

BYC 744 Proteomics-Mass Spectrometry

Qualitative and quantitative burrowing of the proteome

Stephen Barnes, PhD
sbarnes@uab.edu

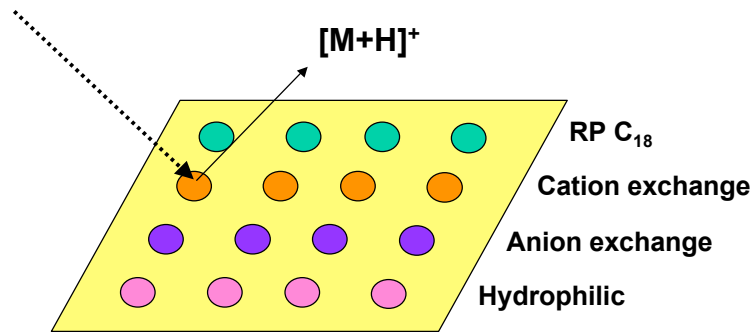
2/10/09

Topics

- **Qualitative changes in the predicted proteome**
 - Digging deeper - SELDI/MALDI
 - Posttranslational modifications
 - Isolation and characterization
 - Making use of the chemistry of the modification
- **Quantitative aspects**
 - Isotope labeling
 - ICAT, $^{18}\text{O}/^{16}\text{O}$, $^{15}\text{N}/^{14}\text{N}$
 - Chemical labeling
 - Absolute measures

2/10/09

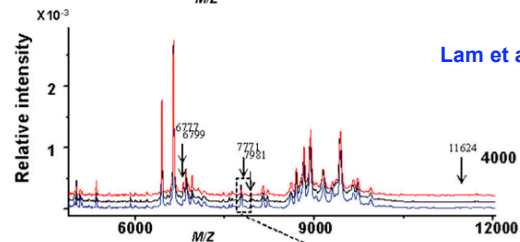
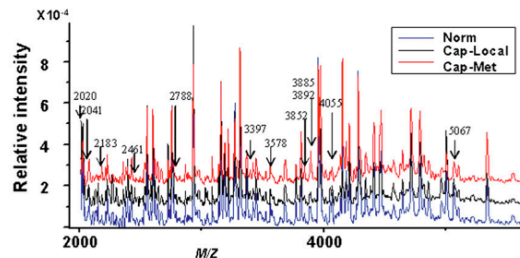
Surface enhanced laser desorption ionization (SELDI)



SELDI chip (plate) with local chemistry on surface causing selective binding of peptides

2/10/09

MALDI-TOF analysis of peptides recovered by C₁₈ extraction of sera from prostate patients, with and without metastases, and controls

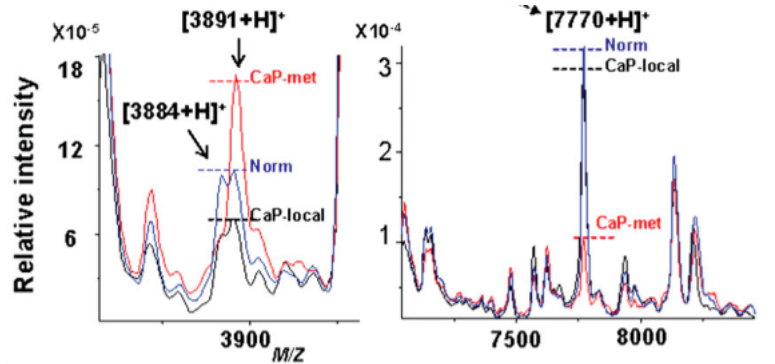


Lam et al., Proteomics 5, 2927

2/10/09

Identification of a significantly altered peptide in men with metastatic prostate cancer

Lam et al., Proteomics 5, 2927

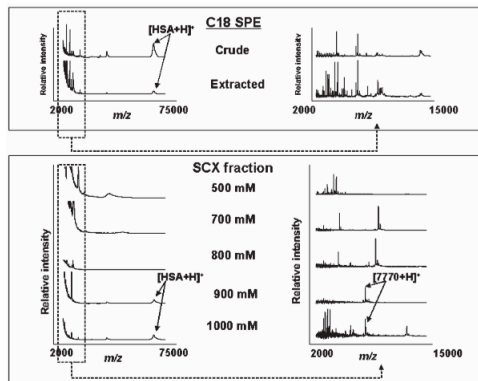


But what is it?

2/10/09

Purification of prostate peptide by SPE extraction and strong cation exchange

Lam et al., Proteomics 5, 2927



Serum was first extracted with a C₁₈ SPE cartridge

SPE fraction was passed over a [H⁺]-form cation exchange resin. Bound peptides were eluted with increasing step gradients of NaCl

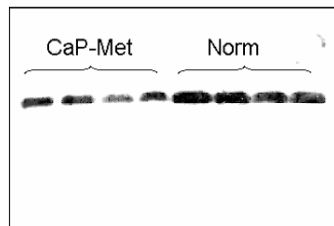
Analyses of fractions were performed by MALDI-TOF MS

The most enriched fraction was contaminated with albumin

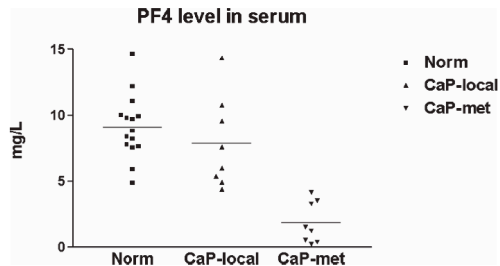
2/10/09

Non-mass spec verification of prostate peptide

Lam et al., Proteomics 5, 2927



Western blot analysis for PF4



ELISA analysis for PF4

Moral: proteomics is a serious business that requires multiple dimensions of separation - glib methods don't work

2/10/09

General classes of modification

- Biochemical events involving peptide processing
- Biochemical events stimulated by enzymes
- Chemical events driven by reactive species
- Chemical events determined by investigator

2/10/09

Examples of peptide processing of polypeptides

- Head groups of membrane proteins
- The family of proteins in HIV that are manufactured as one protein and then hydrolyzed by a protease
- Chromogranin - a brain protein consisting of several bioactive peptides
- Formation of β -amyloid

2/10/09

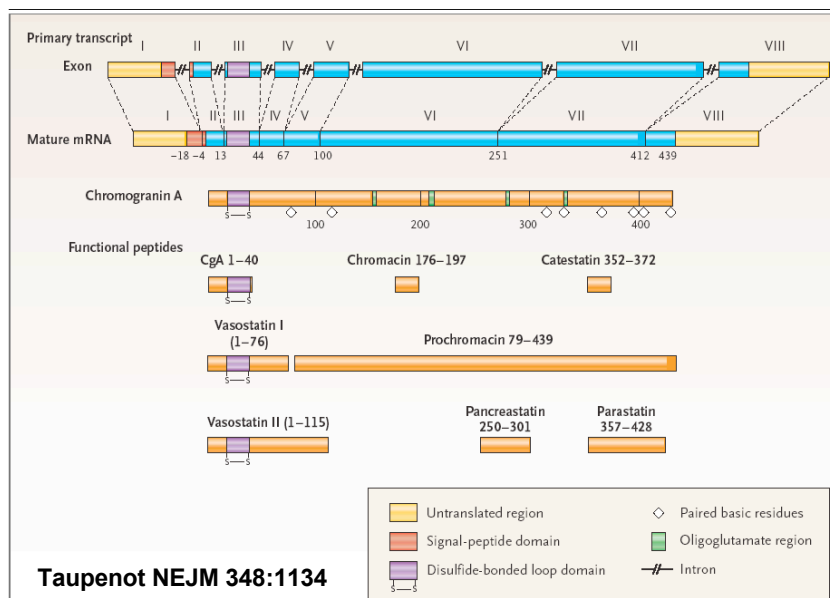


Figure 3. Peptide-Encoding Regions and Putative Functional Domains of Human Chromogranin A (CgA).

Arabic numbers designate amino acids in the mature protein (minus signal peptide). Roman numerals designate exon numbers. The intron-exon structure is not drawn to scale.

Enzymatic modifications

- **Phosphorylation/dephosphorylation**
 - On serine, threonine, tyrosine
- **Glycosylation**
 - N-glycosylation (asparagine-linked)
 - O-glycosylation (serine-, threonine-linked)
- **N-Acetylation/deacetylation**
 - On lysines
- **N-Methylation/demethylation**
 - On lysines - mono-, di- and trimethylation

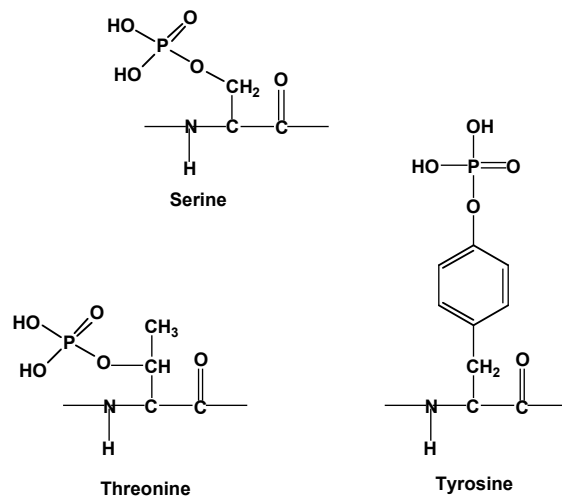
2/10/09

Phosphorylation of proteins

- In some cases, proteins are normally found in a stable, hyperphosphorylated state, e.g., casein
- In many cases, it is a transient event that causes 10-100 fold increase in enzyme activity. This is the way signals are propagated through a signal transduction pathway. However, the molar abundance of phosphorylation at an individual site may only be 1-2%.

