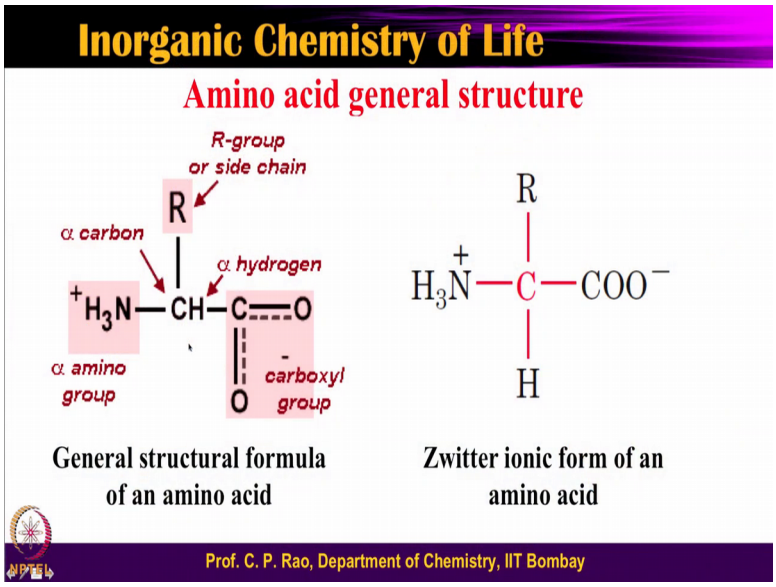
Inorganic Chemistry of Life Principles & Properties
Prof. C. P. Rao
Department of Chemistry
Indian Institute of Technology, Bombay

Lecture - 06
Amino acids, peptides & proteins - An introduction

Welcome you all for the next lecture on Inorganic Chemistry of Life Principles and Perspectives. Let us go through a recapitulation of what I have taught in the previous class. In the previous class the main theme of whatever I talked about is the metal ion is present in the enzymes as a coordination complex, which is surrounded by the protein and as a result of that the metal iron properties do change with respect to the type of protein that you have protein conformation, a protein sequence all different kinds of details is the protein as compared to the simple coordination complex or a anion in aqueous medium so either of these if you compare. And this is a boon in disguise because by changing the protein surroundings and the nature of the protein itself you can change, all together properties of the enzyme ok.

Having been talking a lot about the proteins side chains etcetera, I thought it is appropriate for those who may or may not know more details about this protein. So, I thought I will give some basic details these very basic details of these I mean proteins and ok. So, let us look at first of all starting from the Amino acid.

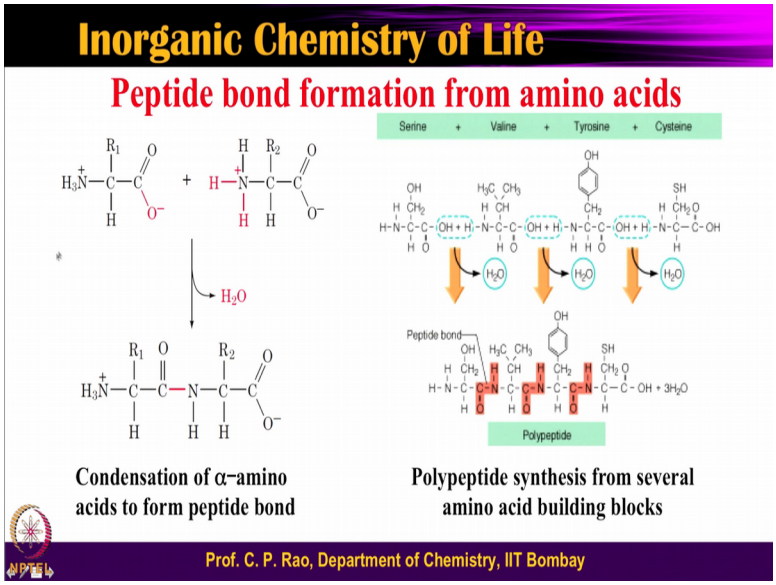
(Refer Slide Time: 01:42)



So, you see the amino acid is also referred as alpha amino acids, you have a carbon centre with one hydrogen and there are 3 substitutions, of which 2 one is the carboxylic group, other is the amine group, this is common for all amino acids. And what is diff what differs from one amino acid to the other is the presence of this particular group ok.

