Amino acids & Proteins – an introduction

Proteins are the most abundant biological macromolecules, occurring in all cells and all parts of cells. Proteins also occur in great variety; thousands of different kinds, ranging in size from relatively small peptides to huge polymers with molecular weights in the millions, may be found in a single cell. Moreover, proteins exhibit enormous diversity of biological function and are the most important final products of the information pathways. Proteins are the molecular instruments through which genetic informationis expressed.

Relatively simple monomeric subunits provide the key to the structure of the thousands of different proteins. All proteins, whether from the most ancient lines of bacteria or from the most complex forms of life, are constructed from the same ubiquitous set of 20 amino acids, covalently linked in characteristic linear sequences. Because each of these amino acids has a side chain with distinctive chemical properties, this group of 20 precursor molecules may be regarded as the alphabet in which the language of protein structure is written. What is most remarkable is that cells can produce proteins with strikingly different properties and activities by joining the same 20 amino acids in many different combinations and sequences. From these building blocks different organisms can make such widely diverse products as enzymes, hormones, antibodies, transporters, muscle fibers, the lens protein of the eye, feathers, spider webs, rhinoceros horn, milk proteins, antibiotics, mushroom poisons, and myriad other substances having distinct biological activities . Among these protein products, the enzymes are the most varied and specialized. Virtually all cellular reactions are catalyzed by enzymes. Proteins are polymers of amino acids, with each **amino acid residue** joined to its neighbor by a specific type of covalent bond. (The term “residue” reflects the loss of the elements of water when one amino acid is joined to another.) Proteins can be broken down (hydrolyzed) to their constituent amino acids by a variety of methods, and the earliest studies of proteins naturally focused on the free amino acids derived from them. Twenty different amino acids are commonly found in proteins. The first to be discovered was asparagine, in 1806. The last of the 20 to be found, threonine, was not identified until 1938. All the amino acids have trivial or common names, in some cases derived from the source from which they were first isolated. Asparagine was first found in

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

asparagus, and glutamate in wheat gluten; tyrosine was first isolated from cheese (its name is derived from the Greek *tyros*, "cheese"); and glycine (Greek *glykos*, "sweet") was so named because of its sweet taste.

Common features of amino acids

All 20 of the common amino acids are α-amino acids. They have a carboxyl group and an amino group bonded to the same carbon atom (the α carbon) . They differ from each other in their side chains, or **R groups,** which vary in structure, size, and electric charge, and which influence the solubility of the amino acids in water. In addition to these 20 amino acids there are many less common ones. Some are residues modified after a protein has been synthesized; others are amino acids present in living organisms but not as constituents of proteins.

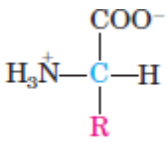
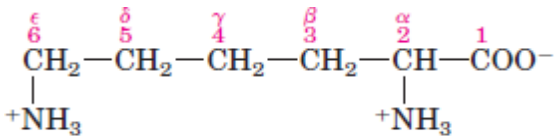


FIGURE 3–2 General structure of an amino acid. This structure is common to all but one of the α-amino acids. (Proline, a cyclic amino acid, is the exception.) The R group or side chain (red) attached to the α carbon (blue) is different in each amino acid.

The common amino acids of proteins have been assigned three-letter abbreviations and one-letter symbol, which are used as shorthand to indicate the composition and sequence of amino acids polymerized in proteins. Two conventions are used to identify the carbons in an amino acid—a practice that can be confusing. The additional carbons in an R group are commonly designated β, γ, δ, and so forth, proceeding out from the α carbon. For most other organic molecules, carbon atoms are simply numbered from one end, giving highest priority (C-1) to the carbon with the substituent containing the atom of highest atomic number. Within this latter convention, the carboxyl carbon of an amino acid would be C-1 and the _ carbon would be C-2. In some cases, such as amino acids with heterocyclic R groups, the Greek lettering system is ambiguous and the numbering convention is therefore used.



Lysine

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Stereoisomers

For all the common amino acids except glycine, the α carbon is bonded to four different groups: a carboxyl group, an amino group, an R group, and a hydrogen atom (Fig. 3–2; in glycine, the R group is another hydrogen atom). The α carbon atom is thus a **chiral center** (p. 17). Because of the tetrahedral arrangement of the bonding orbitals around the α carbon atom, the four different groups can occupy two unique spatial arrangements, and thus amino acids have two possible stereoisomers. Since they are nonsuperimposable mirror images of each other (Fig. 3–3), the two forms represent a class of stereoisomers called **enantiomers** (see Fig. 1–19). All molecules with a chiral center are also **optically active**—that is, they rotate plane-polarized light,

The absolute configurations of simple sugars and amino acids are specified by the **D, L system** (Fig. 3–4), based on the absolute configuration of the three-carbon sugar glyceraldehyde, a convention proposed by Emil Fischer in 1891. (Fischer knew what groups surrounded the asymmetric carbon of glyceraldehyde but had to guess at their absolute configuration; his guess was later confirmed by x-ray diffraction analysis.) For all chiral compounds, stereoisomers having a configuration related to that of L-glyceraldehyde are designated L, and stereoisomers related to D-glyceraldehyde are designated D. L-Amino acids are those with the α amino group on the left, and D-amino acids have the α amino group on the right. Historically, the similar *l* and *d* designations were used for levorotatory (rotating light to the left) and dextrorotatory (rotating light to the right). However, not all L-amino acids are levorotatory, and not all all D-aminoacids are dextro rotatory.. By Fischer's convention, L and D refer *only* to the absolute configuration of the four substituents around the chiral carbon, not to optical properties of the molecule.

Nearly all biological compounds with a chiral center occur naturally in only one stereoisomeric form, either D or L. The amino acid residues in protein molecules are exclusively L stereoisomers. D-Amino acid residues have been found only in a few, generally small peptides, including some peptides of bacterial cell walls and certain peptide antibiotics.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

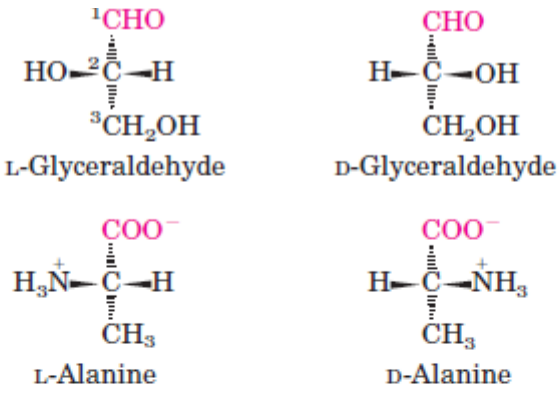


FIGURE 3–4 Steric relationship of the stereoisomers of alanine to the absolute configuration of L- and D-glyceraldehyde. In these perspective formulas, the carbons are lined up vertically, with the chiral atom in the center. The carbons in these molecules are numbered beginning with the terminal aldehyde or carboxyl carbon (red), 1 to 3 from top to bottom as shown. When presented in this way, the R group of the amino acid (in this case the methyl group of alanine) is always below the α carbon. L-Amino acids are those with the α-amino group on the left, and D-amino acids have the α-amino group on the right.

Classification of amino acids

I. Classification on the basis of R-group

II. Classification on the basis of nutrition

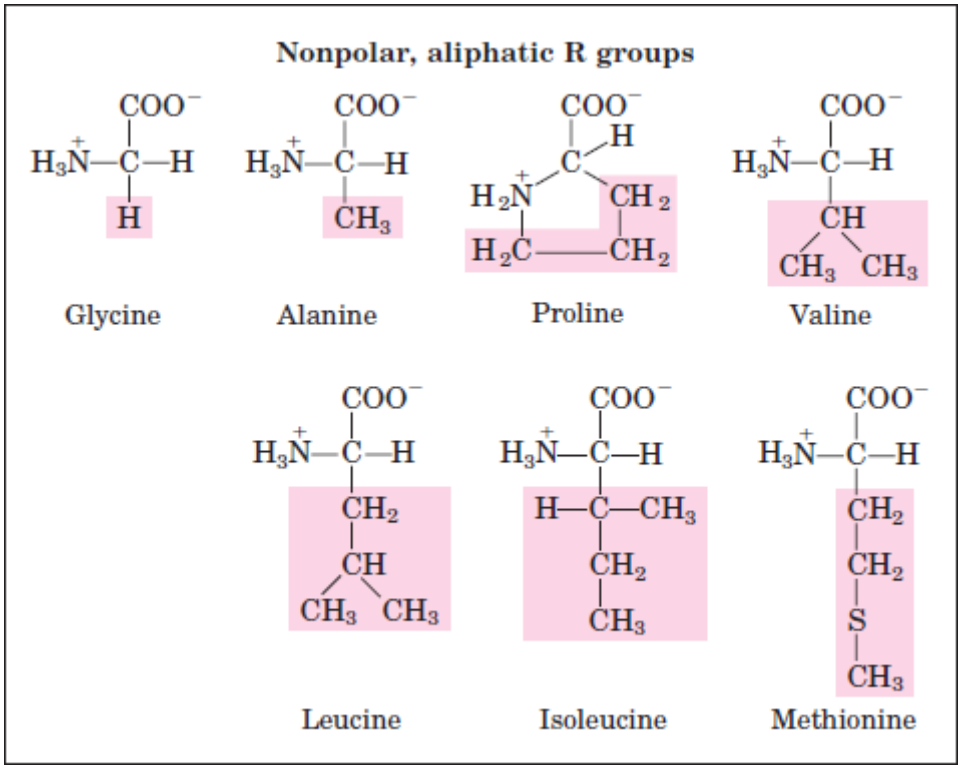
III. Classification on the basis of Catabolism

I) Classification of amino acids on the basis of R-group

Nonpolar, Aliphatic R Groups The R groups in this class of amino acids are nonpolar and hydrophobic. The side chains of alanine, valine, leucine, and isoleucine tend to cluster together within proteins, stabilizing protein structure by means of hydrophobic interactions. Glycine has the simplest structure. Although it is formally nonpolar, its very small side chain makes no real

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

contribution to hydrophobic interactions. Methionine, one of the two sulfur-containing amino acids, has a nonpolar thioether group in its side chain. Proline has an aliphatic side chain with a distinctive cyclic structure. The secondary amino (imino) group of proline residues is held in a rigid conformation that reduces the structural flexibility of polypeptide regions containing proline.



