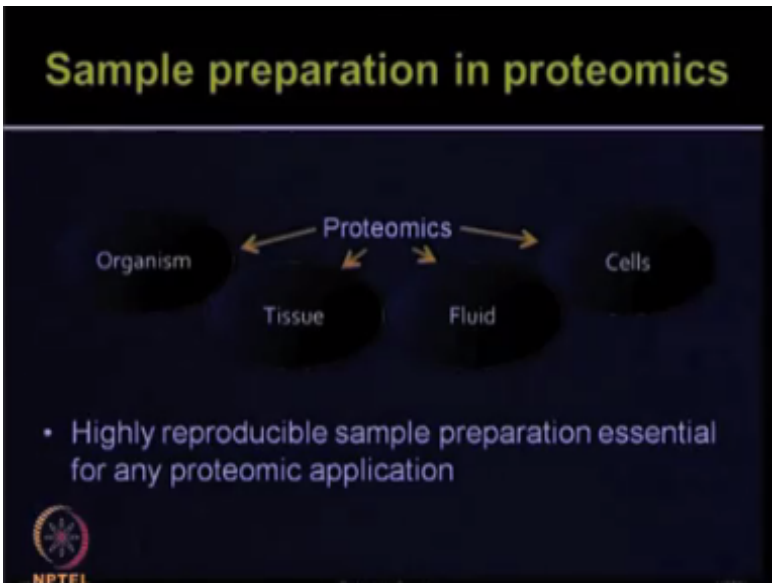
Introduction to Proteomics
Dr. Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology – Bombay

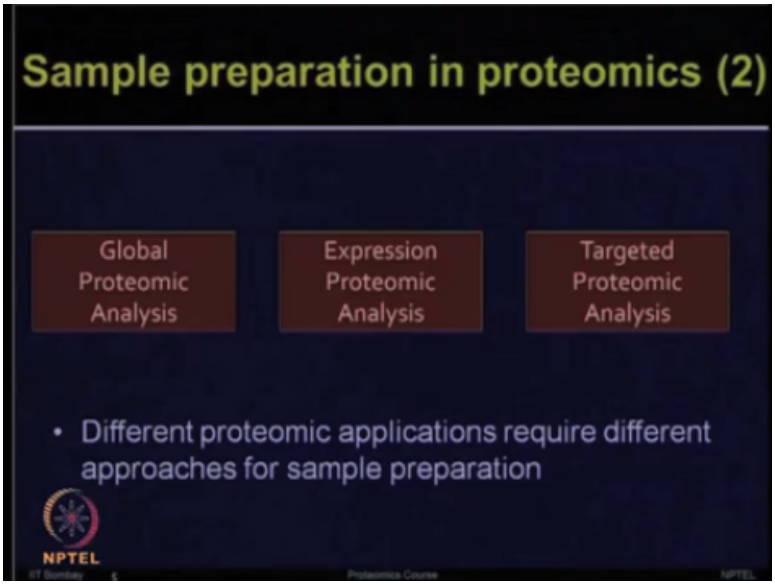
Lecture – 34
TMT: In vitro labeling

(Refer Slide Time: 00:36)



So, the proteome is very complex whether you want to perform your proteomic analysis for the whole organism, it means you want to know all the proteins present in a given organism or in a tissue or in body fluids or in different type of cells. Proteomics can be global or it can be very targeted or expression based, so very highly reproducible samples is very important for performing comparative proteomic analysis.

(Refer Slide Time: 01:26)



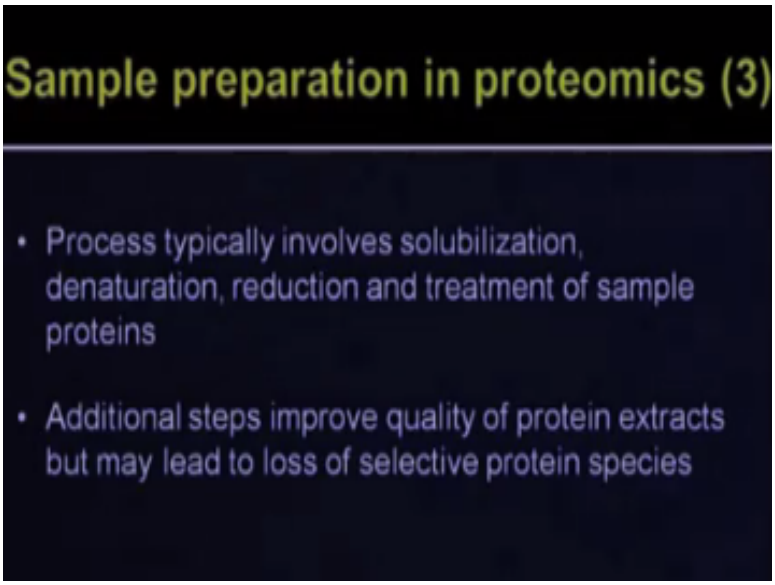
If you want to know the difference in your sample as compared to the controls, you need to ensure that your sample preparation is very reproducible. If you introduce some artifacts to begin with then obviously, you are not going to identify the reproducible biological changes. So, let me give you 3 different terminologies here for proteomic analysis; one is global proteome analysis expression, proteome analysis and targeted proteome analysis.

When I am talking about global proteome analysis, it means your aim is to characterize all the proteins present in the given sample. Expression proteome analysis, it means you are mainly interested to look for those changes, which are due to any chemical or your treatment, those are induced either going up or down the protein amount is changing. So, the protein expression analysis that is most commonly used for various type of clinical and different studies.

Targeted proteomic analysis, if you are very focused for a given organelle or a given sample type often, you would like to know what is happening in that particular proteome for example, mitochondrial proteome, so one need to try different type of strategies when thinking about performing a sample preparation. What is your objective, whether you want to do global profiling or you want to do expression profiling?

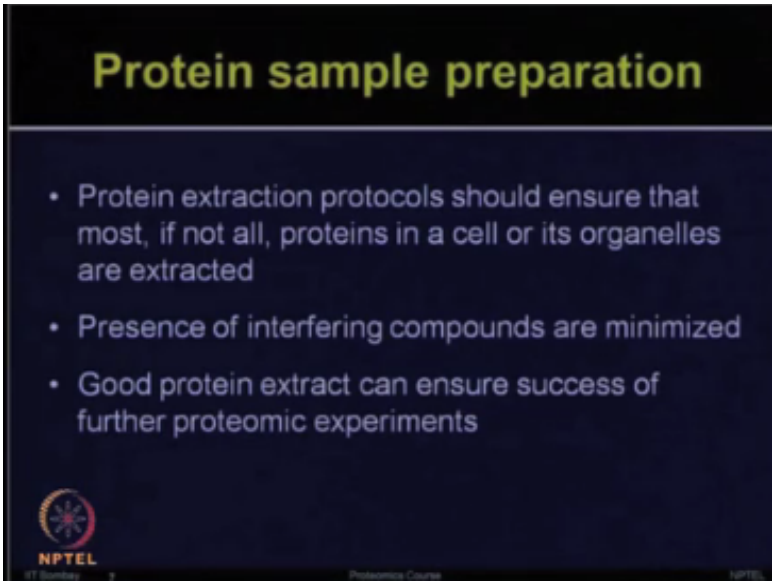
In either case, you need to extract all the proteins present in that particular target sample. Now, when you are looking at targeted proteome analysis, you just want to pre fractionator your sample in such a way that only that particular component is isolated and then all the proteins from that organelle or cell is being extracted. So, different types of strategies need to be used to perform this type of proteomic analysis.

(Refer Slide Time: 03:26)



Now, all of this sample processing involves solubilisation, denaturation, reduction and treatment of sample proteins but you need to involve additional steps depending upon the type of samples and your type of objective, so that the protein quality; the protein extract can be improved and while you are doing this, you have to be very cautious that when you are performing various steps and sequential type of extraction, you may also lose a small fraction of the proteins.

