

# PC235: 2008 Lecture 5: Quantitation

Arnold Falick

[falickam@berkeley.edu](mailto:falickam@berkeley.edu)

# Summary

What you will learn from this lecture:

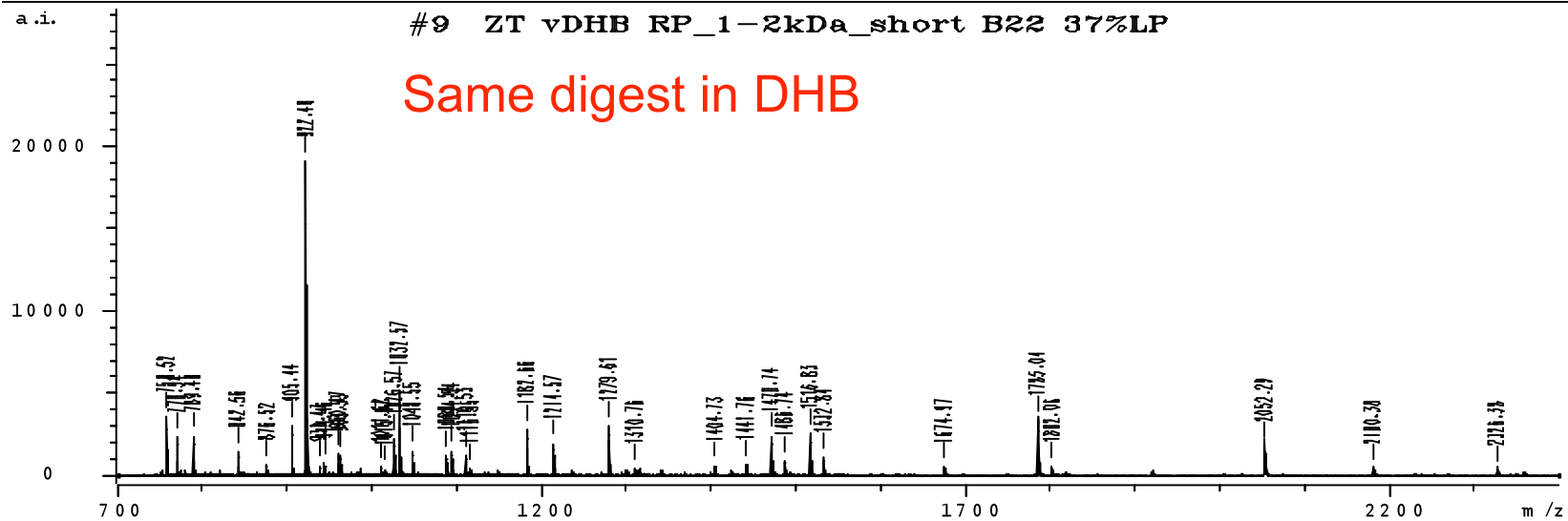
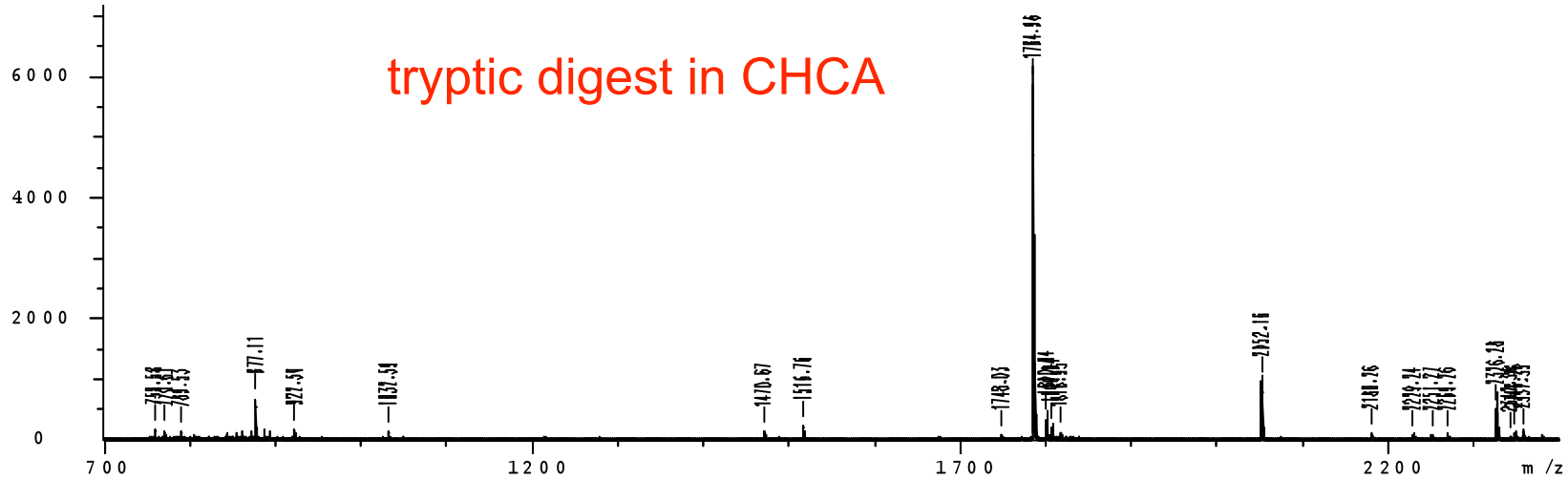
- There are many methods to perform quantitation using mass spectrometry (any method worth mentioning has a trendy acronym)
- Quantitation can be used for studying protein changes, modification changes or as a method for identifying specific results.
- Quantitation is normally relative (between samples), but can also be absolute.
- Software for analysis of this type of data is still being developed.

# Mass Spectrometry is not Inherently Quantitative

- There are a number of factors that contribute to peak intensity in a mass spectrum
  - Absolute amount of component
  - Amount of component in relation to other components in the sample
  - Ease of ionization of component
- Components that are of the same concentration may give drastically different signal intensities
  - E.g., if a protein is digested into peptides, different peptides give very different signal intensities (and some peptides may not be detected at all!)

# Mass Spectrometry is not Inherently Quantitative

- The same sample can look very different based on the analysis technique



# How Can Mass Spectrometry be used for Quantitation?

Different types of quantitation can be performed using mass spectrometry

- Relative quantitation can be performed for chemically similar components present in the same sample
- Relative quantitation can be performed by comparing peak intensities of the same component in similar samples analyzed sequentially.
- Comparison to a similar component at a known concentration can be used to estimate absolute quantities.
- Absolute protein quantities can be crudely estimated without comparison.

# MS Methods used to Produce Quantitative Data

## Stable Isotope Labeling:

- Incorporation of stable isotopes, samples are mixed and then compared to each other.

[Isotopes can be incorporated metabolically or chemically]

## Label-free Quantitation:

- Samples are run sequentially and either peak intensities or numbers of peptides identified from a protein are compared.

# Enzymatic Incorporation of Stable Isotopes

- Serine proteases; e.g., trypsin, will exchange carboxyl terminal oxygens with solution.
  - Incubating a sample in  $\text{H}_2^{18}\text{O}$  will cause incorporation of two oxygens, producing a 4Da mass shift.
  - Mix unlabeled and labeled samples together and analyze.

## Advantages

- Labels all peptides.
- Does not affect peptide fragmentation.
- Can be useful for peptide identification as all C-terminal ions (e.g., y-ions) will give pairs of peaks 4 Da apart.

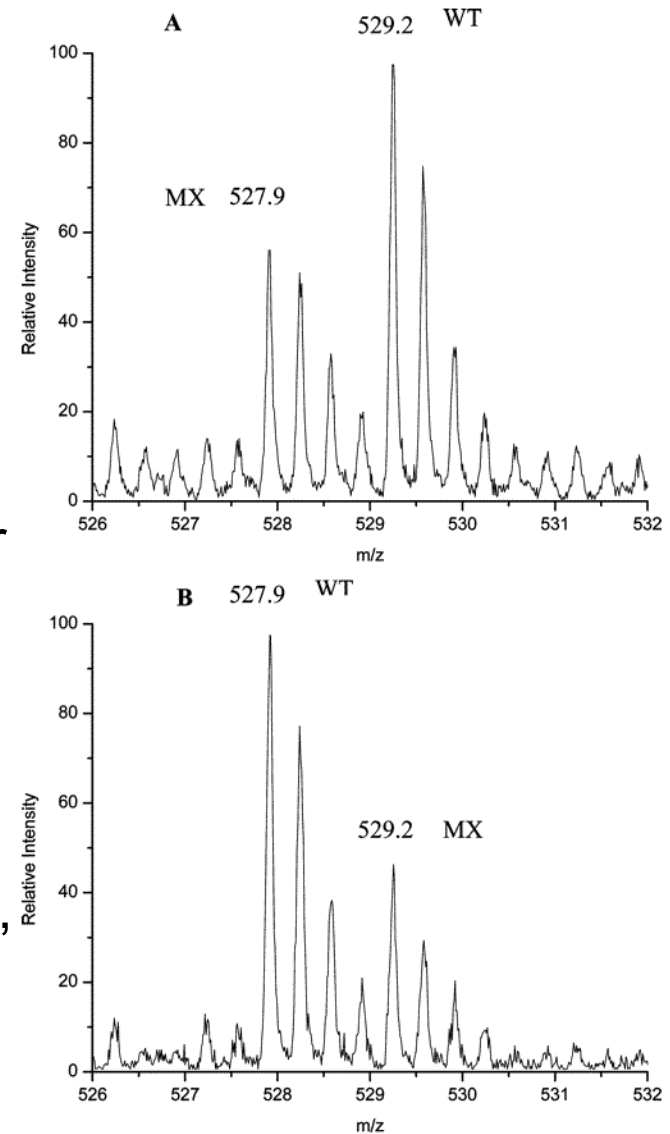
# $^{16}\text{O}/^{18}\text{O}$ Labeling

## Problems

- Somewhat difficult to get complete incorporation of both oxygens.
- Label will slowly exchange back once returned to normal water solution.
- 4 Da is a relatively small mass shift; can get overlap of light and heavy isotope clusters for large peptides.

