The Foundations of Biochemistry

Biochemistry asks how the remarkable properties of living organisms arise from the thousands of different lifeless biomolecules, and how these biomolecules interact to maintain and perpetuate life. The distinguished features of living organisms are:

1. A high degree of chemical complexity and microscopic organization: thousands of different molecules make up a cell's intricate internal structures. Each has its characteristic sequence of subunits, its unique three-dimensioned structure and its highly specific selection of binding partners in the cell.

2. Systems for extracting, transforming and using energy from the environment: enabling organisms to build and maintain their intricate structures and to do mechanical, chemical, osmotic, and electrical work.

3. A capacity for precise self-replication and self-assembly.

4. Mechanisms for sensing and responding to alterations in their surroundings, constantly adjusting to these changes by adapting their internal chemistry.

5. Defined functions for each of their components and regulated interactions among them.

Despite these common properties, and the fundamental unity of life they reveal, there is enormous diversity. Biochemistry describes in molecular terms the structures, mechanisms and chemical processes shared by all organisms and provides organizing principles that underlie life in all its diverse forms.

Cellular Foundations

The unity and diversity of organisms become apparent even at the cellular level. The smallest organisms consist of single cells and are microscopic. Larger, multicellular organisms contain many different types of cells, which vary in size, shape and specialized function. Despite these obvious differences all cells of the simplest and most complex organisms share certain fundamental properties, which can be seen at the biochemical level.

Cells are the Structural and Functional Units of all Living Organisms: Cells of all kinds share certain structural features. The plasma membrane defines the periphery of the cells, separating its contents from the surrounding. It is composed of lipid and protein molecules that form a thin, tough, hydrophobic barrier around the cell. The membrane is a barrier to the free passage of inorganic ions and most other charged or polar compounds. Transport proteins in the plasma membrane allow the passage of certain ions and molecules; receptor proteins transmit signals into the cell; and membrane enzymes participate in some reaction pathways. Because the individual lipids and proteins of the plasma membrane are not covalently linked, the entire structure is remarkably flexible, allowing changes in the shape and size of the cell.

The internal volume bounded by the plasma membrane, the cytoplasm, is composed of an aqueous solution, the cytosol, and a variety of suspended particles with specific functions. The cytosol is a highly concentrated solution containing enzymes and the RNA molecules that encode them; the components (amino acids and nucleotides) from which these

macromolecules are assembled; hundreds of small organic molecules called metabolites, intermediates in biosynthetic and degradative pathways; coenzymes, compounds essential to many enzymes-catalyzed reactions; inorganic ions, and ribosomes, small particles (composed of protein and RNA molecule) that are the sites of protein synthesis.

All cells have for at least part of their life, either a nucleus or a nucleoid, in which the genome- the complete set of genes, composed of DNA is stored and replicated. The nucleoid in bacteria is not separated from the cytoplasm by a membrane. The nucleus in higher organism consists of nuclear materials enclosed with a double membrane, the nuclear envelope. Cells with nuclear envelopes are called eukaryotes (Greek eu, "true" and karyon, "nucleus"); those without nuclear envelopes –bacterial cells are prokaryotes (Greek pro, "before") (Lehninger, Fig. 1-7).

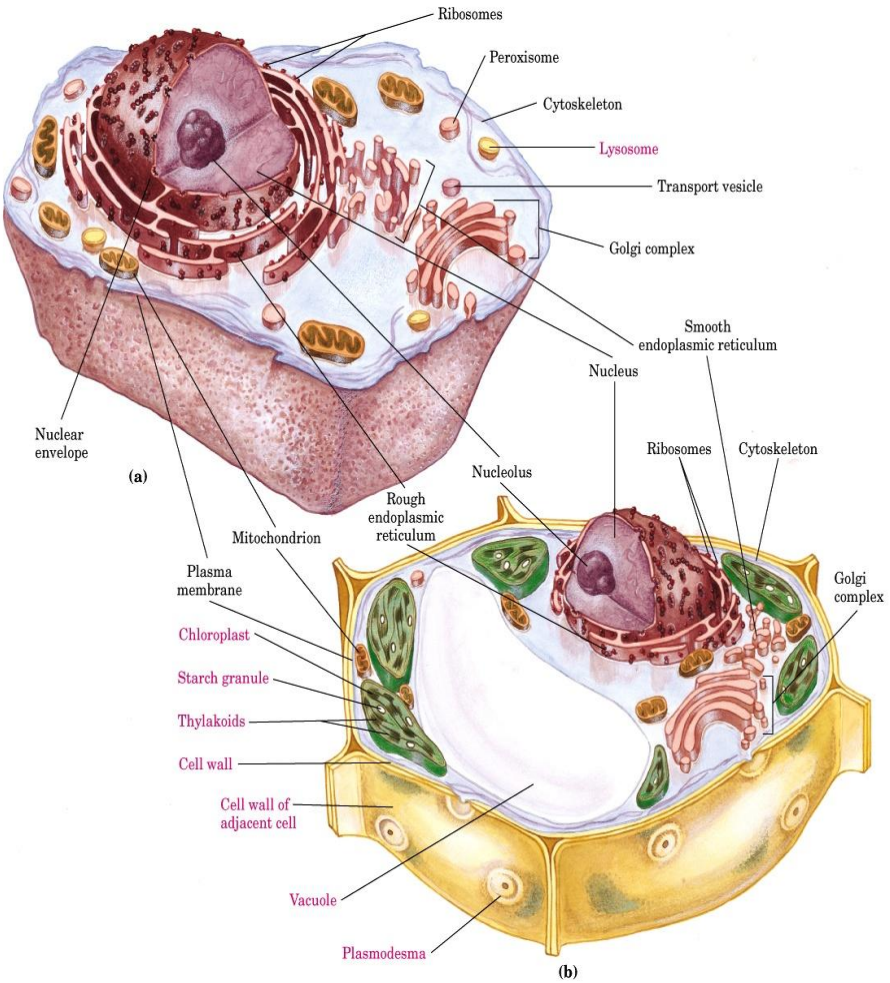


FIGURE 1-7. Eukaryotic cell structure:(a) animal cell and (b) plant cell

Classification of organisms is according to how they obtain energy and carbon they need for synthesizing cellular material (Lehninger, Fig. 1-5).

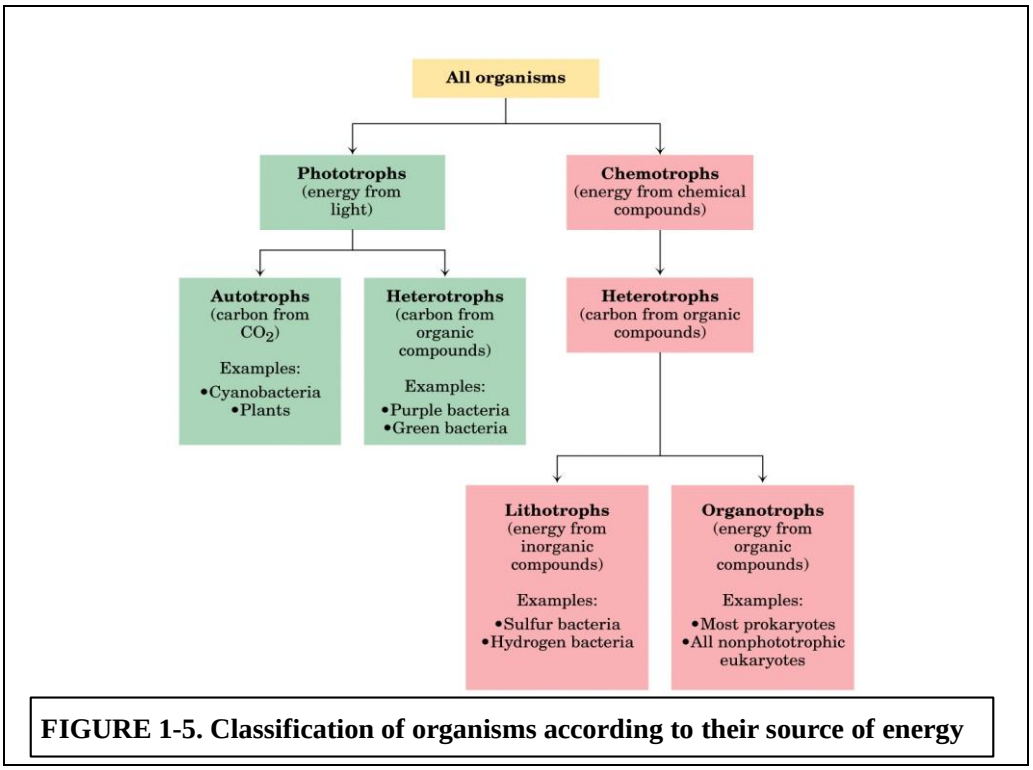


FIGURE 1-5. Classification of organisms according to their source of energy

There are two broad categories based on energy sources. Phototrophs (Greek: trophs, "nourishment") traps and use sunlight, and chemotrophs derives their energy from oxidation of a fuel. All chemotrophs require a source of organic nutrients; they cannot fix CO_2 into organic compounds. The phototrophs can be further divided into those that can obtain all needed carbon from CO_2 (autotrophs) and those that require organic nutrients (heterotrophs). The chemotrophs may be further classified according to a different criterion: whether the fuels they oxidize are inorganic (lithotrophs) or organic (organotrophs).

Most known organisms fall within one of these four broad categories –autotrophs or heterotrophs among the photosynthesizers; lithotrophs or organotrophs among the chemical oxidizers.

Membranous Organelles in Eukaryotic Cells: Eukaryotes have nucleus and a variety of membrane-bound organelles with specific functions: mitochondria, endoplasmic reticulum, Golgi complexes, and lysosomes. Plant cells also contain vacuoles and chloroplasts (Lehninger Fig. 1-7). Also present in the cytoplasm of many cells are granules or droplets containing stored nutrients such as starch and fat.

The Cytoplasm is Organized by the Cytoskeleton and is Highly Dynamic: Electron microscopy reveals several types of protein filaments crisscrossing the eukaryotic cell, forming an interlocking three dimensional meshwork, the cytoskeleton. There are three general types of cytoplasmic filaments – actin filament, microtubules and intermediate filaments differing in width, composition and specific function. All types provide structure and organization to the cytoplasm and shape to the cell. Actin filaments and microtubules also help to produce the motion of organelles or the whole cell.

Each type of cytoskeletal component is composed of simple protein subunits that polymerize to form filaments of uniform thickness. These filaments are not permanent structures; they undergo constant disassembly into their protein subunits and reassembly into filaments. Their locations in cells are not rigidly fixed but may change dramatically with mitosis, cytokinesis,

amoeboid motion or changes in cell shape. The filaments disassemble and then reassemble elsewhere. Membranous vesicles bud from one organelle and fuse with another. Organelles move in the cytoplasm along protein filaments, their motion powered by energy dependent proteins. The endomembrane system segregates specific metabolic processes and provides surfaces on which certain enzyme – catalyzed reactions occur.

Exocytosis and endocytosis, mechanisms of transport (out of and into cells respectively) that involve membrane fusion and fission, provide paths between the cytoplasm and surrounding medium allowing for secretion of substances produced within the cell and uptake of extra cellular material. The interactions between the cytoskeleton and organelles are noncovalent, reversible -and not random- and subject to regulation in response to various intracellular and extracellular signals.

Cells build Supramolecular Structures: The monomeric subunits in proteins, nucleic acids and polysaccharides are joined by covalent bonds. In supramolecular complexes, however, macromolecules are held together by noncovalent interaction much weaker individually, than covalent bonds. Among these noncovalent interactions are hydrogen bounds (between polar groups), ionic interactions (between charged groups), hydrophobic interactions (among nonpolar groups in aqueous solution) and van der Waals interaction. Fig. 1-11 (Lehninger) illustrates the structures hierarchy in cellular organization.

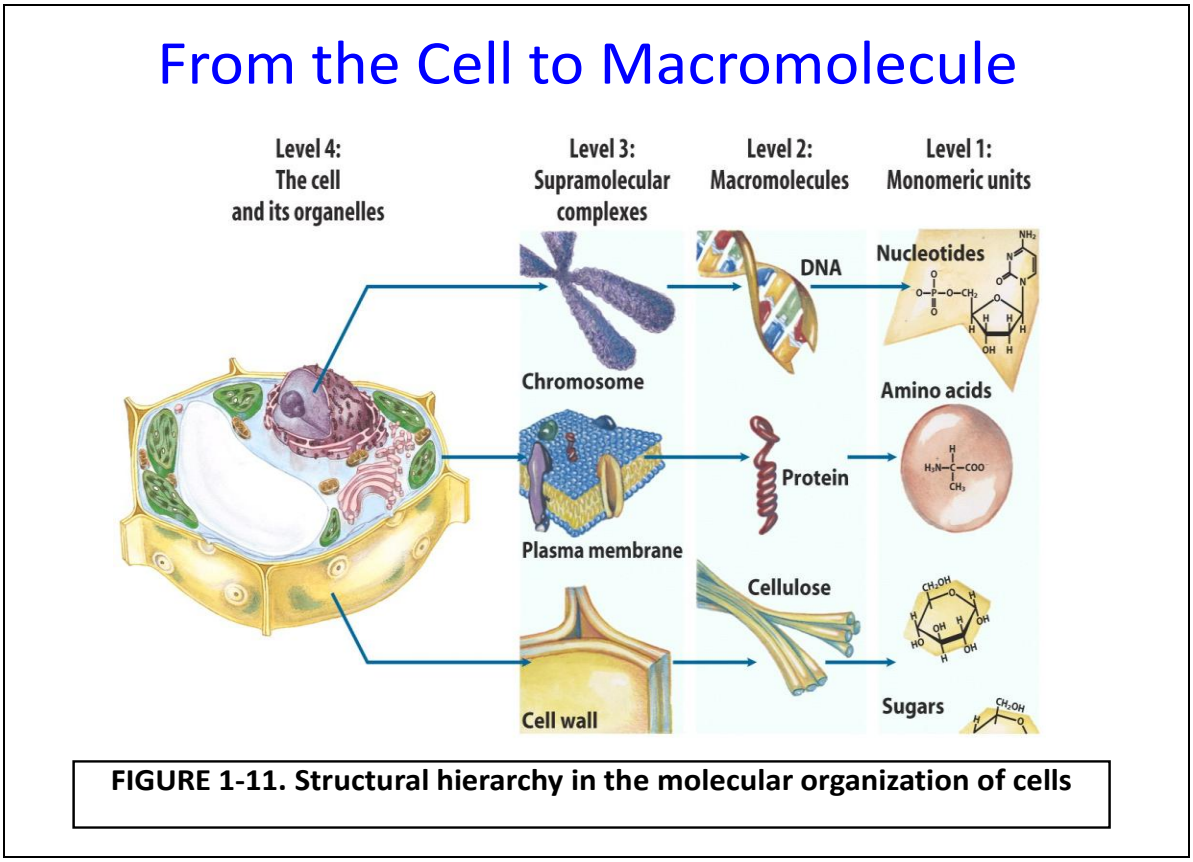


FIGURE 1-11. Structural hierarchy in the molecular organization of cells

Chemical Foundation

Biochemistry aims to explain biological form and function in chemical terms. Glucose breakdown in yeast and in animal muscle cells revealed remarkable chemical similarities (the same 10 chemical intermediates) subsequent studies of many other biochemical processes in

many different organisms have confirmed the generality of this observation, summarized by J. Monod "what is true of E. coli is true of the elephant".

Only about 30 of the more than 90 naturally occurring chemical elements are essential to organisms. The four most abundant elements in living organisms, in terms of percentage of total number of atoms, are H, O, N, C, which together make up more than 99% of the mass of most cells. They are the lightest elements capable of forming one, two, three and four bounds, respectively; in general, the lightest elements form the strongest bonds. The trace elements (Cu, Zn, Co, Ni, Fe, Mo, Se,… etc.) represent a tiny fraction of the weight of the human body, but all are essential to life, usually because they are essential to function of specific proteins, including enzymes. The oxygen–transporting capacity of the hemoglobin molecule, for example, is absolutely dependent on four iron ions that make up only 0.3% of its mass.

Biomolecules are Compounds of Carbon with a Variety of Functional Groups: Covalently linked carbon atoms in biomolecules can form linear chains, branched chains, and cyclic structure. To these carbon skeletons are added groups of other atoms called functional groups which confer specific chemical properties on the molecule.

Most biomolecules can be regarded as derivatives of hydrocarbons, with hydrogen atoms replaced by a variety of functional groups to yield different families of organic compounds. Typical of these are alcohols, which have one or more –OH groups: amines, with amino groups; aldelydes and ketones with carbonyl groups; and carboxylic acids with carboxyl groups. Many biomolecules are polyfunctional, containing two or more different kinds of functional groups.

Cells contain a Universal set of Small Molecules: Dissolved in the aqueous phase (cytosol) of all cells is a collection of 100-200 different small organic molecules (MW ~100 to ~500). This collection of molecules includes the common amino acids, nucleotides, sugars and their phosphorylated derivatives and a number of mono– di– and tricarboxylic acids. They are trapped within the cell because the plasma membrane is impermeable to them.

Macromolecules are the Major Constituents of Cells: Many biological molecules are macromolecules, polymers of high M_r assembled from relatively simple precursors. Proteins, nucleic and polysaccharides are produced by polymerization of relatively small compounds with M_r of 550 or less. The number of polymerized units can range from ten to millions. Macromolecules themselves may be further assembled into supramolecular complexes, forming functional units such as ribosomes.

Proteins, long polymers of amino acids, constitute the largest fraction (besides water) of cells. Some proteins have catalytic activity and function as enzymes; others serve as structural elements, signal receptors or transporters that carry specific substances into or out of cells. Proteins are perhaps the most versatile of all biomolecules. The nucleic acids DNA and RNA are polymers of nucleotides. They store and transmit genetic information. The polysaccharides, polymers of simple sugars such as glucose, have two major functions as energy-yielding fuel stores and as extracellular structural elements with specific binding sites for particular proteins. Shorter polymers of sugar (oligosaccharides) attached to proteins of lipids at the cell surface serve as specific cellular signals. The lipids serve as structural components of membranes, energy-rich fuel stores, pigments and intracellular signals.

Proteins and nucleic acids are informational macromolecules: each protein and each nucleic acid has a characteristic information-rich subunit sequences.

Three–Dimensional Structures is Described by Configuration and Conformation: The covalent bonds and functional groups of a bimolecule are of course, central to its function, but so also is the arrangement of the molecules constituent atoms in three dimensional space -its stereochemistry. A carbon–containing compound commonly exists as stereoisomers, molecules with the same chemical bonds but different stereochemistry -that is, different configuration, the fixed spatial arrangement of atoms.

Configuration is conferred by the presence of either (1) double bonds, around which there is no freedom of rotation, or (2) chiral center around which substituent groups are arranged in a specific sequence. The identifying characteristic of configurational isomers is that they can't be interconnected without temporarily breaking one ore more covalent bonds. Figure 1-18 (Lehninger) shows the configurations of maleic acid and its isomer fumaric acid.

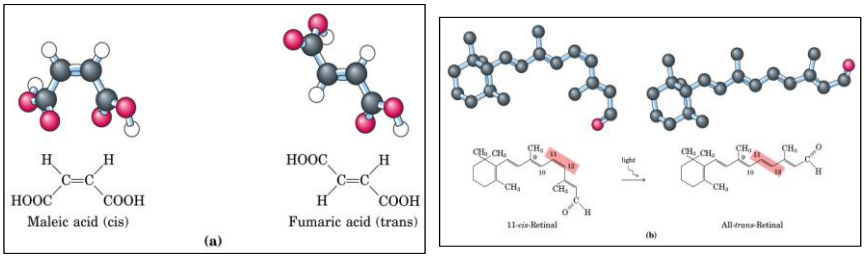


FIGURE 1-18. configurations of geometric isomers

These compounds are geometric or cis-trans isomers; they differ in the arrangement of their substituent groups with respect to the nonrotating double bond (Latin cis, "on this side" groups on the same side of the double bond; trans "across" groups on the opposite side). Maleic acid is the cis isomer and fumaric acid the trans isomer; each is a well-defined compound that can be separated from the other and each has its own unique chemical properties.

In the second type of configurational isomer, four different substituent bonded to a tetrahedral carbon atom may be arranged two different ways in space -that is- have two configuration (Fig 1-19) yielding two stereoisomers with similar or identical chemical properties but differing in certain physical and biological properties. A carbon atom with four different substituents is said to be asymmetric and asymmetric carbons are called chiral center (Greek chiros, "hand"; some stereoisomers are related structurally as the right hand is to the left). A molecule with only one chiral carbon can have two stereo isomers; when two or more (n) chiral carbons are presented, where can be 2^n stereoisomers. Some stereoisomers are mirror images of each other; they are called enantiomers (Lehninger, Fig. 1-19). Pairs of stereoisomer that are not mirror images of each other are called diastereomers (Lehninger, Fig. 1-20).

Enantiomers have nearly identical chemical properties but differ in a characteristic physical property, their interaction with plane-polarized light. In separate solutions, two enantiomers rotate the plane of polarized light in opposite directions, but an equimolar solution of the two enantiomers (a racemic mixture) shows no optical rotation. The most useful system of nomenclature is the RS system. In this system, each group attached to a chiral carbon is assigned a priority. The priorities of some common substations are:

$$—OCH_2 \rangle —OH \rangle —NH_2 \rangle —COOH \rangle —CHO \rangle —CH_2OH \rangle —CH_3 \rangle —H$$

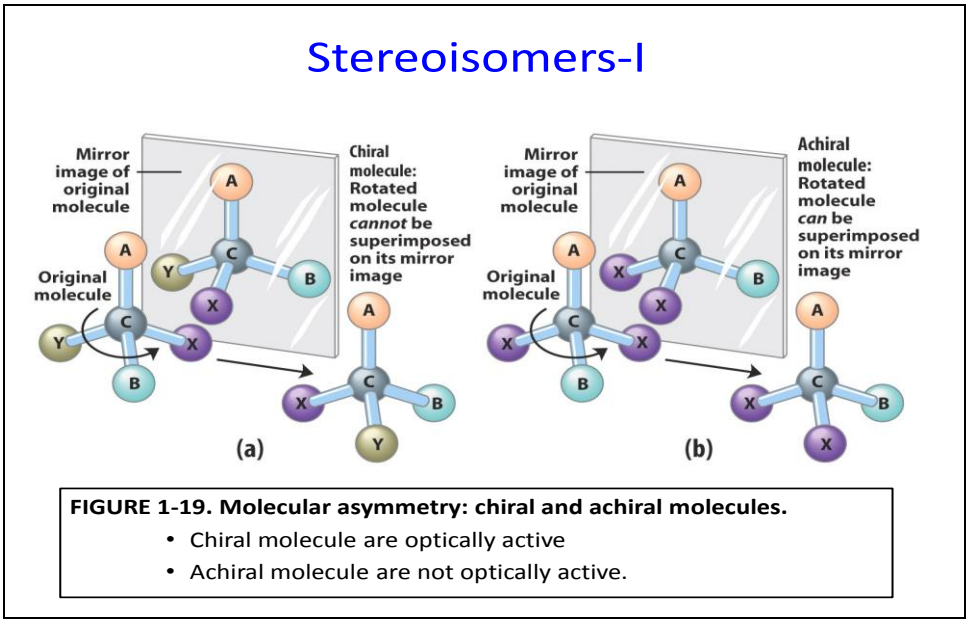


FIGURE 1-19. Molecular asymmetry: chiral and achiral molecules.

- Chiral molecule are optically active
- Achiral molecule are not optically active.

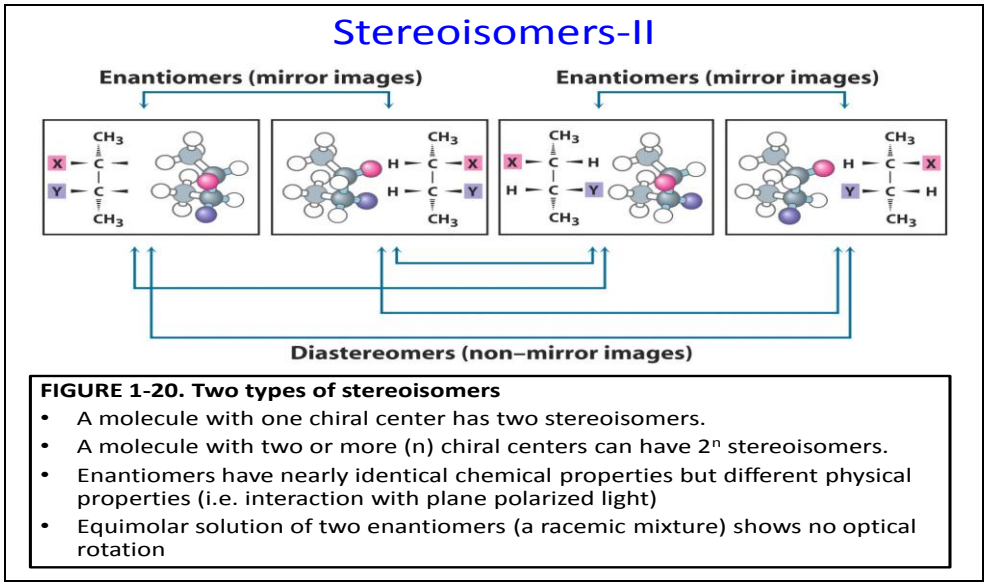


FIGURE 1-20. Two types of stereoisomers

- A molecule with one chiral center has two stereoisomers.
- A molecule with two or more (n) chiral centers can have 2^n stereoisomers.
- Enantiomers have nearly identical chemical properties but different physical properties (i.e. interaction with plane polarized light)
- Equimolar solution of two enantiomers (a racemic mixture) shows no optical rotation

Another naming system for stereoisomers is the D and L system (absolute configuration). The absolute configurations of simple sugars and amino acids are based on the absolute configuration of the three- carbon sugar glyceraldehyde. This fisher's convention L and D refer only to the absolute configuration of the four substituents around the chiral carbon, not optical properties of the molecule. "l" and "d" designation are used for levorotatory (rotating light to the left) and dextrorotatory (rotating light to the right).

Distinct form configuration is molecular conformation. The spatial arrangement of substituent groups that without breaking any bonds, are free to assume different positions in space because of the freedom of rotation about single bonds.

Interactions Between Biomolecules are Stereospicific: Biological interactions between molecules are stereospecific: the three-dimensional structure of biomolecules is of utmost importance in their biological interactions; reactants with enzymes, hormone with its receptor on a cell surface, antigen with its specific antibody. In living organisms chiral molecules are usually present in only one of their chiral forms. For example, the amino acids in proteins occur only as their L isomers; glucose occurs only as its D isomers.

Physical Foundations

Living cells and organisms must perform work to stay alive and to reproduce themselves. The synthetic reactions that occur within cells require the input of energy. Cells have developed highly efficient mechanisms for coupling the energy obtained from sunlight or fuels to the many energy-consuming processes they must carry out. We can consider cellular energy conversions like all other energy conversions in the context of the laws of thermodynamics.

Living Organisms Exist in a Dynamic Steady State Never at Equilibrium with their Surrounding: Although the characteristic composition of an organism changes little through time, the population of molecules within the organism is far from static. Small molecules, macromolecules and supramolecular complexes are continuously synthesized and then broken down in chemical reactions that involve a constant flux of mass and energy through the system. For example, the amount of hemoglobin and glucose in the blood remain nearly constant because the rate of synthesis or intake of each just balances the rate of its breakdown, consumption or conversion into some other products. The constancy of concentration is the result of a dynamic steady state, a steady state that is far from equilibrium. Maintaining the steady state requires the constant investment of energy; when the cell can no longer generate energy, it dies and begins to decay toward equilibrium with its surroundings.

Organisms Transform Energy and Matter from Their Surroundings: If the system exchanges neither matter nor energy with its surroundings, it is said to be isolated. If the system exchanges energy but not matter with its surrounding, it is a closed system; if it exchanges both energy and matter with its surroundings it is an open system.

A living organism is an open system; it exchange both matter and energy with its surroundings. Living organisms derive energy from their surroundings in two ways. First, They take up chemical fuels (such as glucose) from the environment and extract energy by oxidizing them; and second they absorb energy from sunlight.

The first law of thermodynamics developed from physics and chemistry but fully valid for biological systems as well, describes the principle of the conservation of energy; in any physical or chemical change, the total amount of energy in the universe remains constant although the form of energy may change. Cells are capable of interconnecting chemical, electromagnetic, mechanical, and osmotic energy with greet efficiency (Lehninger, Fig. 1-24).

The Flow of Electrons Provides Energy for Organisms: Nearly all living organisms derive their energy directly or indirectly from the radiant energy of sunlight. Photosynthetic cells absorb light energy and use it to drive electrons from water to CO_2, forming energy-rich products such as glucose ($C_6H_{12}O_6$), starch and sucrose and releasing O_2 into the atmosphere:

$$6CO_2 + 6H_2O \xrightarrow{\text{Light}} C_6H_{12}O_6 + 6O_2$$

"Light-driven reduction of CO_2".

Nonphotosynthetic cells and organisms obtain the energy they need by oxidizing the energy-rich products of photosynthesis and then passing electrons to atmospheric O_2 to form water CO_2 and other end products which are recycled in the environment:

$$C_6H_{12}O_6 + 6O_2 \longrightarrow 6CO_2 + 6H_2O + \text{energy}$$

"Energy-yielding oxidation of glucose"

All these reactions involving electron flow are oxidation reduction reactions: one reactant is oxidized (loses electrons) as another is reduced (gains electrons).

Creating and Maintaining Order Requires Work and Energy: DNA, RNA, and proteins are informational macromolecules in addition to using chemical energy to form covalent bonds between the subunits in these polymers; the cell must invest energy to order the subunits in their correct sequence. This would represent increased order in a population of molecules; but according to the second law of thermodynamics, the tendency in nature is toward ever-greater disorder in the universe: the total entropy of the universe is continually increasing. To bring about the synthesis of macromolecules from their monomeric units, free energy must be supplied to the system (in this case the cell).

