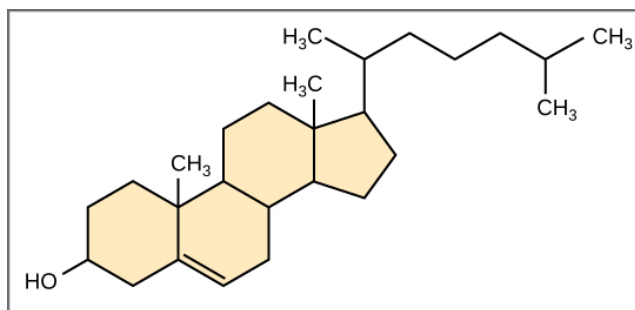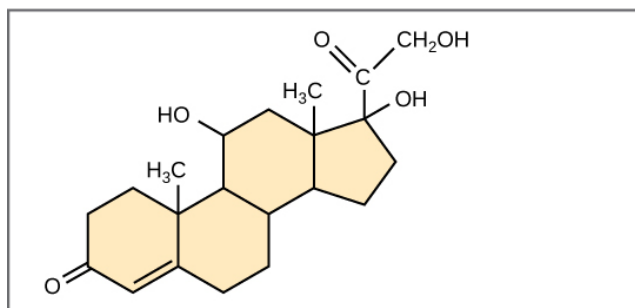insoluble in water. All steroids have four linked carbon rings and several of them, like cholesterol, have a short tail (**Figure 3.21**). Many steroids also have the –OH functional group, which puts them in the alcohol classification (sterols).

Cholesterol

Cortisol

Figure 3.21 Four fused hydrocarbon rings comprise steroids such as cholesterol and cortisol.

Cholesterol is the most common steroid. The liver synthesizes cholesterol and is the precursor to many steroid hormones such as testosterone and estradiol, which gonads and endocrine glands secrete. It is also the precursor to Vitamin D. Cholesterol is also the precursor of bile salts, which help emulsifying fats and their subsequent absorption by cells. Although lay people often speak negatively about cholesterol, it is necessary for the body's proper functioning. Sterols (cholesterol in animal cells, phytosterol in plants) are components of the plasma membrane of cells and are found within the phospholipid bilayer.

LINK TO LEARNING

For an additional perspective on lipids, explore the interactive animation **"Biomolecules: The Lipids" (http://openstaxcollege.org/l/lipids)** .

3.4 | Proteins

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

  • Describe the functions proteins perform in the cell and in tissues
  • Discuss the relationship between amino acids and proteins
  • Explain the four levels of protein organization
  • Describe the ways in which protein shape and function are linked

Proteins are one of the most abundant organic molecules in living systems and have the most diverse range of functions of all macromolecules. Proteins may be structural, regulatory, contractile, or protective. They may serve in transport, storage, or membranes; or they may be toxins or enzymes. Each cell in a living system may contain thousands of proteins, each with a unique function. Their structures, like their functions, vary greatly. They are all, however, amino acid polymers arranged in a linear sequence.

Types and Functions of Proteins

Enzymes, which living cells produce, are catalysts in biochemical reactions (like digestion) and are usually complex or conjugated proteins. Each enzyme is specific for the substrate (a reactant that binds to an enzyme) upon which it acts. The enzyme may help in breakdown, rearrangement, or synthesis reactions. We call enzymes that break down their substrates catabolic enzymes. Those that build more complex molecules from their substrates are anabolic enzymes, and enzymes that affect the rate of reaction are catalytic enzymes. Note that all enzymes increase the reaction rate and, therefore, are organic catalysts. An example of an enzyme is salivary amylase, which hydrolyzes its substrate amylose, a component of starch.

Hormones are chemical-signaling molecules, usually small proteins or steroids, secreted by endocrine cells that act to control or regulate specific physiological processes, including growth, development, metabolism, and reproduction. For example, insulin is a protein hormone that helps regulate the blood glucose level. **Table 3.1** lists the primary types and functions of proteins.

