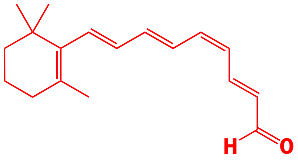
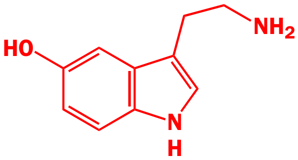
What is organic chemistry?

1

Organic chemistry and you



11-*cis*-retinal
absorbs light when we see



serotonin
human neurotransmitter

▶ We are going to give you structures of organic compounds in this chapter—otherwise it would be rather dull. If you do not understand the diagrams, do not worry. Explanation is on its way.

You are already a highly skilled organic chemist. As you read these words, your eyes are using an organic compound (retinal) to convert visible light into nerve impulses. When you picked up this book, your muscles were doing chemical reactions on sugars to give you the energy you needed. As you understand, gaps between your brain cells are being bridged by simple organic molecules (neurotransmitter amines) so that nerve impulses can be passed around your brain. And you did all that without consciously thinking about it. You do not yet understand these processes in your mind as well as you can carry them out in your brain and body. You are not alone there. No organic chemist, however brilliant, understands the detailed chemical working of the human mind or body very well.

We, the authors, include ourselves in this generalization, but we are going to show you in this book what enormous strides have been taken in the understanding of organic chemistry since the science came into being in the early years of the nineteenth century. Organic chemistry began as a tentative attempt to understand the chemistry of life. It has grown into the confident basis of vast multinational industries that feed, clothe, and cure millions of people without their even being aware of the role of chemistry in their lives. Chemists cooperate with physicists and mathematicians to understand how molecules behave and with biologists to understand how molecules determine life processes. The development of these ideas is already a revelation at the beginning of the twenty-first century, but is far from complete. We aim not to give you the measurements of the skeleton of a dead science but to equip you to understand the conflicting demands of an adolescent one.

Like all sciences, chemistry has a unique place in our pattern of understanding of the universe. It is the science of molecules. But organic chemistry is something more. It literally creates itself as it grows. Of course we need to study the molecules of nature both because they are interesting in their own right and because their functions are important to our lives. Organic chemistry often studies life by making new molecules that give information not available from the molecules actually present in living things.

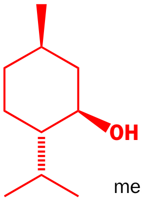
This creation of new molecules has given us new materials such as plastics, new dyes to colour our clothes, new perfumes to wear, new drugs to cure diseases. Some people think that these activities are unnatural and their products dangerous or unwholesome. But these new molecules are built by humans from other molecules found on earth using the skills inherent in our natural brains. Birds build nests; man makes houses. Which is unnatural? To the organic chemist this is a meaningless distinction. There are toxic compounds and nutritious ones, stable compounds and reactive ones—but there is only one type of chemistry: it goes on both inside our brains and bodies and also in our flasks and reactors, born from the ideas in our minds and the skill in our hands. We are not going to set ourselves up as moral judges in any way. We believe it is right to try and understand the world about us as best we can and to use that understanding creatively. This is what we want to share with you.

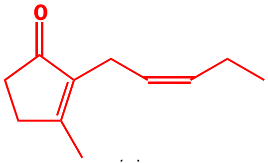
Organic compounds

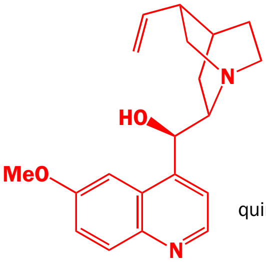
Organic chemistry started as the chemistry of life, when that was thought to be different from the chemistry in the laboratory. Then it became the chemistry of carbon compounds, especially those found in coal. Now it is both. It is the chemistry of the compounds of carbon along with other elements such as are found in living things and elsewhere.

■ You will be able to read towards the end of the book (Chapters 49–51) about the extraordinary chemistry that allows life to exist but this is known only from a modern cooperation between chemists and biologists.

The organic compounds available to us today are those present in living things and those formed over millions of years from dead things. In earlier times, the organic compounds known from nature were those in the 'essential oils' that could be distilled from plants and the alkaloids that could be extracted from crushed plants with acid. Menthol is a famous example of a flavouring compound from the essential oil of spearmint and *cis*-jasmone an example of a perfume distilled from jasmine flowers.

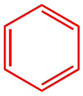

menthol

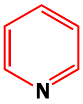

cis-jasmone

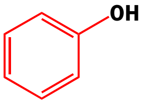

quinine

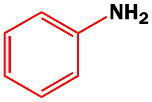
Even in the sixteenth century one alkaloid was famous—quinine was extracted from the bark of the South American cinchona tree and used to treat fevers, especially malaria. The Jesuits who did this work (the remedy was known as 'Jesuit's bark') did not of course know what the structure of quinine was, but now we do.

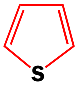
The main reservoir of chemicals available to the nineteenth century chemists was coal. Distillation of coal to give gas for lighting and heating (mainly hydrogen and carbon monoxide) also gave a brown tar rich in aromatic compounds such as benzene, pyridine, phenol, aniline, and thiophene.


benzene

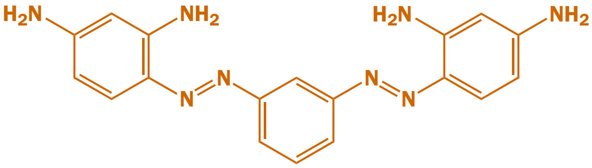

pyridine


phenol


aniline


thiophene

Phenol was used by Lister as an antiseptic in surgery and aniline became the basis for the dyestuffs industry. It was this that really started the search for new organic compounds made by chemists rather than by nature. A dyestuff of this kind—still available—is Bismarck Brown, which should tell you that much of this early work was done in Germany.


Bismarck Brown Y

■ You can read about polymers and plastics in Chapter 52 and about fine chemicals throughout the book.

$CH_3—(CH_2)_n—CH_3$

n = an enormous number
length of molecule is *n* + 2 carbon atoms

$CH_3—(CH_2)_n—CH_2—CH_3$

n = an enormous number
length of molecule is *n* + 3 carbon atoms

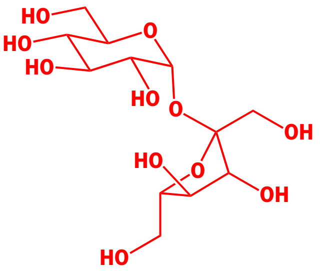
In the twentieth century oil overtook coal as the main source of bulk organic compounds so that simple hydrocarbons like methane (CH_4, 'natural gas') and propane ($CH_3CH_2CH_3$, 'calor gas') became available for fuel. At the same time chemists began the search for new molecules from new sources such as fungi, corals, and bacteria and two organic chemical industries developed in parallel—'bulk' and 'fine' chemicals. Bulk chemicals like paints and plastics are usually based on simple molecules produced in multitonne quantities while fine chemicals such as drugs, perfumes, and flavouring materials are produced in smaller quantities but much more profitably.

At the time of writing there were about 16 million organic compounds known. How many more are possible? There is no limit (except the number of atoms in the universe). Imagine you've just made the longest hydrocarbon ever made—you just have to add another carbon atom and you've made another. This process can go on with any type of compound *ad infinitum*.

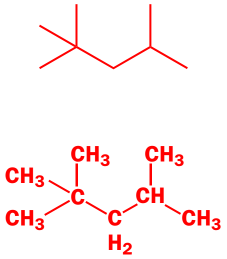
But these millions of compounds are not just a long list of linear hydrocarbons; they embrace all kinds of molecules with amazingly varied properties. In this chapter we offer a selection.

What do they *look* like? They may be crystalline solids, oils, waxes, plastics, elastics, mobile or volatile liquids, or gases. Familiar ones include white crystalline sugar, a cheap natural compound isolated from plants as hard white crystals when pure, and petrol, a mixture of colourless, volatile, flammable hydrocarbons. Isooctane is a typical example and gives its name to the octane rating of petrol.

The compounds need not lack colour. Indeed we can soon dream up a rainbow of organic compounds covering the whole spectrum, not to mention black and brown. In this table we have avoided dyestuffs and have chosen compounds as varied in structure as possible.

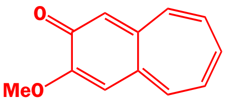
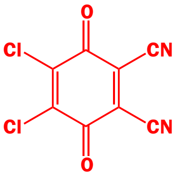
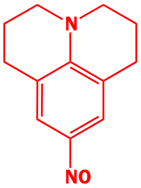
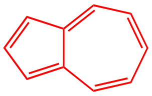
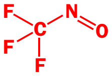


sucrose – ordinary sugar
isolated from sugar cane
or sugar beet
white crystalline solid



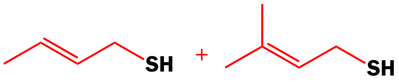
isooctane (2,3,5-trimethylpentane)
a major constiuent of petrol
volatile inflammable liquid

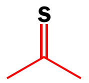
spectrum

Colour	Description	Compound	Structure
red	dark red hexagonal plates	3′-methoxybenzocycloheptatriene-2′-one	
orange	amber needles	dichloro dicyano quinone (DDQ)	
yellow	toxic yellow explosive gas	diazomethane	$CH_2{=}\overset{\oplus}{N}{=}\overset{\ominus}{N}$
green	green prisms with a steel-blue lustre	9-nitroso julolidine	
blue	deep blue liquid with a peppery smell	azulene	
purple	deep blue gas condensing to a purple solid	nitroso trifluoromethane	

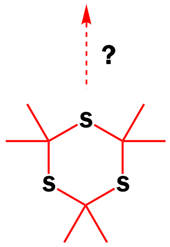
Colour is not the only characteristic by which we recognize compounds. All too often it is their odour that lets us know they are around. There are some quite foul organic compounds too; the smell of the skunk is a mixture of two thiols—sulfur compounds containing SH groups.

skunk spray contains:





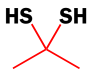
thioacetone



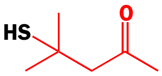
trithioacetone;
Freiburg was evacuated
because of a smell from
the distillation this compound

But perhaps the worst aroma was that which caused the evacuation of the city of Freiburg in 1889. Attempts to make thioacetone by the cracking of trithioacetone gave rise to 'an offensive smell which spread rapidly over a great area of the town causing fainting, vomiting and a panic evacuation…the laboratory work was abandoned'.

It was perhaps foolhardy for workers at an Esso research station to repeat the experiment of cracking trithioacetone south of Oxford in 1967. Let them take up the story. 'Recently…we found ourselves with an odour problem beyond our worst expectations. During early experiments, a stopper jumped from a bottle of residues, and, although replaced at once, resulted in an immediate complaint of nausea and sickness from colleagues working in a building two hundred yards away. Two of our chemists who had done no more than investigate the cracking of minute amounts of trithioacetone…found themselves the object of hostile stares in a restaurant and suffered the humiliation of having a waitress spray the area around them with a deodorant…. The odours defied the expected effects of dilution since workers in the laboratory did not find the odours intolerable . . . and genuinely denied responsibility since they were working in closed systems. To convince them otherwise, they were dispersed with other observers around the laboratory, at distances up to a quarter of a mile, and one drop of either acetone *gem*-dithiol or the mother liquors from crude trithioacetone crystallisations were placed on a watch glass in a fume cupboard. The odour was detected downwind in seconds.'



propane
dithiol



4-methyl-4-
sulfanylpentan-
2-one

two candidates for
the worst smell in the world

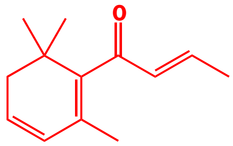
no-one wants to find the winner!

There are two candidates for this dreadful smell—propane dithiol (called acetone *gem*-dithiol above) or 4-methyl-4-sulfanylpentan-2-one. It is unlikely that anyone else will be brave enough to resolve the controversy.

Nasty smells have their uses. The natural gas piped to our homes contains small amounts of deliberately added sulfur compounds such as *tert*-butyl thiol $(CH_3)_3CSH$. When we say small, we mean *very small*—humans can detect one part in 50 000 000 000 parts of natural gas.

the divine smell
of the black truffle
comes from this compound

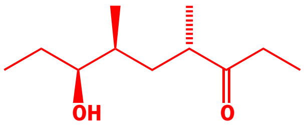
Other compounds have delightful odours. To redeem the honour of sulfur compounds we must cite the truffle which pigs can smell through a metre of soil and whose taste and smell is so delightful that truffles cost more than their weight in gold. Damascenones are responsible for the smell of roses. If you smell one drop you will be disappointed, as it smells rather like turpentine or camphor, but next morning you and the clothes you were wearing will smell powerfully of roses. Just like the compounds from trithioacetone, this smell develops on dilution.



damascenone - the smell of roses

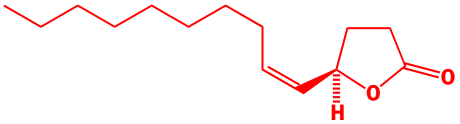
Humans are not the only creatures with a sense of smell. We can find mates using our eyes alone (though smell does play a part) but insects cannot do this. They are small in a crowded world and they find others of their own species and the opposite sex by smell. Most insects produce volatile compounds that can be picked up by a potential mate in incredibly weak concentrations. Only 1.5 mg of serricornin, the sex pheromone of the cigarette beetle, could be isolated from 65 000 female beetles—so there isn't much in each beetle. Nevertheless, the slightest whiff of it causes the males to gather and attempt frenzied copulation.

