Austin Publishing Group

Protein Structure and Function

Rizwan Hasan Khan, Mohammad Khursheed Siddiqi and Parveen Salahuddin*
Interdisciplinary Biotechnology Unit, Aligarh Muslim University, India

***Corresponding author:** Parveen Salahuddin, Interdisciplinary Biotechnology Unit, A.M.U.Aligarh-202002, India

Published Date: March 27, 2017

ABSTRACT

Proteins are made up of hundreds or thousands of smaller units known as amino acids. There are 20 different kinds of amino acids that are linked together by peptide bond to make a protein molecule. The sequence of amino acids determines each protein's unique 3-dimensional structure and its specific function such as catalysis of biochemical reactions, mechanical support and immune protection, movement, transport of ligand, transmits nerve impulses, and control growth and differentiation. The amino acids of a protein have been classified as polar, nonpolar, hydrophilic, hydrophobic, acidic, basic, aliphatic and aromatic. A protein can acquire a regular secondary structure for instance α-helix, β-pleated sheet, β-turn, and coils. The secondary structures are further folded leading to the formation of higher order structure namely, tertiary structure which is stabilized mainly by hydrogen-bonding, electrostatic interactions, hydrophobic interactions and van der Waals interactions. Proteins are not entirely rigid molecules. They undergo conformational changes upon ligand binding. For instance, myoglobin and hemoglobin undergo conformational changes when they bind oxygen, therefore they can transport oxygen to the different tissues and lungs respectively. Thus, these proteins show structure-function relationship. In the views of above, in this chapter we have described briefly from amino acids the building block of proteins to quaternary structure of proteins. Further, we have explained structure-function relationship with special reference to myoglobin, and hemoglobin.

Keywords: Amino acids; Peptide; Secondary structures; Tertiary structure; Quaternary structure; Hemoglobin; myoglobin; Hill equation; Bohr effect

OBJECTIVES

To describe amino acids

To classify amino acids

To explain acid-base character of amino acid

Amino acids

Amino acids are the building block of proteins. Amino acids are important organic compounds that contain amine ($-NH_2$) and Carboxyl (-COOH) functional groups, along with a side-chain (R group) that is specific for each amino acid (Figure 1). Twenty different amino acids are commonly found in proteins. All of these 20 common amino acids are α-amino acids except proline and their general structure is shown below. They have a carboxyl group and amino group which are covalently bonded to a α-carbon atom. They differ from each other in their side chain R groups. Since, the remaining structure are same therefore properties of these amino acids are primarily determined by the side chain groups. The nature of these side chain maybe polar, nonpolar (aliphatic), hydrophilic, hydrophobic, acidic, basic and aromatic. These amino acids have been abbreviated using either three letter word or one letter word (Table 1).

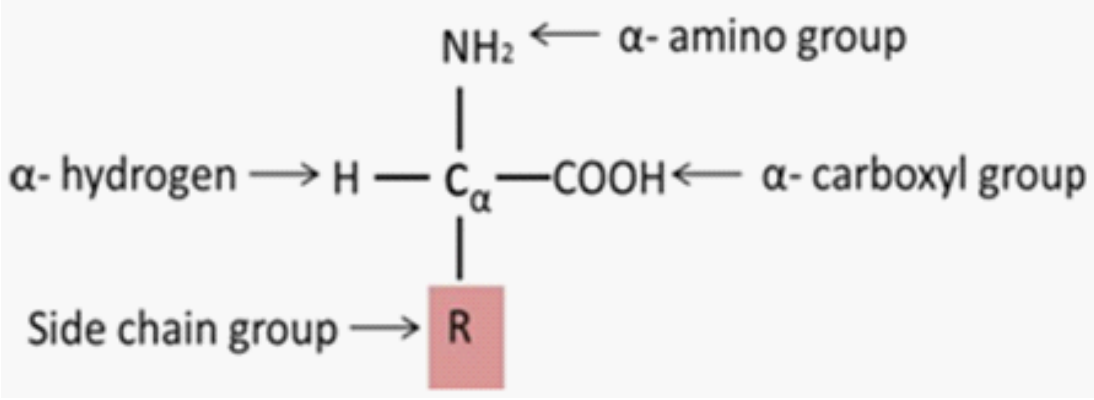


Figure 1: Structure of amino acid containing R side chain.

Table 1: Three-letter and one letter abbreviations.

Amino acid	Three -letter abbreviation	One letter symbol
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asparagine or aspartic acid	Asx	B
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glutamine or Glutamic acid	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Classification

The 20 amino acids have been classified using different criteria by different scientists. For instance, they have been classified as polar, nonpolar, hydrophilic, hydrophobic, acidic, basic, aliphatic and aromatic. Here, we have classified all of these 20 common set of amino acids into six distinct classes.

Nonpolar (Aliphatic) amino acids

The R side chain in this class of amino acids including alanine, valine, leucine and isoleucine (Figure 2) are hydrophobic in nature therefore they stabilize the protein structure through hydrophobic interactions. Glycine is also classified as nonpolar amino acids, but it has very small side chain. Therefore, it does not contribute to hydrophobic interactions. Glycine has the simplest structure. The side chain of proline has a distinctive cyclic structure which is an imino group held in a rigid conformation, therefore it reduces the structural flexibility of particularly that regions of polypeptide chain where it occurs.

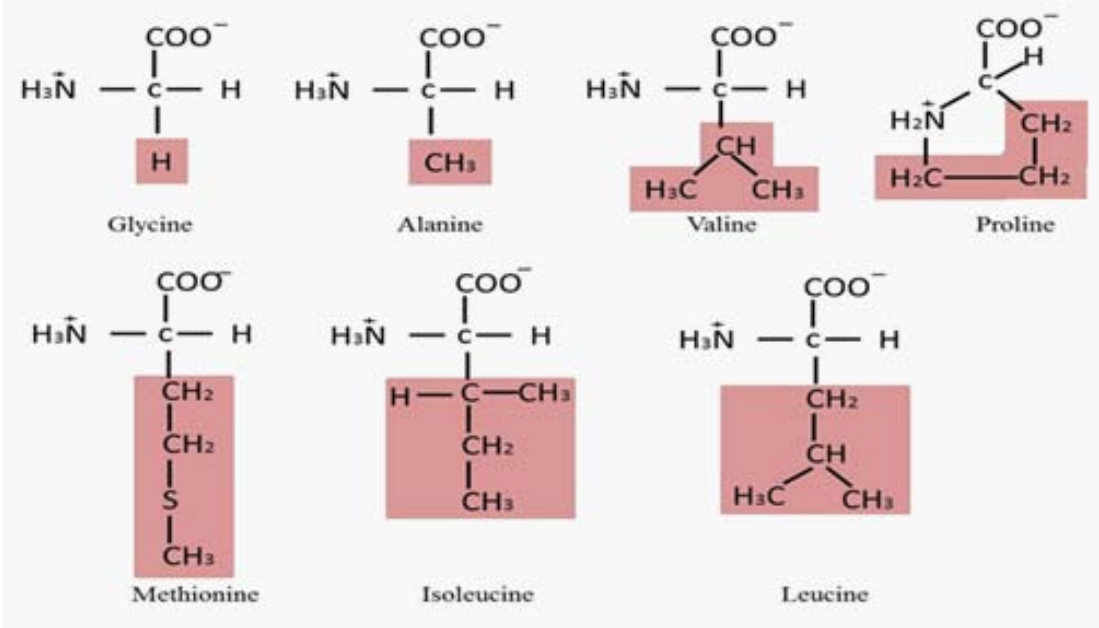


Figure 2: Structure of Nonpolar (aliphatic) amino acids.

Aromatic amino acids (Phenylalanine, Tyrosine and Tryptophan)

The side chain of aromatic amino acids contains an aromatic ring (Figure 3) which are relatively nonpolar (hydrophobic) in nature. These amino acids can participate in hydrophobic interaction. Tyrosine and tryptophan are much more polar than phenylalanine owing to their hydroxyl and nitrogen indole ring respectively. These amino acids show light absorption in the ultraviolet range due to the presence of conjugated double bond-single bond system.

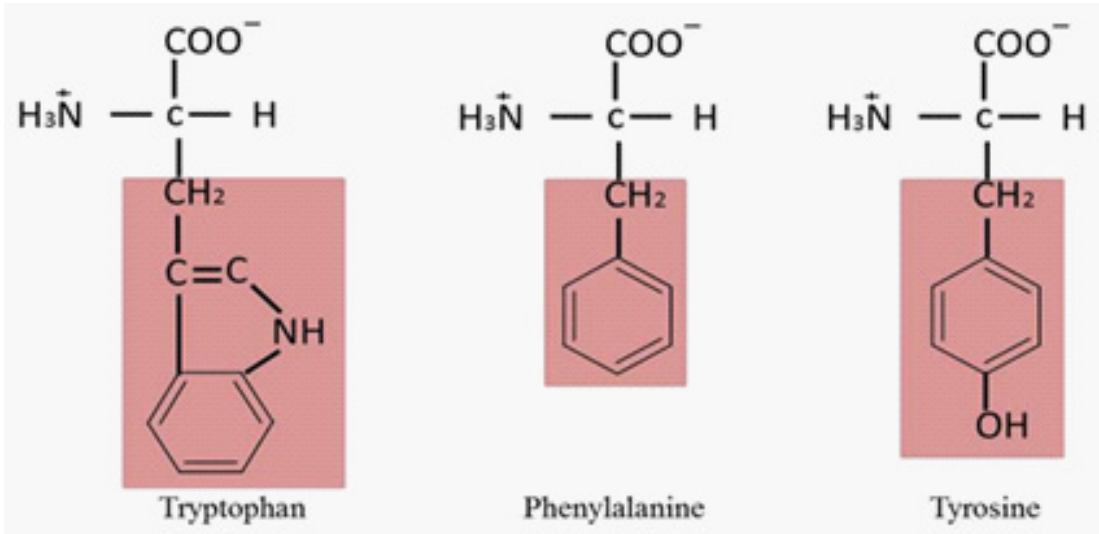


Figure 3: Structure of aromatic amino acids (Tryptophan Phenylalanine and Tyrosine).

Polar, uncharged amino acids

This class of amino acids includes serine, threonine, cysteine, asparagine and glutamine (Figure 4). The R group of these amino acids are more soluble in water or more hydrophilic than those of nonpolar amino acids because they contain functional groups (OH,SH, $CONH_2$) that form hydrogen bonds with water. The polarity of serine and threonine is contributed by their hydroxyl groups, and that of cysteine and tryptophan by sulfhydryl and indole ring respectively which is weakly hydrogen bonded with oxygen and nitrogen respectively. Furthermore, polarity of asparagine and glutamine is contributed by their amide group.

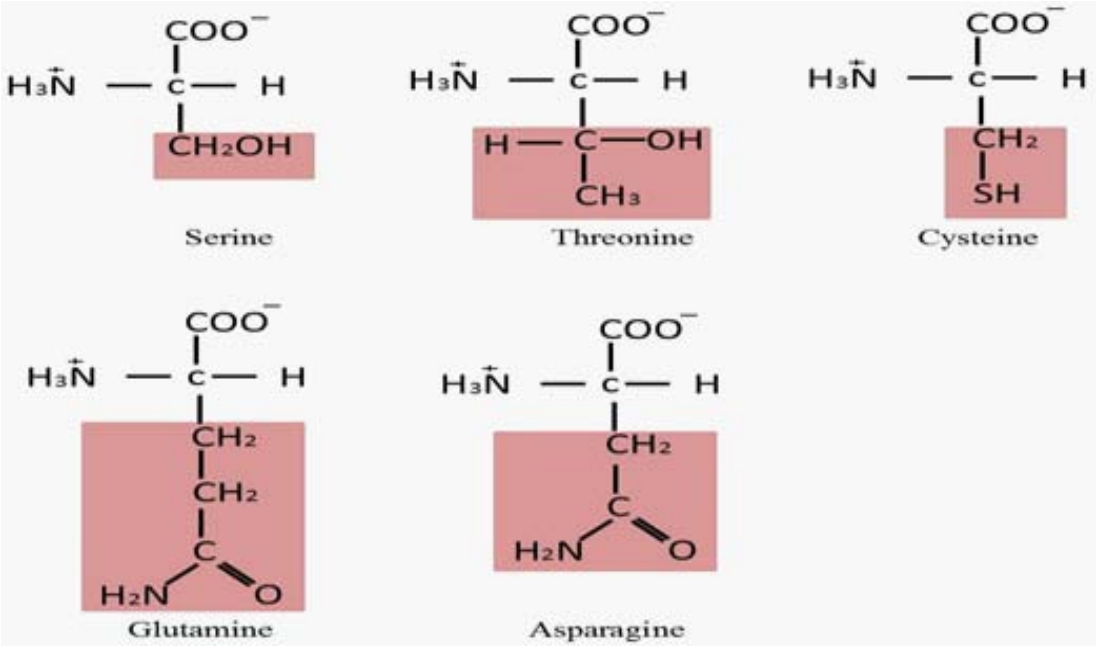


Figure 4: Structure of polar, uncharged amino acids.

