Determining organic structures

3

Connections

Building on

- What sorts of structures organic molecules have ch2

Arriving at

- Determining structure by X-ray crystallography
- Determining structure by mass spectrometry
- Determining structure by ^{13}C NMR spectroscopy
- An introduction to ^{1}H NMR spectroscopy
- Determining structure by infrared spectroscopy

Looking forward to

- How ^{13}C NMR spectroscopy helps locate electrons ch7
- How infrared spectroscopy tells us about reactivity ch10 & ch11
- Using ^{1}H NMR spectroscopy to determine structures ch13
- Solving unknown structures spectroscopically ch13

Introduction

Organic structures can be determined accurately and quickly by spectroscopy

Having urged you, in the last chapter, to draw structures realistically, we now need to answer the question: what is realistic? How do we know what structures molecules actually have? Make no mistake about this important point: *we really do know what shape molecules have*. You wouldn't be far wrong if you said that the single most important development in organic chemistry in modern times is just this certainty, as well as the speed with which we can *be* certain. What has caused this revolution can be stated in a word—spectroscopy.

What is spectroscopy?

Rays or waves interact with molecules	Spectroscopy	Tells us about
X-rays are scattered by atoms	Measures the scattering pattern	Bond lengths and angles
Radio waves make nuclei resonate	Plots charts of resonant frequencies	The symmetry and connectivity of the hydrocarbon skeleton
Infrared waves make bonds vibrate	Plots charts of absorption	The functional groups in the molecule

Structure of the chapter

We shall first consider structure determination as a whole and then introduce three different methods:

- mass spectrometry (to determine mass of the molecule and atomic composition)
- nuclear magnetic resonance (NMR) spectroscopy (to determine symmetry, branching, and connectivity in the molecule)
- infrared spectroscopy (to determine the functional groups in the molecule).

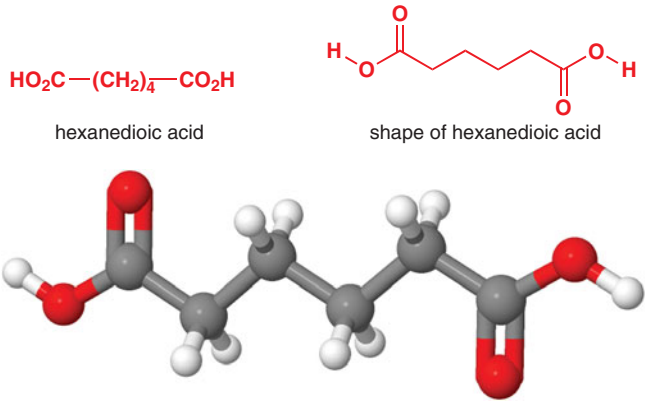
Online support. The icon in the margin indicates that accompanying interactive resources are provided online to help your understanding: just type **www.chemtube3d.com/clayden/123** into your browser, replacing **123** with the number of the page where you see the icon. For pages linking to more than one resource, type **123-1, 123-2** etc. (replacing **123** with the page number) for access to successive links.

■ If you would like more details of any of the spectroscopic methods we discuss, you should refer to one of the specialized books listed in the 'further reading' section at the end of this chapter.

Of these, NMR is more important than all the rest put together and so we shall return to it in more detail in Chapter 13. Then in Chapter 18, after we've discussed a wider range of molecules, there will be a review chapter to bring the ideas together and show you how unknown structures are really determined.

X-ray is the final appeal

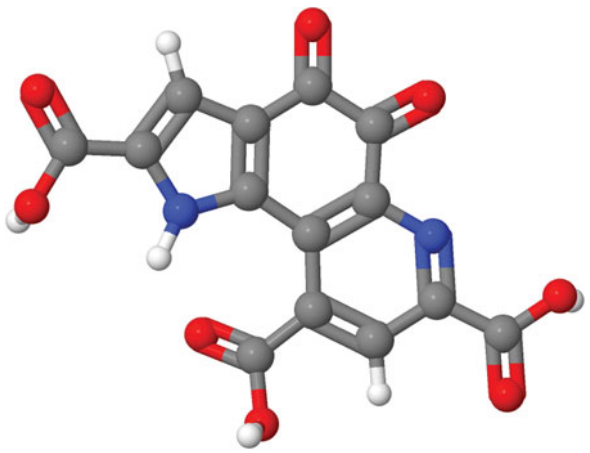
In Chapter 2 we suggested you draw saturated carbon chains as zig-zags and not in straight lines with 90° or 180° bond angles. This is because we know they *are* zig-zags. The X-ray crystal structure of the 'straight' chain diacid, hexanedioic acid, is shown below. You can clearly see the zig-zag chain, the planar carboxylic acid groups, and even the hydrogen atoms coming towards you and going away from you. It obviously makes sense to draw this molecule *realistically*, as in the second drawing.

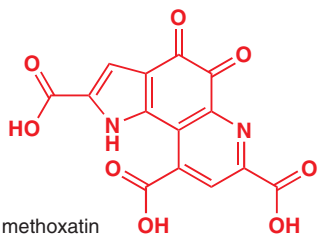


X-ray crystal structures are determined by allowing a sample of a crystalline compound to diffract X-rays. From the resulting diffraction pattern, it is possible to deduce the precise spatial arrangement of the atoms in the molecule—except, usually, the hydrogen atoms, which are too light to diffract the X-rays and whose position must be inferred from the rest of the structure. This is one question that X-ray answers better than any other method: what shape does a molecule have? Another important problem it can solve is the structure of an important new unknown compound. There are subterranean bacteria, for example, that use methane as an energy source. It is amazing that bacteria manage to convert methane into anything useful, and, of course, chemists really wanted to know how they did it. In 1979 it was found that the bacteria use a coenzyme, given the trivial name 'methoxatin', to oxidize methane to methanol. Methoxatin was a new compound with an unknown structure and could be obtained in only very small amounts. It proved exceptionally difficult to solve the structure by NMR but eventually methoxatin was found by X-ray crystallography to be a polycyclic tricarboxylic acid.

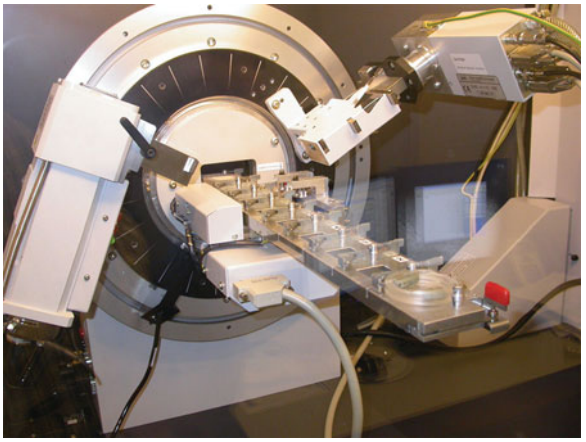
■ **Coenzymes** are biochemical reagents that work hand-in-hand with enzymes to catalyse reactions.

Interactive structure of methoxatin





■ The trivial name 'methoxatin' has a systematic alternative: 4,5-dihydro-4,5-dioxo-1H-pyrrolo[2,3-*f*]quinoline-2,7,9-tricarboxylic acid. Both are valid names. There are no prizes for guessing which one is used more often.



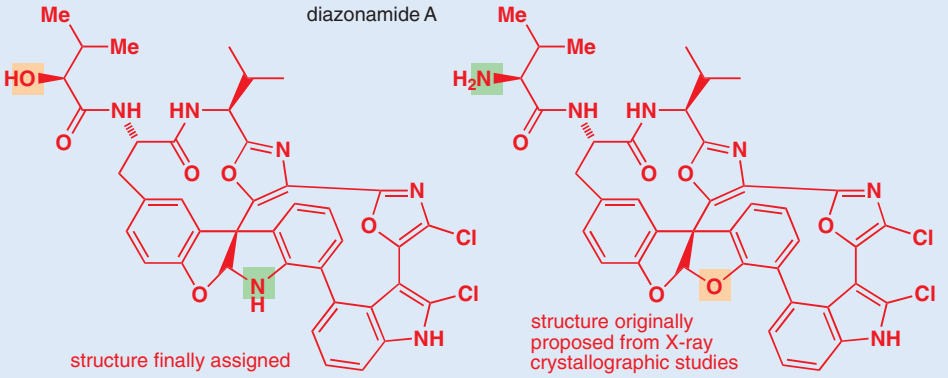
X-ray crystallography has its limitations

If X-ray crystallography is so powerful, why do we bother with other methods? There are two reasons:

- X-ray crystallography works by the scattering of X-rays from electrons and requires crystalline solids. If an organic compound is a liquid or is a solid but does not form good crystals, its structure cannot be determined in this way.
- X-ray crystallography is a science in its own right, requiring specialized skills, and a structure determination can take a long time. Modern methods have reduced this time to a matter of hours or less, but nonetheless by contrast a modern NMR machine with a robot attachment can run more than 100 spectra overnight. We normally use NMR routinely and reserve X-rays for difficult unknown structures and for determining the detailed shape of important molecules.

X-ray crystallography is not infallible!

Because it cannot usually 'see' H atoms, it is important to appreciate that X-ray crystallography is not infallible: it can still get things wrong. A famous example is the antibiotic diazonamide A, which from 1991 (when it was isolated from a marine organism) until 2001 (when the error was realized) was thought to have the structure shown on the right. It has the same mass as the real structure on the left, and X-ray crystallography was unable to tell the O and the N apart. Only when the compound was synthesized did the error become apparent, and the fact that the correct structure was indeed that on the left was confirmed by the fact that synthetic material of this structure made in 2002 was identical with the natural product.



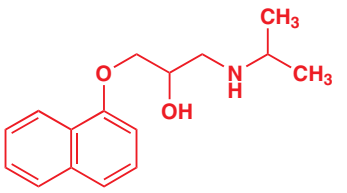
Outline of structure determination by spectroscopy

Put yourself in these situations, regularly encountered by professional chemists:

- You notice an unexpected product from a chemical reaction.
- You discover a previously unknown compound in a plant extract.

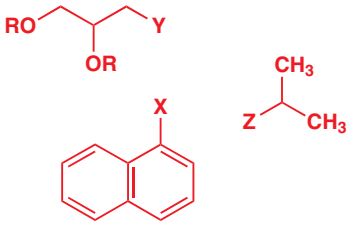
- You detect a suspected food contaminant and need to know what it is.
- You are routinely checking purity during the manufacture of a drug.

In all cases, except perhaps the second, you would need a quick and reliable answer. Suppose you are trying to identify the heart drug propranolol. You would first want to know the molecular weight and atomic composition, and these would come from a *mass spectrum*: propranolol has a molecular weight (relative molecular mass) of 259 and the composition $C_{16}H_{21}NO_2$. Next you would need the carbon skeleton—this would come from *NMR*, which would reveal the three fragments shown below.



propranolol $C_{16}H_{21}NO_2$

fragments of propranolol from the NMR spectrum:



■ NMR does not literally break up the molecule into fragments, but it does view molecules as pieces of hydrocarbon linked together.

There are many ways in which the fragments seen by NMR could be joined together and at this stage you would have no idea whether the oxygen atoms were present as OH groups or as ethers, whether the nitrogen would be an amine or not, and whether Y and Z might or might not be the same atom, say N. More information comes from the infrared spectrum, which highlights the functional groups, and which would show that there is an OH and an NH in the molecule but not other functional groups such as CN or NO_2. This still leaves a variety of possible structures, and these could finally be distinguished by the details revealed by ^{1}H NMR. We will deal with ^{1}H NMR only briefly in this chapter because it is more complicated than ^{13}C NMR, but we will return to it in Chapter 13.

Now we must go through each of these methods and see how they give us information about the propranolol molecule.

● What each spectroscopic method tells us

Method and what it does	What it tells us	Type of data provided
Mass spectrum weighs the molecule	Molecular weight (relative molecular mass) and composition	259; $C_{16}H_{21}NO_2$
^{13}C NMR reveals all the different carbon nuclei	Carbon skeleton	No C=O group; ten carbons in aromatic rings; two carbons next to O; three other saturated C atoms
Infrared reveals chemical bonds	Functional groups	No C=O group; one OH; one NH
^{1}H NMR reveals all the different H nuclei	Distribution of H atoms	Two methyl groups; six H atoms on aromatic rings; three H atoms on carbons next to O; three H atoms on carbons next to N

■ Mass spectro**metry** is different from other forms of spectro**scopy** because it measures mass rather than the absorption of energy.

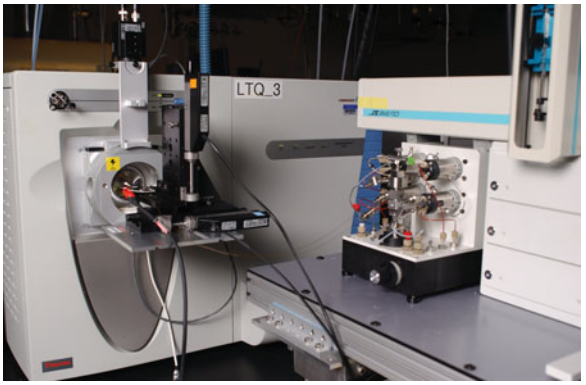
Mass spectrometry

Mass spectrometry weighs the molecule

It's not easy to weigh a neutral molecule, and a mass spectrometer works by measuring the mass of a charged ion instead: the charge makes the molecule controllable by an electric field. A mass spectrometer therefore has three basic components:

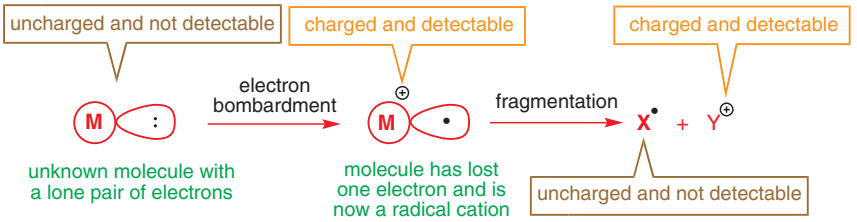
- something to volatilize and ionize the molecule into a beam of charged particles
- something to focus the beam so that particles of the same mass:charge ratio are separated from all others and
- something to detect the particles.

All spectrometers in common use operate in a high vacuum and use one of several methods to convert neutral molecules into cations, the most common being **electron impact**, **chemical ionization**, and **electrospray**.



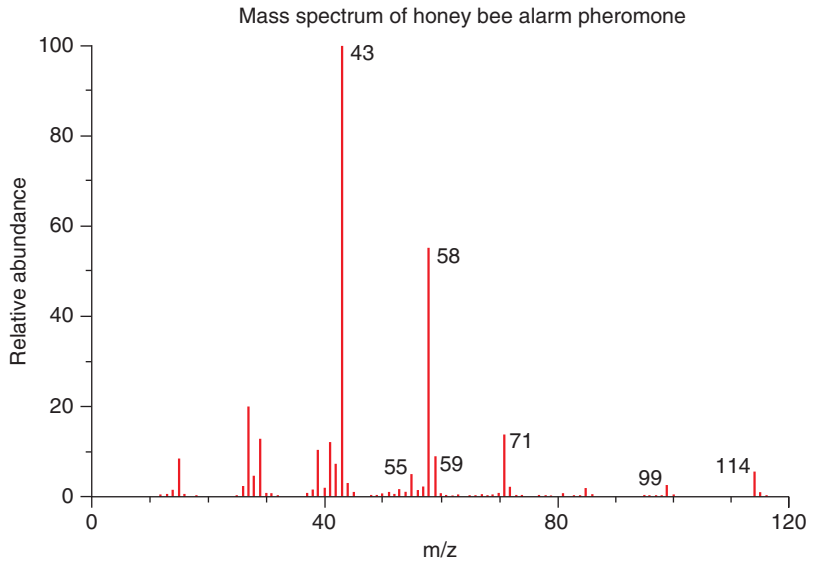
Mass spectrometry by electron impact

In electron impact (EI) mass spectrometry the molecule is bombarded with highly energetic electrons that knock a weakly bound electron out of the molecule. If you think this is strange, think of throwing bricks at a brick wall: the bricks can't stick to the wall but can knock loose bricks off the top of the wall. Losing a single electron leaves behind an unpaired electron and a positive charge. The electron that is lost will be one of relatively high energy (the bricks come from the *top* of the wall), and typically one not involved in bonding, for example an electron from a lone pair.



Thus ammonia gives $NH_3^+\bullet$ and a ketone gives $R_2C{=}O^+\bullet$. These unstable species are known as **radical cations**, and being charged they are accelerated by an electric field and focused onto the detector, which detects the mass of the ion by how far its path has been deflected by the electric field. It only takes about 20 µs for the radical cations to reach the detector, but sometimes they fragment before they get there, in which case other ions will also be detected. These fragments will always have a lower mass than the 'parent' molecular ion, so in a typical mass spectrum we are most interested in the heaviest ion we can see.

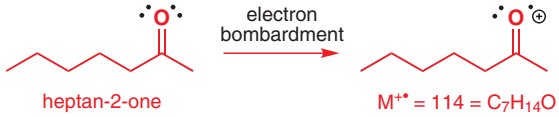
A typical EI mass spectrum looks like this:



Radical cations

Most molecules have all their electrons paired; **radicals** have unpaired electrons. Molecules that carry a negative charge are **anions;** molecules with a positive charge are **cations. Radical cations** and **radical anions** are simply species that are both charged *and* have an unpaired electron.

This compound was identified as a pheromone deposited by worker bees when feeding as a marker to deter their colleagues from visiting the same, now depleted, nectar source. Only minute quantities are available for analysis of course, but that doesn't matter: mass spectrometry is successful even on a microgram scale. The spectrum you see here indicates that the molecule has a mass of 114 because that is the highest mass observed in the spectrum: the molecule is in fact the volatile ketone heptan-2-one.



Mass spectrometry by chemical ionization, electrospray, or other methods

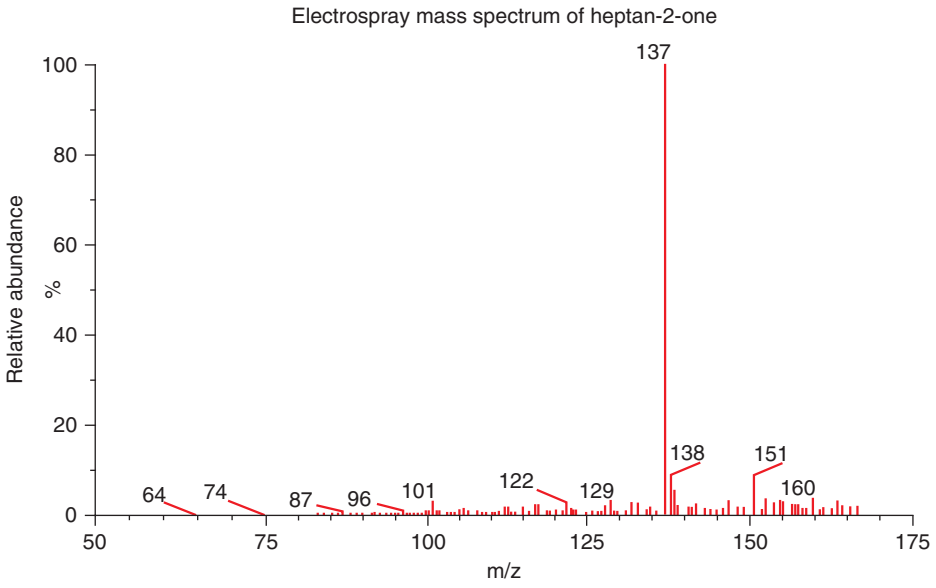
■ If you are interested in how to use fragmentation patterns to establish structure, you should consult one of the specialized textbooks in the bibliography at the end of this chapter.

