Genetic Engineering: Theory and Applications
Professor Vishal Trivedi
Department of Bioscience and Bioengineering
Indian Institute of Technology Guwahati
Module 7 Isolation and Purification of Product Part 1
Lecture 20
Basic of Chromatography Part 1

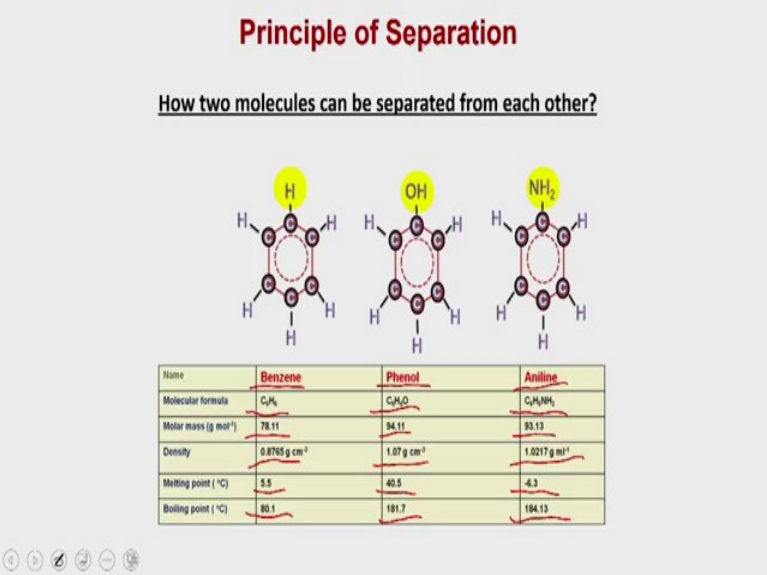
Hello everyone. This is Dr. Vishal Trivedi from Department of bioscience and bioengineering IIT Guwahati. And in what we have discussed so far we have discussed about the extraction of the product from the cells which are over expressing your protein. So in the previous lecture we have discussed about how to break these cells with the different methods.

So we have discussed about the physical methods we have discussed about the chemical method the enzymatic method and as well as at the end of the previous lecture we have discussed about the mechanical process. So all these processes are actually going to release the content which are being intracellular and then the product or the protein which you are going to over Express into these genetically modified organism is released and now the next step is to purify these proteins from the complex mixtures. So once the product is released from the cells which are over expressing the protein you have different types of proteins for you.

So most predominant protein could be the protein which you would like to gratify but the protein of other proteins such as protein from the host will also be in the release in this process and those proteins will going to be make the task very very difficult. And in that process you have to ensure that you will get the your desired protein but at the same time you will going to remove the other protein. So your protein which you are would like to purify is called as the desired protein the protein which are being given from the host are called the impurities are contaminating proteins. So in today's lecture we are going to discuss about how to purify or how to isolate the protein of your interest.

So this process is being extensively being done simply by employing the chromatography techniques but before getting into the details of chromatography we are going to discuss about the basic principle of separation. Let us assume that for explaining you the basic principle of separation we have taken an example of three molecules and we have taken a very simple molecule of the organic molecules.

(Refer Slide Time: 3:15)