Figure 5: Structure of acidic amino acids.

Acidic amino acids

These amino acids contain two carboxyl groups, one α- carboxyl and other β- or γ-carboxyl group (Figure 5). Since they contain two acidic groups (one α-carboxyl group + one β or γ-carboxyl group) and one basic group (α-amino group), the net charge of these amino acids is therefore acidic and they are negatively charged at physiological pH.

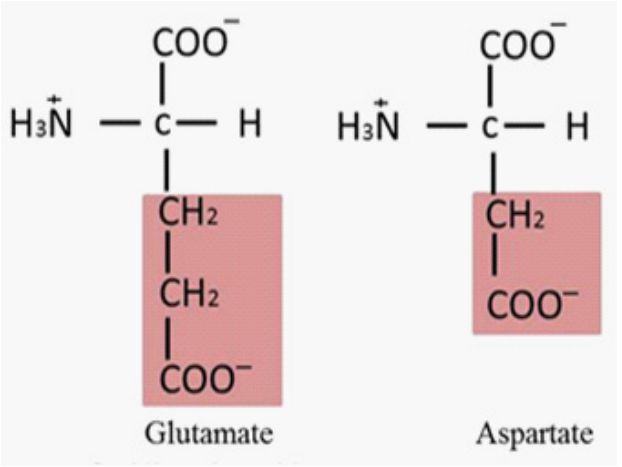


Figure 5: Structure of acidic amino acids.

Basic amino acids

The basic amino acid contains an α-amino group and the side chain contains second amino/imino group (imidazole, ε-amino or guanidine group). These amino acids are histidine, lysine and arginine (Figure 6). Since these amino acids contain two basic groups one acidic group (α-carboxyl group), therefore the net behavior of these amino acids is basic and they are positively charged at physiological pH.

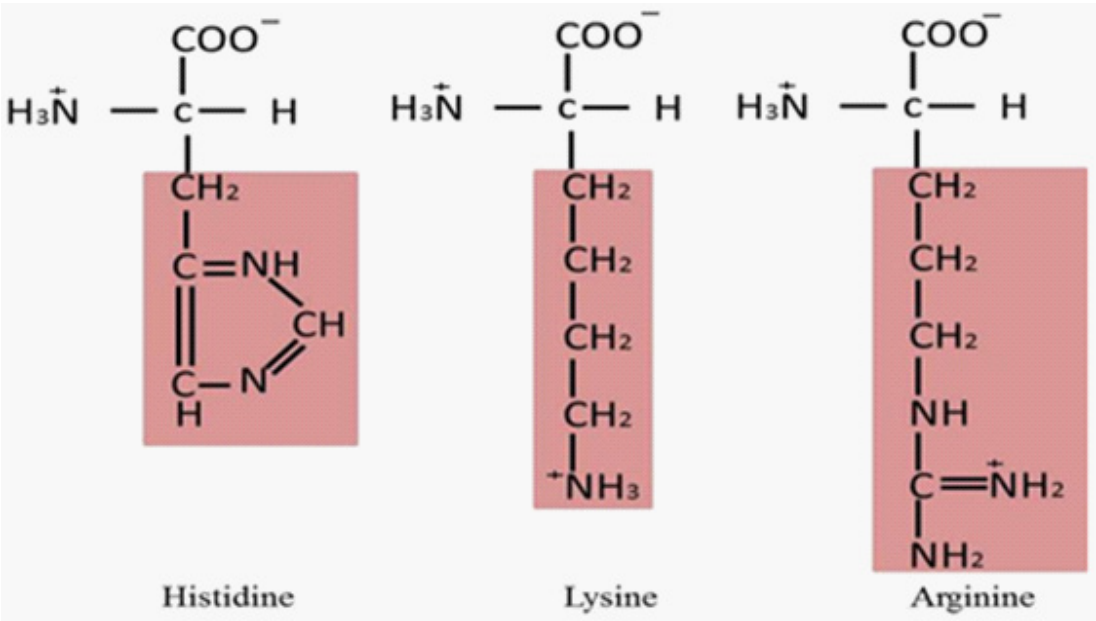


Figure 6: Structure of basic amino acids.

Acid-base character of amino acids

For explaining the acid-base character, let us consider neutral (aliphatic) amino acid alanine which is nonpolar that was discovered in 1923. Its carboxyl group can be deprotonated (Figure 7).

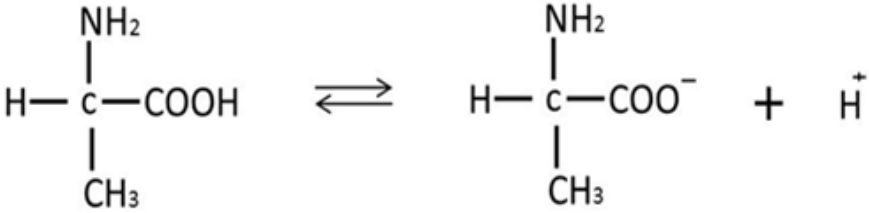


Figure 7: Carboxyl group of amino acid donates a proton.

Similarly, when amino acid accepts a proton (due to the presence of a basic amino group), the amino acid acquires a positive charge (Figure 8).

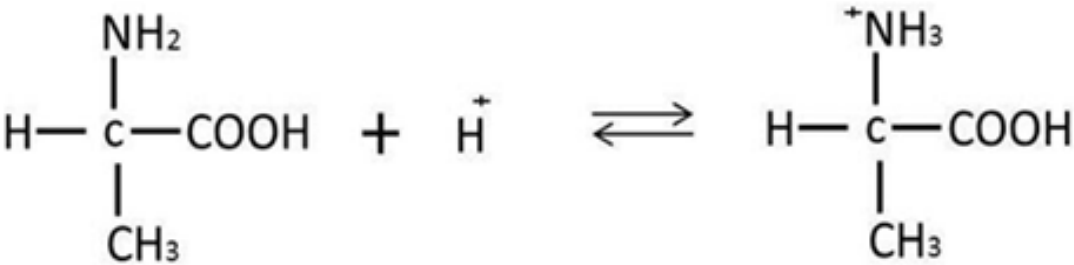


Figure 8: Amino group of amino acid accepts proton.

If we consider the acidity of amino acid, it releases proton which will be taken up by the solvent, water or by the basic amino group available on the amino acid. Since amino group is more basic, it takes the proton donated by carboxyl group. As a result carboxyl group acquires a negative charge whereas a positive charge develops on the amino group. Since this form of amino acid has both positive and negative charges, therefore the net charge of the amino acid is zero. This state of amino acid is known as zwitterionic state (Figure 9).

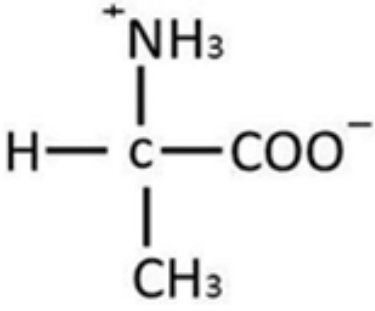


Figure 9: Zwitterionic state of alanine.

In zwitterions form, the carboxylate group acts as a base and the protonated amino group acts as an acid as shown below:

$-COO^- + H^+ \rightarrow -COOH$

$^+-NH_3 + OH \rightarrow -NH_2 + H_2O$

Since this type of amino acid is capable of acting as both acid and base, this implies that amino acid can act as buffer.

Evidences supporting the zwitterionic form of amino acids

The zwitterionic form of neutral (aliphatic) amino acids is supported by their melting points which is higher than organic molecules but comparable to inorganic lattices. This occurs only when amino acids have both positive, negative charges which can interact with several molecules, thus in principle requiring higher energy to break them. Most amino acids are soluble in water which also reflects the zwitterionic form of these amino acids

Before considering the ionization behavior of amino acids, we shall first consider the ionization of weak acid e.g. acetic acid.

$CH_3COOH \leftrightarrow CH_3COO^- + H^+$

The dissociation constant of weak acid K_a can be written as:

$$K_a = \frac{[CH_3COO^-]}{[CH_3COOH]} + [H^+]$$

Taking the logrithimic on both side of the above equation we have:

$$\log K_a = \log \frac{[CH_3COO^-]}{[CH_3COOH]} + \log[H^+]$$

Multiplying both sides by -1 we get:

$$-\log K_a = -\log \frac{\left[CH_3COO^-\right]}{\left[CH_3COOH\right]} + (-\log[H^+])$$

Since $-\log K_a = pK_a$ and $-\log [H^+] = pH$, the above equation can be written as:

$$pK_a = -\log \frac{\left[CH_3COO^-\right]}{\left[CH_3COOH\right]} + pH$$

$$pH = pK_a + \log \frac{[CH_3COO^-]}{[CH_3COOH]}$$

The above equation is known as the Henderson-Hasselbalch equation.

If dissociation of weak acids is represented as:

$HA \leftrightarrow H^+ + A^-$

Then the above Henderson-Haselbalch equation can be written as:

$$pH = pK_a + \log \frac{[A^-]}{[HA]}$$

If we represent the degree of dissociation of a weak acid as α, then Henderson-Haselbalch equation can be written as:

$$pH = pK_a + \log \frac{\alpha}{1-\alpha}$$

This equation tells us about dissociation of a particular weak acid and its dependence on pH. In order to define the pK_a, let us assume the value 'α' is equal to 0.5 which means that at equilibrium, the amino acid is 50% dissociated.

If we substitute the value of 'α' as 0.5, then the above expression becomes:

$$pH = pK_a + \log \frac{0.5}{1.0-0.5}$$

$pH = pK_a + \log 1.0$

$pH = pK_a$

Therefore, pKa can be defined as the pH at which 50% of the acid is dissociated. Alternatively, if we substitute pH=pK, in the Henderson-Hasselbalch equation and the value of 'α, will be 0.5.

Similarly, we can determine the value of 'α' if:

$pH = pK_a+1$; $pH = pK_a+2$; $pH = pK_a-1$; $pH = pK_a-2$

From the Table 2, it is evident that the dissociation of an ionisable group can occur within the pH range of 2 units higher and lower than the pKa value of that group. This means that pH of the medium dictates the charge on the ionisable groups. The deprotonated form of carboxyl group [COO^-] and the protonated form of the amino group [$^+NH_3$] carries one negative and one positive charge respectively. If we consider the dissociation of both the ionisable groups in the amino acid, we can write the different states of alanine at three different pH values (Figure 10).

Table 2: The values of α within the pH range of 2 units higher and lower than the pK_a value of that group.

pH	A	Comments
pK_a + 1	0.91	Acid is 91% dissociated
pK_a + 2	0.99	Acid is 99% dissociated
pK_a - 1	0.09	Acid is 09% dissociated
pK_a - 2	0.01	Acid is 01% dissociated

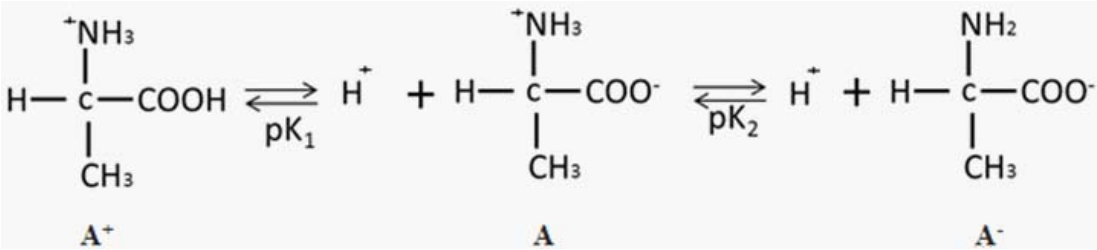


Figure10: Ionization behavior of alanine.

