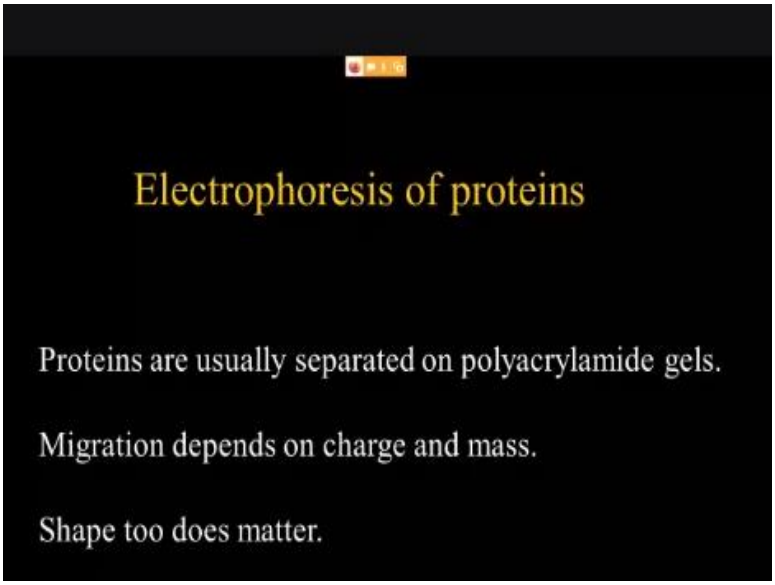
Introduction to Biomolecules
Prof. K. Subramaniam
Department of Biotechnology
Indian Institute of Technology - Madras

Lecture – 6
Electrophoresis of Proteins and Protein Sequencing

(Refer Slide Time: 00:14)



So yesterday we learned about chromatographic methods where you know we looked at 3 different kinds of chromatographic methods; ion exchange chromatography, then gel filtration, then affinity chromatography where we saw how proteins from a mixture can be separated and individual proteins can be purified and then we also learnt about the purification scheme.

How to evaluate and monitor the efficiency of purification looking at the total protein content and the total activity and then how much of the activity we are able to recover at the end. You know that is where we learn to the term specific activity. So today what we are going to do is we are going to look at another method which is primarily an analytical tool okay. So, the chromatography is sort of a preparative tool where you purify proteins of certain quantity for further use.

Electrophoresis also can be used for preparatory purpose where you can purify proteins, but that is really very small scale and its primary goal is actually for analytical purposes. So, we will learn about how electrophoresis is used for that purpose in proteins. So I am sure most of

you already know what is electrophoresis. Electrophoresis is basically migration or movement of charged molecules in an electric field, so that is what is electrophoresis.

So in biological applications what we actually do is we use a solid matrix for separating DNA or RNA we use agarose-based solid matrix and for proteins it is primarily polyacrylamide gels. So the choice of the matrix depends on the kind of molecules we are separating on the size of the molecules. For larger molecules. polyacrylamide gel is not very suitable because the pore size here will be very small.

And so therefore DNA are not usually separated using polyacrylamide, although it is very useful for separating oligonucleotides and for sequencing DNA, so where this sequencing is done by synthesis and those synthesized nucleic acids are short oligonucleotides, but in the case of proteins polyacrylamide gels are the dominant kind of solid matrix we choose. This is the primary way by which electrophoretic separation of proteins are done.

So polyacrylamide is a gel, so it is made up of a monomer called acrylic acid and bisacrylamide and these two form a network like matrix polymer, bisacrylamide acts as the crosslinking molecule that cross-links long polymers of acrylic acid okay, so that is the polyethylene gel. So usually the gel is made from solution of acrylamide and bisacrylamide mixed solution and to that we add molecules that catalyze the polymerization and initiate the polymerization.

So by adding those components, then this solution turns into a gel, once it polymerizes it becomes gel. So in the case of agarose its long chain polysaccharide it is obtained from seaweeds and there it is a powder form, but when you soak it in water and hydrate it usually by heating the agarose gets hydrated and swells into a gel like constituency, so that is how that is made.

And here the primary property that separates the mixture of proteins into individual proteins is by variations in the migration which depends on charge and the mass. So what is the net charge on a given protein and how it compares with the rest of the proteins in that mixture as well as its molecular mass compared to the rest of the proteins and these two properties determine the rate of migration through the gel upon in application of electric charge.

Shape also matters because the pore size is going to have a certain diameter and if you have a long fiber kind of thing with a low diameter then that can migrate while a shorter protein global coating with a bigger diameter is not going to migrate. So shape does matter, but we do a modification to the protein shape such that eventually in the most commonly used version of polyacrylamide gel what matters is the charge by mass ratio that is for unit mass how much charge is present.