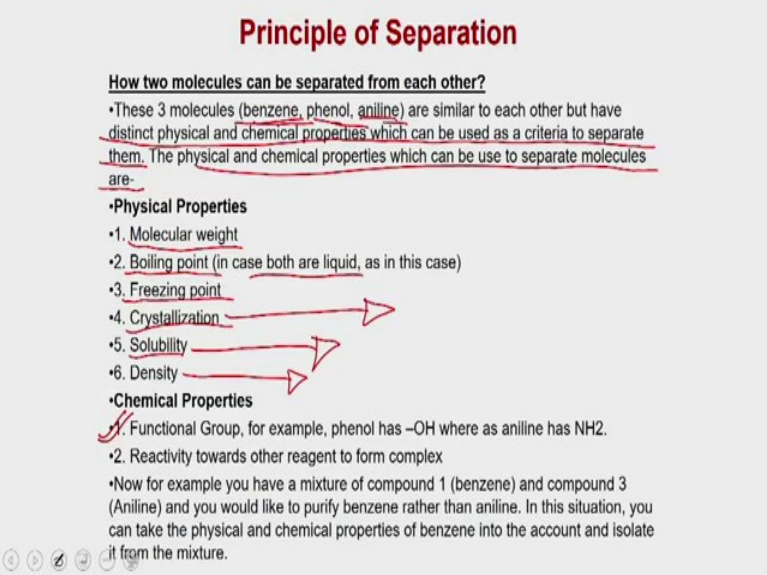
Name	Benzene	Phenol	Aniline
Molecular formula	C_6H_6	C_6H_6O	$C_6H_5NH_2$
Molar mass (g mol^{-1})	78.11	94.11	93.13
Density	0.8765 g cm^{-3}	1.07 g cm^{-3}	1.0217 g ml^{-1}
Melting point (°C)	5.5	40.5	-6.3
Boiling point (°C)	80.1	181.7	184.13

So the molecule number one is known as the benzene, the molecule number two is a phenol and molecule number three is an aniline and imagine that you have a mixture of these three bio organic molecules. First thing what you have to do is if you would like to separate these molecules you have to understand the characteristic of these molecule. Let us see what other characteristics are present in these molecules. So the molecular formula of benzene is C6H6. The molecular formula of phenol is C6H6O. And the molecular formula of an aniline is C6 H5NH2. The mass or the molecular weight of these three molecules are 78 94 and 93.

Similarly the density of the benzene is 0.87 grams per centimeter cube, Phenol 1.07 and the aniline is 1.02. Now melting point melting point is 5.5 degrees Celsius whereas in the case of phenol it is 40.5 and for aniline it is minus 6.3. And the boiling point is 80.1, 181.7 and 184.13.

So what you can see is all these three molecules are varying in terms of the many of the physical properties as well as the chemical properties.

(Refer Slide Time: 4:42)



And this is actually the basis of how to separate the molecules. So we would like to separate the molecules what you are supposed to do is you have to know their physical properties chemical properties as well as the biological properties as long as we are talking about the protein as well.

So that we have taken an example of three molecules that is benzene phenol and aniline and they are looked like very similar to each other but they have the distinct physical and chemical properties which could be used as a criteria to separate them. The physical and the chemical properties which can be used to separate the molecules are, Molecular weight. So if the proteins are if the molecules are of different molecular weight you can employ or you can exploit that criteria to purify that. I would like to have a molecule weight I have I would like to purify a molecule which is of molecular weight of 40 and that actually is going to exclude all the molecules which are above two that or lower to that. So that kind of criteria also can be used.

Similarly you can use the boiling point in case a mulligan can be used only if you are who would like to purify the two molecule which are liquid in nature. Similarly we had the freezing point Crystallization, solubility for the Crystallization is one of the criteria or the physical property which people were using very often to purify.

So what you can do is suppose you have a complex mixture. What you can do is you can let the you can bring the conditions which is after crystallization conditions of the one molecule. So what will happen is if you take the liquid which is actually a mixture of the different

compounds and then you if you bring the Crystallization condition of one of the compound that one compound is going to be crystallized and that can product you can purify simply by filtration and the rest molecule will get filtered out and then you can wash this crystalize product and you can recrystallize again and that actually will give you the very very high purity compounds. This is exactly what people do when they are producing some of the very very high quality chemical compounds.

One of the classical example if sugar itself. So what sugar. What people are doing is they are getting the sugar cane juice. And what they are simply doing is they are just bringing the cancellation conditions of the sugar and that is how they are actually making the sugar from the sugar cane solution or sugar cane juice actually. Similarly you can do a crystallization of the different organic compound because all the organic compounds are having the different cancellation conditions so you if you vary those conditions you could be able to crystallize your compound off your interest and that is how you can actually get that compound from the mixture. Similarly you can use the solubility, solubility and crystallization are very similar properties.

