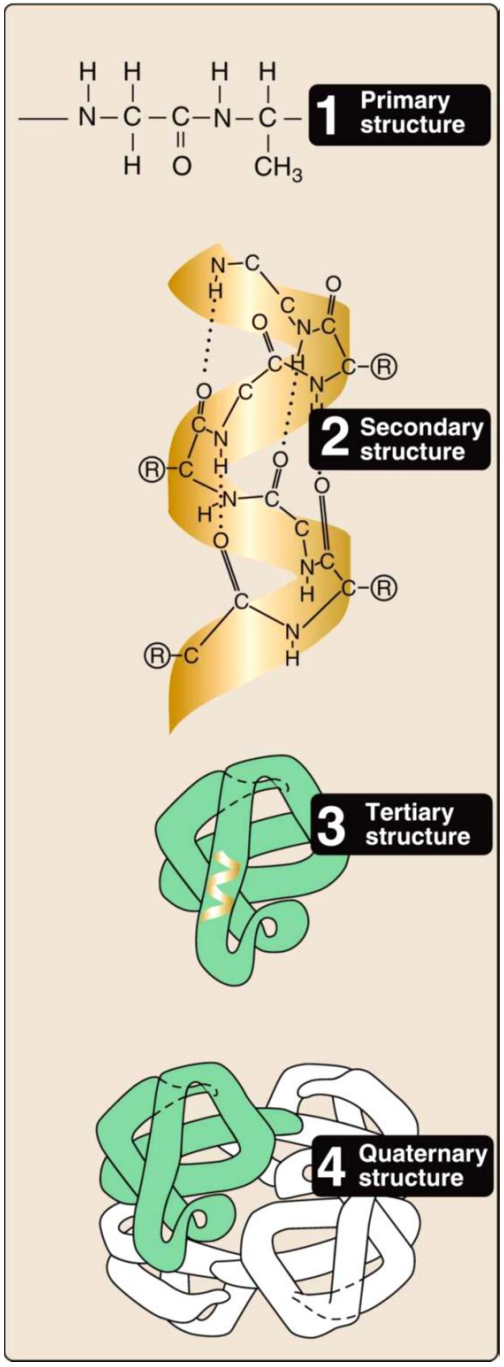
Structure of a Protein

By

Assist.Prof.Dr. Fatimah J Al-Hasani

Structure of a Protein

The 20 amino acids commonly found in proteins are joined together by peptide bonds. The linear sequence of the linked amino acids contains the information necessary to generate a protein molecule with a unique three-dimensional shape that determines function. The complexity of protein structure is best analyzed by considering the molecule in terms of four organizational levels: primary, secondary, tertiary, and quaternary. An examination of these hierarchies of increasing complexity has revealed that certain structural elements are repeated in a wide variety of proteins, suggesting that there are general rules regarding the ways in which proteins achieve their native, functional form. These repeated structural elements range from simple combinations of α-helices and β-sheets forming small motifs to the complex folding of polypeptide domains of multifunctional proteins


1 Primary structure
2 Secondary structure
3 Tertiary structure
4 Quaternary structure

The Primary Structure of a Protein

The sequence of amino acids in a protein is called the primary structure of the protein. Understanding the primary structure of proteins is important because many genetic diseases result in proteins with abnormal amino acid sequences, which cause improper folding and loss or impairment of normal function. If the primary structures of the normal and the mutated proteins are known, this information may be used to diagnose or study the disease.