■ We will not be discussing ionization techniques in detail: it is sufficient for you to realize at this stage that there are several ways of ionizing a molecule gently so that its mass can be determined.

A problem with EI mass spectrometry is that, for fragile molecules, the energy of the bombarding electron can be sufficient to cause it to fragment completely, losing all trace of the molecular ion. Some useful information can be gained from fragmentation patterns, but in general it is more useful to aim to weigh the molecule all in one piece. This can be achieved using any of a number of other techniques, of which the most common are chemical ionization (CI) and electrospray (ES).

Chemical ionization is achieved by mixing a gas such as ammonia with the substrate in the spectrometer. Bombardment of NH_3 with electrons leads to formation of some NH_4^+ by proton transfer, and reaction of this ion with the substrate makes a charged complex, which can be accelerated by the electric field. The masses observed by chemical ionization spectroscopy carried out in this way are usually M + 1 or M + 18 (the mass of NH_4^+) relative to the mass of the substrate. With electrospray mass spectroscopy, an aerosol of the substrate is ionized, and ionization in the presence of sodium ions means that masses of M + 1 and M + 23 are often seen, or, if the ionization forms anions, M – 1.

This is the electrospray mass spectrum of heptan-2-one. Notice how a single molecular ion is clearly visible this time, but that it has a mass of 137, which is 23 more than the mass of 114 (in other words, this is the mass of M + Na^+).

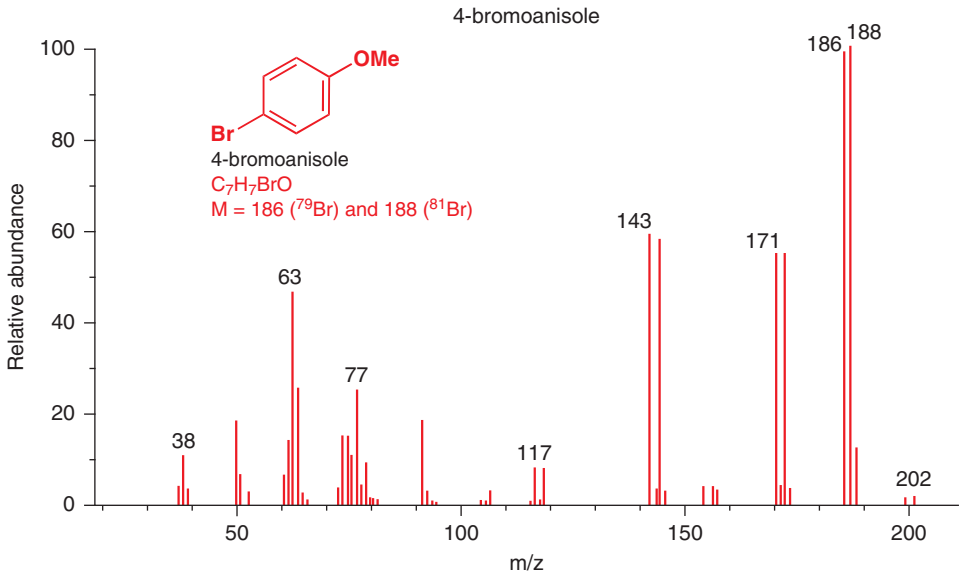


Mass spectrometry detects isotopes

Most elements can exist as more than one isotope. Usually, one isotope accounts for the vast majority (perhaps >99%) of the atoms of an element. But for some elements, atoms of several isotopes make up a significant proportion of the total in a sample. Chlorine, for example, is

normally a 3:1 mixture of ^{35}Cl and ^{37}Cl (hence the averaged relative atomic mass of 35.5 for chlorine), while bromine is an almost 1:1 mixture of ^{79}Br and ^{81}Br (hence the average mass of 80 for bromine). Because mass spectrometry weighs individual molecules, there is no averaging: instead it detects the true weight of each molecule, whatever isotope it contains.

For example, the molecular ion in the EI mass spectrum of this aryl bromide has two peaks at 186 and 188 of roughly equal intensity. Having two molecular ions of equal intensity separated by 2 mass units is indicative of bromine in a molecule.

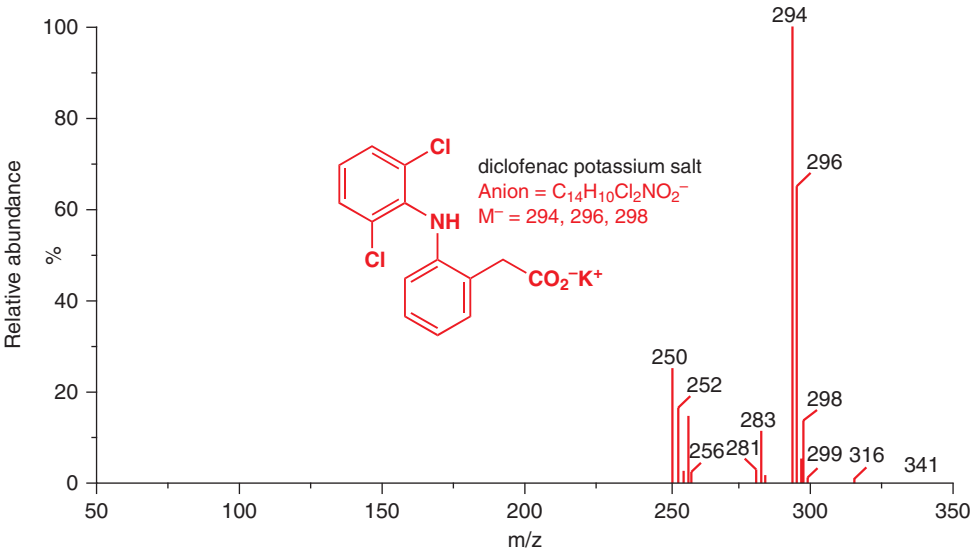


The mass spectrum of a chlorine-containing molecule is likewise easy to identify from two peaks separated by two mass units, but this time in a ratio of 3:1, arising from the 3:1 isotopic ratio of ^{35}Cl and ^{37}Cl.

What happens with more than one Br or Cl? Here's an example: the painkiller diclofenac. This spectrum was obtained from commercial tablets, which contain the potassium salt of the active ingredient (it becomes protonated in the acidic environment of the stomach).

■ Diclofenac behaves like soluble aspirin in this way: see Chapter 8, p. 163.

The ES spectrum shows the mass of the carboxylate anion as three peaks, at 294, 296, and 298. The relative size of the peaks can be worked out from the 75% probability that each Cl atom will be ^{35}Cl and the 25% probability it will be ^{37}Cl. The ratios are therefore $^3/_4 \times {}^3/_4 : 2 \times {}^3/_4 \times {}^1/_4 : {}^1/_4 \times {}^1/_4$ or 9 : 6 : 1.



Summary table of common elements with more than one isotope at >1% abundance

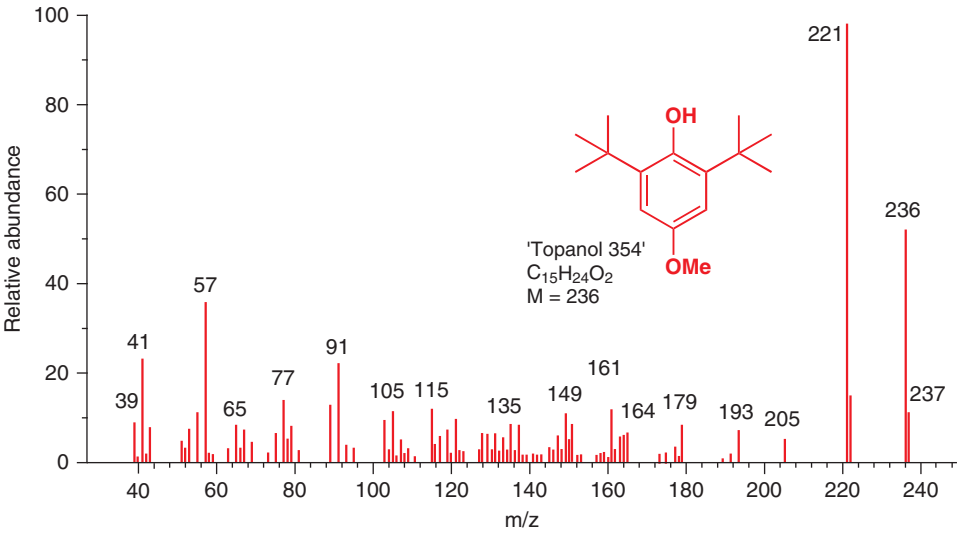
Element	Isotopes	Approximate ratio	Exact ratio
carbon	^{12}C, ^{13}C		98.9:1.1
chlorine	^{35}Cl, ^{37}Cl	3:1	75.8:24.2
bromine	^{79}Br, ^{81}Br	1:1	50.5:49.5

H, N, O, S, P, F, and I have only very small amounts of isotopes other than ^{1}H, ^{14}N, ^{16}O, ^{31}P, ^{32}S, and ^{128}I. The real oddity though is tin, which exists as a mixture of 10 different stable isotopes, the major ones being ^{116}Sn (15%), ^{117}Sn (8%), ^{118}Sn (24%), ^{119}Sn (9%), ^{120}Sn (33%), ^{122}Sn (5%), and ^{124}Sn (6%). In reality the precise ratio of isotopes for any element varies according to its source, a fact which can supply useful forensic information.

Carbon has a minor but important isotope ^{13}C

The minor isotopes of many elements that appear at below the 1% level are not usually important, but one we cannot ignore is the 1.1% of ^{13}C present in ordinary carbon, of which the main isotope is of course ^{12}C. Another isotope, ^{14}C, is radioactive and used in carbon dating, but its natural abundance is minute. The stable isotope ^{13}C is not radioactive, but it is *NMR active*, as we shall soon see. If you look back at all the mass spectra illustrated so far in this chapter, you will see a small peak one mass unit higher than each peak: these are peaks arising from molecules containing ^{13}C instead of ^{12}C. The exact height of these peaks is useful as an indication of the number of carbon atoms in the molecule. Each carbon has a 1.1% chance of being ^{13}C rather than ^{12}C, so the more C atoms there are the bigger this chance becomes. If there are n carbon atoms in a molecular ion, then the ratio of M^+ to $[M + 1]^+$ is $100:(1.1 \times n)$.

Look at the spectrum below: it's the fuel additive Topanol 354, whose structure and molecular formula are shown. With 15 carbons, there's a 16.5% chance there will be one ^{13}C atom in the molecule, and you can clearly see the sizeable M + 1 peak at 237. We can ignore the possibility of having *two* ^{13}C atoms as the probability is so small.



● For any mass spectrum, always look at the heaviest peak first: note whether there is chlorine or bromine in the molecule, and look to check that the ratio of M^+ to $[M + 1]^+$ is about right for the number of carbons you expect.

Atomic composition can be determined by high-resolution mass spectrometry

Ordinary mass spectra tell us the molecular weight (MW) of the molecule: we could easily see, for example, that the bee pheromone on p. 48 had MW 114 even without knowing its structure. When we revealed it was $C_7H_{14}O$, we had to use other information to infer this, because 114 could also be many other things, such as C_8H_{18} or $C_6H_{10}O_2$ or $C_6H_{14}N_2$. These different atomic compositions for the same molecular weight can nonetheless be distinguished if we know the *exact* molecular weight, since individual isotopes have non-integral masses (except ^{12}C by definition). The table below gives these masses to five decimal places, which is the sort of accuracy you need for meaningful results. Such accurate mass measurements are obtained by a technique called *high-resolution mass spectrometry*.

The reason that exact masses are not integers lies in the slight mass difference between a proton (1.67262×10^{-27} kg) and a neutron (1.67493×10^{-27} kg), and in the fact that electrons have mass (9.10956×10^{-31} kg).

Exact masses of common elements

Element	Isotope	Mass number	Exact mass
hydrogen	^{1}H	1	1.00783
carbon	^{12}C	12	12.00000
carbon	^{13}C	13	13.00335
nitrogen	^{14}N	14	14.00307
oxygen	^{16}O	16	15.99492
fluorine	^{19}F	19	18.99840
phosphorus	^{31}P	31	30.97376
sulfur	^{32}S	32	31.97207
chlorine	^{35}Cl	35	34.96886
chlorine	^{37}Cl	37	36.96590
bromine	^{79}Br	79	78.91835
bromine	^{81}Br	81	80.91635

For the bee pheromone on p. 48, the accurate mass turns out to be 114.1039. The table below compares possible atomic compositions for an approximate MW 114, and the result is conclusive. The exact masses to three places of decimals fit the observed exact mass only for the composition $C_7H_{14}O$. You may not think the fit is very good when you look at the two numbers, but notice the difference in the error expressed as parts per million. One answer stands out from the rest. Note that even two places of decimals would be enough to distinguish these four compositions.

Exact mass determination for the bee alarm pheromone

Composition	Calculated M^+	Observed M^+	Error in ppm
$C_6H_{10}O_2$	114.068075	114.1039	358
$C_6H_{14}N_2$	114.115693	114.1039	118
$C_7H_{14}O$	**114.104457**	**114.1039**	5
C_8H_{18}	114.140844	114.1039	369

● **In the rest of the book, whenever we state that a molecule has a certain atomic composition you can assume that it has been determined by high-resolution mass spectrometry on the molecular ion.**

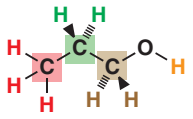
One thing you may have noticed in the table above is that there are no entries with just one nitrogen atom. Two nitrogen atoms, yes; one nitrogen no! This is because any complete molecule with C, H, O, S, and *just one nitrogen in it has an odd molecular weight*. This is because C, O, S, and N all have even atomic weights—only H has an odd atomic weight. Nitrogen is the only element from C, O, S, and N that can form an odd number of bonds (3). Molecules with one nitrogen atom must have an odd number of hydrogen atoms and hence an odd molecular weight.

● **Quick nitrogen count (for molecules containing any of the elements C, H, N, O, and S)**

Molecules with an odd molecular weight must have an odd number of nitrogen atoms. Molecules with even molecular weight must have an even number of nitrogen atoms or none at all.

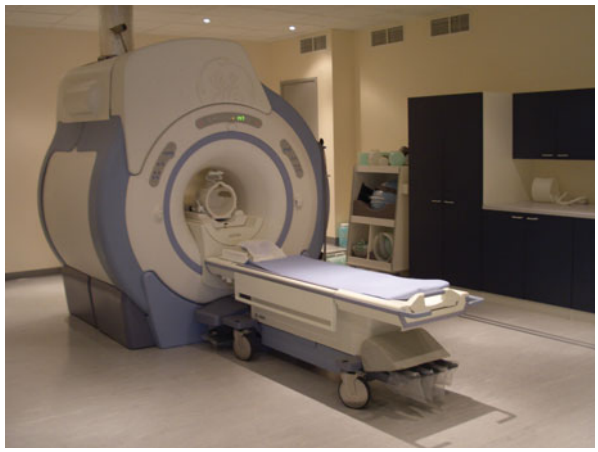
Nuclear magnetic resonance

What does it do?



^{1}H NMR distinguishes the coloured hydrogens
^{13}C NMR distinguishes the boxed carbons

Nuclear magnetic resonance (NMR) allows us to detect atomic nuclei and say what sort of environment they are in within the molecule. In a molecule such as propanol, the hydrogen atom of the hydroxyl group is clearly different from the hydrogen atoms of its carbon skeleton—it can be displaced by sodium metal, for example. NMR (actually ^{1}H, or proton, NMR) can easily distinguish between these two sorts of hydrogens by detecting the environment the hydrogen's nucleus finds itself in. Moreover, it can also distinguish between all the other different sorts of hydrogen atoms present. Likewise, carbon (more precisely ^{13}C) NMR can easily distinguish between the three different carbon atoms. NMR is extremely versatile: it can even scan living human brains (see picture) but the principle is still the same: being able to detect nuclei (and hence atoms) in different environments.



■ When NMR is used medically it is usually called magnetic resonance imaging (MRI) for fear of alarming patients wary of all things *nuclear*.

NMR uses a strong magnetic field

Imagine for a moment that we were able to 'switch off' the earth's magnetic field. Navigation would be made much harder since all compasses would be useless, with their needles pointing randomly in any direction. However, as soon as we switched the magnetic field back on, they would all point north—their lowest energy state. Now if we wanted to force a needle to point south we would have to use up energy and, of course, as soon as we let go, the needle would return to its lowest energy state, pointing north.

In a similar way, some atomic nuclei act like tiny compass needles when placed in a magnetic field and have different energy levels according to the direction in which they are 'pointing'. (We will explain how a nucleus can 'point' somewhere in a moment.) A real compass needle can rotate through 360° and have an essentially infinite number of different energy levels, all higher in energy than the 'ground state' (pointing north). Fortunately, things are simpler with an atomic nucleus: its energy levels are quantized, just like the energy levels of an electron, which you will meet in the next chapter, and it can adopt only certain specific energy levels. This is like a compass which points, say, only north or south, or maybe only north, south, east, or west, and nothing in between. Just as a compass needle has to be made of a magnetic material to feel the effect of the earth's magnetism, so it is that only certain nuclei are 'magnetic'. Many (including 'normal' carbon-12, ^{12}C) do not interact with a magnetic field at all and cannot be observed in an NMR machine. But, importantly for us in this chapter, the minor carbon isotope ^{13}C does display magnetic properties, as does ^{1}H, the most abundant atomic nucleus on earth. When a ^{13}C or ^{1}H atom finds itself in a magnetic field, it has two available energy states: it can either align itself with the field ('north' you could say), which would be the lowest energy state, or against the field ('south'), which is higher in energy.

This picture shows a typical NMR instrument. The fat cylinder is the supercooled magnet. The device hanging over it is an automatic sample changer and the console in the foreground controls the machine.

The property of a nucleus that allows magnetic interactions, i.e. the property possessed by ^{13}C and ^{1}H but not by ^{12}C, is *spin*. If you conceive of a ^{13}C and ^{1}H nucleus spinning, you can see how the nucleus can point in one direction—it is the axis of the spin that is aligned with or against the field.

