CHAPTER 1

Spectroscopy and the Proton NMR Experiment

1 WHAT IS THE STRUCTURE OF A MOLECULE?

There are several levels of understanding what a molecule "looks like" on the scale of individual atoms. The first step is to understand how many of each type of atom make up the collection of atoms that are bonded together to form a molecule. The *molecular formula* is an accounting of the types of atoms in a molecule and the number of each type of atom (*e.g.*, $C_6H_8N_2O_4$). **Mass spectrometry** is used to "weigh" molecules and obtain their exact mass, in atomic mass units (amu). Because atoms have masses that can differ slightly from integer values (*e.g.*, $^{1}H = 1.007825$ amu, $^{12}C = 12.000000$, $^{16}O = 15.994915$, $^{14}N = 14.003074$), a very precise measurement of the mass of a molecule allows us to determine the molecular formula. With a molecular formula, we can start to think about how this group of atoms is connected together. For example, for C_4H_6O (Figure 1.1) we can think of many ways to connect the atoms, while satisfying the valence rules (four bonds to C, two to O, one to H).

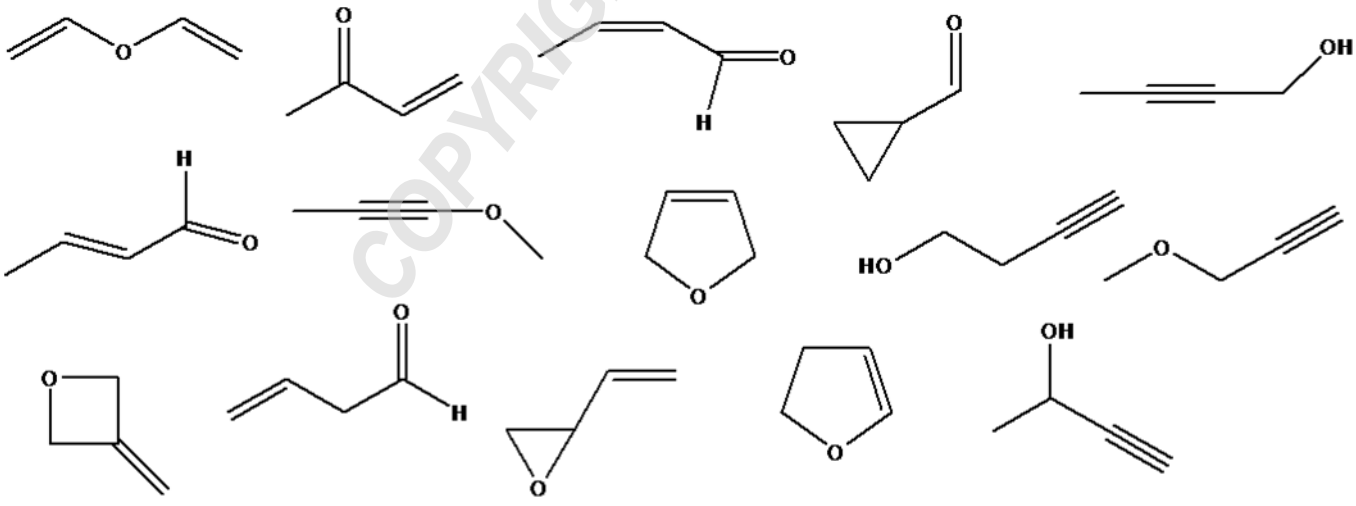

FIGURE 1.1

NMR Data Interpretation Explained: Understanding 1D and 2D NMR Spectra of Organic Compounds and Natural Products, First Edition. Neil E. Jacobsen.

Companion website: www.wiley.com/go/jacobsen/nmrdata

Note that all of the C_4H_6O structures in Figure 1.1 have one thing in common: the total of the number of π bonds plus the number of rings is two in each case. These two "unsaturations" can be determined from the molecular formula by a simple calculation:

1. Discard the oxygen(s): $C_4H_6O \rightarrow C_4H_6$.
2. Any halogens (F, Cl, Br, I) are converted to hydrogens.
3. Any nitrogens (N) are converted to CH (one C and one H for each N). You now have the modified molecular formula: C_4H_6.
4. If **n** is the number of carbon atoms in the modified molecular formula (C_n), calculate the number of hydrogens expected in a saturated hydrocarbon with this number of carbons: $\mathbf{m} = (\mathbf{n} \times 2) + 2 = (4 \times 2) + 2 = 10$.
5. Subtract the number of hydrogens in the modified molecular formula (6) from this saturated hydrocarbon value and divide the result by 2: $\mathbf{m} - 6 = 10 - 6 = 4$; $\mathbf{u} = 4/2 = 2$.

This result (**u**) is equal to the number of π bonds in the molecule *plus* the number of rings. Note that a triple bond (C≡C) is really one σ bond and two π bonds, so it counts as two "unsaturations".

For larger molecules the number of isomers (structures with the same molecular formula) increases very rapidly with the number of atoms. For the formula $C_8H_{11}NO_3$ there are 383 different commercially available compounds! NMR is especially useful for distinguishing between these many possibilities.

In the NMR instrument, each atom (actually the nucleus of each atom) has a precise resonant frequency in the radio frequency spectrum. We can "tune in to the radio channel" of each of these atoms in turn and gather information about the immediate surroundings of that atom in the molecule. There are several kinds of information we can get from each atom:

1. Nearby functional groups change the resonant frequency in predictable ways, so the exact resonant frequency can be used to determine the "chemical environment" of that atom. There are two types of these frequency-shifting effects:
   - **a.** Nearby electronegative atoms (O, N, Br, *etc*.). This effect acts through σ bonds and dies off quickly after 2 or 3 bonds. This is similar to the well-known inductive effect that modifies reactivity in organic chemistry reactions.
   - **b.** Nearby double bonds (C═C or olefin/aromatic, C═O or carbonyl, C≡N or nitrile, *etc*.). This effect acts directly through space and dies off after about 5 Ångstroms (one Ångstrom or Å is approximately the length of a C—H bond). The orientation of the plane of the double bond relative to the atom being observed is also important.
2. Hydrogen atoms are affected by the proximity of other hydrogen atoms in the molecule. So we can look around the immediate vicinity of *our* hydrogen (the one whose radio channel we are tuned to) and see the number and proximity of other hydrogens or groups of hydrogens. This effect manifests itself in two ways:
   - **a.** "Splitting" of the resonant frequency of *our* hydrogen (the one being observed) by a nearby hydrogen into two resonant frequencies very close to each other. The stronger the effect, the wider is the separation of the two frequencies. This effect travels through the bonds and dies off quickly as the number of bonds separating the two hydrogens increases: 2 bonds ≥ 3 bonds > 4 bonds. This effect is sensitive to the angles formed by the bonds connecting the two hydrogens, so we can get information about the relative orientation of groups connected by single bonds. These can either be fixed orientations determined by rigid bonding in rings (stereochemistry) or preferred orientations in a flexible molecule (conformation).
   - **b.** Enhancement of the NMR radio signal received from one hydrogen when we hit the other hydrogen with a radio signal at its precise radio frequency. This enhancement is called an **NOE** and it operates directly through space between hydrogens. The effect dies off quickly with increasing separation and is not seen at all for distances greater than 5 Å. The NOE gives us a molecular ruler for measuring distances between specific pairs of hydrogens in the molecule.

Note that the NMR experiment gives us lots of specific information from the point of view of one atom in the molecule: nearby functional groups and nearby hydrogens, through bonds or directly through space. We can get

the same type of information from each of the atoms in the molecule in turn, especially from the hydrogens. Adding up all of this information (chemical environments, distances, and angles) can give us a covalent structure (which atoms are connected to which by covalent bonds) and a conformation (shape of the molecule in three dimensions).

Determining the structure of a molecule by NMR is a puzzle-solving exercise, and to date it still requires a lot of human judgment and intuition; you don't just feed it into a computer and out pops a structure. The exercise can be exciting and challenging, and it gives the rare human experience of looking straight into the molecular world and getting unambiguous answers to our questions. But it must be emphasized that NMR does not give a **picture** of the molecule. In spite of its close relationship to MRI (magnetic resonance imaging), NMR spectroscopy is not an imaging experiment and it does not give any kind of image or picture of the molecule. You, the person interpreting the NMR data, must put all of these simple pieces of evidence together, along with whatever other information you have, to *propose* a structure of the molecule. Then you have to go over the evidence to make sure all of it is *consistent* with your proposed structure. As in all science, we can gather more and more evidence and be more and more sure of our conclusion, but we can never be absolutely sure. One of the advantages of NMR is that the sheer volume of complimentary information that can be gathered from multiple vantage points (the different atoms in the molecule) makes it a technique with a very high degree of confidence in the conclusions. For small molecules (molecular weight below 500 Da), this confidence comes very close to certainty for experienced users willing to do a number of NMR experiments.