The randomness or disorder of the components of a chemical system is expressed as **entropy, S**. Any change in randomness of the system is expressed as entropy change ΔS, which by convention has a positive value when randomness increases. J. Gibbs who developed the theory of energy change during chemical reactions showed that the **free energy content, G**, of any closed system can be defined in terms of three quantities: **enthalpy, H**, reflecting the number and kinds of bonds; entropy S; and the absolute temperature, T (in degrees Kelvin). The definition of free energy is:

$$\mathbf{G = H - TS}$$

When a chemical reaction occurs at constant temp., the free-energy change ΔG:

$$\mathbf{\Delta G = \Delta H - T\,\Delta S}$$

A process tends to occur spontaneously only if ΔG is negative. To carry out thermodynamically unfavorable energy-requiring (endergonic) reactions like protein and nucleic acid formation, cells couple them to other reactions that liberate free energy (exergonic reactions), so that the overall process is exergonic: the sum of the free energy changes is negative.

In closed systems, chemical reactions proceed spontaneously until equilibrium is reached. When a system is at equilibrium the rate of product formation exactly equals the rate at which product is converted to reactant. Thus there is not net change in the concentrations of reactants and products; a steady-state is achieved. The energy change as the system moves from its initial state to equilibrium, with no changes in temperature or pressure, is given by the free-energy change, **ΔG.**

In living organisms, as in the mechanical example in Fig. 1-26a (Lehninger), an exergonic reaction can be coupled to an endergonic reaction to drive otherwise, unfavorable reactions. Fig. 1-26b (Lehninger) illustrates this principle for the conversion of glucose to glucose 6-phosphate, the first step in the pathway for oxidation of glucose.

Glucose + Pi ⟶ **Glucose 6-phosphate** reaction 1

"ΔG_1 is positive"

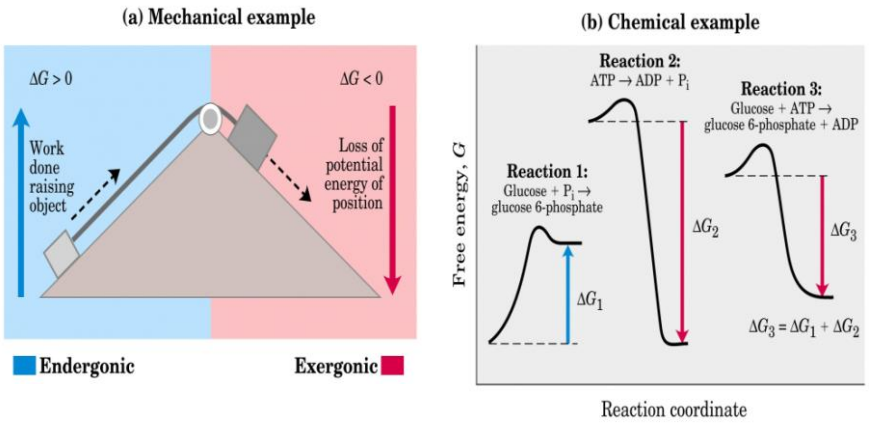


FIGURE 1-26. Energy coupling in mechanical and chemical processes

This reaction does not occur spontaneously; ΔG is positive. A second, very exergonic reaction can occur is all cells:

ATP ⟶ **ADP + Pi** reaction 2

"ΔG_2 is negative"

The two reactions share a common intermediate Pi, which is consumed in reaction 1 and produced in reaction 2. The two reactions can therefore be coupled in the form of a third reaction.

Glucose + ATP ⟶ **Glucose 6-phosphate + ADP** reaction 3

As we shall see, the breakdown of ATP is the exergonic reaction that drives many endergonic processes in cells. In fact, ATP is the major carrier of chemical energy in all cells.

K_{eq} and ΔG^o are Measures of a Reaction Tendency to Proceed Spontaneously: The tendency for a chemical reaction to go to completion can be expressed as an equilibrium constant:

$$\mathbf{aA + bB \longrightarrow cC + dD}$$

$$Keq = \frac{[Ceq]^c [Deq]^d}{[Aeq]^a [Beq]^b}$$

Where (Aeq) is the concentration of A and so on, when the system has reached equilibrium Gibbs showed that ΔG for any chemical reaction is a function of the standard free energy change ΔG^o -a constant that is characteristic of each specific reaction –and a term that expresses the initial concentrations of reactants and products:

$$\Delta G = \Delta G^o + RT \ln \frac{[Ci]^c [Di]^d}{[Ai]^a [Bi]^b}$$

Where [Ai] is the initial concentration of A and so forth; R is the gas constant; and T is the absolute temperature when a reaction has reached equilibrium, no driving force remains and it can do no work: $\Delta G = 0$. For this special case, [Ai] = [Aeq] …..

$$\mathbf{\Delta G^o = - RT \, ln \, k_{eq}} \qquad \mathbf{\Delta G \text{ joules/mole}}$$

Enzymes Promote Sequences of Chemical Reactions

The part from reactant(s) to product(s) involves an energy barrier; called the activation energy (Lehninger, Fig. 1-27), that must be surmounted (تجاوزها) for any reaction to proceed. The breaking of existing bonds and formation of new ones generally requires, first, the distortion of existing bonds, creating a transition state of higher free energy than either reactants or products. The highest point represents the transition state, and the difference in energy between the reactant in its ground state and in its transition state is the activation energy $\Delta G^{\ddagger}$. An enzyme catalyzes a reaction by providing a more comfortable fit for the transition state: a surface that complements the transition state in stereochemistry, polarity and charge. The binding of enzyme to the transition state is exergonic and the energy released by this binding reduces the activation energy for the reaction and greatly increases the reaction rate.

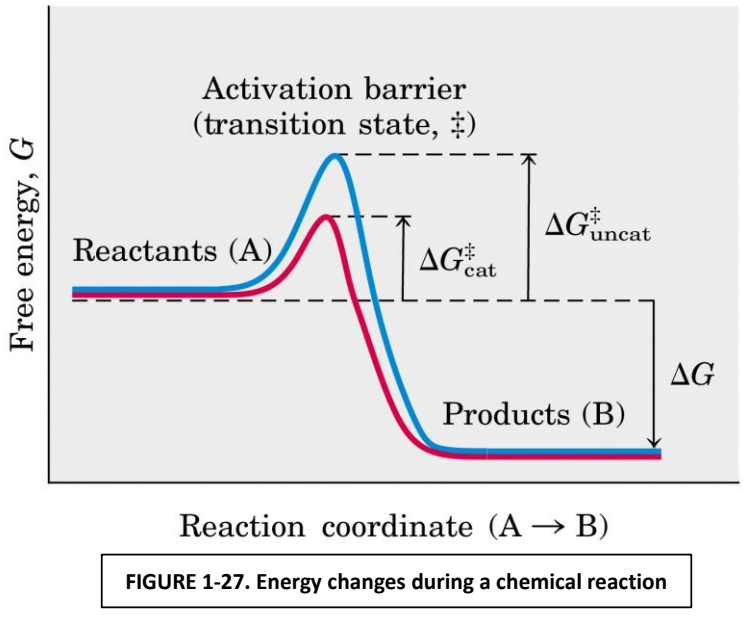


FIGURE 1-27. Energy changes during a chemical reaction

A further contribution to catalysis occurs when two or more reactants bind to the enzyme's surface close to each other with stereospecific orientations that favors the reaction. Enzyme-catalyzed reactions commonly proceed at rates greater than 10^{12} times faster that the uncatalyzed reactions.

Cellular catalysts are with a few exceptions, proteins. Again with a few exceptions each enzyme catalyzed a specific reaction and each reaction in a cell is catalyzed by a different enzyme. Thousands of different enzymes are therefore required by each cell. The multiplicity of enzymes, their specificity, and their susceptibility to regulation give cells the capacity to lower activation barriers selectively. This selectivity is crucial for the effective regulation of cellular processes. By allowing specific reactions to proceed at significant rates at particular times, enzymes determine how matter and energy are channeled into cellular activities.

The thousands of enzyme-catalyzed chemical reactions in cells are functionally organized into many sequences of consecutive reactions called pathways in which the product of one reaction becomes the reactant in the next. Some pathways degrade organic nutrients into simple end products in order to extract chemical energy and convert it into a form useful to the cell; together these degradation, free-energy-yielding reactions are designated catabolism. Other pathways start with small precursor molecules and convert them to progressively lager and more complex molecules, including proteins and nucleic acids. Such synthetic pathways require the inputs of energy are collectively designated anabolism. The overall network of enzyme catalyzed pathways constitute cellular metabolism. ATP is the major connecting link between the catabolic and anabolic components of this network (Lehninger, Fig. 1-28).

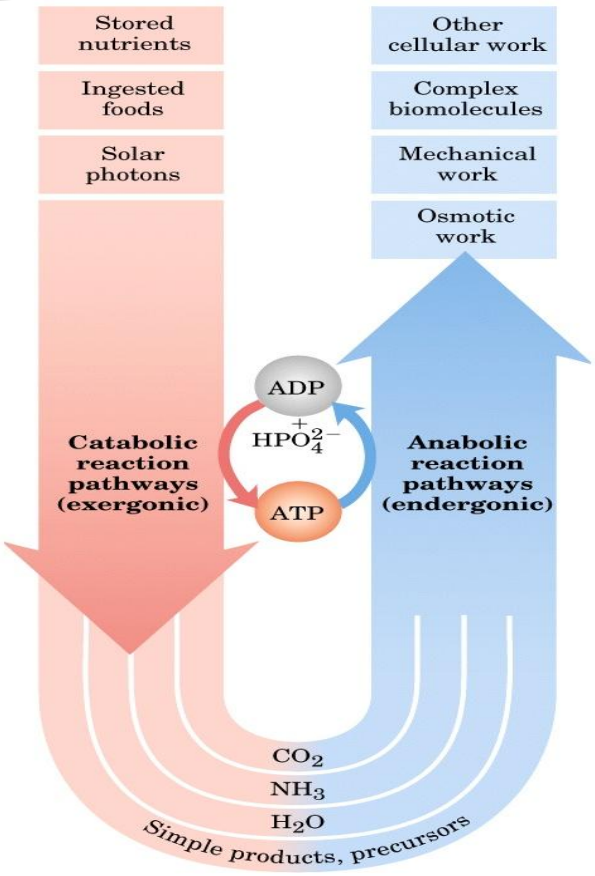
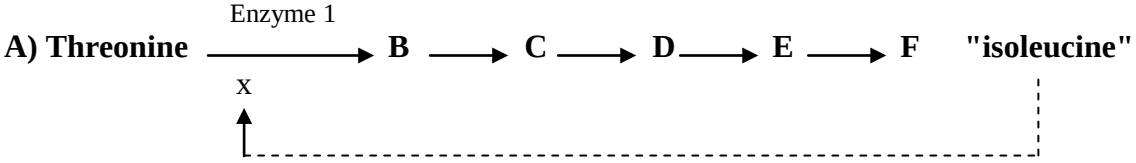


FIGURE 1-28. The central role of ATP in metabolism

Metabolism is Regulated to Achieve Balance and Economy

Not only do living cells simultaneously synthesize thousands of different kinds of carbohydrate, fat, protein and nucleic acid molecules and their simpler subunits, but they do so in the precise proportions required by the cell under any given circumstances. This is achieved through the regulation of key enzymes in the metabolic pathways. Consider the pathway that leads to the synthesis of the amino acid isoleucine in E. coli:



If a cell begins to produce more isoleucine than is needed for protein synthesis, the unused isoleucine accumulates and the increased concentration inhibits the catalytic activity of the first enzyme in the pathway, immediately slowing the production of isoleucine. Such feedback inhibition keeps the production and utilization of each metabolic intermediate in balance.

Genetic Foundation

Perhaps the most remarkable property of living cells and organisms is their ability to reproduce themselves for countless generation with nearly perfect fidelity. The capacity of living cells to preserve their genetic material and to duplicate it for the next generation results from the structural complementarity between the two haves of the DNA molecule (Lehninger, Fig. 1-30). The basic unit of DNA is a linear polymer of four different monomeric subunits, deoxyribonucleotides, arranged in a precise linear sequence.

It is this linear sequence that encodes the genetic information. Before the cell divides, the two DNA strands separate and each serves as a template for the synthesis of a new complementary strand, generating two identical double helical molecules, one for each daughter cell. If one strand is damaged, continuity of information is assured by the information present in the other strand which acts as a template for repair of the damage.

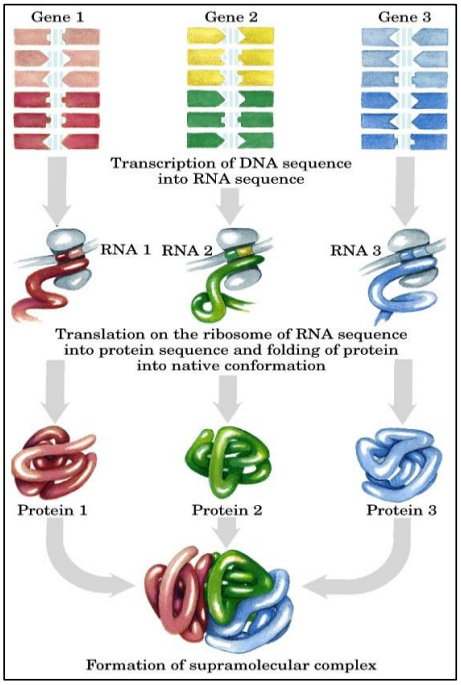


FIGURE 1-31. DNA to RNA to Protein

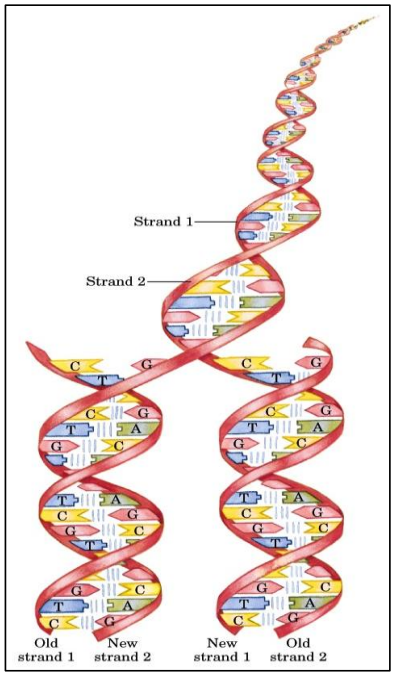


FIGURE 1-30. Complementarity between the two strands of DNA

The Linear Sequence of DNA Encodes Proteins with Three Dimensional Structures:

A linear sequence of deoxyribonucleotides in DNA codes (through an intermediary, RNA) for the production of a protein with a corresponding linear sequence of amino acids (Lehninger, Fig. 1-31). The protein folds into a particular three-dimensional shape,

determined by its amino acid sequence and stabilized primarily by noncovalent interaction. The precise three-dimensional structure or **native conformation** of the protein is crucial to its function. Once in its native conformation, a protein may associate noncovalently with other proteins or with nucleic acids or lipids, to form supramolecular complexes such as chromosomes, ribosomes and membranes.

Water

Water is the most abundant substance in living systems, making up 70% or more of the weight of most organisms. The attractive forces between water molecules and the slight tendency of water to ionize are of crucial importance to the structure and function of biomolecules. We review the topic of ionization in terms of equilibrium constant, pH, and titration curves and consider how aqueous solutions of weak acids or bases and their salts act as buffers against pH changes in biological systems. The water molecule and its ionization products H^+ and OH^-, profoundly influence the structure, self assembly and properties of all cellular components, including proteins, nucleic acids and lipids. The noncovalent interactions among biomolecules are influenced by the solvent properties of water including its ability to form hydrogen bonds with itself and with solutes.

Water Interactions in Aqueous Systems

Hydrogen bonds between water molecules provide the cohesive forces that make water a liquid at room temperature. Polar biomolecules dissolve readily in water because they can replace water-water interactions with more energetically favorable water–solute interaction. In contrast non polar biomolecules interfere with water-water interactions but are unable to form water-solute interactions -consequently non polar molecules are poorly soluble in water, in aqueous solutions non polar molecules tend to cluster together.

Hydrogen bonds and ionic, hydrophobic and van der Waals interactions are individually weak, but collectively they have a very significant influence on the three dimensional structures of proteins, nucleic acids, polysaccharides and membrane lipids.

Hydrogen – Bonding Gives Water its Unusual Properties

Water has a higher melting point, boiling point and heat of vaporization than most other solvents. These unusual properties are a consequence of attractions between adjacent water molecules that give liquid water great internal cohesion. The oxygen nucleus attracts electrons more strongly that does the hydrogen nucleus (a proton); that is oxygen is more electronegative. The sharing of electrons between H and O is therefore unequal; the electrons are more often in the vicinity of the oxygen atom than of the hydrogen. The result of this unequal electron sharing is two electric dipoles in the water molecule, one along each of H-O bonds; each hydrogen bears a partial positive charge (δ^+) and the oxygen atom bears a partial negative charge equal to the sum of the two partial positives ($2\delta^-$). As a result, there is an electrostatic attraction between the oxygen atom of one water molecule and the hydrogen of another called a hydrogen bond (Lehninger, Fig. 2-1).

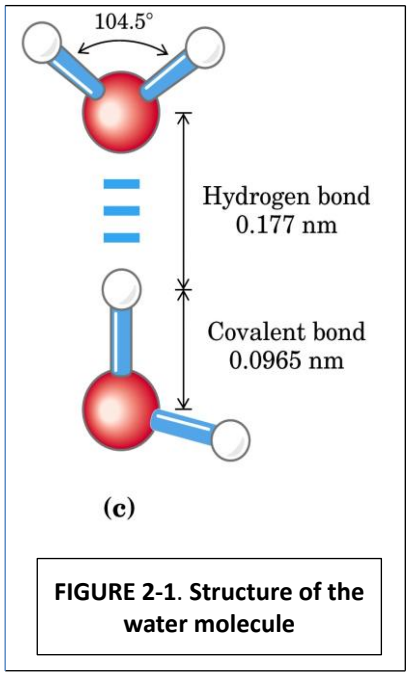


FIGURE 2-1. Structure of the water molecule

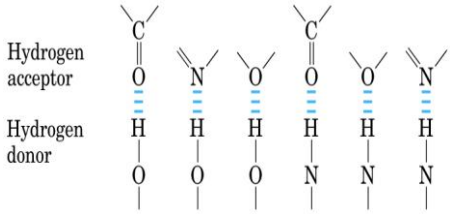


FIGURE 2-3. Common hydrogen bonds in biological systems

Water forms Hydrogen Bonds with Polar solutes

Hydrogen bonds are not unique to water. They readily form between an electronegative atom (the hydrogen acceptor, usually O or N with a lone pair of electrons) and a hydrogen atom covalently bonded to another electronegative atom (the hydrogen donor) in the same or another molecule (Lehninger, Fig. 2-3). Hydrogen atoms covalently bonded to carbon atoms do not participate in hydrogen bonding, because carbon is only slightly more electronegative than hydrogen and thus the C-H bonds is only very weakly polar. Alcohols, aldehydes, ketones and compounds containing N-H bonds all form hydrogen bonds with water molecules (Lehninger, Fig. 2-4) and tend to be soluble in water.

Water Interacts Electrostatically with Charged Solutes

Water is a polar solvent, it readily dissolves most biomolecules, which are generally charged or polar compounds (Lehninger, Table 2-2); compounds that dissolve easily in water are **hydrophilic** (Greek "water loving"). In contrast, non polar solvents such as chloroform and benzene are poor solvents for polar biomolecules but easily dissolve those that are **hydrophobic** non polar molecules such as lipids and waxes.

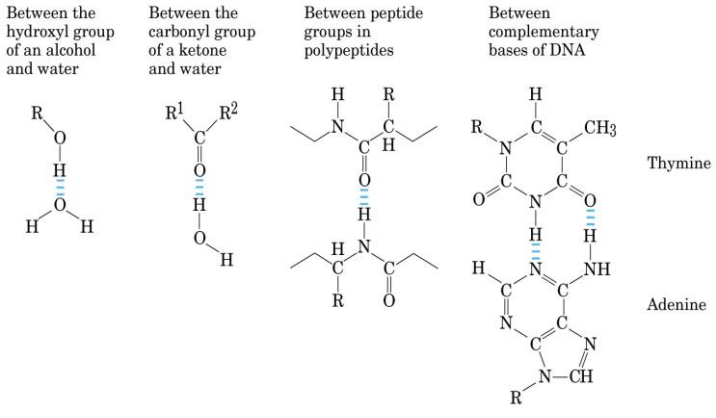


FIGURE 2-4. some biologically important hydrogen bonds

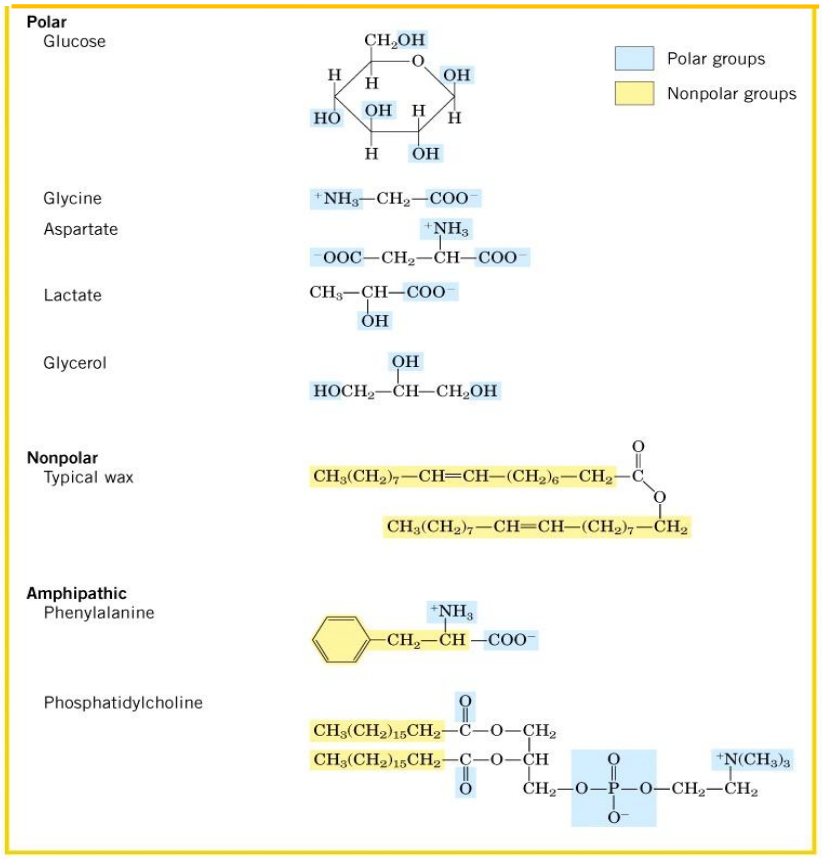
Water dissolves salts such as NaCl by hydrating and stabilizing the Na^+ and Cl^- ions, weakening the electrostatic interactions between them. The same factors apply to charged biomolecules, compounds with functional groups such as ionized carboxylic acids ($-COO^-$), protonated amines ($-NH_4^+$) and phosphate esters or anhydrides. Water readily dissolves such compounds by replacing solute-solute hydrogen bonds with solute-water hydrogen bonds.

As a salt such as NaCl dissolves, the Na^+ and Cl^- ions acquire far greater freedom of motion. The resulting increase in entropy (randomness) of the system is largely responsible for the ease of dissolving salts such as NaCl in water. In thermodynamic terms, formation of the solution occurs with a favorable free energy change:

$$\Delta G = \Delta H - T\Delta S$$

Where ΔH has a small positive value and $T\Delta S$ a large positive value; thus ΔG is negative.

Table 2-2. Some Examples of Polar, Nonpolar, and Amphipathic Biomolecules (shown as ionic forms at pH 7)



Hydrophobic Interactions

Amphipathic compounds contain regions that are polar (or charged) and regions that are non polar. The non polar regions of the molecules cluster together to present the smallest hydrophobic area to the aqueous solvent, and the polar regions are arranged to maximize their interaction with the solvent (Lehninger, Fig. 2-7b). These stable structures of

amphipathic compounds in water, called **micelles**, may contain hundreds of thousands of molecules. The forces that hold the nonpolar regions of the molecules together are called **hydrophobic interactions**. The strength of hydrophobic interactions is not due to any intrinsic attraction between nonpolar moieties. Rather, it results from the system's achieving greatest thermodynamic stability by minimizing the number of ordered water molecules required to surround hydrophobic portions of the solute molecules. Hydrophobic interactions among lipids, and between lipids and proteins, are the most important determinants of structure in biological membranes.

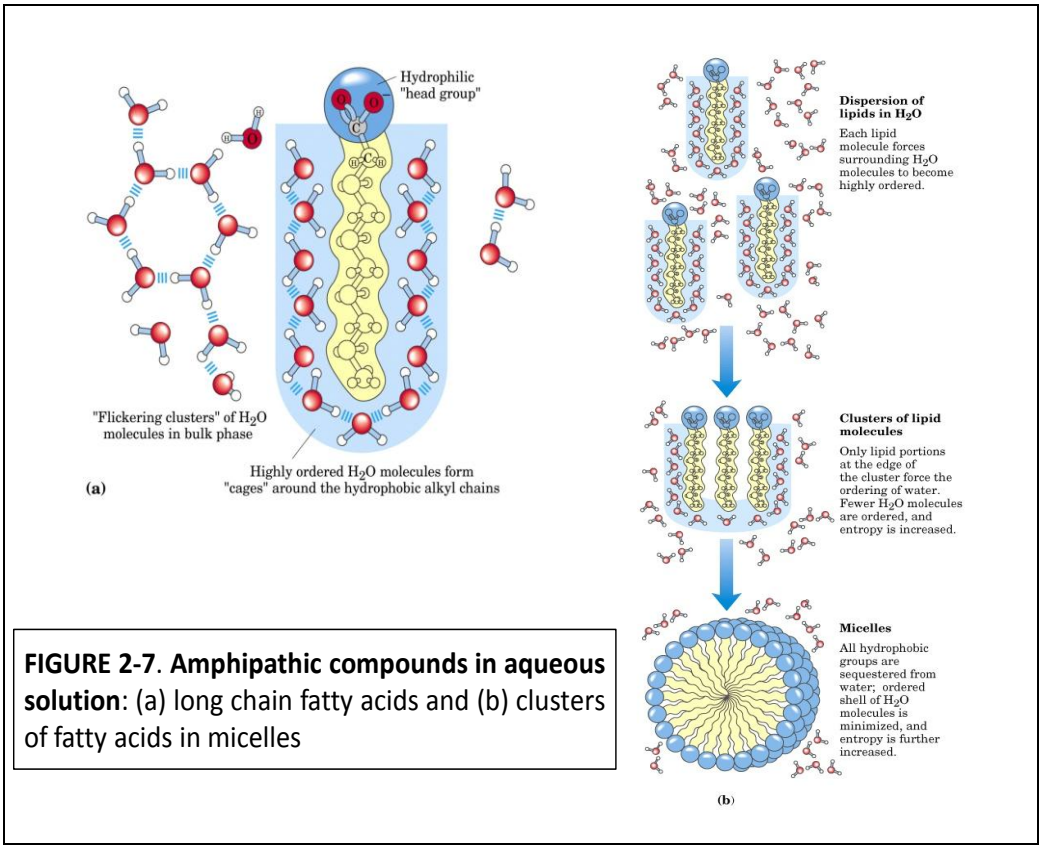


FIGURE 2-7. Amphipathic compounds in aqueous solution: (a) long chain fatty acids and (b) clusters of fatty acids in micelles

Hydrophobic interaction between nonpolar amino acids also stabilizes the three-dimensional structures of proteins. Part of the driving force for binding of a polar substrate (reactant) to the complementary polar surface of an enzyme is the entropy increase as the enzyme displaces, ordered water from the substrate (Lehninger, Fig. 2-8)

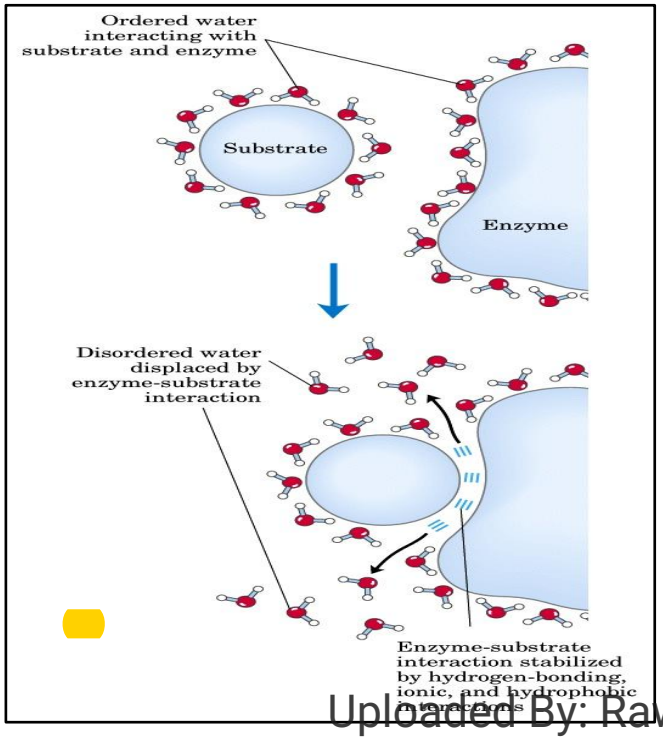


FIGURE 2-8. Release of ordered water favors formation of an enzyme-substrate complex

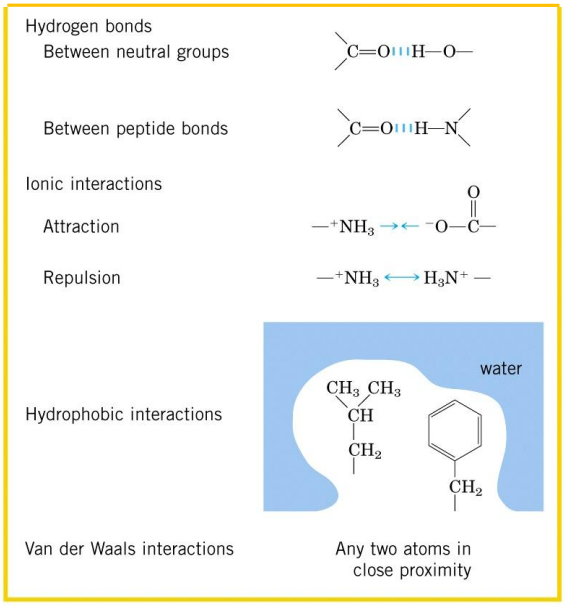
Van der Waals Interactions

When two uncharged atoms are brought very close together, their surrounding electron clouds influence each other. Random variations in the position of the electrons around one nucleus may create a transient, opposite electric dipole in the nearby atom. The two dipoles weakly attract each other, bring the two nuclei closer. These weak attractions are called Van der Waals interaction.

