Comprehensive Molecular Diagnostics and Advanced Gene Expression Analysis

Prof. Arindam Ghosh

Dr. B.C. Roy Multi-Speciality Medical Research Centre

Indian Institute of Technology Kharagpur

Lecture 31 : Label-based Protein Quantification Technologies (Part 1)

Namaskar. Welcome back students, welcome to back to our lecture series on Comprehensive Molecular Diagnostics and Advanced Gene Expression Analysis. We are into our module 7 which deals with some advanced techniques regarding proteomics. And today we will be discussing various protein quantification methods labeled as well as label free. Actually this will be a divided lecture series. So, we will be in the course of the whole lecture series we will be covering mainly quantitative proteomics, the common methods of quantitative proteomics, the various methods that is isotope labeled affinity tag, isobaric tags for relative and acetate quantification ITRAQ tandem mice tag TMT SILAC which stands for stable isotope labeling with amino acids in cell culture.

And we will also be discussing various label free techniques these are all label techniques alright we use isotope labels for that. And thereafter we will also be discussing on data acquisition method. So, we willthis will be divided in a part lecture series alright. So, in today let us see how far we can discuss alright.

So, in case ofquantitative proteomics what is the need of quantitative proteomics you have heard this term right. The basic concept is a proteome the so much protein that are present in the cell the abundance that has got a significant influence on the multitude of the physiological functions that is happening in our body right. We have got so many functions that are going on inside a single cell even if you do not consider an entire tissue. So, the thing is if some of them undergo any form of change alright that can lead to diseases. So, we are mainly focusing on some clinical applications, but the thing is we need to know what is going on alright.

So, if we know the precise quantity or the nature of the protein that is present in the cell then their change we can easily I mean if we can identify the subsequent change that will unravel many questions alright or many problems that are going on inside this complex biological system alright. So, that helps us to understand the need of quantitative proteomics. So, what is quantitative proteomics? The name proteomics you already know quantitative means the amount anything which gives us the amount is known as

quantitative alright. So, quantitative proteomics is a field of study that involves the identification and quantitation of proteins within the biological sample. This is not a new concept we have already covered this in our basic proteomic workflow alright we have already discussed techniques alright.

Now where quantitative proteomics is used what are the applications? I believe you will also you will be able to say even if I do not alright. So, can you predict can you pause the video for a while and you write some of your pointers in a piece of paper and then you can verify whether you have touched these pointers or not. So, any biomarker discovery whenever we are discovering a new biomarkers protein based biomarkers quantitative proteomics very very very important. Understanding the cellular signaling pathways whenever we are treating cell with any condition or any stress how alteration is going on inside a cell again quantitative proteomics have got multiple role over there. Next drug discovery and development very important clinical application drug discovery and development uses multitude of processes that are that falls under quantitative proteomics.

Personalized medicine very very very important if we can identify how precisely the protein is altered in an individual in a specific biological system we can design the drug or a product in such a way that will cater to his or her personal environment. So, personalized medicine is again a very important area where quantitative proteomics has got a big role. Again investigation of post translational modifications as we discussed just by studying genetics we cannot address the phenomena that even after completion of central dogma there are some modification in protein which can lead to various physiological changes as well as diseases. So, again quantitative proteomics very important. And also comparative analysis in systems biology.

So, each system is quite different if you are using an animal model to understand any clinical diseases for specially in translational health research we need should be able to analyze various systems biology and quantitative proteomic techniques help us to understand that. So, these are few of many such applications that quantitative proteomics have got a definite role in all right. So, the thing is how quantitative proteomics is done basically there are two main methods number one is we need to first quantify the analytical techniques which are used in quantification and number two the data analysis. So, data acquisition methods all right. So, quantification and data acquisition these are the two broad processes by which we approach quantitative proteomics.

Now, looking into the quantification methods there are gel based methods and gel free method all right. Now, you already know from last module in detail about two dimensional gel electrophoresis differential gel electrophoresis these two have been discussed in much details and gel free methods are mainly mass spectrometry based methods all right. So, you already know these three we will be continue extending our

concept in application of these methods in further enhancing the mass spectrometric method all right. So, most of these methods ultimately uses mass spectrometry all right. So, with having a concept of mass spectrometry we can now easily understand these methods.

So, these labeled methods. So, you see here label free and labeling methods that is a labeling technology all right. So, we will be discussing label free methods after we have covered the isotope labeled methods or labeled methods all right. Now, this lecture we will be covering some part of the labeled methods and in the next lecture we will try to cover the remaining labeled methods as well as label free methods and then we will move on to data acquisition technique. So, this labeled methods using labeled technology can be again subdivided into in vitro and in vivo ok.

