

polypeptide aggregation that comprise the complete protein structure, and they disassociate from the protein once the target protein is folded.

LINK TO LEARNING

For an additional perspective on proteins, view [this animation \(http://openstax.org//proteins\)](http://openstax.org//proteins) called “Biomolecules: The Proteins.”

3.5 Nucleic Acids

By the end of this section, you will be able to do the following:

- Describe nucleic acids' structure and define the two types of nucleic acids
- Explain DNA's structure and role
- Explain RNA's structure and roles

Nucleic acids are the most important macromolecules for the continuity of life. They carry the cell's genetic blueprint and carry instructions for its functioning.

DNA and RNA

The two main types of nucleic acids are **deoxyribonucleic acid (DNA)** and **ribonucleic acid (RNA)**. DNA is the genetic material in all living organisms, ranging from single-celled bacteria to multicellular mammals. It is in the nucleus of eukaryotes and in the organelles, chloroplasts, and mitochondria. In prokaryotes, the DNA is not enclosed in a membranous envelope.

The cell's entire genetic content is its genome, and the study of genomes is genomics. In eukaryotic cells but not in prokaryotes, DNA forms a complex with histone proteins to form chromatin, the substance of eukaryotic chromosomes. A chromosome may contain tens of thousands of genes. Many genes contain the information to make protein products. Other genes code for RNA products. DNA controls all of the cellular activities by turning the genes “on” or “off.”

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. The DNA molecules never leave the nucleus but instead use an intermediary to communicate with the rest of the cell. This intermediary is the **messenger RNA (mRNA)**. Other types of RNA—like rRNA, tRNA, and microRNA—are involved in protein synthesis and its regulation.

DNA and RNA are comprised of monomers that scientists call **nucleotides**. The nucleotides combine with each other to form a **polynucleotide**, DNA or RNA. Three components comprise each nucleotide: a nitrogenous base, a pentose (five-carbon) sugar, and a phosphate group ([Figure 3.31](#)). Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.

