LECTURE-18

Isotope- Coded Affinity Tagging (ICAT)

HANDOUT

PREAMBLE

The versatility of Mass spectrometry in the field of proteomics led to the emergence of a separate branch of proteomics study – quantitative proteomics. Quantitative proteomics deals with the actual quantification of proteins under given conditions. Need for quantitative proteomic analysis and combination of mass spectrometry led to the emergence of techniques like ICAT (Isotope coded Affinity Tagging), ITRAQ (Isobaric tagging for Relative and Absolute Quantitation) and SILAC (Stable amino acid labeling in cell culture), with ICAT emerging first and the others later. All these techniques employed labeling proteins in some form or the other and then identifying and quantifying them by the tags attached to them. In this lecture we will discuss about ICAT technique and highlight some of its applications in clinical diagnosis.

OUTLINE OF LECTURE

1. Need for gel-free based quantitative proteomics
2. Principle of ICAT
3. ICAT reagent
4. Working protocol of ICAT
5. Data analysis
6. Clinical applications
7. Advantages of ICAT
8. Limitations of ICAT

PROF. SANJEEVA SRIVASTAVA
DEPARTMENT OF BIOSCIENCES & BIOENGINEERING
IIT BOMBAY, MUMBAI, INDIA

BOX FOR TERMINOLOGY

ICAT: Isotope Coded Affinity tagging - ICAT is an *in vitro* labeling technique that modifies peptides or proteins specifically at the cysteine amino acid residue and can be used for accurate quantitation of protein expression.

Light ICAT label: The light ICAT reagent consists of a Cys-reactive group, an ICAT linker consisting of hydrogen atoms and a biotin tag. The chemically reactive group forms covalent bonds with peptides or proteins while the affinity tag enables the protein to be isolated by affinity chromatography in a single step.

Heavy ICAT label: The heavy ICAT label consists of a Cys-reactive group, an ICAT linker consisting of heavy deuterium isotope and a biotin tag.

Affinity purification: A chromatographic purification procedure that makes use of specific interactions between the analyte of interest and the capture reagent immobilized on the column. In ICAT, avidin affinity chromatography is employed due to its specificity of interaction with biotin.

ITRAQ: Isobaric Tagging for Relative and Absolute Quantitation - A technique used to quantify proteins by labeling N terminal residues with a tag.

SILAC: Stable Isotope labeling of Amino Acids in Cell culture - A technique of labeling proteins by providing tagged amino acids in cell culture, so that the cell take them up and label the proteins metabolically.

1. NEED FOR GEL-FREE BASED QUANTITATIVE PROTEOMICS

The term 'gel-free quantitative proteomics' essentially refers to identification and quantification of proteins from complex protein mixtures facilitated by MS based quantitative proteomic approaches (Fig 1). To fully understand the role of a particular protein in the physiology of a cell or an organism, the dynamic changes in the concentration of that protein becomes extremely important. Potential errors in quantitative analysis arise from difference in sample preparation and procedures of sample analysis. The greater the number of steps, greater will be the error, as the error gets multiplied at every stage. Gel-based quantitative proteomics is highly error-prone because of these problems. Gel-based proteomics faces the following drawbacks:

a) Inability to characterize entire proteome
b) Ability to resolve low copy number proteins
c) Reproducibility amongst various gels
d) Membrane proteins get under-represented

The most advanced gel-based quantitative proteomics is DIGE (Difference in Gel Electrophoresis), where different samples are run on the same gel and tagged with different dyes. However, DIGE can not overcome some of the inherent drawbacks of gel-based approaches like limited proteome coverage. Various drawbacks of gel-based quantitative proteomics led to the requirement of other sophisticated yet simple gel-free based proteomic approaches.

Classical techniques for quantitation involves ELISA (Enzyme Linked Immuno-sorbent Assay) or RIA (Radio-immuno Assay), but they are subject to the limitation of

studying one protein at a time because of the requirement of monoclonal antibodies. Thus there is limitation of assays based on availability of antibodies. Highly sophisticated techniques involve MS and microarray. However, protein microarrays are also extremely cumbersome and costly. MS-based quantitative proteomics is the most feasible and cost-effective approach that scientists all over the world rely upon.

MS-based quantitative proteomics is not only able to answer the dynamic changes in protein concentration subject to various external stimuli, but also identifying many biomarkers for diagnosis and identifying target sites for action of drugs. The mass range and sensitivity of a mass spectrometer is able to resolve and quantify even attomoles of proteins that might be present in the sample. In short, MS based quantitative proteomics relies upon tagging proteins with exclusive tags, which are identified during MS analysis and their relative quantification is a measure of the quantity of the protein.

