bonding
(a) DNA
Base pair joined by hydrogen bonding
G C
G C
U A
(b) Transfer RNA

Video: DNA Structure



CONCEPT 5.6: Genomics and proteomics have transformed biological inquiry and applications

- Once the structure of DNA and its relationship to amino acid sequence was understood, biologists sought to “decode” genes by learning their base sequences
- The first chemical techniques for DNA sequencing were developed in the 1970s and refined over the next 20 years

- It is enlightening to sequence the full complement of DNA in an organism's genome
- The rapid development of faster and less expensive methods of sequencing was a side effect of the Human Genome Project
- Many genomes have been sequenced, generating large sets of data

Figure 5.25

- **Bioinformatics** uses computer software and other computational tools to deal with the data resulting from sequencing many genomes
- Analyzing large sets of genes or even comparing whole genomes of different species is called **genomics**
- A similar analysis of large sets of proteins including their sequences is called **proteomics**

Figure 5.26

MAKE CONNECTIONS:
Contributions of Genomics and Proteomics to Biology

Paleontology

Evolution

Hippopotamus

Short-finned pilot whale

Medical Science

Conservation Biology

Species Interactions

DNA and Proteins as Tape Measures of Evolution

- Sequences of genes and their protein products document the hereditary background of an organism
- Linear sequences of DNA molecules are passed from parents to offspring
- We can extend the concept of “molecular genealogy” to relationships between species
- Molecular biology has added a new measure to the toolkit of evolutionary biology

Figure 5.UN02a

Species	Alignment of Amino Acid Sequences of β-globin					
Human	1	VHLTPEEKSA	VTALWGKVNV	DEVGGEALGR	LLVVYPWTQR	FFESFGDLST
Monkey	1	VHLTPEEKNA	VTTLWGKVNV	DEVGGEALGR	LLLVYPWTQR	FFESFGDLSS
Gibbon	1	VHLTPEEKSA	VTALWGKVNV	DEVGGEALGR	LLVVYPWTQR	FFESFGDLST
Human	51	PDAVMGNPKV	KAHGKKVLGA	FSDGLAHLDN	LKGTFATLSE	LHCDKLHVDP
Monkey	51	PDAVMGNPKV	KAHGKKVLGA	FSDGLNHLDN	LKGTFAQLSE	LHCDKLHVDP
Gibbon	51	PDAVMGNPKV	KAHGKKVLGA	FSDGLAHLDN	LKGTFAQLSE	LHCDKLHVDP
Human	101	ENFRLLGNVL	VCVLAHHFGK	EFTPPVQAAY	QKVVAGVANA	LAHKYH
Monkey	101	ENFRLLGNVL	VCVLAHHFGK	EFTPPVQAAY	QKVVAGVANA	LAHKYH
Gibbon	101	ENFRLLGNVL	VCVLAHHFGK	EFTPPVQAAY	QKVVAGVANA	LAHKYH

Data from Human: http://www.ncbi.nlm.nih.gov/protein/AAA21113.1; rhesus monkey: http://www.ncbi.nlm.nih.gov/protein/122634; gibbon: http://www.ncbi.nlm.nih.gov/protein/122616

Figure 5.UN02b

▶ Human

▶ Rhesus monkey

▶ Gibbon

Figure 5.UN03a

	Sample labeled as *O. kisutch* (coho salmon)	5′-CGGCACCGCCCTAAGTCTCT-3′
Known sequences	*O. kisutch* (coho salmon)	5′-AGGCACCGCCCTAAGTCTAC-3′
	O. keta (chum salmon)	5′-AGGCACCGCCCTGAGCCTAC-3′
	Salmo salar (Atlantic salmon)	5′-CGGCACCGCCCTAAGTCTCT-3′

Data from Known sequences are from the International Barcode of Life.

