Application of mass spectrometry to the analysis and identification of peptides proteins and other biological molecules

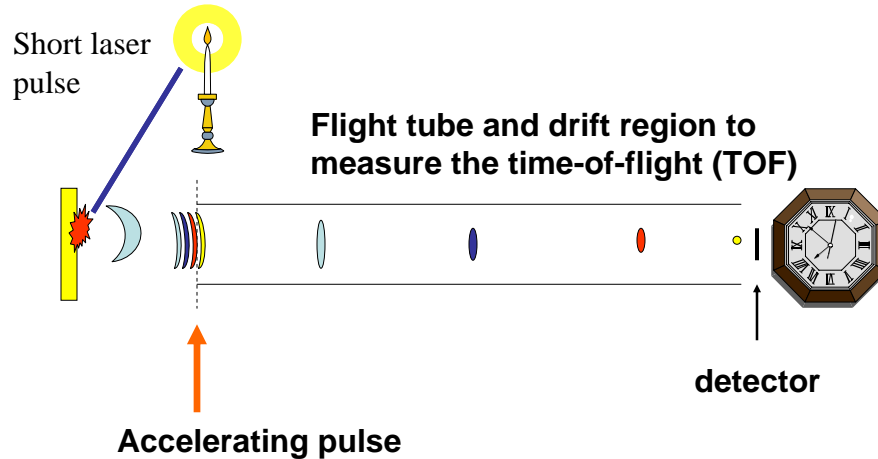## Stephen Barnes, PhD

4-7117

sbarnes@uab.edu

S. Barnes-UAB 1/12/07

Overview

- **MALDI-TOF MS**
  - Protein modifications
  - Peptide mass fingerprinting
- **Electrospray MS**
  - Analysis of intact proteins
  - Molecular weight calculations
  - Max Entropy for MW estimation
- **Peptide analysis**
  - Purity - ESI-MS is a revelation
- **Integration of MS with LC and CE**
  - Multidimensional LC of peptides
- **Tandem MS**
  - Identifying peptide amino acid sequences

S. Barnes-UAB 1/12/07

Matrix-Assisted Laser Desorption Ionization (MALDI)

Short laser pulse

Flight tube and drift region to measure the time-of-flight (TOF)

detector

Accelerating pulse

Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS)

- **Advantages of MALDI-TOF**
  - **More tolerant to common buffers than ESI**
  - **High degree of sensitivity, moderate mass accuracy, and mass resolution**
  - **High mass compounds, i.e. proteins, PEG…**

- **Common Applications of MALDI-TOF**
  - **Masses of large proteins and other compounds**
  - **Enzymatic digestion profiles of proteins to establish their identity**
 - **Peptide sequencing (TOF-TOF)**

Factors from conventional experiments that impact MALDI-TOF analysis

- **Tolerance of buffers/chemicals used in sample preparation**
  - **NaCl up to 150 mM**
  - **Urea up to 2-3 M (carbamoylation can occur!)**
  - **Guanidinium-HCl up to 2 M**
- **Tolerance of detergents**
  - **SDS up to 0.05%**
- **Staining Protocols**
  - **Whole proteins form adducts with Coomassie**
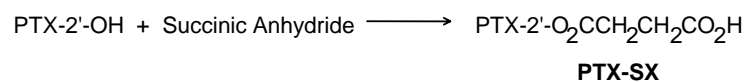 - **Silver staining modifies selected peptides**

S. Barnes-UAB 1/12/07

Benefit of removing salt from tryptic digest
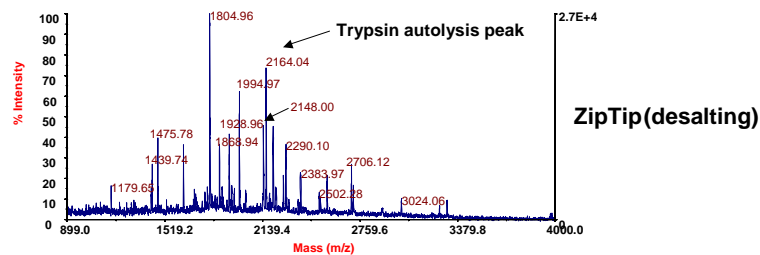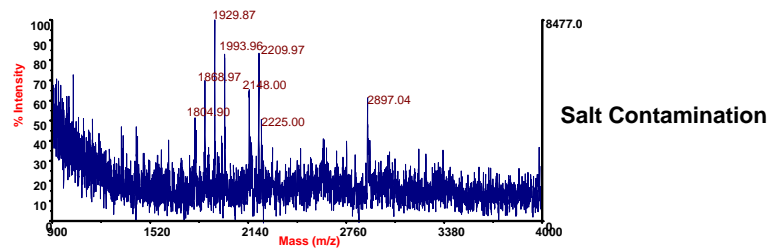
S. Barnes-UAB 1/12/07

Chemically modifying an antibody to bind radioactive metal ions

Scheme 1

PTX-2'-OH + Succinic Anhydride $\longrightarrow$ PTX-2'-O$_2$CCH$_2$CH$_2$CO$_2$H

PTX-SX

Scheme 2

MAb + PTX-SX $\xrightarrow[\text{DMF}]{\text{NHS/EEDQ}}$ NH-PTX

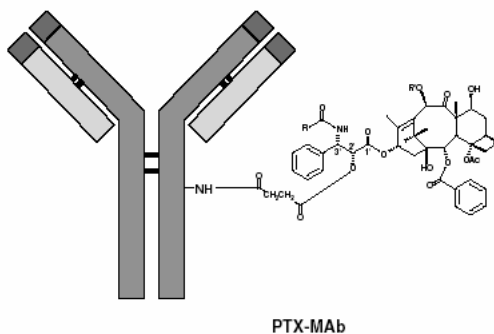**MAb**   **PTXMAb**

S. Barnes-UAB 1/12/07

Ahmad Safavy

Structure of modified antibody

PTX-MAb

S. Barnes-UAB 1/12/07

Ahmad Safavy

Modification of an antibody by MALDI-TOF

BSA

Unreacted Ab

A 76337

B 152223

$[M+2H]^{2+}$

$[M+H]^+$

% Intensity

Mass (m/z)

60000.0 87975.4 115950.8 143926.2 171901.6 199877.0

917.6

Reacted Ab

BSA

D 77246

C 155148

$[M+2H]^{2+}$

$[M+H]^+$

Δ = 2925 Da

% Intensity

Mass (m/z)

60000 87982 115964 143946 171928 199910
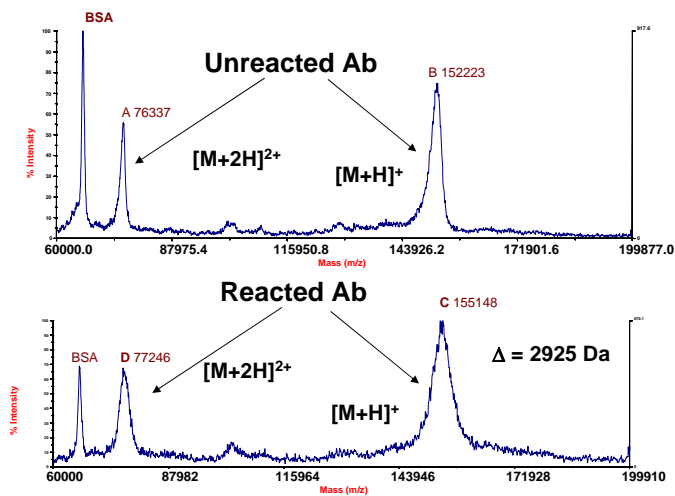
S. Barnes-UAB 1/12/07

Ahmad Safavy

A MALDI-TOF mass spectrum of peptides

% Intensity

100
90
80
70
60
50
40
30
20
10
0

1366.82
2452.17
8401.3
1494.90
1267.76
1388.78
1544.76
2164.05
2186.02
2517.29
2646.33
3141.45
3163.46

