# Quantification by Mass Spectrometry

PC219

Lecture 5

30 April, 2010

sguan@cgl.ucsf.edu

# Applications of Quantitative Proteomics

Qualitative protein identification is NOT sufficient to describe a biological system

***Detecting changes in proteomes***

**in space: cellular and tissue localization**
**in time: signaling cascades or turnover**
**sample perturbations:**

- ***Comparison of normal and diseased tissue samples***
  - Biomarker discovery
    - Drug development
    - Diagnostics
- ***Gene knockdowns or over expression***
- ***Inhibitors: eg antibodies or siRNA***
- ***Growth factors/hormones***
- ***Cell-cell interactions***
- ***Drug treatment***

# Quantitative Proteomics Methodologies

***2D gel electrophoresis***

silver stain
Fluorescence Difference Gel Electrophoresis
Pro-Q Diamond
2,4-dinitrophenylhydrazine

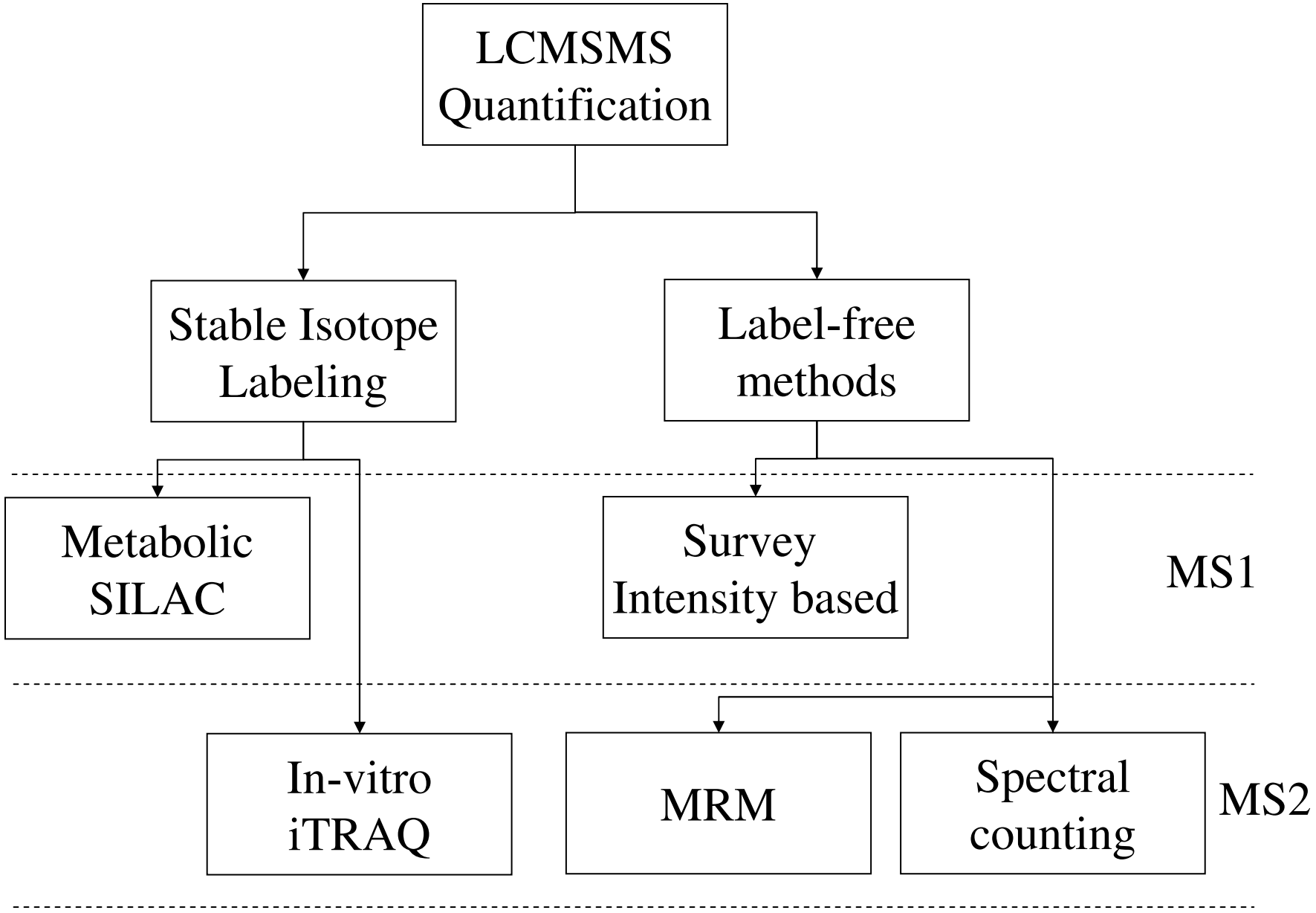
***Protein expression array analysis***

Global epitope tagging
(Nature2003v425p737)

***MS-based quantification methods***

- Quantification
- H/D exchange – protein surface mapping
- Proteomics dynamics

# Basic Classifications of MS Based Quantification

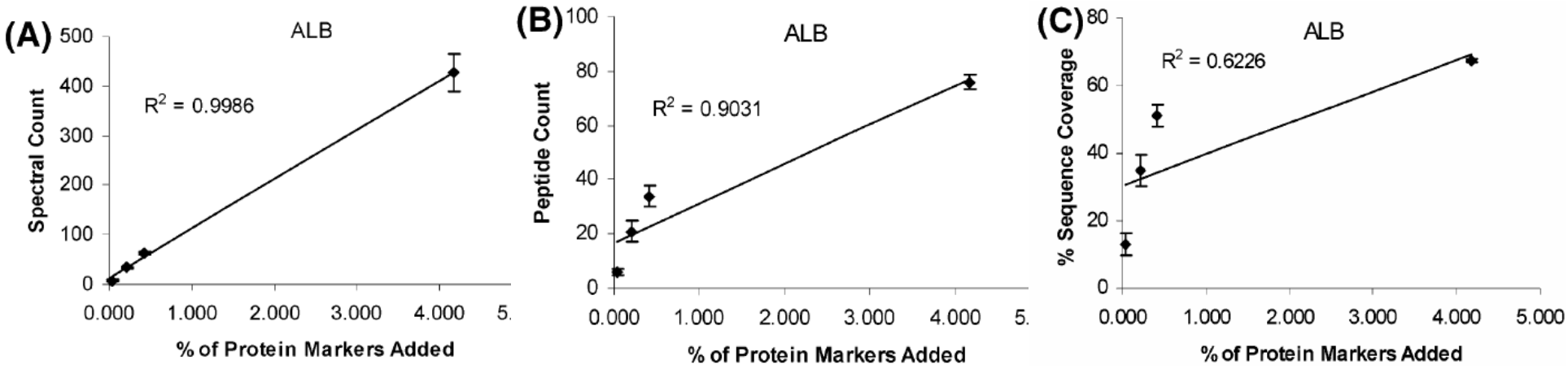


# Label Free Quantification

# Label Free Quantification

*Spectral Counting*

- ***Spectral counts (# of MSMS from a Protein) correlate with protein abundance well***

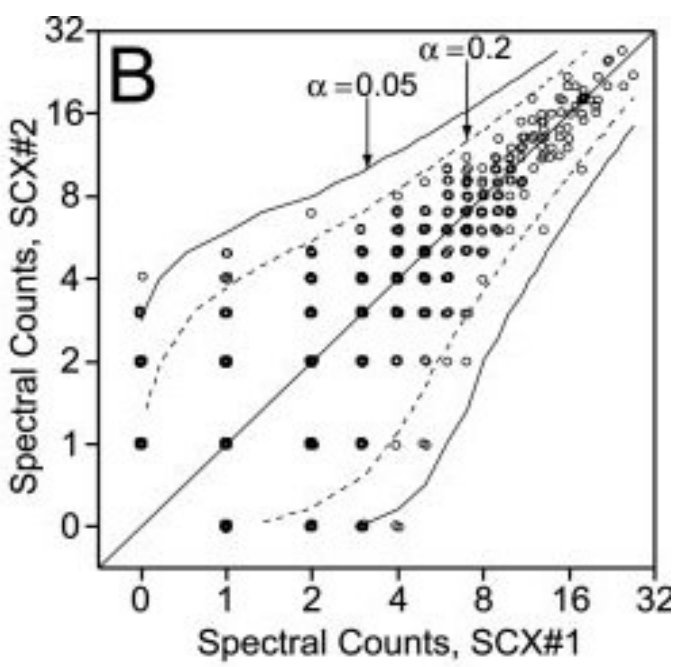


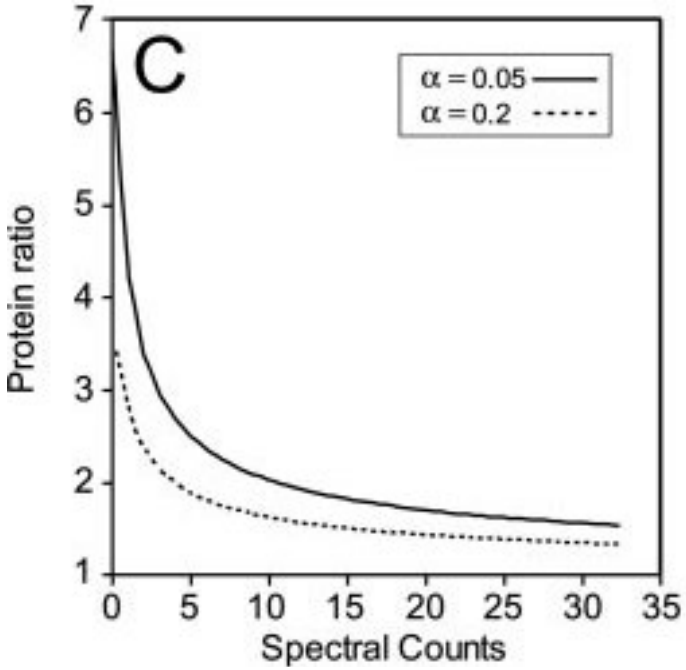

- ***Spectral counting dynamic range (single runs) is about two orders of magnitude***
- ***10 <u>repeated</u> analysis can reach 95% saturation level***

AnalChem2004v76p4193

# Label Free Quantification

*Spectral Counting*




Minimum protein ratios that can be determined at 80 and 95% confidence limits

MCP2005v4p1487

# Label Free Quantification

## *Spectral Counting - Formula*

**$R_{SC}$**: log2(protein ratio) measured from spectral counts
(MCP2005v4p1487)

$$R_{SC} = \log_2[(n_2 + f)/(n_1 + f)] + \log_2[(t_1 - n_1 + f)/(t_2 - n_2 + f)]$$

$n_1$, $n_2$ - spectral counts for sample 1 and 2
$t_1$, $t_2$ – total spectral count (sampling depth) for samples 1 and 2
f – correction factor 0.5 (Bioinformatics2004v20pi31)

**NSAF**: normalized spectral abundance factor
(PNAS2006v103p18928)

$$(\mathrm{NSpC})k = \frac{(\mathrm{SpC}/L)_k}{\sum_{i=1}^{\mathrm{N}} (\mathrm{SpC}/L)_i}$$

SpC: total number of tandem MS spectra matching peptides from protein
L: protein length

# Label Free Quantification

*Extracted Ion Chromatogram Based Methods*

**In-Vitro Phosphorylation of BZR1 by BIN2 Kinase**

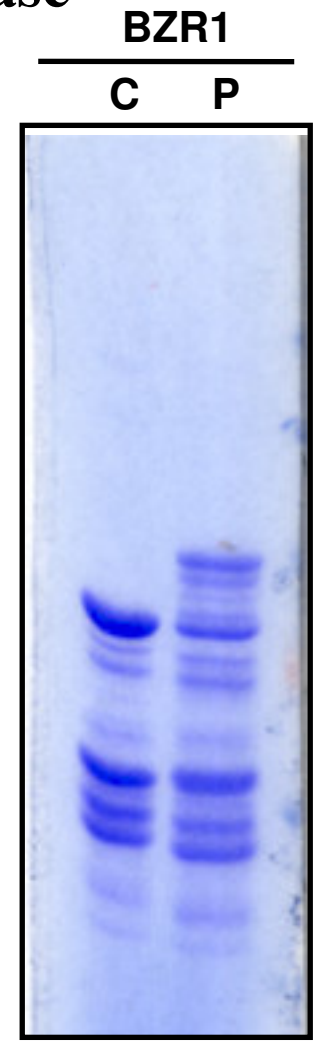
E. coli expressed MBP-BZR1 as substrate

E. coli expressed GST-Bin2 as kinase

Phospho sample: BZR1+Bin2+ATP

Control sample: BZR1+Bin2

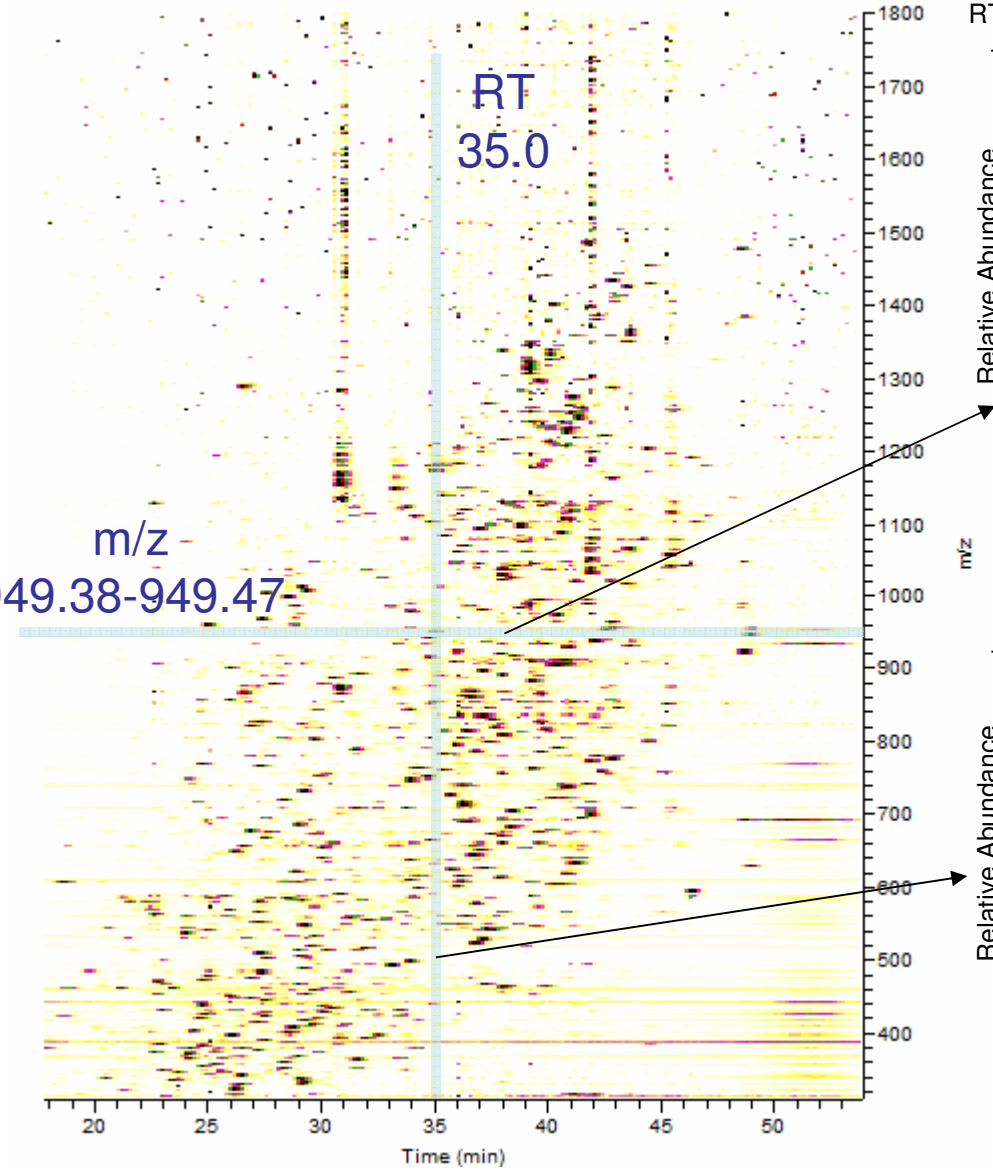
*Goal: Estimate Stoichiometry of PTMs*

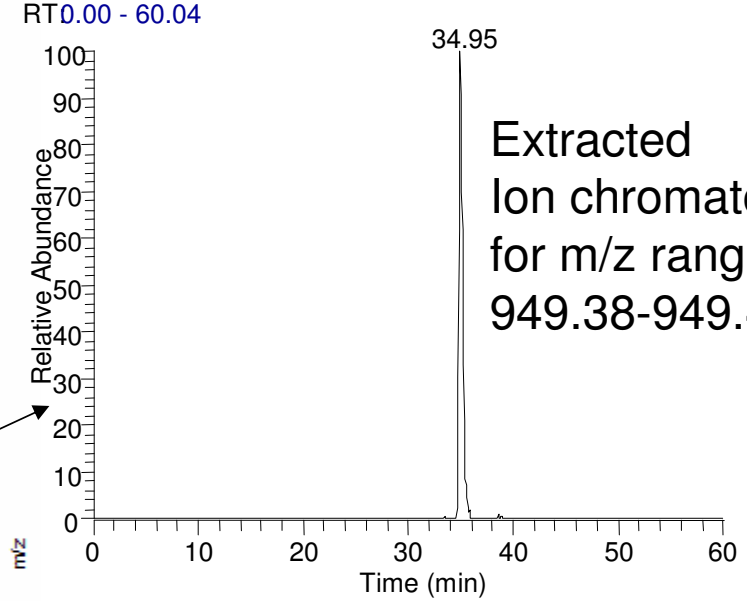


# Label Free Quantification

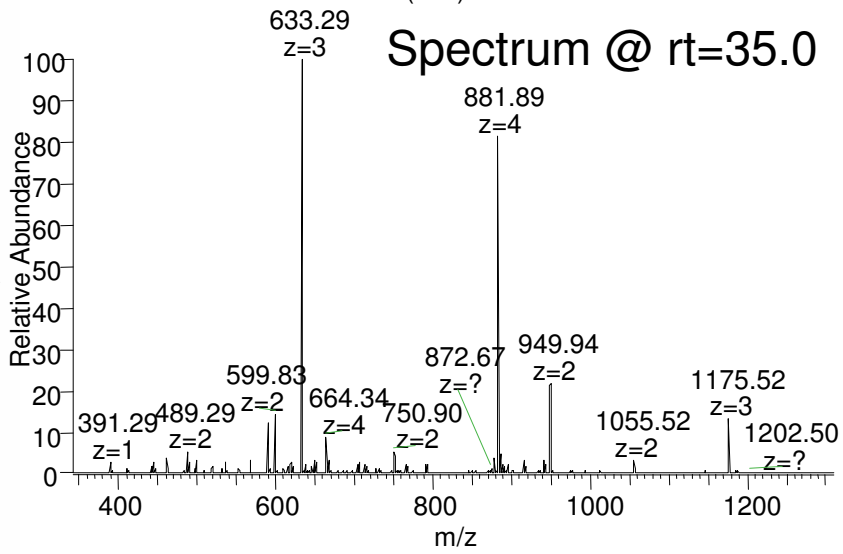
## *Extracted Ion Chromatogram Based Methods*

Retention time - m/z ion map: Phospho sample (BZR1+Bin2+ATP)




Extracted
Ion chromatogram
for m/z range of
949.38-949.47

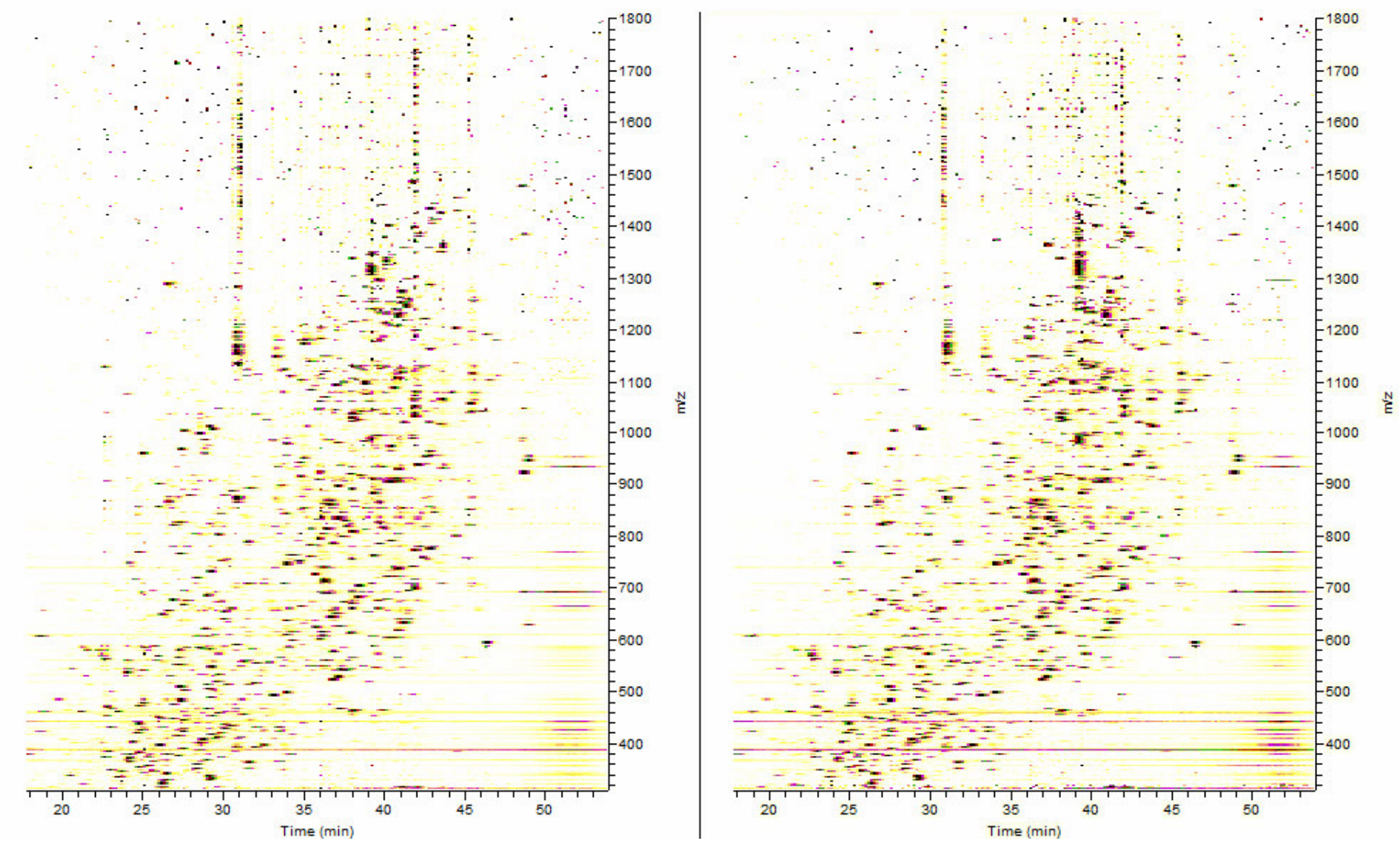


# Label Free Quantification

*Extracted Ion Chromatogram Based Methods*

Phospho Sample

Control Sample



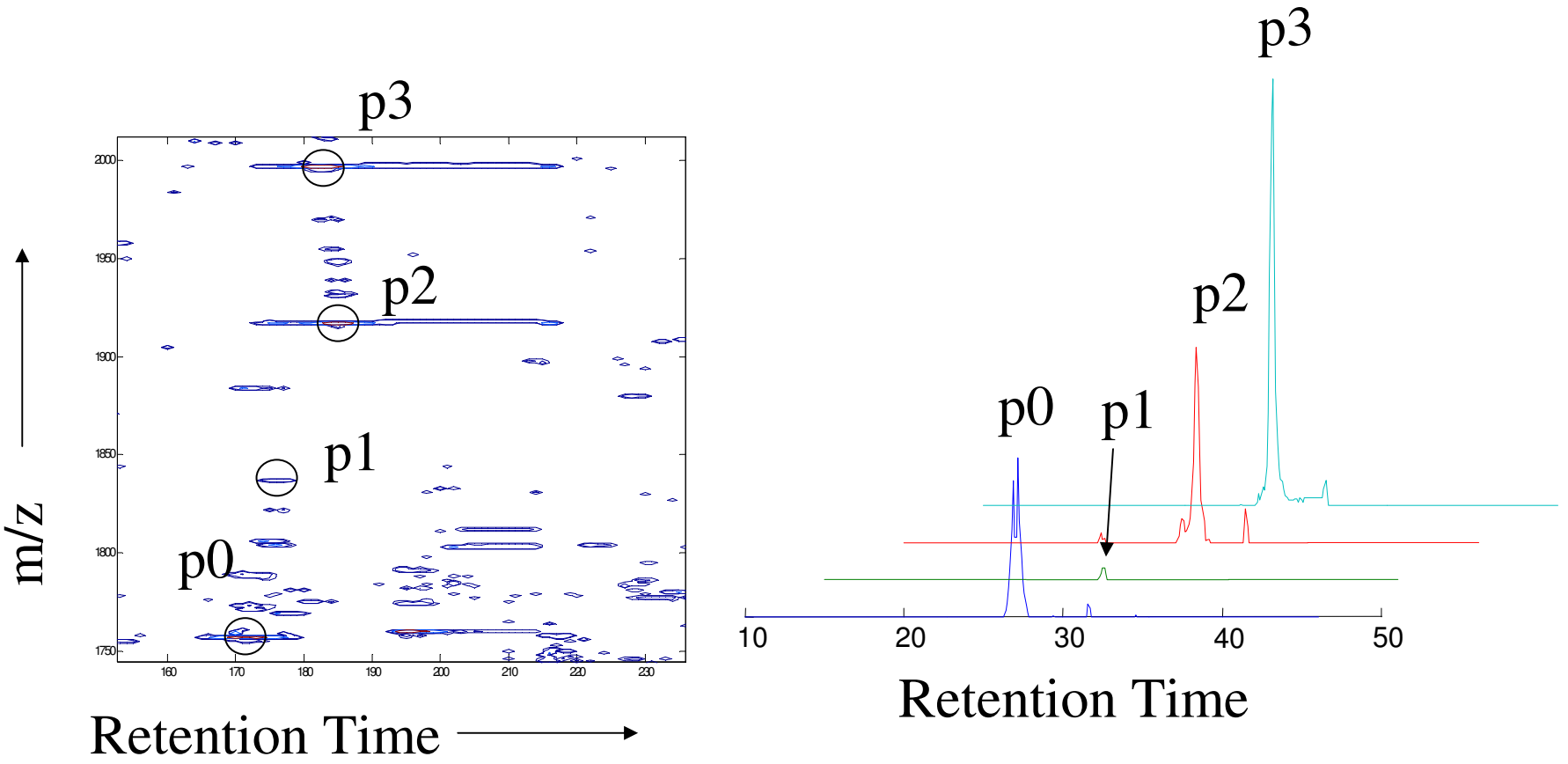
# Label Free Quantification

*Extracted Ion Chromatogram Based Methods*

## Phosphopeptides

170 186

ISN**S**CPV**T**PPVS**S**PTSK



# Label Free Quantification

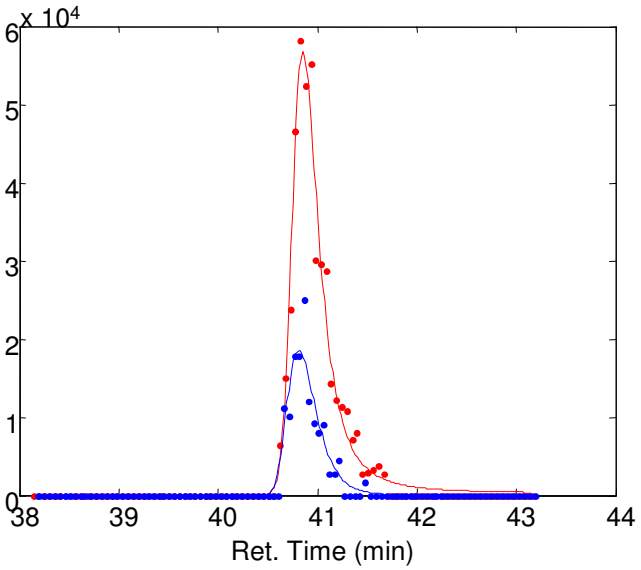
*Extracted Ion Chromatogram Based Methods*