So that is what is going to determine and therefore it is actually the molecular weight, the molecular size of the protein is the sole property based on which proteins are going be separated in this a certain variation of the polyacrylamide gel. So that is what we are going to learn now.

(Refer Slide Time: 06:43)

So let me take you to the pictures that makes it easier to learn. Normally for agarose gel which is the easiest to make what we do is we melt, essentially you measure out a certain amount of agarose powder, you add it into a buffer which is primarily water but you need charged molecules to conduct electric current and therefore you have a buffer there like Tris-acetate-EDTA, so where the acetate ion is going to be the medium for conducting electricity.

So you add the agarose into that, then you put it in the microwave and heat it and then it melts and it dissolves into the liquid, so it becomes clear liquid. Essentially the long polysaccharides are hydrated in this heating process in with buffer. The buffer is primarily water. So now when you allow it to cool, then it slowly becomes a gel, it solidifies and becomes a uniform gel. So what we do is after melting?

We pour on a acrylic sheet based tray and in the tray its solidifies and you make gel. So, the shape of the gel depends on the shape of the tray. So usually it is a rectangular boat like structure with the shorter dimension open, then you keep it in a tank where you have the buffer added, then apply electricity. You will have electrodes on either side of this rectangular tray and when you apply electricity.

Then you have the current passing through the gel and if you have added a mixture of DNA, t gets separated. So that is the simplest one. So this acrylamide bisacrylamide polymerization the way we do is shown in this cartoon on the left. So here you keep two thin glass sheets like this is basically like the window glass. So you have a glass where the top is notched like this, the bottom is straight flat which is not visible which is sitting in this trough.

And you have one more glass sheet that is shown here and in between space so it is basically this the gap between the two glass plates is determined by a plastic sheet that we keep between the two and that is usually 1 millimeter to 1.5 millimeter. So the space between the two glass plates is about that much, maximum it usually use 1.5 millimeters and in some preparatory version they might use 3 millimeters.

What we commonly use for separation to get a gel like what is shown in the right, you use about 1.5 mm spacers. These plastic strips shown here they are the spacers, so they are between the two glass sheets and in between space is where of course there is a space I put at the bottom. So, whatever you are pouring in the space between the two glass plates that does not leak out and once polymerization is over, the bottom one you will remove it, so therefore you can have a tank here with the buffer and also top a tank with the buffer.

So then if you have electrode on the top and bottom as shown here, then electric current will flow through because the buffer in which this acrylamide is polymerized as well as the buffer on the top and bottom they have ions to conduct electricity. Now if the proteins are charged, then let us assume that they are all negatively charged, then they will migrate towards the positive charge okay. So this is how it is done.

So here essentially instead of melting something in buffer and pouring in a tray and allowing to solidify, here actually it is a chemical reaction that makes a polymer from monomer. So,

the monomer used is the acrylamide and it is not simply a long fiber, it gets cross linked by the presence of bisacrylamide that can covalently link on either side to two growing chains of acrylamide polymer and this cross-linked mesh like structure is what finally is the gel.

And mixing these two does not activate the polymerization spontaneously, you need an initiator and you need a catalyst and this is served by 2 compounds, one is called ammonium persulfate and another one is called a TEMED. TEMED is an abbreviation of a long name, so you can look up in the book later. So once you add these two to the solution of acrylamide and bisacrylamide, you mix it and before it solidifies you cast the gel.

Meaning you pour that liquid between the two glass plates and you will keep a comb like structure here, a plastic comb with tooth and after it solidifies if you remove the comb you get these shapes which we call as the well and that is where the protein is loaded. So usually, the protein solution is mixed with the glycerol that increases the density therefore it does not diffuse into the buffer, instead it goes and settles at the bottom of these wells and to visualize the solution as well as its migration you add a colored dye.

So as it migrates you can see that. So once the separation is done, then you open up and take it out and soak it in a solution that actually what we call as fixing. So, it fixes the proteins. The proteins no longer diffuse after separation and they remain in position on the gel. Then you add another solution where usually you have a colored pigment or a dye that is specifically bind to the protein and that process is called staining.

First you fix the protein to the gel so that it does not diffuse while you are staining because it is all in liquid okay. So then protein can readily diffuse if it is soluble in the aqueous buffer. So the fixed separated proteins when you stain with the dye, then the dye binds to all the protein, so that process is called staining. Then you wash with a solution that removes the non-specifically bounded dye non-specifically bound all over the gel.

And only the dye molecules bound to the protein stay there that helps you to visualize the proteins and such a gel is what you are seeing on the right here. These are called the molecular weight markers or molecular weight ladder because it appears like a ladder. So these are our reference proteins whose molecular weight we already know and some of them

are covalently linked to a different colored dye, like for example this appears brown because it is already pre-stained marker.