Here, under the acidic conditions, both of these ionisable groups will be protonated, therefore the net charge on the amino acid will be positive and it will move towards the cathode under the influence of an electric field. Raising the pH of the solution by the addition of OH^-, will remove the proton from an α-carboxyl group which has low pK value. Since the carboxyl group has a low pK value, it will first undergo deprotonation resulting in the creation of negative charge on this group and thus the net charge on the amino acid will become zero. This state [A] will occur at neutral pH and the amino acid will not move under the influence of an electric field. Increasing the pH further towards the alkaline range will result in the deprotonation of amino group. Therefore, the net charge on the amino acid will be negative and it will move towards the anode in an electric field.

If we consider the above equilibria showing the dissociation of alanine, we can see that the neutral (aliphatic) amino acid, in its protonated form acts as a diprotic acid, or in other words it can release two protons upon complete dissociation. If we assume K_1 and K_2 as an equilibrium constants for steps (i) and (ii) respectively, then we have:

$$K_1 = \frac{[H^+][A]}{[A^+]} \text{ and } K_2 = \frac{[H^+][A^-]}{[A]}$$

Multiply K_1 with K_2, we can write as:

$$K_1 * K_2 = \frac{[H^+][A]}{[A^+]} * \frac{[H^+][A^-]}{[A]}$$

$$K_1 * K_2 = \frac{[A^-]}{[A^+]} * [H^+]^2$$

Consider a pH at which amino acid exists in A state. This pH is known as the isoelectric point or more simply pI. At pI, the net charge on the amino acid is zero or number of positive charges is equal to the number of negative charges (or $[A^-]=[A^+]$. The A form of the amino acid is known as an isoelectric form. Thus, we can write equation as:

$$K_1 * K_2 = [H+]^2$$

Taking the logarithm on both sides of the above equation, we have:

$\log K_1 + \log K_2 = 2\log [H^+]$

Multiplying by -1 , we get

$(-\log K_1) + (-\log K_2) = 2(-\log [H^+])$

$pK_1 + pK_2 = 2pH$

Since the pH has been assumed to be equal to pI as mentioned above, we can write the equation as:

$pK_1 + pK_2 = 2pI$

or

$$pI = \frac{pK_1 + pK_2}{2}$$

From the above equation, one can calculate the pI of any amino acid provided that their pK values are known. However, in situation where more than two steps are involved in complete dissociation of amino acid (as occur in the dissociation of acidic and basic amino acids). This will yield two pK values that occur on either sides of the isoelectric form [A] of amino acid. The importance of determining the isoelectric point of an amino acid is to decide the nature of amino acids whether they are acidic, basic or neutral. This can be done in the following manners

If pI < 7.0, the amino acid is acidic in nature.

pI = 7.0, the amino acid is neutral in nature.

pI > 7.0, the amino acid is basic in nature.

Furthermore, after computing the pI, one can determine the net charge present on the amino acid at a given pH.

Titration curve of alanine

Since all amino acids contain at least two ionisable groups (an α-amino and α-carboxyl groups) their titration with a base produce a titration curve which consist of two titration curves (the mid points are shown below) each representing the ionization of a single ionisable group. As shown in the titration curve (Figure 11), at the beginning of the titration curve (point 'X'), only the A^+ form of amino acid exists. Addition of NaOH to the solution results in the conversion of the A^+ form into A. At point 'Y', the amino acid exists in the A form. The curve from point 'X' to point'Y' of a monoprotic acid (acetic acid) shows the complete ionization of one ionisable group (carboxyl group) of the amino acid. At the midpoint 'XY' of this curve, both A^+, A species of the amino acid are present in equal amounts (i.e. 50% each). The corresponding pH value (2.3) at this point represents the pK value of the ionisable group (carboxyl group) which is the pH at which 50% of the amino acid (ionisable group) is dissociated. At point 'Y', if we continue adding OH^- ions, the A species is converted into A^- form and at the point 'Z', A^- form of amino acid exists. The remaining half of the titration curve (from point 'Y' to point'Z') represents the titration of the second ionisable group (amino group) of the amino acid. At the midpoint'YZ' both of the ionic species (A and A^-) are present in equal amounts. Hence, the corresponding pH value 9.7 represents the pK value of the second ionisable group (amino group). At point 'Y', the corresponding pH value is 6.0 which is the isoelectric point (pI) of the amino acid, at this point amino acid present exists in A form (isoelectric form). Furthermore, at points 'XY' and 'YZ', addition of a small amount of base to amino acid solution does not significantly affect pH of the solution. This implies that around these pH values, the amino acid exhibits a buffering action (which is defined as a resistance to a change in pH by the produced by the addition of small amount of acid or base).

From the titration curves of all amino acids, the pK values of the different ionisable groups present in the amino acids (the α-carboxyl, α-amino and side chain groups if ionisable) have been determined and these values are depicted in the Table 3 given below. Analysis of the curve shows that the side chain which does not ionize will behave similar to alanine. In situation where side chain also ionizes, the whole titration curve will exhibit three curves (or midpoints) for monoprotic acid where each curve shows the ionization of a single ionisable group.

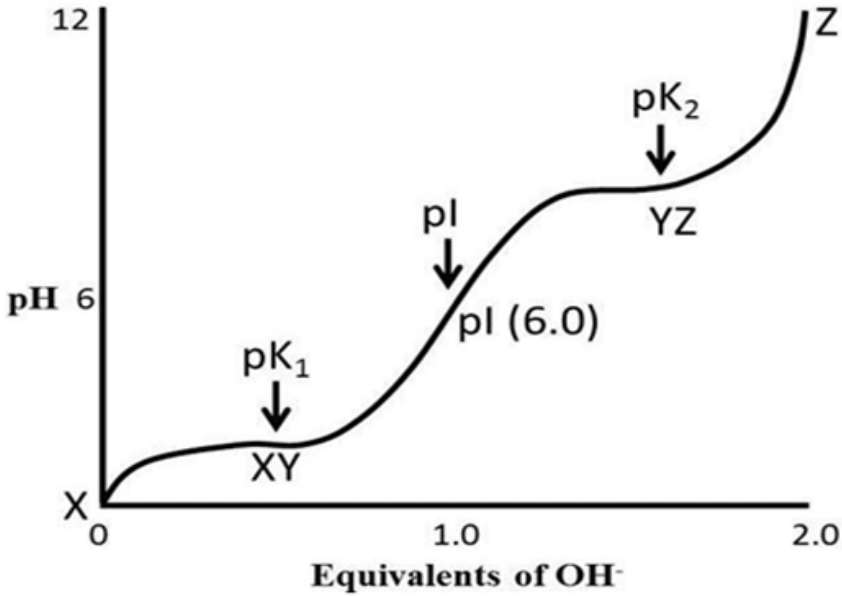


Figure 11: Titration curve of neutral amino acid alanine [1].

Table 3: Ionization of different ionizable groups of amino acids.

Ionisable group	pK
α-Carboxyl group	2.3
α-Amino group	9.7
β-Carboxyl group [D]	3.9
γ-Carboxyl group [E]	4.3
Imidazole group [H]	6.0
E-Amino group [K]	10.5
Guanidino group [R]	12.5
Sulphydryl group [C]	8.3
Hydroxyl group [Y]	10.9

Learning outcomes

Proteins play an important role in various biological processes.

Amino acids are building block of protein.

Amino acids have been classified as polar, nonpolar, hydrophilic, hydrophobic, acidic, basic, aliphatic and aromatic.

Amino acid shows acid-base character.

From the titration curve one can determine total number of ionisable groups, pK values of different ionisable groups, net charge and buffering action of amino acids.

Objectives

Overview of peptides

To explain acid-base properties of peptides

Overview of peptides

Peptides are short polymer of amino acid that is linked by peptide bonds. The peptide bond is formed when the α-carboxyl group of one amino acid reacts with the α-amino group of another amino acid with the release of water molecule to yield dipeptide (Figure 12). These polypeptides vary in size from very small to very large, comprising of two or three to thousands of amino acid residues. A peptide bond is defined as a covalent bond between carbonyl group of one amino acid and imino group of another amino acid. This bond is also known as amide bond. In a polypeptide, the free α-amino group of amino acid is known as amino-terminal (or N-terminal); whereas at the other end has free carboxyl group which is known as carboxyl-terminal (C- terminal) residue. By convention the amino terminal end is placed on the left and the carboxyl terminal end to the right. Thus, the sequence is read from left to right.

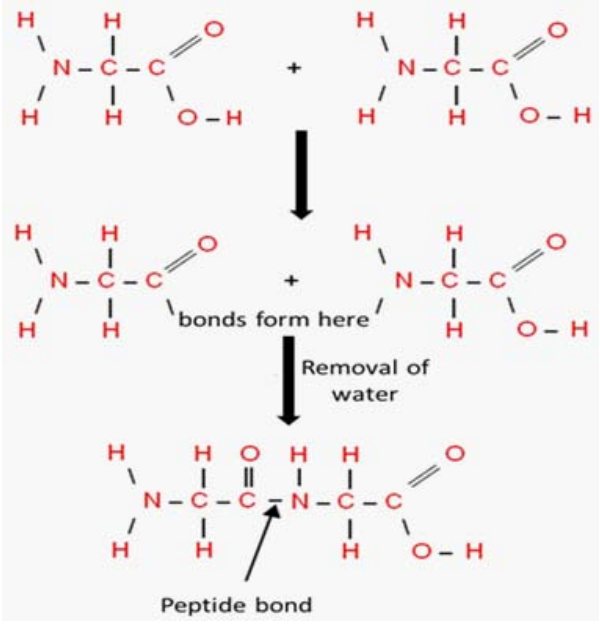


Figure 12: Formation of peptide bond with removal of water molecule.

Acid-base properties of peptides

Peptides contain only one free amino acid and one free α-carboxyl group at the other end of the chain. These groups ionize like free amino acids but their pK value differs. For instance, the pK value of the α-carboxyl group increases to a range between 3 and 4 (i.e. less acidic than α-carboxyl group of free amino acids) whereas the pK of the α-amino group lies in the range

from 7.5 to 8.5 (i.e. it is a weaker base as compared to α-amino group of free amino acids) as the chain length of peptides increases. This occurs because the protonated amino group is far away from protonated free carboxyl group in polypeptide therefore it experiences less electrostatic repulsion among themselves contrary to free α-amino and α-carboxyl group of amino acid. The pK value also depends upon the nature of neighboring ionisable side chain. Variation in the pK values can be also be attributed to other factors including the inductive effect of various groups located in the vicinity of these ionisable group.

Biologically active peptides

Naturally occurring peptides vary in length from two to many thousands of amino acids. These peptides functions as hormones, neurotransmitter, opiate, antibiotics etc. Here we describe the structure and function of some of the biologically active peptides.

Gramicidin

Gramicidin is a circular or cyclic decapeptide containing ten amino acid residues that are linked to each other through peptide bonds. Besides containing amino acids, it has two non-amino acid residues of D-phenylalanine and two residues of L-ornithine (a non-protein amino acid). The structure of gramicidin is depicted below (Figure 13). Gramicidin is an antibiotic.

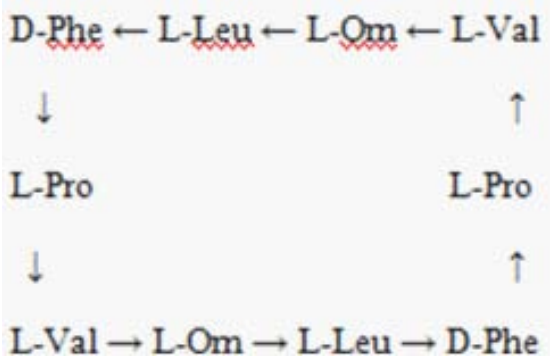


Figure 13: Structure of Gramicidin.

Vasopressin and oxytocin

These are peptide hormones made up of nine amino acid residues containing one disulfide bond. The carboxyl terminal residue is an amide of glycine (Figure 14 and 15). Seven out of nine residues in these hormones are identical. The structure of these peptide hormones are shown below:

Both of these hormones are secreted by the pituitary gland. Oxytocin is involved in the contraction of smooth muscles (uterine contraction during child birth) whereas vasopressin (antidiuretic hormone) helps in the retention of water from urine by kidneys.

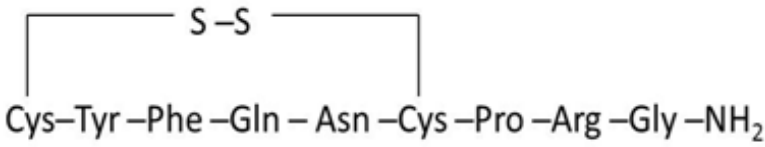


Figure 14: Structure of Oxytocin.

