LECTURE-16

In vivo Quantitative Proteomics using SILAC

HANDOUT

PREAMBLE

Mass-spectrometry based quantitative proteomics is an attractive platform towards studying the dynamic changes in proteins subjected to various stimuli. The identification and quantitation of complex protein mixtures have been facilitated by MS based methods for differential stable isotope labelling. These tags, which can be recognized by MS, provide a basis for quantification. Stable Isotope Labeling by Amino acids in Cell culture **(**SILAC) incorporates specific labelled amino acids into proteins for differential analysis.

OUTLINE OF LECTURE

1. Limitation of iTRAQ and emergence of *in vivo* labeling
2. *In vivo* labeling or metabolic stable isotope labeling
3. Principle of SILAC
4. SILAC Experiment Work flow
    a) Preparation of SILAC media
    b) Cellular adaptation
    c) Differential treatment
    d) Cell lysis and protein estimation
    e) MS analysis
5. Advantages and limitations of SILAC

BOX FOR TERMINOLOGY

SILAC: Stable isotope labeling of amino acids in cell culture - A technique for labeling the amino acids in cell culture medium, where proteins are labeled and then detected and quantified by MS.

ICAT: Isotope coded affinity tag - A technique where cysteine residues of peptides are labeled and then quantified.

iTRAQ: Isobaric tags for relative and absolute quantification - A technique where the N-terminal residues of the peptides are labeled and then detected and quantified using MS.

CDIT: Culture derived isotope tags - A technique where the cells are grown in different medium and then pooled together prior to MS analysis.

1. LIMITATIONS OF iTRAQ AND EMERGENCE OF in-vivo LABELING METHODS

Mass Spectrometry based quantitative proteomics utilizes tags during identification of proteins because of the specified difference in mass they generate. The tags can be introduced into the protein either by chemical, enzymatic or metabolic routes. Chemical introduction of tags, include iCAT (isotope coded affinity tags) and iTRAQ (isobaric tags for relative and absolute quantitation), where, different isotopic or isobaric tags are mixed with trypsin digested peptides and then the labeled peptides are analyzed. The obvious disadvantage of this approach was that some peptides might escape labeling of tag and hence get lost during identification and quantification. Also, these techniques generally depend upon specific amino acid residues, like cysteine for iCAT and primary amine for iTRAQ.

Enzymatic method of incorporating tags includes, usage of heavier oxygen in water, during enzymatic digestion of the peptides. Trypsin hydrolyzes the peptide bond and in the process adds the isotopically labeled water, thereby labeling the peptides. However, this approach does not give enough efficiency as iTRAQ.

To address the above problems, metabolic means of incorporating tags were thought of by scientists, whereby, the proteins are labeled *in vivo*. This can be achieved by labeling the media with contents required by the cells to synthesize proteins, like amino acids. Thus the cells take up the labeled amino acids and hence the proteins formed are labeled. This approach has the benefit of not depending on any particular residue for the label and hence there is no biasness for any amino acid nor the sensitivity of the instrument is a requirement, because the protein gets labeled throughout.

2. *In vivo* labeling or metabolic stable isotope labeling

Stable isotope tagging methods use isotopic nuclei (e.g. 2H, 13C, 15N and 18O) to determine the relative expression level of proteins in two samples. These methods use biological incorporation of isotopically labeled nutrients or amino acids into proteins (Fig 1).