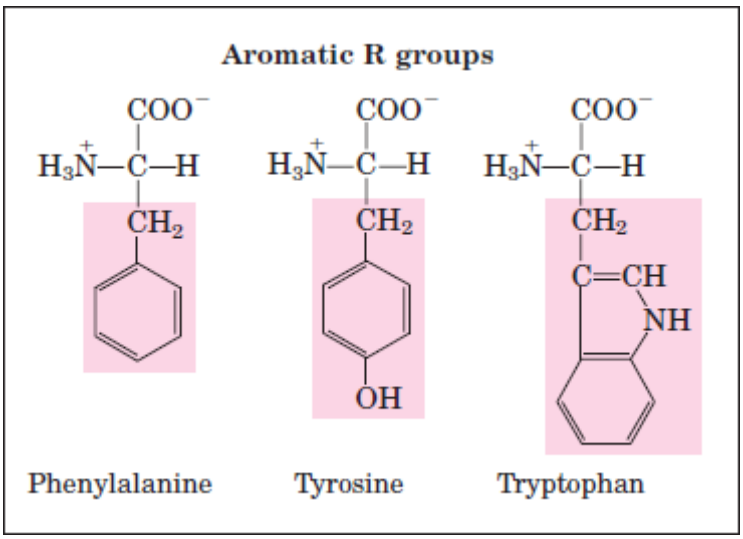
Classification to be continued in next lecture……

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

I) Classification of amino acids on the basis of R-group continued…..

b) Aromatic R Groups :

Phenylalanine, tyrosine, and **tryptophan,** with their aromatic side chains, are relatively nonpolar (hydrophobic). All can participate in hydrophobic interactions. The hydroxyl group of tyrosine can form hydrogen bonds, and it is an important functional group in some enzymes. Tyrosine and tryptophan are significantly more polar than phenylalanine, because of the tyrosine hydroxyl group and the nitrogen of the tryptophan indole ring. Tryptophan and tyrosine, and to a much lesser extent phenylalanine, absorb ultraviolet light This accounts for the characteristic strong absorbance of light by most proteins at a wavelength of 280 nm, a property exploited by researchers in the characterization of proteins.

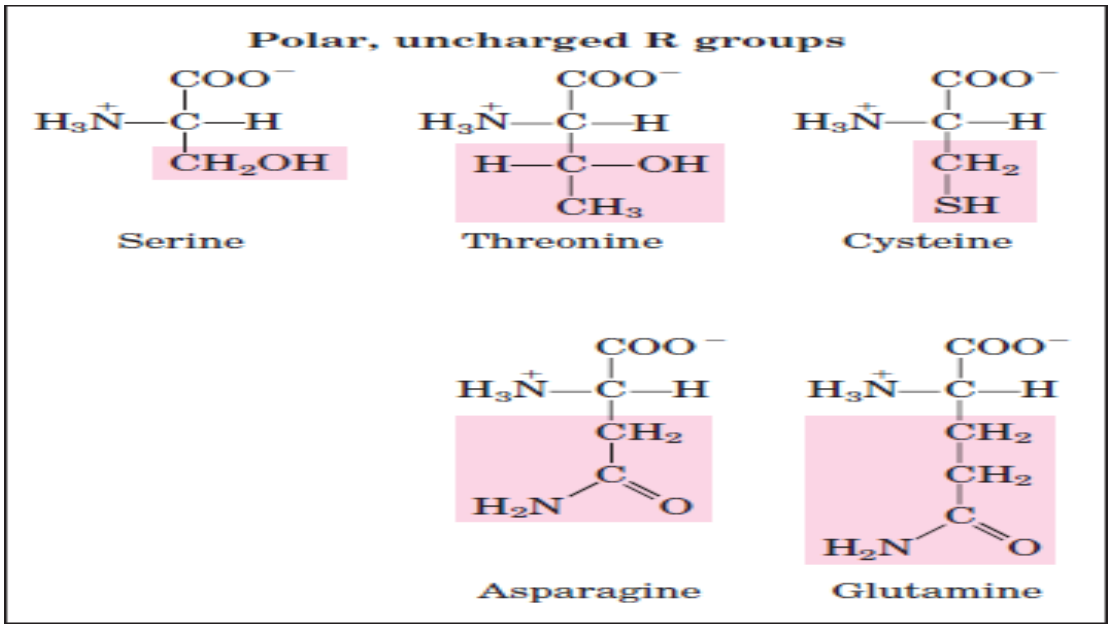


c) Polar, Uncharged R Groups :

The R groups of these amino acids are more soluble in water, or more hydrophilic, than those of the nonpolar amino acids, because they contain functional groups that form hydrogen bonds with water. This class of amino acids includes **serine, threonine, cysteine, asparagine,** and **glutamine.** The polarity of serine and threonine is contributed by their hydroxyl groups; that of cysteine by its sulfhydryl group; and that of asparagine and glutamine by their amide groups.

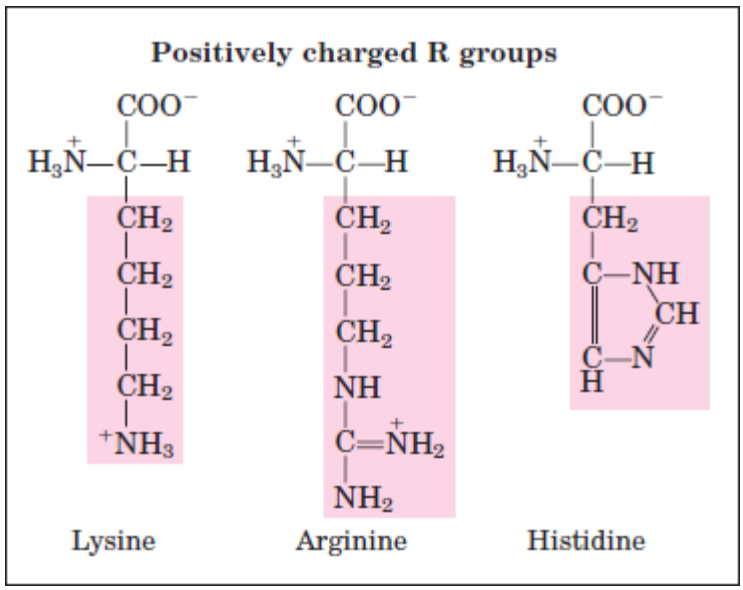
Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Asparagine and glutamine are the amides of two other amino acids also found in proteins, aspartate and glutamate, respectively, to which asparagine and glutamine are easily hydrolyzed by acid or base. Cysteine is readily oxidized to form a covalently linked dimeric amino acid called **cystine,** in which two cysteine molecules or residues are joined by a disulfide bond (Fig. 3–7). 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 protein molecule or between
two different polypeptide chains.



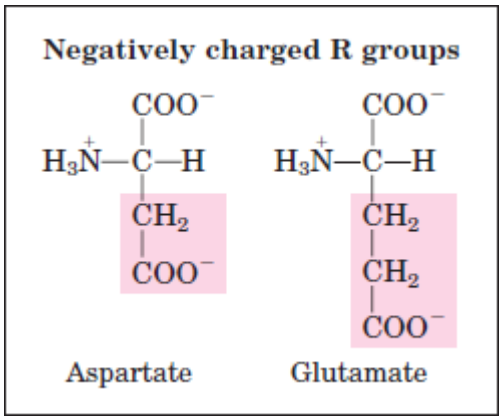
Positively Charged (Basic) R Groups The most hydrophilic R groups are those that are either positively or negatively charged. The amino acids in which the R groups have significant positive charge at pH 7.0 are **lysine,** which has a second primary amino group at the _ position on its aliphatic chain; **arginine,** which has a positively charged guanidino group; and **histidine,** which has an imidazole group. Histidine is the only common amino acid having an ionizable side chain with a p*K*a near neutrality. In many enzyme-catalyzed reactions, a His residue facilitates the reaction by serving as a proton donor/acceptor.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar



Negatively Charged (Acidic) R Groups:

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.



II. Classification of amino acids on the basis of Nutrition:

1. Essential amino acids:

- These amino acids are not synthesized in cells of human beings, so these should be essentially present in diet.
- PVTTIMHALL; Phenylalanine, Valine, Threonine, Tryptophan, Isoleucine, Methionine, Histidine, Arginine*, Leucine, Lysine
- Arginine is conditional amino acids (Essential for infants, non essential for adults)

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

2. Non essential amino acids:

- These aminoacids can be synthesized in body, so need not be included in diet.
- (GASCAGAGTP); Glycine, Alanine, Serine, Cysteine, Asparagine, Glutamine, Aspartic acid, Glutamic acid, Tyrosine, Proline

III. Classification of amino acids on the basis of Catabolism

1. Glucogenic amino acids:

- These aminoacids serves as precursors of **gluconeogenesis** for glucose formation
- GAMD (Glycine, Alanine, methionine, Aspartic acid).