So to visualize these pre-stained markers, you actually do not need to do the staining and washing, de-staining process. So they are visible even while migration happens okay. So here what I have done is I have laced bacteria, so the E coli bacteria that we normally grow in the lab and separated on the gel and each protein appears as a band and so this is stages of purification.

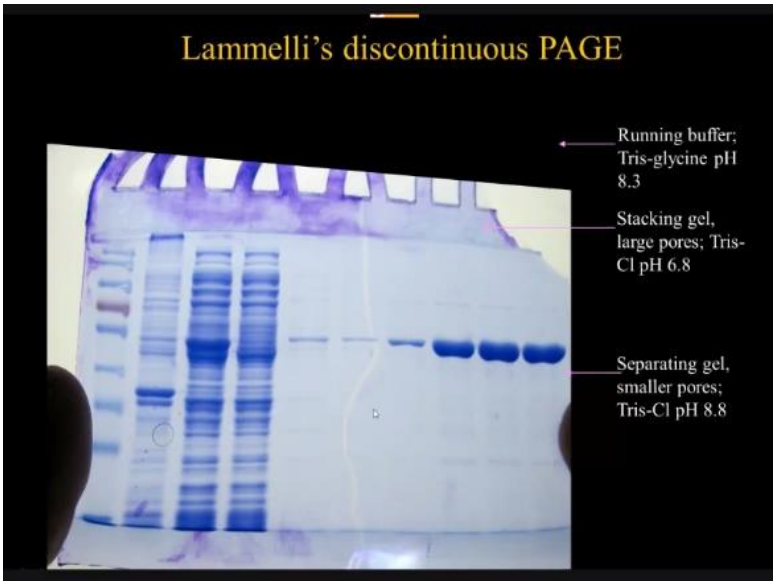
So we see here this is an E coli stain that is genetically engineered to produce a certain protein and that is this protein here, you know this is present here, not in this. So this is normal E coli that does not produce that protein and this one is producing that protein and I have used an affinity chromatography to purify this one out and you see here this pure protein is visible while most of the others are removed.

And in this experiment what I have done is I have checked whether my cloned protein genetically engineered one is it being produced and if it is being produced am I able to purify it out and in which fraction the pure protein has come and you can see here 1, 2, 3, fourth, fifth, sixth fraction contains my protein in binding to the column washing and diluting that I described yesterday with respect to affinity chromatography.

So this all looks cool and nice, but I told you in the previous slide we primarily separate the protein based on the molecular weight alone. So given that see here you have the real well, you know this is the cartoon and this is the well after I have removed the comb you see these swells in which proteins were loaded. So given that it is charge based mobility in electrophoretic field the charge will matter.

And at the same time shape also will matter because you have pores of certain uniform shape and diameter. So a long rod will have a narrow dimension than a shorter big ball, so therefore shape will matter. So we do a certain modification to this basic simple gel such that these two are no longer an issue. The shape is made uniform for all proteins and similarly charge is also made uniform for all proteins and that modified version is called SDS-PAGE.

(Refer Slide Time: 18:07)



So here I am going to explain that in detail and before I go into SDS, I want to explain one more detail about this basic polyacrylamide gel. So this page stands for polyacrylamide gel electrophoresis okay. So in the most common method that we do where the separation is purely based on charge by mass ratio and when charge is constant therefore it is only the mass that matters.

There what we actually do is we boil the protein in a buffer that contains this ionic detergent uh sodium dodecyl sulfate, SDS. So, this is a strong ionic detergent. What it does is it opens up the three-dimensional structure of the protein and that process we call as denaturation. So denaturation is the process by which the three-dimensional structure is disrupted and all molecules assume same rod-shaped structure.

And that is because this SDS based on its strong charge it binds to the amino acids at uniform ratio, say x number of SDS molecules per y number of amino acids. So the number of SDS bound is directly proportional to the length of the polypeptide chain and the charge provided by the SDS is a lot more than the charge on the protein and due to that the net charge of the protein is negligible and the final charge is actually the negative charge produced by SDS.

So after this treatment, the charge of a protein depends on its size because its size determines how many SDS molecules will bind. So therefore, if you take a ratio of the charge by mass that is going to be constant for all proteins. So, this is the treatment we do so that the shape of the protein and the inherent charge on the protein do not affect the separation and the separation is only based on molecular weight.

So this is one modification and second years ago a scientist by name UK Lammelli came up with a very ingenious way of making a polyacrylamide gel. So in this what he did is basically this gel is, his invention is to circumvent a certain difficulty in resolving the proteins that is due to the fact that we make this narrow sheet of gel. So, the gel is made between two vertically held glass plates that I showed you in the previous slide.