There is another technique for molecular structure determination that *does* generate a picture or image of the molecule. **X-ray crystallography** measures the pattern of scattering of X-rays from a solid crystal of the molecule. By analyzing the intensities of thousands of spots from the scattered X-rays, a computer can create a three-dimensional map of the electron density of the molecule. Since atoms are basically dense clouds of electrons, the atoms can be accurately located and you get a three-dimensional structure of the molecule. The main drawback of this technique is that you need a crystal, and even then the crystal may not have the right properties to give good X-ray diffraction. Once you have a good crystal, the process is time consuming and requires a great deal of calculation and refinement of the data by an expert. In contrast, an NMR spectrum can be acquired in a few minutes if a pure sample can be dissolved in a solvent. The analysis of NMR data, as we shall see, is straightforward and can be learned by anyone with a basic understanding of organic chemistry.

Before we look at the NMR experiment in more detail, some of the other tools for organic structure determination will be briefly explained. These give information which is complementary to the NMR data and help to provide the complete picture of the molecule.

2 MASS SPECTROMETRY

Mass spectrometry is essentially a method for weighing individual molecules to determine their mass. Knowing the masses of individual atoms that make up the molecule (H = 1, C = 12, N = 14, O = 16, *etc.*), we can narrow down the possibilities to a small number of possible molecular formulae. For example, for an integer mass of 120 units, we can have the following molecular formulae:

$$C_9H_{12}: [9 \times 12] + [12 \times 1] = 120$$

$$C_8H_8O: [8 \times 12] + [8 \times 1] + [1 \times 16] = 120$$

$$C_7H_4O_2: [7 \times 12] + [4 \times 1] + [2 \times 16] = 120$$

$$C_7H_8N_2: [7 \times 12] + [8 \times 1] + [2 \times 14] = 120$$

Exercise 1.1: Calculate the number of unsaturations (number of π bonds + number of rings) for each of the above molecular formulae. Explain why $C_6H_{16}O_2$ is not a possible molecular formula for a molecular mass of 120.

We will see shortly that with a more accurate molecular mass, like 120.0687 for $C_7H_8N_2$, we can narrow down the possible molecular formulae to a single one.

NMR focuses on the hydrogens and carbons within a molecule, so it has a hard time counting the oxygen and nitrogen atoms, and other atoms like sulfur and halogens can be "invisible" in the NMR data. This makes mass spectrometry an essential complement to NMR data for determination of structure.

2.1 Ionization Methods and Molecular Ions

The basic experiment of a mass spectrometer is to convert a molecule into a charged species, an ion, and move it around in a vacuum using electric and magnetic fields, to determine its mass by the nature of its motion. There are three basic steps in this process:

1. **Ionization**: Convert the neutral molecule into an ion (usually positive).
2. **Mass Analysis**: Separate ions on the basis of their mass.
3. **Detection**: Detect the ion to generate an electrical signal.

From the point of view of the organic chemist, the first step is the most important. There are two main methods of ionization:

2.1.1 Electron Impact (EI)

The spectrometer gets the molecule into the gas phase and hits it with a high energy electron, knocking out an electron. This is a "hard" ionization process because it imparts a lot of energy to the molecule. The result is a radical cation ($M^{+\bullet}$), a very unstable species that quickly fragments to generate more stable pieces of the molecule. This is the oldest and simplest ionization method and is usually used in conjunction with a gas chromatograph (GC). The sample is injected into the GC, the components (if it is not pure) are separated and the peaks emerging from the GC column go directly into the high vacuum of the mass spectrometer, where the electron beam ionizes the molecules. There are a number of disadvantages to this technique:

- The molecule must be at least somewhat volatile. This limits the technique to fairly simple, non-polar molecules.
- The molecular ion ($M^{+\bullet}$) can be a very weak peak in the mass spectrum, which is dominated by the molecular fragments. While this provides useful information about the molecular structure, it limits the usefulness of mass spectrometry for determining the molecular formula.

The EI mass spectrum of caffeine ($C_8H_{10}N_4O_2$) is shown in Figure 1.2.

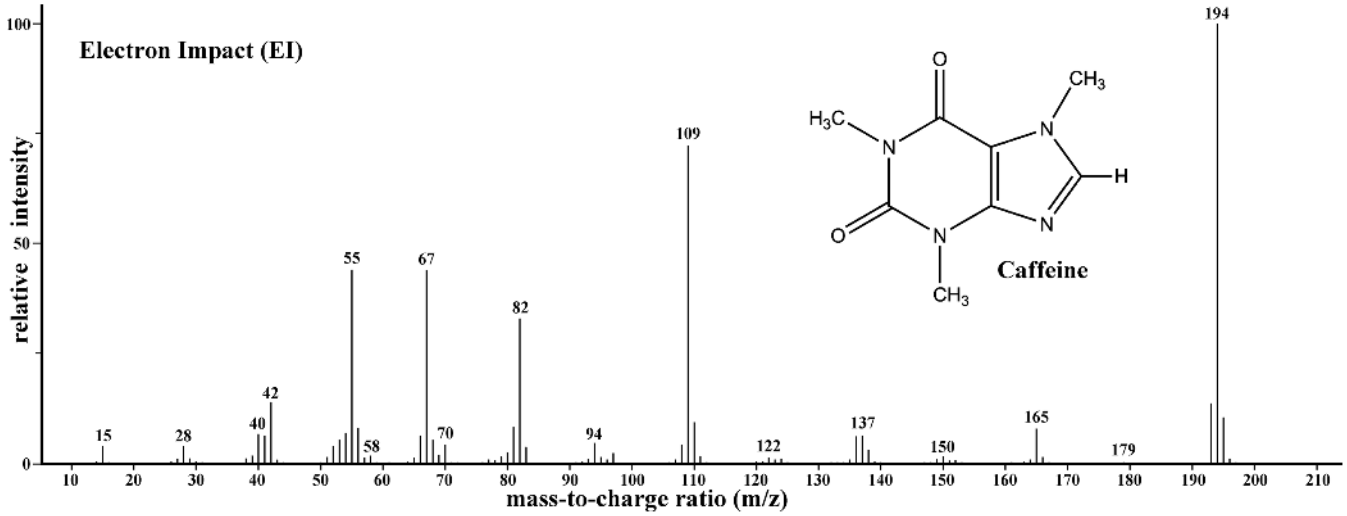

FIGURE 1.2 Courtesy of National Institute of Standards and Technology (NIST).

The horizontal scale is the mass-to-charge ratio (*m/z*), which is essentially the mass in atomic mass units (amu), since virtually all ions are singly-charged (z =1). The vertical scale is the relative intensity of the peaks, relative to the most intense peak, known as the parent ion, as 100%. Note that the peaks are separated between consecutive integer masses. Because caffeine is a very stable aromatic compound, the molecular ion (M^+ at *m/z* 194) is also the most intense peak (the parent ion). The peak at 193 (M – 1) is due to loss of hydrogen ($H^\bullet$), and the peak at 194 (M + 1) is due to the presence of one ^{13}C atom in the molecule (9% intensity). These isotope peaks will be discussed in detail in the next section.

2.1.2 Soft Ionization

This is a general term for low energy ionization methods that essentially just protonate the molecule:

$$M: + BH^+ \rightarrow MH^+ + B:$$

where BH^+ is a proton donor that is supplied to the sample or the mass spectrometer. The molecule has to have at least some basicity, meaning that it needs to have a lone pair that can accept a proton. Some molecules have no basic sites, but these can usually be ionized by negative mode mass spectrometry:

$$MH + B: \rightarrow M:^- + BH^+$$

where B: is a proton acceptor that is supplied to the sample or the mass spectrometer. This mode is less commonly used but is available if positive ion mode gives poor results. Even very weakly basic or acidic sites are amenable to one of these two modes, so the only molecules that would not work would be hydrocarbons.

