Role of Chromatography and Mass Spectrometry in Biomedical Research

Robert J. Cotter
Pharmacology and Molecular Sciences
Biophysics Building Room B5
Johns Hopkins School of Medicine
Baltimore, MD 21205
(410) 955-3022
rcotter@jhmi.edu

Amina S. Woods
NIDA Intramural Program, NIH
Bayview Medical Center
5500 Nathan Schock Drive
Baltimore MD 21224
(410) 550-1507
awoods@intra.nida.nih.gov

Department of Pharmacology

Course: ME 330.804
2nd Quarter 2001/2002

Schedule

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
<th>Instructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>11-05-01</td>
<td>Ionization and Fragmentation</td>
<td>Dr. R. Cotter</td>
</tr>
<tr>
<td>11-07-01</td>
<td>Basic MALDI and Electrospray Theory</td>
<td>Dr. R. Cotter</td>
</tr>
<tr>
<td>11-12-01</td>
<td>Instrumentation: TOF, ion Traps, QTOFs and FTICR</td>
<td>Dr. R. Cotter</td>
</tr>
<tr>
<td>11-14-01</td>
<td>Sample Preparation for MALDI and LC/MS</td>
<td>Dr. A. Woods</td>
</tr>
<tr>
<td>11-19-01</td>
<td>Basic Chromatography: HPLC (RP &amp; IEC) &amp; CE</td>
<td>Dr. A. Woods</td>
</tr>
<tr>
<td>11-26-01</td>
<td>General Overview of Biological Applications</td>
<td>Dr. R. Cotter</td>
</tr>
<tr>
<td>11-28-01</td>
<td>Strategies for Characterization of Proteins</td>
<td>Dr. A. Woods</td>
</tr>
<tr>
<td>12-03-01</td>
<td>Proteomics</td>
<td>Dr. A. Woods</td>
</tr>
<tr>
<td>12-05-01</td>
<td>EXAM I</td>
<td></td>
</tr>
<tr>
<td>12-10-01</td>
<td>Tandem MS &amp; MS^n</td>
<td>Dr. R. Cotter</td>
</tr>
<tr>
<td>12-12-01</td>
<td>Strategies for Characterization of Glycolipids and Lipids</td>
<td>Dr. A. Woods</td>
</tr>
<tr>
<td>12-17-01</td>
<td>Strategies for Characterization of Oligosaccharides</td>
<td>Dr. A. Woods</td>
</tr>
<tr>
<td>12-19-01</td>
<td>Strategies for Characterization of DNA</td>
<td>Dr. R. Cotter</td>
</tr>
<tr>
<td>01-02-02</td>
<td>Quantitation</td>
<td>Dr. A. Woods</td>
</tr>
<tr>
<td>01-07-02</td>
<td>Review</td>
<td></td>
</tr>
<tr>
<td>01-09-02</td>
<td>EXAM II</td>
<td></td>
</tr>
</tbody>
</table>
Texts


http://www.lcms.com
**Ionization and Fragmentation**

1. **The mass spectrometer.**
   a. different ionization sources
   b. different mass analyzers, tandem analyzers, and hybrid instruments
   c. why a high vacuum is necessary

2. **Making ions: electron impact (EI) ionization for small molecules.**
   a. the importance of *heteroatoms* in ionization
   b. positive vs. negative ions
   c. how ionization at functional groups directs fragmentation

3. **Features of the mass spectrum.**
   a. relative intensity, base peak and mass/charge *(m/z)* ratio
   b. nominal, monoisotopic and average mass
   c. molecular weight, isotopic distribution and mass resolution

4. **How fragmentation enables the determination of chemical structures.**
   a. fragmentation of ethyl acetate
   b. fragmentation of amines

5. **Making ions softly: chemical ionization (CI).**
   a. electron impact: too much fragmentation?
   b. chemical ionization reactions of methane
   c. reagent gases and their proton affinities
   d. the importance of *even electron* (closed shell) ions
   e. chemical ionization and fragmentation of a dipeptide

6. **Basic instrumentation.**
   a. electron impact sources
   b. chemical ionization sources
   c. vacuum pumps: mechanical, diffusion and turbomolecular
   d. detectors
   e. mass is an *intrinsic* property of a molecule

7. **How are EI and CI used today?**
1. The mass spectrometer.

Neutral molecules must be converted to charged particles, i.e. ions

Ions can be separated according to mass by electric and/or magnetic fields

Ions impact a detector surface or are sensed by induced currents

Sample inlet system or interface to GC or LC

Vacuum chamber

Vacuum pumps

---

a. different ionization sources
b. different mass analyzers, tandem analyzers, and hybrid instruments

<table>
<thead>
<tr>
<th>Ionization Sources</th>
<th>Mass Analyzer</th>
</tr>
</thead>
<tbody>
<tr>
<td>thermal ionization</td>
<td>magnetic (B)</td>
</tr>
<tr>
<td>spark source</td>
<td>double-focusing (EB)</td>
</tr>
<tr>
<td>electron impact (EI)</td>
<td>reversed geometry (BE)</td>
</tr>
<tr>
<td>photoionization (PI)</td>
<td>ion cyclotron resonance (ICR)</td>
</tr>
<tr>
<td>chemical ionization (CI)</td>
<td>quadrupole (Q)</td>
</tr>
<tr>
<td>field ionization (FI)</td>
<td>quadrupole ion trap (ITMS)</td>
</tr>
<tr>
<td>field desorption (FD)</td>
<td>radio frequency (RF)</td>
</tr>
<tr>
<td>multiphoton ionization (MPI)</td>
<td>time-of-flight (TOF)</td>
</tr>
<tr>
<td>fast atom bombardment (FAB)</td>
<td>Fourier transform (FTMS)</td>
</tr>
<tr>
<td>plasma desorption mass spectrometry (PDMS)</td>
<td></td>
</tr>
<tr>
<td>secondary ion mass spectrometry (SIMS)</td>
<td>triple quadrupole (QQQ)</td>
</tr>
<tr>
<td>therospray (TS)</td>
<td>four sector (EBEB)</td>
</tr>
<tr>
<td>infrared laser desorption (IRLD)</td>
<td>hybrid (EBQQ)</td>
</tr>
<tr>
<td>matrix-assisted laser desorption (MALDI)</td>
<td>hybrid (QTOF)</td>
</tr>
<tr>
<td>electrospray ionization (ESI)</td>
<td>tandem TOF/TOF</td>
</tr>
</tbody>
</table>

---

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
c. Why a vacuum is necessary:

**collision frequency**: \# collisions sec\(^{-1}\) molecule\(^{-1}\)

\[ z = \sqrt{2\pi\sigma^2}\nu n' \]

where \( \sigma \) = molecular diameter
\( \pi\sigma^2 \) = collision cross section
\( \nu \) = average velocity = \((8kT/\pi m)^{1/2}\)
\( n' \) = number of molecules per cubic centimeter

**mean free path**: average distance traveled by a molecule between collisions

\[ \lambda = \frac{\nu}{z} = \frac{1}{\sqrt{2\pi\sigma^2}n'} \]

since \( PV = nRT \) where \( n = \# \) moles

then \( n' = (PV/RT)N_0 \)

where \( N_0 = \) Avagadro’s number

---

What is the mean free path of \( O_2 \) at atmospheric pressure?

\[ n' = \frac{(1 \text{ atm})(10^{-3} \text{ cm}^3 / \text{ liter})(6.023 \times 10^{23} \text{ molecules / mole})}{(0.082 \text{ liter-atm / deg-mole})(298 \text{ deg})} \]

\[ n' = 2.46 \times 10^{19} \text{ molecules / cm}^3 \]

\[ \lambda = \frac{1}{(1.414)(3.14)(3.61 \times 10^{-3})^2(2.46 \times 10^{19})} \]

\[ \lambda = 7.03 \times 10^{-8} \text{ cm} = 703 \text{ angstroms} \]

What is the mean free path of \( O_2 \) at \( 10^{-3} \) torr (mm Hg)?

\[ n' = \frac{(10^{-3} / 760)(10^{-3} \text{ cm}^3 / \text{ liter})(6.023 \times 10^{23} \text{ molecules / mole})}{(0.082 \text{ liter-atm / deg-mole})(298 \text{ deg})} \]

\[ n' = 3.24 \times 10^{13} \text{ molecules / cm}^3 \]

\[ \lambda = \frac{1}{(1.414)(3.14)(3.61 \times 10^{-3})^2(3.24 \times 10^{13})} \]

\[ \lambda = 5.3 \text{ cm} \]

---

The mean free path is shorter for biological macromolecules:

<table>
<thead>
<tr>
<th>Diameter</th>
<th>Oxygen</th>
<th>Actinomycin</th>
<th>Chymotrypsin</th>
<th>Aspartate transcarbamoylase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 meter</td>
<td>3.61 Å</td>
<td>20 Å (σ=11.8)</td>
<td>40 Å (σ=21.8)</td>
<td>70 Å (σ=36.8)</td>
</tr>
<tr>
<td>10⁻³ torr</td>
<td>5.3 cm</td>
<td>4.9 meters</td>
<td>1.4 meters</td>
<td>0.51 meters</td>
</tr>
<tr>
<td>10⁻⁴ torr</td>
<td>53 cm</td>
<td>53 meters</td>
<td>4.9 meters</td>
<td>14 meters</td>
</tr>
<tr>
<td>10⁻⁵ torr</td>
<td>530 meters</td>
<td>49 meters</td>
<td>14 meters</td>
<td>5.1 meters</td>
</tr>
</tbody>
</table>

1 meter TOF mass analyzer: 10⁻⁷ torr
Electron impact ionization source: 10⁻⁴ torr
Quadrupole ion trap: 10⁻³ torr (1 millitorr)
Chemical ionization source: 1-4 torr
Electrospray ion source: 1 atmosphere


2. Making ions: electron impact (EI) ionization for small molecules.

Positive ion formation:
\[
e^- + AB \rightarrow [AB^+ \cdot] + 2e^- \]
\[
[AB^+ \cdot] \rightarrow A + B^+ \]

\([AB^+ \cdot] \) is an odd electron ion carrying excess internal energy
Molecules have different ionization potentials (IP), but their ionization efficiencies are relatively flat around 70 ev.

a. the importance of heteroatoms in ionization

Heteroatoms are those that have non-bonding electrons, which are more easily removed to form ions:
S > N > O.

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
e⁻ + (CH₃CH₂CH₂)₃N → (CH₃CH₂CH₂)₃N⁺ + 2e⁻

e⁻ + CH₃CH₂-Ö-CH₂CH₃ → CH₃CH₂-O⁻⁻-CH₂CH₃

b. positive vs. negative ions

- negative EI is not as useful, since electron capture is resonant at different energies for each compound
- useful for electrophores, compounds used as labels that are ionized at different energies

c. how ionization at functional groups directs fragmentation

(1) ionization occurs by removal of non-bonding electron to form a radical ion:

(2) rearrangement of electrons leads to stable (even-electron) fragment ion:

(3) followed by loss of small neutral molecule:
3. Features of the mass spectrum.

\[
\begin{align*}
\text{H}_3\text{C} & \quad \text{9%} \quad \text{90%} \quad \text{1%} \\
\text{H}_3\text{C} & \quad \text{C} \quad \text{O} \\
\text{CH}_2\text{CH}_3 & \quad \text{CH}_3\text{O} \\
\end{align*}
\]

- \(m/z = 88\)
- \(m/z = 43\)
- \(m/z = 15\)

a. relative intensity, base peak and mass/charge (m/z) ratio

![Relative Abundance Graph]

<table>
<thead>
<tr>
<th>Element</th>
<th>Nominal Mass</th>
<th>Exact Mass</th>
<th>Average Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>12</td>
<td>12.00000</td>
<td>(98.9%)</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>13.00335</td>
<td>(1.1%)</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>1.00783</td>
<td>(99.98%)</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.0140</td>
<td>(0.02%)</td>
</tr>
<tr>
<td>O</td>
<td>16</td>
<td>15.99491</td>
<td>(99.76%)</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>17.9992</td>
<td>(0.20%)</td>
</tr>
<tr>
<td>N</td>
<td>14</td>
<td>14.00307</td>
<td>(99.63%)</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>15.00011</td>
<td>(0.37%)</td>
</tr>
<tr>
<td>S</td>
<td>32</td>
<td>31.97207</td>
<td>(95.0%)</td>
</tr>
<tr>
<td></td>
<td>33</td>
<td>32.97146</td>
<td>(0.76%)</td>
</tr>
<tr>
<td></td>
<td>34</td>
<td>33.96786</td>
<td>(4.22%)</td>
</tr>
<tr>
<td>Cl</td>
<td>35</td>
<td>34.96885</td>
<td>(75.53%)</td>
</tr>
<tr>
<td></td>
<td>37</td>
<td>36.9659</td>
<td>(24.47%)</td>
</tr>
</tbody>
</table>

b. nominal, monoisotopic and average mass
Ethyl acetate (molecular structure):

![Molecular structure of Ethyl acetate]

**Empirical formula:** $C_4H_8O_2$

**Nominal mass** = 88, where C=12, H=1 and O=16

now: 

\[
\begin{align*}
4 \text{ C} & \times 12.00000 = 48.00000 \\
8 \text{ H} & \times 1.00783 = 8.06264 \\
2 \text{ O} & \times 15.99491 = 31.98982 \\
\end{align*}
\]

\[88.05246\]

**Monoisotopic mass** = 88.05

**Exact mass** = 88.05246

**Mass defect** = 0.05 (or 0.05246)

For a small molecule, the mass defect is small, and the nominal mass is usually sufficient.

For large molecules, such as glucagon (described below) the monoisotopic mass may be several mass units higher than the nominal mass.

Note that the mass defect for oxygen is negative.

The monoisotopic mass is so called because there is only one arrangement of isotopes that compose that mass.

---

### c. molecular weight, isotopic distribution and mass resolution

The peak at 89.05 is not a single peak:

\[
\begin{align*}
^{12}\text{C}_3^{13}\text{C}_1\text{H}_8\text{O}_2 & \quad 89.05581 \quad 4.4\% \\
\text{C}_4\text{H}_7\text{D}_1\text{O}_2 & \quad 89.05863 \quad 0.16\% \\
\end{align*}
\]

4.56%

for the peptide glucagon $C_{153}H_{225}N_{42}O_{50}S$

**nominal mass** = 3481

**monoisotopic mass** = 3482.61 (first peak)

**second peak:**

\[
\begin{align*}
^{12}\text{C} & \rightarrow ^{13}\text{C} \quad 153 \times 1.1\% \quad 170 \quad 170\% \\
^{1}\text{H} & \rightarrow ^{2}\text{H} \quad 225 \times 0.02\% \quad 4.5\% \\
^{14}\text{N} & \rightarrow ^{15}\text{N} \quad 42 \times 0.37\% \quad 15.5\% \\
\end{align*}
\]

190 %

therefore, the second peak is larger

The mass resolution required to separate the peak at 89.05 would be:

\[
m/\Delta m = 89/0.003 = 29,667
\]

Or 1 part in 29,667.

Thus, these peaks are not resolved.

The second peak for glucagon has a multiplicity of 3. The multiplicity for the third and following peaks is higher.
c. molecular weight, isotopic distribution and mass resolution

<table>
<thead>
<tr>
<th>NOMINAL, MONOISOTOPIC and AVERAGE masses</th>
</tr>
</thead>
<tbody>
<tr>
<td>153 C 1836</td>
</tr>
<tr>
<td>225 H 225</td>
</tr>
<tr>
<td>42 N 588</td>
</tr>
<tr>
<td>50 O 800</td>
</tr>
<tr>
<td>1 S 32</td>
</tr>
<tr>
<td>3481</td>
</tr>
<tr>
<td>153 C 1837.706</td>
</tr>
<tr>
<td>225 H 226.950</td>
</tr>
<tr>
<td>42 N 588.281</td>
</tr>
<tr>
<td>50 O 799.970</td>
</tr>
<tr>
<td>1 S 32.064</td>
</tr>
<tr>
<td>3484.971</td>
</tr>
</tbody>
</table>

---

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods

4. How fragmentation enables the determination of chemical structures.

a. fragmentation of ethyl acetate, b. fragmentation of amines

\[
\text{CH}_3\text{C} = \text{NH}_2 + e^- \rightarrow \text{CH}_3\text{C} = \text{NH}_2 + 2e^- + 73
\]

\[
\text{CH}_3\text{C} = \text{NH}_2 \rightarrow \text{CH}_3\text{C} = \text{NH}_2 + \cdot\text{CH}_3 + 58
\]

\[
\text{H}_3\text{C} = \text{N} - \text{C}_2\text{H}_7 + e^- \rightarrow \text{H}_3\text{C} = \text{N} - \text{C}_2\text{H}_7 + 73
\]

\[
\text{H}_3\text{C} = \text{N} = \text{C}_2\text{H}_5 \rightarrow \text{H}_3\text{C} = \text{N} = \text{C}_2\text{H}_5 + \cdot\text{C}_2\text{H}_5 + 44
\]

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
5. **Making ions softly: chemical ionization (CI).**

a. electron impact: too much fragmentation?

- Energy of electron = 70 ev
- Ionization potential (removal of electron) = 6-15 ev

b. Chemical ionization reactions of methane

\[
e^{-} + \text{CH}_4 \rightarrow \text{CH}_4^{**} + 2e^{-} \quad \text{electron ionization (IP = 9.95 ev)}
\]

\[
\text{CH}_4^{**} + \text{CH}_4 \rightarrow \text{CH}_5^{+} + \text{•CH}_3 \quad \text{in 1–4 torr methane (AP = 14.4 ev)}
\]

\[
\text{CH}_5^{+} + \text{M} \rightarrow \text{CH}_4 + \text{MH}^{+} \quad \text{protonated molecular ion}
\]

\[\Delta H_{\text{reaction}} = PA(\text{CH}_4) - PA(\text{M})\]

- Electrons are not removed; protons are transferred.
- Protonation reaction is *mildly* exothermic ($\Delta H_{\text{reaction}}$ is negative).
- Ionization is “soft”; i.e. there is little fragmentation.

Yes, if you cannot get a molecular ion!

c. Reagent gases and their proton affinities

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Proton Affinity (Kcal/mole)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH₃</td>
<td>207</td>
</tr>
<tr>
<td>Isobutene</td>
<td>195</td>
</tr>
<tr>
<td>H₂O</td>
<td>165</td>
</tr>
<tr>
<td>CH₄</td>
<td>126</td>
</tr>
<tr>
<td>H₂</td>
<td>101</td>
</tr>
</tbody>
</table>

A large proton affinity means that the reagent ion does not easily transfer its proton to the sample; ionization is soft.

d. The importance of even electron (closed shell) ions

- **Electron impact**

  \[
  \text{CH}_3\text{C} = \text{NH}_2\rightarrow e^{-} \quad \text{m/z 73 (radical ion)}
  \]

- **Chemical ionization**

  \[
  \text{CH}_3\text{C} = \text{NH}_2\rightarrow H^{+} \quad \text{m/z 74 (even electron ion)}
  \]

  \[
  \text{CH}_3\text{C} = \text{NH}_2\rightarrow \text{CH}_4 \rightarrow \text{m/z 58 (even electron ion)}
  \]
e. chemical ionization and fragmentation of a dipeptide

<table>
<thead>
<tr>
<th>Class</th>
<th>Formula</th>
<th>Energy (Kcal/mole)</th>
<th>Charge (ev)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tertiary amines</td>
<td>(CH$_3$)$_3$N</td>
<td>225</td>
<td>9.8</td>
</tr>
<tr>
<td>secondary amines</td>
<td>(CH$_3$)$_2$NH</td>
<td>220</td>
<td>9.5</td>
</tr>
<tr>
<td>primary amine</td>
<td>CH$_3$NH$_2$</td>
<td>214</td>
<td>9.3</td>
</tr>
<tr>
<td>amide</td>
<td>CH$_3$CONH$_2$</td>
<td>210</td>
<td>9.1</td>
</tr>
<tr>
<td>ammonia</td>
<td>NH$_3$</td>
<td>207</td>
<td>9.0</td>
</tr>
<tr>
<td>acids</td>
<td>CH$_3$COOH</td>
<td>196</td>
<td>8.5</td>
</tr>
<tr>
<td>alcohols</td>
<td>CH$_3$OH</td>
<td>189</td>
<td>8.2</td>
</tr>
</tbody>
</table>

A higher proton affinity means a more likely site for protonation, but may also result in more fragmentation.

Carboxylic acids, esters and ethers:

\[
R-C\overset{\cdot}{O}H + H^+ \rightarrow R-C\overset{1H\oplus}{O}H \rightarrow R-C\overset{\cdot}{O}H
\]

\[
R-\overset{\cdot}{O}R + H^+ \rightarrow R\overset{\cdot}{O}H
\]

An amino acid will protonate at the amine nitrogen:

\[
H_2N\overset{\cdot}{C}H-C\overset{\cdot}{O}H + H^+ \rightarrow H_2N\overset{\cdot}{C}H-C\overset{1H\oplus}{O}H
\]

Acetamide may have two protonated forms:

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{NH}_2 + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{1H\oplus}{NH}_3
\]

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{NH}_2 + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{1H\oplus}{OH}
\]

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{NH}_2 + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{\cdot}{OH}
\]

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{NH}_2 + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{\cdot}{OH}
\]

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{OH} + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{\cdot}{OH}
\]

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{OH} + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{\cdot}{OH}
\]

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{OH} + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{\cdot}{OH}
\]

\[
H_3C\overset{\cdot}{C}-\overset{\cdot}{OH} + H^+ \rightarrow H_3C\overset{\cdot}{C}-\overset{\cdot}{OH}
\]
Losses of small neutral molecules involve rearrangements:

loss of side chain can distinguish leucine and isoleucine

\[
\begin{align*}
\text{m/z 132} & \quad \text{y-ion} \\
\end{align*}
\]

loss of HCOOH

\[
\begin{align*}
\text{m/z 74} & \quad \text{y-ion} \\
\text{m/z 86} & \quad \text{a-ion} \\
\end{align*}
\]

Protonation of a dipeptide results in cleavage of the amide bond. The charge may be retained on the N-terminal amino acid:

heterolytic cleavage

\[
\begin{align*}
\text{carbonium ion} & \quad \text{leaving group is a stable neutral amine} \\
\text{b-ion} & \quad \text{A-ion} \\
\end{align*}
\]

\[
\begin{align*}
\text{acylium ion} & \quad \text{m/z 86} \\
\end{align*}
\]
Alternatively, the charge may be retained on the C-terminal amino acid:

\[ \text{H}_2\text{N}-\text{C}-\text{C}-\text{N}+\text{H} \rightarrow \text{H}_2\text{N}-\text{C}-\text{C}-\text{N}+\text{H} \]

hydrogen transfer

The z-ion might be formed more directly by a simple cleavage of the N-C bond:

\[ \text{H}_2\text{N}-\text{C}-\text{C}-\text{N}+\text{H} \rightarrow \text{H}_2\text{N}-\text{C}-\text{C}-\text{N}+\text{H} \]

6. Basic instrumentation.

a. electron impact sources

b. chemical ionization sources

CI sources are similar to EI sources, but are closed to maintain 1-4 torr of reagent gas pressure

c. vacuum pumps: mechanical, diffusion and turbomolecular

Mechanical pumps
- pump from atmosphere to a few microns
- most common type is direct drive
- used to back up diffusion or turbomolecular pumps

Turbomolecular pumps
- high vacuum pumps used to achieve pressures of $10^{-7}$ torr or better
- output is attached to mechanical pump

d. detectors

Discreet dynode electron multipliers
- oldest type of detector used on sector, TOF and many other mass spectrometers
- not as easily saturated as channelplate detectors

Channelplate detector
- used in many TOF mass spectrometers
- coaxial detectors for reflectrons have center hole

Channeltron
- used originally on quadrupole mass spectrometers
- common for quadrupole ion trap mass spectrometers
d. detectors

**Discrete dynode electron multipliers**
- oldest type of detector used on sector, TOF and many other mass spectrometers
- not as easily saturated as channelplate detectors

**Channelplate detector**
- used in many TOF mass spectrometers
- coaxial detectors for reflectrons have center hole

**Channeltron**
- used originally on quadrupole mass spectrometers
- common for quadrupole ion trap mass spectrometers

---

**e. mass is an intrinsic property of a molecule**

<table>
<thead>
<tr>
<th>Mass</th>
<th>ion mobility</th>
<th>retention time</th>
<th>( M_r )</th>
</tr>
</thead>
<tbody>
<tr>
<td>The molecular mass can be calculated from the masses of the elements. ( \text{CH}_3\text{COOCH}_2\text{CH}_3 ) has a mass of: ( 4C = 4\times12 = 48 ) ( 2O = 2\times16 = 32 ) ( 8H = 8\times1 = 8 ) ( \text{----------------} ) ( 88 )</td>
<td>depends on:</td>
<td>depends on:</td>
<td>depends on:</td>
</tr>
<tr>
<td>- electric field</td>
<td>- pressure</td>
<td>- electric field</td>
<td></td>
</tr>
<tr>
<td>- gas pressure</td>
<td>- flow rate</td>
<td>- solid phase</td>
<td></td>
</tr>
<tr>
<td>- molecular shape</td>
<td>- solid phase</td>
<td>- solvent</td>
<td></td>
</tr>
<tr>
<td>- temperature</td>
<td>- solvents</td>
<td>- molecular shape</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- gradients</td>
<td>- polarity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- temperature</td>
<td>- temperature</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- hydrophobicity</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
7. How are EI and CI used today?

a. Forensics
b. Environmental pollutant analysis, using positive/negative CI and GCMS
c. High resolution analysis of dioxins
d. Carbohydrate linkage analysis using permethylated derivatives of hydrolyzed aldito acetates and GCMS (Hakamori method).
e. Identification of drug metabolites
f. Quantitation of metabolites using stable isotope analysis and GCMS
g. Monitoring of drug usage (including both human and equine)
h. Residual gas analysis (including evolved gases in semiconductor manufacture)
i. Space: earth, Martian and other planetary atmospheres; comet flyby
j. Space: detection of organic/biological molecules on planetary surfaces
k. Space: environmental monitoring on space stations
l. Space: humans in space (HIS) monitoring

Example of selected ion monitoring used in gas chromatography mass spectrometry (GC/MS)

Mass spectrometers have been used extensively in space

Viking Landers on Mars
Viking 1, August 20, 1975
Viking 2, September 9, 1975

Neutral Mass Spectrometer (NMS) Alfred O. Nier
• double-focusing (electrostatic and magnetic) mass spectrometer
• measure concentration of species in the Mars atmosphere
• two collectors: 1-7 u and 7-49 u
• mass scanning by sweeping $V_a$ and $±E$

Gas Chromatograph/Mass Spectrometer (GCMS) Klaus Biemann
• substances vaporized from heated soil
• hydrogen carrier gas removed through permeable palladium
• 12-200 u mass range

Mass spectrometers are being used extensively in space

Cassini Mission to Saturn/Titan
Ion & Neutral Mass Spectrometer (INMS)
• quadrupole mass analyzer
• mass range 1-99 amu

Huygens Probe GCMS
• quadrupole mass spectrometer
• aerosol pyrolyzer
Mass spectrometers have also been used in the search for extraterrestrial life

Analysis of interior fracture surface of ALH84001 Martian meteorite from Antarctica (Richard N. Zare, et al.)