2. Ketogenic amino acids:

- These aminoacids breakdown to form ketone bodies.
- Leucine and Lysine

3. Both glucogenic and ketogenic amino acids:

- These amino acids breakdown to form precursors for both ketone bodies and glucose.
- Isoleucine, Phenylalanine, Tryptophan and tyrosine

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Uncommon Amino Acids/Non standard amino acids

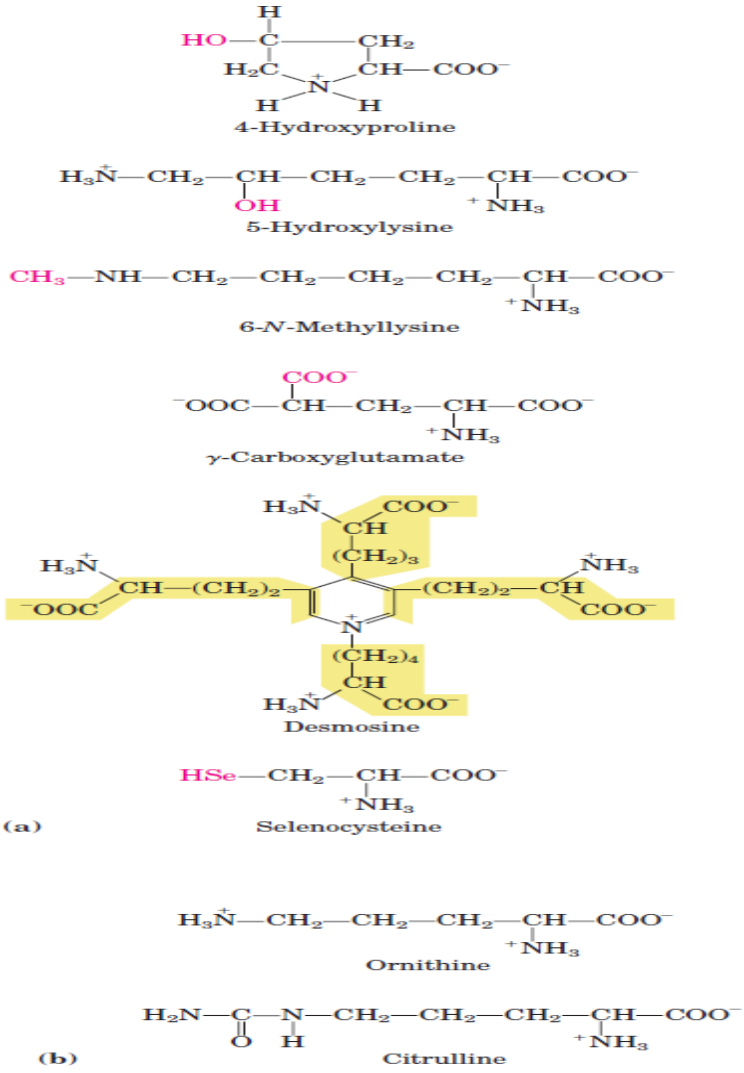
In addition to the 20 common amino acids, proteins may contain residues created by modification of common residues already incorporated into a polypeptide (Fig. below). Among these uncommon amino acids are **4-hydroxyproline,** a derivative of proline, and **5-hydroxylysine,** derived from lysine. The 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 bloodclotting protein prothrombin and in certain other proteins

that bind Ca2_ as part of their biological function. More complex is **desmosine,** a derivative of four Lys residues, which is found in the fibrous protein elastin. **Selenocysteine** is a special case. This rare amino acid residue is introduced during protein synthesis rather than created through a postsynthetic modification. It contains selenium rather than the sulfur of cysteine. Actually derived from serine, selenocysteine is a constituent of just a few known proteins. Some 300 additional amino acids have been found in cells. They have a variety of functions but are not

constituents of proteins. **Ornithine** and **citrulline** deserve special note because they are key

intermediates (metabolites) in the biosynthesis of arginine and in the urea cycle.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

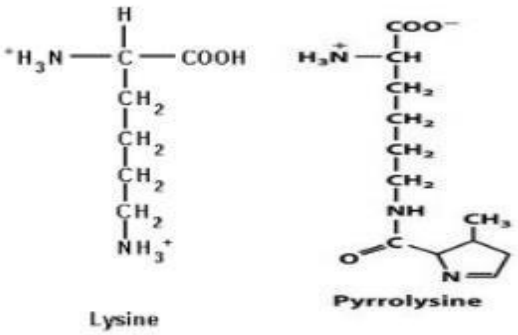


Selenocysteine (Sec) and pyrrolysine (Pyl) are rare amino acids that are cotranslationally inserted into proteins and known as the 21st and 22nd amino acids in the genetic code. Sec and Pyl are encoded by UGA and UAG codons, respectively, which normally serve as stop signals

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Lysine V/s Pyrrolysine

- Pyl is similar to Lys, but with an added ***pyrroline ring*** linked the ***end of Lys side chain*** **(stretching from NH2 toNH)**.



Q. Why do Proteins absorb at 280nm?

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Amino acid Properties

Physical Properties

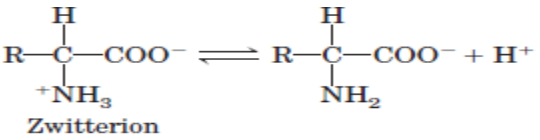
1. Amino acids are colorless, crystalline solid.
2. All amino acids have a high melting point greater than 200^{o}
3. Solubility: They are soluble in water, slightly soluble in alcohol and dissolve with difficulty in methanol, ethanol, and propanol. R-group of amino acids and pH of the solvent play important role in solubility.
4. On heating to high temperatures, they decompose.
5. All amino acids (except glycine) are optically active.

Chemical Properties

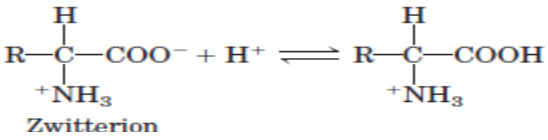
1. Amino Acids Can Act as Acids and Bases

A zwitterion is a molecule with functional groups, of which at least one has a positive and one has a negative electrical charge. The net charge of the entire molecule is zero. Amino acids are the best-known examples of zwitterions. When an amino acid is dissolved in water, it exists in solution as the dipolar ion, or **zwitterion** (German for "hybrid ion"), shown in Figure . A zwitterion can act as either an acid (proton donor):

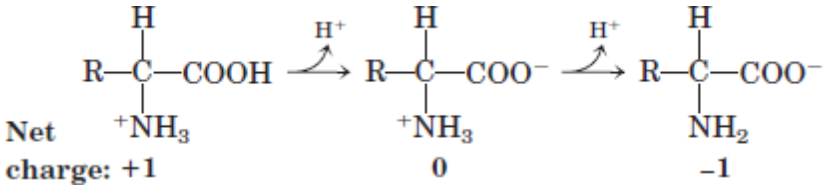
Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar



or a base (proton acceptor):



Substances having this dual nature are **amphoteric** and are often called **ampholytes** (from "amphoteric electrolytes"). A simple monoamino monocarboxylic α-amino acid, such as alanine, is a diprotic acid when fully protonated—it has two groups, the -COOH group and the -NH3 _ group, that can yield protons:



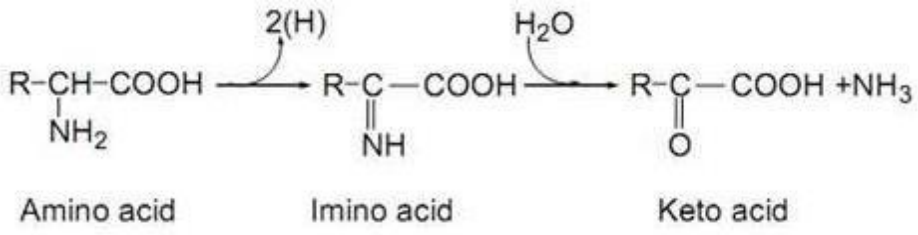
Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Chemical Properties (Chemical reactions) of Amino acids *Continued…..*

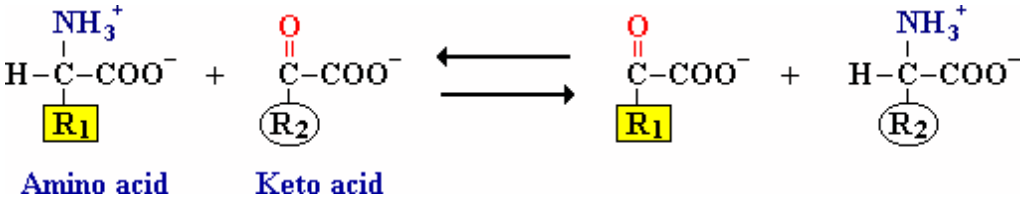
The general reactions of amino acids are due to presence of two functional groups namely carboxyl (-COOH) group and amino (-NH2) group and due to side chains

1. Reactions due to amino group

i) Oxidative deamination-An amino group is removed and corresponding α-keto acid is formed. α -keto acid produced is either converted to glucose or ketone bodies or is completely oxidized.



ii) ii) Transamination-Transfer of an α amino group from an amino acid to an α keto acid to form a new amino acid and a corresponding keto acid.



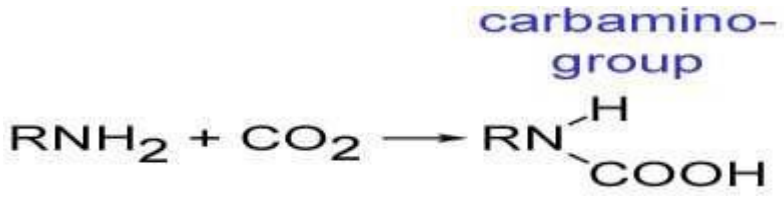
iii) Formation of carbamino compound

CO2 binds to α amino acid on the globin chain of hemoglobin to form carbamino hemoglobin

The reaction takes place at alkaline pH and serves as a mechanism for the transfer of Carbon dioxide from the

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

tissues to the lungs by hemoglobin.



2) Reactions due to carboxyl group

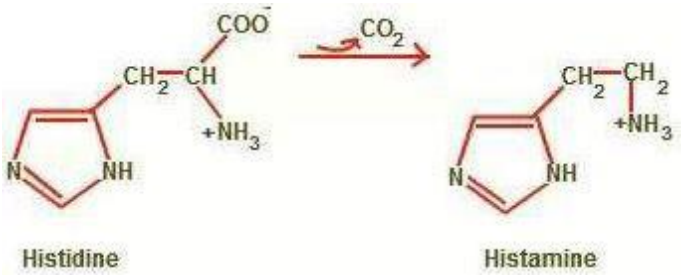
i) Decarboxylation- Amino acids undergo alpha decarboxylation to form corresponding amines.