Figure 15: Structure of vasopressin.

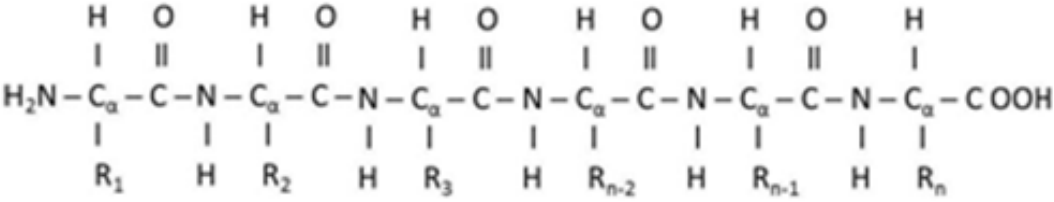


Figure 16: A primary structure of protein containing 'n' number of amino acid residues.

Learning outcomes

Peptides are short polymers of amino acid.

Peptide shows acid-base properties.

Objectives

Overview of proteins

To describe primary structure

To explain conformation of peptide bond

To describe secondary structure

To describe tertiary structure

To describe quaternary structure

Overview

Proteins are large, complex molecules that play many critical roles in the body. They do most of the work in cells and are required for the structure, function, and regulation of the body's tissues and organs. The Swedish chemists J. J. Berzelius in 1938 coined the name 'protein' from the Greek word proteios which means "of the first rank".

Proteins are the most versatile macromolecules in living systems and play important functions in essentially all biological processes. They function as catalysts, they transport and store other molecules such as oxygen, they provide mechanical support and immune protection, they generate movement, they transmit nerve impulses, and they control growth and differentiation. All proteins, whether from the most ancient lines of bacteria or from the most complex forms of life are all made up of from the same universal set of 20 amino acids that are covalently linked to form linear sequence. Each of this amino acid has a side chain that possesses distinctive chemical properties. Cells can produce proteins with distinctly different properties in many combinations and sequences. The sequence of amino acids determines protein's unique 3-dimensional structure and its specific function.

The proteins are of two types: simple proteins and conjugated proteins. Simple proteins contain only amino acid residues whereas conjugated proteins contain amino acids as well as non-amino acid moiety. The non-amino acid moiety of the conjugated protein is called 'prosthetic group'. Conjugated proteins are classified on the basis of chemical nature of their prosthetic group. For instance, glycoproteins contain carbohydrate therefore they are named as glycoprotein. Similarly, lipoproteins and metallo-proteins are so named because they contain lipid, and a specific metal prosthetic group. Some proteins contain more than one prosthetic group. The prosthetic group participates in protein's biological function.

Primary structure

The primary structure of a protein is defined as a linear sequence of amino acids and the location of disulfide (-S-S-) bonds, if any (Figure 16). Working on ribonuclease (an enzyme that catalyzes the hydrolysis of RNA) C.B. Anfinsen in 1957 [1] showed for the first time that amino acid sequence of a protein dictates its three dimensional structure which is biologically active conformation and marginally more stable than unfolded state. The phenomenon of the formation of native biologically functional structure of proteins from the linear sequence of amino acids (primary structure) to three-dimensional structure is known as protein folding. The length of the primary structure varies from protein to protein. It may be as short as 51 amino acid residues for instance insulin or large as 4536 amino acid residues (apolipoprotein B-100).

The figure above shows the primary structure of a protein which is comprised of 'n' number of amino acid residues containing side chains, R1, R2, R3, Rn_{-2}, Rn_{-1} and R_n which differ in different proteins. The backbone or main chain repeat regularly along the length of polypeptide chain and remains same in different proteins. Since proteins are big, it is difficult to write the primary structure or sequence in the above form. Therefore, the primary structures can be written as amino acid residues in the linear form using either three or one letter abbreviations, each separated by dash (-), if amino acids are separated by comma, this signifies amino acid composition (Figure 17).

Val, Leu, Phe, Tyr, Ala, Arg, Cys, Ile, Asp, His (amino acid composition)

Ala-Tyr-Ile -Arg-His-Val-Cys-Asp-Leu-Phe (amino acid sequence)

Figure 17: Amino acid sequence and amino acid composition.

Thus, the analysis of primary structure of a protein from different species helps in the understanding of folding of that protein. Secondly, the knowledge of the primary structure of proteins aids in the understanding of their function, evolution and diagnosis of a diseases.

Conformation of peptide bond

In the late 1930 Linus Pauling and Robert Corey analyzed carefully the peptide structure. 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 the crystal of amino acids and dipeptide and tripeptide showed that the peptide C-N bond is somewhat shorter than the C-N bond in a simple amine and that the atoms associated with the peptide are co-planar. This also indicated by resonance or partial sharing of two pairs of electron between the carbonyl oxygen and the amide nitrogen. The oxygen has a partial negative charge and the nitrogen a partial positive charge (Figure 18), this result in small electric dipole. The six atoms of the peptide group lie in a single plane, with the oxygen atom of the carbonyl group trans to hydrogen atom of the amide nitrogen. From these findings Pauling and Corey concluded that the peptide (C-N) bond because of their partial double bond character cannot rotate freely. Allowing the rotations around the C_α-C and N-C_α bonds, a peptide group can make two dihedral angles namely ψ (psi) and φ (phi) (Figure 19).

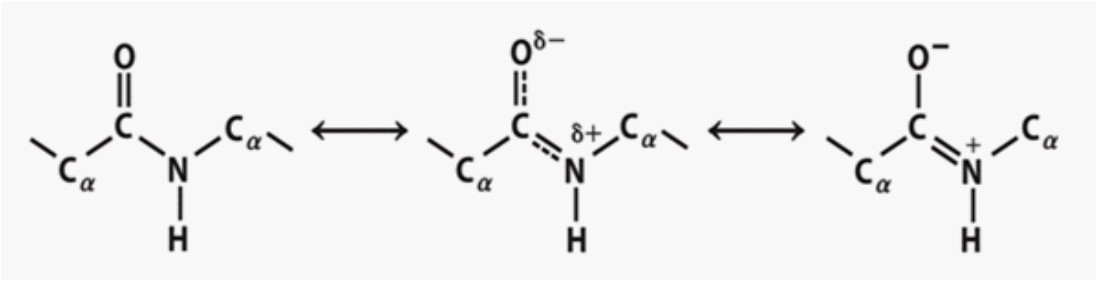


Figure 18: Resonance in the peptide unit.

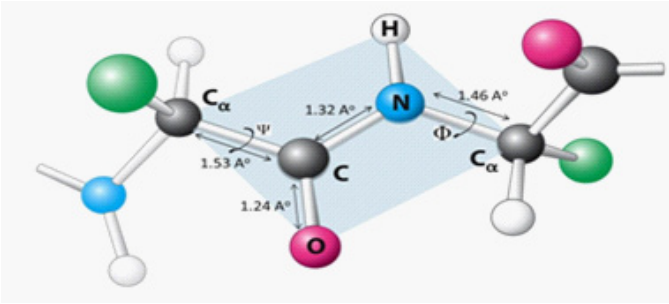


Figure 19: The planar peptide group showing dihedral angles and bond length.

The rotation angle clockwise around the C_α–C bond is known as ψ (psi) whereas angle of rotation around the N-C_α bond is designated as φ (phi). Since rotations around the C_α–C and N–C_α bonds are possible, the allocation of particular values for the ψ and φ angles for each residue in the chain can dictate a particular type of main chain conformation. In fact any value ranging from 0 and 360° can be assigned for the ψ and φ angles. However, G.N. Ramachandran found that certain conditions have to be fulfilled for assigning the values to these dihedral angles. Firstly, assignment of the rotation values should not allow any serious clash between two groups or atoms and it should not produce any steric hindrance. Bond lengths and bond angles should be retained (no internal torsion should occur) and it should maximize the number of non-covalent interactions (mainly hydrogen bonding between the carbonyl oxygen and the imino hydrogen) to stabilize the conformation. As a result of this, certain rotations are completely forbidden and a polypeptide chain can only acquire only a few possible conformations. Ramachandran found that each particular type of secondary structure has a particular value of ψ and φ angles, which can be seen in the energy contour map. The plot between ψ and φ angles is known as Ramachandran plot (Figure 20) [2].

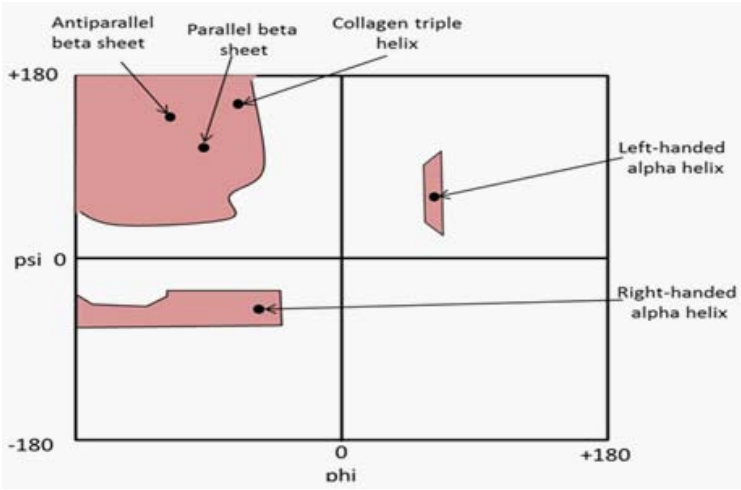


Figure 20: Ramachandran plot for L-alanine residues.

A polypeptide segment can acquire a regular secondary structure only when all the dihederal angles (ψ and φ) within that polypeptide segment are equal. Thus, a polypeptide chain can be folded by allowing either helical twisting, resulting in helical structures or by packing one part of the polypeptide onto another producing β-structures. The hydrogen bond is the main stabilizing force for the secondary structure.

Secondary structures

The secondary folding of the primary structure is known as secondary structure. There are various types of secondary structures mentioned below.

Helical structures

Different types of helical structures can be formed as a result of hydrogen bonding between carbonyl oxygen and imino hydrogen of various residues (Figure 21). If the carbonyl oxygen of the nth residue is hydrogen bonded with the imino hydrogen of the (n+ 3)rd residue the resultant helix is called a 3_{10}-helix.This helix contains 3 residue per turn of the helix with a helical rise per residue of 1.93–2.0 A, and a helical pitch of 5.8–6 A respectively. The number 10 refers to the total number of imino hydrogen per turn of the helix.

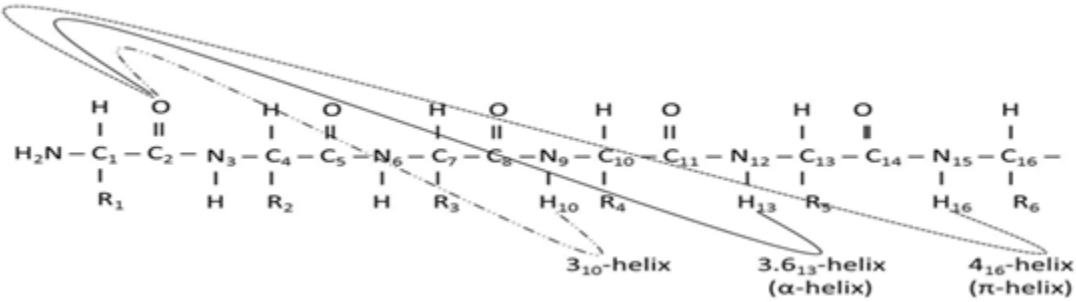


Figure 21: Different helical structures are shown above [3].

Similarly, if the carbonyl oxygen of the nth residue forms a hydrogen bond with the imino hydrogen of (n+4) th residue, the 3.6_{13}-helix is formed, which is also known as the α-helix. It was the first helical structure that was elucidated by Pauling and Corey, therefore its name was given as α-helix. It is right handed twist that bears similarity to a right handed spiral staircase in which the amino acid residues form steps such that the distance between the two amino acid residues (step) is 1.5 A and one turn of the helix contains contain 3.6 amino acid residues therefore the helical pitch is 5.4 A. This helix contains 13 imino hydrogens per turn of the helix which stabilize the helical structure. The ψ and φ values are -47° and -57° respectively. Though α-helix is formed as a result of main chain (CO)-main chain (NH) interaction, side chains also influence the formation of α-helix due to steric hindrance. Some amino acids like Met, Glu, and Leu have the high propensity to form α-helix where as other including Pro, Gly, and Tyr are helix breaker. Some proteins that are rich in α-helix are hemoglobin, myoglobin, myosin, keratins etc.

