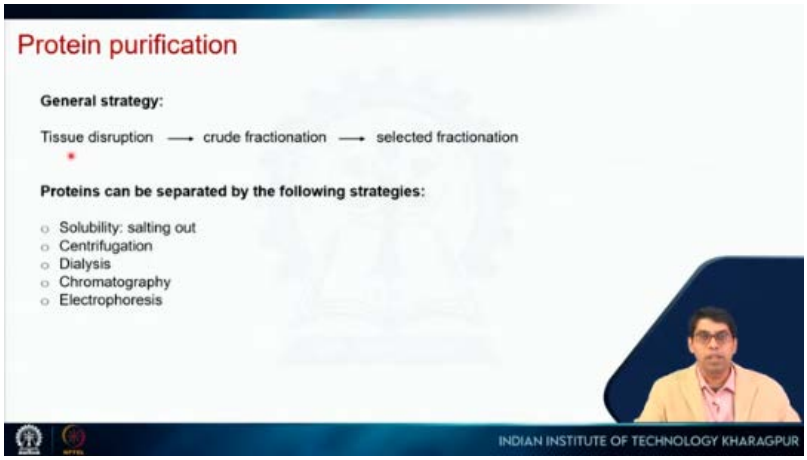
Introduction to Complex Biological Systems
Professor Dibyendu Samanta and Professor Soumya De
Department of Bioscience and Biotechnology
Indian Institute of Technology, Kharagpur

Lecture 57
Chromatographic methods

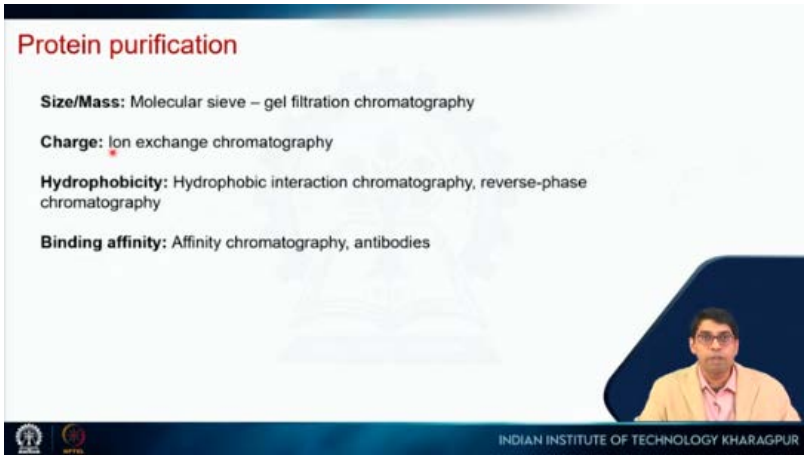
Welcome to Lecture 57 of Introduction to Complex Biological Systems. Today, we are going to talk about chromatographic methods. So, these are the concepts that I am going to cover. So, I am going to talk about protein purification. I will discuss the basics of chromatographic techniques, then I will discuss these three primary techniques that are gel filtration chromatography, ion exchange chromatography, and affinity chromatography.

So, this is the general strategy that is shown, and we have already seen this in different lectures that we may want to purify proteins from some source. It can be a recombinantly expressed protein in bacteria, or it can be something that we are actually identifying or extracting from a tissue. So, we will start with this general strategy. We will disrupt the tissue if it is tissue and get to the cells, then disrupt the cells, do some crude fractionation, and then do some selective fractionation. So, we can do multiple stages of this to get the final protein at the end.

So, proteins can be separated by using the following strategies. The solubility of the protein can be used. So, this is called sorting out. We can use centrifugation, where bigger proteins will precipitate, and smaller proteins will remain in solution. So, we can use some strategies there.



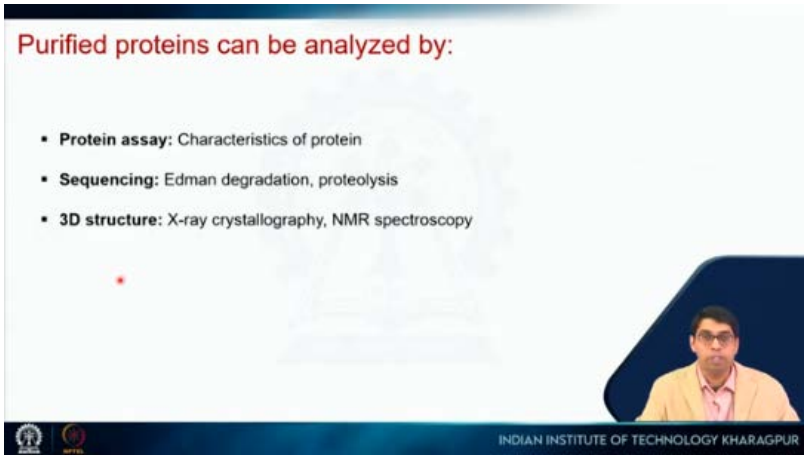
Dialysis to remove some smaller molecules, we can use chromatography and electrophoresis. So, we have already talked about gel electrophoresis, and this is mostly to identify proteins. So, we have already seen that in the previous lecture, and we will see that this is something we use to identify the purified protein, whether we have a purified protein or not. So we can purify proteins according to their size or mass. So, this will be a molecular sieve.



Examples will be filtration chromatography or size-exclusion chromatography. We can take advantage of the charge on the protein. So, we can use ion-exchange chromatography. So, at a particular pH, if your protein is positively charged, then you can remove it or separate it from all other proteins which are negatively charged at that particular pH.

We can take advantage of the hydrophobicity of the protein. If there are hydrophobic groups exposed on the protein, then it will stick to the reverse-phase column more than other proteins. Finally, this is more specific. We can use the binding affinity of the protein

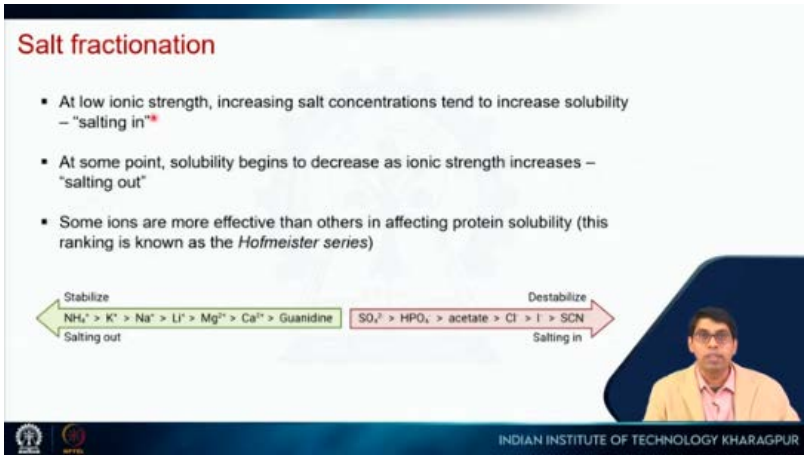
to certain molecules. So, we will see examples of all this in today's lecture. Once we have purified the protein, we would want to know whether It is or how pure it is. Is it just one protein, or is it a mixture? Are there still two or three different proteins present there?



