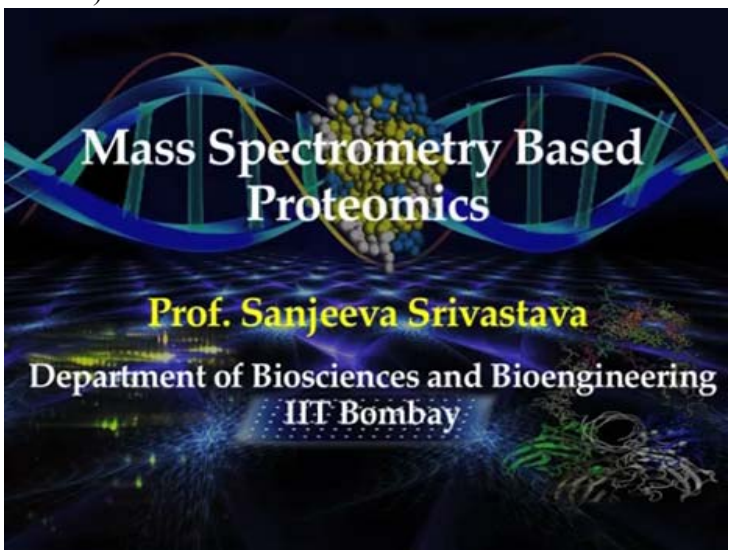
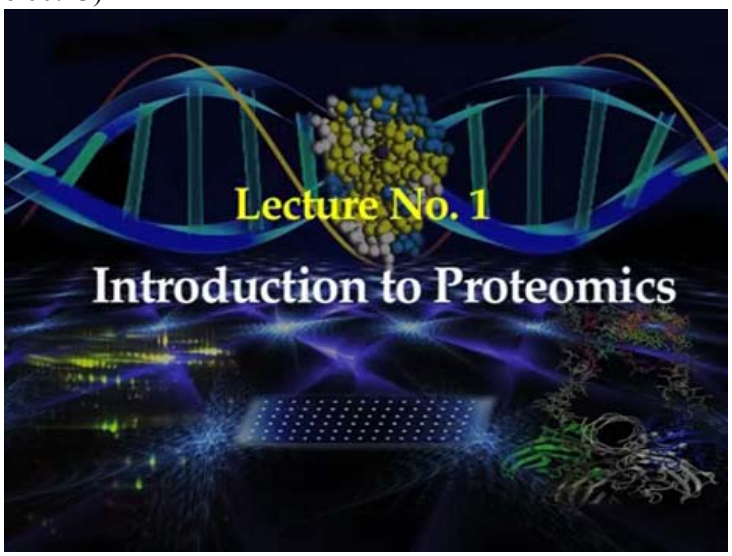
Mass Spectrometry Based Proteomics
Professor Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay
Mod 01 Lecture Number 01

(Refer Slide Time 00:11)



(Refer Slide Time 00:13)

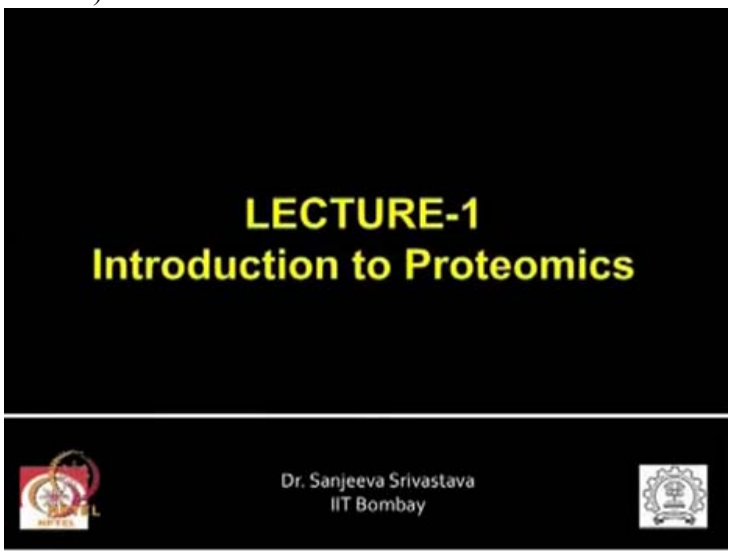


(Refer Slide Time 00:16)

(Refer Slide Time 00:23)

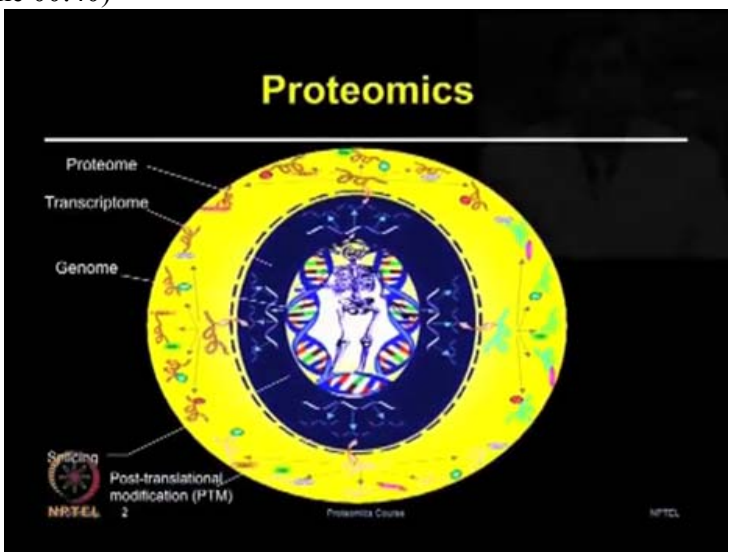
Welcome to course on Proteomics. My name is Sanjeeva Srivastava and I am in Department of Bio-Sciences and Bio-Engineering of IIT Bombay.

(Refer Slide Time 00:35)



In this introductory lecture, I will discuss about proteomics.

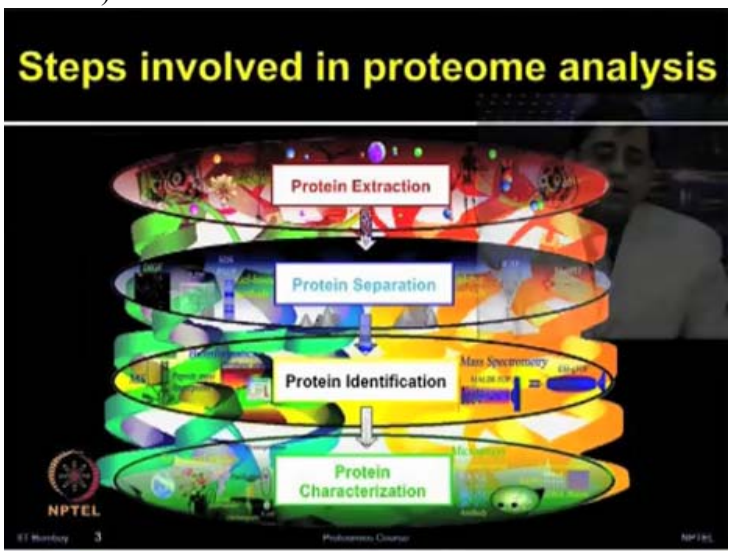
(Refer Slide Time 00:40)



So first of all, what is proteomics? The proteome describes the proteome complement expressed by the genome, or more precisely we can say the protein complement of a given cell at a given time including the set of all protein isoforms and its modifications. The study of entire compendium of proteins which are encoded by the genome is known as proteomics.

In this slide I have illustrated the complexity of human proteome as compared to the genome or transcriptome, the extent of diversity and complexity due to alternative slicing and post-translational modifications is tremendous. Therefore studying proteins and proteomes are very important.

(Refer Slide Time 01:53)



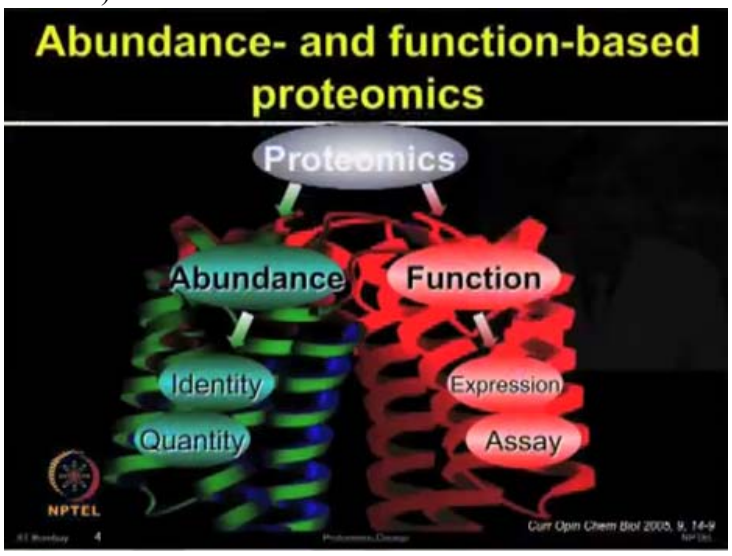
What are different steps involved in the proteome analysis? As shown here, the protein extraction, protein separation, protein identification and protein characterization, these are the major steps which are involved in proteome analysis.

The protein extraction from whole cells, tissue or organism is first requirement for proteome analysis. Protein separation and quantification is achieved by various proteomic techniques including gel-based techniques such as two-dimensional electrophoresis and gel-free techniques such as iTRAQ mass spectrometry based techniques.

The functional characterization of proteins using novel proteomic platforms open new horizon for exploration in biology.

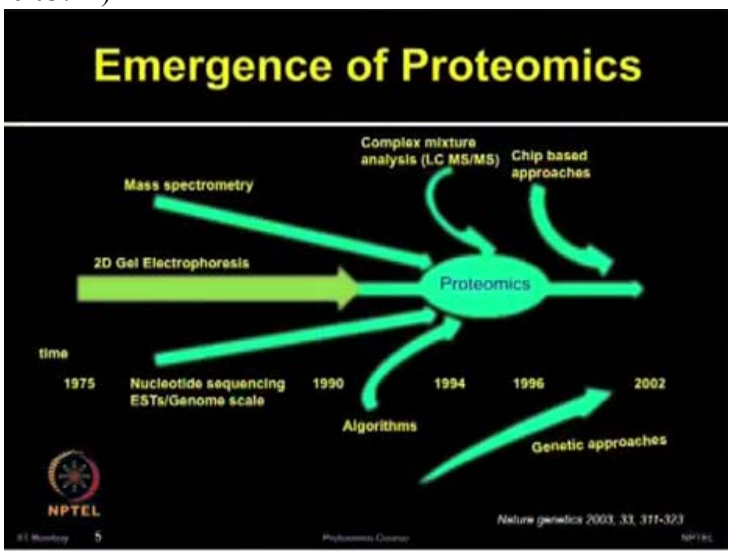
The proteomic discipline can be grouped under two major disciplines, abundance- and function-based proteomics.

(Refer Slide Time 03:11)



The abundance based proteomics aims to measure the abundance of protein expression where as the functional proteomics aims to determine the role of proteins by addressing protein interactions and their bio-chemical activities. So how did proteomics field emerge?

(Refer Slide Time 03:42)

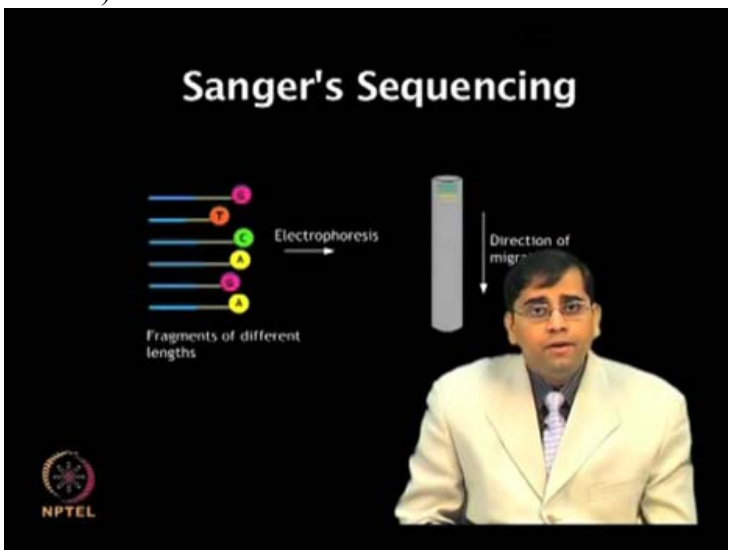


As you can see in the time scale here shown in the slide, advancement of various techniques such as two-dimensional electrophoresis and mass spectrometry, genome sequencing information and computational algorithm together led to the emergence of proteomics field.

(Refer Slide Time 04:01)

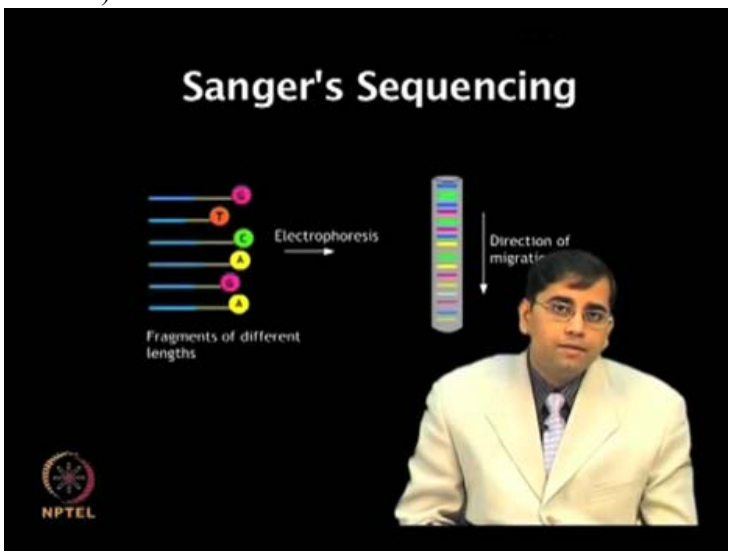
Studying genome of an organism by studying sequencing and genome mapping is known as genomics

(Refer Slide Time 04:16)



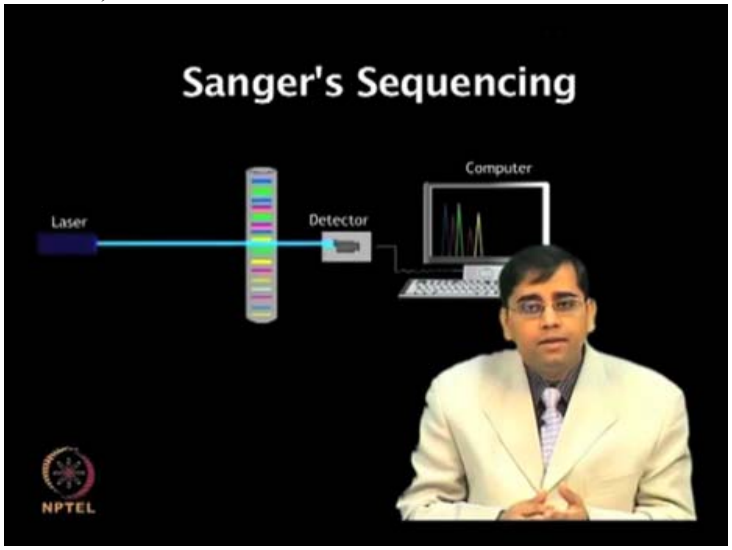
. Several genomics sequencing projects …

(Refer Slide Time 04:18)



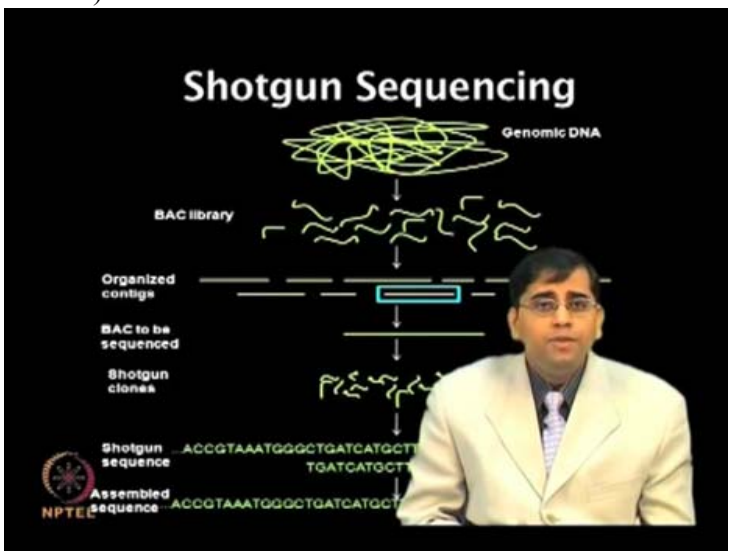
… that aim to elucidate …

(Refer Slide Time 04:21)



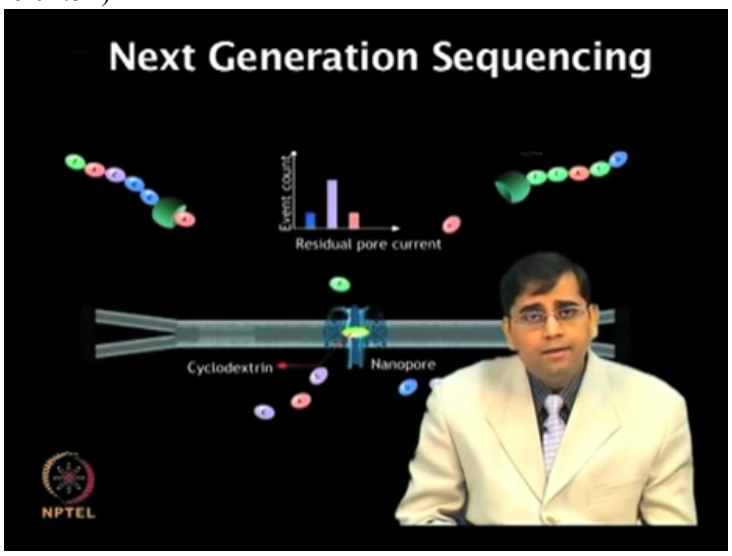
… the complete genome sequence of organisms have been undertaken by several research groups all over the world

(Refer Slide Time 04:33)



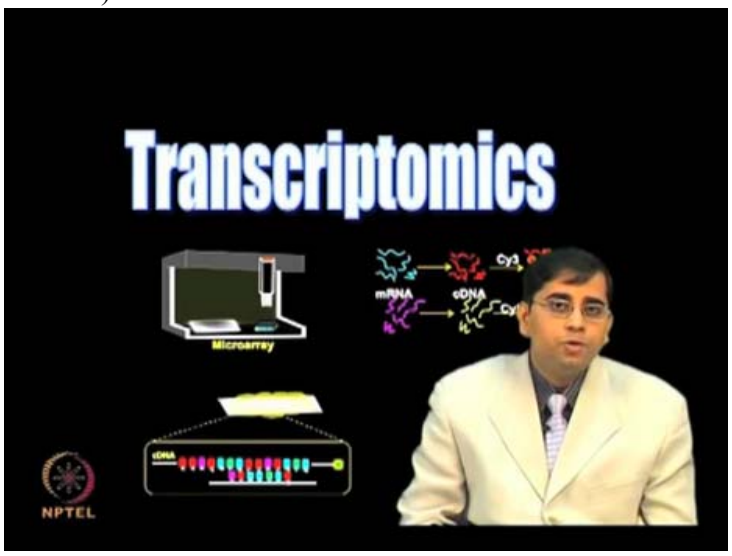
From a genomic library, clones were isolated and ordered into detailed physical map. Further individual clones were sequenced by Shotgun sequencing to provide complete genome sequence

(Refer Slide Time 04:54)



Recently Next Generation Sequencing NGS spreadsheets have dramatically increased the pace of sequencing by several orders of magnitude. Next Generation Sequencing based on nano-pore structures is known as nano-pore sequence.

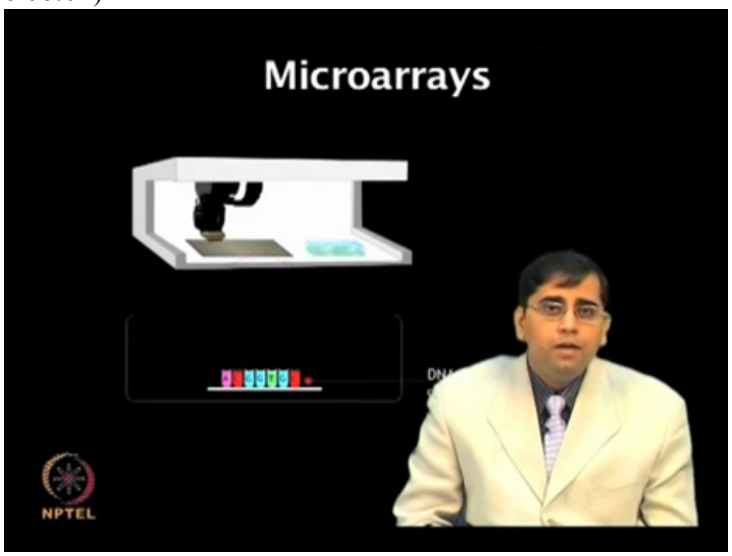
(Refer Slide Time 05:18)



Transcriptomics, the study of all the mRNA molecules expressed by a particular cell type of an organism is known as transcriptomics.

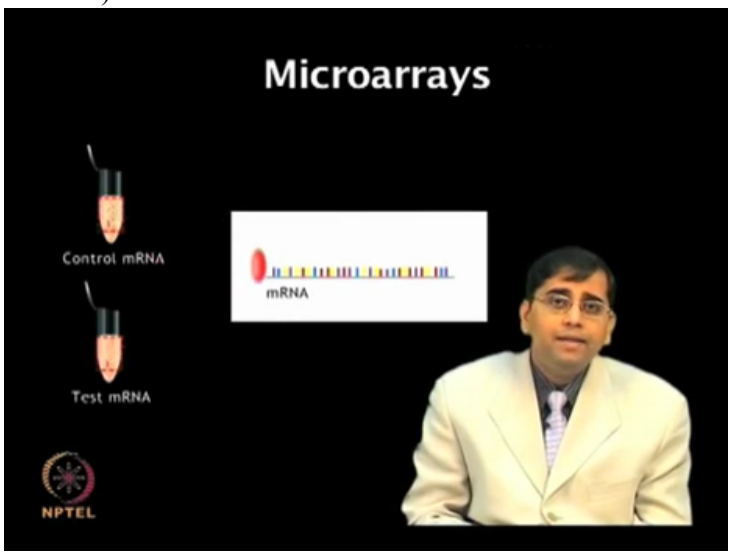
The transcriptomic analysis measures the genes that are being actively expressed at any given time and varies significantly with external environmental conditions. Various techniques such as Microarrays, qRT-PCR etc have been widely used for transcriptional analysis.