***Extracted Ion Chromatogram and Integration***

*Non-constant sampling rate
*Noisy signal
*Initial peak parameters

$m/z = 1225.63, z = 4$

| | Control | Phospho |
|---|---|---|
| Ret. Time (min) | 40.85 | 40.81 |
| Peak Height | $5.67X10^4$ | $1.86X10^4$ |
| Peak Area (min) | $2.26X10^4$ | $7.14X10^3$ |
| Peak Width (min) | 0.319 | 0.346 |
| $R^2$ | 0.982 | 0.923 |




259FAQQQPFSASMVPTSPTFNLVKPAPQQMSPNTAAFQEIGQSSEFK303++++

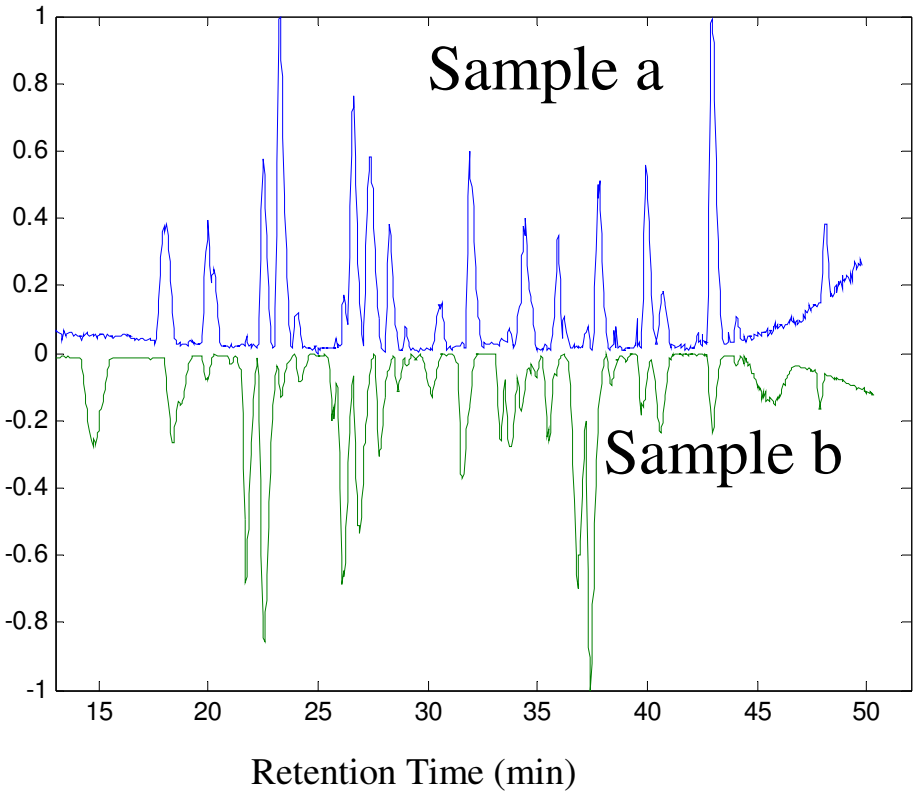
# Label Free Quantification

*Extracted Ion Chromatogram Based Methods*

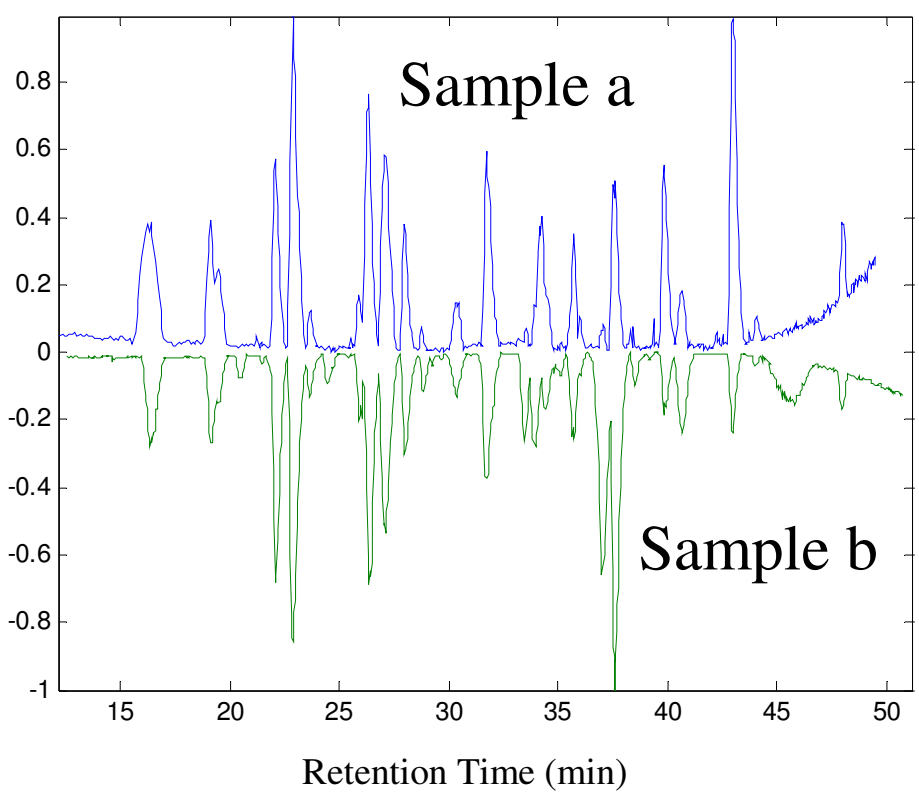
***LC Alignment*** of base-peak chromatograms:

- Dynamic Time Wrapping (similarity in chromatograms)
- Use “landmark” ions (needs MSMS info)

Before



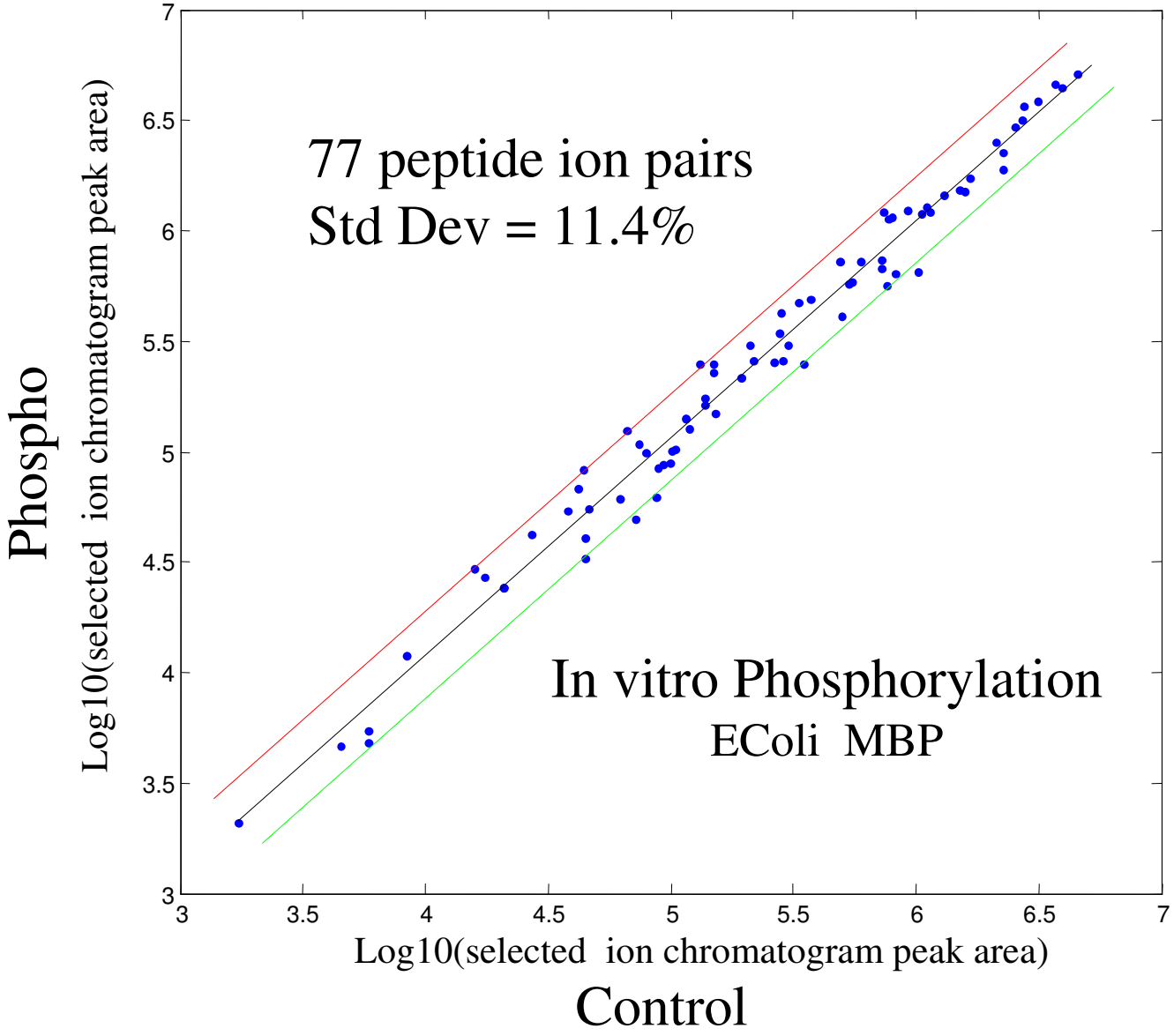

After



# Label Free Quantification

## *Extracted Ion Chromatogram Based Methods*

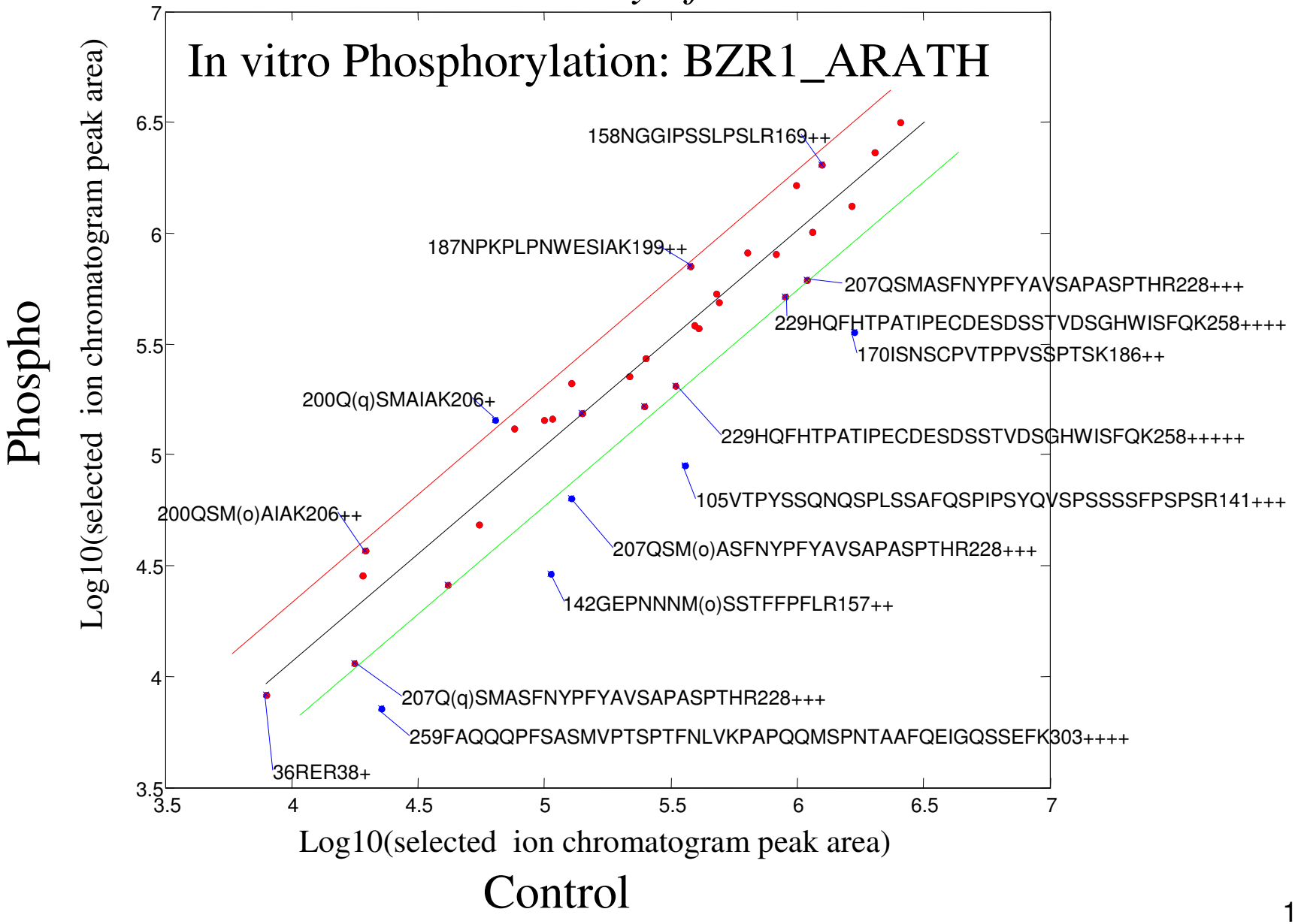
### *Stoichiometry of PTMs*



# Label Free Quantification

## *Extracted Ion Chromatogram Based Methods*

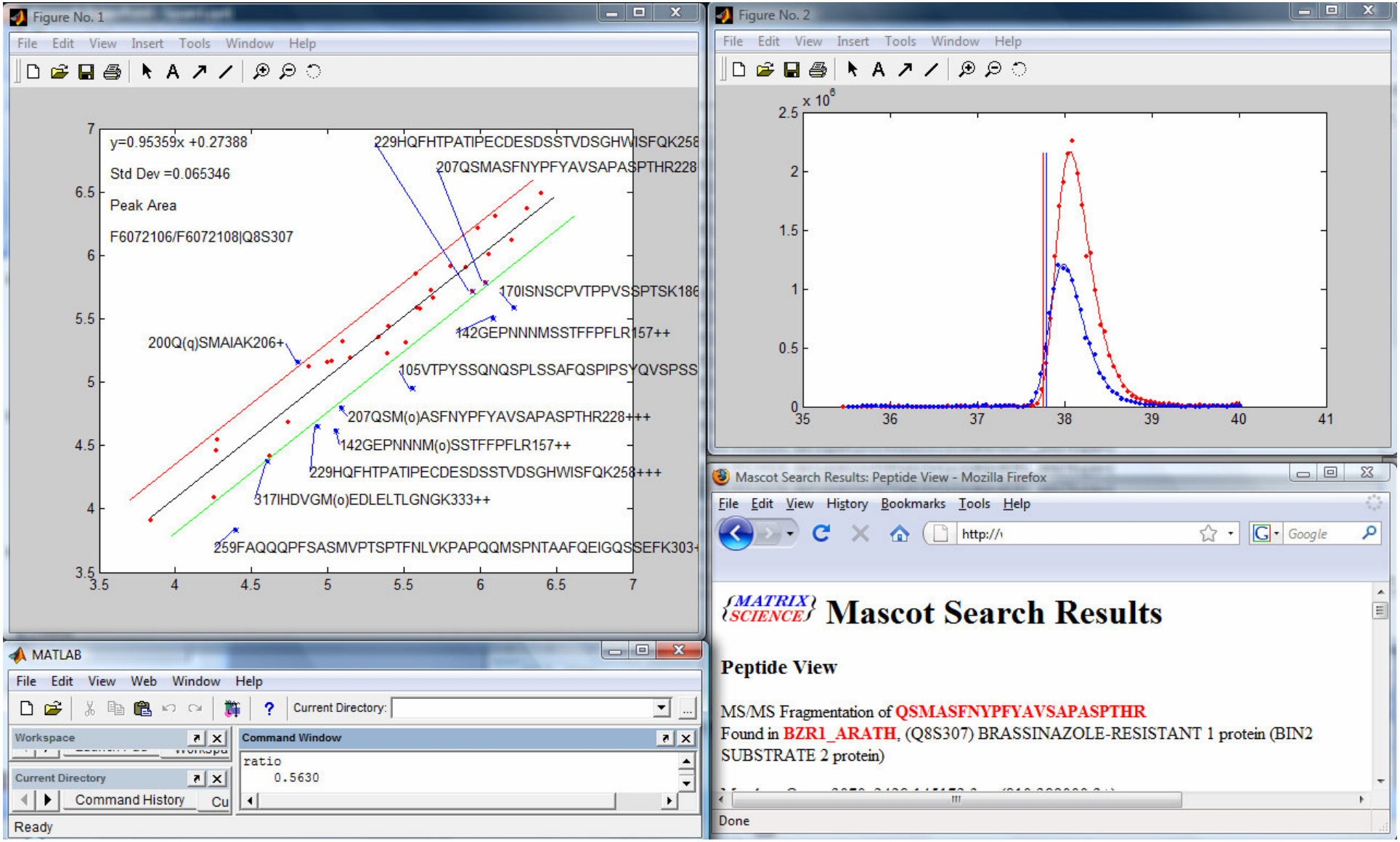
## *Stoichiometry of PTMs*



*Stoichiometry of PTMs*

# User Interface

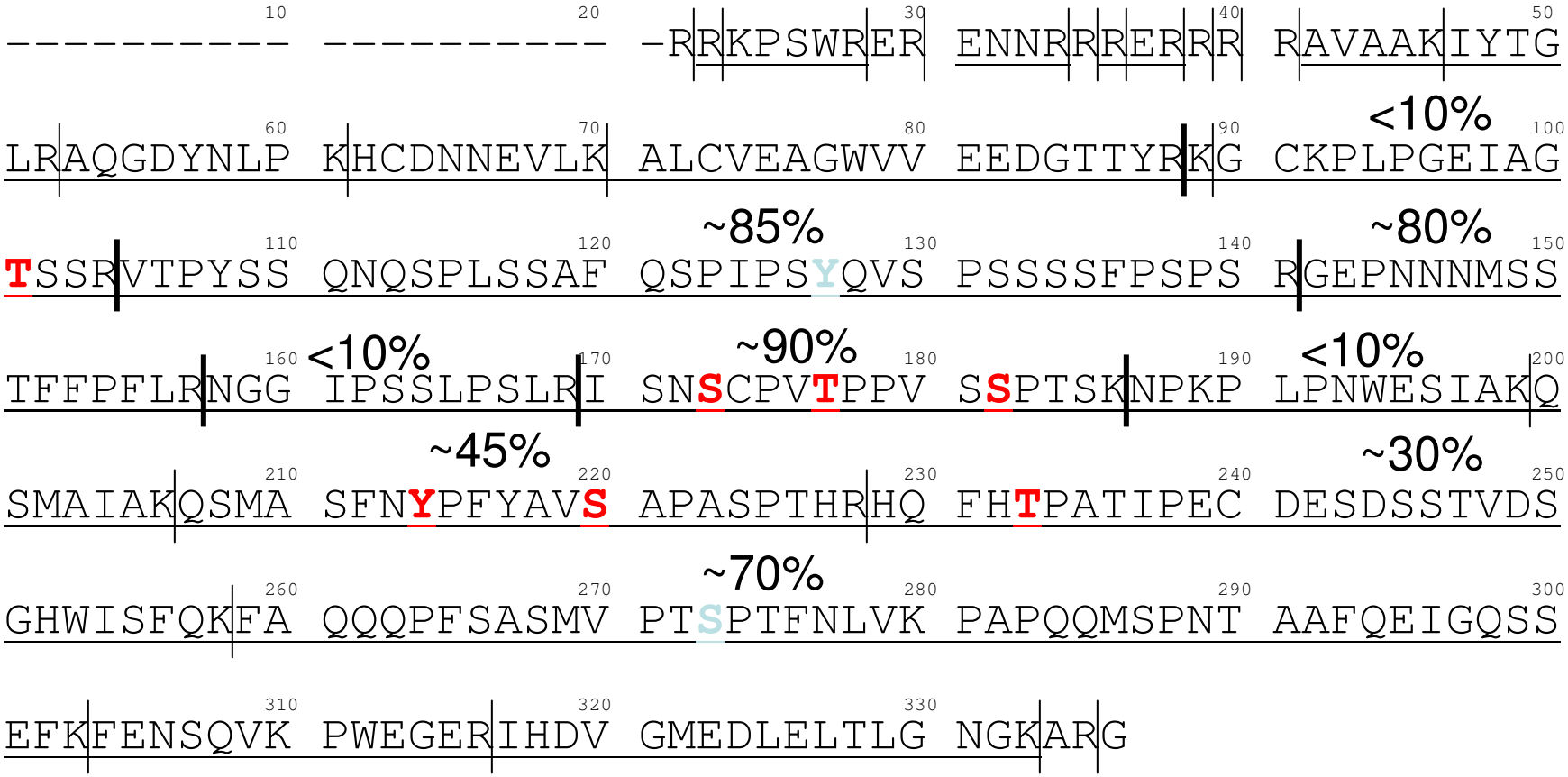
*- Scatter Plot, Extracted IC, Peptide ID Page*