Weak Interactions are Crucial to Macromolecular Structure and Function

The noncovalent interactions (hydrogen bonds and ionic, hydrophobic and Van der Waals interactions) (Lehninger, Table 2-5) are much weaker than covalent bonds. However, the cumulative effect of many such interactions can be very significant. For example, the non covalent binding of an enzyme to its substrate may involve several hydrogen bonds and one or more ionic interaction, as well as hydrophobic and Van der Waals interaction. The formation of each of these weak bonds contributes to a net decrease in the free energy of the system.

TABLE 2-5. Four Types of Noncovalent ("weak") interactions among Biomolecules in aqueous Solvent



Macromolecules such as proteins, DNA and RNA contain so many sites of potential hydrogen bonding or ionic, Van der Waals, or hydrophobic interaction that the cumulative effect of the many small binding forces can be enormous. For macromolecules, the most stable (that is the native) structure is usually that in which weak-binding possibilities are maximized.

The folding of a single polypeptide or polynucleotide chain into its three-dimensional shape is determined by this principle. The binding of an antigen to a specific antibody depends on the cumulative effects of many weak interactions. As noted earlier, the energy released when an enzyme binds non covalently to its substrate is the main source of the enzyme's catalytic power, the binding of a hormone or a neurotransmitter to its cellular receptor protein is the result of weak interactions.

Osmosis

Water movement across a semipermeable membrane driven by differences in osmotic pressure, is an important factor in the life of most cells. Plasma membranes are more permeable to water than to most other small molecular ions and macromolecules. This permeability is due partly to simple diffusion of water through the lipid bilayer and partly to protein channels (aquaporins) in the membrane that selectively permit the passage of water. Solutions of equal osmolarity are said to be isotonic. Surrounded by an isotonic solution, a cell neither gains nor loses water (Lehninger, Fig. 2.13). In a hypertonic solution, one with higher osmolarity that the cytosol, the cell shrinks as water flows out. In a hypotonic solution, with lower osmolarity than the cytosol, the cell swells as water enters. In their natural environments, cells generally contain higher concentration of biomolecules and ions than water into the cell, this would eventually cause bursting of the cell (osmotic lysis).

Several mechanisms have evolved to prevent this catastrophe. Blood plasma and interstitial fluid (the extracellular fluid of tissues) are maintained at an osmolarity close to that of the cytosol. The high concentration of albumin and other proteins in blood plasma contributes to its osmolarity. Cells also actively pump out ions such as Na^+ into the interstitial fluid to stay in osmotic balance with their surroundings.

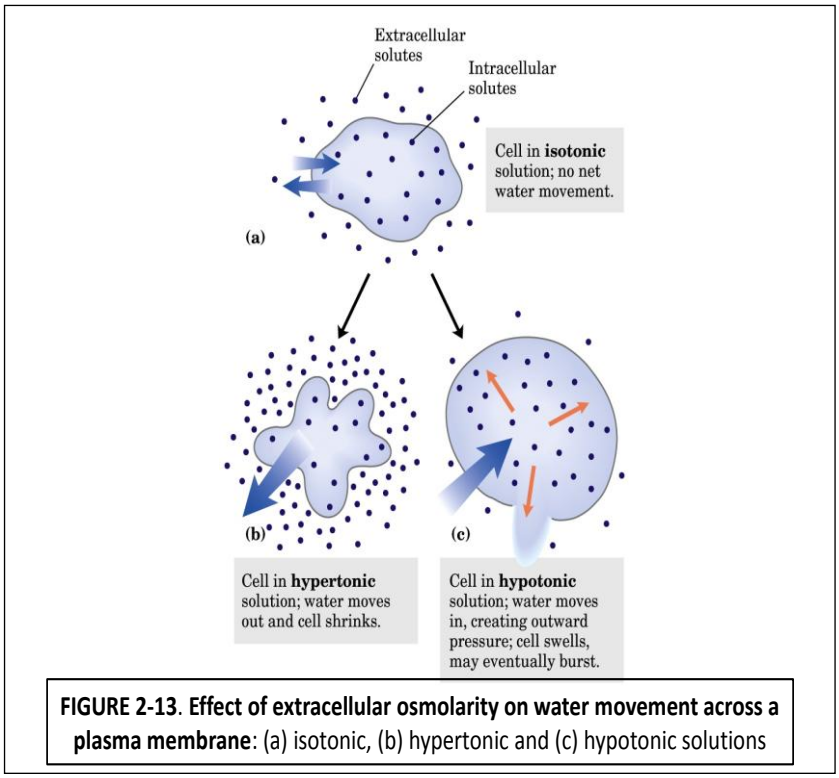


FIGURE 2-13. Effect of extracellular osmolarity on water movement across a plasma membrane: (a) isotonic, (b) hypertonic and (c) hypotonic solutions

Because the effect of solutes on osmolarity depends on the number of dissolved particles, not their mass, macromolecules (proteins, nucleic acids, polysaccharides) have for less effect on osmolarity of a solution than would an equal mass of their monomeric components. For example, a gram of a polysaccharide composed of 1,000 glucose units has the same effect on osmolarity as a mg of glucose. One effect of storage fuel as polysaccharides (starch or glycogen) rather than as glucose or other simple sugars is prevention of an enormous increase in osmotic pressure within the storage cell.

Ionization of Water, Weak Acids and Weak Bases

Although many of the solvent properties of water can be explained in terms of the uncharged H_2O molecule, the small degree of ionization of water to H^+ and OH^- must also be taken into account. When weak acids are dissolved in water, they contribute H^+ by ionizing; weak bases consume H^+ by becoming protonated. These processes are governed by equilibrium constants. The total H^+ concentration from all sources is experimentally measurable and is expressed as the pH of the solution.

Pure Water is Slightly Ionized

Water molecules have a slight tendency to undergo reversible ionization to yield a hydrogen ion (a proton) and a hydroxyl ion.

$$H_2O \rightleftharpoons H^+ + OH^-$$

The degree of ionization of water at equilibrium is small; at 25°C only about two of every 10^9 molecules in pure water are ionized at any instant.

$$Keq = \frac{[H^+][OH^-]}{[H_2O]}$$

In pure water at 25°C, the concentration of water is 55.5M (grams of H_2O in 1L divided by its gram molecular weight: (1000 g/L)/(18.015 g/mol) and is essentially constant in relation to the very low concentration of H^+ and OH^-, namely, 1×10^{-7}M.

$$Keq = \frac{[H^+][OH^-]}{55.5M}$$

(55.5M) (K_{eq}) = $[H^+]$ $[OH^-]$ = K_w

The K_w designates the product (55.5M) (K_{eq}), **the ion product of water** at 25°C.

The value for K_{eq} is 1.8×10^{-16}M at 25°C.

K_w = $[H^+]$ $[OH^-]$ = (55.5M) (1.8×10^{-16}M)

= 1.0×10^{-14} M^2

When there are exactly equal concentrations of H^+ and OH^-, as in pure water, the solution is said to be at neutral pH. At this pH, the concentrations of H^+ and OH^- can be calculated:

$$Kw = [H^+][OH^-] = [H^+]^2$$
$$[H^+] = \sqrt{Kw} = \sqrt{1\times10^{-14}}\,M^2$$
$$[H^+] = [OH^-] = 10^{-7}\,M$$

As the ion product of water is constant, wherever, $[H^+]$ is greater than 1×10^{-7}M, $[OH^-]$ must become less than 1×10^{-7}M and vice versa.

The pH Scale Designates the $[H^+]$ and $[OH^-]$ concentrations

The ion product of water, K_w, is the basis for the pH scale (Lehninger, Table: 2-6)

$$pH = \log\frac{1}{[H^+]} = -\log[H^+]$$

If the $[H^+]$ concentration is 1.0×10^{-7}M

pH = log (1.0×10^{-7}) = 7

Measurement of pH is one of the most important and frequently used procedures in biochemistry. The pH affects the structure and activity of biological macromolecule; for example, the catalytic activity of enzymes are strongly dependent on pH. Measurement of the pH of blood and urine is commonly used in medical diagnoses. The pH of the blood plasma of people with severe, uncontrolled diabetes, for example, is often below the normal value of 7.4; this condition is called acidosis. In certain other disease states the pH of the blood is higher than normal, the condition of alkalosis.

TABLE 2-6. pH Scale

The pH Scale

$[H^+]$ (M)	pH	$[OH^-]$ (M)	pOH*
10^0 (1)	0	10^{-14}	14
10^{-1}	1	10^{-13}	13
10^{-2}	2	10^{-12}	12
10^{-3}	3	10^{-11}	11
10^{-4}	4	10^{-10}	10
10^{-5}	5	10^{-9}	9
10^{-6}	6	10^{-8}	8
10^{-7}	7	10^{-7}	7
10^{-8}	8	10^{-6}	6
10^{-9}	9	10^{-5}	5
10^{-10}	10	10^{-4}	4
10^{-11}	11	10^{-3}	3
10^{-12}	12	10^{-2}	2
10^{-13}	13	10^{-1}	1
10^{-14}	14	10^0 (1)	0

*The expression pOH is sometimes used to describe the basicity, or OH^- concentration, of a solution; pOH is defined by the expression $pOH = -\log [OH^-]$, which is analogous to the expression for pH. Note that in all cases, pH + pOH = 14.

Weak Acids and Bases have Characteristic Dissociation Constants

Hydrochloric, sulfuric and nitric acid commonly called strong acids, are completely ionized in dilute aqueous solutions; the strong bases NaOH and KOH are also completely ionized. Weak acid and bases are not completely ionized when dissolved in water. These are common in biological systems and play important roles in metabolism and its regulation.

Acids may be defined as proton donors and bases are proton acceptors. A proton donor and its corresponding proton acceptor make up a **conjugate acid-base pair** (Lehninger, Fig. 2-16). Acetic acid (CH_3 COOH), a proton donor and the acetate ion (CH_3 COO^-), the corresponding proton acceptor constitute a conjugate acid-base pair, related by the reversible reaction:

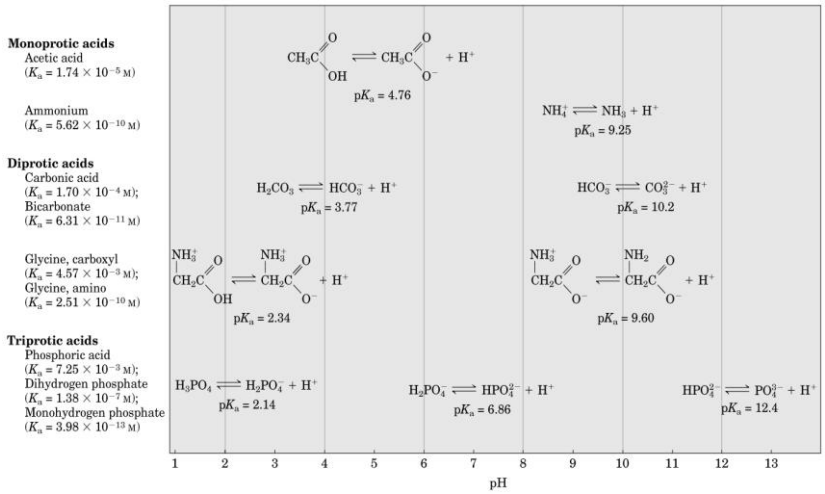


FIGURE 2-16. Conjugate acid-base pairs consist of a proton donor and a proton acceptor

$$CH_3\,COOH \rightleftharpoons CH_3COO^- + H^+$$

The stronger the acid, the greater its tendency to lose its proton. The tendency of any acid (HA) to lose a proton and form its conjugate base (A^-) is defined by the equilibrium constant (K_{eq}) for the reversible reaction

$$HA \rightleftharpoons H^+ + A^-$$

$$Keq = \frac{[H^+][A^-]}{[HA]} = Ka$$

Equilibrium constants for ionization reactions are usually called ionization or **dissociation constants**, often designated Ka.

pKa = $^-$log Ka **(analogous to pH)**

The stronger the tendency to dissociate a proton, the stronger is the acid and the lower its pKa.

Titration Curves

Titration curves are used to determine the amount of an acids in a given solution. A measured volume of the acids is titrated with a solution of a strong base, usually NaOH of known concentration. A plot of pH against the amount of NaOH added (**a titration curve**) reveals the pKa of the weak acid. Consider the titration of 0.1M solution of acetic acid (HAc) with 0.1M NaOH at 25°C (Lehninger, Fig. 2-17).

As Na OH is gradually introduced, the added OH^- combines with the free H^+ in the solution to form H_2O. As free H^+ is removed, HAc dissociates further to satisfy its own equilibrium constant, the net result as the titration proceeds is that more and more HAc ionizes, forming Ac^-, as the NaOH is added. At the midpoint of the titration at which exactly 0.5 equivalent of NaOH has been added, one half of the original acetic acid has undergone dissociation, so that the concentration of the proton donor (HAc) now equals that of the proton acceptor, [Ac^-]. At this midpoint a very important relationship holds: the pH of the equimolar solution of acetic acid and acetate is exactly equal to the pKa of acetic acid (pKa = 4.76) (Lehninger, Fig. 2-16, 2-17).

Fig. 2-18 (Lehninger) compares the titration curves of three weak acids with very different dissociation constants; acetic acid pKa, 4.76); dihydrogen phosphate, H_2PO_4 (pKa = 6.86) and ammonium ion, NH_4^+ (pKa = 9.25). Acetic acid with the highest Ka (lowest pKa) of the three, is the strongest (loses its proton most readily). The most important point about the titration curve of a weak acid is that it shows graphically that a weak acid and its anion -a conjugate acid-base pair- can act as a buffer.

Buffering Against pH Changes in Biological Systems

Almost every biological process is pH dependent; a small change in the pH produces a large change in the rate of the process. The enzymes that catalyze cellular reaction contain ionizable groups with characteristic pKa values. In multicellular organisms, the pH of extracellular fluids is tightly regulated. This is achieved primarily by biological buffers;

mixtures of weak acids and their conjugate bases. Biological buffering is illustrated by the phosphate and carbonate buffering systems of humans.

Buffers are Mixtures of Weak Acids and their Conjugate Bases:

Buffers are aqueous systems that tend to resist changes in pH when small amounts of acid (H^+) or base (OH^-) are added. A buffer system consists of a weak acid (the proton donor) and its conjugate base (the proton acceptor). As an example, a mixture of equal concentrations of acetic acid and acetate ion, found at the midpoint of the titration curve in Fig. 2-17 (Lehninger), is a buffer system. The titration curve of acetic acid has a relatively flat zone extending about 1 pH unit on either side of its midpoint pH of 4.76. In this zone, an amount of H^+ or OH^- added to the system has much less effect on pH than the same amount added outside the buffer range. This relatively flat zone is the buffering region of the acetic acid acetate buffer pair. At the midpoint of the buffering region, where the concentration of acetic acid exactly equals that of the acetate, the buffering power of the system is maximal. The pH at this point in the titration curve of acetic acid is equal to its pKa. Fig. 2-19 (Lehninger) explains how a buffer system works. Whenever H^+ or OH^- is added to buffer, the result is a small change in the ratio of the relative concentration of the weak acid and its anion and thus a small change in pH.

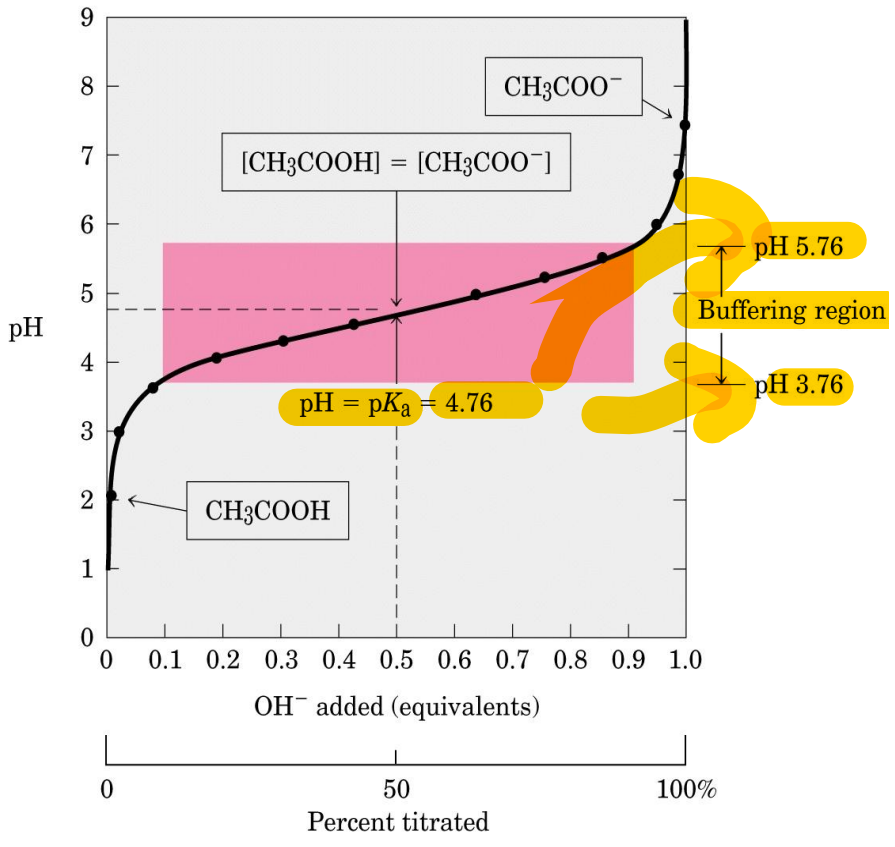


FIGURE 2-17. The titration curve of acetic acid

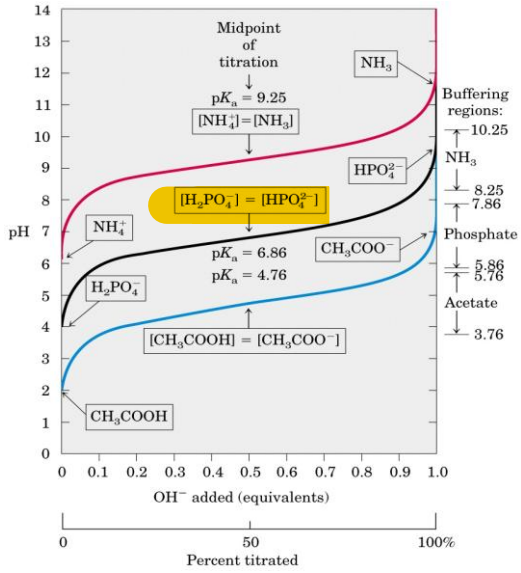


FIGURE 2-18. Comparison of the titration curves of three weak acids

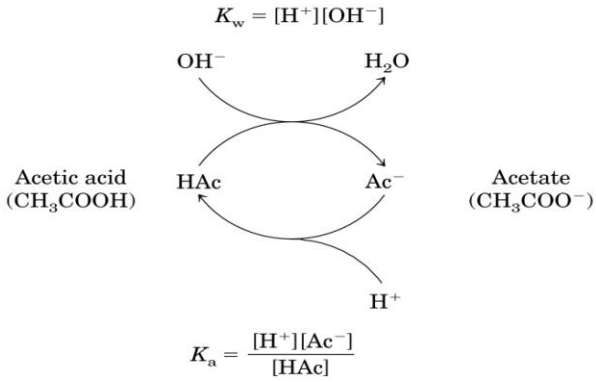


FIGURE 2-19. The acetic acid – acetate pair as a buffer system

Each conjugate acid-base pair has a characteristic pH zone in which it is an effective buffer (Lehninger, Fig. 2-18). The $H_2PO_4^-/HPO_4^{-2}$ pair has a pKa of 6.86 and thus can sever as an effective buffer system between ~ pH 5.9 and pH 7.9.

Henderson-Hasselbalch Equation

$$HA = H^+ + A^-$$

$$Ka = \frac{[H^+][A^-]}{[HA]}$$

$$[H^+] = Ka\,\frac{[HA]}{[A^-]}$$

$$-\log[H^+] = -\log Ka - \log\frac{[HA]}{[A^-]}$$

$$pH = Pka + \log\frac{[A^-]}{[HA]}$$

$$pH = pKa + \log\frac{[proton\ acceptor]}{[proton\ donor]}$$

This Henderson Hasselbalch equation shows why the pKa of a weak acid is equal to the pH of the solution at the midpoint of its titration curve. At that point, [HA] equals [A^-].

pH = pKa + log 1 = pKa + 0 = pKa

The equation also allows us to (1) calculate pKa, given pH and the molar ratio of proton donor and acceptor; (2) calculate pH, given pKa and the molar ratio of proton donor and acceptor; and (3) calculate the molar ratio of proton donor and acceptor, given pH and pKa.

Buffering System in Cells

Two especially important biological buffers are the phosphate and bicarbonate system. The phosphate buffer system, which acts in the cytoplasm of all cells, consists of $H_2PO_4^-$ as proton donor and HPO_4^{-2} as proton acceptor:

$$H_2PO_4^- \rightleftharpoons H^+ + HPO_4^{2-}$$

The phosphate buffer system is maximally effective at pH close to its pKa of 6.86, and thus tends to resist pH changes in the range between about 5.9 and 7.9, it is therefore an effective buffer in biological fluids; in mammals, for example. ECFs and most cytoplasmic comportments have a pH in the range of 6.9 to 7.4.

Blood plasma is buffered in part by the bicarbonate system, consisting of carbonic acid (H_2CO_3) as proton donor and bicarbonate HCO^-_3 as proton acceptor:

$$H_2CO_3 \rightleftharpoons H^+ + HCO_3^-$$

This buffer system is more complex than other conjugate acid bore pairs because one of the components, carbonic acid (H_2CO_3), is formed from dissolved (d) CO_2 and water:

$$CO_2\,(d) + H_2O \rightleftharpoons H_2CO_3$$

CO_2 is a gas under normal conditions, and the concentration of dissolved CO_2 is the result of equilibration with CO_2 of the gas (g) phase:

$$CO_2\,(g) \rightleftharpoons CO_2\,(d)$$

Since the concentration of H_2CO_3 depends on the concentration of CO_2 (d), which in turn depends on the concentration of CO_2 (g), called the **partial pressure** of CO_2. The pH of a bicarbonate buffer exposed to a gas phase is ultimately determined by the concentration of HCO_3^- in the aqueous phase and the partial pressure of CO_2 in the gas phase.

Human blood plasma normally has a pH close to 7.4. Should the pH regulating mechanisms fail or be overwhelmed, as may happen in severe uncontrolled diabetes when an over production of metabolic acids causes acidosis, the pH of the blood can fall to 6.8 or below, leading to irreparable cell damage and death. In other diseases the pH may rise to lethal levels.

The bicarbonate buffer system is an effective physiological buffer near pH 7.4, because the H_2CO_3 of blood plasma is in equilibrium with a large reserve capacity of CO_2(g) in the air space of the lungs. This buffer system involves three reversible equilibria between gaseous CO_2 in the lungs and bicarbonate HCO_3^- in the blood plasma (Lehninger, Fig. 1, box 2-4).

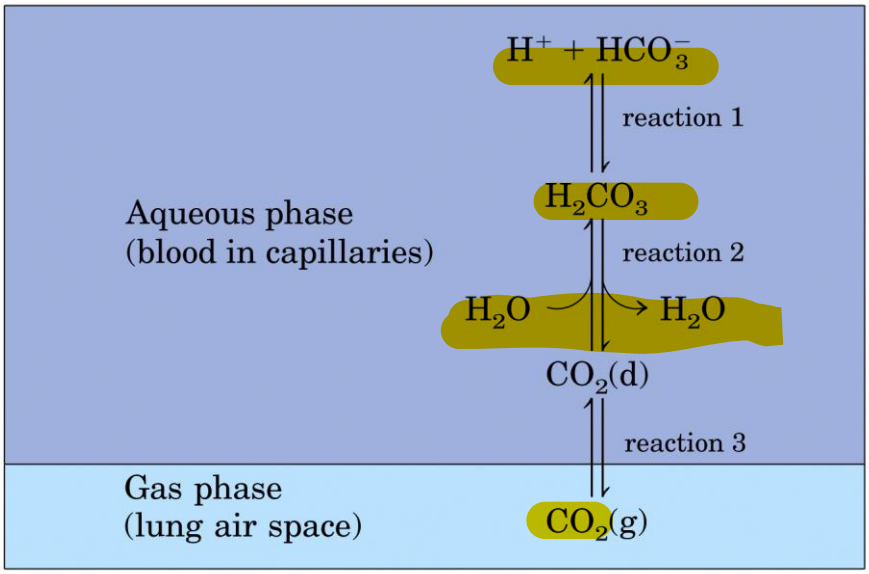


Fig. 1 [Box 2-4]. The CO2 in the air space of the lungs is in equilibrium with the bicarbonate buffer in the blood plasma passing through the lung capillaries

Enzymes typically show maximal catalytic activity at a characteristic pH, called the **pH optimum** (Lehninger, Fig. 2-21). A small charge in pH can make a large difference in the rate of some crucial enzyme-catalyzed reactions. Biological control of pH of cells and body fluids is therefore of central importance in all aspects of metabolism and cellular activities.

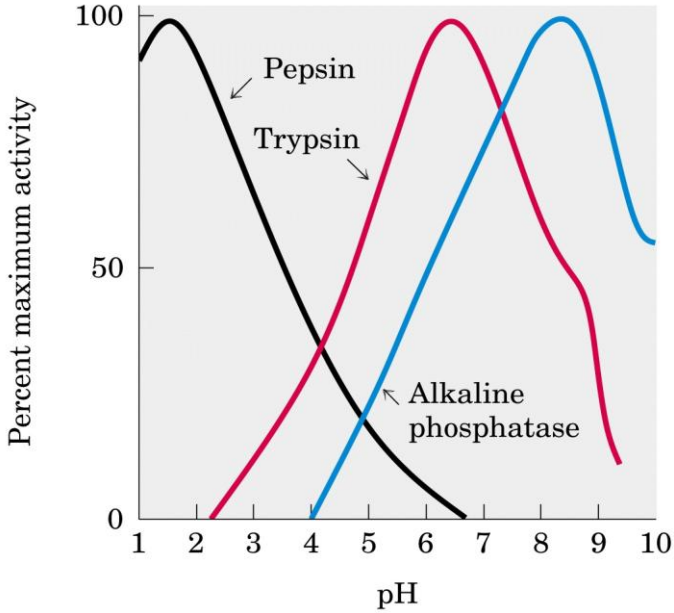


FIGURE 2-21. The pH optima of some enzymes

AMINO ACIDS, PEPTIDES AND PROTEINS

Proteins are the most abundant biological macromolecules, occurring in all cells and all parts of cells. Proteins exhibit enormous diversity of biological function and size. Proteins are the molecular instruments through which genetic information is expressed.

All proteins are constructed from the same set of 20 amino acids, covalently liked in characteristic linear sequences. Cells can produce proteins with different properties and activities by joining the same 20 amino acids in many different combinations and sequences. Different organisms can make diverse products such as enzymes, hormones, antibodies, transporters, muscle fibers, … etc.

Amino Acids Share Common Structural Features

All 20 of the common amino acids are α-amino acids. They have a carboxyl group and an amino group bonded to the same carbon atom (the α carbon) (Lehninger, Fig. 3-2).

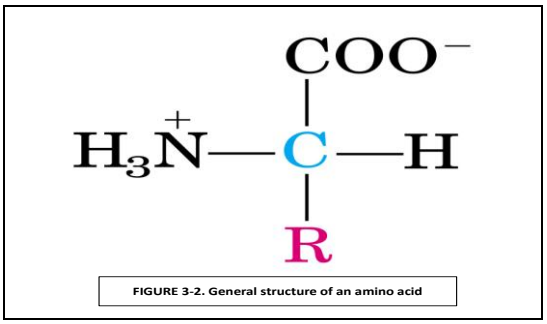
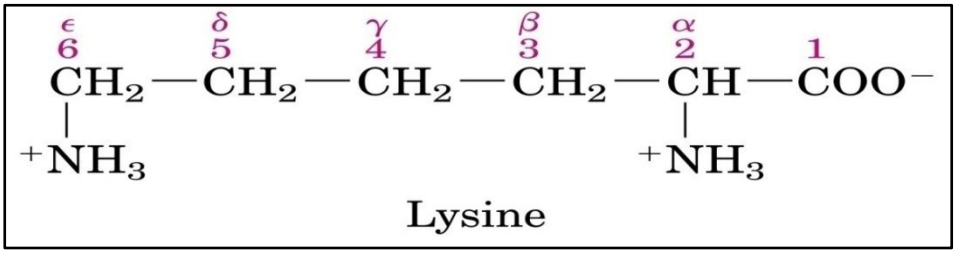


FIGURE 3-2. General structure of an amino acid

They differ from each other in their side chains or R groups, which vary in structure, size and electric charge, and which influence the solubility of the amino acids in water. The carbons in an amino acid are identified by two conventions:



The amino acid residues in protein molecules are exclusively L stereoisomers.