Examples-

Glutamic acid _____GABA

Histidine _____Histamine

Tyrosine_____Tyramine



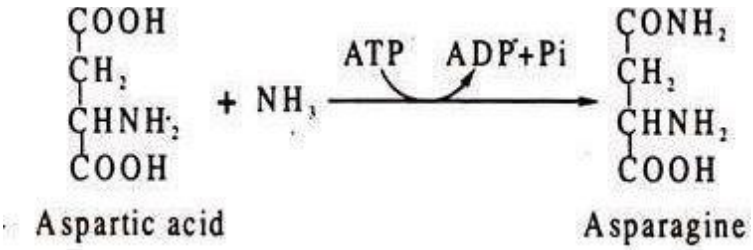
ii) Formation of amide linkage

Non α carboxyl group of an acidic amino acid reacts with ammonia by condensation reaction to form corresponding amides

Aspartic acid→ Asparagine

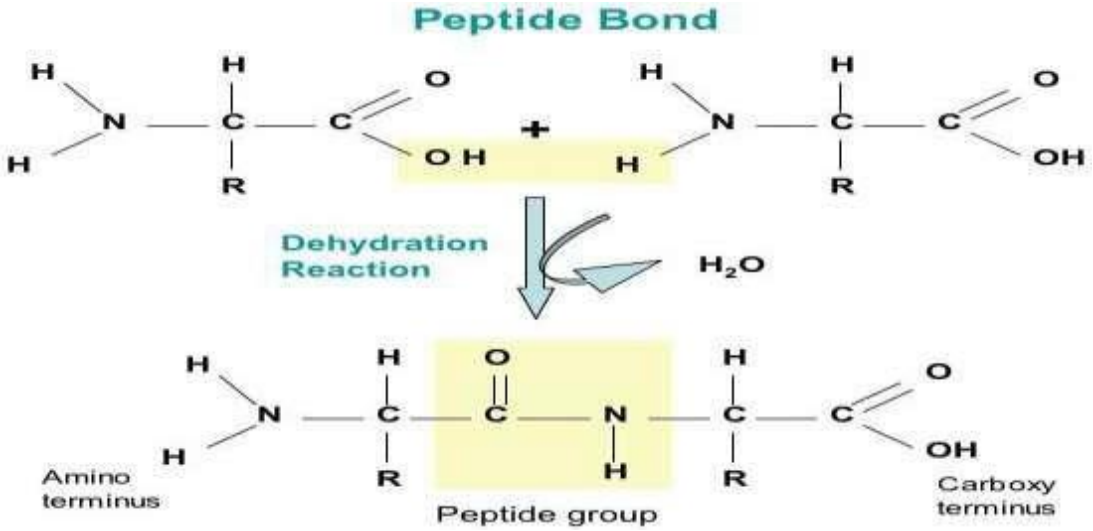
Glutamic acid → Glutamine

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar



3) Reactions due to both amino & carboxyl groups

Formation of peptide bond- Carboxyl group of an amino acid binds with amino group of another amino acid forming a peptide bond with the loss of one molecule of water.



4. Reactions due to side chains

1) Ester formation

 OH containing amino acids e.g. serine, threonine can form esters with phosphoric acid in the formation of phosphoproteins (figure-1)

 OH group containing amino acid can also form: Glycosides – by forming O- glycosidic bond with carbohydrate residues (figure-2)

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

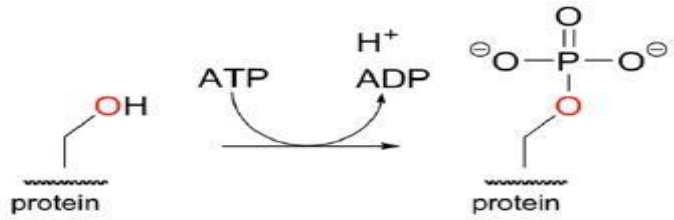


Figure-1

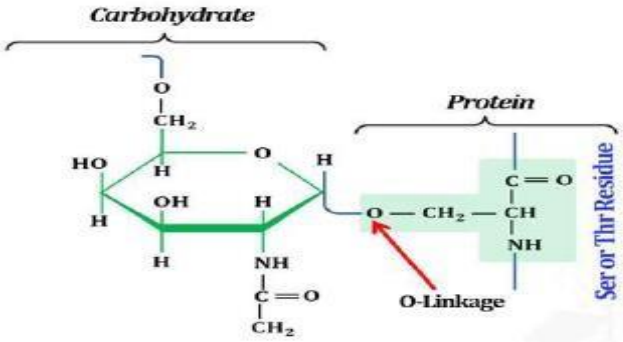
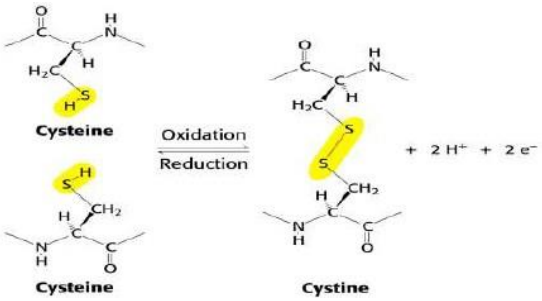


Figure-2

2) Reactions due to SH group (Formation of sulphide bonds)

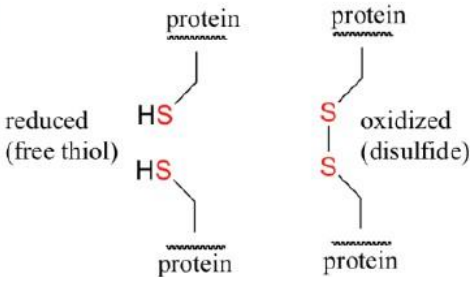
- Cysteine has a sulfhydryl group(SH) group and can form a disulphide (S-S) bond with another cysteine residue.
- The dimer is called Cystine



Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Formation of disulphide bond

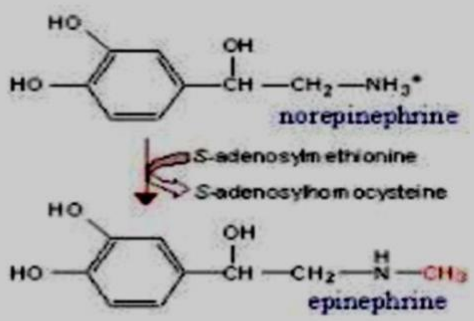
- Two cysteine residues can connect two polypeptide chains by the formation of interchain disulphide chains.



3)Transmethylation

The methyl group of Methionine can be transferred after activation to an acceptor for the formation of important biological compounds.

For example (figure) conversion of norepinephrine to epinephrine by S Adenosyl methionine (active form of methionine



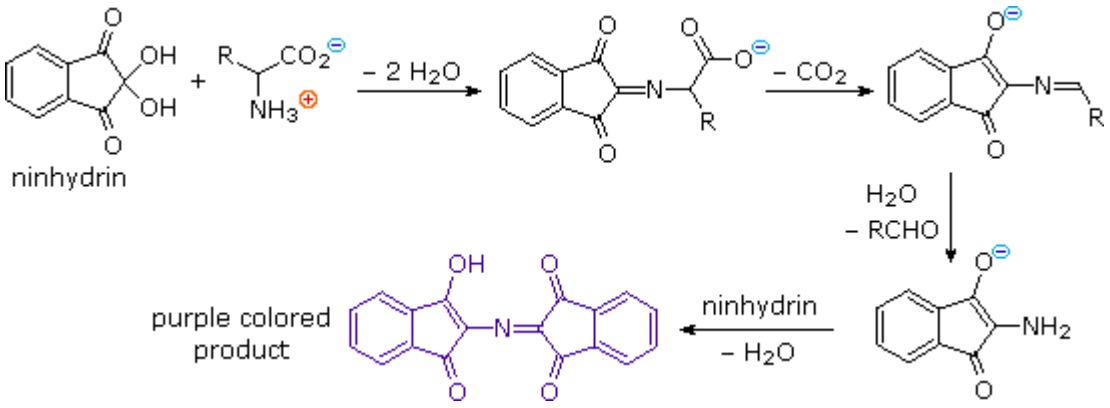
The Ninhydrin Reaction

In addition to these common reactions of amines and carboxylic acids, common alpha-amino acids, except proline, undergo a unique reaction with the triketohydrindene hydrate known as ninhydrin. Among the products of this unusual reaction (shown on the left below) is a purple colored imino derivative (diketohydrin also known as Ruhemann's purple,) which provides as a useful color test for these amino acids, most of which are colorless. A common application of the ninhydrin test is the visualization of amino acids in **paper chromatography**.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Amino acid + Ninhydrin→ Keto acid + NHr+COz+Hydrindantin

Hydrindantin+ NH_3 + Ninhydrin- → Ruhemann's purple



(Note : Proline and hydroxyproline give yellow colour with ninhydrin).

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Introduction to Peptides & Proteins

We now turn to polymers of amino acids, the **peptides** and **proteins.** Biologically occurring polypeptides range in size from small to very large, consisting of two or three to thousands of linked amino acid residues. Our focus is on the fundamental chemical properties of these polymers.

Introduction to Peptides and Proteins (Peptides Are Chains of Amino Acids)

Two amino acid molecules can be covalently joined through a substituted amide linkage, termed a **peptide bond,** to yield a dipeptide. Such a linkage is formed by removal of the elements of water (dehydration) from the α-carboxyl group of one amino acid and the α-amino group of another (Fig). Peptide bond formation is an example of a condensation reaction, a common class of reactions in living cells. Three amino acids can be joined by two peptide bonds to form a tripeptide; similarly, amino acids can be linked to form tetrapeptides, pentapeptides, and so forth. When a few amino acids are joined in this fashion, the structure is called an **oligopeptide.** When many amino acids are joined, the product is called a **polypeptide.**

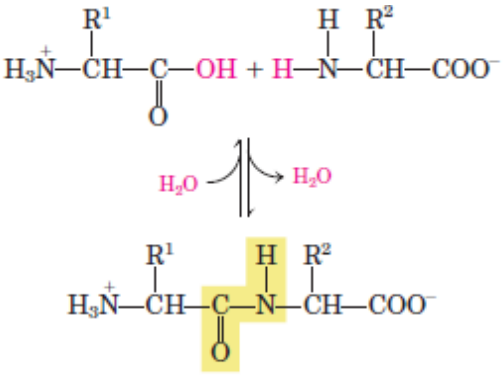


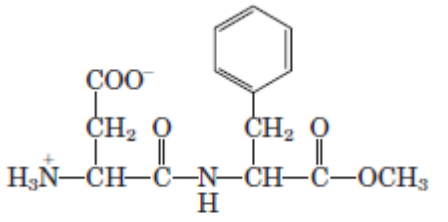
FIGURE 3–13 Formation of a peptide bond by condensation. The α-amino group of one amino acid (with R^2 group) acts as a nucleophile to displace the hydroxyl group of another amino acid (with R^1 group), forming a peptide bond (shaded in yellow). Amino groups are good nucleophiles, but the hydroxyl group is a poor leaving group and is not readily displaced. At physiological pH, the reaction shown does not occur to any appreciable extent.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Proteins may have thousands of amino acid residues. Although the terms “protein” and “polypeptide” are sometimes used interchangeably, molecules referred to as polypeptides generally have molecular weights below 10,000, and those called proteins have higher molecular weights.