Protein Types and Functions

Type	Examples	Functions
Digestive Enzymes	Amylase, lipase, pepsin, trypsin	Help in food by catabolizing nutrients into monomeric units
Transport	Hemoglobin, albumin	Carry substances in the blood or lymph throughout the body
Structural	Actin, tubulin, keratin	Construct different structures, like the cytoskeleton
Hormones	Insulin, thyroxine	Coordinate different body systems' activity
Defense	Immunoglobulins	Protect the body from foreign pathogens
Contractile	Actin, myosin	Effect muscle contraction
Storage	Legume storage proteins, egg white (albumin)	Provide nourishment in early embryo development and the seedling

Table 3.1

Proteins have different shapes and molecular weights. Some proteins are globular in shape; whereas, others are fibrous in nature. For example, hemoglobin is a globular protein, but collagen, located in our skin, is a fibrous protein. Protein shape is critical to its function, and many different types of chemical bonds maintain this shape. Changes in temperature, pH, and exposure to chemicals may lead to permanent changes in the protein's shape, leading to loss of function, or **denaturation**. Different arrangements of the same 20 types of amino acids comprise all proteins. Two rare new amino acids were discovered recently (selenocystein and pirrolysine), and additional new discoveries may be added to the list.

Amino Acids

Amino acids are the monomers that comprise proteins. Each amino acid has the same fundamental structure, which consists of a central carbon atom, or the alpha (α) carbon, bonded to an amino group (NH_2), a carboxyl group (COOH), and to a hydrogen atom. Every amino acid also has another atom or group of atoms bonded to the central atom known as the R group (**Figure 3.22**).
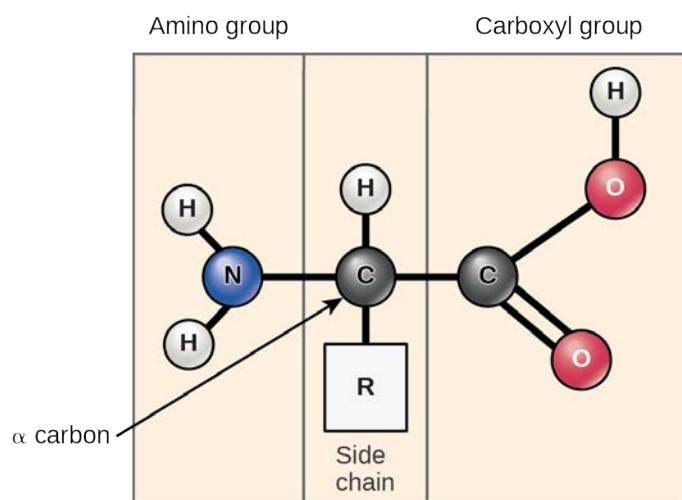
Figure 3.22 Amino acids have a central asymmetric carbon to which an amino group, a carboxyl group, a hydrogen atom, and a side chain (R group) are attached.

