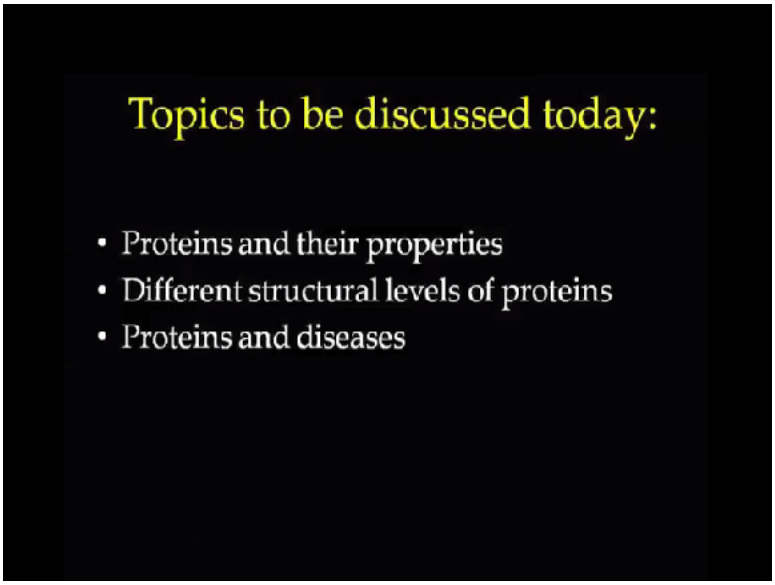
Introduction to Proteomics
Dr. Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology - Bombay

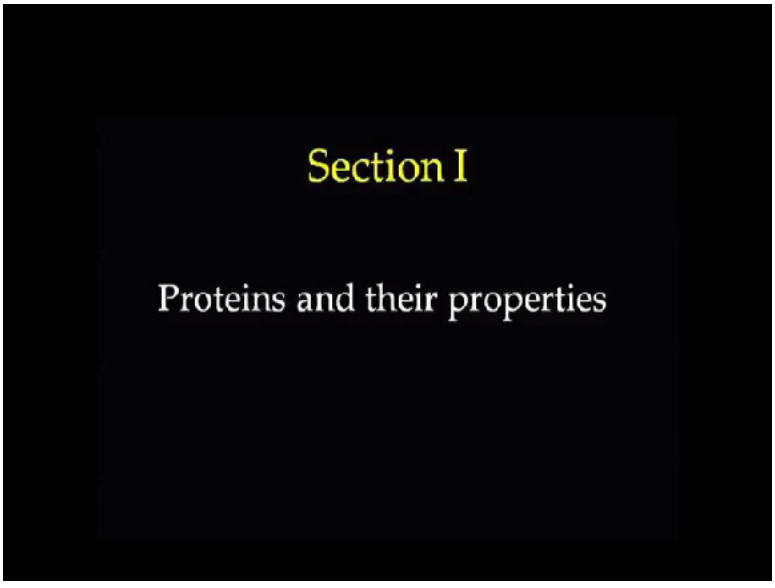
Lecture – 02
Introduction to proteins

(Refer Slide Time: 00:14)



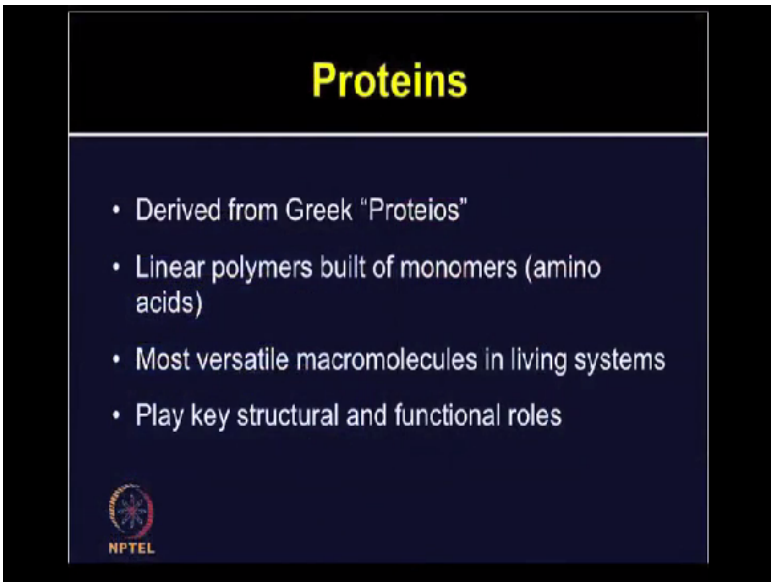
Let me give you the lecture outline. We will first talk about protein and its function and then different levels of protein structure, primary, secondary, tertiary and quaternary.

(Refer Slide Time: 00:34)



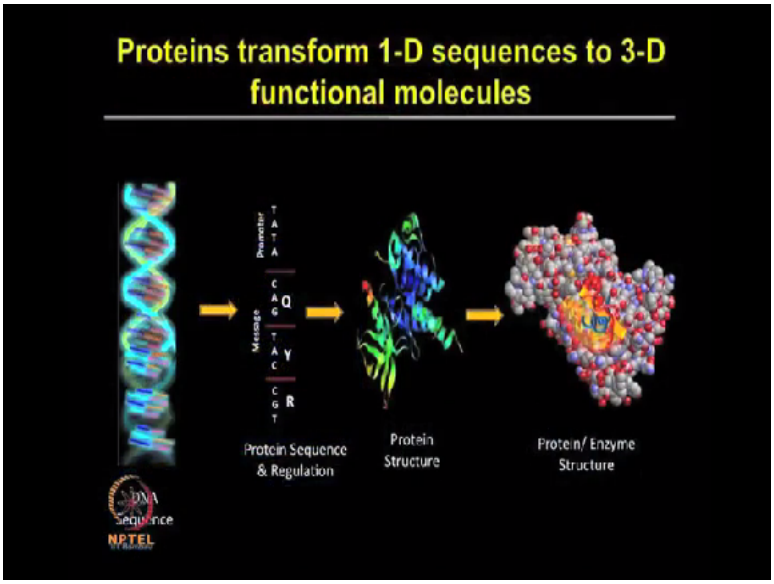
The protein term was derived from the Greek word Proteios which means of the first rank or very important.

(Refer Slide Time: 00:38)



This term was coined by scientist Berzelius in 1833. These are the linear polymers which are built of monomers or amino acid subunits. These are the most versatile macromolecules in any living system. They are crucial for various essential functions of all the biological processes and they play very critical role, both from structural and functional point of view. Therefore, studying about proteins is very important.

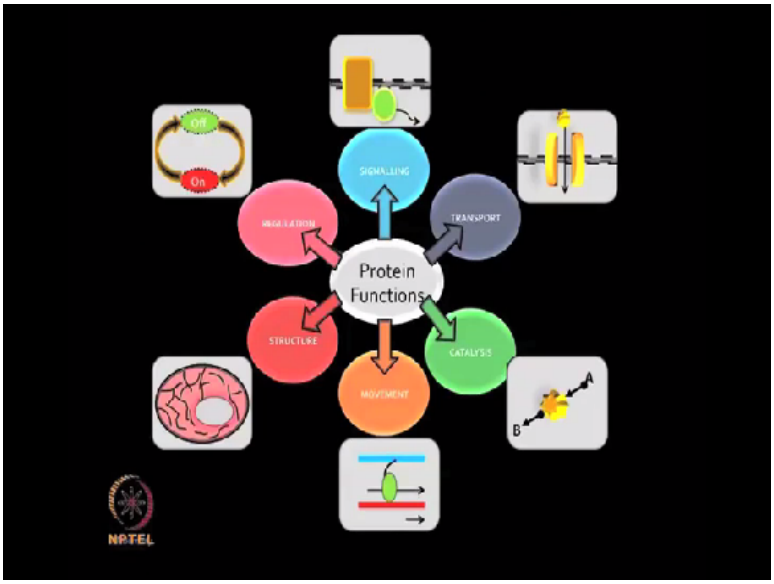
(Refer Slide Time: 01:25)



If you look at a central dogma starting from DNA to RNA and protein, the proteins can transform

the 1-dimensional sequence to the 3-dimensional functional information. Proteins can play wide range of functional properties because of their different functional groups which can account for various protein function and its activity. Protein-protein or protein and their biomolecular interactions, they are generated because of the synergistic capability of these proteins which cannot be obtained from any given individual protein.

(Refer Slide Time: 02:15)



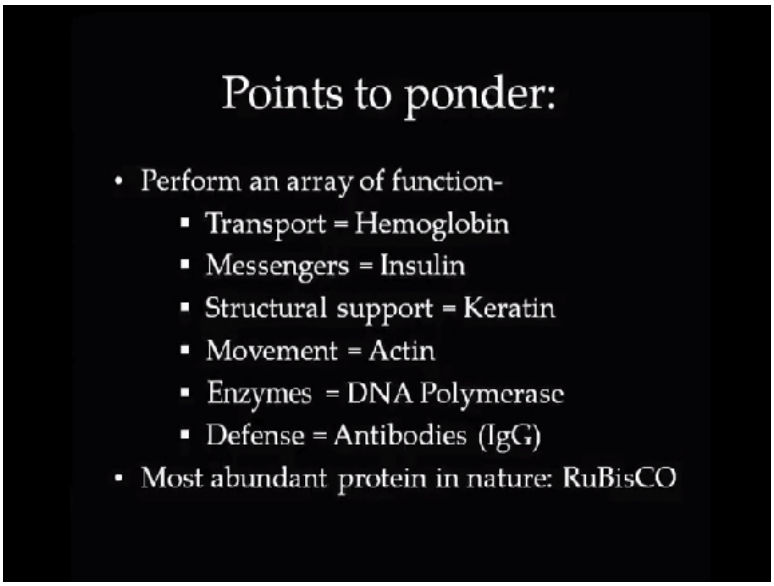
Proteins can perform various type of functions whether it is catalysis, movement, structure, regulation, signalling, transport, etc. As you can see in the slide various type of functions have been shown. Enzyme catalysis, the enzyme catalyse biochemical reactions by increasing the rate of reactions. Transport and storage, proteins can transport small molecule such as oxygen and iron.

Proteins are involved in the movement with muscle contraction if you talk about microorganism in bacteria the chemotaxis, they are responsible for the mechanical strength for example in the skin and bones, the collagen and keratin, all of these are different examples of mechanical strength. Proteins are also present as immunoglobulin responsible for the immunity.

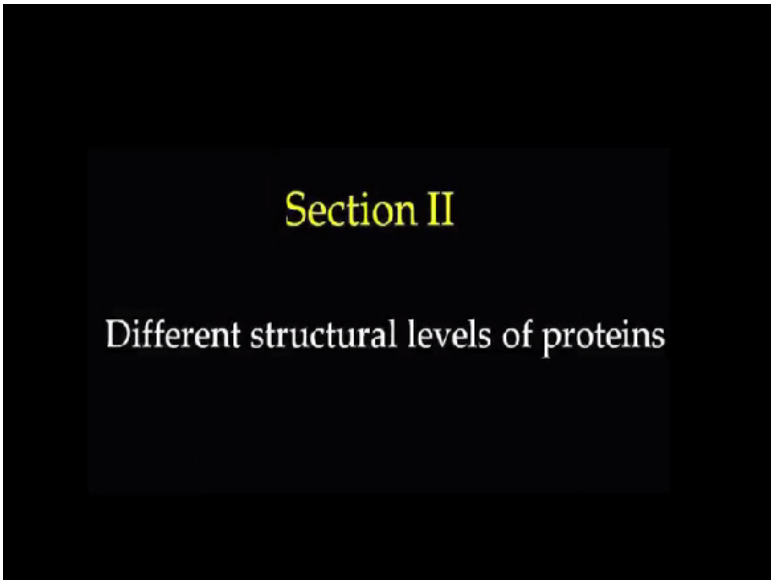
Antibodies are used for various type of protein-protein and protein-ligand interactions. Growth and differentiation, transcription factors, gene expression during growth and development for example nerve growth factors, hormones such as insulin, all of these are various examples of

proteins and the various diverse functions in which they are involved.

