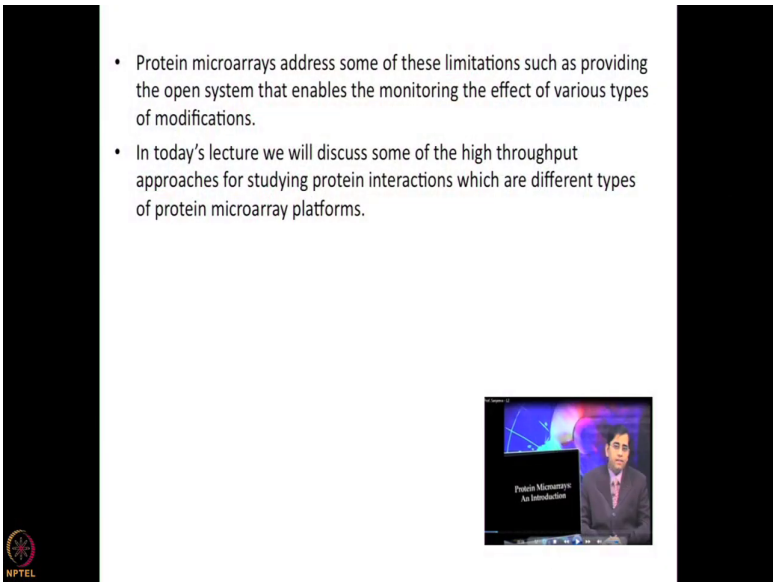
Interactomics Basics and Applications
Prof. Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay

Lecture – 03
High Throughput Platforms of Interactomics: Protein Arrays

Hello students. Today we are going to talk about the fascinating field of protein microarrays. If you remember from the previous lecture, we discussed about different types of traditional methods which have been used for studying the protein-protein interactions. Though many interactions have been discovered by using yeast two hybrid an immuno precipitation study, there has been reports for high false positive rate which poses a challenge on these technologies. Asides from these technical issues, both of these methods immuno precipitation and these two hybrids are primarily the endpoint assay where the whole assay occurs in a closed system inside the cells.

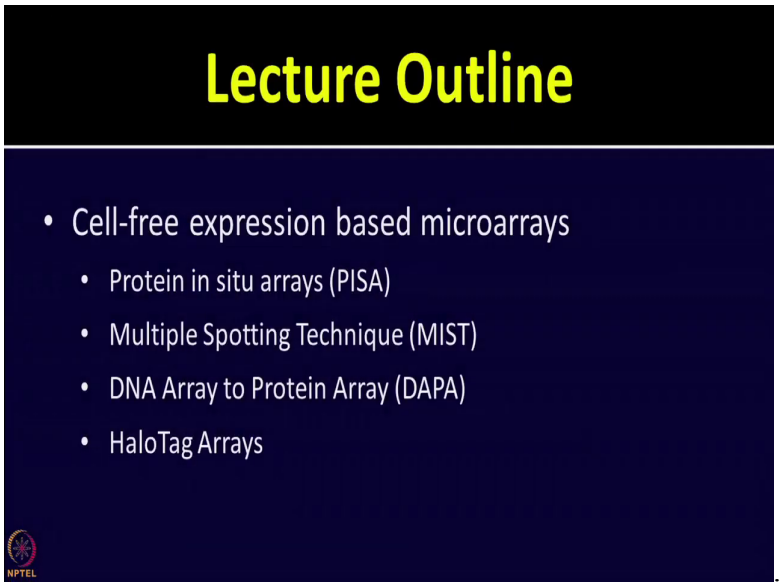
So, modulating the experimental conditions and different type of parameters becomes very challenging. The field of protein microarrays addresses some of these limitations such as providing the open system that enables monitoring the effect of various types of modifications.

(Refer Slide Time: 01:33)



In today's lecture, we will discuss some of the high throughput approaches for studying the protein interactions using different type of protein microarray platforms. I provide you an introduction into the powerful protein microarray platform.

(Refer Slide Time: 01:45)



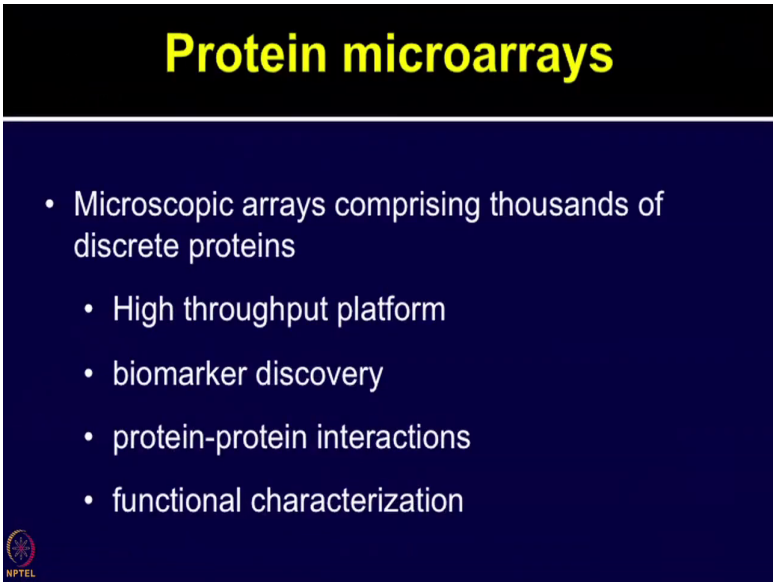
Discuss various types of protein microarrays, understand the basic workflow of any given protein array experiment and finally, we will go over the basic steps involved in data processing of a protein microarray experiment. The high throughput genomic and proteomic projects are so, called because they capture data and the scale of entire organism and incorporating data into relational databases from which insight into various biological systems, organization of physiological networks can be derived. Different types of hypotheses can be made based on these large data sets.

The genomic era has fostered the development of many new methods such as sequencing, SNPs as well as generation of DNA microarrays. The success of DNA microarrays at the time when most of the genes were sequenced, it was almost in the year 2000 and 2003 when we had availability of all the gene sequences. At the time, DNA array technology reached to its

maximum potential because it was very easy to screen thousands of genes and full genome of any organism such as human, for which almost 20,000 genes were already available.

So, by using the DNA microarrays, scientists have shown the potential of high throughput genomic technologies. The success of genomic technologies such as the DNA microarrays have motivated the development of protein microarrays. I think the good idea to add here that genomic is differently advanced and its leading the way and proteomic technologies are trying to follow the same path.

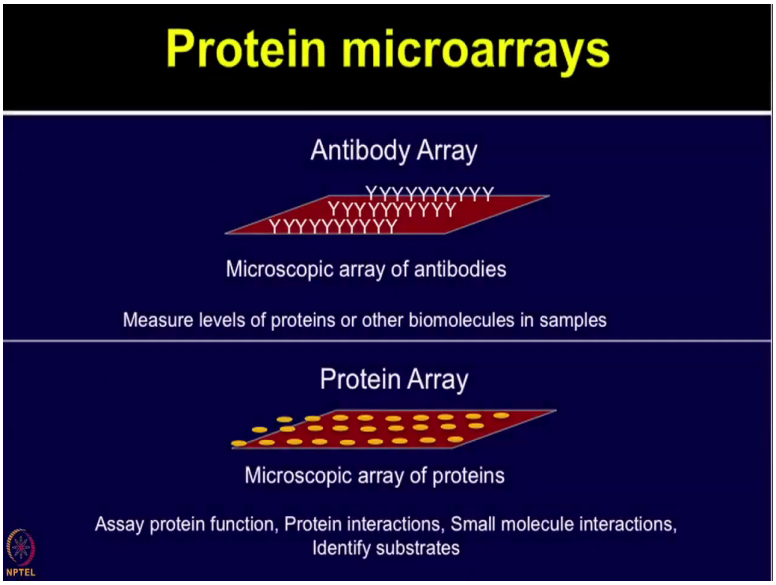
(Refer Slide Time: 04:01)



In this light, scientists thought whether we could also repeat the success of DNA microarrays at the protein array level. So, two protein microarrays; protein microarrays are microscopic arrays which comprise thousands of discreet proteins.

Now, the concept of micro arrays have listed a great deal of excitement in the proteomics community, because it can be applied for several applications such as biomarker discovery, protein-protein interactions, functional characterization of different proteins, identification of different substrates, drug inhibitors studies and all of them is possible in high throughput manner. So, far the protein microarray technology is fully realized, it promises to enable the study of broad variety of protein features at an unprecedented pace and scale.

(Refer Slide Time: 05:05)



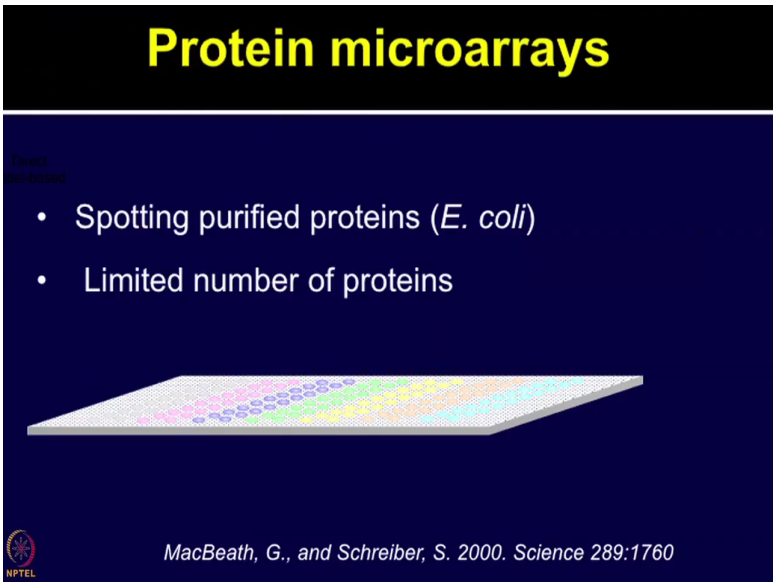
The protein microarrays fall into two general broad classes. The antibody arrays and test protein arrays. As shown in the side the antibody array is an abundance based method which is intended to inform the investigators, how much of each protein is present in each sample or to identify the proteins whose abundance is differentially expressed in one sample as compared to the other sample. There is a always a test control kind of study, where you want to compare the differential expression of the proteins. Antibody arrays, they print thousands

of antibodies on the chip surface and it has been used to measure the proteins or the biomolecules in different samples to compare control versus experimental conditions and looking at their abundance or protein expression.