Do we have the right protein? So, we can do all of that by doing an assay. So, if it is an enzyme, we can set up an enzymatic assay. We have already seen one example here, and I will remind you in the last slide of today's lecture. If you do not know the sequence of the protein, then, of course, you will have to go for sequencing. So, you can use Edman degradation, or you can use mass spec, which is not that difficult.

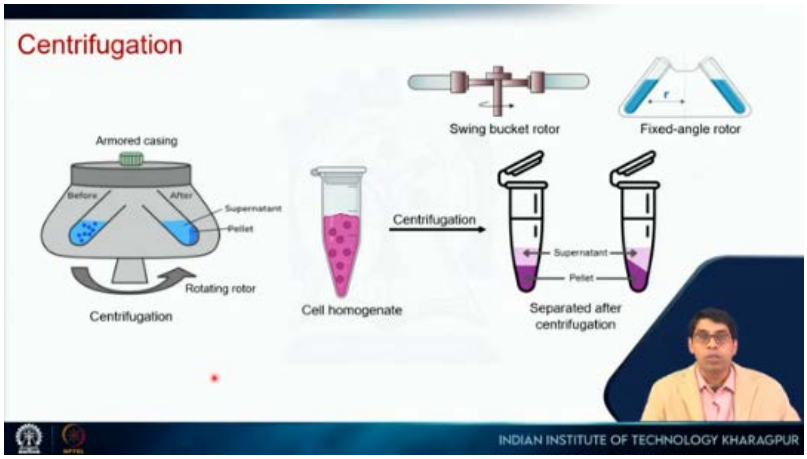
Finally, if you want to know the structure of the protein, you can then go for solving the three-dimensional structure of the protein using X-ray crystallography or NMR spectroscopy, or if it is a big protein, then you can go for cryo-electron microscopy. So, here we are concerned with the purification of the protein. Now, selective precipitation can be done by using different ionic strengths and pH. So, if your protein is insoluble at a certain pH, then you can precipitate it by making your buffer at that pH. You can use certain organic solvents. You can also use crowding agents, which can sometimes precipitate out your protein of interest, such as polyethylene glycol, which will be one such crowding agent. So, the most common method that is used is salt fractionation. So, at low ionic strength, increasing salt concentration tends to increase the solubility of the protein.

If you have a salt concentration, let us say between 50 millimolar to 150 millimolar of sodium chloride or some other salt like potassium chloride, it will improve the solubility of your protein. So, this is called salting in. But if you keep on increasing the salt concentration, let us say you go to 1 molar or 2 molar, then the protein will precipitate out. So, this is called salting out, and it turns out that not all ions are equally effective in this salting out. So, some ions are more effective than others, and this is listed in a series called the Hofmeister series.



So, you see that the cations or the positive ions are shown here, and the anions are shown here. Ions on this side help in salting out, and ions on this side will actually denature the protein. So, we know that if you use guanidinium hydrochloride. So, guanidinium and chloride ions will completely denature your protein if you use a 6-molar concentration. On the other hand, if you take ions from this left-hand side, like ammonium sulfate, that is something used to effectively salt out your protein.

So, ammonium sulfate is the first choice people use to salt out proteins. If that does not work, then you can try some other combination. So, once it salts out, you actually have to precipitate it. So, you can let it sit and hope that it will precipitate out, but that can take time. However, we can speed up the process by using something called a centrifuge.

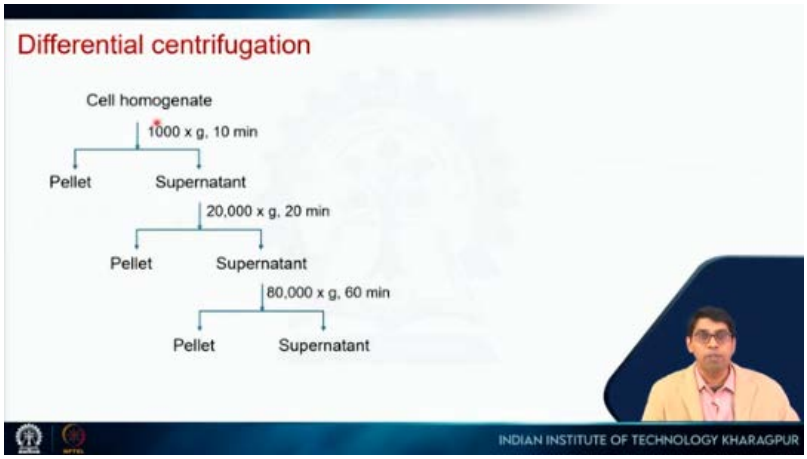


So, the process will be called centrifugation. So, all you have to do is, let's say, you have your protein of interest in a tube; you just spin it, and then these larger clumps that are formed will all precipitate out like this. So, you have already seen this setup when we discussed the Meselson-Stahl experiment. So, there, a gradient was used to separate out the different DNA, N14 and N15 DNA. So, a centrifuge is a very useful instrument, and it turns out that it comes in two flavors: one is called a fixed-angle.

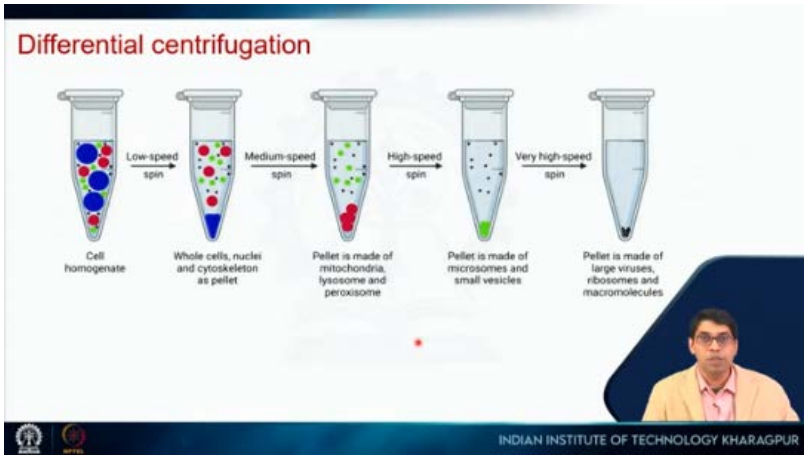
So, this is called a rotor. This is something that will rotate at very high speed. You can control the speed, and you can also control the temperature if it is a temperature-controlled centrifuge. Now, the tubes you will put in and they go in a certain angle. So, this is called a fixed angle rotor.