Let's return to the compass for a moment. If you want to move a compass needle away from pointing north, you have to push it—and expend energy as you do so. If you put the compass next to a bar magnet, the attraction towards the magnet is much greater than the attraction towards the north pole, and the needle now points at the magnet. You also have to push much harder if you want to move the needle. Exactly how hard it is to turn the compass needle depends on how strong the magnetic field is and also on how well the needle is magnetized—if it is only weakly magnetized, it is much easier to turn it round and if it isn't magnetized at all, it is free to rotate.

Likewise, for a nucleus in a magnetic field, the difference in energy between the nuclear spin aligned with and against the applied field depends on:

- how strong the magnetic field is, and
- the magnetic properties of the nucleus itself.

The stronger the magnetic field, the greater the energy difference between the two alignments of the nucleus. Now there is an unfortunate thing about NMR: the energy difference between the nuclear spin being aligned with the magnetic field and against it is really *very* small—so small that we need a very, very strong magnetic field to see any difference at all.

Nuclear spin is quantized and has the symbol *I*. The exact number of different energy levels a nucleus can adopt is determined by the value of I of the particular isotope. The nuclear spin I can have various values such as 0, $^{1}/_{2}$, 1, $^{3}/_{2}$... and the number of energy levels is given by $2I + 1$. Some examples are ^{1}H, $I = {}^{1}/_{2}$; ^{2}H (= D), $I = 1$; ^{11}B, $I = {}^{5}/_{2}$; ^{12}C, $I = 0$.

NMR machines contain very strong electromagnets. The earth's magnetic field has a field strength of between 30 and 60 microtesla. A typical magnet used in an NMR machine has a field strength of between 2 and 10 tesla, some 10^5 times stronger than the earth's field. These magnets are dangerous and no metal objects must be taken into the rooms where they are: stories abound of unwitting workmen whose metal toolboxes have become firmly attached to NMR magnets. Even with the immensely powerful magnets used the energy difference is still so small that the nuclei only have a very small preference for the lower energy state. Fortunately, we can just detect this small preference.

NMR also uses radio waves

A ^{1}H or ^{13}C nucleus in a magnetic field can have two energy levels, and energy is needed to flip the nucleus from the more stable state to the less stable state. But since the amount of energy needed is so small, it can be provided by low-energy electromagnetic radiation of radio-wave frequency. Radio waves flip the nucleus from the lower energy state to the higher state. Turn off the radio pulse and the nucleus returns to the lower energy state. When it does so, the energy comes out again, and this (a tiny pulse of radio frequency electromagnetic radiation) is what we detect.

We can now sum up how an NMR machine works.

1. The sample of the unknown compound is dissolved in a suitable solvent, placed in a narrow tube, and put inside a very strong electromagnet. To even out imperfections in

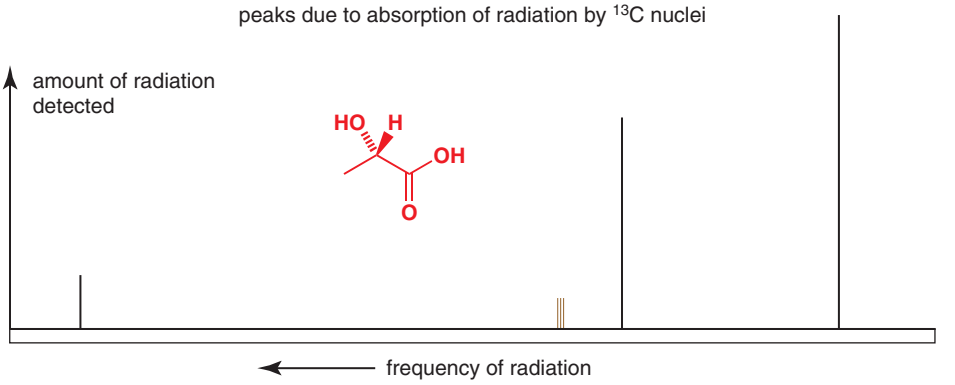
Radio waves are very, very low in energy. You may know—and if not, you will need to in the future—that the energy associated with electromagnetic radiation is related to its wavelength λ by the formula:

$$E = hc/\lambda$$

where h and c are constants (Planck's constant and the speed of light). Radio waves, whose wavelengths are measured in metres, are millions of times less energetic than rays of visible light, with wavelengths between 380 nm (violet) and 750 nm (red).

the sample, the tube is spun very fast by a stream of air. Inside the magnetic field, any atomic nuclei with a nuclear spin now possess different energy levels, the exact number of different energy levels depending on the value of the nuclear spin. For ^{1}H and ^{13}C NMR there are two energy levels.

2. The sample is irradiated with a short pulse of radiofrequency energy. This disturbs the equilibrium balance between the two energy levels: some nuclei absorb the energy and are promoted to a higher energy level.
3. When the pulse finishes, the radiation given out as the nuclei fall back down to the lower energy level is detected using what is basically a sophisticated radio receiver.
4. After lots of computation, the results are displayed in the form of intensity (i.e. number of absorptions) against frequency. Here is an example, which we shall return to in more detail later:

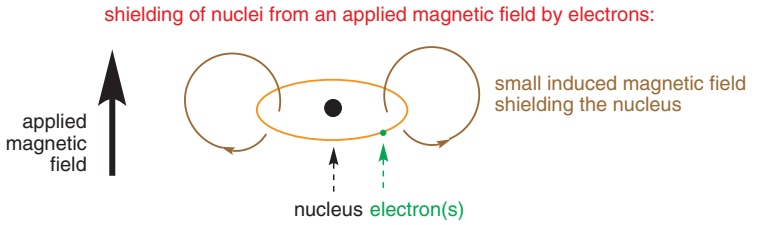


■ 'Resonance' is a good analogy here. If you find a piano and hold down a key to release a single string then give the piano lid a good thwack, you will hear the note you are holding down, and only that note, continuing to sound—it resonates. The thwack provides the piano with sound energy of a range of frequencies, but only sound energy with the right frequency is absorbed and then re-emitted by the vibrating string. There is another chemical use of the word resonance, mentioned in Chapter 7, which is much less appropriate: the two have nothing to do with one another.

Why do chemically distinct nuclei absorb energy at different frequencies?

In the spectrum you see above, each peak represents a different kind of carbon atom: each one absorbs energy (or resonates—hence the term 'nuclear magnetic *resonance*') at a different frequency. But why should carbon atoms be 'different'? We have told you two factors that affect the energy difference (and therefore the frequency)—the magnetic field strength and what sort of nucleus is being studied. So you might expect all ^{13}C nuclei to resonate at one particular frequency and all protons (^{1}H) to resonate at one (different) frequency. But they don't.

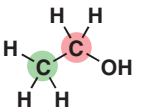
The variation in frequency for different carbon atoms must mean that the energy jump from 'nucleus-aligned-with' to 'nucleus-aligned-against' the applied magnetic field must be different for each type of carbon atom. The reason is that the ^{13}C nuclei in question experience a magnetic field that is not quite the same as the magnetic field that we apply. Each nucleus is surrounded by electrons, and in a magnetic field these will set up a tiny electric current. This current will set up its own magnetic field (rather like the magnetic field set up by the electrons of an electric current moving through a coil of wire or solenoid), which will oppose the magnetic field that we apply. The electrons are said to **shield** the nucleus from the external magnetic field. If the electron distribution varies from ^{13}C atom to ^{13}C atom, so does the local magnetic field experienced by its nucleus, and so does the corresponding resonating frequency.



- **Changes in the distribution of electrons around a nucleus affect:**
  - **the local magnetic field that the nucleus experiences**
  - **the frequency at which the nucleus resonates**
  - **the chemistry of the molecule at that atom**

This variation in frequency is known as the chemical shift. Its symbol is δ.

As an example, consider ethanol (right). The red carbon attached to the OH group will have a smaller share of the electrons around it compared to the green carbon since the oxygen atom is more electronegative and pulls electrons towards it, away from the red carbon atom.

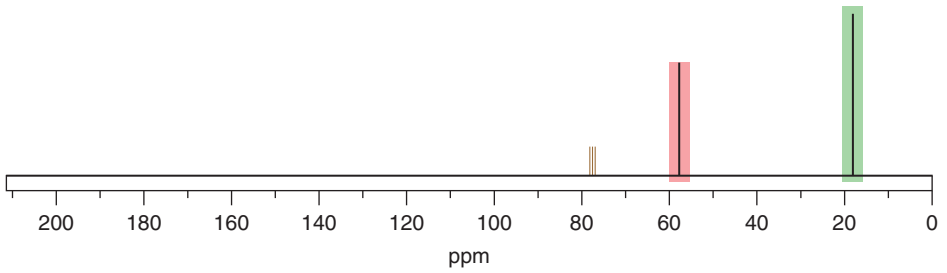


ethanol

The magnetic field that the red carbon nucleus feels will therefore be slightly greater than that felt by the green carbon, which has a greater share of the electrons, since the red carbon is less shielded from the applied external magnetic field—in other words it is **deshielded**. Since the carbon attached to the oxygen feels a stronger magnetic field (it is more 'exposed' to the field as it has lost some of its electronic shielding) there will be a greater energy difference between the two alignments of its nucleus. The greater the energy difference, the higher the resonant frequency (energy is proportional to frequency). So for ethanol we would expect the red carbon with the OH group attached to resonate at a higher frequency than the green carbon, and indeed this is exactly what the ^{13}C NMR spectrum shows.

■ We wouldn't usually draw all the Cs and Hs of course, but we have done so here because we want to talk about them.

^{13}C NMR spectrum of ethanol

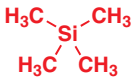


■ The peaks at 77 ppm, shaded brown, are those of the usual solvent ($CDCl_3$) and can be ignored for the moment. We shall explain them in Chapter 13.

The chemical shift scale

When you look at a real NMR spectrum you will see that the scale does not appear to be in magnetic field units, nor in frequency, nor yet even energy, units, but in 'parts per million' (ppm). There is a very good reason for this. The exact frequency at which the nucleus resonates depends on the external applied magnetic field. This means that if the sample is run on a machine with a different magnetic field, it will resonate at a different frequency. It would make life very difficult if we couldn't say exactly where our signal was, so we say how far it is from some reference sample, as a fraction of the operating frequency of the machine. We know that all protons resonate at approximately the same frequency in a given magnetic field and that the *exact* frequency depends on what sort of chemical environment it is in, which in turn depends on its electrons. This approximate frequency is the operating frequency of the machine and simply depends on the strength of the magnet—the stronger the magnet, the larger the operating frequency. The precise value of the operating frequency is simply the frequency at which a standard reference sample resonates. In everyday use, rather than actually referring to the strength of the magnet in tesla, chemists usually just refer to its operating frequency. A 9.4 T NMR machine is referred to as a 400 MHz spectrometer since that is the frequency in this strength field at which the protons in the reference sample resonate; other nuclei, for example ^{13}C, would resonate at a different frequency, but the strength is arbitrarily quoted in terms of the proton operating frequency.

The reference sample—tetramethylsilane, TMS



tetramethylsilane, TMS

■ Silicon and oxygen have opposite effects on an adjacent carbon atom: electropositive silicon shields; electronegative oxygen deshields. Electronegativities: Si: 1.8; C: 2.5; O: 3.5.

The compound we use as a reference sample is usually tetramethylsilane, TMS. This is silane (SiH_4) with each of the hydrogen atoms replaced by methyl groups to give $Si(CH_3)_4$. The four carbon atoms attached to silicon are all equivalent and, because silicon is more electropositive than carbon, they are fairly electron-rich (or *shielded*), which means they resonate at a frequency a little less than that of most organic compounds. This is useful because it means our reference sample is not bang in the middle of our spectrum!

The chemical shift, δ, in parts per million (ppm) of a given nucleus in our sample is defined in terms of the resonance frequency as:

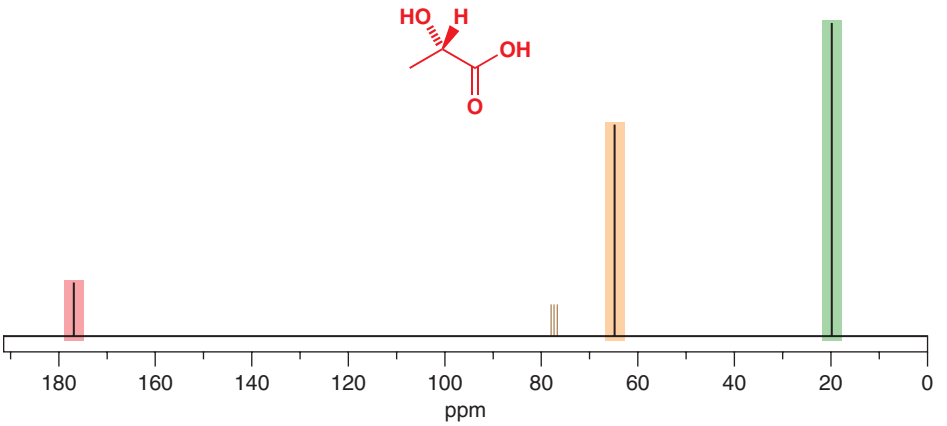
$$\delta = \frac{\text{frequency (Hz)} - \text{frequency TMS (Hz)}}{\text{frequency TMS (MHz)}}$$

No matter what the operating frequency (i.e. strength of the magnet) of the NMR machine, the signals in a given sample (e.g. ethanol) will always occur at the same chemical shifts. In ethanol the (red) carbon attached to the OH resonates at 57.8 ppm whilst the (green) carbon of the methyl group resonates at 18.2 ppm. Notice that by definition TMS itself resonates at 0 ppm. The carbon nuclei in most organic compounds resonate at greater chemical shifts, normally between 0 and 200 ppm.

Now, let's return to the sample spectrum you saw on p. 54 and which is reproduced below, and you can see the features we have discussed. This is a 100 MHz spectrum; the horizontal axis is actually frequency but is usually quoted in ppm of the field of the magnet, so each unit is one ppm of 100 MHz, that is, 100 Hz. We can tell immediately from the three peaks at 176.8, 66.0, and 19.9 ppm that there are three different types of carbon atom in the molecule.

■ Again, ignore the brown solvent peaks at 77ppm—they are of no interest to us at the moment. You also need not worry about the fact that the signals have different intensities. This is a consequence of the way the spectrum was recorded and in ^{13}C spectra signal intensity is usually of no consequence.

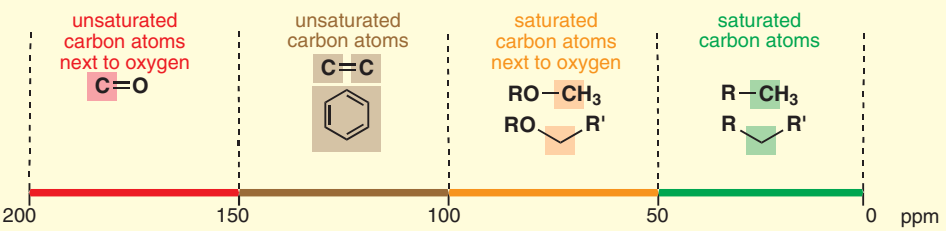
^{13}C NMR spectrum of lactic acid



Regions of the ^{13}C NMR spectrum

But we can do better than this: we can also work out what sort of chemical environment the carbon atoms are in. All ^{13}C spectra can be divided into four major regions: saturated carbon atoms (0–50 ppm), saturated carbon atoms next to oxygen (50–100 ppm), unsaturated carbon atoms (100–150 ppm), and unsaturated carbon atoms next to oxygen, i.e. C=O groups (150 to about 200 ppm).

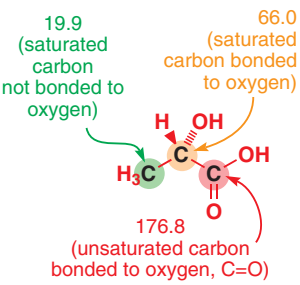
● **Regions of the ^{13}C NMR spectrum**



The spectrum you just saw is in fact that of lactic acid (2-hydroxypropanoic acid). When you turned the last page, you made some lactic acid from glucose in the muscles of your arm—it is the breakdown product from glucose when you do anaerobic exercise. Each of lactic acid's carbon atoms gives a peak in a different region of the spectrum.

But hang on one moment, you may say—don't we only see signals for carbon-13 nuclei and not carbon-12, which make up most of the carbon atoms in any normal sample of lactic acid? The answer is yes, and indeed only about 1.1% (the natural abundance of ^{13}C) of the C atoms in any sample are 'visible' by ^{13}C NMR. But since those ^{13}C atoms will be distributed more or less randomly through the sample, this fact does not affect any of the arguments about the appearance of the spectrum. What it does mean, however, is that ^{13}C NMR is not as sensitive as ^{1}H NMR, for example, where essentially all of the H atoms in the sample will be 'visible'.

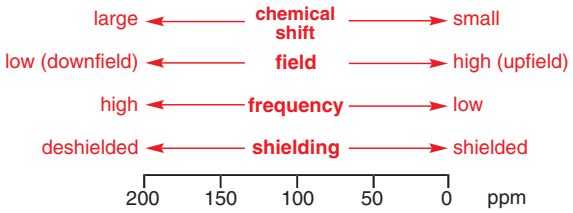
lactic acid
(2-hydroxypropanoic acid)



Different ways of describing chemical shift

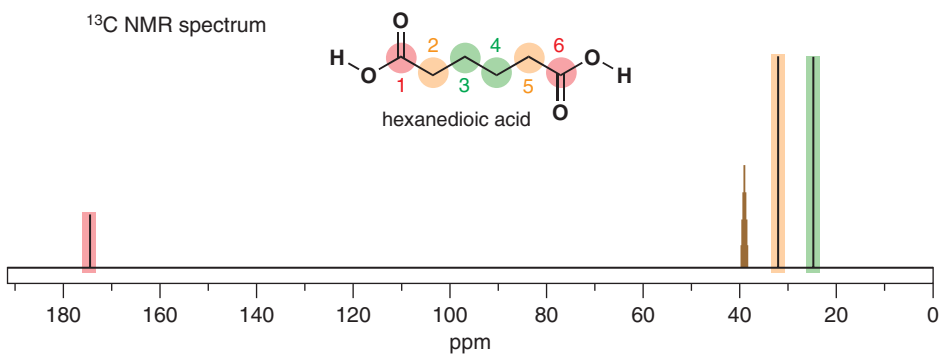
The chemical shift scale runs to the left from zero (where TMS resonates), i.e. backwards from the usual style. Chemical shift values around zero are obviously small but are confusingly called 'high field' because this is the high magnetic field end of the scale. We suggest you say 'large' or 'small' chemical shift and 'large' or 'small' δ, but 'high' or 'low' field to avoid confusion. Alternatively, use 'upfield' for high field (small δ) and 'downfield' for low field (large δ).

