**COURSE BOOK OF**

**CHEMISTRY 2**

**(BIOCHEMISTRY)**

**Desert land and reclamation program**

**Level 1**

**Department of Biochemistry**

**Benha University, Agriculture College**

**Theoretical**

**Theoretical of Chemistry 2**

**BIOCHEMISTRY**

**AC 0102**

The word “**protein**” is derived from the Greek word ***“proteios”,*** which means “of primary importance”. In fact, proteins plays an important role in all biochemical and physiological body processes; they act as enzymes, hormones, receptors, antibodies and are required for the structural integrity of cells.

**AMINO ACIDS, PEPTIDES AND PROTEINS**

Proteins are the most abundant biological macromolecules, occurring in all cells and all parts of cells. Proteins also occur in great variety; thousands of different kinds, ranging in size from relatively small peptides to huge polymers with molecular weights in the millions, may be found in a single cell. Moreover, proteins exhibit enormous diversity of biological function and are the most important final products of the information pathways. Proteins are the molecular instruments through which genetic information is expressed.

Relatively simple monomeric subunits provide the key to the structure of the thousands of different proteins. All proteins, whether from the most ancient lines of bacteria or from the most complex forms of life, are constructed from the same ubiquitous set of 20 amino acids, covalently linked in characteristic linear sequences. Because each of these amino acids has a side chain with distinctive chemical properties, this group of 20 precursor molecules may be regarded as the alphabet in which the language of protein structure is written. What is most remarkable is that cells can produce proteins with strikingly different properties and activities by joining the same 20 amino acids in many different combinations and sequences. From these building blocks different organisms can make such widely diverse products as enzymes, hormones, antibodies, transporters, muscle fibers, the lens protein of the eye, feathers, spider webs, rhinoceros horn, milk proteins, antibiotics,

mushroom poisons, and myriad other substances having distinct biological activities. Among these protein products, the enzymes are the most varied and specialized. Virtually all cellular reactions are catalyzed by enzymes. We begin with a description of the fundamental chemical properties of amino acids, peptides, and proteins.

**Amino Acids**

Proteins are polymers of amino acids, with each **amino acid residue** joined to its neighbor by a specific typeof covalent bond. (The term “residue” reflects the lossof the elements of water when one amino acid is joinedto another.) Proteins can be broken down (hydrolyzed)to their constituent amino acids by a variety of methods,and the earliest studies of proteins naturally focused on the free amino acids derived from them. Twenty different amino acids are commonly found in proteins. The first to be discovered was asparagine, in 1806. The last of the 20 to be found, threonine, was not identified until 1938. All the amino acids have trivial or common names, in some cases derived from the source from which they were first isolated. Asparagine was first found in asparagus and glutamate in wheat gluten; tyrosine was first isolated from cheese (its name is derived from the Greek *tyros,* “cheese”); and glycine (Greek *glykos,*“sweet”) was so named because of its sweet taste

**Amino Acids Share Common Structural Features**

All 20 of the common amino acids are *\_*-amino acids. They have a carboxyl group and an amino group bonded to the same carbon atom (thecarbon). 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.

In addition to these 20 amino acids there are many less common ones. Some are residues modified after a protein has been synthesized; others are amino acids present in living organisms but not as constituents of proteins. The common amino acids of proteins have been assigned three-letter abbreviations and one-letter symbols which are used as shorthand to indicate the composition and sequence of amino acids polymerized in proteins.

**Protein structure**

**Proteins** are organic compounds made of **“amino acids”** joined together by “**peptide linkages**”.





**Anatomy of an amino acid. Except for proline and its derivatives, all of the amino acids commonly found in proteins possess this type of structure.**



**The *α*-COOH and *α*-NH3+ groups of two amino acids can react with the resulting loss of a water molecule to form a covalent amide bond.**

For all the common amino acids except glycine, the carbon is bonded to four different groups: a carboxyl group, an amino group, an R group, and a hydrogen atom in glycine, the R group is another hydrogen atom. The carbon atom is thus a chiral center. Because of the tetrahedral arrangement of the bonding orbitals around the carbon atom, the four different groups can occupy two unique spatial arrangements, and thus amino acids have two possible stereoisomers. Since they are nonsuper imposable mirror images of each other, the two forms represent a class of stereoisomers called enantiomers. All molecules with a chiral center are also optically active that is, they rotate plane-polarized light.

Special nomenclature has been developed to specify the absolute configuration of the four substituents of asymmetric carbon atoms. The absolute configurations of simple sugars and amino acids are specified by the D, L system based on the absolute configuration of the three-carbon sugar glyceraldehyde, a convention proposed by Emil Fischer knew what groups surrounded the asymmetric carbon of glyceraldehyde but had to guess at their absolute configuration; his guess was later confirmed by x-ray diffraction analysis.) For all chiral compounds, stereoisomers having a configuration related to that of L-glyceraldehyde are designated L, and stereoisomers related to D-glyceraldehyde are designated D. The functional groups of L-alanine are matched with those of L glyceraldehyde by aligning those that can be interconverted by simple, one-step chemical reactions. Thus the carboxyl group of L-alanine occupies the same position about the chiral carbon as does the aldehyde group of L-glyceraldehyde, because an aldehyde is readily converted to a carboxyl group via a one-step oxidation.

Historically, the similar l and d designations were used for levorotatory (rotating light to the left) and dextrorotatory (rotating light to the right). However, not all L-amino acids are levorotatory, and the convention was needed to avoid potential ambiguities about absolute configuration. By Fischer’s convention, L and D refer only to the absolute configuration of the four substituents around the chiral carbon, not to optical properties of the molecule. Another system of specifying configuration around a chiral center is the RS system, which is used in the systematic nomenclature of organic chemistry and describes more precisely the configuration of molecules with more than one chiral center.