Figure 5.UN03b

UNF05-cr

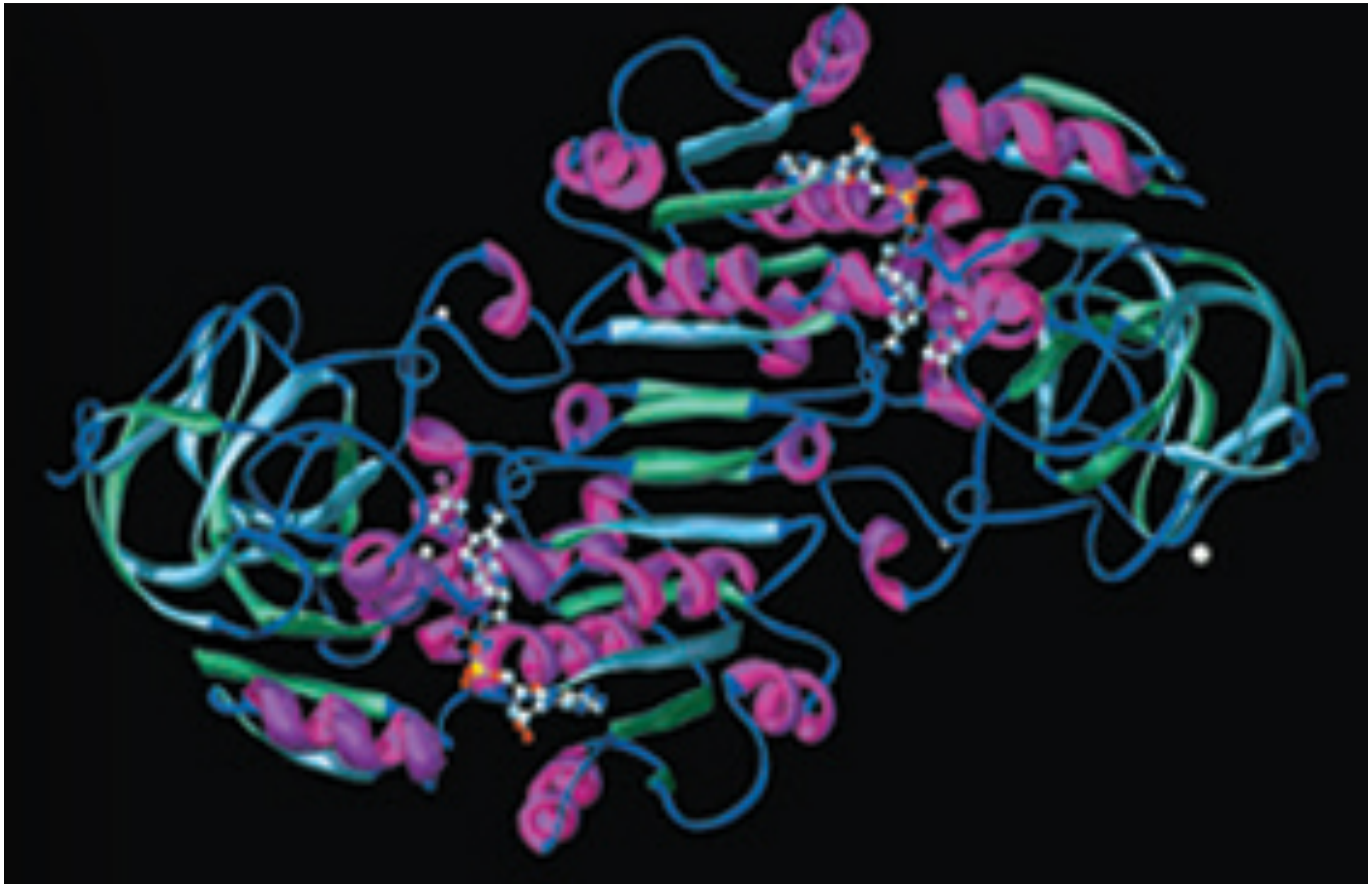


Figure 5.UN04

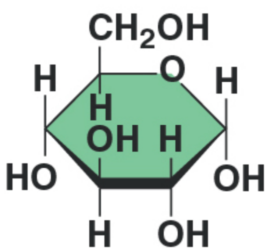
Components	Examples	Functions
 Monosaccharide monomer	Monosaccharides: glucose, fructose	Fuel; carbon sources that can be converted to other molecules or combined into polymers
	Disaccharides: lactose, sucrose	
	Polysaccharides: • Cellulose (plants) • Starch (plants) • Glycogen (animals) • Chitin (animals and fungi)	• Cellulose strengthens plant cell walls • Starch stores glucose for energy in plants • Glycogen stores glucose for energy in animals • Chitin strengthens animal exoskeletons and fungal cell walls

Figure 5.UN05

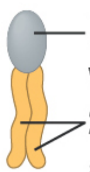
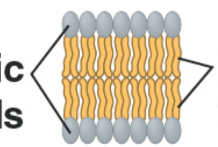
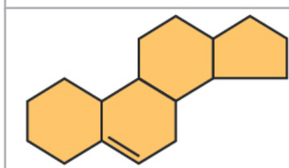
Components	Examples	Functions
Glycerol  3 fatty acids	**Triacylglycerols (fats or oils): glycerol + three fatty acids**	**Important energy source**  
 Head with P 2 fatty acids	**Phospholipids: glycerol + phosphate group + two fatty acids**	**Lipid bilayers of membranes** Hydrophilic heads  Hydrophobic tails
 Steroid backbone	**Steroids: four fused rings with attached chemical groups**	• **Component of cell membranes (cholesterol)** • **Signaling molecules that travel through the body (hormones)**

Figure 5.UN06

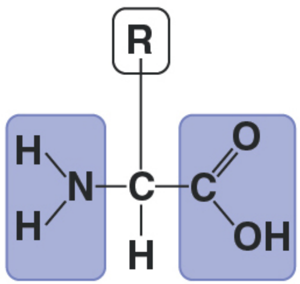
Components	Examples	Functions
 Amino acid monomer (20 types)	• Enzymes • Defensive proteins • Storage proteins • Transport proteins • Hormones • Receptor proteins • Motor proteins • Structural proteins	• Catalyze chemical reactions • Protect against disease • Store amino acids • Transport substances • Coordinate organismal responses • Receive signals from outside cell • Function in cell movement • Provide structural support

Figure 5.UN07

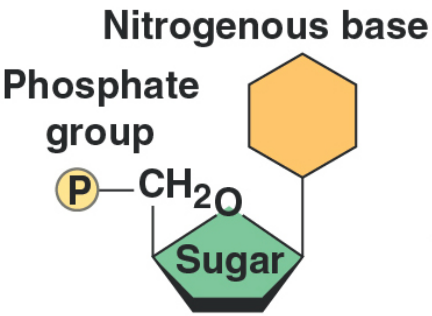
Components	Examples	Functions
 Nucleotide (monomer of a polynucleotide)	DNA:  • Sugar = deoxyribose • Nitrogenous bases = C, G, A, T • Usually double-stranded	Stores hereditary information
	RNA:  • Sugar = ribose • Nitrogenous bases = C, G, A, U • Usually single-stranded	Various functions in gene expression, including carrying instructions from DNA to ribosomes

Figure 5.UN08