BMG 744 Proteomics-Mass Spectrometry

Qualitative burrowing of the proteome

Stephen Barnes, PhD
sbarnes@uab.edu

1/23/12

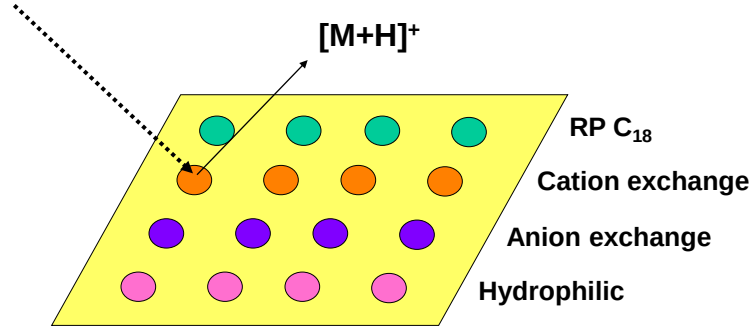
Topics

Qualitative changes in the predicted proteome

- **Digging deeper - SELDI/MALDI**
- **Posttranslational modifications**
  - **Isolation and characterization**
  - **Making using of the chemistry of the modification**
  - **Ubiquitination/SUMOylation**
  - **Global discovery**

1/23/12

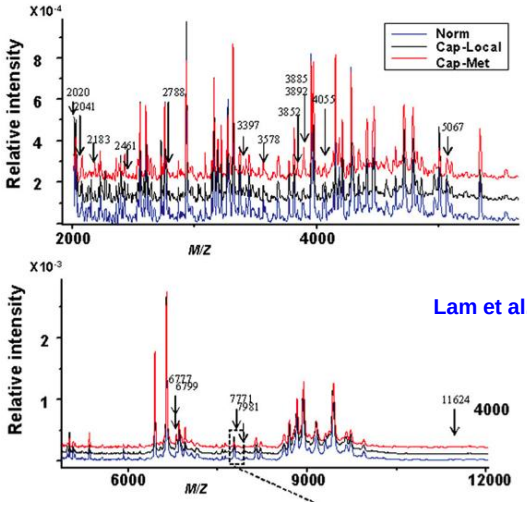
Surface enhanced laser desorption ionization (SELDI)



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

1/23/12

MALDI-TOF analysis of peptides recovered by C_{18} extraction of sera from prostate patients, with and without metastases, and controls

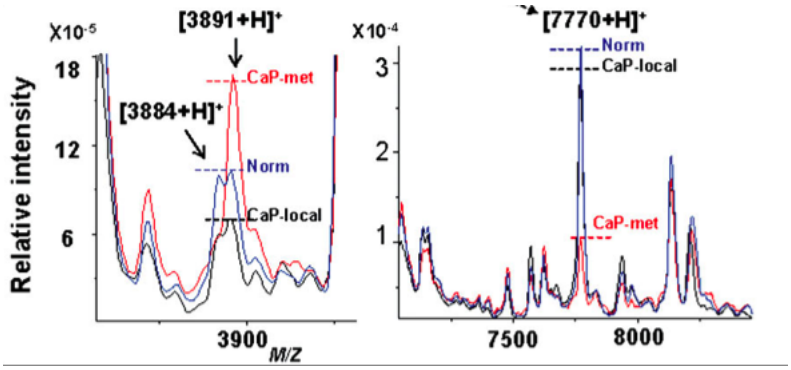


Lam et al., Proteomics 5, 2927

1/23/12

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

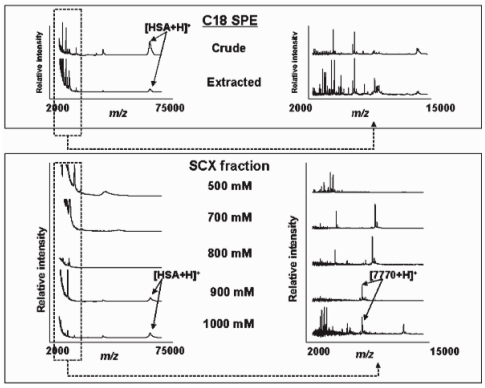
Lam et al., Proteomics 5, 2927



But what is it?

Purification of prostate peptide by SPE extraction and strong cation exchange

Lam et al., Proteomics 5, 2927



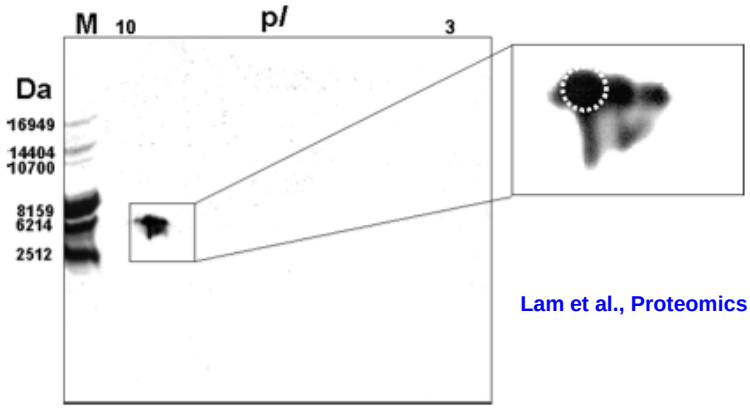
Serum was first extracted with a C_{18} SPE cartridge