The sex pheromone of the Japanese beetle, also given off by the females, has been made by chemists. As little as 5 μg (micrograms, note!) was more effective than four virgin females in attracting the males.



serricornin

the sex pheromone of the cigarette beetle
Lasioderma serricorne

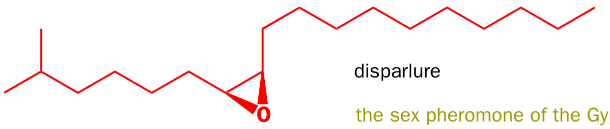


japonilure

the sex pheromone of the Japanese beetle
Popilia japonica

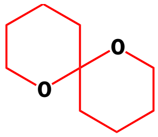
The pheromone of the gypsy moth, disparlure, was identified from a few μg isolated from the moths and only 10 μg of synthetic material. As little as 2×10^{-12} g is active as a lure for the males in field tests. The three pheromones we have mentioned are available commercially for the specific trapping of these destructive insect pests.

Don't suppose that the females always do all the work; both male and female olive flies produce pheromones that attract the other sex. The remarkable thing is that one mirror image of the molecule attracts the males while the other attracts the females!



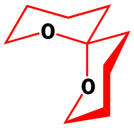
disparlure

the sex pheromone of the Gypsy moth *Portheria dispar*

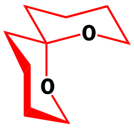


olean

sex pheromone of the olive fly *Bacrocera oleae*



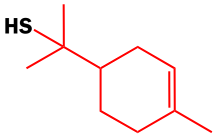
this mirror image isomer attracts the males



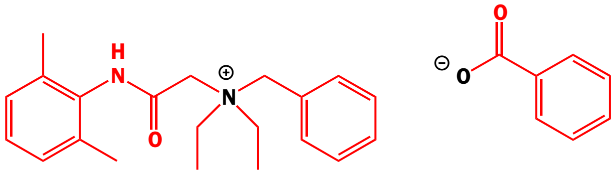
this mirror image isomer attracts the females

What about taste? Take the grapefruit. The main flavour comes from another sulfur compound and human beings can detect 2×10^{-5} parts per billion of this compound. This is an almost unimaginably small amount equal to 10^{-4} mg per tonne or a drop, not in a bucket, but in a good-sized lake. Why evolution should have left us abnormally sensitive to grapefruit, we leave you to imagine.

For a nasty taste, we should mention 'bittering agents', put into dangerous household substances like toilet cleaner to stop children eating them by accident. Notice that this complex organic compound is actually a salt—it has positively charged nitrogen and negatively charged oxygen atoms—and this makes it soluble in water.



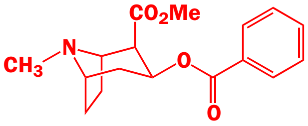
flavouring principle of grapefruit



bitrex
denatonium benzoate
benzyldiethyl[(2,6-xylylcarbamoyl)methyl]ammonium benzoate

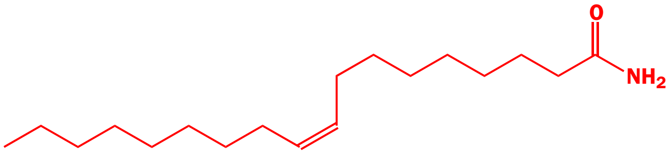
Other organic compounds have strange effects on humans. Various 'drugs' such as alcohol and cocaine are taken in various ways to make people temporarily happy. They have their dangers. Too much alcohol leads to a lot of misery and any cocaine at all may make you a slave for life.

alcohol
(ethanol)



cocaine
- an addictive alkaloid

Again, let's not forget other creatures. Cats seem to be able to go to sleep at any time and recently a compound was isolated from the cerebrospinal fluid of cats that makes them, or rats, or humans go off to sleep quickly. It is a surprisingly simple compound.

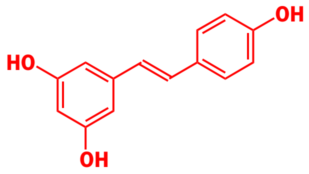


a sleep-inducing fatty acid derivative
cis-9,10-octadecenoamide

This compound and disparlure are both derivatives of fatty acids, molecules that feature in many of the food problems people are so interested in now (and rightly so). Fatty acids in the diet are a popular preoccupation and the good and bad qualities of saturates, monounsaturates, and polyunsaturates are continually in the news. This too is organic chemistry. One of the latest molecules to be recognized as an anticancer agent in our diet is CLA (conjugated linoleic acid) in dairy products.

CLA (Conjugated Linoleic Acid)
cis-9-*trans*-11 conjugated linoleic acid
dietary anticancer agent

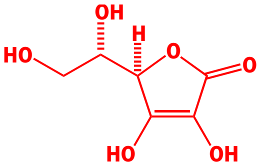
Another fashionable molecule is resveratrole, which may be responsible for the beneficial effects of red wine in preventing heart disease. It is a quite different organic compound with two benzene rings and you can read about it in Chapter 51.



resveratrole from the skins of grapes
is this the compound in red wine which helps to prevent heart disease?

For our third edible molecule we choose vitamin C. This is an essential factor in our diets—indeed, that is why it is called a vitamin. The disease scurvy, a degeneration of soft tissues, particularly in the mouth, from which sailors on long voyages like those of Columbus suffered, results if we don't have vitamin C. It also is a universal antioxidant, scavenging for rogue free radicals and so protecting us against cancer. Some people think an extra large intake protects us against the common cold, but this is not yet proved.

▶ Vitamin C (ascorbic acid) is a vitamin for primates, guinea-pigs, and fruit bats, but other mammals can make it for themselves.



vitamin C (ascorbic acid)

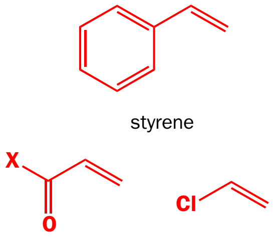
Organic chemistry and industry

Vitamin C is manufactured on a huge scale by Roche, a Swiss company. All over the world there are chemistry-based companies making organic molecules on scales varying from a few kilograms to thousands of tonnes per year. This is good news for students of organic chemistry; there are lots of jobs around and it is an international job market. The scale of some of these operations of organic chemistry is almost incredible. The petrochemicals industry processes (and we use the products!) over 10 million litres of crude oil every day. Much of this is just burnt in vehicles as petrol or diesel, but some of it is purified or converted into organic compounds for use in the rest of the chemical industry. Multinational companies with thousands of employees such as Esso (Exxon) and Shell dominate this sector.

Some simple compounds are made both from oil and from plants. The ethanol used as a starting material to make other compounds in industry is largely made by the catalytic hydration of ethylene from oil. But ethanol is also used as a fuel, particularly in Brazil where it is made by fermentation of sugar cane wastes. This fuel uses a waste product, saves on oil imports, and has improved the quality of the air in the very large Brazilian cities, Rio de Janeiro and São Paulo.

Plastics and polymers take much of the production of the petrochemical industry in the form of monomers such as styrene, acrylates, and vinyl chloride. The products of this enormous industry are everything made of plastic including solid plastics for household goods and furniture, fibres for clothes (24 million tonnes per annum), elastic polymers for car tyres, light bubble-filled polymers for packing, and so on. Companies such as BASF, Dupont, Amoco, Monsanto, Laporte, Hoechst, and ICI are leaders here. Worldwide polymer production approaches 100 million tonnes per annum and PVC manufacture alone employs over 50 000 people to make over 20 million tonnes per annum.

monomers for polymer manufacture



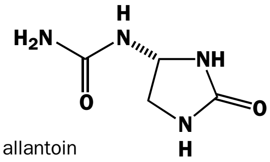
styrene
acrylates
vinyl chloride

The washing-up bowl is plastic too but the detergent you put in it belongs to another branch of the chemical industry—companies like Unilever (Britain) or Procter and Gamble (USA) which produce soap, detergent, cleaners, bleaches, polishes, and all the many essentials for the modern home. These products may be lemon and lavender scented but they too mostly come from the oil industry. Nowadays, most products of this kind tell us, after a fashion, what is in them. Try this example—a well known brand of shaving gel along with the list of contents on the container:

Ingredients
aqua, palmitic acid, triethanolamine, glycereth-26, isopentane, oleamide-DEA, oleth-2, stearic acid, isobutane, PEG-14M, parfum, allantoin, hydroxyethyl-cellulose, hydroxypropyl-cellulose, PEG-150 distearate, CI 42053, CI 47005

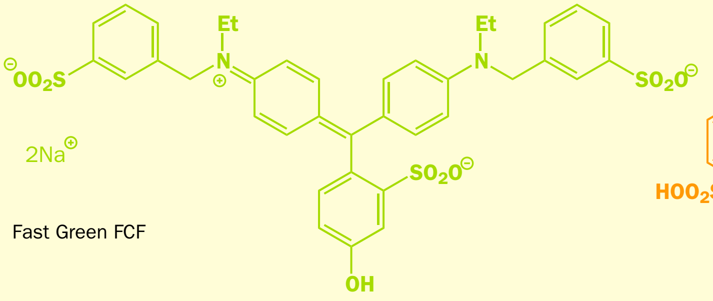
Does any of this make any sense?

It doesn't all make sense to us, but here is a possible interpretation. We certainly hope the book will set you on the path of understanding the sense (and the nonsense!) of this sort of thing.

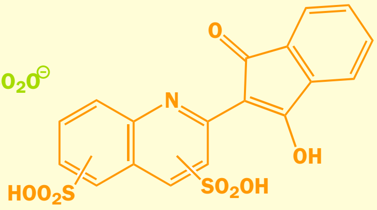
Ingredient	Chemical meaning	Purpose
aqua	water	solvent
palmitic acid	$CH_3(CH_2)_{14}CO_2H$	acid, emulsifier
triethanolamine	$N(CH_2CH_2OH)_3$	base
glycereth-26	glyceryl$(OCH_2CH_2)_{26}OH$	surfactant
isopentane	$(CH_3)_2CHCH_2CH_3$	propellant
oleamide-DEA	$CH_3(CH_2)_7CH{=}CH(CH_2)_7CONEt_2$	
oleth-2	Oleyl$(OCH_2CH_2)_2OH$	surfactant
stearic acid	$CH_3(CH_2)_{16}CO_2H$	acid, emulsifier
isobutane	$(CH_3)_2CHCH_3$	propellant
PEG-14M	polyoxyethylene glycol ester	surfactant
parfum	perfume	
allantoin	 allantoin	promotes healing in case you cut yourself while shaving
hydroxyethyl-cellulose	cellulose fibre from wood pulp with $–OCH_2CH_2OH$ groups added	gives body
hydroxypropyl-cellulose	cellulose fibre from wood pulp with $–OCH_2CH(OH)CH_3$ groups added	gives body
PEG-150 distearate	polyoxyethylene glycol diester	surfactant
CI 42053	Fast Green FCF (see box)	green dye
CI 47005	Quinoline Yellow (see box)	yellow dye

The structures of two dyes

Fast Green FCF and Quinoline Yellow are colours permitted to be used in foods and cosmetics and have the structures shown here. Quinoline Yellow is a mixture of isomeric sulfonic acids in the two rings shown.

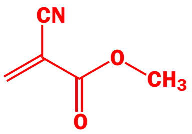


Fast Green FCF



Quinoline Yellow

The particular acids, bases, surfactants, and so on are chosen to blend together in a smooth emulsion when propelled from the can. The result should feel, smell, and look attractive and a greenish colour is considered clean and antiseptic by the customer. What the can actually says is this: 'Superior lubricants within the gel prepare the skin for an exceptionally close, comfortable and effective shave. It contains added moisturisers to help protect the skin from razor burn. Lightly fragranced.'



Superglue bonds things together
when this small molecule
joins up with hundreds of its fellows
in a polymerization reaction

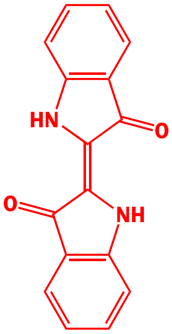
■ The formation of polymers is discussed in Chapter 52.

Another oil-derived class of organic chemical business includes adhesives, sealants, coatings, and so on, with companies like Ciba–Geigy, Dow, Monsanto, and Laporte in the lead. Nowadays aircraft are glued together with epoxy-resins and you can glue almost anything with ‘Superglue’ a polymer of methyl cyanoacrylate.

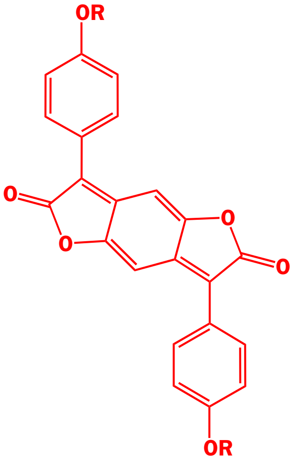
There is a big market for intense colours for dyeing cloth, colouring plastic and paper, painting walls, and so on. This is the dyestuffs and pigments industry and leaders here are companies like ICI and Akzo Nobel. ICI have a large stake in this aspect of the business, their paints turnover alone being £2 003 000 000 in 1995.

