Structural Analysis

Chem 314

Prof. Oleg Boyarkine

Prof. Lyndon Emsley

Prof. Christoph Bostedt

Email: oleg.boiarkin@epfl.ch

Always write in the subject line: Chem314

Organizational details

Classes:

- Mondays, 13h15-16h, Rm. CH B3 30

Exercises:

- Mondays, 16h15-17h, Rm. CH B3 30

Weeks 1, 3-5:

- Professor O. Boyarkine - Mass Spectrometry

Weeks 6-10:

- Professor L. Emsley - NMR

Weeks 11-14:

- Professor C. Bostedt - X-ray Crystallography

Moodle

http://moodle.epfl.ch/

Course category: Chimie, Génie chimique CGC

Sub-category: Bachelor

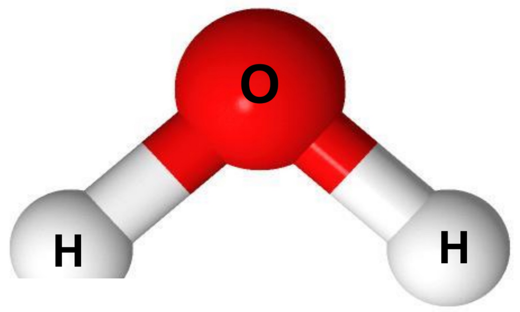
Course: CH-314

Password: structure314

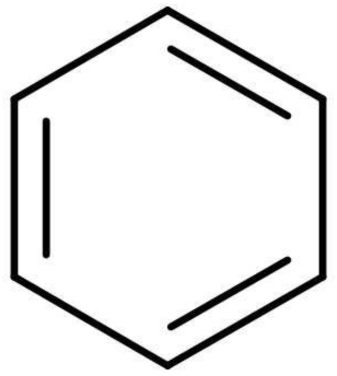
- All slides as pdf
- Exercises and solutions
- Important announcements
- Zoom link: https://epfl.zoom.us/skype/2070226542

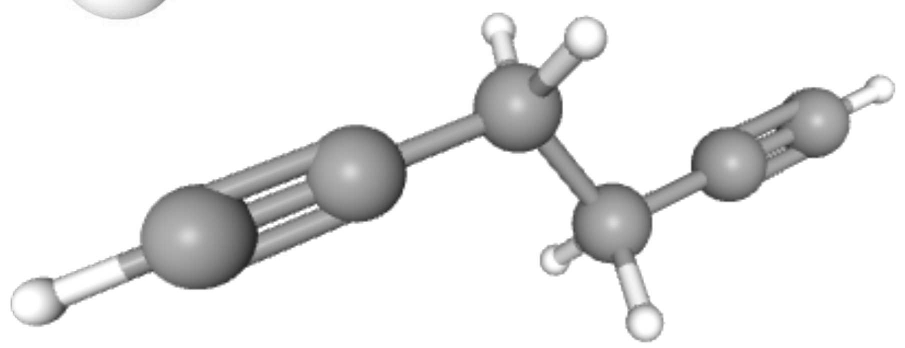
What do we mean by structure?

I. Chemical composition: Number of atoms of each sort in a molecule:



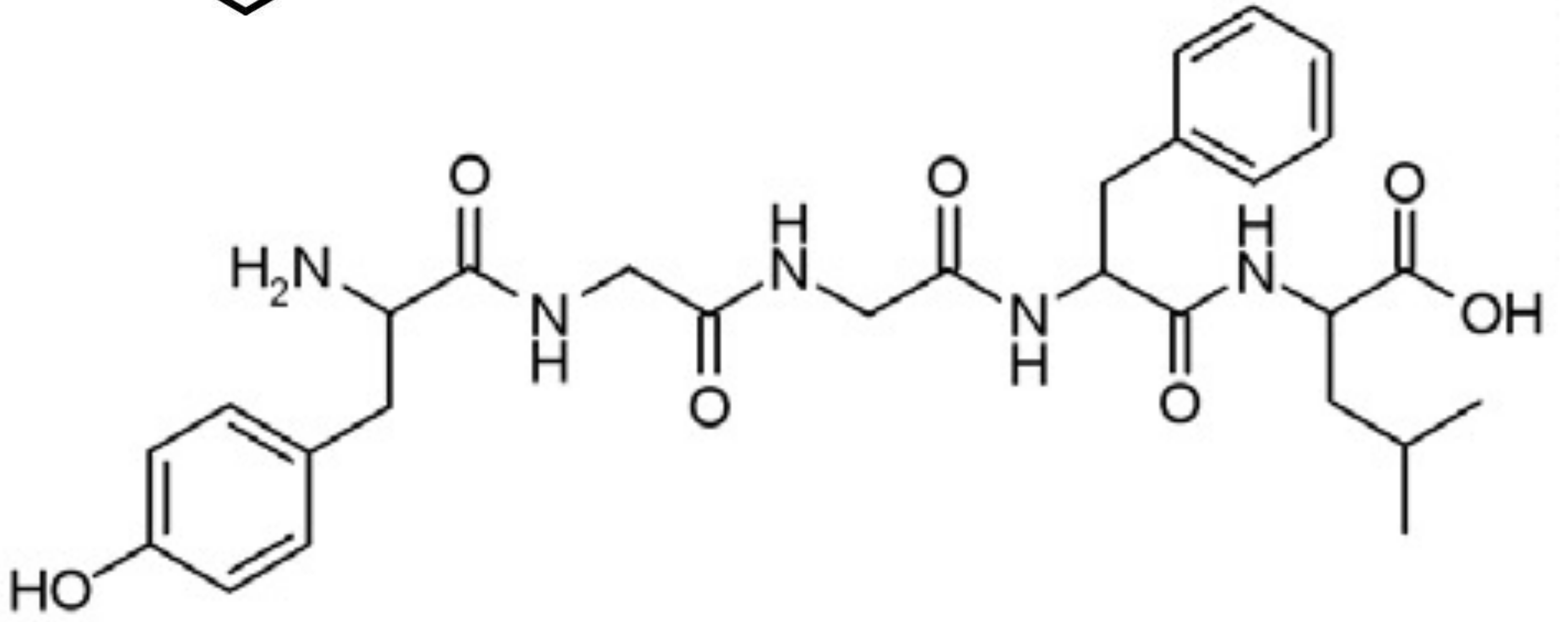
H_2O



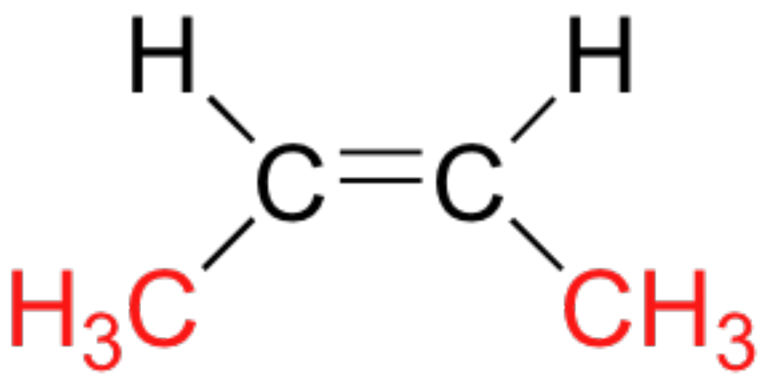


C_6H_6

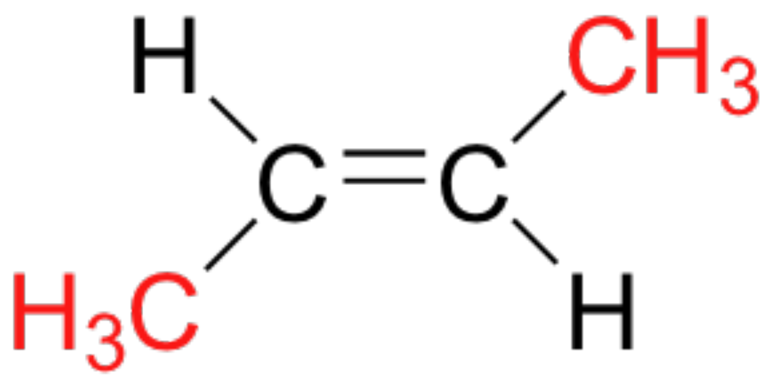
$C_{28}H_{37}N_5O_7$???



II. Arrangement: Linking of atoms (groups) by chemical and non-covalent bonds.

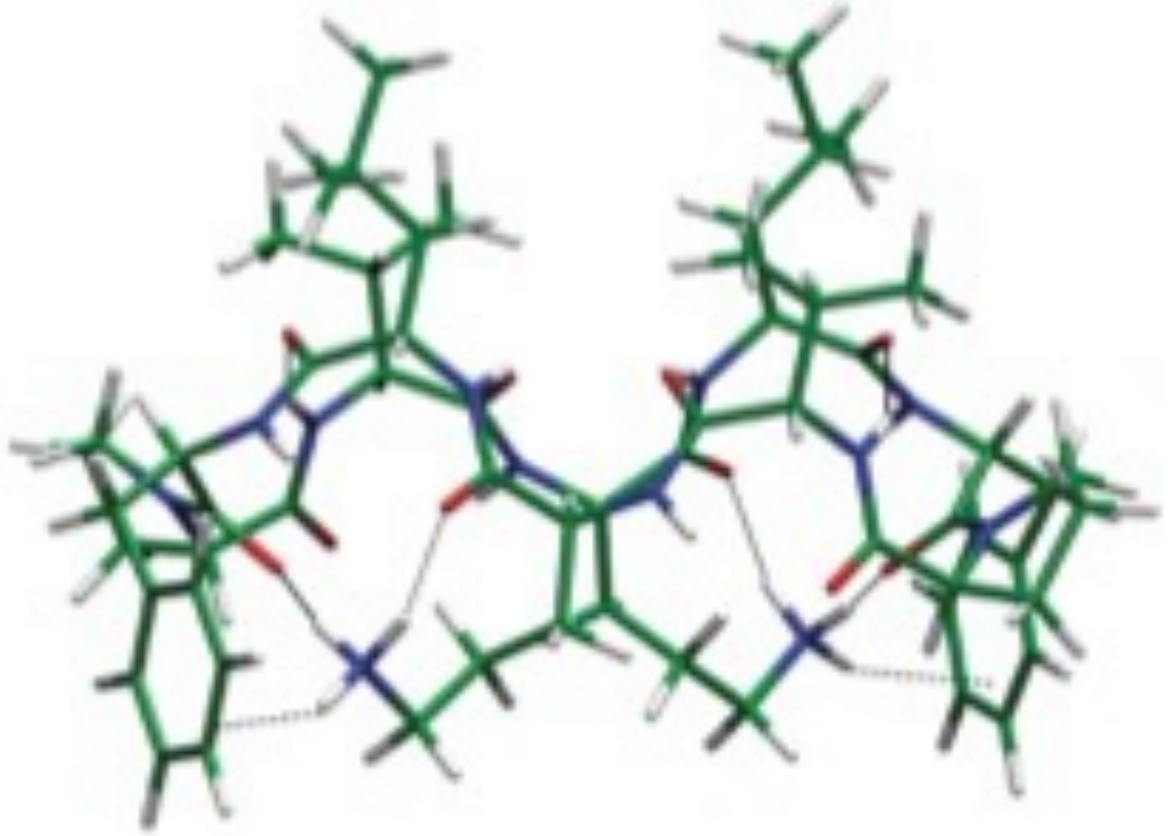


cis-2-butene



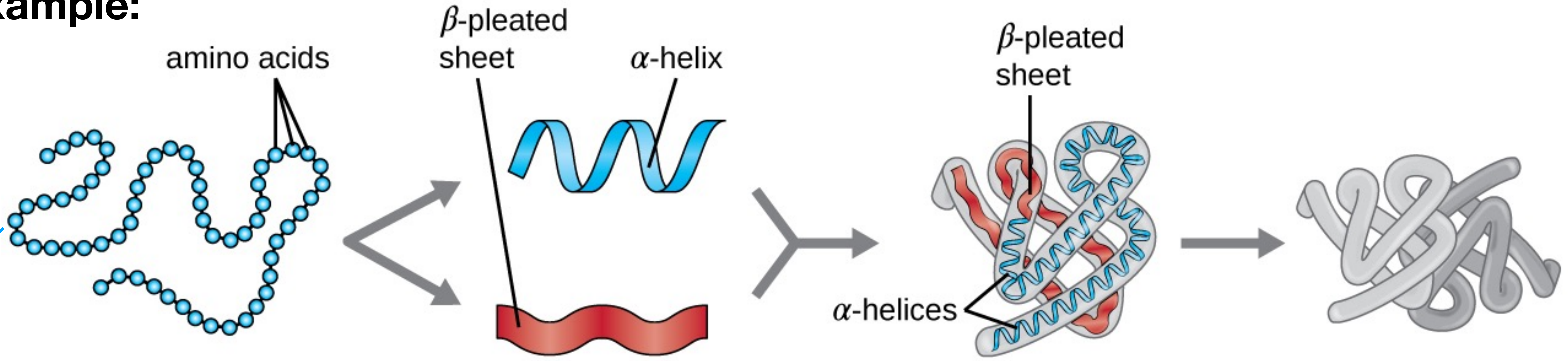
trans-2-butene

III. Geometry: relative positions of atoms in space.



What do we mean by structure?

For example:

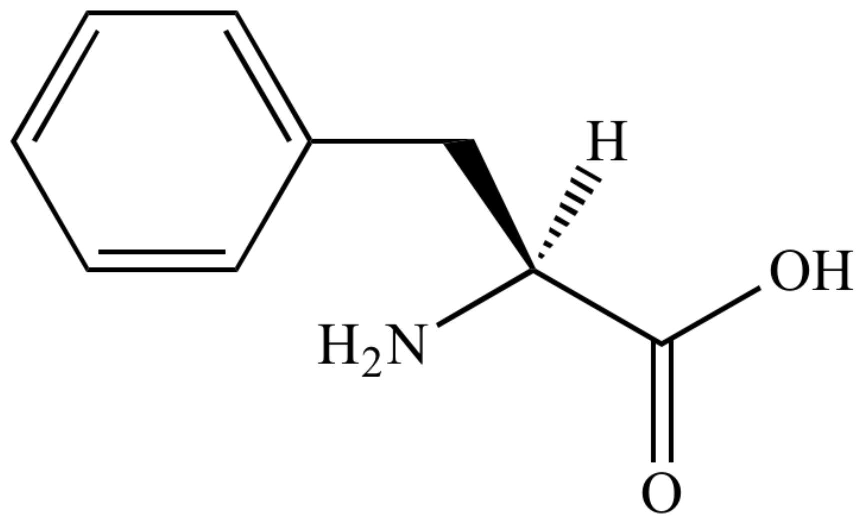


Primary Protein Structure
Sequence of a chain of amino acids

Secondary Protein Structure
Local folding of the polypeptide chain into helices or sheets

Tertiary Protein Structure
three-dimensional folding pattern of a protein due to side chain interactions

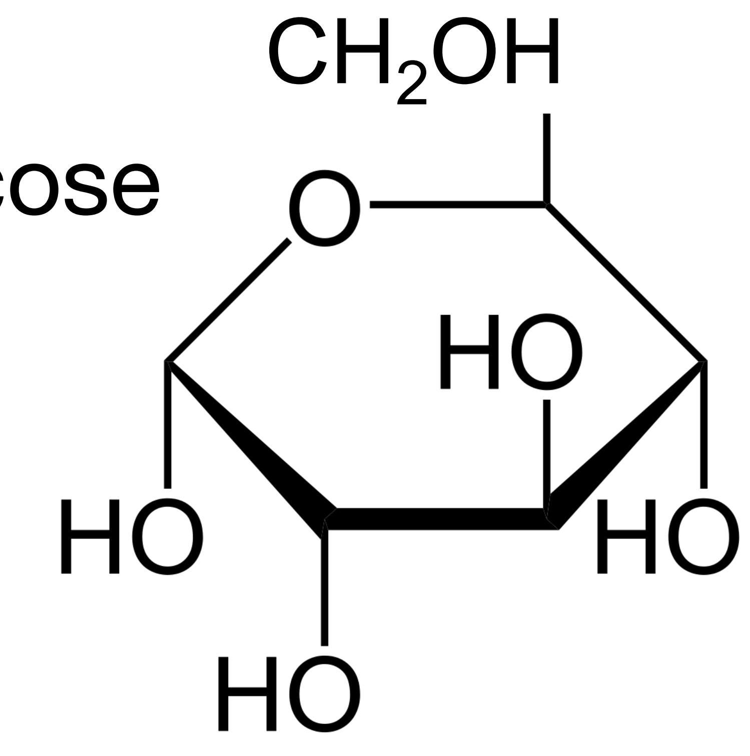
Quaternary Protein Structure
protein consisting of more than one amino acid chain

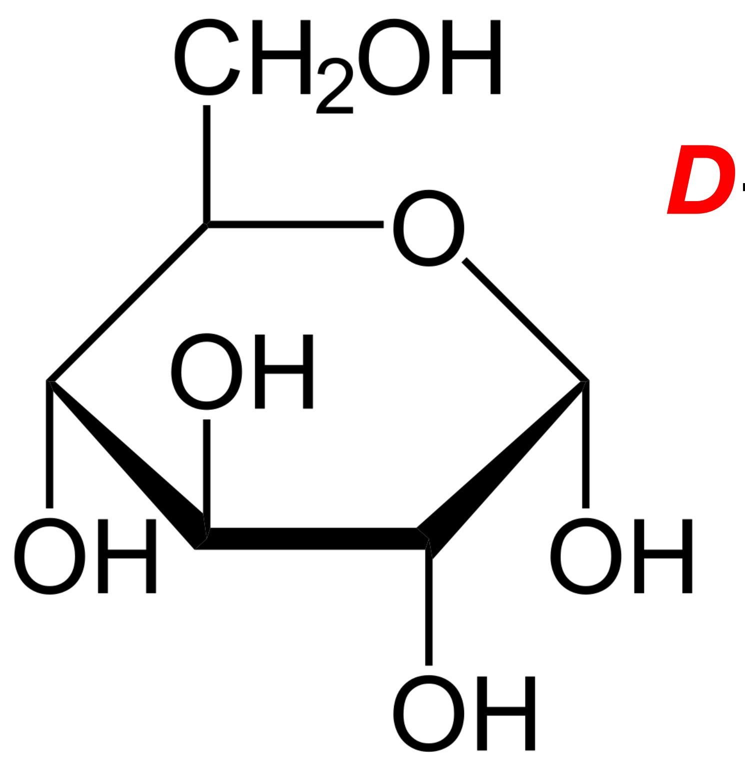


Why do we care about structure?

Structure determines function

L-glucose



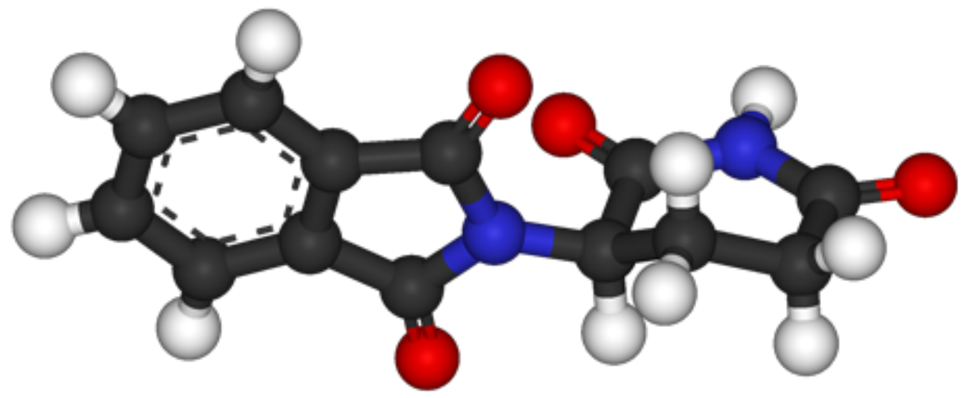


D-glucose

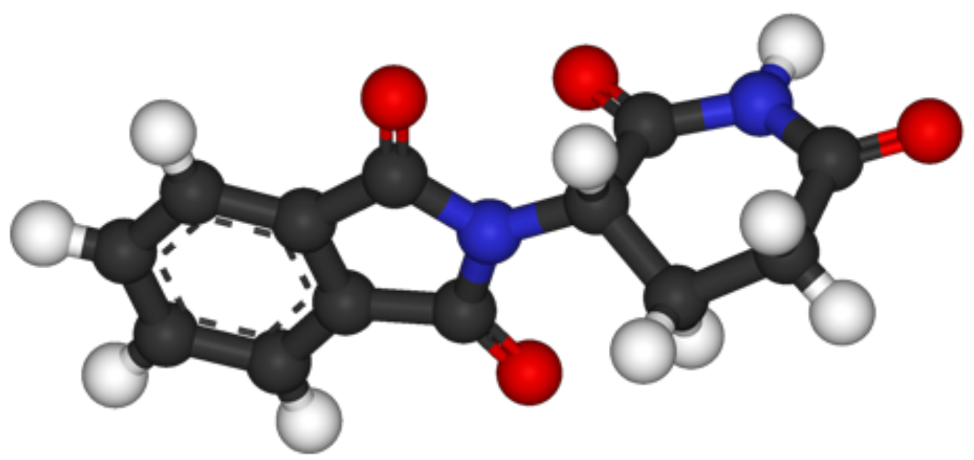
Both are sweet, but only ***D*** (natural) gives energy;

L cannot be metabolized

<u>Pain relief drug for children</u>



(R)-Thalidomide:
Has desired pharmacological effect.



