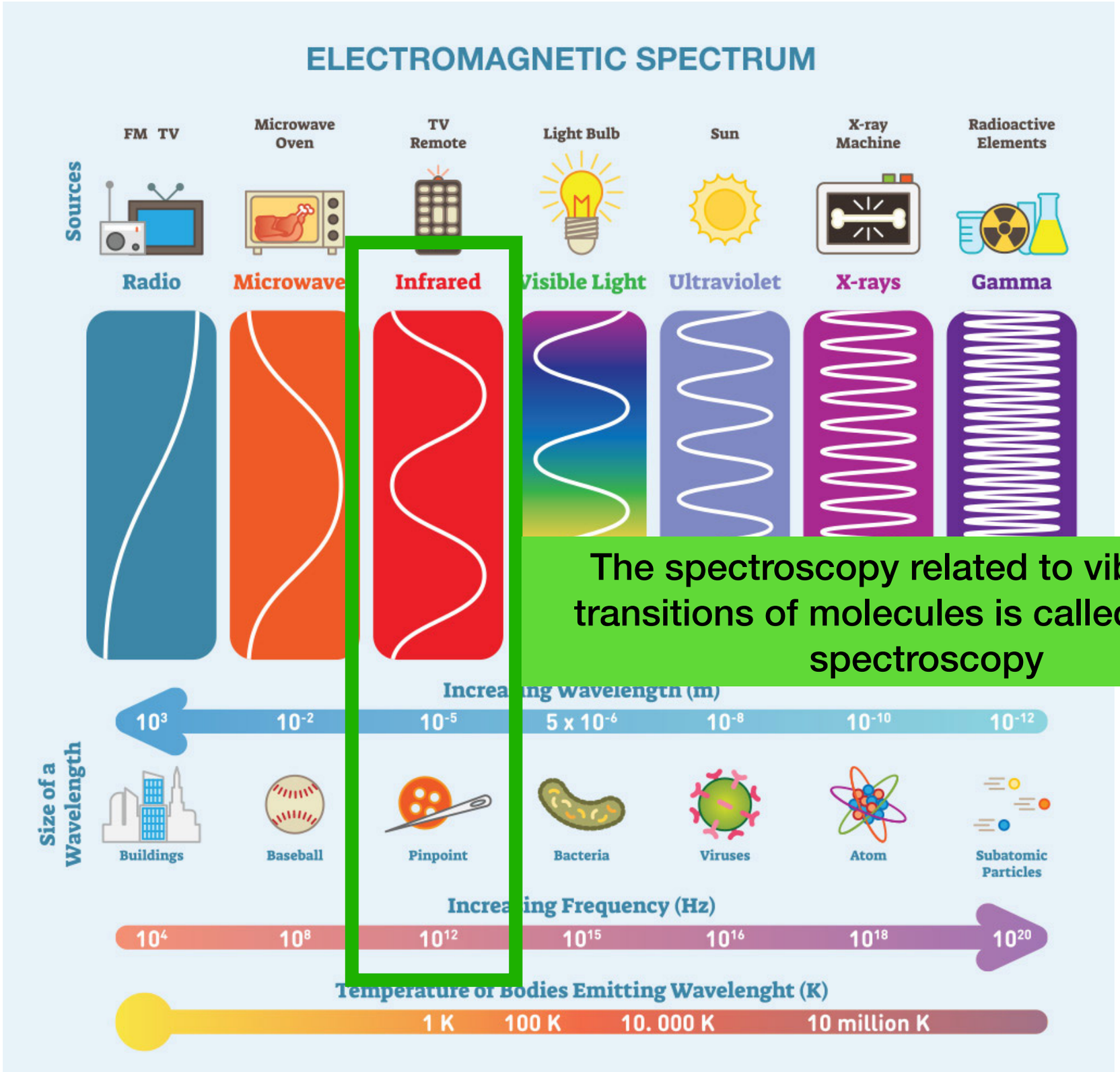
Today's class:

Infrared Spectroscopy and Mass spectrometry

This lecture follows the materials from the following books

- *Physical Chemistry for Life Sciences, by PW Atkins and JD Paula, Oxford, 2006*
- *Fundamentals of Biochemistry by Voet, Voet and Pratt, 5th Ed, Wiley*

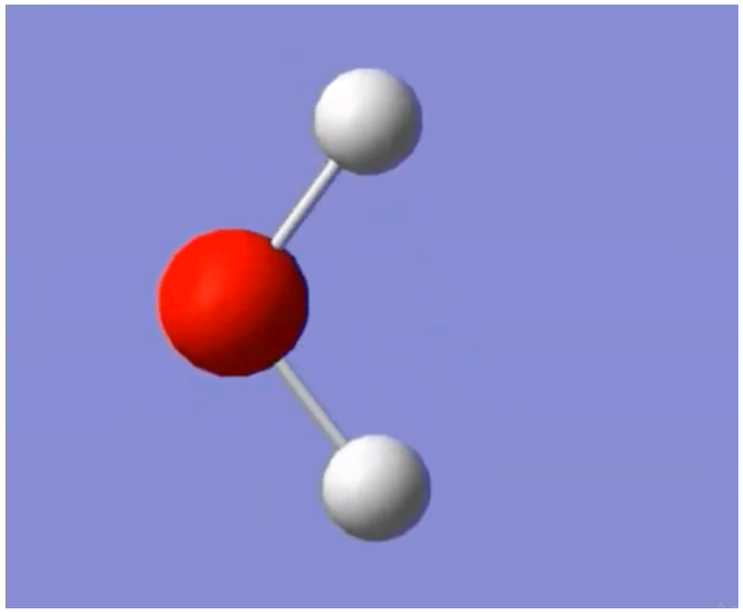
IR spectroscopy is the study of molecular vibrations

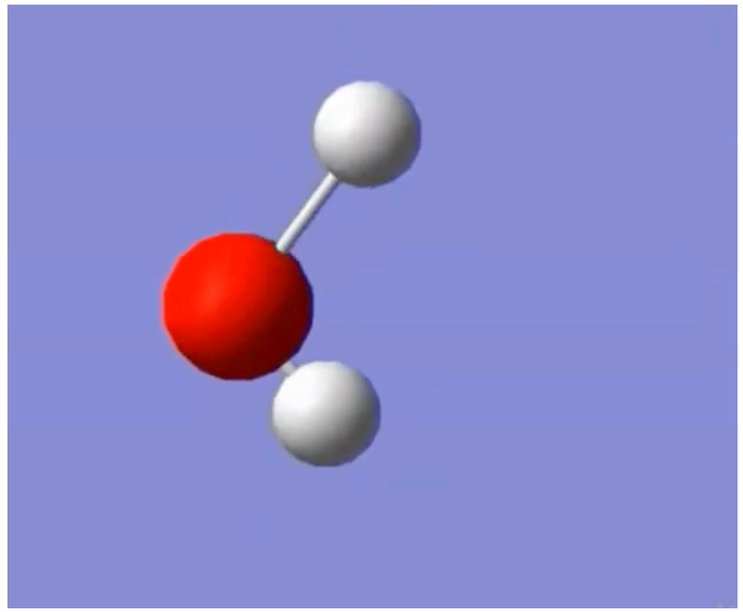


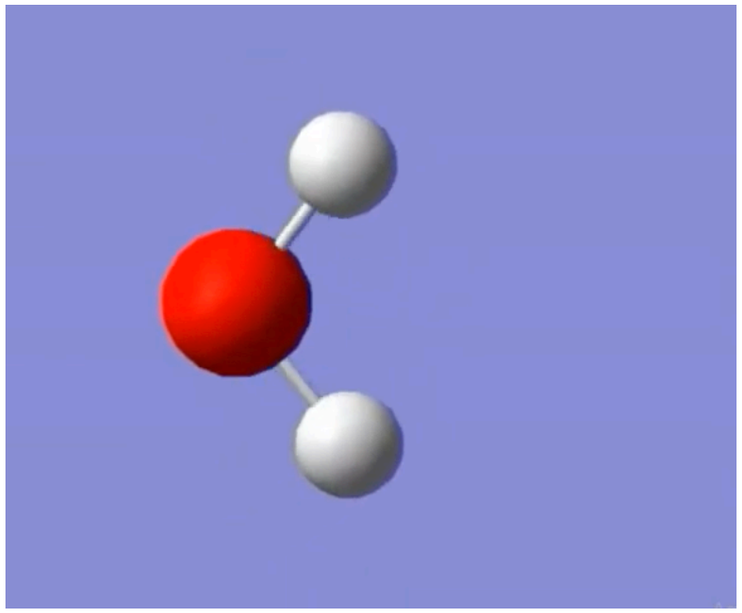
The spectroscopy related to vibrational transitions of molecules is called Infrared spectroscopy