## In Summary

A quick, easy and cheap quantitation method, but not the most accurate.



# ICAT (Isotope-Coded Affinity Tag)

Affinity Tag (Biotin)	Linker Region Light: 8 $^1\text{H}$ Heavy: 8 $^2\text{H}$	Cysteine Reactive Group (Iodoacetamide)
-----------------------	---	--

## Advantage:

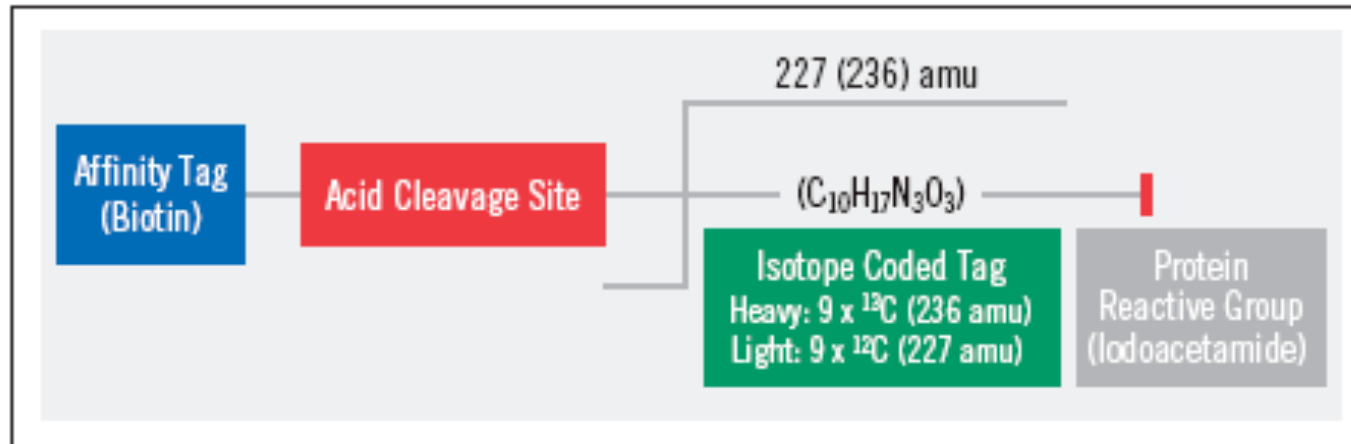
Can selectively enrich labeled peptides.

## Problems:

- Light and heavy reagents do not co-elute in reverse phase, making quantitation of LC-MS data more complicated.
- Protein must contain Cys
- The biotin tag causes problems:
  1. Difficult to elute quantitatively from streptavidin
  2. Tag yields intense fragment ions
    - Fragment biotin rather than peptide
    - Often difficult to identify peptide

Gygi, S.P. et al. *Nat. Biotechnol.* (1999) **17** 994-999

# cICAT (cleavable ICAT)



1. Affinity enrich modified peptides.
2. Cleave tag using TFA, removing the biotin.
  - The smaller isotopically labeled tag fragments less than the whole protein, making peptide identification simpler.
  - Light and heavy reagents co-elute, making quantitation easier.

## Issue

- Most proteins are identified by a single peptide, requiring high accuracy and reproducibility in measurements to be able to draw conclusions.

Hansen, K. C. et al. *Mol Cell Proteomics* (2003) **2** 299-314

# SILAC (Stable Isotope Labeling of Amino Acids in Culture)

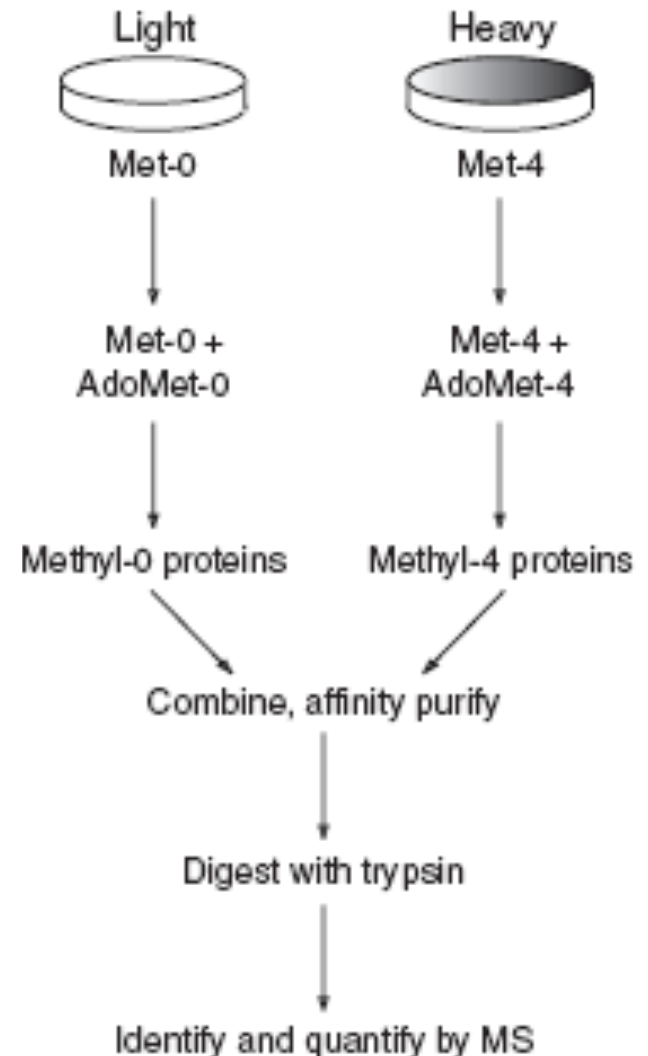
In cell culture, replace one (or two) amino acids in the culture medium with  $^{13}\text{C}$  and/or  $^{15}\text{N}$  equivalent.

- Which amino acid/s should you choose?
  - You want to label as many peptides as possible
  - It needs to be an amino acid that is big enough to give a mass shift of at least 4Da.
  - For tryptic digests, lysine (C6, N2) and/or arginine (C6, N4) are sensible choices.
- Need to be aware that some amino acids can be synthesized by the cell, or can be converted from one amino acid to another.
  - Safest bet is to keep cells well fed, so there is never a shortage of any amino acid.

Ong, S-E. et al. *Mol Cell Proteom.* (2002) **1** 5 376-386

# SILAC for Modification Quantitation

- Also possible to label modifications through addition of a labeled version of the donor.
- S-adenosyl methionine (AdoMet) is the sole donor of methyl groups for protein (and DNA, RNA) methylation.
- It is synthesized from methionine
- By adding  $^{13}\text{C}^2\text{H}_3$ methionine in culture one can label lysine and arginine methylations e.g. in histones.



# Multiplexing SILAC

- Three samples can be compared in the same experiment.

