Proteomics Workshop

Stanford MS users' meeting

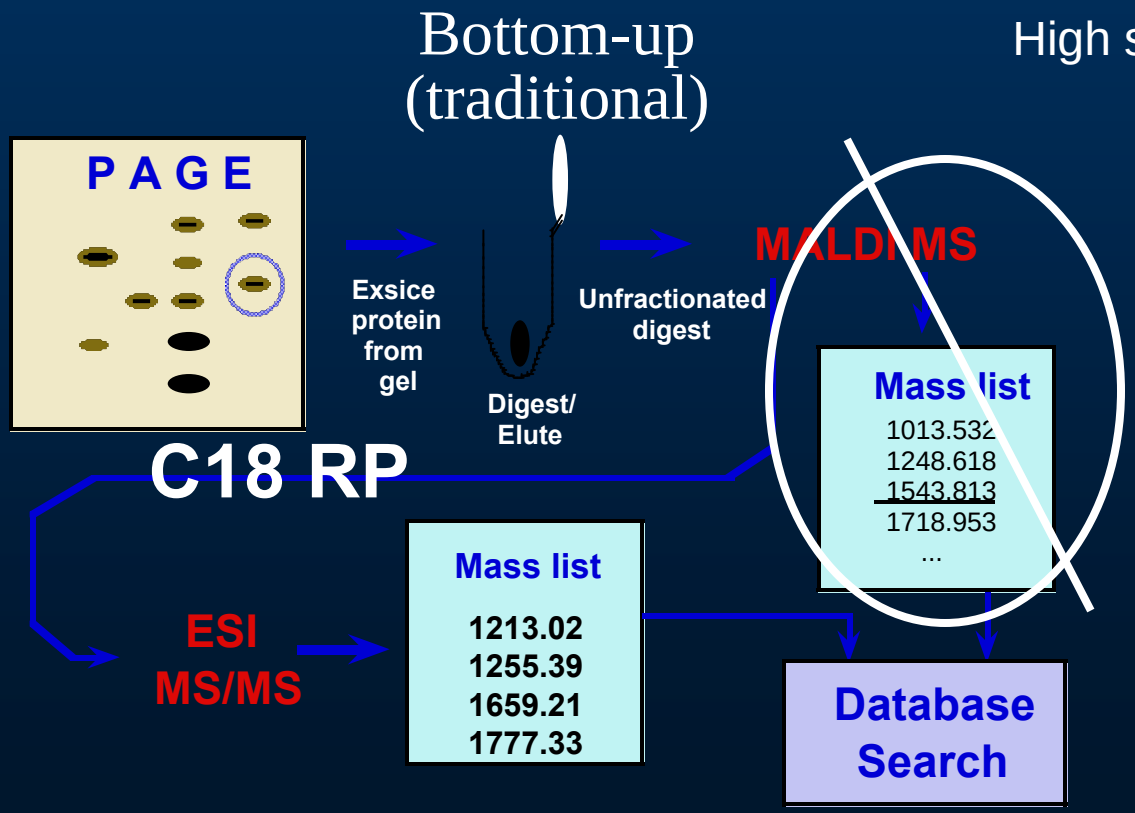
Chris Adams, Ph.D.

Thursday, August 21, 2008

Stanford University
MASS SPECTROMETRY

The Proteomics Workflow

Chris Adams, PhD
Stanford University Mass Spectrometry
http://mass-spec.stanford.edu/



