

## INTRODUCTION

Biochemistry is the science concerned with the various molecules that occur in living cells and organisms and with their chemical reactions. Anything more than an extremely superficial comprehension of life—in all its diverse manifestations—demands a knowledge of biochemistry. In addition, medical students who acquire a sound knowledge of biochemistry will be in a strong position to deal with two central concerns of the health sciences:

- (1) the understanding and maintenance of health
- (2) the understanding and effective treatment of disease

### **Biochemistry Is the Chemistry of Life**

Biochemistry can be defined more formally as *the science concerned with the chemical basis of life* (Gk *bios* "life").

The **cell** is the structural unit of living systems. Consideration of this concept leads to a functional definition of biochemistry as *the science concerned with the chemical constituents of living cells and with the reactions and processes that they undergo*. By this definition, biochemistry encompasses large areas of **cell biology**, of **molecular biology**, and of **molecular genetics**.

**The Aim of Biochemistry Is to Describe and Explain, in Molecular Terms, All Chemical Processes of Living Cells**

## **A Knowledge of Biochemistry Is Essential to All Life Sciences**

The biochemistry of the nucleic acids lies at the heart of genetics; in turn, the use of genetic approaches has been critical for elucidating many areas of biochemistry.

Physiology, the study of body function, overlaps with biochemistry almost completely.

Immunology employs numerous biochemical techniques, and many immunologic approaches have found wide use by biochemists.

Pharmacology and pharmacy rest on a sound knowledge of biochemistry and physiology; in particular, most drugs are metabolized by enzyme-catalyzed reactions, and the complex interactions among drugs are best understood biochemically.

Poisons act on biochemical reactions or processes; this is the subject matter of toxicology.

Biochemical approaches are being used increasingly to study basic aspects of pathology (the study of disease), such as inflammation, cell injury, and cancer.

Many workers in microbiology, zoology, and botany employ biochemical approaches almost exclusively.

These relationships are not surprising, because life as we know it depends on biochemical reactions and processes. In fact, the old barriers among the life sciences are breaking down, and biochemistry is increasingly becoming their common language.

**Some uses of biochemical investigations and laboratory tests in relation to diseases.**

Use	Example
1. To reveal the fundamental causes and mechanisms of diseases	Demonstration of the nature of the genetic defects in cystic fibrosis.
2. To suggest rational treatments of diseases based on (1) above	Use of a diet low in phenylalanine for the treatment of phenylketonuria.
3. To assist in the diagnosis of specific diseases	Use of the plasma enzyme creatine kinase MB (CK-MB) in the diagnosis of myocardial infarction.
4. To act as screening tests for the early diagnosis of certain diseases	Use of measurement of blood thyroxine or thyroid-stimulating hormone (TSH) in the neonatal diagnosis of congenital hypothyroidism.
5. To assist in monitoring the progress (eg, recovery, worsening, remission, or relapse) of certain diseases	Use of the plasma enzyme alanine aminotransferase (ALT) in monitoring the progress of infectious hepatitis.
6. To assist in assessing the response of diseases to therapy	Use of measurement of blood carcinoembryonic antigen (CEA) in certain patients who have been treated for cancer of the colon.

# "NATURE" of PROTEINS

## **A. Structure**

1. Proteins are linear, unbranched polymers constructed from 20 different  $\alpha$ -amino acids that are encoded in the DNA of the genome.
2. All living organisms use the same 20 amino acids and, with few exceptions, the same genetic code (see Chapter 10 I B).

**B. Size.** Proteins are diverse in size. The mass of single-chain proteins is typically 10-50 kilodaltons (kdal), although proteins as small as 350 dal and greater than 1000 kdal are known to exist. Multichain protein complexes of greater than 200 kdal are frequently encountered.