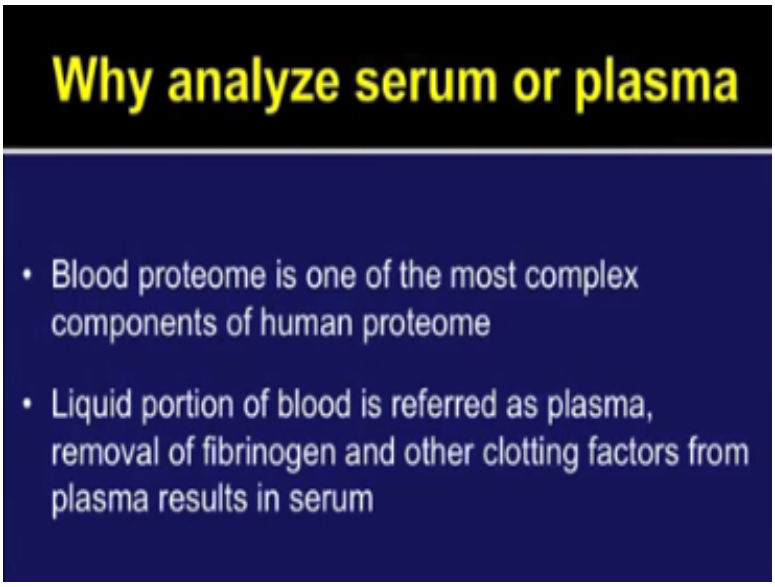
(Refer Slide Time: 04:13)



So, one has to be careful when adding various additional steps during the sample preparation. Now, protein extraction protocols they need to ensure that most; if not all the proteins in a cell or its organelle are extracted. The presence of interfering compounds should be minimized, so if you have optimized a very good protein extraction procedure that should ensure that you have a

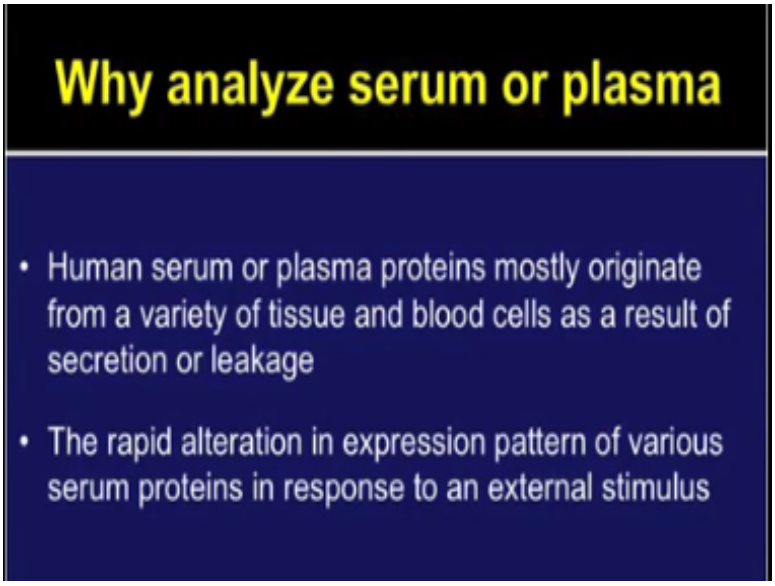
very wide proteome coverage and that is ultimately going to determine the success of your proteomic experiment.

(Refer Slide Time: 04:49)



So, first of all you may ask, why to analyze serum or plasma for any proteomic applications, so as you know blood proteome is one of the most complicated components of the human proteome. The liquid portion of the blood is referred to as plasma and removal of fibrinogen as well as other clotting factors from the plasma result into serum.

(Refer Slide Time: 05:13)



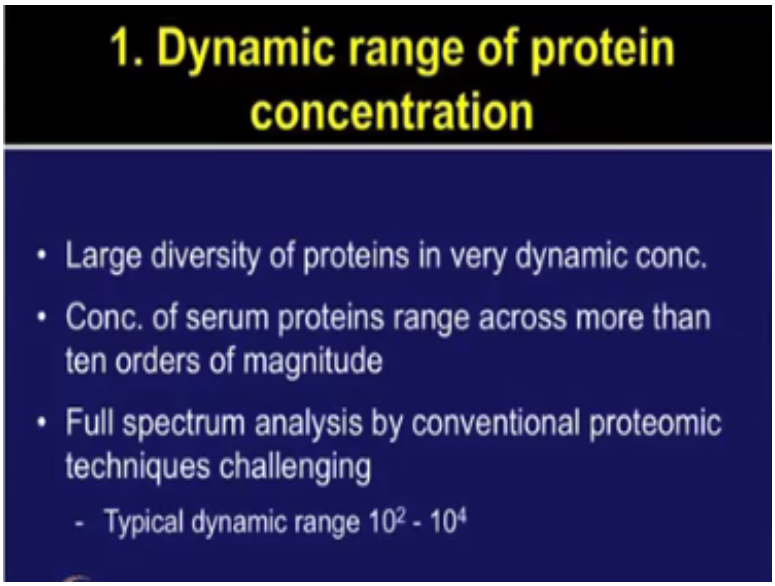
So, human serum or plasma proteins they mostly originate from a variety of tissue and blood cells as a result of secretion or leakage from the neighbouring tissue or the blood cells. The rapid alteration in the expression pattern of various serum proteins due to response of a diseased

condition or an external stimulus is true reflection of physiological changes occurring in an individual.

So, to get a feel about what are all the physiological changes happening in a patient due to a disease, people analyze serum or plasma proteome very oftenly, because blood removal for various type of tests is being performed in clinics, so blood is very easily accessible sample and performing the serum hot plasma proteome becomes very easy as compared to dissecting out a tissue for further analysis.

Although, sample removal is really but sample analysis; the serum or proteome analysis is not so easy. There are major challenges in serum or plasma proteome analysis. Let us talk about some of these challenges point wise; the first point the dynamic range of the protein concentration. In serum, there is a large diversity of proteins, which provide a very dynamic environment of almost 10 to the power 10 magnitude.

(Refer Slide Time: 07:07)

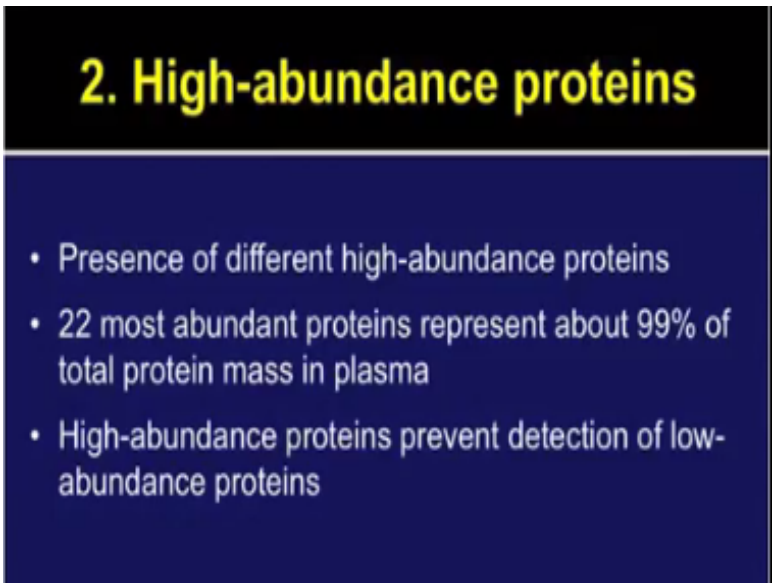


So, the concentration of serum proteins ranges more than 10 orders of magnitude, if you want to obtain the full spectrum of serum or plasma by applying any of the conventional proteomic techniques, it is very challenging because the typical dynamic range for any of these techniques, any these platforms will be much smaller which in 10 to the power 2 to 10 to the power 4, so how to capture all the dynamic events, which are happening in the serum.

If the dynamic range of the proteins are very large and your detection techniques are not able to capture that whole dynamic range, so to avoid these issues, people try to remove some of the

abundant proteins from very complex serum proteome, so that overall the dynamic range can be reduced and minimized. So, the second point the high abundance proteins; there are different high abundant proteins which are present in serum and plasma, which makes its analysis very very complicated.

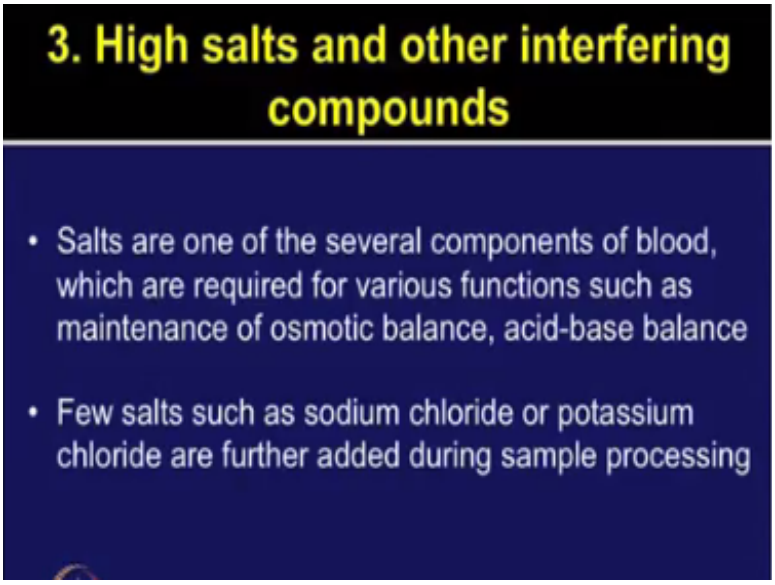
(Refer Slide Time: 08:22)



There are almost 22 highly abundant proteins present in serum, which represent about 99% of total protein mass of serum or plasma. These high abundance proteins prevent the detection of very low abundant proteins and often these low abundant proteins could be the target, which you are probably looking for as a part of biomarker discovery. So, how to get rid of the high abundant proteins?