Image courtesy: pasco.com

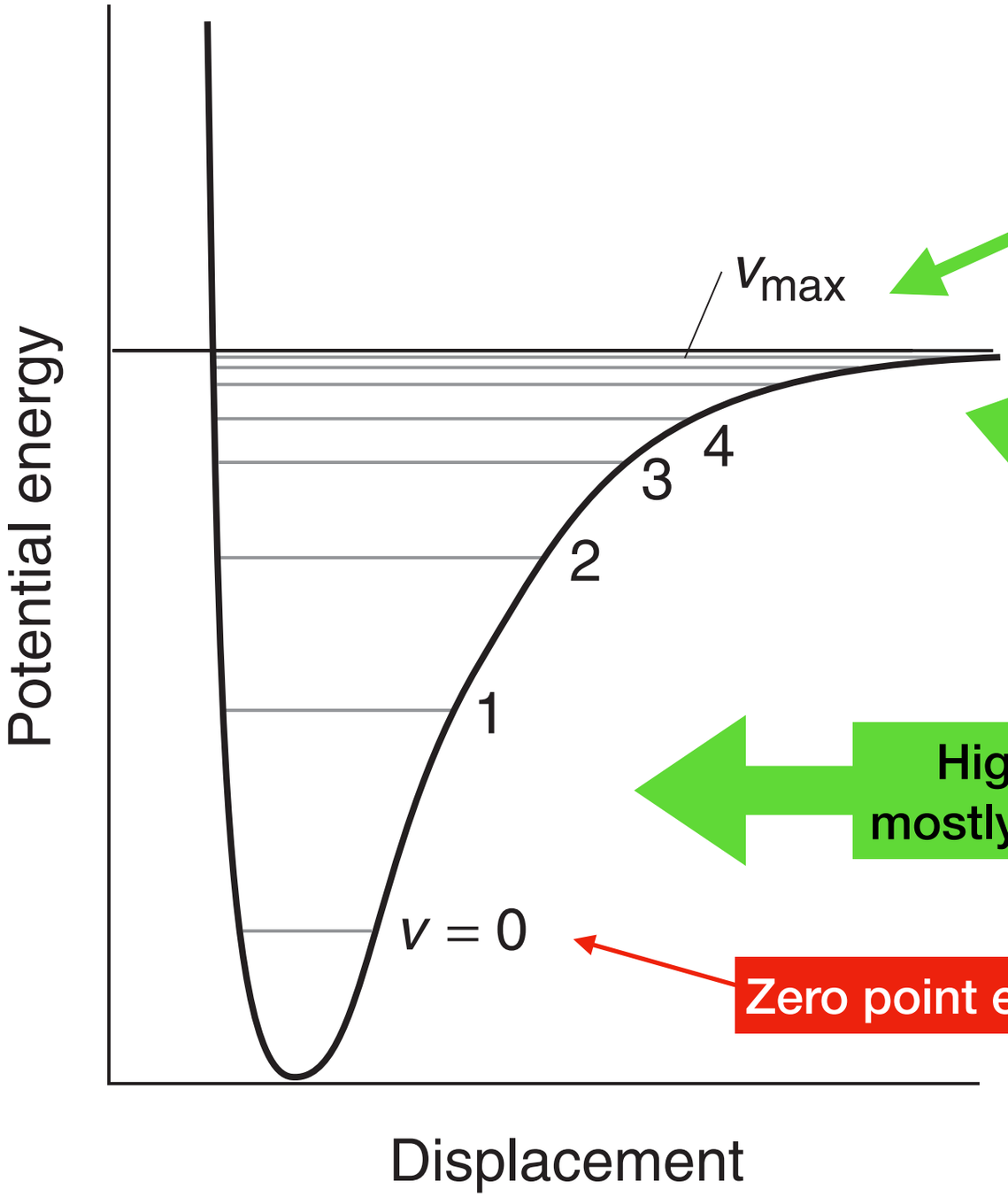
Molecular vibrations in water







Molecular vibrational energy levels



Highest frequency vibration beyond which the molecule dissociates

Lower energy excitation bands, mostly bending vibrational modes

High energy excitation bands, mostly stretching vibrational modes

Zero point energy

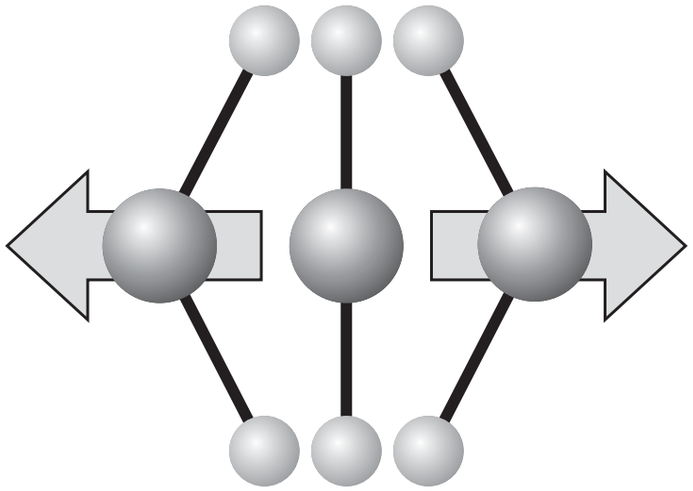
At room temperature, most molecules are at $v = 0$ state

Most important vibrational transition is the transition from $v = 0$ to $v = 1$ state

$\bar{\nu}_{0\rightarrow1}$ = fundamental wavenumber

$\nu_{0\rightarrow1}$ = fundamental frequency

Selection rules for IR spectroscopy



The gross selection rule for IR absorption spectra = the electric dipole moment of the molecule must change during the vibration.