The most famous dyestuff is probably indigo, an ancient dye that used to be isolated from plants but is now made chemically. It is the colour of blue jeans. More modern dyestuffs can be represented by ICI’s benzodifuranones, which give fashionable red colours to synthetic fabrics like polyesters.

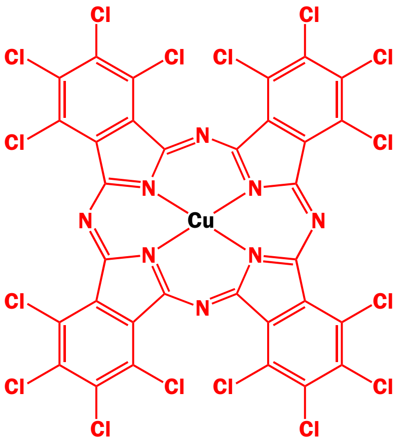
We see one type of pigment around us all the time in the form of the colours on plastic bags. Among the best compounds for these are the metal complexes called phthalocyanines. Changing the metal (Cu and Fe are popular) at the centre and the halogens round the edge of these molecules changes the colour but blues and green predominate. The metal atom is not necessary for intense pigment colours—one new class of intense ‘high performance’ pigments in the orange–red range are the DPP (1,4-diketopyrrolo[3,4-c]pyrroles) series developed by Ciba–Geigy. Pigment Red 254 is used in paints and plastics.



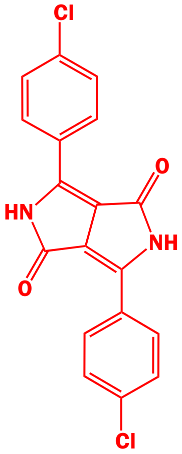
indigo
the colour of blue jeans



ICI’s Dispersol
benzodifuranone
red dyes for polyester



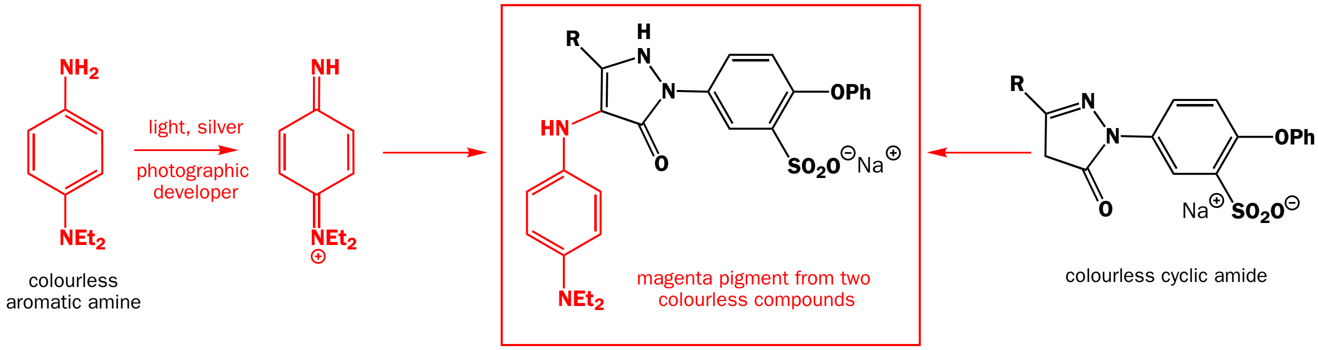
ICI’s Monastral Green GNA
a good green for plastic objects



Ciba Geigy’s Pigment Red 254
an intense DPP pigment

■ You can read in Chapter 7 why some compounds are coloured and others not.

Colour photography starts with inorganic silver halides but they are carried on organic gelatin. Light acts on silver halides to give silver atoms that form the photographic image, but only in black and white. The colour in films like Kodachrome then comes from the coupling of two colourless organic compounds. One, usually an aromatic amine, is oxidized and couples with the other to give a coloured compound.



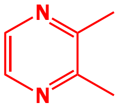
That brings us to flavours and fragrances. Companies like International Flavours and Fragrances (USA) or Givaudan–Roure (Swiss) produce very big ranges of fine chemicals for the perfume, cosmetic, and food industries. Many of these will come from oil but others come from plant sources. A typical perfume will contain 5–10% fragrances in an ethanol/water (about 90:10) mixture. So the perfumery industry needs a very large amount of ethanol and, you might think, not much perfumery material. In fact, important fragrances like jasmine are produced on a >10 000 tonnes per annum scale. The cost of a pure perfume ingredient like *cis*-jasmone, the main ingredient of jasmine, may be several hundred pounds, dollars, or euros per gram.

cis-jasmone
the main compound in jasmine perfume

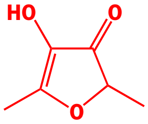
The world of perfumery

Perfume chemists use extraordinary language to describe their achievements: 'Paco Rabanne pour homme was created to reproduce the effect of a summer walk in the open air among the hills of Provence: the smell of herbs, rosemary and thyme, and sparkling freshness with cool sea breezes mingling with warm soft Alpine air. To achieve the required effect, the perfumer blended herbaceous oils with woody accords and the synthetic aroma chemical dimethylheptanol which has a penetrating but indefinable freshness associated with open air or freshly washed linen'. (J. Ayres, *Chemistry and Industry*, 1988, 579)

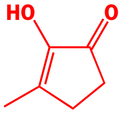
Chemists produce synthetic flavourings such as 'smoky bacon' and even 'chocolate'. Meaty flavours come from simple heterocycles such as alkyl pyrazines (present in coffee as well as roast meat) and furonol, originally found in pineapples. Compounds such as corylone and maltol give caramel and meaty flavours. Mixtures of these and other synthetic compounds can be 'tuned' to taste like many roasted foods from fresh bread to coffee and barbecued meat.



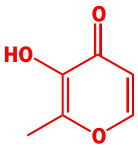
an alkyl pyrazine
from coffee and roast meat



furonol
roast meat



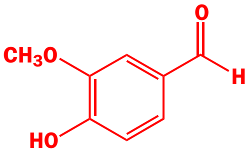
corylone
caramel
roasted taste

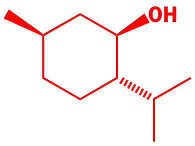


maltol
E-636 for cakes and biscuits

Some flavouring compounds are also perfumes and may also be used as an intermediate in making other compounds. Two such large-scale flavouring compounds are vanillin (vanilla flavour as in ice cream) and menthol (mint flavour) both manufactured on a large scale and with many uses.

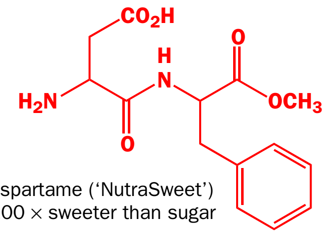
vanillin
found in vanilla pods; manufactured on a large scale





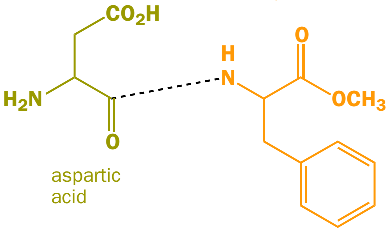
menthol
extracted from mint; 25% of the world's supply manufactured

Food chemistry includes much larger-scale items than flavours. Sweeteners such as sugar itself are isolated from plants on an enormous scale. Sugar's structure appeared a few pages back. Other sweeteners such as saccharin (discovered in 1879!) and aspartame (1965) are made on a sizeable scale. Aspartame is a compound of two of the natural amino acids present in all living things and is made by Monsanto on a large scale (over 10 000 tonnes per annum).



aspartame ('NutraSweet')
200 × sweeter than sugar

is made from two amino acids –



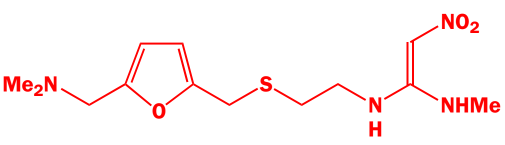
methyl ester of phenylalanine

aspartic acid

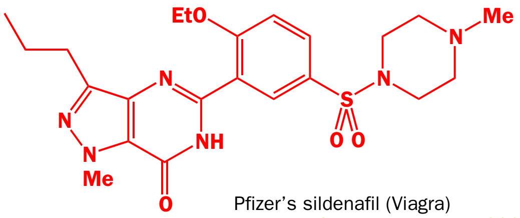
The pharmaceutical businesses produce drugs and medicinal products of many kinds. One of the great revolutions of modern life has been the expectation that humans will survive diseases because of a treatment designed to deal specifically with that disease. The most successful drug ever is ranitidine (Zantac), the Glaxo–Wellcome ulcer treatment, and one of the fastest-growing is Pfizer's sildenafil (Viagra). 'Success' refers both to human health and to profit!

You will know people (probably older men) who are 'on β-blockers'. These are compounds designed to block the effects of adrenaline (epinephrine) on the heart and hence to prevent heart disease. One of the best is Zeneca's tenormin. Preventing high blood pressure also prevents heart disease and certain specific enzyme inhibitors (called 'ACE-inhibitors') such as Squibb's captopril work in this way. These are drugs that imitate substances naturally present in the body.

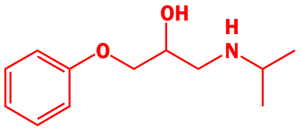
The treatment of infectious diseases relies on antibiotics such as the penicillins to prevent bacteria from multiplying. One of the most successful of these is Smith Kline Beecham's amoxycillin. The four-membered ring at the heart of the molecule is the 'β-lactam'.



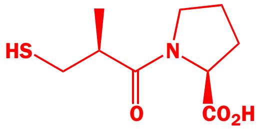
Glaxo-Wellcome's ranitidine
the most successful drug to date
world wide sales peaked >£1,000,000,000 per annum



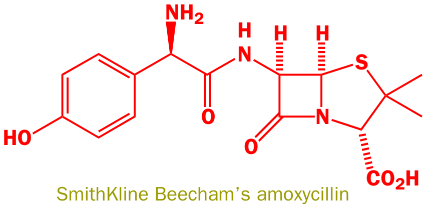
Pfizer's sildenafil (Viagra)
three million satisfied customers in 1998



Zeneca's tenormin
cardioselective β-blocker
for treatment and prevention
of heart disease



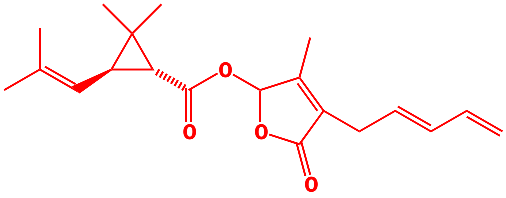
Squibb's captopril
specific enzyme inhibitor
for treatment and
prevention of hypertension



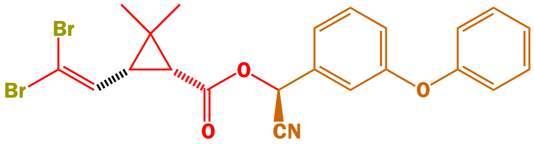
SmithKline Beecham's amoxycillin
β-lactam antibiotic
for treatment of bacterial infections

We cannot maintain our present high density of population in the developed world, nor deal with malnutrition in the developing world unless we preserve our food supply from attacks by insects and fungi and from competition by weeds. The world market for agrochemicals is over £10 000 000 000 per annum divided roughly equally between herbicides, fungicides, and insecticides.

At the moment we hold our own by the use of agrochemicals: companies such as Rhône-Poulenc, Zeneca, BASF, Schering–Plough, and Dow produce compounds of remarkable and specific activity. The most famous modern insecticides are modelled on the natural pyrethrins, stabilized against degradation by sunlight by chemical modification (see coloured portions of decamethrin) and targeted to specific insects on specific crops in cooperation with biologists. Decamethrin has a safety factor of >10#000 for mustard beetles over mammals, can be applied at only 10 grams per hectare (about one level tablespoon per football pitch), and leaves no significant environmental residue.



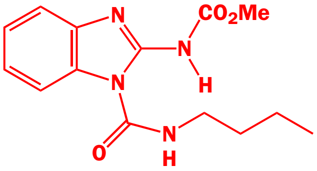
a natural pyrethin
from *pyrethrum* - daisy-like flowers from East Africa



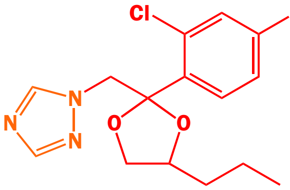
decamethrin
a modified pyrethrin - more active and stable in sunlight

As you learn more chemistry, you will appreciate how remarkable it is that Nature should produce three-membered rings and that chemists should use them in bulk compounds to be sprayed on crops in fields. Even more remarkable in some ways is the new generation of fungicides based on a five-membered ring containing three nitrogen atoms—the triazole ring. These compounds inhibit an enzyme present in fungi but not in plants or animals.