(Refer Slide Time 06:02)



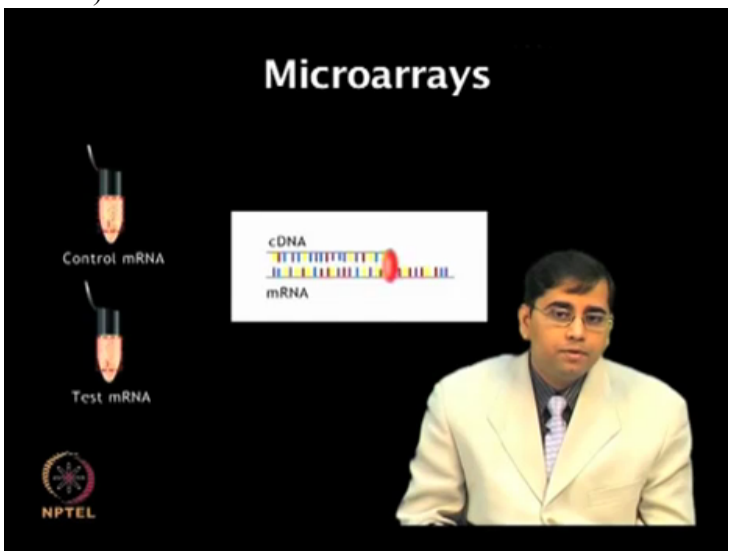
In a Microarray experiment, the mRNA from control and test samples …

(Refer Slide Time 06:08)



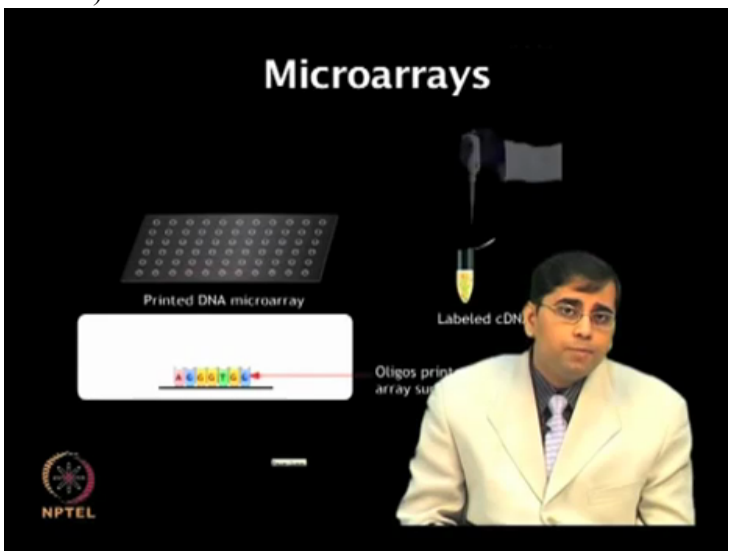
… are extracted and …

(Refer Slide Time 06:12)



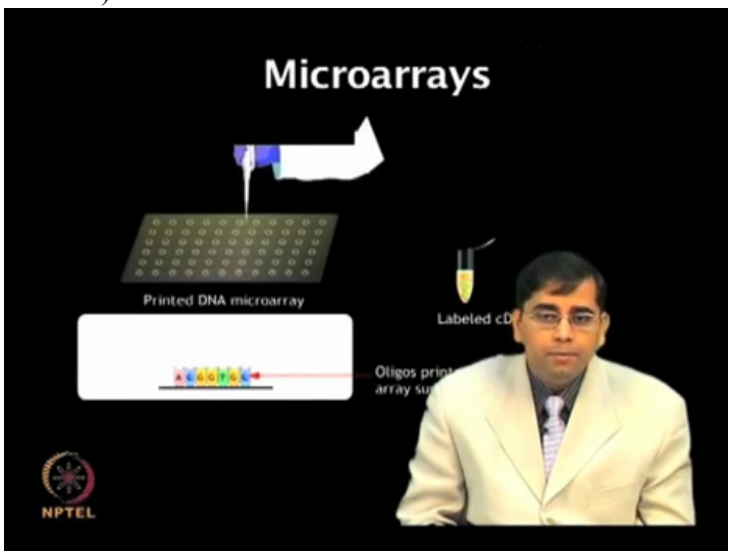
… reverse transcribed into ….

(Refer Slide Time 06:14)



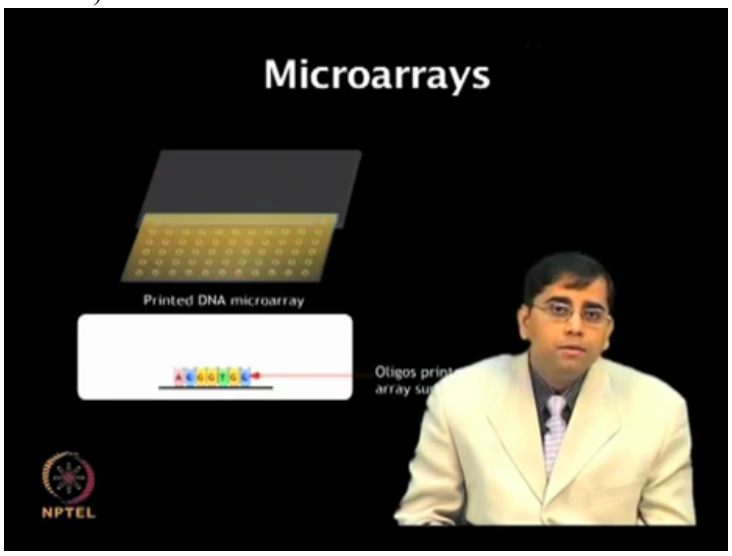
… its corresponding cDNA

(Refer Slide Time 06:19)



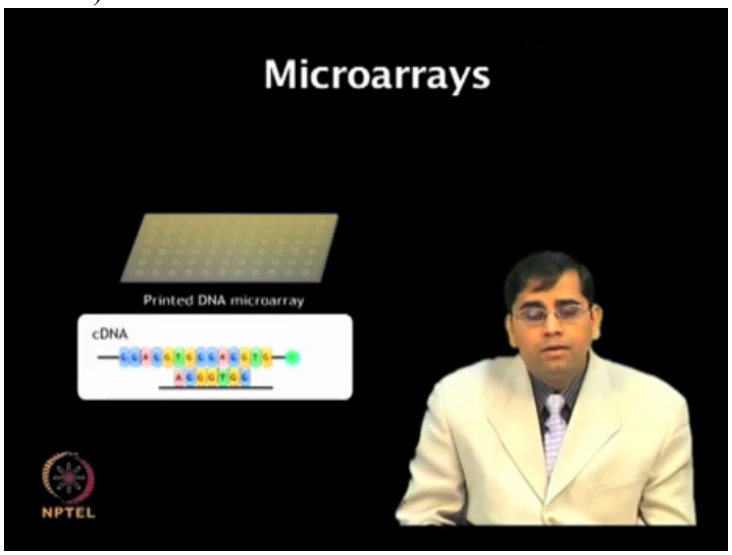
The cDNA samples are labeled with Cy5 and Cy3 dyes and mixed cDNA sample is incubated on …

(Refer Slide Time 06:30)



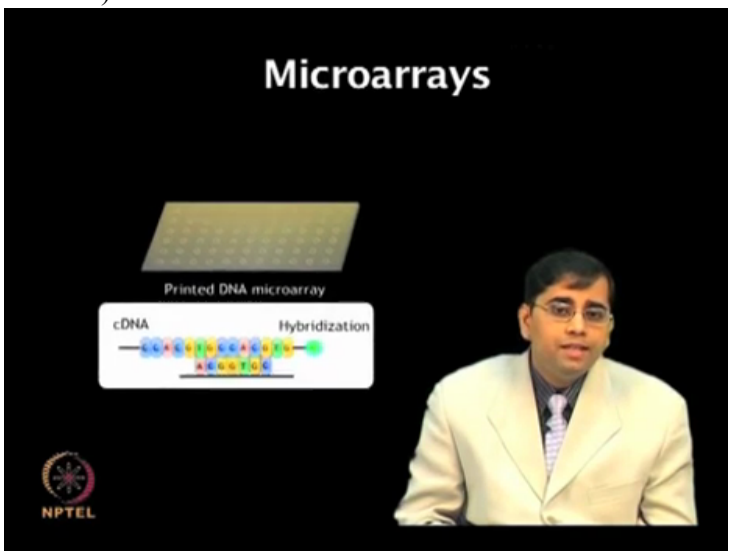
… printed DNA Microarray

(Refer Slide Time 06:34)



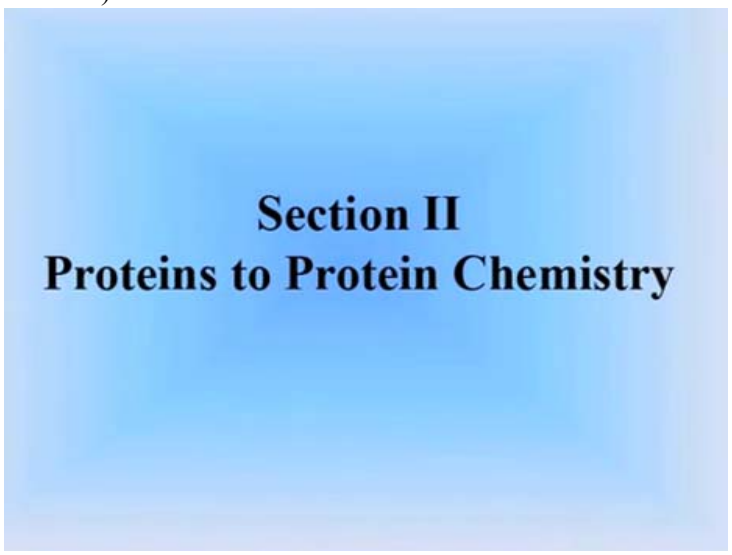
This allows hybridization to occur between the probed oligonucleotide on array surface and the labeled cDNA sample of interest. In this manner, expression level of ….

(Refer Slide Time 06:52)

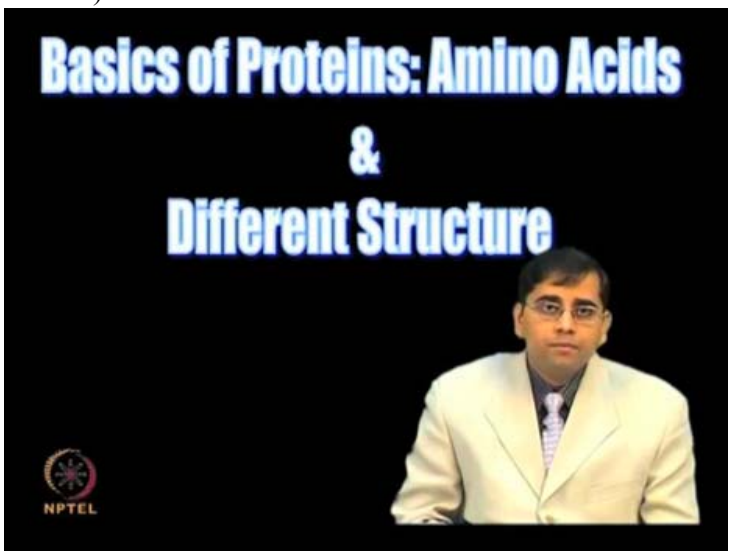


… thousands of genes can be measured and analyzed simultaneously

(Refer Slide Time 06:59)

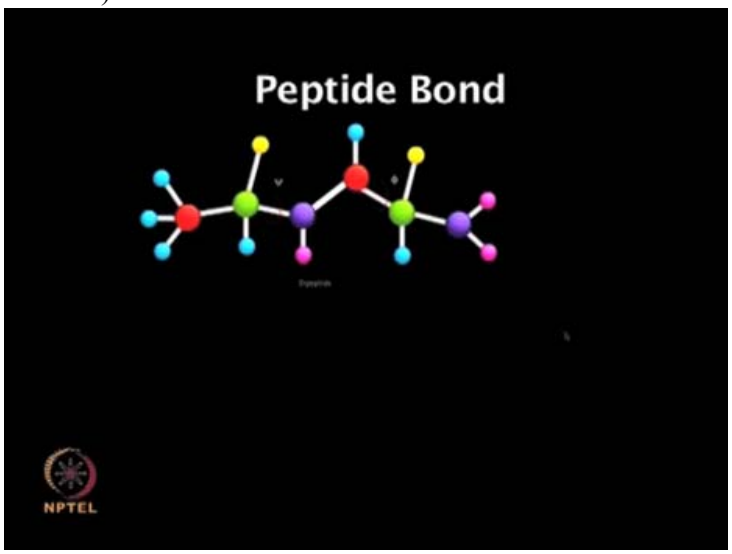


(Refer Slide Time 07:04)



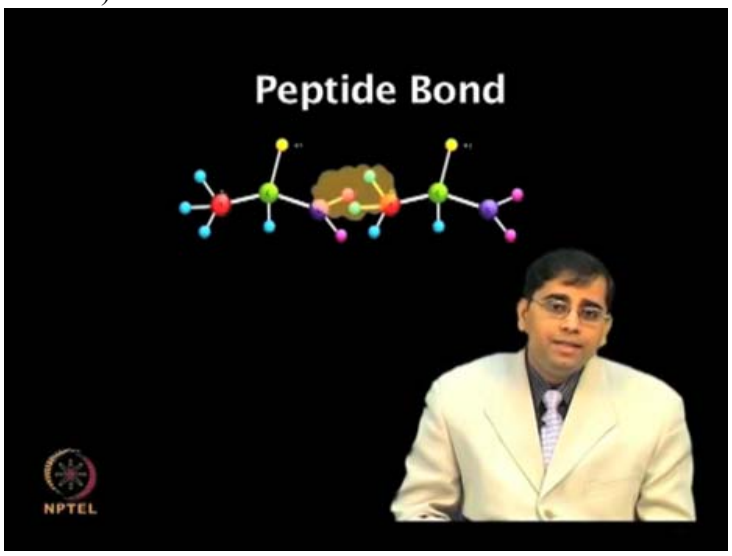
Proteins play an important role in essential characteristics of living systems, how they function …

(Refer Slide Time 07:14)



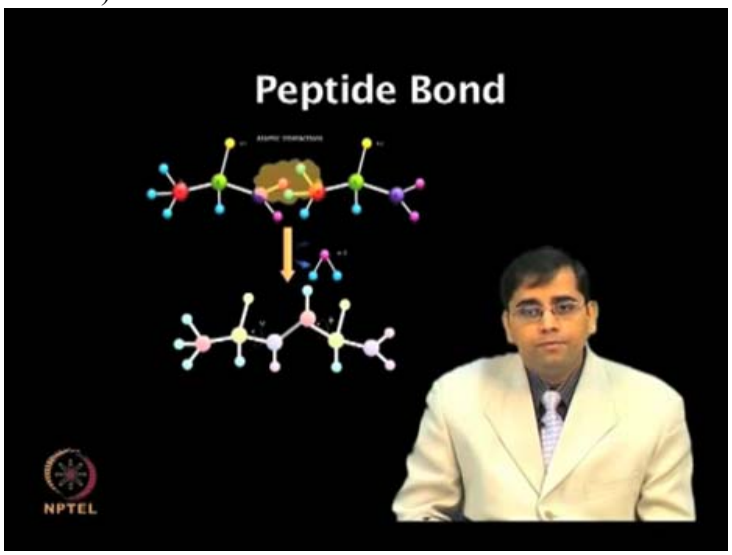
…and replicate themselves through intricate molecular interactions

(Refer Slide Time 07:22)



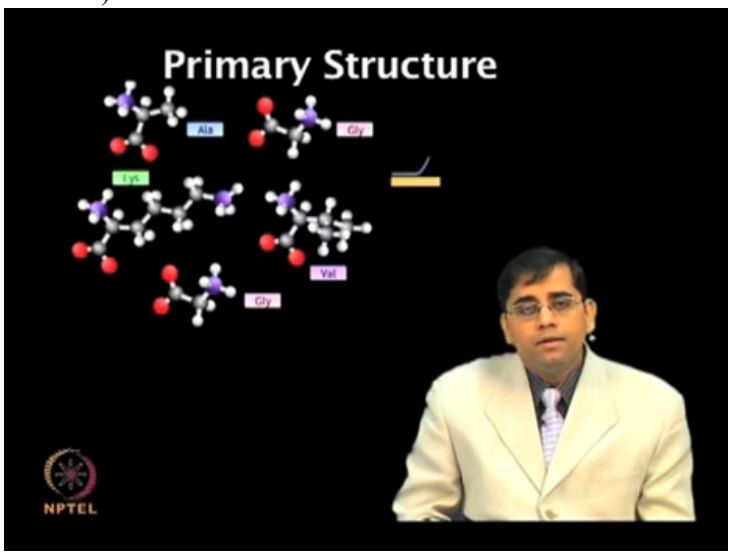
Amino acids constitute the basic monomeric units of proteins ….

(Refer Slide Time 07:28)



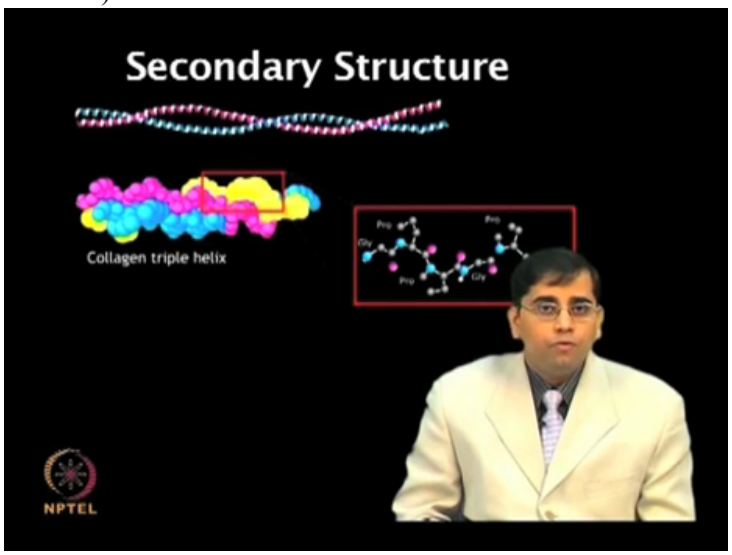
... which are joined together by peptide bonds

(Refer Slide Time 07:35)



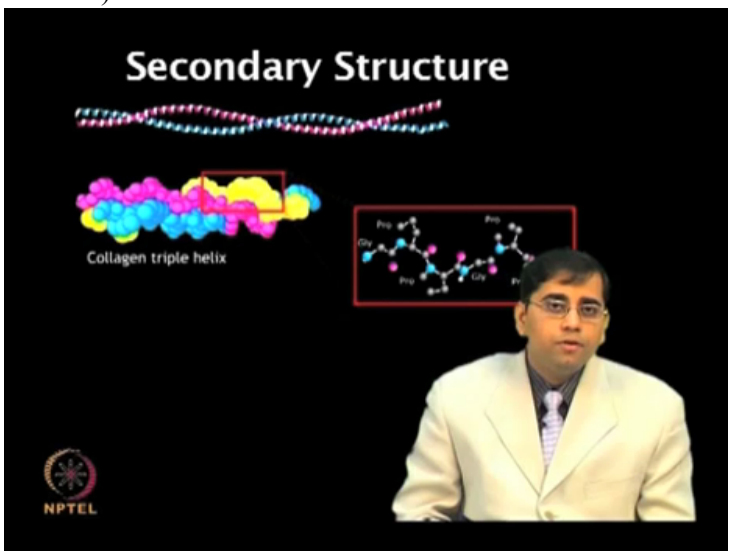
The linear sequence of amino acids constitutes the primary structure.

(Refer Slide Time 07:43)



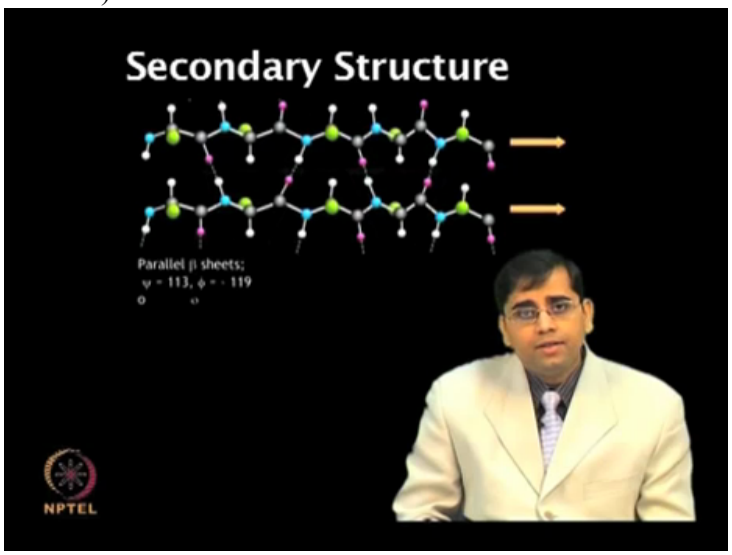
Folding of polypeptides …

(Refer Slide Time 07:44)



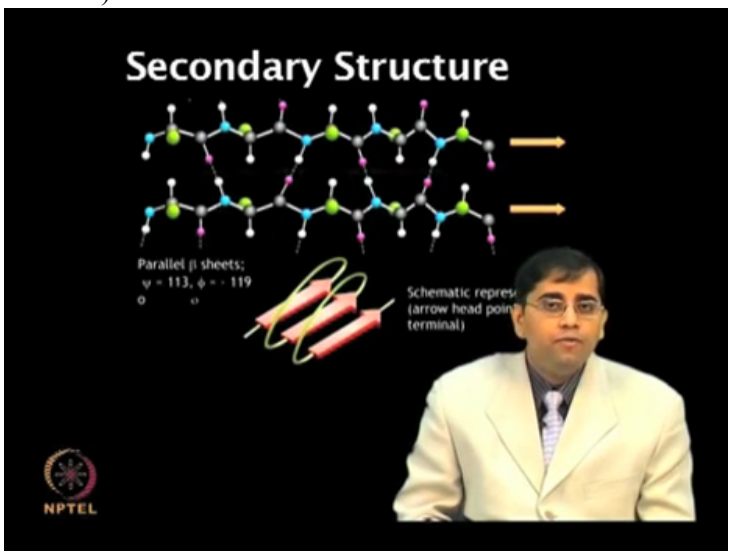
… or protein chains into regular structures

(Refer Slide Time 07:49)



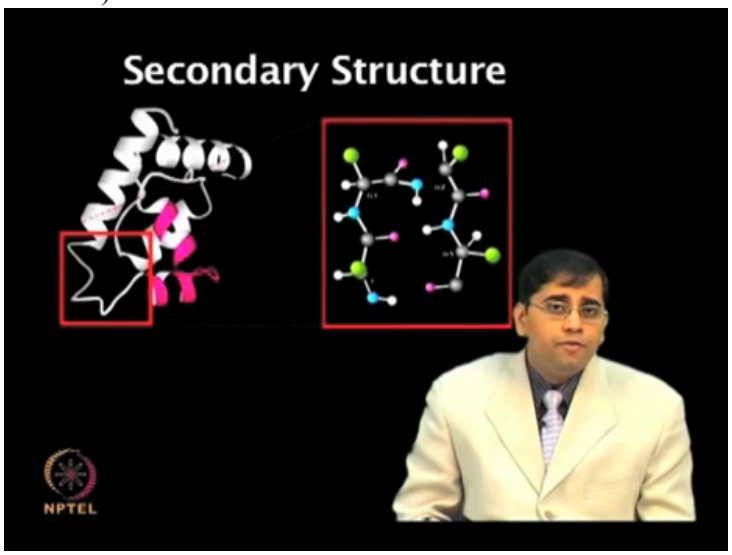
…. like alpha helices ….

