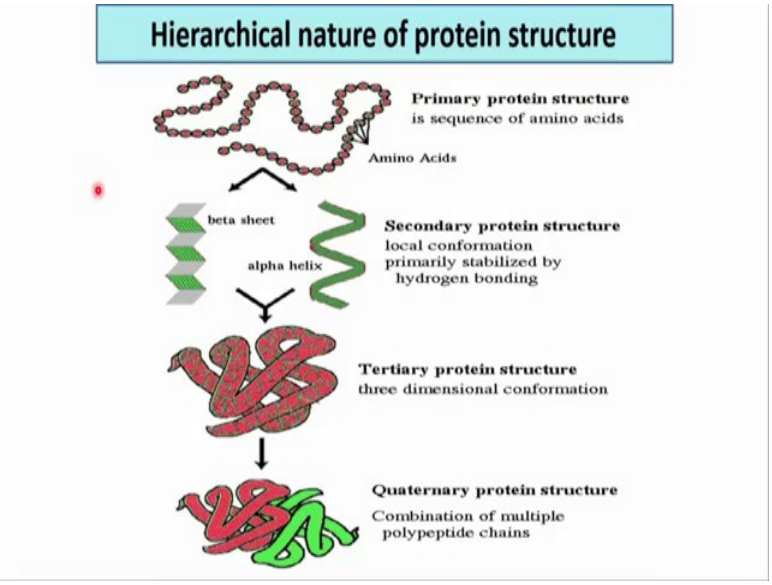
Organic Chemistry In Biology And Drug Development
Prof. Amit Basak
Department of Chemistry
Indian Institute of Technology, Kharagpur

Lecture - 10
Hierarchical Structure of Proteins: Secondary, Tertiary and Quaternary Structure

Welcome back to this new session. In the last session, we have discussed about the peptide synthesis and the related issues. Now, we will go on to theof the peptide We have already covered the primary structure of a peptide or a proteins (large peptides are proteins). We have done the primary structure which means that how the amino acids are linked to one another; that is called the amino acid sequence.

Now, proteins have a hierarchical structure; that means, that beyond the primary structure, there are other kinds of structure. So, the first layer of structural analysis is the primary structure; that means, the amino acid sequence, then there comes the second (higher) order and that is called the second degree structure.

(Refer Slide Time: 01:31)



If you look at the slides, see this is the primary structure of a protein and; these small blocks are representing one amino acid each. So, that is the protein chain. So this is primary structure. Then what happens, as the protein is synthesized? It tries to fold at different locations and at different local points, you will see different types of

geometries. That means, the geometry or the three dimensional structure of the protein at different locations of the protein are not same.

There are distinct features of structural patterns at different locations. And some of these distinct features are known, you have already heard of them possibly that there may be something which is called a helical structure like some portions of the protein may look like a helix as is shown here; and this is called α helix; or some portion may look like a β sheet; or some portion may look like a loop which means kind of a turn shape; u turn; or some portion maybe completely random. So, all these possibilities are there.

Now, these different geometries at different locations together make up what is called a secondary structure. And here the definition is already given; what is a secondary structure of a protein? Itt is the local conformation; that means, the local geometries which are present at different locations of the protein throughout the entire protein molecule. And remember the primary structure is obtained by connection of the amino acids through peptide bonds.

So, this in a primary structure, the covalent bonds are the important ones and they are the only ones to be considered (the peptide bonds). In a secondary structure, now you are not connecting any amino acids with each other through covalent bonds. What you are doing is that the protein is taking different shapes at different locations ok.

So, what are those forces that stabilize the different geometries at different locations?. The primary stabilizing forces of this α helix, β sheet or β turn etc, are the hydrogen bondings; this is occurring within the same molecule, hence the hydrogen bonding that is taking place here is intramolecular hydrogen bonding. And also it is intrastrand hydrogen bonding; if you take the α helix the hydrogen bonding is present within this strand itself.

β sheet has a little bit different connotation about the hydrogen bond, but the most important point is that the force that plays a dominant role in forcing the protein to take up different geometries at different locations are the hydrogen bondings. The hydrogen bonding between the NH and the CO; NH because it is the hydrogen bonding donor and the acceptor is the carbonyl; that means, the amide bond. Those are the points of interest for considering the hydrogen bond.

Now, after we analyze the local conformations, we will consider the overall geometry of the molecule. This overall geometry may take different shapes, but here we are not talking about local conformations. It is the molecule in total; that means, the whole molecule itself; how the 3D geometry of the molecule look like. Taking everything together, the β sheet the α helix, everything together. Thus, it is a complete three dimensional conformation of the molecule and that is what is called the tertiary structure. There is one more higher order stucture that is called the quaternary structure.