- multiphoton ionization (MPI) to find organic compounds
- time-of-flight (TOF) mass spectrometry
- polyaromatic hydrocarbons (PAHs):
  - 178 phenanthrene
  - 202 pyrene
  - 228 chrysene
  - 252 benzopyrene
  - 278 anthanthracene

Gases trapped in fullerenes found in 4.6 billion year old Allende Meteorite (Luann Becker)
- time-of-flight mass spectrometry

ME 330 L804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amini S. Woods
Basic MALDI and Electrospray Ionization Theory

1. Desorption/Ionization and its Mechanisms
   a. chemical ionization processes
   b. preformed ions
   c. matching the charge sign

2. Matrix-Assisted Laser Desorption/Ionization (MALDI)
   a. basic mechanisms
   b. pulsed lasers and their wavelengths
   c. laser optics and instrumentation
   d. matrices
   e. mass spectra of proteins
   f. MALDI discourages the formation of multiply-charged ions; survivor theory
   g. fragmentation nomenclature for peptides
   h. masses of the 20 common amino acid “residues”

3. Electrospray Ionization (ESI)
   a. basic instrumentation and theory of operation
   b. mass spectra of proteins

4. Electrospray Ionization (ESI) continued.
   c. ionization and fragmentation of carbohydrates
   d. fragmentation nomenclature for carbohydrates
   e. masses of the common carbohydrate residues
   f. ABEE and other derivatives
   g. examples of ESI mass spectra of carbohydrates

5. MALDI and Electrospray: when to use each.

5. MS and MS/MS: observing fragmentation.
1. Desorption/Ionization and its Mechanisms

Both electron impact and chemical ionization required that a sample molecule be first volatilized and then ionized (as a separate event) in the gas phase. Thus, these methods are limited to relatively small volatile molecules.

In desorption/ionization methods the movement of molecules from the condensed (either solid or solution) phase to the gas phase and ionization are (operationally) indistinguishable events. On a microscopic scale, ions may pre-exist in the condensed phase prior to desorption (preformed ions) or be ionized immediately upon entering the gas phase (chemical ionization).

a. chemical ionization processes

Amino acids or peptides entering the gas phase would encounter an ample number of proton donors originating from the matrix (MALDI) or solution (ESI) to be chemically ionized:

\[
\begin{align*}
\text{H}_2\text{N} & - \text{C} - \text{H} - \text{C} - \text{OH} \\
+ & \quad \text{H}^+ \\
\rightarrow & \quad \text{H}_3\text{N} - \text{C} - \text{H} - \text{C} - \text{OH}
\end{align*}
\]

However, amino acids and peptides exist as zwitterions in solution and, since either positive (protonated) and negative (deprotonated) species can be recorded by a mass spectrometer, one may be simply desorbing ions already formed:

\[
\begin{align*}
\text{H}_2\text{N} & - \text{C} - \text{O}^- \\
\rightarrow & \quad \text{H}_2\text{N} - \text{C} - \text{OH} \\
\text{H}_3\text{N} & - \text{C} - \text{OH} \\
\rightarrow & \quad \text{H}_3\text{N} - \text{C} - \text{OH}
\end{align*}
\]

b. preformed ions

Quaternary ammonium ions are clearly preformed. Basic residues (such as arginine and lysine) in peptides are also likely to result in direct desorption of preformed ions.

\[
\begin{align*}
\text{H}_3\text{C} - \text{N} - \text{CH}_3 & \\
\text{Cl}^- & \\
\rightarrow & \quad \text{H}_3\text{C} - \text{N} - \text{CH}_3
\end{align*}
\]

Most likely, both preformed and chemical ionization processes occur, as well as chemical reactions in the condensed phase/gas phase interface (selvedge).
c. matching the charge sign

Mass spectrometers can be made to record either positive or negative ions by making the source voltage positive or negative, respectively. The recording mode should match the charge sign of the analyte:

**peptides** are generally best analyzed as positive ions, particularly those containing arginine or lysine; because they are zwitterions they are best analyzed at low pH

**phosphorylated or sulfated peptides** may be analyzed in the -ve ion mode

**oligonucleotides**, which contain a large number of anionic phosphate groups may nonetheless be best analyzed as positive ions at low pH, in order to promote the formation of singly-charged species

**fatty acids** are best analyzed as fatty acyl anions

**carbohydrates** are more easily protonated than deprotonated, although they may be observed as their oxyanions

In general ions are all even-electron and follow the fragmentation patterns typical of chemically ionized species

---

### 2. Matrix-Assisted Laser Desorption/Ionization

The technique which we now know as matrix-assisted laser desorption/ionization (MALDI) was developed simultaneously in two laboratories in 1987. The first report of high mass ions (above m/z 10,000) was a paper presented by Koichi Tanaka of the Shimadzu Corporation (Kyoto, Japan) at the *Second Japan-China Joint Symposium on Mass Spectrometry*, held September 15-18, 1987 in Takarazuka, Japan. Using a pulsed N₂ laser (337 nm) and a time-of-flight mass spectrometer equipped with a coaxial reflectron, they recorded molecular ions at m/z 34,529 from carboxypeptidase-A dissolved in a slurry of glycerol and an ultra-fine metal powder. In addition, they reported a mass spectrum of lysozyme (MW 14,307) containing multimeric ions up to the pentamer recorded at m/z 71,736. At the same time, Michael Karas and Franz Hillenkamp from the University of Muenster (Germany) had developed a matrix-assisted technique using a frequency-quadrupled (266 nm) Q-switched Nd:YAG laser to desorb intact molecular ions from proteins dissolved in matrix solution containing nicotinic acid. Their first high mass results were reported at the *International Mass Spectrometry Conference* (IMSC) in Bordeaux, France in August 1988, and included molecular ions for bovine serum albumin observed in their mass spectrum at m/z 66,750. Results from both of these groups were first published in 1988, followed by a number of other reports by Hillenkamp and Karas for proteins with molecular weights in excess of 100 kDa.


2. *Matrix-Assisted Laser Desorption/Ionization*

**a. basic mechanisms**

1. The matrix absorbs UV or IR energy from the pulsed laser

2. The matrix ionizes and dissociates; it undergoes a phase change to a supercompressed gas; charge is passed to some of the analyte molecules

3. Matrix expands at supersonic velocity; additional analyte ions are formed in the gas phase; ions are entrained in the expanding plume

**b. pulsed lasers and their wavelengths**

<table>
<thead>
<tr>
<th>Type</th>
<th>Wavelength</th>
<th>Laser</th>
</tr>
</thead>
<tbody>
<tr>
<td>IR</td>
<td>10.6 microns</td>
<td>CO$_2$</td>
</tr>
<tr>
<td>IR</td>
<td>2.94 microns</td>
<td>Er:YAG</td>
</tr>
<tr>
<td>IR</td>
<td>1.06 microns</td>
<td>Nd:YAG</td>
</tr>
<tr>
<td>VIS</td>
<td>532 nm</td>
<td>Nd:YAG (2$^{nd}$ harmonic)</td>
</tr>
<tr>
<td>UV</td>
<td>355 nm</td>
<td>Nd:YAG (3$^{rd}$ harmonic)</td>
</tr>
<tr>
<td>UV</td>
<td>337 nm</td>
<td>N$_2$</td>
</tr>
<tr>
<td>UV</td>
<td>266 nm</td>
<td>Nd:YAG (4$^{th}$ harmonic)</td>
</tr>
</tbody>
</table>

Ions may be preformed and released by the expanding plume

And, the matrix supplies proton donors for gas-phase chemical ionization

---


---

*ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research*, Robert J. Cotter and Amina S. Woods
c. laser optics and instrumentation

Laser optics usually includes:
- variable attenuator, generally rotating optical density filter
- mirror
- lens (material dependent on wavelength) with 1-5 inch focal length
- window to vacuum chamber (material dependent on wavelength)

Laser pulse width: 300 ps to 3 nanoseconds
Laser energy: 10 microjoules to 10 millijoules (depending on pulse width)
Laser power density: $10^6 - 10^7$ watts/cm$^2$

---

d. matrices

<table>
<thead>
<tr>
<th>Chemical Name</th>
<th>Molecular Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>nicotinic acid (NA)</td>
<td><img src="image" alt="Nicotinic Acid" /></td>
</tr>
<tr>
<td>caffeic acid (CA)</td>
<td><img src="image" alt="Caffeic Acid" /></td>
</tr>
<tr>
<td>sinapinic acid (SA)</td>
<td><img src="image" alt="Sinapinic Acid" /></td>
</tr>
<tr>
<td>3,4-dihydroxycinnamic acid</td>
<td><img src="image" alt="3,4-Dihydroxycinnamic Acid" /></td>
</tr>
<tr>
<td>3,5-dimethoxy-4-hydroxycinnamic acid</td>
<td><img src="image" alt="3,5-Dimethoxy-4-Hydroxycinnamic Acid" /></td>
</tr>
<tr>
<td>gentisic acid (DHBA)</td>
<td><img src="image" alt="Gentisic Acid" /></td>
</tr>
<tr>
<td>2,5-dihydroxybenzoic acid</td>
<td><img src="image" alt="2,5-Dihydroxybenzoic Acid" /></td>
</tr>
<tr>
<td>3-hydroxypicolinic acid (HPA)</td>
<td><img src="image" alt="3-Hydroxypicolinic Acid" /></td>
</tr>
<tr>
<td>α-cyano-4-hydroxycinnamic acid (CHCA)</td>
<td><img src="image" alt="α-Cyano-4-Hydroxycinnamic Acid" /></td>
</tr>
</tbody>
</table>

---
e. mass spectra of proteins
MALDI mass spectra of proteins obtained on a 4-inch endcap time-of-flight mass spectrometer. Mass resolution ≈ 500

f. MALDI discourages the formation of multiply-charged ions; survivor theory
• multiply-charged ions less stable in gas phase, though charge-state rises with increasing mass
• multiply-charged ions less easily desorbed from the surface because of electrostatic attraction
• charge-state of desorbed ions are reduced by collision in expanding plume. We see the surviving ions.

IR MALDI mass spectrum of cytochrome C (MW=12,327 Da)
UV MALDI mass spectrum of bovine serum albumin (MW=66,429 Da)

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods

How much resolution do you need?
Theoretical peak shapes for a peptide with an elemental composition of $C_{900}H_{1323}O_{294}N_{247}S_6$ with a molecular weight = 20,488.5

Resolution = 500
Resolution = 1,000
Resolution = 10,000
Resolution = 20,000

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
e. mass spectra of proteins

High resolution mass spectrum obtained on a Kratos AXIMA time-of-flight mass spectrometer with curved-field reflectron, with expansion of molecular ion region for human insulin.

Mixture of sheep, bovine, porcine and human insulins

High resolution mass spectrum obtained on a Kratos AXIMA time-of-flight mass spectrometer with curved-field reflectron, with expansion of molecular ion region for human insulin.

Rs = 12979

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods


g. fragmentation nomenclature for peptides

C-terminal ions: charge is retained on the carboxy terminal fragment

\[
\begin{align*}
\text{H}_3\text{N} & \quad \text{R}_1 \quad \text{C} \quad \text{N} \quad \text{C} \quad \text{O} \\
& \quad \text{R}_2 \\
\end{align*}
\]

\[
\begin{align*}
\text{H}_3\text{N} & \quad \text{R}_1 \quad \text{C} \quad \text{N} \quad \text{C} \quad \text{O} \\
& \quad \text{R}_2 \\
\end{align*}
\]

\[
\begin{align*}
\text{H}_3\text{N} & \quad \text{R}_1 \quad \text{C} \quad \text{N} \quad \text{C} \quad \text{O} \\
& \quad \text{R}_2 \\
\end{align*}
\]

\[
\begin{align*}
\text{H}_3\text{N} & \quad \text{R}_1 \quad \text{C} \quad \text{N} \quad \text{C} \quad \text{O} \\
& \quad \text{R}_2 \\
\end{align*}
\]

\[
\begin{align*}
\text{H}_3\text{N} & \quad \text{R}_1 \quad \text{C} \quad \text{N} \quad \text{C} \quad \text{O} \\
& \quad \text{R}_2 \\
\end{align*}
\]

- \text{NH}_3

- \text{y-ion}

- \text{z-ion}

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
g. fragmentation nomenclature for peptides

**N-terminal ions:** charge is retained on the amino terminal fragment

\[
\text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+ \quad \rightarrow \quad \text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+ \quad \rightarrow \quad \text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+ \quad \rightarrow \quad \text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+
\]

**b-ion**

\[
\text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+ \quad \rightarrow \quad \text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+ \quad \rightarrow \quad \text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+
\]

**a-ion**

\[
\text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+ \quad \rightarrow \quad \text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+ \quad \rightarrow \quad \text{H}_2\text{N}-\text{R}_1\text{H}-\text{R}_2\text{H}-\text{R}_3\text{H}-\text{COOH} \quad + \quad \text{H}^+
\]

**y-ions** for the peptide \textit{DAEFR} are calculated by summing the masses to the right of each cleavage point including the mass of a hydrogen ion:

\[
\begin{align*}
1 & \quad 115.1 & \quad 71.1 & \quad 129.1 & \quad 147.1 & \quad 156.2 & \quad 17 & \quad 1 \\
\text{H} & \quad \text{Asp} & \quad \text{Ala} & \quad \text{Glu} & \quad \text{Phe} & \quad \text{Arg} & \quad \text{OH} & \quad + \quad \text{H}^+
\end{align*}
\]

\[
\begin{align*}
+\text{H} & \quad +\text{H} & \quad +\text{H} & \quad +\text{H} & \quad +\text{H} & \quad +\text{H}
\end{align*}
\]
### h. masses of the 20 common amino acid “residues”

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
<td>71</td>
<td>71.079</td>
<td>71.079</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg</td>
<td>R</td>
<td>156</td>
<td>156.188</td>
<td>156.188</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>N</td>
<td>114</td>
<td>114.104</td>
<td>114.104</td>
</tr>
<tr>
<td>Aspartic Acid</td>
<td>Asp</td>
<td>D</td>
<td>115</td>
<td>115.089</td>
<td>115.089</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>C</td>
<td>103</td>
<td>103.143</td>
<td>103.143</td>
</tr>
<tr>
<td>Glutamic Acid</td>
<td>Glu</td>
<td>E</td>
<td>129</td>
<td>129.116</td>
<td>129.116</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln</td>
<td>Q</td>
<td>128</td>
<td>128.131</td>
<td>128.131</td>
</tr>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>G</td>
<td>57</td>
<td>57.052</td>
<td>57.052</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td>H</td>
<td>137</td>
<td>137.141</td>
<td>137.141</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Ile</td>
<td>I</td>
<td>113</td>
<td>113.160</td>
<td>113.160</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
<td>113</td>
<td>113.160</td>
<td>113.160</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys</td>
<td>K</td>
<td>128</td>
<td>128.175</td>
<td>128.175</td>
</tr>
<tr>
<td>Methionine</td>
<td>Met</td>
<td>M</td>
<td>131</td>
<td>131.197</td>
<td>131.197</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phe</td>
<td>F</td>
<td>147</td>
<td>147.177</td>
<td>147.177</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
<td>97</td>
<td>97.117</td>
<td>97.117</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>S</td>
<td>87</td>
<td>87.078</td>
<td>87.078</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr</td>
<td>T</td>
<td>101</td>
<td>101.105</td>
<td>101.105</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>Trp</td>
<td>W</td>
<td>186</td>
<td>186.214</td>
<td>186.214</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td>Y</td>
<td>163</td>
<td>163.176</td>
<td>163.176</td>
</tr>
<tr>
<td>Valine</td>
<td>Val</td>
<td>V</td>
<td>99</td>
<td>99.133</td>
<td>99.133</td>
</tr>
</tbody>
</table>

**b-ions** for the peptide **DAEFR** are calculated by summing the masses to the left of each cleavage point:

\[
\begin{align*}
115.1 + 1 + 71.1 + 129.1 + 147.2 + 156.2 + 17 + 1 & \rightarrow m/z = 637.7 \\
71.1 + 1 + 129.1 + 147.2 + 156.2 + 17 + 1 & \rightarrow m/z = 522.6 \\
129.1 + 1 + 147.2 + 156.2 + 17 + 1 & \rightarrow m/z = 451.5 \\
147.2 + 1 + 156.2 + 17 + 1 & \rightarrow m/z = 322.4 \\
156.2 + 1 + 17 + 1 & \rightarrow m/z = 175.2
\end{align*}
\]

---

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
**a-ions** for the peptide **DAEFR** are **b-28**:

![Diagram of peptide fragmentation](image)

**CID mass spectra of peptides:**

**NFNRHLHFTLVKDR** and **LLSYDDEAFFRDVAK**


**Homework problem:**
calculate the masses of **b** and **y** ions for the two peptides shown and compare with results obtained in their mass spectra.
3. Electrospray Ionization (ESI)

a. basic instrumentation and theory of operation

**Introduced by John Fenn:**

**Based on an ion evaporation model of Iribarne and Thomson:**
b. mass spectra of proteins

Multiply-charged ions are favored in ESI mass spectra, with ions having the formula:

\[(M + nH)^+n\]

Average \(m/z\) are around 1000, which is compatible with quadrupole mass spectrometers.


![Mass spectrum of proteins](image)

**Fig. 1.** ESI mass spectrum of equine myoglobin. Peaks \(m/z\) 707.25 through \(m/z\) 1305 are centered values for these multiply protonated \((24+\) through \(13+\) ions). The simultaneous relation of one pair, \(m/z\) 898.25 and \(m/z\) 1060.7, is illustrated where the integer value of \(n\) is 17. Calculation of \(M\), as discussed in the text.

---

c. ionization and fragmentation of carbohydrates

![Carbohydrate structures](image)

Oxonium ion

**Y-ion**

**B-ion**

**C-ion**

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
d. fragmentation nomenclature for carbohydrates


e. masses of the common carbohydrate residues

Similar to the amino acids, these residue masses are the masses of monomers minus water. The mass of a single water is added to any collection of residues (sugar, peptide, phosphate, etc.) for a biological molecule.

---

**e. masses of the common carbohydrate residues**

<table>
<thead>
<tr>
<th>monosaccharide</th>
<th>abbreviations</th>
<th>monoisotopic mass</th>
<th>average mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>D-fructose</td>
<td>Fru</td>
<td>162.052</td>
<td>162.143</td>
</tr>
<tr>
<td>L-fucose</td>
<td>Fuc / F</td>
<td>146.058</td>
<td>146.14</td>
</tr>
<tr>
<td>D-galactose</td>
<td>Gal / GL</td>
<td>162.052</td>
<td>162.143</td>
</tr>
<tr>
<td>N-acetyl-D-galactosamine</td>
<td>GalNAc</td>
<td>203.079</td>
<td>203.19</td>
</tr>
<tr>
<td>D-glucose</td>
<td>Glc / G</td>
<td>162.052</td>
<td>162.143</td>
</tr>
<tr>
<td>D-glucuronic acid</td>
<td>GlcA</td>
<td>176.032</td>
<td>176.126</td>
</tr>
<tr>
<td>N-acetyl-D-glucosamine</td>
<td>GlcNAc</td>
<td>203.079</td>
<td>203.19</td>
</tr>
<tr>
<td>hexose (non-specific)</td>
<td>Hex</td>
<td>162.052</td>
<td>162.143</td>
</tr>
<tr>
<td>D-mannose</td>
<td>Man / M</td>
<td>162.052</td>
<td>162.143</td>
</tr>
<tr>
<td>N-acetyl-D-mannosamine</td>
<td>ManNac</td>
<td>203.079</td>
<td>203.19</td>
</tr>
<tr>
<td>N-acetyleneuraminic acid</td>
<td>NANA/NeuAc/SA</td>
<td>291.095</td>
<td>291.26</td>
</tr>
<tr>
<td>Sialic acid</td>
<td>SA</td>
<td>291.095</td>
<td>291.26</td>
</tr>
</tbody>
</table>
f. ABEE and other derivatives

Glycopeptides: glycosylation is one of the most common post-translational modifications of peptides. **N-linked oligosaccharides** are generally attached via a terminal GlcNAc to an asparagine residue found in the consensus sequence: **Asn-X-Ser/Thr**.

![Image of N-linked oligosaccharide structure]

O-linked oligosaccharides are attached to serine or threonine residues by an acyl linkage formed between two hydroxyl groups (loss of H₂O), such as the linkage between GalNAc and serine above.

Neutral carbohydrates: the structures of carbohydrates not attached to peptides and not containing charged (sialic acid, KDO, etc.) groups or hydrophobic (fatty acyl) groups are often derivatized to improve ionization efficiency. One example is the **amino-isobuteric acid ethyl ester (ABEE)** and related derivatives:

![Image of ABEE derivatization]

g. examples of ESI mass spectra of carbohydrates

Electrospray ionization mass spectrum of high mannose sugars derivatized with seven different amines.