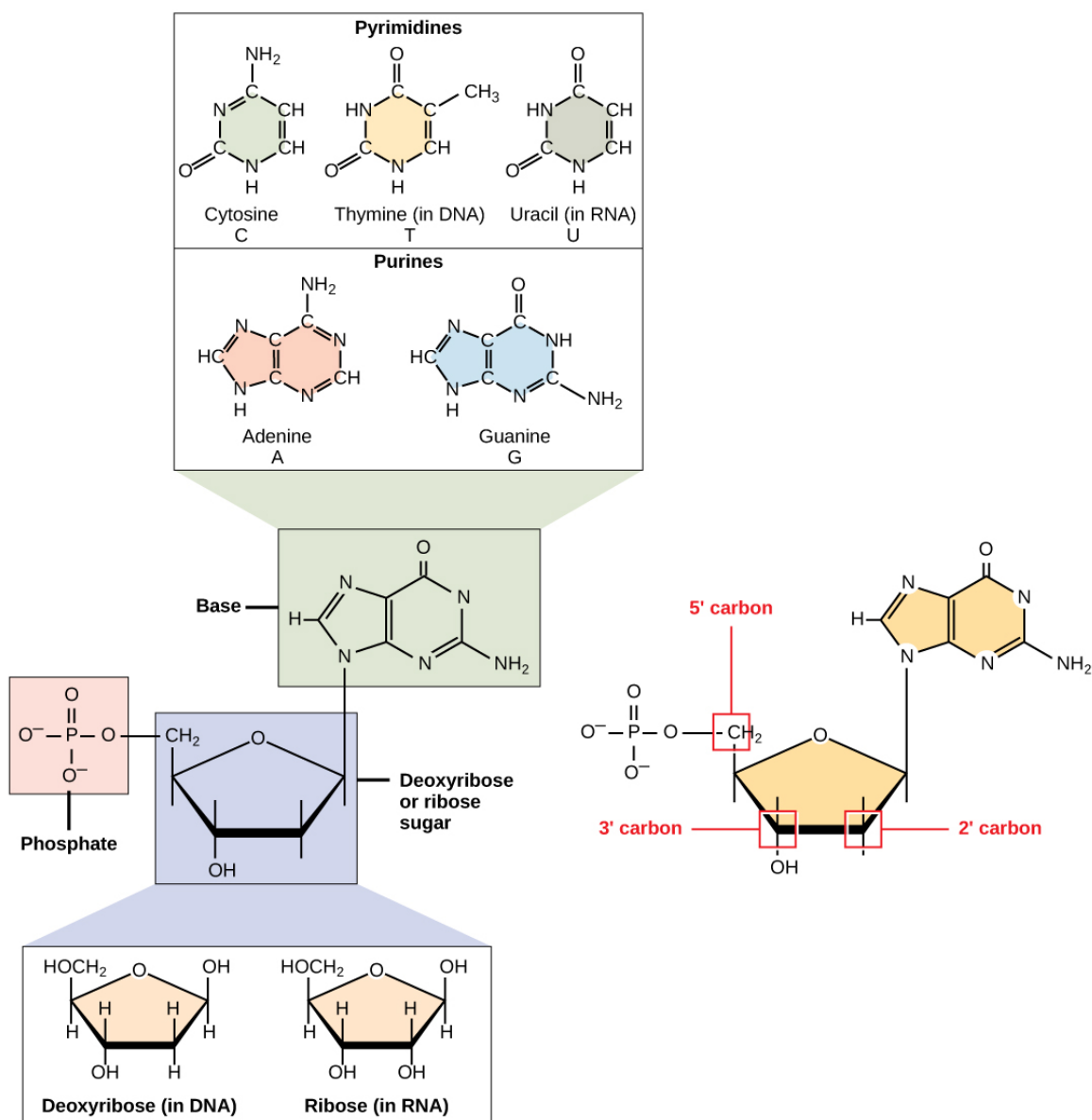


Figure 3.31 Three components comprise a nucleotide: a nitrogenous base, a pentose sugar, and one or more phosphate groups. Carbon residues in the pentose are numbered 1' through 5' (the prime distinguishes these residues from those in the base, which are numbered without using a prime notation). The base is attached to the ribose's 1' position, and the phosphate is attached to the 5' position. When a polynucleotide forms, the incoming nucleotide's 5' phosphate attaches to the 3' hydroxyl group at the end of the growing chain. Two types of pentose are in nucleotides, deoxyribose (found in DNA) and ribose (found in RNA). Deoxyribose is similar in structure to ribose, but it has an H instead of an OH at the 2' position. We can divide bases into two categories: purines and pyrimidines. Purines have a double ring structure, and pyrimidines have a single ring.

The nitrogenous bases, important components of nucleotides, are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus decreasing the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G), cytosine (C), and thymine (T).

Scientists classify adenine and guanine as **purines**. The purine's primary structure is two carbon-nitrogen rings. Scientists classify cytosine, thymine, and uracil as **pyrimidines** which have a single carbon-nitrogen ring as their primary structure ([Figure 3.31](#)). Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, we know the nitrogenous bases by their symbols A, T, G, C, and U. DNA contains A, T, G, and C; whereas, RNA contains A, U, G, and C.

The pentose sugar in DNA is deoxyribose, and in RNA, the sugar is ribose ([Figure 3.31](#)). The difference between the sugars is the presence of the hydroxyl group on the ribose's second carbon and hydrogen on the deoxyribose's second carbon. The carbon atoms of the sugar molecule are numbered as 1', 2', 3', 4', and 5' (1' is read as "one prime"). The phosphate residue attaches to the hydroxyl group of the 5' carbon of one sugar and the hydroxyl group of the 3' carbon of the sugar of the next nucleotide, which forms a 5'–3' **phosphodiester** linkage. A simple dehydration reaction like the other linkages connecting monomers in macromolecules does not form the phosphodiester linkage. Its formation involves removing two phosphate groups. A polynucleotide may have thousands of such phosphodiester linkages.