(Refer Slide Time: 03:49)

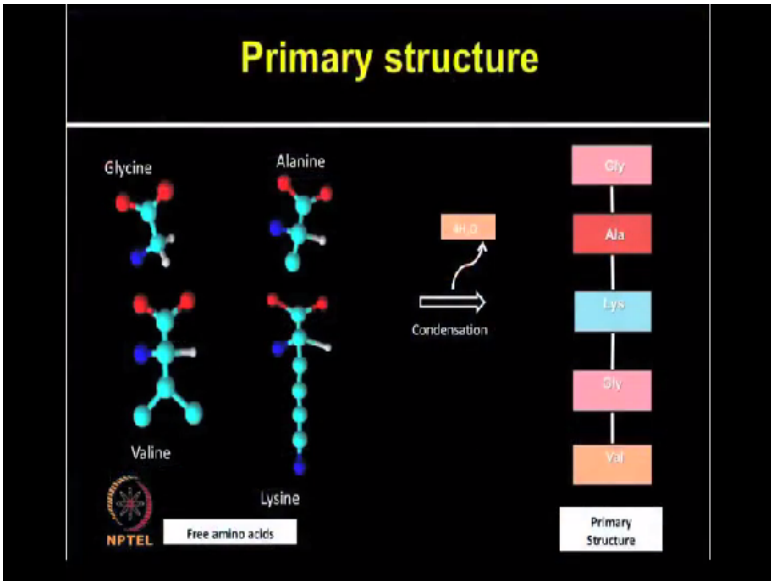


(Refer Slide Time: 04:06)



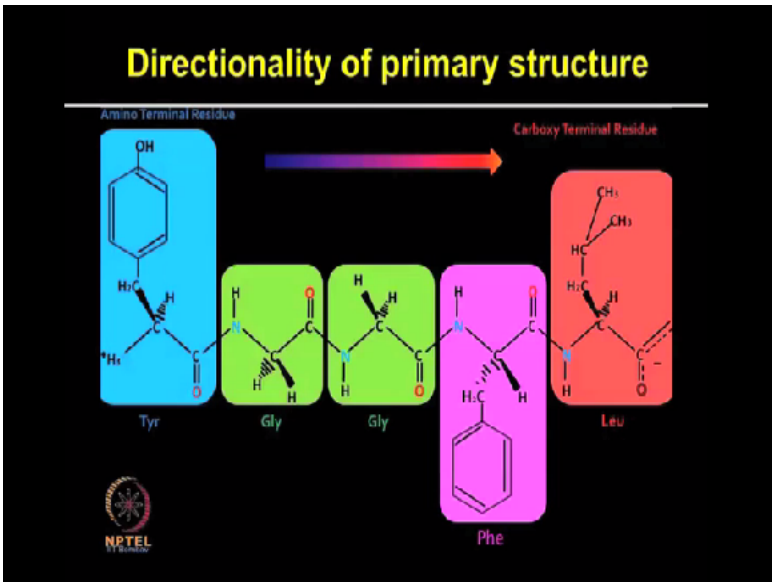
Now let us move on to different structural level of proteins starting with the primary structure. Amino acids constitute the basic monomeric unit of proteins and they are joined together by the peptide bonds. These 20 standard amino acids can be arranged in several ways and therefore, it can give rise to unique structural and functional properties. So primary structure refers to the sequence of amino acids.

(Refer Slide Time: 04:44)



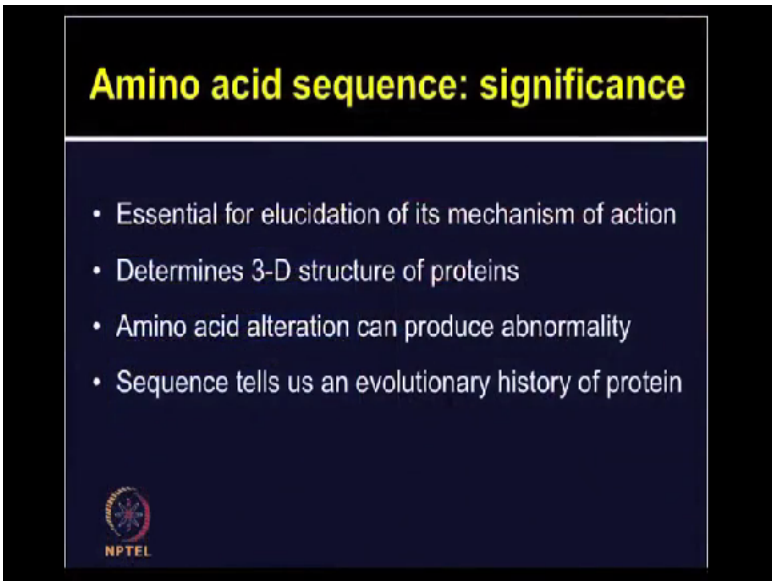
As you can see on the next slide different amino acids can come together and the linear sequence of amino acids constitute the primary structure with loss of water molecule.

(Refer Slide Time: 05:01)



Now what are directionality of primary structure. The polypeptide chain has polarity, so one end is alpha amino group and other end is alpha carboxyl group. The amino marks the beginning of any polypeptide chain. So what is the significance of amino acid sequence. As we discussed, this represents the primary structure.

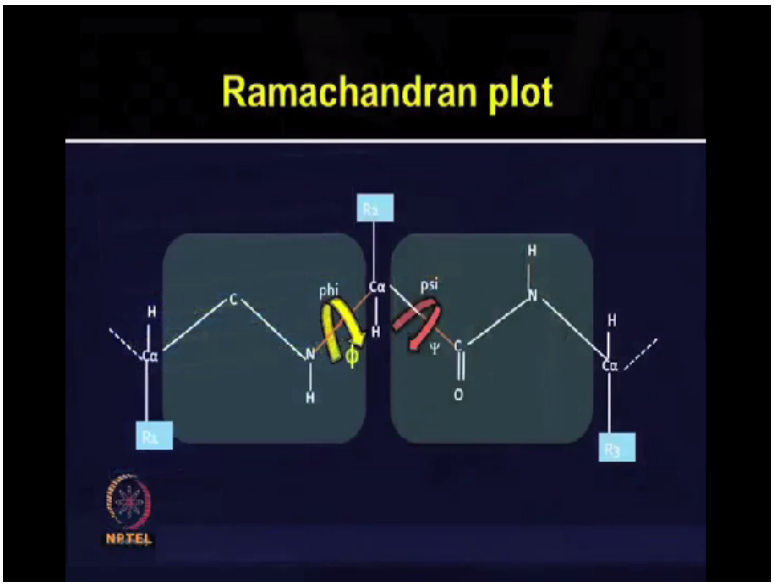
(Refer Slide Time: 05:29)



So amino acid sequence is essential for elucidation of its mechanism of action and enzyme's catalytic action can be determined. It determines 3-dimensional structure of the proteins linked in the functional 3-dimensional protein structure and the genetic message obtained from the DNA. The amino acid alterations can produce various type of disease abnormality, for example cystic fibrosis, change in only 1 amino acid can give rise to the abnormality.

And these amino acid sequences can also tell us various type of evolutionary aspect of the protein. So various type of information can be obtained from amino acid sequence.

(Refer Slide Time: 06:24)

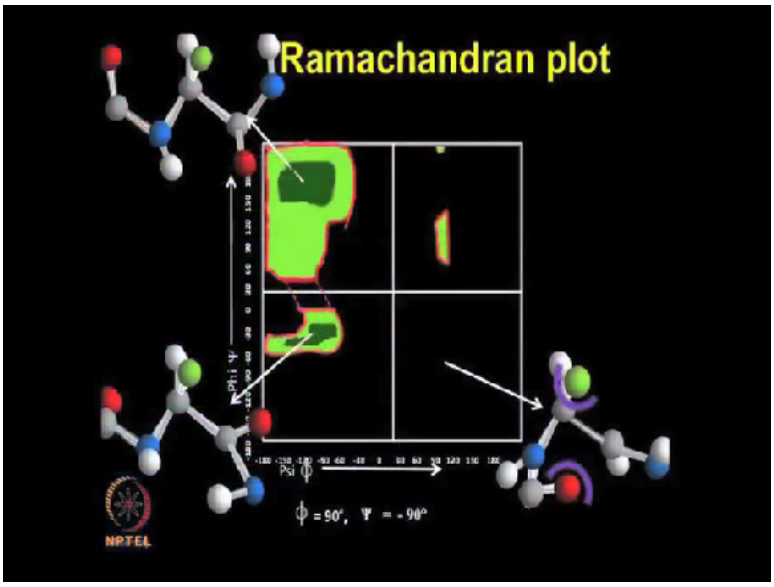


Now let us move on to concept of Ramachandran plot but before that, it is important to know

about the phi at the psi angles. As you can see here in the slide, rotation of 2 single bonds at the structure of each amino acids in polypeptides. The phi angle between the nitrogen and the alpha carbon atom and the psi angle between alpha carbon and carbonyl carbon. This phi and psi angles determine the path of polypeptide chain.

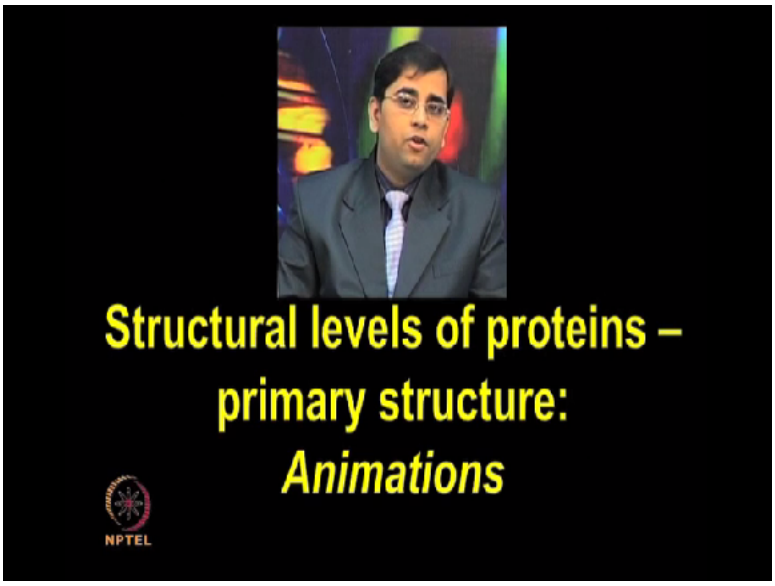
All the combination of phi and psi angles are not possible. So the allowed combinations can be viewed on 2-D plot which is known as a Ramachandran plot or Ramachandran diagram. There could be many possibilities of various type of combination but all the combinations are not allowed because of the steric hindrances and collision between the various atoms; therefore, steric exclusion can be a powerful governing factor for organising such plot.

(Refer Slide Time: 07:37)



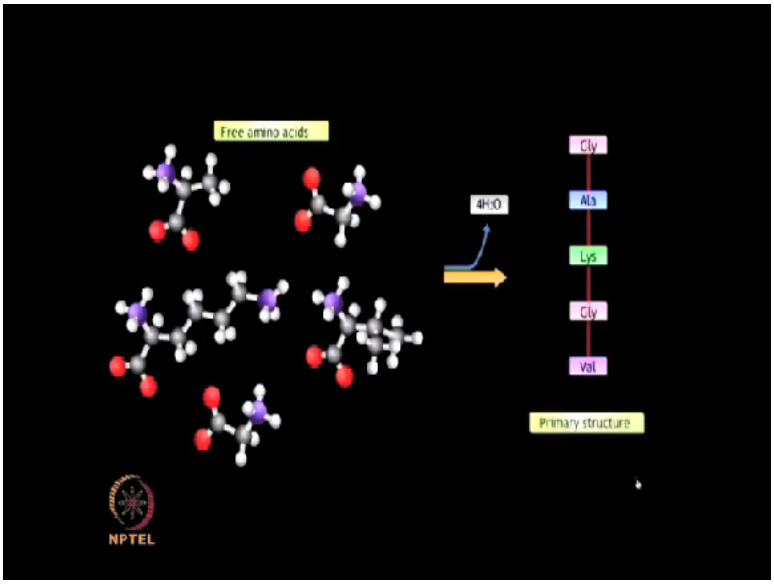
Now you can see it more clearly in the slide, the Ramachandran plot where more favourable regions are shown in the dark green colour and less favoured regions are shown in the light green.

(Refer Slide Time: 07:54)



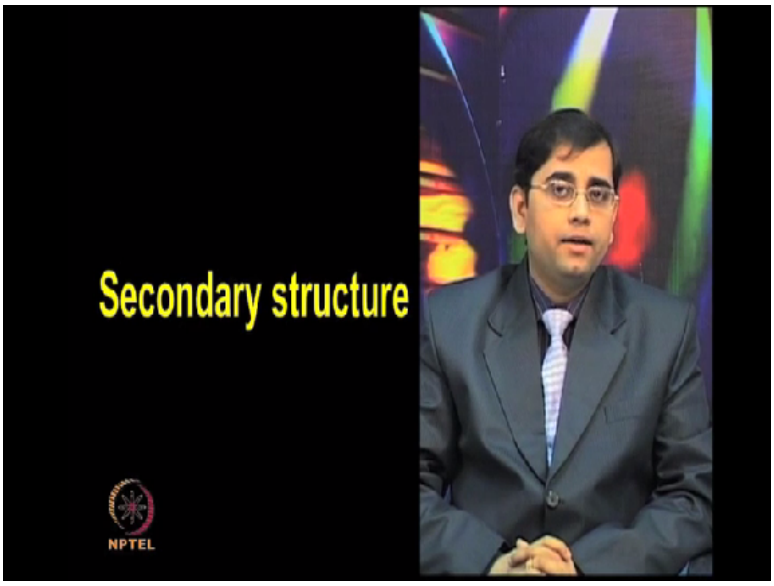
The structure level of proteins, the primary structure, few concepts will be discussed in following animation.