2/10/09

Chemistry of phosphorylation



2/10/09

Mass spectrometry of phosphorylated proteins

- Adds H₃PO₄ (+98)
- Eliminates water (-18)
- Net change +80
- if the phosphate ion is released (i.e., in negative ion spectra), it is seen as *m/z* 79 and/or *m/z* 63
- Phosphate can be confused with sulfate and bromide (sulfate, 79.9568 Da; phosphate, 79.9663 Da; and ⁸¹Br, 79.9083 Da)

2/10/09

Limitations of mass spectrometry

- Although it can deliver sensitivity in the low fmol range (similar to immunological methods), because it is a universal detection method, finding the needle in the “haystack” of all the other peptides is a challenge
- Recovering the phosphopeptides from the matrix of the sample is more important than the mass spectrometry measurement

2/10/09

Finding a phosphate group

Several methods are in current use for detection of phosphopeptides

- use of parent ion or neutral loss scanning
- phosphatase sensitivity
- affinity methods for enrichment of phosphopeptides
 - anti-phospho-Ser/Thr/Tyr antibodies
 - metal ion affinity
 - chemical reaction/biotin affinity

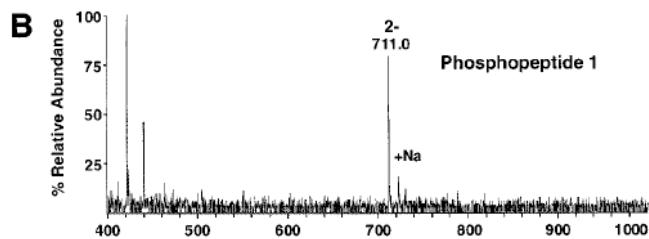
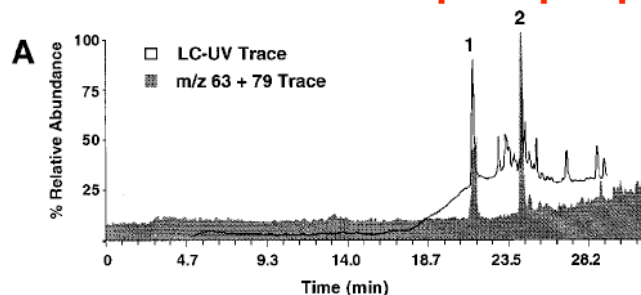
2/10/09

Parent ion scanning to detect phosphopeptides

- The procedure depends on the detection of the m/z 79 ion fragment (PO_3^-) during collision-induced dissociation in a triple quadrupole instrument operating in the negative ion mode
- *Parent ion scanning* is a reversal of the more familiar daughter ion MS-MS where the parent ion is selected (in Q1) and a mass spectrum of the daughter ion fragments is obtained by scanning in Q3
- In *parent ion scanning*, the daughter ion fragment (in this case m/z 79) is held constant in Q3 and a mass spectrum of parent ions that give rise to the daughter ion obtained by scanning in Q1.
- Having identified the phosphopeptides, the sample can be re-analyzed to obtain daughter ion MS-MS spectra on selected ions in the positive ion mode

2/10/09

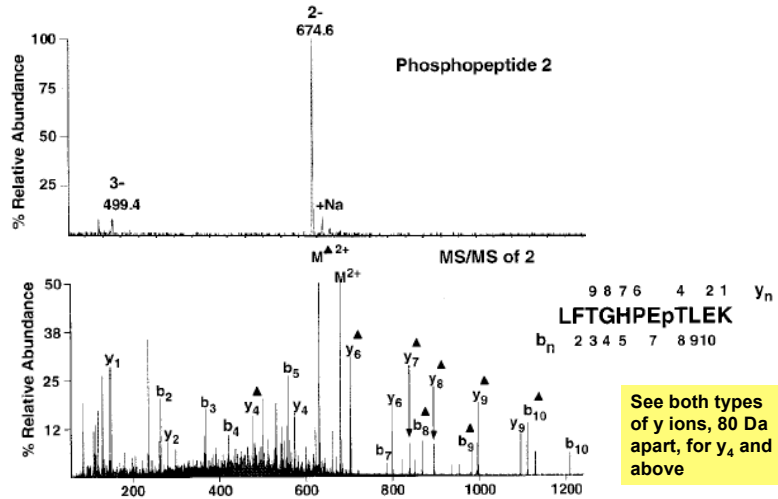
Parent ion scan to detect phosphopeptide



2/10/09

Annan Anal Chem 73:393

Phosphopeptide 2



2/10/09

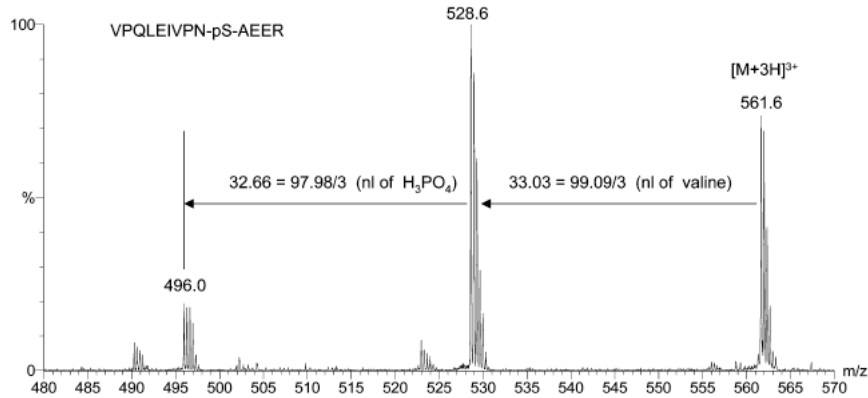
Annan Anal Chem 73:393

Neutral loss for phosphopeptides

- Based on the loss of phosphoric acid (H_3PO_4) - 97.98 Da
- If the peptide is doubly charged, then it's m/z 48.88
- If it's triply charged, then it's m/z 32.66
- Problems can occur if the peptide contains an N-terminal proline (97 Da) or valine (99 Da) and a low resolution instrument is used for the analysis

2/10/09

Casein peptide - mono- or diphosphate?



Lehmann et al., J Prot Res 7-2866

2/10/09

Problems of low mass resolution for parent ion scanning for phosphate

Lehmann et al., J Prot Res 7-2866

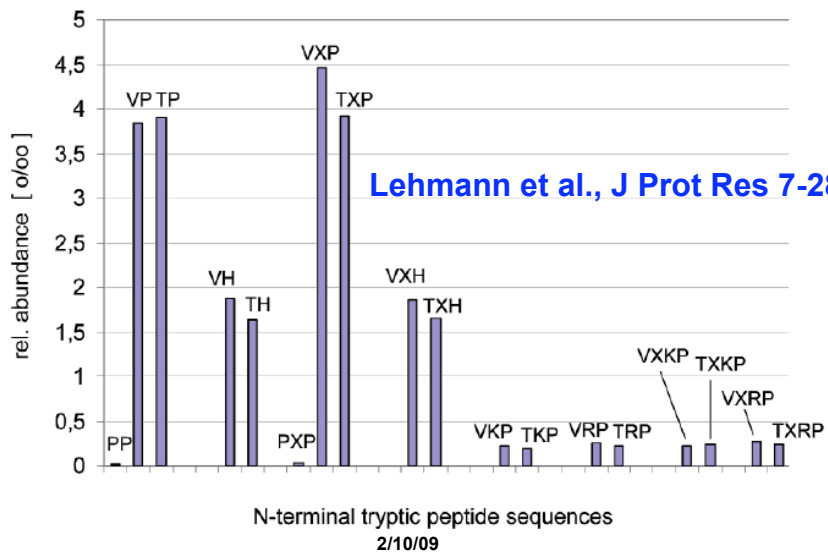
Table 1. Peptide Modifications Showing an Abundant 'Close-to-98/z' Neutral Loss Compared to the Specific Loss of H₃PO₄ from pSer/pThr Phosphopeptides^a

structure	neutral loss	1+	2+	3+	4+
phosphoS/phosphoT	H ₃ PO ₄	97.9769	48.9884	32.6590	24.4942
P -P... (N-term)	Pro	97.0528	48.5264	32.3509	24.2632
C-sulfo	SO ₃ + H ₂ O	97.9674	48.9837	32.6558	24.4918
V -P... (N-term)	Val	99.0684	49.5342	33.0228	24.7671
T -P... (N-term)	Thr	101.0477	50.5238	33.6826	25.2619
...-dhBA (C-term)	dhBA	101.0477	50.5238	33.6862	25.2619
C -P... (N-term)	Cys	103.0092	51.5046	34.3364	25.7523
...- S (C-term)	Ser + H ₂ O	105.0426	52.5213	35.0142	26.2606
M-acetamido	MTA	105.0248	52.5124	35.0083	26.2562

^a Neutral loss *m/z* values for the charge states +1 to +4 are listed (dhBA, dehydrobutyric acid; MTA, 2-(methylthio)acetamide; all amino acid symbols refer to the amino acid mass minus water).

