Lecture – 1

Introduction to Interactomics and Protein Arrays

Applications of Interactomics using

Genomics and Proteomics technologies

Welcome to more course, on a precations of Interactomics, using Genomics and Proteomics technology. Before we start this course, let me give you, the genesis of, why we are going to offer you this course. This course is mainly to update you, about various application of, advanced, high through put technologies, like, Micro arrays, Next generation sequencing, Mass Spectrometry, Label free platforms. If you understand these technologies, their principles and possible applications, then it can actually help you, to address, many biological questions. Irrespective of, which field, which discipline, of life sciences you work with, you can definitely get highly benefited. Let me also mention, even if you are not a biology student, you are student from, Computational background or Bioinformatic background, even you will get highly benefited, for attending this course, because, you will get to know, how the big data is being generated, and what are different pipelines, which could be utilized for, it's data analysis. So again, this course is definitely going to broaden, your scope about, current latest technologies available, which could be utilized for, various types of valuable applications, especially, in Interactomics applications, in the areas of, Genomics and Proteomics.

At IIT Bombay, we conducted a workshop, in 2018 and this course is actually a modified version of, this workshop, where we invited, many eminent scientists and application scientists, especially one of the distinguished, foreign faculty, Dr Joshua LaBaer, he visited us and participated in this course and workshop. And many other, academy and industry leaders and speakers were also involved in giving lectures and providing, hands on sessions, during this course. We realise that, it was a major effort, to really educate community about, the latest advancements in this area, directly from this experts. So by attending this course, you will get highly benefitted by, listening the lecture, directly from the experts and also getting a feel of, how to do the experiments, in this area. And this course, we are going to cover in the next 8 weeks. Where various lecture, hands on sessions and assignments will be given to you periodically. Before each lecture or even before we have any expert or the invited speaker, I will provide, a brief overview of, what the lecture can concern about. And also give you, my summary of, at the end of lecture, what was concluded from that lecture. So this will also help you, to bring the perspective and try to, understand the whole lecture and course, in a systematic manner. Apart from the intensive lecture series and the demonstrations of experiments, this course will also include, weekly assignments and interactive sessions. We'll also have, the live sessions, directly from my Protobics laboratory, at IIT Bombay. Where, I will be involved, directly in showing you some experiments, along with my teaching assistants and we will take your query live and try to address your comments, you know, any concerns you have or any curiosity you have. This will definitely, you know, open up and broaden your, understanding, as a scope, of how to do this experiment, directly in the lab. This session, when we have offered many other more courses, have been very stimulating and very lively and lot of you know, participants like this kind of personal interaction. I am sure, that this course also, we will continue this effort and you will get highly benefited, by directly interacting with us. So now let's talk about, today's class. So, you are going to see the recordings which we had taken for my lecture during the workshop.

The very first lecture, where I provided, an overview of, Micro Arabious technologies, So imagine that you want to understand a biological system, complex system and you have a very small amount of clinical sample available for testing or very small amount of protein or drug available, to do the testing, for the thousands of protein. So Microarray technology, where you have, thousands of proteins, printed on the chip, came here, very powerful platform, for high support applications and its feeling. So in this lecture, I am going to talk to you about, the applications of, Proteo Arabious

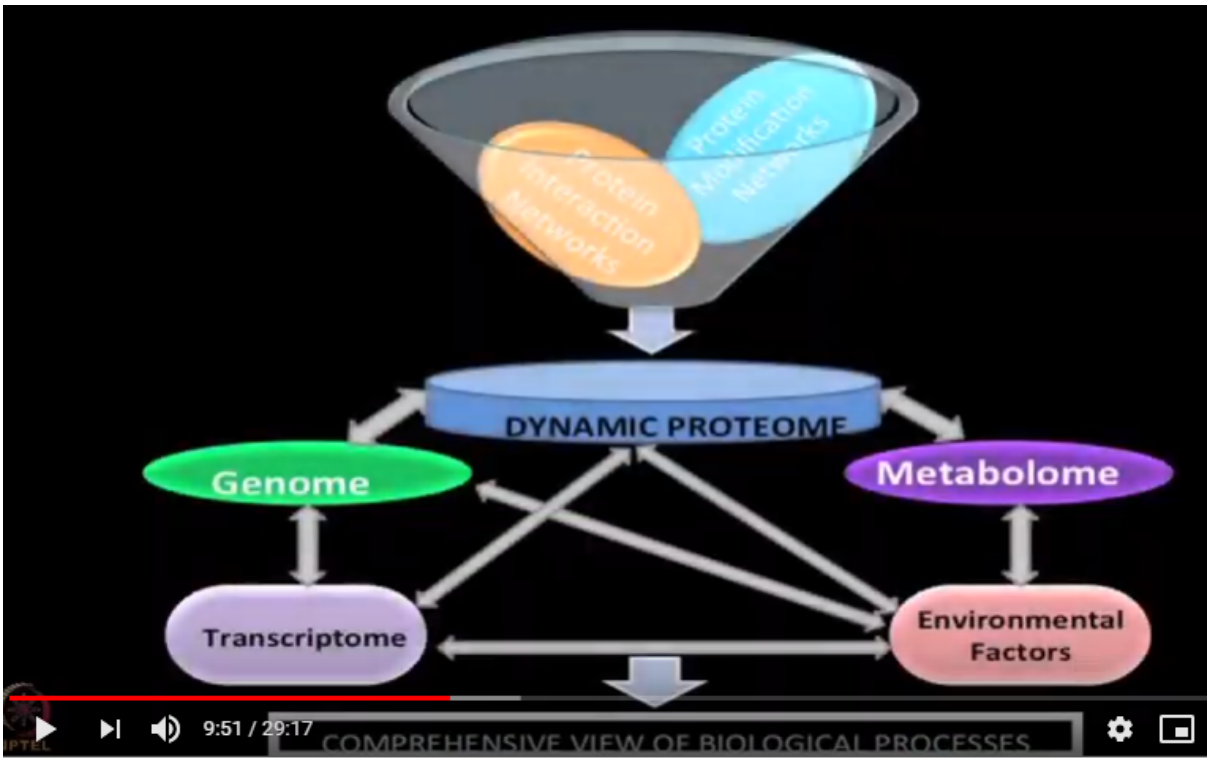
technology, its Genesis, it's advancement and it's various applications, in the areas of life sciences and translation biology. I hope, you will enjoy this first lecture.
Let me start the course detail. The course, will actually revolve around, in Tactomics, where Dr Joshua LaBaer, he is a expert scientist, who works in that area. And you know, fortunately, I was actually, I did my post Doc, when I was at Harvard School, under Dr Joshua LaBaer. Then he moved to, Arizona Institute, Arizona University, where, he is currently, a Director of Bio Design Centre. So I feel really fortunate, that you know, I got training from, Dr Joshua LaBaer and today he said, he's coming here, to India, to, conduct this kind of course, 5 days' on core. Which is quite a significant of time contribution, for somebody of his, busy schedule. When I was doing my post Doc, we were offering some course, in this area, of Tactromics and Proteomics, at Colds berg Harvard and we went almost 3 times and those courses were almost, 14 days long. So, you know just imagine that, you know, 14 days' you're, a kind of conducting a course, 2 weeks long course. And, and courses are you know, very intense, like full day course. And the plan is that, people come without any prior background, any knowledge and then over the period, then kind of they feel, they're really confident about taking those things. And many of them are, you know, like you, like many of them are actually dependant PI's, who want to setup their groups, who want to bring up new technologies to the lab. So that was a very good experience, you know, both, working with josh, as well as many as experts and also how to, you know, conduct quality courses. And I must say, that you know, I was always thinking and you know hoping, that can we conduct some of these kind of, you know quality courses back in India, you know when I joined back to, reached to India. Can we do, some of these kinds of courses? where we have real experts, who will talk to you directly. Because many times, we do courses, but you know, may not be the person or the, be the best experts in the field. So this is one of the initiative, I may not say that this is probably going to live up, to the same level Colds berg Harvard, what we would have done it. But atleast this is one effort, in that direction, to give you, some training and experience, with the direct experts of the field.

So as you can see, that you know, on this image, we are going to talk about, several things linked to both Micro Arabious platforms and Label-free biosensors. So, imagine that you know, if you want to label your proteins, then for detection you need some sort of, you know readouts, you personal waste readouts. And those are very powerful and many labs have those kinds of scanners and the ability to perform, these kinds of experiments. So therefore, Micro Arabious waste and specially label waste, which has been quite popular. Nevertheless, you know, whenever you're modifying a protein, whenever you're adding certain, tax or you know, some sort of, you know labels, then there is always a chance, that protein gets modified. And what signal you see, may not be true, it could be artefacts. And, you know the whole process of doing labelling and doing the entire experiment, is also very tedious, it takes lot of time. And again when you do Micro array experiment, it will be, you know, a powerful platform, but what you may rely, the whole day, you were pretty much, working in the blind area and then at the end of the day, then probably you will, kind of sly.