(Refer Slide Time: 08:04)



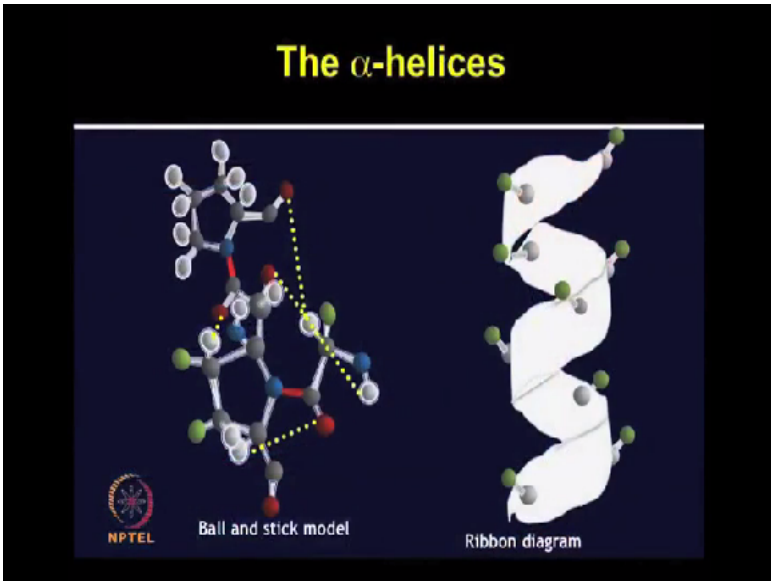
Amino acids are joined together in a head to tail arrangement by means of peptide bonds with the release of water molecules. This linear sequence of amino acids constitutes the primary structure.

(Refer Slide Time: 08:31)



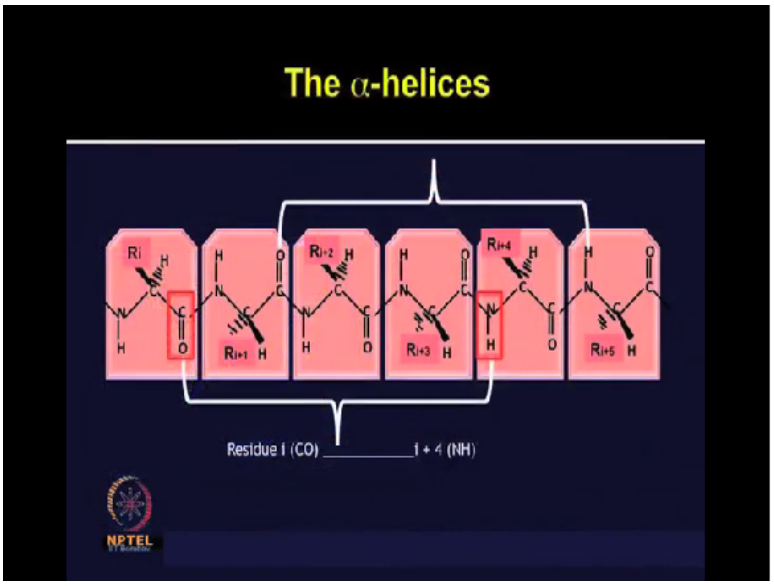
Let us now discuss about secondary structure which refers to locally folded regions. The folding of polypeptide or protein chain into regular structures like alpha helices, beta sheets, turns and loops, all these represent the secondary structures.

(Refer Slide Time: 08:49)



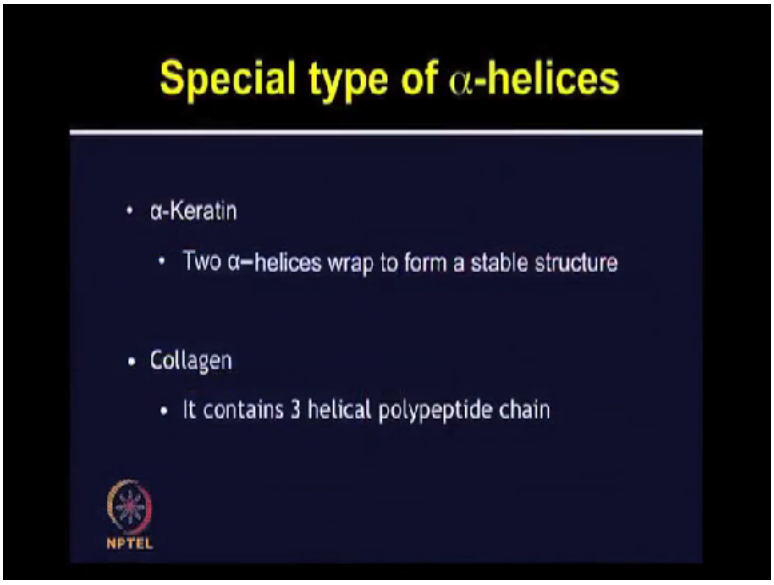
Let us first start with the alpha helices. So proteins have variable helix content. Alpha helix shows rod-like destruction. It has a main chain and a side chain. The main chain is tightly coiled around helical axis and a side chain is extended outward away from the helical axis as you can see in the ribbon diagram on the right side and the ball and a stick model on left side. Success with hydrogen bond stabilise these helical core.

(Refer Slide Time: 09:30)



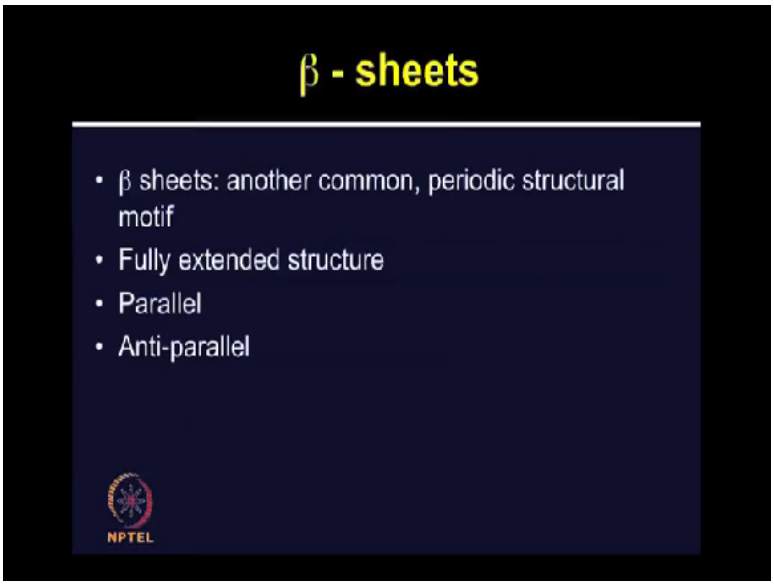
The alpha helix can be stabilised by the hydrogen bond. So the carbonyl group of each amino acid with NH group of amino acid which are 4 residues ahead in the sequence, they form these hydrogen bonds as you can see in this figure here. There are special type of alpha helices where 2 alpha helices can wrap up or 3 alpha helices can come together.

(Refer Slide Time: 10:02)



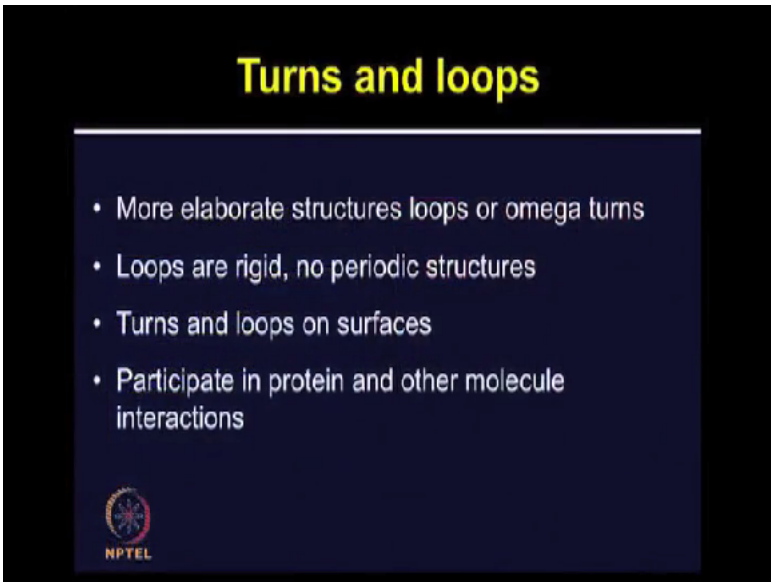
So the first example is alpha-keratin where 2 alpha helices can wrap to form a stable structure. It is (()) (10:14) component of hair and consists of 2 helical coiled around each other and forms a left-handed super helix which is known as a coiled coil. Another example is collagen which is fibrous component of skin, bone, etc. It is also most abundant in mammal. It contains 3 helical polypeptide chains.

(Refer Slide Time: 10:44)



Beta sheet is another common periodic structural motif which was discovered by scientist Pauling and Corey. It is fully extended structure unlike the tightly coiled alpha helices. It can be having 2 directions, the parallel or anti-parallel. Parallel when they are running in the same direction and anti-panel when they are running in opposing directions. Another category of secondary structures are turns and loops.

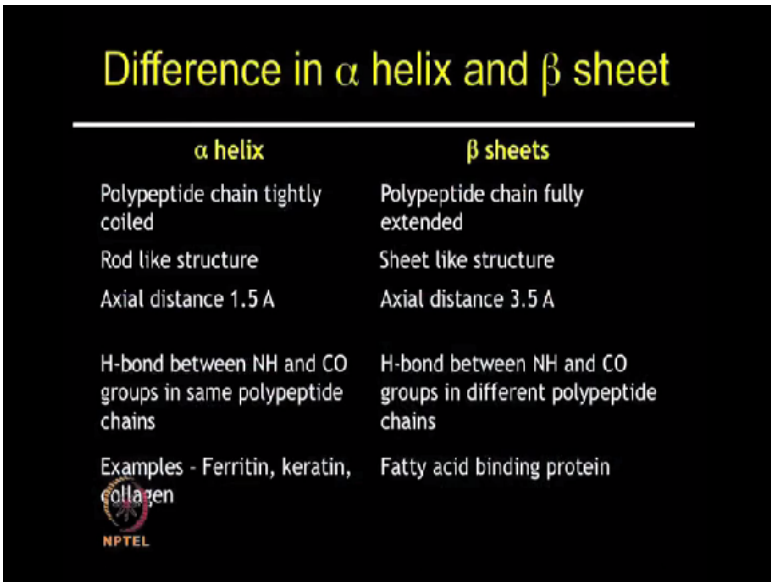
(Refer Slide Time: 11:18)



These are more elaborate structures; loops or omega turns which also perform chain reversal. Loops are rigid, well-defined. These are not periodic structures. The turns and loops are present on surface and participate in various important properties of proteins and other biomolecule

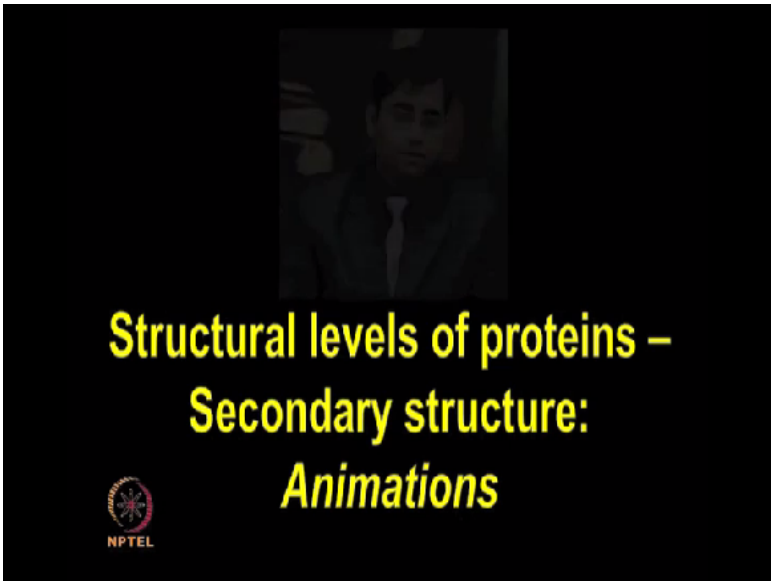
interactions. The differences in alpha helix and beta sheet is summarised in this slide.