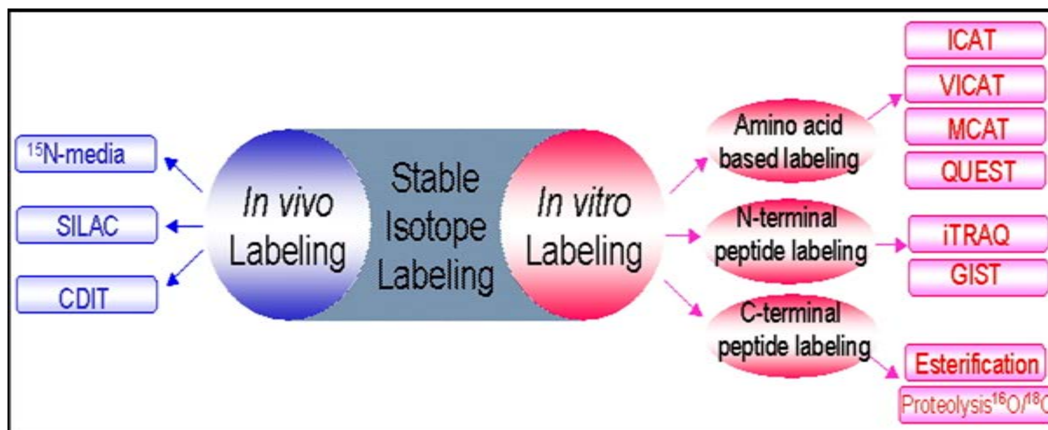

Fig 1. Stable-isotope tagging techniques, based on *in vivo* and *in vitro* labeling for differential proteomics.

^{15}N LABELING

In this method yeast or bacterial cultures are grown in two separate media, one containing 15N. Cells are pooled together, proteins extracted and quantified on MS. One of the drawbacks associated with this method is mass shift in the MS spectra. Since different proteins incorporate unequal amount of stable isotopes; therefore, labeled and unlabeled peptides exhibit variable mass shift in the MS spectra. Moreover, this method is difficult and expensive for mammalian systems, which poorly incorporates stable isotope.

CULTURE DERIVED ISOTOPE TAGS (CDIT)

For both relative and absolute quantitative proteome studies, in CDIT method cells cultured in a stable isotope-enriched medium are mixed with tissue samples to serve as an internal standard. After protein extraction and separation, digested proteins are analyzed by MS to identify and quantify peptides. The ratio between the two isotopic distributions (from a tissue sample and from isotope-labeled cells) can then be determined by using MS.

EMERGENCE OF SILAC

The usage of stable isotopes of nitrogen for labeling proteins worked well with prokaryotic systems but mammalian systems would incorporate the tag poorly. Hence the idea developed into providing isotopically labeled amino acids in the media for growth of specific mammalian cells. This led to the emergence of SILAC – stable isotopically labeled amino acids in cell culture.

3. PRINCIPLE OF SILAC

SILAC is a metabolic strategy of incorporating tags into proteins by usage of isotopically labeled amino acids in the cell culture media. It depends on the cellular protein synthesis machinery for the uptake of amino acids from the media and incorporating them into the proteome. Ideally, in SILAC a wide range of isotopes can be used. Mostly, ^{13}C is used as a stable isotope for labeling the amino acids. The usage of carbon as a stable isotopic tag was limited earlier due to the presence of carbon in other bio-molecules. But in SILAC due to the ability to specifically label amino acids, it becomes easier to use either nitrogen or carbon or oxygen as a stable isotopic tag.

As the cell grows in labeled lysine or arginine containing media, slowly the cellular machinery starts incorporating the tags into the proteins, and with natural turnover of cell and the proteins, ultimately after a few generations, the entire proteome gets labeled. As a control, cells are also grown in light media for comparison. SILAC provides a basis for absolute quantification of proteins because the tags are naturally incorporated during protein synthesis and hence there is no chance of fabrication as is the case for iCAT or iTRAQ, where the labels are added externally after protein digestion. The difference in the relative abundance of peptides between heavy and light samples is a measure of the degree of differential expression of that particular protein on external stimuli (Fig 2).