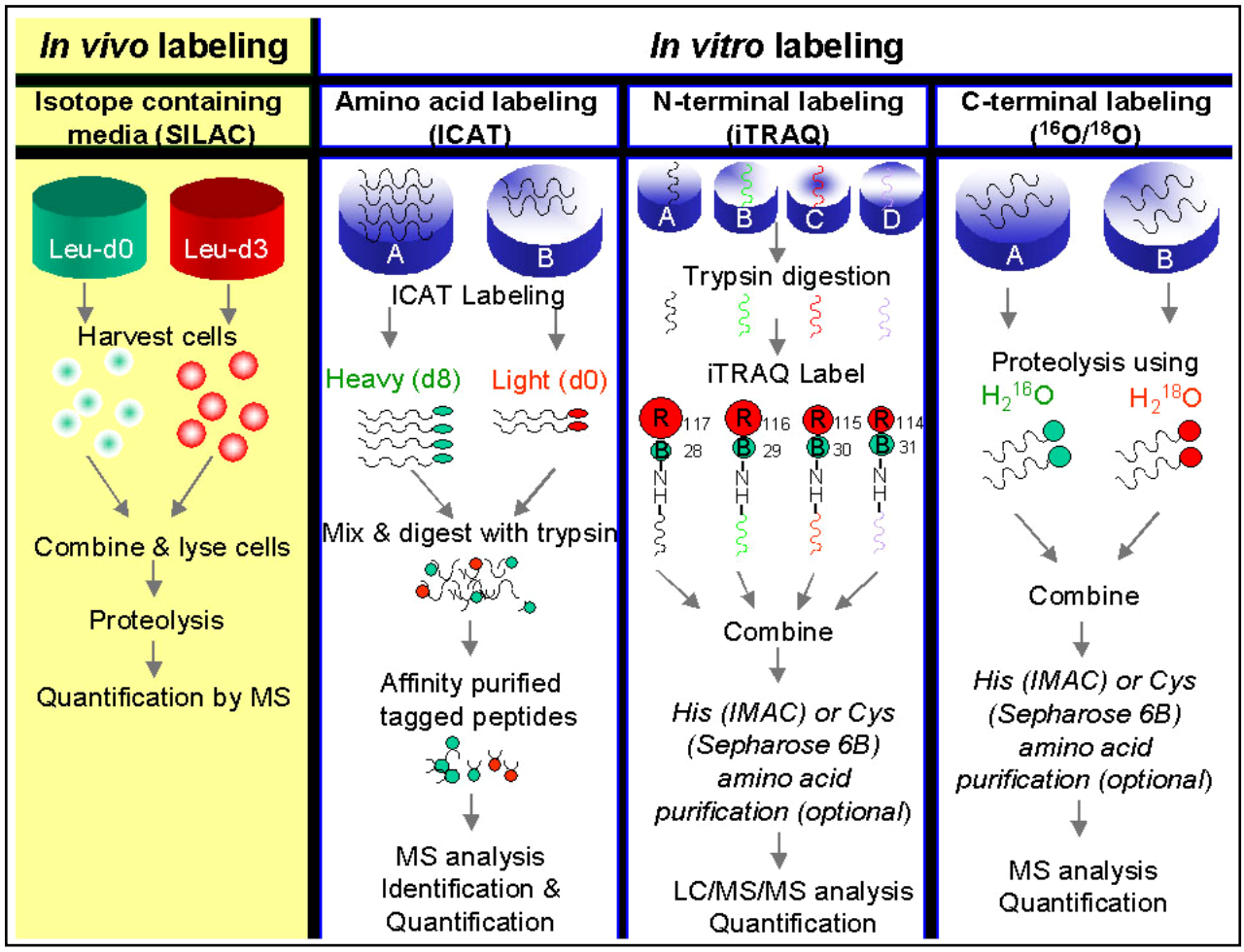

Fig 1. An overview of stable-isotope tagging methods. SILAC uses *in vivo* labeling whereas ICAT, iTRAQ, $^{16}O/^{18}O$ uses *in vitro* labeling. ICAT enables tagging of specific proteins chemically in two separate samples with isotopes. ICAT relies on tagging cysteine residues whereas in iTRAQ method tagging is on primary amines.