One consequence of soft ionization is that it's possible to get multiply charged ions. For example, in positive ion mode there may be multiple basic sites on the molecule so there can be a number of different charge states (*e.g.*, $[MH]^+$, $[MH_2]^{+2}$, $[MH_3]^{+3}$, and so forth). This brings up an interesting point about mass spectrometry: it is not actually the *mass* that is measured, but rather the *mass-to-charge ratio* (*m/z*). For small molecules, the organic molecules of this book, this is usually not a big issue, but for biological molecules like peptides and proteins multiply charged ions are very common. For example, if the molecular mass is 1000 (M) and the charge is +6 ($[MH_6]^{+6}$), the measured mass-to-charge ratio (*m/z*) would be 1006/6 = 167.7. In this book, we will give mass spectral data (*m/z*) for singly-charged positive ions only:

M^+	Molecular ion (radical cation)	Electron impact
$[MH]^+$	Protonated molecular ion	Soft ionization

In this case, the mass is the same as the mass-to-charge ratio.

2.2 High-Resolution Mass Spectrometry and Exact Mass

Some mass analyzers are low-resolution, essentially giving only the integer mass. In the example above (*m/z* 120), it would be impossible to distinguish between the possible molecular formulae:

$$C_9H_{12} \quad C_8H_8O \quad C_7H_4O_2 \quad C_7H_8N_2 \quad C_7H_4S$$

All of these give an integer mass of 120 for the molecular ion (or 121 for $[MH]^+$). But high resolution mass analyzers, such as FTICR (Fourier Transform Ion Cyclotron Resonance) and TOF (Time of Flight), can give *m/z* measurement accurate to less than 1 ppm (1 part per million of the measured *m/z*). For example, for a mass of 120, 1 ppm is 0.00012 mass units, or about one unit in the fourth decimal place. At this level of accuracy it's possible to distinguish between

different molecular formulae with the same integer mass. The exact masses of the major isotopes of common atoms are given below:

^{1}H	1.007825	^{32}S	31.972072
^{12}C	12.000000	^{31}P	30.973763
^{16}O	15.994915	^{28}Si	27.976928
^{14}N	14.003074	^{19}F	18.998403

One isotope is exactly equal to the integer mass (^{12}C, by definition), others are slightly more than the integer value (^{1}H, ^{14}N), and others are slightly below the integer values (^{16}O, ^{32}S, ^{31}P, ^{28}Si and ^{19}F). These slight differences from the integer mass allow us to distinguish different molecular formulae if the mass measurement is made with very high accuracy.

It's important to understand that in mass spectrometry we are observing specific isotopic species. For example, for C_8H_8O the major molecular ion peak (M^+) is really $^{12}C_8{}^{1}H_8{}^{16}O$:

$$m/z = (8 \times 12.000000) + (8 \times 1.007825) + (1 \times 15.994915) = 120.057515$$

This is a different mass from $C_7H_8N_2$, which is really $^{12}C_7{}^{1}H_8{}^{14}N_2$:

$$m/z = (7 \times 12.000000) + (8 \times 1.007825) + (2 \times 14.003074) = 120.068748$$

The difference in mass between these two formulae is 0.011233 mass units, or 94 ppm ($0.011233/120 = 94 \times 10^{-6}$). These two formulae can easily be distinguished with a high resolution mass analyzer. High resolution not only requires special equipment, but takes more time and costs more money. Careful calibration using calibrant molecules is required to get this kind of accuracy.

Exact mass calculations can be made easily using this calculator from Scientific Instrument Services:

http://www.sisweb.com/referenc/tools/exactmass.htm

In this book, molecular masses (M^+ or $[MH]^+$) are given in many of the problems and examples to simplify the structural problem of unknowns. Most of these values are not experimental values; they are calculated and, in the case of exact masses, a random error is added or subtracted to give a simulation of actual data. Always remember to subtract 1 mass unit (1.007825 for exact mass) from the protonated molecular ion ($[MH]^+$) *m/z* value to get the mass that corresponds to the (neutral) molecular formula. Sometimes with chemical ionization methods a sodium or potassium ion can take the place of H^+ in creating a positive ion: $[M{\cdot}Na]^+$ or $[M{\cdot}K]^+$. In this case, instead of subtracting the mass of hydrogen to obtain the molecular mass, one has to subtract the mass of sodium ($^{23}Na = 22.989770$) or potassium ($^{39}K = 38.963708$).

Exercise 1.2: For each of the following protonated molecular ion ($[MH]^+$) exact mass values, find the molecular formula, using only H, C, N and O. The number of carbons (obtained by NMR) is given to limit the number of possibilities. Compare the observed *m/z* value to the calculated value and give the error in ppm. Calculate the number of unsaturations (π bonds plus rings) in the molecule.

a. 167.1075 (C_{10}) **b**. 136.1124 (C_9) **c**. 210.1497 (C_{12}) **d**. 195.1500 (C_{11})

The **nitrogen rule** (or odd/even rule) is a simple consequence of the odd number of valences of nitrogen (3 bonds) combined with its even mass (14). The other common atoms have either an odd mass and odd number of bonds (H, Br, Cl) or an even mass and even number of bonds (C, O, S). The rule can be stated simply:

- If the neutral molecular mass (M) is **even**, there is an **even** number of nitrogens in the molecular formula: 0, 2, 4, 6, . . .

- If the neutral molecular mass (M) is **odd**, there is an **odd** number of nitrogens in the molecular formula: 1, 3, 5, 7, . . .

Most importantly, an odd molecular mass means that we probably have nitrogen. An even mass means either we have no nitrogen, or we have at least two nitrogens.

Because the calculation of the number of unsaturations is based on the valencies of the various atoms (C = 4, N = 3, O = 2, H = 1, *etc.*), any molecular formula that violates the nitrogen rule will give a half-integer number of unsaturations. For example, in Exercise 1.2a a formula of $C_{12}H_8N$, with mass of 120, would violate the rule because the mass is even and the number of nitrogens (1) is odd. The calculated number of unsaturations ($C_{12}H_8N \rightarrow C_{13}H_9$, [28 − 9]/2 = 9.5) is not an integer.

2.3 Isotope Patterns and the Halogens Br and Cl

So far we have dealt with atoms that have one isotope with almost 100% abundance (^{1}H, 99.99%, ^{12}C, 98.9%, ^{16}O, 99.76%, ^{14}N, 99.63%). Of these only carbon gives a significant M + 1 isotope peak (^{13}C = 1.11% of the ^{12}C abundance). The intensity of this isotope peak depends on the probability of an ion containing one ^{13}C atom:

$^{12}C_8H_8O$:	Probability of $^{12}C_8 = 0.989^8 = 0.915$	*m/z* 120 (100%)
$^{13}C^{12}C_7H_8O$:	Probability $= 8 \times 0.011 \times 0.989^7 = 0.081$	*m/z* 121 (8.9%)

The probability of all 8 carbons being ^{12}C is the product of all 8 individual probabilities: $0.989 \times 0.989 \times 0.989 \times 0.989 \times 0.989 \times 0.989 \times 0.989 \times 0.989 = 0.989^8$ (0.989 to the eighth power)

Here we are ignoring the slightly less than one probabilities for ^{16}O and ^{1}H, and multiplying by 8 in the second case because there are 8 different ways we can have one ^{13}C and seven ^{12}C atoms in a molecule with 8 carbons. In the mass spectrum, there will be the major molecular ion peak (M^+) at 120 and another peak one mass unit higher (M + 1) with intensity 8.9% of the main (M^+) peak. If there were a large enough number of carbons in the molecule, the probability of one ^{13}C would be larger than the probability of all carbon atoms being ^{12}C. For example, with 100 carbons (C_{100}), the M + 1 ion would have an intensity of 111% relative to the molecular (all ^{12}C) ion. Organic molecules are small enough that this is never a problem. In general, the molecular ion is defined as the lowest mass ion in a cluster of isotope peaks, the ion in which all atoms have their lowest mass: ^{1}H, ^{12}C, ^{14}N, ^{16}O, ^{35}Cl, ^{79}Br, *etc*. In our discussion, this will be defined as 100% intensity, even though there may be other isotope peaks, or fragment ion peaks, that are more intense.

Carbon can be described as an **M + 1 atom** because of the significant isotope with one extra mass unit (^{13}C *vs*. ^{12}C). Similarly, oxygen, sulfur, chlorine and bromine can be described as **M + 2 atoms**:

O	^{16}O	100%	^{18}O	0.21%
S	^{32}S	100%	^{34}S	4.52%
Cl	^{35}Cl	100%	^{37}Cl	31.96%
Br	^{79}Br	100%	^{81}Br	97.28%

As always with mass spectral intensities, the most abundant isotope is expressed as 100% intensity, leading to the perplexing fact that the total is more that 100%. Remember that these are not percentages of the total intensity, but only percent of the parent ion intensity. In low resolution mass spectrometry, the relative intensity of the M + 2 ion can be used to estimate the number of oxygens or, if it is greater than 4%, the number of sulfur atoms.