High sensitivity, throughput, but:
No intact MW information

Modifications, sequence errors are easily missed

False positives

Proteomic Applications and Mass Spectrometry

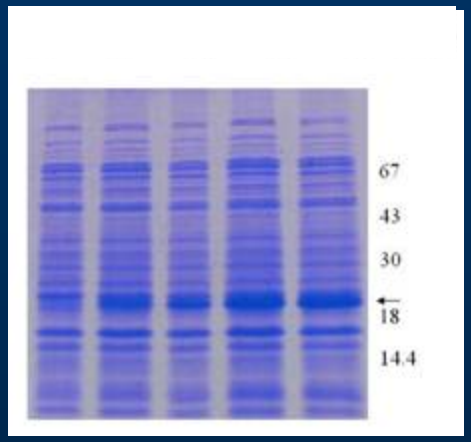
Protein ID- 1D, 2D gels

Complex Mixtures-
Cell lysates, IP's

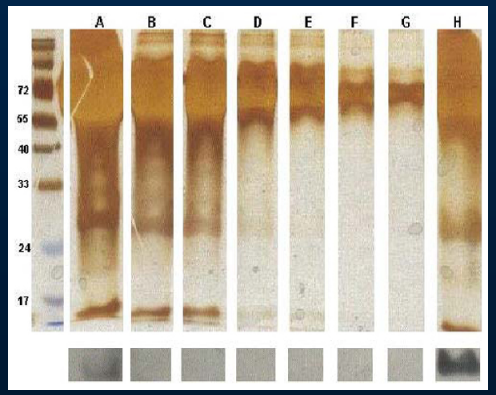
PTM's- In Vivo/ In Vitro
Increased Sequence Coverage
Custom Labels, Phosphorylation,
Acetylation, Ubiquitination ect.