(Refer Slide Time 07:51)



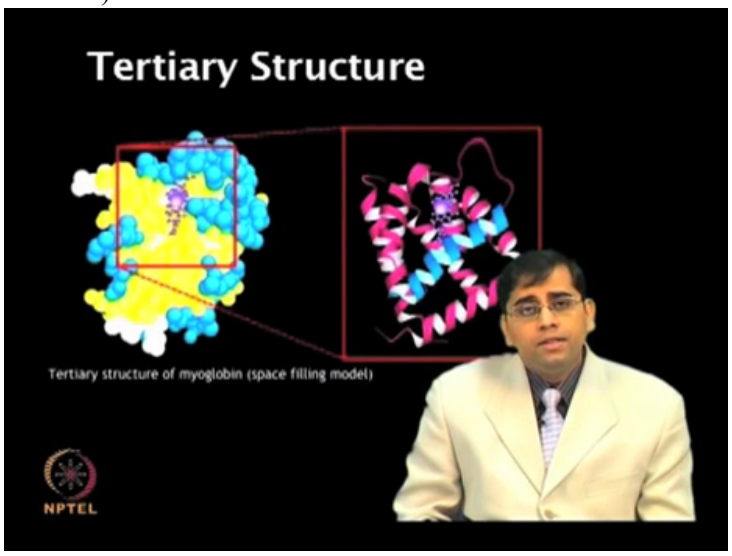
… beta sheets, turns and loops …

(Refer Slide Time 07:56)



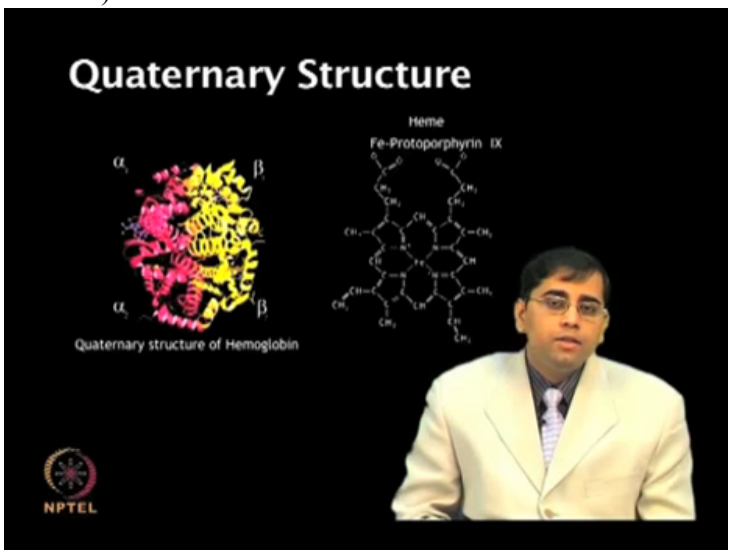
… give rise to secondary structure.

(Refer Slide Time 08:01)



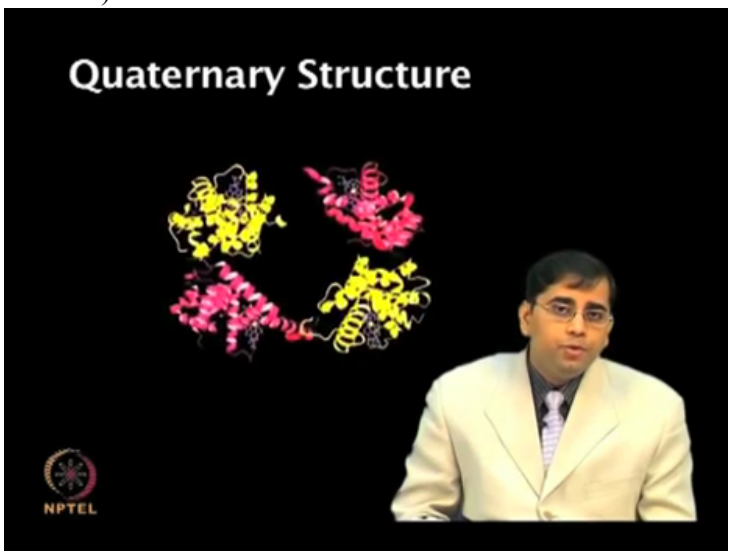
The three-dimensional compactly folded structure of proteins makes tertiary structure which represents overall organization of secondary structural elements in three-dimensional space.

(Refer Slide Time 08:21)



The quaternary structure refers to the interactions between individual proteins …

(Refer Slide Time 08:28)

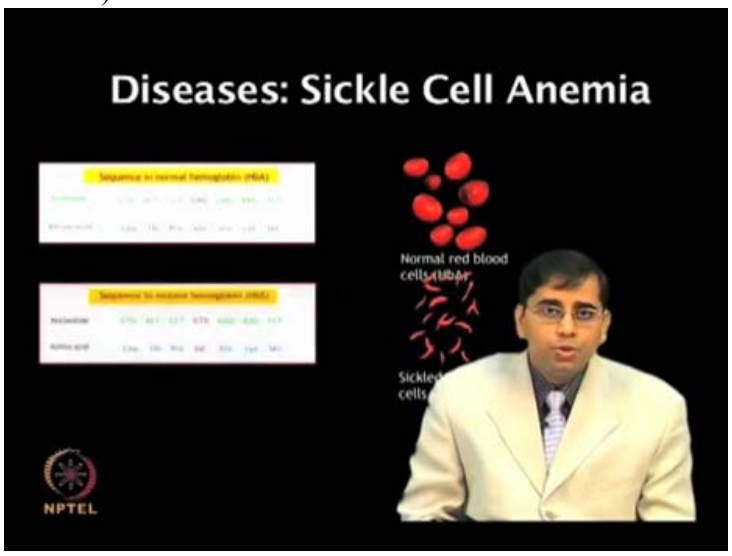


… subunits in a multi-sub-unit complex

(Refer Slide Time 08:34)

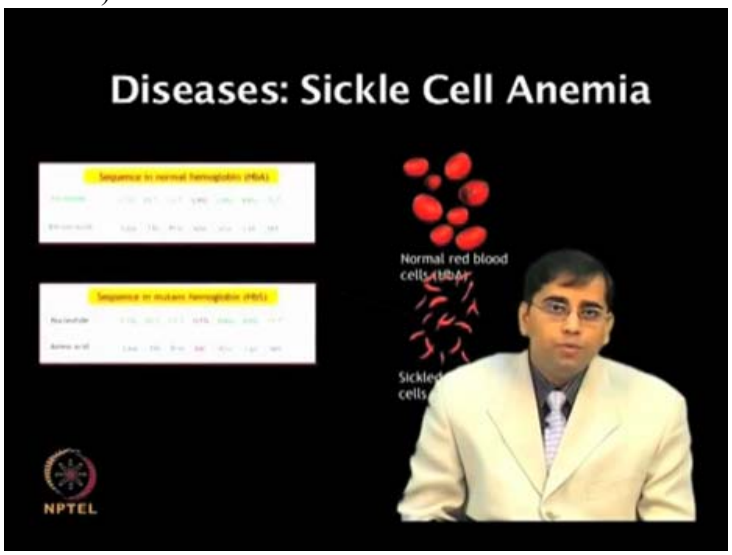
Sickle cell anemia is caused due to ….

(Refer Slide Time 08:38)



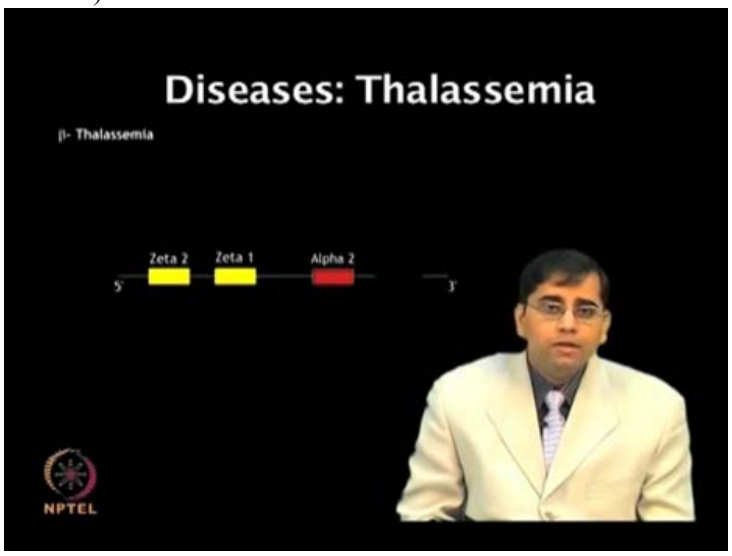
…a single nucleotide substitution …

(Refer Slide Time 08:39)



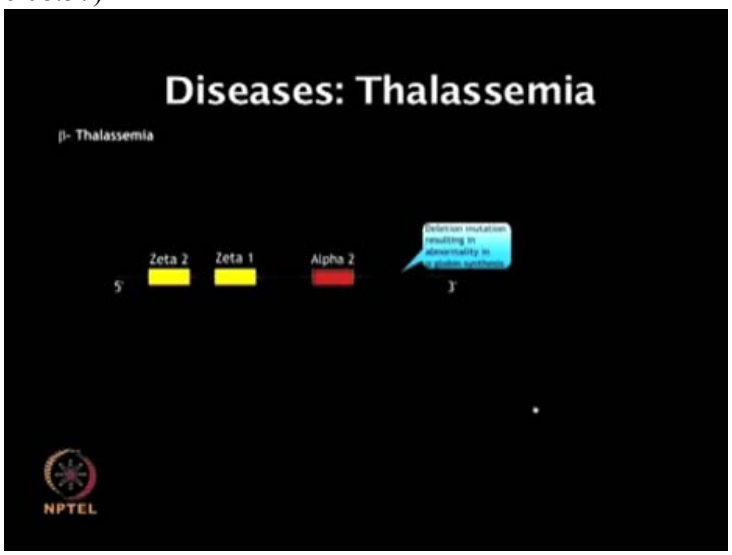
… which converts glutamic acid residue to valine in the beta chain of hemoglobin

(Refer Slide Time 08:51)



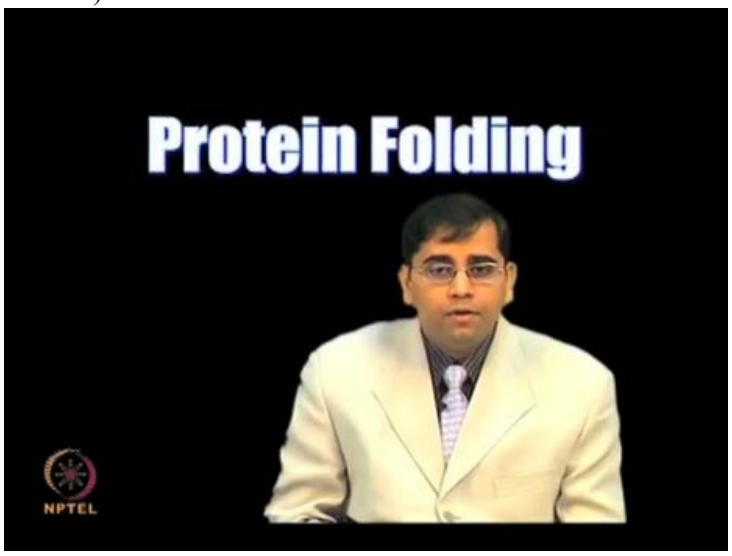
Thalassemia is caused due …

(Refer Slide Time 08:57)



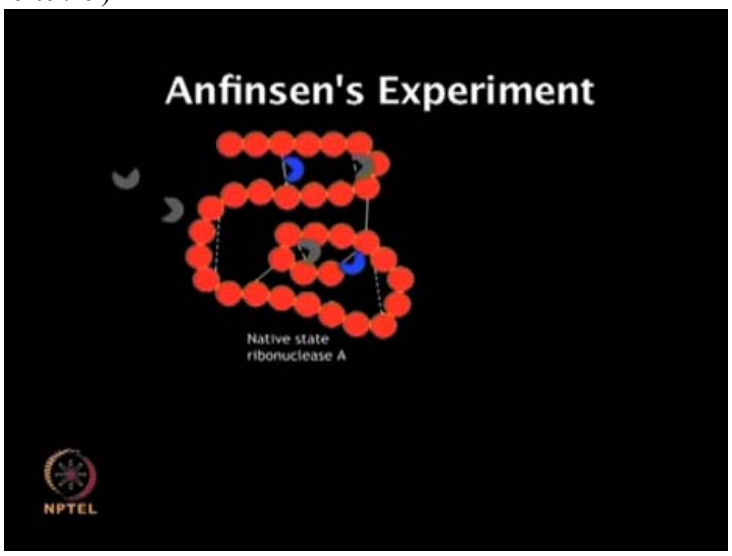
… to abnormalities in hemoglobin synthesis

(Refer Slide Time 09:02)



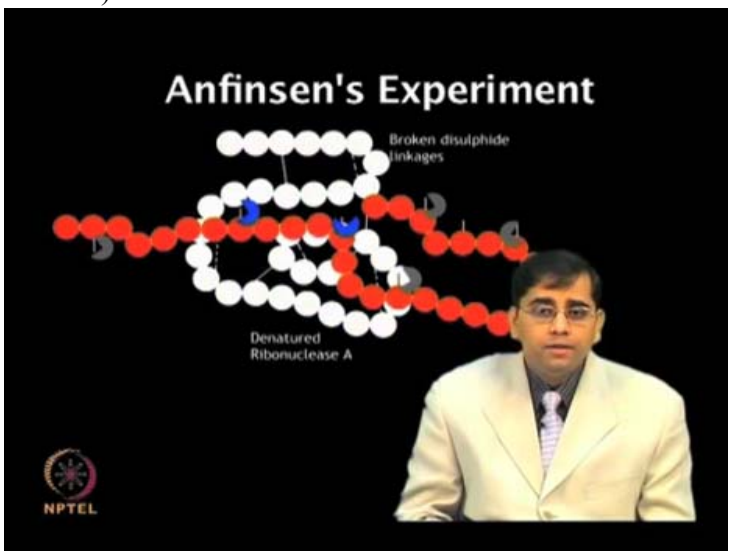
Protein folding is an elegant example of biological self-assembly. Understanding the mechanisms through which protein folding takes place remains challenging for the scientific community.

(Refer Slide Time 09:19)



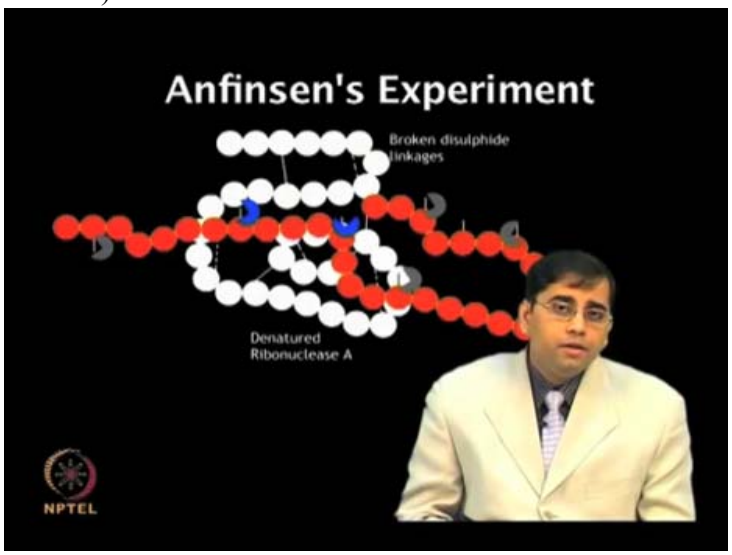
Anfinsen tested the ability of reduced and …

(Refer Slide Time 09:26)



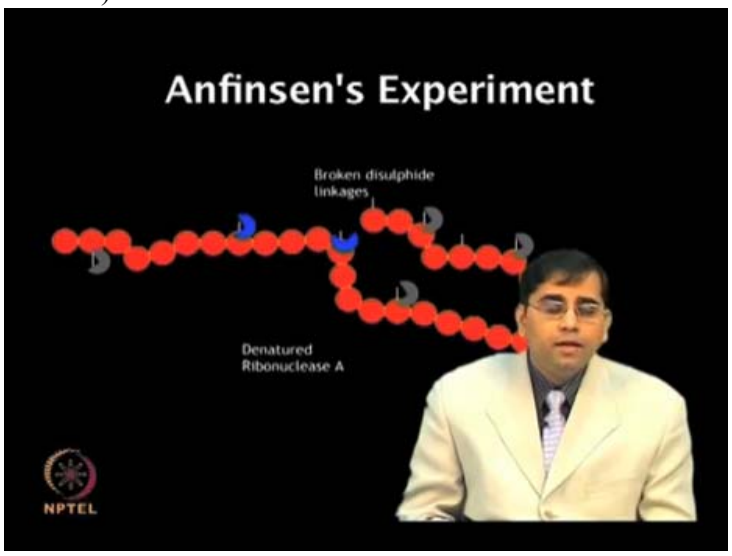
…unfolded protein ….

(Refer Slide Time 09:28)



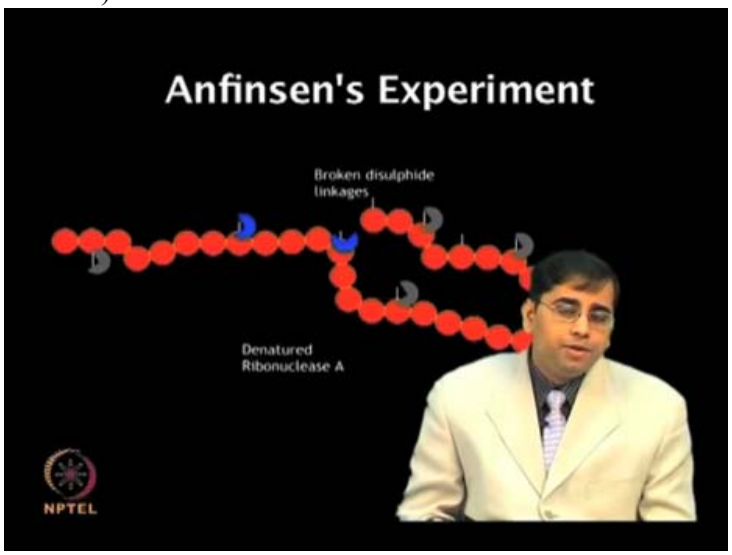
…to spontaneously fold into its native state by using …

(Refer Slide Time 09:34)



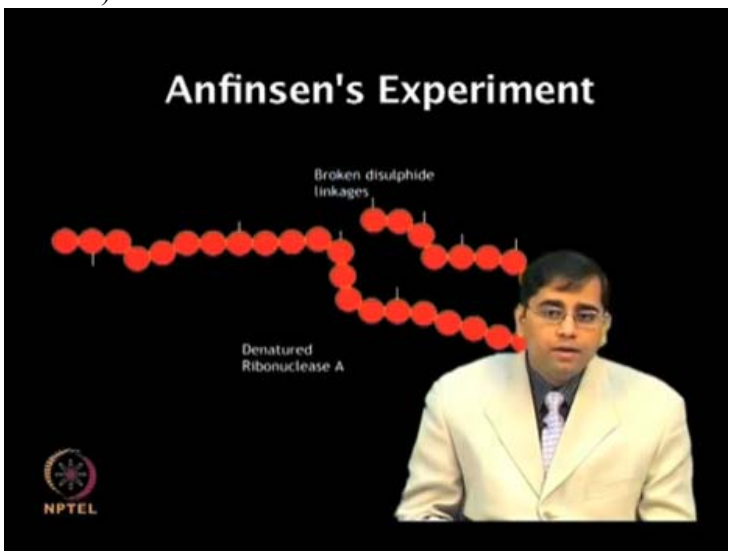
… protein ribonuclease A

(Refer Slide Time 09:38)



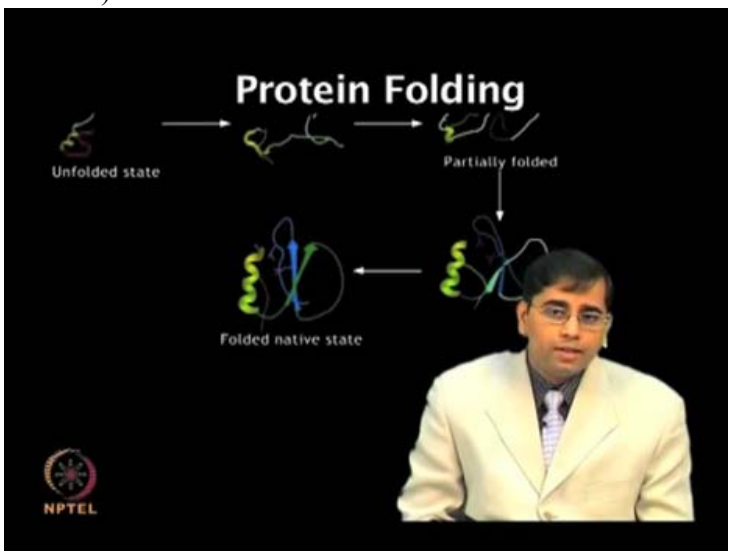
The study established that a primary amino acid sequence of a protein contain all of the information necessary for the proper folding

(Refer Slide Time 09:51)



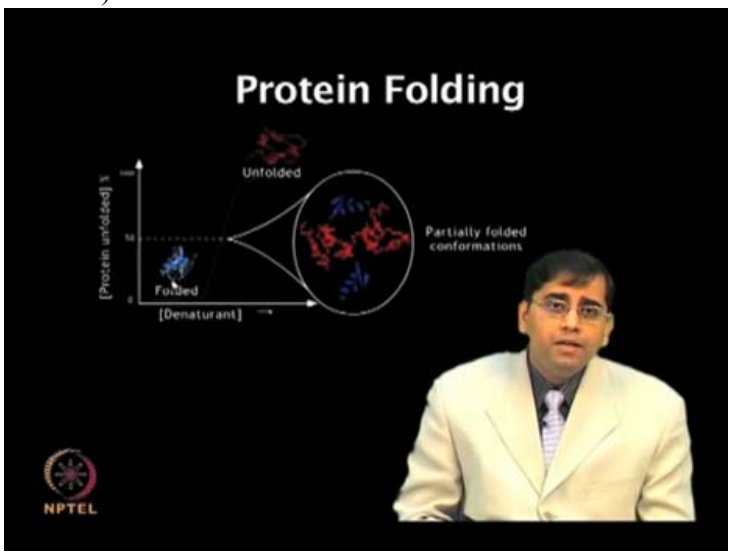
...into its native form Protein folding is a cooperative process…

(Refer Slide Time 09:57)



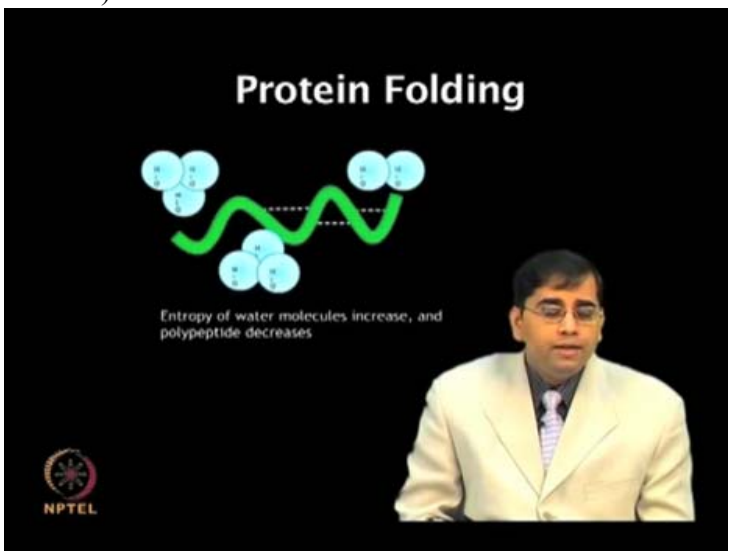
… which arises from simultaneous formation of ….