One helpful description we have already used is shielding. Each carbon nucleus is surrounded by electrons that shield the nucleus from the applied field. Simple saturated carbon nuclei are the most shielded: they have small chemical shifts (0–50 ppm) and resonate at high field. One electronegative oxygen atom moves the chemical shift downfield into the 50–100 ppm region. The nucleus has become deshielded. Unsaturated carbon atoms experience even less shielding (100–150 ppm) because of the way in which electrons are distributed around the nucleus. If they are also bonded to oxygen (the most common unsaturated carbons bonded to oxygen are those of carbonyl groups), then the nucleus is even more deshielded and moves to the largest chemical shifts around 200 ppm. The next diagram summarizes these different ways of talking about NMR spectra.



A guided tour of the ^{13}C NMR spectra of some simple molecules

So, on to some real ^{13}C NMR spectra. Our very first compound, hexanedioic acid, has the simple NMR spectrum shown here. The first question is: why only three peaks for six carbon atoms? Because of the symmetry of the molecule, the two carboxylic acids are identical and give one peak at 174.2 ppm. By the same token C2 and C5 are identical, and C3 and C4 are identical. These are all in the saturated region 0–50 ppm but the carbons next to the electron-withdrawing CO_2H group will be more deshielded than the others. So we assign C2/C5 to the peak at 33.2 ppm and C3/C4 to 24.0 ppm.



■ In fact, the low abundance of ^{13}C in natural carbon makes ^{13}C spectra simpler than they would otherwise be—we go into this in more detail in Chapter 13.

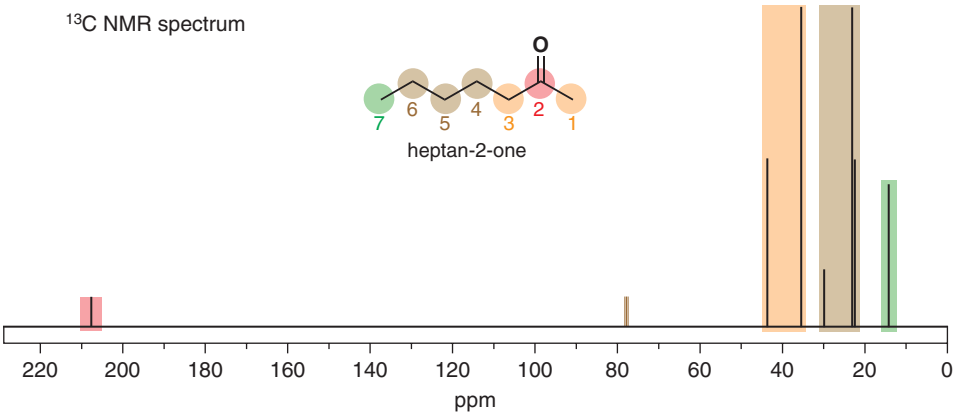
■ NMR spectra were originally recorded by varying the applied field. They are now recorded by variation of the frequency of the radio waves and that is done by a pulse of radiation. The terms 'high field' and 'low field' are a relic from the days of scanning by field variation.

If you are coming back to this chapter after reading Chapter 4 you might like to know that unsaturated C atoms are more deshielded than saturated ones because a π bond has a *nodal plane*, i.e. a plane with no electron density in at all. Electrons in π bonds are less efficient at shielding the nucleus than electrons in σ bonds.

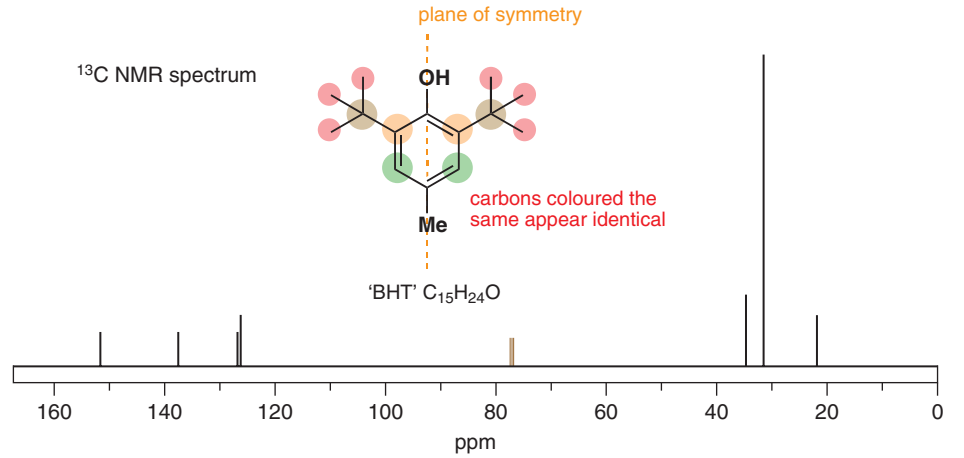
■ Why isn't this compound called 'hexane-1,6-dioic acid'? Well, carboxylic acids can only be at the end of chains, so no other hexanedioic acids are possible: the 1 and 6 are redundant.

■ This spectrum was run in a different solvent, DMSO (dimethylsulfoxide), hence the brown solvent peaks are in a different region and have a different form. Again, we will deal with these in Chapter 13.

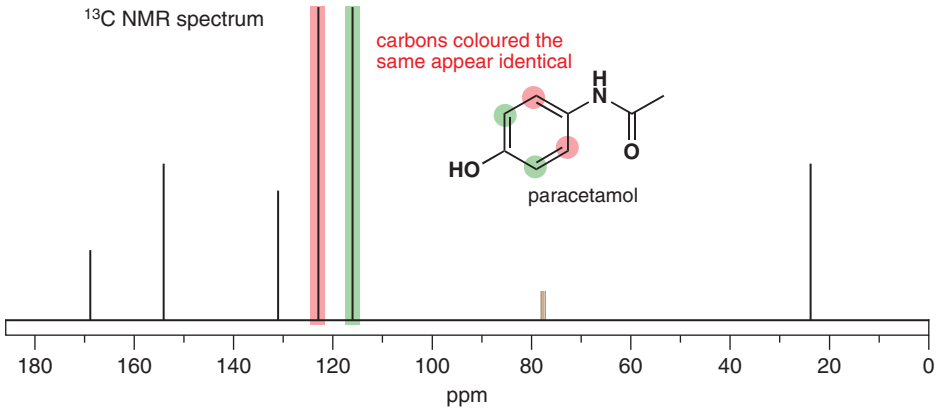
Heptan-2-one is the bee pheromone mentioned on p. 48. It has no symmetry so all its seven carbon atoms are different. The carbonyl group is easy to identify (208.8 ppm) but the rest are more difficult. The two carbon atoms next to the carbonyl group come at lowest field, while C7 is at highest field (13.9 ppm). It is important that there is the right number of signals at about the right chemical shift. If that is so, we are not worried if we cannot assign each frequency to a precise carbon atom (such as atoms 4, 5, and 6, for example). As we said before, don't be concerned with the *intensities* of the peaks.



You met BHT on p. 8: its formula is $C_{15}H_{24}O$ and the first surprise in its NMR spectrum is that there are only seven signals for the 15 carbon atoms. There is obviously a lot of symmetry; in fact the molecule has a plane of symmetry vertically as it is drawn here, and the coloured blobs indicate pairs or groups of carbons related to each other by symmetry which therefore give only one signal. The very strong signal at $\delta = 30.4$ ppm belongs to the six identical methyl groups on the *t*-butyl groups (coloured red) and the other two signals in the 0–50 ppm range are the methyl group at C4 and the brown central carbons of the *t*-butyl groups. In the aromatic region there are only four signals as the two halves of the molecule are the same. As with the last example, we are not concerned with exactly which is which—we just check that there are the right number of signals with the right chemical shifts.

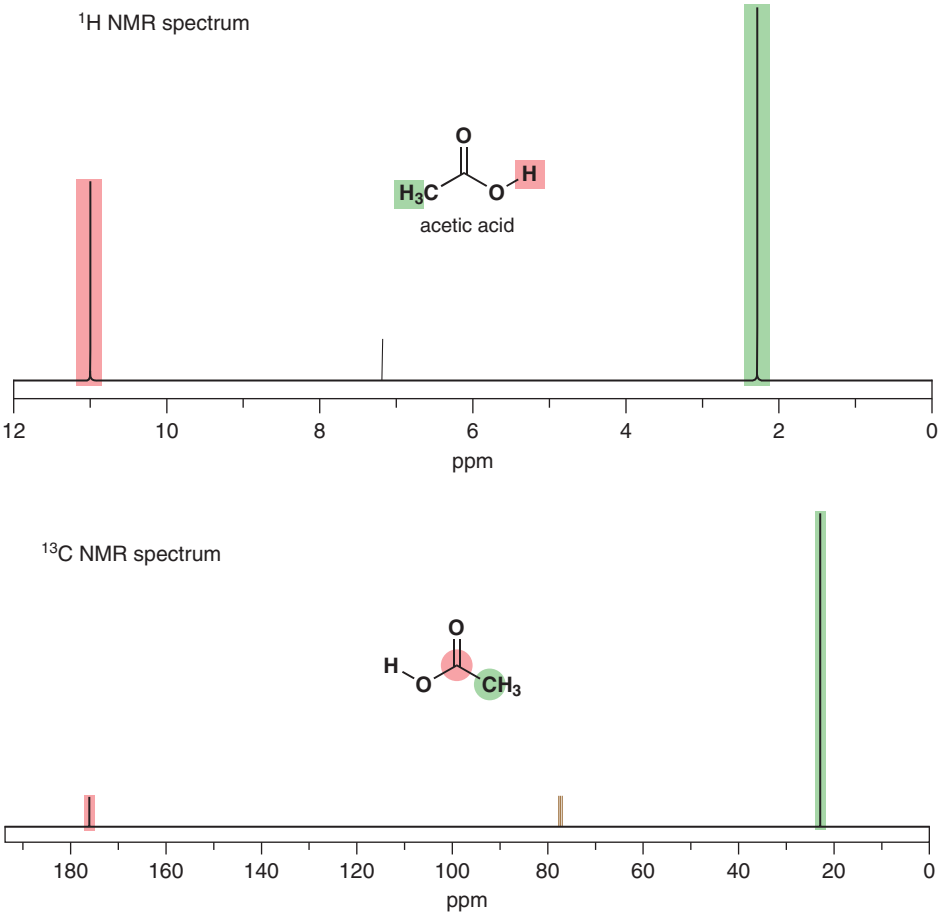


Paracetamol is a familiar painkiller with a simple structure—it too is a phenol but in addition it carries an amide substituent on the benzene ring. Its NMR spectrum contains one saturated carbon atom at 24 ppm (the methyl group of the amide side chain), one carbonyl group at 168 ppm, and four other peaks at 115, 122, 132, and 153 ppm. These are the carbons of the benzene ring. Why four peaks? The two halves of the benzene ring must be the same (only one signal for each pair of carbons coloured red and green), which tells us that the $NHCOCH_3$ group doesn't really lie just to one side as shown here, but rotates rapidly, meaning that *on average* the two sides of the ring are indistinguishable, as in BHT. Why is one of these aromatic peaks in the C=O region at 153 ppm? This must be C4 because it is bonded to oxygen, a reminder that carbonyl groups are not the only unsaturated carbon atoms bonded to oxygen (see the chart on p. 56), although it is not as deshielded as the true C=O group at 168 ppm.



The ^{1}H NMR spectrum

^{1}H NMR (or 'proton NMR') spectra are recorded in the same way as ^{13}C NMR spectra: radio waves are used to study the energy level differences of nuclei, but this time they are ^{1}H and not ^{13}C nuclei. Like ^{13}C, ^{1}H nuclei have a nuclear spin of 1/2 and so have two energy levels: they can be aligned either with or against the applied magnetic field. Here, as an example, is the ^{1}H NMR spectrum of acetic (ethanoic) acid, $MeCO_2H$, and below it the ^{13}C NMR spectrum.



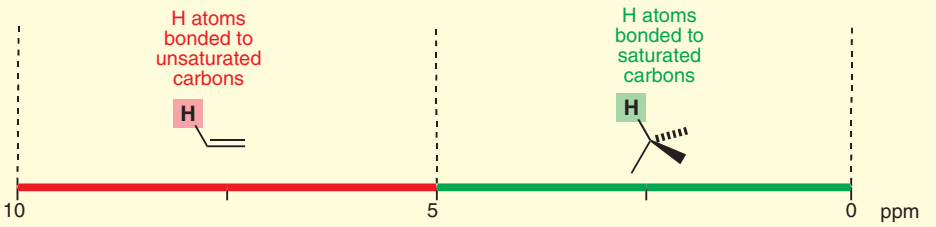
■ The brown peak at 7.25 ppm is a solvent peak and can be ignored.

^{1}H NMR spectra have many similarities with ^{13}C NMR spectra: the scale runs from right to left and the zero point is given by the same reference compound, though it is the proton resonance of Me_4Si rather than the carbon resonance that defines the zero point. However, as you immediately see in the spectrum above, the scale is much smaller, ranging over only about 10 ppm instead of the 200 ppm needed for carbon. This makes perfect sense: the variation in the chemical

shift is a measure of the shielding of the nucleus by the electrons around it. There is inevitably less change possible in the distribution of two electrons around a hydrogen nucleus than in that of the eight valence electrons around a carbon nucleus. Nonetheless the acetic acid spectrum above shows you that, just as you would expect, the H atom of the carboxylic acid group, directly attached to an oxygen atom, is more deshielded than the H atoms of acetic acid's methyl group.

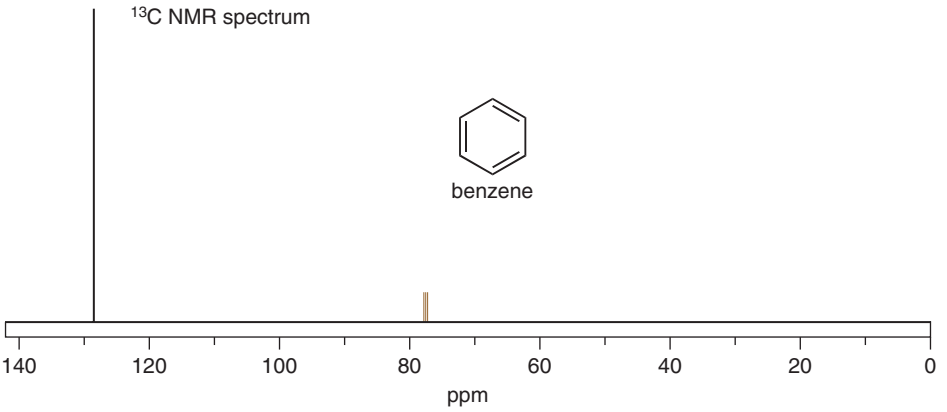
We can also divide up the ^{1}H NMR spectrum into regions that parallel the regions of the ^{13}C NMR spectrum. Hydrogen atoms bonded to saturated carbon atoms appear in the right-hand, more shielded (between 5 and 0 ppm) region of the spectrum, while those bonded to unsaturated carbon atoms (alkenes, arenes, or carbonyl groups primarily) appear in the left-hand, less shielded region between 10 and 5 ppm. As with the ^{13}C spectrum, nearby oxygen atoms withdraw electron density and make the signals appear towards the left-hand end of each of these regions.

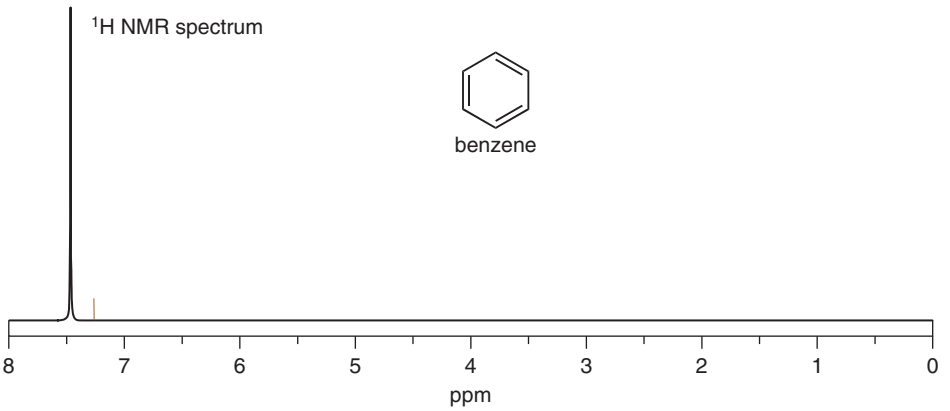
Regions of the ^{1}H NMR spectrum

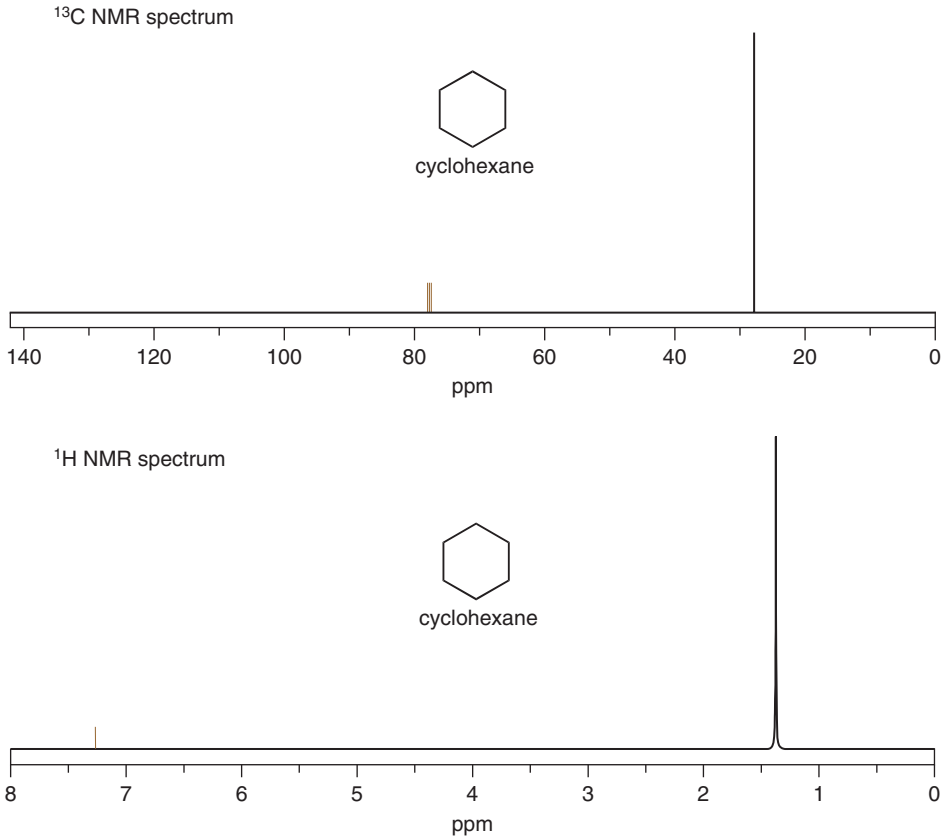


Some examples of ^{1}H NMR spectra

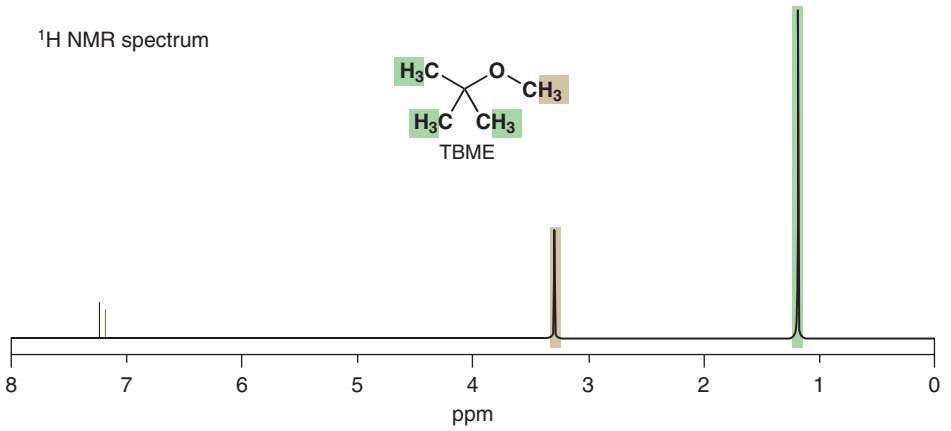
You can see exactly how ^{1}H NMR signals fall into these regions in the following collection of spectra. The first two spectra each contain only one peak because every proton in benzene and in cyclohexane is identical. In benzene the peak is at 7.5 ppm, where we expect a proton attached to an unsaturated C atom to lie, while in cyclohexane it is at 1.35 ppm because all the cyclohexane protons are attached to saturated C atoms. Again, to help comparisons, we have also included the ^{13}C spectra of benzene and cyclohexane. For benzene, the signal falls in the unsaturated C region (100–150 ppm), at 129 ppm, while for cyclohexane it is in the saturated C region, at 27 ppm.