Stains

Coomassie



Silver



Detection Limits

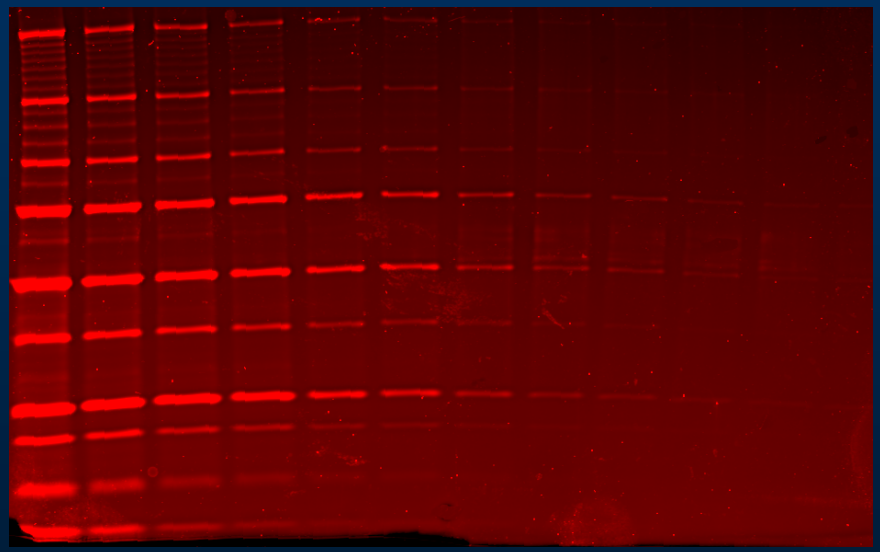
Brilliant Blue 50 ng
Colloidal 10-20 ng

Mass Spec Compatible* 1-5 ng

***No fixing/staining steps involving formaldehyde/glutaraldehyde**

The Compromise

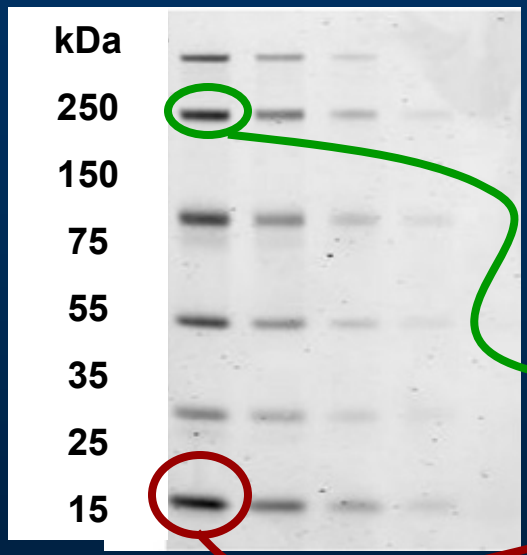
Sypro Ruby



Detection Limit

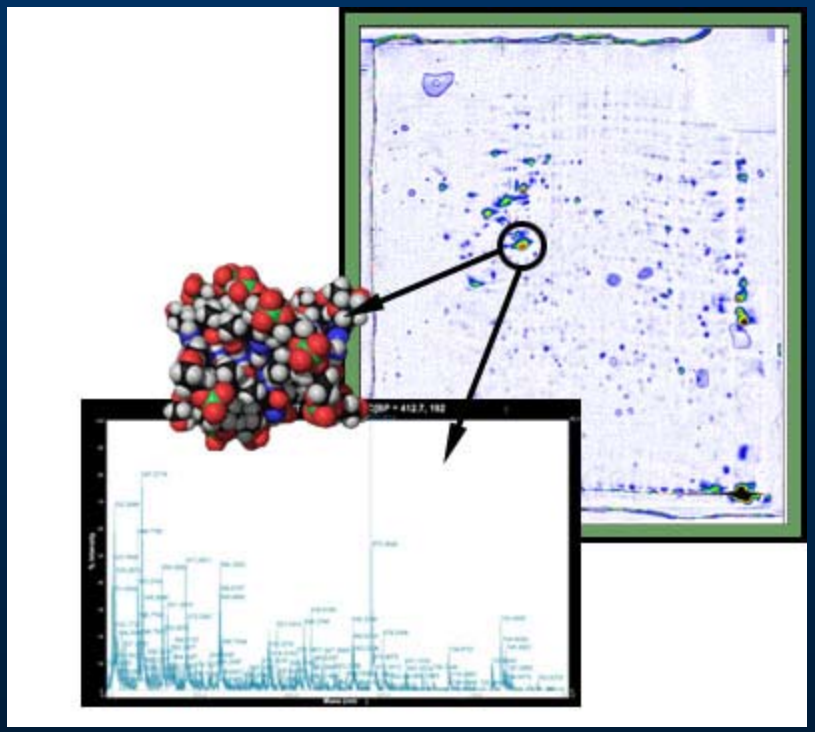
5-10 ng

Size Matters? Why



- detection limit of protein staining is on a weight basis
- detection limit of protein with the mass spectrometer is on a molar basis
- higher the molecular weight, at the same mass, the higher the detection limit will be for the mass spectrometer
- 1.0ng of a 15kd protein is 67 fmol, while 1.0ng of a 250kd protein is only 4 fmol.
- Both proteins will have similar stain intensities, but there is 15 times less protein on a molar basis from the 250kd protein.
- Protein stains detect total protein, mass spectrometer detects proteins individually.

2D Gel Spots



Are OK at low mass and preferentially Coomassie stained

Band Excision, Sample Handling
Yours/Mine/Mouse/Lab Partners Hair and Skin

Ideas to Keep Gels Clean

Gloves

Clean scapel

Clean “cutting surface”

Hood

Eppendorfs not exposed to atmospheric particles

Pippet tips not exposed to atmospheric particles

In Solution Digestion - Consistency Counts

-20 Acetone precipitation followed by

Reduction (DTT) alkylation (IAA) and tryptic digest

Should not/cannot contain detergents (tween) or surfactants (SDS)

