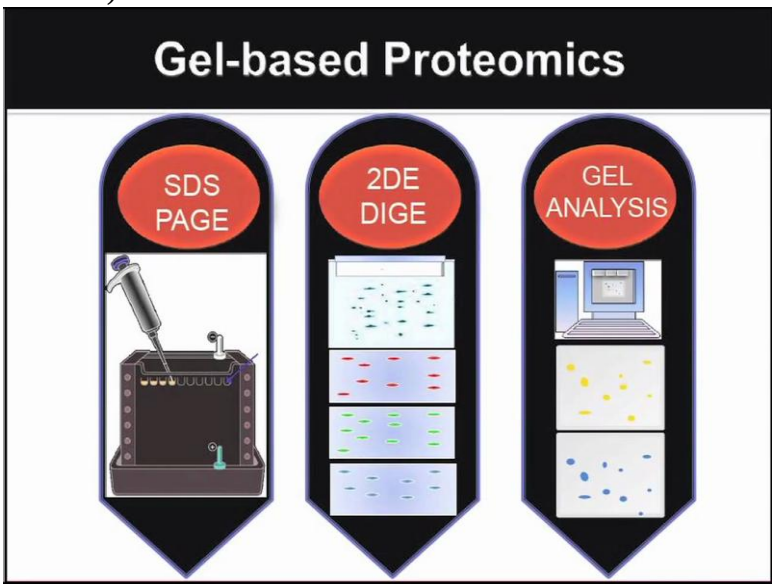
Proteins and Gel-Based Proteomics
Professor Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay
Mod 01 Lecture Number 1

(Refer Slide Time: 00:13)

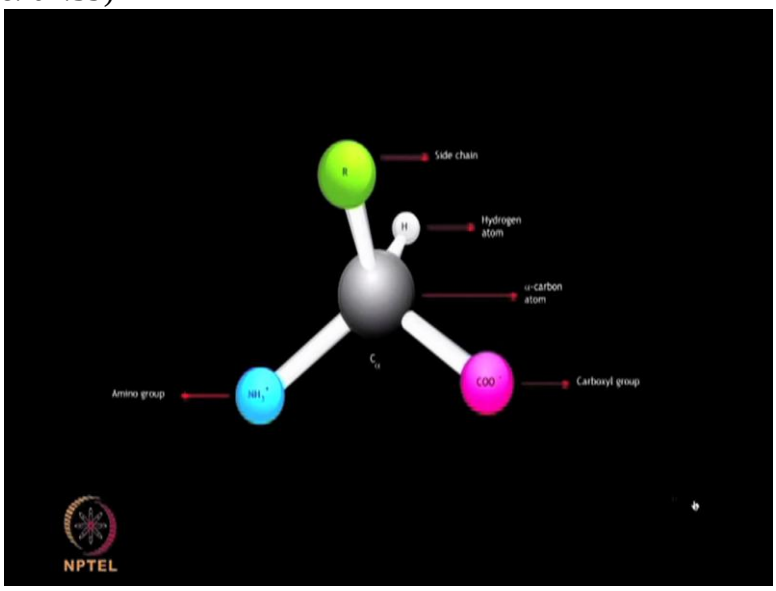


Welcome to the MOOC NPTEL course on Proteins and gel-based proteomics. Before we move on to gel-based proteomics and discuss various techniques and concepts involved in proteomics let me give you an overview of whole course.

In week 1, we will talk about basics of Amino acids and then we will move on to the structural level of proteins, their folding and misfold. Further we will study various chromatography techniques employed for the protein purification.

In week 1, the first lecture we will start with basics of amino acids and then we will discuss about the structural backbone and their classification.

(Refer Slide Time: 01:39)



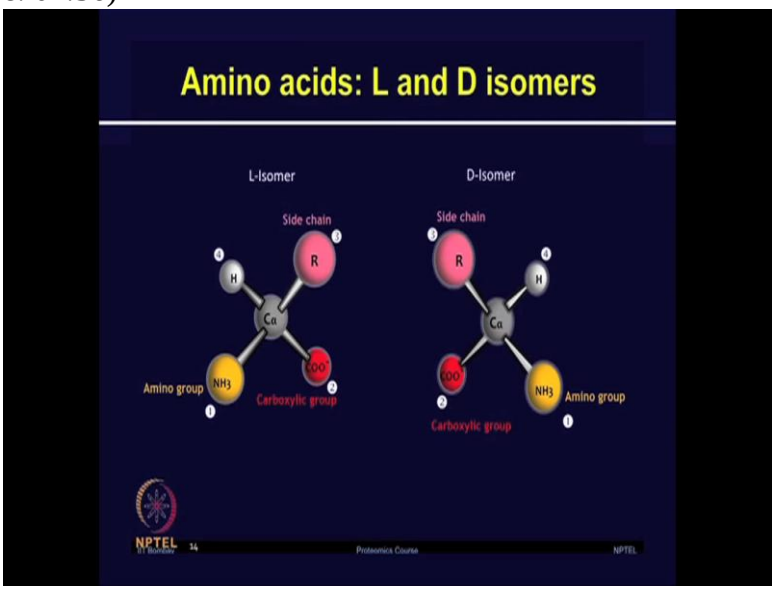
Amino acids are the building blocks or the monomers that make up proteins. They consist of central alpha carbon atom, bonded covalently to an amino group, a carboxyl group, a hydrogen atom, and variable side chain called as R group.

Amino acids are the basic monomeric constituents of proteins found in varying amount depending upon the type of protein. They are classified based on the properties of their side chains, or R groups which vary in size, in structure and charge.

Polarity of side chain is one of the main basis for the classification; amino acids having non-polar aliphatic side chains including glycene, alanine, proline, valine, leucene, isoleucene and metheonine.

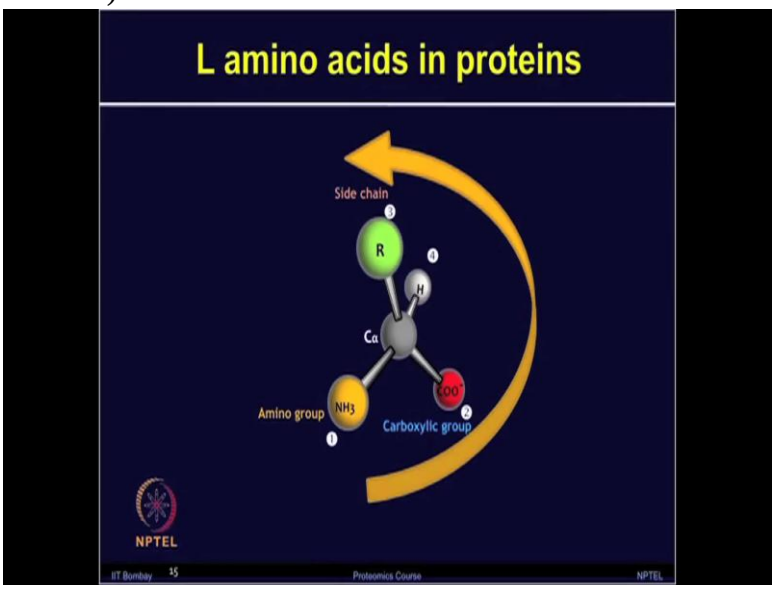
We will then discuss about different types of amino acid isomers, L and D isomers with respect to the conformation stability and their natural propensity to be found in the L isomeric form.

(Refer Slide Time: 02:56)



Amino acids have 4 different groups which are connected to the alpha carbon atom. It can form 2 mirror images which can exist in L or the D isomers which are shown in the slide here.

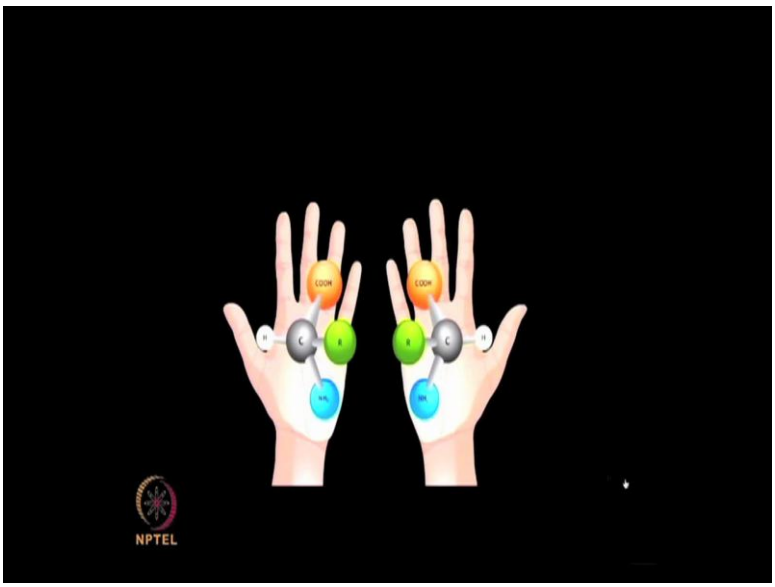
(Refer Slide Time: 03:18)



The alpha amino acids are chiral. There could be R or S configurations in the amino acids depending on the priority groups but only L amino acids are present in the proteins.

