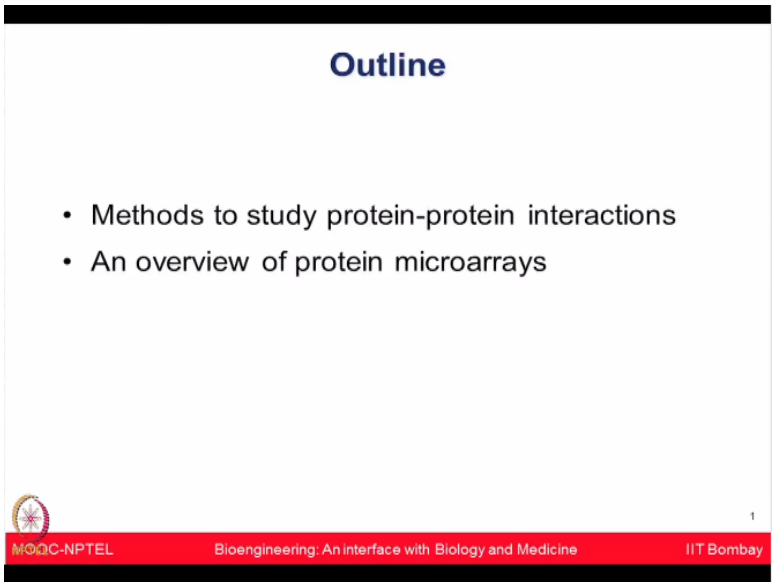
Bioengineering: An Interface with Biology and Medicine
Prof. Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology – Bombay

Lecture - 37
Protein Interactions & Microarrays

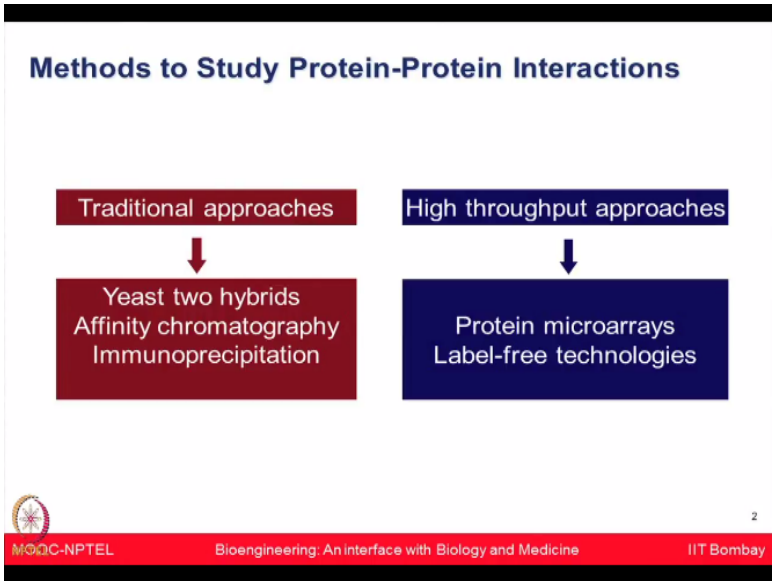
Hello everyone. Today, I am going to discuss the finer facets of functional proteomics and how it can be deciphered using an advanced protein microarray platforms. First, we are going to talk to you about methods to study protein-protein interactions.

(Refer Slide Time: 00:33)



And then I am going to provide you an overview of protein microarray platforms, how it can be used to study various interactions and other applications. Conventionally, there are 10 different methods which have been used to study protein-protein interactions.

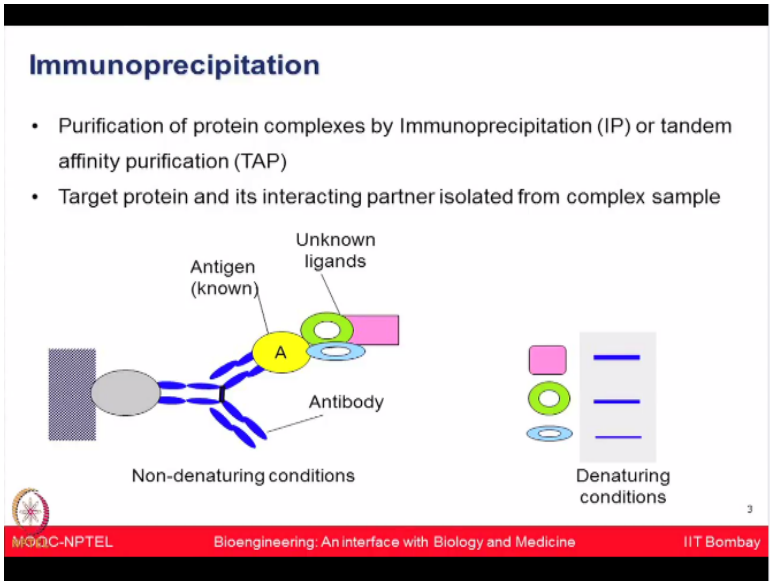
(Refer Slide Time: 00:52)



For example, traditional approaches including yeast two hybrids, affinity chromatography, immunoprecipitation methods were used. More recently there are high throughput methods like protein microarrays, surface plasmon resonance and other label-free technologies which have been used for studying biomolecular interactions.

So let me kind of you know give you a very brief overview of the conventional methods which have been used over long time for studying biomolecular interactions especially protein-protein interactions using immunoprecipitation method.

(Refer Slide Time: 01:23)



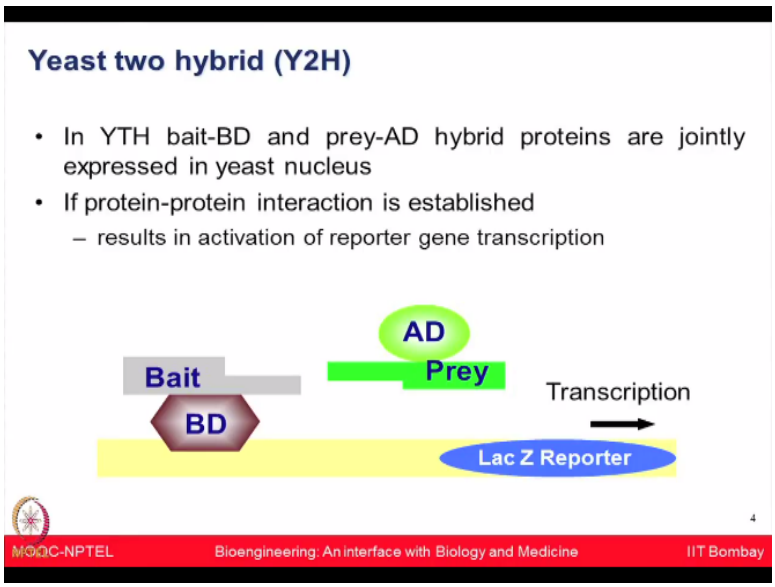
In immunoprecipitation method, the purification of protein complexes achieved by immunoprecipitation or tandem affinity purification or TAP method. The target protein and its interacting partners they are isolated from the complex tissue lysates or complex samples.

Then, they are separated on the denaturing gel SDS-PAGE and now you can see those are the shown on the image.

Different bands can be seen which are potentially interactors. How it is not very clear that you know these are the direct interactors or they could be sticky proteins which are also interacting with your proteins of interest. So therefore you will have the potential list but you are not very sure these are the real interactors or these are the even potentially sticky proteins. Another method which has been used in the past is yeast two-hybrid method.

In yeast two-hybrid, the best binding domain and the prey activation domain hybrid proteins are jointly expressed in the yeast nucleus.

(Refer Slide Time: 02:30)



And if the protein-protein interaction is established then it results into the activation of reporter gene and you can see the transcription initiation to happen. So that is something conventionally used. A lot of people have used this kind of you know yeast two-hybrid method; however, it results into you know large number of false positive as well because you will get you know big list of interactors but you are not entirely sure how many of them are actually directly interacted.

So these methods were used traditionally, they are quite easy to achieve. Nevertheless, you know they do give you the larger list from which you are very ensured that you know which are the right interactors for the given protein or which could be the potential protein but there

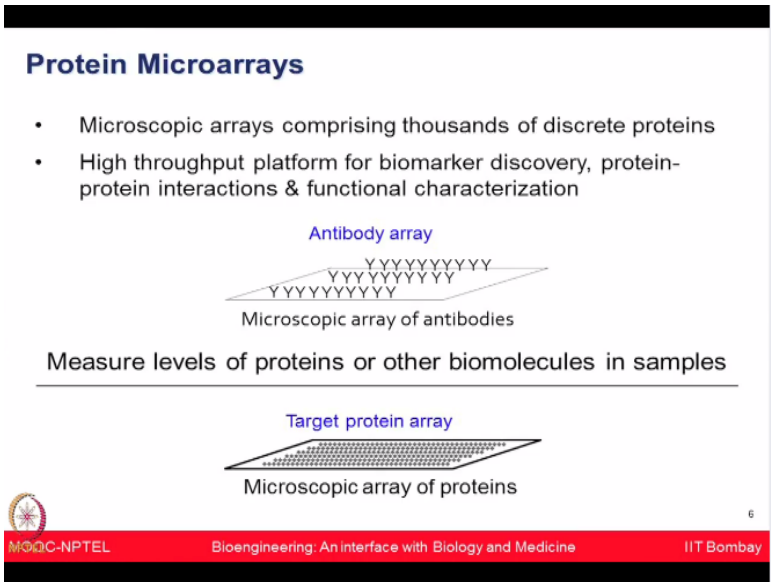
could be sticky proteins and they may not be the right proteins and could be false positives as well.

So with the advent of new technologies in the area of protein microarrays, there are different type of microarray platforms have been established to look into protein-protein interactions as well as many other clinical applications alright.

(Refer Slide Time: 03:31)

So let me give you an overview of a protein microarray platforms. So what are protein microarrays? These are microscopic arrays which comprised of 1000s of discreet proteins printed on the array surface.

(Refer Slide Time: 03:42)

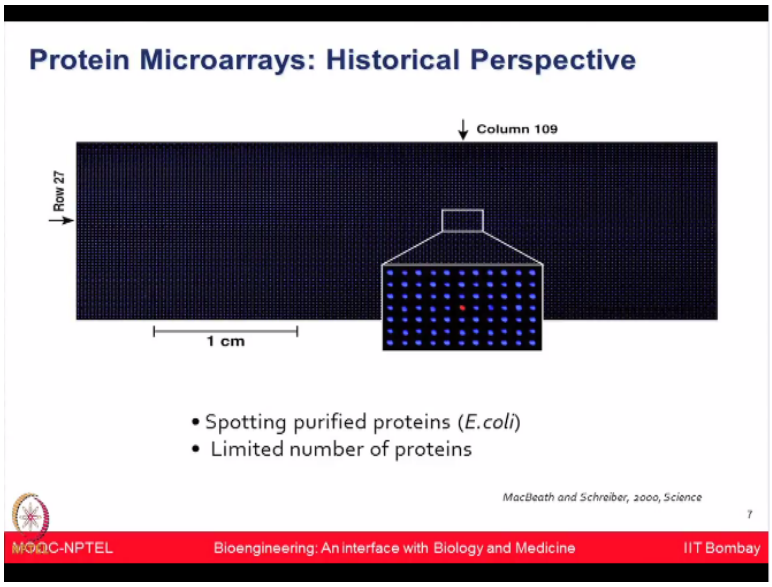


This high-throughput platforms could be used for many applications such as biomarker discovery, protein-protein interactions and various type of functional characterization. You can use different type of contents and accordingly your array platform could be termed as antibody arrays or protein arrays. For example, if antibodies are immobilized on the array surface that is an antibody array which can be used to measure the abundance of biomolecule.

