Consortium for
Educational
Communication

Module on

Amino acids

By

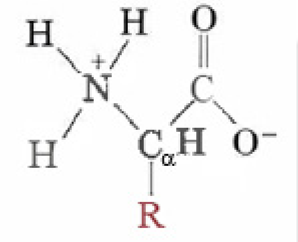
Dr. Shaista Qadir
Sr. Assistant Professor
Govt. Degree College for Women
M.A. Road, Srinagar
9469106443

Introduction

Every living cell contains protein; indeed, they are the most abundant macromolecules in living cells. Proteins also occur in great variety; thousands of different kinds may be found in a single cell. Moreover, proteins exhibit great diversity in their biological function. In animals they help to form supporting and protective structures such as cartilage, skin, nail, hair and muscle. They are major constituents of enzymes and hormones in all organisms. The name "protein" comes from the Greek word "protos" which means first or foremost. While plants are able to utilize inorganic sources of nitrogen such as ammonia, nitrates and nitrites, man and other animals are for the most part dependent on a source of amino acids to build their body proteins. All proteins, whether from the most ancient line of bacteria or from the most complex forms of life, are polymers of amino acids, with each amino acid residue joined to its neighbor by a specific type of covalent bond. (The term residue reflects the loss of the elements of water when one amino acid is joined to another). Proteins can be reduced 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 amino acid to be discovered in protein 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, glutamate from wheat gluten tyrosine from cheese (Greek tyros "cheese" and glycine (Greek *glycos* "sweet") was so named because of its sweet taste.

Amino Acids have common structural features

The naturally occurring amino acids have a common structure. Amino acids, as the name implies, have two functional groups, an amino group ($-NH_2$) and a carboxyl group ($-COOH$). These groups are joined to a single (aliphatic) carbon. In organic chemistry, the carbon directly attached to a carboxyl group is the alpha (α) position, so the amino acids in proteins are all **alpha-amino acids.** The side chains that distinguish one amino acid from another are attached to the alpha carbon, so the structures are often written as:



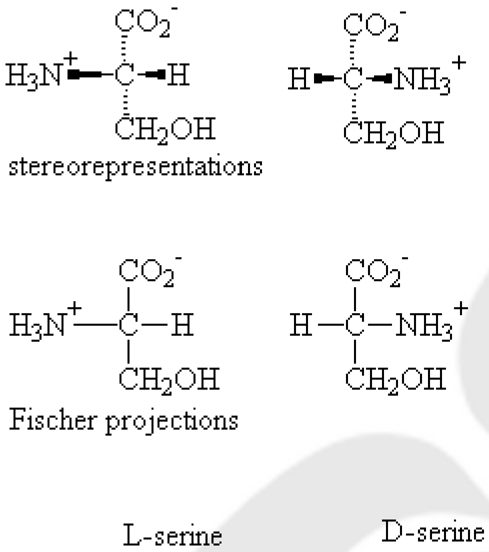
In the simplest case, R=H (amino acetic acid or glycine). In other amino acids, R is an aliphatic, aromatic or heterocyclic residue and may incorporate other functional groups which vary in structure, size, and electric charge, and which influence the solubility of the amino acids in water. In addition to these 20 amino acids there are many less common ones which are not present as constituents of proteins in living organisms. The common amino acids of proteins have been assigned three-letter abbreviations and one-letter symbol, which are used as short hand to indicate the composition and sequence of amino acids polymerised in proteins.

Table of the Amino Acids

Name		Linear Structure
Alanine	**ala** A	CH3-CH(NH2)-COOH
Arginine	**arg** R	HN=C(NH2)-NH-(CH2)3-CH(NH2)-COOH
Asparagine	**asn** N	H2N-CO-CH2-CH(NH2)-COOH
Aspartic Acid	**asp** D	HOOC-CH2-CH(NH2)-COOH
Cysteine	**cys** C	HS-CH2-CH(NH2)-COOH
Glutamic Acid	**glu** E	HOOC-(CH2)2-CH(NH2)-COOH
Glutamine	**gln** Q	H2N-CO-(CH2)2-CH(NH2)-COOH
Glycine	**gly** G	NH2-CH2-COOH
Histidine	**his** H	**N**H-CH=N-CH=**C**-CH2-CH(NH2)-COOH
Isoleucine	**ile** I	CH3-CH2-CH(CH3)-CH(NH2)-COOH
Leucine	**leu** L	(CH3)2-CH-CH2-CH(NH2)-COOH
Lysine	**lys** K	H2N-(CH2)4-CH(NH2)-COOH
Methionine	**met** M	CH3-S-(CH2)2-CH(NH2)-COOH
Phenylalanine	**phe** F	Ph-CH2-CH(NH2)-COOH
Proline	**pro** P	**N**H-(CH2)3-**C**H-COOH
Serine	**ser** S	HO-CH2-CH(NH2)-COOH
Tryptophan	**trp** W	**Ph**-NH-CH=**C**-CH2-CH(NH2)-COOH
Tyrosine	**tyr** Y	HO-Ph-CH2-CH(NH2)-COOH
Valine	**val** V	(CH3)2-CH-CH(NH2)-COOH