---

Electrospray ionization CID mass spectrum of the [M+Na]+ ion from the 3-AQ derivative of (GlcNAc)_2Man_3.
4. MALDI and Electrospray: when to use each.

MALDI: singly-charged ions best used with a high mass range analyzer such as the TOF mass analyzer. Singly-charged ions are an advantage where there is a complex mixture accompanied by significant fragmentation; an example is microheterogenous mixtures of glycopeptides.

ESI: multiply charged ions; can be used with quadrupoles and quadrupole ion traps. These instruments are more readily configured as tandem mass spectrometers for mass-selecting and fragmenting single components of a mixture.

MALDI: a solid-phase technique that will be utilized for high throughput microarrays on silicon chips, imaging of tissue or selection of individual cells or microorganisms

ESI: a liquid techniques compatible with on-line chromatographic (reversed-phase HPLC, anion exchange, etc.) chromatography and capillary electrophoresis.

5. MS and MS/MS: observing fragmentation.

Fragmentation in normal MS: structurally informative fragmentation is sometimes observed in normal mass spectra, but generally only for relatively small compounds; cannot distinguish fragmentation from mixtures.

In-source decay: in TOF instruments observable fragmentation is increased by lengthening the time ions spend in the source; again, cannot distinguish fragmentation from mixtures.

Post-source decay: in TOF instruments this provides the opportunity to mass-select molecular ions and to observe fragmentation specific to each molecular ion.

Tandem mass spectrometer: quadrupoles, ion traps, Fourier transform mass spectrometers and hybrids (such as the QTOF) provide the opportunity for mass-selection of molecular ions and observation of their specific fragment ions; in addition, they increase the extent of fragmentation through collision-induced dissociation (CID).
**Instrumentation: TOF, Ion Traps, QTOF and FTICR**

1. **Time-of-Flight Mass Spectrometers**
   a. basic principles and equations
   b. spatial/energy distributions and their effects on mass resolution
   c. single-stage and dual-stage reflectrons
   d. time-lag focusing/delayed extraction
   e. orthogonal extraction/acceleration TOF
   f. quadrupole ion injection: the QTOF

2. **Quadrupole Mass Spectrometers**
   a. basic instrumentation and stability diagram

3. **Quadrupole Ion Traps**
   a. basic instrumentation
   b. mass selective instability mode
   c. resonance ejection mode

4. **Fourier Transform Mass Spectrometers**

5. **Conclusions**

---

**1. Time-of-Flight Mass Spectrometers**

a. basic principles and equations

Short source region (s) with a high field for extracting the ions

Longer field-free drift region (D)

Flight times follow a simple square root dependence on mass

\[
 t = \left( \frac{m}{2eV} \right)^{1/2} D
\]
b. spatial/energy distributions and their effects on mass resolution

The actual time-of-flight is more complex:

\[ t = \frac{(2m)^{1/2}}{eE} \left[ \left( U_0 + eE_{\text{th}} \right)^{1/2} + U_0^{1/2} \right] + \frac{(2m)^{1/2} D}{2(U_0 + eE_{\text{th}})^{1/2}} + t_0 \]

- Distribution in initial position in the source
- Distribution in time of ion formation
- Turn-around time
- Initial kinetic energy distribution
- Time in ion source
- Time in flight tube

Mass resolution is reduced by the spread in initial ion kinetic energy prior to acceleration.

How is it improved?
(a) low laser power
(b) high accelerating voltage
(c) reflectron

FIGURE 2.3: Two ions with different initial kinetic energies, and the effects of kinetic energy distributions on the mass spectrum. (In this and subsequent figures, the differences in the lengths of the arrows are intended to represent differences in ion velocities.)
b. spatial/energy distributions and their effects on mass resolution

Mass resolution is reduced by the spread in initial positions of the ions in the ion source.

How is it improved?
(a) thin sample
(b) dual stage extraction to move the space-focus plane

![Diagram of ion source and drift region]

**FIGURE 3.4** Two ions formed in different locations with respect to the focusing plane, and the effects of spatial distributions on the mass spectrum.

---

c. single-stage and dual-stage reflectrons

Mass resolution is improved with a reflectron

**Single-stage**

\[ i = \left( \frac{m}{2eV} \right)^{2/3} (L_1 + L_2 + 4d) \]

**Dual-stage**

![Diagram of single-stage and dual-stage reflectrons]

---

ME 330L.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
d. time-lag focusing/delayed extraction

**Time-lag focusing, pulsed extraction and delayed extraction methods:**
- the first extraction field is pulsed
- a narrow range of mass is focused
- delay time and/or pulse amplitude are mass-dependent

**Pulsed extraction**

\[ E_1 \]

Heavy ions focus behind the detector

\[ E_2 \]

Lighter ions focus before the detector

\[ E_D = 0 \]


**Delayed extraction with MALDI**

**Instrumental configurations**


d. time-lag focusing/delayed extraction

Improved mass resolution for protein molecular weight measurements using delayed extraction


FIGURE 7.28: Delayed-extraction MALDI mass spectra of the molecular ion region of the peptides (a) bradykinin (FMV, MW = 1066.8) obtained at an extraction pulse potential of 0.9 kV; (b) LYS-LYS-PRF, (FMV, MW = 1259.7) at 1.0 kV; (c) Asp-Val-Lys-Lys-Leu-Glu-Asp-Enk, (FMV, MW = 1545.3) at 1.3 kV; (d) Asp-Asp-Met-Leu-Lys-Val-Leu-Asp-Enk, (FMV, MW = 1846.9) at 1.75 kV; (e) Asp-Asp-Met-Leu-Lys-Val-Leu-Asp-Enk, (FMV, MW = 2146.9) at 2.05 kV; and (f) Asp-Asp-Met-Leu-Lys-Val-Leu-Asp-Enk, (FMV, MW = 2577.7) at 2.3 kV (Reproduced with permission from reference 41).
d. time-lag focusing/delayed extraction

Resolution is even better when combined with a reflectron

Resolution is mass-dependent: tuned on glucagon-like peptide molecular ion

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods

c. single-stage and dual-stage reflectrons
d. time-lag focusing/delayed extraction

Insulin B-chain using delayed extraction/time-lag focusing on a linear instrument (no reflectron)

Bovine insulin using delayed extraction/time-lag focusing on a reflectron instrument

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
e. orthogonal extraction/acceleration TOF


What does it do?
- improves mass resolution: ions are extracted at right angles to the initial velocity distribution
- improves duty cycle: ions are stored between pulsed extraction cycles

f. quadrupole ion injection: the QTOF

Orthogonal acceleration from the ion storage region

Quadrupole ion guide operated in rf-only mode (passes all masses)

What does it do?
- improves mass resolution: collisions collimate the beam along the axis; reduces spatial spread.
- collisions reduce kinetic energies, provides an interface to high pressure source region
- makes mass analyzer independent of source; can use any source
f. quadrupole ion injection: the QTOF

Electron impact mass spectrum of PFTBA

Signal averaging of 200 transients
2GHz sampling
Repetition rate of 4 kHz
Accumulation time 0.05 s
Mass resolution for \( m/z \) 264 is increased

Berkout, V.; Cotter, R.J., unpublished results

Table 2. Determination of the number of N and Q in peptide LEIPFEDCDXN (X = Q or K) by comparison between the experimentally determined mass of 1437.8194a and the theoretically possible masses

\[
\begin{array}{cccc}
\text{Number of} & \text{M}_{\text{calc}} \text{ (Da)} & \text{Dev} \text{ (Da)} \\
4 & 1437.8184 & +0.0 \\
3 & 1437.8287 & -5.7 \\
2 & 1437.8390 & -3.0 \\
1 & 1437.8493 & -2.3 \\
0 & 1437.8596 & -3.6 \\
\end{array}
\]

Figure 2. ESI-QTOF mass spectra of peptide LEIPFEDCDXN (X = Q or K) showing the tryptically protonated molecule peak. The spectrum is internally calibrated with the tryptically charged molecular peaks of angiotensin I and neurokinin. The deconvoluted peaks show the mass for the uncharged molecule.

(Note: used here in MS mode only; MS/MS mode described in tandem lecture)
2. **Quadrupole Mass Spectrometers**

   a. basic instrumentation and stability diagram

   **dc voltage:** $U$

   **rf voltage:** $V_0 \cos(\omega t)$

   \[
   a = \frac{8eU}{mr_0^2 \omega^2} \quad q = \frac{4eV_0}{mr_0^2 \omega^2}
   \]

   If $U/V_0 = 0.167$, ions of $m/z$ are transmitted (have stable trajectories) when:

   \[
   m/z = 0.136 V_0/r_0^2 \omega^2
   \]

   where $f$ is 1MHz

   **rf-only mode** transmits all ions and is used for quadrupole injection

---

3. **Quadrupole Ion Traps**

   a. basic instrumentation

   b. mass selective instability mode

   c. resonance ejection mode

   **Trapping cycle:** fixed amplitude of the fundamental rf voltage

   **Ion ejection cycle:** remove unwanted ions by symmetric or asymmetric pulses on endcaps; *stored waveform inverse Fourier transform (SWIFT)* or filtered noise field techniques

   **Mass analysis cycle:**
   - mass selective instability mode: scan fundamental rf voltage; mass range approx. 650
   - resonance ejection mode: set supplemental rf voltage on endcaps and then scan fundamental rf voltage

   **MS/MS mode:** low amplitude supplemental rf voltage applied to endcaps

---

*ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods*
Mathieu parameters:

\[ q_z = \frac{8eV}{m(r_0^2 + 2z_0^2)\Omega_0^2} \]

\[ a_z = \frac{-16eU}{m(r_0^2 + 2z_0^2)\Omega_0^2} \]

**Mass selective instability mode:** If dc voltage on the endcaps is zero, then scan along the \( a_z \) line (by varying the rf voltage); ion ejection occurs at the stability boundary when \( a_z = 0.908 \)

The mass ejected is then given by:

\[ \frac{m}{z} = \frac{8V}{(r_0^2 + 2z_0^2)\Omega_0^2} \]

Where \( z \) is the number of charges and \( \Omega_0 \) is the angular drive frequency

\( (\Omega_0/2\pi \approx 1.1 \text{ MHz}) \)

---

**Resonance ejection mode:**

A supplementary rf voltage is applied to the endcaps

Scan the fundamental rf voltage on the ring electrode

Ions are ejected “through a hole in the stability region”

**Extension of mass range through axial modulation**

Supplementary rf = 69.9 kHz

\( q_{\text{eject}} = 0.182 \)

\[ m/z = (0.91/0.182) \times 650 = 5 \times 650 = 3,250 \]

Supplementary rf = 35.2 kHz

\( q_{\text{eject}} = 0.091 \)

\[ m/z = (0.91/0.091) \times 650 = 10 \times 650 = 6,500 \]

---

4. Fourier Transform Mass Spectrometers

- pulsed instruments that were immediately compatible with MALDI
- very high mass resolution
- ability to do multiple tandem-in-time experiments, MS^n
- with quadrupole ion guides, can accommodate ESI


Electrospray FTMS Analysis of Oligosaccharides

5. Conclusions

When MALDI and ESI were first introduced:

- **MALDI worked best on the TOF mass spectrometer:** It is a pulsed technique which required a very high mass range.

- **ESI worked best on quadrupoles (Q) and quadrupole ion traps (ITMS):** It is a continuous technique compatible with a scanning instrument (Q) and the multiple charges put it with the limited mass range of these two analyzer.

Quadrupole injection and orthogonal acceleration make ESI compatible with TOF mass spectrometers.

Quadrupole injection TOF mass spectrometers can also be used as tandem (MS/MS) mass spectrometers, e.g. the QTOF.

Ion trap mass spectrometers can provide multiple tandem-in-time measurements, or MS^n.

MALDI can be used on QTOFS and ion traps to take advantage of MS/MS capabilities, but quadrupole fields limit the m/z range.
Sample Preparation for MALDI and LC/MS

1. General
2. Sample Preparation for MALDI
   Methods
   Matrix selection
   Sample cleanup for MALDI
   Matrix cleanup
   Tips for better sample preparation
3. Sample Preparation for ESI
   Reverse phase chromatography
   Liquid-liquid extraction
   Immunoaffinity extraction (IAE)

1. General Sample Preparation

ESI and MALDI have made analysis of trace amounts of biomolecules a routine occurrence.

The information obtained from the spectrum of a compound is usually dependent on how the sample was prepared. Certain simple rules can greatly improve results.

Avoid or minimize the sample content of the following ingredients:

- Salts in general and sodium and phosphate salts in particular. If buffers are needed try to use ammonium salts at concentrations of 5-25 mM.

- Detergents especially polyethylene glycols (e.g. Triton, Nonidet P-40). If detergents are needed use CHAPS (3-[3-Cholamidopropyl)dimethyl ammonio]-1-propanesulfonate) or Octyl b-D-Glucopyranoside.

- Viscous compounds (DMSO, glycerol).
2. Sample Preparation for MALDI

a. Dried drop
   Apply 0.3 µl of analyte solution to the probe
   Add 0.3 µl matrix solution
   Let dry at RT

b. Thin film
   Dissolve matrix in acetone (fast evaporating solvent) with 1-2% water and apply
   to probe. Rinse dry layer with water.
   Apply sample.

c. Sandwich
   Prepare a dry layer as in (b), then apply a mixture of analyte and matrix to the dry layer. The base layer serves as seed crystals.

2. Sample Preparation for MALDI (cont)

Matrix Selection:

a-cyano-4-hydroxycinnamic acid for peptides (CHCA)
Sinnapinic acid (3,5-dimethoxy-4-hydroxy cinnamic acid) for proteins (SA)
Gentisic acid for lipids and sugars (GA)
6-Aza-2-thithymine (ATT) to detect non-covalent complexes
Hydroxypicolinic Acid and ATT for oligonucleotides

Sample Clean Up for MALDI:

Desalting, on the probe clean up:
Add 1-2 µl deionized water to the dried sample spot
After 2-3 minutes aspirate the water.
This simple and fast rinsing will remove contaminants such as salts and small molecules.
Should only be used with water insoluble matrices.

HPLC: will be discussed in detail later.
Sample Cleanup for MALDI (continued):

Polyethylene and polypropylene membranes:

Activate membrane by depositing 2 µl methanol.  
While membrane is still wet add 1 µl sample and allow it to dry.  
Rinse spot with 50-100 µl of 50% methanol solution. Carefully aspirate solution.

When membrane dries, add 1 µl matrix solution, let it dry, affix membrane on probe with two sided tape and acquire spectrum.


Example showing cleanup of SDS
2. Sample Preparation for MALDI (cont)

Matrix Clean up:

a. Water insoluble matrices:

Fill lower third of an eppendorf tube with matrix, add water, vortex, centrifuge, and discard.

This will remove a significant amount of sodium and potassium ions.

b. Water soluble matrices:

Can be dissolved and desalted by adding cation exchange beads charged with ammonium ions.

Tips For better sample preparation

• Vacuum dry sample-matrix spot on the target before acquiring a spectrum.

• Raise the pH of sample-matrix mixture by adding ammonium citrate or bicarbonate.

• To decrease suppression, add a drop of ammonium sulfate to the sample before matrix addition.
3. Sample Preparation for ESI

Suitable sample preparation is an important prerequisite for obtaining reliable results.

Sample can be injected into an ESI source using a Harvard pump. However more often than not liquid chromatography (HPLC) or Capillary zone electrophoresis are coupled with ESI to directly analyze biomolecules from solutions.

The presence of non-volatile reagents such as alkylating agents, reductants and buffers used to extract and digest biological compounds, as well as contaminants such as polyacrylamide, SDS and various other detergents can lead to suboptimal mass spectrometric results.

Your aim should be to remove these substances and at the same time minimize analyte losses.

Reverse Phase Chromatography:

If using HPLC as the introduction system, the use of C4 or C18 columns will accomplish a dual purpose, it will remove most of the contaminants and separate the various biological components.

If an HPLC system is not available one can use pipet tip columns such as "ZipTip™" which can concentrate and purify 1 to 100 µl of sample.

Mixed phase Column such as C18+ cation exchange or C18+ anion exchange, can be very useful with biological samples that contain glycolipids in addition to peptides or proteins.
3. Sample Preparation for ESI (cont)

**Liquid-Liquid extraction:**

Mostly used for drug extraction to isolate drugs after cleavage of the biochemical component to which they are conjugated.

Followed by separation on a reverse phase column.

**Immunoaffinity extraction (IAE):**

IAE can be coupled with HPLC. The first column is packed with the IEA bound antibodies.

The second column is a trapping column and the third one is an analytical column. The sample is loaded, the analyte is bound, then the analyte is eluted off the column, and trapped onto the trapping column, then the analyte is back-flushed onto the appropriate analytical column.
Basic Chromatography: HPLC (RP) & CE

1. High performance liquid chromatography (HPLC)
2. Components of an HPLC system
3. Mechanism of reverse phase HPLC (RP HPLC)
4. Gradient Elution
5. Ion pairing agents
6. Organic modifiers
7. UV detection
8. Typical gradients
9. Other types of chromatography
10. Capillary electrophoresis (CE)

1. High performance liquid chromatography (HPLC)

The various techniques used for High Pressure Liquid Chromatography are:

-Reverse Phase Chromatography.
-Size Exclusion Chromatography.
-Ion Exchange (cation or anion) Chromatography.
-Hydrophobic interaction Chromatography.

The high resolving power of Reverse Phase [RP] Chromatography has made it the dominant mode of HPLC for both analytical and preparative separation of peptides and proteins, as well as other biomolecules.
1. HPLC (cont)

The term “Reverse Phase Chromatography” was used because RP is a form of partition chromatography where chemically bonded phase is hydrophobic or non-polar (e.g. octadecyl group), and the starting mobile phase (e.g. water) must be more polar than the stationary phase.

This is “reversed” from normal phase chromatography, where the stationary phase is polar or hydrophilic and the starting mobile phase is more non-polar or hydrophobic than the stationary phase, hence the term “Reverse Phase Chromatography”.

2. Components of an HPLC system

Basic Components of an HPLC System
3. **Mechanism of reverse-phase HPLC**

Peptides and proteins are adsorbed onto the hydrophobic surface of the column and remain there until the concentration of the organic modifier is high enough to elute the molecules from the hydrophobic surface. The elution order is related to the increasing hydrophobic nature of the solute, the more soluble a solute is in water or the more hydrophilic the solute, the faster it will be eluted.

<table>
<thead>
<tr>
<th>#</th>
<th>Peptide Sequence</th>
<th>MW</th>
<th>RT</th>
<th>Hydrophobicity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RGGGGIGIGK</td>
<td>871</td>
<td>10.5</td>
<td>-0.180</td>
</tr>
<tr>
<td>2</td>
<td>RGGGGIGGLGK</td>
<td>871</td>
<td>15.0</td>
<td>-0.250</td>
</tr>
<tr>
<td>3</td>
<td>RGGGGLGLGK</td>
<td>871</td>
<td>20.5</td>
<td>-0.320</td>
</tr>
</tbody>
</table>

Retention Time: The time between injection and the appearance of the peak maximum.
4. Gradient elution

Technique for decreasing separation time by increasing mobile phase strength over time during the chromatographic separation. Gradient can be linear or stepwise.

Binary, ternary and quaternary solvent gradients can be used. The most widely used is the linear binary gradient.

<table>
<thead>
<tr>
<th>Types of reversed phase Column:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>MW in Da</td>
</tr>
<tr>
<td>C\textsubscript{18}</td>
<td>&lt; 5000</td>
</tr>
<tr>
<td>C\textsubscript{8}</td>
<td>5000-10000</td>
</tr>
<tr>
<td>C\textsubscript{4}</td>
<td>&gt; 10000</td>
</tr>
</tbody>
</table>

5. Ion-pairing agents

Ion-pairing agents are ionic compounds that contain a hydrocarbon chain that imparts a certain hydrophobicity so that the ion pair can be retained on a reversed-phase column. Ion Pairing agents are added at concentrations of 0.05 to 0.2. All ion-pairing agents are potentially capable of ion-pairing with the positively charged basic residues of peptides or proteins, thus reducing hydrophilicity and increasing their retention time. Hydrophobic counterions such as TFA and HFBA in addition to ion-pairing with the positively charged solute also increase the affinity of the solute (peptide or protein) for the hydrophobic stationary phase. While hydrophilic counterions such as H\textsubscript{3}PO\textsubscript{4} following ion-pair formation with positive charged residues would be unlikely to interact with the stationary phase.
5. Ion-pairing agents (cont)

-Trifluoroacetic acid (TFA).
-Heptafluorobutyric acid (HFBA).
-Hexafluoroacetone (HFA).
-Phosphoric Acid.
-Hydrochloric Acid.
-Triethylamine Phosphate (TEAP).

6. Organic modifiers

Additive that changes the character of the mobile phase. In RP chromatography water is the weak solvent, and acetonitrile, the strong solvent is added gradually to generate a gradient.

Acetonitrile.
Isopropanol.
Methanol.
Ethanol

Acetonitrile is the most widely used organic modifier. Isopropanol is used either alone or in combination with acetonitrile to elute large or hydrophobic proteins.
7. **UV detection**

Peptides and proteins are detected by UV absorption at wavelength from 210-220 nm which detect the amide bond. The aromatic side chains of tyrosine, phenylalanine and tryptophan absorb light in the 250 to 290 nm ultraviolet range.

![Amide Bond](image)

8. **Typical gradients**

When dealing with an unknown mixture, the following is a good starting gradient

**Buffer A:** 0.1% TFA water  
**Buffer B:** 0.1% TFA acetonitrile

<table>
<thead>
<tr>
<th>Time</th>
<th>% A</th>
<th>% B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>75</td>
<td>30</td>
<td>70</td>
</tr>
<tr>
<td>85</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>95</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>96</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>120</td>
<td>100</td>
<td>0</td>
</tr>
</tbody>
</table>
9. Other types of chromatography

Size Exclusion chromatography:

Mainly used for very large proteins.

Cation exchange chromatography:

Very useful when peptides or proteins contain an inordinate number of
negatively charged residues (Asp and Glu). Also useful for proteins
containing posttranslational modifications that make them more
hydrophilic.

Mixed mode chromatography:

uses a mixture of reverse-phase and cationic sorbents.

10. Capillary electrophoresis (CE)

CE is a cross between gel electrophoresis and high pressure liquid
chromatography. Separation is based on size to charge ratio. It uses high
voltage generating electroosmotic [EOF] and electrophoretic flow of buffer
solutions and ionic species, respectively within the capillary.

The basic instrument is made of a fused silica capillary, a controllable
high voltage power supply, two electrode assemblies, two buffer
reservoirs, a UV detector and a data acquisition system. The ends of the
capillary are placed in the buffer. After filling the capillary with buffer
the sample can be introduced by
dipping the end of the capillary in the
sample solution. In CE nothing is
retained so the analogous term to
retention time is migration time.
10. CE (cont)

The fused silica capillaries have ionizable silanol group in contact with the buffer within the capillary. The pI of the silica is about 1.5. The degree of ionization is controlled by the pH of the buffer. The negatively charged wall attracts positively charged ions from the buffer, creating an electrical double layer. When voltage is applied across the capillary, cations in the diffuse portion of the double layer migrate in the direction of the cathode carrying water with them. The result is a net flow of buffer solution in the direction of the negative electrode.

\[ + \quad + \quad + \quad + \quad + \quad + \quad + \quad + \quad EOF \quad - \quad - \quad - \quad - \quad - \quad - \quad - \]

The EOF makes possible the simultaneous analysis of cations, anions and neutral species in a single analysis. At neutral to alkaline pH, the EOF is sufficiently stronger than electrophoretic migration, such that all species are swept towards the negative electrode.

The order of migration is: cations, neutrals, and anions

Buffers for capillary electrophoresis:

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH Range</th>
<th>Zwitterionic Buffer</th>
<th>pH Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phospate</td>
<td>1.14 - 3.14</td>
<td>MES</td>
<td>5.15 - 7.15</td>
</tr>
<tr>
<td>Acetate</td>
<td>3.76 - 5.76</td>
<td>PIPES</td>
<td>5.80 - 7.80</td>
</tr>
</tbody>
</table>

At high pH, the EOF is large and the peptide is negatively charged. Despite the peptide migration towards the positive electrode, the EOF is overwhelming and the peptide migrates towards the negative electrode. At low pH the peptide is positively charged EOF is very small. Thus peptide electrophoretic migration and EOF are towards the negative electrode. In silica capillaries most solutes migrate towards the negative electrode regardless of charge when the buffer pH is above 7.0. At acidic pH, most zwitterions and cations will also migrate towards the negative electrode.