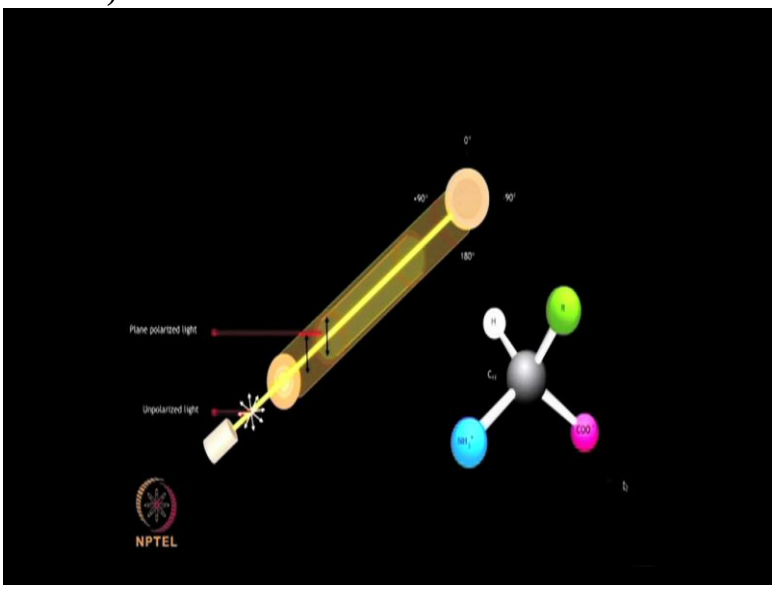
All the L amino acids have S configuration which describes the counterclockwise direction as shown here from the highest to the lowest priority group which is an indicative of the chiral center with the S configurations. Some of the isomerism properties of amino acids will be discussed in the following animation.

Before learning about the isomerism, let's first know what is chirality. The term chirality arises from the Greek term cheir meaning handedness. (Refer Slide Time: 04:11)



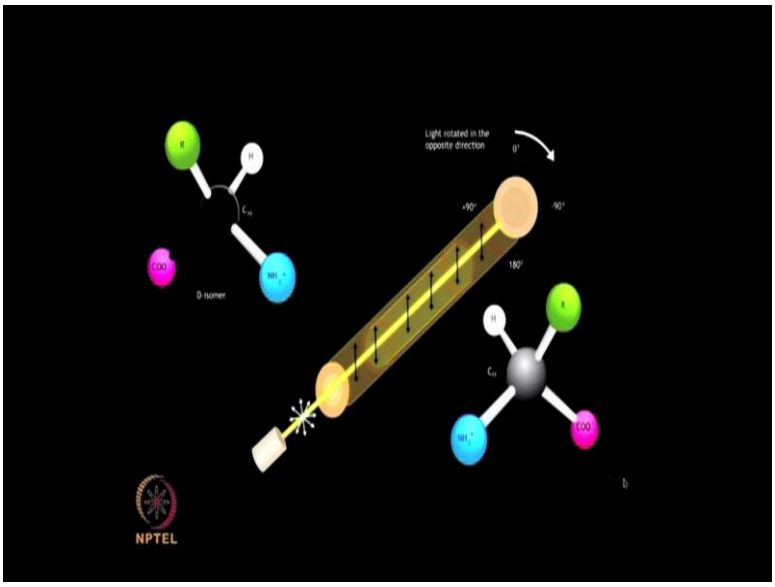
Just like 2 hands are non superimposable mirror images of each other, amino acid molecules are also non-superimposable due to their chiral alphacarbon center.

(Refer Slide Time: 04:29)



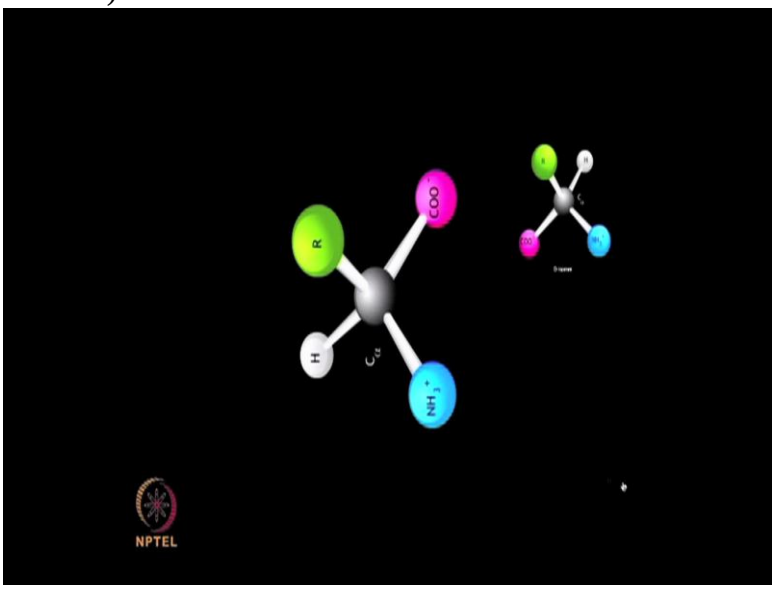
All amino acids except lysine contain an asymmetric center that makes them chiral in nature due to which they can rotate the plane of polarized light.

(Refer Slide Time: 04:48)



The two enantiomers designated as D and L rotate the plane of polarization in opposite directions. The two enantiomers of amino acid are non-superimposable mirror images due to the special arrangements of 4 different groups about the chiral carbon atom.

(Refer Slide Time: 05:23)

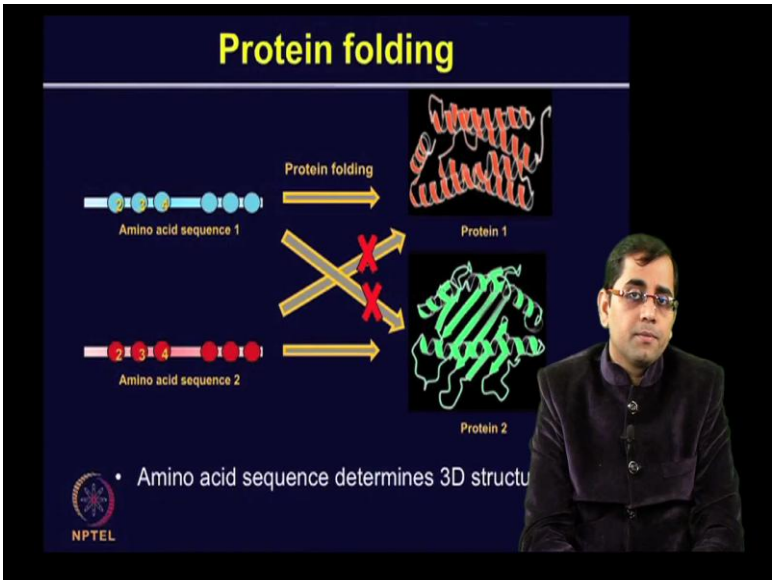


Rotation of either isomer about its central axis will never give rise to the other isomeric structure.

In next lecture, we will talk about the versatility of proteins as macro-molecules and the building blocks of the body. Further we will scrutinize the structural hierarchy of the proteins. We talk about secondary structural elements commonly found in proteins. Understanding

different levels of structural organization of proteins can help us to extrapolate their functional significance to the cell or the organism and it will be covered in this lecture.

(Refer Slide Time: 06:37)



In third lecture we will talk about the tertiary structure of proteins which is formed due to folding of proteins. We will also discuss about the improper protein folding and aggregation which can be corrected by the molecular chaperons which facilitates the process of folding in case of large proteins.

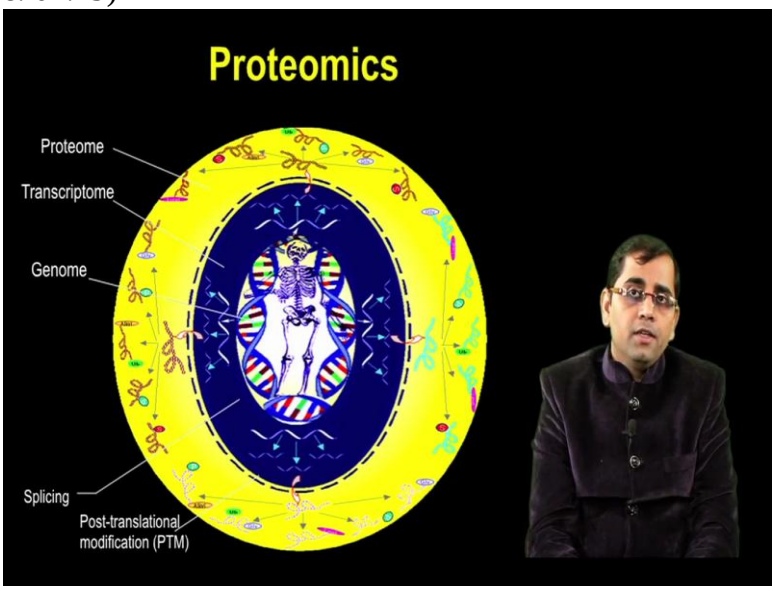
I will also discuss about the significance of protein folding with respect to diseases and functions like cellular trafficking, regulation of cell cycle and cell growth.