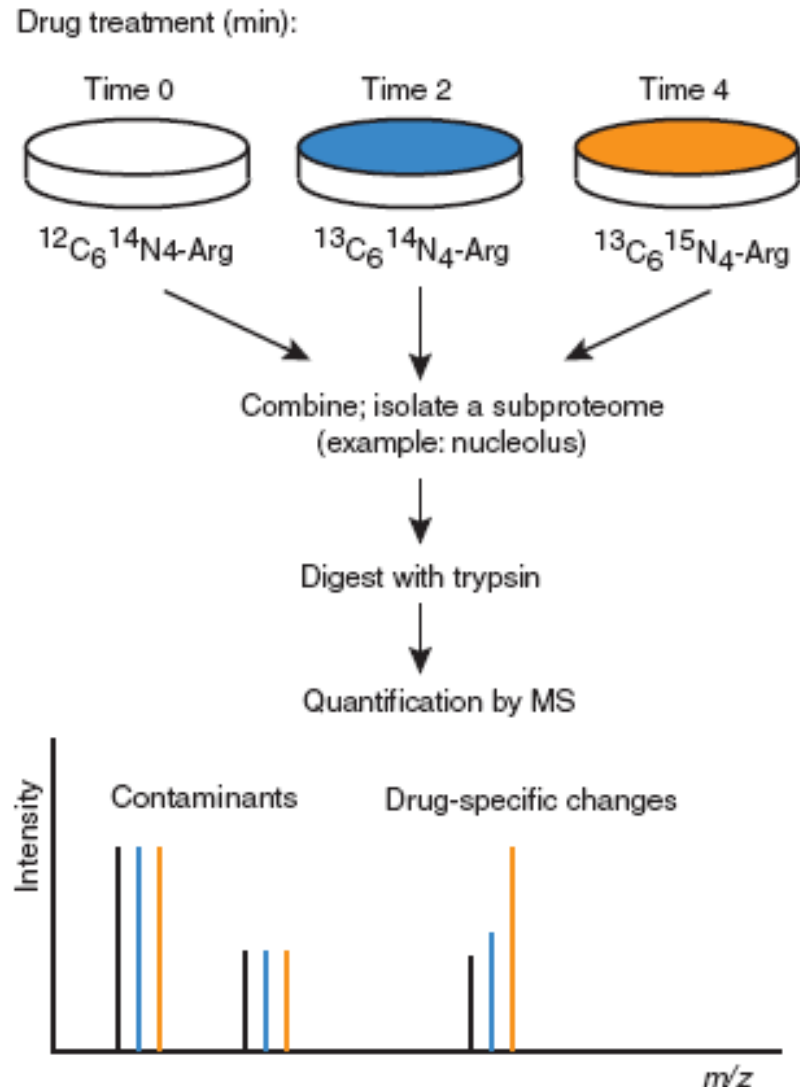
## SILAC Summary

### Strengths:

- Can get essentially complete labeling (need about 5 population doublings for complete incorporation).
- Labeling does not affect fragmentation.

### Weakness:

- Requires cell culture.



# Heavy Methyl Esterification

- Can methyl esterify side-chain carboxylic acids
  - Used to reduce non-specific binding in IMAC for phosphopeptide enrichment.
- Can use normal and deuterated methanol, mix samples and compare.

## Problem

- Reaction is sometimes difficult to get to completion.
- 3 Da mass shift is smaller than ideal (get overlap of peptide isotope patterns).

# Amine-Reactive Labels

•Deuterated versions of many reagents have been used for labeling peptide N-termini and lysine side-chains.

E.g.:

- Acetic anhydride
- Propionic anhydride
- Succinic anhydride
- Phenyl isocyanate

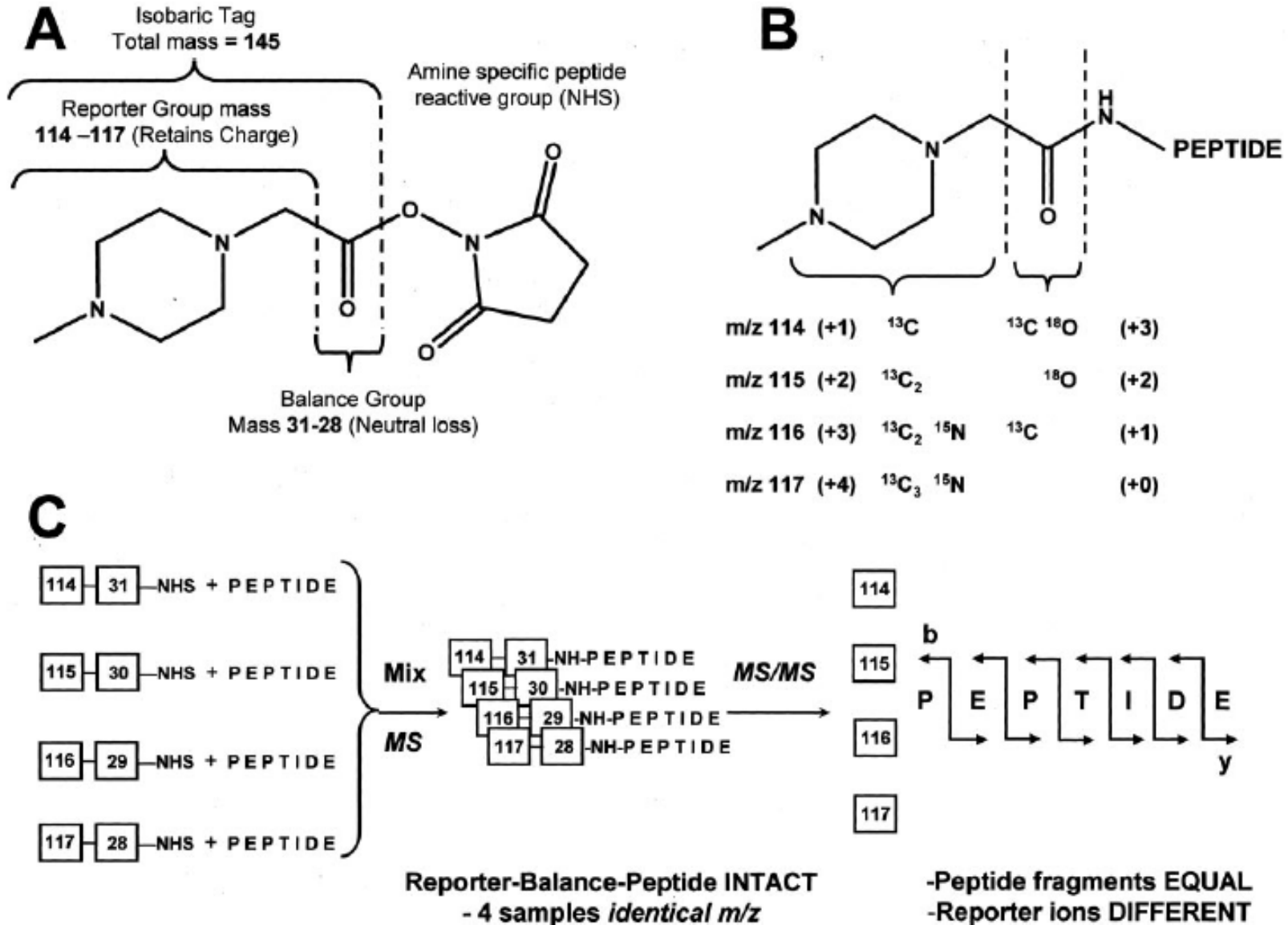
# Isotopic Tag for Relative and Absolute Quantitation (iTRAQ)

The cleverest and most elegant amine-reactive tagging reagents.

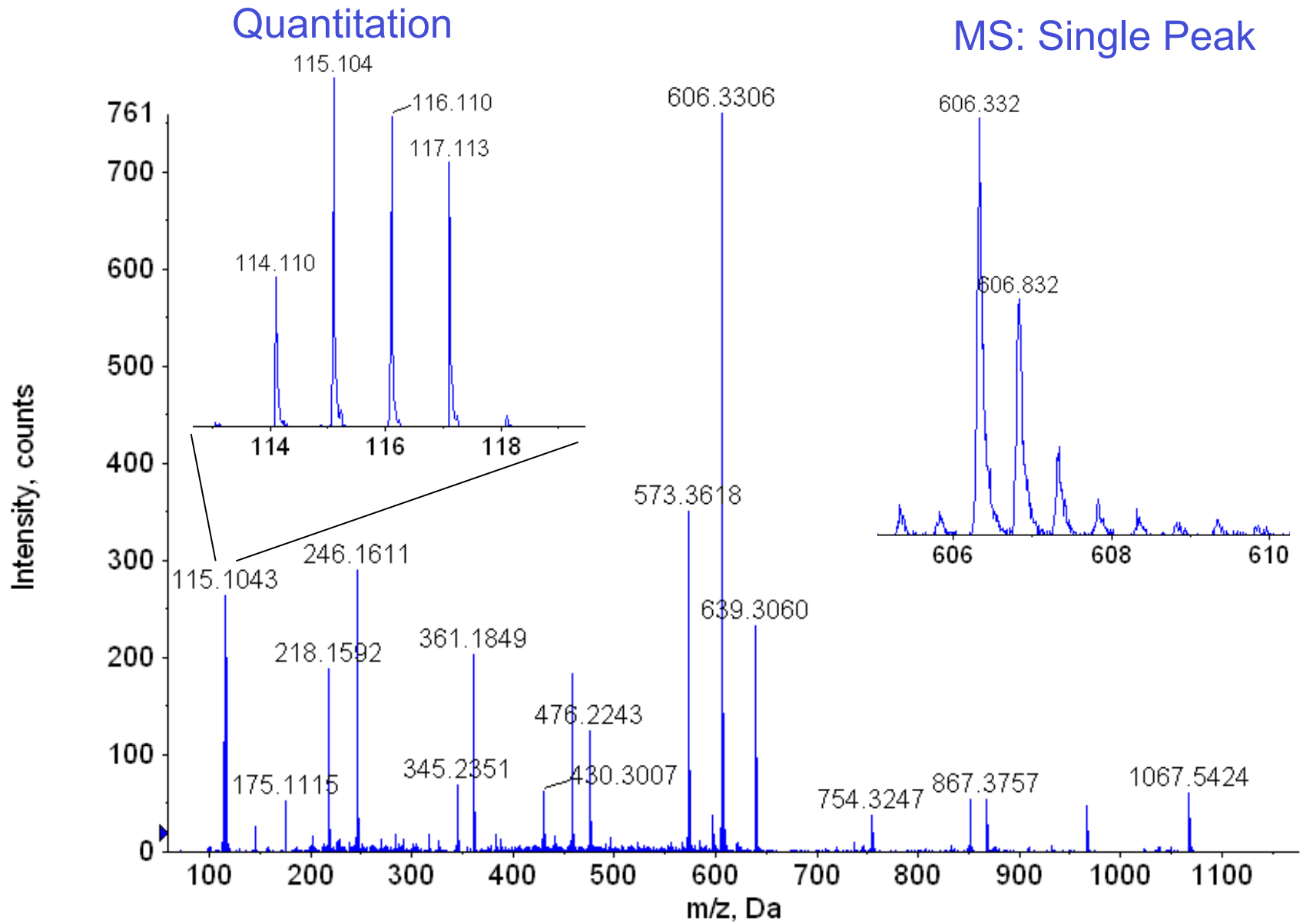
- Amino-groups are modified (N-terminus, Lys)
- A set of four reagents, allowing simultaneous comparison of four samples.
- All reagents attach nominally same mass tag, giving one peak in the MS, but produce different 'reporter ions' in MSMS, allowing relative quantitation.