The π helix is an extremely rare secondary structural element in proteins. Although once thought to be rare, short π-helices are found in 15% of known protein. Hydrogen bonding between the carbonyl oxygen of the nth residue and imino hydrogen of (n+5th) residue gives rise to the formation of 4_{16}-helix or π-helix. The π-helix contains 4 amino acid residues per turn of the helix where each amino acid corresponds to an 87° turn in the helix and a translation of 1.15 A along the helical axis with pitch of the helix 4.6 A. The π-helix contains 16 imino hydrogen per turn of the helix. Among all helical structures mentioned above, α-helix structure is thermodynamically most stable and therefore the predominant helical structure found in proteins. All of these helical structures are shown in the above figure.

β -Sheet structure

The β-sheet was the second structure proposed by Pauling and Corey. In this type of secondary structure one segment or β-strand folds on the top of the another segment. These segment/strands are joined to each other through hydrogen bonds involving the carbonyl oxygen of one strand and the imino hydrogen of another strand. These segment of a polypeptide can be oriented either parallel or antiparallel giving rise to the parallel β-sheet structure or antiparallel β-sheet structure (Figure 22). The β-sheet exists as fully extended (stretched) structure. Since large number of segments/strands assemble through hydrogen bonds, it resemble β-pleated sheet. The distance between 2 amino acids in parallel and antiparallel β-sheet is 6.5 A and 7A respectively. The dihedral angles ψ and φ for parallel β-structure have been found to be +113^{o} -119^{o} respectively and +135^{o} and -139^{o} respectively for the antiparallel β-structure. The antiparallel β-structure is more prevalent as it allows far more close packing whereas in the parallel β-structure there is more steric hindrance, therefore it is less tightly packed hence it is less stable. Some β-sheet former amino acids are Val, Ile, Phe and Tyr. Whereas β-sheet breakers are Pro, Gly and Tyr. Proteins that are rich in beta sheet structures are silk fibroin, chymotrypsin, ribonuclease and cytochrome c.

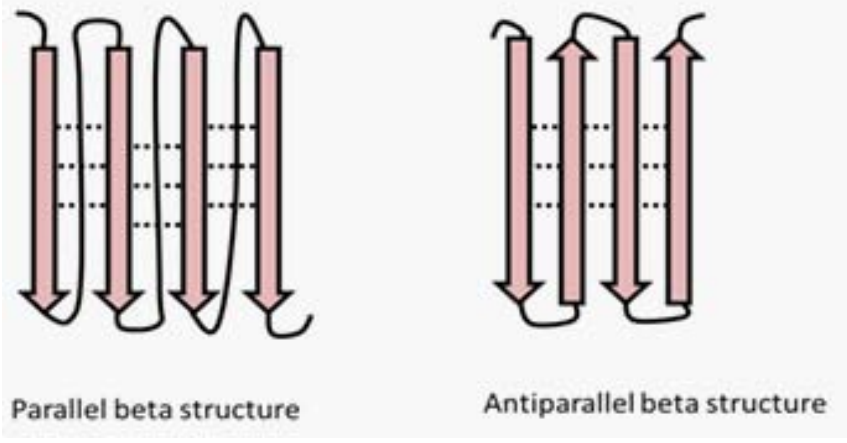


Figure 22: The structure of parallel and antiparallel β-sheet.

β-Turn

β-Turn is also known as β-bend, reverse turn or hairpin structure (Figure 23). For the proper folding and packing into globular shaped structure, it is necessary that polypeptide chain reverse its direction several times. This is achieved only by the presence of another secondary structure known as β-turn or reverse turn. Since the structure resemble hairpin, hence it is named as a hairpin structure. The structure contains 4 residue stabilized by an intramolecular hydrogen bond between the carbonyl of the nth residue and the imino hydrogen of the (n+3) residue. The amino acid residue favoring the formation of the β-turn are Pro, Gly, Ser, Asp and Asn. In all β-turns, the two residues generally present are Pro and Gly and the remaining residues can be either Ser, Asp or Asn. Since Pro cannot participate in hydrogen bond formation it introduces a kink in the polypeptide chain wherever it occurs resulting in the formation of a β-turn structure. To overcome the steric hindrance produced by Pro, a Gly residue is always found in the middle of β-turn. This structure is usually found at the surface of globular proteins and often connects β-sheet. There are three types of β-turns (type I, type II and Type III) which have been found in proteins in which type III is the least common. The general structure of a β-turn is shown in Figure 23.

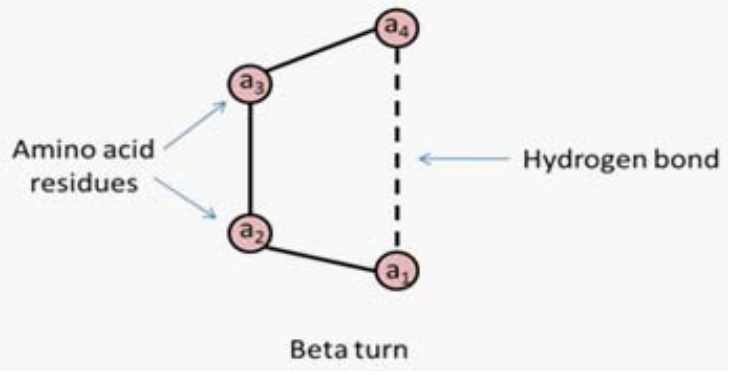


Figure 23: Structure of β-turn [3].

Coils or loops

In addition to the secondary structures described above there is significant amount of unordered structure in various segment of polypeptide chain in which psi and phi angles are not equal. These unordered structures have coil or loop conformation.

Tertiary structure

When secondary structure undergoes tertiary folding this gives rise to the formation of tertiary structure. The tertiary structure is stabilized by long range interactions such as disulfide cross-link, hydrogen bonding, hydrophobic interactions, electrostatic interactions and van der Waals interactions. In this type of structure amino acid that are far apart in the polypeptide sequence are brought together in close proximity and different types of secondary structure may interact to

form tertiary structure (Figure 24) [3,4]. Tertiary structures of protein can be analyzed by NMR spectroscopy and X-ray diffraction. In the tertiary structure hydrophobic amino acids are buried in the interior of the protein while polar residues occur at the surface.

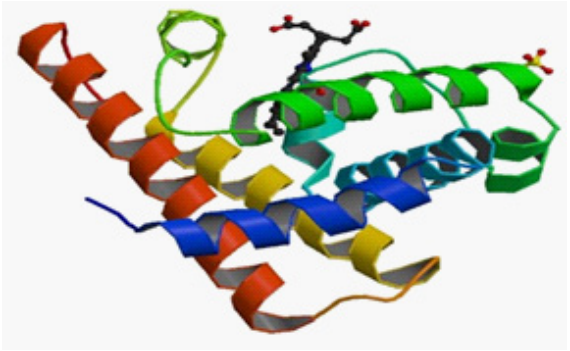


Figure 24: Tertiary structure of oxymyoglobin (**PDB ID 1MBO**) [4].

Quaternary Structure

Some proteins contain two or more separate polypeptide chains or subunits which may be identical or different. The arrangement of these protein subunits in three dimensions is known as quaternary structure which is biologically functional. These subunits or polypeptide chains may be similar or different thus produce homogeneous and heterogeneous quaternary structures respectively. The hemoglobin proteins possess heterogeneous quaternary structure because it is made up of 2 α chains and two β-chains (Figure 25) [5]. Similarly aspartate transcarbamylase possess twelve polypeptide chains or subunits (six catalytic and six regulatory) that are arranged in the form of two catalytic trimers and three regulatory dimers. Based on higher level of structure, proteins are classified as fibrous and globular. Fibrous proteins contain single type of secondary structure such as strands or sheets and their tertiary structure is relatively simple where as in globular proteins, the polypeptide chains are arranged into spherical or globular shape and it often contain several types of secondary structures.

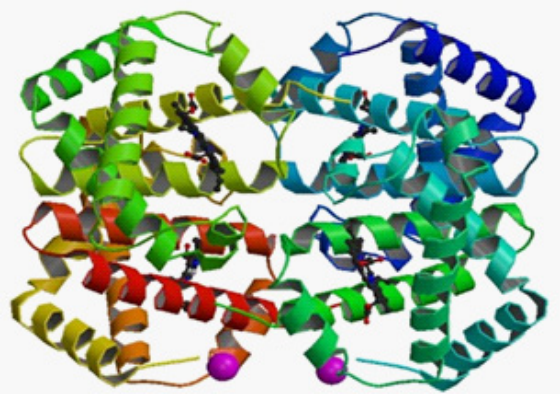


Figure 25: Quaternary structure of human deoxyhaemoglobin (PDB ID 2HHB) [5].

Learning outcomes

The primary structure of a protein is defined as a linear sequence of amino acids and the location of disulfide (-S-S-) bonds if any.

A peptide group can make two dihederal angles namely ψ (psi) and φ (phi).

A polypeptide segment can acquire a regular secondary structure such as different helical structure, β-pleated sheet, β-turn, coils or loop, tertiary structure, quaternary structure.

Objectives

To provide structure-function relationships of keratins, collagen and silk fibroin

α-Keratins

α-Keratins provide mechanical strength to cell. It is basically found in hair, wool, nails, claws, quills, horn, hoves and much of the outer layer of the skin where it accounts almost the entire dry weight. The α-keratins belong to broad family of protein known as Intermediate Filament (IF) proteins. α-keratins are rich in hydrophobic amino acids. The α-keratin is a right-handed α-helix which is arranged as a coiled coil. The two strands of α-keratin are oriented in parallel and are wrapped about each other to form super twisted coiled coil (Figure 26). The super twisting increases the strength of the overall structure. Because of twisting of the axis of α-helix the pitch of the helix varies from 5.15-5.25A rather than 5.4A.The helical path of super twist is left handed. The surface where two α-helix interact with each other are hydrophobic.

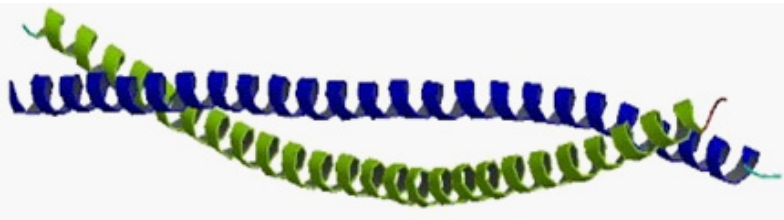


Figure 26: Structure of keratin.

Collagen

Collagen also provides strength to the cell. It is found in the connective tissue such as tendon, cartilage, bone matrix and cornea. The collagen helix is quite different from α-helix. It is a left-handed helix and has three amino acid residues per turn of the helix. The dihedral angles ψ and φ are $+153^0$ and -51^0 respectively. Collagen is also a coiled coil protein in which three separate polypeptides α-chains are super twisted about each other (Figure 27) [6]. But the super twisting is right-handed in collagen opposite to that found in keratin. The collagen helix is found in different vertebrates and contains 35% Gly, 11% Ala and 21% Pro and 4-Hyp (4-hyroxyprolone). These unusual contents of the amino acids impose the structural constraints in collagen helix. The collagen contains Gly-X-Y motif, where Pro is often found at X position and Y is 4-Hyp. At the

tight junction of α-chain only Gly occurs. The Pro and 4-Hyp residues permit sharp twisting of the collagen helix. The amino acid sequence and the super twisted quaternary structure of collagen allow a very tight packing of its three polypeptide chain whereas 4-hydroxyproline has a structural role. The α-chains of collagen molecule of fibrils are cross-linked by covalent bonds involving Lys, HyLys (5-hydroxylysine), or His residue that are present at a few of the X and Y positions. These links create an uncommon amino acid residue such as *dehydrohydroxy-lysinonorleucine*. The tight wrapping of the α-chains in the collagen triple helix provides tensile strength which is greater than that of a steel wire of equal cross-section area. The triple helical collagen molecules are also sometimes referred to as tropocollagen molecules. The collagen molecules are arranged in various ways to provide different degrees of tensile strength.