*Stoichiometry of PTMs*

# BRASSINAZOLE-RESISTANT 1 protein (BZR1)

## Extent of Phosphorylation



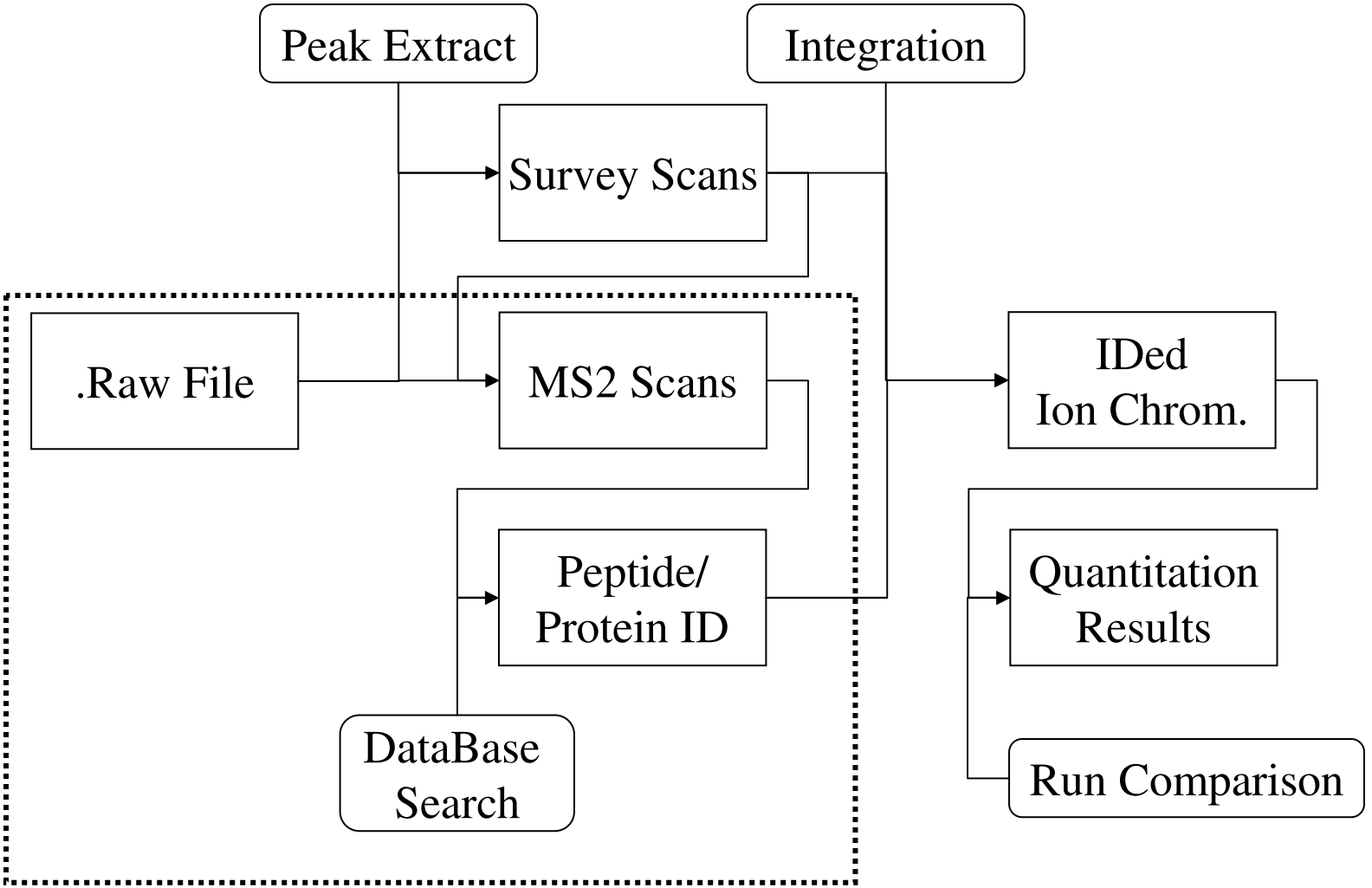

S, T, and Y are conformed assignments in this experiment

S and Y are tentative assignment

# Label Free Quantification

## *Extracted Ion Chromatogram Based Methods*

*Data Processing Workflow*



# Label Free Quantification

## *Extracted Ion Chromatogram Based Methods*

**Table 1.** Overview of LC-MS alignment software for proteomics solutions

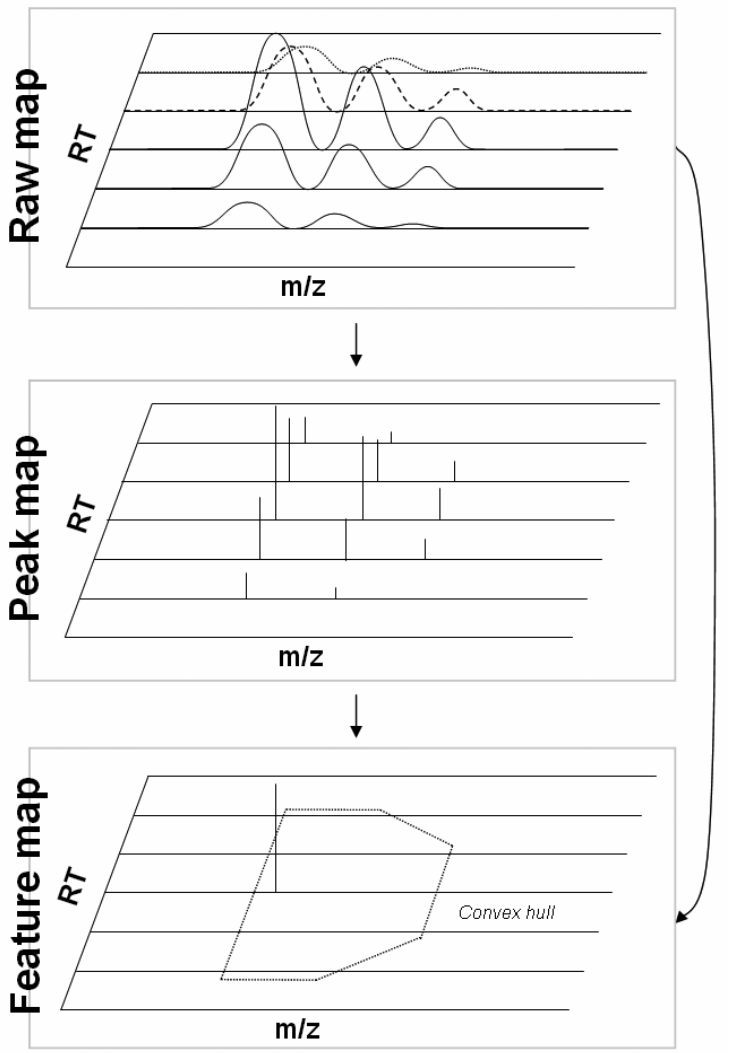
| Software name | Supplier / author | Database/ environment | availability | Functionality | website | reference |
|---|---|---|---|---|---|---|
| **PLGS IdentityE Expression Informatics** | Waters Corp | PLGS | commercial | f, h,i,b,a,r,s,l | http://www.waters.com/ | [15, 16, 58, 59] |
| **SIEVE** | Thermo Scientific | BioWorks | commercial | p,h,v,b,a,r,s,l | http://www.thermo.com/ | |
| **DeCyderMS** | GE Healthcare | | commercial | f,h,i,b,v,a,r,s,l | http://www.gelifesciences.com/ | [50] |
| **Rosetta Elucidator** | Rosetta Biosoft-ware | | commercial | f,h,i,b,v,a,r,s,l | http://www.rosettabio.com/products/elucidator/default.htm | |
| **MS-Xelerator** | MsMetrix | | commercial | f,l,i,b,a,r,s, | http://www.msmetrix.com | |
| **MassView** | SurroMed | | custom | f,l,i,b,a,r,s,l | | [61–63] |
| **MetAlign** | WUR | | free for acad. | p,l,b,a,s | www.metalign.wur.nl | [33] |
| **MzMine** | VTT Finland | | open source | f,h,v,a,r | http://mzmine.sourceforge.net/index.shtml | [42, 43] |
| **MSight** | SIB | | open source | f,h,i,v,(a) | http://www.expasy.org/MSight/ | [64] |
| **MS Inspect** | CPL (Fhcrc) | CPAS | open source | f,h,v,a,r (l,d) | http://proteomics.fhcrc.org/CPL/msinspect.html | [41, 65] |
| **SpecArray** | ISB /SPC | TPP | open source | f,h,i,v,a,r,s | http://tools.proteomecenter.org/SpecArray.php | [49] |
| **PePPer** | BROAD MIT | Genepattern | open source | h,a,r,s,l | http://www.broad.mit.edu/cancer/software/genepattern/desc/proteomics.html | [60, 66] |
| **VIPER** | PNNL | PRISM | open source | f,h,i,b,v,a,r,s,l,d | http://ncrr.pnl.gov/software | [57, 67] |
| **OpenMS** | Berlin Saarland Tubingen Univ. | TOPP | open source | (f,h,i,b,v,a,r,s,l,d) | www.openMS.de | [68–70] |
| **SuperHirn** | IMSB @ETH | Corra | open source | f,i,b,v,a,r,s | http://tools.proteomecenter.org/SuperHirn.php | [53] |
| **CPM (continous profile models)** | Listgarten/Emili | MatLab | free for acad. | l,a | http://www.cs.toronto.edu/~jenn/CPM/ | [52] |
| **Xalign** | Purdue Univ | Xmass | upon request | (f,h),i,a,s | zhang100@purdue.edu | [71] |
| **Fischer et.al.** | ETH | | not described | h,a | http://people.inf.ethz.ch/befische/ | [46] |
| **CRAWDAD** | Washington Univ | | upon request | f,l,i,a,r,s,l,d | http://proteome.gs.washington.edu/software/crawdad | |
| **CHAMS** | Inst Pasteur, Paris | | web server | h,a,s | http://www.pasteur.fr/recherche/unites/BioIsys/chams/index.htm | [51, 72] |
| **OBI-WARP** | Univ. Texas | | open source | a,r,l | http://bioinformatics.icmb.utexas.edu/obi-warp/ | [73] |
| **LCMSWARP** | PNNL | PRISM | open source | h,a | http://ncrr.pnl.gov/software | [74] |
| **LCMS2D** | Albert Einstein College of Medicine | | | | http://www.bioc.aecom.yu.edu/labs/angellab/ | [75, 76] |
| **PETAL** | CPL (Fhcrc) | CPAS | open source | a | http://peiwang.fhcrc.org/research-project.html | [77, 78] |

p/f: peak/feature detection; h/l: high/low resolution; i: de-isotoping; b: batch processing; v: LCMS 2-D visualization; a: alignment; r: result visualization; s: statistical analysis; l: link MS to MS/MS; d: results database

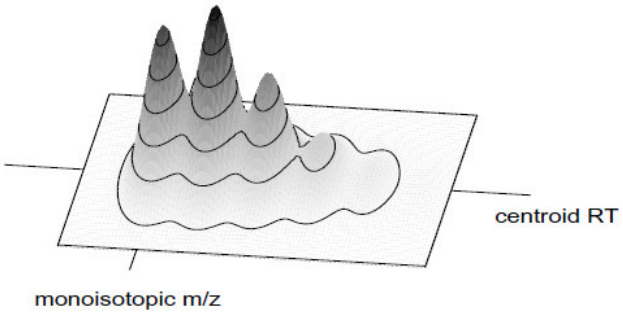
JProteomeRes2008v7p51
Proteomics2008v8p731

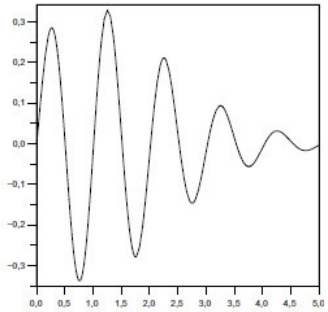
# Label Free Quantification

*Extracted Ion Chromatogram Based Method - OpenMS*

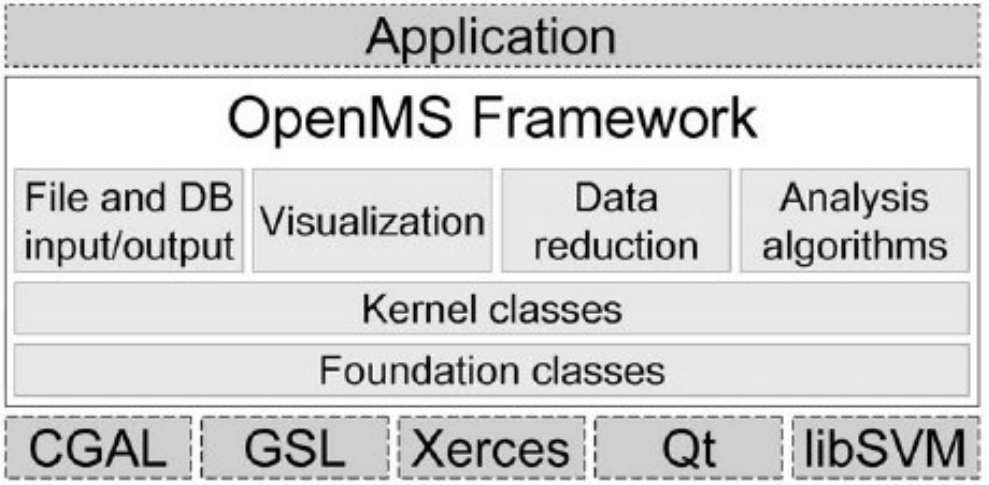



Isotope distribution modeling: wavelet



Software architecture: C++/Linux




BMCBioinformatics2008v9p163

# Label Free Quantification

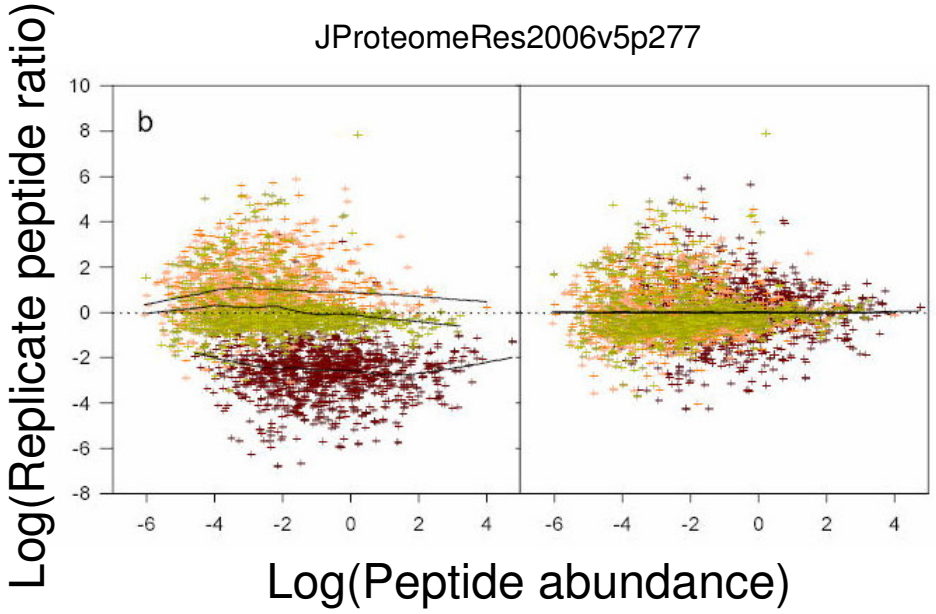
*Extracted Ion Chromatogram Based Method*

***Experimental***

Consistency in sample collection and processing
LC and MS stability
Replicate analysis

***Data Prcessing***

- Data format
- Data visualization
- Normalization
  - Mass recalibration
  - LC retention time alignment
  - Abundance normalization
- Data quality assessment
- Result analysis
  - Difference detection
  - Multivariate analysis

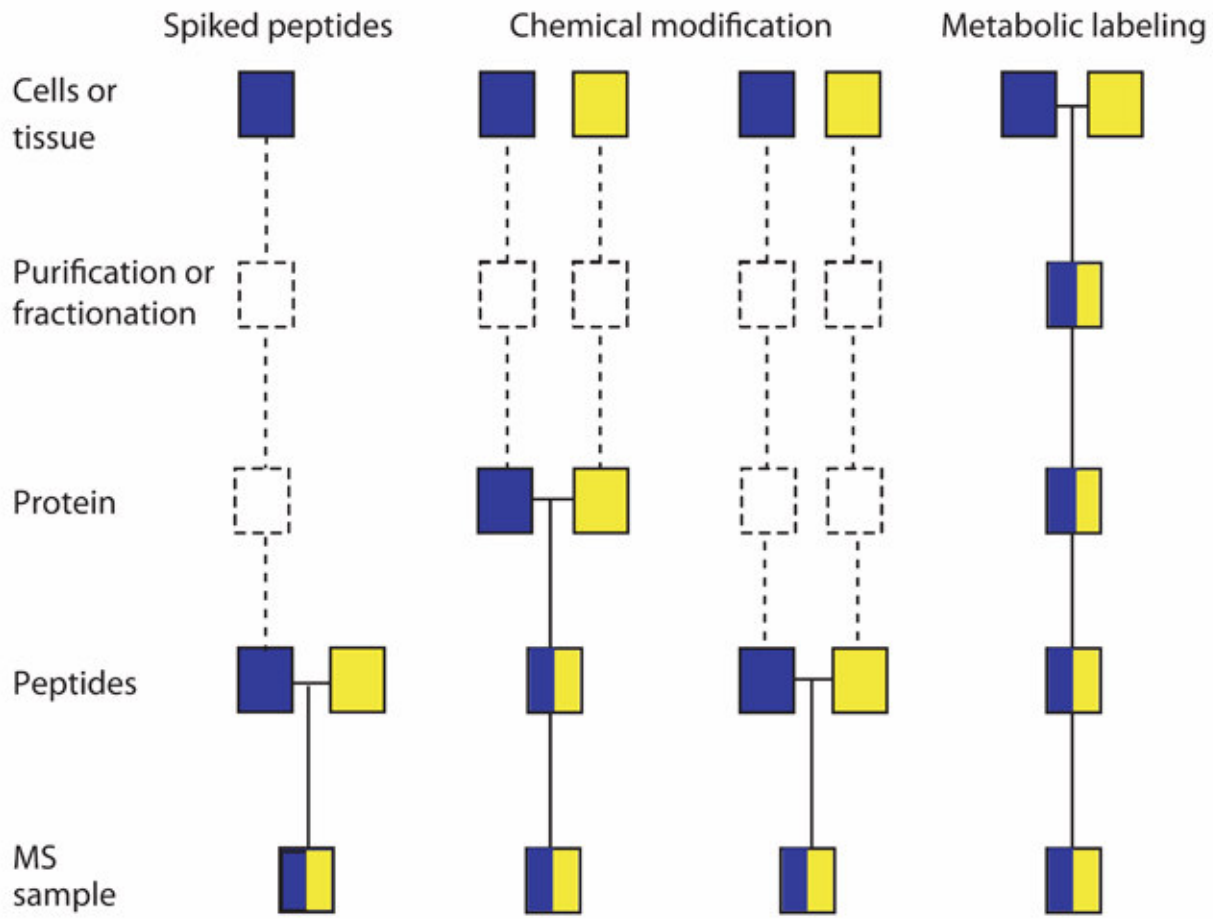


# Stable Isotope Labeling

# Experimental Strategies

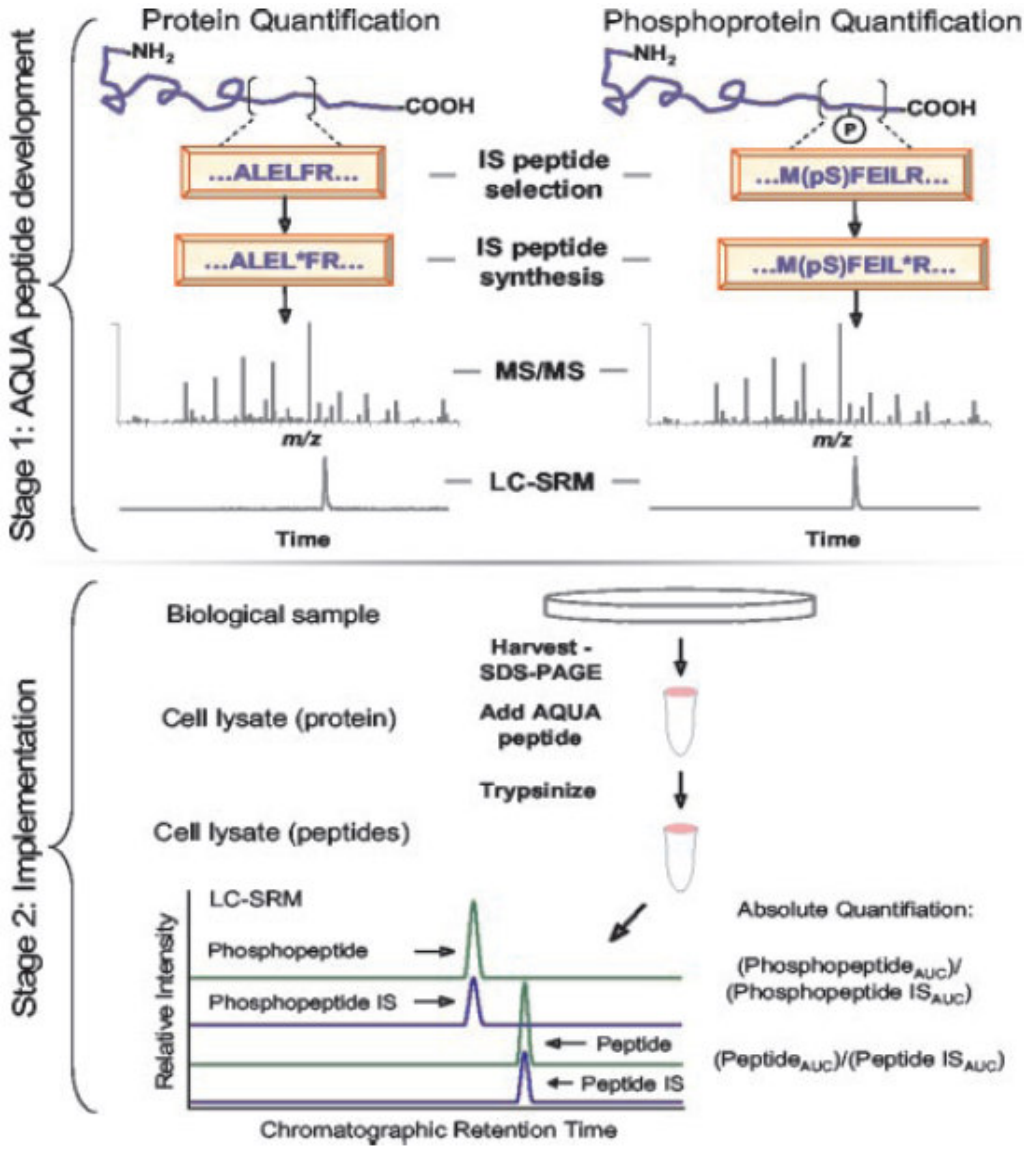
*Introduction of Controls in Stable Isotope Labeling*

(QconCAT:
Spiked Q-proteins)




Proteomics2008v8p4873

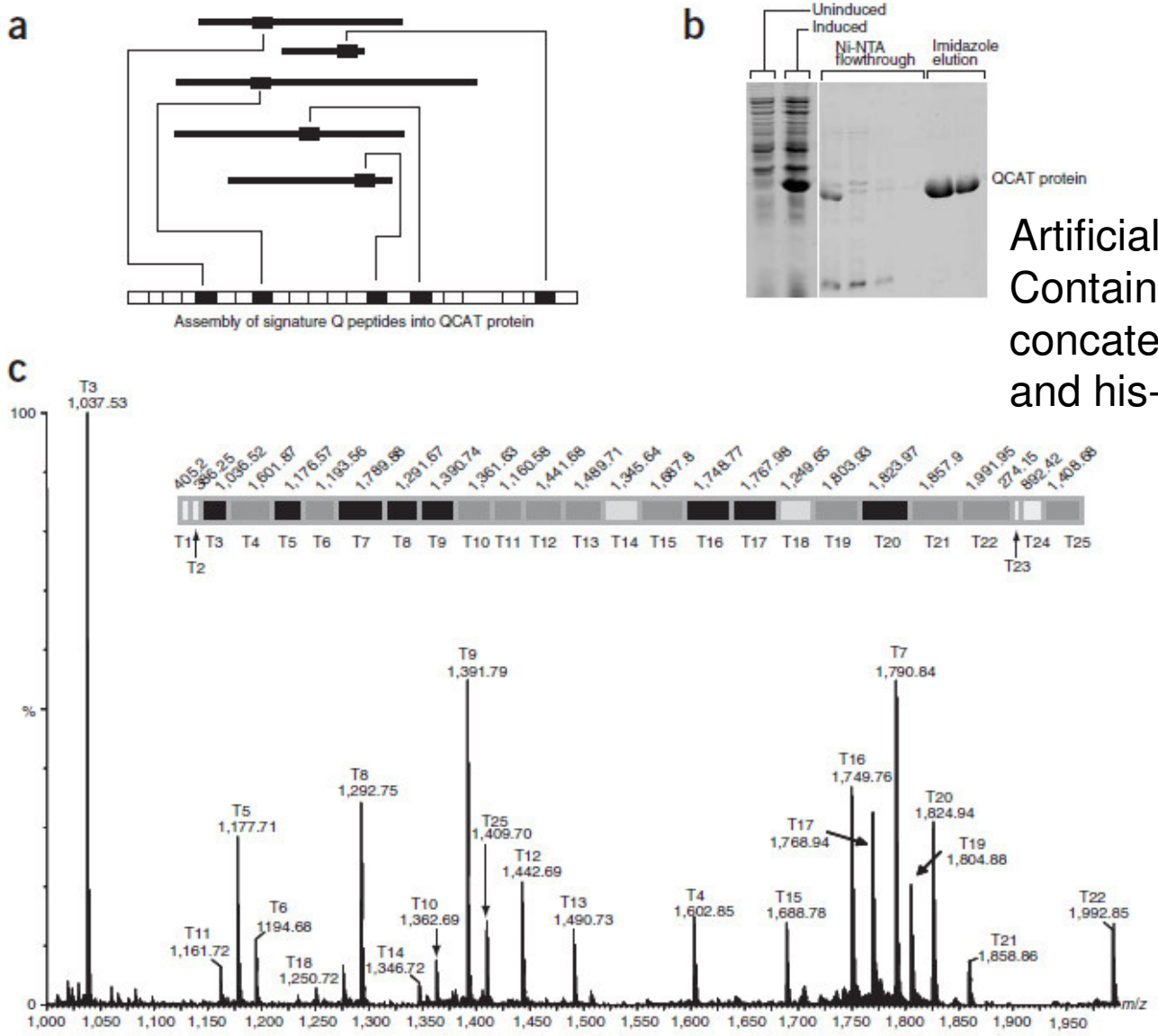
# Absolute Quantification (AQUA) with Internal Standard Peptides




PNAS2003v100p6940

# QconCAT

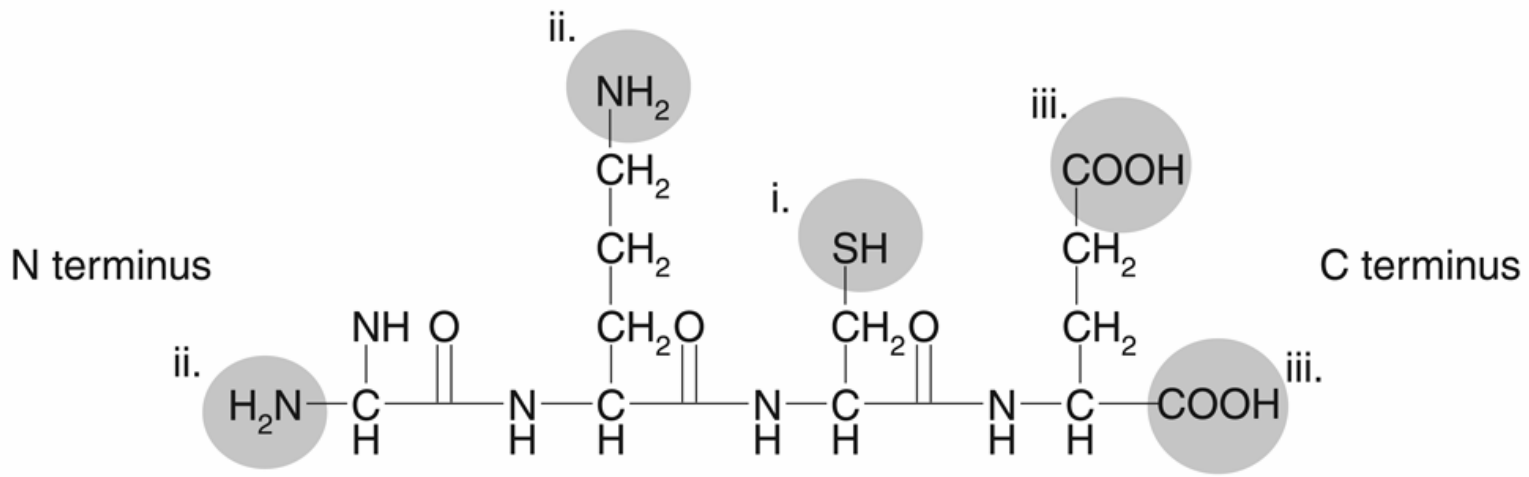
*Tryptic standard peptides from artificial proteins*