One fungus (potato blight) caused the Irish potato famine of the nineteenth century and the various blights, blotches, rots, rusts, smuts, and mildews can overwhelm any crop in a short time. Especially now that so much is grown in Western Europe in winter, fungal diseases are a real threat.



benomyl
a fungicide which controls many plant diseases



propiconazole
a *triazole* fungicide

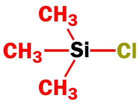
You will have noticed that some of these companies have fingers in many pies. These companies, or groups as they should be called, are the real giants of organic chemistry. Rhône–Poulenc, the French group which includes pharmaceuticals (Rhône–Poulenc–Rorer), animal health, agrochemicals, chemicals, fibres, and polymers, had sales of about 90 billion French Francs in 1996. Dow, the US group which includes chemicals, plastics, hydrocarbons, and other bulk chemicals, had sales of about 20 billion US dollars in 1996.

Organic chemistry and the periodic table

All the compounds we have shown you are built up on hydrocarbon (carbon and hydrogen) skeletons. Most have oxygen and/or nitrogen as well; some have sulfur and some phosphorus. These are the main elements of organic chemistry but another way the science has developed is an exploration of (some would say take-over bid for) the rest of the periodic table. Some of our compounds also had fluorine, sodium, copper, chlorine, and bromine. The organic chemistry of silicon, boron, lithium, the halogens (F, Cl, Br, and I), tin, copper, and palladium has been particularly well studied and these elements commonly form part of organic reagents used in the laboratory. They will crop up throughout this book. These 'lesser' elements appear in many important reagents, which are used in organic chemical laboratories all over the world. Butyllithium, trimethylsilyl chloride, tributyltin hydride, and dimethylcopper lithium are good examples.

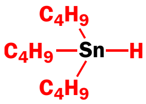
The halogens also appear in many life-saving drugs. The recently discovered antiviral compounds, such as fialuridine (which contains both F and I, as well as N and O), are essential for the fight against HIV and AIDS. They are modelled on natural compounds from nucleic acids. The naturally occurring cytotoxic (antitumour) agent halomon, extracted from red algae, contains Br and Cl.

BuLi
butyllithium



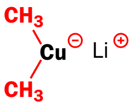
$CH_3-\overset{CH_3}{\underset{CH_3}{Si}}-Cl$

Me_3SiCl
trimethylsilyl chloride



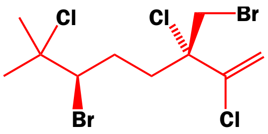
$C_4H_9-\overset{C_4H_9}{\underset{C_4H_9}{Sn}}-H$

Bu_3SnH
tributyltin hydride

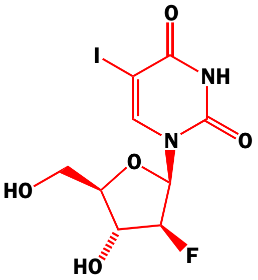


$\overset{CH_3}{\underset{CH_3}{Cu^{\ominus}}}\ Li^{\oplus}$

Me_2CuLi
dimethylcopper lithium



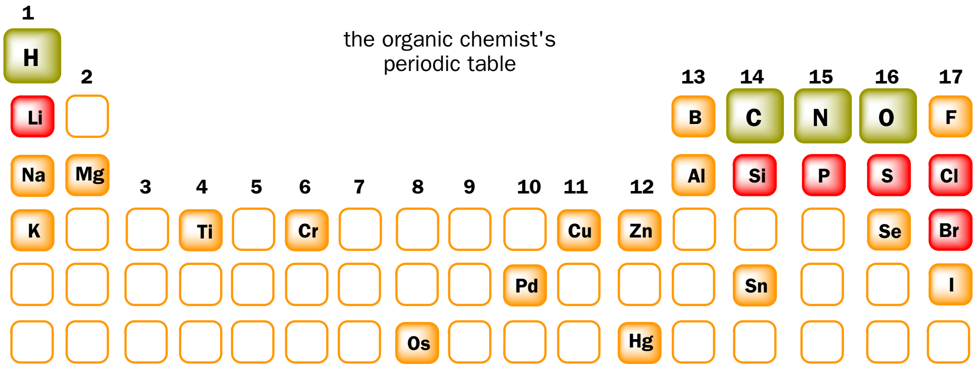
halomon
naturally occurring antitumour agent



fialuridine
antiviral compound

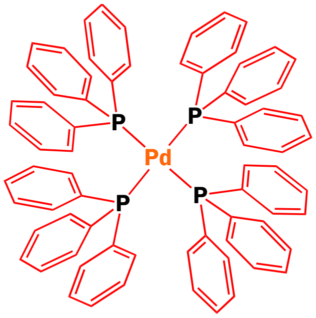
Another definition of organic chemistry would use the periodic table. The key elements in organic chemistry are of course C, H, N, and O, but also important are the halogens (F, Cl. Br, I), p-block elements such as Si, S, and P, metals such as Li, Pd, Cu, and Hg, and many more. We can construct an organic chemist's periodic table with the most important elements emphasized:

▶ You will certainly know something about the periodic table from your previous studies of inorganic chemistry. A basic knowledge of the groups, which elements are metals, and roughly where the elements in our table appear will be helpful to you.



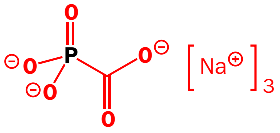
the organic chemist's periodic table

So where does inorganic chemistry end and organic chemistry begin? Would you say that the antiviral compound foscarnet was organic? It is a compound of carbon with the formula CPO_5Na_3 but is has no C–H bonds. And what about the important reagent tetrakis triphenyl phosphine palladium? It has lots of hydrocarbon—twelve benzene rings in fact—but the benzene rings are all joined to phosphorus atoms that are arranged in a square around the central palladium atom, so the molecule is held together by C–P and P–Pd bonds, not by a hydrocarbon skeleton. Although it has the very organic-looking formula $C_{72}H_{60}P_4Pd$, many people would say it is inorganic. But is it?



tetrakis
triphenylphosphine
palladium

$[(C_6H_5)_3P]_4Pd$
$(Ph_3P)_4Pd$



foscarnet – antiviral agent

The answer is that we don't know and we don't care. It is important these days to realize that strict boundaries between traditional disciplines are undesirable and meaningless. Chemistry continues across the old boundaries between organic chemistry and inorganic chemistry on the one side and organic chemistry and biochemistry on the other. Be glad that the boundaries are indistinct as that means the chemistry is all the richer. This lovely molecule $(Ph_3P)_4Pd$ belongs to *chemistry*.

Organic chemistry and this book

We have told you about organic chemistry's history, the types of compounds it concerns itself with, the things it makes, and the elements it uses. Organic chemistry today is the study of the structure and reactions of compounds in nature of compounds, in the fossil reserves such as coal and oil, and of those compounds that can be made from them. These compounds will usually be constructed with a hydrocarbon framework but will also often have atoms such as O, N, S, P, Si, B, halogens, and metals attached to them. Organic chemistry is used in the making of plastics, paints, dyestuffs, clothes, foodstuffs, human and veterinary medicines, agrochemicals, and many other things. Now we can summarize all of these in a different way.

● **The main components of organic chemistry as a discipline are these**

- **Structure determination—how to find out the structures of new compounds even if they are available only in invisibly small amounts**
- **Theoretical organic chemistry—how to understand those structures in terms of atoms and the electrons that bind them together**
- **Reaction mechanisms—how to find out how these molecules react with each other and how to predict their reactions**
- **Synthesis—how to design new molecules—and then make them**
- **Biological chemistry—how to find out what Nature does and how the structures of biologically active molecules are related to what they do**

This book is about all these things. It tells you about the structures of organic molecules and the reasons behind them. It tells you about the shapes of those molecules and how the shape relates to their function, especially in the context of biology. It tells you how those structures and shapes are discovered. It tells you about the reactions the molecules undergo and, more importantly, how and why they behave in the way they do. It tells you about nature and about industry. It tells you how molecules are made and how you too can think about making molecules.

We said 'it tells' in that last paragraph. Maybe we should have said 'we tell' because we want to speak to you through our words so that you can see how we think about organic chemistry and to encourage you to develop your own ideas. We expect you to notice that four people have written this book and that they don't all think or write in the same way. That is as it should be. Organic chemistry is too big and important a subject to be restricted by dogmatic rules. Different chemists think in different ways about many aspects of organic chemistry and in many cases it is not yet possible to be sure who is right.

We may refer to the history of chemistry from time to time but we are usually going to tell you about organic chemistry as it is now. We will develop the ideas slowly, from simple and fundamental ones using small molecules to complex ideas and large molecules. We promise one thing. We are not going to pull the wool over your eyes by making things artificially simple and avoiding the awkward questions. We aim to be honest and share both our delight in good complete explanations and our puzzlement at inadequate ones. So how are we going to do this? The book starts with a series of chapters on the structures and reactions of simple molecules. You will meet the way structures are determined and the theory that explains those structures. It is vital that you realize that theory is used to explain what is known by experiment and only then to predict what is unknown. You will meet mechanisms—the dynamic language used by chemists to talk about reactions—and of course some reactions.

The book starts with an introductory section of four chapters:

1 What is organic chemistry?
2 Organic structures
3 Determining organic structures
4 Structure of molecules

In Chapter 2 you will look at the way in which we are going to present diagrams of molecules on the printed page. Organic chemistry is a visual, three-dimensional subject and the way you draw molecules shows how you think about them. We want you too to draw molecules in the best way available now. It is just as easy to draw them well as to draw them in an old-fashioned inaccurate way.

Then in Chapter 3, before we come to the theory of molecular structure, we shall introduce you to the experimental techniques of finding out about molecular structure. This means studying the interactions between molecules and radiation by **spectroscopy**—using the whole electromagnetic spectrum from X-rays to radio waves. Only then, in Chapter 4, will we go behind the scenes and look at the theories of why atoms combine in the ways they do. Experiment comes before theory. The spectroscopic methods of Chapter 3 will still be telling the truth in a hundred years time, but the theories of Chapter 4 will look quite dated by then.

We could have titled those three chapters:

2 What shapes do organic molecules have?
3 How do we know they have those shapes?
4 Why do they have those shapes?

You need to have a grasp of the answers to these three questions before you start the study of organic reactions. That is exactly what happens next. We introduce organic reaction mechanisms in Chapter 5. Any kind of chemistry studies **reactions**—the transformations of molecules into other molecules. The dynamic process by which this happens is called **mechanism** and is the language of organic chemistry. We want you to start learning and using this language straight away so in Chapter 6 we apply it to one important class of reaction. This section is:

5 Organic reactions
6 Nucleophilic addition to the carbonyl group

Chapter 6 reveals how we are going to subdivide organic chemistry. We shall use a mechanistic classification rather than a structural classification and explain one type of reaction rather than one type of compound in each chapter. In the rest of the book most of the chapters describe types of reaction in a mechanistic way. Here is a selection.

9 Using organometallic reagents to make C–C bonds
17 Nucleophilic substitution at saturated carbon
20 Electrophilic addition to alkenes
22 Electrophilic aromatic substitution
29 Conjugate Michael addition of enolates
39 Radicals

Interspersed with these chapters are others on physical aspects, organic synthesis, stereochemistry, structural determination, and biological chemistry as all these topics are important parts of organic chemistry.

‘Connections’ section

Chemistry is not a linear subject! It is impossible simply to start at the beginning and work through to the end, introducing one new topic at a time, because chemistry is a network of interconnecting ideas. But, unfortunately, a book is, by nature, a beginning-to-end sort of thing. We have arranged the chapters in a progression of difficulty as far as is possible, but to help you find your way around we have included at the beginning of each chapter a ‘Connections’ section. This tells you three things divided among three columns:

(a) what you should be familiar with before reading the chapter—in other words, which previous chapters relate directly to the material within the chapter (‘Building on’ column)

(b) a guide to what you will find within the chapter (‘Arriving at’ column)

(c) which chapters later in the book fill out and expand the material in the chapter (‘Looking forward to’ column)

The first time you read a chapter, you should really make sure you have read any chapter mentioned under (a). When you become more familiar with the book you will find that the links highlighted in (a) and (c) will help you see how chemistry interconnects with itself.

Boxes and margin notes

The other things you should look out for are the margin notes and boxes. There are four sorts, and they have all appeared at least once in this chapter.

● Heading

The most important looks like this. Anything in this sort of box is very important—a key concept or a summary. It’s the sort of thing you would do well to hold in your mind as you read or to note down as you learn.

Heading

Boxes like this will contain additional examples, amusing background information, and similar interesting, but inessential, material. The first time you read a chapter, you might want to miss out this sort of box, and only read them later on to flesh out some of the main themes of the chapter.

▶ Sometimes the main text of the book needs clarification or expansion, and this sort of margin note will contain such little extras to help you understand difficult points. It will also remind you of things from elsewhere in the book that illuminate what is being discussed. You would do well to read these notes the first time you read the chapter, though later, as the ideas become more familiar, you might choose to skip them.

■ This sort of margin note will mainly contain cross-references to other parts of the book as a further aid to navigation. You will find an example on p. 000.