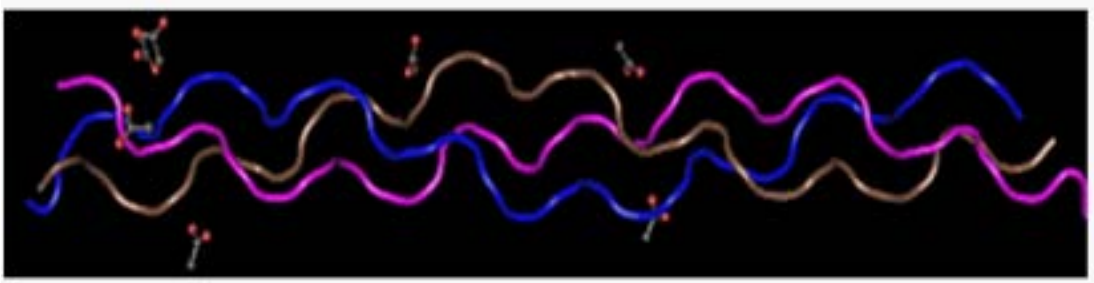


Figure 27: Hydration structure of a collagen peptide. (PDB ID: 1CGD) [6].

Silk fibroin

Silk fibroin is produced by insects and spiders. It has β-conformation and is rich in Ala and Gly residues that facilitate close packing of β-sheet and meshing of their side chains (Figure 28). The silk fibroin structure is stabilized by extensive hydrogen bonding and van der Waals interactions. The silk fibroin structure does not stretch because it already exists as extended structure.

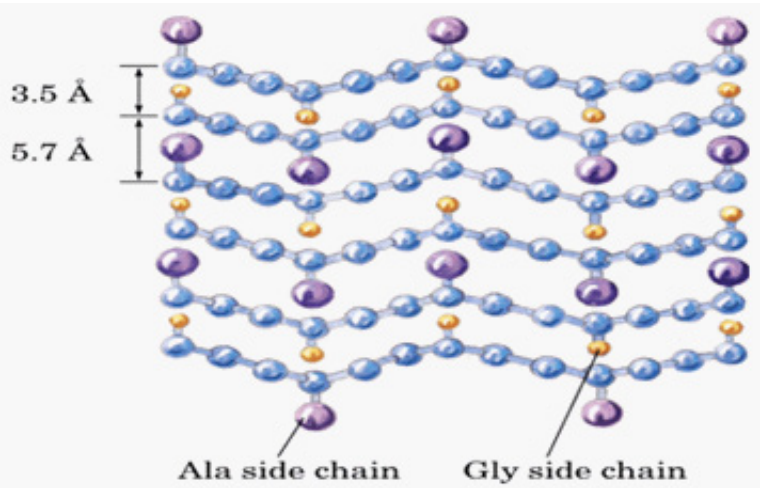


Figure 28: Structure of fibroin.

Learning outcomes

Fibrous proteins such as keratins, collagen and silk fibroin provide strength to cells. These proteins are found in hair, wool, nails, claws, quills, horn, hove, connective tissue such as tendon, cartilage, and bone matrix and cornea. These protein showed structure-function relationship.

Objectives

To describe overview of myoglobin

To provide detail informations pertaining to binding equilibria

To describe overview of hemoglobin

To derive Hill equation

To explain models for allostery

To show Bohr effect

Myoglobin and Hemoglobin

Myoglobin

Myoglobin is a globular protein found primarily in the cardiac and red skeletal muscles, functions in the storage of oxygen and facilitates the transport of oxygen to the mitochondria for the oxidative phosphorylation. Myoglobin is relatively small (Mr 16700) oxygen-binding protein of muscle cells. It functions both to store and to facilitate oxygen diffusion in rapidly contracting muscle tissues. Myoglobin contains a single polypeptide chain of 153 amino acid residues and a single protoporphyrin or heme group. The heme group is responsible for the red-brown color of myoglobin. Its three-dimensional structure was first solved by John Kendrew and his colleagues in the 1950's.

Figure 29 shows the structure of myoglobin. The backbone of the myoglobin consists of eight straight segments of α-helix which are interrupted by bend some of which are β-turns. The longest α-helix has 23 amino acid residues whereas shortest helix contains 7 amino acid residues and all helices are right handed. More than 70% of the residues in myoglobin are α- helix.

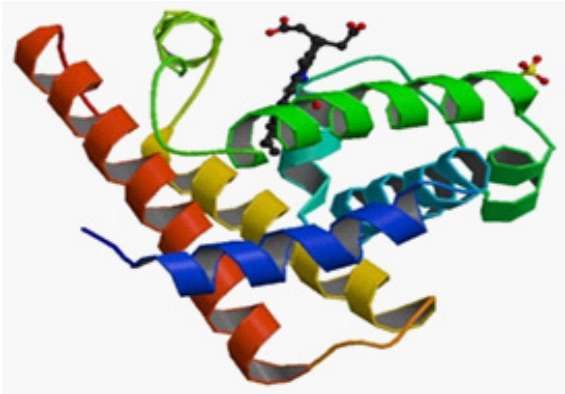


Figure 29: Structure of oxymyoglobin (PDB ID 1MBO) [4].

Hydrophobic amino acid residues are buried which stabilize the structure of myoglobin through hydrophobic interactions. Only two of the polar group lies at the surface. The α-helices wrap around the central pocket containing heme group which is capable of binding various ligand molecules including oxygen, carbon monoxide and nitric acid. The iron atom is in the center of the heme group. Oxidation of the iron atom (Fe^{2+} -> Fe^{3+}) is mainly responsible for the color of muscle and blood. At the center of protporphyrin, the iron atom is bonded to nitrogen atoms of four pyrrole rings. The iron atom can form two additional bonds, one on each side of the heme plane. These binding sites are called the fifth and sixth coordination sites. In myoglobin, the fifth coordination site is occupied by the imidazole ring of a histidine residue on the protein. This histidine is referred to as the proximal histidine. The sixth coordination site is available to bind oxygen. The iron atom in deoxymyoglobin lies about four angstrom out of the plane of the protoporphyrin plane because it is too big to fit into the well-defined hole. Analysis of the different myoglobin structures showed that structural changes occur in making the binding site more favorable for oxygen binding. This provided for the first time structure-function relationship. The function of myoglobin and hemoglobin depends on its ability not only to bind oxygen but also to release it. Therefore, function of myoglobin pertains to protein ligand interaction. Myoglobin and hemoglobin's structures and ligand binding properties have evolved differently for the different functions of these two proteins.

Binding equilibria

Binding equilibrium has same general concept as proton binding/dissociation mentioned earlier. Analysis of binding equilibria requires a few simple algebraic equations:

It is more appropriate to consider dissociation constant here for protein-ligand interaction as $K_d = 1/K_a$ and its unit is in molar concentration (M)

$$P{\cdot}L \leftrightarrow P+L \qquad (1)$$

Deriving the expression for the equilibrium dissociation constant K_d for the reaction

we have:

$$kd = \frac{[P][L]}{[PL]} = \frac{kd}{ka} \qquad (2)$$

$$[PL] = \frac{[P][L]}{kd} \qquad (3)$$

Useful parameter for plotting the data is θ (greek letter theta) which is defined as the fraction of total binding sites on the protein ($[P]_{total}$) that are actually occupied by ligand under the given conditions, [occupied sites]/[total sites]:

$$\theta = \frac{[L]_{bound}}{[P]_{bound}} = \frac{[PL]}{[P][L]} \qquad (4)$$

The range of values that θ can varies from 0 (all sites empty, at [L] = 0) to 1.0 (max. θ is where all sites are occupied that is saturation conditions is met [L] >>K_d)

θ = 1.0 isn't experimentally quite achievable, since approach to site saturation is asymptotic. Substituting the expressions for K_d in the above equation we have the following expression for θ:

$$è = \frac{[L]}{Kd + [L]} \qquad (5)$$

When the [L] equals K_d half of the ligand binding sites are occupied. As the L value falls below K_d, lesser amounts of ligands are bound to the protein. In order for 90% of the available ligand-binding sites to be occupied, [L] must be nine times greater than K_d. Thus, K_d is equivalent to the molar concentration of ligand at which half of the available ligand binding sites are occupied. This is the point where the protein is said to have reached half-saturation with respect to ligand binding. More tightly a protein binds to ligand, the lower the concentration of ligand is required for half of the binding sites to be occupied and thus the lower the value of K_d or in other words the lower the value of K_d corresponds to a higher affinity of ligand for the protein. However, because oxygen is a gas, we can substitute the concentration of dissolved oxygen for [L]

$$MbO_2 \rightarrow Mb + O_2 \qquad (6)$$

$$\text{è} = \frac{[O_2]}{K_d + O_2} = \frac{PO_2}{P_{50} + PO_2} \qquad (7)$$

Substituting the value of Kd in the expression 5 for θ we have:

$$\text{è} = \frac{[MbO_2]}{[Mb] + [MbO_2]} \qquad (8)$$

In experiment using oxygen as ligand we get the saturation curve for myoglobin as shown above (Figure 30), it is the partial pressure of oxygen (PO_2) of the gas phase which is much easier to measure than the concentration of oxygen dissolved in solution. Thus, concentration of oxygen in the blood is proportional to the local pressure of the gas. So we can define partial pressure of oxygen as P_{50}:

$$kd = \frac{[Mb][O_2]}{MbO_2} \qquad (9)$$

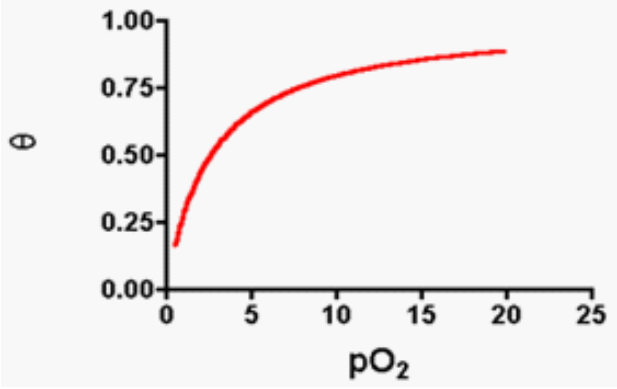


Figure 30: O_2 saturation curve for myoglobin.

The units of pO_2 on plot above are torr. Gas partial pressure units: 1 torr = 1 mm Hg = 0.133 kilopascals (kPa). The partial pressure of O_2, pO_2 is the fraction of gas phase that is O_2, another way to express the concentration of O_2 is P_{50} which is the partial pressure of O_2 required to give 50% saturation.

Hemoglobin

Hemoglobin also abbreviated as Hb, is an iron-containing oxygen-transport protein that is present in the red blood cells of nearly all vertebrates as well as in the tissues of some invertebrates. Hemoglobin has a molecular weight of approximately 64500 kDa. It is spherical in shape with diameter of approximately 5.5nm. Hemoglobin transports oxygen from

the lungs to the rest of the body (i.e. the tissues). There, it releases the oxygen to permit aerobic respiration to provide energy for carrying out the metabolic reaction. In tissues hemoglobin picks up carbon dioxide and transport back to the lungs. In the arterial circulation, hemoglobin has a higher affinity for oxygen and a lower affinity for carbon dioxide, organic phosphates, and hydrogen and chloride ions. In the venous circulation, these relative affinities are reversed.

Hemoglobin comprises of four subunits, where each subunit is made up of single polypeptide chain and one heme group (Figure 31*)*. All types of hemoglobins carry the same prosthetic heme group iron protoporphyrin IX which is associated with α-chain containing 141 amino acids residues and β-chain of 146 amino acid residues. The ferrous ion of the heme is linked to the nitrogen of a histidine. The porphyrin ring is fixed into the pocket by a phenylalanine of the polypeptide chain. The polypeptide chains of adult hemoglobin are of two kinds, one is alpha chain and other is beta, which are similar in length but differ in amino acid sequence. The alpha chain of all human hemoglobins, embryonic and adult, is similar. The other chains are beta chain of normal adult hemoglobin ($\alpha_2\beta_2$), the gamma chain of fetal hemoglobin ($\alpha_2\beta_2$), and the delta chain of HbA_2 ($\alpha_2\delta_2$). In some variants of hemoglobin, the gamma genes are duplicated, giving rise to two types of gamma chains. The quaternary structure of hemoglobin shows strong interaction between it α and β subunits. More than 30 residues in hemoglobin are involved in $\alpha_1\beta_1$ and $\alpha_2\beta_2$ interface whereas $\alpha_1\beta_2$ and $\alpha_2\beta_1$ interface involves only 19 residues. The hydrophobic interactions predominate in all types of interfaces.

