An Introduction to Proteins and amino acids

By

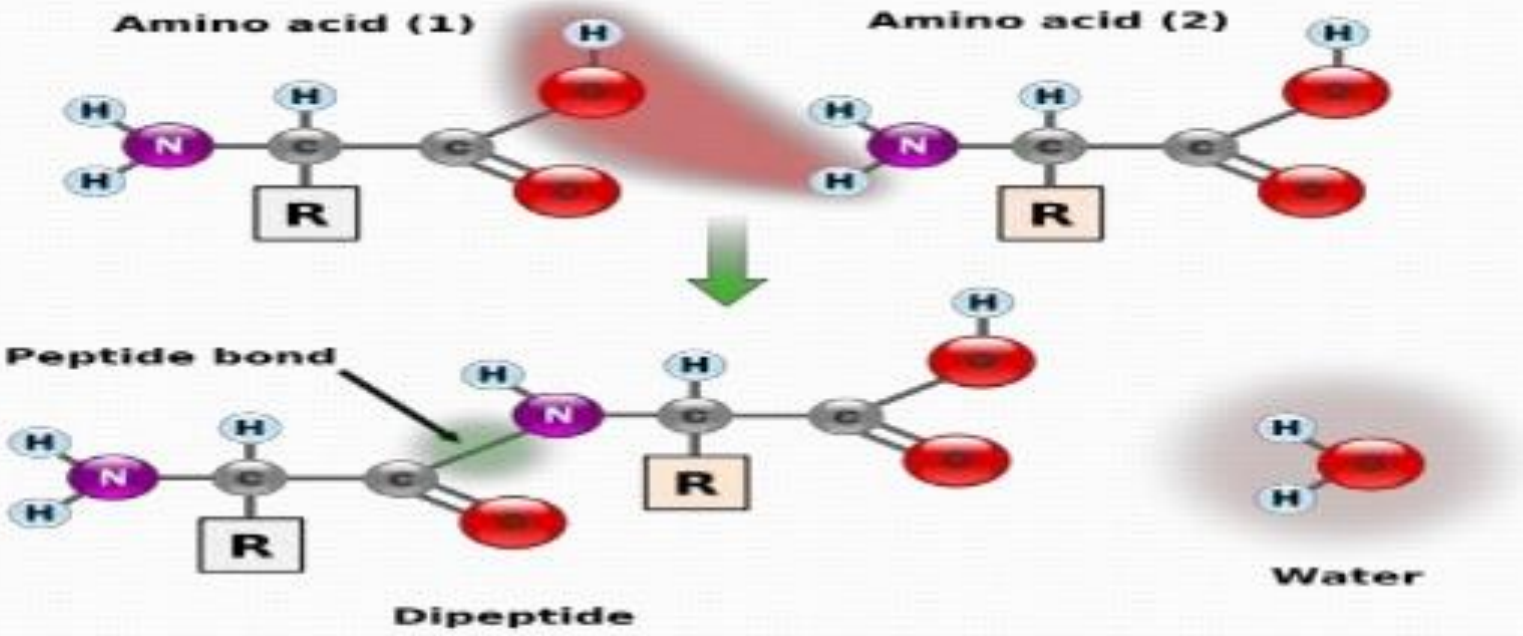
Dr. Burhan Balaky

lecturer

- The word **PROTEIN** comes from Greek language (prota) which means "of primary importance". This name was introduced by Jons Jakob Berzelius in 1838 for large organic compounds with almost equivalent empirical formulas. This name was used because the studied organic compounds were primitive but seems to be very important for animal nutrition.

Protein-Definition

Proteins may be defined as the high molecular weight mixed polymers of alpha-amino acids joined together with peptide linkage.



Definition

Proteins are cellular macromolecules made up of amino acid polymers (polypeptides). The sequence of amino acids, or primary structure of the protein, is dictated by the nucleotide sequence of the gene coding for that particular protein.

- ALL PROTEINS CONTAIN:
  * Carbon (C)
  * Oxygen (O)
  * Nitrogen (N)
  * Hydrogen (H)
  * Sulfur (S)

- Generally speaking, proteins do everything in the living cells. All functions of the living organisms are related with proteins. Each protein or group of proteins are responsible for their own specific function.

