# Amino acid Structure and Functional Group Properties

- The amino acids are regarded as building blocks of proteins.
- Amino acids are nitrogenous compounds made of a central α-*carbon* atom attached with four different groups; an acidic carboxyl (— COOH), a basic amino (— $NH_2$) group, a hydrogen and a R group.
- R can be as simple as a hydrogen atom (H) or a methyl group (— $CH_3$) or a more complex structure.
- There are 20 amino acids (differ in their –R group) found in almost all the proteins called as standard amino acids
- The **α-carbon** of all the amino acids is asymmetric except in glycine where the α-carbon is symmetric.
- Amino acids (except glycine) are optically active and exist in two isomeric forms: those having — $NH_2$ group to the right are designated as D-forms and those having — $NH_2$ group to the left as L-forms.
- Naturally occurring amino acids in proteins are predominantly existed in L form although rare cases of D-amino acids are also observed

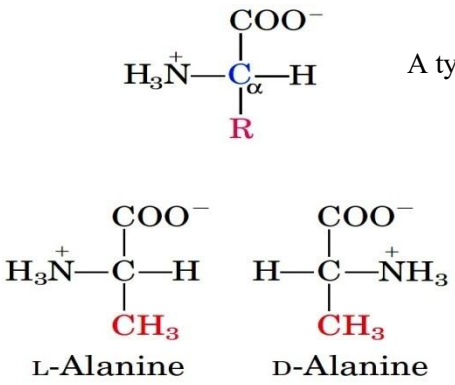



Figure 1: Structure of a typical amino acid and two isomeric forms of alanine

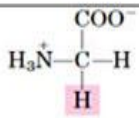
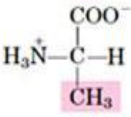
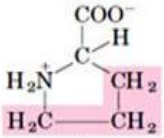
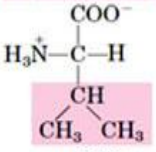
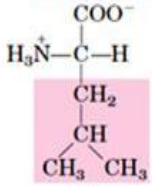
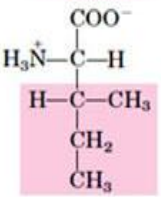
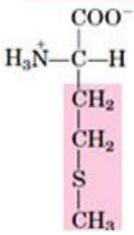
**Classification of Amino acid**

- Amino acid classification is based on the polarity of the R groups (*i.e.*, their tendency to interact with water at biological pH) (around pH 7.0).
- The polarity of the R groups varies widely, from non-polar and hydrophobic (water-insoluble) to highly polar and hydrophilic (water-soluble).
- **20 amino acids are divided into five major groups**
  - Non-polar, aliphatic R groups amino acids
  - Aromatic R group amino acids
  - Polar, uncharged R groups amino acids
  - Positive charged R group amino acids
  - Negative charged R group amino acids

**Non-polar, aliphatic R groups amino acids**

- The R groups in this class of amino acids are nonpolar and hydrophobic
- The side chain R groups of Ala, Leu, Val and Ile are branched and highly hydrophobic. They tend to remain clustered in the interior of the proteins in order to stabilize 3D structure through hydrophobic interactions
- Glycine has the simplest structure with no optical activities
- The aliphatic side chain of the Pro is covalently linked to the amino group forming a cyclic structure (pyrrolidine group). The imino group of Pro residues is held in a rigid conformation that lowers the structural flexibility of polypeptide regions containing Pro.

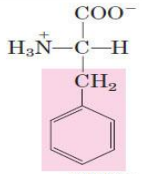
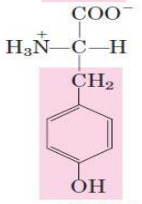
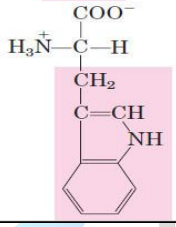
**Table 1: Structure and properties of Non-polar, aliphatic R group amino acids**

| Name | Symbol | | Mr | Structure | $pK_1$ (-COOH) | $pK_2$ (-NH3) | $pK_R$ (R group) | PI | Hydropathy Index | R Group |
|---|---|---|---|---|---|---|---|---|---|---|
| Glycine | Gly | G | 75 |  | 2.34 | 9.60 | | 5.97 | -0.4 | Hydrogen |
| Alanine | Ala | A | 89 |  | 2.34 | 9.69 | | 6.01 | 1.8 | Methyl |
| Proline | Pro | P | 115 |  | 1.99 | 10.96 | | 6.48 | 1.6 | Pyrrolidine |
| Valine | Val | V | 117 |  | 2.32 | 9.62 | | 5.97 | 4.2 | Branched Chain |
| Leucine | Leu | L | 131 |  | 2.36 | 9.60 | | 5.98 | 3.8 | Branched Chain |
| Isoleucine | Ile | I | 131 |  | 2.36 | 9.68 | | 6.02 | 4.5 | Branched Chain |
| Methionine | Met | M | 149 |  | 2.28 | 9.21 | | 5.74 | 1.9 | Thioether |