For example, if you are looking at level of a given protein for which you have antibody, now if you are passing your biomolecule on the chip then you can measure the level or the concentration or the abundance of that particular protein using this antibody-based approach. Additionally, there has been target protein arrays where different forms of the functionalized proteins are printed on the array surface and those are known as the target protein arrays.

Historically, long back in 2000 Gavin MacBeath first time showed that you know it is possible to print the proteins on the chip. They printed only two proteins, one protein was printed in the large number of times, one protein printed once only and wanted to show the specificity of detecting even one protein on the chip among 1000s of features printed using E. coli purified protein.

(Refer Slide Time: 05:03)

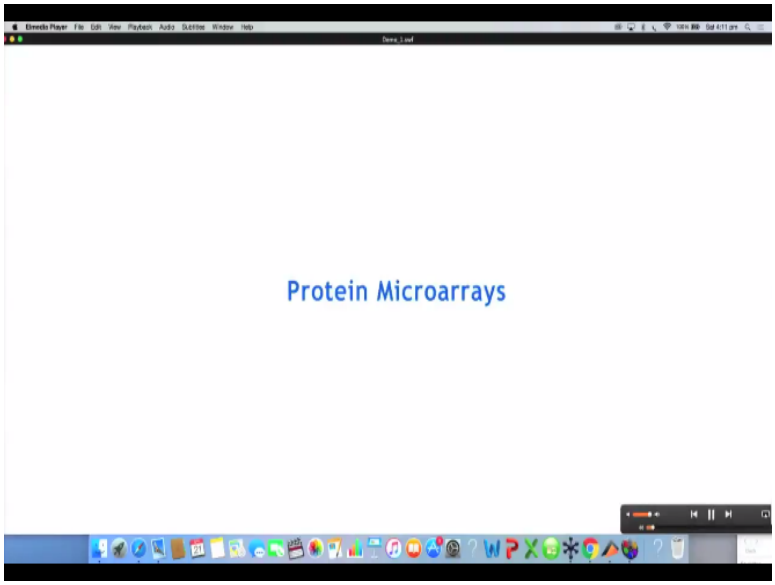


And that first time established that you know the protein microarrays in principle are achievable.

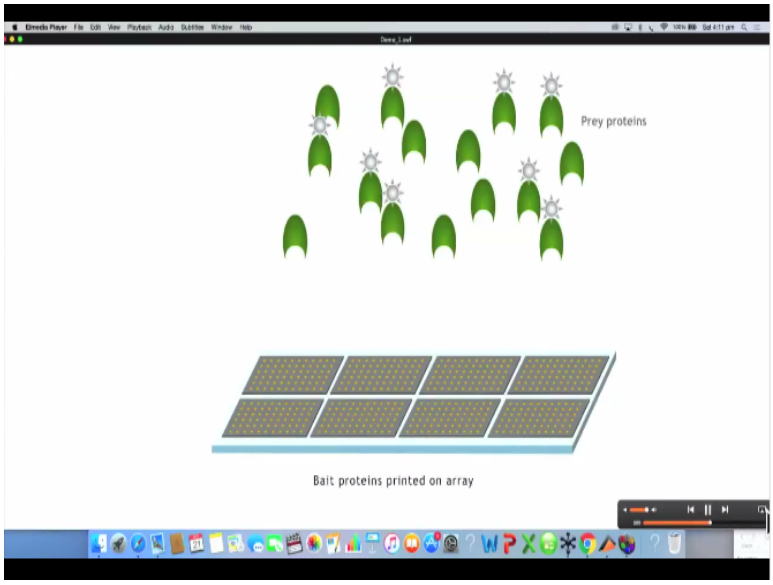
(Refer Slide Time: 05:10)

Let us now review some of these concepts in an animation.

(Refer Slide Time: 05:14)

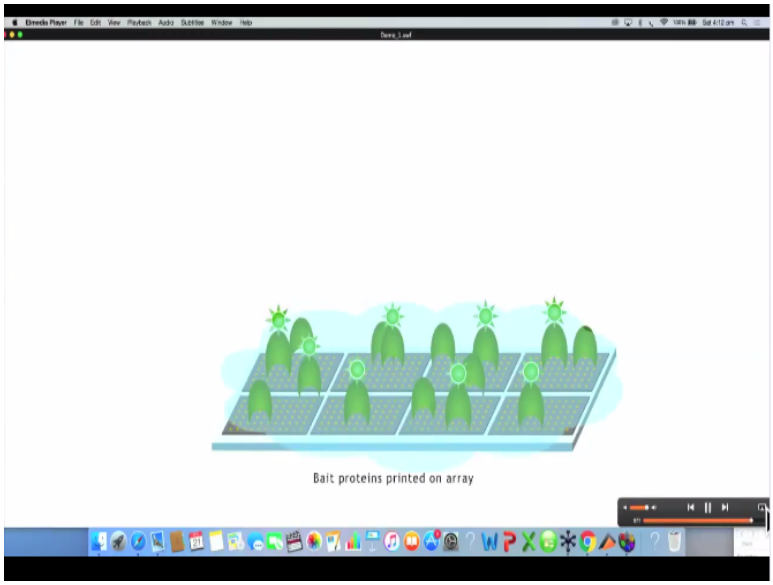


(Refer Slide Time: 05:15)



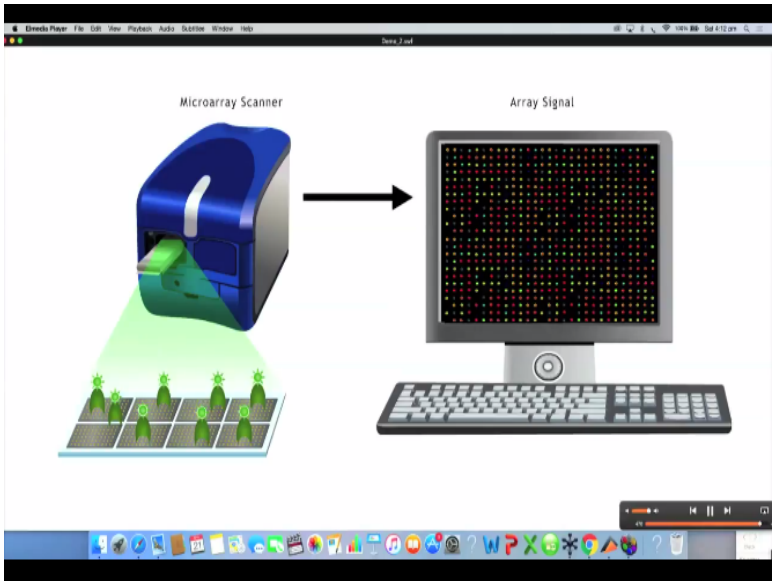
Protein microarrays are widely used for protein interaction studies. One of the proteins to be analyzed is printed onto a microarray surface usually made of glass. The proteins known as bait proteins get immobilized onto the array surface that is functionalized with reagents like nickel or aldehyde compounds that interact with groups present in the protein. This bait protein is then probed for interactions with suitably labeled query or prey proteins.

(Refer Slide Time: 06:00)



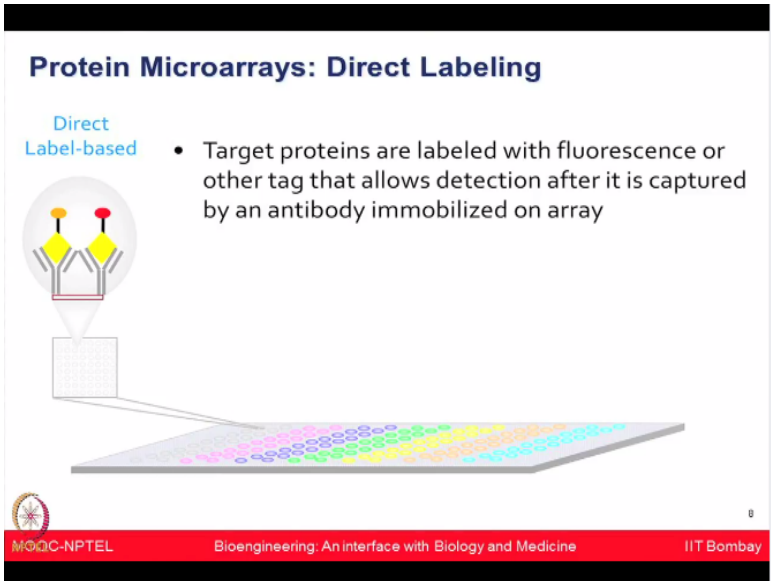
Any unbound proteins are washed off the array surface.

(Refer Slide Time: 06:19)



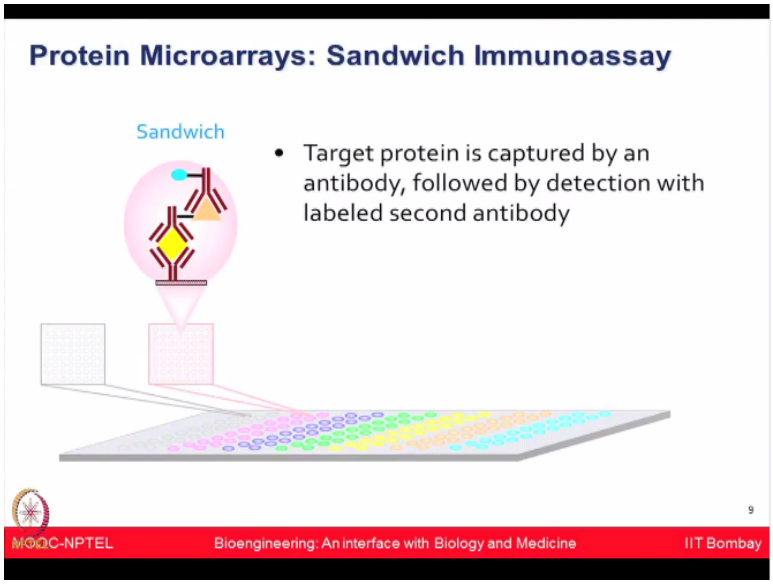
Once the unbound proteins are washed off the array surface, the protein interactions are detected by means of an array scanner. These protein microarrays are extremely useful in studying interactions with other proteins as well as small molecules DNA or RNA. Let us now talk about a different type of protein microarrays platform. For example, the direct labeling based method.

