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

Stephen Barnes, PhD

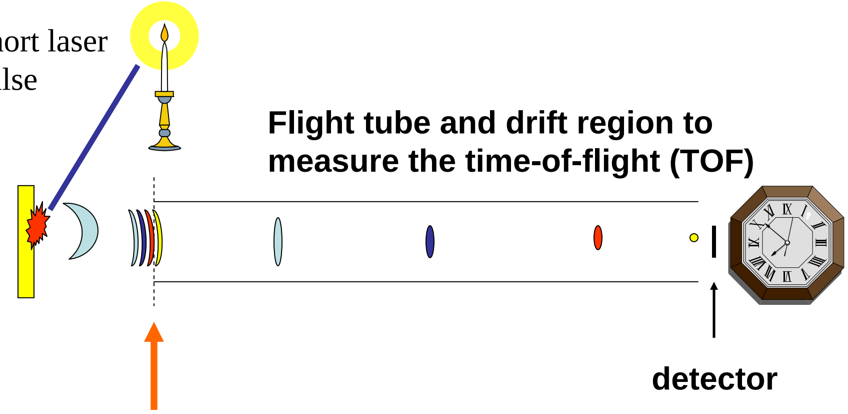
4-7117

sbarnes@uab.edu

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**

Matrix-Assisted Laser Desorption Ionization (MALDI)



Short laser pulse

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

detector

Accelerating pulse

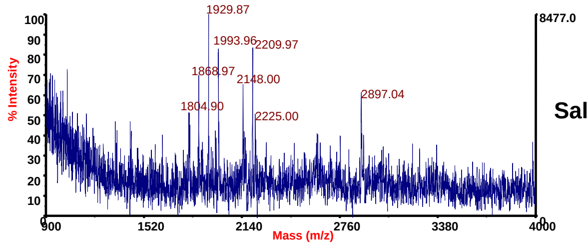
Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS)

- 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)

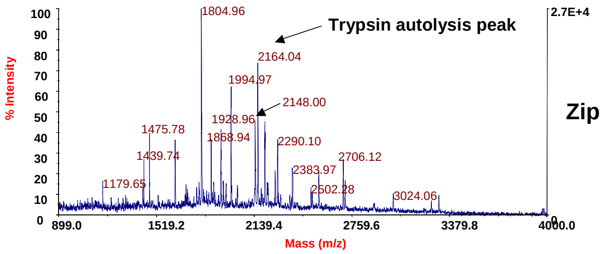
Factors from conventional experiments that impact MALDI-TOF analysis

- **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**

Benefit of removing salt from tryptic digest



Salt Contamination



ZipTip (desalting)

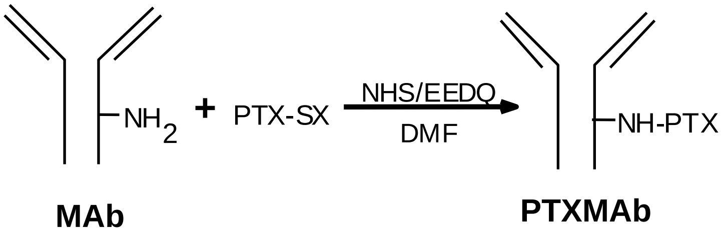
Chemically modifying an antibody to bind radioactive metal ions

Scheme 1

$PTX\text{-}2'\text{-}OH + \text{Succinic Anhydride} \longrightarrow PTX\text{-}2'\text{-}O_2CCH_2CH_2CO_2H$

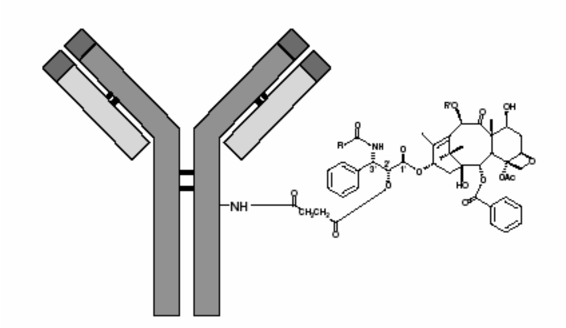
PTX-SX

Scheme 2



Ahmad Safavy

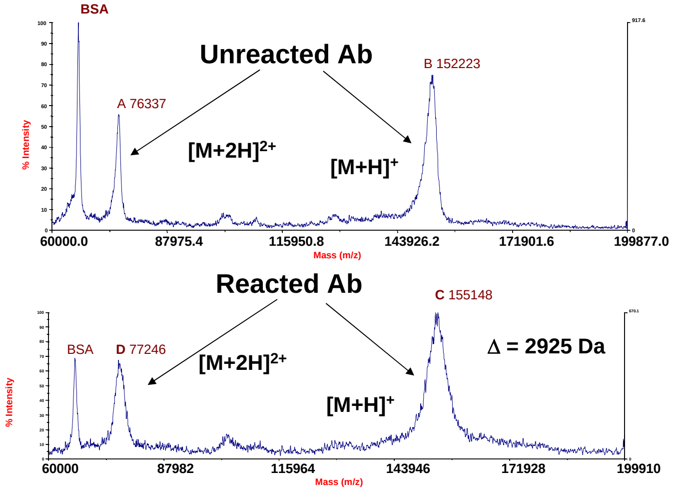
Structure of modified antibody



PTX-MAb

Ahmad Safavy

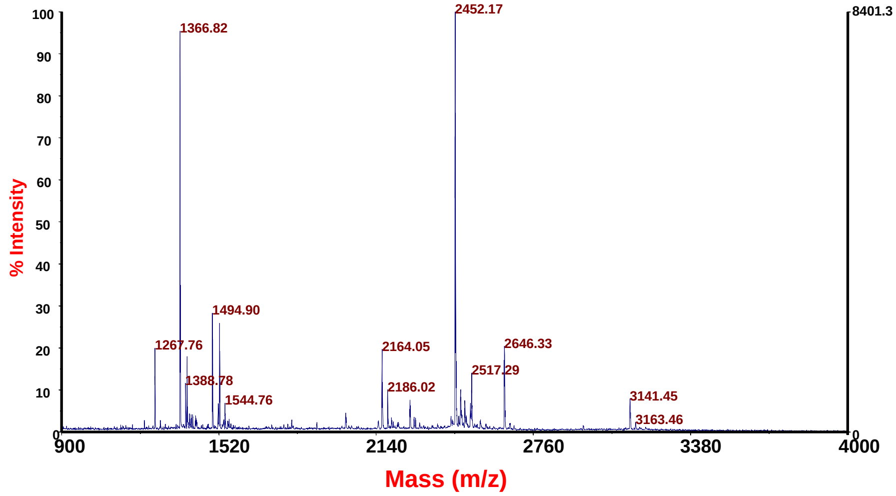
Modification of an antibody by MALDI-TOF