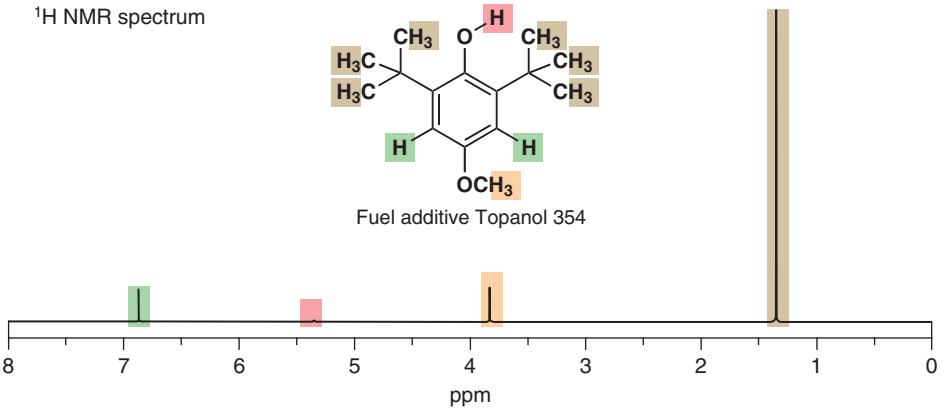




tert-Butyl methyl ether is a solvent and fuel additive whose ^{1}H spectrum illustrates the effect of a nearby oxygen atom: the large peak at 1.1 ppm comes from the nine H atoms making up three identical methyl groups of the *tert*-butyl part of the molecule, while the three H atoms of the methyl part of the ether are at 3.15 ppm. These three hydrogen atoms are all bonded directly to a C atom, which itself is bonded to O, whose electronegativity attracts their electrons, deshielding the ^{1}H nuclei and shifting them to larger chemical shift.



The plane of symmetry we noted in the ^{13}C NMR spectrum of BHT means that the ^{1}H NMR spectrum of the related compound Topanol 354 is relatively simple for a compound with 26 H atoms: a large peak and two small peaks between 5 and 0 ppm for the 18 protons of the *tert*-butyl groups and the three protons of each methyl group, and another small peak between 5 and 10 ppm for the two protons attached to the aromatic ring.

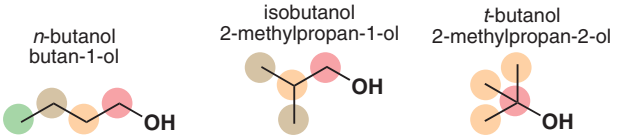


^{1}H NMR has many more features, which we will leave aside for the moment, and it is no exaggeration to say it is in general more important for the routine determination of structure than all the other methods put together. We will come back to ^{1}H NMR in more detail in Chapter 13.

NMR is a powerful tool for solving unknown structures

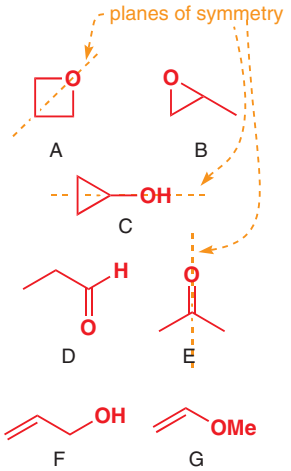
To illustrate the power of NMR, consider these three alcohols of formula $C_4H_{10}O$, each of which has a quite different ^{13}C NMR spectrum. Peaks from the spectra are shown in the table below.

➡ The meanings of *n*-, *iso*-, and *tert*- were covered in Chapter 2 (p. 26).



Chemical shift (δ, ppm)

Carbon atom	*n*-butanol	isobutanol	*tert*-butanol
	62.9	70.2	69.3
	36.0	32.0	32.7
	20.3	20.4	–
	15.2	–	–



Each alcohol has a saturated carbon atom next to oxygen, all appearing in the region typical of saturated carbon atoms next to oxygen (p. 56). Then there are carbons next door but one to oxygen: they are back in the 0–50 ppm region but at its low field end—about 30–35 ppm—because they are still deshielded by the nearby oxygen atom. Two of the alcohols have carbon(s) one further away still at yet smaller chemical shift (further upfield, more shielded) at about 20 ppm, but only the *n*-butanol has a more remote carbon still at 15.2. The *number* and the *chemical shift* of the signals identify the molecules very clearly.

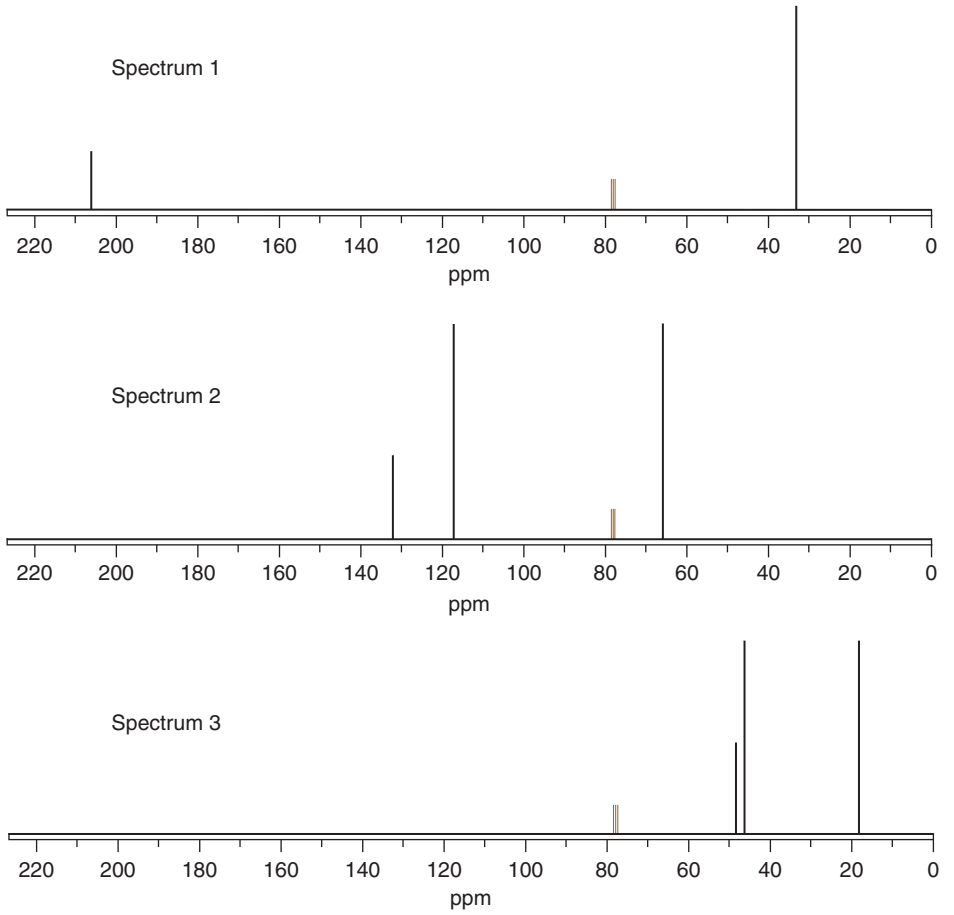
A common situation chemists find themselves in is that they have some idea about a molecular formula—from high-resolution mass spectrometry, for example—and need to match a structure to NMR data. Here's an example: the formula C_3H_6O is represented by seven reasonable structures, as shown in the margin. The three ^{13}C NMR spectra below represent three of these compounds. The challenge is to identify which three. We will give you some clues, and then we suggest you try to work out the answer for yourself before turning the page.

Simple symmetry can distinguish structures A, C, and E from the rest as these three have only two types of carbon atom. The two carbonyl compounds, D and E, will have one peak in the 150–200 ppm region but D has two different saturated carbon atoms while E has only one. The two alkenes, F and G, both have two unsaturated carbon atoms (100–200 ppm) but in ether G one of them is joined to oxygen—you would expect it therefore to be deshielded and to appear between 150 and 200 ppm.

The three saturated compounds (A, B, and C) present the greatest problem. The epoxide, B, has two different carbon atoms next to oxygen (50–100 ppm) and one normal saturated carbon atom (0–50 ppm). The remaining two both have one signal in the 0–50 ppm region and one in the 50–100 ppm region, and only the more powerful techniques of ^{1}H NMR and, to a certain extent, infrared spectroscopy (which we will move on to shortly) will distinguish them reliably.

■ An epoxide is a three-membered cyclic ether, such as B.

Here are NMR spectra of three of these molecules. Before reading further see if you can assign them to the structures on the previous page. Try also to suggest which signals belong to which carbon atoms.

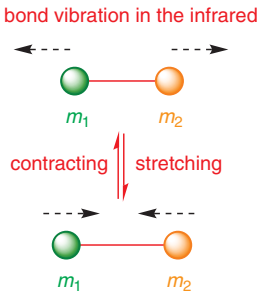


We hope these didn't give you too much trouble. The only carbonyl compound with two identical carbons is acetone (E) so spectrum 1 must be that one. Notice the very low field signal (206.6 ppm) typical of a simple ketone C=O carbon atom. Spectrum 2 has two unsaturated carbons and a saturated carbon next to oxygen so it must be F or G. In fact it has to be F as both unsaturated carbons are similar (137 and 116 ppm) and neither is next to oxygen (>150 ppm). This leaves spectrum 3, which appears to have no carbon atoms next to oxygen as all chemical shifts are less than 50 ppm. No compound fits that description and the two signals at 48.0 and 48.2 ppm are suspiciously close to the arbitrary 50 ppm borderline. They are, of course, both next to oxygen and this is compound B.

Infrared spectra

Functional groups are identified by infrared spectra

^{13}C and ^{1}H NMR spectra tell us a lot about the hydrocarbon skeleton of a molecule, and mass spectroscopy weighs the molecule as a whole. But none of these techniques reveal much about functional groups. Some functional groups, for example C=O or C=C, can be seen in the ^{13}C NMR spectrum because they contain carbon atoms, but many, such as ethers or nitro groups, cannot be seen at all by NMR—they show their presence only by the way they affect the chemical shifts of nearby H or C atoms.



Infrared (IR) spectroscopy, however, provides a direct way of observing these functional groups because it detects the stretching and bending of bonds rather than any property of the atoms themselves. It is particularly good at detecting the stretching of unsymmetrical bonds of the kind found in functional groups such as OH, C=O, NH_2, and NO_2, and for this reason IR spectroscopy complements NMR beautifully as a method for structural analysis.

NMR requires electromagnetic waves in the radio-wave region of the spectrum to make nuclei flip from one state to another. The amount of energy needed for stretching and bending individual bonds, while still very small, is rather greater, and therefore corresponds to much shorter wavelengths. These wavelengths lie in the infrared, just to the long wavelength side of visible light (wavelengths between 10 and 100 mm). When the carbon skeleton of a molecule vibrates, all the bonds stretch and relax in combination and by and large these absorptions are unhelpful. However, some bonds stretch essentially independently of the rest of the molecule, and we can use these to identify functional groups. This occurs if the bond is either:

- much stronger or weaker than others nearby, or
- between atoms that are much heavier or lighter than their neighbours

Indeed, the relationship between the frequency of the bond vibration, the mass of the atoms, and the strength of the bond is essentially the same as Hooke's law for a simple harmonic oscillator. Hooke's law shows that the frequency of the vibration ν is proportional to the square root of a force constant *f*—more or less the bond strength—and inversely proportional to the square root of a reduced mass μ, that is, the product of the masses of the two atoms forming the bond divided by their sum:

$$\mu = \frac{m_1 m_2}{m_1 + m_2}$$

The precise maths is less important to us as chemists than the simple result.

Hooke's law describes the movement of two masses attached to a spring. You may have met it if you have studied physics. You need not be concerned here with its derivation, just the result. It takes the following form:

$$\nu = \frac{1}{2\pi c}\sqrt{\frac{f}{\mu}}$$

where ν is the frequency, f is the force constant and μ is the reduced mass. c is a constant needed to make the units work.

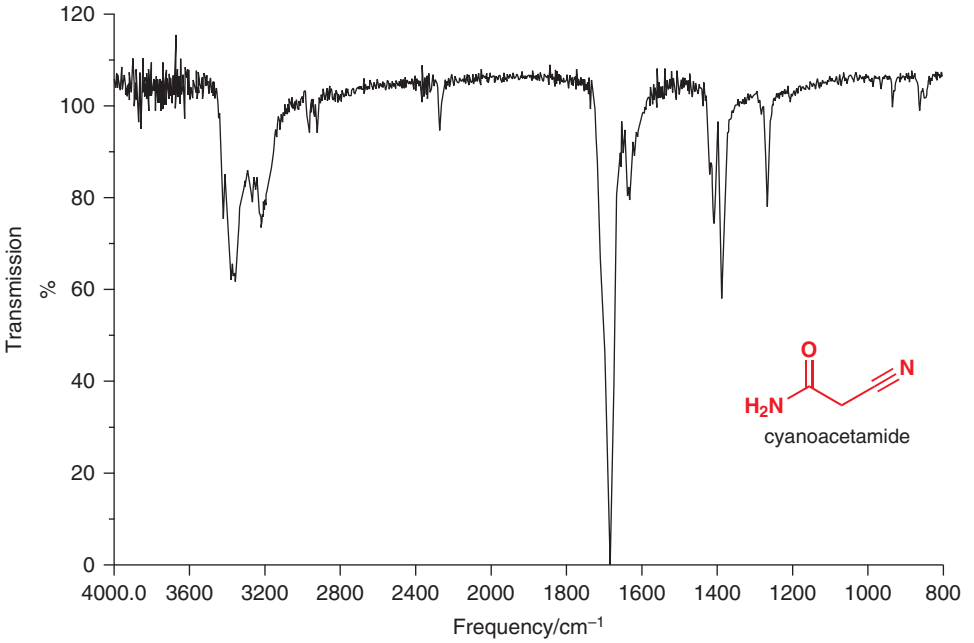
Stronger bonds vibrate faster and so do lighter atoms.

Infrared spectra are simple absorption spectra. The sample is dissolved in a solvent (or sometimes deposited on the surface of an inert NaCl plate) and exposed to infrared radiation. The wavelength scanned across the spectrum and the amount of infrared energy able to pass through the sample are plotted against the wavelength of the radiation. Just to make the numbers work out nicely, IR spectra don't usually indicate the wavelength but instead a value known as the 'wavenumber', in cm^{-1}, which is simply the number of wavelengths in one centimetre. For a typical bond this will fall between 4000 (short wavelengths, i.e. high frequency) and 500 (long wavelengths, i.e. low frequency). Strong bonds, and light atoms, vibrate fast, so you expect to see these bonds at the high wavenumber end of the spectrum, always plotted at the left-hand end.

To illustrate what we mean, here are some typical values for the IR frequencies of a selection of bonds grouped in two ways. Firstly, a series of bonds to increasingly heavy atoms (D, deuterium, has twice the mass of H, and Cl has about twice the mass of O) and secondly a series of bonds of increasing strength.

Values chiefly affected by mass of atoms (lighter atom, higher frequency)			
C—H	C—D	C—O	C—Cl
3000 cm^{-1}	2200 cm^{-1}	1100 cm^{-1}	700 cm^{-1}
Values chiefly affected by bond strength (stronger bond, higher frequency)			
C≡O	C=O	C—O	
2143 cm^{-1}	1715 cm^{-1}	1100 cm^{-1}	

Here's what a typical IR spectrum actually looks like: notice that the wavenumber scale runs from high to low but also that absorption maxima are shown upside down (IR spectra plot 'transmission')—you might say that IR spectra are upside down and back to front. If you look carefully you will also see that the scale changes in the middle to give more space to the more detailed right-hand half of the spectrum.



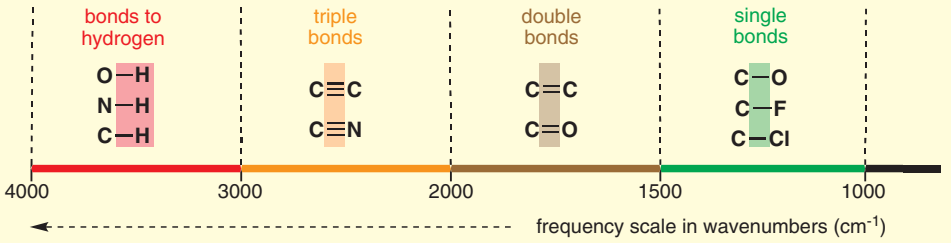
This is the spectrum of cyanoacetamide, the compound shown on the right. The overall shape of the spectrum is characteristic of this compound, but as chemists we need to be able to *interpret* the spectrum, and we can do this by dividing it up into regions, just as we did with the NMR spectra.