Scientists use the name "amino acid" because these acids contain both amino group and carboxyl-acid-group in their basic structure. As we mentioned, there are 20 common amino acids present in proteins. Nine of these are essential amino acids in humans because the human body cannot produce them and we obtain them from our diet. For each amino acid, the R group (or side chain) is different (**Figure 3.23**).
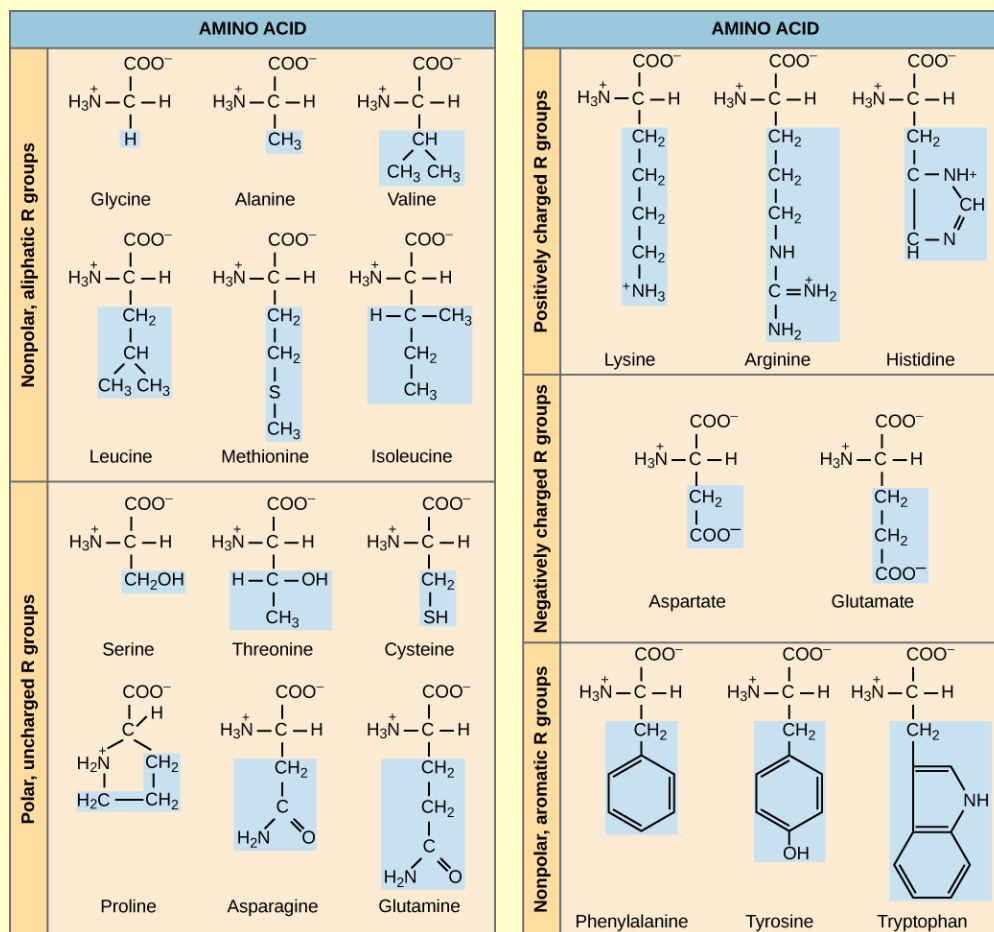
Figure 3.23 There are 20 common amino acids commonly found in proteins, each with a different R group (variant group) that determines its chemical nature.

Which categories of amino acid would you expect to find on a soluble protein's surface 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?

The chemical nature of the side chain determines the amino acid's nature (that is, whether it is acidic, basic, polar, or nonpolar). For example, the amino acid glycine has a hydrogen atom as the R group. Amino acids such as valine, methionine, and alanine are nonpolar or hydrophobic in nature, while amino acids such as serine, threonine, and cysteine are polar and have hydrophilic side chains. The side chains of lysine and arginine are positively charged, and therefore these amino acids are also basic amino acids. Proline has an R group that is linked to the amino group, forming a ring-like structure. Proline is an exception to the amino acid's standard structure since its amino group is not separate from the side chain (**Figure 3.23**).

A single upper case letter or a three-letter abbreviation represents amino acids. For example, the letter V or the three-letter symbol val represent valine. Just as some fatty acids are essential to a diet, some amino acids also are necessary. These essential amino acids in humans include isoleucine, leucine, and cysteine. Essential amino acids refer to those necessary to build proteins in the body, but not those that the body produces. Which amino acids are essential varies from organism to organism.