The next lecture will be focused on protein purification, peptide isolation and analysis. We will discuss in detail about the different types of chromatographic techniques like gel filtration, ion exchange chromatography, affinity chromatography, HPLC, strong cation exchange and reversed phase chromatography.

These chromatography techniques are often used for various proteomics applications and it is good to cover their basic understanding in the first week of this course.

The lecture five will explore the field of proteomics and give an introduction to proteomics, how it is different from conventional protein chemistry.

(Refer Slide Time: 07:49)

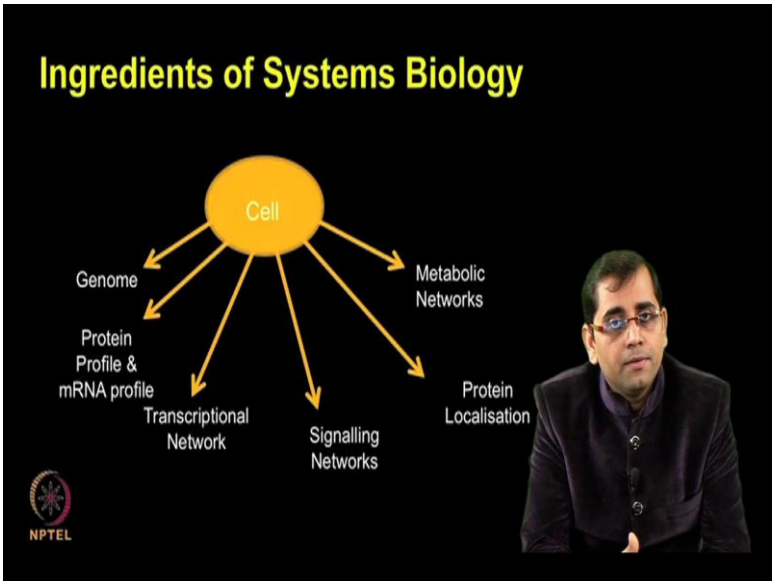


In this lecture we will discuss about how proteomics can provide holistic overview of studying the proteins and protein complexes. We will discuss about how protein chemistry and proteomics differ in the techniques which are employed in these approaches as well as their scale of usability.

The traditional protein chemistry has employed some of these techniques like electrophoresis and mass spectrometry. So what is different?

The advancement of these techniques along with genome sequence availability has made that difference which will be discussed in this lecture. The sensitivity, resolution and robustness of high throughput approaches which has made this transition will be discussed in detail.

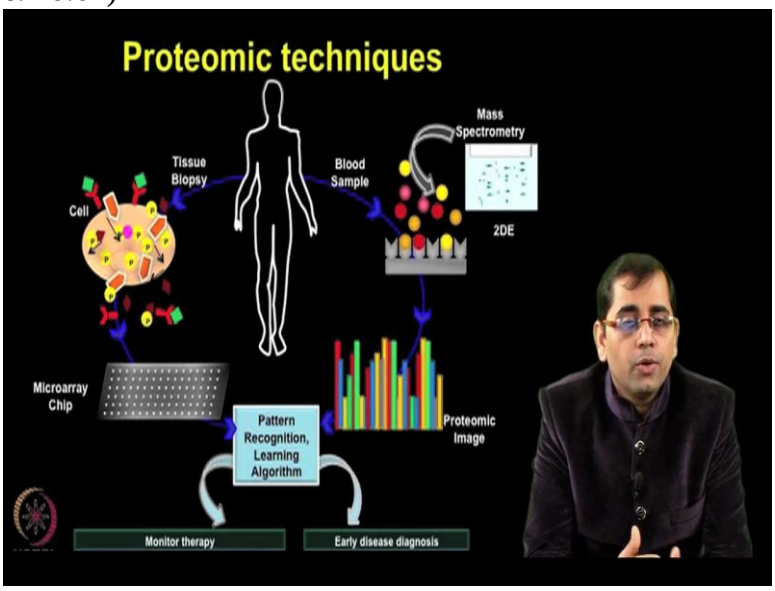
(Refer Slide Time: 09:31)



The second week we will focus on system biology and proteomics as well as sample preparation strategies. The first lecture we will start on system biology examination of a biological entity as an integrated system rather than studying its individual characteristic reactions and components.

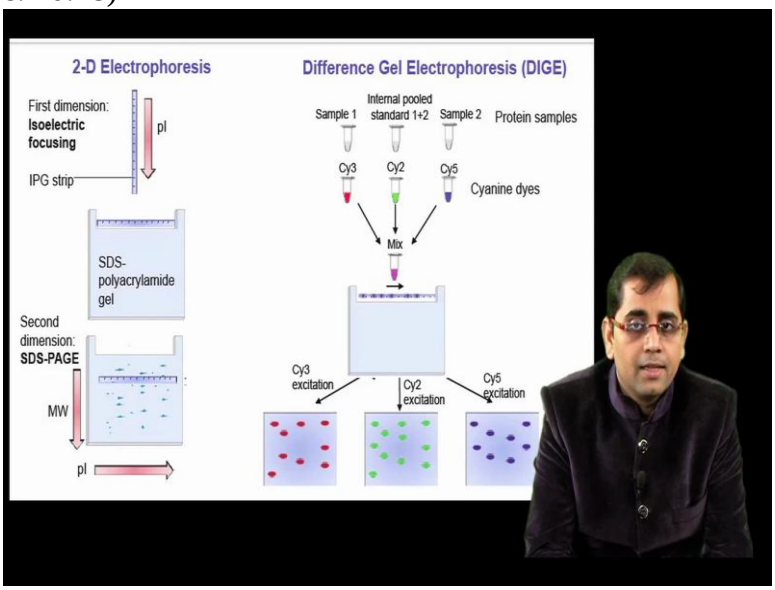
The various ingredients of system biology, for example...if you are studying about a cell and system behavior, one needs to look at the genome, transcriptome and proteome profile, how protein DNA and different type of transcriptional networks are altered, protein-protein signaling networks, multimeric complexes how are they formed and how protein is localized in the intra-cellular dynamics and metabolic networks. All of these are ingredients of studying about the system.

(Refer Slide Time: 10:04)



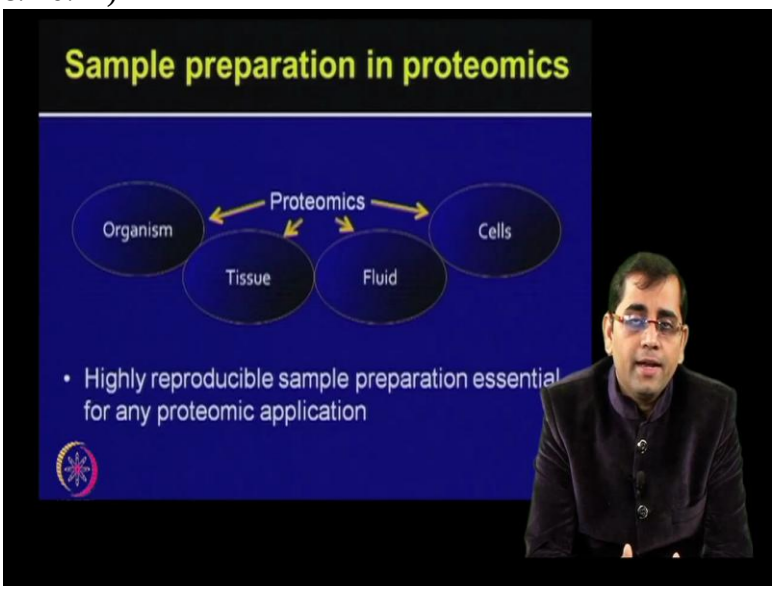
There are several techniques which have been used in proteomics which typically aim to elucidate the expression, localization, interaction and cellular function of proteins.