(S)-Thalidomide:
Has embryo-toxic and teratogenic effects.

A molecule's structure is not the *only* factor that determines its function:
Flexibility or ability to change its structure (conformers).

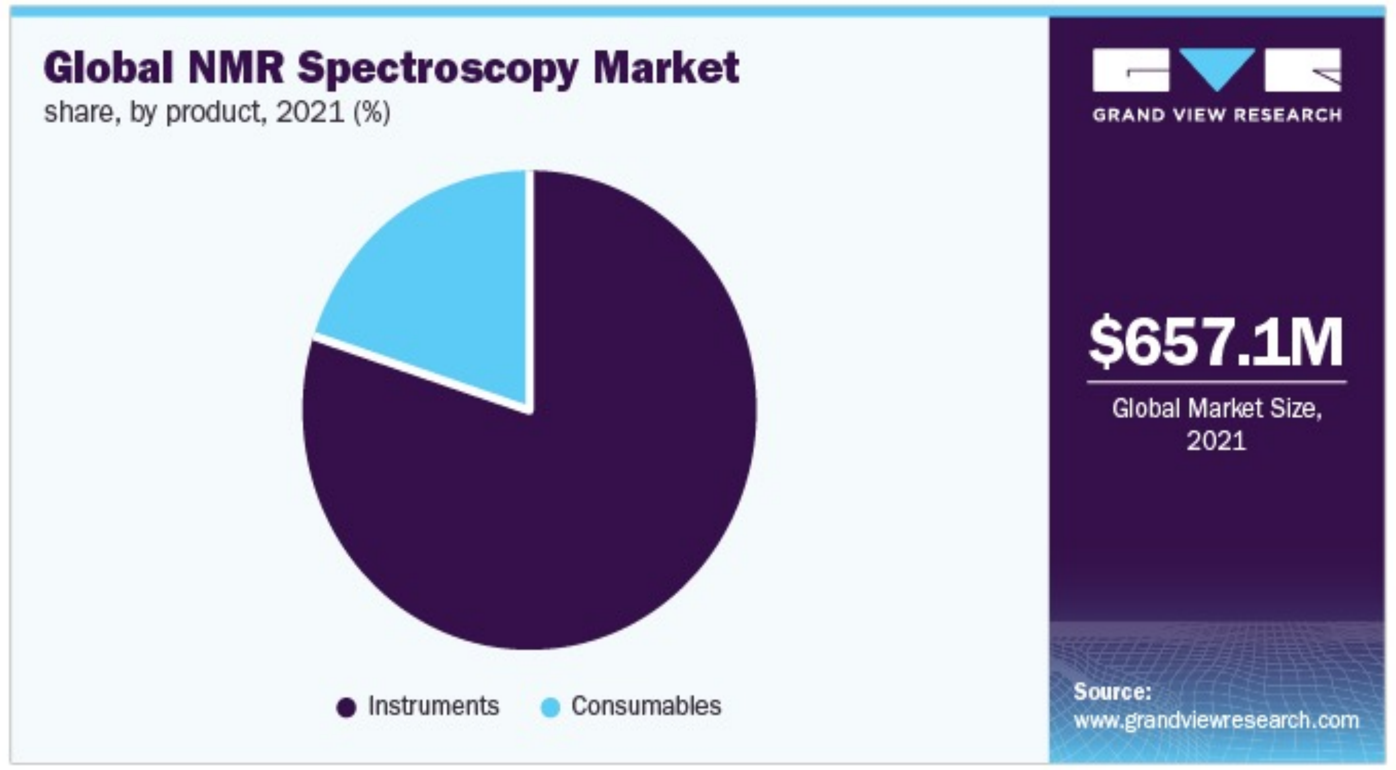
This is a course on structural analysis techniques

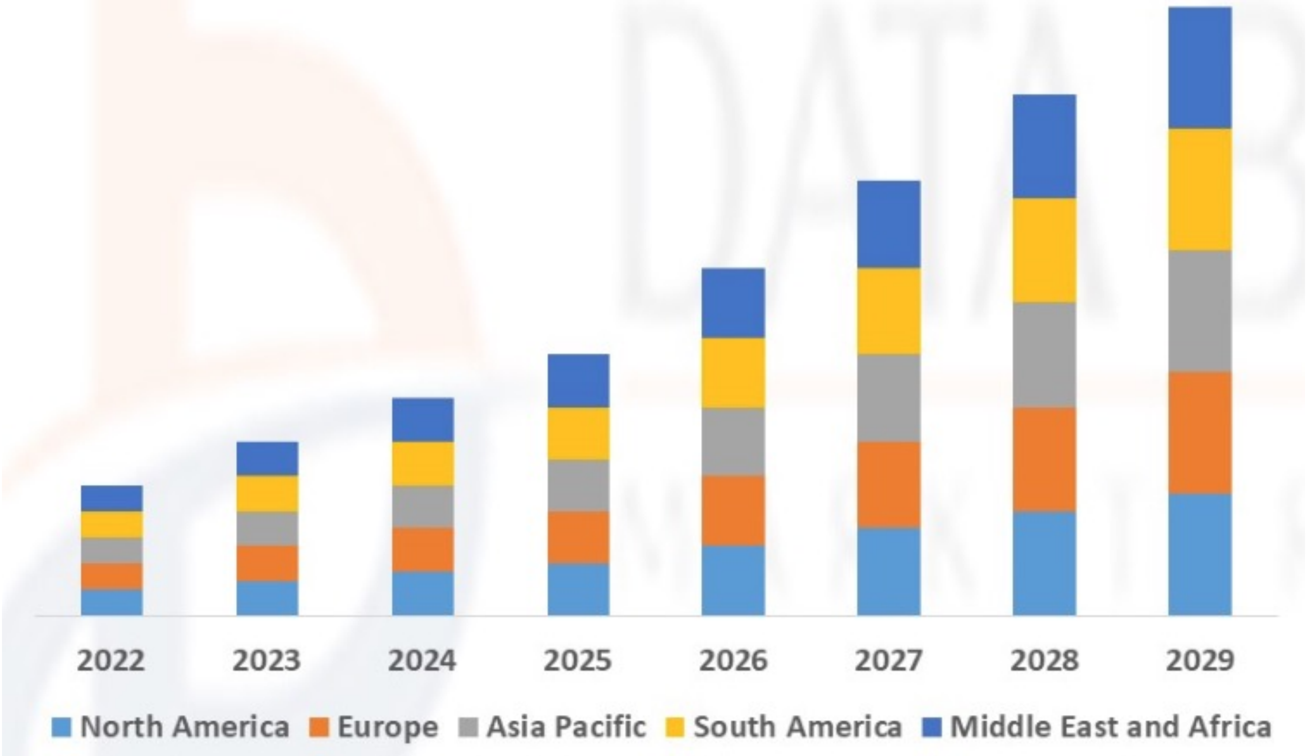
- Mass spectrometry
- NMR
- X-ray crystallography

The physical chemistry (and physics) of how these techniques work.

Why should I study this course ???

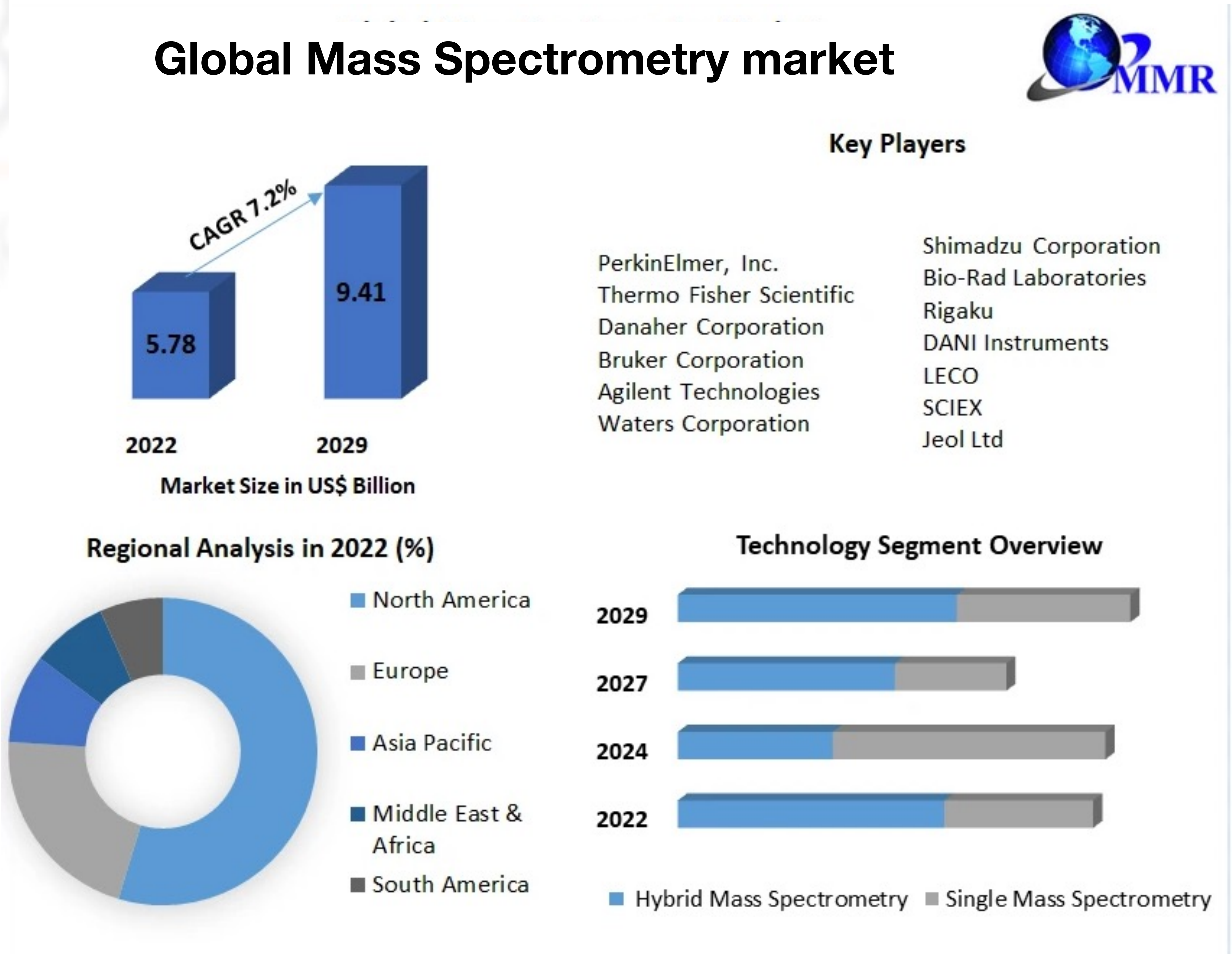
Worldwide market for NMR, x-ray crystallography and mass spectrometry





NMR = $ 0.66 Bln

Cryo-EM = $ 0.73 Bln



X-ray crystallography = $ 1.6 Bln

Mass spectrometry = $ 5.8 Bln

These techniques, MS in particularly, are not only for research and labs, but largely for Industries

Structure depends upon environment

Gas phase

- Isolated. No solvent. OK for small rigid molecules. Unnatural environment for biological molecules.
- Best suited tool: Mass spectrometry
  - For biological molecules: primary structure, quaternary structure, elements of secondary and tertiary structure.

Solution phase

- Closer to natural environment for biological molecules
- Best suited tool: NMR
  - Precise 3-dimensional structure.

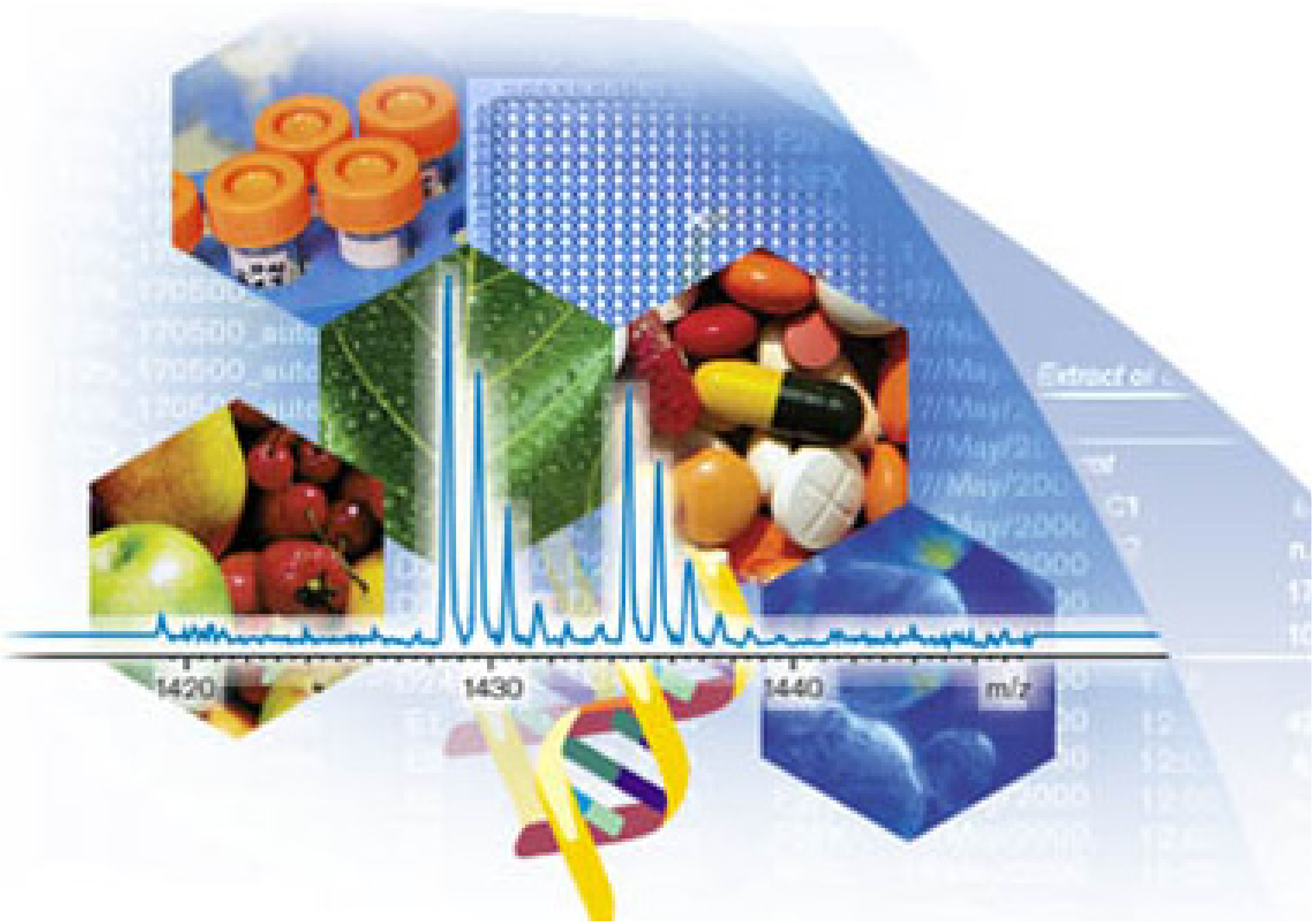
Solid phase

- Unnatural for biological systems
- Best suited tool: NMR and X-ray crystallography

In vivo

- Natural biological environment. Crowding effects influence structure.
- Best suited tool: NMR (for imaging rather than structure)

Mass Spectrometry



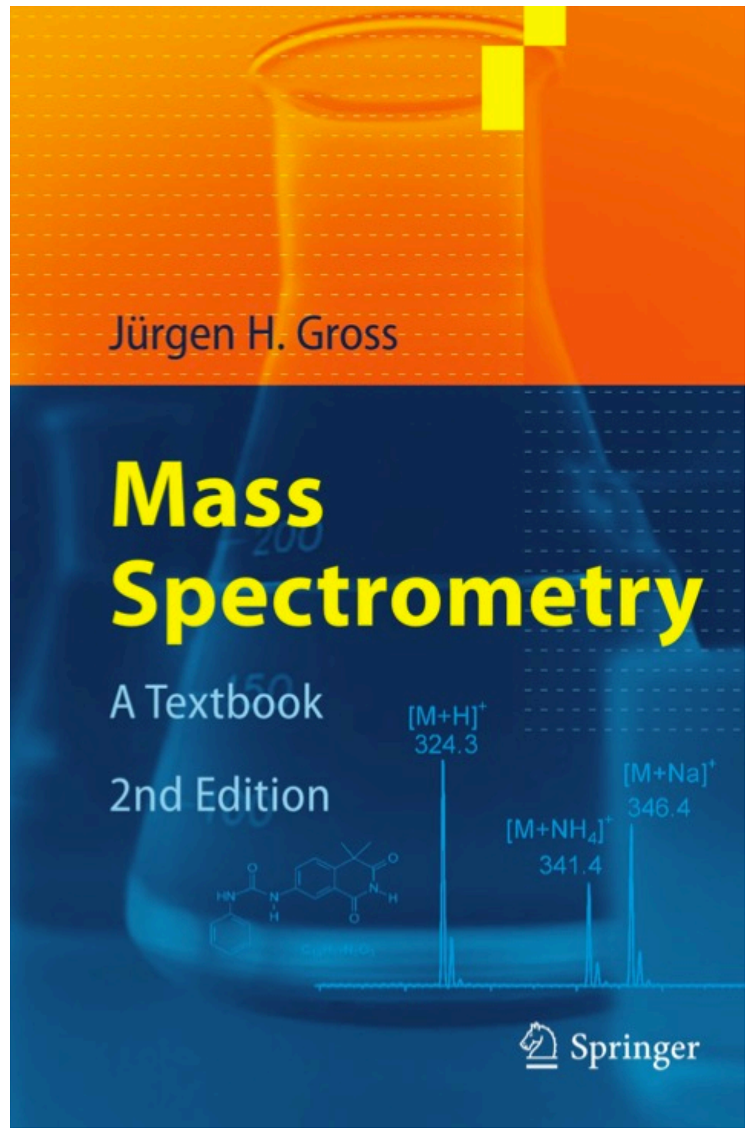
Course outline, weeks 1, 3-5

1. Introduction to mass spectrometry
2. Masses of elements and molecules
3. Isotopes and isotope distributions
4. Mass accuracy and resolution
5. Mass spectrometry instrumentation
    (a) Ion sources
    (b) Mass analyzers
    (c) Detectors
6. Tandem MS

Textbooks

Mass-spectrometry. A textbook. 2^{nd} edition, by Jürgen H. Gross (available as e-book through EPFL library,