900 1520 2140 2760 3380 4000

Mass (m/z)
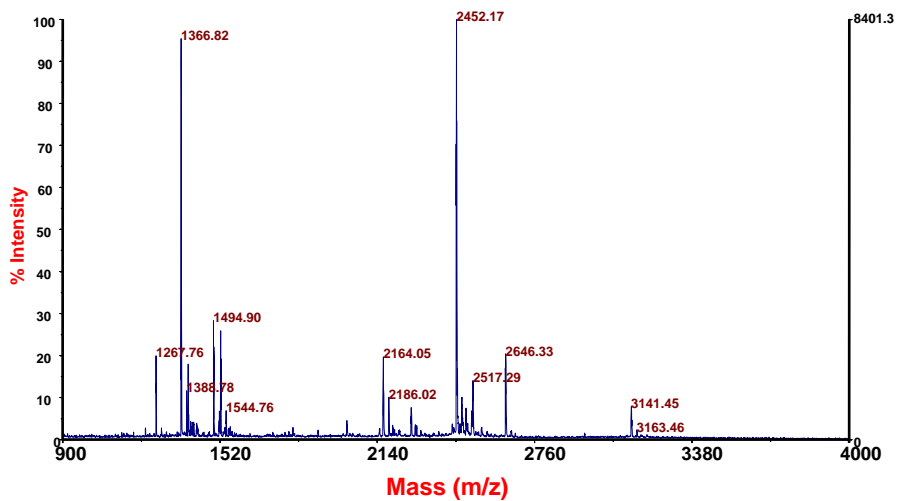
S. Barnes-UAB 1/12/07

Isotope profile of an individual peptide ion

S. Barnes-UAB 1/12/07
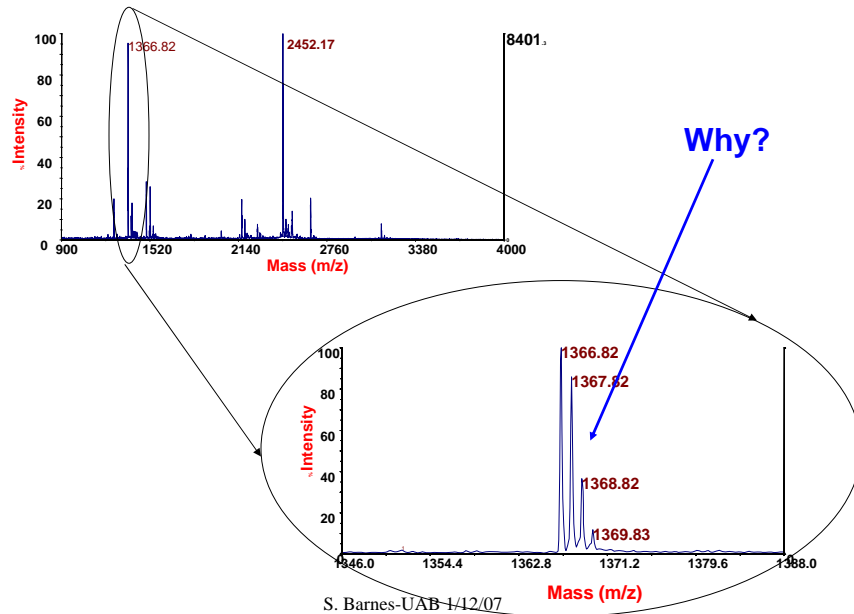
Stable isotopes of the most abundant elements found in peptides

Element	Mass	Abundance
H	1.0078	99.985%
	2.0141	0.015%
C	12.0000	99.89%*
	13.0034	1.11%*
N	14.0031	99.64%*
	15.0001	0.36%*
O	15.9949	99.76%*
	16.9991	0.04%*
	17.9992	0.20%*
S	31.9721	94.93%*
	32.9715	0.76%*
	33.9679	4.29%*
	35.9671	0.02%*

*Varies according to its source

S. Barnes-UAB 1/12/07

Isotope pattern for a larger peptide (207 C-atoms)

Two ^{12}C

4361.45
One ^{12}C

Three ^{12}C

Four ^{12}C

Five ^{12}C

4360.45
All ^{12}C

Six ^{12}C

Seven ^{12}C

m/z
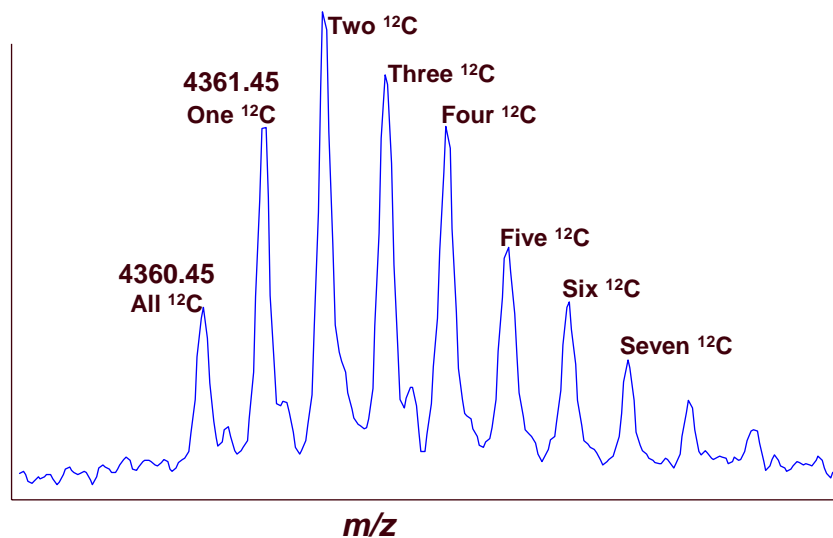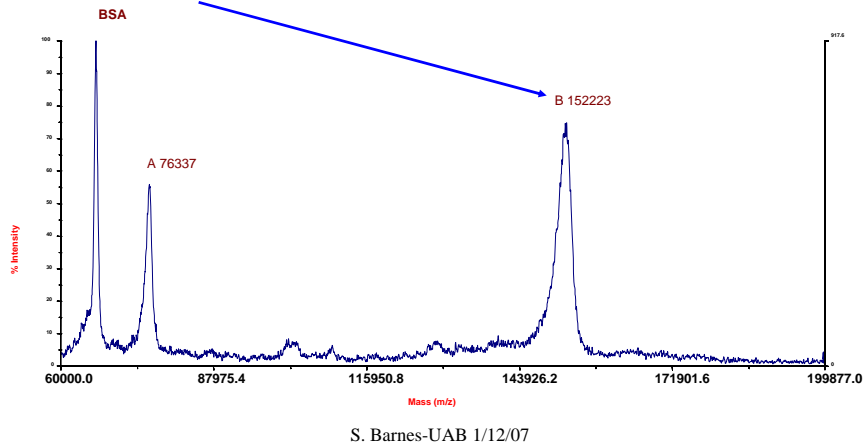
S. Barnes-UAB 1/12/07

How to represent the mass of compound?

- **At high resolution where the isotopic peaks are fully resolved, then we can determine the *monoisotopic mass* for each one**

- **At low mass resolution (where the isotope peaks cannot be resolved) what is observed is the *average mass***
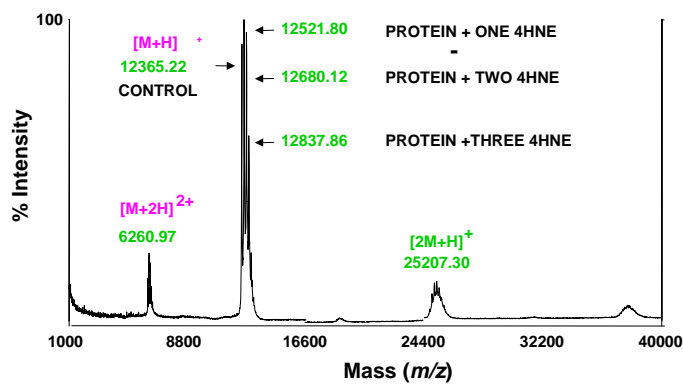
S. Barnes-UAB 1/12/07

Average mass

For MALDI-TOF spectra above *m/z* 5000, it becomes difficult/impossible to resolve the isotope profile. Instead, the average mass is calculated.