Artificial protein
Containing
concatenated tryptic Q peptides
and his-tag

NatMethods2005v2p587

# Stable Isotope Chemical Labeling




i. Sulfhydryl group–directed (cysteine)

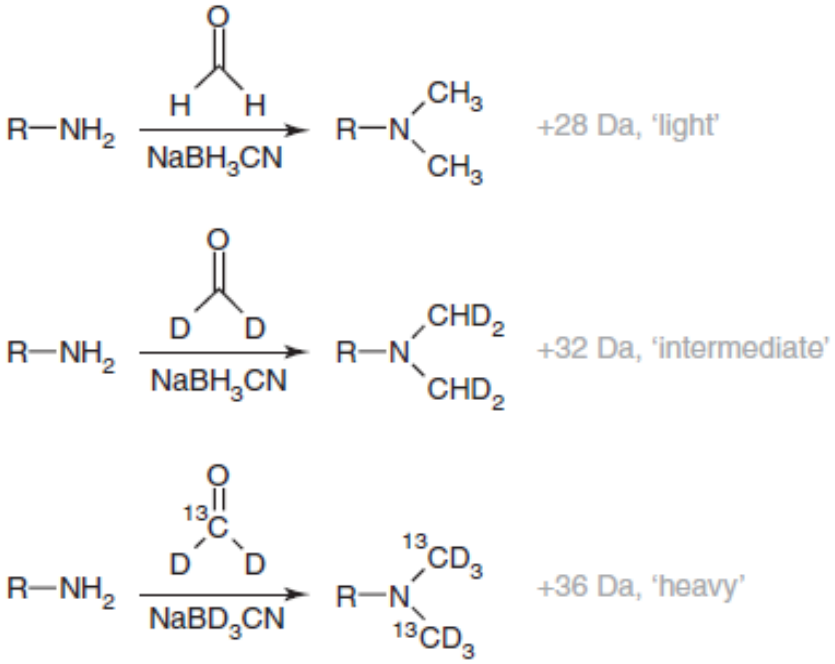
ii. Amine-directed (amino terminus and ε-amino group of lysine)

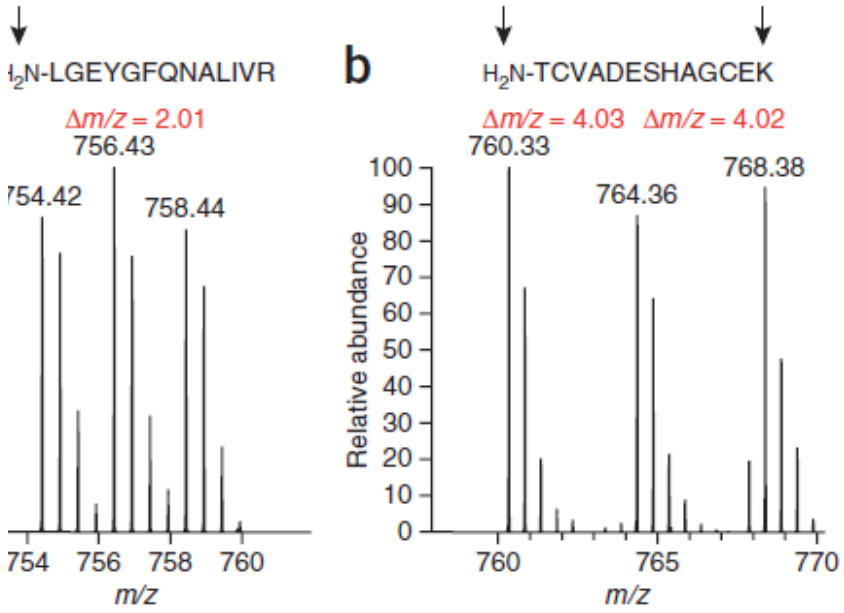
iii. Carboxyl-directed (carboxyl terminus, aspartic and glutamic acid)

# stable isotope chemical labeling

*dimethyl labeling of ε-amine*

- Can be performed (a) in-solution, (b) online with LC-MS or (c) on-column(SepPak)
- 2 carbons and 6 hydrogens for labeling – in principle
- Quantification in MS or survey scan level



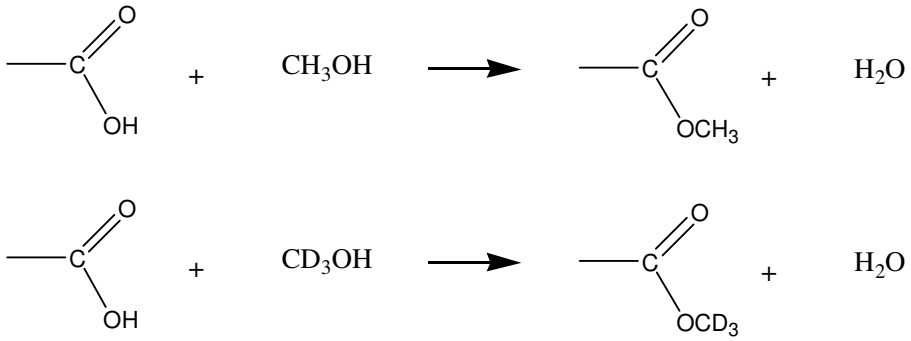



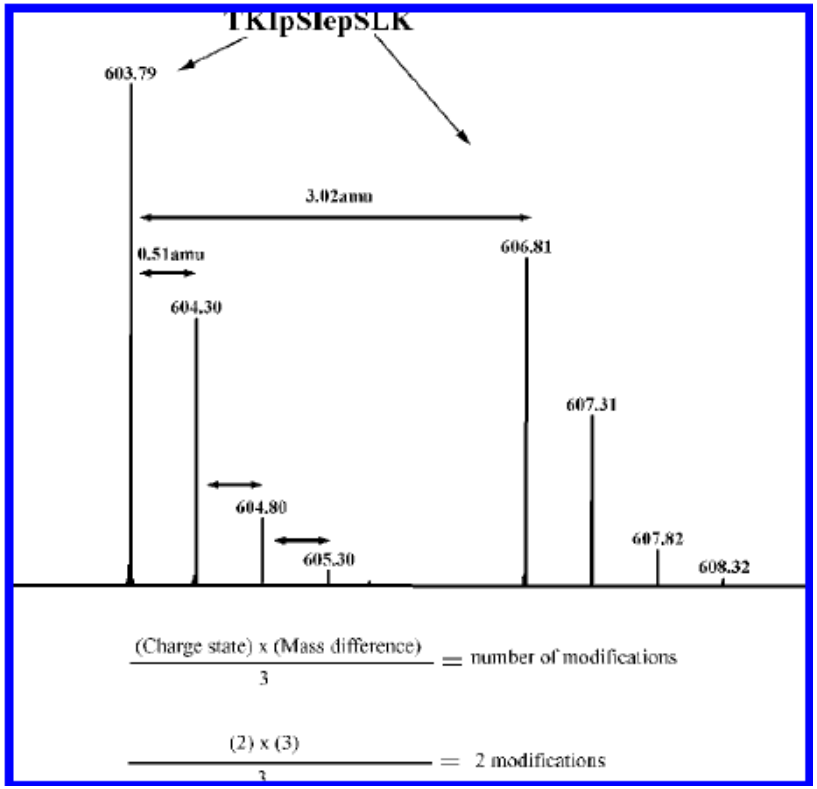

NatProtocol2009v4p484

# stable isotope chemical labeling

*Methyl esterification of carboxyl groups*

- Perform in dry condition (sample, glassware, etc)
  - if a peptide has 10 Es or Ds and single site labeling efficiency is 99%, only 90% of peptide is fully labeled
- Help to increase charge states (better sensitivity)
  - decrease acidic groups (better fragmentation)
  - improve phosphopeptide enrichment
- Quantification in MS or survey scan level

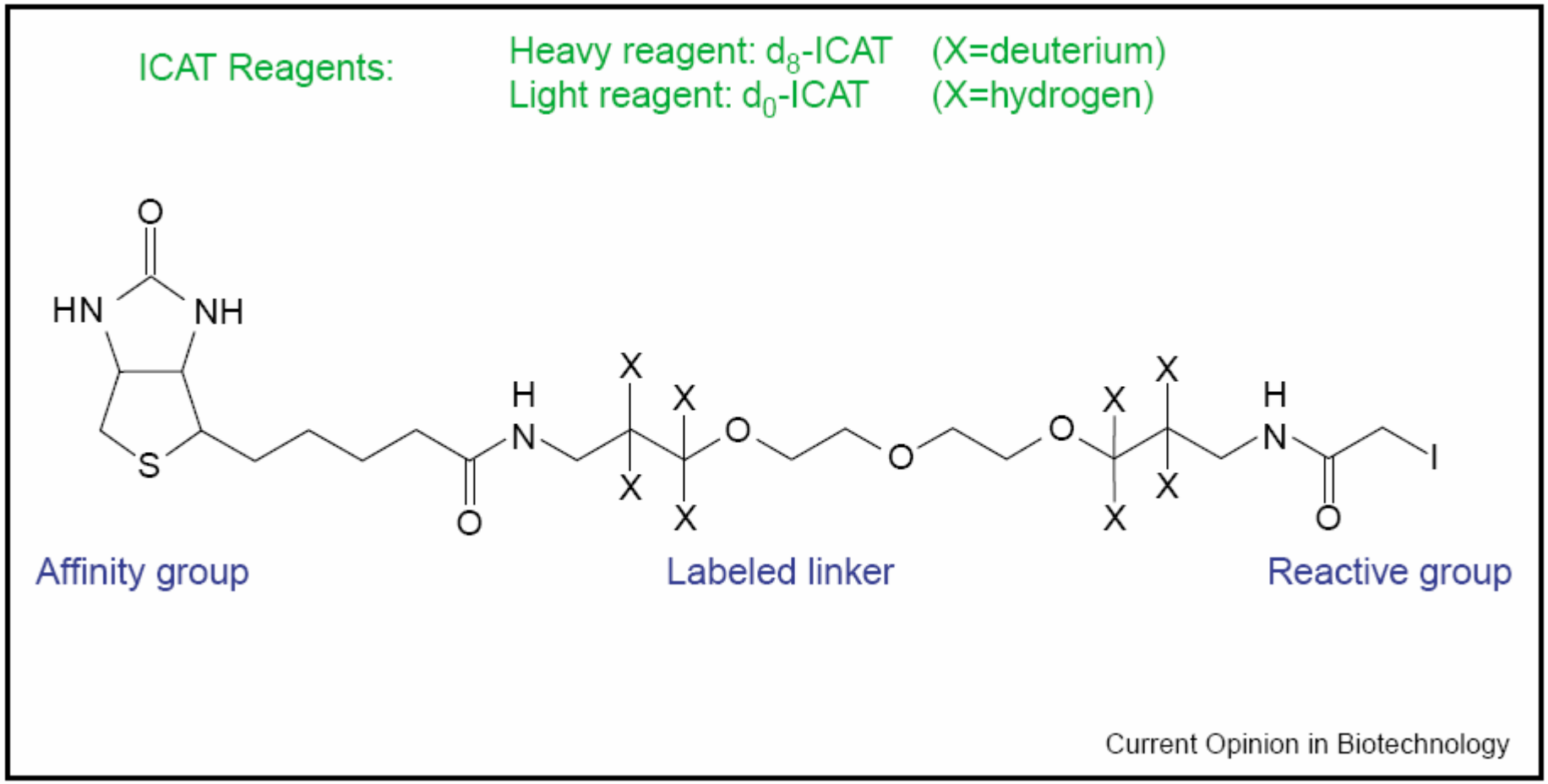



JProteomeRes2009v8p1431

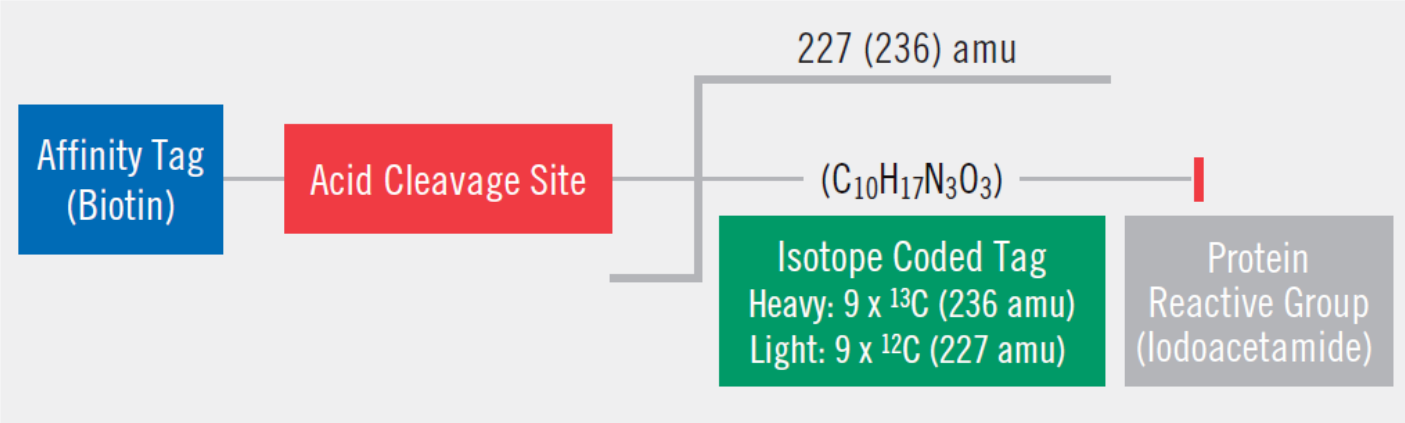
# Isotope coded affinity tag (ICAT)

- Purify peptides by streptavidin beads
- Only labels cysteines
- The tag is too large for MS
- Difficult to elude peptides




CurOpinionBiotechnol2000v11p396

# Isotope coded affinity tag (ICAT)

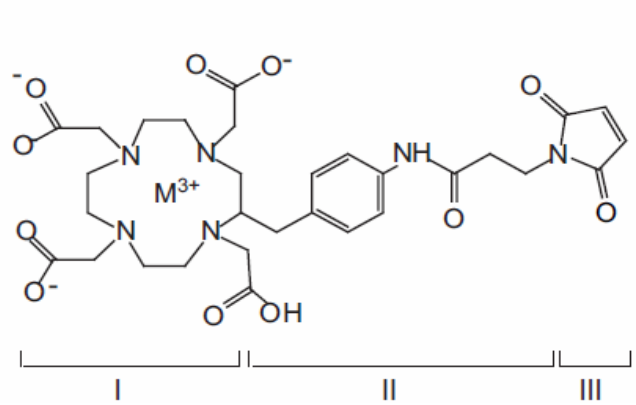



- Acid cleavable linker facilitates release
- Heavy and light carbon allows for co-eluting peptides
- The structure is not disclosed

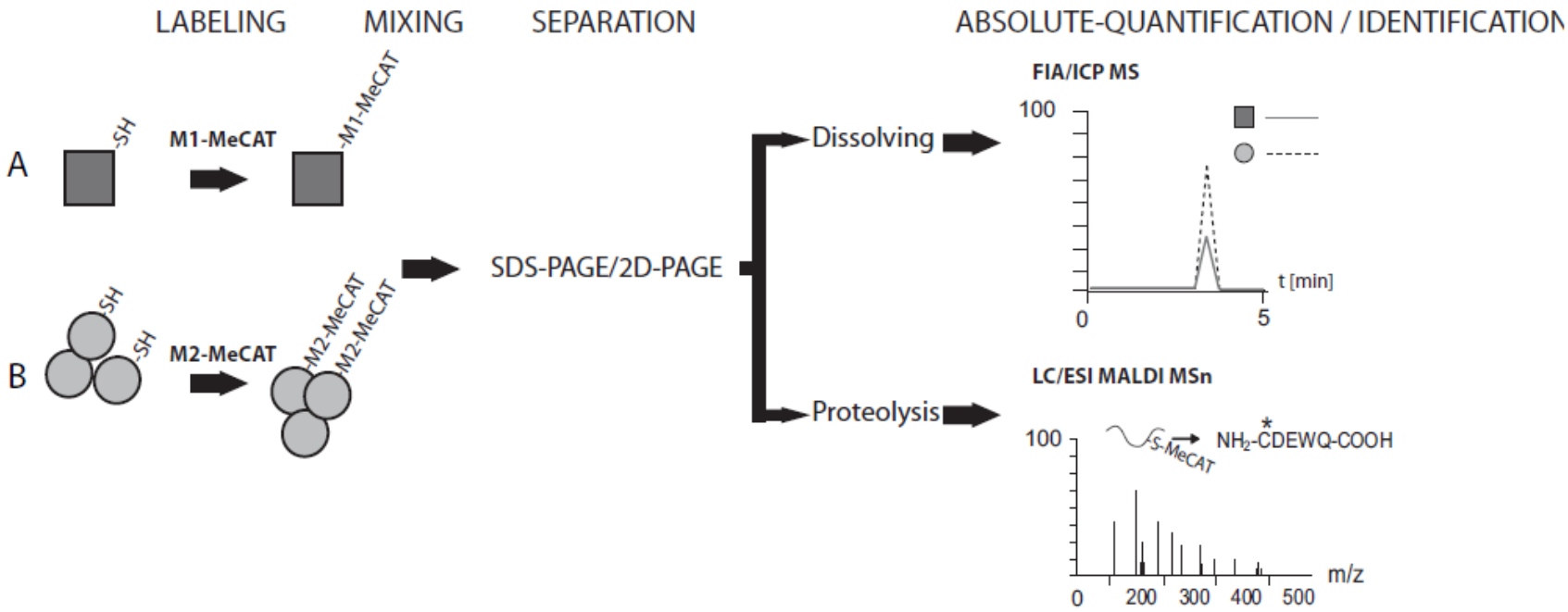
Early usage: AnalChem2006v78p4543
http://www.absciex.com/LITERATURE/cms_040324.pdf

# metal-coded affinity tag (MeCAT)

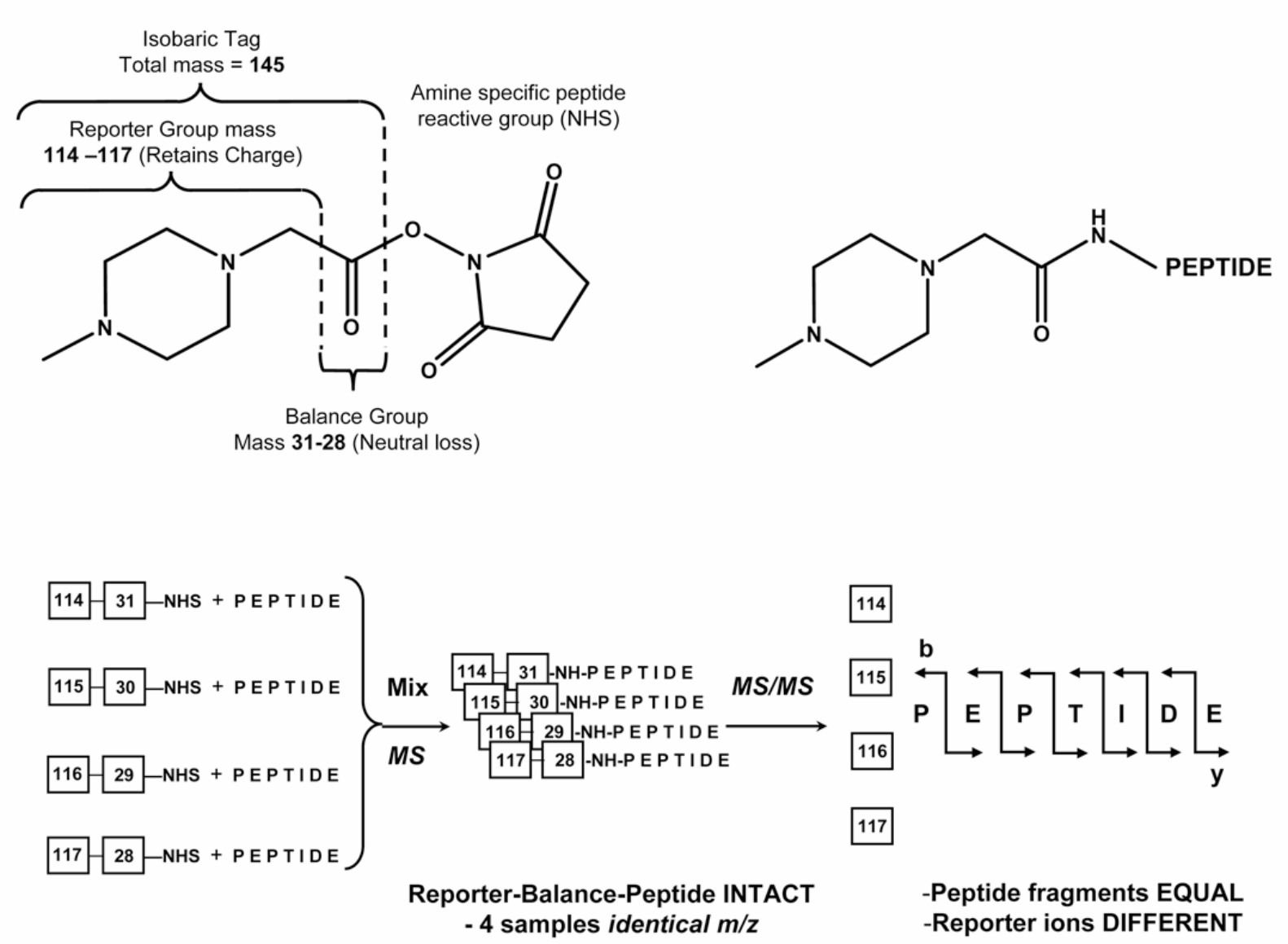
- Labeling group: cysteines
- Metals used: $^{159}Tb$ $^{165}Ho$ $^{169}Tm$ and $^{175}Lu$
- Detection limit: 100amole