In the test protein arrays, the proteins are spotted not the antibodies. As we talked in the earlier case, it is done by purifying the proteins first and then printing them on the chip surface. The goal of these test protein arrays is to perform functional studies. So, that different type of functions could be assigned, different type of biological questions which are related to the protein activities and its functions can be studied using test protein arrays.

So, these have been used assaying the protein function, protein interactions in studying the small molecular interactions as well as identification of substrates. As compared to the DNA microarrays which have shown its promises and potential in various biological applications. There are relatively fewer studies published using protein microarrays, but over the past few years there has been many studies many promising studies which have shown the potential of using protein microarrays for various biological and clinical applications.

(Refer Slide Time: 07:25)



The protein microarrays still remains very challenging just because the way of generating the content; the protein is very challenging. I am sure you agree that even purifying a single protein in the lab setting is not easy and if you think about protein purification for thousands of protein that is of course, going to be very challenging. So, scientists Gavin Macbeth at Harvard first demonstrated the feasibility of printing the proteins on the chip surface in a high density array similar to the DNA microarrays.

In 2000, at the time when DNA microarray technologies had started showing its potential, Gavin Macbeth published his work in science and showed that proteins could be printed on the chip very similar to what is being done in the DNA array technologies. So, why he was able to first time demonstrate?

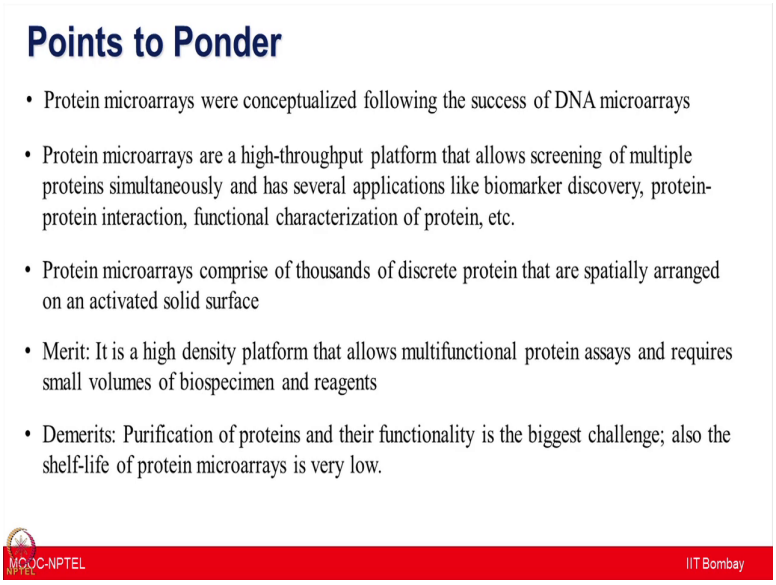
The utility of protein arrays and protein arrays could be made like DNA arrays, but still he was able to only spot two proteins one in the large number and one just to show the specificity of the assay, which also opened up the discussion that how challenging it will be to print thousands of proteins on the array especially, if you have to purify one of those proteins separately and then use those for the printing the chip.

So, protein contain generation definitely is one of the major challenges behind the success of protein microarray field. The success of Gavin Macbeth study motivated other scientists also to start doing the protein microarrays based research Doctor Mike Schneiders lab who which is also pioneer in the protein microarray field, they showed that whole each proteome could be printed on the chip by purifying 5800 yeast clones which were Histidine tagged.

And then they also showed that novel calmodulin and lipid binding proteins could be identified using this novel whole yeast proteome arrays. This was a full scale yeast protein array showed in 2001 the potential of these chip technologies for protein interactions and different type of other functional applications, I must say this was kind of a beginning of the very exciting field of protein microarrays when many investigators who had access to the large number of clones or access to the purified protein, they started thinking about using the miniaturised platform.

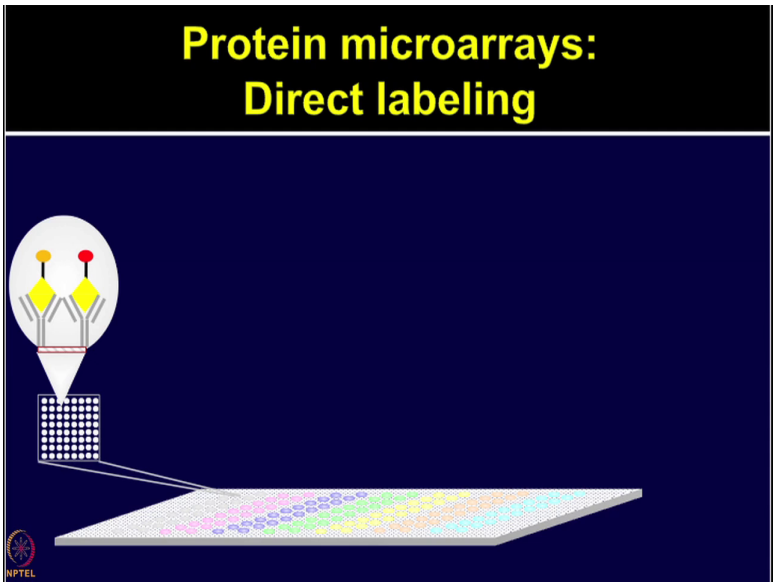
It starts printing the proteins on the chip and then performing some very interesting biochemistry at a very very small and minitaturised scale. But of course, the throughput will be very large because from the very small volume of the sample, you can now investigate the thousands of proteins which are printed on the chip surface.

(Refer Slide Time: 10:39)



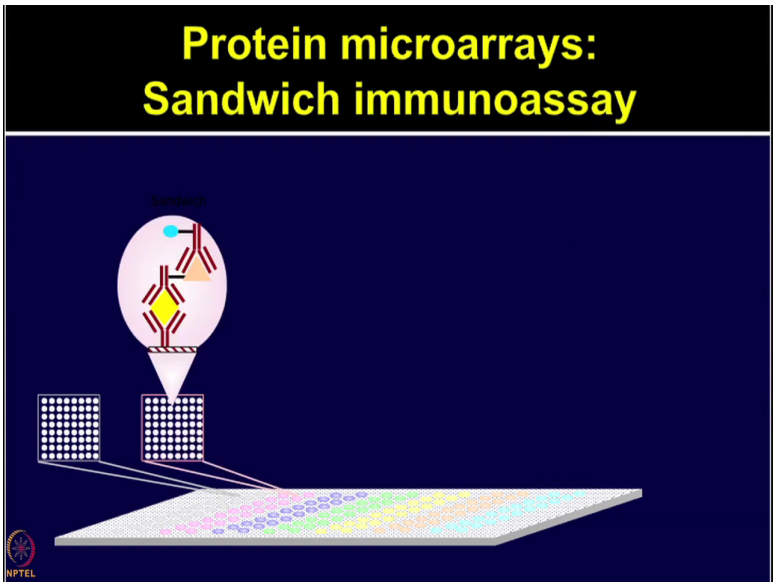
So, there are different types of platforms which are available for studying the proteins using protein microarrays. Let us have a quick look on some of these available platforms..

(Refer Slide Time: 11:29)



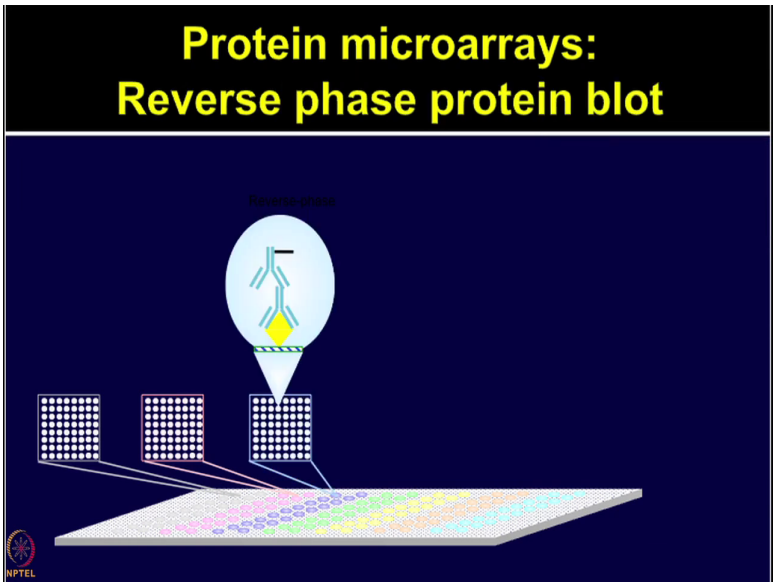
Let us first start with the abundance based arrays specifically the use of antibodies for making protein arrays. So, antibodies have been used to print on the chip surface for various proteomics applications in different type of orientations. The one shown here in the slide shows the direct labeling where the target proteins are labeled with fluorescence or other tag, which allows detection after it is captured by the antibody immobilized on the array surface.

(Refer Slide Time: 12:05)



What is shown here is the sandwich immunoassay, in which the target protein is captured by an antibody followed by detection with labeled secondary antibody.

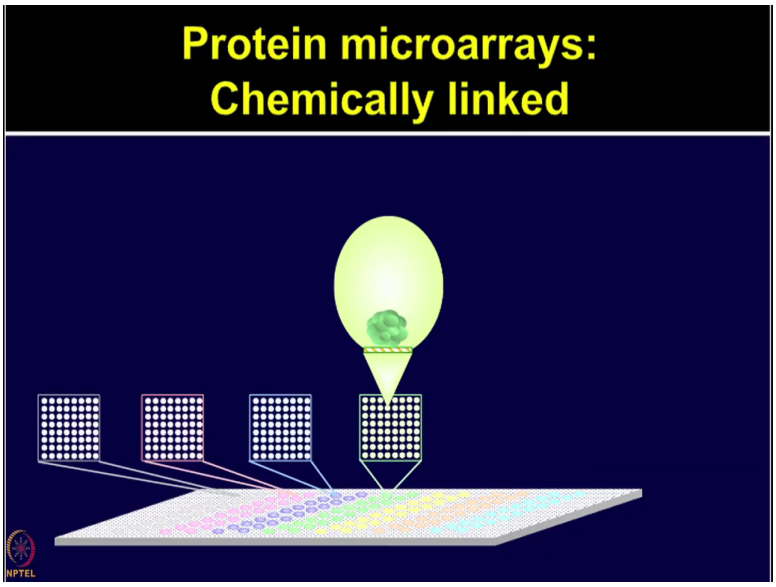
(Refer Slide Time: 12:17)



In the reverse phase protein arrays, the complex mixture such as a cell lysate or the tissue lysate is printed and are probed with a specific detection label. All of these three methods direct labeling sandwich immunoassay and reverse phase protein arrays they rely on antibodies. While antibodies and antigen interactions are very strong and gives a very easy way of doing biochemical based assay, very strong biochemistry can be done if you have access to good antibodies. But access of good antibodies especially both the quality and the cost in the large number will definitely a challenge.