**Function.** Proteins serve a wide range of functions in living organisms. A few of their functions include:

1. **Enzymatic catalysis**—Most enzymes are proteins.
2. **Transport and storage** of small molecules and ions
3. **Structural elements of the cytoskeleton.** Proteins make up the cytoskeleton, which:
  - a. Provides **strength and structure** to cells
  - b. Forms the **fundamental mechanistic components for intracellular and extracellular movement**
4. **Structure of skin and bone.** Proteins such as **collagen**, the most abundant protein in the body, give these structures high tensile strength.
5. **Immunity.** The **immune defense system** is composed of proteins such as **antibodies**, which mediate a protective response to pathogens.
6. **Hormonal regulation.** Hormones coordinate the metabolic actions within the body
  - a. Some **hormones** are proteins [e.g., somatotropin (pituitary growth hormone) and insulin].
  - b. **The cellular receptors** that recognize hormones and neurotransmitters are proteins.
7. **Control of genetic expression.** Activators, repressors, and many other regulators of gene expression in prokaryotes and eukaryotes are proteins.

## **Unique conformation**

1. **Specificity.** Proteins show an exquisite specificity of biologic function—a consequence of the uniqueness of the three-dimensional structural shape, or **conformation**, of each protein.
2. In humans, **disease states** are often related to the altered function of a protein, which is often attributed to an anomaly in the protein's structure. Examples of diseases caused by abnormal protein structure and function include:
  - a. **Hemoglobinopathies**, in particular **sickle cell anemia**
  - b. **Marfan syndrome**, which appears to be caused by single amino acid changes in an elastic connective tissue protein called **fibrillin**
  - c. **Cystic fibrosis**, the major form of which arises because of a single amino acid deletion in the adenosine triphosphate (ATP)-binding domain of a transmembrane conductance regulatory protein

**AMINO ACIDS** are the fundamental units of proteins.

## **Composition**

1. Amino acids are composed of an **amino group** ( $\text{—NH}_2$ ), a **carboxyl group** ( $\text{—COOH}$ ), a hydrogen atom, and a distinctive **side chain**, all bonded to a carbon atom (the  **$\alpha$ -carbon**). Table lists the 20 fundamental amino acids according to their side chains.
2. One of the 20 amino acids, **proline**, is an **imino acid** ( $\text{—NH—}$ ), **not** an  **$\alpha$ -amino acid** as are the other 19.

**Post-translational modification.** Other amino acids are found in a number of proteins but are not coded for in DNA; they are derived from some of the 20 fundamental amino acids after these have been incorporated into the protein chain (i.e., post-translational modification). More than 100 different kinds of amino acids that arise from post-translational modifications have been identified. Examples of a few of the major post-translational modifications of amino acids are:

- a. Addition of **hydroxyl** ( $\text{—OH}$ ) groups to some prolines and lysines in collagen and gelatin
- b. Addition of **methyl** ( $\text{—CH}_3$ ) groups to some lysines and histidines in muscle myosin
- c. Addition of **carboxyl** ( $\text{—COOH}$ ) groups to glutamates in blood clotting and bone proteins
- d. Addition of **phosphate** ( $\text{—PO}_3$ ) groups to some serine, threonine, and tyrosine molecules. **Reversible phosphorylation** is a common method of regulating the activity of many enzymes, cell-surface receptors, and other regulatory molecules.

There are many **nonprotein amino acids** found throughout nature. In some cases, these amino acids serve as antibiotics or toxins.



## Optical activity

1. With the exception of glycine, all amino acids contain at least one **asymmetric carbon** atom and are, therefore, optically active.
2. **Enantiomers.** Amino acids exist as stereoisomeric pairs called enantiomers. These amino acid isomers are typically called **L** (levorotatory) or **D** (dextrorotatory) depending on the direction they rotate plane-polarized light.
  - a. **L-Amino acids** are the only optically active amino acids that are incorporated into proteins.
  - b. **D-Amino acids** are found in bacterial products (e.g., in cell walls) and in many peptide antibiotics, but they are not incorporated into proteins via the ribosomal protein synthesizing system.