Amino Acids can be Classified by R Group

There are five main classes of amino acids based on the properties of their R groups (Lehninger, Table 3-1), in particular, their **polarity**, or tendency to interact with water at biological pH (near pH 7.0).

The structures of the 20 common amino acids are shown in Fig. 3-5 (Lehninger), and some of their properties are listed in Table 3-1 (Lehninger).

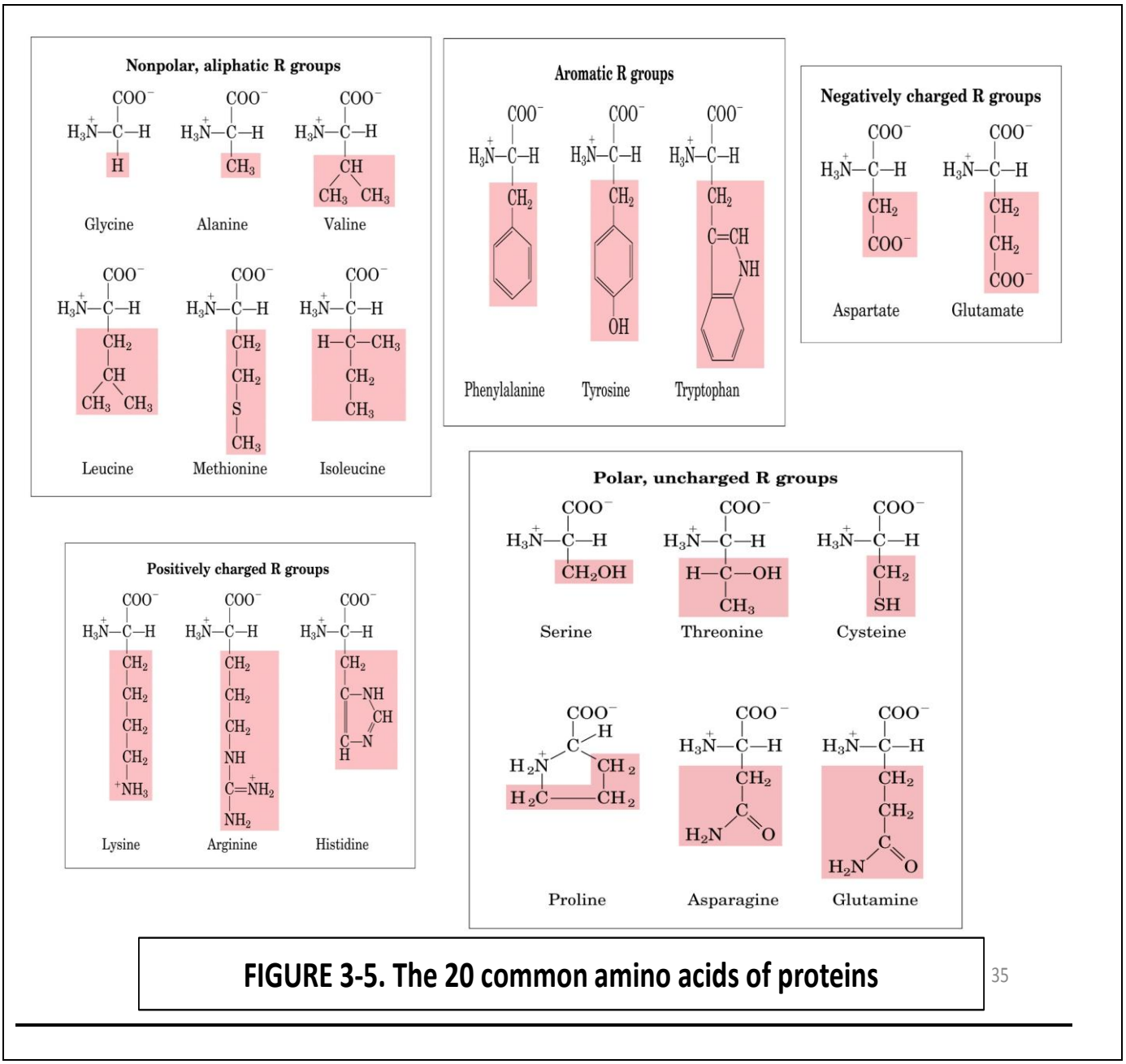


FIGURE 3-5. The 20 common amino acids of proteins

35

TABLE 3-1. Properties of common amino acids found in proteins

Properties and Conventions Associated with the Standard Amino Acids

Amino acid	Abbreviated names		M_r	pK_a values: pK_1 (—COOH)	pK_2 ($—NH_3^+$)	pK_R (R group)	pI	Hydropathy index*	Occurrence in proteins (%)
Nonpolar, aliphatic R groups									
Glycine	Gly	G	75	2.34	9.60		5.97	−0.4	7.2
Alanine	Ala	A	89	2.34	9.69		6.01	1.8	7.8
Valine	Val	V	117	2.32	9.62		5.97	4.2	6.6
Leucine	Leu	L	131	2.36	9.60		5.98	3.8	9.1
Isoleucine	Ile	I	131	2.36	9.68		6.02	4.5	5.3
Methionine	Met	M	149	2.28	9.21		5.74	1.9	2.3
Aromatic R groups									
Phenylalanine	Phe	F	165	1.83	9.13		5.48	2.8	3.9
Tyrosine	Tyr	Y	181	2.20	9.11	10.07	5.66	−1.3	3.2
Tryptophan	Trp	W	204	2.38	9.39		5.89	−0.9	1.4
Polar, uncharged R groups									
Serine	Ser	S	105	2.21	9.15		5.68	−0.8	6.8
Proline	Pro	P	115	1.99	10.96		6.48	1.6	5.2
Threonine	Thr	T	119	2.11	9.62		5.87	−0.7	5.9
Cysteine	Cys	C	121	1.96	10.28	8.18	5.07	2.5	1.9
Asparagine	Asn	N	132	2.02	8.80		5.41	−3.5	4.3
Glutamine	Gln	Q	146	2.17	9.13		5.65	−3.5	4.2
Positively charged R groups									
Lysine	Lys	K	146	2.18	8.95	10.53	9.74	−3.9	5.9
Histidine	His	H	155	1.82	9.17	6.00	7.59	−3.2	2.3
Arginine	Arg	R	174	2.17	9.04	12.48	10.76	−4.5	5.1
Negatively charged R groups									
Aspartate	Asp	D	133	1.88	9.60	3.65	2.77	−3.5	5.3
Glutamate	Glu	E	147	2.19	9.67	4.25	3.22	−3.5	6.3

*A scale combining hydrophobicity and hydrophilicity of R groups; it can be used to measure the tendency of an amino acid to seek an aqueous environment (− values) or a hydrophobic environment (+ values). See Chapter 12. From Kyte, J. & Doolittle, R.F. (1982) *J. Mol. Biol.* **157,** 105 – 132.

†Average occurrence in over 1150 proteins. From Doolittle, R.F. (1989) Redundancies in protein sequences. In *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G.D., ed) Plenum Press, NY, pp. 599–623.

Nonpolar, Aliphatic R Groups: the R groups in this class of amino acids are nonpolar and hydrophobic. The side chains of **alanine, valine, leucine and isoleucine** tend to cluster together within proteins, stabilizing protein structure by means of hydrophobic interactions. **Glycine** has the simplest structure. **Methionine**, one of the two sulfur-containing amino acids, has a nonpolar thioether group in its side chain. Proline has an aliphatic side chain with a distinctive cyclic structure. **Proline** reduces the structural flexibility of the polypeptide regions containing proline.

Aromatic R Groups: Phenylalanine, tyrosine, and tryptophan, with their aromatic side chains, are relatively nonpolar (hydrophobic). The –OH group of tyrosine can form hydrogen bond, and it is an important functional group in some enzymes. Tyrosine and tryptophan, and to a much lesser extent phenylalanine, absorb U.V. light (Lehninger, Fig. 3-6). This accounts for the characteristic strong absorbance of light by most proteins at a wavelength of 280 nm.

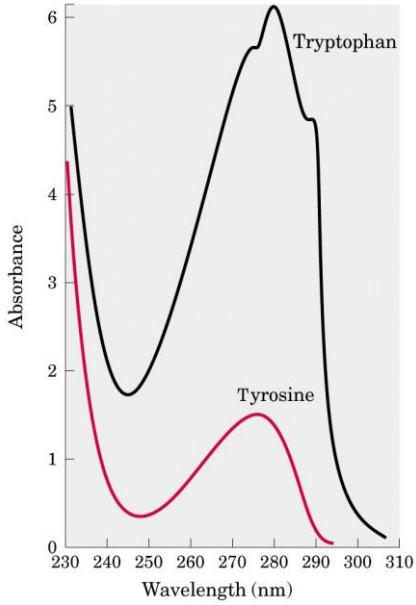


FIGURE 3-6. Absorption of U.V. light by aromatic amino acids

Polar, Uncharged R groups: The R groups of these amino acids are more soluble in water, or more hydrophilic than nonpolar amino acids, because they contain functional groups that form H- bonds with water. This class of amino acids includes: **serine, threonine, cysteine, asparagine, and glutamine**. Asparagine and glutamine are the amides of aspartate and glutamate. Cysteine is readily oxidized to form a covalently linked dimeric amino acid called **cystine** (Lehninger, Fig. 3-7).

Positively Charged (Basic) R Groups: The most hydrophilic R groups are those that are either positively or negatively charged. **Lysine** has a second primary amino group at the ε position. **Arginine** has a positively charged guanidino group, and **histidine**, which has an imidazole group. Histidine is the only common amino acid having an ionizable side chain with a pKa near neutrality. In many enzyme catalyzed reactions, a histidine residue facilitates the reaction by serving as a proton donor/acceptor.

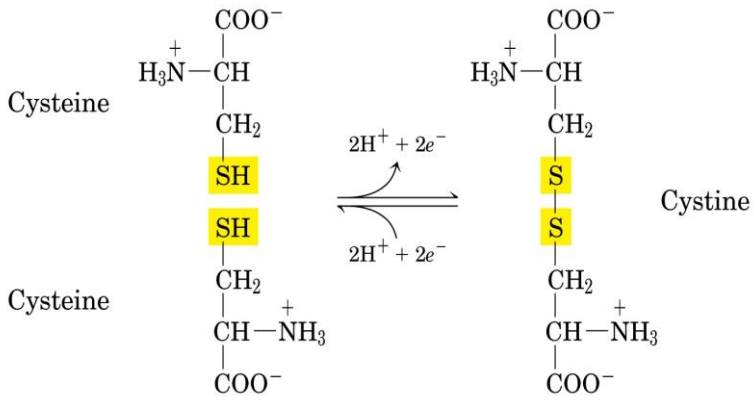


FIGURE 3-7. Reversible formation of a disulfide bond by the oxidation of two molecules of cycteine

Negatively Charged (Acidic) R Groups

The two amino acids having R groups with a net negative charge at pH 7.0 are **aspartate** and **glutamate**, each of which has a second carboxyl group.

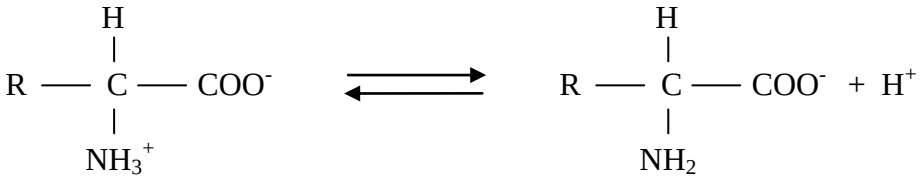
Uncommon Amino Acids

In addition to the 20 common amino acids, proteins may contain residues created by modification of common residues. Among these uncommon amino acids are 4-hydroxyproline, a derivative of proline, and 5-hydroxyl lysine, derived from lysine. The former is found in plant cell wall proteins, and both are found in collagen. 6-**N-Methyllysine** is a constituent of myosin. **γ–Carboxyglutamate** found in blood–clotting protein prothrombin and in certain other proteins that bind Ca^{++} as part of their biological function.

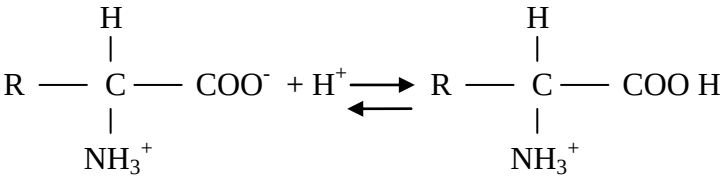
Some 300 additional uncommon amino acids have been found in cells. They have a variety of functions but are not constituents of proteins. Ornithine and Citrulline are key intermediates in the biosynthesis of arginine and in the urea cycle.

Amino Acids Can Act as Acids and Bases

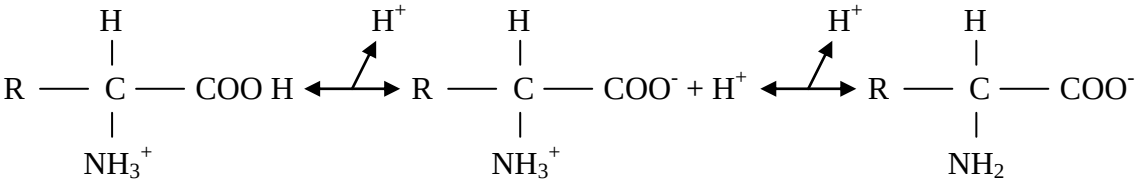
When an amino acid is dissolved in water, it exists in solution as the dipolar ion-or **zwitterions**. A zwitterion can act as either an acid (proton donor)



or a base (proton acceptor)



Substances having this dual nature are amphoteric and are called ampholytes.



Net charge +1 **O** **-1**

Titration Curves

Acid-base titration involves the gradual addition or removal of protons. Fig. 3-10 (Lehninger) shows the titration curve of glycine. The plot has two distinct stages corresponding to deprotonation of two different groups of glycine. At very low pH, the predominant ionic species of glycine is the fully protonated form: $^{+}H_3N - CH_2 - COOH$. At the midpoint in the first stage of the titration, where pH = pKa of the carboxyl group, glycine is present as: $^{+}H_3N - CH_2 - COOH$ and $^{+}H_3N - CH_2 - COO^{-}$ of equimolar concentration the pH at the midpoint is 2.34.

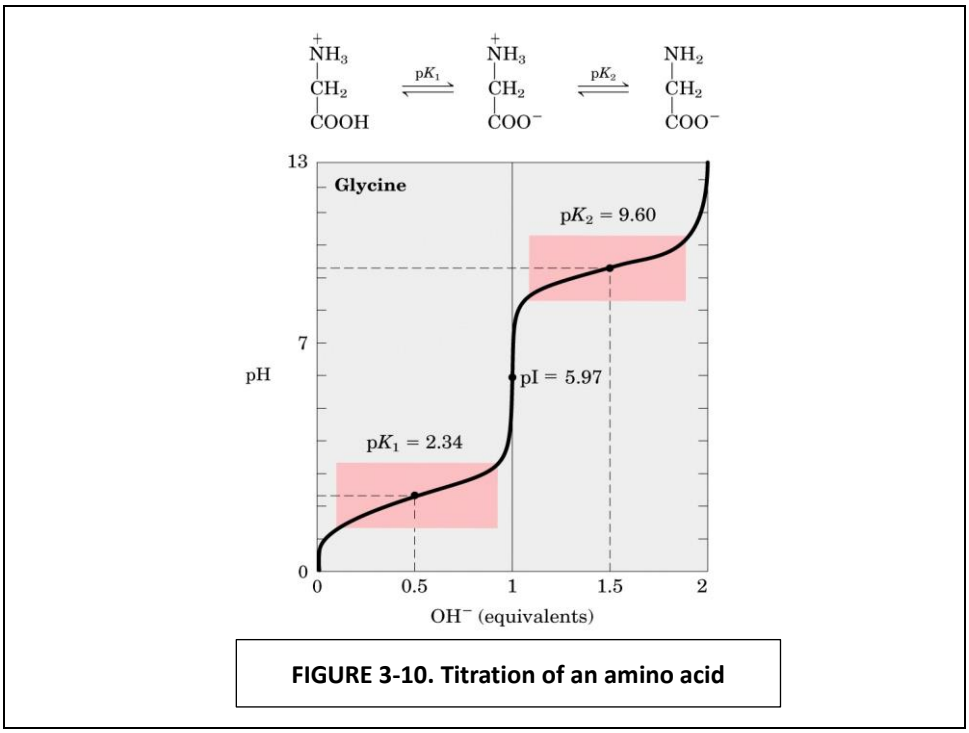


FIGURE 3-10. Titration of an amino acid

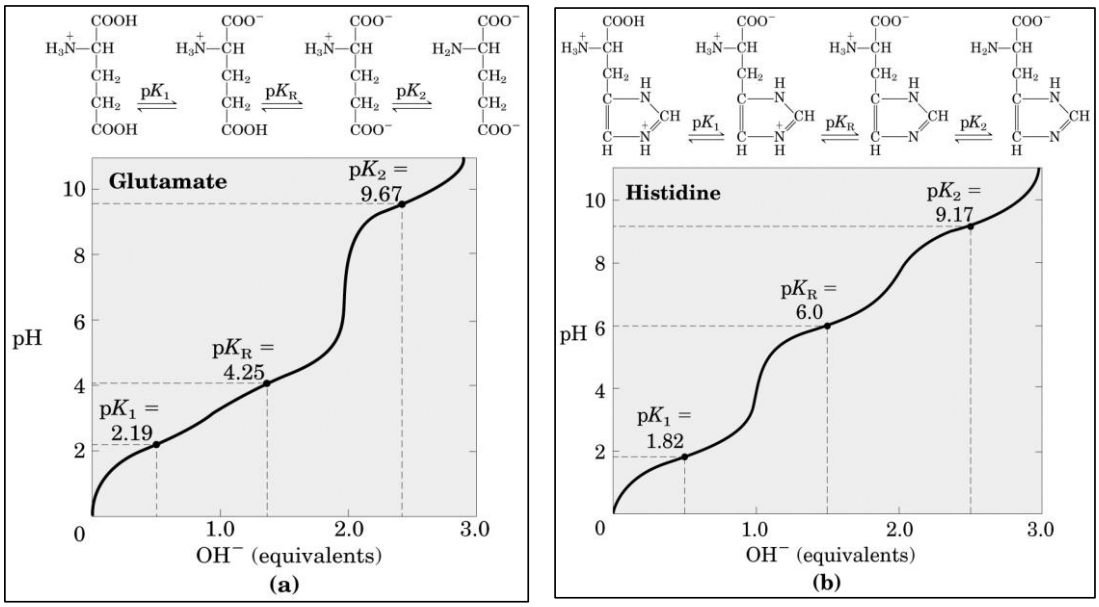


FIGURE 3-12. Titration curve for (a) glutamic acid and (b) histidine

The pKa of the -COOH group of glycine (pK_1=2.34). The pKa is a measure of the tendency of a group to give up a proton, with that tendency decreasing tenfold as the pKa increases by one unit.

The second stage of the titration corresponds to the removal of a proton from $–NH_3^+$ group of glycine. The pH at the midpoint of this stage is 9.60, equal to the pKa for the $–NH_3^+$ group (pK_2). The titration is essentially complete at a pH of about 12, at which point the predominant form of glycine is $NH_2 – CH_2 – COO^-$.

From the titration curve of glycine we can derive several important pieces of information:

- it gives a quantitative measure of the pKa of each of the two ionizing groups. Note that the –COOH group of glycine is 100 times more acidic than the carboxyl group of acetic acid (pKa of 2.34 vs. 4.76). This is due to the repulsion between departing H^+ and the nearby positively charged amino group on the α-carbon atom. The pKa of a functional group is greatly affected by its chemical environment. (This is important in active sites of enzymes).
- This amino acid has two regions of buffering power (maximum buffering activity at pH = pKa).
- The relationship between the net electric charge of the amino acid and the pH of the solution. At pH 5.97 glycine is fully ionized with no net electric charge. The pH at which the net electric charge is zero is called the **isoelectric point** or **isoelectric pH** designated **pI**. For glycine:

$$\mathbf{pI} = \frac{1}{2}\ \mathbf{(Pk_1 + Pk_2)} = \frac{1}{2}\ \mathbf{(2.34 + 9.60) = 5.97}$$

Glycine has a net negative charge at pH > pI
Glycine has a net positive charge at pH < pI

Acid –Base Properties of Amino Acids

- All amino acids with a single α-amino group, a single α-carboxyl group, and an R-group that does not ionize have titration curves resembling that of glycine.
- Amino acids with an ionizable R group have more complex titration curves, with three stages corresponding to three ionizable groups. (Lehninger, Fig. 3-12).
- Only histidine has significant buffering power near the neutral pH found in the intracellular and extracellular fluids of most animals.

Peptides and proteins

Peptides and proteins are polymers of amino acids. Biologically occurring polypeptides range in size from small to very large, consisting of 2 or 3 to thousands of linked amino acid residues.

Peptides are Chains of Amino Acids

Two amino acids can be covalently joined by a **peptide bond** to yield a dipeptide (Lehninger, Fig.3-13), when a few amino acids are joined in this fashion; the structure is called an **oligopeptide**. When many amino acids are joined, the product is called a **polypeptide**. Polypeptides may have generally M_r below 10,000 and proteins have higher M_r.

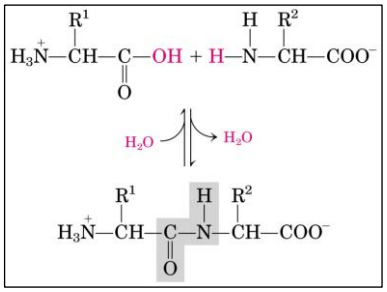


FIGURE 3-13. Formation of a peptide bond by condensation

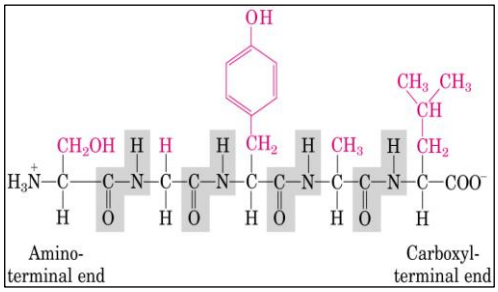


FIGURE 3-14. The pentapeptide (Ser-Gly-Tyr-Ala-Leu)

Fig. 3-14 (Lehninger) shows the structure of a pentapeptide. In a peptide, the amino acid residue at the end with a free α-amino group is the **amino–terminal** (or N-terminal) residue; the residue at the other end, which has a free carboxyl group, is the **carboxyl-terminal** (C-terminal) residue.

The acid-base behavior of a peptide can be predicted from its free α-amino and α-carboxyl groups as well as the nature and number of its ionizable R groups. Therefore, peptides have characteristic titration curves and a characteristic pI.

Naturally occurring peptides range in length from two to many thousands of amino acid residues. Even the smallest peptides can have biologically important effect like the commercially synthesized dipeptide L-asp–L-phe methyl ester, the artificial sweetener better known as aspartame or NutraSweet.

Each type of protein yields a characteristic proportion of mixture of the different amino acids, when completely hydrolyzed. Some amino acids may occur only once or not at all in a given type of protein; others occur in large numbers.

Some Proteins Contain Chemical Groups

Many proteins, for example the enzyme ribonuclease A and chymotrypsinogen, contain only amino acid residues and no other chemical constituents; these are considered **simple proteins**. However, some proteins contain permanently associated chemical components; these are called **conjugated proteins**. These chemical components are called **prosthetic groups** (Lehninger, Table 3-4). For example, **lipoproteins** contain lipids, **glycoproteins** contain sugar groups and **metalloproteins** contain a specific metal. A number of proteins contain more than one prosthetic group. Usually the prosthetic group plays an important role in the protein's biological function.

TABLE 3-4. Conjugated Proteins

Conjugated Proteins

Class	Prosthetic group(s)	Example
Lipoproteins	Lipids	β_1-Lipoprotein of blood
Glycoproteins	Carbohydrates	Immunoglobulin G
Phosphoproteins	Phosphate groups	Casein of milk
Hemoproteins	Heme (iron porphyrin)	Hemoglobin
Flavoproteins	Flavin nucleotides	Succinate dehydrogenase
Metalloproteins	Iron	Ferritin
	Zinc	Alcohol dehydrogenase
	Calcium	Calmodulin
	Molybdenum	Dinitrogenase
	Copper	Plastocyanin

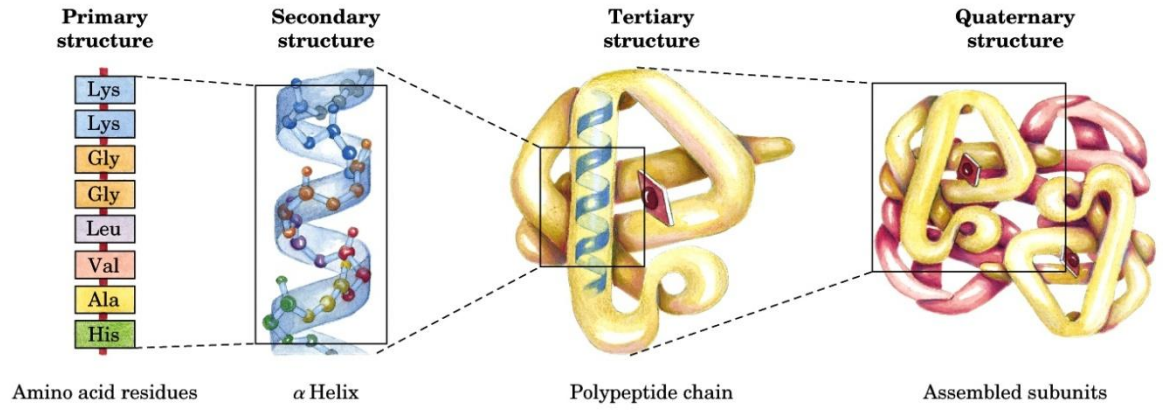


FIGURE 3-16. Levels of structure of proteins

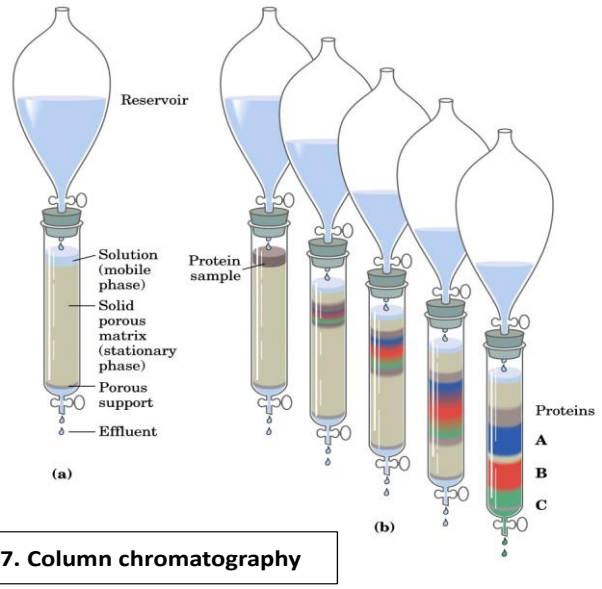


FIGURE 3-17. Column chromatography

Levels of Protein Structure

Four levels of protein structure are commonly defined (Lehninger, Fig. 3-16). **Primary structure** describes all covalent bonds with the most important element of primary structure is the sequence of amino acid residues. **Secondary structure** refers to particularly stable arrangements of amino acid residues giving rise to recurring structural patterns (α-helix, β-sheet). **Tertiary structure** describes all aspects of the three-dimensional folding of a polypeptide. When a protein has two or more polypeptide subunits, their arrangement in space is referred to as **quaternary structure**.

Separation and Purification of Proteins

A pure preparation is essential before a protein's properties and activities can be determined. Cells contain thousands of different kinds of proteins. Methods for separation of proteins are based on protein size, charge and binding properties.

The first step in any protein purification procedure is to break open the cells, releasing their proteins into a solution called a **crude extract**. If necessary, differential centrifugation can be used to prepare subcellular fractions or to isolate specific organelles.

Separation of proteins into different fractions is based on a property such as size or charge, a process referred to as **fractionation**. Early fractionation steps in purification utilize differences in protein solubility, which is a complex function of pH, temperature, salt concentration and other factors. The solubility of proteins is generally lowered at high salt concentration, an effect called "salting out". $(NH_4)_2\ SO_4$ is generally used for this purpose because of its high solubility in water.

Dialysis is a procedure that separates proteins from solvents by taking advantage of the proteins' larger size. The partially purified extract is placed in a bag made of semipermeable membrane.

Column Chromatography: one of the most powerful methods for fractionating proteins based on protein charge, size, binding affinity and other properties (Lehninger, Fig. 3-17). A porous solid material with appropriate chemical properties (the stationary phase) is held in a column and a buffered solution (the mobile phase) percolates (**يتخلل**) through it.

Individual proteins migrate faster or more slowly through the column depending on their properties. For example in **cation-exchange chromatography** (Lehninger, Fig. 3-18a), the solid matrix has negatively charged groups. In the mobile phase, proteins with a net positive charge migrate through the matrix more slowly than those with a net negative charge.

Size-exclusion chromatography (Lehninger, Fig. 3-18b): separates protein according to size. In this method, large proteins emerge from the column sooner than small ones. The solid phase consists of beads with pores of particular size. Large proteins can't enter the pores so they move around the beads. Small proteins enter the cavities and migrate slowly.

Affinity Chromatography is based on the binding affinity of a protein. The beads in the column have a covalently attached chemical group. A protein with affinity for this particular chemical group will bind to the beads in the column and its migration will be retarted as a result (Lehninger, Fig. 3-18c).