**Amino Acids Can Be Classified by R Group**

Knowledge of the chemical properties of the common amino acids is central to an understanding of biochemistry. The topic can be simplified by grouping the amino acids into five main classes based on the properties of their R groups in particular, their polarity, or tendency to interact with water at biological pH (near pH 7.0). The polarity of the R groups varies widely, from nonpolar and hydrophobic (water-insoluble) to highly polar and hydrophilic (water-soluble). The structures of the 20 common amino acids. Within each class there are gradations of polarity, size, and shape of the R groups. 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. Although it is formally nonpolar, its very small side chain makes no real contribution to hydrophobic interactions. 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. The secondary amino (imino) group of proline residues is held in a rigid conformation that reduces the structural flexibility of polypeptide regions containing proline.

**Aromatic R Groups**

Phenylalanine, tyrosine, and tryptophan, with their aromatic side chains, are relatively nonpolar (hydrophobic). All can participate in hydrophobic interactions. The hydroxyl group of tyrosine can form hydrogen bonds, and it is an important functional group in some enzymes. Tyrosine and tryptophan are significantly more polar than phenylalanine, because of the tyrosine hydroxyl group and the nitrogen of the tryptophan indole ring. Tryptophan and tyrosine, and to a much lesser extent phenylalanine, absorb ultraviolet light. This accounts for the characteristic strong absorbance of light by most proteins at a wavelength of 280 nm, a property exploited by researchers in the characterization of proteins.

**Polar, Uncharged R Groups**

The R groups of these amino acids are more soluble in water, or more hydrophilic, than those of the nonpolar amino acids, because they contain functional groups that form hydrogen bonds with water. This class of amino acids includes serine, threonine, cysteine, asparagine, and glutamine. The polarity of serine and threonine is contributed by their hydroxyl groups; that of cysteine by its sulfhydryl group; and that of asparagine and glutamine by their amide groups. Asparagine and glutamine are the amides of two other amino acids also found in proteins, aspartate and glutamate, respectively, to which asparagine and glutamine are easily hydrolyzed by acid or base. Cysteine is readily oxidized to form a covalently linked dimeric amino acid called cystine, in which two cysteine molecules or residues are joined by a disulfide bond. The disulfide-linked residues are strongly hydrophobic (nonpolar). Disulfide bonds play a special role in the structures of many proteins by forming covalent links between parts of a protein molecule or between two different polypeptide chains. Positively Charged (Basic) R Groups The most hydrophilic R groups are those that are either positively or negatively charged. The amino acids in which the R groups have significant positive charge at pH 7.0 are lysine, which has a second primary amino group at the position on its aliphatic chain; arginine, which 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 His residue facilitates the reaction by serving as a proton donor/acceptor. 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.




















****

****

**Esential and non-essential amino acids:**

There are 20 standard amino acids which differ in their side chain (R).

Some of them are considered “essential” since they cannot be synthesized in our body and must be therefore provided in the diet (e.g. tryptophan & phenylalanine), while others are “non-esential” and can be synthesized in the body (e.g. alanine & cysteine).

**Uncommon Amino Acids Also Have Important Functions**

In addition to the 20 common amino acids, proteins may contain residues created by modification of common residues already incorporated into a polypeptide. Among these uncommon amino acids are 4-hydroxyproline, a derivative of prolin and 5-hydroxylysine, derived from lysine. The former is found in plant cell wall proteins, and both are found in collagen, a fibrous protein of connective tissues. 6-NMethyllysine is a constituent of myosin, a contractile protein of muscle. Another important uncommon amino acid is carboxyglutamate, found in the bloodclotting protein prothrombin and in certain other proteins that bind Ca2+ as part of their biological function.

More complex is desmosine, a derivative of four Lys residues, which is found in the fibrous protein elastin. Selenocysteine is a special case. This rare amino acid residue is introduced during protein synthesis rather than created through a postsynthetic modification. It contains selenium rather than the sulfur of cysteine. Actually derived from serine, selenocysteine is a constituent of just a few known proteins. Some 300 additional amino acids have been found in cells. They have a variety of functions but are not constituents of proteins Ornithine and citrulline.





**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 zwitterion (German for “hybrid ion”). A zwitterion can act as either an acid (proton donor):





Substances having this dual nature are amphoteric and are often called ampholytes (from “amphoteric electrolytes”). A simple monoamino monocarboxylic amino acid, such as alanine, is a diprotic acid when fully protonated it has two groups, the COOH group and the NH3 group, that can yield protons:



Consider the acid–base behavior of glycine, the simplest amino acid. At low pH, both the amino and carboxyl groups are protonated and the molecule has a net positive charge. If the counterion in solution is a chloride ion, this form is referred to as **glycine hydrochloride.** If the pH is increased, the carboxyl group is the first to dissociate, yielding the neutral zwitterionic species Gly0. A further increase in pH eventually results in dissociation of the amino group to yield the negatively charged **glycinate.** If we denote these three forms as Gly+, Gly0, and Gly-, we can write the first dissociation of Gly+ as



and the dissociation constant *K*1 as





**The ionic forms of the amino acids, shown without consideration of any ionizations on the side chain. The cationic form is the low pH form, and the titration of the cationic species with base yields the zwitterion and finally the anionic form.**

**References:**

* Lehninger Principles of Biochemistry (Nelson W. H. Freeman. 4th Ed, 2004).
* Biochemistry (Reginald H. Garrett and Charles M. Grisham, University of Virginia 4th Ed, 2010).