Mass spectra of peptides and proteins - and LC analysis of proteomes

Stephen Barnes, PhD
4-7117
sbarnes@uab.edu

S. Barnes-UAB 1/28/05
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 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
Matrix-Assisted Laser Desorption Ionization (MALDI)

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

Accelerating pulse

Short laser pulse

detector

S. Barnes-UAB 1/28/05
Chemically modifying an antibody

Scheme 1

PTX-2'-OH + Succinic Anhydride \[\rightarrow\] PTX-2'-O\_2CCH\_2CH\_2CO\_2H

PTX-SX

Scheme 2

MAb + PTX-SX \[\rightarrow\] PTXMAb

S. Barnes-UAB 1/28/05

Ahmad Safavy
Structure of modified antibody

PTX-MAb

S. Barnes-UAB 1/28/05
Modification of an antibody by MALDI-TOF

\[ \Delta = 2925 \text{ Da} \]
A mass spectrum of peptides from a tryptic digest
Isotope profile of individual peptide ion

S. Barnes-UAB 1/28/05
Increased sensitivity in reflector vs. linear mode

Reflector Mode

Mass (m/z)

% Intensity

1570.84
1571.82
1572.82
1573.78
1574.63
1575.66

Linear Mode

Mass (m/z)

% Intensity

1570.97
1571.97
1573.98
1575.05

S. Barnes-UAB 1/28/05
Benefit of removing salt from tryptic digest

Salt Contamination

ZipTip (desalting)

Trypsin autolysis peak

S. Barnes-UAB 1/28/05
Protein analysis by MALDI 2004

1. **Peptide extraction**
2. **Desalting Ziptip**
3. **MALDI plate**
4. **Incubate overnight**
5. **Eppendorf tube**
6. **destain**
7. **Speed-Vac**

**Mass (m/z)**

- 956.53
- 1219.2
- 1539.4
- 1859.6
- 2179.8
- 2500.0

**% Intensity**

- 0
- 10
- 20
- 30
- 40
- 50
- 60
- 70
- 80
- 90
- 100

S. Barnes-UAB 1/28/05
Peptide mass fingerprinting

• This method has been developed because of the availability of predicted protein sequences from genome sequencing.

• Proteins do 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.
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
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 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
MALDI-TOF mass spectrum of tryptic digest of p22 band purified by 6xHis-tag

S. Barnes-UAB 1/28/05
Probability Based Mowse Score

Score is \(-10\times\log(P)\), where \(P\) is the probability that the observed match is a random event.

Protein scores greater than 71 are significant (\(p<0.05\)).

<table>
<thead>
<tr>
<th>Accession</th>
<th>Mass</th>
<th>Score</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gi</td>
<td>548939</td>
<td>20840</td>
<td>108 FKBP-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE SLYD (PPIASE) (ROTAMA</td>
</tr>
<tr>
<td>gi</td>
<td>13384624</td>
<td>46931</td>
<td>45 myocyte enhancer factor 2C [Mus musculus]</td>
</tr>
<tr>
<td>gi</td>
<td>5257384</td>
<td>43424</td>
<td>44 (AF137308) phytochrome B [Lolium perenne]</td>
</tr>
<tr>
<td>gi</td>
<td>4505147</td>
<td>50305</td>
<td>44 MADS box transcription enhancer factor 2, polypeptide C (myocyte enhan</td>
</tr>
<tr>
<td>gi</td>
<td>1515365</td>
<td>44552</td>
<td>43 (U52596) nucleocapsid protein [Avian infectious bronchitis virus]</td>
</tr>
<tr>
<td>gi</td>
<td>6093850</td>
<td>49443</td>
<td>42 PRESENILIN 2 (PS-2)</td>
</tr>
<tr>
<td>gi</td>
<td>15225198</td>
<td>47999</td>
<td>42 hypothetical protein [Arabidopsis thaliana]</td>
</tr>
<tr>
<td>gi</td>
<td>113854</td>
<td>58376</td>
<td>41 NITROGENASE IRON-IRON PROTEIN ALPHA CHAIN (NITROGENASE COMPONENT I)</td>
</tr>
<tr>
<td>gi</td>
<td>13928425</td>
<td>13831</td>
<td>40 (AB040419) envelope protein [Bovine immunodeficiency virus]</td>
</tr>
<tr>
<td>gi</td>
<td>4389228</td>
<td>56064</td>
<td>40 Chain Z, Crystal Structure Of The Complex Between Escherichia Coli Glycerol</td>
</tr>
</tbody>
</table>
### MASCOT SEARCH SUMMARY