2/10/09

Residues in the proteome with phosphate-like neutral loss



Recovery and enhancement of phosphopeptides

The biggest problem in the detection of phosphopeptides is how to convert the initial sample matrix into a form suitable for mass spectrometry analysis.

- how to handle minute samples with minimal losses
- how to recover and detect all the phosphopeptides
- how to recover and detect the non-phosphorylated proteins to determine the extent of phosphorylation at individual sites

2/10/09

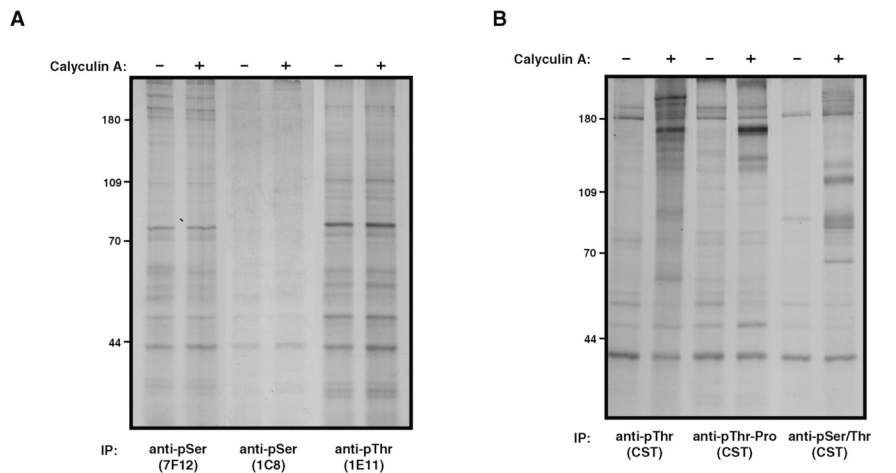
Antibodies and phosphopeptides

- In this approach, both the phosphorylated and non-phosphorylated forms of a protein may be recovered from the sample matrix
- This can most easily be achieved by immunoprecipitation of the protein with an antibody that recognizes epitope(s) that is(are) in common with both forms

NOTE *that as with all immunoprecipitation methods, the best results will be achieved if the antibody is coupled to agarose beads. This allows selective immuno-absorption of the antigen, washing free of contaminating proteins, AND following elution, minimization of the amount of antibody that is in the eluate and therefore would be analyzed by mass spectrometry*

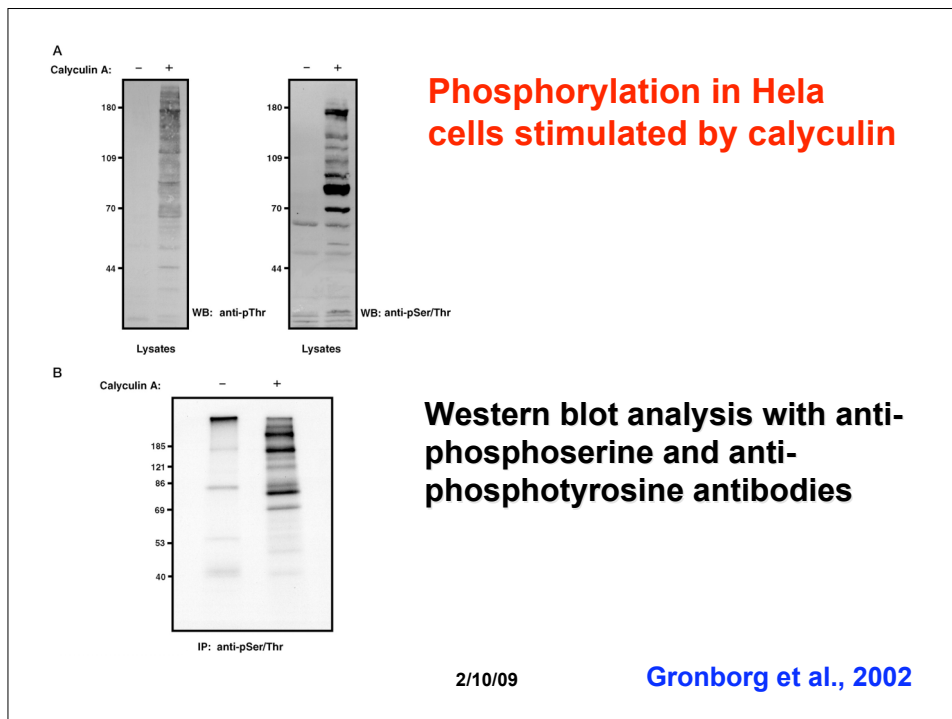
2/10/09

Variability of anti-phosphoserine and anti-phosphothreonine antibodies



2/10/09

Gronborg et al., 2002



Detection of phosphopeptides based on their sensitivity to phosphatase

- An alternative source of potentially phosphorylated proteins are individual spots on 2D-IEF/SDS gels. The protein preparation so isolated is either hydrolyzed by trypsin in solution (or in the gel piece) or using solid-phase trypsin
- One portion of the resulting tryptic peptides (in 50% acetonitrile:water) is analyzed by MALDI-TOF-MS. A second portion is diluted into 50 mM NH_4HCO_3 buffer and reacted with 0.5 U calf intestinal alkaline phosphatase at 37°C for 30 min. Sample is dried with a SpeedVac, redissolved in 50% acetonitrile:water, and re-analyzed by MALDI-TOF-MS

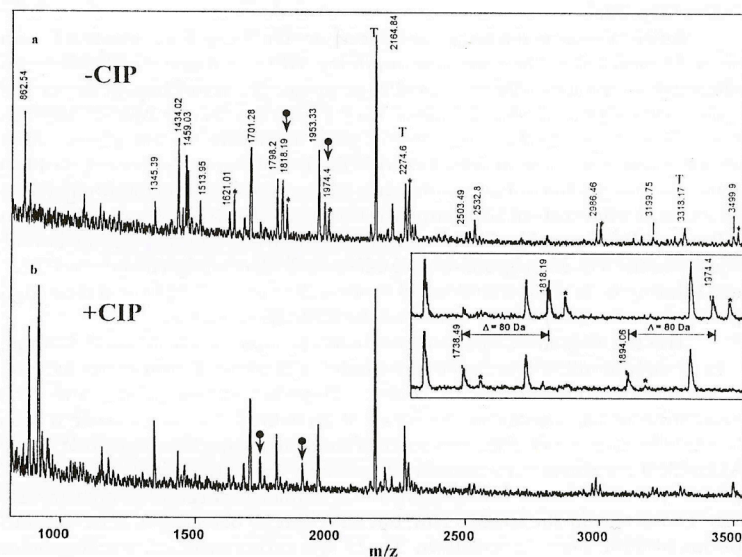
2/10/09

Identifying phosphopeptides using alkaline phosphatase

- Phosphopeptides shift down by m/z 80 (or units of 80 in the case of multiply phosphorylated peptides)
- The peaks identified as phosphopeptides can then be analyzed in a nanoelectrospray experiment where collision-induced dissociation is used to determine the identity of the peptide and the phosphorylation site in the sequence

2/10/09

Detecting a phosphopeptide with alkaline phosphatase



2/10/09

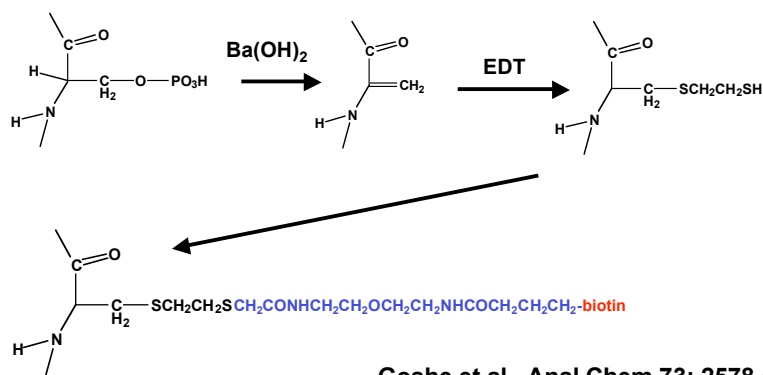
Zhang et al., 2000
MS in Biology & Medicine

Selective enhancement of phosphopeptides in tryptic digests

- Immobilized metal affinity chromatography (IMAC). Similar to Ni-affinity resins used in the purification of 6xHis-tagged proteins. The affinity phase can be charged with different metal ions (as their chlorides)
- Fe(III) and Ga(III), and to a lesser extent Zr(IV), were the most effective for the recovery of two synthetic phosphopeptides
- A tryptic digest containing both phosphorylated and non-phosphorylated peptides is passed over the IMAC column at acid pH (pH 2.5-3)
- The column is washed with 0.1 M acetic acid to remove unbound peptides
- Elute with sodium phosphate (have to desalt) or with NH_4OH
- Esterification may prevent Asp- or Glu-containing peptides from binding

2/10/09

Selective biotinylation of phospho-groups



Goshe et al., Anal Chem 73: 2578 (2001)
Adamcyk et al., Rapid Commun Mass Spec 15: 1481 (2001)

2/10/09

How to identify phosphorylated peaks by searching databases

- The databases you have used earlier in this class have some ability to predict the expected masses for a limited number of posttranslational modifications
 - MASCOT (<http://www.matrixscience.com>)
 - PROTEIN PROSPECTOR (<http://prospector.ucsf.edu/>)

2/10/09

How to identify posttranslational modifications

FindMod at

<http://www.expasy.org/tools/findmod/>

It examines mass fingerprinting data for mass differences between empirical and theoretical peptides. If the mass difference corresponds to a known modification, it also makes intelligent guesses as to the site of modification.