S. Barnes-UAB 1/08/08

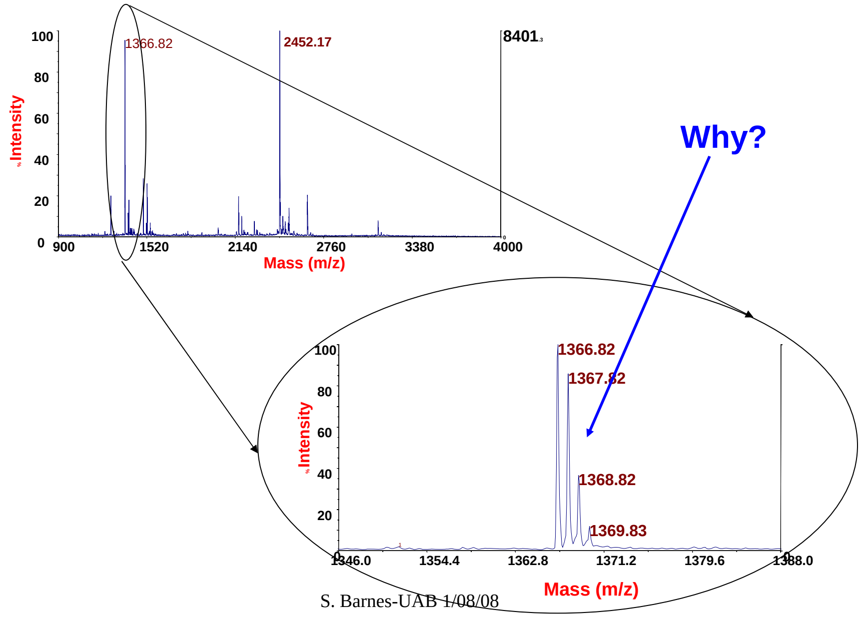
Ahmad Safavy

A MALDI-TOF mass spectrum of peptides



S. Barnes-UAB 1/08/08

Isotope profile of an individual peptide ion



S. Barnes-UAB 1/08/08

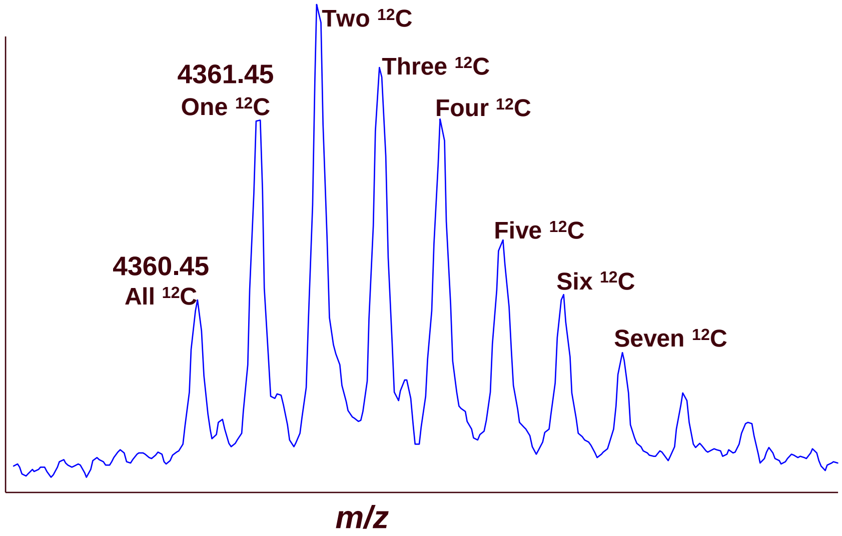
Stable isotopes of the most abundant elements found in peptides

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

*Varies according to its source

S. Barnes-UAB 1/08/08

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

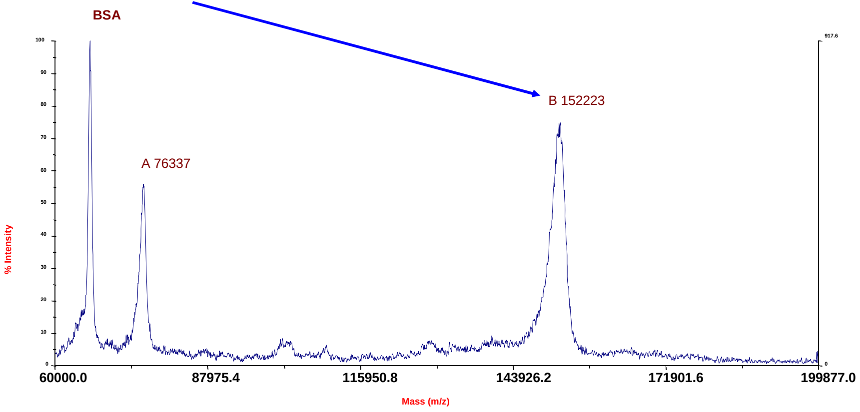


How to represent the mass of compound?

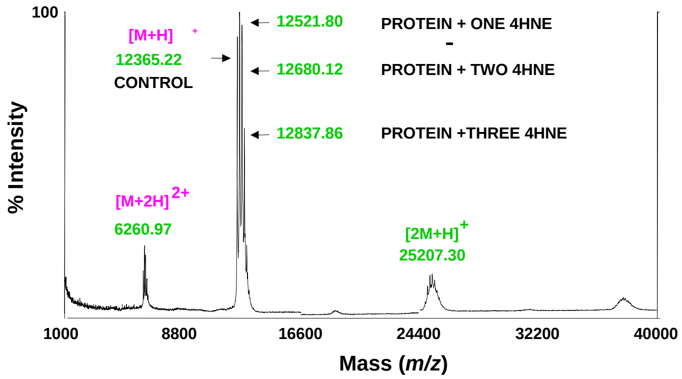
- At high resolution where the isotopic peaks are fully resolved, then we can determine the ***monoisotopic mass*** for each one
- At low mass resolution (where the isotope peaks cannot be resolved) what is observed is the ***average mass***

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.