There are four important regions of the infrared spectrum

The first region, from **4000 to 2500 cm^{-1} is the region for C–H, N–H, and O–H** bond stretching. Most of the atoms in an organic molecule (C, N, O, for example) are about the same weight (12, 14, 16…). Hydrogen is an order of magnitude lighter than any of these and so it dominates the stretching frequency by the large effect it has on the reduced mass, so any bond to H comes right at the left-hand end of the spectrum.

Even the strongest bonds between non-H atoms—triple bonds such as C≡C or C≡N—absorb at slightly lower frequencies than bonds to hydrogen: these are in the next region, the **triple bond region from about 2500 to 2000 cm^{-1}**. This and the other two regions of the spectrum follow in logical order of bond strength as the reduced masses are all about the same: C=C and C=O **double bonds appear about 2000–1500 cm^{-1}** and at the right-hand end of the spectrum come **single bonds, below 1500 cm^{-1}**. These regions are summarized in this chart, which you should memorize.

Regions of the IR spectrum



Reduced mass and atomic mass

We introduced the idea of reduced mass on p. 64. To illustrate the effect of H on reduced mass, consider this: the reduced mass of a C–C bond is $(12 \times 12)/(12 + 12)$, i.e. $144/24 = 6.0$. If we change one of these atoms for H, the reduced mass changes to $(12 \times 1)/(12 + 1)$, i.e. $12/13 = 0.92$, but if we change it instead for F, the reduced mass changes to $(12 \times 19)/(12 + 19)$, i.e. $228/31 = 7.35$. There is a small change when we increase the mass to 19 (F), but an enormous change when we decrease it to 1 (H).

■ Absorptions in the IR are frequently referred to as 'peaks'—on the spectrum of course they are 'troughs'!

Looking back at the spectrum of cyanoacetamide on p. 65, we see peaks in the X–H region at about 3300 and 2950 cm^{-1}, which are the N–H and C–H stretches of the NH_2 and CH_2 groups. The one rather weak peak in the triple bond region (2270 cm^{-1}) is the C≡N group and the strong peak at about 1670 cm^{-1} belongs to the C=O group. We shall explain soon why some IR peaks are stronger than others. The rest of the spectrum is in the single bond region. This region is not normally interpreted in detail but is characteristic of the compound as a whole rather in the way that a fingerprint is characteristic of an individual human being—similarly, it cannot be 'interpreted'. It is indeed called the fingerprint region. The useful information from this spectrum is the presence of the C≡N and C=O groups and the exact position of the C=O absorption.

The X–H region (4000–3000 cm^{-1}) distinguishes C–H, N–H, and O–H bonds

The reduced masses of the C–H, N–H, and O–H combinations are all about the same. Any difference between the positions of the IR bands of these bonds must then be due to bond strength. In practice, C–H stretches occur at around 3000 cm^{-1} (although they are of little use in identifying compounds, it's a rare organic compound that has *no* C–H bonds), N–H stretches occur at about 3300 cm^{-1}, and O–H stretches higher still at around 3500 cm^{-1}. We can immediately deduce that the O–H bond is stronger than N–H, which is stronger than C–H. IR is a good way to measure such bond strengths.

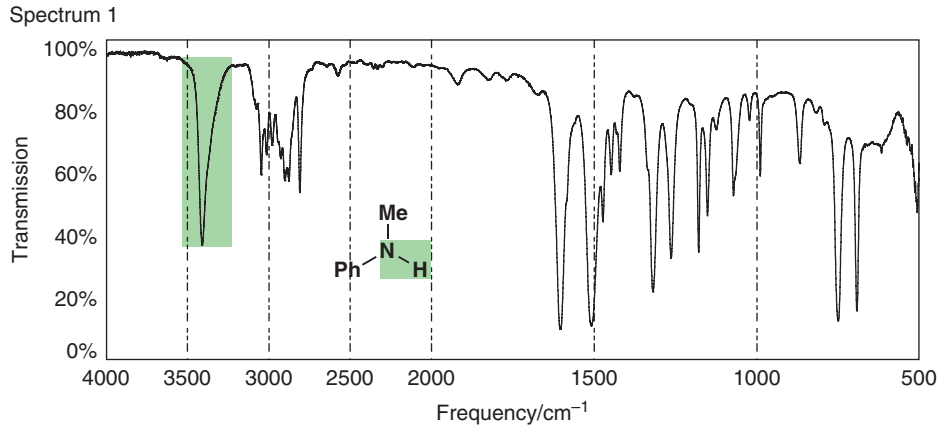
■ This may surprise you: you may be used to thinking of O–H as more reactive than CH. This is, of course, true but, as you will see in Chapter 5, factors other than bond strength control reactivity. Bond strengths will be much more important when we discuss radical reactions in Chapters 35 and 39.

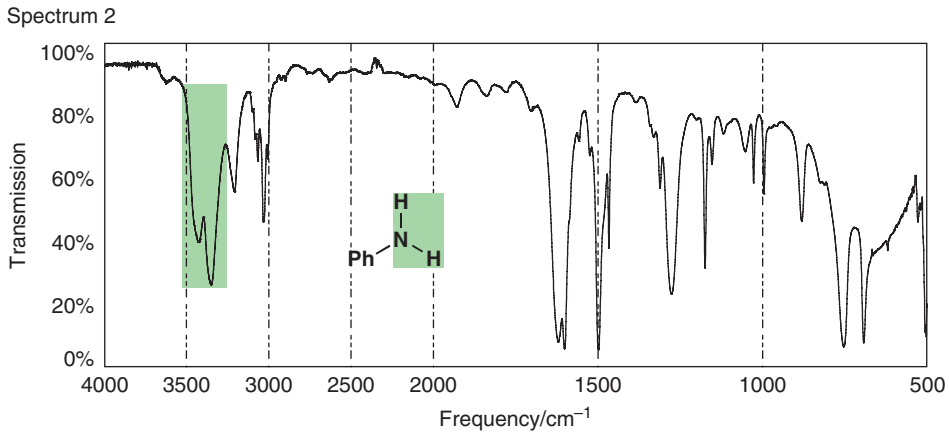
IR bands for bonds to hydrogen

Bond	Reduced mass, μ	IR frequency, cm^{-1}	Typical bond strength, kJ mol^{-1}
C–H	12/13 = 0.92	2900–3200	CH_4: 440
N–H	14/15 = 0.93	3300–3400	NH_3: 450
O–H	16/17 = 0.94	3500–3600[a]	H_2O: 500

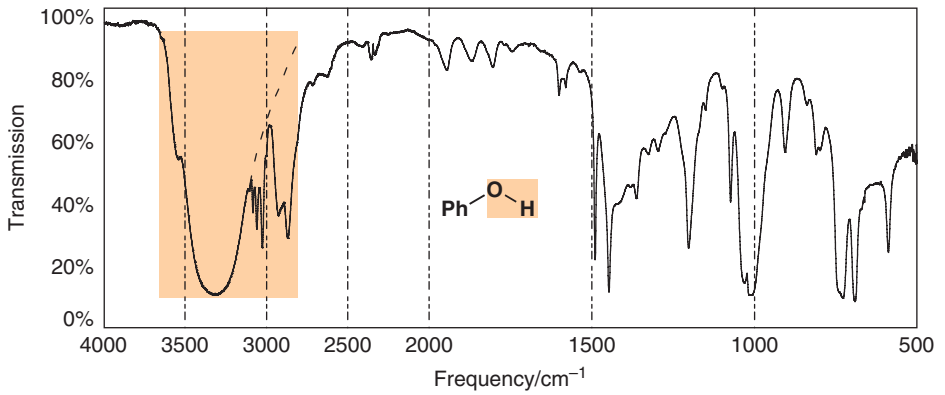
[a]When not hydrogen-bonded: see below.

The *form* of the absorption bands resulting from X–H IR stretches are very different in these four compounds. Have a look at the shaded portions of the following spectra:

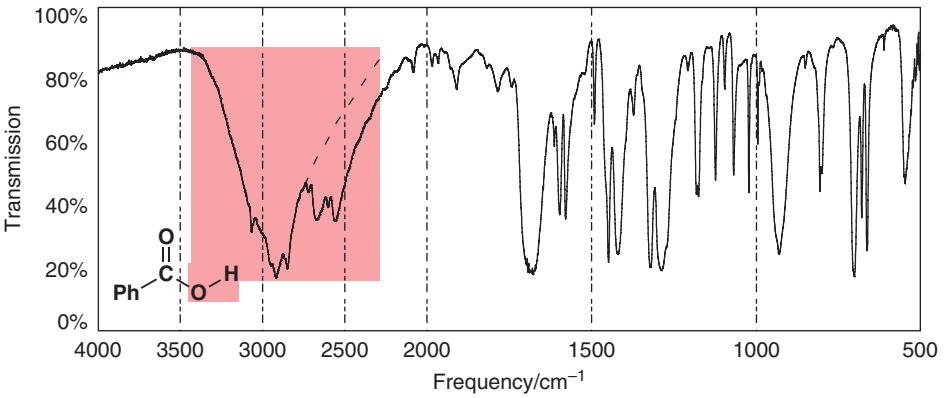


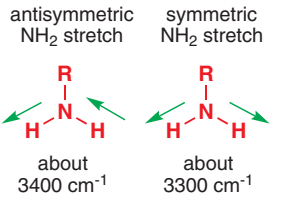


Spectrum 3



Spectrum 4





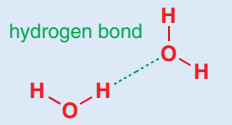
Interactive vibrations of methylamine

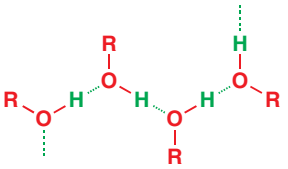
The IR peak of an NH group looks different (spectrum 1) from that of an NH_2 group (spectrum 2). A bond gives an independent vibration only if both bond strength and reduced mass are different from those of neighbouring bonds. In the case of an isolated N–H group, this is likely to be true and we usually get a sharp peak at about 3300 cm^{-1}, whether the NH group is part of a simple amine (R_2NH) or an amide (RCONHR). The NH_2 group is also independent of the rest of the molecule, but the two NH bonds inside the NH_2 group have identical force constants and reduced masses, and so vibrate as a single unit. Two equally strong bands appear: one for the two N–H bonds vibrating in phase (symmetric) and one for the two N–H bonds vibrating in opposition (antisymmetric). The antisymmetric vibration requires more energy and is at slightly higher frequency.

The O–H bands occur at higher frequency, sometimes as a sharp absorption at about 3600 cm^{-1}. More often, as in spectra 3 and 4, you will see a broad absorption at anywhere from 3500 to 2900 cm^{-1}. This is because OH groups form strong hydrogen bonds that vary in length and strength. A sharp absorption at 3600 cm^{-1} indicates a non-hydrogen-bonded OH group; the lower the absorption frequency the stronger the H bond.

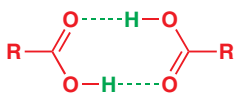
Alcohols form hydrogen bonds between the hydroxyl oxygen of one molecule and the hydroxyl hydrogen of another. These bonds are variable in length (although they are usually rather longer than normal covalent O–H bonds) and they slightly weaken the true covalent O–H bonds by varying amounts. When a bond varies in length and strength it will have a range of stretching frequencies distributed about a mean value. Alcohols, including the phenol shown in spectrum 3, typically give a rounded absorption at about 3300 cm^{-1} (contrast the sharp shape of the N–H stretch in the same region you see in the spectra above). Carboxylic acids (RCO_2H) form hydrogen-bonded dimers with two strong H bonds between the carbonyl oxygen atom of one molecule and the acidic hydrogen of the other. These also vary considerably in length and strength, and usually give the very broad V-shaped absorbance you see in the benzoic acid spectrum 4.

Hydrogen bonds are weak bonds formed from electron-rich atoms such as O or N to hydrogen atoms also attached by 'normal' bonds to the same sorts of atoms. In this diagram of a hydrogen bond between two molecules of water, the solid line represents the 'normal' bond and the green dotted line the longer hydrogen bond. The hydrogen atom is about a third of the way along the distance between the two oxygen atoms.



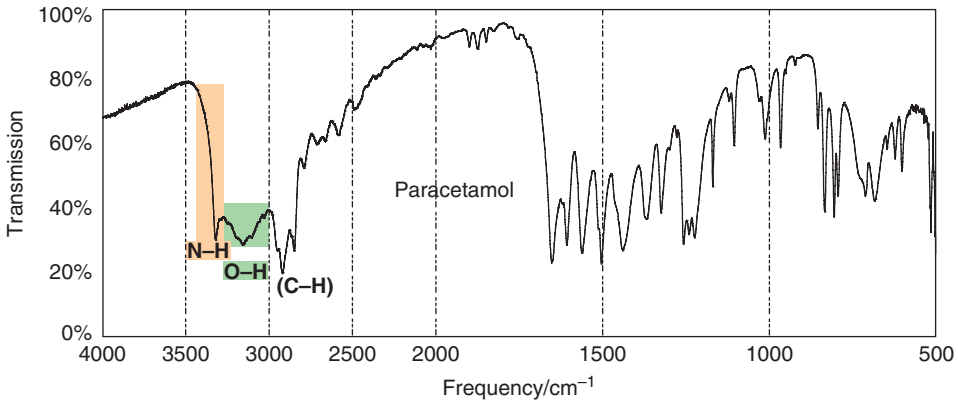


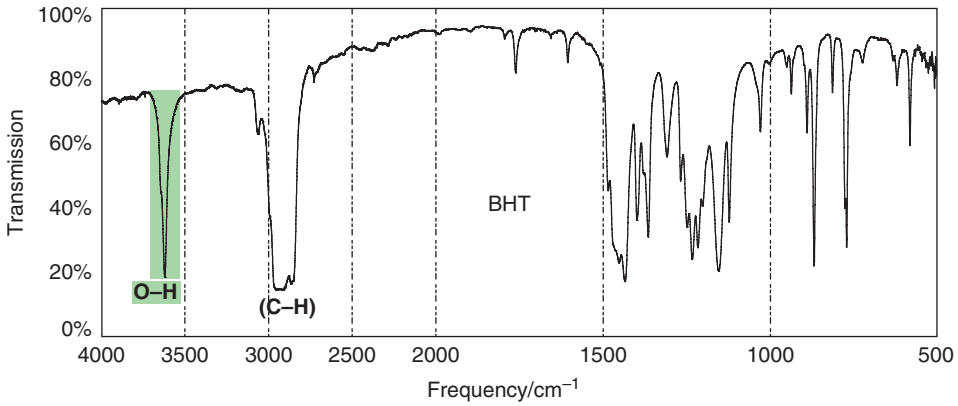
hydrogen bonding in an alcohol

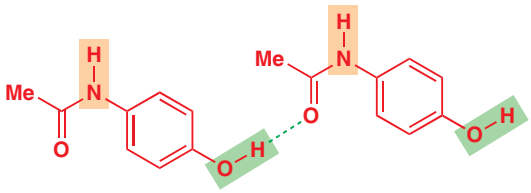


the hydrogen-bonded dimer of a carboxylic acid

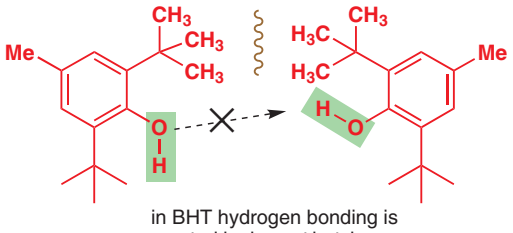
The spectra of paracetamol and BHT (which you met on pp. 58–59) illustrate the effect of hydrogen bonding on peak shape. Paracetamol has a typical sharp peak at 3330 cm^{-1} for the N–H stretch and then a rounded absorption for the hydrogen-bonded O–H stretch from 3300 down to 3000 cm^{-1} in the gap between the N–H and C–H stretches. By contrast, BHT has a sharp absorption at 3600 cm^{-1} as the two large *t*-butyl groups prevent the typical hydrogen bond from forming.







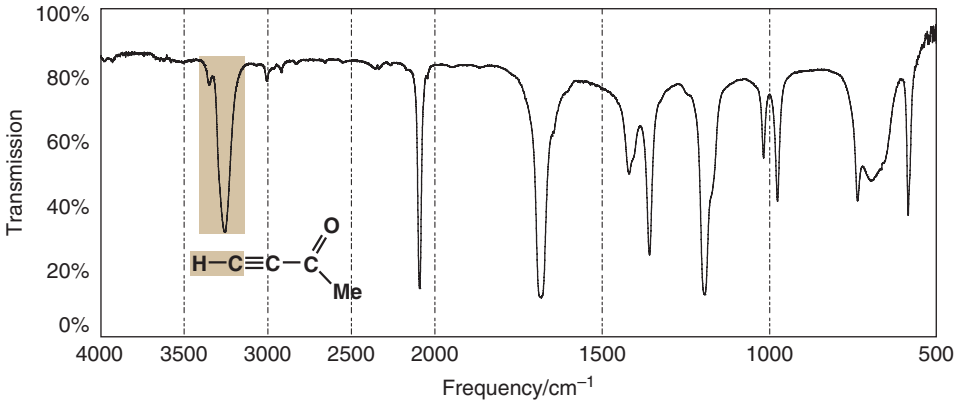
the hydrogen-bonded OH group in paracetamol



in BHT hydrogen bonding is prevented by large *t*-butyl groups

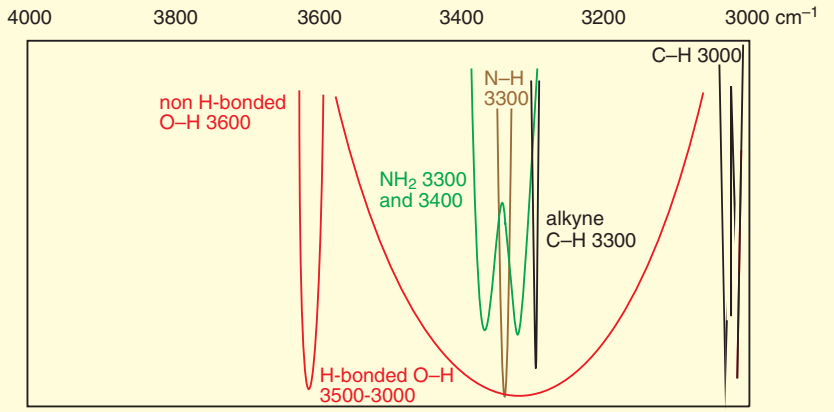
You may be confused the first time you see the IR spectrum of a terminal alkyne, R–C≡C–H, because you will see a strongish sharp peak at around 3300 cm^{-1} that looks just like an N–H stretch—the spectrum below (of methyl propynoate, also known as methyl propiolate) illustrates this. The displacement of this peak from the usual C–H stretch at about 3000 cm^{-1}

cannot be due to a change in the reduced mass and must be due to a marked increase in bond strength. The alkyne C–H bond is shorter and stronger than alkane C–H bonds.



