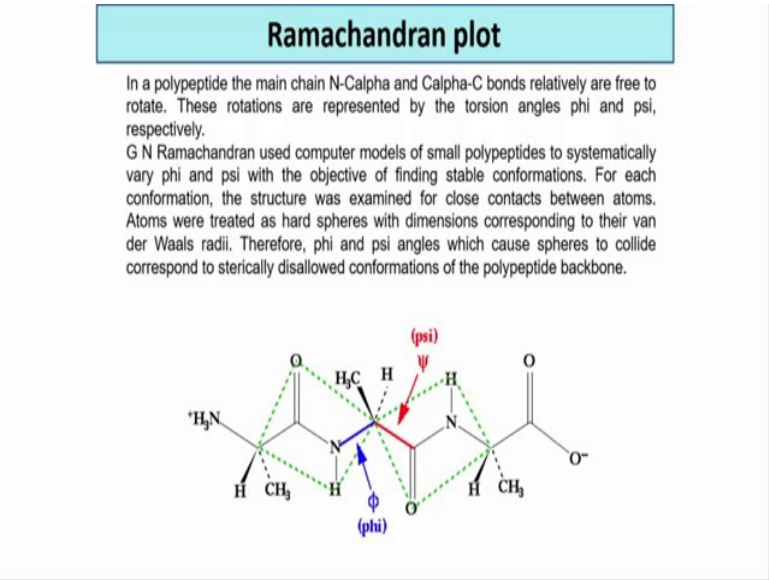
Organic Chemistry In Biology And Drug Development
Prof. Amit Basak
Department of Chemistry
Indian Institute of Technology, Kharagpur

Lecture - 11
Ramachandran Plot and Protein Purification Techniques



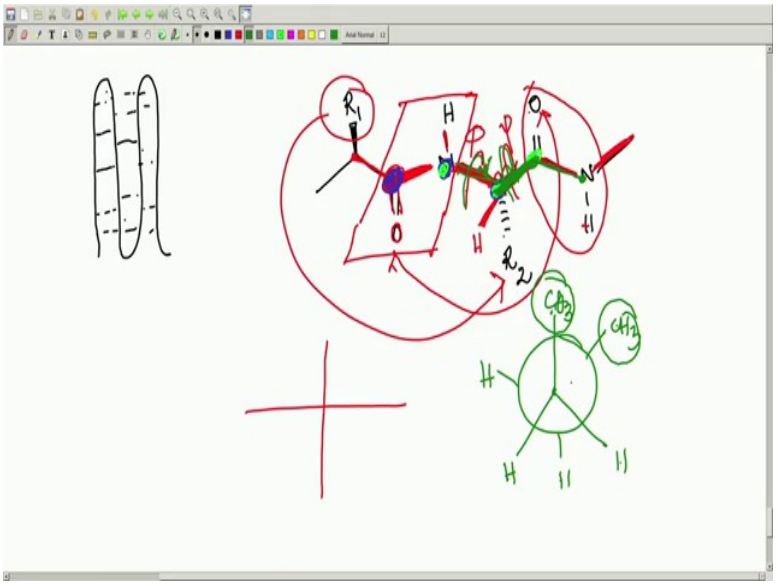
We have just discussed the different hierarchical structures of a protein or polypeptides: Primary, secondary, tertiary and quaternary. Remember all proteins will not have quaternary structure; it is only for those proteins which are present in multimeric forms.

We have so far talked about these local conformations, the local geometries, α helix then β sheet, β turn; all these things; but the issue is that if you want to make different geometries, what you need is rotation around bonds. Unless you are able to rotate the bonds, you cannot get different shapes.

So, there must be a lot of rotational flexibility in a polypeptide chain. So, in this session, we will first try to show that which bonds in the polypeptide are free to rotate. And which bonds are rigid in the polypeptide chain. Because it must be so that some bonds must be rigid because you are not getting infinite number of conformations; you are getting only a finite number of conformations and hence ultimately the protein takes a particular unique shape.

So; that means, there are many bonds which could be rigid. At the same time there are different bonds which are flexible, which allow the protein to ultimately come to the energy minimized three dimensional conformation. So, we should try to identify these bonds.

(Refer Slide Time: 02:27)



Now, a protein is basically having a peptide bond. First we will say this is carbonyl. Then you have this NH and then a side chain again here and there is stereochemistry here, suppose R^1 and there is a stereochemistry which is L .

Then again the whole thing repeats. So, there will be a CO here, then NH. This CONH bond is rigid, because in this NH is anti to the carbonyl. This takes a transoid conformation and that is quite rigid because of the double bond character between the carbon and the nitrogen. That means, you have again another rigid part. So, this is continuously coming, one after another, but in between them, there is the α carbon. This α carbon is connected by single bonds and there is no question of any resonance or any double bond character here.

So, these are the bonds which can rotate. So, this can rotate, but keeping the whole thing together, you cannot disturb that dihedral angle between this carbonyl and the NH; that is still there. That should be in plane and that the dihedral angle should be 180, that anti conformation you should be maintained here, but what you can do? You can take the whole thing and then start rotating around this C α and the carbonyl.

At the same time, the same thing is applicable that this C α and this nitrogen, this is a single bond. So, you can now take this whole plane containing the CO and NH and you can rotate this in one direction, clockwise or anti clockwise. So, these are the bonds which can now rotate. And that is the genesis of different conformations of polypeptides.

Now, there was one famous Indian biochemist Sir G. N. Ramachandran. He tried to see, that whether all types of rotations; or any angle of rotation is allowed or not. Because when you do these rotations, what will happen is that the inter-relation between these carbonyls will change and also the interrelations between these R^1 and R^2 will also change. Because if you start rotating, it will be found that all conformations may not be permissible. Ultimately, he came to a conclusion that all types of rotations are not permissible.

Now, before doing that, he defined two angles. You have a framework which is planar; that means, this carbonyl and this carbon nitrogen, this carbon-carbon bond; this is the backbone of the protein. We are just neglecting the substituents. We are talking only about the backbone; With this backbone, we are considering the rotation about this carbon nitrogen bond.