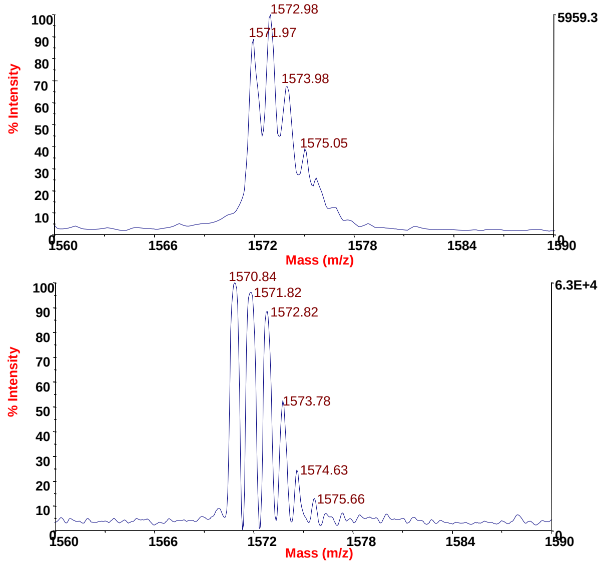


Modification of cytochrome C by the reactive aldehyde 4HNE



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.

Increased resolution and sensitivity in reflector vs. linear mode



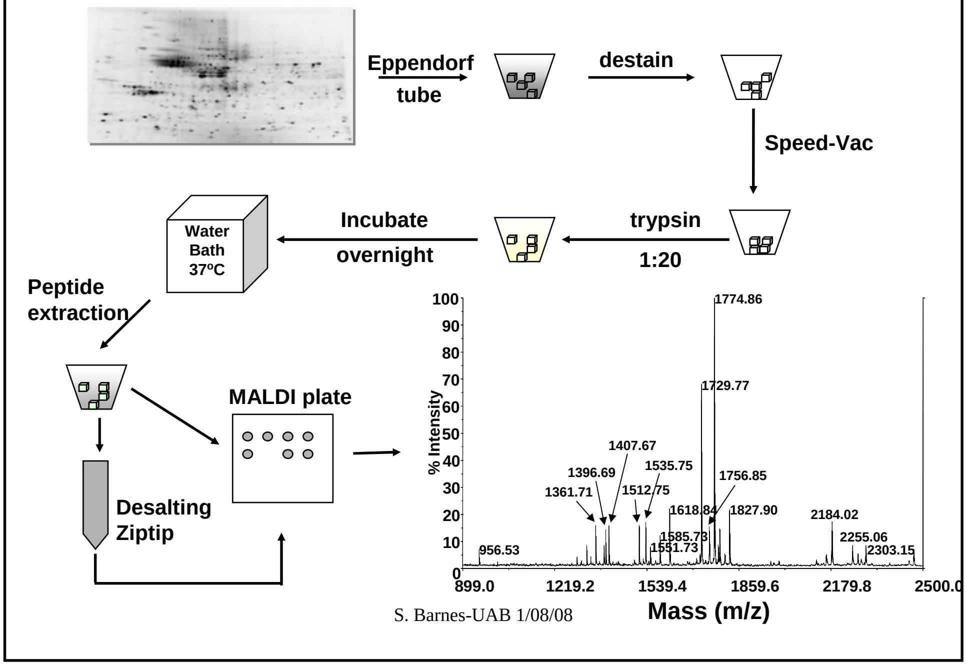
Linear Mode

Reflector Mode

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**

Protein analysis by MALDI 2008



Eppendorf tube

destain

Speed-Vac

trypsin 1:20

Incubate overnight

Water Bath 37ºC

Peptide extraction

MALDI plate

Desalting Ziptip

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**

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

S. Barnes-UAB 1/08/08

Choice of peptidase

- **Analogous to DNA restriction enzymes**
- **Tryptic peptide fingerprinting may identify, not one, but 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**

S. Barnes-UAB 1/08/08

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 1/08/08

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 1/08/08

Peptides from digestion with Glu-C

MKCLLLALALTCGAQALIVTQTMK
GLD

IQKVAGTWYSLAMAASD ISLLD AQSAPLRVYVE

E LKPTPE GD LE ILLQKWE NGE CAQKKIIAE

KTKIPAVFKID ALNE NKVLVLD TD YKKYLLFCME

NSAE PE QSLACQCLVRTPE VD D E ALE KFD

KALKALPMHIRLSFNPTQLE E QCHI

Amino acid residue masses

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

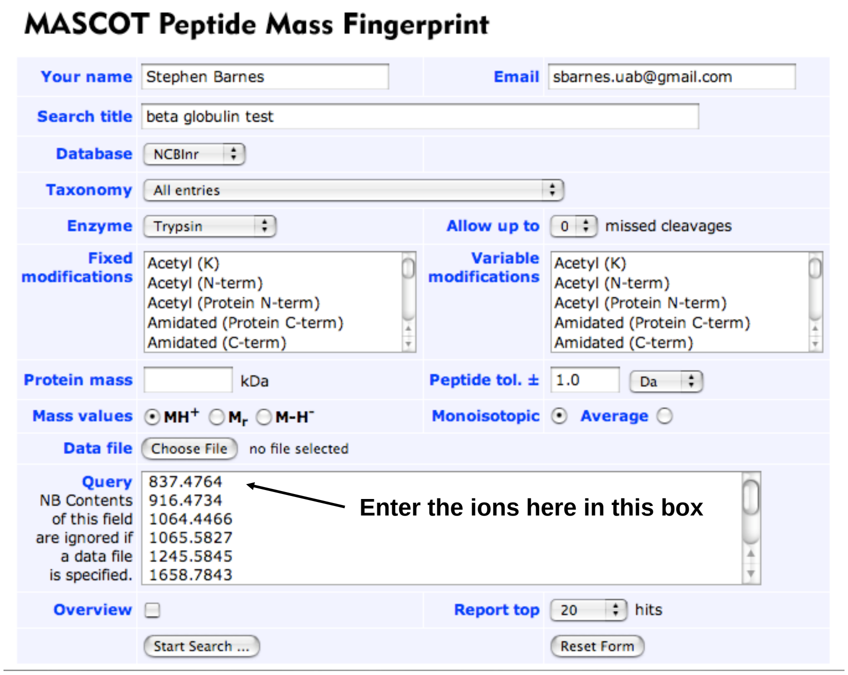
Expected peptides from trypsin and Glu-C digestion of bovine β-lactoglobulin

837.4764
916.4734
1064.4466
1065.5827
1245.5845
1658.7843
2275.2586
2313.2588
2647.2023
2707.3760

800.4876
929.5455
1003.5605
1232.6634
1259.7722
1337.6632
1447.7032
1811.8996
2307.3006
2819.5265