2. **PRINCIPLE OF ICAT**

Isotope Coded Affinity Tagging (ICAT) is an *in vitro* method of tagging proteins with specific tags. *In vitro* labeling relies on use of labeling reactions at specific sites in proteins and peptides. *In vitro* tagging can be classified into various sub-categories based on the chemistry involved. Ideally, there are three broad *in vitro* labeling methods:

a) Specific amino acid tagging: Example: ICAT
b) N – Terminal tagging: Example: ITRAQ
c) C – Terminal tagging: Example: O^{18} labeling by proteolysis

ICAT is an example of amino acid tagging. Cysteine is an amino acid present in approximately all proteins and comprises about 4% in proteins. Using this information, tags have been generated, which specifically interact with the cysteine residues of the protein and hence tag them. ICAT was first developed by Gygi et al in 1999, as a method for absolute quantitation of protein expression in *Saccharomyces cervisiae* subjected to growth in ethanol or galactose. The tags differ from each other in a particular mass and hence when MS analysis is done, the same protein from two different conditions give peaks at two different m/z values. The ratio of the intensity of the two peaks is a measure of the differential protein expression in response to the external stimuli.

3. ICAT REAGENT

The ICAT method is based on a chemical reaction between the cysteine residues of the protein and a chemical tag. Hence the tag should have a reactive group that can specifically react with the SH group of cysteine and shows no other cross reactivity. The ICAT reagent consists of three parts (Fig 2):

a) Iodoacetamide residue (chemically reactive group), which binds to cysteine thiol group.

b) Linker region, containing hydrogen or deuterium, imparting the increase in mass.

c) Biotin residue, which helps in the enrichment of peptides using Streptavidine assisted affinity chromatography.

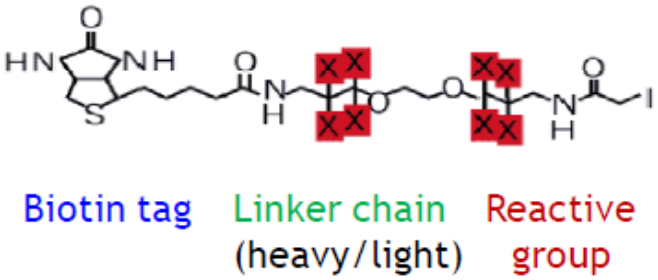

Fig 2. A general structure of the ICAT reagent

The ICAT reagent comes in two forms: one containing no deuterium while the other with eight deuterium atoms. The ICAT reagent binds to the cysteine groups of the protein via the iodoacetamide residue and depending on the number of deuterium ion, there is an increase in the mass of the peptide.

4. WORKING PROTOCOL OF ICAT

ICAT is not only used for quantifying proteins but also for studying the degree of expression of proteins subject to external stimuli. Ideally, the isolated proteins are first treated with the ICAT reagent in either of the two ways:

a) Control with ICAT reagent having no deuterium and treated with ICAT reagent containing eight deuterium
b) Tag swapping

Ideally, both the type of tagging should be done to reproduce the data so obtained. The tagged proteins are then trypsinized. Since cysteine is present in approximately 4% in approximately all proteins, it is assumed that the multiple peptides generated from a protein will get tagged at least multiple times if not all. Since the ICAT tag also contains a biotin residue, using Streptavidine based affinity chromatography; the peptides are enriched and then subjected to MS analysis and MS/MS analysis for identification (Fig 3). The ratio of the peaks of the control and the treated peptides is a measure of the differential expression of the protein.

