

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

LECTURE-25

Quantitative proteomics: iTRAQ and TMT

TRANSCRIPT

Welcome to the proteomics course. Today we will talk about quantitative proteomics and discuss about iTRAQ and TMT techniques. The quantitative proteomics aims to answer various questions including the identification of biomarkers for various diseases and target identification for early diagnostic intervention. Various techniques or strategies have been developed for quantitative proteomics including the gel-based approaches such as DIGE which we have discussed in the last module and MS-based measurements. In the last class we talked about ICAT which was first generation MS-based quantitative approach for quantification of relative levels of proteins in various control and treatment samples. As we discussed the ICAT reagent consists of three thiol functional groups, a linker group and a biotin moiety. Two samples treated with ICAT reagent, light or heavy isotope, were selectively alkylated for cysteine residues. These samples were combined, tryptic digested, affinity purified based on avidin and then analyzed by LC-MS based approach. So the ion abundance of light and heavy isotope-labeled peptides could provide the relative abundance of these proteins. ICAT was the first quantitative MS-based approach which was initiated. But because only two samples could be analyzed through this approach new approaches came considering the sample multiplexing. Because one limitation of ICAT was because only two labels are available, therefore more samples can not be compared in a given experiment. So there is a need to compare large number of treatments that led to the development of two or four-plex isotope coded protein label (ICPL), four or eight-plex isobaric tagging for relative and absolute quantification (iTRAQ) and two or six-plex tandem mass tag (TMT) based techniques. Today we will continue our discussion on iTRAQ and TMT techniques which can compare upto 8 and 6 samples in a single analysis. So these techniques can increase the throughput. iTRAQ: the identification and quantification of complex protein mixtures have been facilitated by MS-based quantitative proteomics

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

techniques. The iTRAQ reagent consists of amine specific stable isotope reagents that can label peptides of upto 4 or 8 different biological samples. Although both ICAT and iTRAQ are similar in overall concept, ICAT depends on tagging cysteine residues, whereas iTRAQ is based on tagging of primary amines.

So we will continue today's lecture on quantitative methods; the in vitro labeling based approaches. We will discuss in detail about iTRAQ method, I will talk about the composition of iTRAQ reagent, then I will try to give an experimental overview for the iTRAQ procedure, we will compare iTRAQ and ICAT techniques, then I will talk about TMT and then we will compare iTRAQ and TMT.

So let's talk about quantitative proteomics and different in vitro labeling methods. The in vitro labeling methods rely on use of labeling reactions at a specific site in proteins or peptides. Based on various labeling chemistry different types of strategies have been developed to introduce isotopes at either protein or peptide level. We will discuss these strategies in the next slides.

There are 3 different types of in vitro labeling; amino acid based labeling, N-terminal peptide labeling and C-terminal peptide labeling. In the amino acid based labeling such as isotope coded affinity tag (ICAT), visual isotope coded affinity tag (VICAT), mass coded abundance tagging (MCAT) and quantitation using enhanced signal tags (QUEST). Then there are different type of N-terminal peptide labeling methods such as iTRAQ, TMT and global internal standard technology (GIST). Then there are C-terminal peptide labeling methods such as esterification and proteolysis using ^{16}O or ^{18}O , light and heavy form of oxygen. So there are variety of in vitro based labeling approaches available and depending upon your quantitative application there is a possibility of using different type of strategies.

In today's discussion we will mainly focus on iTRAQ and TMT. So I discussed that MS has played a major role in proteomics and now it is becoming a very essential tool to study the complex biological system in various diseases. iTRAQ is a MS-based technique for relative and absolute quantitation of proteins present in upto four samples

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

or upto eight samples depending upon the iTRAQ tags and these labels can be provided in the proteins at the N-terminal. The iTRAQ labels are available from Applied Biosystems and TMT tags are available from Thermo Fisher. There are the only commercially available tagging techniques, currently, where quantitation can be carried out in the MS/MS mode. The iTRAQ technique was described for the first time by Ross et al. in 2004 and it was subsequently commercialized by Applied Biosystems.

iTRAQ method provides multiplexing capability of analysis which was not possible using ICAT. In ICAT, only two different reagents are possible, light and heavy form. Only one light and one heavy can be compared, whereas iTRAQ was the first method which gave an opportunity to compare more than 4 samples. So 4 samples and 8 samples can be analyzed by using iTRAQ method. So these iTRAQ reagents are set of multiplexed, amine-specific, stable isotopes. It enables simultaneous identification and quantitation, both relative and absolute. There are two types of iTRAQ reagent available, 4-plex for processing upto 4 samples and 8-plex for analysis of upto 8 samples. In iTRAQ method the derivatized peptides of a given sequence are isobaric and are co-eluted because they are derived from control and treatment biological samples. So in MS upon collision induced dissociation during the MS/MS experiment, it provides reporter /signature ions that differ in m/z value.