2/10/09

FindMod

Some of the modifications considered are:

acetylation	amidation	biotinylation
C-mannosylation	deamidation	flavinilylation
farnesylation	formylation	geranyl-geranylation
γ -carboxyglutamic acid	hydroxylation	lipoylation
methylation	myristoylation	N-acyl diglyceride
O-GlcNac	palmitoylation	phosphorylation
pyridoxal phosphate	phospho-pantetheine	pyrrolidone-carboxylic acid
sulfation		

NOTE that none of the common chemical modifications (alkylation of sulfhydryl groups with iodoacetic acid, iodoacetamide, 4-vinylpyridine, and acrylamide) were included.

Wilkins M.R., Gasteiger E., Gooley A., Herbert B., Molloy M.P., Binz P.A., Ou K., Sanchez J.-C., Bairoch A., Williams K.L., Hochstrasser D.F. *High-throughput Mass Spectrometric Discovery of Protein Post-translational Modifications*. [Journal of Molecular Biology](#), 289, p. 645-657 (1999)

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; *Protein Identification and Analysis Tools on the ExPASy Server*; (In) John M. Walker (ed): [The Proteomics Protocols Handbook](#). Humana Press (2005).

2/10/09

Site for compilation of PTMs

<https://www.abrf.org/index.cfm/dm.home>

This site was put together by Ken Mitchelhill, Len Packman and friends

Currently ranges from dephospho (-79) to (Hex)3-HexNAc-(dHex)HexNAc (+1,039)

2/10/09

Detecting PTMs

- If we don't have any idea about what the modification(s) is(are), how do we proceed?
- We won't have the modification on MASCOT or Protein Prospector
- No antibodies or convenient affinity phases

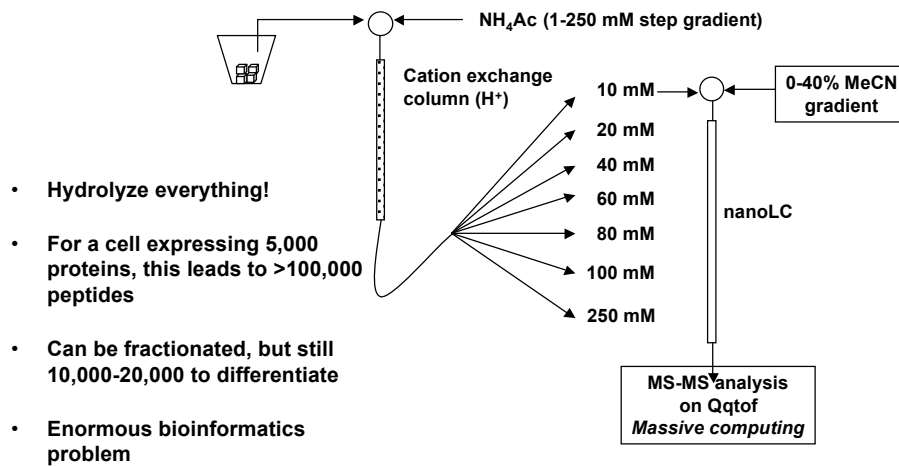
2/10/09

Can we approach this globally?

- It's asking a lot
 - Too many degrees of freedom
 - Endless modifications
- Should we try?
 - John Yates' group has
 - [MacCoss et al. PNAS 99:7900 \(2002\)](#)

2/10/09

MudPIT - Multi-dimensional Protein Identification Technology



2/10/09

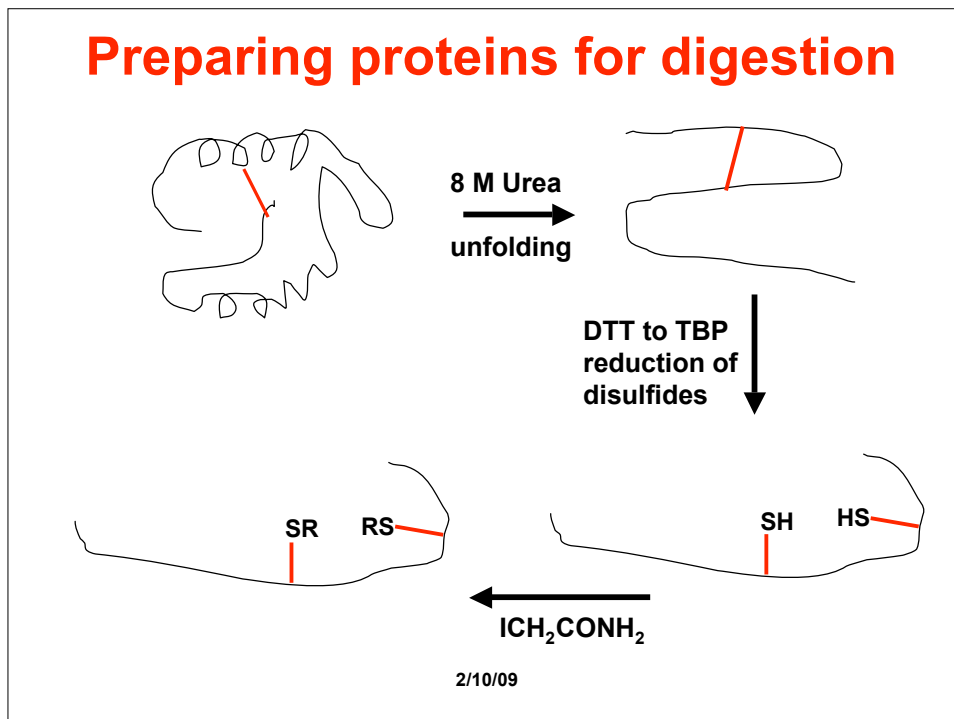
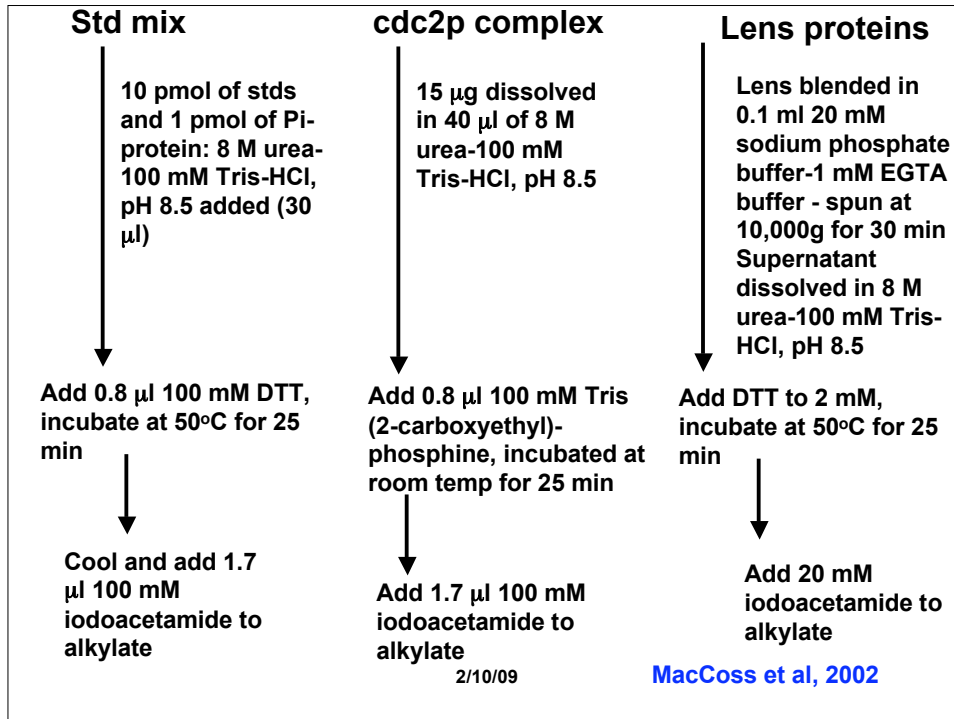
John Yates

The Yates approach - selection of “proteomes”

- BioRad mixed molecular weight standards (10 pmol each) plus 1 pmol of a phosphorylated glycogen phosphorylase
- Cdc2 protein complex isolated with TAP (15 μg)
- Lens proteins from 4-yr old