Buffers for capillary electrophoresis:

<table>
<thead>
<tr>
<th>Buffer</th>
<th>pH Range</th>
<th>Zwitterionic Buffer</th>
<th>pH Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phospate</td>
<td>1.14 - 3.14</td>
<td>MES</td>
<td>5.15 - 7.15</td>
</tr>
<tr>
<td>Acetate</td>
<td>3.76 - 5.76</td>
<td>PIPES</td>
<td>5.80 - 7.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tricine</td>
<td>7.15 - 9.15</td>
</tr>
</tbody>
</table>
10. CE (cont)

**Advantage of CE:**

- Separation takes minutes rather than hours.
- Uses much less reagents.
- Better separation of peptide with similar hydrophobicity index.

**Disadvantage:**

- Requires much more skill and technical ability.
General Overview of Biological Applications

1. Molecular Weight Strategies for Proteins and Peptides
   a. Protein molecular weights
   b. Using enzymes to determine structure and sequence
   c. Peptide mapping: tryptic and other enzymatic digests
   d. Identification from protein/genome databases
   e. Amino acid sequencing using carboxypeptidases
   f. Amino acid sequencing using aminopeptidases
   g. Locating post-translational modifications
   h. For example: phosphorylation

2. Molecular Weight Strategies for Carbohydrates
   a. Carbohydrate heterogeneity
   b. Exoglycosidases and endoglycosidases
   c. Sequence and branching
   d. Chemical degradation

3. Fragmentation of Peptides and Glycoconjugates

1. Molecular Weight Strategies for Proteins and Peptides
   • strategies in which the molecular mass, or changes in molecular mass due to post-translational modifications or chemical/enzymatic reactions are utilized to determine structure
   • low laser power (in MALDI) or low skimmer voltages (in ESI) are used to reduce fragmentation so that only molecular ions are recorded
   • the mass spectrometer is operated in the MS mode, rather than the MS/MS mode used to record fragmentation
a. Protein molecular weights

Two ways to measure the molecular weight!

- SDS-PAGE depends upon shape and size as well as mass.
- MALDI mass spectrometry measures mass with an accuracy of 0.1% to 0.01%.

With this accuracy, it is difficult to identify a protein based only on its measured molecular weight.

However, it provides an important method to limit database identification using tryptic maps.


---

b. Using enzymes to determine structure and sequence

<table>
<thead>
<tr>
<th>endopeptidases</th>
<th>cleavage sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>trypsin</td>
<td>after K/R</td>
</tr>
<tr>
<td>endoproteinase Lys-C</td>
<td>after K</td>
</tr>
<tr>
<td>endoproteinase Glu-C</td>
<td>after E</td>
</tr>
<tr>
<td>endoproteinase Asp-N</td>
<td>before D</td>
</tr>
<tr>
<td>endoproteinase Arg C</td>
<td>after R</td>
</tr>
<tr>
<td>chymotrypsin</td>
<td>after F/W/Y/L</td>
</tr>
<tr>
<td>pepsin</td>
<td>after F/W/Y/L</td>
</tr>
<tr>
<td>thermolysin</td>
<td>before L/I/M/F/W</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>exopeptidases</th>
<th>selectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>carboxypeptidase A</td>
<td>stops at R/PX, sometimes at G/S/D/E</td>
</tr>
<tr>
<td>carboxypeptidase B</td>
<td>cleaves at R/K</td>
</tr>
<tr>
<td>carboxypeptidase P</td>
<td>cleaves PX/D/E, sometimes stops at S/G</td>
</tr>
<tr>
<td>carboxypeptidase Y</td>
<td>cleaves at PX/E, sometimes stops at K/R/S/G</td>
</tr>
<tr>
<td>aminopeptidase M</td>
<td>non-specific</td>
</tr>
<tr>
<td>leucine aminopeptidase</td>
<td>stops at or near K/R/P</td>
</tr>
</tbody>
</table>

b. Using enzymes to determine structure and sequence

<table>
<thead>
<tr>
<th>chemical reagents</th>
<th>cleavage sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>cyanogen bromide (CNBr)</td>
<td>after M homoserine -30.1 Da</td>
</tr>
<tr>
<td></td>
<td>homoserine lactone - 48.1 Da</td>
</tr>
<tr>
<td>BNPS-skatole or DMSO+HCl acid hydrolysis</td>
<td>after W D/P then random</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
\text{N} & \text{H} \\
\text{H} & \text{C \ O} \\
\text{H}_2\text{C} & \text{O} \\
\text{N} & \text{H} \\
\text{H} & \text{C \ C} \\
\text{CH}_2 & \text{CH}_2 \\
\text{S} & \text{CH}_3 \\
\end{align*}
\]

homoserine lactone

\[
\begin{align*}
\text{N} & \text{H} \\
\text{H} & \text{C \ CN} \\
\text{H} & \text{C \ H} \\
\text{CH}_2 & \text{CH}_2 \\
\end{align*}
\]

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods

---

c. Peptide mapping: tryptic and other enzymatic digests

**Mass balancing:** the sum of the molecular weights of \( n \) enzymatic fragments, minus \( n-1 \) water molecules, add up to the molecular weight of the intact protein.

**Example.** \( \beta \)-amyloid peptide (\( \beta \)A\(1-40\)): 

DAEFRHDSGYEVHHQKLVVFAEDVGSNKGAIIIGLMGVV

**tryptic digest:**

- **DAEFR**  \( MW = 636.7 \)  \( \beta \text{A}_{1-5} \) MH\(^+\) observed = 637.8
- **HDGYEVHHQK**  \( MW = 1336.5 \)  \( \beta \text{A}_{6-16} \) MH\(^+\) observed = 1337.1
- **LVFFAEDVGSNK**  \( MW = 1325.7 \)  \( \beta \text{A}_{17-28} \) MH\(^+\) observed = 1326.7
- **GAIIGLMVGVV**  \( MW = 1085.5 \)  \( \beta \text{A}_{29-40} \) MH\(^+\) observed = 1086.1

**cyanogen bromide:**

- **VGGVV**  \( MW = 429.6 \)  \( \beta \text{A}_{36-40} \) MH\(^+\) observed = 431.1
- **VGGVVIA**  \( MW = 613.8 \)  \( \beta \text{A}_{36-42} \) MH\(^+\) observed = 614.2
- **GAILLM**  \( MW = 673.9 \)  \( \beta \text{A}_{29-35} \) MH\(^+\) observed = 626.0

- **homoserine**  \( MW = 643.8 \)
- **homoserine lactone**  \( MW = 625.8 \)

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
d. Identification from protein/genome databases

Table 2. Nine Proteins Identified from HEL Cell CBB 2.D Gel

<table>
<thead>
<tr>
<th>Spot</th>
<th>enzyme</th>
<th>MW / pI</th>
<th>SwissProt access. No.</th>
<th>protein name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gl</td>
<td>trypsin</td>
<td>18012.6/7.68</td>
<td>PO5092</td>
<td>PPIase</td>
</tr>
<tr>
<td>G2</td>
<td>trypsin</td>
<td>26669.6/6.45</td>
<td>PO0098</td>
<td>AMP deaminase, isomerase</td>
</tr>
<tr>
<td>G3</td>
<td>trypsin</td>
<td>26669.6/6.45</td>
<td>PO0098</td>
<td>TIM</td>
</tr>
<tr>
<td>G8</td>
<td>trypsin</td>
<td>29032.8/4.75</td>
<td>P12324</td>
<td>tropomyosin, cytoskeletal type</td>
</tr>
<tr>
<td>G10</td>
<td>trypsin</td>
<td>32575.2/4.64</td>
<td>PO6748</td>
<td>NPM</td>
</tr>
<tr>
<td>G11</td>
<td>trypsin</td>
<td>32575.2/4.64</td>
<td>PO6748</td>
<td>β-actin</td>
</tr>
<tr>
<td>G12</td>
<td>trypsin</td>
<td>41737.0/5.29</td>
<td>PO02570</td>
<td>HSP-60</td>
</tr>
<tr>
<td>G13</td>
<td>trypsin</td>
<td>56782.7/5.99</td>
<td>P30101</td>
<td>ERG60</td>
</tr>
<tr>
<td>G14</td>
<td>trypsin</td>
<td>47169.2/7.01</td>
<td>PO6733</td>
<td>α-enolase</td>
</tr>
</tbody>
</table>

Larger proteins will give more false hits since they have more peptide fragments; restrict search by MW of protein


ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods

e. Amino acid sequencing using carboxypeptidases

Carboxypeptidase digestion results in a mixture of peptides that are presented to the mass spectrometer

The mass spectrum contains only molecular ions of this “ladder” mixture

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
Example of ladder sequencing using carboxypeptidase digest

In situ digestion: several aliquots of the intact peptide are placed on different locations on the sample probe or slide. Enzyme (in excess) is added to each spot and quenched after a predetermined time by addition of the matrix solution.

Timed-course digests provide a means to maximize amino acid sequence information.


e. Amino acid sequencing using aminopeptidases

The amino acid sequence information is being generated before the sample goes into the mass spectrometer.

Fragment ions are not observed or necessary by this method.
Example of ladder sequencing using aminopeptidase digest

| A. | CTL assay and (insert) MALDI mass spectrum. |
| B. | Aminopeptidase M digest of fraction 45. |
| C. | Carboxypeptidase P digest of fraction 45. |
| D. | Chymotrypsin digest of fraction 45. |

Example: MHC Class I associated peptide from $5 \times 10^9$ Salmonella-infected murine L cells.

A. RP HPLC and MALDI mass spectrum of bioactive fraction.
B/C. Carboxy and amino peptidase digests consistent with Salmonella GroEL protein $\text{GMQFDRGYL}$ or murine heat shock protein $\text{GMKFDRGYI}$.
C. Tryptic digest suggests bacterial GroEl protein.

**TABLE IV. COMMON POST-TRANSLATIONAL MODIFICATIONS**

<table>
<thead>
<tr>
<th>name</th>
<th>site</th>
<th>modification</th>
<th>DM</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-terminal acetylation</td>
<td>terminal NH$_2$</td>
<td>replaced by CH$_3$CONH$-$</td>
<td>+42</td>
</tr>
<tr>
<td>N-terminal formylation</td>
<td>terminal NH$_2$</td>
<td>replaced by HCONH$-$</td>
<td>+28</td>
</tr>
<tr>
<td>N-terminal myristylation</td>
<td>terminal NH$_2$</td>
<td>rep'd by CH$_3$(CH$<em>2$)$</em>{12}$CONH$-$</td>
<td>+210</td>
</tr>
<tr>
<td>N-terminal palmitoylation</td>
<td>terminal NH$_2$</td>
<td>rep'd by CH$_3$(CH$<em>2$)$</em>{14}$CONH$-$</td>
<td>+238</td>
</tr>
<tr>
<td>C-terminal amidation</td>
<td>terminal -COOH</td>
<td>replaced by -CONH$_2$</td>
<td>-1</td>
</tr>
<tr>
<td>disulfide bonds</td>
<td>2 Cys -SH</td>
<td>replaced by -S-S-</td>
<td>-2</td>
</tr>
<tr>
<td>glycosylation (N-linked)</td>
<td>N-X-S/T</td>
<td>see TABLE VII</td>
<td></td>
</tr>
<tr>
<td>glycosylation (O-linked)</td>
<td>S/T</td>
<td>see TABLE VII</td>
<td></td>
</tr>
<tr>
<td>sulfation</td>
<td>-OH of Y</td>
<td>replaced by -OSO$_3$H</td>
<td>+80</td>
</tr>
<tr>
<td>phosphorylation</td>
<td>-OH of Y/S/T</td>
<td>replaced by -OP$_3$H$_2$</td>
<td>+80</td>
</tr>
<tr>
<td>N-methylation</td>
<td>-NH$_2$ of K/R/H/Q</td>
<td>replaced by -NHCH$_3$</td>
<td>+14</td>
</tr>
<tr>
<td>O-methylesterification</td>
<td>-COOH of E/D</td>
<td>replaced by -COOCH$_3$</td>
<td>+14</td>
</tr>
<tr>
<td>carboxylation</td>
<td>-NH$_2$ of E/D</td>
<td>replaced by -NHOCH$_3$</td>
<td>+30</td>
</tr>
<tr>
<td>hydroxylation</td>
<td>-NH$_2$ of P/K/D</td>
<td>replaced by -NOH</td>
<td>+16</td>
</tr>
</tbody>
</table>


ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
h. For example: phosphorylation

Phosphorylated peptides are 80 mass units higher than their unphosphorylated counterparts, as shown in this synthetic phosphopeptide.

This 80 mass unit difference can be observed for:
- the molecular weights of intact protein
- tryptic or other enzymatic fragments
- in carboxy and aminopeptidase ladders

For phosphopeptides, amino and carboxypeptidase ladder sequencing is carried out in the same manner, but the sequence includes a phosphoserine which has a residue mass of:

87 + 80 = 167
2. Molecular Weight Strategies for Carbohydrates

a. Carbohydrate heterogeneity

Structural analysis of carbohydrates is very different from peptides:

• there are a limited number of residue masses for a number of isomers
• there is considerable branching (bianeunary, trianenary, fucosylated sugars, etc.)
• glycopeptides are generally heterogeneous.
• glycosidases are specific for linkage


b. Exoglycosidases and endoglycosidases

**endoglycosidases**

| endo-β-N-acetylglucosaminidase F (Endo F) | GlcNAc(β1→4)GlcNAc |
| endo-β-N-acetylglucosaminidase H (Endo H) | GlcNAc(β1→4)GlcNAc |
| peptide-N-glycosidase F (PNGase F) | GlcNAc-Asn (converts Asp to Asn) |

**exoglycosidases**

| β-galactosidase (jack bean) | Gal(β1→6)[4>>3]GlcNAc |
| β-galactosidase (S. pneumoniae) | Gal(β1→4)GlcNAc,GalNAc |
| β-galactosidase (bovine testes) | Gal(β1→3)[4>>3]GlcNAc,GalNAc |
| β-N-acetylglucosaminidase (S. pneumoniae) | GlcNAc(β1→2)Man, GlcNAc(β1→3,6)Gal |
| a-mannosidase (jack bean) | Man(α1→3,6)Man(β1→4) |

Figure 1. Example of a hybrid carbohydrate structure. Monosaccharide names are given in bold on the left with arrows pointing to proximal monosaccharide names in parentheses. Glycosidase cleavage sites are indicated with arrows.
c. Sequence and branching

Figure 10.16 (a) is the MALDI mass spectrum of the intact glycopeptide, revealing five peaks at \( m/z \) 4820 (I), 5023 (II), 5185 (III), 5224 (IV) and 5388 (V). The glycopeptides were resistant to cleavage of the oligosaccharide from the peptide using PNGase F, suggesting the presence of a fucosyl residue on the GlcNAc proximal to asparagine. Proton-catalyzed defucosylation using 70% trifluoroacetic acid, resulted in shifting all five peaks by 146 Da, indicating that all species contained a fucosyl residue. Subsequent reaction with PNGase F resulted in removal of the oligosaccharides, giving the single peak at \( m/z \) 3651 shown in (b). This was consistent with the molecular weight calculated from the amino acid sequence (3649 Da), conversion of asparagine to aspartic acid, and protonation to form the MH+ ion.

Digestion of the intact glycopeptide mixture with \( \beta \)-galactosidase from jack bean (c) and bovine testes reduced the glycopeptides to I, II and IV (while that from \( S. \) pneumoniae did not), suggesting a Gal(\( \beta \)1→3)HexNAc in III and V. Subsequent digestion of this mixture with \( \beta \)-N-acetylglucosaminidase removed a single GlcNAc residue from structure II and two GlcNAc residues from structure IV, reducing all of the glycopeptides to structure I (d). Digestion of this species with \( \alpha \)-mannosidase at low enzyme concentration removed a single mannose residue, corresponding to cleavage of Man(\( \alpha \)1→6)Man(\( \beta \)1→4), resulting in the peak at \( m/z \) 4659 in (e). Subsequent digestion with the same enzyme at high concentration resulted in loss of a second mannose, corresponding to cleavage of Man(\( \alpha \)1→6)Man(\( \beta \)1→4). The peak at \( m/z \) 4497 in (f) corresponds to a core glycan (consisting of 2 GlcNAc, 1 Man, 1 Fuc and 1 Xyl residues).

Oligosaccharide structures for tryptic peptides from KBPase

d. Chemical degradation

3. Fragmentation of Peptides and Glycoconjugates

Fragmentation mass spectra will be covered in the section on Tandem MS & MS^n.
Strategies for Solving Protein Structures

1. Identification of a binding site
2. Location of a phosphorylation site
3. Non-covalent interactions
4. Peptide-peptide interactions analyzed by MALDI
5. Problems in locating phosphorylation sites
1. Identification of a binding site

Tetanus toxin enters vertebrate motorneurons by binding neuronal surface gangliosides “1b”. The domain of tetanus toxin involved in the binding resides in the carboxyl terminal half of the toxin’s heavy chain. A photoaffinity reagent based on the ganglioside structure was developed and used to label the tetanus domain (Fig. 2).

1. Identification of a binding site (cont)

Strategy:

1. A synthetic peptide representing the carboxyl terminal 34 residues of the tetanus toxin (fragment C) was made. Fragment C was labeled with the photoaffinity label.
2. Spectra of the labeled and non-labeled peptides were acquired (Fig.6).
**Table II**

Peptides and photoaffinity labeled peptides detected by MALDI mass spectrometry

The indicated synthetic peptides or photoaffinity labeled peptide masses were detected by MALDI mass spectrometry (see Figs. 6–8). Enzymatic cleavage was performed on the spectrometer probe as detailed under "Materials and Methods." Enzymes used were: Tryp, trypsin; GluC, endoproteinase GluC; Chym, chymotrypsin; and CP, carboxypeptidase P. Peptide residues 1–34 correspond to tetanus toxin amino acids 1282–1315.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Residues</th>
<th>Enzyme</th>
<th>Azido</th>
<th>Calculated mass</th>
<th>Observed mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>DILLISNWFNKLKKILGCWWFVFTDEGDN</td>
<td>1–34</td>
<td>–</td>
<td>–</td>
<td>4091.5</td>
<td>4091.6</td>
</tr>
<tr>
<td>DILLISNWFNKLKKILGCWWFVFTDEGDN (Na)</td>
<td>1–34</td>
<td>–</td>
<td>+</td>
<td>4323.8</td>
<td>4329.2</td>
</tr>
<tr>
<td>DILLISNWFNKLKK (Na)</td>
<td>1–16</td>
<td>Tryp</td>
<td>+</td>
<td>2188.5</td>
<td>2189.2</td>
</tr>
<tr>
<td>ILISNWFNKLKK (Na)</td>
<td>2–14</td>
<td>GluC</td>
<td>+</td>
<td>1945.2</td>
<td>1943.8</td>
</tr>
<tr>
<td>NLLDKILGCWW (Na)</td>
<td>11–22</td>
<td>Chym</td>
<td>+</td>
<td>1816.1</td>
<td>1815.5</td>
</tr>
<tr>
<td>NLLDKILGCWW (Na)</td>
<td>11–22</td>
<td>Chym</td>
<td>+</td>
<td>1652.9</td>
<td>1652.3</td>
</tr>
<tr>
<td>DILLISNWFNKLKK (Na)</td>
<td>1–15</td>
<td>Tryp/CP</td>
<td>+</td>
<td>2060.3</td>
<td>2059.3</td>
</tr>
<tr>
<td>DILLISNWFNKL (Na)</td>
<td>1–14</td>
<td>Tryp/CP</td>
<td>+</td>
<td>1945.2</td>
<td>1944.0</td>
</tr>
<tr>
<td>DILLISNWFNKL (Na)</td>
<td>1–12</td>
<td>Tryp/CP</td>
<td>+</td>
<td>1817.1</td>
<td>1812.7</td>
</tr>
<tr>
<td>DILLISNWFN (Na)</td>
<td>1–11</td>
<td>Tryp/CP</td>
<td>+</td>
<td>1703.9</td>
<td>1702.7</td>
</tr>
<tr>
<td>DILLISNWFN (Na)</td>
<td>1–11</td>
<td>Tryp/CP</td>
<td>–</td>
<td>1566.8</td>
<td>Not found</td>
</tr>
<tr>
<td>DILLISNWFN (Na)</td>
<td>1–11</td>
<td>Tryp/CP</td>
<td>–</td>
<td>1378.5</td>
<td>1377.5</td>
</tr>
</tbody>
</table>

**Fig. 7.** MALDI-MS of proteolyzed photoaffinity labeled polypeptide.

The synthetic 34-amino acid carboxyl-terminal polypeptide of tetanus toxin derivatized with azido-GluC (spectrum 1), trypsin (spectrum 2), or chymotrypsin (spectrum 3). Masses corresponding to expected polypeptides derivatized with a single 2-(p-aminosalicylamido)ethanethiol group were detected in each case (see Table II for interpretation).
2. Locating a phosphorylation site

Strategy:

1. CMV B-capsids were labeled with $^{32}$P and the assembly protein separated by RP HPLC using a C$_4$ column.
2. The $^{32}$P-labeled assembly protein is digested with trypsin.
3. The peptides were separated by RP HPLC using a C$_{18}$ column, fractions collected and counted. The fractions containing the highest concentration of $^{32}$P counts are lyophilized and resuspended in 20 µl diluent, and a spectrum is acquired.
4. When looking for phosphorylation sites acquire spectra in both positive and negative ion modes, because phosphorylation consensus sites are surrounded by basic or acidic residues. Acidic sites are usually not detected by positive ion mode.
5. **Phosphorylation at serines 156 and 157 is characteristic of caseine kinase II consensus sites.**

Tip:

If sequence is known look for consensus sites (acidic or basic residues surrounding the three amino acid residues Ser, Thr and Tyr that could be phosphorylated) before designing your strategy.

In very rare cases the amino group of lysine is phosphorylated.
## Table 1. Acidic regions, including NLS1 and PGE motifs, in homologs of CMV pAP

<table>
<thead>
<tr>
<th>Group</th>
<th>Sequence</th>
<th>Acidity/sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSV-1</td>
<td>KRRR-----YEASFSEQYCDQDEPDAA---YPTPPGE</td>
<td>7/22</td>
</tr>
<tr>
<td>VZV</td>
<td>RRPDRER--DFKSMQREDLSDSFYSGDQMD--FPSNYYPGE</td>
<td>7/30</td>
</tr>
<tr>
<td>ILTV</td>
<td>RRARESR--TISENE</td>
<td>29</td>
</tr>
<tr>
<td>HSV-2</td>
<td>KRRR-----HEVQPEYDORRORRDR--FPTYPGE</td>
<td>8/22</td>
</tr>
<tr>
<td>EHV-1</td>
<td>KRRR-----HWDNADETRED--GIYYPGE</td>
<td>5/16</td>
</tr>
<tr>
<td>PRV</td>
<td>YDDVAGD-----NAYYPGE</td>
<td>3/11</td>
</tr>
<tr>
<td>Avg %</td>
<td></td>
<td>29 (34)</td>
</tr>
<tr>
<td>Beta</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HCMV</td>
<td>KRRK------ETAAASSSSSDEE---LSYPGE</td>
<td>4/16</td>
</tr>
<tr>
<td>SCMV</td>
<td>KRRRRE------DASSDEER------MSFPGE</td>
<td>6/12</td>
</tr>
<tr>
<td>HHV-6</td>
<td>KTLKRR-----HPOQDEDE------LSFPQD</td>
<td>4/12</td>
</tr>
<tr>
<td>HHV-7</td>
<td>RGSQKR-----CAPTQSDDE------MSPFQD</td>
<td>4/12</td>
</tr>
<tr>
<td>MCMV?</td>
<td>[KRRRERRGAAPDDEGGLSL]×(SDDDODDDEE + 18)---PGE</td>
<td>4/13</td>
</tr>
<tr>
<td>Avg %</td>
<td></td>
<td>34 (45)</td>
</tr>
<tr>
<td>Gamma</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EBV</td>
<td>RSNKKEKR-----DPEDEEE------GGLFPE</td>
<td>6/11</td>
</tr>
<tr>
<td>EHV-2</td>
<td>RPKGKER-----DCDEEPS------GMLFPE</td>
<td>5/11</td>
</tr>
<tr>
<td>HVS</td>
<td>RPNKKEKR-----EDFDEE------CVFPE</td>
<td>4/8</td>
</tr>
<tr>
<td>HHV-8</td>
<td>RTGKKRR-----GAEDE------GHLFPE</td>
<td>4/10</td>
</tr>
<tr>
<td>AHV-1</td>
<td>RPKKAR-----EDFDEE------VSPFPE</td>
<td>5/9</td>
</tr>
<tr>
<td>MVH68</td>
<td>RAGKER-----ECRRED------QVVFPE</td>
<td>5/10</td>
</tr>
<tr>
<td>Avg %</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>Unclassified</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CCV</td>
<td>RYKRRA-----PEPKTAVE------AYVRAPYG</td>
<td>2/13</td>
</tr>
<tr>
<td>Avg %</td>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>

* NLS1 and PGE/D were chosen as reference motifs because of their proximity to the SCMV CKII site and because each has a counterpart in all pAP homologs. Sequence sources have been identified before (29). Abbreviation not given in succeeding footnotes: VZV, varicella-zoster virus; ILTV, infectious laryngotracheitis virus; PRV, pseudo rabies virus; HCMV, human CMV; EBV, Epstein-Barr virus; HVS, herpesvirus saimiri; AHV-1, wildebeest herpesvirus; MVH68, murine herpesvirus 68.