1. gi|548939  Mass: 20840  Score: 108

**FKBP-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE SLYD (PPIASE) (ROTAMA)**

<table>
<thead>
<tr>
<th>Observed</th>
<th>Mr(expt)</th>
<th>Mr(calc)</th>
<th>Delta</th>
<th>Start</th>
<th>End</th>
<th>Miss</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>1046.38</td>
<td>1045.37</td>
<td>1045.59</td>
<td>-0.22</td>
<td>132</td>
<td>140</td>
<td>0</td>
<td>FNVEVVAIR</td>
</tr>
<tr>
<td>1262.47</td>
<td>1261.46</td>
<td>1261.70</td>
<td>-0.24</td>
<td>6</td>
<td>16</td>
<td>0</td>
<td>DLVVSLAYQVR</td>
</tr>
<tr>
<td>2343.88</td>
<td>2342.87</td>
<td>2343.08</td>
<td>-0.20</td>
<td>58</td>
<td>78</td>
<td>0</td>
<td>FDVAVGANDAYGQYDENLVQR</td>
</tr>
<tr>
<td>3857.71</td>
<td>3856.70</td>
<td>3856.89</td>
<td>-0.19</td>
<td>96</td>
<td>131</td>
<td>0</td>
<td>FLAETDQGPVPVEITAVEDDHVVVDGNHMLAGQNLK</td>
</tr>
</tbody>
</table>

2. gi|13384624  Mass: 46931  Score: 45

**myocyte enhancer factor 2C [Mus musculus]**

<table>
<thead>
<tr>
<th>Observed</th>
<th>Mr(expt)</th>
<th>Mr(calc)</th>
<th>Delta</th>
<th>Start</th>
<th>End</th>
<th>Miss</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>1046.38</td>
<td>1045.37</td>
<td>1045.50</td>
<td>-0.13</td>
<td>263</td>
<td>271</td>
<td>0</td>
<td>NTMPSVNQR</td>
</tr>
<tr>
<td>3857.71</td>
<td>3856.70</td>
<td>3856.76</td>
<td>-0.06</td>
<td>178</td>
<td>218</td>
<td>0</td>
<td>NSMSPGVTHRPPSA NTGGLMGDLTSGAGTSAGNGYGNPR</td>
</tr>
</tbody>
</table>

No match to: 1262.47, 2343.88

3. gi|5257384  Mass: 43424  Score: 44

**(AF137308) phytochrome B [Lolium perenne]**

<table>
<thead>
<tr>
<th>Observed</th>
<th>Mr(expt)</th>
<th>Mr(calc)</th>
<th>Delta</th>
<th>Start</th>
<th>End</th>
<th>Miss</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>1046.38</td>
<td>1045.37</td>
<td>1045.54</td>
<td>-0.17</td>
<td>380</td>
<td>389</td>
<td>0</td>
<td>GI DELSSVAR</td>
</tr>
<tr>
<td>3857.71</td>
<td>3856.70</td>
<td>3856.72</td>
<td>-0.02</td>
<td>86</td>
<td>122</td>
<td>0</td>
<td>SPHGCH AQYM ANMGSIASLVM AVI SSGGEDEHN MGR</td>
</tr>
</tbody>
</table>

No match to: 1262.47, 2343.88

4. gi|4505147  Mass: 50305  Score: 44

**MADS box transcription enhancer factor 2, polypeptide C (myocyte enhan**

<table>
<thead>
<tr>
<th>Observed</th>
<th>Mr(expt)</th>
<th>Mr(calc)</th>
<th>Delta</th>
<th>Start</th>
<th>End</th>
<th>Miss</th>
<th>Peptide</th>
</tr>
</thead>
<tbody>
<tr>
<td>1046.38</td>
<td>1045.37</td>
<td>1045.50</td>
<td>-0.13</td>
<td>265</td>
<td>273</td>
<td>0</td>
<td>NTMPSVNQR</td>
</tr>
<tr>
<td>3857.71</td>
<td>3856.70</td>
<td>3856.76</td>
<td>-0.06</td>
<td>180</td>
<td>220</td>
<td>0</td>
<td>NSMSPGVTHRPPSAGNTGGLMGDLTSGAGTSAGNGYGNPR</td>
</tr>
</tbody>
</table>