So, you are changing the angle between this bond, and this bond (which is now anti to each other). Now that is what is called the dihedral angle. Dihedral angle is basically between two planes. Now usually books will write that the dihedral angle involving carbon, nitrogen is called the Φ angle and the dihedral angle involving the carbonyl carbon and the α carbon is called the Ψ angle. But remember these are dihedral angles. Dihedral angles are basically the angles between two planes. Two planes intersect and that creates a dihedral angle.

So, two atoms cannot make a dihedral angle. Although, we will say that this carbon nitrogen makes an Φ angle and this carbon carbonyl makes a Ψ angle. But what are these Φ and Ψ? We take this carbon, this nitrogen and this carbon which makes one plane. And we take this one, this one and that one to make another plane. And the common bond between these two planes is this carbon nitrogen bond. So, this dihedral angle is called the Φ angle.

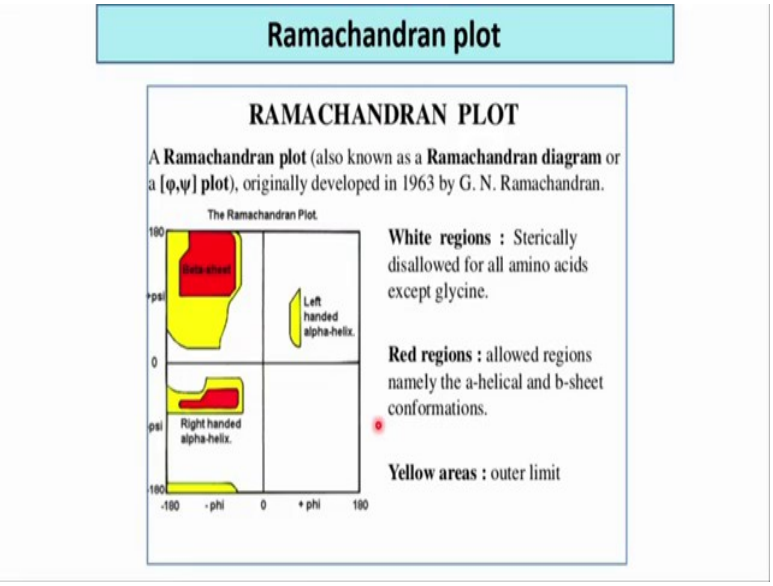
Similarly, when you talk about the Ψ angle, it is basically the angle between the two planes comprised of the nitrogen, α carbon and the carbonyl carbon; and the nitrogen, carbonyl carbon and this carbon respectivelySo, the common bond is between the carbonyl carbon and the α carbon. Just saying that merely mentioning the angle between the carbon and the nitrogen is not enough because dihedral angle is between the planes.

Again I repeat what are these planes? for the Φ angle, the planes are basically this carbonyl carbon, the nitrogen and the α carbon, and the nitrogen the α carbon and the carbonyl carbon on this side. And for the Ψ angle, it is the nitrogen, then the α carbon and the carbonyl carbon and here on the right side the nitrogen, the carbonyl carbon and this carbon.

Remember when I am checking the two planes we have a common bond in between, that is the way to do it. Now let us show you the model that what happens here. If you go to the slide see, it says that this is the Ψ angle and this is the Φ angle. But what I am saying is that actually we should say it is the angle between the planes containing this carbon, this nitrogen, this carbon; and the plane containing this carbonyl carbon, this carbon and this nitrogen; that is the angle..

When you consider the Φ angle, the common bond between these two planes is actually this carbon-nitrogen bond. When you consider the Ψ dihedral angle, the common bond is this bond connecting the carbonyl carbon and the α carbon bond. I can show you very simply that when we talked about the geometry of just ^{n}-butane, remember what was the dihedral angle between the methyls? We used to say that the dihedral angle in the gauche butane form 60°°. And sometimes casually we used to say that the methyl-methyl dihedral angle is 60°; but truly speaking, we should say that the dihedral angle between this plane containing this methyl, this carbon and the back carbon; and the other plane containing this methyl, the back carbon and the front carbon; that is the dihedral angle. So, that needs to be very clear. Now, the question is whether all Φ and Ψ angles are allowed or not.

(Refer Slide Time: 11:45)



Sir Ramachandran plotted the Φ angle along one axis; and the Ψ angle along the other axis. And you see that he basically bisected two horizontal lines; this is the 0 line that is the other 0 line; that means, here Φ is 0 and here Ψ is 0. A rotation means a complete rotation which will

be 360° rotation. He considered it that rotation from - 180° to plus 180° for the Ψ angle. And this axis is - 180° to + 180° for the Φ angle.

So, from this side to that side, Φ angle is plus, from this side to that side, it is minus; and from this downwards, the Ψ angle is minus and from the half the Ψ angle is plus. Now he has shown some of these regions in red color, some regions are yellow in color, some are not colored at all, only white; what does it say? .

He studied all these different structures (the β sheets structure, the α helix, the turn structure etc) and then he determined what are the Φ Ψ angles in that. And what he found that a point in the red region represents a particular Φ and Ψ combination. He said that this red region represents the most allowed combination of Φ Ψ angles; that means, the peptide will assume the Φ Ψ angle in such a way that it falls in this region, thus red region is the ideal case.

So, the red ones are the allowed regions namely for the α helix. Here it is α helix and β sheet conformations. So, now this one is for β sheet; that means, for β sheet, the Φ angle should be here. Atleast we can say that it cannot be 0 because the white region is sterically disallowed for all amino acids, except glycine; glycine does not have any steric effect. So, these are the β sheets; that means β sheet formation will be favoured, when the dihedral anle is close to 180; thus dihedral angle close to 180 is allowed for beta sheet formation.

And for the α helix, this is the red region. That means, α helix is allowed when the Φ Ψ combination is close to about 40-50. And on this side you see this is about 90 this is about 60-70 and that is close to 90 because this is 0 this is 180; so, the middle point is 90; that means, 90 not being taken up, close to that is the ideal region.