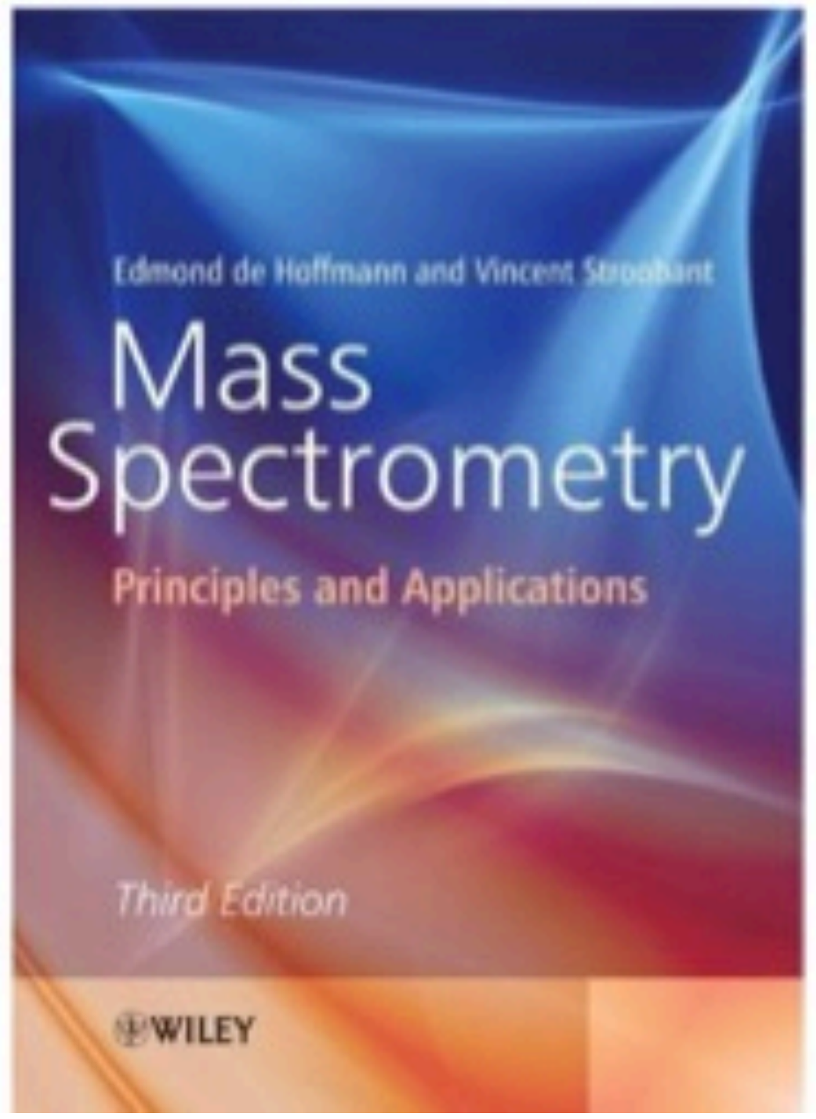
http://link.springer.com/book/10.1007/978-3-642-10711-5/page/1)



Literature

In French:
Spectrométrie de masse,
3^{e} edition, par Edmond de Hoffmann et Vincent Stroobant

In English:
Mass spectrometry: principles and applications,
3rd edition, by Edmond de Hoffmann and Vincent Stroobant

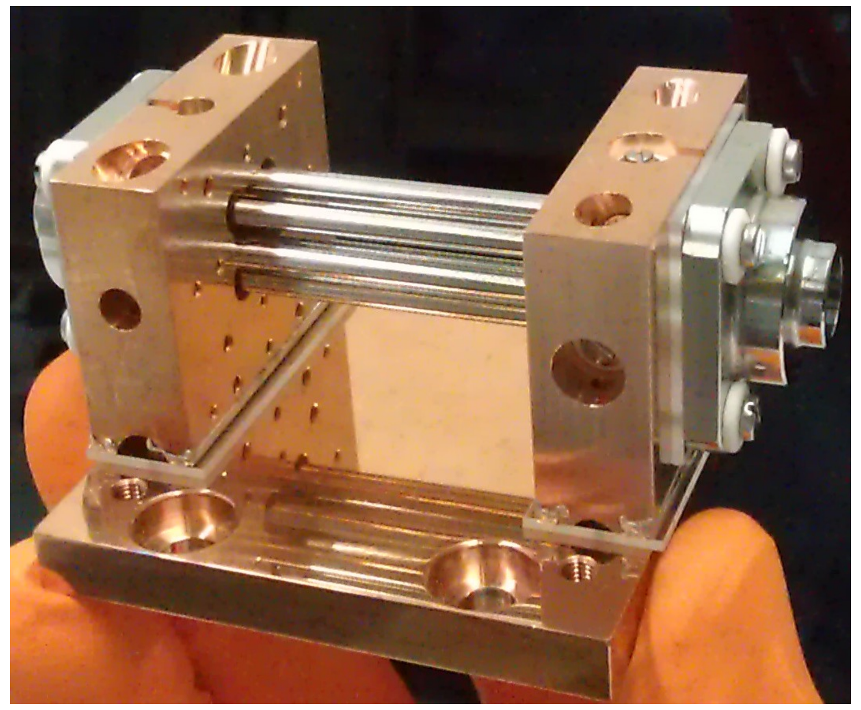


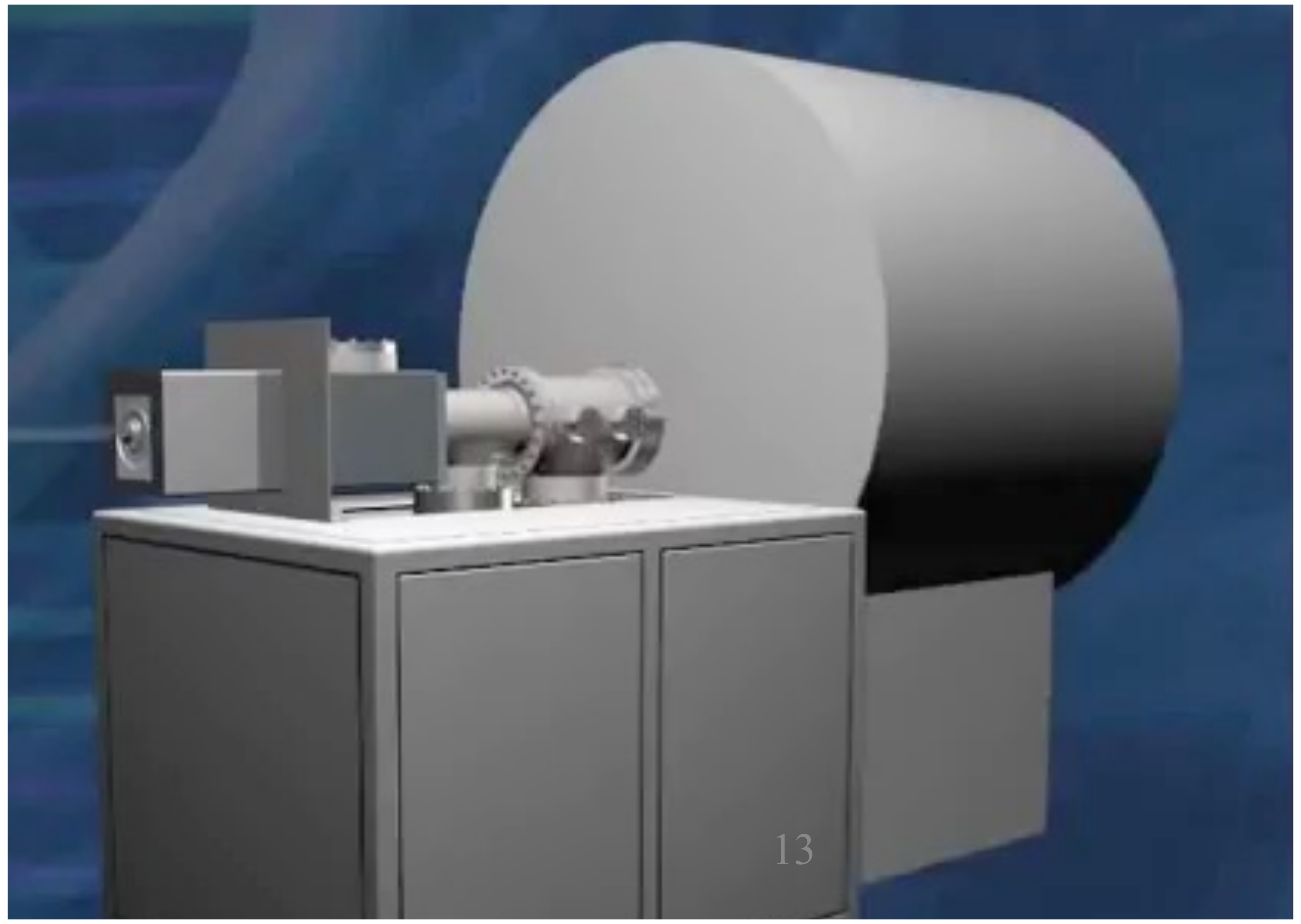
What is Mass Spectrometry?

The study of matter based on the mass of charged molecules and their fragments

Mass Spectrometer

is an instrument that measures the *mass-to-charge ratio* of individual molecules that have been converted to ions.

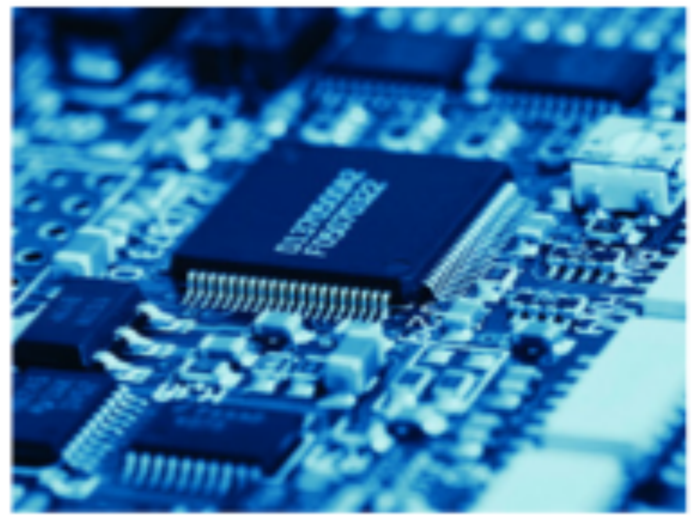




Where is it used?

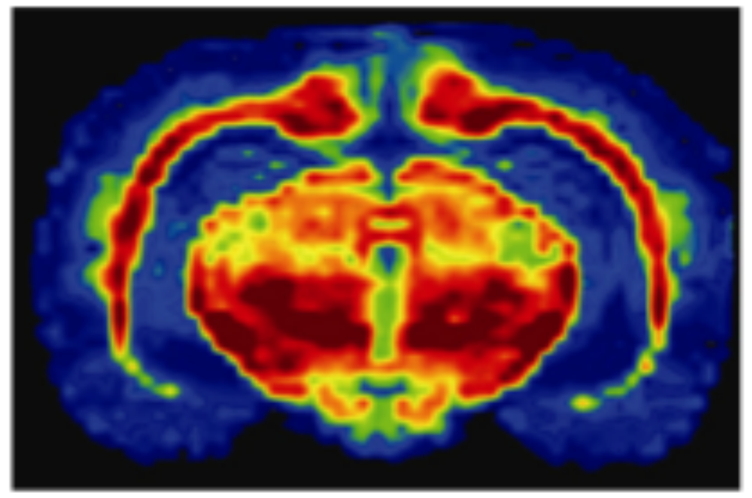
clinical

environmental



materials

pharma

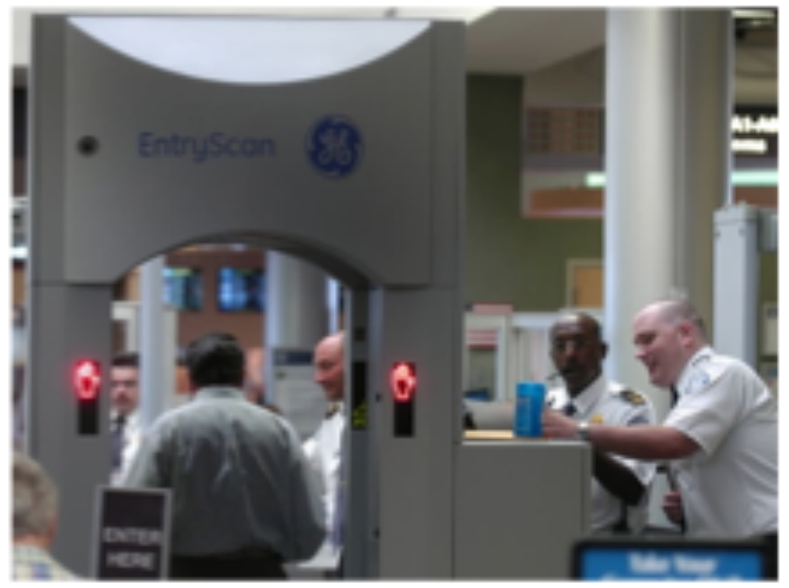


life sciences

energy

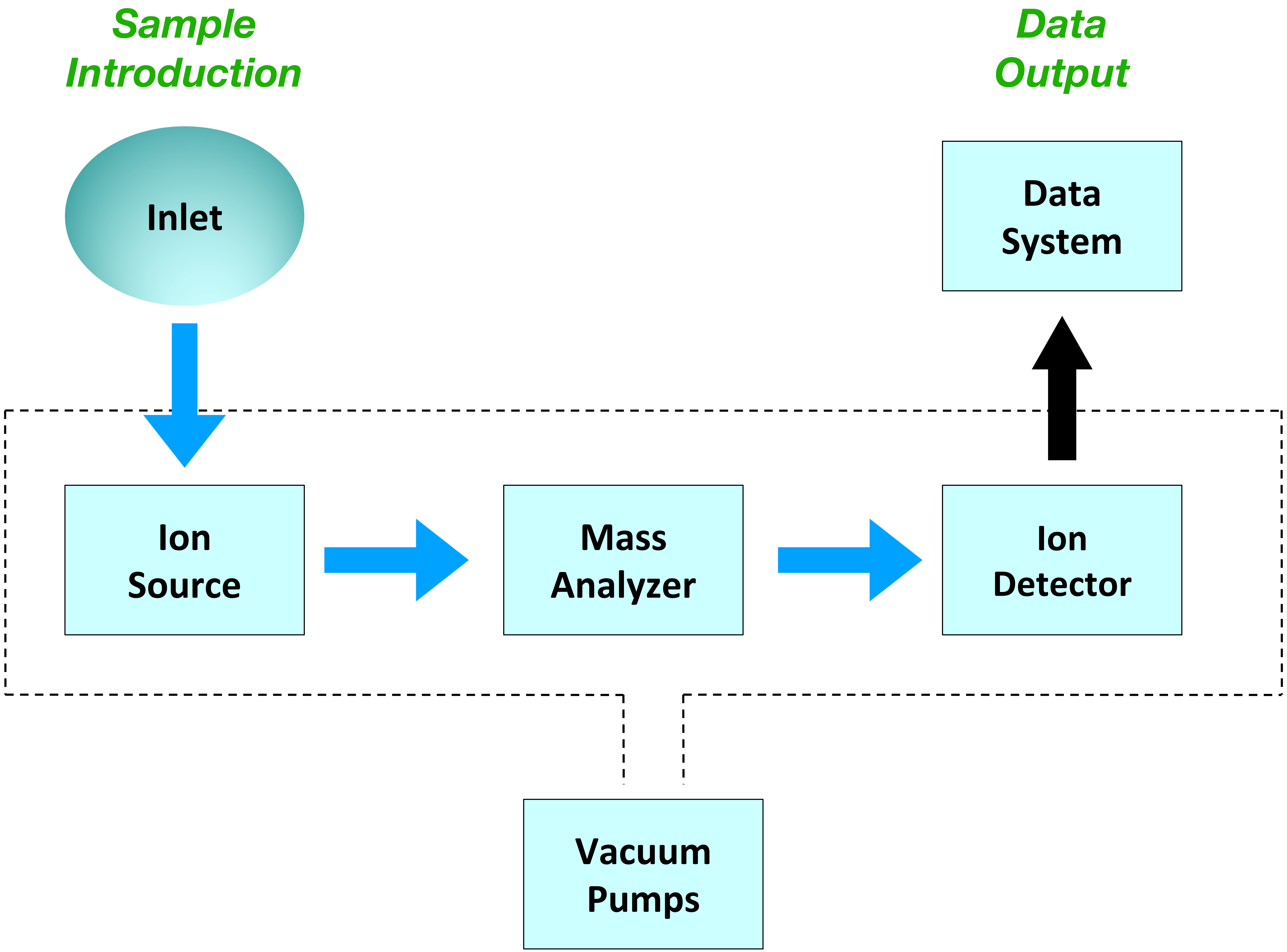
forensic

fundamentals



security

How does a mass spectrometer work?



What does a mass spectrometer do?

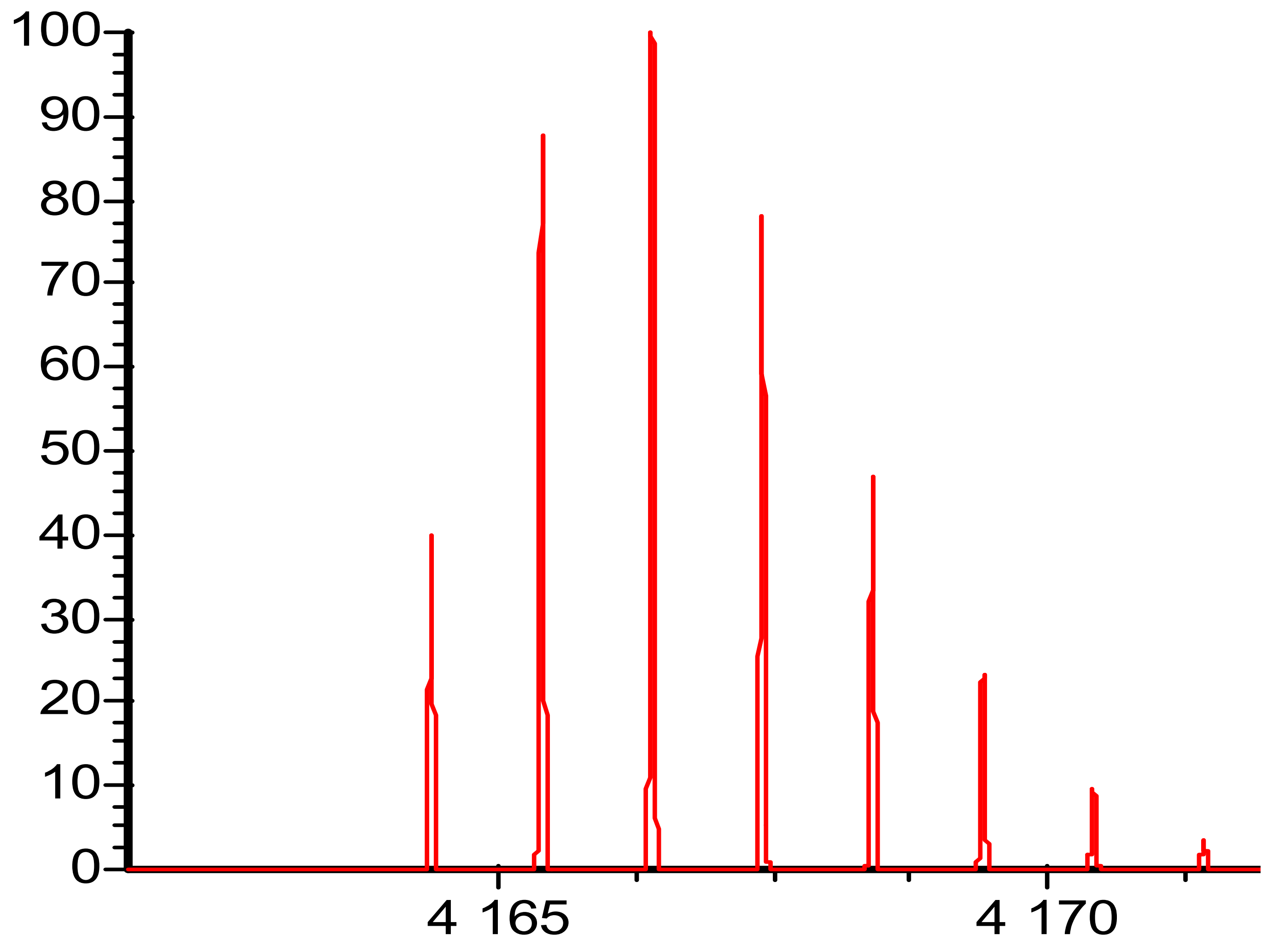
- Generates ions (positive or negative) from samples introduced directly or from a GC or HPLC

- Measures number of ions art each m/z within a desired range (at single m/z for different conditions of generating ions).