Two conventions are used to identify the carbons in an amino acid. The additional carbons in an R group are commonly designated as α, β,γ,δ,ε, and so forth, proceeding out from the α,carbon. For most other organic molecules, carbon atoms are simply numbered from one end, giving highest priority (C-1) to the carbon with the substituent containing the atom of highest atomic number. Within this latter convention, the carboxyl carbon of an amino acid would be C-1 and the α carbon would be C-2. In some cases, such as amino acids with heterocyclic R groups, the Greek lettering system is ambiguous and the numbering convention is used

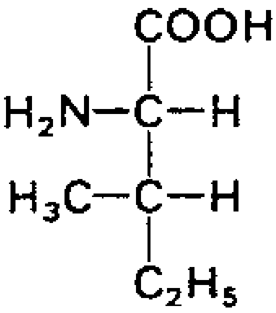
Amino acid chilarity Amino acids (except for glycine have a chiral carbon atom i.e α- carbon is bounded to four different groups; a carboxyl group, an amino group; an R group, and a hydrogen atom, in glycine, the R group is another hydrogen atom. The α- carbon is thus a chiral center. Because of the tetrahedral arrangement of the bonding orbitals around the α-carbon atom, the four different groups can occupy two unique spatial arrangements, and thus amino acids have two possible stereoisomers. Since they are non-superimposable mirror images of each other, the two forms represent a class of stereoisomers called enantiomers. All molecules with a chiral centre are also optically active-that is, they rotate plane polarized light, with the direction of the rotation differing for different stereoisomers. The classification and naming of stereoisomers is based on absolute configuration of the four substituents of the asymmetric carbon atom. For this purpose a reference compound , the 3-carbon sugar, glyceraldehyde, has been chosen, to which all other optically active compounds are compared. Glyceraldehyde is the smallest sugar to have an asymmetrical carbon atom, the naming of its configuration is established by ×-ray diffraction analysis.The **st**ereoisomers of all chiral compounds having a configuration related to that of L-glyceraldehyde are referred to as **L** (for levorotatory, meaning left), and the stereoisomers related to D- glyceraldehyde are designated D (The D comes from dextrorotatory, meaning right.) The amino acids found in proteins are L-alpha amino acids. D-amino acids are not naturally found in proteins and are not involved in the metabolic pathways of eukaryotic organisms, although they are important in the structure and metabolism of bacteria. For example, D-glutamic acid and D-alanine are structural components of certain bacterial cell walls. Similarly some peptide antibiotics, such as bacitracin, also contain D-amino acids. All amino acids found in proteins occur in the L-configuration about the chiral carbon atom. For example, take a look at the stereo-representations of serine.



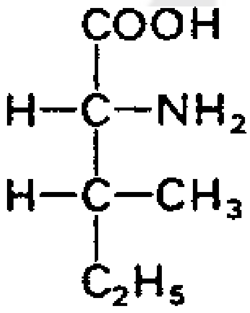
Enantiomers of Serine

2001 A.M. Helmenstine
Licensed to About, Inc.

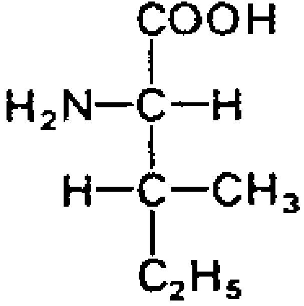
Isoleucine, threonine and 4-hydroxyproline have two asymmetric C-atoms, thus each has four isomers:



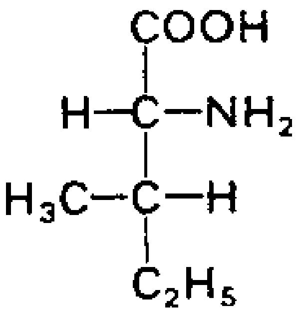
L-Isoleucine (2S: 3S)-Isoleucine (Common in proteins)



D-Isoleucine (2R: 3R)-Isoleucine



L-allo-Isoleucine (2S: 3R)-Isoleucine



D-allo-Isoleucine (2R: 3S)-Isoleucine

Amino acids can be classified by R Group

While there are several methods of categorizing them, one of the most common is to group them according to the nature of their side chains (R-groups), particularly their polarity (distribution of electric charge) or tendency to interact with water at biological pH (near pH 7.0). The polarity of the R-groups varies widely, from totally nonpolar or hydrophobic (water-insoluble) to highly polar or hydrophilic (water-soluble)

There are 3 main classes of amino acids

Nonpolar Side Chain: The R groups of these amino acid have either aliphatic or aromatic groups. Glycine, alanine,valine, leucine, isoleucine and proline have aliphatic R groups. **Proline** is unique (the cyclic amino acid) among these amino acids. It has four carbons, with the alpha amino group bonded not only to the alpha carbon but also to the last side chain carbon. The cyclic side chain means that proline is conformationally rigid. That is, the carbon-carbon bonds of proline do not rotate in solution. Other amino acids are more flexible in solution. Phenylalanine, tyrosine and tryptophan have aromatic R groups.