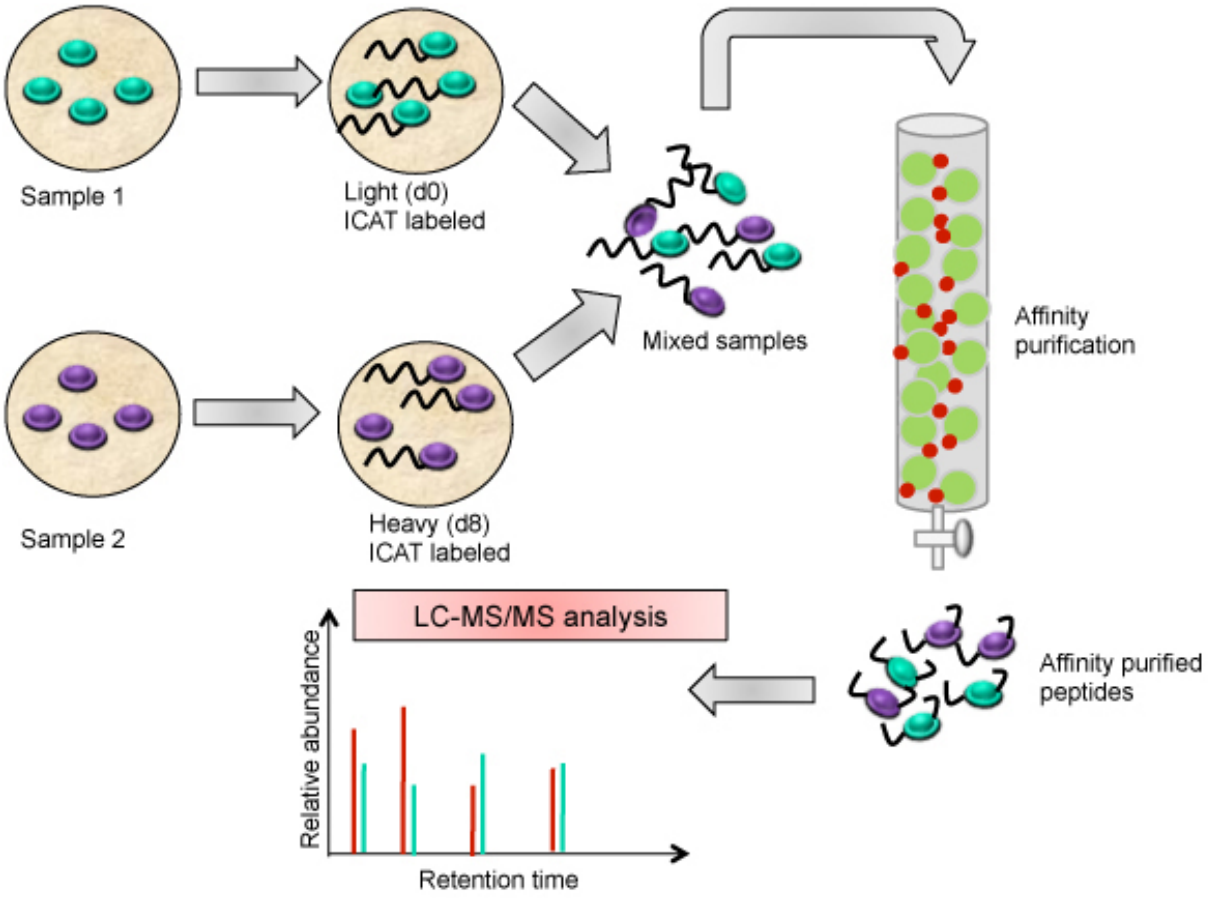

Fig 3: Working scheme for ICAT

Illustration: ICAT sample labeling

ICAT is an in vitro labeling procedure that involves tagging of protein or peptide samples with the ICAT reagent specifically at their Cys residues. The ICAT reagent consists of a biotin tag, a light or heavy linker chain and a Cys-reactive group. One sample is tagged with the light ICAT reagent while the other is tagged with heavy ICAT.

The labeled samples are mixed together and then cleaved enzymatically to generate peptide fragments, some of which will contain the ICAT tag. These are purified by means of affinity chromatography, which makes use of the specific interaction between immobilized avidin on the column and biotin of the linker.

Illustration: ICAT LC-MS/MS analysis

Further purification of the affinity purified peptides is carried out by 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 peptide fragments are then analyzed by MS/MS. Both the quantity and sequence identity of the proteins from which the tagged peptides were obtained can be determined. This technique can be used for complex protein mixtures.

5. DATA ANALYSIS

The linker region in the ICAT reagent accounts for the difference in the peak obtained for the same protein subjected to two different experimental conditions. The linker region contains either no or eight deuterium atoms. Hence the shift in the peptide peak occurs after 8 Da (Fig 4). At a large MS scale, only one peak is observed. Zooming into the peak gives two subsequent peaks, which are 8 Da apart. The ratio of intensity of the peaks is a measure of the amount of differential expression.

