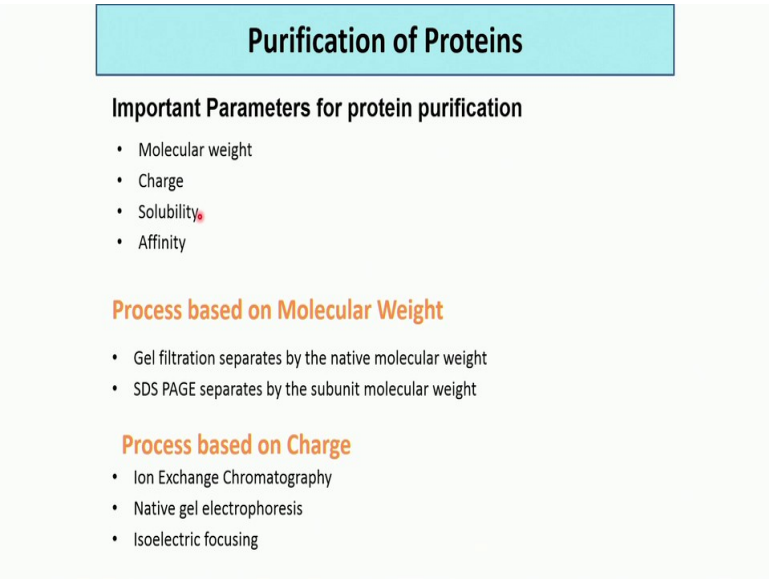
Organic Chemistry In Biology And Drug Development
Prof. Amit Basak
Department of Chemistry
Indian Institute of Technology, Kharagpur

Lecture - 12
Protein Purification Techniques (Contd.)

Welcome back to this course on Organic Chemistry in Biology and Drug Design.

(Refer Slide Time: 00:38)



Let us recapitulate what we were discussing; we were discussing about the methods to purify the protein molecules from a mixture of other proteins. The properties that can be used to separate the proteins are discussed here. The properties include molecular weight as different proteins have different molecular weights.

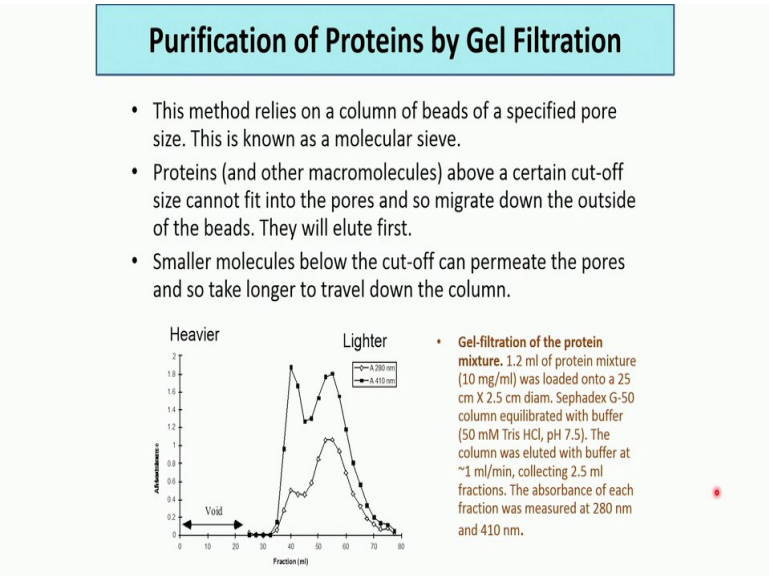
Based on the molecular weight, there are purification methods which separate on the basis of molecular weight. There are also methods which separate proteins on the basis of the charge that they possess; you must remember that the proteins are made up of amino acids and it is not only the terminal ends that are charged as NH_3 plus and COO minus, but at the same time, the side chains can also have positive charge or negative charge. So, the overall charge depends on the primary structure of the protein. Primary structure means the sequence of the amino acids.

And different proteins have different charges which can be utilized to separate them via electrophoresis. Then different proteins have different solubilities. Usually the solubility of proteins in water will decrease as you add inorganic salts which increase the ionic strength of the solution. So, there are precipitation methods like adding ammonium sulfate can precipitate proteins and you can maintain different concentration of ammonium sulfate to specifically precipitate various proteins.

And the fourth one is called affinity chromatography. In this process, you actually attach a molecule, which has a high affinity towards a particular protein and then as you elute the protein solution, the protein which has got higher affinity with the affinity column will stick to the column while others will pass through. And then you can apply some buffer solution to detach the protein which is attached to the affinity column. In this way, you can purify the protein.

Next, we are going to discuss two methods; one which is on the basis of molecular weight and another is on the basis of charge. I think yesterday I showed this to you.

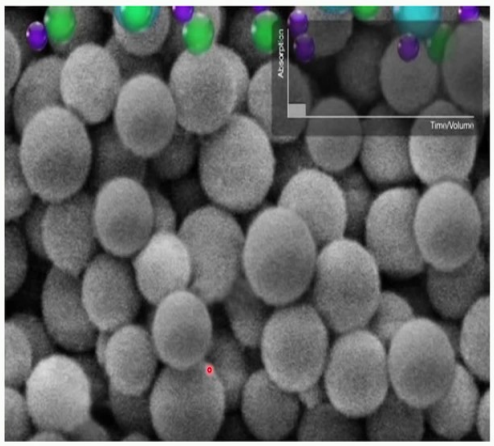
(Refer Slide Time: 03:18)



Now, let us just have a quick glance of what we have discussed; first thing that we discussed was the purification of proteins by gel filtration. This is also called size exclusion chromatography, because here the column is a glass column which is filled up with macromolecules which are basically polymeric beads. They have pores in them. When you