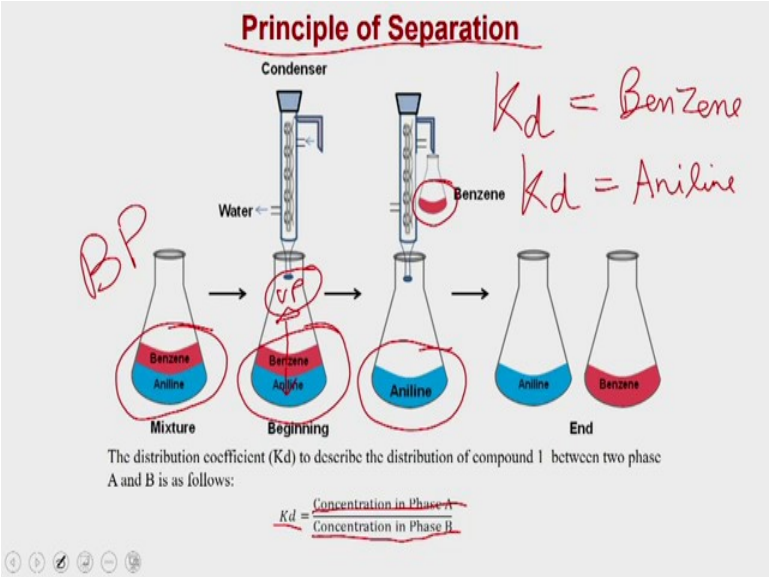
The molecule which is going to crystallize in a particular condition is also going to have the different solubility so you can also use the solubility as a criteria for example if some molecule is more soluble in water but less soluble in benzene you can just simply dump that particular complex mixture into the water and as a result what will happen is all other molecules are going to be not soluble so they will make the different layer where as your compound will get fascinated and it will get go into the aqueous layer and that is how you can actually be able to purify.

If you remember when we were doing the plasmid isolation. That time we were doing the the phenol chloroform treatment and in that phenol chloroform treatment you were getting the two layer one is Aqueous layer and other one was the phenol layer and the DNA which is actually the polar molecules it was migrating to aqueous layer whereas the protein and all other molecules they were been going into the phenol layer and that is how you are actually separating the DNA from the protein as well as the other contaminating molecules. So that can be used as a criteria to purify the molecule as well. Then you can use the density you can use density also then sometime when the physical properties are not good enough people are also using the chemical properties for example the functional group, If you see all these three

molecules benzene does not contain any functional group whereas phenol contains a functional group of OH and an aniline contains the functional group of NH2.

So once you have the functional groups actually you can exploit the functional groups because these functional group could have a differential reactivity and that differential reactivity could bring the some kind of alteration in terms of its physical properties. For example if you if you make a complex of make some compound that actually will reduce its solubility then what will happen is that particular compound is going to be precipitated if you add that reagent and the other two molecule are going to remain in the solution and that is how you can be able to purify remaining compounds. And at the end we also can explode the biological properties for the biological properties are mostly been associated with either the DNA or the proteins which are biological in nature the molecules that we will discuss anyway in detail in the subsequent lectures.

(Refer Slide Time: 10:24)



So let us see a clear example how you can do a separation. For example if you have a mixture of benzene and aniline so you have a complete mixture of benzene aniline and you are using the boiling point as a criteria.

So if you remember the benzene and Aniline are having the different boiling point. So what we have done is we have started putting it into a condensation reactions or condensation apparatus and we started heating up this liquid and as soon as we start heating up the both the benzene as well as the aniline is going to form the vapour phase so that benzene is going to be vaccinated into the two phases one is called as the vapor phase. The other one is called as

the liquid phase. So the both molecules the benzene as well as the aniline is going to distribute themselves into the two phases. One is called as the vapor phase and the other one is called as a liquid phase.

