B.Sc. Semester-VI
Organic Chemistry
Paper-XIV

1. Amino Acids, Peptides, Proteins and Nucleic Acids

Coverage:

1. **Amino Acids and Peptides-1**

Dr. Rajeev Ranjan

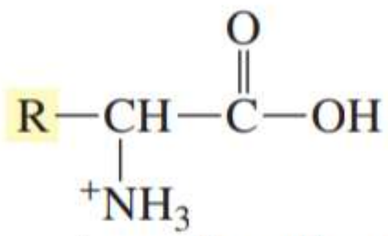
University Department of Chemistry

Dr. Shyama Prasad Mukherjee University, Ranchi

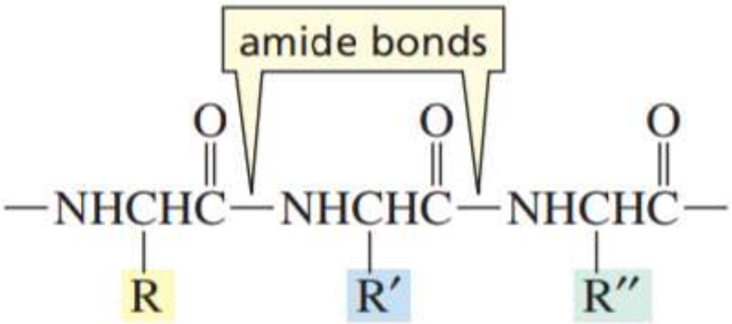
Amino Acids and Peptides

Amino acid polymers can be composed of any number of monomers. A **dipeptide** contains two amino acid residues, a **tripeptide** contains three, an **oligopeptide** contains three to 10, and a **polypeptide** contains many amino acid residues. Proteins are naturally occurring polypeptides that are made up of 40 to 4000 amino acid residues.

From the structure of an amino acid, we can see that the name is not very precise. The compounds commonly called amino acids are more precisely called α-amino-carboxylic acids.



α-aminocarboxylic acid
an amino acid



amino acids are linked together by amide bonds

Proteins and peptides serve many functions in biological systems. Some protect organisms from their environment or impart strength to certain biological structures. Hair, horns, hooves, feathers, fur, and the tough outer layer of skin are all composed largely of a **structural protein** called keratin. Collagen, another structural protein, is a major component of bones, muscles, and tendons. Some proteins have other protective functions. Snake venoms and plant toxins, for example, protect their owners from other species, blood-clotting proteins protect the vascular system when it is injured, and antibodies and protein antibiotics protect us from disease. A group of proteins called **enzymes** catalyzes the chemical reactions that occur in living systems, and some of the hormones that regulate these reactions are peptides. Proteins are also responsible for many physiological functions, such as the transport and storage of oxygen in the body and the contraction of muscles.

Classification and Nomenclature of Amino Acids

The structures of the 20 most common naturally occurring amino acids and the frequency with which each occurs in proteins are shown in Table 1. Other amino acids occur in nature, but only infrequently. All amino acids except proline contain a primary amino group. Proline contains a secondary amino group incorporated into a five-membered ring. The amino acids differ only in the substituent (R) attached to the α-carbon. The wide variation in these substituents (called side chains) is what gives proteins their great structural diversity and, as a consequence, their great functional diversity.

Table 1. The Most Common Naturally Occurring Amino Acids

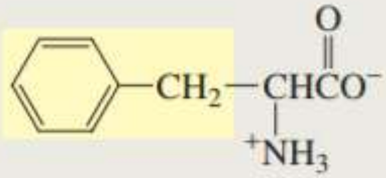
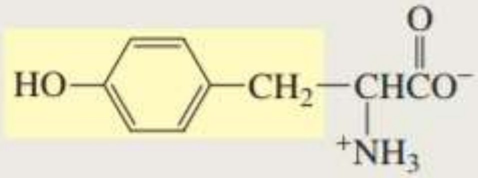
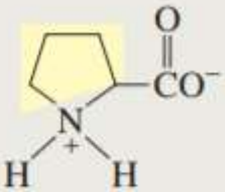
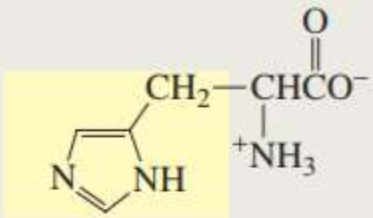
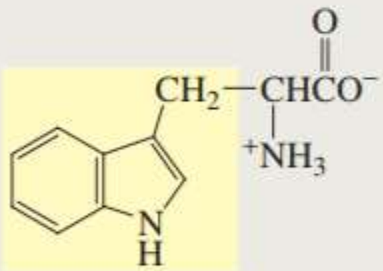
	Formula	Name	Abbreviations		Average relativ abundance in proteins
Aliphatic side chain amino acids	$H-CH(^{+}NH_3)C(=O)O^-$	Glycine	Gly	G	7.5%
	$CH_3-CH(^{+}NH_3)C(=O)O^-$	Alanine	Ala	A	9.0%
	$CH_3CH(CH_3)-CH(^{+}NH_3)C(=O)O^-$	Valine*	Val	V	6.9%
	$CH_3CH(CH_3)CH_2-CH(^{+}NH_3)C(=O)O^-$	Leucine*	Leu	L	7.5%
	$CH_3CH_2CH(CH_3)-CH(^{+}NH_3)C(=O)O^-$	Isoleucine*	Ile	I	4.6%
Hydroxy-containing amino acids	$HOCH_2-CH(^{+}NH_3)C(=O)O^-$	Serine	Ser	S	7.1%
	$CH_3CH(OH)-CH(^{+}NH_3)C(=O)O^-$	Threonine*	Thr	T	6.0%

