Organic Chemistry

with a Biological Emphasis

By
Tim Soderberg
University of Minnesota, Morris

UC San Diego Edition

Edited with permission from the author by
Valerie A. Schmidt & Haim Weizman

AY23-24

This work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. https://creativecommons.org/licenses/by-nc-sa/4.0/

Chapter 4

Structure determination, part I:

Mass spectrometry

(Photo credit: https://www.flickr.com/photos/vamapaull/)

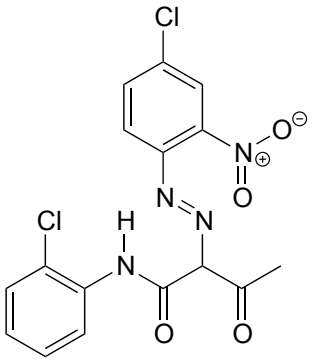
In the end, it was a 'funky yellow color' that led to the demise of Charles Heller's not-so-illustrious career in the world of collectable art.

William Aiken Walker was a 19th-century 'genre' painter, known for his small scenes of sharecroppers working the fields in the post-Civil War south. For much of his career, he traveled extensively, throughout the southern states but also to New York City and even as far as Cuba. He earned a decent living wherever he went by setting up shop on the sidewalk and selling his paintings to tourists, usually for a few dollars each. While he never became a household name in the

art world, he was prolific and popular, and his paintings are today considered collectible, often selling for upwards of ten thousand dollars.

In August 1994, Robert Hicklin, an art gallery owner in Charleston, South Carolina, was appraising a Walker painting brought to him by another South Carolina art dealer named Rick Simons. Hicklin's years of experience with Walker paintings told him that something just wasn't right with this one - he was particularly bothered by one of the pigments used, which he later described in a story in the *Maine Antique Digest* as a 'funky yellow color'. Reluctantly, he told Simons that it almost certainly was a fake.

Hoping that Hicklin was wrong, Simons decided to submit his painting to other experts for analysis, and eventually it ended up in the laboratory of James Martin, whose company *Orion Analytical* specializes in forensic materials analysis. Using a technique called infrared spectroscopy, Martin was able to positively identify the suspicious yellow pigment as an organic compound called 'Pigment Yellow 3'.



pigment yellow 3

As it turns out, Pigment Yellow 3 had not become available in the United States until many years after William Aiken Walker died.

Simons had purchased his painting from a man named Robert Heller for $9,500. When Heller approached him again to offer several more Walker paintings for sale, Simons contacted the FBI. A few days later, with FBI agents listening in, Simons agreed to buy two more Walker paintings. When he received them, they were promptly analyzed and found to be fake. Heller, who turned out to be a convicted felon, was arrested and eventually imprisoned.

(You may want to check out a short article from the September 10, 2007 issue of *Chemical and Engineering* News for more about this story).

In the first three chapters of this text, we have focused our efforts on learning about the structure of organic compounds. Now that we know what organic molecules look like, we can begin to address, in the next two chapters, the

question of *how* we get this knowledge in the first place. How are chemists able to draw with confidence the bonding arrangements in organic molecules, even simple ones such as acetone or ethanol? How was James Martin at Orion Analytical able to identify the chemical structure of the pigment compound responsible for the 'funky yellow color' in the forged William Aiken Walker painting?

This chapter is devoted to important techniques used by chemists to learn about the structures of organic molecules. First, we will learn how **mass spectrometry** can provide us with information about the mass of a molecule as well as the mass of fragments into which the molecule has been broken. Then, we will begin our investigation of **molecular spectroscopy**, which is the study of how electromagnetic radiation at different wavelengths interacts in different ways with molecules - and how these interactions can be quantified, analyzed, and interpreted to gain information about molecular structure

Looking ahead, Chapter 5 will be devoted to **nuclear magnetic resonance (NMR) spectroscopy**, where we use ultra-strong magnets and radio frequency radiation to learn about the electronic environment of individual atoms in a molecule and use this information to determine the atom-to-atom bonding arrangement. For most organic chemists, NMR is one of the most powerful analytical tools available in terms of the wealth of detailed information it can provide about the structure of a molecule.