(Refer Slide Time: 11:50)



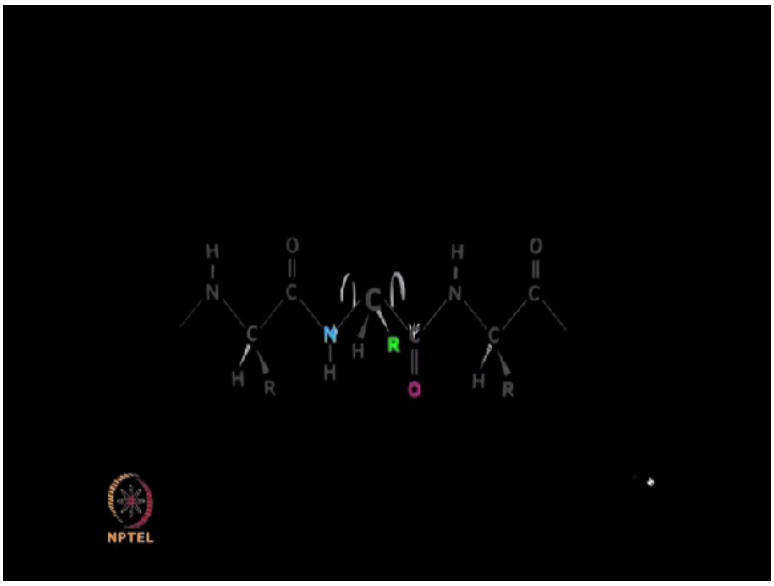
The alpha helix polypeptide chain is tightly coiled whereas in beta sheet, it is fully extended. Alpha helix rod-like structure and beta sheets sheet-like structure. The axial distance between adjacent amino acids is 1.5 angstrom whereas it is 3.5 angstrom in beta sheets. In alpha helix, hydrogen bond between NH and CO groups in same polypeptide chains whereas in beta-sheet, it is in different polypeptide chains. Alpha helix examples include ferritin, keratin, collagen, etc. In beta sheet, it is fatty acid binding protein.

(Refer Slide Time: 12:38)



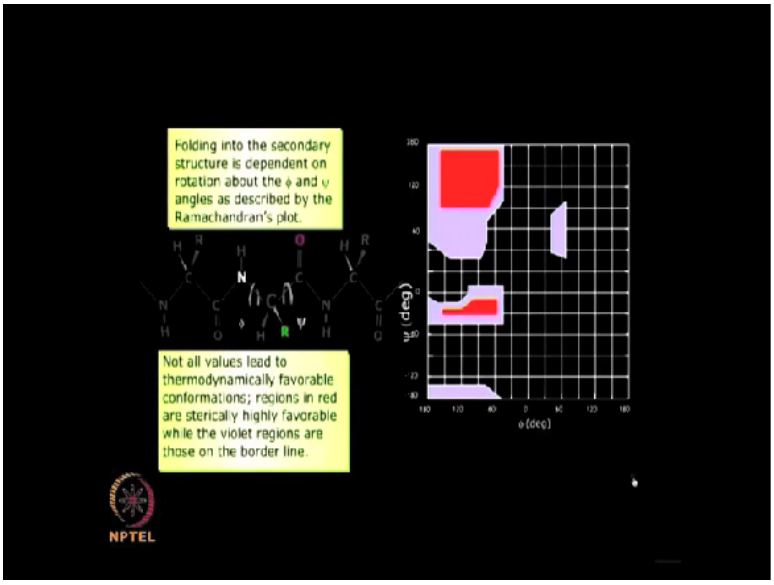
Some of the properties of secondary structures will be described in following animation.

(Refer Slide Time: 12:45)



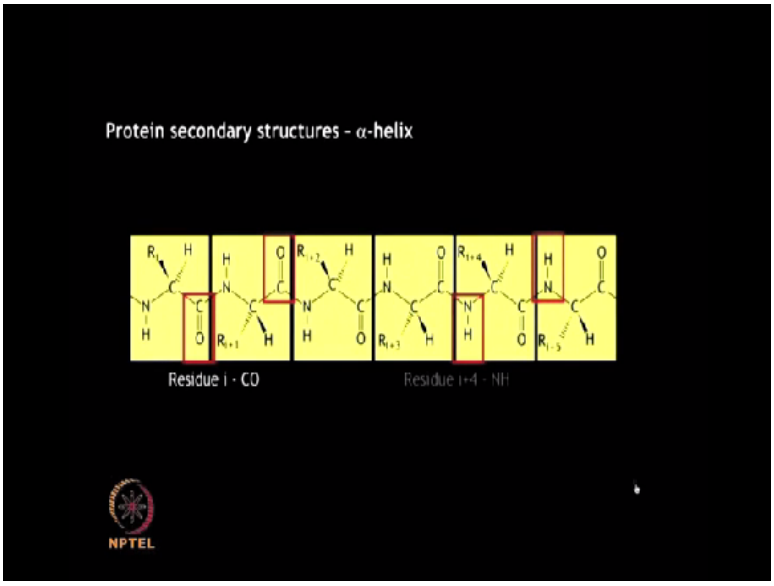
The folding of the primary structure into the secondary is governed by the permissible rotations about the phi and psi angles. Not all value of these angles lead to sterically favourable conformations.

(Refer Slide Time: 13:07)



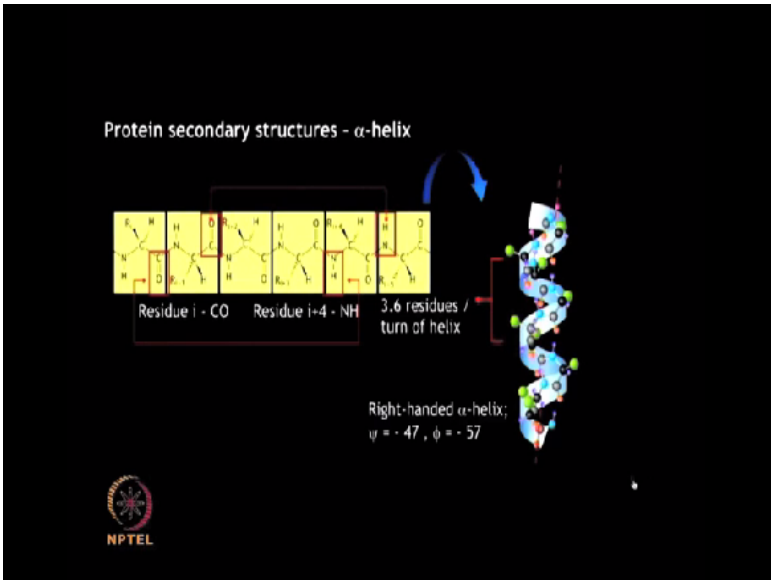
D. Ramachandran plot defines these regions of favourability.

(Refer Slide Time: 13:18)



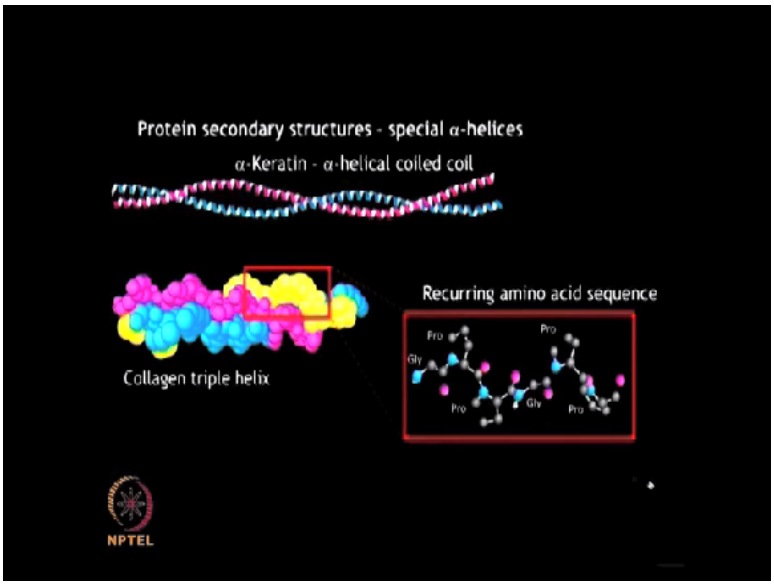
Amino acids along the polypeptide backbone interact through hydrogen bonds leading to the secondary structures. The alpha helix has intra-chain hydrogen bonds between the hydrogen of NH and oxygen of CO in every 4th residue. Most alpha helices are right-handed since the conformation is energetically more favourable.

(Refer Slide Time: 13:53)



Here you can see more energetically favourable alpha helix structure. The amino acid proline, which has a cyclic side chain, does not fit into the regular alpha helix structure and thereby limits flexibility of the backbone. It is commonly referred to as the helix breaker.

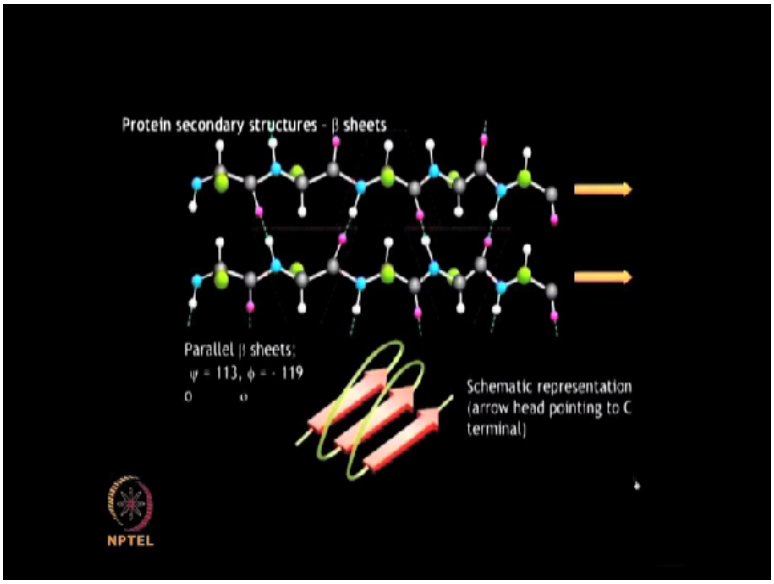
(Refer Slide Time: 14:23)



Alpha helices can also wind around each other to form stable structures such that their hydrophobic residues are buried inside while their polar side chains are exposed to the aqueous involvement. Alpha keratin, the major protein component of hair consists of 2 such coiled coils forming a left-handed super-helix. Collagen, which is the fibrous component of skin, muscles, etc. consist of 3 such coiled alpha helices.

It has a characteristic recurring amino acid sequence of glycine, proline, hydroxyproline with glycine appearing at every 3rd residue.

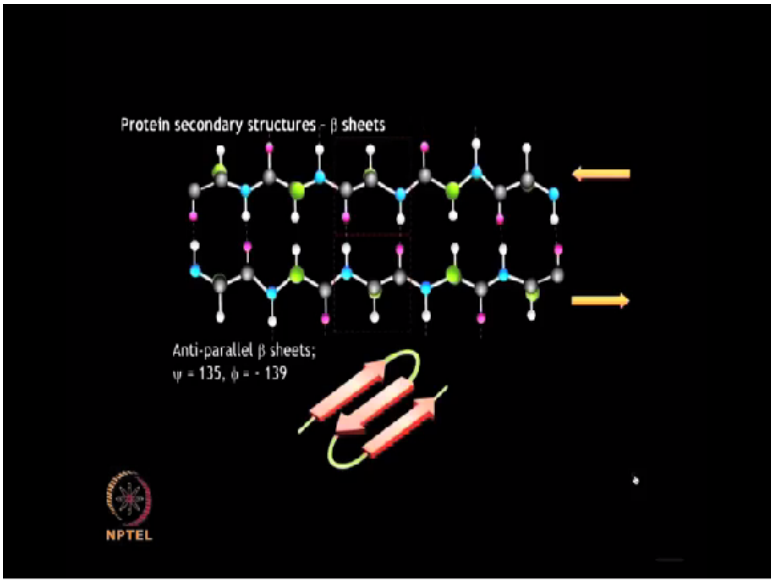
(Refer Slide Time: 15:26)



Beta-pleated sheets discovered by Pauling and Corey is another common secondary structure

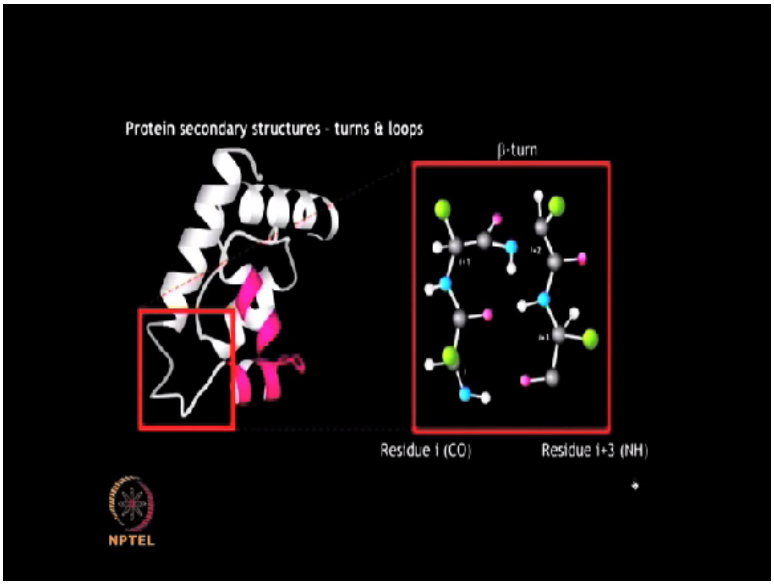
with periodic repeating units. It is composed of 2 or more polypeptide chains with their side chains oriented above and below the plane. It is an extended structure with hydrogen bonds between the chains stabilising it. Amino acids in parallel beta sheets which run in the same direction interact with 2 different amino acids on their adjacent screen through the hydrogen bonds.