Sometimes some proteins can exist in more than one monomeric form. That means, say one monomer can combine with another monomer to make a dimer or it could be a tetramer; all sorts of possibilities are there. So, when a single molecular protein, attaches through weak interactions to another protein molecule and remain present as an ensemble, that is called the quaternary structure of the protein.

So, quaternary structure means combination of multiple polypeptide chains; each polypeptide chain represents a protein. Now this quaternary structure can be between the same polypeptide chains; in that case if it is a called a homodimer. If that is between different polypeptide chains that will be called a heterodimer.

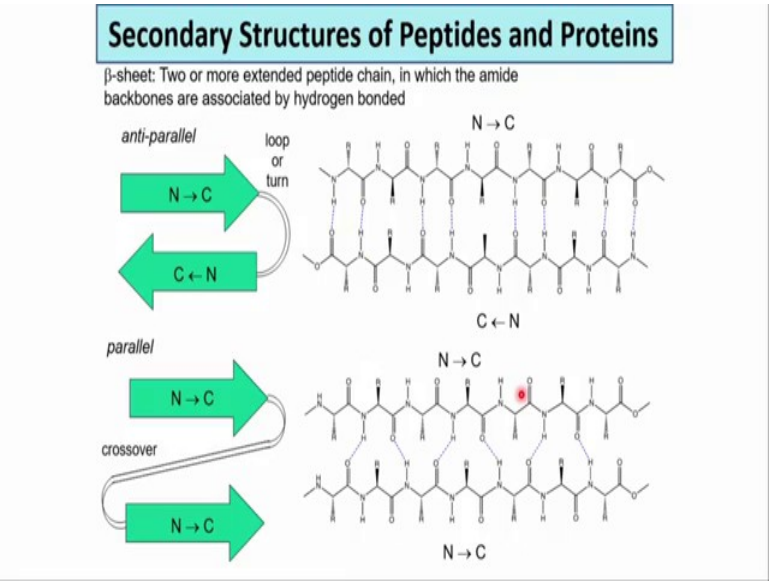
So, now we have all these four layers of structures; primary is the first one if you isolate a protein the first thing, you need to determine is its sequence. Once you know the sequence, the second layer comes which addresses that what are the local conformations at different positions. After that is analysed, then finally, that is still not final. For some proteins, which are only monomeric in nature, upto tertiary structure, that is a final. But for proteins which are multimeric in nature, you have to go for the quaternary structure.

Now, for all these secondary, tertiary and quaternary structures, weak forces take the role to induce particular geometries for the proteins and even to have multimeric nature of the proteins. So, the covalent bond formation is basically involved in the amide bond; However, there is one more covalent bond we should be aware of; and that is by the cysteine molecule which has got S-H, you know that it is a sulphur containing amino acid with the free thiol group that can be easily oxidized to a disulphide. And if two cysteines happen to be very close by, then what happens? They can be oxidized and form a disulphide bond (S-S). So, the disulphide are the only other covalent bond that is possible in a protein structure apart from the amide between the amino acids.

Other than that, the disulphide plays a role in the tertiary structure and in the quaternary structure. Sometimes, the tertiary structures are stabilized by formation of disulphide bonds; suppose there is a cysteine here at this location and another cysteine at that location. They are close by. So, there may be a disulphide bond formation which holds these two parts together in this shape. And the disulphide can also be there between the different monomeric forms to make the multimeric ensemble; that is also there.

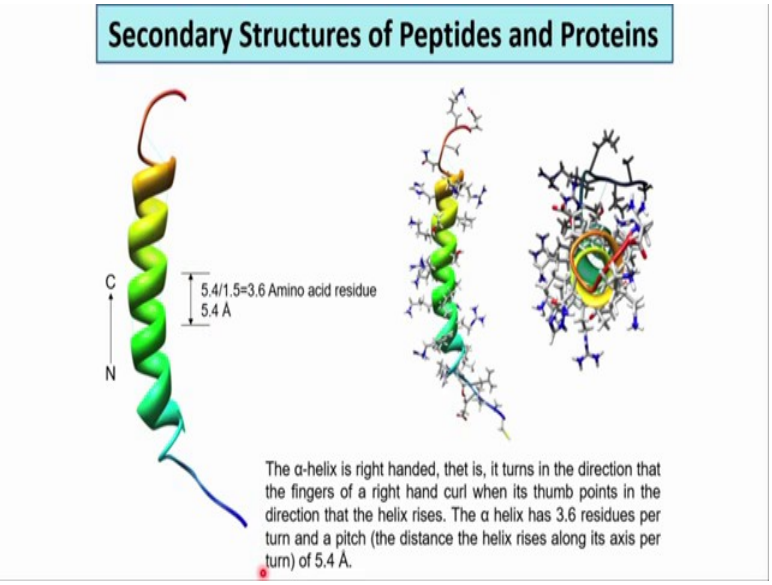
But apart from disulphide, leave aside disulphide, the others are all weak interactions, specially hydrogen bonds. In the secondary structure, in the tertiary structure; it is hydrophobic interactions, it is salt bridges and it could be π-stacking interactions; all these things will be there. So, we may go to the next slide.