And then you realise, okay, can I see a signal or I don't see a signal? So then whole day you have **[8:54]** no control on your experiment, it is like western dot, whereas in label-free technology, you have the ability to modulate experiment, to change experiment, in the live manner. So in the real time way, if you see a binding is not happening, then probably you can say, 'Okay, now the right consultation is not correct, my mobilisation was not good, I should actually change the Ph condition, probably I should not change the temperature, because that interaction will only happen in that condition'. So in the label-free manner, there are approaches, including Surface Carbon resonance, including Bioleral Infermatory. Many other new platforms are coming, Bio sensors are coming, which could be used, to do the interaction analysis in the, label-free manner. So again, these are some of the

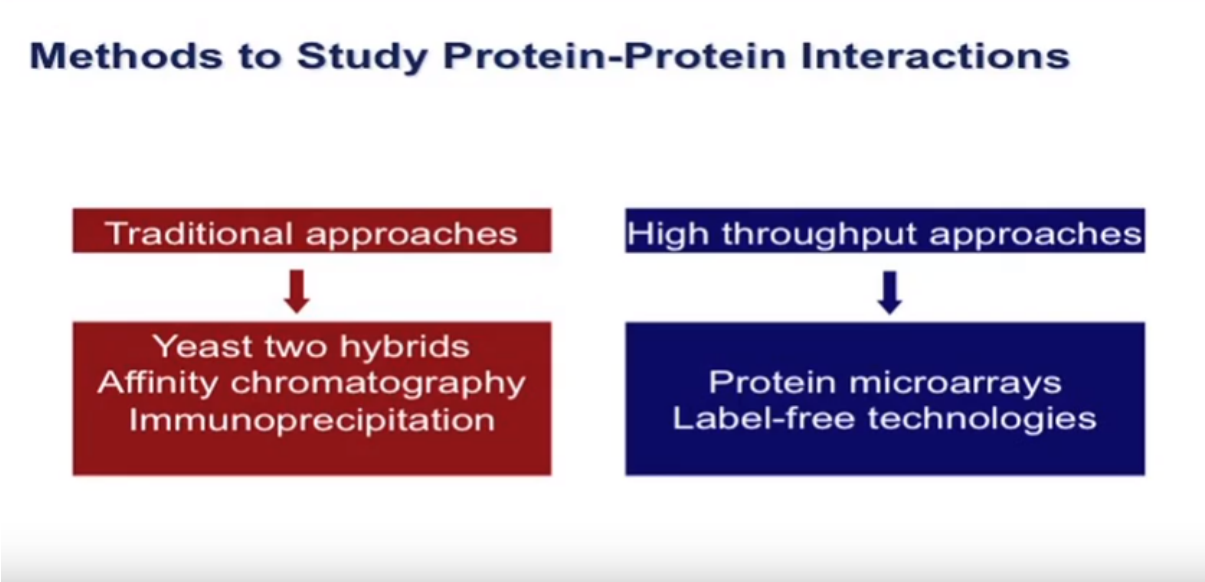
newer ways of, thinking about of, how to study Bio Molecular Interactions. It's not limited to only protein, protein. You know protein, small molecules, even protein DNA; all of those interactions are possible.

Refer slide time (09:51)



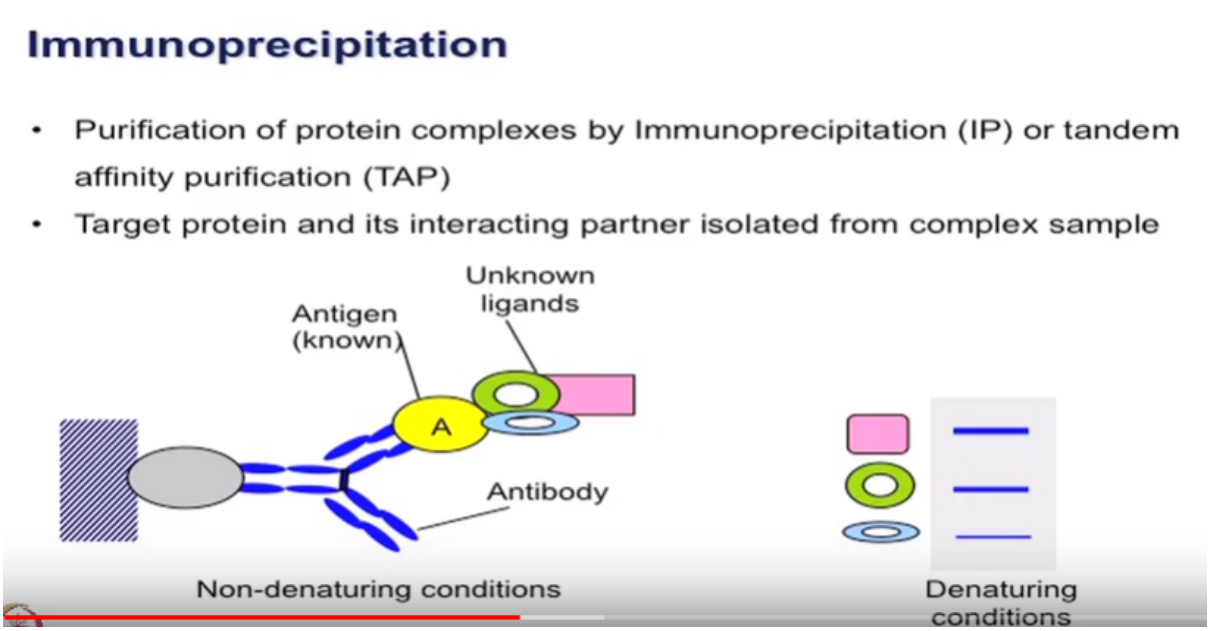
So, this slide just conveys you that you know, how dynamic the Proteome is. And of course, if you really, want to study, any physiological system, just studying one bio molecule is not sufficient. So, you really need to know, you know, what are genes involved, which are the transcripts involved, what could be Metabolites, how would the mental factors, affect them and then of course the proteins which are very dynamic molecule, by different modifications happening. To go there, once you study these things, including that, what network of proteins and bio molecules, which are, you know, triggering different cascade of events, the activities, those things becomes very crucial for us to really identify it's **[10:30]** physiological system and then try, to get information, in much more systems way. So, therefore a system Biology field is really growing much, where intention is not only to look at, one field at a time, or one property at a time. But rather look at the dynamic molecules together and see how a system works. And many times, you know, when you, try to extra collate information, at the protein level or DNA level, RNA level, you're seeing, you know, small bit of picture. But when you look at the systems network, probably you're seeing a, much bigger, much different picture, of the whole system. And that is something which is, very much computational biology driven field, where lot of, you know, computer scientists are now, getting involved, to, to come with a big data and make some sort of notes and some hypo sols based on those, which can be tested back in the lab. Alright, so, let me start with, you know, couple of classical ways of studying, a Protein, Protein Interactions,

Refer slide time (09:51)



which ideally, you are, you know, briefly familiar with. So, the conventional approaches of using protein, protein, an interaction varies to hybrids, different type of human precipitation methods, Affinity Chromatography etc. More, latest approaches and more high support, technologies have emerged, which includes protein micro arrays and different type of label-free technologies. So this kind of a broad picture, which you can keep in your mind. That all of these techniques in one another way are going to give you same formation. But, you are, there are some classical approaches and by learning from those, there some newer approaches, which have tried to overcome those approaches.