In some cases, there is no fixed angle, but the tubes are connected to something that is a bucket and this bucket has a hinge. So, when the rotor rotates these tubes will fly out and become almost horizontal. So, this is a swing bucket rotor. So, both are very common and they have their different applications. So, you will see that we will use centrifugation at various stages of protein purification.

Suppose we have the cell homogenate. So, you have disrupted the tissue, the cells. So, you want to precipitate the cells from whatever slurry that you have. So, you can centrifuge it. You will do that at a lower speed. You will get the pellet, everything will be in the supernatant.



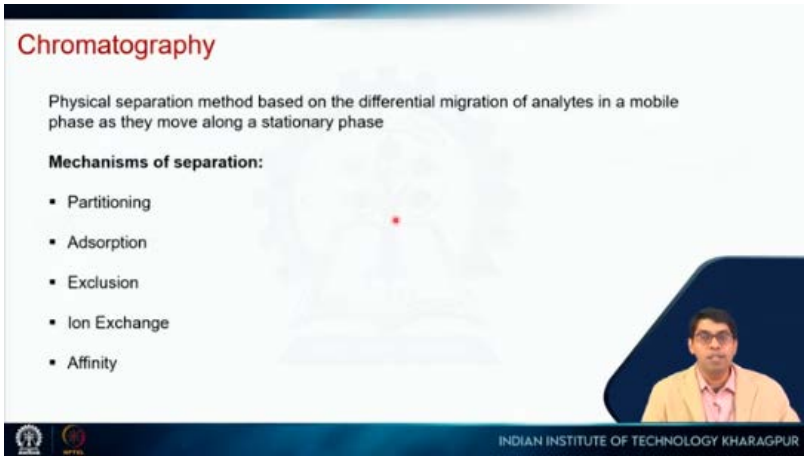
So, there you have to choose if you want to take the cell, you will take the pellet and then you go from there, or if you do not want to take that, you want to take something else, then you take that. Centrifugation is done at multiple steps. In some cases, we take the pellet, in some cases we take the supernatant, then we will go for a higher speed of centrifugation. So one example is shown here. So we have the cell homogenate here. In this case you spin it at a low speed as shown in the previous slide.



So now your Precipitate will form. In this case, the larger pieces or the larger entities, these are whole cells, nuclei, and cytoskeleton, as pellets. All of these things will go to the pellet. If you are interested in this, you will take this and proceed further. If you are not interested

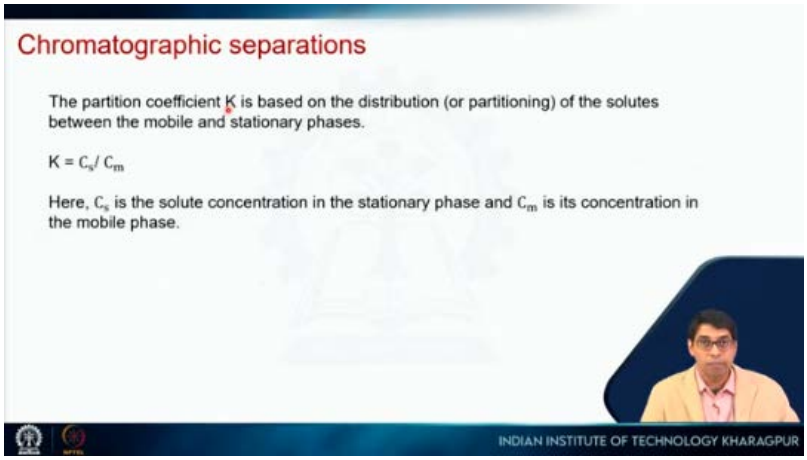
in this and are interested in the supernatant, then you will take that and proceed further. So, now in this case we will spin it at a higher speed. So, now it is a medium speed and now the precipitate pellet is made up of mitochondria, lysosomes, and peroxisomes. If you are interested in this, you take it; otherwise, you go for the supernatant.

So, now in this case we separate out these microsomes and small vesicles and then finally you are left with the proteins. So, let us say we have the proteins, for example; if we do this for bacterial cells, we crack open the bacteria, DNA will come out, RNA will come out, all the proteins will come out, and there will be cell debris. So, we can use precipitation and centrifugation to precipitate the cell debris and the DNA and other things and let us say only proteins are there in the supernatant and some small molecules, but still we have to purify our protein of interest from a mixture of almost 5000 different proteins that are expressed by the bacteria. So, now we have to use chromatography techniques to separate our protein of interest from a mixture of proteins or mixture of molecules.



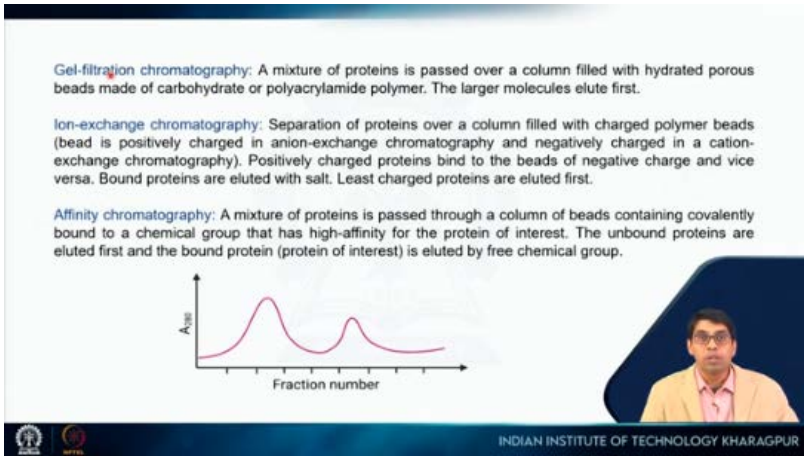
So, what is chromatography? It is a physical separation method based on differential migration of analytes. So, in this case, these are all the molecules which are dissolved in some solution. So, differential migration of these analytes in a mobile phase, which is your buffer. As they move along a stationary phase, which will be some solid matrix and the method of separation can be partitioning, adsorption, exclusion, ion exchange or affinity. So, we are going to talk about these three in detail today. The partition coefficient K is a term that is used quite often in chromatographic techniques. So, it is based on the

distribution or partitioning of solutes between the mobile phase and the stationary phase. So, you have two phases.