(Refer Slide Time: 09:59)



Let us talk about the secondary structure now. The secondary structure is made up of the local conformations and these are distinct structural features that can be seen in a protein if looked at different regions.

(Refer Slide Time: 10:23)



The first one that strikes our mind is the helical form of this protein chain. This is called an α-αhelix and this is a right handed α-αhelix. Now, whenever you draw this helix, there will be an N terminus somewhere and a C terminus at the other end. In this case, this arrow is shown here; N going to C; that means, the N terminus is somewhere here and the C terminus is here. So, this helix is going from this direction to that direction; from the N to the C terminus.

Now, this right handed α-αhelix. Now I will show you some very simple models that how does it really look. If you look from the top side or from different angles, and what is the driving force that keeps it in this helical form? As I already said, hydrogen bond is the driving force, but the question is which amino acid is forming hydrogen bond with which other amino acid; because an amino acid here forms a hydrogen bond with an amino acid just at the top of it.

Now, if you take a line, just at the top of it, it can make a hydrogen bond if the orientations are favourable like a carbonyl pointing downwards and then NH pointing upwards. That can form the hydrogen bond. Now let us consider the various questions that come. First of all, why is it forming this right handed α-αhelix, why not the left handed? that is number one. Number two is, if we look from this side it is right handed helix whether it will be right handed, if we look from this side or not?That is the second question.

I have seen many students they face difficulty; they think that it is like a rotation of a fan; that if the fan rotates in an anti clockwise direction from the bottom then from the top, if we see that will become clockwise. But in cae of helix, interestingly it is not like that; it appears right handed upon looking from the bottom side, it will be also right handed if we look from the top side.

And then the question comes that what is the pitch of this helix? Helix is like a screw. What is the pitch of the screw? Screw pitch means the distance that is travelled that if there is a complete turn of a screw. So, here the distance is basically between a point here and a point there; that what is pitch. And then the next question is how many amino acids are there in a complete turn? So, these are some of the issues.

(Refer Slide Time: 13:39)

Now, let usshow you the model. We can make a model of these very simple things like aluminum foil. I can make an α helix. αIf you see from the middle, you can see that there is an axis. So, the helix is basically perpendicular to the helical axis.

Now, the question is how do you know this is right handed or left handed helix? In the right handed screw, what happens? You push it in a clockwise direction. So, if you do that, you also see here that as I do it, it is going in this direction. So, this is a right handed helix and you will see from both sides it will be the same; if you look from this side or if you look from the other side, it will look like the same. So, it is a right handed helix.

Now, what is holding this helix? You see these different bonds shown here which are marked as red these are those hydrogen bonds. So, the hydrogen bonds are actually maintaining the helical shape of this polypeptide chain. So, these are the hydrogen bonds. The hydrogen bond contains hydrogen and then nitrogen and a carbonyl which is acceptor.

So, if this is your first amino acid; then as you go here, that will be the fifth amino acid whose carbonyl will be coordinating with the NH. So, it is basically the hydrogen bond between the first amino acid and the fifth amino acid. So, after each four amino acids, you will have a hydrogen bond with the bottom one. I hope this is clear and then the distance between this and that; that means, the pitch of the screw, the pitch of the helix is 5.4 Å. And another thing you have to consider is that when you start from the nitrogen suppose, then come to the α carbon of the amino acid, then come to the carbonyl,

there is a rise in the helix because it is not on a horizontal plane, it is actually a spiral thing which is going up slowly. So, if you start from the nitrogen of an amino acid and go to the carbonyl, how much rise it is? how much you are elevating the surface? The rise is 1.5 Å. So, now, we know that the pitch of the helix is 5.4 Å and the rise per amino acid is one 1.5 Å.

So, then what will be the number of amino acids in one complete turn? That will be obtained by just 5.4 divided by 1.5, that is 3.6, roughly 3.6 amino acids per complete turn. Again I just go back to the questions; we have mentioned first of all that it is a right handed helix. Why it is right handed and not left handed? Because if it assumes this type of helical form, then the one thing that is not shown here, that is the substituents of the helix; the amino acid side chains. The substituents can actually project out.

They cannot project inside because then there will be too much steric crowding. Since they are in the L configuration, so the stereochemistry comes into play because they are in L configuration; that means, the side chains are projected outwards of the helix. So, that is why this is a right handed helix.

