Structural organization of proteins

Objectives:

The main objectives of this discussion are

- to understand the primary structure of protein.
- to understand the secondary-helical as well as pleated sheet structure of protein.
- to understand overall 3dimensional nature of protein and
- to understand the arrangement of two or protein in 3dimensional form.

Introduction

The spatial arrangement of atoms in a protein is called its **conformation.** The possible conformations of a protein includes any structural state without breaking covalent bonds. A change in conformation could occur by rotation about single bonds in the polypeptide chain. Hence theoretically numerous conformations are possible in a protein containing hundreds of single bonds. The conformations of protein exist in thermodynamically the most stable which having the lowest Gibbs free energy (G). Proteins in any of their functional, folded conformations are called native proteins.

Primary Structure of protein

The primary structure of a protein or first level of protein structure deals with the sequence of amino acid residue and spatial arrangement of atoms in a polypeptide chain which is linked by covalent -peptide bond. The backbone of polypeptide can be pictured as a series of rigid planes sharing a common point of rotation at alpha carbon (Cα).We can read as **C-C-N-C, C-C-N-C, C-C-N-C so on,** 1st C and last C is alpha carbon,

Here, the alpha carbon of each adjacent amino acid residues are separated by **three covalent bonds**

Cα—C—N — Cα

In the figure;
First covalent bond is form between 1st alpha carbon and carbonyl carbon (Cα—C)
2nd covalent bond is in between carbonyl carbon and amide nitrogen C—N,(the peptide bond) and
3rd covalent bond is in between amide nitrogen and 2nd alpha carbon N — Cα
In this series of plane, X ray studies found that there are six atoms directly or indirectly participating in the peptide bond. Here the four atom namely, 1st Alpha carbon (Cα), Carbonyl carbon(C), Amide nitrogen, (N), 2nd Alpha carbon (C) are in the main chain, and **two atom** which the **oxygen** atom of carbonyl group and **hydrogen** atom of amide nitrogen in the peptide; altogether six atoms (CNHCOCare lies in the same plane (co-planer).

In the peptide bond (CONH)-**second bond in the figure**, the **oxygen atom** of cabonyl group and **hydrogen atom** of amide nitrogen are in trans configuration that means when O is up position, H is down position. And, the oxygen atom of cabonyl group has partial negative charge and nitrogen atom of amide has partial positive charge hence setting up a small electric dipole. Partial sharing of two electrons occured between carbonyl **oxygen** and amide **nitrogen**therefore the peptide bonds is double bond character with bond length of 1.32Å which is shorter than Cα—N bonds with bond length of 1.45Å.

Since peptide bonds(C-N) has double bond character, the rotation of such peptide bond is strictly prohibited however, first covalent bond between alpha carbon and carbonyl carbon (Cα—C) and **third covalent bond** betweenalpha carbon andamide nitrogen (Cα—N) is purely single covalent, hence these bonds are free to rotate.By convention, the angle of rotation between Cα—N is called **phi(φ)**and angle of rotation between Cα—C is called **psi(ψ)**.The **phi (φ)** and **psi(ψ)** angles of each amino acids allows proteins to fold in many different directions.

Secondary structures of proteins

Secondary structures of proteins or 2nd level of protein structure deals with common regular folding pattern of a polypeptide/protein. These structure are mainly arises from hydrogen bonding between electronegative oxygen atom of **one peptide bond** and amide nitrogen of **another peptide bond** either in same or different polypeptide chain.

There are four types of secondary structures are fond in proteins. Let us discuss each of them one by one

1. **α-helical structure (first studied so called alpha):**

   α-helical is a rod like structure. In this structure, the polypeptide containing rigid peptide bond is tightly wound around an imaginary axis drawn longitudinally in the middle of the helix and

   The R-groups i.e. the side chain of the amino acids protrudes outward from the helical backbone.
   Generally the length of α-helix is 10-15 amino acid residues.

