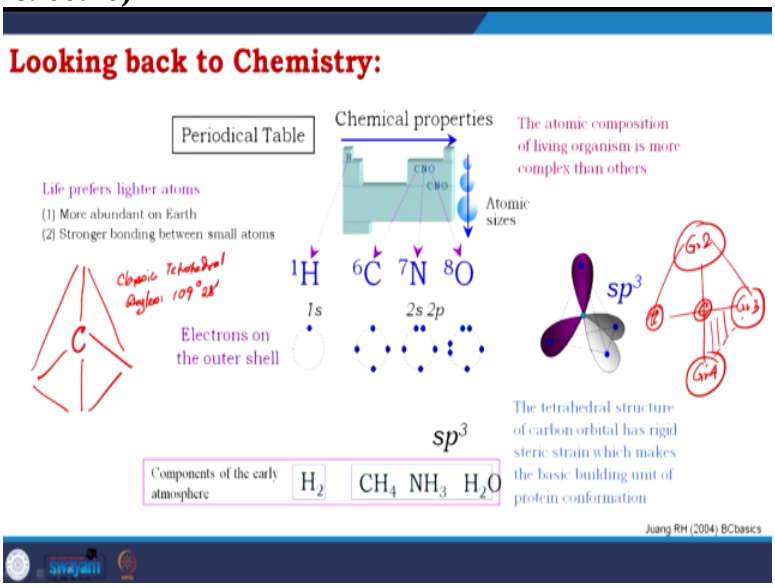
Structural Biology
Prof. Saugata Hazra
Department of Biotechnology
Indian Institute of Technology - Roorkee

Lecture - 07
Protein: Protein Chemistry, Chirality, Peptide Bond and Levels of Protein Structure

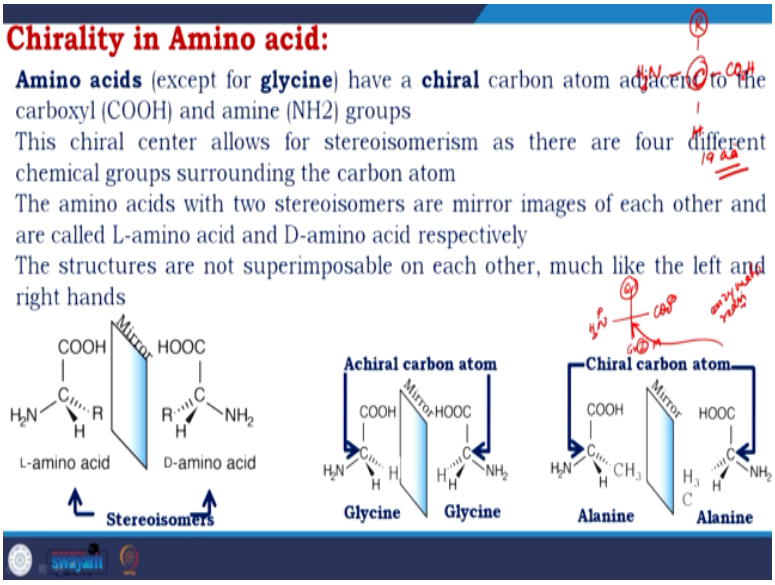
Welcome to the structural biology course again today. We are continuing our discussion on protein chemistry, chirality, peptide bonds, and the level of different protein structures.

(Refer Slide Time: 00:40)



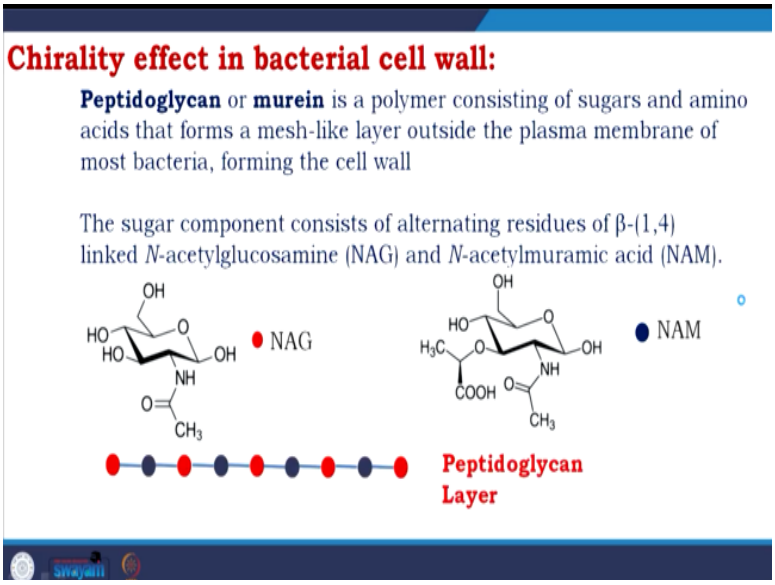
In the last class, I talked about the chemistry, the position of the atoms in the periodic table, the involvement of hydrogen, carbon, nitrogen, oxygen as well as sulfur and phosphorus, we have looked at their electronic outer cells, and we also have seen that how they develop sp^3 hybridization to develop a tetrahedral structure.A very interesting phenomenon we discussed is called chirality.

(Refer Slide Time: 03:43)



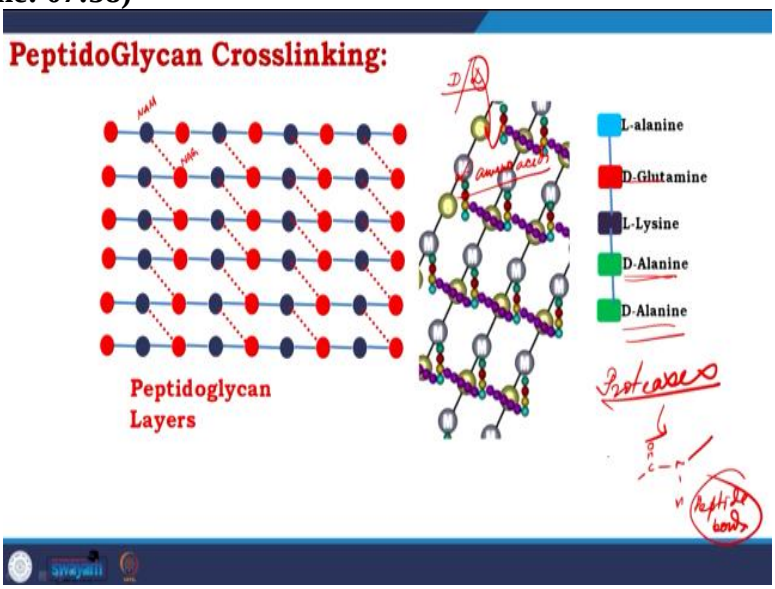
And amino acids have chirality because it has carbon, a COOH group, an NH_2 or NH_3 group it has hydrogen, and 19 of the amino acids among 20 have an R different than hydrogen. Only glycine has hydrogen. Amino acids, except for glycine, have a chiral carbon atom adjacent to the carboxyl and amino groups. This chiral center allows for stereoisomerism as four different chemical groups surround the carbon atom. The amino acids with two stereoisomers are mirror images of each other and are called L amino acids and D amino acids. So if you see their stereoisomers, they could not be superimposed. How or why this chirality is affecting is what you have to understand. So, if you see two groups, one is above the plane, and another is below the plane. When some ligand approached this carbon, it had to come on the side of the above or below the plane. So, if you think it is a chemical reaction, the most likely side is having less steric. So, because of the chirality, there is a directionality in the enzymatic reaction, which would contribute to the reaction specificity. And I would follow this with a very interesting story, which will help you understand the importance of chirality.

(Refer Slide Time: 07:08)



Peptidoglycan or murein is a polymer consisting of sugars and amino acids that form a mesh-like layer in a network outside that plasma membrane of most bacteria forming the major component of the cell wall. The sugar component consists of altering residues of beta 1, 4 linked N-acetyl glucosamine called NAG and N-acetyl muramic acid called NAM. So, this is the structure of N-acetyl glucosamine, and this is the structure of N-acetyl muramic acid, they form linear chains, which are known as peptidoglycan layers.

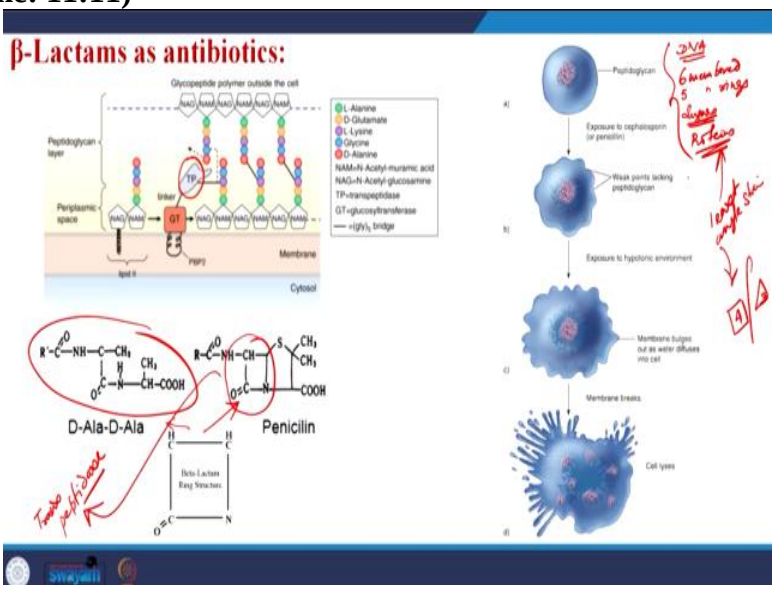
(Refer Slide Time: 07:58)



There are many such layers in the cell wall, but the layer would be strengthened only when cross-linked. So, the layers are cross-linked NAM to NAG, but it could be the opposite way too, and a pentapeptide does this cross-link. They might vary, but mostly they are L alanine, D glutamine, L lysine, D alanine, and D alanine. I talked about that the amino acids could be D or L, and they normally have the L amino acids, and when they cross-link, gives very good stability to the cell wall.