I will describe some of the strategies, which can be used to remove high abundant proteins in next couple of slides.

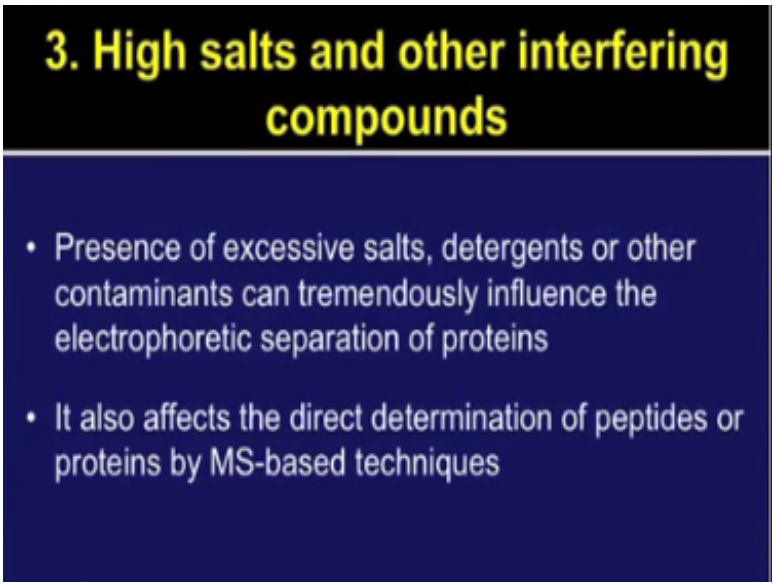
(Refer Slide Time: 09:14)



But first let us talk about which are other challenging factors for the serum or plasma analysis. The third point the presence of high salt and other interfering compounds, as we have talked about different type of interfering compounds, so now you are familiar that salt are one of the several components, which could be very much interfering during your proteomic applications. Now, salts are present in the blood which are required for various function such as the maintenance of osmotic balance, acid base balance etc.

Few salt such as sodium chloride or potassium chloride are also added, when you are processing the serum sample. During this whole sample manipulation due to the intrinsic salt present in the blood as the less extrinsic start added during the sample processing, the overall salt component becomes very high and that is; that creates problem for various type of proteomic applications.

(Refer Slide Time: 10:48)

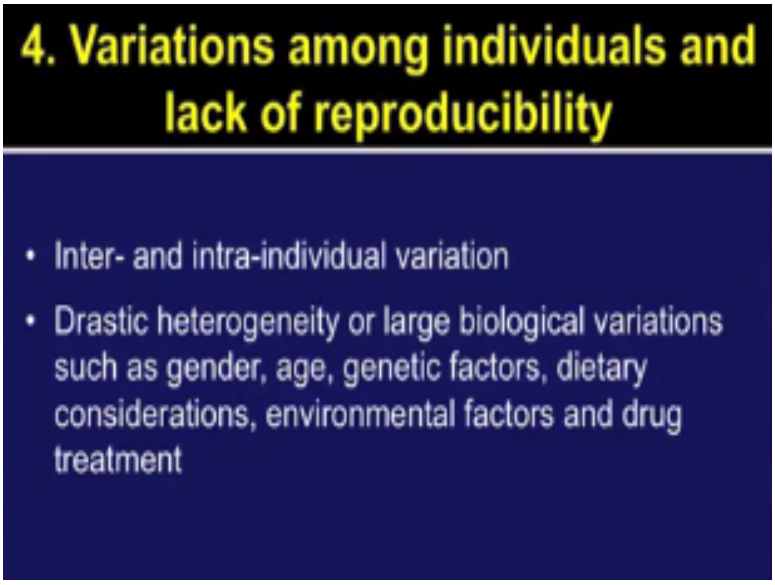


So, high salt and interfering compounds should be removed, now these presence of excessive salt, detergents or other contaminants can tremendously influence the electrophoretic separation of proteins, if your target technology is 2 dimensional electrophoresis or other gel based method you have to really ensure that salt is very low in the serum or plasma components. It also affects the direct determination of proteins or peptides by mass spectrometry based techniques.

So, regardless of whether you use gel based or gel free method, we have to ensure that the overall salt component is removed efficiently from the serum or plasma. Now, let us talk about fourth point, other challenge, so variations among the individuals and lack of reproducibility, these are some of the very major issues in clinical studies, where you have inter and intra individual variation.

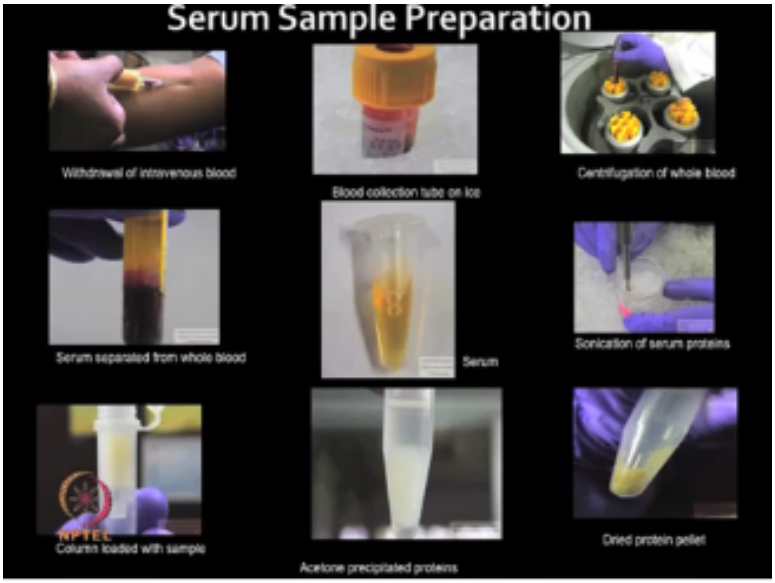
Intra individual variation is obviously more expected but even within one person due to the dye, due to different type of medication, the serum or plasma proteins can be changed. So, how to avoid these inter and intra individual variations and as if you refer to the previous section, when we talked about how to minimize different type of these factors which are going to ensure the success for clinical studies.

(Refer Slide Time: 12:16)



Probably, you will be able to keep an eye on different type of extrinsic and intrinsic influences which one should try to remove as much as one can during the clinical proteomic analysis, so the drastic heterogeneity or large biological variations such as gender, age, the genetic factors dietary considerations, environmental factors and drug treatments are going to affect the reproducibility of your experiments.

(Refer Slide Time: 12:55)

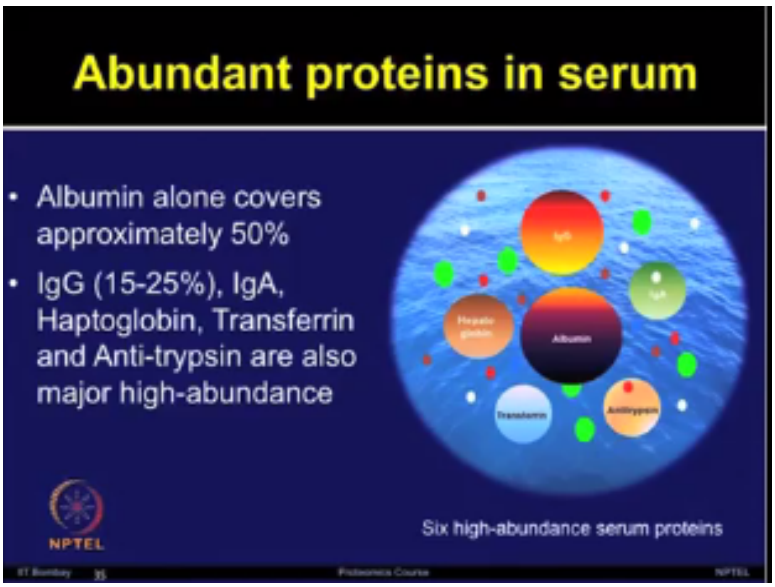


So, if you are careful in designing experiment, one can at least minimize these variations and increase the reproducibility of the proteome analysis. So, I am giving you an overview here for the serum sample preparation, the various steps are shown in images such as withdrawal of intravenous blood, blood collection in the tube, centrifugation step of whole blood, how to remove the serum from the whole blood, serum can be transferred into the fresh tube.