(Refer Slide Time 10:03)



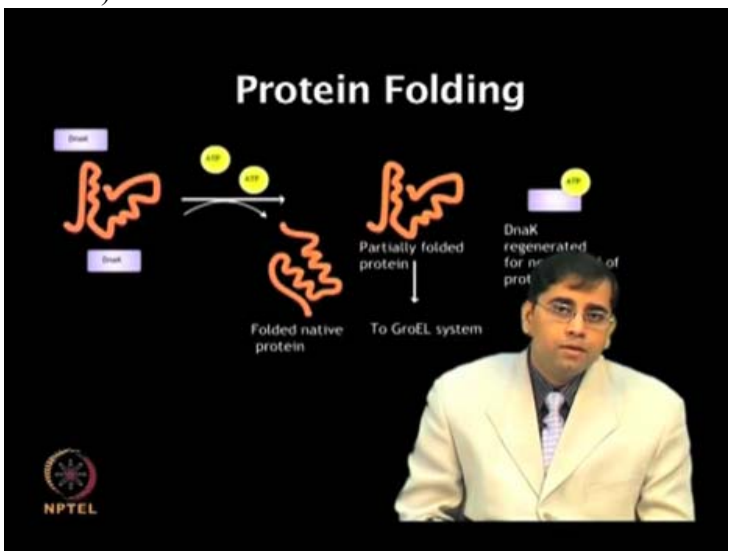
… multiple interactions within a polypeptide chain

(Refer Slide Time 10:10)



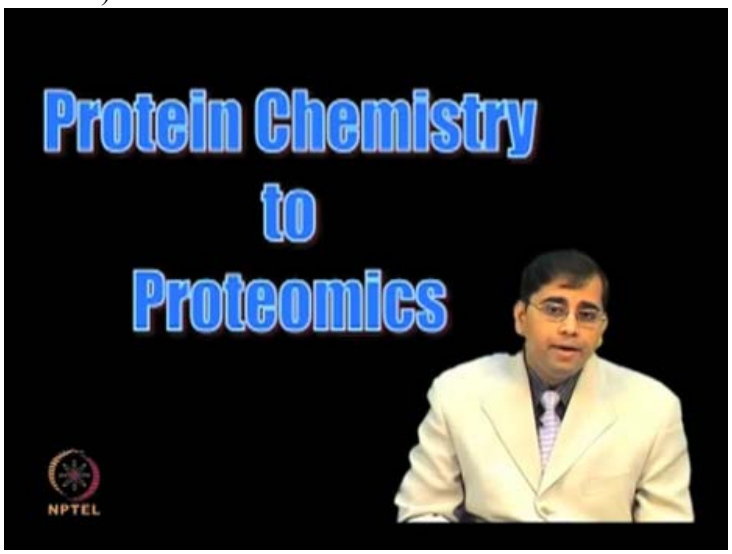
Protein folding is thermodynamically favorable and is spontaneous process.

(Refer Slide Time 10:17)



The folding efficiency could be limited by processes such as protein aggregation. The molecular chaperones are designed to assist in protein refolding.

(Refer Slide Time 10:43)

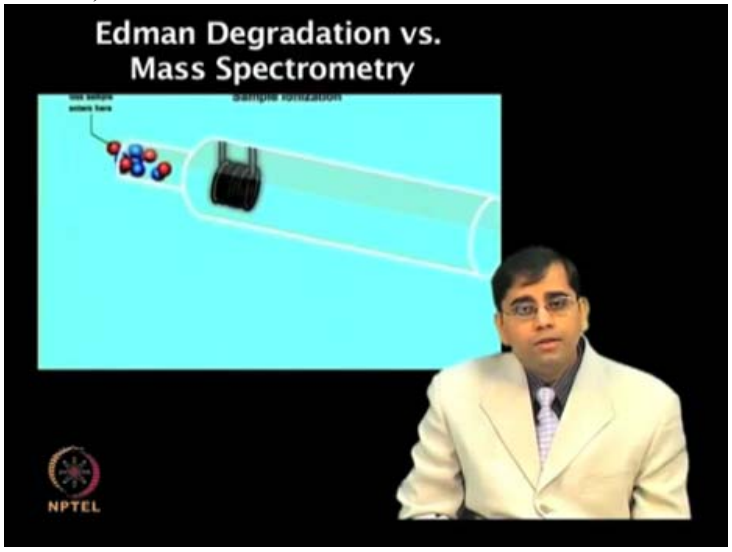


Proteomics research originates from classical protein chemistry and it has embraced new high throughput techniques to analyze complex samples. Many of the techniques used under the modern proteomic umbrella, for example two-dimensional electrophoresis, mass spectrometry have actually originated several years ago. So what is new?

The technological advancements in protein analysis with increased sensitivity, resolution and capability to carry out high throughput studies has led to the transition from protein chemistry

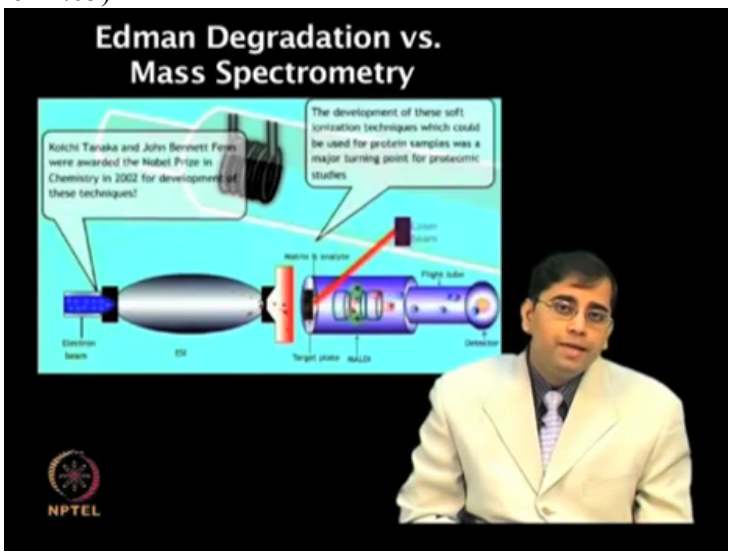
to new field of proteomics. Protein analysis by mass spectrometry was challenging due to complete degradation of samples with available hard ionization techniques.

(Refer Slide Time 11:49)



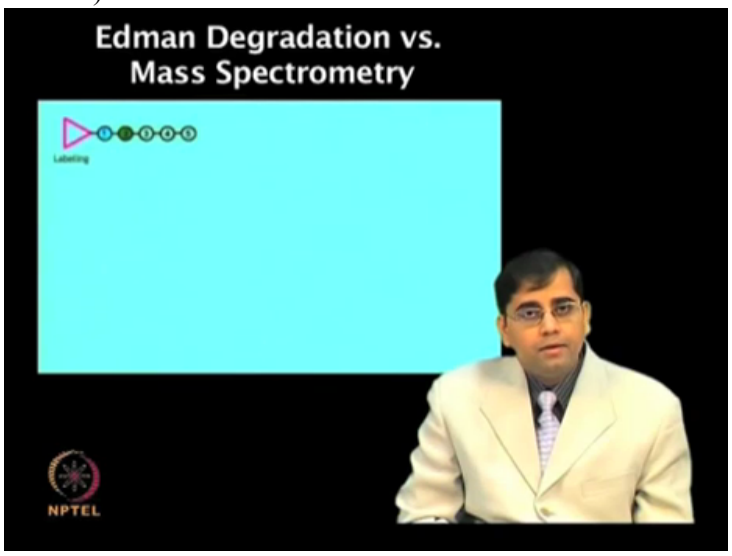
This limitation was overcome by soft ionization techniques such as MALDI and electro spray ionization. These techniques have greatly improved the proteomic studies …

(Refer Slide Time 12:05)



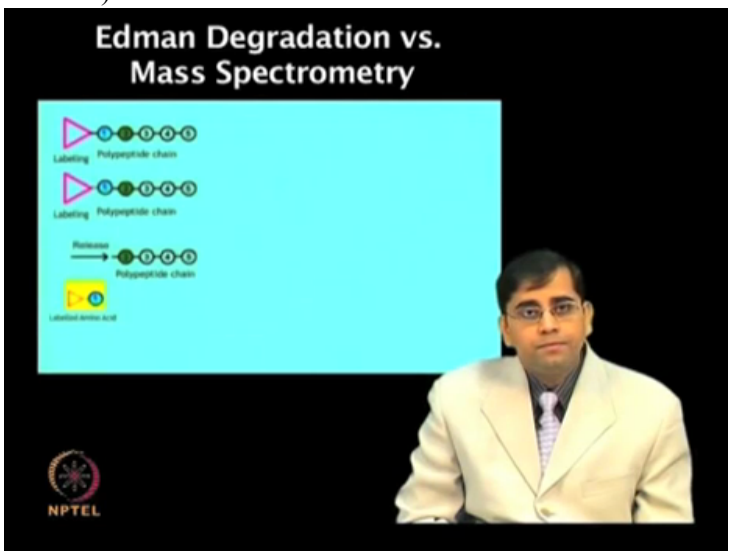
… as they facilitated mass spectrometry of protein samples

(Refer Slide Time 12:16)



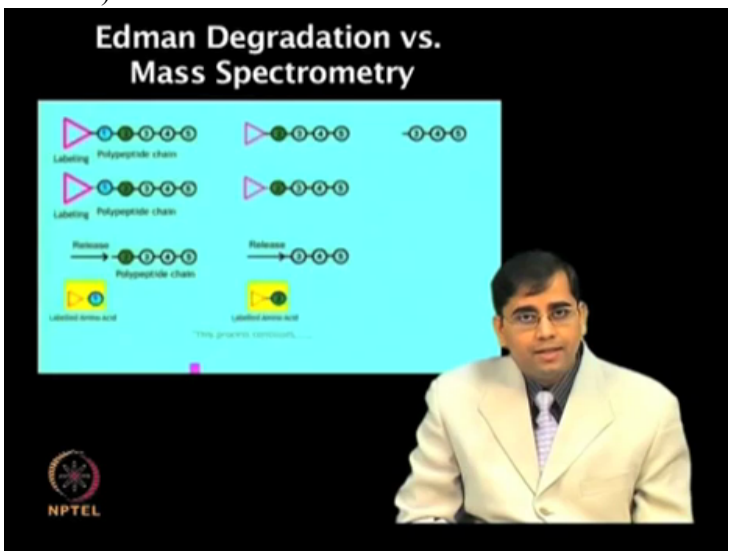
Protein sequencing by Edman degradation is time consuming and cumbersome.

(Refer Slide Time 12:23)



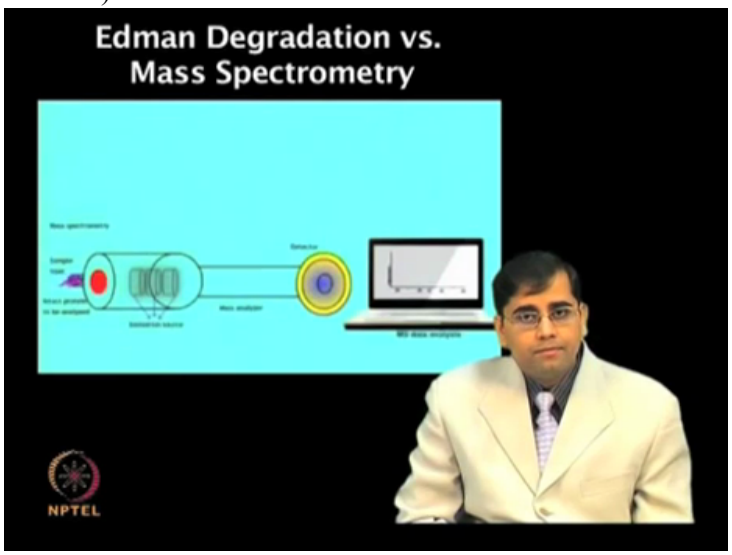
Several rounds of sequencing are required for analysis …

(Refer Slide Time 12:30)



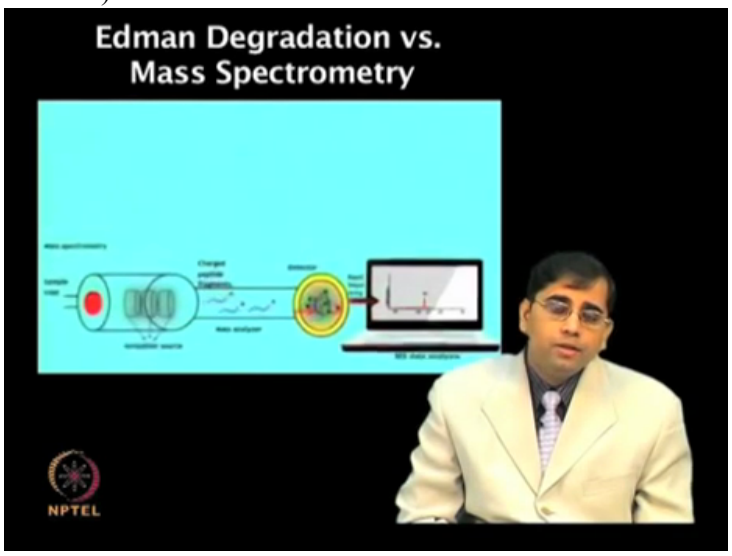
… of polypeptide chains

(Refer Slide Time 12:31)



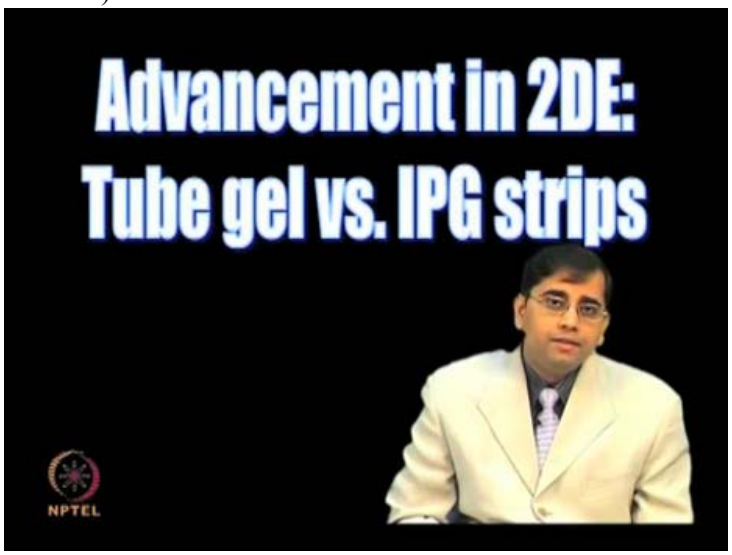
However, peptide sequencing by mass spectrometry is much faster and allows

(Refer Slide Time 12:40)



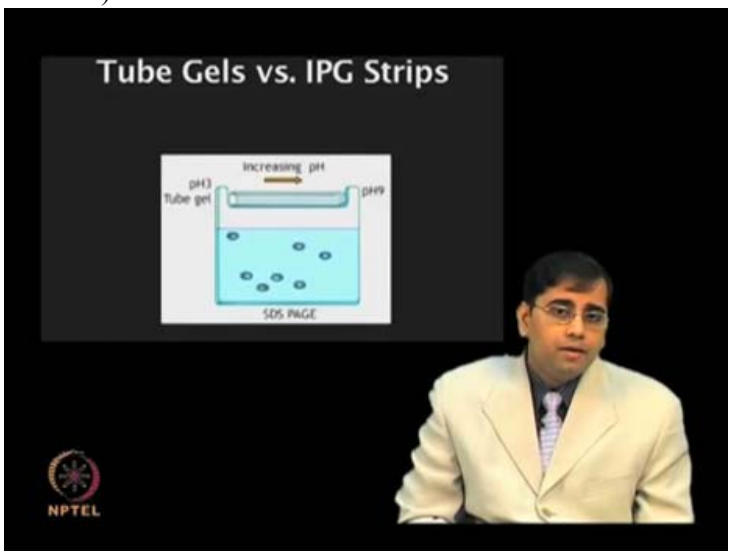
… large number of samples to be analyzed in a short time

(Refer Slide Time 12:56)



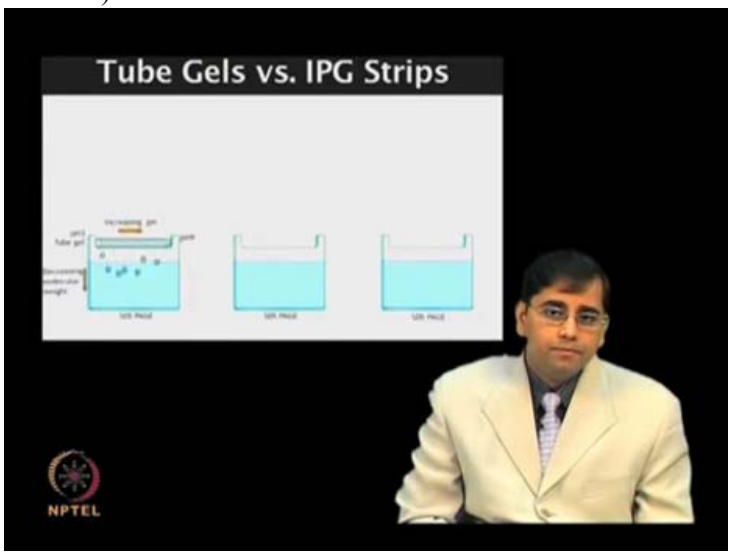
Another aspect, development of immobilized pH strips facilitated proteomic analysis using two-dimensional electrophoresis. The pH gradients in tube gels are established by ampholyte gradients which are not always very stable

(Refer Slide Time 13:14)



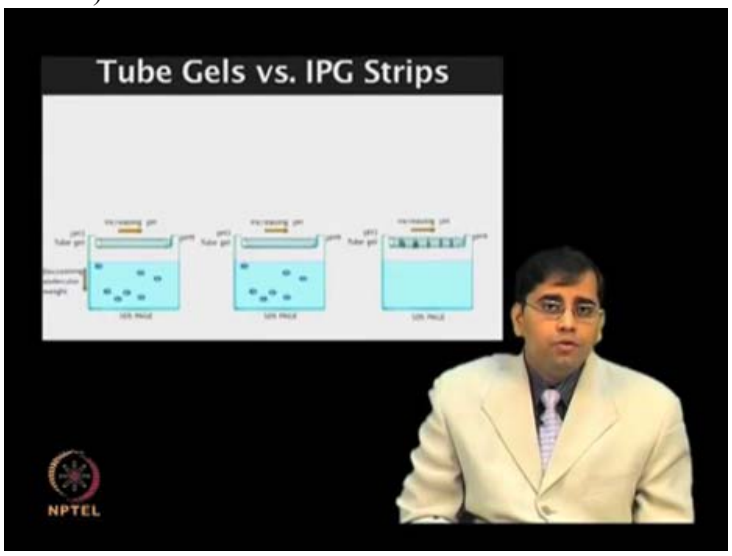
…. and tend to break down upon addition of …

(Refer Slide Time 13:21)



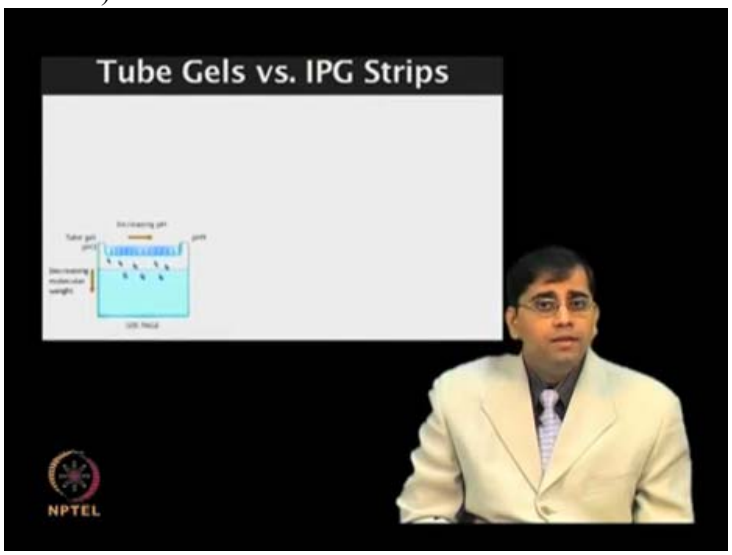
...the concentrated samples Analysis of protein mixture by two-dimensional electrophoresis using tube gels often resulted into …

(Refer Slide Time 13:32)



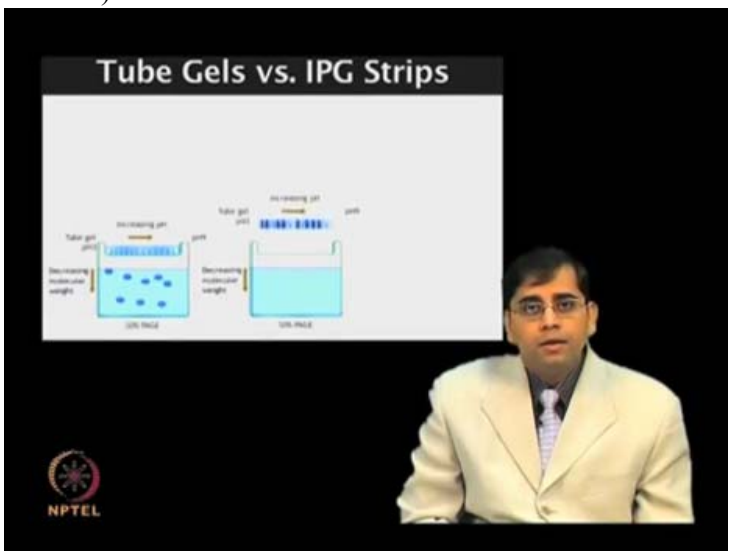
… variation in the gels The problem of reproducibility was overcome to a large extent

(Refer Slide Time 13:44)