No match to: 1262.47, 2343.88

S. Barnes-UAB 1/28/05
E. coli: FKBP-TYPE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE

Nominal mass of protein (Mr): 20840

1  MKVAKDLVVS  LAYQVRTEDG  VLVDESPVSA  PLDYLHGHGS
41  LISGLETAL  GHEVGDKFDV  AVGANDAYGQ  YDENLVRQRVP
81  KDVFMSGVDE  QVGRMFLAET  DQGPVPEIET  AVEDDHVVVD
121  GNHMLAGQNL  KFNVEVVAIR  EATEEELAHG  HVHGAHDHHH
161  DHDHDGCCGG  HGHDHGEHG  GEGCCGGKGN  GCCGCH

Tryptic fragments detected by MALDI-TOF-MS

132–140  FNVEVVAIR
6– 16  DLVVSLSAYQVR
58– 78  FDVAVGANDAYGQYDENLVR
96–131  FLAETDQGPVPEITAVEDDHVVVDGNHMLAGQNLK

S. Barnes-UAB 1/28/05
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 this example it is gi|548939
  - 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 1IX5
  - Go to Structure (top of web page) and enter 1IX5 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 Feb 8th)

• Identify the following proteins from these MALDI ions (corrected for isotope effects):
  - 910.46, 1350.81, 1515.66, 1632.87, 1800.92, 1853.96 (human)
  - 965.46, 998.56, 1001.45, 1068.48, 1581.81, 1677.96 (rat)
  - 937.52, 972.49, 1049.56, 1209.64, 1508.70, 1844.89 (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?
How to represent the mass of compound?

• At low mass resolution (where the isotope peaks cannot be resolved) what is observed is the *average mass*

• At high resolution where the isotopic peaks are fully resolved, then we can determine the *monoisotopic mass* for each one
Take home question (due Feb 4th)

1. What is the monoisotopic mass of human myoglobin?
   - Hint: workout the empirical formula of hMyoglobin - its sequence can be obtained from record P02144 at http://www.ExPasy.org

2. What is the molecular weight of the most abundant species of human myoglobin?
   - Hint: assume that the abundance of $^{13}$C is 1.00% of total carbon atoms - do not worry about $^2$H or other isotopes
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

- $^+\text{H}_3\text{NCHR}_1\text{CO(NHCHR}_n\text{CO})_n\text{NHCHR}_2\text{COOH}$ is the ion that’s found in dilute acid solution.
- If there are internal basic residues, then the ions will be of the form $[\text{M+nH}]^{n+}$, where $n = 1, 2, \text{etc.}$.
- 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

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

Cumulative MW estimate = 13,680.29

SD = 2.94

S. Barnes-UAB 1/28/05
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

Junlong Shao
Deconvolution of oxidized forms of β-lactoglobulin
ESI spectrum of bacterially expressed protein

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

For 50+ charge state of a 50 kDa protein,

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

S. Barnes-UAB 1/28/05

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 1/28/05

Courtesy of Mindan Sfakianos
Cytochrome C Modified by HNE
MALDI-TOF Mass Spectrum

MALDI spectra usually contain only the molecular ion \([M+H]^+\). This is an advantage since it maximizes the signal, but is a disadvantage in that it gives no clues as to structure.
ESI-MS of 4HNE-Modified Cytochrome C

- 12356 Native Cytochrome C
- 12512 Cytochrome C + One 4HNE Michael Addition (+156)
- 12494 Cytochrome C+One 4HNE Schiff Base (+138)
- 12668 Cytochrome C+Three 4HNE Michael Addition (+467)
- 12823 Cytochrome C+Two 4HNE Michael Addition (+312)

S. Barnes-UAB 1/28/05

Courtesy of Amanda Foxwell
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
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