Now, for the serum proteome analysis, different type of modifications can be performed such as sonication can be used for disrupting the high abundant proteins and it also helps in better resolution, then depletion strategies to remove the abundant proteins, precipitation of these proteins by adding acetone and then drying out the protein pellet and reconstituting for the suitable buffer for further proteomic applications, okay.

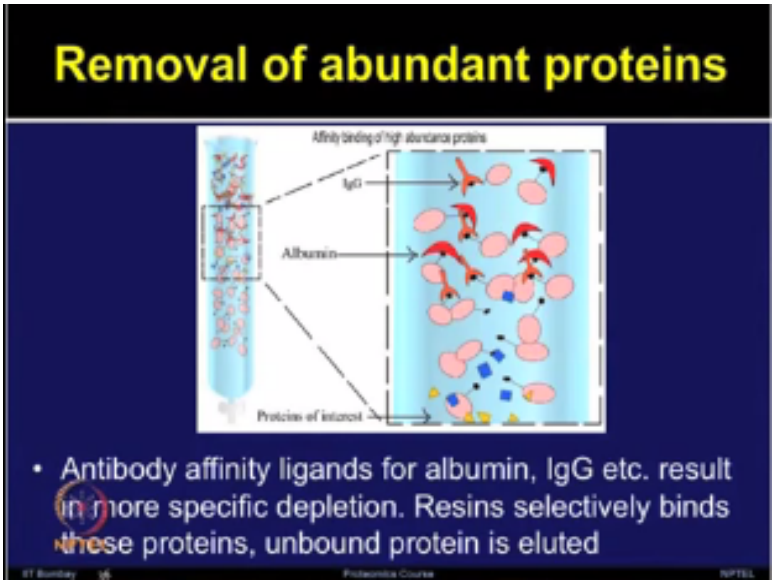
So, I hope in the animation you are able to understand how to go step by step to perform the serum proteome analysis. Now, as I promised previously, I will talk to you about how to remove the high abundant proteins which are present in the serum. So, as we talked there are more than 20 abundant proteins which are present in the serum and albumin alone covers approximately 50% of these abundant proteins.

(Refer Slide Time: 14:39)



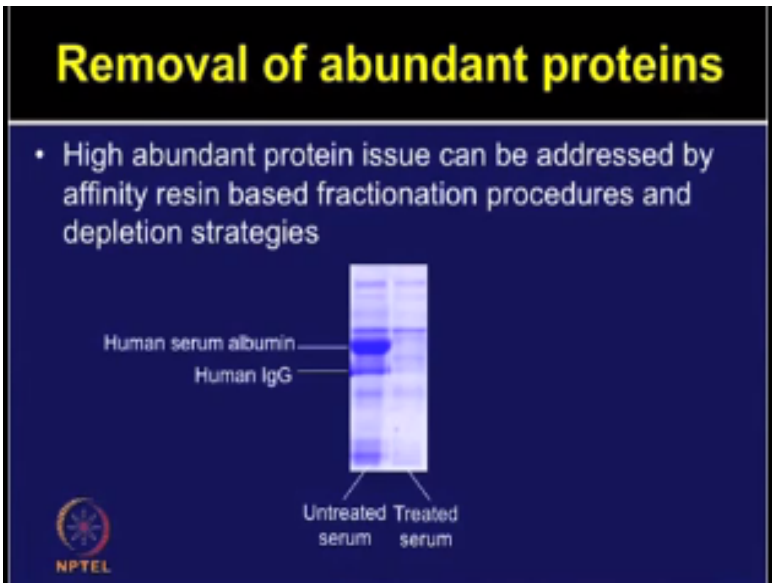
Immunoglobulin G contains 15 to 25% of the abundant proteins then IgA haptoglobin, transferrin and antitrypsin, these are also major high abundant proteins. So, in the diagram I have shown you the 6 high abundance serum proteins. The different ways people have tried to remove these abundant proteins, molecular weight cut off and different type of chromatography methods have been used.

(Refer Slide Time: 15:43)



The affinity chromatography based methods are highly efficient for its specific removal of these abundant proteins, which are present in the serum because there is always that fear with the molecular weight cut off whether along with albumin and other high body weight proteins, you will also get rid of your various high molecular weight protein, which are non-abundant. So, the affinity based methods ensure the specific targeting of albumin IgG and other specific abundant proteins.

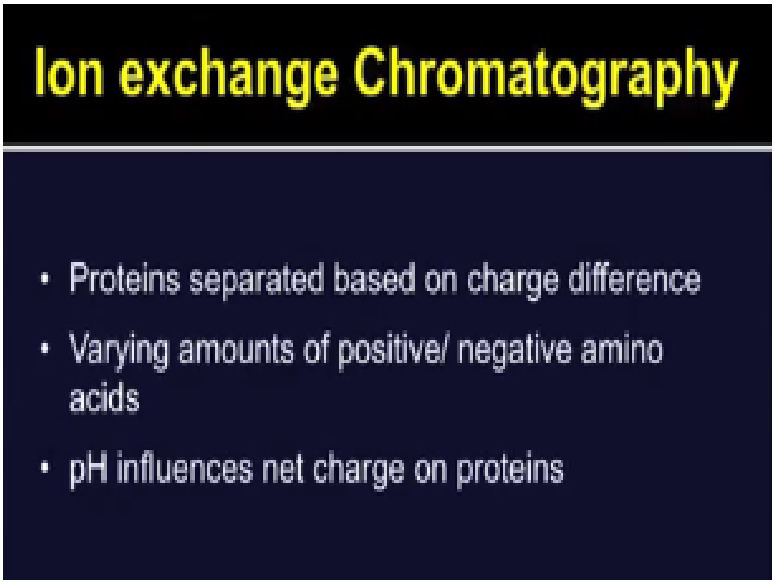
(Refer Slide Time: 16:15)



The antibody affinity ligands are used for albumin IgG and other abundant proteins and it result into very specific depletion. These resins can selectively bind to these proteins and unbound proteins can be eluted in suitable buffer. By applying the affinity resin based fractionation method and different type of depletion strategies, one can get rid off the high abundant proteins as shown in this image of SDS page gel.

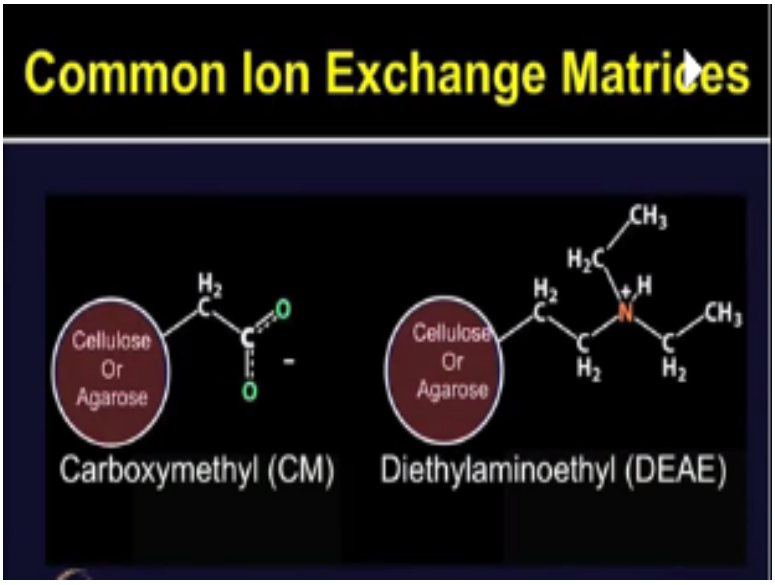
The left lane is loaded with the untreated serum and the right lane is loaded with the treated serum, as you can see some of the very high abundant proteins such as albumin IgG, those are efficiently removed and which allowed for some low abundant proteins to appear on the gel. Let us now move on to next chromatography method which is ion exchange chromatography, this is one of the most versatile chromatographic separation method which relies on differences between number of charges and distribution of charge groups in defined pH and solvent conditions.