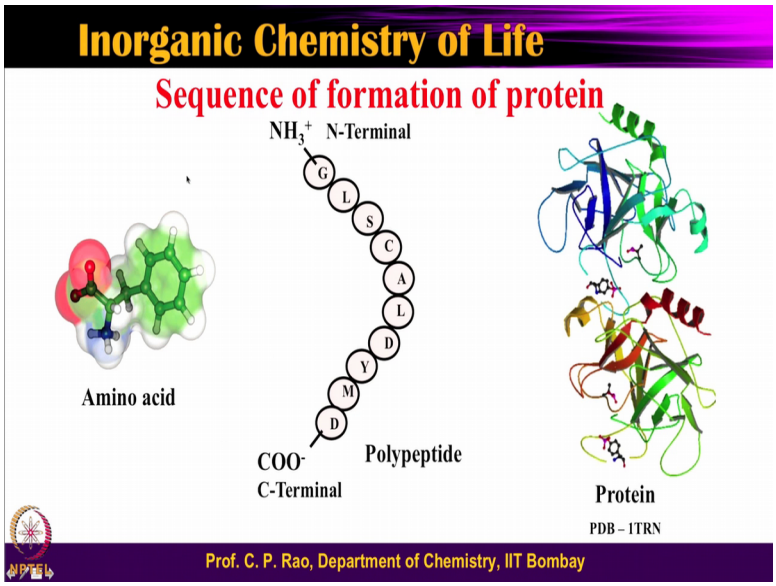
So, this particular group is R group. So, this is just referred as R, in other words the side chain. So, this side chain varies from one amino acid to the other among the 20 naturally occurring amino acids. Of course, today 21 is also added which is Selenocysteine in the literature. Otherwise we can take just as 20 amino acids. So, that is not a problem and so, you can see that the same thing as shown over there and the carboxylic group losing a proton, amine group gaining a proton this is referred such kind of a structure it is called the Zwitter ionic structure, and this structure of this Zwitter ionic or otherwise is dependent on the medium pH. So, the pH has the medium we will have a control whether you have a COOH or COO minus NH 2 or NH 3 plus these. So, generally these are represented by the kind of a structure these Zwitter ionic. Now the different amino acids come out because of the change in the R group. So, let us look at the next slide.

(Refer Slide Time: 03:21)

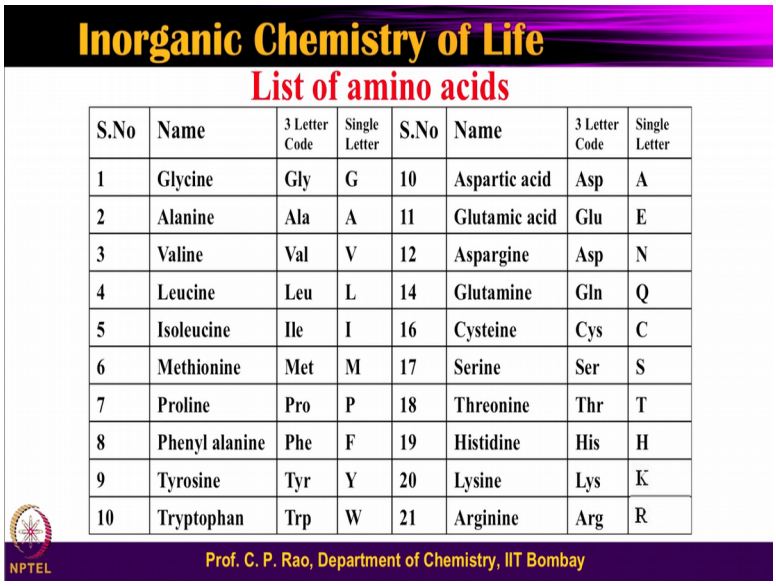


So, in this slide so, what we are seen? Having known that there are so many different kinds of a amino acids are there and these amino acids together can form yeah.

(Refer Slide Time: 03:29)



(Refer Slide Time: 03:33)



S.No	Name	3 Letter Code	Single Letter	S.No	Name	3 Letter Code	Single Letter
1	Glycine	Gly	G	10	Aspartic acid	Asp	A
2	Alanine	Ala	A	11	Glutamic acid	Glu	E
3	Valine	Val	V	12	Aspargine	Asp	N
4	Leucine	Leu	L	14	Glutamine	Gln	Q
5	Isoleucine	Ile	I	16	Cysteine	Cys	C
6	Methionine	Met	M	17	Serine	Ser	S
7	Proline	Pro	P	18	Threonine	Thr	T
8	Phenyl alanine	Phe	F	19	Histidine	His	H
9	Tyrosine	Tyr	Y	20	Lysine	Lys	K
10	Tryptophan	Trp	W	21	Arginine	Arg	R

Let us let us come once for these set of the amino acids that we have, we have glycine, we have alanine, we have valine, leucine, isoleucine, methionine, proline, phenyl, alanine, tyrosine, tryptophan and aspartic acid, glutamic acid, asparagine, glutamine, cysteine, serine, threonine, histidine, lysine, arginine. I have already showed explained you the kind of a side chains that are present into you know lectures prior to this, if you look at those slides you will find what are the side chains that you have.

I have given in the form of side chains the non-polar polar, polar with charged positively charged, negatively charged all these. I hope I have also explained those 6 7 8 different amino acids which are capable of bind into the metal synthesis is already talked to you. And in the protein literature, it is commonly the amino acids are referred by a 3 letter code or a one letter code and it is better that you all familiar with these once.

The glycine is written as Gly, alanine is written as Ala, valine is written Val, leucine is written as Lu Leu etcetera. Aspartic acid is written as Asp etcetera. Now if you look at this one code letters then you would see that the glycine is given as G, alanine see is A, valine is V, leucine is L etcetera. Now if you come to glutamic acid it cannot have G because G is already there. So, they are given E as the one, aspartic acid A that is fine, asparagine A is already given so, therefore, N is given here. So, glutamine and cysteine is C which is not used earlier so serine S, threonine T, histidine H. So, therefore, you have 3 letter code and one letter code and when you write protein sequence, you start using the one single letter code and when you write small peptides you generally use the 3 letter code, that is all the difference is.