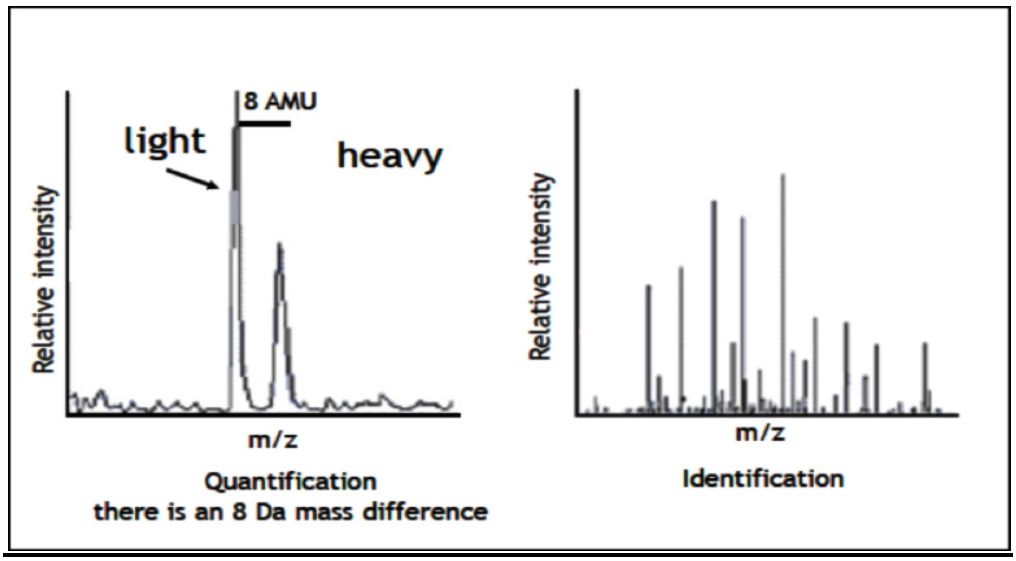

Fig 4. A typical MS spectrum of ICAT

The above MS spectrum is a zoomed in view of a single peak. As is seen in the above spectrum, the two peaks differ from each other by 8 Da. For identification of the protein, MS/MS is performed. The peak is chosen and selected for further fragmentation into the second MS. The sequence so obtained from the second MS is searched against the list of proteins available on the database and hence the identity of the protein is established.

Illustration: ICAT LC-MS/MS 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.

6. CLINICAL APPLICATIONS

The ability to quantify and hence compare the dynamic concentration of proteins under two conditions (normal vs diseased) is one of the most promising clinical applications of ICAT. The ICAT technique was first quantitative proteomic technique via mass spectrometry to be developed and it revolutionized the field of clinical proteomics. Biomarker discovery in the field of clinical proteomics is an important application (Fig 5) and ICAT due to its high sensitivity has the ability to detect low levels of proteins in all types of samples. However, the inability to multiplex, reduced the ability of ICAT to be used in monitoring the progress of drug action in some diseased states.

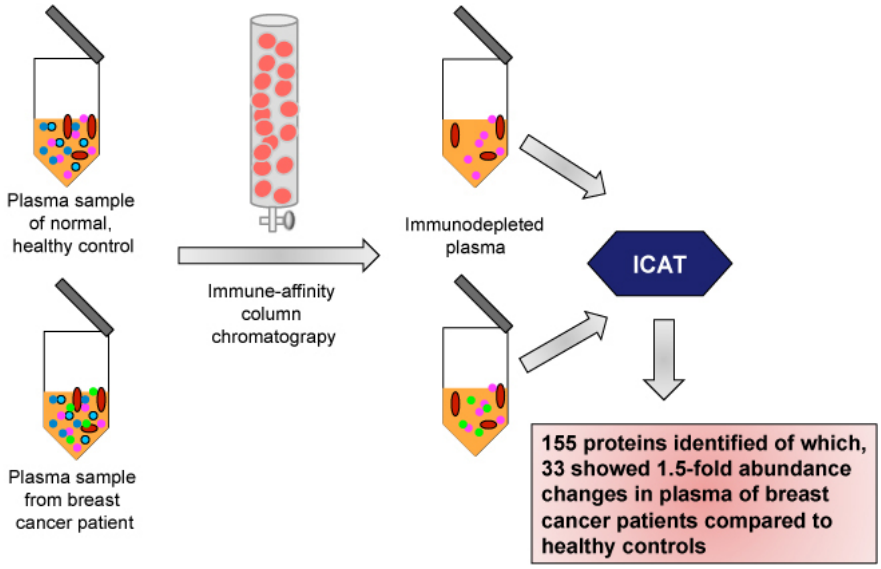

Fig 5. ICAT application – breast cancer plasma proteome analysis

Illustration: Application of ICAT

Kang et al. obtained plasma proteomes from 6 breast cancer patients and 6 healthy controls. These plasma samples were first treated on an immune-affinity column in order to deplete them of their high abundance proteins. The immunodepleted serum samples were then used for further analysis by ICAT. 155 proteins were identified of which 33 showed 1.5-fold abundance changes in plasma of breast cancer patients as compared with the healthy controls. Biotinidase was found to be significantly down-regulated in breast cancer patients.