If some vibrational mode of a molecule has no changing dipole moment that mode is IR inactive

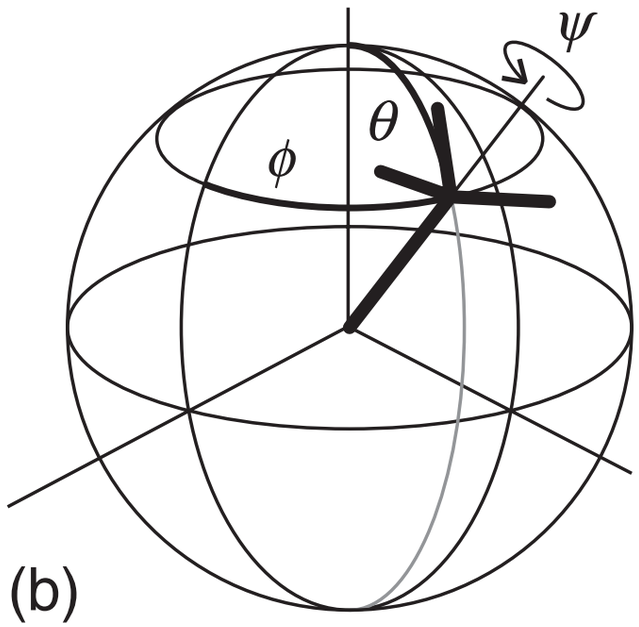
A homonuclear diatomic molecule is IR inactive

A heteronuclear diatomic molecule is IR active

The specific selection rule for infrared absorption spectra is

$$\Delta v = \pm 1$$

The vibrations of polyatomic molecules



For a non-linear molecule of N atoms,

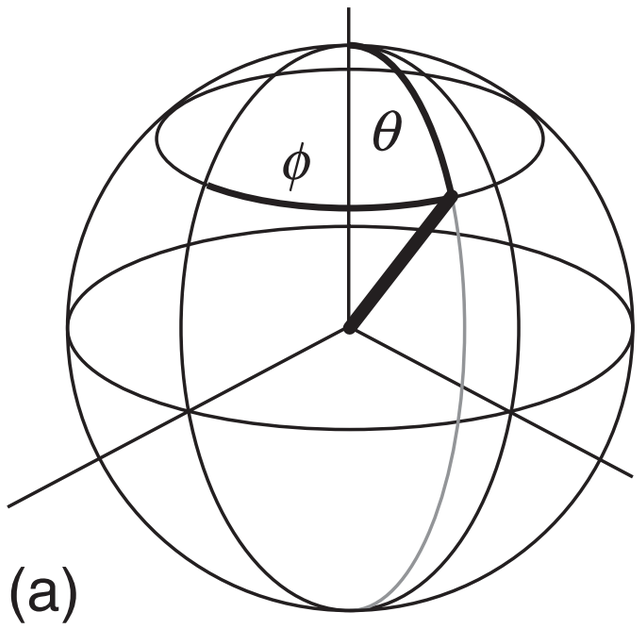
Total degrees of freedom = $3N$

How many are not vibrational modes?