The halogens chlorine and bromine have huge M + 2 ions, making them very easy to pick out in a mass spectrum. Bromine occurs naturally as ^{79}Br (50.7% of total) and ^{81}Br (49.3% of total). Converting to percentages relative to the lower mass isotope (100%) makes it easier to calculate relative probabilities: ^{79}Br = 100%; ^{81}Br = 97.28%. Just convert the percentage to a fractional intensity (100% = 1.00) and multiply:

$$\underset{\text{M (100\%)}}{Br_2 = {}^{79}Br_2(1.00 \times 1.00)} + \underset{\text{M + 2 (94.6\%)}}{{}^{79}Br^{81}Br(2 \times 1.00 \times 0.9728)} + \underset{\text{M + 4 (94.6\%)}}{{}^{81}Br_2(0.9728 \times 0.9728)}$$

Fractional intensity is converted back to percent at the end. For the M + 2 ion there are two ways that ^{79}Br^{81}Br can happen, so we multiply by two for this statistical factor. An isotope pattern like this (M: M + 2: M + 4 = 100:195:95) can

be directly read out as two bromine atoms in the molecule. This pattern is clearly seen in the EI mass spectrum of 3,5-dibromotoluene ($C_7H_6Br_2$, Figure 1.3), particularly in the expansion of the region around the molecular ion.

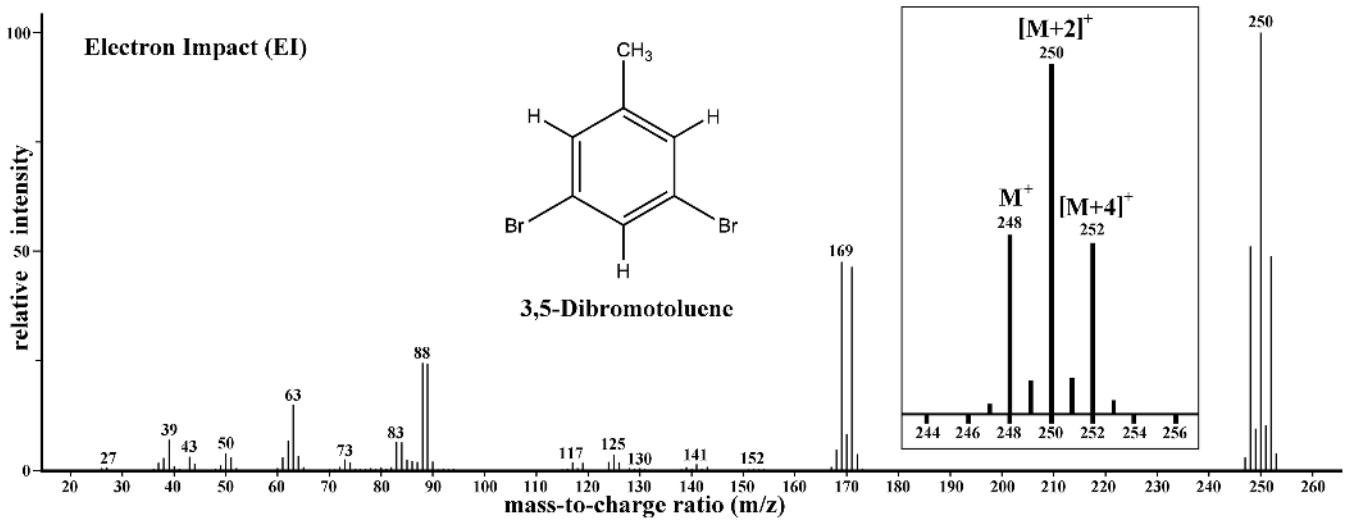

FIGURE 1.3 Courtesy of NIST.

The molecular ion ($M^+ = 248$) is the lowest mass ion in the cluster of isotope peaks, representing the isotopic composition $^{79}Br_2$ for the bromines. The roughly 1 : 2 : 1 ratio of peaks, each one two mass units higher than the previous peak, is a dead giveaway that there are two bromines in the molecule. The fragment ion at *m/z* 169 has a companion peak at *m/z* 171 of nearly equal intensity (1 : 1 ratio), so it contains only one bromine atom. In fact, this fragment represents the loss of Br from the molecular ion (248 − 79 = 169). This fragment can be described as "M-79" (molecular ion minus 79 mass units) or "M-Br" (loss of bromine).

Exercise 1.3: Calculate the number of unsaturations in the caffeine and 3,5-dibromotoluene molecular formulae, and count the number of π bonds and the number of rings in each structure.

Exercise 1.4: For the following isotope combinations, calculate the M + 2 (and M + 4, M + 6, *etc.*) intensities in the mass spectral isotope pattern.

a. Cl_2 **b.** BrCl **c.** Cl_2Br **d.** Br_3

All of these isotope abundances are expressed as percent of the lowest mass isotopic species. In mass spectrometry, the peak intensities are usually expressed in percent of the most intense peak (the parent peak), which can easily be the M + 2 or M + 4 peak if there are multiple bromine or chlorine atoms present. As long as the peak intensities are expressed as ratios this should not cause any confusion. For example:

$$m/z\ 210/212/214\ (100{:}\,195{:}\,95)$$

implies that there are two bromines in the molecular formula, with the 210 mass corresponding to $^{79}Br_2$, the 212 mass corresponding to $^{79}Br^{81}Br$, and the 214 mass corresponding to $^{81}Br_2$. Alternatively, this could be written as:

$$m/z\ 210/212/214\ (51{:}\,100{:}\,49)$$

All intensity ratios in this book will be expressed using 100 for the lowest mass isotopic species in the molecular ion.

3 INFRARED (IR) SPECTROSCOPY

This is an important tool for the organic chemist that measures the absorption of electromagnetic radiation (light) in the frequency range of 500 to 4000 wavenumbers (cm^{-1}), corresponding to a wavelength range of 2.5 to 20 microns (1 micron or $\mu = 10^{-6}$ meters). This is beyond the low energy (red) edge of the visible light spectrum, hence the term "infra – red" (below red). Absorption of light in this frequency range corresponds to stretching vibrations and bending motions of the chemical bonds in a molecule. The IR spectrum can give useful information about the functional groups of a molecule: carbonyl (C═O), hydroxyl (OH), nitrile (C≡N), olefin (C═C), *etc.* The IR spectrum is presented with a frequency scale at the bottom and the baseline at the top, with absorption **bands** appearing as dips in the baseline towards the bottom of the display. Some of these bands can be narrow ("sharp") and well-defined, and others can be very wide ("broad") and amorphous. We are concerned here only with a few of the most useful absorption bands in the IR spectrum, those that are easily interpreted and give information that may be difficult to obtain from the NMR spectrum.

The IR spectrum of *para*-acetyl-benzonitrile (*p*-CH_3CO-C_6H_4-CN) is shown in Figure 1.4.

