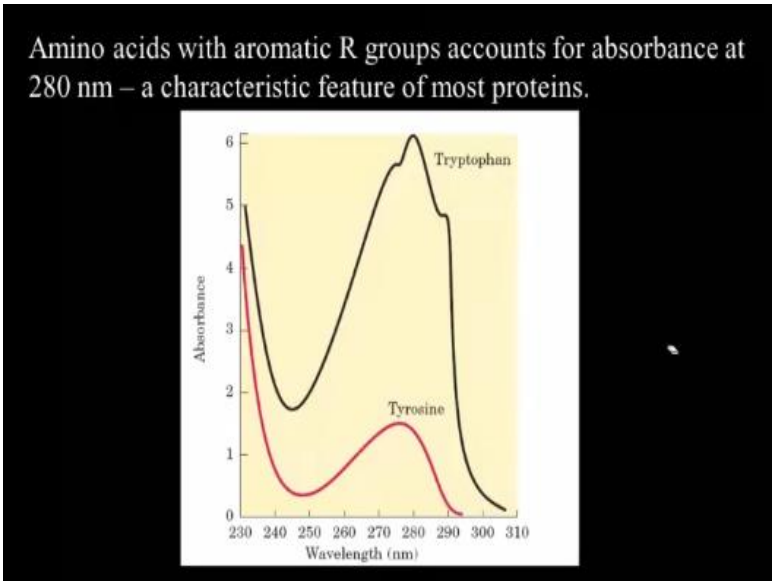
Introduction to Biomolecules
Prof. K. Subramaniam
Department of Biotechnology
Indian Institute of Technology - Madras

Lecture – 5
Characteristics of Proteins and Chromatography Techniques

(Refer Slide Time: 00:14)



Okay so we will continue our discussions on the proteins. So in the last lecture in the last slide I was listing you the 20 amino acids and the structures and their groupings. We had a non-polar aliphatic side chain amino acids, then aromatic side chain amino acids and then we had polar amino acids, then we had the charged ones. So we saw the amino acid structures. So today we are going to very briefly look at some of the characteristics of amino acids.

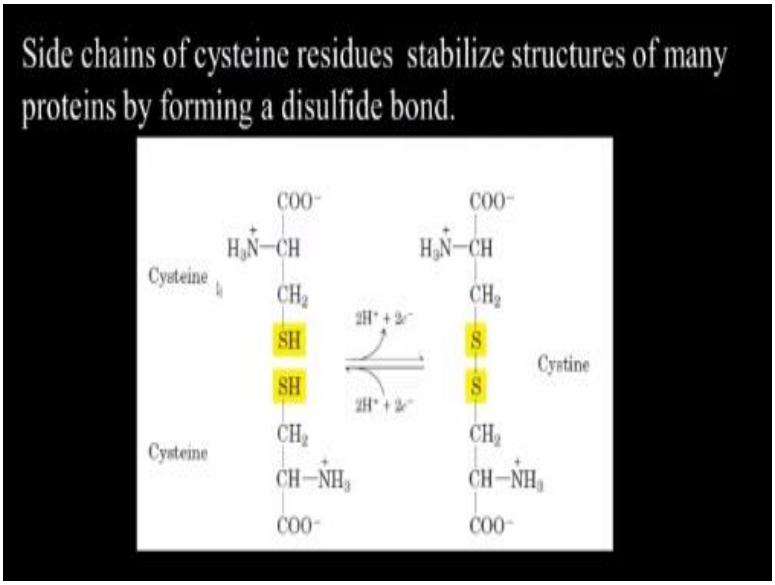
And then proteins and we will go into how to work with proteins, some experimental methods. So the first thing is one of the characteristics of proteins is that proteins have an absorbance be maximum around 280 nanometers. So, a colorless solution of protein can be judged as a solution of protein based on this. They will have a characteristic absorbance peak around 280, so 278 to be precise and this characteristic absorbance comes from the presence of aromatic rings like for example tyrosine and tryptophan.

So this is an indole ring and this is a phenol group, benzene the hydroxyl group. So, presence of these two amino acids give rise to this 280 nanometer characteristic absorbance maximum for proteins and this is often used to measure presence of proteins. Most of the time we use

this in the context of nucleic acid purification. So nucleic acids for example DNA and RNA they have the absorbance peak around 260.

So 260-280 ratios tells you whether you have significant protein contamination in the nucleic acid preparation or not. So that is the context in which we often use this, but otherwise to quantitate proteins or for proper estimation of proteins we go for other reagents that give absorbance at color visible range and they are more sensitive, so very small quantities can be quantitative. But for now, this 280 nanometers absorbance characteristic comes from the presence of the aromatic R groups.