Ross *et al.* *Mol Cell Proteomics* (2004) **3**, 1154.

# iTRAQ Reagents



# Example iTRAQ Data

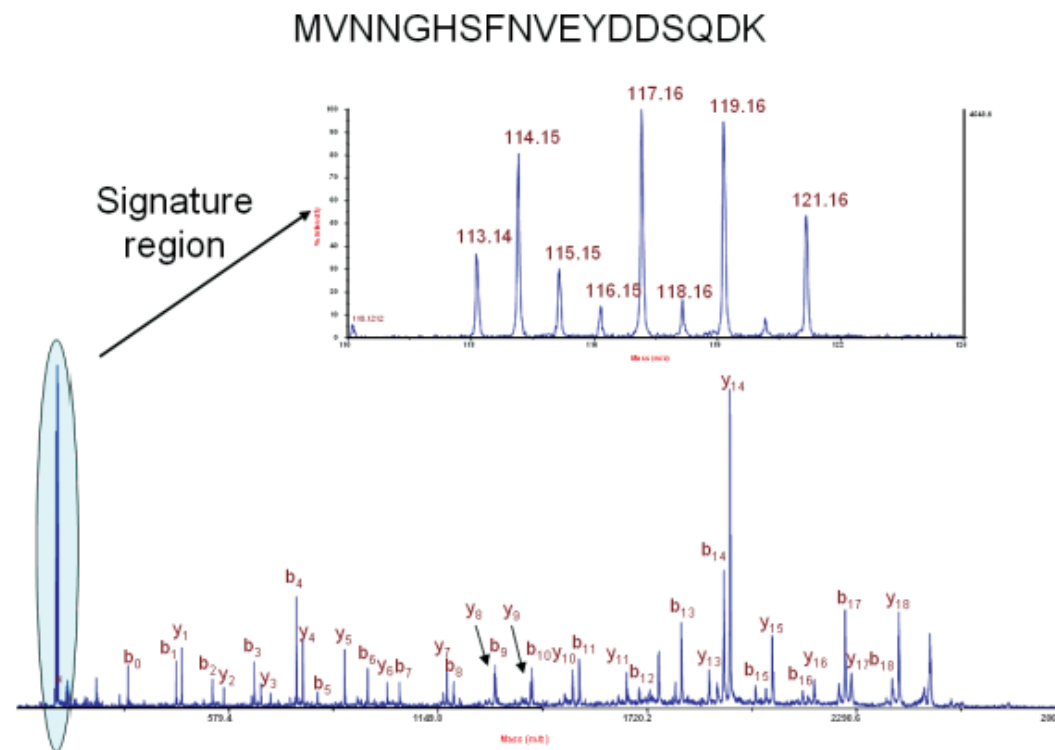


# 8-plex iTRAQ

When four is not enough...

8-plex version produces reporter ions at  $m/z$  113, 114, 115, 116, 117, 118, 119 and ... 121

Have to pick  $m/z$  that are not normally observed in peptide CID spectra.



Choe, L., et al. *Proteomics* (2007) **7** (issue#20) 3651-60

# $\beta$ -elimination followed by Michael addition of DTT (BEMAD)

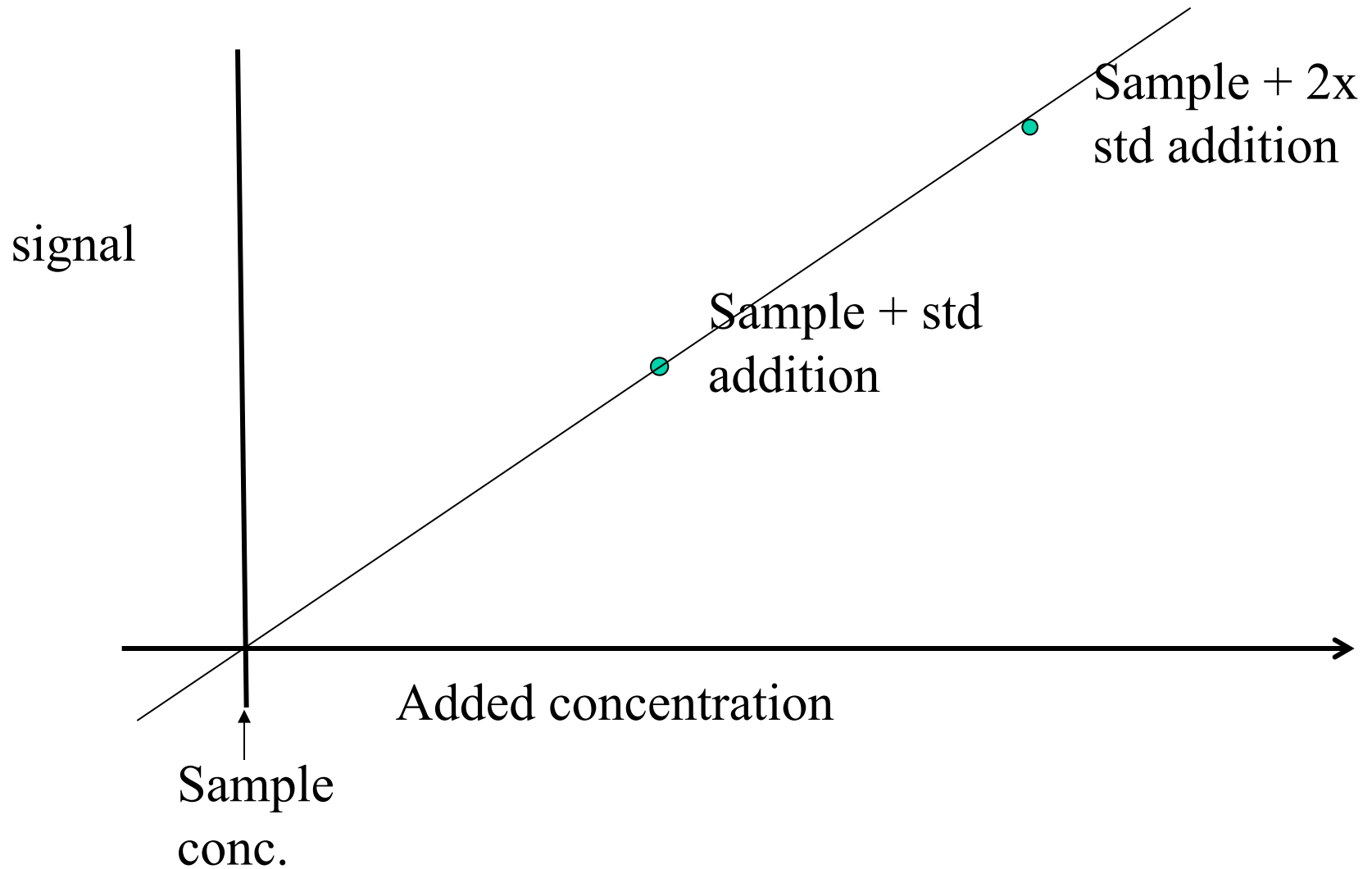
- Primarily aimed at Ser/Thr post-translational modifications
  - Phosphorylation
  - O-Glycosylation (e.g. O-GlcNAcylation)
  - O-Sulfation
- Alkylated cysteines also undergo reaction.
- Using  $^2\text{H}_6$ -DTT, can get quantitative information.
- Isotopic labeling can also be used to identify type of O-linked modification.