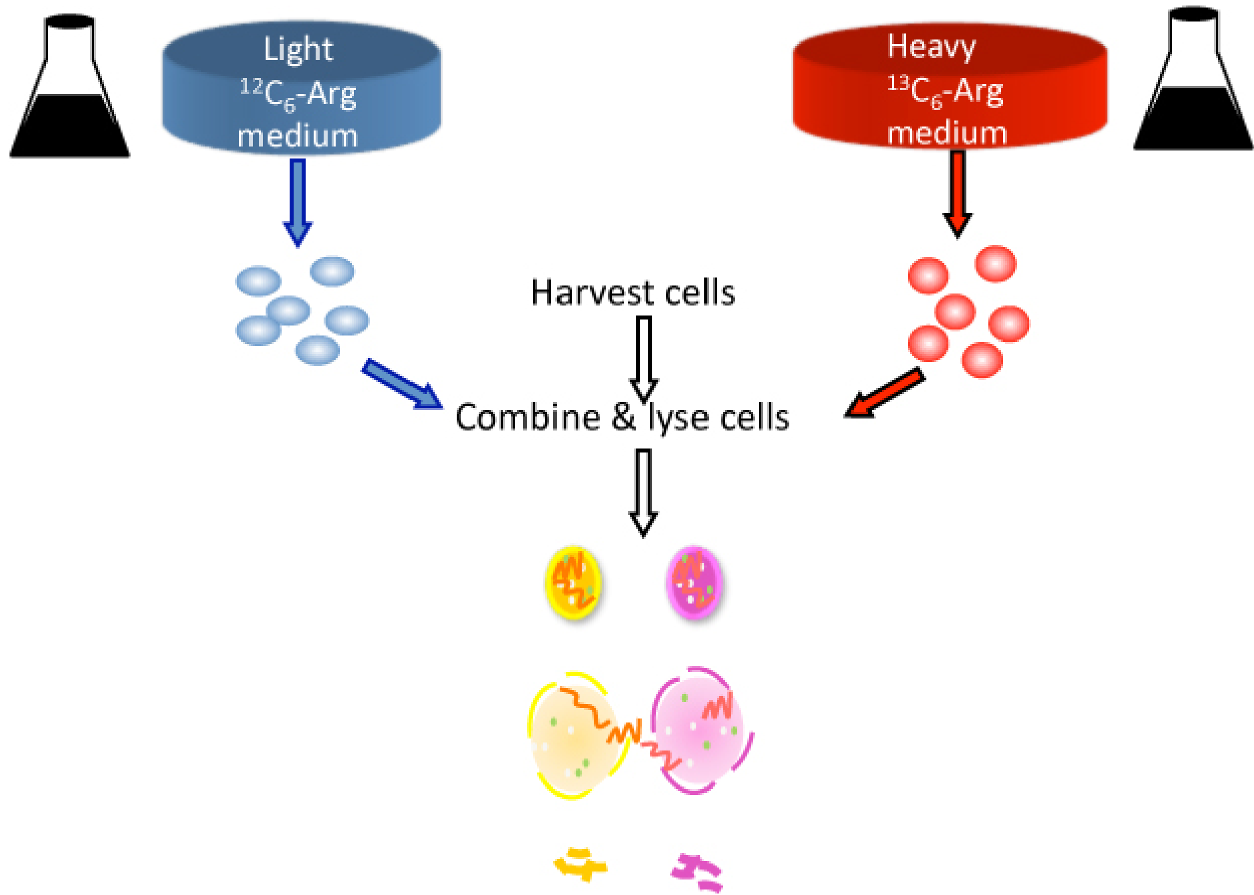

Fig 2. An overview of SILAC-based experiment

4. SILAC EXPERIMENTAL WORK FLOW

1. Preparation of SILAC labeling medium
2. Adaptation of cells: from DMEM to SILAC labeling media
3. Differential treatment applied to the SILAC cells
4. Cell lysis and protein estimation
5. MS analyses and quantitation

(a) PREPARATION OF SILAC MEDIA

The cell culture media used for SILAC strictly depends on the cells to be cultured. For most mammalian cells, DMEM or RPMI media are used. The media is prepared by usual formulations without adding the supplementary amino acids. At this stage the pre-media is divided into two parts, on one part the natural amino acids are added, while in the other one, isotopically labeled amino acids are added. Certain amino acids like lysine, arginine are isotopically labeled and added into the heavier medium, keeping all the other amino acids isotopically light. Both lysine and arginine have a backbone of six carbon atoms and hence provide a heavier mass of 6 Da to the labeled protein, which is the basis for quantification. Other components like fetal bovine serum, for growth factors and other vitamins, and antibiotics, to prevent contamination, are added to make up the final media.

There is an inherent ability of the cellular machinery to recycle some of the amino acids. Sometimes, apart from synthesizing proline by the biosynthetic pathway, some

amount of proline is also formed by breakdown of arginine. This proline to arginine conversion is inevitable and at the same time diminishes the quality of spectrum so obtained and hence hampers the data analysis. The problem can be mitigated by adding proline into the media. Still some amount of arginine to proline conversion occurs. These days software has been developed, which can exactly calculate the probability and percentage of this conversion. Accordingly, experimental conditions can be tuned in to achieve maximum reproducibility.