**Aromatic R group amino acids**

- Trp, Tyr and Phe have a non-polar aromatic group in their side chain R.
- Trp and Tyr are significantly more polar than Phe, because of the tyrosine hydroxyl group and the nitrogen of the tryptophan indole ring.
- Trp and Tyr, and to a much lesser extent Phe, absorb ultraviolet light at 280 nm and this property is exploited by researchers in the characterization of proteins

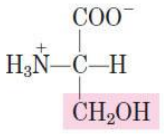
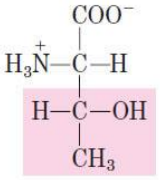
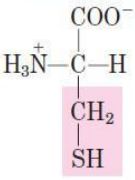
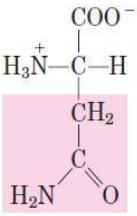
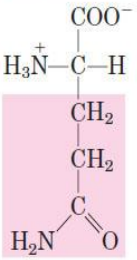
**Table 2: structure and properties of Aromatic R group amino acids**

| Name | Symbol | | Mr | Structure | $pK_1$ (-COOH) | $pK_2$ (-NH3) | $pK_R$ (R group) | PI | Hydropathy Index | R Group |
|---|---|---|---|---|---|---|---|---|---|---|
| Phenylalanine | Phe | F | 165 |  | 1.83 | 9.13 | | 5.48 | 2.8 | Phenyl group |
| Tyrosine | Tyr | Y | 181 |  | 2.20 | 9.11 | 10.07 | 5.66 | -1.3 | Phenol group |
| Tryptophan | Trp | W | 204 |  | 2.38 | 9.39 | | 5.89 | - 0.9 | Indol group |

**Polar and uncharged R groups amino acids**

- The R groups Ser (-OH), Thr (-OH), Cys (-SH), Asn ($-CONH_2$) and Gln ($-CONH_2$) are polar and soluble in water.
- Cysteine is readily oxidized to form a covalently linked dimeric amino acid called cystine, in which two cysteine 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 polypeptide molecule or between two different polypeptide chains

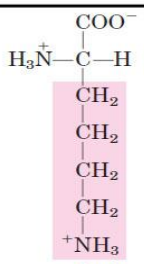
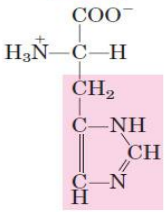
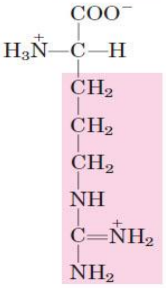
**Table 3: Structure and properties of polar R group amino acids**

| Name | Symbol | | Mr | Structure | $pK_1$ (-COOH) | $pK_2$ (-NH3) | $pK_R$ (R group) | PI | Hydropathy Index | R Group |
|---|---|---|---|---|---|---|---|---|---|---|
| Serine | Ser | S | 105 |  | 2.21 | 9.15 | | 5.68 | - 0.8 | Hydroxyl |
| Threonine | Thr | T | 119 |  | 2.11 | 9.62 | | 5.87 | - 0.7 | Hydroxyl |
| Cysteine | Cys | C | 121 |  | 1.96 | 10.28 | 8.18 | 5.07 | 2.5 | Sulfhydryl |
| Asparagine | Asn | N | 132 |  | 2.02 | 8.80 | | 5.41 | - 3.5 | Amide |
| Glutamine | Gln | Q | 146 |  | 2.17 | 9.13 | | 5.65 | - 3.5 | Amide |

**Positive charged R group amino acids**

- Lys, His and Arg possessed hydrophilic positively charged R groups (amino, imidazol and guanidinium group respectively) at physiological pH-7.0.
- His residues facilitate many enzyme-catalyzed reactions by serving as proton donors/acceptors