Now, let us go back to this slide, as I said these are the amino acid side chains and the side chains differ as polar nonpolar charged positively negatively charged and they are all having ligating centers, they can bind to the metal centers. Now having known individual amino acid let us look at the peptide so, peptide a protein; protein is what? Protein is the one which is formed with peptide bonds. So, protein is the one which is formed from the peptide bonds. How is a peptide bond is formed? Peptide bond is formed between the 2 amino acids.

So, carboxylic end of the one amino acid reacting with the amino end of the another amino acid, and there is a elimination of water which is called dehydration reaction and as a result of that you will get a bond C O N H this is this bond is referred as the peptide bond or amide bond. If it is in a proteins; it is called peptide bond, if it is in simple amides then they called as an amide bond. So, they are one and the same; so, amino acid 1 amino acid 2. So, you can see R 1 and R 2. Sometimes this is R 1 and R 2 can be the same or R 1 and R 2 can be different. And as you can see that the carboxylic of this and the amine of this interacting together through the loss of water, you have an amine left over here and a carboxylic left over here. So, therefore, this is referred as the amine

terminal or amino terminal, carboxylic terminal or carboxylic C terminal so, N terminal and C terminal.

Let us look at little bit more or longer a bigger case. So, we have four different amino acids, the amino acid serine, valine, 3 tyrosine and cysteine. So, these once if you see here, the carboxylic of this amino acid, reacting with the amine of this amino acid, carboxylic of this one amine of that, carboxylic of this amine of that. So, if you have a condensation here, condensation here, condensation here; loss of water, loss of water, loss of water will give what 3 peptide bonds; one here, one here, one here. So, in this whole thing has got four amino acids, which is also referred as a tetra peptide; amino acid 1, amino acid 2, amino acid 3, amino acid 4 how do you recognize the amino acid; from the C alpha center.

So, C alpha center you can easily identify these once this is a tetra peptide. So, if you extend it becomes a poly peptide. So, from several so, now, what you say? Each one of these amino acid is called a building block in this. So, you have a building block.

Now, after this is formed what you have? This end you have a amine, this end you have a carboxylic, but otherwise there is no other amine and carboxylic you only have C O N H, C O N H. So, therefore, what is important? Only the side chains. In a huge protein let us say having 100 amino acids, 200 amino acids, 1000 amino acids then what is important? The important thing is that side chains in all those cases. So, let us re look at this is a tetra peptide, it has a peptide bonds over here and side chains in a huge poly peptide like protein having 100, 200, 1000 amino acids these will become terminals and the entire protein will have the side chains coming up.

So, the same thing is you as I talked to you earlier, you start with the amino acid combined with one more amino acid, combined with one more amino acid, combined with one more. So, you can starts stitching the amino acids together to form a peptide bond and the peptide bond connects these amino acids chains. So, here you can see one such poly peptide having 1 2 3 4 5 6 7 8 9 10 amino acid residues. So, you can call this as a deca peptide

So, in this you have glycine, leucine etcetera all are there. So, in the junction to these is a peptide bond, in junction to this peptide bond peptide bond. So, you have left with one amine part and one carboxylic part. So, the amine part is referred as N terminal,

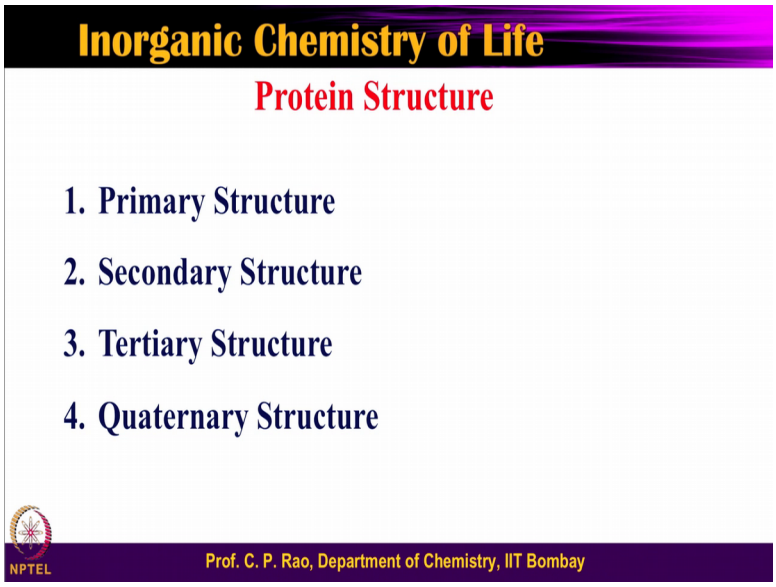
carboxylic part is referred as a C terminal. This is important because later stage we will come across the N terminal and C terminal situations where we talk about what is happening at this over the other.

Now you have taken the amino acids, you have connected them together like you take flowers and make a garland. So, let us take a different kinds of flowers, flowers with different a nature of, flowers with a different kind of a colors, then you take flower one with red color then put next one is a green colored, the next one may be yellow color or again go back to red etcetera. So, you are connecting them together like this one.

So, the this polypeptide is nothing like a nothing but like a garland, which is made from flowers of different colors or flowers of different types. Now such a polypeptide it converts into some kind of a matured level of the protein, which I will come in a while this is called structure. So, the structure this or the protein is formed and this is really ready for any function. So, I will explain the things connecting these one so. An amino acid a poly peptide to a protein and that is how we move from one to the other to the other.