Now, apart from that we have a yellow region. This is the limit means sometimes you get values which are within the yellow region, but the ideal ones are these red ones. Sometimes you get a left handed α helix. That is the red, but these are much rarer, this is the limit.

Limit means when he draws a line, that means, if you determine the structure of your protein and see that the Φ Ψ angle falls here; that means, your structure analysis is wrong, that cannot be the structure. So, whenever somebody gets a structure of a protein, immediately he computes all the Φ and Ψ angles and these angles have to satisfy this Ramachandran plot. Otherwise that structure will not be accepted as correct; because if your Φ Ψ angle falls within the white region; that means, there is something wrong. So, this is a very interesting

and a landmark work from an Indian biochemist or biologist, Sir G. N. Ramachandran. In every biochemistry book you will see the Ramachandran plot.

Now, let us see that why the 0 0 combination is not allowed? We will just explore some extreme cases; why are these 0 0 combination or 180 180 combination not allowed?. That means, there must be some serious steric effect or electronic effect that is operating in these combinations.

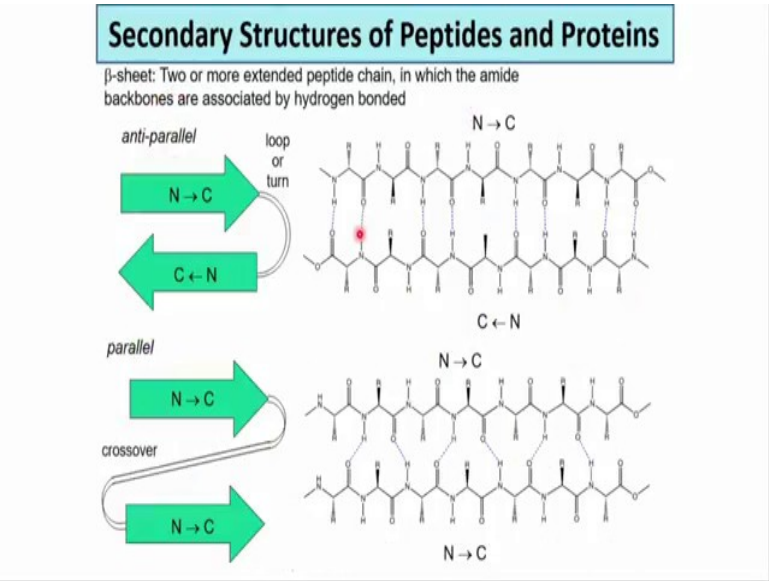
(Refer Slide Time: 17:25)

So, let us see; now I have a small peptide here usually the N terminus starts from the left side. So, I take the NH here, usually it starts from CONH if you have the amino acid from the left, amine is free then the CO then NH. So, I will do that way. So, CO then NH maybe this way let me see, let me; first fix the direction in which you will be comfortable, yes I think this is correct. So, this is your CO; one thing that one should remember is that the camera is looking at me. So, what is left in my case is right in the actual image that you are seeing. So, that creates a little bit of confusion.

So, let me try to again correct. So, this is a carbonyl, that is C double bond O. You can see two bonds here. And then there is this NH. Then there is the α carbon with the substituent here and this is the hydrogen, the green one is the hydrogen. Then you have again another carbonyl and this is NH. So, we have CONH here and we have CONH here so; that means, there are two peptide bonds so; that means, we are considering two peptide bonds, hence we are considering a tripeptide,

Ramachandran said that you first make the extended conformation. Extended means like when we discuss the β sheet structure, we drew the peptide chain in an extended form; I think I can quickly go back to that when you showed the β sheet structure.

(Refer Slide Time: 19:35)



You see, we wrote the peptide in the fully extended form like a linear hydrocarbon. So, similarly according to Ramachandran, first you consider the peptide in a fully extended form. If you do that, then this carbonyl this is the NH they are anti to each other; that you know why; and then you have this α carbon, single bond, you have this carbon, that is another single bond, this is the carbonyl, and this is the NH and then the again the α carbon.

Now, the question is, which one is Φ angle and which one is Ψ angle? The Φ angle is between the carbon nitrogen containing bond. So, this is the carbon, now I rotate it this way. Now the Φ angle will be between the carbon and the nitrogen (NH). And the Ψ angle will be considered between the α carbon and the carbonyl, see this is a single bond. And, I can rotate without disturbing the coplanarity of the carbonyl and the NH; also they are retaining the transoid geometry. So, I can start rotating this.

So, as I rotate, NH and CO they still remain anti to each other. So, this is the full complete rotation. So, this is the extended conformation. Here are the ΦΦ and the Ψ angles. They will be 180. If this is fully anti to each other, then they are 180 because, if you look at this dihedral angle, like the two methyls when they are opposite to each other. I think if you look

at this way, this bond and that bond they are anti to each other and that makes it a Φ angle. Now this is the Ψ angle this is the carbonyl and this is the Φ angle this is the Φ.