- A data system processes the recorded data (normalization, background subtraction, mass range displayed, etc.)

- The mass spectrum is a plot of the normalized,

  relative abundance against m/z.



Why the charge?

Ion in electric and magnetic fields experience forces:

$$\vec{F}_{el} = q \cdot z \cdot \vec{E}; \quad \vec{F}_{m} = q \cdot z \cdot \vec{\mathrm{v}} \times \vec{B}$$

Because $\vec{F} = m \cdot \vec{a}$: $\quad \vec{a}_e = \frac{q \cdot z}{m} \cdot \vec{E}$ and $\vec{a}_e = \frac{q \cdot z}{m} \cdot \vec{\mathrm{v}} \times \vec{B}$

- Ions can be deflected, accelerated or focused, analysis of the motion can be used to analyze their mass-to-charge ratio, m/z.

- Neutral molecules at room temperature are almost unaffected by electromagnetic field.

Electric force is usually much stronger than magnetic and gravitational

Advantages of Mass Spectrometry

NEAR UNIVERSAL APPLICABILITY

- Almost all substances give a mass spectrum. Wide range of molecular weights.

SELECTIVITY

- High **resolving power** allows selection of one component from a complex mixture

SPECIFICITY

- Exact monoisotopic mass often identifies chemical composition and limits the number of possible candidate compounds; observation of a particular fragmentation pattern in MS often identifies a given component.

SENSITIVITY

- Detection levels as low as 1 femtomole (10^{-15} moles). Complete "structure" from less than 100 femtomoles

SPEED

- From a few μs to tens of seconds

Limitations of Mass Spectrometry

- MS-only is fundamentally insensitive to isomers; fragmentation MS may sometimes distinguish them from each other.
- It can distinguish isobaric compounds (e.g. CO and N_2) only if the resolving power is sufficiently high.
- Compounds may decompose or isomerize during the ionization process. Secondary or higher structures observed in solution may be lost during the ionization process.
- Detailed mechanism of ionization and fragmentation processes are not fully understood. It is often difficult to predict fragmentation MS of a particular molecule from first principles.
- Getting detailed structural information from the spectra requires a solid understanding of fragmentation mechanisms.
- Quantitation requires use of an internal standard such as a deuterated analog.

Masses of atoms and molecules

Units of mass and mass/charge ratio

- SI: kilogram (mg, g,…)
- In physics:

  atomic mass unit (a.m.u.) =1/12 of ^{12}C

1 a.m.u. = $1.6605655 \cdot 10^{-27}$ kg.

- In mass spectrometry:

Dalton (1 Da = 1 amu)

- Mass to charge ratio (m/z)

Thomson (1 Th = 1 Da/atomic charge)

- Electron charge: $e = -1.602 \times 10^{-19}$ coulomb
- Charge of an ion: $q = z \cdot e$ (z, number of charges + or –)

Electron and proton mass

- Mass of electron = 0.0005489 Da
  - That's 9.11 x 10^{-31} kg
  - Small but measurable by HRMS

- Gain/lose of proton H^+
  - H atom mass: 1.007824
  - H^+ mass: 1.007276

- In high-resolution mass spectrometry, the electron mass must be considered!

Masses of chemical elements

An atom consist of a nuclear and electrons

- Z - atomic number, number of protons
- N – number of neutrons
- A=Z+N – atomic mass number

$$m_A \cong (Z \cdot m_P + N \cdot m_N) + Z \cdot m_e$$

- Nuclei of the same chemical element (Z) may have different masses due to different number of neutrons (**N**):

  These atoms are called ***ISOTOPES*** of the chemical element with **Z**.

  Relative concentration of an isotope of the same element is called ***Isotopic abundance.***

- Chemical element: ID=Z (usually symbol of chemical element: Z=29 = Cu)

- Atom - any one of the isotopes of the chemical element: ID= ^{A}Z (^{63}Cu)

*Masses of atoms cannot be accurately calculated from **m_p, m_N** and **m_e** (because of binding energy) and must be measured*

$$m_A = (Z \cdot m_P + N \cdot m_N) + Z \cdot m_e - \frac{\varepsilon_b}{c^2}$$

Isotopes in nature

Oxygen

O-13	0.0089 s
O-14	70.599 s
O-15	122.24 s
O-16 stable,	**99.762%**
O-17 stable,	**0.038%**
O-18 stable,	**0.200%**
O-19	26.91 s
O-20	13.57 s
O-21	3.4 s

^{13}C natural abundance:

on Earth: ^{13}C - 1.07%, ^{12}C - 98.9% (on some stars ^{13}C - 22%).

Natural isotopic abundance of some elements

Isotope	% nat. abundance	atomic mass
^{1}H	99.985	1.007825
^{2}H	0.015	2.014101
^{12}C	98.89	12 (definition)
^{13}C	1.11	13.00335
^{14}N	99.64	14.00307
^{15}N	0.36	15.00011
^{16}O	99.76	15.99491
^{17}O	0.04	16.99913
^{18}O	0.2	17.99916
^{19}F	100	18.99840
^{23}Na	100	22.98976

Note that these are stable isotopes. There are many other radioactive isotopes.

Isotope	% nat. abundance	atomic mass
^{28}Si	92.232	27.97693
^{29}Si	4.699	28.97649
^{30}Si	3.092	29.97377
^{31}P	100	30.97376
^{32}S	95.0	31.97207
^{33}S	0.76	32.97146
^{34}S	4.22	33.96786
^{36}S	0.014	35.96709
^{35}Cl	75.77	34.96885
^{37}Cl	24.23	36.96590
^{79}Br	50.69	78.91838
^{81}Br	49.31	80.91629
^{127}I	100	126.90447

Mass of a chemical element

- **Isotopic mass:**

Exact mass of an isotopic atom ^{A}X

^{1}H= 1.00787825 Da; ^{15}N=15.0001 Da; ^{235}U=235.044 Da.

- **Monoisotopic mass (Mm, A_0):**

Mass of the most abundant isotope

H= 1.00787825 Da; N=14.0031 Da; U=238.050 Da

- **Nominal mass (Nm):**

Integer mass of the most abundant stable isotope:
H=1, N=14, U=238 Da

Mass of a chemical element

- **Relative mass (A_r) or atomic weight (AW):**

Weighted average isotopic mass

$$A_r = \sum_i \alpha_i \cdot m_i$$

α_i - abundance of the i^th^ isotope

m_i - mass of the i^th^ isotope

$$\sum_i \alpha_i = 1$$

For carbon:

$$A_r = \alpha_{12} \cdot m_{12} + \alpha_{13} \cdot m_{13}$$

A_r = 0.9889 x12.0000 + 0.0111 x 13.00335 = 12.011

Mass of a molecule

is the sum of masses of all atoms: $$XM = \sum_i n_i \cdot Xm_i$$

n_i – number of i-th chemical element;
m_i – mass of the i-th chemical element;
X – type of the mass (M_0, N, W).

H_2O: i=1, 2; n_1=2, n_2=1;
Mm_1=1.007878 Da, Mm_2=15.994915 Da
Nm_1= 1 Da, Nm_2=16 Da

- ***Monoisotopic mass (MM, M_0)*:** $$M_0 = \sum_{i=1}^{l} n_i A_0^i$$

  Mass of the molecule with all most abundant isotopes:

  H_2O = $^1H_2{}^{16}O$= 18.010565 Da; H_2O^+= 18.010002 Da (m_e=5.4858·10^{-4} Da)

- ***Nominal mass (NM):*** Integer mass of the molecule with all most abundant stable isotopes (nearest integer of MM): H_2O =18 Da, (the sum of atomic mass numbers)

- ***Average mass or Molecular weight (MW):***

  Weighted average mass over all natural isotopes of all atoms of the molecule.

How to calculate Average mass (MW) ?

The mass calculated using a weighted average of the natural isotopes for the atomic mass of each chemical element

$$MW = \sum_{i=1}^{l} n_i A_r^i$$

A_r^i - average atomic mass (atomic weight);

m_k^i - mass of k-th isotope of i-th chemical element;

α_k^i - isotopic abundance of k-th isotope of i-th chemical element;

n_i - number of i-th chemical element;

l - number of different chemical elements

CO_2: $M_0 = 12 + 15.995 \cdot 2 = 43.990$ Da; NM= 44 Da;

MW=12.011+ 15.9994·2= 44.0098 Da

$$MW = \sum_{i=1}^{l} n_i A_r^i.$$

CCl_4: $^{12}C=12$ (98.9%), $^{13}C= 13.0034$; $^{35}Cl = 34.9688$ (75.7%), $^{37}Cl=36.9659$ (24.3%); $n_C=1$, $n_{Cl}=4$

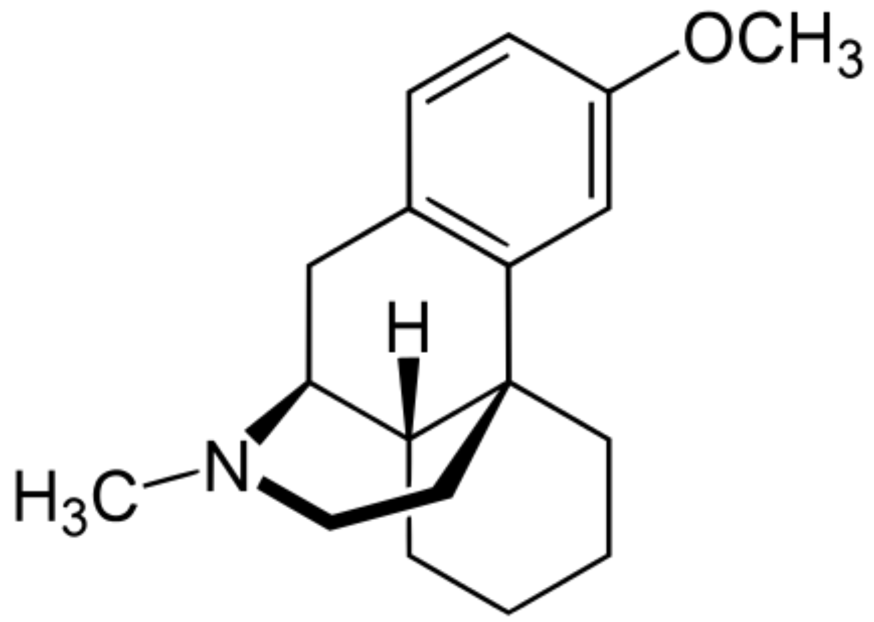
M_0= 12+ 34.96885·4 =151.8754 Da

$$MW = \sum_{i=1}^{l} \sum_{k} n_i m_k^i \alpha_k^i$$

MW= (12·0.989+13.0034·0.011) + (34.9688·0.757+36.9659·0.243)·4 =153.8225 Da

More examples

Consider the molecule Levomethorphan ($C_{18}H_{25}NO$), an opioid analgesic



Nominal mass:

(18*12)+(25*1)+(1*14)+(1*16) = 271

Average molecular mass:

This is what we weight, but not what we measure in a mass spectrometer!

(18*12.0107)+(25*1.0079)+(1*14.0067)+(1*15.9994) = 271.396

More examples

Consider the molecule Levomethorphan ($C_{18}H_{25}NO$), an opioid analgesic

Monoisotopic mass:

Consider the protonated form $[M+H]^+$: $\longrightarrow C_{18}H_{25}NOH^+$

This is calculated with the most abundant isotope of each element

(18*12.0000)+(25*1.0078)+(1*14.0031)+(1*15.9949) +1.0073 = 272.2003

Exact mass of a specific isotope:

Consider again the protonated form $[M+H]^+$. Must specify the isotopes.

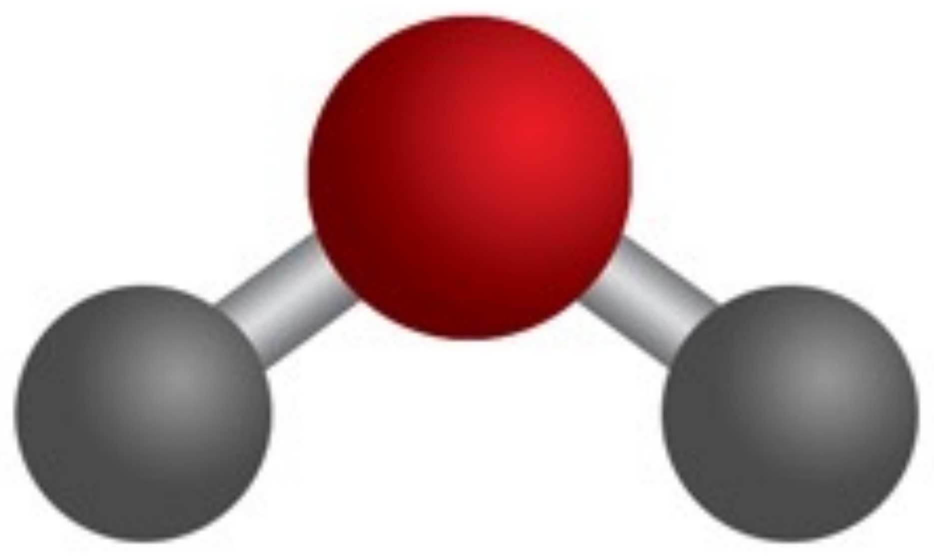
Assume $C_{18}H_{25}{}^{15}NOH^+$

(18*12.0000)+(25*1.0078)+(1*15.0001)+(1*15.9949) +1.0073 = 273.1973

Mass of molecules

$$M^n = \sum_{i=1}^{n} C_i$$

Chemical bond has a mass!



$$E = mc^2 \;\; => \;\; m = E/c^2$$

H—OH: D_0 = 41145.92 cm^{-1}=8.2·10^{-19} J; c= 3·10^8 m/s;

δm=10^{-35} kg = 5.5·10^{-9} Da.

The required resolution R= 3·10^9…. (currently 31·10^6 at best)

For a small protein with 10^4 bonds: δm ≈ 10^{-5} Da.

Not yet feasible, but may come…

Types of molecular ions

1. Ionization: $M - e^- = M^+$: Electron removed; *open-shell* positive ion.

2. Electron attachment: $M + e^- = M^-$: *open-shell* negative ion.

3. Protonation: $M + p^+ = [M+H]^+$: a proton added; *closed-shell* positive ion.

4. Deprotonation: $M - p^+ = [M-H]^-$: a proton removed; *closed-shell* negative ion.

Properly account for the type of ion when calculating its mass!

$M_p = 1.007276$ Da

$m_e = 5.4858 \cdot 10^{-4}$ Da

Isotopic profiles of molecules

- May allow for determination of the charge of an ion => ionic mass => mass of a molecule.

- May allow for determination of chemical composition (the number of certain chemical elements) of a molecule.

Some definitions

Isobars: Ions with the same nominal mass. **Example:** CO and N_2 (NM= 28 Da)

Isotopologues

- Isotopologues are molecules that differ only in the isotopic composition of one or more of their atoms
- Isotopologues have *different* mass

Examples:

CH_4/ CDH_3/CD_4/$^{14}CH_4$

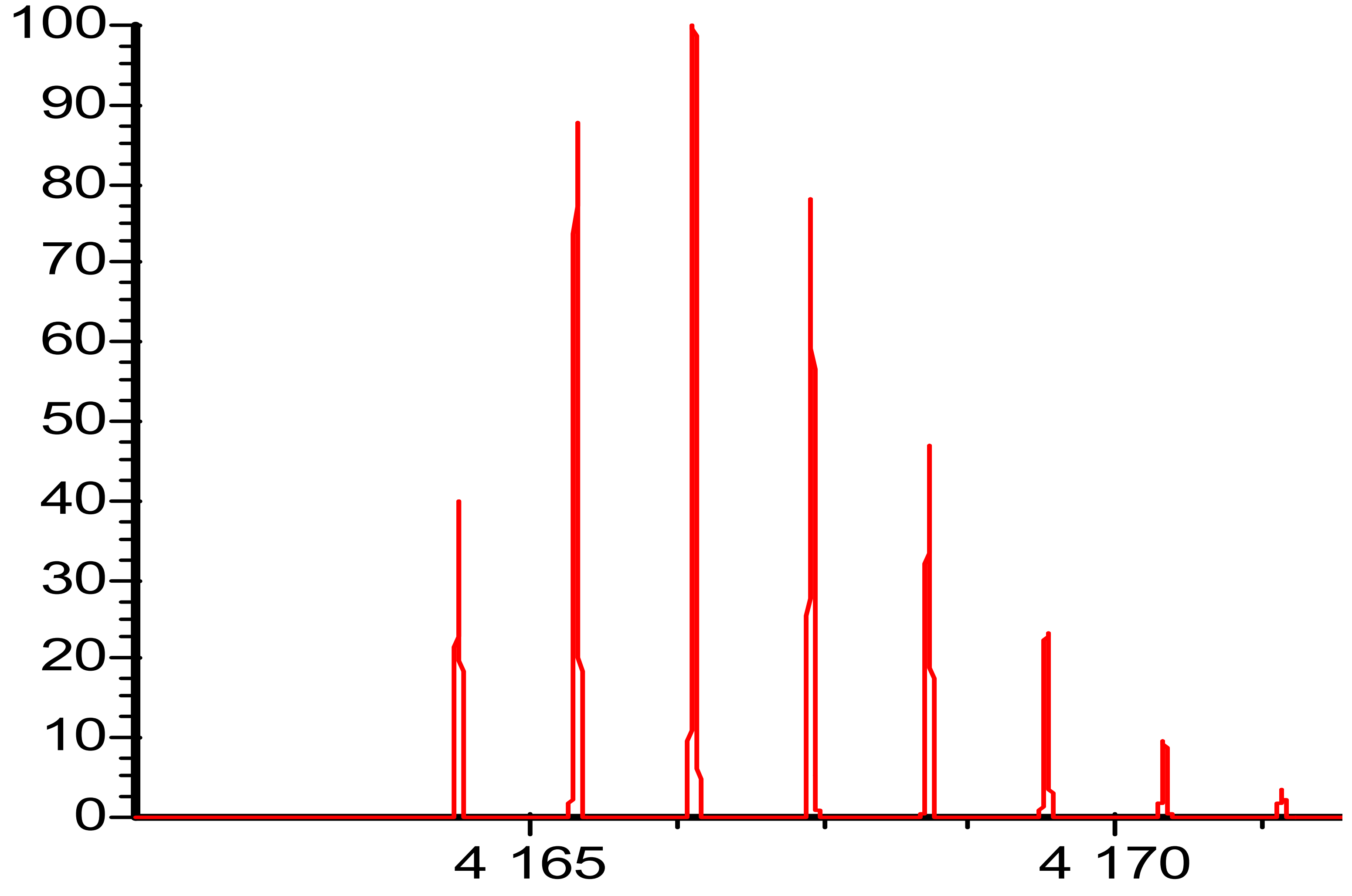
H_2O / HDO / D_2O

$^{10}BF_3$ / $^{11}BF_3$

Isotopic distribution

- Relative abundance of all natural isotopologues of the same molecule. = Relative intensities of mass peaks in an MS spectrum of single compound.