# Absolute Quantitation (AQUA)

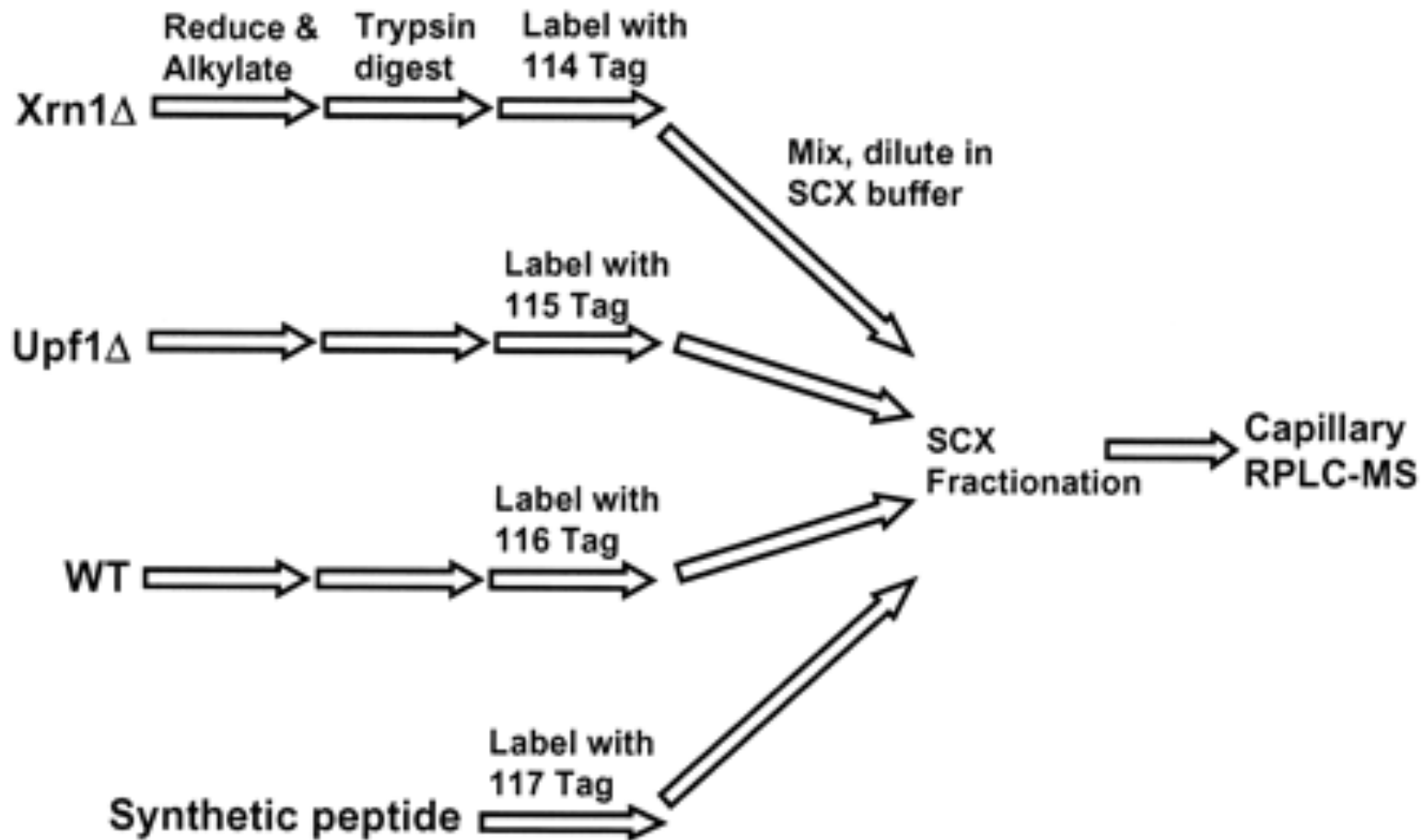
- All mass spectrometry quantitation so far has been relative.
  - What if it is relative to a known amount?
    - Absolute quantitation!
- For highest accuracy, a calibration curve of known amounts spiked into a sample should be obtained, then from this a peak intensity can be converted into an amount.
- Peak intensity for a given component is roughly linear with respect to intensity
- Add isotopic labeled version of peptide of known amount to sample, then compare intensity to unlabeled version in the sample<sup>1</sup>.

<sup>1</sup>Gerber et al. *PNAS* (2003) **100** (12) 6940-6945

# Method of standard addition

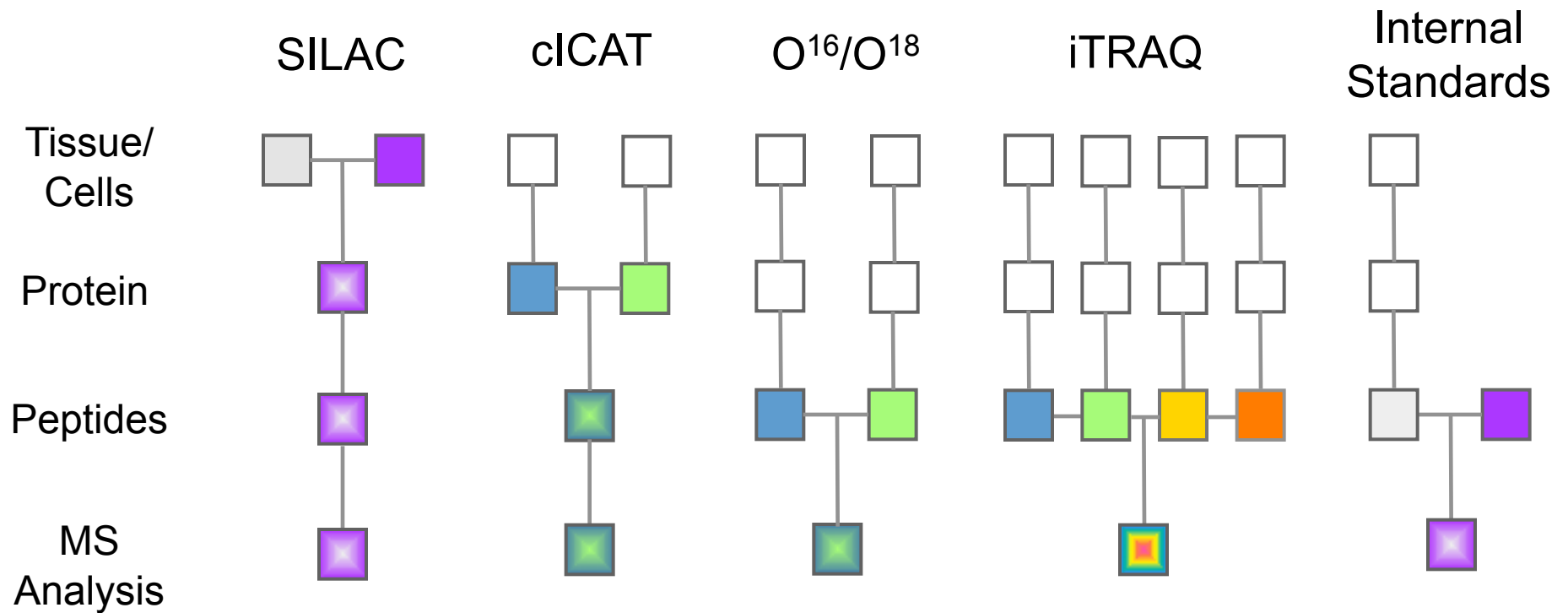


# iTRAQ for Absolute Quantitation



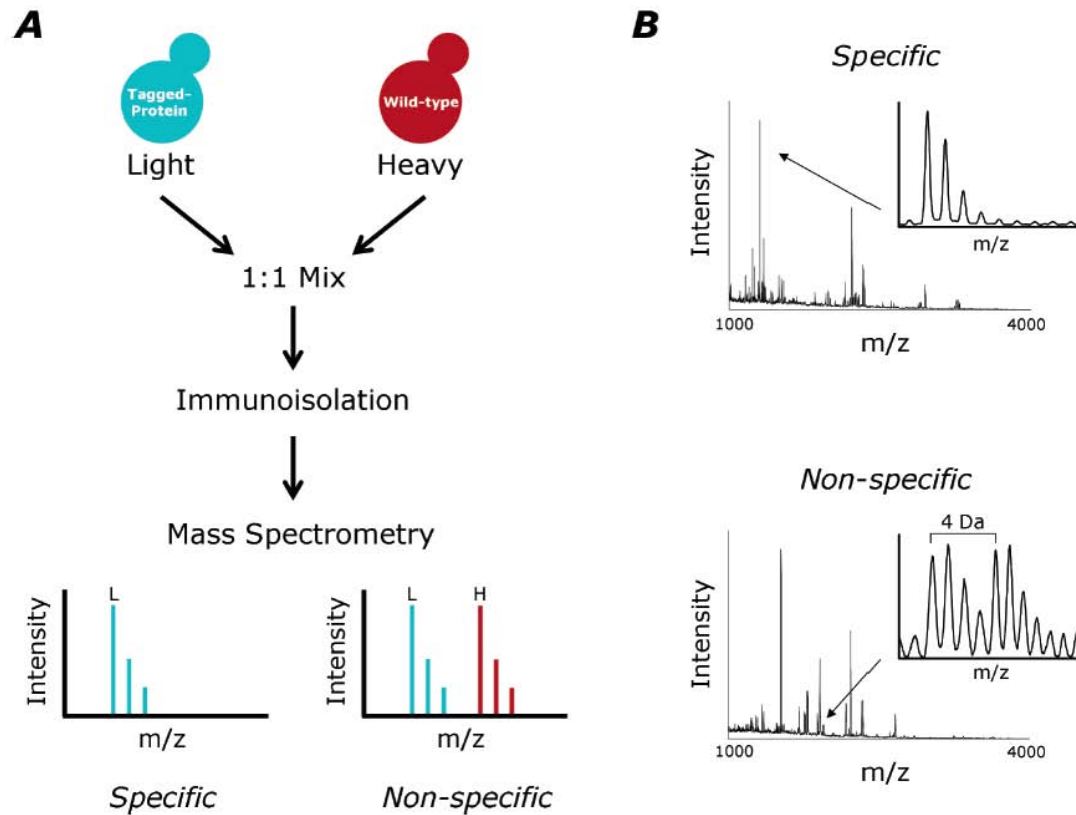
# Mass Spectrometry-based Protein Quantification Methods

At what point are the samples combined?