CLASSIFICATION OF PROTEINS BY COMPOSITION

- **Simple Proteins**
  - **Albumins:** blood (serum albumin); milk (lactalbumin); egg white (ovalbumin); lentils (legumelin); kidney beans (phaseolin); wheat (leucosin).
  - Globular protein.
  - Soluble in water and dilute salt solution; precipitated by saturation with ammonium sulfate solution; coagulated by heat; found in plant and animal tissues.

- **Globulins:** blood (serum globulins); muscle (myosin); potato (tuberin); Brazil nuts (excelsin); lentils (legumin).
- Globular protein; sparingly soluble in water; soluble in neutral solutions; precipitated by dilute ammonium sulfate and coagulated by heat; distributed in both plant and animal tissues.

Conjugated Proteins

- **Nucleoproteins:** cytoplasm of cells (ribonucleoprotein); nucleus of chromosomes (deoxyribonucleoprotein) viruses, and bacteriophages. Contains nucleic acids, nitrogen, and phosphorus. Present in chromosomes and in all living forms as a combination of protein with either RNA or DNA.
- **Mucoprotein:** saliva (mucin); egg white (ovomucoid). Proteins combined with amino sugars, sugar acids, and sulfates.
- **Glycoprotein:** bone (osseomucoid); tendons (tendomucoid); carilage (chondromucoid).

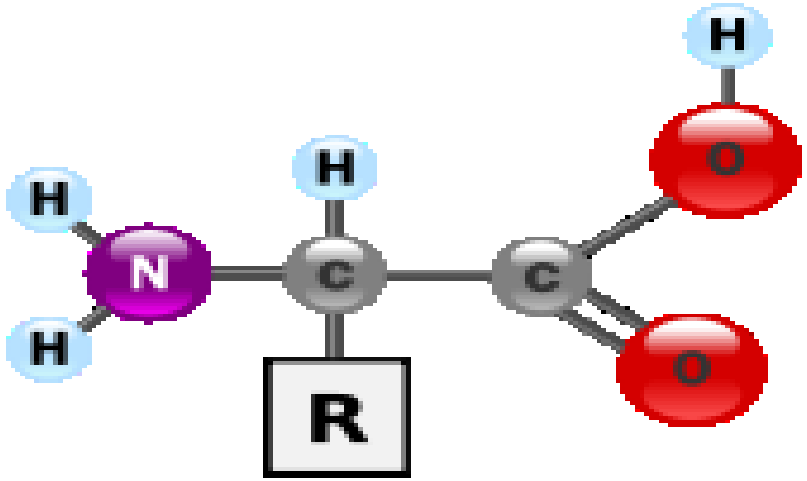
- **Phosphoproteins:** milk (casein); egg yolk (ovovitellin). Phosphoric acid joined in ester linkage to protein.
- **Chromoproteins:** hemoglobin; myoglobin; flavoproteins; respiratory pigments; cytochromes. Protein compounds with some nonprotein pigments as heme; colored proteins.
- **Lipoproteins:** serum lipoprotein; brain, nerve tissues, milk, and eggs. Water-soluble protein conjugated with lipids; found dispersed widely in all cells and all living forms.
- **Metalloproteins:** ferritin; carbonic anhydrase; ceruloplasmin. Proteins combined with metallic atoms that are not parts of a nonprotein prosthetic group.

CLASSIFICATION BASED UPON FUNCTION

- Enzymes (catalytic proteins) *Lactase, ribonuclease, pyruvate dehydrogenase, fumarase, proteinase*
- Structural proteins *Collagen, elastin, keratin*
- Regulatory or hormonal proteins *Insulin, adrenaline*
- Transport proteins *Hemoglobin, myoglobin*
- Genetic proteins *Nucleoproteins, histones*
- Immune Proteins *Gamma Globulin, Ig's*, (Ab's)
- Contractile Proteins *Actin, myosin*
- Storage Proteins *Zein, ovalbumin, casein*