2/10/09

MacCoss et al, 2002



Hydrolysis procedures

Reduced, carboxymethylated in 8 M urea

Diluted x 3 with
100 mM Tris-HCl,
pH 8.5

100 mM CaCl₂
added to 1 mM

Roche **trypsin** 1:50,
incubated 12-24 h
at 37°C

quenched with
90% formic acid
to 4%

Diluted x 3 with 4.8
M Urea-100 mM Tris-
HCl, pH 8.5

subtilisin added
1:50, incubated 2-3
hr at 37°C

quenched with
90% formic acid
to 4%

2/10/09

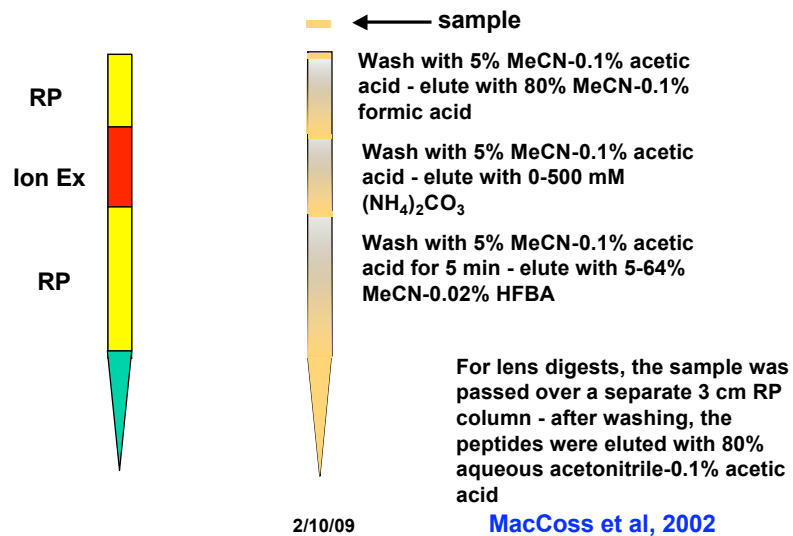
Diluted x 3 with
100 mM Tris-HCl,
pH 8.5

incubated with
elastase 1:50 12 h
at 37°C

quenched with
90% formic acid
to 4%

MacCoss et al, 2002

Elution from a triphasic column

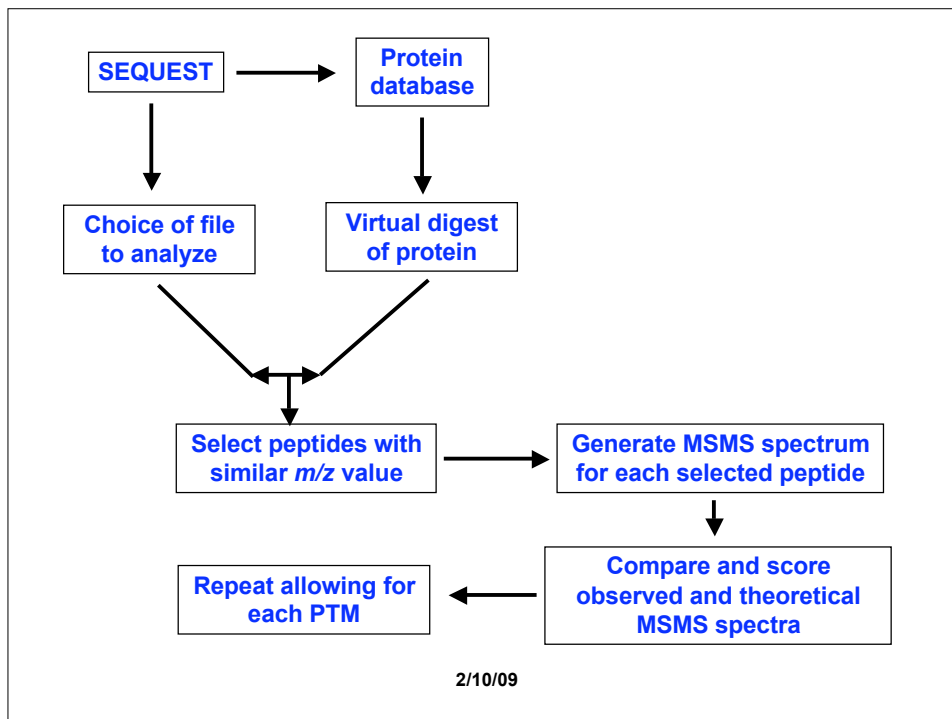


Automated MS-MS analysis

- Limit analysis to 2⁺ or 3⁺-charged peptides
- Delete poor quality spectra
- Identify peptides with 98-Da neutral loss
- Analyze the remaining spectra with SEQUEST (operating on a 31-node beowulf computer cluster) to identify proteins
- Search sequences of identified proteins for PTMs by 80 (STY phosphorylation), 42 (K acetylation), 16 (MWY oxidation) and 14 (K methylation)

2/10/09

MacCoss et al, 2002



2/10/09

Results

Protein standard mixture:

- The three digests were combined - 83.7% and 95.4% coverage for glycogen phosphorylase and BSA
- Identified the glycogen phosphorylase phosphorylation site as well as the two known sites (S69/S345) in ovalbumin
- N-acetylation found plus many sites of methionine oxidation (? due to work up or real)
- New sites found - phosphorylation at S237/S241 in ovalbumin and methylation at R652

2/10/09

MacCoss et al, 2002

Results

Cdc2p complex:

- 200 proteins - 20 showed >40% sequence coverage
- Expected Y15 and T167 phosphorylation sites on Cdc2p found
- New phosphorylation sites found on cyclin partners Cdc13 and Cig1p
- Multiple methylation sites on Cdc2p

2/10/09

MacCoss et al, 2002

Results

Lens proteins:

- Found 272 proteins - 52 had >40% sequence coverage
- 90% are crystallins
- PTMs accumulate over your lifetime
- Used 18-step MudPIT because of complexity
- 73 different PTMs found on the 11 crystallins
- Found the 13 of 18 PTMs previously described in all species
- Found 60 other new PTMs in phosphorylation, oxidation, acetylation and methylation

2/10/09

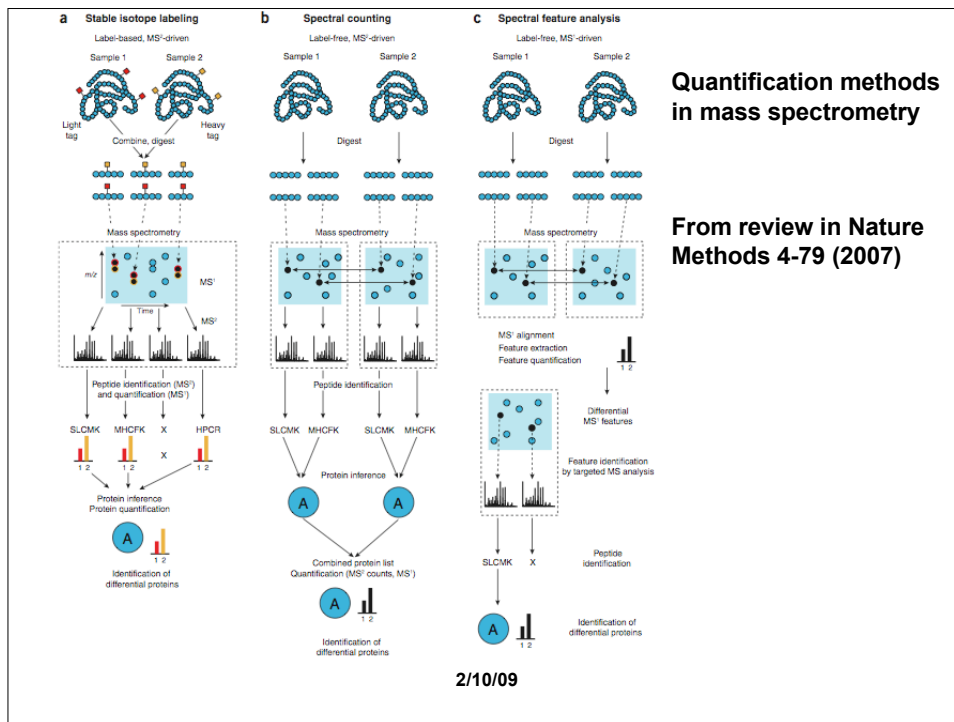
MacCoss et al, 2002

Quantitative proteomics

Use of isotopes

- ICAT (d_0/d_8) and ICAT $^{13}C_0/^{13}C_8$
- d_0/d_{10} propionic anhydride (N-terminal labeling)
- $^{15}N/^{14}N$ (whole cell labeling)
- $^{18}O/^{16}O$ (trypsin)
- iTRAQ labeling
- Non-isotope methods
 - Peptide coverage
 - Classical triple quadrupole methods

2/10/09

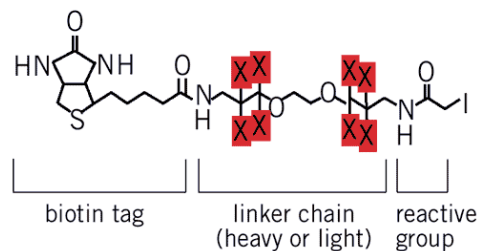


Isotope-coded affinity technology

Isotope-Coded Affinity Tags

heavy reagent: D8-ICAT Reagent (X=deuterium)