DNA Double-Helix Structure

DNA has a double-helix structure ([Figure 3.32](#)). The sugar and phosphate lie on the outside of the helix, forming the DNA's backbone. The nitrogenous bases are stacked in the interior, like a pair of staircase steps. Hydrogen bonds bind the pairs to each other. Every base pair in the double helix is separated from the next base pair by 0.34 nm. The helix's two strands run in opposite directions, meaning that the 5' carbon end of one strand will face the 3' carbon end of its matching strand. (Scientists call this an antiparallel orientation and is important to DNA replication and in many nucleic acid interactions.)

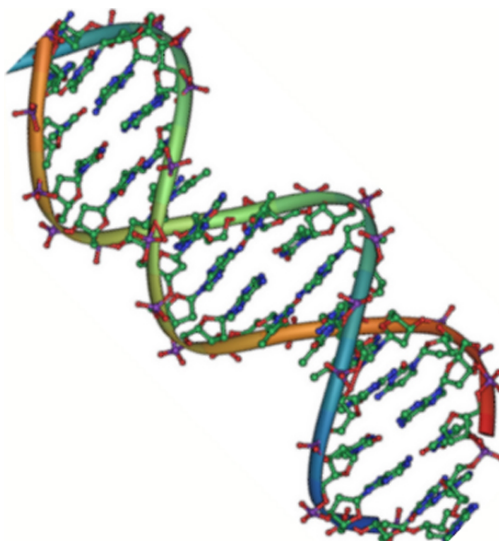


Figure 3.32 Native DNA is an antiparallel double helix. The phosphate backbone (indicated by the curvy lines) is on the outside, and the bases are on the inside. Each base from one strand interacts via hydrogen bonding with a base from the opposing strand. (credit: Jerome Walker/Dennis Myts)

Only certain types of base pairing are allowed. For example, a certain purine can only pair with a certain pyrimidine. This means A can pair with T, and G can pair with C, as [Figure 3.33](#) shows. This is the base complementary rule. In other words, the DNA strands are complementary to each other. If the sequence of one strand is AATTGGCC, the complementary strand would have the sequence TTAACCGG. During DNA replication, each strand copies itself, resulting in a daughter DNA double helix containing one parental DNA strand and a newly synthesized strand.



VISUAL CONNECTION

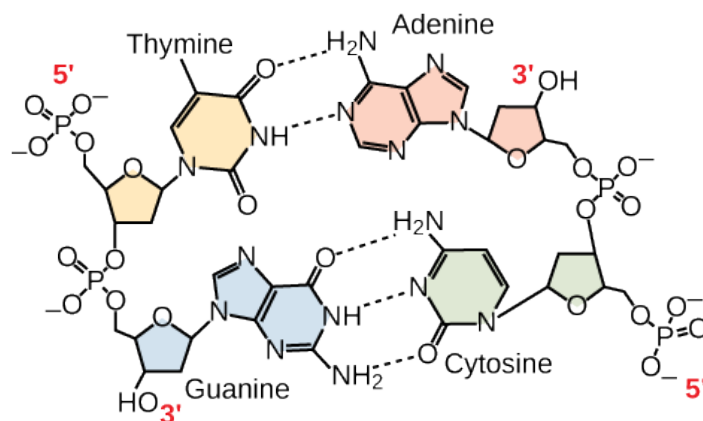


Figure 3.33 In a double stranded DNA molecule, the two strands run antiparallel to one another so that one strand runs 5' to 3' and the other 3' to 5'. The phosphate backbone is located on the outside, and the bases are in the middle. Adenine forms hydrogen bonds (or base pairs) with thymine, and guanine base pairs with cytosine.

A mutation occurs, and adenine replaces cytosine. What impact do you think this will have on the DNA structure?