BSA

A 76337

B 152223

917.6

% Intensity

| 60000.0 | 87975.4 | 115950.8 | 143926.2 | 171901.6 | 199877.0 |

Mass (m/z)

S. Barnes-UAB 1/12/07

Modification of cytochrome C by the reactive aldehyde 4HNE

100

[M+H] $^+$
12365.22
CONTROL

12521.80 ← PROTEIN + ONE 4HNE
-
12680.12 ← PROTEIN + TWO 4HNE

12837.86 ← PROTEIN +THREE 4HNE

% Intensity

[M+2H] $^{2+}$
6260.97

[2M+H] $^+$
25207.30

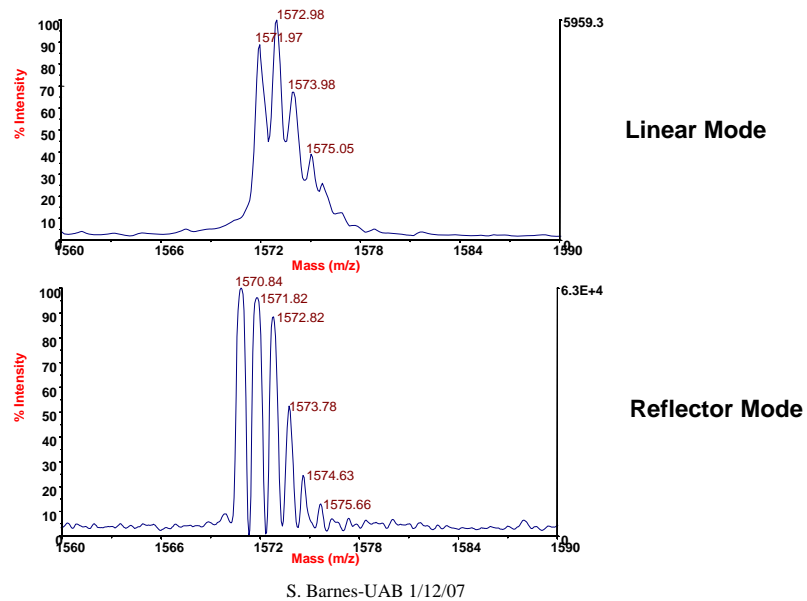| 1000 | 8800 | 16600 | 24400 | 32200 | 40000 |

Mass (*m/z*)

MALDI spectra usually contain only the **molecular ion [M+H]$^+$**. The mass accuracy is approximately $\pm$1-2 Da. Hard to pick out the ^{12}C isotope peak. The 4HNE Michael adduct should be 156 Da.

S. Barnes-UAB 1/12/07

Increased resolution and sensitivity in reflector vs. linear mode

Linear Mode

Reflector Mode

S. Barnes-UAB 1/12/07

Peptide mass fingerprinting

- **This method was developed because of the availability of predicted protein sequences from genome sequencing**

- **Proteins did not have to have been previously sequenced - only that the open reading frame in the gene is known - the rest is a virtual exercise in the hands of statisticians, bioinformaticists and computers**

- **However, remember the matching is only as good as the database content - this can change**

S. Barnes-UAB 1/12/07

9

Protein analysis by MALDI 2007

S. Barnes-UAB 1/12/07
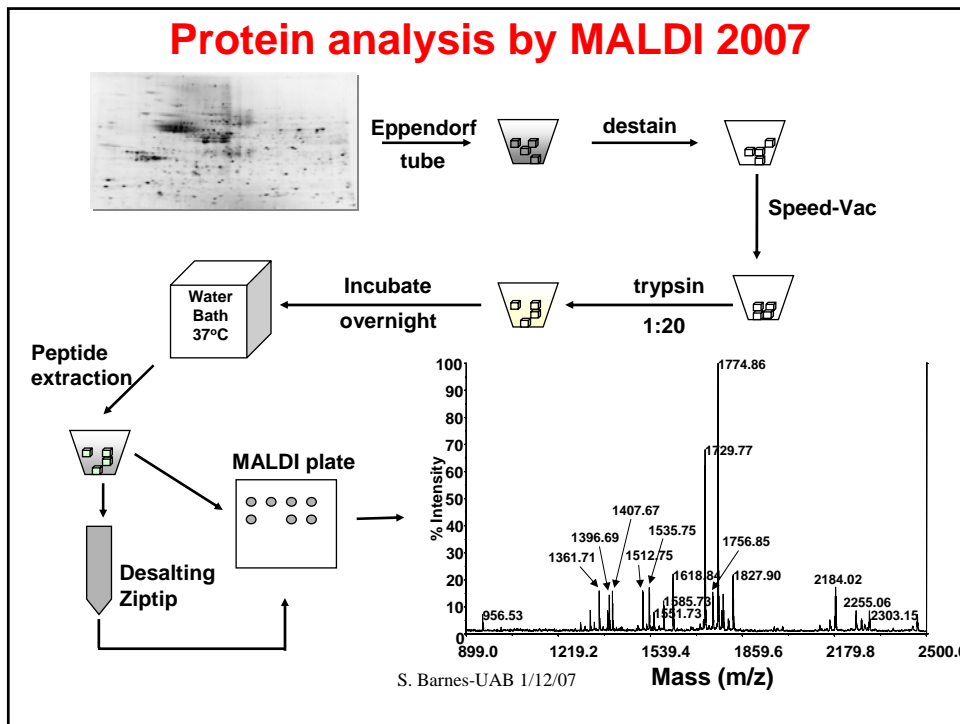
Proteolytic enzymes used to hydrolyze proteins

The choice of enzyme largely depends on the nature of the amino acid sequence and the specific issue that is being addressed

- **Trypsin -** *cleaves at arginine and lysine residues*

- **Chymotrypsin -** *cleaves hydrophobic residues*

- **Arg-C** *- cleaves at arginine residues*

- **Glu-C** *- cleaves at aspartate/glutamic acid residues*

- **Lys-C** *- cleaves at lysine residues*

- **V8-protease** *- cleaves at glutamic acid residues*

- **Pepsin** *- cleaves randomly, but at acid pH*

See **http://www.abrf.org/JBT/1998/September98/sep98m_r.html**

S. Barnes-UAB 1/12/07

Searching databases with peptide masses to identify proteins

Best site is at www.matrixscience.com

The program (MASCOT) can search the OWL or NCBI databases using a set of tryptic peptide masses, or the fragment ions (specified or unspecified) of peptides

Presents the expected set of tryptic peptides for each matched protein

Choice of peptidase

- **Analogous to DNA restriction enzymes**

- **Tryptic peptide fingerprinting may identify several highly related protein candidates (e.g., actins)**

- **Inspection of the sequences may reveal that there is a difference at one residue that distinguishes between two candidates.**

- **If for instance it is a glutamate, then use of Glu-C or V8-protease may enable the two proteins to be correctly identified**

- **INSPECT sequences carefully**

Sequence of β-lactoglobulin

```
MKCLLLALAL TCGAQALIVT QTMKGLDIQK
VAGTWYSLAM AASDISLLDA QSAPLRVYVE
ELKPTPEGDL EILLQKWENG ECAQKKIIAE
KTKIPAVFKI DALNENKVLV LDTDYKKYLL
FCMENSAEPE QSLACQCLVR TPEVDDEALE
KFDKALKALP MHIRLSFNPT QLEEQCHI
```

S. Barnes-UAB 1/12/07

GLDIQK   CLLLALALTCGAQALIVTQTMK   VYVEELK

M   VAGTWYSLAMAASDISLLDAQSA   TK
K   PLR

ALK   YLLFCMENSAEPEQSLACQC   K   K
          LVR

WENGECAQK   PTPEGDLEILLQK   VLVLDTDY
                                   K

FDK   IPAVFK   ALPMHIR   IIAEK

TPEVDDEALEK   IDALNEN   LSFNPTQLEEQCHI
                K

S. Barnes-UAB 1/12/07

Peptides from digestion with Glu-C

MKCLLLALALTCGAQALIVTQTMK
IQKVAGTWYSLAMAASD GLD ISLLD AQSAPLRVYVE

E LKPTPE GD LE ILLQKWE NGE CAQKKIIAE

KTKIPAVFKID ALNE NKVLVLD TD YKKYLLFCME

NSAE PE QSLACQCLVRTPEVD D E ALE KFD

KALKALPMHIRLSFNPTQLEE QCHI

Amino acid residue masses

Amino acid	Mass	Amino acid	Mass
Alanine	71.037	Leucine	113.084
Arginine	156.101	Lysine	128.094
Asparagine	114.043	Methionine	131.040
Aspartic acid	115.027	Phenylalanine	147.068
Cysteine	103.009	Proline	97.053
Glutamic acid	129.043	Serine	87.032
Glutamine	128.058	Threonine	101.048
Glycine	57.021	Tryptophan	186.079
Histidine	137.059	Tyrosine	163.063
Isoleucine	113.084	Valine	99.068

The m/z value of a peptide [M+H]$^+$ is the sum of the residue masses plus 18.015 for H_2O plus 1.008. So, what is it for ISLLD?