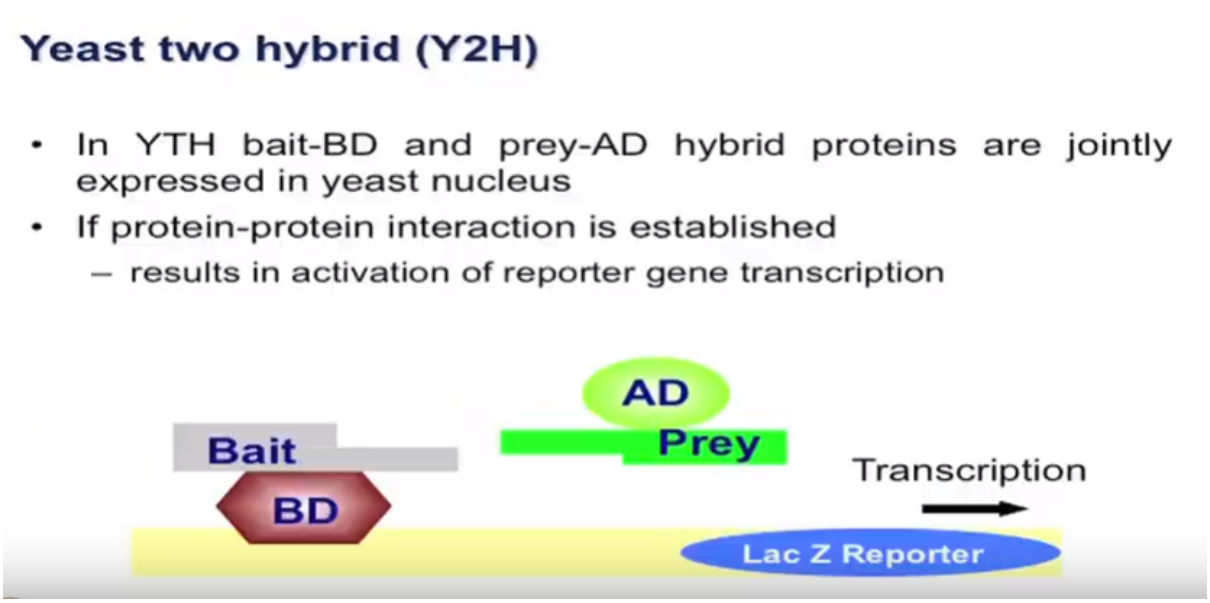
Refer slide time (12:10)



So Immunoprecipitation, you identify the protein of your interest and now you want to identify, which are the possible interactors. So to do that, let's say you're using, you know, this is case, when you

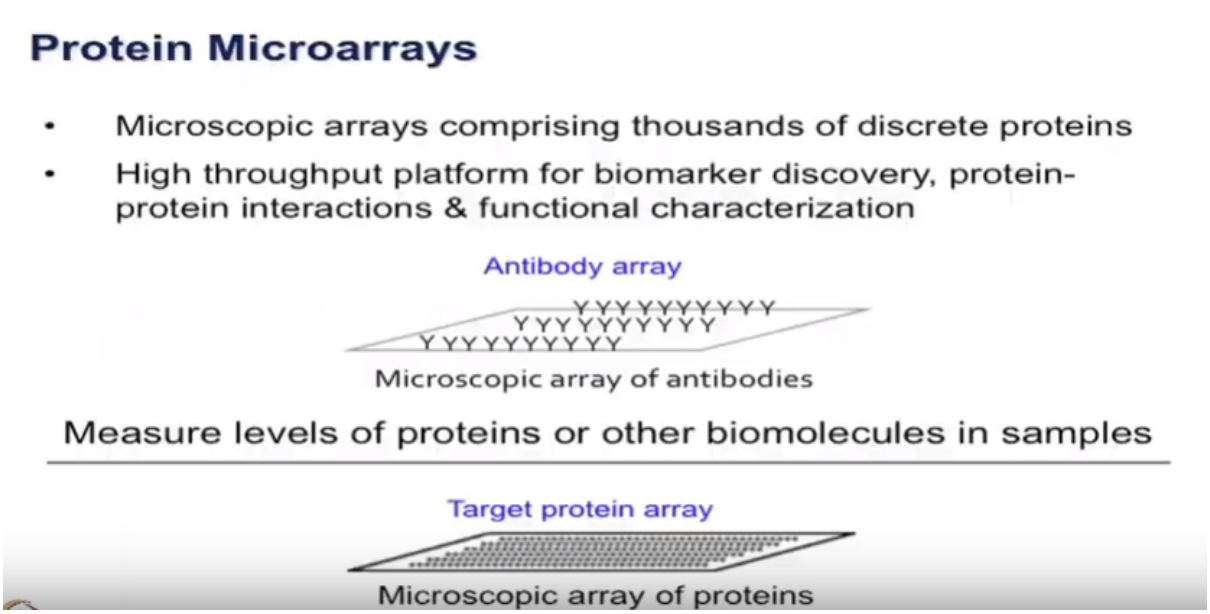
have an Antigen, which is, known for your interest, this is Antibody which is binding. And now, many other potential interactors are actually binding, which are having potential interaction, which you do not know, direct interaction or indirect interaction. But they are the potential interactors. But then, when you are resolving them on the gel, then, on the denaturing condition, then you can separate them, those well. So here you're trying to provide non-denaturing condition, so that interactions could happen. And then you're trying to allow, the denaturing condition, so that you can separate with interactors and this the where, you can then use, Mass Spectrometry, technologies to identify, the known, orpotation interactors. Many times, you will find out that, many proteins are very sticky proteins. They are not the real interactors, and that's where I think then, you will feel to need to have, more you know, latest approaches and more kind of you know, high throughput approaches, to look for more, direct interactions.

Refer slide time (13:07)



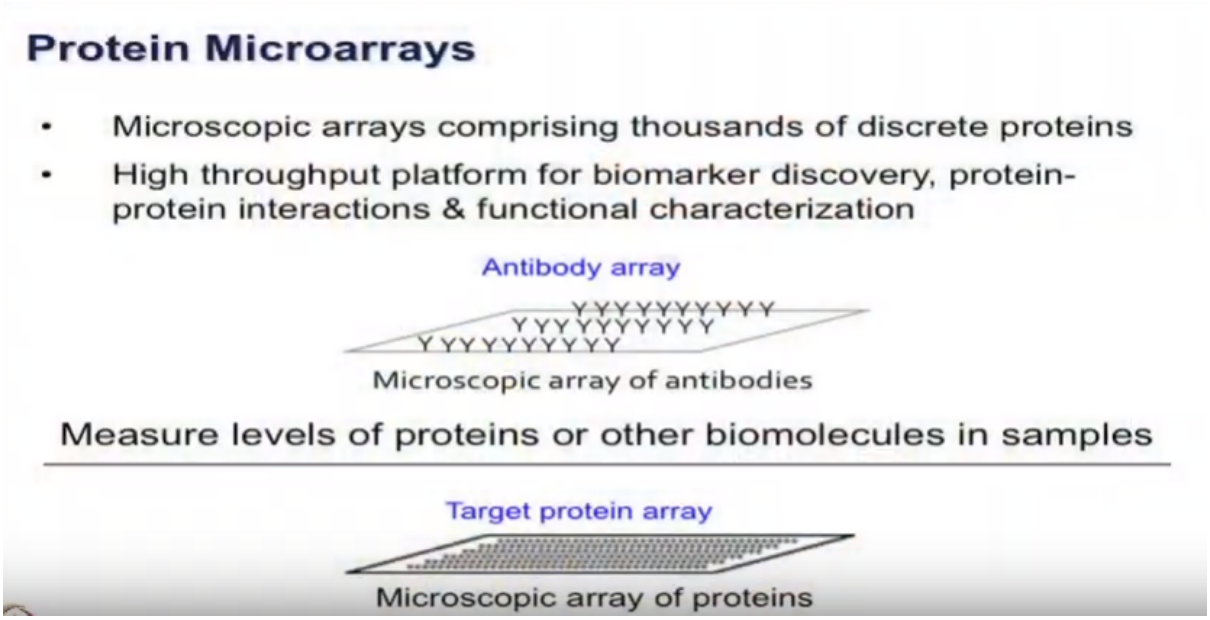
The Classical approach of the Yeast Two Hybrid (Y2H), has been in field for very long time. It is still, being used very heavily. Where interactions, are you know, being used, especially in the Yeast kind of involvement, when you have a weight and have a prey, binding domain and activation domain, when they come together, then the transcription can happen. And this kind of approach is very powerful. But major problems, which you people have seen, there are lot of false positives, which comes from this clinic. So when you identify, hundreds of interactors, with Yeast two hybrids, you are very not sure, you know, how many those are really relevant for you to, really take it forward. So, this can be good starting point, but, of course not the most powerful way, of doing the protein, protein or Bio Molecular Interactions. So therefore, some of the newer approaches have come forward, which includes Protein Micro Arabious, platforms and that kind of stuff, which we want to talk in more detail today.