So in the beginning the both molecules are going to distribute themselves into the two phases but, as the process will continue because the benzene is going to have the different boiling point the benzene will prefer to be more into the vapor phase whereas the aniline is going to be preferred more into the Aqua phase.

So as a result what will happen is that the benzene will prefer to be into the vapor phase and it will come into the condenser and then within the condenser it is going to be condensed and then you can collect the benzene into the separate flask whereas the aniline the mostly aniline is going to be remain into the liquid phase as long as you are keeping the temperature up to the boiling point of benzene and at that point when you are bringing it to the boiling point of benzene the more and more amount of benzene is going to be collected into the other flask.

So if you continue this process in the beginning there will be a mixture of benzene aniline in the top flask as well. But if you continue this process the amount of analine is going to be reduced in a subsequent distillations. And at the end what you are going to have is the pure benzene into the top flask and the pure aniline into the lower flask.

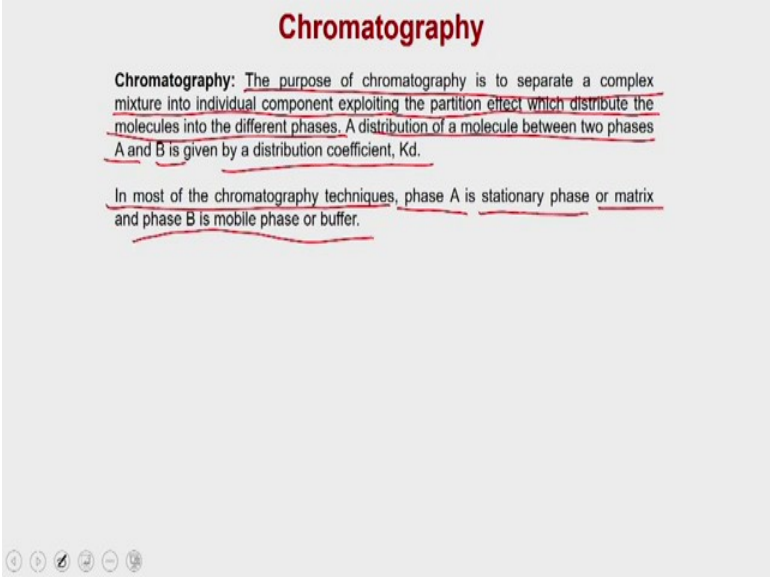
And this phenomena of distributing the molecule into the two phases is actually the basic principle of separation and a parameter which actually measures the distribution of the two molecules or distribution of molecules into the different phases is called as the distribution coefficient. So the formula for distribution coefficient is that the KD is equivalent to the concentration of the molecule in Phase A versus the molecule in of the constellation in phase B so in this particular case you are going to have 2 KD molecule 2 KD values 1 for benzene and the other one is for the aniline, so aniline is going to have the separate KD values and the benzene is going to be the separate KD values.

And because of that the KD values for benzene would be more towards the vapor phase whereas the KD values for aniline would be more towards the liquid phase and as a result the benzene will go and get distilled into separate flask whereas the aniline will get into remain into the original flask. So this is actually the basic principle of separation. But as long as the protein is concerned you cannot use the boiling point or melting point and all and any such criteria to separate the protein because the proteins are very very delicate molecules or they

are gentle molecules they are actually having a defined predefined three dimensional structures.

So the benzene… The protein has to be a protein or the molecules of biological in nature which actually require some kind of activity and three dimensional structures and all those kind of parameters. Those molecule has to be purified in a more sophisticated techniques and those technique is known as the chromatography.