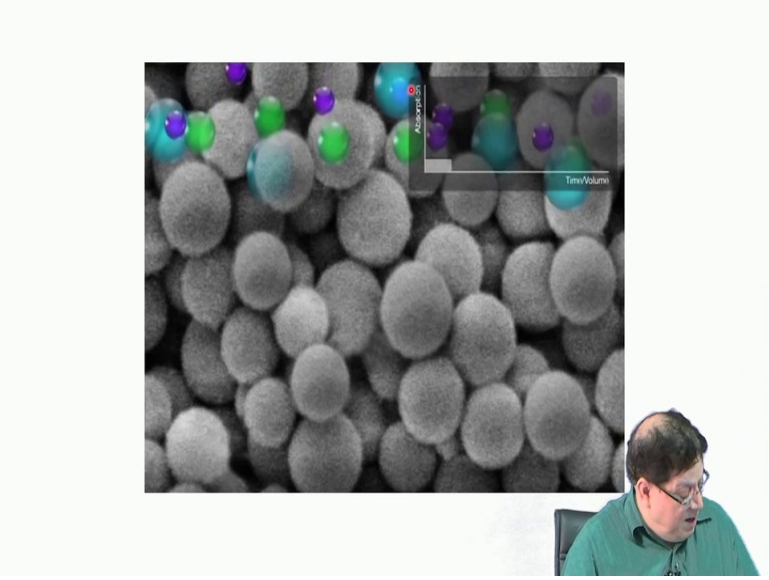
apply the protein solution, the large proteins cannot go through the pores of these beads instead it will just flow through them, straight away. Whereas the medium or the very small sized ones will interact with the pores and that will slow down their movement. So, ultimately what will happen? If you have a chromatogram like this, it is the high molecular weight proteins that will first come out; that means, the protein molecules with higher sizes will come out first, followed by the proteins of the lower sizes.

(Refer Slide Time: 04:38)



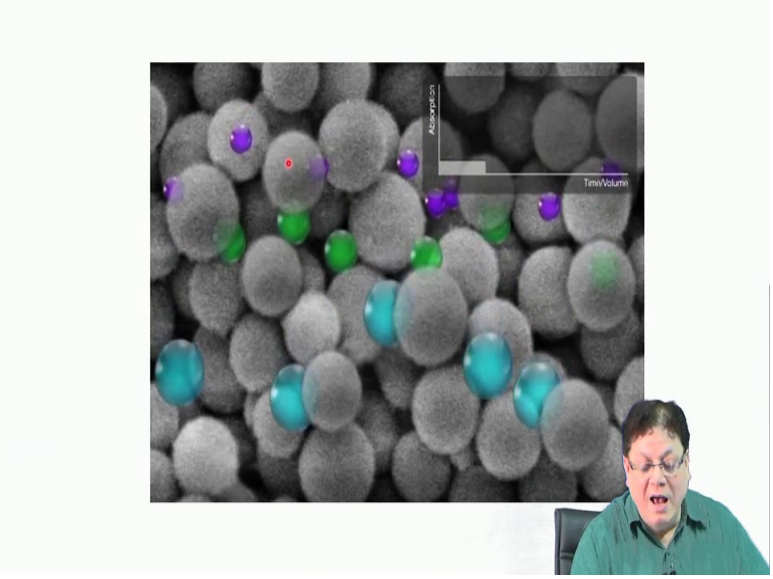
And this was demonstrated by looking in these slides. These are the beads and there are many hollow positions in the beads. And then when you apply the protein solution, which is in this case is indicated by these colored balls, the green ones and the violet ones. The violet ones have a small size. And then there were some balls which were blue in color; the complete picture is not there, but the blue balls are the biggest ones of this protein mixture.

(Refer Slide Time: 05:12)



So, as it percolates through the gel, you will see that already there is a difference between them; the blue (big sized) ones are already ahead of the green(medium sized) and violet(small sized) ones .

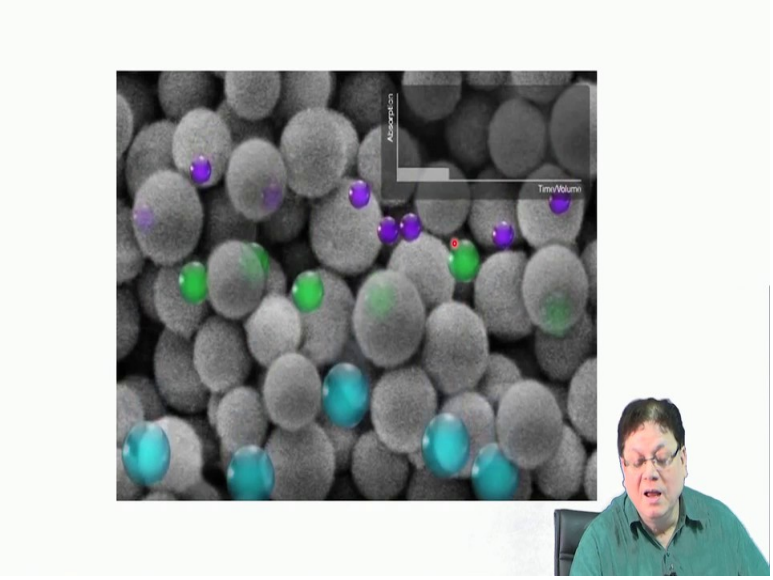
(Refer Slide Time: 05:26)



As you allow time more time, the pattern will become more prominent. Ultimately this is nothing, but a gravity filtration. Here the gravity filtration is differentiating between the large

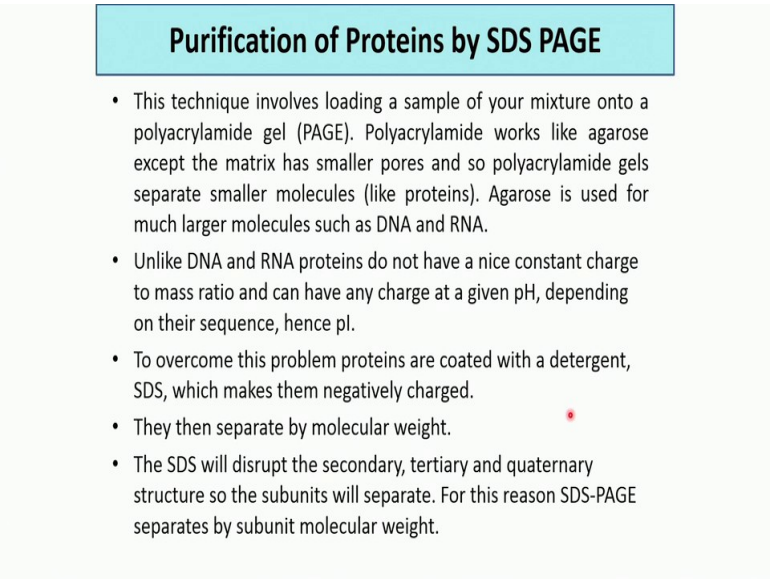
and small molecules. And that differentiation is coming because the small proteins are interacting with the pores that are available in these gels.

(Refer Slide Time: 05:51)



So ultimately, what will happen? The blue ones will elute first followed by the green ones and then the violet ones.