Translational modes of the whole molecule = 3

Rotational modes of the whole molecule = 3

Vibrational modes = $3N - 6$

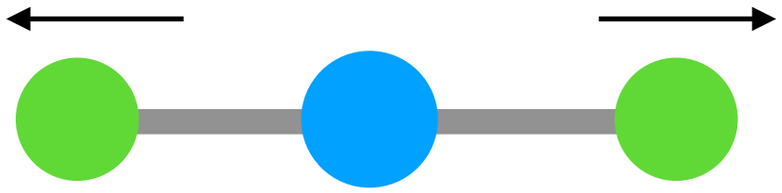


For linear molecules there is one less rotational mode

So, vibrational modes for linear molecules = $3N - 5$

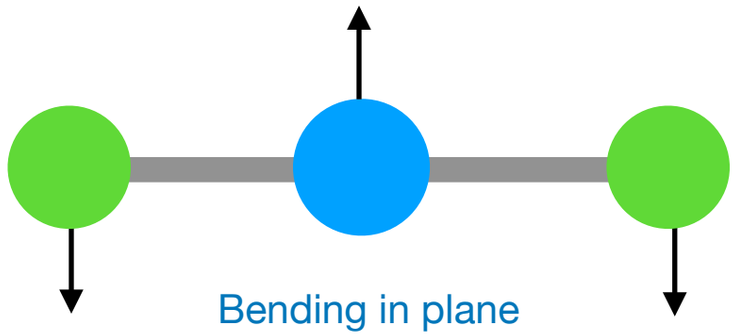
The vibrational modes of carbon dioxide

No of vibrational modes = $3N - 5 = 3 \times 3 - 5 = 4$



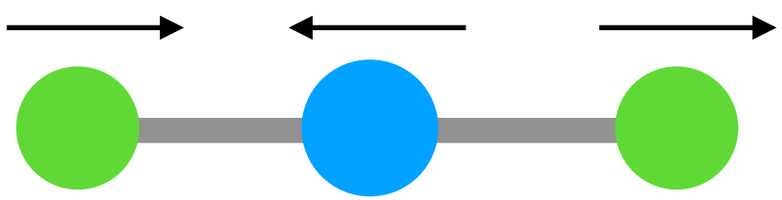
Symmetric stretching

IR inactive



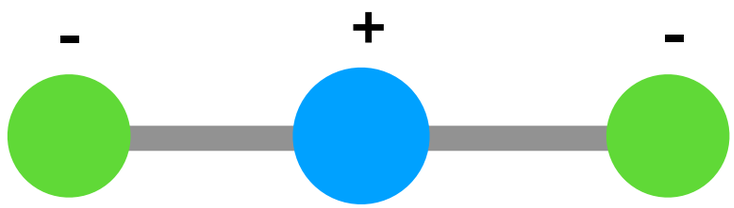
Bending in plane

IR active



Antisymmetric stretching

IR active

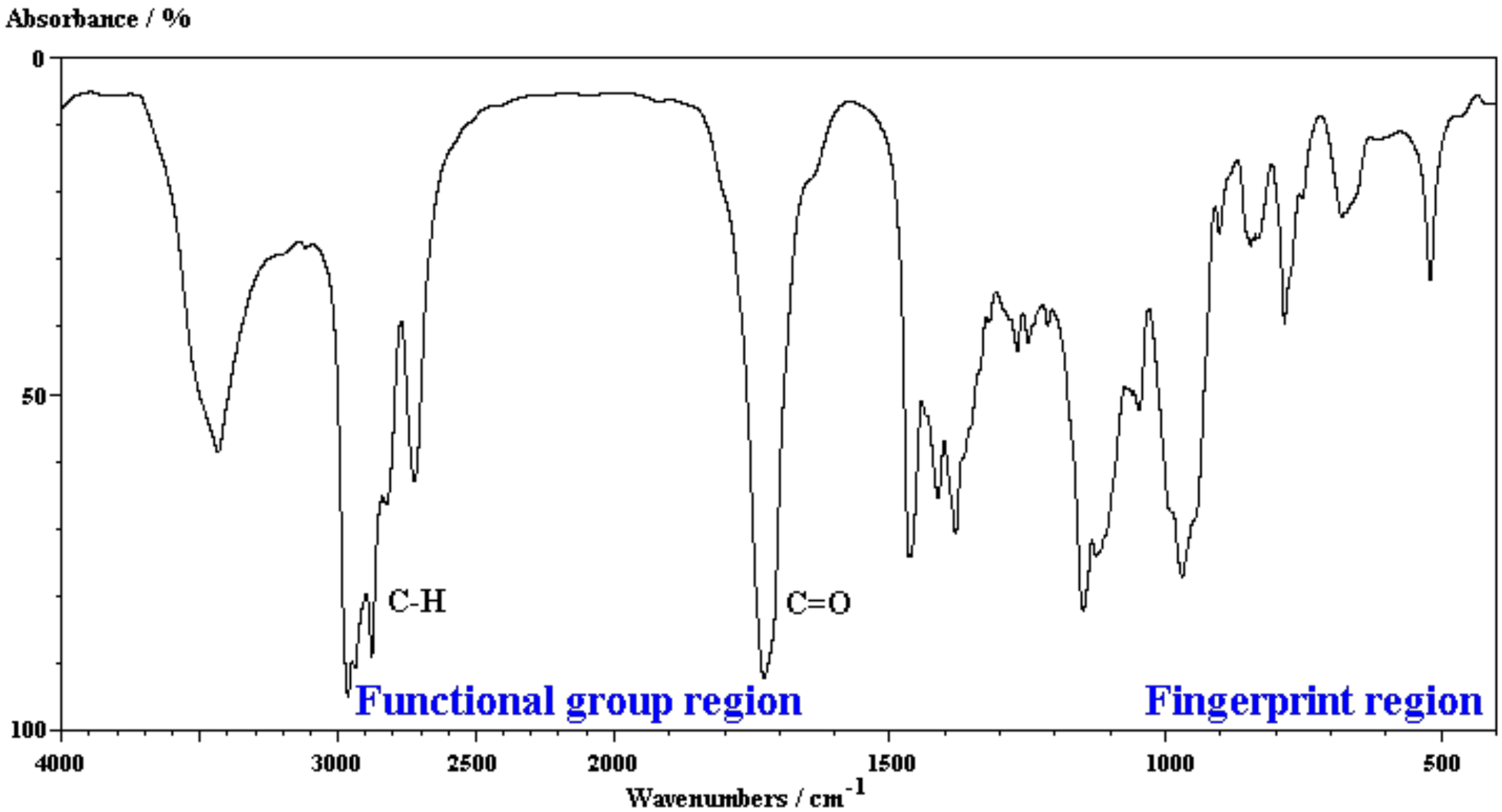


Bending out of plane

‘+’ = front on the screen move
‘-‘ = behind on the screen move

IR active

Analyzing an IR spectrum



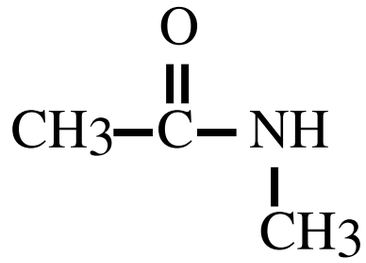
- Typically $\bar{\nu} \sim$ 4000-1600 cm^{-1} known as the functional group region
- $\bar{\nu} \lesssim$ 1600 cm^{-1} known as the fingerprint region = IR signature characteristic of a molecule
- Presence of a molecule can be confirmed by the fingerprint region
- Unknown molecules can be identified by characteristic vibrations of functional groups that occur outside the fingerprint region.

Typical vibrational wavenumbers

Vibration type	$\tilde{\nu}$/cm^{-1}
C—H	2850–2960
C—H	1340–1465
C—C stretch, bend	700–1250
C═C stretch	1620–1680
C≡C stretch	2100–2260
O—H stretch	3590–3650
C═O stretch	1640–1780
C≡N stretch	2215–2275
N—H stretch	3200–3500
Hydrogen bonds	3200–3570

Vibration spectroscopy of proteins

IR spectrum of proteins mostly focus on signatures of secondary structures on the vibrations of the peptide link -CONH-



N-methylacetamide is often used as a model compound for peptide link

Functional group IR peaks

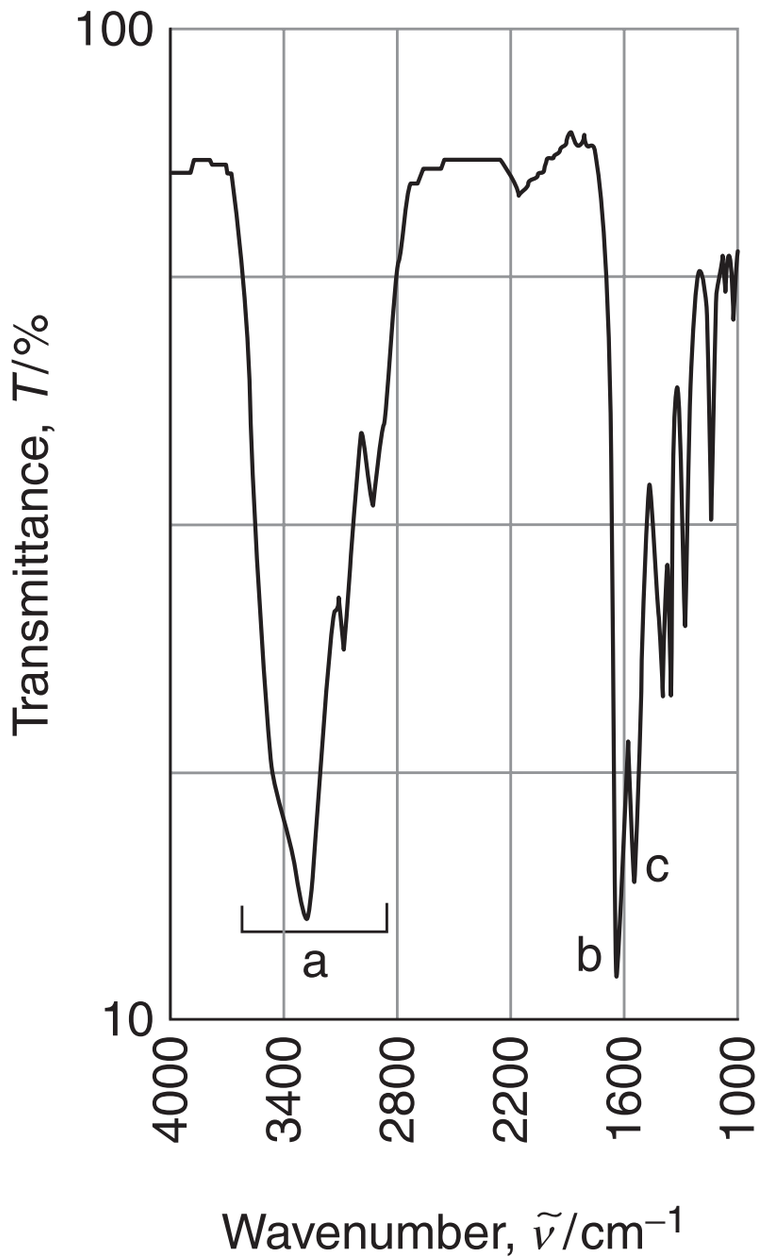
a : three bands above 2800 cm^{-1} (right to left)

Finger print region

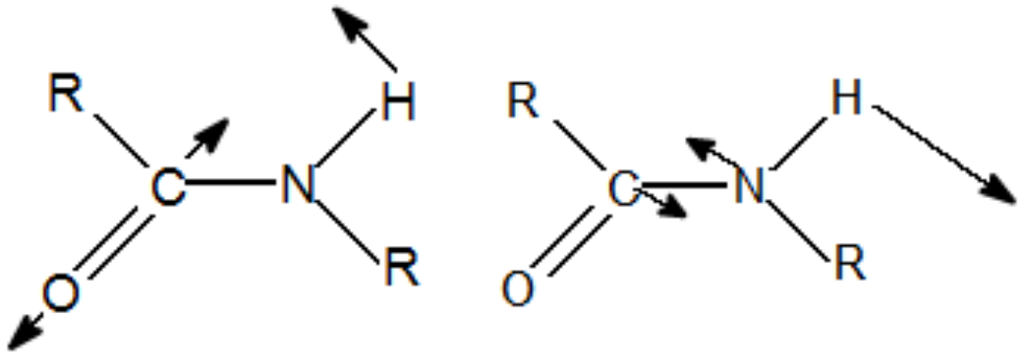
b : ~ 1650 cm^{-1} - amide I band

c : ~ 1600 cm^{-1} - amide II band

These two amide bands change significantly for different secondary structures



Amid I and II are the main protein fingerprint bands



amide I amide II

Range of amid I bands = 1600-1700 cm^{-1}

Range of amid II bands = 1510-1580 cm^{-1}

Vibration type	Vibrational wavenumber ($\tilde{\nu}$/cm^{-1}) for α Helix	β Sheet	Random coil
Amide I	1653	1640	1656
Amide II	1545	1525	1535