End-of-chapter problems

You can’t learn organic chemistry—there’s just too much of it. You can learn trivial things like the names of compounds but that doesn’t help you understand the principles behind the subject. You have to understand the principles because the only way to tackle organic chemistry is to learn to work it out. That is why we have provided end-of-chapter problems. They are to help you discover if you have understood the material presented in each chapter. In general, the 10–15 problems at the end of each chapter start easy and get more difficult. They come in two sorts. The first, generally shorter and easier, allow you to revise the material in that chapter. The second asks you to extend your understanding of the material into areas not covered by the chapter. In the later chapters this second sort will probably revise material from previous chapters.

If a chapter is about a certain type of organic reaction, say elimination reactions (Chapter 19), the chapter itself will describe the various ways (‘mechanisms’) by which the reaction can occur and it will give definitive examples of each mechanism. In Chapter 19 there are three mechanisms and about 65 examples altogether. You might think that this is rather a lot but there are in fact millions of examples known of these three mechanisms and Chapter 19 only scrapes the surface. Even if you totally comprehended the chapter at a first reading, you could not be confident of your understanding about elimination reactions. There are 13 end-of-chapter problems for Chapter 19. The first three ask you to interpret reactions given but not explained in the chapter. This checks that you can use the ideas in familiar situations. The next few problems develop specific ideas from the chapter concerned with why one compound does one reaction while a similar one behaves quite differently.

Finally there are some more challenging problems asking you to extend the ideas to unfamiliar molecules.

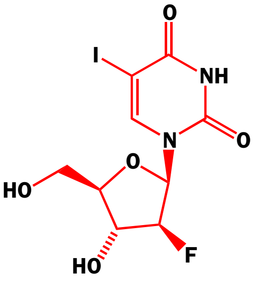
The end-of-chapter problems should set you on your way but they are not the end of the journey to understanding. You are probably reading this text as part of a university course and you should find out what kind of examination problems your university uses and practise them too. Your tutor will be able to advise you on suitable problems for each stage of your development.

The solutions manual

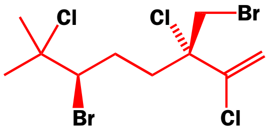
The problems would be of little use to you if you could not check your answers. For the maximum benefit, you need to tackle some or all of the problems as soon as you have finished each chapter without looking at the answers. Then you need to compare your suggestions with ours. You can do this with the solutions manual (*Organic Chemistry: Solutions Manual*, Oxford University Press, 2000). Each problem is discussed in some detail. The purpose of the problem is first stated or explained. Then, if the problem is a simple one, the answer is given. If the problem is more complex, a discussion of possible answers follows with some comments on the value of each. There may be a reference to the source of the problem so that you can read further if you wish.

Colour

You will already have noticed something unusual about this book: almost all of the chemical structures are shown in red. This is quite intentional: emphatic red underlines the message that structures are more important than words in organic chemistry. But sometimes small parts of structures are in other colours: here are two examples from p. 000, where we were talking about organic compounds containing elements other than C and H.

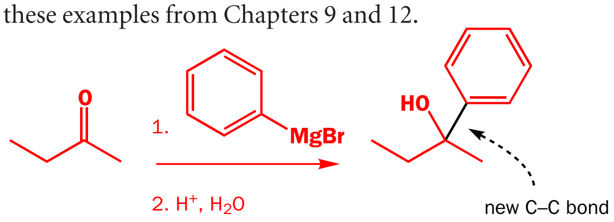


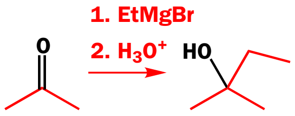
fialuridine
antiviral compound



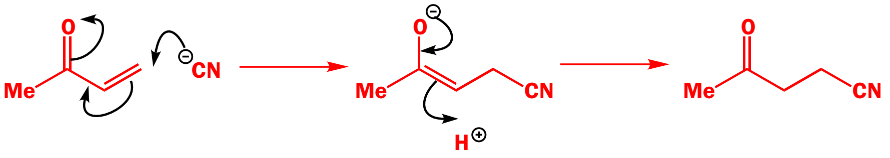
Halomon
naturally occurring antitumour agent

Why are the atom labels black? Because we wanted them to stand out from the rest of the molecule. In general you will see black used to highlight important details of a molecule—they may be the groups taking part in a reaction, or something that has changed as a result of the reaction, as in these examples from Chapters 9 and 12.

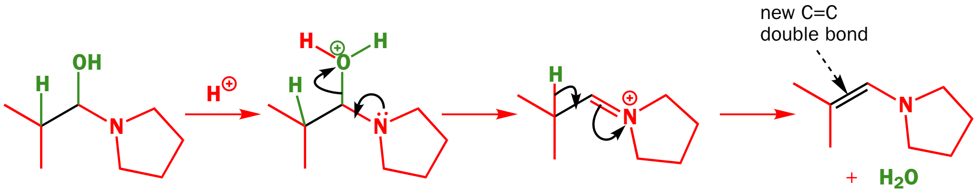




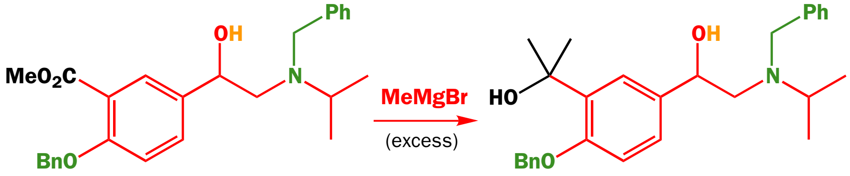
We shall often use black to emphasize ‘curly arrows’, devices that show the movement of electrons, and whose use you will learn about in Chapter 5. Here is an example from Chapter 10: notice black also helps the ‘+’ and ‘–’ charges to stand out.



Occasionally, we shall use other colours such as green, or even orange, yellow, or brown, to highlight points of secondary importance. This example is part of a reaction taken from Chapter 19: we want to show that a molecule of water (H_2O) is formed. The green atoms show where the water comes from. Notice black curly arrows and a new black bond.



Other colours come in when things get more complicated—in this Chapter 24 example, we want to show a reaction happening at the black group in the presence of the yellow H (which, as you will see in Chapter 9, also reacts) and also in the presence of the green 'protecting' groups, one of the topics of Chapter 24.



And, in Chapter 16, colour helps us highlight the difference between carbon atoms carrying four different groups and those with only three different groups. The message is: if you see something in a colour other than red, take special note—the colour is there for a reason.

That is all we shall say in the way of introduction. On the next page the real chemistry starts, and our intention is to help you to learn real chemistry, and to enjoy it.

Organic structures

2

Connections

Building on:

- This chapter does not depend on Chapter 1

Leading to:

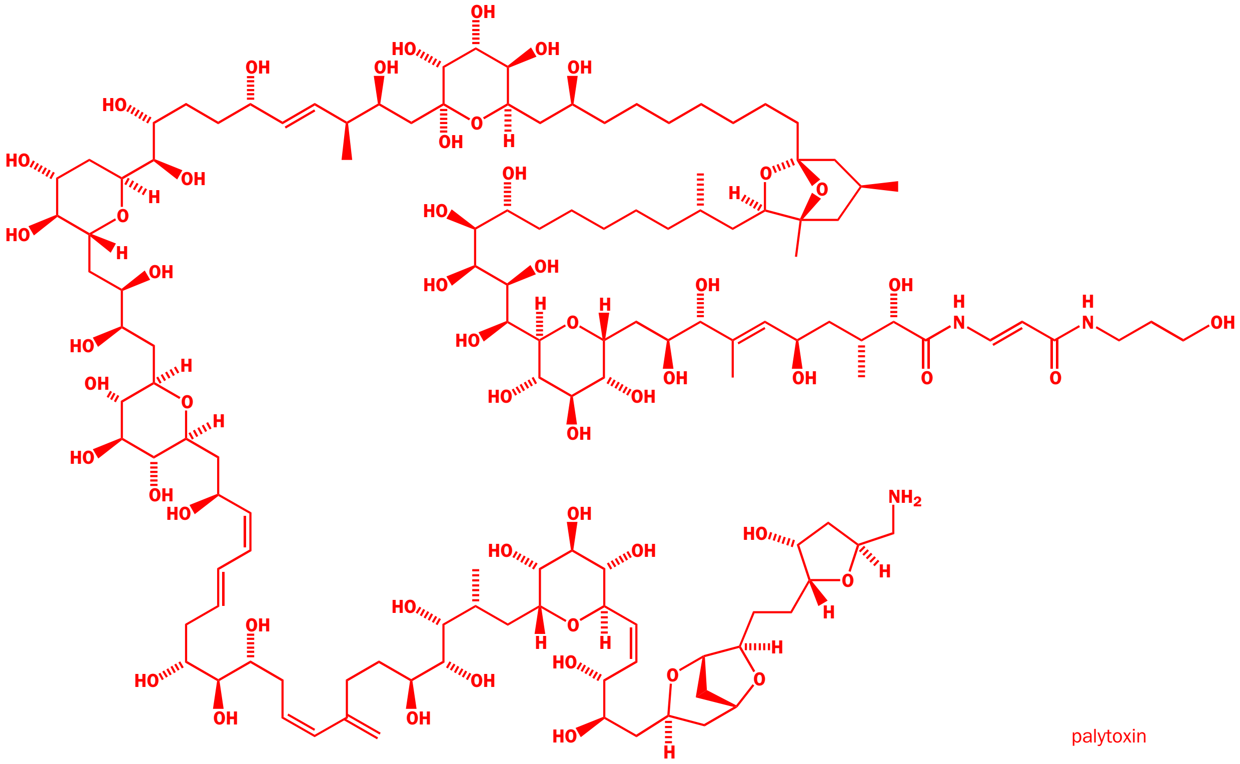
- The diagrams used in the rest of the book
- Why we use these particular diagrams
- How organic chemists name molecules in writing and in speech
- What is the skeleton of an organic molecule
- What is a functional group
- Some abbreviations used by all organic chemists
- Drawing organic molecules realistically in an easily understood style

Looking forward to:

- Ascertaining molecular structure spectroscopically **ch3**
- What determines a molecule's structure **ch4**

There are over 100 elements in the periodic table. Many molecules contain well over 100 atoms—palytoxin, for example (a naturally occurring compound with potential anticancer activity) contains 129 carbon atoms, 221 hydrogen atoms, 54 oxygen atoms, and 3 nitrogen atoms. It's easy to see how chemical structures can display enormous variety, providing enough molecules to build even the most complicated living creatures. But how can we understand what seems like a recipe for confusion? Faced with the collection of atoms we call a molecule, how can we make sense of what we see? This chapter will teach you how to interpret organic structures. It will also teach you how to draw organic molecules in a way that conveys all the necessary information and none of the superfluous.

■ Palytoxin was isolated in 1971 in Hawaii from *Limu make o Hane* ('deadly seaweed of Hana') which had been used to poison spear points. It is one of the most toxic compounds known requiring only about 0.15 microgram per kilogram for death by injection. The complicated structure was determined a few years later.

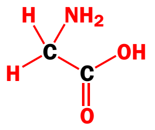


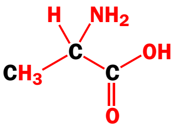
palytoxin

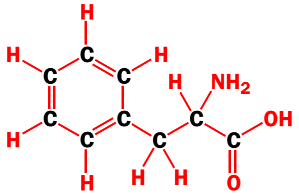
Hydrocarbon frameworks and functional groups

As we explained in Chapter 1, organic chemistry is the study of compounds that contain carbon. Nearly all organic compounds also contain hydrogen; most also contain oxygen, nitrogen, or other elements. Organic chemistry concerns itself with the way in which these atoms are bonded together into stable molecular structures, and the way in which these structures change in the course of chemical reactions.

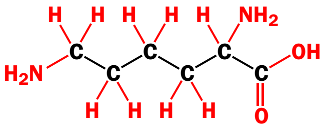
Some molecular structures are shown below. These molecules are all amino acids, the constituents of proteins. Look at the number of carbon atoms in each molecule and the way they are bonded together. Even within this small class of molecules there's great variety—*glycine* and *alanine* have only two or three carbon atoms; *phenylalanine* has nine.

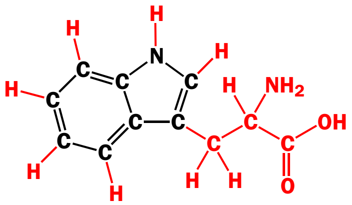




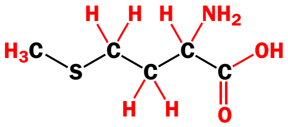


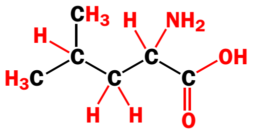
Lysine has a chain of atoms; *tryptophan* has rings.

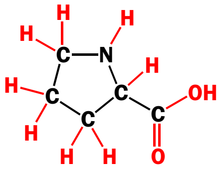




In *methionine* the atoms are arranged in a single chain; in *leucine* the chain is branched. In *proline*, the chain bends back on itself to form a ring.



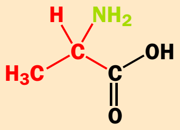




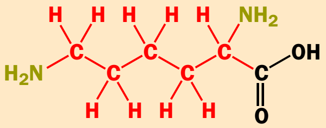
■ We shall return to amino acids as examples several times in this chapter, but we shall leave detailed discussions about their chemistry till Chapters 24 and 49, when we look at the way in which they polymerize to form peptides and proteins.