113.084 + 87.032 + 113.084 +113.084 + 115.027 + 18.015 + 1.008 = 560.334

13

Expected peptides from trypsin and Glu-C digestion of bovine β-lactoglobulin

837.4764	800.4876
916.4734	929.5455
1064.4466	1003.5605
1065.5827	1232.6634
1245.5845	1259.7722
1658.7843	1337.6632
2275.2586	1447.7032
2313.2588	1811.8996
2647.2023	2307.3006
2707.3760	2819.5265

Assumes all cuts are complete, there is no oxidation of Met residues, and Cys residues are unmodified

S. Barnes-UAB 1/12/07


MASCOT Peptide Mass Fingerprint

Your name	Stephen Barnes	**Email**	sbarnes.uab@gmail.com

Search title beta globulin test

Database NCBInr

Taxonomy All entries

Enzyme Trypsin   **Allow up to** 0 missed cleavages

Fixed modifications
Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (Protein C-term)
Amidated (C-term)

Variable modifications
Acetyl (K)
Acetyl (N-term)
Acetyl (Protein N-term)
Amidated (Protein C-term)
Amidated (C-term)

Protein mass ____ kDa   **Peptide tol. ±** 1.0 Da

Mass values ⦿ MH⁺ ◯ Mr ◯ M-H⁻   **Monoisotopic** ⦿ **Average** ◯

Data file Choose File no file selected

Query
NB Contents of this field are ignored if a data file is specified.
837.4764
916.4734
1064.4466
1065.5827
1245.5845
1658.7843

Enter the ions here in this box

Overview ☐   **Report top** 20 hits

Start Search ...   Reset Form

Mascot Search Results

{MATRIX SCIENCE}

```
User          : Stephen Barnes
Email         : sbarnes.uab@gmail.com
Search title  : beta globulin test
Database      : NCBInr 20061230 (4378862 sequences; 1508892933 residues)
Timestamp     : 1 Jan 2007 at 02:17:51 GMT
Top Score     : 210 for gi|87196497, lactoglobulin, beta [Bos taurus]
```

Probability Based Mowse Score

Protein score is $-10*Log(P)$, where P is the probability that the observed match is a random event.
Protein scores greater than 79 are significant ($p<0.05$).
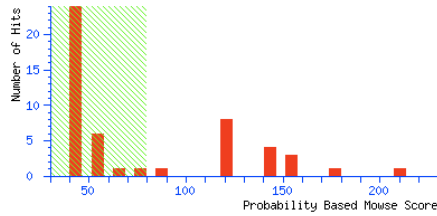
Protein Summary Report

Format As	Protein Summary ▾	Help

Significance threshold p< 0.05   Max. number of hits 20

Protein records provided by MASCOT search

```
    Accession    Mass   Score  Description
 1. gi|87196497  19870  210    lactoglobulin, beta [Bos taurus]
 2. gi|4388846   18269  179    Chain , Bovine Beta-Lactoglobulin Complexed With Palmitate, Lattice Z
 3. gi|223780    18165  152    lactoglobulin beta
 4. gi|72079     18255  152    beta-lactoglobulin - water buffalo
 5. gi|520       19908  150    beta-lactoglobulin [Bos taurus]
 6. gi|20178290  20010  148    Beta-lactoglobulin precursor (Beta-LG)
 7. gi|165839    19934  148    beta-lactoglobulin
 8. gi|2194088   18297  147    Chain A, Bovine Beta-Lactoglobulin, Lattice X
 9. gi|110612608 19891  144    beta-lactoglobulin [Bubalus bubalis]
10. gi|162748    17156  126    beta-lactoglobulin
11. gi|125912    19962  125    Beta-lactoglobulin precursor (Beta-LG)
12. gi|7245834   18363  124    Chain A, Structural Changes Accompanying Ph-Induced Dissociation Of The
13. gi|229460    18355  124    lactoglobulin beta
14. gi|4388939   18355  124    Chain , Structural Basis Of The Tanford Transition Of Bovine Beta-Lacto
15. gi|49259423  18339  124    Chain X, The Cys121ser Mutant Of Beta-Lactoglobulin
16. gi|54037712  18139  120    Beta-lactoglobulin (Beta-LG)
17. gi|57164367  19908  117    beta-lactoglobulin [Ovis aries]
18. gi|90108547  18264   82    Chain A, Reindeer Beta-Lactoglobulin
19. gi|71980384  20035   80    beta-lactoglobulin [Rangifer tarandus tarandus]
20. gi|26352113  13020   60    unnamed protein product [Mus musculus]
```

S. Barnes-UAB 1/12/07

15

Comparison of observed and predicted tryptic peptides

```
gi|87196497        Mass: 19870     Score: 210     Expect: 4.4e-15  Queries matched: 10
lactoglobulin, beta [Bos taurus]
 Observed    Mr(expt)    Mr(calc)    Delta Start    End Miss Peptide
  837.4764   836.4691    836.4691   0.0001   158 - 164    0   K.ALPMHIR.L
  916.4734   915.4661    915.4661  -0.0000   100 - 107    0   K.IDALNENK.V
 1064.4466  1063.4393   1063.4393   0.0001    77 -  85    0   K.WENGECAQK.K
 1065.5827  1064.5754   1064.5753   0.0001   108 - 116    0   K.VLVLDTDYK.K
 1245.5845  1244.5772   1244.5772   0.0000   141 - 151    0   R.TPEVDDEALEK.F
 1658.7843  1657.7770   1657.7770   0.0000   165 - 178    0   R.LSFNPTQLEEQCHI.-
 2275.2586  2274.2513   2274.2513   0.0000     3 -  24    0   K.CLLLALALTCGAQALIVTQTMK.G
 2313.2588  2312.2515   2312.2515   0.0001    57 -  76    0   R.VYVEELKPTPEGDLEILLQK.W
 2647.2023  2646.1950   2646.1950   0.0001   118 - 140    0   K.YLLFCMENSAEPEQSLACQCLVR.T
 2707.3760  2706.3687   2706.3686   0.0001    31 -  56    0   K.VAGTWYSLAMAASDISLLDAQSAPLR.V
```

S. Barnes-UAB 1/12/07

Search against SwissProt database

```
     Accession   Mass   Score   Description
 1.  LACB_BOVIN  19870   271    Beta-lactoglobulin precursor (Beta-LG) (Allergen Bos d 5) - Bos taurus (Bovine)
 2.  LACB_BUBBU  20010   197    Beta-lactoglobulin precursor (Beta-LG) - Bubalus bubalis (Domestic water buffalo)
 3.  LACB_CAPHI  19962   135    Beta-lactoglobulin precursor (Beta-LG) - Capra hircus (Goat)
 4.  LACB_OVIMU  18139   104    Beta-lactoglobulin (Beta-LG) - Ovis orientalis musimon (Mouflon)
 5.  LACB_SHEEP  19908   102    Beta-lactoglobulin-1/B precursor (Beta-LG) - Ovis aries (Sheep)
 6.  YKH5_YEAST  52055    41    Hypothetical 52.1 kDa protein in SMY1-MUD2 intergenic region - Saccharomyces cerevi
 7.  TCPH_VIBCH  16247    40    Toxin coregulated pilus biosynthesis protein H (TCP pilus biosynthesis protein tcpH
 8.  POLG_DEN22  54291    39    Genome polyprotein [Contains: Envelope protein E] (Fragment) - Dengue virus type 2
 9.  PFPI_PYRKO  18404    38    Intracellular protease 1 (EC 3.2.-.-) (Intracellular protease I) - Pyrococcus kodak
10.  VIRB9_AGRT5 32181    38    Protein virB9 precursor - Agrobacterium tumefaciens (strain C58 / ATCC 33970)
11.  VPS71_YEAST 32001    36    Vacuolar protein sorting-associated protein 71 (SWR complex protein 6) - Saccharomy
12.  YHR7_YEAST  71812    34    TPR repeat-containing protein YHR117W - Saccharomyces cerevisiae (Baker's yeast)
13.  LGB_LOTJA   15745    34    Leghemoglobin - Lotus japonicus
14.  YH13_VACCV  13615    34    Hypothetical 13.6 kDa HindIII-C protein - Vaccinia virus (strain Western Reserve /
15.  MKT1_YEAST  94435    34    Protein MKT1 - Saccharomyces cerevisiae (Baker's yeast)
16.  H2A1_LILLO  12165    34    Histone H2A.1 (GcH2A) - Lilium longiflorum (Trumpet lily)
17.  RS12_BRUA2  13863    33    30S ribosomal protein S12 - Brucella abortus (strain 2308)
18.  RS12_BRUAB  13863    33    30S ribosomal protein S12 - Brucella abortus
19.  RS12_BRUME  13863    33    30S ribosomal protein S12 - Brucella melitensis
20.  RS12_BRUSU  13863    33    30S ribosomal protein S12 - Brucella suis
```