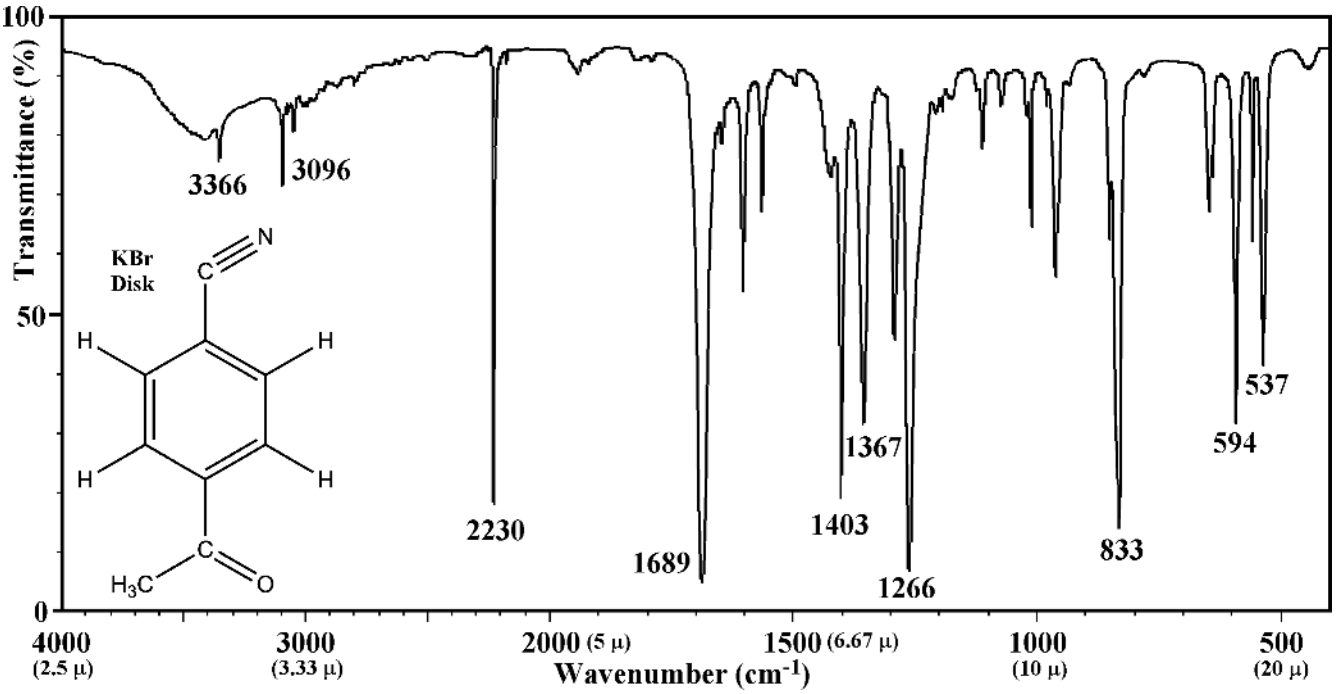

FIGURE 1.4 Courtesy of National Institute of Advanced Industrial Science and Technology (AIST), Japan.

The frequency (horizontal) scale is expanded in the 1500–500 cm^{-1} range. Wavenumbers are the reciprocal of wavelength ($10^4/\lambda$ in microns), so they are proportional to frequency (Hz = speed of light/λ). The vertical scale is transmittance, the percentage of light that makes it through the sample without being absorbed. The solid sample was ground with solid potassium bromide (KBr) in a mortar and pestle and then pressed into a thin disc. The infrared light passes through this disc and ends up at a detector. Solid KBr does not absorb infrared light, so this is a very clean way to record the IR spectrum. Liquid samples (neat liquids without solvent) can be pressed in a thin film between two salt plates. There are also solution sample cells using solid salt, but in solution some regions will be wiped out by the IR absorption bands of the solvent.

The most important single band in an IR spectrum is the **carbonyl (C═O) stretching** vibration, a strong band in the region of 1700 cm^{-1}. This is particularly useful because this is a "quiet" region of the spectrum with little interference from other stretching or bending motions. The generic value for a ketone is 1715 cm^{-1}. In Figure 1.4 this band is at 1689 cm^{-1}, shifted to lower frequency from the generic value due to conjugation of the ketone with the aromatic ring. Esters (R—C(═O)—OR′) show this band shifted to higher frequency (~1750 cm^{-1}).

The nitrile triple bond gives rise to a very distinctive band at 2230 in *p*-acetyl-benzonitrile, due to the **C≡N stretching** vibration (Figure 1.4). This strong, narrow band is also in a quiet region of the spectrum. The only other

band in this region is the weak C≡C stretching vibration of alkynes (2260–2100 cm^{-1}). Beware of contamination with deuterated chloroform (the most common solvent used in NMR) because the C-D stretching vibration is at 2256 cm^{-1}.

The **C—H bond stretching** vibrations give rise to bands in the region near 3000 cm^{-1}. These are fairly weak bands, and the general rule is that aliphatic (sp^3-hybridized) C—H bands occur to the right (lower frequency) of 3000 cm^{-1} and aromatic and olefinic (sp^2-hybridized) C-H bands occur to the left (higher frequency) of 3000 cm^{-1}. Note the band at 3096 cm^{-1} in Figure 1.4. The distinction between aliphatic, aromatic and olefinic hydrogens is easily made by NMR.

The **O—H bond stretching** gives rise to a very broad band around 3300 cm^{-1}. If salt plates are not protected from moisture this band will appear due to H_2O even if the sample contains no O—H bonds. The N—H stretching band of amines and amides appears in the same region and is also quite broad.

Since the advent of NMR spectroscopy in the 1960s, infrared spectroscopy has gradually diminished in importance in the elucidation of organic structures. The important difference between IR and NMR is that the frequency of infrared absorption bands depends on the vibrational modes of the molecule as a whole, whereas the resonant frequency of each nucleus (*e.g.*, a specific H or C in the molecule) is a local phenomenon that responds to the immediate environment (within 3–4 bonds or within 5 Å) of that particular atom within the molecule.

4 ULTRAVIOLET (UV) AND VISIBLE SPECTROSCOPY

The visible light spectrum extends from a wavelength of 390 nm (0.39 microns, violet) to 700 nm (0.7 μ, red). The infrared spectrum is lower energy (longer wavelength) than the low energy (red) side of the visible light spectrum, and the ultraviolet (UV) spectrum is higher energy (shorter wavelength) than the high energy (violet) side of the visible spectrum. The recorded ranges of spectrometers are shown below:

$$20\ \mu \underset{IR}{\rightarrow} 2.5\ \mu \qquad 700\ nm \underset{VIS}{\rightarrow} 390\ nm \qquad 390\ nm \underset{UV}{\rightarrow} 210\ nm$$

An example of an ultraviolet (UV) absorption spectrum is shown in Figure 1.5 for oxybenzone, a major ingredient in sunscreen creams.

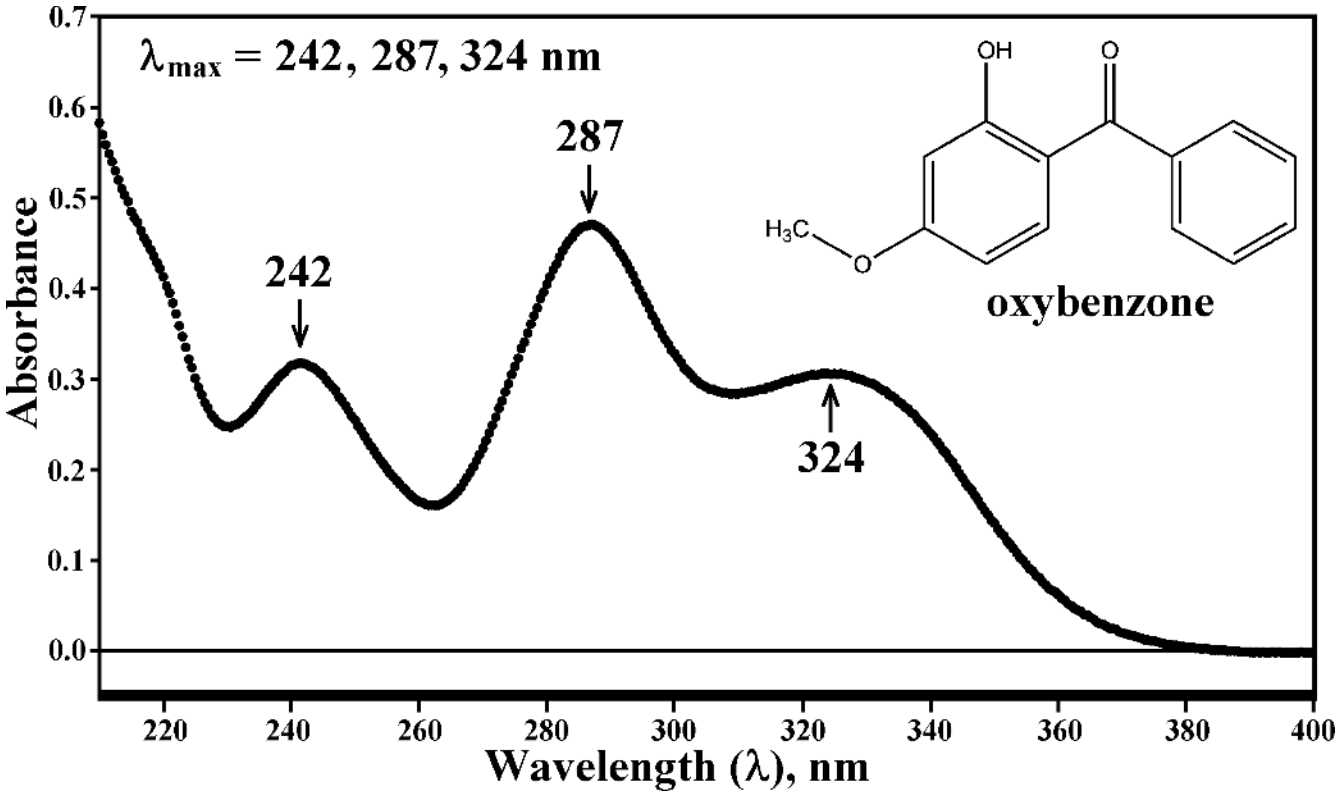

FIGURE 1.5

Ultraviolet light from the sun reaches the earth primarily as UVB (315–400 nm wavelength), since the more energetic UVA (280–315 nm) is largely absorbed by the ozone layer in the earth's atmosphere. Oxybenzone is particularly useful as a sunblock agent because it blocks the longer wavelength UVB as well as UVA. The UV spectrum of oxybenzone shows three absorption maxima, at 242, 287 and 324 nm. The data is usually reported by listing these maxima:

UV spectrum (CH_3OH): λ_{max} 242, 287, 324 nm.