Peptide bond

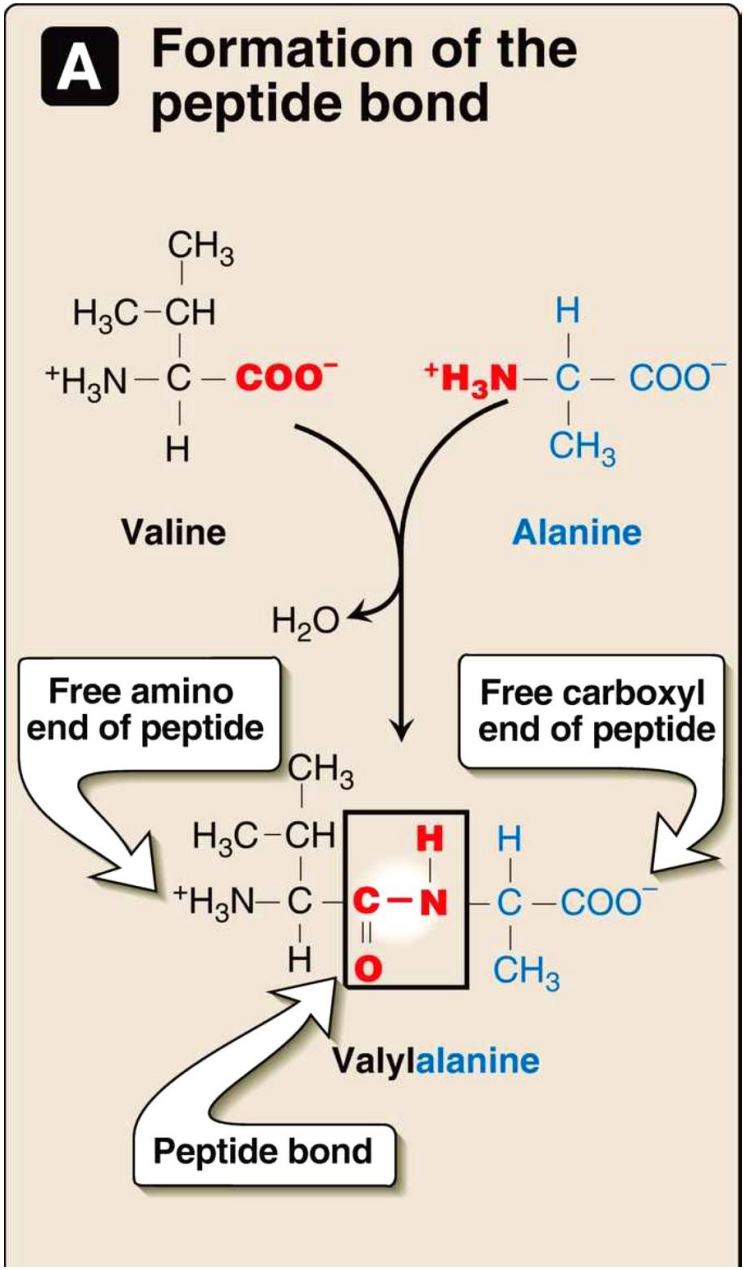
In proteins, amino acids are joined covalently by peptide bonds, which are amide linkages between the α-carboxyl group of one amino acid and the α-amino group of another. For example, valine and alanine can form the

dipeptide valylalanine through the formation of a peptide bond Peptide bonds are resistant to conditions that denature proteins, such as heating and high concentrations of urea

What is a peptide bond in simple terms?

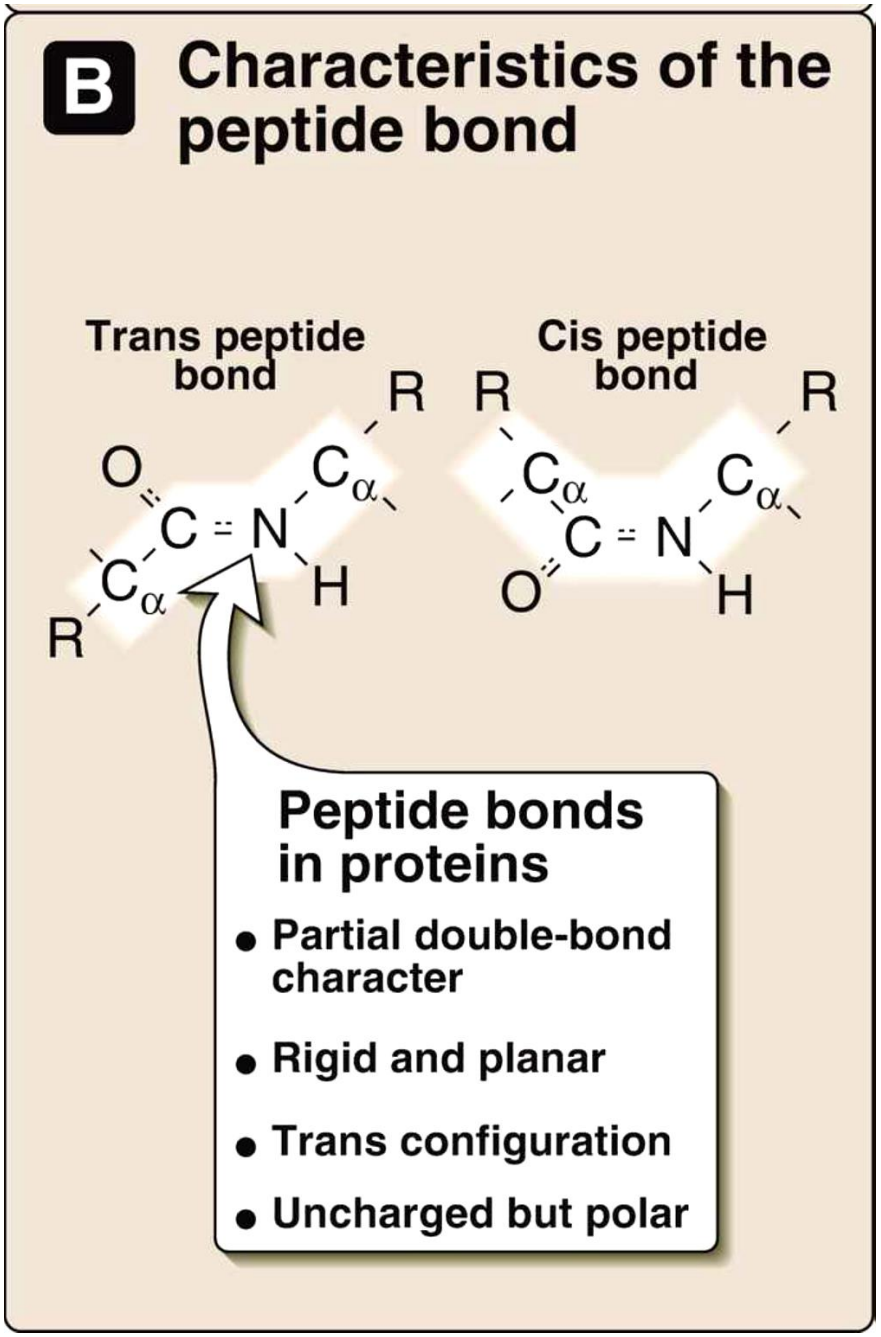
A peptide bond is a chemical bond formed between two molecules when the carboxyl group of one molecule reacts with the amino group of the other molecule, releasing a molecule of water (H2O).