And this is done due to primarily to avoid exposure to air because oxygen in the air interferes with the polymerization of the acrylamide, it gets oxidized to acrylic acid. As a result, we pour this gel between two glass plates and due to that it is a thin sheet of gel and it will load sufficient volume of any liquid, here sufficient volume of what I mean is I am talking in the range of 10 microliters to 50 microliter volume.

So this is going to be a long comb like structure, so you can see here it extends some more, it is usually about a 1 to 2 centimeter depth well. So now when you add liquid, let us say the liquid let us take this one, you know this is a uniform shape. Suppose if the liquid fills this much and when it is homogeneous, the same protein molecule will be present here, here, here, everywhere to up to the top.

So you have a column whose height is about 1 centimeter let us say. Now when the protein solution migrates it will be 1 centimeter column that is migrating. So a given molecular weight protein is present all through this 1 centimeter column. So, every band will be exactly the same height of the original amount of liquid or original column of liquid that was there. So due to that you are not going to get this fine sharp bands like this.

So look at that this well was filled, but this protein is only a sharp band. So if it is going to be wider like the original size, then next one, next one all will overlap and separation will not be good. So this problem or this limitation is primarily because you are making a thin gel between two glass plates and that is because you want to prevent exposure to air and to overcome this problem and increase the resolution what UK Lammelli did was he introduced a modification and that we call as the discontinuous polyacrylamide gel electrophoresis.

So here what is discontinuous is there are two things that are discontinuous. One you see here a line and this line is because I poured originally a certain concentration of polyacrylamide

mixture here and allowed it to form a gel there. It solidified and made a neat gel with the sharp straight flat line here. So do not worry about this, this one is not another separator, this is the migrating front that is the electrophoretic front.

So, the fastest molecule that was here when I applied it has reached here when I stopped the gel. So here no molecules are counter. So this discontinuous is because I poured one gel of a certain concentration. So usually the concentration here is higher, in this particular gel it is 12% gel, meaning 12% of acrylamide-bisacrylamide in basically water. It has some buffer to allow ion for current, but primarily it is water.

And then I poured another gel on top of it after keeping the comb, so in that gel I have the comb shape formed. So this gel is a low percentage gel. The acrylamide here is about 4%. So here the pore size is larger, so the size is definitely not going to matter in this. While the pore size is smaller and here size is going to matter. So there is a discontinuity in the gel pore size itself, so that is one thing.

In addition, here the acrylamide and bisacrylamide I mixed in a buffer which had a pH of 8.8, it is a higher pH and here I have used a lower pH 6.8. So, there is a discontinuity in the proton concentration, hydrogen ion concentration as well and the buffer used for making these two gels is Tris HCl okay, Tris is the base and the HCl is the acid group here. So, we simply call it as Tris Cl, Tris chloride or Tris HCl or whatever and once this is done it is mounted on tank.

There is one tank on the top, one tank at the bottom and the electricity is applied and the tanks are filled with the buffer and that is going to supply the ions for continuous passage of the electric current through this and that buffer is Tris glycine okay. So here the pH is 8.3. So pay attention to all these values, these are all important and these are directly connected to the PKA concept that we learned.

So now when you apply electricity when this Tris glycine you know this is a weak acid, it is not going to totally dissociate while chloride would have completely dissociated and when you apply current the chloride ion will move very fast. So there is a set of molecules that are moving very fast at the bottom of the gel and on the top the partially dissociated glycine enters but it gets right away protonated because the pH is low here.

So it starts to move very slow and due to that a voltage gradient or potential gradient is set up in the gel and that makes all the molecules move fast and line up here at the junction between the two gels. The voltage gradient created by the difference in the pH between this and this and the molecule chosen and its pK a value. So this glycine is pushed down because there is a potential gradient which is actually going to apply a lot of force down.

But it does not move because it does not have a charge and as a result it all get concentrated in this junction and that is how while all the liquid has moved, all the protein molecules have all concentrated as a narrow band here. Now once they slowly enter here, they quickly get ionized because the pH is high. So now glycine gets ionized and now it starts to move steadily and at the end once the all the chloride ions are gone.

And it is all simply 8.3 pH all through because due to the continuous flow these two are going to get disrupted, then you have a steady movement of the protein molecules as sharp bands. So this whole drama is done primarily to concentrate all the molecules in that taller column of liquid into sharp band at this junction. It forms at this junction primarily due to the pH difference and the pore size difference that is why it gets concentrated here.

So this is the basic principle of discontinuous polyacrylamide gel electrophoresis. So this gel can be used without adding SDS as well, but when you add SDS we call it as SDS-PAGE. So I told you the sample mean the mixture of proteins is boiled in SDS and therefore the protein is denatured and then you have a certain number of SDS bound to the protein and the number of SDS molecules bounded depends on the length of the polypeptide chain.