(Refer Slide Time: 05:58)

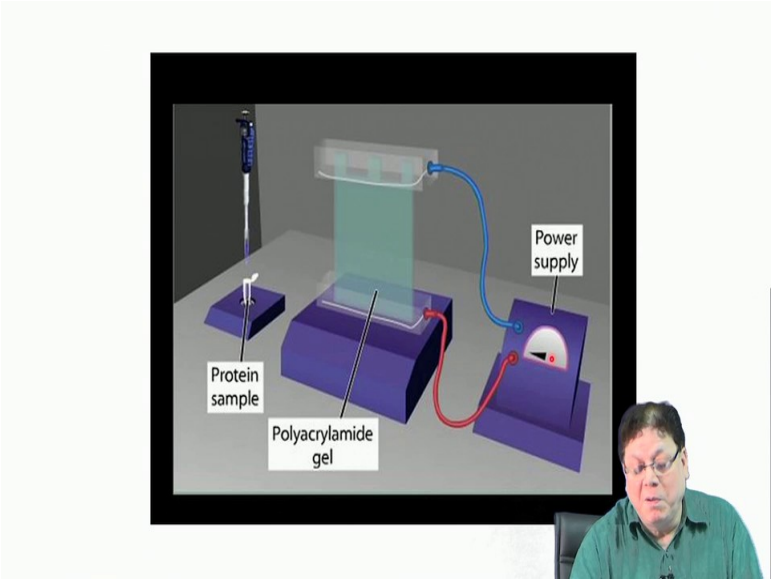


So, that was the basis of size exclusion chromatography which was just dependent on the molecular weight and the size. Now the next one that we discussed was SDS PAGE. SDS full name is sodium dodecyl sulfate; it is a long chain sulfonate. So, it is a negatively charged anion; and this PAGE means polyacrylamide gel electrophoresis. So, by this technique, one can separate the proteins. Here the basis is that, the proteins in the native state remains folded; and they have a native charge and the charge depends on the types of side chains that are present in the proteins.

So, first thing that you want to do is to unfold the protein and make them linearized; this is needed otherwise a small protein may have higher charge and a large protein may have lower charge. So, it would be non-uniform and you cannot really have a process, where the small ones come first and the bigger ones come later or the bigger ones coming first and the smaller ones coming later; as was in the case of size exclusion chromatography.

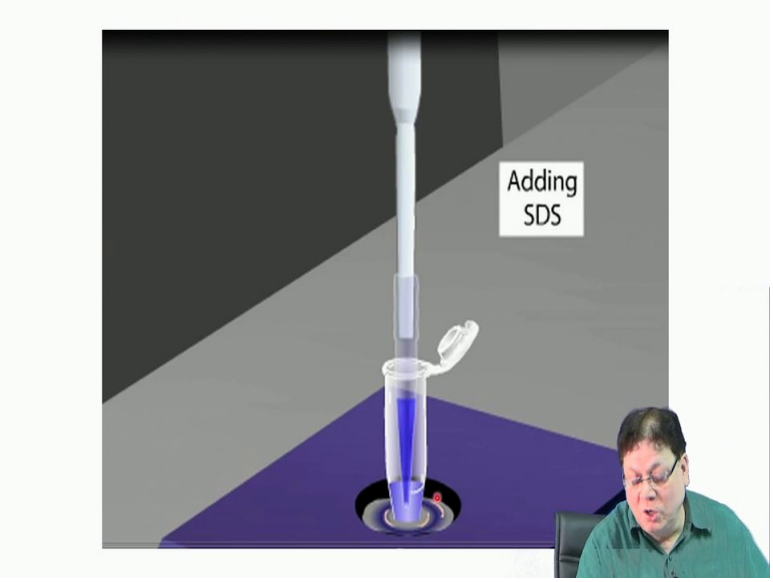
So, first you allow the protein to unfold completely; such that now the protein molecule is having a linear shape. And now coat the surface of the proteins with SDS molecules which are highly negatively charged. So, the charge density per unit area of the surface of the protein remains the same.

(Refer Slide Time: 07:50)



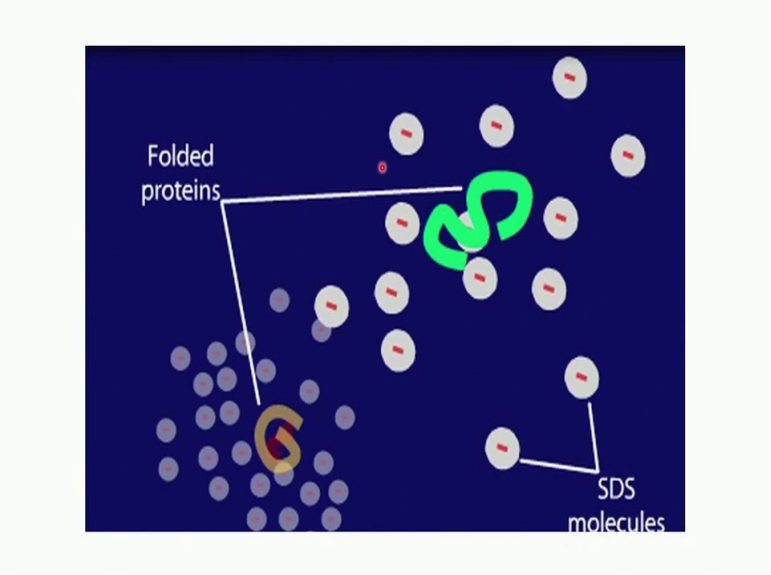
This is the apparatus to apply the current; this is the power supply which supplies the current; and this is where you make the sample, sample means you mix the protein with SDS. The figure shows mixing.

(Refer Slide Time: 08:08)



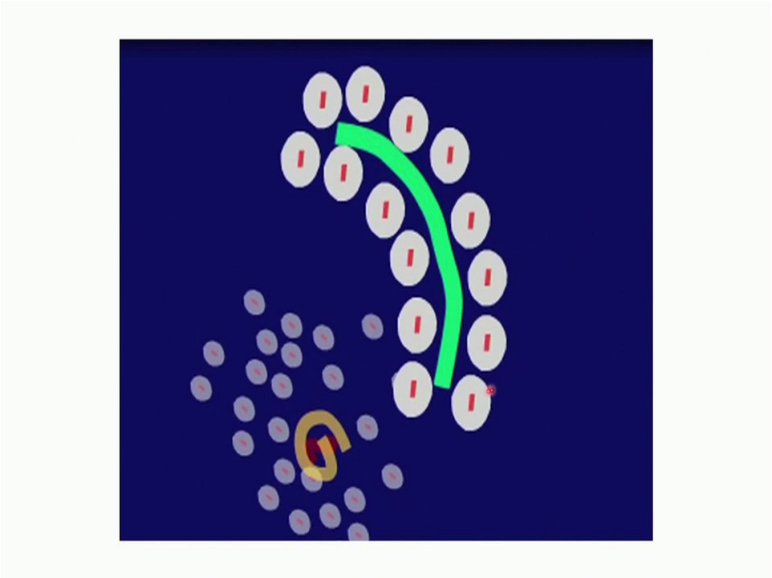
And then you have to also heat it to ensure that the protein has completely unfolded.