Will include a “stage tip” cleanup prior to LC MSMS

Complex Samples/Mixtures

Lysates
& maybe
IP's

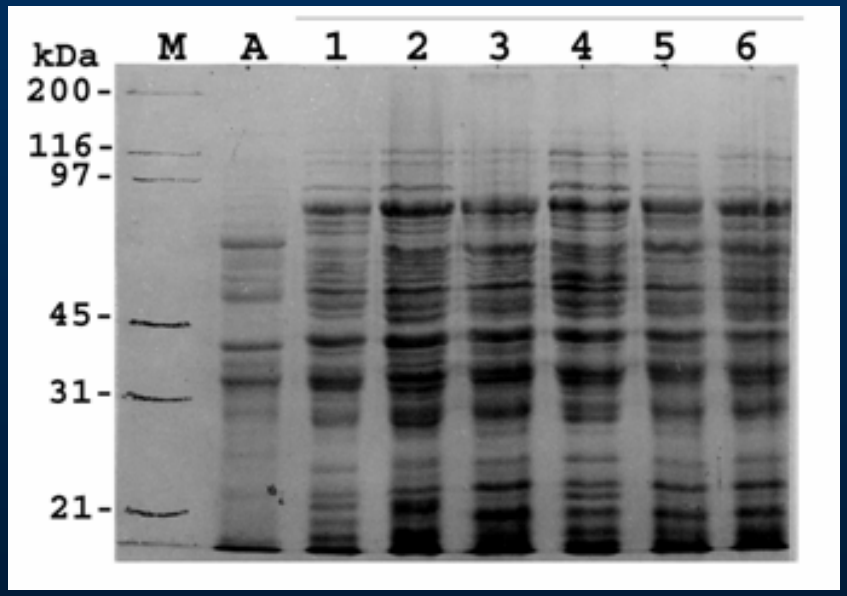
In-Solution Digest
then decide how do we reduce complexity?

Mudpit (offline)

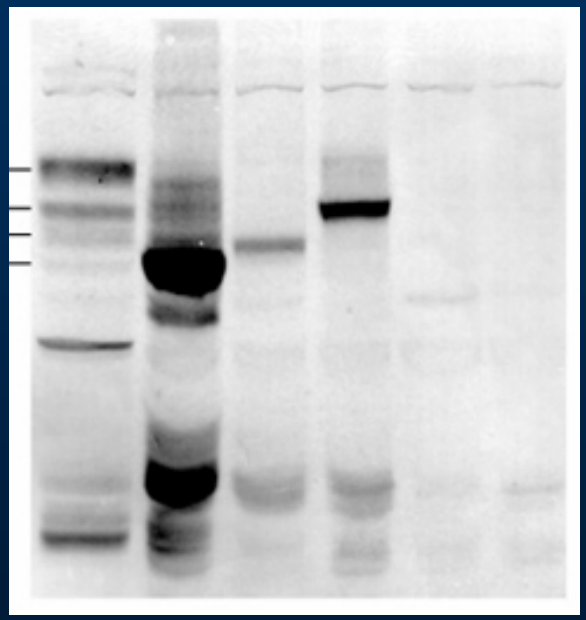
Fractionate by Hydrophobicity

Increase LC MSMS Gradient

How Complex??

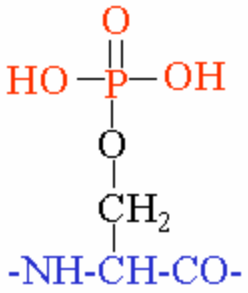


vs

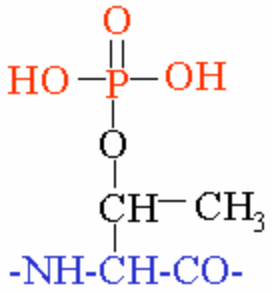


Modifications - know yours

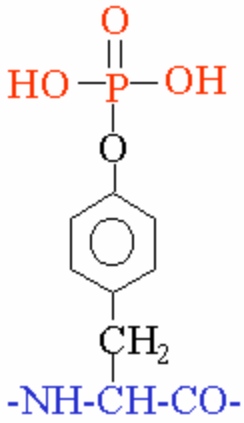
Phosphorylation



PhosphoSerine

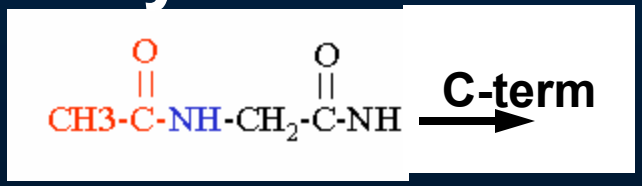


PhosphoThreonine



PhosphoTyrosine

Alkylation



ECT…….

Pictures Ionsource.com

Modifications are Ubiquitous

TABLE I

List of detected base (unmodified) peptides of the abundant protein actin in a human proteome sample (A431 cell line)

Mowse scores (M-score) of unique base peptides and numbers of unique dependent peptides detected for each base peptide by ModifiComb are given. The data are pooled from three independent LC/MS/MS runs.

