Interactomics Basics and Applications

Prof. Sanjeeva Srivastava
Dr. Joshua LaBaer
Department of Biosciences and Bioengineering
Arizona State University, USA
Indian Institute of Technology, Bombay

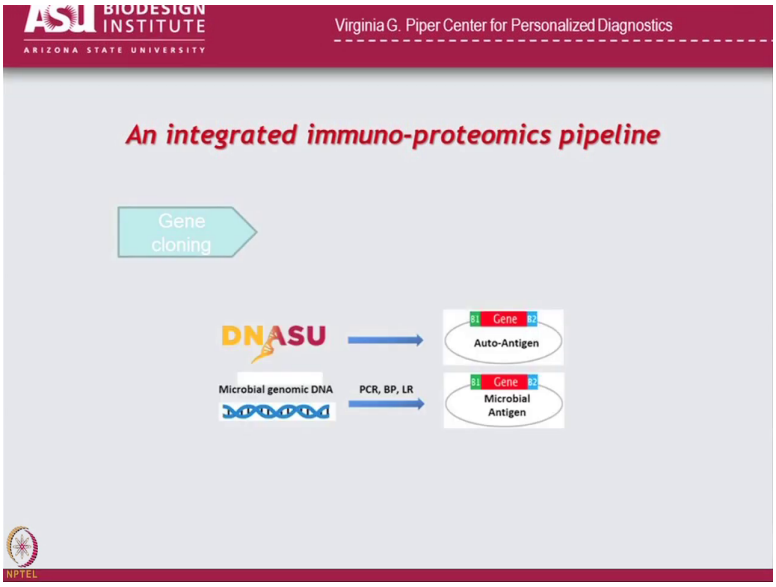
Lecture - 07
Nappa Technology and Protein Arrays-II

In last lecture, Dr. Joshua Labaer give you an overview of Proteomics field and to perform high throughput proteomics based experiment, they need to generate the clone repositories. And to achieve the high throughput biology and performing proteomic experiment, what are the key considerations you need to pay attention for generating those cloned repositories. Once you have obtained large number of gene clones, now you are ready to perform many experiments and one such experiment was developed in his lab which is a Novel Protein Microarray Technology which is NAPPA or Nucleic Acid Programmable Protein Arrays.

To perform NAPPA, if you have these cloned repositories available, you have large number of c DNA clones available, you can print them on the chip and imagine that what goes on in our body to do the central dogma the transcription and translation process from the genes to RNA and the proteins. The same cascade of event could we try to reproduce, could we try to replicate on the chip itself and that was a concept of NAPPA from the DCDNA sequences printed on the glass slides, could we add the machinery which is required for in vitro transcription and translation. And use those to synthesize the protein on the chip itself; looks science fiction, but it was the reality. Dr. Joshua Labaer and when of his senior postdocs DR. Niroshin Ramachandran, they made this technology for performing high throughput protein microarrays using NAPPA.

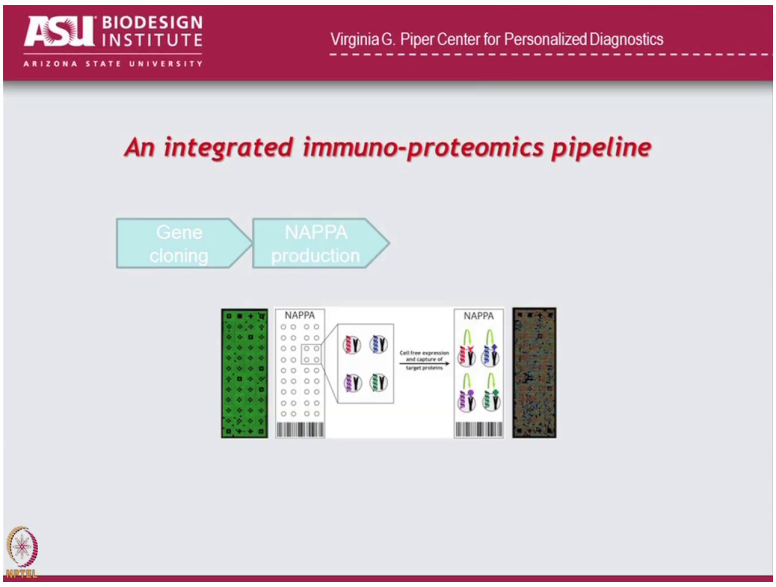
Today, Dr. Labaer will talk to you about the development of this NAPPA technology and how to use this for various microarray-based applications. So, let us welcome Dr. Labaer for a lecture on NAPPA technologies.

(Refer Slide Time: 02:36)



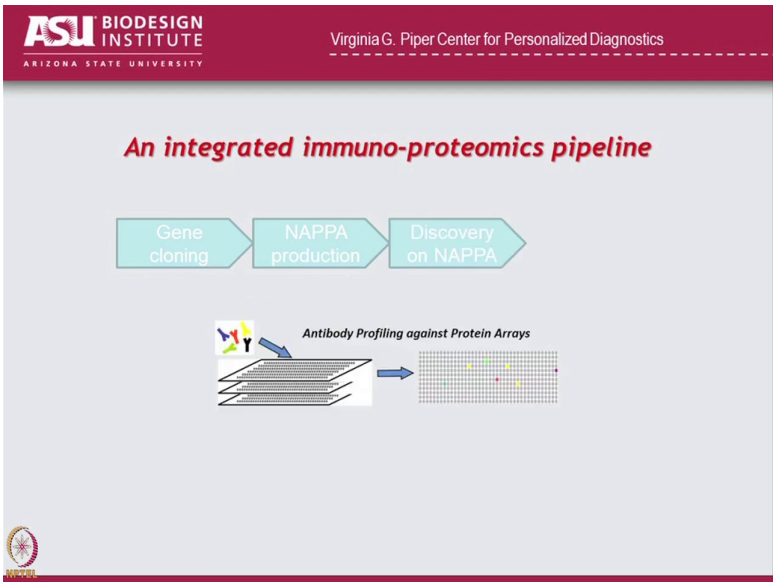
All right. So, we are going to we are going to start now by talking a little bit about, then the NAPPA method right and I already spent some time talking to you about the gene cloning part right.

(Refer Slide Time: 02:51)



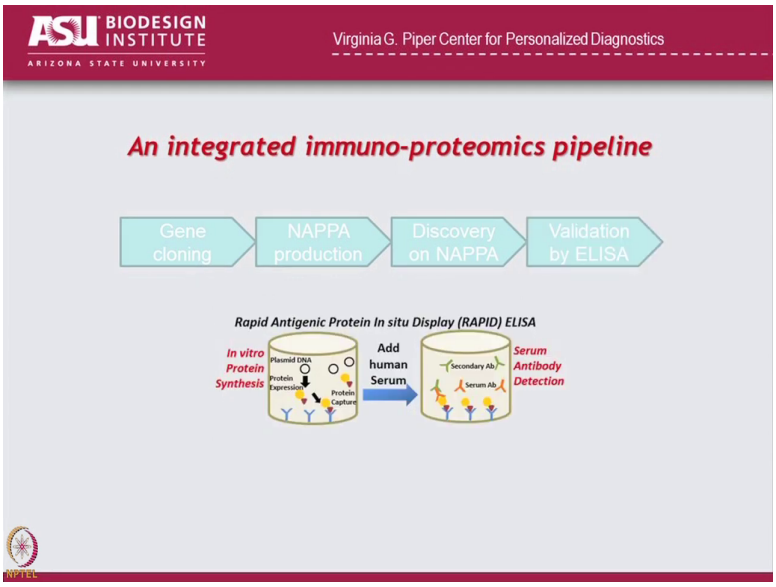
So, that is how do you make the clones for the genes that you are going to put on your protein arrays. We will talk a little bit today about NAPPA production and I am going to do it a little bit from a historical perspective. So, how did we come about this method and how does it generally work.

(Refer Slide Time: 03:02)



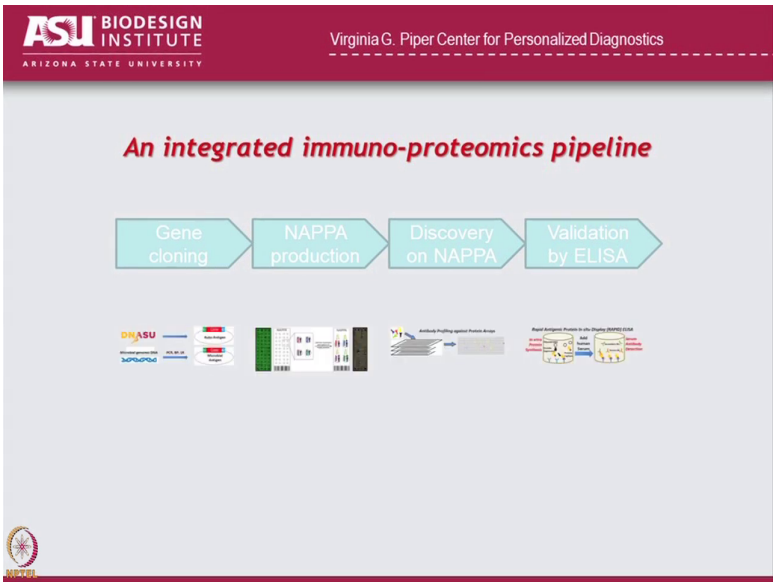
Then, we will talk a little bit about how to do discovery on the platform.

(Refer Slide Time: 03:06)

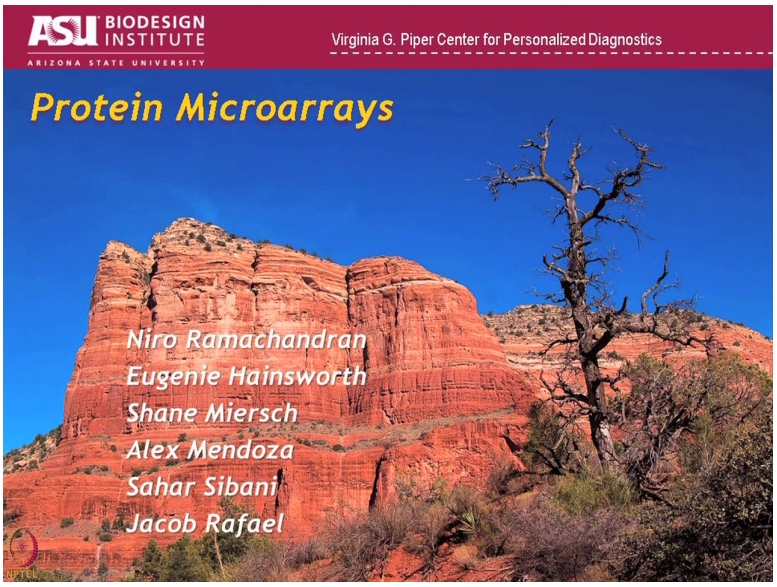


And then finally, nothing you do works if you do not go back and validate it.

(Refer Slide Time: 03:12)



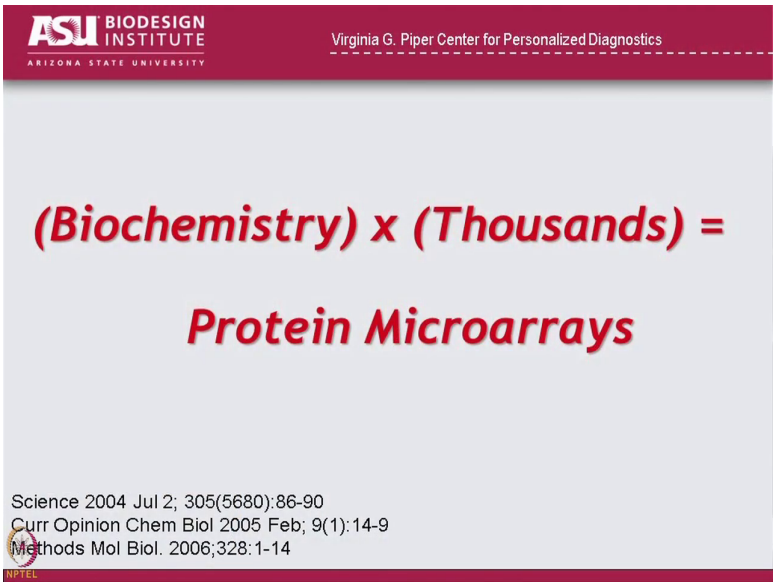
(Refer Slide Time: 03:15)



So, that is kind of like the whole end to end process. So, so let us let us talk a little bit about protein microarrays right; what are they and why do we want to do them. And I am going to point out that these are a number of the people that were most responsible for doing this work. Niro Ramachandran was really the leader in our group that really pioneered this methodology. Eugenie Hainsworth was the engineer, who did a lot of the work and then, some of these folks also contributed to some of the early methodology.

So, we talked earlier about how having a library of expression clones would allow you to do all kinds of different studies. Well, here is one that we got very interested in which is biochemistry. If you want to study the biochemistry of a protein, you need to be able to make that protein and do experiments.