(Refer Slide Time: 08:13)



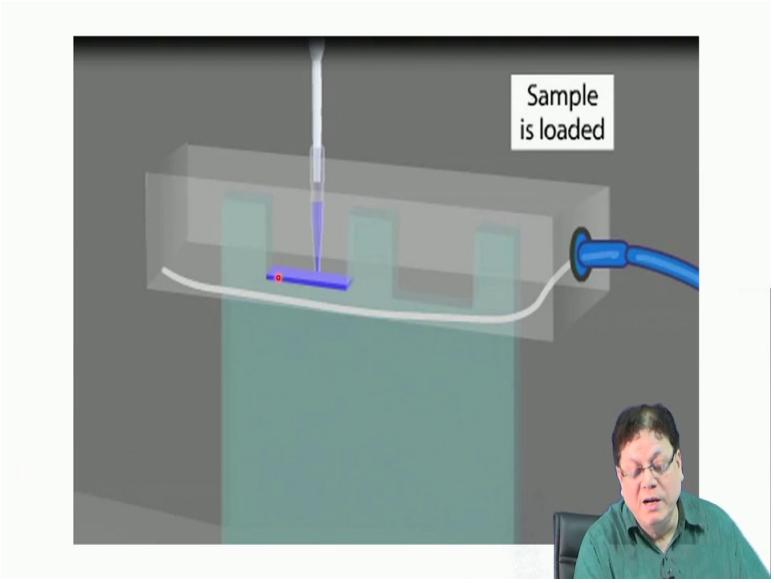
And then as you as you heat it, in presence of SDS,

(Refer Slide Time: 08:25)



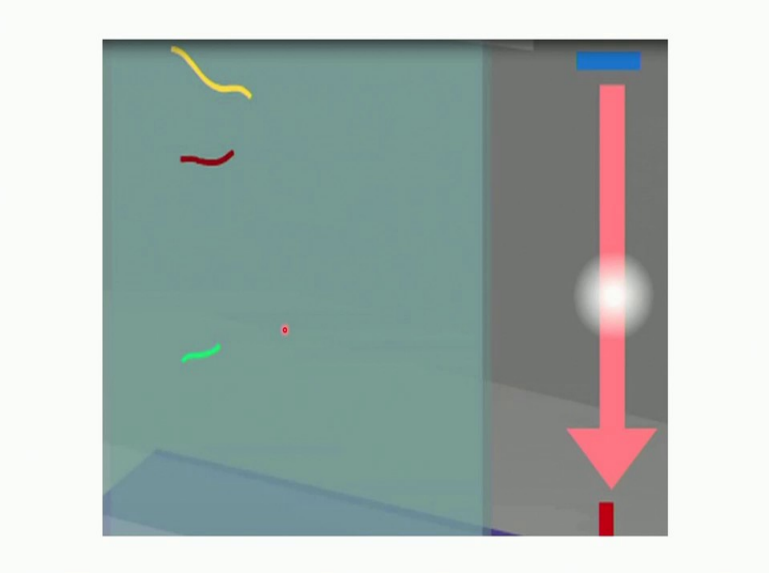
it will become linearized and it will be completely surrounded by these SDS anion molecules. The same thing will happen with this yellow one. So, whatever proteins are there, they will be unfolded and they will be just covered by the negative charges. So, now, if you apply the voltage, what will happen? The smaller ones, because of their small molecular weight, will come out first, their mobility will be high and the larger ones will come later.

(Refer Slide Time: 08:57)



So, you are now applying the solution of the protein which is denatured with SDS.

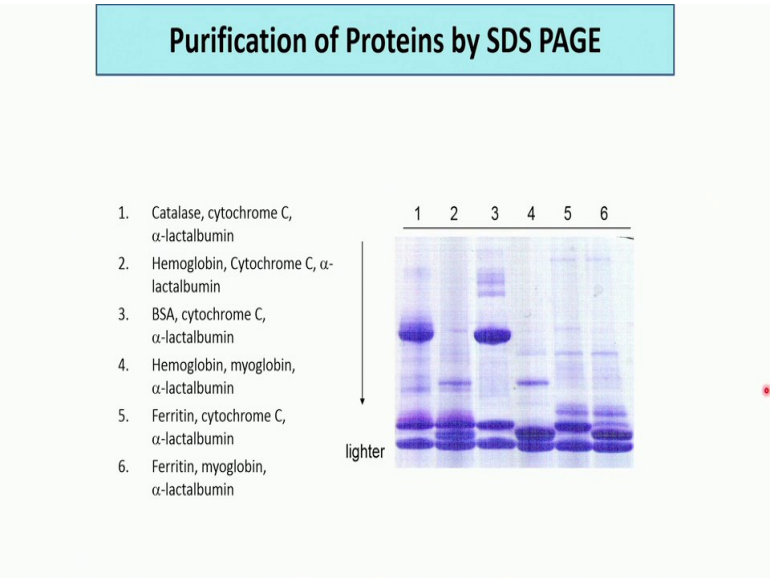
(Refer Slide Time: 09:05)



And then as you apply the voltage, current flows. And there you see the bigger ones fall behind, then the medium ones move little faster and the smaller one are the fastest. And usually what happens? You put a colored dye, which is known to move the fastest; that is just an indicator that how much time you will allow for the electrophoresis.

So, when you add the protein solution and you add a dye; then the dye moves the fastest. So, when the dye comes here at the top, then you stop the electrophoresis by stopping the current; and then you take this gel out and you dip it in a solution of another dye which is called coomassie blue.

(Refer Slide Time: 09:47)



You first soak it in coomassie blue solution and then you put it in acetic acid methanol solution. Where there is no protein that is not bound to the gel, that will be just washed off; and where ever there are proteins, they will appear as dark blue spots or bluish violet type of spots that you get.