* Essential amino acids

Table 1. The Most Common Naturally Occurring Amino Acids

	Formula	Name	Abbreviations		Average relative abundance in proteins
Sulfur-containing amino acids	$HSCH_2-CH(^+NH_3)-C(=O)O^-$	Cysteine	Cys	C	2.8%
	$CH_3SCH_2CH_2-CH(^+NH_3)-C(=O)O^-$	Methionine*	Met	M	1.7%
Acidic amino acids	$^-OC(=O)CH_2-CH(^+NH_3)-C(=O)O^-$	Aspartate (aspartic acid)	Asp	D	5.5%
	$^-OC(=O)CH_2CH_2-CH(^+NH_3)-C(=O)O^-$	Glutamate (glutamic acid)	Glu	E	6.2%
Amides of acidic amino acids	$H_2NC(=O)CH_2-CH(^+NH_3)-C(=O)O^-$	Asparagine	Asn	N	4.4%
	$H_2NC(=O)CH_2CH_2-CH(^+NH_3)-C(=O)O^-$	Glutamine	Gln	Q	3.9%
Basic amino acids	$H_3\overset{+}{N}CH_2CH_2CH_2CH_2-CH(^+NH_3)-C(=O)O^-$	Lysine*	Lys	K	7.0%

Table 1. The Most Common Naturally Occurring Amino Acids

	Formula	Name	Abbreviations		Average relative abundance in proteins
Benzene-containing amino acids		Phenylalanine*	Phe	F	3.5%
		Tyrosine	Tyr	Y	3.5%
Heterocylic amino acids		Proline	Pro	P	4.6%
Heterocyclic amino acids (continued)		Histidine*	His	H	2.1%
		Tryptophan*	Trp	W	1.1%

* Essential amino acids

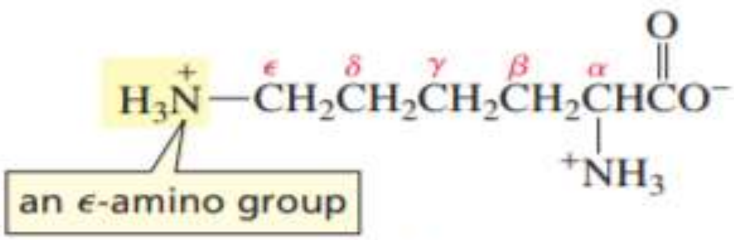
The amino acids are almost always called by their common names. Often, the name tells you something about the amino acid. For example, glycine got its name because of its sweet taste (*glykos* is Greek for "sweet"), and valine, like valeric acid, has five carbon atoms. Asparagine was first found in asparagus, and tyrosine was isolated from cheese (*tyros* is Greek for "cheese").

Dividing the amino acids into classes makes them easier to learn. The aliphatic side chain amino acids include glycine, the amino acid in which R = H, and four amino acids with alkyl side chains. Alanine is the amino acid with a methyl side chain, and valine has an isopropyl side chain. Can you guess which amino acid—leucine or isoleucine—has an isobutyl side chain? If you gave the obvious answer, you guessed incorrectly. Isoleucine does *not* have an "iso" group; it is leucine that has an isobutyl substituent—isoleucine has a *sec*-butyl substituent. Each of the amino acids has both a three-letter abbreviation (the first three letters of the name in most cases) and a single-letter abbreviation.