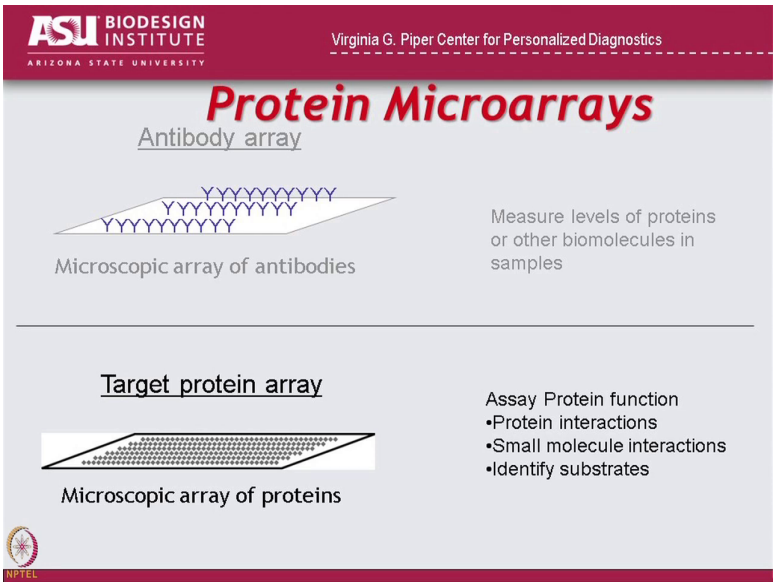
(Refer Slide Time: 04:05)



And if you want to be able to do it times thousands, then that is what led us to the idea of protein microarrays and just to give you some perspective when we first began this work, we actually started by trying to do high throughput protein purification. And so, we were developing methods to make proteins and bacteria. Lyse bacteria, capture the proteins on columns, loot the proteins and then, study them in high throughput and believe it now not we still do some of that work for other reasons.

But we realized very quickly that it is hard to do high throughput protein purification that if you if you try to isolate lots and lots of proteins. First of all, a lot of the proteins do not purify well. Secondly, the ones most of the proteins do not give you very high yield and then third, you do not really know if the proteins that you are purifying are going to be of high quality folded you know an active. So, that is what led us to this idea of protein microarrays.

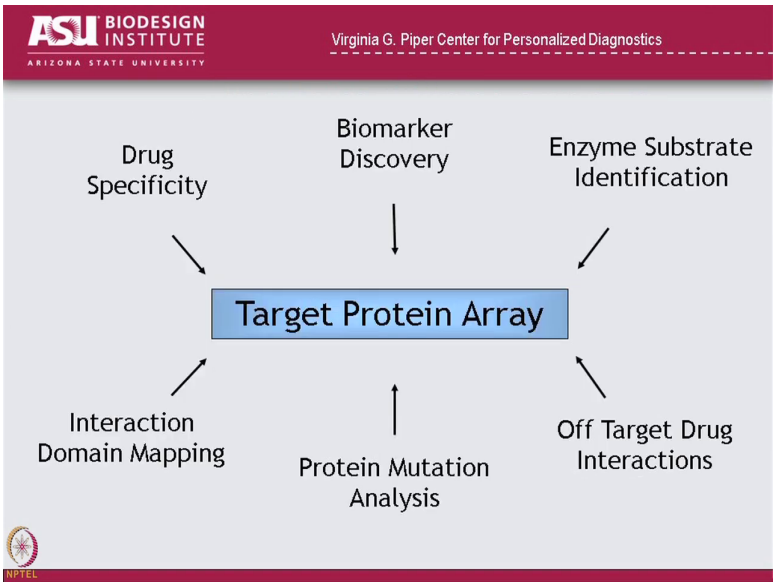
(Refer Slide Time: 05:19)



Now, there are two kinds of protein arrays; the first kind which I am not going to spend a lot of time on or these antibody arrays, where you print an array on a microscope slide, you put antibodies down that recognize different proteins on them and then, you use those arrays to probe a sample to capture whatever proteins are in that sample as a way of measuring the levels of those proteins in the sample.

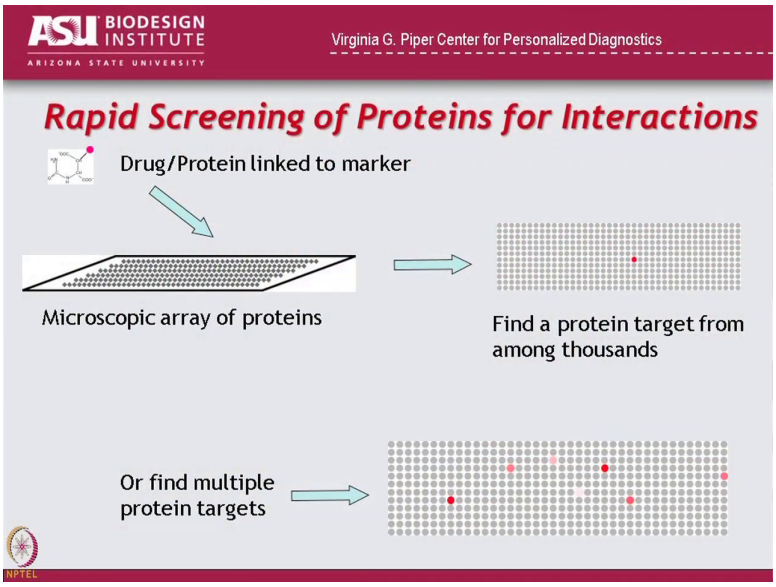
So, the goal of this array is to measure the levels of proteins. The protein arrays that I am going to talk about are called target protein arrays and the goal for these protein these arrays are to look at the proteins themselves. What do they do; who do they interact with; how do they fold what is their function? So, the idea on these slides is that you have a slide and each of these different spots represents a different protein on the array.

(Refer Slide Time: 06:10)



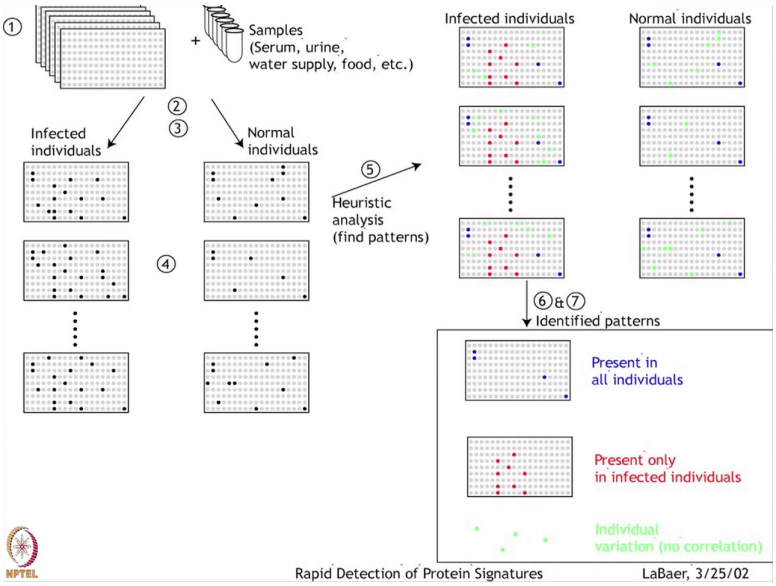
So, lots of things that you can do with a target protein array, you can look at drug specificity I will show you an example of that later. You can do biomarker discovery. You can do enzyme substrate sub identification, you can do interaction domain mapping, you can do analysis of how protein mutate gene mutations affect the function of the proteins, you can look at off target protein interactions.

(Refer Slide Time: 06:37)



So, here is an example of what that might look like. Imagine if you had a fluorescently tagged molecule and you probe the array, you can see which protein that that molecule targets or if it turns out that it binds to multiple proteins, you might get it binding to multiple proteins and you might see the differences in the binding and that would give you some sense of it specificity.

(Refer Slide Time: 07:01)

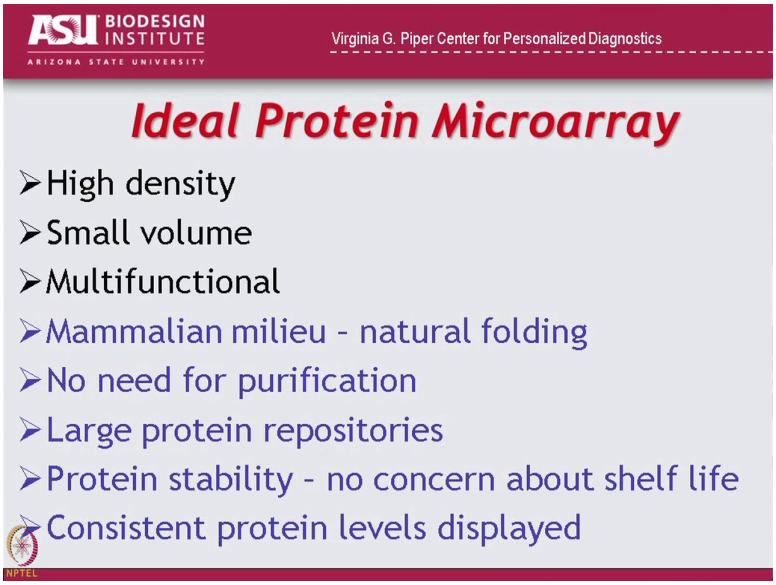


This is the area that that I have spent a lot of micro viewer on and that is looking for patterns of binding that indicate the presence of disease. I told you yesterday that we were looking to do look for markers for breast cancer right and we did that by looking at the immune response those markers.

So, imagine you take this serum and you apply it against in the affected individuals or the normal individuals. Now, I have shown that a number of spots light up and I did that for a reason, it turns out that even normal people develop antibodies against some proteins. The problem is that the word normal in quotes; it is in quotes because all of us have medical histories. We may not have cancer, but we have had other things in our lifetime and those things can affect your immune system and those things can get responses.

So, the key here is to know which responses correlate to cancer and which responses are not related to cancer right and so, you have to accept the fact that there will be these other responses and then, the idea is you do variety of informatics processing to take these, compare these to these and look for patterns like this one here that is present in everybody. A look for these guys that occur only in the affected individuals and then, there will be some that are just random variation that occur from person to person.

(Refer Slide Time: 08:39)



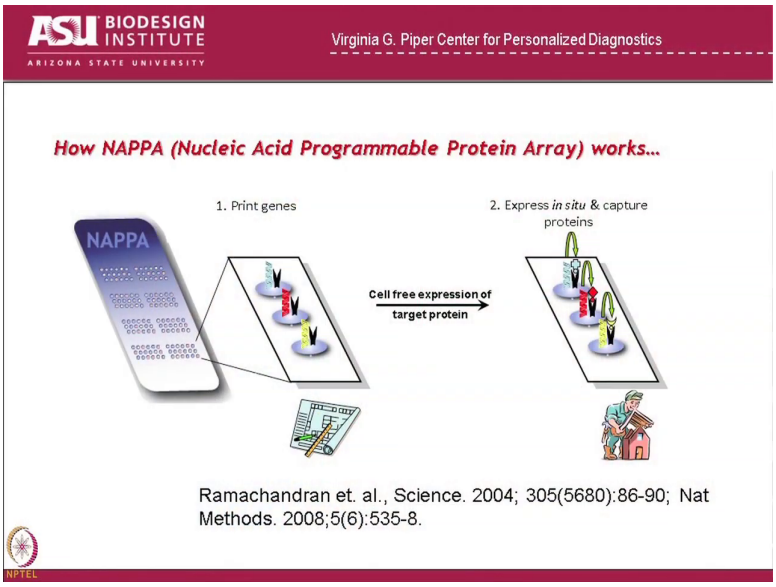
So, that is so that is in another approach that we are doing so. So, let us start here by asking the question, what are the ideal qualities that we want in a protein array; what would make an array a really good array. So, I would argue the first thing it has to be high density. The whole point of a protein array is to get lots of different proteins in a very small space. So, you can study it.

Of course, you want to be able to work with small volumes. The advantage of a protein array is that you can take only a couple 100 microliters or a few microliters of serum and test thousands of proteins with that information with that amount. Of course, it needs to be multifunctional, you would be up to be able to test it with lots of activities; but then you also want it to have natural folding, you want the proteins on the array to look like they do in normal circumstances and ideally, they would be made in a milieu similar to the one in which they normally occur.

You do not want to have to purify the proteins because if you have to purify the proteins, then you are going to end up with all the things that can occur during protein purification. Proteins could lose their folding, they could get low yield, they may not be in a proper conformation. You obviously, want to be able to test as many different repositories as possible. So, you want to build to test any proteins that you would like and you would like the to not worry about the shelf life of the protein. Once you print the protein on a chip, there is always this worry that the clock is starting and the longer it sits there, the more likely it is to stop being active or being well folded. All right.

So, and of course, you want the levels of protein from one to the other to be very consistent. Those are the things that you would really like all right. So, this is the array type that we are going to talk about today which is called NAPPA for Nucleic Acid Programmable Protein Array.

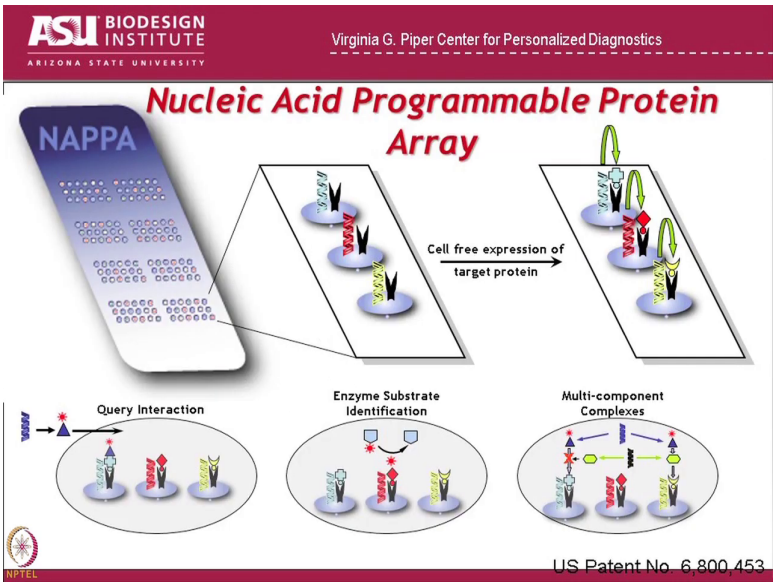
(Refer Slide Time: 10:32)



And the idea for NAPPA is that we print the gene for the protein on the chip and we store the chip as a DNA chip.