(Refer Slide Time: 17:24)



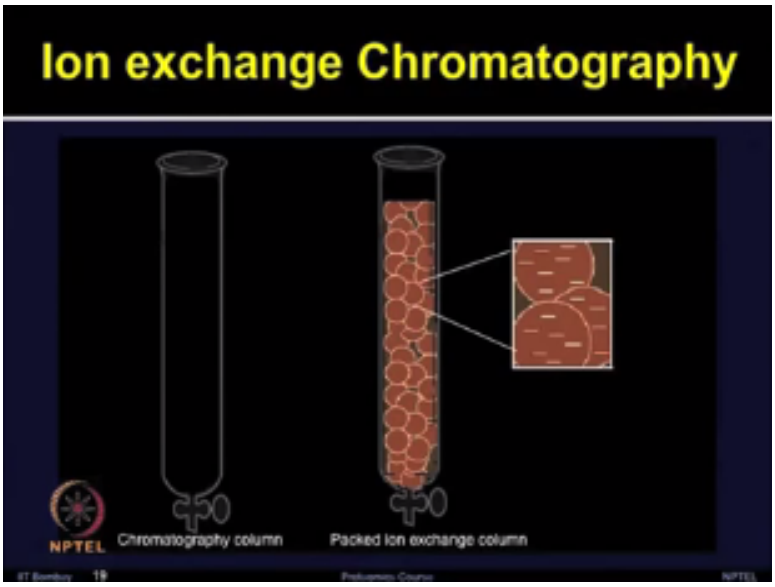
In ion exchange chromatography, the proteins are separated based on charge difference, the proteins with overall negative charge will interact with positive charges or the vice versa, so by varying the amount of positive and negative amino acids and even pH can influence the net charge on proteins.

(Refer Slide Time: 17:53)



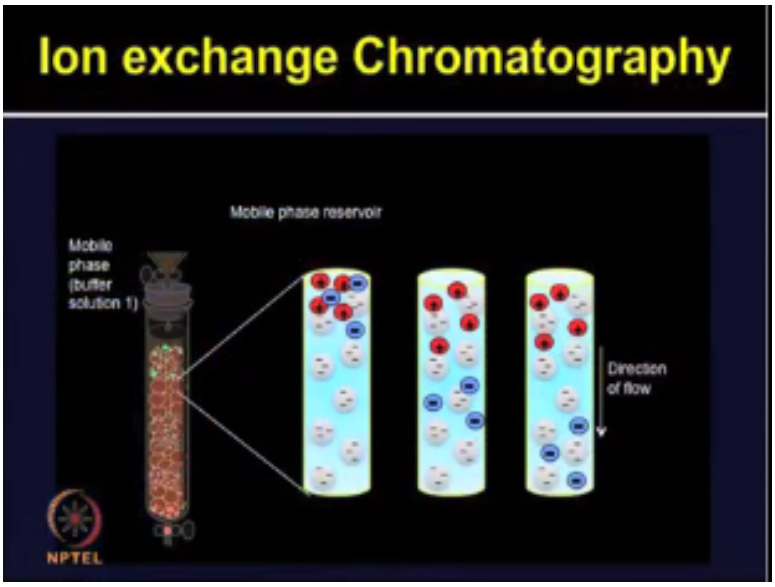
So, in this slide some of the common ion exchange matrices are shown like carboxymethyl CM and diethylaminoethyl, DEAE, so when a desired protein is positively charged, the cation exchange chromatography should be used when a desired protein is negatively charged, the anion exchange chromatography method should be used.

(Refer Slide Time: 18:22)



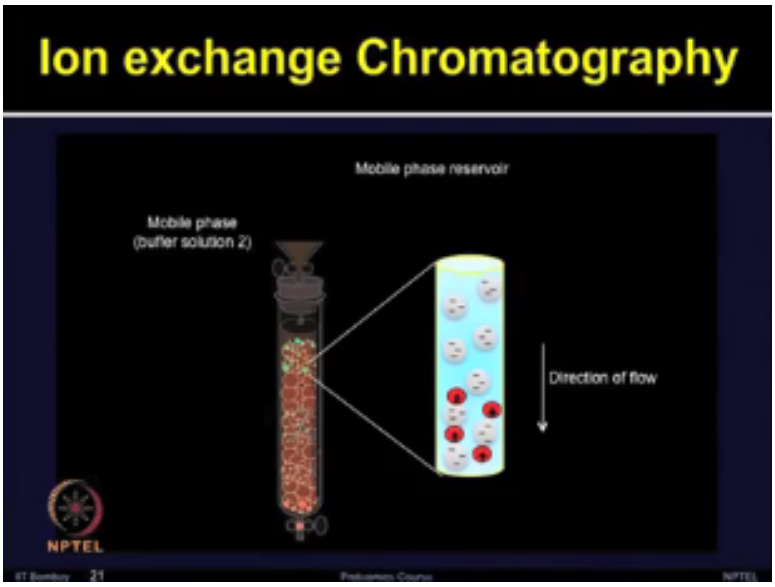
So, in ion exchange chromatography, the column is packed with a resin, whether its cation or anion exchanger depending upon the charge of the protein that needs to be bound to the column and purified.

(Refer Slide Time: 18:38)



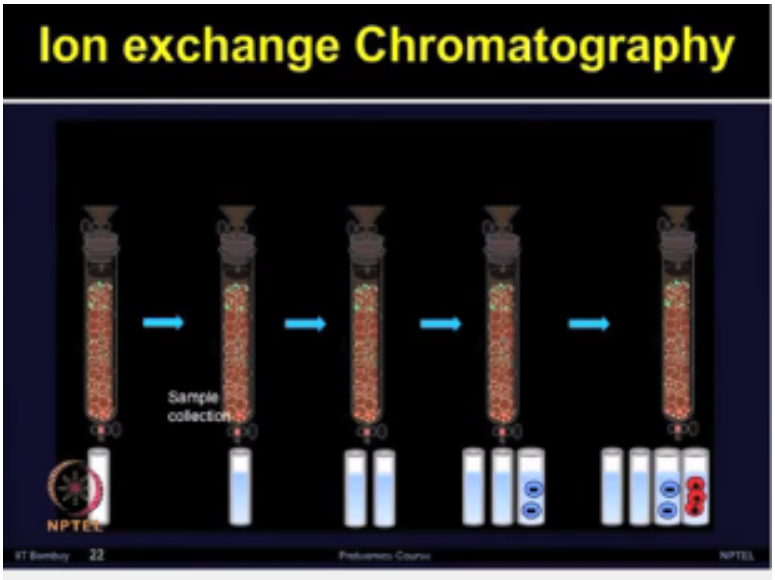
So, proteins are adds up to the ion exchange column and then it can be dissolved by increasing the salt or altering the pH of the buffer, which can change the charge on protein. So, various anionic buffers such as acetate and phosphate are used for cation exchange and cationic buffers such as Trischloride or ethanol amine are used for the anion exchange.

(Refer Slide Time: 19:11)



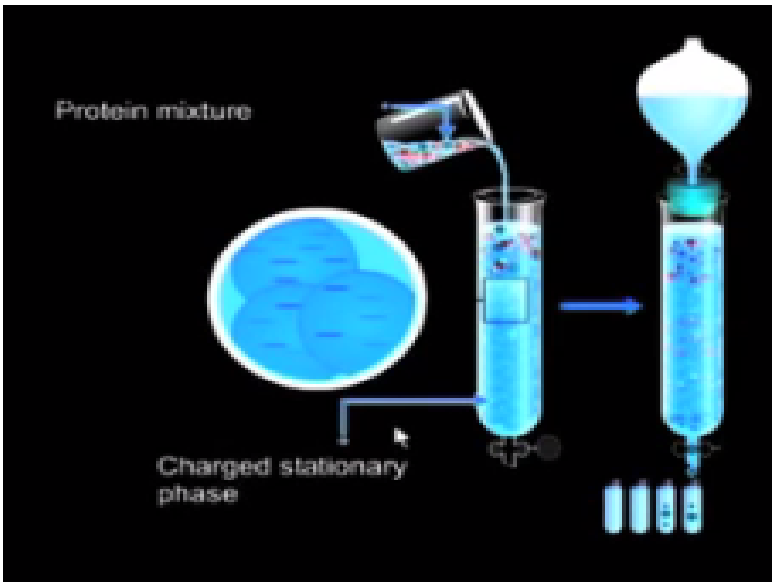
Now, the buffered solution is changed, so that the net pH of the protein of interest can be modified and it no longer binds to the ion exchange resin. Therefore, the bound protein can be eluted out as shown in this slide.

(Refer Slide Time: 19:35)



So, if you have negatively charged protein which gets eluted first will be present in the initial fractions, while the positively charged protein that bound to the column will be eluted in the latter fraction or it will be vice versa. So, let me describe how an exchange chromatography works is step by step in following animation.