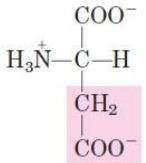
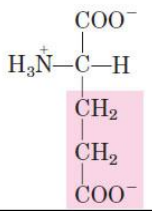
**Table 4: structure and properties of positively charged R group amino acids**

| Name | Symbol | | Mr | Structure | $pK_1$ (-COOH) | $pK_2$ (-NH3) | $pK_R$ (R group) | PI | Hydropathy Index | R Group |
|---|---|---|---|---|---|---|---|---|---|---|
| Lysine | Lys | K | 146 |  | 2.18 | 8.95 | 10.53 | 9.74 | - 3.9 | ε-amino |
| Histidine | His | H | 155 |  | 1.82 | 9.17 | 6.00 | 7.59 | - 3.2 | Imidazol |
| Arginine | Arg | R | 174 |  | 2.17 | 9.04 | 12.48 | 10.76 | - 4.5 | Guanidinium |

**Negative charged R group amino acids**

- 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

**Table 5: structure and properties of egatively charged R group amino acids**

| Name | Symbol | | Mr | Structure | $pK_1$ (-COOH) | $pK_2$ (-NH3) | $pK_R$ (R group) | PI | Hydropathy Index | R Group |
|---|---|---|---|---|---|---|---|---|---|---|
| Aspartate | Asp | D | 133 |  | 1.88 | 9.60 | 3.65 | 2.77 | - 3.5 | β-carboxyl |
| Glutamate | Glu | E | 147 |  | 2.19 | 9.67 | 4.25 | 3.22 | - 3.5 | γ-carboxyl |

**Metabolic classification**

- ***Ketogenic amino acids*** are broken down into ketone bodies. Lysine and Leucine are the only pure ketogenic amino acids.
- ***Mixed ketogenic and glucogenic amino acids*** degraded into both ketone bodies and glucose precursors. These include isoleucine, phenylalanine, tyrosine and tryptophan.
- ***Glucogenic amino acids*** converted into precursors of glucose. They include the rest of amino acids.

**Nutritional classification**

- Humans are incapable of synthesizing half of the 20 common amino acids known as essential amino acids.
- These essential amino acids must be provided in the diet.

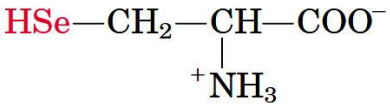
Table 6: **Nonessential and Essential Amino Acids for Humans and the Albino Rat**

| Nonessential | Conditionally essential* | Essential |
|---|---|---|
| Alanine | Arginine | Histidine |
| Asparagine | Cysteine | Isoleucine |
| Aspartate | Glutamine | Leucine |
| Glutamate | Glycine | Lysine |
| Serine | Proline | Methionine |
| | Tyrosine | Phenylalanine |
| | | Threonine |
| | | Tryptophan |
| | | Valine |

*Required to some degree in young, growing animals, and/or sometimes during illness.

**Selenocystein: the 21st amino acid**

- **Selenocysteine** is a rare amino acid residue that is introduced during protein synthesis rather than created through a posttranslational modification.
- It contains selenium rather than the sulfur of cysteine.
- Derived from serine, selenocysteine is a constituent of a few known proteins such as glutathione peroxidase, glycine reductase 5' –deiodinase, thioredoxin reductase.



Selenocysteine

**Non-standard 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*-Methyllysine** is a constituent of myosin, a contractile protein of muscle. Another important uncommon amino acid is **γ-carboxyglutamate-**, found in the blood-clotting protein prothrombin and in certain other proteins that bind $Ca^{+2}$ as part of their biological functions. More complex is **desmosine**, a derivative of four Lys residues, which is found in the fibrous protein elastin.
- Some amino acid residues in a protein may be modified transiently to alter the protein's function.
- The addition of phosphoryl, methyl, acetyl, ADP-ribosyl, or other groups to particular amino acid residues can increase or decrease a protein's activity.
- Phosphorylation is a particularly common regulatory modification.
- Some 300 additional amino acids have been found in cells. They have a variety of functions but are not all constituents of proteins.
- **Ornithine** and **citrulline are not found in the proteins. They are** the key intermediates (metabolites) in the biosynthesis of arginine and in the urea cycle.