Refer slide time (14:03)



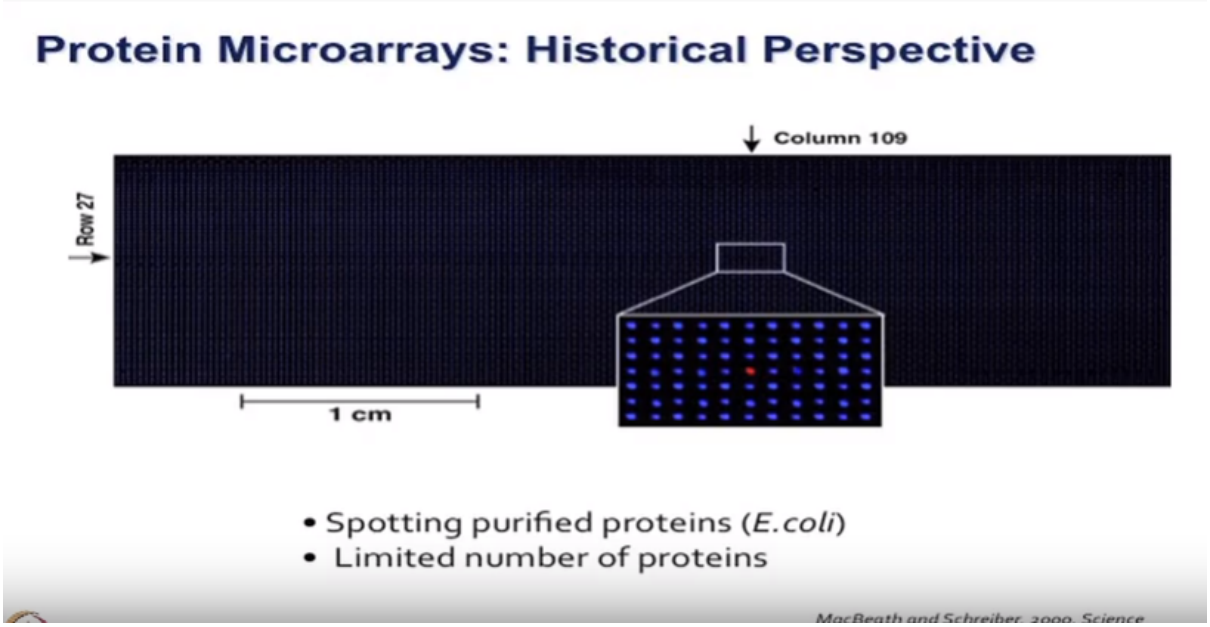
So, like you know, other type of Alibo arrays, which were in field, in 1995 or so, more. At that time, many type of, Genome sequencing projects were happening. So, looking at the Genome sequencing success, people were able to get all the aligos, printed on the chips and therefore, at that time, between 1995 to, let's say 2005, that was the time, when many type of chips were produced. People started thinking about screening, you know, lot of genes simultaneously. From that success, which you know, people got inspired in the protein field and they thought, can we, you know, replicate the same success at the protein level. Of course, you know, you all are aware that protein production itself is very challenging. You don't have, you know, capability to produce proteins, at straight forward way, you know, as you can do, using PCR for the aligos. Right? So, but nevertheless people thought to, to use those kinds of approaches as well. And what the different type of arrays came in the field, which are listed here;

Refer slide time (15:02)



one approach could be, Antibody based arrays, if you have, you know, antibodies available, good antibodies available, if they are the mobilised on the glass slide on a **[15:10]** or different types of sub straights, then that is termed as the, 'Antibodies arrays'. Or if you have purified proteins or proteins produced, in any, any possible manner, those can be termed as the, 'Target Protein Arrays' so this about classification, which started with, to the field. Anti body arrays are very robust, as long as they have access to good antibody, then you know, you have very clean and very neat signal. But, I'm sure, you are aware of, that there is not very good anti bodies are available and you know, purchasing level is very costly. So the overall, you know, doing the experiment with this anti body arrays, are very limited. So, then many times, you're only looking at handpicked proteins or antibodies, which are available for it and you only want to limit your screen to that.
You cannot afford to do, high throughput screening, for all kind of proteins of interest. Right? So, nevertheless, all of this intention was to, print the contents, print the proteins or antibodies, on the chip, which are on the glass slides, in a very high density approach, so that, even with the very small amount of contents, a small amount of your clinical sample or any kind of molecule, which you want to screen, you can still do the screening on thousands of molecules, simultaneously. So that was the, genesis of the whole field.

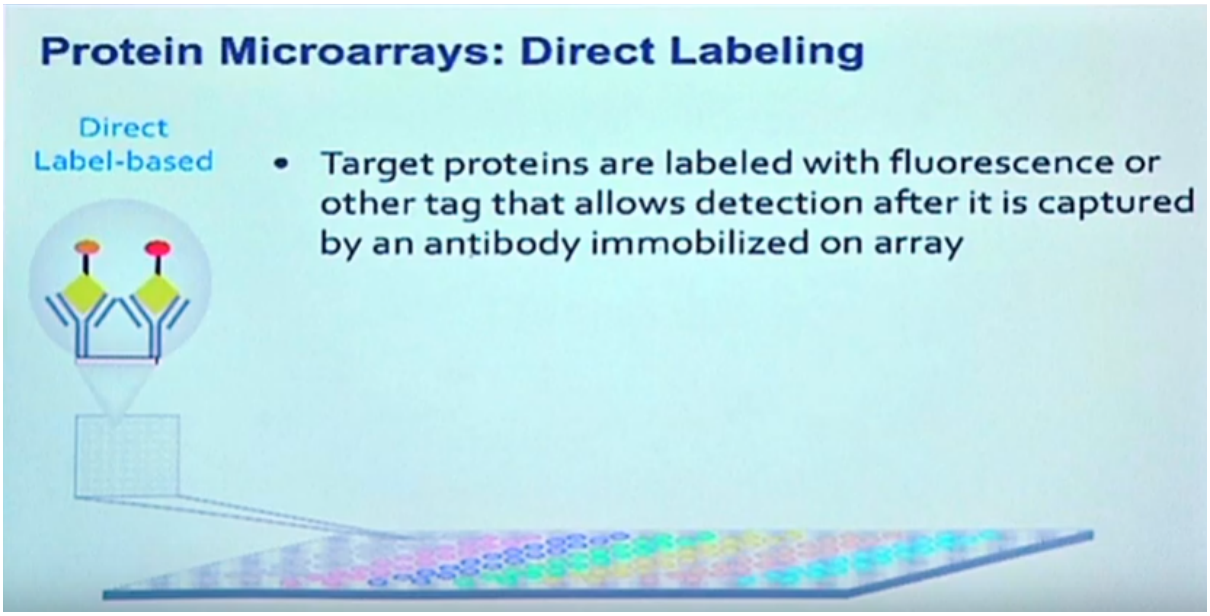
Refer slide time (16:24)



The very first approach, of taking this particular type of, Protein array, field forward was done by Gavin Macbeth, at Harvard. As I mentioned, you know, that time 2000 or so. That time, lot of, you know, people were publishing the word, based on Oligo arrays, different type of; chips were available, at that time. And he made the first attempt, that can we print the proteins on the chip and can we use that, for a screening. And that is the first time, he made an effort to extra collate understanding, from the genes to the protein level. And as a result, so this was, you know, very smart experiment, where only two proteins were used. One of the protein tissue on the red is different protein and rest everything is one, same protein. And he just showed that you know, if I'm incubating with the, you know, specific antibody, I can that particular type of signal, which can distinguish from other protein. So, despite, you know, the fact that he did not have access to large number of purified protein, but still, you know, consumption he showed, that proteins could be printed on the chip and you can achieve specific signal, out of these kinds of, small arrays. So that would, that's why, he got

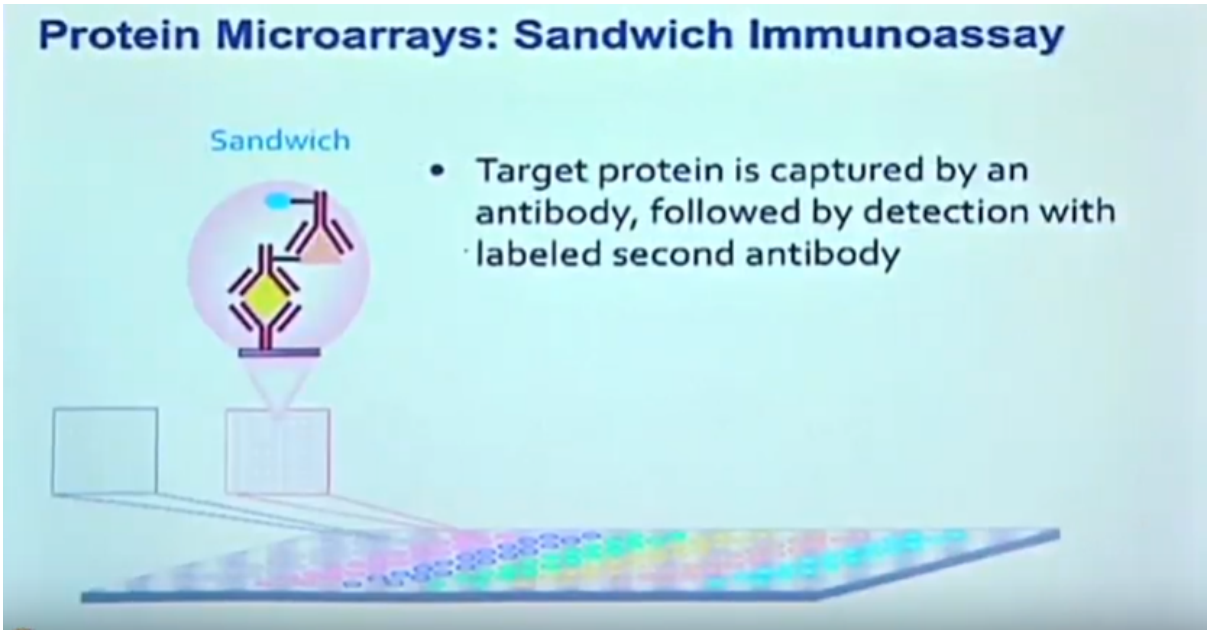
publication in science, at that time. And that, just kind of started the field. Many people who had access that time, of clones, many clones who had in their lab, into special vectors, they started purifying the proteins or those guys who had already had access to pure proteins. You know, they started, you know, printing on the chips. And immediately from 2000 to 2005, many good papers in Nature and Science came, especially from Mike Stydus, when they had access to many of the Yeast Proteins and then they immediately printed 56 hundred Yeast proteins on the chip. And then they did, some sort of, you know, interactional studies on those. So, you know, this, concept just brought forward, many scientists and many approaches that now we can do, high throughput screening, from the, for the protein and do many essays, on the chip itself. So, I'm showing you now, you know, several approaches, which can be used, for different type of micro arrays.