IR spectrum of whole cell

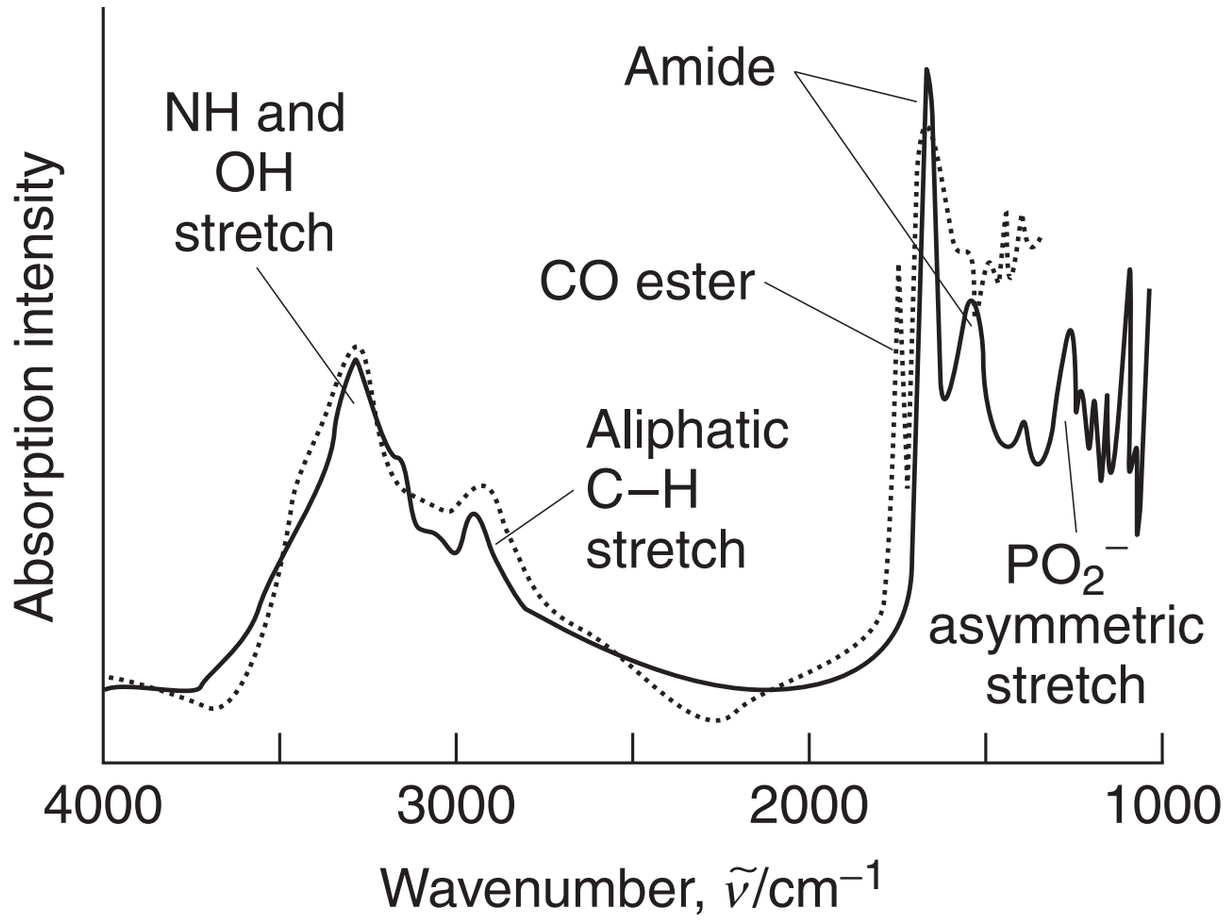
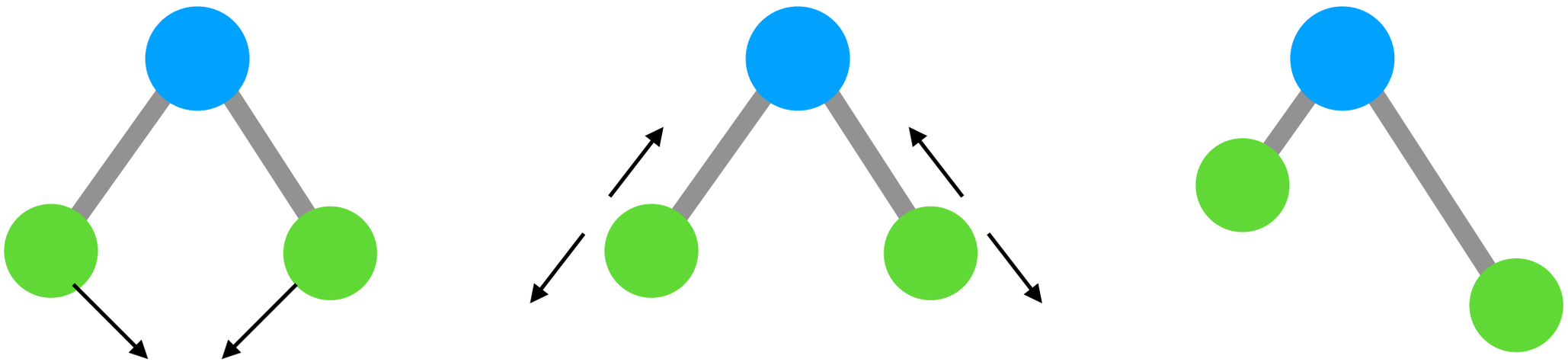


Fig. 13.26 Infrared absorption spectra of a single mouse cell: (solid line) living cell, (dotted line) dying cell. (Adapted from N. Jamin *et al.*, *Proc. Natl. Acad. Sci. USA* **95**, 4837 [1998].)

The dying cell shows an additional absorption at 1730 cm^{-1}, which is due to the ester carbonyl group from an unidentified compound.

Solvent selection for recording IR spectrum

No of vibrational modes in water = $3N - 6 = 3 \times 3 - 6 = 3$



Bending | Symmetric stretching | Antisymmetric stretching

Usually water is not chosen as solvent for IR spectroscopy

- Water has IR active vibration modes which show absorbance
- Water concentration in liquid water is very high ~55 M, so absorbance is very high
- Water's IR absorption dilutes the low intensity modes of solutes, like proteins

The most common solvents chosen for IR are
Carbon Tetrachloride (CCl_4), Carbon Disulfide (CS_2) and Chloroform ($CHCl_3$)

They are usually transparent in the important absorption regions of the spectrum

Mass spectrometry

What is mass spectrometry?

It is the most precise technique for the determination of molar mass

What does a mass spectrometer do?