(Refer Slide Time: 06:53)



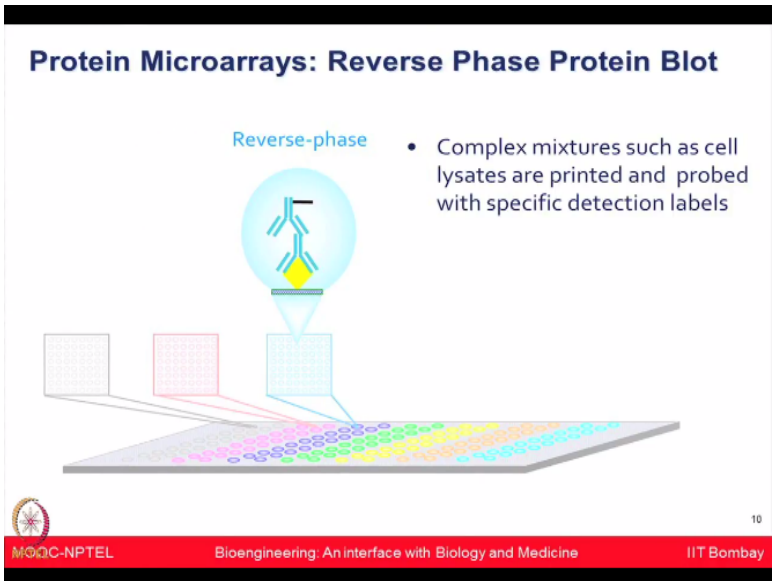
In direct labeling, the target proteins are labeled with fluorescence or some other type of tag molecules which allows detection after it is captured on the antibody which is immobilized on the array surface as shown in the slide on the image.

(Refer Slide Time: 07:07)



Additional you can also use the sandwich immunoassay where the target protein is captured by an antibody and it followed by the detection with labeled secondary antibodies. So that is a sandwich immunoassay.

(Refer Slide Time: 07:21)

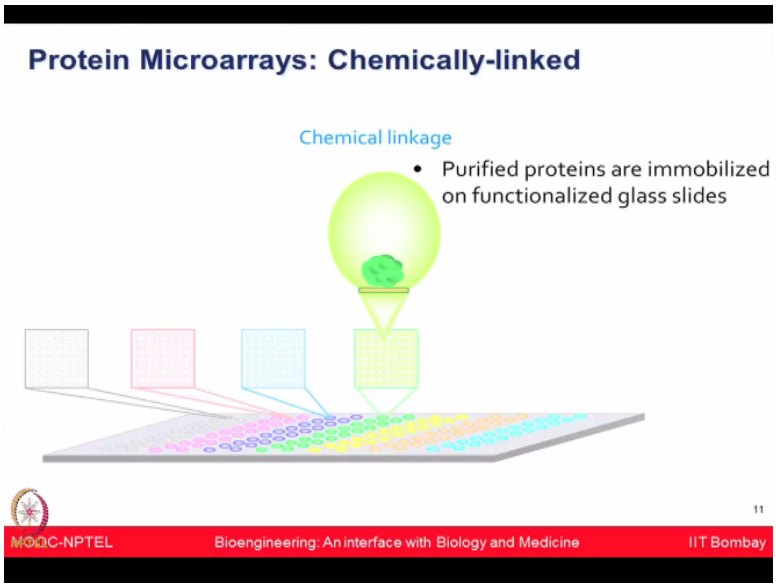


Another approach which is clinically quite relevant is reverse phase protein blots in which the complex mixtures such as cell lysates are printed and probed with the specific detection labels. So many times clinicians have these complex tissue lysate or cell lysate and they want to just detect one of the specific antibody and specific proteins out of those samples and then reverse phase arrays could be very quickly employed for that purpose.

Protein microarrays ideally can be done with the purified proteins, although obtaining large number of purified proteins is the challenge and therefore making protein arrays from the

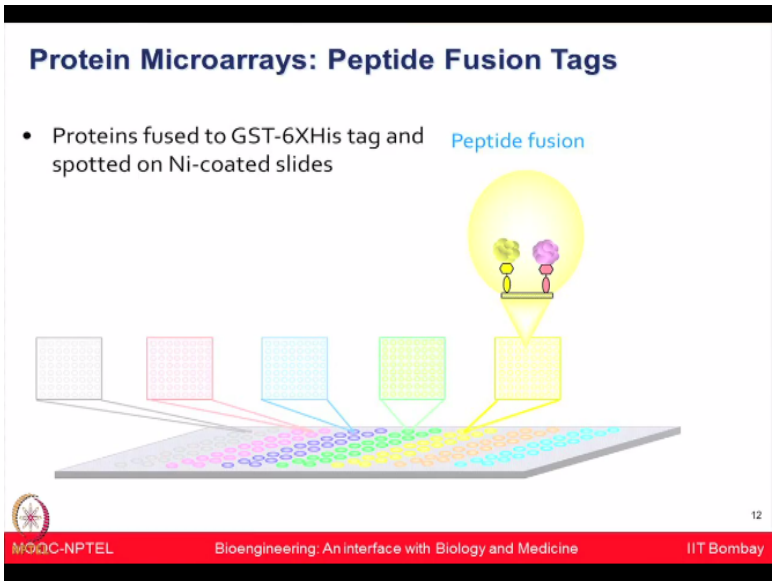
purified proteins has not been you know so rapid not been used by many labs and many companies are not around to print the entire protein arrays but still if ideally you have access to the protein contents then using the chemical linkage of immobilizing the purified protein on the functional slide can be very powerful.

(Refer Slide Time: 08:16)



Additionally, if you have identified some peptide sequences let say you know we have done the mass spectrometry experiments from the previous lectures and now you have identified some peptides of you know your interest which you want to now further confirm and validate.

(Refer Slide Time: 08:35)

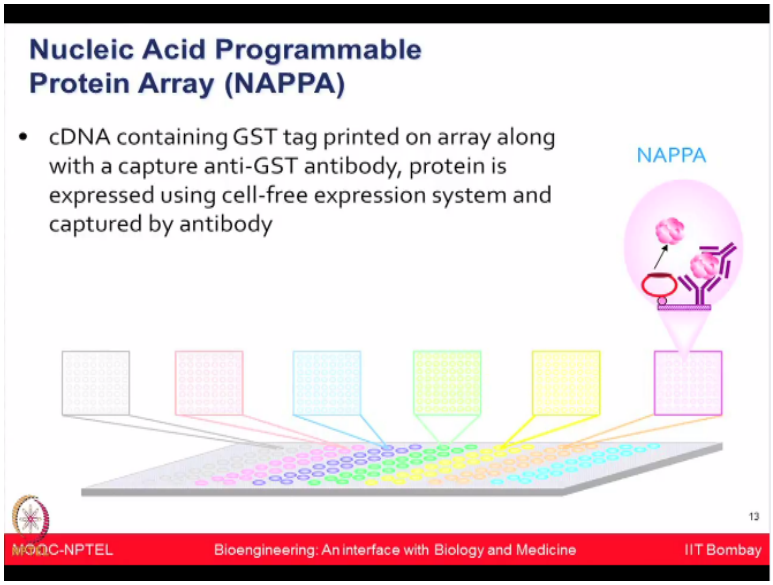


So then those could be synthesized artificially, one could make the peptide fusions and then print those on the array surface and then now screen those in the array format. In addition to

looking at the contents directly from the protein, scientists have also thought about can we use in-vitro transcription and translation based methods where proteins could be synthesized on the chip directly without having need to purify them.

So the whatever machinery in the body makes the protein which involves the process for transcription and translation, can we provide those material on the chip itself and from those contents now the proteins can be synthesized on the chip. So people have taken let say the cDNA containing as the gene of interest from which you want to make protein.

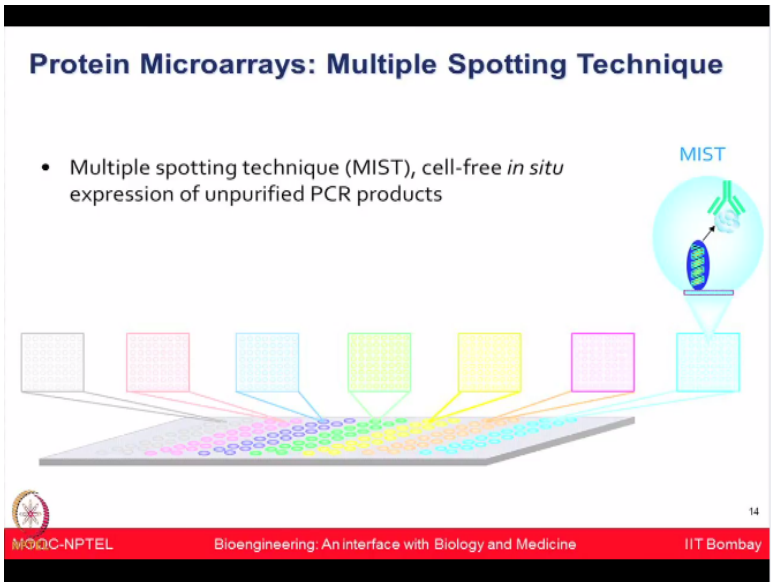
(Refer Slide Time: 09:24)



And now an approach which is known as NAPPA nucleic acid programmable protein array has been used where a captured antibody is immobilized along with your DNA or the gene of interest. Now let say if your gene of interest contains a tag GST, so the captured antibody will be anti-GST capture antibody and then to immobilize them you have to use some other molecules like you have to use a cross linker like BS3, to help the strong binding you can use a BSA.

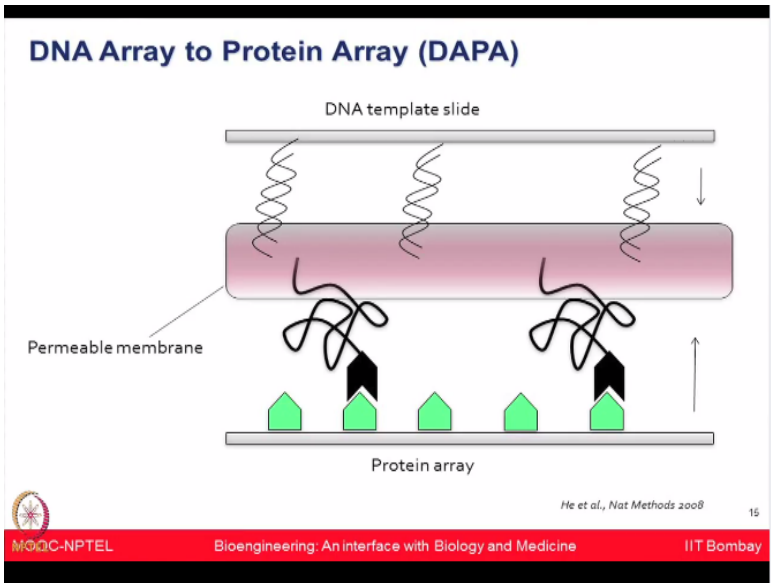
So those are part of the chemistry of how to make the NAPPA arrays but conceptually you can start from the gene and make the protein directly on the chip by using the cell-free in-vitro transcription and translation with it.