   The repeating unit in a single turn includes 3.6 amino acid residues and the distance between corresponding point per turn is 5.4A^{o} along the axis. **α-helix** structure is stabilized by hydrogen bonds in the main polypeptide chain. Particularly, the CO group of each amino acid (N-terminal) forms a hydrogen bond with the NH group of the amino acid that is situated four residue ahead (C-terminal) in the sequence. All the CO and NH groups formed the hydrogen bond except those amino acid which are close to both ends of the helix.

α-helix may be right handed- if the helical structure appear screw type or clockwise or left handed -counterclockwise conformation. Essentially, all α-helix found in proteins are right handed. The right handed are energetically more favourable, because the steric clash between side chains and the polypeptide backbone is less. The α-helix content of each protein is widely differ and may range from zero to 100%. For example; ferritin (an iron binding protein) – contain 75% of amino acid residue as α-helix.

Single α-helix is usually less than 5 Å however, 2 or more α-helix can interlink to form a very stable structure which can have length of 1000 Å or more. For example; myosin and tropomyosin in muscle protein and keratin in hair are made up from interlink of α-helix.

Factor affecting stability of α-helix structure

Stability of α-helix structure is mainly affected by amino acid sequence in the α-helix region, for example presence of long ⁻ve charge glutamic acid orpositive charge lysine in the polypeptide chain will repel each other hence affected helix formation. And presence of proline and glycine causes constrain in the α-helix where, proline residue in peptide linkage introduce a kink in a α-helix, hence it cannot formed any H-bond with others.

2. **β-Pleated sheet structure- second type of secondary structure of protein :**
   In this structure, the polypeptide chain is extended in zig-zag. Therefore they are relatively flat. This zig- zag chains can be arranged as side by side to form a structure resembling a series of pleats and R-groups of adjacent amino acids protrude from the zig-zag structure.
   Two or more β-Pleated sheets are layered close together within a protein. The layered polypeptide chain may be **parallel** or **antiparallel. In parallel,** the adjacent polypeptide segments are in **same orientation** whereas in **antiparallel,** the adjacent polypeptide segments are in **opposite orientation.**
   In parallel arrangement, the CO group of each amino acid is form a hydrogen bond with NH group of the 2 amino acid residue farther along the adjacent polypeptide segment.
   In antiparallel arrangement, the NH group of each amino acid is form a hydrogen bond with CO group of adjacent polypeptide segment.

   In beta sheets, many structure typically four or five but as many as 10 or more may organised together. Such beta sheets can be purely antiparallel, purely parallel or mixed.

Generally β-Pleated sheet contain high amount of glycine and alanine residue which has smallest R group. For example; fibron protein of Silk and spider web are rich in glycine and alanine residue.

3. **β-turn/ reverse turn**

   β-turn structure is a connecting element which link the ends of two adjacent segment of antiparallel β-sheet. The structure is 180^{o} turn that involved **four amino** acid residues and also known as reverse turn.

   In this structure, the Oxygen atom of CO group of first residue form a hydrogen bond with nitrogen atom of NH group of the fourth amino acid. However, the peptide group of the central two residue do not participate any inter residue hydrogen bonding. Glycine and proline residue often occur in **β-turn** where the proline make the tight turn.

   β-turn structure are mostly found in near the surface of a protein and the central two residue which do not participate any inter residue can form a hydrogen bond with water.

4. **Loops :** Loops are random coil like irregular structure found in protein. Loops are well known for reversal of polypeptide chain and comprises of 2-16 amino acid residue. They are generally found to connect the α-helix and β-sheet structure. Similar to **β-turn, they** found in near the surface of a protein and participate in interaction between protein and other molecules.

Tertiary structure of proteins / the 3rd level of protein structure

Each protein is made up of a variety of helices, β-sheets and non-regular regions, which are folded in a unique three dimensional structure. The overall 3-dimensionalstructure of a protein is referred to as tertiary structure of proteins. This folding of protein into native 3D conformation includes not only positions of primary backbone atoms, but also includes all the interaction between side chain atoms of distant places. The interaction is just an application of the solubility rule ‘like dissolve likes’.