(Refer Slide Time: 10:18)



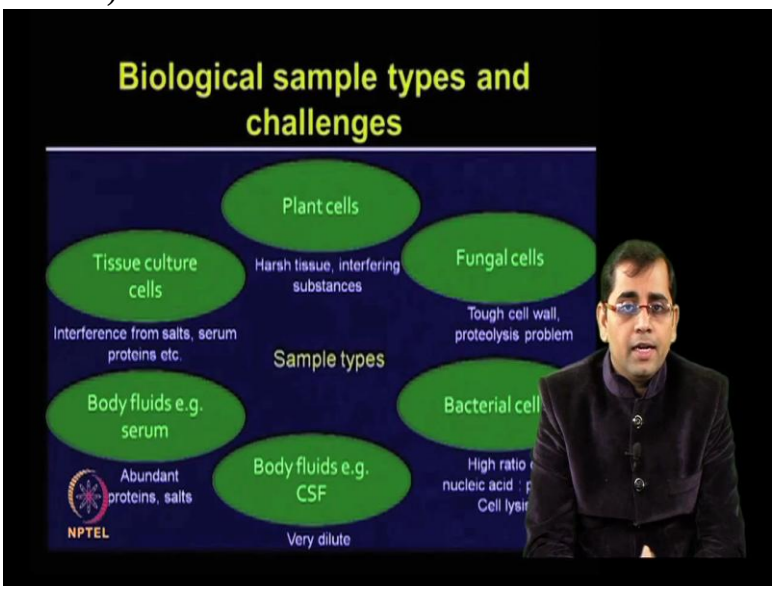
Lecture 2 will provide an overview of various gel based techniques such as SDS PAGE, two-dimensional electrophoresis and DIGE which are commonly used for proteomics applications.

(Refer Slide Time: 10:44)



Next we will talk about the sample preparation strategy for proteomic analysis. The sample preparation is very essential for any proteomic analysis. The protein extraction methods aimed to extract all the proteins in a cell or its organelle and the presence of interfering components have to be minimized.

(Refer Slide Time: 11:06)



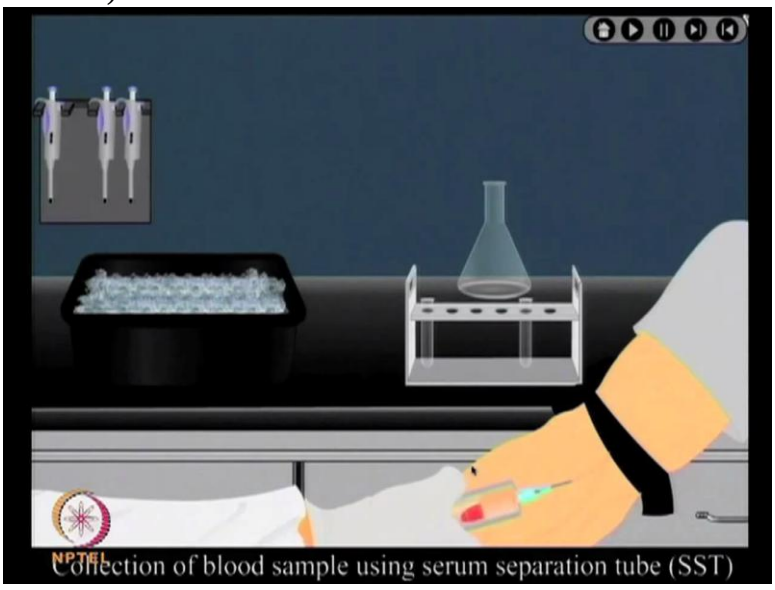
Different biological samples pose different challenges. For instance serum proteomic analysis is very complex because there are different proteins of different order of magnitude present in serum. Some are very highly abundant, some are very low abundant. Proteins such as albumin and immunoglobulin are highly abundant and it is preferred to remove some of these abundant proteins prior to any proteomic analysis.

In bacterial protein sample preparation, sonication becomes very important to disrupt the bacterial membrane. Sonication process breaks open the cellular membrane to release intra-cellular components.

Accurate protein quantification becomes very crucial for the quantitative protein analysis. Protein quantification is sensitive to detergents and certain ions therefore it becomes crucial to select the correct quantification method. Absorbance of standard and unknown protein samples can be measured and protein concentration can be determined from the standard plot of absorbance values.

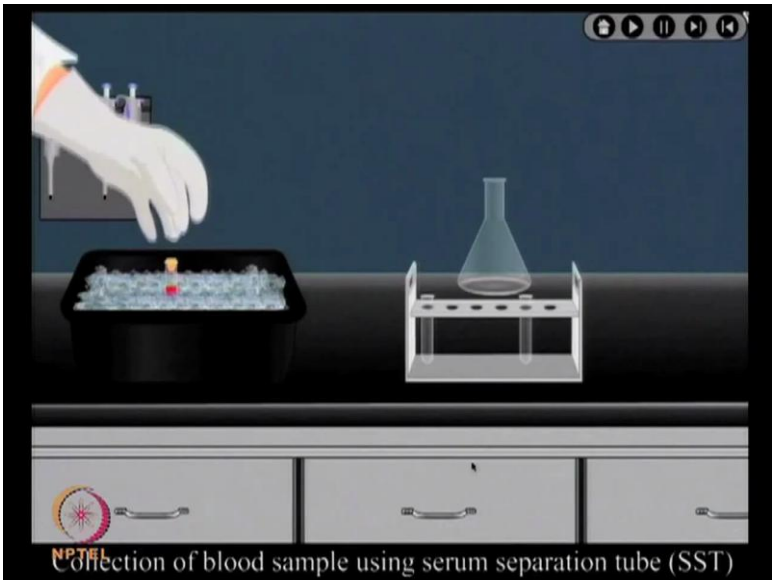
Serum proteome analysis: In this interactive animation, I will show stepwise how to collect serum, and how to perform the proteome sample preparation.

(Refer Slide Time: 12:42)



So first click on the syringe so that blood can be collected. The blood proteome is one of the most complex components of the human proteome. It fluctuates depending on the physiological and pathological conditions of the patient. So collect almost 5 ml of blood from the vein of healthy or diseased participants using serum separation tubes.

(Refer Slide Time: 12:54)

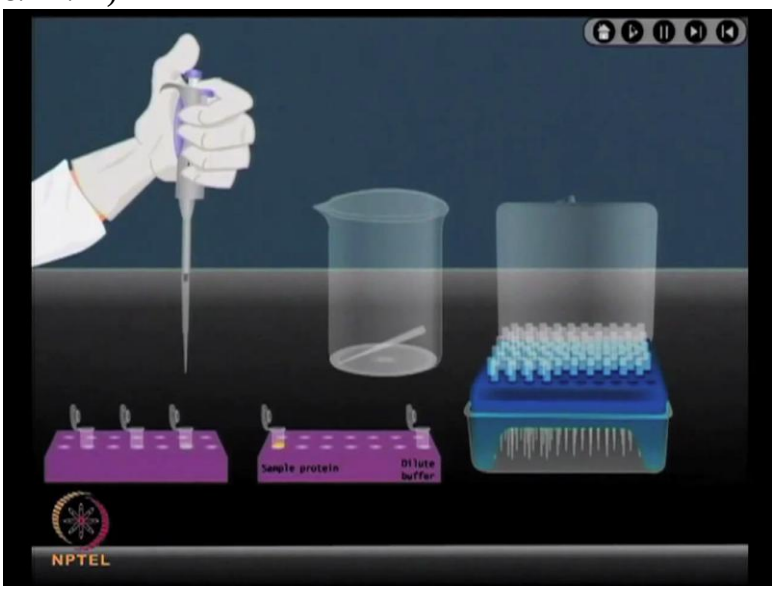


Immediately after the blood collection, the tube should be kept in ice for 30 minutes for clotting.

(Refer Slide Time: 13:03)

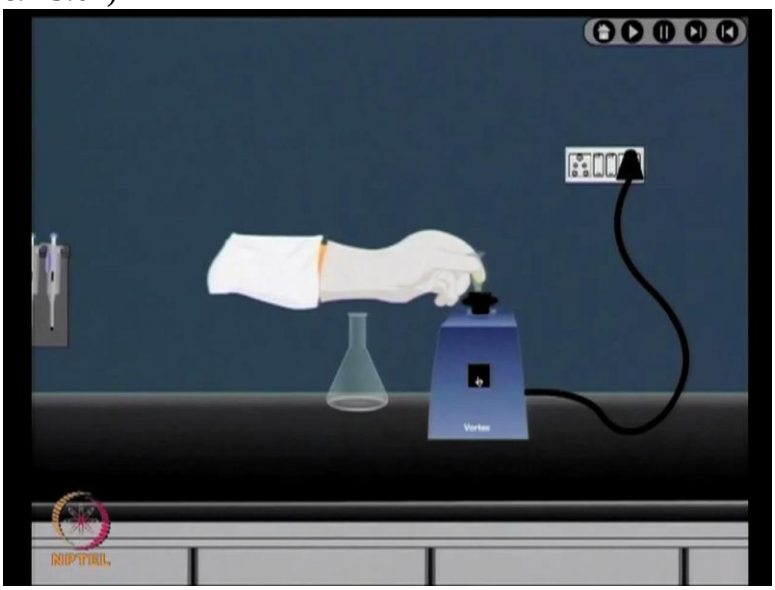
After clotting, the sample is centrifuged at 12000 rpm at twenty degrees for 10 minutes. Serum can be separated and multiple aliquots can be made for further storage at -80 degrees. So during this animation for the sample preparation, I will describe one method which we are currently using in my laboratory where we have tried combining different types of methods being available from different published protocols.