* Region of pAP and homologs containing CKII consensus sequence (double underlining) and a comparatively high content of acidic residues (29 to 32% for alpha- and betaherpesviruses; 15% for channel catfish virus [CCV]).

* Highly conserved sequence (seq) near the middle of pAP and its homologs; only published variants are in human herpesvirus (HHV-6), HHV-7, and CCV.

* Numbers of residues between the end of NLS1 and proline in PGE sequence divided by number of acidic residues (D + E) in the same sequence. group is shown. Values in parentheses are average in percentages calculated by including potential CKII-site phosphoserines and phosphothreonines (only in equine herpesvirus 1 [EHV-1]).

* CKII consensus sequence of murine cytomegalovirus (MCMV) pAP precedes NLS1. The double-headed arrow indicates that the sequences are shown in reverse order for purposes of presentation. +18, 18 amino acids are not shown.
3. Non-covalent interactions by MS

Non-covalent interaction between substrates and enzymes, proteins and metal, receptors and ligands, proteins, and peptides can be studied by both ESI and MALDI.

Hepatitis C viruses-metal complex: ESI spectrum of HCV-Zn (A) At a pH 3.0. only the peptide is seen. (B) At a pH of 7.0 both the peptide and the complex are seen

For an example of a protein-Zn complex seen by MALDI check the following paper:


Figure 7: Shows the ESI spectrum of a drug-ras-GDP non-covalent complex.

In general the pH of the solution containing the non-covalently bound compounds is crucial. Whether ESI or MALDI are used to detect non-covalent complexes. Optimum results are obtained in the pH range 6-7.5.

In ESI in addition to pH, adjusting the orifice voltage could be a factor. The stronger the binding energy of the non-covalent complex, the less is the effect on the stability of the complex produced by the increasing the orifice potential.
Arginine side chain  \[ \text{NH}_2 \quad \text{O} \quad \text{Glu or Asp side chain} \]

\[
\begin{array}{c}
\text{C} \quad \text{NH}_2 \quad \text{C} \\
+ \quad - \\
\end{array}
\]

\[
\begin{array}{c}
\text{NH}_2 \quad \text{O} \\
\end{array}
\]
The side chains of Glutamic and Aspartic Acid differ only in having one and two methylene group respectively. The slight difference in length causes them to have different interaction with the peptide backbone. Hence Glu and Asp have different effects on conformation and chemical reactivity of the backbone.

The carboxyl group of Glu and Asp ionize with pKa values of 4.3 and 3.9 respectively. They are ionized and very polar under physiological conditions.
The Arginine side chain consists of three non-polar methylene groups and a strongly basic d-guanido group:

1. It has a pKa of 12.
2. It is ionized over the entire pH range in which proteins exist naturally.
3. It is planar as a result of resonance and the positive charge is effectively distributed over the entire group.

3. Non-covalent interactions by MS (cont)
4. Peptide-peptide interactions by MALDI

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
The interaction was seen when 6-aza-2-thiothymine [ATT] was used as a matrix (pH 5.4), but was disrupted with a more acidic matrix, a-cyano-4-hydroxycinnamic acid [CHCA] (pH 2.0).

Dynorphin, an opioid peptide, and five of its fragments that contain two adjacent basic residues (Arg$_6$ - Arg$_7$), all interact non-covalently with peptides that contain two to five adjacent acidic residues (Asp or Glu).

Non-related peptides containing two Arg or three (Arg-Lys-Arg) adjacent basic amino acid residues were studied and exhibited the same behavior. However, peptides containing adjacent Lys or His did not form non-covalent complexes with acidic peptides.

The non-covalent bonding was sufficiently stable that digestion with trypsin only cleaved Arg and Lys residues that were not involved in hydrogen bonding with the acidic residues.
4. Peptide-peptide interactions by MALDI (cont)

Positive ion mode spectrum of the 5 minute incubation time of a tryptic digest of an equimolar mixture of mini-gastrin (1647.9 amu) and basic peptide 1 [RKRARKE] (944.0 amu). The MH\(^+\) at 686.0 amu is the tryptic fragment [RKRAR]. The MH\(^+\) at 2590.6 amu is the non-covalent complex of mini-gastrin and basic peptide 1 and the MH\(^+\) at 2333.6 amu is the non-covalent complex of mini-gastrin and the tryptic fragment [RKRAR].
Positive ion mode spectrum of the 10 minute incubation time of a tryptic digest of an equimolar mixture of mini-gastrin (1647.7 amu) and basic peptide 1 [RKRARKE]. The MH$^+$ at 2106.0 amu is the non-covalent complex of mini-gastrin and basic peptide 1 tryptic fragment [RKR].
Peptide-peptide interaction could be explained through the physical and chemical properties of the side chains of the residues involved.

**Arg and Trp are the only residues whose side chain can serve as hydrogen donors; while Asp and Glu can both receive and donate hydrogen.**

The carboxyl groups of Glu and Asp differ marginally in their physical and chemical properties. The carboxyl groups of Asp and Glu side chains ionize with intrinsic pKa values of 3.9 and 4.3, respectively; these residues are ionized and very polar under physiological conditions. The Arg side chain consists of three non-polar methylene groups and a strongly basic d-guanidinium group, which is dissociated over the entire pH range [pKa is 12.0].

The chemistry of basic residues such as arginine, and acidic residues such as aspartic and glutamic acid, seem to support the likelihood that **electrostatic attraction of salt bridges and hydrogen bonding between residues is at the base of the observed peptide-peptide interaction.** It is obvious that a pH range close to physiological pH is needed [A.S. Woods and M.A. Huestis, A Study of Peptide Peptide Interaction by MALDI, JASMS 12, 88-96 (2001)].
Negative Ion ESI MS of MG and D_{1-7}

INSET: Dissociation Products of their Complex

\[
\begin{align*}
[D_{1-7} - H]^- \\
[MG - 2H]^2- \\
[D_{1-7} + MG - 3H]^3- \\
[D_{1-7} - H]^+ \\
\end{align*}
\]

MSMS

\[
\begin{align*}
[D_{1-7} - H]^+ \\
[MG - 2H]^2- \\
[D_{1-7} + MG - 3H]^3- \\
\end{align*}
\]

(* Mass Selected for MSMS)

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
Negative Ion ESI MS of MG and RKRARKE
INSET: Dissociation Products of their Complex

![Graph showing ion signals and mass spectra](image-url)

- **[MG - 2H]^2-**
- **[RKRARKE - H]^−**
- **[RKRARKE + MG - 3H]^3-**

(* Mass Selected for MSMS)
4. Peptide-peptide interactions by MALDI (cont)

Negative Ion Electro spray Mass Spectra of dynorphin 1-7 (D₁₋₇) Mini Gastrin I (MG) and D₁₋₇-MG complex at pH 7 (A) and D₁₋₇ and MG only at pH 2 (B).

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
5. Example of problems in locating phosphorylation sites

Detection of phosphorylated peptides by various mass spectrometric techniques has often been an arduous and pains taking task. It was observed that if the residue (S, T or Y) to be phosphorylated were:

1. Surrounded by basic residues (R/K) [such as in the consensus sites of protein kinase C], detection of the phosphorylated peptides could easily be done by MALDI or ESI in both positive or negative ion mode.

2. However if the residues to be phosphorylated were surrounded by negatively charged residues (E/D) [such as in the consensus sites of casein kinase I and II], the peptides especially if in a mixture were often not detected.
5. Phosphorylation example (cont)

We made a mixture of the following peptides:

RRREEEpTEEEAA Peptide 3 (P3)
RRREEEpSEEEAA Peptide 4 (P4)
RRREEETEEE Peptide 5 (P5)

and acquired spectra using a MALDI-TOF, a Q-star and an MALDI-IM-TOF. We had no problem detecting the peptides. However if we added trypsin to our mixture and cleaved the RRR leader we started seeing suppression of some or all the phosphorylated peptides.

To circumvent suppression of the negatively charged peptides we buffered the solution with ammonium buffers and used MALDI matrices that had a pH range of 4-6 rather than 1-2, and were able to squelch suppression.
5. Phosphorylation example (cont)

Enzymatic Digests:

i. Buffered digest: 5 ml peptide or peptide mixture (P3, P4, P5)  
     + 10 ml 25 mM Ammonium bicarbonate buffer pH 7.8  
     + 5 ml Trypsin.

ii. Non-Ammonium Buffered digest:

     5 ml peptide or peptide mixture (P3, P4, P5)  
     + 10 ml 25 mM Tris pH 7.5  
     + 5 ml Trypsin.

Two 0.3 ml aliquots of each digest were spotted on a sample plate at 1, 5, 10, 20 and 30 minutes. The enzymatic reaction was stopped by addition of 0.3 ml matrix to one spot and addition of 0.3 ml 100 mM Ammonium sulfate and 0.3 ml matrix to the other spot.
5. Phosphorylation example (cont)

**Q-star settings:** The Q-star is fitted with a Protana (Odense, Denmark nanoelectrospray ion source. All spectra were acquired in positive ion mode. The spray voltage was varied to optimize signal, but usually 1050 V was used. The nozzle skimmer potential was minimized (5 V) to avoid fragmentation. The mixtures are prepared with 10 pmol/ul of each peptide dissolved in 50% water / 50% isopropanol with 5 mM ammonium acetate.
### 5. Phosphorylation example (cont)

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
<th>MW</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1</td>
<td>DRVpYIHPF</td>
<td>1126.2</td>
</tr>
<tr>
<td>P2</td>
<td>DRVYIHPF</td>
<td>1046.2</td>
</tr>
<tr>
<td>P3</td>
<td>RRREEEepTEEEA</td>
<td>1584.5</td>
</tr>
<tr>
<td>P4</td>
<td>RRREEEepSEEEA</td>
<td>1570.5</td>
</tr>
<tr>
<td>P5</td>
<td>RRREEETEEE</td>
<td>1362.4</td>
</tr>
<tr>
<td>P6</td>
<td>KRpTIRR</td>
<td>909.0</td>
</tr>
<tr>
<td>P7</td>
<td>VRKRTLRL</td>
<td>1197.5</td>
</tr>
<tr>
<td>P8</td>
<td>RRApSPVA</td>
<td>835.9</td>
</tr>
</tbody>
</table>
Positive ion mode ESI-TOF spectrum of 8 peptides.

Low relative intensity for phosphorylated peptides surrounded by acidic residues.

Labeled Ion Signals Correspond to:
1. DRVpYIHPF
2. DRVYIHPF
3. RRREEEpTEEEAA
4. RRREEEpSEEEAA
5. RRREEETEEE
6. KRpTIRR
7. VRKRTLRRL
8. RRApSPVA
Positive ion mode ESI-TOF spectrum of phosphorylated angiotensin 2 (P1) and angiotensin 2 (P2), shows that the relative intensity of P1 is a third that of P2.
Positive mode ESI-TOF spectrum of peptides all containing 6 Glu and 3 Arg (P3, P4, P5) the relative intensity of the phosphorylated ones (P3, P4) is a third to half the unphosphorylated peptides.
Positive ion mode spectrum, tryptic digest of P3,P4,P5

Conditions:
Buffered with ammonium bicarbonate, Ammonium sulfate added to spot
MALDI Negative ion mode spectrum, tryptic digest of P3,P4,P5

Conditions: Buffered with ammonium bicarbonate, Ammonium sulfate added to spot

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
Conclusions

1. Phosphorylation is an important post-translational modification. However, it is often difficult to detect because sites are usually partially phosphorylated and peptides containing acidic consensus sites are difficult to detect especially in positive ion mode. In the positive ion mode spectra of the peptide mixture (P3, P4, P5) tryptic digest [30 min]:

   Only fragments containing one or more Arg are seen

2. Addition of Ammonium sulfate, and the use of an ammonium buffer, improved the resolution of the spectra and the relative intensity of the fragments. In the negative ion mode spectrum. All fragments were seen

3. However fragments in the spectra acquired from the spot to which ammonium sulfate was added had a higher relative intensity.
MALDI Positive ion mode spectrum of a tryptic digest of P3

Conditions: Digest buffered with Tris, Ammonium sulfate added to sample spot

Trypsin fragment EEEpTEEEAAA

m/z (Dalton) 650 870 1090 1310 1530 1750

655.9 809.1 1115.6 1627.7

1137.1

1583.8

1605.8

1627.7

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
MALDI Negative ion mode spectrum of a tryptic digest of P3

Conditions: Digest buffered with Tris, Ammonium sulfate added to sample spot

EEEpTEEEAA

Trypsin fragment

809.2
1115.5
1137.7
1584.4

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
Spectra of the mixture digest where Tris was used as a buffer did not show any fragments in the positive or negative ion mode at 30 min, however if ammonium sulfate was added when spotting the sample most fragments are seen in negative ion mode, although the relative intensity of these fragments is low.

The spectra of the tryptic digest of P3 or P4 alone in Tris buffer showed no fragments, however if ammonium sulfate was added to the sample spot, fragment EEEpTEEEMAA or EEEpSEEEMAA were detected in both positive and negative ion mode, although the relative intensity in negative mode is more prominent.

In conclusion the addition of ammonium buffers seems to be beneficial in diminishing peptide suppression caused by the presence of negatively charged residues or phosphorylation.
MALDI-IM-TOF 2D spectrum; P1 and P2 Positive ion Mode

Conditions: a mixture of P1 & P2 was spotted on the probe followed by 100 mM Ammonium sulfate and saturated HCHA in ethanol:water 1:1. The dye is an internal standard.
Proteomics

1. Definition of proteomics
2. Gel electrophoresis
3. Historical milestones
4. Analytical Strategies and protocols
5. The ICAT method for quantitation
6. 2D gel approach
7. Large scale approach
8. Protein identification and analysis tools
9. Primary structure analysis
**Definition of proteomics**

**What is proteomics?**

The word *proteomics* combines the word *protein* with *omics* such as used in the word *genomics*.

One definition of *proteomics* is the quantification of all proteins expressed at any individual time in a cell.

*Functional proteomics* may be defined as the determination of the function of all the proteins encoded by an organism’s entire genome.

*Structural proteomics* is the determination of the structures of proteins, which may have multiple functions that can only be identified in the context of their interactions with other proteins, DNAs, small molecules, cofactors, etc.
**Gel electrophoresis:**

Polyacrylamide gel electrophoresis [PAGE]: is the most widely used protein separation and purification technique.

Two-dimension-gel electrophoresis [2D-gel]: is applied to separation of crude protein mixtures. Blotting protein digests from 2D-gels onto a PVDF membrane is followed by Edman sequencing or amino-acid analysis of the generated peptides. However these techniques need a minimum of 10 pmoles and usually use as much as 100 pmoles.
**Historical milestones**

In the past ten years, two scientific developments in very different fields have resulted in a new approach to solving protein structure.

1. **Development of easy-to-use yet high-performance mass spectrometers:** Techniques such as MALDI and ESI that have high resolution, sensitivity and accuracy have allowed the measurement of proteins and the peptides generated from the digests of these proteins, followed by fragmentation of these peptides from as little as a few femtmoles of protein.

2. **The human genome project:** as well as other bacterial and animal genome sequencing has provided huge databases that can be searched by feeding into these programs data such as molecular weight [MW], isoelectric point [pl], hydrophobicity index, amino acid sequence or composition.
Before the development of the new strategy of identifying and sequencing proteins with the help of mass spectrometric analysis and data base searches, the conventional approach consisted of electroblotting proteins from gels followed by antibody labeling of the protein, amino-acid analysis and Edman degradation.

- In 1987 Ruedi Aebersol showed that the internal amino acid sequence of a protein can be obtained by *in situ* enzymatic digestion on the membrane, RP HPLC separation and Edman degradation.
- In 1994 Vestling and Fenselau, and Strupat *et al.* showed that proteins electroblotted on PVDF could be directly analyzed by MALDI MS.
- In 1997 Rachel Ogorzalek-Loo demonstrated that analysis of intact proteins could be analyzed directly from thin gel slabs by MALDI MS.
In 1997 Woods et al. performed CNBr digests and enzymatic digests in situ on polypropylene membranes followed by MALDI MS analysis.

Through the 1990s Aebersol and Yates eluted proteins from gels or membranes and analyzed them by ESI or MALDI.

However in 1996 A. Shevshenko and M. Mann refined the technique and published what is considered the “Standard Strategy” for sequencing unknown proteins.
Analytical Strategy for Mass Spectrometric Characterization of Gel isolated Proteins

Excise gel pieces with protein spots, *in-gel* reduce and alkylate (25 spots in parallel) (6 hours)

↓

Digest *in-gel* with trypsin (overnight)

↓

High mass accuracy MALDI MS analysis of 0.3 µl of supernatant (5 minutes/sample)

↓

If Protein is identified Analysis is done

Search database with peptide map

↓

If Protein is not identified Proceed

If Protein is identified Analysis is done
Analytical Strategy for Mass Spectrometric Characterization of Gel isolated Proteins (cont)

Extract gel pieces, dry down pooled extracts. Perform a single desalting-concentration step using Zip-Tips™ or Sep-Pack™ to extract peptide mixture. (3 hours)

Nano ES tandem MS sequencing (one hour per sample)

If Protein is identified → Analysis is done

If Protein is not identified → Proceed

Search database with peptide map
Analytical strategies and protocols

The analytical strategy recommended by A. Shevshenko and M. Mann has become the standard because:

1. It side-steps electroblotting which can waste a lot of protein, and is often protein dependent.

2. Handling proteins directly in the gel makes it easier to detect proteins present in sub picomole amounts because the protein is fixed in the gel during sample preparation.

3. Sodium dodecyl sulfate interferes with mass spectrometric analysis, but can be effectively removed by repeated rinsing with buffer.
If the protein cannot be identified by the Shevshenko-Mann protocol it could be due to:

1. Large number of posttranslational modifications [e.g. heavily glycosylated proteins, proteins with lipid or glycolipid modifications].

2. Poor calibration leading to inaccurate mass measurements.

3. Suppression.

4. This particular protein has not been sequenced yet.

5. The data base you are using has not yet incorporated the results from the latest proteins to be sequenced.

6. Not enough protein available in the preparation to be detected by Coomassie stain (100 ng needed for visualization).

7. This problem can be remedied by excising gel bands from silver stained gels.

8. Reverse staining methods such as Zinc/imidazole can also be used.
Analytical strategies and protocols (cont)

Mann’s tips for achieving success:

1. Rinse gel slab in water for at least one hour after the gel has been fixed.

2. If the protein of interest appears to be contaminated with other proteins migrating nearby then excise the control blank below the band of interest to obtain a representative pattern of the contaminants.

3. Trypsin is the enzyme of choice because it generates peptide fragments in the 800-2500 Da range which can be easily fragmented by PSD or tandem MS.

4. Trypsin autolytic peptides can be used as internal calibrants (see tables).

5. For proteins identified by MALDI peptide mass map and database search
   
a. The peptides mass map measurements should be within 50ppm.
   
b. The sequence coverage should be better than 15 %. For 100000 or larger protein, a coverage of 10 % could be enough.
Analytical strategies and protocols (cont)

c. Peptide mass maps should always be inspected for the presence of cysteine containing peptides which are S-alkylated with acrylamide owing to the presence of unreacted reagents in the gel.

d. The N-terminal tryptic peptide should be detected. If it is not detected at the expected mass, try to remove the initiating Methionine residue or allow for N-acetylation or other posttranslational modifications.

6. Small hydrophilic or hydrophobic proteins are difficult to find, because not enough peptides are detected.

7. When low femtomolar amounts of proteins are present, PSD fragmentation of peptides can be inaccurate and analysis by nanospray using instruments such as the Q-TOF should give a better result.
### Porcine Trypsin Fragments

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Mono.</th>
<th>Avg.</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>53</td>
<td>261.14</td>
<td>261.28</td>
<td>SR</td>
</tr>
<tr>
<td>54</td>
<td>57</td>
<td>514.32</td>
<td>514.63</td>
<td>IQVR</td>
</tr>
<tr>
<td>108</td>
<td>115</td>
<td>841.50</td>
<td>842.01</td>
<td>VATVSLPR</td>
</tr>
<tr>
<td>209</td>
<td>216</td>
<td>905.50</td>
<td>906.05</td>
<td>NKPGVYTK</td>
</tr>
<tr>
<td>148</td>
<td>157</td>
<td>1005.48</td>
<td>1006.15</td>
<td>APVLSDDSSCK</td>
</tr>
<tr>
<td>98</td>
<td>107</td>
<td>1044.56</td>
<td>1045.16</td>
<td>LSSPATLNSR</td>
</tr>
<tr>
<td>134</td>
<td>147</td>
<td>1468.72</td>
<td>1469.68</td>
<td>SSGSSYPSLLQCLK</td>
</tr>
<tr>
<td>217</td>
<td>231</td>
<td>1735.84</td>
<td>1736.97</td>
<td>VCNYVNWIQQTIAAN</td>
</tr>
<tr>
<td>116</td>
<td>133</td>
<td>1767.79</td>
<td>1768.99</td>
<td>SCAAAGTECLISGWGNTK</td>
</tr>
<tr>
<td>158</td>
<td>178</td>
<td>2157.02</td>
<td>2158.48</td>
<td>SSYPGQITGNMICVGLEGGK</td>
</tr>
<tr>
<td>58</td>
<td>77</td>
<td>2210.10</td>
<td>2211.42</td>
<td>LGEHNIDVLEGNEQFINAAK</td>
</tr>
<tr>
<td>78</td>
<td>97</td>
<td>2282.17</td>
<td>2283.63</td>
<td>IITHPNFNGNTLDNDIMLIK</td>
</tr>
<tr>
<td>179</td>
<td>208</td>
<td>3012.32</td>
<td>3014.33</td>
<td>DSCQGDSDGG...SWGYGCAQK</td>
</tr>
<tr>
<td>9</td>
<td>51</td>
<td>4474.09</td>
<td>4477.04</td>
<td>IVGGYTCAA...VVSAAHCYK</td>
</tr>
</tbody>
</table>
## Bovine Trypsin Fragments

<table>
<thead>
<tr>
<th>From</th>
<th>To</th>
<th>Mono.</th>
<th>Avg.</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>110</td>
<td>111</td>
<td>259.19</td>
<td>259.35</td>
<td>LK</td>
</tr>
<tr>
<td>157</td>
<td>159</td>
<td>362.20</td>
<td>362.49</td>
<td>CLK</td>
</tr>
<tr>
<td>238</td>
<td>243</td>
<td>632.31</td>
<td>632.67</td>
<td>QTIASN</td>
</tr>
<tr>
<td>64</td>
<td>69</td>
<td>658.38</td>
<td>658.76</td>
<td>SGIQVR</td>
</tr>
<tr>
<td>112</td>
<td>119</td>
<td>804.41</td>
<td>804.86</td>
<td>SAASLNSR</td>
</tr>
<tr>
<td>221</td>
<td>228</td>
<td>905.50</td>
<td>906.05</td>
<td>NKPGVYTK</td>
</tr>
<tr>
<td>160</td>
<td>169</td>
<td>1019.50</td>
<td>1020.17</td>
<td>APIILSDSSCK</td>
</tr>
<tr>
<td>229</td>
<td>237</td>
<td>1110.55</td>
<td>1111.33</td>
<td>VCNYVSWIK</td>
</tr>
<tr>
<td>146</td>
<td>156</td>
<td>1152.57</td>
<td>1153.25</td>
<td>SSGTSYPDVLK</td>
</tr>
<tr>
<td>207</td>
<td>220</td>
<td>1432.71</td>
<td>1433.65</td>
<td>LQGIVSWGSGCAQK</td>
</tr>
<tr>
<td>191</td>
<td>206</td>
<td>1494.61</td>
<td>1495.61</td>
<td>DSCQGDSGGPVVCSGK</td>
</tr>
<tr>
<td>70</td>
<td>89</td>
<td>2162.05</td>
<td>2163.33</td>
<td>LGEDNINVVEGNEQFISASK</td>
</tr>
<tr>
<td>170</td>
<td>190</td>
<td>2192.99</td>
<td>2194.47</td>
<td>SAYPGQITSNMFCAGYLEGGK</td>
</tr>
<tr>
<td>90</td>
<td>109</td>
<td>2272.15</td>
<td>2273.60</td>
<td>SIVHPSYNSTLNNDIMLIK</td>
</tr>
<tr>
<td>120</td>
<td>145</td>
<td>2551.24</td>
<td>2552.91</td>
<td>VASISLPTS...LISGWGNTK</td>
</tr>
<tr>
<td>21</td>
<td>63</td>
<td>4550.12</td>
<td>4553.14</td>
<td>IVGGYTCGA...VVSAAHCYK</td>
</tr>
</tbody>
</table>
Analytical strategies and protocols (cont)

My favorite in gel-digest protocol for Coomassie Blue In-gel Digestion

1. Place gel slice into 1.5 mL Eppendorf tube.

2. Destaining: Add 150 µL of 50:50 (v/v) ACN: 25 mM NH₄CO₃ and let sit for 30 min. Repeat destaining step two more times; each time removing the solvent to waste. The majority of blue dye should be removed (or at least be very faint).