1-Naming the peptide: By convention, the free amino end (N-terminal) of the peptide chain is written to the left and the free carboxyl end (C-terminal) to the right. Therefore, all amino acid sequences are read from the N- to the C-terminal end. For example, in Figure, the order of the amino acids in the dipeptide is valine, alanine. Linkage of ≥50 amino acids through peptide bonds results in an unbranched chain called a polypeptide, or protein. Each component amino acid is called a residue because it is the portion of the amino acid remaining after the atoms of water are lost in the formation of the peptide bond. When a peptide is named, all amino acid residues have their suffixes (-ine, -an, -ic, or -ate) changed to -yl, with the exception of the C-terminal amino acid. For example, a tripeptide composed of an N-terminal valine, a glycine, and a C-terminal leucine is called **valylglycylleucine.**



2. Peptide bond characteristics:

The peptide bond has a partial double-bond character, that is, it is shorter than a single bond and is rigid and planar .This prevents free rotation around the bond between the carbonyl carbon and the nitrogen of the peptide bond. However, the bonds between the α-carbons and the α-amino or α-carboxyl groups can be freely rotated (although they are limited by the size and character of the R groups). This allows the polypeptide chain to assume a variety of possible conformations. The peptide bond is almost always in the trans configuration,in large part because of steric interference of the R groups (side chains) when in the cis position.



3. Peptide bond polarity:

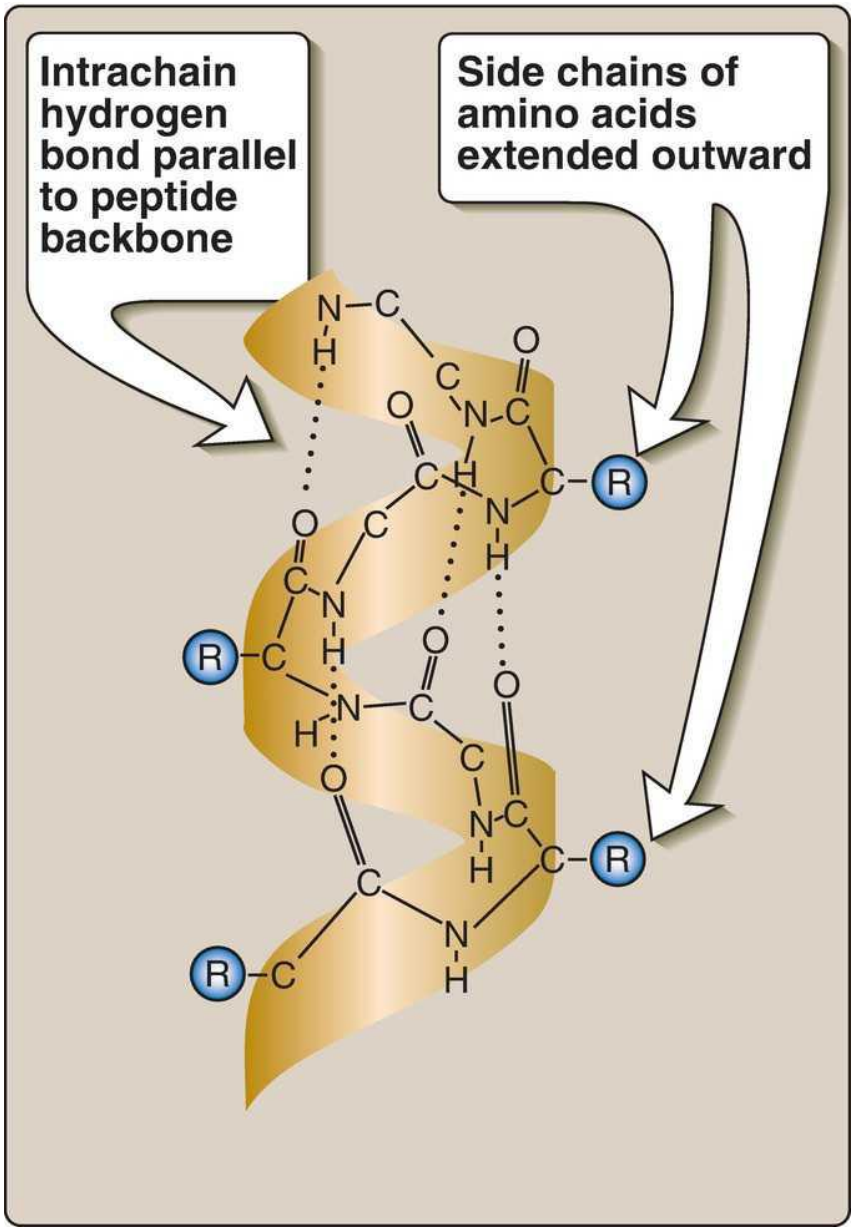
Like all amide linkages, the −C = O and −NH groups of the peptide bond are uncharged, and neither accept nor release protons over the pH range of 2–12. Thus, the charged groups present in polypeptides consist solely of the N-terminal (α-amino) group, the C terminal (α-carboxyl) group, and any ionized groups present in the side chains of the constituent amino acids. The −C = O and −NH groups of the peptide bond are polar, however, and are involved in hydrogen bonds (for example, in α-helices and β-sheets),