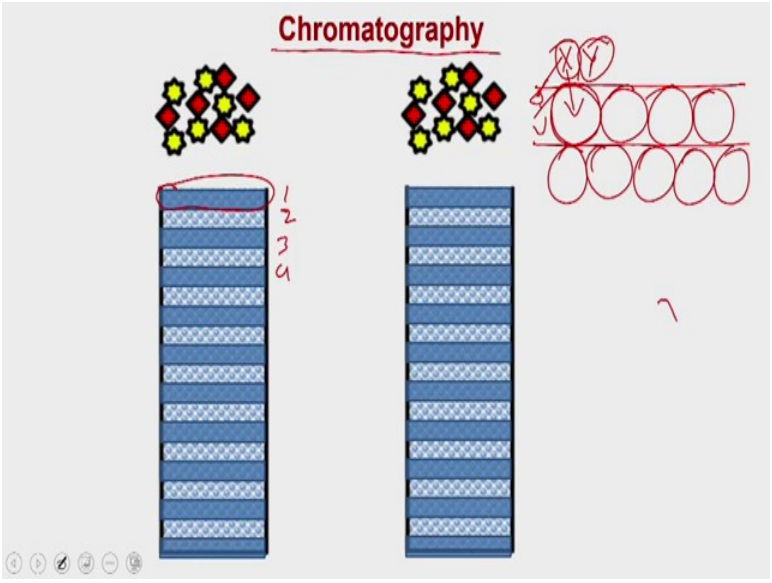
(Refer Slide Time: 14:45)



So in a chromatography the purpose of the chromatography is to separate a complex mixture into the individual component exploiting the partition effects which distributes the molecule into the different phases. So the basic principle of separation remains the same except that the process or the way in which you distribute the molecule into the two phases are different. As in chromatography compared to the boiling point or melting point and all other kind of those harsh harsh processes.

The distribution of a molecule between the two phases a and b is given by a distribution coefficient KD if you remember in a previous slide the KD is the concentration of the molecule in Phase A divided by concentration of the molecule in phase B, but compared to that in most of the chromatography technique the phase A is considered to be a stationary phase or the mattress whereas the phase B is the mobile phase or the buffer which you use for the purification.

(Refer Slide Time: 15:57)



Now let us understand how the distribution or the separation techniques works in the case of chromatography techniques or chromatography procedures. So in most of chromatography procedures or the column chromatography procedures you are actually filling a column with a bead or some metrics so metrics are nothing but the beats. For example these beats could be of silica beats or it could be of agarose beadsbeats or it could be some very ferolose beats or you can imagine that some circular beats like are being filled and that actually is making a column of a chromatography technique. Now this beats are actually been arranged in this fashion and as a result the beats are actually making a layer after layer. So these beats are sitting on to each other and as a result they are actually making a multiple layers. And what happened is that once the molecule X or Y is interacting with these beats they are actually experiencing the distribution.

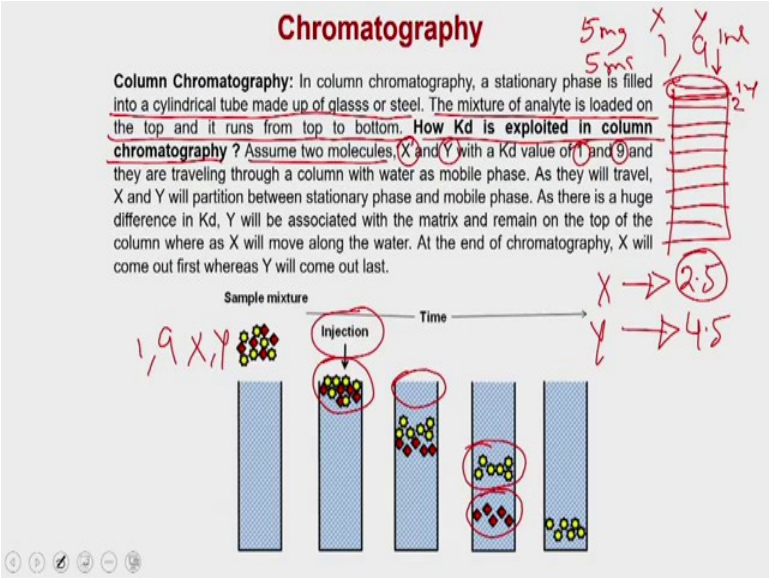
So as soon as the X or the Y or Z or ABC or any molecule where it is interacting with these beats because the beats are not alone they are also having the water. So the X which is actually as soon as X will interact with the first layer of beats. It is actually going to distribute themselves between the aqueous phase and the beat phase but so depending on their distribution they will distribute themselves towards the beat phase versus the aqueous phase and that actually will continue from the first layer. So you can imagine that the beats are having the first layer.