(Refer Slide Time: 14:44)



So the crude serum can be diluted five times with phosphate buffer pH 7.4. Ensure that mixing is uniform.

(Refer Slide Time: 15:07)

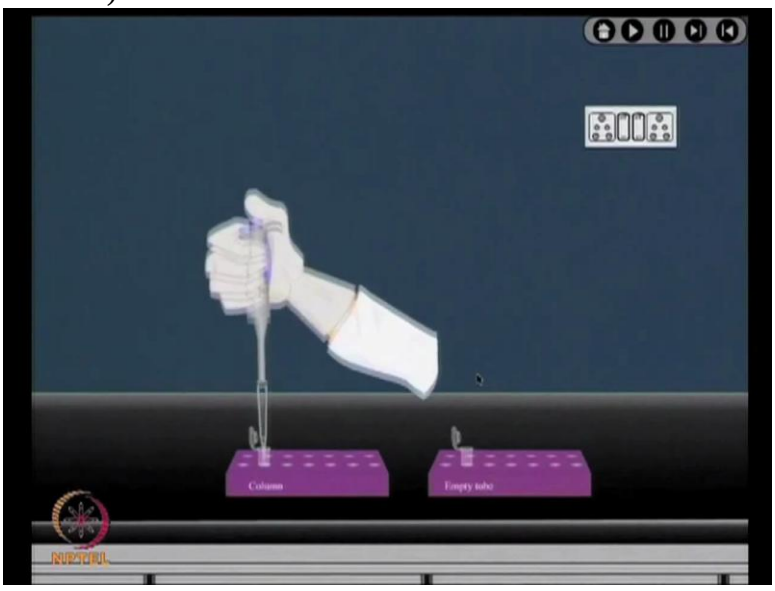


And then, Vortex by clicking on this Vortex for 30 seconds…

(Refer Slide Time: 15:18)

Mild sonication is performed by using a sonicator for 6 cycles of 5-second pulse with 30 second gap in between at 20% of amplitude. In mild sonication step serum sample is found to be effective in improving the gel quality and resolution. Now you can start depletion strategies to remove the abundant proteins.

(Refer Slide Time: 16:19)

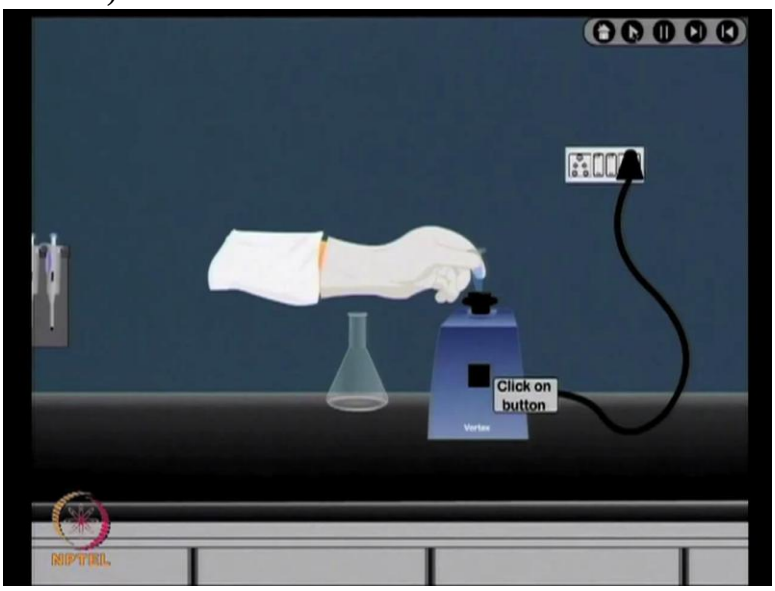


There are commercially available depletion columns. The affinity binders or ligands are immobilized to a solid support the chromatographic medium and used to specifically bind abundant proteins from a complex protein solution.

The depletion strategy effectively enhances the resolution and provides the maximum number of spots on two-dimensional gels thereby a better coverage of the whole serum proteome. The

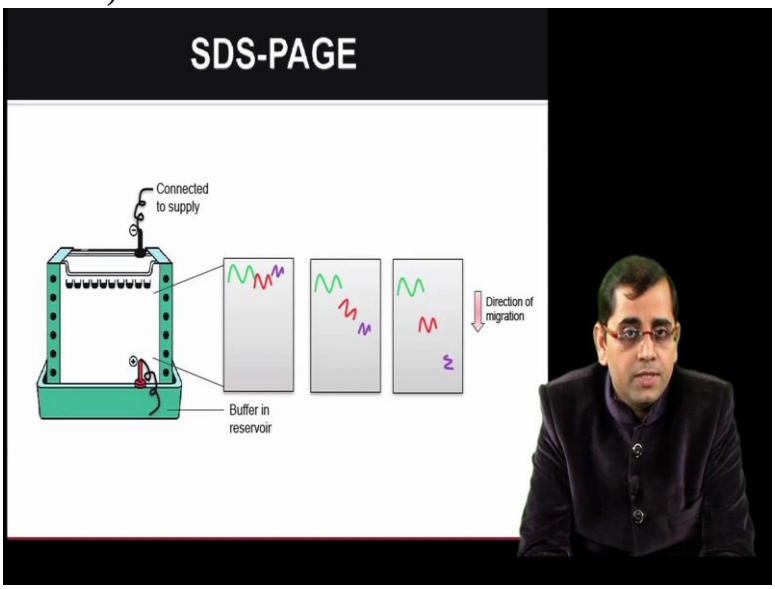
deputed serum is mixed with ice-cold acetone containing 10 % TCA and after Vortexing; the mixture is incubated at -20 degrees for protein precipitation followed by centrifugation. After incubation, the tubes can be centrifuged at 12000 rpm for 15 minutes at 4 degrees.

(Refer Slide Time: 17:42)



The palette can be dissolved in lysis buffer containing 8 molar urea, 2 molar thiourea, 4% CHAPS, 2% IPG buffer, 40 milli-molar of DTT and trace amount of bromophenol blue. You can then again perform the Vortexing step and keep this extract in -80 degrees until further use.

(Refer Slide Time: 18:20)



Fifth lecture of week two will focus on SDS PAGE gel where protein gets separated on the basis of their molecular weight. The low molecular weight proteins having high mobility

migrate further through the gel and high molecular weight proteins remain close to the points of application.

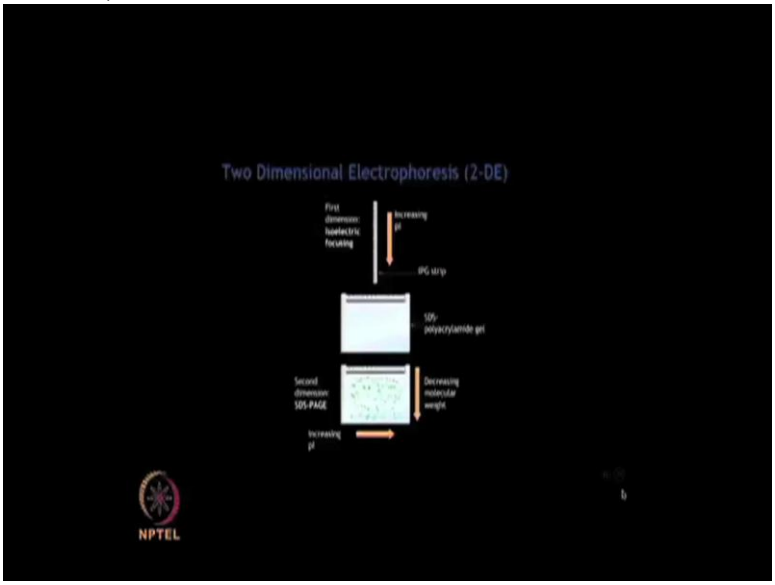
Week 3 will be entirely focused on two-dimensional electrophoresis principle, methodology and challenges involved in performing good two-D gels as well as various applications.

In two-dimensional electrophoresis the first dimensional separation occurs based on the isoelectric point of the proteins and second dimensional separation occurs based on the molecular weight using SDS PAGE.

Isoelectrical focusing works on the principle of isoelectric point which is pH at which the net charge of protein is zero. Two-dimensional electrophoresis remains one of the core technologies to study the proteome and is applied for various biological applications.