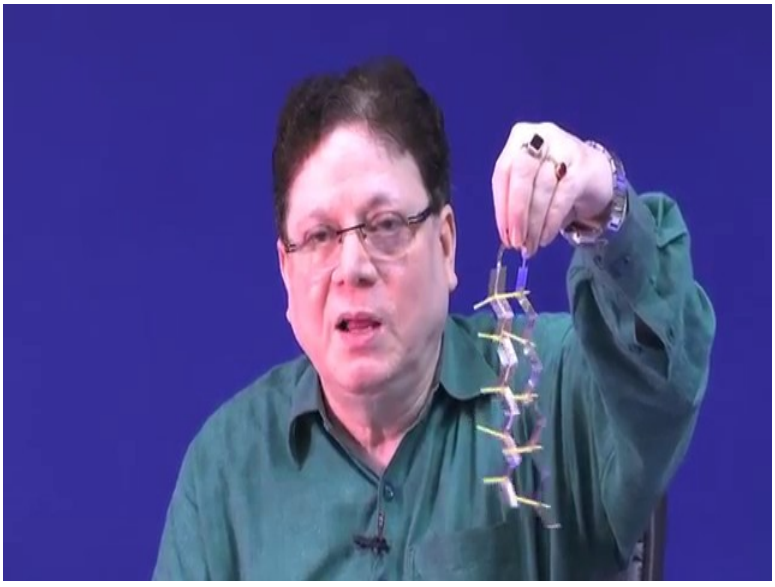
Now, if it is only glycine which does not have any substituent, then we have to remember that glycine is much more flexible and it is not constrained to interrupt this type of helical form. And there is one more amino acid which is called proline. Proline because of the secondary nature of the amine, it cannot participate in hydrogen bond

because when it forms the amide bond with another amino acid; the amide here is a tertiary amide where the N doesn't have any H; So, it is no longer a hydrogen doner.

So, proline and glycine are exceptional amino acids; glycine does not have any substituent. So, it is much more flexible. So, whenever you want to bring flexibility in a in a polypeptide chain, you incorporate glycine; you will see that at flexible regions, there are glycines. And if you want to bring in rigidity; means if you want to disrupt this kind of system which is stabilized by intramolecular hydrogen bond, then incorporate proline since proline cannot participate in hydrogen bonding.

So, proline is basically is also called a helix breaker; it breaks the helix. So, if some protein is going in a helical shape and suddenly changes into another shape; at that changeover point, proline may be present because proline is the one which does not allow any helix formation. Now, apart from this αright handed α helix, there is another one which is also very interesting, that is called the β sheet structure.

(Refer Slide Time: 20:07)



What is the β sheet structure? These are the amino acid planes these are the amino acid side chains, these are pointing upwards and then downwards; we have not put any substituent here. So, when the ridge is on the top, then the R group is going to the top side. Ridge means the point where these two surfaces are meeting, that meeting point. If the ridge is going downwards, then the substituent will be on the left side.

So, this is the β sheet. Now these two β sheets are actually interconnected by hydrogen bonds between the two sheets. Basically now I have shown the two sheets, now what is actually connecting these two sheets? The two sheets are connected by hydrogen bonds. So, this is the sheet structure where the hydrogen bonds are like this; see this is the extended form of the polypeptide and NH is pointing downward, CO is pointing downward.

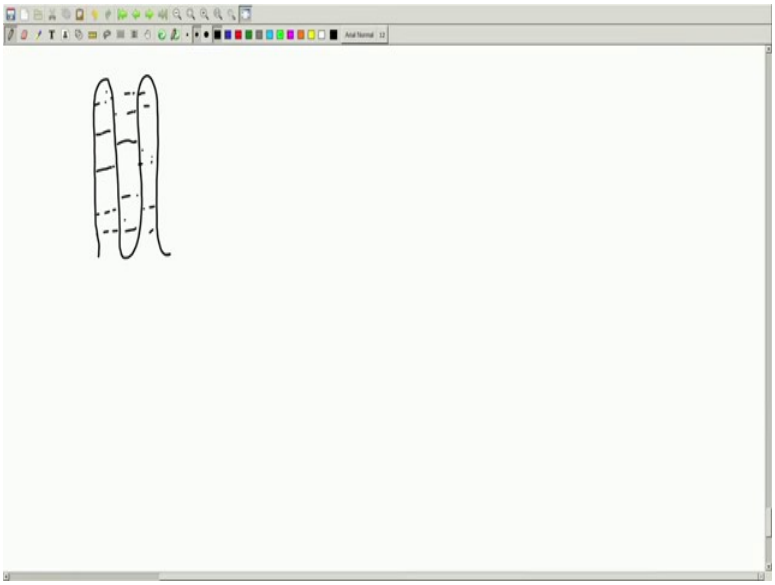
What did I say that when this is pointing upwards this carbon α carbon containing the R group is going upwards. And in the next amino acid, theR group is going downwards., thus it just alternates. If you have a very similar chain like this that there is a CO here; NH here and then alternating CONH is like that which are present in amino acids by the way.