Then this is the second layer third layer fourth layer so you have one layer second layer third layer fourth layer like that and when you flow the water from the top the molecules which are present into the into the Layer number one are again going to distribute with the beats which

are present under Layer number two. That is why. And that will continue until that distribution will going to make the separation between the two molecules and.

That actually is going to make wider separation between the two molecule because the molecules are going to experience not only one distribution or one KD values they are going to experience multiple KD values depending on how many layers of these beats are being packed in that particular column. And as a result that is why if you have hundreds of molecules these hundreds of molecules are going to experience hundreds of KD values or hundreds of distributions and as a result you will see that a molecule A is being separated from the molecule B. Let us see how this is being done in the real value or real real way.

(Refer Slide Time: 19:06)



Now imagine that you have a column okay where you have filled the stationary phase is filled into a cylindrical tube made up of a glass of steel the matrics mixture of analyte is loaded onto the top and it runs from top to bottom. Now let us see how the KD is exploited in a column chromatography. Assume two molecules X and Y with a KD value of 1 and 9 is being loaded onto this. So you have two molecules which which are having a KD value of 1 and 9 against this particular column. Okay.

Now if they are traveling through this column as to what is mean by 1 And 9 is that if7 they are travelling in this column. In the... suppose you make the different layers of these beats and I assume that every beat or every beat volume is one ml actually. So what will happen is that suppose you loaded five mg of the X. And five mg of Y. Okay. And you can see that X is having a lower KD values whereas the Y is having the higher KD values which means that

the X is preferring to be remain with the acquiesces whereas Y is preferring or Y is distributing themselves to the metric phase.

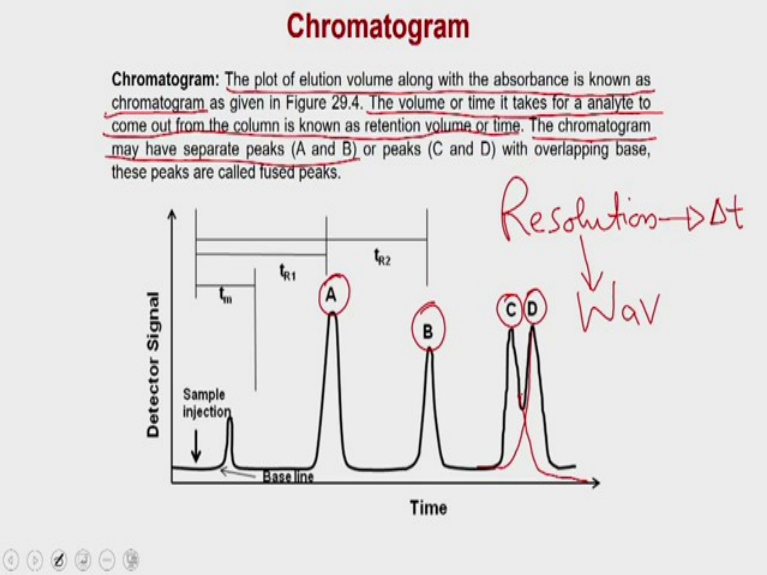
So as soon as they will interact with the first layer of metrics and you have flown one ml of buffer what will happen is that they will distribute this 5 milligrams which is of X and Y according to their distribution plane according to the distribution coefficient. So what will is that X. What is the distribution. One right for X is going to distribute between 50 50. One is means the KD value of one means that the concentration of material in the matrix verses the contradiction of material into the aqua phase is going to be equal at any any distributions.