In summary, the structure determination techniques we will be studying in this chapter and the next primarily attempt to address the following questions about an organic molecule:

<u>Chapter 4:</u>

Mass spectrometry (MS): *What is the atomic weight of the molecule and its common fragments?*

<u>Chapter 5:</u>

Nuclear magnetic resonance spectroscopy (NMR): *What is the overall bonding framework of the molecule?*

Learning Outcomes

- Use MS to obtain molecular mass/weight.
- Given a mass spectrum, recognize the molecular ion peak along with M+1 and M+2 peaks and be able to explain what these are and why they are there.
- Recognize a peptide sequence based on a given MS
- Describe, in 'generic' terms, a molecular spectroscopy experiment, and understand its elements: a spectrum of wavelengths passing into a sample, a quantum transition from a ground to an excited state in the sample molecules with absorbance of corresponding wavelength(s) of radiation, and detection of wavelengths absorbed along with the intensity of each absorbance.

Section 4.1: Mass Spectrometry

Mass spectrometry (MS) is a powerful analytical technique widely used by chemists, biologists, medical researchers, and environmental and forensic scientists, among others. With MS, we are looking at the mass of a molecule, or of different fragments of that molecule.

4.1A: Overview of mass spectrometry

There are many different types of MS instruments, but they all have the same three essential components:

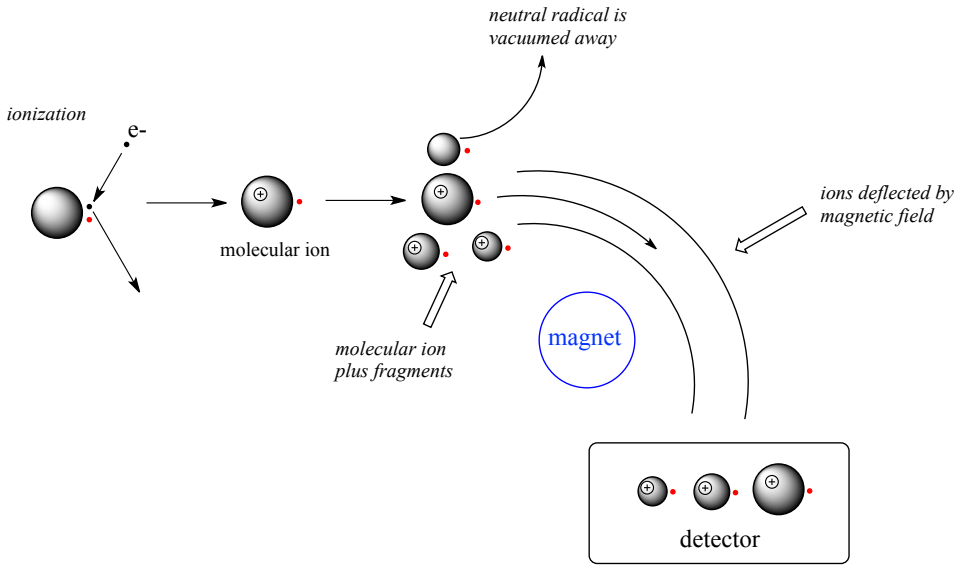
1) First, there is an ionization source, where the molecule is given a positive electrical charge, either by removing an electron or by adding a proton.

2) Depending on the ionization method used, the ionized molecule may or may not break apart into a population of smaller fragments

3) Next in line there is a mass analyzer, where the positively-charged fragments are separated according to mass.

4) Finally, there is a detector, which detects and quantifies the separated ions.