So, what I want to tell here is that if the amino acids are in L, then why do you see different or opposite chirality. This is because in the cell, there are proteases, they cut the bond between which we are going to talk about two amino acids called peptide bonds, but with the presence of opposite chirality, they are not supposed to cut it. So nature understands the importance of this cross-linking in the survival of bacteria. So they have incorporated opposite chirality so that the common proteases could not cut this. So that is what I was talking about: chirality is extremely important towards biochemical reactivities and enzymatic activities. There are enzymes, which are specific to some conformers like L amino acids, there are some that have broad specificity. And depending on those differences, many different biochemical biophysical phenomena are altered.

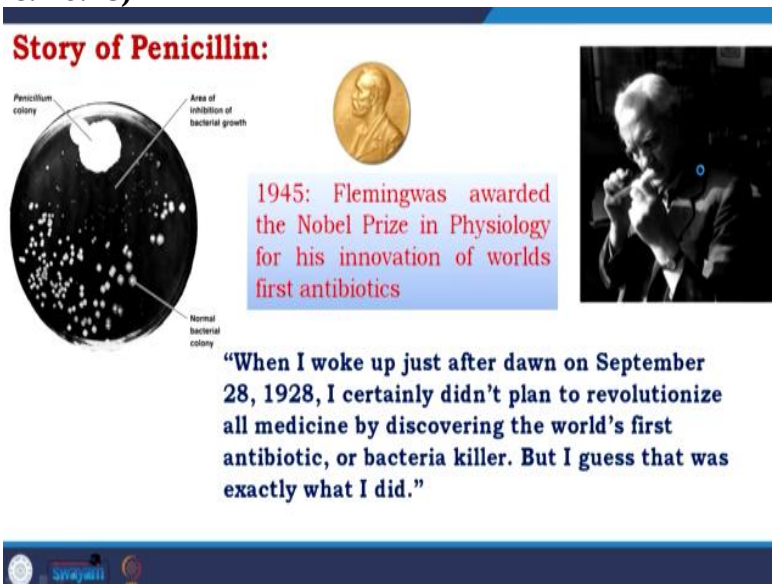
(Refer Slide Time: 11:11)



And when I am talking about these peptidoglycans, I would not leave the opportunity to introduce the mechanism of beta-lactams and beta-lactams if you guys do not know the most prescribed antibiotics. Beta lactams are antibiotics. Why are they so important? The peptidoglycan layer has D Ala-D Ala cross-linking. And when you look at the D Ala-D Ala moiety, this cross-linking is done by an enzyme called transpeptidase. If you look at penicillin, the first beta-lactam, you will see that it has remarkable similarities with the D Ala-D Ala. Because of that, penicillin, or other beta-lactams, sit down in the active site of that transpeptidase that is not the end. If you look at biology, you will look at DNA. You will see six-membered rings, and you will see five-membered rings, the same you will see in sugars in proteins. Biology prefers 5 or 6 membered rings because the strain generated because of the angular changes is least here.

And the strain is most in a four-membered ring or a three-membered ring because of ring strain, and they are very reactive. So when they sit down at the active site of the transpeptidase, they react, and ring breakage happens because of that. Remember a very, very unique phenomenon of biochemistry I am talking about. I am talking about the reaction between an enzyme and a substrate, and if you guys remember the definition of an enzyme, you will recall that the enzyme does not take part in the reaction. And this is a very good example where the enzyme itself takes part in the reaction to form a covalent bond. As a result, the enzyme transpeptidase becomes inactive because it forms a covalent bond with the substrate. So, it could not further participate in the peptidoglycan layer's cross-linking. The cell membrane of the bacteria becomes weak because they do not have the cross-linked peptidoglycan. So, with exposure to a hypertonic environment, the membrane will bulge out as water will be diffused membrane breaks, ultimately resulting in the death of bacteria.

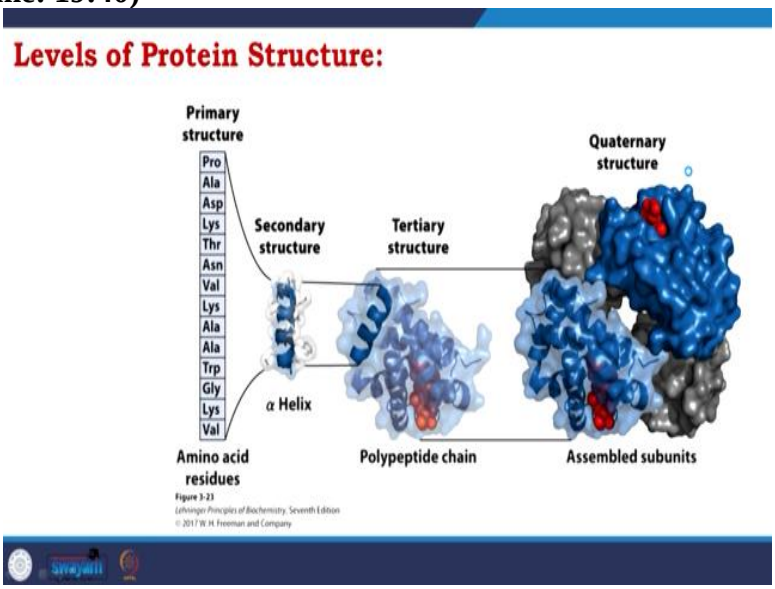
(Refer Slide Time: 16:28)



When I talk about beta-lactams, I will finish by talking about the story of penicillin. You all know about Fleming. And his experiment, he was working on bacteria S. aureus. And the picture you are looking at here is a Petri plate where he grows the bacteria. When bacteria grow, it becomes visible, called a colony, so these are normal bacterial colonies. He was working on them, and he went on vacation keeping the plates there. When he came back, he found a penicillium colony in one part of the petriplate. What makes him excited around the penicillium colony? If you see here in the picture that the normal bacterial colonies did not exist, that gives him an idea that these penicillium colonies might be giving something or doing something so that the bacteria are killed. So he picked up that penicillium colony,

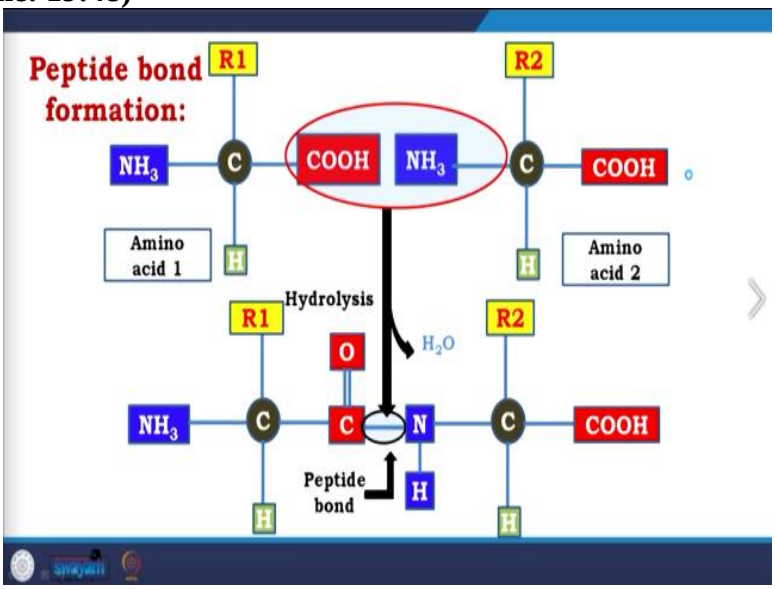
applied it to different bacterial stocks, and confirmed that yes, this penicillium colony has antibacterial activity in his own language because it was exciting when I woke up just after dawn on September 28, 1928. I certainly did not plan to revolutionize all medicine by discovering the world's first antibiotic or bacteria killer. But I guess that was exactly what I did.

(Refer Slide Time: 19:40)



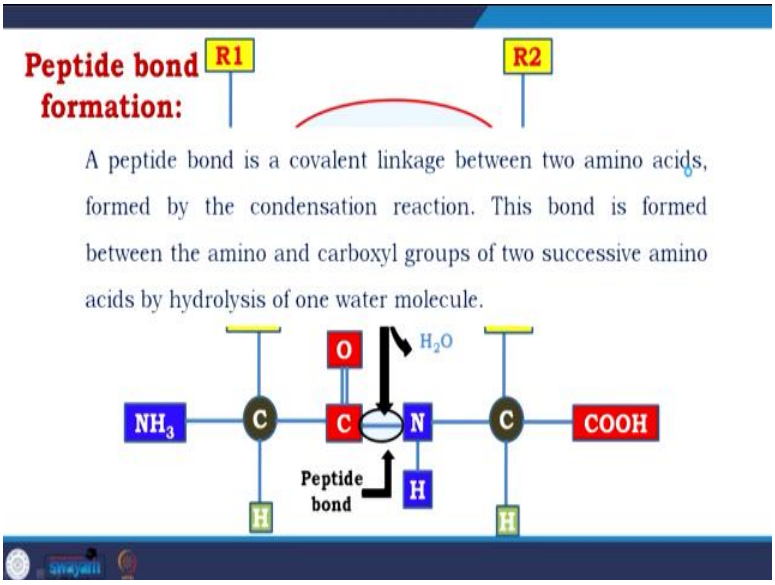
We talked about protein structure, primary structure, secondary structure, tertiary structure, and quaternary structure.