(b) **CELLULAR ADAPTATION**

The cells are first grown in normal DMEM or RPMI medium and then sub-cultured by splitting into two petriplates containing heavy and light media. Usually, 10-15% of the master culture is split into heavy and light media and kept for cell doublings. This is done specifically because, mammalian cells are prone to contact inhibition dependent cell density and so a minimum amount of cells are added to the media to allow the stable integration of tags into the protein by allowing at least five cell divisions before contact inhibition stops cell division.

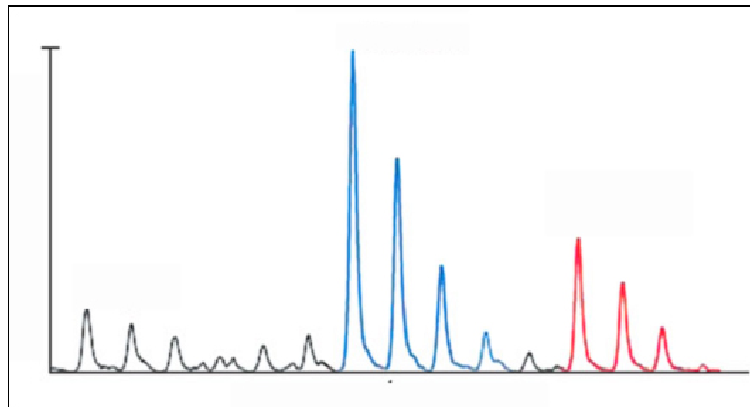

Fig 3. MS spectrum for cellular adaption

The above figure (Fig 3) shows a representative spectrum of protein isolated from the 5th generation of cell division. Like the famous DNA replication experiment by Messelson and Stahl, the time when spectrum becomes consistent and high, due to incorporation and recycling of the labeled amino acids, it is then can we conclude that stable incorporation of amino acid in the entire proteome has taken place.

(c) DIFFERENTIAL TREATMENT

SILAC and other quantitative proteomic approaches were developed to analyze the dynamic variation in protein concentration due to some external stimuli. The external stimuli can be the effect of some drugs that needs to be tested or some other physiological conditions that the protein might face. Differential treatment in SILAC is only provided when all the proteins have stably incorporated the tags into them. This cannot be manually checked, but using MS it can be ensured that after 4 - 5 generations all the proteins are tagged.

(d) CELL LYSIS AND PROTEIN ESTIMATION

Cell lysis is done by any normal procedure, specific to the cell line in question. Enzymatic lysis is preferred over any harsh method. The proteins are isolated, precipitated and then reduced using DTT (dithiothreitol) or Beta mercaptoethanol. Iodoacetamide (IAA) is added to alkylate the reduced disulfide bonds, and to prevent re-oxidation of the bonds. The proteins are then trypsinized, desalted by zip tipping and analyzed by mass spectrometry.

Illustration: SILAC experiment

SILAC is a simple method for in vivo incorporation of a label into proteins for quantitative proteomic purposes. Two groups of cells are cultured in media that are identical in all respects except that one contains a heavy, isotopic analog of an essential amino acid while the other contains the normal light amino acid.

The essential amino acids, which are obtained from the cell culture medium are incorporated into the corresponding newly synthesized proteins during cell growth and replication. Medium containing the heavy amino acids will give rise to heavy, isotopic proteins. After a number of cell divisions, all instances of the particular amino acid will be replaced by its isotopic analog. The grown cells are then combined together and harvested. Centrifugation of the mixture will result in the pelleting of cells, which can then be used for further analysis. The grown cells are then lysed using a suitable lysis buffer and the proteins degraded using a proteolytic enzyme like trypsin. This results in a mixture of light and heavy peptide fragments, which MS can quantify suitably. The complex mixture of peptide fragments is further separated by SDS-PAGE to simplify the analysis. Each band of the gel is cut out and re-dissolved in a suitable buffer solution. These simplified peptide fragments are then used for further analysis.