Yet all of these molecules have similar properties—they are all soluble in water, they are all both acidic and basic (amphoteric), they can all be joined with other amino acids to form proteins. This is because the chemistry of organic molecules depends much less on the number or the arrangement of carbon or hydrogen atoms than on the other types of atoms (O, N, S, P, Si…) in the molecule. We call parts of molecules containing small collections of these other atoms **functional groups**, simply because they are groups of atoms that determine the way the molecule works. All amino acids contain two functional groups: an amino (NH_2 or NH) group and a carboxylic acid (CO_2H) group (some contain other functional groups as well).

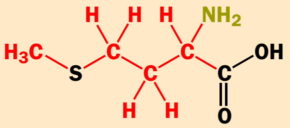
● The **functional groups** determine the way the molecule works both chemically and biologically.



alanine
contains just the amino and carboxylic acid functional groups



lysine
has an additional amino group



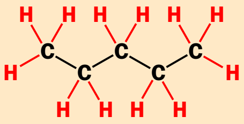
methionine
also has a sulfide functional group

That isn't to say the carbon atoms aren't important; they just play quite a different role from those of the oxygen, nitrogen, and other atoms they are attached to. We can consider the chains and rings of carbon atoms we find in molecules as their skeletons, which support the functional groups and allow them to take part in chemical interactions, much as your skeleton supports your internal organs so they can interact with one another and work properly.

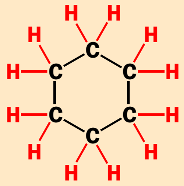
Organic skeletons

Organic molecules left to decompose for millions of years in the absence of light and oxygen become literally carbon skeletons—crude oil, for example, is a mixture of molecules consisting of nothing but carbon and hydrogen, while coal consists of little else but carbon. Although the molecules in coal and oil differ widely in chemical structure, they have one thing in common: no functional groups! Many are very unreactive: about the only chemical reaction they can take part in is combustion, which, in comparison to most reactions that take place in chemical laboratories or in living systems, is an extremely violent process. In Chapter 5 we will start to look at the way that functional groups direct the chemical reactions of a molecule.

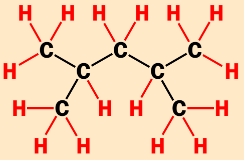
● The **hydrocarbon framework** is made up of chains and rings of carbon atoms, and it acts as a support for the functional groups.



a chain



a ring



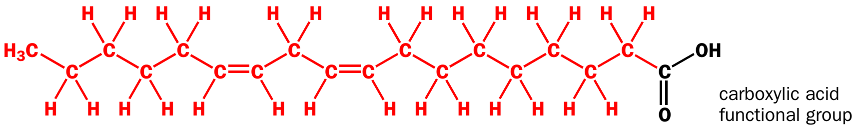
a branched chain

We will see later how the interpretation of organic structures as hydrocarbon frameworks supporting functional groups helps us to understand and rationalize the reactions of organic molecules. It also helps us to devise simple, clear ways of representing molecules on paper. You saw in Chapter 1 how we represented molecules on paper, and in the next section we shall teach you ways to draw (and ways not to draw) molecules—the handwriting of chemistry. *This section is extremely important,* because it will teach you how to communicate chemistry, clearly and simply, throughout your life as a chemist.

Drawing molecules

Be realistic

Below is another organic structure—again, you may be familiar with the molecule it represents; it is a fatty acid commonly called linoleic acid.



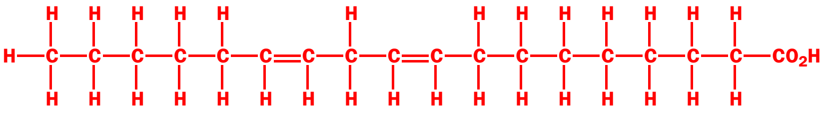
linoleic acid

We could also depict linoleic acid as

$CH_3CH_2CH_2CH_2CH{=}CHCH_2CH{=}CHCH_2CH_2CH_2CH_2CH_2CH_2CH_2CO_2H$

linoleic acid

or as

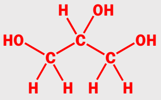


$H{-}CH_2{-}CH_2{-}CH_2{-}CH_2{-}CH_2{-}CH{=}CH{-}CH_2{-}CH{=}CH{-}CH_2{-}CH_2{-}CH_2{-}CH_2{-}CH_2{-}CH_2{-}CH_2{-}CO_2H$

linoleic acid

Three fatty acid molecules and one glycerol molecule combine to form the fats that store energy in our bodies and are used to construct the membranes around our cells. This particular fatty acid, linoleic acid, cannot be manufactured in the human body, and is an essential part of a healthy diet found, for example, in sunflower oil.

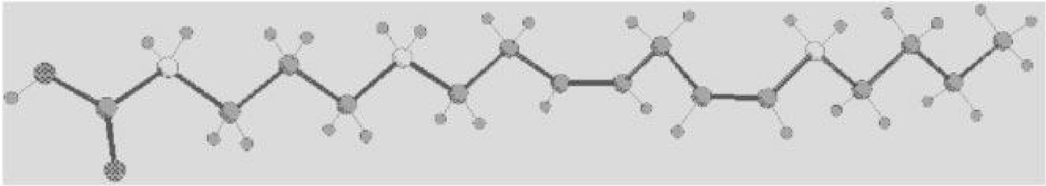
Fatty acids differ in the length of their chains of carbon atoms, yet they have very similar chemical properties because they all contain the carboxylic acid functional group. We shall come back to fatty acids in Chapter 49.



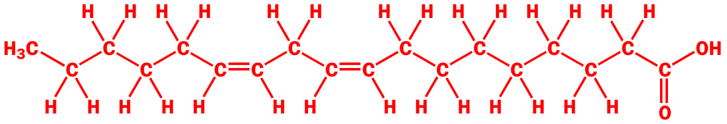
glycerol

You may well have seen diagrams like these last two in older books—they used to be easy to print (in the days before computers) because all the atoms were in a line and all the angles were 90°. But are they realistic? We will consider ways of determining the shapes and structures of molecules in more detail in Chapter 3, but the picture below shows the structure of linoleic acid determined by X-ray crystallography.

▶ X-ray crystallography discovers the structures of molecules by observing the way X-rays bounce off atoms in crystalline solids. It gives clear diagrams with the atoms marked a circles and the bonds as rods joining them together.



You can see that the chain of carbon atoms is not linear, but a zig-zag. Although our diagram is just a two-dimensional representation of this three-dimensional structure, it seems reasonable to draw it as a zig-zag too.



linoleic acid

This gives us our first guideline for drawing organic structures.

● Guideline 1

Draw chains of atoms as zig-zags

Realism of course has its limits—the X-ray structure shows that the linoleic acid molecule is in fact slightly bent in the vicinity of the double bonds; we have taken the liberty of drawing it as a 'straight zig-zag'. Similarly, close inspection of crystal structures like this reveals that the angle of the zig-zag is about 109° when the carbon atom is not part of a double bond and 120° when it is. The 109° angle is the 'tetrahedral angle', the angle between two vertices of a tetrahedron when viewed from its centre. In Chapter 4 we shall look at why carbon atoms take up this particular arrangement of bonds. Our realistic drawing is a projection of a three-dimensional structure onto flat paper so we have to compromise.

Be economical

When we draw organic structures we try to be as realistic as we can be without putting in superfluous detail. Look at these three pictures.

1

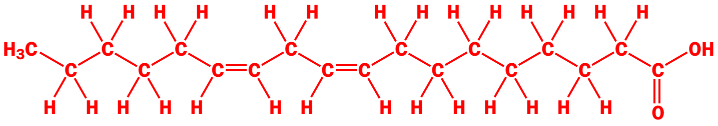
2

3

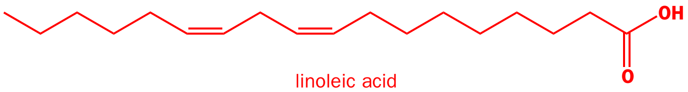
(1) is immediately recognizable as Leonardo da Vinci's Mona Lisa. You may not recognize (2)—it's also Leonardo da Vinci's Mona Lisa—this time viewed from above. The frame is very ornate, but the picture tells us as much about the painting as our rejected linear and 90° angle diagrams did about

our fatty acid. They're both correct—in their way—but sadly useless. What we need when we draw molecules is the equivalent of (3). It gets across the idea of the original, and includes all the detail necessary for us to recognize what it's a picture of, and leaves out the rest. And it was quick to draw—this picture was drawn in less than 10 minutes: we haven't got time to produce great works of art!

Because functional groups are the key to the chemistry of molecules, clear diagrams must emphasize the functional groups, and let the hydrocarbon framework fade into the background. Compare the diagrams below:



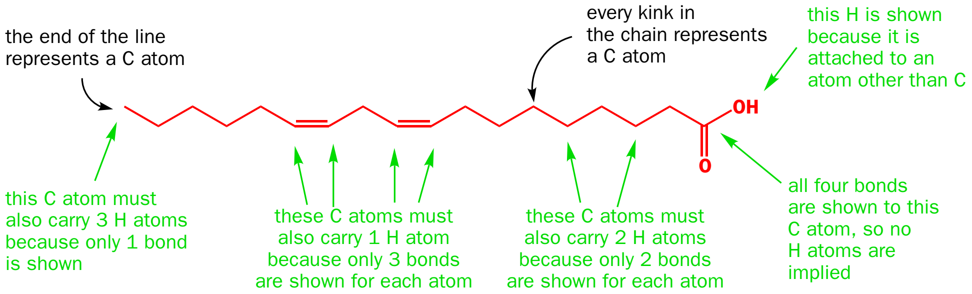
linoleic acid



linoleic acid

The second structure is the way that most organic chemists would draw linoleic acid. Notice how the important carboxylic acid functional group stands out clearly and is no longer cluttered by all those Cs and Hs. The zig-zag pattern of the chain is much clearer too. And this structure is much quicker to draw than any of the previous ones!

To get this diagram from the one above we've done two things. Firstly, we've got rid of all the hydrogen atoms attached to carbon atoms, along with the bonds joining them to the carbon atoms. Even without drawing the hydrogen atoms we know they're there—we assume that any carbon atom that doesn't appear to have its potential for four bonds satisfied is also attached to the appropriate number of hydrogen atoms. Secondly, we've rubbed out all the Cs representing carbon atoms. We're left with a zig-zag line, and we assume that every kink in the line represents a carbon atom, as does the end of the line.



We can turn these two simplifications into two more guidelines for drawing organic structures.

● **Guideline 2**

Miss out the Hs attached to carbon atoms, along with the C–H bonds (unless there is a good reason not to)

● **Guideline 3**

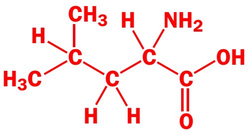
Miss out the capital Cs representing carbon atoms (unless there is a good reason not to)

What is 'a good reason not to'? One is if the C or H is part of a functional group. Another is if the C or H needs to be highlighted in some way, for example, because it's taking part in a reaction. Don't be too rigid about these guidelines: they're not rules. Better is just to learn by example (you'll find plenty in this book): if it helps clarify, put it in; if it clutters and confuses, leave it out. One thing you must remember, though: if you write a carbon atom as a letter C then you *must* add all the H atoms too. If you don't want to draw all the Hs, don't write C for carbon.

Be clear

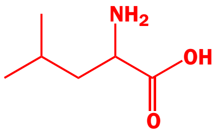
Try drawing some of the amino acids represented on p. 000 in a similar way, using the three guidelines. The bond angles at tetrahedral carbon atoms are about 109°. Make them look about 109° projected on to a plane! (120° is a good compromise, and it makes the drawings look neat.)

Start with leucine — earlier we drew it as the structure to the right. Get a piece of paper and do it now; then see how your drawing compares with our suggestions.

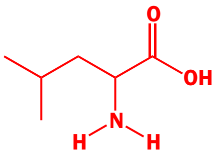


leucine

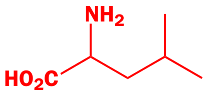
It doesn't matter which way up you've drawn it, but your diagram should look something like one of these structures below.



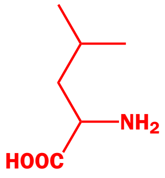
leucine



leucine



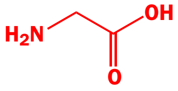
leucine



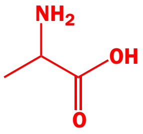
leucine

The guidelines we gave were only guidelines, not rules, and it certainly does not matter which way round you draw the molecule. The aim is to keep the functional groups clear, and let the skeleton fade into the background. That's why the last two structures are all right—the carbon atom shown as 'C' is part of a functional group (the carboxyl group) so it can stand out.

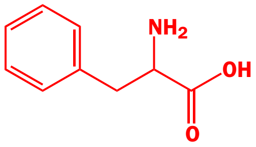
Now turn back to p. 000 and try redrawing the some of the other eight structures there using the guidelines. Don't look at our suggestions below until you've done them! Then compare your drawings with our suggestions.