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

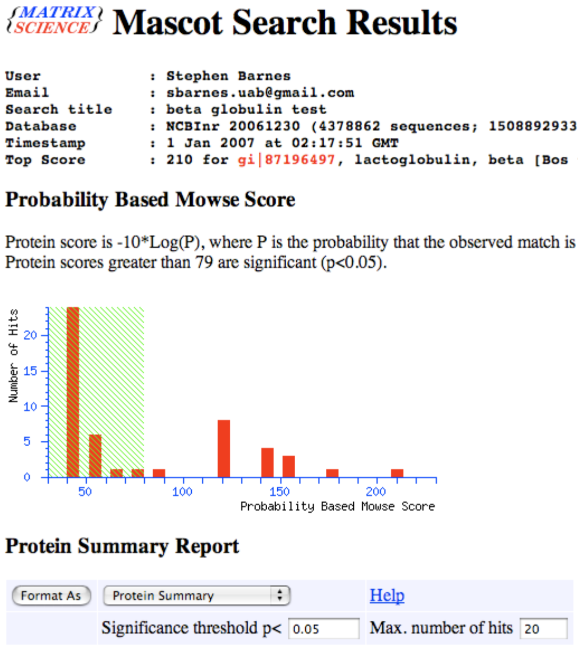
---



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^+ ○ M_r ○ $M\text{-}H^-$	**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		
Overview	☐	**Report top**	20 hits
	Start Search ...		Reset Form

Enter the ions here in this box



Mascot Search Results

```
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]
```

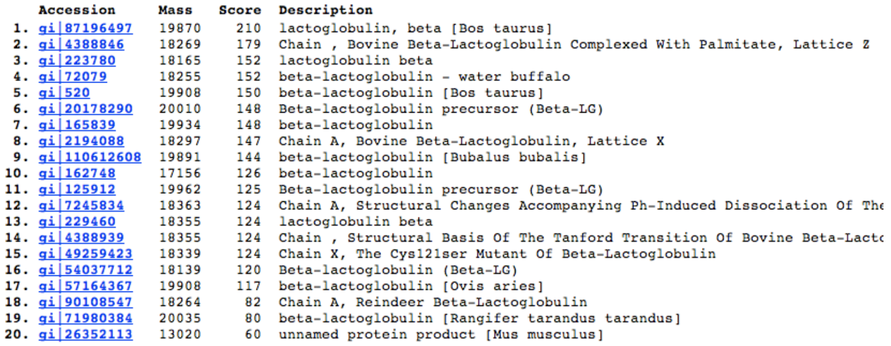
Probability Based Mowse Score

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

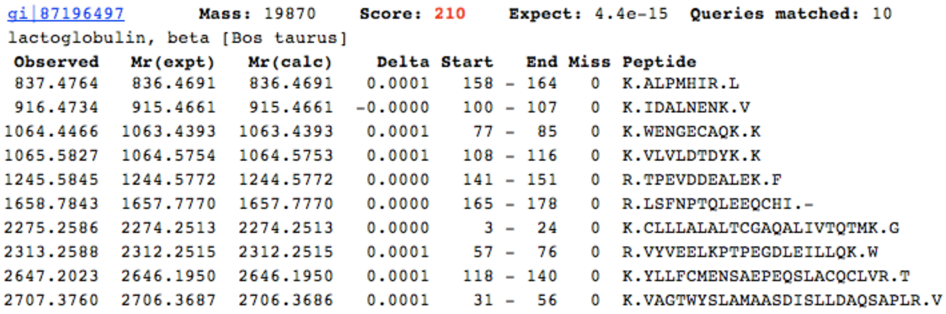
Significance threshold p< 0.05
Max. number of hits 20

Protein records provided by MASCOT search



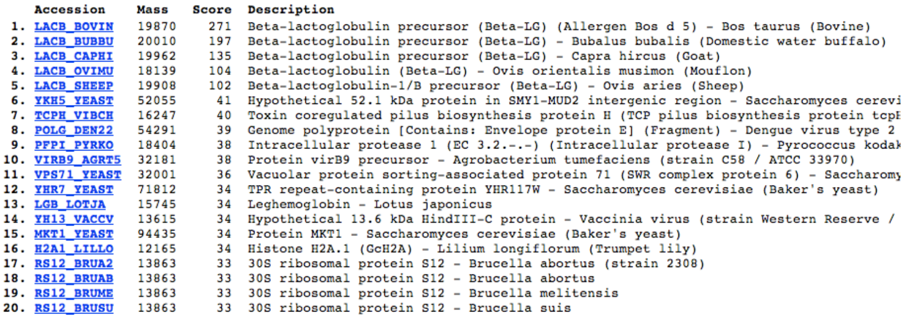
```
     Accession      Mass   Score  Description
 1. gi|87196497    19870    210  lactoglobulin, beta [Bos taurus]
 2. gi|4388846     18269    179  Chain , Bovine Beta-Lactoglobulin Complexed With Palmitate, Lattice Z
 3. gi|223780      18165    152  lactoglobulin beta
 4. gi|72079       18255    152  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-Lact
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]
```

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
```

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 tcp
 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 koda
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) - Saccharom
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
```

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. Cysteines 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

---

Other web sites for peptide analysis

- http://prowl.rockefeller.edu/
  - Choose ProFound
- http://prospector.ucsf.edu/
  - Choose MS-fit

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

Examples for homework (due Jan 22)

- Identify the following proteins from these MALDI ions (corrected for isotope effects):
  - 2280.11, 1700.89, 1393.71, 1385.74, 1164.60, 1087.49, 1007.54, 969.44, 916.50, 865.44, 813.40 (human)
  - 2582.21, 2511.17, 1677.96, 1647.77, 1581.81, 1068.48, 1018.49, 1011.58, 1001.45 (rat)
  - 1993.90, 1964.91, 1802.88, 1756.90, 1738.67, 1716.91, 1668.68, 1581.82, 1510.72, 1497.76, 1428.66, 1386.68, 1383.69, 1250.62, 1241.69, 1209.60, 1172.60, 1161.69, 1140.56, 1131.65, 1128.62, 1107.49, 1093.52, 1027.46, 1026.51, 1006.50 (mouse)
- Set the number of missed tryptic cuts to 0 and try varying the mass accuracy from 0.02 to 1.0 Da. How does this alter the MOWSE score?
- Do any of these have crystal structures?