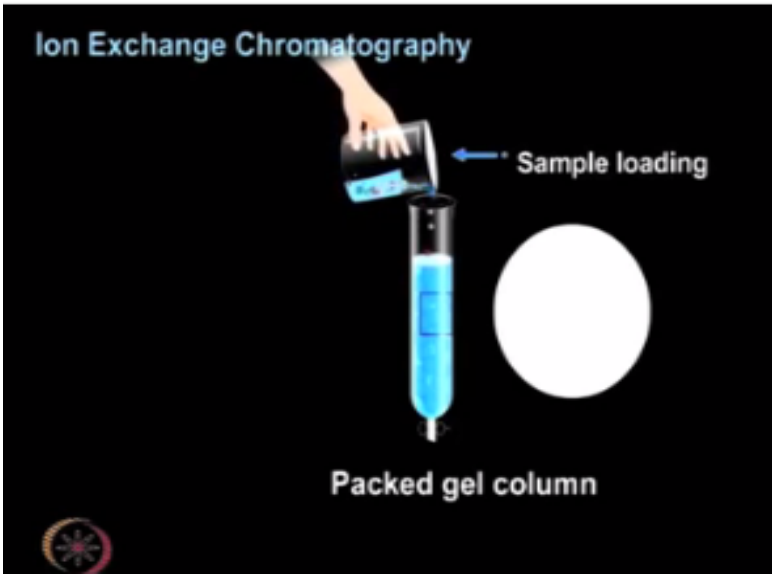
(Refer Slide Time: 19:54)



The charged stationary phase, the colonization rephase consists of a positively or negatively charged polymeric matrix which will bind molecules of the opposite charge. Commonly used ion exchangers include negatively charged carboxymethyl cellulose or CM cellulose, which is the cation exchanger and positively charged DEA cellulose which is an annoying exchange. The protein mixture and purified protein mixture which consists of proteins of different net charges are loaded onto the column.

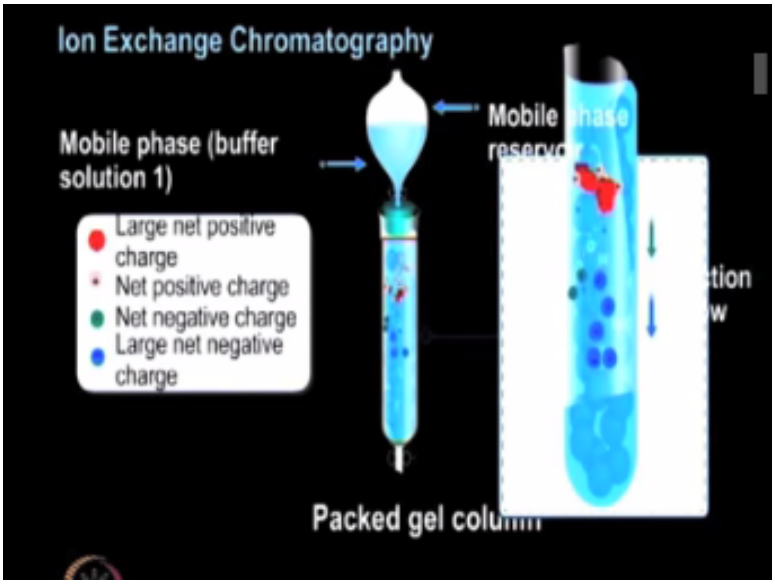
The proteins having charges opposite to that of each stationary matrix will bind to it while remaining proteins will be eluted. Mobile phase, the proteins are eluted out of the columns by using suitable mobile phase and then samples are collected by using different sample fractions. The solution leaving the column can be collected in suitably sized fractions for further analysis.

(Refer Slide Time: 21:15)



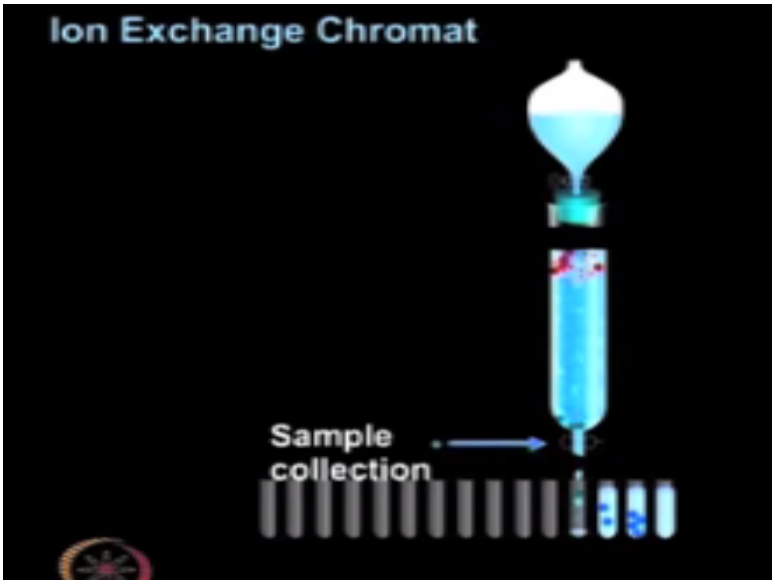
After giving you a brief description of the components, let me show you the process in animation, the column is packed with the suitable cation or anion exchange resin depending upon the charge of the protein that needs to be bound to the column and purified. The anion exchange column is then loaded with the impure protein mixture consisting of various positively and negatively charged proteins.

(Refer Slide Time: 21:59)



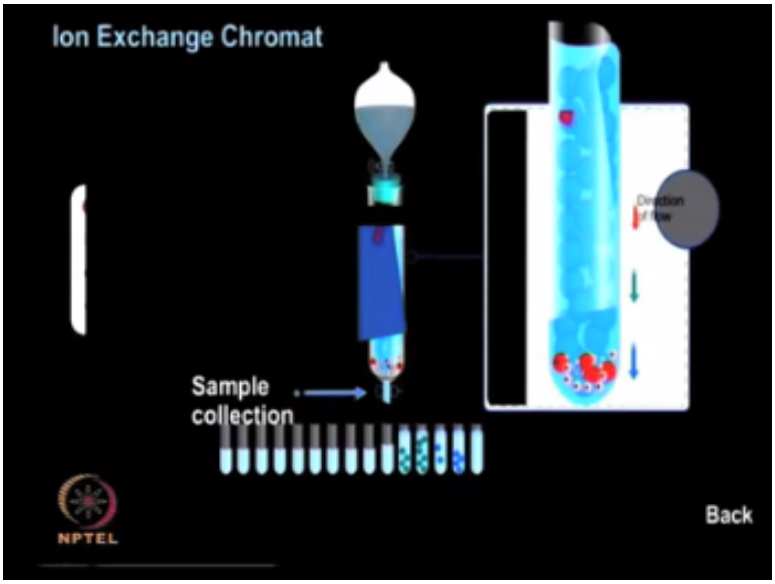
The column is eluted with a buffered solution of suitable pH such that the negatively charged molecules are removed from the column, while the positively charged molecules remain bound to the anion exchange resin. The buffered solution is then changed such that the net pH of the protein of interest is modified and no longer binds the ion exchange resin.

(Refer Slide Time: 22:19)



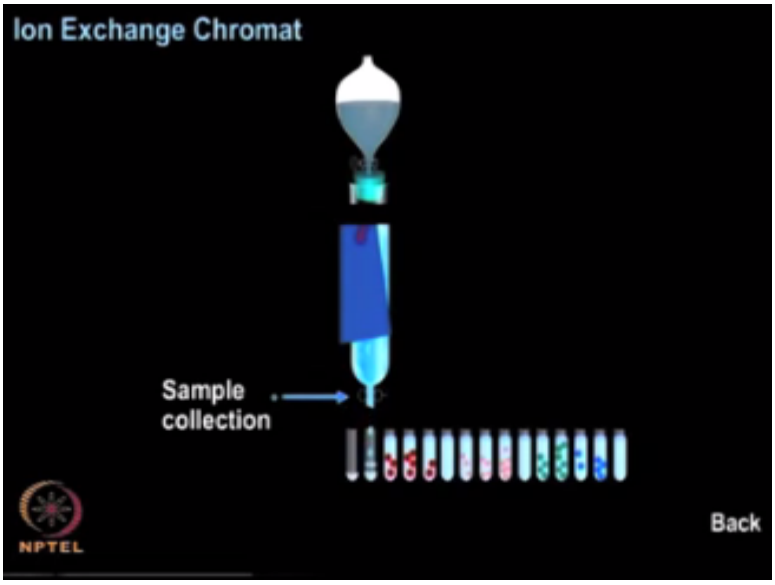
Therefore, the bound protein also gets eluted out of the column in this manner. The fraction of appropriate size must be collected and analyzed for their protein contents.

