I'm Jennifer Van Eyk, and today I'm going to be discussing proteomics and heart disease and, really, the movement from technology — because proteomics is still very technology-orientated — but to look at how that technology can be used to understand clinical applications, as well as to how it can be used in the clinic itself.

So, we know that the phenotype of a disease actually is constituted by the cellular components. So, in other words, in the heart, the myocyte — when the proteins that are inside the myocyte, as they change with disease, is the actual phenotype of the disease. The two parts of that are really the hereditary and the environment, which will alter the protein content of the cell. The protein content of the cell is actually called the proteome. And it was originally thought to really involve just quantitatively the proteins that were there, but now we know, of course, that proteins themselves can be drastically modified both in their physical characteristics, as well as their location within the cell.

So, the proteome is really the third layer of the cell. We have the genome, the transcriptomes, the mRNA, which then translate into the proteins themselves. And it's the proteins that produce the phenotypes. And that's why proteomics is such a powerful strategy, or technology, in that it's the one that is directly studying the phenotype.

So, if you look at it, DNA is the genome, and it's quite static. mRNA can be varied. Then protein, which is a gene product and is, of course, very varied depending on the organ, as well as the disease, and the age of a person, and then the proteins themselves which, through a cellular process, can be modified, which will alter the structure and function of that protein and, collectively, all the proteins within the cell that have been modified produce a phenotype. And the phenotype can be that of the cell itself and, globally, as an organ or species. What's very true today, and what we know, is that one gene no longer produces one protein, but it's much more complex than that.

So, if you're to take a strip of DNA and look at the multiple genes that are present, here we have two genes, and they are produced by different exons and introns. They will produce a different number of mRNA transcripts. Those transcripts will produce an array of different proteins. So, one gene can produce four mRNAs, for example, which can, of course, produce ten protein species. But, as I said, proteins themselves can be modified post-translationally.

Suppose translation modifications include things like glycosylation, oxidation and phosphorylation, as well as the concept of these proteins moving around the cell. So, what are proteins? Well, they primarily start with twenty amino acid residues. Of course, there are some unusual amino acids as well. But these twenty amino acid residues are put down into a certain sequence, and those produce a unique protein. There are often common functional motifs within a protein such as a phosphorylation side, or a cleavage side, or a site where a protein interacts with another one, like a PT zed zone but, collectively, how these different functional motifs go together, as well as unique sequence, make each protein unique.

The proteins can have quaternary and tertiary structure so they'll fold and interact with other sub-units, or other proteins. Collectively, that will make the functional attributes of the protein and protein complex. And the regulation of the function is through the modification, either phosphorylation, or oxidation, or glycosylation, as well as through different protein protein interactions.

So, even though at the proteome, we are asking the question, what proteins are there and what are their modification it is still an indirect measure of the functional attributes. And so, proteomics is really the tip of the iceberg and it then has to be followed up by many functional analyses if you are going after molecular mechanism of a cell or disease.

So, there are four intrinsic characteristics of proteins — mass, the size of the protein; the charge, the number of positive and negative charged amino acid residues; the hydrophobicity, or really the solubility of the protein or is it a membrane protein or is it one that likes to be in the cytosol; and then the biospecificity, really the ability of the proteins to interact with one another. It is these four characteristics that we actually use to separate proteins and to distinguish them from one another.

In the next slide is the overall proteomic approach. We would start with the physiological model or, of course, a set of human tissues. In heart disease, that can be hard because we don't have access to a lot of human heart tissue. Usually, that's going to be at the end stage of heart failure using an explanted heart, but then you're only looking at the end stage of a disease and, if you're interested in acute models, this is not appropriate.

So, normally, we have to study physiological models, different animal models, that are very well defined in order to get enough tissue for proteomic analysis. The only caveat to that is when we're working with biological fluids. There we can have access to human populations, but in those cases we're often going after a different question than a mechanistic one. When we're using serum plasma, we're actually going very specifically for a diagnostic marker although, as you'll see later on, I'll be talking about how we can use serum in plasma in order to ask even more global questions about how the human body works.