RNA

Ribonucleic acid, or RNA, is mainly involved in the process of protein synthesis under the direction of DNA. RNA is usually single-stranded and is comprised of ribonucleotides that are linked by phosphodiester bonds. A ribonucleotide in the RNA chain contains ribose (the pentose sugar), one of the four nitrogenous bases (A, U, G, and C), and the phosphate group.

There are four major types of RNA: messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA), and microRNA (miRNA). The first, mRNA, carries the message from DNA, which controls all of the cellular activities in a cell. If a cell requires synthesizing a certain protein, the gene for this product turns “on” and the messenger RNA synthesizes in the nucleus. The RNA base sequence is complementary to the DNA’s coding sequence from which it has been copied. However, in RNA, the base T is absent and U is present instead. If the DNA strand has a sequence AATTGCGC, the sequence of the complementary RNA is UUAACGCG. In the cytoplasm, the mRNA interacts with ribosomes and other cellular machinery ([Figure 3.34](#)).

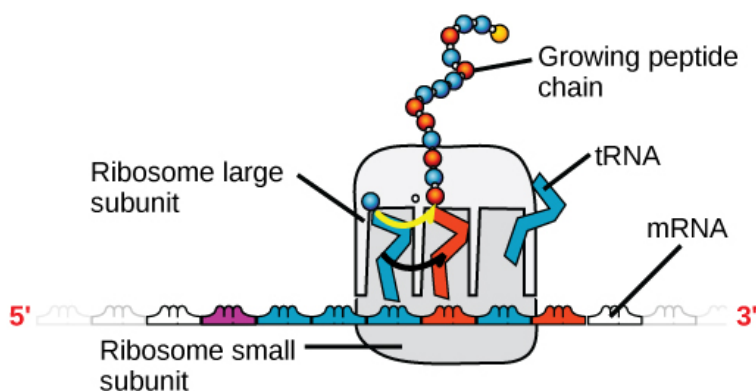


Figure 3.34 A ribosome has two parts: a large subunit and a small subunit. The mRNA sits in between the two subunits. A tRNA molecule recognizes a codon on the mRNA, binds to it by complementary base pairing, and adds the correct amino acid to the growing peptide chain.

The mRNA is read in sets of three bases known as codons. Each codon codes for a single amino acid. In this way, the mRNA is read and the protein product is made. **Ribosomal RNA (rRNA)** is a major constituent of ribosomes on which the mRNA binds. The rRNA ensures the proper alignment of the mRNA and the Ribosomes. The ribosome’s rRNA also has an enzymatic activity (peptidyl transferase) and catalyzes peptide bond formation between two aligned amino acids. **Transfer RNA (tRNA)** is one of the smallest of the four types of RNA, usually 70–90 nucleotides long. It carries the correct amino acid to the protein synthesis

site. It is the base pairing between the tRNA and mRNA that allows for the correct amino acid to insert itself in the polypeptide chain. MicroRNAs are the smallest RNA molecules and their role involves regulating gene expression by interfering with the expression of certain mRNA messages. [Table 3.2](#) summarizes DNA and RNA features.

DNA and RNA Features

	DNA	RNA
Function	Carries genetic information	Involved in protein synthesis
Location	Remains in the nucleus	Leaves the nucleus
Structure	Double helix	Usually single-stranded
Sugar	Deoxyribose	Ribose
Pyrimidines	Cytosine, thymine	Cytosine, uracil
Purines	Adenine, guanine	Adenine, guanine

Table 3.2

Even though the RNA is single stranded, most RNA types show extensive intramolecular base pairing between complementary sequences, creating a predictable three-dimensional structure essential for their function.

As you have learned, information flow in an organism takes place from DNA to RNA to protein. DNA dictates the structure of mRNA in a process scientists call **transcription**, and RNA dictates the protein's structure in a process scientists call **translation**. This is the Central Dogma of Life, which holds true for all organisms; however, exceptions to the rule occur in connection with viral infections.