(Refer Slide Time: 16:24)



Amino acids in anti-parallel strands on the other hand interact with only one amino acid or an adjacent strand.

(Refer Slide Time: 16:43)



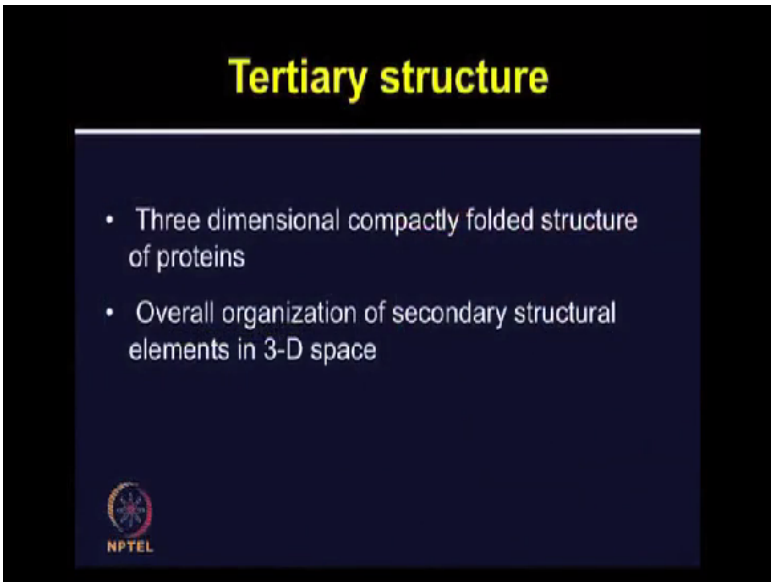
Almost all proteins exhibit a compact globular structure which is possible only if there are turns

or loops between the various regions. Beta turns which are the most commonly observed turn structures consist of rigid well-defined structures that usually lie on the surface of the protein molecule and interact with other molecules.

(Refer Slide Time: 17:19)

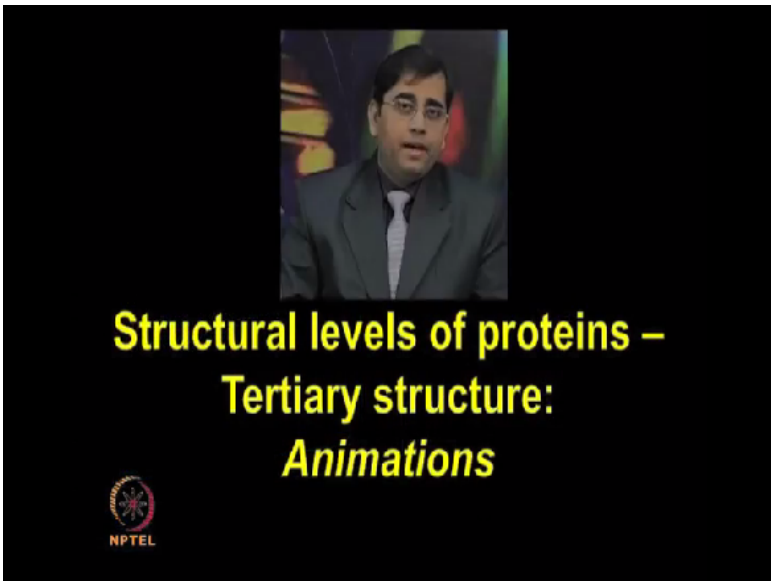
Let us now move on to tertiary structure which refers to overall folded structure. NMR and x-ray crystallography provides detailed 3-dimensional structures.

(Refer Slide Time: 17:32)



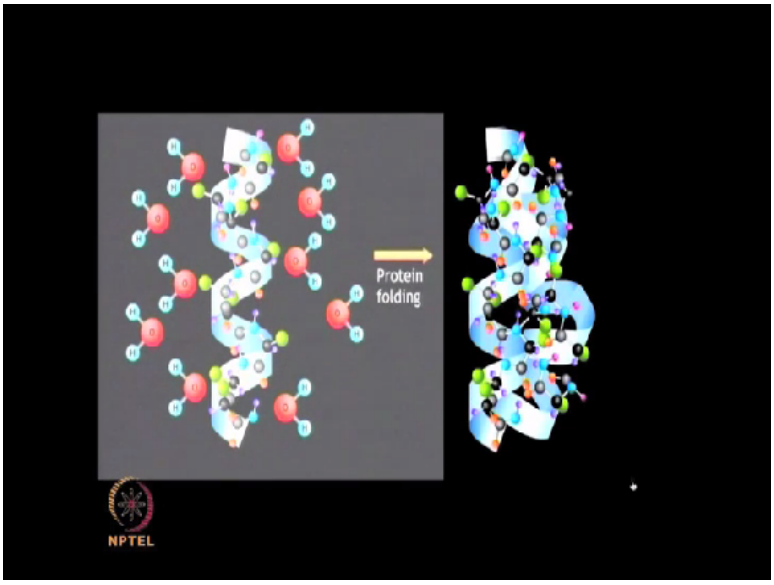
The 3-dimensional structure as compactly folded structure of proteins and it represents overall organisation of secondary structural elements in 3-D space. There are numerous interactions which stabilise the tertiary structure of proteins.

(Refer Slide Time: 17:52)



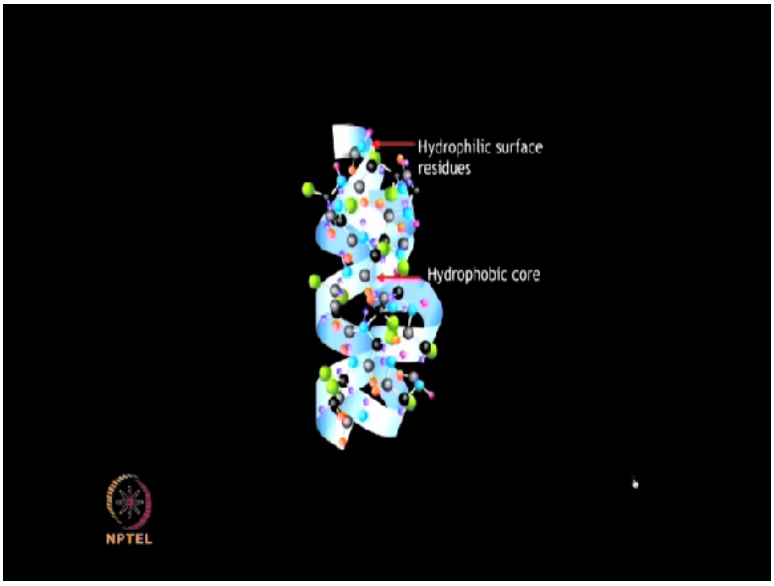
We will take myoglobin as an example and describe some of the properties of tertiary structure in following animation.

(Refer Slide Time: 18:03)

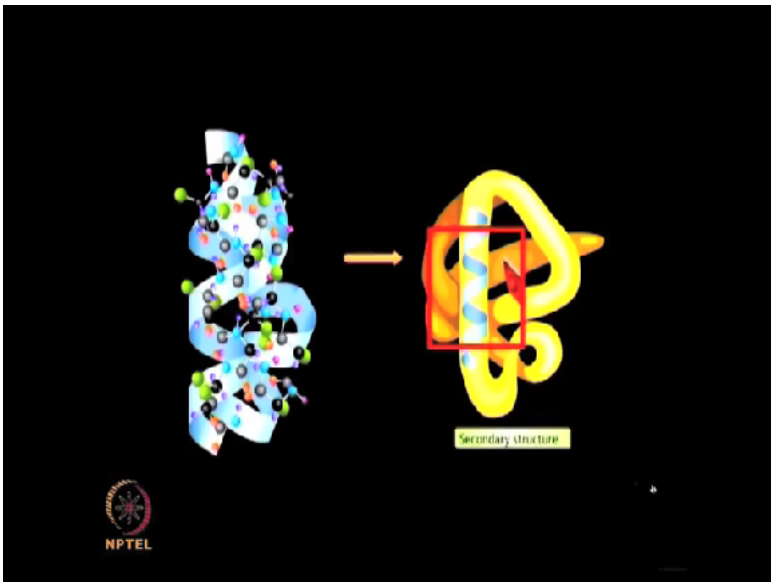


Amino acids located far apart on the polypeptide chain interact with each other by means of hydrogen bonds, electrostatic interactions, disulphide bridges, etc. which allows the protein to fold 3-dimensionally in the space giving rise to the tertiary structure.

(Refer Slide Time: 18:35)

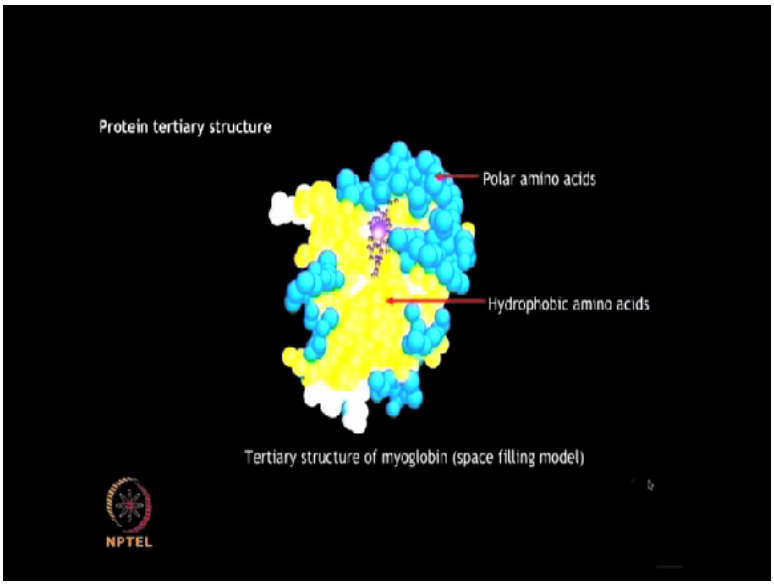


(Refer Slide Time: 18:53)



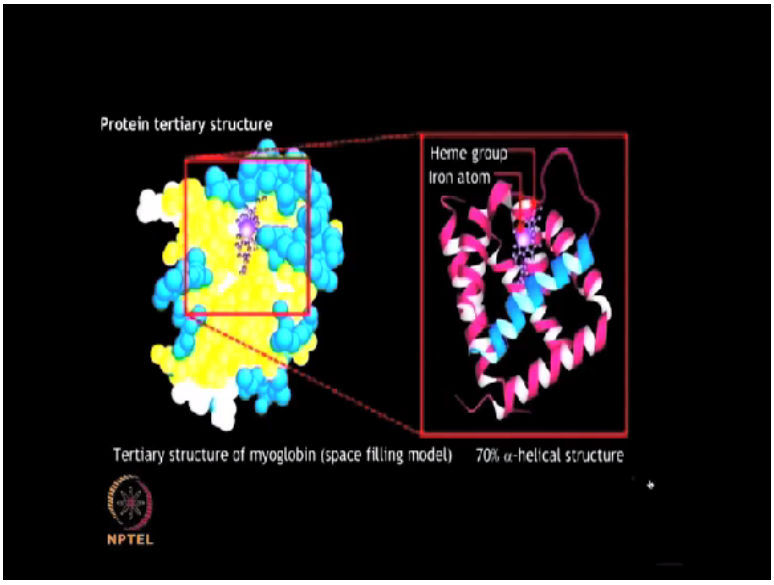
Protein folding takes place such that the hydrophobic residues are buried inside the structure while the polar residues remain in contact with the surroundings.