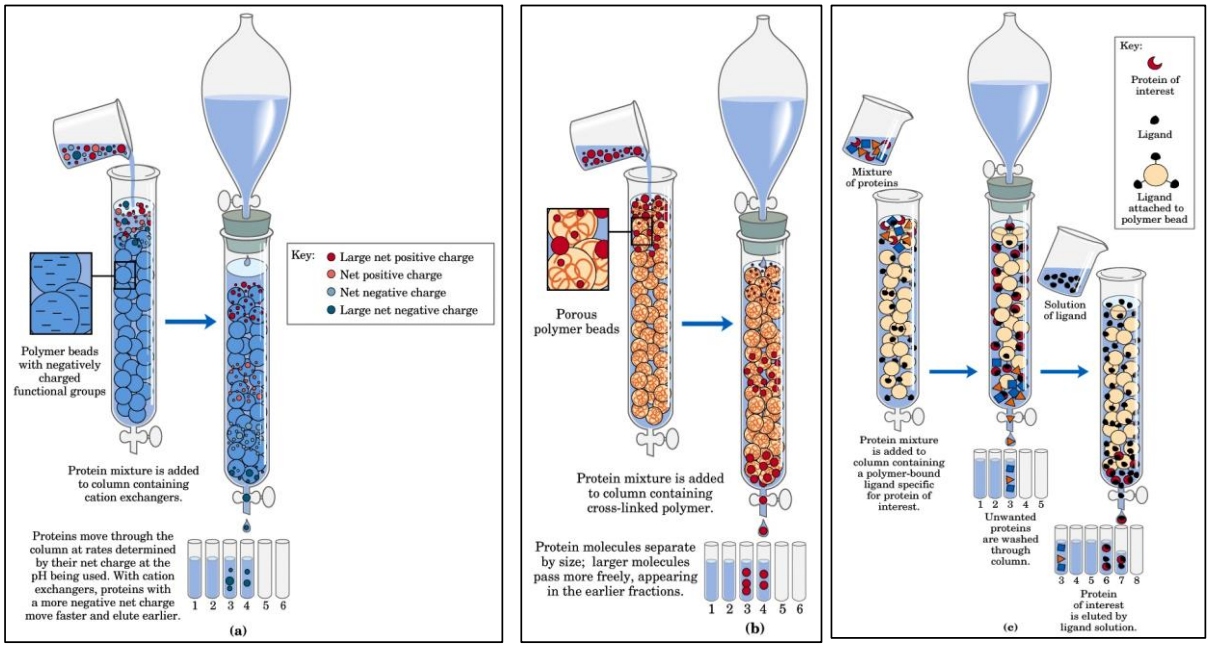


FIGURE 3-18. Three chromatographic methods used in protein purification: (a) ion-exchange chromatography; (b) size-exclusion chromatography; (c) Affinity chromatography

High-Performance Liquid Chromatography (HPLC): HPLC makes use of high-pressure pumps that speed the movement of the protein molecules down the columns. By reducing the transit time of the column, HPLC can limit diffusional spreading of protein bands and thus greatly improve resolution.

In most cases, several different methods must be used sequentially to purify a protein completely.

Electrophoresis: separates proteins according to their charge in an electric field. It is mainly used as an analytical method. It is used to estimate quickly the number of different proteins in a mixture or the degree of purity of a particular protein preparation. It is also used to determine crucial properties of a protein such as its pI and molecular weight.

Electrophoresis of proteins is generally carried out in gels made up of the cross-linked polymer polyacrylamide (Lehninger, Fig. 3-19), migration is based on the charge -to- mass ratio and is affected by protein shape.

An electrophoretic method commonly used for estimation of purity and molecular weight makes use of the detergent **Sodium Dodecyl Sulfate (SDS)**. SDS binds to most proteins, giving proteins a net negative charge and causes denaturation of proteins, thus most proteins assume similar shape. SDS-electrophoresis separates protein almost exclusively on the basis of mass (molecular weight), with smaller polypeptides migrating more rapidly.

After separation, proteins are visualized by adding a dye such as Comassie blue. When compared with the positions to which protein of known molecular weight migrate in the gel, the protein's M_r can be estimated (Lehninger, Fig. 3-20). If the protein has two or more different subunits the subunits will general be separated by the SDS treatment and a separate band will appear for each.

Isoelectric Focusing (IEF): Is a procedure used to determine the isoelectric point (pI) of a protein (Lehninger, Fig. 3-21). A pH gradient is established by allowing a mixture of low M_r organic acids and bases (ampholytes) to distribute themselves in an electric field generated across the gel. When a protein mixture is applied, each protein migrates until it reaches the pH that matches its pI.

Two-Dimensional Electrophoresis: Combining IEF and SDS gel electrophoresis permits the resolution of complex mixtures of proteins (Lehninger, Fig. 3-22). This method separates proteins of identical M_r that differ in pI, or proteins with similar pI values but different molecular weights.

Protein Sequencing

The two major discoveries in biochemistry in 1953 were the double helical structure of DNA by Watson and Crick and the amino acid sequence of insulin by Sanger. A decade later, the role of the nucleotide sequence of DNA in determining the amino acid sequence of protein was revealed.

Sanger discovered the reagent 1-flouro 2,4-dinitrobenzene (FDNB) to label the amino-terminal residue. Dansylchloride and dabsylchloride are more easily detectable. This procedure is used only to determine the N-terminal amino acid residue. It can also help to determine the number of subunits (or distinct polypeptides).

The other chemical methods were devised by Pehr Edman. The **Edman degradation** procedure labels and removes only the amino-terminal residue from a peptide, leaving all other peptide bonds intact. The reagent used is phenylisothiocyanate.

Large polypeptide must be broken down into smaller pieces to be sequenced efficiently. **Disulfide bonds** are broken by either oxidation using performic acid or reduction by dithiothreitol. Enzymatic method is also used to break peptide bonds. **Proteases** cleave specifically peptide bonds. Trypsin cleaves after Lys or Arg.

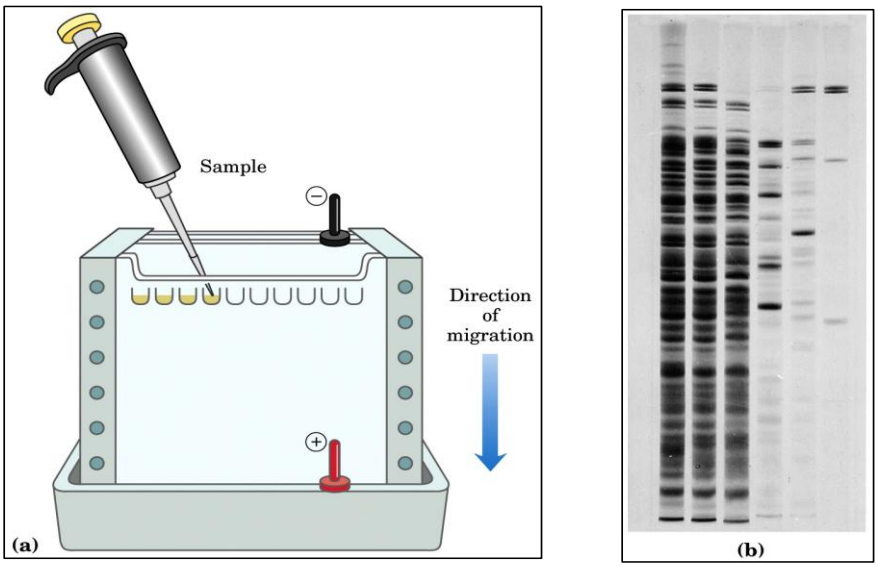


FIGURE 3-19. Electrophoresis: (a) different samples are loaded in wells at the top of the polyacrylamide gel; (b) proteins are visualized after electrophoresis with a stain such as Coomassie blue

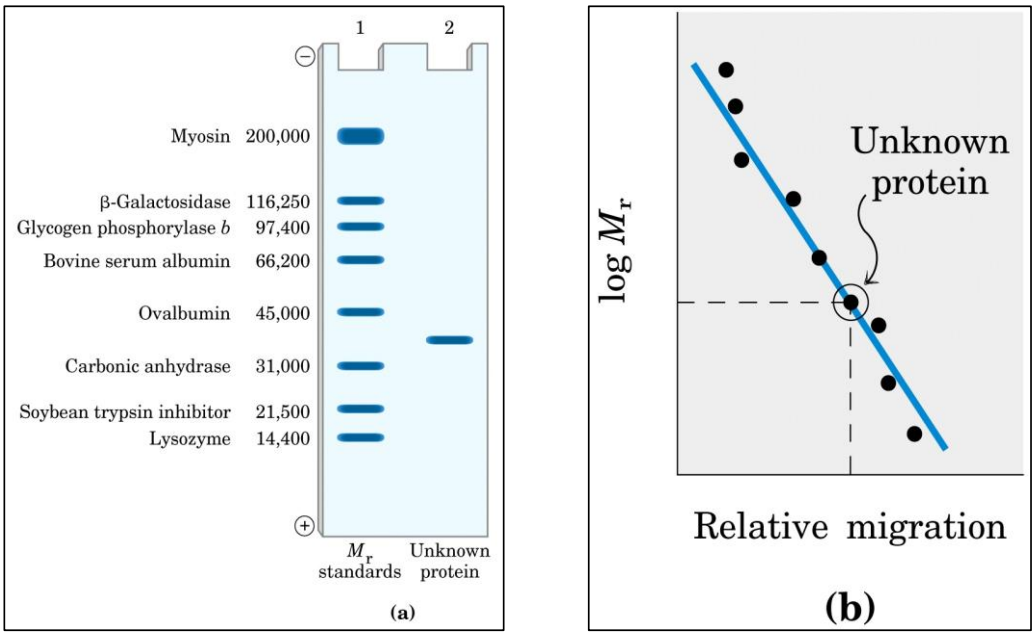


FIGURE 3-20. Estimating the molecular weight of a protein: (a) Standard proteins of known M.wt. in lane 1, unknown protein in lane 2; (b) a plot of log M.wt. of maker proteins vs. relative migration during electrophoresis.

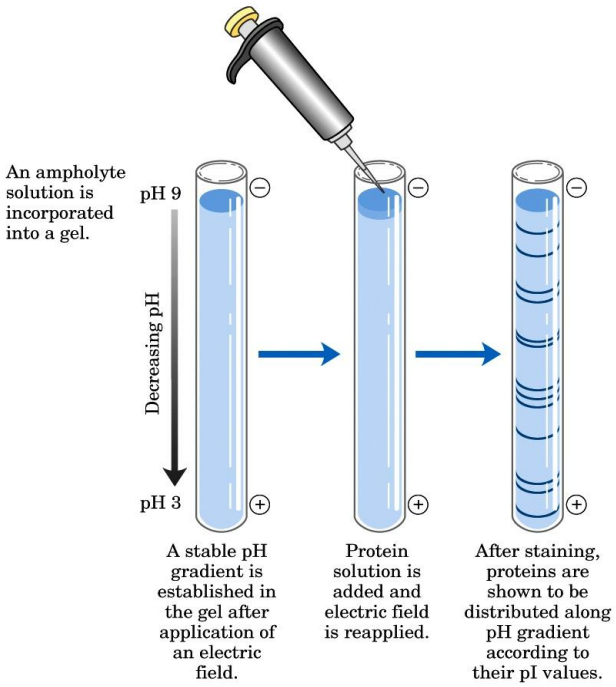


FIGURE 3-21. Isoelectric focusing

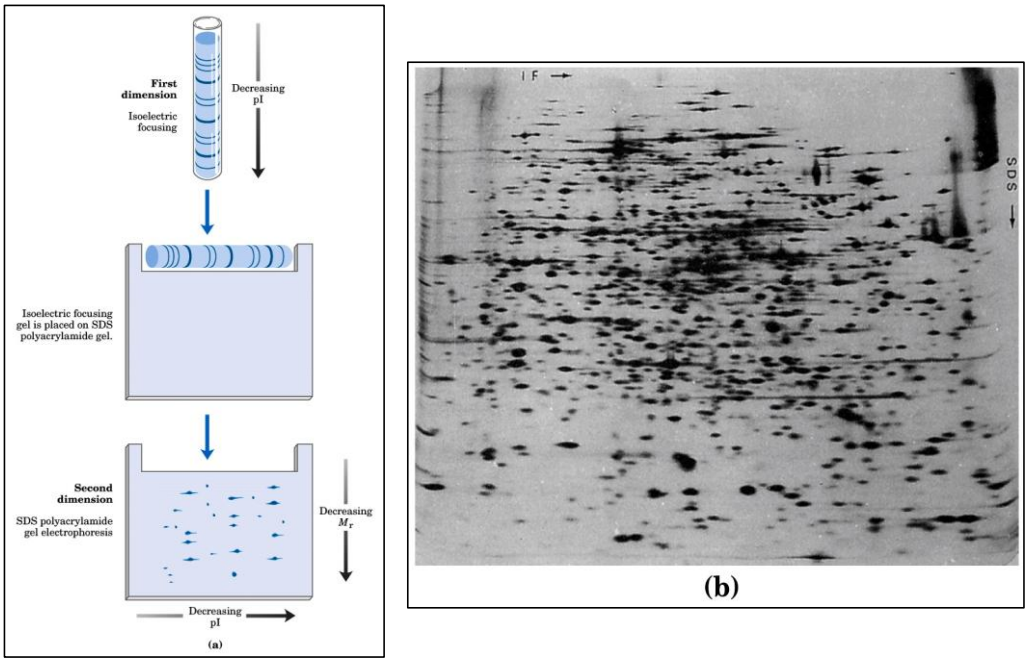


FIGURE 3-22. Two-dimensional electrophoresis

Protein Structure

The spatial arrangement of atoms in a protein is called its **conformation**. A change in conformation could occur, for example, by rotation about single bonds. The need for multiple stable conformations reflects the changes that must occur in most proteins as they bind to other molecules or catalyze reactions. Proteins in any of their functional, folded conformations are called **native** proteins.

Primary Structure

The Peptide Bond

Covalent bonds also place important constraints on the conformation of a polypeptide. The α-carbons of adjacent amino acid residues are separated by three covalent bonds, arranged as Cα – C – N – Cα. A resonance of partial sharing of two pairs of electrons occurs between the carbonyl oxygen and the amide nitrogen (Lehninger, Fig. 4-2a). The oxygen has a partial negative charge and the nitrogen a partial positive charge, setting up a small electric dipole. The six atoms of the peptide group lie in a single plane, with the oxygen atom of the carbonyl group and the hydrogen atom of the amide nitrogen trans to each other. From these findings Pauling and Corey concluded that the peptide C-N bonds are unable to rotate freely because of their partial double bond character. Rotation is permitted about the N-Cα and the Cα-C bonds. The backbone of a polypeptide chain can thus be pictured as a series of rigid planes with consecutive planes sharing a common point of rotation at C α (Lehninger, Fig. 4-2b). The rigid peptide bonds limit the range of conformation that can be assumed by a polypeptide chain.

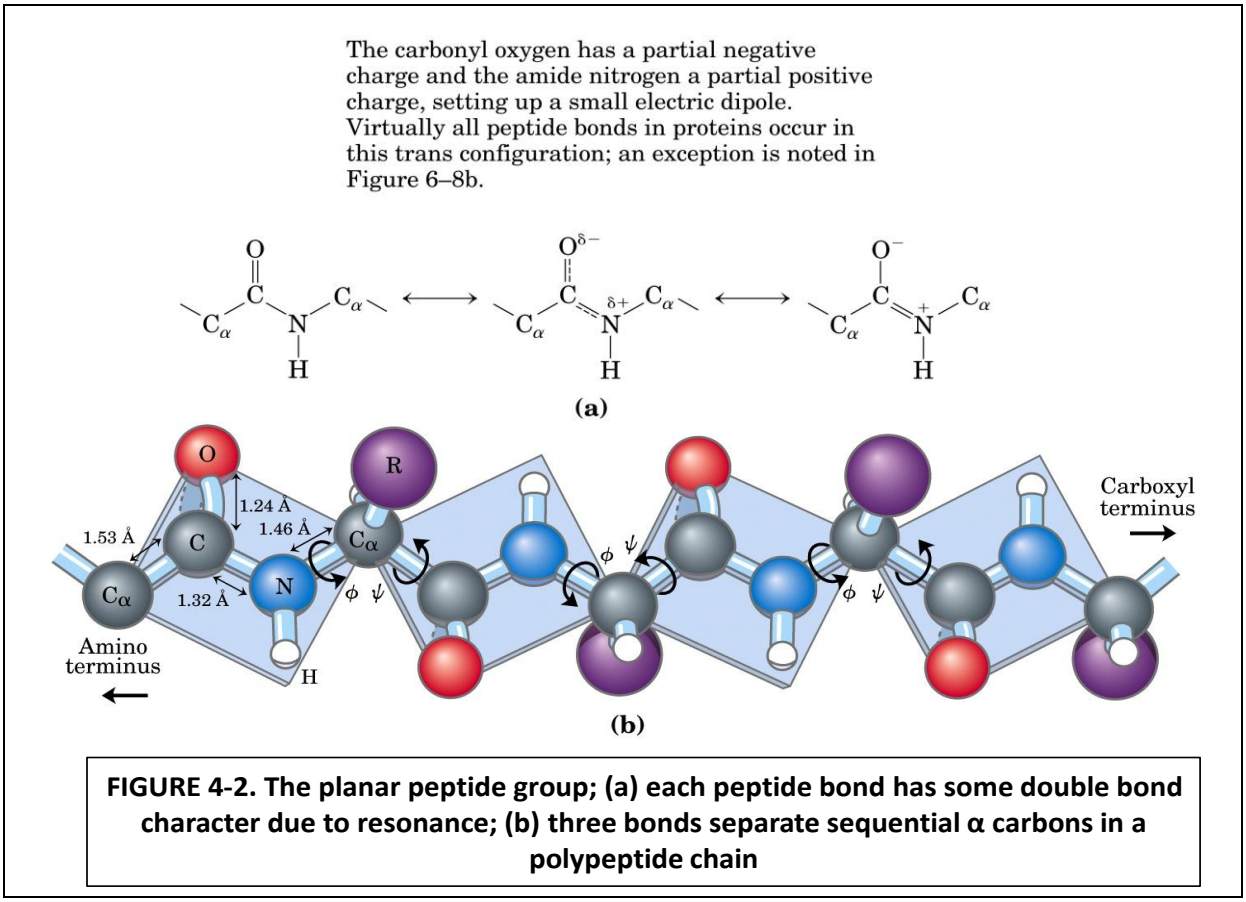


FIGURE 4-2. The planar peptide group; (a) each peptide bond has some double bond character due to resonance; (b) three bonds separate sequential α carbons in a polypeptide chain

Protein Secondary Structure

The term secondary structure refers to the local conformation of some part of the polypeptide. A few types of secondary structure are particularly stable and occur widely in proteins. The most prominent are the α helix and β conformation.

α Helix

The protein that makes up hair (the fibrous protein α–keratin) has a regular structure that repeats every 5.4A^{o} (A^{o} = 0.1nm). The simplest arrangement the polypeptide chain could assume with the rigid peptide bonds is a helical structure (α-helix). In this structure the polypeptide backbone is tightly wound around an imaginary axis drawn longitudinally through the middle of the helix, and the R groups of the amino acid residues protrude outward from the helical backbone (Lehninger, Fig. 4-4). Each helical turn includes 3.6 amino acid residues. The helical twist of the α helix proved to be the predominant structure in α keratins.

The α helix form more readily because the α-helix makes optimal use of internal hydrogen bonds. The structure is stabilized by a hydrogen bond between the hydrogen atom attached to the electronegative nitrogen atom and the electronegative carbonyl oxygen atom of the fourth amino acid on the N-terminal side of that peptide bond (Lehninger, Fig. 4-4b).

A long block of Glu residues, Lys and/or Arg residues prevent formation of the α-helix. The bulk and shape of Asn, Ser, Thr and Gys residues can also destabilize an α–helix if they are close together in the chain. A constraint on the conformation of α–helix is the presence of Pro or Gly residues.

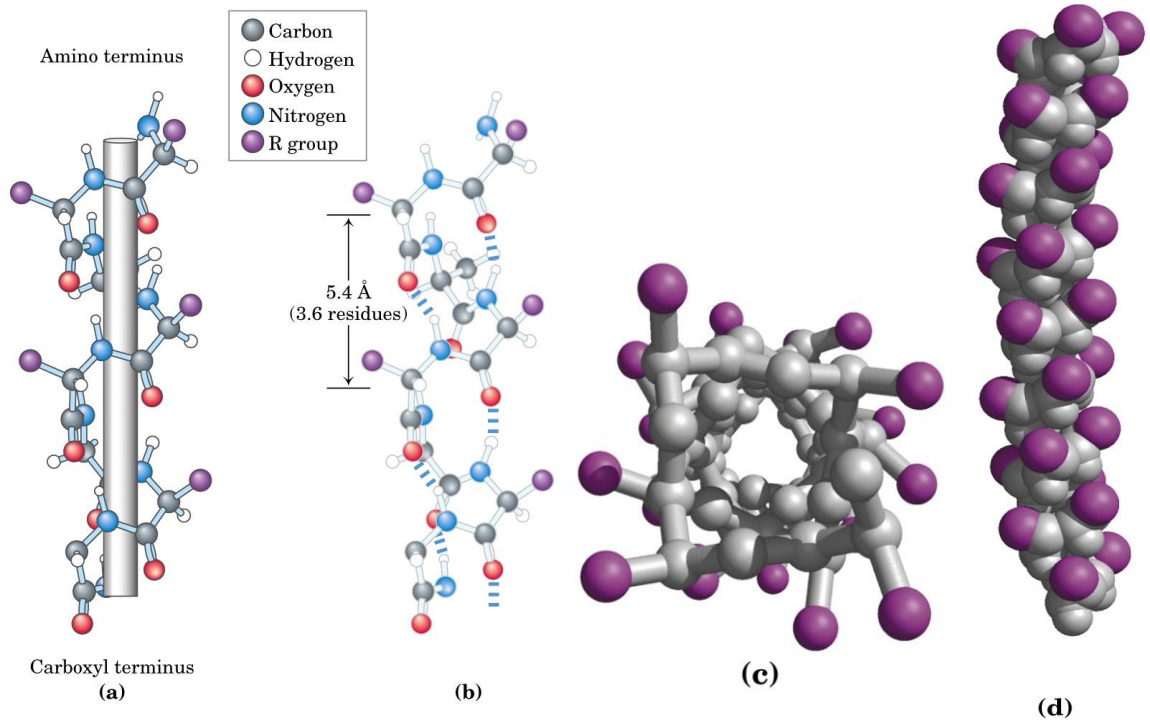


FIGURE 4-4. Four models of the α helix, showing different aspects of its structure

β-Sheet

The **β-conformation** is a more extended conformation of polypeptide chains. In the β-conformation, the backbone of the polypeptide chain is extended into a zigzag rather than helical structure (Lehninger, Fig. 4-7). The zigzag polypeptide chains can be arranged side by side to form a structure resembling a series of pleats (طيّة). In this arrangement, called a **β-Sheet** hydrogen bonds are formed between adjacent segments of polypeptide chain. The individual segments that form a β-sheet are usually nearby on the polypeptide chain, but can also be quite distant from each other in the linear sequence of the polypeptide; they may even be segments in different polypeptide chain. The R group of adjacent amino acids protrudes from the zigzag structure in opposite directions.

The adjacent polypeptide chains in a β-sheet can be either parallel or antiparallel (having the same or opposite amino-to-carboxyl orientation, respectively). β-keratins such as silk fibroin and the fibroin of spider webs have a very high content of Gly and Ala residues (the two amino acids with the smallest R groups).

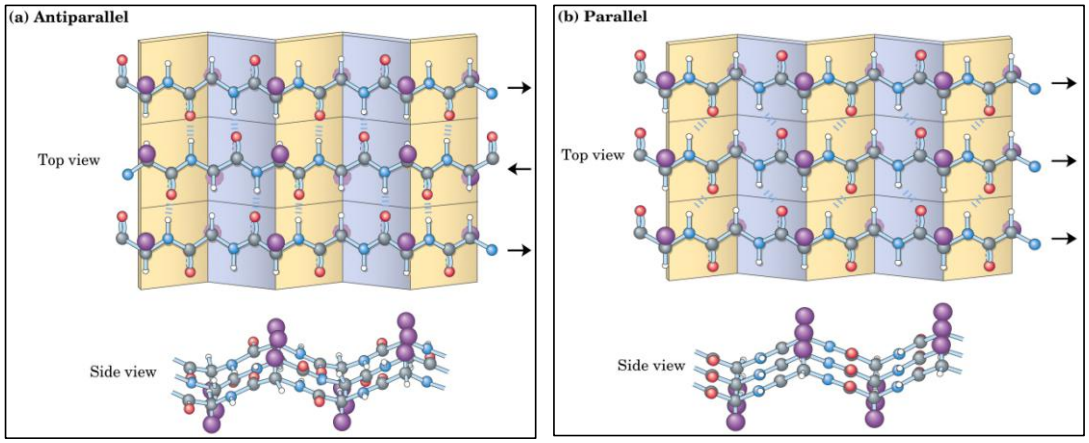


FIGURE 4-7. The β conformation of polypeptide chains: (a) antiparallel and (b) parallel

β-Turn

In globular proteins, which have a compact folded structure, nearly ⅓ of the amino acid residues are in turns or loops where the polypeptide chain reverses direction. These are the connecting elements that link successive runs of α-helix or β-conformation. Particularly common are **β turns** that connect the ends of two adjacent segments of an antiparallel β sheet. Gly and Pro residues often occur in β turns.

Protein Tertiary and Quaternary Structures

The overall three-dimensional arrangement of all atoms in a protein is referred to as the protein's **tertiary structure**. Amino acids that are far apart in the polypeptide sequence and that reside in different types of secondary structure may interact within the completely folded structure of a protein. The location of bends in the polypeptide chain and the direction and angle of these bends are determined by the number and location of specific bend-producing residues, such as Pro, Thr, and Gly. Interacting segments of polypeptide chains are held in their characteristic tertiary positions by different kinds of weak bonding interactions (and sometime by covalent bonds such as disulfide cross-links) between the segments.

Some proteins contain two or more separate polypeptide chains, or subunits, which may be identical or different. The arrangement of these protein subunits in three-dimensional complexes constitutes **quaternary structure**.

In considering these higher levels of structure, proteins are classified into two major groups, fibrous proteins, having polypeptide chains arranged in long strands or sheets, and globular proteins, having polypeptide chains folded into a spherical or globular shape. The two groups are structurally distinct: fibrous proteins usually consist largely of a single type of secondary structure; globular proteins often contain several types of secondary structures. The two groups differ functionally in that structures that provide support, shape and external protection to vertebrates are made of fibrous proteins, whereas most enzymes and regulatory proteins are globular proteins.

Fibrous Proteins

Fibrous Proteins such as α-keratin, collagen and silk fibroin nicely illustrate the relationship between protein structure and biological function (Lehninger, Table 4-1). Fibrous proteins share properties that give strength and /or flexibility to the structures in which they occur. In each case, the fundamental structural unit is a simple repeating element of secondary structures. All fibrous proteins are insoluble in water, a property conferred by a concentration of hydrophobic amino acids.

TABLE 4-1. Secondary Structure and Properties of Fibrous Proteins

Secondary Structures and Properties of Fibrous Proteins

Structure	Characteristics	Examples of occurrence
α Helix, cross-linked by disulfide bonds	Tough, insoluble protective structures of varying hardness and flexibility	α-Keratin of hair, feathers, and nails
β Conformation	Soft, flexible filaments	Silk fibroin
Collagen triple helix	High tensile strength, without stretch	Collagen of tendons, bone matrix

α -Keratin

The α–keratin provide strength, found in mammals, these proteins constitute almost the entire dry weight of hair, wool, nails, claws, horns and much of the outer layer of skin.

The α–keratin helix is a right–handed α–helix. Two strands of α–keratin, oriented in parallel are wrapped about each other to form a super twisted coiled coil. The supertwisting amplifies the strength of the overall structure (Lehninger, Fig. 4-11). The helical path of the supertwist is left-handed. The intertwining of the two α-helical polypeptides is an example of quaternary structure. Coiled coils of this type are common structural elements in filamentous proteins and in muscle protein myosin.

Many coiled coils can be assembled into large supramolecular complexes, such as the arrangement of α–keratin to form intermediate filament of hair (Lehninger, Fig. 4.11b). The strength of fibrous proteins is enhanced by covalent cross-links (-S-S- bonds). In the hardest and toughest α–Keratin, such as those of rhinoceros horn up to 18% of the residues are cysteines involved in disulfide bonds.