One of the more common types of MS techniques used in the organic laboratory is **electron ionization** (EIMS). Although biomolecules are usually too polar for this technique, it is useful for the analysis of intermediates and products in laboratory synthesis and is a good starting place to learn about MS in general. In a typical EIMS experiment, the sample molecule is bombarded by a high-energy electron beam, which has the effect of knocking a valence electron off the molecule to form a **radical cation**. Because a great deal of energy is transferred by this bombardment process, the radical cation quickly begins to break up into smaller **fragments**, some of which are cations (both radical and non-radical) and some of which are neutral. A subset of the initially formed radical cations do not fragment at all - these are referred to as **molecular ions**. The neutral fragments are either adsorbed onto the walls of the chamber or are removed by a vacuum source. The molecular ion and all other cationic fragments are accelerated down a curved tube by an electric field.

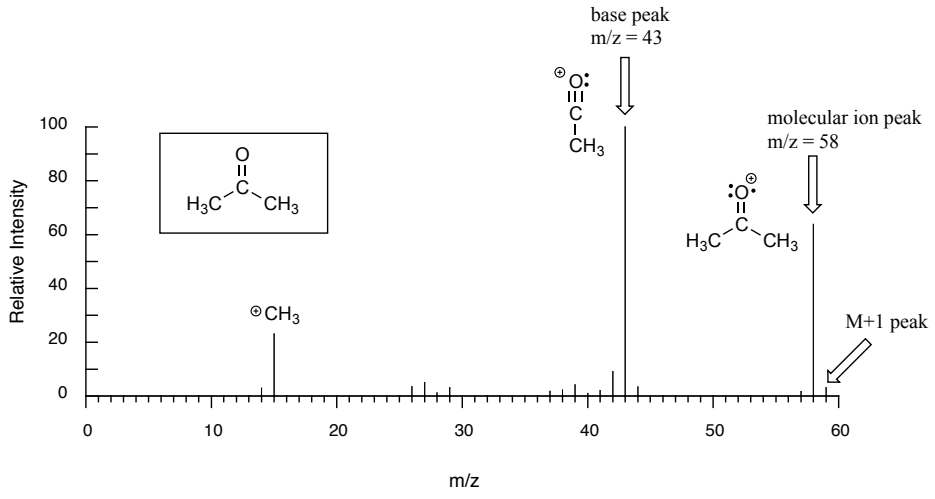


another image

As they travel down the curved tube, the ions are deflected by a strong magnetic field. Cations of different **mass to charge (*m/z*) ratios** are deflected to a different extent, resulting in a sorting of ions by mass (virtually all ions have charges of *z* = +1, so sorting by the mass to charge ratio is essentially the same thing as sorting by mass). A detector at the end of the curved flight tube records and quantifies the sorted ions.

4.1B: Looking at mass spectra

Below is typical output for an electron-ionization MS experiment (MS data in the section is derived from the Spectral Database for Organic Compounds, a free, web-based database of spectra.)