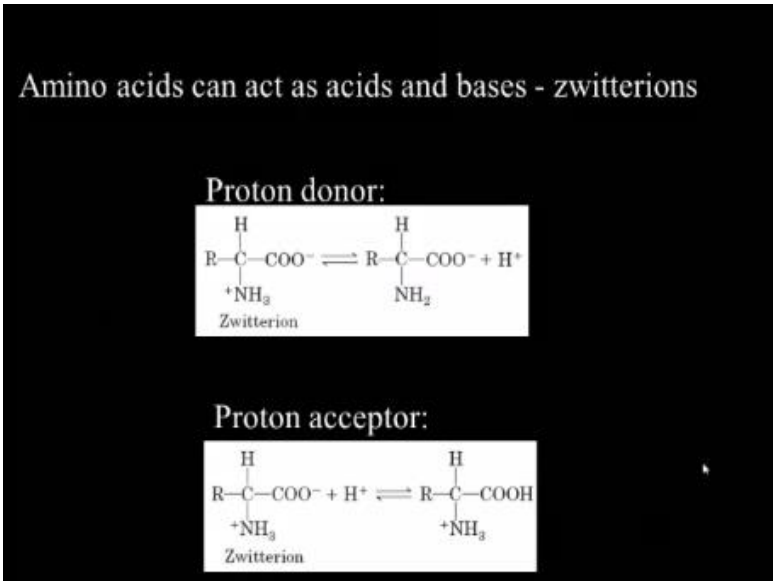
(Refer Slide Time: 03:07)



And the other one I mentioned earlier the cysteine amino acid has a thiol group in its side chain and two cysteine amino acids can form a disulfide covalent bond and this helps in stabilizing protein change in the three-dimensional structure. So in the folded form in the three-dimensional structure if two cysteines come next to each other close enough for the disulfide bond to form, then this is a covalent linkage and that strengthens the structure.

So this amino acid with the free thiol group is pronounced as cysteine and this disulfide form is pronounced as cystine, okay ti, cysteine and this is cystine.

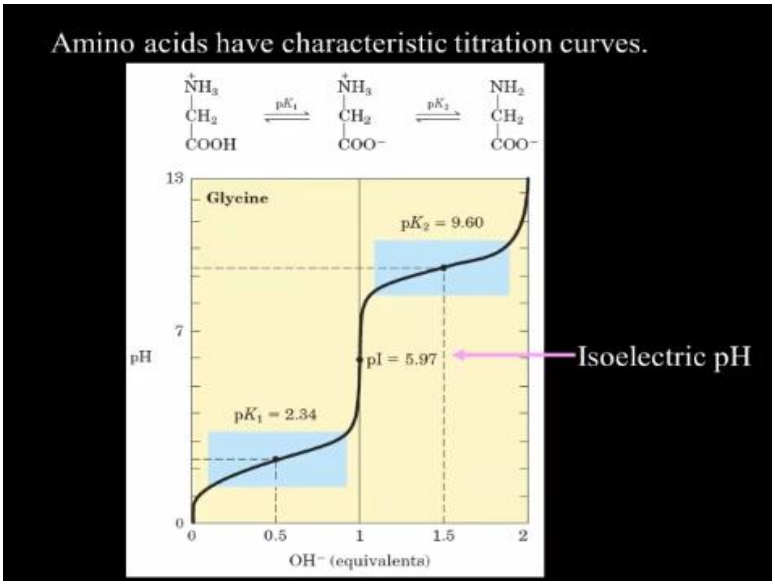
(Refer Slide Time: 04:07)



And the third point to mention here is the zwitterion forms, this I already told you in the last class while talking about weak acids and bases. So here the amino group may be protonated so it is charged and the carboxyl group may dissociate and become negatively charged. So when you have one group positively charged and another one negatively charged when these charges are equal, so we call that as zwitterion form.

And so this one may exist in equilibrium with the dissociated, sorry protonated and unprotonated amino group and you can have this acts as a proton donor and this acts as a proton acceptor. So here both of them accepts, so this can take a proton and this can like for example if this is the one it can accept and become this. So they have both acid and base group on the same molecule and this actually affects their properties which is shown in the next to next slide.