So, then this chain and this adjacent chain can be interlocked by hydrogen bond which can be formed between this NH and CO. Now this is, by the way, suppose this is your N terminus say, and this is your C terminus so; that means, the direction of peptide bond peptide is generally as we write it from N to C. So, this is going from N to C and here also, you see the amine at this position and the carboxy at this position; that means, this is the N terminus and this is the C terminus.

So, here the direction of the peptide is N to C from this side. So, this is N to C from that side and this is N to C from this side. So, they are basically called anti-parallel β sheets. There may be other sheets or other chains that can also be linked to this chain; Because here you see, there is this NH and then followed by a carbonyl. They are not coordinately saturated means they are hydrogen bond capacity is still there that is not satisfied; only this carbonyl and this NH are satisfied, then the next NH and CO are still available for hydrogen bonding.

So, what can happen? Another chain can come on the top of it and then, they can assemble βto form a β sheet. So, β sheet is nothing, but you have first an extended chain of peptide, then there is some turn there and then another extended sheet of peptide and then there will be interchain hydrogen bonding. And then it can again take a turn and come to that side.

(Refer Slide Time: 24:13)



So, basically if I draw it here, what is happening is that there is something like this. So, the β sheet can be like this. So, there is hydrogen bond; all this interlinked. Here I have only two chains in this model; I have shown only two chains; I think you can see. So, what I am saying is that at at the top, there has to be a turn, otherwise how can we get this shape because if there is no turn, then that will go to the top of it and there will be no hydrogen bond between the two chains. So, there should be a turn like this which is also called the β turn; I will come to β turns slightly later.

And then again there can be another β turn which can form another layer of polypeptide and this can continue. And this whole thing becomes a bundle of a large sheet structure. This is the one which shows the β sheet. Here, one chain is going up and down. Then another chain in the middle going up and down and these chains are connected by hydrogen bonds and then there is another chain.

Now, what are the major differences between α and β sheet? One difference is manifested through and Ψ angles that are there according to the Ramachandran plot; we will come to that little bit later. And the second one is also very important; in α helix, there is hydrogen bond within the helix within the same chain (intrachain).

But in β sheet, the hydrogen bond that is taking place is between the two chains (interchain); that means, it is the same molecule, but this part or that part is not individually stabilized by hydrogen bonds. These two are interlinked by hydrogen bonds

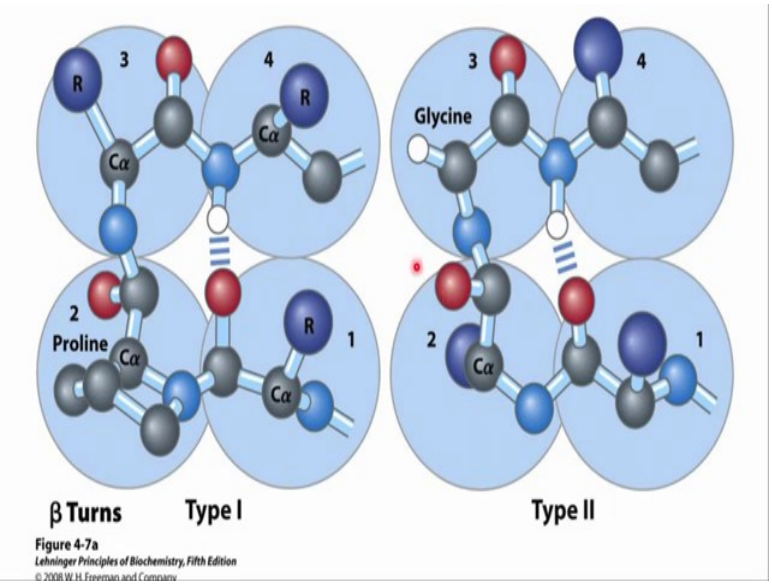
from here to there. So, there are hydrogen bonds in between these chains. So, that is the major difference between these two and in fact, in this β sheet, the hydrogen bond directions are very favourable. So, hydrogen bond strengths are more in β sheet than the helix. So, the β sheet structure is more stable than the α structure.

However this delicacy of α and β; their equilibrium from one to the other is a major issue in controlling the activity of the proteins that we have in our body. We will come to that when we discuss medicinal chemistry.

So, we have discussed the secondary structure; then we talk about the bends say; what are these bends? Now again I should say that there are parallel β sheets; parallel means you have N to C in this direction and N to C in this direction. If you want to have parallel sheet, then there should be two turns that you need; one is a turn on this side, the chain goes like this and then you get another turn and then the two directions become same N to C (same direction).

