Protein Structure Basics

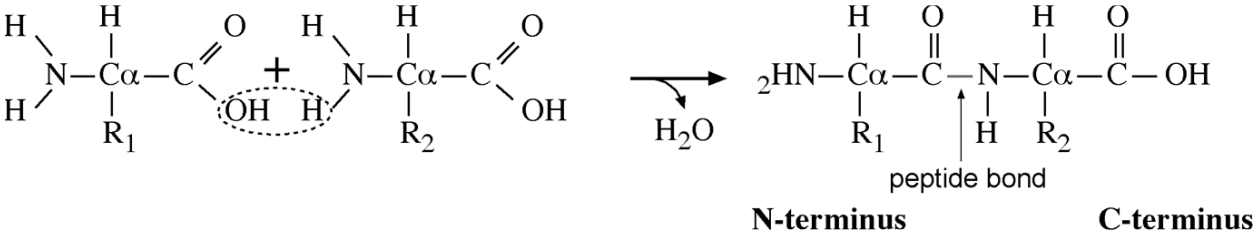
Proteins perform most essential biological and chemical functions in a cell. They play important roles in structural, enzymatic, transport, and regulatory functions. The protein functions are strictly determined by their structures.

AMINO ACIDS

The building blocks of proteins are twenty naturally occurring amino acids, small molecules that contain a free amino group (NH2) and a free carboxyl group (COOH). Both of these groups are linked to a central carbon (Cα), which is attached to a hydrogenand a side chain group (R). Amino acids differ only by the side chain R group. Amino acids can be grouped into several categories based on the chemical and physical properties of the side chains, such as size and affinity for water. According to these properties, the side chain groups can be divided into small, large, hydrophobic, and hydrophilic categories. Within the hydrophobic set of amino acids, they can be further divided into aliphatic and aromatic. *Aliphatic side chains* are linear hydrocarbon chains and *aromatic side chains* are cyclic rings. Within the hydrophilic set, amino acids can be subdivided into polar and charged. *Charged amino acids* can be either positively charged (basic) or negatively charged (acidic). Of particular interest within the twenty amino acids are glycine and proline. Glycine, the smallest amino acid, has a hydrogen atom as the R group. It can therefore adopt more flexible conformations that are not possible for other amino acids. Proline is on the other extreme of flexibility. Its side chain forms a bond with its own backbone amino group, causing it to be cyclic. The cyclic conformation makes it very rigid, unable to occupy many of the main chain conformations adopted by other amino acids

PEPTIDE FORMATION

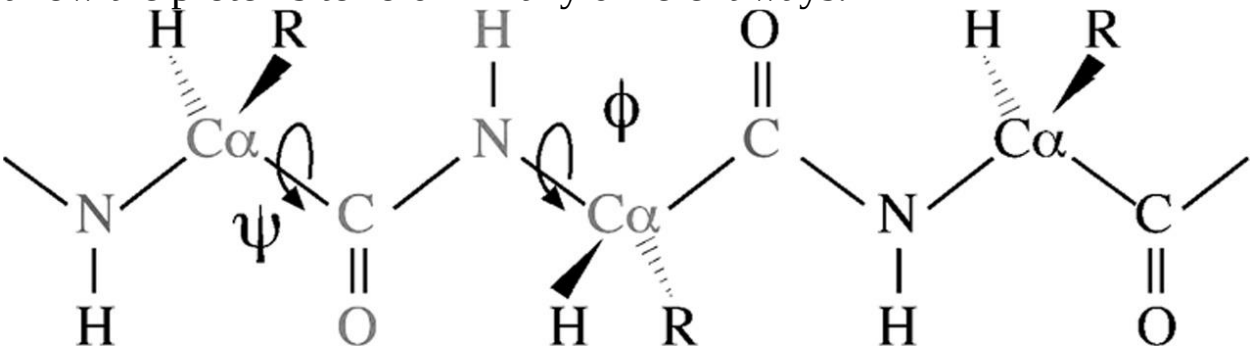
The peptide formation involves two amino acids covalently joined together between the carboxyl group of one amino acid and the amino group of another.