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

Purification of prostate peptide needed 2D-electrophoresis

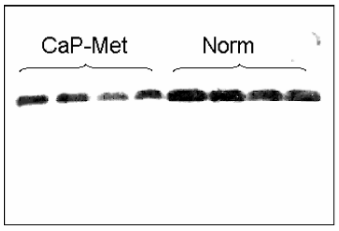


Lam et al., Proteomics 5, 2927

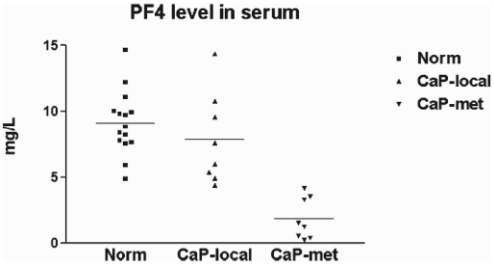
1/23/12

Non-mass spec verification of prostate peptide

Lam et al., Proteomics 5, 2927



Western blot analysis for PF4

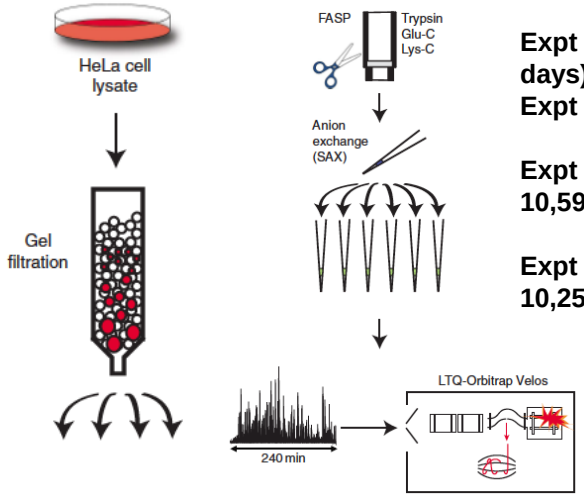


ELISA analysis for PF4

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

1/23/12

How deep can we go?



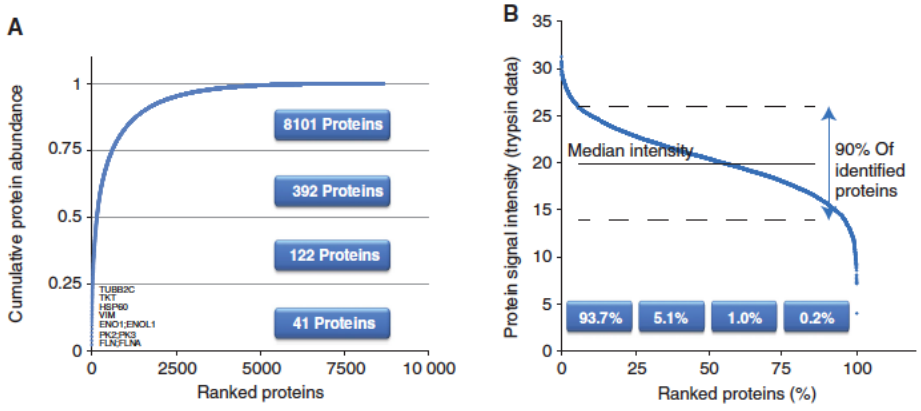
Expt 1 – 126 fractions (21 days)
Expt 2 – 72 fractions (12 days)

Expt 1 – 187,006 peptides and 10,596 proteins

Expt 2 – 163,784 peptides and 10,255 proteins

Nagara et al., Mol Sys Biol 7:548, (2011)

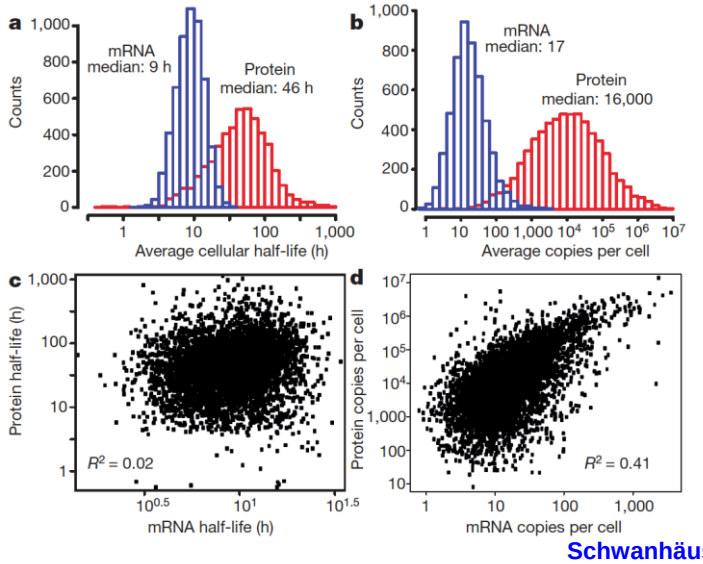
Description of protein abundances



Just 41 proteins accounted for 25% of the combined tryptic peptide ion intensities – 122 accounted for 50% and 392 for 75%.