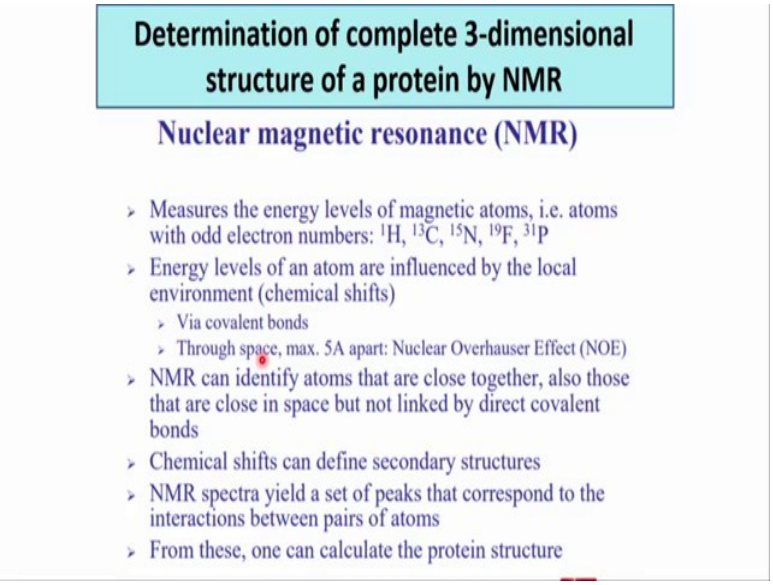
So, when you talk about Φ, you see this carbon-nitrogen bond carbon-carbon bond and this other this bond. And then you can say it; it is little complicated; I think the complication is even more because you are seeing the mirror image of what I am holding; that creates more problem. But again I repeat you see it first of all in the fully extended form, if you do that then it will be in the zigzag. And here that ideal angle combination Φ Ψ is 180 180. And that you see that in the β sheet conformation, according to Ramachandran, that around dihedral angle 180 and 180 is allowed. Where is Ramachandran?You see this is the region that is the 180 region and this is the 180 region. So, that is allowed.

Now, if you start rotating, just follow me, if I am rotating, this is changing the Φ angle. So, I rotate the Φ angle, this is 90° rotation, and then ultimately it comes to 180° rotation. So, when I do 180° rotation, I see one thing that this carbonyl and this carbonyl are very close to each other. So, what will happen? Oxygen has lone pair of electrons. There will be a lot of the electronic repulsion between the lone pairs of oxygen.. So, that is why, this 180 and 0 is not allowed. You see in this plot, at 180 and 0, you do not see any point here, so that is not allowed.

Again I come back to the original conformation, that is the anti conformation we started. So, I rotated this by 180. So, it goes there. And now I rotate the other one, this is the Ψ angle. I rotate, without changing the NHCO, I rotate it by 180, if I do that, see what happens.

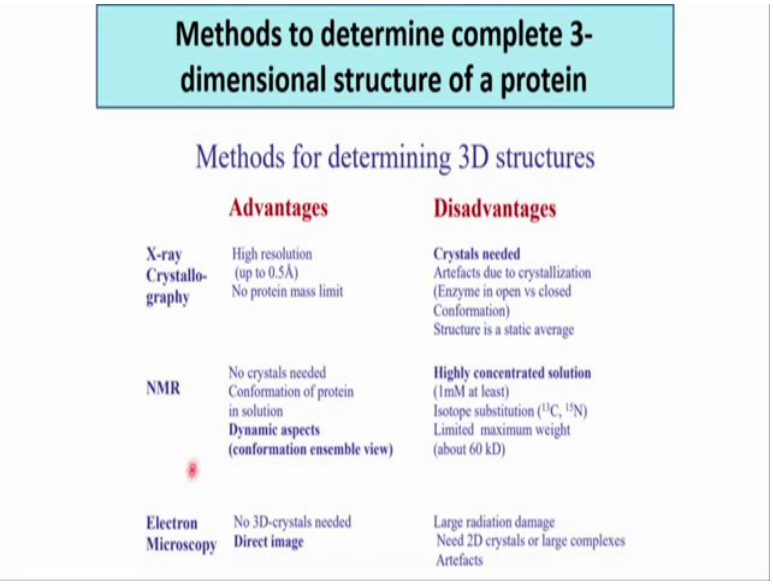
Now, there is even more problem. So, I have rotated from 180, I went to 0; also from 180, I went to 0; because Ramachandran said that the extended form has 180° angle. So, when I rotate by 180, I go to 0 and this is also going to 0. Now 0 0 combination is not allowed and you see 0 0 combination is not allowed. ,So he actually , computed all these and determined the crystal structure, analysed it and from that, he came out with this plot. If there is any difficulty, again we can come back and show what this Ramachandran plot is.

(Refer Slide Time: 25:15)



Now, regarding the determination of the 3-D structure of proteins, there are different techniques, to determine the 3-D structure of proteins. One of the best ways is X-ray crystallography. If you can isolate a crystal structure of a protein then that is the ultimate structural evidence. Methods of determine complete 3-D structure.

(Refer Slide Time: 25:39)



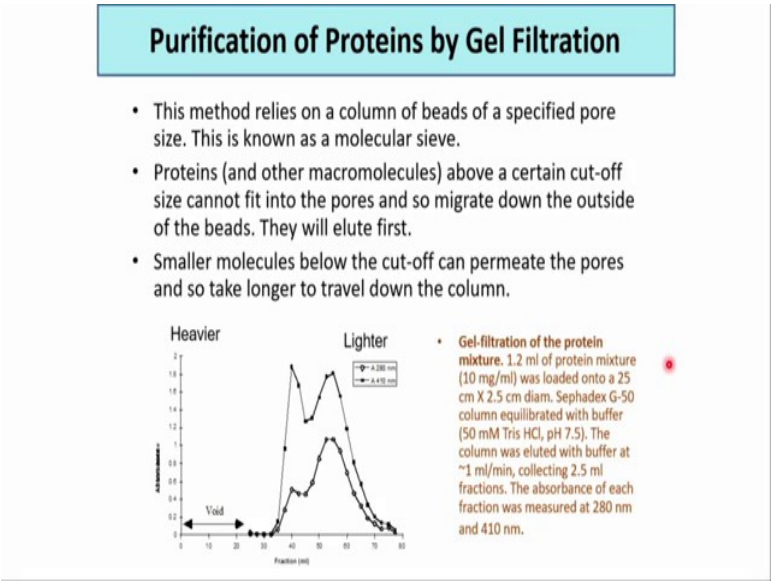
Usually there are three types. One is X-crystallography and that is the most reliable one; then there is NMR spectroscopy, very high field NMR machine is needed; through NMR spectroscopy you can determine the 3-D structure of a protein; however, there are some

disadvantages of NMR spectroscopy. You cannot handle very large molecules in NMR, but at the same time there are some advantages as well. In X-ray crystallography, if you cannot crystallize a protein then you cannot get the crystal structure. . X-ray crystallography needs crystals, and this structure is a static one, we get the average structure in the crystal.