SECONDARY STRUCTURE

The polypeptide backbone does not assume a random three-dimensional structure but, instead, generally forms regular arrangements of amino acids that are located near each other in the linear sequence. These arrangements are termed the secondary structure of the polypeptide. The α-helix, β-sheet, and β- bend (or, β-turn) are examples of secondary structures commonly encountered in proteins. Each is stabilized by hydrogen bonds between atoms of the peptide backbone.

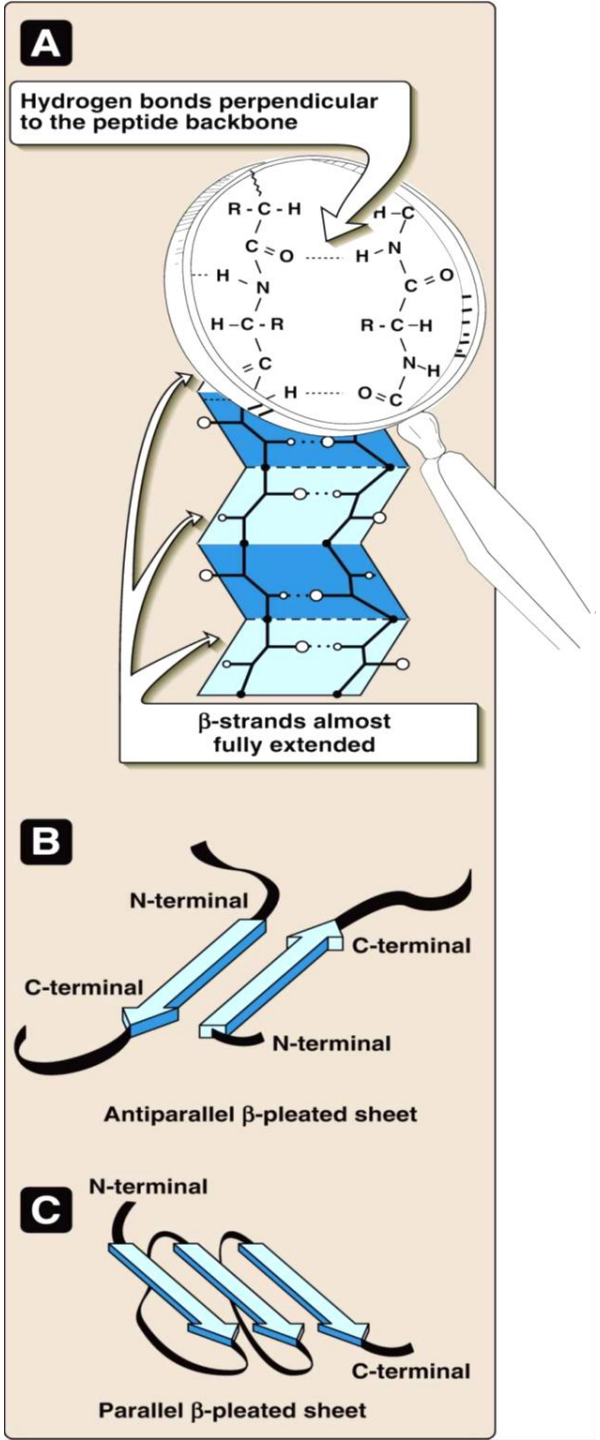
α-Helix

Several different polypeptide helices are found in nature, but the α-helix is the most common. It is a rigid, right-handed spiral structure, consisting of a tightly packed, coiled polypeptide backbone core, with the side chains of the component L-amino acids extending outward from the central axis to avoid interfering sterically with each other. A very diverse group of proteins contains α-helices. For example, the keratins are a family of closely related, rigid, fibrous proteins whose structure is nearly entirely α- helical. They are a major component of tissues such as hair and skin. In contrast to keratin, myoglobin, whose structure is also highly α-helical, is a globular, flexible molecule found in muscles.