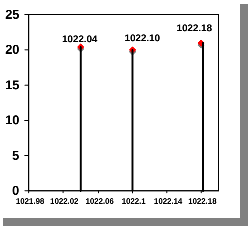
Take home question (due Jan 22)

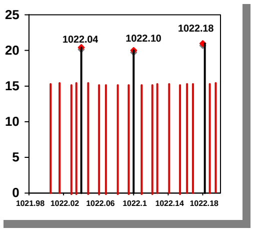
1. **What is the monoisotopic mass of human calgranulin A?**
   - **Hint: workout the empirical formula of human calgranulin A - its sequence can be obtained from record P05109 at http://www.ncbi.nlm.nih.gov**
2. **What is the molecular weight of the most abundant species of calgranulin A ?**
   - **Hint: assume that the abundance of ^{13}C is 1.11% of total carbon atoms - do not worry about ^{2}H or other isotopes**

S. Barnes-UAB 1/08/08

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 at that time
  - Now, the protein information space has become more dense and MALDI-TOF is no longer adequate
  - Previously identified proteins may not be correct

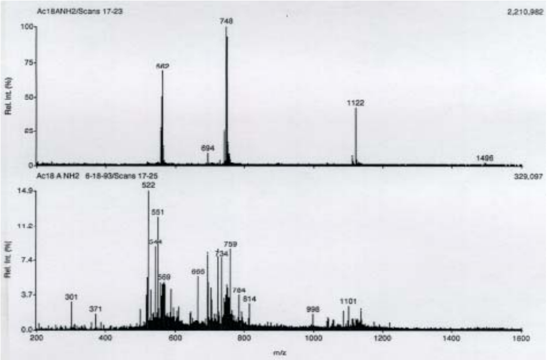




Electrospray ionization

- **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

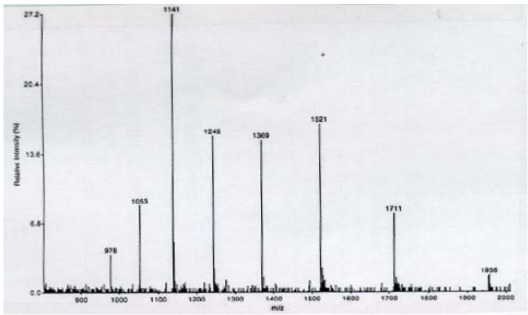


Guarantees of purity based on observation of "a single peak by reverse-phase HPLC" and by "it gave the correct sequence when analyzed by Edman degradation" are hollow.The lower spectrum was of a "pure" HPLC peak. The method of purification was amended and the upper spectrum was obtained

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**

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

Calculation of molecular weights and ion states

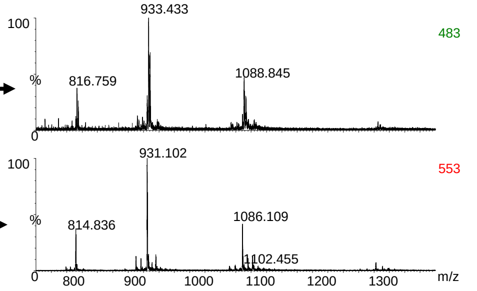
- **For two ions in a series for a peptide of molecular weight M, the lower *m/z* value (x) will be for the n+1 ion state and the larger *m/z* value (y) will be for the n+ ion state.**
  - **(1) $(M+n)/n = y$**
  - **(2) $(M+n+1)/(n+1) = x$**
- **Hence**
  - **(3) $M+n = ny$ and $M = ny-n$**
  - **(4) $M+n+1 = (n+1)x$ 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**