But in order to have anti-parallel chains, it will have just one turn and there is no cross over like this and obviously, if you look at the structure of the two, here they are aligned perfectly well interms of H and CO (orientation for hydrogen bonding), but here that alignment is not there. So, the anti parallel sheet is more stable than the parallel β sheet.

(Refer Slide Time: 29:09)



Now, how to get this turn or what makes this turn? There are primarily two types of turns. Say this is a turn comprised of amino acid 1 suppose, then amino acid 2 amino acid 3, amino acid 4. So, basically the peptide is now starting from here, going like this; this is the backbone and then it is taking a turn here.

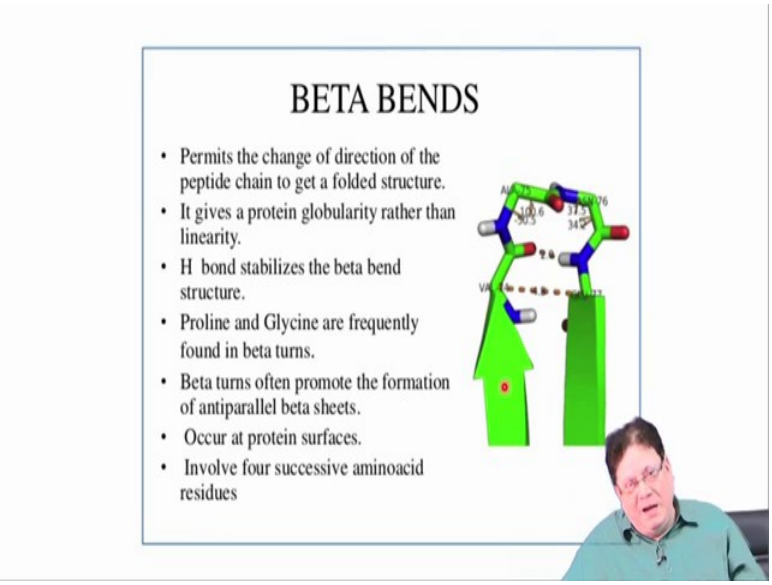
Now, in order to induce turn again, you should have a hydrogen bond and that hydrogen bond now is between the carbonyl of the first amino acid and the NH of the fourth amino acid. So, if it is i then this is i + 3; because this is C number 1, but this may not be the starting point of the peptide that is why it is better that you say that this is the i^{th} amino acid and this is the i + 3 amino acid.

So, there is the formation of intramolecular hydrogen bond and if you calculate the size of this, that will come to be 10 membered ring. So, because of the presence of this 10 membered hydrogen bonded network, the chain takes a turn. It was going in this direction and then ultimately comes in the opposite direction. So, that induces the turn; and there are two types of turns and these are actually again based on Ramachandran plot (Φ and Ψ angles), but the most important aspect that what are the amino acids which can induce these turns; that means, the amino acids which does not give any stability to the α helix or the amino acids which does not give any stability to the β sheet.

Now, there are two ways to do that on the basis of conformational flexibility of the amino acids; one is by incorporatingamino acid is very flexible; that is one possibility. The other is to have amino acid that is very rigid and devoid of hydrogen bonding capability and that is proline. So, you see in type I, the number 2 amino acid; that means, the i + 1 amino acid is proline and in type II turn, the i + 2 amino acid is glycine. You see this flexible amino acid makes the formation of the β turn and the hydrogen bonding inability of proline makes the possibility of the type I β turns. So, these are the major types.

Other β turns are there like type III, then type II a prime; all these are there; but for our level, I think it is sufficient to know that there are primarily these two types: I and type II turns.

(Refer Slide Time: 32:31)

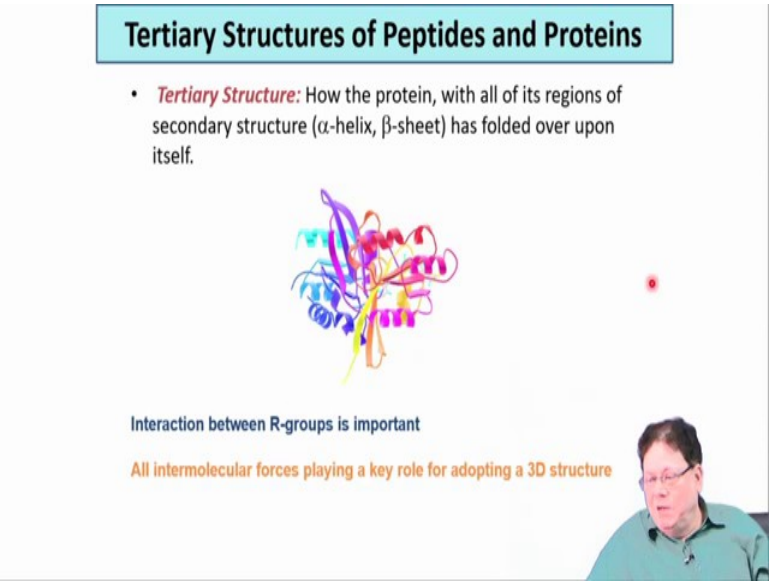


So, these are sometimes also called β bends, β turns or Ω loop. It is also said that this is the β sheet; the sheet structure that was going in this direction and then it takes a turn and then the sheet structure goes in the opposite direction. So, that is anti-parallel. So, that is how the different sheets are connected.