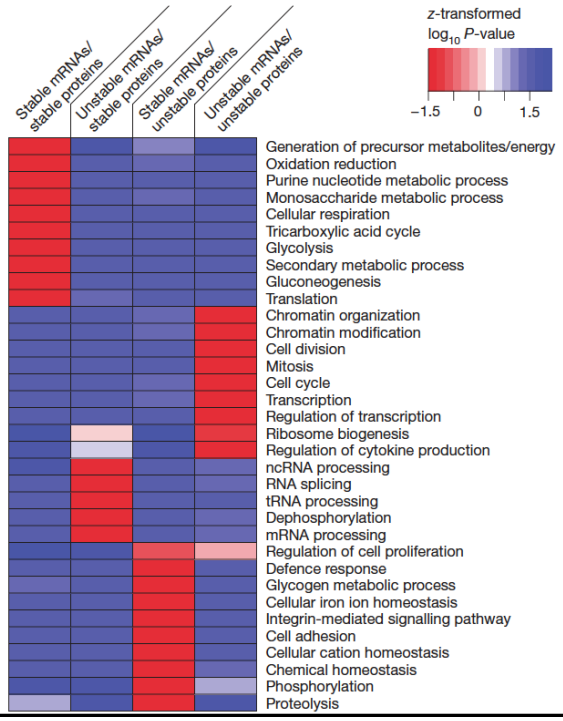
Nagara et al., Mol Sys Biol 7:548, (2011)

Protein – mRNA correlations



Schwanhäusser et al.
Nature 473: 337 (2011)

1/23/12



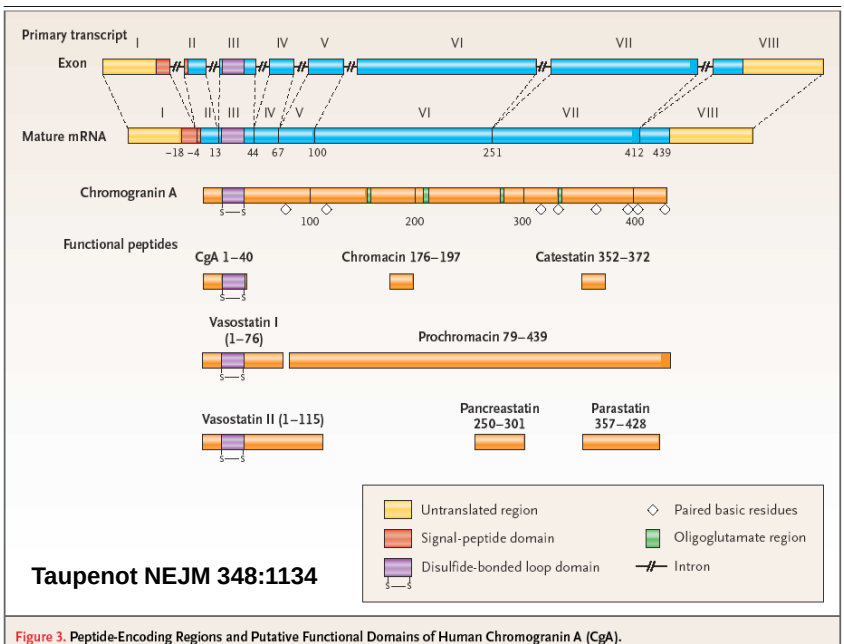
Proteins and mRNA have different stabilities and can be divided into 4 quadrants. The major cellular functions are divided as shown in the table.

Schwanhäusser et al.
Nature 473: 337 (2011)

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**

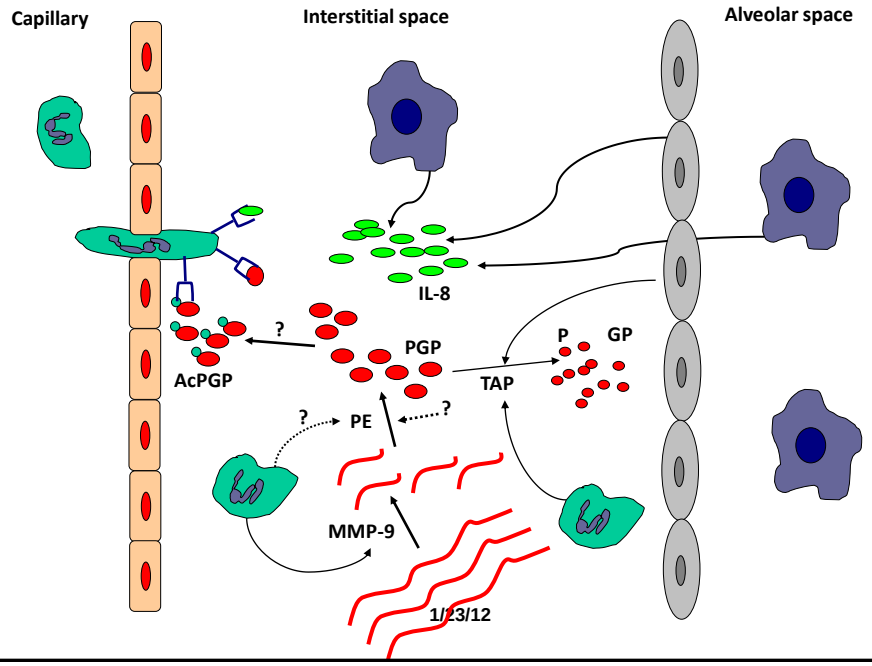
1/23/12



Taupenot NEJM 348:1134

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.