The stability of tertiary structure of protein is maintained by many non-covalent interactions in their side chain in addition to covalent peptide and disulfide bond. Non-covalent interactions includes :

1. hydrogen bond- hydrogen bond may form between two alcohol groups,(e.g. serine, threonine) or between an alcohol and an acid (serine, aspartic acid) etc.
2. hydrophobic bond- the hydrophobic bond may be found between nonpolar side chains. (e.g. leucin and isoleucin)
3. van der Waals forces- van der Waals forces may be found between tightly packed hydrocarbon side chains
4. Electrostatic force (Salt bridges)- it may be form between positive and negative charge of side chains. (e.g. lysine and aspartic acid)

Now let us discuss, how R-groups are interacted in tertiary structure of Water soluble globular proteins.
For example; Myoglobin-an oxygen binding muscle protein

Myoglobin is a water-solublecomplex globular proteinmade up from 153 amino acid residue. *x-ray diffraction studies show that* about 70% of the main chains is folded into eight helices and rest of the chain forms turn and loops between helix. It is very compact structure and almost no empty space inside.

The interior of the myoglobin is made up of nonpolar residue such as leucine, valine, methiononie and phenylalanine however charge residue such as aspartate, glutamate, lysine and arginine are absent. Only two histidine residues found to contain in theinteriorthat play a critical role in binding of iron and oxygen.Polypeptide chain therefore folds, so that its hydrophobic side chains areburied inside and its polar, charged side chains are on the surface.

The secret of burying a segment of main chain in a hydrophobic environment arises from pairing all the NH and CO groups by hydrogen bonding in an helix or sheet. An unpaired hydrogen bonding in peptide NH or CO group markedly prefers water environment. The hydrocarbon side chains are tightly packed and stabilized by van der Waals forces which require intimate contact.

But in the case of hydrophobic protein, for example; In porin, (*an* outer membranes protein of many bacteria), the distribution of hydrophobic and hydrophilic amino acid residues are reversed in compared to water soluble protein.

Domains

Domains are structural independent folding units looking like a separate globular proteins but they are all part of same polypeptide chain connected in primary structure.

Large proteins often have 2 or more domains and the size of domain range from about 30 to 400 amino acid residues. For example, Troponin C, a protein found in muscle has 2 domains, which are present in one polypeptide chain. And extracellular part of CD4, (the cell surface protein) which the HIV attached, has four domains

Quaternary structure

Many proteins contain two or more separate polypeptide chain and each polypeptide chain is called a subunits. The arrangement and nature of their interaction of such 2 or more subunits in 3 dimensionals conformation is known as Quaternary structure of protein. The proteins with only one chain have no quaternary structure, therefore quaternary structure of protein deals with only proteins with more than one polypeptide chain. Quaternary structure is stabilized by noncovalent interactions such as hydrophobic interaction, electrostatic interaction(salt bridge), and hydrogen bonds as well as covalent disulfide bonds.

Naming of quaternary protein

Quaternary proteins are name as dimer for 2 subunits, trimer for 3 subunits, tetramer for 4 subunits and so on. And if the subunits are of same kind, then prefix homo is used and if it is different kinds then called hetero. Example; heamoglobin is made up of four subunit with 2 different kind, hence known as ***heterotetramer*** **($\alpha_2\beta_2$)**, where 2subunits of one type (α) and 2subunits of another type (β).

Conclusion:

The primary structure of a protein or first level of protein structure is the sequence of amino acid residue in a polypeptide chain. Secondary structures of proteins or 2^{nd} level of protein structure is the common regular folding pattern of a polypeptide/protein chain due to hydrogen bonding between

electronegative oxygen and nitrogen atom of adjacent peptide bonds either in same or different polypeptide chain. Secondary structure includes α-helical, β-sheet, turn and loop. Tertiary structure of proteins refers to the overall 3-dimensional conformation of a whole polypeptide chain in its folded state. Tertiary structure is stabilized by many noncovalent bonds viz; hydrogen bond, hydrophobic bond, salt bridge and covalent disulfide bond etc. and the arrangement of 2 or more subunits in 3dimentionalconformation is the Quaternary structure of protein.