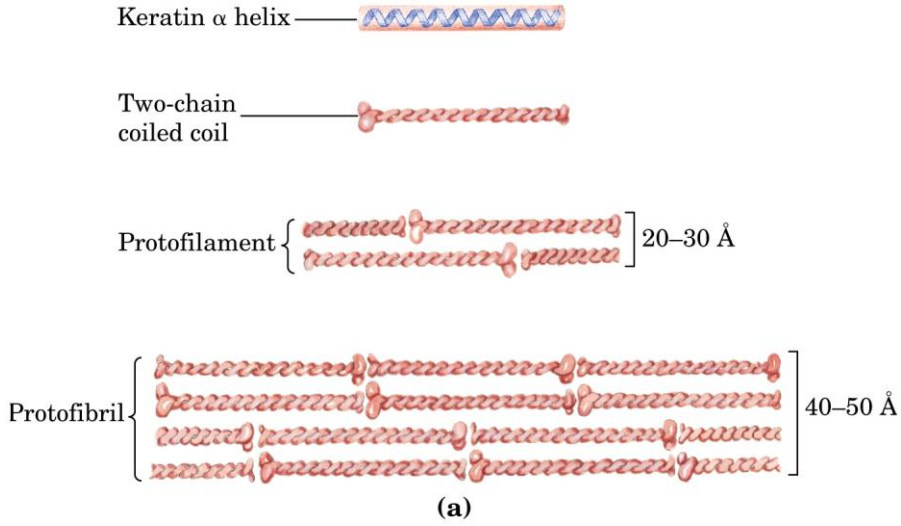


FIGURE 4-11. Structure of hair: (a) hair α-keratin is an elongated α-helix

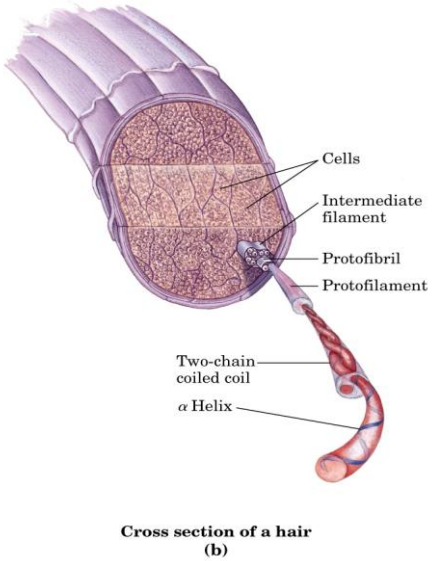


FIGURE 4-11. Structure of hair: (b) a hair is an array of many α-keratin filaments

Collagen

Like α–keratin, collagen provides strength. It is found in connective tissue such as tendons, cartilage, and the organic matrix of bone and the cornea of the eye. The collagen helix is a left handed and has three amino acids per turn. Collagen is also a coiled coil. Three separate polypeptides, called α-chains, are supertwisted about each other. The suprahelical twisting is right-handed in collagen. Collagen typically contains ~35% Gly, 11% Ala, 21% Pro and hydroxyproline (Hyp).

The tight wrapping of the α-chains in the collagen triple helix provides tensile strength greater than that of a steel wire of equal cross section. The increasingly rigid and brittle character of aging connective tissue results from accumulated covalent cross-links in collagen fibrils.

The amino acid sequence in collagen is generally a repeating tripeptide unit, Gly-X-Y, where X is often Pro and Y is often 4-Hyp. Genetic defects in collagen structure are catastrophic and lethal. Gly can't be replaced by another amino acid residue without disrupting collagen structure and function.

Silk Fibroin

Fibroin, the protein of silk, is produced by insects and spiders. Its polypeptide chains are predominantly in the β conformation. Fibroin is rich in Ala and Gly residues (Lehninger, Fig. 4-14). The overall structure is stabilized by extensive H-bonding. Silk does not stretch, because the β conformation is already highly extended. However, the structure is flexible because the sheets are held together by numerous weak interactions rather than by covalent bonds such as the disulfide bonds in α-keratins.

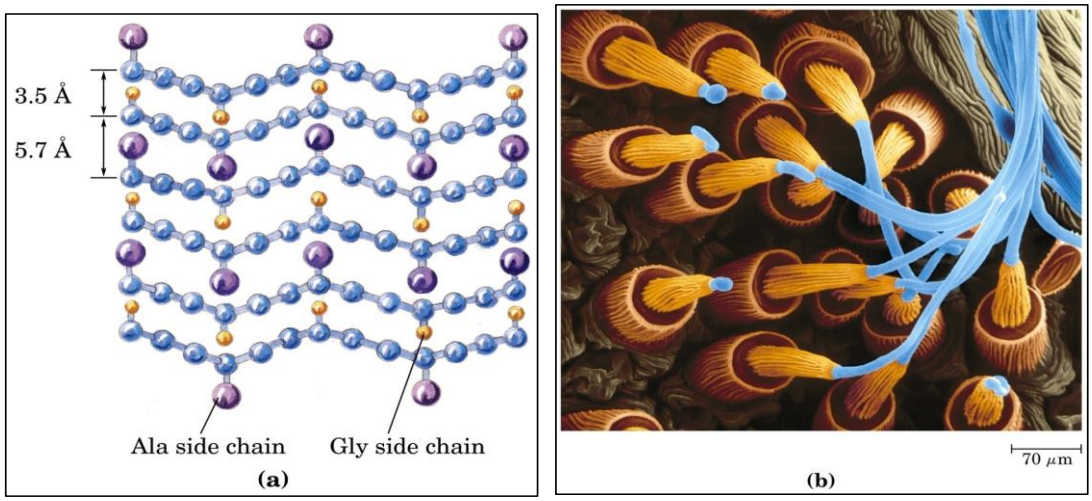


FIGURE 4-14. Structure of silk: (a) fibroin consists of layers of antiparallel β sheet rich in Ala; (b) strands of fibroin emerge from the spinnerets of a spider

Globular Proteins

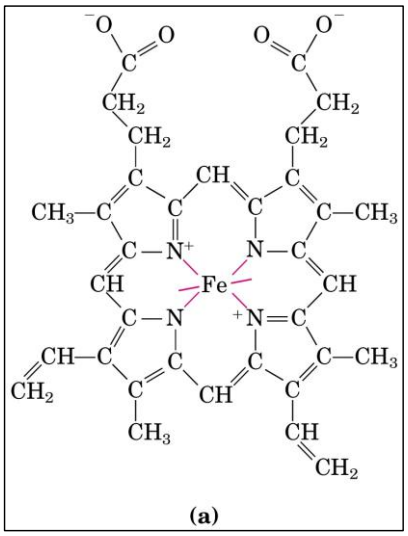
In a globular protein, different segments of a polypeptide chain fold back on each other (Lehninger, Fig. 4-15). This folding generates a compact form relative to polypeptides in a fully extended conformation. The folding also provides the structural diversity necessary for proteins to carry out a wide array of biological functions. Globular proteins include enzymes, transport proteins, motor proteins, regulatory proteins, immunoglobulins and proteins with many other functions.

FIGURE 4-15. Globular protein structures are compact and varied

Myoglobin is a relatively small (M_r 16,700), oxygen-binding protein of muscle cells. It functions both to store oxygen and to facilitate oxygen diffusion in rapidly contracting muscle tissue. Myoglobin contains a single polypeptide chain of 153 amino acid residues and a single iron protoporphyrin, or heme groups. The same heme group is found in hemoglobin, the oxygen-binding protein of erythrocytes. Myoglobin is particularly abundant in the muscles of diving mammals such as whales and seals, that permits these animals to remain submerged for long periods of time.

The backbone of myoglobin is made up of 8 relatively straight segments of α helix, interrupted by bends, some of which are β terns. More than 70 % of the residues in myoglobin are in these α-helical regions. Hydrophobic interactions provide stability to myoglobin. Most of the hydrophobic R groups are in the interior of myoglobin molecule.

The flat heme group rests in a crevice, or pocket, in the myoglobin molecule. The iron atom in the center of the heme group has two bonding (coordination) positions perpendicular to the plane of the heme (Lehninger, Fig. 4-17). One of these is bound to the R group of the His residue at position 93, the other is the site at which an O_2 molecule binds. Within the pocket, the accessibility of the heme group to solvent is highly restricted. This is important for function, because free heme groups in an oxygenated solution are rapidly oxidized from the Fe^{2+} from to Fe^{3+} form which does not bind O_2.



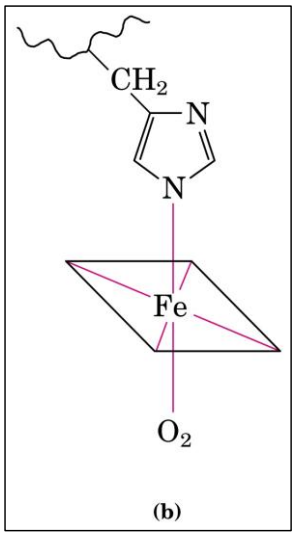


FIGURE 4-17. The heme group: (a) heme cosists of protoporphyrin to which is bound an iron atom in ferrous state; (b) in myoglobin and hemoglobin one of the bonds is bound to a nitrogen atom of a His residue

Common Structural Pattern of Globular Proteins

Supersecondary structures, also called **motifs** or simply **folds**, are particularly stable arrangements of several elements of secondary structure and the connections between them. Polypeptides with more than a few hundred amino acid residues often fold into two or more stable, globular units called **domains**. Different domains often have distinct functions, such as the binding of small molecules or interaction with other proteins.

Protein Quaternary Structure

Many proteins have multiple polypeptide subunits. The association of polypeptide chains can serve a variety of functions. Many multisubunit proteins have regulatory roles; the binding of small molecules may affect the interaction between subunits, causing large changes in the protein's activity in response to small changes in the concentration of substrate or regulatory molecules.

In other cases, separate subunits can take on separate but related functions, such as catalysis or regulation. Some associations (fibrous proteins or viral coat proteins) serve primarily structural roles. Some very large assemblies are the site of complex, multistep reactions. One example is the ribosome site of protein synthesis, which incorporates dozens of protein subunits along with a number of RNA molecules.

A multi subunit protein is also referred to as a **multimer**. Multimeric protein can have from two to hundreds of subunits. A multimer with just a few subunits is often called an **oligomer**. Hemoglobin (M_r 64,500) contains 4 polypeptide chains and 4 heme prosthetic groups, in which the iron atoms are in the ferrous (Fe^{2+}) state. The protein portion is called globin and consists of two α-chains and two β chains.

Protein Folding

All proteins begin as a linear sequence of amino acid residues. This polypeptide must fold during and following synthesis to take up its native conformation. A loss of three-dimensional structure sufficient to cause loss of function is called **denaturation**. Most proteins can be denatured by heat which affects the weak interactions in a protein primarily hydrogen bonds. Other denaturing factors are extreme pH, certain organic solvents, urea and guanidine HCl or by detergents such as SDS. Denaturation of some proteins is reversible (renaturation) which proofs that the tertiary structure of proteins is determined by the amino acid sequence.

A classic example is the denaturation and renaturation of ribonuclease. Denaturation by urea and reducing agent causes a complete loss of catalytic activity. When urea and reducing agent are removed, the randomly coiled, denatured ribonuclease refolds into its correct tertiary structure (Lehninger, Fig. 4-27) with full restoration of catalytic activity.

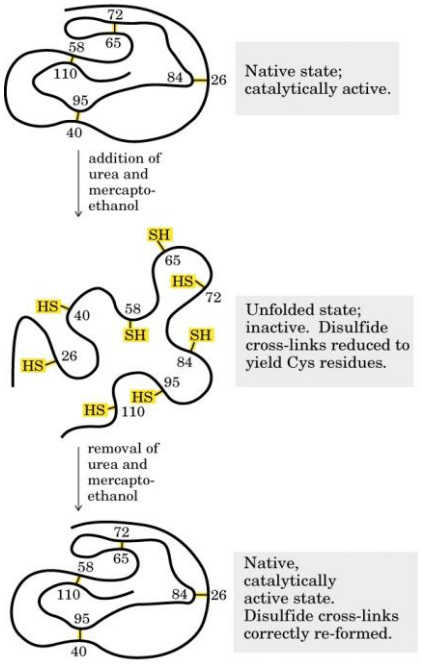


FIGURE 4-27. Renaturation of unfolded, denatured ribonuclease

In living cells, proteins are assembled from amino acids at a very high rate. The folding pathway is complicated and several models have been developed. In one, the folding process is envisioned as hierarchical. Local secondary structures form first. Certain amino acid sequences fold readily into α-helices or β-sheets. This is followed by longer-range interactions. The process continues until complete domains form and the entire polypeptide is folded.

Defects in protein folding may be the molecular basis for a wide range of human genetic disorders. For example, cystic fibrosis is caused by defects in a membrane-bound protein called cystic fibrosis transmembrane conductance regulator (CFTR), which acts as a channel for chloride ions. A deletion of Phe at position 508 in CFTR causes improper protein folding. Many of the disease related mutations in collagen also cause defective folding.

Assisted folding

Not all proteins fold spontaneously as they are synthesized in the cell. Folding is facilitated by the action of specialized proteins called **molecular chaperones**. They interact with partially folded or improperly folded polypeptides, facilitating correct folding pathways or provide microenvironments in which folding can occur. Example of these chaperones are a family of proteins called **Hsp 70** (heat shock protein) of M_r 70,000.

The second classes of chaperones are called **chaperonins** like the two proteins found in E.cole Gro EL and Gro ES (Lehninger, Fig. 4-31).

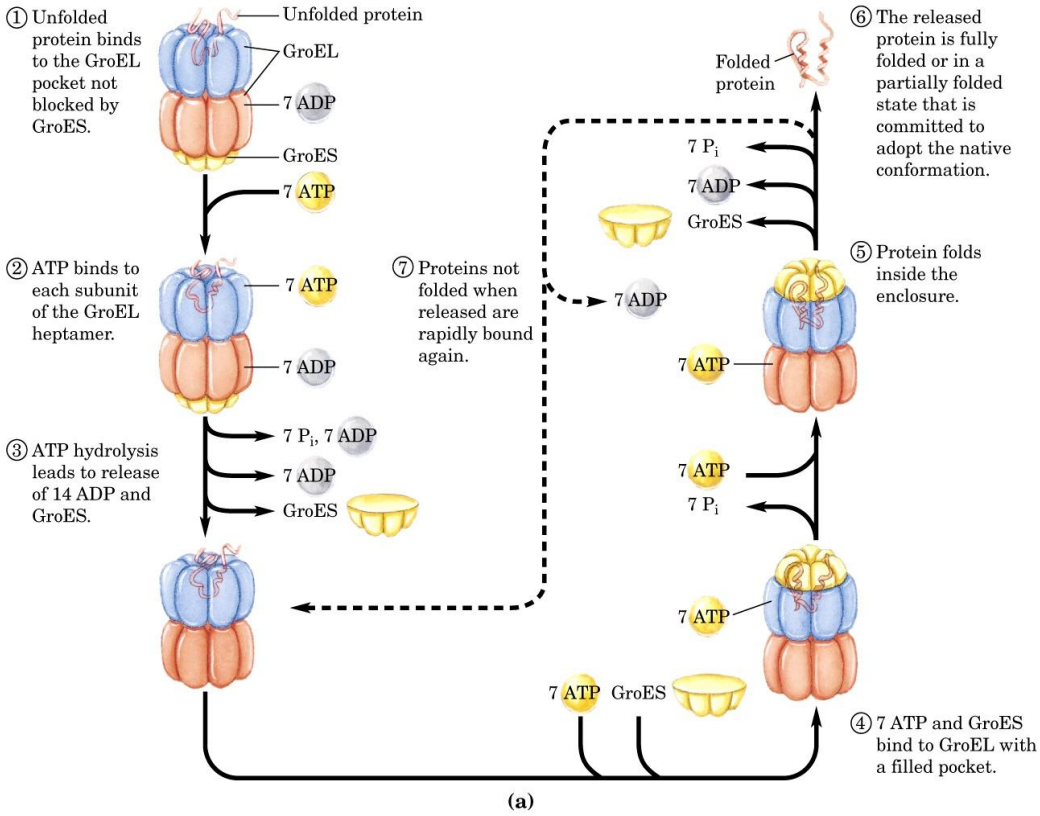


FIGURE 4-31. Chaperonins in protein folding

Finally, the folding pathways of a number of proteins require two enzymes that catalyze isomerization reactions. **Protein Disulfide Isomerase (PDI)** catalyzes the interchange or shuffling of disulfide bonds until the bonds of the native conformation are formed. Among its functions, PDI catalyzes the elimination of folding intermediates with inappropriate disulfide cross-links.

PROTEIN FUNCTION

Proteins are dynamic molecules whose functions depend on interaction with other molecules, and these interactions are affected by changes in protein conformation. The function of many proteins involves the reversible binding of other molecules (**ligands**). The transient nature of protein ligand interactions is critical to life allowing an organism to respond rapidly and reversibly to changing environmental and metabolic circumstances.

A ligand binds at a site on the protein called the **binding sites** which is complementary to the ligand in size, shape, charge, and hydrophilic or hydrophobic character. This interaction is specific: the protein can discriminate among the thousands of different molecules in its environment and selectively bind only one or a few. A given protein may have separate binding sites for several different ligands.

The binding of a protein and ligand is often coupled to a conformational change in the protein that makes the binding site more complementary to the ligand, permitting tighter binding. The structural adaptation that occurs between protein and ligand is called **induced-fit.** In a multisubunit protein, a conformational change in one subunit often affects the conformation of other subunits. Interactions between ligands and proteins may be regulated, usually through specific interactions with one or more additional ligands. These other ligands may cause conformational changes in the protein that affect the binding of the first ligand.

Enzymes represent a special case of protein function. Enzymes bind and chemically transform other molecules, they catalyze reactions. The molecules acted upon by enzymes are called reaction **substrates** rather than ligands, and the ligand-binding site is called the **catalytic site** or **active site**.

Reversible Binding of a Protein to a Ligand: Oxygen-Binding Proteins

Oxygen is poorly soluble in aqueous solutions and can't be carried to tissues in sufficient quantity if it is simply dissolved in blood serum. Diffusion of oxygen through tissues is also ineffective over distances greater than a few mm. Multicellular animals have proteins that transport and store oxygen. None of the amino acid side chains in proteins is suited for the reversible binding of oxygen molecules. This role is filled by certain transition metals, among them Fe and Cu, that have a great tendency to bind O_2.

However, free Fe promotes the formation of highly reactive O_2 species such as hydroxyl radicals that can damage DNA and other macromolecules. Iron is often incorporated into a protein bound prosthetic group called **heme** (a prosthetic group is a compound permanently associated with a protein that contributes to the protein's function).

Heme consists of a complex organic ring structure, **protoporphyrin**, to which is bound a single iron atom in its ferrous (Fe^{2+}) state (Lehninger, Fig. 5-1). The Fe atom has six coordination bonds, four to nitrogen atoms that are part of the flat **porphyrin ring** system. The other two coordination bonds are occupied by a side-chain nitrogen of a His residue and O_2 (Lehninger, Fig. 5-2). When O_2 binds, the electronic properties of heme iron change; these account for the change in color from the dark purple of oxygen-depleted venous blood to the bright red of oxygen-rich arterial blood. Some small molecules, such as CO and NO, coordinate to heme Fe with greater affinity than does O_2.

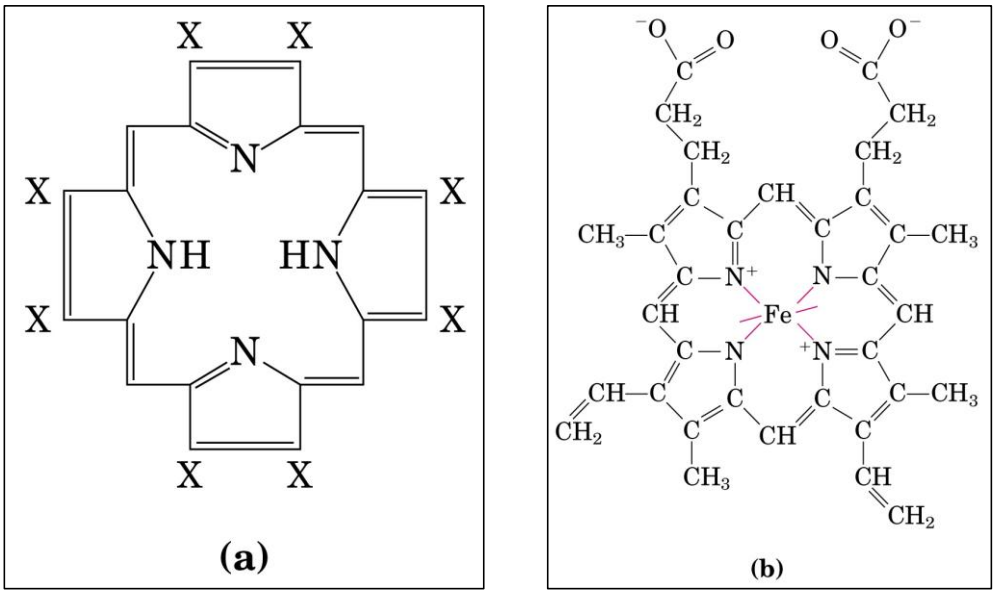


FIGURE 5-1. Heme: (a) porphyrins; (b) the iron atom of heme has six coordination bonds

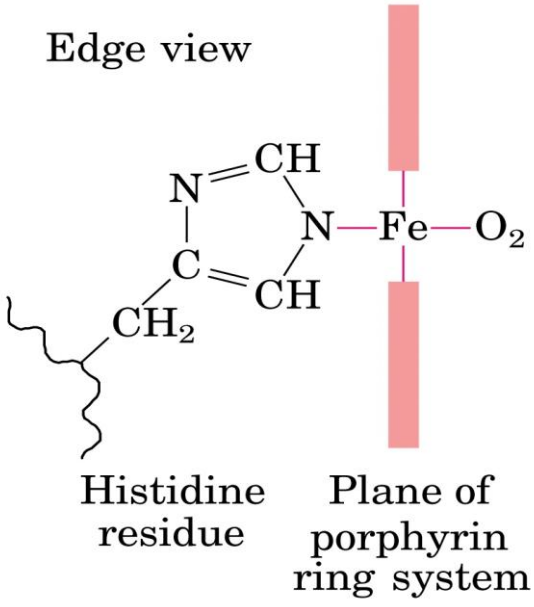


FIGURE 5-2. The heme group viewed from the side

Myoglobin

Myoglobin (Mb, M_r 16,700) is an O_2-binding protein found primarily in muscle tissue. It facilitates O_2 diffusion in muscle. Myoglobin is particularly abundant in the muscles of diving mammals such as seals and whales, where it also has an O_2-storage function.

Myoglobin is a single polypeptide of 153 amino acid residues with one molecule of heme. It is typical of the family of proteins called globins, all of which have similar primary and tertiary structures. The polypeptide is made up of 8 α-helical segments connected by bends (Lehninger, Fig. 5-3).

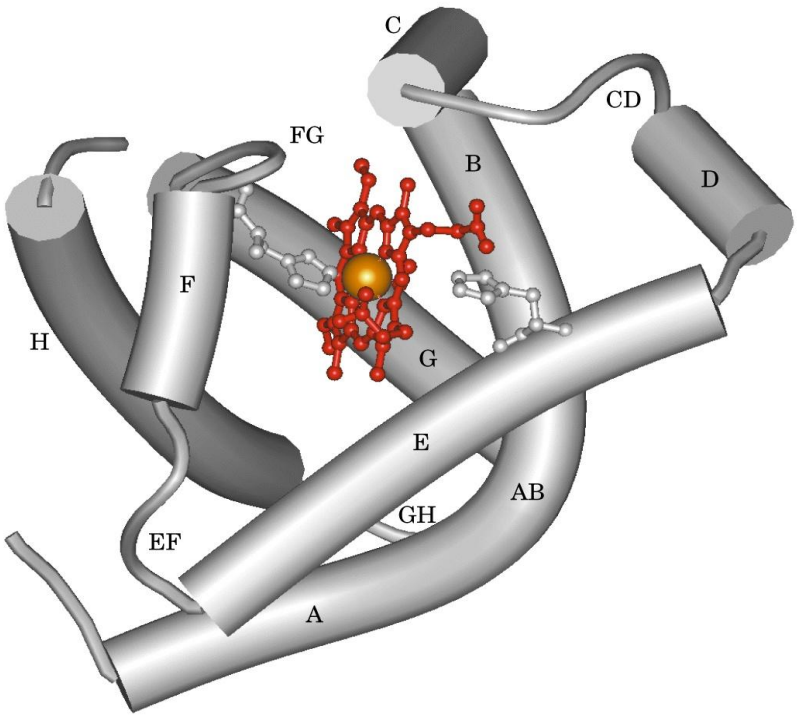


FIGURE 5-3. The structure of myoglobin

Protein-Ligand Interaction

The function of myoglobin depends on the protein's ability not only to bind oxygen but also to release it when and where it is needed.

In general, the reversible binding of a protein (P) to a ligand (L) can be described by a simple equilibrium expression:

$$\mathbf{P + L} \rightleftharpoons \mathbf{PL}$$

$$K_a = \frac{[PL]}{[P][L]}$$

The term K_a is an association constant that provides a measure of the affinity of the ligand L for the protein. K_a has units of M^{-1}. It is more common to consider the **dissociation constant**, K_d, which is the reciprocal of K_a ($K_d=1/K_a$). K_d is equivalent to the molar concentration of ligand at which half of the available ligand-binding sites are occupied.

$$\theta(theta) = \frac{binding\ \ sites\ occupied}{total\ \ binding\ \ sites} = \frac{[PL]}{[PL]+[P]}$$

$$\theta = \frac{[L]}{[L]+\frac{1}{K_a}}$$

$$Kd = \frac{[P][L]}{[PL]}$$

$$\theta = \frac{[L]}{[L]+K_d}$$

The binding of O_2 to myoglobin

Because O_2 is gas, the concentration of oxygen is expressed as the partial pressure of oxygen. Using oxygen as a ligand, the partial pressure of oxygen, pO_2, is a measure of oxygen concentration. P_{50} is the oxygen concentration that equals K_d.

$$\theta = \frac{pO_2}{pO_2 + P_{50}}$$

A binding curve for myoglobin that relates θ to pO_2 is shown in Figure 5-4b.

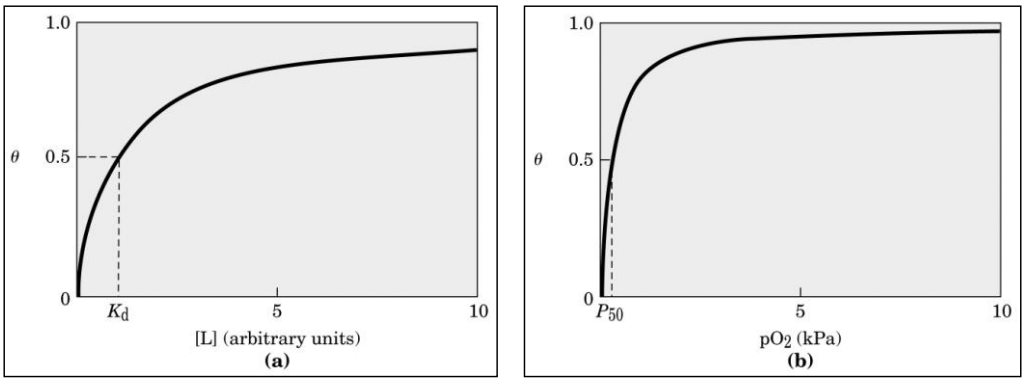


Figure 5-4. Graphical representations of ligand binding: (a) A hypothetical binding curve for a ligand L; (b) A curve describing the binding of oxygen to myoglobin.

Protein Structure Affects How Ligands Bind

The interaction or ligands to proteins is greatly affected by protein structure and is often accompanied by conformational changes. For example, the specificity with which heme binds its various ligands is altered when heme is a component of myoglobin. Carbon monoxide binds to free heme molecules more than 20,000 times better than does O_2 (that is, the K_d or P_{50} for CO binding to free heme is more than 20,000 times lower than that for O_2, but it binds only about 200 times better when the heme is bound in myoglobin. The difference may be partly explained by steric hindrance.

Oxygen Is Transported in Blood by Hemoglobin

Nearly all the oxygen carried by whole blood in animals is bound and transported by hemoglobin in erythrocytes (red blood cells). Normal human erythrocytes are small (6 to 9 μm in diameter), biconcave disks. They are formed from precursor stem cells called hemocytoblasts. In the maturation process, the stem cell produces daughter cells that form large amounts of hemoglobin and then lose their intracellular organelles-nucleus, mitochondria, and endoplasmic reticulum. Erythrocytes are unable to reproduce and, in humans, destined to survive for only about 120 days. Their main function is to carry hemoglobin, which is dissolved in the cytosol at a very high concentration (~34% by weight).

In arterial blood passing from the lungs through the heart to the peripheral tissues, hemoglobin is about 96% saturated with oxygen. In the venous blood returning to the heart, hemoglobin is only about 64% saturated. Thus; each 100 mL of blood passing through a tissue releases about one-third of the oxygen it carries.

Myoglobin, with its hyperbolic binding curve for oxygen (Fig. 5-4b), is relatively insensitive to small changes in the concentration of dissolved oxygen and so functions well as an oxygen-storage protein. Hemoglobin, with its multiple subunits and O_2-binding sites, is better suited to oxygen transport. As we shall see, interactions between the subunits of a multimeric protein can permit a highly sensitive response to small changes in ligand concentration. Interactions among the subunits in hemoglobin cause conformational changes that alter the affinity of the protein for oxygen. The modulation of oxygen binding allows the O_2-transport protein to respond to changes in oxygen demand by tissues.

Hemoglobin Subunits Are Structurally Similar to Myoglobin

Hemoglobin (M_r 64,500; abbreviated Hb) is roughly spherical. It is a tetrameric protein containing four heme prosthetic groups, one associated with each polypeptide chain. Adult hemoglobin contains two types of globin, two α chains (141 residues each) and two β chains (146 residues each). Although fewer than half of the amino acid residues in the polypeptide sequences of the α and β subunits are identical, the three-dimensional structures of the two types of subunits are very similar. Furthermore, their structures are very similar to that of myoglobin, even though the amino acid sequences of the three polypeptides are identical at only 27 positions. All three polypeptides are members of the globin family of proteins.

The quaternary structure of hemoglobin features strong interactions between unlike subunits. The $\alpha_1\beta_1$ interface (and its $\alpha_2\beta_2$ counterpart) involves more than 30 residues, and its interaction is sufficiently strong. Hydrophobic interactions predominate at the interfaces, but

there are also many hydrogen bonds and a few ion pairs (sometimes referred to as salt bridges).