7. ADVANTAGES OF ICAT

The biggest advantage of ICAT was the ability to quantify proteins with reasonable accuracy of within 10%. ICAT is based on quantifying proteins by labeling the cysteine residues with specific tags. Almost 90% of all proteins contain cysteine residues and that is in almost 4% of the protein. Thus the rationale behind labeling cysteine and quantifying protein was novel and highly sophisticated. ICAT can thus be used to quantify proteins from complex protein mixtures, which gel-based technique could not do. The inclusion of Mass spectrometry helps in protein identification also simultaneously, which was not possible with gel-based techniques.

Overall, the major advantages of ICAT involve its high specificity and sensitivity and its applicability to all proteins from all sources. Moreover, the ICAT reagent is insensitive to reagents like guanidine, SDS or urea, which hampers labeling in DIGE. Also, the entire procedure is extremely automated and simple, thus giving an edge over conventional gel-based proteomics.

8. LIMITATIONS OF ICAT

Despite its several advantages, ICAT has several limitations, which led to emergence of other techniques like iTRAQ and SILAC. The biggest disadvantage of ICAT is the requirement of cysteine residues for labeling. Although almost all proteins contain cysteine residues, the biasness for cysteine affects quantification results.

Another disadvantage of ICAT reagent was its huge size. This interfered with the usual fragmentation pattern of the peptides, because the peptides are fragmented by trypsin post-labeling. Also, the presence of the large tag lead to further anomalous fragmentation, which might lead to loss of the peptide ions. To overcome this problem, cleavable ICAT reagents were introduced, where the biotin tag and the iodoacetamide groups were separated by a spacer of nine carbon atoms.

Another disadvantage of ICAT is availability of only two reagents, hence multiplexing is not possible. Hence its usage in clinical proteomics becomes restricted. Also, the reagent is extremely costly, adding to another limitation. All these limitations of ICAT led to the emergence of more quantitative techniques like iTRAQ and SILAC which depend on labeling N-terminal residues of proteins and amino acids in cell culture, respectively. The problem of multiplexing was also solved using these techniques as different kinds of reagents were available.

Nonetheless, the usage of ICAT still remains important in case of quantitative proteomics and is used for several applications.

Illustration: ICAT-advantages and limitations

ICAT is a promising quantitative proteomic technique that can be useful for determining complex protein mixtures. It is highly automated allowing peptides to be sequenced directly using tandem MS. However, the major limitation for ICAT is its bias for cysteine rich proteins. The large ICAT reagent often interferes with MS fragmentation and reduces the quality of the MS spectra obtained. This limitation has however, been successfully overcome through development of a cleavable ICAT (cICAT) reagent, which is removed prior to MS fragmentation and analysis.

REFERENCE

- Proteomics- Introduction to Methods and Applications. Ed. Agniezka kraj and jerzy silberring. Wiley publication, 2012.
- Gygi, S. P., et al. "Quantitative analysis of complex protein mixtures using isotope-coded affinity tags." Nat.Biotechnol. 17.10 (1999): 994-99.
- Jiang, Y. and X. Wang. "Comparative mitochondrial proteomics: perspective in human diseases." J.Hematol.Oncol. 5 (2012): 11.
- Kapoor, I., et al. "Proteomics approaches for myeloid leukemia drug discovery." Expert.Opin.Drug Discov. 7.12 (2012): 1165-75.
- Gygi, S. P. et al., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotech.* 1999, 17:994-999.
- Kang, U. B. et al., Differential profiling of breast cancer plasma proteome by isotope-coded affinity tagging method reveals biotinidase as a breast cancer biomarker. *BMC Cancer* 2010, 10:114.
- Rivera-Monroy, Z. et al., Fluorescent isotope-coded affinity tag (FCAT). I: Design and synthesis. *Bioorg. Chem.* 2008, 36(6):299-311.
- Gasparri C. et al., Proteomics reveals high levels of vitamin D binding protein in myocardial infarction. Front Biosci (Elite Ed). 2010, 2:796-804.
- Butler, G. S. et al., Identification of cellular MMP substrates using quantitative proteomics: isotope-coded affinity tags (ICAT) and isobaric tags for relative and absolute quantification (iTRAQ). Methods Mol. Biol. 2010, 622:451-70.