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**

Oxidized Aprotinin ESI mass spectrum

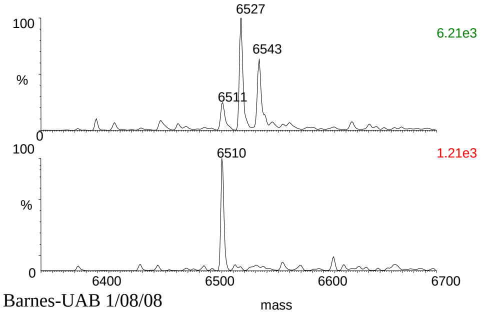
Control Aprotinin ESI mass spectrum



Deconvoluted mass spectra

Oxidized

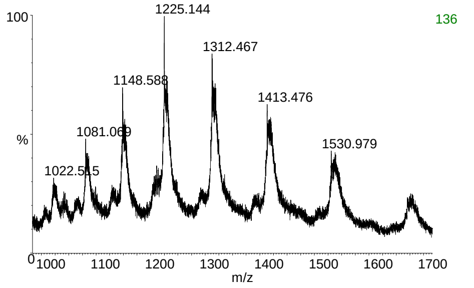
Control

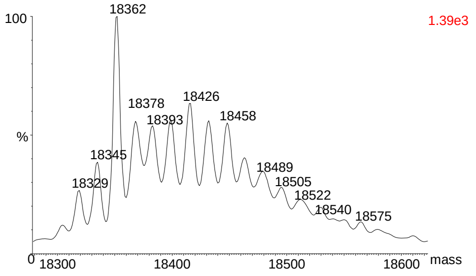


Junlong Shao

S. Barnes-UAB 1/08/08

Deconvolution of oxidized forms of β-lactoglobulin

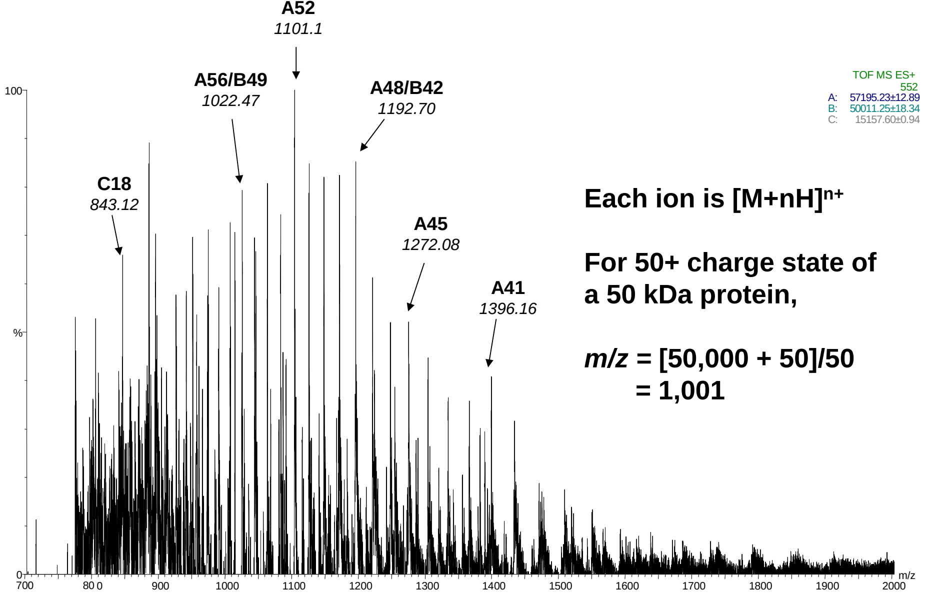




S. Barnes-UAB 1/08/08

Junlong Shao

ESI spectrum of bacterially expressed protein



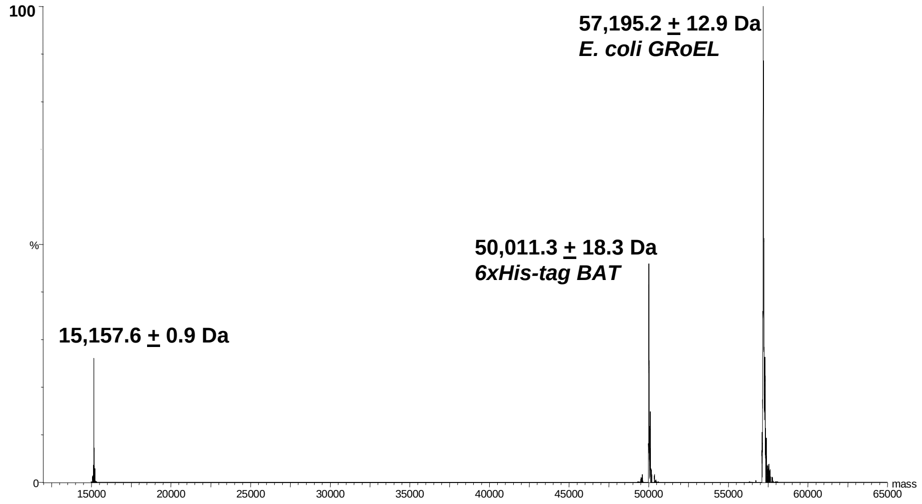
Each ion is $[M+nH]^{n+}$

For 50+ charge state of a 50 kDa protein,

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

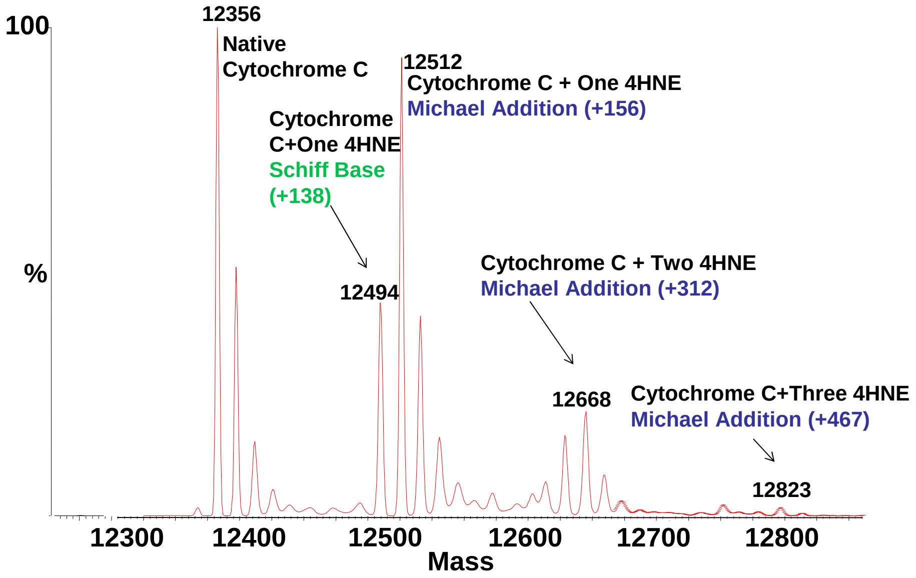
Courtesy of Mindan Sfakianos

MaxEnt deconvolution of MWs



Courtesy of Mindan Sfakianos

ESI-MS of 4HNE-Modified Cytochrome C



S. Barnes-UAB 1/08/08 *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**
  - **It's a 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**

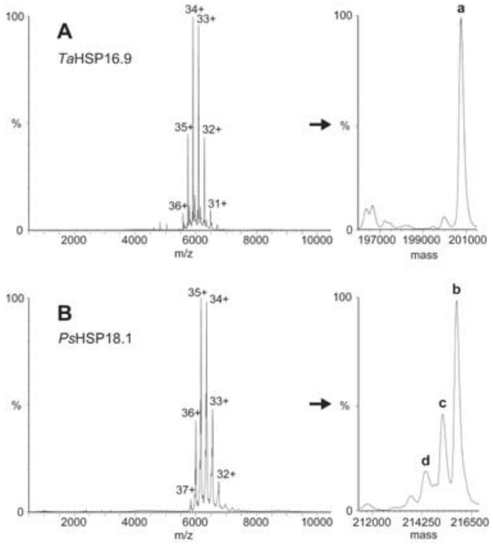
S. Barnes-UAB 1/08/08

Studying high molecular weight complexes by ESI

- **Most instrument ESI interfaces have a limited *m/z* range - up to 3,000**
- **In protein complexes water, and hence H+ ions, is “squeezed” out, thereby substantially increasing observed *m/z* values**
- **Interfaces that pass ions with *m/z* values above 10,000 have been designed**

---

nanoESI-MS of HMW complexes of small heat shock proteins



Note the large *m/z* values (6,000-7,000) for the observed ions

The ESI data were deconvoluted to reveal the distribution of the masses of the complexes