Hemoglobin Undergoes a Structural Change on Binding Oxygen

X-ray analysis has revealed two major conformations of hemoglobin: the R state and the T state. Although oxygen binds to hemoglobin in either state, it has a significantly higher affinity for hemoglobin in the R state. Oxygen binding stabilizes the R state. When oxygen is absent experimentally, the T state is more stable and is thus the predominant conformation of deoxyhemoglobin. T and R originally denoted "tense" and "relaxed," respectively, because the T state is stabilized by a greater number of ion pairs. The binding of O_2 to a hemoglobin subunit in the T state triggers a change in conformation to the R state. When the entire protein undergoes this transition, the structures of the individual subunits change little, but the αβ subunit pairs slide past each other and rotate, narrowing the pocket between the β subunits. In this process, some of the ion pairs that stabilize the T state are broken and some new ones are formed.

Hemoglobin Binds Oxygen Cooperatively

Hemoglobin must bind oxygen efficiently in the lungs, where the pO_2 is about 13.3 kPa, and release oxygen in the tissues, where the pO_2 is about 4 kPa. Myoglobin, or any protein that binds oxygen with a hyperbolic binding curve, would be ill-suited to this function, for the reason illustrated in Figure 5-12. A protein that binds O_2 with high affinity would bind it efficiently in the lungs but would not release much of it in the tissues. If the protein bound oxygen with a sufficiently low affinity to release it in the tissues, it would not pick up much oxygen in the lungs.

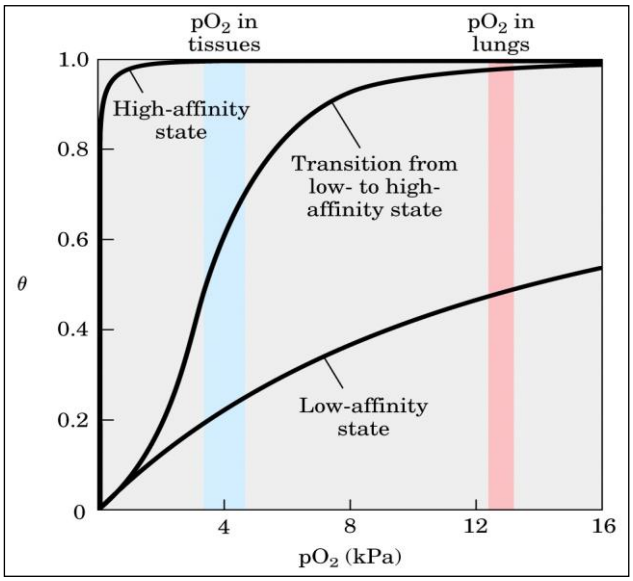


FIGURE 5-12. A sigmoid (cooperative) binding curve

Hemoglobin solves the problem by undergoing a transition from a low-affinity state (the T state) to a high-affinity state (the R state) as more O_2 molecules are bound. As a result, hemoglobin has a hybrid S shaped, or sigmoid, binding curve for oxygen (Fig. 5-12). A

single-subunit protein with a single ligand binding site cannot produce a sigmoid binding curve even if binding elicits a conformational change because each molecule of ligand binds independently and cannot affect the binding of another molecule. In contrast, O_2 binding to individual subunits of hemoglobin can alter the affinity for O_2 in adjacent subunits. The first molecule of O_2 that interacts with deoxyhemoglobin binds weakly, because it binds to a subunit in the T state. Its binding, however, leads to conformational changes that are communicated to adjacent subunits, making it easier for additional molecules of O_2 to bind. In effect, the T → R transition occurs more readily in the second subunit once O_2 is bound to the first subunit. The last (fourth) O_2 molecule binds to a heme in a subunit that is already in the R state, and hence it binds with much higher affinity than the first molecule.

An **allosteric protein** is one in which the binding of a ligand to one site affects the binding properties of another site on the same protein. The term "allosteric" derives from the Greek *allos,* "other," and *stereos,* "solid" or "shape." Allosteric proteins are those having "other shapes," or conformations, induced by the binding of ligands referred to as modulators. The conformational changes induced by the modulator(s) interconvert more-active and less-active forms of the protein. The modulators for allosteric proteins may be either inhibitors or activators. When the normal ligand and modulator are identical, the interaction is termed **homotropic.** When the modulator is a molecule other than the normal ligand the interaction is **heterotropic.** Some proteins have two or more modulators and therefore can have both homotropic and heterotropic interactions.

Cooperative binding of a ligand to a multimeric protein, such as we observe with the binding of O_2 to hemoglobin, is a form of allosteric binding often observed in multimeric proteins. The binding of one ligand affects the affinities of any remaining unfilled binding sites, and O_2 can be considered as both a ligand and an activating homotropic modulator. There is only one binding site for O_2 on each subunit, so the allosteric effects giving rise to cooperativity are mediated by conformational changes transmitted from one subunit to another by subunit-subunit interactions. A sigmoid binding curve is diagnostic of cooperative binding. It permits a much more sensitive response to ligand concentration and is important to the function of many multisubunit proteins. The principle of allostery extends readily to regulatory enzymes.

Cooperative conformational changes depend on variations in the structural stability of different parts of a protein. The binding sites of an allosteric protein typically consist of stable segments in proximity to relatively unstable segments, with the latter capable of frequent changes in conformation or disorganized motion. When a ligand binds, the moving parts of the protein's binding site may be stabilized in a particular conformation, affecting the conformation of adjacent polypeptide subunits. If the entire binding site were highly stable, then few structural changes could occur in this site or be propagated to other parts of the protein when a ligand binds. As is the case with myoglobin, ligands other than oxygen can bind to hemoglobin. An important example than is carbon monoxide, which binds to hemoglobin about 250 times better than does oxygen. Human exposure to CO can have tragic consequences.

Two Models Suggest Mechanisms for Cooperative Binding

Biochemists now know a great deal about the T and R states of hemoglobin, but much remains to be learned about how the T → R transition occurs. Two models for the cooperative binding of ligands to proteins with multiple binding sites have greatly influenced thinking about this problem. The first model was proposed by Jacques **M**onod, Jeffries **W**yman, and Jean-Pierre **C**hangeux in 1965, and is called the **MWC model** or the **concerted**

model (Fig. 5-15a). The concerted model assumes that the subunits of a cooperatively binding protein are functionally identical, that each subunit can exist in (at least) two conformations, and that all subunits undergo the transition from one conformation to the other simultaneously. In this model, no protein has individual subunits in different conformations. The two conformations are in equilibrium. The ligand can bind to either conformation, but binds each with different affinity. Successive binding of ligand molecules to the low-affinity conformation (which is more stable in the absence of ligand) makes a transition to the high-affinity conformation more likely.

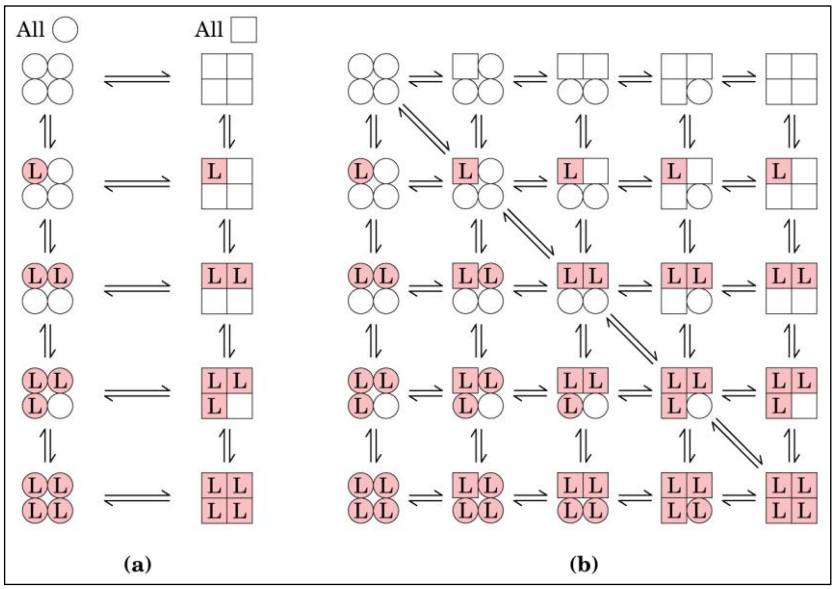


FIGURE 5-15. Two general models for the interconversion of inactive and active forms of cooperative ligand-binding protein: (a) concerted model; (b) sequential model

In the second model, the **sequential model** (Fig. 5-15b), proposed in 1966 by Daniel Koshland and colleagues, ligand binding can induce a change of conformation in an individual subunit. A conformational change in one subunit makes a similar change in an adjacent subunit, as well as the binding of a second ligand molecule, more likely. There are more potential intermediate states in this model than in the concerted model. The two models are not mutually exclusive; the concerted model may be viewed as the "all-or-none" limiting case of the sequential model.

Hemoglobin Also Transports H^+ and CO_2

In addition to carrying nearly all the oxygen required by cells from the lungs to the tissues, hemoglobin carries two end products of cellular respiration - H^+ and CO_2 from the tissues to the lungs and the kidneys, where they are excreted. The CO_2, produced by oxidation of organic fuels in mitochondria, is hydrated to form bicarbonate:

$$CO_2 + H_2O \rightarrow H^+ + HCO_3^-$$

This reaction is catalyzed by **carbonic anhydrase**, an enzyme particularly abundant in erythrocytes. Carbon dioxide is not very soluble in aqueous solution, and bubbles of CO_2 would form in the tissues and blood if it were not converted to bicarbonate. As you can see from the equation, the hydration of CO_2 results in an increase in the H^+

concentration (a decrease in pH) in the tissues. The binding of oxygen by hemoglobin is profoundly influenced by pH and CO_2 concentration, so the interconversion of CO_2 and bicarbonate is of great importance to the regulation of oxygen binding and release in the blood.

Hemoglobin transports about 40% of the total H^+ and 15% to 20% of the CO_2 formed in the tissues to the lungs and the kidneys. (The remainder of the H^+ is absorbed by the plasma's bicarbonate buffer; the remainder of the CO_2 is transported as dissolved HCO_3^- and CO_2). The binding of H^+ and CO_2 is inversely related to the binding of oxygen. At the relatively low pH and high CO_2 concentration of peripheral tissues, the affinity of hemoglobin for oxygen decreases as H^+ and CO_2 are bound, and O_2 is released to the tissues. Conversely, in the capillaries of the lung, as CO_2 is excreted and the blood pH consequently rises, the affinity of hemoglobin for oxygen increases and the protein binds more O_2 for transport to the peripheral tissues. This effect of pH and CO_2 concentration on the binding and release of oxygen by hemoglobin is called the **Bohr effect**.

The binding equilibrium for hemoglobin and one molecule of oxygen can be designated by the reaction:

$$Hb + O_2 \rightleftharpoons HbO_2$$

To account for the effect of H^+ concentration on this binding equilibrium we rewrite the reaction as

$$HHb^+ + O_2 \rightleftharpoons HbO_2 + H^+$$

where HHb^+ denotes a protonated form of hemoglobin. This equation tells us that the O_2-saturation curve of hemoglobin is influenced by the H^+ concentration (Fig. 5-16). Both O_2 and H^+ are bound by hemoglobin, but with inverse affinity. When the oxygen concentration is high, as in the lungs, hemoglobin binds O_2 and releases protons. When the oxygen concentration is low, as in the peripheral tissues, H^+ is bound and O_2 is released.

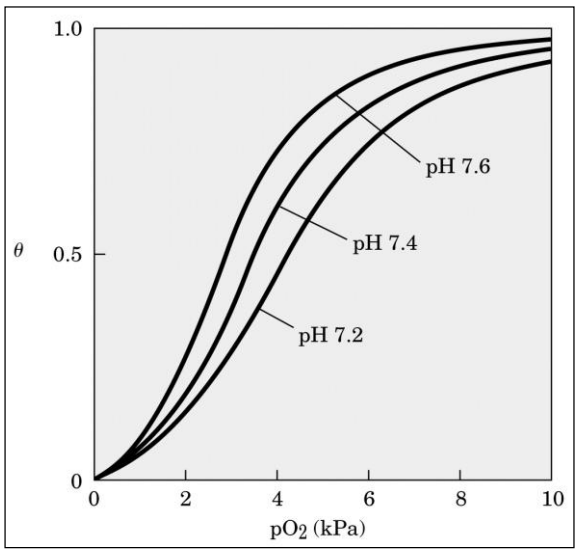


FIGURE 5-16. Effect of pH on the binding of oxygen to hemoglobin

When the concentration of carbon dioxide is high, as in peripheral tissues, some CO_2 binds to hemoglobin and the affinity for O_2 decreases, causing its release. Conversely, when hemoglobin reaches the lungs, the high oxygen concentration promotes binding of O_2 and release of CO_2, It is the capacity to communicate ligand-binding information from one polypeptide subunit to the others that makes the hemoglobin molecule so beautifully adapted to integrating the transport of O_2, CO_2, and H^+ by erythrocytes.

Oxygen Binding to Hemoglobin is Regulated by 2,3-Bisphosphoglycerate

The interaction of **2,3-bisphosphoglycerate (BPG)** with hemoglobin provides an example of heterotropic allosteric modulation. BPG is present in relatively high concentrations in erythrocytes. When hemoglobin is isolated, it contains substantial amounts of bound BPG, which can be difficult to remove completely. 2,3-Bisphosphoglycerate is known to greatly reduce the affinity of hemoglobin for oxygen -there is an inverse relationship between the binding of O_2 and the binding of BPG. We can therefore describe another binding process for hemoglobin:

$$HbBPG + O_2 \rightleftharpoons HbO_2 + BPG$$

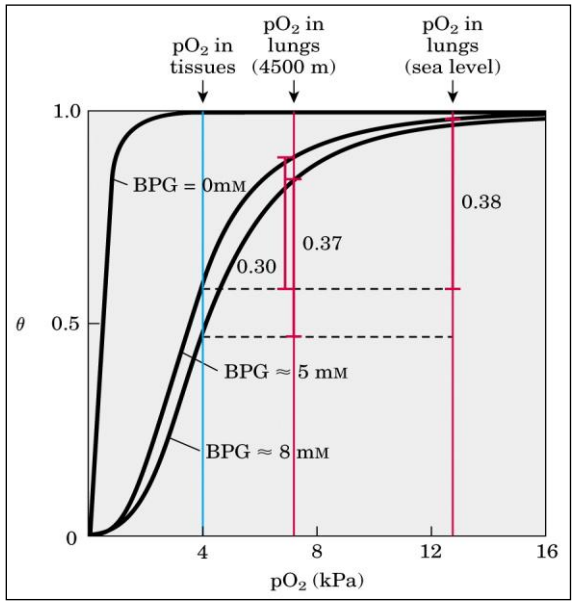


FIGURE 5-17. Effect of BPG on the binding of oxygen to hemoglobin

BPG binds at a site distant from the oxygen-binding site and regulates the O_2-binding affinity of hemoglobin in relation to the pO_2 in the lungs. BPG plays an important role in the physiological adaptation to the lower pO_2 available at high altitudes. For a healthy human at sea level, the binding of O_2 to hemoglobin is regulated such that the amount of O_2 delivered to the tissues is equivalent to nearly 40% of the maximum that could be carried by the blood (Fig. 5-17). Imagine that this person is quickly transported to a mountainside at an altitude (ارتفاع) of 4,500 meters, where the pO_2 is considerably lower. The delivery of O_2 to the tissues is now reduced. However, after just a few hours at the higher altitude, the BPG concentration in the blood has begun to rise, leading to a decrease in the affinity of hemoglobin for oxygen. This adjustment in the BPG level has only a small effect on the binding of O_2 in the lungs but a considerable effect on the release of O_2 in the tissues. As a result, the delivery of oxygen to the tissues is restored

to nearly 40% of that which can be transported by the blood. The situation is reversed when the person returns to sea level. The BPG concentration in erythrocytes also increases in people suffering from **hypoxia,** lowered oxygenation of peripheral tissues due to inadequate functioning of the lungs or circulatory system.

Unlike O_2, only one molecule of BPG is bound to each hemoglobin tetramer. BPG lowers hemoglobin's affinity for oxygen by stabilizing the T state. In the absence of BPG, hemoglobin is converted to the R state more easily.

Regulation of oxygen binding to hemoglobin by BPG has an important role in fetal development. Because a fetus must extract oxygen from its mother's blood, fetal hemoglobin must have greater affinity than the maternal hemoglobin for O_2. The fetus synthesizes γ subunits rather than β subunits, forming $\alpha_2\gamma_2$ hemoglobin. This tetramer has a much lower affinity for BPG than normal adult hemoglobin, and a correspondingly higher affinity for O_2.

Sickle-Cell Anemia Is a Molecular Disease of Hemoglobin

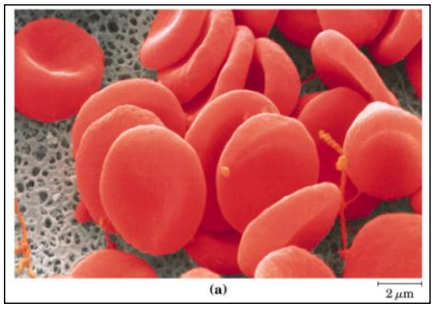
The great importance of the amino acid sequence in determining the secondary, tertiary, and quaternary structures of globular proteins, and thus their biological functions, is strikingly demonstrated by the hereditary human disease sickle-cell anemia. Almost 500 genetic variants of hemoglobin are known to occur in the human population; all but a few are quite rare. Most variations consist of differences in a single amino acid residue. The effects on hemoglobin structure and function are often minor but can sometimes be extraordinary. Each hemoglobin variation is the product of an altered gene. The variant genes are called alleles. Because humans generally have two copies of each gene, an individual may have two copies of one allele (thus being homozygous for that gene) or one copy of each of two different alleles (thus heterozygous).

Sickle-cell anemia is a genetic disease in which an individual has inherited the allele for sickle-cell hemoglobin from both parents. The erythrocytes of these individuals are fewer and also abnormal. In addition to an unusually large number of immature cells, the blood contains many long, thin, crescent-shaped erythrocytes that look like the blade of a sickle (Fig. 5-19). When hemoglobin from sickle cells (called hemoglobin S) is deoxygenated, it becomes insoluble and forms polymers that aggregate into tubular fibers. Normal hemoglobin (hemoglobin A) remains soluble on deoxygenation. The insoluble fibers of deoxygenated hemoglobin S are responsible for the deformed sickle shape of the erythrocytes, and the proportion of sickled cells increases greatly as blood is deoxygenated.

The altered properties of hemoglobin S result from a single amino acid substitution, a Val instead of a Glu residue at position 6 in the two β chains. The R group of valine has no electric charge, whereas glutamate has a negative charge at pH 7.4. Hemoglobin S therefore has two fewer negative charges than hemoglobin A, one for each of the two β chains. Replacement of the Glu residue by Val creates a "sticky" hydrophobic contact point at position 6 of the β chain, which is on the outer surface of the molecule. These sticky spots cause deoxyhemoglobin S molecules to associate abnormally with each other, forming the long, fibrous aggregates characteristic of this disorder.

Sickle-cell anemia occurs in individuals homozygous for the sickle-cell allele of the gene encoding the β subunit of hemoglobin. Individuals who receive the sickle-cell allele from only one parent and are thus heterozygous experience a milder condition called sickle-cell trait; only about 1% of their erythrocytes become sickled on deoxygenation. These

individuals may live completely normal lives if they avoid vigorous exercise or other stresses on the circulatory system.



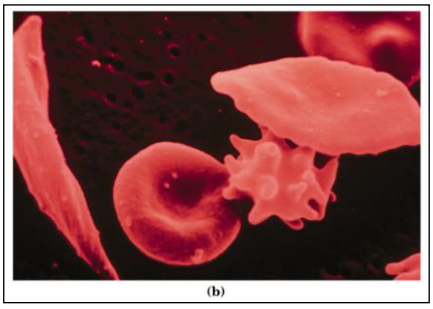


FIGURE 5-19. A comparison of normal erythrocytes (a) with erythrocytes seen in sickle-cell anemia (b)

Sickle-cell anemia is a life-threatening and painful disease. People with sickle-cell anemia suffer from repeated crises brought on by physical exertion (**جهد جسدي**). They become weak, dizzy, and short of breath, and they also experience heart murmurs and an increased pulse rate. The hemoglobin content of their blood is only about half the normal value of 15 to 16 g/100 mL, because sickled cells are very fragile and rupture easily; this results in anemia ("lack of blood"). An even more serious consequence is that capillaries become blocked by the long, abnormally shaped cells, causing severe pain and interfering with normal organ function -a major factor in the early death of many people with the disease.

Without medical treatment, people with sickle-cell anemia usually die in childhood. Nevertheless, the sickle cell allele is surprisingly common in certain parts of Africa. Investigation into the persistence of an allele that is so obviously deleterious in homozygous individuals led to the finding that in heterozygous individuals, the allele confers a small but significant resistance to lethal forms of malaria. Natural selection has resulted in an allele population that balances the deleterious effects of the homozygous condition against the resistance to malaria afforded by the heterozygous condition.

Complementary Interactions between Proteins and Ligands: The Immune System and Immunoglobulins

Our discussion of oxygen-binding proteins showed how the conformations of these proteins affect and are affected by the binding of small ligands (O_2 or CO) to the heme group. However, most protein-ligand interactions do not involve a prosthetic group. Instead, the binding site for a ligand is more often like the hemoglobin binding site for BPG-a cleft in the protein lined with amino acid residues, arranged to render the binding interaction highly specific. Effective discrimination between ligands is the norm at binding sites, even when the ligands have only minor structural differences.

All vertebrates have an immune system capable of distinguishing molecular "self'" from "nonself'" and then destroying those entities identified as "nonself". In this way, the immune system eliminates viruses, bacteria, and other pathogens and molecules that may pose a threat to the organism. On a physiological level, the response of the immune system to an invader is an intricate and coordinated set of interactions among many classes of proteins, molecules, and cell types. However, at the level of individual proteins, the immune response demonstrates how an acutely sensitive and specific biochemical system is built upon the reversible binding of ligands to proteins.

The Immune Response Features a Specialized Array of Cells and Proteins

Immunity is brought about by a variety of **leukocytes** (white blood cells), including **macrophages** and **lymphocytes**, all developing from undifferentiated stem cells in the bone marrow. Leukocytes can leave the bloodstream and patrol the tissues, each cell producing one or more proteins capable of recognizing and binding to molecules that might signal an infection.

The immune response consists of two complementary systems, the humoral and cellular immune systems. The **humoral immune system** (Latin *humor*, "fluid") is directed at bacterial infections and extracellular viruses (those found in the body fluids), but can also respond to individual proteins introduced into the organism. The **cellular immune system** destroys host cells infected by viruses and also destroys some parasites and foreign tissues.

The proteins at the heart of the humoral immune response are soluble proteins called **antibodies** or **immunoglobulins**, often abbreviated **Ig**. Immunoglobulins bind bacteria, viruses, or large molecules identified as foreign and target them for destruction. Making up 20% of blood protein, the immunoglobulins are produced by **B lymphocytes**, or **B cells**, so named because they complete their development in the bone marrow.

The agents at the heart of the cellular immune response are a class of **T lymphocytes**, or **T cells** (so called because the latter stages of their development occur in the thymus), known as **cytotoxic T cells** (**Tc cells**, also called **killer T cells**). Recognition of infected cells or parasites involves proteins called **T-cell receptors** on the surface of Tc cells. Receptors are proteins, usually found on the outer surface of cells and extending through the plasma membrane; they recognize and bind extracellular ligands, triggering changes inside the cell.

In addition to cytotoxic T cells, there are **helper T cells** (**T_H cells**), whose function it is to produce soluble signaling proteins called cytokines, which include the interleukins. T_H

cells interact with macrophages. Table 5-2 summarizes the functions of the various leukocytes of the immune system.

TABLE 5-2

Some Types of Leukocytes Associated with the Immune System

Cell type	Function
Macrophages	Ingest large particles and cells by phagocytosis
B lymphocytes (B cells)	Produce and secrete antibodies
T lymphocytes (T cells)	
Cytotoxic (killer) T cells (T_C)	Interact with infected host cells through receptors on T-cell surface
Helper T cells (T_H)	Interact with macrophages and secrete cytokines (interleukins) that stimulate T_C, T_H, and B cells to proliferate.

Each recognition protein of the immune system, either an antibody produced by a B cell or a receptor on the surface of a T cell, specifically binds some particular chemical structure, distinguishing it from virtually all others. Humans are capable of producing more than 10^8 different antibodies with distinct binding specificities. This extraordinary diversity makes it likely that any chemical structure on the surface of a virus or invading cell will be recognized and bound by one or more antibodies. Antibody diversity is derived from random reassembly of a set of immunoglobulin gene segments through genetic recombination mechanisms.

Any molecule or pathogen capable of eliciting an immune response is called an **antigen**. An antigen may be a virus, a bacterial cell wall, or an individual protein or other macromolecule. A complex antigen may be bound by a number of different antibodies. An individual antibody or T-cell receptor binds only a particular molecular structure within the antigen, called its **antigenic determinant** or **epitope**.

It would be unproductive for the immune system to respond to small molecules that are common intermediates and products of cellular metabolism. Molecules of M_r < 5,000 are generally not antigenic. However, small molecules can be covalently attached to large proteins in the laboratory, and in this form they may elicit an immune response. These small molecules are called **haptens**.

Self Is Distinguished from Nonself by the Display of Peptides on Cell Surfaces

The immune system must identify and destroy pathogens, but it must also recognize and *not* destroy the normal proteins and cells of the host organism -the "self." Detection of protein antigens in the host is mediated by MHC (major histocompatibility complex) proteins. MHC proteins bind peptide fragments of proteins digested in the cell and present them on the outside surface of the cell. These peptides normally come from the digestion of typical cellular proteins, but during a viral infection viral proteins are also digested and presented on the cell surface by MHC proteins. Peptide fragments from foreign proteins that are displayed by MHC proteins are the antigens the immune system recognizes as nonself. T-cell receptors bind these fragments and launch the subsequent steps of the immune response. There are two classes of MHC proteins (Lehninger, Fig. 5-21), which differ in their distribution among cell types and in the source of digested proteins whose peptides they display.

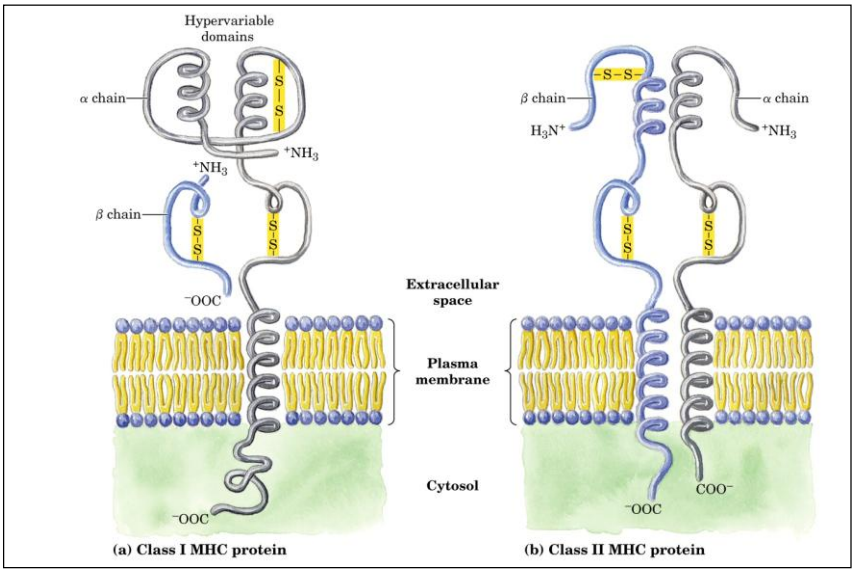


FIGURE 5-21. MHC proteins

Antibodies Have Two Identical Antigen-Binding Sites

Immunoglobulin G (IgG) is the major class of antibody molecule and one of the most abundant proteins in the blood serum. IgG has four polypeptide chains: two large ones, called heavy chains, and two light chains, linked by noncovalent and disulfide bonds into a complex of M_r 150,000. The heavy chains of an IgG molecule interact at one end, then branch to interact separately with the light chains, forming a Y-shaped molecule (Lehninger, Fig. 5-23). At the "hinges" separating the base of an IgG molecule from its branches, the immunoglobulin can be cleaved with proteases. Cleavage with the protease papain liberates the basal fragment, called **Fc** because it usually crystallizes readily, and the two branches, called **Fab,** the antigen-binding fragments. Each branch has a single antigen-binding site.

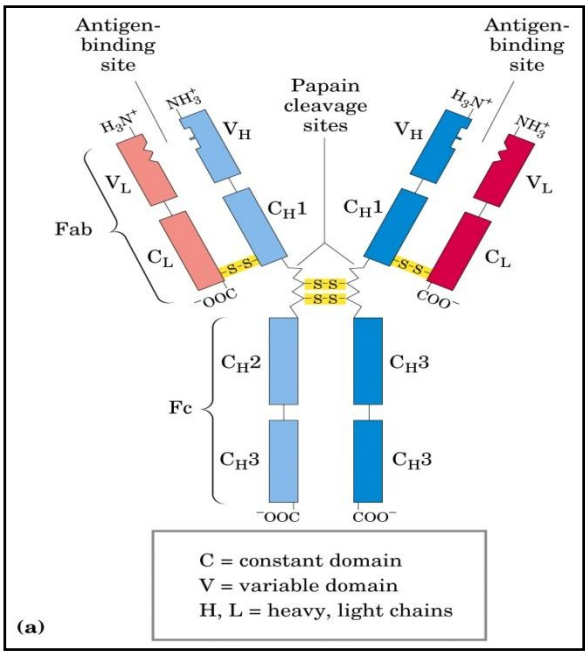


FIGURE 5-23. The structure of immunoglobulin G

Each chain of IgG is made up of identifiable domains; some are constant in sequence and structure from one IgG to the next, others are variable. The heavy and light chains also have one variable domain each, in which most of the variability in amino acid residue sequence is found. The variable domains associate to create the antigen-binding site (Lehninger, Fig. 5-24) of immunoglobulins. Each class has a characteristic type of heavy chain, denoted α, δ, ε, γ, and μ for IgA, IgD, IgE, IgG, and IgM, respectively. Two types of light chain, κ and λ, occur in all classes of immunoglobulins. The overall structures of **IgD** and **IgE** are similar to that of IgG. **IgM** occurs either in a monomeric, membrane bound form or in a secreted form that is a cross-linked pentamer of this basic structure (Lehninger, Fig. 5-25). **IgA,** found principally in secretions such as saliva, tears, and milk, can be a monomer, dimer, or trimer. IgM is the first antibody to be made by B lymphocytes and is the major antibody in the early stages of a primary immune response. Some B cells soon begin to produce IgD (with same antigen-binding site as the IgM produced by the same cell), but the unique function of IgD is less clear.