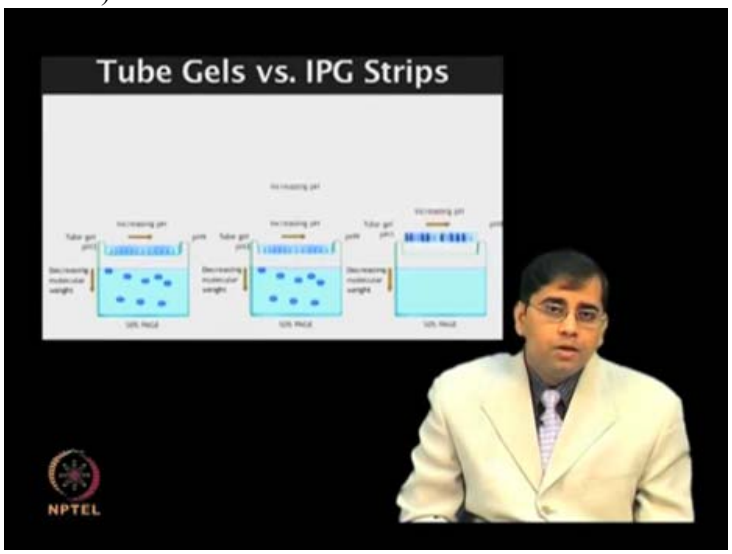
… by the development of immobilized pH gradient strips or IPG strips

(Refer Slide Time 13:52)



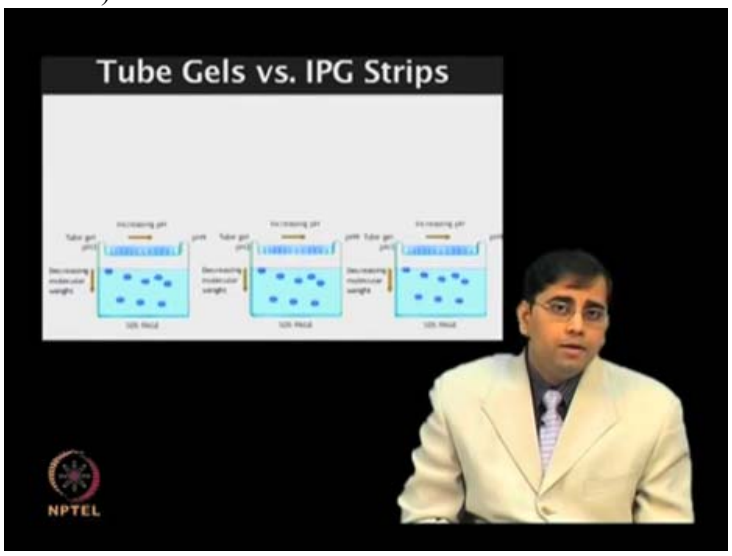
Minimal gel to gel variations was observed where samples were run by two-dimensional electrophoresis employing IPG strips

(Refer Slide Time 14:04)



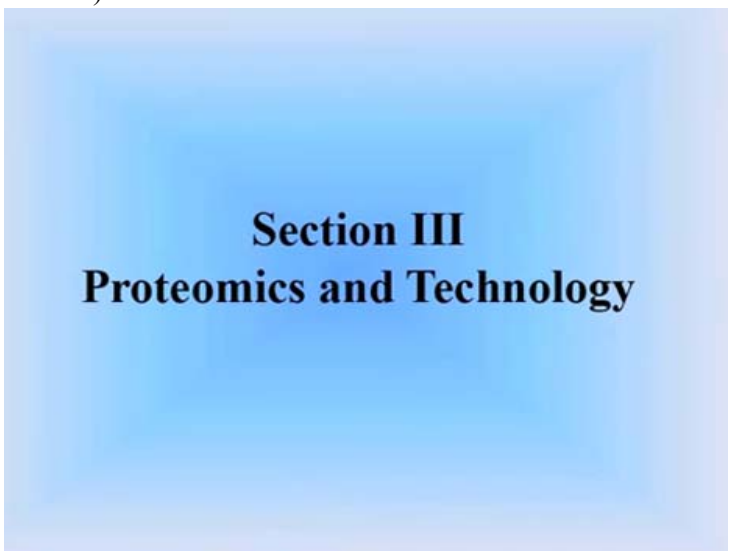
… which made this technique suitable for the large scale …

(Refer Slide Time 14:09)

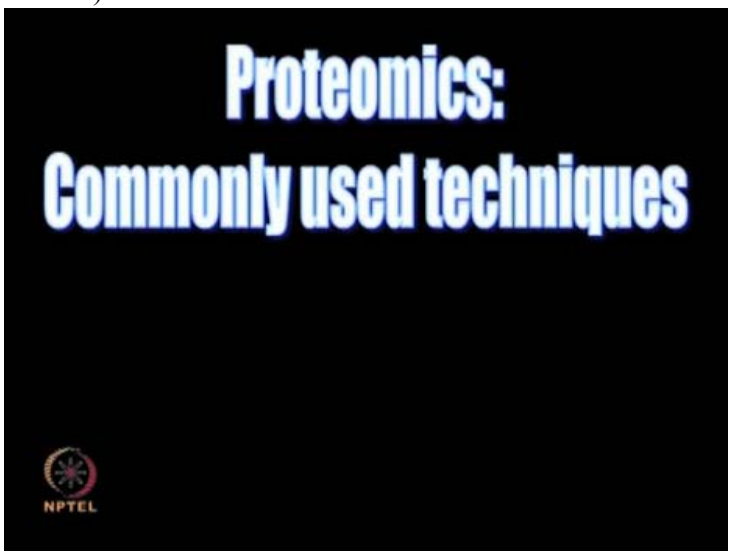


… proteomic applications

(Refer Slide Time 14:14)

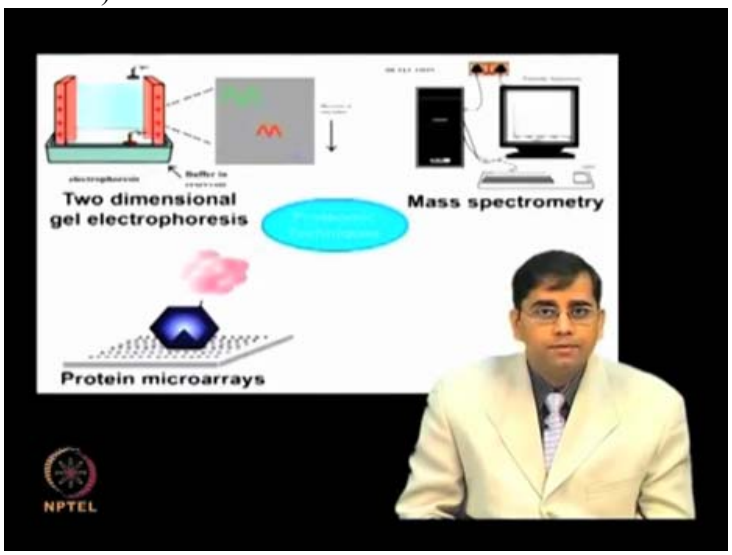


(Refer Slide Time 14:20)



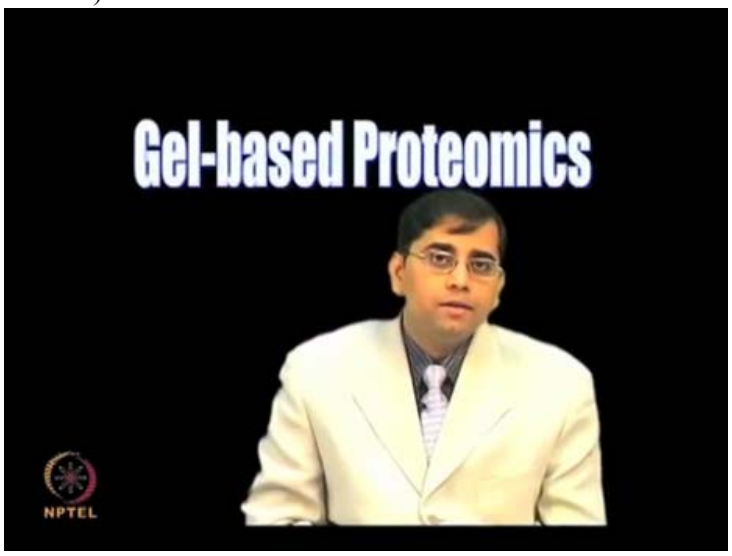
Different types of proteomic technologies …

(Refer Slide Time 14:33)



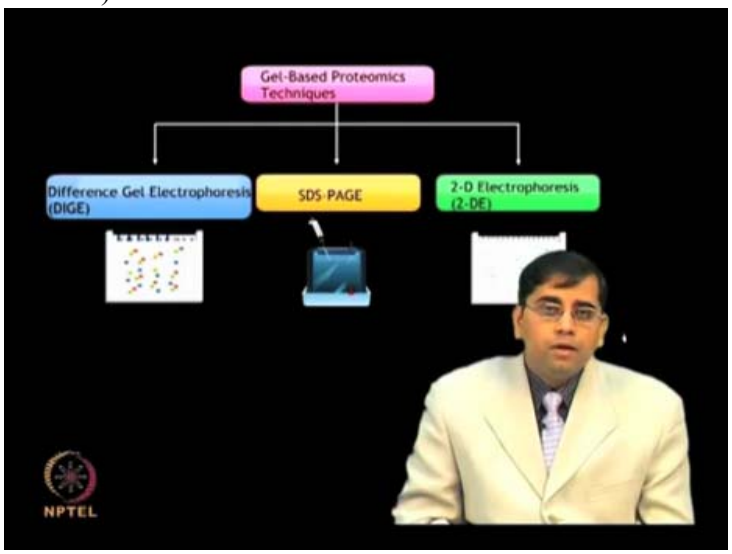
…such as two-dimensional electrophoresis, mass spectrometry, Microarrays and label-free techniques will be discussed in more detail later

(Refer Slide Time 14:36)



In gel-based proteomics, proteins are commonly analyzed using SDS PAGE and two-dimensional electrophoresis.

(Refer Slide Time 14:58)

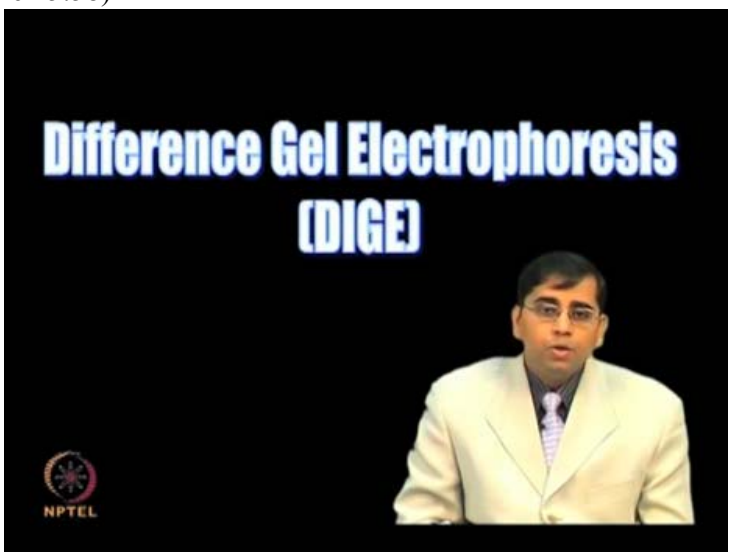


Separation in SDS PAGE occurs almost excessively on the basis of molecular weight where as in 2DE, the first dimensional separation is based on isoelectric point and second dimension separation based on the molecular weight.

Some of the limitations of two-dimensional electrophoresis can be overcome by Difference gel electrophoresis or DIGE technique. 2DE or DIGE in combination with mass spectrometry has been the standard technique for proteome analysis.

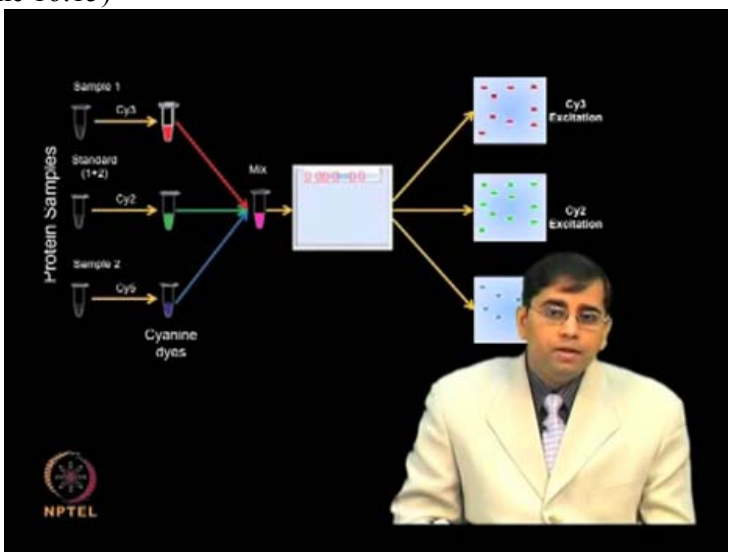
The two-dimensional electrophoresis involves the protein separation on pH gradient based on their isoelectric points using Isoelectric Focusing followed by separation in second dimension using SDS PAGE.

(Refer Slide Time 15:58)



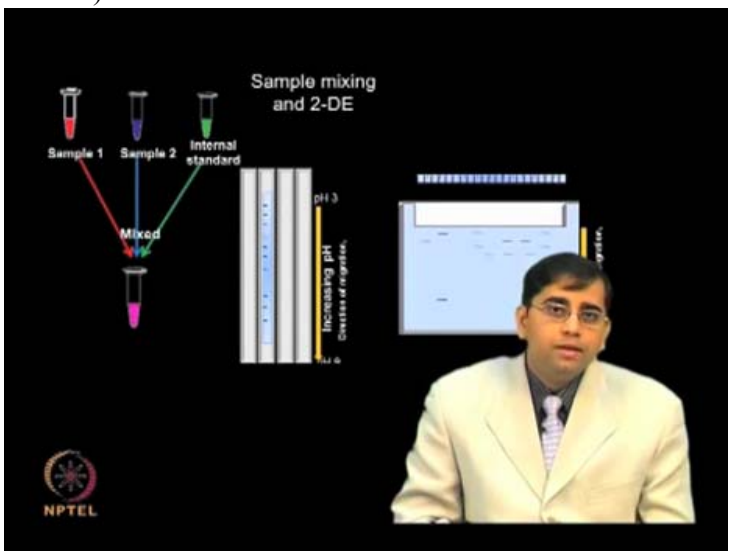
Two-dimensional electrophoresis has high resolving power but it has several limitations such as staining artifacts and reproducibility in gel to gel.

(Refer Slide Time 16:15)



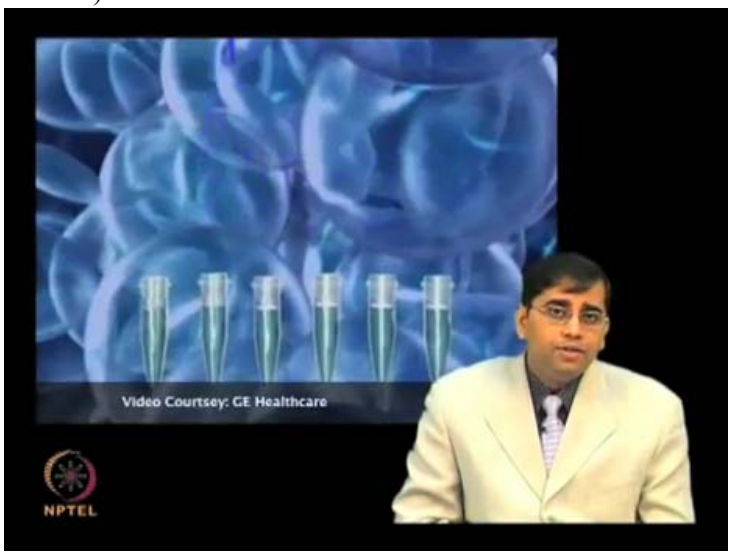
Fluorescence two-dimensional difference in gel electrophoresis or 2D DIGE is advanced 2D technique that allows for accurate quantitation …

(Refer Slide Time 16:28)



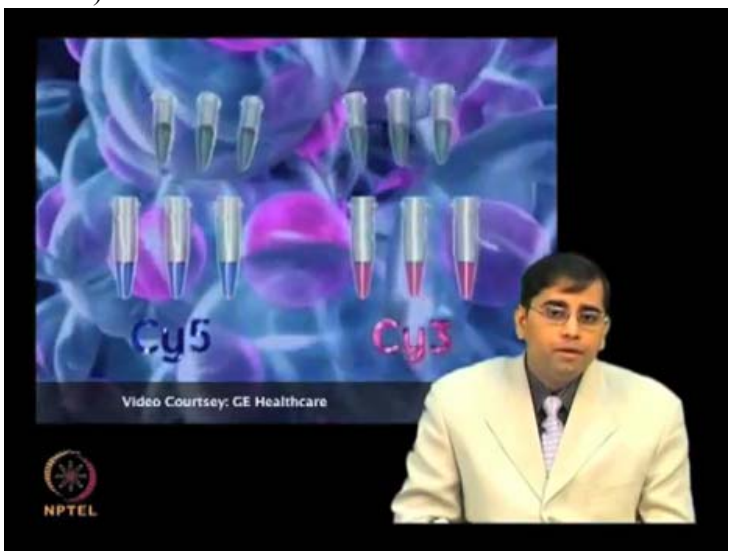
…with statistical confidence while controlling non-biological variations

(Refer Slide Time 16:36)



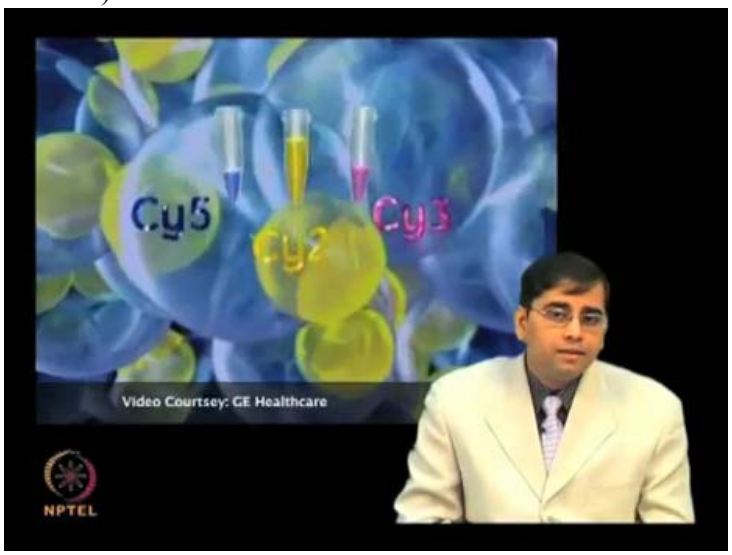
In DIGE, proteins extracted from different types of cells or tissue samples are

(Refer Slide Time 16:44)



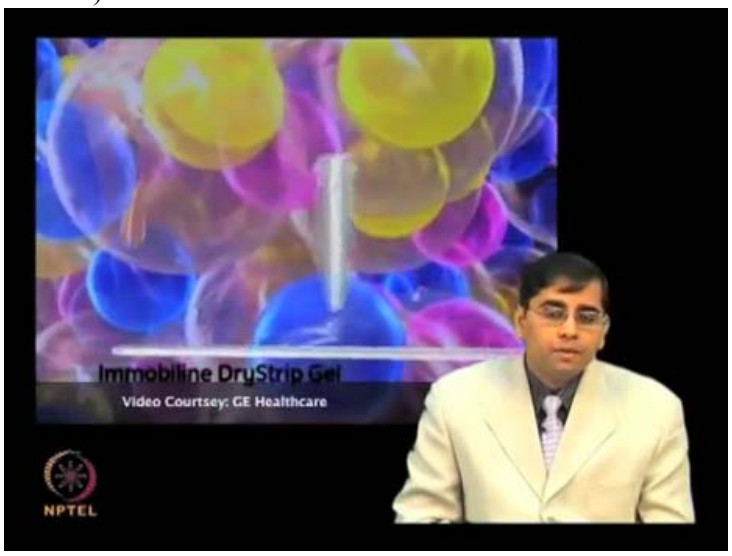
…labeled with different fluorescent reagents …

(Refer Slide Time 16:48)



… such as CY2, CY3 and CY5 …

(Refer Slide Time 16:53)

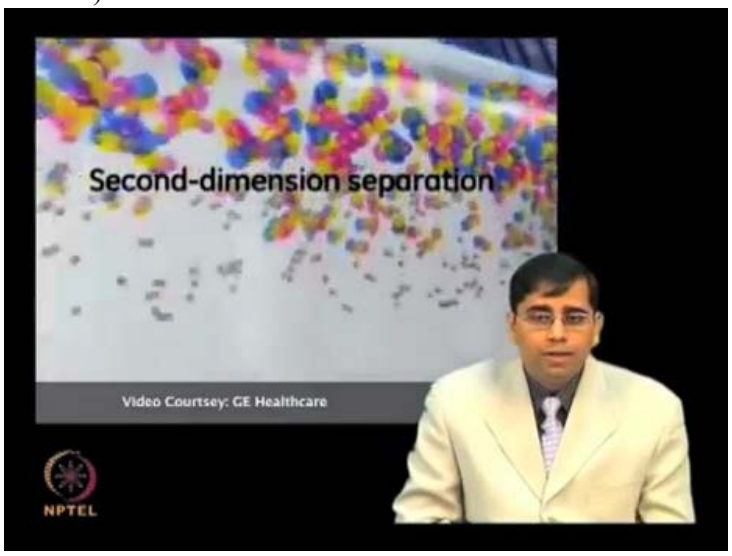


… mixed and …

(Refer Slide Time 16:55)

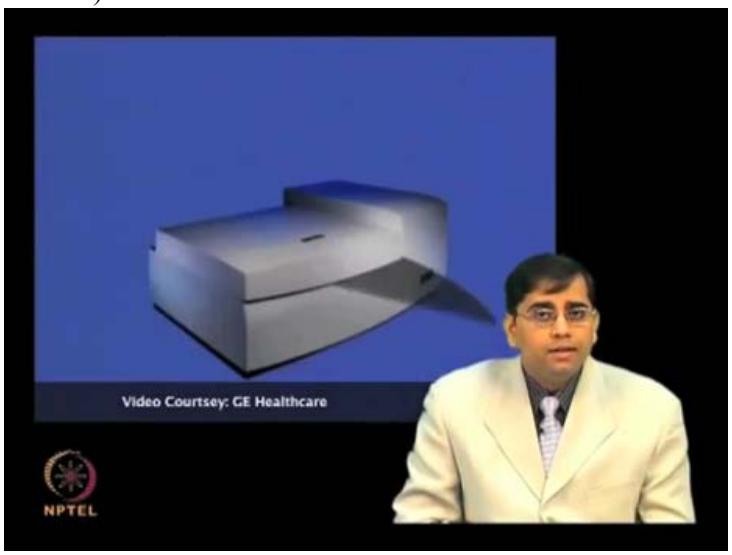
… then separated by two-dimensional electrophoresis …