So the reporter ions are used to track the quantitation and can be used to monitor the relative quantitation for proteins.

Now let's discuss about the iTRAQ reagent. There are set of 4 isobaric amine specific labeling reagents, 114, 115, 116 or 117. The iTRAQ reagent consists of a reporter group, a balancer group and a peptide reactive group (PRG). PRG labels the N-terminus of all the peptides as well as the free amine groups of lysine side chains. The neutral balance portion and reporter group provide total mass of 145. So this method can allow the multiplexing of upto 4 or 8 different samples in a single LC-MS/MS experiment. The different distribution of isotopes between the reporter and balance

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

group makes the label isobaric and it enables the detection upon fragmentation and their release in MS.

So as I briefly talked, there are three major components of the iTRAQ reagent. 1. Reporter group which is based on N,N-dimethylpiperazine (DMP). 2. Mass balance carbonyl group. 3. Peptide-reactive group which is an ester of N-hydroxysuccinimide (NHS).

So in the iTRAQ reagent, the m/z value range from 114.1 to 117.1 if you are using a 4-plex reagent. The balance group of mass is 28 to 31 Da. The overall mass of reporter plus balance components remains constant. So 145.1 Da will remain for all the four reagents.

So when reacted with a peptide iTRAQ tag forms an amide linkage to any peptide amine, N-terminal or lysine amino group.

This is the structure of iTRAQ reagent where you can see that it consists of reporter group in neutral balance group and a peptide reactive group and overall mass remains 145.

The reporter group provides signature ions in MS/MS mode. It provides good b and y-ion series and it maintains the charge state and ionization efficiency of peptides. There is a balancer group. It balances the mass change of reporter to provide total mass of 145 and neutral loss in MS/MS mode.

The iTRAQ label is an isobaric tagging compound consisting of reporter ion series. Here I have shown you a 4-plex iTRAQ reagent with 114 to 117 Da, a balancer region with 191 to 188 Da, so that the total mass of isobaric tag is 305.

Now let's look at 8-plex iTRAQ reagent. In 8-plex reporter ions, the variable masses from 113 to 120 Da and the balancer region is from 185 to 192 Da so that the total mass of isobaric tag remains 305.

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

We will now talk about performing an iTRAQ experiment. In an iTRAQ experiment, the control and treated proteins samples are reduced, alkylated and digested with trypsin. The digested samples are reacted with different iTRAQ reagents, 4- or 8-plex. The samples are then combined. In fact it can control and 3 treatments or it can be 4 different time-point samples. So four samples can be combined and analyzed by LC-MS/MS. So as per the guidelines provided by Applied Biosystems the protein samples can be prepared and now we will discuss that step-by-step in the next few slides.

So in the sample preparation procedure first part is protein reduction and blocking of cysteine. Firstly dissolve the protein samples in 0.5 M triethyl ammonium bicarbonate at pH 8.5. Then perform a reduction step by adding a reducing agent. Incubate the samples at 60°C for an hour. After that add a cysteine blocking reagent so that the cysteine residues are blocked.

Once reduction and cysteine blocking is done then the protein digestion is performed. Add trypsin solution, incubate overnight at 37°C so that these proteins can be digested into peptides. We may have discussed this thing briefly in the previous module when we talked about in-gel digestion for doing gel-based proteomics. Similar types of concepts are applied here but this one is in-solution digestion. Once you have done the over-night digestion then a clean-up step can be performed by using ZipTip so that some salts and other contaminants can be removed.

Now next step is labeling. Adding these iTRAQ reagents to four different samples. Firstly, reconstitute the iTRAQ reagent in isopropanol. Add iTRAQ reagent to the digested protein samples. If you have four samples you can use 4-plex iTRAQ reagents, 114, 115, 116 and 117, as I have shown in this slide.