metal coding | spacer | thiol-specific labeling

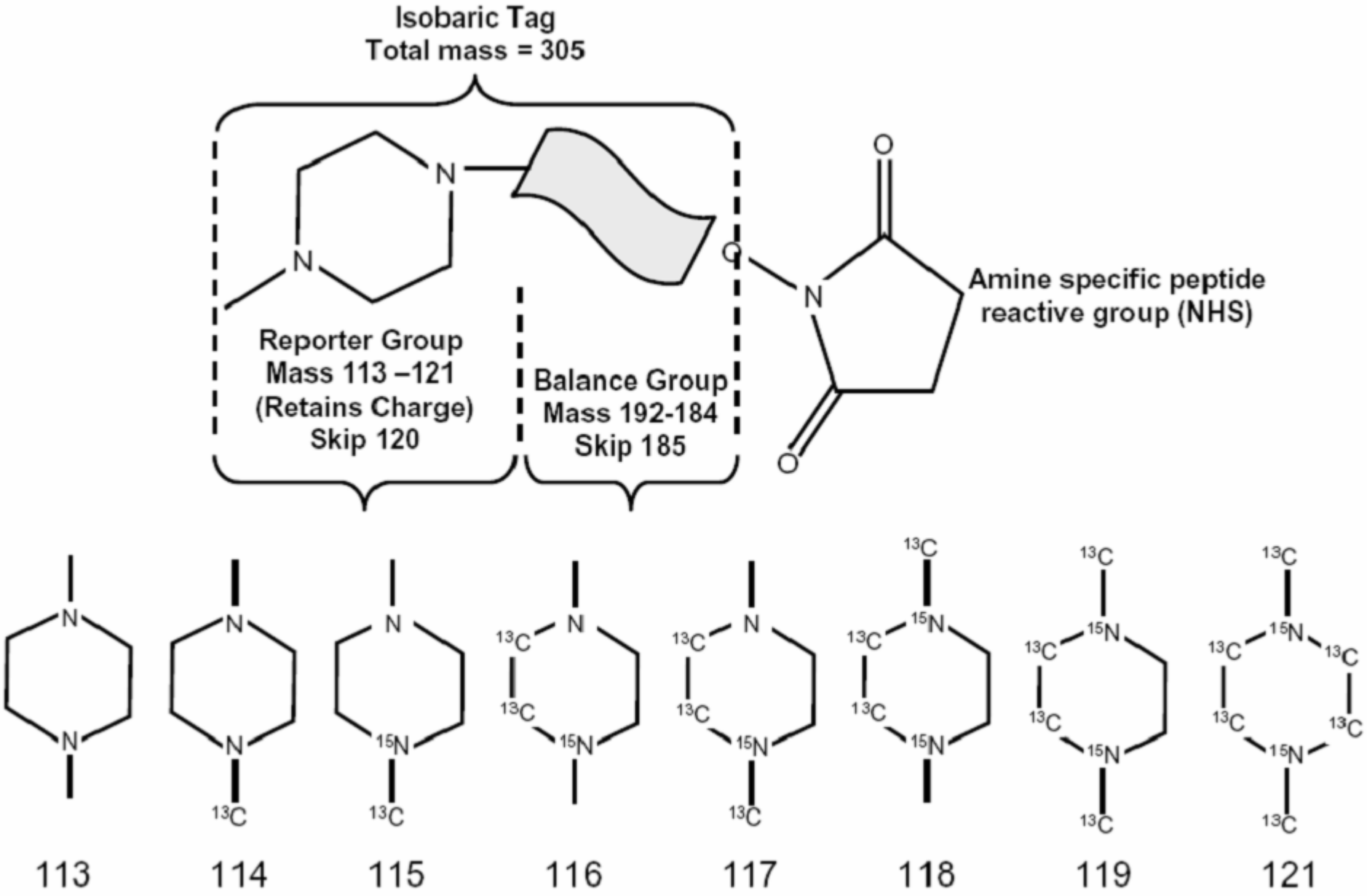


# Isobaric Tags for Relative and Absolute Quantitation (iTRAQ)




MolCelProteomics2004v3p1154

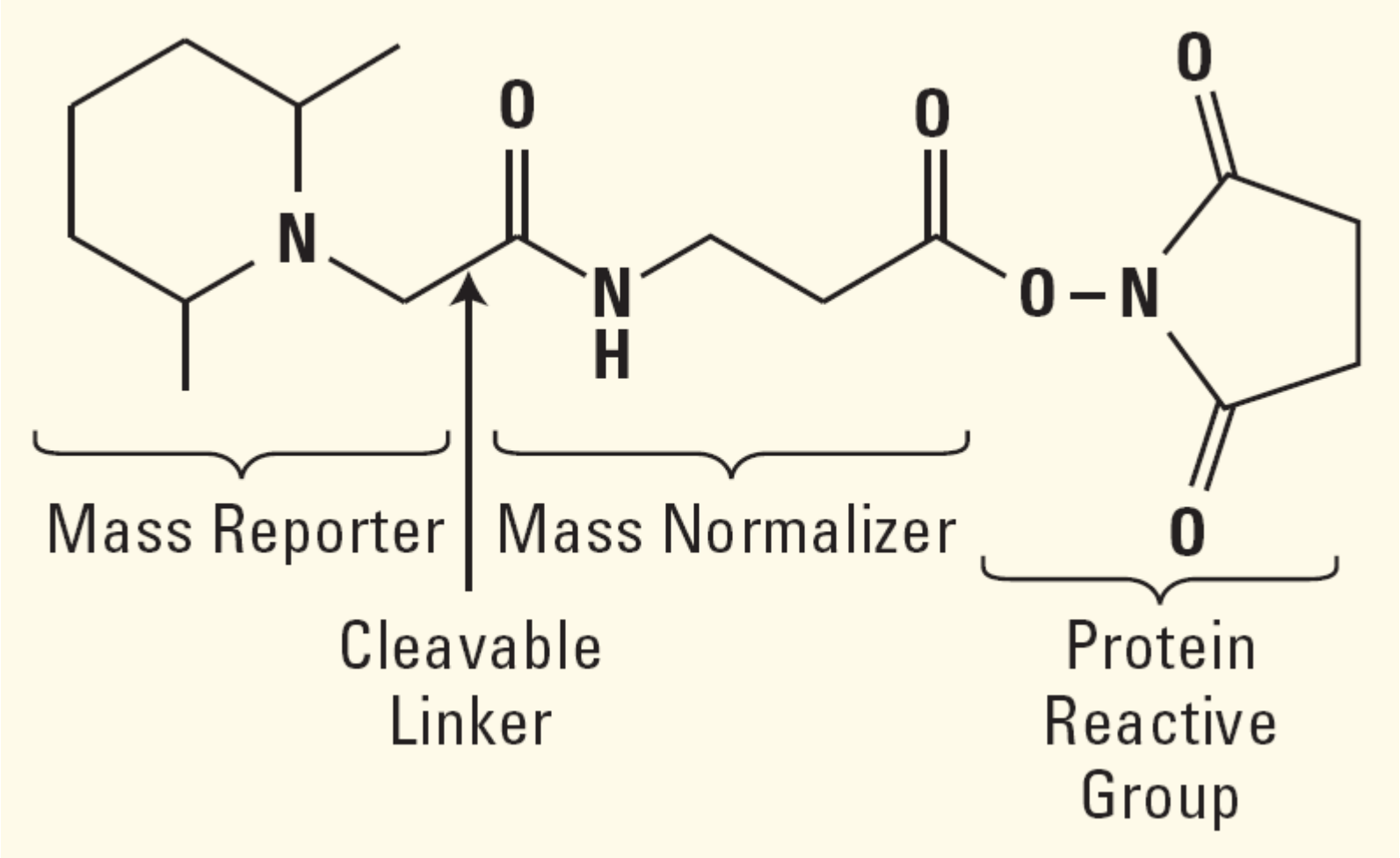
# Isobaric Tags for Relative and Absolute Quantitation (iTRAQ)




MolCelProteomics2008v7p853

# Tandem Mass Tag (TMT)

6 Channels




http://www.piercenet.com/Objects/View.cfm?type=Page&ID=B60171CD-5056-8A76-4E36-958F5C186495#tmtstruct

# MS Instrumentation for iTRAQ

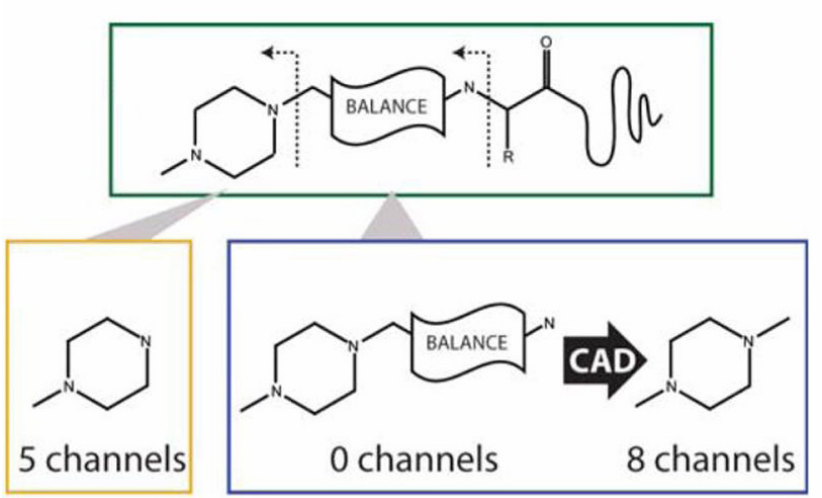
*iTRAQ was initially introduced on Q-TOF type instruments*

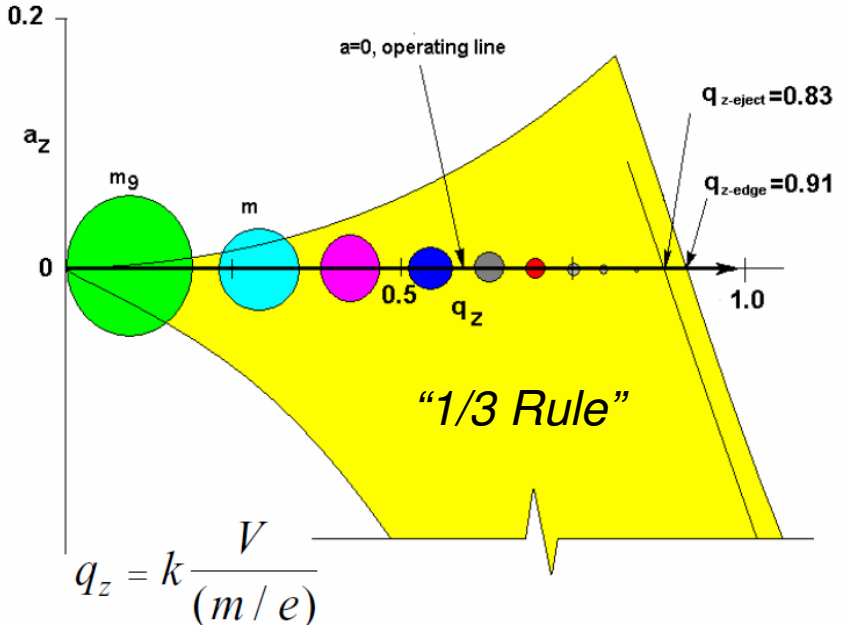
- MS3 in ion traps
(http://www.thermo.com/eThermo/CMA/PDFs/Various/File_27402.pdf)

- PQD for detecting fragment ions on ion traps
(MCP2008v7p1702)

- HCD for detection on Orbitraps
LTQ Velos Oritrap

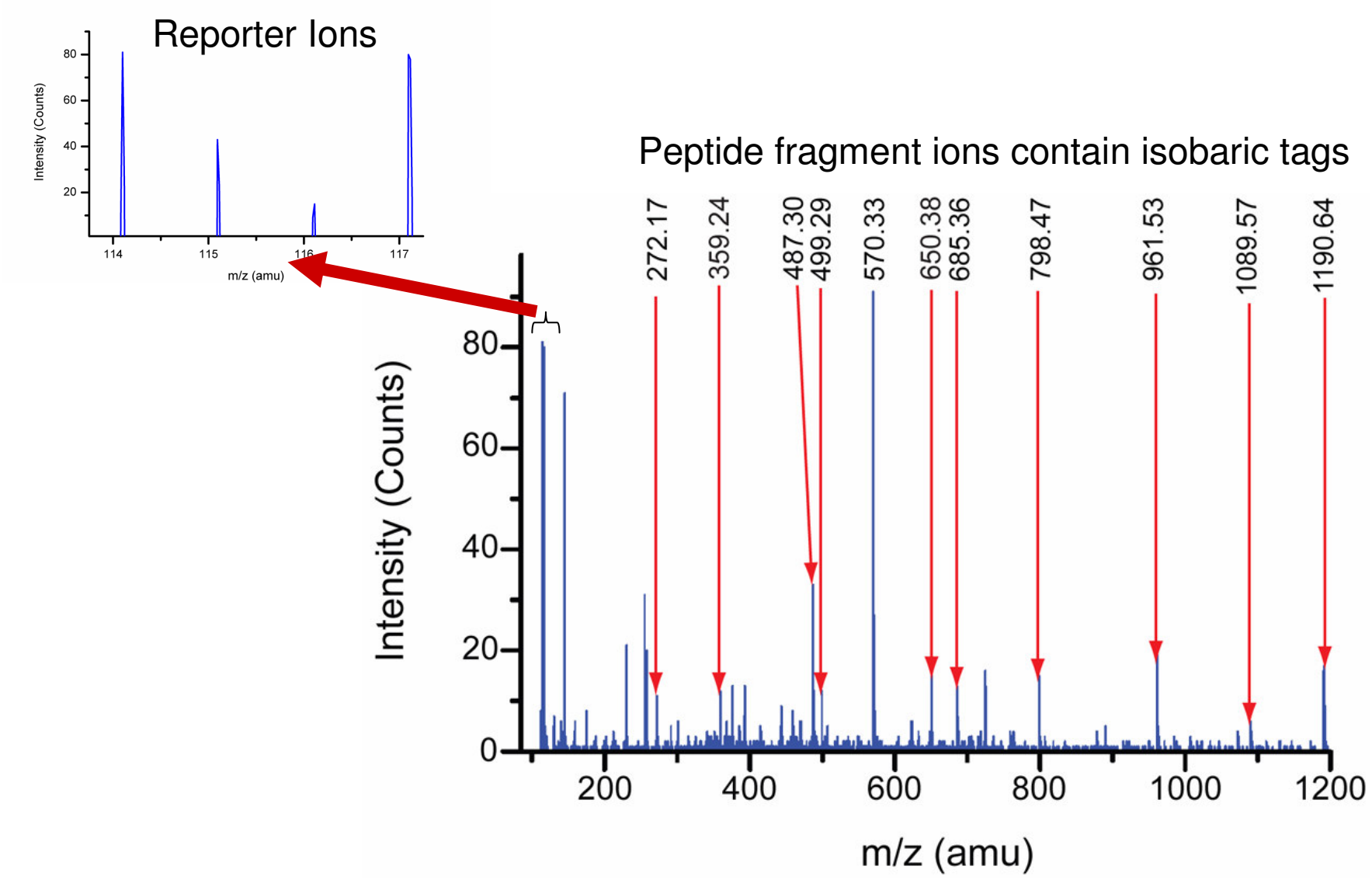
- ETD (or combined with CID) can generate reporter ions
(AnalChem2009v81p1693)



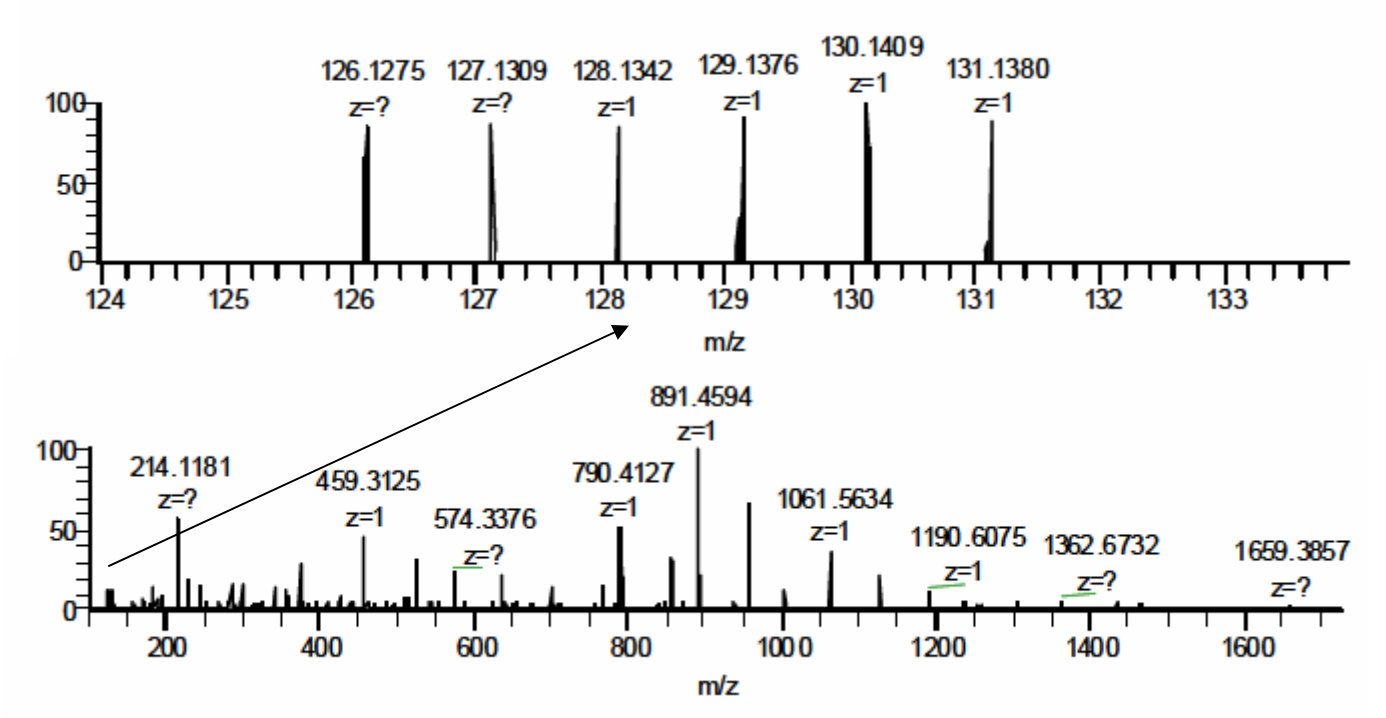



# iTRAQ

*Both quantification and identification info is contained in MSMS*



# Quantification with TMT




**E. Coli digest**

Quantifiable peptides

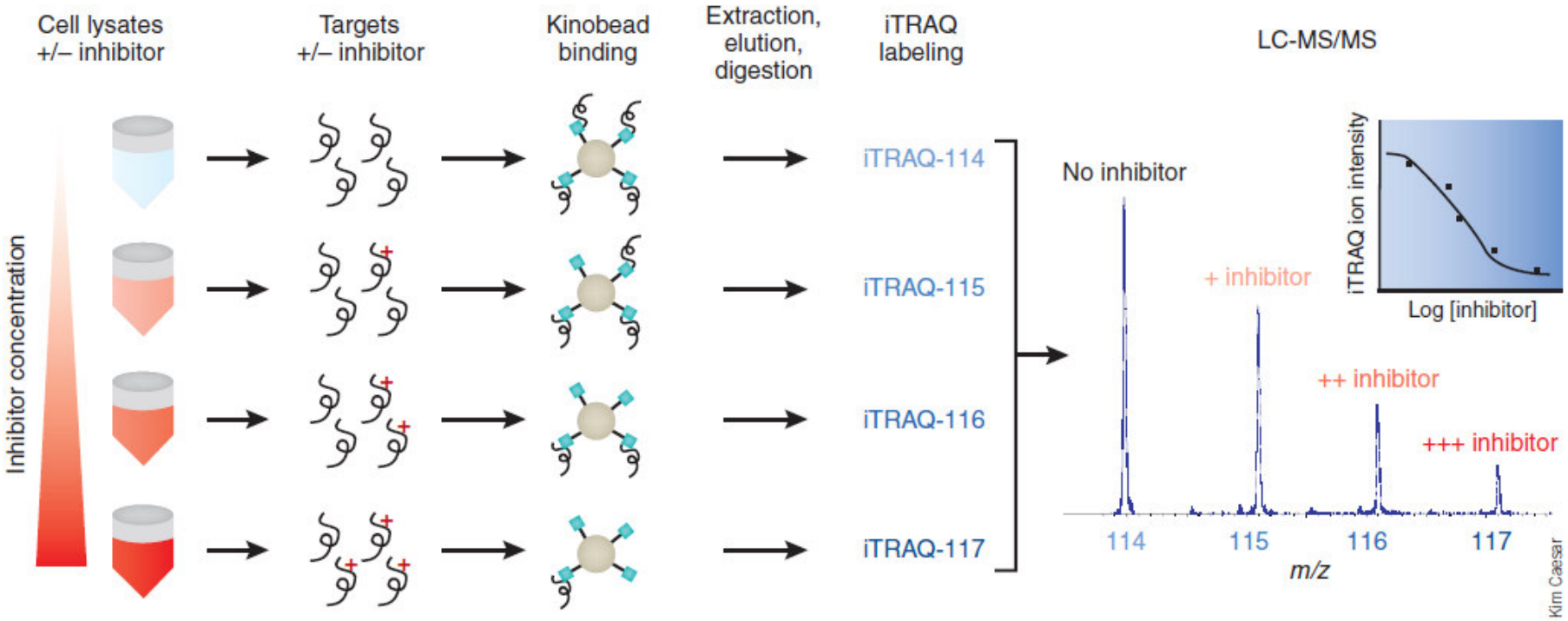
- 74% (40ng)
- 81% (80ng)

Quantification Errors

- 20% (40ng)
- 15% (80ng)

Mass Spectrometer: Thermo Scientific LTQ Orbitrap Velos
MS Resolution: 60000
MS2 Resolution: 7500
MS AGC target: 1e6
MS/MS AGC target: 5e4
Exclusion mass tolerance: 10 ppm
Injection Time FTMS/MS: 200 ms
Full MS mass range: 400–1400 *m/z*
MS/MS Mass range: 100–2000 *m/z*
Isolation 1.2 Da
MS/MS Events: Full MS in Orbitrap followed by top ten data dependent HCD MS/MS
CE for HCD: 45%

# iTRAQ
## *Quantification of Phosphorylation Levels*



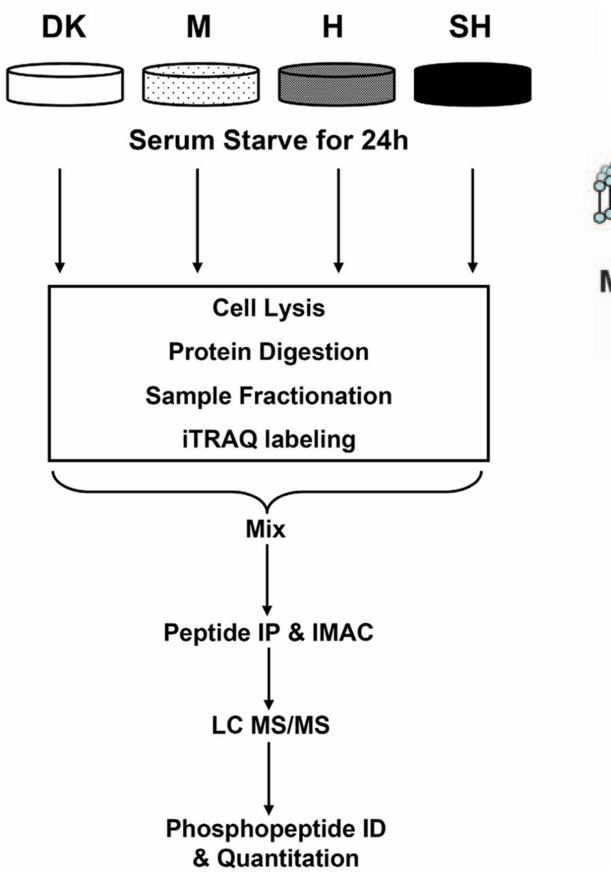

Kinobeads: ~1000 kinase inhibitors
(http://www.cellzome.com/files/kinobeads.pdf)