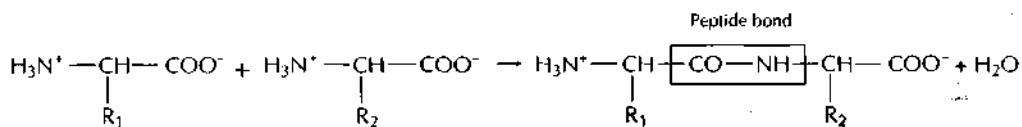
### **Amphoteric properties of amino acids**

1. Amino acids are **amphoteric molecules**; that is, they have both basic and acidic groups.
2. Monoamino-monocarboxylic acids exist in aqueous solution as **dipolar molecules (zwitterions)**, which means that they have both positive and negative charges.
  - a. The  **$\alpha$ -carboxyl group** is dissociated and negatively charged.
  - b. The  **$\alpha$ -amino group** is protonated and positively charged.
  - c. Thus, the overall molecule is **electrically neutral**.
3. At **low pH** (i.e., high concentrations of hydrogen ion), the carboxyl group accepts a proton and becomes uncharged, so that the overall charge on the molecule is positive.
4. At **high pH** (i.e., low concentrations of hydrogen ion), the amino group loses its proton and becomes uncharged; thus, the overall charge on the molecule is negative.
5. Some amino acids have **side chains** that contain **dissociating groups**,
  - a. **Side chains**
    - (1) Those of aspartate and glutamate are acidic; those of histidine, lysine, and arginine are basic.
    - (2) Two others, cysteine and tyrosine, have a negative charge on the side chain when dissociated.
  - b. **Dissociating groups**
    - (1) Whether these groups are dissociated depends on the prevailing pH and the apparent dissociation constant ( $pK_i$ ) of the dissociating groups.
    - (2) These dissociating amino acids also exist in solution as zwitterions. For example, glutamate has three dissociable protons with  $pK_i$  values of 2.19, 4.25, and 9.67. As the pH increases above each of these  $pK_i$  values, protons dissociate and the charge changes

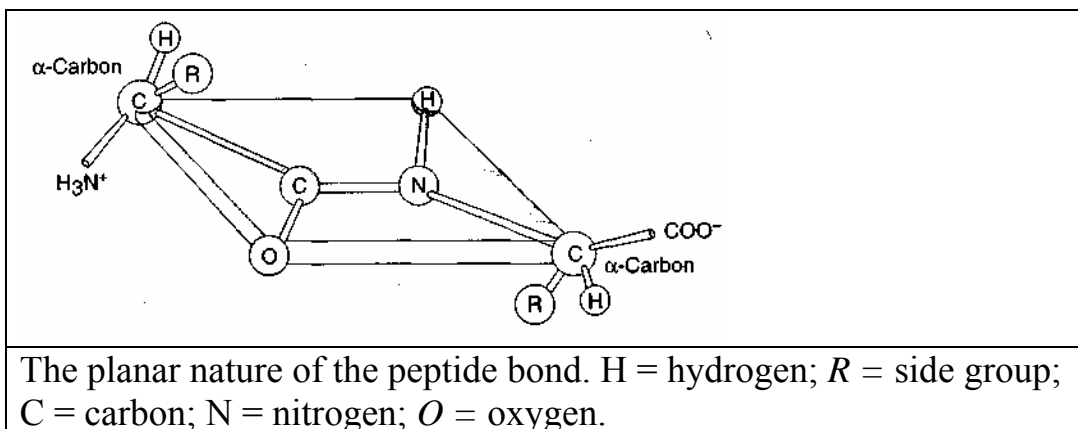
### III PEPTIDES AND POLYPEPTIDES

**Formation.** The linking together of amino acids produces **peptide** chains, also called **polypeptides** if many amino acids are linked.

1. The **peptide bond** is the bond formed between the  $\alpha$ -carboxyl group of one amino acid and the  $\alpha$ -amino group of another. In the process, water is removed.



2. Peptide bond formation is highly **endergonic** (i.e., energy-requiring) and requires the concomitant hydrolysis of high-energy phosphate bonds.
3. The peptide bond is a **planar** structure with the two adjacent  $\alpha$ -carbons, a carbonyl oxygen, an  $\alpha$ -amino nitrogen and its associated hydrogen atom, and the carbonyl carbon all lying in the same plane (Figure 2-1). The —CM— bond has a partial double-bond character that prevents rotation around the bond axis.
4. Amino acids, when in polypeptide chains, are customarily referred to as **residues**.