Now combine these labeled samples in one tube so that later on sample processing will not have any manual artifact. So the labeling is performed separately, after that all the labeled samples are pooled in one tube.

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

These pooled samples can be further purified on a strong cation exchange column so that the excess unbound reagent can be removed. This is an optional step but this does help in proper cleaning and getting a better signal. This step facilitates sample clean-up. This is although not essential but recommended to perform this step.

Let's have an overview of the protocol. We have control and treatment populations. Each of those samples were first reduced, cysteine blocked and digested. After that iTRAQ labeling was performed for each sample. Then control and treatment populations were all mixed together with iTRAQ labels. Fractionation and clean-up steps were performed and then LC-MS/MS analysis can be performed for protein identification and quantification.

I think in this slide it will be more clear to you the overall step. You have four different populations A, B, C and D. each population is reduced, cysteine blocked, trypsin digested and then further labeled with 4 different types of iTRAQ labels, combined, fractionated and further LC-MS analysis can be performed and quantitation can be obtained.

Once the sample preparation is done then further MS analysis can be started. The peptides which are differentially labeled can be mixed together and measured by MS. This method enables simultaneous identification and protein quantification. The labels react with the N-terminus and the reporter group is lost during fragmentation. This method can be used to determine the relative abundance of selected peptides of interest from 4 or 8 samples. Therefore, high-multiplexing can be obtained by using iTRAQ method.

Four independent reagents of same mass (145) can give rise to four unique reporter ions ($m/z=114-117$) in MS/MS analysis and subsequently these reporter ions can be used for quantification of different samples.

In iTRAQ the quantification occurs at the level of fragment ion spectrum (MS/MS). The peptides with identical sequence but derivatized with different isobaric reagents are

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

indistinguishable during the MS spectra alone. That is where only during the MS/MS phase the quantitation can happen. So it shows the identical fragment ion series for the peptide derived fragments but it shows low mass reagent derived fragment ions that indicate that sample of origin and permit the quantitation. So the identification and quantification of peptides during the MS/MS analysis. The iTRAQ enables the quantitative protein profiling of multiplex samples without making the MS spectra complicated. It reduces the redundancy in selecting the MS precursor ions for collision induced dissociation and enhances the efficiency for MS/MS analysis and peptide identification.

This is a representative spectra for the 4-plex MS analysis and the MS/MS spectrum is showing the reporter region, signature of these 4-plex iTRAQ labeled peptides 114, 115, 116 and 117.

Now we will have a look on the 8-plex data. In the MS/MS spectrum we are showing the reporter region signature of a 8-plex iTRAQ reagent showing 113.14, 114.15, 115.15, 116.15, 117.16, 118.16, 119.16 and 121.16 reporter ions.

I will now describe you the overall iTRAQ technique by showing you an animation

The protein samples to be analyzed are first digested with trypsin into smaller peptide fragments. The trypsin cleaves the proteins at the C-terminal of lysine and arginine residues unless they are followed by a proline residue.

The peptide fragments generated are separated by SDS-PAGE to simplify the mixture and then tagged with the iTRAQ label. The iTRAQ reagent consists of a reporter group, a balance portion and a peptide reactive group that interacts with the N-terminus of the peptide or free amino group of Lysine residues, giving it an overall mass of 145. The reporter group used to label each peptide sample is unique, with mass varying between 114-117, thereby enabling the labeling and quantification of four samples simultaneously. This has been further improved to allow labeling of eight samples simultaneously.

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

As you can see the first sample has a reporter group 117 and balancer group 28. Similarly for the 2nd, 3rd and 4th samples the reporter groups are 116, 115 and 114 and the balancer group 29, 30 and 31, respectively. Overall 145 Da.

The labeled samples are then pooled together. The pooled samples are purified on a strong cation exchange column to remove any excess unbound iTRAQ reagent. This facilitates sample clean-up prior to further finer separation and purification using reverse phase chromatography.

Further purification of the SCX purified peptides is carried out by reverse phase liquid chromatography wherein the sample is passed through a column containing a packed stationary phase matrix that selectively adsorbs only certain analyte molecules. The eluted fractions are further characterized by MS.

The purified labeled peptide fragments are then analyzed by MS/MS. The different masses of the reporter groups allows the peptide fragments to be identified. The reporter group is lost during fragmentation. Relative quantification of up to eight samples can now be performed using iTRAQ.