# Isotopic Differentiation of Interactions as Random or Targeted (I-DIRT)

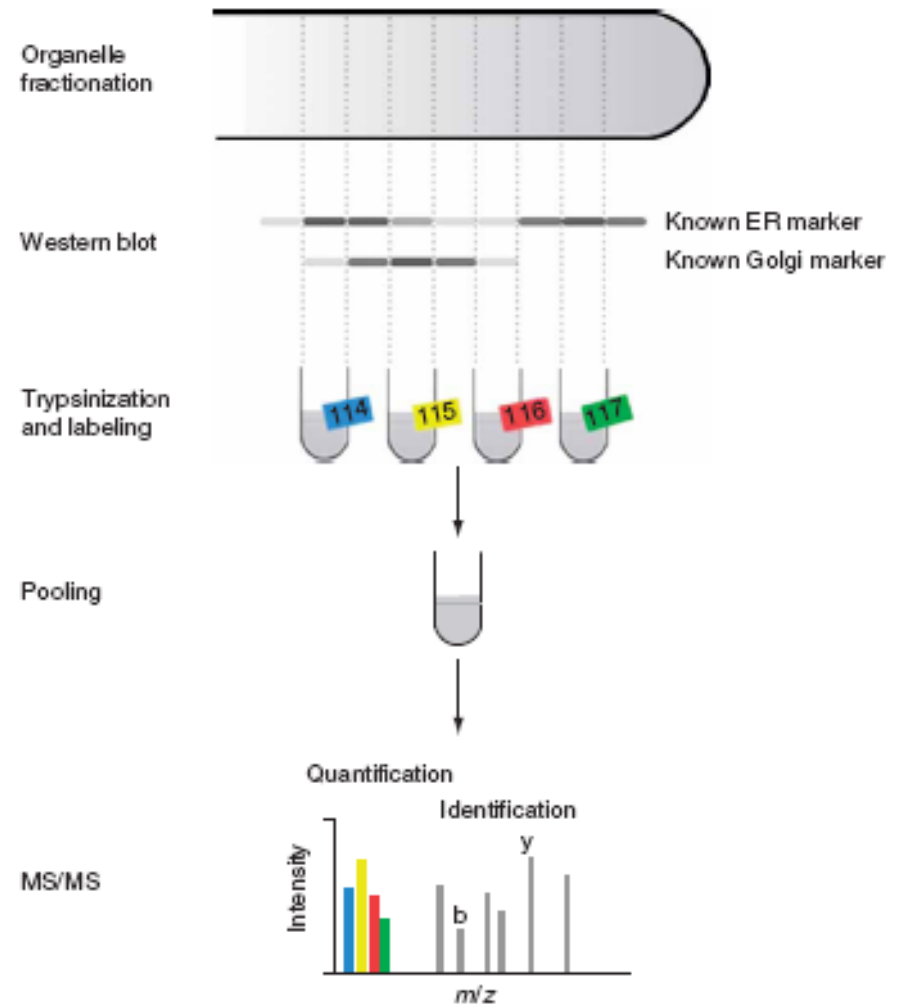
Identifying non-specific interactors isolated as part of an IP/tag pulldown.



Tackett et al, *J. Prot. Res.* (2005) **4**, 1752.

# Localization of Organelle Proteins by Isotope Tagging (LOPIT)

- Density centrifugation is used for organelle separation.
  - A given organelle will span several fractions.
  - Most fractions contain proteins from more than one organelle.
- Distribution of proteins between fractions should be the same for those in the same organelle.
- If separate fractions are labeled with iTRAQ, the reporter peak intensity pattern can be used to identify organellar location.



# Label-Free Quantitation

All approaches so far require purchase of isotopically labeled reagents (can be expensive).

- What if you want to compare large numbers of samples (10+)
- What if you can't afford lots of reagents?
  
- Peak/Spectral counting
  
- Peak area comparison (XIC)

# Spectral Counting

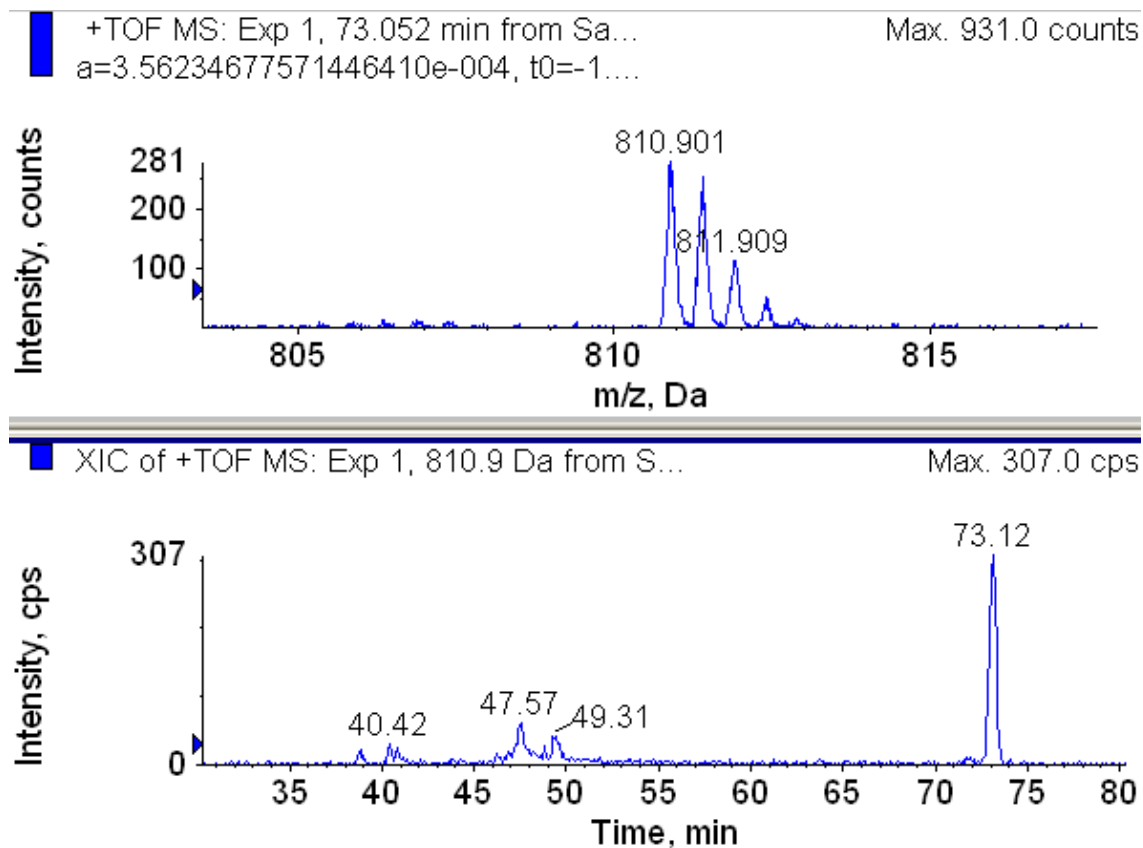
- Count the number of peptides identified from a protein in each sample.
  - Do you count repeat identifications of the same peptide?
- Not accurate at quantifying magnitude of change, but can be used to determine if there is a difference.
- In general, need a spectral count difference of about 4 peptides in order to be confident of a difference being real.<sup>1</sup>
  - Most proteins in complex mixtures are identified by less than 4 peptides.

<sup>1</sup>Old et al. Mol Cell Proteomics (2005) 4 10 1487-1502

# XIC

## (Extracted Ion Chromatogram)

- Measure intensity of peak during its elution off HPLC column and into the mass spectrometer.
- Measure area of peak in XIC.
  - More accurate than selecting peak intensity for one given scan.



# XIC with Internal Standards

- When comparing samples for label-free quantitation, how do you adjust for different overall amounts of protein loaded?
  - Add an internal standard to each sample and normalize against this.
  - Adding a protein digest will give multiple internal standards throughout the run.
    - Can normalize to standard eluting at similar time
      - Minimizes effects due to different levels of ion suppression in different samples.