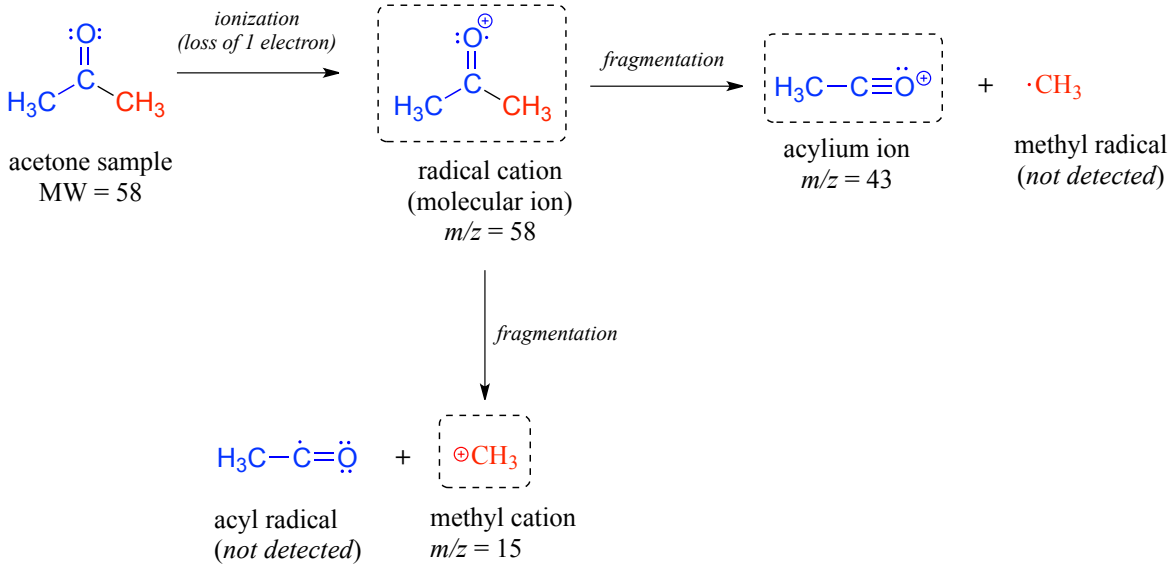


The sample is acetone. On the horizontal axis is the value for *m/z* (as we stated above, the charge *z* is almost always +1, so in practice this is the same as mass).

On the vertical axis is the relative abundance of each ion detected. On this scale, the most abundant ion, called the **base peak**, is set to 100%, and all other peaks are recorded relative to this value. In the acetone spectrum above, the base peak is at *m/z* = 43, representing a fragment of acetone with a mass of 43 amu. The molecular weight of acetone is 58, so we can identify the peak at *m/z* = 58 as that corresponding to the **molecular ion peak**, or **parent peak**. In some mass spectra, the molecular ion peak is the most abundant peak, and so it is also the base peak.

The data collected in a mass spectrum - a series of m/z values, each associated with a relative abundance value - is a unique 'fingerprint' of the compound being analyzed, so the spectrum of an unknown sample can be used to search a computer database of MS spectra to produce a 'hit', just as fingerprint information can be used by the police to identify a suspect from a fingerprint database. In fact, mass spectrometry is a powerful tool in criminal forensics.

An experienced chemist can gain a wealth of useful structural information from a mass spectrum, even without access to a reference database. The parent peak provides the molecular weight of the compound being analyzed, which of course is an especially useful piece of information. However, much of the utility in electron-ionization MS comes from the fact that the radical cations generated in the electron-bombardment process tend to fragment in predictable ways. The base peak at *m/z* = 43 in the acetone spectrum, for example, is the result of the molecule breaking apart into what is termed an acylium cation and a methyl radical. The methyl radical fragment has a mass of 15 amu, but is not detected because it is neutral, not a cation.



Notice that we do in fact see a peak in the spectrum at m/z = 15, which corresponds to a methyl cation, formed in a different fragmentation event in which the second product is a neutral (and thus undetected) radical fragment.

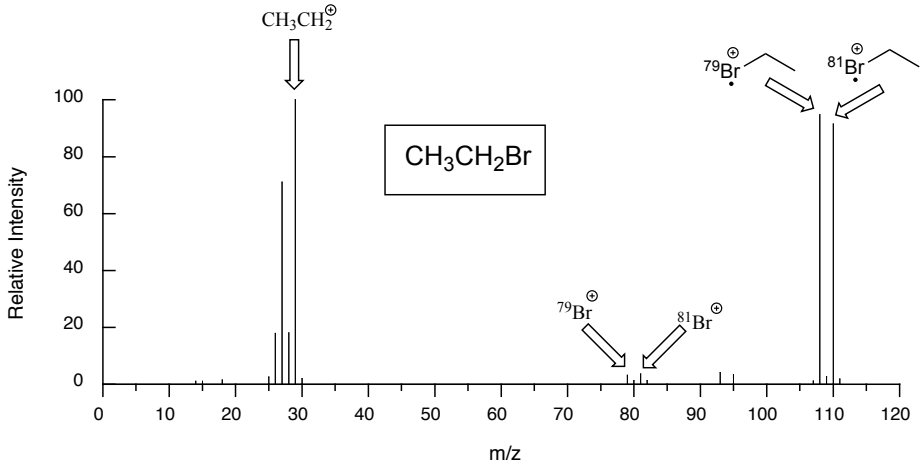
Exercise 4.1: Using the fragmentation patterns for acetone above as a guide, predict the signals that you would find in the mass spectra of:

a) 2-butanone; b) 3-hexanone; c) cyclopentanone.

Many other organic functional groups and bonding arrangements have predictable fragmentation patterns. As a rule, the cationic fragments that form in greatest abundance are those which are, relatively speaking, the most stable (we will learn about carbocation stability in Chapter 8). A discussion of the mechanisms for fragmentation in mass spectrometry is outside the scope of this book but will be covered if you take more advanced courses in organic or analytical chemistry.

Notice in the mass spectrum of acetone that there is a small peak at *m/z* = 59: this is referred to as the **M+1 peak**. How can there be an ion that has a greater mass than the molecular ion? A small fraction - about 1.1% - of all carbon atoms in nature are the ^{13}C rather than the ^{12}C isotope. The ^{13}C isotope has an extra neutron in its nucleus, and thus is heavier than ^{12}C by 1 mass unit. The M+1 peak corresponds to those few acetone molecules in the sample which contained a ^{13}C.

Molecules with several oxygen atoms sometimes show a small **M+2 peak** (2 *m/z* units greater than the parent peak) in their mass spectra, due to the presence of a small amount of ^{18}O (the most abundant isotope of oxygen is ^{16}O). Because there are two abundant isotopes of both chlorine (about 75% ^{35}Cl and 25% ^{37}Cl) and bromine (about 50% ^{79}Br and 50% ^{81}Br), chlorinated and brominated compounds have very large and recognizable M+2 peaks. Fragments containing both isotopes of bromine can be seen in the mass spectrum of bromoethane: notice that the equal abundance of the two isotopes is reflected in the equal height of their corresponding peaks.



The spectrum for chloroethane shows a parent peak at *m/z* = 64, and an M+2 peak at 1/3 the height of the parent peak. These characteristic M+2 patterns are useful for identifying the presence of bromine or chlorine in a sample.

Exercise 4.2: Predict some signals that you would expect to see in a mass spectrum of 2-chloropropane.

Exercise 4.3: The mass spectrum of an aldehyde shows a parent peak at *m/z* = 58 and a base peak at *m/z* = 29. Propose a structure and identify the two species whose *m/z* values were listed.