So, this is just simple garland this is not that this is a decorated one. So, this decorations are coming from where I will explain in a while in this.

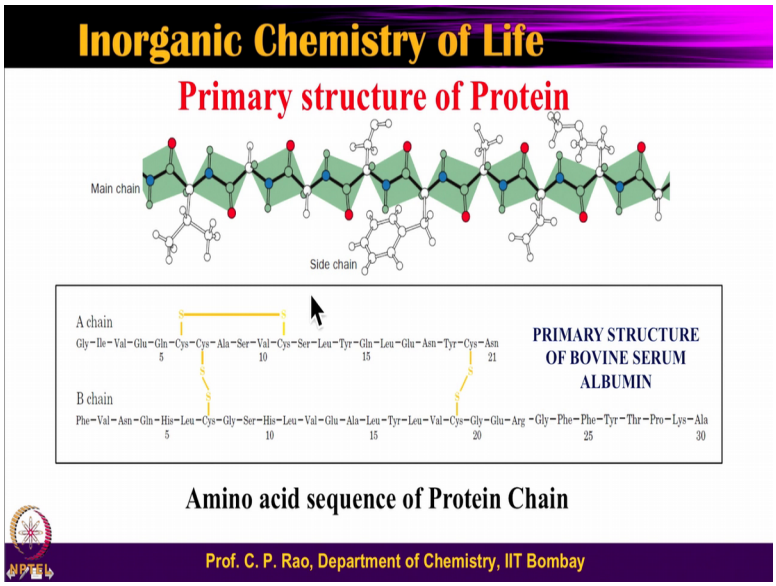
(Refer Slide Time: 12:00)



So in order to explain the protein structure that you have seen in the previous slide or in other words I referred it as a decoration and we need to introduce the four parameters; one is the primary structure, other is secondary structure, third is tertiary structure, fourth is quaternary structure. So, these are the different stages of these protein and we will go through this.

So, the primary structure is the one where you connect the amino acids together to make a poly peptide. The secondary structure is the one where such a poly peptide will not stay like a long thread there will be some kind of a rearrangements; structural rearrangements. And these structural rearrangements in various forms, I will explain in a while and such structural rearrangements taking shape in the entire protein and the protein takes a topological structure is called the tertiary structure. And such tertiary structured units joined together in a particular symmetry or symmetry element will result in a quaternary structure. I will again explain these things as we keep going through the next few slides ok.

(Refer Slide Time: 13:15)



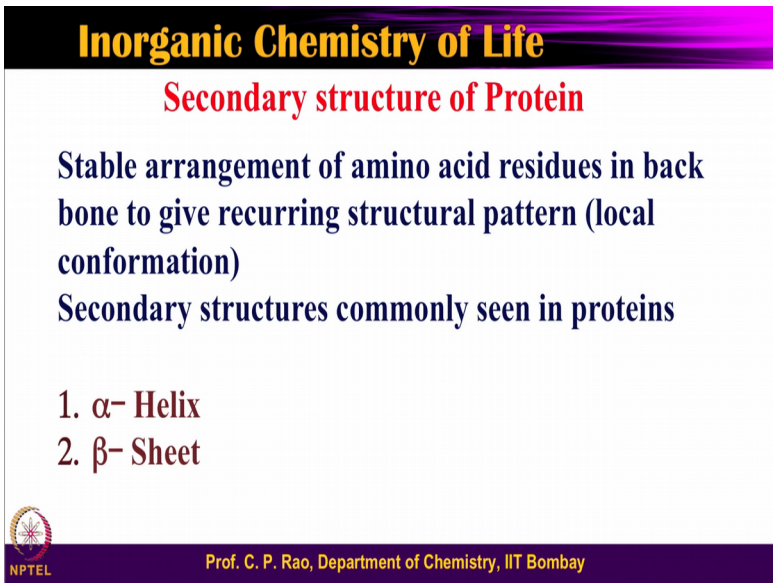
Let us look at the first part of the structure which is called primary structure, the protein. So, what the primary structure the protein has got? This is coming from one amino acid, second amino acid, third amino acid etcetera. So, what do you have? You have a C alpha center another C alpha center C alpha center C alpha center, C alpha center, C alpha center and each one of these have got a peptide bond you see that C O N H, C O N H, C

O N H, C O N H. So, this whole thing is referred is the protein chain. So, A is connected to B is connected to C is connected to D etcetera. So, that is that kind of a sequence is referred as the primary structure.

So, in a protein when we knows primary structure means we know the order in which the amino acids are connected. There are ways in means by which one can do that. This is not the stage where I will explain I would explain some other stage because you take a protein, you hydrolyze you start detecting you can find selectively hydrolyze and find which is the first amino acid which is the next amino acid which is the third amino acid etcetera etcetera. These are all very well known in the literature, but those are outside the scope of this particular course, let us not too much worry about it.

So, the here is a particular structure of the protein taken, primary structure of a Bovine Serum Albumin. So, it is a serum albumin protein from the bovine source. So, bovine is cow, the cow source and serum is blood and the albumin is the albumin like protein, and in that you have a chain structure shown here, you have the each amino acid is connected here and this number shows the 5th 10th 12th 15th 28th etcetera and there is another one. So, this is just to indicate the primary structure enough this much information this is the primary structure.