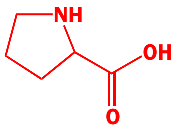
glycine



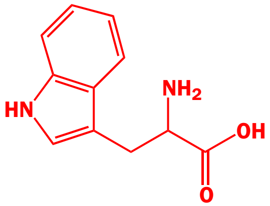
alanine



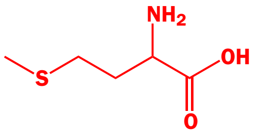
phenylalanine



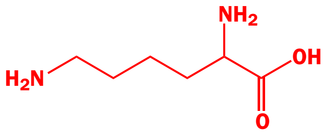
proline



tryptophan



methionine

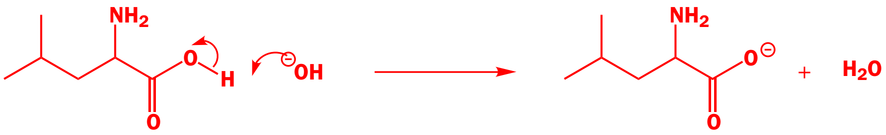


lysine

Remember that these are only suggestions, but we hope you'll agree that this style of diagram looks much less cluttered and makes the functional groups much clearer than the diagrams on p. 000. Moreover, they still bear significant resemblance to the 'real thing'—compare these crystal structures of lysine and tryptophan with the structures shown above, for example.

Structural diagrams can be modified to suit the occasion

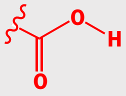
You'll probably find that you want to draw the same molecule in different ways on different occasions to emphasize different points. Let's carry on using leucine as an example. We mentioned before that an amino acid can act as an acid or as a base. When it acts as an acid, a base (for example, hydroxide, OH^-) removes H^+ from the carboxylic acid group in a reaction we can represent as



Not all chemists put circles round their plus and minus charges—it's a matter of personal choice.

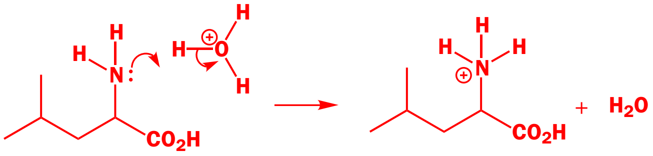
The product of this reaction has a negative charge on an oxygen atom. We have put it in a circle to make it clearer, and we suggest you do the same when you draw charges: +'s and –'s are easily mislaid. We shall discuss this type of reaction, the way in which reactions are drawn, and what the 'curly arrows' in the diagram mean in Chapter 5. But for now, notice that we drew out the CO_2H as the fragment left because we wanted to show how the O–H bond was broken when the base attacked. We modified our diagram to suit our own purposes.

The wiggly line is a graphical way of indicating an incomplete structure: it shows where we have mentally 'snapped off' the CO_2H group from the rest of the molecule.



When leucine acts as a base, the amino (NH_2) group is involved. The nitrogen atom attaches itself to a proton, forming a new bond using its *lone pair*.

We can represent this reaction as



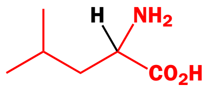
A lone pair is a pair of electrons that is not involved in a chemical bond We shall discuss lone pairs in detail in Chapter 4. Again, don't worry about what the curly arrows in this diagram mean—we will cover them in detail in Chapter 5.

Notice how we drew the lone pair at this time because we wanted to show how it was involved in the reaction. The oxygen atoms of the carboxylic acid groups also have lone pairs but we didn't draw them in because they weren't relevant to what we were talking about. Neither did we feel it was necessary to draw CO_2H in full this time because none of the atoms or bonds in the carboxylic acid functional group was involved in the reaction.

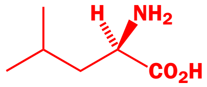
Structural diagrams can show three-dimensional information on a two-dimensional page

Of course, all the structures we have been drawing only give an idea of the real structure of the molecules. For example, the carbon atom between the NH_2 group and the CO_2H group of leucine has a tetrahedral arrangement of atoms around it, a fact which we have so far completely ignored.

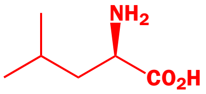
We might want to emphasize this fact by drawing in the hydrogen atom we missed out at this point as in structure 1 (in the right-hand margin).



1



2



3

We can then show that one of the groups attached to this carbon atom comes towards us, out of the plane of the paper, and the other one goes away from us, into the paper. There are several ways of doing this. In structure 2, the bold, wedged bond suggests a perspective view of a bond coming towards you, while the hashed bond suggests a bond fading away from you. The other two 'normal' bonds are in the plane of the paper.

Alternatively we could miss out the hydrogen atom and draw something a bit neater though slightly less realistic as structure 3.

We can assume the missing hydrogen atom is behind the plane of the paper, because that is where the 'missing' vertex of the tetrahedron of atoms attached to the carbon atom lies. These conventions allow us to give an idea of the three-dimensional shape (stereochemistry) of any organic molecule—you have already seen them in use in the diagram of the structure of palytoxin at the beginning of this chapter.

When you draw diagrams like these to indicate the three-dimensional shape of the molecule, try to keep the hydrocarbon framework in the plane of the paper and allow functional groups and other branches to project forwards out of the paper or backwards into it.

Reminder

Organic structures should be drawn to be *realistic, economical, clear.*

We gave three guidelines to help you achieve this when you draw structures:

- Guideline 1: Draw chains of atoms as zig-zags
- Guideline 2: Miss out the Hs attached to carbon atoms, along with the C–H bonds
- Guideline 3: Miss out the capital Cs representing carbon atoms

We shall look in more detail at the shapes of molecules—their *stereochemistry*—in Chapter 16.

The guidelines we have given and conventions we have illustrated in this section have grown up over decades. They are used by organic chemists because they work! We guarantee to follow them for the rest of the book—try to follow them yourself whenever you draw an organic structure. Before you ever draw a capital C or a capital H again, ask yourself whether it's really necessary!

Now that we have considered how to draw structures, we can return to some of the structural types that we find in organic molecules. Firstly, we'll talk about hydrocarbon frameworks, then about functional groups.

Hydrocarbon frameworks

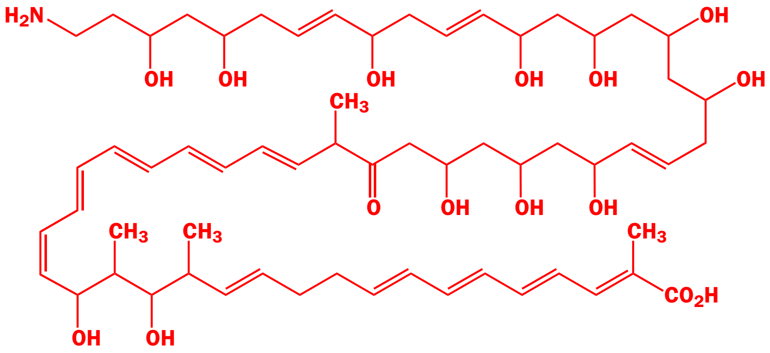
Carbon as an element is unique in the variety of structures it can form. It is unusual because it forms strong, stable bonds to the majority of elements in the periodic table, including itself. It is this ability to form bonds to itself that leads to the variety of organic structures that exist, and indeed to the possibility of life existing at all. Carbon may make up only 0.2% of the earth's crust, but it certainly deserves a whole branch of chemistry all to itself.

Chains

The simplest class of hydrocarbon frameworks contains just chains of atoms. The fatty acids we met earlier have hydrocarbon frameworks made of zig-zag chains of atoms, for example. Polythene is a polymer whose hydrocarbon framework consists entirely of chains of carbon atoms.

a section of the structure of polythene

At the other end of the spectrum of complexity is this antibiotic, extracted from a fungus in 1995 and aptly named linearmycin as it has a long linear chain. The chain of this antibiotic is so long that we have to wrap it round two corners just to get it on the page.



We haven't drawn whether the CH_3 groups and OH groups are in front of or behind the plane of the paper, because (at the time of writing this book) no one yet knows. The stereochemistry of linearmycin is unknown.

▶ Notice we've drawn in four groups as CH_3—we did this because we didn't want them to get overlooked in such a large structure. They are the only tiny branches off this long winding trunk.

Names for carbon chains

It is often convenient to refer to a chain of carbon atoms by a name indicating its length. You have probably met some of these names before in the names of the simplest organic molecules, the alkanes. There are also commonly used abbreviations for these names: these can be very useful in both writing about chemistry and in drawing chemical structures, as we shall see shortly.

▶ The names for shorter chains (which you must learn) exist for historical reasons; for chains of 5 or more carbon atoms, the systematic names are based on Greek number names.

Names and abbreviations for carbon chains

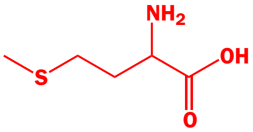
Number of carbon atoms in chain	Name of group	Formula†	Abbreviation	Name of alkane (= chain + H)
1	methyl	$-CH_3$	Me	methane
2	ethyl	$-CH_2CH_3$	Et	ethane
3	propyl	$-CH_2CH_2CH_3$	Pr	propane
4	butyl	$-(CH_2)_3CH_3$	Bu	butane
5	pentyl	$-(CH_2)_4CH_3$	—‡	pentane
6	hexyl	$-(CH_2)_5CH_3$	—‡	hexane
7	heptyl	$-(CH_2)_6CH_3$	—‡	heptane
8	octyl	$-(CH_2)_7CH_3$	—‡	octane
9	nonyl	$-(CH_2)_8CH_3$	—‡	nonane
10	decyl	$-(CH_2)_9CH_3$	—‡	decane

† This representation is not recommended.

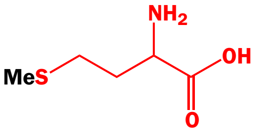
‡ Names for longer chains are not commonly abbreviated.

Organic elements

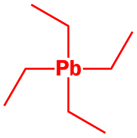
You may notice that the abbreviations for the names of carbon chains look very much like the symbols for chemical elements: this is deliberate, and these symbols are sometimes called 'organic elements'. They can be used in chemical structures just like element symbols. It is often convenient to use the 'organic element' symbols for short carbon chains for tidiness. Here are some examples. Structure 1 to the right shows how we drew the structure of the amino acid methionine on p. 000. The stick representing the methyl group attached to the sulfur atom does, however, look a little odd. Most chemists would draw methionine as structure 2, with 'Me' representing the CH_3 (methyl) group. Tetraethyllead used to be added to petrol to prevent engines 'knocking', until it was shown to be a health hazard. Its structure (as you might easily guess from the name) is shown as item 3. But it's much easier to write as $PbEt_4$ or Et_4Pb



1 methionine

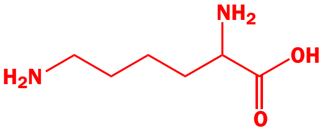


2 methionine



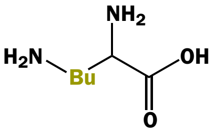
3 tetraethyllead

Remember that these symbols (and names) can only be used for terminal chains of atoms. We couldn't abbreviate the structure of lysine from



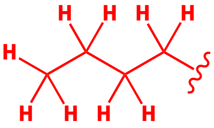
lysine

to

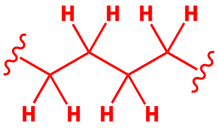


NOT CORRECT

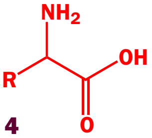
for example, because Bu represents



and not



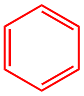
Before leaving carbon chains, we must mention one other very useful organic element symbol, R. R in a structure can mean *anything*—it's a sort of wild card. For example, structure 4 would indicate any amino acid, where R = H is glycine, R = Me is alanine… As we've mentioned before, and you will see later, the reactivity of organic molecules is so dependent on their functional groups that the rest of the molecule can be irrelevant. In these cases, we can choose just to call it R.



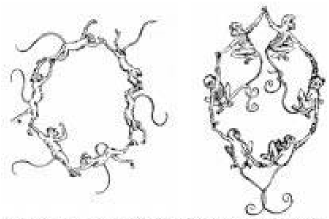
4

Carbon rings

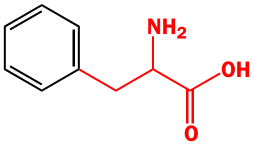
Rings of atoms are also common in organic structures. You may have heard the famous story of Auguste Kekulé first realizing that benzene has a ring structure when he dreamed of snakes biting their own tails. You have met benzene rings in phenylalanine and aspirin. Paracetamol also has a structure based on a benzene ring.