Refer slide time (18:25)



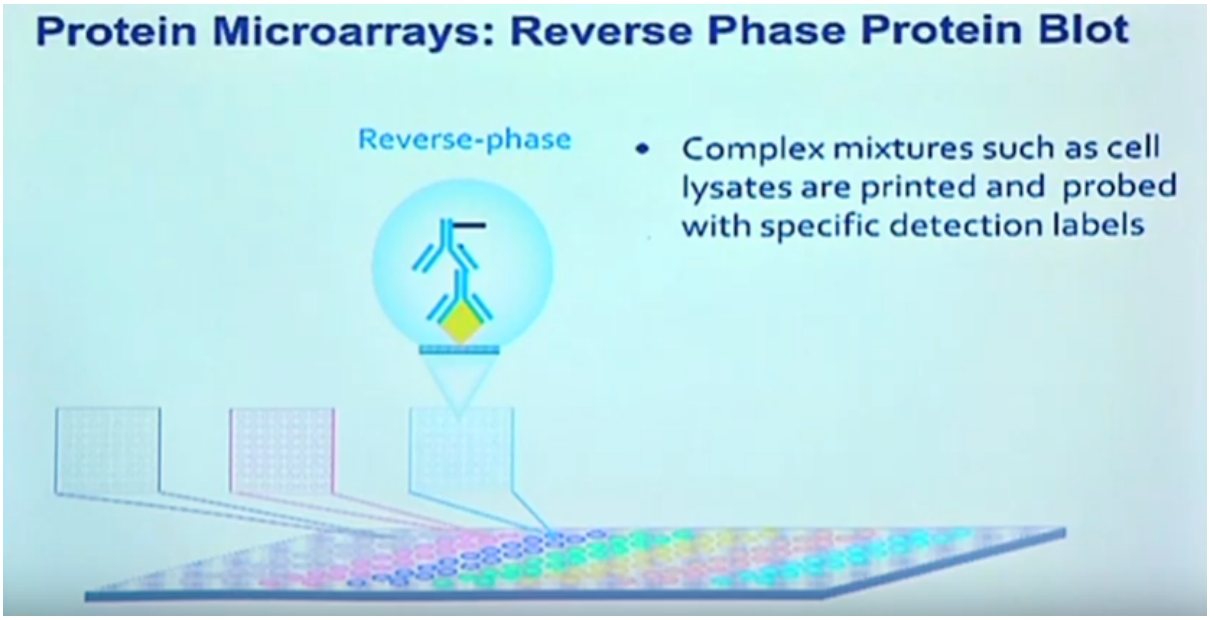
Let's start with here, with the direct label based methods. Now here anti bodies are printed. So initially I'm showing you couple of anti body based approaches, which can determine the abundance based protein arrays. Now, if anti bodies are printed and now you have the target proteins, which are labelled with different type of fluorescent molecules, that can be used as one of the array, which is known as the,

Refer slide time (18:48)



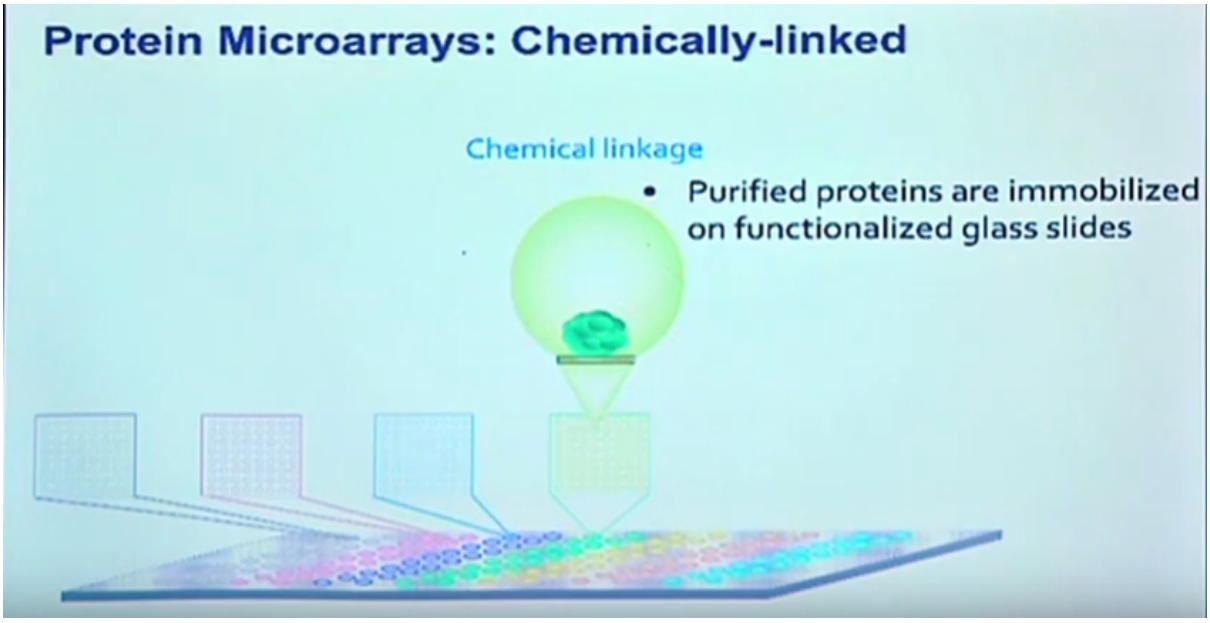
'Direct Labelling Based Method'. Or you can have, you know, Anti bodies and let's say, you have the secondary anti body, which is going to, be labelled with the, Capture Molecule and that is known as the, 'Sandwich Immunoassay', like the way you do for, you know, Elisa Western Blots. So this can be much more powerful, much more, specific signal can be seen?

Refer slide time (19:08)



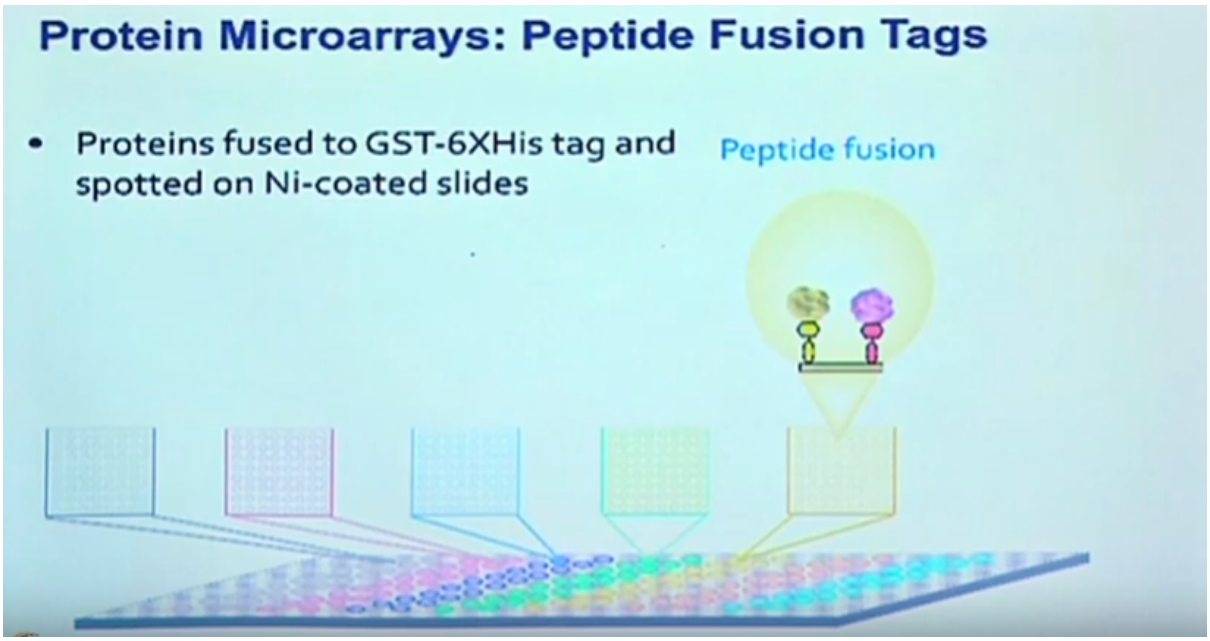
But then, you can also have the, the tissue or your cell lysates, printed on the chip. And then you can have the, specific anti body, which you want to probe. And that is Reverse Phase Arrays, which we will talk in, much more detail today.