Therefore, scientists have started exploring different methods of printing protein on the chip surface for different types of applications.

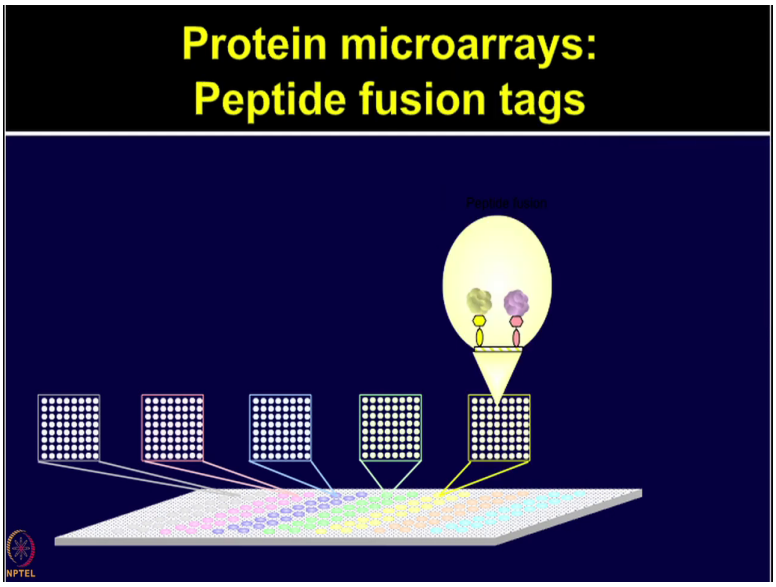
(Refer Slide Time: 13:15)



The very conventional or most widely used method for printing the proteins involved chemical linkage. The purified proteins are immobilized on the functionalized glass light and it can be used for various applications. If one can purify large number of proteins then of course, this will be the most ideal system when the purified proteins could be printed and now you can use it for different type of applications.

But because of the challenges involved in doing the protein purification at a scale; this method remains very very challenging. So, scientists have thought about more alternative ways of making protein arrays.

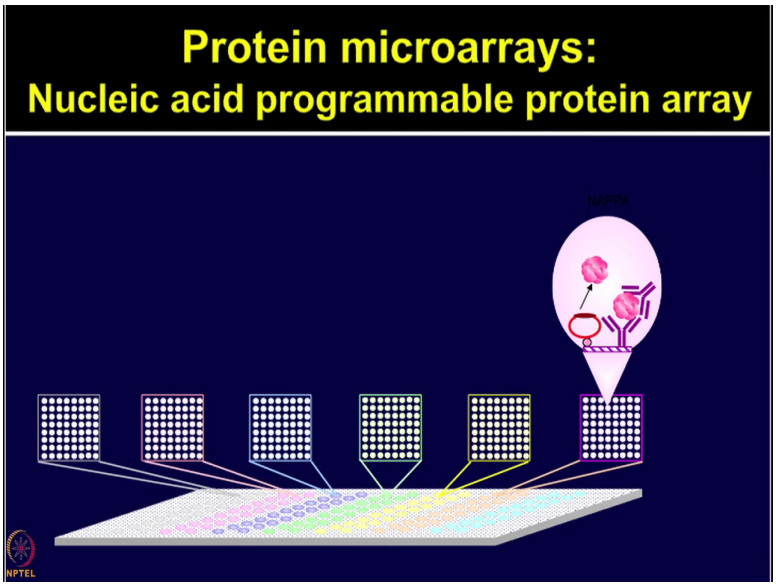
(Refer Slide Time: 14:01)



Peptide fusion tax is another approach. Many times if you are doing mass spectrometry based proteomics, you are a specifically identifying only peptide sequences that some of these peptides which are relevant for a given protein are showing differential expression. So, now, you would like to only take those peptide sequences synthesized peptide fusion tags and those could be printed on the array platform to make peptide arrays.

So, these peptides can be synthesized artificially, there are many company manufacturers or even different labs where these peptides could be synthesized then these are fused to the different type of tags like GST or Histidine tags and then spotted on the array platform. Depending on what type of tag you are using, you can also think about different chemistry for measuring those biomolecules and different type of interactions using these peptide arrays.

(Refer Slide Time: 15:03)



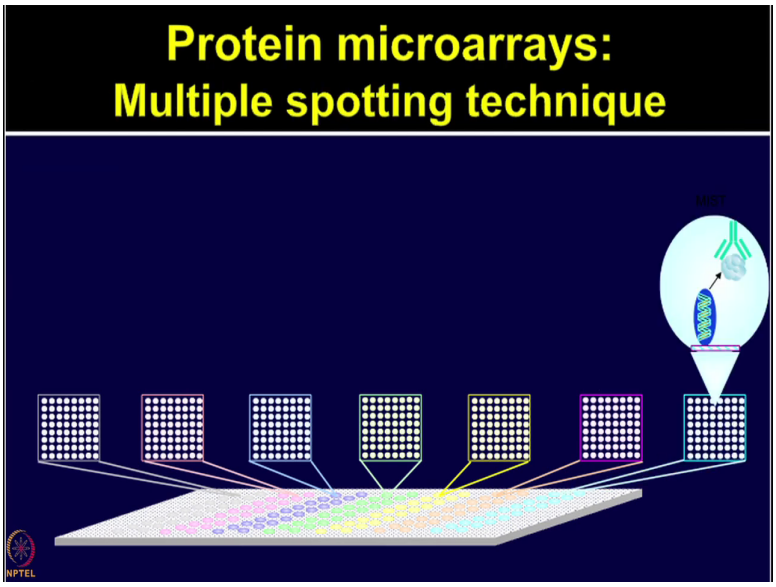
So, while have discussed about antibody arrays and purified proteins and peptide arrays.

However, due to the challenges involved in having good antibodies purified proteins or peptides, scientists thought about exploring alternative ways to eliminate the protein purification process. Doctor Joshua Labaers group at the Harvard at that time in 2003 2004 while now they have you know moved to Arizona State University, but this work was done at Harvard when they developed Nucleic Acid Programmable Protein Arrays or NAPPA in which the CDNA containing the GST tags are printed on the array surface along with capture and anti GST antibody.

Protein was expressed using cell free expression system and captured by the antibody. I will of course, talk about the cell free expression systems NAPPA and some of these concepts in much more detail later, but just you know to give the context here, we are simply trying to

take the DNA or the CDNA molecules adding the machinery which can do transcription and translation on the chip itself and then synthesizing the protein directly. This has very very revolutionary concept and of course, the work was published in science in 2004.

(Refer Slide Time: 16:33)

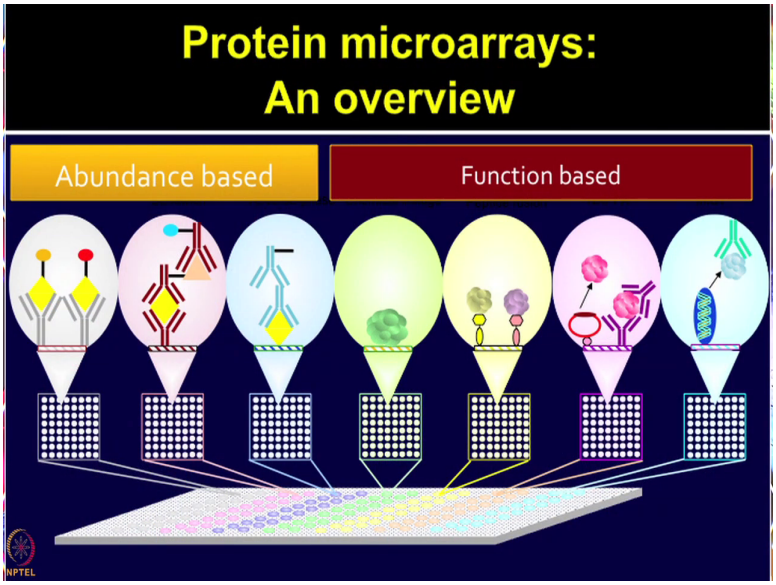


Another cell free expression based method which also tried to overcome the limitations of NAPPA.

Trying to print the CDNA along with the in vitro transcription translation mix, which is a process known as multiple spotting technique or mist. So, mist involves cell free expression in c two expression of the unpurified PCR products and the cell free lysates are printed on top of the spot so, that both in vitro transcription and translation can be performed in the chip surface. So, hope I have given you glimpse there is a different type of microarray platforms are available both using the purified proteins or even antibodies or one could also utilize even

cell free expression based systems and make use of the CDNA or even unpurified PCR products.

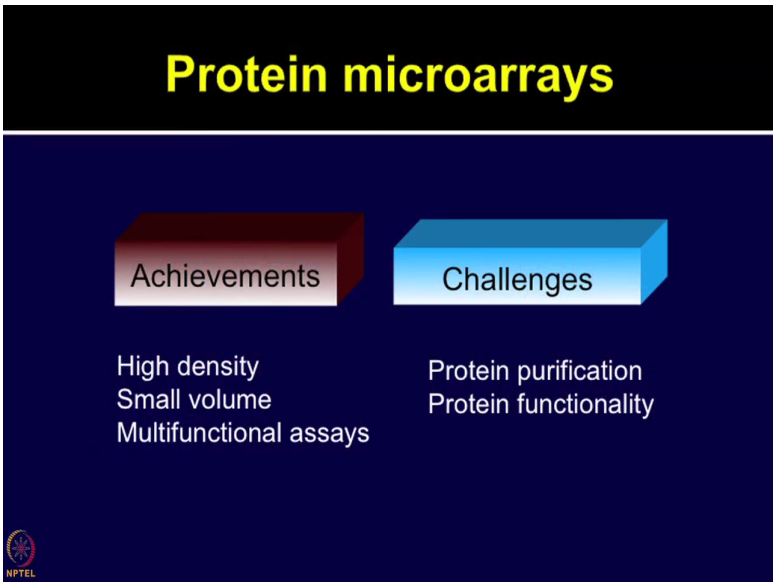
(Refer Slide Time: 17:43)



Broadly we can divide all of these things into two different type of protein microarray based system, one is abundance based arrays one is function based arrays. But we discussed about use of antibodies for direct labeling, sandwich inmmunoassays, reverse phase arrays in abundance based method and then we talked about chemically linked protein printing, peptide fusions and nucleic acid programmable protein arrays and multiple spotting techniques in function based protein microarrays. I hope now you got the sense that protein microarray field is very vast.