The spectrum was recorded using a dilute solution of oxybenzone in methanol. The vertical scale (*y* axis) in the spectrum is in absorbance units, defined as the logarithm of the light intensity ratio:

$$A = \log_{10}(I_o/I)$$

As light of a single wavelength (monochromatic light) passes through the sample (a liquid solution), the intensity of light coming into the sample (I_o, "I-zero") is reduced by absorption to give a lower intensity coming out (I). For example, if only 1/10 of the light makes it though the sample to come out the other side, we have an absorbance of 1.0 ($\log_{10}$ of 10).

In addition to reporting the wavelengths of the maxima in the UV spectrum, often the intensity of absorption is reported at each maximum. This is a measure of the molecule's ability to absorb light at that wavelength, known as the molar absorptivity or **extinction coefficient**. The word "extinction" refers to the dimming of the light as it passes through the sample solution. The absorption of light is proportional to the concentration (molarity) of the sample solution, according to the Beer-Lambert law:

$$A = \varepsilon c\ L$$

where A is the absorbance, c is the concentration (M, moles per liter) of the sample solution, L is the path length in centimeters (cm) of the light passing through the sample, and ε (epsilon) is the extinction coefficient. In other words, ε is the proportionality constant between concentration (and path length) and the absorbance measured. Since A has no units (it is a logarithm), the units of ε are $M^{-1}\ cm^{-1}$. The extinction coefficient is a physical property of the molecule at any particular wavelength, as long as the same solvent is used. The extinction coefficient is only reported at the maximum of the spectrum, and if there are multiple maxima (like the three observed for oxybenzone), an extinction coefficient can be reported for each maximum.

None of this would be particularly interesting to the organic chemist interested in determining the structure of a molecule, except that the wavelength of the UV maximum (the λ_{max}) is related to the amount of **conjugation** of double bonds in a molecule. A double bond (*e.g.*, C═C) is conjugated if it is next to another double bond in the structure, for example:

C=C—C=O

Now we have an alkene, C═C, directly attached to a carbonyl group, C═O, so that the four p orbitals of the two π bonds form a continuous series, or an extended π system. Electrons in the π system are **delocalized**, meaning that any one π electron can roam over the entire region of the four p orbitals (four atoms: C, C, C and O). The rule of thumb for the UV spectrum is that **the longer the extended π system, the longer the wavelength of UV absorption**, meaning the farther the curve extends to the right in the UV spectrum. A brief look into how light is absorbed will give some insight into this simple relationship.

Absorption of light in the UV and visible range is due to electronic transitions in the molecule, promoting an electron from the highest occupied molecular orbital (HOMO) to the empty lowest unoccupied molecular orbital (LUMO), above it on the energy scale. The electron starts in a lone pair (n) or bonding (σ or π) orbital and ends up in an *antibonding* (σ* or π*) orbital, absorbing a photon of light. The wavelength of light that is absorbed depends on the energy separation (the "gap") between the ground state (lower energy state) and the excited state (higher energy level) molecular orbitals. If the energy separation is large, the wavelength is short (higher energy photons), and if the energy separation is smaller, the wavelength moves toward the low energy (longer wavelength) edge of the UV range (390 nm). Eventually if the energy gap is small enough the molecule will absorb visible light, appearing yellow to the

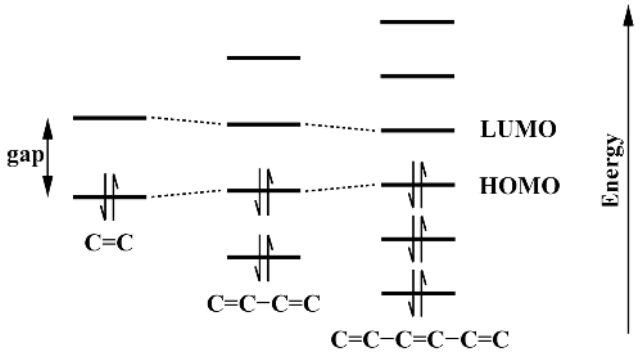

FIGURE 1.6

eye because the blue side of the visible spectrum is being removed (absorbed) by the molecule. This relationship between the HOMO-LUMO gap and the wavelength of absorbed light is illustrated in Figure 1.6.

A simple alkene (C═C) has two π electrons, and there are two molecular orbitals: the lower energy (bonding or π) orbital and the higher energy (antibonding or π*) orbital. The two electrons pair together in the lower energy (highest occupied or HOMO) orbital, and the absorption of light promotes one electron up to the higher energy (lowest unoccupied or LUMO) orbital. The wavelength of light absorbed is determined by the energy gap (HOMO-LUMO gap), and the wider the gap the higher the light energy (shorter wavelength).

If two alkenes are connected in conjugation (*e.g.*, 1,3-butadiene), the four p orbitals combine to form four molecular orbitals, and the four π electrons are paired and put into the two lowest energy orbitals (Figure 1.6, center). These four orbitals are closer together in energy than the two orbitals of the simple alkene, resulting in a smaller HOMO-LUMO gap, and a lower energy (longer wavelength) of light that can be absorbed. In other words, the measured λ_{max} in the UV spectrum is a larger number, corresponding to a longer wavelength of light absorbed. Adding another conjugated alkene unit results in three π bonds in a conjugated system (*e.g.*, 1,3,5-hexatriene), with six π electrons paired into the three lowest energy molecular orbitals (Figure 1.6, right). These levels are even more closely spaced, so the gap is smaller and the wavelength of light absorbed (λ_{max}) is even longer. This relationship is illustrated in Figure 1.7 with a series of conjugated olefins with increasing length of the extended π system.

The λ_{max} values shown are for the longest-wavelength maximum in the UV spectrum of each compound. Eventually as the π system gets longer and longer the light absorbed moves out of the UV range and into the visible light spectrum, resulting in colored organic compounds. With the visible light spectrum starting at 390 nm on the violet end, it's clear that the bottom two examples in Figure 1.7 are colored compounds. Synthetic dyes are an example of these extended π systems. Since UV absorption peaks are usually broad, any strong absorbance with a maximum of 350 nm or more will lead to absorbance of visible light and a colored compound.

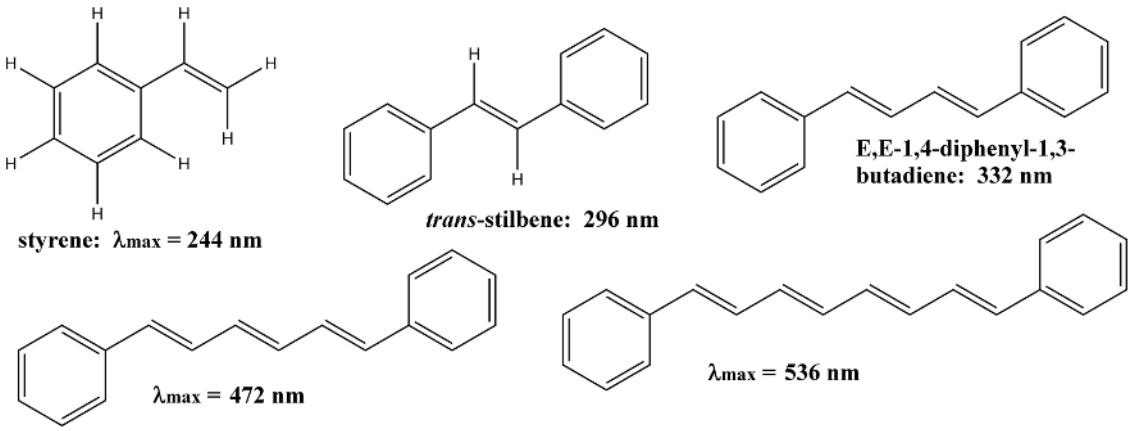

FIGURE 1.7

The relationship between the type of extended π system (linear, branched, involving O or N atoms, aromatic, *etc*.) and the λ_{max} and ε of the UV spectrum has been studied in detail, and there are rules that make it possible to distinguish between various structural possibilities using the data in the UV spectrum. Some sense of the typical values of UV absorption parameters can be gained by looking at some common functional groups in organic compounds (Figure 1.8).