This reaction is a condensation reaction involving reaction is a condensation reaction involving removal of elements of water from the two molecules. The resulting product is called a *dipeptide.* The newly formed covalent bond connecting the two amino acids is called a *peptide bond*. Once an amino acid is incorporated into a peptide, it becomes an amino acid residue. Multiple amino acids can be joined together to forma longer chain of amino acid polymer. A linear polymer of more than fifty amino acid residues is referred to as a *polypeptide*.A polypeptide, also called a protein, has a well-defined three-dimensional arrangement. On the other hand, a polymer with fewer than fifty residues is usually called a peptide without a well-defined three-dimensional structure. The residues ina peptide or polypeptide are numbered beginning with the residue containing the amino group, referred to as the *N*-terminus, and ending with the residue containing the carboxyl group, known as the *C*-terminus. The actual sequence of amino acid residues in a polypeptide determines its ultimate structure and function.The atoms involved in forming the peptide bond are referred to as the *backbone atoms*. They are the nitrogen of the amino group, the *α*carbon to which the side chainis attached and carbon of the carbonyl group.

DIHEDRAL ANGLES

A peptide bond is actually a partial double bond owing to shared electrons between O=C–N atoms. The rigid double bond structure forces atoms associated with the peptide bond to lie in the same plane, called the *peptide plane*. Because of the planar nature of the peptide bond and the size of the R groups, there are considerable restrictions on the rotational freedom by the two bonded pairs of atoms around the peptide bond. The angle of rotation about the bond is referred to as the *dihedral angle* (also called the *tortional angle*). For a peptide unit, the atoms linked to the peptide bond can be moved to a certain extent by the rotation of two bonds flanking the peptide bond. This is measured by two dihedral angles. One is the dihedral angle along the N–Cα bond, which is defined as phi (φ); and the other is the angle along the Cα–C bond, which is called psi (ψ). Various combinations of φ and ψ angles allow the proteins to fold in many different ways.



SECONDARY STRUCTURES

As mentioned, local structures of a protein with regular conformations are known as secondary structures. They are stabilized by hydrogen bonds formed between carbonyl oxygen and amino hydrogen of different amino acids. Chief elements of secondary structures are α-helices and β-sheets.

α-Helices

Anα-helix has a main chain backbone conformation that resembles a corkscrew. Nearly all known α-helices are right handed, exhibiting a rightward spiral form. In such a helix, there are 3.6 amino acids per helical turn. The structure is stabilized by hydrogen bonds formed between the main chain atoms of residues I and $I + 4$. The hydrogen bonds are nearly parallel with the helical axis. The average φ and ψ angles are $60^\circ$and 45°, respectively, and are distributed in a narrowly defined region in the lower left region of a Ramachandran plot. Hydrophobic residues of the helix tend to face inside and hydrophilic residues of the helix face outside. Thus, every third residue along the helix tends to be a hydrophobic residue. Ala, Gln, Leu, And Met are commonly found in anα-helix, but not Pro, Gly, and Tyr. These rules are useful in guiding the prediction of protein secondary structures.

β-Sheets

A β-sheet is a fully extended configuration built up from several spatially adjacent regions of a polypeptide chain. Each region involved in forming the β-sheet is a β-strand. The β-strand conformation is pleated with main chain backbone zigzagging and side chains positioned alternately on opposite sides of the sheet.β-Strandsare stabilized by hydrogen bonds between residues of adjacent strands. β-strandsnear the surface of the protein tend to show an alternating pattern of hydrophobic and hydrophilic regions, whereas strands buried at the core of a protein are nearly all hydrophobic. The β-strands can run in the same direction to form a parallel sheet or can run every other chain in reverse orientation to form an antiparallel sheet, or a mixture of both. The hydrogen bonding patterns are different in each configuration. The φ and ψ angles are also widely distributed in the upper left region in a Ramachandran plot. Because of the long-range nature of residues involved in this type of conformation, it is more difficult to predict β-sheets than α-helices.