(Refer Slide Time 16:58)



.. on a single gel

(Refer Slide Time 17:02)

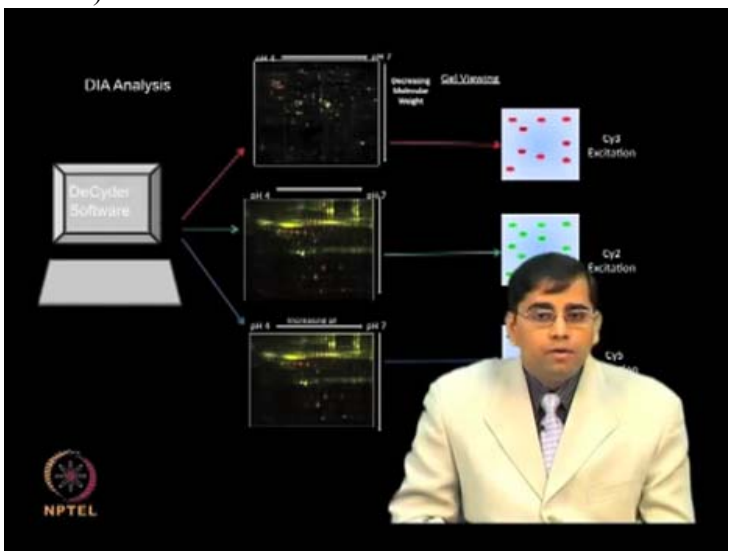


The proteins are detected separately …

(Refer Slide Time 17:06)

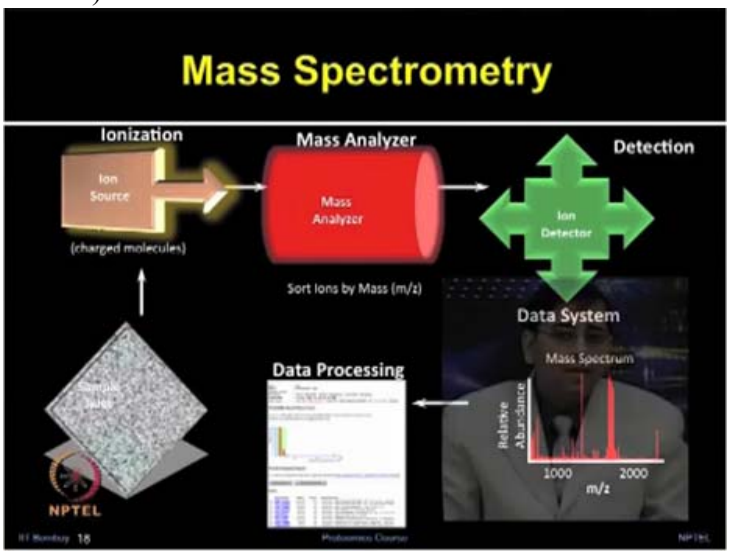
… using the excitation wavelength specific to the …

(Refer Slide Time 17:11)



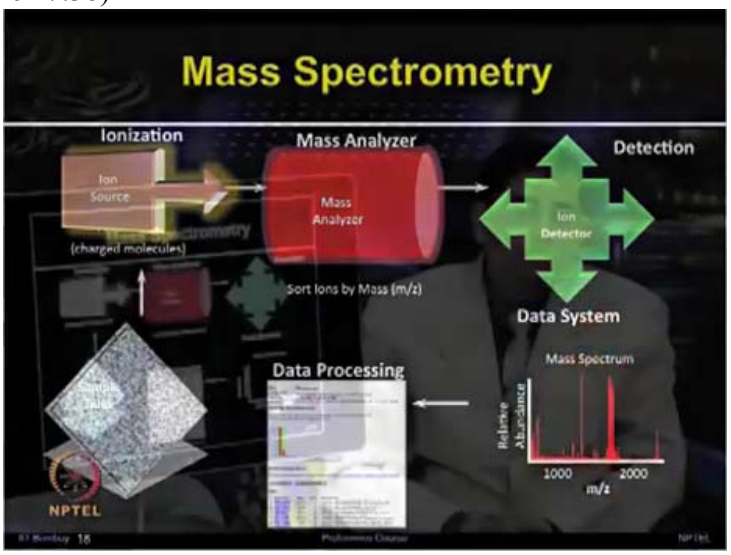
… different sourcing reagents, Cy2, Cy3 and Cy5

(Refer Slide Time 17:18)



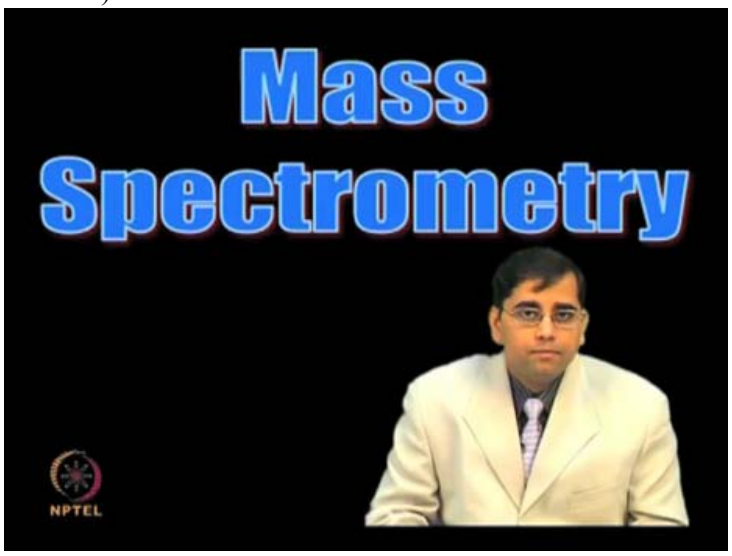
Mass spectrometry is technique for protein identification and analysis by production of charged molecular species in vacuum and its separation in magnetic and electric fields based on mass to charge ratio.

(Refer Slide Time 17:38)



Mass spectrometry has become the method of choice for analysis of complex protein samples in proteomics study due to its ability to identify thousands of proteins.

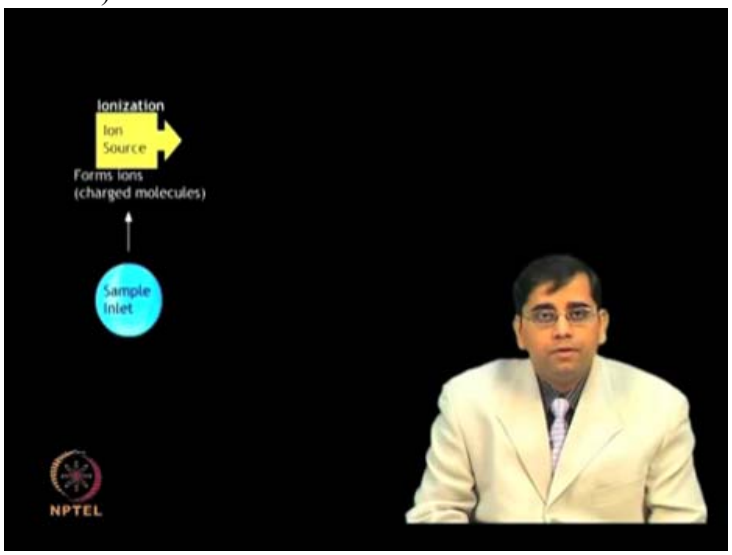
(Refer Slide Time 18:20)



The gel-based techniques typically resolve only the products of few hundred genes at best, had low throughput and low dynamic range. To overcome such issues, mass spectrometry has become an important analytical tool in proteomics and in biology in general. It offers high throughput, sensitive and specific analysis for many applications.

In this module, we will discuss basic concepts of mass spectrometry, ionization sources, mass analyzers, hybrid MS configurations and quantitative proteomic techniques such as SILAC and iTRAQ.

(Refer Slide Time 18:58)

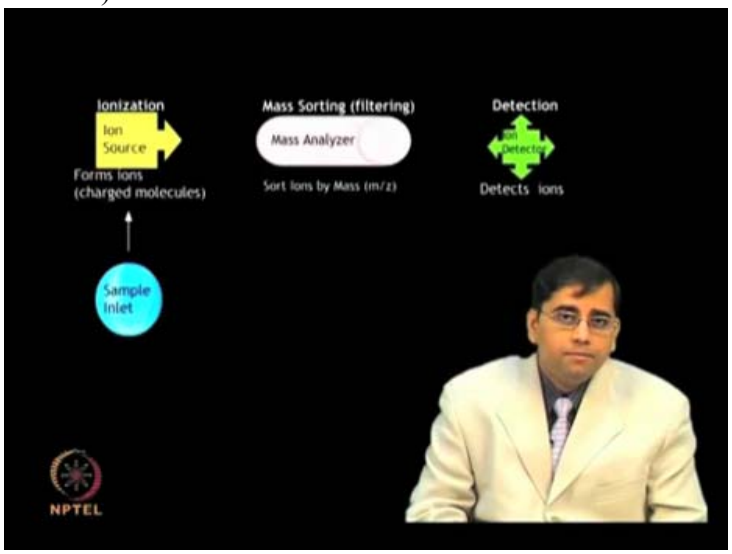


The basic components of mass spectrometry involve sample inlet to transfer the sample into

(Refer Slide Time 19:06)

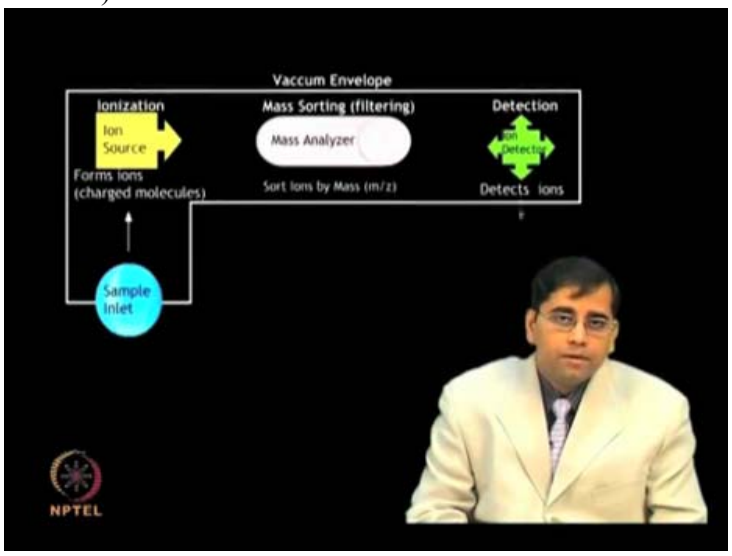
….the ion source, ionization source which converts neutral sample molecules into the gas phase ions

(Refer Slide Time 19:16)



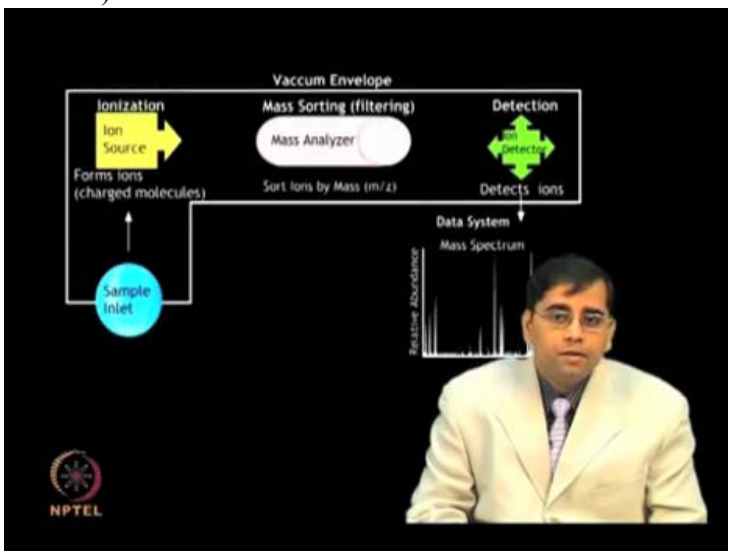
A mass analyzer separates and analyzes mass of ionic species.

(Refer Slide Time 19:25)



Detector which measures and amplifies ion current of mass resolved ion …

(Refer Slide Time 19:32)



… and data system to process and analyze data

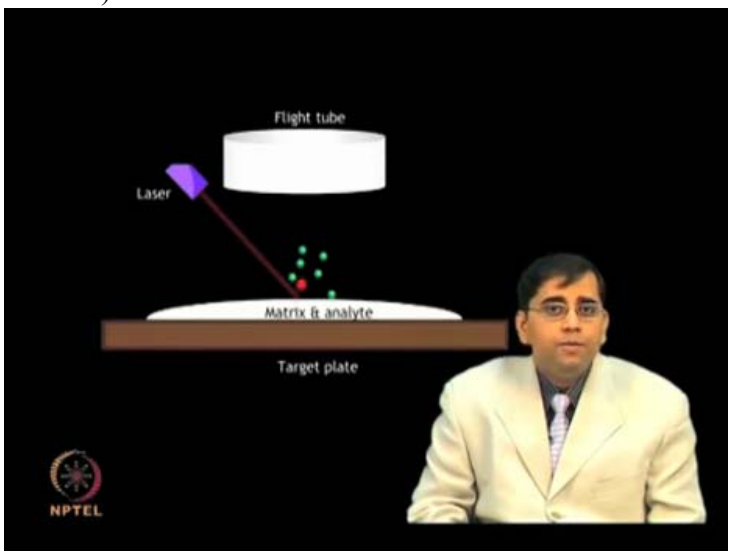
(Refer Slide Time 19:45)

Soft ionization techniques such as Matrix Assisted Laser Desorption/Ionization MALDI and Electro Spray Ionization ESI are now widely used for proteomic applications.

(Refer Slide Time 19:59)

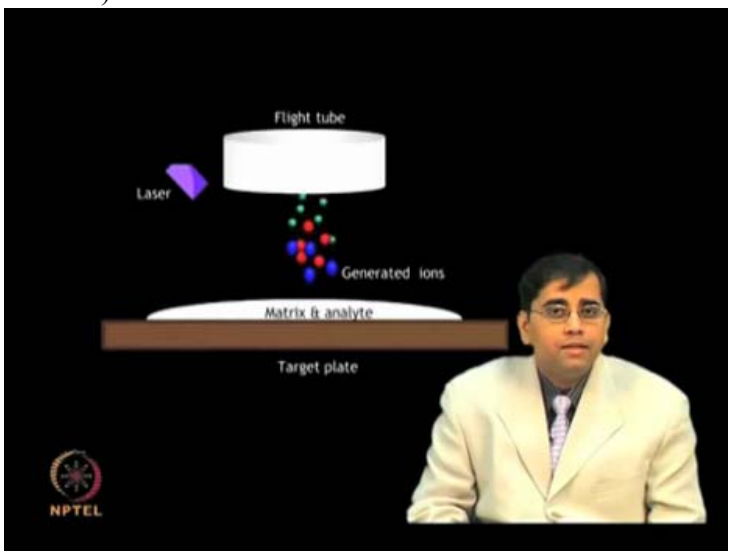
In MALDI, protein is mixed with matrix and laser beam ionizes matrix molecules.

(Refer Slide Time 20:10)



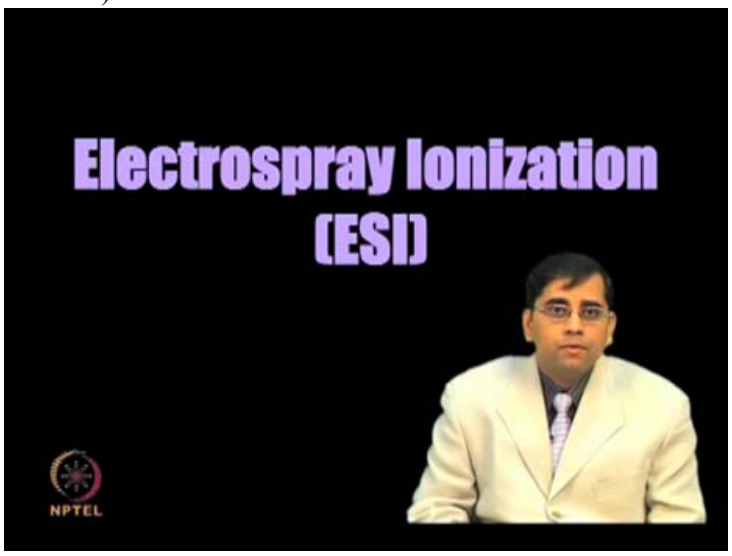
It is an efficient process for generating gas phase ions of peptides and proteins …

(Refer Slide Time 20:18)



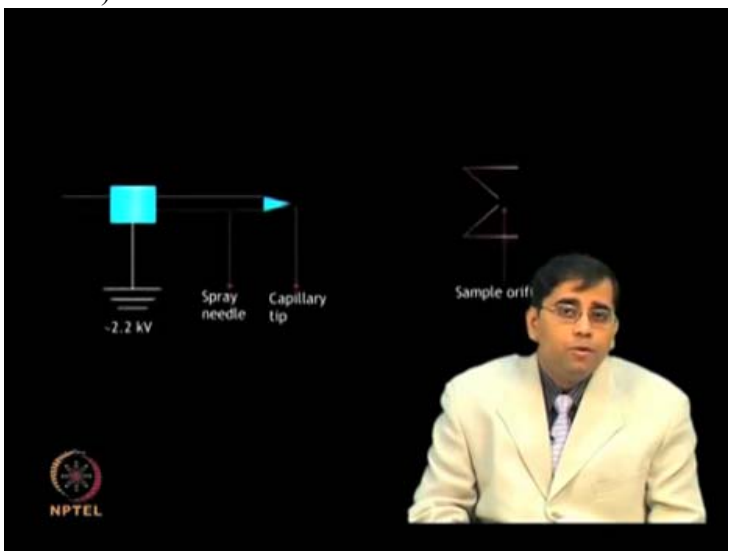
….for mass spectrometry detection

(Refer Slide Time 20:24)



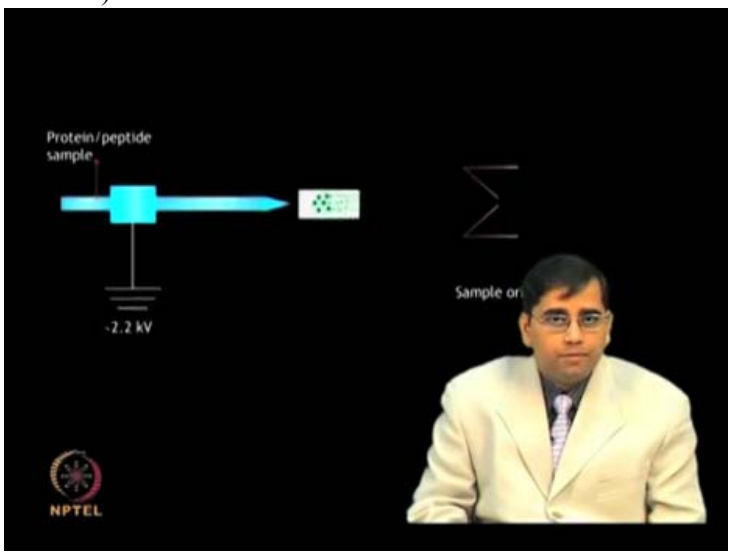
Electro-Spray Ionization requires sample of interest to be in solution, produces gas phase ions

(Refer Slide Time 20:36)



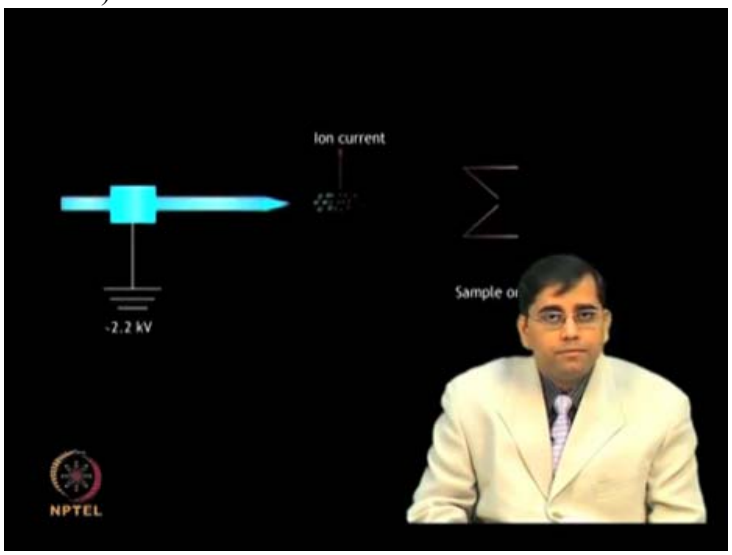
…from solution obtained from the …

(Refer Slide Time 20:42)



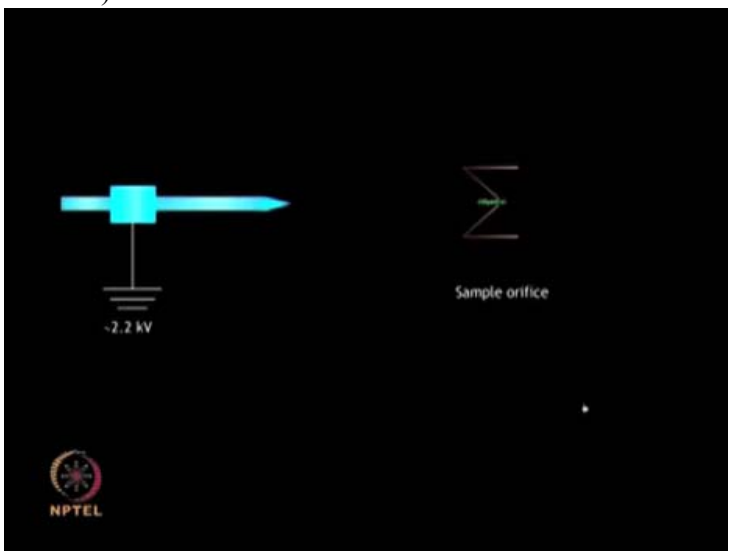
… chromatographic steps

(Refer Slide Time 20:43)



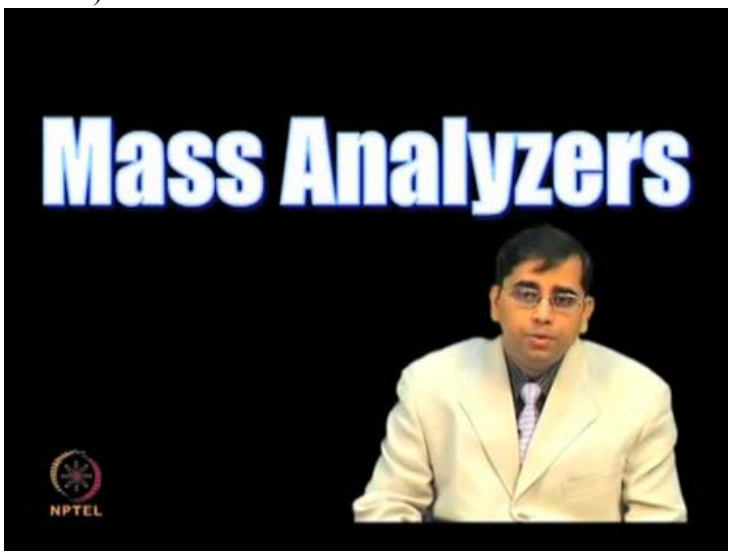
The distinguishing feature of Electro-Spray Ionization is its ability to …

(Refer Slide Time 20:53)



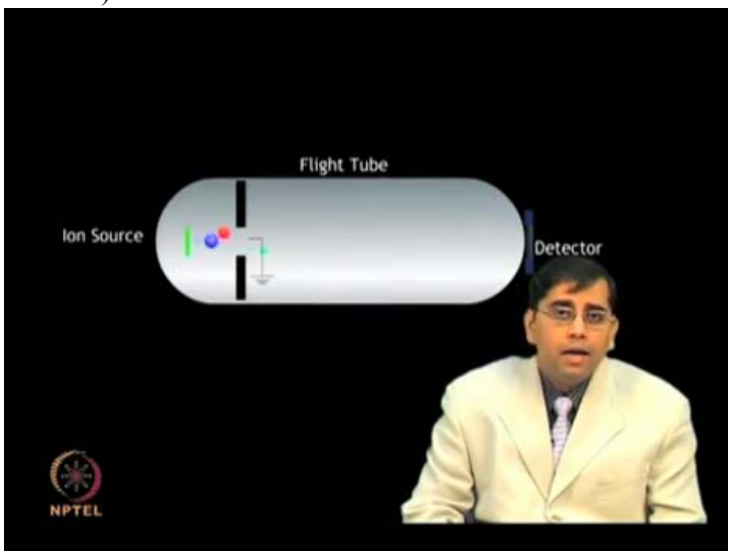
…produce multiple charged species.