The sequence and the number of amino acids ultimately determine the protein's shape, size, and function. A covalent bond, or **peptide bond**, attaches to each amino acid, which a dehydration reaction forms. One amino acid's carboxyl group and the incoming amino acid's amino group combine, releasing a water molecule. The resulting bond is the peptide bond (**Figure 3.24**).
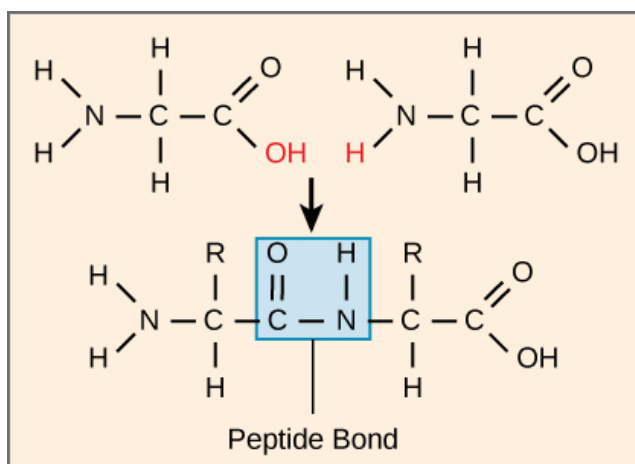
Figure 3.24 Peptide bond formation is a dehydration synthesis reaction. The carboxyl group of one amino acid is linked to the incoming amino acid's amino group. In the process, it releases a water molecule.

The products that such linkages form are peptides. As more amino acids join to this growing chain, the resulting chain is a polypeptide. Each polypeptide has a free amino group at one end. This end the N terminal, or the amino terminal, and the other end has a free carboxyl group, also the C or carboxyl terminal. While the terms polypeptide and protein are sometimes used interchangeably, a polypeptide is technically a polymer of amino acids, whereas the term protein is used for a polypeptide or polypeptides that have combined together, often have bound non-peptide prosthetic groups, have a distinct shape, and have a unique function. After protein synthesis (translation), most proteins are modified. These are known as post-translational modifications. They may undergo cleavage, phosphorylation, or may require adding other chemical groups. Only after these modifications is the protein completely functional.

LINK TO LEARNING

Click through the steps of protein synthesis in this **interactive tutorial (http://openstaxcollege.org/l/ protein_synth)** .

e**v**olution CONNECTION

The Evolutionary Significance of Cytochrome c

Cytochrome c is an important component of the electron transport chain, a part of cellular respiration, and it is normally located in the cellular organelle, the mitochondrion. This protein has a heme prosthetic group, and the heme's central ion alternately reduces and oxidizes during electron transfer. Because this essential protein's role in producing cellular energy is crucial, it has changed very little over millions of years. Protein sequencing has shown that there is a considerable amount of cytochrome c amino acid sequence homology among different species. In other words, we can assess evolutionary kinship by measuring the similarities or differences among various species' DNA or protein sequences.

Scientists have determined that human cytochrome c contains 104 amino acids. For each cytochrome c molecule from different organisms that scientists have sequenced to date, 37 of these amino acids appear in the same position in all cytochrome c samples. This indicates that there may have been a common ancestor. On comparing the human and chimpanzee protein sequences, scientists did not find a sequence difference. When researchers compared human and rhesus monkey sequences, the single difference was in one amino acid. In another comparison, human to yeast sequencing shows a difference in the 44th position.

Protein Structure

As we discussed earlier, a protein's shape is critical to its function. For example, an enzyme can bind to a specific substrate at an active site. If this active site is altered because of local changes or changes in overall protein structure, the enzyme may be unable to bind to the substrate. To understand how the protein gets its final shape or conformation, we need to understand the four levels of protein structure: primary, secondary, tertiary, and quaternary.

Primary Structure

Amino acids' unique sequence in a polypeptide chain is its **primary structure**. For example, the pancreatic hormone insulin has two polypeptide chains, A and B, and they are linked together by disulfide bonds. The N terminal amino acid of the A chain is glycine; whereas, the C terminal amino acid is asparagine (**Figure 3.25**). The amino acid sequences in the A and B chains are unique to insulin.
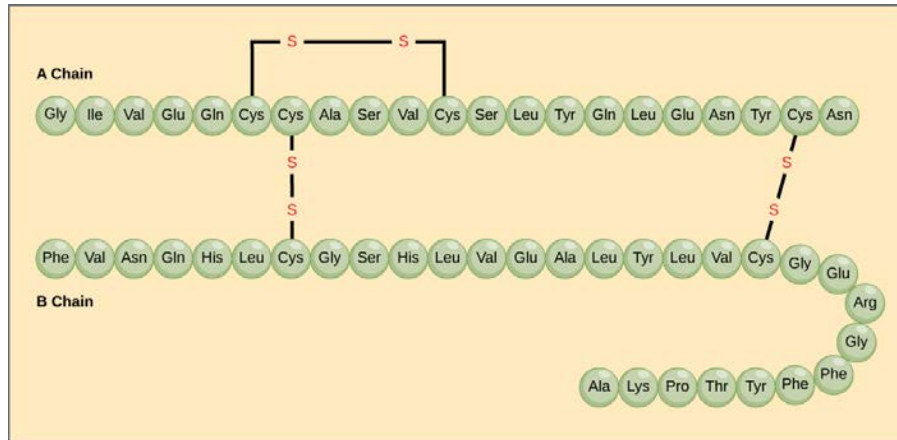