Refer slide time (19:08)



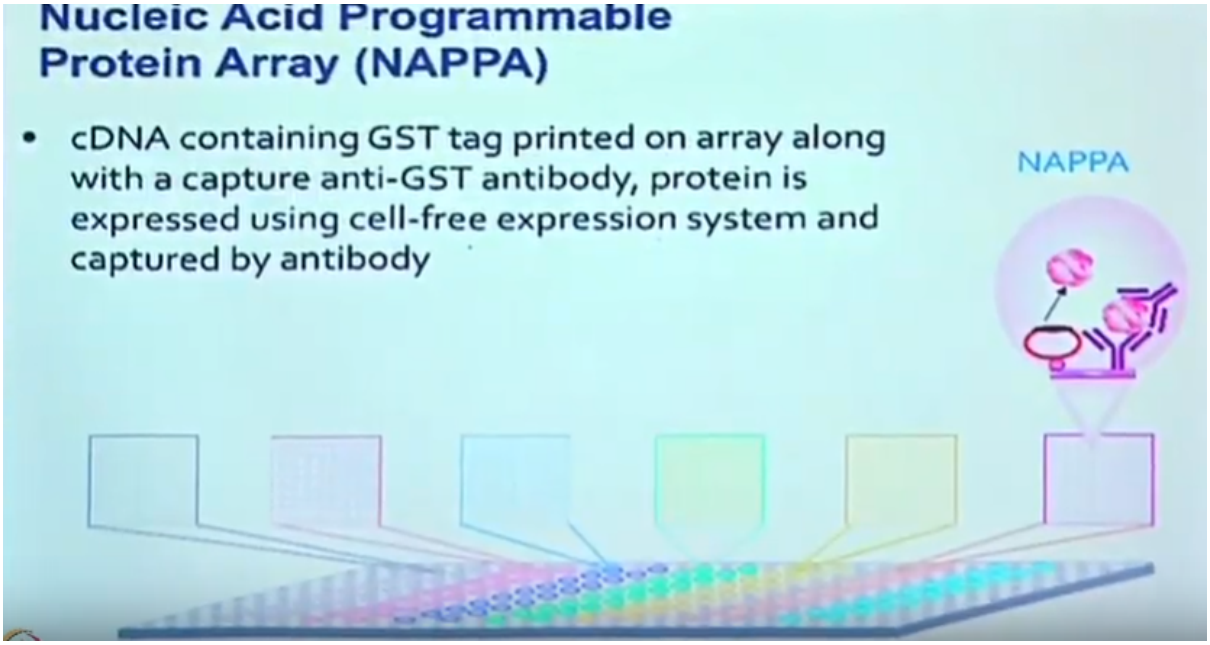
Or if you have access to the purified proteins in the lab, that is the best thing, which you can then try to immobilize those on the chip surface directly. And that is the, you know, using chemical linkage, that is the, purified protein based arrays.

Refer slide time (19:33)



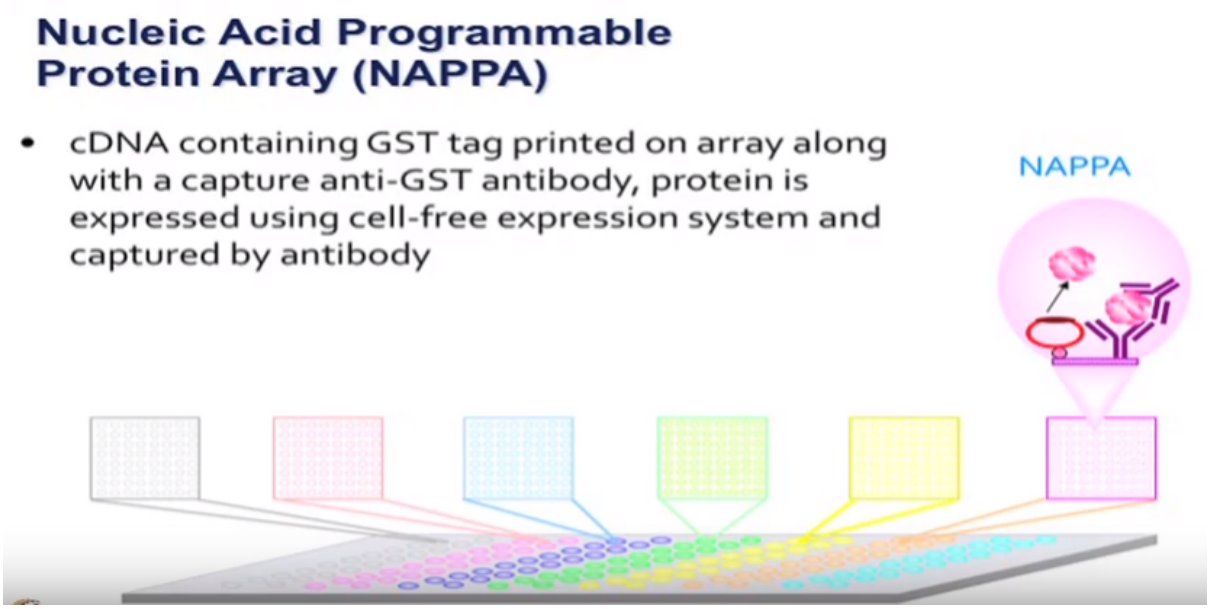
Many times you may not be interested in looking at the entire proteins. You just want to, see that, you know, certain domains, can you, you know, get the peptides for those and then you want to put peptides, itself on the chip. And those can be studied using, Peptide arrays. So, many time, you know, when you have, the clones which are having, let's say, a histidine tag and if you have the nickel NTA (Ni-NTA) type of coating, so this concept can be used here for this array based platform which is shown here for Peptide Fusion based arrays.

Refer slide time (20:07)



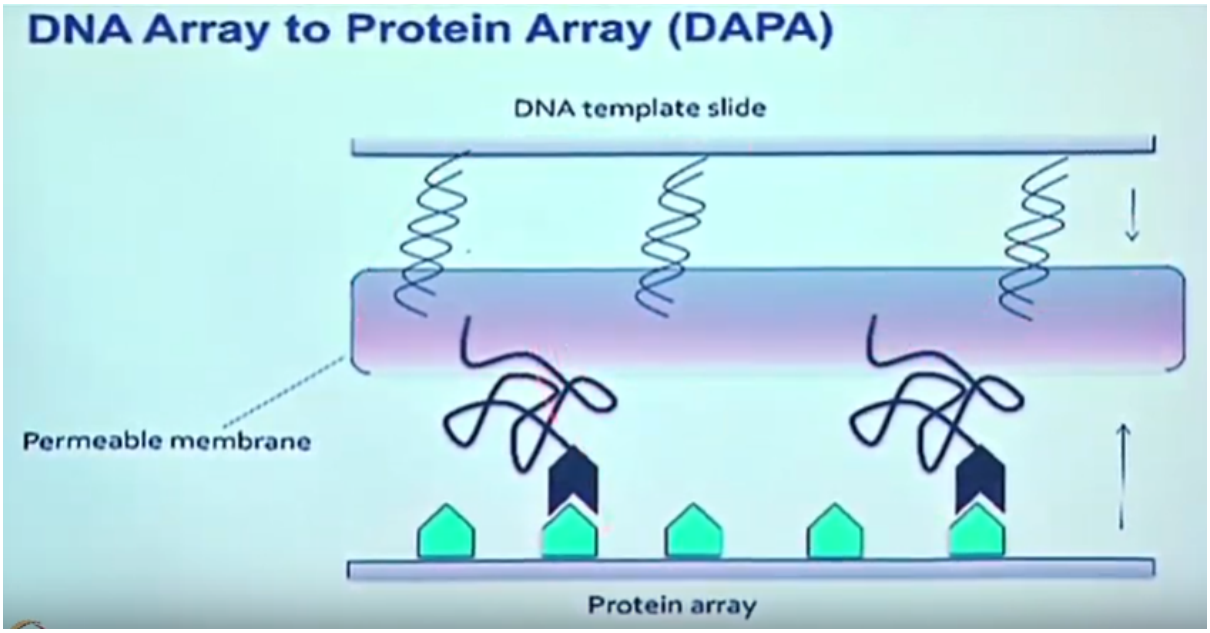
Then we can also think about generating the protein contents, which could be used for the for the making the protein directly from the DNA itself. That is another approach, which you can think about, making the protein directly on the chip, from the DNA and that is known early in Vitro, transcription, translation based method and different type of cell-free expression based methods, can be used for that. Alright, so this part is something which Dr Joshua LaBaer is going to cover in much more detail. So his lab, was first time, they showed, that you can take CDNA and then from those CDNA, if you can do the transcription and translation on the chip itself, then you can make the protein directly on the chip. And that was you know, very revolutionary concept. Because, now everybody has access to laptops, you know, the CDNA clones and if, from those CDNA directly if you can synthesis sufficient amount of protein the chip, then probably that can be very powerful, for doing lot of protein based synthesis. And the concept of in which your transcription translation was not novel, like, it is not done by Josh Lab. It is already in the field from a long time. But thinking about how to use that on the protein array, was the first novel concept. And they came up with an approach, which is, ‘Nucleic Acid Programmable Protein Arrays’.