So, let us say we have a column and in that we have some gel matrix. So, the gel matrix will be the stationary phase, and the buffer or the solution in which your protein is dissolved is the mobile phase because it will pass through the column. So K is defined as C_s/C_m where C_s is the solute concentration in the stationary phase. So this s stands for the stationary phase, and C_m is the concentration of the solute in the mobile phase. So this m stands for mobile phase.

So, the concentration in the mobile phase and the concentration in the stationary phase, you take the ratio of these two, and you get your partition coefficient. So, I will discuss these three chromatography techniques, primarily gel filtration. So, in this case, a mixture of proteins is passed over a column filled with hydrated porous beads made up of carbohydrate or polyacrylamide polymer. So, these beads are polymers. Now, larger molecules will pass through faster, and smaller molecules will go slower.

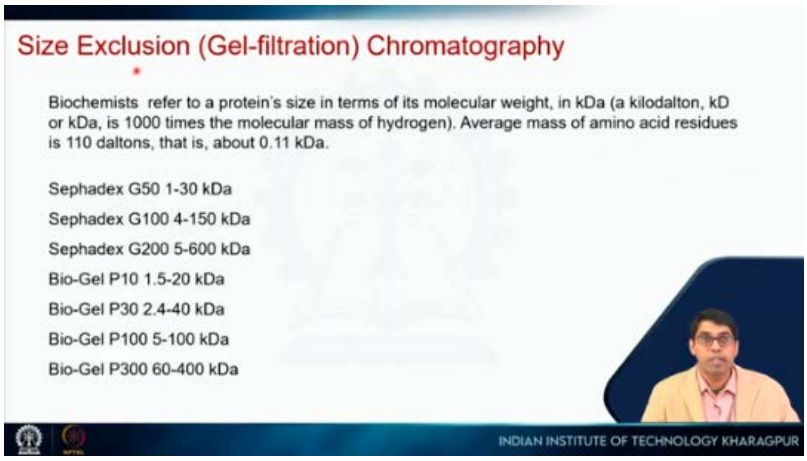


So, the larger molecules will come out first, and the smaller molecules will come out next. Ion exchange chromatography, here, we use charge. So, the separation of proteins over a column filled with charged polymer beads. Now, the bead is positively charged in anion exchange chromatography and negatively charged in cation exchange chromatography. So, positively charged proteins will bind to the beads of negative charge, and negatively charged proteins will bind to the beads of positive charge. So, once they are bound, how do you elute them? So, you can elute them by increasing the salt concentration. So, again, in this case, let us say we have a mixture of proteins. Now, your protein of interest has, let us say, a pH of 8. It has a highly negative charge. So, you will pass it through an anion exchange column.

So, the negatively charged protein will stick to it. Now, there will be other proteins which will be positively charged. So, they will elute out first. They will go out first. Then, there will be proteins which will not have that much of a negative charge because of their PI. So, they will also elute out at different positions and then, finally, when you increase the salt concentration, your protein will come out.

So, you can separate your protein from all other proteins, but these two methods might not give you completely pure proteins. A more sophisticated method is affinity chromatography. So, in this case, a mixture of proteins is passed through a column of beads containing covalently bound chemical groups for which your protein has high affinity. So, let us say your protein can bind to glutathione.

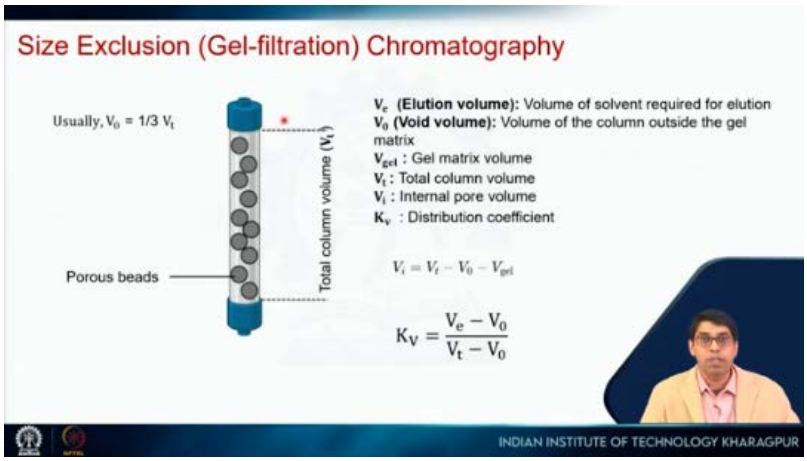
Then what you will do is you will add glutathione to the column. So, glutathione will be covalently linked to the column, and then your protein will bind to it. Everything else will flow through and just go out because they will not bind to glutathione and then you can pass dissolved glutathione through your column. So, your protein will bind to this soluble glutathione and then it will come off the column. So, let us look at size exclusion chromatography, the first one.



So, in this case, protein size is used to separate the proteins. Now, we refer to protein size in terms of molecular weight, and the typical unit that is used is kilodaltons. So, we write it as kD or kDa. These are different columns that are typically used, and you can see that these columns will have ratings like this, which means they can separate molecules in this range. So, smaller proteins can be separated using this column, whereas medium-sized proteins can be separated by this column, and larger proteins can be separated by this column.

So, depending on your expectation, I will have a mixture of proteins which will be in the range of 10 kilodaltons to 100 kilodaltons. So, you will go for a column like this or a column like this. So, this is how a typical column or size exclusion column or gel filtration column will look. These are the beads. It will be a glass column and then it will be connected by tubes at the two ends. So, the protein will enter in this and then it will pass through and then it will come out from this end and then there are of course automated machines which will pass these proteins through these columns at a particular speed. If you want to control temperature you can put all of these things inside a refrigerator. Now there

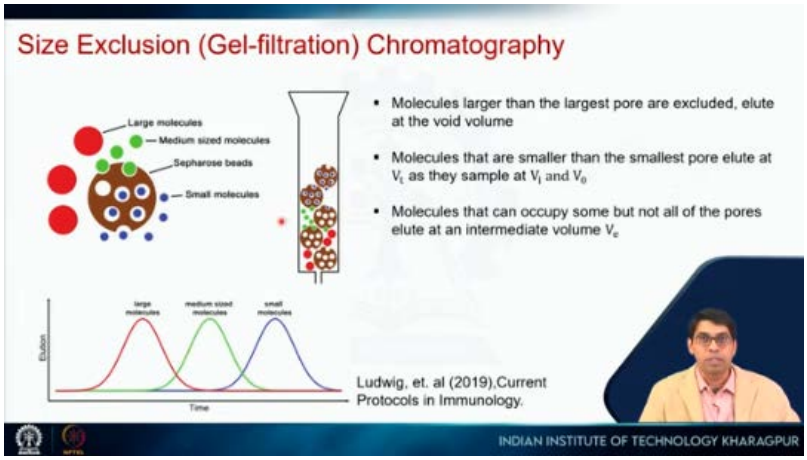
are certain terms that are very useful. We are going to look at these terms. So V_e is the elution volume so that is the volume at which your protein is eluted, V_0 is the volume of the column outside the gel matrix, and V_t is the total column volume. So, we will use these terms to define something called the distribution coefficient. What is the distribution coefficient? Let us say the total volume of the column is 120 milliliters.