(Refer Slide Time 21:02)



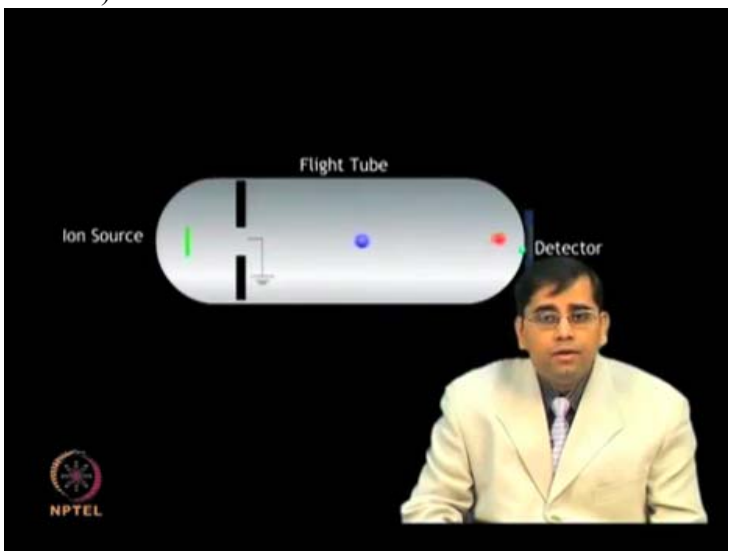
The mass analyzer disperses all the ions based on their mass to charge ratio and focuses all the mass resolved ions at a single focal point and maximizes their transmission.

(Refer Slide Time 21:17)



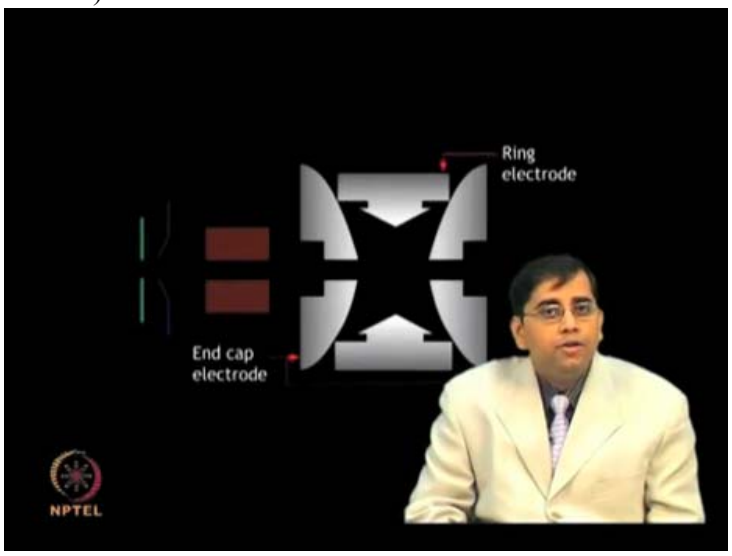
The Time of Flight measures the m by z ratio of ions…

(Refer Slide Time 21:21)



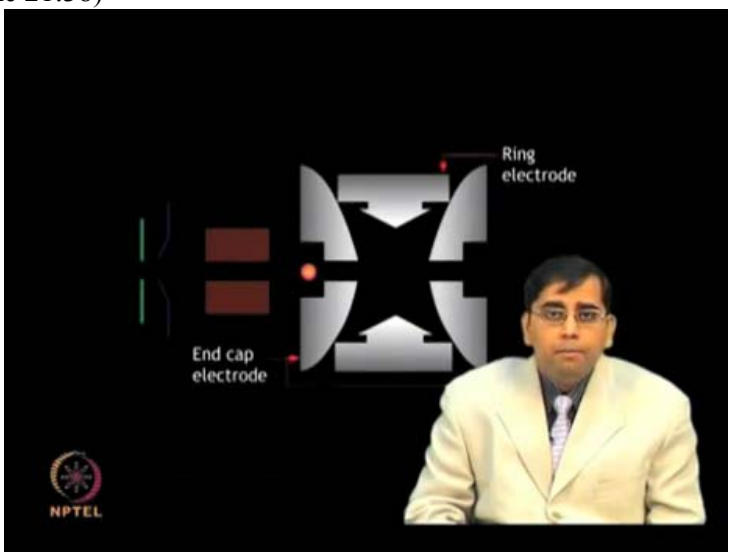
… based on the time it takes for the ions to fly in the analyzer and strike to the detector. .

(Refer Slide Time 21:32)



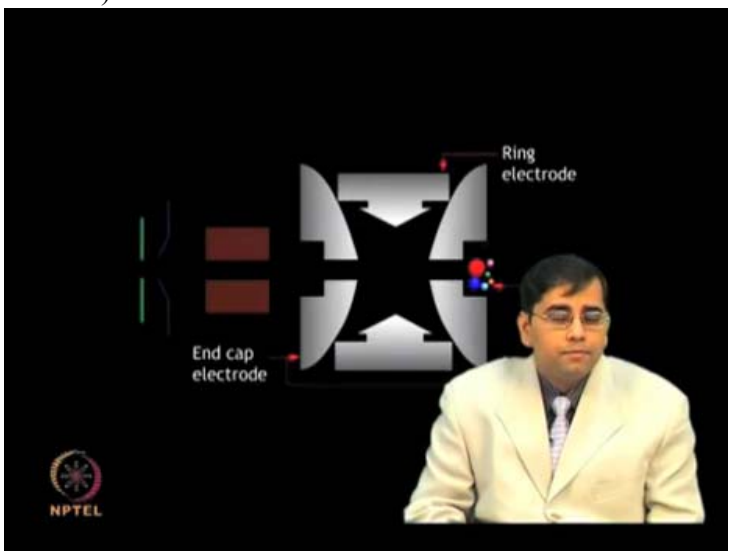
The Ion Trap, it traps ions using electrical field and ….

(Refer Slide Time 21:38)



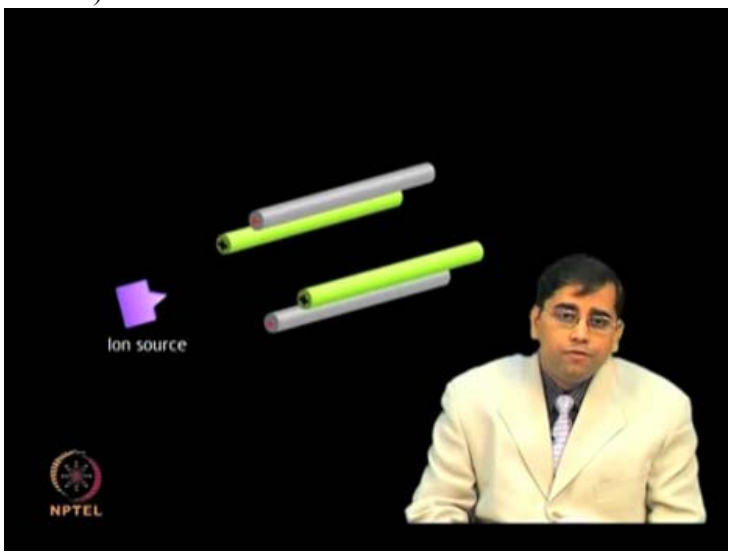
…. measures mass by selectively ejecting them …

(Refer Slide Time 21:45)



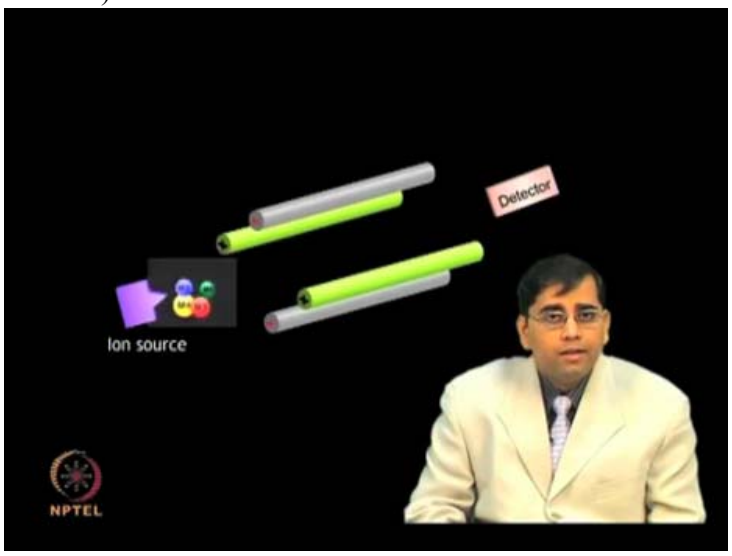
… to the detector

(Refer Slide Time 21:48)



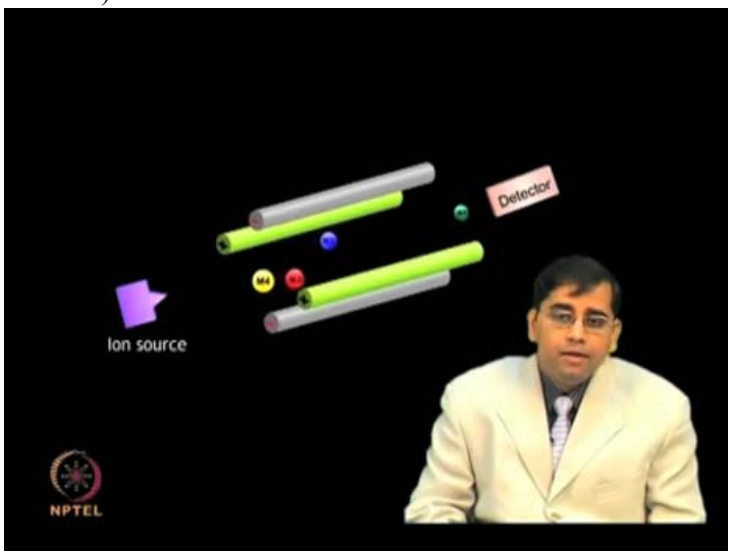
Quadrapole, it consists of 4 parallel metal rods …

(Refer Slide Time 21:51)



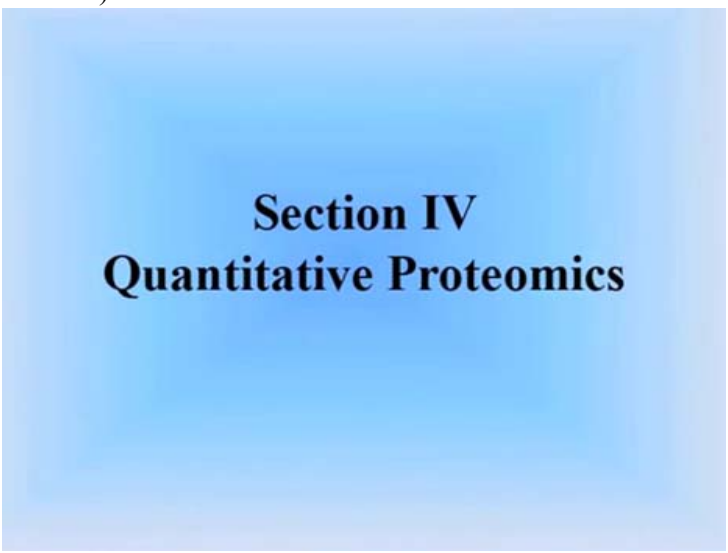
… and mass separation is accomplished by …

(Refer Slide Time 21:56)

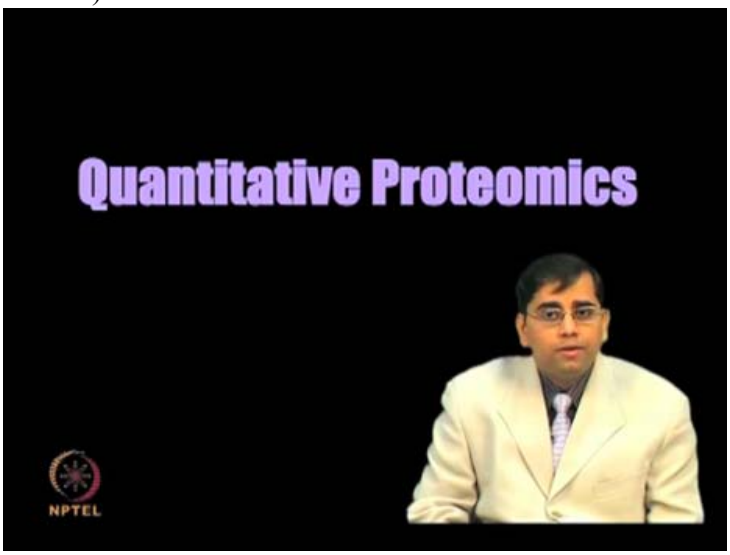


… stable, vibratory motion of ions in high frequency oscillating electric field that is created by applying direct current and radio frequency potentials to these electrodes.

(Refer Slide Time 22:17)



(Refer Slide Time 22:21)

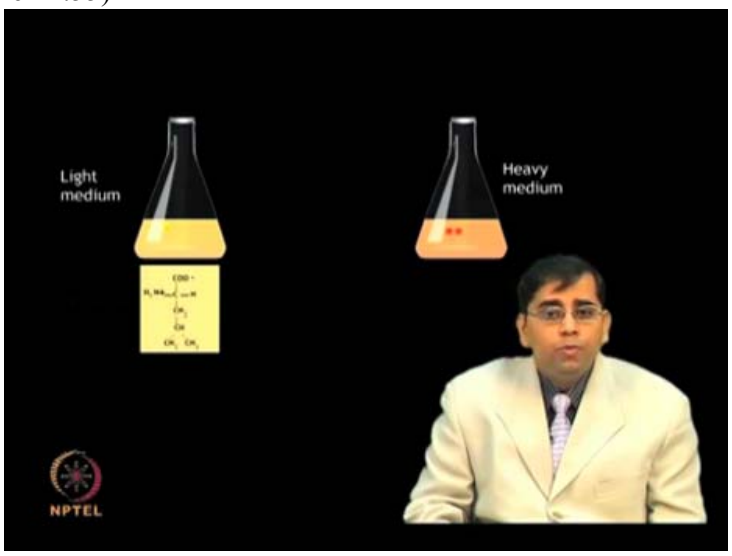


Protein labeling with stable isotopes are effective methods for quantitative proteome profile using mass spectrometry.

(Refer Slide Time 22:33)

Stable Isotope Labeling by Amino acids in Cell culture or SILAC which is a metabolic labeling strategy to encode whole cellular proteome is widely used method for quantitative proteomics.

(Refer Slide Time 22:55)



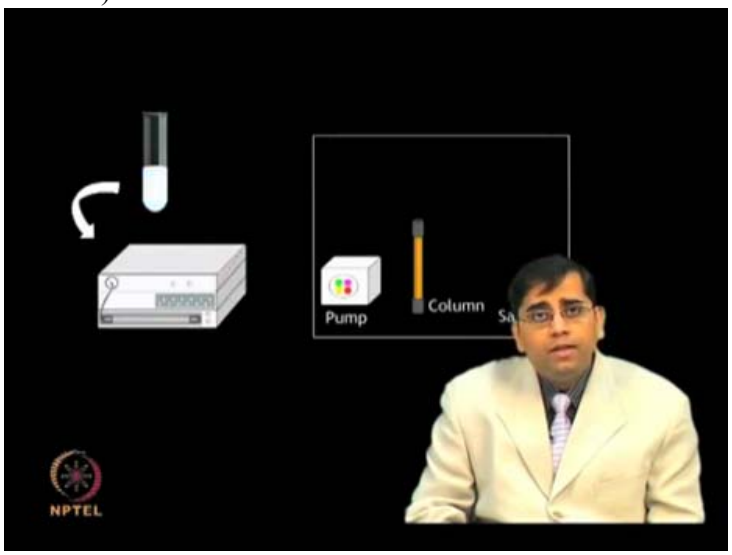
In SILAC, two groups of cells are cultured in media that are identical in all the respects except that one contains a heavy isotopic analog of an essential amino acid while the other contains the normal light amino acid.

After a number of cell divisions, the grown cells are combined and digested using Trypsin.

(Refer Slide Time 23:29)

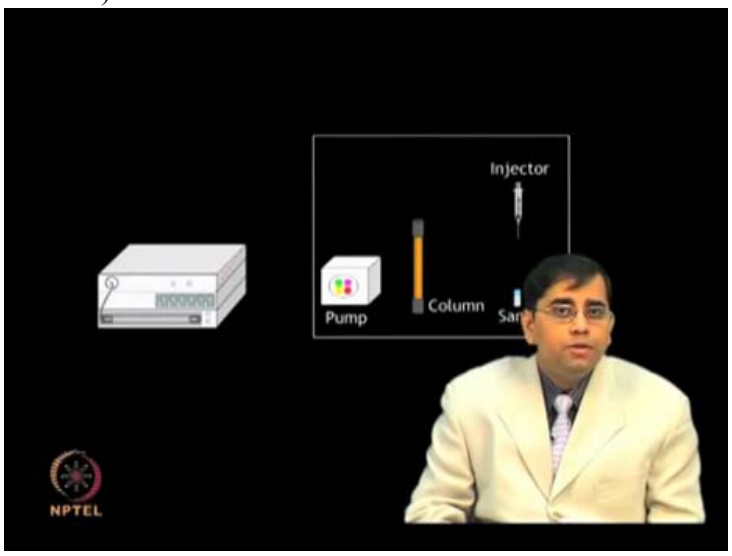
The complex protein mixture is further separated by SDS PAGE to simplify the analysis.

(Refer Slide Time 23:34)



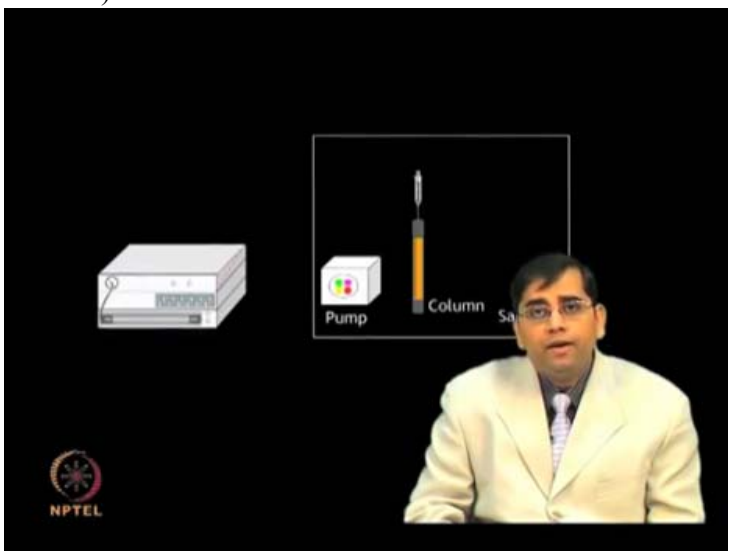
Further application is carried out ….