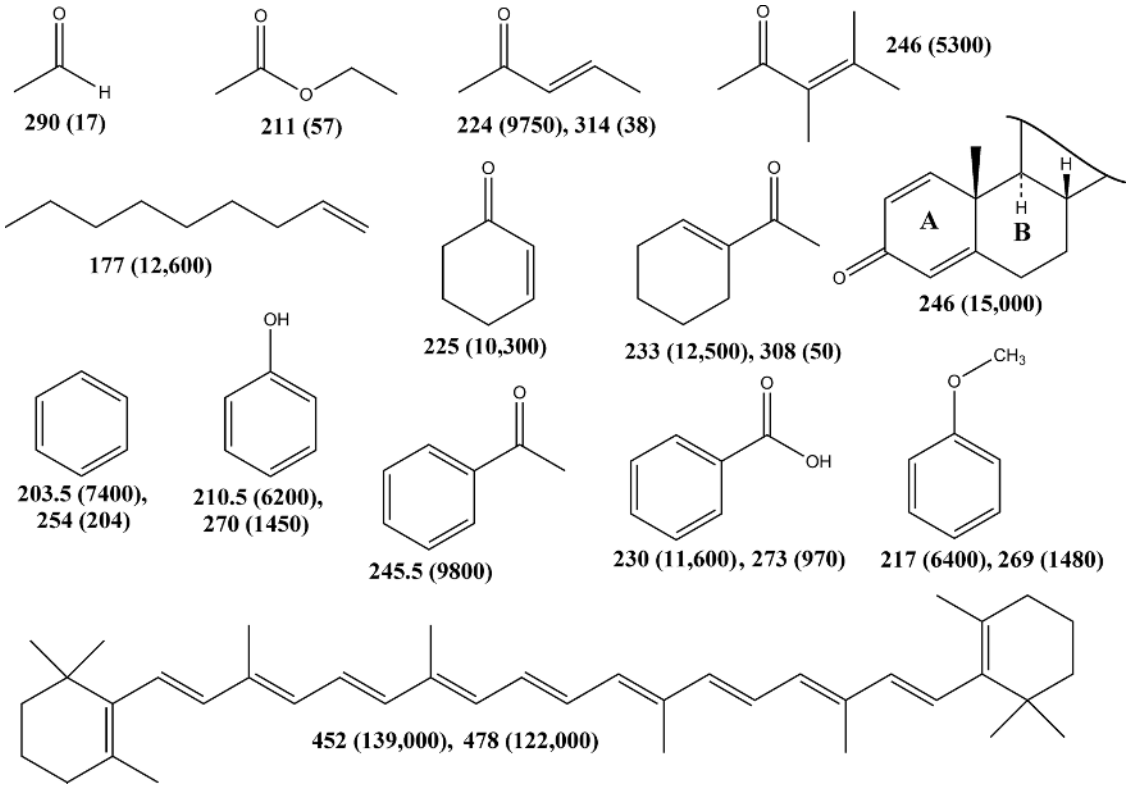

FIGURE 1.8

Next to each structure the λ_{max} value in nm is given, with the extinction coefficient (ε) in parentheses. Where there is more than one absorption maximum the two most important ones are listed. Maxima that are shorter than 220 nm do not provide much useful information to the organic chemist, and may be difficult to measure. Note that the "branched" enone system of the steroid (*e.g*., prednisolone, Figure 5.26) has a longer wavelength maximum and larger extinction coefficient than the simple α,β-unsaturated ketones. The long conjugated system of *trans*-β-carotene (Figure 1.8, bottom, 11 double bonds in a linear system) has strong absorbance in the visible light region, giving it an orange color.

5 A HIGHLY SIMPLIFIED VIEW OF THE NMR EXPERIMENT

NMR stands for **N**uclear **M**agnetic **R**esonance. This technology is based on the atom's nucleus (the **N** in NMR), which for our purposes can be viewed as a positively charged sphere. The nucleus has a property called *spin*, meaning that it is permanently spinning on its axis just like the Earth (Figure 1.9, left). The rotating charge creates a magnetic field, so we can think of the nucleus as a tiny bar magnet (Figure 1.9, right). This magnetic property of the nucleus accounts for the **M** (magnetic) in NMR.

In the NMR spectrometer, the sample is placed in a very strong magnetic field (Figure 1.10). Each nuclear magnet can line up *with* the magnetic field, like a compass needle in the Earth's magnetic field (Figure 1.10, arrow pointing up), or it can line up *against* the magnetic field (Figure 1.10, arrow pointing down). Its orientation is limited by quantum

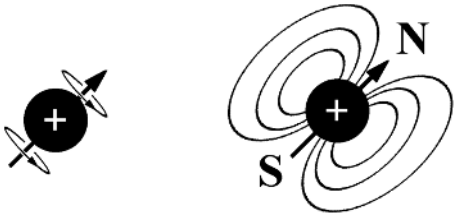

FIGURE 1.9

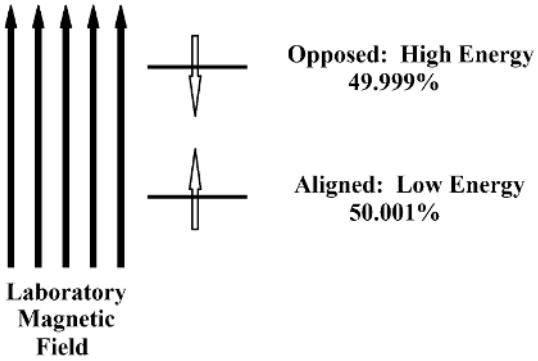

FIGURE 1.10

mechanics to only these two possibilities. The aligned (pointing up) state is lower in energy because the magnet likes to be aligned with the magnetic field. The opposed (pointing down) state is higher in energy because it is fighting the instrument's magnetic field.

The energy difference, or energy gap, between these two possible states corresponds to the energy of radio waves (radio frequency) in the range of hundreds of millions of Hertz (one million cycles per second or Hz is one Megahertz or MHz). By comparison, the frequencies of your car radio's FM scale correspond to 88–104 MHz. The size of the energy gap is very small (radio waves have very low energy), so there is only a slight preference for the nuclei in a sample to be in the lower energy (aligned) state: slightly more than 50% find themselves in this state. The exact size of the energy gap is affected slightly by the chemical environment that the nucleus finds itself in within the molecule. For example, for a hydrogen nucleus (a proton), being close to an oxygen atom (H-C-O) in an organic molecule will increase the energy gap very slightly relative to a hydrogen that is far away from the oxygen. The radio frequency corresponding to the increased energy gap is also higher. We call this precise radio frequency the *resonant frequency* of that particular proton within the molecule, and this gives us the R in NMR (for resonance). It should be clear, then, that the exact value of this resonant frequency is different for each proton within the molecule and carries valuable information about the chemical environment (types of bonding, nearby atoms, nearby double bonds, *etc.*) in the immediate vicinity of that hydrogen atom.

The purpose of the NMR experiment is to measure the resonant frequencies of all of the nuclei of a particular type (*e.g.*, all of the hydrogen nuclei or protons) contained in a molecule. The job of the chemist is to interpret these frequencies and use that information to determine the structure of the molecule, or to confirm a structure that is already known or suspected.

In the NMR experiment, a powerful pulse of radio frequency energy is applied to the sample (Figure 1.11, left), causing some of the nuclei in the aligned (lower energy) state to jump up to the opposed (higher energy) state (Figure 1.11, right).

Over the next one or two seconds, the nuclei drop back down or "relax" to the lower energy state, each time shooting out a burst of radio frequency energy at the exact resonant frequency of the nucleus that is relaxing (Figure 1.12).