Figure 3.25 Bovine serum insulin is a protein hormone comprised of two peptide chains, A (21 amino acids long) and B (30 amino acids long). In each chain, three-letter abbreviations that represent the amino acids' names in the order they are present indicate primary structure. The amino acid cysteine (cys) has a sulfhydryl (SH) group as a side chain. Two sulfhydryl groups can react in the presence of oxygen to form a disulfide (S-S) bond. Two disulfide bonds connect the A and B chains together, and a third helps the A chain fold into the correct shape. Note that all disulfide bonds are the same length, but we have drawn them different sizes for clarity.

The gene encoding the protein ultimately determines the unique sequence for every protein. A change in nucleotide sequence of the gene's coding region may lead to adding a different amino acid to the growing polypeptide chain, causing a change in protein structure and function. In sickle cell anemia, the hemoglobin β chain (a small portion of which we show in **Figure 3.26**) has a single amino acid substitution, causing a change in protein structure and function. Specifically, valine in the β chain substitutes the amino acid glutamic. What is most remarkable to consider is that a hemoglobin molecule is comprised of two alpha and two beta chains that each consist of about 150 amino acids. The molecule, therefore, has about 600 amino acids. The structural difference between a normal hemoglobin molecule and a sickle cell molecule—which dramatically decreases life expectancy—is a single amino acid of the 600. What is even more remarkable is that three nucleotides each encode those 600 amino acids, and a single base change (point mutation), 1 in 1800 bases causes the mutation.
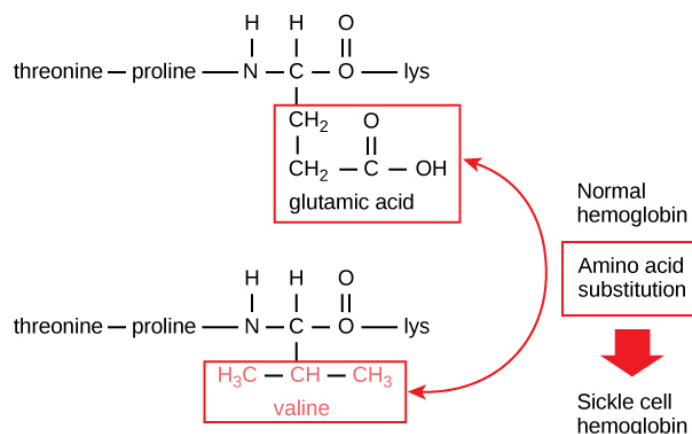
Figure 3.26 The beta chain of hemoglobin is 147 residues in length, yet a single amino acid substitution leads to sickle cell anemia. In normal hemoglobin, the amino acid at position seven is glutamate. In sickle cell hemoglobin, a valine replaces glutamate.