(Refer Slide Time 23:37)



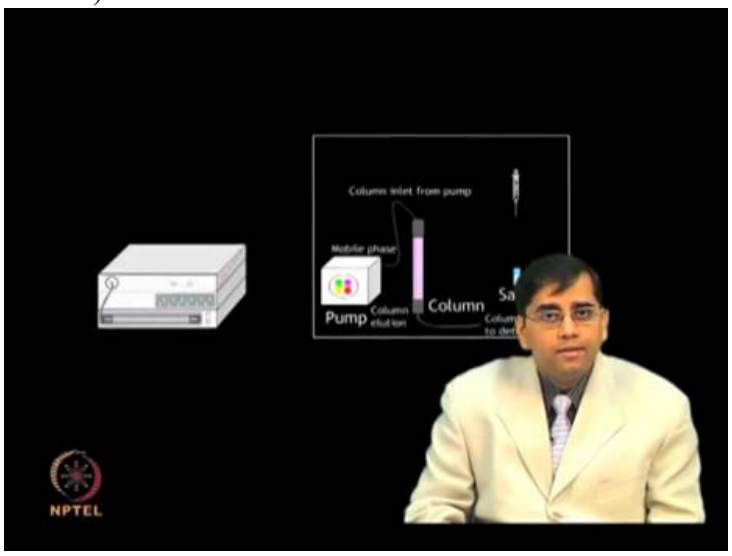
... by liquid chromatography and purified peptide fragments …

(Refer Slide Time 23:42)



… are analyzed…

(Refer Slide Time 23:44)

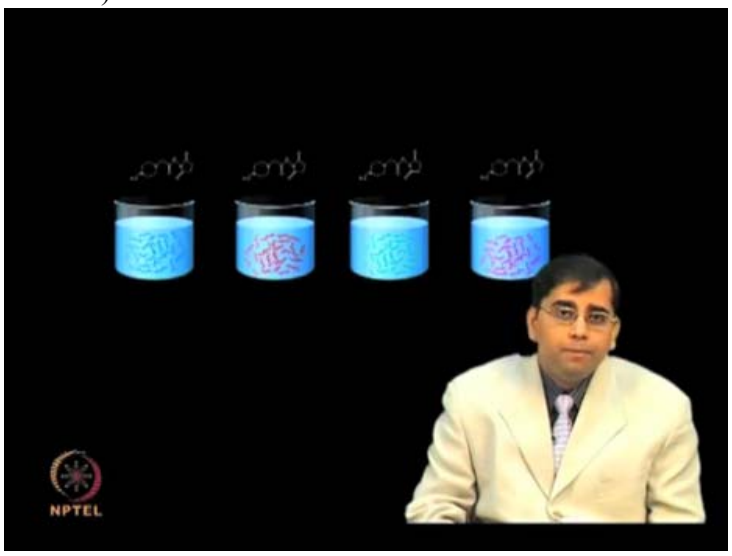


… by MS/MS

(Refer Slide Time 23:48)

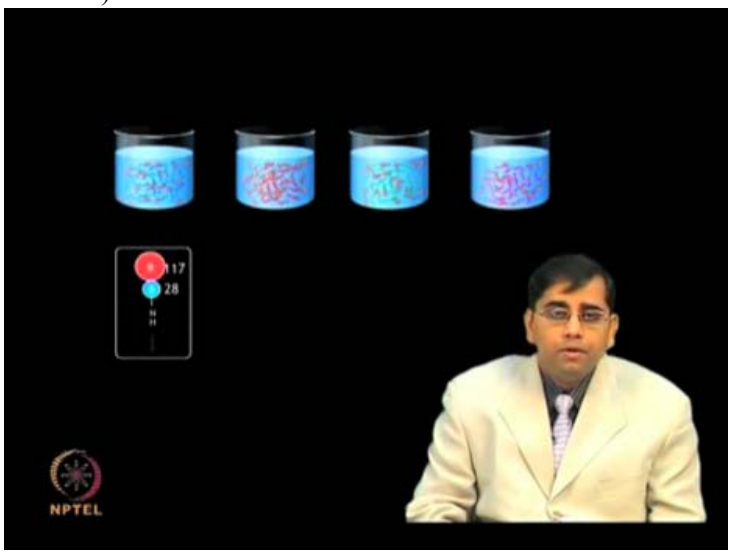
iTRAQ, it is a MS based technique for relative and absolute quantitation of proteins. iTRAQ reagents are a set of 4 isobaric amino-specific …

(Refer Slide Time 24:03)



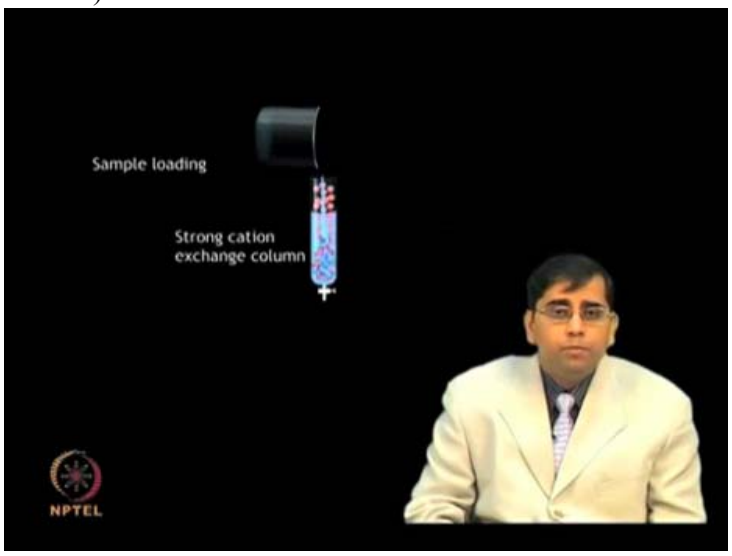
labeling reagents; 114, 115, 116 and 117. An iTRAQ reagent consists of a reporter group, a balancer group. …

(Refer Slide Time 24:18)



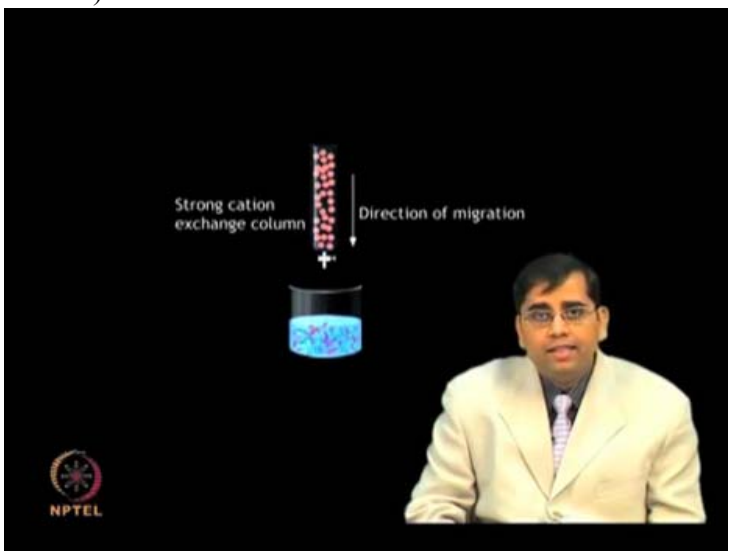
… and a peptide reactor group

(Refer Slide Time 24:21)



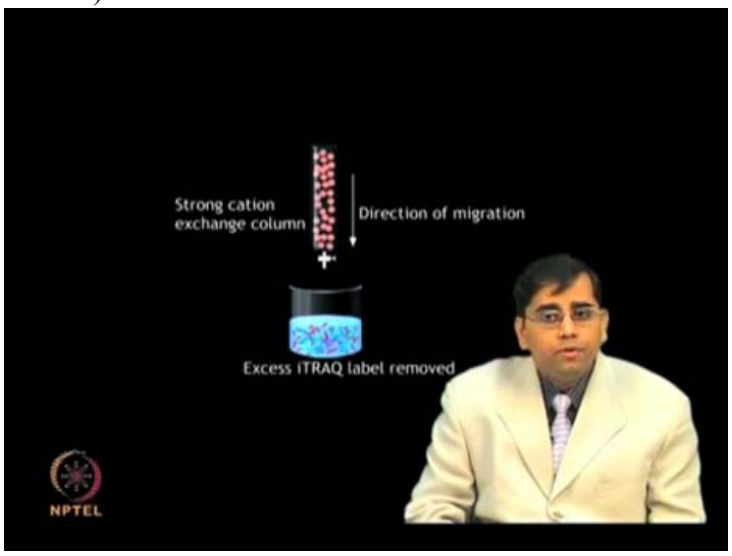
Pooled samples are verified on Strong Cation eXchange SCX column …

(Refer Slide Time 24:27)



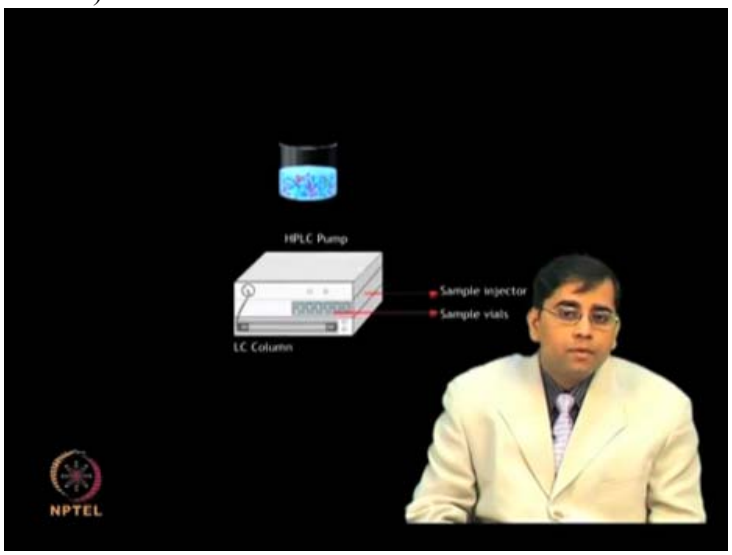
… to remove …

(Refer Slide Time 24:29)



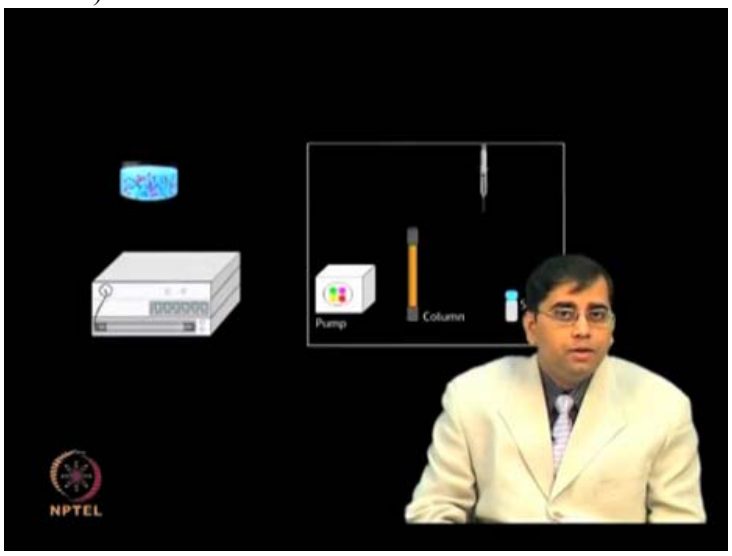
…. excess unbound reagent.

(Refer Slide Time 24:33)



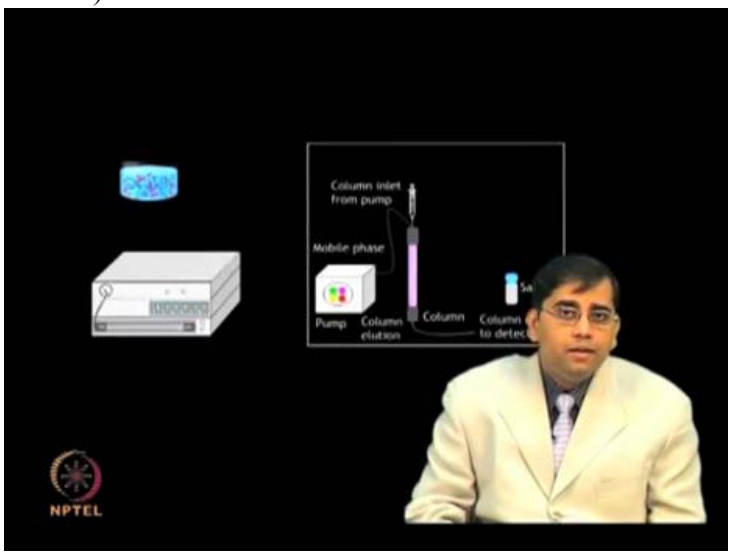
These isobaric reagents are detected upon ….

(Refer Slide Time 24:39)



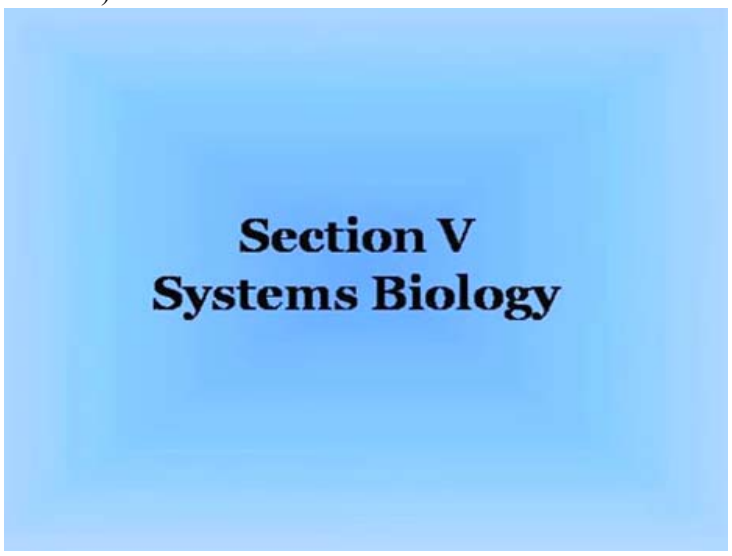
… fragmentation and release in ….

(Refer Slide Time 24:40)

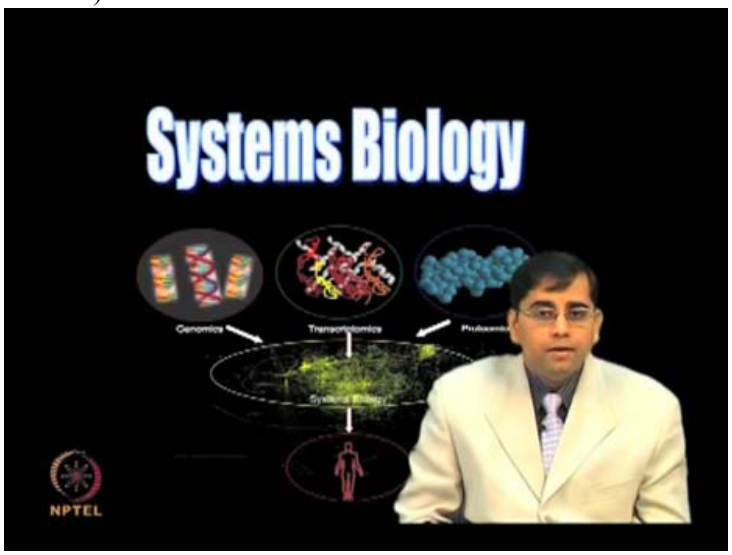


… mass spectrometry

(Refer Slide Time 24:43)

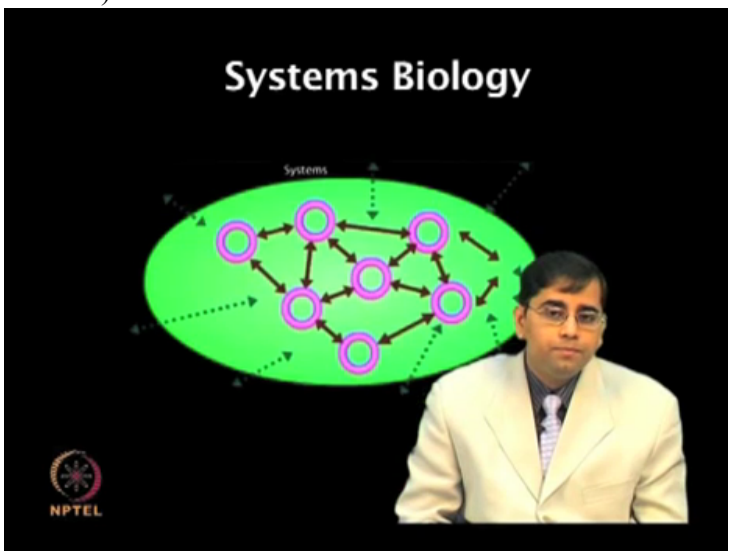


(Refer Slide Time 24:47)



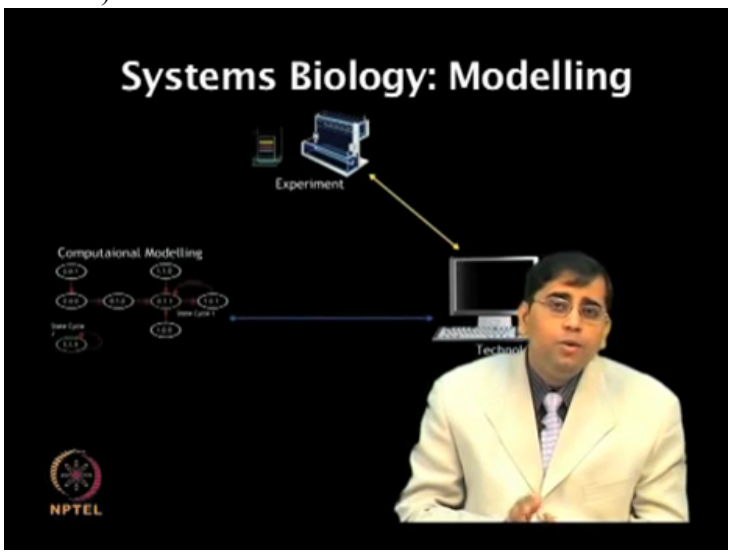
In Omics era, technological advancements in genomics, proteomics and metabolomics have generated large scale datasets in all the aspects of biology. These large datasets have motivated the computational biology and systems approaches with objective of understanding the biological system as a whole.

(Refer Slide Time 25:19)



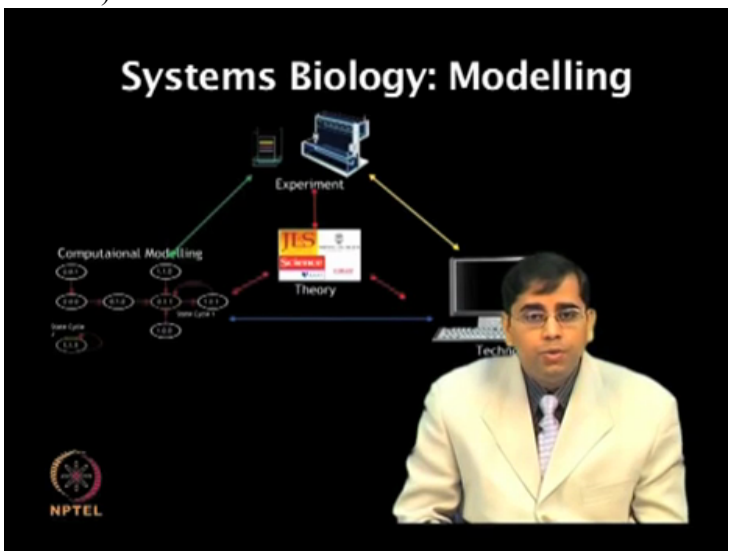
The Systems Biology and biological network modeling aims to understand the biological processes as whole system rather than isolated parts…

(Refer Slide Time 25:33)



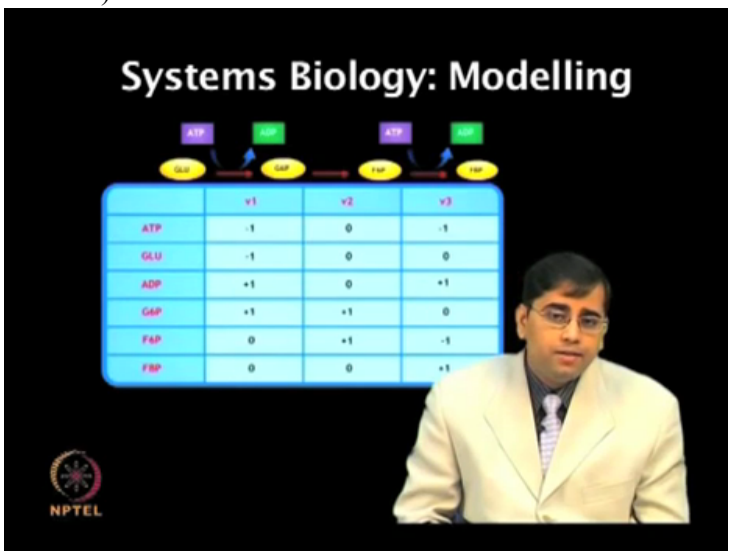
…by synergistic application of experiment, theory, technology…

(Refer Slide Time 25:39)



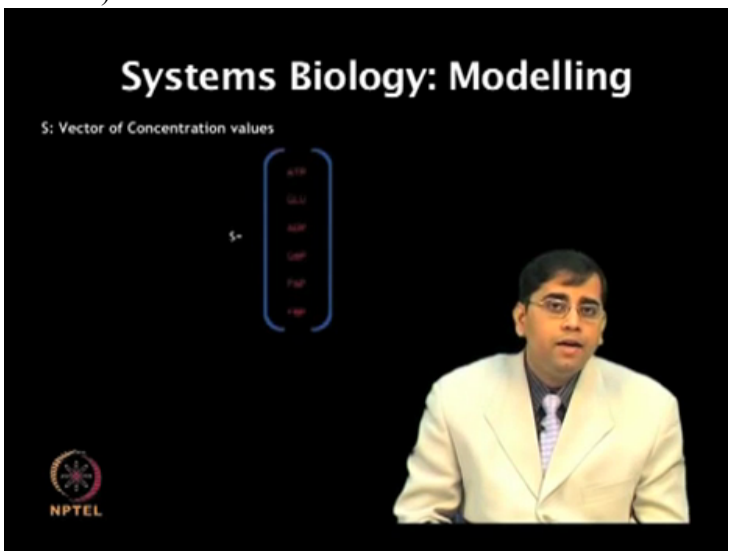
… and modeling

(Refer Slide Time 25:44)



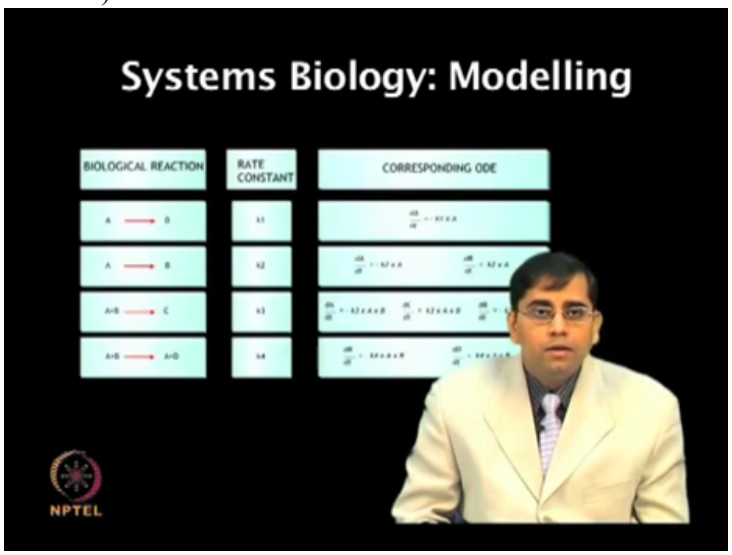
The Systems level studies aims to develop …

(Refer Slide Time 25:50)



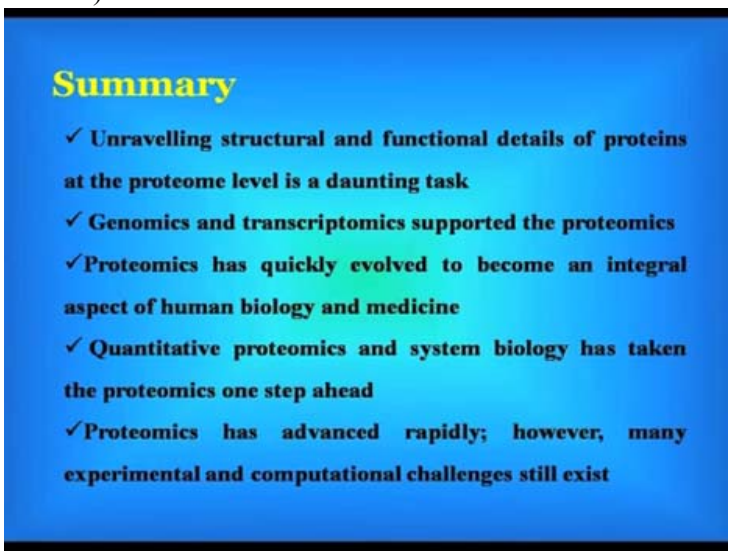
…. computationally efficient and reliable models of underlying gene regulatory networks

(Refer Slide Time 25:57)



The quantitative analysis measures and aims to make models for precise kinetic parameters of a system's network component. It also uses properties of network connectivity.

(Refer Slide Time 26:18)



(Refer Slide Time 26:28)

I hope it will enthuse you to learn about proteomic techniques and proteomic concepts. Thank you for your attention.