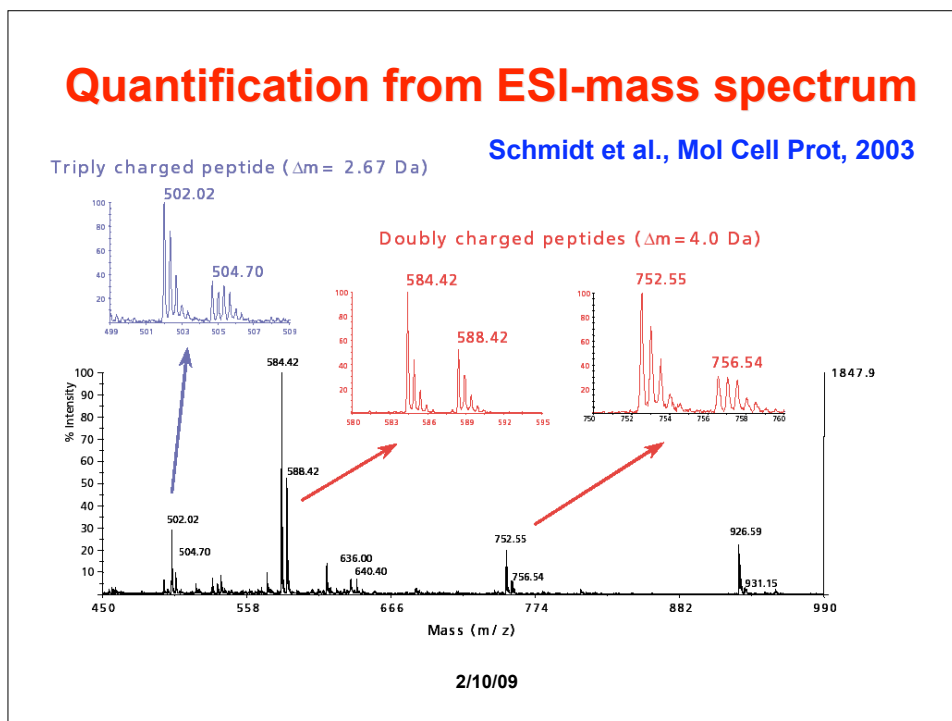
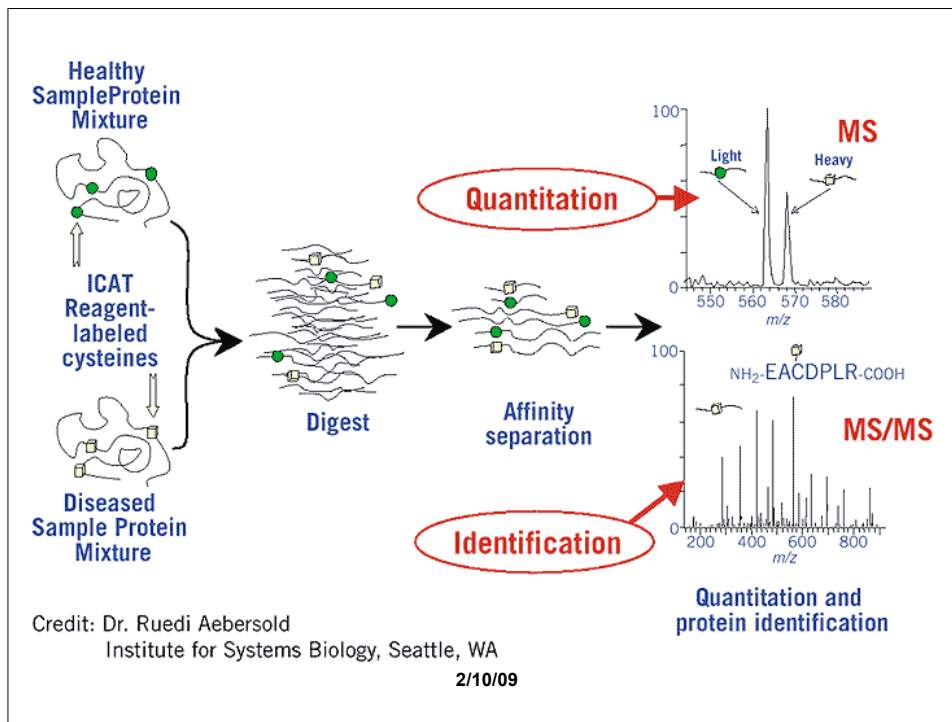
light reagent: D0-ICAT Reagent (X=hydrogen)



This reagent reacts with cysteine-containing proteins (80-85% of proteome)

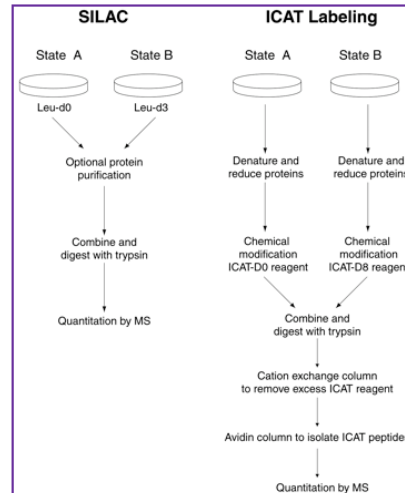
Labeling can be replacement of hydrogens (X) with deuterium, or better to exchange ¹²C with ¹³C in the linker region (this avoids chromatography issues)

2/10/09



SILAC

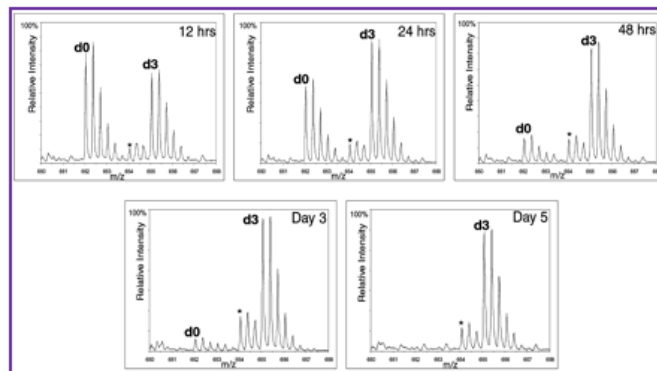
- **SILAC, stable isotope labeling by amino acids in cell culture, is being used to quantify proteins.**



2/10/09

Ong et al., MCP 1:367, 2002

Time-dependent leucine incorporation with SILAC



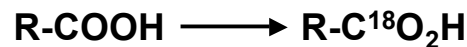
The cells are pre-labeled with leucine-d₀. Leucine-d₃ is added to the medium and cells sampled at various times later. The peaks annotated with d0 and d3 are the triply charged peaks of the peptide VAPEEHPVLLTEAPLNPK, which contains three leucines.

2/10/09

Ong et al., MCP 1:367, 2002

¹⁸O-labeling

- Trypsin catalyzes the transfer of ¹⁸O in ¹⁸O-enriched water to both the carboxylate oxygens of the C-terminus of tryptic peptides

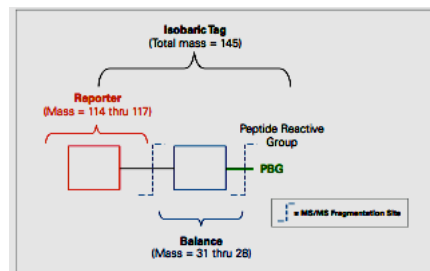


- The peptides have an increase in mass of 4 Da
- Generally not considered a large enough mass difference

2/10/09

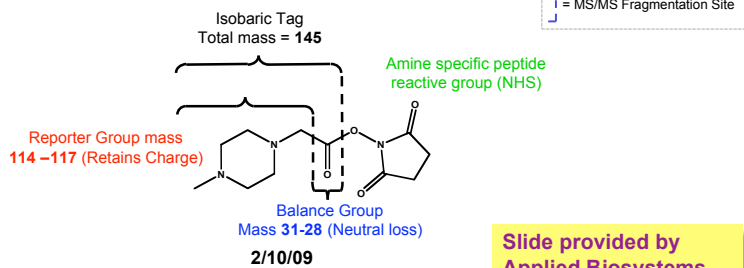
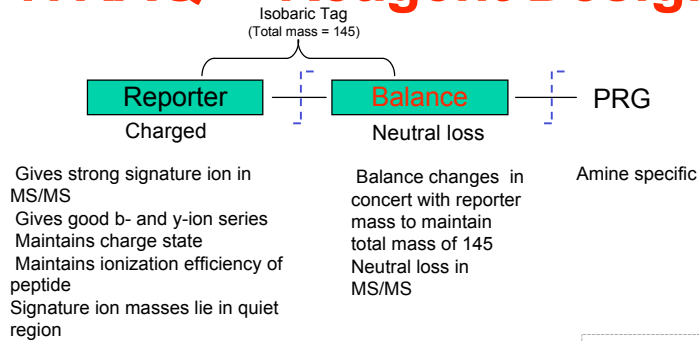
iTRAQ quantification