Because of this change of one amino acid in the chain, hemoglobin molecules form long fibers that distort the biconcave, or disc-shaped, red blood cells and causes them to assume a crescent or "sickle" shape, which clogs blood vessels (**Figure 3.27**). This can lead to myriad serious health problems such as breathlessness, dizziness, headaches, and abdominal pain for those affected by this disease.
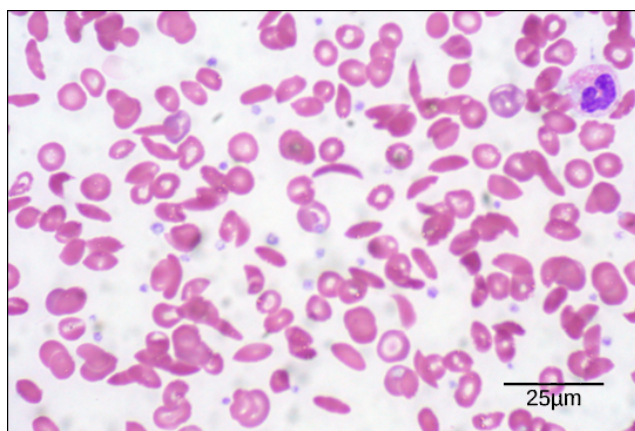

Figure 3.27 In this blood smear, visualized at 535x magnification using bright field microscopy, sickle cells are crescent shaped, while normal cells are disc-shaped. (credit: modification of work by Ed Uthman; scale-bar data from Matt Russell)

Secondary Structure

The local folding of the polypeptide in some regions gives rise to the **secondary structure** of the protein. The most common are the **α-helix** and **β-pleated sheet** structures (**Figure 3.28**). Both structures are held in shape by hydrogen bonds. The hydrogen bonds form between the oxygen atom in the carbonyl group in one amino acid and another amino acid that is four amino acids farther along the chain.
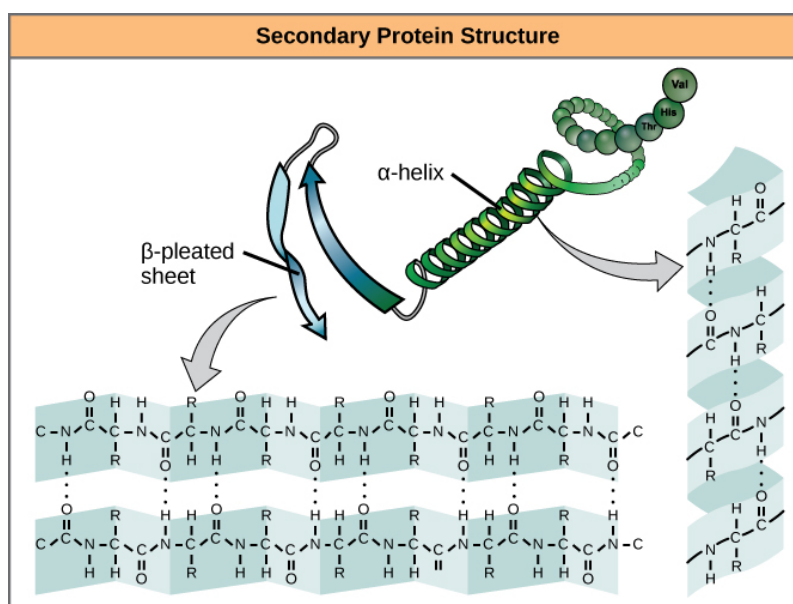
Figure 3.28 The *α*-helix and *β*-pleated sheet are secondary structures of proteins that form because of hydrogen bonding between carbonyl and amino groups in the peptide backbone. Certain amino acids have a propensity to form an *α*-helix, while others have a propensity to form a *β*-pleated sheet.

Every helical turn in an alpha helix has 3.6 amino acid residues. The polypeptide's R groups (the variant groups) protrude out from the *α*-helix chain. In the *β*-pleated sheet, hydrogen bonding between atoms on the polypeptide chain's backbone form the "pleats". The R groups are attached to the carbons and extend above and below the pleat's folds. The pleated segments align parallel or antiparallel to each other, and hydrogen bonds form between the partially positive nitrogen atom in the amino group and the partially negative oxygen atom in the peptide backbone's carbonyl group. The *α*-helix and *β*-pleated sheet structures are in most globular and fibrous proteins and they play an important structural role.