But, essentially, you start with an animal model, and then you take your tissue sample. That can be the whole heart tissue, you can make isolated myocytes, whatever you need to do, depending on the actual question you are asking.

The big part of proteomics is concentrated really on the separation methodologies, and you'll see throughout this talk that I'll spend considerable time here. The goal is to be able to see as many proteins as possible within the context of the question, be it a broad-based question, you know, asking what proteins are present, or what proteins are changed in disease, or one that's much more focused, asking what proteins are associated with this, I'd say, a particular signaling complex. If you're looking at disease, you would then do differential analysis comparing disease versus non-disease over a time course. And then you have to go down to identify the proteins. And that's where really the second big step in proteomics came

about several years ago, and that is the development of very refined and user friendly mass spectrometers. Mass spectrometry is a technology that's very powerful and allows us not only to identify proteins, but to go on and characterize the different post-translational modifications.

So, these are some of the protein separation technologies that are currently available and, as you can see, they involve primarily using one or more of the intrinsic characteristics of the proteins in fact themselves. For example, one dimensional electrophoresis, SDS PAGE electrophoresis is a very common, very powerful technique that separates proteins based on mass, the size of the protein. Something like two dimensional gel electrophoresis, which I'll be talking about later on, is when we're going to separate proteins in the first dimension based on their overall charge, or pI, and in the second dimension, by mass. So, there are electrophoresis methods — liquid chromatography methods and methods that involve really biospecificity, or protein chip methods, where very commonly one would use amino precipitation, but there's also taking proteins where you would tag them, recombinant proteins, and then pull them out using the tag, or cell de-technology which, again, I'll be discussing later.

So, during this talk, I'll be talking really about two parts. One is going after mechanistic proteomic strategies. We'll be looking at very common strategies used in proteomics. The first is a database, and the database is really trying to, well really asking the question, what proteins are present in that tissue, or what protein is really present in that sub-proteome, let's say the mitochondria, or the ER, or the nucleus. The second approach is going after candidate proteins. This is a much more focused approach. You'll go in and say, what proteins are changed; we're only interested in so many, and then we're going to use that and screen a larger population. The second one is a broad-based mechanistic approach where we're going to try to see literally every protein present in the cell, if that was possible today, which it's not. But we try to use as many technologies as we have and go in as deeply as possible.

The second part, and probably the one that will hit the clinical realm first, is the ability to use this information one gathers from proteomics, either from the tissue or from body fluids, to develop biomarkers. And I'll briefly go over some of the aspects of biomarker discovery.

So, ischemic heart disease is what we're going to be focusing on as our system. And, it's really a spectrum of diseases. It can be a mild ischemic episode and produce myocardial stunning. Of course, with time, the cell is able to repair itself. It's been thought for a long time that that would involve a post-translational modification. If the ischemia is too long, you'll end up having cellular necrosis and a myocardial infarction. Those that survive an ischemic event and have a myocardial infarction may go on to develop heart failure. Heart failure itself, of course, is also a spectrum of disease. It's very progressive, going from compensation to decompensation, again, onto cell death.

So, what we're going to do when we're studying tissue, is to look at the tissue throughout the progress, from stunning all the way to heart failure and through that process.

The other side of the coin that one can view as proteomics, as I said earlier, is really looking at body fluids. And, for example, in serum or plasma. If you were to just take a serum sample right now of someone going through one of the stages of heart disease, you would be able to look at circulating proteins and peptides that are present in the serum and plasma because of the failing or injured heart. These biomarkers will be really systemic monitoring biomarkers. It will be very hard to find specific proteins that are being released from the heart in the serum, mainly, well, not that they're not there, they are, but because our technology is not good enough to get deep enough to see them in the proteome.