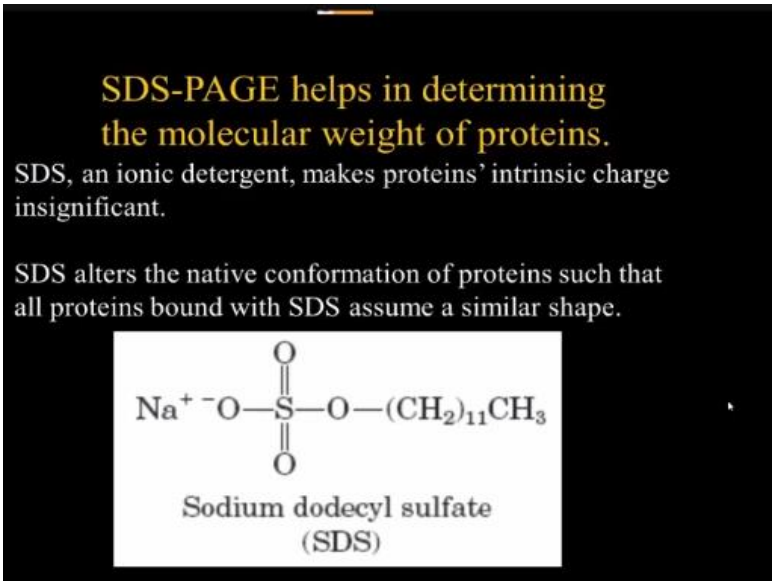
So with these two modifications we get a separation like this okay. So, I hope you have understood this concept quite clearly. This is very important because this will keep reappearing in biochemistry or in doing experiments in multiple contexts. So therefore pay attention to remembering the basic chemistry behind the SDS-PAGE. So if you look at this gel, you probably able to count about 100 bands in this. So it is in the order of hundreds.

So maximum if you increase the length of the gel some more and allow more theoretical plates and increase the percentage and so on and when you increase the resolution probably you can go to 300 to 400 bands and that is a stretch, normally it is under 200 bands that are

identifiable. So any normal eukaryotic cells like cells of our body make about 20,000 different proteins.

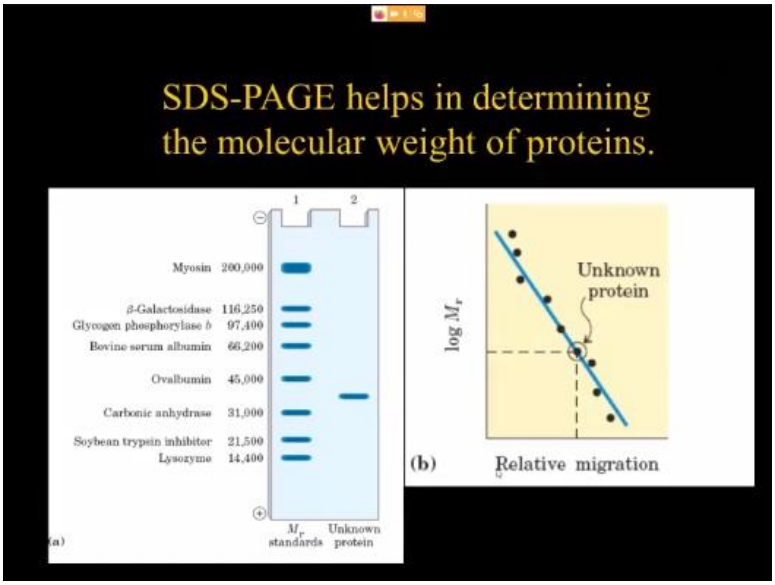
So obviously there will be multiple proteins that are of the same molecular mass and as a result a band that you see here like this may not be a single protein, it may be a mixture of proteins. So you want to separate them as well, so for that we move on to another kind of gel.

(Refer Slide Time: 31:56)



So this is the sodium dodecyl sulfate. This structure is no threat to you because you already know free fatty acids, long aliphatic side chain we know, instead of carboxylic acid group you have a sulfate group. So this is sodium dodecyl sulfate.