## **Amphoteric properties of polypeptide**

- 1. The formation of the peptide bond** removes two dissociating groups, one from the  $\alpha$ -amino and one from the  $\alpha$ -carboxyl, per residue.
- Although the N-terminal and C-terminal  $\alpha$ -amino and  $\alpha$ -carboxyl groups can play important roles in the formation of protein structures, and thus in protein function, the amphoteric properties of a polypeptide are mainly governed by the **dissociable groups** on the amino acid **side chains**.
- 3. Laboratory use.** These properties of proteins are not only important in terms of protein structure and function but are also useful in a number of analytic procedures, such as ion exchange or high-performance liquid chromatography, for the purification and identification of proteins.

## CONFORMATION OF PROTEINS.

Every protein in its **native state** has a unique **three-dimensional structure**, which is referred to as its **conformation**. The function of a protein arises from its conformation. Protein structures can be classified into four levels of organization: primary, secondary, tertiary, and quaternary.

The **primary structure** is the covalent "backbone" of the polypeptide formed by the **specific amino acid sequence**.

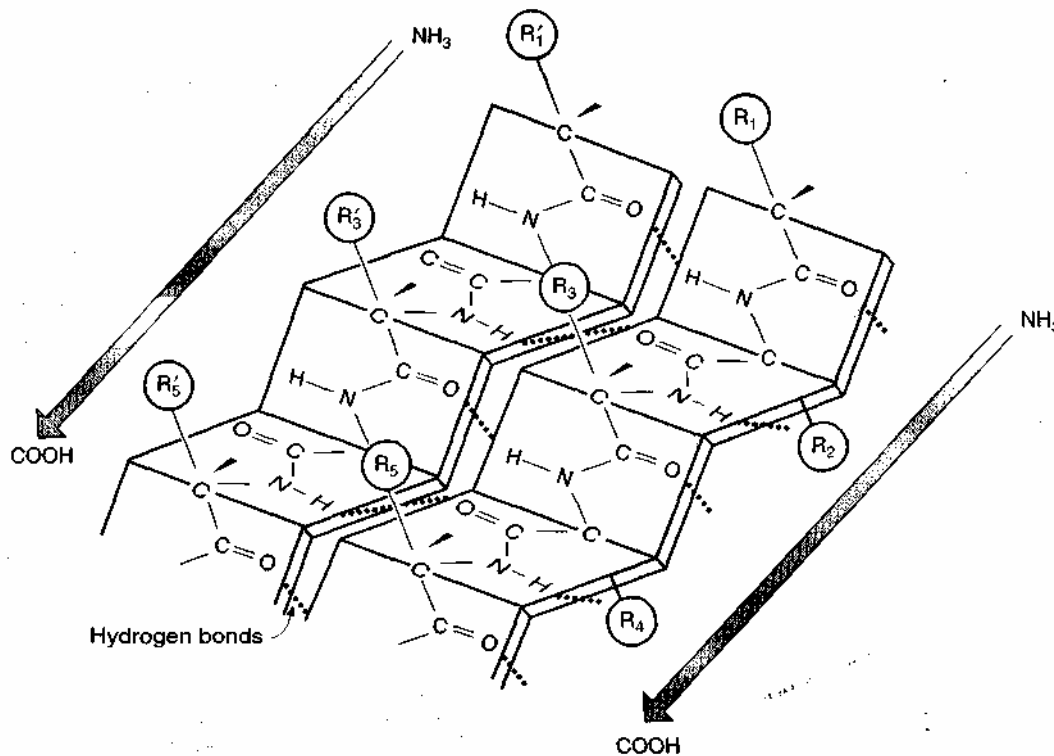
1. This sequence is coded for by DNA and determines the final three-dimensional form adopted by the protein in its native state.
2. By convention, peptide sequences are written from left to right, starting with the amino acid residue that has a free  $\alpha$ -amino group (the so-called **N-terminal amino acid**) and ending with the residue that has a free  $\alpha$ -carboxyl group (the **C-terminal amino acid**). Either the three-letter abbreviations (e.g., Ala-Glu-Lys) or, for long peptides, the single-letter abbreviations are used.