(Refer Slide Time: 19:48)



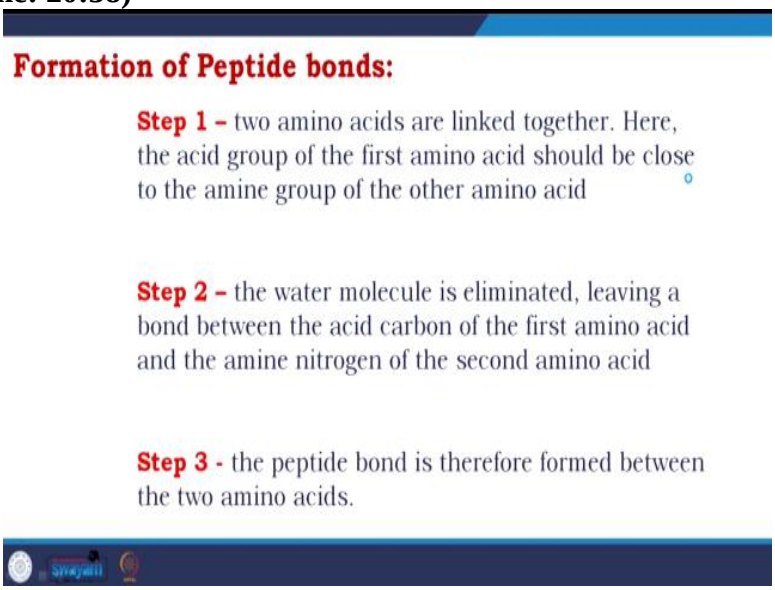
So, how the primary structure forms come to form something called a peptide bond from the primary structure. We have amino acid one amino acid two, then hydrolysis and peptide bond formation. So, the C=OH NH_3 they combine leaves a water molecule and connected.

(Refer Slide Time: 20:19)



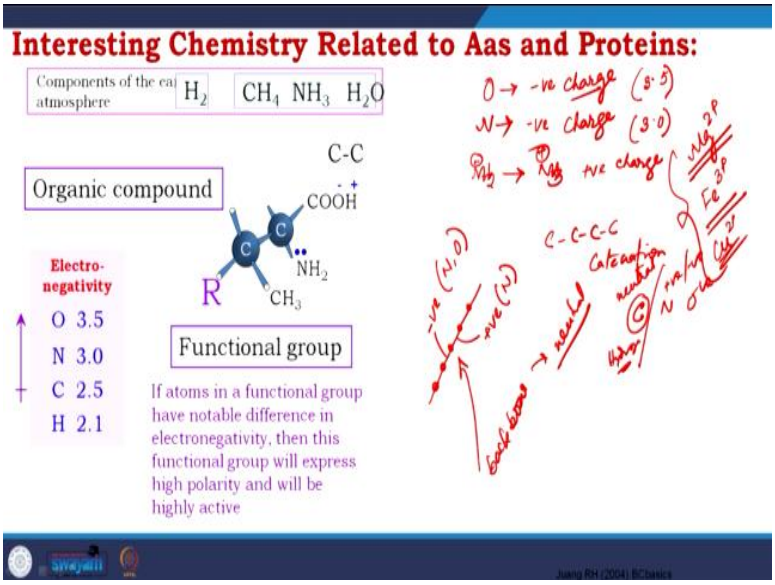
A peptide bond is a covalent linkage between two amino acids formed by the condensation reaction. This bond is formed between the amino and carboxyl groups of two successive amino acids by hydrolysis of one water molecule.

(Refer Slide Time: 20:38)



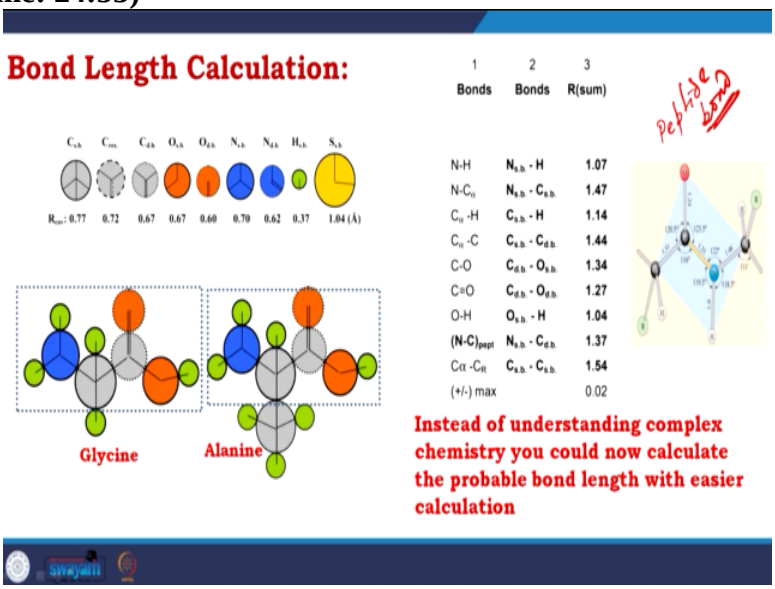
How do the peptide bonds form? In the first step, two amino acids are linked together. Here the acid group of the first amino acid should be close to the amine group of the other amino acid. Step 2, the water molecule is eliminated, leaving a bond between the acid carbon of the first amino acid and the amine nitrogen of the second amino acid. Step 3, the peptide bond is formed between the two amino acids.

(Refer Slide Time: 21:11)



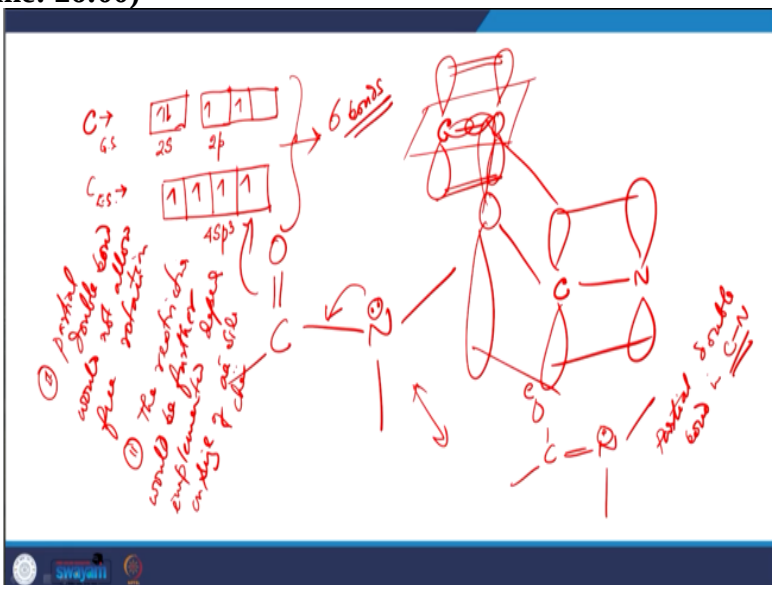
So, we talked about in the last class how the presence of those fundamental elements, hydrogen, carbon, nitrogen, and oxygen, has developed biology. I talked about two things: oxygen, which provides us negative charge. We showed that oxygen is mostly electronegative with 3.5 in the electronegativity scale, but very interestingly, there is nitrogen that comes up with electronegativity 3. It definitely could provide us negative charge, but it also forms NH_3^+ provides us positive charge.

(Refer Slide Time: 24:53)



We also talked about the bond length calculation shown in the last class. How we have taken the bond length of single carbon bond, carbon double bond, carbon resonating bond, single oxygen bond, double bond, nitrogen single and double bond, single hydrogen bond, and sulfur single bond, we have shown the bond lengths. We said that we do not have to do much when a bond is formed. We have to add the bond length here.

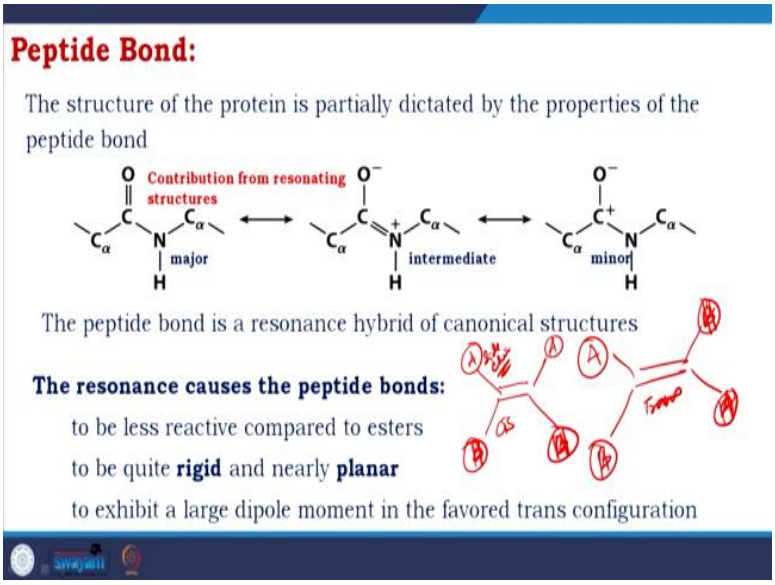
(Refer Slide Time: 26:00)



Now, coming to another aspect, we have carbon. We talked about carbon is having 2s2, 2p2 in its ground state. In its excited state, it forms 4sp3 orbitals, and they form the sigma bonds, but now, we will consider the pi bond. So, this is the sigma bond between carbon and oxygen where they have the overlapping now, and the pi bond would be formed where they will use the pi electrons to develop the bond and sigma bond because the overlap is straightforward is having higher strength. Whereas pi bond would have relatively lower strength now, the situation comes like this, but here it has a lone pair, which would be used to form a double bond.