LINK TO LEARNING

To learn more about DNA, explore the [Howard Hughes Medical Institute BioInteractive animations \(http://openstax.org/l/DNA\)](http://openstax.org/l/DNA) on the topic of DNA.

KEY TERMS

alpha-helix structure (α -helix) type of secondary protein structure formed by folding the polypeptide into a helix shape with hydrogen bonds stabilizing the structure

amino acid a protein's monomer; has a central carbon or alpha carbon to which an amino group, a carboxyl group, a hydrogen, and an R group or side chain is attached; the R group is different for all 20 common amino acids

beta-pleated sheet (β -pleated) secondary structure in proteins in which hydrogen bonding forms "pleats" between atoms on the polypeptide chain's backbone

biological macromolecule large molecule necessary for life that is built from smaller organic molecules

carbohydrate biological macromolecule in which the ratio of carbon to hydrogen and to oxygen is 1:2:1; carbohydrates serve as energy sources and structural support in cells and form arthropods' cellular exoskeleton

cellulose polysaccharide that comprises the plants' cell wall; provides structural support to the cell

chaperone (also, chaperonin) protein that helps nascent protein in the folding process

chitin type of carbohydrate that forms the outer skeleton of all arthropods that include crustaceans and insects; it also forms fungi cell walls

dehydration synthesis (also, condensation) reaction that links monomer molecules, releasing a water molecule for each bond formed

denaturation loss of shape in a protein as a result of changes in temperature, pH, or chemical exposure

deoxyribonucleic acid (DNA) double-helical molecule that carries the cell's hereditary information

disaccharide two sugar monomers that a glycosidic bond links

enzyme catalyst in a biochemical reaction that is usually a complex or conjugated protein

glycogen storage carbohydrate in animals

glycosidic bond bond formed by a dehydration reaction between two monosaccharides with eliminating a water molecule

hormone chemical signaling molecule, usually protein or steroid, secreted by endocrine cells that act to control or regulate specific physiological processes

hydrolysis reaction that causes breakdown of larger molecules into smaller molecules by utilizing water

lipid macromolecule that is nonpolar and insoluble in water

messenger RNA (mRNA) RNA that carries information from DNA to ribosomes during protein synthesis

monomer smallest unit of larger molecules that are polymers

monosaccharide single unit or monomer of carbohydrates

nucleic acid biological macromolecule that carries the cell's genetic blueprint and carries instructions for the cell's

functioning

nucleotide monomer of nucleic acids; contains a pentose sugar, one or more phosphate groups, and a nitrogenous base

omega fat type of polyunsaturated fat that the body requires; numbering the carbon omega starts from the methyl end or the end that is farthest from the carboxylic end

peptide bond bond formed between two amino acids by a dehydration reaction

phosphodiester linkage covalent chemical bond that holds together the polynucleotide chains with a phosphate group linking neighboring nucleotides' two pentose sugars

phospholipid membranes' major constituent; comprised of two fatty acids and a phosphate-containing group attached to a glycerol backbone

polymer chain of monomer residues that covalent bonds link; polymerization is the process of polymer formation from monomers by condensation

polynucleotide long chain of nucleotides

polypeptide long chain of amino acids that peptide bonds link

polysaccharide long chain of monosaccharides; may be branched or unbranched

primary structure linear sequence of amino acids in a protein

protein biological macromolecule comprised of one or more amino acid chains

purine type of nitrogenous base in DNA and RNA; adenine and guanine are purines

pyrimidine type of nitrogenous base in DNA and RNA; cytosine, thymine, and uracil are pyrimidines

quaternary structure association of discrete polypeptide subunits in a protein

ribonucleic acid (RNA) single-stranded, often internally base paired, molecule that is involved in protein synthesis

ribosomal RNA (rRNA) RNA that ensures the proper alignment of the mRNA and the ribosomes during protein synthesis and catalyzes forming the peptide linkage

saturated fatty acid long-chain hydrocarbon with single covalent bonds in the carbon chain; the number of hydrogen atoms attached to the carbon skeleton is maximized