There are many ways of printing and measuring the proteins and utilizing this high throughput platform to study your biology and interesting biological questions.

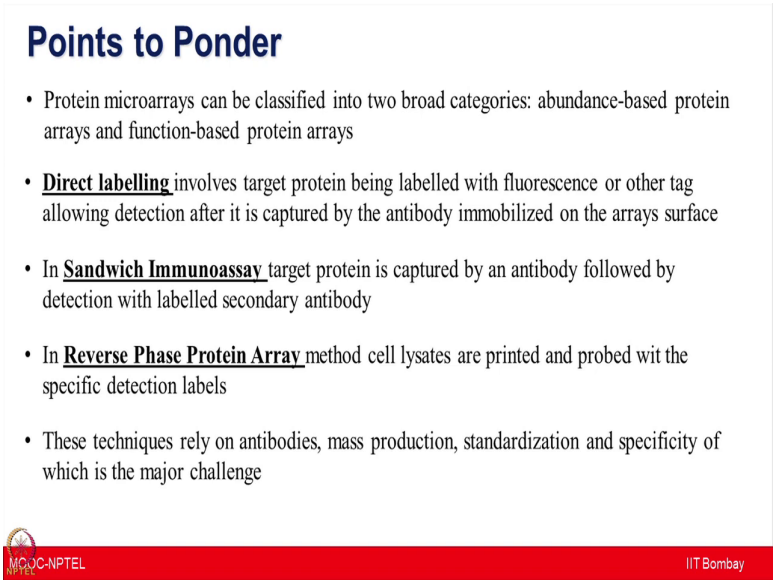
(Refer Slide Time: 18:29)



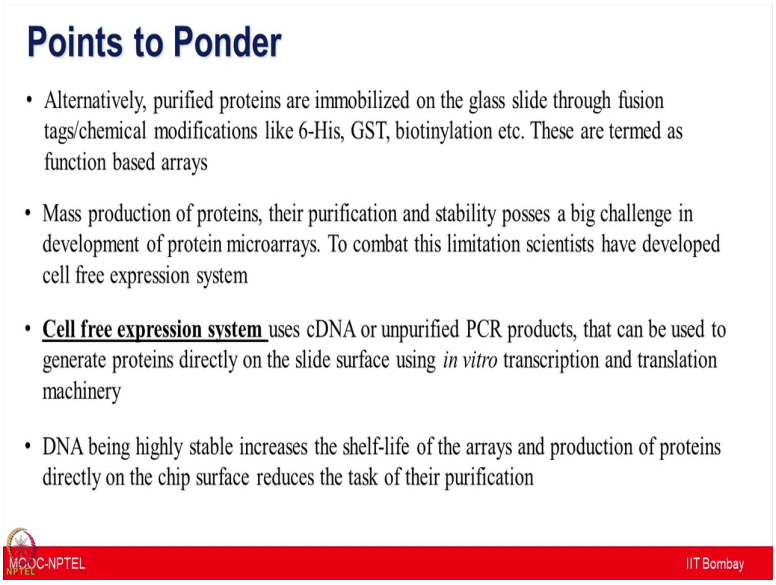
Protein microarrays have provided the high density, high throughput platform, which was one of the major achievements of this technology. The very small volume of clinical or biological sample which is usually the challenge or the drug molecule the pharmaceutical products can be used on these array surfaces and multi functional assays can be performed..

However, there are limitations and challenges if using the protein microarray which includes generating the protein contents, protein purification, how long we can store the proteins, keep it functional, whether they are going to degrade or the time period the printing quality whether that is going to get affected. So, many of these things are is still the major challenges and roadblocks in the success of protein microarrays.

(Refer Slide Time: 19:23)

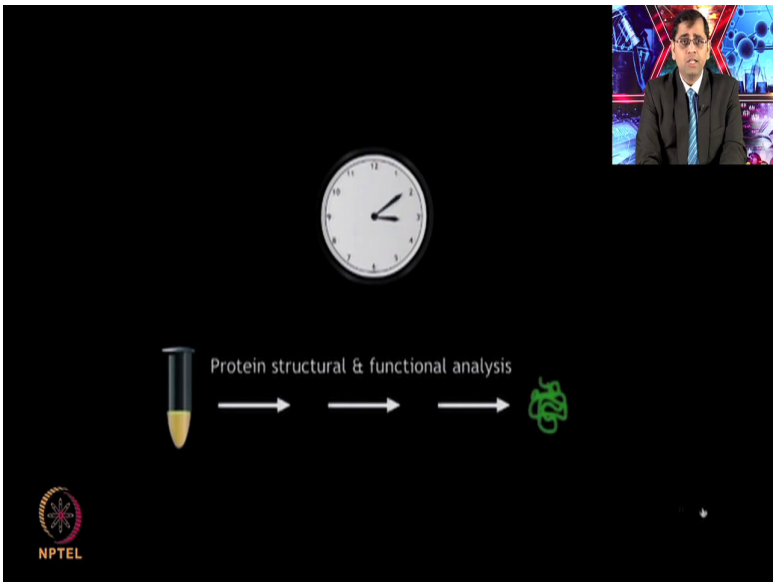


(Refer Slide Time: 20:01)



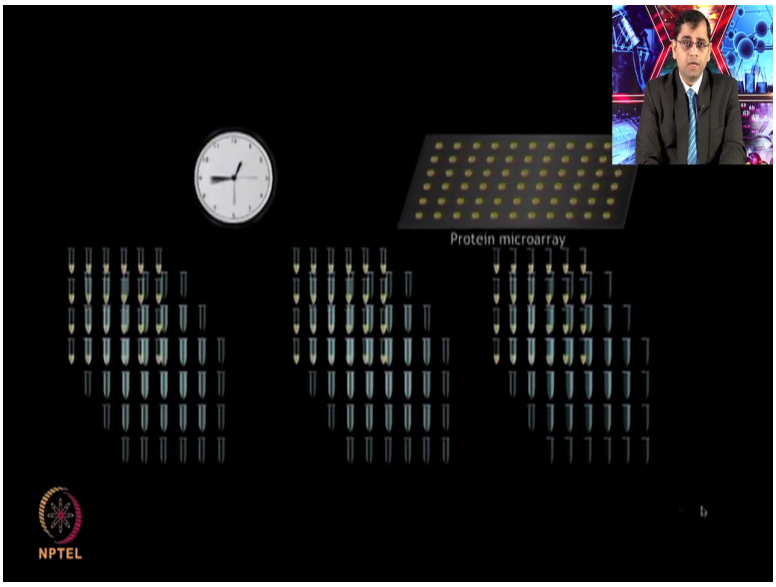
So, the development of protein microarrays on which thousands of discrete proteins are printed at high spatial density, definitely offers a novel tool to interrogate the protein function in high throughput manner. In this animation I will discuss different type of features, different type of processes in the protein microarrays and the need for protein microarrays.

(Refer Slide Time: 21:07)



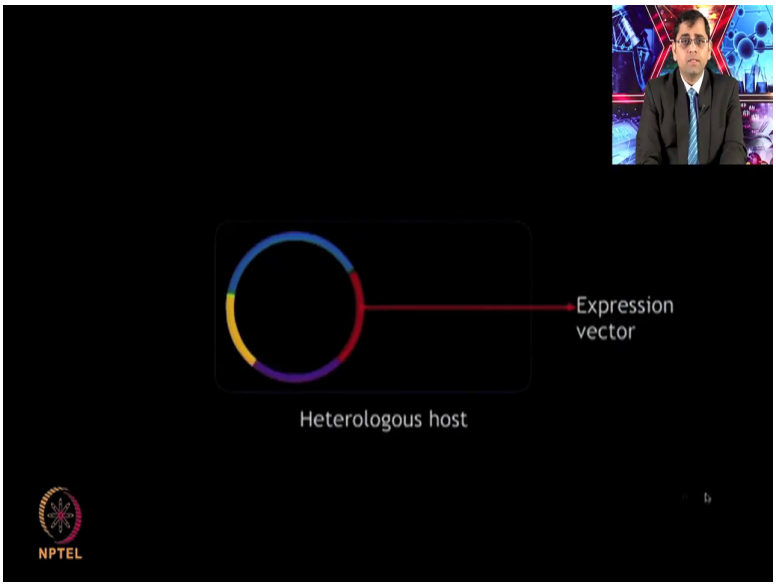
Functional analysis of proteins is a time consuming process that requires many expression. Analysis of a single protein at a time would be a tedious and laborious process. Analysis of several protein samples will undoubtedly take a long time if they are analyzed at a time.

(Refer Slide Time: 21:25)



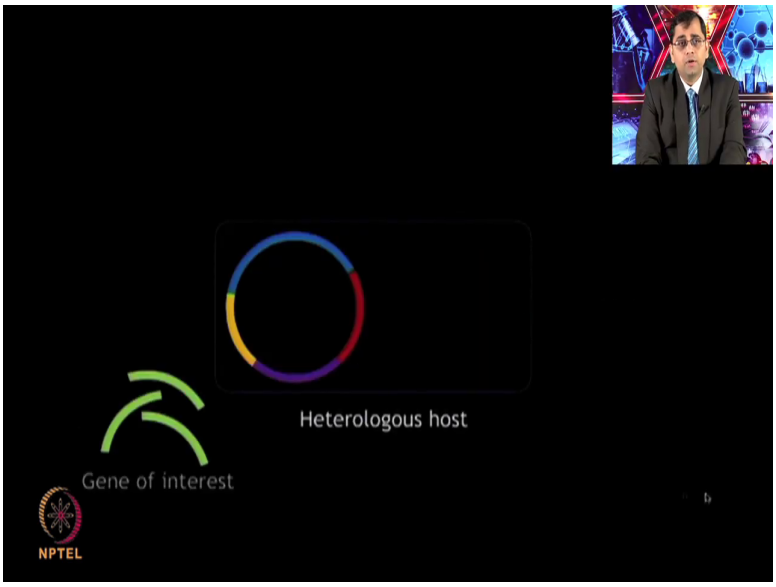
Protein microarrays have successfully overcome this hurdle by allowing analysis of several samples simultaneously..