- The iTRAQ™ reagents
 - React with Lys amino groups and each one adds 145 Da to the molecular weight of the peptide
 - Fragmentation produces reporter ions from *m/z* 114, 115, 116 and 117
 - New iTRAQ kit contains 8 forms with reporter fragment ions of *m/z* 114, 115, 116, 117, 118, 119 and 121



2/10/09

iTRAQ™ Reagent Design

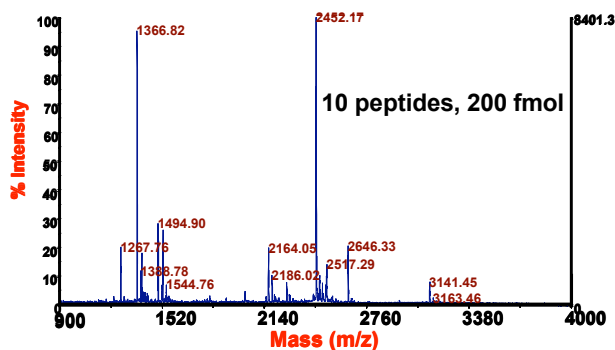


Slide provided by Applied Biosystems

Other non-isotopic quantitative methods in proteomics

The coverage (the number of peptides observed for a protein) is sensitive to the amount of the protein

- This can be used to calculate whether a treatment affects the abundance of a protein where fold-change > 2
- Applies to LC-MS (MUDPIT methods)

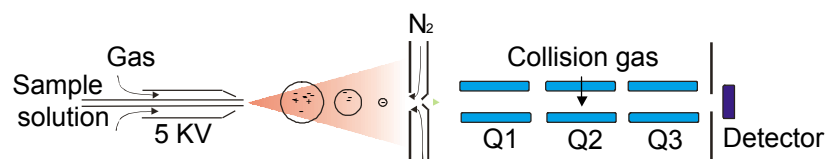


2/10/09

Triple quad MRM analysis

Peptides of interest can be analyzed like small molecules

- Choose the parent molecular ion, collide with argon gas and select a unique fragment



- **Multiple reaction ion scanning**

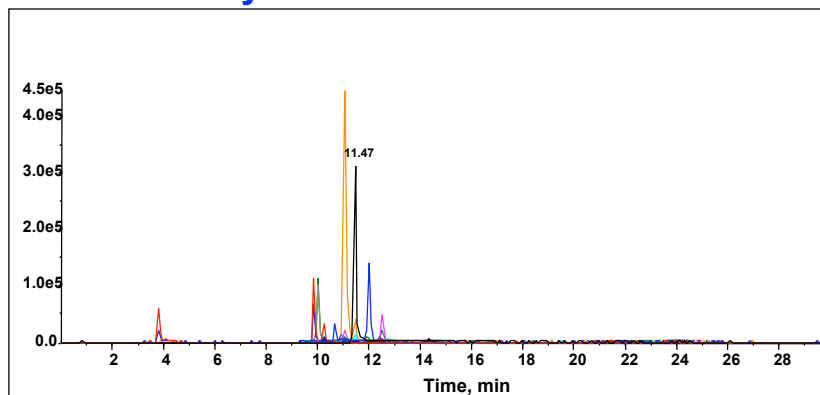
First filter the [M-H]⁻ molecular ion of the analyte (Q1)

Fragment the molecular ion with N₂ gas (Q2)

Select a specific (and unique) fragment ion (Q3)

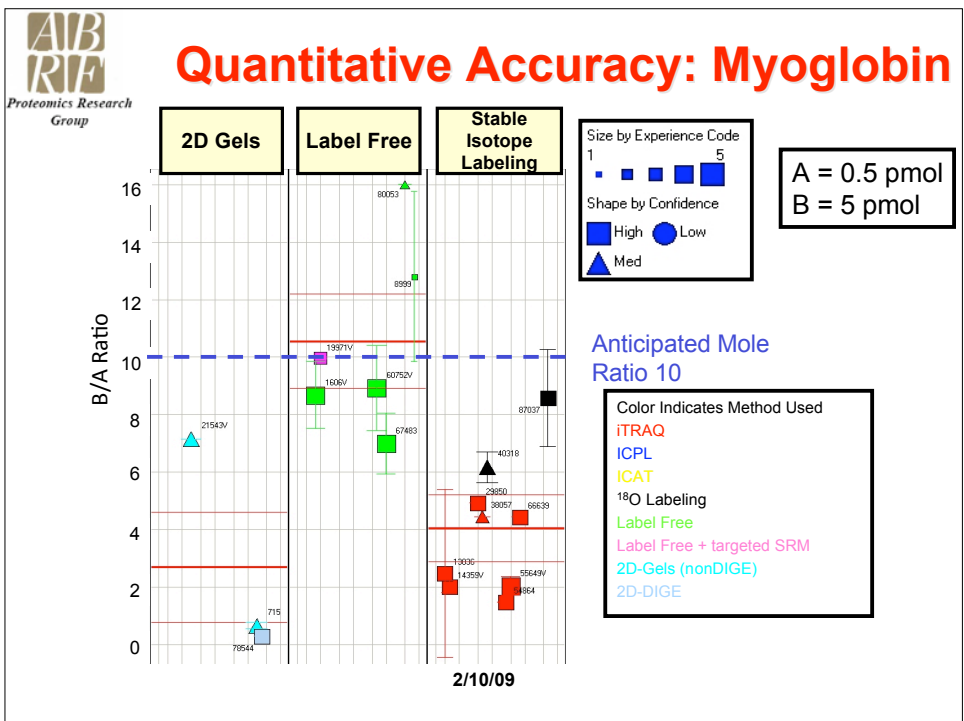
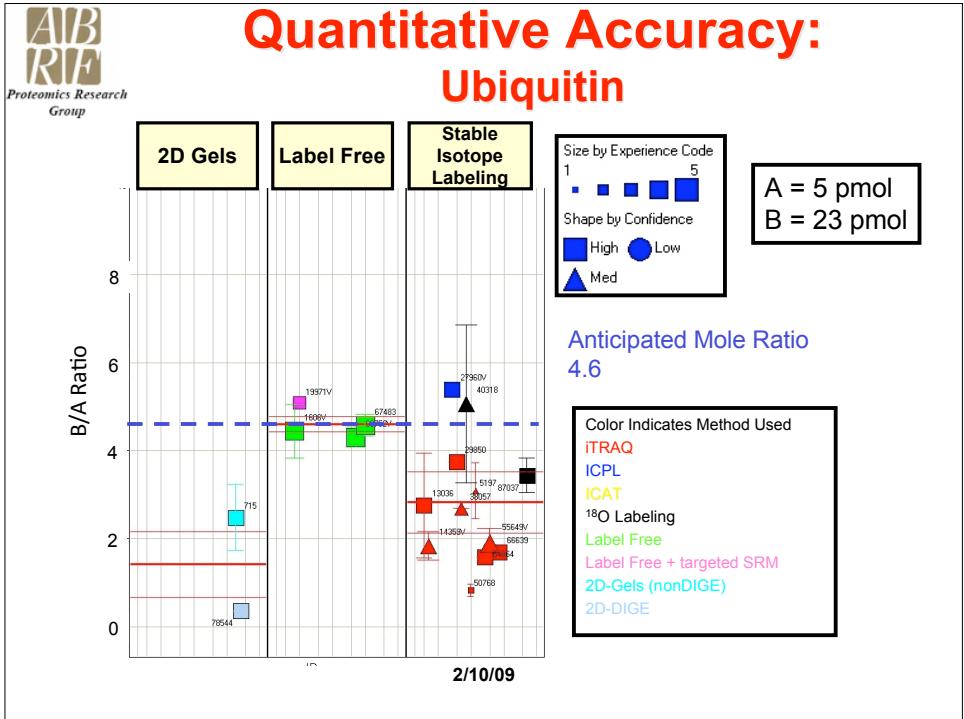
2/10/09