(Refer Slide Time: 05:25)



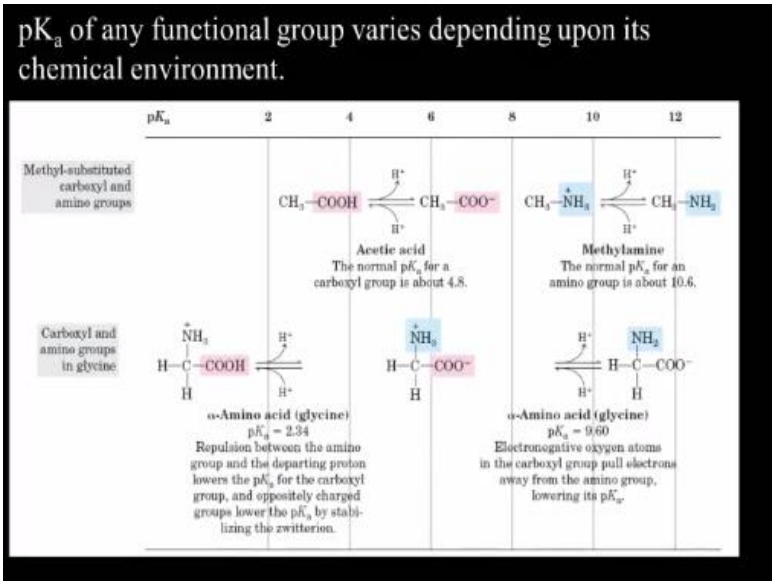
But this is an important characteristic that comes from this presence of two groups that can donate or accept a proton and therefore you have two pK a values. For example, the acetic acid or the carboxyl group has its pK a value for it the dissociation of the carboxyl group, so you have a pK a value. And similarly for the dissociation of this protonated amino group you have its own pK a value.

So therefore.,you have pK 1 and pK 2 and so it is going to behave like a weak acid for both the groups and since it is on the same molecules you get a peak like this when you do titration. So first when you reach near the pK 1 like for this carboxyl group then it resists pH change in this region where it is highlighted in blue here and then you get the corresponding the pH where it equals pK 1 and then once that is done then pH rapidly increases as you add more hydroxyl ions.

And once you are near the range in which the other group is going to dissociate then again it is going to resist pH change and the midpoint of that pH will be the second pK 2 and once that dissociation is also complete then the pH is going to go up. So you have two like a biphasic graph is what you get for these zwitterion molecules. The midpoint of that where you have positive and negative charges are equal that is when you call it as isoelectric pH.

So isoelectric pH is the pH at which the total charge is 0, so the molecule is in this form. So this applies to amino acids, but since proteins are multiple amino acids it applies to proteins as well.

(Refer Slide Time: 07:53)



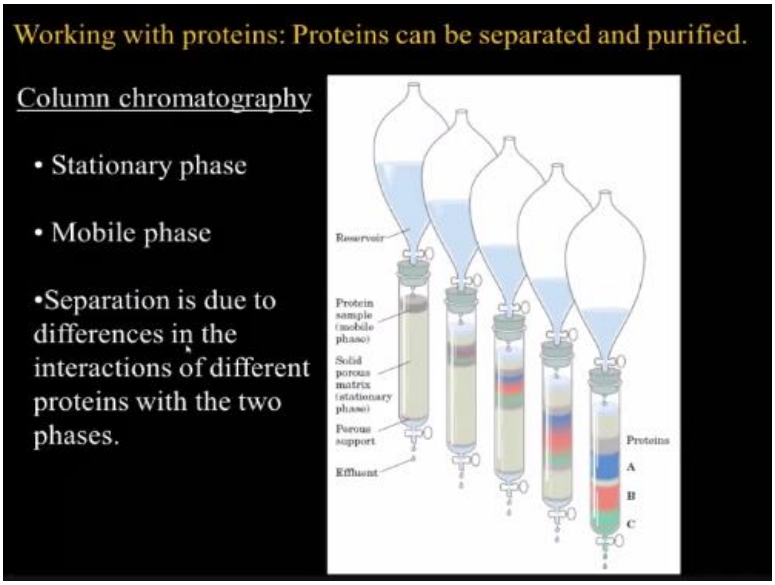
So now a small discussion on the consequences of the presence of two charged groups on the same molecule. So this I actually mentioned earlier and I asked you to do it as a homework and I am not sure you have figured it out or not but the answer is here in this slide. So the carboxyl group or as part of acetic acid, acetic acid has a pK a value which is 4.8 and if you look at the pK a for the same carboxyl group but in the context of glycine you have a much lower pK a value.

What it means is at much lower pH, this will readily dissociate already, in acidic condition itself it will start dissociating. The reason is because of this molecular environment. So in this case the leaving group the proton has a charge repulsion with this positively charged amino group, so that promotes the proton to readily leave one and second the product of the dissociation will be this negative charge and this negative charge and positive charge they stabilize each other.

So due to these two reasons, the dissociation is promoted here and as a result the pK a value is lower for the carboxyl group in glycine compared to the carboxyl group in acetic acid. And the same logic holds good for the amino group as well, so I will not get into the details but it is the same logic. So therefore, the concept here on the title pK a of any functional group varies depending upon its chemical environment.