β-Sheet

The β-sheet is another form of secondary structure in which all of the peptide bond components are involved in hydrogen bonding (Fig. A). Because the surfaces of β-sheets appear "pleated," they are often called β-pleated sheets. [Note: Pleating results from successive α-carbons being slightly above or below the plane of the sheet.] Illustrations of protein structure often show β-strands as broad arrows (FigB).



β-Bends (reverse turns, β-turns)

β-Bends reverse the direction of a polypeptide chain, helping it form a compact, globular shape. They are usually found on the surface of protein molecules and often include charged residues. [Note: β-Bends were given this name because they often connect successive strands of antiparallel β- sheets.] β-Bends are generally composed of four amino acids, one of which

may be proline, the amino acid that causes a kink in the polypeptide chain. Glycine, the amino acid with the smallest R group, is also frequently found in β-bends. β-Bends are stabilized by the formation of hydrogen bonds between the first and last residues in the bend.

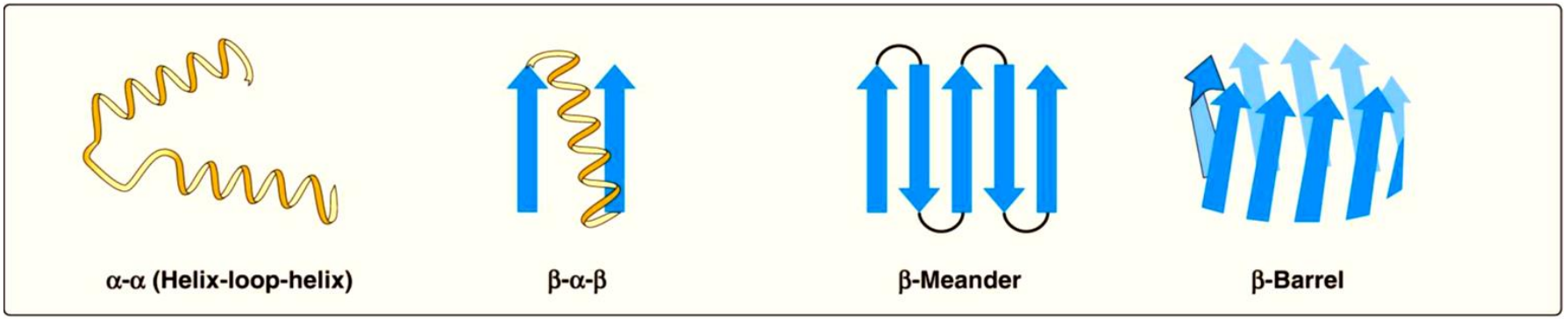
D. Nonrepetitive secondary structure

Approximately one half of an average globular protein is organized into repetitive structures, such as the α-helix and β-sheet. The remainder of the polypeptide chain is described as having a loop or coil conformation. These nonrepetitive secondary structures are not random but rather simply have a

less regular structure than those described above. [Note: The term “random coil” refers to the disordered structure obtained when proteins are denatured

E. Supersecondary structures (motifs)

Globular proteins are constructed by combining secondary structural elements (that is, α-helices, β-sheets, and coils), producing specific geometric patterns, or motifs. These form primarily the core (interior) region of the molecule. They are connected by loop regions (for example, β-bends) at the surface of the protein. Supersecondary structures are usually produced by the close packing of side chains from adjacent secondary structural elements. For example, α-helices and β-sheets that are adjacent in the amino acid sequence are also usually (but not always) adjacent in the final, folded protein.



TERTIARY STRUCTURE

The primary structure of a polypeptide chain determines its tertiary structure.“Tertiary” refers both to the folding of domains (the basic units of structure and function; see A. below) and to the final arrangement of domains in the polypeptide. The tertiary structure of globular proteins in aqueous solution is compact, with a high density (close packing) of the atoms in the core of the molecule. Hydrophobic side chains are buried in the interior, whereashydrophilic groups are generally found on the surface of the molecule.