So, generally the methods that can be used in vitro are isotope coded AFNI tag or ICAT, isobaric tag for relative and absolute quantitation ITRAQ and tandem mass tag all right. The application the principle that can be used in vivo is stable isotope labeling with amino acids in cell culture all right. Mind it cell culture is an in vitro technique, but we can use this principle in vivo inside the body inside animals ok. However, this has to be done in test tubes ok. Anyway so, just mind it in vivo if a multiple choice question comes out of all these methods which can be used in vivo the answer will be SILAC all right.

Mind it this process can also be used in vitro all right fine. So, with basic knowledge of all of them these might seem a very I mean big names, but we will be decoding them one by one all right. So, first let us discuss isotope coded affinity tag. We will for convenience we will try to keep these lecture series as short in duration so, that it will be very easy for you to grasp the concept, complete the assignments and move on to the next class all right. So, in vitro isotope leveling technique in quantity proteomics by mass spectrometry mind it I told you mass spectrometry will have definite roles because these are the techniques which are used by mass spectrometry.

So, that uses chemical labeling reagents fine. So, each of them are keywords all of them can prop in multiple choice questions in an assignment or proctored exam. So, be very careful or you are reading through the lines. Now chemical probes consist of three elements. What are the three elements? A reactive group for leveling an amino acid side chain all right mainly iodoacetamide.

To modify cysteine residue, cysteine residues are mainly targeted or modified all right. So, one is the this one reactive group an isotopically coded linker and a tag for example, biotin. The principle that is used is affinity isolation. So, we first label our protein of interest with all these three and then use affinity chromatography to separate them ok.

This is the basic principle again you already know chromatography.

So, we are extending one concept into another ok. So,basically the principle is the samples are combined we are going to the principles we will be discussing workflow again. So, again you might find the information being repeated, but as I always say repeating is learning ok. The more youhear an song in loop the more easily you tend to remember right. So, we are also trying to do the same.

So, the samples are combined and then separated through chromatography. What chromatography already discussed affinity affinity chromatography. Then sent through a mass spectrometer to determine the mass to charge ratio between the proteins fine. So, here first samples are mixed then they are separated and then sent to mass spectrometer. The the special thing about this is only cysteine containing proteins or peptides can be analyzed all right.

Why because our tag only tags or labels the cysteine residues of the protein ok. So, since only cysteine containing peptides are analyzed often the post translational modification is lost. So, with the help of ICAT it is not a method of choice to study post translational modification ok. This we need to keep in mind. So, what is the development? So, I hope you have understood the principle we again will be discussing the workflow then it will again become clear.

So, in every step if you feel the video to be rushed which I honestly feel you would not still you can rewatch it again. So, the thing isoriginally you know deuterium the heavy hydrogen isotope was used to develop the tags, but deuterium tags what happened they tend to interact with the stationary phase of the column deuterium is heavy water right D2O. So, that led to a problem in peak separation during liquid chromatography. We are using affinity chromatography liquid chromatography. So, if we have got a component the tag heavy hydrogen which interacts with the column we will not be able to properly comment whether the peak is due to our analyte that has been separate or due to the interaction of the tag with the column.

Therefore, after many such engineering innovations it was redesigned using 13 carbon to address the issues. So, nowadays we use 13 carbon tags for ICAT alright. So, what is the workflow? For basic quantity comparison what we do is two proteome for example, two different sample or two different cell condition isolates one is control or one is treated ok. We label one with isotopically light probe and we label another with isotopically heavy version ok. So, both are labeled differently one is light labeled and one is heavy isotope labeled.

Then the combined proteins are digested mind it it is very important for any proteomic

procedure when you are preparing the sample this digestion the protease for example, trypsin is an inherent component of it. Whenever we are trying to analyze we need to break down the larger proteins into smaller fragments ok. Here the labeling is done first and then we are breaking down with the help of protease that is trypsin. Then they are subjected to affinity chromatography using the tag avidin alright. So, avidin has got affinity for biotin.

So, we use that biochemical principle to exploit the nature or the separation in chromatography alright. So, we use affinity chromatography to isolate the peptide with the help of isotope coded tagging reagents alright. So, I hope you can follow. Then these peptides are analyzed by I mean as these components are separated ok. These are analyzed using liquid chromatography and then subjected to mass spectrometry.

Basically the thing that I discuss once they are subjected to affinity chromatography and then they are analyzed by mass spectrometry it is nothing, but LCMS. As I told you LCMS is basically HPLC tagged with or combined with mass spectrometry. So, this is being done this is the phase of labeling and tagging, this is the phase of chromatography or affinity chromatography use HPLC using an affinity column and this is mass spectrometry. So, this is basically LCMS that is going on over here ok. Then the ratios of the signal intensity of different mass tag peptide pairs are quantified to determine the relative levels of protein in two samples.