S. Barnes-UAB 1/12/07

16

Things to consider when doing peptide mass fingerprinting

- **Proteins can be oxidized both biologically (real data) and during the workup**
- **Treat the protein or the peptide digest with a reagent that reacts with Cys sulhydryl groups - e.g., iodoacetamide, iodoacetic acid, N-ethylmaleimide or 4-vinylpyridine. These groups may also have reacted with acrylamide in the gel.**
- **Set the options in the fixed or variable modification boxes before searching**
- **Allow for at least one missed cleavage - trypsin does not cut when Lys or Arg are followed by a Pro residue**

S. Barnes-UAB 1/12/07

Other web sites for peptide analysis

- **http://prowl.rockefeller.edu/**
  - **Choose ProFound**

- **http://prospector.ucsf.edu/**
  - **Choose MS-fit**

S. Barnes-UAB 1/12/07

Further information on identified protein

- **Take the protein identifier number:**
  - **For bovine β-lactoglobulin it is gi|520**
  - **Go to http://www.ncbi.nlm.nih.gov**
  - **Under Entrez, paste in the gi number**
  - **A link to the protein will appear**
  - **Click on Blink - this is similar to BLAST, but better**
  - **Select 3D-structures on this page to get Protein Data Base record(s) of crystal structure data of the nearest protein - this will yield 1CJ5**
  - **Go to Structure (top of web page) and enter 1CJ5 and click on its icon on the next page**
  - **To view a 3D-image of the protein, first download Cn3D from the NCBI site**

S. Barnes-UAB 1/12/07

Examples for homework (due Jan 23)

- **Identify the following proteins from these MALDI ions (corrected for isotope effects):**
  - **818.44, 1071.55, 1252.71, 1529.73, 1833.89, 2967.61, 2996.49 (human)**
  - **899.43, 940.43, 1048.56, 1254.62, 1316.64, 1570.76, 1830.05, 2427.28 (rat)**
  - **847.50, 1190.53, 1396.68, 1803.04, 1965.91, 2047.89, 2871.32 (mouse)**

- **Set the number of tryptic cuts to 0 and try varying the mass accuracy from 0.02 to 1.0 Da. How does this alter the MOWSE score?**

S. Barnes-UAB 1/12/07

Take home question (due Jan 23)

1. **What is the monoisotopic mass of bovine β-lactoglobulin alpha chain?**
   – **Hint: workout the empirical formula of bovine β-lactoglobulin - its sequence can be obtained from record P02754 at http://www.ExPasy.org**

2. **What is the molecular weight of the most abundant species of β-lactoglobulin ?**
   – **Hint: assume that the abundance of ^{13}C is 1.11% of total carbon atoms - do not worry about ^2H or other isotopes**

S. Barnes-UAB 1/12/07

Blurring of protein space

- **Identification using MALDI-TOF with MASCOT depends on:**
  - Number of peptides recognized as being part of the protein
  - The mass accuracy of the peptides that are recognized
  - Pre-2000, an accuracy of better than 0.05 Da in a 1000 Da peptide (i.e., 50 ppm) was sufficient to distinguish the unknown protein from the other proteins in the databases <u>at that time</u>
  - Now, the protein information space has become more dense and MALDI-TOF is no longer adequate
 - Previously identified proteins may not be correct
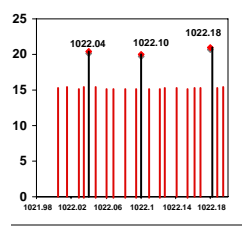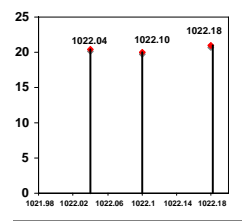
Electrospray ionization

- **ESI-MS is very sensitive to the presence of electrolyte species -**
    - **these ionize more easily than solutes and may also form adducts with solutes**
- **In ESI-MS, multiple charge states are possible**
    - **These lead to more accurate MWs**
- **This is a softer ionization than MALDI where the UV laser at 337 nm alters the chemistry of modifications such as Tyr-NO$_2$ and Cys-SNO**

ESI-MS and purity of peptides
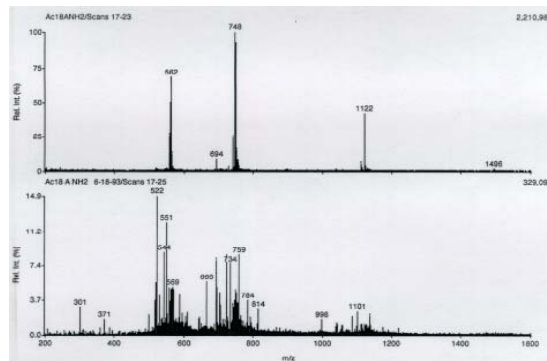
Guarantees of purity based on observation of "a single peak by reverse-phase HPLC" and by "it gave the correct sequence when analyzed by Edman degradation" are hollow.The lower spectrum was of a "pure" HPLC peak. The method of purification was amended and the upper spectrum was obtained

Ionizing proteins and peptides

- $^+H_3NCHR_1CO(NHCHR_nCO)_nNHCHR_2COOH$ is the ion that's found in dilute acid solution

- If there are internal basic residues, then the ions will be of the form $[M+nH]^{n+}$, where n = 1, 2, etc.

- A tryptic peptide will have a N-terminal amino group and an amino group from Arg or Lys
    - If the peptide has a mol. wt. of 1000 Da, then the singly charged ion will have a *m/z* of 1001, whereas the doubly charged ion has a *m/z* of 501

S. Barnes-UAB 1/12/07

ESI mass spectrum of ribonuclease
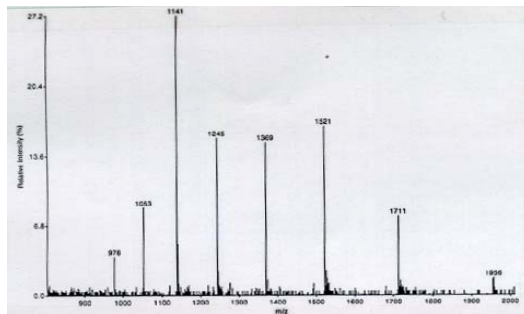
Cumulative MW estimate = 13,680.29

SD = 2.94

Peak (m/z)	Intensity	Charge (est.)	Mol. Wt. (Est.)
978.00	7,778	14.00000	13,677.89
1,053.00	18,532	13.02656	13,675.90
1,141.00	59,087	11.95446	13,679.91
1,245.00	33,275	10.96146	13,683.91
1,369.00	32,390	10.03219	13,679.92
1,521.00	35,668	8.99995	13,679.93
1,711.00	16,624	7.99996	13,679.94
1,956.00	3,333	6.97955	13,684.94