The number of spots will you tell you the minimum number of proteins that you have in your solution. I am saying minimum number because this spot looks like as if the number of spot is one, but you never know that how many protein are there; there may be several proteins which will have similar molecular weight and they may be coming as one spot. So, if you want to know whether other proteins are here, then this SDS PAGE cannot distinguish between the proteins with similar molecular weight; because SDS PAGE is under denaturing condition.

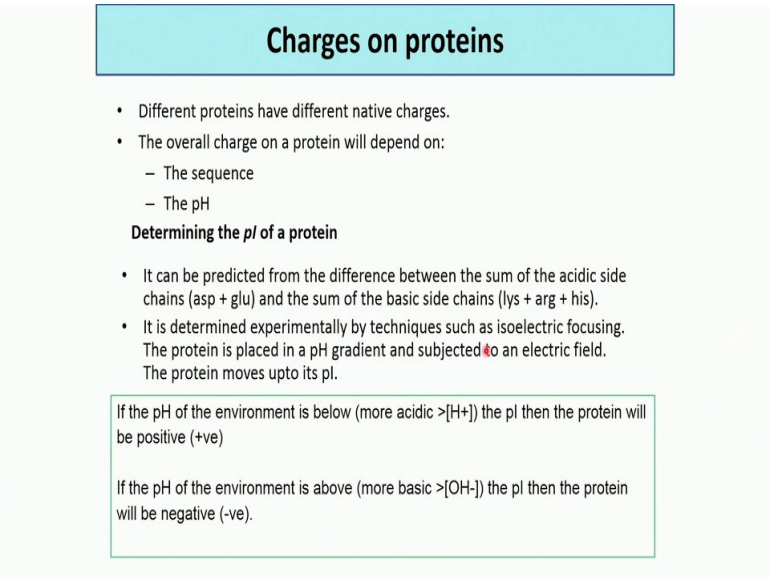
So, if you want to see whether there is more than 1 protein; then you have to do another type of electrophoresis; that is based on the native charge of the protein; in the native form and that will differentiate between different proteins with similar molecular weight.

So, you need to do native electrophoresis. That means, you drop the SDS and do only PAGE (polyacrylamide gel electrophoresis); that is called the native gel electrophoresis; it involves electrophoresis under denatured condition. And then you can tell whether this spot actually represents one protein or multiple proteins that are present here. You have to apply different

techniques in order to really know the purity of a protein. Just appearance of a single band in the coomassie blue, does not tell the number of proteins that are present.

There are other techniques also, but you we are not going into that; the best is to do the native PAGE and the also the SDS PAGE; if you are not very sure how many proteins are there here.

(Refer Slide Time: 12:00)



Native charge is basically the charge that is present in the protein itself; and these proteins will move according to their charge. This charge depends on the pH of the medium.

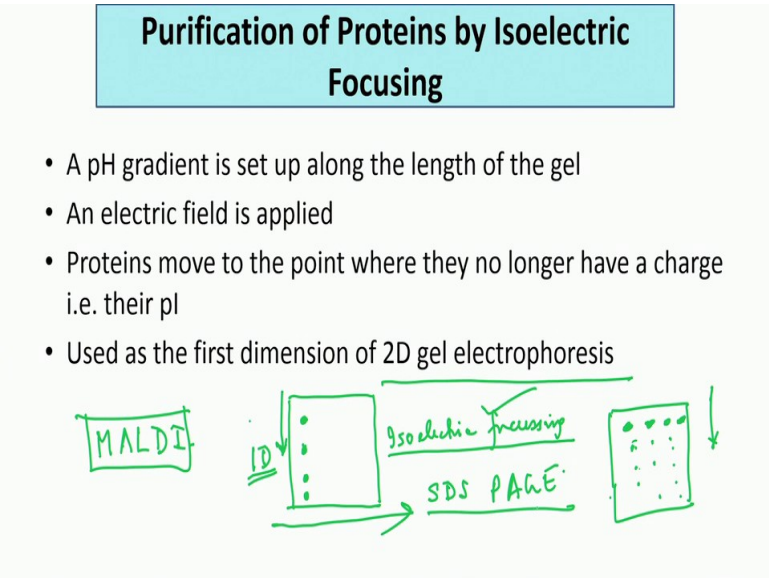
So, the pH is also needed to be adjusted. By adjusting the pH, you can vary the native charge on the proteins and that can also help you to separate the proteins by native gel electrophoresis.

There is a pH where amino acids become neutral; that means, there is no charge on the amino acid. The same thing happens with proteins. There is a pH which is called the isoelectric point where the net charge on the protein molecule is 0. So, the protein will not move at that pH.

Now, the question is how to determine the isoelectric point of a protein? It will be very difficult to really calculate the pI of a protein. You have so many pK_a values now in the protein. So, it will be very difficult to really pinpoint what is isoelectric point. Now there is a

technique which is called isoelectric focusing, that is also a gel electrophoresis method, and it is very interesting.

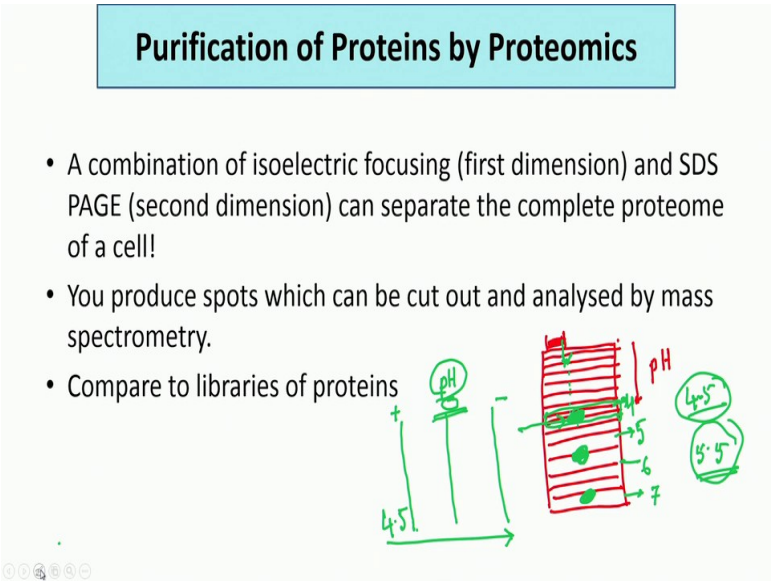
(Refer Slide Time: 13:48)



We will now discuss isoelectric focusing; Let us first finish up the electrophoresis part.