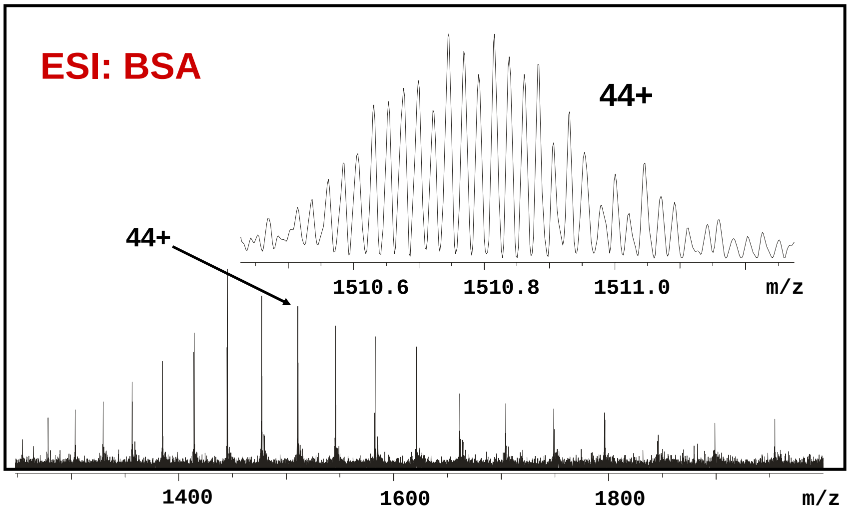
Sobbott et al., J Biol Chem 277:38921

Use of FT-MS in ESI of proteins

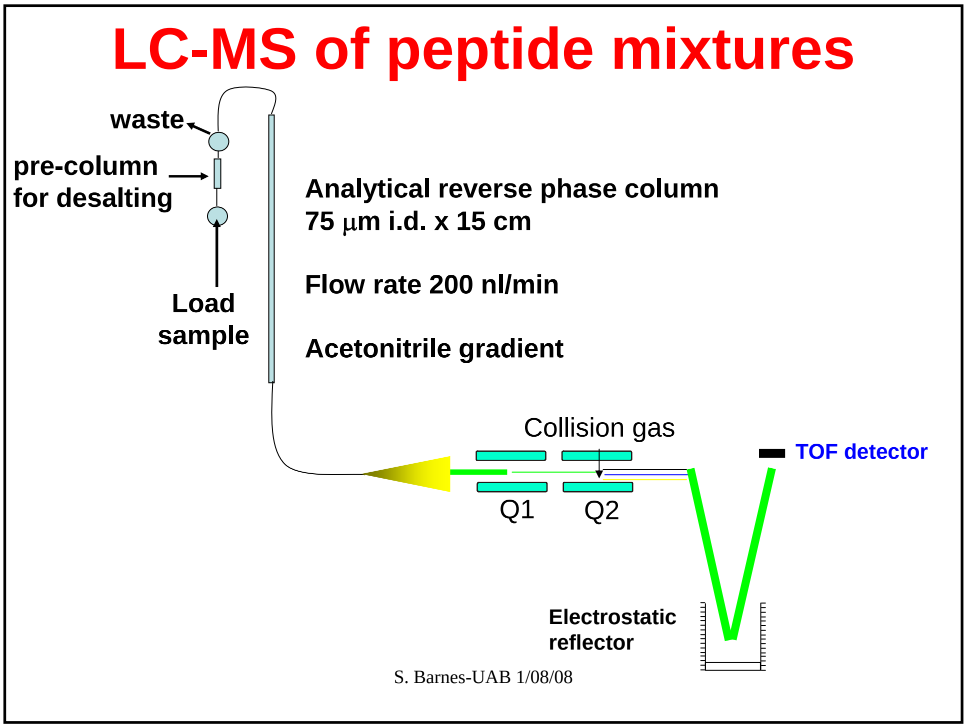
- The very high resolving power of FT-MS enables a direct measure of charge state of an individual ion since each peptide or polypeptide will have several/many isotope peaks
- The distance in Da between successive isotope peaks of a multiply charged ion is the reciprocal of the number of charges