Now, if you're interested in very specific cardiac markers, one would have to go to the tissue analysis and go from tissue to serum. For example, upon cell death, the myocyte releases all of its components into the blood. It's vastly diluted. But, if we can detect those proteins, those will be cardiac specific biomarkers. Troponin-I is an excellent example of this. It is a myocardial protein that's released upon cell death into the serum. Troponin-I, as I'll show you, undergoes many modifications, and they are disease-specific.

So, what you need to understand when we're talking about cardiac-specific biomarkers is that, as the cardiac myocyte proteome changes during the course of disease, when the cell dies, these modified proteins are what's released into the serum and, if they can be captured and detected, can give us insight into this stage of the disease. Again, that contrasts to systemic biomarkers where what we're monitoring really is the response of the whole cardiovascular pulmonary system to the injury of the heart. There, we're going to see probably proteins like BNP and other proteins involving the acute phase response.

So, the first strategy we'll talk about as far as tissue is really data basing, again, asking the question, what proteins are present. And when you're doing this, there are many different approaches one can use. Some will just run one-dimensional gels; some will use something called a peptide shotgun approach. And this is where you take the whole sub-proteome, or your whole proteome, and you digest it with an enzyme. And this is going to create thousands and thousands and thousands of peptides which one attempts to separate using two dimensions in a liquid phase. The first dimension is often based on pI, or charge of the protein, using ion exchange chromatography, and the second one is reverse phase, or hydrophobicity separation. One then analyzes each peptide to obtain amino acid sequence and matches that to the database. The second approach is to use whole proteins. And instead of digesting the protein is actually to separate the proteins first, trying to either use two-dimensional gel electrophoresis or two DLC, the two examples I'll show you right now, in order to separate the proteins prior to their digestion and identification. And the example I'll be showing you today is really the analysis of the inner mitochondria membrane alone.

So, what are two-dimensional gels? Well, two-dimensional gels are really the cornerstone of proteomics still today. It was developed back in the 1970's but underwent one major change recently, and that was the production of IPG strips, immobilized pH gradient first-dimension strips. What this means is that, in the first dimension, we're going to separate

proteins based on their charge. Remember, each amino acid has a charge and, collectively, that will produce the overall charge of the protein. When you apply an electrical gradient, the proteins will migrate in this pH gradient until neutralized. That's when it's focused. Previously to the invention of IPG strips, we used to have to do this all by hand, and there was a great variation between lab to lab, and in reproducibility. With the invention of immobilized strips, we are able to get high reproducibility.

After we've focused in our first dimension, we literally place the first strip on top of the second dimension, which is SDS acrylamide gel electrophoresis, one-dimensional gels, in which we're going to separate proteins based on mass. What you can see as the advantage is, as you go down, you are able to, at the same mass, have three different spots. For example, each of those spots in the middle could represent three different proteins, or protein with two different modifications. It's our job to be able to compare that between control and disease states and then identify what the proteins are. And, in databasing, you would end up having to try to identify every spot you have on a gel.

The gel you see in front of you, the black and white image, is a real gel of the myofilament proteins, the very important contractile proteins that are found in the heart. Of course, this is a very simple proteome in that there are not that many proteins, and so the proteins can be very well resolved. And if we take out a spot to identify it, we can pretty much be assured it's going to be a single protein. But that's not the case with whole tissue. Whole tissue has thousands and thousands and thousands of proteins, which one cannot very well resolve onto two-dimensional gels.

So, what we do is we go into what's called sub-proteomes. These are different groups of proteins based on either their physical characteristics, their overall charge, or the pH dependence, or location within the cell. This is an example of a pH dependent sub-proteomics sub-fractionation in which we took cardiac myocytes and sequentially went from neutral pH down to pH of 2.3, first extracting out the cytosolic proteins and then the second extract, the myofilament and mitochondria proteins. This allows us to be able to see deep into the proteome.

Now, the other trick one can use is to use different pH gradients. And the pH gradient just spread out the protein further and further. As you can see, you are able to observe more of the proteome.