So, this is something you need to remember whether you are going to learn biology or not but this is an important chemistry concept applicable in many different contexts.

(Refer Slide Time: 10:14)



So these are the few things that we generally discussed about amino acids and their properties, primarily it is all coming from having two charged groups on their structure and other than that we looked at the aromatic amino acid having the 280 nanometer characteristic absorbance and then the disulfide bond possibility. So having seen a few of these characteristics of proteins, now we will move on to a few techniques that help us to work with proteins.

You know proteins are really important because they are important for pretty much all biological activities because they are bio catalysts. So, every biochemical reaction is dependent on enzymes and therefore a whole lot of biology experiments actually deal with proteins. So therefore, we will learn the theoretical part of few protein-based techniques primarily focusing on protein separation like separating individual proteins from a mixture of proteins.

The first technique we are going to learn is column chromatography. So before I get into column chromatography, I want to briefly discuss about the basic principles behind the chromatography technique. So, first of all the word itself chromatography. So chroma, chroma refers to color, so it is primarily because this method was originally used to separate plant pigments, colored pigments and because it was used for separating colored pigments it got the name chromatography.

Graph means a drawing or a picture showing different things, here the different colors, display of different colors, so a graph of colors that is how chromatography the word theme

okay and later same technique applied for separating any kind of molecules are still being called the same thing chromatography. The basic principle of chromatography comes from an earlier technique called counter current distribution.

So counter current distribution depends on differences in the partition coefficient of different liquids, counter current distribution was originally used for separating liquids. For example you take two liquids that are immiscible like let us say you take butanol and the water. So if you take equal volume of butanol and water and if you mix them, initially some amount of water will mix with butanol like butanol will get hydrated.

I do not mean mixing like the emulsion are obtained, it really becoming a stable solution and that is a very, very small quantity of water, but once butanol is water saturated then if you add butanol, water saturated butanol and water and mix it and let the tube stay on the bench for some time the two liquids will separate. The less dense butanol will go to the top phase and water will be at the bottom, aqueous phase and butanol phase they will separate.

So now let us assume we have a solute dissolved in one of the two and the other one does not have the solvent. Let us say we take some nucleic acids, RNA purified from a biological material. So RNA being hydrophilic we already know it has the phosphoric acid group and that has one acid group that is free and it dissociates and as a result is negatively charged, so it is a charged molecule.

And we have polar groups on the nitrogenous base and then we have alcohol group on the sugar as well. So it is charged as well as polar, so it readily dissolves in water. Now if you have the RNA dissolved in water and then you add this water saturated butanol to it equal volume and mix them, vigorously mix them so that they form a completely emulsified. Now let it separate into water and butanol phase, now what will happen?

In the case of RNA since it is highly hydrophilic all of the RNA will dissolve in as stay in water and nothing will go into butanol. Now let us say actually in the RNA you have a mixture of solutes that have some solubility in butanol as well and some solubility in water and let us say that the solubility of the different solutes in this mixture vary, some more readily go into butanol than water than others.

So if you have that kind of a situation, now what will happen? If the entire mixture if you originally have it in water and if you add butanol and mix it something that likes to be in butanol would have after one separation, one cycle of this mixing and allowing the phase separation you will find more of that molecule let us say that is solute A. So solute A would have now gone more into butanol than in water.

And if you have solid B that prefers water more than butanol then you will after one round of this mixing and phase separation you will find the ratio of A, B having changed in the water and if you repeat it multiple times eventually what you will find is A would have mostly gone into butanol and B will be mostly in water. So the A to B ratio in water would have drastically decreased this is what we call as counter current distribution.

And this differing ability to dissolve is what you call as partition coefficient, partitions between the two solvents. So now the separation of the A and B is critically dependent on the number of times you mixed and allowed the phase separation and each one of that events we call as a theoretical plate. So one separation event is a one theoretical plate. So this number of separating events is very critical in properly resolving the two molecules.

So that is what we call as resolution here in terms of separation. So chromatography is exactly based on this principle. So here you saw you needed two solvents, water and butanol example I used, but it could be any two different solvents which are immiscible. So this two solvent thing in chromatography metamorphosis as stationary phase and mobile phase. So you have one face which is stationary that stays in the tube.

For example, here you see a cylindrical tube with a stopper at the bottom and you have a porous material packed here and that itself does not go through so you have a fit added and then it does not get disturbed because we have added another thing. Now you have a liquid that is going through. So if a set of molecules applied on the top, so in this cartoon you have a gravity phase to flow.

So this flow equals in the counter current distribution you are mixing and allowing the phase to separate. So instead of doing it in discrete events, here it is a continuous process. The mobile phase enters and goes through continuously, and if the molecules applied here if they

more readily partition into the mobile phase and they do not like the stationary phase so they can either dissolve like this is not purely solid.