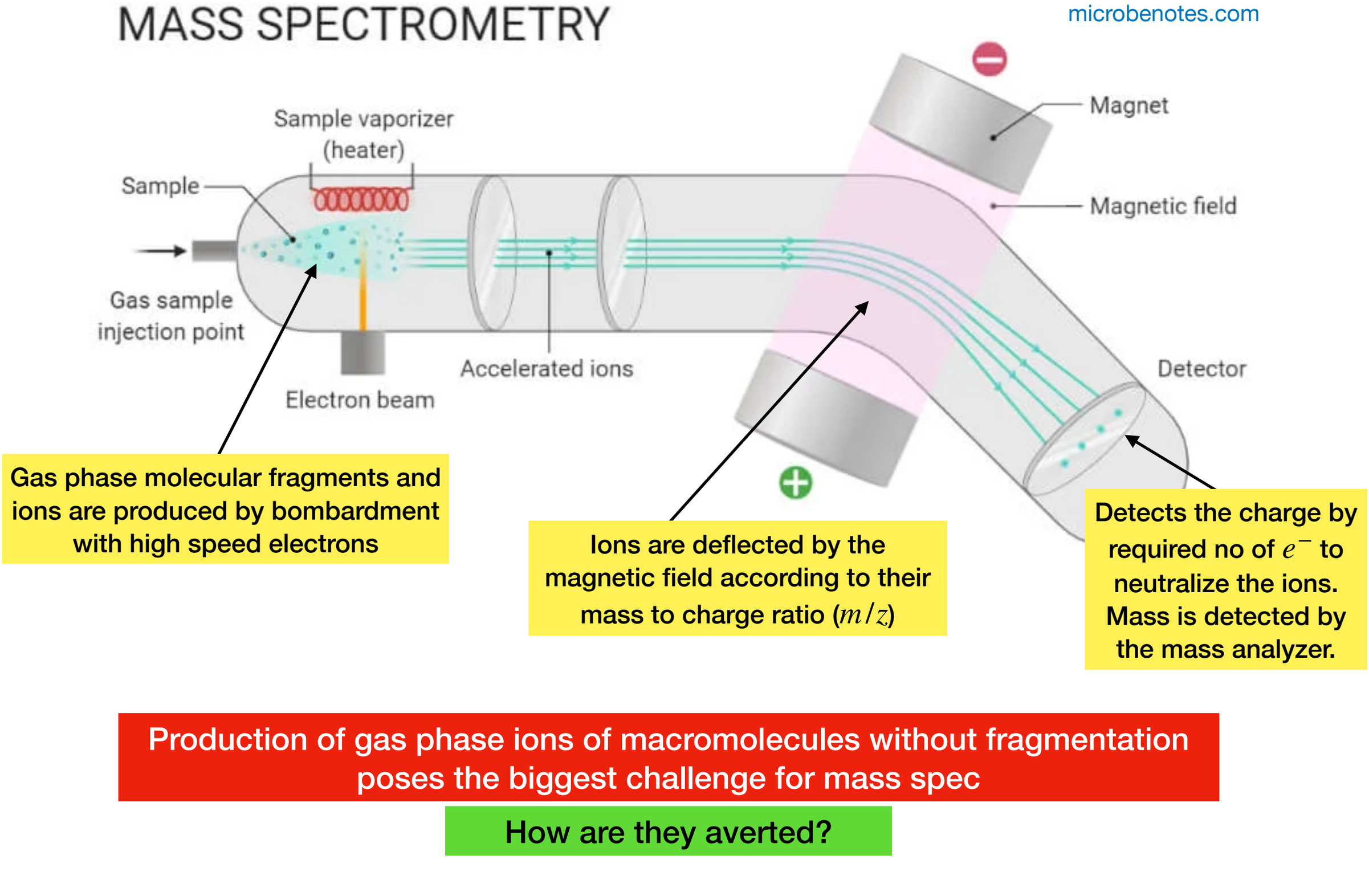
- It measures mass better than any other technique.
- It can give information about chemical structures or sequence.

What are the mass measurements used for?

To identify, verify, and quantitate:

- metabolites,
- recombinant proteins,
- proteins isolated from natural sources,
- oligonucleotides,
- drug candidates, peptides,
- synthetic organic chemicals, polymers

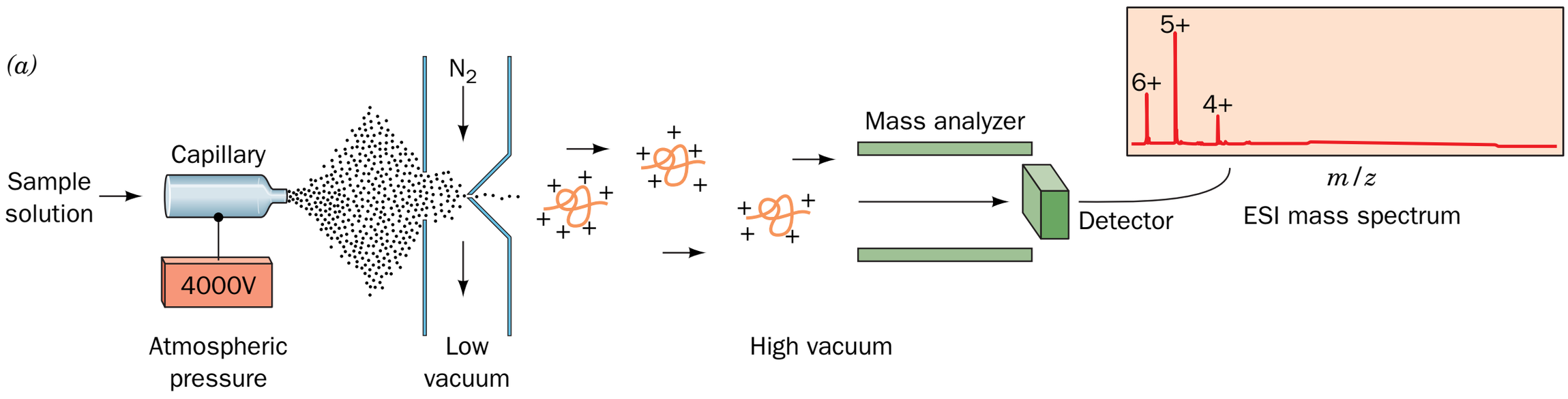
Basic principle of mass spectrometry



Production of gas phase ions of macromolecules without fragmentation poses the biggest challenge for mass spec

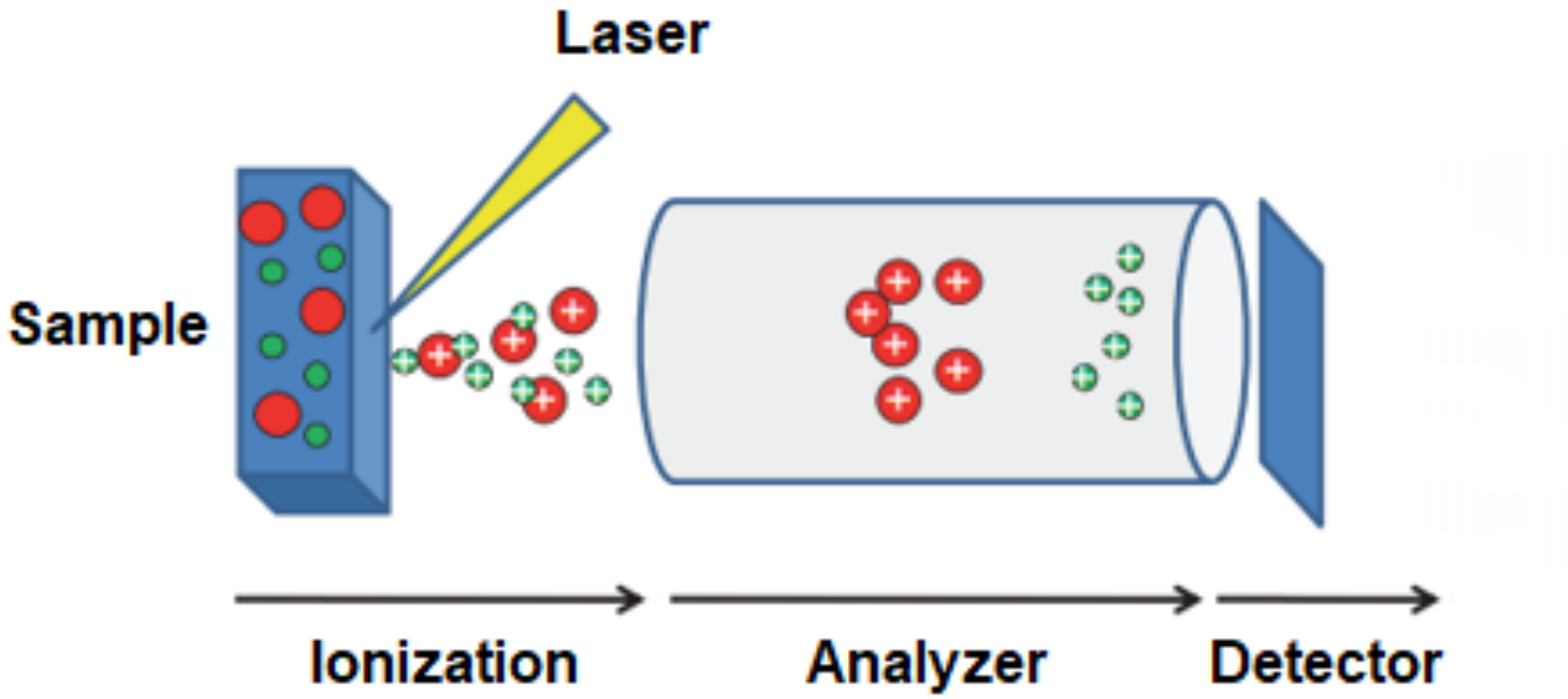
How are they averted?