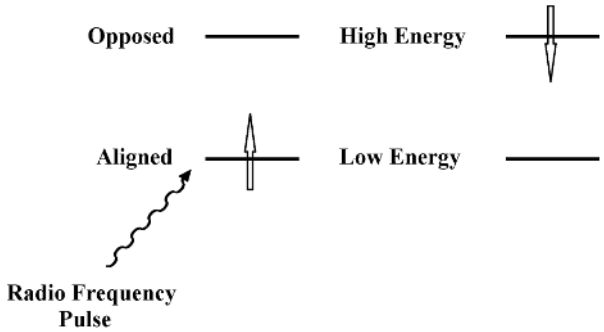

FIGURE 1.11

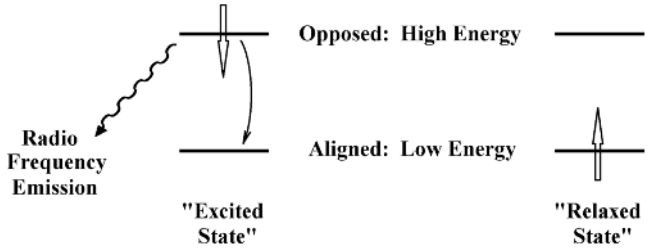

FIGURE 1.12

These bursts of radio waves are received by a radio receiver and converted to digital information in the spectrometer. For a proton (^{1}H) experiment, one precise frequency is detected in the radio receiver for each proton (hydrogen) in the molecule. These frequencies are all mixed together in the signal received by the spectrometer. For example, a sample of *n*-propyl acetate gives four distinct proton frequencies (four different chemical environments for hydrogen):

$$\underset{a}{CH_3}—C(O)—O—\underset{b}{CH_2}—\underset{c}{CH_2}—\underset{d}{CH_3}$$

a.	400.130821 MHz	–	400.13 MHz	=	821 Hz
b.	400.131610 MHz	–	400.13 MHz	=	1610 Hz
c.	400.130661 MHz	–	400.13 MHz	=	661 Hz
d.	400.130378 MHz	–	400.13 MHz	=	378 Hz

Just like in your car's radio receiver, the station frequency (in NMR this is the frequency of the nucleus you are "tuned to") is subtracted out, leaving the small differences that contain the chemical information. For the *n*-propyl acetate sample, these four frequencies are mixed together and the complex signal gradually dies out as all of the excited nuclei relax back to their original (equilibrium) state. A simulation of this sum of four frequencies is shown in Figure 1.13. An enlargement of the first part of this raw NMR data (called the FID, or free induction decay) is shown in Figure 1.14.

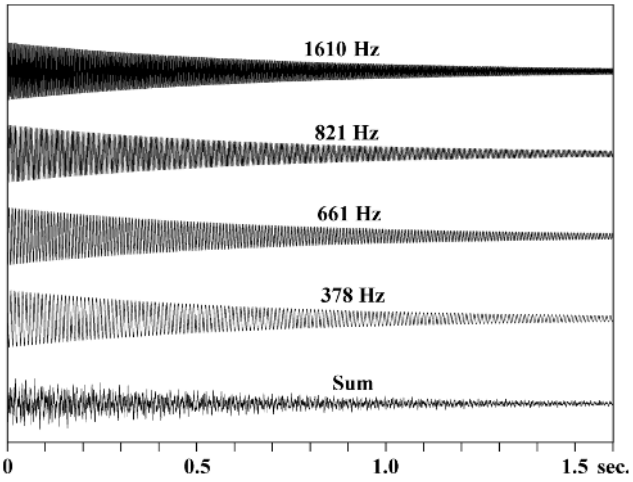

FIGURE 1.13

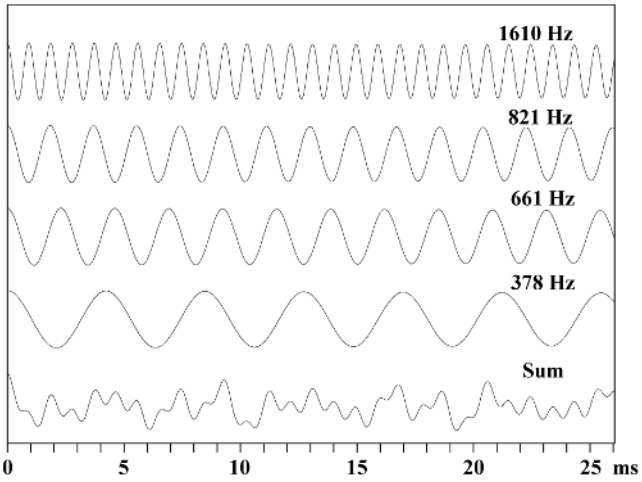

FIGURE 1.14

The complex waveform of the sum (the raw NMR data or FID) cannot be interpreted by us because all the sine waves of different frequencies are mixed together. Instead, a mathematical calculation called the Fourier transform is used to "pull out" the frequency information. After the Fourier transform, we have a graph of frequency (horizontal axis) *vs.* intensity (vertical axis) that shows the four distinct frequencies as lines or peaks in the graph (Figure 1.15). This is called the NMR spectrum, and it is now possible to interpret the NMR data and get information about the structure of the molecule in the sample.

Different NMR spectrometers can have different magnet strengths, and the stronger the magnet the larger the differences in frequency between different protons within a molecule. To standardize these differences, or chemical shifts, we normalize the frequencies by dividing by the "radio station" frequency that was subtracted out:

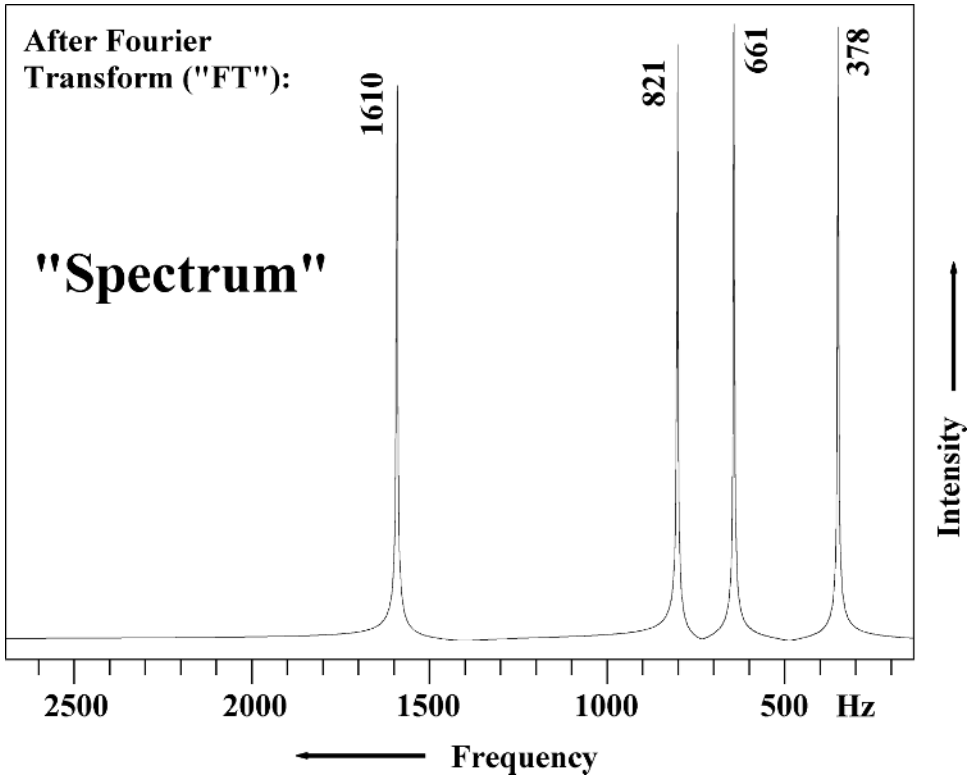

FIGURE 1.15

$$\underset{\mathbf{a}}{CH_3}—CO—O—\underset{\mathbf{b}}{CH_2}—\underset{\mathbf{c}}{CH_2}—\underset{\mathbf{d}}{CH_3}$$

a. 821 Hz/400.13 MHz = 2.051 parts per million (ppm)
b. 1610 Hz/400.13 MHz = 4.024 ppm
c. 661 Hz/400.13 MHz = 1.651 ppm
d. 378 Hz/400.13 MHz = 0.944 ppm

If we buy a bigger magnet, the "station frequency" goes up by the same factor as the frequency difference in Hz, so the chemical shift in ppm (Hz/MHz = 10^{-6} = parts per million) remains the same and can be considered a characteristic of the chemical environment of that proton within the molecule. Through simple rules, tables of chemical shifts and, more recently, large databases of chemical shifts of thousands of different molecules, we can learn a lot about the chemical environments in a molecule from these chemical shift values.