So, the gene is there; the clone the ones that we talked about in the first half of today is on the chip along with an agent that is going to capture the protein when it is made, but the protein has not been made yet. Then, on the day of the experiment, we add a cell free extract make the protein and capture it and then, we display the protein at that time fresh.

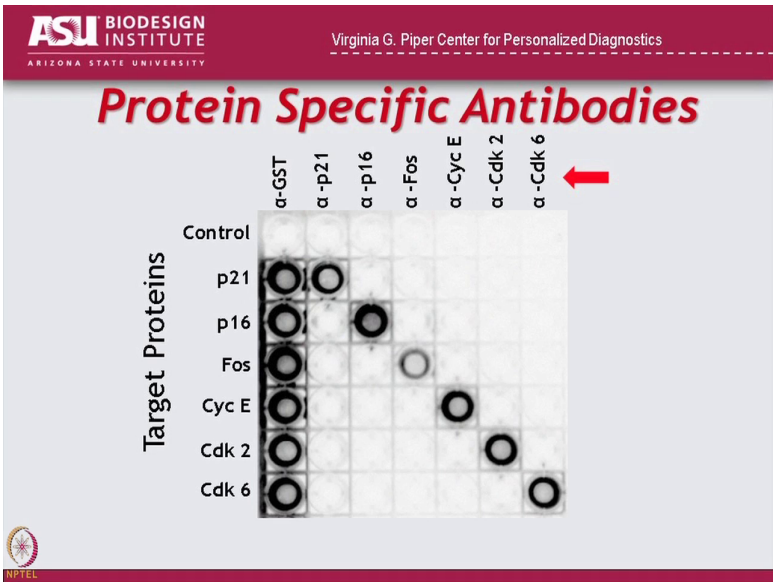
(Refer Slide Time: 11:00)



And so, the idea of NAPPA is that we can do all kinds of studies on it. We can do in our actions with specific protein queries, we can do an enzyme substrate modification, we can even build multi protein complexes.

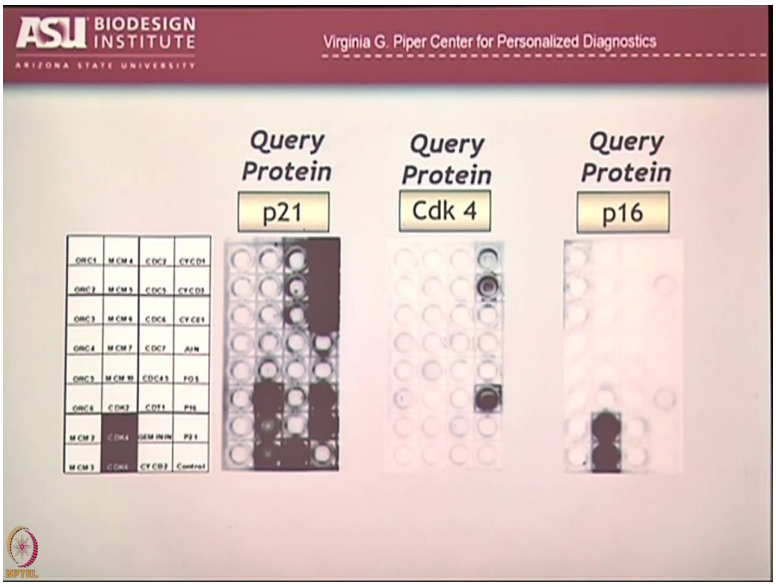
So, let me begin at the beginning by showing you the idea behind NAPPA, before it was actually on a protein array.

(Refer Slide Time: 11:20)



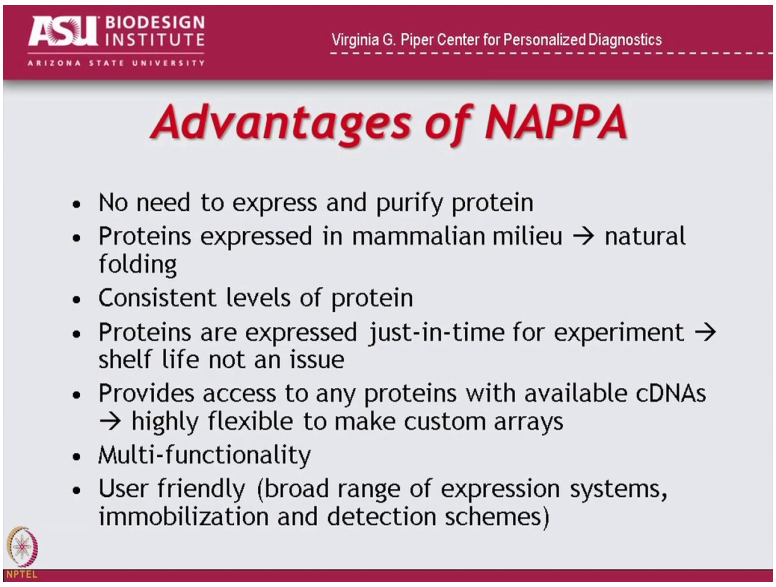
So, here what we are looking at is just making proteins using cell free extract in the wells of a 96 well dish. So, the proteins that we make all have a GST tag at the C terminus. Remember we talked about you always have to have a tag. So, these guys all have a tag. And so, if we if we make the protein in these wells and we probe them with an antibody that recognizes the tag, they all light up right. But if we probe them with antibodies that are specific to each individual protein, then only the p 21 lights up at p 21 only; the 16 lights up at 16 only the foss and so on and so forth. So, depending on what you use as your antibody you only get that signal.

(Refer Slide Time: 12:14)



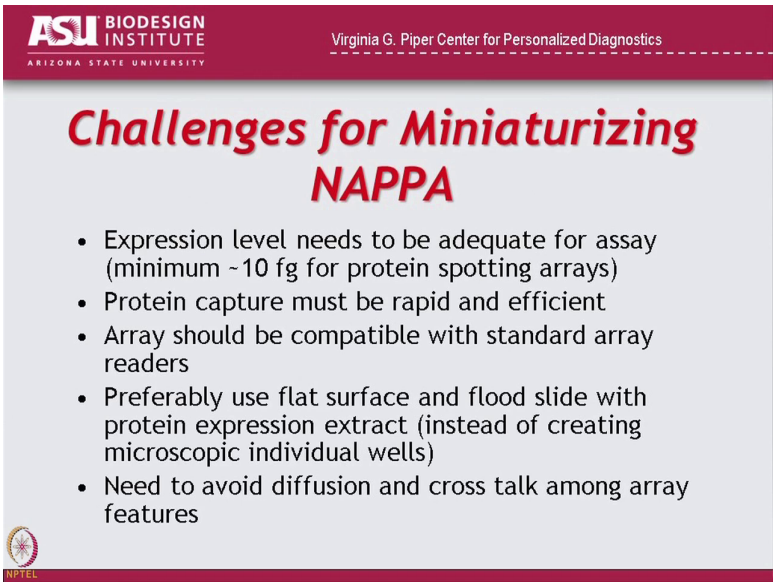
Now, you can use the same approach to look at interaction queries. In this case, here is a bunch of different proteins on a chip right and in chip, I mean in quotes because these are actually the wells of a 96 well dish. If I probe with the protein p 21, it binds to all the Cdk's. It binds to these cyclones; it binds to these guys here and it binds actually to itself again. So, those are the interactives repeat 21, if you probe with a different query, in this case Cdk 4, its going to bind to the cyclin ds and its going to bind to p 16 and if I probe with p 16, it is going to its going to pick up these Cdk's. So, you can see using different queries; I get different interactors and again, this is not a microscopic array, this is a nice as well dish well of course, the advantage of this is.

(Refer Slide Time: 13:07)



Now, I do not have to express these proteins or purify them. They were made using mammalian extracts. The levels of proteins were consistent on this chip; they were made at the time of the experiment. So, I made the proteins and I tested them minutes later. And of course, I can do this general approach using any kind of cDNA. If I can clone the gene and make the cDNA, I can make the protein array. And then, we talked about the multi functionality and of course, users can modify this as they need to, but there is still some challenges

(Refer Slide Time: 13:43)

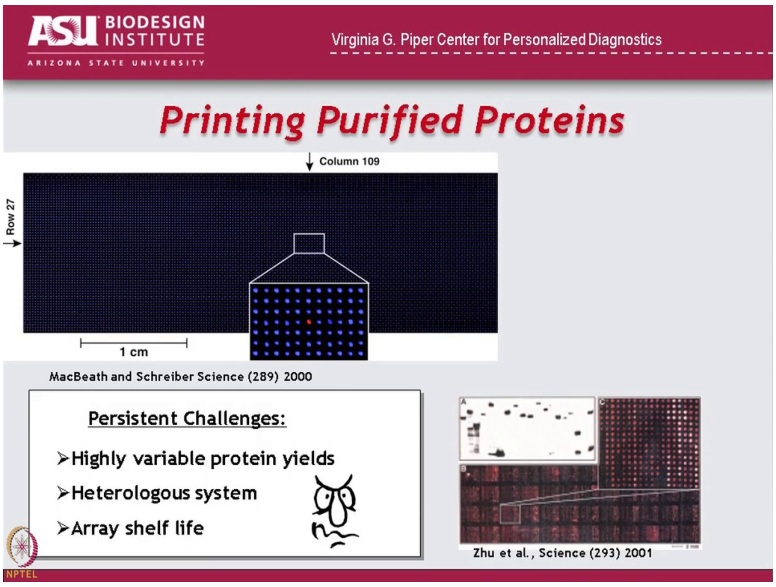


. So, we need to make we need to now take it from this format which is on 96 well dishes and we had to get it onto a microscopic chip, we had to be able to print it on slides so that we could do thousands at a time.

So, we had to be able to deal with very low amounts of protein, we need to make at least ten m two grams of protein. We had to get capture that was rapid; we had to find arrays that were compatible with standard array readers. At the time, we really wanted to work on a single slide format that that where we could add the extract of the whole slide. We did not want to have to manufacture these specialized methods for expressing in little tiny wells or something like that. I am going to come back to that because we now are moving in that direction, but at the time we really wanted this to be very simple and of course, we had to avoid cross talk from spot to spot.

(Refer Slide Time: 14:38)

(Refer Slide Time: 14:41)



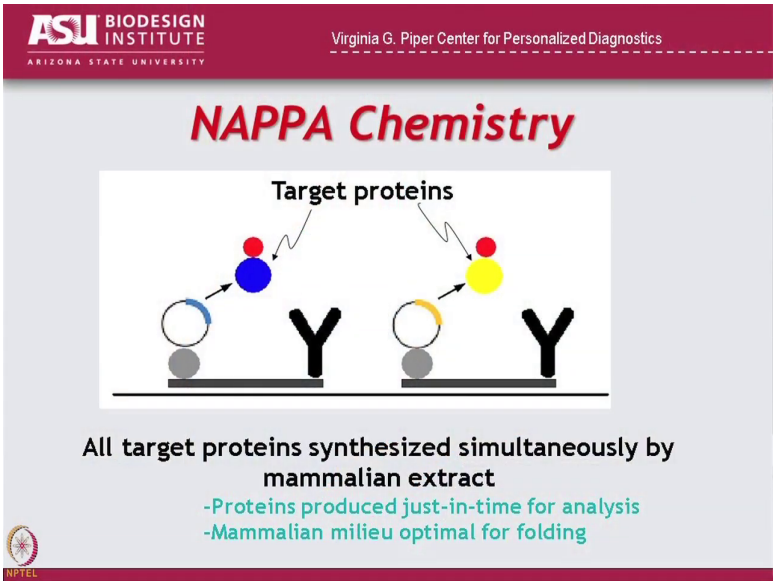
So, let me talk about the technology. So, most people who built protein arrays do it by purifying the proteins first. They do we started with which was this high throughput purification kind of technology and they do it in 96 well plates or even 384 well plates ah, but they have a couple of problems.

First of all, you get very highly variable protein yields. So, the amount of protein that you get from one protein to the next can change over 4 logarithms. So, 4 orders of magnitude of difference; they get they working typically in heterologous system. So, they are either purifying in bacteria or they are purifying from yeast or they are purifying from insect cells and that in and of itself introduces some differences.

And of course, the biggest concern I have is array shelf life. So, you purify the protein, you store the protein, then you take the protein out, then you print the protein and then, you store

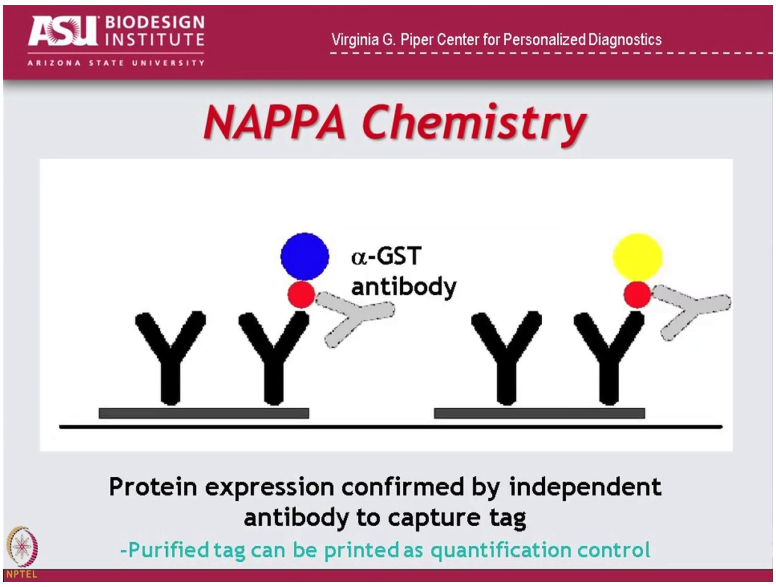
the printed protein. So, you have a lot of steps in there and all those are opportunities for proteins to lose their shape lose their folding and not be as functional. And of course, some proteins will stay fine during all that process and other proteins will not and you never know which ones they are. There is no way to tell which ones are the good ones and the bad ones.