In the next slide, you can see a very well resolved cytoplasmic extract from pH 4 to 7. In the database, each one of these spots is cut out and analyzed to try to identify the protein using mass spectrometry. To do this, there are a series of steps one has to take. You cut out the band, or the spot that is of interest, and we digest it using an enzyme or a chemical that has a known cleavage site for an amino acid residue. If we use trypsin, for example, we know that it will probably clip behind an arginine, or lysine amino acid residue. Therefore, one is able to create a theoretical mass spectrum of all the proteins known in the database. You compare the theoretical mass spectrum to that of your observed mass spectrum. When you

take your digested protein, you then put it into your mass spectrometer, which is able to measure mass very accurately, of each of these fragments. Well, in fact, you don't see each of the fragments. In fact, you only see some of them. But you try to match the fragments that you have accurate mass on to those of the theoretical mass spectrum that you've obtained from *in silica* digest of all the known protein database. This is why protein databases are very important — the ones that are generated from the genome database that we have. Sometimes, if you're working in a species that's not well represented in the database, like pig or rabbit or dog, then one has to use homology. And, in this case, you're wanting to match based on homology between the species and the protein of interest. This can often raise problems, and so you need to go to the next level of confirmation, and that is obtaining the amino acid sequence itself.

So, what one does is you isolate in your mass spectrometer a single fragment. You add extra energy to this fragment, and it starts to break down to its amino acid sequence, and you measure the mass as it starts to break down. So, from a mass spectrometer, one is able to obtain a peptide mass fingerprint, as well as amino acid sequencing.

Now, when you're working on this, you often have to get to the sub-proteome if your goal is to see as much as possible, or to answer the question, what proteins are present in their certain sub-proteome, such as the nucleus, or the inner mitochondria membrane, or the sarcoplasmic reticulum. So, we've done that in our lab, and many other labs have gone in to try to find and isolate what we call proteomic perfect sub-proteomes. These are proteomes that we know should only consist of the inner mitochondria membrane, or the nucleus.

Sub-proteomes of the myocyte can consist of one individual organelle versus that of a chemical property. The organelle, such as a mitochondria is, of course, of interest to many diseases in cardiac muscle.

The inner mitochondria membrane is a sub-sub-proteome in some ways in that you would first isolate the mitochondria and then go in and isolate the inner mitochondria membrane. So, the question that we are giving the example today is to be able to ask the question, what proteins are associated with, or in, the inner mitochondria membrane.

So, the first way we're going to analyze the protein is to resolve the proteins from the inner mitochondria membrane on a two-dimensional gel. This is what it looks like on a pH 4 to 7, 12 percent SDS PAGE gel. You can see there are many spots. Here's a blow up. Very well designed spots, round circles. Each one of those spots is excised and taken to mass spectrometry. And you can see also it's very crowded at the top part of the gel. So, one would then go and use a 4 to7 gel again, but now a 6 percent SDS PAGE to spread out the top region of the proteins and get the high mass proteins of interest. Again, one might want to look at the basic proteins, and here's a pH 6 to 11 gel. Again, each spot is removed and analyzed by mass spectrometry, trying to identify each protein.

For example, aralar2 circled in green is a protein that has 6 transmembrane domains. It's a very hydrophobic protein. When we digest it with trypsin, we were able to identify and match with the known aralar protein 46 percent of the coverage of the amino acid sequence. This means that this protein, we're pretty confident, is aralar2. But see, now you've just learned something about mass spectrometry. It is really a probability of identification. The more information you have, the more sequence coverage, or amino acid sequence you have, the more sure you can be that you have the correct protein identified.

The other part that we do is to go in and look at proteins that are integral membranes versus those that are just associated with the membrane. And to do that, we used a very old biochemical trick to take the membrane homogenate and wash it with sodium carbonate to remove the lipid sheets and just look at those proteins that are integral. To do this for running one-dimensional gels, in part, because of the simplicity and, again, you can see aralar is sitting there.