secondary structure regular structure that proteins form by intramolecular hydrogen bonding between the oxygen atom of one amino acid residue and the hydrogen attached to the nitrogen atom of another amino acid residue

starch storage carbohydrate in plants

steroid type of lipid comprised of four fused hydrocarbon

rings forming a planar structure

tertiary structure a protein's three-dimensional conformation, including interactions between secondary structural elements; formed from interactions between amino acid side chains

trans fat fat formed artificially by hydrogenating oils, leading to a different arrangement of double bond(s) than those in naturally occurring lipids

transcription process through which messenger RNA forms on a template of DNA

transfer RNA (tRNA) RNA that carries activated amino

acids to the site of protein synthesis on the ribosome

translation process through which RNA directs the protein's formation

triacylglycerol (also, triglyceride) fat molecule; consists of three fatty acids linked to a glycerol molecule

unsaturated fatty acid long-chain hydrocarbon that has one or more double bonds in the hydrocarbon chain

wax lipid comprised of a long-chain fatty acid that is esterified to a long-chain alcohol; serves as a protective coating on some feathers, aquatic mammal fur, and leaves

CHAPTER SUMMARY

3.1 Synthesis of Biological Macromolecules

Proteins, carbohydrates, nucleic acids, and lipids are the four major classes of biological macromolecules—large molecules necessary for life that are built from smaller organic molecules. Macromolecules are comprised of single units scientists call monomers that are joined by covalent bonds to form larger polymers. The polymer is more than the sum of its parts: it acquires new characteristics, and leads to an osmotic pressure that is much lower than that formed by its ingredients. This is an important advantage in maintaining cellular osmotic conditions. A monomer joins with another monomer with water molecule release, leading to a covalent bond forming. Scientists call these dehydration or condensation reactions. When polymers break down into smaller units (monomers), they use a water molecule for each bond broken by these reactions. Such reactions are hydrolysis reactions. Dehydration and hydrolysis reactions are similar for all macromolecules, but each monomer and polymer reaction is specific to its class. Dehydration reactions typically require an investment of energy for new bond formation, while hydrolysis reactions typically release energy by breaking bonds.

3.2 Carbohydrates

Carbohydrates are a group of macromolecules that are a vital energy source for the cell and provide structural support to plant cells, fungi, and all of the arthropods that include lobsters, crabs, shrimp, insects, and spiders. Scientists classify carbohydrates as monosaccharides, disaccharides, and polysaccharides depending on the number of monomers in the molecule. Monosaccharides are linked by glycosidic bonds that form as a result of dehydration reactions, forming disaccharides and polysaccharides with eliminating a water molecule for each bond formed. Glucose, galactose, and fructose are common monosaccharides; whereas, common disaccharides include lactose, maltose, and sucrose. Starch and glycogen, examples of polysaccharides, are the storage forms of glucose in plants and animals,

respectively. The long polysaccharide chains may be branched or unbranched. Cellulose is an example of an unbranched polysaccharide; whereas, amylopectin, a constituent of starch, is a highly branched molecule. Glucose storage, in the form of polymers like starch or glycogen, makes it slightly less accessible for metabolism; however, this prevents it from leaking out of the cell or creating a high osmotic pressure that could cause the cell to uptake excessive water.

3.3 Lipids

Lipids are a class of macromolecules that are nonpolar and hydrophobic in nature. Major types include fats and oils, waxes, phospholipids, and steroids. Fats are a stored form of energy and are also known as triacylglycerols or triglycerides. Fats are comprised of fatty acids and either glycerol or sphingosine. Fatty acids may be unsaturated or saturated, depending on the presence or absence of double bonds in the hydrocarbon chain. If only single bonds are present, they are saturated fatty acids. Unsaturated fatty acids may have one or more double bonds in the hydrocarbon chain. Phospholipids comprise the membrane's matrix. They have a glycerol or sphingosine backbone to which two fatty acid chains and a phosphate-containing group are attached. Steroids are another class of lipids. Their basic structure has four fused carbon rings. Cholesterol is a type of steroid and is an important constituent of the plasma membrane, where it helps to maintain the membrane's fluid nature. It is also the precursor of steroid hormones such as testosterone.