(Refer Slide Time: 16:09)



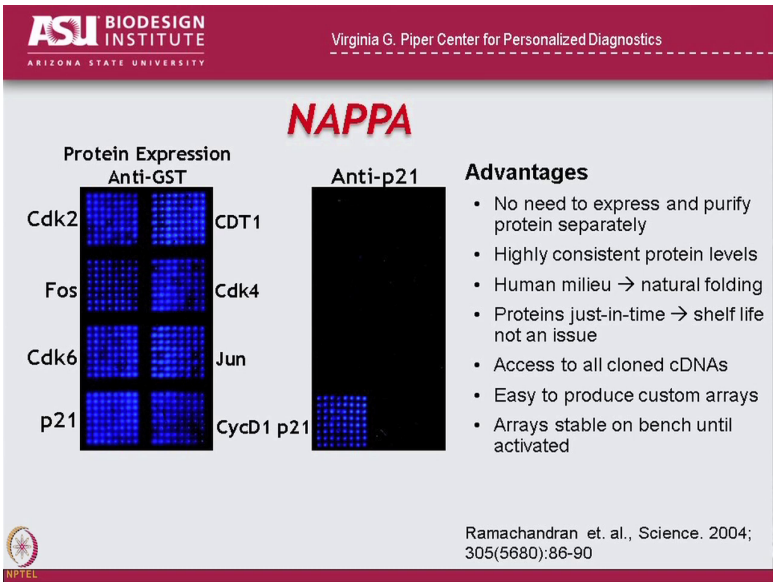
So, this is the idea behind NAPPA. In the in the case of NAPPA what we do is we print the gene on a plasmid, we add cell free extract that makes the protein and this is meant to show that we have the protein in blue and the GST tag in red and then, here is a different protein in yellow and a GST tag in red.

(Refer Slide Time: 16:31)



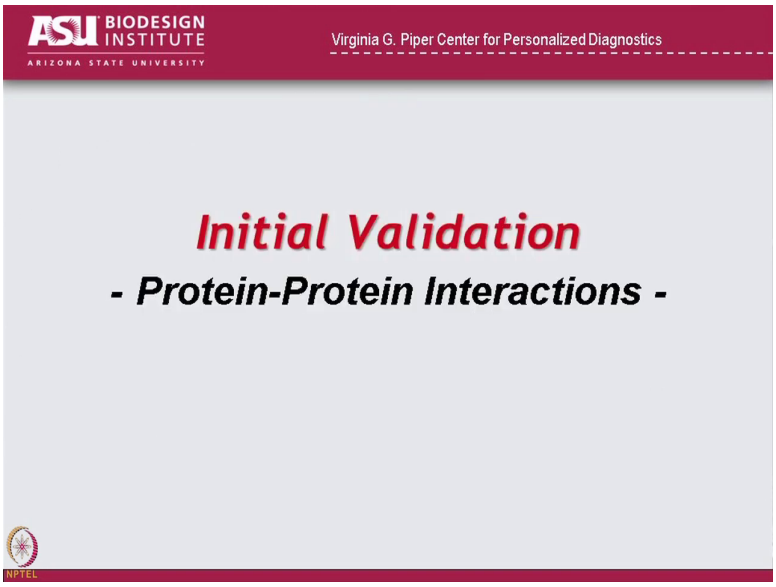
So, we make this at the time of the experiment and then, what is going to happen is yeah, I just told you that is that the GST is going to get captured by the entire GST antibody and now you are displaying the protein on the surface. So, it flips upside down. So, that the protein part is what facing up; makes sense.

(Refer Slide Time: 16:47)



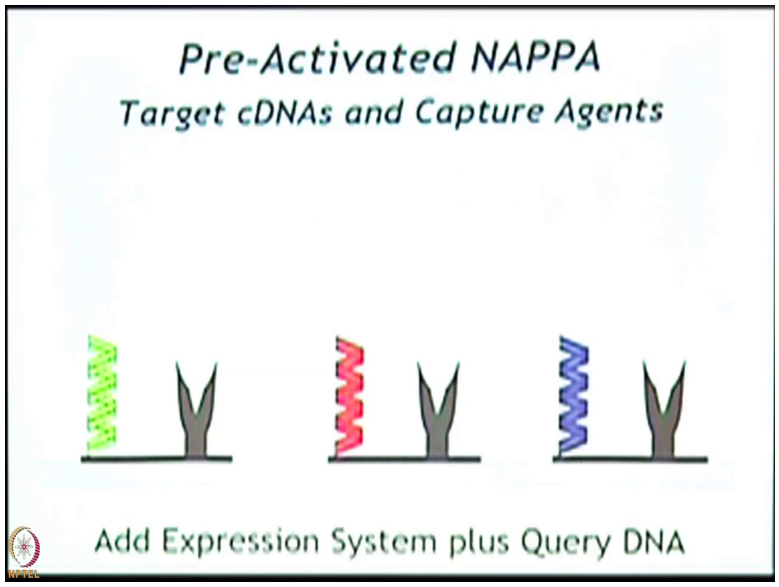
And then, this is what it looks like here is an early array printed 8 different protein 64 times each and here, we probe with anti GST that is that is a way of measuring how much total protein, we have on the array and so, all of them wide up and then if we probe it with anti p 21. Just remember the little 96 well plate, same idea. Now, we only get the p 21 lighting up alright. And I already kind of covered these inner these advantages pretty much the same ones here, so I won't go over them again alright.

(Refer Slide Time: 17:19)



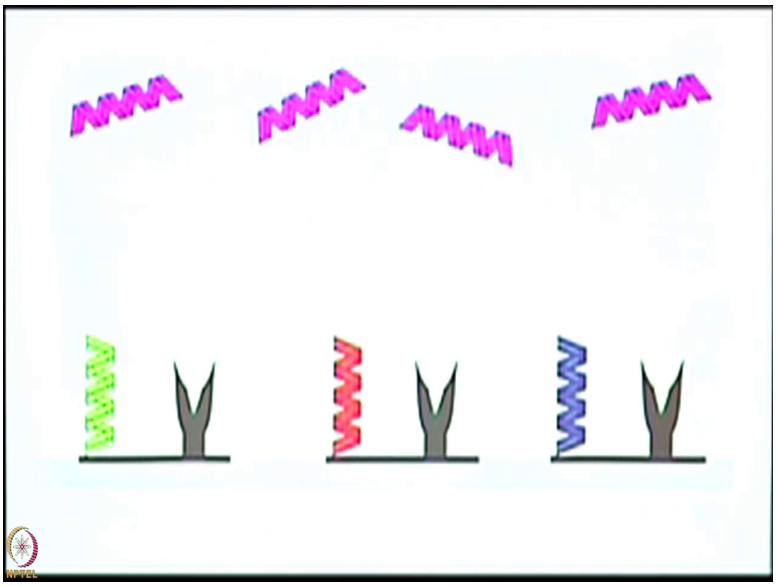
So, here is how we first tested this. We decide decided to do Protein-Protein Interactions.

(Refer Slide Time: 17:27)



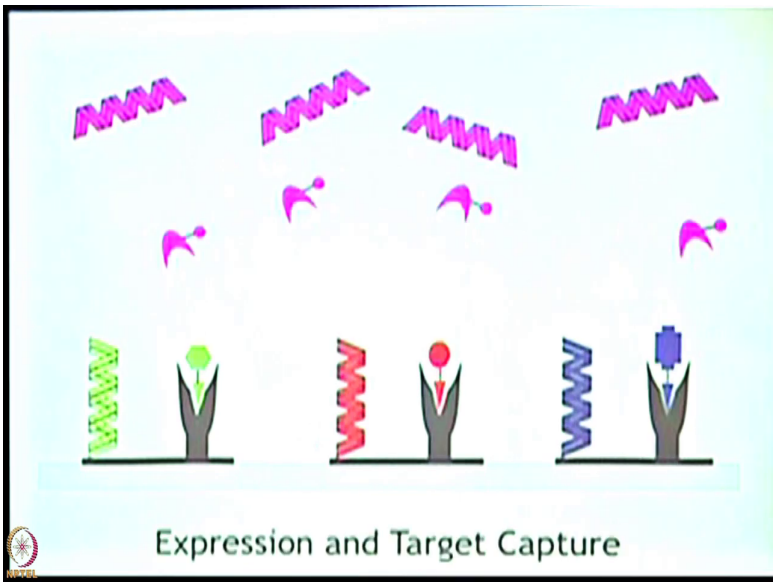
So, imagine that you have three different spots on your array. In fact, we had we had many more than that, but let us just talk about three; this one makes the yellow protein, the red protein and the blue protein and these are the jeans these are DNA and then here is the antibody that is going to capture the GST.

(Refer Slide Time: 17:49)



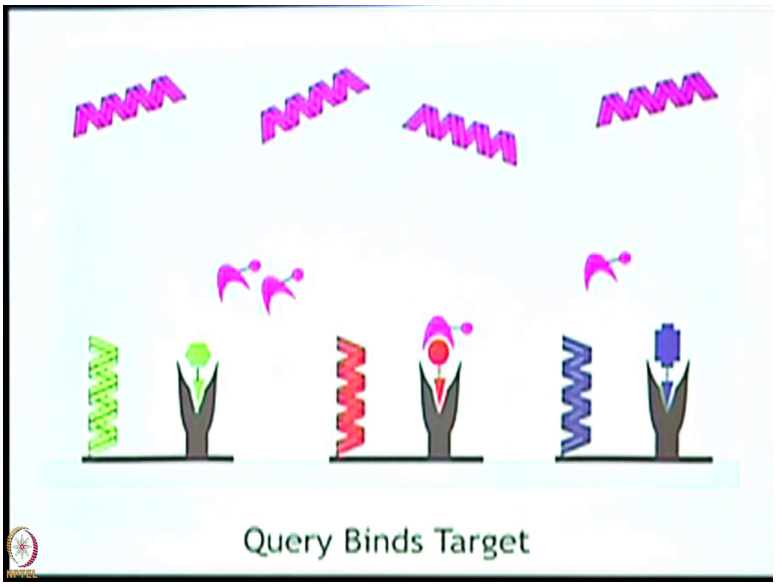
Now, we want to we want to ask do any of these three proteins interact with the pink protein alright? And the case in this case what we are doing is we are going to add the gene for the pink protein in the solution along with these guys which are attached to the surface. So, these guys are bound to the spot, this is free to go anywhere it wants. We then express the protein right.

(Refer Slide Time: 18:13)



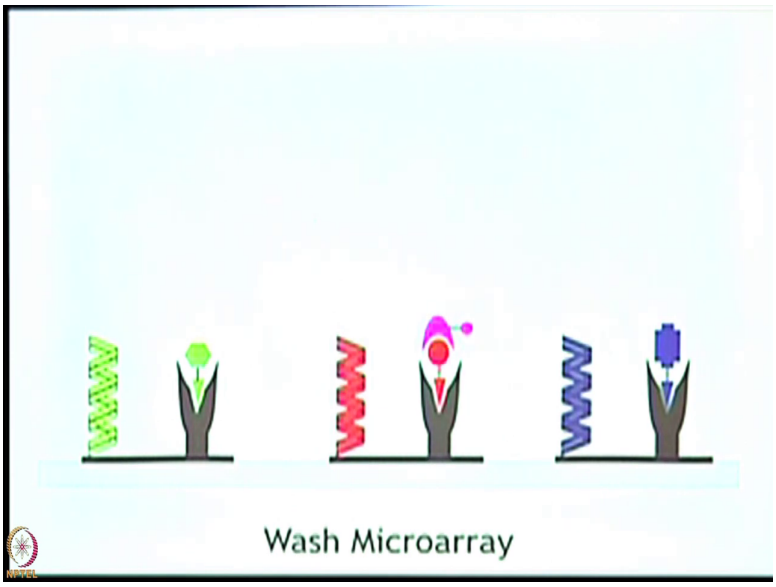
And these three proteins are going to get captured to the surface of the array because they have the GST tag. So, that is going to lock them down to their spot. But this does not have a GST tag. So, it is going to float everywhere on the across the array.

(Refer Slide Time: 18:32)



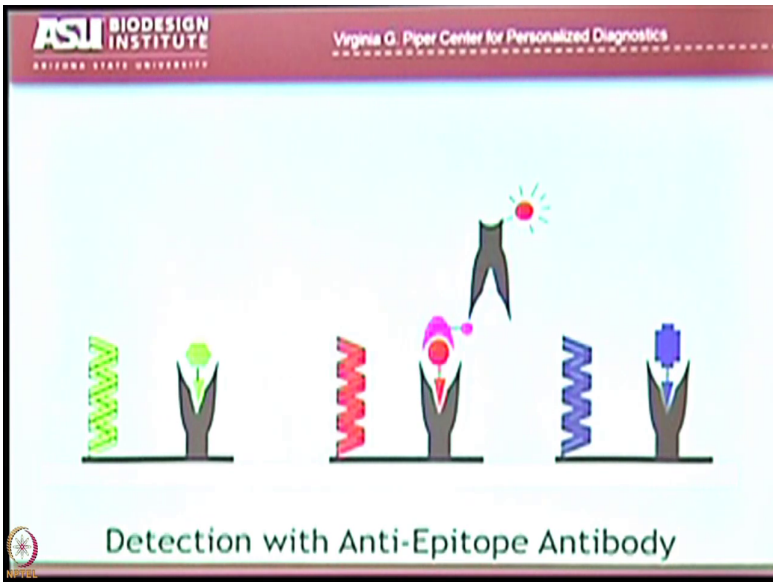
And then, over time if you give a time, the query protein will bind to the target proteins if it recognizes them. In this case it binds to these guys, but not these guys right.