Electro Spray Ionization (ESI) for biomacromolecules



- A protein solution is sprayed from a narrow capillary tube maintained at high voltage (~4000 V), forming fine, highly charged droplets
- Dry nitrogen gas promotes the evaporation of solvent from the droplets leaving gas-phase ions.
- The series of gas-phase macromolecular ions contains ionic charges in the range +0.5 to +2 per kD.
- The charges result from the protonation of basic side chains such as Arg and Lys.
- The ions are directed into the mass analyzer tube and the ion detector, which measures their m/z values with an accuracy of ~0.01% and generates the mass spectrum.
- The mass spectrum consists of a series of peaks corresponding to ions that differ by a single ionic charge and the mass of one proton.
- Determination of an ion's charge z permits its molecular mass m to be determined with far greater accuracy than by any other method.

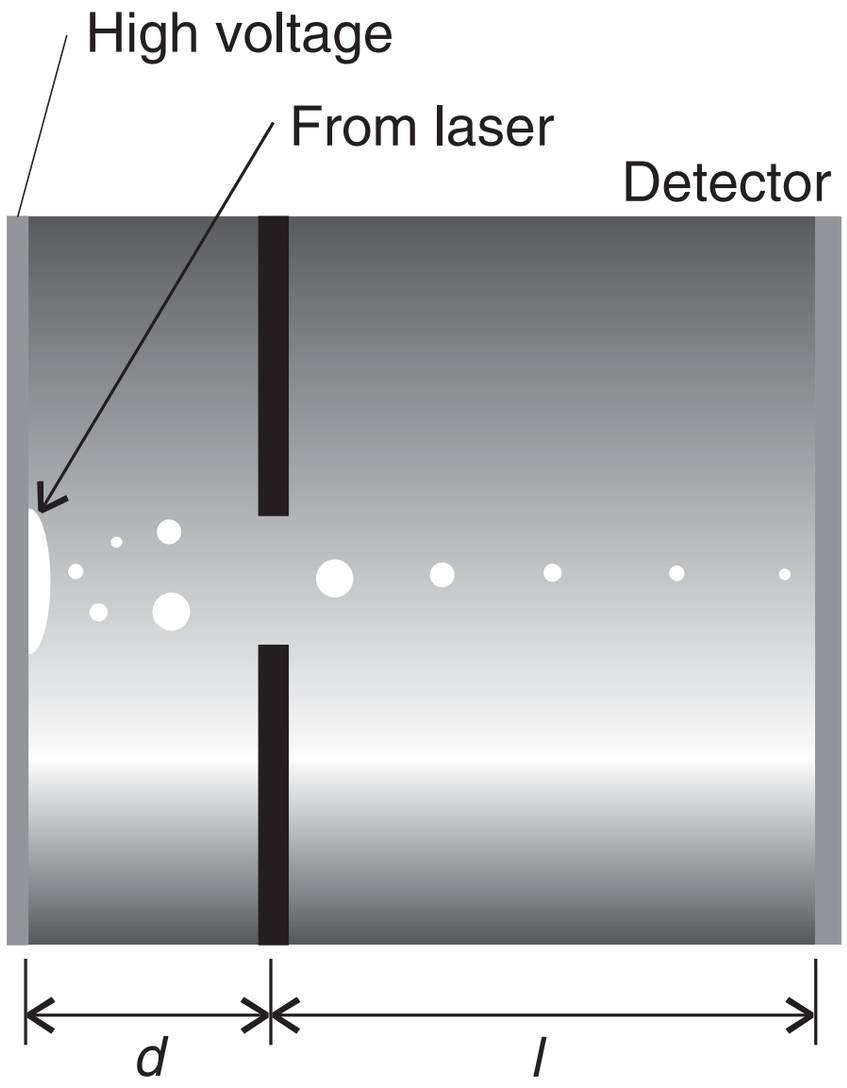
Matrix Assisted Laser Desorption Ionization (MALDI) for biomacromolecules



- The macromolecule is first embedded in a solid matrix that often consists of an organic acid such as 2,5-dihydroxybenzoic acid, nicotinic acid, or α-cyanocarboxylic acid.
- This sample is then irradiated with a laser pulse that ejects matrix ions, cations, and neutral macromolecules
- Thus a dense gas plume is created above the sample surface.
- The macromolecule is ionized by collisions and complexation with H^+ cations.

The MALDI-TOF mass spectrometer

MALDI technique coupled to a time-of-flight (TOF) mass analyzer/ion detector



Working principle of a TOF detector

- The ions are accelerated over a short distance d by an electrical field of strength E and then travel through a drift region of length l.
- For an ion of charge ze and mass m that is accelerated from rest we can write