(Refer Slide Time: 15:15)

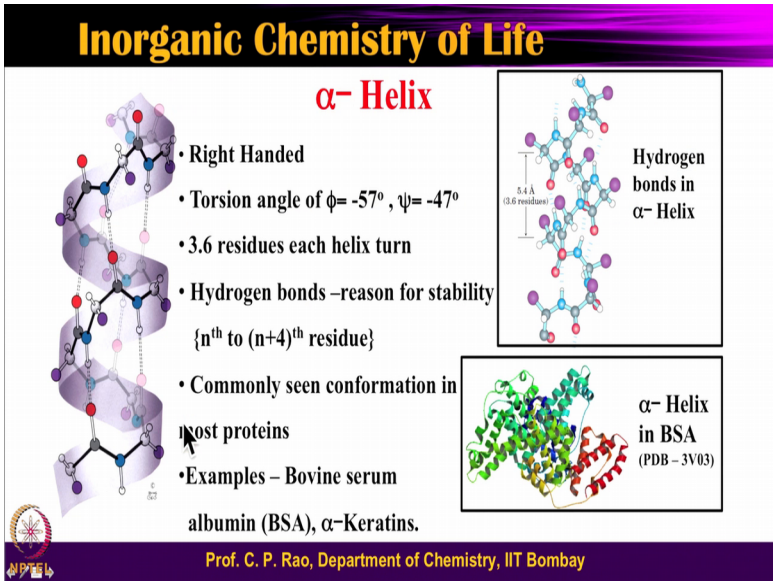


And now once the primary structure is formed, you can have the secondary structure of protein. So, what is the secondary structure? So, this will brings table arrangement amino

acid residues, in the backbone to give some kind of a structural pattern and that will bring to the local confirmation tool. So, the secondary structures which are commonly seen in proteins are 2 major works, one is the alpha Helix other is a the beta Sheet. The third one is left out portion which is called the random quail kind of a portion.

So, means no specific structure it is open kind of thing, so alpha helix, beta sheet and then the portions which are connecting these ones. Now we will look at each of these; what is alpha helix, what is beta sheet etcetera in the next few slides ok.

(Refer Slide Time: 16:03)



Alpha helix as you can see is that this particular protein, we have seen earlier a polypeptide. Now please take the poly peptide like this and goes this way, and goes this way and goes this way. Why should it twist? So, this twisting as you can see this some connectivity here and there is some connectivity here and there is some connectivity here. So; that means, these connectivity is or dominating. So, these connectivity is for dictating this particular poly peptide chain, to take this kind of a shape and this shape is referred as the alpha helix. So, what are these connectives? These connectivities are C O it is carbonate N H is the N H moiety.

So, each peptide amide bond has got C O N H. So, C O of this with N H of this, C O of this with N H of this ok, so similarly of this. So, you have C O N H kind of a hydrogen bonds; where are these hydrogen bonds? If you take nth residue and to the n plus 4th residue these 2 are connected and such a kind of things are formed in the alpha helix and

the one which is shown here is the right handed alpha helix because we are talking about all the l amino acids.

If you go to d amino acids, the situation will be changed let us not too much worry and what is the you see when you go from here to here, you are bringing a kind of a bend. So, this bend is also referred as a torsion. So, and some bonds you have to you have to twist and torsions. So, if you do such kind of a twist or torsion, then you will get this bend and those angles are given over here.

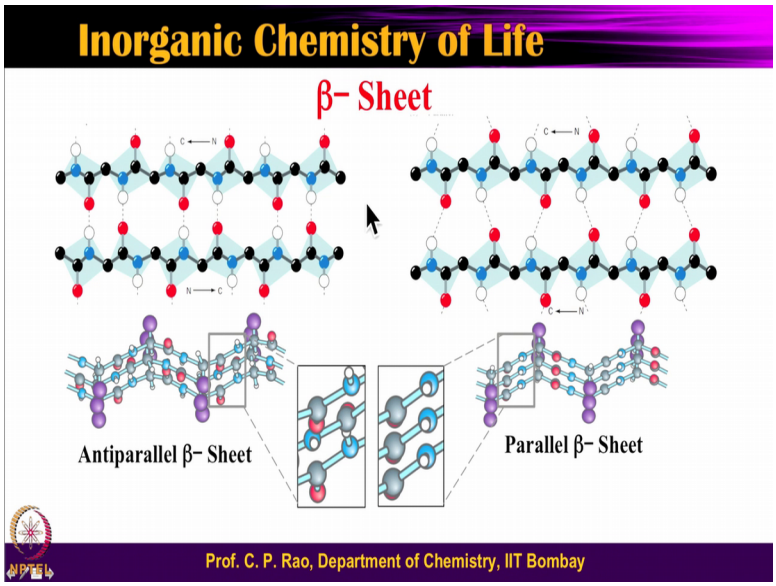
So, one other angle is CC alpha other angle is N C alpha ok. So, these are the 2 angles that you get because these 2 bonds are the single bonds. So, this single bonds will have a rotation, because of this rotations this kind of a structure takes. So, in this alpha helix there is 3.6 residues each helix turn has got 3.6 residues ok. And there are a 4 D atom hydrogen atoms in this if you start counting between this C O and the N H of that you will find.