(Refer Slide Time: 18:45)



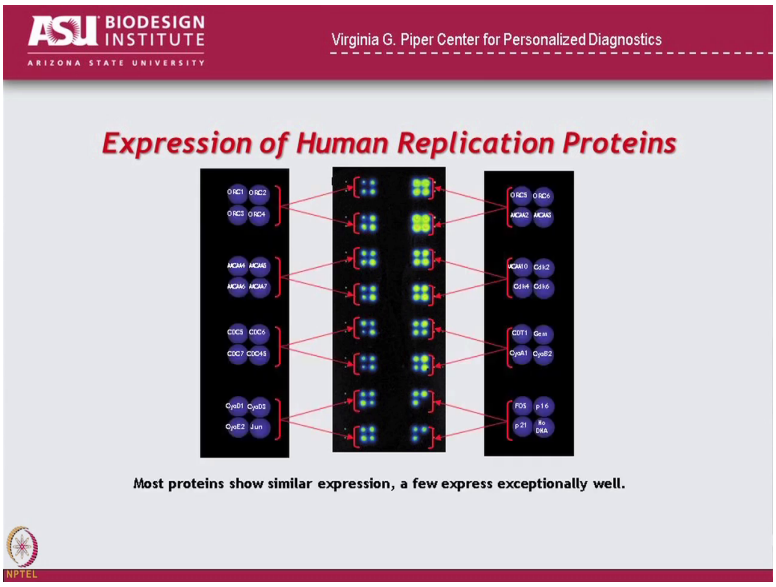
And so, now, I can wash away anything that is not bound and I am left now with this guy bound to this spot. Now, how do I know where it binds? Well, I know the identity of every spot on the array. I know whatever position it is which gene it is and so, I know that if this spot lights up that pink binds to red right.

(Refer Slide Time: 19:05)



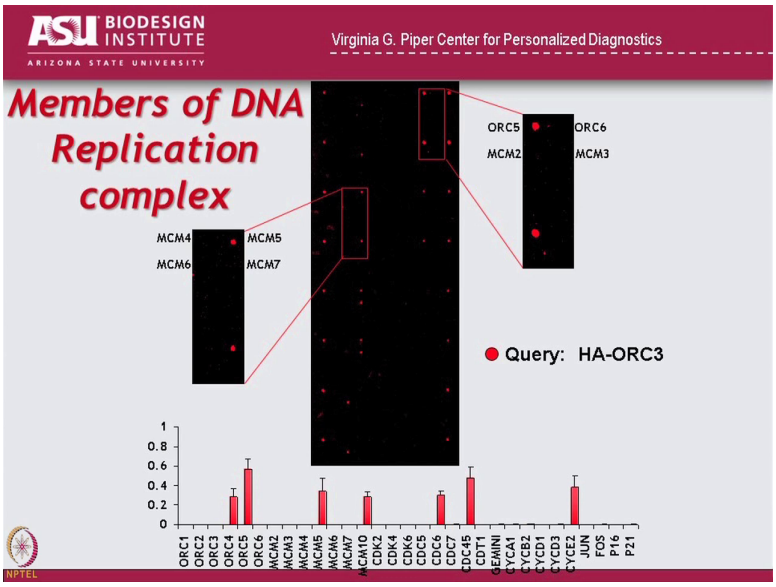
And I can detect that interaction with an antibody that has a fluorescent marker on it that we will recognize either a tag on the pink protein or it can recognize the pink protein itself. There is all kinds of variations you can do here ah, you could use click chemistry to look at interactions and we have done that, you can have you know other molecules that interact with this guy avidin and biotin lots of different strategies. But the bottom line is as long as you can recognize the query protein; you can determine where the binding occurred.

(Refer Slide Time: 19:41)



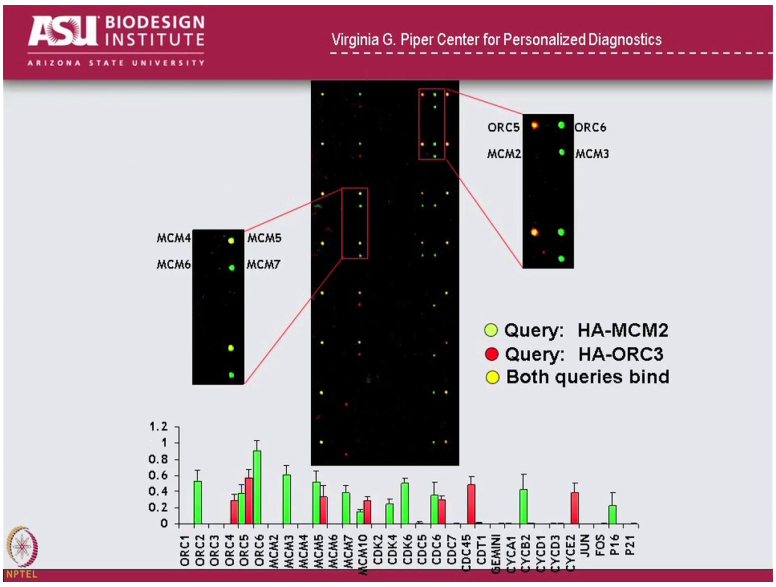
So, our first experiment was this guy we took all the proteins in the human DNA replication complex, these are the this is that collection there. Cloned all those genes and then, we printed them and express them and this is measuring with GST just to show that they all got made and everybody was done in duplicate. So, they were all there in two spots.

(Refer Slide Time: 20:03)



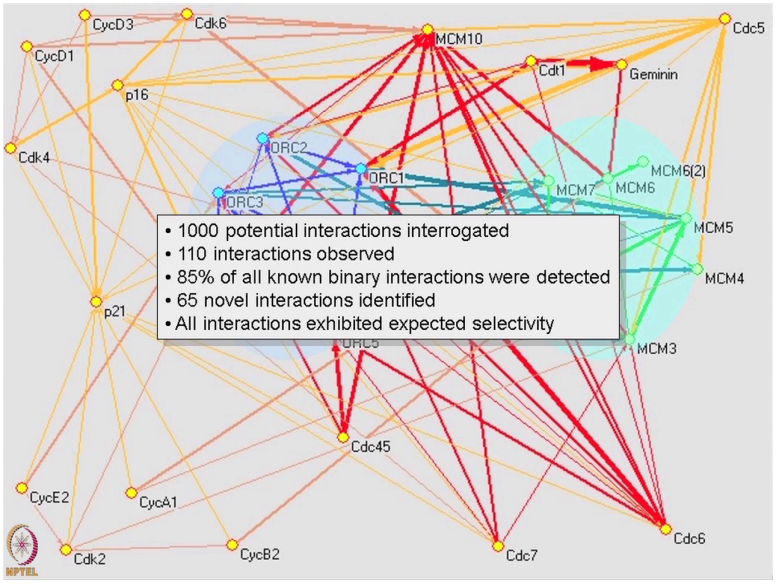
Then, we query the array using an antibody using this protein here, one of the proteins in the set MCM2 and then and then, you can see that MCM2 binds to ORC5, ORC6 and MCM3. And here, it is binding to these guys over here and of course, it does not in duplicate, so we know it is. We were confident of the result. You can do the same thing with a different protein. This is ORC three and again, its binding to certain proteins, but not other proteins. So, you have every other protein on the array is sort of a negative control right.

(Refer Slide Time: 20:37)



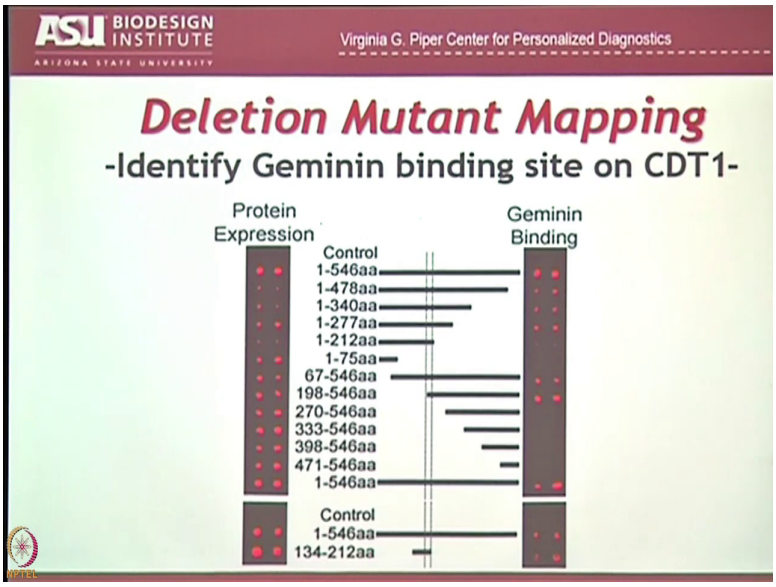
And you can merge the two images if you wanted to and even build, use that kind of thing to build an interaction map.

(Refer Slide Time: 20:45)



For all the proteins in the complex with all the other proteins and that is effectively what we did we identified. We queried over a 1000 possible interactions, we identified 110 of them including many new ones, using this general approach. So, you can use this to kind of look at Protein-Protein or actions.

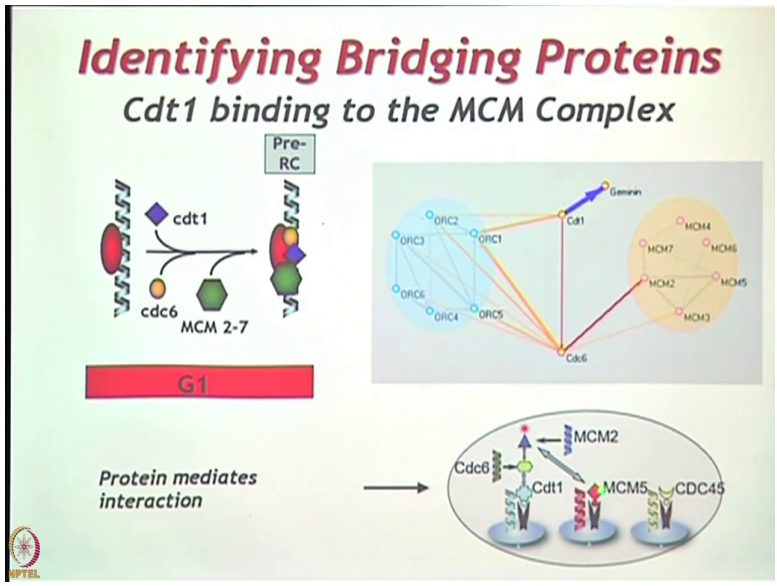
(Refer Slide Time: 21:05)



And of course, you are not restricted to looking at full length proteins. If you want to map the binding domain of specific parts of proteins, you can do that. So, in this case, we were looking at where does this protein geminin bind to this protein CDT1.

So, we took CDT1 and we made a series of different deletions right and we showed that all of them were expressed on the array and then, we probed them with geminin which interacts with them and you can see that geminin binds to some of them, but it does not bind to others. So, that gives you some sense of where the binding site is right, this line here, if that if this part of the protein was present, then it always bound. So, that map quickly where two proteins may talk to each other. And then, neuro impact and made a very small version of this guy and showed that it was sufficient for binding.

(Refer Slide Time: 22:07)

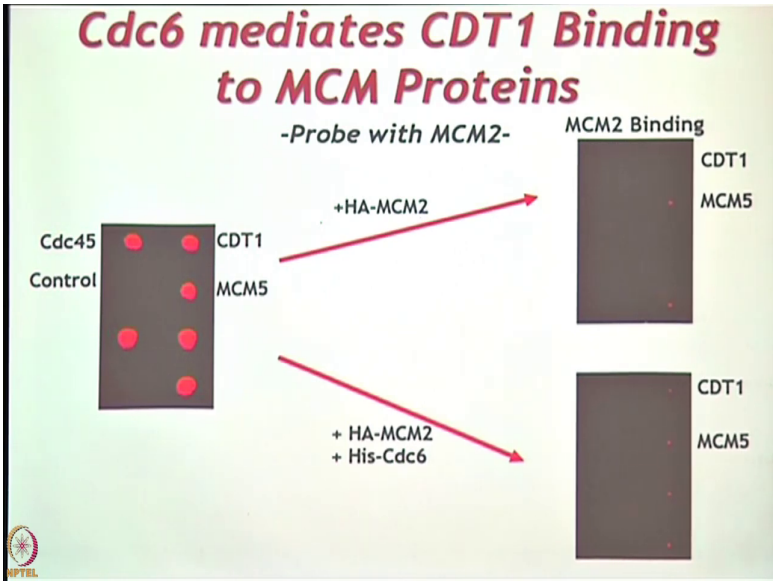


The other thing that you can do if you want to play with these arrays is you can actually do that you can look at the possibility of multiple proteins interacting together. So, we knew that CDT1 bound to the MCM complex, we could tell that by looking because of biochemical studies that have been done before we got involved. But we did not see CDT1 directly interacting with any of the proteins over here. This is a map that came from that big map I showed you and what we did what we figured out was that although CDT did not bind to any of these proteins. It did bind to this protein and this protein bound to that protein.