So, I am showing you a peptide bond, and if it is a single bond, there would be free rotation, but because it is a partial double bond, that rotation is also restricted. How is it restricted? It is restricted depending on the size of the amino acid side chain. So, what is apparent here is that one partial double bond would not allow free rotation to the restriction would be further implemented, depending on the size of the amino acid side chain. This makes protein different from other polymers. Why? Because in other polymers, if you imagine be it carbohydrate, be it DNA, be it RNA, all the connections are made through a single bond. Now, here in the junction, we are getting a double bond. So, there are restrictions, and again, the restrictions are different with different amino acid sizes, giving us a chance to model to control. You have a carbohydrate, and it could have any confirmation, so you cannot make any rule. But because here we have the partial double bond and the difference in size and the existence of only L amino acids for the 19 chiral amino acids. You could dream of controlling, modeling, and having the rules, which is what happens in the case of proteins.

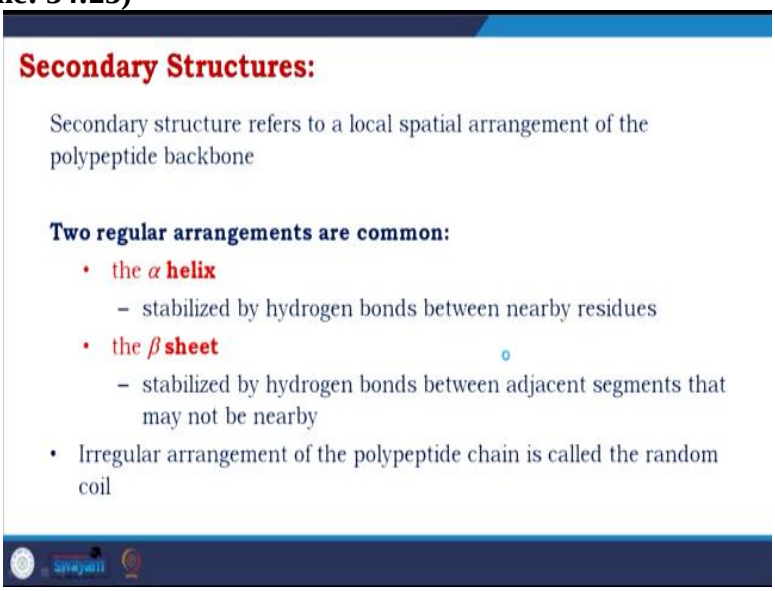
(Refer Slide Time: 32:40)



The properties of the peptide bond partially dictate the structure of the protein. A resonance hybrid of the canonical structures that you could see here causes the peptide bonds to be less reactive than esters to be quite rigid and nearly planar to exhibit a large dipole moment in the favored trans configuration.

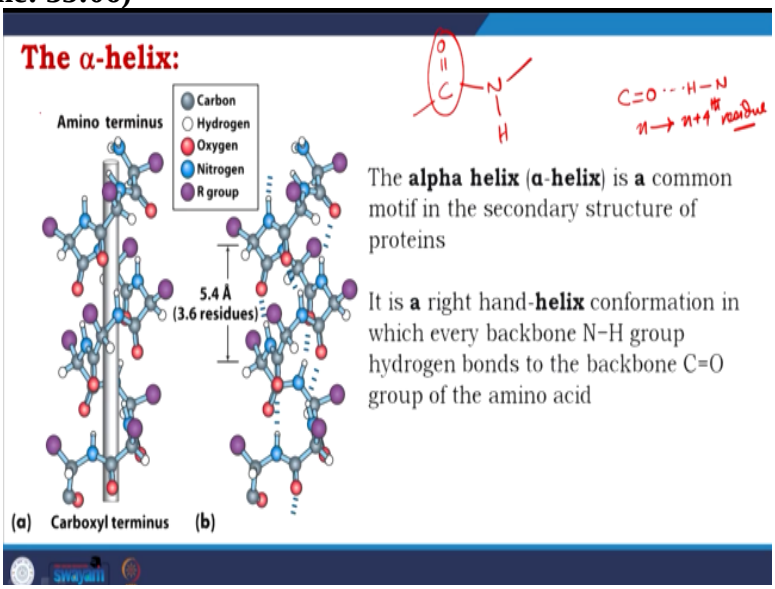
So, it is a double bond. There is a possibility of having a double bond, and you have a possibility of cis-trans. So, side chains of the first amino acid and second amino acid in the peptide bond would be returning in a trans configuration.

(Refer Slide Time: 34:23)



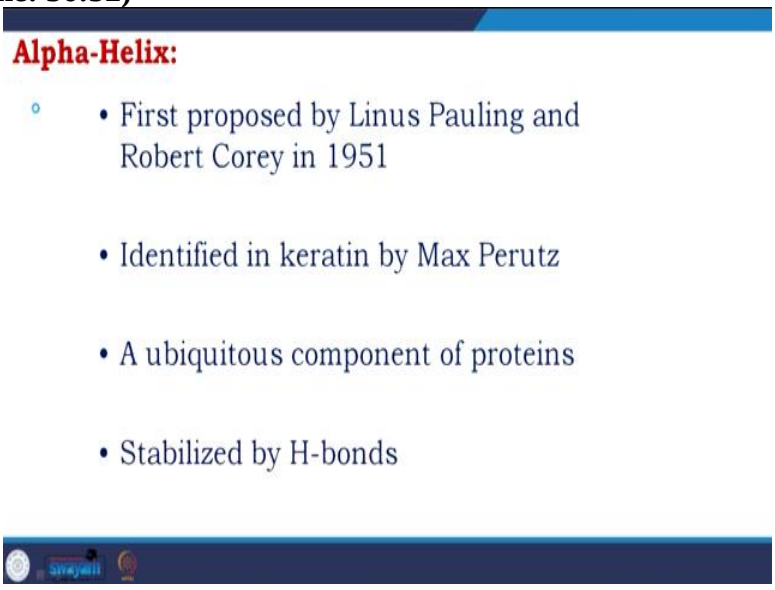
The secondary structure refers to the local spatial arrangement of the polypeptide backbone. Two regular arrangements are common one is the alpha helix, which is stabilized by hydrogen bonds between nearby residues. The beta-sheets are stabilized by hydrogen bonds between adjacent segments that may not be nearby. The irregular arrangement of the polypeptide chain is called a random coil, loop or turn.

(Refer Slide Time: 35:06)



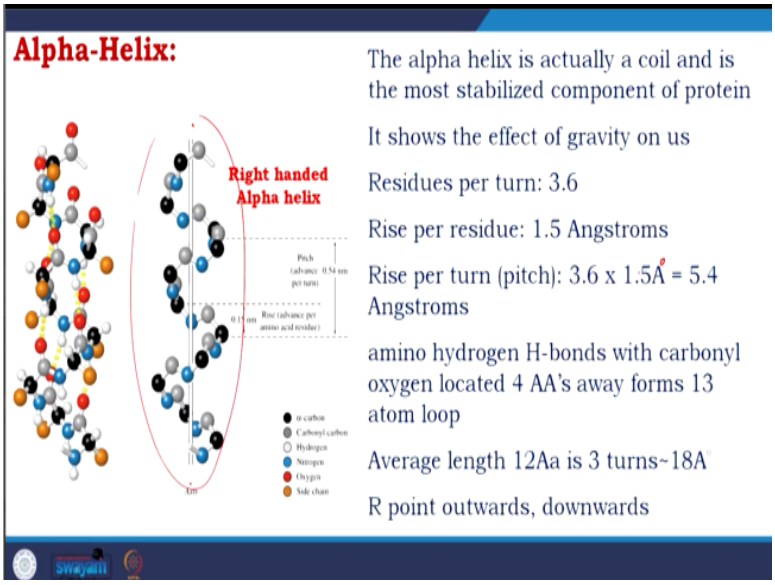
Alpha helix is a common motif in the secondary structure of the protein. It is a right-hand helix conformation in which every backbone NH group hydrogen bonds to the backbone C double bond O group of the amino acid. In the case of alpha-helix, it is n to n + 4th residue. So, if you see here, it always starts from the amino terminus remember when I was showing you, it starts from the amino terminus and goes to the carboxy terminus.

(Refer Slide Time: 36:52)



Alpha-helix was first proposed by Linus Pauling and Robert Corey in 1951, identified in keratin by Max Perutz, a ubiquitous component of proteins and stabilized by hydrogen bonds.

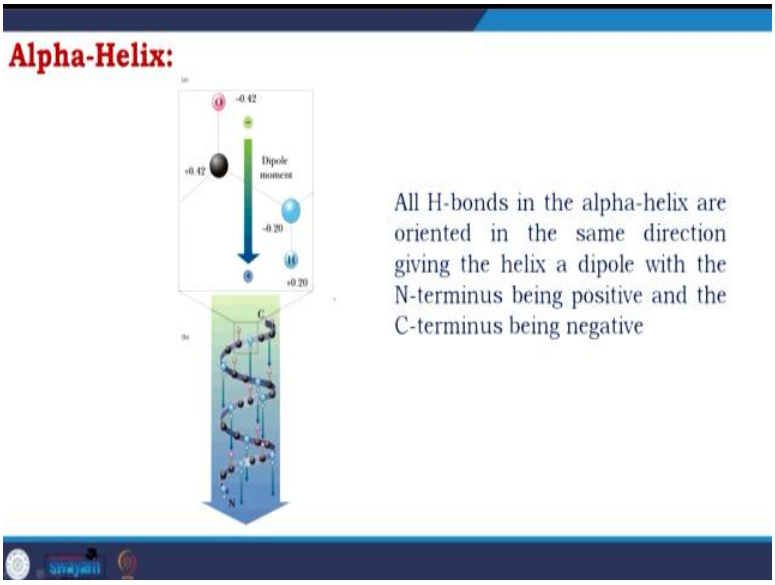
(Refer Slide Time: 37:12)



The alpha helix is a coil and is the most stabilized protein component. As I told it shows the effect of gravity on us what I mean by that, what I mean is if you guys could take a drop of water and you allow the water in the top of your finger and allow that to drop and watch carefully, you will see that the water becomes spherical. That is the effect of gravity, and because of that, every entity for their stability, especially polymers, shapes spherically.