3.4 Proteins

Proteins are a class of macromolecules that perform a diverse range of functions for the cell. They help in metabolism by acting as enzymes, carriers, or hormones, and provide structural support. The building blocks of proteins (monomers) are amino acids. Each amino acid has a central carbon that bonds to an amino group, a carboxyl group, a hydrogen atom, and an R group or side chain. There are 20 commonly occurring amino acids, each of which differs in the R group. A peptide bond links each amino acid

to its neighbors. A long amino acid chain is a polypeptide.

Proteins are organized at four levels: primary, secondary, tertiary, and (optional) quaternary. The primary structure is the amino acids' unique sequence. The polypeptide's local folding to form structures such as the α -helix and β -pleated sheet constitutes the secondary structure. The overall three-dimensional structure is the tertiary structure. When two or more polypeptides combine to form the complete protein structure, the configuration is the protein's quaternary structure. Protein shape and function are intricately linked. Any change in shape caused by changes in temperature or pH may lead to protein denaturation and a loss in function.

3.5 Nucleic Acids

Nucleic acids are molecules comprised of nucleotides that direct cellular activities such as cell division and protein synthesis. Pentose sugar, a nitrogenous base, and a

phosphate group comprise each nucleotide. There are two types of nucleic acids: DNA and RNA. DNA carries the cell's genetic blueprint and passes it on from parents to offspring (in the form of chromosomes). It has a double-helical structure with the two strands running in opposite directions, connected by hydrogen bonds, and complementary to each other. RNA is a single-stranded polymer composed of linked nucleotides made up of a pentose sugar (ribose), a nitrogenous base, and a phosphate group. RNA is involved in protein synthesis and its regulation. Messenger RNA (mRNA) copies from the DNA, exports itself from the nucleus to the cytoplasm, and contains information for constructing proteins. Ribosomal RNA (rRNA) is a part of the ribosomes at the site of protein synthesis; whereas, transfer RNA (tRNA) carries the amino acid to the site of protein synthesis. The microRNA regulates using mRNA for protein synthesis.

VISUAL CONNECTION QUESTIONS

1. [Figure 3.5](#) What kind of sugars are these, aldose or ketose?
2. [Figure 3.23](#) Which categories of amino acid would you expect to find on the surface of a soluble protein, and which would you expect to find in the interior? What distribution of amino acids would you expect to find in a protein embedded in a lipid bilayer?
3. [Figure 3.33](#) A mutation occurs, and cytosine is replaced with adenine. What impact do you think this will have on the DNA structure?

REVIEW QUESTIONS

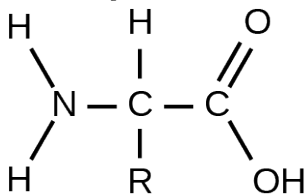
4. Dehydration synthesis leads to formation of
 - a. monomers
 - b. polymers
 - c. water and polymers
 - d. none of the above
5. During the breakdown of polymers, which of the following reactions takes place?
 - a. hydrolysis
 - b. dehydration
 - c. condensation
 - d. covalent bond
6. The following chemical reactants produce the ester ethyl ethanoate ($C_4H_8O_2$):
 $C_2H_6O + CH_3COOH$
 What type of reaction occurs to make ethyl ethanoate?
 - a. condensation
 - b. hydrolysis
 - c. combustion
 - d. acid-base reaction
7. An example of a monosaccharide is _____.
 - a. fructose
 - b. glucose
 - c. galactose
 - d. all of the above
8. Cellulose and starch are examples of:
 - a. monosaccharides
 - b. disaccharides
 - c. lipids
 - d. polysaccharides
9. Plant cell walls contain which of the following in abundance?
 - a. starch
 - b. cellulose
 - c. glycogen
 - d. lactose