So, maybe this protein here is acting like a bridge protein. It is holding CDT1 in connection with that complex right. So, the question was could we test that on the arrays right and we did that by doing a couple of things. We could probe MCM 2; we knew was in this complex. So,

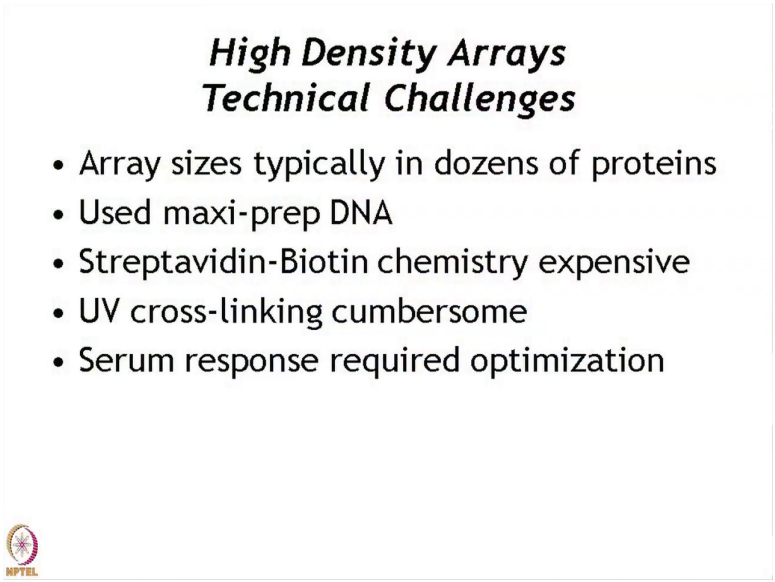
we probed MCM 2 against CDT1 either with CDC6 or without it and we as the control had MCM 5 and we also had a negative control CDC45 and.

(Refer Slide Time: 23:29)



And this is just to show you what that looks like. This just shows you that all the proteins were made on the array. If we add MCM 2 without CDC6, you do not see any binding here; but you do see the positive control MCM 5 indicated. If you add MCM 2 plus CDC6, now you can see the CDT1 binding. It is pretty faint. I am not sure, you can see it where you are, but we definitely observed it ok.

(Refer Slide Time: 23:55)



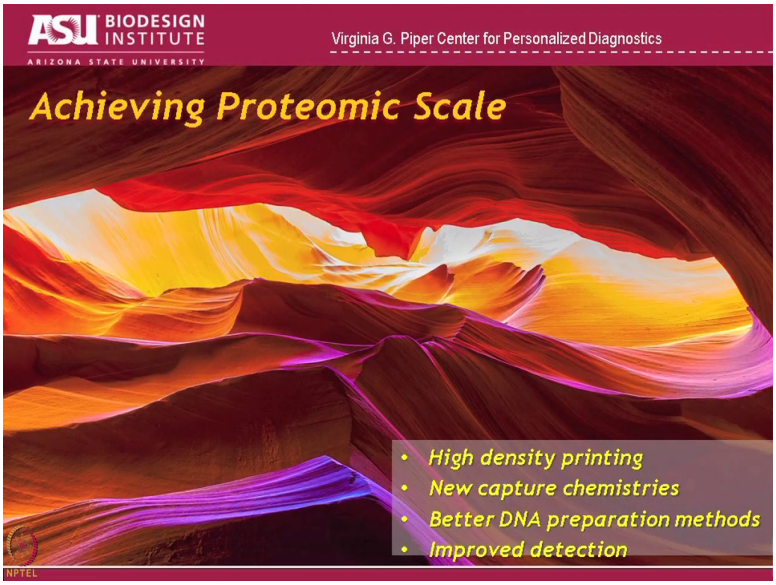
So, we are going to spend more time later in the course talking about the high density arrays, but I want to give you a flavor of what we had to do to get. Now, from these arrays which I showed you showed a hat around you know 50 proteins on them. The goal of course, I told you from the beginning was to get to thousands. So, how do we adapt a platform to get to thousands.

So, yeah; so, we were working in sort of dozens range, we were we at those days we were using maxi prep DNA. If you are going to do thousands of proteins and you remember that what we need to make is DNA, not protein. And that is advantageous because it is easier to make DNA than protein easier to purify it and much easier to quantify how much you have made. But still as easy as it is to make DNA, if you want to do an array of 10000 proteins, then you have to prepare 10000 DNA's.

So, you need to you can't do that by Maxi-prep. You have to be working at small scale that is reproducible and reliable. Yeah, we were using this Streptavidin-Biotin Chemistry to do our arrays at the beginning and that clearly was not going to work because its expensive and it involved having to stick them in the UV box for a while it was really kind of a pain in the butt. Yeah, the UV cross linking part was a cumbersome and we needed to do some optimization and we needed to increase our content.

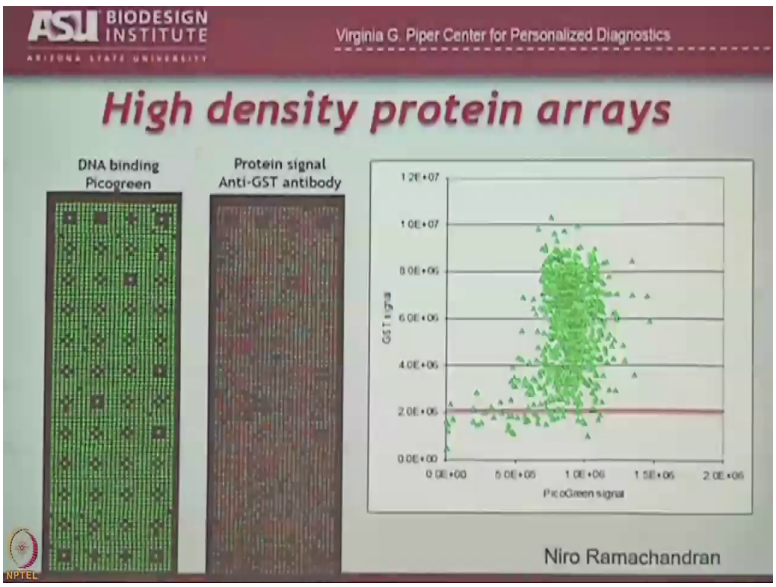
So, I am going to fast forward a little bit. I am not going to go through all of the optimization experiments that we are done. You are going to learn how to make these arrays using our current approach.

(Refer Slide Time: 25:41)



But we spent a lot of time working on the high density printing, developing a better capture chemistry figuring out how to make DNA and high throughput and figuring out how to detect the interactions in a in a more ready and easy way.

(Refer Slide Time: 25:53)



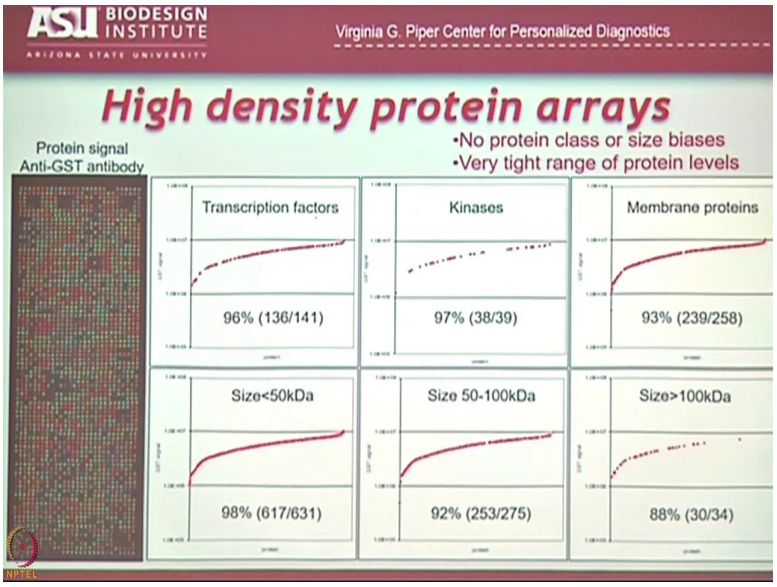
And this is sort of the result of all that work, what you are looking at is kind of our current version of what we do most of the time for NAPPA. Now, on the left is one of our typical arrays around 2300 genes or proteins printed on the array. The signal, you are looking at here is the DNA signal. So, we stain the chip. Every time we make one for DNA and that tells us that our printing was good right because remember what we are printing is DNA. So, if we stain for peacock with Pico green and we see even staining; that means, that we did a good job of printing even amounts of DNA.

Then we convert it to protein and we measure with anti GST and that tells us how much protein we have on the array and that tells us that not only to be print well, but we can express and capture the protein as well alright. And that is what this plot shows you here in the x axis is DNA signal that is the Pico green; in the y axis is GST signal that is the what protein made.

A couple of things that you can observe here: first of all in terms of the x axis the vast majority of our spots line up very evenly around this area here. That means, that we do a pretty good job, printing the same amount of DNA for every spot and that is encouraging it means our printing is good. You will see that there is a few down here that did not print well and so, it is good to know that. Secondly, if you look at that from the top of the expression to the bottom of the expression, the entire range is within one order of magnitude.

So, instead of those protein arrays where you would have the lowest amount of protein to the highest amount of protein being over 4 orders of magnitude. Now, all of our protein is even in a single order of magnitude. In fact, 93 percent of these spots are within two fold of the mean which means that we are getting very even levels of protein on the chip and that is exactly remember that one of the things, we wanted from an ideal array was to have very consistent even levels of protein.

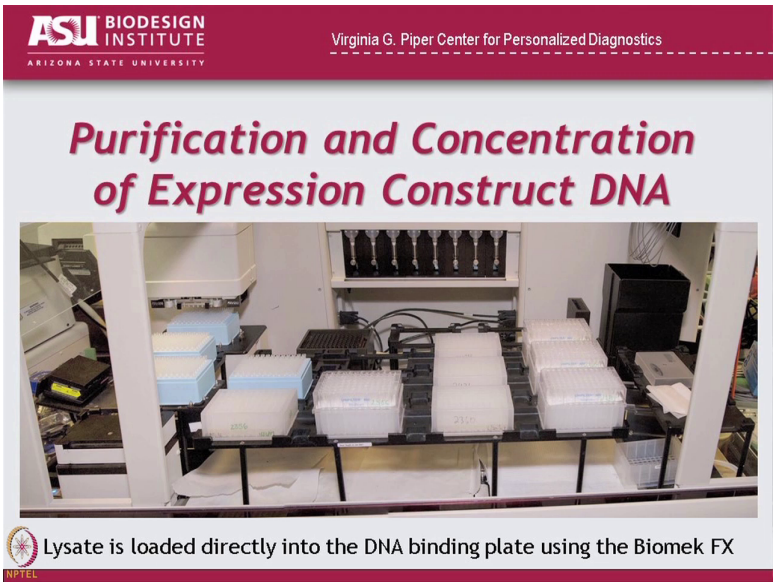
(Refer Slide Time: 27:57)



And just to give you some better sense of that here is quantification based on different types of proteins going from the weakest to strongest and then, this line here represents the lower end of detection and this is the higher energy detection. Ah No, that is not true here. I am sorry. the Its right here, I do not have that on this graph. This is just to signal intensity. But you can see that we get by a large 96 percent of transcription factors were detectable, 97 percent of kinases membrane proteins are very detectable and then, small medium and large proteins are all detectable.

So, roughly speaking about 97 percent or whatever we print we can get good expression of. Occasionally, we will run into proteins that have unusual amino acid sequences that make it difficult to get high yield, but that is by far the exception.

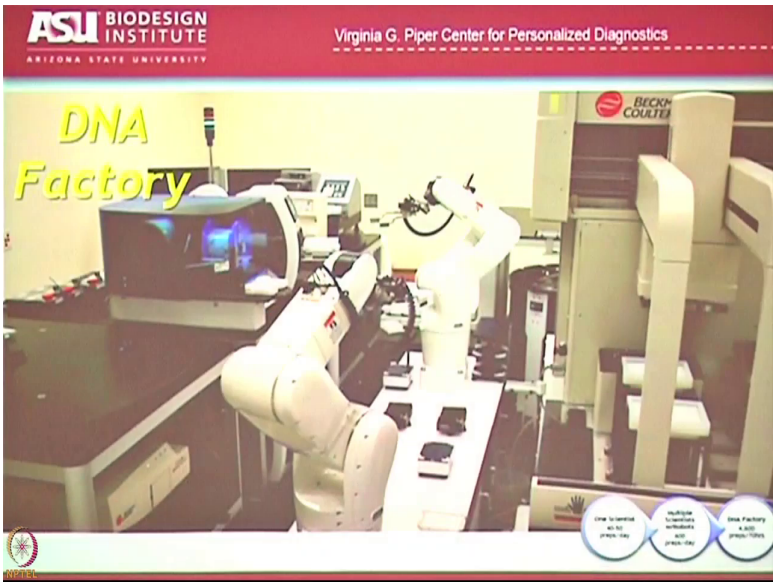
(Refer Slide Time: 28:53)



So, this is the method that we use to purify the DNA or and that I should say the method that we used to use those of you who know Sanjeeva, this is when he was in the lab this is how he did it. This is an automated platform we had worked out using robot, how did you do DNA mini preps using robots and that allowed us to do if you were really working hard about 600 a day in a sort of a team approach. That that was not easy, but you could do it just to give you some frame of reference when I did DNA mini preps in my day. If you did 50 in a day, you were working your butt off ah; but with robots you could get up to around 600 ok.

Since then, in Arizona, now we have this technology.

(Refer Slide Time: 29:43)



So, we have taken that robot technology which was we used to call it sneaker net which is you connect one robot to the other by a graduate student, who runs from one to the other to an a fully automated platform and let me see if I can make this go did that is that going yeah. So, in this is what we have in Arizona now, in the basement where my lab is located. This is a fully automated platform for growing DNA and purifying it.