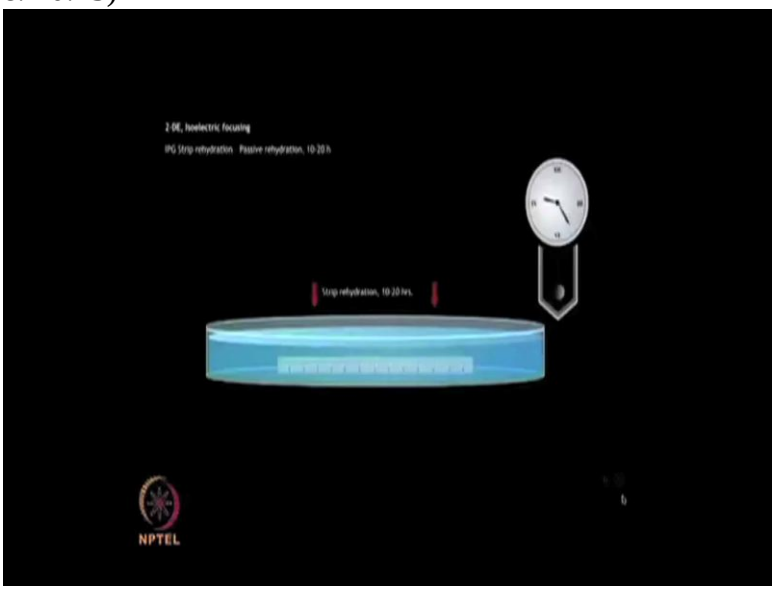
It is relatively easy and affordable technique and it is highly sensitive method to visualize the proteins which makes it widely used method for protein separation. So in this animation I will describe you the two-dimensional electrophoresis process.

(Refer Slide Time: 19:58)



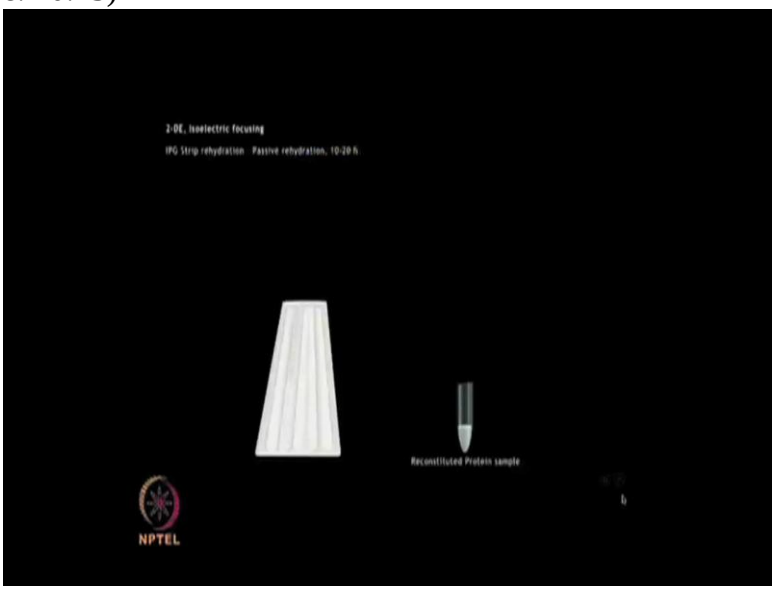
Prior to isoelectric focusing in two-dimensional electrophoresis, the commercially available IPG strip must be rehydrated. This process can be done either by the passive rehydration or active rehydration.

(Refer Slide Time: 20:15)



In passive rehydration, the IPG strip is placed with its gel side downwards in a well containing the protein sample reconstituted with a suitable buffer.

(Refer Slide Time: 20:25)

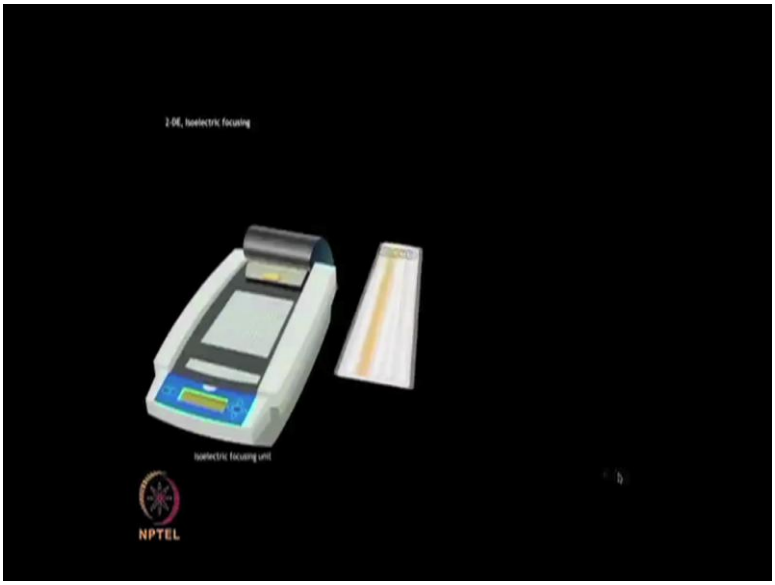


As you can see in the animation, you can apply the protein sample and then add the IPG strip so that it can absorb the protein solutionuvgjetyersb This process can be done for 10 to 20 hours depending on your length of the IPG strip.

This is then covered with mineral oil to prevent the gel from drying up and left overnight; another process known as active rehydration, where the protein sample is added to the strip via a sample cup, followed by cover fluid to prevent the gel from drying up.

The protein sample being applied, a very low voltage is provided and then cover fluid can be added.

(Refer Slide Time: 21:47)



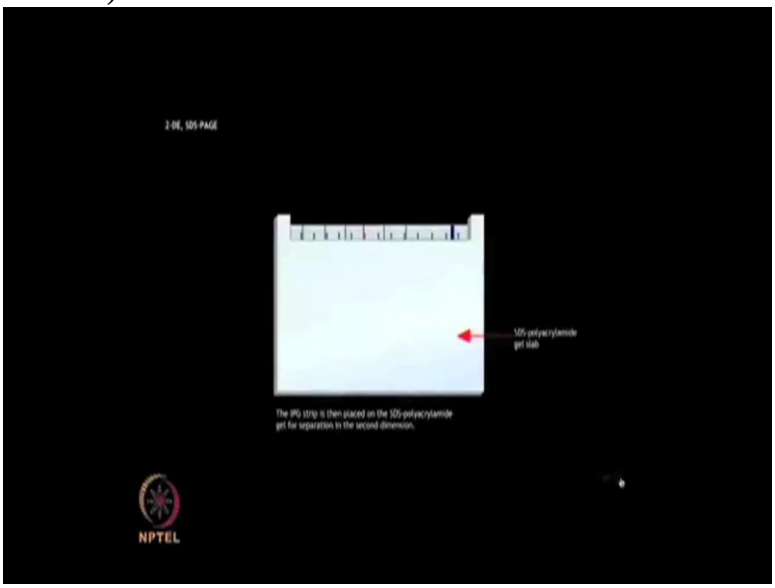
This is then placed in the isoelectric focusing equipment, and low voltage is applied. Process can be performed from 10 to 20 hours. These loaded strips are then focused on an isoelectric focusing unit by passing the current.

(Refer Slide Time: 22:30)

The various proteins of the sample mixture migrate in the electric field and come to rest when the pH is equal to their isoelectric point or pI. So they become neutral and are no longer being affected by the electric field.

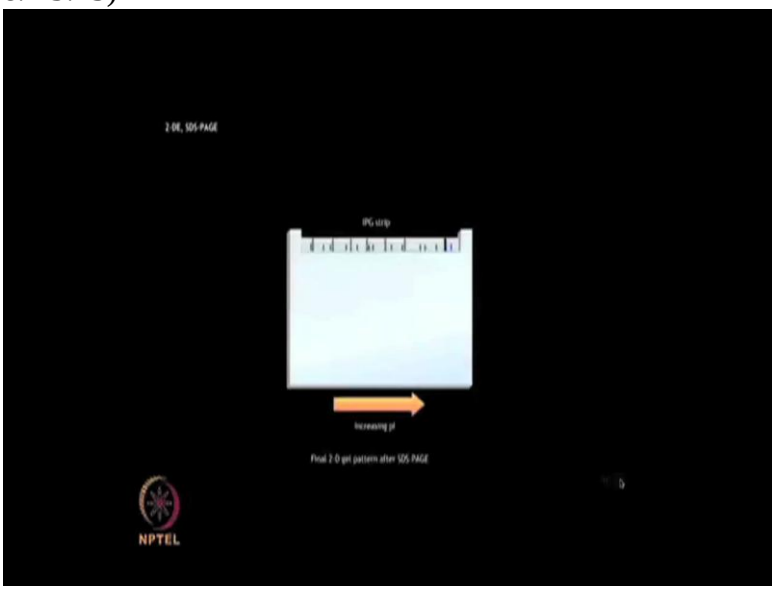
The progress of electrophoresis is monitored by adding a tracking dye which you can see in the animation moving ahead of the proteins. The IPG strip is then equilibrated in a reducing agent such as DTT followed by alkylating agent iodoacetamide which prevents reformation of reduced bonds.