So, you see, there's a lot of work to be able to get a database just based on one dimensional and two-dimensional gels. And, in fact, the database of the inner mitochondria membrane is over 150 proteins. But, of course, that's only one technology, and it only is separating proteins based on pI and mass. Can we use other intrinsic characteristics of proteins to separate them and maybe observe more? Well, one technology that we use in our lab is a two-dimensional liquid chromatography that separates proteins. And, in the first dimension, it's also going to separate proteins based on pI just like two-dimensional gels. But, in the second dimension, instead of separating on mass, it separates based on hydrophobicity.

Here are some examples of the inner mitochondria membrane. So, after being fractionated on pI, each fraction is analyzed by hydrophobicity. The longer the protein is retained on the column, the more hydrophobic the protein is. But, what we do to get the identification is, each fraction collected off this phase is then also digested and taken to the mass spectrometer for identification just in the same manner as we saw with two-dimensional gels.

The goal is to find as many proteins as possible in the inner mitochondria membrane, proteins that have been identified and that have a high probability of being correct. Here's some list of some of the proteins. And in the inner mitochondria membrane, you find proteins associated with the matrix inner mitochondria membrane. And one of the main reasons one does this is to be able to discover novel proteins, proteins that either one did not know existed in cardiac muscle, or proteins that you did not know existed, let's say, in the inner mitochondria membrane, although one has to be a bit careful that that's not a contamination, or a novel protein that was not known to be in mammalian systems and maybe, for example, only known to be in drosophila, or down, let's say, in one of the plants.

To discover novel proteins and finding them associated with a particular sub-proteome within the heart, or within any cell type, can lead to insights. But, of course, we don't

necessarily know the function of those novel proteins. And so, it's really the beginning step of a whole long range of experiments that have to be taken and carried out.

In the second strategy called the candidate protein screen, we already know what proteins are present in a sub-proteome, or in the tissue, and you're now going to ask what proteins are changed with disease. In this one, we often use when we're dealing with human tissue because we never have that much sample. In this one, we have used the genomic data to sub-classify a small group of patients with heart failure. Now, one can use anything — clinical parameters, echo, some biomarker that they have — to try to cluster ahead of time to get patients that you think are clinically closer to one another. And the reason one might choose to do this is that we know patient to patient variation can be huge. And you either have the choice of running hundreds and hundreds of human samples, which can be very difficult to get, or to gather really a group of patients that are closer together, analyze them in depth, and then take those proteins, knowing some of them will really be false positive, and screen the large cohort for a very specific sub-set of proteins.

So, what you see here is we took patients that had the genomics carried out, clustered them into, and those patients with cardiomyopathy, and analyzed them by two-dimensional gels, many different pH ranges, many sub-proteomes, and came up with a list of proteins that are changed. Now, the next step is really the validation, taking those proteins and carrying them out on a larger screen of patients. To do that, we're going to use a technology called SELDI, Surface-Enhanced Laser Desorption Ionization. But, really this area of protein chip is an up and coming area in technology, and SELDI right now works quite well, but it doesn't have all the characteristics of some of the new technologies coming downstream and, over the next few years, one will find many different protein chip technologies out there and available for us to use.

The way that we use SELDI in this approach is to capture the candidate protein using an antibody or a second protein that it can interact with. Capture it from a very complex proteomic mixture, be it tissue or tissue cytosolic extract, or whatever you are analyzing in this case or, of course, in a biomarker discovery, which I'll show later, you could be capturing it from serum, or from plasma. Once you've captured the protein based on biospecificity, one then puts this chip into a mass spectrometer. Now, I told you that normally what we do with mass a spectrometer is to digest the protein, put the peptides into the mass spectrometer and get accurate mass. That is true. That is one thing we do with mass spectrometer. But mass spectrometer, what it does is measure mass. It can do it on small mass, and it can do it on larger mass. The MALDI machine is very good at measuring whole mass at larger proteins. Therefore, one can use a MALDI mass spectrometer to analyze whole mass of the proteins that you have captured. You can't identify the protein, but you're able to determine the whole mass. For example, I'm going to give you an example of troponin-I. Troponin I is a key myofilament protein that's involved in the heart. And it really acts as a light switch for contraction and relaxation of the heart itself. Troponin I is a very positively charged protein and has a pI of over 9.5, binds incredibly tight to two other sub-units, troponin C and troponin T. And it's lying there on the myofilaments.