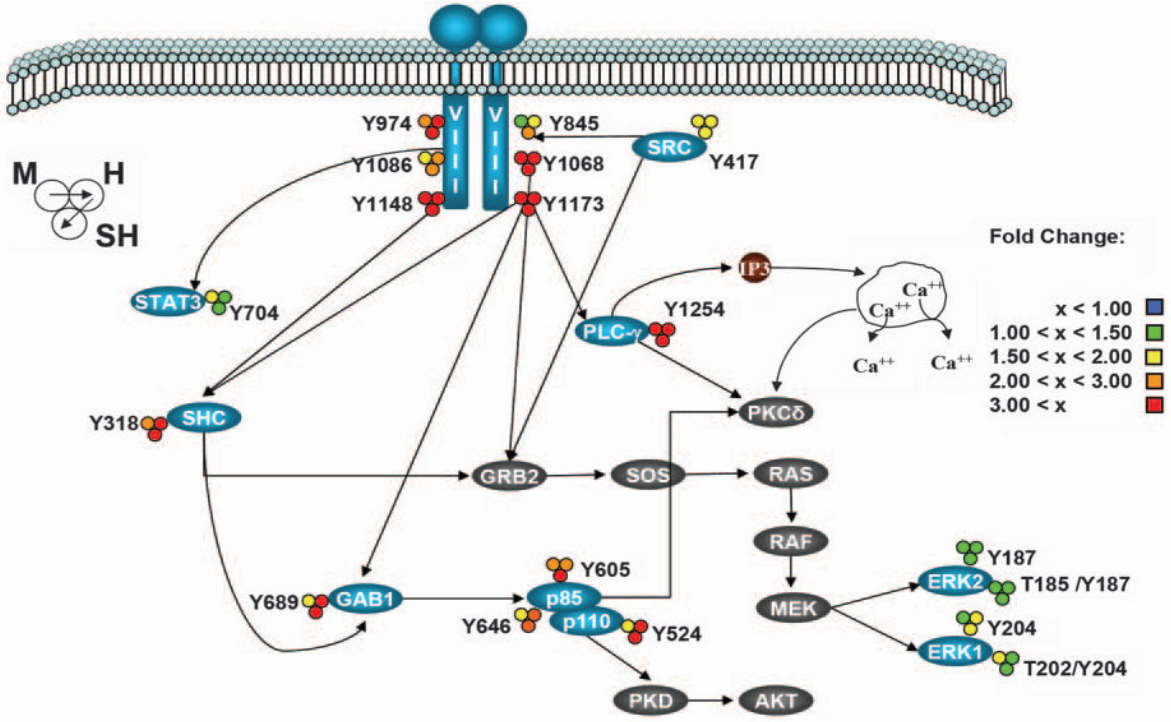
NatBiotech2007v25p994

# iTRAQ

*Phosphorylation Levels and Signaling Pathways*

EGFR pathways: Different receptor level in U87MG glioblastoma cell lines

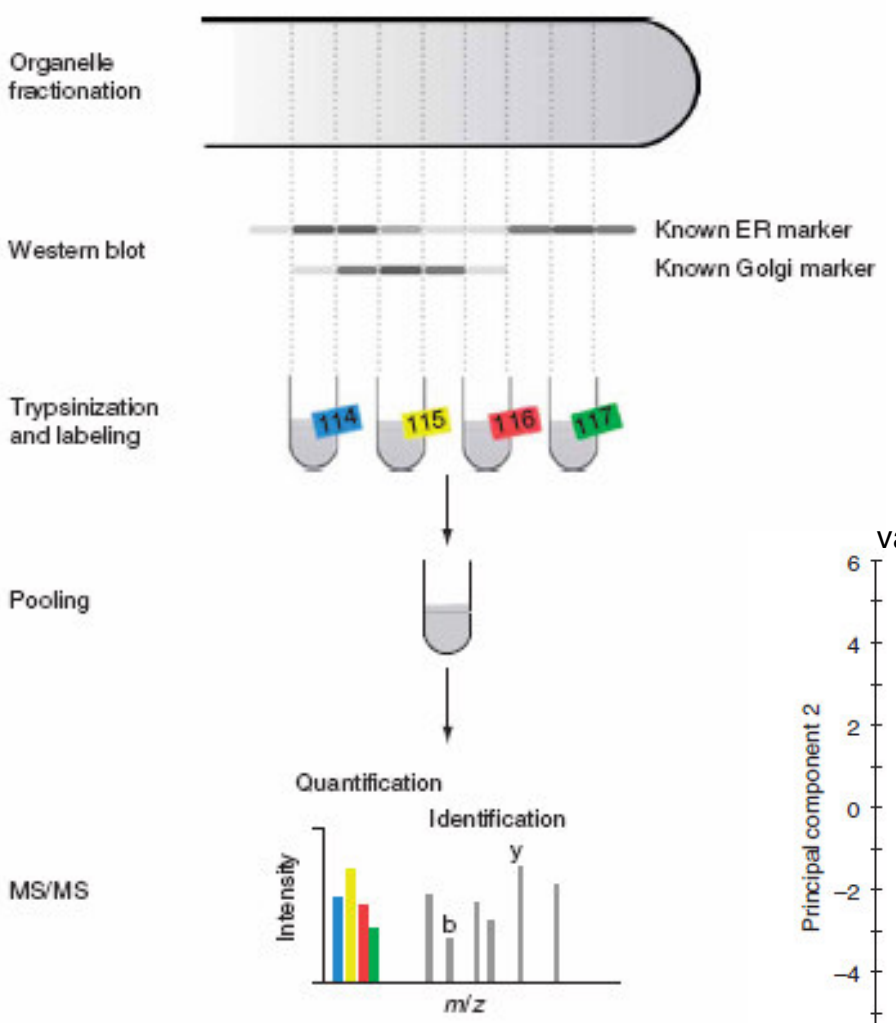


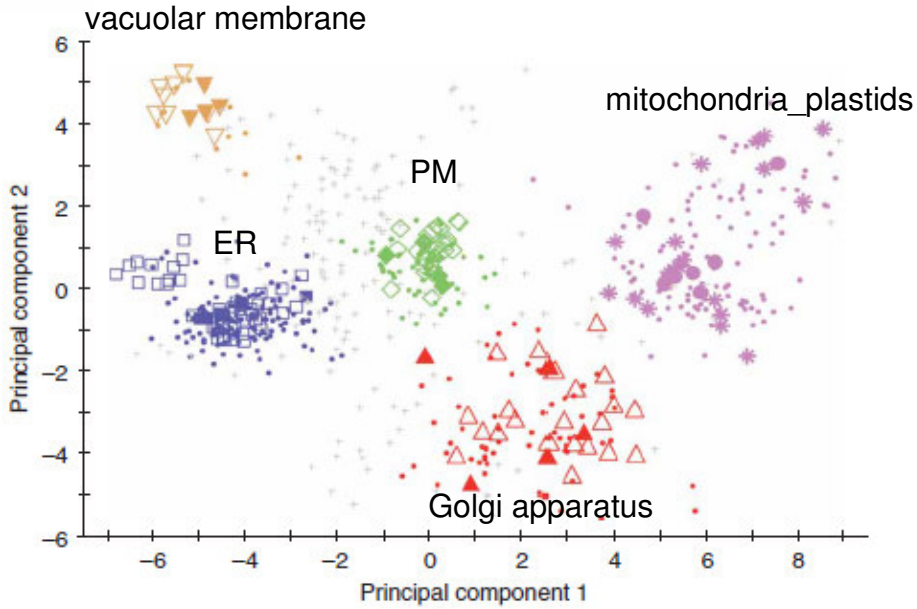

PNAS2007v104p12867

# iTRAQ

*localization of organelle proteins by isotope tagging (LOPIT)*

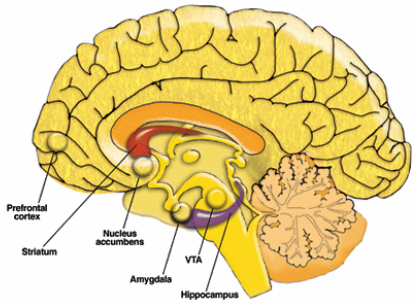



Principal component analysis (PCA)
Reduction of dimensions

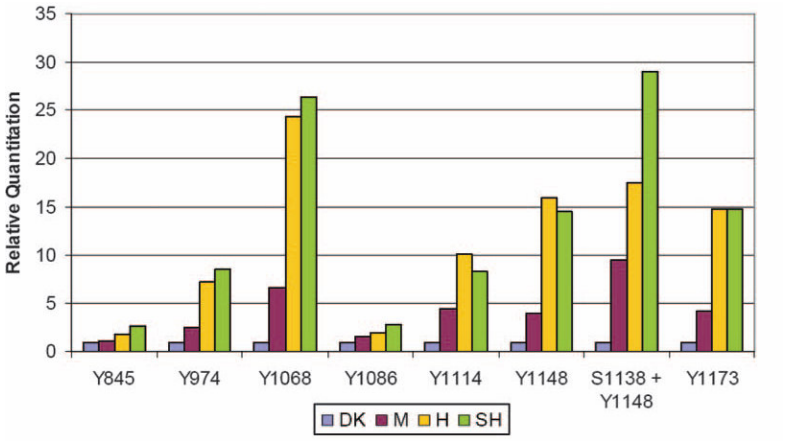



NatProtocols2006v1p1778

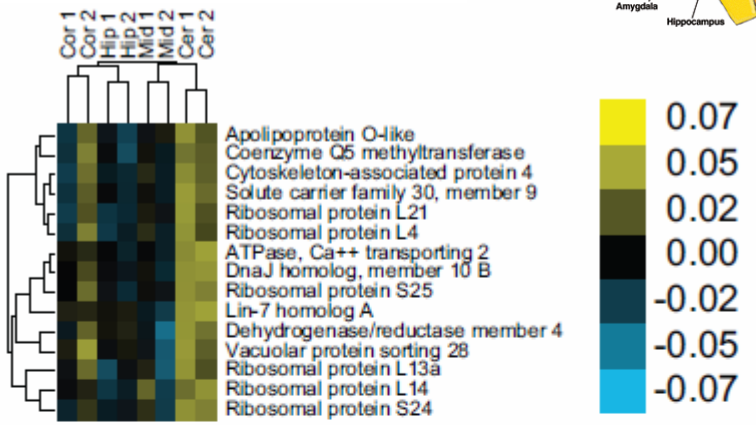
# Quantification
## *Data Analysis – Clustering Analysis*



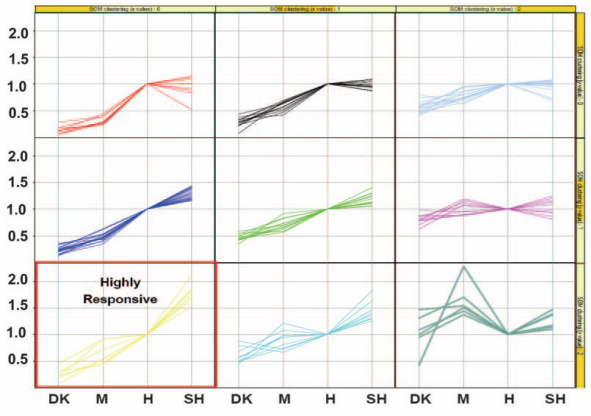

PNAS2007v104p12867




MCP2008v7p684




## ***Self Organized Maps (SOM)***



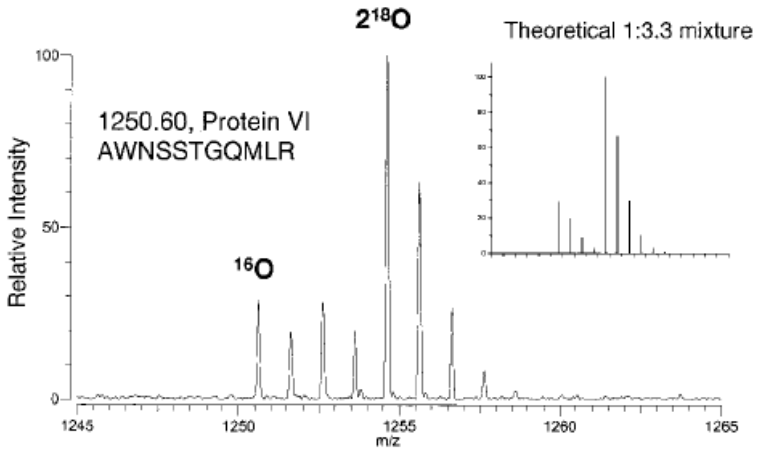

## ***Hierarchical Clustering***

| Unigene | Candidate Interactors | Pearson Correlation |
|---|---|---|
| Mm.334658<br>Mm.260938 | Calcium channel, alpha 1A<br>Plakophilin 4 | 0.963 |
| Mm.4424<br>Mm.34977 | Calcium channel, alpha 1B<br>Neuron navigator 1 | 0.989 |
| Mm.235182<br>Mm.384353 | CaMK II, gamma<br>Neurobeachin | 0.994 |
| Mm.336569<br>Mm.277235 | Brain-specific angiogenesis inhibitor 3<br>Liprin, alpha 3 | 0.990 |
| Mm.330223<br>Mm.339755 | Calcium channel, beta 4<br>Flightless I homolog | 0.959 |
| Mm.240421<br>Mm.27005 | Ubiquitin specific peptidase 30<br>Visinin-like 1 | 0.964 |
| Mm.89320<br>Mm.38009 | Glycine receptor, alpha 1<br>Phospholipase C, beta 4 | 0.994 |
| Mm.200497<br>Mm.291463 | enoyl-Coenzyme A hydratase, alpha<br>enoyl-Coenzyme A hydratase, beta | 0.995 |

Cor 1 Cor 2 Hip 1 Hip 2 Mid 1 Mid 2 Cer 1 Cer 2

Reference: PNAS1998v95p14863
Freeware: http://rana.lbl.gov/EisenSoftware.htm

# Proteolytic $^{18}O$ Labeling




AnalChem2001v73p2836

Trypsin introduces a mixture of one and two 18O incorporated peptides

LysN incorporates a single 18O at high pH

Electrophoresis1996v17p945

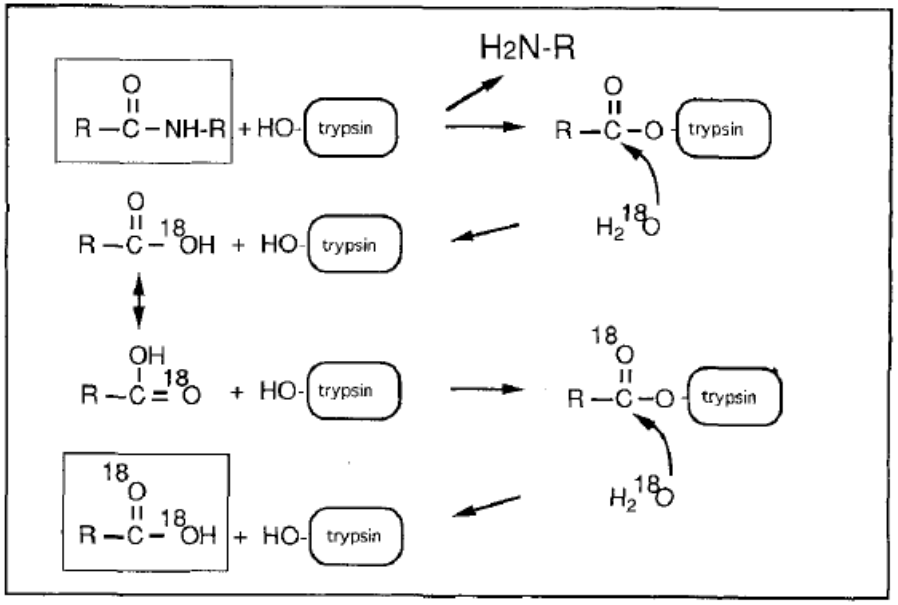


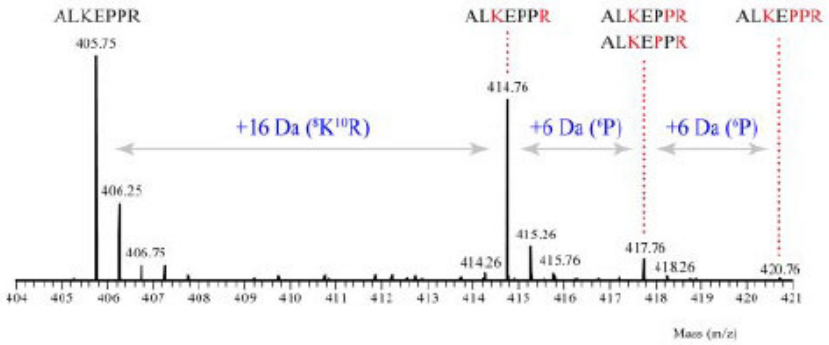

*Figure 2.* Mechanism of trypsin-catalyzed incorporation of $^{18}O$ into proteolytic fragments. One atom of $^{18}O$ is introduced by the reaction leading to cleavage of the peptide bond. A second $^{18}O$ atom is introduced by repeated binding/hydrolysis cycles with the proteolytic peptide fragment as a pseudosubstrate.

JProteomeRes2005v4p507
MCP2005v4p1550

# Metabolic Stable Isotope Labeling

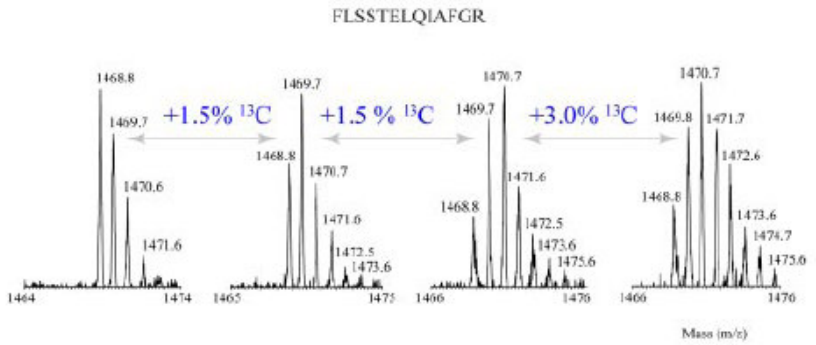
A Stable Isotope Labeling with Amino Acids in Cell Culture

MCP2002v1p376



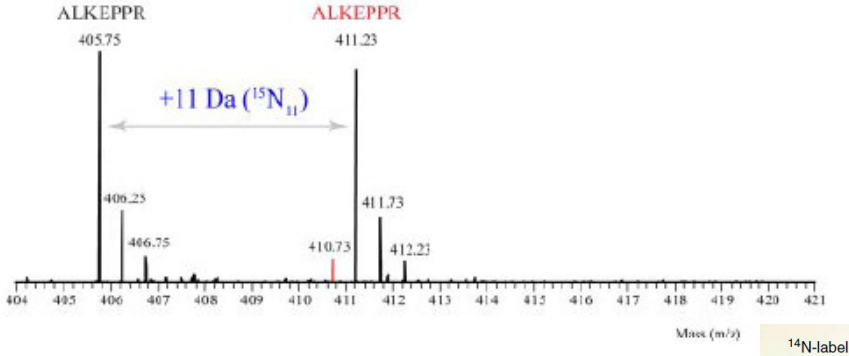

C Subtle Modification of Isotope Ratio Proteomics

Phytochemistry2004v65p1507

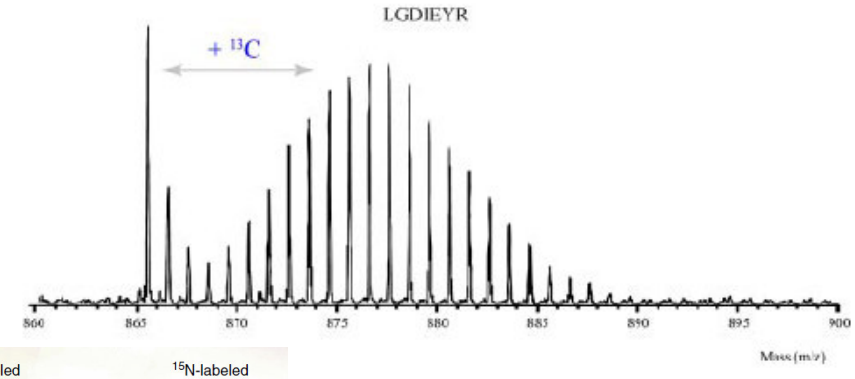


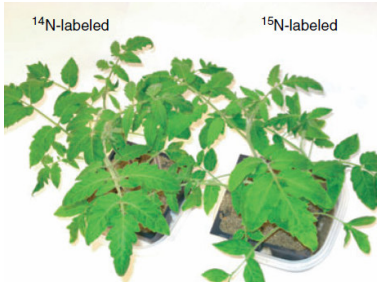

B Heavy Nitrogen Labeling

PNAS1999v96p6591




D Synthesis/Degradation Ratio Mass Spectrometry



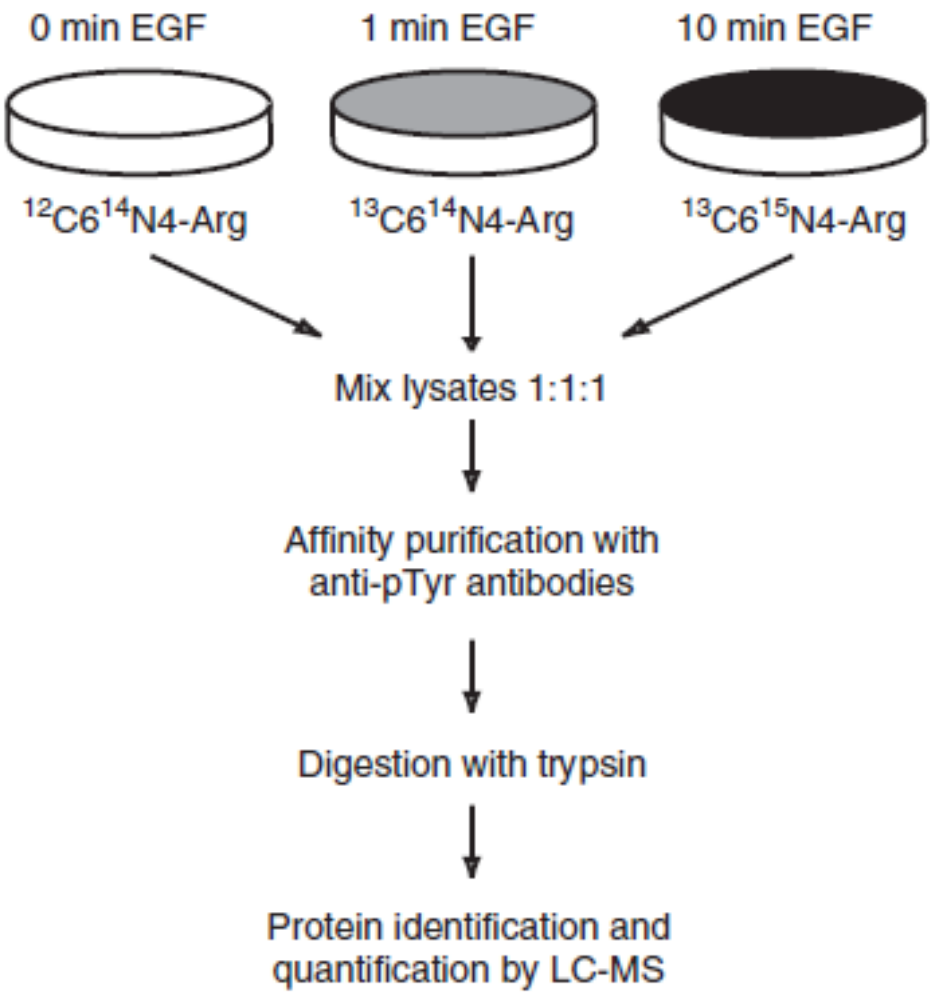



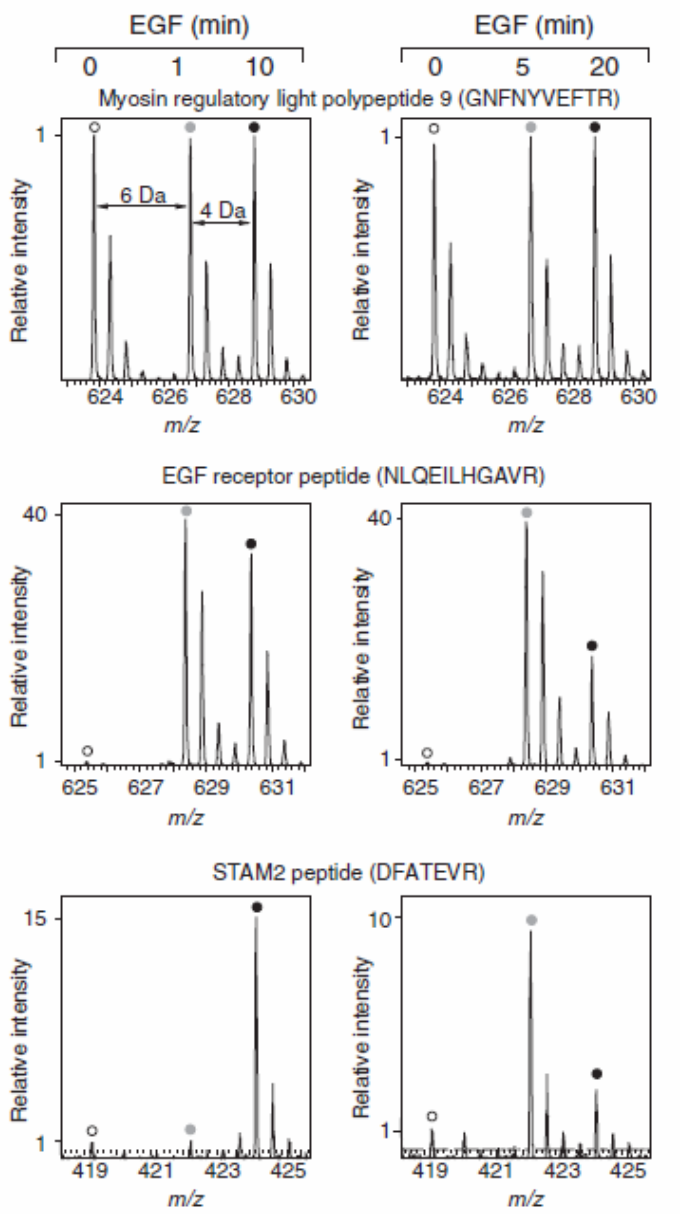

MCP2010v9p11

PlantJ2008v56p840

# Metabolic Stable Isotope Labeling

*Stable Isotope Labeling by Amino Acids in Cell Culture (SILAC)*




NatBiotechnol2004v22p1139

# Multiple Reaction Monitoring(MRM)

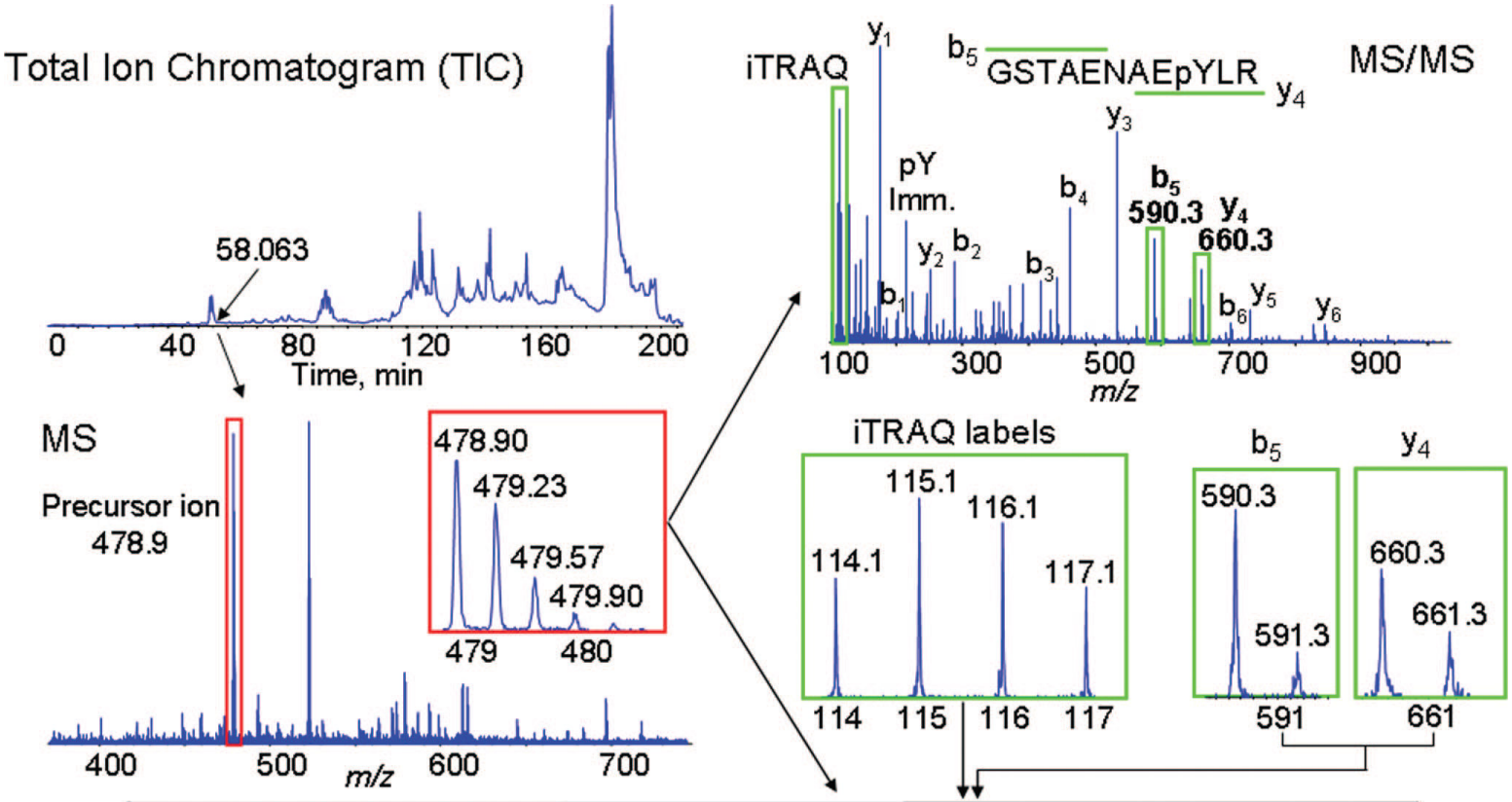
Term coined by Cooks et al in 1978
(AnalChem1978v50p2017)

Used by pharmaceutical industry for years
Drug metabolites analysis

Performance improved by introduction of Qtrap instruments
(DDTtarget2004v3pS31)
Modern instruments have MRM dwell time of 5msec
Sensitivity of <1fmole