But NMR is recorded in solution. So, what you see is exactly the dynamics of the system. But you need highly concentrated solution to get the signals; the peaks are actually very complicated ones; you should have isotopically labeled carbons, isotopically labeled nitrogens; there are techniques to do that. But by NMR you can do the structure analysis, but not for very large proteins there is a limit; at present, up to 60 kilo Dalton, you can do the structural elucidation.

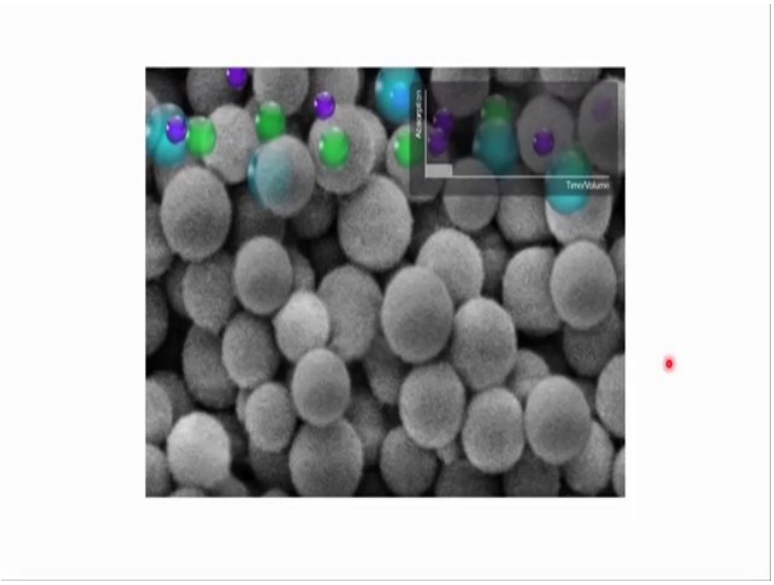
Electron microscopy started long back, but recently there is a technique which has won the Nobel Prize 3 or 4 years back; that is called Cryo Electron Microscopy. And through that many of the crystal structures have been solved these days.

(Refer Slide Time: 27:29)



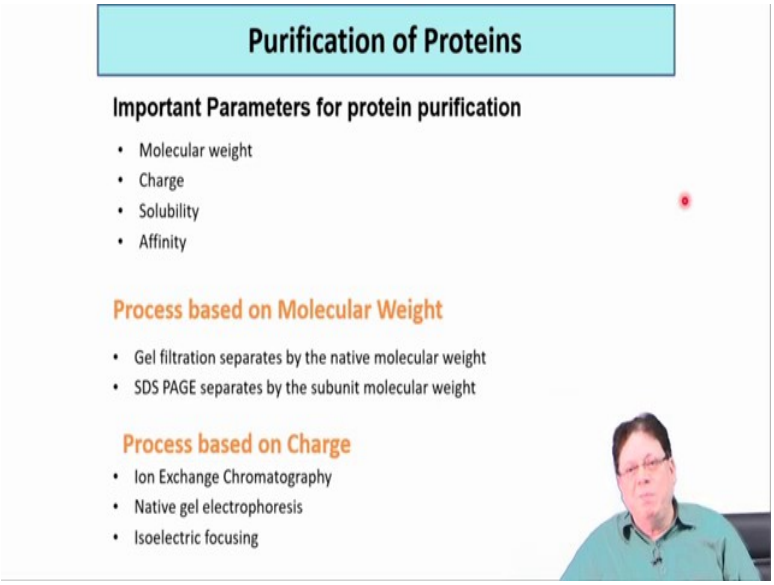
Now, let us come to the purification of proteins. So, we have completed the structure analysis. Now we come to purification of proteins.

(Refer Slide Time: 27:35)



First of all, for any purification method, you have to see that what type of property, these molecules have. So, that you can make use of them during the purification.

(Refer Slide Time: 27:49)



So, let us see what are the properties that a protein possesses? One is the molecular weight. Some proteins are small; some proteins are big. Thus, protein purification can be based on molecular weight if the proteins differ in size and in molecular weights. Or there could be charge difference because we know that proteins have different side chains which are charged. So, there may be difference in charges of different proteins.

There can be a difference in solubility in buffer or in water; or there could be some affinity for a particular agent and we can utilize that affinity for that agent to separate protein molecules. Like if you are trying to catch a particular fish in a pond, you should use a particular fish hook which the fish has affinity towards. Similarly if you have affinity for a protein, you can use a selective reagent to capture the protein.

So, these are the four parameters that one can use. Now there are a few processes that are based on molecular weight; that is equivalent size of the protein (in most of the cases). One is called the Gel filtration and the other is called the SDS PAGE (that is Sodium Dodecyl Sulfate and polyacrylamide gel electrophoresis). And by these two, actually we can separate a mixture of proteins as per their size and the size is nothing, but the molecular weight.

Of course there is an important point here. That when I talk about size, some proteins maybe really globular in shape and some proteins may be little bit stretched in a linear fashion. So, their size maybe different; a high molecular weight molecule may have lower size then a low molecular weight molecule, we will come back to that. Then the processes based on charge are Ion exchange chromatography, Native gel electrophoresis and Isoelectric focusing. So, we will consider all these five. Let us see how many we can do today, otherwise we will do it in the next session.

The first one is protein purification by gel filtration. Gel filtration is also called size exclusion chromatography. What it does? You fill a column with some matrix which are like molecular sieves. Sieve means which can filter out something. They look like small beads also, very tiny beads.

Now, you add the protein solution of different proteins; or any other macro molecule also you can do polymers, even chemical polymers can be separated by this process.This is done depending on the type of bead that you are using, you can have a cut off. First of all, you have a series of beads in a column and then you add the protein solution and what will happen now? There will be separation between the large proteins and the small proteins.