α-Amino Acids

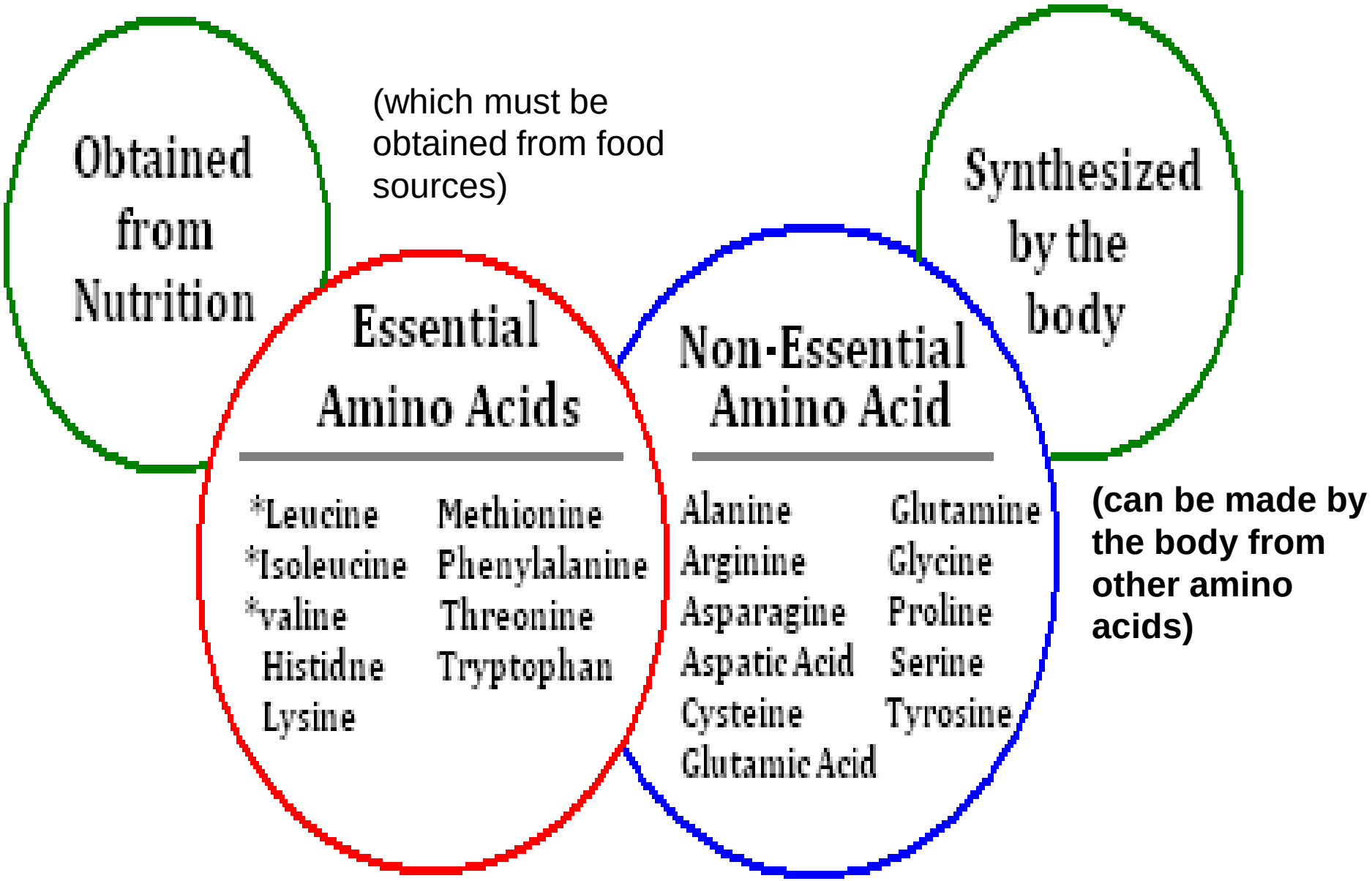
Amino acids are molecules containing an amine group, a carboxylic acid group and a side chain that varies between different amino acids. These molecules contain the key elements of carbon, hydrogen, oxygen, and nitrogen.