First one was SDS PAGE; again I just revisit SDS PAGE; that is under denatured condition. The second one was the native gel electrophoresis; that is whatever the intrinsic charge that is present in the protein, but I told that the charge can vary according to the different pHs. Now the question is how to know the isoelectric point of a protein? Because, if you know the isoelectric point then you can adjust the pH to move the proteins to different directions and that will allow you to separate the proteins.

(Refer Slide Time: 14:44)



Now, this isoelectric electric focusing is a technique that allows you to know what is the isoelectric point; that means the pH where the protein does not have any charge.

I can show here what is basically done. You are running a gel electrophoresis, but the gel is little different; here the gel has a pH gradient; that means, the pH gradually increases from say here to there, there are ways to do that. So, the pH of the gel slowly increases from say top to bottom. And now you apply your protein solution and they will move under application of voltage across the gel.

Usually the proteins start from a very low pH (highly acidic pH); it is very unlikely that the protein will have an isoelectric point where the pH is 1. So, suppose you start with pH 1. In this direction the pH is increasing. If the isoelectric point of a protein is here; that means, a pH lesser than the pI is maintained in this left zone; and a pH greater than pI is maintained on the right zone. So the protein will be positively charged in the left zone and negatively charged in the right zone; and it will be neutral where pH = pI.Suppose here the pH is 4; and suppose here the pH is 5; suppose this is 6. Now what will happen? Suppose there are 2 proteins; one is having a pI of 4.5; and the other protein has a pI of 5.5.

When the proteins move; the protein which has got a pI of 4.5, as soon as it reaches the region where pH is 4.5 (somewhere here), it stops there; because now it will be electrically

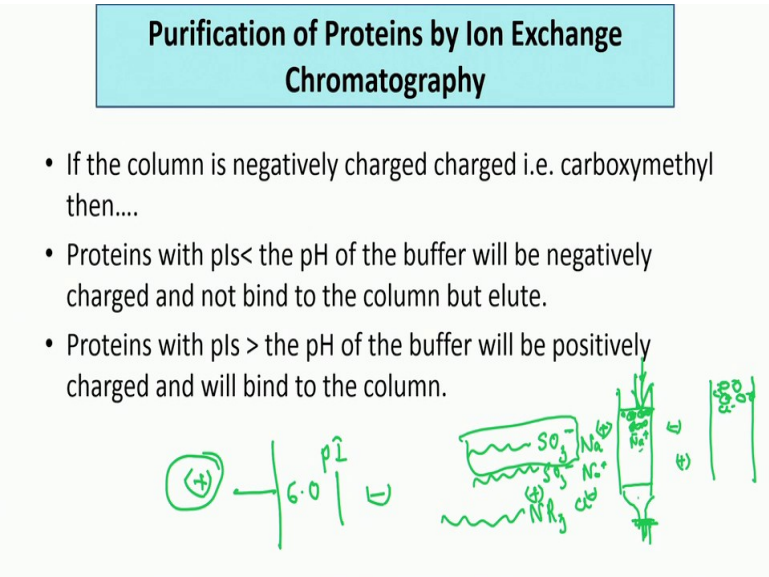
neutral here at this pH. What about the other protein which has a pI of 5.5? That will now come up to 5.5 where it will have charge 0; at that point it will be neutral.

So, basically now depending on different pIs, we can separate the proteins. This is what is called isoelectric focusing. So, you are basically collecting all the proteins at a particular pH, which is equal to its isoelectric point.

So, that is why this is called isoelectric focusing; focusing means you are pin-pointing towards a particular region. The protein which has got an isoelectric point of 4.5, will stop there where the pH is 4.5; if there is another protein which has got a pI of 7, that will stop where the pH is 7. So, this is how you determine the pI of proteins. If you know the pH of this region (suppose here it is 4.5), you can say that the pI (isoelectric point) of the protein is 4.5.

So, basically where ever protein stops, the pH of that zone is the isoelectric point of the system. So, by this way, you determine the isoelectric point.

(Refer Slide Time: 19:42)



So, now, we have discussed the SDS PAGE, the gel filtration, isoelectric focusing and the native gel electrophoresis. We will discuss one more separation technique and that is called the ion exchange chromatography.

In ion exchange chromatography, basically you have a glass column; it looks like this and there is a stopper here and you fill that with polymeric beads. These polymeric beads are either negatively charged or could be positively charged or could be neutral (both positive and negative). One cannot have only negative charges because the whole thing is electrically neutral.

So, if you have some polymer; these polymers end up with SO_3 minus and Na plus. So that means, these beads are mainly made up of this aliphatic sulphonic acid and these sulphonic acids have a surface which is negatively charged and then there is a sodium ion which is attached to these beads.

So, sodium ions are attached to the beads which are negatively charged. So, that is why these are called the ion exchange resins. There could be again another type of polymer, where there could be NR_3 plus and say Cl minus. Now what will happen? It will be just the reverse. You have a column, where there are beads and these beads are now all positively charged and this is surrounded by all the negatively charged chloride ions. Now what will happen? When you add the protein solution here, say at an acidic pH, then proteins are expected to be positively charged. Because, I said that at an acidic pH, the proteins will generally have net positive charges. The isoelectric points of proteins are usually greater than 4 or 5; it ranges from 4 to 10; even 10 is a little unusual.

The pI depends on the number of arginines, number of lysines; whatever the number of basic residues and acidic residues that are present in the protein. Considering all these, usually the range varies from 3 to 8. So, if you take a protein solution which has got a pH of 2;, it is expected that all the protein molecules will be positively charged. Suppose you add that solution to the column. So, what will happen now? Because, the beads are negatively charged which are surrounded by the sodium ions; so, now the sodium ions will be released from the beads and the proteins will get stuck to these beads replacing the sodium ions. So, now the positive-negative attractive charge interactions will take place between the protein and the beads.