(Refer Slide Time: 23:03)



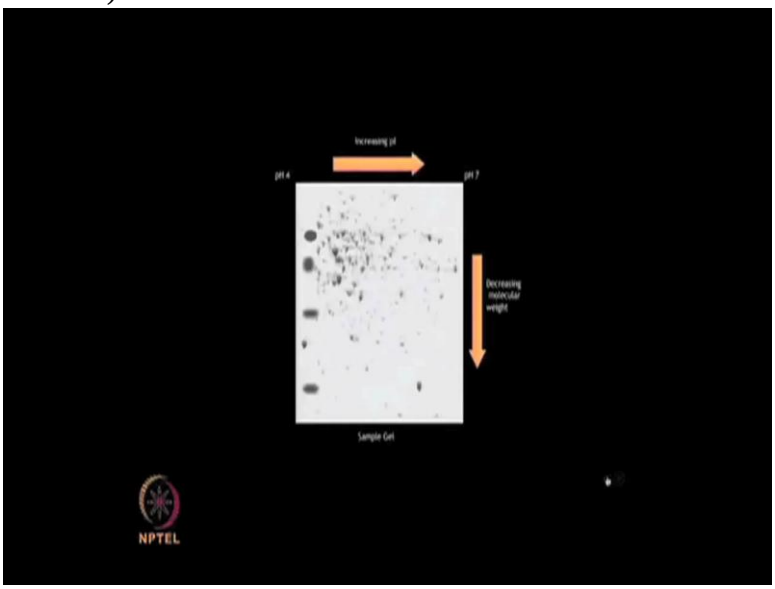
This strip containing the separated proteins is then placed on SDS PAGE gel slab for further protein separation in the second dimension based on the molecular weight. The proteins on the IPG strip are subjected to SDS PAGE by applying a direct current between 100 to 350 volts depending upon the size of the gel.

(Refer Slide Time: 23:25)



Any protein that may have been present as a single band on the IPG strip as you can see in the red circle, due to similar isoelectric point can then be separated based on the molecular weight where the smaller proteins will migrate the farthest.

(Refer Slide Time: 24:05)



This is view of a sample gel which has been run by two dimensional electrophoresis and stained with coomassie blue. Each spot provides the information about the molecular weight and isoelectric point of the protein.

(Refer Slide Time: 24:26)

If protein separation and staining is fine, then at the end we should be able to see good protein separation on gel. But many times the different types of gel issues which could give rise to very bad gel image. These issues could be because of the sample preparation method,

reagents involved into the process, isoelectric focusing, various parameters and different types of staining methods being used.

I will also show you different substandard or bad gels and what could be the possible reasons behind them which can help us in troubleshooting these kinds of artifacts. Some of these issues of gel-based proteomics can be overcome by using advanced method which is known as two dimensional difference electrophoresis or DIGE.

The two-dimensional electrophoresis can be used for large number of biological applications. These include studying the global protein expression which is the separation and identification of all the proteins present in the given sample, comparison of protein abundance, fitting the test and control samples. This is one of the very common applications also known as differential protein expression or abundance based proteomics.

Resolution of proteins from the complex mixture is separated using 2DE. Studying different types of isoforms or the post translation modification is another application which is commonly used. One of the major advantages of this technique is the visual analysis of the proteins on gel which is not possible from the gel-free approaches such as mass spectrometry.

In the last module in week 4, we will talk about more advanced and sensitive gel-based technique known as two dimensional fluorescence difference electrophoresis or 2D DIGE. There are 3 different Cy dyes, Cy2, Cy3 and Cy5 which are used for the DIGE experiments. Excitation and emission wavelength of these three are defined and used during the scanning of the gels. In this lecture we will talk about two different types of labeling strategies, lysine labeling or system labeling.

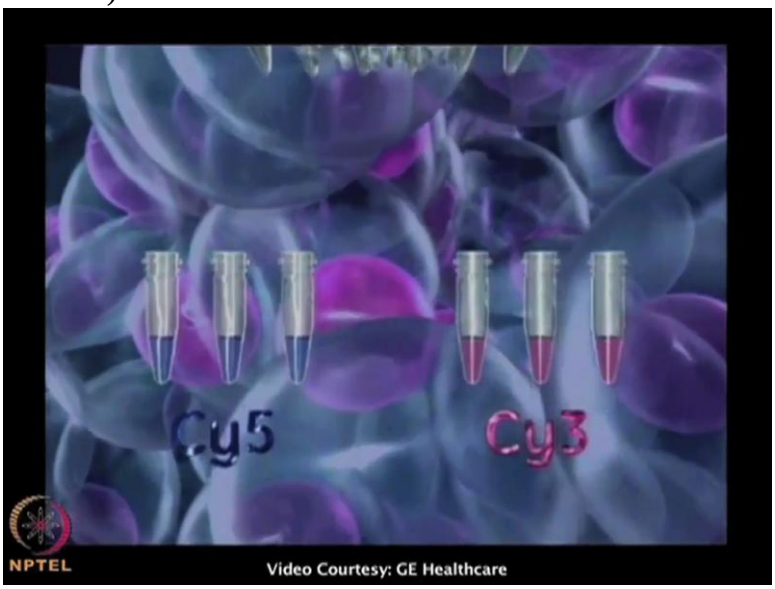
Lysine labeling method is also known as minimum labeling method, where the charge and size of dyes are matched. When to use which type of labeling method and the rationale of using these labels will be discussed in this lecture.

In next lecture we will talk about overall steps involved in performing DIGE. While most of the steps are exactly the same as what we have learnt in two dimensional electrophoresis but major variation will be in the preparation of internal standard.

Internal standard all your control and treatment samples will be mixed together and it will be separated on one IPG strip. The DIGE gel can be scanned using a fluorescence scanner such as DIGE imager and Cy3, Cy2 and Cy5 patterns can be obtained. After scanning from same gel, three separate images are generated.

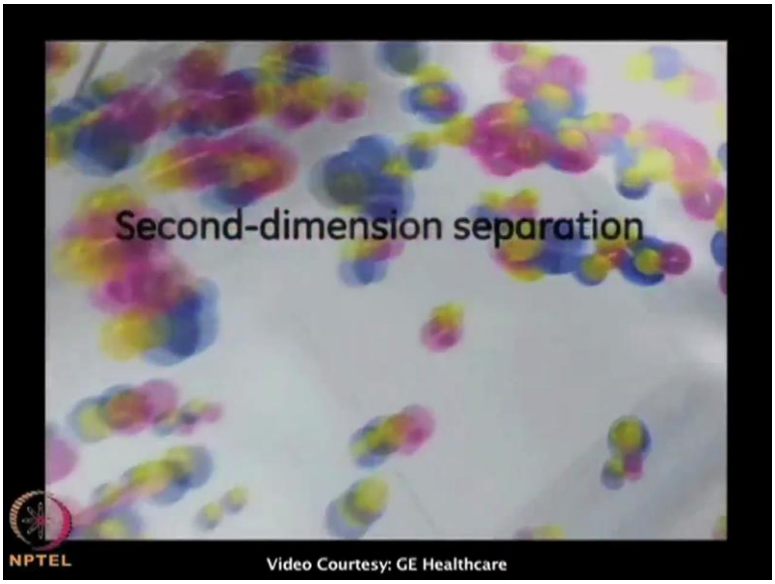
Further a combined image from mixing all the 3 dyes and the overlap images are generated. In overlap image, the red and green are indicative of over expression or under expression of a given protein, and the white or yellow color represent there is no differential response. The DIGE image analysis is very robust and automated as compared to the two-dimensional electrophoresis analysis.

(Refer Slide Time: 29:35)



Yes, as we were talking about the three samples we can label with Cy2, Cy3, Cy5, now we have 2 different samples from two different source, now you are taking 3 samples into 3 different eppendorf tubes, and you are labeling with Cy3 dye and as well as again another with Cy5, then you are taking the two samples, cooling them and labeling them with Cy2, then you are mixing those 3 into a single tube, then you are running in a single strip.

(Refer Slide Time: 30:00)

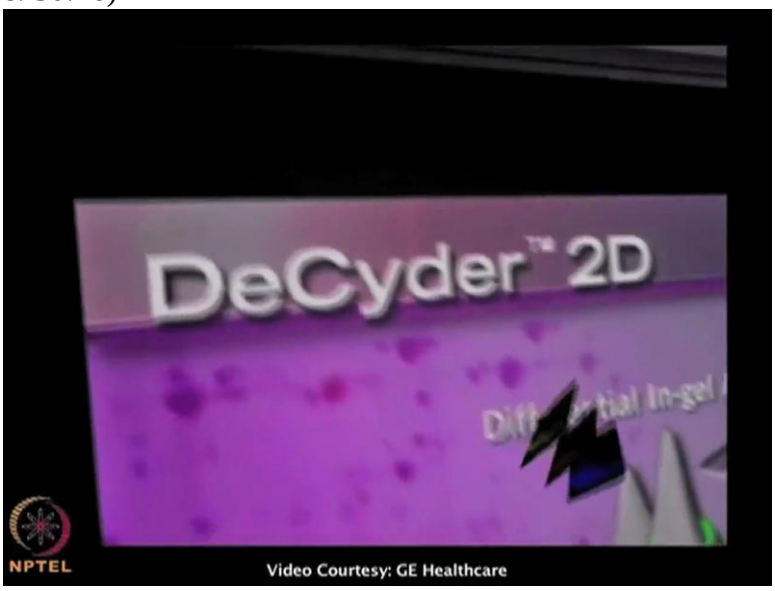


There you will be analyzing first dimension, the IEF where as in first dimension, after analyzing IEF, you will be placing it to second dimension where you have completely analyzed the second dimension.