This means in the first layer you are going to have 2.5 mg of X and a 2.5 mg of X will go into the aqua phase. Similarly for y... It is actually be distributed in terms of 4.5 versus 0.5 because the query value is nine. This means the vi the 4.5 mg of y will be bound to the first layer of your beats or the first layer of the Beats whereas the 2.5 mg of X is going to bind to that particular beat. Now if you follow one more ml what will happen is that the two 1.25 mg of X is going to be transferred to the layer number two. That is from the X.

And the 1.25 from the aqueous is going to be transferred into the first layer. So because of that if you see that the X is actually preferentially going towards the aqua phase which means more and more X is going into the aqua phase and you know that aqua phase is running faster than the matrix because matrix is a stationary.

So the Y will remain that and so if you if you start in the beginning the X and Y will remain on to the top but once you run this column for a longer period of time the Y will go and bind to the beats and that is how they will be they will be going to be remain on top of the column whereas the X proteins or X molecules are going to travel faster with the aqua phase and that is how the first your X will come out and subsequently the Y will come out and that is why the KD values are actually distributing the molecule between the two phases and that is how they are actually controlling the overall separation. Now if you see how and monitor this particular type of illusion of these molecule from the column you will get a pattern which is called as the chromatogram.

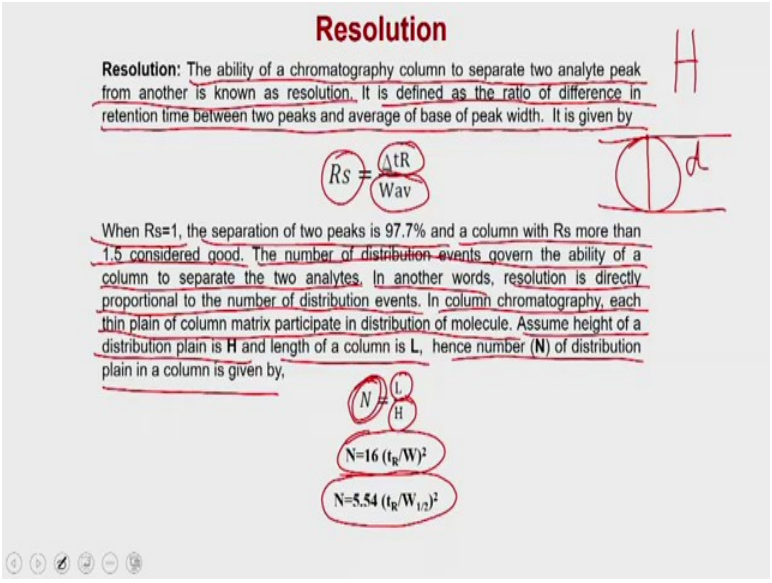
(Refer Slide Time: 23:46)



So what is the chromatogram the plat of illusion volume along with the observers are you can use any parameter is known as the chromatogram. Which is given in this particular figure the volume or time it takes for an analyte to come out from the column is known as the retention volume or the retention time. The chromatogram may have separate peaks for example in this case the peak A and B are the two separate peaks which are not having the overlapping base whereas you can have the peaks which are having the overlapping base which is actually the C and D. So what you see is the C is actually if the D is not present the C is illuting in this way where as the D is eluting in this way. That is why since they are actually sharing the base that is why these two peaks are not being separated.

So the column is been able to separate A and B whereas the column could not be able to separate C and D. So what actually determine or what actually can be used as a criteria to see whether my column is good or bad so the ability of a column to separate the two molecule is known as the resolution And. The resolution is depending on two parmeter. One is called as the distances between the two peaks which means the peak between A and B and. So it is actually dependent on the Delta T which means actually the distances between the two peak whereas the resolution is also dependent on the average width of the two peaks which are neighboring peaks.