As already noted, an amino acid unit in a peptide is often called a **residue** (the part left over after losing a hydrogen atom from its amino group and the hydroxyl moiety from its carboxyl group). In a peptide, the amino acid residue at the end with a free α- amino group is the **amino-terminal** (or *N*-terminal) residue; the residue at the other end, which has a free carboxyl group, is the **carboxyl-terminal** (*C*-terminal) residue. Although hydrolysis of a peptide bond is an exergonic reaction, it occurs slowly because of its high activation energy. As a result, the peptide bonds in proteins are quite stable, with an average half-life ($t1/2$) of about 7 years under most intracellular conditions.

No generalizations can be made about the molecular weights of biologically active peptides and proteins in relation to their functions. Naturally occurring peptides range in length from two to many thousands of amino acid residues. Even the smallest peptides can have biologically important effects. Consider the commercially synthesized dipeptide L-aspartyl-L-phenylalanine methyl ester, the artificial sweetener better known as aspartame or NutraSweet.



L-Aspartyl-L-phenylalanine methyl ester (aspartame)

Many small peptides exert their effects at very low concentrations. For example, a number of vertebratehormones are small peptides. These include **oxytocin** (nine amino acid residues), which is secreted by the posterior pituitary and stimulates uterine contractions; bradykinin (nine

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

residues), which inhibits inflammation of tissues; and thyrotropin-releasing factor (three residues), which is formed in the hypothalamus and stimulates the release of another hormone, thyrotropin, from the anterior pituitary gland. Some extremely toxic mushroom poisons, such as amanitin, are also small peptides, as are many antibiotics.

Slightly larger are small polypeptides and oligopeptides such as the pancreatic hormone insulin, which contains two polypeptide chains, one having 30 amino acid residues and the other 21. Glucagon, another pancreatic hormone, has 29 residues; it opposes the action of insulin. Corticotropin is a 39-residue hormone of the anterior pituitary gland that stimulates the adrenal cortex.

Length of the polypeptide chains vary considerably in proteins? Human cytochrome *c* has 104 amino acid residues linked in a single chain; bovine chymotrypsinogen has 245 residues. At the extreme is titin, a constituent of vertebrate muscle, which has nearly 27,000 amino acid residues and a molecular weight of about 3,000,000. The vast majority of naturally occurring proteins are much smaller than this, containing fewer than 2,000 amino acid residues.

Some proteins consist of a single polypeptide chain, but others, called **multisubunit** proteins, have two or more polypeptides associated noncovalently. The individual polypeptide chains in a multisubunit protein may be identical or different. If at least two are identical the protein is said to be **oligomeric,** and the identical units (consisting of one or more polypeptide chains) are referred to as **protomers.** Hemoglobin, for example, has four polypeptide subunits: two identical α- chains and two identical β-chains, all four held together by noncovalent interactions. Each α-subunit is paired in an identical way with a β subunit within the structure of this multisubunit protein, so that hemoglobin can be considered either a tetramer of four polypeptide subunits or a dimer of $\alpha\beta$ protomers. A few proteins contain two or more polypeptide chains linked covalently. For example, the two polypeptide chains of insulin are linked by disulfide bonds. In predominate such cases, the individual polypeptides are not considconsidered subunits but are commonly referred to simply as chains.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

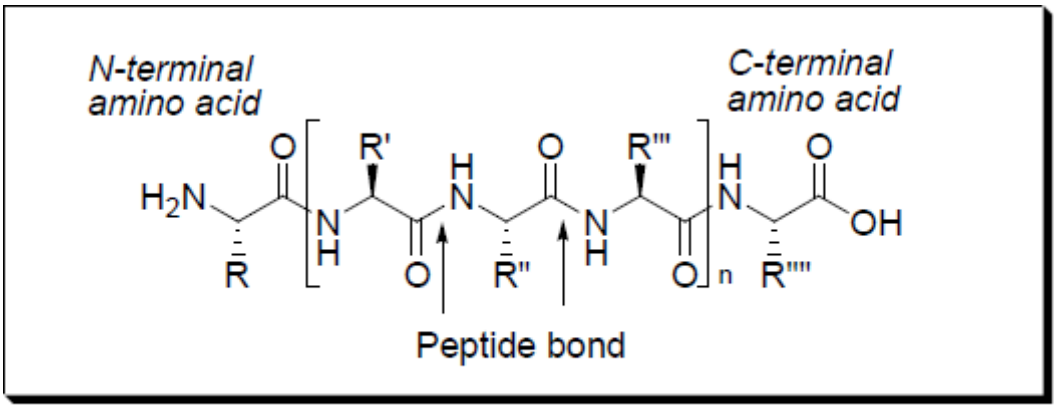
We can calculate the approximate number of amino acid residues in a simple protein containing no other chemical constituents by dividing its molecular weight by 110. Although the average molecular weight of the 20 common amino acids is about 138, the smaller amino acids in most proteins. If we take into account the proportions in which the various amino acids occur in protein the average molecular weight of protein amino acids is nearer to 128. Because a molecule of water (Mr 18) is removed to create each peptide bond, the average

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Protein Structure

In the previous we discussed several aspect of amino acids, the building block of proteins. Now in the current lecture, we will discuss more about the protein structure and its function. Proteins are polymers of amino acids, joined by the covalent bonds, known as pepide bond. A peptide bond is an amide linkage formed between carboxyl group of first and amino group of second amino acid with release of water (Figure 32.1, A,B). it is a dehydration synthesis or condensation reaction. . Every polypeptide chain has a free N- and C terminals For large macromolecules such as proteins, the tasks of describing and understanding structure are approached at several levels of complexity, arranged in a kind of conceptual hierarchy. Four levels of protein structure are commonly defined

- ✓ Primary Structure
- ✓ Secondary Structure
- ✓ Tertiary structure and
- ✓ Quaternary Structure.



Scheme 1

1. Primary structure & Peptide Bond

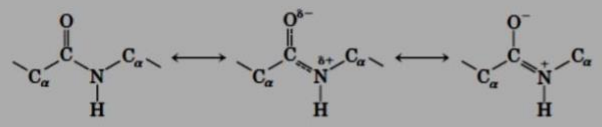
A description of all covalent bonds (mainly peptide bonds and disulfide bonds) linking amino acid residues in a polypeptide chain is its **primary structure.** The most important element of primary structure is the *sequence* of amino acid residues. Thus primary structure may be defined

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

as the linear sequence of amino acids in a polypeptide chain that are joined by peptide bonds. Covalent bonds also place important constraints on the conformation of a polypeptide. In the late 1930s, Linus Pauling and Robert Corey embarked on a series of studies that laid the foundation

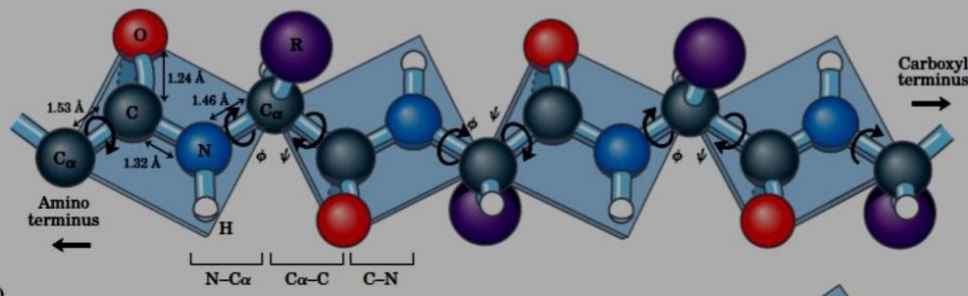
for our present understanding of protein structure. They began with a careful analysis of the peptide bond. The α carbons of adjacent amino acid residues are separated by three covalent bonds, arranged as Cα-C-N-Cα. X-ray diffraction studies of crystals of amino acids and of simple dipeptides and tripeptides demonstrated that the peptide CON bond is somewhat shorter than the CON bond in a simple amine and that the atoms associated with the peptide bond are coplanar. This indicated a resonance or partial sharing of two pairs of electrons between the carbonyl oxygen and the amide nitrogen (Fig. 4–2a). The oxygen has a partial negative charge and the nitrogen a partial positive charge, setting up a small electric dipole. The six atoms of the **peptide group** lie in a single plane, with the oxygen atom of the carbonyl group and the hydrogen atom of the amide nitrogen trans to each other. From these findings Pauling and Corey concluded that the peptide C-N bonds are unable to rotate freely because of their partial double-bond character. Rotation is permitted about the N-Cα *a*nd the Cα-C bonds. The backbone of a polypeptide chain can thus be pictured as a series of rigid planes with consecutive planes sharing a common point of rotation at Cα (Fig. 4–2b). The rigid peptide bonds limit the range of conformations that can be assumed by a polypeptide chain. By convention, the bond angles resulting from rotations at Cα are labeled φ (phi) for the N-Cα bond and ψ (psi) for the Cα-C bond. Again by convention, both φ and ψ are defined as 180º when the polypeptide is in its fully extended conformation and all peptide groups are in the same plane (Fig. 4–2b). In principle, φ and ψ can have any value between -180º and +180º, but many values are prohibited by steric interference between atoms in the polypeptide backbone and amino acid side chains. The conformation in which both φ and ψ are 0_ (Fig. 4–2c) is prohibited for this reason; this conformation is used merely as a reference point for describing the angles of rotation. Allowed values for φ and ψ are graphically revealed when ψ is plotted versus φ in a **Ramachandran plot** (Fig. below), introduced by G. N. Ramachandran.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

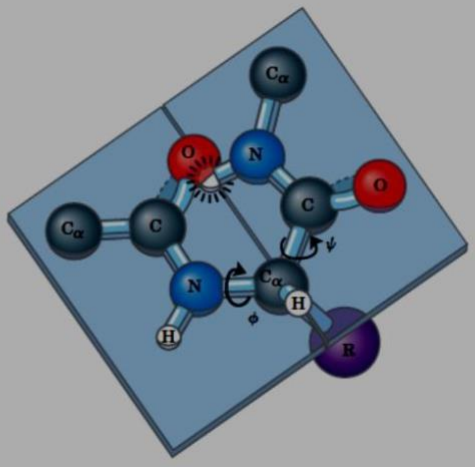


The carbonyl oxygen has a partial negative charge and the amide nitrogen a partial positive charge, setting up a small electric dipole. Virtually all peptide bonds in proteins occur in this trans configuration; an exception is noted in Figure 4–8b.