(Refer Slide Time: 32:19)



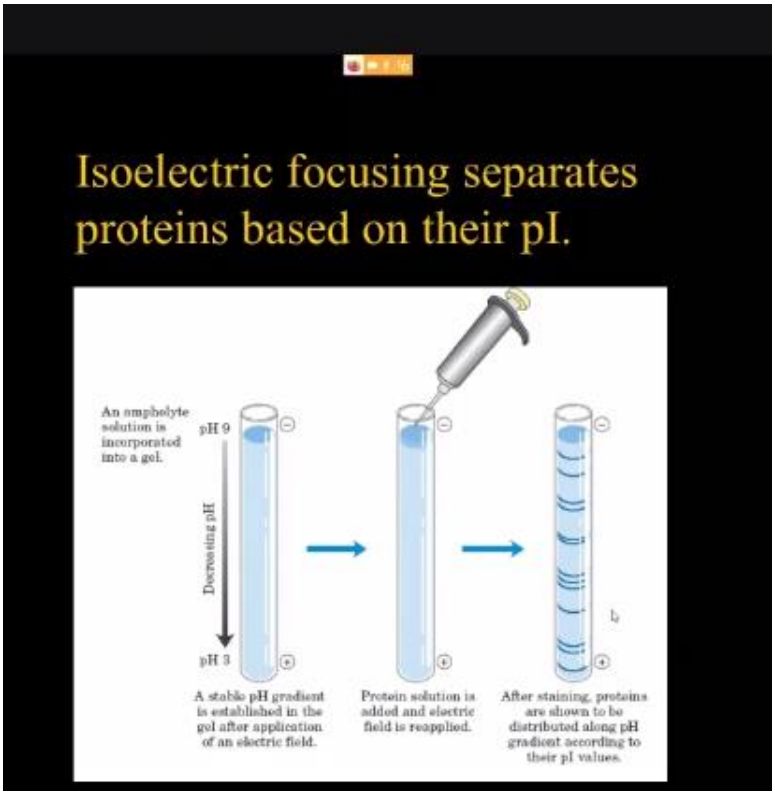
So this is how the separation happens. Another important point that I want to point out is the distance migrated is directly proportional to the log of the, sorry inversely proportional to the log of the mass. So y axis is log and x axis is the distance in centimeters usually. So if you measure and plot if you know the molecular mass of these reference proteins, the molecular weight of an unknown protein can be readily determined.

There are other more sophisticated methods, but his is the most practical and reliable way of determining molecular weight of proteins. So molecular weight may be affected by other, sorry the shape everything will affect certain methods and there are sophisticated methods like ultracentrifugation where you can the other factors like size, shape none of them influence and you accurately determine the molecular weight.

But for most practical purposes this SDS-PAGE is the easiest and practical way of determining molecular weight and usually it's very accurate. So you measure the distance migrated by your reference and plot on a semi log sheet, then you look at the distance your unknown protein has migrated and then from that graph you will get the molecular weight straight.

If this is the distance it has migrated and look at here what is the log of the molecular weights from that you get the molecular weight of the protein. So this one you know readily without even plotting by knowing the sizes you can tell yeah this is somewhere between 30 and 45 probably about 40 kd protein. So, this is an important application of SDS-PAGE.

(Refer Slide Time: 34:23)



So now to circumvent this issue of not being able to separate all proteins, **you** we exploit two distinct properties of proteins. So here we are exploiting the molecular weight differences among the proteins for separation. Then we already know depending on the kind of amino acid sequence a protein has its isoelectric pH can be varying and I told you isoelectric pH is the isoelectric point or pI is the pH at which a protein's net charge is 0.

Which means that protein is not going to move in an electric field if the gel pH at that given position equals its pI. So that is in no way related to its molecular weights, so it is a completely independent distinct property, so that is exploited in separation and such a electrophoresis is called isoelectric focusing. So this isoelectric focusing, focusing means the protein migrates gets focused at the pH on the gel where its charge is 0, meaning that pH equals its pI.

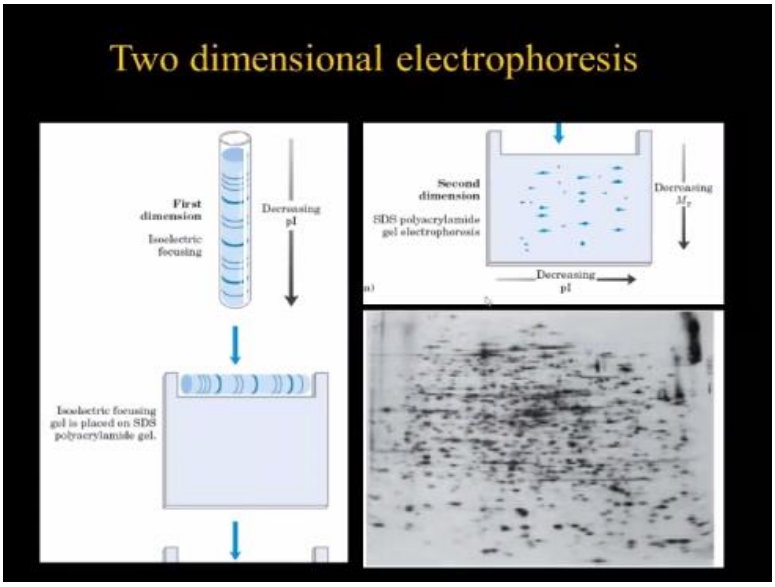
So you are actually focusing the molecules of that particular protein from different parts of the gel to that particular place because up to that it will migrate. Then it is going to stop migrating. So this method separates the proteins based on that pI and this is usually made in a tube gel. So this is called a slab gel because it is a slab-like structure a sheet and this one is made in a tube. So tube, do not imagine a test tube here okay.