A. Domains

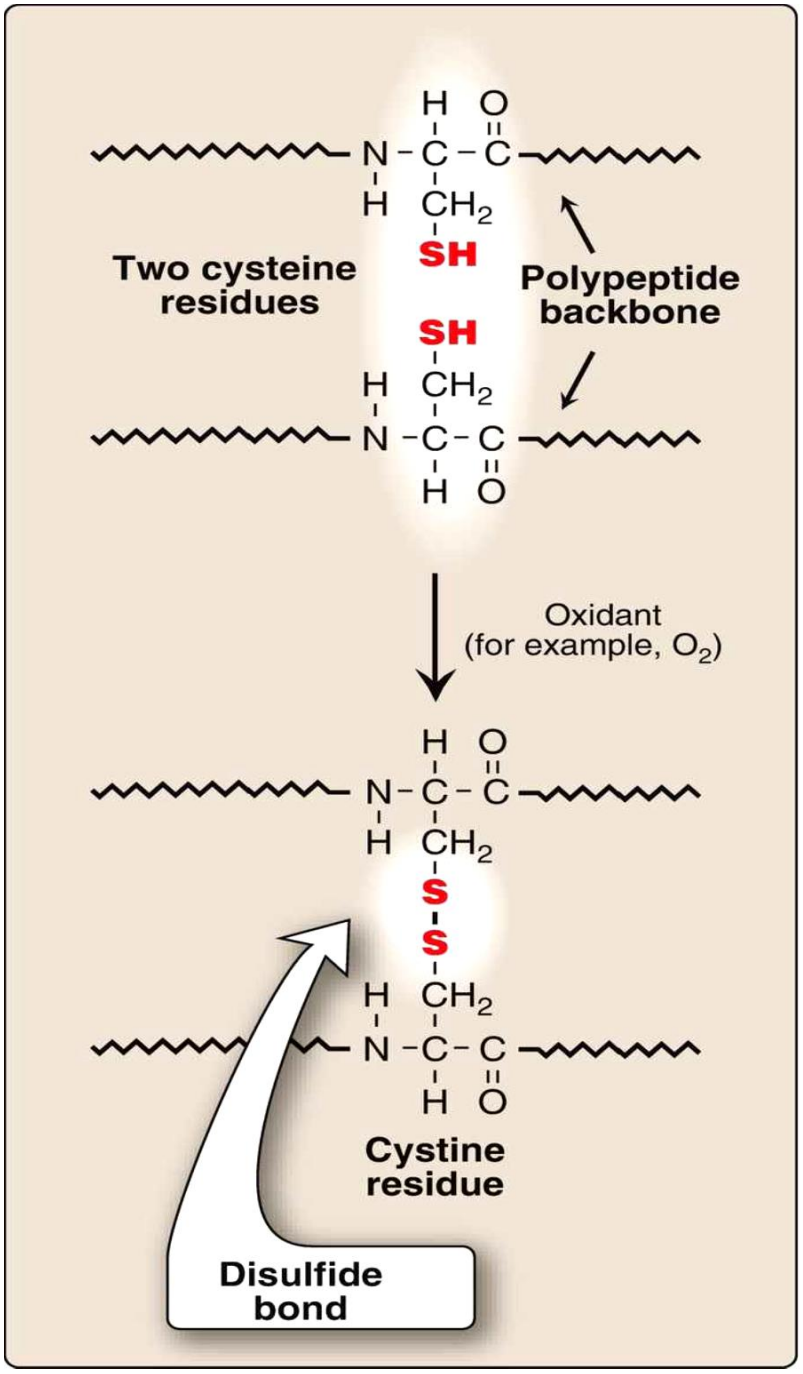
Domains are the fundamental functional and three-dimensional structural units of polypeptides. Polypeptide chains that are >200 amino acids in length generally consist of two or more domains. The core of a domain is built from combinations of supersecondary structural elements (motifs). Folding of the peptide chain within a domain usually occurs independently of folding in other domains. Therefore, each domain has the characteristics of a small, compact globular protein that is structurally independent of the other domains in the polypeptide chain.

B. Stabilizing interactions

The unique three-dimensional structure of each polypeptide is determined by its amino acid sequence. Interactions between the amino acid side chains guide the folding of the polypeptide to form a compact structure. The following four types of interactions cooperate in stabilizing the tertiary structures of globular proteins.