Polar, Uncharged Side Chains: The R-Groups of these amino acids are more soluble in water, or 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, methionine, asparagine and glutamine. The polarity of Serine and threonine is contributed by their hydroxyl groups; that of cysteine and methionine by their sulfur atom and Asparagine and glutamine by their amide groups.

Charged Side Chains: These include amino acids with negatively charged (Acidic) R groups and amino acids with positively charged (basic) R groups. The former group includes aspartate and glutamate, each with a second carboxyl group. The latter includes lysine, arginine and histidine.

Uncommon Amino Acids 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 proline 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-N-Methylysine is a constituent of myosin, a contractile protein of muscle.. Another important uncommon amino acid is γ-carboxyglutamate, found in the blood clotting protein prothombin and in certain other proteins that bind Ca^{2+} as part of their biological function. More complex is desmocine, 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 systeine. 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 constituent of proteins. Ornithine and citrulline deserve special note because they are key intermediates (metabolites) in the biosynthesis of arginine and in the urea cycle

Based on their nutritional/physiological roles, amino acids can be differentiated as:

Essential amino acids: Eight of the amino acids are designated as essential amino acids. These are Valine, leucine, isoleucine, phenylalanine, tryptophan, methionine, threonine and histidine. These amino acids cannot be synthesized by humans in adequate amounts to sustain growth and health and must be supplied by the diet.

Nonessential amino acids:

These are also necessary for health but can be synthesized by humans from other amino acids and nitrogenous compounds. These include Glycine, alanine, proline, serine, cysteine, tyrosine, asparagine, glutamine, aspartic acid, and glutamic acid.

Semi- essential amino acids. These are **essential for infants, e.g.,** lysine and arginine.

Properties of Amino acids

Amino acids are colourless, crystalline solids, water-soluble high melting solids and behave like salts.

Amino Acids Can Act as Acids and Bases

The carboxyl and amino groups of the amino acids can respectively donate a proton(acid) to and accept a proton (base) from water. This exchange happens simultaneously in solution so that the amino acids form doubly ionized species, termed **zwitterions** (from German *zwei,* meaning hybrid ion) in solution as:

The zwitter ions are dipolar, when the proteins are present in an aqueous solution, the carboxyl group can loose a proton and amino group can accept a proton giving rise to a dipolar ion which is neutral but has both positive and negative charges.

In the ionic form, amino acids show slow amphoteric behavior as they react with acids and bases. In the acidic solution, the carboxyl group (- COO^-) accepts a proton and gets converted to carboxyl substituent (- COOH) while in basic solution the ammonium ($+NH_3$) loses a proton and gets converted to amino group ($-NH_2$).

The movement of the amino acid in an electric field is different in acid and alkaline solution. In acidic solution, an amino acid exists as a positive ion and migrates towards the cathode in an electric field, while in alkaline solution it exists as a negative ion and migrates towards anode. At a certain pH zwitterions exists as a neutral ion and do not migrate to either electrode. This pH is known as the isoelectric point of the amino acid.

Solubility of amino acids

The solubilities of amino acids in water are highly variable. Besides the extremely soluble proline, hydroxyproline, glycine and alanine are also quite soluble. Other amino acids are significantly less soluble, with cysteine and tyrosine having particularly low solubilities. Addition of acids or bases improves the solubility through salt formation. The presence of other amino acids, in general, also brings about an increase in solubility. Thus, the extent of solubility of amino acids in a protein hydrolysate is different than that observed for the individual components. The solubility in organic solvents is not very good because of the polar characteristics of the amino acids. All amino acids are insoluble in ether. Only cysteine and proline are relatively soluble in ethanol (1.5 g/100 g at 19 ◦C). Methionine, arginine, leucine

(0.0217 g/100 g;25 ◦C), glutamic acid (0.00035 g/100 g; 25 ◦C),phenylalanine, hydroxy-proline, histidine and tryptophan are sparingly soluble in ethanol. The solubility of isoleucine in hot ethanol is relatively high (0.09 g/100 g at 20 ◦C; 0.13 g/100 g at 78–80 ◦C).

UV absorption

Amino acids such as phenylalanine, tyrosine and tryptophan absorb in the UV range of the spectrum with absorption maxima at 200–230 nm and 250–290 nm .Dissociation of the phenolic HO-group of tyrosine shifts the absorption curve by about 20 nm towards longer wavelength. Absorption readings at 280 nm are used for the determination of proteins and peptides. Histidine, cysteine and methionine absorb between 200 and 210 nm.