(Refer Slide Time: 10:13)



That is one of the very powerful method and couple of technologies have also come forward which have used a similar type of concept like a multiple spotting technique or MIST technology where they have used PCR products and purified PCR products, added cell-free lysates and then try to make the chip which is protein arrays eventually.

(Refer Slide Time: 10:30)

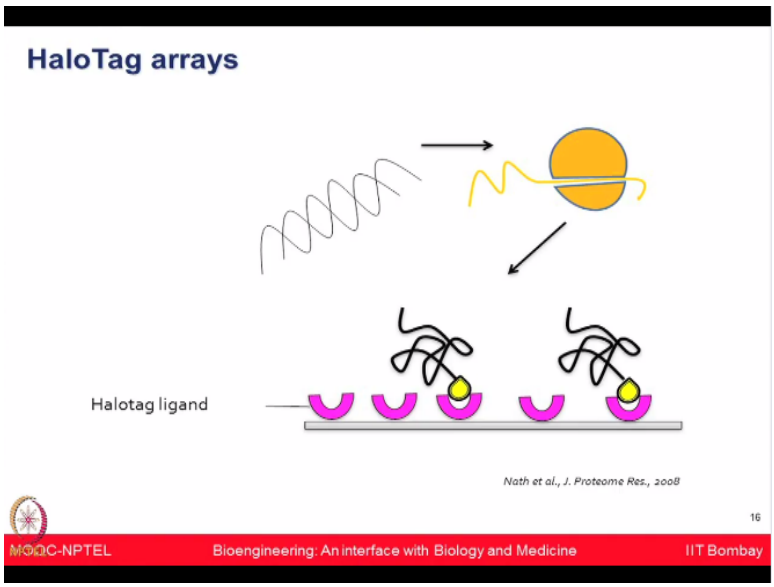


Another promising approach in the field of making the protein microarrays has been DNA arrays to protein arrays or DAPA where you start with the DNA printed on the glass slide and imagine that your DNA has the you know the histidine tag. Now once you add the in-vitro transcription and translation to make the protein on the chip, then these proteins which are you know you can add one permeable membrane.

The protein which are synthesized they will pass through permeable membrane and on another glass slide you have added you know let say Nickel NTA on the glass slide. So all the proteins which are you know coming out of this particular DNA slide which are having the histidine tag could be now immobilized on to second slide which is having the your Nickel NTA coated on the slide.

So as shown in the image that you know now you can have the pure proteins coming out from the membrane and they could be immobilized on the new chip and from the same glass slide actually you can from the same DNA slide you can actually make couple of slides for the protein array. So you can increase your efficiency of generating these you know the protein arrays from the same DNA template slide.

(Refer Slide Time: 11:43)

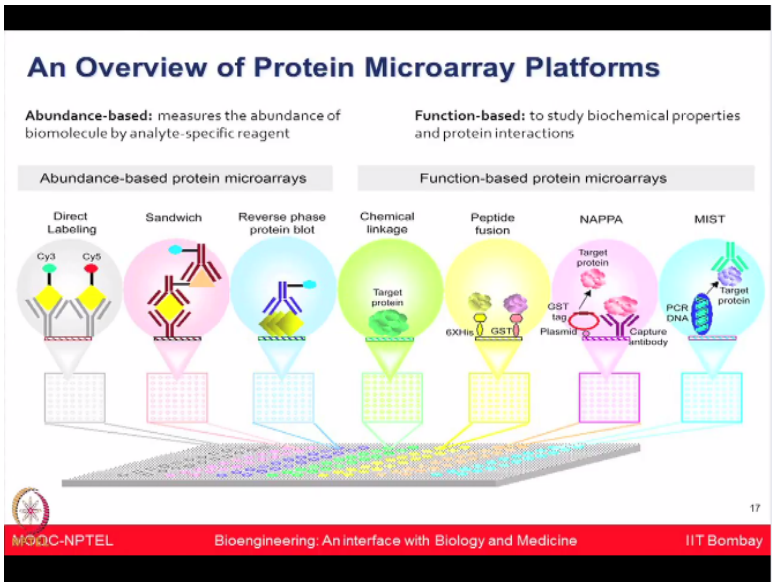


A more commercial approach has come forward recently which is HaloTag arrays where on the glass slide these HaloTag ligands the chloroalkanes are printed and then your DNA of gene of interest contains this particular tag which is HaloTags and now once the proteins are synthesized after doing cell-free expression based approach then these HaloTags are going to bind to the ligands on the chip and you can have the very strong covalent interaction. So these kind of you know cell-free expression approaches are very relevant.

Because when you are doing the lot of microarray based studies you have to do many you know washing steps and you have to do many processing steps, those could be easily avoided if you can use these HaloTag arrays. Although, you know the limitations are that you know

you have to clone these, you have to add the HaloTag in the gene of interest. You have to use this kind of chip platforms which will having the ligands coated on the slide.

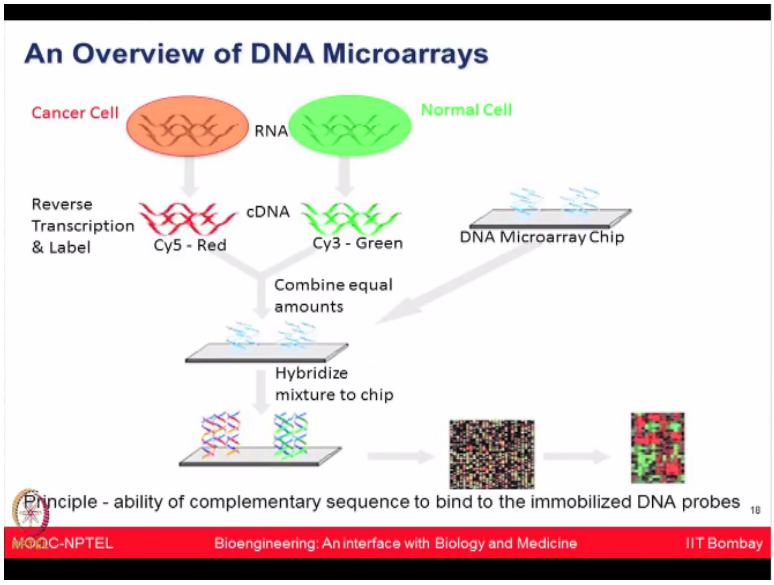
(Refer Slide Time: 12:44)



So briefly I have shown you kind of you know the glimpse of an overview of protein microarray platforms so far which could be broadly you know classified into abundance-based as well as the function-based microarrays. In the abundance-based microarray, we talked about the antibodies which could be linked with the different type of antibody-based platform like direct labeling, sandwich immunoassays and reverse phase protein arrays.

In the function-based arrays, we talked about purified protein or the peptide arrays or looking into the NAPPA arrays or the MIST approach. All of these things are various possibility for doing protein microarray based platforms. I must would like to remind you, you know based on our previous lecture discussion that the DNA microarray has been very powerful and which has been used from long time probably from 1990s you know it was very actively used and the DNA microarrays conceptually involves.

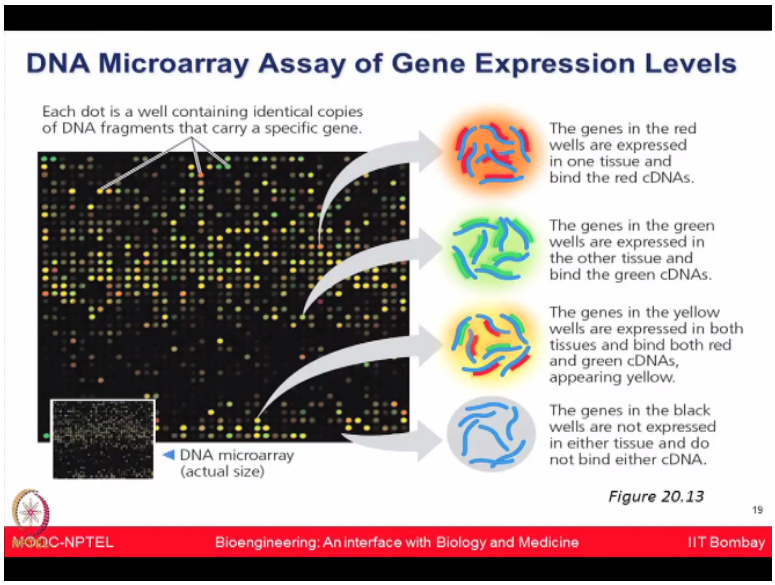
(Refer Slide Time: 13:27)



Let say we have the cancer cell and the normal cell from which the RNAs has been extracted, then you are converting RNA to the cDNA forms which we have learnt in the molecular biology modules and now each of that cDNA from the control and the cancer cell could be labeled with Cy3 and Cy5 tags as you can see the schematic here. They are combined equally now and then they are allowed to hybridize on the chip surface.

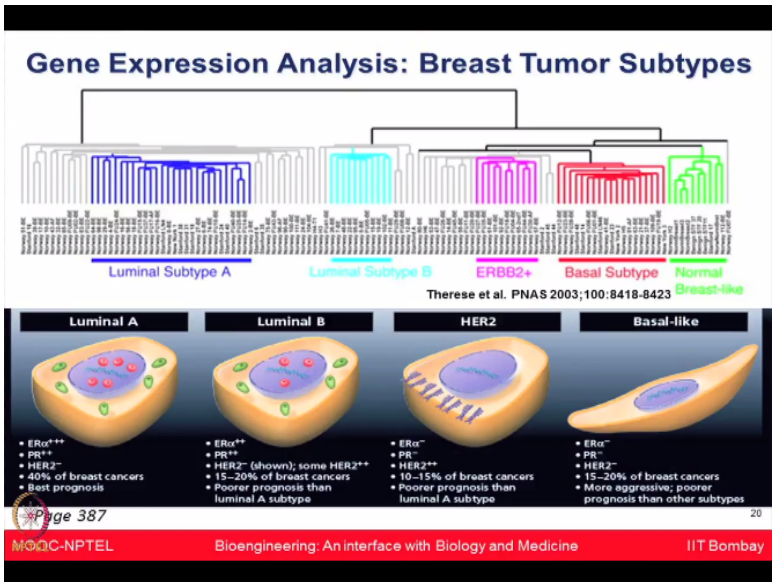
And wherever they you know hybridize looking at the you know ability of the complementary sequences to bind to the immobilized DNA probes, now you can detect those signals on the DNA microarrays which is conceptually different than the protein microarray but I am just trying to give you the parallelism that how DNA microarrays have been used for different applications shown in the slide here again to remind you.

(Refer Slide Time: 14:21)



That you know different dots, different colors can be seen which are red in the color, green in the color, yellow in the color or the gray color. So looking at that if you have labeled with Cy3 or Cy5 one condition like control or your cancer cells. Then, accordingly now you can decide that you know what is the over expression or the down regulation of the given genes in a given system.