So here we are talking about capillary like 200 microliter capillary or something like that where the diameter of the lumen is in terms of millimeters, so one millimeter or less than that that sort of a small tube we are talking. So here the polyacrylamide is mixed with a group of molecules called ampholytes. So these have different kind of charged moieties on them such that when they are distributed they generate a pH gradient.

So for example in this particular cartoon we have a gradient of pH 9 to pH 3. Now when you apply proteins what they do is they migrate, it can migrate in both directions because it depends on the net charge on the protein, you have not saturated it with a molecule like SDS, so the net native charge only is there. So now they migrate on this gel and stop migrating when they reach a pH where that pH equals that pI.

So it creates bands like this along the high pH to low pH. Now using a thin needle inserted between the glass and the gel and gently pushing some water through it, you can extrude this gel out, the tube will come out.

(Refer Slide Time: 38:22)



And that tube what we do is place it on a SDS-PAGE slab gel. So here the comb is like a flat single plastic sheet and therefore it will make a single big comb and in that space you push this tube into it and now you soak it in SDS buffer and this is of course made in SDS buffer. Although the line is not drawn in this, this is discontinuous gel just like the one that I showed you here, the discontinuity is there, exactly it is normal SDS-PAGE that we just learnt elaborately except that the comb is big to fit this tube.

Now you apply current. Now what happens all these bands get concentrated here and they are denatured and they now separate based on the molecular weight and each of this band, you know cylindrical band, here when it gets concentrated it becomes a spot and if that spot has a mixture of proteins of different molecular weights, they get separated and that is what you are seeing here okay.

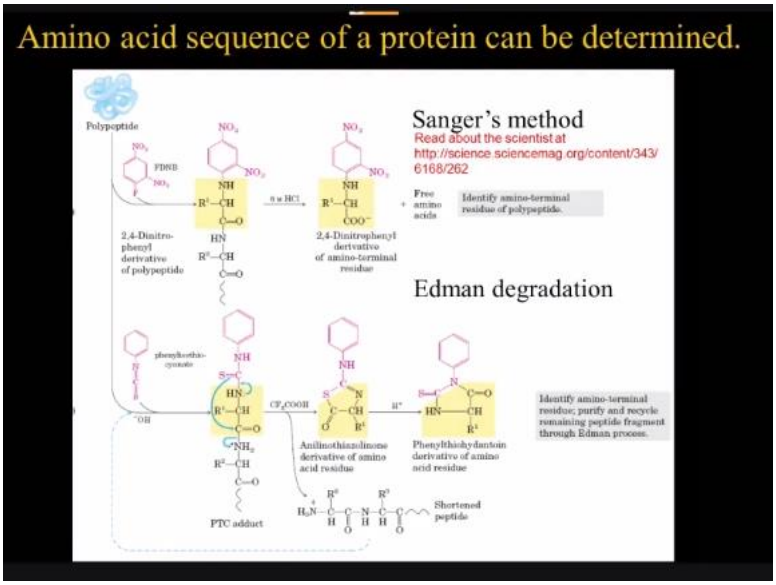
So all dots in one vertical column, they all started together because they all have the same pI and therefore they were in one band and they are in different rows now when you look horizontally that is because their molecular weight is different and by doing this instead of seeing a couple of hundred bands, now you see about a thousand two thousand spots and this is called a two-dimensional electrophoresis.

First dimension is separation is based on pI and this method is called isoelectric focusing and the second dimension is based on the charge and that is SDS-PAGE. So this isoelectric focusing coupled SDS-PAGE is what we call two-dimensional electrophoresis. So this technique is very powerful in separating the protein molecules. In modern times, the technology has advanced well enough that we can generate reliably identical patterns reproducibly when you repeat with the same sample.

And we have the ability to excise out a given spot and identify what protein that is and doing this kind of a thing is a bedrock of a modern approach called proteomics where you study all proteins of a particular cell or an organism in one go, like genomics it is proteomics. So, this two-dimensional electrophoresis is one of the really powerful techniques of proteomics okay. So, this completes our discussion on the protein based methods.