So, the hydrogen bonds are shown over there with more clarity here. So, this is very commonly found in many proteins, large number of proteins example is shown here on the right side is bovine serum albumin you can see this. So, the same thing if I put a color it look like this, and there are different kinds of colors because they are coming from different regions, you do not need to worry at this stage why the color is different. So, all that you need to see is you can see this spiral kind of thing.

So, this spiral is nothing, but helical. So, these helical are alpha helical. So, as you would see that total protein is dominant by the helical structures, this small threat you can see. This small threats have no specific structure, they can be called as a random quail structure. And this structure anybody can see in the p d b protein data bank with this number if you go, you will get this particular protein structure you can get this one to ok.

So, I hope you understand the one of the secondary structures is alpha helix, and the alpha helix is stabilized by these C O N H hydrogen bonds of n to n plus 4th that is the driving force, and these hydrogen bonds as you know is stabilizes the structure because the hydrogen bond when is formed there is a energy being released ok. So, delta h is minus. So, therefore, total delta g will come to the minus favorable. So, therefore, such structure is formed favorably in the proteins.

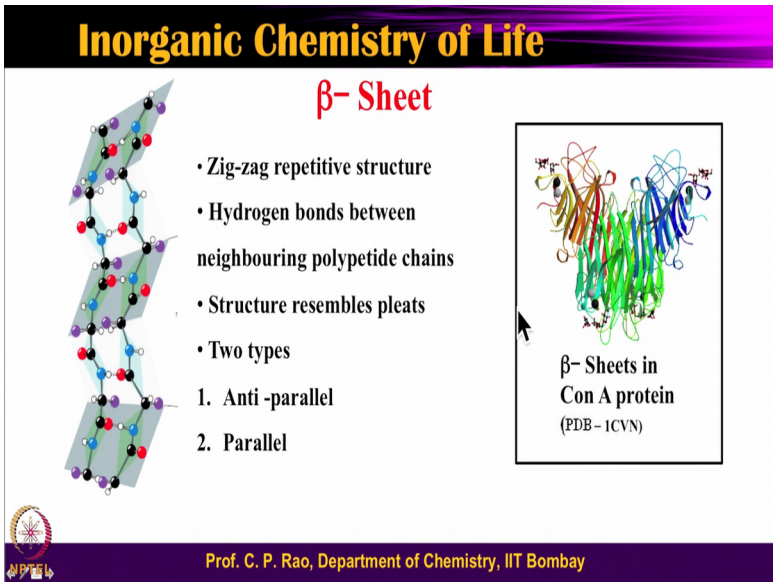
(Refer Slide Time: 20:13)



And go to the second category of this mentioned beta sheets. Beta sheets are formed not by the a twist, they are they just in their struggle in their zigzag kind of a fashion, but between one such unit another such strand you can say. So, therefore, you have a C O N H, another C O N H, another C O N H C O N H this side this side you have all this kind of things. So, such kind of a structures are called the beta sheets. So, the beta sheets you can see the n for this thing this side is n this side is carboxylic and for this strand this side is nitrogen or N terminal this side is carboxyl C terminal so, therefore, N terminal to C terminal and C terminal to N terminal. So, that kind of a thing gives you the sheet type of a structure which is called the antiparallel beta sheet. On the other hand both the N to C and N to C or C terminal N terminal C terminal N terminal they are exactly aligned it is called parallel. So, it is called parallel beta sheet.

Just because of this parallel and anti-parallel; obviously, you will find the corresponding sheet has a different properties. As you can see that this looks quite different from this one, this is coming from the parallel one this is comes in the anti-parallel one. So, parallel beta sheet anti parallel beta sheet kind of a structures. So, what I have been telling you about the beta sheet structures parallel beta sheet and anti parallel beta sheet anti parallel beta sheet and this is only their alignment is N terminal to C terminal and C terminal to N terminal here, and this side both are aligned the both the N terminals and both the C terminals are aligned together, and this also stabilizes the structure and this is also formed in number of proteins to.

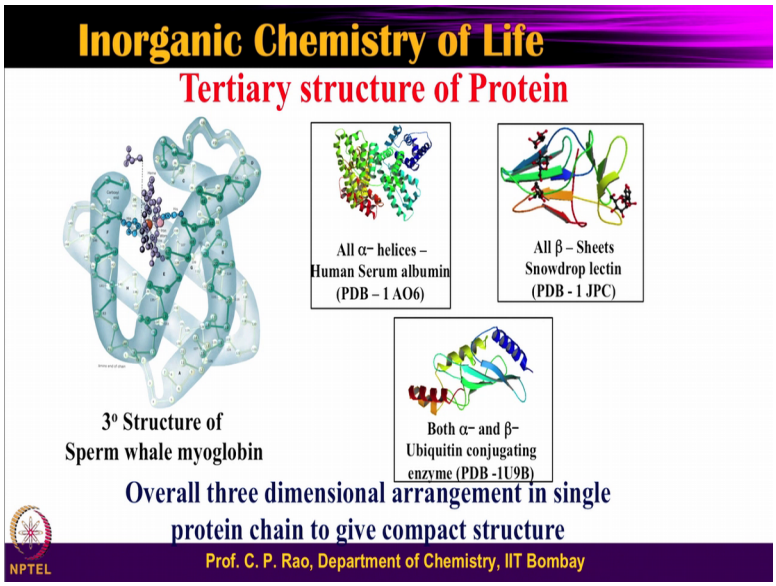
(Refer Slide Time: 22:25)