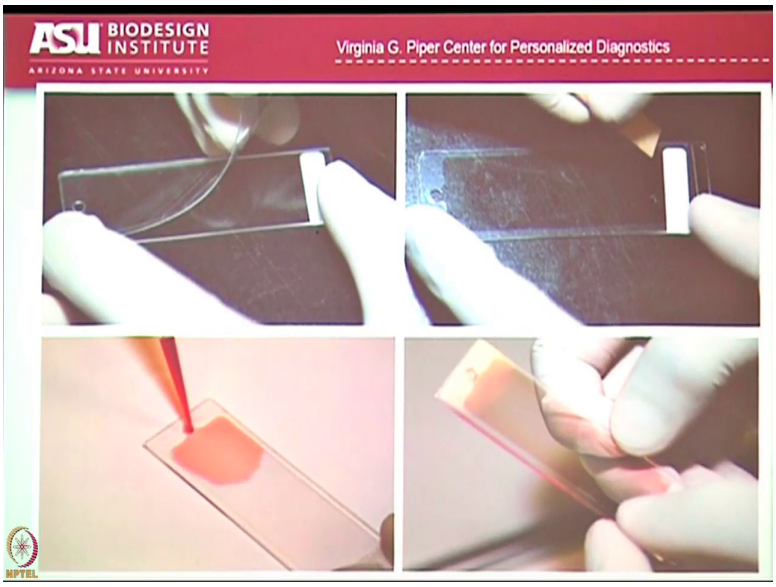
The camera is sitting on an incubator shaker that grows the bacteria with the DNA in it. This particular shaker has a specialized door in the back. So, that these robot arms here can reach in and take out each one at a time as it needs to and they can put they pass it back and forth on that platform there. So, hey this guy is handing it off to this guy. So, he is going to pick it up now, but they can pass it back to the centrifuge which is right there ah. There is also a freezer which is right over here that stores the pellets after they have been grown and then, this liquid

robot over here will purify the DNA from the bacterial pellets. You can see this guys mixing by turning it upside down, just the way you would invert a plate right.

So, the robots can do that for you if you want them to and of course, we have sealers and peelers to seal the plates and they do not spill. Barcode reader, there is a barcode reader right down here which you can see that checks the barcode of every plate to make sure that it is what is supposed to be and then, that this device over here will read the optical density of the DNA after you have made the DNA the OD260.

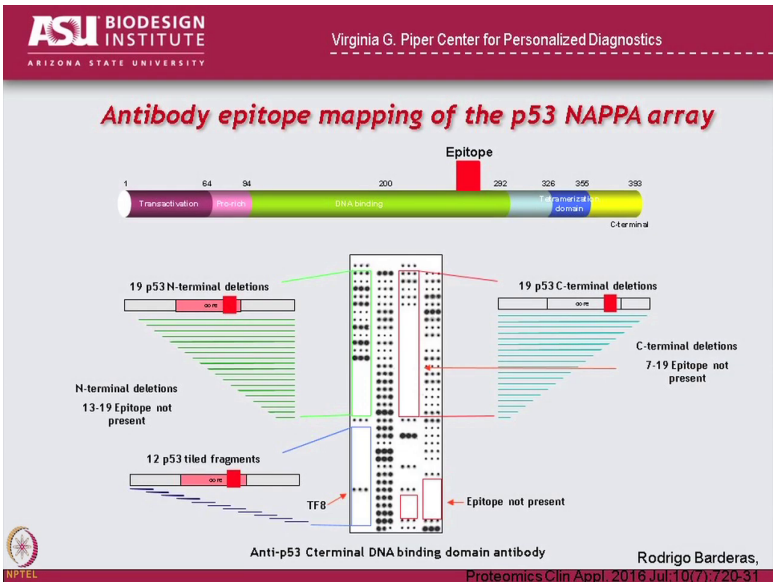
So, we can actually measure how much DNA, we are making and we can the robots will automatically adjust the concentration to make them what you want them to be. So, all in all we went from doing 50 mini preps a day 600 mini preps a day to now 4600 mini preps in 70 hours start to finish growing bacteria to getting purified DNA. So, really accelerates what you do and it also gives you a little bit more certainty that you were what you are doing is working.

(Refer Slide Time: 31:51)



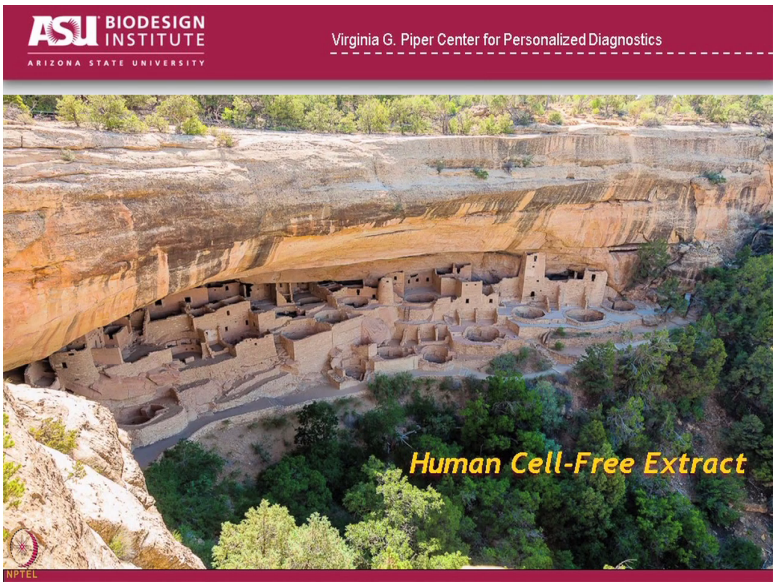
So, this is how we actually make the chips and I think you guys have examples of this at your desks ah. You have this coverslip here that you put on the chip, you there is a little hole here and a little hole here. You inject your lysate in here, you fill it all up. Make sure you do not get any bubbles that is always the trick one of the things you have to learn how to do and then, you can make the proteins on the chip using this approach.

(Refer Slide Time: 32:17)



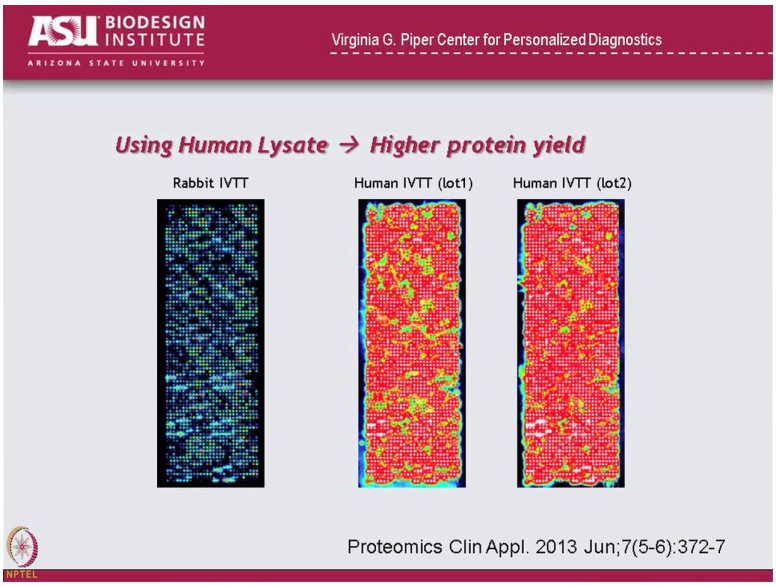
This just indicates that you can map a more high a higher throughput version of mapping, we are binding occurs. Here is an example of an antibody that was binding to the p 53 protein. We did a series of in terminal deletions and you can see it binds to all of them until it gets to here. We can see terminal deletions and again binds to here, then you do not see it and we did fragments that walk across the protein and of course, it binds to just that one there.

(Refer Slide Time: 32:47)



So, there is been a couple of advances that that have occurred in the last I would say several years, when we first started this work we were using reticulocyte lysate from rabbits to make proteins. That was the expression lysate that we use to make proteins and I will say that still works ok.

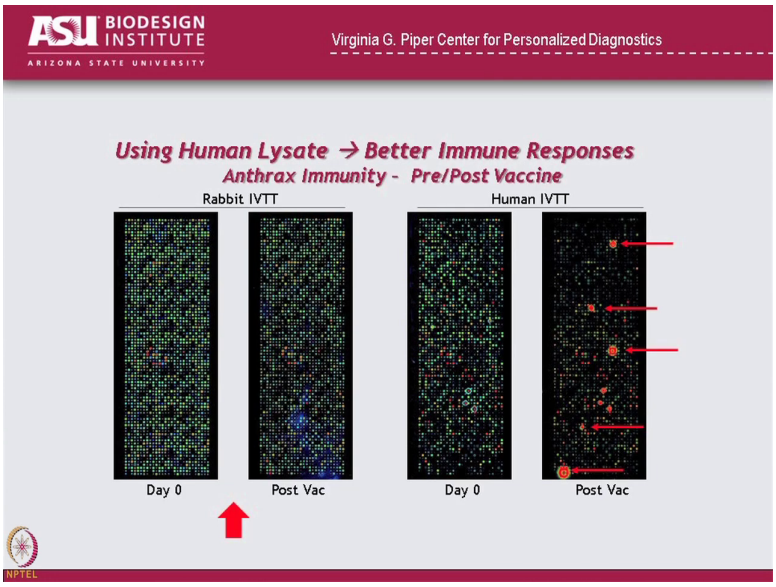
(Refer Slide Time: 33:07)



This is what that looks like. So, you have been seeing a rays look like this now most of the day; blue is expression, green is better expression, orange is really good expression, red is like amazing expression ok. So, that gives you some flavor of the thing and we were quite happy with this ah. But then, the patent ended on that and new companies came out and a new version of lysate came out that was made from human cells, the advantage of the human lysate was that it came from a purified cell line.

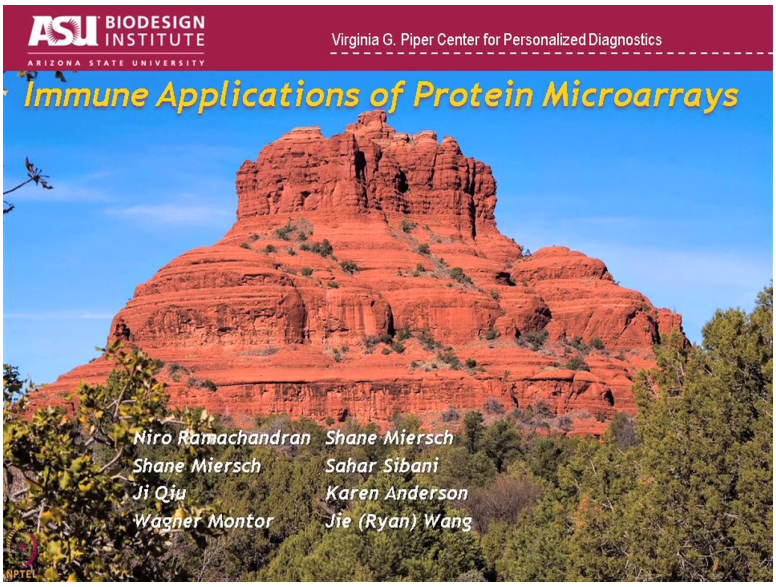
So, you did not have to get. So, much variation from animal to animal and this is what the signal looked like. Just unbelievable signal about 15 times stronger then we could get from the rabbit lysate. This is whole human extract that includes human ribosomes and human chaperone proteins which means that there are proteins in the lysate that help these proteins fold in their natural in their natural shape.

(Refer Slide Time: 34:09)



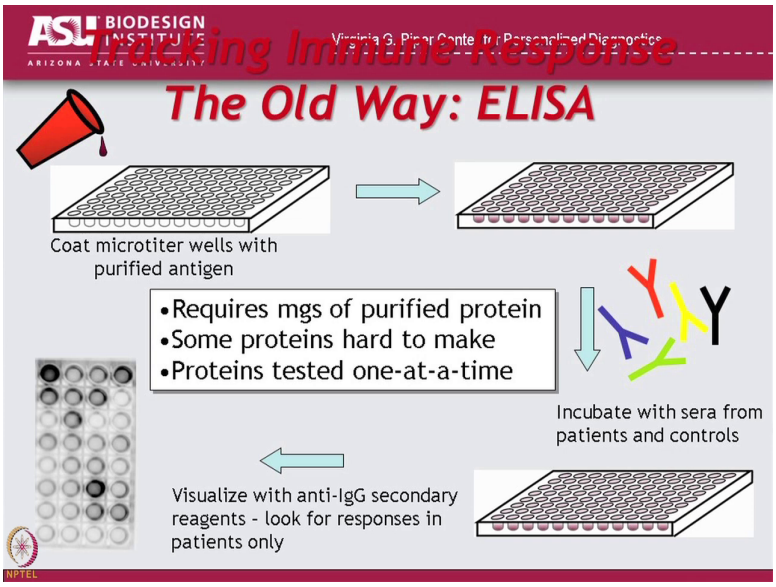
The other advantage of the human lysate is that its less likely to inhibit immune reactions. So, one of the problems that we used to see with rabbit reticulocyte was every once in a while when we were looking at this is a person who was vaccinated with an antibody to these put to this anthrax, that even after vaccine we saw no response on the chips. If we made the same chip with human lysate, we saw very good responses. Probably, what is happening is in the human lysate, I mean in the rabbit lysate there are because it is from blood there may be inhibitors of immune response there and those were blocking it on the chips, but the human lysate comes from a purified cell line, there is no immune system around and so, you do not get that problem.