(a)



(b)



(c)

FIGURE 4–2 The planar peptide group. **(a)** Each peptide bond has some double-bond character due to resonance and cannot rotate. **(b)** Three bonds separate sequential α carbons in a polypeptide chain. The N—C_α and C_α—C bonds can rotate, with bond angles designated ϕ and ψ, respectively. The peptide C—N bond is not free to rotate. Other single bonds in the backbone may also be rotationally hindered, depending on the size and charge of the R groups. In the conformation shown, ϕ and ψ are 180° (or −180°). As one looks out from the α carbon, the ψ and ϕ angles increase as the carbonyl or amide nitrogens (respectively) rotate clockwise. **(c)** By convention, both ϕ and ψ are defined as 0° when the two peptide bonds flanking that α carbon are in the same plane and positioned as shown. In a protein, this conformation is prohibited by steric overlap between an α-carbonyl oxygen and an α-amino hydrogen atom. To illustrate the bonds between atoms, the balls representing each atom are smaller than the van der Waals radii for this scale. 1 Å = 0.1 nm.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Protein Structure ***Continued……***

2. Secondary Structure

The secondary structure of a protein is the local fold of the protein backbone. Can a polypeptide chain fold into a regularly repeating structure? In **1951**, Linus Pauling and Robert Corey proposed two periodic structures called the *a helix* (alpha helix) and the *β pleated sheet* (beta pleated sheet). Subsequently, other structures such as the *β turn* and *omega* (Ω) *loop* were identified. Although not periodic, these common turn or loop structures are well defined and contribute with *α* helices and βsheets to form the final protein structure. Alpha helices, β strands, and turns are formed by a regular pattern of hydrogen bonds between the peptide N-H and C=O groups of amino acids that are *near one another* in *the linear sequence.* Such folded segments are called *secondary structure.*

i. α helix

In evaluating potential structures, Pauling and Corey considered which conformations of peptides were sterically allowed and which most fully exploited the hydrogen-bonding capacity of the backbone NH and CO groups. The first of their proposed structures, the *a helix,* is a rod like structure (Figure). A tightly coiled backbone forms the inner part of the rod and the side chains extend outward in a helical array. The α helix is stabilized by hydrogen bonds between the NH and CO groups of the main chain. In particular, the CO group of each amino acid forms a hydrogen bond with the NH group of the amino acid that is situated four residues ahead in the sequence (Figure 2.30). Thus, except for amino acids near the ends of an α helix, all *the main-chain* CO *and NH groups are hydrogen bonded.* Each residue is related to the next one by a *rise,* also called *translation,* of 1.5 A along the helix axis and a rotation of 100 degrees, which gives 3.6 amino acid residues per turn of helix. Thus, amino acids spaced three and four apart in the sequence are spatially quite close to one another in an α helix. In contrast, amino acids spaced two apart in the sequence are situated on opposite sides of the helix and so are unlikely to make contact. The *pitch* of the α helix, which is equal to the product of the translation 1.5 Aº and the number of residues per turn (3.6), is 5.4 A. The *screw sense* of a helix can be right-handed (clockwise) or left -handed (counterclockwise). The Ramachandran diagram reveals that both the right-handed and the left handed helices are among allowed conformations . However, right -

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

handed helices are energetically more favorable because there is less steric clash between the side chains and the backbone. *Essentially all a helices found in proteins are right-handed.* In schematic representations of proteins, α helices are depicted as twisted ribbons or rods (figure 2.32).

Pauling and Corey predicted the structure of the α helix 6 years before it was actually seen in the x-ray reconstruction of the structure of myoglobin n. *The elucidation of the structure of the a helix* is *a landmark in biochemistry because it demonstrated that the conformation of a polypeptide chain could be predicted if the properties of its components are rigorously and precisely known* .

The α helical content of proteins ranges widely, from none to almost 100%. for example, about 75% of the residues in ferritin, a protein that helps store iron, are in IX helices (Figure 2.33). Indeed, about 25% of all soluble proteins are composed of α helices connected by loops and turns of the polypeptide chain. Single α helices are usually less than 45 Aº long. Many proteins that span biological membranes also contain α helices.

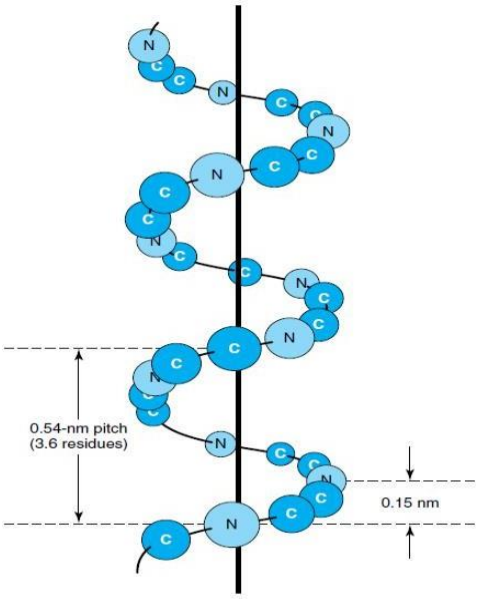


Figure 5–2. Orientation of the main chain atoms of a peptide about the axis of an α helix.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

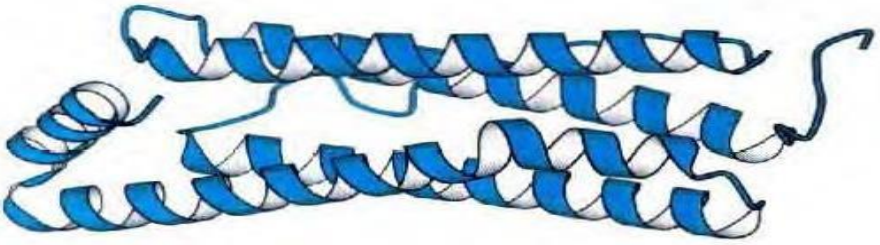


Figure 2.33 A largely α-helical protein. Ferritin, an iron-storage protein, is built from a bundle of α helices.

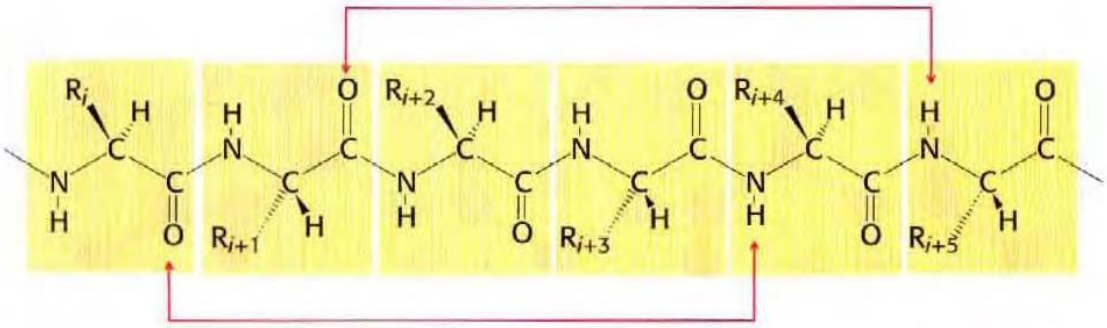


Figure 2.30 Hydrogen-bonding scheme for an α helix. In the α helix, the CO group of residue i forms a hydrogen bond with the NH group of residue $i + 4$.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

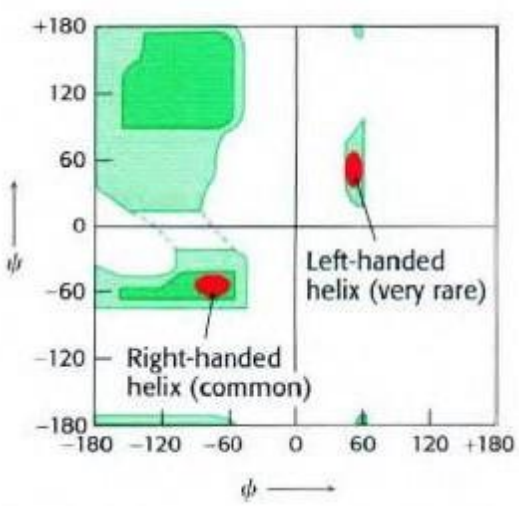


Figure 2.31 Ramachandran diagram for helices. Both right- and left-handed helices lie in regions of allowed conformations in the Ramachandran diagram. However, essentially all α helices in proteins are right-handed.

Screw sense

Describes the direction in which a helical structure rotates with respect to its axis. If, viewed down the axis of a helix, the chain turns in a clockwise direction, it has a right-handed screw sense. If the turning is counterclockwise, the screw sense is left-handed.

Torsion angle

A measure of the rotation about a bond, usually taken to lie between -180 and +180 degrees. Torsion angles are sometimes called dihedral angles example ψ &φ.