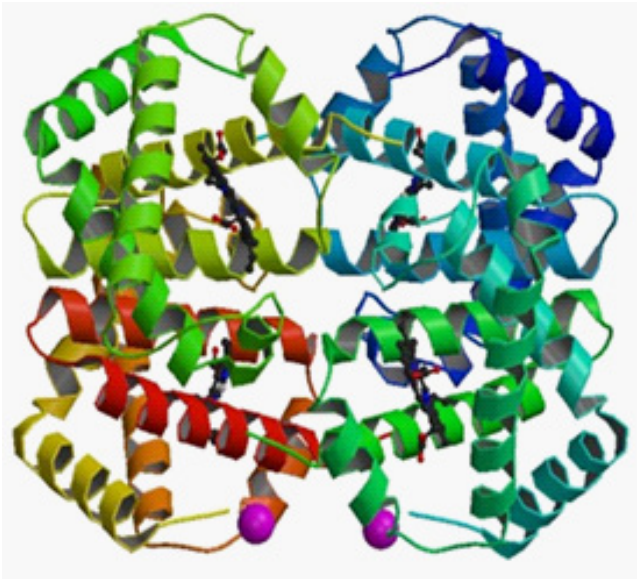


Figure 31: Crystal structure of human deoxyhaemoglobin (PDB ID 2HHB) [5].

Hemoglobin undergoes structural change upon binding to oxygen. The X-ray crystallographic data have shown that hemoglobin exists in two major conformational states: the stressed T state and relaxed R state (Figure 32) [7,8]. The R state has significantly higher affinity for oxygen than

the T state. When oxygen is not present, T state is more stable and is predominant conformation of deoxyhemoglobin. Binding of O_2 to hemoglobin subunit in the T state (low affinity state) triggers conformational changes leading to the formation of R state (high affinity state). As more and more O_2 molecules are bound all of the subunits are transformed to the T state through subunit-subunit interactions as a result the curve is sigmoidal or S shaped and protein shows T→R transition (Figure 33) [9]. Contrary to this, single subunit protein like myoglobin cannot produce sigmoid binding curve even if binding to ligand elicits conformational changes. Max Perutz showed that in T→R transition produce changes in the positions of key amino acid residues surrounding the heme group. In the T state the porphyrin is slightly puckered causing the heme iron to protrude somewhat on the proximal side of His8 residue. Binding of O_2 causes the heme to assume a more planar conformation thus shifts the position of proximal His8. These changes produce conformational rearrangements in the ion pairs at the interface of the $\alpha_1\beta_2$ subunit. Because of cooperativity as shown by sigmoidal binding curve (Figure 33) of hemoglobin which is much more sensitive to small differences in concentration between the tissues and the lungs. It binds oxygen in the lungs where pO_2 is high and releases the oxygen in tissues where pO_2 is low. Since myoglobin has single subunit and single binding site, it cannot produce a sigmoidal binding curve when it binds oxygen. When the first molecule of O_2 interacts with deoxyhemoglobin it binds weakly, because it binds only one subunit which is in the T state. Its binding induces the conformational changes to adjacent subunits making it easier for additional molecules of O_2 to bind. In effect the T→R transition occurs much more readily in the second subunit and when the last (fourth) O_2 molecule binds to a heme, all of the subunit exists in the R state. Hence it binds with much higher affinity than the first molecule. This type of protein is known as allosteric protein derived from the Greek word allos "other" and stereos" other shapes". Its binding to ligand at one site affects the binding properties of another site. Ligands of this type are referred as modulators. The modulator for allosteric protein maybe an activator or an inhibitor. When the normal ligand and modulator are identical the interaction is said to be homotropic and when the modulator is a molecule other than the normal ligand, then this type of interaction is known as heterotropic.

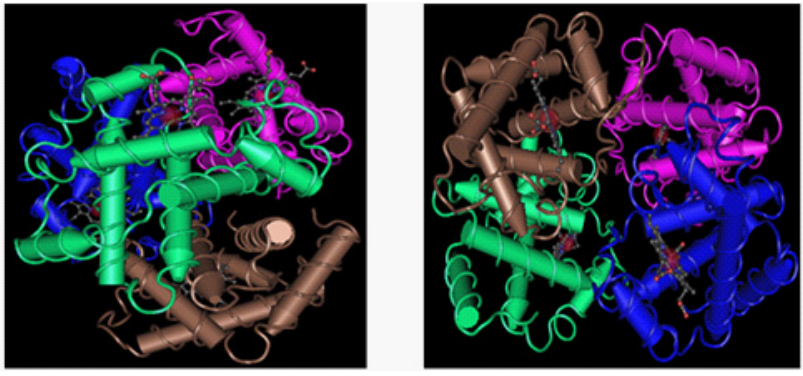


Figure 32: T state of deoxyhemoglobin (PDB ID 1HGA) (left) [7] and R-state of oxyhemoglobin (PDB ID 1BBB) (right) [8].

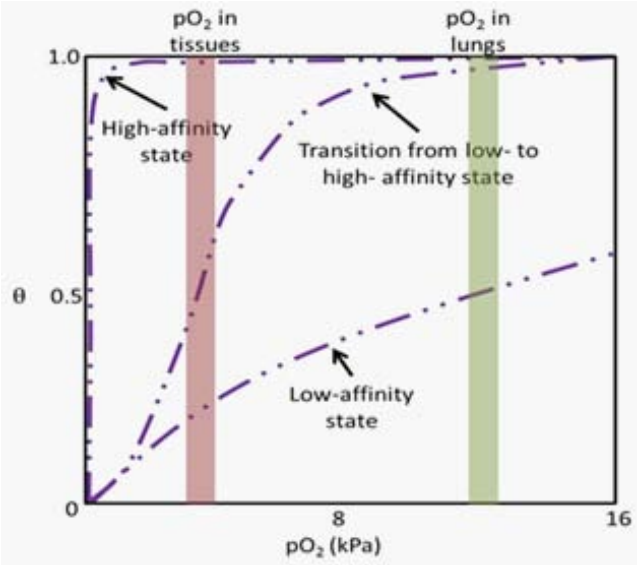


Figure 33: A sigmoid binding (cooperative) curve of hemogolobin [9].

The Hill Plot

Since hemoglobin binds oxygen cooperatively. This cooperative ligand binding can be shown by by Hill plot [10] using the Hill equation which is similar to the one dicussed above (Henderson-Hasselbalch equation).

For a protein with n binding sites, the equilibrium constant can be written as:

$$P + nL \rightarrow PLn \qquad (10)$$

$$Kd = \frac{[P][L]^n}{[PL_n]} = \frac{K_d}{K_a} \qquad (11)$$

$$[PL_n] = \frac{[P][L]^n}{K_d} \qquad (12)$$

Useful parameter for plotting the data is θ (greek letter theta) which is the fraction of the total binding sites on the protein ($[P]_{total}$) molecule that are actually occupied by ligand under the given conditions, [occupied sites]/[total sites]:

$$\frac{\theta}{1-\theta} = \frac{[L]^n}{K_d} \qquad (13)$$

Rearranging the above equation and taking the log on both sides, yields an equation:

$$\theta = \frac{[L]^n}{[L]^n + K_d} \qquad (14)$$

$$\log \frac{\theta}{1-\theta} = n\log[L] - \log k_d \quad (15)$$

where $K_d = [L]^n\ 0.5$

The above equation is Hill equation and a plot of log [θ /(1- θ) vs log[L] is known as Hill plot. The plot has a slope of n. However, experimentally determined slope actually reflects not the number of binding sites but the degree of interaction between them. The slope of the Hill plot is therefore denoted by nH, the Hill coefficient which is a measure of the degree of cooperativity. If nH equals 1 this means that ligand binding is not cooperative whereas nH greater than 1 indicates positive cooperativity. Under this condition binding of one molecule facilitates the binding of other and when nH=n, the binding site of hemoglobin is fully saturated. When nH is less than 1, it implies negative cooperativity.

For the binding of oxygen to hemoglobin, the Hill equation pO_2 was substituted for [L] and P^n_{50} for $K_{d,}$ so we have:

$$\log \frac{\theta}{1-\theta} = n\log O_2 - n\log P_{50} \quad (16)$$

Models for Allostery, Cooperative Ligand Binding (equilibria between high-affinity and low-affinity conformations)

Two models for explaining the cooperativity are known. The first model was proposed by Jacqueous Monod, Jefferies Wyman, and Jean-Pierre Changuex in 1965 and it is known as MWC model [11] or the concerted model and the second model is sequential model (Figure 34). This model describes that when allosteric protein is made up of identical subunits, in this scenario each of the subunit can exist in at least two conformational state either low affinity state (□) or high affinity state (▫) and that all of its subunits are postulated to be in the same conformation sate (either □ or ▫). The two conformations are said to be in equilibrium. The second model is known as sequential model which was proposed by Daniel Koshland and colleagues. In the sequential model, each subunit can be either low affinity (□) or high affinity (▫). In this model when ligand binds it induces conformational changes of that subunit which makes a similar change in adjacent subunit thereby it facilitate binding of second ligand molecule more likely which in turn further facilitate third ligand molecule to bind much more easily and when the last ligand molecule binds to the last subunit, all of the subunits exist in the T state.

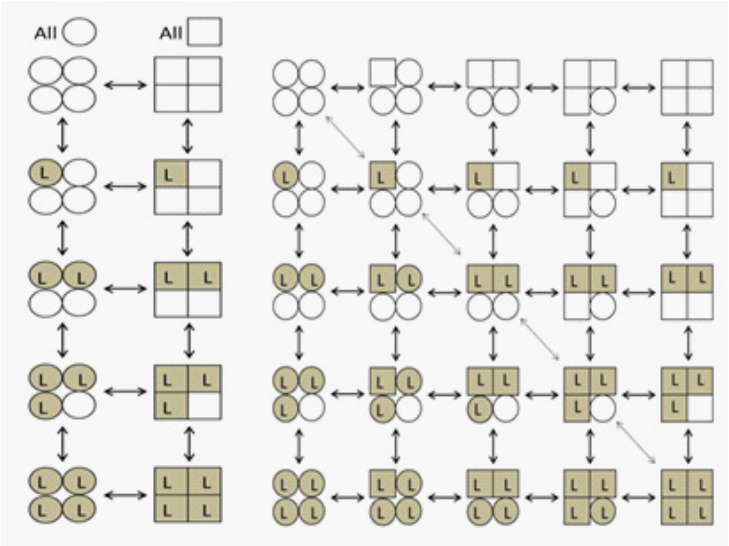


Figure 34: The MWC or concerted model (left) and sequential model for allostery (right).

The Bohr Effect

Besides carrying oxygen from the lungs to the tissue, hemoglobin also carries H^+ and CO_2 from the tissue to the lungs and the kidneys where they are excreted. Binding of H^+ and CO_2 is inversely related to the binding of oxygen. The effect of pH and CO_2 concentration on the binding and release of oxygen by hemoglobin is known as Bohr Effect which is named after Christian Bohr, the Danish Physiologist in 1904.

Protons and CO_2 are negative heterotropic effectors (allosteric inhibitors) of O_2 binding to Hb. In the tissues, pH is low due to metabolic production of CO_2 which release the protons (H^+) upon hydration to form bicarbonate.

$$CO_2 + H_2O \leftrightarrow HCO_3^- + H^+$$

This reaction is catalyzed by carbonic anhydrase. At low pH more proton binds to hemoglobin and stabilizes the deoxy state which is a low-O_2 affinity form of Hb (T state), therefore it release oxygen readily to the muscle (Figure 35). The origin of the Bohr Effect lies in the fact that deoxy Hb has several functional groups that are weak acids than they are in oxy-hemoglobin, so oxygen binding reduces H^+ affinity and H^+ binding reduces oxygen affinity.