(Refer Slide Time: 18:57)



The tertiary structure of myoglobin determined by John Kendrew clearly revealed that the nature of amino acid side chains dictate their location in the tertiary structure. The hydrophobic residues are found buried inside the structure while the polar amino acids are present in the surface.

(Refer Slide Time: 19:25)

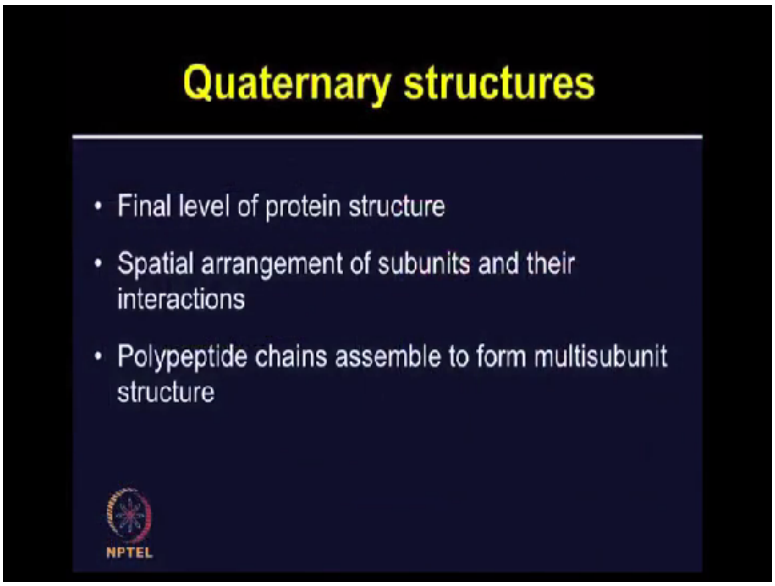


70% of main chain of myoglobin is folded into alpha helices with the rest being present in the form of turns and loops which are essential to provide it the compact structure.

(Refer Slide Time: 19:52)

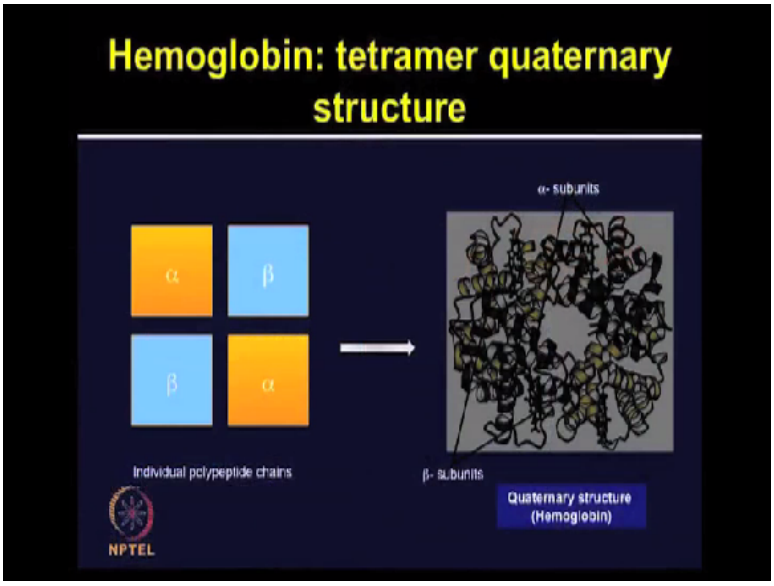
What is quaternary structure. It refers to interaction between individual protein subunits in a multi subunit complex.

(Refer Slide Time: 20:01)



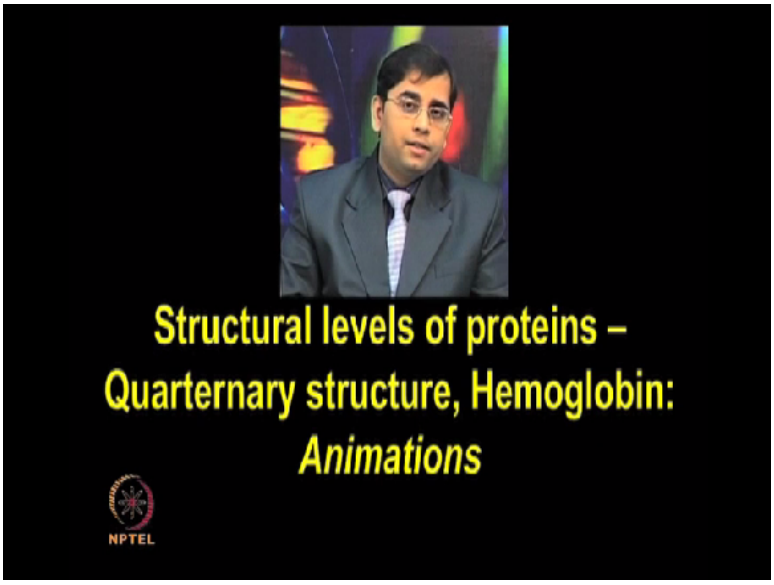
So quaternary structures represent final level of protein structure which has spatial arrangement of subunits and their interactions. The polypeptide chains assemble to form multi-subunit structure and each polypeptide chain is known as subunit. The different examples such as DNA binding cro protein of bacteriophage lambda which is representing simplest quaternary structure. Then you have classical example of haemoglobin.

(Refer Slide Time: 20:38)



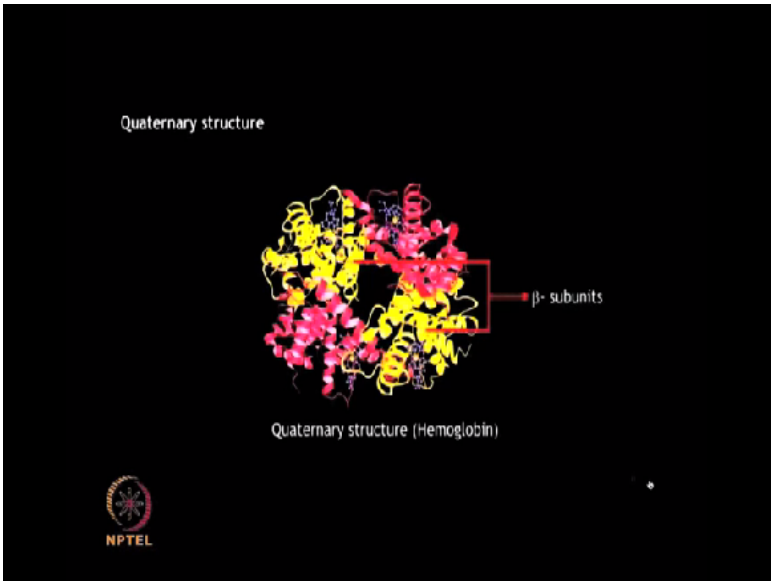
So let us look at haemoglobin, the tetramer quaternary structure which has 2 alpha subunits and 2 beta subunits. The individual polypeptide chains, they come together and form quaternary structure.

(Refer Slide Time: 20:55)



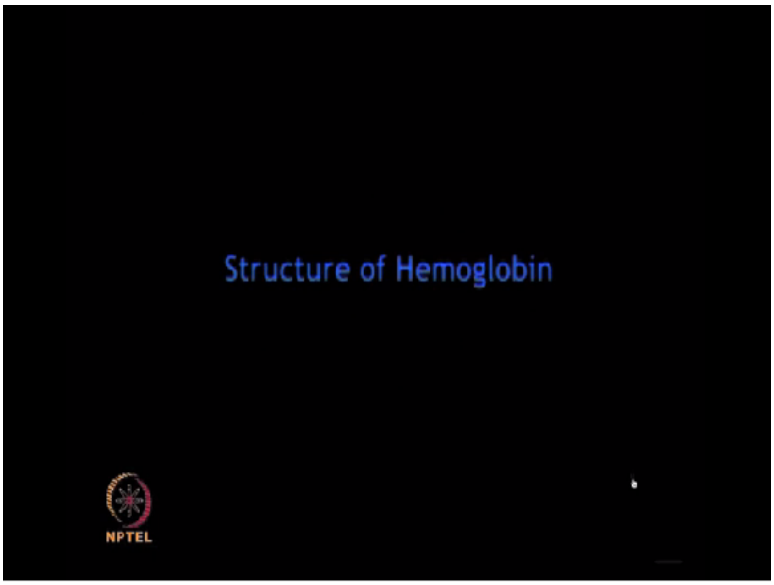
Properties of quaternary structures and some details about haemoglobin and comparison of myoglobin and haemoglobin will be discussed with following animation.

(Refer Slide Time: 21:08)

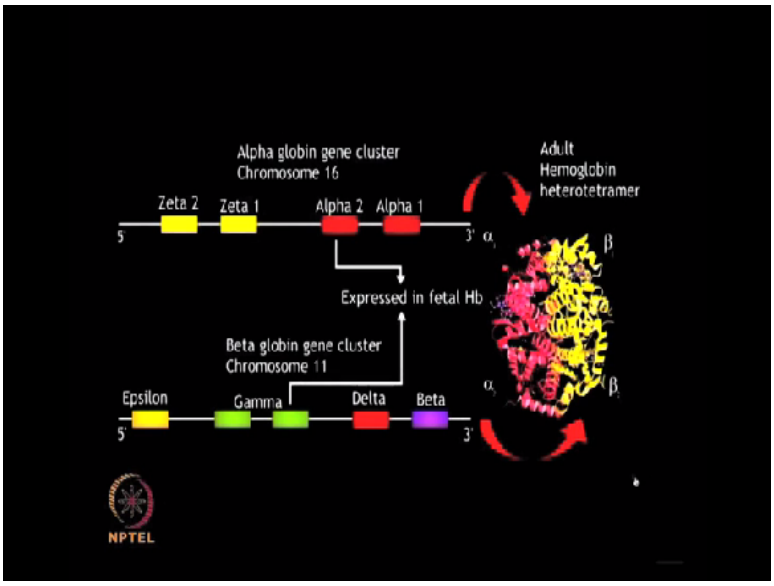


Different subunits or polypeptide chains interact with one another and are held together by means of ionic, electrostatic van der Waals, etc. interactions. Such multi-subunit proteins having a quaternary structure which is the final level of protein structure.

(Refer Slide Time: 21:46)

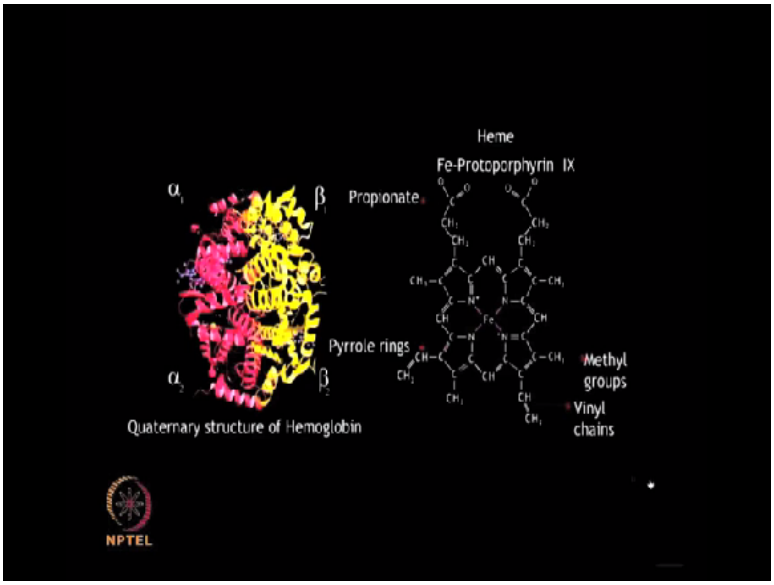


(Refer Slide Time: 21:48)



Hemoglobin is a heterotetramer composed of 2 alpha and 2 beta chains. The alpha globin gene locus resides on chromosome 16 with each gene contributing to the synthesis of the alpha globin chain. The beta globin gene locus resides on chromosome 11 and consist of all genes that are expressed from the time of embryonic development of haemoglobin to that of adult beta globin gene. The globin chains are synthesised by ribosomes in the cytosol.