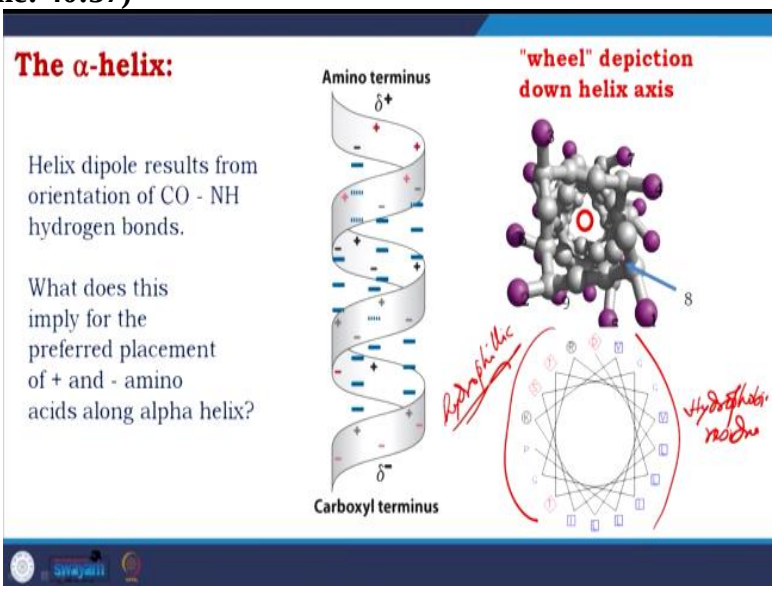
This spherical apparent spherical tendency of the polymer allows them to become helix. So, all the biological polymer have a tendency, how the residues stay there, per turn there are 3.6 residues and rise per residue is 1.5 angstrom, which provides rise per turn that is called pitch 3.6 into 1.5 angstroms that are 5.4 angstroms: the amino hydrogen, hydrogen bonds with carbonyl oxygen located four amino acids away and forms 13 atom loop. The average length of 12 amino acid is 3 turns and approximately 18-angstrom length. The R group(the side chains) point outward generally downwards.

(Refer Slide Time: 40:01)



If you look at alpha-helices, they have a directional motif. So, all the hydrogen bonds in the alpha-helix are oriented in the same direction, giving a helix a dipole with the N terminus positive and the C terminus negative.

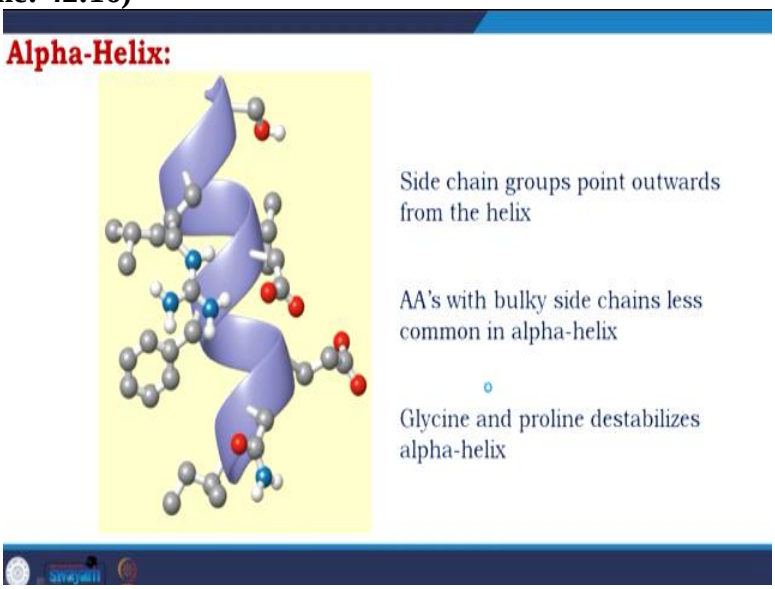
(Refer Slide Time: 40:37)



Now, the helix dipole result from the orientation of the CO-NH hydrogen bonds. You see, you get delta plus in the amino acid terminus delta minus the carboxy terminus. What does this imply for the preferred placement of alpha-helix positive and negative amino acids? So, you come to the quill, and you see that it looks like a sphere, and you could see that the hydrophobic residues come in one direction of the quill and the other residues, including the hydrophilic, come in the other side. Why are these important, and what is the implication? This will tell you how one helix will interact, and residues should be kept in one position. Suppose you are doing enzyme engineering, protein engineering. In that case, we will see

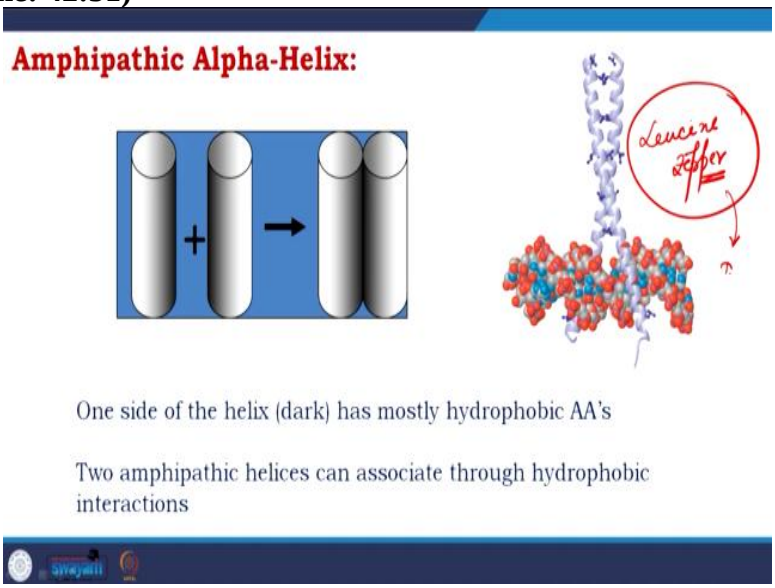
them in later sections, but the way the helix is developed, definitely there are rules more you understand the rules better you understand the protein.

(Refer Slide Time: 42:16)



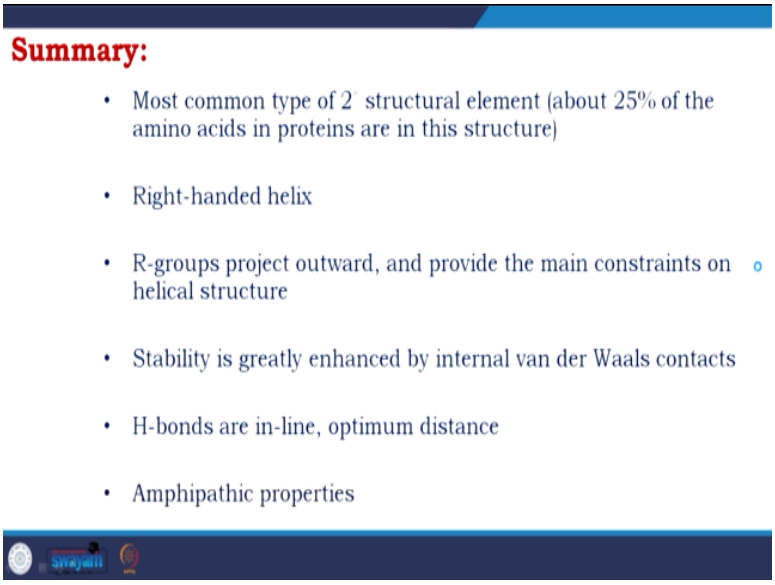
Sidechain groups point outwards from the helix, amino acids with bulky side chains less common in alpha-helix, glycine, and proline destabilize alpha helix. Glycine being very small brings flexibility and proline if you remember the locking they forced to make a turn.

(Refer Slide Time: 42:51)



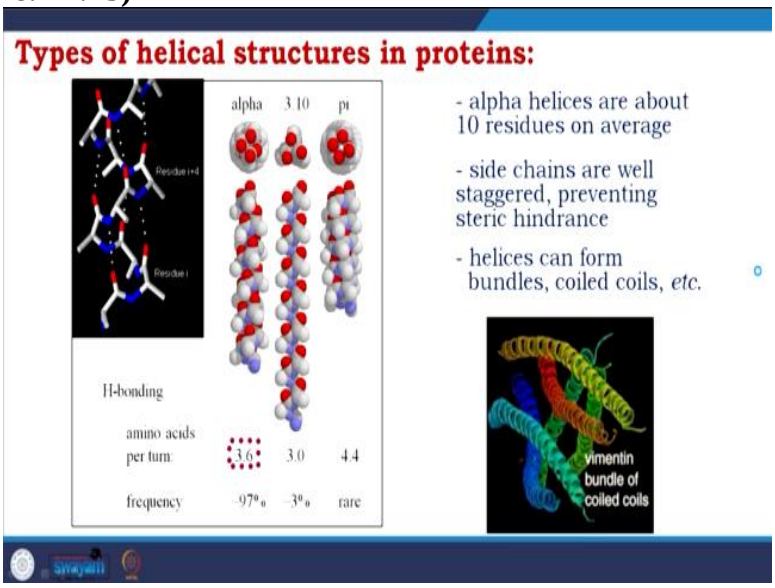
Amphipathic alpha-helix: One side of the helix has mostly hydrophobic amino acids. Two amphipathic helices can associate with hydrophobic interactions. Here you see the example called leucine zipper. Leucine zipper is a motif super secondary structure which we will talk about later here that is a helix-helix interaction stabilizing the motif, and it could bind to DNA.