---

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



Bruker Daltonics

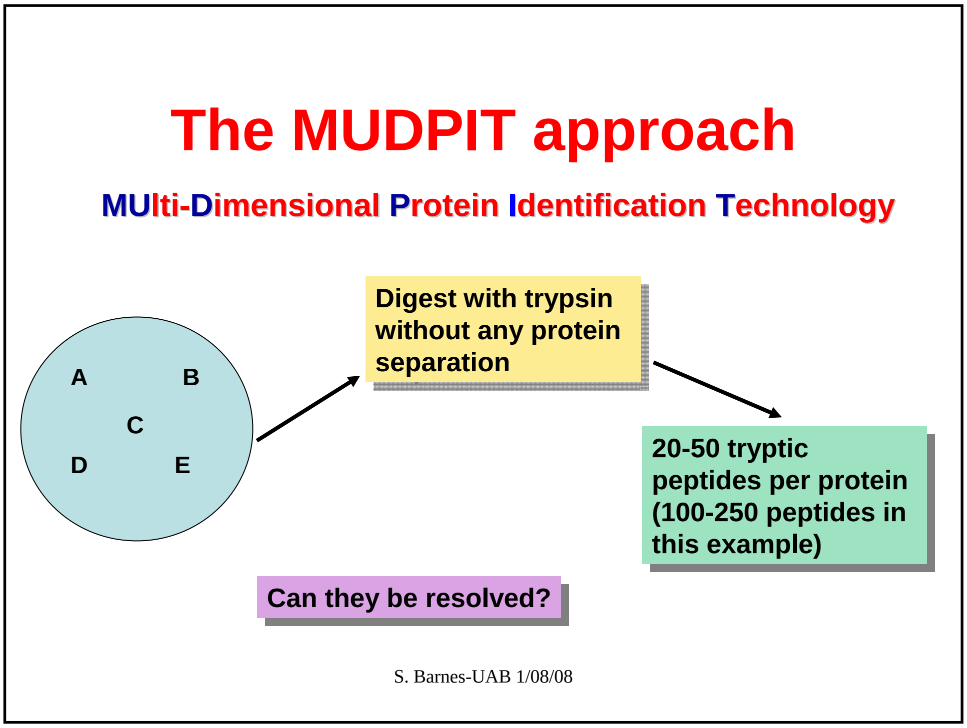


LC-MS of peptide mixtures

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

Flow rate 200 nl/min

Acetonitrile gradient



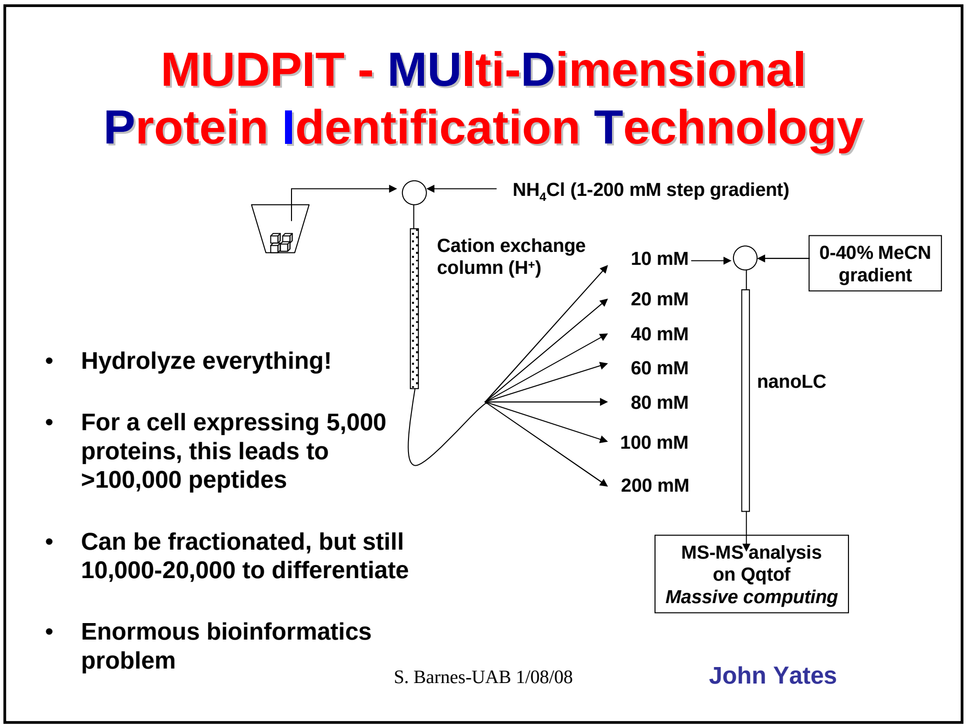
The MUDPIT approach

MUlti-Dimensional Protein Identification Technology

Digest with trypsin without any protein separation

20-50 tryptic peptides per protein (100-250 peptides in this example)

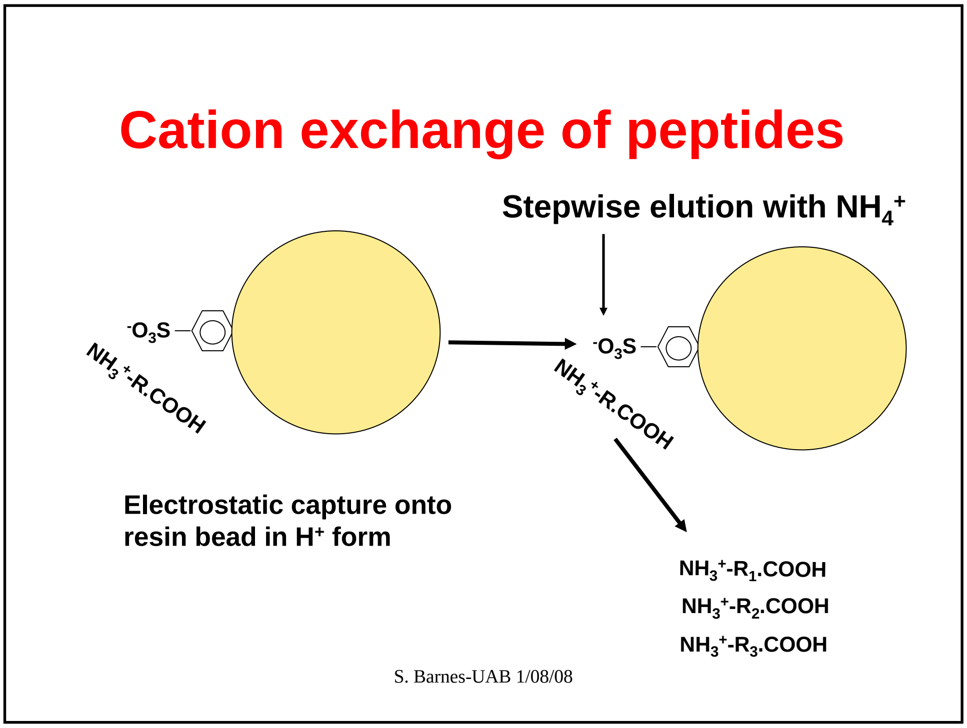
Can they be resolved?



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

John Yates



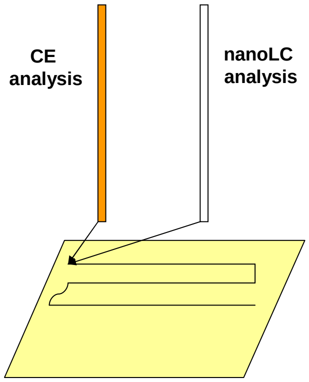
Cation exchange of peptides

Stepwise elution with NH_4^+

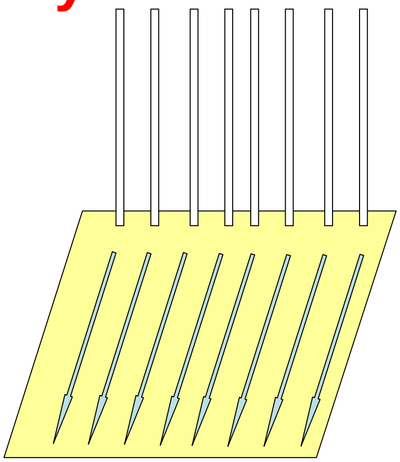
Electrostatic capture onto resin bead in H^+ form

NH_3^+-R_1.COOH
NH_3^+-R_2.COOH
NH_3^+-R_3.COOH

Connecting CE and LC to MALDI 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 1/08/08

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?

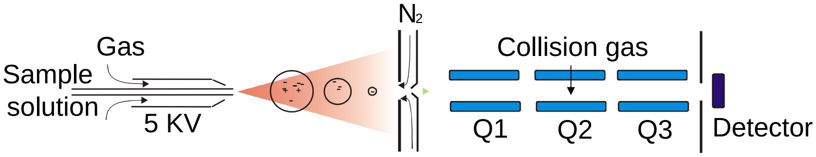
S. Barnes-UAB 1/08/08

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 1/08/08

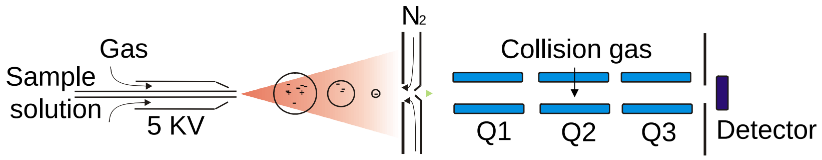
Tandem mass spectrometry on a triple quadrupole instrument



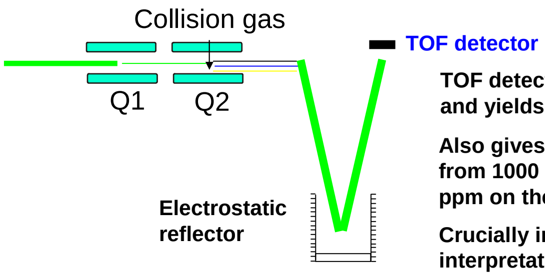
- **Daughter ion spectra**
- **Parent ion spectra**
- **Multiple reaction ion scanning**

S. Barnes-UAB 1/08/08

Triple quad versus Q-tof and sensitivity



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



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

---

Fragmentation of a peptide