# emPAI

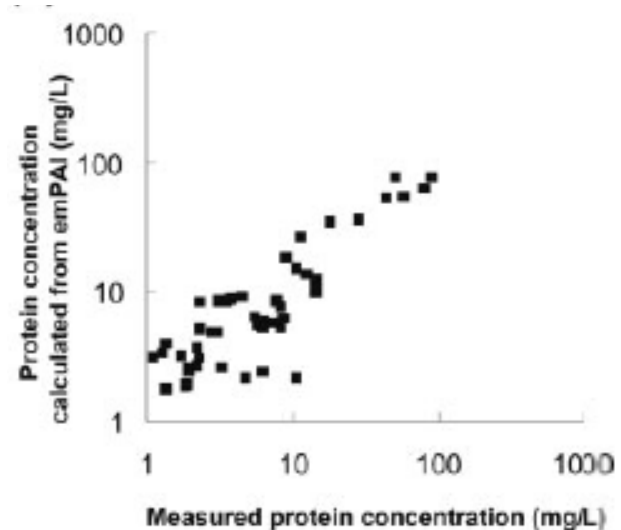
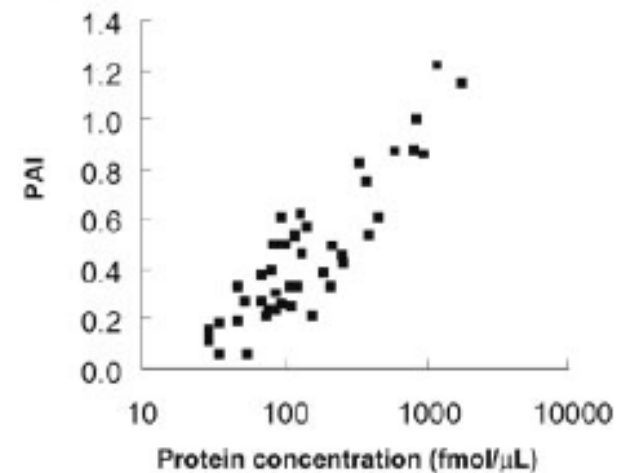
## (Exponentially Modified Protein Abundance Index)

$$\text{emPAI} = 10^{\text{PAI}} - 1$$

Where  $\text{PAI} = N_{\text{observed}} / N_{\text{observable}}$

What is an 'observable' peptide?

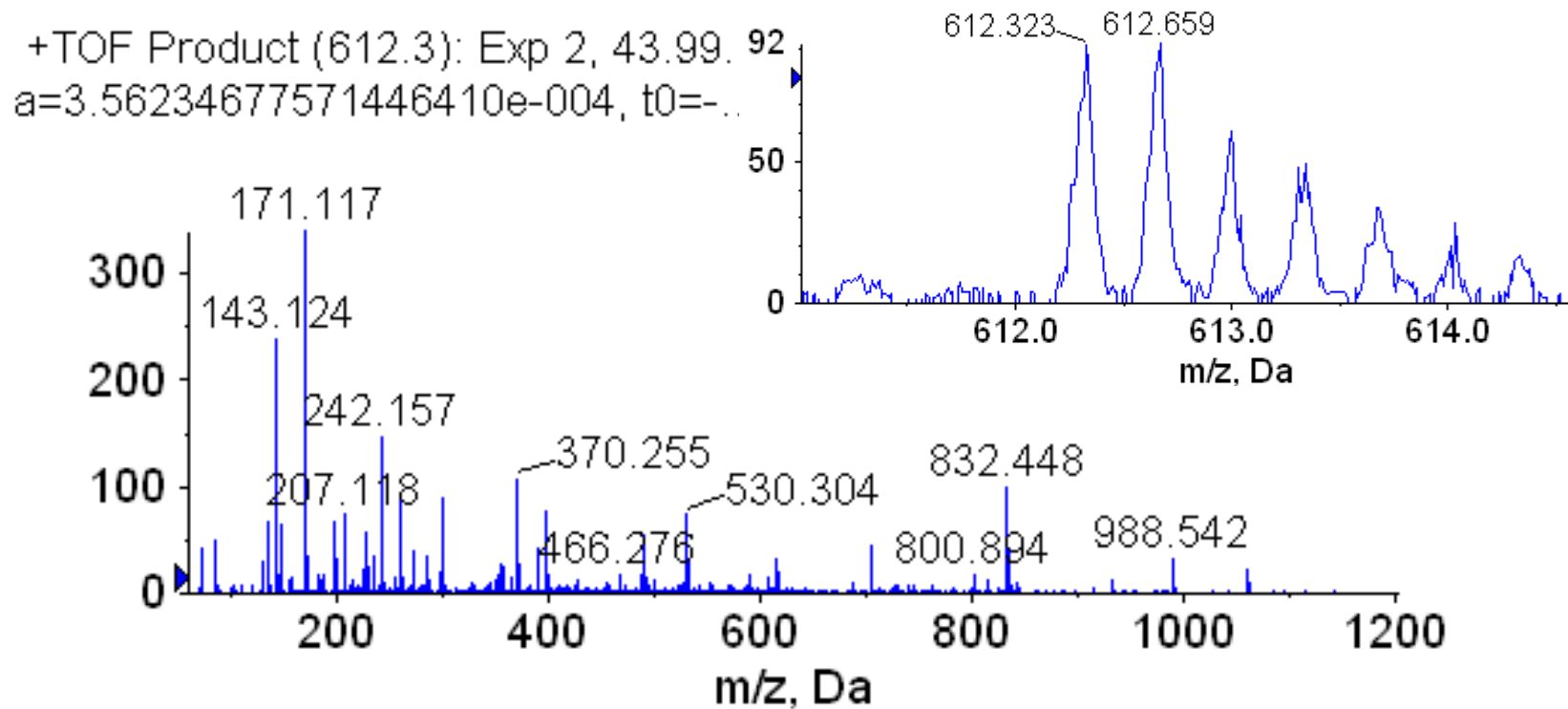
- Peptides with a precursor mass between 800-2400Da.
- There is a roughly linear relationship between log protein concentration and the ratio of 'observable' peptides observed in range of 3-500 fmoles.
- If you know how much total protein you analyzed you can derive absolute abundancies.
- You can observe more peptides than are observable!



# MRM

## (Multiple Reaction Monitoring)

Look for a component of a specific mass that when fragmented forms a fragment of another specific mass.



Transition: precursor m/z 612.32  $\longrightarrow$  fragment m/z 832.45

- Very sensitive and specific.

# MRM

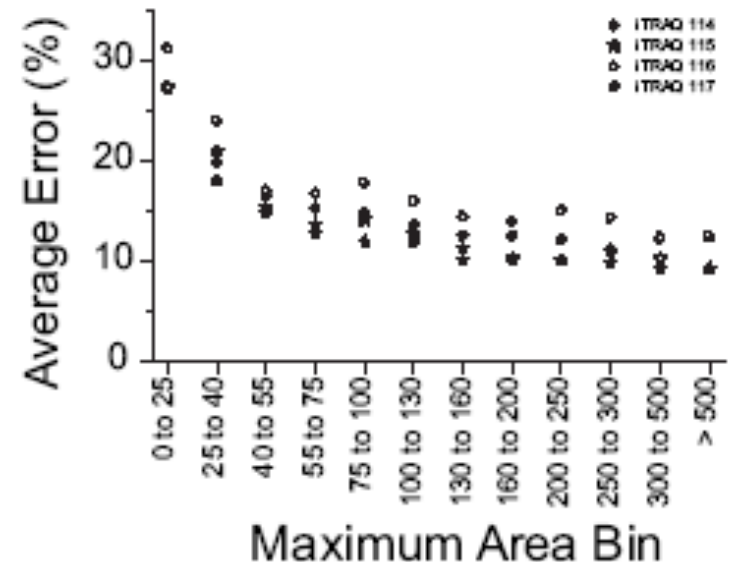
- Best performed on a triple quadrupole instrument.
- Scans are very fast, so can perform multiple transition scans on a chromatographic time-scale.
- If you know retention times, you can schedule different scans at different times during the LC-MS run.
  - It is possible to analyze a few hundred 'transitions' in a one hour run.
- If you perform MRM analysis of isotopically labeled (light and heavy) peptides can perform quantitation analysis at very high sensitivity.
- With a calibration curve of amount vs intensity, can approximate absolute concentration.
  - This is an approach that is likely to be powerful for biomarker analysis.

# Software for MS Quantitation Analysis

- Different software is required for different quantitation methods.
- For most quantitation methods, software needs to be able to read instrument manufacturers raw format.
  - Each manufacturer uses own proprietary format
- Most quantitation software only works for one type of quantitation on one type of instrument, using one search engine.
- ProQuant (Applied Biosystems): ICAT, cICAT and iTRAQ data acquired on their instruments, searched with their own search engine (Protein Pilot)
- MSQuant: SILAC data acquired from one of three instrument manufacturers, searched using Mascot
- Protein Prospector: Any type of isotopic labeling acquired on instruments from two manufacturers.