3. Dry down sample completely in Speed-Vac (~ 1 hour).

4. Digest: Add 5 µL of trypsin solution (0.1 µg/µL in 1% acetic acid) and 25 µL of 25 mM ammonium bicarbonate pH 8.0. Incubate at 37°C overnight.

5. Extraction: Add 30 µL of 50:50 (v/v) ACN: H₂O, 5% TFA and let sit for 30 minutes. Remove the eluant and place into new 1.5 µL Eppendorf tube. Repeat step 5.
6. Dry down eluant in Speed-Vac and then reconstitute by adding 5 µL 0.1 % TFA 20 % acetonitrile.

7. Place 0.3 µL of eluant, 0.3 µL saturated ammonium sulfate solution and 0.3 µL CHCA matrix in 50% ethanol on MALDI probe. Prepare another spot with 0.3 µL of eluant and 0.3 µL CHCA matrix only

However this protocol is lacking two steps: “reduction and alkylation”. It generally works quite well for some small proteins. But in some cases reduction is necessary to get a more complete digest. Alkylation can be helpful in some cases.
Analytical strategies and protocols (cont)

My tips:

1. If you have a second gel slice available, digest with a second enzyme and compare coverage.

2. If the protein is very large the following protocol could be used as an alternative:
   
   a. Start with a chemical digest using CNBr (cleaves at Met) or N-Chloro-succinamide (cleaves at Trp) that will cut the protein into 5 to 10 large fragments. Separate the fragments by running another gel or RP-HPLC.
   
   b. Digest and sequence a few of the large fragments.
   
   c. Do a database search.
Analytical strategies and protocols (cont)

2D-gel electrophoresis is a powerful tool.

- It can separate 2-4000 proteins in a single analysis.
- It separates a complex mixture of proteins into single spots on a gel.
- The separation is based on physical properties related to the isoelectric point [pI is the pH at which the peptide has a net charge of zero] and molecular weight of each protein.
- The amount of proteins in a 2D spot usually reflects its in-vivo abundance. However spot quantitation has been difficult.
The ICAT method for quantitation

Isotope Coded Affinity Tags


1. site-specific, covalent labeling of protein with isotopically normal or heavy ICAT reagents.
2. Proteolysis of the combined, labeled protein mixture.
3. Isolation and Purification by multidimensional chromatography or by SDS-PAGE.
4. ESI-MS/MS.
5. Proteins Identification and quantification. The relative abundance of the peptide and therefore the relative abundance of the protein from which the peptide originated is determined by the ratio of signal intensities of the isotopically normal and heavy forms of the peptide.
The ICAT method for quantitation (cont)

1. Labeling
2. Proteolysis
3. Isolation
4. MS
5. ID

Figure 1. Schematic representation of the quantitative proteomics procedure. Proteins in the microsomal fraction of naive or PMA-treated HL-60 cells were labeled with ICAT reagents, combined, and analyzed as described in the text.
**The ICAT method for quantitation (cont)**

---

**The ICAT reagents:**

Consist of 3 functional components (see above fig.).

1. A thiol reactive group that is selective for the sulphydryl group reduced cysteines.

2. An ethylene glycol linker that occurs in deuterated and isotopically normal form and provides the basis for quantification.

3. Biotin which provides an affinity tag for the selective isolation of tagged peptides.
The **ICAT** method for quantitation (cont)

Figure 2: Multidimensional liquid chromatography tandem mass spectrometric analysis of a complex peptide mixture.

(A) Distribution of peptides contained in tryptic-digested HL-60 microsomal fraction on a strong cation-exchange chromatography column. Peptides were detected by absorbance at 214 nm (blue line) and 280 nm (green line). Solvent gradient (red line), and pressure (pink line) are also indicated. Collected fraction numbers are shown on the x-axis. (B) Analysis of the biotinylated, cysteine-containing peptides contained in cation-exchange fraction 18 by LC-ESI-MS/MS. Ion chromatogram displaying the base peak (most intense ion signal in each MS scan) as a function of retention time. Dotted line indicates the percentage acetonitrile solvent gradient used to develop the reverse-phase capillary column. (C) MS spectrum of peptides detected in the 30 s time window indicated in (B). Signals indicated with asterisk (*) were ICAT-labeled peptides that were subjected to automated CID. The data obtained from the peptides numbered as 1 and 2 are shown in figure 3. (D) Base peak ion chromatogram of all of the cation-exchange fractions collected, indicating ICAT peptide exchange distribution. Red line indicates solvent gradient.
The ICAT method for quantitation (cont)

Figure 3: Post-MS processing of data from selected peptides.

Two peptides numbered as 1 (A) and 2 (B) in Figure 2C were subjected to alternating MS and MS/MS scans. For each peptide we show the CID spectrum, the identity of the parent protein determined by sequence database searching using SEQUEST\(^{16}\) (C\(^{*}\) designates cysteine labeled with the heavy form of ICAT reagent, while C# designates cysteine labeled with the light form of ICAT reagent), the data indicating the relative abundance, and the calculated d0:d8 ratio obtained using XPRESS software. (C) Flowchart indicating the steps applied to process the initial CID spectra to identify and quantify microsomal proteins.
Figure 4: Consistency of redundant protein quantification.

A total of 13 peptides from the CD45 transmembrane protein tyrosine phosphatase CD45 were identified and quantified. (A) CD45 protein sequence with the identified, ICAT reagent-labeled peptides indicated in boxes with yellow background. (B) Redundancy and accuracy of quantitation: the amino acid sequence of the identified peptides, the cation-exchange fraction(s) in which each peptide was detected, the number of times each peptide was identified and quantified, and the abundance ratio and its standard deviations are indicated. Results are shown from three independent experiments: one large scale (first experiment) and two smaller scale (second and third experiments).
The ICAT method for quantitation (cont)


This requires the specific activity (i.e. cpm/pmole protein/methionine) for the cell lysate to be calculated. This is done by measuring the protein abundance by amino acid analysis (AAA) for a few known spots, then the radioactivity of these AAA quantitated spots is counted. As the number of methionine residues in each one of these spots is known, this information gives a specific activity for each of the small number of selected spots.

The mean value for this number can be used to quantitate spots of lesser intensity by measuring the counts present in the spot and the identity of the protein
35S-labeled methionine lysate

Run 2-D gel

Stain, and dry gel. Excise spots

Scintillation counting of spots

Cold lysate

Run 2D-gel

Excise identical spots

Perform in gel tryptic digests

Identify proteins in digests by using MS to do a peptide map and sequence some peptides

Calculate number of methionines

The specific activity of the cell lysate is calculated (cpm/pmol protein/methionine)

Calculate abundance of individual spots

Aebersold’s procedure to quantitate and identify proteins from a total cell extract
Disadvantages of the method:

1. The protein sequence must be known.
2. Does not work for proteins that do not contain methionines.
3. Spots should be well resolved, because if two or more proteins are present in the same sample, they will contribute to errors for both the AAA and radioactivity measurement.
4. A time-course for the incorporation of radioactive material should be done to determine that labeling is done at equilibrium. The background should be measured and subtracted from all values prior to calculation.
5. The potential cleavage of the initiator methionine should be taken into account.

Calculation example: cpm

If the mean specific activity of the cell lysate is 500 cpm/pmol protein/methionine, and a 2D protein spot containing 6 methionine and 1145 cpm is counted, then 1145 cpm divided by 6 methionines is 191 cpm/methionine.

191 cpm/methionine X 1 pmole protein / 500 cpm/methionine = 0.382 pmol protein in the spot. If 1.0 X 10^6 cells were used to prepare the lysate then 0.382 pmol protein / 1.0 X 10^6 cells is 3.82 X 10^7 pmol protein/cell or 230000 copies/cell
In-gel digests from: Normal cells    Diseased cells

Excise gel pieces with protein spot (including control piece with no protein).

Chop into small pieces

Wash then dry

Add enzyme + appropriate buffer, incubate at 37° C overnight

Remove digest buffer

Extract peptides with organic buffer

Clean with Zip-tip

Determine masses with ESI or MALDI MS

Compare results from diseased cells to those from normal cells. Often in diseased cells a protein is missing, modified or present in inadequate quantities. Such proteins are called markers.
How realistic are our expectations of protein identification?

According to Aebersold et al. [PNAS, 97, 9390-9395 (2000)] :

“Although this technique is powerful, mature, and sensitive questions remain concerning its ability to characterize all of the elements of a proteome.”

The 2D-gel Approach:

- In this study they visualized over 1500 proteins from 0.5 mg of total soluble yeast.

- Proteins present in low abundance were not detected, although about half of the yeast proteins are lower abundance proteins, i.e. that they are expressed from genes with codon bias values of <0.1.

- Such proteins are found if protein quantities exceeding the capacity of a 2D-gel are fractionated and analyzed.
The Large Scale Approach:

- They used 50 mg (100 times more than average cell lysate) of total soluble yeast, loaded into a single large well, ran the gel.

- Cut gel strip (3 mm X 10 mm). Cut strip into 1 mm$^3$ pieces.

- In-gel digested with trypsin.

- Eluted peptides by HPLC using a polysulfoethyl A column and a 60 min gradient.

- Some of the fractions were analyzed by micro-capillary LC-MS/MS using a 120 min gradient.

The large scale approach (50 mg) allowed the detection of some of the low abundance proteins, none of which had been detected by the 2D-gel approach.
The take home message is: although the 2D-gel approach is an efficient and effective tool that requires a picomole or less for the detection of a large number of proteins

• It does not necessarily give the whole picture as to the protein content of a particular cell.

• Sometimes it is necessary to use lysates containing 10 to 100 picomoles to be able to visualize the low abundance proteins.

• Certain proteins are only expressed when needed and are rapidly metabolized, so they are usually present in low abundance.
Protein Identification and Analysis tools

NOW THAT YOU HAVE ALL THAT DATA WHAT DO YOU DO WITH IT?

Several sites on the web are available to search databases and identify proteins. The SwissProt site has a server called “Expasy”. This server offers numerous tools:

• Protein identification and characterization

• DNA -> Protein

• Similarity searches

• Pattern and profile searches

• Post-translational modification prediction

• Primary structure analysis
Protein Identification and Analysis tools (continued)

• Secondary structure prediction

• Tertiary structure

• Transmembrane regions detection

• Sequence alignment

Each of these headings has sub-headings that can take you to one or more sites where you can get parameters or extra data that will help with visualizing the structure of your protein and even inferring a function to it. The next page shows what information can be gleaned from “the primary structure” heading.
Primary structure analysis

InterPro Scan - Integrated search in PROSITE, Pfam and PRINTS

ScanProsite - Scans a sequence against PROSITE or a pattern against SWISS-PROT and

TrEMBL ProfileScan - Scans a sequence against protein profile databases (including PROSITE)

Frame-ProfileScan - Scans a short DNA sequence against protein profile databases (including PROSITE)

Pfam HMM search; scans a sequence against the Pfam protein families db [At Washington University or at Sanger Centre]

FPAT - Regular expression searches in protein databases

PRATT - Interactively generates conserved patterns from a series of unaligned proteins

PPSEARCH - Scans a sequence against PROSITE (allows a graphical output); at EBI
Primary structure analysis (continued)

PROSITE scan - Scans a sequence against PROSITE (allows mismatches); at PBIL

PATTINPROT - Scans a protein sequence or a protein database for one or several pattern(s); at PBIL

SMART - Simple Modular Architecture Research Tool; at EMBL

TEIRESIAS - Generate patterns from a collection of unaligned protein or DNA sequences; at IBM

Hits - Relationships between protein sequences and motifs
Tandem MS and MS\textsuperscript{n}

1. Prompt Fragmentation

   a. In-source decay (ISD)
   b. Post-source decay (PSD)
   c. The curved-field reflectron
   d. Nested PSD

3. Tandem Mass Spectrometry
   a. Mass filters vs. mass analyzers
   b. Mass selection
   c. Collision-induced dissociation (CID)

4. Triple Quadrupoles

5. Quadrupole Ion Traps
   a. Tandem-in-time
   b. Mass isolation and excitation
6. Atmospheric Pressure MALDI: examples of MS, MS/MS and MS^n

7. Fourier transform mass spectrometers

8. Tandem Time-of-Flight Mass Spectrometers
   a. High energy vs. low energy collisions
   b. Dual reflectron tandem time-of-flight
   c. Delayed extraction focusing as MS1
   d. Photodissociation
   e. Pulsed extraction in MS2

9. Hybrid instruments
   a. Sector/orthogonal TOF
   b. Trap/TOF
   c. QTOF

10. Electrospray ionization CID mass spectra of multiply-charged ions

11. Conclusions: Analysis of Proteins and Glycoproteins
1. Prompt Fragmentation

- Fragmentation that occurs during the ionization process

- Fragment ions appear along with molecular ions in the same spectrum and follow the same square root law, i.e. for $m_1 \rightarrow m_2 + n$:

$$t_1 = \left( \frac{m_1}{2eV} \right)^{1/2} D$$

$$t_2 = \left( \frac{m_2}{2eV} \right)^{1/2} D$$

- Advantage: simplest way to obtain structural information from a pure compound

- Disadvantage: cannot distinguish fragment ions from molecular ions, or which fragment belongs to which molecule

- Disadvantage: cannot be used for mixtures

- Disadvantage: methods for increasing fragmentation, such as raising the laser power, also reduce mass resolution

a. In-source decay (ISD)

- Fragmentation that occurs in the ion source during the time prior to application of a delayed extraction pulse

- Increases the time that ions can fragment in the source; larger mass ions take longer to fragment

- Fragmentation takes place very close to sample surface; therefore has all the same properties as prompt fragmentation

- Advantage: large increase in fragmentation compared with prompt fragmentation

- Advantage: ISD has been used quite successfully to obtain amino acid sequences from purified peptides

- Disadvantage: cannot select ions from a mixture
In-source decay mass spectrum of the oxidized chain of bovine insulin obtained using 350 ns delayed extraction


Like prompt fragmentation ISD fragments follow a square root law
b. Post-source decay (PSD)

- Fragmentation occurs after ions leave the source, i.e. in the drift region
- Information from post-source decay can only be obtained from a reflectron instrument

- Before entering the reflectron, fragment ions have the same velocity and positions of their molecular precursors
- After they exit the reflectron, fragment ions have the same velocity as their precursors, but lead them in space and time
- Advantage: can select molecular ions for fragmentation from a mixture of molecular ions
- Disadvantage: cannot directly control the amount of fragmentation
b. Post-source decay (PSD) cont.

The single-stage reflectron focuses molecular ions at flight times:

\[ t = \left( \frac{m}{2eV} \right)^{1/2} \left[ L_1 + L_2 + 4d \right] \]

Fragment ions formed by post-source decay are dispersed but not all focused on a linear scale:

\[ t_b = \left( \frac{m_a}{2eV} \right)^{1/2} \left[ L_1 + L_2 + 4 \frac{m_b}{m_a} d \right] \]
When single-stage or dual-stage reflectrons are used to record post-source decay mass spectra, it is necessary to “step” the reflectron voltage in order to focus different mass regions.

The mass spectral segments are then “stitched” to form a finished mass spectrum.

CID/PSD mass spectrum of a wasp neurotoxin obtained by introducing argon gas. Mass spectra were acquired on a PE BioSystems Voyager Elite.

Within each segment mass follows a linear dependence on time, but with different slopes. Computer presents a single mass scale.

c. The *curved-field* reflectron

The voltage vs. reflectron depth follows the arc of a circle.

Fragment ions formed by PSD are focused without scanning the reflectron voltage.
The curved-field reflectron can be used to obtain amino acid sequences from peptides in a mixture. After setting the gate for each peptide, the laser power is raised slightly to induce fragmentation. The mass spectrum is obtained without scanning the reflectron voltage.

c. The *curved-field* reflectron (cont.)

High performance *curved-field reflectron* TOF mass spectrometer: high resolution *post-source decay* mass spectrum of angiotensin II obtained on a Kratos AXIMA-CFR.
c. The *curved-field* reflectron (cont.)

Expansion of high fragment ion regions from high resolution *post-source decay* mass spectrum of angiotensin II obtained on a Kratos AXIMA-CFR, showing isotopic mass resolution.
c. The *curved-field* reflectron (cont.)

PSD sensitivity: high resolution *post-source decay* mass spectrum of 100 fmol of insulin B chain obtained on a Kratos AXIMA-CFR.
d. Nested PSD

PSD of intact peptide and one carboxypeptidase fragment

Peaks in both spectra with a common mass are \( a \), \( b \) or \( c \) ions; peaks with shifted masses are \( x \), \( y \) or \( z \) ions


3. Tandem Mass Spectrometry

a. Mass filters vs. mass analyzers
b. Mass selection
c. Collision-induced dissociation (CID)

In a normal configuration:

Mass filter: passes a single mass  
Collision gas: He or Xe  
Collision-induced dissociation  
High energy (1-20 Kev) single collisions, or  
Low energy (10-50 ev multiple collisions.  
Mass analyzer: records a mass spectrum

Normal mode = product ion scan
4. Triple Quadrupoles

The quadrupole is a *mass filter*. To record a mass spectrum it must be scanned.

Because both mass spectrometers are mass filters, the duty cycle is low.
Scan modes of the triple quadrupole:

**Product ion (normal) scan:**
- mass filter: single precursor mass
- RF-only mode: mass analyzer: Scan product masses for selected precursor

**Precursor ion scan:**
- mass analyzer: Scan precursor masses for selected product
- RF-only mode: mass filter: single product mass

**Constant neutral loss (CNL) scan** (*reaction ion monitoring*):
- mass analyzer: scan precursor masses
- RF-only mode: mass analyzer: scan product masses differing by constant mass difference
5. Quadrupole Ion Traps

a. Tandem-in-time

trapping cycle: fundamental (1.1 MHz) RF voltage on ring electrode
mass isolation cycle (MS1): resonant ejection of all but selected ion, using
high amplitude supplementary RF on ring electrode
excitation cycle (low energy CID): low amplitude supplementary RF voltage
on endcaps
mass analysis cycle (MS2): resonance ejection mode, high amplitude
supplementary RF voltage on endcaps while scanning the amplitude of the
fundamental RF voltage on the ring electrode
b. Mass isolation and excitation

Isolation and excitation can be carried out by scanning the resonant frequency or by a stored-waveform inverse Fourier transform (SWIFT) pulse.

6. **Atmospheric Pressure MALDI on an ion trap: examples of MS, MS/MS and MS^n**

**Mass Spectra:**

- bombesin fragment 8-14 (812.0 Da)
- amyloid ß-protein fragment 25-35 (1060.3 Da)
- angiotensin I (1296.5 Da)
- α-MSH (1664.9 Da)
6. Atmospheric Pressure MALDI on an ion trap: examples of MS, MS/MS and MS^n (cont.)

MS/MS spectrum:
mass selection of the MH^+ ion of angiotensin I (m/z 1297.5)

MS/MS spectrum of four peptide mixture (50 pg/µL of each peptide) at m/z 1297.5.

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
6. Atmospheric Pressure MALDI on an ion trap: examples of MS, MS/MS and MS$^n$ (cont.)

**MS$^3$ spectrum:**

mass selection of the b$_9$ ion of angiotensin I (m/z 1165.6)

MS$^3$ spectrum of four peptide mixture (0.5 ng/µL of each peptide) at m/z 1297.5 → 1165.5 →.
7. Fourier transform mass spectrometry

Electrospray ionization


Figure 2. ESI-FT-ICR mass spectrum of the oligomannose 6 hydrolysate products after a 24-h resin hydrolysis. Spectrum displays the ladder sequence of the N-linked oligosaccharide. The dashed lines in the oligosaccharide structure represent the glycosidic linkages that have been cleaved by acid hydrolysis.
8. **Tandem Time-of-Flight Mass Spectrometers**

a. High energy vs. low energy collisions

**High energy collisions**

precursor ion kinetic energy: $E_1 = 1-20$ keV

single collision conditions

product ion energy:

$$E_2 = \frac{m_2}{m_1} E_1$$

used in instruments in which mass measurement depends upon kinetic energy, i.e. magnetic sector and TOF mass spectrometers

produces more abundant side-chain fragmentation than low energy collisions

can be used to distinguish leucine and isoleucine
8. *Tandem Time-of-Flight Mass Spectrometers*

a. High energy vs. low energy collisions (cont.)

High energy collisions (cont.)

collision energy ($E_{\text{rel}}$) in the center-of-mass frame

\[
E_{\text{rel}} = \frac{m_n}{m_n + m_M} E_M
\]

where $M$ is the molecular ion and $n$ is the collision gas

Low energy collisions

precursor ion kinetic energy = 10-50 ev

activation through multiple collisions

used in instruments in which mass measurement is not a function of kinetic energy, i.e. quadrupoles and ion traps
b. Dual reflectron tandem time-of-flight

In a dual reflectron instrument, the collision chamber was replaced with a pulsed valve and a *curved-field reflectron* in MS2 was used to focus the wide range of product ions.

c. delayed extraction focusing as MS1

Kratos AXIMA-CFR uses delayed extraction as MS1, with PSD on a curved-field reflectron as MS2.

In effect, a tandem TOF, but without a collision chamber.
c. delayed extraction focusing as MS1
d. photodissociation


1. Note that this avoids relative energy problem
2. But reflectron voltage must still be stepped if a single or dual stage reflectron is used.