■ In Chapter 4 you will see that carbon uses an sp^3 orbital to make a C–H bond in a saturated structure but has to use an sp orbital for a terminal alkyne C–H. This orbital has one-half s character instead of one-quarter s character. The electrons in an s orbital are held closer to the carbon's nucleus than in a p orbital, so the sp orbital makes for a shorter, stronger C–H bond.

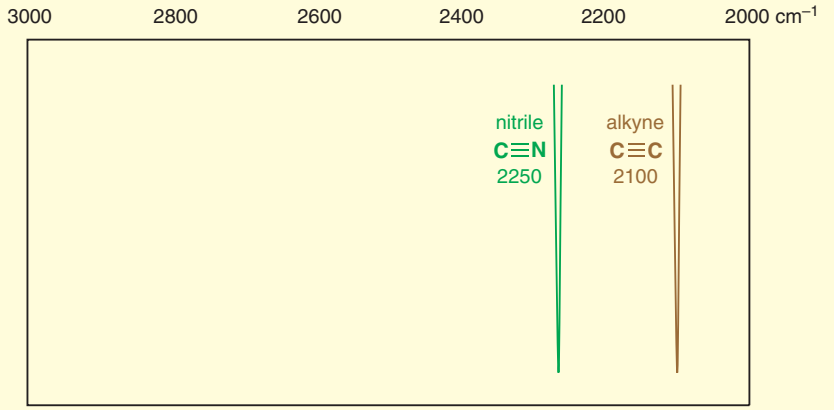
● **Typical peak shapes and frequencies for X–H bonds in the region 4000–3000 cm⁻¹.**



The triple bond region (3000–2000 cm⁻¹)

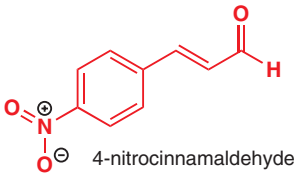
This region is often empty, meaning that when you do see a peak between 2000 and 2500 you can be absolutely certain that the compound is an alkyne (usually at around 2100) or a nitrile (at 2250 cm⁻¹). There are examples above and on p. 65.

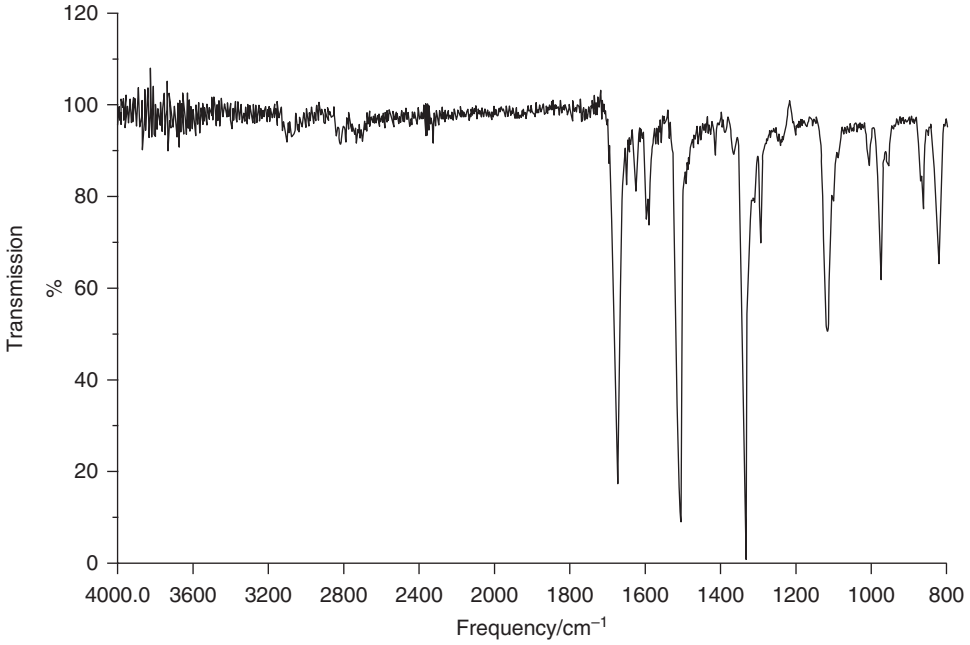
● **The only two peaks in the triple bond region**



The double bond region is the most important in IR spectra

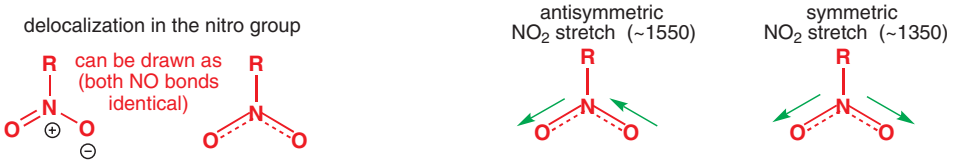
The most important absorptions in the double bond region are those of the carbonyl (C=O), alkene or arene (C=C), and nitro (NO_2), groups. All give rise to sharp bands, C=O gives one strong (i.e. intense) band anywhere between 1900 and 1500 cm^{-1}; alkene C=C gives one weak band at about 1640 cm^{-1}, and NO_2 gives two strong (intense) bands in the mid-1500s and mid-1300s cm^{-1}. Arenes usually give two or three bands in the region 1600–1500 cm^{-1}. We can illustrate several of these features in the spectrum shown below, which is that of 4-nitrocinnamaldehyde, shown in the margin.





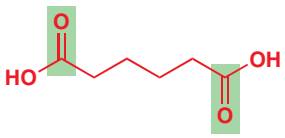
Why the nitro group gives two bands is easily understood. Just as with OH and NH_2, it is a matter of how many identical bonds are present in the same functional group. Carbonyl and alkene clearly have one double bond each. The nitro group at first sight appears to contain two different groups, N^+–O^- and N=O, but delocalization means they are identical and we see absorption for symmetric and antisymmetric stretching vibrations. As with NH_2, more energy is associated with the antisymmetric vibration and it occurs at higher frequency (>1500 cm^{-1}).

➡ Delocalization is covered in Chapter 7; for the moment, just accept that both NO bonds are the same.

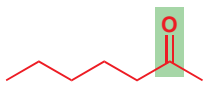


Arenes, being rings, have a much more complex pattern of vibration that cannot be analysed simply, However, it's worth noting that arene C=C bonds come at lower frequency (<1600 cm^{-1}) than alkene C=C bonds (>1600 cm^{-1}). Why? Well the individual C–C bonds in benzene are of course not full C=C double bonds—all six bonds are the same, and have the averaged character of one-and-a-half bonds each. Not surprisingly, the absorptions of these bonds fall right on the boundary between the single and double bond regions.

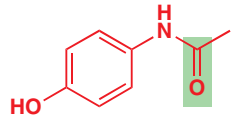
You've already seen the IR spectra of the three carbonyl compounds below in this chapter. It's easy to identify the C=O peak in each spectrum—C=O peaks are always intense (you will see why in a minute) and come somewhere near 1700 cm^{-1}.



hexanedioic acid
1720 cm^{-1}



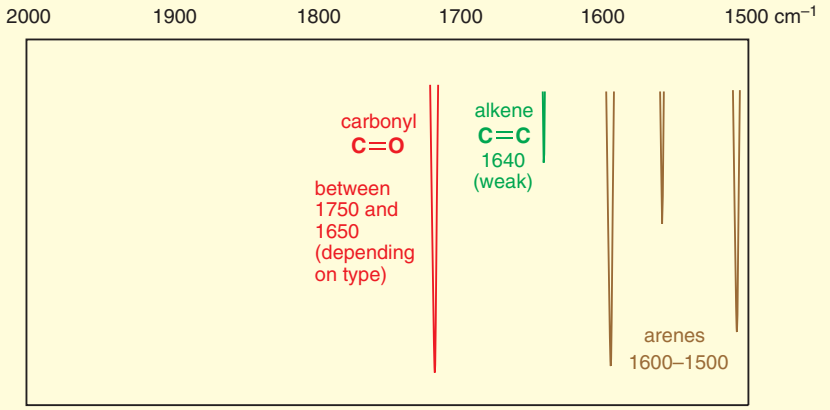
heptan-2-one
1710 cm^{-1}



paracetamol
1667 cm^{-1}

Why the positions of the peaks vary, and what we can make of this information, will be discussed in Chapter 18.

● **Important absorptions in the double bond region**



The strength of an IR absorption depends on dipole moment

If you look back at the X–H regions (3000–4000 cm^{-1}) of the four spectra on pp. 66–67. you'll notice something that at first sight seems odd. The N–H and O–H absorptions are stronger than the C–H absorptions at 3000 cm^{-1}, despite there being more C–H bonds in these molecules than O–H or N–H bonds. The reason for this is that the strength of an IR absorption varies with the change of *dipole moment* (see the box below for a definition) when the bond is stretched. If the bond is perfectly symmetrical, there is no change in dipole moment and there is no IR absorption. Obviously, the C=C bond is less polar than either C=O or N=O and its absorption is less intense in the IR. Indeed it may be absent altogether in a symmetrical alkene. By contrast the carbonyl group is very polarized, with oxygen attracting the electrons away from carbon, and stretching it causes a large change in dipole moment. C=O stretches are usually the strongest peaks in the IR spectrum. O–H and N–H stretches are stronger than C–H stretches because C–H bonds are only weakly polarized.

■ Contrast the term 'strength' applied to *absorption* and to *bonds*. A stronger absorption is a *more intense* absorption. A strong *bond* on the other hand has a *higher frequency* of absorption (other things being equal).

Dipole moments

Dipole moment depends on the variation in distribution of electrons along the bond and also its length, which is why stretching a bond can change its dipole moment. For bonds between unlike atoms, the larger the difference in electronegativity, the greater the dipole moment and the more it changes when stretched. For identical atoms (C=C, for example) the dipole moment, and its capacity to change with stretching, is much smaller. Stretching frequencies for symmetrical molecules can be measured using an alternative method known as Raman spectroscopy. This is an IR-based technique using scattered light that relies on the polarizability of bonds. Raman spectra are outside the scope of this book.

This is a good point to remind you of the various deductions we have made so far about IR spectra.

● **Absorptions in IR spectra**

Position of band depends on:	reduced mass of atoms bond strength	light atoms give high frequency strong bonds give high frequency
Strength (intensity) of band depends on:	change in dipole moment	large dipole moment gives strong absorption
Width of band depends on:	hydrogen bonding	strong H bond gives broad peak

The single bond region is used as a molecular fingerprint

The region below 1500 cm^{-1} is where the single bond vibrations occur. Here our hope that individual bonds may vibrate independently of the rest of the molecule is usually doomed to disappointment. The atoms C, N, and O all have about the same atomic weight and C–C, C–N, and C–O single bonds all have about the same strength.

Single bonds

Pair of atoms	Reduced mass	Bond strength
C–C	6.0	350 kJ mol^{-1}
C–N	6.5	305 kJ mol^{-1}
C–O	6.9	360 kJ mol^{-1}

■ A matching fingerprint is used to link a suspect to a crime, but you can't *interpret* a fingerprint to deduce the height, weight, or eye-colour of a criminal. Likewise with the fingerprint region: a matching fingerprint confirms that two compounds are identical, but without a 'suspect' you have to rely on the rest of the spectrum, above 1500 cm^{-1}, for analysis.

In addition, C–C bonds are often joined to other C–C bonds with virtually identical strength and reduced mass, and they have essentially no dipole moments. The only one of these single bonds of any value is C–O, which is polar enough to show up as a strong absorption at about 1100 cm^{-1}. Some other single bonds, such as C–Cl (weak and with a large reduced mass, so appearing at low frequency), are quite useful at about 700 cm^{-1}. Otherwise the single bond region is usually crowded with hundreds of absorptions from vibrations of all kinds used as a 'fingerprint' characteristic of the molecule but not really open to interpretation.

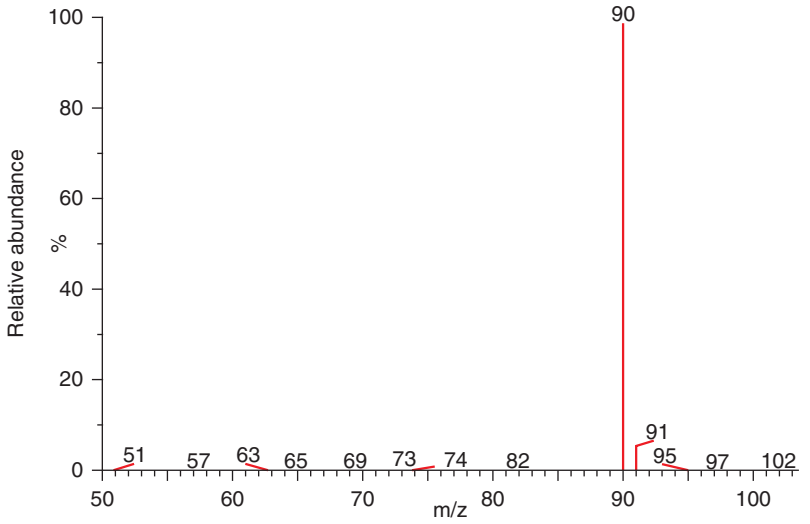
Among those hundreds of peaks in the fingerprint region there are some of a quite different kind. Stretching is not the only bond movement that leads to IR absorption. Bending of bonds, particularly C–H and N–H bonds, also leads to quite strong peaks. These are called *deformations*. Bending a bond is easier than stretching it (which is easier, stretching or bending an iron bar?). Consequently, bending absorptions need less energy and come at lower frequencies than stretching absorptions for the same bonds. These bands may not often be useful in identifying molecules, but you will notice them as they are often strong (they are usually stronger than C=C stretches, for example) and may wonder what they are.

Deformation frequencies

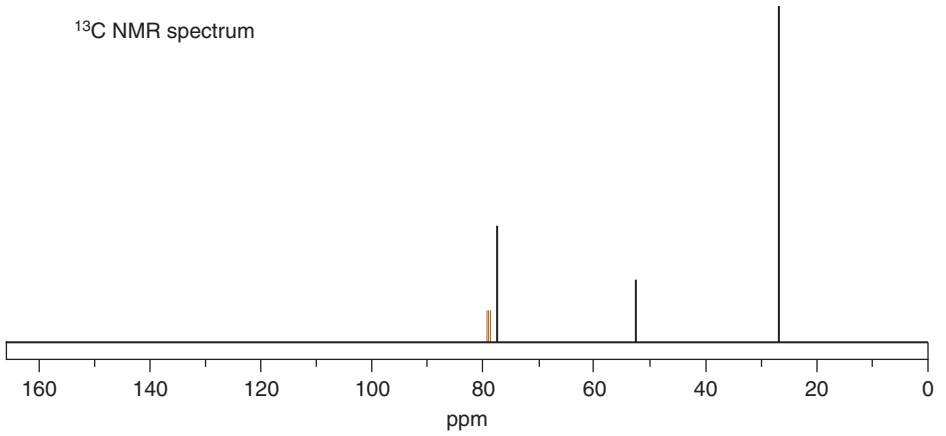
Group	Frequency, cm^{-1}
CH_2	1440–1470
CH_3	~1380
NH_2	1550–1650

Mass spectra, NMR, and IR combined make quick identification possible

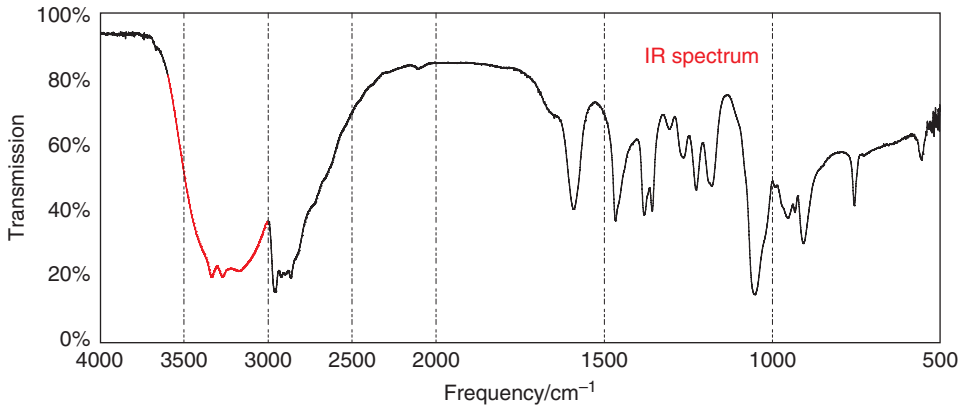
If these methods are each as powerful as we have seen on their own, how much more effective they must be together! We shall finish this chapter with the identification of some simple unknown compounds using all three methods. The first is an industrial emulsifier used to blend solids and liquids into smooth pastes. Its elecrospray mass spectrum shows it has M + H with a mass of 90, so an odd molecular weight (89) suggests one nitrogen atom. High-resolution mass spectrometry reveals that the formula is $C_4H_{11}NO$.



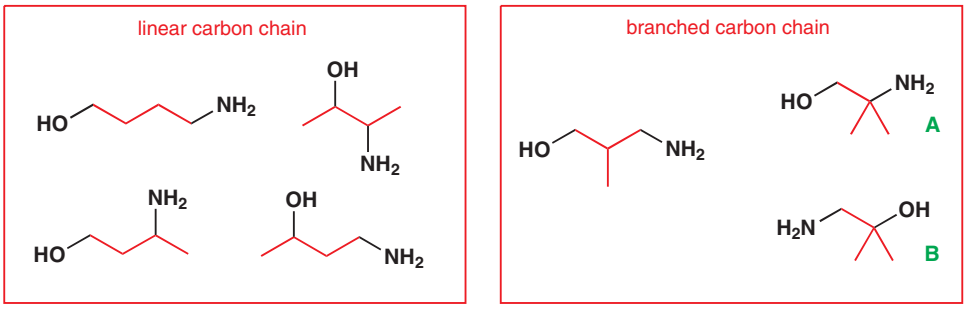
The ^{13}C NMR spectrum has only three peaks so two of the carbon atoms must be the same. There is one signal for saturated carbon next to oxygen, and two for other saturated carbons, one more downfield than the other.



The IR spectrum reveals a broad peak for an OH group with two sharp NH_2 peaks just protruding. If we put this together, we know we have C–OH and C–NH_2. Neither of these carbons can be duplicated (as there is only one O and only one N) so it must be the other two C atoms that are the same.