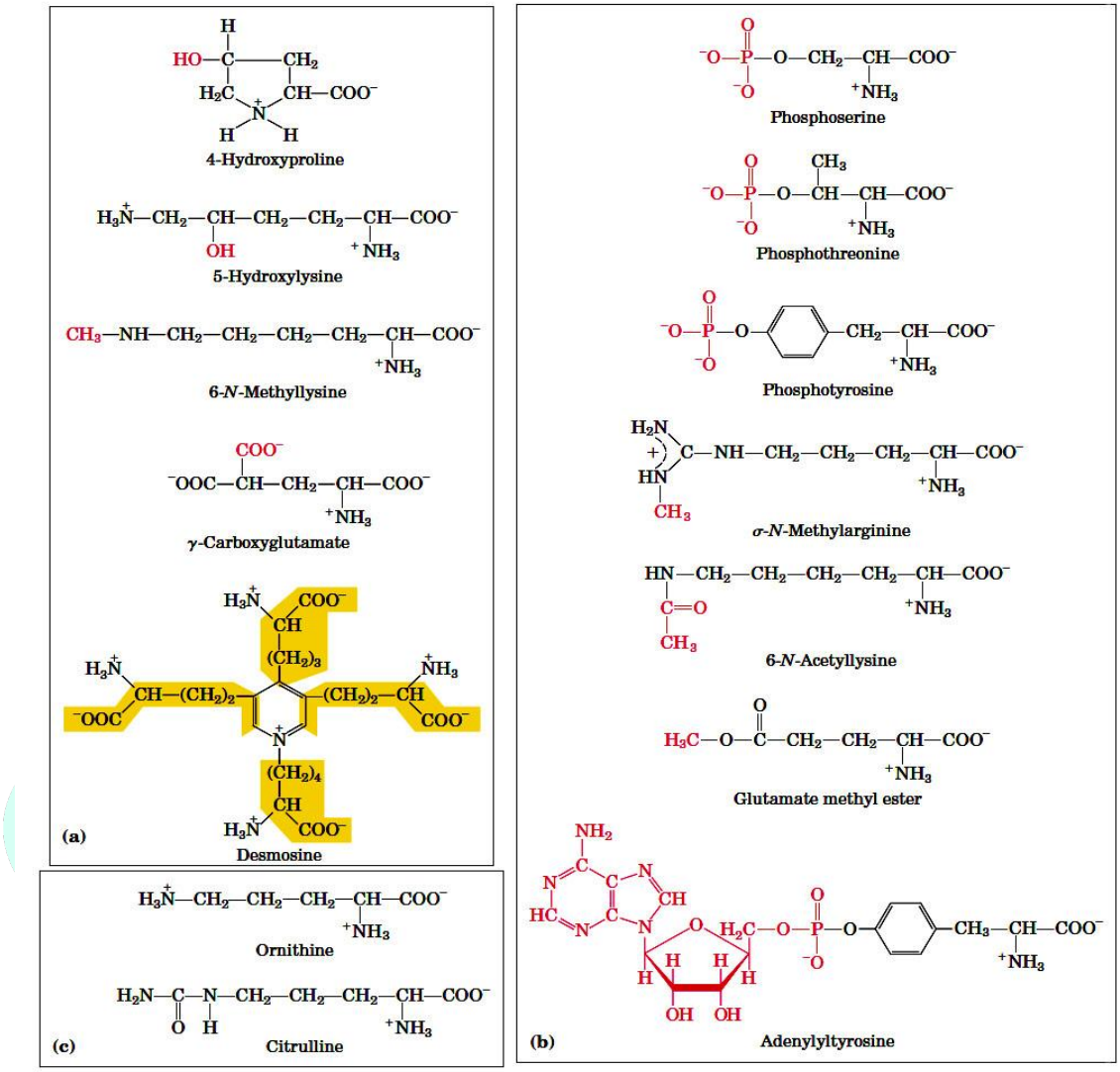



Figure 2: Non-standard amino acids (a) some uncommon amino acids found in proteins are derived from common amino acids. Extra functional groups (red in color) added by modification reactions. Desmosine is formed from four Lys residues (b) Reversible amino acid modifications involved in regulation of protein activities. Phosphorylation, acetylation, methylation and ADP-ribosylation are the most common type of regulatory modification. (c) Ornithine and citrulline are not found in proteins, they are intermediates in the urea cycle.

## References


- Nelson DL and Cox MM. *Lehninger Principles of Biochemistry*, *5th Edition* 2008, W.H. Freeman and Company, New-York
- Voet & Voet *Biochemistry 2nd Edition*, John Wiley and Sons., Inc. Canada
- Berg, Tymoczko and Stryer *Biochemistry 5th Edition,* W.H. Freeman and Company, New-York
- Jain and Jain *Principles of Biochemistry 6th Edition*, S. Chand and Company Ltd, New Delhi, India