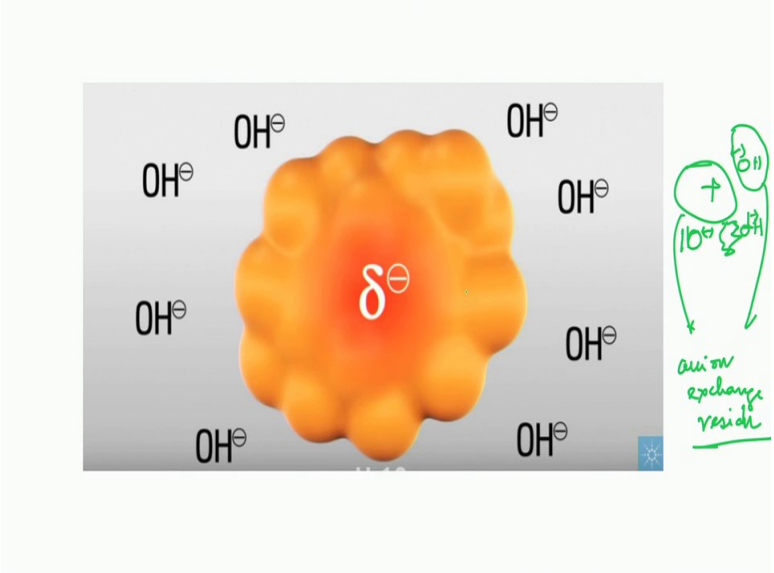
The extent of charge on a protein depends on the pH of the medium. When the pH is equal to the isoelectric point of the protein, then the charge is 0; but before that, it is positively charged. However, the extent of positive charge will be different at different pHs. That means, suppose the pI of a protein is say 6.0. Below 6.0, we know that it is positively charged

and above 6.0, it is negatively charged. But the extent of positive charge will also depend on what pH you are maintaining; lower the pH, the more positively charged species we will get.

So depending on the pH, the charge densities of the proteins will vary. So the interaction of proteins with the solid matrix (polymeric beads or what are called resins) will be accordingly different. As already discussed that initially the sodium ions go out; and the proteins get attached to the resin beads. Then you gradually increase the pH of that solvent (eluent) that you want to elute your column with.

Different proteins will have different charge densities. So, their interactions with the beads will be different. So some proteins will be stripped off from these solid resins and they will come out faster than the others. I think the next few slides will clear the whole thing that I have said.

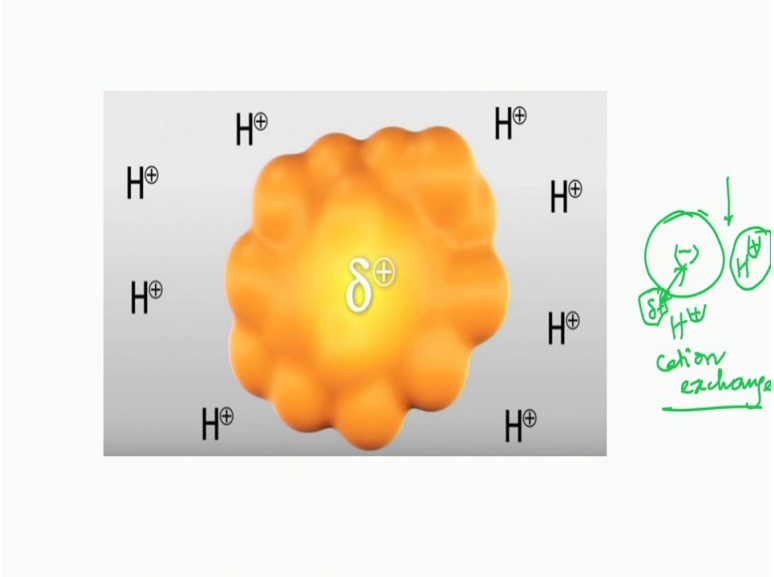
(Refer Slide Time: 25:10)



This is a protein which is having an overall negative charge; that means, this protein is kept at a pH which is above the isoelectric point of the protein. So, if it is negatively charged, it should now compete with the hydroxide ions (or other anions which were previously bound to the positively charged resin) when it binds to the bead. Earlier it was positively charged resin beads surrounded by all these hydroxide anions.

Now, when you add the protein which is also negatively charged; that means, you are having a pH which is higher than the pI of the protein, then it will start displacing the hydroxide and the protein molecule is going to attach to this bead instead.

(Refer Slide Time: 26:18)

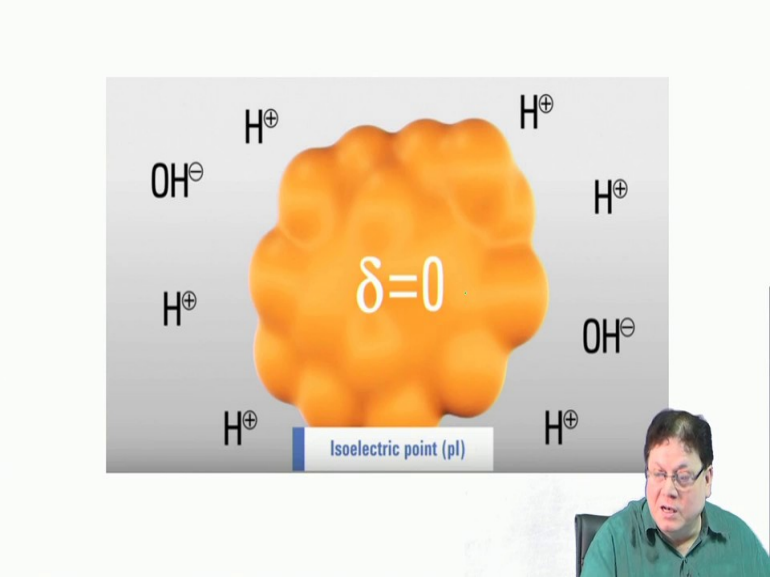


On the other hand, if you go to the pH which is lower than the isoelectric point, then the protein will compete with the H plus or it will compete with the other counter cation like sodium plus or potassium. The protein will displace the H plus because now the bead that you are going to use is negatively charged and earlier this was surrounded by H plus and when you add this protein, the H plus will be knocked out and you will get the protein which is now δ positively charged and hence that will be attached to the negatively charged bead.

As you increase the pH of the eluent; then what will happen? The proteins where the charge density is less they will start coming out faster than the proteins where the extent of positive charge is more. So, if that happens then you have a separation of the proteins according to their extent of positive charge at different pHs. This exchange process, when your bead is negatively charged and your protein is exchanged in place of the hydrogen ion is known as cation exchange chromatography. The hydrogen ions are displaced and the bead is negatively charged. Thus it is a cation exchanger; the cations which were earlier H plus are now replaced by the cationic proteins. And the earlier description involving positively charged resin will be the anion exchanger; that was anion exchange chromatography and the resin will be called anion exchange resin.