In the approximately-20 amino acids found in our bodies, what varies is the side chain

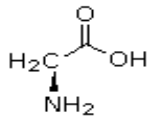
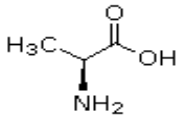
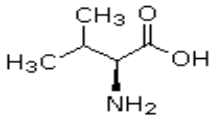
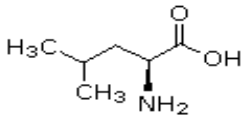
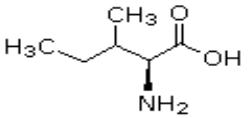
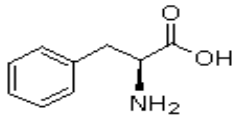
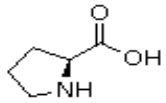
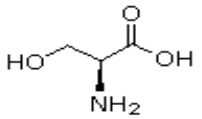
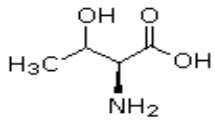
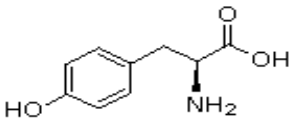
AMINO ACIDS

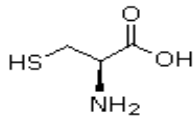
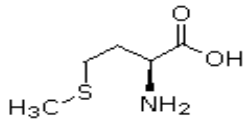
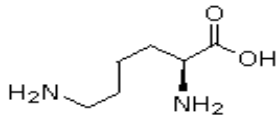
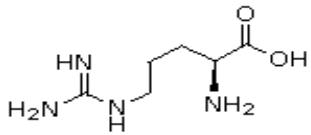
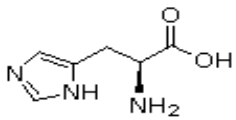
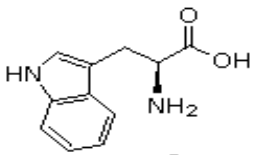
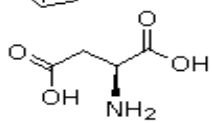
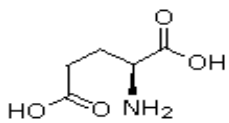
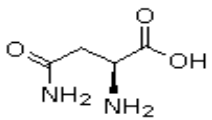
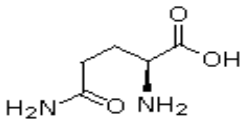
- Amino acids are organic solvents.
- Have two functional groups $-NH_2$ and -COOH group.
- The amino group is basic while carboxylic group is acidic in nature.
- Soluble in water but insoluble in organic solvents e.g chloroform,acetone,ether,etc.
- All amino acids which make up proteins are L-α-aminoacids.
- All amino acids have chiral carbon, exept Glycine.



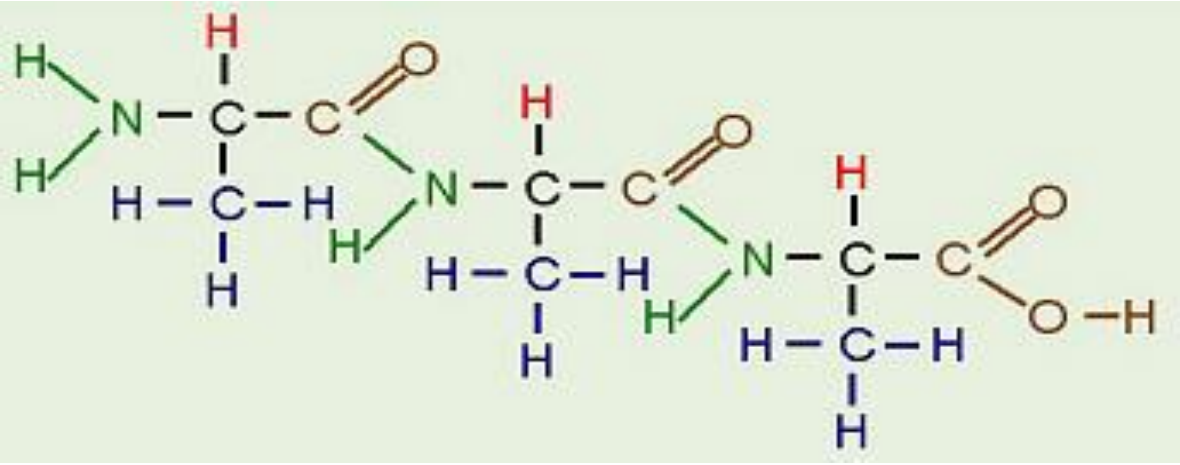
Amino acid in human body

- There are only 22 standard amino acids that exist in living organism. Sometimes these amino acids are chemically modified in the protein after protein synthesis. In total the number of different proteins, which it is possible to produce from 22 amino acids is enormous. For example for 10 amino acid sequence it is possible to have 20^{10} different sequences, which is approximately equal to 10^{13} or 10 trillions of different structures.