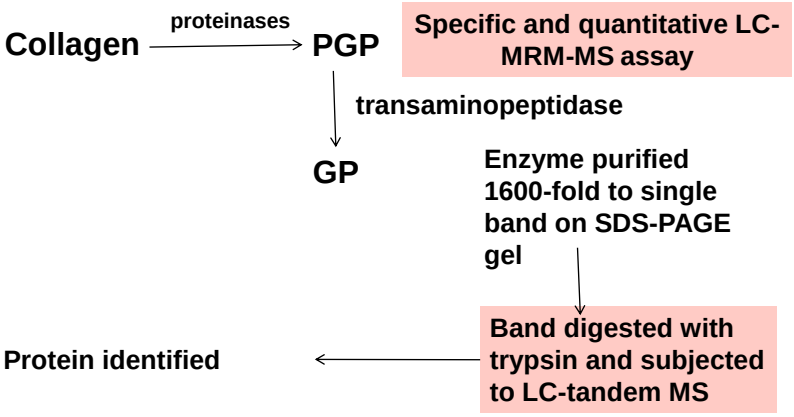
A homeostatic circuit controlling neutrophil recruitment



PGP is a common peptide in human collagen

MFSFVDLRLLLLLAATALLTHGQEEGQVEGQDEDIPPITCVQNGLRYHDRDVWK
PEPCRI
CVCDNGKVLCDDVICDETKNCPGAEVPEGECCPVCPDGSESPTDQETTGVEGP
KGDTGPR
GPRGPAGPPGRDGIPGQPGLPGPPGPPGPPGPPGPLGGNFAPQLSYGYDEKST
GGISVPGP
MGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKNGD
DGEAGKPGR
PGERGPPGPQGARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEPGS
PGENGAPGQ
MGPRGLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAVGA
KGEAGPQGP
RGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGF
PGARGPSGP
QGPGGPPGPKGNSGEPGAPGSKGDTGAKGEPGPVGVQGPPGPAGEEGKRGA
RGEPGPTGL
PGPPGERGGPGSRGFPGADGVAGPKGPAGERGSPGPAGPKGSPGEAGRPGE
AGLPGAKGL
TGSPGSPGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVMGFPGPKGAAGE
PGKAGERGV
PGPPGAVGPAGKDGEAGAQGPPGPAGPAGERGEQGPAGSPGFQGLPGPAGP
PGEAGKPGE
QGVPGDLGAPGPSGARGERGFPGERGVQGPPGPAGPRGANGAPGNDGAKGD
AGAPGAPGS
QGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGP

Mass spec contribution to PGP story



(Robert Snelgrove et al., Science 2010)

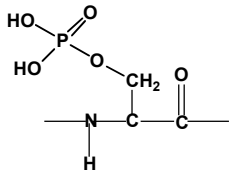
1/23/12

Enzymatic modifications

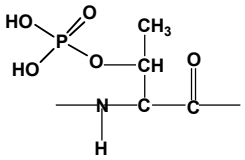
- Phosphorylation/dephosphorylation
  - On serine, threonine, tyrosine
- Glycosylation
  - N-glycosylation (asparagine-linked)
  - O-glycosylation (serine-, threonine-linked)
- N-Acetylation/deacetylation
  - On lysines
- N-Methylation/demethylation
  - On lysines - mono-, di- and trimethylation

1/23/12

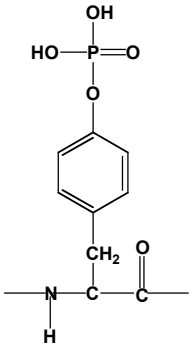
Chemistry of phosphorylation



Serine



Threonine



Tyrosine

1/23/12

Mass spectrometry of phosphorylated proteins

- Adds H_3PO_4 (+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 ^{81}Br, 79.9083 Da)