Base peptide	Position	M-score	Dependent peptides
AGFAGDDAPR	18–27	83	3
AVFPSIVGRPR	28–38	75	4
HQGVMVGMGQK	39–49	111	18
DSYVGDEAQSK	50–60	103	13
DSYVGDEAQSKR	50–61	63	8
GILTLK	62–67	35	1
YPIEHGIVTNWDDMEK	68–83	91	10
MWHHTFYNELR	84–94	81	18
VAPEEHPVLLTEAPLNPK	95–112	95	16
TTGIVMDSGDGVTHTVPIYEGYALPHAILR	147–176	135	16
LDLAGR	177–182	35	3
DLTDYLMK	183–190	67	10
GYSFTTTAER	196–205	61	8
DIKEK	210–214	35	1
EKLCYVALDFEQEMATAASSSSLEK	213–237	29	1
LCYVALDFEQEMATAASSSSLEK	215–237	98	2
SYELPDGQVITIGNER	238–253	86	10
CPEALFQPSFLGMESCGIHETTFNSIMK	256–283	39	1
CDVDIRK	284–290	42	5
KDLYANTVLSGGTTMYPGIADR	290–311	149	14
DLYANTVLSGGTTMYPGIADR	291–311	54	17
EITALAPSTMK	315–326	93	9
IKIIAPPER	326–335	35	2
IIAPPER	328–334	31	1
IIAPPERK	328–335	30	1
QEYDESGPSIVHR	359–371	116	10

21mer

M. Nielsen, et al, *Mol. Cell. Prot.*, 2006, 5:2384

Modifications the Need for Enrichment

Extent of Modifications in Human Shotgun Proteomics Samples

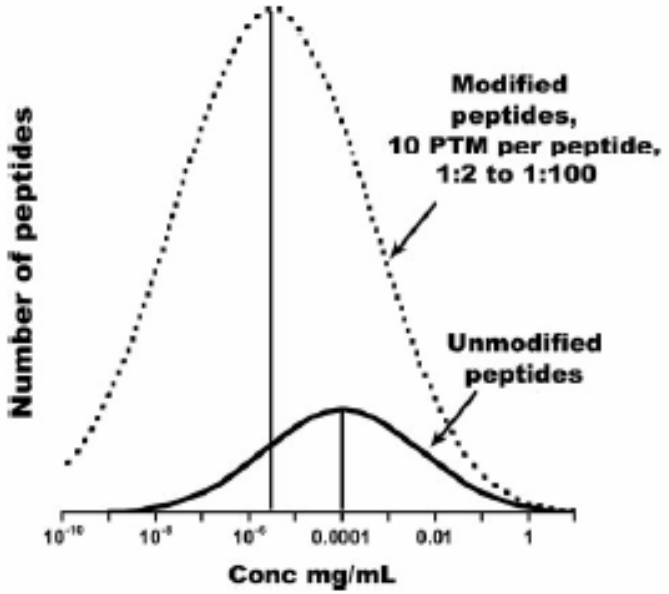


FIG. 4. Theoretical distribution of unmodified tryptic peptide concentrations in a complex biological sample (*solid line*) and the resulting distribution of modified peptide concentrations (*dashed line*) assuming 10 modifications per peptide at a substoichiometric range of 1:2 to 1:100.