S. Barnes-UAB 1/12/07

Calculation of molecular weights and ion states
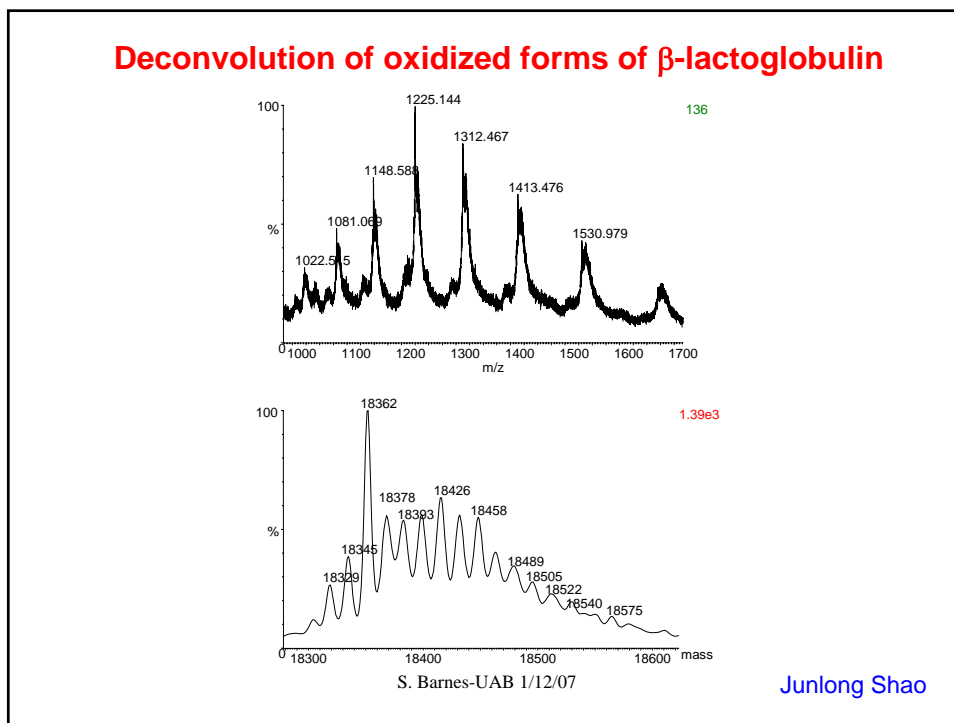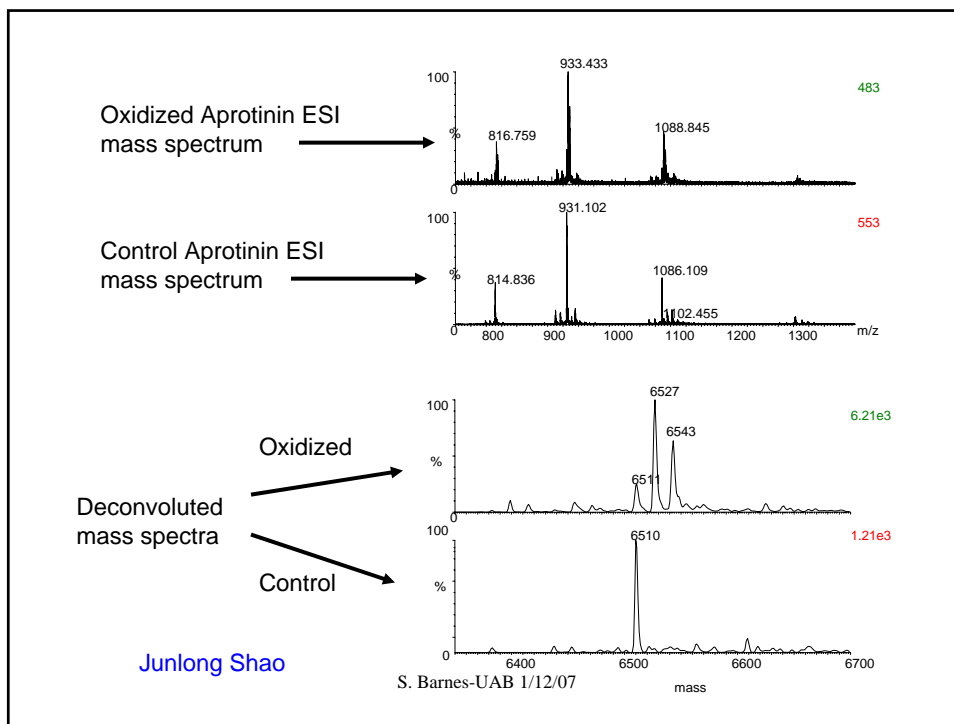
- For two ions in a series for a peptide of molecular weight M, the lower *m/z* value (x) will be for the n+1 ion state and the larger *m/z* value (y) will be for the n+ ion state.
  - (1) $(M+n)/n = y$
  - (2) $(M+n+1)/(n+1) = x$
- Hence
  - (3) $M+n = ny$ and $M = ny-n$
  - (4) $M+n+1 = (n+1)x$ and $M = (n+1)x-(n+1)$
- Hence
  - $ny-n = (n+1)x - (n+1)$
  - $ny-n-xn+n = x-1$
  - $n(y-x) = x-1$
  - $n = (x-1)/(y-x)$
- The value of n can then be substituted in equation (1) to obtain the molecular weight of the peptide

S. Barnes-UAB 1/12/07

Deconvolution of MS data

- **When several proteins are present, then their multiply charged ion clusters overlap**

- **Can this be overcome? - yes, use the MaxEntropy program provided by Micromass**

S. Barnes-UAB 1/12/07

Oxidized Aprotinin ESI mass spectrum

100
%
816.759
933.433
1088.845
483
0

Control Aprotinin ESI mass spectrum

100
%
814.836
931.102
1086.109
1102.455
553
0
800 900 1000 1100 1200 1300 m/z

Deconvoluted mass spectra

Oxidized

100
%
6511
6527
6543
6.21e3
0

Control

100
%
6510
1.21e3
0
6400 6500 6600 6700 mass

Junlong Shao

S. Barnes-UAB 1/12/07

Deconvolution of oxidized forms of β-lactoglobulin

100
%
1022.545
1081.069
1148.588
1225.144
1312.467
1413.476
1530.979
136
0
1000 1100 1200 1300 1400 1500 1600 1700 m/z

100
%
18329
18345
18362
18378
18393
18426
18458
18489
18505
18522
18540
18575
1.39e3
0
18300 18400 18500 18600 mass

S. Barnes-UAB 1/12/07

Junlong Shao

23

ESI spectrum of bacterially expressed protein

A52
1101.1

A56/B49
1022.47

A48/B42
1192.70

C18
843.12

A45
1272.08

A41
1396.16

TOF MS ES+
552
A: 57195.23±12.89
B: 50011.25±18.34
C: 15157.60±0.94

Each ion is $[M+nH]^{n+}$

For 50+ charge state of a 50 kDa protein,

$m/z = [50,000 + 50]/50$
$= 1,001$

S. Barnes-UAB 1/12/07

Courtesy of Mindan Sfakianos
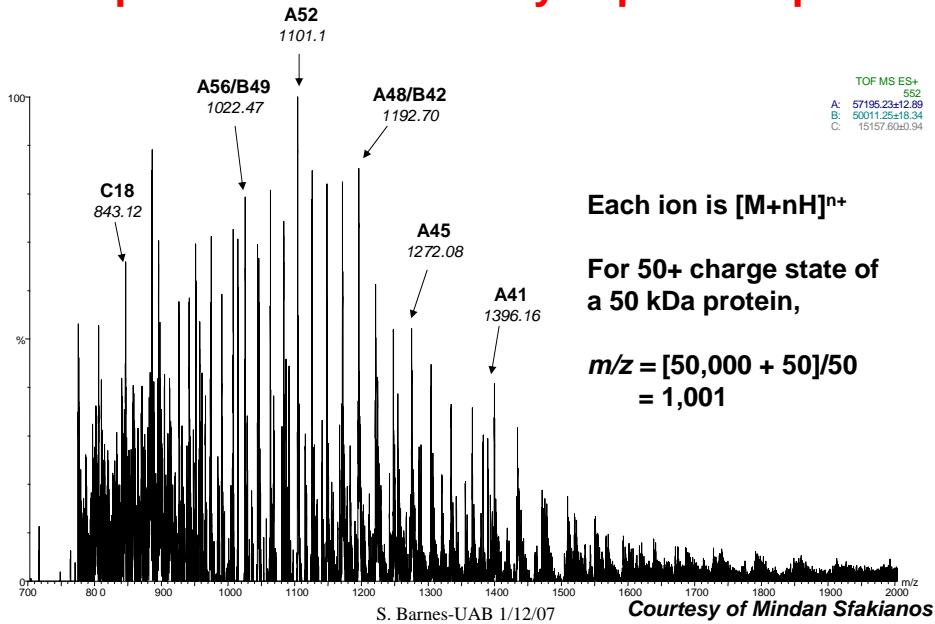
MaxEnt deconvolution of MWs