# Multiple Reaction Monitoring(MRM)

*information-dependent acquisition (IDA) to MRM Transitions*



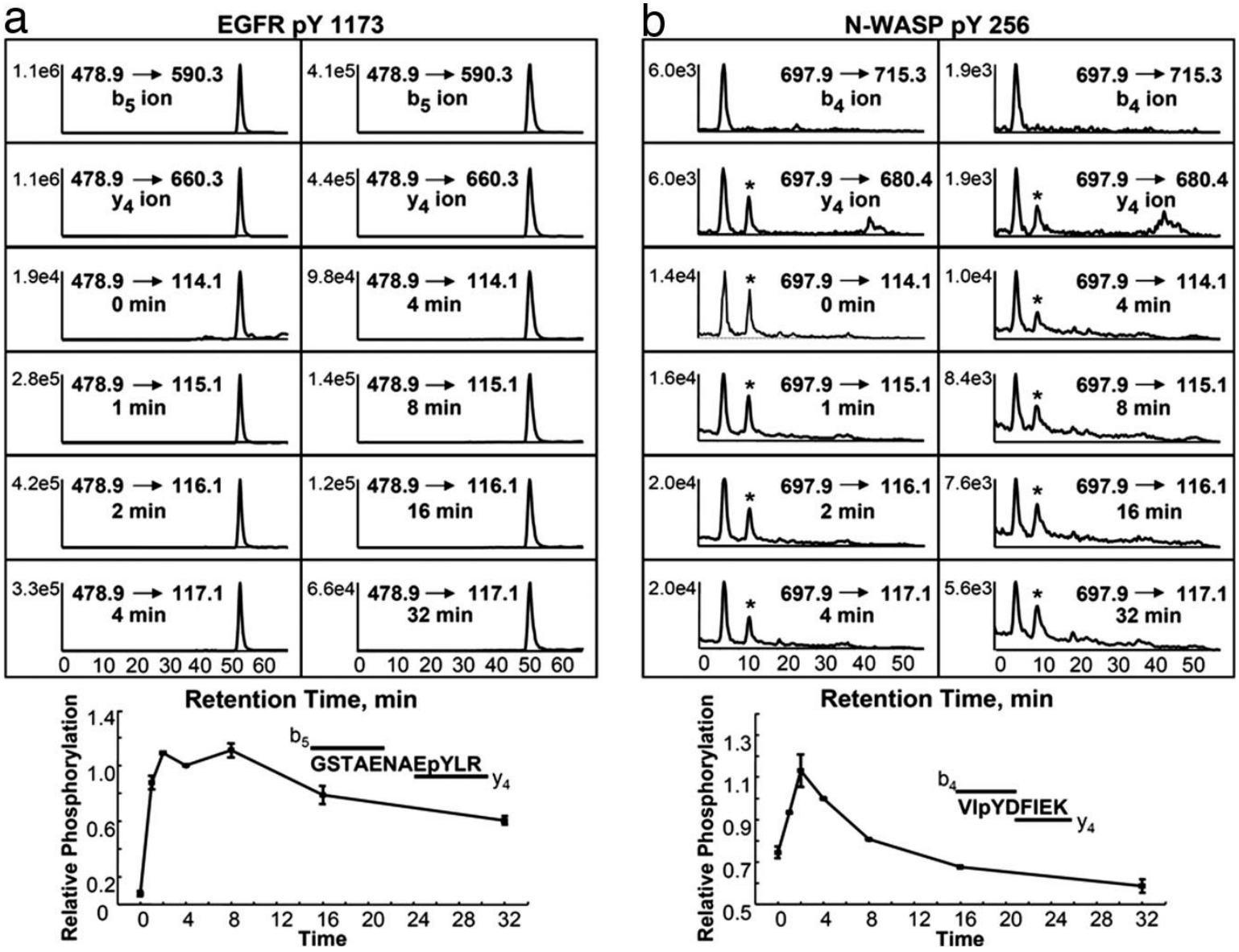

**Parameters to build MRM Acquisition Method**

| Protein | Peptide | pY | Elution Order | Prec. Ion | Frag. Ion | CE (V) | Time (ms) | Ion |
|---|---|---|---|---|---|---|---|---|
| EGFR | GSTAENAEpYLR | 1173 | 15 | 478.9 | 590.3 | 27 | 30 | b5 |
| EGFR | GSTAENAEpYLR | 1173 | 15 | 478.9 | 660.3 | 27 | 30 | y4 |
| EGFR | GSTAENAEpYLR | 1173 | 15 | 478.9 | 114.1 | 41 | 30 | iTRAQ |
| EGFR | GSTAENAEpYLR | 1173 | 15 | 478.9 | 115.1 | 41 | 30 | iTRAQ |
| EGFR | GSTAENAEpYLR | 1173 | 15 | 478.9 | 116.1 | 41 | 30 | iTRAQ |
| EGFR | GSTAENAEpYLR | 1173 | 15 | 478.9 | 117.1 | 41 | 30 | iTRAQ |

PNAS2007v104p5860

# Multiple Reaction Monitoring(MRM)

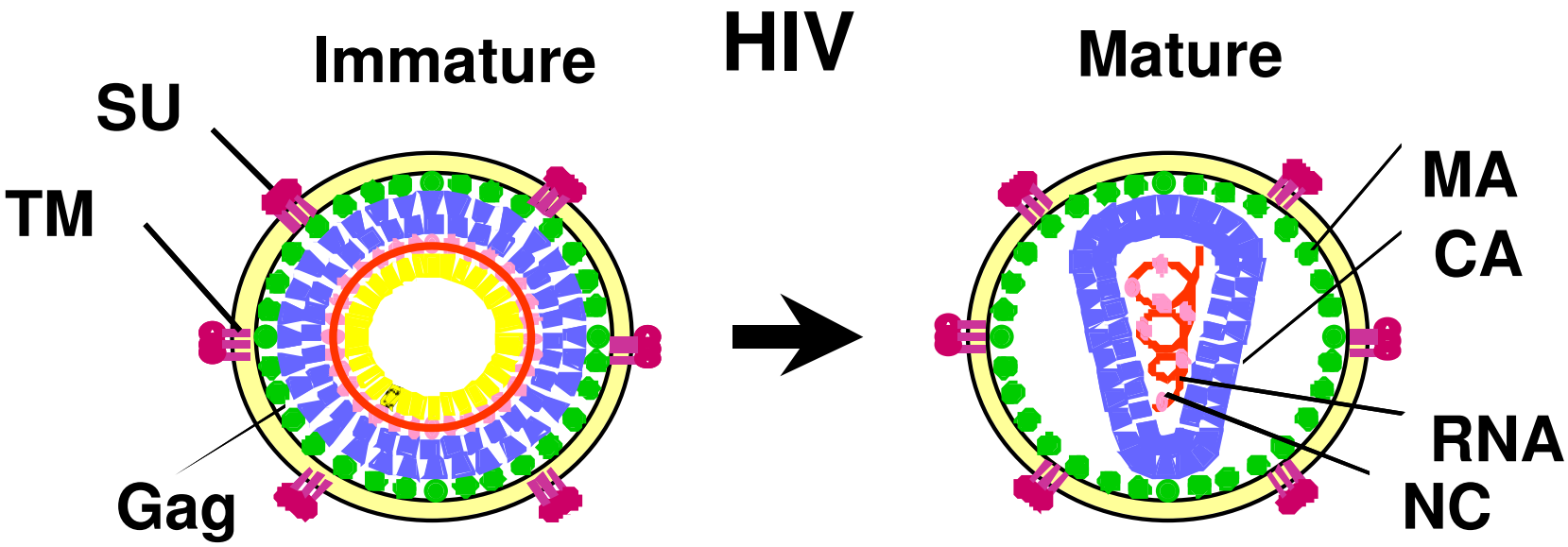
*Specificity of MRM Transitions – Fragment Ion Coelution*

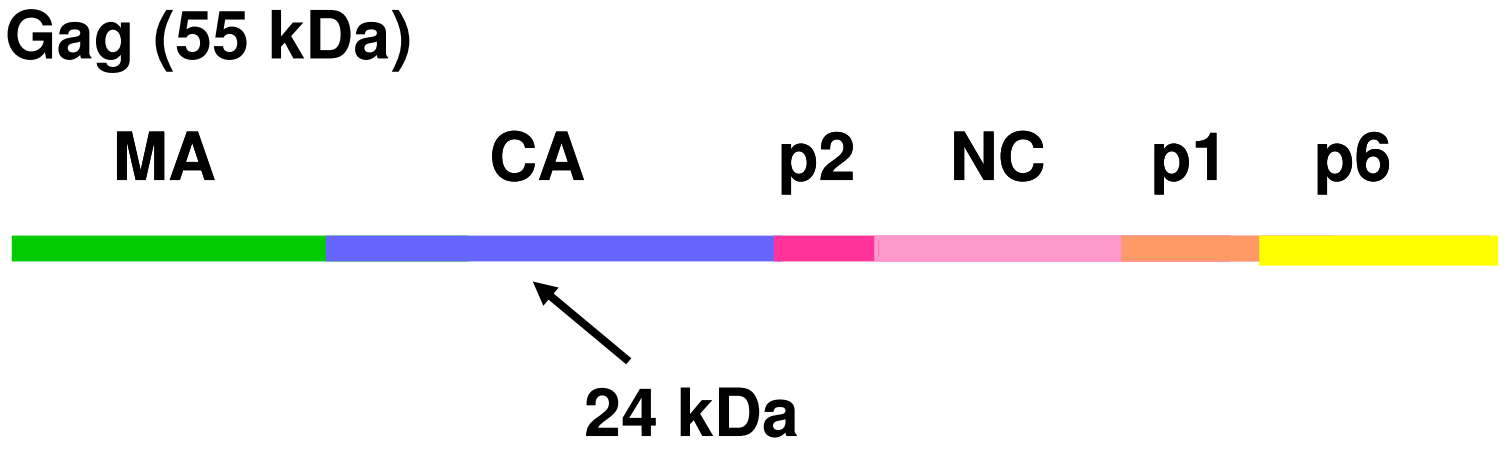



PNAS2007v104p5860

# Hydrogen/Deuterium Exchange

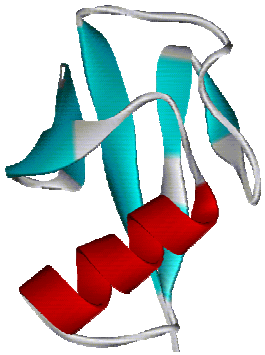
*Study of High Order Protein Structures*



# Hydrogen/Deuterium Exchange

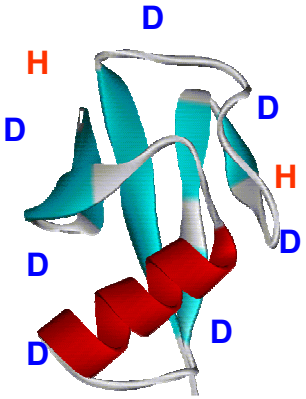
*Study of High Order Protein Structures*



Protein or Complex

Dilute 10 fold

w/ $D_2O$ buffer




H/D exchange

Quench

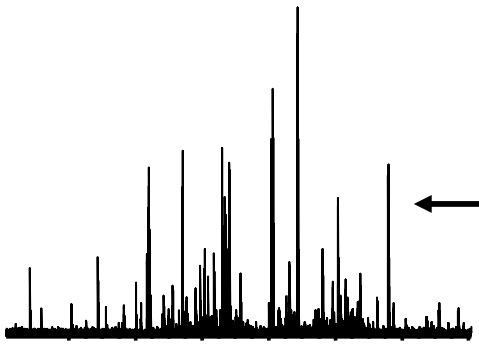
pD 2.5 ~ 3.0

Freeze in liquid $N_2$ & store at -70 °C

Thaw on ice

Digest with pepsin in ice water for 3 min

Online LC/FT-ICR MS



Splitter

to waste

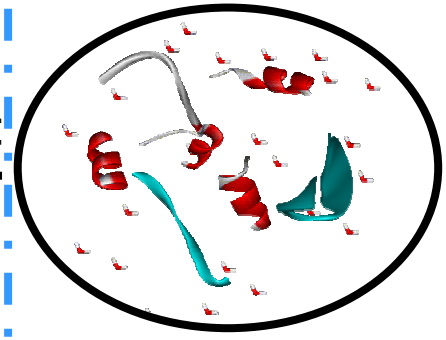
Ice water

Micro-bore c8 column

30 μL/min

Inject

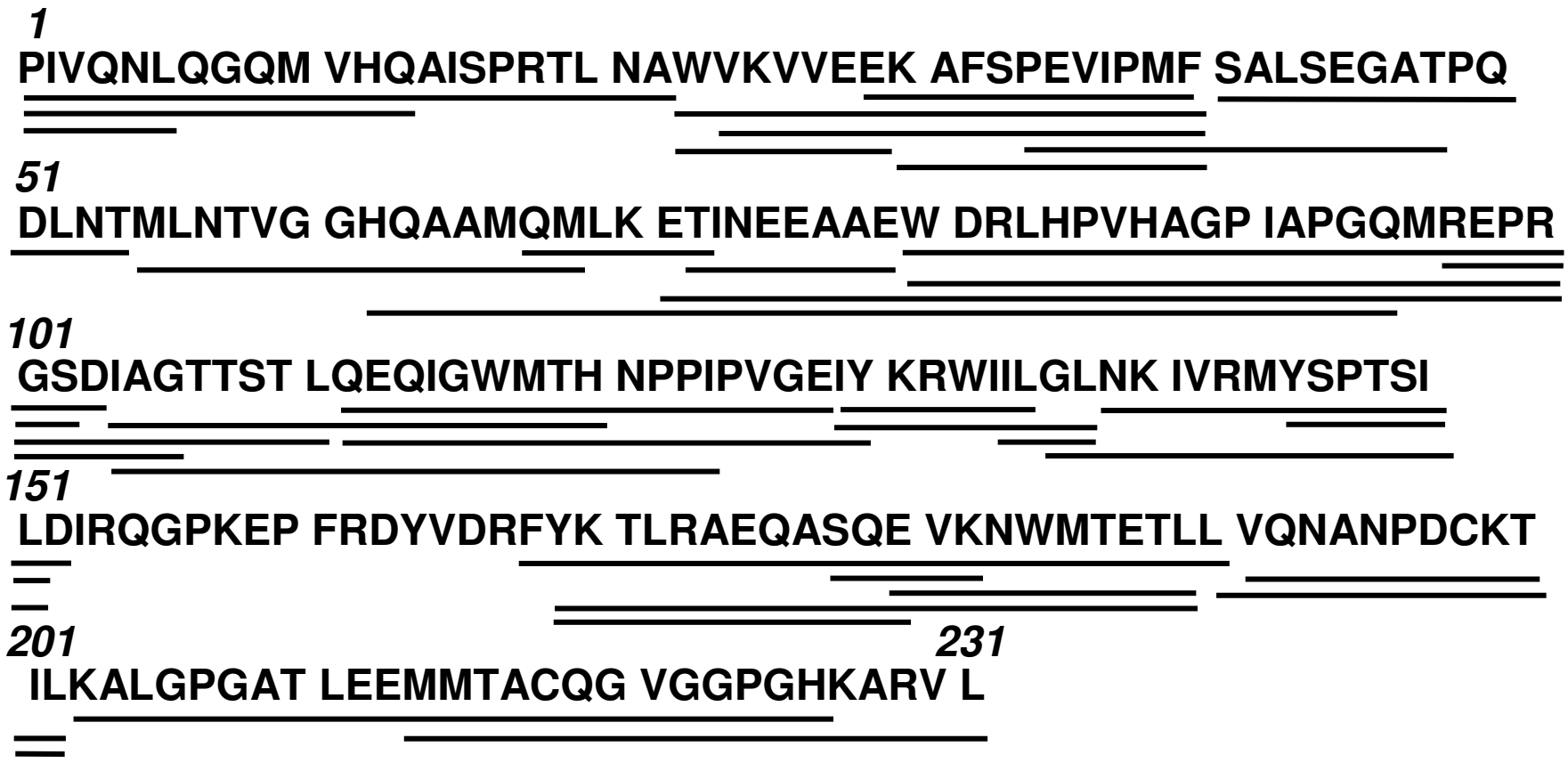
Ice water



# Hydrogen/Deuterium Exchange

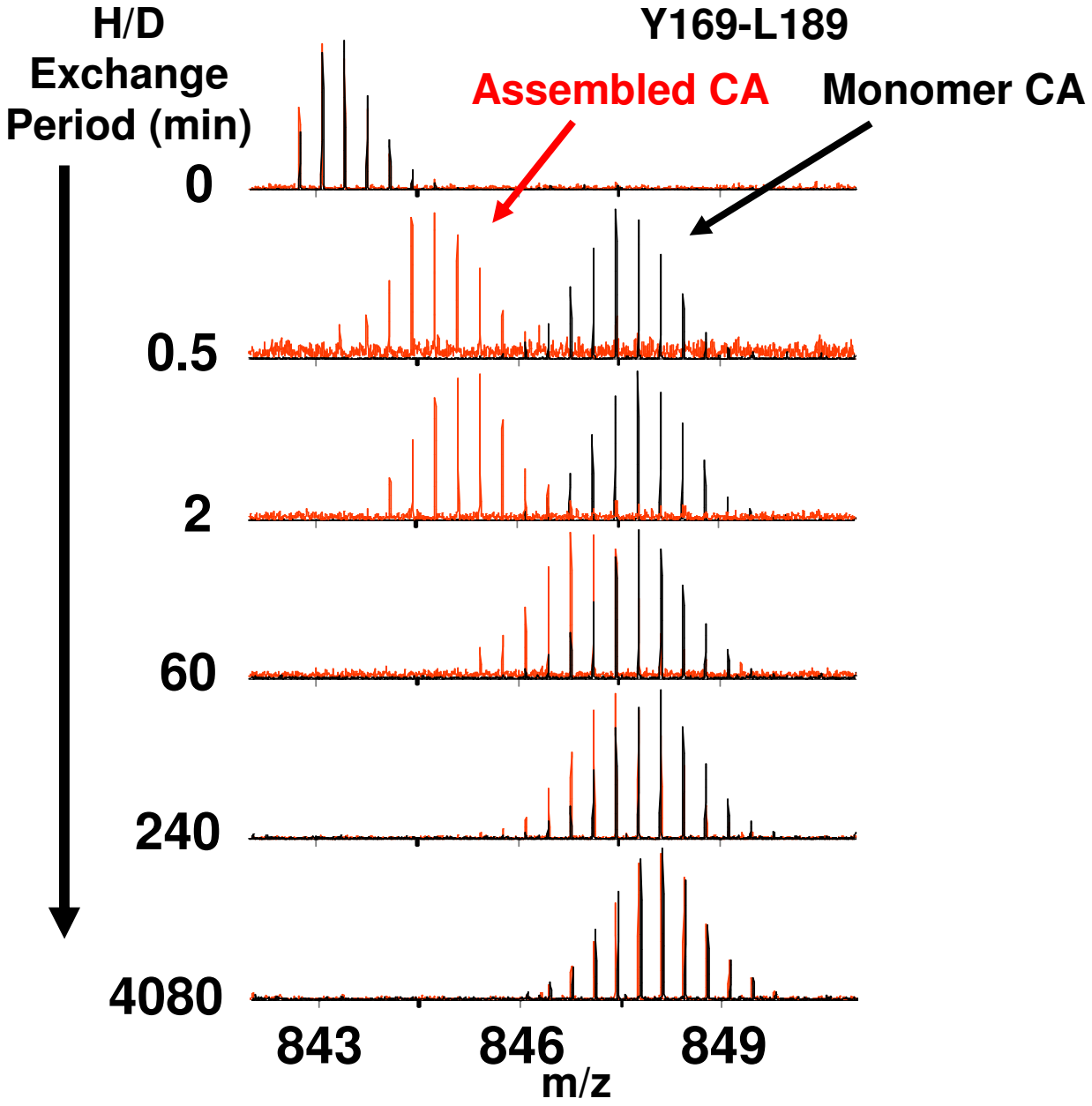
*Study of High Order Protein Structures*

## Peptic Peptides Cover 95% of CA Primary AA Sequence



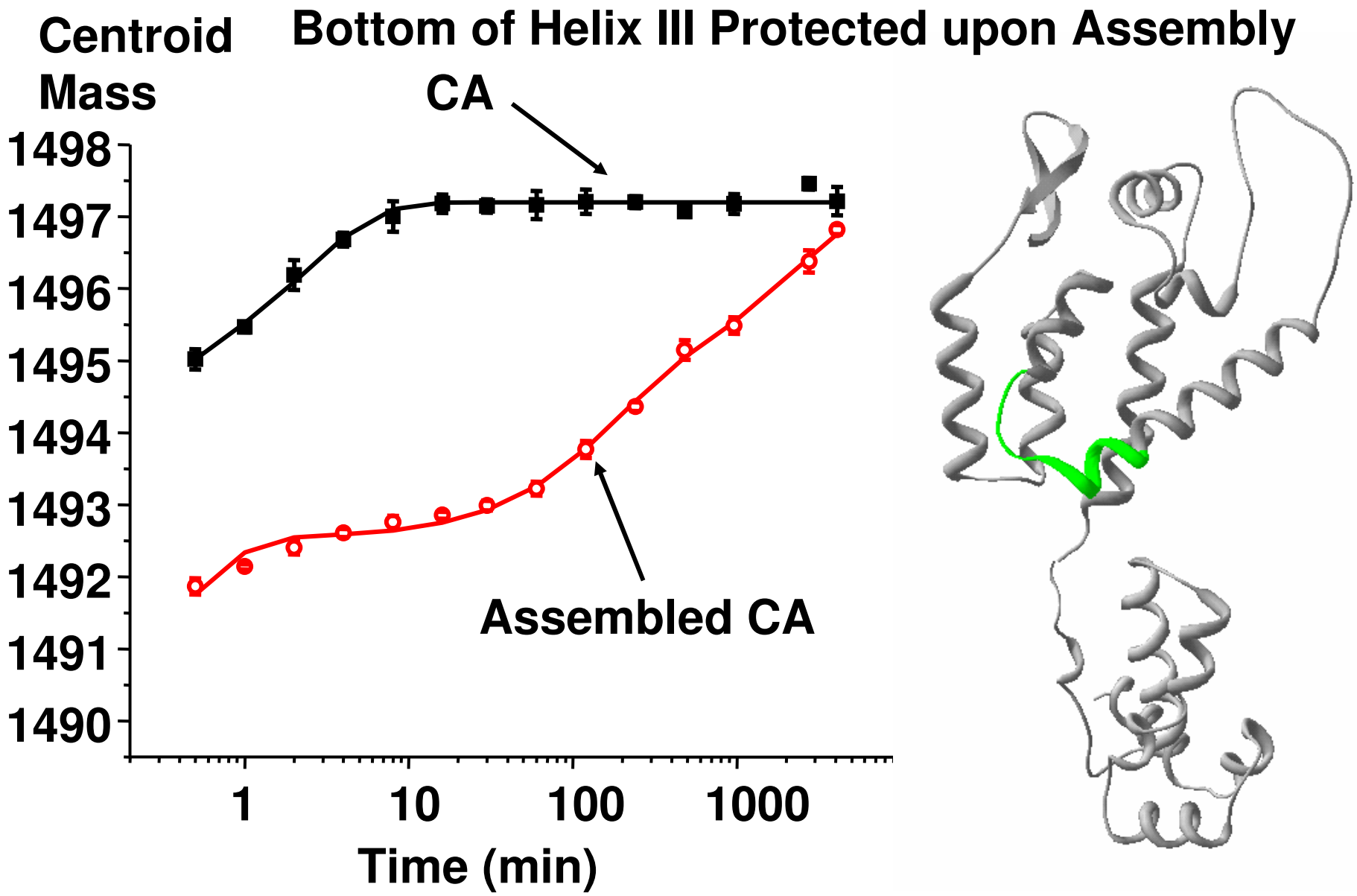
# Hydrogen/Deuterium Exchange

*Study of High Order Protein Structures*



# Hydrogen/Deuterium Exchange

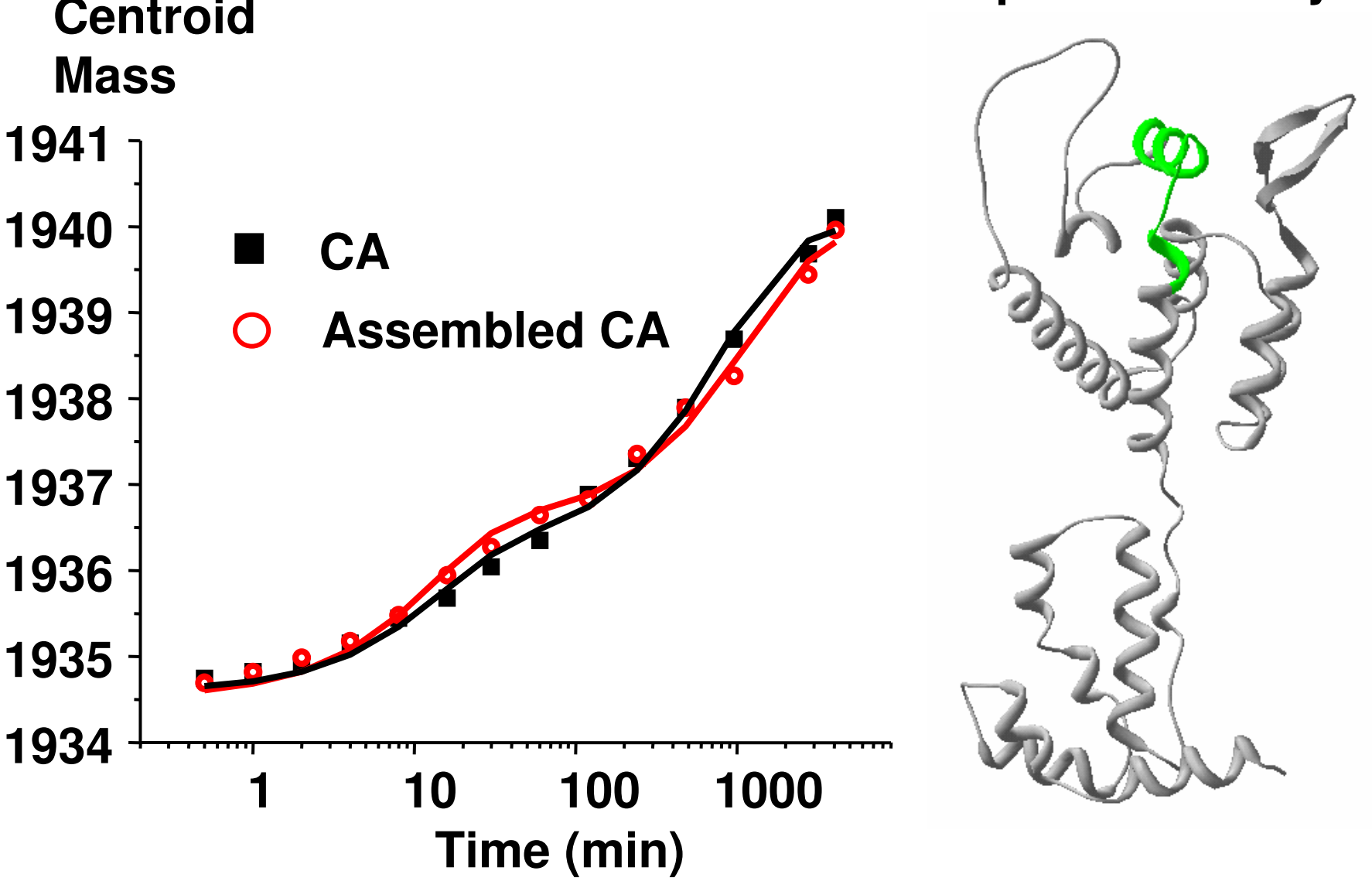
*Study of High Order Protein Structures*