Tertiary Structure

The polypeptide's unique three-dimensional structure is its **tertiary structure** (**Figure 3.29**). This structure is in part due to chemical interactions at work on the polypeptide chain. Primarily, the interactions among R groups create the protein's complex three-dimensional tertiary structure. The nature of the R groups in the amino acids involved can counteract forming the hydrogen bonds we described for standard secondary structures. For example, R groups with like charges repel each other and those with unlike charges are attracted to each other (ionic bonds). When protein folding takes place, the nonpolar amino acids' hydrophobic R groups lie in the protein's interior; whereas, the hydrophilic R groups lie on the outside. Scientists also call the former interaction types hydrophobic interactions. Interaction between cysteine side chains forms disulfide linkages in the presence of oxygen, the only covalent bond that forms during protein folding.
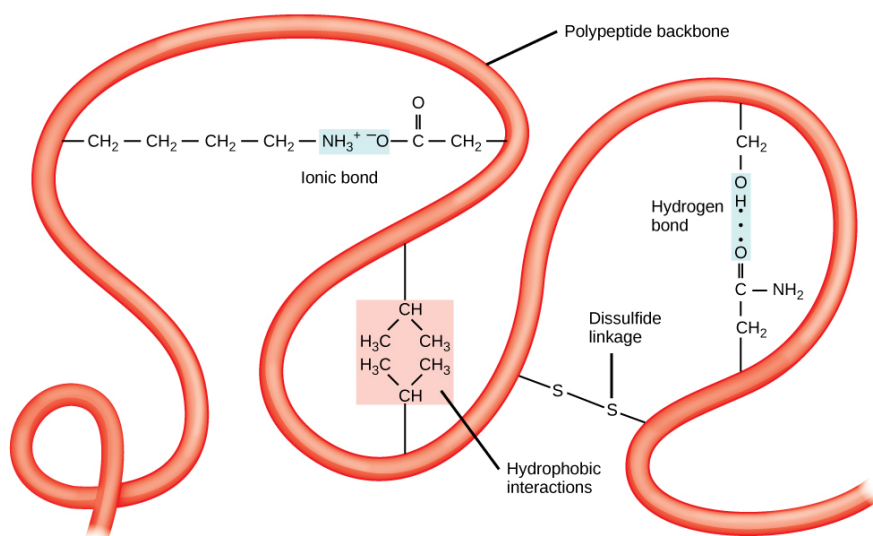
Figure 3.29 A variety of chemical interactions determine the proteins' tertiary structure. These include hydrophobic interactions, ionic bonding, hydrogen bonding, and disulfide linkages.

All of these interactions, weak and strong, determine the protein's final three-dimensional shape. When a protein loses its three-dimensional shape, it may no longer be functional.

Quaternary Structure

In nature, some proteins form from several polypeptides, or subunits, and the interaction of these subunits forms the **quaternary structure**. Weak interactions between the subunits help to stabilize the overall structure. For example, insulin (a globular protein) has a combination of hydrogen and disulfide bonds that cause it to mostly clump into a ball shape. Insulin starts out as a single polypeptide and loses some internal sequences in the presence of post-translational modification after forming the disulfide linkages that hold the remaining chains together. Silk (a fibrous protein), however, has a β-pleated sheet structure that is the result of hydrogen bonding between different chains.