Now, that will be V_t and the void volume is V_0. So, what is void volume? Any molecule which is so big that it does not interact with the bead or enter the bead will just pass through these interstitial spaces. So, it will just go through like this. That is the void volume, and it turns out that typically the void volume will be around one-third of the total volume.

So, $V_t - V_0$ is a constant. V_e is something that will change depending on the protein. If your protein is very small, then V_e will be very close to V_t because it will enter all the pores and nooks in the beads. So, it will pass through a very large volume. So, it will elute very late. But if the protein is big, then it will not do much of that. So, it will elute faster.

So, your elution volume will be smaller. For example, if we say my V_t is 120ml and V_0 is 40 ml, then this will be $120 - 40 = 80$. So, the denominator is 80 milliliters. V_e for a bigger protein will be, let us say, 60. So, 60 minus 40 is 20. So, 20 by 80; my K_v will be one-fourth. If it is a small protein, then V_e will be, let us say, 100. So, this is $100 - 40 = 60$; $60/80$ is three-fourths. So, this K_v, the distribution coefficient, will depend on the size of the protein. So, this is what happens. These are the beads. The beads are packed inside the column. Now, we have three different-sized molecules: large-sized molecules shown in red, medium-sized molecules shown in green, and very small molecules shown in blue.

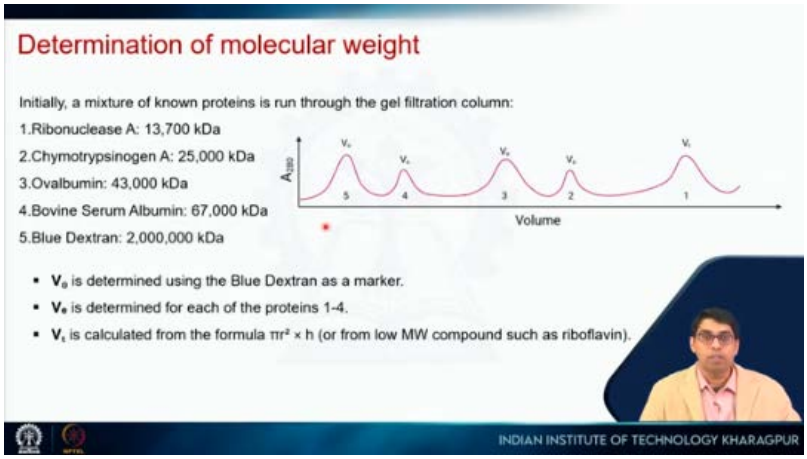


So, you can see that this bead has its size, and then there are these small pores in which these medium-size molecules can interact, and these small-size molecules can also interact. Now, the medium-size molecules will interact only with these pores which are on the surface, but these small-size molecules are so small that they can enter some of these pores which are inside. So now, if we pour this mixture, then what happens to these large molecules? They will not enter or interact with any of these pores, so they will just pass through the spaces that are between these beads and it will come out. So, it will come out very quickly. So, this is how I am looking at the protein; it can be absorbance at 280, and this is the time or volume. So, the large molecule will come out first, or they will come out at a lower elution volume, let us say 60 ml.

The medium size molecules will not enter these internal pores, but they will interact with these surface pores. So, they will spend more time in the column, but less time compared to the small molecules. So, then they will come out next which is shown in green and then the small molecules will come out because they are passing through all the pores. So, then the small molecule comes out.

So, this is how we can separate a mixture of these three proteins using these beads using this size exclusion or gel filtration chromatography. Once we have done that, we have done our separation, but this technique not only gives you a purer protein, but it can also tell you what the molecular mass of your protein is, approximately. So, how do we do that? For that, what we do is we run a mixture of proteins. So, these proteins will depend on the actual separation of the column. So, if the column separates something between 10

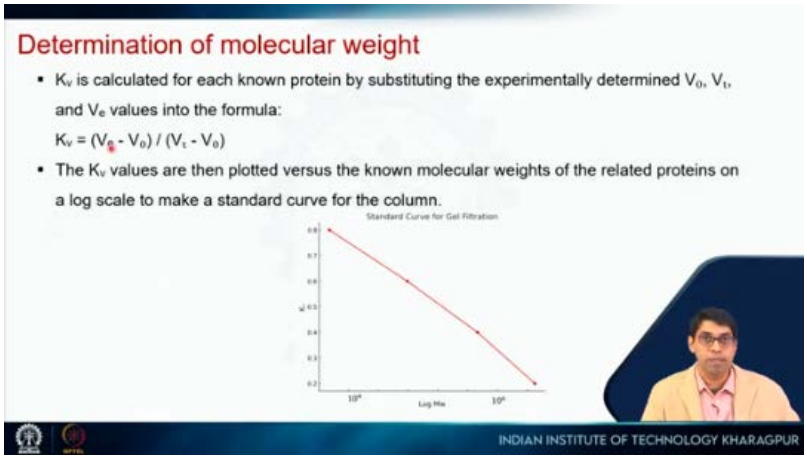
kilodaltons and 100 kilodaltons, then maybe this is fine, but if it separates something between 10 kilodaltons and 600 kilodaltons, then you will want to use some proteins that are bigger than bovine serum albumin.



So, you have this mixture of proteins. You pass it through, you know, their mass. So, what you do is you get the elution volume of those proteins. Now, blue dextran is a very big molecule. So, it will not go through any of the pores. So, it will just pass out so that will give you the void volume.