(Refer Slide Time: 14:45)

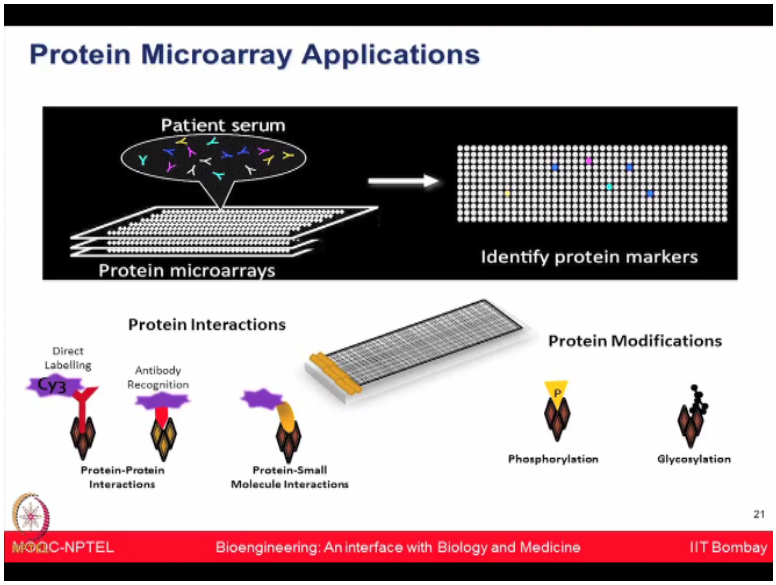


And a similar kind of concepts have been very powerful clinically in fact where clinicians and the you know molecular biology scientist work together and looked into the breast cancer one of the major problem where all the women suffering from breast tumors do not you know respond to the same kind of drugs. So when people had looked into their gene expression signature using this kind of protein microarray another expression analysis.

Then they found out that you know all the women who are having the breast cancer they could be broadly under 4 subtypes. So there are not only one disease alone, there are 4 major subtypes could be you know luminal subtype A, it could be luminal subtype B, it could be ERBB2, it can be basal subtype and of course you are comparing with the normal breast-like appearance.

So based on these you know whether it belongs to the luminal A, luminal B, HER2 or the basal-like then the treatment modalities could be defined. This just shows you illustrates you the practical usage of the technologies how it can be so helpful for the clinicians to make the decision that what kind of drugs to be given. Coming back to the protein microarrays.

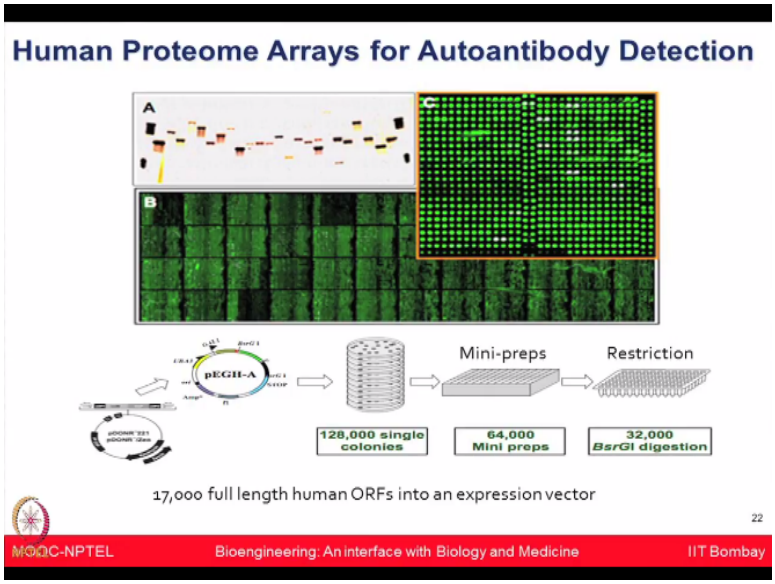
(Refer Slide Time: 15:55)



So protein microarrays have been used for many applications especially looking at identifying the protein biomarkers, doing the protein-protein interaction studies, protein modifications like phosphorylation, glycosylation, acetylation, etc. All of these have been studied using protein microarrays. One of the other you know the strong application of this is looking for autoantibodies.

Many times for you know different autoimmune disorders and cancer, you know body will start producing some antibodies to try to combat and fight the disease.

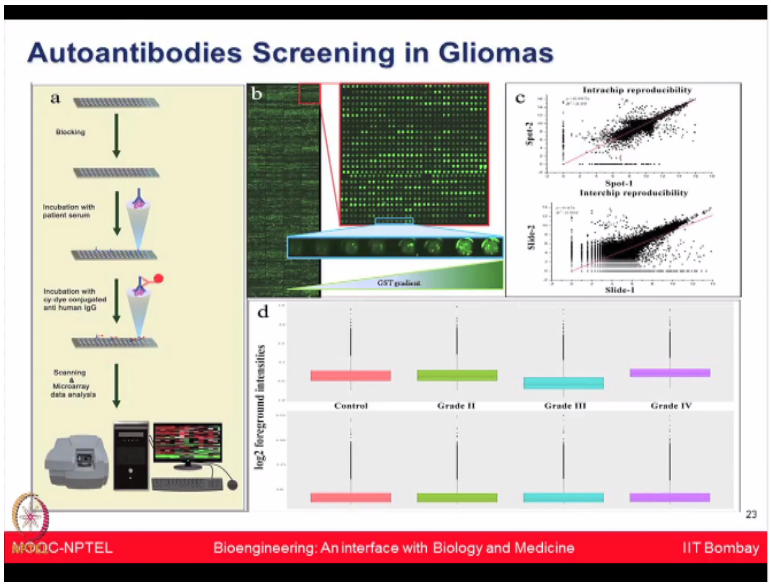
(Refer Slide Time: 16:16)



And those could be detected on the protein arrays and you know one such array platform you know we have used for our own research in the brain tumor where we have used the human

proteome arrays which contains all the you know human proteins possible till date which is 19,000 proteins.

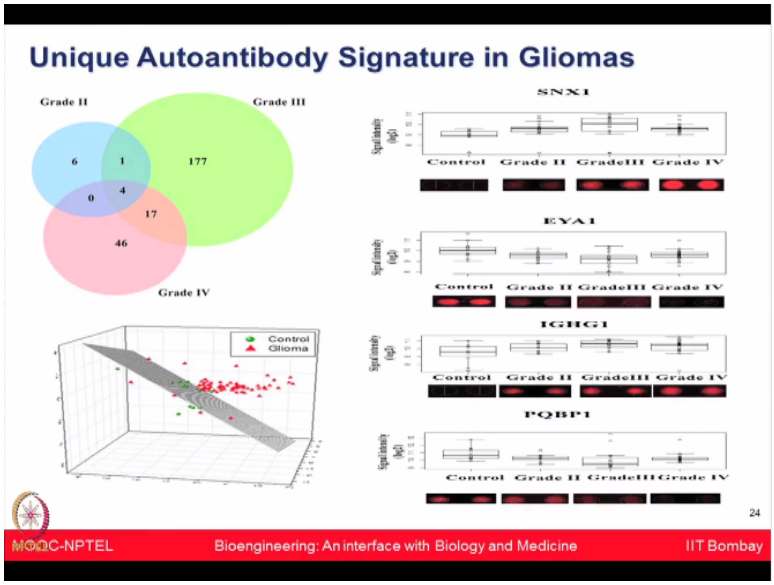
(Refer Slide Time: 16:29)



And after you know we took the serum samples from different patients suffering from the grade II, grade III, and grade IV of the brain tumors gliomas and then you know along with the healthy individuals comparison with the control after looking at a signal then we could identify you know some unique biomarkers which are appearing in the grade II patients or grade III patients or grade IV patients and as shown here.

Couple of images which shows that you know the some of the biomarkers are showing the sequential increase of the you know over expression of the proteins and probably those could be used for the early indicator of the disease.

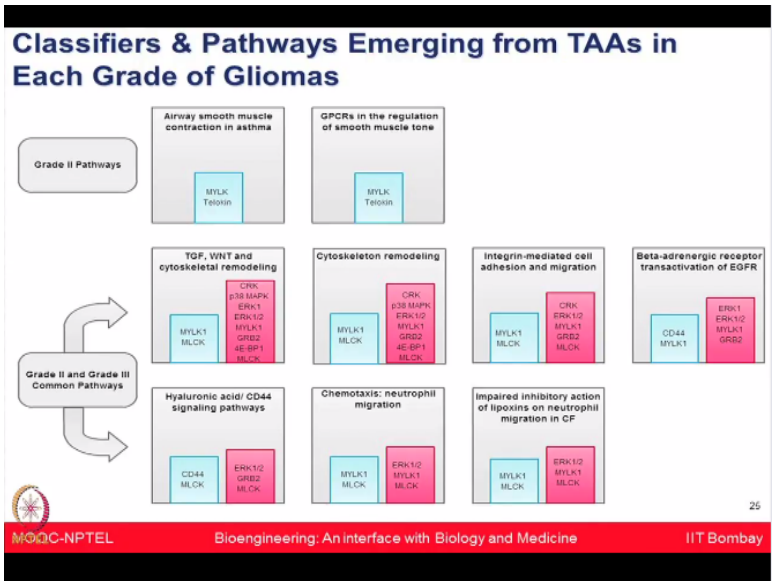
(Refer Slide Time: 17:12)



And some so like as next one as shown you know show you increased continuous so as IGHG1 which also shows you from the you know control to the higher grade continuous increase. Some proteins like IA1 and PQBP1 shows the reverse response like it is higher in the control and it goes down regulated as you know over the period in the higher grades. This kind of you know the analysis also gives us a much better understanding of what is happening physiologically in these patients.

For example, when the disease progresses from the low grade to the higher grade from the grade II, III and IV.

(Refer Slide Time: 17:45)

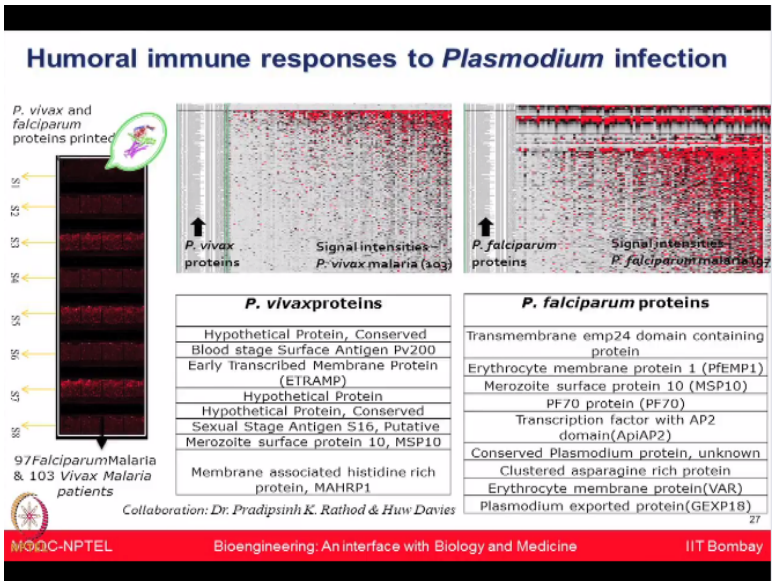


Then, what kind of you know the genes and what kind of proteins are being expressed, are there some new unique proteins appearing as the disease progresses from the low to the high

grade and what kind of major pathways are you know being changed. So those kind of studies even basic molecular biologists can do and obtain by using these kind of high-throughput screening.