Figure 3.30 illustrates the four levels of protein structure (primary, secondary, tertiary, and quaternary).
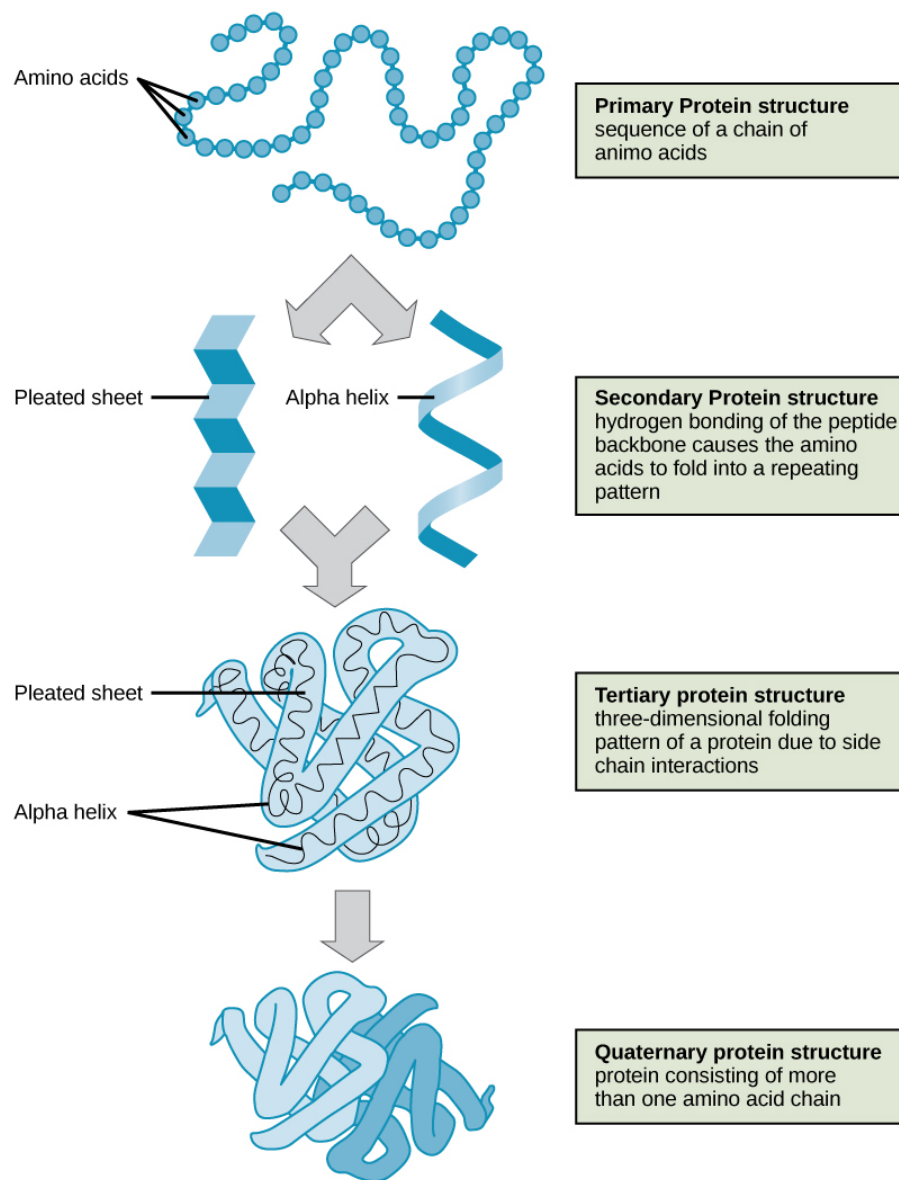
Figure 3.30 Observe the four levels of protein structure in these illustrations. (credit: modification of work by National Human Genome Research Institute)

Denaturation and Protein Folding

Each protein has its own unique sequence and shape that chemical interactions hold together. If the protein is subject to changes in temperature, pH, or exposure to chemicals, the protein structure may change, losing its shape without losing its primary sequence in what scientists call denaturation. Denaturation is often reversible because the polypeptide's primary structure is conserved in the process if the denaturing agent is removed, allowing the protein to resume its function. Sometimes denaturation is irreversible, leading to loss of function. One example of irreversible protein denaturation is frying an egg. The albumin protein in the liquid egg white denatures when placed in a hot pan. Not all proteins denature at high temperatures. For instance, bacteria that survive in hot springs have proteins that function at temperatures close to boiling. The stomach is also very acidic, has a low pH, and denatures proteins as part of the digestion process; however, the stomach's digestive enzymes retain their activity under these conditions.

Protein folding is critical to its function. Scientists originally thought that the proteins themselves were responsible for the folding process. Only recently researchers discovered that often they receive assistance in the folding process from protein helpers, or **chaperones** (or chaperonins) that associate with the target protein during the folding process. They act by preventing 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://openstaxcollege.org/l/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.