Why is that the, I think that will be clear in this slide. See these are the beads that you are using to separate the proteins. Now you have a mixture of proteins you can see some green balls here, that is representing a protein; a green is bigger one that is a bigger protein; this is a smaller protein there may be others also.

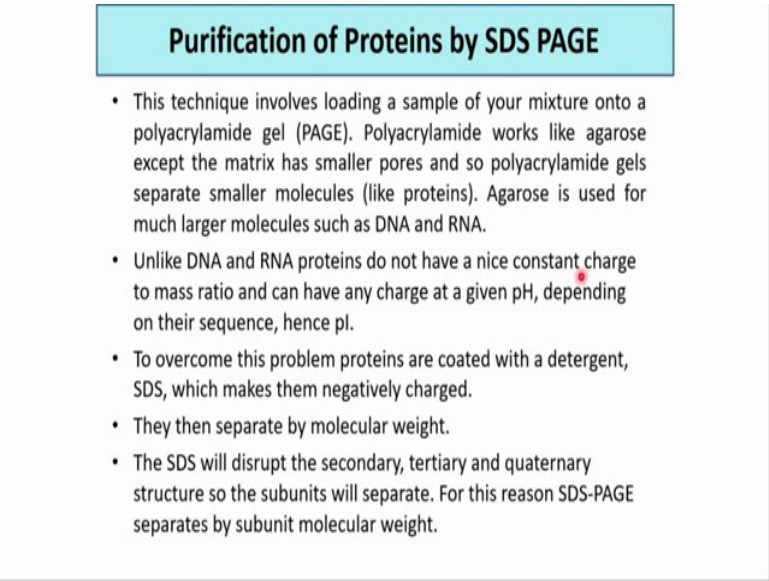
Now, you have these beads all packed in a glass column. Now as you add the protein solution, what will happen?After a few minutes, you will see that they have started dropping. Because this is by gravity; they will slowly come down; you are adding a solvent and maybe water and then this is coming down.

Now, what will happen? There are many holes here as you see, there are many pockets. The smaller ones because of their small size, will fit into those pockets, but the larger ones cannot fit into those pockets. So, basically what will happen? The smaller ones will fill the pocket and then as the solvent goes down, they will slowly come down. So, their migratory aptitude, their velocity of coming down will be much less than the larger ones because larger ones cannot fit into the small holes and then as time goes on, you will see now the larger ones have come ahead of the medium ones and the medium ones are ahead of the smaller ones or the smallest ones.

So, then ultimately what will happen? The larger ones are almost now coming down, coming away from the column. So, this is the principle of gel filtration chromatography. And you have different sizes of these beads. And basically if you want to say a cut off, means they will say that a 10 kD cut off gel; that means, molecules which are size less than 10 kD (kilo Dalton) they will be attached to the bead and the others will flow through very quickly. Or you can get 40 kD cut off; that means, your protein of interest has a molecular weight of more than 40 kD, then you use the 40 kD cut off.

So, this is a primary purification step that every protein chemist does. First you determine the molecular weight and then once you determine the molecular weight, you can determine which type of bead you should use. They are made up of carbohydrates, called sepharose. Sepharose gels you can use. There are different cut off levels that are available and then do the size exclusion chromatography. Because the basis of separation is actually size.

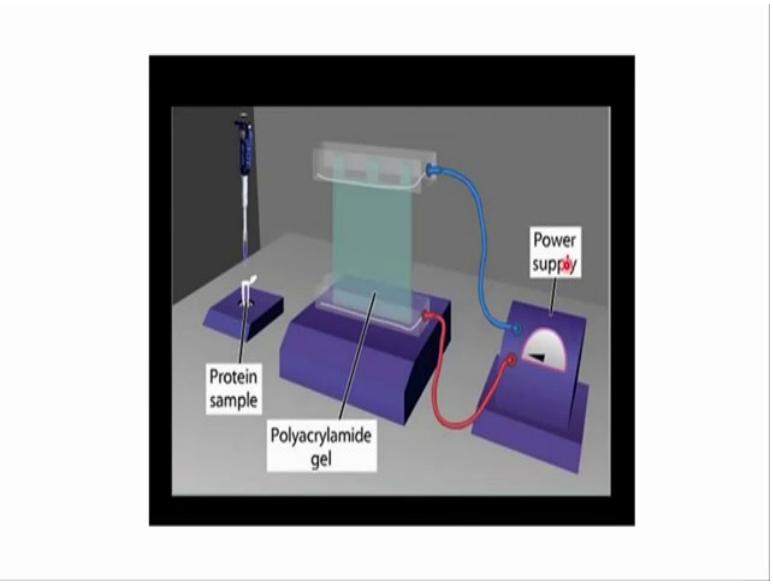
(Refer Slide Time: 34:33)



In SDS PAGE, SDS is sodium dodecyl sulphate. It is a long fatty chain with a sulfonic acid and PAGE is polyacrylamide gel. Acrylamide is polymerized along with the cross linker and that makes it like gel; it is said polyacrylamide gel.

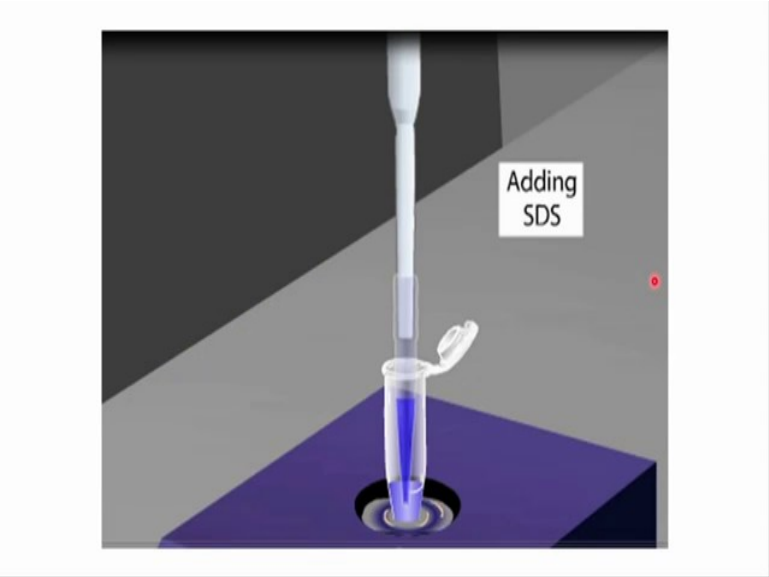
Agarose is another gel; there are different types of gels available. Anyway, when you have this gel, in this case you do electrophoresis; remember I said earlier; electrophoresis is basically separation by applying a voltage difference and when some compounds are placed in a matrix. So, according to the charge as well as their mass, they will move.