(Refer Slide Time: 22:38)



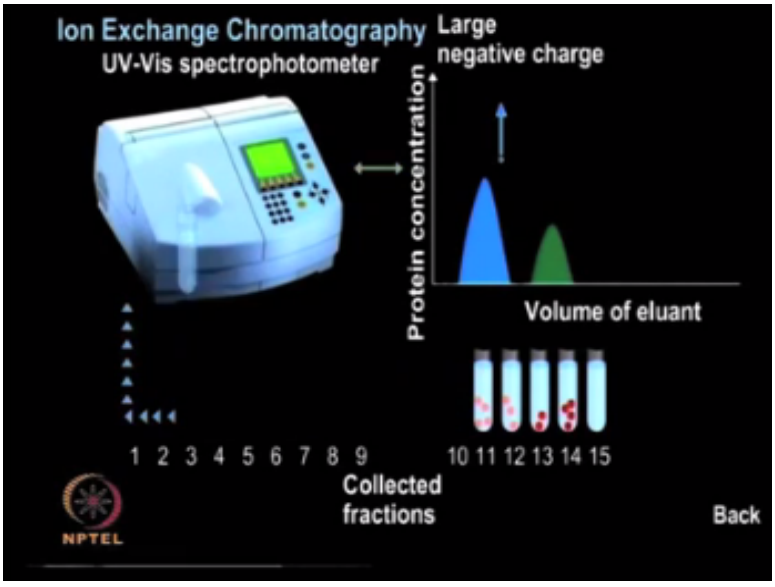
The negatively charged proteins which get eluted first will be present in the initial fractions while the positively charged protein that bound to the column is eluted in the latter fractions.

(Refer Slide Time: 22:56)



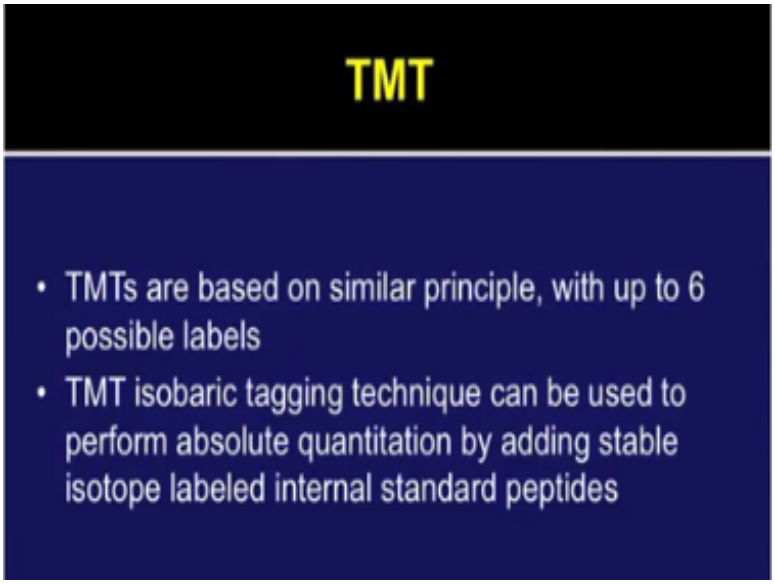
Once all the fractions are collected, then the protein content can be analyzed by using a spectrophotometer. So, analyse the contents, these fractions for their protein content by using a UV visible spectrophotometer at 280 nanometers.

(Refer Slide Time: 23:28)



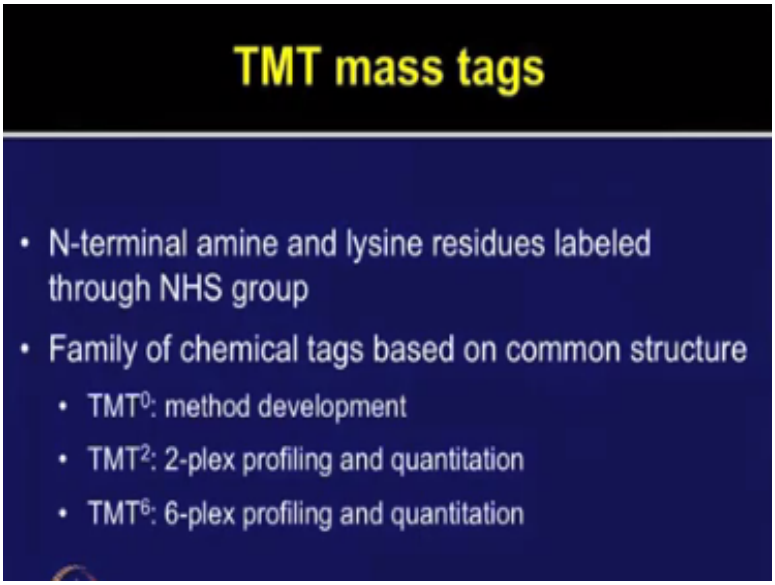
A graph of eluant volume versus protein concentration can then be plotted. In this particular example, the negative charge large molecules coming first and then positively charged fractions are coming later, okay. So, now let us talk about tandem mass tag or TMT, this method is similar to the iTRAQ which we just discussed. TMT is also MS/MS based quantitative technique, which uses the isotopomer levels referred as tandem mass tags.

(Refer Slide Time: 24:35)



It also provides the accurate quantification of peptides and proteins. The tandem mass tags have been developed by the proteome sciences and currently commercialized by thermo efficient, I have given you the reference for the original study on tandem mass tag in the slide. So, these tender mass tags; they are based on the similar principle of iTARQ, here the possibility for multiplexing is up to 6 possible labels.

(Refer Slide Time: 25:15)

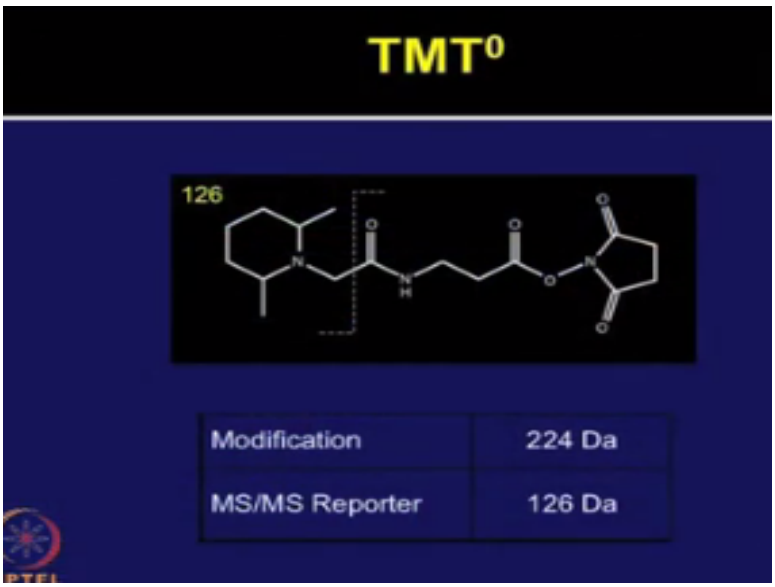


The TMT isobaric tagging technique can be used to perform absolute quantification by adding a stable isotope labeled internal stereo peptides. It can be done by comparing the peptides from a target protein to a known amount of labeled standard peptide are spiked into a sample. In that way, the absolute quantification can be obtained. The N-terminal amine and lysine residues are labeled through the NHS group.

There are family of chemical tags which are based on the common structures. The series of TMT tags available TMT0, TMT 2 plex, TMT 6 plex, so these TMT's are an innovative set of isobaric mass tags for labeling the proteins and peptides at amine functions and mixing of up to 6 different protein samples are possible, while duplex and 6 plex labels, TMT differ by the number of isotopic substitutions.

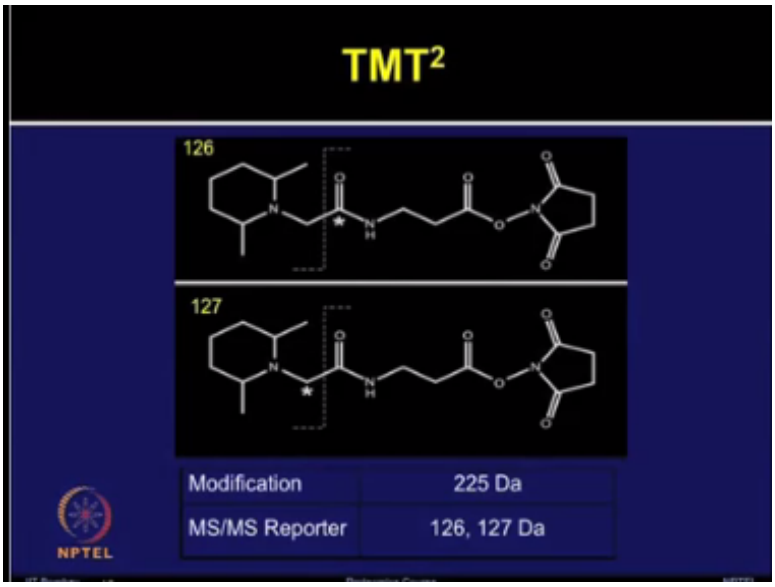
The TMT 0 is a non isotropically substituted a structure that has been produced for only method development. During the MS/MS analysis, the TMT tag give rise to 6 reporter ions from 126 to 131 Dalton, therefore it allows for the relative quantification. The TMT 6 plex each tag adds a mass of 229 Daltons for labeled amine to the protein. The TMT duplex and TMT 0 share the TMT complex structures.