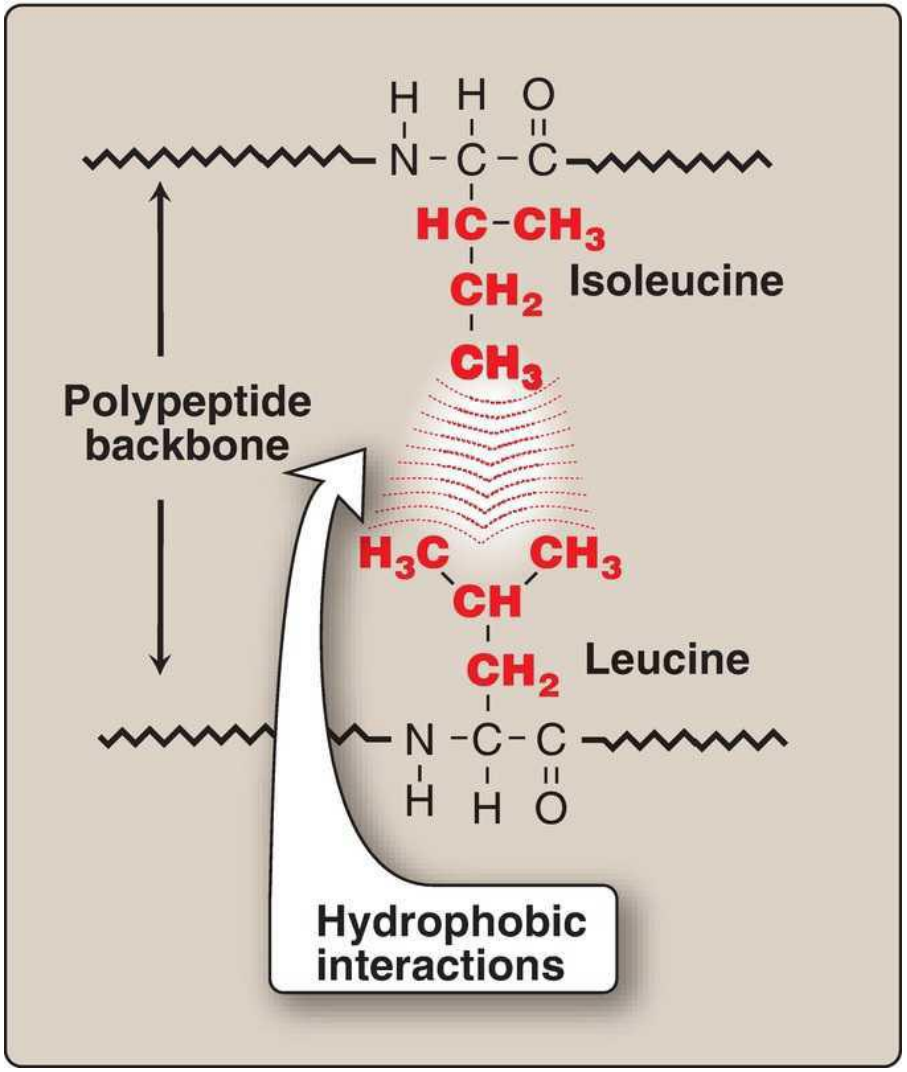
1. Disulfide bonds:

A disulfide bond (–S–S–) is a covalent linkage formed from the sulfhydryl group (−SH) of each of two cysteine residues to produce a cystine residue. The two cysteines may be separated from each other by many amino acids in the primary sequence of a polypeptide or may even be located on two different polypeptides. The folding of the polypeptide(s) brings the cysteine residues into proximity and permits covalent bonding of their side chains. A disulfide bond contributes to the stability of the three-dimensional shape of the protein molecule and prevents it from becoming denatured in the extracellular environment. For example, many disulfide bonds are found in proteins such as immunoglobulins that are secreted by cells. [Note: *Protein disulfide isomerase* breaks and reforms disulfide bonds during folding.]



2. Hydrophobic interactions:

Amino acids with nonpolar side chains tend to be located in the interior of the polypeptide molecule, where they associate with other hydrophobic amino acids. In contrast, amino acids with polar or charged side chains tend to be located on the surface of the molecule in contact with the polar solvent. [Note: Recall that proteins located in nonpolar (lipid) environments, such as a membrane, exhibit the reverse arrangement case, a segregation of R groups occurs that is energetically most favorable.