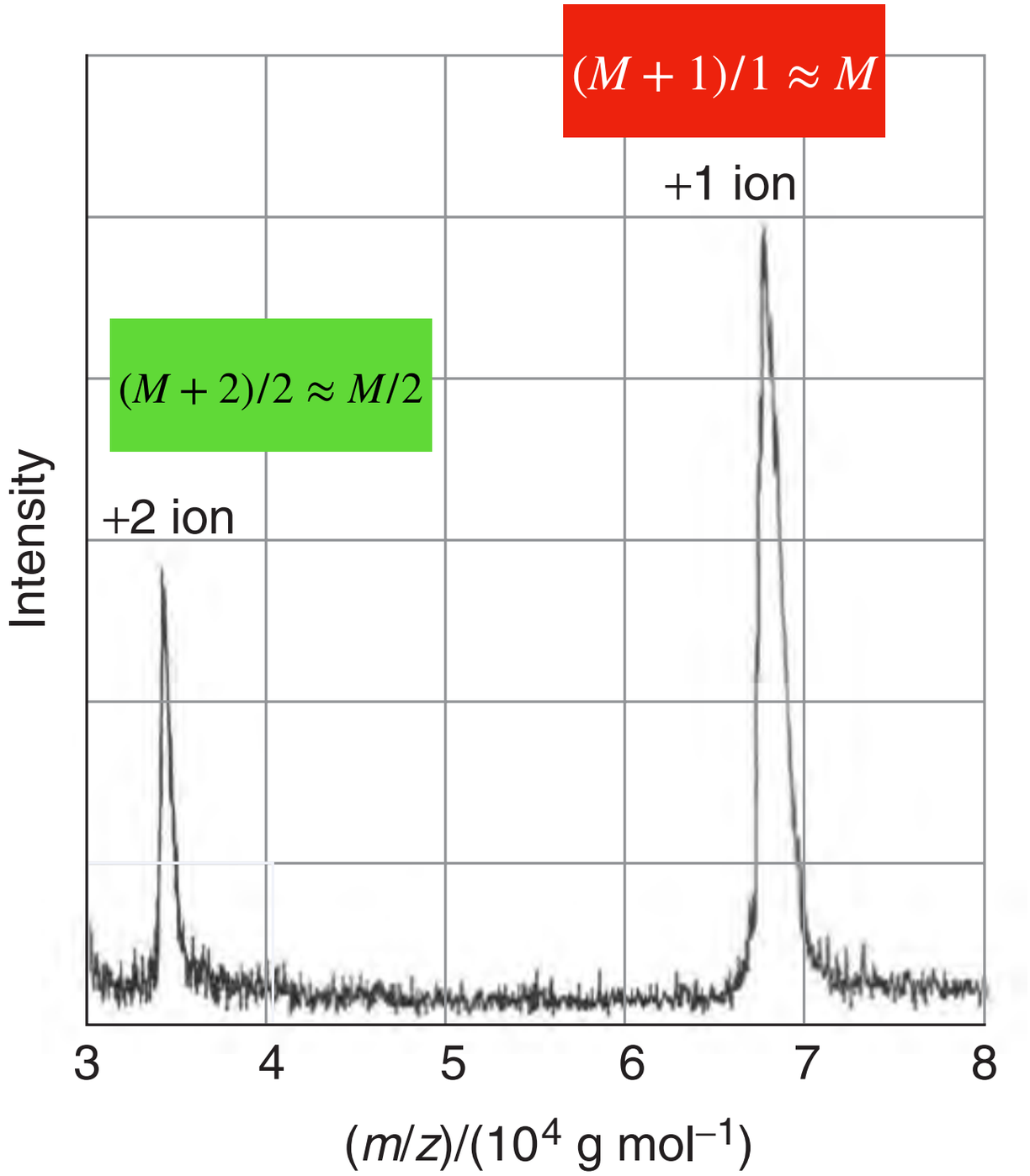
  Kinetic energy, $E_K = \frac{1}{2}mv^2 = zeEd$

  where, v = velocity of the ion
- Ignoring any acceleration in the flight tube since the length and time taken are very short, we can further get the time t taken by the ion to reach the detector

$$t = \frac{l}{v}, \implies \frac{m}{z} = 2eEd\left(\frac{t}{l}\right)^2$$

- Thus, t is a direct measure of m/z ratio for the ions.

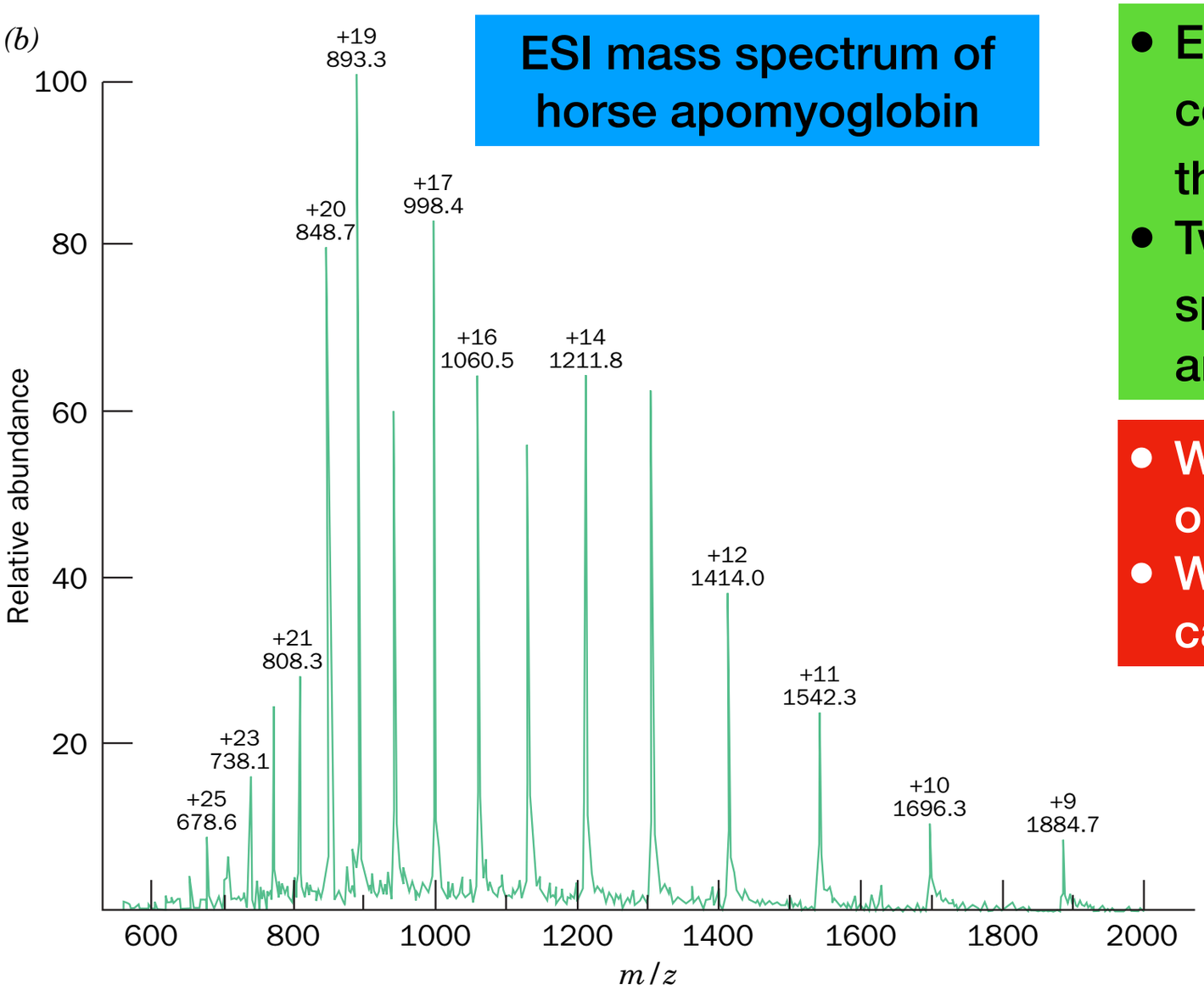
MALDI-TOF mass spec: example



The MALDI-TOF mass spectrum of bovine albumin, a protein with molar mass $m = 66430\ \text{g mol}^{-1}$

- **During the MALDI process, the protein takes up one or two protons with $z = 1\ \&\ 2$.**
- **Because the protein does not fragment, the $z = 2$ ion gives rise to a peak in the spectrum at a m/z that is one-half the value for the peak associated with the $z = 1$ ion.**

m/z to molar mass for biomacromolecules: example



ESI mass spectrum of horse apomyoglobin

- Each peak on the spectrum corresponds to the m/z ratio of an ion that looks like: $(M + nH)^{n+}$
- Two successive peaks in this mass spectrum have measured m/z of 1414.0 and 1542.3.

- What is the molecular mass of the original apomyoglobin molecule?
- What are the charges of the ions causing these peaks?

Peak 1: $p_1 = (\mathrm{M} + z)/z$

Peak 2: $p_2 = (\mathrm{M} + z - 1)/(z - 1)$

Solutions:

$$z = (p_2 - 1)/(p_2 - p_1)$$

$$M = \frac{(p_1 - 1)(p_2 - 1)}{p_2 - p_1}$$

$$\mathrm{M} = (1542.3 - 1)(1414.0 - 1)/(1542.3 - 1414.0) = 16,975\mathrm{D}$$

$$z = (1542.3 - 1)/(1542.3 - 1414.0) = +12$$