**Figure 1.** Diagram of the home-built DE MALDI-TOFMS for photodissociation studies. The distance from the source to the ion selection optics and photodissociation region are 0.52 and 0.58 m, respectively. The distance from the ion selection region to the entrance of the reflectron is 0.9 m.

**Photofragment ion mass spectrum of Substance P-amide (RPKPQQFFGLM-NH₂)**
e. Pulsed extraction in MS2

If a single-stage reflectron is used, then ions can be re-accelerated after collision to limit the need for scanning the reflectron voltage.

In the Applied BioSystems tandem (TOF/TOF) instrument, re-acceleration is accomplished by pulsed extraction.
9. Hybrid instruments

Because the first mass spectrometer acts as a mass filter, hybrid tandem instruments usually use a scanning instrument, such as a magnetic sector or quadrupole for MS1.

Conversely, the second mass spectrometer should be one that has the multichannel recording advantage, such as a TOF mass spectrometer.

a. Sector/orthogonal TOF

It is difficult to achieve high energy collisions in an orthogonal configuration.

This configuration uses 800 ev collisions with a heavy gas such as Xe to achieve high energy collisions while limiting the lateral motion of the ions.

9. Hybrid instruments (cont.)

b. Trap/TOF

Why a trap/TOF, when the trap is itself a tandem?

The TOF has higher mass range, can record product ions of higher m/z than precursor.

9. Hybrid instruments (cont.)

b. QTOF

Q₁ is an RF-only quadrupole filter that collimates ions from high pressure source.

Q₂ is a quadrupole mass filter that selects the precursor mass.

Q₃ is an RF-only quadrupole filter used as a collision chamber.

Collisions are low energy.

The TOF makes it possible to analyze product ions with higher m/z than their precursors.

___________________________
___________________________
___________________________
___________________________
___________________________
___________________________

Electrospray CID mass spectra of peptides: NFNRHLHFTLVKDR and LLSYDDEAFIRDVAK on a QTOF mass spectrometer


**Homework problem:** calculate the masses of b and y ions for the two peptides shown and compare with results obtained in their mass spectra.
Electrospray ionization CID mass spectrum of the [M+Na]^+ ion from the 3-AQ derivative of (GlcNAc)_2Man_3.

10. Electrospray ionization CID mass spectrum of multiply-charged ions

- Because their internal energies are increased by charge repulsion, multiply-charged ions fragment easier.
- For peptides, doubly-charged molecular ions produce singly-charged fragment ions.

- Singly-charged fragment ions may have m/z values greater than the molecular ion.
10. Electrospray ionization CID mass spectrum of multiply-charged ions (cont.)

Fragment ions from the doubly-charged molecular ion of phospholipids

11. Conclusions: Analysis of proteins and glycoproteins


![Figure 4](attachment:image.png)

*Figure 4.* Mass Spectrometric analysis of fraction 36. (A) MALDI-TOF spectrum of the fraction before digestion. (B) MALDI-TOF spectrum of the tryptic digest. (C) MALDI-TOF spectrum of the tryptic digest after ladder sequencing. (D) ESI-MS/MS spectrum of a peptide at m/z 380.5 (2− ion).
Lipids

1. Introduction to lipids.
2. Sample handling
3. Phospholipids
4. GPI and phospholipases
5. Other polar lipids
6. GPI example
7. LPS and lipid A
8. Glycolipids
1. Introduction

Lipids constitute about 50% of the mass of most animal cell membranes. Lipid molecules in membranes are amphipathic. The most abundant are phospholipids. The most abundant phospholipids are phosphatidylethanolamine, phosphatidylcholine, phosphatidylserine and sphingomyelin.

The study of membranes from pathogenic bacteria is also very important because the membrane allows bacteria to colonize its hosts tissues and evade the immune system. Most antibiotics target bacterial membranes. The study of biomolecules such as bacterial surface glycolipids or lipooligosaccharides (LOS) allows researchers to better understand bacterial pathogenesis and how to prevent it.
1. Introduction (cont)

LOS consist of 2 parts:

- A hydrophobic and anionic lipid A moiety consisting of two 1-6 linked glucosamine sugars substituted with two phosphate groups and up to six or more N- and O-linked fatty acids.

- A complex oligosaccharide that is linked to lipid A through an acidic sugar. The oligosaccharide region extends outside the membrane and usually interacts with the host cell surface and molecules.
2. Sample Handling

The same matrices used for oligosaccharides, should be used for lipids.

DHB in water or in 50% ethanol or acetonitrile is the most widely used. Negative in mode usually gives better spectra.

When using ESI negative ion mode also gives better spectra. Samples are injected in a stream of water:acetonitrile. ESI-ion trap instruments such as the LCQ are ideal for figuring the sequence of lipids because of the ability to do MS\textsuperscript{n}. It fragments fatty acid and gives very characteristic spectra.
The metabolism of carbohydrates and lipids are interrelated, as seen in the diagram below.
3. **Phospholipids**

- Phosphatidylinositol
- Phosphatidylcholine
- Phosphatidylethanolamine
- Phosphatidylserine
4. GPI and phospholipases

Glycosylphosphatidylinositol (GPI) anchors proteins to the non-cytoplasmic side of the membrane. The proteins can be released by adding the enzyme Phosphatidylinositol specific phospholipase C which cleaves the protein from their anchor.

Various phospholipases can be used to help characterize the various components that are attached to the glycerol molecule.

Phospholipase A1
Phospholipase A2
Phospholipase C
Phospholipase D
5. Other polar lipids

Triacyl Glycerol

Stearic Acid

Hydrophilic

Hydrophobic
6. GPI Example

a. RP-HPLC profile at $A_{210}$ of CD1d1 and control endogenous HLA-associated ligands. The arachidonic acid has conjugated double bonds in its chain that absorb at 210 nm, thus giving the large peak indicated by the arrow in the chromatograph.

b. MALDI mass spectrum of the fraction containing CD1d1-associated ligand

6. GPI Example (cont)

c. Positive ion MALDI mass spectrum of fraction containing CD1d1-associated ligand after aminopeptidase digest which is dissolved in saturated ammonium sulfate.

d. Negative ion MALDI mass spectrum of the same sample.

A combination of the Ammonium Sulfate added during sample preparation and In source decay hydrolyzed and fragmented the GPI.
6. **GPI Example (cont)**

a. RP-HPLC profile at A$_{210}$ of the soluble CD1d1-associated and control ligands.

b. Positive ion MALDI mass spectrum of sCD1d1-associated ligand.

**A combination of the Ammonium Sulfate added during sample preparation and In source decay hydrolyzed and fragmented the GPI.**

7. LPS and lipid A

Lipid A (endotoxin), the hydrophobic anchor of lipopolysaccharide (LPS), is a glucosamine-based phospholipid that makes up the outer monolayer of the outer membranes of most Gram-negative bacteria.
There are ~10^6 lipid A residues and ~10^7 glycerophospholipids in an *Escherichia coli* cell. The minimal LPS required for the growth of *E. coli* consists of lipid A and two Kdo units. In wild type strains, additional core and O-antigen sugars are present. These are not required for growth, but enhance survival during environmental stresses, and help bacteria evade the immune system. In *Haemophilus influenzae*, the heptose residues of the core are needed for virulence in infant rats.
7. LPS and lipid A (cont)

FIG. 4. MALDI-TOF mass spectrometry of lipid A isolated from E. coli RO138 (lpxA2 recA) complemented by expression of LpxA variants with reversed chain length selectivity. A, RO138/pTO17 (P. aeruginosa M169G lpxA); B, RO138/pTO16 (E. coli G173M lpxA); C, RO138/pTO12 (E. coli G173M lpxA-high copy vector). pTO12 contains E. coli G173M lpxA on the Xba1-HindIII fragment from pTO101 (Table 1) ligated into pBluescriptSK (Stratagene) cut with the same two enzymes. pTO16 contains E. coli G173M lpxA on a SacI fragment from pTO12 ligated into the lower copy plasmid pNGH1 (8), also cut with SacI. The gene is oriented for expression using the lac promoter in pNGH1. pTO17 was constructed in a similar manner by ligating SacI-digested pNGH1 and the relevant SacI fragment of pTO13. pTO13 contains P. aeruginosa M169G lpxA on the Xba1-HindIII fragment from pTO201 (Table 1) ligated into pBluescriptSK cut with the same two enzymes. Each plasmid was transformed into RO138 (lpxA2 recA) (8). Cultures of RO138/pTO16 and RO138/pTO17 (150 ml each) were grown at 42 °C (225 rpm) in LB broth (29), containing 20 μg/ml chloramphenicol, to A_{660} = 2. A 250-ml culture of RO138/pTO12 was grown at 30 °C (225 rpm) in LB broth, containing 100 μg/ml ampicillin, to A_{600} = 0.6 and then shifted to 42 °C for 3 more h. Lipid A was prepared from each culture and stored at -80 °C prior to mass spectrometry (8). The proposed structures for the major lipid A species of A and C are indicated. The peak at m/z 1741.9 in B is attributed to a mixture of lipid A species containing one O-linked hydroxymyristate and one O-linked hydroxydecanate residue (structures not shown). The error in the determination of m/z is approximately 1 atomic mass unit.

Diphosphoryl lipid A in 1 M NaCl.

1. Polyethylene Membrane
2. Polypropylene membrane

In each experiment 0.5 pmol of lipid A was spotted.

7. **LPS and lipid A (cont)**

Diphosphoryl lipid A in 1% Triton-X 100.

1. Polyethylene membrane

2. Polypropylene membrane


In each experiment 0.5 pmol of lipid A was spotted.

7. LPS and lipid A (cont)

Diphosphoryl lipid A in 10% SDS.

1, Polyethylene Membrane
2, Polypropylene membrane;
3, Stainless steel probe.

In each experiment 0.5 pmol of lipid A was spotted.


---

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
7. LPS and lipid A (cont)

Fig. 5. A. Diagrammatic representation of the lgt locus Neisseria gonorrhoeae strain 1291. Coding sequences are indicated by solid arrows. Frameshift mutations in the homopolymeric tract, which put the gene out of frame for expression, are indicated by striped arrows. In each case where a homopolymeric tract of guanosine (G) residues is present in a gene, its position and the number of residues in the tract are indicated below the gene.

B. Negative ion MALDI spectra of the O-deacylated LOS from 1291 wild type (top) and 1291 lgtB mutant (bottom). Partial structures for the most abundant LOS glycoform in each strain are shown above each spectrum, where Lipid A represents the various phosphorylated forms of O-deacylated Lipid A (see text and Table 1). In addition, origin and masses of the associated prompt fragments are indicated for the two major LOS-glycoforms via cleavage between the Kdo and Lipid A glycosidic bond, a process described in detail elsewhere (Gibson et al., 1997). In the molecular ion regions, peaks marked with asterisks are salt adducts and other unlabelled (M-H)^- species are listed in Table 1. Additional peaks are also present in the prompt fragment regions that are not discussed in the text. These peaks are further decompositions of the oligosaccharide fragments, including loss of CO₂ (-44 Da) and/or Kdo (-220 Da) and have been previously described (Gibson et al., 1987). It should be noted that the hydrazine treatment used to prepare the O-deacylated LOS would have also removed the O-acetate group previously identified on the 1291 wild-type oligosaccharide (John et al., 1991).


ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
7. LPS and lipid A (cont)


Figure 9. Molecular ion region of intact LOS from S. typhimurium Ra in the negative-ion mode. Note the dominance of the expected hexaacyl substituted lipid A with the small amount of tetraacyl species at lower masses corresponding to the loss of myristic acid (ΔM = 210.36 u, C₁₄H₂₉O) and β-hydroxymyristic acid (ΔM = 226.36 u, C₁₄H₂₆O₂), for example ~438.7 u, m/z 4028 (hexaacyl LOS) and 3992 (tetraacyl LOS). Impurity peaks are marked with asterisks (see also the corresponding O-deacylated LOS spectrum in Figure 9).
7. LPS and lipid A (cont)

Table 2. Structures and molecular weights of intact LOS

<table>
<thead>
<tr>
<th>Bacterial Strain Major LOS Structure(s)</th>
<th>Expt. (M_r)</th>
<th>(+/-) Moiety(^a)</th>
<th>Calc. (M_r)</th>
<th>(\Delta M)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>S. typhimurium</em> Ra LPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gal (\downarrow) P</td>
<td>3825.4</td>
<td>(-P, PEA)</td>
<td>3827.0</td>
<td>(-1.6)</td>
</tr>
<tr>
<td>GlcNAc (\rightarrow) Glc (\rightarrow) Gal (\rightarrow) Glc (\rightarrow) Hep (\rightarrow) Hep (\rightarrow) (\text{(Kdo)}_2) \rightarrow) lipid A(^b)</td>
<td>3906.4</td>
<td>(-PEA)</td>
<td>3906.9</td>
<td>(-0.5)</td>
</tr>
<tr>
<td>H(\uparrow) PPEA</td>
<td>3949.5</td>
<td>(-P)</td>
<td>3960.0</td>
<td>(-0.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-EA)</td>
<td>3986.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4029.4</td>
<td>(-)</td>
<td>4030.0(^c)</td>
<td>(-0.6)</td>
</tr>
<tr>
<td>where lipid A = diphosphorylhexaacyl</td>
<td>4068.8</td>
<td>(+EA)</td>
<td>4066.9</td>
<td>(+1.9)</td>
</tr>
<tr>
<td></td>
<td>4108.8</td>
<td>(+P)</td>
<td>4110.0</td>
<td>(-1.2)</td>
</tr>
<tr>
<td>where lipid A = diphosphoryltetraacyl</td>
<td>(\approx) 3470</td>
<td>(-P)</td>
<td>3470.2</td>
<td>(\leq 1.0)</td>
</tr>
<tr>
<td></td>
<td>(\approx) 3514</td>
<td>(-EA)</td>
<td>3513.3</td>
<td>(\leq 1.0)</td>
</tr>
<tr>
<td></td>
<td>3590</td>
<td>(-)</td>
<td>3593.3</td>
<td>(-0.4)</td>
</tr>
</tbody>
</table>

\(^a\) Abbreviations: P, phosphate; PEA, phosphoethanolamine; Gal, galactose; Glc, glucose; GlcNAc, N-acetylglucosamine; Hep, heptose; Kdo, 2-keto-3-deoxyoctulosonic acid.

\(^b\) Lipid A is present primarily in its hexaacyl form with an average \(M_r\) of 1798.4. A second major LOS series is also observed with the lipid A in the tetraacyl form (see text; average \(M_r = 1361.7\)).

\(^c\) The \(M_r\)s of the major predicted O-deacetylated LOS or LOS from each strain are shown in bold type.

8. Glycolipids

The figure shows an example of a glycosphingolipid. This is the acetyl derivative. The PSD-spectrum gives information on the glycan chain. However CID had to be used to obtain the structure between the two chains of the ceramide.

The best work in the carbohydrates and glycolipids field has been done by the following researchers:
Catherine Costello
David Harvey
Vernon Reinhold

References:
Carbohydrates

1. Introduction
2. Structures of carbohydrates
3. MALDI MS of carbohydrates
4. Carbohydrate fragmentation in the mass spectrometer
5. Post-source decay
6. Chemical methods of sequencing
1. Introduction

Carbohydrates are one of the most abundant and structurally diverse compounds in nature. Proteins and oligonucleotides are linear polymers while carbohydrates form branched structures, as the glycosidic bond can occur at a number of positions on each monosaccharide.

-Carbohydrates are polyhydroxy ketones and aldehydes (reducing group).
   Aldoses have an aldehyde at the carbonyl group.
   Ketoses have a ketone at the carbonyl group.
   Each C except for the carbonyl C has a OH group.

-Carbohydrates consist of compounds with the general formula

\[ C_nH_{2n}O_n \]

Where n= 3 to 9 with n = 6 being the most abundant.

-Deoxy: absence of one or more oxygen.
-Uronic Acid: Oxidation of one of the primary OH group to a carboxylic acid.
In the case of glucose-like sugars two types of rings can be formed:

α-D-Glucopyranose: α means the OH group attached to C-1 is below the plane of the ring.

β-D-Glucopyranose: β means the OH group attached to C-1 is above the plane of the ring.

**O-glycosidic bonds:** links sugars between C-1 of one sugar and hydroxyl oxygen atom on C-4 of the adjacent sugar. The C-1 carbon is called the **anomeric carbon** atom.

**N-glycosidic bonds:** links sugars between C-1 of one sugar and N atom of an amine of an adjacent aminated sugar or an oligonucleotide base.

**Oligosaccharides:** are formed by O-glycosidic bonds between 2-10 monosaccharides.

**Polysaccharides:** are formed by O-glycosidic bonds between 11 or more monosaccharides.
Anomeric C atom

α-D-glucose

β-D-Glucose

β-D-mannose

α-D-glucose

Boat

α-D-galactose

Chair
2. Structures of carbohydrates

Most proteins are glycosylated. They contain three types of glycans:

- **N-linked oligosaccharides**
- **O-linked oligosaccharides**
- **Glycosylphosphatidylinositol (GPI) lipid anchors**

**N-linked oligosaccharides** are linked to the amide N in the side chain of Asn in the consensus sequence Asn-Xaa-Ser/Thr, where Xaa can be any amino acid besides Pro and Asp.

**N-glycans** can be subdivided into three groups

- 'high mannose'
- 'hybrid'
- 'complex'
with the common pentasaccharide core –

\[
\text{Man}(\alpha 1,6)-(\text{Man}(\alpha 1,3))-\text{Man}(\beta 1,4)-\text{GlcNAc}(\beta 1,4)\text{GlcNAc}(\beta 1,N)-\text{Asn} –
\]

in all three groups. (Besides GlcNAc the following sugars have been found N-linked to Asn, mainly on bacterial glycoproteins: Glc, GalNAc and L-Rha.

\[
\begin{align*}
\text{Man}\alpha 1,2-\text{Man}\alpha 1,6 \\
\text{Man}\alpha 1,2-\text{Man}\alpha 1,3 \\
\text{Man}\alpha 1,2-\text{Man}\alpha 1,2 — — \text{Man}\alpha 1,3 \\
\text{Man}\alpha 1,3 \rightarrow \text{Man}\alpha 1,6 \\
\text{Man}\alpha 1,4-\text{GlcNAc}\beta 1,4-\text{GlcNAc}\beta 1,N
\end{align*}
\]

*High Mannose Type*
2. Structures of carbohydrates (cont.)

**complex type chains:** occur also as mono-, bi-, tri (2,4- and 2,6-branched), and tetra-antennary structures and differ in the amount and type of sialylation.

All three types originate from one precursor oligosaccharide which contains the described

- **common pentasaccharide core Man3-GlcNAc2**
- **some additional sugar residues**
- **the non-reducing end**

It is then processed enzymatically to yield these three types.
2. Structures of carbohydrates (cont.)

The hydroxyl group of Ser/Thr is involved in hydrogen bonding during the enzymatic attachment of the oligosaccharide precursor molecule to yield a favorable conformation for the action of the oligosaccharyltransferase.

The steric hindrance of Proline might be too large, thus preventing glycosylation at Pro containing sites.

The negative influence of aspartic acid may be due to the negative charge on the side chain of this residue.

In some cases Ser/Thr is replaced by cysteine. While Ser replacement by Cys generally leads to decreased glycosylation, substitution by Thr at a given potential glycosylation site can lead to increased glycosylation.

Although there are many potential glycosylation sites in a protein, glycosylation occurs only at one third of them, mostly at those sites where the surrounding amino acids allow the formation of a beta turn.
2. Structures of carbohydrates (cont.)

**O-linked Oligosaccharides**: are smaller, varied and there is no common core structure, and no consensus motif for locating it on the protein. They are attached to **Serine or Threonine**. Several O-linked sites are usually found in close proximity.

**GPI Anchors**: are glycolipids attached to the carboxy terminus of some proteins whose function is to anchor those proteins to membranes.
Blood type is determined by the carbohydrate on the surface of the RBC

**TYPE O**

- Gal
- GlcNAC
- R
- Fuc

**TYPE A**

- GalNAc
- Gal
- GlcNAC
- R
- Fuc

**TYPE B**

- Gal
- GlcNAC
- R
- Fuc
3. MALDI MS of Carbohydrates


Matrices:

2,5-dihydroxybenzoic acid (DHB): The most effective and widely used matrix. I usually use a saturated solution in 50% ethanol, however DHB is also used in water or in 50-% acetonitrile. It produces mostly MNa⁺ rather than MH⁺.

3-aminoquinoline: Has a lower background and an increased resolution, however it sublimes too rapidly.

B-Carbolines: such as harmane, nor-harmane and harmol are comparable to DHB, however they are much more expensive. They produce more MH⁺ than MNa⁺. Harmane is the recommended matrix with sulfated glycans.

Ferrulic Acid: Has the disadvantage of being insoluble in water. So it is hard to use with soluble sugars.
6-Aza-2-Thiothymine: is very effective with acidic glycans.

2,4,6-Trihydroxyacetophenones: in a 1:1 acetonitrile: 20 mM ammonium citrate is a very good matrix for sialylated glycans, and glycoproteins.

2,6-Dihydroxyacetophenone is a good matrix for glycoproteins.

Derivatization:

Carbohydrates are not ionized as efficiently as proteins or transferred to the vapor phase as effectively. To remedy these problems derivatization has been used.

a. **Derivatives formed by reductive amination:** The sugar is dissolved in mild acid to open the ring and reacted with excess of amine to form a **Shiff base** in the presence of a reducing agent.

b. **Carbonyl derivatives:** Substituted oxime formation by adding a basic peptide residue in its aminooxyacetyl form to the reducing terminus of neutral N-linked glycans.

c. **Methylation of the carboxylic acid of sialic acids** to produce a neutral sugar. Makes it easier to detect carbohydrates in positive mode.
3. MALDI MS of Carbohydrates (cont.)

<table>
<thead>
<tr>
<th>Oligosaccharide</th>
<th>Residue formula</th>
<th>Residue mass</th>
<th>Methyl Residue Mass</th>
<th>Methyl No. Me</th>
<th>Acetyl Residue Mass</th>
<th>Acetyl No. Ac</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deoxy-pentose</td>
<td>C₅H₁₀O₃</td>
<td>116.047</td>
<td>130.063</td>
<td>1</td>
<td>158.058</td>
<td>1</td>
</tr>
<tr>
<td>Pentose</td>
<td>C₆H₁₀O₄</td>
<td>116.117</td>
<td>130.144</td>
<td>2</td>
<td>158.154</td>
<td>1</td>
</tr>
<tr>
<td>Deoxy-hexose</td>
<td>C₆H₁₆O₄</td>
<td>132.042</td>
<td>160.074</td>
<td>2</td>
<td>216.063</td>
<td>2</td>
</tr>
<tr>
<td>Hexose</td>
<td>C₆H₁₆O₅</td>
<td>132.116</td>
<td>160.170</td>
<td>2</td>
<td>216.191</td>
<td>2</td>
</tr>
<tr>
<td>Hexosamine</td>
<td>C₆H₁₆N₂O₄</td>
<td>146.078</td>
<td>174.089</td>
<td>2</td>
<td>230.079</td>
<td>2</td>
</tr>
<tr>
<td>HexNAc</td>
<td>C₆H₁₆N₂O₅</td>
<td>146.143</td>
<td>174.197</td>
<td>2</td>
<td>230.218</td>
<td>2</td>
</tr>
<tr>
<td>Hexuronic-Acid</td>
<td>C₆H₁₆O₅</td>
<td>162.053</td>
<td>204.100</td>
<td>3</td>
<td>288.084</td>
<td>3</td>
</tr>
<tr>
<td>Kάο</td>
<td>C₄H₂O₂</td>
<td>162.142</td>
<td>204.233</td>
<td>3</td>
<td>288.254</td>
<td>3</td>
</tr>
<tr>
<td>1-N-acetyl neuraminic acid</td>
<td>C₁₁H₁₇N₂O₄</td>
<td>161.069</td>
<td>217.131</td>
<td>4</td>
<td>287.100</td>
<td>3</td>
</tr>
<tr>
<td>1-N-glycol neuraminic acid</td>
<td>C₁₁H₁₇N₂O₄</td>
<td>161.157</td>
<td>217.265</td>
<td>3</td>
<td>287.269</td>
<td>3</td>
</tr>
<tr>
<td>2-N-acetyl neuraminic acid</td>
<td>C₁₁H₁₇N₂O₄</td>
<td>203.079</td>
<td>245.126</td>
<td>3</td>
<td>287.100</td>
<td>3</td>
</tr>
<tr>
<td>2-N-glycol neuraminic acid</td>
<td>C₁₁H₁₇N₂O₄</td>
<td>203.179</td>
<td>245.276</td>
<td>3</td>
<td>287.269</td>
<td>3</td>
</tr>
</tbody>
</table>


Masses of intact glycans can be obtained by addition of the residue masses plus the figure for the relevant terminal groups that are listed in Table 3.