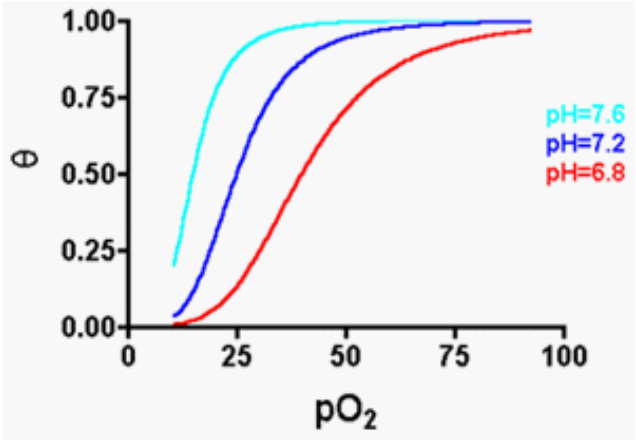


Figure 35: Effect of H^+ on oxygen saturation curve for hemoglobin.

$Hb{\cdot}4O_2 + nH^+ \leftrightarrow Hb{\cdot}nH^+ + 4O_2$

In the Bohr effect, the His residue is protonated which forms salt bridge to Asp 94, thereby it stabilizes the structure, leading to an increase in pK_a (weaker acid) of other functional groups and promotion of deoxy state (T state) and releases the oxygen. Salt bridge does not form in the oxy state (R state). Further, protonation of His-146 favors the deoxy conformation.

Carbon Dioxide Transport

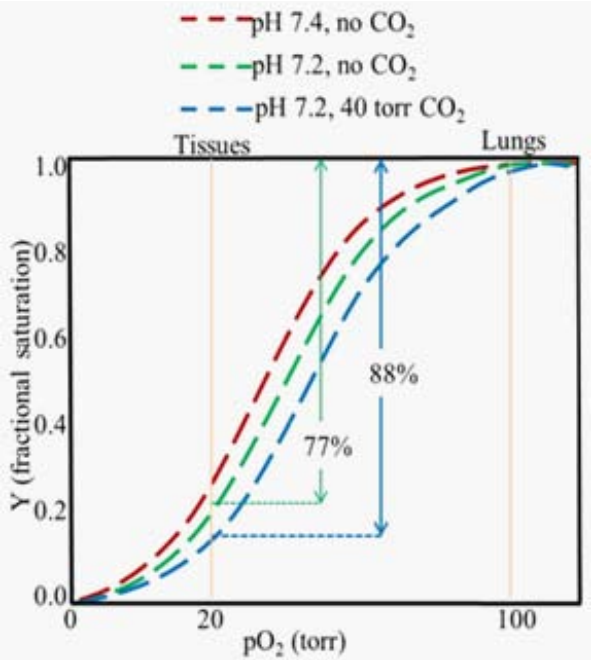
CO_2 is a negative heterotropic effector or allosteric inhibitor of O_2 binding to Hb. The presence of CO_2 in the tissues reduces affinity of Hb for O_2 thereby it favors deoxy, T state in two ways:

CO_2 lowers the pH (Bohr Effect) of the blood by reacting with the N-terminal α-amino groups of Hb and thereby it forms carbamate to give rise to the formation of carbaminohemoglobin and release the protons as well.

$R{-}NH_2 + CO_2 \leftrightarrow R{-}NH{-}COO^- + H^+$

The negative charge introduced on the protein (carbamate) allows formation of additional salt bridges, but only in the deoxy state (T state). Similarly the released H^+ also contribute to Bohr effect (as mentioned above). Thus, carbamate formation (CO_2 binding) and released H^+ favor the

deoxy state and promote the release of oxygen (or shift the curve to the right) (Figure 36). Conversely, when hemoglobin reaches the lungs, the high oxygen concentration promotes oxygen binding and release the CO_2.



High altitude adaptation

Humans are capable of adaptation to high altitude involving several physiological processes including an increase in the number of erythrocytes and also increase in the amount of hemoglobin per erythrocyte, which generally takes several weeks to accomplish. Thus, increase in the amount of 2, 3-BPG occurs within 24 hours at high altitude. Increasing the concentration of 2, 3-BPG reduces the affinity of hemoglobin for O_2 (or shifts the curve to right) (Figure 37) thereby it facilitates delivery of the O_2 to tissues.

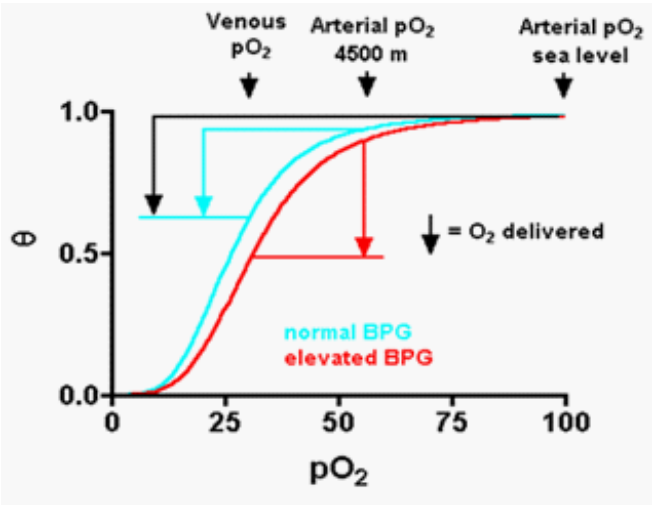


Figure 37: Effect of BPG on delivery of oxygen at high altitude.

Fetal Hemoglobin

Regulation of oxygen binding to hemoglobin by BPG has an important role in fetal development. The fetus synthesizes γ subunits rather than β-subunits thus form $\alpha_2\gamma 2$ tetramer of hemoglobin. This tetramer containing γ subunit has a much lower affinity for BPG than normal adult which contains β-subunit. Similarly, the $\alpha_2\gamma 2$ subunit has a higher affinity for oxygen than adult hemoglobin (Figure 38). Therefore fetus extracts oxygen from its mother's blood which diffuses across the placenta.

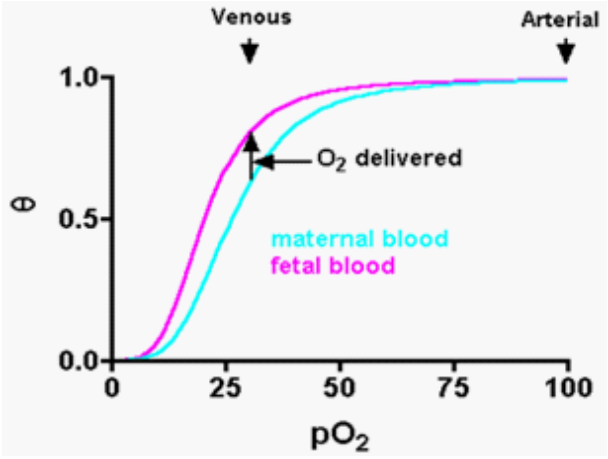


Figure 38: Diagram shows the oxygen saturation curve of fetal hemoglobin and how it extracts oxygen from the maternal blood.

Learning outcomes

Myoglobin is a globular protein found primarily in the cardiac and red skeletal muscles, functions in the storage of oxygen and facilitates the transport of oxygen to the mitochondria for the oxidative phosphorylation.

Hemoglobin is an iron-containing oxygen-transport protein that transports oxygen from the lungs to the rest of the body (i.e. the tissues).

Two models for explaining the cooperativity are known one is concerted and second is sequential.

Hemoglobin undergoes T→R transition upon binding to oxygen and shows cooperativity.

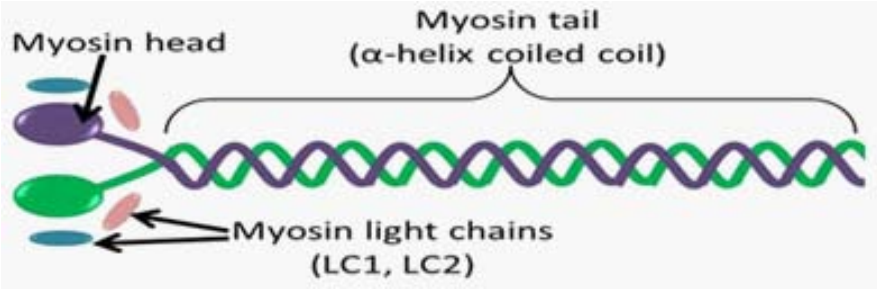
Effect of pH and CO_2 concentration on the binding and release of oxygen by hemoglobin is known as Bohr Effect.

Objectives

To describe myosin and actin

Myosin and Actin

Myosins comprise a super family of ATP-dependent motor proteins**.** Myosin plays a structural and enzymatic role in muscle contraction and intracellular motility and is best known for their role in muscle contraction and their involvement in a wide range of other motility processes in eukaryotes. Myosin belongs to super family of protein which binds actin, hydrolyzes ATP and transduces force. Thus, most of myosins are located in the muscle cells. Myosin has molecular mass of 540,000. It contains six subunit: two heavy chains (each has a molecular mass of 220,000 and four light chains each of molecular mass of 20,000). These chains are wrapped around each other in a fibrous left-handed coiled coil (tail) and the amino terminus is folded into globular domains (heads) (Figure 39). The structure of myosin is similar to that of α-keratin.



Actin is a family of globular multi-functional proteins that forms microfilaments. It is found in nearly all eukaryotic cells (except nematode sperm), where it may be present at a concentration of over 100 μM. An actin protein's molecular mass is roughly 42000Da, with a diameter of 4 to 7 nm, and it is the monomeric subunit of microfilaments, and thin filaments, that are part of the contractile apparatus in muscle cells. It can exists as free monomer called G-actin (globular) or as part of a linear polymer microfilament called F-actin (filamentous), both of which are essential for cellular functions such as mobility and contraction of cells during cell division (Figure 40).

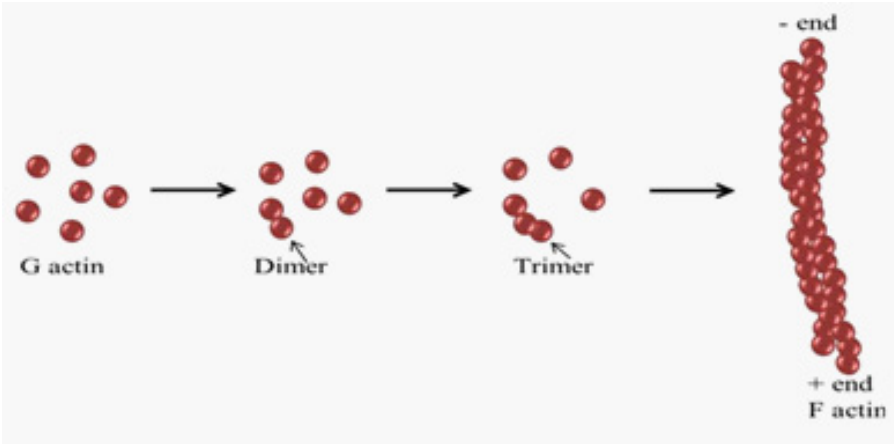


Figure 40: Structure of G and F actin.

References

1. Sela M, White FH, Anfinsen CB. Reductive Cleavage of Disulfide Bridges in Ribonuclease. Science. 1957; 125: 691-692.
2. Ramachandran GN, Ramakrishnan C, Sasisekharan V. Stereochemistry of polypeptide chain configurations. J Mol Biol. 1963; 7: 95-99.
3. Tayyab S, Nasrulhaq A. A journey from amino acids to proteins. University of Malaya press. 2006.
4. Phillips SE. Structure and refinement of oxymyoglobin at 1.6 A resolution. J Mol Biol. 1980; 142: 531-554.
5. Fermi G, Perutz MF, Shaanan B, Fourme R. The crystal structure of human deoxyhaemoglobin at 1.74 A resolution. J Mol Biol. 1984; 175: 159-174.
6. Bella J, Brodsky B, Berman HM. Hydration structure of a collagen peptide. Structure. 1995; 3: 893-906.
7. Liddington R, Derewenda Z, Dodson E, Hubbard R, Dodson G. High resolution crystal structures and comparisons of T-state deoxyhemoglobin and two liganded T-state hemoglobin's: T(alpha-oxy)hemoglobin and T(met)hemoglobin. J Mol Biol. 1992; 228: 551-579.
8. Silva MM, Rogers PH, Arnone A. A third quaternary structure of human hemoglobin A at 1.7-A resolution. J Biol Chem. 1992; 267: 17248-17256.
9. Hill AV. The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. J Physiol. 1910; 40: 4-6.
10. Cox MM, Nelson DL. Lehninger Principles of Biochemistry 2011. Fifth Edition. WH Freeman and Company. New York (USA).
11. Monod J, Wyman J, Changeux JP. On the nature of allosteric transitions: A plausible model. J M Biol. 1965; 12: 88-118.