(e) MS DATA ANALYSIS

The peaks so obtained after MS analysis are identified by searching databases. The relative quantification or fold change in mass between the normal and treated cells is 6 Da. Hence, two closely associated peaks differing in 6 Da are invariably from the same protein (Fig 4). Since the properties of the peptides are the same, they co-elute and

hence are co-detected. The ratio of abundance of two peaks is a measure of amount of fold change that has occurred due to differential treatment.

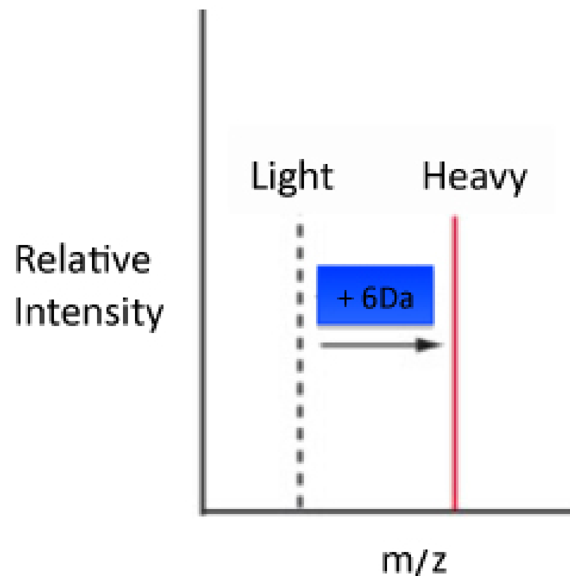

Fig 4. MS spectrum obtained from SILAC experiment

Multiplexing in SILAC is also possible with the usage of wide variety of substitution in the arginine. E.g. 5-plexing using arginine has been made possible in SILAC. The five different forms of arginine are:

a) Arginine with all the four nitrogen labeled
b) Arginine with all the six carbon labeled
c) Arginine with all the six carbon and four nitrogen labeled
d) Arginine with all the six carbon, four nitrogen and seven hydrogen labeled
e) Unlabeled arginine

These arginines vary from each other by a mass of 4, 6, 10 and 17 Da, respectively. The difference in the protein at subsequent levels are possible by taking the ratio of the various peaks.

Illustration: LC-MS/MS analysis

Further purification is carried out by liquid chromatography, wherein the sample is passed through a column containing packed stationary phase matrix that selectively adsorbs only certain analyte molecules. Reverse phase and strong cation exchange chromatography are the most commonly used. The eluted fractions are further characterized by MS.

The purified peptide fragments are analyzed by MS/MS. Peptides containing the heavy amino acid show higher m/z than the corresponding light peptide fragments. The pairs of identical peptides can be differentiated due to the mass difference and the ratio of peak intensities can be correlated to the corresponding protein abundance.

Illustration: SILAC data analysis

The MS/MS data analysis shareware has some extra inputs such as Quantitation, MS/MS tolerance, peptide charge, instrument etc. in addition to the fields for PMF. They require inputs from the user regarding the experimental parameters used such as enzyme cleavage, protein name, modifications 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.