Isotopic distribution carries two pieces of information about chemical composition of a molecule



Natural isotopic abundance of some elements

Element	A		A+1		A+2		Element type
	Mass	%	Mass	%	Mass	%	
H	1	100	2	0.015			A
C	12	100	13	1.1			A+1
N	14	100	15	0.37			A+1
O	16	100	17	0.04	18	0.20	A+2
F	19	100					A
Si	28	100	29	5.1	30	3.4	A+2
P	31	100					A
S	32	100	33	0.79	34	4.4	A+2
Cl	35	100			37	32.0	A+2
Br	79	100			81	97.3	A+2
I	127	100					A

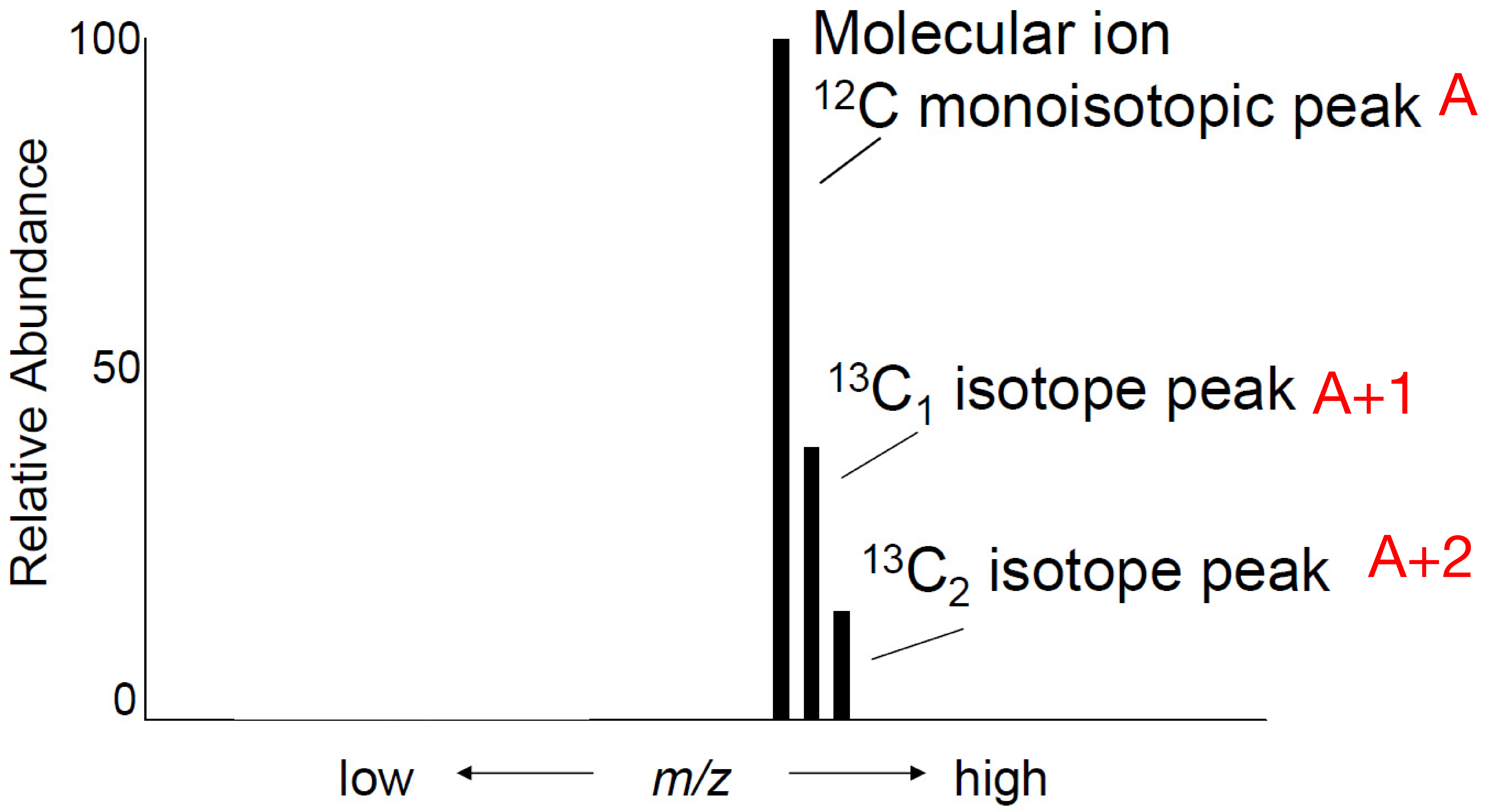
- Stable isotopes only
- For most of large organic molecules isotopic peaks due to carbon are the most abundant.

Isotopic distribution due to ^{13}C

Observations

- A peak (monoisotopic) is the most abundant.
- Relative intensity of A+1 peak is higher than natural abundance of ^{13}C.
- A+2 peaks is present, although there are only two stable isotopes.

Mass spectrum of a pure compound (bar plot)



- ➢ Why ?
- ➢ What can we determine for the chemical composition of the molecule ?

Statistics !

Calculating a relative isotopic distribution

Consider only carbon isotopes (abundances α_{12}= 0.989 and α_{13}= 0.011) for a molecule with n- number of carbon atoms

Probability for a molecule to have all carbons as ^{12}C (Monoisotopic peak A):

$$P_A = \alpha_{12}^n$$

Probability for a molecule to have only one ^{13}C (First isotopic peak, A+1):

$$P_{A+1} = n \cdot \alpha_{13} \cdot \alpha_{12}^{n-1}$$

Probability to have two ^{13}C atoms (Second isotopic peak, A+2):

$$P_{A+2} = \frac{1}{2} \cdot n(n-1) \cdot \alpha_{13}^2 \cdot \alpha_{12}^{n-2}$$

Probability to have m ^{13}C atoms (m-th isotopic peak, A+m):

$$P_{A+m} = \frac{n!}{(n-m)!m!} \cdot \alpha_{13}^m \cdot \alpha_{12}^{n-m}.$$

For a molecular ion the A+***m*** peak means ***m*** number of neutrons, compare with the monoisotopic peak.

!!! The sum of all probabilities (m= 0,…n) is equal to 1.

Calculating a relative isotopic distribution

Relative intensity of the peak in % for ion with ***m*** number of ^{13}C and ***n*** total number of carbons: ***I***(A+m):

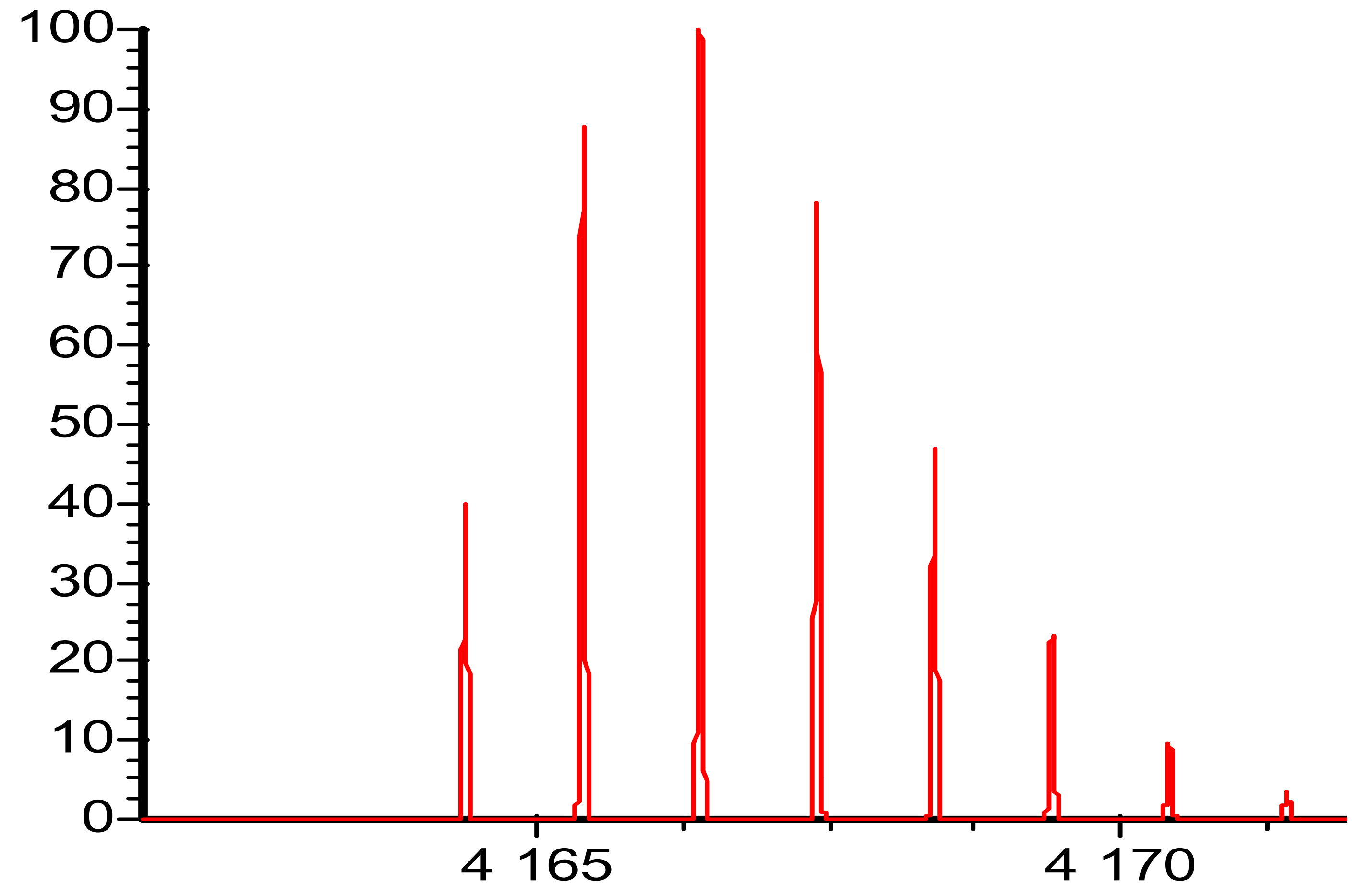
$$\boxed{I(A+m) = \frac{P_{A+m}}{P_A} \cdot 100\%}$$

Note that: $$\frac{I_{A+m+1}}{I_{A+m}} = \frac{(n-m)!m!}{(n-m-1)!(m+1)!} \cdot \frac{\alpha_{13}^{m+1}\alpha_{12}^{n-m-1}}{\alpha_{13}^{m}\alpha_{12}^{n-m}} = \frac{(n-m)\cdot\alpha_{13}}{(m+1)\cdot\alpha_{12}} \simeq 0.011 \cdot \frac{(n-m)}{(m+1)};$$

For m=0 : $$\frac{I_{A+1}}{I_A} \simeq 0.01112 \cdot n \quad \Rightarrow \quad \boxed{n_C \approx 90 \cdot I_{A+1}/I_A}$$

For m=1 : $$\frac{I_{A+2}}{I_{A+1}} \simeq 0.00556 \cdot (n-1) \quad \Rightarrow \quad n_C \approx 180 \cdot I_{A+2}/I_{A+1} + 1$$

....



!!! *Number of carbons can be estimated from ratios of observed subsequent isotopic peaks.*

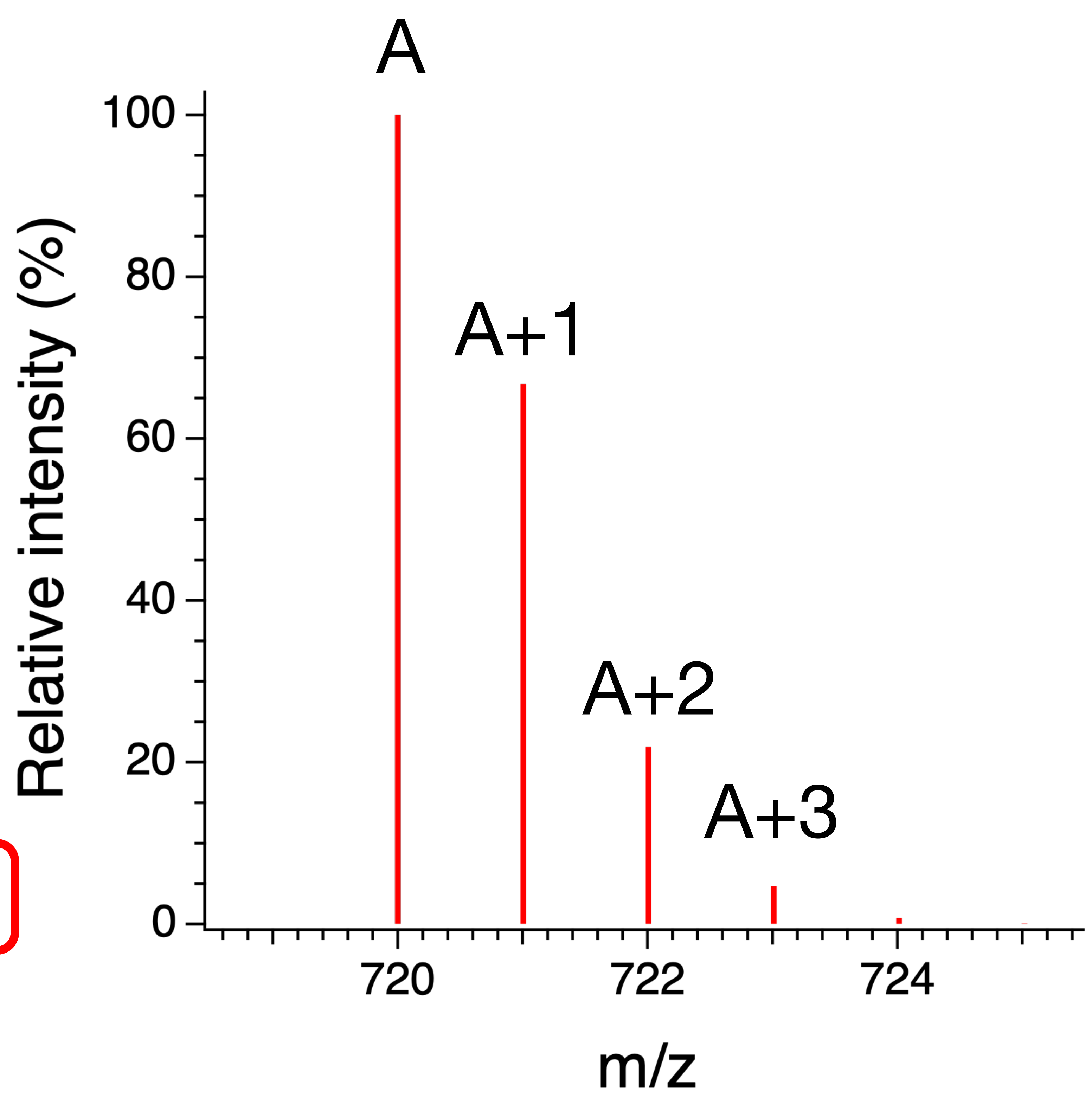
Fit of all intense peaks may improve the accuracy

Calculating a relative isotopic distribution

If the abundance of peak A is 100%,
then the contributions from the isotopic peaks A+1, A+2, and A+3 are:

n	(A+1)	(A+2)	n	(A+1)	(A+2)	(A+3)
C_1	1.1	0.00	C_{16}	18	1.5	0.1
C_2	2.2	0.01	C_{17}	19	1.7	0.1
C_3	3.3	0.04	C_{18}	20	1.9	0.1
C_4	4.4	0.07	C_{19}	21	2.1	0.1
C_5	5.5	0.12	C_{20}	22	2.3	0.2
C_6	6.6	0.18	C_{22}	23	2.8	0.2
C_7	7.7	0.25	C_{24}	26	3.3	0.3
C_8	8.8	0.34	C_{26}	29	3.9	0.3
C_9	9.9	0.44	C_{28}	31	4.5	0.4
C_{10}	11.0	0.54	C_{30}	33	5.2	0.5
C_{11}	12.1	0.67	C_{35}	39	7.2	0.9
C_{12}	13.2	0.80	C_{40}	44	9.4	1.3
C_{13}	14.3	0.94	C_{50}	55	15	2.6
C_{14}	15.4	1.1	C_{60}	66	21	4.6
C_{15}	16.5	1.3	C_{100}	110	60	22

Fullerene C_{60}



How to calculate an isotopic distribution considering essential isotopes of all elements

1. Start from the molecular formula, and calculate the accurate masses for all possible isotopic combinations: A, A+1, A+2, . . .
   - Do not consider ^{2}H and ^{17}O isotopes — they are not sufficiently abundant.
   - Do not consider any more than 3-5 isotope peaks, unless explicitly asked.

2. For each isotopic peak (A, A+1, A+2), calculate the **abundance** A separately for each chemical element using the following formula:

$$A = \frac{n!}{(a)!(b)!(c)!\ldots}(r_1)^a(r_2)^b(r_3)^c\ldots$$

where ***n*** is the number of atoms of a given element; ***a***, ***b***, and ***c*** are the numbers of each type of isotopes of this element (a+b+c=n), and $\mathbf{r_1}$, $\mathbf{r_2}$, $\mathbf{r_3}$ are the abundances of each isotope.

3. For each isotopic peak (A, A+1, A+2), calculate the total abundance (combination of all participating elements) using the following formula:

Total A = A(element 1) x A(element 2) x A(element 3) . . .

How to calculate an isotopic distribution: example

MF: $C_{112}H_{165}N_{27}O_{36}$ Pic: $^{12}C_{110}{}^{13}C_2H_{165}{}^{14}N_{27}{}^{16}O_{36}$ (A_2)

Accurate mass: 2466.1978 Da

$$Pr(^{12}C_{110}{}^{13}C_2) = \frac{112!}{110! \times 2!} Pr(^{12}C)^{110} \times Pr(^{13}C)^2 = 0.2180$$

$$Pr(^{1}H_{165}{}^{2}H_0) = \frac{165!}{165! \times 0!} Pr(^{1}H)^{165} \times Pr(^{2}H)^0 = 0.9812$$

$$Pr(^{14}N_{27}{}^{15}N_0) = \frac{27!}{27! \times 0!} Pr(^{14}N)^{27} \times Pr(^{15}N)^0 = 0.9057$$

$$Pr(^{16}O_{36}{}^{17}O_0{}^{18}O_0) = \frac{36!}{36! \times 0! \times 0!} Pr(^{16}O)^{36} \times Pr(^{17}O)^0 \times Pr(^{18}O)^0 = 0.9161$$

Peak abundance: 0.2180 x 0.9812 x 0.9057 x 0.9161 = 0.1775

(total of all = 1)