Depending on the type of chromatography it may be simply charged solid or it may be hydrated solid where water will be there as a stationary phase to interact with. So in any case, if the applied mixture here like let us say molecule A now let us move to protein, let us say protein A more readily partitions into the mobile phase liquid that is flowing through, then adhering to this solid phase.

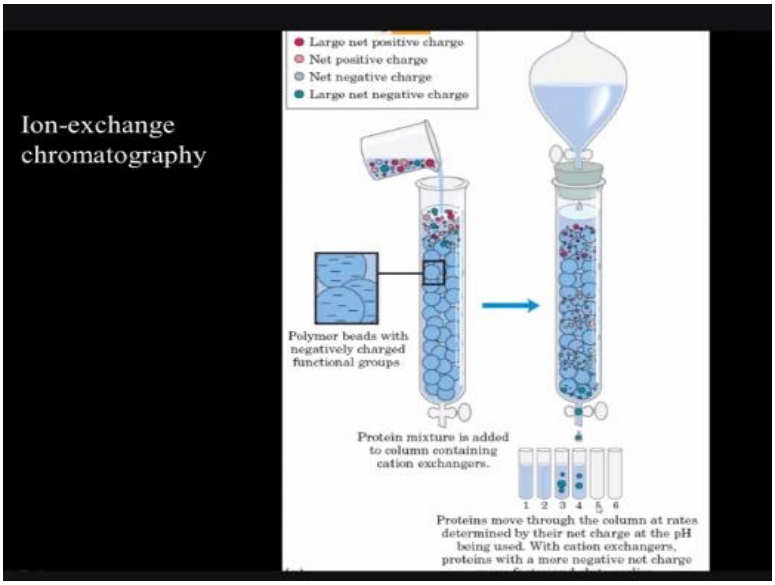
Then it would readily go through and it will come out faster compared to another one that prefers interaction with the solid phase over the mobile phase that will come through slowly. So the mixture as shown here are having different colors then they get separated and if you have a collection while where you collect fixed volume or for a fixed time and if that automatically rotates and moves to the next collection while and so on you get the fractions collected.

And then you can ask say individual fraction to see which fraction has your particular protein. So the ability to detect a particular protein based on its activity is what I called as by using the word as assay, you assay for a protein activity. So, this is the basic principle of all chromatography. So you have a stationary phase which is usually a porous solid matrix a resin, we will get into the details when we get into individual chromatography, then another one is a mobile phase.

So the differences in the partitioning between the two is the basic principle by which the molecules get separated. And to get a good resolution like a non-overlapping separation of this green and red and this blue is based on the number of separating events or number of theoretical plates that are happening during the process. So in this particular example you can imagine the longer the column and long longer this mobility, then more the separation.

But there are limits to how long you can have a column. So therefore, there is a maximum limit to the number of theoretical plates. So this is the basic principle of chromatography.

(Refer Slide Time: 22:48)



So now let us look at different types. The first one we have here is ion exchange chromatography. So let me check whether this is the first one, yeah yes. So in ion exchange chromatography we rely on differences in the net charges of the different protein molecules. So at a given pH one particular protein may be more positively charged than the other one or one may be more negatively charged than the other one.

So depending on that, they will get separated when it goes through a matrix that has charged groups in them and that is why this is called ion exchange chromatography. It is primarily the ions that are there that dissociate and associate with your protein or the opposite, so that is why it is an ion exchange. Oftentimes, we use this method even for purifying water. Water purification for the laboratory purpose, not your drinking water.

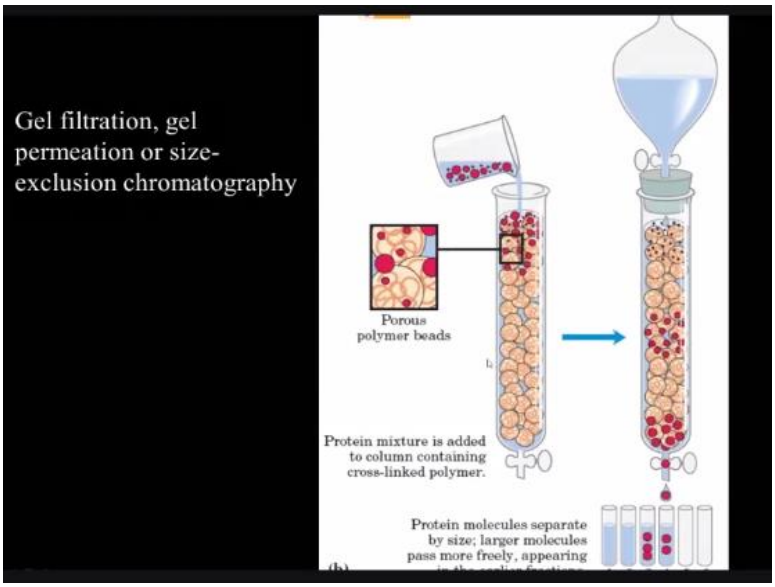
Drinking water is dependent on reverse osmosis where you are not interested in getting rid of all the charged salts present there, charged ions like sodium, magnesium and so on. Calcium is what you would normally like to remove in water because that gives hardness to water. So a very similar principle here. So here you have a mixture, you see the color codes here, you know a protein that has large net positive charge on net positive charge slightly less than the first one and negative net charge and large net negative charge.