So there are many more methods, we are not going to get into all the details, this kind of gives you a flavor of what kind of things people do with proteins.

(Refer Slide Time: 42:04)



So next what we are going to do is I am going to open a different power point. So here our goal is to how do I find out the amino acid sequence. So what is in one end and what is next and what is next it is a polymer and I would like to know all the amino acids in the sequence they appear in the given polypeptide chain. So sequencing a protein so that is the topic here. So a long time ago, Frederick Sanger found a way to determine the amino acid sequence.

So he got a Nobel prize for that and then several years later he again found another breakthrough sort of method that is determining the nucleotides in DNA and for that he got one more Nobel prize, so he is one of the few scientists or maybe the only one, my memory is not sure, who got 2 Nobel prizes unshared and his life is very interesting and inspiring. So I urge you to go to this link and read about him, read about the scientist.

So this will appear here and there when I describe certain specific discoveries pertaining to a certain scientist who contributed immensely to the field of biology. So Sanger is one such scientist, so please when you have time go to this website and read about Sanger. So now what is the Sanger's method of sequencing proteins? So essentially what he does is the protein chain has an N-terminal residue.

Meaning at one end of the chain the amino group of that amino acid at the end is not linked to any other amino acid because it is the end there and that amino acid that has the free amino group that is the end terminus of the chain, this I told you already and the other end will have the carboxyl group free that we call C terminus. So the N terminus amino acid is derivatized using a reactive group.

Here the group that we use is the FDNB, this is a fluoro dinitrobenzene okay, FDNB this group is strongly reactive with this amino group and it is covalently linked. Now you break the peptide bonds by adding 6 molar HCl and boiling. Then all amino acids, you know the peptide bonds get hydrolyzed into individual amino acids. Now you identify the amino acid that is the derivative with FDNB.

So because this amino acid is going to have this group which does not exist in the normal polypeptide and from this particular amino acid whatever be the R group is the N-terminal amino acid. So this is how you can identify the N-terminal amino acid and then by doing partial overlapping cleavage in a very laborious way you can identify the subsequent amino acids as well, but to make it lot easier to ensure only the N-terminal amino acid is cleaved at a time.

Edman came out with a method called Edman degradation. So this is the method that is currently used for sequencing proteins if at all protein needs to be sequenced. Most of the time now we directly translate from the coding sequence of the nucleic acid because nucleic acids are a lot easier to sequence and based on that we determine the amino acid sequence purely by theoretical translation of the triplet codon.

But if a particular experiment demands that the actual protein is sequenced and for that currently used automated method uses this chemistry called Edman degradation. So here in the FDNB he uses phenyl isothiocyanate. So this is cyanide and an oxygen would have been cyanate and since it is sulfur so it is thiocyanate and it has phenyl group, so therefore it is phenyl isothiocyanate or commonly PTC.

So PTC forms an adduct with the amino group and then through electron flow indicated by these arrows it cyclizes and forms this carboxyl group gets hydrolyzed, this peptide bond gets hydrolyzed and that forms a link with this sulfur and this is the thiazolinone derivative of the N-terminal amino acid and this can be separated and this hydrolysis is done by trifluoroacetic acid, instead of boiling with 6 molar HCl this is a milder but effective treatment for this cleavage and the cyclization.

So now this derivative can be identified while the rest of the polypeptide chain is left intact.

So this is a polypeptide chain that is shortened by only one amino acid. Now this you can purify and repeat the process again and by doing this, this sequence can be done. So, these are usually done by fixing the protein first to a solid matrix, a disc like paper like matrix on which the protein is added and this is at the bottom of a tube.

So the solutions are added, reaction happens and then you elude this molecule out, then you wash it and the shortened polypeptide is still there in the matrix, then you add again. So like this a machine can repeatedly do this cycle and each of this that comes out can be separated by a chromatographic method and based on the absorbance you can get the peaks and based on the distance migration you immediately know what amino acid that is based on the reference distances migrated by known amino acid derivatives with the PTC.

So this is how the proteins get sequenced uh in modern time, so this is Edman degradation. The primary improvement over the Sanger's method is in Sanger's method the entire protein gets hydrolyzed into individual amino acids. So that is laborious to go to find them second one, third one and so on. On the other hand, this PTC adduct formation cleaves only the N-terminal amino acid leaving the rest of the polypeptide chain intact and ready for a next cycle.

So, I will stop here with the amino acid sequencing, then tomorrow we will see how do we do the opposite like if I give you a mixture of amino acids and if I want you to make a polypeptide chain having a certain sequence, how do we synthesize that? We do not artificially synthesize a whole protein having few hundred amino acids, but we can make peptides having 20 amino acids or you know oligopeptides of that range chemically can be synthesized. So, we will look at that tomorrow.