(Refer Slide Time: 30:05)

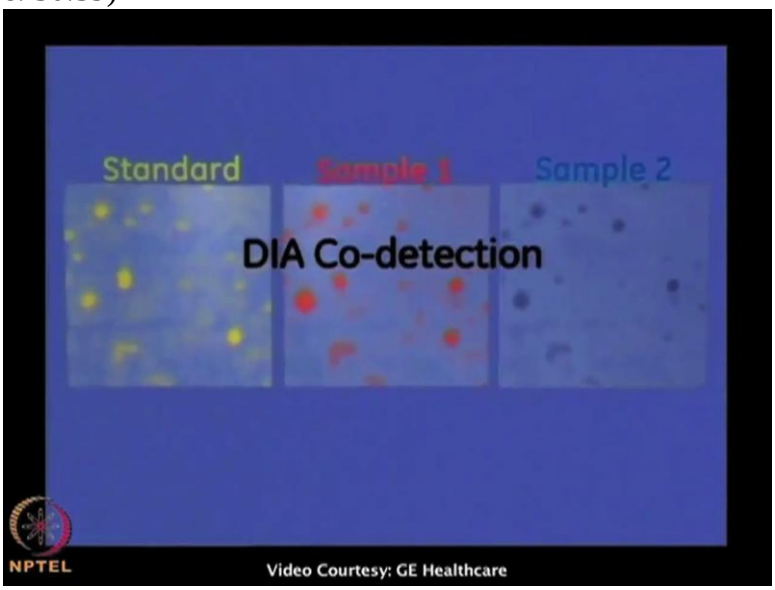
Now this gel can be scanned in the Typhoon Triplus where you will have 3 images from a single gel.

(Refer Slide Time: 30:16)



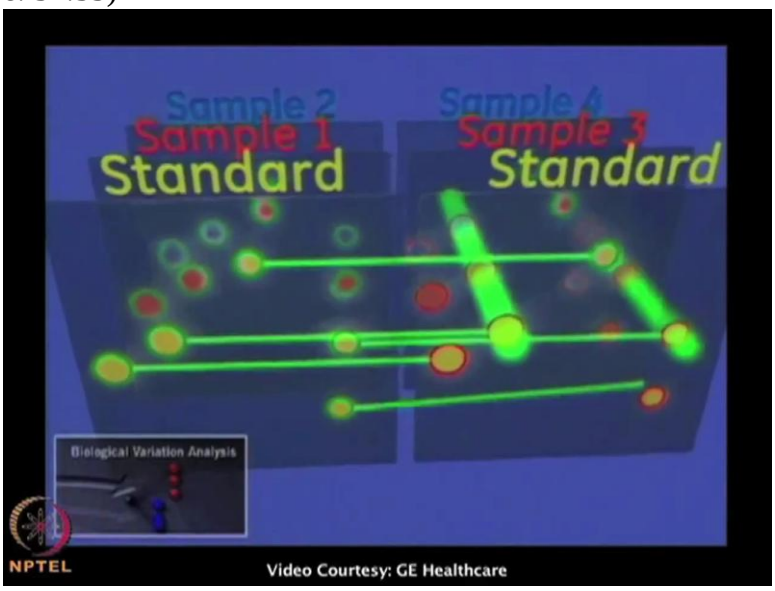
Then this can be analyzed with the help of Decyder software, then...where the complete analysis can be done with the help of Decyder.

(Refer Slide Time: 30:39)



Then in DIA there will be a co-detection. The co-detection nothing but there is a standard gel labeled with Cy2 from the same area where expanding remain 2 images also. So the artifacts can be minimized. Again this is fully automated analysis. There will be very less manual interference would be there in this analysis.

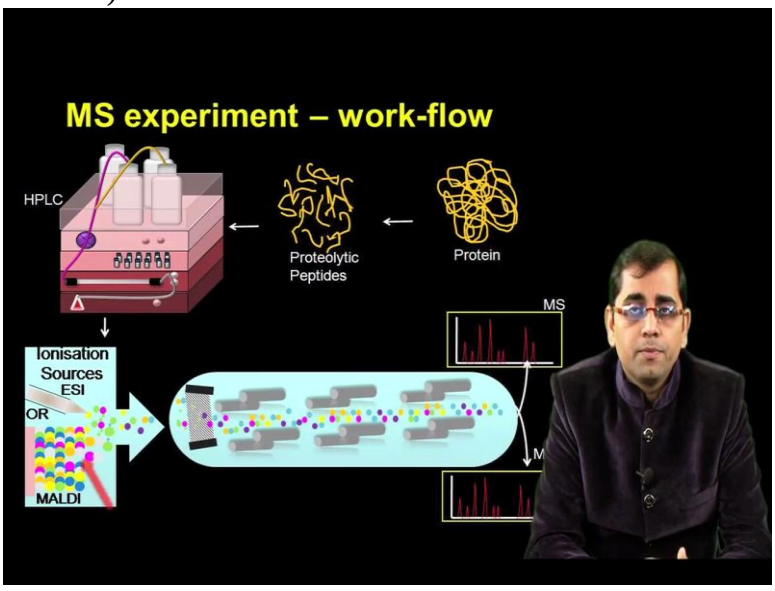
(Refer Slide Time: 31:35)



And after finishing of DIA, we can go to BVA. In BVA we can compare all DIAs together as like you can see in this video, this complete...first of all it matches between the standard gel and from standard gel to again to the corresponding DIA it matches each and individual spot and it will give you the final data.

Then these BVAs can be analyzed in the further EDAs, that is, Extended Data Analysis. Here we will have lots of stringent...stringent statistical data, that is Differential Expression Analysis, Pattern analysis, Discriminant analysis and PCA analysis.

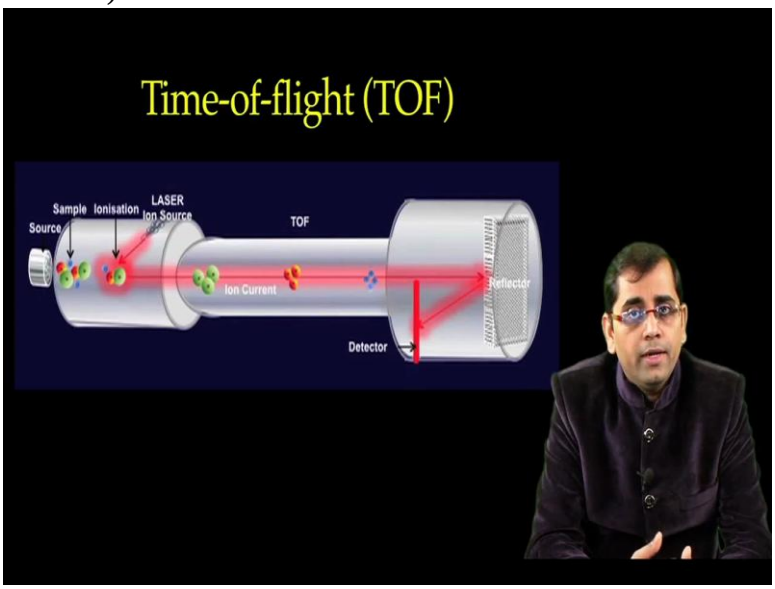
(Refer Slide Time: 32:54)



Once you have identified the protein of interest using gels you would like to identify them using mass spectrometry. So the last 2 lectures of this course will focus on MALDI based mass spectrometry.

The mass spectrometry has the ability to provide the platform for comprehensive coverage of proteome and it has become a very important analytical tool in biology in general and in proteomics during the last decade. In this lecture we will talk about the concepts of MALDI.

(Refer Slide Time: 33:15)



How the time of flight analyzer can be used to resolve the ions produced by ionisation source on the basis of their mass to charge ratio. Briefly we will also talk about the data analysis and protein identification using Mascot

Overall this course will provide you the basic knowledge of gel based proteomic techniques with a special emphasis on two-dimensional electrophoresis and difference in gel electrophoresis or DIGE.

Hope this information and knowledge acquired from this course will be useful to you whether or not you work in proteomic research but these techniques can be used for several applications. I hope this course will be a stimulating experience for you. Thank you.