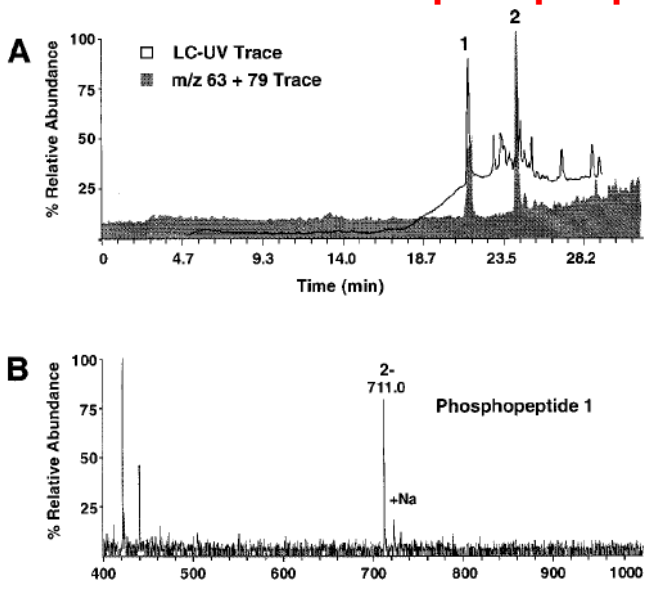
1/23/12

Finding a phosphate group

Several methods are in current use for detection of phosphopeptides

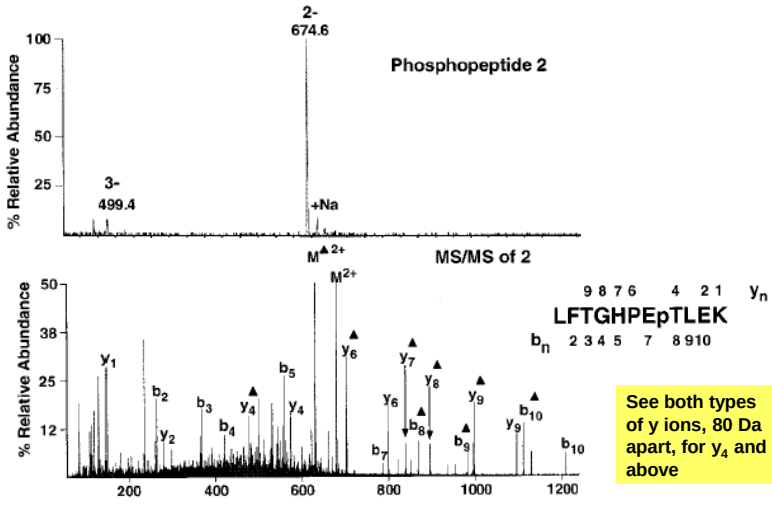
- use of parent ion or neutral loss scanning
- phosphatase sensitivity
- affinity methods for enrichment of phosphopeptides
  - anti-phospho-Ser/Thr/Tyr antibodies
  - metal ion affinity
  - chemical reaction/biotin affinity

Parent ion scan to detect phosphopeptide



Annan Anal Chem 73:393

Phosphopeptide 2



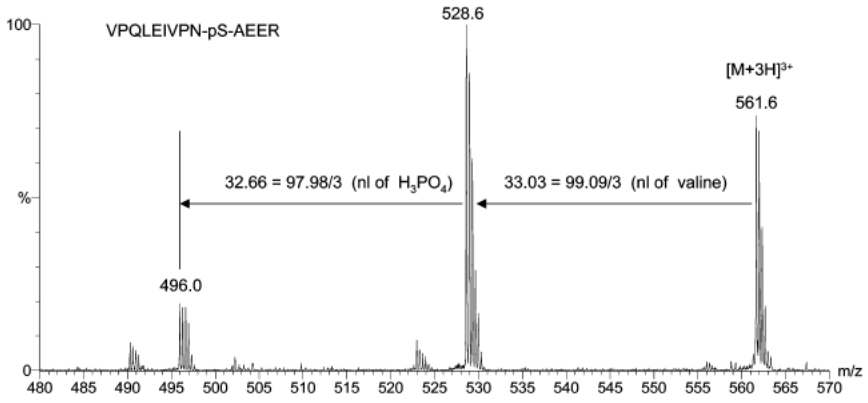
See both types of y ions, 80 Da apart, for y_4 and above

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

Casein peptide - mono- or diphosphate?



Lehmann et al., J Prot Res 7-2866

1/23/12

Problems of low mass resolution for parent ion scanning for phosphate

Lehmann et al., J Prot Res 7-2866

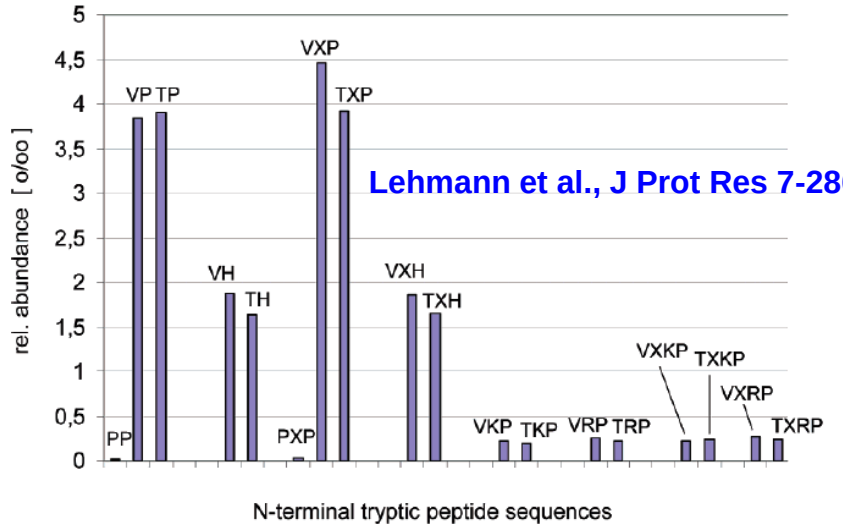
Table 1. Peptide Modifications Showing an Abundant 'Close-to-98/z' Neutral Loss Compared to the Specific Loss of H_3PO_4 from pSer/pThr Phosphopeptides[a]

structure	neutral loss	1+	2+	3+	4+
phosphoS/phosphoT	H_3PO_4	97.9769	48.9884	32.6590	24.4942
P -P... (N-term)	Pro	97.0528	48.5264	32.3509	24.2632
C-sulfo	$SO_3 + H_2O$	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_2O	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).

1/23/12

Residues in the proteome with phosphate-like neutral loss



Lehmann et al., J Prot Res 7-2866

Recovery and enhancement of phosphopeptides

The biggest problem in the detection of phospho-peptides 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**