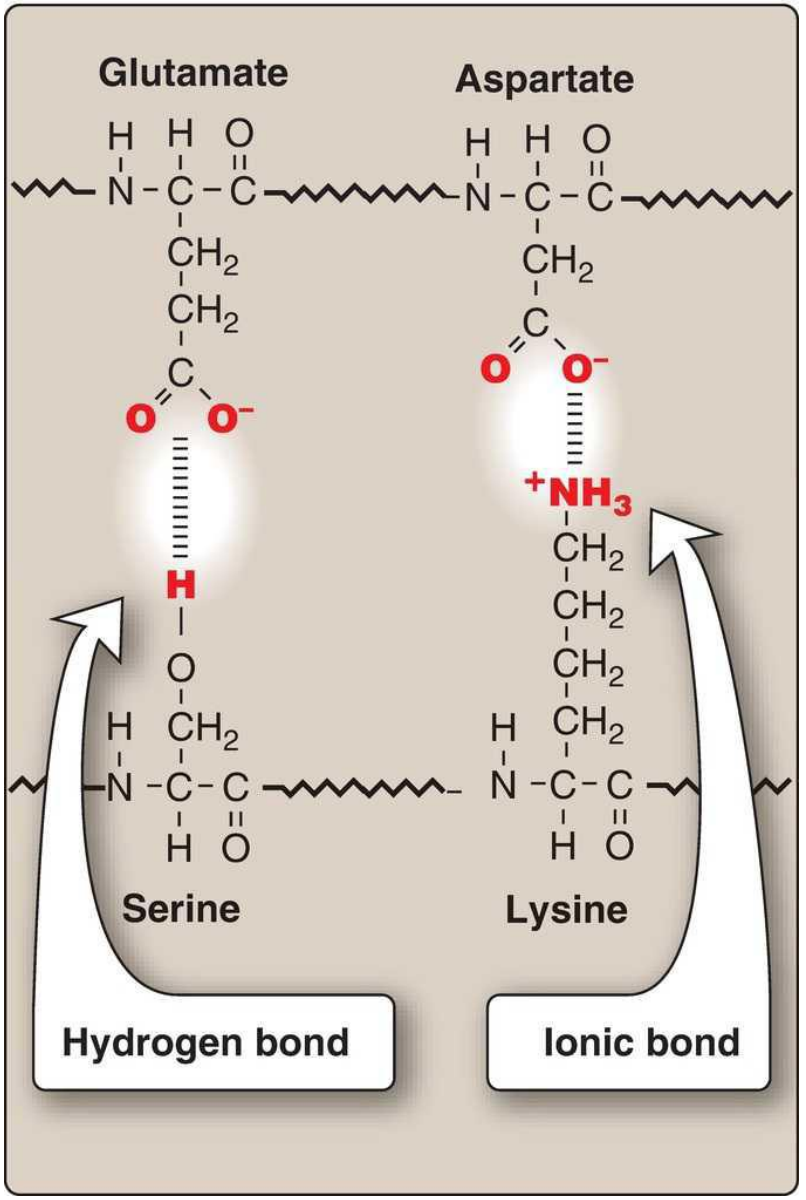


3. Hydrogen bonds:

Amino acid side chains containing oxygen- or nitrogen-bound hydrogen, such as in the alcohol groups of serine and threonine, can form hydrogen bonds with electron-rich atoms, such as the oxygen of a carboxyl group or carbonyl group of a peptide bond . Formation of hydrogen bonds between polar groups on the surface of proteins and the aqueous solvent enhances the solubility of the protein.

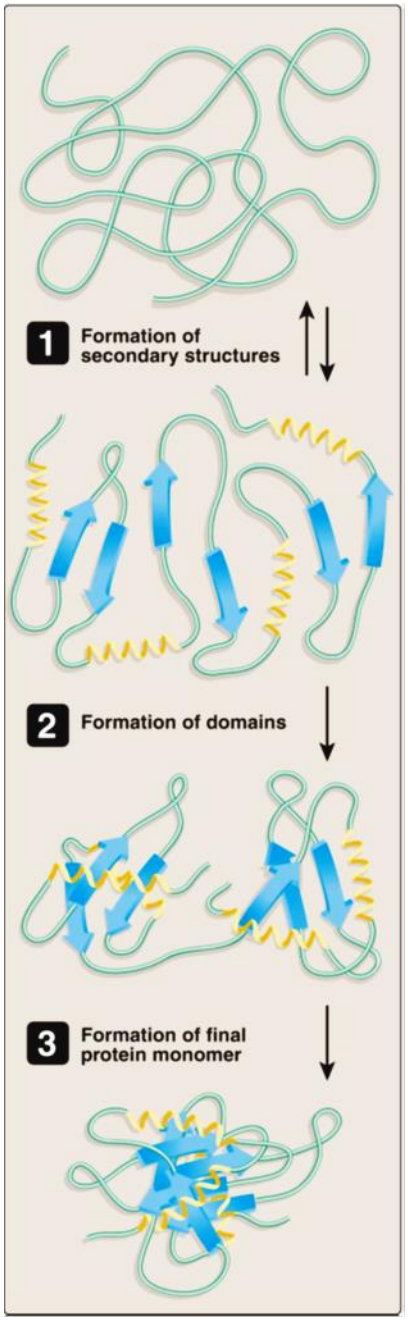
4. Ionic interactions:

Negatively charged groups, such as the carboxylate group (−COO–) in the side chain of aspartate or glutamate, can interact with positively charged groups such as the amino group (−NH_{3+}) in the side chain of lysine



Protein folding

Interactions between the side chains of amino acids determine how a linear polypeptide chain folds into the intricate three-dimensional shape of the functional protein. Protein folding, which occurs within the cell in seconds to minutes, involves nonrandom, ordered pathways. As a peptide folds, secondary structures form, driven by the hydrophobic effect (that is, hydrophobic groups come together as water is released). These small structures combine to form larger structures. Additional events stabilize secondary structure and initiate formation of tertiary structure. In the last stage, the peptide achieves its fully folded, native (functional) form characterized by a low-energy state [Note: Some biologically active proteins or segments thereof lack a stable tertiary structure. They are referred to as intrinsically disordered proteins.]



QUATERNARY STRUCTURE

Many proteins consist of a single polypeptide chain and are defined as monomeric proteins. However, others may consist of two or more polypeptide chains that may be structurally identical or totally unrelated. The arrangement of these polypeptide subunits is called the quaternary structure of the protein.

Subunits are held together primarily by noncovalent interactions (for example, hydrogen bonds, ionic bonds, and hydrophobic interactions). Subunits either may function independently of each other or may work cooperatively, as in hemoglobin, in which the binding of oxygen to one subunit of the tetramer increases the affinity of the other subunits for oxygen