(Refer Slide Time: 26:50)



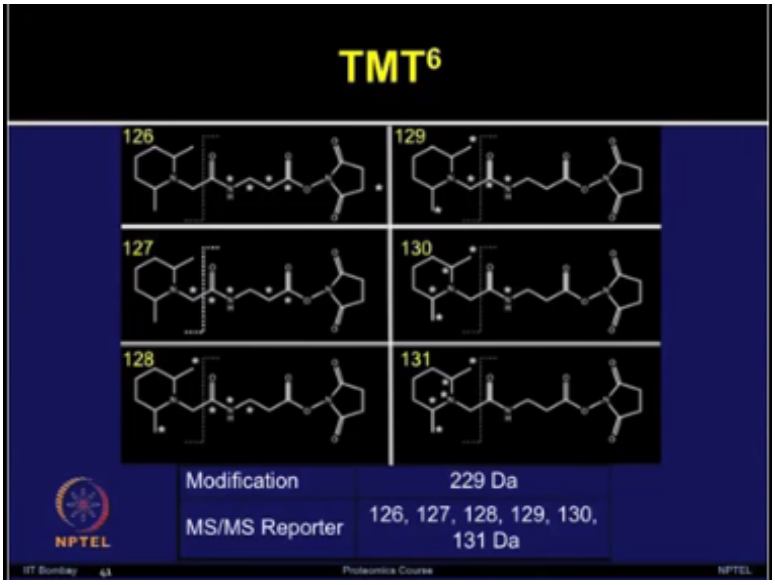
Let us look at the TMT label structure in more detail. The TMT0 tag is used for testing and optimization of the sample preparation, labelling, fractionation and MS fragmentation for peptide identification and reporter detection, the modification is 224 Daltons and MS/MS reporter ion is 126 Daltons.

(Refer Slide Time: 27:20)



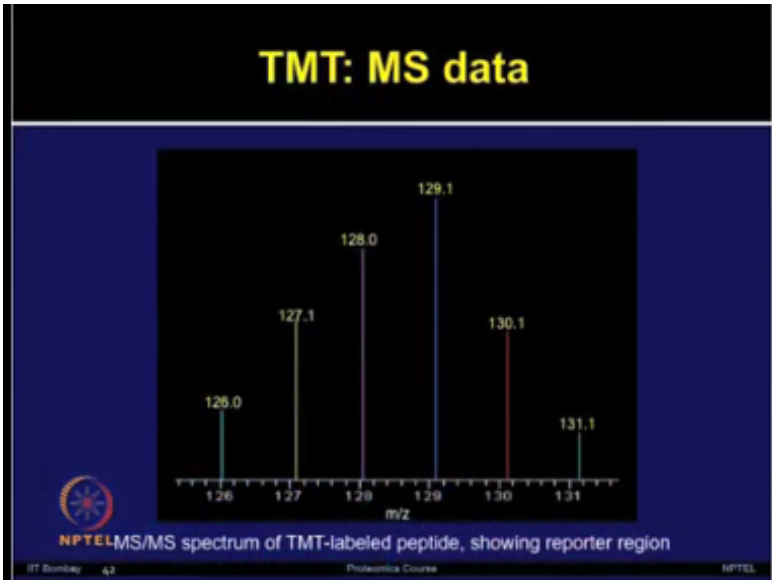
Now, let us look at TMT duplex. The TMT duplex reagent allows for the comparison of 2 samples; 126 and 127, these are 2 different MS/MS reporter ions available and modification is 225 Daltons.

(Refer Slide Time: 27:37)



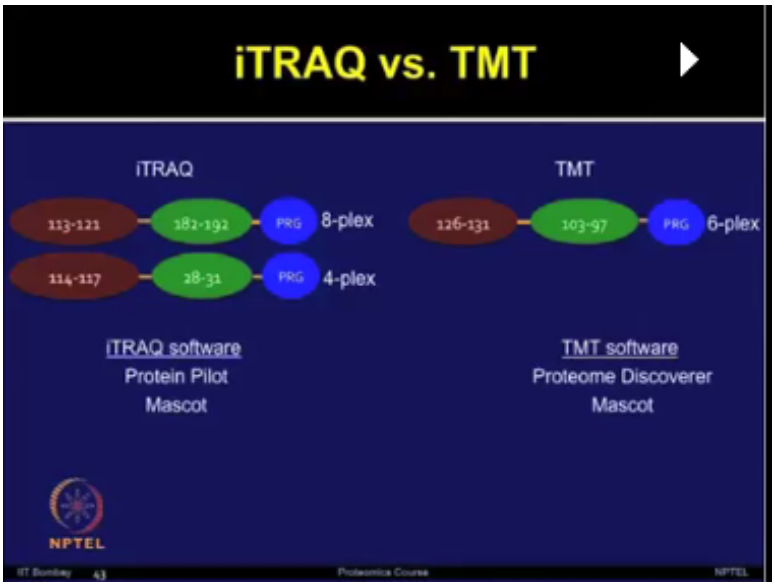
Let us now look at the TMT 6 plex reagent, it allows the comparison of up to 6 conditions. The MS/MS reporter ions, as you can see in the structure are from 126, 127, 128,129, 130 and 131 Daltons. So, the TMT 6 plex regent allows comparison of up to 6 conditions, it could be useful for studying about time codes, drug, those responses replicates or looking for multiple sample disease comparison.

(Refer Slide Time: 28:22)



The modification is 229 Daltons, I am showing you one representative MS/MS spectrum of TMT labeled peptide which is showing a reporter region, the relative abundance of target protein or peptide fragment in 6 different samples can be easily measured by comparing these signature mass peaks, which are generated by the different mass tags.

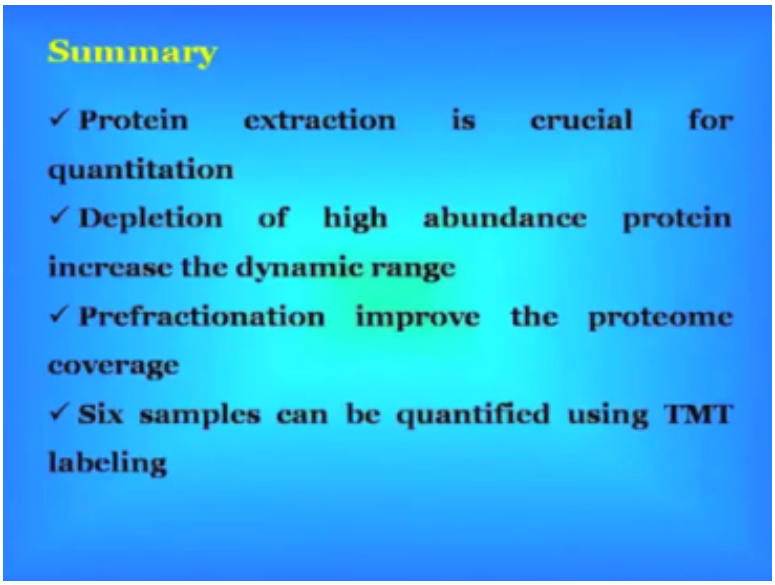
(Refer Slide Time: 28:53)



Let us now look at the comparison of iTRAQ and TMT tags. When iTARQ, as we talked, there are 2 different type of regents available, 4 plex and 8 plex, in both there is a reporter group, a balancer group and then there is a protein reactive group. Same concept is also in the TMT tags, where we have a reporter group, a balancer group and PRG; I have shown you a comparison with a 6 plex TMT tag.

But, as we have seen earlier, there is TMT duplex and TMT 0 tags are also available. Now, in iTRAQ, the reporter in the 4 plex consist of either from 114 to 117 Dalton, the balancer is between 28 to 31 Dalton, whereas in 8 plex, it is from 113 to 121 and the balancer is from 182 to 192. In case of TMT, the reporter groups having 126 to 132 Daltons and the balancer consists of 97 to 103, so tags are quite similar in the overall structure.

(Refer Slide Time: 30:24)



The iTRAQ analysis can be performed by using the software such as protein pilot and also the mascot, the TMT based analysis can be performed from software such as proteome discoverer and mascot.