Name	Formula	Abbreviations	
Glycine		Gly	G
Alanine		Ala	A
Valine		Val	V
Leucine		Leu	L
Isoleucine		Ile	I
Phenylalanine		Phe	F
Proline		Pro	P
Serine		Ser	S
Threonine		Thr	T
Tyrosine		Tyr	Y

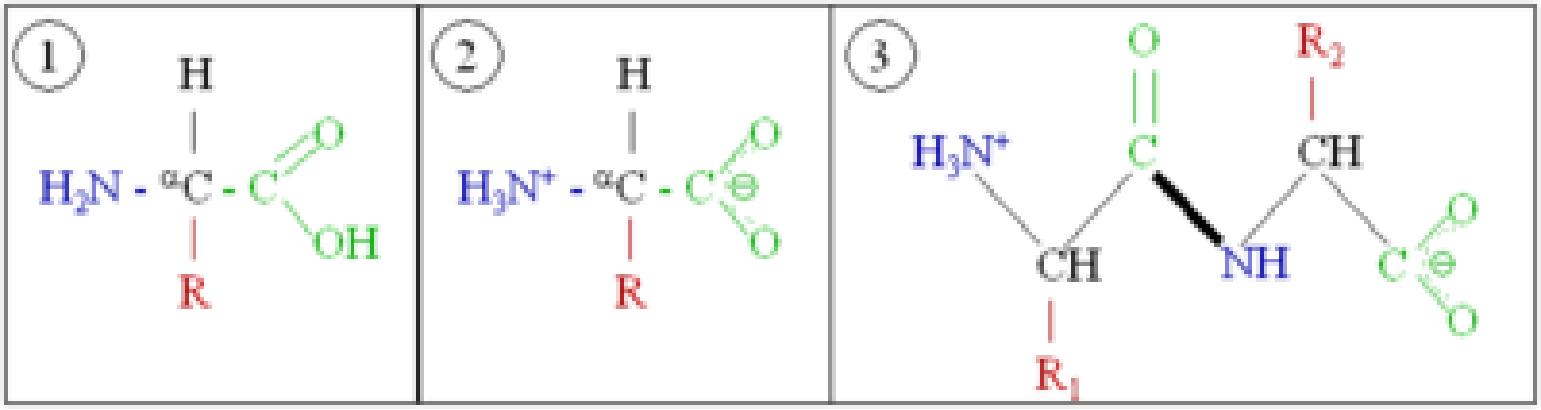
Name	Formula	Abbreviations	
Cysteine		Cys	C
Methionine		Met	M
Lysine		Lys	K
Arginine		Arg	R
Histidine		His	H
Tryptophan		Trp	W
Aspartic Acid		Asp	D
Glutamic Acid		Glu	E
Asparagine		Asn	N
Glutamine		Gln	Q

- To form protein, the amino acids are linked by dehydration synthesis to form peptide bonds. The chain of amino acids is also known as a polypeptide.



The peptide bond

- Proteins are made of amino acids linked into linear chains, called polypeptide chains. Amino acids links between each other by peptide bonds - this peptide bond is formed between the carboxyl and amino groups of neighbouring amino acids. Proteins are formed by one or several polypeptide chains. The sequence of the polypeptide chain is defined by a gene with genetic code.



oligopeptide

- This process can then continue to join other amino acids and yield in an amino acid chain. When there are few amino acids in a chain, it is called an **oligopeptide**,
-