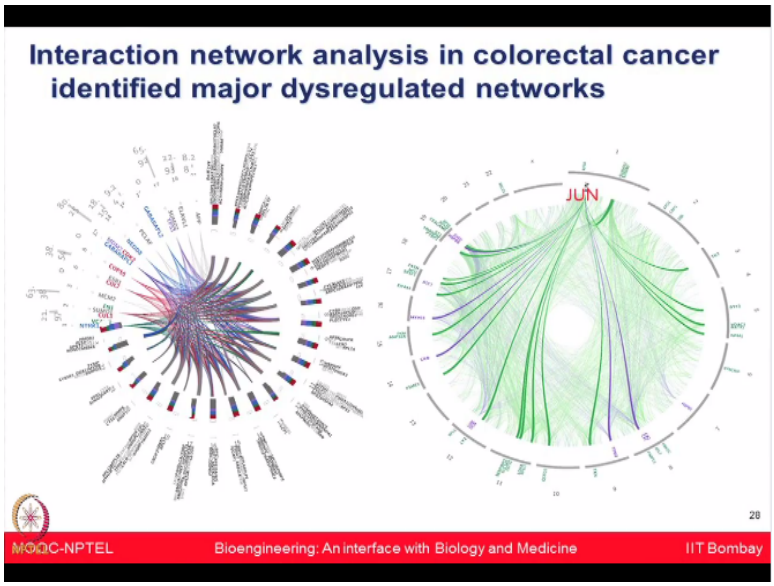
We have also used similar kind of platform for looking at the infectious disease responses.

(Refer Slide Time: 18:06)



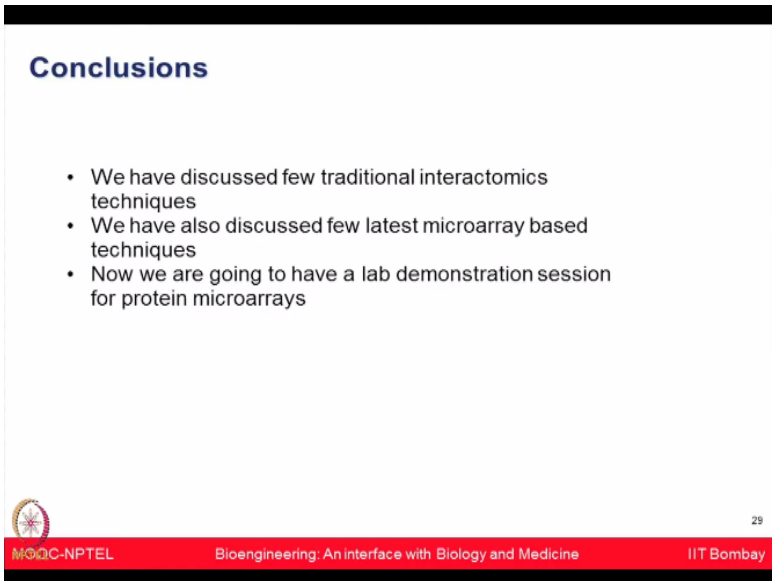
And in this case you know it is shown for the malaria proteins when we have both falciparum and vivax, antigens printed on the chip and we are looking at the humoral immune response to the patient suffering from falciparum or vivax malaria. I just want to illustrate you that these kind of platforms can be used for many projects, many different diseases, many biological questions can be addressed.

(Refer Slide Time: 18:27)



We have also looked at the colorectal cancer and identify certain kind of you know network and hubs looking at the network analysis obtained from our mass spec data and some of the microarray data.

(Refer Slide Time: 18:37)



So in conclusion we have today discussed you know broadly the interactomics field, some of the traditional approaches being used especially immunoprecipitation, looking at the yeast two-hybrid and then kind of I moved to describe you about different type of protein microarray platforms which are used for different applications.

And this protein microarray platform could be from the you know antibodies to the purified proteins to the DNA content itself by using the cell-free expression based microarrays which could be used for many applications whether it is cancer or infectious disease or to understand different type of you know basic biological questions.

So now I am going to stop here and we will take you now for the lab demonstration session where we are going to show you how to do this kind of microarray experiment especially the protein microarray-based experiment directly inside the laboratory. Thank you.

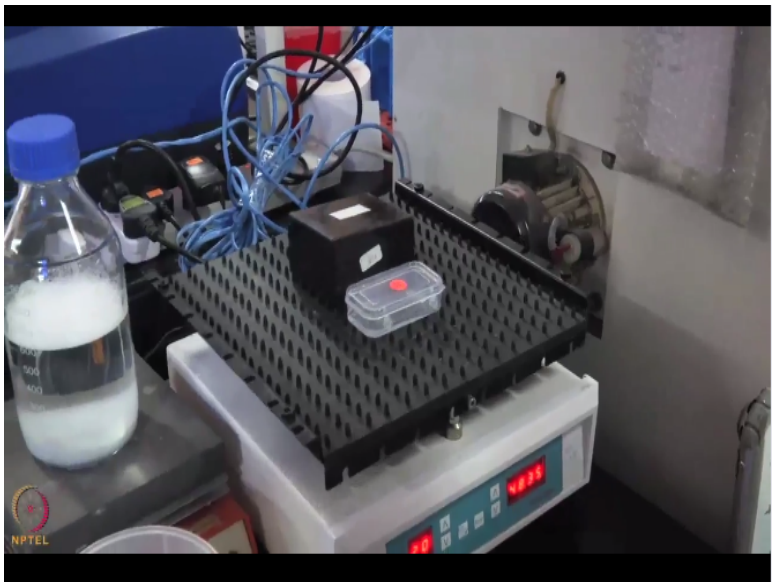
(Refer Slide Time: 19:27)

So I will be showing how to do autoantibody screening using HuProt arrays.

(Refer Slide Time: 19:33)

So after adding blocking solution you have to keep it for two-hour incubation in room temperature with gentle shaking, so we are keeping in the rocker.

(Refer Slide Time: 19:45)



So after 2-hour incubation, we have to wash the solution. So for that we are using 1X TBST of pH 7.4. So first you have to discard the blocking solution. Then, you add 1X TBST. Then, you keep at rocker. This step has to be repeated for 4 times for 5 minutes. So after the last wash, you have to discard the washing solution completely.

(Refer Slide Time: 20:40)

Then, you have to incubate the primary antibody which is a mixture of 1 in 500 dilution of serum and 1 in 5000 dilution of anti-GST. So now you are adding the 10 ml volume of 1X TBST containing 1 in 500 dilution of your serum and anti-GST antibody. Then, you keep it for 2 hour incubation. So after incubating with primary antibody, we again repeated the washing step.

Now the next step is incubation with secondary antibody. So now we are discarding the washing solution.

(Refer Slide Time: 21:27)

And this step is very cautious because the secondary antibodies are labeled with the fluorescent dye. So you have to do it in the dark and or cover it with the foil. So this is secondary antibody, it is a mixture of 1 in 1000 dilution of anti-human IgG which is labeled with Cy5 and 1 in 5000 dilution of anti-rabbit antibody which is Cy3 labeled.

So once you added the solution that is the antibody solution you have to cover the box with a foil or any other dark I mean or you can use a dark box also. Now we will keep it for two-hour incubation. So I am keeping the secondary antibody for two-hour incubation. After secondary antibody incubation we again washed the slides the same way. Then, after washing we will dry the slides using centrifuge and make sure that all the steps has to be performed in dark because the slide is already incubated with Cy3 and Cy5 fluorescent labeled dyes.

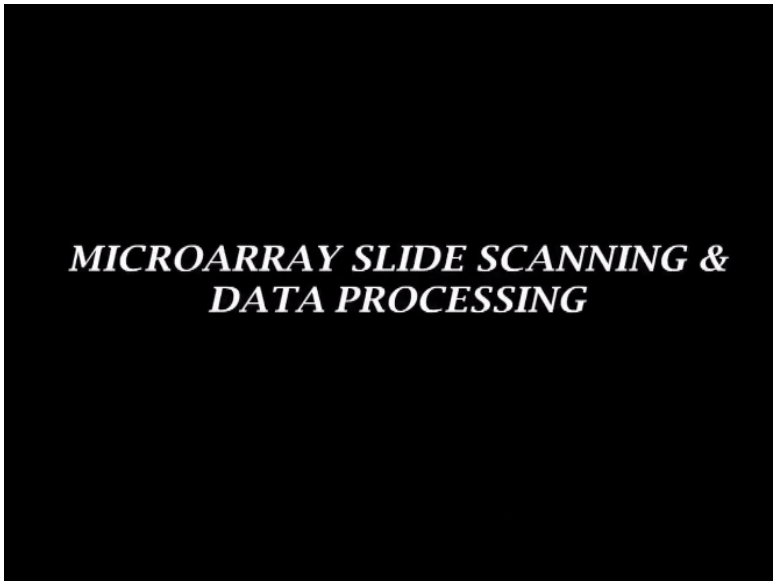
You just have to open the centrifuge, take the slide, keep it in the cassette which is designed for the microarray experiment.

(Refer Slide Time: 23:16)

So once you keep the slide, close the lid and here the parameters are speed is 900 Rpm, time is 2 minutes, acceleration and deceleration is 8 and 8 and the temperature is room temperature. So once you keep the slides then spin it. So just you have to start the spinning. So after centrifugation now we are going to take out the slides. So you can see now the slide is completely dry.

So after complete drying now you have to scan the slides in two different wavelengths. One is 635 which is for the red channel and the 532 for green channel. So the 635 will the red channel will show them anti-human IgG that is the autoantibodies in patient serum and Cy3 that is the green channel which is a QC check. It will show all the printed protein spot.