(Refer Slide Time: 34:55)



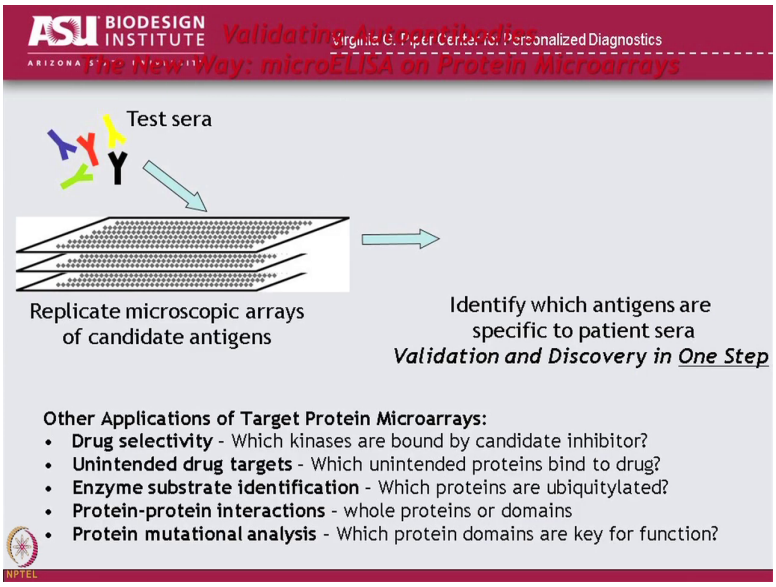
So, I am going to show you just a couple of applications of this approach and then, maybe even just one. So, we have talked about immune responses.

(Refer Slide Time: 35:07)



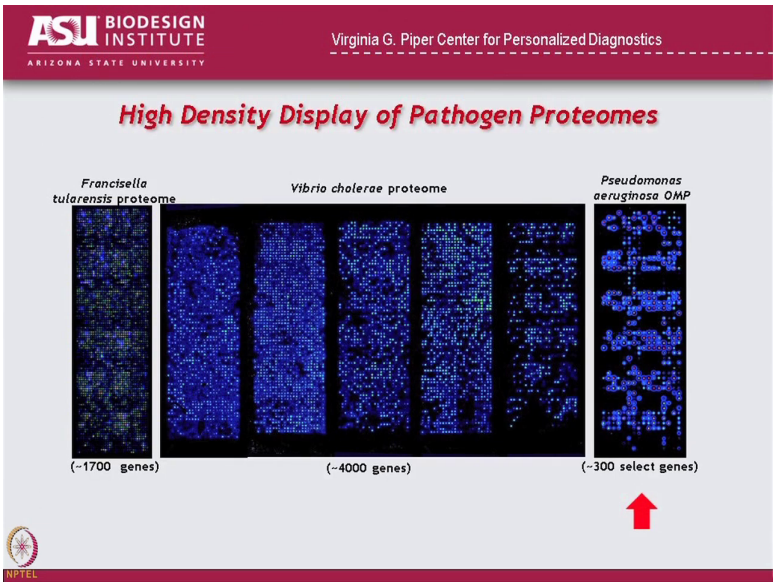
The classic immune responses by ELIZA, where you coat the well of a 96 well dish with your protein, you add serum to the wells and then here is a patient who had a strong response and here is a patient who had no response.

(Refer Slide Time: 35:35)



Of course, if you do ELIZA's you are doing one protein at a time, it typically requires a lot of protein to do that and some proteins are very hard to make to begin with. So, of course, what we would like to do is this. Take a chip that displays thousands of proteins that ok. So, the idea is you probe a chip with serum and then various spots on the chip light up and we have talked about all these different types of assays. So, I won't believe it that.

(Refer Slide Time: 35:51)



So, let me give you one example, we will talk about more as the course goes on. This is a pathogen proteins right and what you are looking at here I remember I mentioned tularemia, when you ask your question that is the entire proteome of Franciselle Tulerensis. In fact, it is the entire proteome and duplicate. So, we got them all into a single array. So, because we in the end we got that to work. These 5 chips are cholera. So, vibrio cholera and then, this chip here is just outer membrane proteins from an organism called Pseudomonas Aeruginosa which is the organism that causes pseudomonas, it is the leading cause of death and patients who have a disease called Cystic fibrosis.

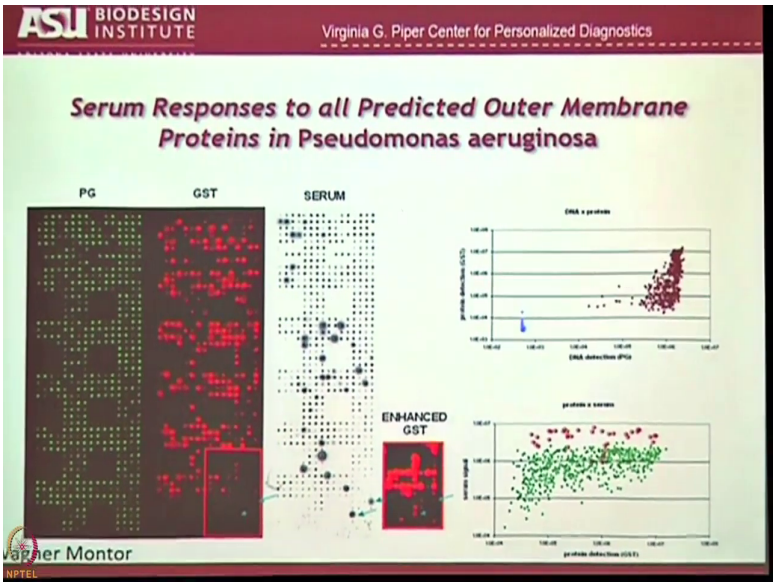
I am going to spend a little bit more time on this guy. So, there is around 300 proteins on there. We were working with a collaborator when we were in Boston, Steve Lawrie was interested in identified developing a vaccine against pseudomonas because this was the

leading cause of death in patients with CF and also a major cause of death of patients in the hospital who are intubated or otherwise immune compromised.

This is an organism by the way that we have all been exposed to. It is in the environment all the time. Most of us if we are healthy do not get infections, but under certain circumstances you get infections. So, he his idea was the proteins on the outer membrane of the bacteria are the most likely to be inducing an immune response and to respond to a vaccine. So, he wanted to look at which of those proteins was immunogenic. His idea was he was going to purify those proteins and then test them.

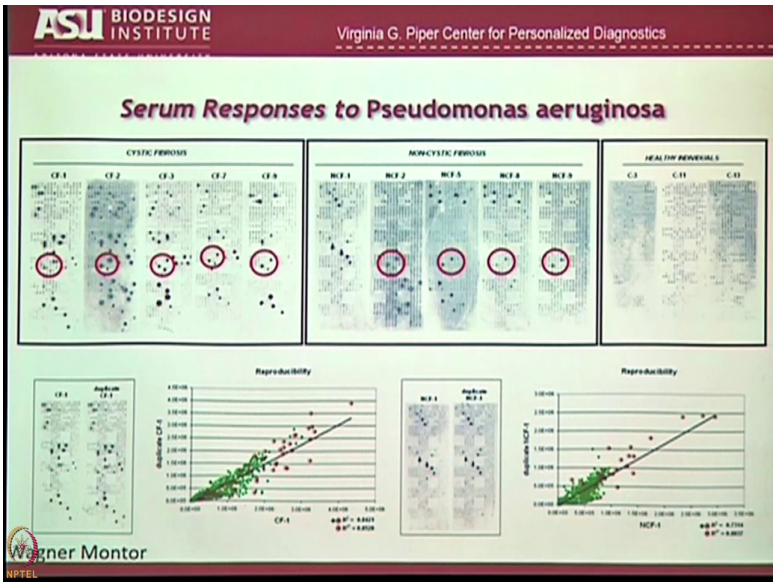
Now, if you ever tried to purify a membrane protein, you know how difficult that is. It is hard to purify proteins in general, but purifying membrane proteins is a is a nightmare and so, he boldly went ahead to clone to purified 300 of these and I think he got 4 all right. So, we suggested try the array because we knew the arrays can make membrane proteins pretty well.

(Refer Slide Time: 37:59)



And in fact, they did. So, here is the chip the DNA stain, here is the chip the protein stain. You notice that they are all red and they are all expressed. So, the membrane proteins were well made on the chip and then we probed them with patient's serum and you can see that this patient is responding to certain proteins on the chip.

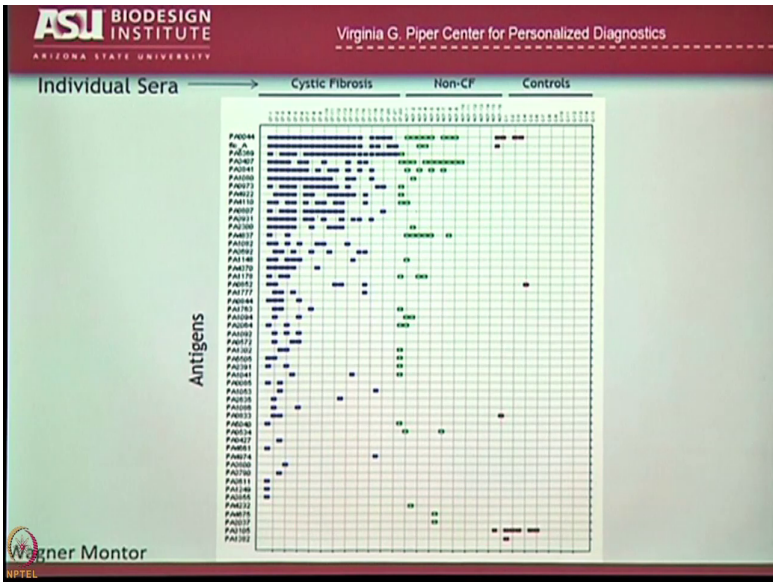
(Refer Slide Time: 38:19)



So, then you can ask the question well are there common responses. Because if you are going to make a vaccine you want it to be one that is common that works for most people yeah. So, here he took a number of patients with cystic fibrosis, who had documented pseudomonas infections. Here is a group of non CF patients, who had who still had documented infections. But just they did not have CF and then, here are healthy controls this is just to show you that the responses were very reliable from chip to chip and if you start looking carefully at this, you will start to see a pattern emerge certain features show up over and over again.

So, that pair there is there, it is there, it is there, it ids there and there and there and there same its those two spots are the same protein everybody's on duplicate. They show up repeatedly that is a sign that particular protein is what we call immunodominant.

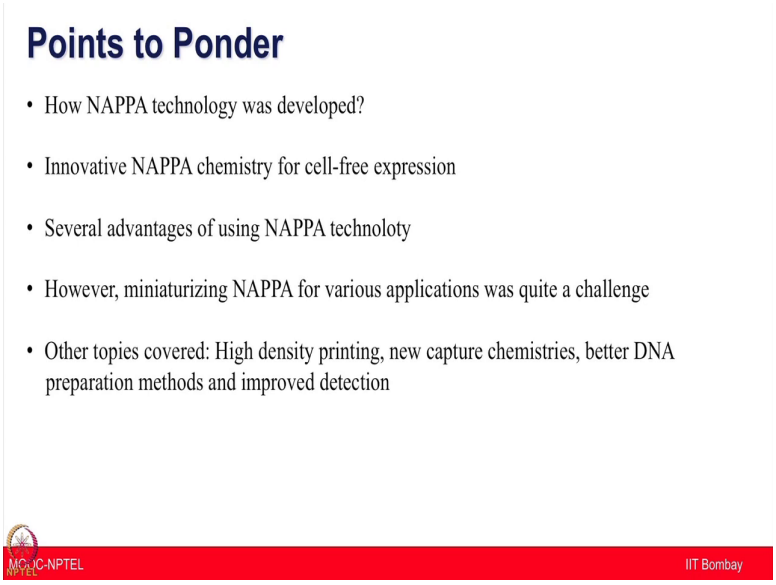
(Refer Slide Time: 39:12)



And if you look at patients and columns and antigens and rows, you will see that these top twelve or. So, antigens show up in numerous patients and so, this is the group of proteins that we should be looking at to think about developing a vaccine.

Now, there is a couple of things to remember. First of all, you cannot make a vaccine until you know that those proteins actually induce a protective response. So, we have not done that part yet. The second thing is that no single protein worked for everybody. It turned out that to get everybody. You have to get a mixture of a few proteins and I think that is going to be a common, but I know that is going to be a common theme moving forward in developing biomarkers is that it is going to be rare that a single biomarker will work for everybody. Eventually, you are going to need multiple biomarkers because different people have different responses.

(Refer Slide Time: 40:09)



So, I hope you got a very good overview of how this fascinating technology, nucleic acid, programmable protein array was developed. The kind of thought process of generating these resources especially protein without having the protein expression and purification was differently one of the revolutionary concepts in the proteomics field.

The NAPPA chemistry was explained by Dr. Labaer in detail and you are also now, familiar with what are the advantages of using this technology platform. Of course, there are challenges in miniaturizing these are sales these features to do the high density printing, but those were overcome with many innovative ways. And Dr. Labaer has talked to you about high density printing, the new capture chemistry's, the different modified ways of DNA preparation and the improve detection technologies which have really progressed the initial

versions of NAPPA technology to the very latest much more easy and reproducible and high throughput NAPPA based platforms.

I hope now, you are very convinced that using cell free expression microarrays could overcome many limitations of protein expression and purification. You need not to limit yourself to express and purify each protein of interest and large number of proteins to be purified before you can do a protein microarray experiment. Even if you have cDNA for the genes of interest, you can still do the protein microarray based experiments and NAPPA could be one of the very powerful technologies to do these kind of experiments.

I hope you got some understanding of this novel technology and a basics of some applications which could be performed using NAPPA arrays.

Thank you very much.