So, what we're going to do is we're going to capture troponin I using an antibody against that particular protein. But it's not just going to pull up troponin I. Because troponin I is bound so tightly to troponin C and troponin T, it also pulls out both of those proteins. So, bound on the antibody will be the complex of troponin I, troponin T and troponin C. When we put the chip into the mass spectrometry, we get three protein peaks with the three different masses of the protein. One advantage of this methodology is that one can also see post-translation of modifications. For example, if you are able to look at just the resolution of troponin I, for example, you might see three peaks. This is actually representing the native form and two phosphorylated forms. Troponin I can be phosphorylated by many kinases that regulate its function in the cell. Knowing its phosphorylation state, or whether or not it's degraded or modified in some way, which it is, very much, in disease, can help us diagnose and know the stage of the disease in the patient. But, of course, as a functional implication to this protein, we know so much about the function of this protein that any change to it will have a drastic effect on the outcome and contractility of the heart.

The third strategy I'd like to touch on is really the broad-based discovery. And it doesn't differ so much in the technology that one uses, but the strategy. Here, we're trying to really get it down, really down to the mechanistic, and asking what proteins cause disease. These are big studies where one would go in and try to see every part of the sub-proteomes and all the proteomes using every technology one can think of. The example I'm going to give you is really an acute disease of pre-conditioning. In this study, we took adenosine and diazoxide, two drugs known to cause pre-conditioning. Pre-conditioning provides protection of the heart against a very severe ischemic insult. Ischemia, when it's prolonged, results in necrosis, and cell death. So, when you do pre-conditioning, one can prevent the irreversible cell death occurring with severe ischemia. In the first window, these are very fast changes in protection lasting only a few hours and are thought to involve primarily post-translation modifications. In this study, we analyzed using two-dimensional gels of different sub-proteomes and many different pH's to find and identify 150 proteins, well over 150 proteins, that were found not to be changed. Surprisingly, only 26 proteins were identified to change with pre-conditioning, both with adenosine and/or diazoxide. The idea of using two different drugs was that we knew that they had very different mechanistic pathways, but both ended up with pre-conditioning. So, the assumption was that the common proteins, the proteins that change with both drugs might give us more insight into the pre-conditioned status. But what happened, to our surprise, is that most of the proteins were not the same. However, they all came down to changes within the mitochondria, or at least the vast majority of them. So, in this study, in the broad-based scheme, we are able to identify many protein changes all focusing really down into the mitochondria proteins and into the TCA cycle and the oxidation phosphorylation cycle.

So, what does proteomics do? Well, it was able to let us look across the many different sub-proteomes within the cell and see far-reaching changes, those involved in second window preconditioning in the nucleus and transcription factors to calcium handling proteins, but yet narrow it down to a cellular system because the majority of the protein changes resided again

in the mitochondria. Here's an example of changes that are happening. Again, the cell is not static, but it undergoes many changes. What you see here is a small region of a two-dimensional gel outlining two proteins that are undergoing change over time. If you look, each protein spot is actually a volume, which is what you see in the panel below. Two spots on the gel are two volume curves for adenosine control and diazoxide. And at five minutes, both proteins are present in all three samples. However, the drugs work at different times. What you can see is at 15 minutes, the proteins are starting to change. And at 30 minutes, the protein on the left has completely disappeared with adenosine, and it's only at 60 minutes that the same protein disappears with diazoxide. So, although the same protein is being changed with both adenosine and diazoxide, the time course in which they're doing this is very different. This provides us information about the high dynamics and the changes that are happening within the cell.