# Issues with MS Quantitation Analysis

- Should you use all data for quantitation?
- Minimum peak intensity?
  - Peaks near to signal to noise will have much higher variability in quantitation accuracy.
- Are proteins identified by a single peptide accurately quantified?
- What if a peptide contains a modification; e.g. oxidized methionine?
- What if a peptide is observed with and without a missed cleavage?



Results are normally reported with a mean and standard deviation.

# Software for Analysis of Label-free Data

- Quantitation of label-free data is much more difficult from a bioinformatic perspective.
  - When comparing LC-MS runs, elution times are not identical from run-to-run.
  - Mass measurements may vary slightly.
  - Software needs to be able to distort (stretch / squeeze) chromatogram to be able to compare runs.

# Comparison of Quantitation Accuracy

What level of quantitation accuracy can you expect with these methods?

Dynamic Range:

Most MS methods have a dynamic range accuracy of only about 1 order of magnitude:

e.g., cannot distinguish reliably between 20 fold and 100 fold increase.

Problem is usually the lower value in the comparison is too close to the signal to noise level.

What is the minimum change that can be reliably measured?

For label-free methods: 2 fold change

For isotopic labeling methods: 30% change

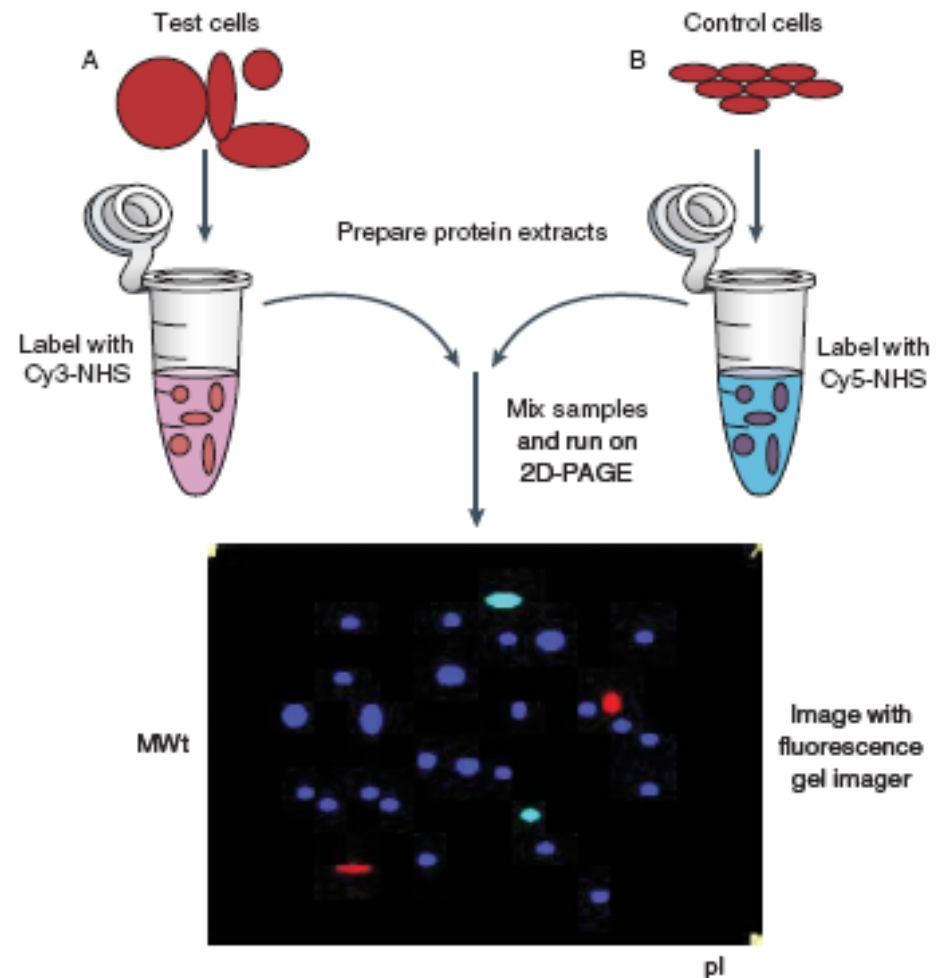
# Quantitate then MS: Gels

- Coomassie and Sypro stain densities are relatively linear with respect to amount.
  - Silver stain is less quantitative.
- Complication: How do you know there is only one major protein component in the band.
- Greater separation with 2D gels reduces chances of multiple components per spot
  - Multiple components may still be present if starting mixture was complex.
  - 2D gels allow quantitation of a specific-modified version of protein (providing you know what modification state the spot corresponds to!)
  - Will separate different splice variants.

Quantitative Westerns: specific, but not very accurate quantitation.

# Difference Gel Electrophoresis (DIGE)

- Fluorescent dyes which react with lysine residues.
- Fluoresce at different wavelengths, so can be differentially visualized.
- Labeling based on 'minimal labeling': only 5% of protein is singly labeled.

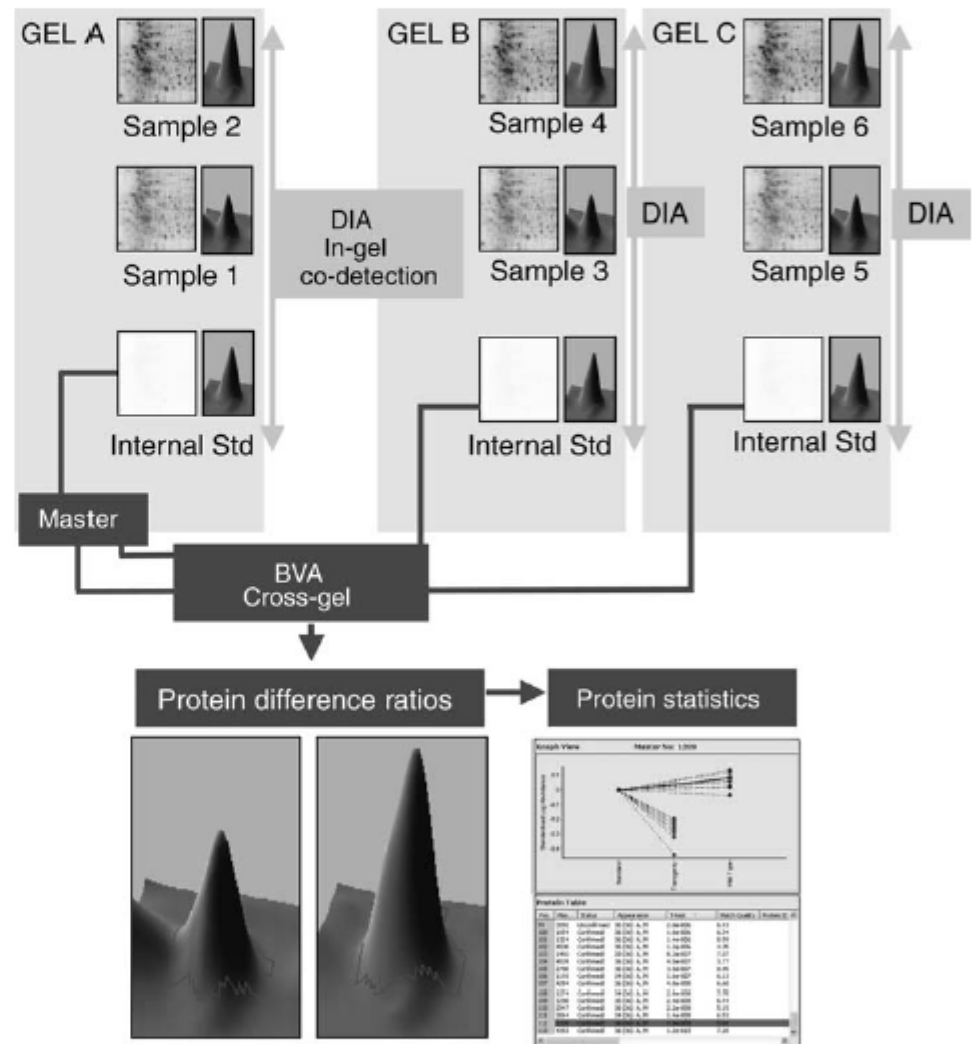


# Difference Gel Electrophoresis (DIGE)

- Label alters migration, so majority of protein (unlabeled) does not co-elute
  - Cutting out spot for ID is more complicated.

Technology sold by GE Healthcare.

- Can only detect spots using their scanner. 😞
- Very powerful analysis software available 'Decyder'. 😊
- Scanner and software cost a small fortune. 😞



# Conclusions

- Although not intrinsically quantitative, mass spectrometry can be used for many types of quantitation.
- Isotopic labeling is most accurate.
- Label-free approaches are simpler (and cheaper) and more appropriate for comparison of large numbers of samples, but are less accurate.
- Most quantitation is relative, but if one component is at a known concentration then absolute quantitation can be performed.
- Software for analyzing this type of data is still being developed.