Following iTRAQ, the data obtained from tandem mass spectrometry can be analyzed by means of the Mascot search engine. The MS/MS data analysis shareware requires inputs from the user regarding the experimental parameters used such as enzyme cleavage, protein name, modifications, instrument used, peptide charge etc. and the desired search criteria like taxonomy, peptide tolerance etc. Commonly used protein databases against which the MS information is processed to retrieve sequence data include NCBI, MSDB and SwissProt. The data file generated from MS is uploaded and the search carried out.

I hope that the animation was informative and now you are able to understand all the steps involved in iTRAQ experiment. So now let's talk about advantages of using iTRAQ method. This method performs relative and absolute quantification upto 4 or 8 samples. It gives good multiplexing capability, it increases the analytical precision and accuracy.

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

The expanded proteome coverage can be obtained by iTRAQ method by tagging the tryptic peptides. It eliminates the limitations of previously discussed ICAT for dependence on cysteine residues.

But this method has several disadvantages as well. First of all there's possibility of errors in the quantification due to the differences in the efficiency of enzymatic digestion because we are performing that step separately for four samples. Earlier we talked about sample A, B, C and D, each of those was treated separately and if during the tryptic digestion the efficiency was not exact for the digestion process. May be that could lead to some difference. Although if you start with the same enzyme lot and if you have done the quantification for the proteins properly, ideally that should not matter. But there's some possibility that some errors could be coming from that efficiency of enzymatic digestion. The peptides pre-fractionation step is another possible way of introducing some variations. The variability in the initial protein digestion and then tagging is performed only after the individual sample processing is done. So if we can do the multiplexing in some way that the tryptic digestion can not be performed separately, that can reduce the overall chances of error. But that will actually be not possible because we need to label each sample separately with iTRAQ reagents. So these are still of the possible demerits and disadvantages of iTRAQ method.

These reagents are quiet costly. That is also a limiting factor for many labs to be able to use that. And various search algorithms and databases are required. Obviously more and more studies are happening. So we have now better software and information available for performing this experiment. But still that is one of the limitation.

So we will now talk about comparison of iTRAQ with ICAT techniques. Both iTRAQ and ICAT allow identification and quantification of proteins. The ICAT is an amino acid based labeling method as we discussed earlier and iTRAQ is based on primary amino group. So iTRAQ potentially allows tagging of all tryptic peptides. ICAT labeling has advantage to reduce sample complexity by eliminating the non-labeled or non-cysteine containing peptides. iTRAQ provides multiplexing compatibility whereas ICAT provides

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

only 2 sample comparison. For example if you want to compare various stages tumour progression simultaneously from the normal to the cancerous stage. Now iTRAQ is a good system which can be used if you want to look at different developmental stages of an organism then iTRAQ again is a good choice as multiplexing as it is not possible using ICAT.

iTRAQ method provides more complete coverage of original protein sequence than the ICAT method because every proteolytic peptide is tagged at the N-terminus. So tagging is not limited with cysteine residues as in ICAT. In iTRAQ has increased confidence in the identification of proteins. In iTRAQ all the precursor ions are isobaric so it saves some MS run-time. There is no need to run LC-MS/MS to sequentially select the differentially labeled precursor over time course of HPLC peak. In order to obtain MS/MS spectra to identify the proteins the same MS/MS spectrum can be provide information on peptide sequence despite of the label it presents. So the data required for relative quantification can be obtained where relative signal intensities of reporter fragment ions.

Just briefly look at the iTRAQ applications and I will show you this in an animation.

Animation - In this animation we will look at one application of iTRAQ method. A study by Boylan et al. 2010 used iTRAQ for the identification of candidate biomarkers in ovarian cancer serum.

A multiple affinity removal system was used to carry out immunodepletion of the serum samples from normal controls as well as ovarian cancer patients. This helped in removing the high abundance proteins, leaving behind only the medium and low abundance proteins for iTRAQ analysis.

The immunodepleted serum samples were then labeled with the iTRAQ reagent and analyzed. The authors detected a total of 220 unique proteins of which 14 were found to be elevated in the ovarian cancer serum samples compared to the healthy controls and four novel candidate biomarkers were detected. Results were validated by Western

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

immunoblotting. This just gives you an overview of how iTRAQ reagents can be used for various types of applications including biomarker discovery.