(Refer Slide Time: 35:37)



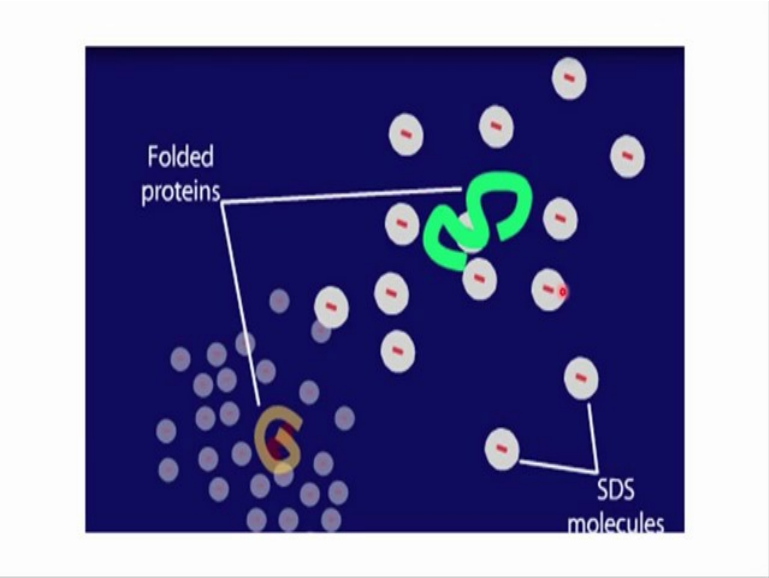
So, this is the apparatus that is it called the polyacrylamide gel apparatus. What it doesThis is the gel and these are called grooves, where you can add your protein solution and on this side there is a tiny test tube like thing which are made up of actually plastic. They are called eppendorf tubes and this is your micro pipette and there is a power supply. This is the anode and that is the cathode.

(Refer Slide Time: 36:15)



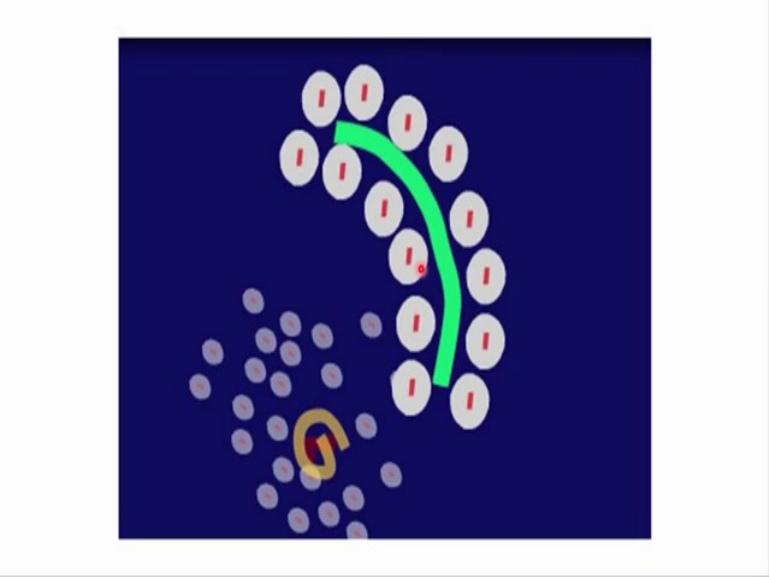
So, what you do here? It is your protein solution, suppose lot of proteins are there. What you do? You add your SDS (sodium dodecyl sulphate), then heat this to a certain temperature; not very much maybe 60 to 70°. What happens as you heat?.

(Refer Slide Time: 36:35)



The proteins were earlier in a globular form, in a kind of folded state. You add these folded proteins into SDS which are negatively charged. Sodium dodecyl sulfate; that means, it is sodium salt of a big anion; hydrophobic chain followed by a head group carboxylate. And now what will happen? As they are being mixed and little bit warmed, this protein will unfold totally.

(Refer Slide Time: 37:07)

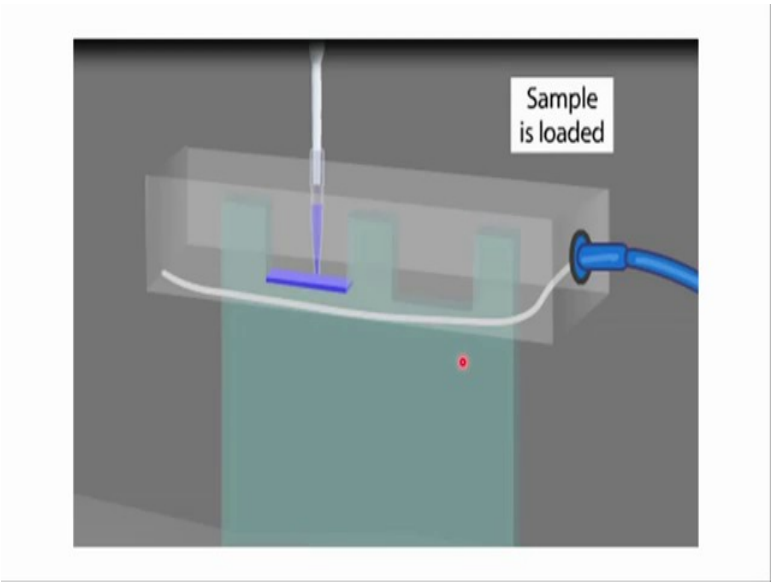


And it will be surrounded by all these SDS. The surface is surrounded by the SDS. So, this protein will also do that; that will also become linear and that will be surrounded by SDS

PAGE. So, now what will happen? There will be a difference in the length of the large proteins and the small proteins. But the charge density will remain the same.