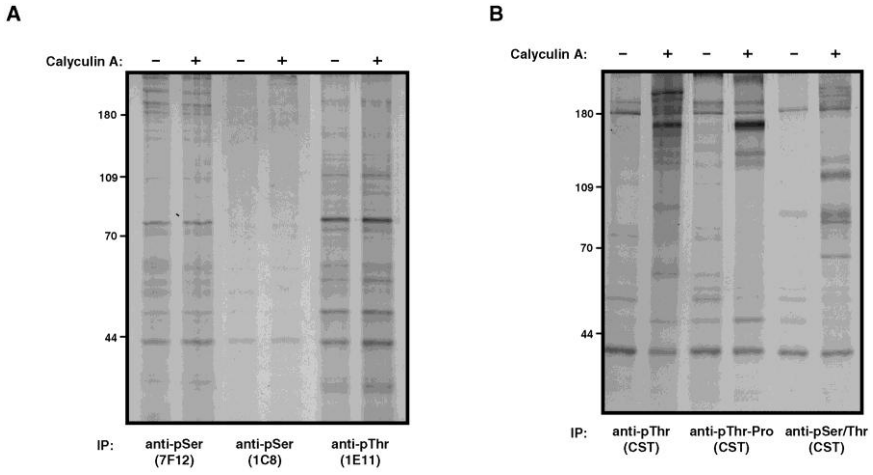
Antibodies and phosphopeptides

- **In this approach, both the phosphorylated and non-phosphorylated forms of a protein may be recovered from the sample matrix**
- **This can most easily achieved by immuno-precipitation 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***

1/23/12

Variability of anti-phosphoserine and anti-phosphothreonine antibodies



1/23/12

Gronborg et al., 2002

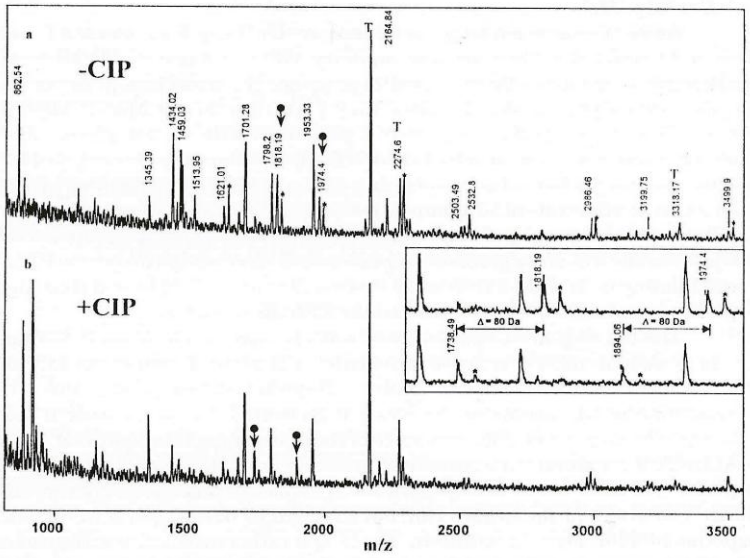
Detection of phosphopeptides based on their sensitivity to phosphatase

- 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

1/23/12

Detecting a phosphopeptide with alkaline phosphatase



1/23/12

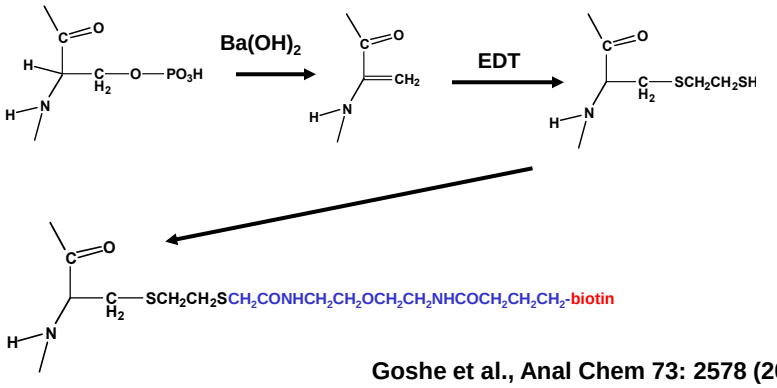
Zhang et al., 2000
MS in Biology & Medicine

Selective enhancement of phospho-peptides 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
- TiO_2 is now being used with success

1/23/12

Selective biotinylation of phospho-groups



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

1/23/12

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/)**
  - **Phosphosite is more focused http://www.phosphosite.org/staticAboutPhosphosite.do**

1/23/12

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.

1/23/12

FindMod

Some of the modifications considered are:

acetylation	amidation	biotinylation
carboxymethyl cysteine	carboxyamidomethyl cysteine	
C-mannosylation	deamidation	flavinylation
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		

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).**

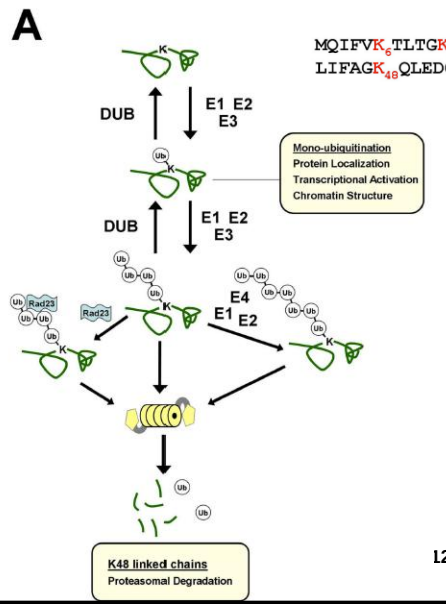
Site for compilation of PTMs

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)