# Hydrogen/Deuterium Exchange

*Study of High Order Protein Structures*

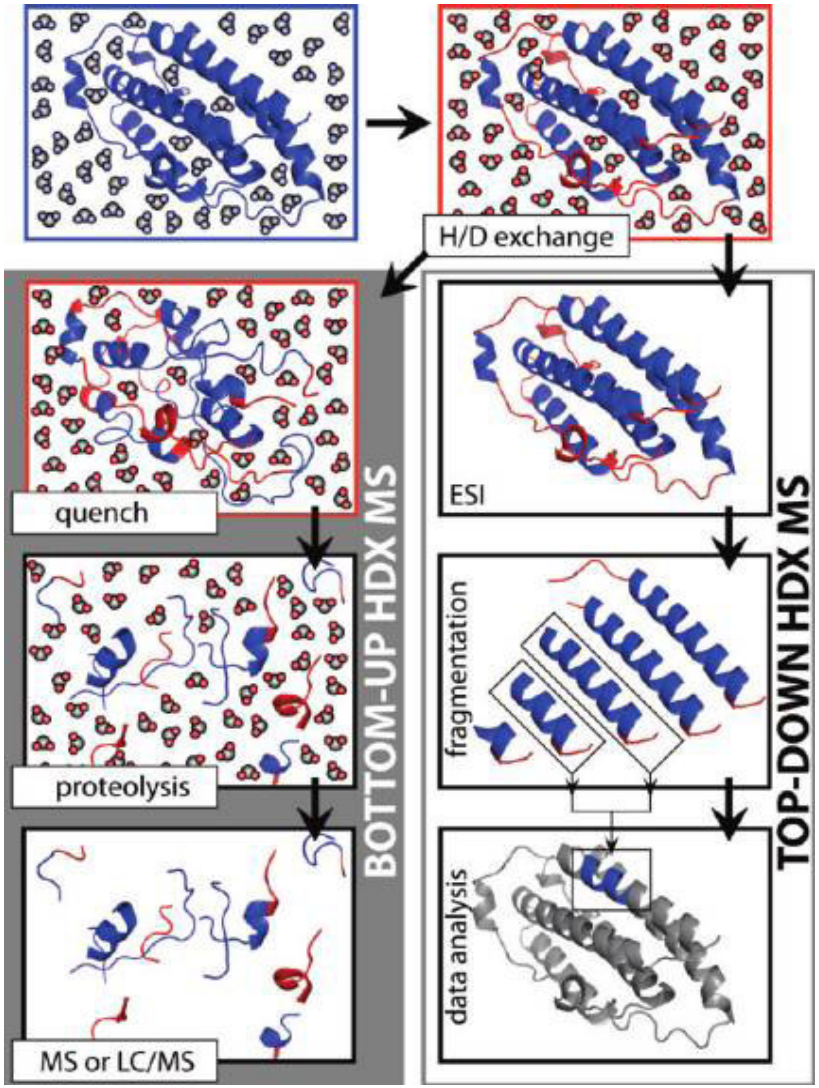
## Helices VI and VII Not Protected upon Assembly



# Hydrogen/Deuterium Exchange

*Study of High Order Protein Structures*

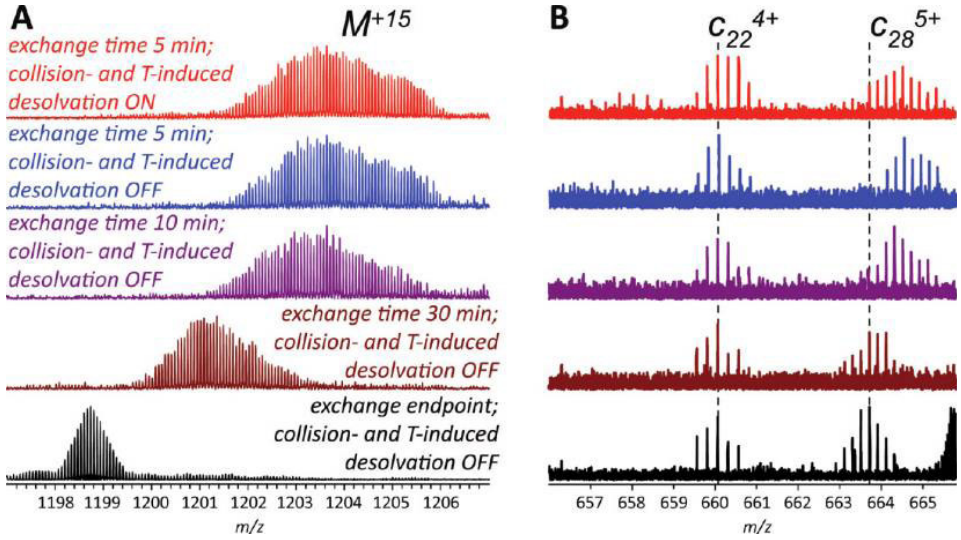
H/D by Topdown Analysis




Fully deuterated protein

+

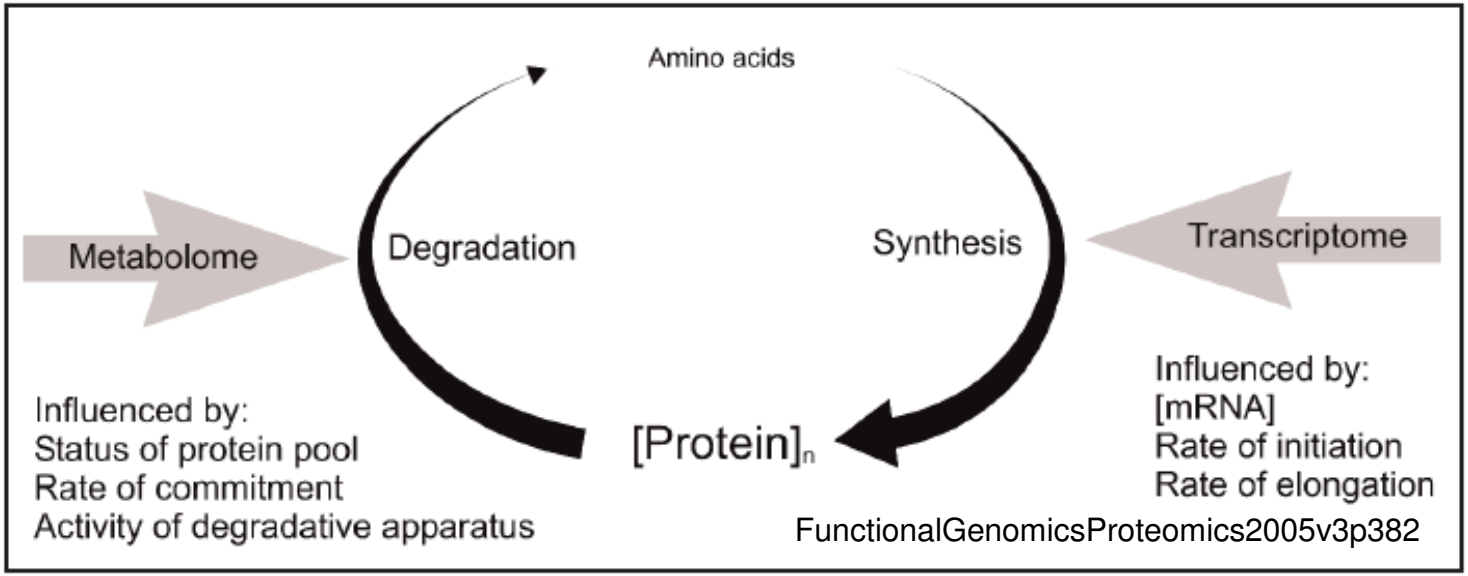
$^{1}H_2O/CH_3CO_2N^{1}H_4$ at pH 3.5




AnalChem2009v81p7892

# Dynamic Proteomics

*Study Protein Turnover on A Proteomic Scale*

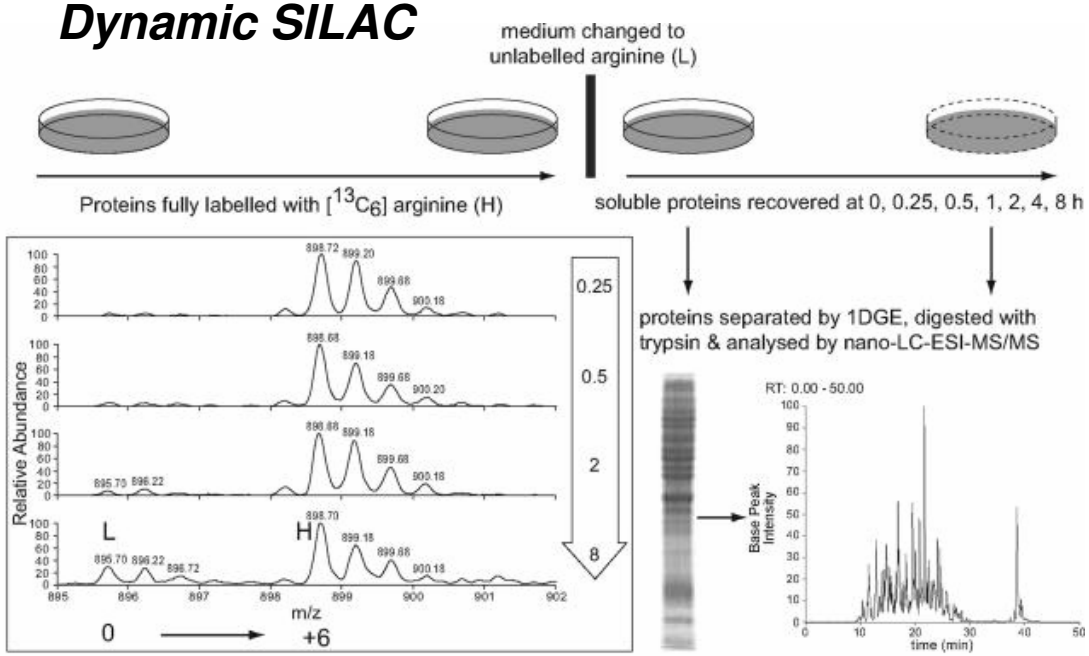



FunctionalGenomicsProteomics2005v3p382

# Dynamic Proteomics

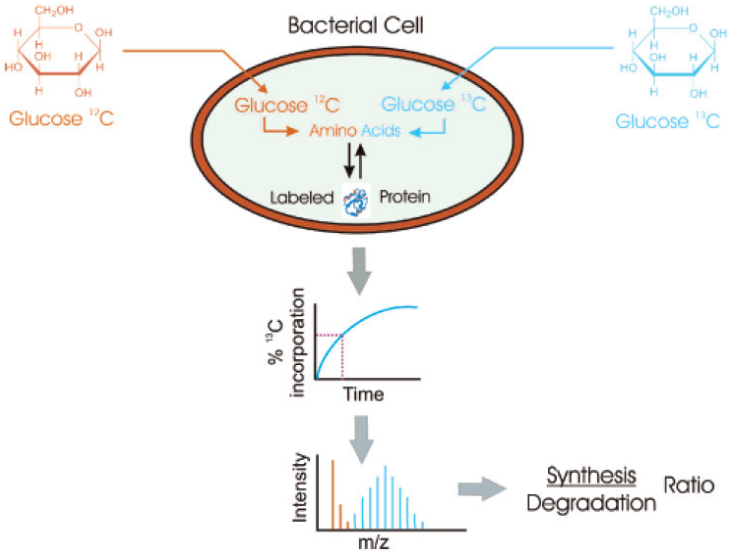
*Experimental Strategies*

## *Dynamic SILAC*




JProteomeRes2009v8p104

## *Synthesis/Degradation Ratio Mass Spectrometry*

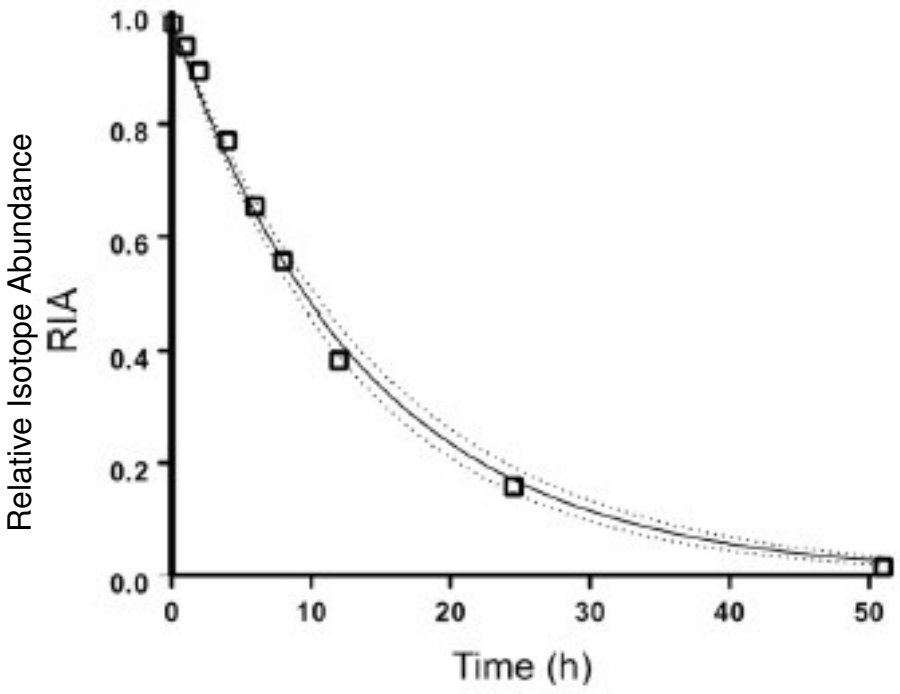


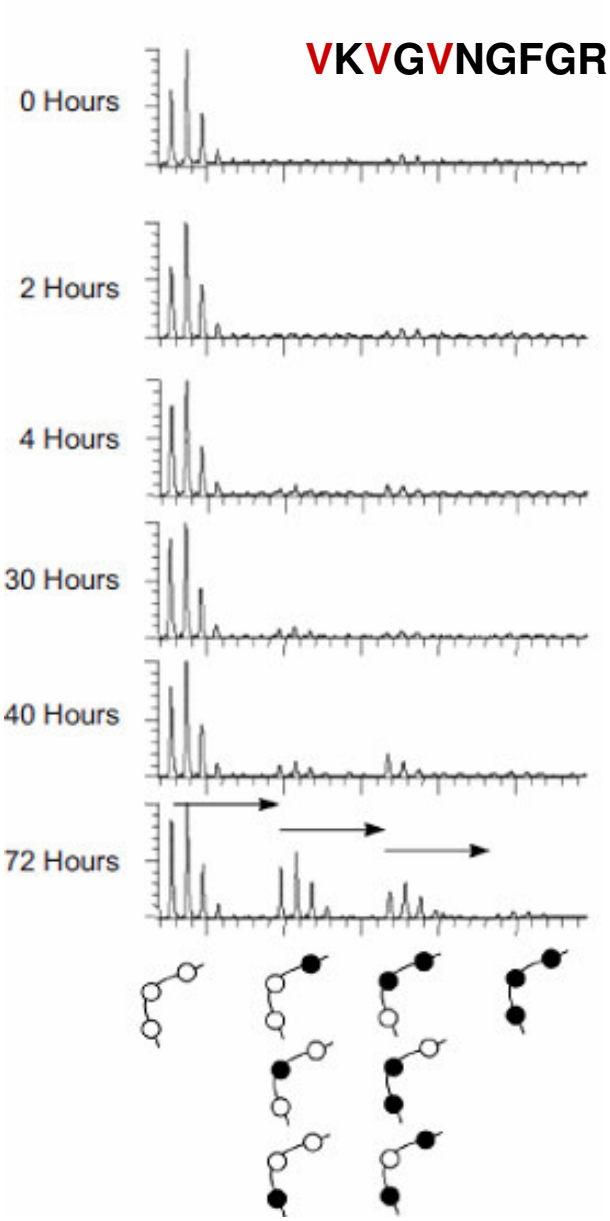

AnalChem2004v76p86

# Dynamic Proteomics

*Dynamic SILAC*

**Chicken fed with stable isotope labeled valine**

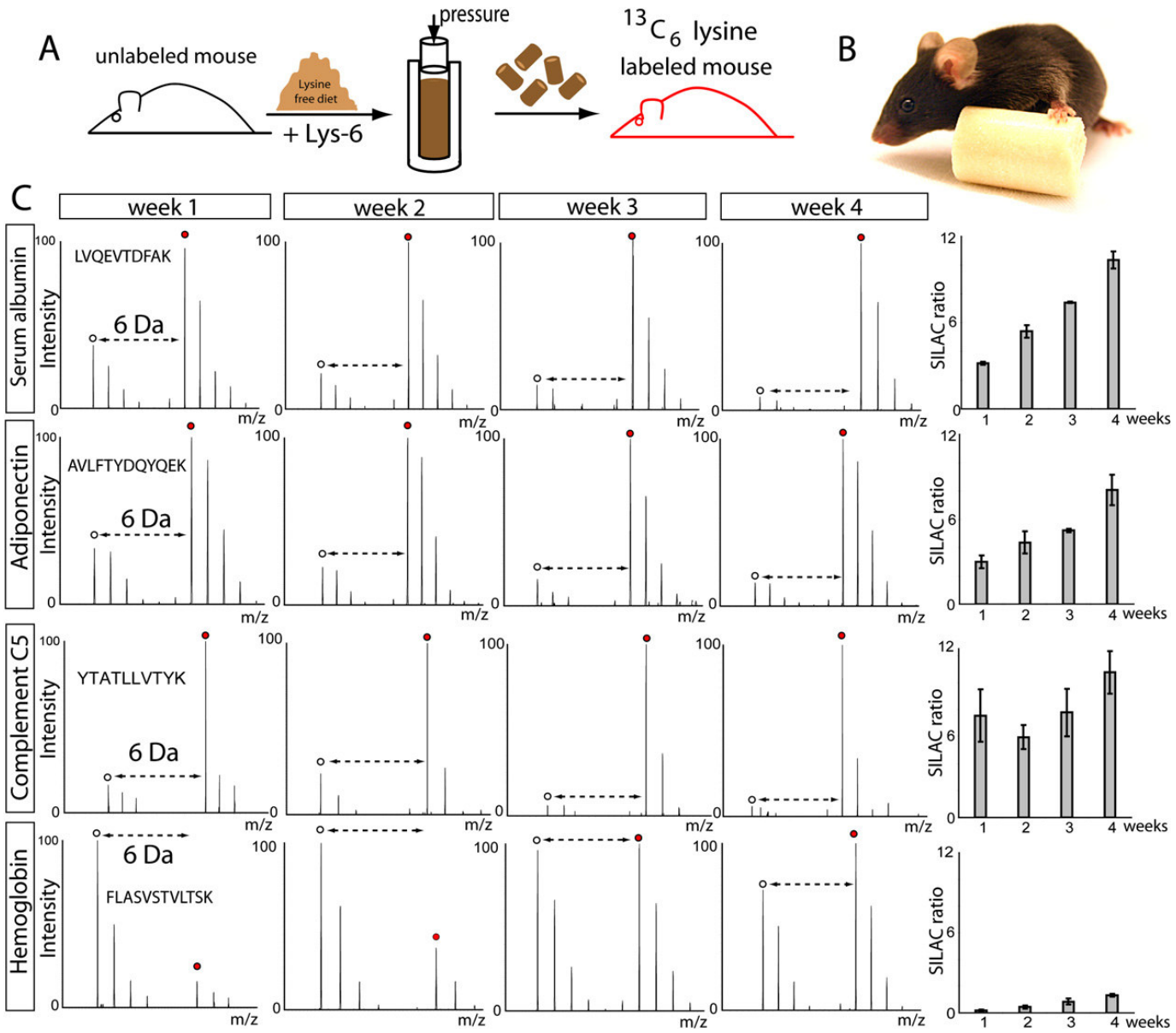



FunctionalGenomicsProteomics2005v3p382

# Dynamic Proteomics

## *SILAC Mouse*

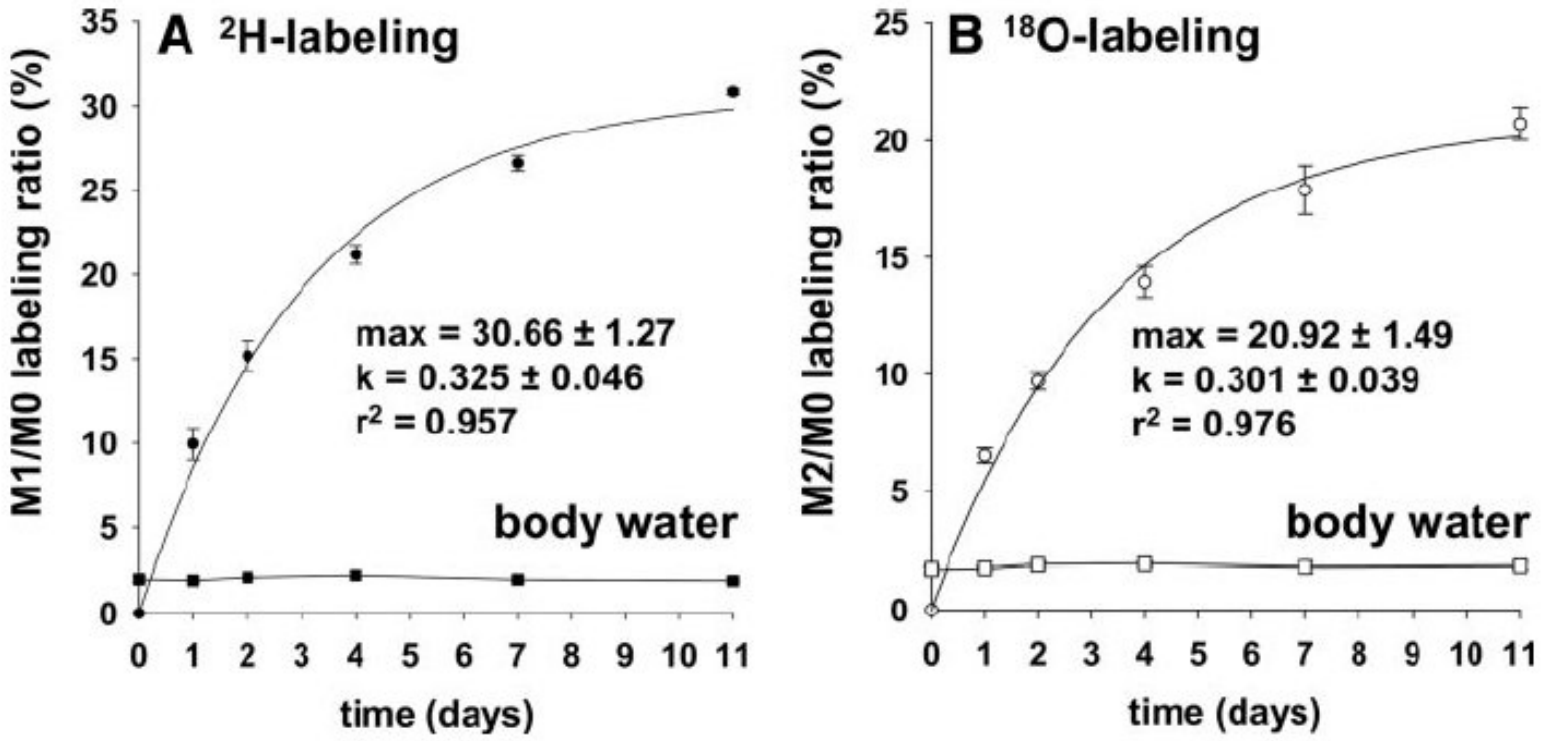



Cell2008v134p353

# Dynamic Proteomics

*Water Labeling*

incorporation of $^{2}H$ or $^{18}O$ into plasma albumin
by bolus injection




MCP2009v8p2653

# Conclusions

- Label free quantification was originated and is still driven by biomarker discovery
- Absolute quantification is possible with spiked-in or QconCAT calibrant peptides
- SILAC allows mixing at protein level
- iTRAQ has been widely used for studying signaling pathways
- Dynamics of a proteome can be probed by metabolic labeling