ESI: BSA
LC-MS of peptide mixtures

Pre-column for desalting

Load sample

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

Flow rate 200 nl/min

Acetonitrile gradient

Collision gas

Q1

Q2

Electrostatic reflector

TOF detector

S. Barnes-UAB 1/28/05
• 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

S. Barnes-UAB 1/28/05

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 1/28/05
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 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 1/28/05
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.
Tandem mass spectrometry on a triple quadrupole instrument

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

S. Barnes-UAB 1/28/05
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.
Fragmenting a peptide

\[
\text{NH}_3^+ - \text{CHR}^1 - \text{CO-NH-CHR}^2 - \text{CO-NH-CHR}^3 - \text{CO-NH-CHR}^4 - \text{COOH}
\]

\[
x_2 \quad y_2 \quad z_2
\]

\[
\begin{align*}
R^1 \quad O & \quad R^2 \\
| & | & | \\
H_2N--C--C--N^+=C & | & | \\
| & | & | \\
H & H & H & a_2
\end{align*}
\]

\[
\begin{align*}
R^3 \quad O & \quad R^4 \\
| & | & | \\
+O\equiv C--HN--C--C--N--C--COOH & | & | \\
| & | & | \\
H & H & H
\end{align*}
\]

\[
\begin{align*}
R^1 \quad O & \quad R^2 \\
| & | & | \\
H_2N--C--C--N=N=C--C\equiv O^+ & | & | \\
| & | & | \\
H & H & H & b_2
\end{align*}
\]

\[
\begin{align*}
R^3 \quad O & \quad R^4 \\
| & | & | \\
H_3N^+--C--C--N--C--COOH & | & | \\
| & | & | \\
H & H & H & y_2
\end{align*}
\]

\[
\begin{align*}
R^1 \quad O & \quad R^2 \quad O \\
| & | & | | \\
H_2N--C--C--N=N=C--NH_3^+ & | & | \\
| & | & | \\
H & H & H & c_2
\end{align*}
\]

\[
\begin{align*}
R^3 \quad O & \quad R^4 \\
| & | & | \\
+C--C--N--C--COOH & | & | \\
| & | & | \\
H & H & H & z_2
\end{align*}
\]

S. Barnes-UAB 1/28/05

http://www.matrixscience.com/help/fragmentation_help.html
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Mass</th>
<th>Amino Acid</th>
<th>Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>71.037</td>
<td>Leucine</td>
<td>113.084</td>
</tr>
<tr>
<td>Arginine</td>
<td>156.101</td>
<td>Lysine</td>
<td>128.094</td>
</tr>
<tr>
<td>Asparagine</td>
<td>114.043</td>
<td>Methionine</td>
<td>131.040</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>115.027</td>
<td>Phenylalanine</td>
<td>147.068</td>
</tr>
<tr>
<td>Cysteine</td>
<td>103.009</td>
<td>Proline</td>
<td>97.053</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>129.043</td>
<td>Serine</td>
<td>87.032</td>
</tr>
<tr>
<td>Glutamine</td>
<td>128.058</td>
<td>Threonine</td>
<td>101.048</td>
</tr>
<tr>
<td>Glycine</td>
<td>57.021</td>
<td>Tryptophan</td>
<td>186.079</td>
</tr>
<tr>
<td>Histidine</td>
<td>137.059</td>
<td>Tyrosine</td>
<td>163.063</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>113.084</td>
<td>Valine</td>
<td>99.068</td>
</tr>
</tbody>
</table>
Identification of daughter ions and peptide sequence

b ions: 262 375 446 503 632 760 875 989 1088 1187 1343

y ions: 1361 1247 1100 987 916 859 730 602 487 373 274 175
Sequencing O-GlcNAc peptides by ECD FT-ICR-MS

Casein kinase II - AGGSTPVSSANMMSG

[M+2H]2+

c8 fragment with sugar attached

b ion cleavage
c ion cleavage

S. Barnes-UAB 1/28/05
Fragment ions of a small 5-mer peptide

Homework - write down the masses of the b and y ions

\[ b_n = \text{[residue masses} + 1] \]

\[ y_n = \text{[residue masses} + \text{H}_2\text{O} + 1] \]

\[ \text{A Q Y E K} \]