(Refer Slide Time: 43:47)



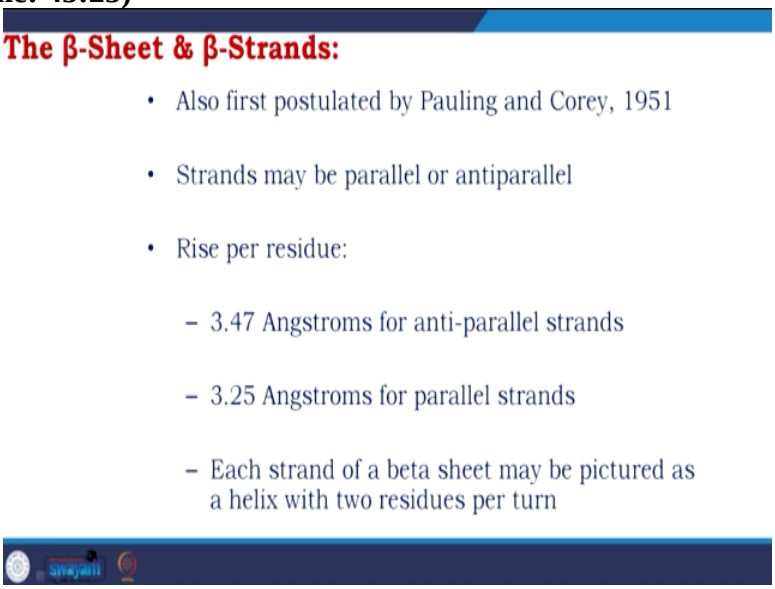
So, in summary, the most common type of secondary structural element, about 25% of the amino acid in protein, is in the right-handed structure helix. R groups project outward and provide the main constraints on the helical structure. Internal Vander Waal contacts greatly enhance stability. Hydrogen bonds are in line with optimum distance and have amphipathic properties.

(Refer Slide Time: 44:23)



There are commonly three types of helices. We talked about alpha helix. We talked about amino acid per turn 3.6, and the frequency of existence is 97%. There are others: 3_{10} helices and pi helix, 3_{10} helix amino acid per turn is three and existence is 3%, and pi helix amino acid per turn is 4.4 and existence is rare. So, there are alpha-helical structures like vimentin, the bundle of alpha helixes coil-coil motif. So, alpha helices are about ten residues on average. Sidechains are all staggered, preventing steric hindrance. Helices can form bundles coiled coils.

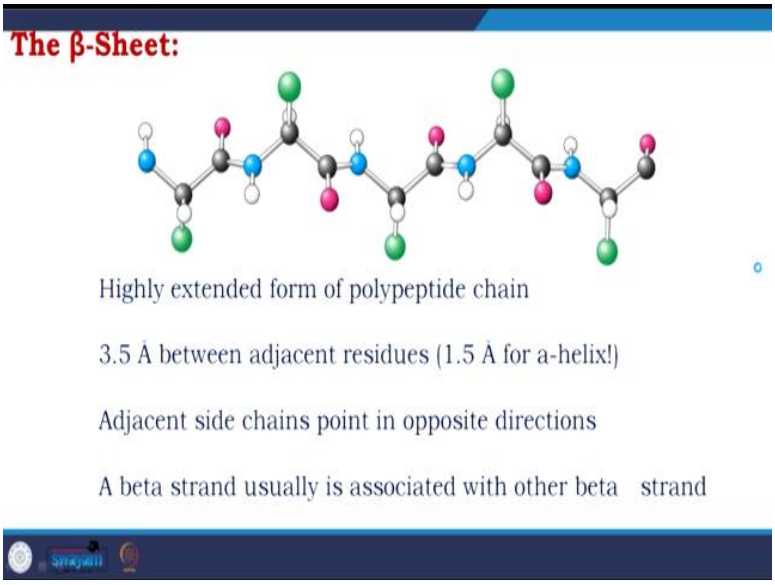
(Refer Slide Time: 45:25)



Now, we will talk about beta-sheet and beta strands individually. It is called a beta strand, and if you look at a beta-strand, it cannot form hydrogen bonds like an alpha helix. So, you cannot calculate the stability or energy of a beta-strand. What you have to do is to consider a beta-sheet which is formed by multiple beta strands, and they have hydrogen bonds between each other.

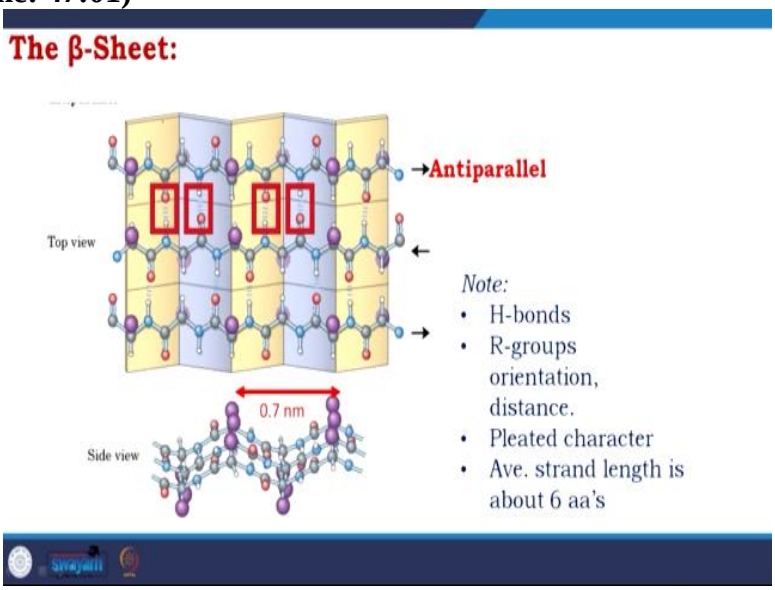
It is not individualistic. It is a hydrogen bond between one beta-strand and another forming a beta-sheet. It was also first postulated by Pauling and Corey in 1951. As I told may be parallel or antiparallel, the rise per residue is 3.4 angstrom for the antiparallel strands and 3.25 angstrom per parallel strands. Each strand of a beta-sheet may be pictured as a helix with two residues per turn.

(Refer Slide Time: 46:43)



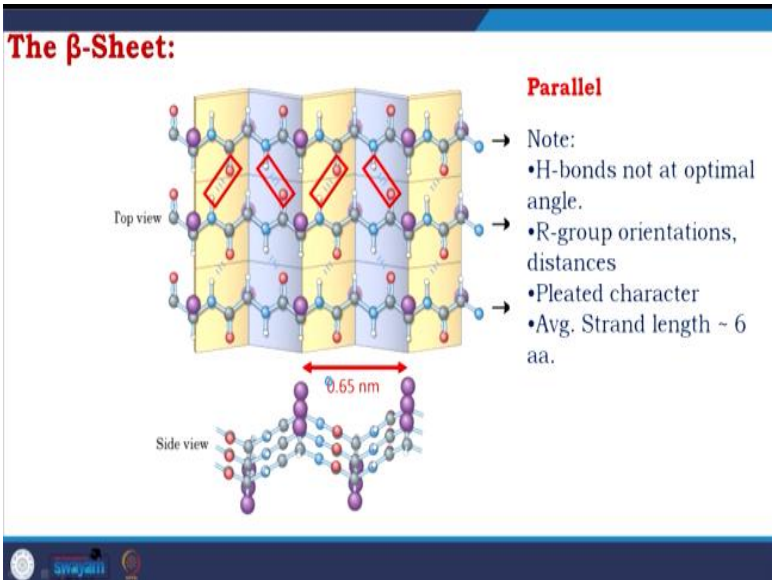
So, beta sheets are a highly extended polypeptide chain—3.5 angstroms between adjacent residues (1.5 angstroms for alpha helix). Adjacent side chains point in opposite directions. A Beta strand usually is associated with another beta-strands.

(Refer Slide Time: 47:01)



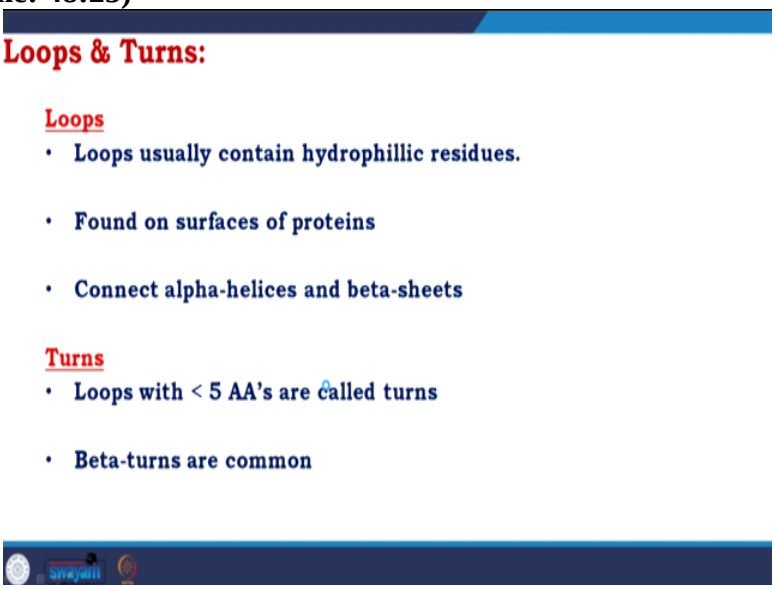
So, when they are antiparallel, they are stabilized by their hydrogen bonds, which are head to head generally. The R group orientation maintains a certain distance. It has a pleated character, and the average strand length is about six amino acids.