**B. The secondary structure** is the spatial relation of neighboring amino acid residues.

- 1. Secondary structure is dictated by the primary structure.** The secondary structure arises from interactions of neighboring amino acids. Because the DNA-coded primary sequence dictates which amino acids are near each other, secondary structure often forms as the peptide chain comes off the ribosome (see Chapter 10 V B; VI).
- 2. Hydrogen bonds.** An important characteristic of secondary structure is the formation of hydrogen bonds (H bonds) between the —CO group of one peptide bond and the —NH group of another nearby peptide bond.
  - a. If the H bonds form between peptide bonds in the same chain, either **helical** structures, such as the  **$\alpha$ -helix**, develop or turns, such as  **$\beta$ -turns**, are formed.
  - b. If H bonds form between peptide bonds in different chains, **extended** structures form, such as the  **$\beta$ -pleated sheet**.
- 3. The  $\alpha$ -helix** is a rod-like structure with the peptide bonds coiled tightly inside and the side chains of the residues protruding outward (Figure 2-2).
  - a. Each —CO is hydrogen bonded to the —NH of a peptide bond that is **four residues away** from it along the **same chain**.
  - b. There are 3.6 amino acid residues per turn of the helix, and the helix is **right-handed** (i.e., the coils turn in a clockwise fashion around the axis).
- 4.  $\beta$ -Pleated sheet structures** are found in many proteins, including some globular, soluble proteins, as well as some fibrous proteins (e.g., silk fibroin).
  - a. They are more extended structures than the  $\alpha$ -helix and are "pleated" because the C—C bonds are tetrahedral and cannot exist in straight lines (Figure).
  - b. The chains lie side by side, with the **hydrogen bonds** forming between the —CO group of one peptide bond and the —NH group of another peptide bond in the **neighboring chain**.
  - c. The chains may run in the same direction, forming a **parallel  $\beta$ -sheet**, or they may run in opposite directions, as they do in a globular protein in which an extended chain is folded back on itself, forming an **antiparallel  $\beta$ -structure**.
- 5. A  $\beta$ -turn** is the tightest turn a polypeptide chain can make, although there are many ways a polypeptide chain can turn.  $\beta$ -Turns result in a complete reversal in the direction of a polypeptide chain in just four amino acid residues.

**Tertiary structure** refers to the spatial relations of more distant residues.

**Folding.** The secondarily ordered polypeptide chains of soluble proteins tend to fold into globular structures with the hydrophobic side chains<sup>1</sup> in the interior of the structure away from the water and the hydrophilic side chains on the outside in contact with water. This folding is due to associations between segments of  $\alpha$ -helix, extended  $\beta$ -chains, or other secondary structures and represents a state of lowest energy (i.e., of **greatest stability**) for the protein in question.



Parallel  $\beta$ -pleated sheet. Planar peptide bonds are shown with hydrogen bonds between parallel adjacent peptide chains. The numbered circles represent the varied side chains (R and R') of the different amino acids on each peptide chain. The adjacent arrows indicate the directionality of the peptide chains from their amino to carboxyl ends.

The conformation results from:

- a. **Hydrogen-bonding** within a chain or between chains
- b. **The flexibility** of the chain at **points of instability**, allowing water to obtain maximum entropy and thus govern the structure to some extent
- c. The formation of **other noncovalent bonds** between side-chain groups, such as salt linkages, or  $\pi$ -electron interactions of aromatic rings
- d. The sites and numbers of **disulfide bridges** between Cys residues within the chain (Cys residues linked by disulfide bonds are termed **cystine** residues)

---

<sup>1</sup> Amino acids with long, nonpolar side chains (e.g., valine, leucine, isoleucine, phenylalanine, trypto-phan, and methionine) are hydrophobic in nature.

A peptide chain free in solution will not achieve its biologically active tertiary structure as rapidly or properly as within the cell. Within the cell, some of the proteins that facilitate proper folding are:

- **Protein disulfide isomerase.** This protein catalyzes the formation of proper disulfide bond formation between cysteine residues.
- **Chaperones.** This family of proteins catalyzes the proper folding of proteins in part by inhibiting improper folding and interactions with other peptides.



**Quarternary structure** refers to the spatial relations between individual polypeptide chains in a multichain protein; that is, the characteristic noncovalent interactions between the chains that form the native conformation of the protein as well as occasional disulfide bonds between the chains.

1. Many proteins larger than 50 kdal have more than one chain and are said to contain **multiple subunits**, with individual chains known as **protomers**.
2. Many multisubunit proteins are composed of different kinds of **functional subunits** [e.g., the regulatory and catalytic subunits of regulatory proteins].