(Refer Slide Time: 24:26)

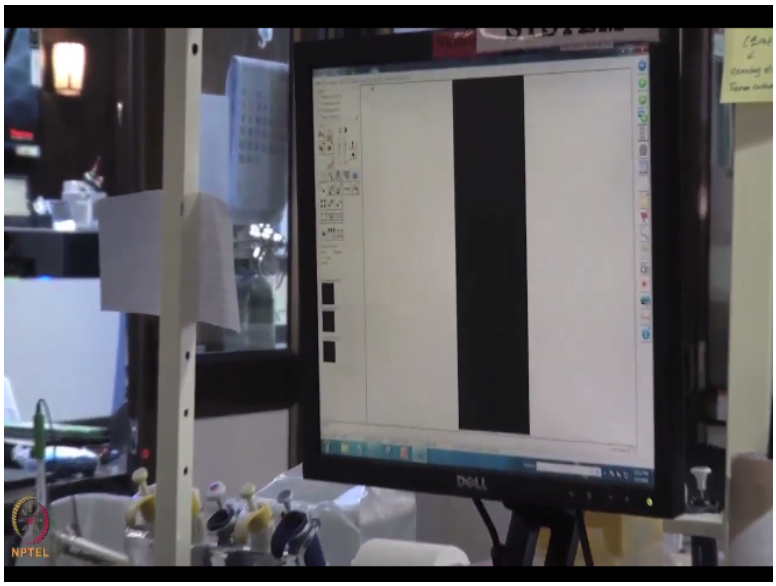


So once you have the dry slide, what you have to do is you have to take the slide and put it for scanning. This is the scanner; it is GenePix 4000B. It scans at two wavelengths. So slides are always kept facing downwards that is the proteins that are printed should always face downwards.

(Refer Slide Time: 24:53)

So this is how we keep the slide inside the cassette. Make sure the slide is aligned properly and is fixed properly otherwise it might get broken inside the scanner. So once we keep the slides properly, you have to close the door and then just close it. Make sure that the scanner is kept on for 20 minutes to stabilize all the lasers. So once we keep the slide inside the scanner, we open GenePix Pro software. It will search for the scanner, it will calibrate it and then the slides are ready for scanning.

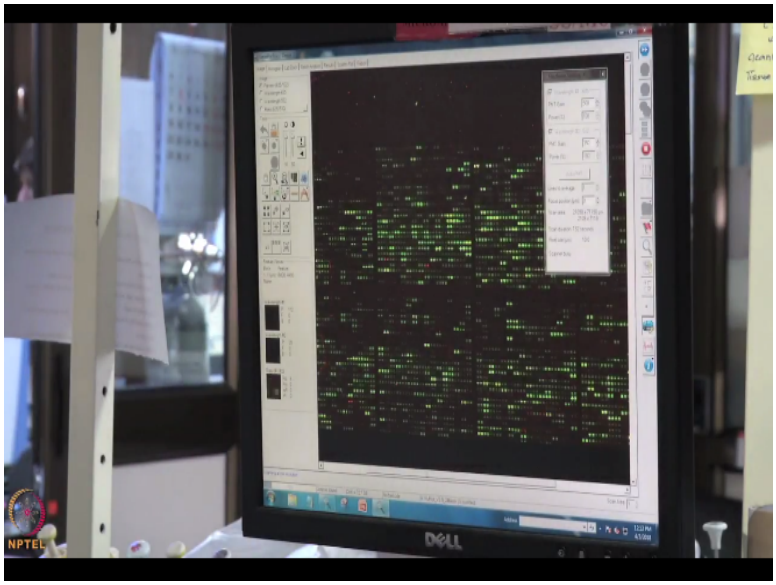
(Refer Slide Time: 25:27)



So what you do is you set the wavelength because it is a dual-channel thing, you set it for red channel, we set it at 635 nanometers. For green channel, we set the wavelength at 532 nanometers and then we go for the PMT settings. So usually if we keep PMT settings at red channel at 500 PMT gain 100% laser power and PMT gain at 350. This settings you can optimize using auto PMT.

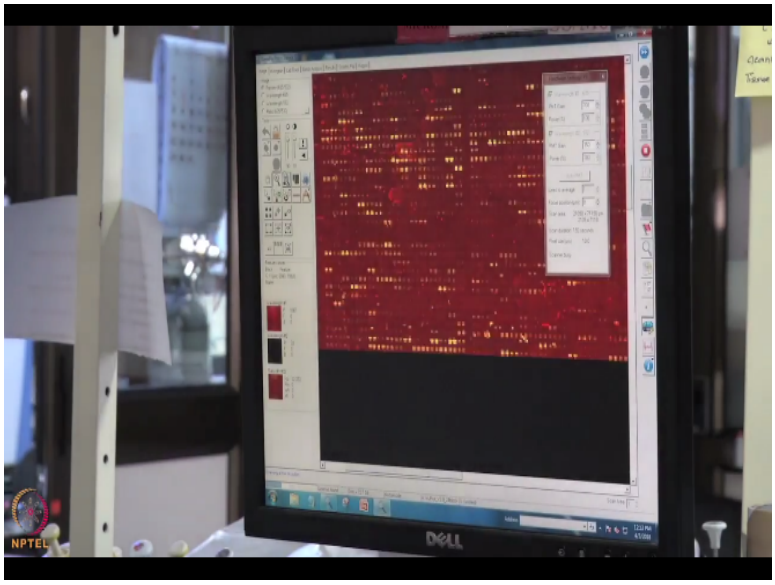
I will show you how this PMT settings work. So let us let us start the preview scan for the same. The scanning is more or less like you scan your documents in a scanner. So here what you do is you scan the area for the slide and then you can just look here if you zoom into it.

(Refer Slide Time: 27:17)



You can see how the panel looks like. Now when you change the PMT settings, say I put it at 700, you can see that the power for this channel is increasing and you are getting more of the brighter spots and more of the saturated spots.

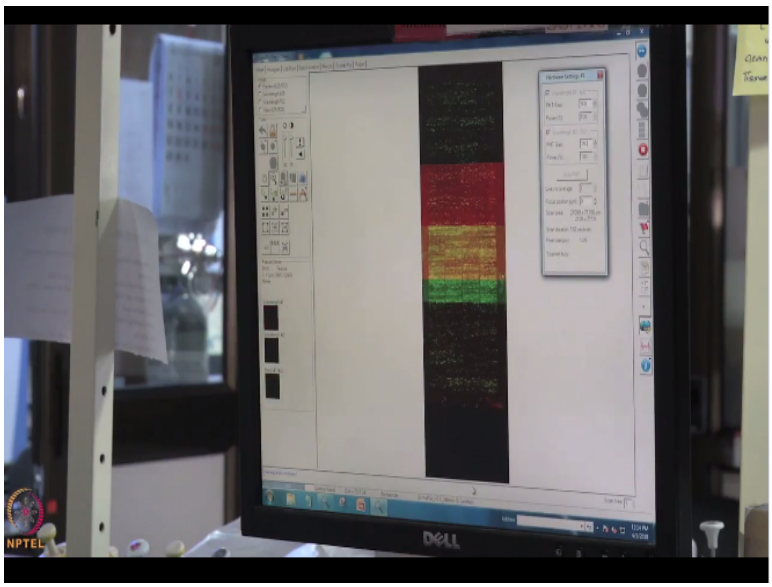
(Refer Slide Time: 26:33)



Now if I increase this also to say 600, then you can see that the whole intensity is increasing. So PMT's parameters are very important because when you have to scan a slide, all the spots should be scanned at a particular wavelength and you do not want to miss out on a lot of spots as well as you do not want most of the spots to be saturated. So auto PMT is one of the options where you can go and it will just scan the slide at 3 different wavelengths.

And it will try to fix at 5% saturated spots and you can get a PMT at which you should scan the slide at both the channels.

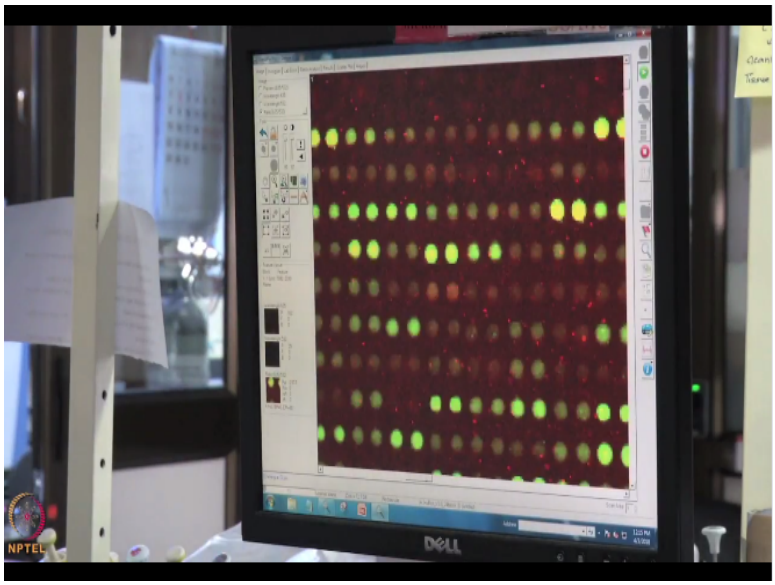
(Refer Slide Time: 27:19)



So this is how the preview scan looks like. So once you have done it, you can just lay over the grid like the way you put you set an area for your paper scanning, you can set the area for scanning of the slide, so if you want only this much of area to be scanned, you can just resize

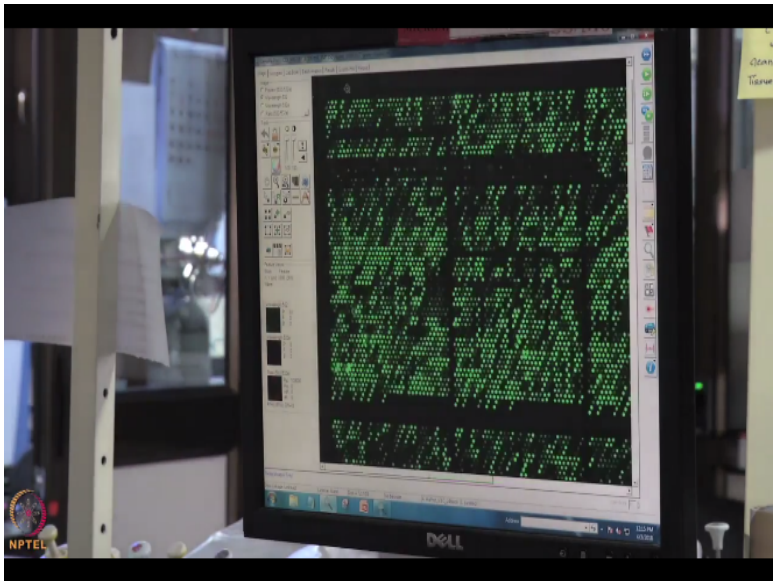
the frame and then you can start the scanning of the slides. So this you can see that the scanning is a comparatively slow process and if you zoom into it you can see that the spots are like quite proper now.

(Refer Slide Time: 28:22)



So this is how the scanning is performed. Once you have the scan slide you can just stop and save this filet the scanned image. I will just show you how to further process the image. So once you have this you just to go to GenePix Pro analysis software. You open that, you go to the saved images. You can open images and select your image that you want to see. I will just zoom into one of the blocks here.