So, if one protein of interest is expressed differently into different samples the way we design the tags it will be easily able to identify the differences ok. So, the labeling reagent has got a biotin, a linker region and the cysteine. This is the figure of the labeling reagent as I told you avidin and biotin have got a interaction in chemical inside the column and thus they help to separate the analyte on in case peptide of interest ok. So, what are the advantages of ICAT ok? Mind it whenever you are answering a question in exam or whenever you are writing any descriptive answer your first you should able you should actually explain the acronym isotope coded affinity tag. Do not start with your answer with ICAT yes the examiner knows, but if when you are writing a theoretically theoretical answer I mean answer in a descriptive theoretical exam it is always a must to first write the full form establish the acronym and then use it for subsequent explanation that which I have done right.

So, what are its advantages comparison of protein expression levels between different biological sample that thus has been shown in the graph or the figure. Next enables selective enrichment of cysteine containing peptides and proteins alright. So, if we already know a specific protein has got cysteine rich acid residues we can focus on that. So, thus allows us researchers to focus on a subset of the proteome fine. Next by tagging proteins with isotopically distinct labels it enables identification of identification

quantification of interacting proteins in complex biological samples.

This is a basic principle that we are using. So, since how a protein so suppose in a biological samples there are two different proteins that are interacting together. So, if we tag two different proteins and if they are interacting together it may so happen that they can be separated as a n block or as a dual peptide or a dipeptide and then we can easily come to a conclusion that these two proteins are interacting together. So, these are the various applications of isotope coded affinity tag which is one of the important method of in vitro labeling techniques in mass spectrometry and quantitative proteomics. Next we will discuss these methods iTRAQ and TMT, but before that these actually fall under the broad category of isobaric labeling.

So, what is isobaric labeling? Again mass spectrometry these are the I mean special variety or the special technique we are now going in depth into mass spectrometry detection methods. So, isobaric labeling is a mass spectrometry strategy used in quantitative proteomics definitely. So, where peptides or proteins see protein we refer is a single protein peptide. See amino acids are connected by peptide bonds ok. So, one peptide bond connecting two amino acid amino acid 1 and amino acid 2 this make a dipeptide I am going into very basics over here if any student is having some confusion regarding the term peptide and protein.

So, dipeptide whenever another amino acid comes and binds with a peptide bond it is known as tripeptide. So, dipeptide has got a single peptide bond tripeptide has got two bond. Similarly from 2 to 10 we call it as an oligopeptide multiple amino acid we call it polypeptide all right and ultimately and with various biological forces and interaction various bond forming these coil on to itself to form a full functional protein. So, if you break down a protein it has got peptides. So, we can work on fractionated peptides or we can work on proteins as a whole.

So, the basically this is all about knowing the basic things or the terminologies of proteins. So, proteins or peptides are labeled with chemical groups that have identical mass, but vary in terms of distribution of heavy isotopes in their structure all right. So, the tags that are used to label the proteins have got similar mass identical mass all right, but they vary in distribution of heavy isotopes in their structure. So, we actually try to target I mean separate proteins that have a similar mass it has says at least nominally ok and this chemical groups should also have nearly identical mass all right. So, again read it proteins or peptides are labeled with chemical groups ok.

The proteins can or cannot have identical mass even if they are of identical mass we can separate them. Now, the chemical groups we are using to label them should have ideally identical mass or nearly identical. So, nominally variable a variation is allowed ok. So,

they vary in terms of distribution of, but they should they actually do vary in terms of distribution of heavy isotopes in the structure this is the basic basis of which by which we can separate them. So, under these two will be focusing on discussion on isobaric tags for relative and absolute quantitation termed as ITRAQ and tandem mass tags also termed as TMT ok.

So, let us first discuss ITRAQ ok, isobaric tags for relative and absolute quantitation. So, ITRAQ definitely is a quantitative proteomic research technology it was discovered in 2004 which helps us helped us in comparing proteins of various different samples very basic thing we are basically introducing the ITRAQ technology. So, what is it is power again just like ICAT it can help us study difference in protein expression levels in tissue sample under pathological condition or various developmental stages ok. It can accurately quantify and identify all proteins expressed in a complex biological system or in one genome ok. So, these are all the praises about ITRAQ that we have learnt even before we have learnt ITRAQ.