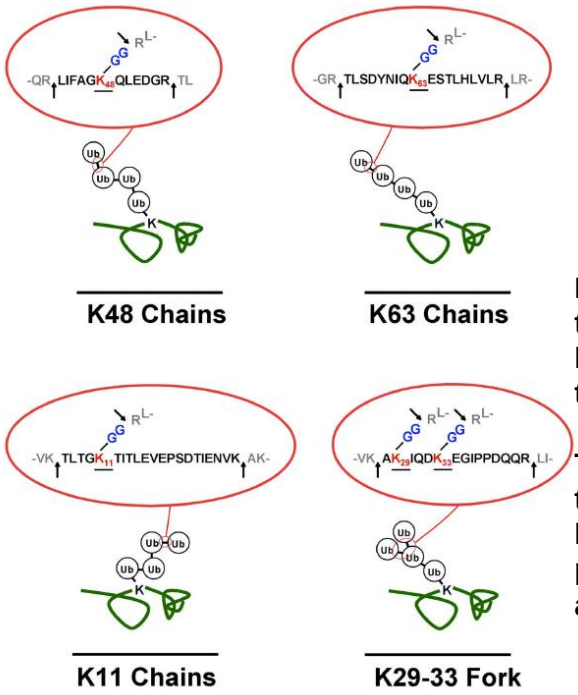
Ubiquitination



Ubq adds to lysine residues on proteins; it can also add to itself at K_6, K_{11}, K_{21}, K_{29}, K_{33}, K_{46} and K_{66} to form polyubiquinated proteins

Kirkpatrick et al., 2005

12

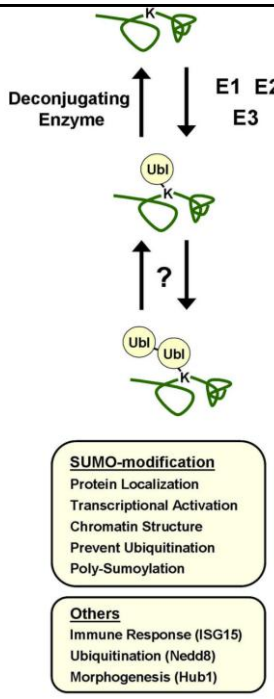


Sites of attachment of Ubq groups

Note that Ubq is attached through the side chain of lysine via the two C-terminal glycines of Ubq.

The next residue is Arg – therefore, trypsin hydrolysis will yield a peptide modified by the addition of 114 (*m/z* 261)

Kirkpatrick et al., 2005



SUMOylation

There are ubiquitin-like proteins (Ubls), one of which is Small Ubiquitin-like Modifier (SUMO)

Have different sequence, but similar structure. Recognize the consensus motif Ψ-K-x-D/E for SUMOylation. Connected via a Gly-Gly group to the lysine residue (just like ubiquitin)

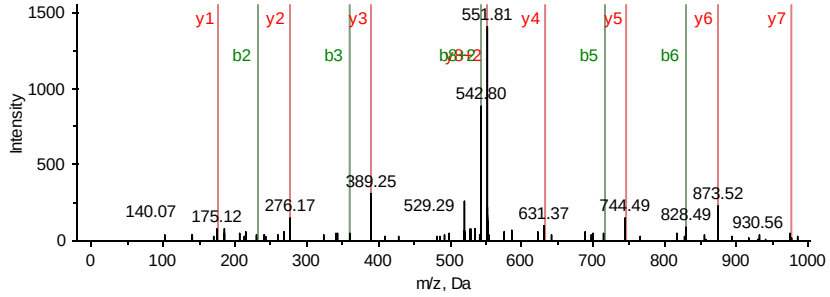
http://sumosp.biocuckoo.org/faq.php

1/23/12

Kirkpatrick et al., 2005

Residue	b	y
Q		1102.622
T	230.1135	974.5629
E	359.1561	873.5152
L	472.2402	744.4726
K[UGG]	714.3781	631.3886
L	827.4621	389.2507
T	928.5098	276.1666
R	1084.611	175.119

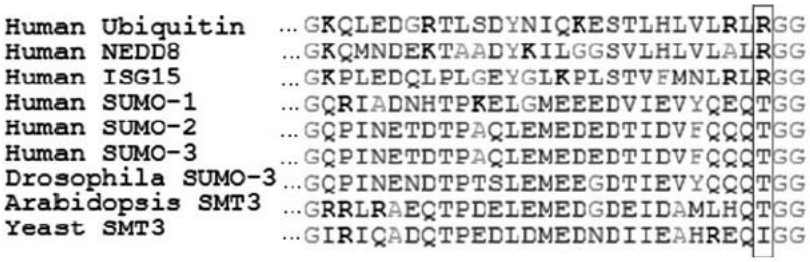
QTELK$_{(SUMO)}$LTR



1/23/12

McClure et al., unpublished

Ubiquitin versus SUMO



So, trypsin will not cut SUMO sites like for ubiquitin to generate a Gly-Gly modified Lysine

Knuesel et al., 2005

Detecting PTMs

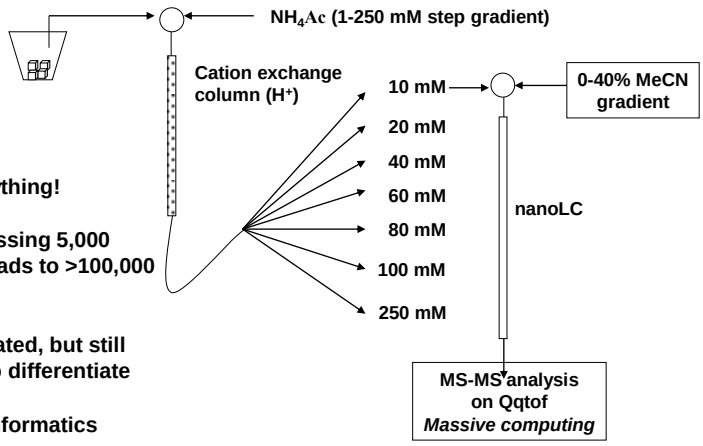
- If we don't have any idea about what the modification(s) is(are), how do we proceed?
- We won't have the modification on MASCOT or Protein Prospector
- No antibodies or convenient affinity phases

Can we approach this globally?