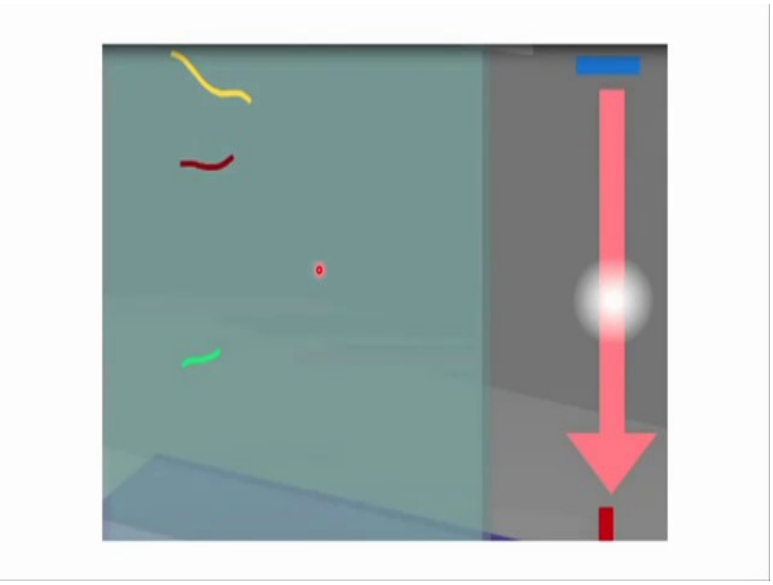
Because at every point, there is a sodium dodecyl sulphate. Suppose this polypeptide has got ten amino acids. So, it has got ten sodium dodecyl sulphate on one side; that means, per amino acid you have one. If this has got five amino acids it will also have five dodecyl sodium dodecyl sulphate surrounding it. So, ultimately the total charge density of the surface will be same, for all the proteins.

(Refer Slide Time: 38:17)



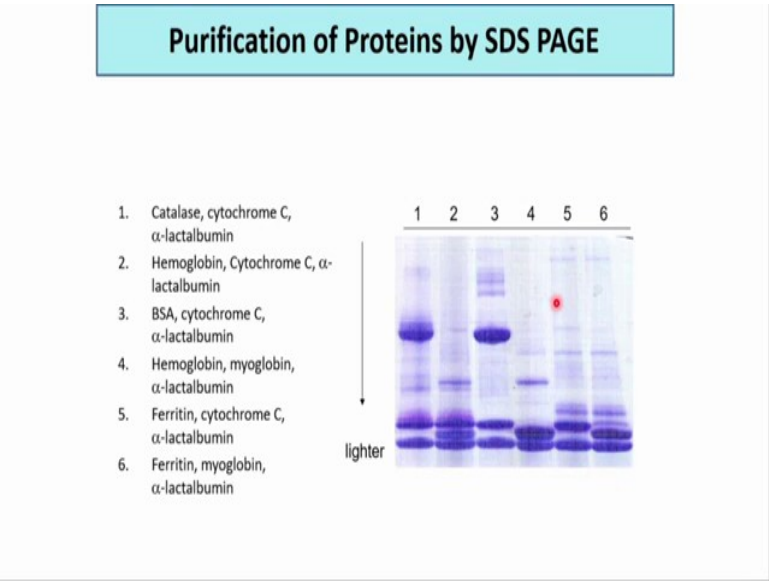
So, now with similar charges on the proteins, you add it to the groove. These grooves are called wells. Now after adding the protein, you apply the voltage.

(Refer Slide Time: 38:27)



As you apply the voltage, now what will happen? As the charge density is same, so only the molecular weight will be the deciding factor. This is a small molecular weight protein and that is a medium and finally that is large. So, the large ones will come most slowly then the middle one, and the small protein is the one which is moving ahead of the others. Remember this is your annode (the positively charged electrode), because everything is negative bridges surrounded by SDS. So, that is why this is called SDS PAGE, but this is also called the polyacrylamide gel electrophoresis under denaturing conditions.

(Refer Slide Time: 39:23)



Because you have denatured the protein; you have destroyed the inherent 3-D structure of the protein. So, after this electrophoresis is over, usually you add another dye which moves ahead of all the proteins. So, when the dye reaches at the front, you stop the current and then you take the gel and put it in a dye which is called coomassie blue. And then you stain it with coomassie blue, keep it overnight in acetic acid methanol and then wherever there are proteins, you will get blue bands. So, these are the proteins that you see so; that means, if you have different eppendorf tubes, your tube 1 has got a major protein here, another major protein here, another major protein here.

So, depending on the intensity, you usually put a molecular weight marker. You can give standard proteins on one lane; where the molecular weights are known. So, from that, you can tell, what is the molecular weight of the protein. Because you have molecular weight markers on one side; either this side or that side. So, that you know the molecular weight of each protein. That is the SDS PAGE purification.

You also have very similar method called native gel purification. In native gel, you do not add SDS, you just apply voltage across the gel. Here the protein maintains its 3-D structure; it is in its conformationally active form. But native gel requires a lot of voltage because first of all SDS gives a lot of negative charge, hence the charge density is very high. For native gel, the intrinsic charge ultimately will decide which way it will move, how fast it will move; but native gel will not be dependent on the size of the protein. It depends only on the charge.

If you want to separate in the native form, then you have to do the native gel electrophoresis. Otherwise generally for identification purpose, we do the SDS PAGE. I think the principle is now clear. So, SDS PAGE depends on the size of the protein and native gel electrophoresis depends on the on the intrinsic charge of the protein, at the pH that you are doing the gel electrophoresis.

Thank you very much.