So, what are the characteristics of β bends? First of all, it permits the change of direction of the peptide chain to get a folded structure. It gives the protein globularity; that means, rather than linearity; now it can be ircular when you have a turn; that means, it is going through a kind of a circular direction. So, a globule is kind of a circular thing. So, the proteins generally try to adapt a globular shape and these turn structures assist the protein to adapt the globular state.

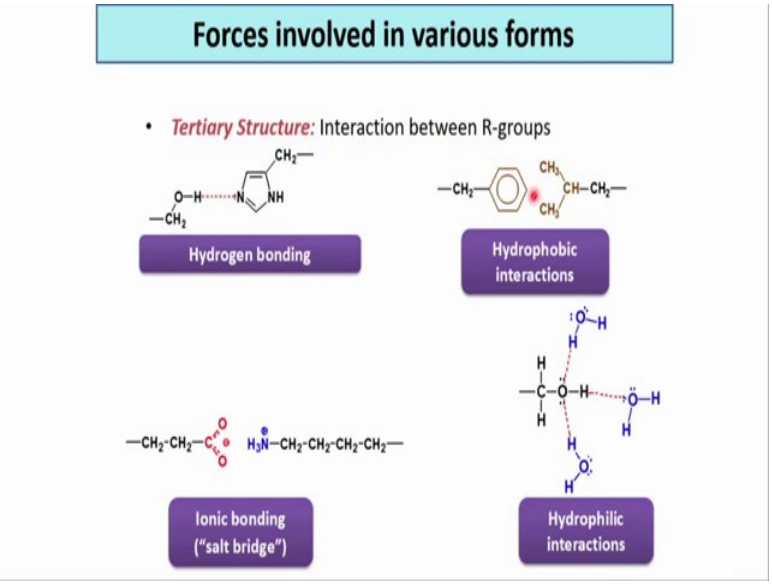
Hydrogen bond stabilizes the β-bends; again hydrogen bond plays the dominant role, proline and glycine are frequently found in β turns. β turns often promote the formation of anti-parallel β sheets as shown here, it usually occurs at protein surfaces and not at inside and involve four successive amino acid residues; i and up to i + 3. So, that completes the discussion on secondary structure of protein.

(Refer Slide Time: 34:01)



The tertiary structure is a little bit easier; tertiary structure, as I said it is the complete three dimensional structure of a single monomeric protein molecule and ultimately it can be a combination of α helix, some portion may be β sheet, some portion may have loops. And incidentally now another class of proteins are there where some portion may be intrinsically disordered; that means, they do not have any order, they are vibrating from one geometry to another geometry. So, they are called intrinsically disordered, but they are much less in number; a major class of proteins actually have a well defined 3D structures.

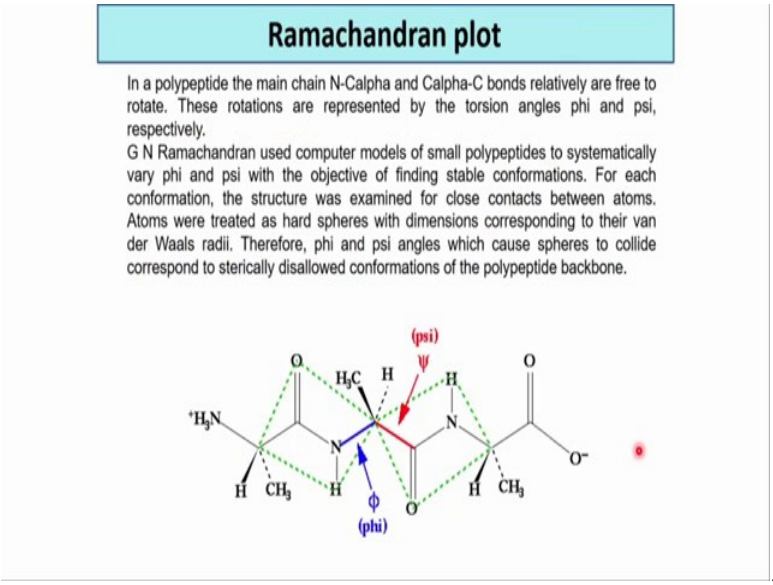
(Refer Slide Time: 34:59)



These are the interactions that take place when a protein adopts the tertiary structure. Remember again the primary structure is stabilized by covalent bonds, the secondary structure is stabilized by hydrogen bond. It could be in the same chain; it could be in the same locality; that means, α helix have intrachain or it could be interchain as found in β-sheets. And then the tertiary structure is stabilized by hydrogen bonding; there are lots of hydrophobic interactions; that means, water repellent groups or it could be salt bridges like if you have an aspartate or glutamate and if you have a lysine, then lysine will be present at the biological pH as NH_3 plus this will be CO_2 minus.