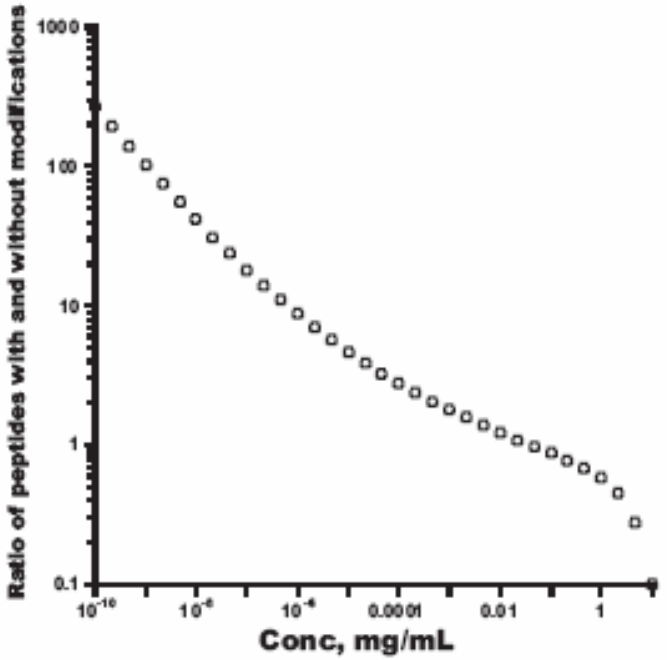


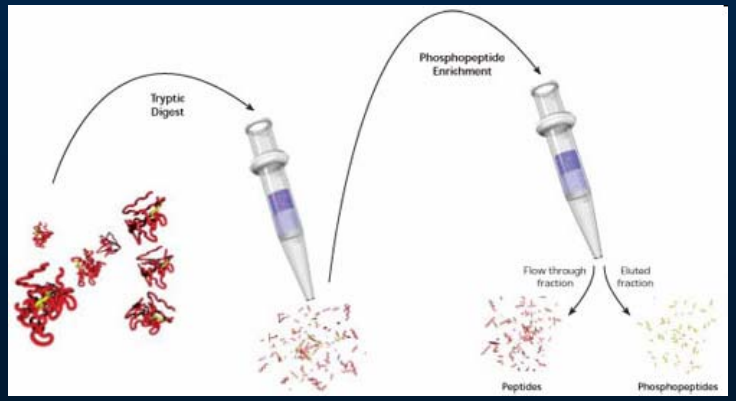
FIG. 5. The average number of modified peptides per single unmodified peptide at a given concentration. The distributions of modified and unmodified peptides are shown in Fig. 4.

M. Nielsen, et al, *Mol. Cell. Prot.*, 2006, 5:2384

Enrichments Help

Bottom Up (Peptide) or Top Down (Protein)

In-house Galium phosphopetide enrichment and TiO2 Enrichment Capabilities



***Affinity purification works but be cautious about your antibody**

Want over abundance of starting material

Increased Sequence Coverage - Multiple Enzymes

Knowing your protein before hand and potential sites of interest

In-Silica Digest →

Favorites ↑

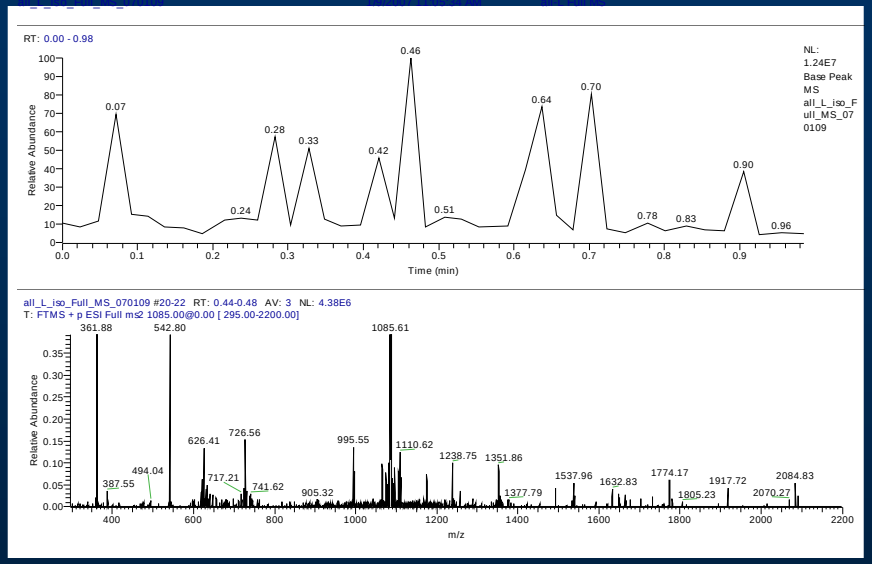
Trypsin
LysC
ArgC
Chymotrypsin
GluC
AspN