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

1/23/12

MudPIT - Multi-dimensional Protein Identification Technology



- Hydrolyze everything!
- For a cell expressing 5,000 proteins, this leads to >100,000 peptides
- Can be fractionated, but still 10,000-20,000 to differentiate
- Enormous bioinformatics problem

1/23/12

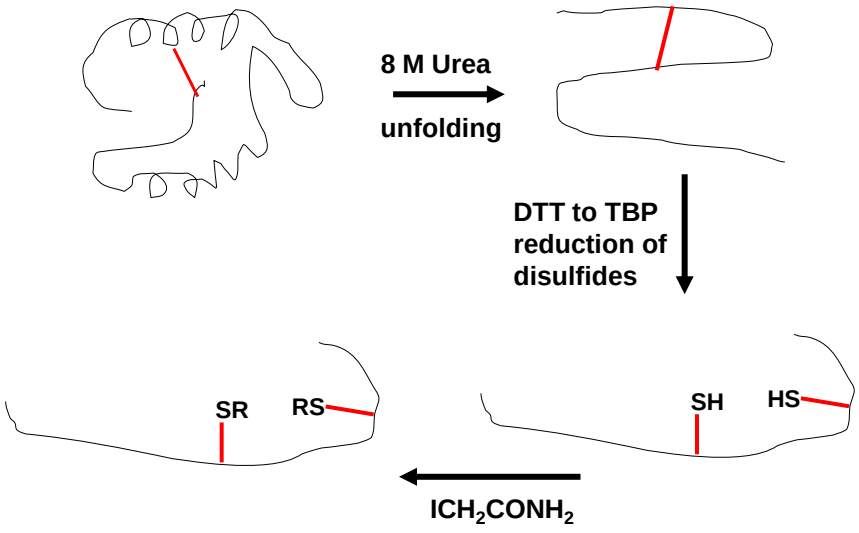
John Yates

Std mix	cdc2p complex	Lens proteins
10 pmol of stds and 1 pmol of Pi-protein: 8 M urea-100 mM Tris-HCl, pH 8.5 added (30 µl)	15 µg dissolved in 40 µl of 8 M urea-100 mM Tris-HCl, pH 8.5	Lens blended in 0.1 ml 20 mM sodium phosphate buffer-1 mM EGTA buffer - spun at 10,000g for 30 min Supernatant dissolved in 8 M urea-100 mM Tris-HCl, pH 8.5
Add 0.8 µl 100 mM DTT, incubate at 50°C for 25 min	Add 0.8 µl 100 mM Tris (2-carboxyethyl)-phosphine, incubated at room temp for 25 min	Add DTT to 2 mM, incubate at 50°C for 25 min
Cool and add 1.7 µl 100 mM iodoacetamide to alkylate	Add 1.7 µl 100 mM iodoacetamide to alkylate	Add 20 mM iodoacetamide to alkylate

1/23/12

MacCoss et al, 2002

Preparing proteins for digestion



1/23/12

Hydrolysis procedures

Reduced, carboxymethylated in 8 M urea

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

↓

100 mM $CaCl_2$ 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%

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

↓

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

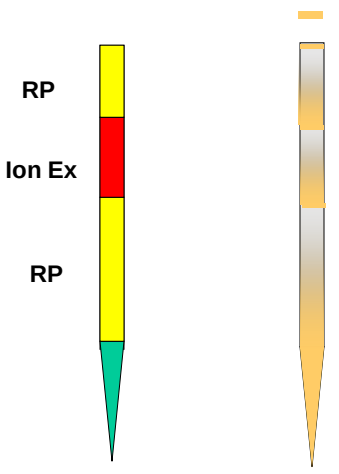
↓

quenched with 90% formic acid to 4%

1/23/12

MacCoss et al, 2002

Elution from a triphasic column



← sample

Wash with 5% MeCN-0.1% acetic acid - elute with 80% MeCN-0.1% formic acid

Wash with 5% MeCN-0.1% acetic acid - elute with 0-500 mM $(NH_4)_2CO_3$

Wash with 5% MeCN-0.1% acetic acid for 5 min - elute with 5-64% MeCN-0.02% HFBA

For lens digests, the sample was passed over a separate 3 cm RP column - after washing, the peptides were eluted with 80% aqueous acetonitrile-0.1% acetic acid

1/23/12

MacCoss et al, 2002

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)

MacCoss et al, 2002

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

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

 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

 MacCoss et al, 2002

Summary

- The proteome is very complex
- Some proteins are in pro-forms and can be found in many, often small, but active fragments
- In addition, proteins are heavily modified
- Modification can account for marked differences in turnover
- Selective fractionation is needed to demonstrate PTMs that are in low abundance

1/23/12

Paper to read

- McClure M, DeLucas LJ, Wilson L, Ray M, Rowe S, Wu X, Dai Q, Hong JS, Sorscher EJ, Kappes JC, Barnes S. Palmitoylation and other post-translational modifications of CFTR with relevance to gating and processing. *Protein Eng Des Sel. 2012 Jan;25(1):7-14*.

1/23/12