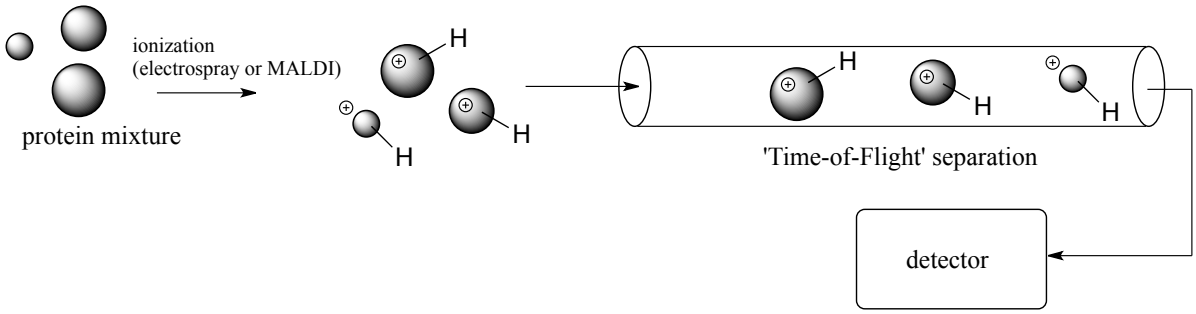
4.1D: Mass spectrometry of biomolecules

As was mentioned earlier, electron ionization mass spectrometry is generally not very useful for analyzing biomolecules: their high polarity makes it difficult to get them into the vapor phase, the first step in EIMS. Mass spectrometry of biomolecules has undergone a revolution over the past few decades, with many new ionization and separation techniques being developed. The strategy for biomolecule analysis involves **soft ionization**, in which much less energy (compared to techniques such as EIMS) is imparted to the molecule being analyzed during the ionization process. Usually, soft ionization involves adding protons rather than removing electrons: the cations formed in this way are significantly less energetic than the radical cations formed by removal of an electron. The result of soft ionization is that little or no fragmentation occurs, so the mass being measured is that of an intact molecule. Typically, large biomolecules are digested into smaller pieces using chemical or enzymatic methods, then their masses are determined by 'soft' MS.

New developments in soft ionization MS technology have made it easier to detect and identify proteins that are present in very small quantities in biological

samples. **In electrospray ionization** (ESI), the protein sample, in solution, is sprayed into a tube and the molecules are induced by an electric field to pick up extra protons from the solvent. Another common 'soft ionization' method is 'matrix-assisted laser desorption ionization' (**MALDI**). Here, the protein sample is adsorbed onto a solid matrix, and protonation is achieved with a laser.

Typically, both electrospray ionization and MALDI are used in conjunction with a time-of-flight (TOF) mass analyzer component.



another schematic of MALDI-TOF

The proteins are accelerated by an electrode through a column, and separation is achieved because lighter ions travel at greater velocity than heavier ions with the same overall charge. In this way, the many proteins in a complex biological sample (such as blood plasma, urine, etc.) can be separated and their individual masses determined very accurately. Modern protein MS is extremely sensitive – recently, scientists were even able to detect the presence of *Tyrannosaurus rex* protein in a fossilized skeleton! (*Science* **2007**, *316*, 277).

Soft ionization mass spectrometry has become in recent years an increasingly valuable tool in the field of **proteomics**. Traditionally, protein biochemists tend to study the structure and function of individual proteins. Proteomics researchers, in contrast, want to learn more about how large numbers of proteins in a living system interact with each other, and how they respond to changes in the state of the organism. One important subfield of proteomics is the search for protein 'biomarkers' for human disease: in other words, proteins which are present in greater quantities in the tissues of a sick person than in a healthy person. Detection in a healthy person of a known biomarker for a disease such as diabetes or cancer could provide doctors with an early warning that the patient may be especially susceptible to the disease, so that preventive measures could be taken to prevent or delay onset.

In a 2005 study, MALDI-TOF mass spectrometry was used to compare fluid samples from lung transplant recipients who had suffered from tissue rejection to samples from recipients who had not suffered rejection. Three peptides (short proteins) were found to be present at elevated levels specifically in the tissue rejection samples. It is hoped that these peptides might serve as biomarkers to identify patients who are at increased risk of rejecting their transplanted lungs.

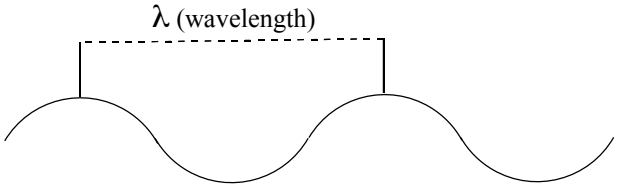
(*Proteomics* **2005**, *5*, 1705). More recently, MALDI-TOF MS was used to identify protein biomarkers for Alzheimer's disease in the blood, which could lead to an early diagnosis method for Alzheimer's which is much less expensive than brain scans, the only detection method currently available (*Nature* **2018**, *554*, 249).