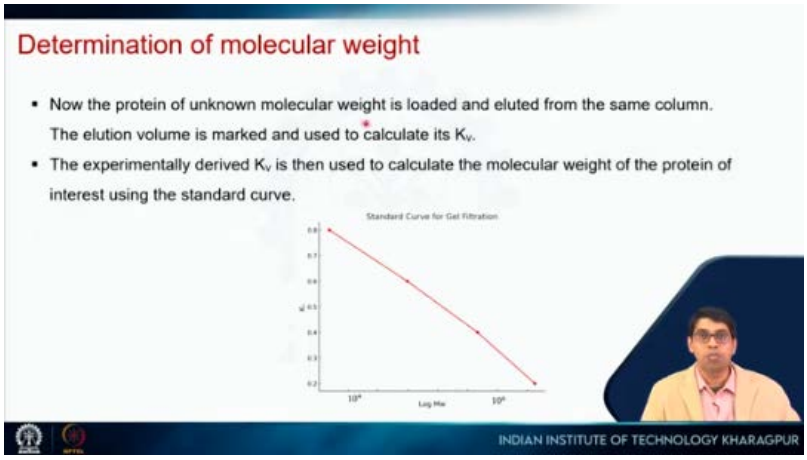
So, you can determine the void volume using this blue dextran, and then you have these 4 proteins which will pass out at different times. So, the smallest one will come out later and the largest one will come out fast and you know the molecular mass of these proteins. So, then what we can do is we can calculate this K_v as shown in the previous slide.

$$K_v = (V_e - V_0)/(V_t - V_0)$$

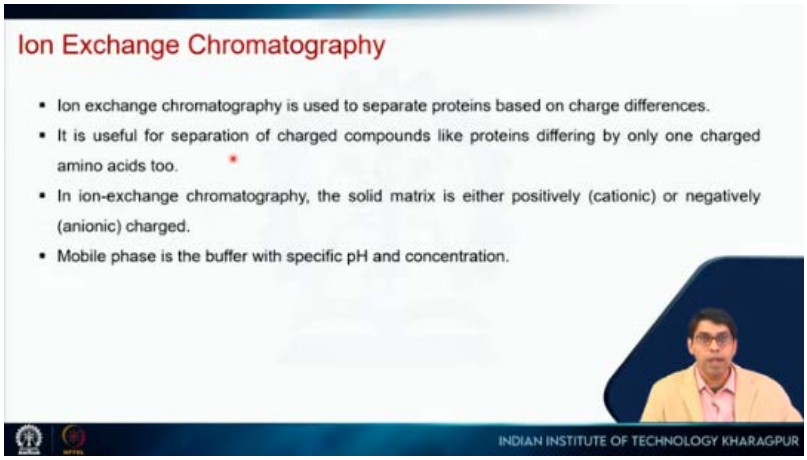


So, you plot that K_v for each of these proteins and on the y axis and the log of the molecular mass because we know this in the x axis. So, this gives us our calibration curve. Now, let us see your protein of interest elutes with a K_v of 0.5. So, then what you do is you go straight here and then you come down here.

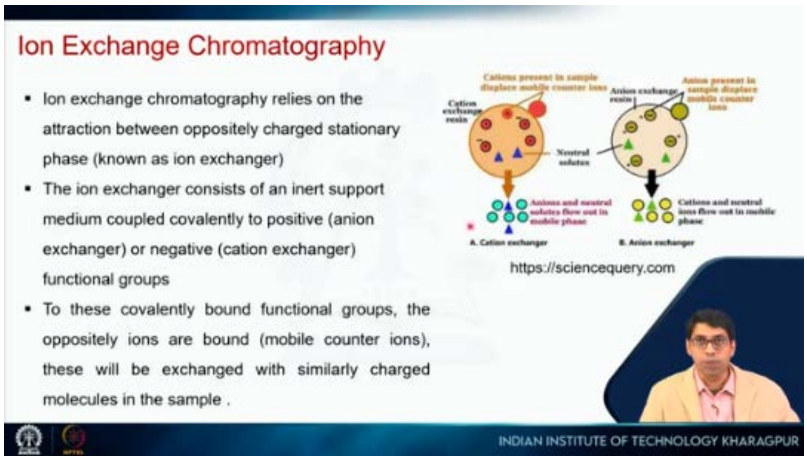
So, your molecular mass will be this. So, you can get up. It is not accurate to the dalton, but in kilodalton, it gives you a very good mass. So, you can determine the mass of your protein. So, an unknown protein is loaded, and then you calculate its K_v and determine that K_v. Then, determine the molecular mass from the standard curve from this curve.



Now, ion exchange chromatography. So, ion exchange chromatography is used to separate proteins based on their charge differences. It is useful for the separation of charged compounds like proteins differing by only one charged amino acid. For example, let us say you have mutated glutamic acid to an alanine, so you have made a change in charge. You can separate out a mixture of these two proteins using ion exchange chromatography.

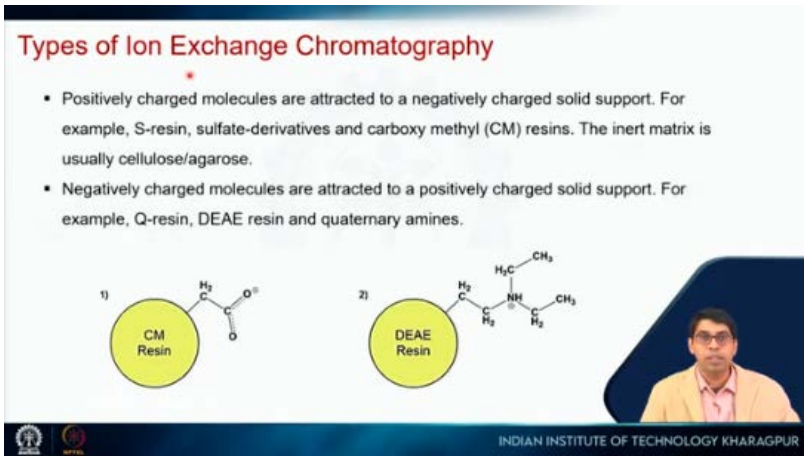


So the solid matrix is either cationic, positively charged, or anionic, negatively charged. The mobile phase is the buffer, and we will determine the pH at which you are going to dissolve your protein; that will depend on the isoelectric point of your protein or the pI of your protein. So, let us look at this. Ion exchange chromatography relies on the attraction of oppositely charged molecules. So, this is the bead.

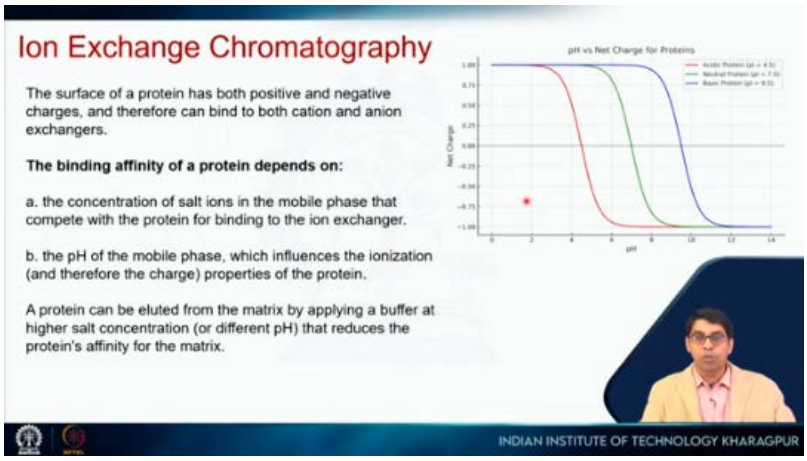


So, in this case, it is a cation exchange resin. So, cations present in the sample displace mobile counter ions. It will bind to the cations or the positively charged proteins, and this one will bind to the negatively charged proteins. So, when you pour a mixture of proteins into this column, only those proteins which are positively charged will bind, and all those proteins which are neutral or negatively charged will flow out, and then, by using a salt gradient, you can slowly replace these proteins which are positively charged. So, proteins which have a higher positive charge will stay bound for a longer period of time, and proteins which have a lower positive charge will elute first.