57,195.2 ± 12.9 Da
E. coli GRoEL

50,011.3 ± 18.3 Da
6xHis-tag BAT

15,157.6 ± 0.9 Da

S. Barnes-UAB 1/12/07

Courtesy of Mindan Sfakianos
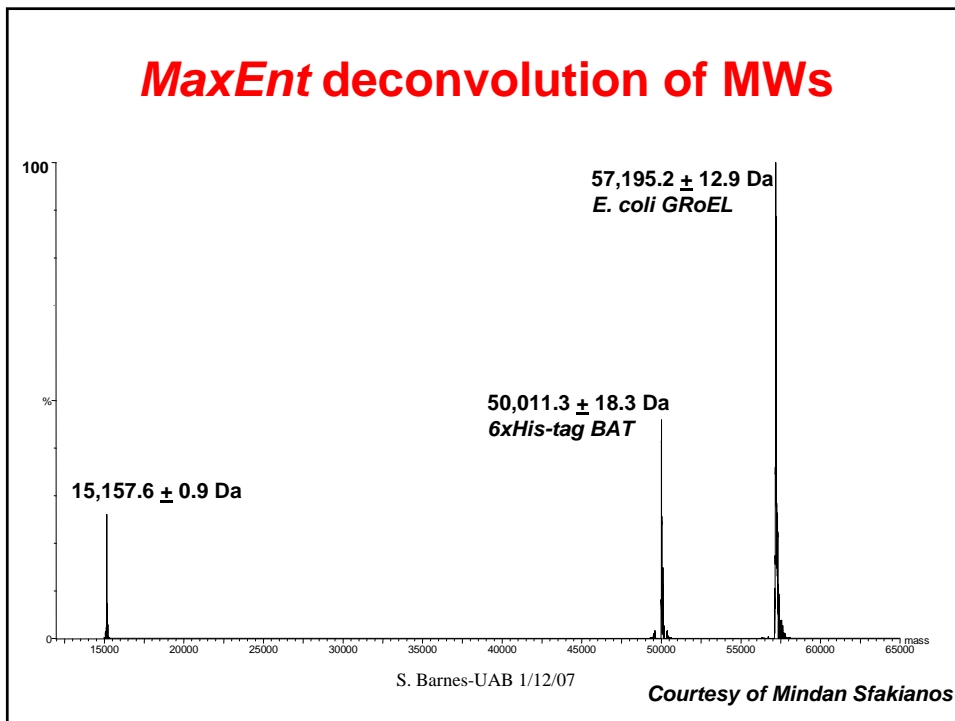
ESI-MS of 4HNE-Modified Cytochrome C

S. Barnes-UAB 1/12/07   *Courtesy of Amanda Isom (d. 2005)*
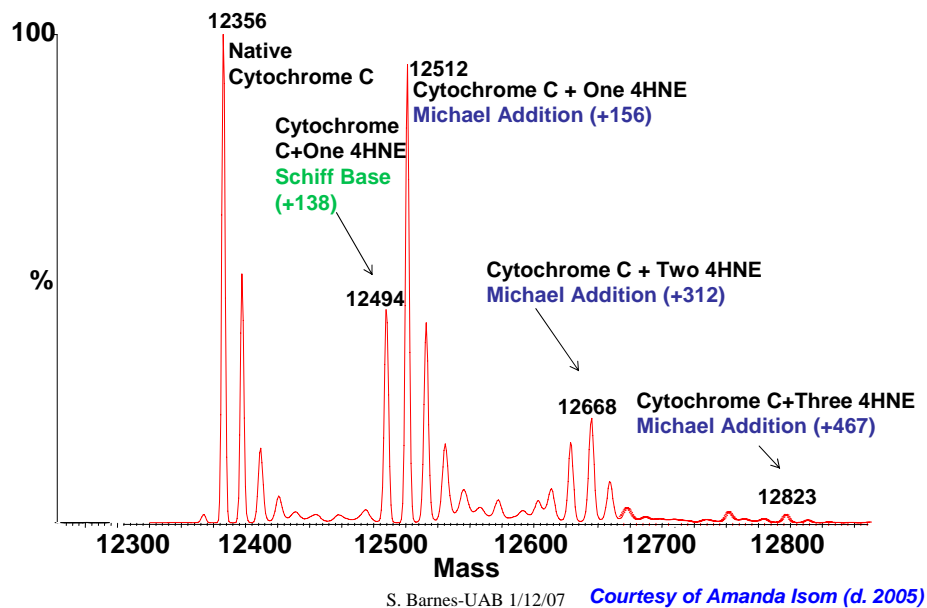
Summary of determining MW by ESI

- **The multiple charge states of a protein allow:**
    - **Mol Wt of large proteins to be estimated**
    - **accurate estimation of mol wt (super SDS-PAGE gel)**
- **Important to remember that the protein sample must be free of salt**
    - **Typically, a sample is cleaned up on a short reverse-phase column prior to electrospray**
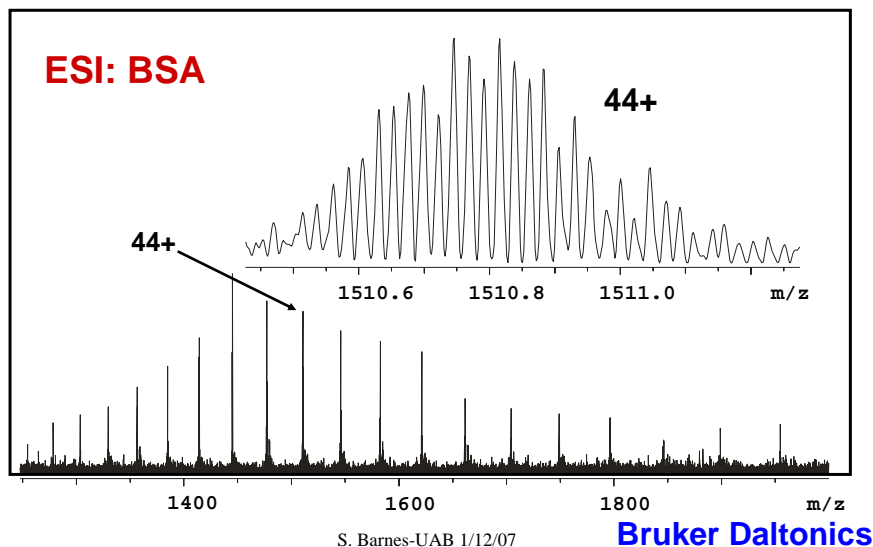   - **Alternative, use ammonium acetate as buffer**

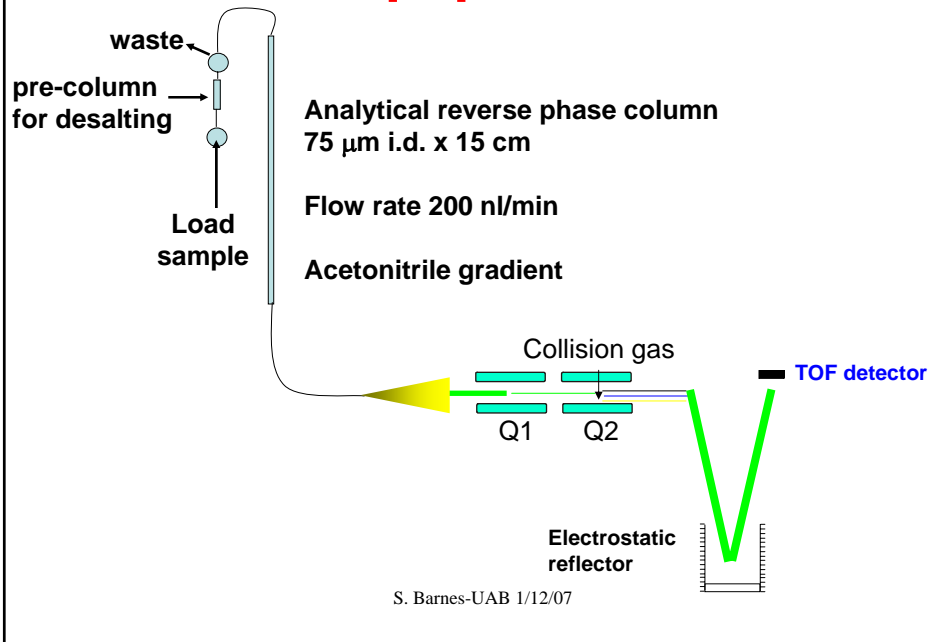S. Barnes-UAB 1/12/07

Use of FT-MS in ESI of proteins

- **The very high resolving power of FT-MS enables a direct measure of charge state of an individual ion since each peptide or polypeptide will have several/many isotope peaks**

- **The distance in Da between successive isotope peaks of a multiply charged ion is the reciprocal of the number of charges**
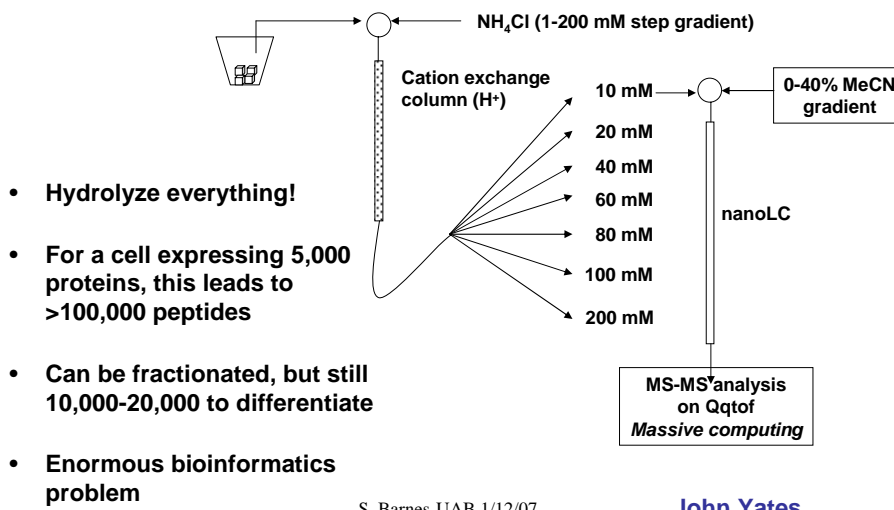
S. Barnes-UAB 1/12/07