Section 4.2: Review of molecular spectroscopy

4.2A: The electromagnetic spectrum

Electromagnetic radiation, as you may recall from a previous chemistry or physics class, is composed of oscillating electrical and magnetic fields. Visible light is electromagnetic radiation. So are the gamma rays that are emitted by spent nuclear fuel, the x-rays that a doctor uses to visualize your bones, the ultraviolet light that causes a painful sunburn when you forget to apply sun block, the infrared light that the army uses in night-vision goggles, the microwaves you use to heat up your frozen burritos, and the radio-frequency waves of your cell phone signal.

While the speed of a wave in the ocean can vary, the speed of electromagnetic waves – commonly referred to as the speed of light – is a constant, approximately 300 million meters per second, whether we are talking about gamma radiation, visible light, or FM radio waves. Electromagnetic radiation is defined by its **wavelength**, which is the distance between one wave crest to the next.



Because electromagnetic radiation travels at a constant speed, each wavelength corresponds to a given **frequency**, which is the number of times per second that a crest passes a given point. Longer waves have lower frequencies, and shorter waves have higher frequencies. Frequency is commonly reported in hertz (Hz), meaning 'cycles per second', or 'waves per second'. The equivalent standard unit for frequency is s^{-1}.

When talking about electromagnetic waves, we can refer either to wavelength or to frequency - the two values are inversely proportional:

equation 4.1 $$\lambda\nu = c$$

where λ (Greek '*lambda*') is wavelength, ν (Greek '*nu*') is frequency in s^{-1} (or Hz), and c is the speed of light, a constant value of 3.0×10^8 $m \cdot s^{-1}$.

Electromagnetic radiation transmits energy in discreet quantum 'packages' called **photons**. Shorter wavelengths (and higher frequencies) correspond to higher energy.

High energy radiation, such as gamma radiation and x-rays, is composed of very short waves – as short as 10^{-16} m. Longer wavelengths are far less energetic, and thus are less dangerous to living things. Visible light waves are in the range of 400 – 700 nm (nanometers, or 10^{-9} m), while radio waves can be several hundred meters in length.

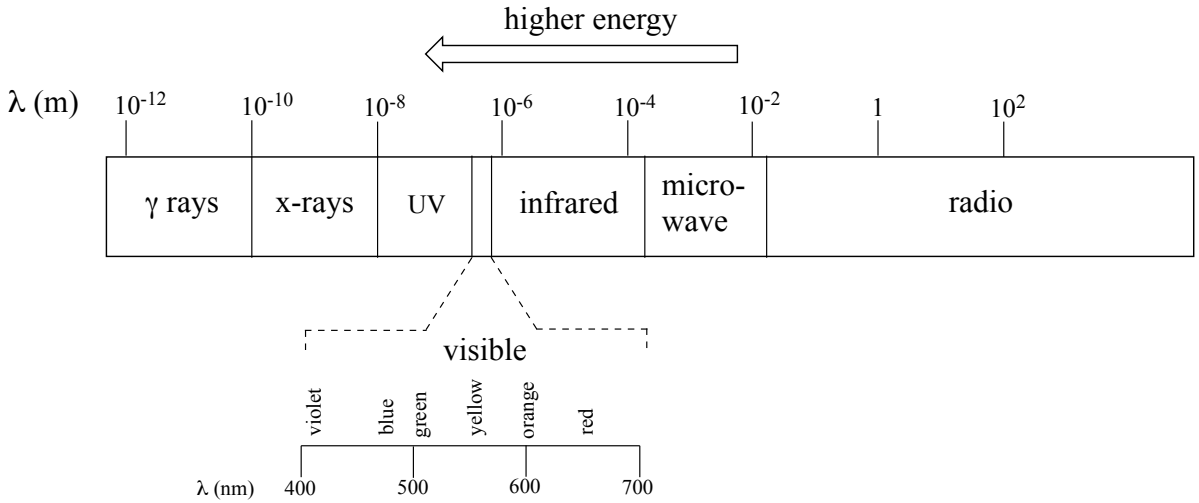
The energy of a particular wavelength of electromagnetic radiation can be expressed as:

equation 4.2 $$E = hc/\lambda = h\nu$$

where E is energy in kJ/mol of photons and *h* is 3.99 x 10^{-13} kJ·s·mol^{-1}, a number known as **Planck's constant**.