You can see in little better way here this zigzag repetitive structure and you have seen the hydrogen bonds between the neighboring poly peptide chains. So, this is referred as a structure, which resembles the pleats, the pleat kind of a structure ok. So, therefore, the two types anti-parallel and parallel. There is one example is shown over there this protein has hardly any kind of helix and it is mainly has sheets these sheets are shown over there here the sheets in the form of green color, in the blue color light blue color and the orange red etcetera you see that so, but all of them are beta sheet kind of a structures. So, this is a protein which is lentil and which is from lectin protein,

Concanavalin A with right now you do not need to worry about the name of this protein, is a beta sheet structure is dominating. And this can also be obtained from the protein data bank protein data bank PDB number is the number is given here protein data bank. So, if you go in to the internet and run this number under the protein data bank you will get this structure you will see that. So, this structure is primarily the beta sheet structure.

(Refer Slide Time: 24:00)

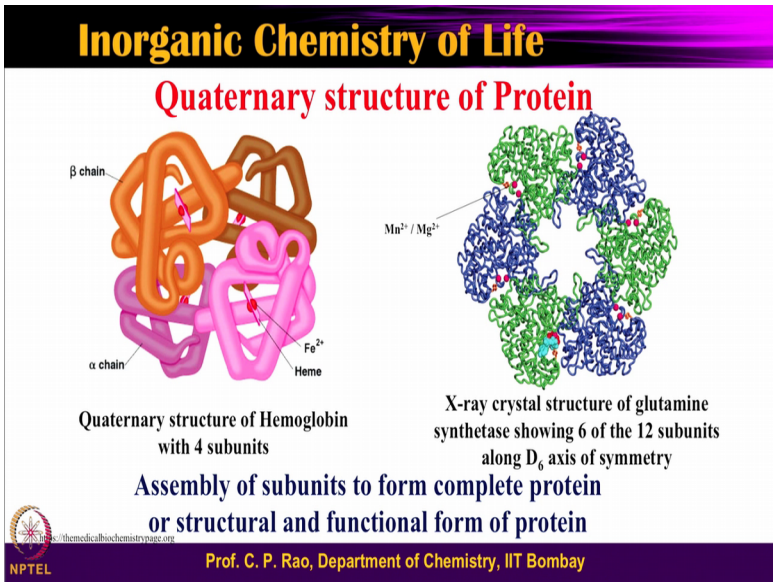


Is it that you should have in a protein all alpha helixes and you should have only beta sheets, that is not correct. There are proteins where both the alpha helix or beta sheets both are present to you can see in an example here. So, this protein has got not only alpha helical structures, but also beta sheets. This has primarily alpha helical this has primarily beta sheet, but this one has got both of these two. Now so, having seen these alpha helical beta sheets; that means, they are folding the protein in various forms. So, once you fold the entire protein you are getting to a kind of topological structure, and which is in turn stabilized by secondary structures. And the secondary structures are dictated by the protein sequence that we have and the turns that it can take place.

So, therefore, in the in the tertiary structure you have a topological features of the total structure. So, this is shown as a tertiary structure. So, tertiary structure is a super structure having it is subset is the secondary structures together, and the primary structure of course, is involved ok. So, several secondary structures together gives the tertiary structure for the total protein, the example shown over here is a myoglobin protein.

Of course we are going to look into the details of this at later stage and this is the portion where you have a heme containing portion. So, these not that the appropriate time to highlight the heme part of it, but otherwise you can see this is the tertiary structure of that ok.

(Refer Slide Time: 25:47)



Let us look at the quaternary structure. Now you have looked at the tertiary structure in the previous slide for the myoglobin and where it has one polypeptide chain and it has interior heme protein. Now you take four such units and try to arrange one of this is the this light orangish and a peach dark you know brownish orangish kind of color, and then you have a pink and darker pink.

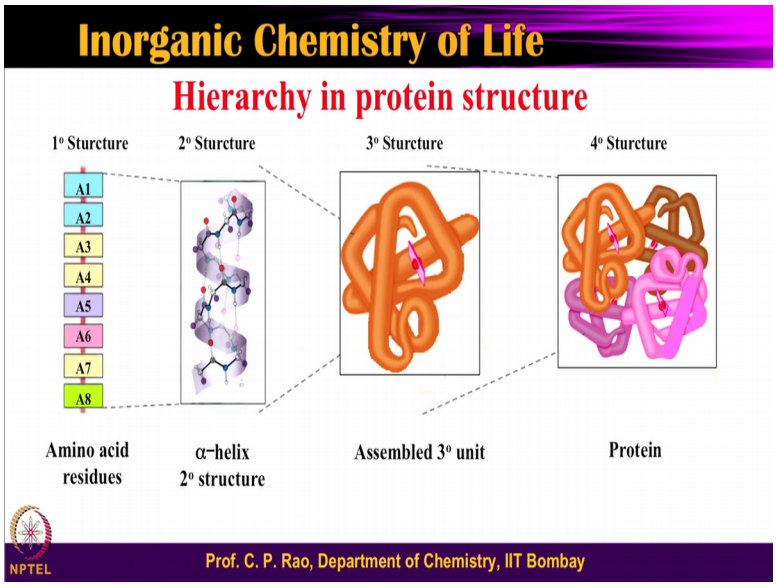
So, there are 4. So, these four each one of this is the polypeptide chain that have shown you in case of myoglobin turn. So, therefore, you have put the four of these but these are coordinated in a particular specific manner. So, you have a symmetry related, 2 of them are towards us or in towards us and 2 of them are towards the back side.