So, again, you will see these proteins getting eluted at different times, and hopefully, your protein will come out as a separate peak; then you have your pure protein. So, the same principle holds for the anion exchange column. There are different resins that are used; positively charged molecules are attracted to negatively charged solid support. So, in this case, the resins will be S resin, which are sulfate derivatives or carboxymethyl resins. The inert resin matrix is usually made up of cellulose or agarose.



So, this matrix is made up of cellulose or agarose, and then there are these groups which are covalently linked and these groups are negatively charged or positively charged. So, this will bind to positively charged proteins, and this will bind to negatively charged proteins. So, in this case, the resins will be Q-resin or DEAE resin or quaternary amines. So, this just shows how we are going to choose the pH.

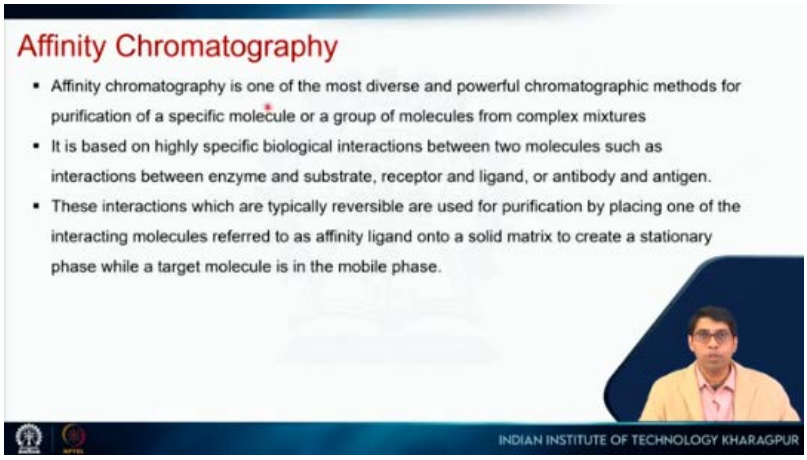


So, for example, we have let us say there is a mixture of three proteins, and they have three different pIs. So, one has a pI of 4.4, the other one has a pI of 7, and the third one has a pI

of 9.5, and they are shown as red, green, and blue. So, here you see pH and their net charge. So, net charge as a function of pH. So, suppose if I want to separate out these three proteins, I am interested in this red protein.

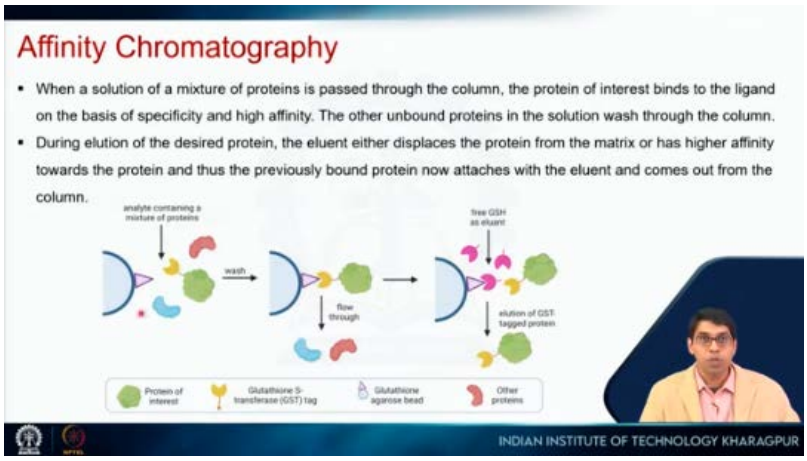
So, what I will do is I will use up pH of 6. So, if I use a pH of 6, if I dissolve these three proteins in a buffer which has a pH of 6, what will happen? This protein will have almost a net charge of -1. So slightly less than -1, but these two proteins will have a net positive charge. So this one will be almost around +0.8, and this blue one is practically +1. So, +0.8, +1, and this is -0.9. So, now, if I use an ion exchange column, I can separate out these two proteins from this protein. So, this is the basic principle.

So, you have to be very careful about choosing the pH, and you will choose the pH depending on the pI of your protein. So, you will have to have some idea about the pI of your protein. If you do not know the pI, then you can still do it, but you have to then use some sort of assay to find out where your protein is. Is it going in the flow-through, or is it sticking to the column and coming out in different fractions? So, affinity chromatography is the most powerful method.

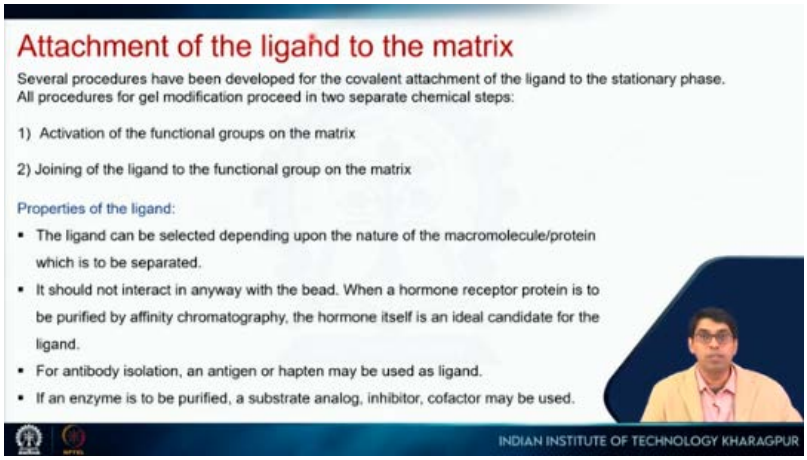


So, in this case, what we are going to do is use some highly specific biological interaction to purify the protein. For example, if it is an enzyme, we can use its substrate, or its inhibitor or substrate analog. If it is a receptor, we can use its ligand or ligand analog. If it is an antibody, we can use its antigen or antigen analog.