(Refer Slide Time: 21:45)



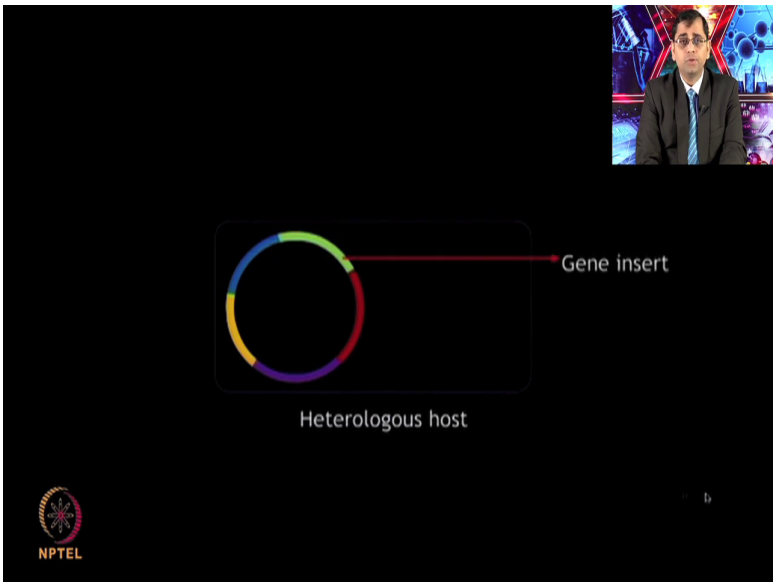
How to express the proteins and purify it?.

(Refer Slide Time: 21:47)

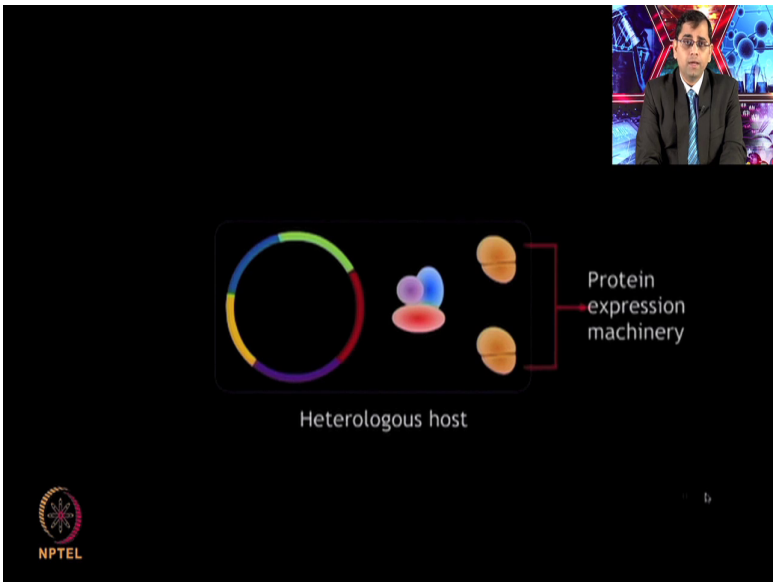


The gene coding for the protein of interest is expressed in a suitable heterologous system.

(Refer Slide Time: 21:51)

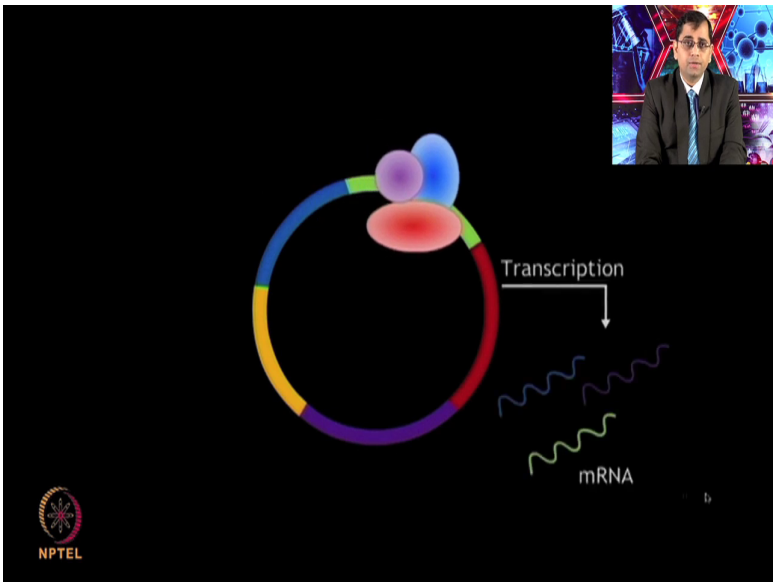


(Refer Slide Time: 21:55)



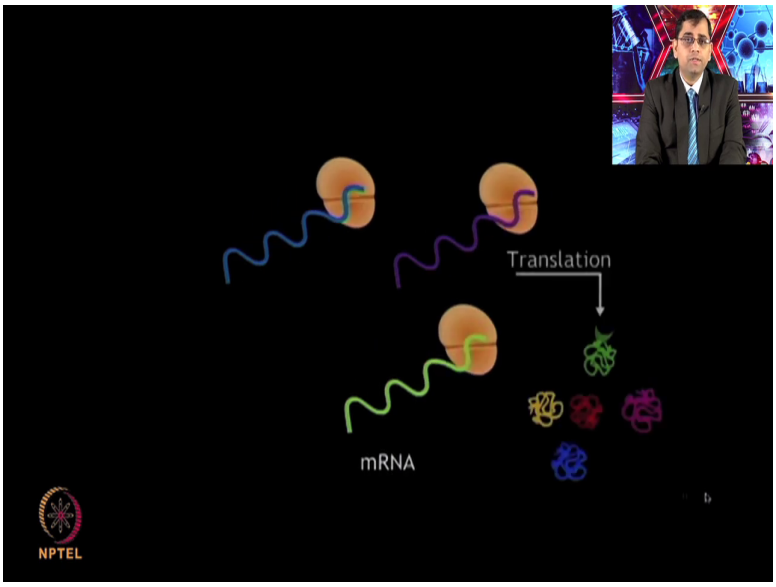
Such as ecoli by means of expression lipids like plasmids.

(Refer Slide Time: 22:01)



The host cell machinery is used for transcription and translation which results in a mixture of proteins consisting of the target protein along with other host proteins.

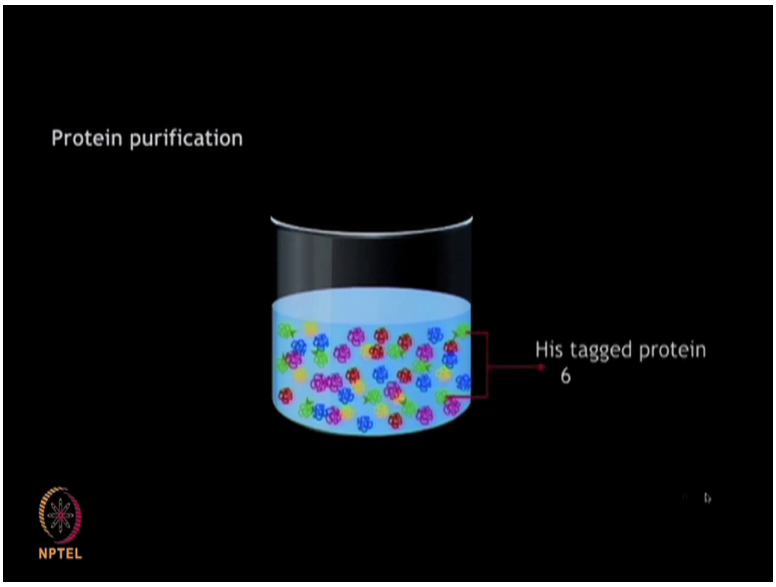
(Refer Slide Time: 22:07)



Since the protein of interest is expressed along with other proteins.

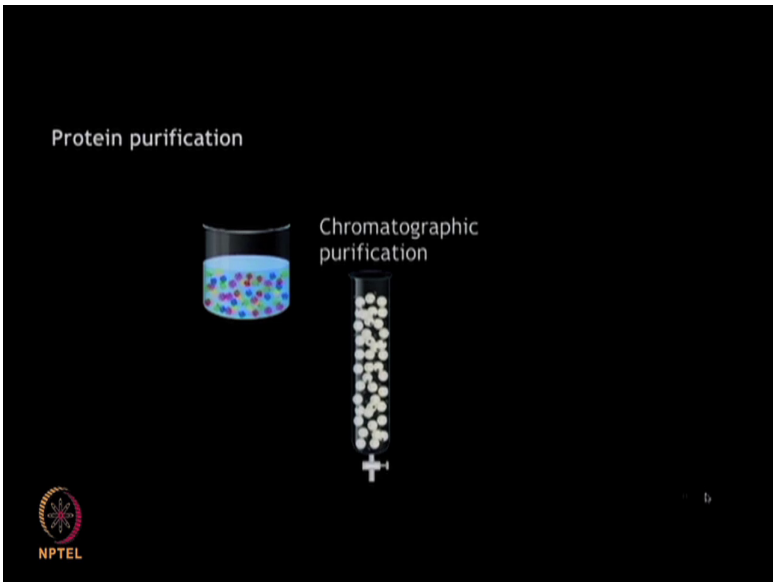
(Refer Slide Time: 22:13)

(Refer Slide Time: 22:15)



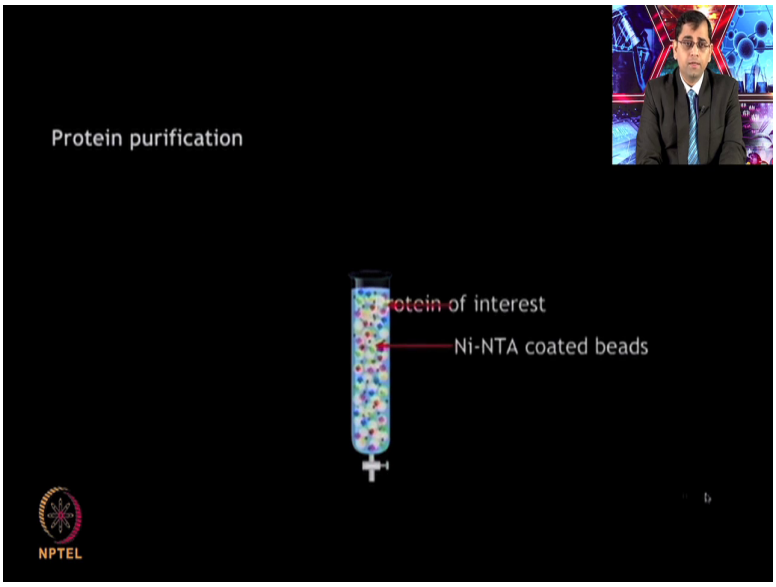
Native to the host it is essential to purify the target protein before it can be used for protein microarray applications.