So, therefore, you have a structure which is being built in the form of a symmetry. So, this is symmetrically based arrangement of tertiary structures is nothing, but a quaternary structure. In some proteins the tertiary structure and quaternary structures are same sub units only one sub unit in some proteins there can be 2 sub units 3 sub unit there can be 4 there can be 5 there can be 6 variety of things are there. So, therefore, when you have more than one sub unit, they are arranged together in the form of a part in a particular in the in a particular symmetry and results in a quaternary structure.

Another example is shown over here you can see as a one, second, the third, the fourth the fifth and sixth and there are 6 move on the back side. So, this protein has got has 12 sub units and the 12 sub units of which you can see the top 6 and the bottom 6 on the

back and these are also coming from the metallo protein called glutamine synthase synthetase glutamine synthetase, where manganese and magnesium ions are present manganese and magnesium ions are present. So, this is again the quaternary structure of another protein. Now, what have we learned; the primary structure the secondary structure, the tertiary structure and the quaternary structure. So, all of these; obviously, are important ok.

(Refer Slide Time: 28:18)



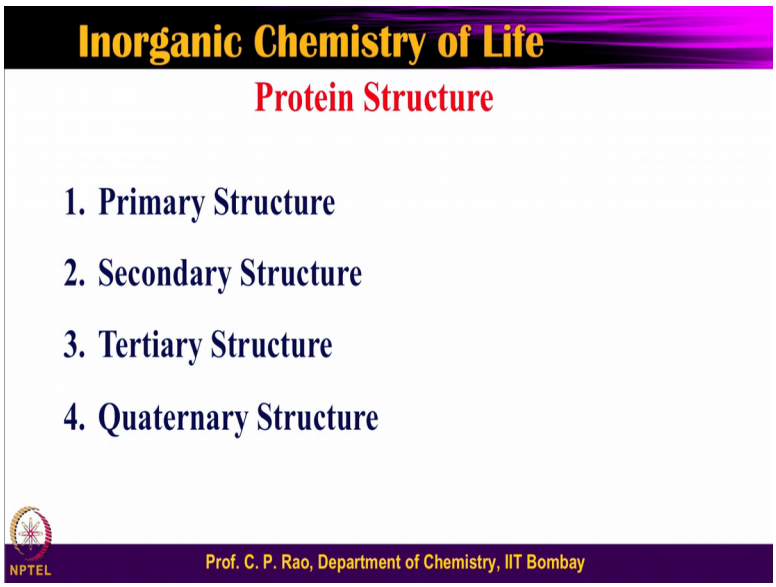
Now, let me put them together in one slide and explain you. So, what do you see on the left end left extreme is shown like A 1, A 2, A 3, A 4 means amino acids 1, amino acid 2, amino acid 3, they are connected; that means, it is a poly peptide. So, it is a poly peptide they are connected in this particular fashion therefore it is a primary structure; so primary structure or sequence, one and the same so, the amino acid residues are connected in this fashion. The secondary structure is the one where in this particular example where shown alpha helical, but you can have beta sheet also so, which I explained you. So, they take this turns because of stabilizing forces. And to make this you have to make a twist and which is something some diadem angle twist and these twists are made with respect to N C alpha, CC alpha ok.

So, cc alpha or N C alpha these 2 are single bonds therefore, single bond rotation does not cost much energy. There for that is what that energy is supplied by these stabilization in the hydrogen bond. So, hydrogen bond gains the energy and twisting requires the

energy. So, there is some expenditure there is some gain so. So, therefore, net there is a gain in this once. So, therefore, secondary structures are build and tertiary structure is, these kind of a secondary structures running across the entire protein is called the tertiary structure and such proteins are joined together in a particular kind of a the symmetry, and that results in the quaternary structure.

So, secondary, turn primary, secondary, tertiary, quaternary is as an example only one is shown secondary structure, but you can have other ones also. This is protein structure hierarchy this is very important because we will be later on talking about the protein function everything. Protein function comes not from primary structure not from secondary structure, it comes from the tertiary structure it can even come from the quaternary structure, but not from primary and secondary protein functions coming into these things.

(Refer Slide Time: 30:40)



So, in this lecture primarily I have focused and explaining you what is an amino acid, what is a peptide bond, what is the primary sequence, what is a secondary see structure, what is a tertiary structure, what is a quaternary structure. And I have also told you just by the having the primary structure and secondary structure function is not come, function comes from the tertiary structure and the quaternary. In some case for the poly peptides or the proteins, tertiary structure is quaternary structures same where only single poly peptide chain is there, otherwise they are different.

So, you can have tertiary structure showing the function, quaternary structure showing the function. So, the next in the next class we will look at the other bio molecules like nucleic acids etcetera and we will also look at the processes like protein synthesis and other biological parameters.

Thank you.