(Refer Slide Time: 47:50)



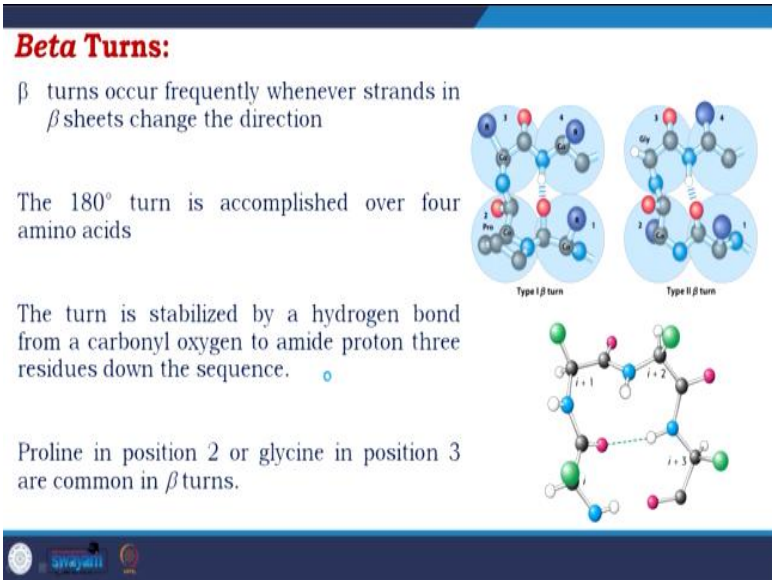
Parallel hydrogen bonds are not at an optimal angle because if you see they have an angle, they are not head to head the R group orientation is also different. They also have pleated character, and the average standard length is six amino acids. The distance between the two turns is reduced from 0.7 nanometers to 0.65 nanometers.

(Refer Slide Time: 48:25)



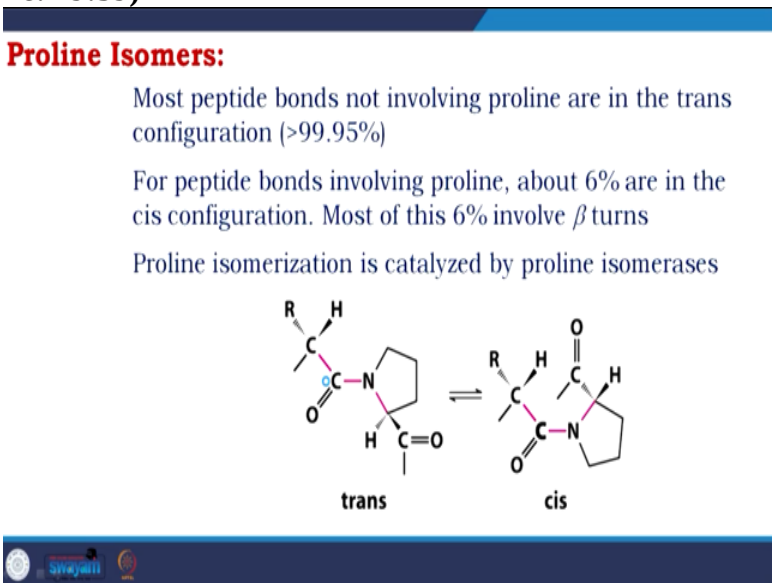
Loops usually contain hydrophilic residues found on surfaces of the proteins that connect alpha helices and beta sheets. Loops with less than five amino acids are called turns, and beta turns are common.

(Refer Slide Time: 48:44)



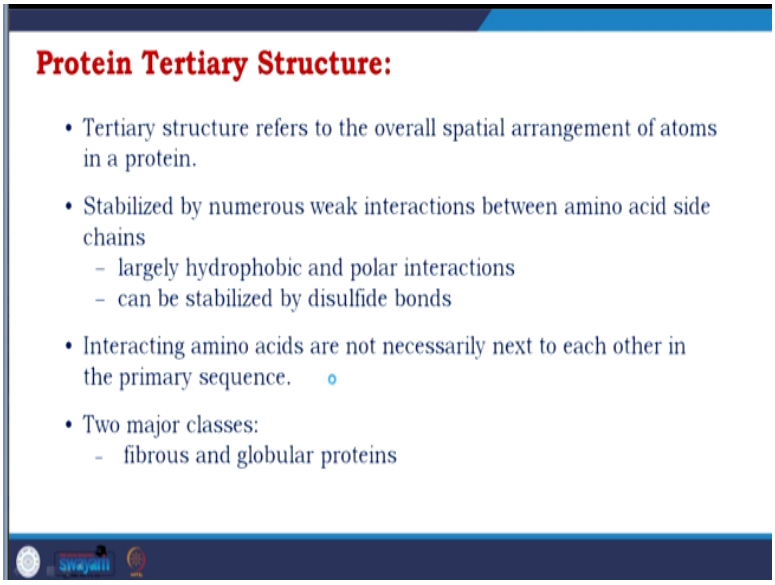
Beta turns to occur frequently whenever strands in beta sheets change the direction. The 180-degree turn is accomplished over four amino acids. The turn is stabilized by a hydrogen bond for carbonyl oxygen to amide proton 3 residues down the sequence, so it is similar to alpha helix, but instead of n to n + 4, it is three residues, proline in position two or glycine in position three are common.

(Refer Slide Time: 49:39)



Most peptide bonds not involving proline are in trans configuration. For peptide bonds involving proline, about 6% are in the cis configuration. Most of these involve beta turns, and the enzyme proline isomerases catalyze proline isomerization.

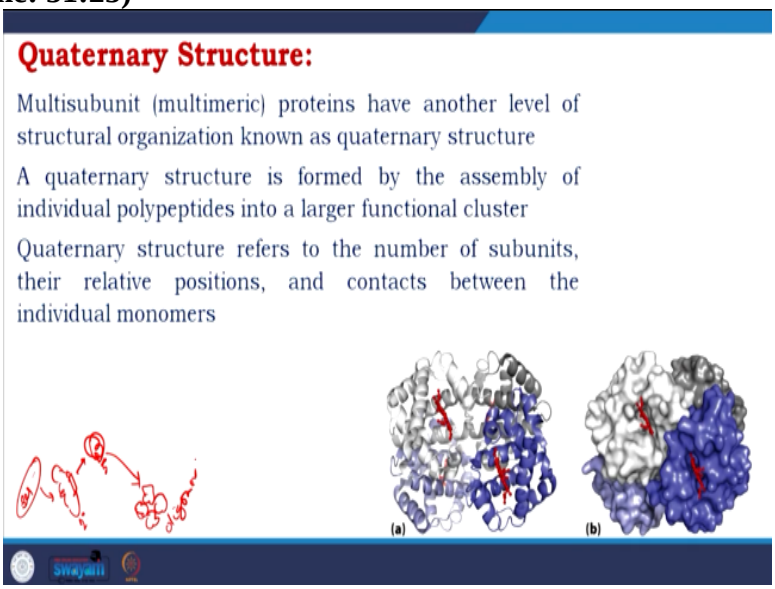
(Refer Slide Time: 50:11)



The tertiary structure refers to the overall spatial arrangement of atoms in a protein, stabilized by numerous weak interactions between amino acid side chains. Largely hydrophobic and polar interactions can be stabilized by a disulfide bond. Now, starting from the secondary structure to the tertiary structure to the quaternary structure, it is the dominance of the non-covalent bonds that we will especially study after this section.

We will see how they are working, their nature, and their influence. Still, we could say that the hydrogen bonds solid beads, hydrophobic patches base-base stacking, have an immense role in protein folding. Interacting amino acids are not necessarily next to each other in the primary sequence. Two major fibrous and globular proteins, mostly fibrous proteins, are structural and globular proteins that work as enzymes.

(Refer Slide Time: 51:25)

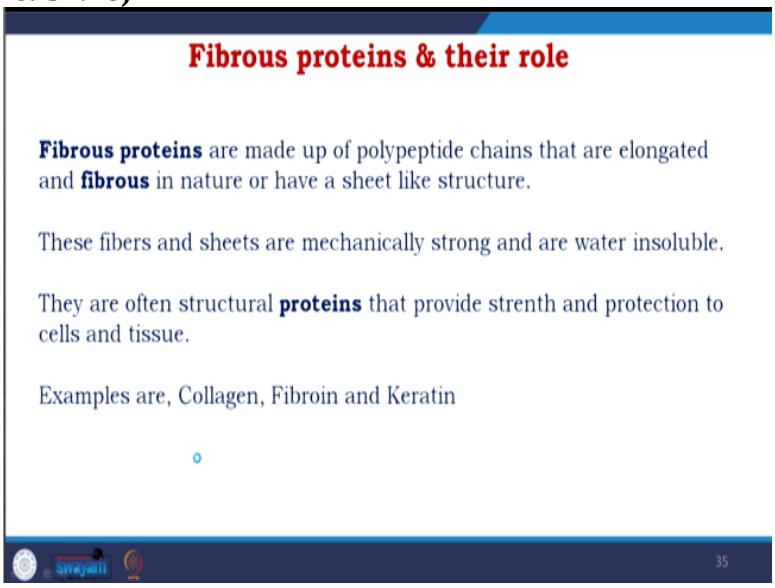


Quaternary structures multi-subunit proteins have another level of the structural organization known as quaternary structure. So, if your protein is only one subunit in the sequence, the

sequence becomes alpha helical beta sheets and then goes to secondary, tertiary. Now, if they have multi subunits, they are coming together, forming oligomers, dimer, trimer, tetramer, and many others could be possible.