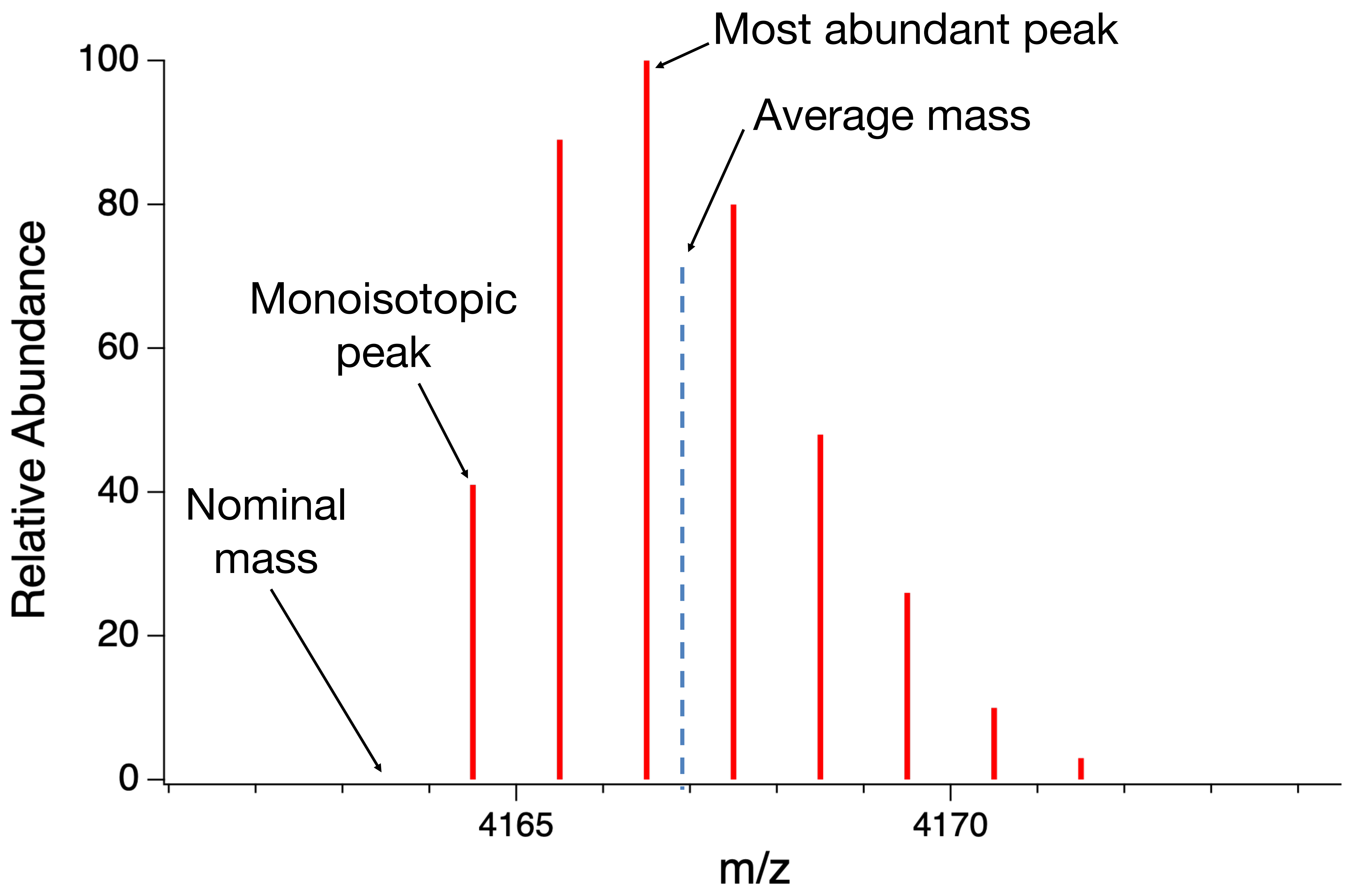
Isotope distribution calculators

https://www.envipat.eawag.ch/index.php

https://www.sisweb.com/mstools/isotope.htm

http://www.chemcalc.org

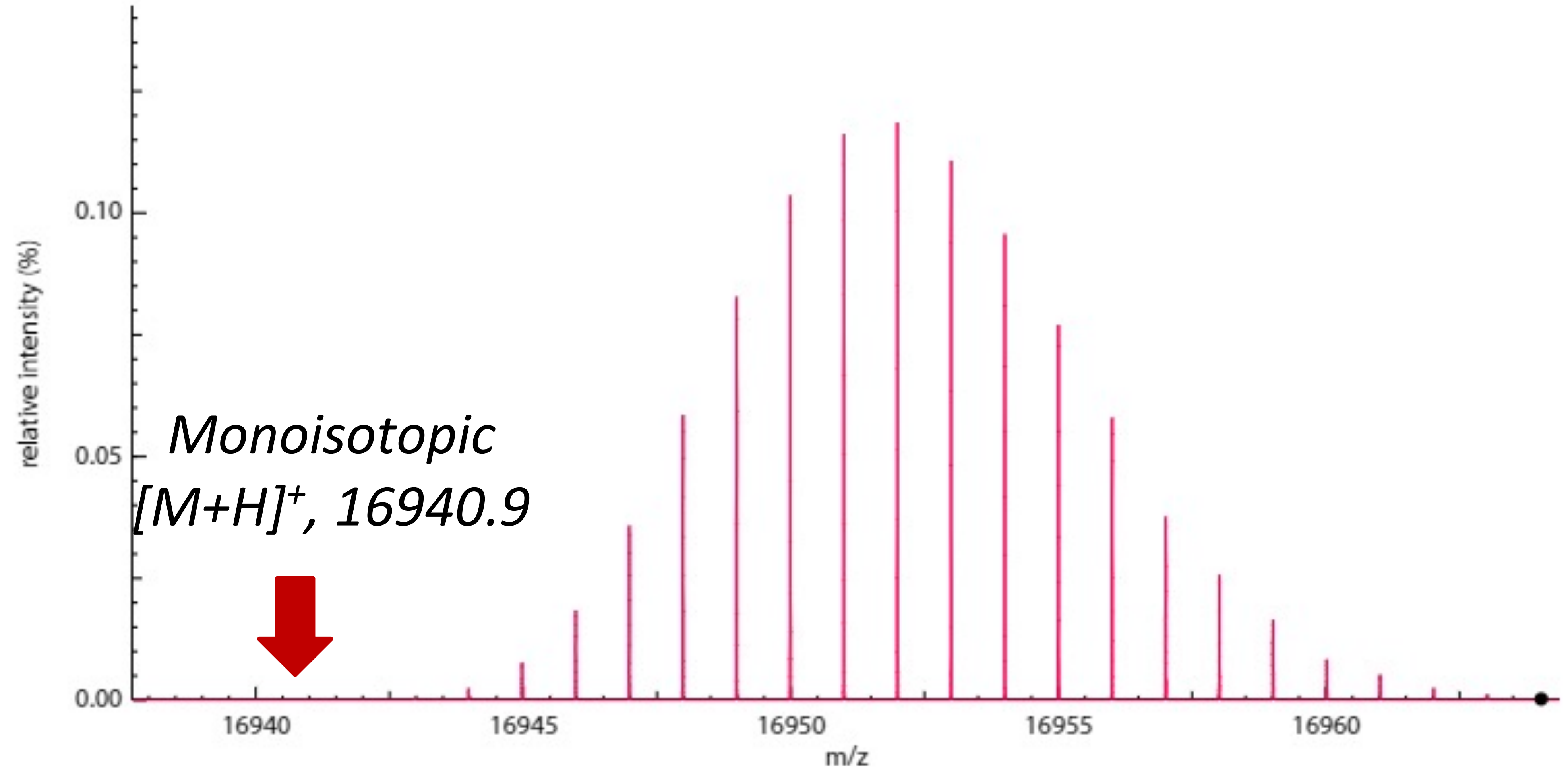
Example: Mass of a peptide



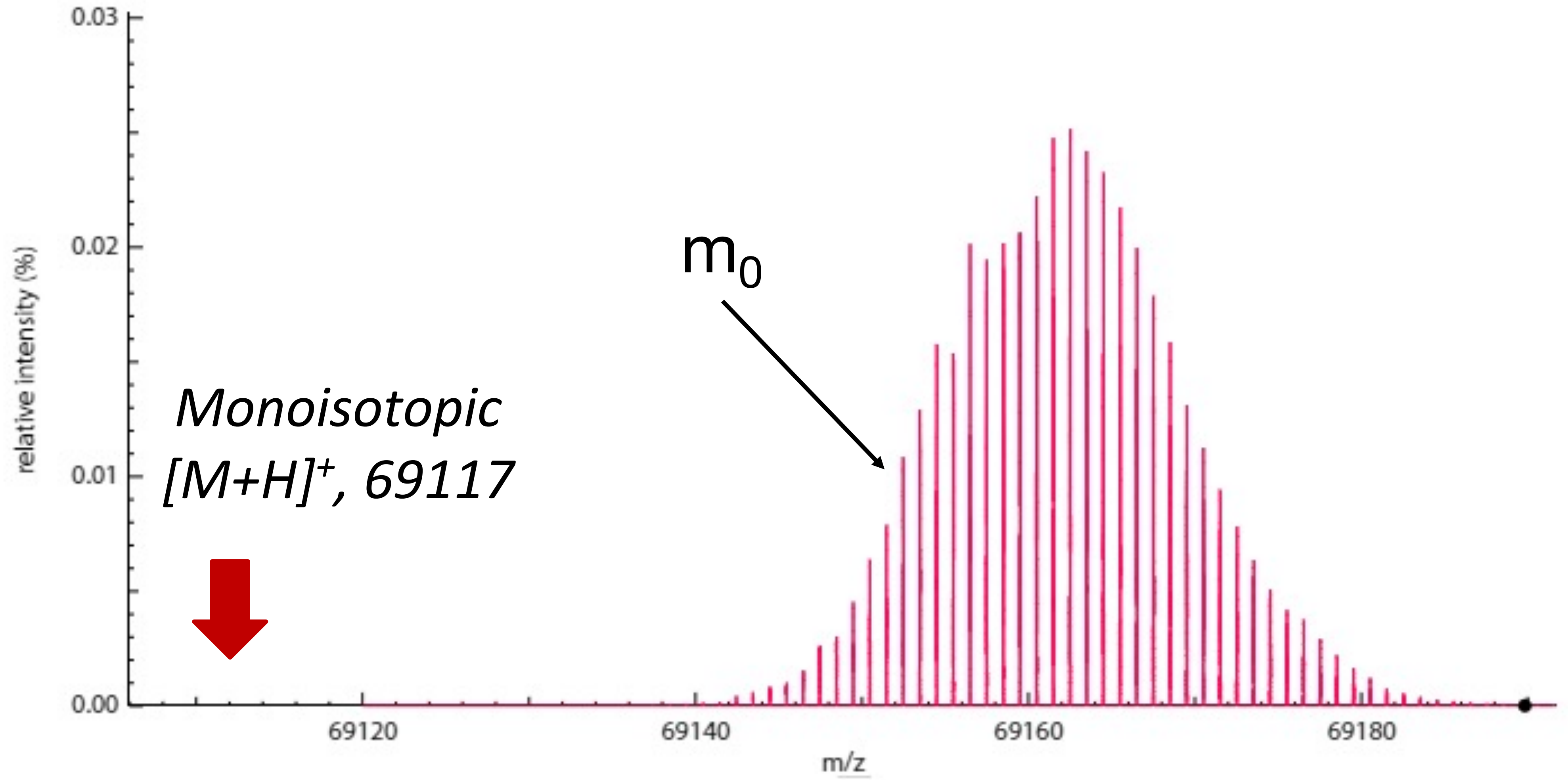
- Isotopic distribution for peptides is mainly due to ^{13}C
- Isotopic distribution allows for calculation of average mass
- Isotopic distribution allows for determination the number of carbons
- Other isotopes slightly change isotopic distribution

Monoisotopic mass for large molecules

Myoglobin, $C_{769}H_{1212}N_{210}O_{218}S_2$, 16.9 kDa



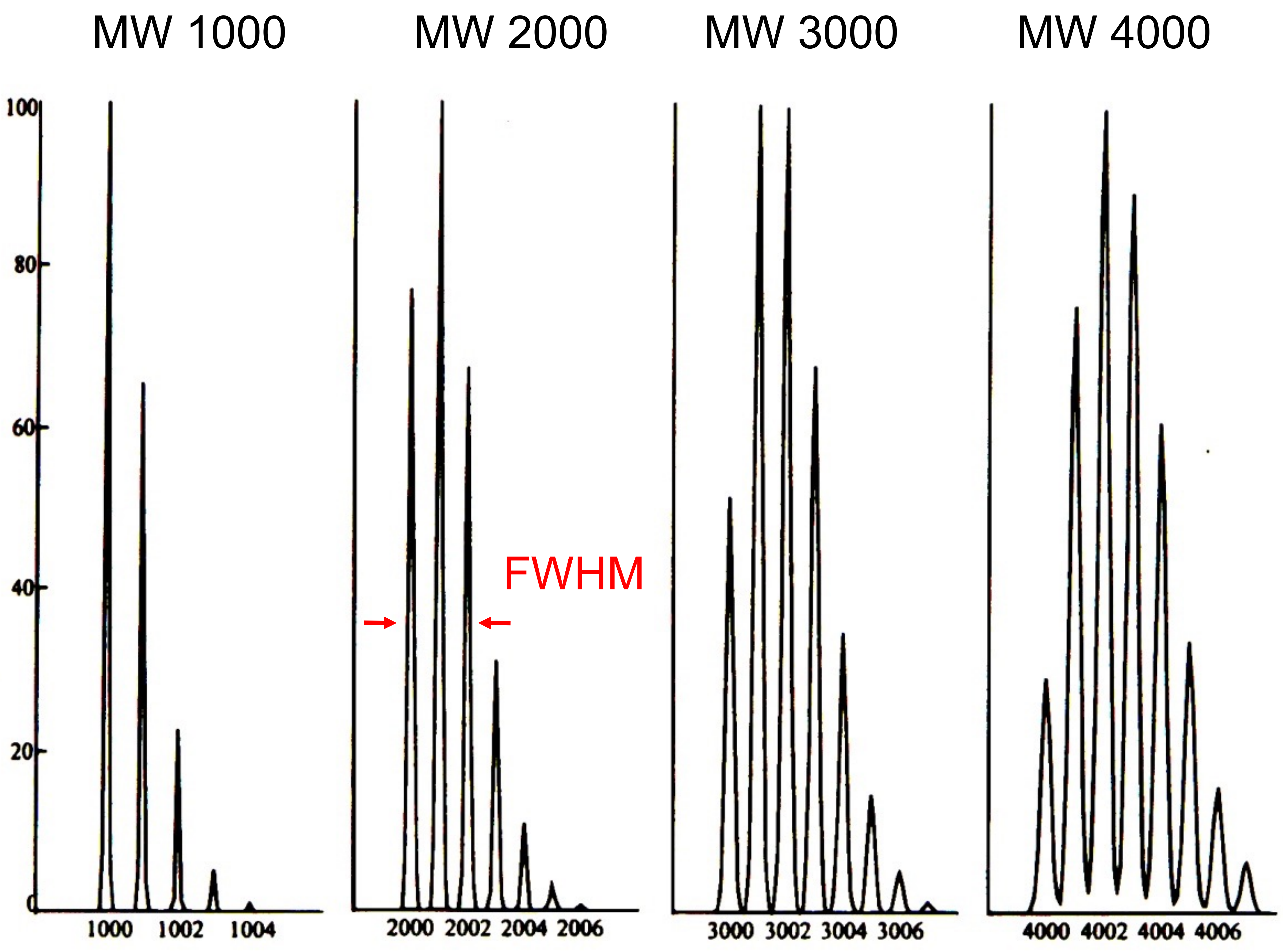
Serum Albumin, $C_{3066}H_{4817}N_{815}O_{926}S_{39}$, 69.1 kDa



Make a fit with

$$\frac{I_{A+m+1}}{I_{A+m}} \simeq 0.011 \cdot \frac{(n-m)}{(m+1)};$$

How do isotopic distributions change with mass?



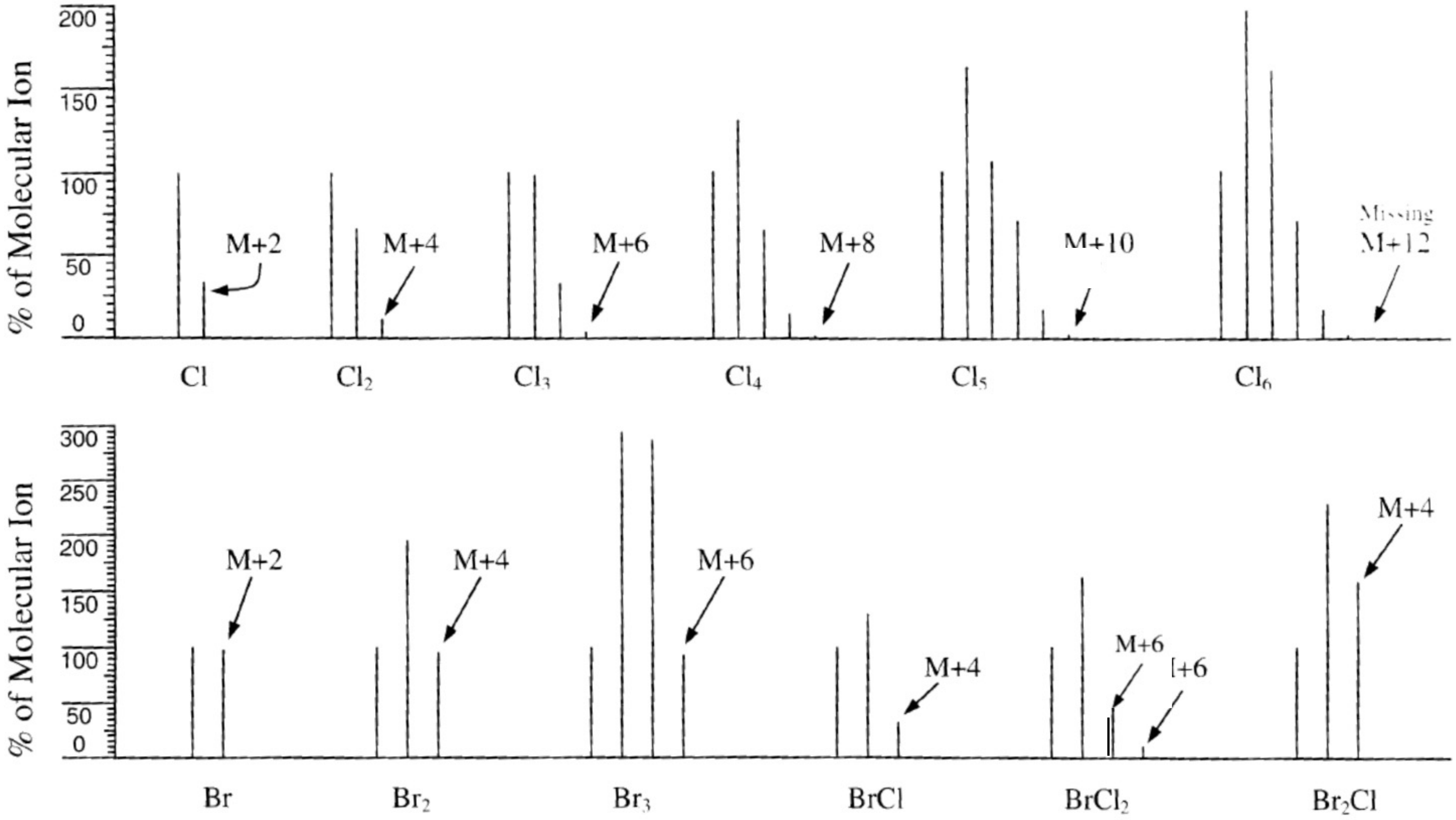
Theoretical isotope distributions of peptides of 1000, 2000, 3000 and 4000 Da.

Yergey J, Heller D, Hansen G, Cotter RJ, Fenselau C.
Anal. Chem. 1983, *55,* 353-356.

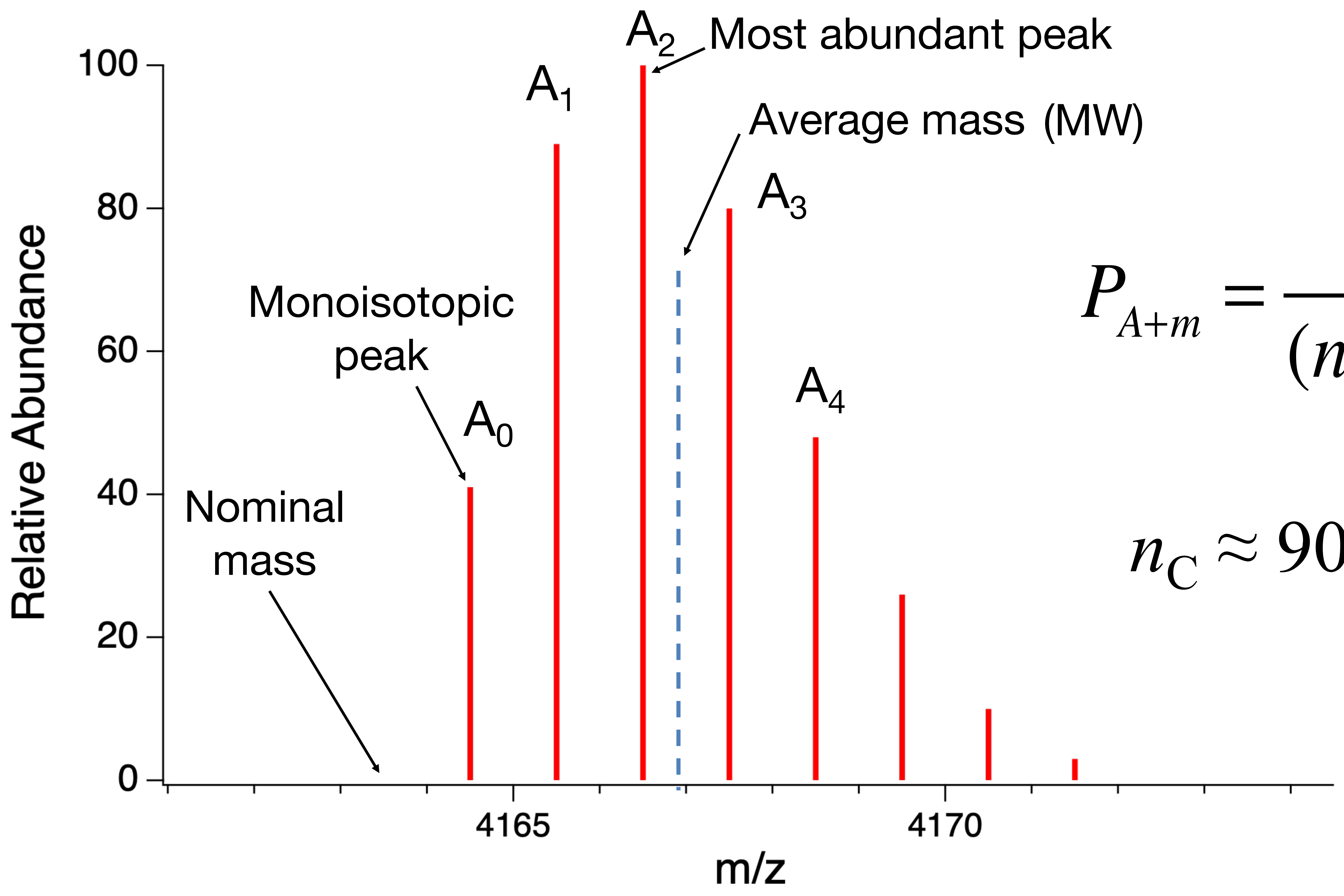
- Monoisotopic mass dominates up to MW ~1100 - 1500
- Above MW ~7000, the monoisotopic peak is vanishingly small
- Becomes more symmetric
- The width grows sublinearly.
  For <3 kDa, MW/FWHM ~1100
  For 10 kDa, MW/FWHM ~2000
- The most abundant mass is 0 to 1 Da below the average mass
- Fine structure for all peaks but monoisotopic.