So you have a range and then you pass them through the column. So now if the column has primarily a negative charge, then what is going to happen? The positively charged ones are going to stay there, you see the large net positive that is not have entered yet while this one

has made some distance through it while these two are moving out already. So this is going to come out first and this next and this next and this next.

So like that you can get it separated, So this is charge based separation. So here the partitioning is since the stationary phase is negatively charged, the positively charged molecules have preference, they partition better with the solid phase than with the mobile phase and as a result they come out last.

(Refer Slide Time: 25:42)



So the next technique is called gel filtration chromatography or gel permeation chromatography or size exclusion chromatography. So all the names will become very familiar once we go through this. So this is an extremely simple easy to perform and it is most widely used in protein purification. So another thing I want to mention is that not one technique is sufficient to totally purify a given protein, so these are used in combination.

For example, the purified fraction, a desirable fraction of ion exchange chromatography elevate may be loaded on a gel filtration chromatography for further separation. For example, you might have two different proteins of similar charge but different sizes, so they would have come here together, although their sizes vary but charge is very similar and as a result they come in same fraction.

Now you use a chromatography that separates them based on sizes, so first based on charge and then based on size, then you would have separated them. So they are often used in combination in a sequential or tandem manner. So here we have gravity as the main source

for separation. So here you have protein molecules that are globular in shape, so here shape matters.

If you have non-globular proteins, this method is not the choice, for that we have to resort to other methods. This method primarily is for globular proteins and most of the proteins we work with are globular except structural proteins like collagen or keratin or fibrin which are long fiber like structures, otherwise most of the cellular proteins particularly enzymes are globular in nature, so this method is widely used for separating them.

So here the solid matrix, in ionic change chromatography they had a particular charge in the example we saw it was negatively charged. Here instead these resins are going to have pores in them, each one is a sphere-shaped structure and within those sphere-shaped globular structure you have characteristic size pores in them. So you say 200 micron size, 50 micron size, 10 micron size, 100 micron size like that specific pore size resins are there.

So these are like resins. These particles themselves are very small fine particles. So when you pack them in a column, so you will have each of these beads having pores and a protein molecule that can readily enter the pore, meaning its diameter is smaller than the pore size then that will go into the pores and then come out and it will keep repeating that process from top to bottom of this chromatography column.

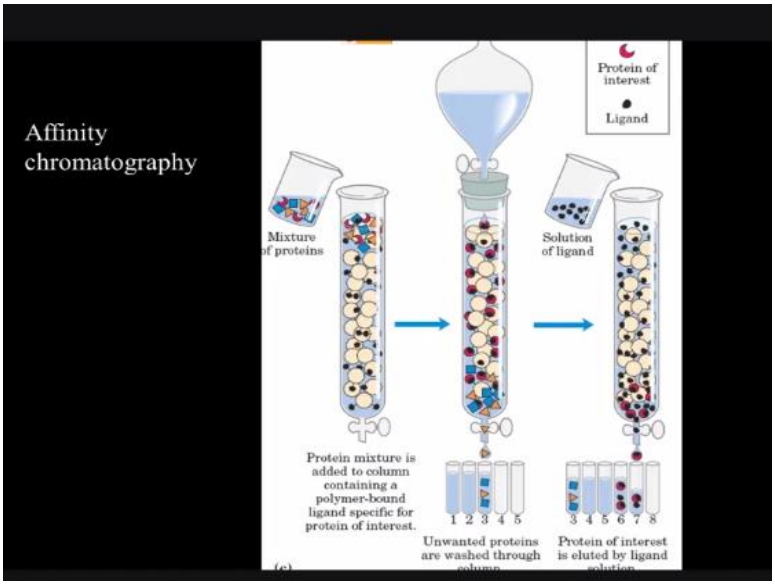
While a protein like this big red one whose diameter is bigger than the pore size, this will not be able to enter into the pore, as a result it will prefer only the mobile phase, so it will continuously readily go through the mobile phase. So it prefers, it partitions better with the mobile phase, its partition coefficient for mobile phase is higher than for the solid phase and that is in contrast to the smaller one.