Refer slide time (21:23)



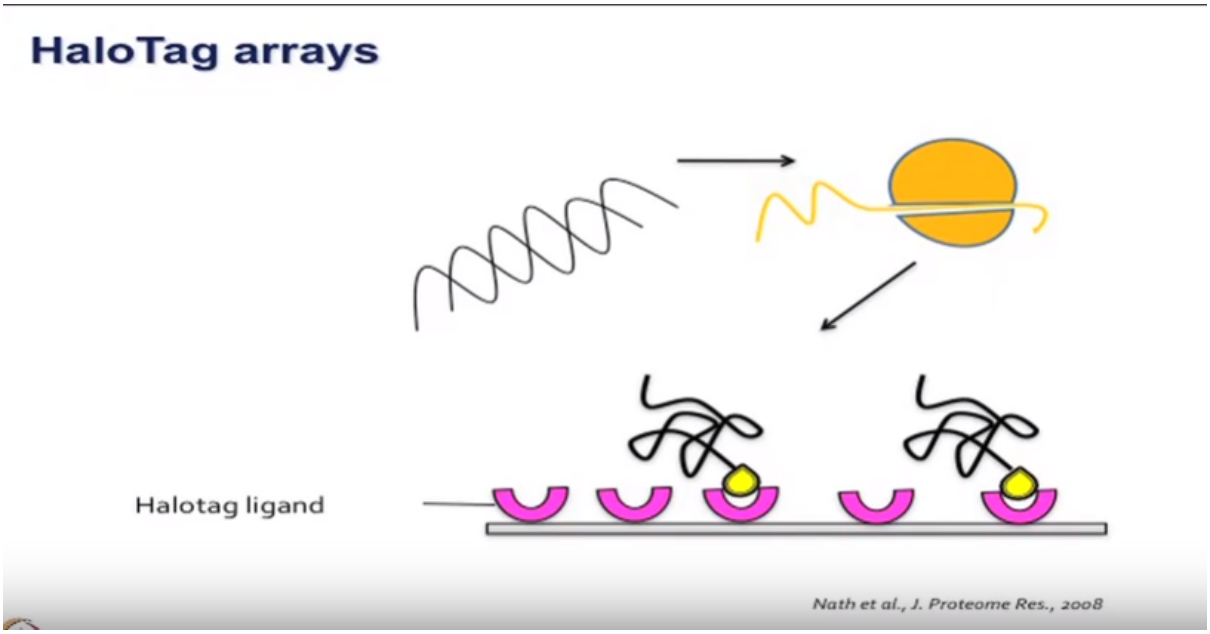
In this case, each of the cDNA , has a GST tag and on the same chip, then if you have immobilize anti GST molecule, so if you're adding in Vitro transcription translation machinery, all the amino acids and you know, the polimary etc, so that you're, you know, the proteins can be synthesized directly, so if any protein synthesized from this particular clone, then if that goes and binds to the anti GST body, then probably you have, a way to detect that particular protein. Of course protein production is very little, it's not too much amount present, but that much amount is sufficient for, detection for different type of Micro Arabious synthesizes. So this talk, you know, this part will be covered of course in much more detail, eventually. I'm just giving you the, the feel of doing different type of approaches. Looking at the previous experiment from NAPPA, with Josh Pepper, then another group came forward with, Multiple Spotting Technology, your Missed Technology, where intention was, 'Can we even take some of the, you know, PCR based products'? Which are unpurified and directly use those products, to print on the chip and then we can still do Vitro transcription translation. Again in, Intent is, that without purifying the protein, which is a difficult thing in the field, can we, you know, generate a, still the protein content and do the analysis, in very high throughput and reverse manner. But each of the method which I've been talking, has its own pros and cons, which as we go along, we'll keep talking.

Refer slide time (22:50)



One other, another approach of using Shelf free, session based arrays, was DAPA, or DNA arrays to Protein Arrays. In which case, let's say, you have template slide here, where you have the DNA printed. And now you have a membrane, a real membrane, in which you have added, now the in Vitro transcription translation mix. Proteins are synthesized from there and they're getting passed. Assume that you have you have histidine tag in these and on another chip you have coated nickel nta, then they're going to bind to the, this chip, which is for the protein arrays. So this is much more pure protein arrays, because you have actually removed, the DNA part of it, only purify protein that is leashing out and that they are getting printed on the chip. So it was again a good concept. Of course it has its own problem of, you know, diffusion and all that, which did not make it so popular.

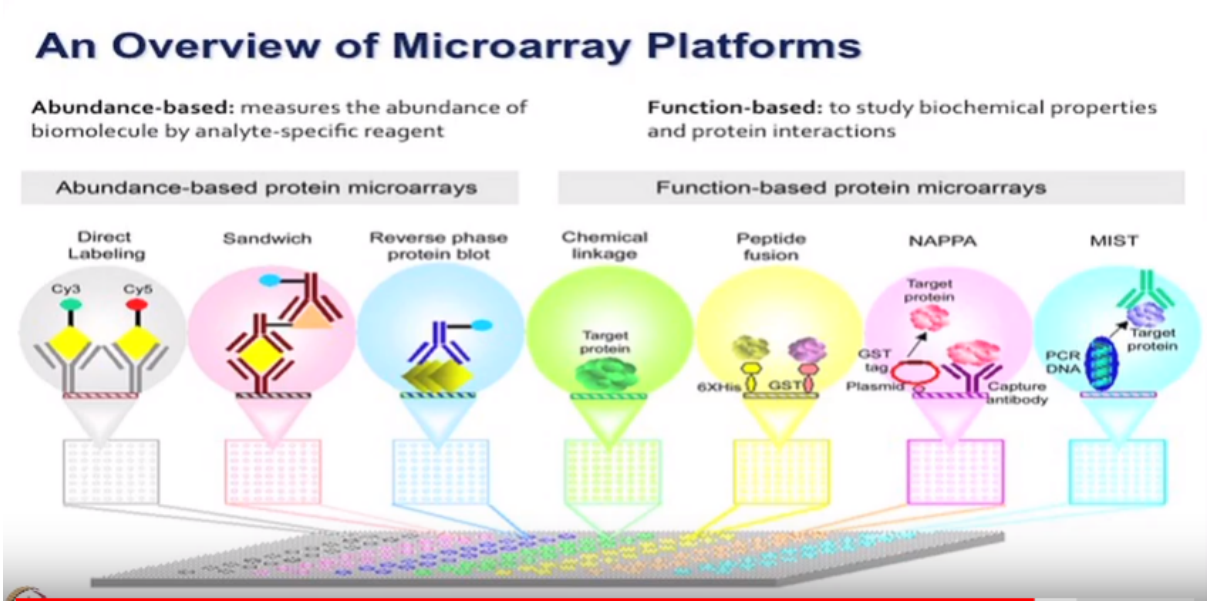
Refer slide time (23:46)



When more approach came, which is, 'Halo Tag based arrays'. So looking at the academic success, even companies came forward. And, Promega thought to use, their existing IVT mix, which they had