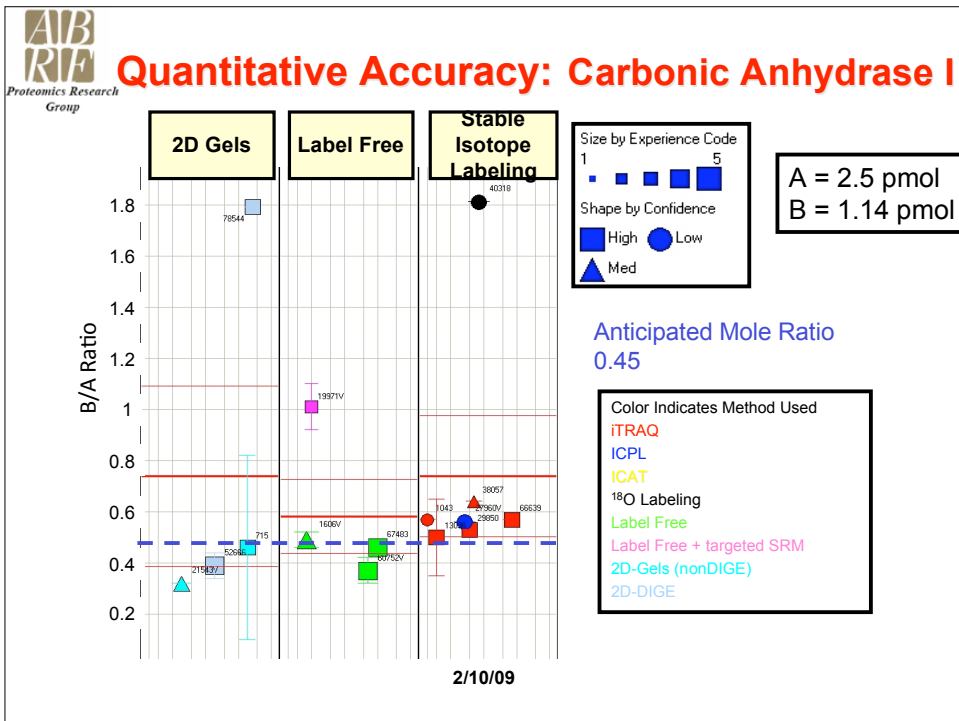
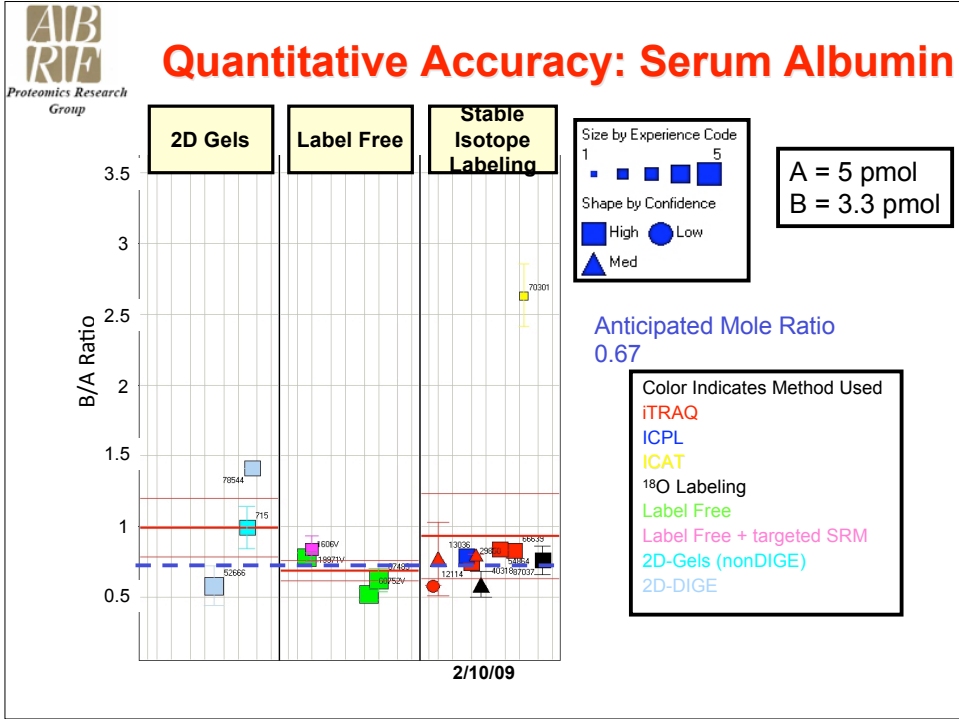
Quantitation experiment for biotinylated cytochrome c MRM analysis monitored in 50 channels



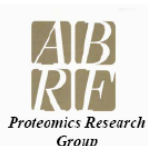
Each colored peak represents a different biotinylated peptide

2/10/09





2008 ABRF Study - identification of three truncated peptides



Conclusions

- Proteomics offers a wide arrange of approaches for the qualitative analysis of proteins
- Many methods and approaches were used successfully identify and sequence the truncated sites
- In many cases, the combination of two complementary approaches (e.g., 1D SDS PAGE to resolve protein components followed by LC-MS/MS for sequence information) gave a higher success rate than use of a single experimental approach.
- As expected, experience remains a key factor in this study

2/10/09

Application of LC-MRM-MS to the plasma proteome

Identify the proteins of interest



In silico, generate the tryptic peptides from each protein



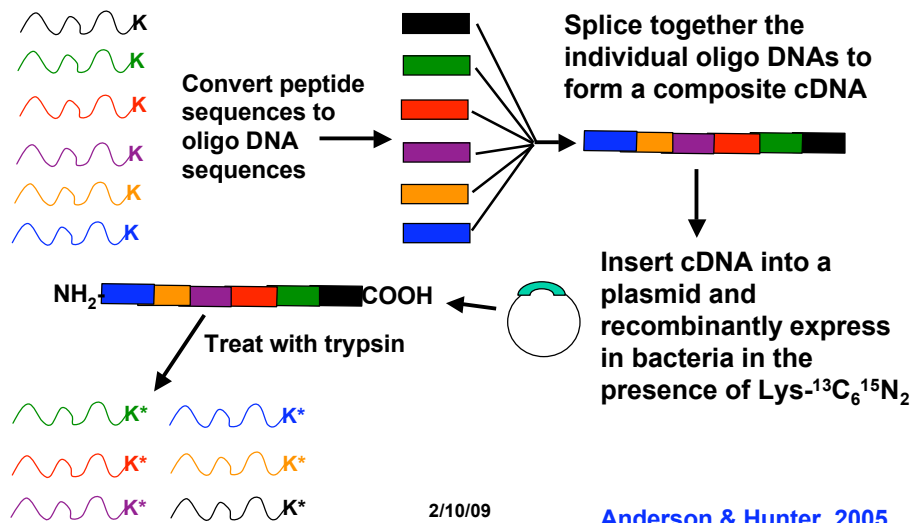
Select the best combination of parent peptide and fragment y ion



Determine the expected y ions for each peptide and compare to y ions of all other tryptic peptides of known human proteins that have masses within $\pm 1\text{Da}$

2/10/09

Concatenation - making ^{13}C -labeled peptide internal standards



Quantitative peptide MRM-MS

- The albumin-depleted plasma proteome is mixed with the composite ^{13}C , ^{15}N -labeled protein internal standard and then treated with trypsin
- The molecular ions (doubly charged) and the specific y ions for each peptide and its labeled form are entered into the MRM script one channel at a time
- A single run may consist of 30 peptides in 60 channels
- Sensitivity is compromised by “sharing out” measurement time, but can be compensated for by carrying out nanoLC

2/10/09

Advantage of a C-terminal labeled lysine

186 301 448 505 642 755 886 987 1115 b ions
A D E F G H I M T K
1133 1062 948 833 686 629 492 379 248 147 y ions

With the labeled lysine at the C-terminus, only the b₁₀ ion contains the isotope atoms

186 301 448 505 642 755 886 987 1123 b ions
A D E F G H I M T K*
1141 1070 956 841 694 637 500 387 256 155 y ions

2/10/09

References for this talk (1)

- Annan RS, Hudleston MJ, Verma R, Deshaies RJ, Carr SA. A multidimensional electrospray MS-based approach to phosphopeptide mapping. *Anal. Chem.* 73:393, 2001.
- Flory MR, Griffin TJ, Martin D, Aebersold R. Advances in quantitative proteomics using stable isotope tags. *Trends in Biotechnology* 20: S23, 2002.
- Taupenot L, Harper KL, O'Connor DT. The chromogranin-secretogranin family. *New Engl. J. Med.* 348: 1134, 2003.
- Lam YW, Mobley JA, Evans JE, Carmody JF, Ho S-M. Mass profiling-directed isolation and identification of a stage-specific serologic protein biomarker of advanced prostate cancer. *Proteomics* 5: 2927, 2005.
- Lehmann WD, Krüger R, Salek M, Hung CW, Wolschin F, Weckwerth W. Neutral loss-based phosphopeptide recognition: a collection of caveats. *J Proteome Res.* 6:2866-73, 2007.

2/10/09

Bibliography (2)

- Ong SE, Mann M. Mass spectrometry-based proteomics turns quantitative. *Nature Chemical Biology*. 1:252-262, 2005.
- Gruhler A, Schulze WX, Matthiesen R, Mann M, Jensen ON. Stable isotope labeling of *Arabidopsis thaliana* cells and quantitative proteomics by mass spectrometry. *Molecular & Cellular Proteomics*. 4:1697-1709, 2005.
- Anderson L, Hunter CL. Quantitative Mass Spectrometric Multiple Reaction Monitoring Assays for Major Plasma Proteins. *Molecular & Cellular Proteomics* 5:573-588, 2006.
- Yao X, Freas A, Ramirez J, Demirev PA, Fenselau C. Proteolytic ¹⁸O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Analytical Chemistry* 73, 2836-42, 2001.
- Wang G, Wu WW, Zeng W, Chou C-L, Shen R-F. Label-Free Protein Quantification Using LC-Coupled Ion Trap or FT Mass Spectrometry: Reproducibility, Linearity, and Application with Complex Proteomes. *Journal of Proteome Research* 5: 1214-1223, 2006.

2/10/09