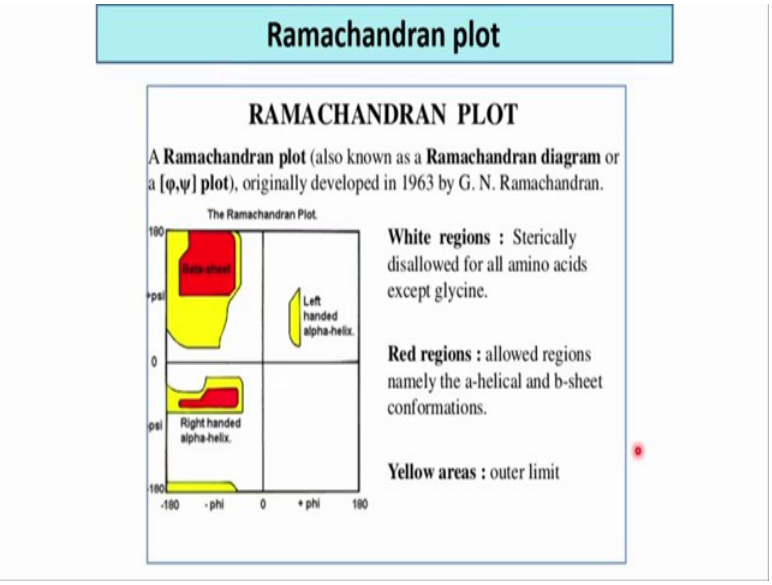
So, there can be electrostatic interaction that is sometimes called the salt bridge formation; water plays a dominant role that in some of the pockets water may be there which also can participate in hydrogen bonds with the R groups.

(Refer Slide Time: 36:13)



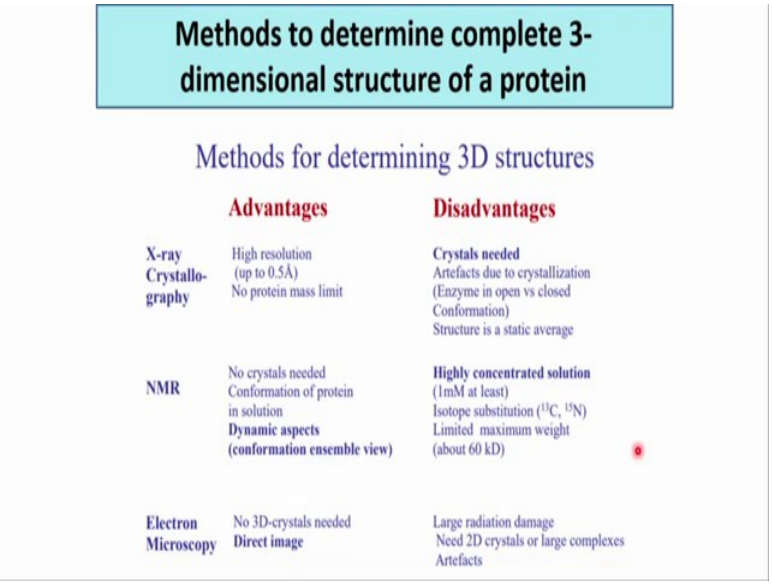
So, these are the stabilizing factor of for the tertiary structure.

(Refer Slide Time: 36:21)



And quaternary structure are very similar; quaternary structures are stabilized by hydrogen bonds and salt bridges.

(Refer Slide Time: 36:25)



And disulphide; disulphide it is not mentioned here; disulphides are very important. They play a very important role; to give a well defined shape to the protein; like a very simple example, that some people have curly hair and some people have straight hair. The people who have curly hair have a lot of cysteines in their keratin and cysteines are actually present as disulphide (oxidized dimeric form of cysteine containing the S-S

linkage is known as cystine). Now when they are present in disulphide, you get a bend shape. So, the hair becomes very curly.

So, now there is a tradition that curly hairs can be straightened and how to do that? You break those disulphide bonds with reducing agents and you break that bond and then make the hair straight. So, basically these disulphide bonds also play their role in controlling the geometry of the tertiary as well as quaternary structure. Like in many antibodies which are also proteins, there are a lot of disulphide bonds that stabilizes the monomeric systems.

Thank you.