10. Lactose is a disaccharide formed by the formation of a _____ bond between glucose and _____.
 - a. glycosidic; lactose
 - b. glycosidic; galactose
 - c. hydrogen; sucrose
 - d. hydrogen; fructose
11. Which of the following is not an extracellular matrix role of carbohydrates?
 - a. protect an insect's internal organs from external trauma
 - b. prevent plant cells from lysing after the plant is watered
 - c. maintain the shape of a fungal spore
 - d. provide energy for muscle movement
12. Saturated fats have all of the following characteristics except:
 - a. they are solid at room temperature
 - b. they have single bonds within the carbon chain
 - c. they are usually obtained from animal sources
 - d. they tend to dissolve in water easily
13. Phospholipids are important components of _____.
 - a. the plasma membrane of cells
 - b. the ring structure of steroids
 - c. the waxy covering on leaves
 - d. the double bond in hydrocarbon chains
14. Cholesterol is an integral part of plasma membranes. Based on its structure, where is it found in the membrane?
 - a. on the extracellular surface
 - b. embedded with the phospholipid heads
 - c. within the tail bilayer
 - d. attached to the intracellular surface
15. The monomers that make up proteins are called _____.
 - a. nucleotides
 - b. disaccharides
 - c. amino acids
 - d. chaperones
16. The α -helix and the β -pleated sheet are part of which protein structure?
 - a. primary
 - b. secondary
 - c. tertiary
 - d. quaternary
17. Mad cow disease is an infectious disease where one misfolded protein causes all other copies of the protein to begin misfolding. This is an example of a disease impacting _____ structure.
 - a. primary
 - b. secondary
 - c. tertiary
 - d. quaternary
18. A nucleotide of DNA may contain _____.
 - a. ribose, uracil, and a phosphate group
 - b. deoxyribose, uracil, and a phosphate group
 - c. deoxyribose, thymine, and a phosphate group
 - d. ribose, thymine, and a phosphate group
19. The building blocks of nucleic acids are _____.
 - a. sugars
 - b. nitrogenous bases
 - c. peptides
 - d. nucleotides
20. How does the double helix structure of DNA support its role in encoding the genome?
 - a. The sugar-phosphate backbone provides a template for DNA replication.
 - b. tRNA pairing with the template strand creates proteins encoded by the genome.
 - c. Complementary base pairing creates a very stable structure.
 - d. Complementary base pairing allows for easy editing of both strands of DNA.

CRITICAL THINKING QUESTIONS

21. Why are biological macromolecules considered organic?
22. What role do electrons play in dehydration synthesis and hydrolysis?

23. Amino acids have the generic structure seen below, where R represents different carbon-based side chains.



Describe how the structure of amino acids allows them to be linked into long peptide chains to form proteins.

24. Describe the similarities and differences between glycogen and starch.
25. Why is it impossible for humans to digest food that contains cellulose?
26. Draw the ketose and aldose forms of a monosaccharide with the chemical formula $\text{C}_3\text{H}_6\text{O}_3$. How is the structure of the monosaccharide changed from one form to the other in the human body?
27. Explain at least three functions that lipids serve in plants and/or animals.
28. Why have trans fats been banned from some restaurants? How are they created?
29. Why are fatty acids better than glycogen for storing large amounts of chemical energy?
30. Part of cortisol's role in the body involves passing through the plasma membrane to initiate signaling inside a cell. Describe how the structures of cortisol and the plasma membrane allow this to occur.
31. Explain what happens if even one amino acid is substituted for another in a polypeptide chain. Provide a specific example.
32. Describe the differences in the four protein structures.
33. Aquaporins are proteins embedded in the plasma membrane that allow water molecules to move between the extracellular matrix and the intracellular space. Based on its function and location, describe the key features of the protein's shape and the chemical characteristics of its amino acids.
34. What are the structural differences between RNA and DNA?
35. What are the four types of RNA and how do they function?