Knowing the Right Hand from the Left

There is a simple method for determining whether a helical structure is right-handed or left-handed. Make fists of your two hands with thumbs outstretched and pointing straight up. Looking at your right hand, think of a helix spiraling up your right thumb in the direction in which the other four fingers are curled as shown (counterclockwise). The resulting helix is right-handed. Your left hand will demonstrate a left-handed helix, which rotates in the clockwise direction as it spirals up your thumb.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Secondary structure *contd*…….

B-Sheets

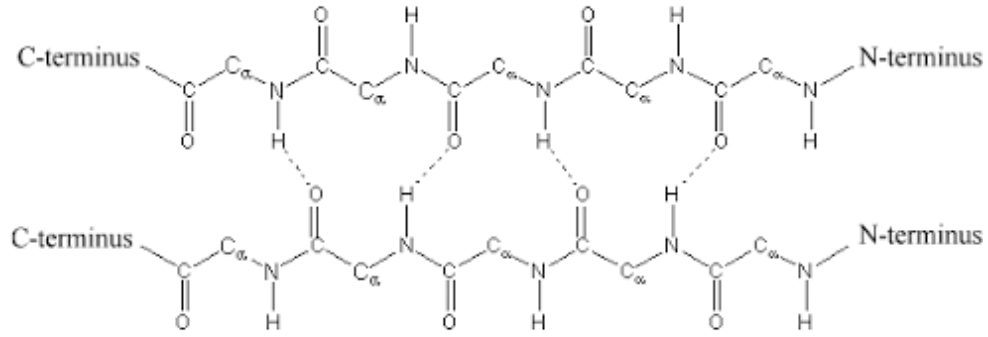
Pauling and Corey proposed another periodic structural motif, which they named the β- *pleated sheet* (β because it was the second structure that they elucidated, the α helix having been the first). The β pleated sheet (or, more simply, the β sheet) differs markedly from the rod like α helix. It is composed of two or more polypeptide chains called *β strands.* A β strand is almost fully extended rather than being tightly coiled as in the α helix. A range of extended structures are sterically allowed (Figure 2.34). The distance between adjacent amino acids along a β strand is approximately 3.5 A, in contrast to a distance of 1. 5 A along an α helix. The side chains of adjacent amino acids point in opposite directions. A β sheet is formed by linking two or more β strands lying next to one another through hydrogen bonds. Adjacent chains in a β- sheet can run in opposite directions (antiparallel β sheet) or in the same direction (parallel β sheet). In the antiparallel arrangement, the NH group and the CO group of each amino acid are respectively hydrogen bonded to the CO group and the NH group of a partner on the adjacent chain (Figure 2.36). In the parallel arrangement, the hydrogen-bonding scheme is slightly more complicated.

For each amino acid, the NH group is hydrogen bonded to the CO group of one amino acid on the adjacent strand, whereas the CO group is hydrogen bonded to the NH group on the amino acid two residues farther along the chain (Figure 2.37). The repeat period is shorter for the parallel conformation (6.5 Å, versus 7 Å for antiparallel) .Many strands, typically 4 or 5 but as many as 10 or more, can come together in β-sheets. Such β-sheets can be purely antiparallei, purely parallel , or mixed (Figure 2.3H). In schematic representation β - strands are usually depicted by broad arrows pointing in the direction of the carboxyl-terminal end to indicate the type of β- sheet formed parallel or antiparallel.

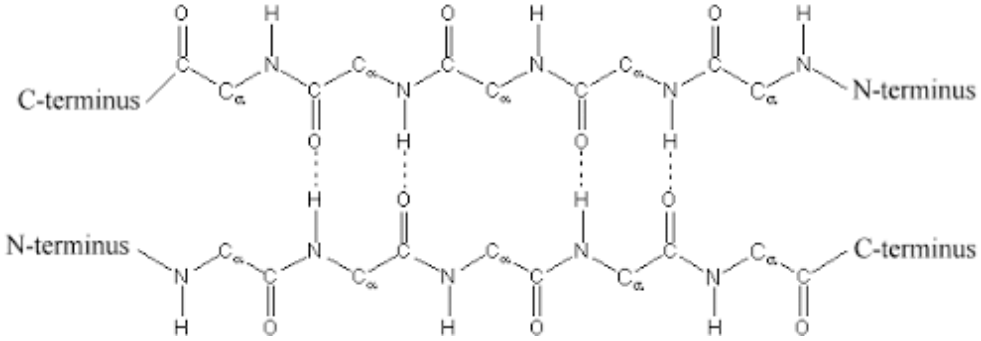
Some protein structures limit the kinds of amino acids that can occur in the *β*-sheet. When two or more *B*- sheets are layered close together within a protein, the R groups of the amino acid residues on the touching surfaces must be relatively small. β-Keratins such as silk fibroin and the fibroin of spider webs have a very high content of Gly and Ala residues, the two amino acids with the smallest R groups. Indeed, in silk fibroin Gly and Ala alternate over large parts of the sequence.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Parallel β Sheet



Antiparallel β Sheet



Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

iii) *β*-turns

In globular proteins, which have a compact folded structure, nearly one-third of the amino acid residues are in turns or loops where the polypeptide chain reverses direction. These are the connecting elements that link successive runs of *α* helix or *β* conformation. Particularly common are *β*-**turns** that connect the ends of two adjacent segments of an antiparallel *β* sheet. The structure is a 180º turn involving four amino acid residues, with the carbonyl oxygen of the residue i forming a hydrogen bond with the amino-group hydrogen of residue i+3. The peptide groups of the central two residues do not participate in any inte-rresidue hydrogen bonding. *Gly and Pro residues often occur in β-turns, the former because it is small and flexible, the latter because peptide bonds involving the imino nitrogen of proline readily assume the cis configuration* , a form that is particularly amenable to a tight turn. In other cases, more elaborate structures are responsible for chain reversals. These structures are called *loops* or sometimes **Ω**- *loops* (omega loops) to suggest their overall shape. Unlike α- helices and β strands, loops do not have regular, periodic structures. Nonetheless, loop structures are often rigid and well defin ed . Turns and loops invariably lie on the surfaces of proteins and thus often participate in interactions between proteins and other molecules. Considerably less common is the γ- turn, a three residue turn with a hydrogen bond between the first and third.

The term random coil, is used to describe a protein that has lost its secondary structure (the protein is then said to be denatured).

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

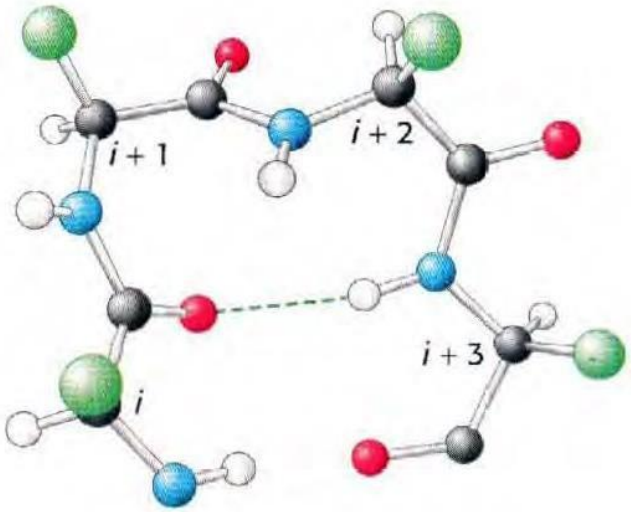


Figure 2.41 Structure of a reverse turn. The CO group of residue *i* of the polypeptide chain is hydrogen bonded to the NH group of residue *i* +3 to stabilize the turn.

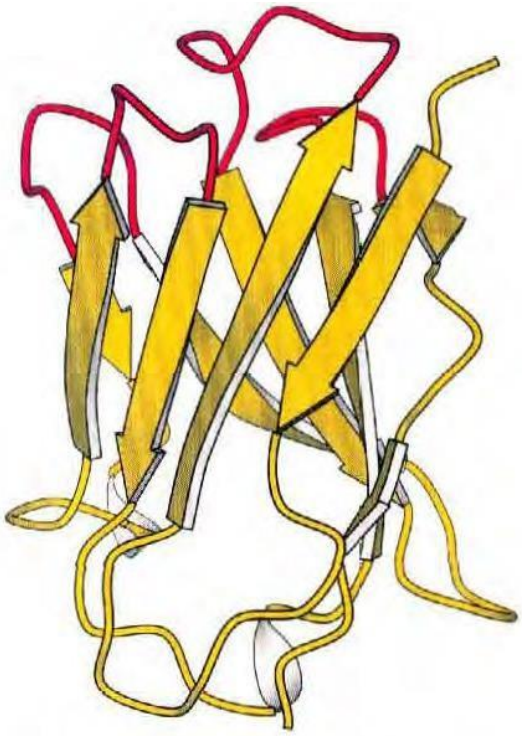


Figure 2.42 Loops on a protein surface. A part of an antibody molecule has surface loops (shown in red) that mediate interactions with other molecules. [Drawn from 7FTP.pdb.]

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

Forces Stablising Protein Structure

Tertiary structure is the three-dimensional structure of a protein. While individual amino acids in the primary sequence can interact with one another to form secondary structures such as helices and sheets and individual amino acids from distant parts of the primary sequence can intermingle via charge-charge, hydrophobic, disulfide, or other interactions, the formation of these bonds and interactions will serve to change the shape of the overall protein. The folding that we end up with for a given polypeptide is the tertiary structure.

There are four types of bonding interactions between "side chains" including: hydrogen bonding, salt bridges, disulfide bonds, and non-polar hydrophobic interactions.

1. **Covalent bonds -Disulfide Bridges**

Covalent bonds are the strongest chemical bonds contributing to protein structure. Covalent bonds arise when two atoms share electrons.In addition to the covalent bonds that connect the atoms of a single amino acid and the covalent peptide bond that links amino acids in a protein chain, covalent bonds between cysteine side chains can be important determinants of protein structure. Cysteine is the sole amino acid whose side chain can form covalent bonds, yielding disulfide bridges with other cysteine side chains: --CH2-S-S-CH2

2. **Electrostatic Interactions- Ionic Bonds – (Salt Bridges)**

Ionic bonds are formed as amino acids bearing opposite electrical charges are juxtaposed in the hydrophobic core of proteins. Ionic bonding in the interior is rare because most charged amino acids lie on the protein surface. Although rare, ionic bonds can be important to protein structure because they are potent electrostatic attractions that can approach the strength of covalent bonds.