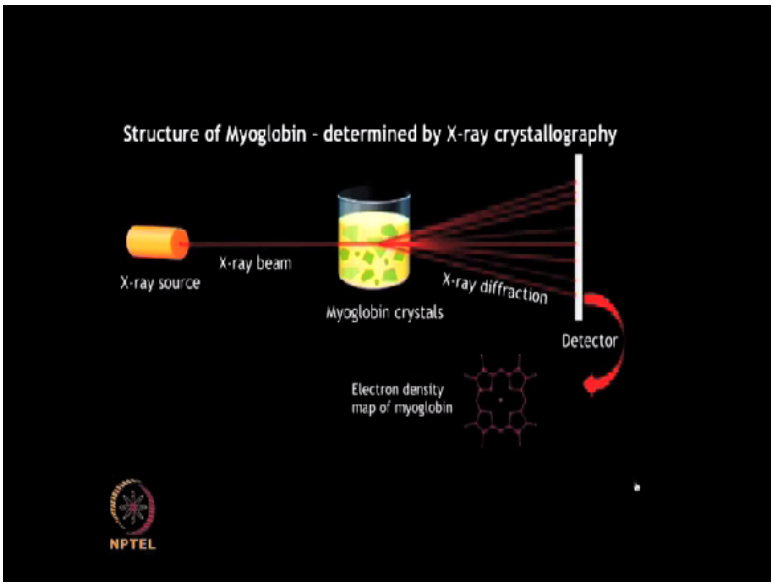
(Refer Slide Time: 23:00)



Every subunit of haemoglobin is bound to the prosthetic group known as heme. This consist of a central iron atom in its ferrous straight surrounded by a complex organic ring structure known as protoporphyrin. The heme group is essential for the oxygen-binding property of haemoglobin. The iron atom forms 6 coordinate bonds, 4 of which are to the nitrogen atom of protoporphyrin

ring, 1 to histidine side chain in the global subunit and the other being the binding site for oxygen.

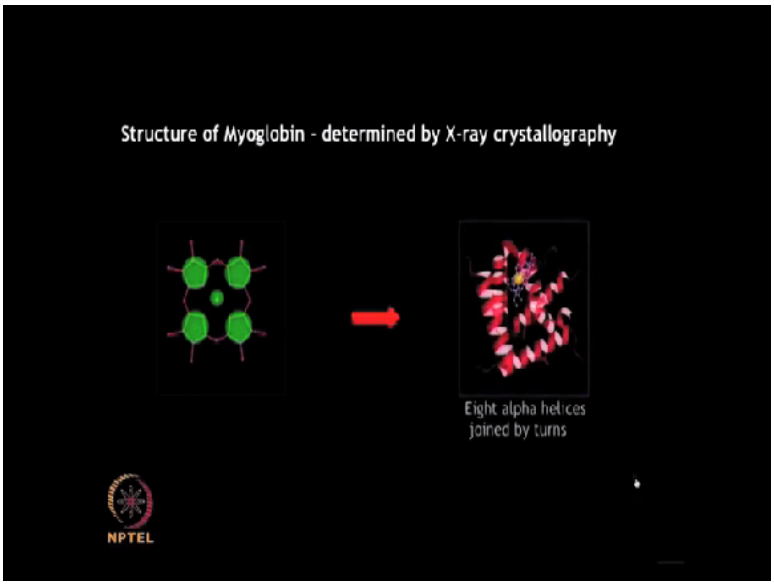
(Refer Slide Time: 24:05)



X-ray crystallography is a very useful visualisation technique that facilitates the determination of 3-dimensional coordinates of atoms in a protein. Myoglobin was the first protein whose structure was determined by x-ray crystallographic studies. When a beam of x-ray was passed through the crystals of myoglobin, some part of the beam was found to pass straight through while the others were scattered in different directions.

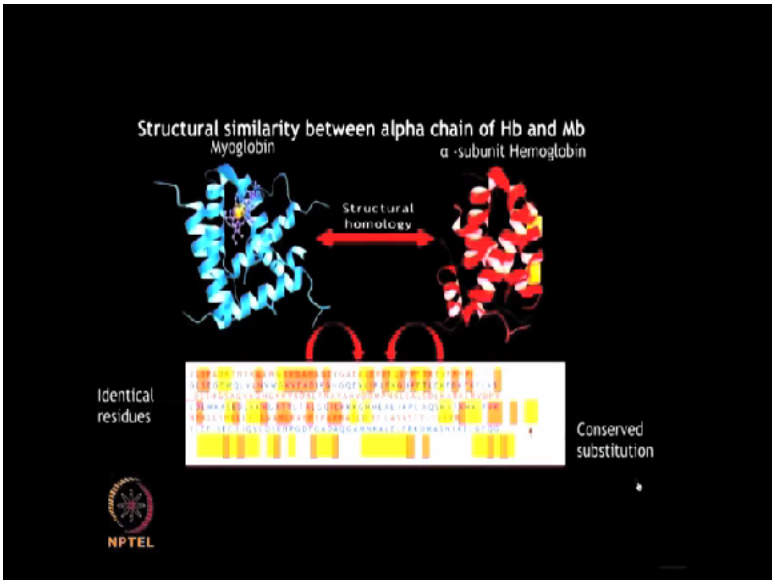
These scattered beams were detected by means of an x-ray film. After the spot intensity calculations, it provided an electron density map of the protein.

(Refer Slide Time: 25:20)



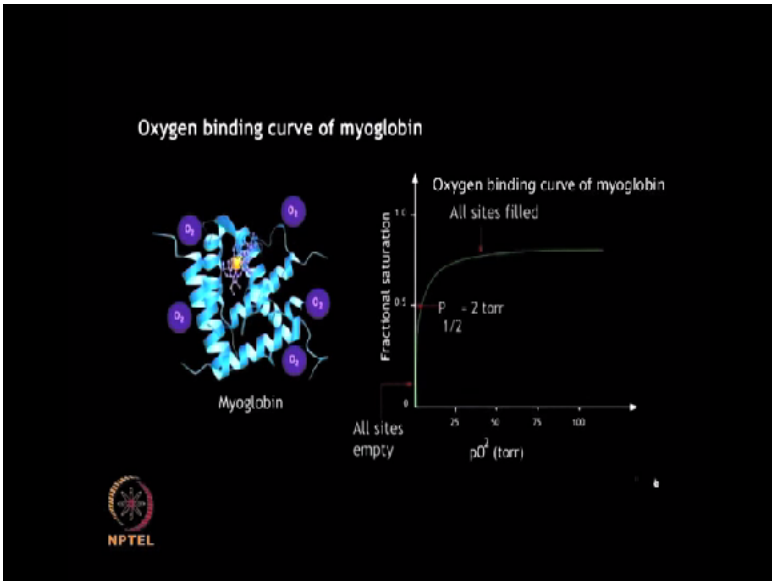
The protein was found to consist of a single polypeptide chain having 8 alpha helices along with a heme group in the centre similar to the haemoglobin.

(Refer Slide Time: 25:46)



Myoglobin found largely in muscle tissues has been found to be structurally similar to the alpha subunit of haemoglobin. The alpha helix arrangement of both proteins has been found to be the same with the recurring structures being known as global fold. The haemoglobin chain having 141 amino acids and myoglobin having 153 residues have also been found to have very high sequence homology. Despite the similarities, their oxygen binding capacities are different.

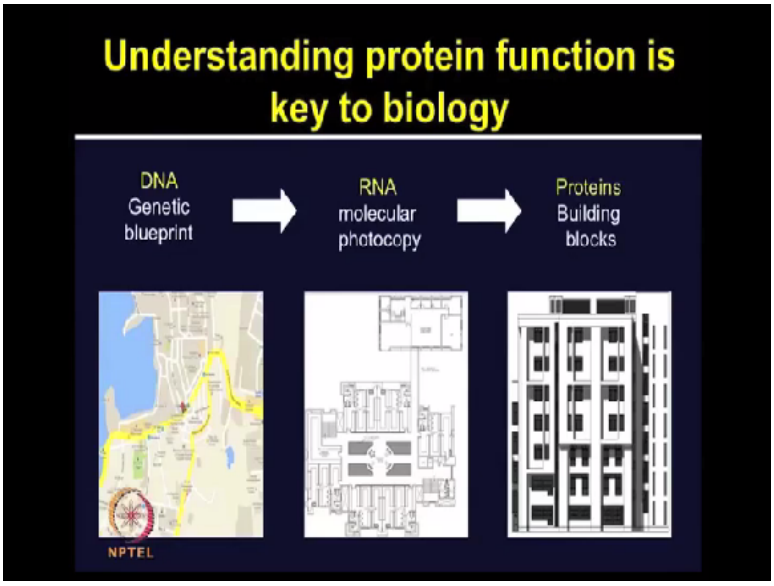
(Refer Slide Time: 26:55)



Myoglobin functions largely as an oxygen binding protein that stores a regional supply of oxygen in muscle tissues while haemoglobin serves to transport oxygen. Myoglobin binds strongly to the oxygen and acts as an oxygen storage protein rather than a transporter. It shows 50% saturation at a pressure as low as 2 torr and get saturated even under low oxygen pressure conditions that prevail in the muscle.

Myoglobin can use only 7% of the oxygen carrying capacity as opposed to haemoglobin which can utilise nearly 90% of the oxygen carrying capacity. Unlike haemoglobin which has a sigmoidal oxygen binding curve, myoglobin has a hyperbolic which indicates that it binds to oxygen irrespective of a surrounding partial pressure of oxygen in the tissue.

(Refer Slide Time: 28:28)

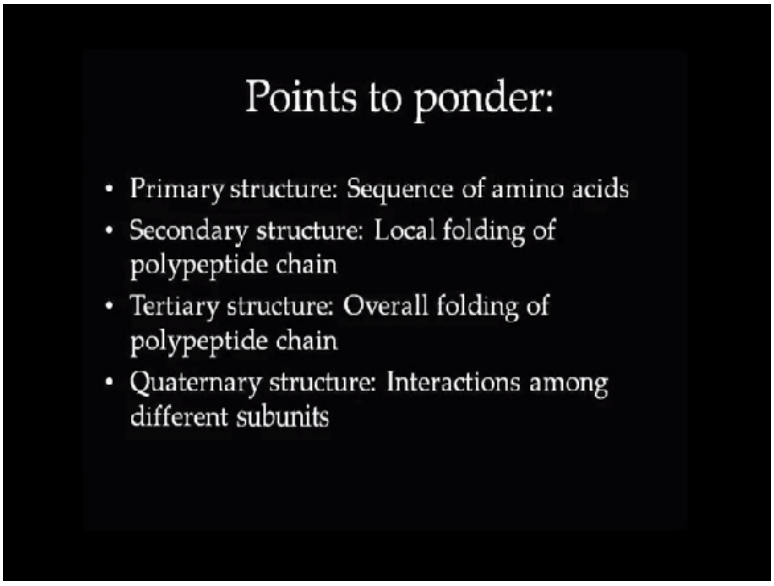


After discussing various properties of amino acids and different level of structural proteins, primary, secondary, tertiary and quaternary, now let us just touch upon why understanding protein function is key to the biology. So diseases result as due to the protein malfunction; therefore, all the current drugs, they are either targeting the protein function or they are protein themselves which demonstrates the significance of studying about proteins.

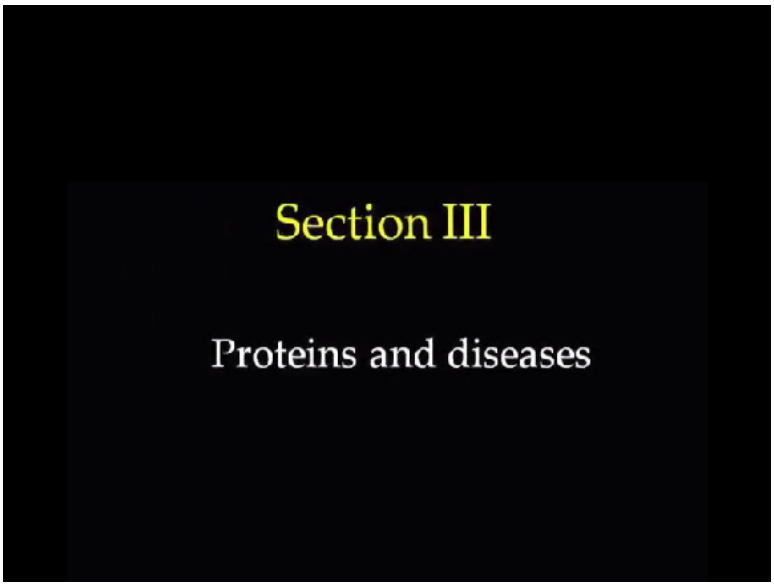
Let me describe you this slide where I have shown the various biomolecules of central dogma, DNA, RNA and protein. For example, if you look at the map here where IIT is located in Powai area, so it is like DNA which is the genetic blueprint. It contains only information. Now if you want to make a building in this area, it is you have to define that area which is like iron molecule which is a molecular photocopy and it is used on the site of construction by the such contractors.

Now proteins are like the building which you want to create on that site. These are the building blocks or building materials. They are bits and mortars of engines of biology. So it shows you how various type of biomolecules have their significance but it is the protein which ultimately defines the function.