So due to this with a certain number of theoretical plates, you will see this kind of separation. The biggest one comes out faster than the intermediate one which comes faster than the smaller ones. So now depending on your protein size, you select the right beads with the right pore size and that helps in this kind of separation. So, this sort of it is a filtration, you are filtering using the gel, so this matrix itself we call as gel.

Because they swell in water when you soak them and it does that therefore gel-like properties, so we informally call them as gel. So here you are using a gel to filter out, so therefore it is gel filtration and since it has a permeability like it permeates through the pores of these beads or of the gel so this is also called gel permeation. And here since the size-based exclusion.

Like for example this large protein is excluded from these beads or the solid phase. Therefore, it is a size-based exclusion so it is also called size exclusion chromatography. All three terms are continuously commonly used, so you should be familiar with any of them if you encounter, alright.

(Refer Slide Time: 31:28)



So in both these ion exchanges and gel filtration there are minor variations and we are not getting into it because this course is not a separation technique course, it is biochemistry and therefore it is only a preliminary introduction to the main concepts. So, the third major type of chromatography that we often encounter are called affinity chromatography procedures. So in this method, the solid phase has a certain affinity to the protein.

Based on what we have already learnt about characteristics of a given protein. For example if you are purifying an antigen for which antibodies are already available, then the antibody may be covalently linked to these beads, these resins beads or gel beads. So resins, gel, matrix or these are all commonly used terms to refer to the solid phase in chromatography techniques.

So if it is covalently linked and when you pass a mixture of proteins in which one of the proteins is your antigen and this antigen is going to bind to the antibody that is coupled to the column. So here you are not worried about the ion exchange property where you depend on the charge difference or the size difference and here the theoretical plates can really be increased a lot because the antigen antibody affinity is going to retain the protein gear for very long time duration.

During which the mobile phase continuously flows through, and as a result the other molecules that do not have affinity for this antibody can all be washed away. So due to this, this is a very very powerful technique where the column length can be really short and it is based on the biological interaction, affinity of one molecule for the other. Most of the time these affinities are very high.

For example the dissociation constant for an antigen and antibody binding ligand receptor interaction kind of thing is usually in nanomolar range. What I mean is in nanomolar concentration, the two molecules will already be 50% bound and 50% unbound, so that is what is the dissociation constant. The equilibrium constant K here will refer to what is the numerator divided by denominator.

And if it is nanomolar meaning in nanomolar concentration of the two molecules already 50% will be bound. So, when you have that sort of a high affinity and very high specificity as well. So an antigen that binds to a particular antibody, other antigens are not going to bind to it at all, and due to this huge difference in the partitioning between the solid phase and the mobile phase, this helps in really resolving, the resolving power of this technique is very high.

So you can very readily get rid of all other proteins and you can have really highly purified form of your protein of interest. So, I use antigen antibody as the example, but there are many other interacting molecules can be used for this purpose, ligand receptor in the case of signaling molecules and so on. So, there are variety of interactions that are known and they are often used for affinity purification.

(Refer Slide Time: 35:45)

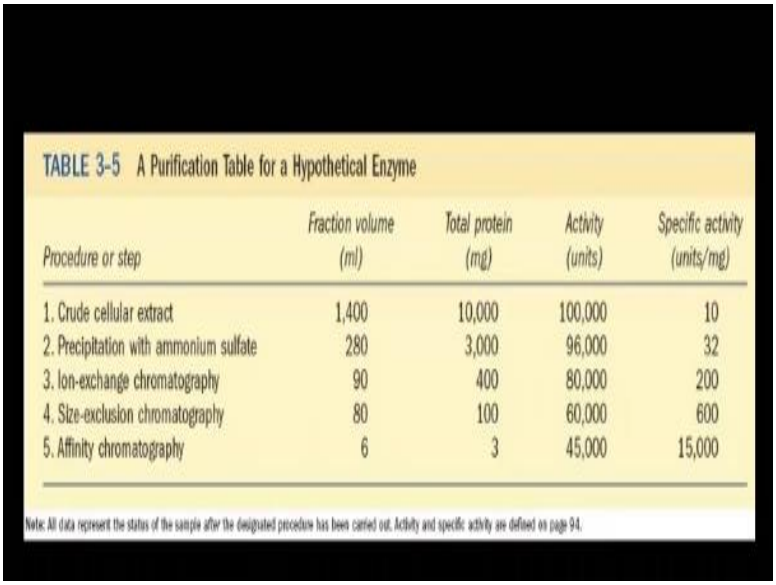


TABLE 3-5 A Purification Table for a Hypothetical Enzyme

Procedure or step	Fraction volume (ml)	Total protein (mg)	Activity (units)	Specific activity (units/mg)
1. Crude cellular extract	1,400	10,000	100,000	10
2. Precipitation with ammonium sulfate	280	3,000	96,000	32
3. Ion-exchange chromatography	90	400	80,000	200
4. Size-exclusion chromatography	80	100	60,000	600
5. Affinity chromatography	6	3	45,000	15,000

Note: All data represent the status of the sample after the designated procedure has been carried out. Activity and specific activity are defined on page 94.