(Refer Slide Time: 22:29)

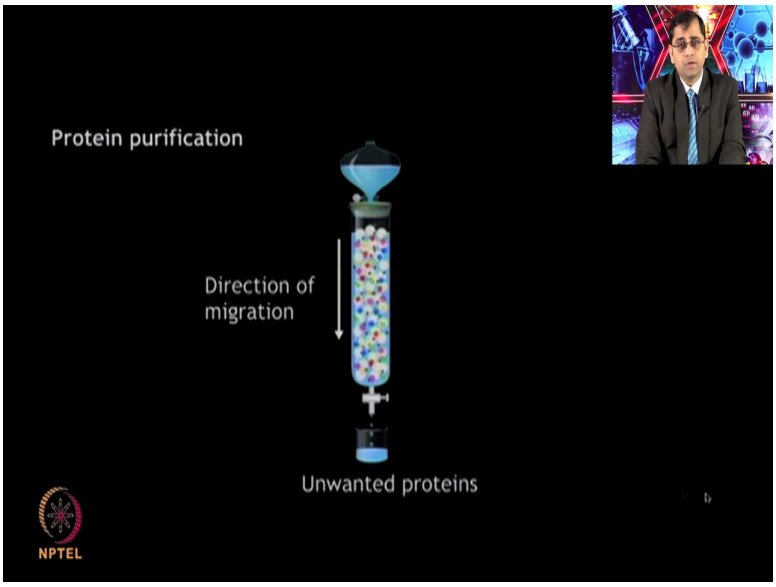


This can be done by chromotographic procedures to obtain the pure target protein.

(Refer Slide Time: 22:35)

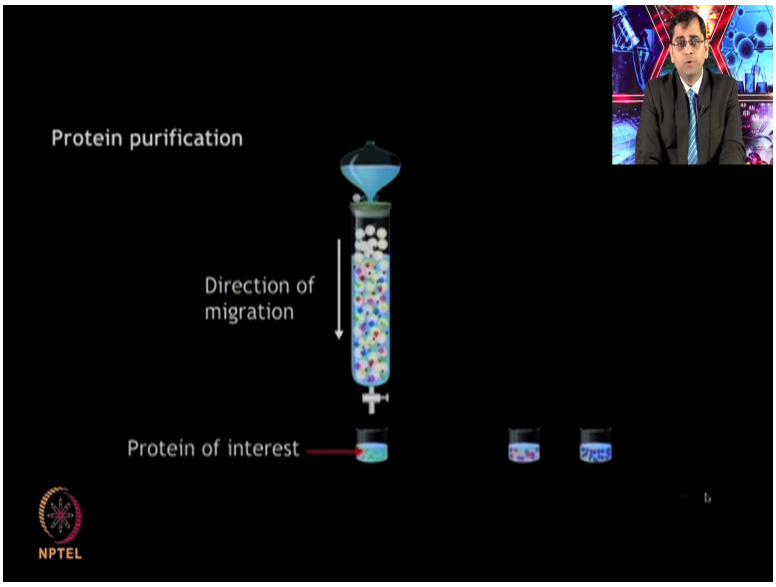


(Refer Slide Time: 22:39)

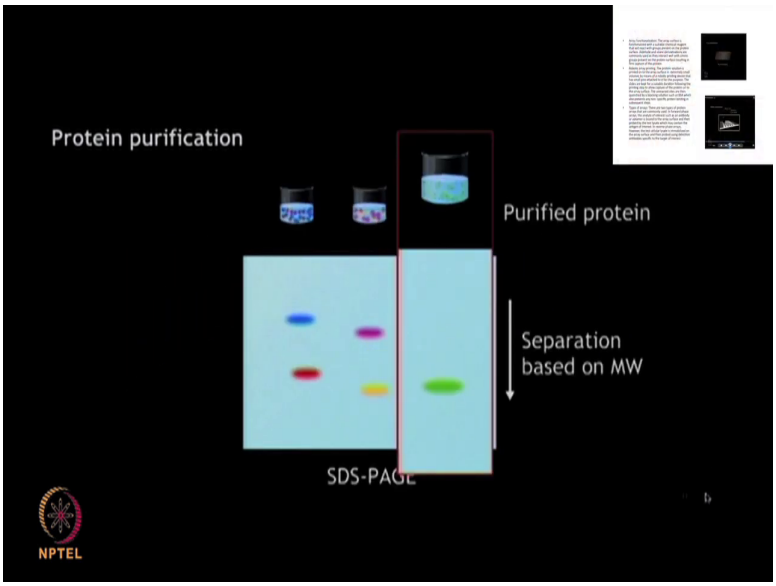


Protein purity is tested on HDS based gels tags such as histidine (Refer Time: 22:46) are often fused with the protein of interest to facilitate the purification process due to it is a specific affinity towards nickel.

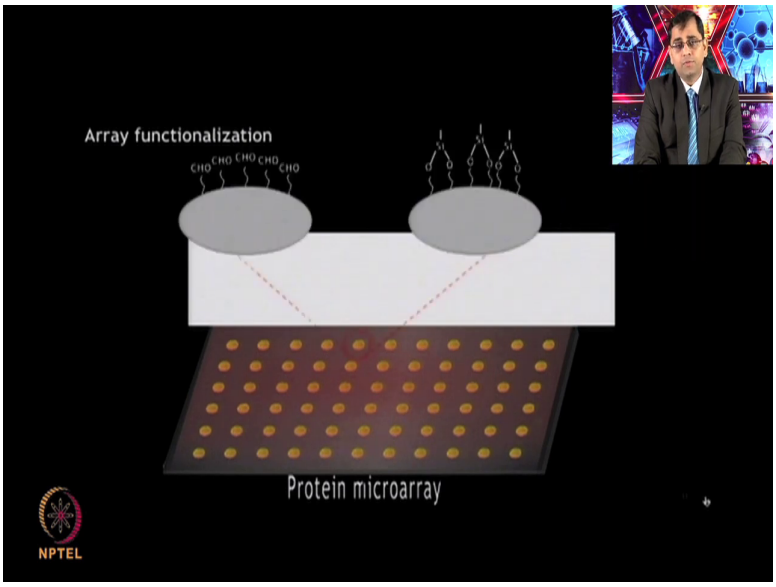
(Refer Slide Time: 22:49)



(Refer Slide Time: 22:57)

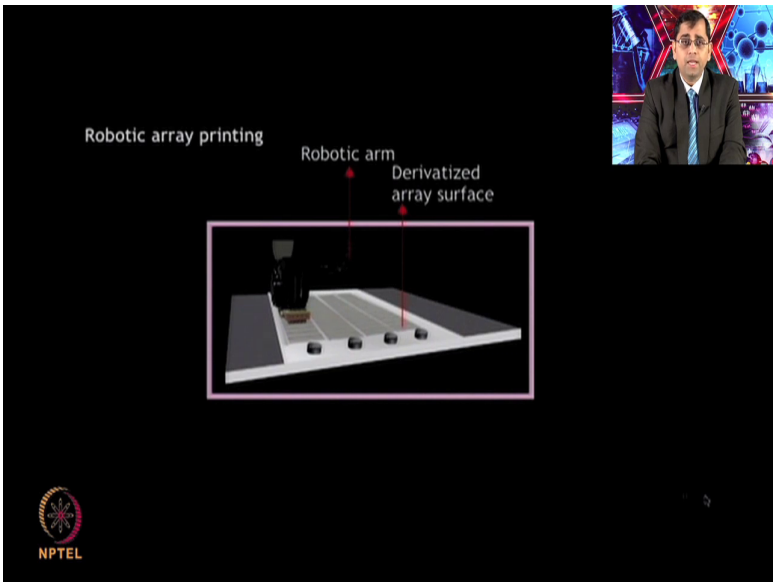

Protein purification
Purified protein
Separation
based on MW
SDS-PAGE
NPTEL

(Refer Slide Time: 23:01)



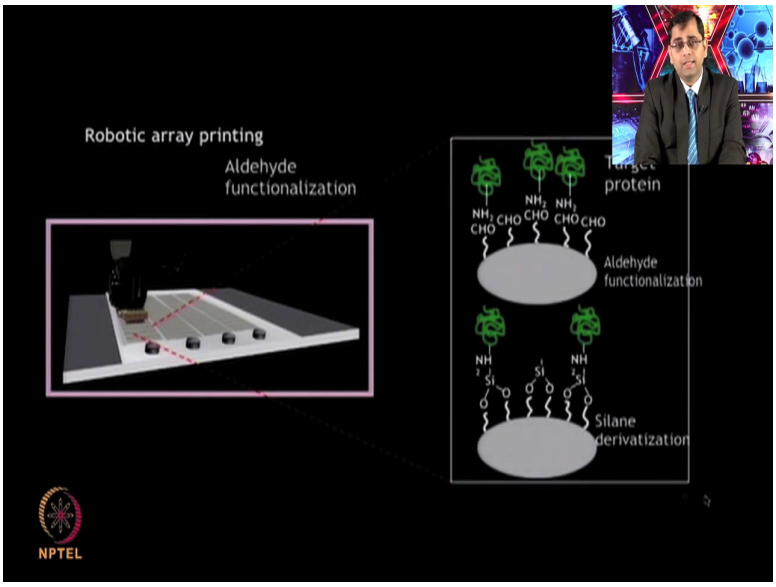
The array functionalisation: the array surface is functionalized with a suitable chemical reagent, that will react with groups present on the protein surface. Aldehyde and silane the rigidization are commonly used as they interact well with amino groups present on the protein surface resulting in firm capture of the proteins.

(Refer Slide Time: 23:25)



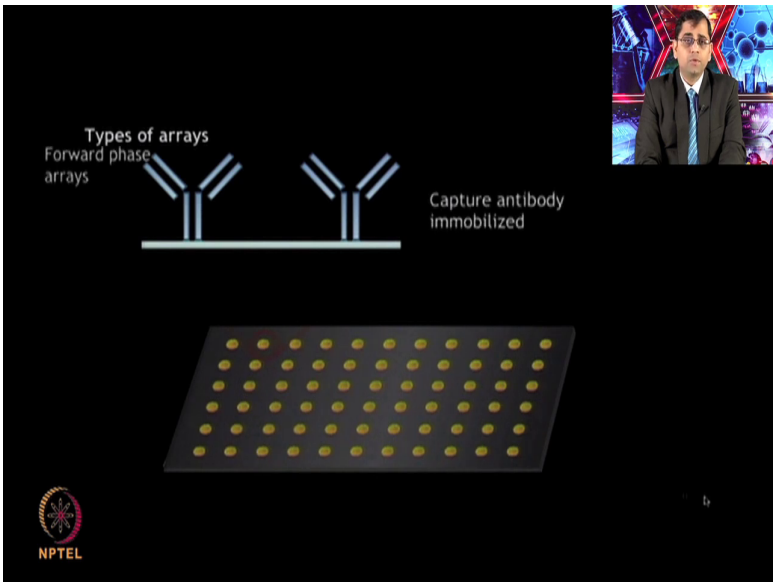
Robotic array printing:. Tyhe printing solution is printed on to the array surface in extremely small volume by means of a robotic printing device, that has a small prints attached to it for this purpose. Histidines are kept for a suitable duration following the printing steps to allow capture of the protein onto the array surface.