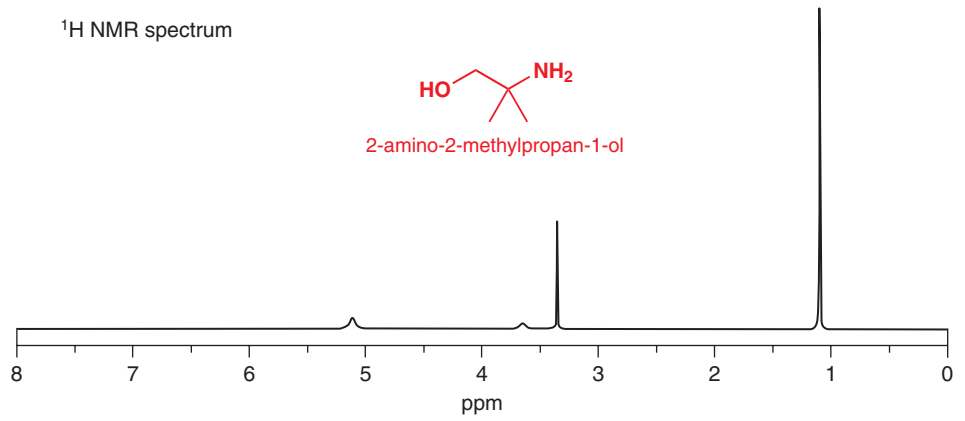
The next stage is one often overlooked. We don't seem to have much information, but try and put the two fragments together, knowing the molecular formula, and there's very little choice. The carbon chain (shown in red) could either be linear or branched and that's it!



There is no room for double bonds or rings because we need to fit in the 11 hydrogen atoms. We cannot put N or O in the chain because we know from the IR that we have the groups OH and NH_2, which can each be joined only to one other group. Of the seven possibilities only the last two, A and B, are possible since they alone have two identical carbon atoms (the two methyl groups in each case); all the other structures would have four separate signals in the NMR.

2-amino-2-methylpropan-1-ol

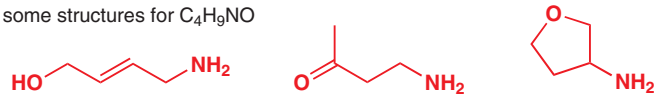
So, how can we choose between these? The solution is in the ^{1}H NMR spectrum, which is shown below. There are only two peaks visible: one at 3.3 and one at 1.1 ppm. It's quite common in ^{1}H NMR spectra not to see signals for protons attached to O or N (you will see why in Chapter 13) so we can again rule out all structures with more than two different types of H attached to C. Again, we are left with A and B, confirming our earlier deductions. But the chemical shift of the signal at δ 3.3 tells us more: it has to be due to H atoms next to an oxygen atom because it is deshielded. The industrial emulsifier must therefore be A: 2-amino-2-methylpropan-1-ol.



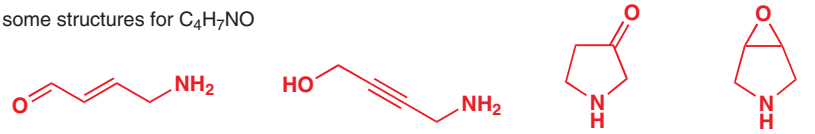
Double bond equivalents help in the search for a structure

The last example was fully saturated but it is usually a help in deducing the structure of an unknown compound if, once you know the atomic composition, you immediately work out how much unsaturation there is. It may seem obvious to you that, as $C_4H_{11}NO$ has no double bonds, then C_4H_9NO (losing two hydrogen atoms) must have one double bond, C_4H_7NO two double bonds, and so on. Well, it's not quite as simple as that. Some possible structures for these formulae are shown below.

some structures for C_4H_9NO



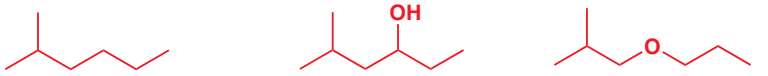
some structures for C_4H_7NO



Some of these structures have the right number of double bonds (C=C and C=O), one has a triple bond, and three compounds use rings as an alternative way of 'losing' some hydrogen atoms. Each time you make a ring or a double bond, you have to lose two more hydrogen atoms. So double bonds (of all kinds) and rings are called **double bond equivalents** (DBEs).

You can work out how many DBEs there are in a given atomic composition just by making a drawing of one possible structure for the formula (all possible structures for the same formula have the same number of DBEs). Alternatively, you can calculate the DBEs if you wish. A saturated hydrocarbon with n carbon atoms has $(2n + 2)$ hydrogens. Oxygen doesn't make any difference to this: there are the same number of Hs in a saturated ether or alcohol as in a saturated hydrocarbon.

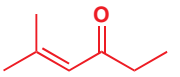
saturated hydrocarbon C_7H_{16} saturated alcohol $C_7H_{16}O$ saturated ether $C_7H_{16}O$



All have $(2n + 2)$ H atoms

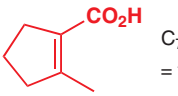
So, for a compound containing C, H, and O only, take the actual number of hydrogen atoms away from $(2n + 2)$ and divide by two. Just to check that it works, for the unsaturated ketone $C_7H_{12}O$ the calculation becomes:

1. Maximum number of H atoms for 7Cs: $2n + 2 = 16$
2. Subtract the actual number of H atoms (12): $16 - 12 = 4$
3. Divide by 2 to give the DBEs: $4/2 = 2$



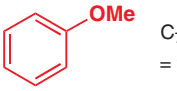
$C_7H_{12}O$ = two DBE

Here are two more examples to illustrate the method. This unsaturated cyclic acid has: $16 - 10 = 6$ divided by $2 = 3$ DBEs and it has one alkene, one C=O, and one ring. Correct.



$C_7H_{10}O_2$ = three DBE

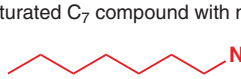
The aromatic ether has $16 - 8 = 8$ divided by 2 gives 4 DBEs and it has three double bonds in the ring and the ring itself. Correct again. A benzene ring always gives four DBEs: three for the double bonds and one for the ring.



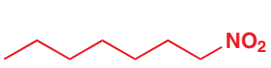
C_7H_8O = four DBE

Nitrogen makes a difference. Every nitrogen adds *one extra hydrogen* atom because nitrogen can make three bonds. This means that the formula becomes: subtract actual number of hydrogens from $(2n + 2)$, *add one for each nitrogen atom*, and divide by two. We can try this out too. Here are some example structures of compounds with seven C atoms, one N and an assortment of unsaturation and rings.

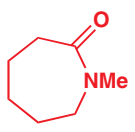
saturated C_7 compound with nitrogen



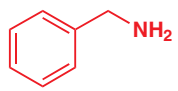
$C_7H_{17}N$ = (2n + 3) H atoms



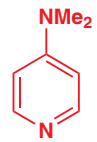
$C_7H_{15}NO_2$ = one DBE



$C_7H_{13}NO$ = two DBE



C_7H_9N = four DBE



$C_7H_{10}N_2$ = four DBE

The saturated compound has $(2n + 3)$ Hs instead of $(2n + 2)$. The saturated nitro compound has $(2n + 2) = 16$, less 15 (the actual number of Hs) plus one (the number of nitrogen atoms) = 2.

Divide this by 2 and you get 1 DBE, which is the N=O bond. We leave the third and fourth examples for you to work out, but the last compound (we shall meet this later as DMAP) has:

1. Maximum number of H atoms for 7Cs: $2n + 2 = 16$
2. Subtract the actual number of H atoms (10): $16 - 10 = 6$
3. Add number of nitrogens: $6 + 2 = 8$
4. Divide by 2 to give the DBEs: $8/2 = 4$

There are indeed three double bonds and a ring, making four in all. Make sure that you can do these calculations without much trouble.

If you have other elements too it is simpler just to draw a trial structure and find out how many DBEs there are. You may prefer this method for all compounds as it has the advantage of giving you one possible structure before you really start. One good tip is that if you have few hydrogens relative to the number of carbon atoms (at least four DBEs) then there is probably an aromatic ring in the compound.

Knowing the number of double bond equivalents for a formula derived by high-resolution mass spectrometry is a quick short cut to generating some plausible structures. You can then rule them in or rule them out by comparing with IR and NMR data.

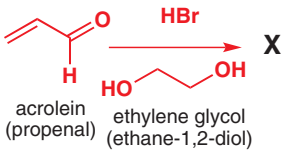
■ Do not confuse this calculation with the observation we made about mass spectra that the molecular weight of a compound containing one nitrogen atom must be odd. This observation and the number of DBEs are, of course, related but they are different calculations made for different purposes.

● Working out the DBEs for an unknown compound

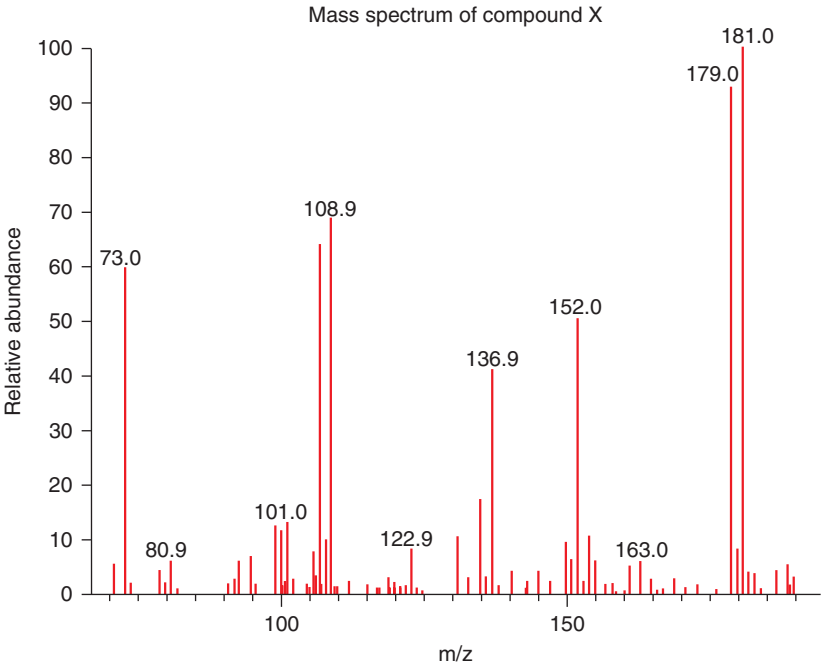
1 Calculate the expected number of Hs in the saturated structure
- **(a) For C_n there would be $2n + 2$H atoms if C, H, O only.**
- **(b) For C_nN_m there would be $2n + 2 + m$H atoms.**

2 Subtract the actual number of Hs and divide by 2. This gives the DBEs.
3 If there are other atoms (Cl, B, P, etc.) it is best to draw a trial structure.
4 A DBE indicates either a ring or a double bond (a triple bond is two DBEs).
5 A benzene ring has four DBEs (three for the double bonds and one for the ring).
6 If there are few Hs, e.g. less than the number of Cs, suspect a benzene ring.
7 A nitro group has one DBE only.

An unknown compound from a chemical reaction



Our last example addresses a situation very common in chemistry—working out the structure of a product of a reaction. The situation is this: you have treated propenal (acrolein) with HBr in ethane-1,2-diol (or glycol) as solvent for 1 hour at room temperature. Distillation of the reaction mixture gives a colourless liquid, compound X. What is it?

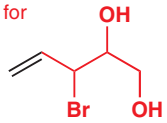


The mass spectrum shows a molecular ion (181) much heavier than that of the starting material, $C_3H_4O = 56$. Indeed it shows two molecular ions at 181 and 179, typical of a bromo compound, so it looks as if HBr has added to the aldehyde somehow. High resolution mass spectrometry reveals a formula of $C_5H_9BrO_2$, and the five carbon atoms make it look as though the glycol has added in too. If we add everything together we find that the unknown compound is the result of the three reagents added together less one molecule of water.

■ It's often very helpful, when you have an unknown product, to subtract the molecular mass of the starting material from its molecular mass to find out what has been added (or taken away).

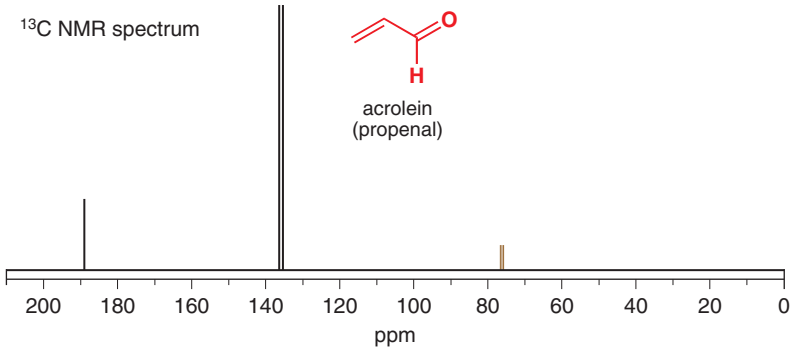
Now, how many DBEs have we got? With a formula like this the safest bet is to draw something that has the right formula—*it need not be what you expect the product to be*. Here's something in the margin—we just added atoms till we got there, and to do so we had to put in *one* double bond. $C_5H_9BrO_2$ has one DBE.

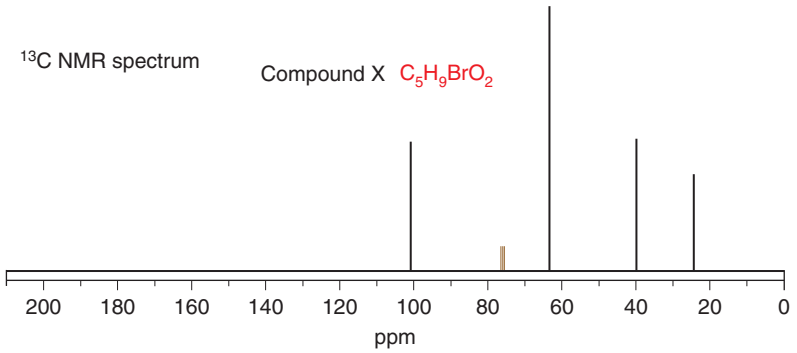
trial structure for $C_5H_9BrO_2$: one DBE



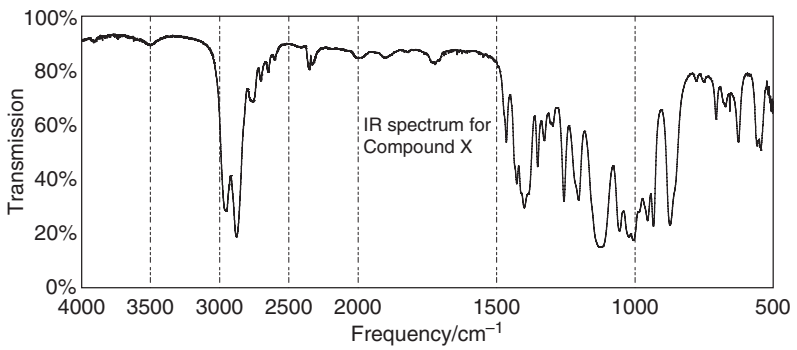
Just a trial structure, not a suggestion for the real product!

The next thing is to see what remains of the hydrocarbon skeleton of propenal by NMR. The ^{13}C NMR spectrum of $CH_2{=}CH{-}CHO$ clearly shows one carbonyl group and two carbons on a double bond. These have all disappeared in the product and for the five carbon atoms we are left with four signals, two saturated, one next to oxygen, and one at 102.6 ppm, just creeping into the double bond region.

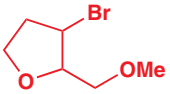




The IR spectrum gives us another puzzle—there appear to be no functional groups at all! No OH, no carbonyl, no alkene—what else can we have? The answer is an ether, or rather two ethers as there are two oxygen atoms. Now that we suspect an ether, we can look for the C–O single bond stretch in the IR spectrum and find it at 1128 cm^{-1}.

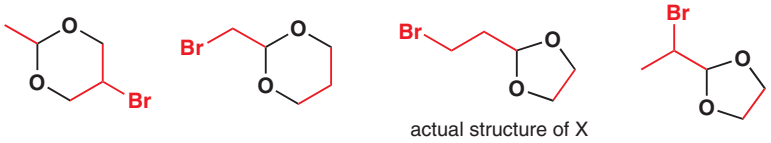


Each ether oxygen must have a carbon atom on each side of it, but we seem to have only one carbon in the saturated C next to O region (50–100 ppm) in the ^{13}C NMR. Of course, as you've already seen, these limits are arbitrary, and in fact the peak at 102 ppm is also a saturated C next to O. It is unlikely to be an alkene anyway as it takes two carbons to make an alkene. What would deshield a saturated C as much as this? The answer is two oxygen atoms. We can explain the ^{13}C spectrum if we assume a symmetrical fragment C–O–C–O–C accounts for three of the five carbon atoms.



So, where is our double bond equivalent? We know we haven't got a double bond (no alkene and no C=O) so the DBE must be a ring. You might feel uncomfortable with rings, but you must get used to them. Five-, six-, and seven-membered rings are very common. In fact, most known organic compounds have rings in them. We could draw many cyclic structures for the formula we have here, such as this one in the margin.

But this won't do as it would have five different carbon atoms. It is much more likely that the basic skeletons of the organic reagents are preserved, that is, that we have a two-carbon (from the ethylene glycol) and a three-carbon (from the propenal) fragment joined through oxygen atoms. This gives four possibilities, all containing the C–O–C–O–C fragment we deduced earlier (highlighted in black).



These are all quite reasonable, although we might prefer the third as it is easier to see how it derives from the reagents. The product is in fact this third possibility, and to be sure we would have to turn to the fine details of ^{1}H NMR spectroscopy, which we return to in Chapter 13.

Looking forward to Chapters 13 and 18

We have only begun to explore the intricate world of identification of structure by spectroscopy. It is important that you recognize that structures are assigned, not because of some theoretical reason or because a reaction 'ought' to give a certain product, but because of sound evidence from spectra. You have seen four powerful methods—mass spectra, ^{13}C and ^{1}H NMR, and IR spectroscopy—in this chapter. In Chapter 13 we delve more deeply into the most important of all (^{1}H NMR) and, finally, in Chapter 18 we shall take each of these methods a little further and show how the structures of more complex unknown compounds are really deduced. The last problem we have discussed here is not really solvable without ^{1}H NMR and in reality no-one would tackle any structure problem without this most powerful of all techniques. From now on spectroscopic evidence will appear in virtually every chapter. Even if we do not say so explicitly every time a new compound appears, the structure of this compound will in fact have been determined spectroscopically. Chemists make new compounds, and every time they do they characterize the compound with a full set of spectra. No scientific journal will accept that a new compound has been made unless a full description of all of these spectra are submitted with the report. Spectroscopy lets the science of organic chemistry advance.

Further reading

You will find it an advantage to have one of the short books on spectroscopic analysis to hand as they give explanations, comprehensive tables of data, and problems. We recommend *Spectroscopic Methods in Organic Chemistry*, 6th edn, by D. H. Williams and Ian Fleming, McGraw-Hill, London, 2007, and the Oxford Primer *Introduction to Organic Spectroscopy* by L. M. Harwood and T. D. W. Claridge, OUP, Oxford, 1996.

Check your understanding

To check that you have mastered the concepts presented in this chapter, attempt the problems that are available in the book's Online Resource Centre at **http://www.oxfordtextbooks.co.uk/orc/clayden2e/**