Bovine Serum Albumin (66 kDa)
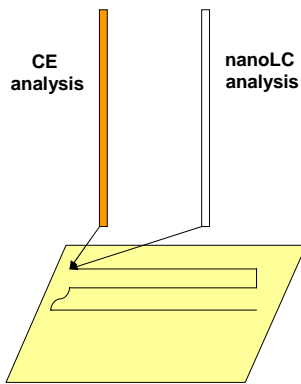# 4.7 T Actively Shielded Magnet

ESI: BSA

44+

44+

1510.6   1510.8   1511.0   m/z

1400   1600   1800   m/z

S. Barnes-UAB 1/12/07

Bruker Daltonics

LC-MS of peptide mixtures

waste

pre-column
for desalting

Analytical reverse phase column
75 μm i.d. x 15 cm

Flow rate 200 nl/min

**Load
sample**

Acetonitrile gradient

Collision gas

TOF detector

Q1   Q2

**Electrostatic
reflector**

S. Barnes-UAB 1/12/07

MUDPIT - MUlti-Dimensional
Protein Identification Technology

NH$_4$Cl (1-200 mM step gradient)

**Cation exchange
column (H$^+$)**

10 mM

0-40% MeCN
gradient

20 mM

- **Hydrolyze everything!**

40 mM

60 mM

nanoLC

- **For a cell expressing 5,000
  proteins, this leads to
  >100,000 peptides**

80 mM

100 mM

- **Can be fractionated, but still
  10,000-20,000 to differentiate**

200 mM

- **Enormous bioinformatics
  problem**

**MS-MS analysis
on Qqtof**
Massive computing

S. Barnes-UAB 1/12/07

John Yates

Connecting CE and LC to MALDI analysis

CE analysis

nanoLC analysis

Creates 20 mm wide tracks that can be scanned by MALDI laser for MS analysis
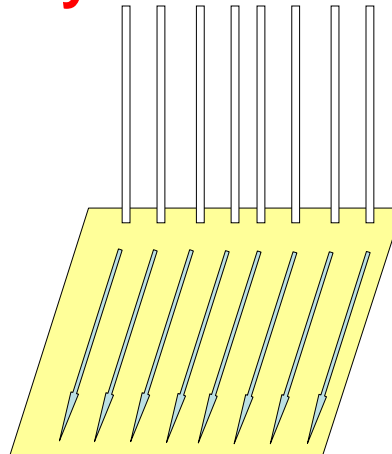
Parallel capture of effluents of 8 nanoLC separations on Mylar - can be scanned simultaneously by fast laser

S. Barnes-UAB 1/12/07

Pros/Cons of laying down LC or EC separations on matrix plate

- **Allows off-line analysis both in real time and then in a retrospective mode**

- **MALDI-TOF analysis is very fast**

- **Can also do TOF-TOF MS-MS analysis**

- **BUT what happens chemically on the acidic environment on the surface of the plate during storage?**

- **Also, can the laser beam cause chemical changes?**

S. Barnes-UAB 1/12/07

Sequencing of peptides

- **Using tandem mass spectrometry in a triple quadrupole, Q-tof, or ion trap instrument, the parent ion is first selected in the first quadrupole**

- **The parent ion is collided with argon gas and it breaks into fragments (daughter ions)**

- **By identifying the daughter ions, the peptide amino acid sequence is inferred**
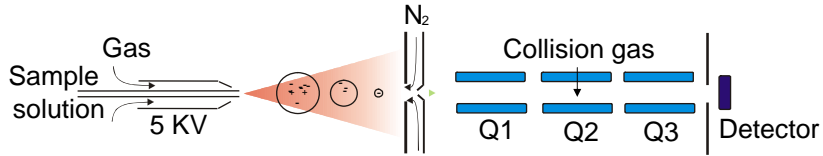
S. Barnes-UAB 1/12/07

# Tandem mass spectrometry on a triple quadrupole instrument
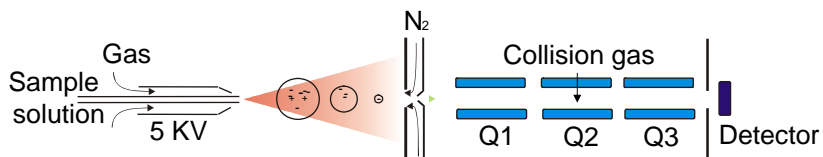
- **Daughter ion spectra**
- **Parent ion spectra**
- **Multiple reaction ion scanning**

S. Barnes-UAB 1/12/07

Triple quad versus Q-tof and sensitivity

Gas

Sample solution

5 KV

N₂

Collision gas

Q1   Q2   Q3   Detector

The quadrupole analyzer (Q3) is slow and insensitive - it's a filter - thus throws away large amounts of data

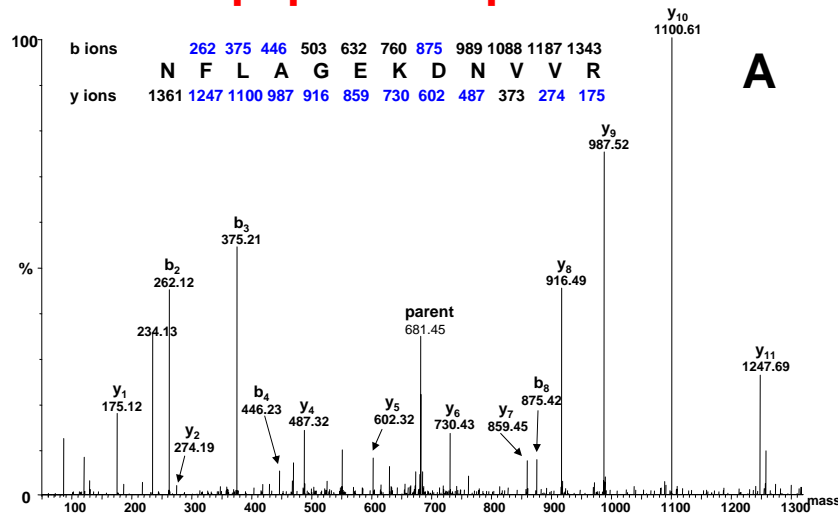Collision gas

Q1   Q2
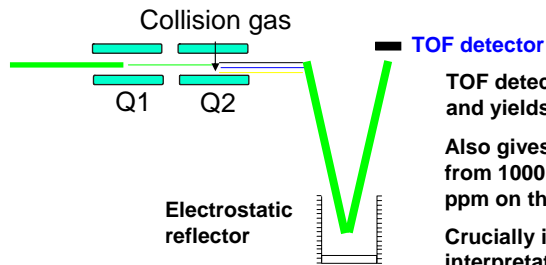
TOF detector

TOF detector collects all ions generated and yields fmol rather than pmol sensitivity

Also gives far greater mass accuracy - from 1000 ppm on the triple quad to <20 ppm on the Q-tof

Crucially important for automated interpretation of MS-MS spectra to yield amino acid sequence

Electrostatic reflector

S. Barnes-UAB 1/12/07

Identification of daughter ions and peptide sequence

b ions	262 375 446 503 632 760 875 989 1088 1187 1343
	N F L A G E K D N V V R
y ions	1361 1247 1100 987 916 859 730 602 487 373 274 175

A

100

%

0

y_{10} 1100.61

y_9 987.52

b_3 375.21

b_2 262.12

234.13

y_8 916.49

parent 681.45

y_{11} 1247.69

y_1 175.12

b_4 446.23

y_4 487.32

y_5 602.32

y_6 730.43

y_7 859.45

b_8 875.42

y_2 274.19

100 200 300 400 500 600 700 800 900 1000 1100 1200 1300 **mass**

S. Barnes-UAB 1/12/07