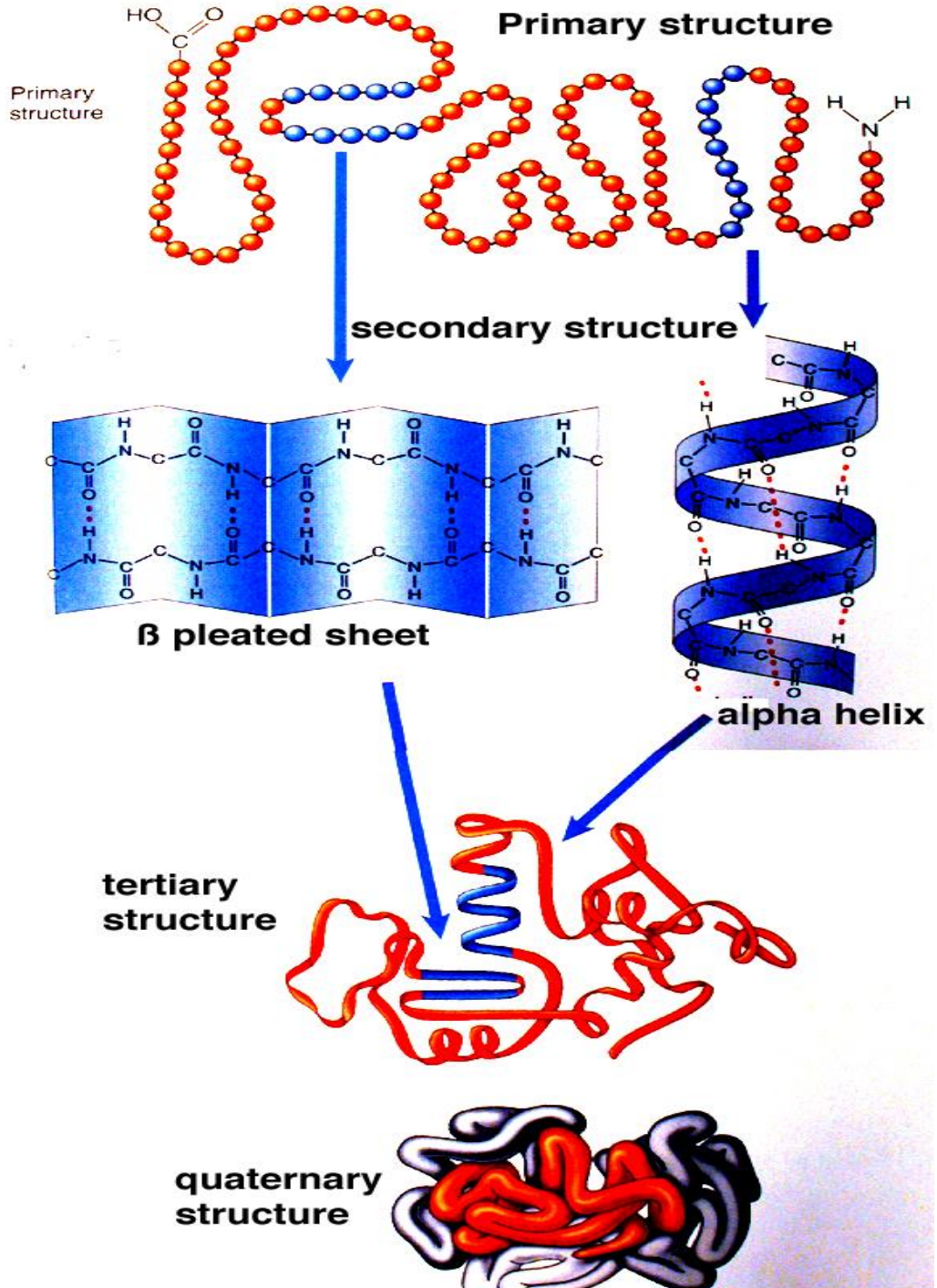
polypeptide

- when there are many it is called a **polypeptide**. Although the terms "protein" and "polypeptide" are sometimes used to describe the same thing, the term polypeptide is generally used when the molecular weight of the chain is below 10,000. An amino acid unit in a peptide is often called a residue.

Proteins structure:

- Proteins are about 50% of the dry weight of most cells, and are the most structurally complex macromolecules known. Each type of protein has its own unique structure and function.
- Polymers are any kind of large molecules made of repeating identical or similar subunits called monomers

- Proteins are composed of a linear (not branched and not forming rings) polymer of amino acids. The twenty genetically encoded amino acids are molecules that share a central core: The α-carbon is bonded to a primary amino (-NH2) terminus, a carboxylic acid (-COOH) terminus, a hydrogen atom, and the amino acid side chain, also called the "R-group".