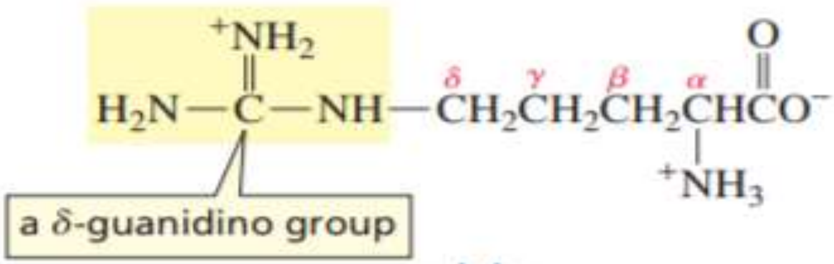
Two amino acid side chains—serine and threonine—contain alcohol groups. Serine is an HO-substituted alanine and threonine has a branched ethanol substituent. There are also two sulfur-containing amino acids: Cysteine is an HS-substituted alanine and methionine has a 2-methylthioethyl substituent.

There are two acidic amino acids (amino acids with two carboxylic acid groups): aspartate and glutamate. Aspartate is a carboxy-substituted alanine and glutamate has one more methylene group than aspartate. (If their carboxyl groups are protonated, they are called aspartic acid and glutamic acid, respectively.) Two amino acids—asparagine and glutamine—are amides of the acidic amino acids; asparagine is the amide of aspartate and glutamine is the amide of glutamate. Notice that the obvious one-letter abbreviations cannot be used for these four amino acids because A and G are used for alanine and glycine. Aspartic acid and glutamic acid are abbreviated D and E, and asparagine and glutamine are abbreviated N and Q.

There are two basic amino acids (amino acids with two basic nitrogen-containing groups): lysine and arginine. Lysine has an ϵ-amino group and arginine has a δ-guanidino group. At physiological pH, these groups are protonated. The ϵ and δ can remind you how many methylene groups each amino acid has.



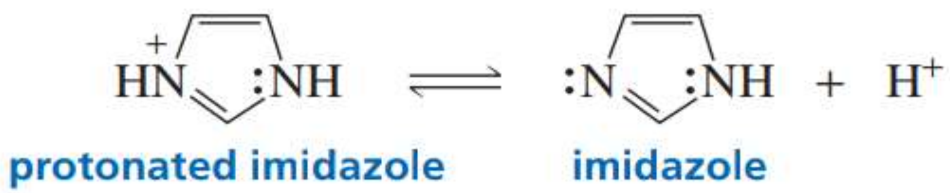
lysine



arginine

Two amino acids—phenylalanine and tyrosine—contain benzene rings. As its name indicates, phenylalanine is phenyl-substituted alanine. Tyrosine is phenylalanine with a *para*-hydroxy substituent.

Proline, histidine, and tryptophan are heterocyclic amino acids. Proline has its nitrogen incorporated into a five-membered ring—it is the only amino acid that contains a secondary amino group. Histidine is an imidazole-substituted alanine. Imidazole is an aromatic compound because it is cyclic and planar and has three pairs of delocalized π electrons. The pK_a of a protonated imidazole ring is 6.0, so the ring will be protonated in acidic solutions and nonprotonated in basic solutions .



Tryptophan is an indole-substituted alanine . Like imidazole, indole is an aromatic compound. Because the lone pair on the nitrogen atom of indole is needed for the compound's aromaticity, indole is a very weak base. (The pK_a of protonated indole is −2.4.) Therefore, the ring nitrogen in tryptophan is never protonated under physiological conditions.

Ten amino acids are *essential amino acids*. We humans must obtain these 10 **essential amino acids** from our diets because we either cannot synthesize them at all or cannot synthesize them in adequate amounts. For example, we must have a dietary source of phenylalanine because we cannot synthesize benzene rings. However, we do not need tyrosine in our diets, because we can synthesize the necessary amounts from phenylalanine. The essential amino acids are denoted by red asterisks (*) in Table 1. Although humans can synthesize arginine, it is needed for growth in greater amounts than can be synthesized. So arginine is an essential amino acid for children, but a nonessential amino acid for adults. Not all proteins contain the same amino acids. Bean protein is deficient in methionine, for example, and wheat protein is deficient in lysine. They are *incomplete* proteins: They contain too little of one or more essential amino acids to support growth. Therefore, a balanced diet must contain proteins from different sources.

Dietary protein is hydrolyzed in the body to individual amino acids. Some of these amino acids are used to synthesize proteins needed by the body, some are broken down further to supply energy to the body, and some are used as starting materials for the synthesis of nonprotein compounds the body needs, such as adrenaline, thyroxine, and melanin .

THANK YOU