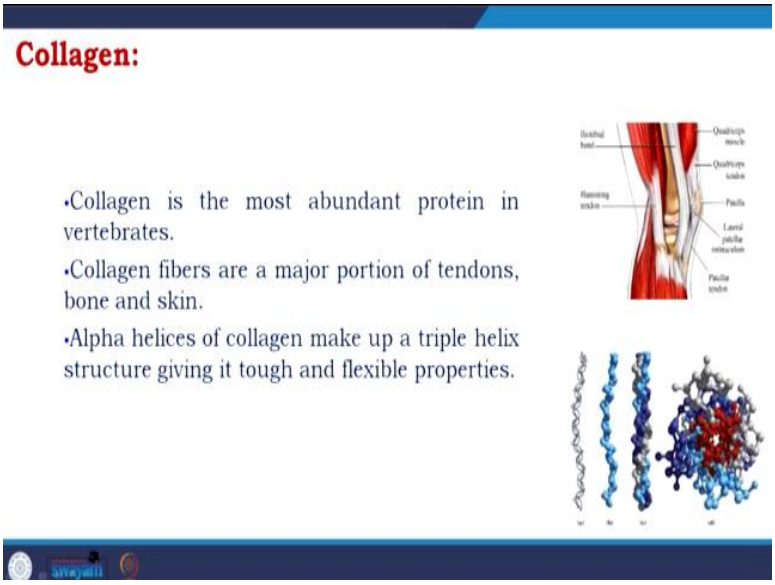
The assembly of individual polypeptides forms a quaternary structure into a larger functional cluster. The quaternary structure refers to sub-unit, relative positions, and contacts between the individual monomers.

(Refer Slide Time: 52:28)



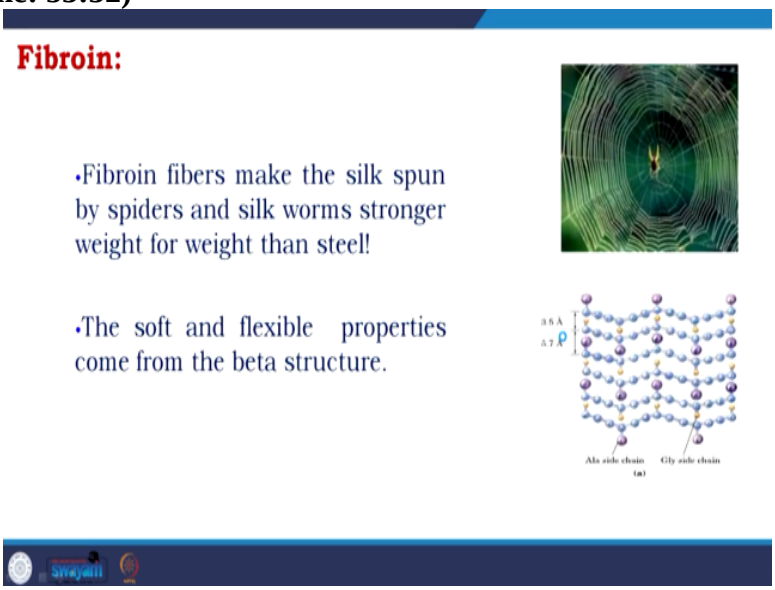
So, there are two types of proteins, fibrous and globular proteins. Fibrous proteins are made up of polypeptide chains that are elongated and fibrous or have a sheet-like structure. These fibers and sheets are mechanically strong and are water-insoluble. They are often structural proteins that provide strength and protection to cells, and tissues. Examples are collagen, fibrin, keratin.

(Refer Slide Time: 52:56)



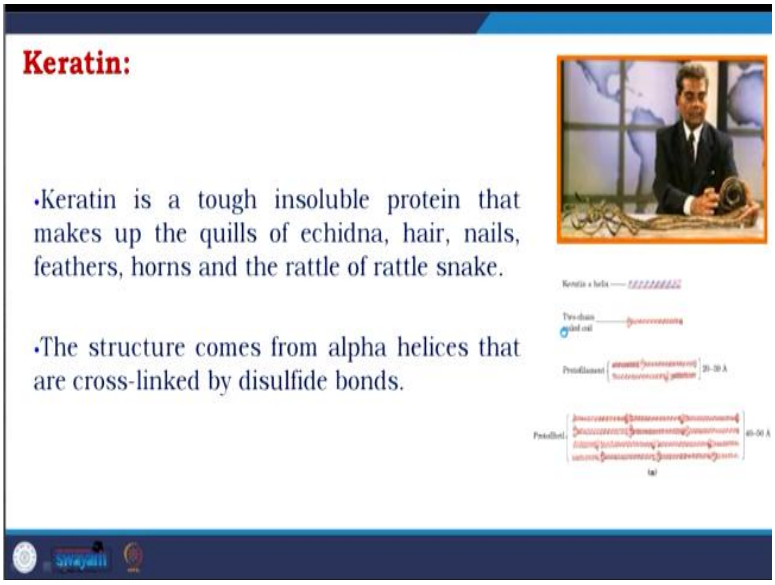
Collagen is the most abundant protein in vertebrates. Collagen fibers are a major portion of the tendons, bones, and skin. Alpha helices of collagen makeup, a triple helix structure giving it tough and flexible properties.

(Refer Slide Time: 53:32)



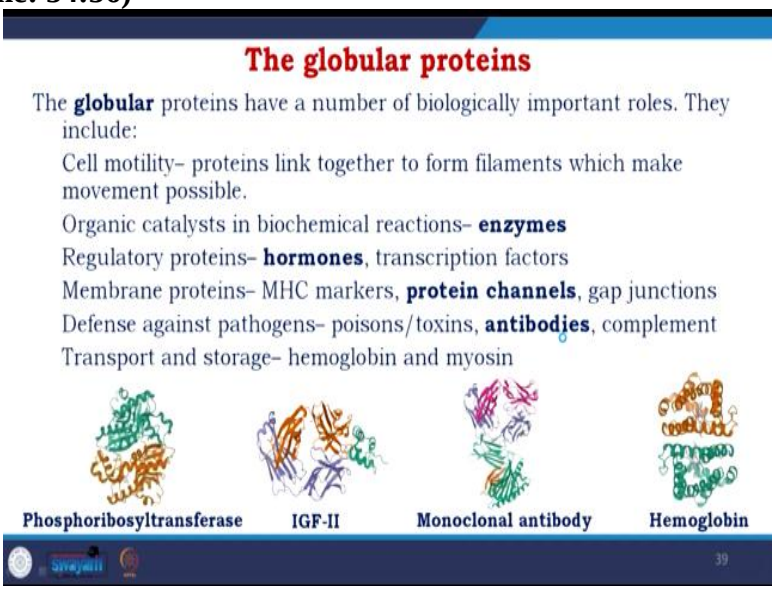
Fibroin fibers make the silk spun by spiders, and silkworm stronger weight for weight than steel show you all have seen how the spiders are developing their nets. These are fibroin proteins whose soft and flexible properties come from the beta structure. So this is the molecular structure where you see the presence of glycine side chain presence of alanine side chains. They are giving the structural rigidity.

(Refer Slide Time: 54:14)

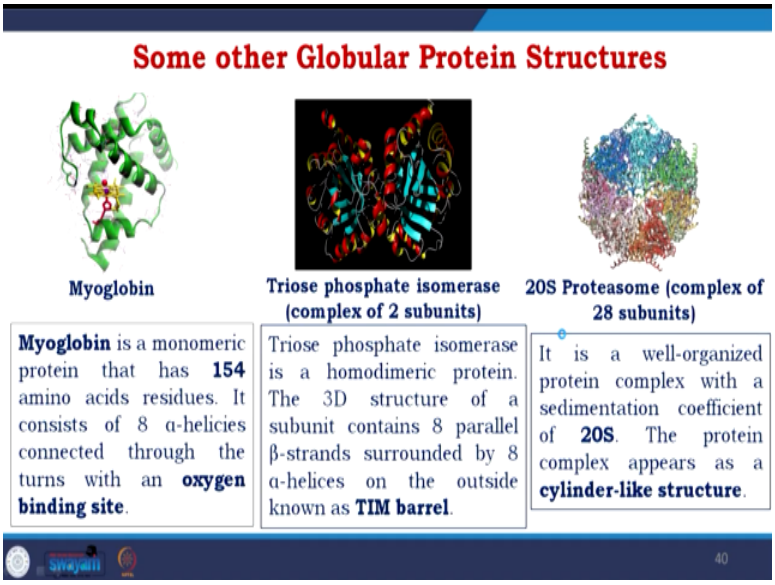


Keratin is a tough insoluble protein that makes up the quills of echidna, hairs, nails, feathers, horns, and the rattle of the rattlesnake. The structure comes from alpha-helices that are cross-linked by a disulfide bond.

(Refer Slide Time: 54:56)



The globular proteins have many biologically important roles. They include cell motility, proteins linked together to form filaments, which make movement possible, organic catalyst in biochemical reactions, enzymes, regulatory proteins, hormones, transcription factors, etc. Membrane proteins, MHC markers, protein channels, gap junctions, defense against pathogens, poisons, toxins, antibodies, complements, transport and storage, hemoglobin, and myosin.



Some other globular protein, myoglobin, is a monomeric protein that is 154 amino acid residues. It consists of 8 alpha-helices connected with an oxygen binding site through the turns. So, we talked about hemoglobin, hemoglobin is a multimer protein, myoglobin is a monomer protein, myoglobin is one of the first proteins studied and crystallized. Triosephosphate isomerase is called TIM. It is a homodimeric protein with a 3D structure of subunit, it has eight parallel beta-strands surrounded by eight alpha strands. TIM barrel proteins are available in many domains of life. Triosephosphate isomersae is the first protein where TIM barrel was found. The first structure was solved because this is part of our most abundant metabolic pathway, glycolysis.

20S proteasome it is a complex of 28 subunits. It is a well-organized protein complex with a sedimentation coefficient of the 20S. The protein complex appears as a cylinder-like structure. So, these are three structures, we have talked about them, one is a monomer, one is dimmer, and another is multimer. So, in that way, organizational complexity enhances, and protein from one protein structure becomes molecular machine will study about them.