3. **Hydrogen Bonds**

When two atoms bearing partial negative charges share a partially positively charged hydrogen, the atoms are engaged in a hydrogen bond (H-bond). The correct 3-D structure of a protein is often dependent on an intricate network of H-bonds. Individual hydrogen bonds are much weaker than a covalent bond, but collectively, they can exert strong forces.

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

These can occur between a variety of atoms, involving:

- ✓ atoms on two different amino acid sidechains
- ✓ atoms on amino acid sidechains and water molecules at the protein surface
- ✓ atoms on amino acid sidechains and protein backbone atoms
- ✓ backbone atoms and water molecules at the protein surface
- ✓ backbone atoms on two different amino acids

Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

4. Van der Waals Forces

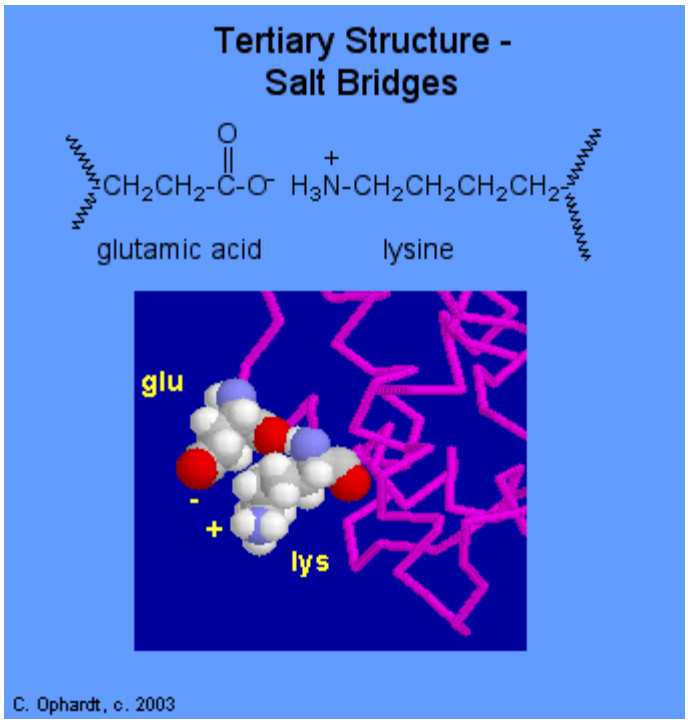
The Van der Waals force is a transient, weak electrical attraction of one atom for another. Van der Waals attractions exist because every atom has an electron cloud that can fluctuate, yielding a temporary electric dipole. The transient dipole in one atom can induce a complementary dipole in another atom, provided the two atoms are quite close. These short-lived, complementary dipoles provide a weak electrostatic attraction, the Van der Waals force. Of course, if the two electron clouds of adjacent atoms are too close, repulsive forces come into play because of the negatively-charged electrons. The appropriate distance required for Van der Waals attractions differs from atom to atom, based on the size of each electron cloud, and is referred to as the Van der Waals radius. The dots around atoms in this and other displays represent Van der Waals radii.Van der Waals attractions, although transient and weak, can provide an important component of protein structure because of their sheer number. Most atoms of a protein are packed sufficiently close to others to be involved in transient Van der Waals attractions.

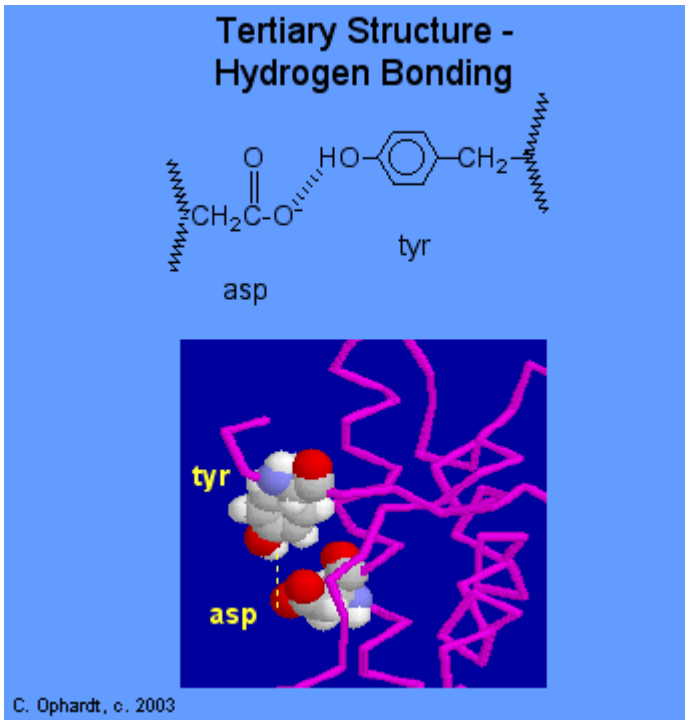
5. Hydrophobic interactions

The hydrophobic interactions of non-polar side chains are believed to contribute significantly to the stabilizing of the tertiary structures in proteins. This interaction is really just an application of the solubility rule that "likes dissolve likes". The non-polar groups mutually repel water and other polar groups and results in a net attraction of the non-polar groups for each other. Hydrocarbon alkyl groups on ala, val, leu, and ile interact in this way. In addition, benzene (aromatic) rings on phe and tyr can "stack" together. In many cases this results in the non-polar

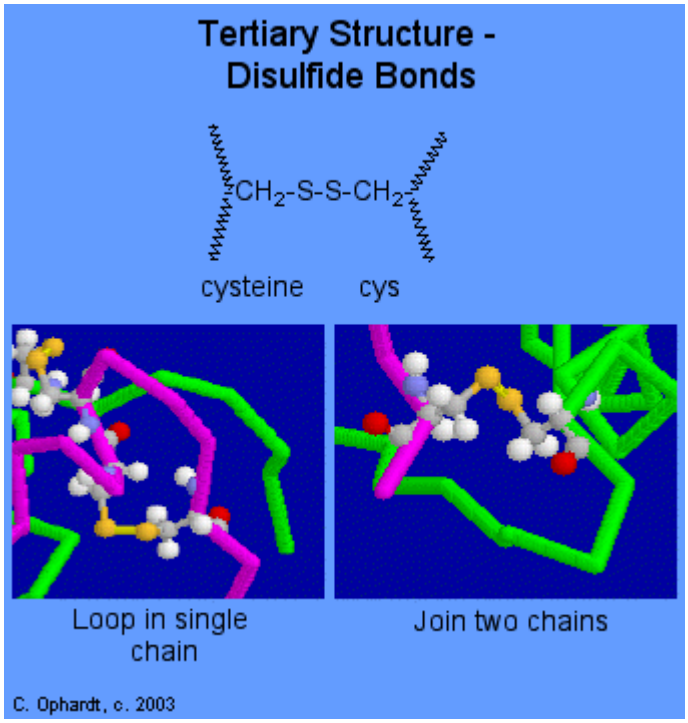
Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar

side chains of amino acids being on the inside of a globular protein, while the outside of the proteins contains mainly polar groups.

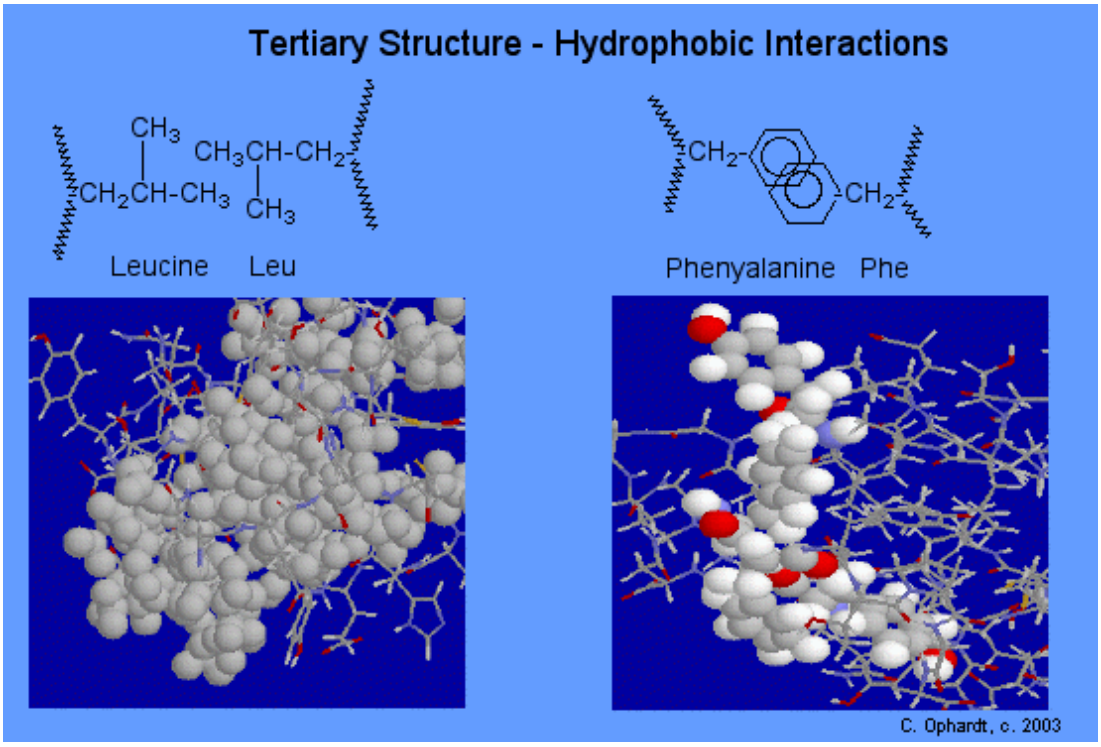




Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar



Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar



Dr. Adfar Yousuf
Assistant professor Dept. of Biotechnology
SP College Srinagar