So, let us see what basically is happening. So, it uses this method is based on covalent labeling of N terminus and side chain amines of peptides from protein digestion. The earlier method ICAT was targeting cysteine residues here it is targeting the amines ok. So, terminal N terminal amine group or the side chain amine group where we are getting them if when we are digesting a protein suppose we have got a complex protein we need to analyze we are digesting the protein with a protease such as trypsin then those proteins are subjected to labeling by these tags alright. Tags with nominally varying mass using stable isotope labeled molecules ok. The butase it can label 8 different reagents simultaneously ok, the protein can be labeled with 8 different reagents and compare 8 different protein samples ok.

So, we are using 8 different reagents it can happen in 4 or 8, but it can up to 8 possible even up to 10 are possible using modern methods ok. Now when the tagging is done these reagents are composed of a report group and balance group and specific amine reactive group just like ICAT which was using a target group, a linker group and a biotin alright this is also comprised of these tags are composed of 3 different components ok the report group, balance group and one amine specific reactive group. So, as I discussed we when we are discussing the workflow these are the 4 different samples. Let us follow the figure and we will be easilyunderstanding the workflow. So, they are parallelly dealing denatured and digested ok.

So, all the 4 samples are parallelly denatured and digested. So, the currently 2 main reagents are used in come for commercial purpose for clinical detection as well as research purposes 4 plex and 8 plex meaning either 4 dyes or 8 dyes alright to label all peptides from different samples and treatments ok. So, here we are using 4 different

groups you can see the isotopes are varying. So, one is labeled 114, 115, 116 and 117 alright and the chemical groups are also varying nominally in their masses ok due to the fact that they are using different isotopes ok. Now they are tagged and then they are subjected to MS-MS or tandem mass spectrometry alright.

There we see that the peptide fragments are equal, but the reporter ions are different. So, let us understand the workflow again all the 4 samples are digested and labeled with these ITRAQ labeling compounds isotopes, then these samples are pooled and then fractionated by liquid chromatography and analyzed by tandem mass spectrometry. So, if you have understood LC-MS you can extrapolate that concept to almost all of these labeling methods. So, first we are digesting we are treating them with all the labeling compound various different 4 to 8 labeling compounds, then we are putting them all together then are fractionated using liquid chromatography and then subjected to tandem mass spectrometry to generate a data alright because mass spectrometry generates peaks ok. Then with the help of that characteristic curve or chromatogram we are searching a database alright because we are getting molecular information about various groups, then a database searching should be done in order to understand what chemical groups have been separated alright.

So, a fragmentation data is subjected to database searching to identify the level peptides and hence the corresponding proteins alright. So, we can back calculate the to know what type of proteins are tagged by what reporting dye or here the tags. Next the fragmentation of this attached tag generates a low molecular mass reporter iron very important that can be used to relatively quantify the proteins and the proteins from which they have originated ok. So, since these reporter this tags that were I attached with the proteins after tandem mass spectrometry we can again fragment them to know their nature that will give us an idea about what protein from which they have been regenerated. So, this area is very important the fragmentation of this attached tag generates a low molecular mass reporter.

So, it is generated inside the reaction ok and this different information about different low mass reporter will give idea from what protein they have been generated ok. Now, the thing is how what are the advantages of iTRAQ? You have already understood till now what are the advantages. So, let us even if we just recall the process you will be able to see that the process itself are the advantages. So, see this is in vitro labeling technique performed at peptide level and is suitable for many type of biological samples why if you have multitude of biological sample the more the merrier why it can simultaneously mark 4 to 8 sample in one experiment this is the major advantage over the previous method there we were using two samples one was tagged with a normal light or one was tagged with a heavy ok because we are targeting the cis tine residue. Here since we are labeling the terminal amines and we can label multiple such sample we can analyze

multiple data observe data from multiple condition in one single experiment this is very simple and fast.

So, since we can analyze or report multiple data or analyze do multiple experiments in a single unit time this means this process has got high throughput. So, it is very suitable for high throughput proteomics for direction of multiple sample and hence the experiment design remains much more flexible and the thing is it has got characteristic of good repeatability and high sensitivity means the data we have got from one reading we can rely upon. So, there is less need of repeating our experiments alright. So, both qualitative and quantitative analysis can be performed simultaneously this is also another advantage of eye track alright.

So, we will stop our discussion here. So, in today's class we have discussed what is quantity proteomics we have discussed the overview of quantity proteomics what is the need. So, we have discussed what are the approach or major steps by which quantity proteomics is done we have discussed the common methods and we have started discussing two methods in this class isotope coded affinity tag and isobaric tags for relative and absolute quantitation I cat and eye track and both the cases we have discussed the principle the workflow as well as the advantages and we have also touched on the developmental history that initially deuterium was used and now it is replaced by 13 carbon for designing the tags in case of eye cat. So, that is it for today these are the references for today's class and I thank you for your patient listening. Thank you.