(Refer Slide Time: 23:49)



The unreacted side are then quenched by the blocking solution such as BSE which also prevents in a nonspecific protein binding in subsequent steps.

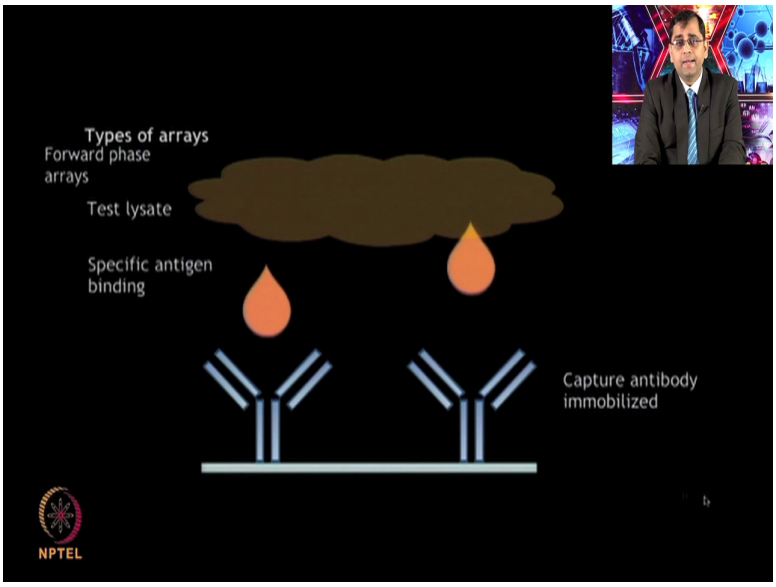
(Refer Slide Time: 24:09)



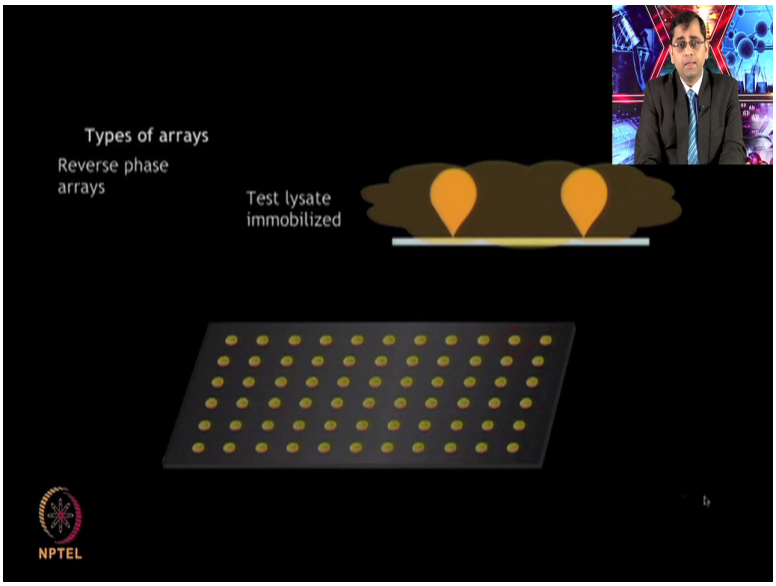
Types of arrays; there are two types of protein arrays that are commonly used.

In forward phase arrays the analytes of interest such as an antibody or optimer is bound to the array surface and then probed by the test lysate which may contain the antigen of interest.

(Refer Slide Time: 24:25)

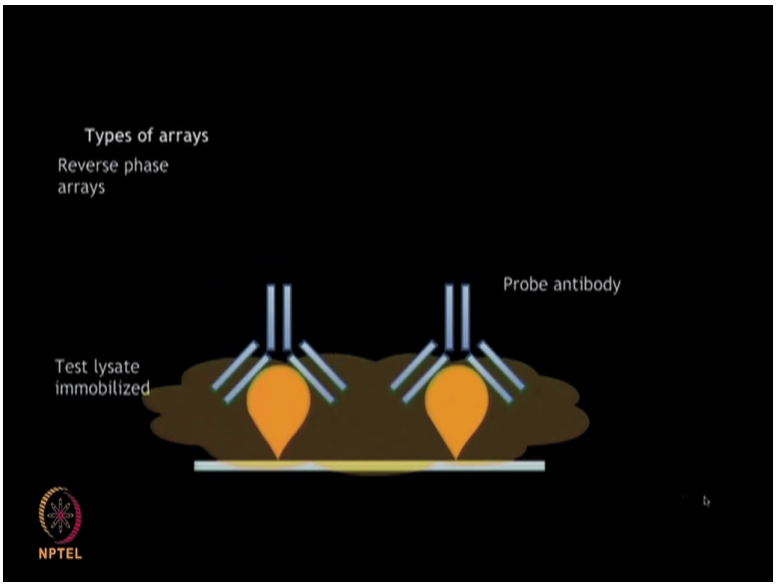


(Refer Slide Time: 24:33)



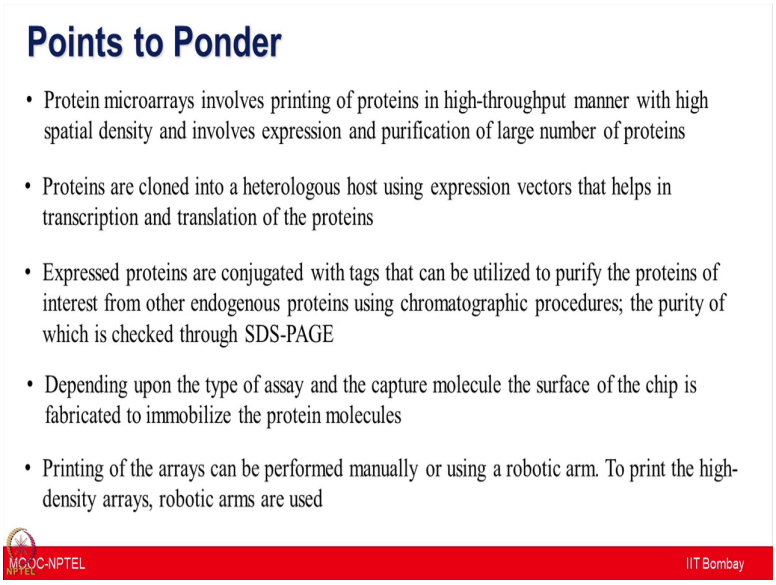
In reverse phase arrays however, the test cellular lysate is immobilized on the eraser case and then probed using detection antibody expected to the targets of interest.

(Refer Slide Time: 24:43)

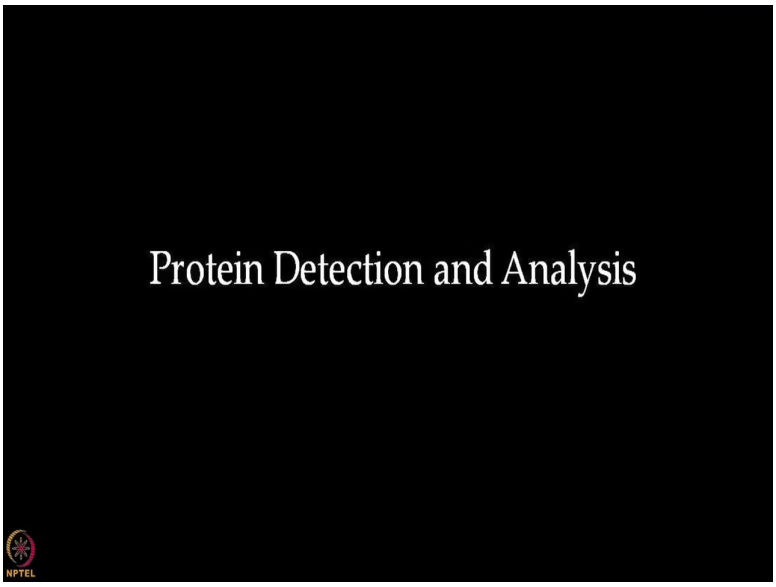


.

(Refer Slide Time: 24:47)

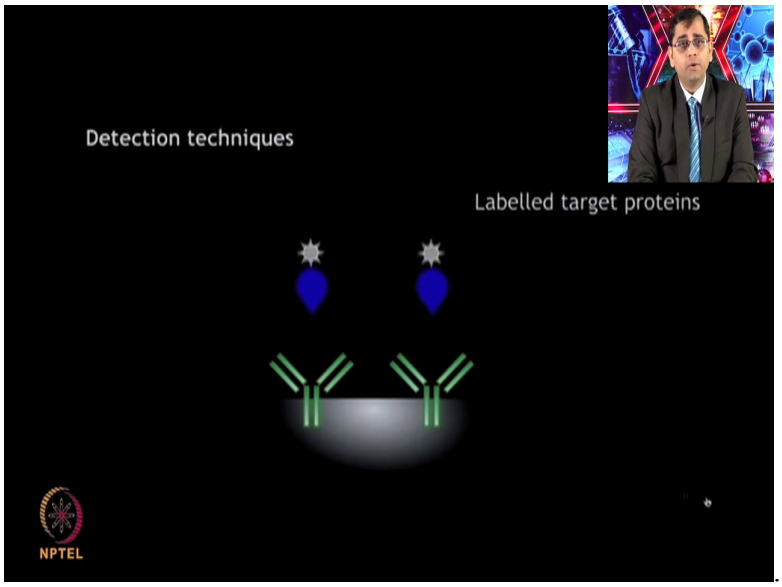


(Refer Slide Time: 25:27)

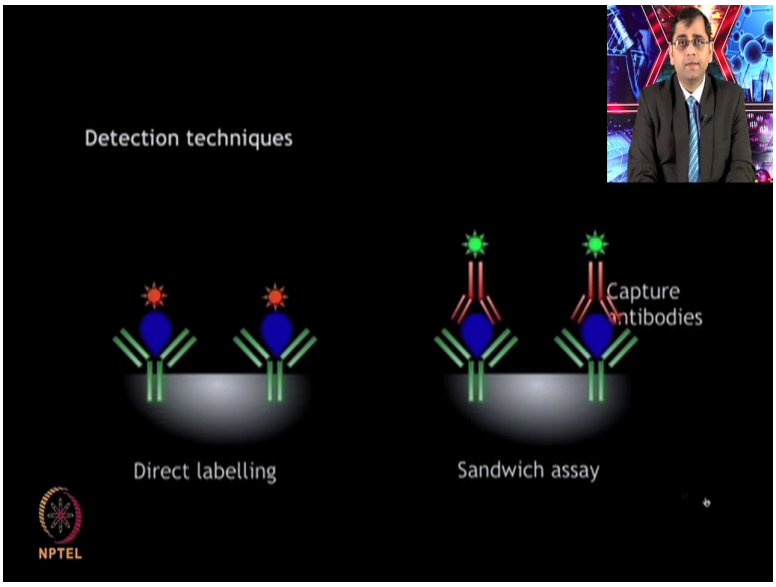


Protein detection and analysis in the direct labeling detection techniques all the target proteins are labeled with the fluorescence or radioactive tag that facilitates easy detection upon binding to the immobilized capture antibody on the array surface.