(Refer Slide Time: 25:40)



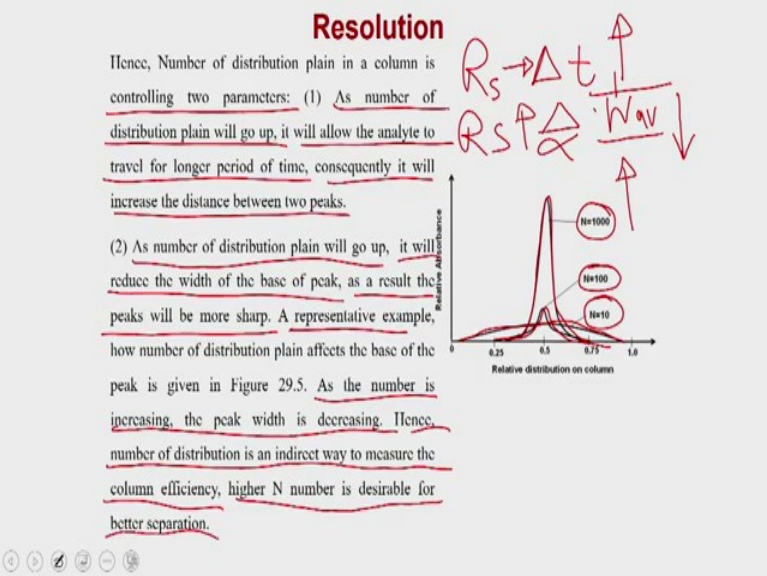
So the ability of what is resolution ability of a chromatography column to analyte or separate two analyte peak from another is known as a resolution it is defined as the ratio of differences in the retention time between the two peak and average of the base of the peak. Which means that evolution is directly proportional to Delta TR versus Wav. If you have resolution value of one you would be able to see ninety seven percent separation of two peaks. But if you have a resolution value or resolution value of 1.5 that is considered good because that is going to give you the two separate peaks or that will give you the separation of more than ninety nine percent.

The number of distributions for what it actually is governing the resolution. The resolution is being governed completely by the number of distribution planes because if you have the number of distributions it is going to having the higher chances of separating the two molecules into two separate peaks. That is why the number of distribution events govern the ability of a column to separate the two analyte. In another words that resolution is directly proportional to the number of distribution event in column chromatography. Each tin plane of column metrix participate in distribution of a molecule. Now you can assume that the. If you have a bead and its diameter is D.

This means the individual. Planes which are going to make the distribution is actually of the D height. So if you have a column of H. Height you can actually be able to calculate the total therotically how many number of distribution planes are present. So you can imagine that if you if the height of a distribution plane is H which means the diameter of these beads and the length of a column is l the number of distribution plain in a column is given by the N equal to

l by H. And if you solve this you can get the N equal to 16 TR by W whole squared or N is equal to 5.54 TR by W halfwhole Square which means that N is going to control the resolution and how you could be able to how at the N is controlling the resolution to N is controlling the resolution in two ways.

(Refer Slide Time: 28.20)



It is controlling the two parameter as the number of distribution playing will go up. It will allow that analytic travel for longer period of time. Consequently it will going to increase the distances between the two peak which means it is actually going to increase the delta T between the two peaks. If you increase the delta between the two peak you are actually going to increase the resolution because resolution is directly proportional to the Delta T.

Now as the number of distribution plane will go up it will going to reduce the width of the base of a peak. As a result the peak will become more sharp which means it is actually going lower down the W. W of a peak which means it is going to make the peak with lower which means eventually it is going to lower down the b Wav between the two peaks which means. It is going to increase to resolution because if you reduce this you are going to increase this because the resolution is inversely proportional to t one by aW. Now you can see how the number of distribution plane is affecting the the peak width okay. Now you can see in this figure when you have the number of diffusion plane to 10 you are not going to get a peak. In Fact you are going to get a flat peak or flat plate actually.

If you increase the number N equal to hundred you are going to see a peak which is very very having vary with the very big base, but if you increase N to the thousand you will see a

decent peak with a narrow base which means if you increase N you are actually going to reduce the base of width and as a result the peaks are going to be sharper and sharper. As the number is increasing the peak width is decreasing, hence the number of distribution in an indirect way to measure the column efficiency higher the number is desirable for the better separation. So this is all about the separation of the two molecules under the chromatography so with this we would like to conclude our lecture here thank you.