→ Ideal peptide 8-12mer, w/ modif. site centrally located

MS/MS and Data Validation

Data validation takes time, patience and N ≥ 2

Mass accuracy matters



I & L isomeric

K (128.09496) v. Q (128.05858)

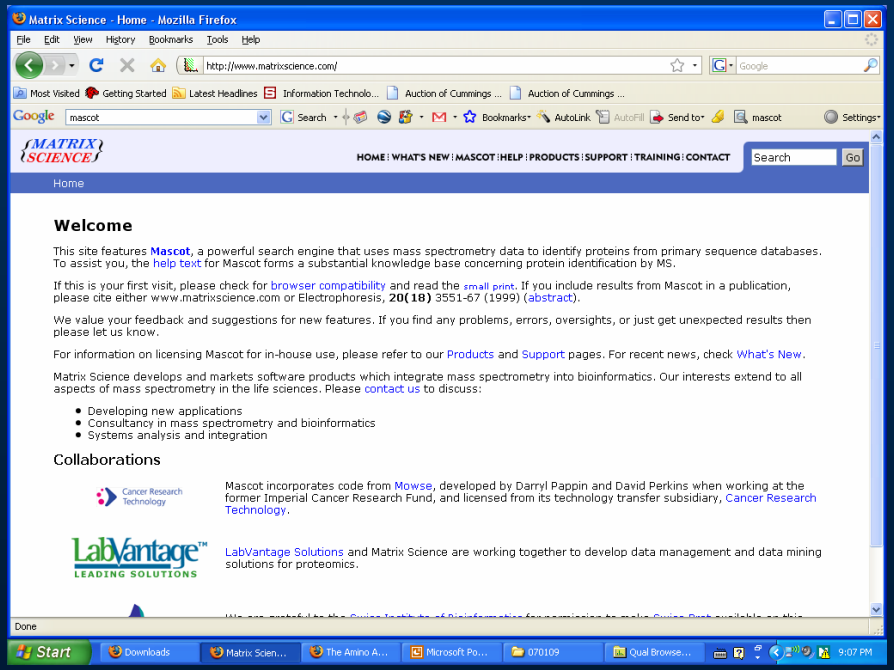
Deamidation N → D + 0.98 Da

N v. D Δ 1 Da

Q v. E Δ 1 Da

Database Searching

Mascot



Sequest

Jimmy K. Eng, Ashley L. McCormack, and John R. Yates, III (1994). "An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database".

J.Am Soc Mass Spectrom **5: 976–989**

Knowing taxonomy reduces search time dramatically

Push for Higher Mass Accuracy

With 1 ppm accuracy, elemental composition of a peptide up to 600-700 Da (6-7 residues) can be derived (Zubarev, RA Marshall, A)

Distinguishing K/Q (36.4 mDa) :

With 1 ppm accuracy, K/Q can be distinguished in peptides up to 3800 Da.

Identification of post-translational modifications :

Tryptic peptide, m = 997.514 +/- 0.001 Da (1 ppm accuracy)

Did not match any peptide sequences predicted for methionine-tRNA ligase

Assumption: it is a N-terminal peptide with Met removed and Ser$_1$ acetylated$_2$.

Testing: calculated mass m=997.5126 Da, agrees with the measured value

(Gibson, B.W.; Biemann, K. Proc. Natl. Acad. Sci. U.S.A. 81 (1984) 1956)

Data base search :

Tryptic peptide, m = 1513.794 +/- 0.001 (1 ppm accuracy)

Unique sequence, GAAFICAIHSPTLR, is found in the yeast genome.

Four sequences are found in a non-redundant database of 203 000 entries;

including TFHRIQQMLPDK with the same elemental composition.

With 1 ppm accuracy, tryptic peptides may be unique for a small genome

(Jensen, O.N.; Podtelejnikov, A.; Mann, M. Rapid Commun. Mass Spectrom. 10 (1996) 1371)

Acknowledgments

- SUMS Group
  - Allis Chien
  - Karolina Krasinska
  - Theresa McLaughlin
  - Pavel Aronov