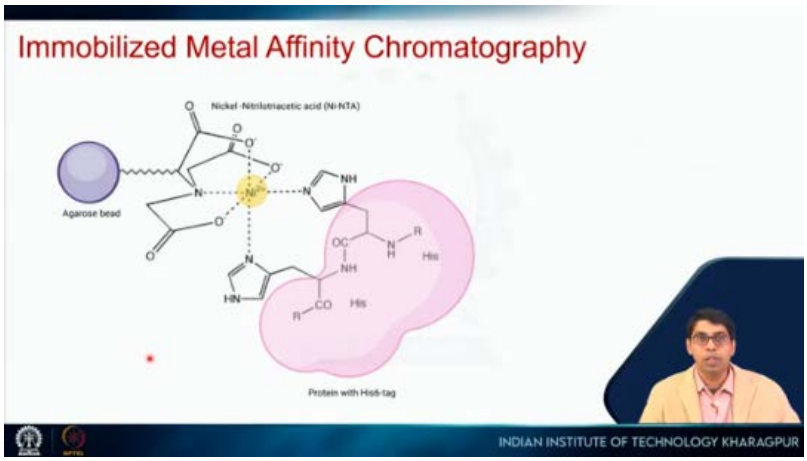
So, these molecules will be coherently linked to the solid matrix and then will be used to separate out the protein. So, one example is shown here. This is an agarose bead, so you can see the lesions down here. That's the agarose bead, and this triangle is glutathione. What we want to do is purify our protein of interest. Now, this protein, using recombinant DNA technology, is fused with another protein. This protein is glutathione S-transferase. So, a chimeric protein is produced where this glutathione S-transferase is covalently linked to our protein of interest. Now, you crack open the cell; it can be a mammalian cell or a bacterial cell, so there are all sorts of proteins, but only your protein has this glutathione S-transferase tag. So, once you pass it through the column, this GST protein will bind to the glutathione and since your protein is covalently linked to it, it will also stay there. Everything else will flow out. So, you will wash your column thoroughly; everything else goes out, and only your protein sticks to the column. Now, what you do is flow free glutathione through this. So, free glutathione you can flow, or you can flow the free protein.



So, that will bind, and your protein will come off. So, there are several ways we can do this; several processes have been developed for the covalent attachment of the ligand to the stationary phase. So, typically, we do not have to worry about this because you will get these columns commercially, like the GST column, which is available commercially. However, if there is something specific you want to do, then you will have to do it yourself. Now, the property of the ligand should be such that the ligand can be selected depending on the macromolecular protein.



It should not interact with the bead in any way. So, there should not be any interaction and you should be able to purify your protein. So, break that interaction using some other common molecule. Another very common chromatography that is used there is also affinity chromatography, but in this case instead of GST or glutathione S transferase, 6 histidine amino acids are used. So, your protein of interest will be attached to 6 histidines.



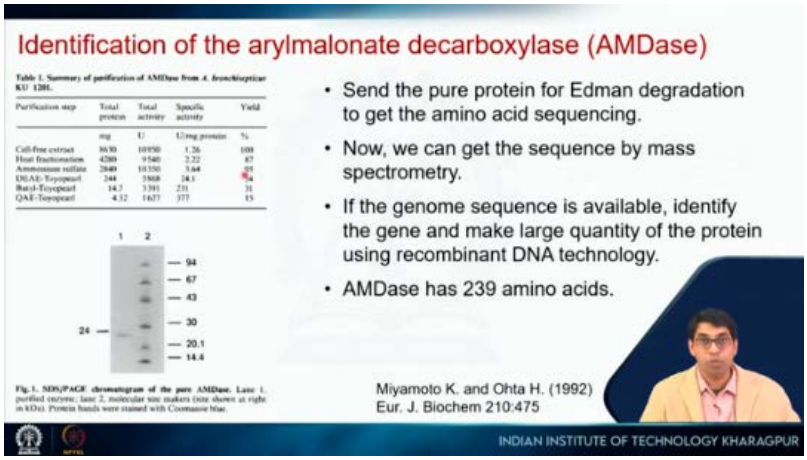
So, again you can do that using recombinant DNA technology. Now, these 6 consecutive histidines interact with nickel ions like this. So, nickel ions are immobilized on an agarose base like this. So, when you have your protein which has this consecutive histidine, it will interact with the nickel ion and become immobilized.

So, other proteins, chances of having 6 consecutive histidines is very less in any protein. So, all other proteins will elute out; only your protein will stick to this. So how do you elute it? How do you get your protein out of the column? You use imidazole, which is only this molecule that is present in the side chain of histidine. So you pass a solution of imidazole,

and it will come and interact with this nickel, and your protein will elute out. So what are the applications? There are many applications, so these are used for the isolation and purification of all proteins, enzymes, etcetera, and we have already seen one example, so let me go to that slide. So, before that once you have purified your protein you want to know whether you have your protein or not right or what is the concentration of that protein or whether that protein is pure or not.

So, we use all the techniques that we have learned earlier. So, if you want to know if the protein is pure or not, you can do that using an SDS-PAGE. If you want to know whether the protein is there or not, you can use a Bradford assay. If you want to quantify it, you can do so using a Bradford assay. If you know the extinction coefficient, then you can also use the absorbance at 280 nm to determine the concentration of your protein.

So, we have already seen this example. So, this was aryl malonate decarboxylase or AMDase. So, this scientist was purifying this from a particular bacterium. So, first they had a cell-free extract and they saw that yes, there is activity. So, they were using this assay to identify whether their protein was present in a particular fraction or not.



So, then they did heat fractionation. So, they heated their cell lysate. A lot of proteins precipitated, but some stayed in solution. So, you can see that the amount goes down. So, now, they again took that supernatant.

So, they centrifuged it, took the supernatant, checked for activity, and saw that there is activity. So, this is good. It means that their protein did not precipitate. So, then they used to salt out with ammonium sulfate. So, again, many proteins precipitated.

So, again, it depends. You have to fine-tune the concentration of ammonium sulfate. So, they used such a concentration after several trials where a lot of proteins precipitated. So, you can see that almost half of them are but their proteins still remain in solution, so there is activity. Then you can see they used anion exchange, cation exchange, and also hydrophobic interactions. So, they used all these columns, and you can see that progressively their protein is getting purer and purer. So, less and less amount of protein is there, but the activity is there, and the specific activity, which is activity divided by the mass of the protein, is increasing, which is very good. It means that you are throwing out other proteins and keeping your protein of interest, which is the enzyme AMDase. Then, finally, they used SDS-PAGE to see that, yes, what they have at the end is a pure protein. They do not see any other band; they see only their protein and then go for Edman degradation to get the amino acid sequence, but these days we would go for mass spectrometry.

So, that is all for today. Thank you.