(Refer Slide Time: 29:00)

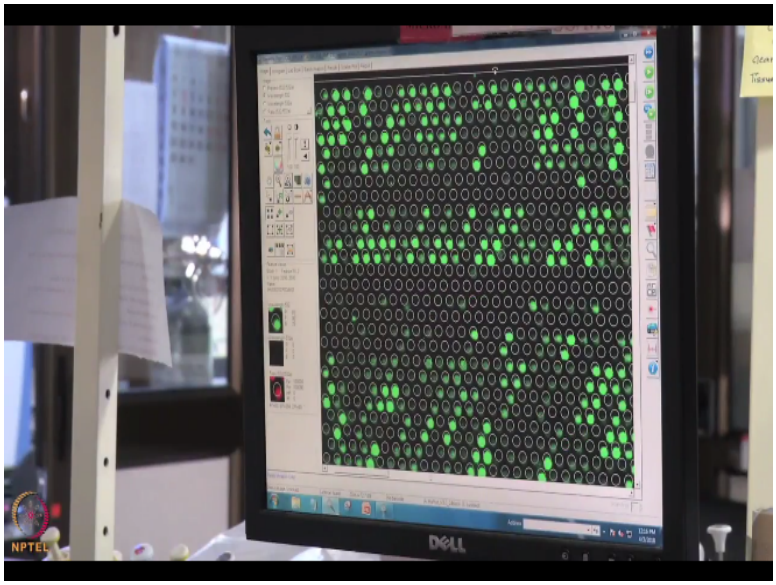


This is how in green channel the slide looks like and green channel is majorly done for checking the quality of the slide, for checking the printing quality of the slide. So now if you

look here, it is these spots have no identity here but when you put a GAL file over it, GAL file is the file which is given with each and every slide vendor gives you a GAL file which tells you the location of these proteins on a slide.

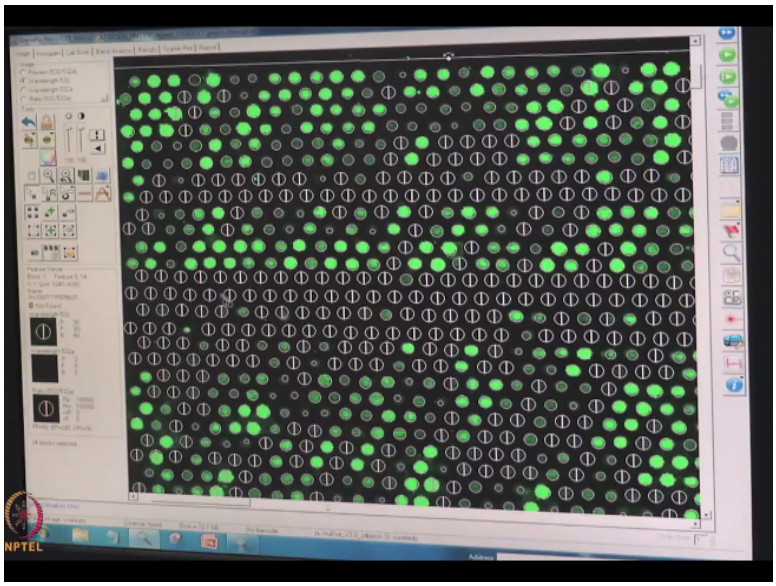
So you can just adjust all the blocks together and you can try to just fix it onto the slide. I will further zoom it here.

(Refer Slide Time: 29:49)



Now if you look each and every spot will have a name, a ID because this is the identity of the protein. So now you can just try to rephrase it, try to align it properly.

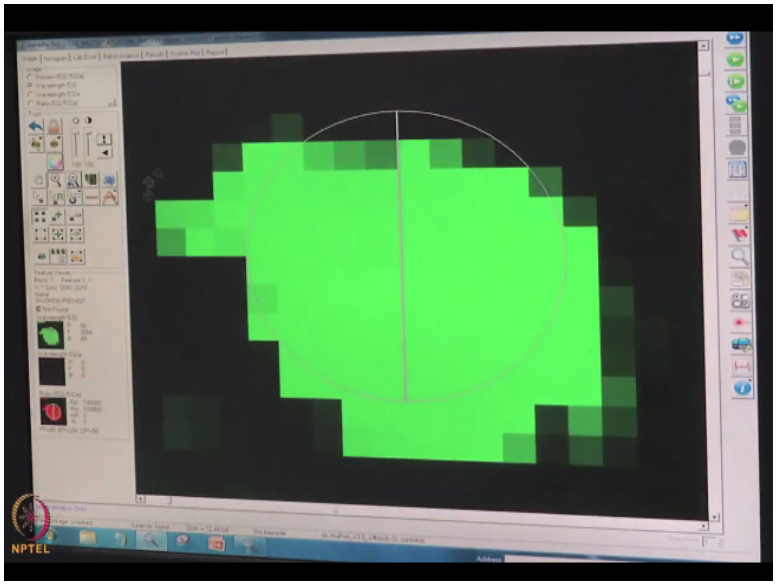
(Refer Slide Time: 30:23)



So now you see that this slide is properly aligned but some of the features because so here are the settings which are given to the slide here. Yeah here in the setting you can define the

minimum diameter which should consider for a spot and the maximum diameter which should consider, so that I do not pick any background as a spot. So and we also tell what should be your comparative background intensity above which only a threshold above which only you should consider a spot. So if you zoom in here.

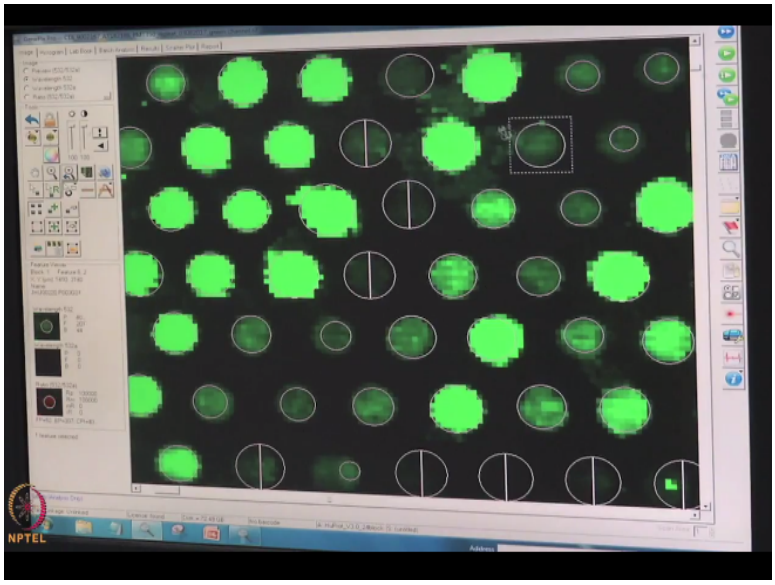
(Refer Slide Time: 31:11)



Some spots say this spot is not considered as a spot because of the background intensity is way more as compared to the actual spot but here you see these are the spots in duplicate and here you can see that this spot can be is a spot and it is not an artifact. So you just have to press F and try to resize it and take it as a spot. Again here if you see that these two spots are the diagonal duplicate spots.

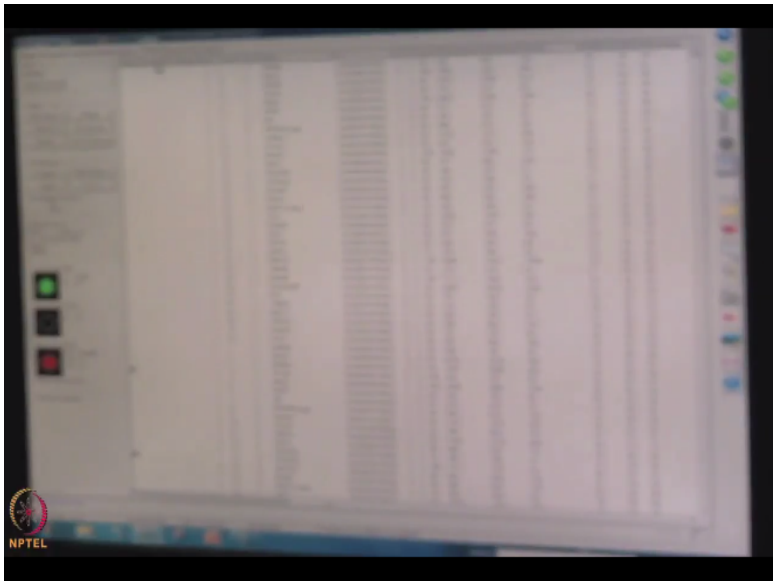
So this is also a spot but because of this background it is not considered as a spot, so now you will again resize the spot and take the exact intensity.

(Refer Slide Time: 31:54)



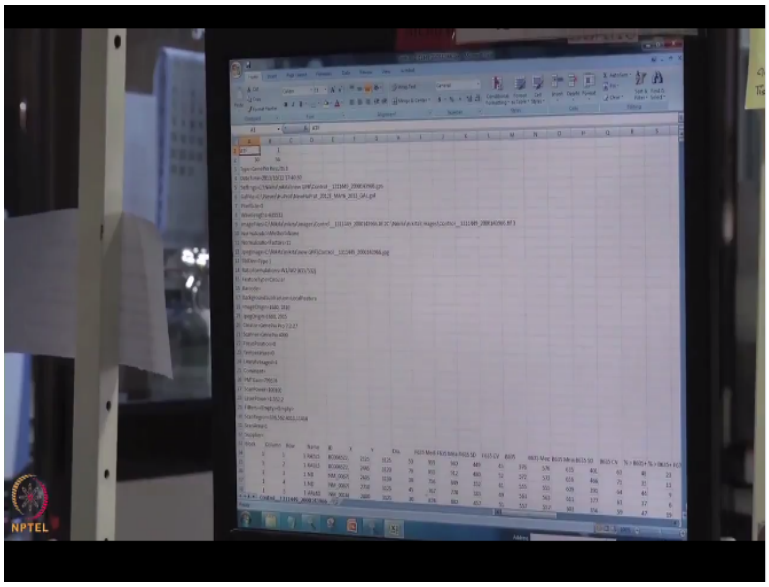
And the encircled area is the intensity which it tries to average out and give you a spot intensity, so that is why it should be really compact spot consisting only of your protein of interest. So once that is done, once you lay the grid for all the blocks in the slide, you will just go to data and then it will analyze the intensity at green channel as well as at red channel and it will give you a dot GPR file which is called as dot GenePix Pro the GenePix result files.

(Refer Slide Time: 32:32)



And this is how all the parameters are given like diameter of the, at what location it is present, the diameter of the spot, what is the foreground value, what is the background value and then you can just export the result. You can select it all by control A and then you can save the results as dot GPR file. So once you saved file as dot GPR file, the files can be opened in excel and it gives you all the information of the PMT gain, scan power and the date on which it has been scanned and what all GAL file features have been used for scanning.

(Refer Slide Time: 33:11)



It also gives you a value of foreground and background and this is how for all the protein it gives you the ID, name and all the diameter and the intensity at which it has acquired and then you can of all the intensity which it has acquired and then you can just select these files, you can use foreground-background value. Foreground median-background median value for the statistical analysis.

So you take all the details from each and every slide and then you do statistical analysis for getting the autoantibodies if there are any in some of the patients and to look at the differential expression of autoantibodies amongst different groups of patients.

(Refer Slide Time: 33:57)