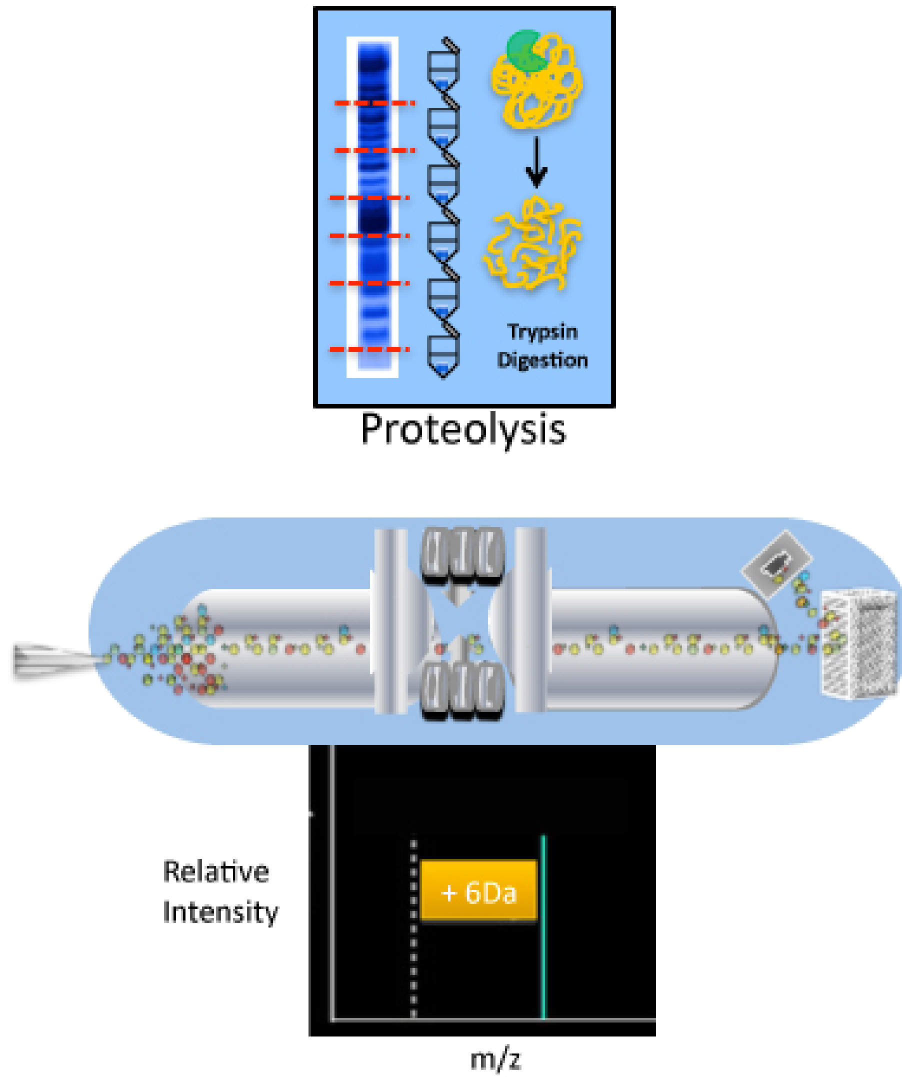

Fig 6. A typical work-flow of SILAC

5. ADVANTAGES AND LIMITATIONS OF SILAC

SILAC is extremely efficient in incorporating tags, unlike chemical methods like iCAT and iTRAQ. Since, SILAC involves the stable integration of tags into the proteins naturally, there is 100% incorporation of tags. However, to ensure that the tags are incorporated into the entire proteome, at least 4 to 5 generations of cell divisions is necessary. In addition, proteome analysis of every generation needs to be done to be on the safer side. This increases the cost and the time for a particular SILAC experiment.

Since the cells are grown in exactly the same medium, there is less chance of variation, because it is only the amino acids in the medium that varies. Also, since the samples are pooled before quantification, there is less chance of handling errors. The ability for multiplexing has also increased the high-throughput level of protein quantification.

The biggest disadvantage of SILAC is its restriction to only culturable cells. Serum or other body fluids hence cannot be assessed by SILAC. This limits the usage of SILAC in clinical proteomics, especially in biomarker discovery. Another disadvantage of SILAC is requirement for full incorporation of amino acids in the proteome for analysis, unlike radioactivity. Nevertheless, SILAC finds excellent applications in various *in vivo* applications and hence is an extremely valuable tool despite its limitations.

REFERENCE

1. Proteomics- Introduction to Methods and Applications. Ed. agniezka kraj and jerzy silberring. wiley publication, 2012.
2. Harsha, H. C., H. Molina, and A. Pandey. "Quantitative proteomics using stable isotope labeling with amino acids in cell culture." Nat.Protoc. 3.3 (2008): 505-16.
3. Ong, S. E. and M. Mann. "A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC)." Nat.Protoc. 1.6 (2006): 2650-60.
4. Sato, T., Y. Ishihama, and Y. Oda. "Quantitative proteomics of mouse brain and specific protein-interaction studies using stable isotope labeling." Methods Mol.Biol. 359 (2007): 53-70.
5. Yates, J. R., C. I. Ruse, and A. Nakorchevsky. "Proteomics by mass spectrometry: approaches, advances, and applications." Annu.Rev.Biomed.Eng 11 (2009): 49-79.