(Refer Slide Time: 25:31)

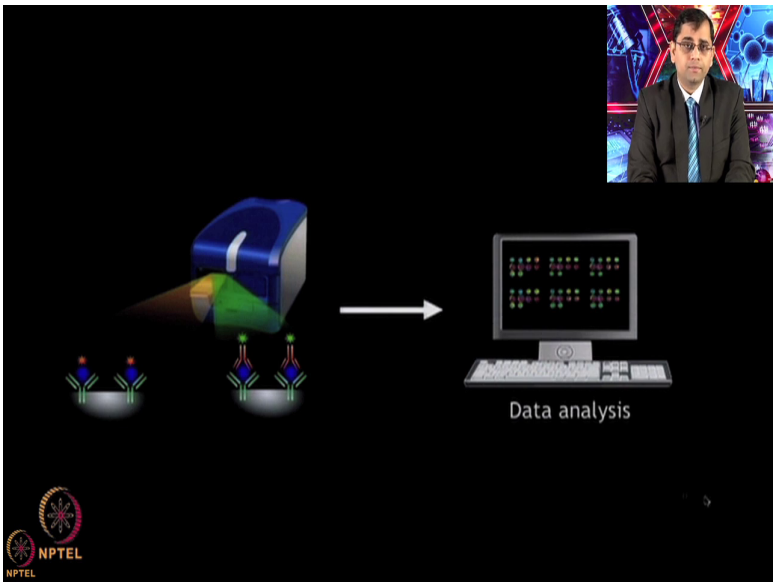


(Refer Slide Time: 25:53)



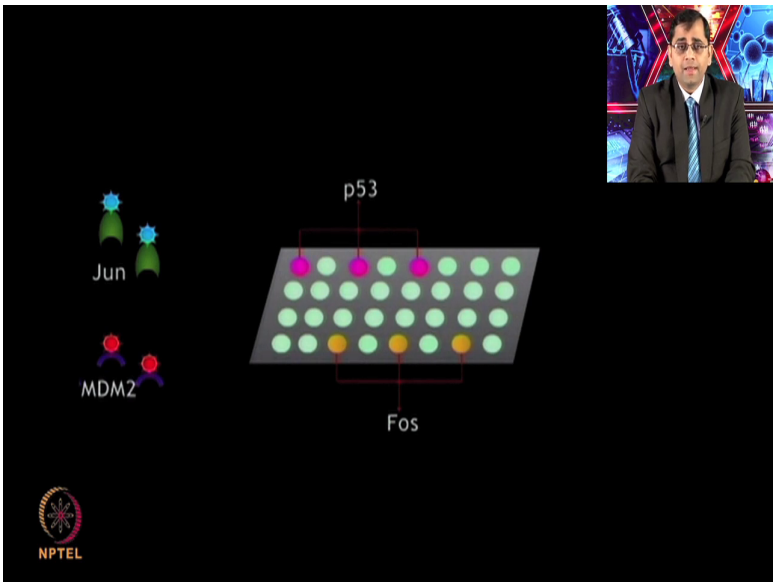
In the sandwich assay, however, the fluorescent tag secondary antibody that recognizes a different epitope and the target antigen binds to it and is detected by means of fluorescence.

(Refer Slide Time: 26:09)



The protein microarray is then scanned in a microarray scanner, that allows detection of the fluorescently labeled proteins or antibodies. The output from this scanner is then received by an appropriate software on which the data can be analyzed.

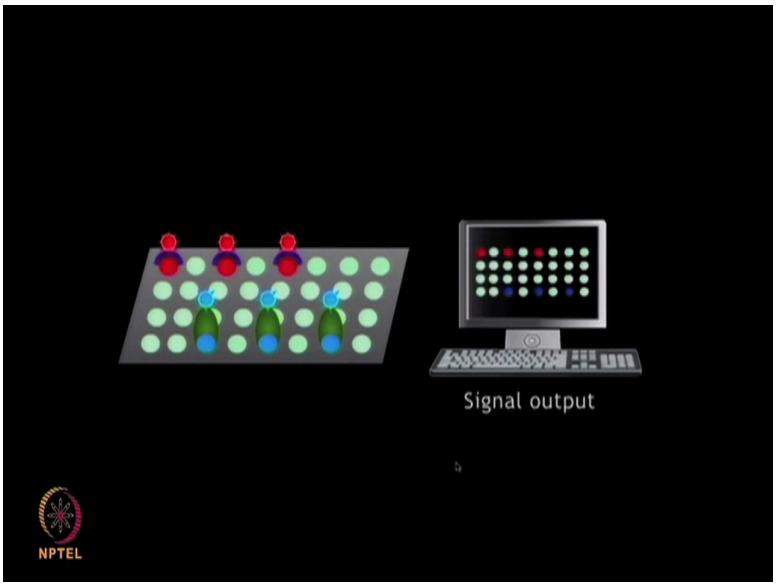
(Refer Slide Time: 26:29)



Certain well characterized proteins can be printed on the arrays as shown in the animation.

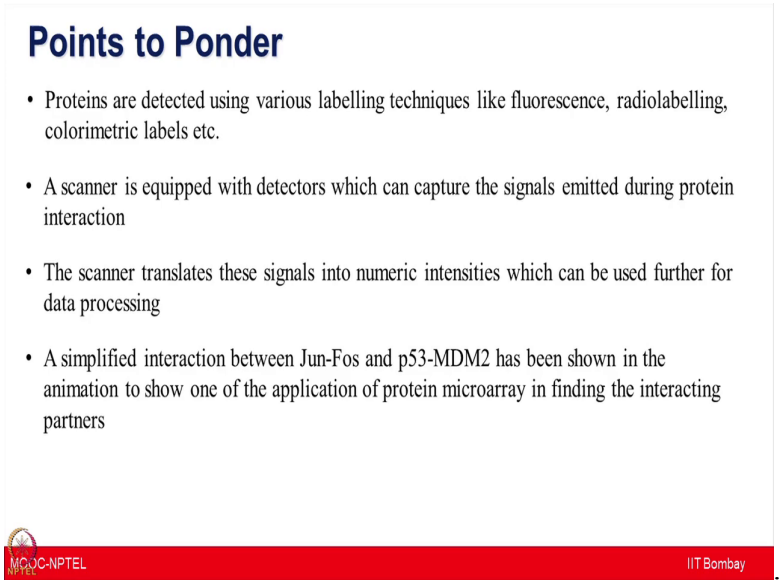
Now a proof-of-concept array is shown here where well characterized proteins are printed on the array surface along with their corresponding carry molecules shown on the left labeled with different fluorescent. (Refer Time: 26:55) Now by using this interactions let us match the protein interacting pair, Jun and for p 53 and MDM 2 by dragging the query to the correct protein on the array surface in order to see the signal output. So, the array surface there are both p 53 and for proteins present.

(Refer Slide Time: 27:21)



Now, if we drag the Jun protein it should interact with for protein. As you can see by the interactions here MDM 2 protein interact with p 53 protein while the interaction is established then these signals can be directed using it a micro array scanner.

(Refer Slide Time: 27:45)

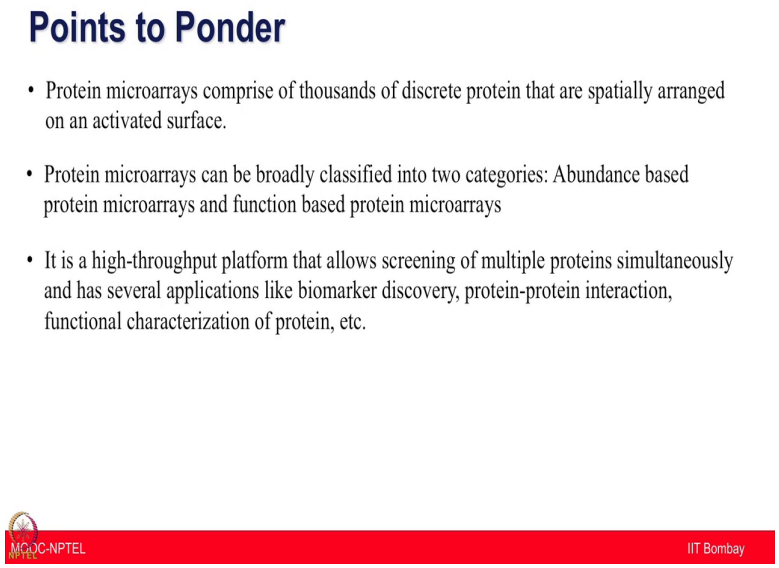


So, in summary today you got a glimpse of different type of features which can be printed on the array surface to make protein microarrays such as antibodies, purified proteins, peptides CDNA or even PCR products by using the cell free expression system.

We also got a glimpse of the history of the development of protein array field, which really got motivation from the DNA arrays. The entire workflow involves very very high throughput robotic ways of doing the work where large number of features has to be generated they have to be then printed on the chip and once you have made these arrays then it is very small volume of the samples, now you can test their antibody response or protein interactions, the substrate identification and various type of interesting multiple questions could be probed using these arrays in very very high throughput manner.

In the next lecture, we will talk more about the different type of cell free expression based protein microarray platforms.

(Refer Slide Time: 29:25)



Thank you.