Primary structure
Primary structure
secondary structure
ß pleated sheet
alpha helix
tertiary structure
quaternary structure

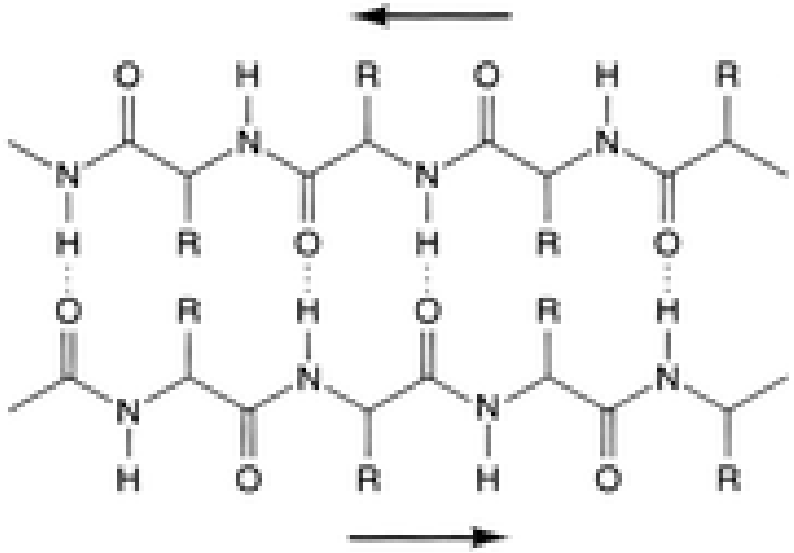
- The R-group determines the identity of the amino acid. In an aqueous solution, at physiological pH (~6.8), the amino group will be in the protonated -NH3+ form, and the carboxylic acid will be in the deprotonated -COO- form, forming a zwitterion. Most amino acids that make up proteins are L-isomers, although a few exotic creatures use D-isomers in their proteins

Protein structural levels

- Biochemists refer to four distinct aspects of a protein's structure:
- **Primary structure**
- Primary structure is practically a synonym of the amino acid sequence. It can also contain information’s on amino acids linked by peptide bonds. Primary structure is typically written as a string of three letter sequences, each representing an amino acid. Peptides and proteins must have the correct sequence of amino acids.

Secondary structure

- *Secondary structure* elements are elementary structural patterns that are present in most,if not all,known proteins. These are highly patterned sub-structures -- alpha helix and beta sheet-- consisting of loops between elements or segments of polypeptide chain that assume no stable shape..



Secondary structure

- Secondary structure elements, when mapped on the sequence and depicted in the relative position they have in respect to each other, define the*topology* of the protein. It is also relevant to note that hydrogen bonding between residues is the cause for secondary structure features; secondary structure is usually described to beginning biochemists as (almost) entirely independent of residue side-chain interactions.

Tertiary structure

- *Tertiary structure* is the name given to refer to the overall shape of a single protein molecule. Although tertiary structure is sometimes described (especially to beginning biology and biochemistry students) as being a result of interactions between amino acid residue side chains, a more correct understanding of tertiary structure is the interactions between elements of secondary protein structure, i.e. alpha-helices and beta-pleated sheets. Tertiary structure is often referred to as the "fold structure" of a protein, since it is the result of the complex three-dimensional interplay of other structural and environmental elements.
-

- **Quaternary structure**
- Quaternary structure is the shape or structure that results from the union of more than one protein molecule, usually called *subunit proteins subunits* in this context, which function as part of the larger assembly or protein complex.
- And it refers to the regular association of two or more polypeptide chains to form a complex. A multi-subunit protein may be composed of two or more identical polypeptides, or it may include different polypeptides. Quaternary structure tends to be stabilized mainly by weak interactions between residues exposed on surfaces polypeptides within a complex.

Protein Properties

Proteins have optimal conditions at which they function.

When exposed to extreme conditions, proteins begin to unfold – denature.

Changes in pH, temperature and salt concentration can denature proteins.

If denaturation occurs moderately over time, returning to the original conditions may result in renaturation.

Denaturation:

- When a protein loses its native conformation. For example, egg white contains a protein called albumin which is water-soluble. However, if heated, albumin becomes denatured and loses its ability to be water-soluble.