So, these are the 3 major methods and there are variations which we are going to ignore and we know this much and these basic concepts the idea of partition coefficient what is counter current distribution, what is theoretical plate and the differences in the partition coefficient between molecules in a mixture for the mobile phase versus solid phase, then you have gotten the main idea.

So now let us look at a certain parameter that are often used for describing purification. So, let us say we are purifying an enzyme and that enzyme has a certain activity, meaning it converts certain moles of substrate per second, so that we call as its activity. So one more of the enzyme converts one mole of the substrate in a second would be a way to define enzyme activity. So now the crude cellular extract which is mixture of everything that was in the cell.

Your enzyme is also present but your enzyme is one of the many things that are there. So in the crude extract or let us say you are starting with 1.4 liters and then you have about 10 grams of protein and in that the entire amount of your protein is also part of the 10 gram and you have let us say 100,000 units but when you divide by the number of milligrams of protein then you get 10 meaning per milligram you have 10 units of your enzyme there.

So now when you go through the purification when you are getting rid of the other ones. let us take an ideal scenario, in ideal scenario from this 10,000 milligrams or 10 grams you have gotten rid of all other proteins and then you have pure your single protein. Now what will happen? The specific activity would have increased, let us say here you have 10 units per milligram because the total of 10, 000 milligrams had 100,000 units.

Now let us say this 10 milligram is actually your protein, so now you will find at the end if there is absolutely no loss of your protein or reduction in protein activity during the process of purification, then you will retain the entire 100,000 units activity in that 10 milligram because after all that 10 milligram present here only contributed to this activity because other molecules are not catalyzing this particular reaction.

So what will happen this 100,000 would have come from the 10 milligrams, so 100,000 divided by 10, the specific activity will be 10.000. So, it would have increased from 10 to 10,000. So that shows you as you are going to in purify, the specific activity is going to increase. So, the increase in specific activity is an indication that you are actually enriching the fraction for your particular protein.

But in reality, you do not get that ideal 100% purification that is 100% recovery of your protein of interest and that without losing any activity at all without any loss of activity that does not happen, so therefore we use this as a parameter to measure the efficiency of our purification methods. So if a purification method gives very high specific activity, then you know that that is really good.

So the first step here is the precipitation with ammonium sulphate, we did not go into the details of how ammonium sulphate helped here, so do not worry about it that is something we will not have time to get into it. So basically if you add ammonium sulphate to a solution of proteins since this is a charged molecule, it is going to more readily interact with water than the proteins dissolved.

So depending on the solubility of different proteins as the ammonium sulfate concentration increases, these proteins will precipitate out of the solution. So the percentage of ammonium sulphate at which a given protein gets precipitated is a way to initially purify proteins. So this is the initial step before getting into chromatography. So this is often called salting out. You may have learnt in your chemistry class already.

So after that step, the fraction comes down because now from large volume you have precipitated and you are taking the precipitating in a new solvent, so the volume has significantly come down and you have removed a whole lot of proteins that did not

precipitate at a certain percentage of ammonium sulfate, so the protein amount has come down, but you have some loss of activity because some of your protein actually remained in solution after that ammonium sulphate precipitation step.

So you have some loss, so you are not recovering the way ideal scenario I was first mentioning. So you end up getting increasing specific activity to some extent and this process gets repeated as you go through the next step. So the ammonium sulphate percentage at which most of your protein precipitated, you know most of it based on a minimal loss in the activity, total activity and that fraction now you applied to the ionic change column.

Then the volume comes down because in a smaller fraction you are able to collect most of the activity. And now more proteins have been removed, therefore the total protein comes down with some more loss of activity, but specific idea has increased. What it means is in this 400 milligrams you have more of your protein than what was there in these 10,000 milligrams. Although some amount is lost, this is purer, meaning this is more enriched for your protein.

And then you take that and put in size exclusion, now you have gotten three-fourth of the protein out but then you have lost some more, but the specific activity has significantly increased for this loss and when you repeat this step like this finally you get to really small quantity of protein that is mainly your protein and as a result specific activity has significantly increased.

So this is the way we kind of monitor how the purification scheme is progressing, alright. So I will stop here today and we will continue on other separation techniques in the next class.