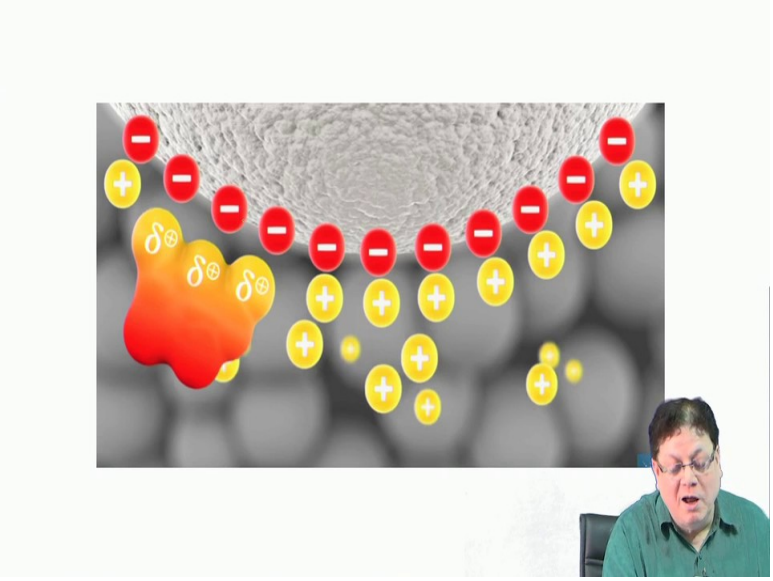
Here the exchange is taking place between the hydroxide and the negatively charged protein.

(Refer Slide Time: 28:16)



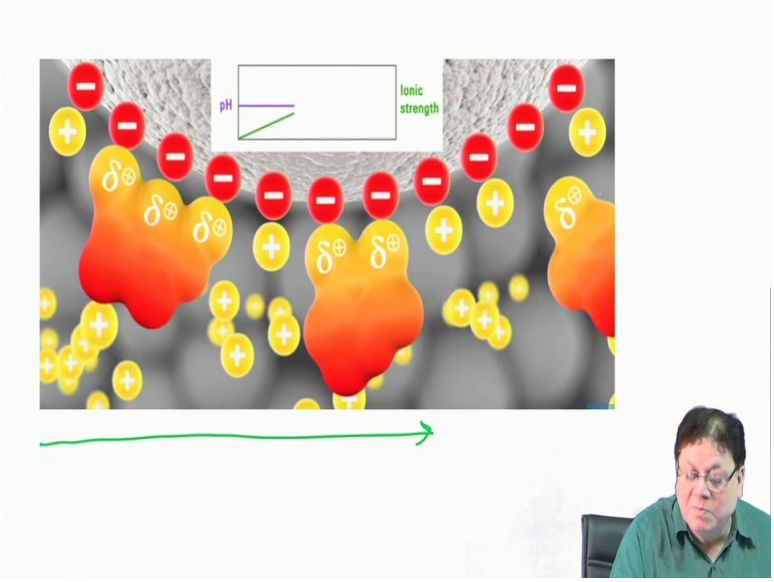
Now, I will show you some slides that how it works. At the isoelectric point, the protein basically does not compete with anything; it will be detached from the bead because the overall charge is 0.

(Refer Slide Time: 28:28)



So, this is suppose the whole bead; this is anionic as the bead has a negative charge. This is cation exchanger because, this is surrounded by all these cations to start with. Then, you add proteins, which are positively charged. The protein solution was in acidic pH; so that it is positively charged. Hence the protein will be bound to this resin displacing this positively charged counter ions.

(Refer Slide Time: 29:06)



And then, what will happen as you increase the pH? The movement of the proteins is in this direction. So, now, what will happen? This has got 2 δ+. So, as the solvent flows through it, the one which has the lowest positive charge (δ+), that will first move out because the interaction with the resin will be minimum there.

So, you raise the pH, it goes out followed by the one, which has got 2δ+ (two positive charge means here it is showing that it has got more positive charges);. Finally, the last one to elute is the one which has got the highest positive charges (3δ+). There is a relation between the pI and the extent of positive charge. That means, how much you are below the isoelectric point.

This is actually very close to the isoelectric point of this protein so that will come first; then, as you increase the pH, the pI of another protein is reached; then that will come out. As we increase the pH further, you reach the pI of the third protein and then it will detach from here and it will go out of the column. So, that is the principle of ion exchange chromatography. So, now, we have all these different techniques. I will just tell you little bit about 2

dimensional gel electrophoresis. Earlier what used to happen was that the biologist were trying to identify the proteins individually one by one. If I ask you what is the number of proteins in our body? It is very difficult to answer. But after the human genome project, when we came to know the entire genome, now we can say that about 30,000 proteins are present in the human body. But if you have an unknown organism and there are thousands and thousands of proteins; and if you want to identify the proteins one by one, then it will take hundreds of years to really identify and characterize all the proteins. So, what is done is called the proteomics study. Proteomics means the study of all the proteins which are together; now you are not identifying one by one. So, what you are doing? You do a 2 dimensional gel electrophoresis.

First, in 1 dimension you do the isoelectric focusing. By this method, the proteins are separated by their isoelectric point. Since you have done this isoelectric focusing; that means, you have a pH gradient here. Suppose there are these spots that you get after the isoelectric focusing. I earlier told you that each spot maybe again containing hundreds of protein. So, then basically what you are thinking to be one protein, may not be actually correct; there should be other methods to resolve (separate) these proteins.

Next what you do? You carry out in-plane rotation by 90° and now all the points or all the bands are here;I am taking just 4 bands. And now what you do? You do an SDS PAGE. According to molecular weight, the proteins will be separated. SDS PAGE is only dependent upon the molecular weight; the length of the protein. So, if you do that, now these are the different spots that you will get.

There are methods by which you can take this gel out from here and directly put it into the mass spectrometer; there is a special technique called MALDI (Matrix Assisted Laser Desorption Ionization) spectroscopy. So, if you now do mass spectrometric analysis, you can know what is the structure of this protein. So, you take the band from here, you know the structure of the protein that was here at this point.

Thus you are doing 2 electrophoresis; in two different directions, two directions mean if you do not make it upside down then basically you are applying the field in this direction so that the first dimension is used for the isoelectric focusing and the second dimension is used for the SDS PAGE. By this technique, you can separate all the proteins in a gel; and then you can pick up individual spots; go to the mass spec. and do the mass spec. analysis.

Ultimately it is actually pretty complicated though it appears to be simple in technique. During the ultimate analysis, when you have huge mass spectral data, then it is the job of the bio informatics people to know the structure of the protein. So, that is the present day technology which they do. This process of identifying all the proteins together at the same time is called the proteomics.

Thank you.