So, in proteomics, what one tries to do is to identify all the protein changes that occur in heart disease. Primarily, they are using animal models, again, only because of the inability to obtain a large amount of human tissue. If we start with an animal model, we need to validate that those protein changes are present in human disease if possible. And the reason one goes through this is really to get at the therapeutic side. By understanding the molecular mechanism of the disease and all of its broad and communicative interactions across the cell, one is able to select candidates based on our knowledge. Now, this doesn't mean that one of the proteins we find will actually be therapeutic but, in fact, will probably be the whole network one will want to go in to promote or inhibit as a new therapeutic target.

Because of the amount of work and the knowledge we need to gain in order to get to therapeutics, it will take a long time for proteomics to actually make an impact on that area. But what it will first be able to make an impact on is really diagnostics. Any protein that changes in the myocyte and that is released into the blood has the potential to be a biomarker. Any protein that changes in the serum has a potential to be a biomarker. What is important in biomarkers is not the function, but rather that the protein can very specifically be linked to a stage of the disease, and that that protein is only altered in that one stage of the disease. Therefore, the strategy to develop biomarkers either through tissue or serum, although we use the same technology, has a very different focus.

Why would you even want to go after biomarkers? Well, the game plan for many laboratories and really the future of diagnostic markers is the concept of panels, using more than one biomarker to finely tune the clinical diagnostic. For example, being able to use biomarkers such as BNP and troponin I collectively is able to more specifically define patients at risk, or at high risk, for heart disease, than those that are positive only for BNP or those that are only positive for troponin I. So, the idea is to be able to develop very specific biomarkers, single or panels, for each of the different parts of heart disease. Looking over all of them, one is able to then optimally, well hopefully, optimally design the therapeutic regiments, or treatment, of these patients.

There are two ways to get at biomarkers as I have alluded. How we go about it is to analyze both the myocardium to try to increase the probability of finding organ-specific proteins and, to do this, we'll want high abundant proteins such as the myofilament proteins or other cytosolic proteins like creatine kinase for they are released very quickly from the dying cell and can, therefore, be detected in the serum. The other way is to analyze serum specifically looking for a protein such as BNP. The strategy is really three-part — the discovery, validation, and then translation into the clinical test. In discovery, one can start with either animal models or discovery cohorts from patient cohorts. In discovery of animal models, one can start with tissue, and that should lead to a higher probability of obtaining heart-specific diagnostics. Working with serum or plasma, one probably increases the probability of finding cardiovascular and pulmonary monitoring diagnostics.

What we go to next is if you've worked on the animal model just as we did with the functional studies is to validate it in human tissue, or human serum. We then go on to validate the biomarkers, or potential biomarkers, against a large cohort. Now, I like to think about these cohorts in two phases. One is a large cohort that is within the disease state itself. So, if you were looking for an ischemic biomarker, you might be analyzing patients that have had angioplasty over time and looking for proteins that change there. But that's a very specific population. You would then want to validate your potential biomarker against the cohort of many patients entering emergency with chest pain. But, again, chest pain due to heart disease.

But if you wanted to differentiate between a patient with heart failure versus those patients that might have a response to a differnt disease — COPD, or asthma, or renal failure — you have to get rid of general disease markers. Therefore, one has to validate against a large mix cohort, mix cohorts having people or patients for many different disease types. What you're trying to do in the validation process is to narrow out and to identify those biomarkers that very specifically identify a particular disease state or disease process.

Also, during validation, because you'll be screening large numbers of proteins, you have to come up with the best panel, or maybe you'll end up only with one protein, but you hope that you'll end up with a panel of proteins. And those all have to be translated into the clinical test. This, of course, is a whole area involving both industry and clinical chemists, emergency physicians, and cardiologist for the education and the use of these biomarkers.

The discovery and validation aspects of this schematic definitely can use proteomic technology. The example I'm going to give you is the one I have been talking about a lot — troponin I. Troponin I is currently the gold standard, or one of the gold standards, for the detection of myocardial infarction. That's because cardiac troponin I is only found in the heart. Troponin I does exist in skeletal muscle, but it's a different isoform. So, if we find the cardiac-specific isoform of troponin I in your blood, it should only come from your heart. Therefore, one of the heart cells has died. At least that's the belief.