<table>
<thead>
<tr>
<th>TABLE 3. Masses of terminal groups to be used with masses listed in Table 2 to obtain the molecular weights of carbohydrates.</th>
<th>Methyl</th>
<th>Acetyl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reducing end</td>
<td>Underivatized</td>
<td>Methyl</td>
</tr>
<tr>
<td>Free</td>
<td>18.011</td>
<td>46.042</td>
</tr>
<tr>
<td>Reduced</td>
<td>18.153</td>
<td>46.069</td>
</tr>
</tbody>
</table>

1Top figure = monoisotopic mass; lower figure = average mass.

<table>
<thead>
<tr>
<th>TABLE 4. Masses of common adducts for carbohydrates ionized by MALDI.</th>
<th>Monoisotopic mass</th>
<th>Average mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adduct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrogen</td>
<td>1.007825</td>
<td>1.00794</td>
</tr>
<tr>
<td>Sodium</td>
<td>22.9898</td>
<td>22.9896</td>
</tr>
<tr>
<td>Lithium</td>
<td>7.0160</td>
<td>6.941</td>
</tr>
<tr>
<td>Potassium</td>
<td>38.9637</td>
<td>39.0983</td>
</tr>
<tr>
<td>Cesium</td>
<td>132.9054</td>
<td>132.9054</td>
</tr>
</tbody>
</table>

1Most abundant isotope.

<table>
<thead>
<tr>
<th>TABLE 5. Masses of additional groups found on carbohydrates.</th>
<th>Methyl</th>
<th>Acetyl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
<td>Underivatized</td>
<td>Methyl</td>
</tr>
<tr>
<td>Sulfate</td>
<td>79.957</td>
<td>65.941²</td>
</tr>
<tr>
<td>Phosphatate</td>
<td>79.966</td>
<td>66.037</td>
</tr>
<tr>
<td>Inositol</td>
<td>162.053</td>
<td>218.115</td>
</tr>
</tbody>
</table>

1Top figure = monoisotopic mass; lower figure = average mass. ²Assuming that no derivatization occurs.
4. Carbohydrate fragmentation in the mass spectrometer

Fragmentation depends on:

• The ion formed [MNa\(^+\), MH\(^+\)]. Protonated species decompose more readily than cationized species.
• Linear carbohydrates give more fragments than branched ones.
• The charge state.
• The energy deposited into the ion.
• The time available for fragmentation.
4. Carbohydrate fragmentation in the mass spectrometer (cont.)

There are two types of fragmentation:

- **Glycosidic bond cleavages** that are the result of the cleavage of a bond linking two sugar rings. They provide information on sequence and branching.

- **Cross ring cleavages** that involve the breaking of two bonds, and provide information on linkage.

  - X [cross ring], Y and Z Ions: are ions that retain the charge at the reducing terminus (glycosidic).
  - A [cross ring], B and C Ions: are complementary ions.

  - Sugar ions are numbered from the reducing end for X, Y and Z, and from the non-reducing end for A, B and C.
4. Carbohydrate fragmentation (cont.)

Fragmentation occurring during spectrum acquisition can help solve structure. Spectrum b is an example of how derivatization can improve and increase fragmentation.

**SCHEME 4.** Proposed structure for ion D. Symbols are defined in the legend to Fig. 11.

**FIGURE 11.** (a) Positive ion MALDI mass spectrum of a sialylated biantennary glycan (200 pmol) recorded on a magnetic sector mass spectrometer from 2,5-DHB, (b) MALDI spectrum of the same compound after methylation of the carboxylic acid groups. Major glycosidic (B and Y) and cross-ring (A) fragment ions are labelled. Key to symbols: ■ GlcNAc, □ galactose, ○ mannose, ● stialic acid. (From Powell & Harvey, 1996.)
Derivative Formation by Reductive Amination

\[ \text{GlcNac} = \square \]
\[ \text{Gal} = \bigcirc \]

**FIGURE 18.** MALDI–PSD spectra of the benzylamino derivatives of (a) the MNa\(^+\) and (b) the MH\(^+\) ions of the linear glycan LNT ionized from 2,5-DHB. Symbols for the monosaccharides are as in Fig. 14. (From Lemoine, Chirat, & Domon, 1996).

**FIGURE 19.** MALDI–PSD spectra of the benzylamino derivatives of three isomeric pentasaccharides ionized from 2,5-DHB. Symbols for the monosaccharides are as in Fig. 14 plus \(\bullet\) or \(\bigcirc\) = fucose. (From Lemoine, Chirat, & Domon, 1996).
5. Post Source Decay

In PSD fragmentation is dependent on:

Laser power: More ions are seen at lower mass with high laser power due to successive cleavages.

Matrix: In linear mode, if DHB is used there is more loss of sialic acid from a carbohydrate molecule than if ATT is used. In reflectron mode there is significant sialic acid loss with both matrices.

In the example in the figure, CHCA matrix efficiently protonates the 2-aminopyridine used to derivatize the carbohydrate giving an MH$^+$ ion, which fragments resulting in a Y series of fragments (fig. a), while DHB generates a MNa$^+$, with non-localised charge, which fragment to give a mixture of Y and B ions (fig. B).
5. Post Source Decay (cont.)

PSD spectra of MNa\(^+\) ions from neutral carbohydrates mostly show glycosidic and internal cleavage reactions with few contributions from cross ring products.

B and Y cleavages give information on sequence and branching. Linkage information is derived from cross-ring cleavage ions.

PSD of a 3-branched triantennary N-linked glycan

![MALDI-PSD spectra of a complex 3-branched triantennary N-linked glycan (a) and a complex 6-linked glycan (b). Symbols for the constituent monosaccharides are: ■ = GlcNAc, □ = fucose, ○ = mannose, ◇ = core mannose, ● = galactose. (From Rouse et al., 1998, with permission from Academic Press).]
6. Chemical methods of sequencing

N-linked glycans can be released by alkaline degradation (figure). A strong base causes successive loss of residues from the reducing end by glycosidic cleavage. This is also called a peeling reaction; because the cleavages are from the reducing end it gives information about sequence and branching.

N-linked glycan released by alkaline degradation
O-Linked Glycans are released by β-elimination with alkali reagents. The resulting reducing sugar is unstable at high pH, so sodium borohydride is added to reduce them to alditol which leads to:

- release of N-linked sugars.
- the reducing terminus is removed.

A better way to release O-linked sugars is with hydrazine treatment at 65°C under anhydrous conditions to keep the reducing terminus intact.
<table>
<thead>
<tr>
<th>Enzyme</th>
<th>EC number</th>
<th>Source</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-D-Sialidase</td>
<td>3.2.1.18</td>
<td><em>Arthrobacter ureafaciens</em></td>
<td>NeuNAc/Gcα2 → 6 &gt; 3,8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Newcastle disease virus</td>
<td>NeuNAc/Gcα2 → 3,8</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Clostridium perfringes</em></td>
<td>NeuNAc/Gcα2 → 3,6,8</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Vibrio cholerae</em></td>
<td>NeuNAc/Gcα2 → 3,6,8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bovine testis</td>
<td></td>
</tr>
<tr>
<td>β-D-Galactosidase</td>
<td>3.2.1.23</td>
<td><em>Streptococcus pneumoniae</em></td>
<td>Galβ1 → 3 &gt; 4 &gt; 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Jack bean</em></td>
<td>Galβ1 → 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>(Canavalia ensiformis)</em></td>
<td>Galβ1 → 6 &gt; 4 ≥ 3</td>
</tr>
<tr>
<td>α-D Galactosidase</td>
<td>3.2.1.22</td>
<td><em>Escherichia coli</em></td>
<td>Galβ1 → 4Glc</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Green coffee bean</td>
<td>Galα1 → 3,4,6</td>
</tr>
<tr>
<td>β-N-Acetyl-d-hexosaminidase</td>
<td>3.2.1.30</td>
<td>Jack bean (0.01 U/mL)</td>
<td>Glc(Gal)NAcβ1 → 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(10 U/mL)</td>
<td></td>
</tr>
<tr>
<td>α-N-Acetyl-d-hexosaminidase</td>
<td>3.2.1.49</td>
<td>Chicken liver</td>
<td>Glc(Gal)NAcβ1 → 2,3,4,6</td>
</tr>
<tr>
<td>α-D-Mannosidase</td>
<td>3.2.1.24</td>
<td><em>Aspergillus phoenicus</em></td>
<td>GalNAcα1 →</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Helix pomatia</em></td>
<td>Manα1 → 2,3,6</td>
</tr>
<tr>
<td>β-D-Mannosidase</td>
<td>3.2.1.25</td>
<td><em>Achatina fulica</em></td>
<td>Manα1 → 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Charonia lampas</em></td>
<td>Manβ1 → 4</td>
</tr>
<tr>
<td>α-L-Fucosidase</td>
<td>3.2.1.51</td>
<td>Bovine epididymis</td>
<td>Fucα1 → 6 &gt; 2,3,4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Almond emulsin II</td>
<td>Fucα1 → 2</td>
</tr>
<tr>
<td>β-L-Xylosidase</td>
<td>3.2.1.103</td>
<td><em>Charonia lampas</em></td>
<td>Xylβ1 → 2</td>
</tr>
<tr>
<td>Endo-β-D-galactosidase</td>
<td></td>
<td><em>Escherichia freundii</em></td>
<td>RGalβ1 → 4GlcNAc</td>
</tr>
</tbody>
</table>
Strategies for Characterization of DNA

1. Molecular Weight Measurements: How big?
   a. Sodium adducts
   b. Base losses
   c. IR vs UV lasers

2. Ladder Sequencing
   a. Sanger-type sequencing
   b. exonuclease ladders
   c. hydrolysis ladders

3. Hybridization probes
   a. example: tissue typing
   b. multiplexing
   c. quantitation: allelic frequencies
1. Molecular Weight Measurements: How big?

a. Sodium adducts

The presence of multiple sodium counter ions can result in poor mass measurement accuracy, particularly when not resolved.

Purification by cation exchange, or the addition of ammonium salts, can promote protonated (M+H⁺) and deprotonated (M-H⁻) species.

b. Base losses

Mass measurement accuracy is also effected by random losses of bases, a non sequence-specific fragmentation that results in broadening at the base of the peak when it occurs during ion acceleration.

This base loss can be observed more clearly on a reflectron instrument.

1. Molecular Weight Measurements: How big? (cont)

c. UV and IR lasers

UV

MALDI-TOF mass spectra of (a) 246-bp DNA amplified from pLB129 and (b) 500-bp DBA amplified from bacteriophage lambda genome. The laser wavelength was 266 nm and the matrix HPA/3-HPA.

1. Molecular Weight Measurements: How big? (cont)

c. UV and IR lasers

IR

IR-MALDI TOF mass spectra of (a) synthetic 21-nt DNA (MW = 6398), (b) plasmid DNA restriction enzyme digest (pBluescript-KS+; Bgl I and Rsa I; 87, 112, 285 and 433 kDa), (c) plasmid DNA restriction enzyme digest (pBluescript-KS−; Not I and Ssp I; 70, 170 and 673 kDa), and (d) 1206 nt RNA transcript (387 kDa). Wavelength 2.94 µm (Er:YAG laser) with glycerol matrix.

2. **Ladder Sequencing**

a. Sanger-type sequencing

b. exonuclease ladders

c. hydrolysis ladders

Demonstration of mock sequencing: (a) MALDI mass spectra of synthetic nucleotides corresponding to the first 24 DNA fragments generated in typical sequencing reactions using a standard M13mp19 template and the primer d(GTAAAACGACGGCCAGT) and (b) overlay of the spectra shown in (a). 0.5 pmol of each component was used and the matrix was 3-HPA.

3. Hybridization probes

a. Example: tissue typing

Class II HLA (Expressed in specialized immune cells)

Antigen Presenting Cell

Class II HLA
T Cell Receptor

The HLA must be the same in both donors and recipients

Helper T Lymphocyte

HLA between donors and recipients must be matched prior to transplantation.

Allele-Specific DNA Typing by MALDI Mass Spectrometry, Timothy Worrall, Barbara Schmeckpeper, Joseph Corvera and Robert J. Cotter.
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry

The Major Histocompatibility Complex

HLA are encoded in the Major Histocompatibility Complex (MHC)

MHC  Human Chromosome 6

Class II  Class I

DP  DQ  DR  B  C  A

DRB1  DRB3,4, or 5  DRA

or

HLA Class II

Macrophage
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry

**Allelic Variation Within the DRB Gene**

```
DRB
  /\  
  |  
  1  2  3  intron  4  5  6

hypervariable codons

6-14  19  27-35  47  57-58  67-78  85-87
```

**Current Methods Used to Determine HLA Alleles**

- Serologic Methods
- Molecular (DNA-based) methods
  - Sequence Specific PCR
  - Hybridization (dot blot or reverse dot blot)
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry

DNA Isolation

Target Amplification

Biotinylated Template Strand

Variable bases are determined unambiguously
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry


ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry

Extension reactions, products and possible extension masses using the sense strand template.

MALDI-TOF Mass Spectrometry
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry

Codon 35 base 1 using the sense strand template

ddNTPs and dNTPs used in extension reaction
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry

Codon 37 bases 1-2
using the sense strand template

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
a. Example: Allele-Specific DNA Typing by MALDI Mass Spectrometry

Codon 47 base 2 using the sense strand template

| Codon 47 base 1 | 5' - AG - G C G G G G C G A G G C G G CTG 3' |
| ddATP          | dTTP                                         |
| 3' - A G C C C C C A C C T C C C T C G A C 5' | mass = 6009.0 amu |
| 3' - A T G C C C C C A C C T C C C T C G A 5' | mass = 6313.2 amu |
Other reactions and spectra using the sense strand template
Extension reactions and spectra using the anti-sense template

**Codon 27:**
- **Base 2:** 5' - GTC CA G - 3'

**Reaction 1:**
- dTTP
- dCTP
- dGTP

**Reaction 2:**
- ddTTP
- ddCTP
- ddGTP

**Reaction 3:**
- dTTP
- dCTP
- dATP

**Codon 67:**
- 5' - CAG GA GTC CT C G GCT GTT CCA GAA - 3'

**Reaction 2:**
- ddCTP
- ddTTP
- ddGTP

**Reaction 3:**
- ddTTP
- dCTP
- dATP

**Codon 67 Base 1:**
- 5' - CAG GA GTC CT C G GCT GTT CCA GAA - 3'

**Reaction 1:**
- dTTP
- dCTP
- dATP

**Reaction 3:**
- ddCTP
- ddTTP
- ddGTP

**Reaction 3:**
- ddTTP
- dCTP
- dATP

**Masses:**
- Mass = 4914.2 amu
- Mass = 5228.4 amu
- Mass = 6771.4 amu
- Mass = 8263.4 amu
- Mass = 8278.4 amu
- Mass = 6771.4 amu
- Mass = 7051.6 amu
- Mass = 7075.6 amu
- Mass = 7355.8 amu
- Mass = 7364.8 amu
- Mass = 7355.8 amu
- Mass = 7364.8 amu
Table 1. Comparison of observed PROBE product masses with calculated PROBE product masses for a series of unknown samples having a DR2 allele.

<table>
<thead>
<tr>
<th>Allele</th>
<th>c27b2</th>
<th>c30b1</th>
<th>c37b1-2</th>
<th>c47b2</th>
<th>c50b2</th>
<th>c67b1 rRxn1</th>
<th>c67b1 rxn2</th>
<th>c67b1 rxn3</th>
<th>c72b1-3</th>
<th>c74b1-2</th>
<th>c86b2-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRB1*15011</td>
<td>4913.2</td>
<td>5497.6</td>
<td>6456.2</td>
<td>6008.2</td>
<td>6601.4</td>
<td>7363.8</td>
<td>6770.4</td>
<td>7074.6</td>
<td>5789.8</td>
<td>5195.4</td>
<td>4801.2</td>
</tr>
<tr>
<td>DRB1*15012</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Unknown 2</td>
<td>4917.4</td>
<td>5507.7</td>
<td>6457.5</td>
<td>6016.3</td>
<td>6603.7</td>
<td>7366.8</td>
<td>6770.6</td>
<td>7066.9</td>
<td>5789.1</td>
<td>5203.4</td>
<td>4805.4</td>
</tr>
<tr>
<td>Unknown 7</td>
<td>4911.7</td>
<td>5506.6</td>
<td>6452.3</td>
<td>6010.1</td>
<td>6598.0</td>
<td>7369.2</td>
<td>6775.3</td>
<td>7078.1</td>
<td>5792.8</td>
<td>5202.5</td>
<td>4803.7</td>
</tr>
<tr>
<td>DRB1*15021</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*15022</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Unknown 3</td>
<td>4910.7</td>
<td>5508.9</td>
<td>6456.0</td>
<td>6018.5</td>
<td>6614.5</td>
<td>7365.8</td>
<td>6772.2</td>
<td>7068.5</td>
<td>5788.3</td>
<td>5208.5</td>
<td>5116.1</td>
</tr>
<tr>
<td>DRB1*1503</td>
<td>------</td>
<td>5817.8</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Unknown 1</td>
<td>4915.4</td>
<td>5804.7</td>
<td>6452.2</td>
<td>6017.0</td>
<td>6609.8</td>
<td>7360.3</td>
<td>6770.6</td>
<td>7081.2</td>
<td>5794.7</td>
<td>5194.0</td>
<td>4811.3</td>
</tr>
<tr>
<td>DRB1*1504</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1505</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1506</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*16011</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Unknown 4</td>
<td>4911.7</td>
<td>5502.2</td>
<td>6459.8</td>
<td>6301.5</td>
<td>6600.1</td>
<td>7356.1</td>
<td>8275.2</td>
<td>6771.2</td>
<td>6118.2</td>
<td>5197.0</td>
<td>5118.2</td>
</tr>
<tr>
<td>DRB1*16012</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*16021</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>Unknown 5</td>
<td>4911.1</td>
<td>5502.6</td>
<td>6472.8</td>
<td>6300.0</td>
<td>6601.9</td>
<td>6752.8</td>
<td>8258.5</td>
<td>7046.6</td>
<td>6125.1</td>
<td>5203.3</td>
<td>5118.9</td>
</tr>
<tr>
<td>Unknown 6</td>
<td>4913.1</td>
<td>5506.0</td>
<td>6457.5</td>
<td>6298.5</td>
<td>6593.4</td>
<td>6753.6</td>
<td>8256.8</td>
<td>7062.4</td>
<td>6119.2</td>
<td>5201.0</td>
<td>5109.4</td>
</tr>
<tr>
<td>DRB1*16022</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1603</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1604</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1605</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1606</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1607</td>
<td>5227.4</td>
<td>------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>DRB1*1608</td>
<td>------</td>
<td>6118.0</td>
<td>6132.4</td>
<td>7354.8</td>
<td>8277.4</td>
<td>6761.4</td>
<td>6119.0</td>
<td>6119.0</td>
<td>5114.2</td>
<td>5114.2</td>
<td>5114.2</td>
</tr>
</tbody>
</table>

The DRB1*15011 allele is used as a reference, with only those masses that differ shown for the other alleles. Unknowns and their masses are listed below their corresponding allele.
3. Hybridization probes

b. Multiplexing

*Time-of-flight (TOF) mass spectrometry* makes it possible to multiplex DNA analyses by recording broad mass ranges simultaneously.

In this case, the extension products from five primers are recorded.

Broad mass range recording is equally important for *proteomics*.

SEQUENOM, La Jolla CA
### 3. Hybridization probes

#### b. Multiplexing

<table>
<thead>
<tr>
<th>Protein</th>
<th>Probe Sequence</th>
<th>Mass (a.m.u.)</th>
<th>Expected Allele</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Apolipoprotein A IV</strong></td>
<td>5’-AGCCAGGACAAAG-3’ (A347)</td>
<td>3688.40</td>
<td>primer</td>
</tr>
<tr>
<td></td>
<td>5’-AGCCAGGACAAAGTC-3’</td>
<td>4265.80</td>
<td>347Ser</td>
</tr>
<tr>
<td></td>
<td>5’-AGCCAGGACAAAGA-3’</td>
<td>3985.60</td>
<td>347Thr</td>
</tr>
<tr>
<td></td>
<td>5’-ACAGCAGGAACAGCA-3’ (A360)</td>
<td>4604.00</td>
<td>primer</td>
</tr>
<tr>
<td></td>
<td>5’-ACAGCAGGAACAGCATC-3’</td>
<td>5181.40</td>
<td>360His</td>
</tr>
<tr>
<td></td>
<td>5’-ACAGCAGGAACAGCAG-3’</td>
<td>4917.20</td>
<td>360Gln</td>
</tr>
<tr>
<td><strong>Apolipoprotein E</strong></td>
<td>5’-GCGGACATGGAGGACGTG-3’ (E112)</td>
<td>5629.60</td>
<td>primer</td>
</tr>
<tr>
<td></td>
<td>5’-GCGGACATGGAGGACGTG-3’</td>
<td>6247.00</td>
<td>112Cys</td>
</tr>
<tr>
<td></td>
<td>5’-GCGGACATGGAGGACGTGC-3’</td>
<td>5902.80</td>
<td>112Arg</td>
</tr>
<tr>
<td></td>
<td>5’-GATGCCGATGACCTGCAGAAG-3’ (E158)</td>
<td>6480.20</td>
<td>primer</td>
</tr>
<tr>
<td></td>
<td>5’-GATGCCGATGACCTGCAGAAGC-3’</td>
<td>6753.40</td>
<td>158Arg</td>
</tr>
<tr>
<td></td>
<td>5’-GATGCCGATGACCTGCAGAAGTG-3’</td>
<td>7097.60</td>
<td>158Cys</td>
</tr>
<tr>
<td><strong>Apolipoprotein B-100</strong></td>
<td>5’-GTGCCCTGCGCTCTTTAAGACC-3’ (B3500)</td>
<td>7313.80</td>
<td>primer</td>
</tr>
<tr>
<td></td>
<td>5’-GTGCCCTGCGCTCTTTAAGACTG-3’</td>
<td>7931.20</td>
<td>3500Gln</td>
</tr>
<tr>
<td></td>
<td>5’-GTGCCCTGCGCTCTTTAAGACC-3’</td>
<td>7587.00</td>
<td>3500Arg</td>
</tr>
</tbody>
</table>

---

ME 330.804: Role of Chromatography and Mass Spectrometry in Biomedical Research, Robert J. Cotter and Amina S. Woods
3. Hybridization probes

c. Quantitation: allelic frequencies


![Diagram](image)

Figure 2. Mass spectra obtained at representative intervals from a quantitation study of CP450 2D6 SNP site in pooled PCR products. Oligonucleotide sequences used were: PCR F: 5'-CCAAGGGAGCCCTGAGACG-3', PCR R: 5'-GTGTCAGAGGTCAGTCACCAC-3', genotyping primer: 5'-AATGATGAGAACCTG-3'. Mass spectra were generated from homozygous wild-type (A), heterozygous mutant (B) and mixtures of PCR from wild-type individuals and heterozygous mutant individuals in wild-type:mutant ratios ranging from 5:1–50:1 (C–P).
3. Hybridization probes

c. Quantitation: allelic frequencies


Figure 3. Quantitation curve for analysis of CP450 2D6 SNP site in pooled PCR products. Peak area ratios (T allele:C allele) were measured from five replicate mass spectra and averaged, including spectra shown in Figure 2. Vertical bars show standard deviation at each concentration point. Area ratios were corrected for the relative response between T and C alleles as determined from analysis of replicate heterozygous mutant samples.