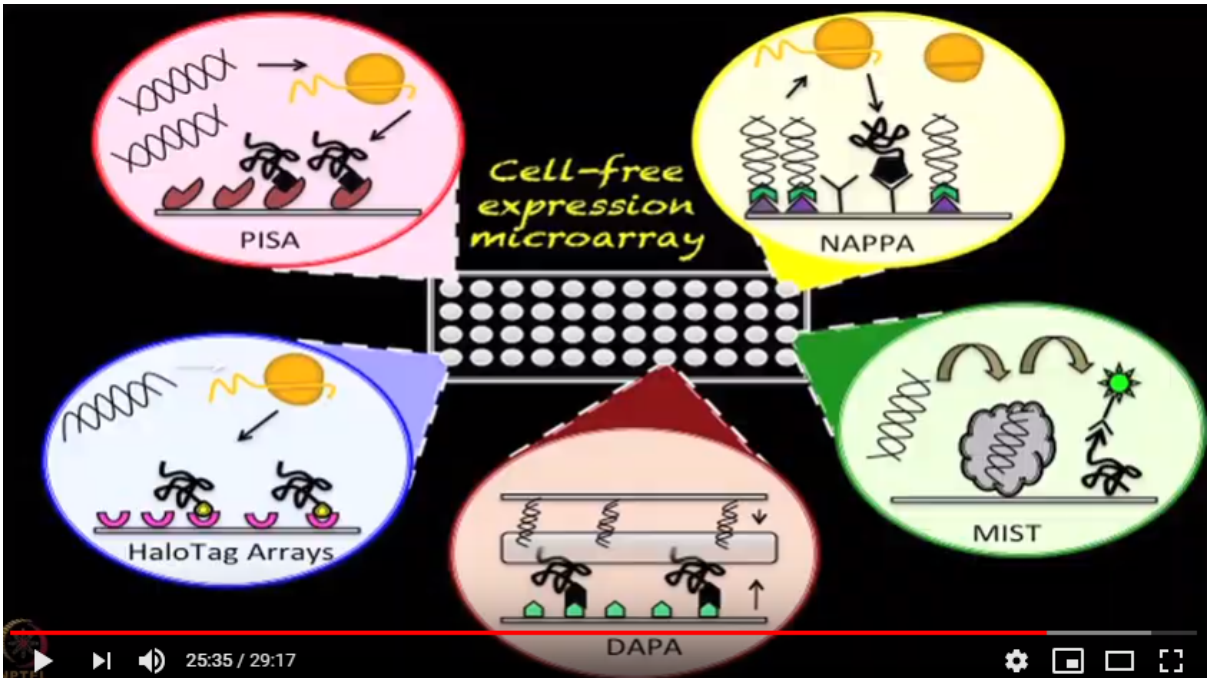
in Vitro transcription translation mix, which they had. And Halo Tag technology, which they already been using for structural work. How we can use those two together, for doing the, protein arrays work. Halo Tag actually showed very covalent, you know, very strong binding, with the lygades on the chip. And if that is the case, probably, when you're using it for protein micro array experiment, then at that time, you're doing lot of washing steps, so your protis will not be washed off. So this was, was actually, you know, very strong, signal you can see with Halo Tag Arrays and they're also providing this particular thing in a small chip format, where you can do, your own in house type of printing. So small, you will get a glass slide and a gasket will be given to you, in the gasket there will be some space, now you can put your, you know, DNA material, which is having Halo Tag contents. So then, now you're actually printing automatically, without having a micro array printer. And then when you add this particular, you know the, the, the mix, proteins are being synthesized, then you can take that, using the Halo Tag, Anti bodies. So these are sometimes, very powerful approach. Some of this, we are also going to show you and demonstrate you, as we go along in the course.

Refer slide time (25:00)



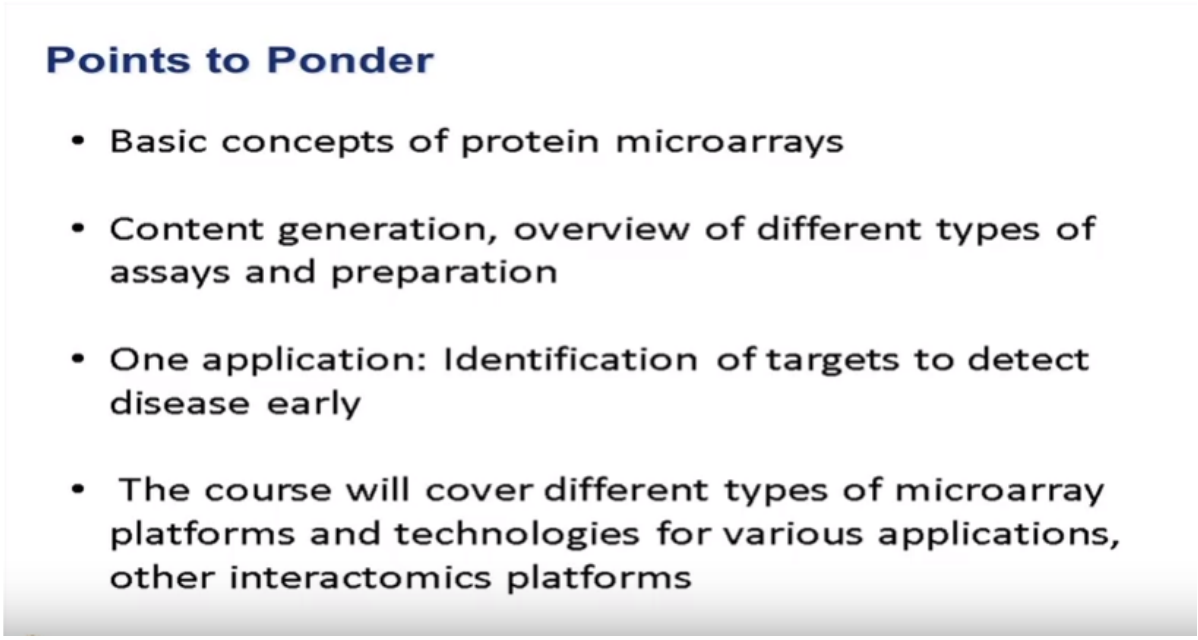
So broadly, you know, as I said, I may not have time enough, right now, to talk to you about, Yeast technology, in much more detail. But I've given you an overview, broad overview, to appreciate that there are many type of approaches that are already in place, started from various type of anti body based approach. Looking both direct and indirect ways of using anti bodies for array based approaches. Purified protein, peptides, and even using your cDNA molecules, to directly produce the proteins on the chip itself. So there are whole lot of things happening in this area.

Refer slide time (25:35)



And within the cell-free expression waste approach itself, there are many technology which are in place. For example, we have nucleic acid programmable Protein array or NAPPA. We have multiple spotting techniques, DNA arrays to Protein arrays, Halo Tag arrays. And one of the very older methods was Protein in situ array (PISA). So, some of these are just kind of, you know, a glimpse to you, to convey that, it is not, a field which is, very limited or which is very, you know, have very small end of, end users. Is actually you know, a grown field, but of course has some limitation of the contents and the, you know, how much density can you produce. Those things, the, they access to the chip. Some of these are the limitations. But there are many ways. And when you go to Europe, actually you will find lot of labs are using different type of Peptide Arrays and different type of, you know, array base, based approaches. Many companies are actually producing those, those kinds of contents now. So this is one of the growing fields. And of course if you're in a tensionless to really something, you know, functional and formation, for your unknown proteins and uncharacterized proteins, this can be one of the powerful platforms. In addition to just, you know, many times your context and you know a protein and you want to understand more about it. But you know, lot of time you have a protein, which is totally unknown protein. Right? And after doing your entire, you know, big screening discovery work, you identify the protein for which, now you want to know, what is the function of it?

Refer slide time (26:53)



So today we tried to convey you, the basic concepts of Protein Micro Arrays. I'm sure you're aware of the basics of, DNA based Micro arrays. Which are the still much simpler technology, much robust technology. But when we talk about Protein Micro Arrays, the content generation itself is very challenging. And then you're talking about, various assets, where you want to address, many biologically valid questions. So having, you know, certain biological questions, like screening of auto nt body, could be very powerful platform, using micro arrays. And for different cancer and different auto immune disorders, it has been shown, that using Protein Micro Arabious technology and platform, one could identify, the targets, which could help us to detect the disease early. And just imagine, that if you're going to a clinic and from the routine blood base test, these kind of testing could be done, using auto nt bodies, where doctors can predict, that this person might be, suffering from this disease, early stage. I think that could be, very valuable information. And Protein micro arrays could definitely give, atleast some clues, in this light. So I hope you got, a glimpse of some of these advanced high throughput technologies, which I discussed, in todays lecture. These will also be covered again, in more detail, in the following lectures, where other speakers, are also going to give you a glimpse of, different type of Micro Arabious platform, different technologies and different applications. See you in next lecture.