What we know about troponin I was actually discovered first using an animal model. What we found was that, as we increased the severity of ischemia of an isolated rat heart, that

troponin I very specifically and selectively was degraded. First off one end, the C terminus and then the N terminus. So, you can imagine, if in the heart, troponin I is all different lengths, that when the cell necrosed, that T and I is going to be all different lengths floating around in the serum.

So, what we did throughout the discovery of the troponin I was actually find out and go through that whole pathway. We found out that it was degraded in the heart. We confirmed that it happened in cardiac myocardium from bypassed patients. What was interesting is that it does not appear that troponin I is necessarily present in large mammalian species such as dog and pig, but tht other proteins such as troponin T are degraded. Remember, troponin T and troponin I are both sub-units of the same complex. We then determined the functional consequences of the degraded form of troponin I, and then we went on to look at it as a biomarker and asked the question, is it present in patients with AMI.

So, the concept again is that troponin I, or troponin T, or whatever protein of interest, will degrade very specifically and selectively with increasing degree of ischemic severity. So, if you detect a single intact protein in the patient's serum, that person is having a heart attack, but perhaps it's not as severely compromised as someone in whom you'd find many degradation products. The first question was to find out whether we could see T and I in the serum in very, very low concentrations, because it's diluted greatly in the blood. And so we had to develop a technology to allow us to do this. In this one, we are separating proteins based on one dimension, then blotting it on to nitrocellulose and using an antibody specifically for troponin I to light up the troponin I bands. There are two microliters of serum from a patient obtained over 5 days. This patient is undergoing a large heart attack. You can see that from the commercial biomarkers, CK CK-MB, and the commercial troponin I test. But what cannot be seen are degradation products. Upon admission, this patient had many degradation products of the troponin I present in the blood, and we are fairly confident that these degradation products are arising from the heart itself.

This patient, however, using the same technology, western blot, we were able to see and detect troponin I when the other commercial kits were unable to. So, this method turned out to not only be able to detect the degradation product, but was very sensitive for detection of the troponin protein itself.

So, western blot is one way of detecting the troponin I in the degradation product. But it's very slow and time-consuming. So, one can resort back to SELDI, Surface-Enhanced Laser Desorption Ionization, which I talked to you about earlier. And this is one of the methods that one can use to validate in serum, just as we were trying to validate proteins in the candidate protein technology what we were talking about with tissue.

Here, we're going to, again, capture troponin I using an antibody and put it into mass spectrometry and look for the mass. In this case, the whole protein is present, but so are a lot of degradation products. That's why troponin I at 23,000 daltons is not observed in this mass; it's outside the mass range, but you can see many smaller masses at 8,923 or 6,452. Those are

troponin I degradation products that we can detect in serum of patients with AMI. One interesting aspect of this is that one can see different degradation products depending on the severity of their ischemic event in patients.

So, what we propose we're seeing is, in fact, people experiencing stuttering events before they have an AMI and small amounts of troponin I, or troponin T, are being released into the serum. Remember, this is because a cell has necrosed. If we are able to detect it in a more sensitive manner than the current test, one will be able to go down and see patients with minor injury. The degradation products appear to only occur in those patients with very severe AMI, but you can have one or many, many degradation products, and the outcomes will be different for these patients.

So, what can proteomics provide? Well, when you're studying tissue and getting at mechanism, it can give you two pieces of information. It can give you a very broad overview of the cell. It can allow you to look at the communication between sub-proteomes and what's happening to the cell. It can look at both the gene and protein responses, and it can help you define the molecular mechanism of the disease process itself. If you're doing time courses, as you can with animal models, you are able to look at both the cause and effect over time. And, if you look at multiple models of the same disease, be able to identify those protein changes that are unique to a particular disease versus those that result from common disease mechanisms.

So, diagnostic markers will very rapidly hit the clinical field. Being able to have specific markers for staging and for the severity of ischemic heart disease will be incredibly useful and herald in a new way of analyzing and really optimizing the therapy of patients.