Predicted patterns for compounds with various combination of Cl and Br

	Abundance	AM
^{35}Cl	75.77	34.96885
^{37}Cl	24.23	36.96590
^{79}Br	50.69	78.91838
^{81}Br	49.31	80.91629



Summary L1



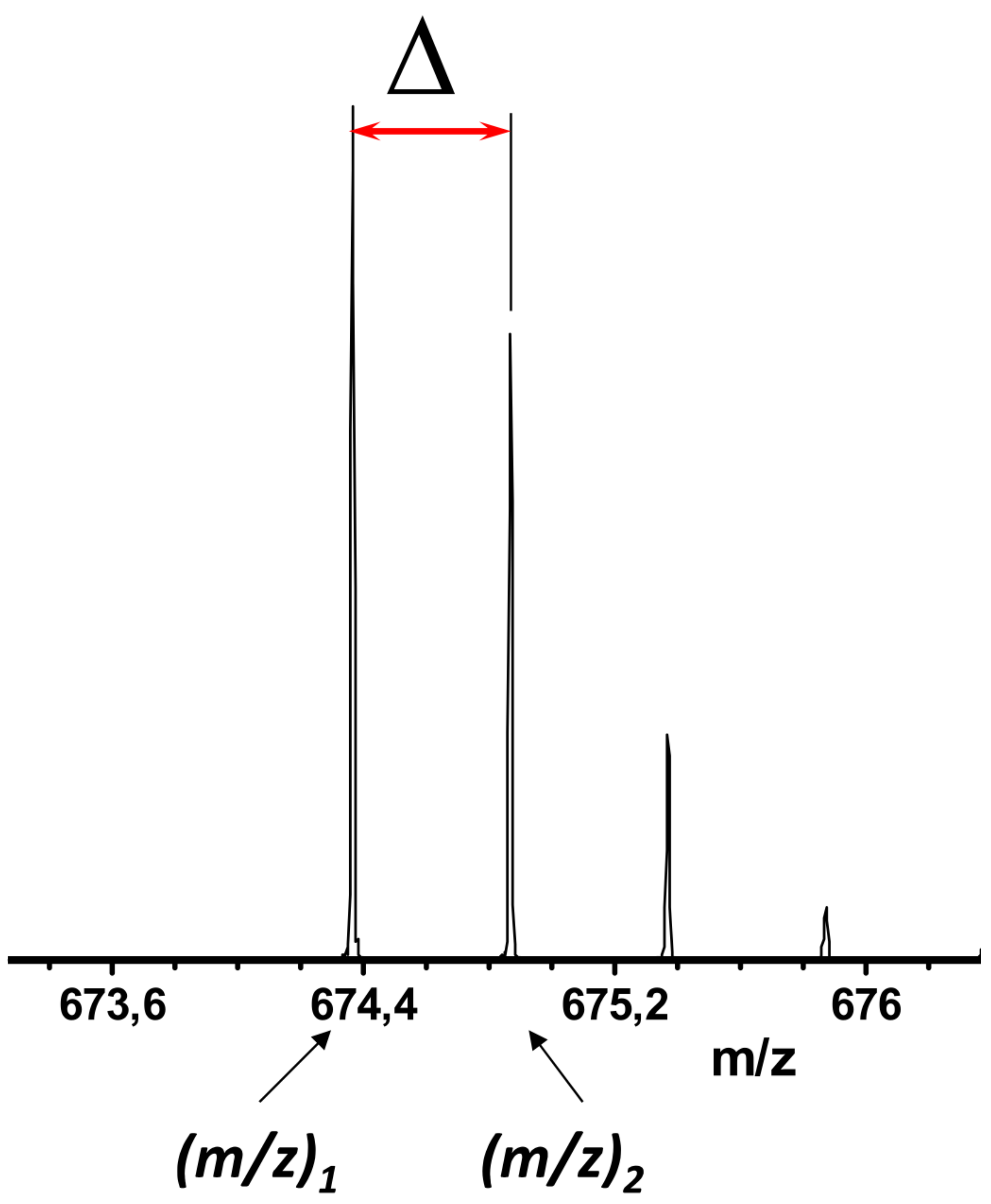
$$P_{A+m} = \frac{n!}{(n-m)!m!} \cdot \alpha_{13}^{m} \cdot \alpha_{12}^{n-m}.$$

$$n_C \approx 90 \cdot I_{A+1}/I_A$$

- Isotopic distribution for peptides is mainly due to ^{13}C
- Isotopic distribution allows for calculation of average mass
- Isotopic distribution allows for determination the number of carbons
- Other isotopes can be accounted for similar to ^{13}C

Isotope spacing: multiple charges

- As the charge state increases, the spacing between istoope peaks decreases.



$$\Delta = (m/z)_2 - (m/z)_1$$
$$= (m_2/z) - (m_1/z)$$
$$= (m_2 - m_1)/z = 1/z$$

$$z = 1/\Delta$$

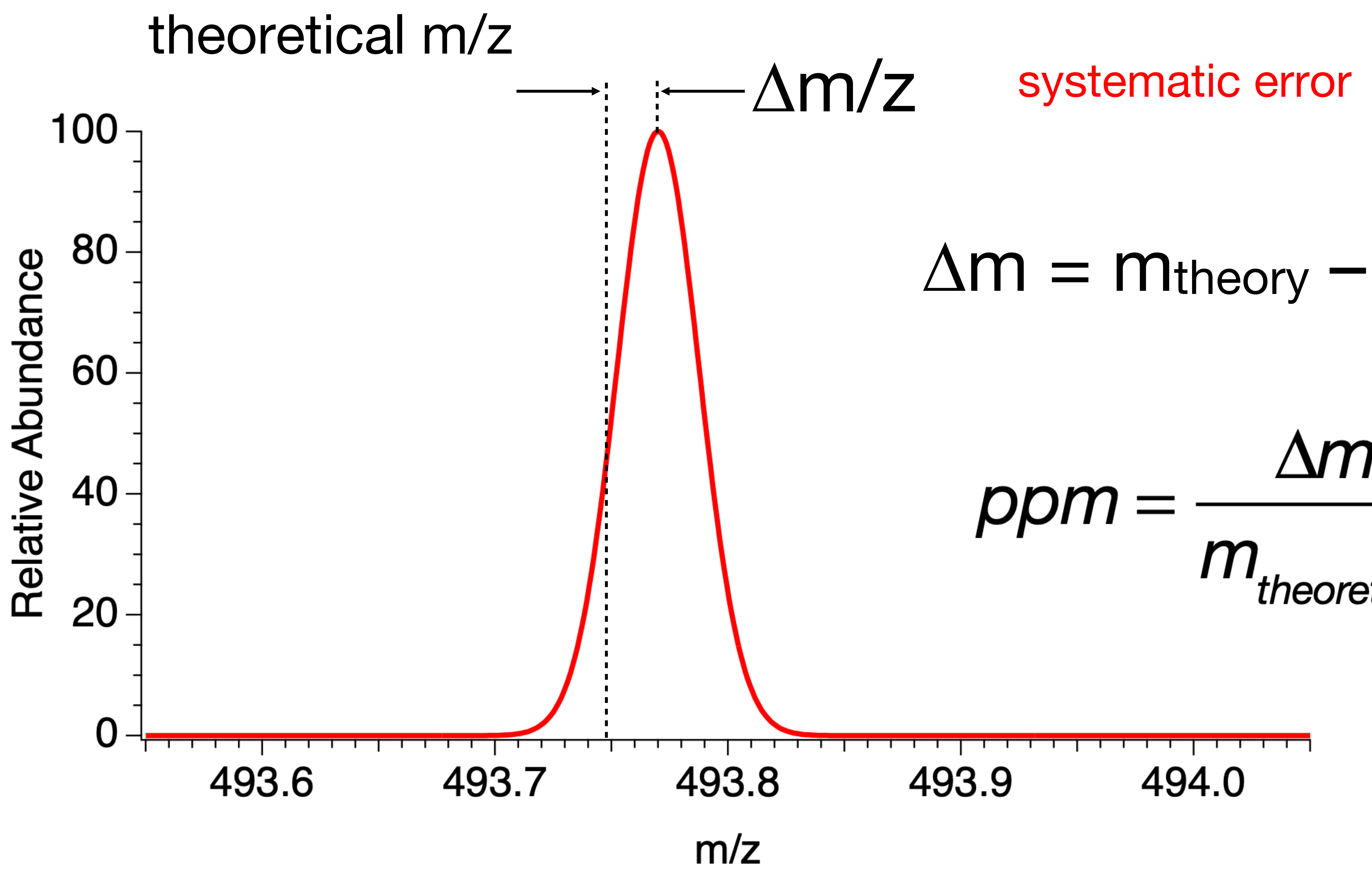
$\Delta = 1$ Da	z=1
$\Delta = 0.5$ Da	z=2
$\Delta = 0.33$ Da	z=3
$\Delta = 0.25$ Da	z=4

- Resolution needs to be increased to make fine structure visible for multiply-charged ions

Mass Accuracy and Resolution

Mass accuracy

Mass accuracy: Difference between the theoretical and experimental masses



$$\Delta m = m_{theory} - m_{experiment} \text{ (Da)}$$

$$ppm = \frac{\Delta m}{m_{theoretical}} \cdot 10^6$$

Mass accuracy

Examples of mass deviation

	m/z 100	m/z 500	m/z 1000
± 1 ppm	± 0.0001	± 0.0005	± 0.001
± 10 ppm	± 0.001	± 0.005	± 0.01
± 100 ppm	± 0.01	± 0.05	± 0.1

Mass accuracy

- Achieving good mass accuracy depends on:
  - Quality of calibration
  - Resolution
  - Signal intensity (S/N ratio)
  - Possible chemical interferences
- Accurate mass can be used to help determine **elemental composition.**
- Higher mass accuracy allows fewer possible compositions.
- A 1 ppm mass accuracy is sometimes sufficient for elemental formula assignment of molecules < 300 Da.
- The isotopic pattern provides additional information to discriminate between structures with similar masses.
- http://www.cheminfo.org/flavor/mass/Calculations/Find_a_MF_from_a_monoisotopic_mass/index.html

Mass accuracy

From monoisotopic mass to molecular formula: example of Reserpine

see http://www.cheminfo.org/flavor/mass/index.html for an exact mass calculator

$C_{33}H_{40}N_2O_9$ exact mass $[M+H]^+$ *m/z* 609.28066

Number of possible MFs (hits) with only C, H, N & O

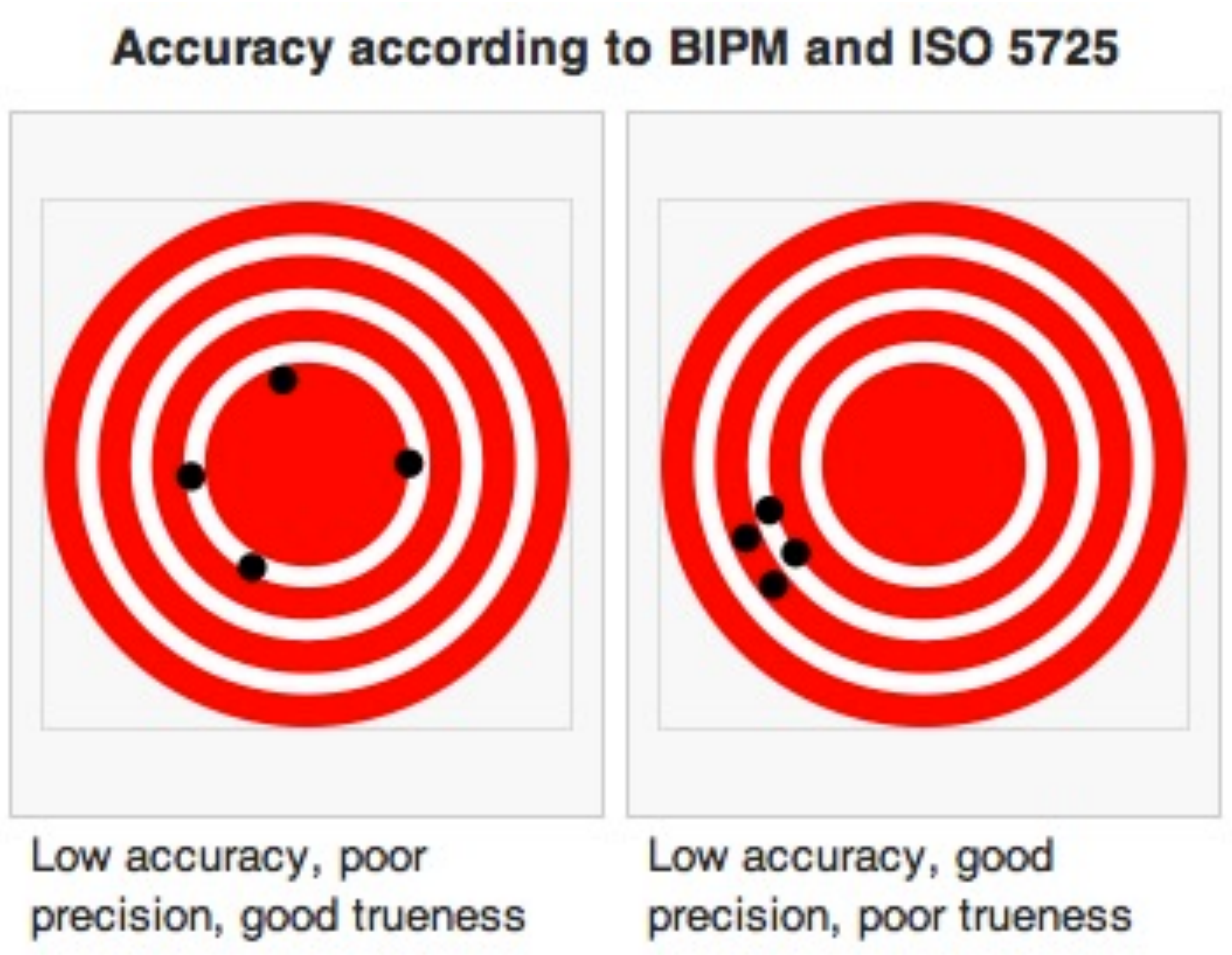
- Can't always reach a single MF by high-resolution MS only. Not always unambiguous.
- Can combine with other info to narrow possibilities
  - fine isotope structure
  - MS/MS
  - NMR

Mass Accuracy	number of compounds
200 ppm	53
10 ppm	4
5 ppm	1
3 ppm	1
1 ppm	1

This would be higher if we allowed more elements

Precision

Definition: Closeness of agreement among a set of repetitive measurements



Repeatability: *Short term stability of results on the same instrument*

Reproducibility: *Long term stability of results on any instrument*

Resolution

FWHM = width of a peak measured at the half of its height;

Peaks in MS are often of Gaussian shape (statistics)

Raley criteria of resolution:

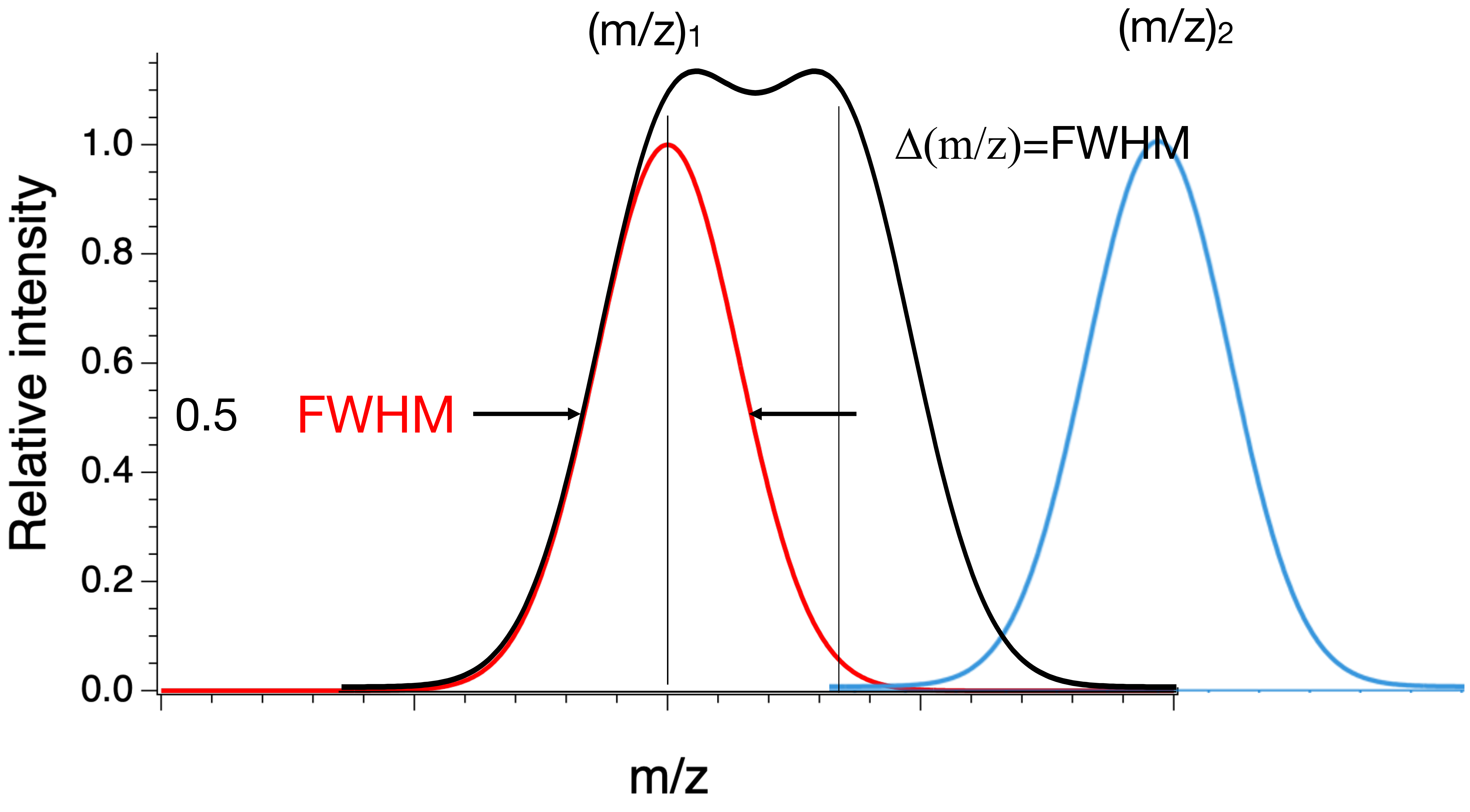
$$\Delta(m/z) = FWHM;$$

Definition:

Mass-resolution **R** is the ratio of ***m/z*** over the FWHM of the peak:

$$R = \frac{(m/z)}{FWHM};$$

Resolution is characteristic of a peak



Resolution determines FWHM of a mass-peak:

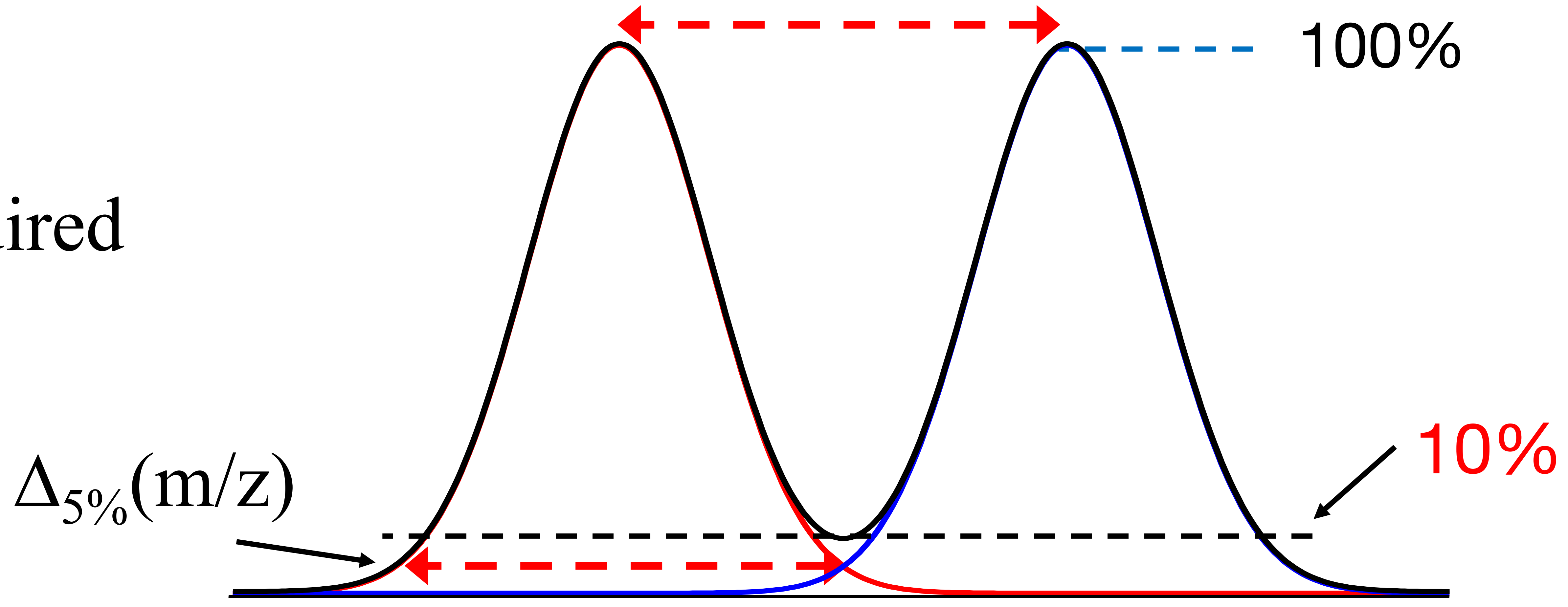
$$FWHM = \frac{(m/z)}{R};$$

Resolving power

Definition:

Resolving power **RP** is the ratio of (m/z) over the minimum separation of two equal peaks Δ(m/z) required for a MS instrument to distinguish them.

Resolving Power relates to two peaks



Most popular in MS criteria of RP is 10% valley:

$$\Delta(m/z) \equiv \Delta_{5\%}(m/z) = 1.8 \cdot FWHM;$$

$$RP_{10\%} = \frac{(m/z)}{\Delta_{5\%}(m/z)} \simeq \frac{(m/z)}{1.8 \cdot FWHM} = 0.56 \cdot R;$$

***RP** of an instrument may change with **m/z**.*

$$\boxed{R = 1.8 \cdot RP_{10\%} = 1.8 \cdot \frac{(m/z)}{\Delta(m/z)}.}$$

*With this **R** two peaks will be resolved with 10% valley.*

Example: CO^+ and N_2^+

$M_0(CO^+)$=12.00000 + 15.99492 - 0.00055 = 27.99437

$M_0(N_2^+)$=2 x 14.00307 - 0.00055 = 28.00559

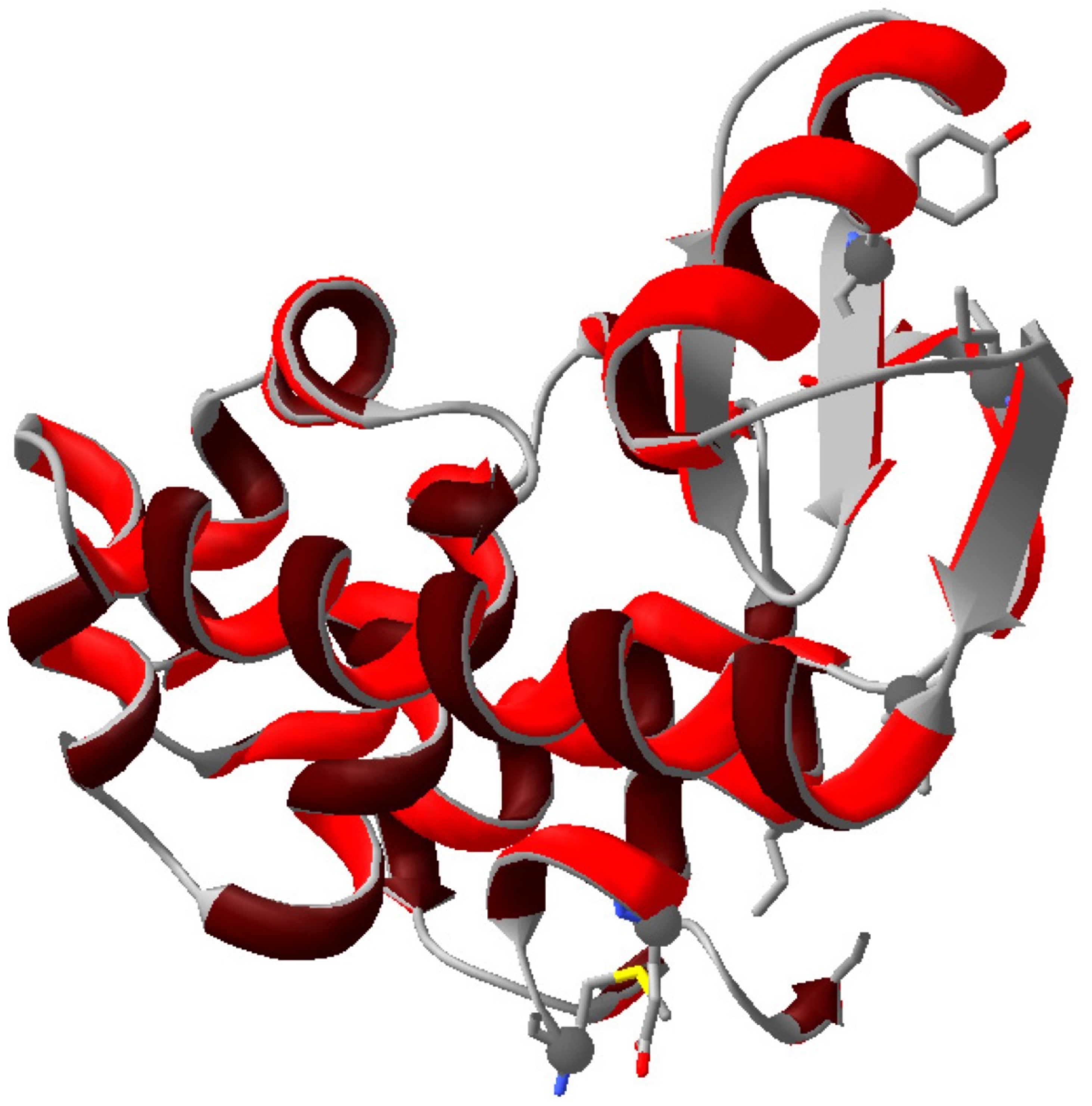
$$\Delta(m/z) = 0.01122; \quad \text{NM=28;}$$

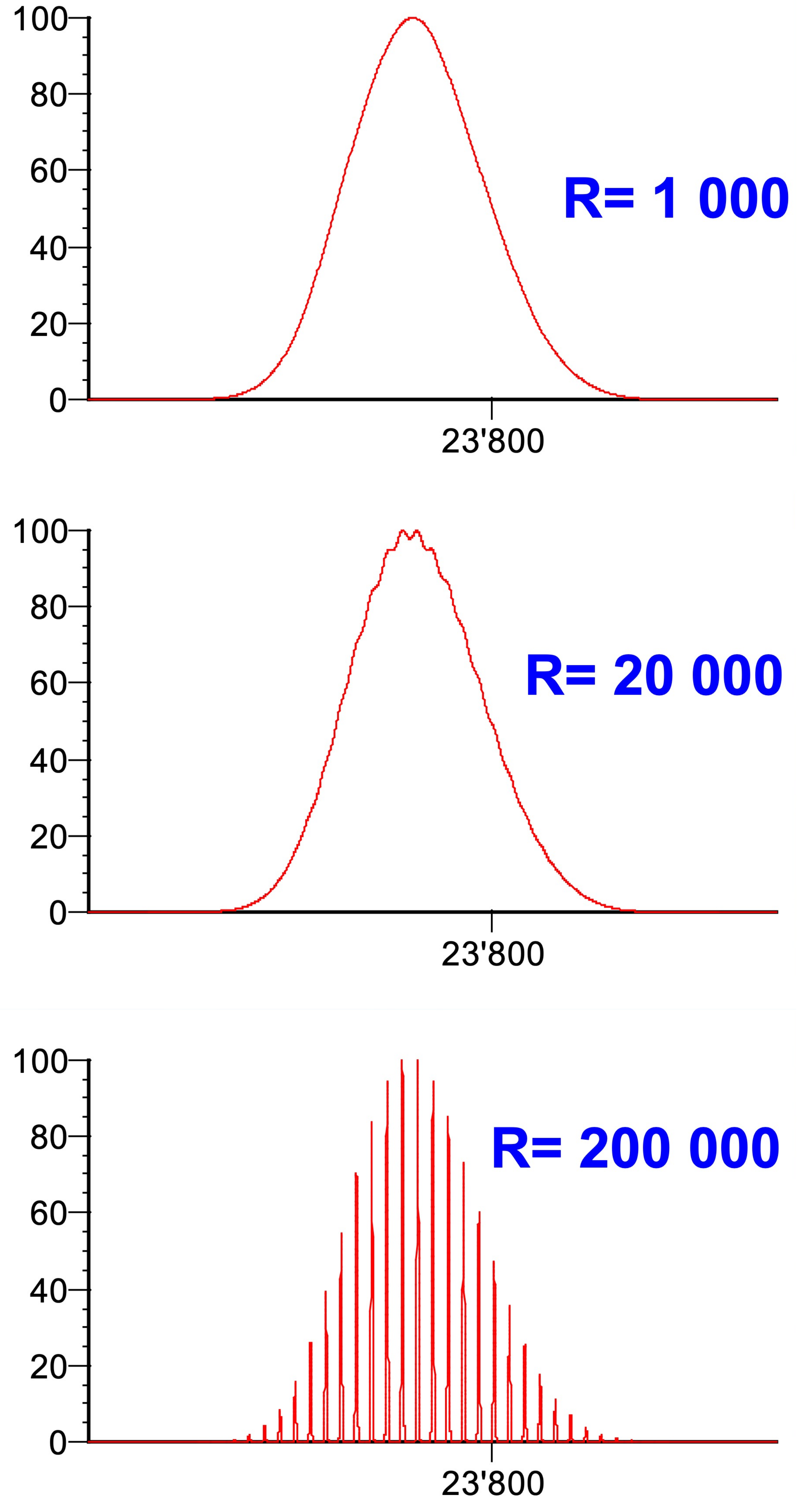
$$R = 1.8 \cdot \frac{28}{0.01122} \simeq 4500.$$

Keep only the first two non-zero digits

Resolution example

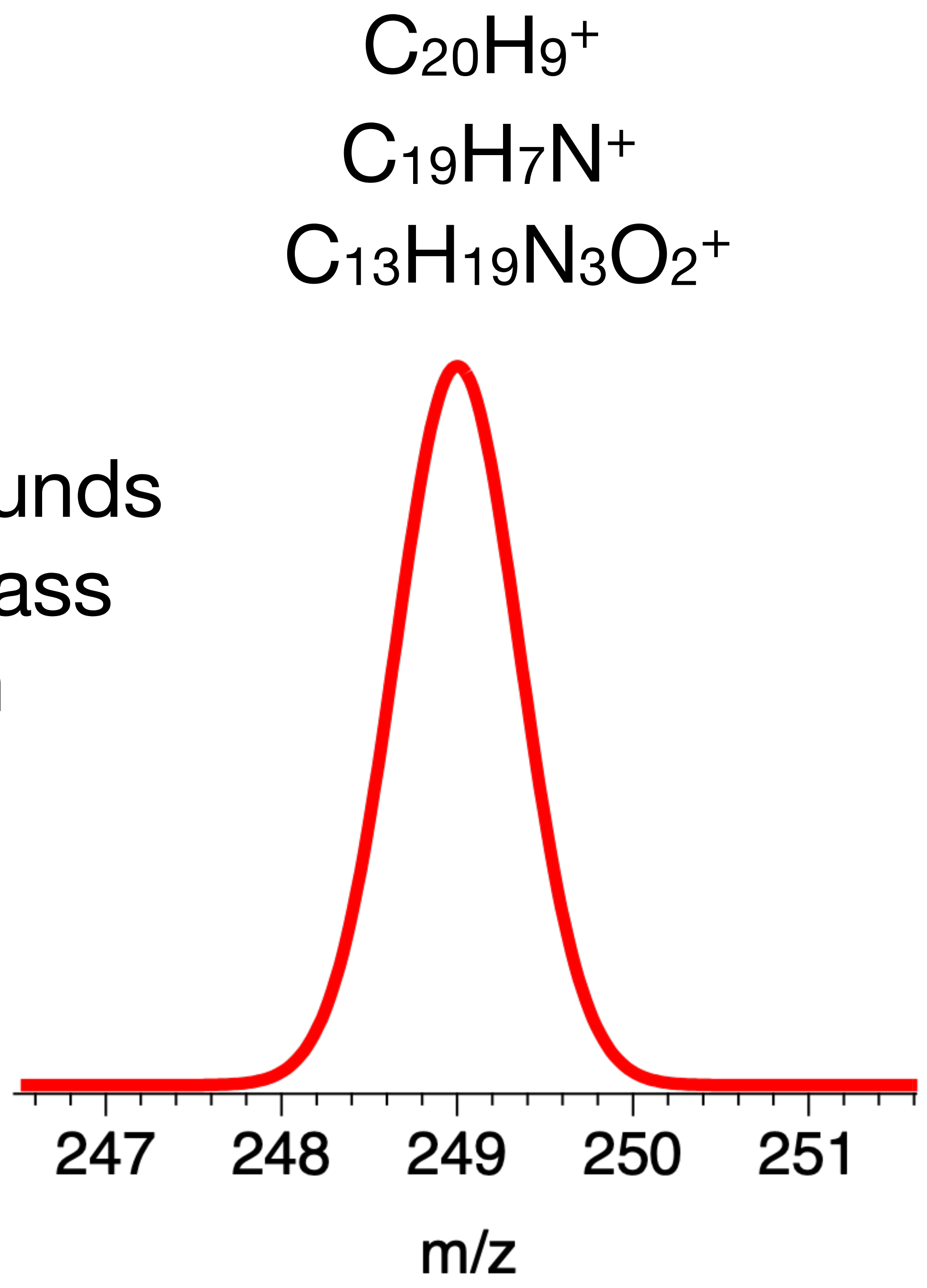
Protein of 24 kDa



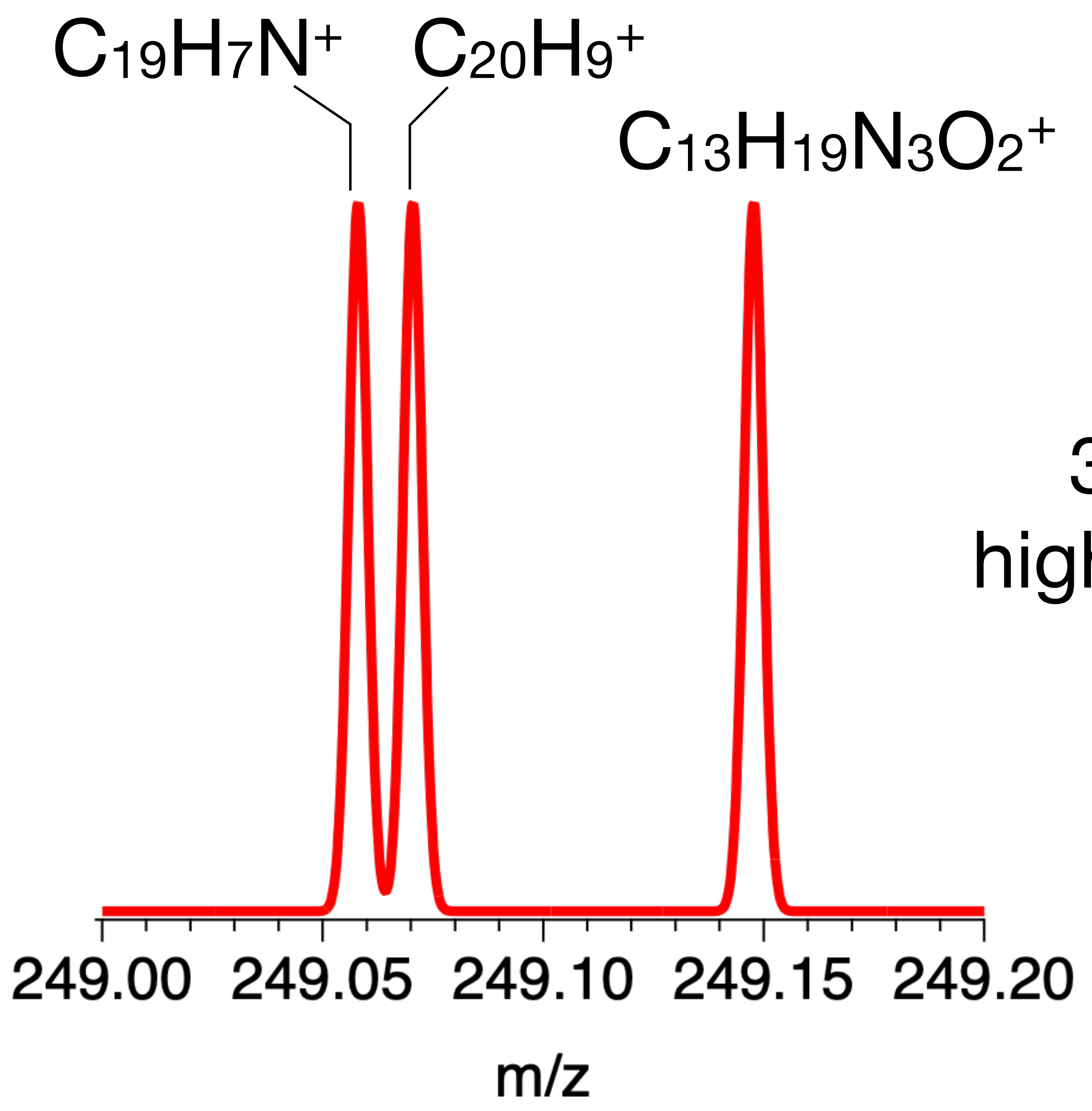


Mass accuracy and resolution

Isobaric ions:
Ions that have the same nominal mass



3 different compounds
same nominal mass
low resolution



3 different compounds
3 different exact masses
high resolution, high accuracy

M_0 ($C_{19}H_7N^+$) = 249.0578 Da

M_0 ($C_{20}H_9^+$) = 249.0704 Da

M_0 ($C_{13}H_{19}N_3O_2^+$) = 249.1477 Da