(Refer Slide Time: 30:23)

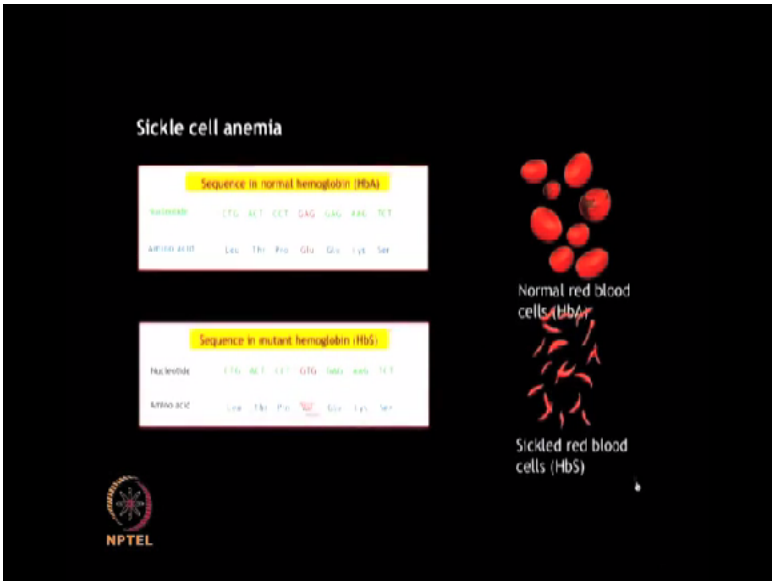


(Refer Slide Time: 30:43)



The malfunction of proteins can result into various diseases and so the disease will be described in following animation.

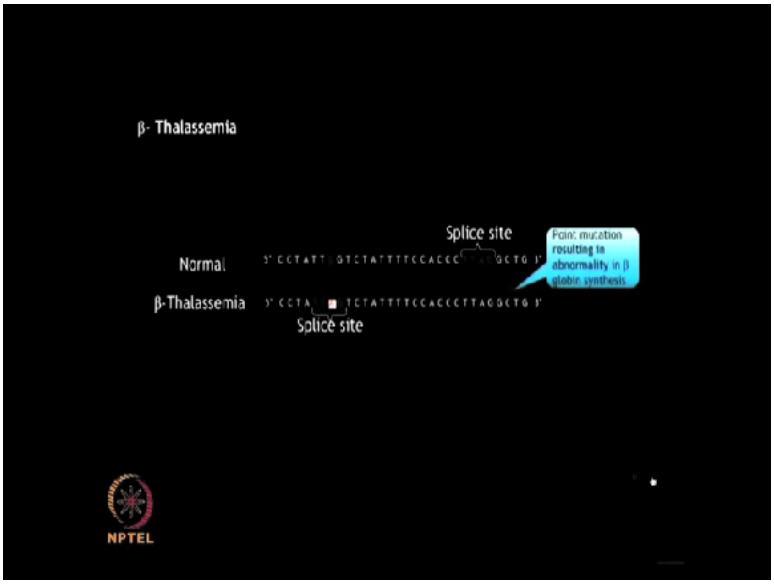
(Refer Slide Time: 30:57)



A large number of mutations have been described in the global gene and mutation causing sickle cell anaemia is a single nucleotide substitution of A-T in the codon for amino acids sets of the beta chain. This change converts a glutamic acid residue to valine in the corresponding amino acid sequence.

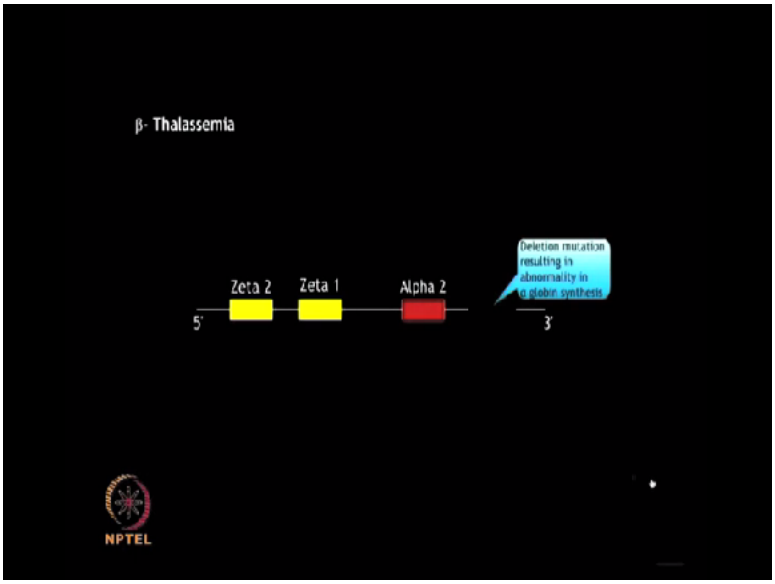
Replacement of the glutamine residue by valine creates a sticky hydrophobic contact point at position 6 of the beta chain. These sticky spots cause deoxyhemoglobin S molecules to associate abnormally with each other leading to clumping of the cells. Their oxygen carrying capacity is greatly reduced and these patients require frequent transfusion.

(Refer Slide Time: 32:16)



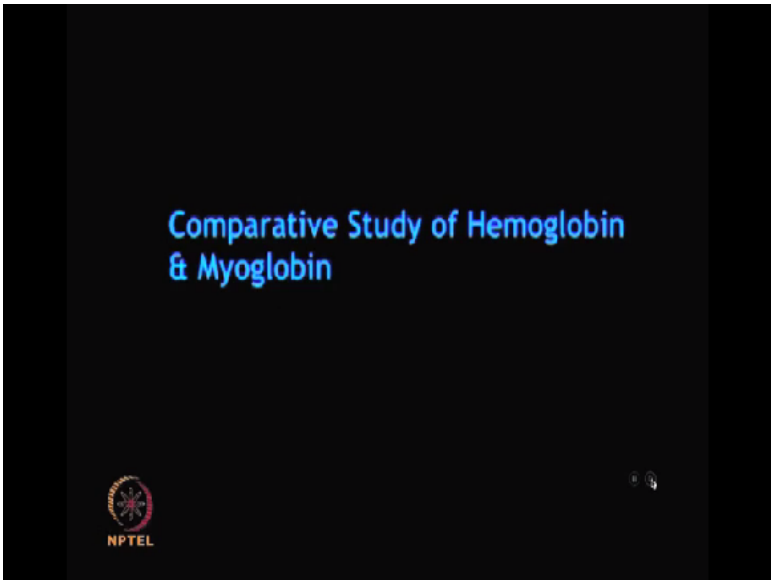
Thalassemia is a result of abnormalities in haemoglobin synthesis. Deficiency in beta globin synthesis results in beta thalassemia. Mutation of a single base from G-A in an intron of beta globin gene generates a new splice site. The resulting mRNA contains a stop codon further upstream and leads to premature translation termination thereby producing aberrant protein.

(Refer Slide Time: 33:08)

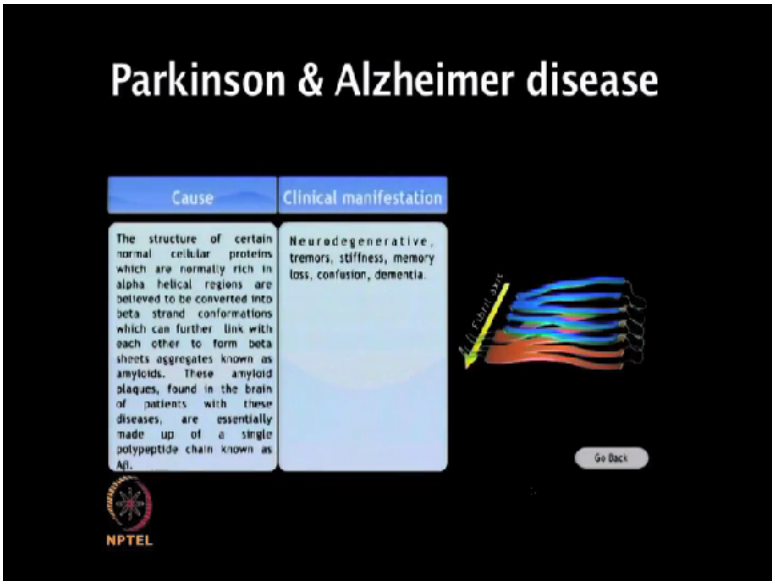


Deficiencies in alpha global synthesis due to inactivation of one or all the 4 alpha globin gene results in the alpha thalassemia.

(Refer Slide Time: 33:26)



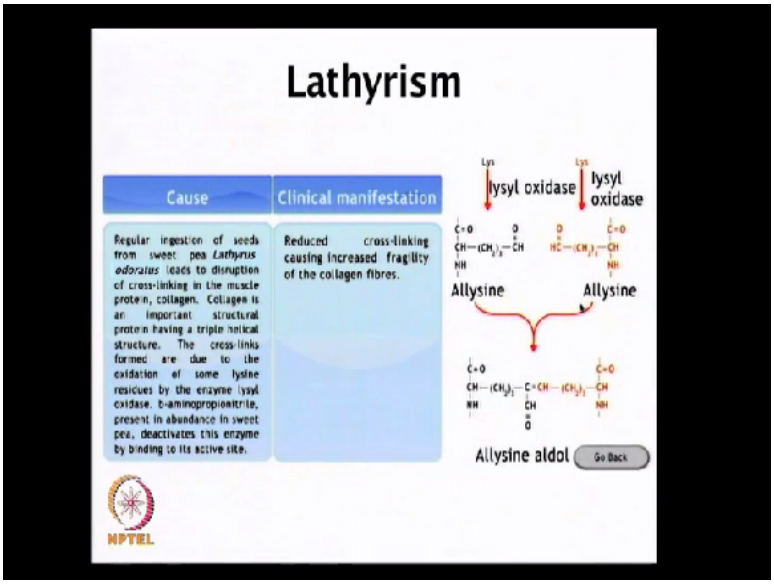
(Refer Slide Time: 33:27)



Parkinson and Alzheimer disease. The structure of certain normal cellular proteins which are normally rich in alpha helical regions are believed to be converted into beta-strand conformations which can further link with each other to form beta sheet aggregates known as amyloids. These amyloid plates are found in the brain of patients with these diseases, are essentially made up of a single polypeptide chain.

The clinical manifestation includes neurodegenerative, tremors, stiffness memory loss, confusion, dementia, etc.

(Refer Slide Time: 34:19)

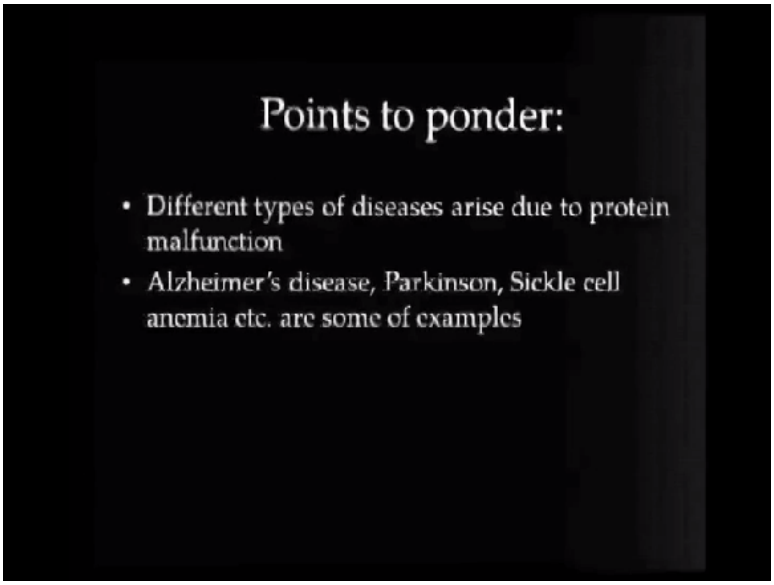


Lathyrism cause regular ingestion of seeds from sweet pea, Lathyrus odoratus which leads to

disruption of cross-linking in the muscle protein collagen. Collagen is an important structural protein having a triple helical structure. The cross links formed due to the oxidation of some lysine residue by the enzyme Lysyl oxidase. Beta-aminopropionitrile present in abundance in sweet pea deactivates this enzyme by binding to its active site.

A clinical manifestation includes reduced cross-linking causing increased fragility of the collagen fibers.

(Refer Slide Time: 35:16)



(Refer Slide Time: 35:24)

In summary, today we talked about proteins and its function. We have talked about different

levels of protein structures, primary, secondary, tertiary and quaternary. We discussed in little more detail about myoglobin and haemoglobin, the model proteins and then briefly we touched upon significance of studying proteins and its malfunction may result in to various diseases. We will continue our discussion on some basic concepts of proteins in the next class. Thank you.

(Refer Slide Time: 36:00)

(Refer Slide Time: 36:05)