The IgG described above is the major antibody in secondary immune responses, which are initiated by memory B cells. As part of the organism's ongoing immunity to antigens already encountered and dealt with, IgG is the most abundant immunoglobulin in the blood. When IgG binds to an invading bacterium or virus, it activates certain leukocytes such as macrophages to engulf and destroy the invader, and also activates some other parts of the immune response. Yet another class of receptors on the cell surface of macrophages recognizes and binds the Fc region of IgG. When these Fc receptors bind an antibody-pathogen complex, the macrophage engulfs the complex by phagocytosis (Lehninger, Fig. 5-26).

IgE plays an important role in the allergic response, interacting with basophils (phagocytic leukocytes) in the blood and histamine-secreting cells called mast cells that are widely distributed in tissues. This immunoglobulin binds, through its Fc region, to special Fc receptors on the basophils or mast cells. In this form, IgE serves as a kind of receptor for antigen. If antigen is bound, the cells are induced to secrete histamine and other biologically active amines that cause dilation and increased permeability of blood vessels. These effects on the blood vessels are thought to facilitate the movement of immune system cells and proteins to sites of inflammation. They also produce the symptoms normally associated with allergies. Pollen or other allergens are recognized as foreign, triggering an immune response normally reserved for pathogens.

Antibodies Bind Tightly and Specifically to Antigen

The binding specificity of an antibody is determined by the amino acid residues in the variable domains of its heavy and light chains. Many residues in these domains are variable, but not equally so. Some, particularly those lining the antigen-binding site, are hypervariable -especially likely to differ. Specificity is conferred by chemical complementarity between the antigen and its specific binding site, in terms of shape and the location of charged, nonpolar, and hydrogen-bonding groups. For example, a binding site with a negatively charged group may bind an antigen with a positive charge in the complementary position. In many instances, complementarity is achieved interactively as the structures of antigen and binding site are influenced by each other during the approach of the ligand. Conformational changes in the antibody and/or the antigen then occur that allow the complementary groups to interact fully.

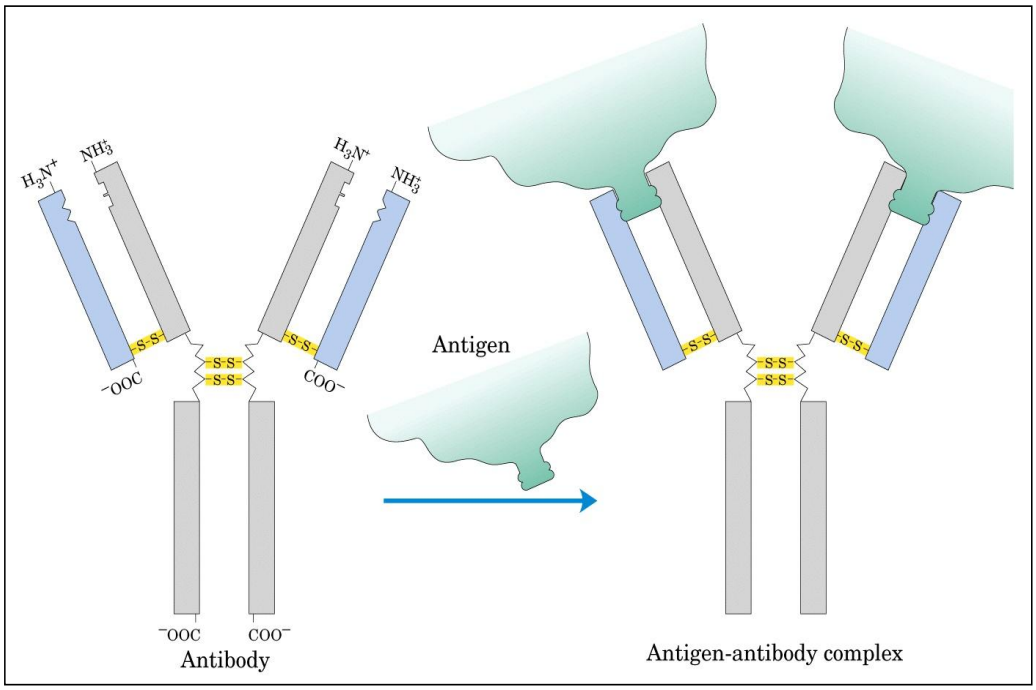


FIGURE 5-24. Binding of IgG to antigen

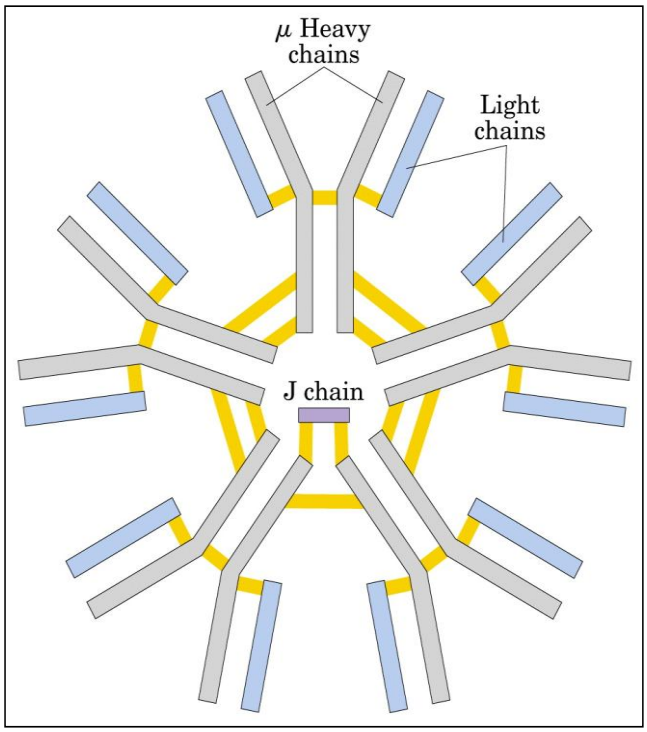


FIGURE 5-25. IgM pentamer of immunoglobulin units

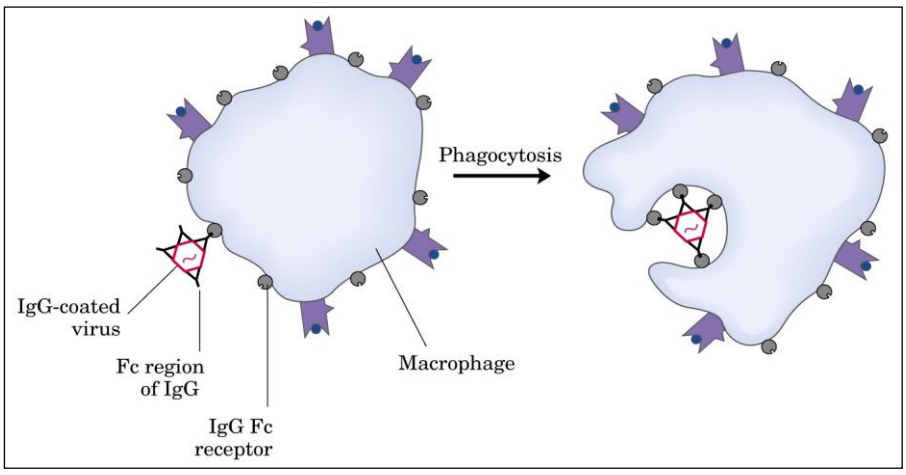


FIGURE 5-26. Phagocytosis of an antibody-bound virus by a macrophage

A typical antibody-antigen interaction is quite strong, characterized by K_d values as low as 10^{-10} M (recall that a lower K_d corresponds to a stronger binding interaction). The K_d reflects the energy derived from the various ionic, hydrogen-bonding, hydrophobic, and van der Waals interactions that stabilize the binding.

The Antibody-Antigen Interaction is the Basis for a Variety of Important Analytical Procedures

The extraordinary binding affinity and specificity of antibodies make them valuable analytical reagents. Two types of antibody preparations are in use: polyclonal and monoclonal. **Polyclonal antibodies** are those produced by many different B lymphocytes responding to one antigen, such as a protein injected into an animal. Cells in the population of B lymphocytes produce antibodies that bind specific, different epitopes within the antigen. Thus, polyclonal preparations contain a mixture of antibodies that recognize different parts of the protein. **Monoclonal antibodies,** in contrast, are synthesized by a population of identical B cells (a **clone)** grown in cell culture. These antibodies are homogeneous, all recognizing the same epitope.

The specificity of antibodies has practical uses. A selected antibody can be covalently attached to a resin and used in a chromatography column of the type shown in Figure 3-18c (Lehninger). When a mixture of proteins is added to the column, the antibody specifically binds its target protein and retains it on the column while other proteins are washed through. The target protein can then be eluted from the resin by a salt solution or some other agent. This is a powerful tool for protein purification.

In another versatile analytical technique, an antibody is attached to a radioactive label or some other reagent that makes it easy to detect. When the antibody binds the target protein, the label reveals the presence of the protein in a solution or its location in a gel or even a living cell. An **ELISA** (enzyme-linked inununosorbent assay) allows for rapid screening and quantification of the presence of an antigen in a sample (Lehninger, Fig. 5-28b). Proteins in a sample are adsorbed to an inert surface, usually a 96 well polystyrene plate. The surface is washed with a solution of an inexpensive nonspecific protein (often casein from nonfat dry milk powder)

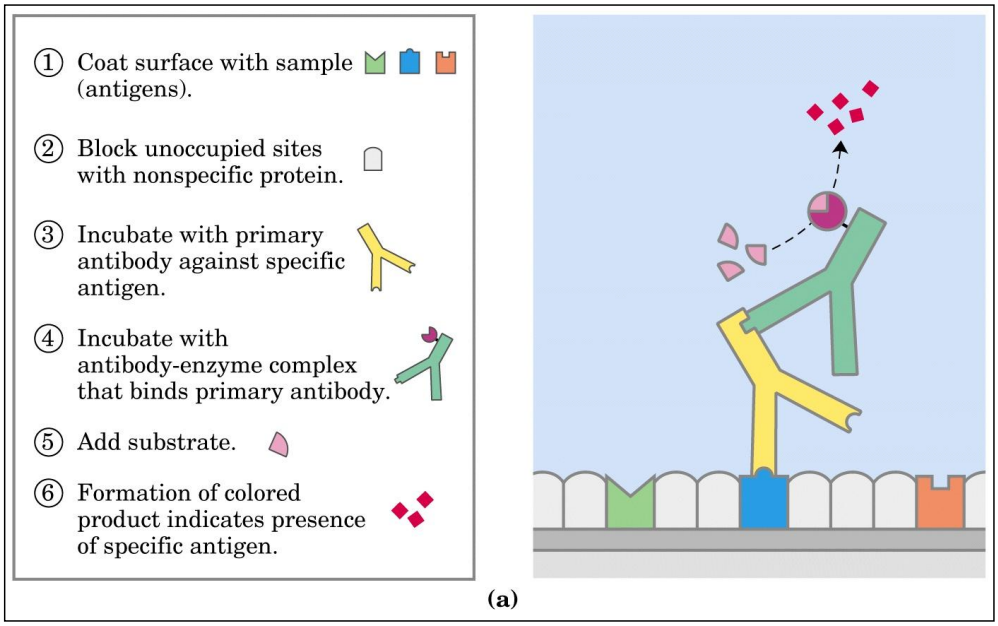


FIGURE 5-28. Antibody techniques: (a) schematic representation of the general method

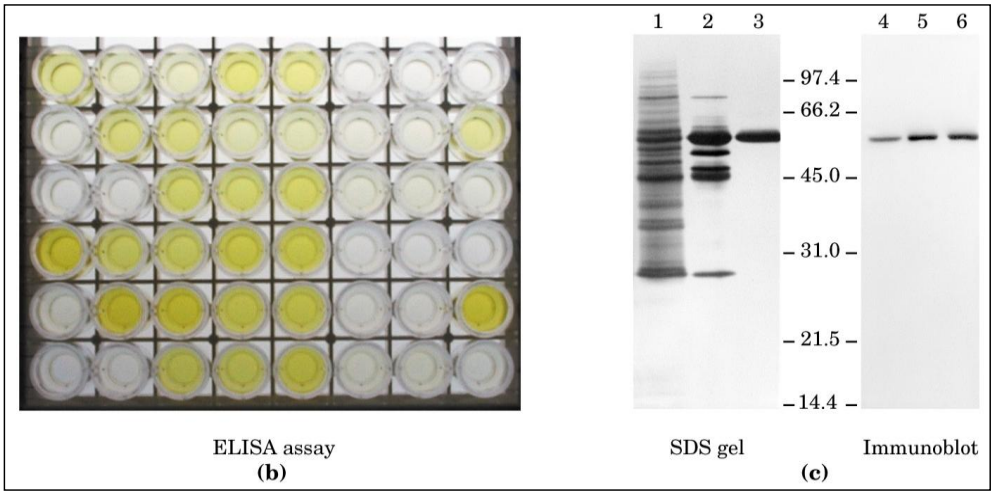


FIGURE 5-28. Antibody techniques: (b) an ELISA; (c) an immunoblot

to block proteins introduced in subsequent steps from also adsorbing to these surfaces. The surface is then treated with a solution containing the primary antibody -an antibody against the protein of interest. Unbound antibody is washed away and the surface is treated with a solution containing antibodies against the primary antibody. These secondary antibodies have been linked to an enzyme that catalyzes a reaction that forms a colored product. After unbound secondary antibody is washed away, the substrate of the antibody-linked enzyme is added. Product formation (monitored as color intensity) is proportional to the concentration of the protein of interest in the sample.

In an **immunoblot** assay (Lehninger, Fig. 5-28c), proteins that have been separated by gel electrophoresis are transferred electrophoretically to a nitrocellulose membrane. The membrane is blocked (as described above for ELISA), then treated successively with primary antibody, secondary antibody linked to enzyme, and substrate. A colored precipitate forms only along the band containing the protein of interest.

Protein Interactions Modulated by Chemical Energy: Actin, Myosin, and Molecular Motors

Organisms move, cells move, organelles and macromolecules within cells move. Most of these movements arise from the activity of a fascinating class of protein-based molecular motors. Fueled by chemical energy, usually derived from ATP, large aggregates of motor proteins undergo cyclic conformational changes that accumulate into a unified, directional force -the tiny force that pulls apart chromosomes in a dividing cell, and the immense force that levers a pouncing, quarter-ton jungle cat into the air.

Motor proteins underlie the contraction of muscles, the migration of organelles along microtubules, the rotation of bacterial flagella, and the movement of some proteins along DNA.

The Major Proteins of Muscle are Myosin and Actin

The contractile force of muscle is generated by the interaction of two proteins, myosin and actin. These proteins are arranged in filaments that undergo transient interactions and slide past each other to bring about contraction. Together, actin and myosin make up more than 80% of the protein mass of muscle.

Myosin *(M_r* 540,000) has six subunits: two heavy chains (each of M_r 220,000) and four light chains (each of M_r 20,000). The heavy chains account for much of the overall structure. At their carboxyl termini, they are arranged as extended α helices, wrapped around each other in a fibrous, left-handed coiled coil similar to that of α-keratin (Lehninger, Fig. 5-29a). At its amino terminus, each heavy chain has a large globular domain containing a site where ATP is hydrolyzed. The light chains are associated with the globular domains.

In muscle cells, molecules of myosin aggregate to form structures called **thick filaments** (Lehninger, Fig. 5-30a). These rodlike structures serve as the core of the contractile unit. Within a thick filament, several hundred myosin molecules are arranged with their fibrous "tails" associated to form a long bipolar structure. The globular domains project from either end of this structure, in regular stacked arrays. The second major muscle protein, **actin,** is abundant in almost all eukaryotic cells. In muscle, molecules of monomeric actin, called G-actin (globular actin; M_r 42,000), associate to form a long polymer called F-actin (filamentous actin). The **thin filament** (Lehninger, Fig. 5-30b) consists of F-actin, along with the proteins troponin and tropomyosin. The filamentous parts of thin filaments assemble

as successive monomeric actin molecules add to one end. On addition, each monomer binds ATP, then hydrolyzes it to ADP, so every actin molecule in the filament is complexed to ADP. This ATP hydrolysis by actin functions only in the assembly of the filaments; it does not contribute directly to the energy expended in muscle contraction. Each actin monomer in the thin filament can bind tightly and specifically to one myosin head group (Lehninger, Fig. 5-30c).

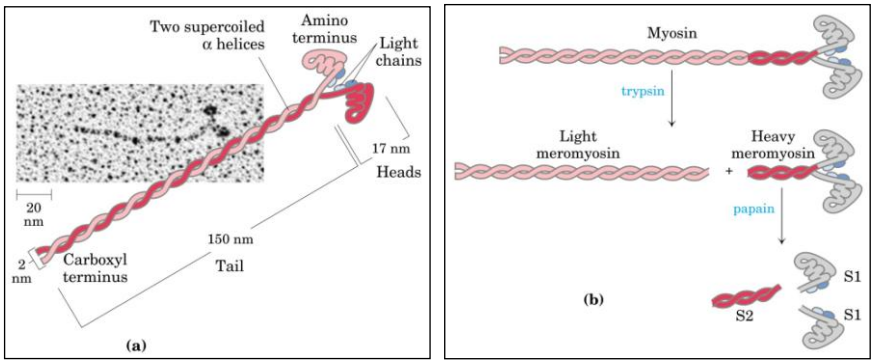
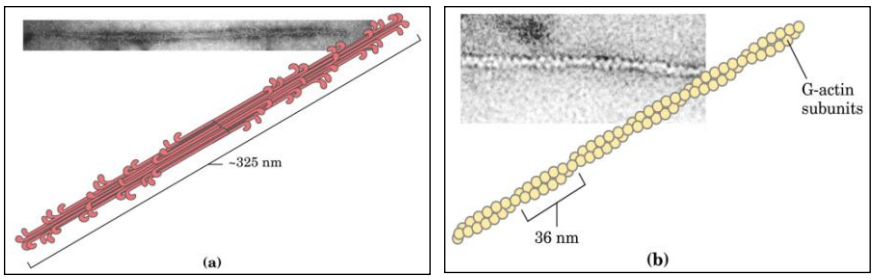


FIGURE 5-29. Myosin



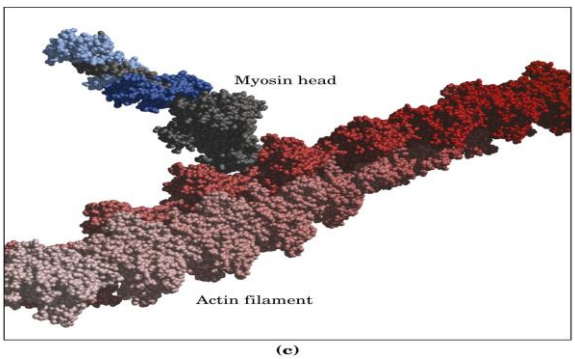


FIGURE 5-30. The major component of muscle: (a)myosin aggregates; (b) F-actin; (c) association of actin filament with myosin head

Additional Proteins Organize the Thin and Thick Filaments into Ordered Structures

Skeletal muscle consists of parallel bundles of **muscle fibers,** each fiber a single, very large, multinucleated cell, 20 to 100 μm in diameter, formed from many cells fused together and often spanning the length of the muscle. Each fiber, in turn, contains about 1,000 **myofibrils,** 2 μm in diameter, each consisting of a vast number of regularly arrayed thick and thin filaments complexed to other proteins (Lehninger, Fig. 5-31). A system of flat membranous vesicles called the **sarcoplasmic reticulum** surrounds each myofibril. Examined under the electron microscope, muscle fibers reveal alternating regions of high and low electron density, called the **A bands** and **I bands** (Lehninger, Fig. 5-31b, c). The A and I bands arise from the arrangement of thick and thin filaments, which are aligned and partially overlapping. The I band is the region of the bundle that in cross section would contain only thin filaments. The darker A band stretches the length of the thick filament and includes the region where parallel thick and thin filaments overlap. Bisecting the I band is a thin structure called the **Z disk,** perpendicular to the thin filaments and serving as an anchor to which the thin filaments are attached. The A band too is bisected by a thin line, the **M line** or M disk, a region of high electron density in the middle of the thick filaments. The entire contractile unit, consisting of bundles of thick filaments interleaved at either end with bundles of thin filaments, is called the **sarcomere.** The arrangement of interleaved bundles allows the thick and thin filaments to slide past each other causing a progressive shortening of each sarcomere (Lehninger, Fig. 5-32).

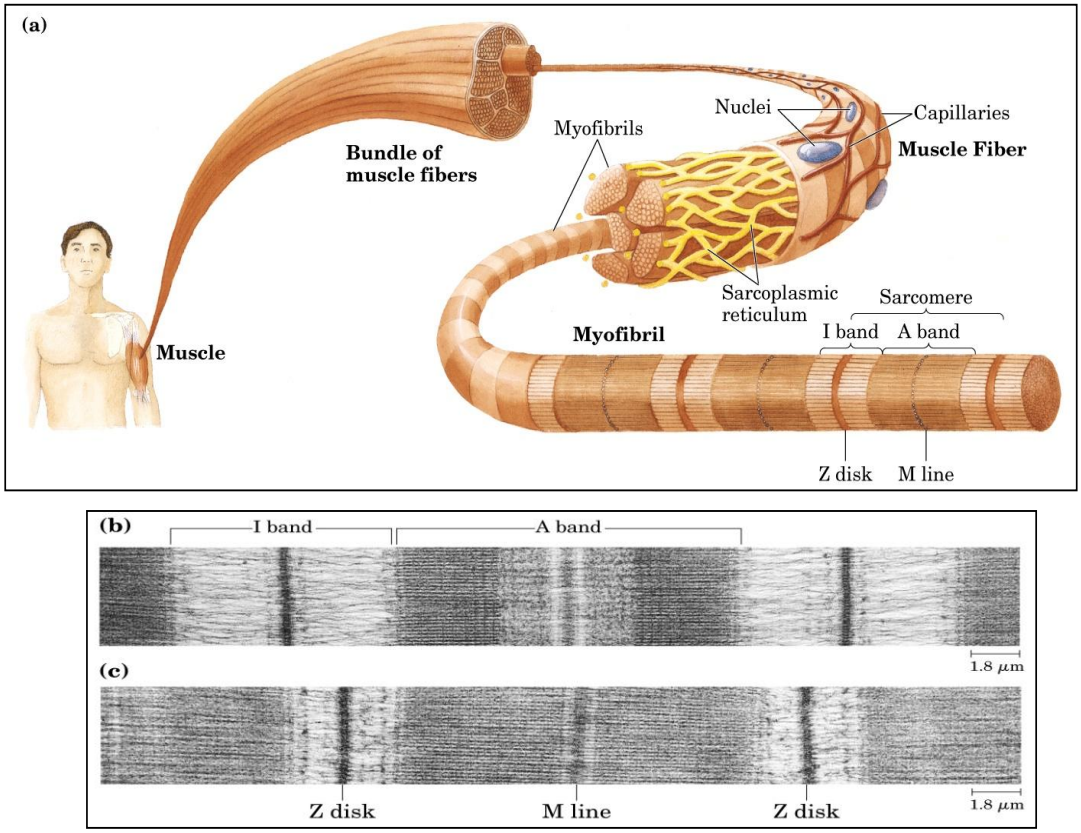


FIGURE 5-31. Structure of skeletal muscle: (a) muscle fibers; (b) electron micrographs of (b) relaxed and (c) contracted muscle

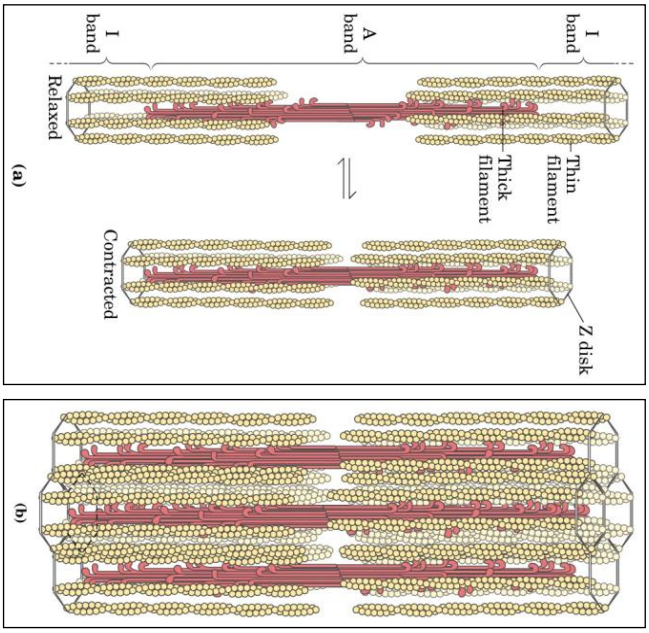


FIGURE 5-32. Muscle contraction: (a) Muscle contraction occurs by the sliding of the thick and thin filaments; (b) Thick filaments surrounded by six thin filaments

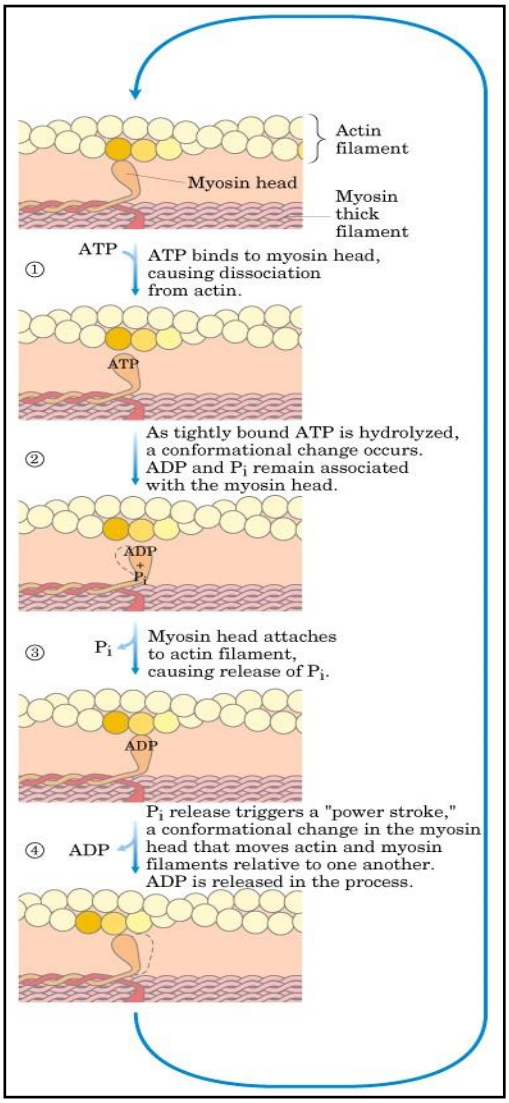


FIGURE 5-33. Molecular mechanism of muscle contraction

Myosin Thick Filaments Slide along Actin Thin Filaments

The interaction between actin and myosin, like that between all proteins and ligands, involves weak bonds. When ATP is not bound to myosin, a face on the myosin head group binds tightly to actin (Lehninger, Fig. 5-33). When ATP binds to myosin and is hydrolyzed to ADP and phosphate, a coordinated and cyclic series of conformational changes occurs in which myosin releases the F-actin subunit and binds another subunit farther along the thin filament.

The cycle has four major steps (Lehninger, Fig. 5-33). In step 1 ATP binds to myosin and a cleft in the myosin molecule opens, disrupting the actin-myosin interaction so that the bound actin is released. ATP is then hydrolyzed in step 2, causing a conformational change in the protein to a "high-energy" state that moves the myosin head and changes its orientation in relation to the actin thin filament. Myosin then binds weakly to an F-actin subunit closer to the Z disk than the one just released. As the phosphate product of ATP hydrolysis is released from myosin in step 3, another conformational change occurs in which the myosin cleft closes, strengthening the myosin-actin binding. This is followed quickly by step 4, a "power stroke" during which the conformation of the myosin head returns to the original resting state, its orientation relative to the bound actin changing so as to pull the tail of the myosin toward the Z disk. ADP is then released to complete the cycle. The thick filament thus actively slides forward past the adjacent thin filaments. This process, coordinated among the many sarcomeres in a muscle fiber, brings about muscle contraction.

The interaction between actin and myosin must be regulated so that contraction occurs only in response to appropriate signals from the nervous system. The regulation is mediated by a complex of two proteins, **tropomyosin** and **troponin**. Tropomyosin binds to the thin filament, blocking the attachment sites for the myosin head groups. Troponin is a Ca^{2+} -binding protein. A nerve impulse causes release of Ca^{2+} from the sarcoplasmic reticulum. The released Ca^{2+} binds to troponin (another protein-ligand interaction) and causes a conformational change in the tropomyosin-troponin complexes, exposing the myosin-binding sites on the thin filaments. Contraction follows.

Working skeletal muscle requires two types of molecular functions that are common in proteins-binding and catalysis. The actin-myosin interaction, a protein ligand interaction like that of immunoglobulins with antigens, is reversible. When ATP binds myosin, however, it is hydrolyzed to ADP and Pi. Myosin is not only an actin binding protein, it is also an ATPase -an enzyme.