Summary:

The primary structure of a protein is the sequence of amino acid residue which is linked by peptide bond. In the peptide bond, the partial sharing of two electrons occurred between carbonyl oxygen and amide nitrogen hence the bond is double bond character. Therefore, rotation of peptide bond is strictly prohibited however bonds adjacent to peptide bonds are free to rotate. Hence, proteins are allow to fold in many different directions. Secondary structures of proteins deals with common regular folding pattern of a protein which are arises from hydrogen bonding between electronegative oxygen atom of one peptide bond and nitrogen atom of peptide in a polypeptide chain. In α-helical, the CO group of each amino acid (N-terminal) forms a hydrogen bond with the NH group of the amino acid that is situated four residue ahead (C-terminal) in the sequence except that amino acid which are close to both ends of the helix. In β-Pleated sheet structure,the polypeptide chain is extended in zig zag therefore the chains are arranged as side by side to form a series of pleats. The layered polypeptide chain may be parallel (same orientation) or antiparallel (opposite orientation)**.** β-turn and Loops are connecting element which link the ends of two adjacent segment of antiparallel β-sheet and they are generally found to connect the α-helix and β-sheet structure.

The overall 3-dimensional structure of a protein is referred to as tertiary structure of proteins and it includes all the interaction between side chain atoms of distant places. E.g. Myoglobin;70% of the main chains is folded into eight helices and rest forms turn and loops between helix. Domains arestructurally independent folding unitslooking like separate globular proteins but they are all part of same polypeptide chain. Many proteins contain two or more separate polypeptide chain and the arrangement of such

2 or more subunits in 3dimentional conformation is known as Quaternary structure of protein. Quaternary structure is stabilized by noncovalent interactions such as hydrophobic interaction, electrostatic interaction and hydrogen bonds as well as covalent disulfide bonds.Example; heamoglobin **($\alpha_2\beta_2$)** is made up of four subunit with 2 different kind.

GLOSSARY:

Peptide bond : Special type of covalent bond found in protein where the amino acids are linked together in a protein**.**

Partial charge : Fraction of charge carried by an atom due unequal sharing of electron among the molecule, it is denoted by "**δ"**

Co-planer : It is a phenomenon where the atoms or groups are in the same plane.

Resonance : It is a phenomenon to contribute the true structure of a molecule due to presence of unpaired electron.

Double bond : The covalent bond form by sharing of two electrons

Conformation : The three dimensional shape or arrangement of a macromolecule e.g protein

Covalent bond : A chemical bond formed by the sharing of electron pairs

Dipole : A molecule having both positive and negative charges

Protein : A biomacromolecule which is made up from amino acid monomer

R-groups : The side chain of the amino acid is called R-groups that differ for different amino acid.

Angstrom (Å) : It is a unit of length which is equal to 10^{-10}m (metre).

FAQs

1. How secondary structure of protein is formed?

Ans: The secondary structure of protein arises from hydrogen bonding

between electronegative oxygen atom of **one peptide bond** and nitrogen atom of **another peptide bond** either in same or different polypeptide chain.

2. **Name the amino acids which are usually present in the interior of the water soluble globular protein**.

Ans: The amino acids that are usually present in the interior of the water soluble globular protein are leucine, valine, methionine and phenylalanie etc.

3. How does water soluble globular protein differ from membrane protein in structural organization?

Ans: In water soluble globular protein, non-polar amino acids are in the hydrophobic core of the protein however in case of membrane protein such amino acid residues are on the outer surface of protein in contact with the neighboring alkane chains of the lipid membrane.

4. Define the term protein domain.

Ans: Domains are structurally independent folding units looking like separate globular proteins but they are all part of same polypeptide chain connected in same primary structure.

5. **Define quaternary structure of protein?**

Ans: The arrangement of two or more separate polypeptide chain in 3dimentional conformation is known as Quaternary structure of protein.

6. Define the primary structure of protein.

Ans: The primary structureof a protein is the sequence of amino acid residue in a polypeptide chain which is linked by covalent -peptide bond.