Using equations 4.1 and 4.2, we can calculate, for example, that visible red light with a wavelength of 700 nm has a frequency of 4.29 x 10^{14} Hz, and an energy of 171 kJ per mole of photons.

The full range of electromagnetic radiation wavelengths is referred to as the **electromagnetic spectrum**.



another image

Notice in the figure above that visible light takes up just a narrow band of the full spectrum. White light from the sun or a light bulb is a mixture of all the visible wavelengths plus some UV. You see the visible region of the electromagnetic

spectrum divided into its different wavelengths every time you see a rainbow: violet light has the shortest wavelength, and red light has the longest.

Exercise 4.4: Visible light has a wavelength range of about 400-700 nm. What is the corresponding frequency range? What is the corresponding energy range, in kJ/mol of photons?

4.2B: Overview of the molecular spectroscopy experiment

In a molecular spectroscopy experiment, electromagnetic radiation of a specified range of wavelengths is allowed to pass through a sample containing a compound of interest. The sample molecules absorb energy from some of the wavelengths, and as a result jump from a low energy 'ground state' to some higher energy 'excited state.' Other wavelengths are *not* absorbed by the sample molecule, so they pass on through. A detector records which wavelengths were absorbed, and to what extent they were absorbed.

General schematic of molecular spectroscopy

As we will see in this chapter, we can learn a lot about the structure of an organic molecule by quantifying how it absorbs (or does not absorb) different wavelengths in the electromagnetic spectrum. Three of the most useful types of molecular spectroscopy for organic chemists involve absorption of radiation in the infrared, ultraviolet/visible, and radio regions of the electromagnetic spectrum. We will focus first on infrared spectroscopy.

End of Chapter Self-Check List

Check your progress towards success. Verify that after completing this chapter you can:

Mass spectrometry:

- ☐ Explain the fundamental principles of an MS experiment.
- ☐ Recognize on a mass spectrum the molecular ion peak along with M+1 and M+2 peaks and be able to explain what these are.
- ☐ Recognize characteristic peaks for chloro- and bromo-alkanes (M+2 peaks in 1:1 or 3:1 ratios, respectively).

Spectrometry basics:

- ☐ Order the regions of the electromagnetic spectrum from shortest to longest wavelength: gamma radiation, x-rays, ultraviolet, blue visible light, red visible light, infrared, microwave, and radio wave.
- ☐ Convert mathematically among the following: wavelength expressed in meters, wavelength expressed in wavenumbers (cm^{-1}), frequency, and energy in kJ/mol of photons.
- ☐ Recognize qualitatively the relationships between wavelength, wavenumbers, frequency, and energy. Shorter wavelengths correspond to higher wavenumbers, higher frequency, and higher energy.
- ☐ Describe in 'generic' terms the elements of molecular spectroscopy experiment: a spectrum of wavelengths passing into a sample, a quantum transition from ground to an excited state in the sample molecules with absorbance of corresponding wavelength(s) of radiation, and detection of wavelengths absorbed along with their intensity.

If you didn't check off all items on this list, practice more and reach out to your instructional team for additional help.

Problems

P4.1: Which represents a higher energy frequency of electromagnetic radiation, 1690 cm^{-1} or 3400 cm^{-1}? Express each of these in terms of wavelength (meters) and frequency (Hz).

P4.2: One would expect the mass spectrum of cyclohexanone to show a molecular ion peak at m/z = 98. However, the m/z = 98 peak in the cyclohexanone spectrum is unusually abundant, compared to the molecular ion peaks in the mass spectra of other ketones such as 2-hexanone or 3-hexanone. Explain.