Now let's talk about TMT. This method is similar to iTRAQ. TMT is also MS/MS based quantitative technique which uses the isotope labeled models referred as tandem mass tags. It provides the accurate quantification of peptides and proteins. TMT has been developed by Proteome Sciences and currently commercialized by Thermo Fisher. I have given you the reference to the original study on TMT in this slide.

These TMTs are based on similar principle of iTRAQ. Here the multiplexity of upto 6 multiple labels is possible. The TMT isobaring technique can be used to perform absolute quantification by adding stable isotope labeled internal standard peptides. It can be done by comparing the peptides from a target protein to a known amount of labeled standard peptide spiked into sample. In that way the absolute quantification can be obtained.

The N-terminal amine and lysine residue are labeled through NHS group. There are families of chemical tags which are based on common structures. The series of TMT tags available TMT0, TMT duplex and TMT 6-plex. So these TMTs are innovative set of isobaric mass tags for labeling the proteins and peptides at amine functions and mixing of upto 6 different protein samples are possible. While duplex and 6-plex labels TMT differ by the number of isotopic substitutions. TMT0 is non-isotopically substituted that has been produced for method development only. During the MS/MS analysis the TMT tags give rise to 6 reporter ions 126 to 131 Da, therefore it allows for the relative quantitation. In TMT 6-plex, each tag adds a mass of 229 Da per labeled amine to the protein. The TMT duplex and TMT0 share the TMT complex structures.

Let's look at TMT0 labeled structure in more detail. The TMT0 tag is used for testing and optimization of sample preparation, labeling, fractionation and MS fragmentation for peptide identification and reporter detection. The modification 224 Da and MS/MS reporter ion is 126 Da.

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

Now let's look at TMT duplex. The TMT duplex reagent allows for the comparison of two samples. 126 and 127 are two different MS/MS reporter ions available and modification is 225 Da.

Now let's look at TMT 6-plex reagent. It allows the comparison of upto 6 samples. The MS/MS reporter ions, as you can see in the structure are from 126, 127, 128, 129, 130 and 131 Da. So the TMT 6-plex allows the comparison of upto 6 conditions. It could be useful for studying time-course, drug-dose responses, replicates or looking for multiple disease comparison. The modification is 229 Da.

I am showing one representative MS/MS spectrum of TMT labeled peptide which is showing a reporter region. The relative abundance of target protein or peptide fragment of 6 different samples is easily measured by comparing these signature mass peaks which are generated by different mass tags.

Let's now look at the comparison of iTRAQ and TMT tags. In iTRAQ, as we talked, there are two different type of reagents available, 4-plex and 8-plex. In both there is a reporter group, a balancer group and then there is a protein-reactive group. Same concept is in TMT tags where we have a reporter group, a balancer group and PRG. I have shown a comparison with 6-plex TMT tag but as you have seen earlier there are TMT0 and TMT duplex tags available. Now in iTRAQ the reporter in the 4-plex consists of either from 114 to 117 Da, the balancer is between 28 to 31 Da whereas in 8-plex it is from 113 to 121 and the balancer is from 182 to 192. In case of TMT, the reporter group has 126 to 132 Da and the balancer consists of 97 to 103. So tags are quite similar in overall structure. The iTRAQ analysis can be performed by using the software such as ProteinPilot and Mascot. The TMT based analysis can be performed by using software such as Protein Discoverer and Mascot.

So overall in today's lecture we talked about iTRAQ technique, we compared iTRAQ with ICAT and TMT. During the discussion we talked in detail about performing the iTRAQ experiment. So we know that there's been steady advancement in the field of MS, the quantitative proteomics has progressed dramatically in the past few years. The

NPTEL VIDEO COURSE – PROTEOMICS

PROF. SANJEEVA SRIVASTAVA

efforts to analyze proteome of many species, both qualitative and quantitative have generated abundance of data from a variety of biological samples, from bacteria to human. The chemical labeling using the isobaric tag for relative and absolute quantification iTRAQ or TMT reagents is based on the extraction of reporter ions from tandem-MS, MS/MS spectra. These methods can be used on all kinds of biological samples and provide high level of reliability for the quantitative data. So we will continue our discussion on quantitative proteomics techniques in the next lecture on in vivo based stable isotope labeling methods and then we will try to compare in vitro and in vivo based labeling methods.