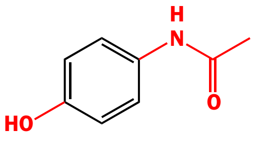
benzene



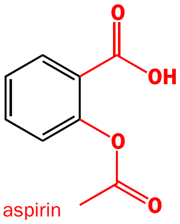
Kekul's snake dream inspired these figures that appeared in a spoof edition of the German chemical Journal, *Berichte der Deutschen Chemischen Gesellschaft* in 1886



phenylalanine

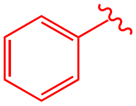


paracetamol

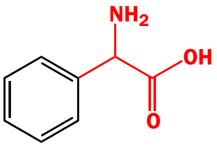


aspirin

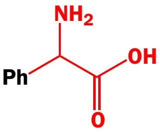
When a benzene ring is attached to a molecule by only *one* of its carbon atoms (as in phenylalanine, but not paracetamol or aspirin), we can call it a 'phenyl' group and give it the organic element symbol Ph.



the phenyl group, Ph

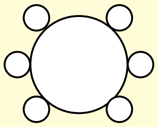


is equivalent to



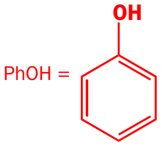
Benzene has a ring structure

In 1865, August Kekulé presented a paper at the Academie des Sciences in Paris suggesting a cyclic structure for benzene, the inspiration for which he ascribed to a dream. However, was Kekulé the first to suggest that benzene was cyclic? Some believe not, and credit an Austrian schoolteacher, Josef Loschmidt with the first depiction of cyclic benzene structures. In 1861, 4 years before Kekulé's dream, Loschmidt published a book in which he represented benzene as a set of rings. It is not certain whether Loschmidt or Kekulé—or even a Scot named Archibald Couper—got it right first.



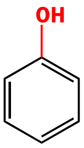
Any compound containing a benzene ring, or a related (Chapter 7) ring system is known as 'aromatic', and another useful organic element symbol related to Ph is Ar (for 'aryl'). While Ph always means C_6H_5, Ar can mean any *substituted* phenyl ring, in other words, phenyl with any number of the hydrogen atoms replaced by other groups.

Of course, Ar = argon too, but so few argon compounds exist that there is never any confusion.

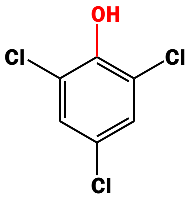


phenol

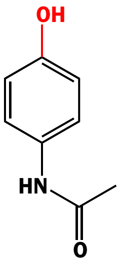
For example, while PhOH always means phenol, ArOH could mean phenol, 2,4,6-trichlorophenol (the antiseptic TCP), paracetamol or aspirin (among many other substituted phenols). Like R, the 'wild card' alkyl group, Ar is a 'wild card' *aryl* group.



phenol

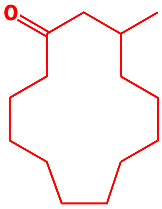


2,4,6-trichlorophenol



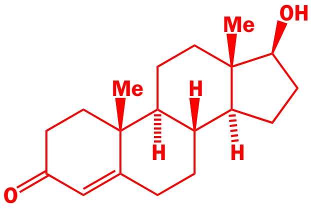
paracetamol

The compound known as muscone has only relatively recently been made in the lab. It is the pungent aroma that makes up the base-note of musk fragrances. Before chemists had determined its structure and devised a laboratory synthesis the only source of musk was the musk deer, now rare for this very reason. Muscone's skeleton is a 13-membered ring of carbon atoms.

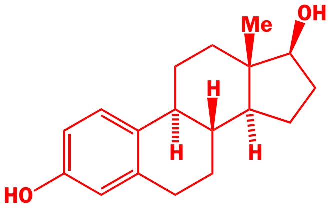


muscone

The steroid hormones have several (usually four) rings fused together. These are testosterone and oestradiol, the important human male and female sex hormones.

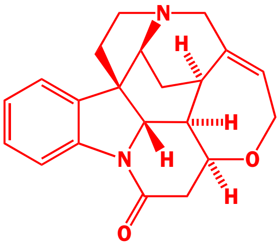


testosterone

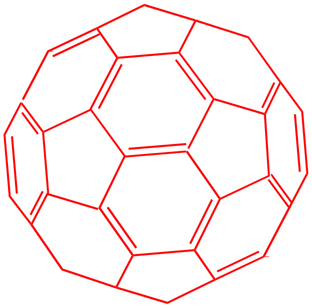


oestradiol

Some ring structures are much more complicated. The potent poison strychnine is a tangle of interconnecting rings.



strychnine


Buckminsterfullerene

Buckminsterfullerene

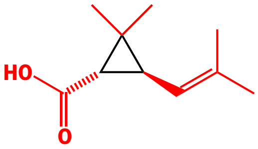
Buckminsterfullerene is named after the American inventor and architect Richard Buckminster Fuller, who designed the structures known as 'geodesic domes'.

One of the most elegant ring structures is shown above and is known as Buckminsterfullerene. It consists solely of 60 carbon atoms in rings that curve back on themselves to form a football-shaped cage.

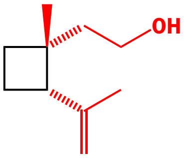
Count the number of bonds at any junction and you will see they add up to four so no hydrogens need be added. This compound is C_{60}. Note that you can't see all the atoms as some are behind the sphere.

Rings of carbon atoms are given names starting with 'cyclo', followed by the name for the carbon chain with the same number of carbon atoms.

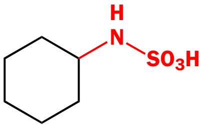
To the right, structure 1 shows chrysanthemic acid, part of the naturally occurring pesticides called pyrethrins (an example appears in Chapter 1), which contains a cyclopropane ring. Propane has three carbon atoms. Cyclopropane is a three-membered ring. Grandisol (structure 2), an insect pheromone used by male boll weevils to attract females, has a structure based on a cyclobutane ring. Butane has four carbon atoms. Cyclobutane is a four-membered ring. Cyclamate (structure 3), formerly used as an artificial sweetener, contains a cyclohexane ring. Hexane has six carbon atoms. Cyclohexane is a six-membered ring.



1 chrysanthemic acid



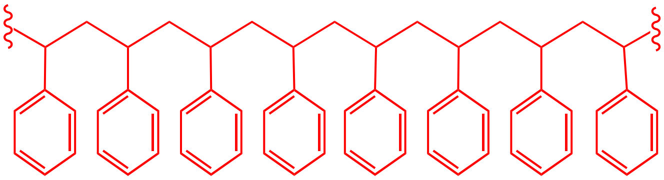
2 grandisol



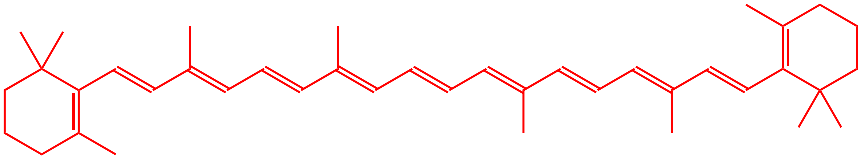
3 cyclamate

Branches

Hydrocarbon frameworks rarely consist of single rings or chains, but are often branched. Rings, chains, and branches are all combined in structures like that of the marine toxin palytoxin that we met at the beginning of the chapter, polystyrene, a polymer made of six-membered rings dangling from linear carbon chains, or of β-carotene, the compound that makes carrots orange.

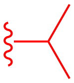


part of the structure of polystyrene

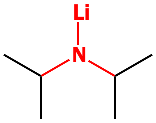


β-carotene

Just like some short straight carbon chains, some short branched carbon chains are given names and organic element symbols. The most common is the isopropyl group. Lithium diisopropylamide (also called LDA) is a strong base commonly used in organic synthesis.



the isopropyl group
i-Pr

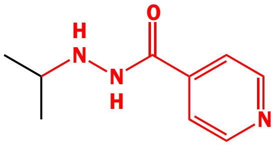


is equivalent to LiN*i*-Pr$_2$

lithium diisopropylamide (LDA)

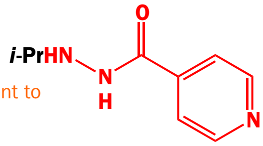
Iproniazid is an antidepressant drug with *i*-Pr in both structure and name.

Notice how the 'propyl' part of 'isopropyl' still indicates three carbon atoms; they are just joined together in a different way—in other words, as an *isomer* of the straight chain propyl group. Sometimes, to avoid confusion, the straight chain alkyl groups are called '*n*-alkyl' (for example, *n*-Pr, *n*-Bu)—*n* for 'normal'—to distinguish them from their branched counterparts.

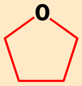


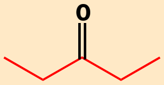
iproniazid

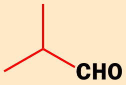
is equivalent to



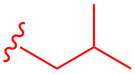
● ***Isomers*** **are molecules with the same kinds and numbers of atoms joined up in different ways** ***n*****-propanol,** ***n*****-PrOH, and isopropanol,** ***i*****-PrOH, are isomeric alcohols. Isomers need not have the same functional groups, these compounds are all isomers of C_4H_8O.**





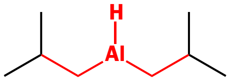


■ 'Isopropyl' may be abbreviated to *i*-Pr, *i*Pr, or Pr^i. We will use the first in this book, but you may see the others used elsewhere.



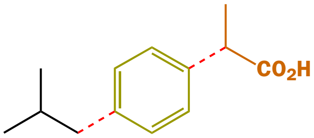
the isobutyl group
i-Bu

The isobutyl (*i*-Bu) group is a CH_2 group joined to an *i*-Pr group. It is *i*-PrCH_2–
Two isobutyl groups are present in the reducing agent diisobutyl aluminium hydride (DIBAL).



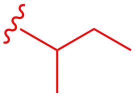
diisobutyl aluminium hydride (DIBAL)
is equivalent to **HAl*i*-Bu$_2$**

The painkiller ibuprofen (marketed as Nurofen®) contains an isobutyl group.



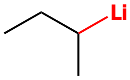
Ibuprofen

■
Notice how the invented name ibuprofen is a medley of 'ibu' (from i-Bu for isobutyl) + 'pro' (for propyl, the three-carbon unit shown in gold) + 'fen' (for the phenyl ring). We will talk about the way in which compounds are named later in this chapter.

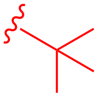


the *sec*-butyl group
s-Bu

There are two more isomers of the butyl group, both of which have common names and abbreviations. The *sec*-butyl group (*s*-butyl or *s*-Bu) has a methyl and an ethyl group joined to the same carbon atom. It appears in an organolithium compound, *sec*-butyl lithium, used to introduce lithium atoms into organic molecules.

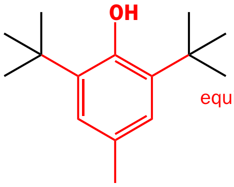


is equivalent to *s*-**BuLi**



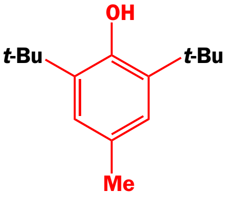
the *tert*-butyl group
t-Bu

The *tert*-butyl group (*t*-butyl or *t*-Bu) group has three methyl groups joined to the same carbon atom. Two *t*-Bu groups are found in BHT ('butylated hydroxy toluene'), an antioxidant added to some processed foods.



BHT

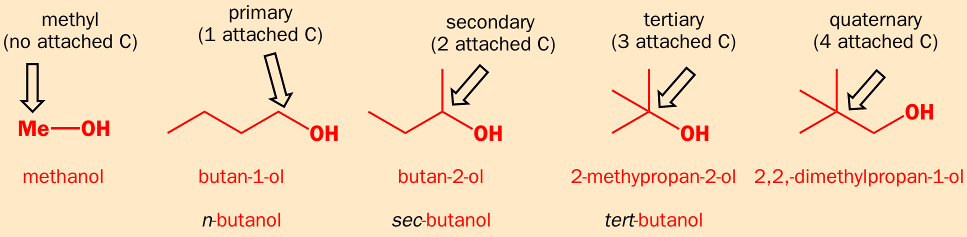
is equivalent to



BHT

Primary, secondary, and tertiary

The prefixes *sec* and *tert* are really short for secondary and tertiary, terms that refer to the carbon atom that attaches these groups to the rest of the molecular structure.

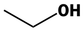


A primary carbon atom is attached to only one other C atom, a secondary to two other C atoms, and so on. This means there are five types of carbon atom.

These names for bits of hydrocarbon framework are more than just useful ways of writing or talking about chemistry. They tell us something fundamental about the molecule and we shall use them when we describe reactions.

This quick architectural tour of some of the molecular edifices built by nature and by man serves just as an introduction to some of the hydrocarbon frameworks you will meet in the rest of this chapter and of this book. Yet, fortunately for us, however complicated the hydrocarbon framework might be, it serves only as a support for the functional groups. And, by and large, a functional group in one molecule behaves in much the same way as it does in another molecule. What we now need to do, and we start in the next section, is to introduce you to some functional groups, and to explain why it is that their attributes are the key to understanding organic chemistry.

Functional groups

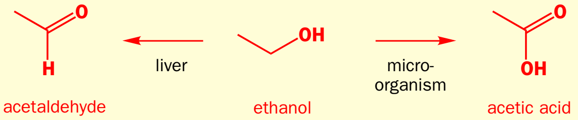
If you can take ethane gas (CH_3CH_3, or EtH, or even , though a single line like this doesn't look much like a chemical structure) and bubble it through acids, bases, oxidizing agents, reducing agents—in fact almost any chemical you can think of—it will remain unchanged. Just about the only thing you can do with it is burn it. Yet ethanol (CH_3CH_2OH, or , or preferably EtOH) not only burns, it reacts with acids, bases, and oxidizing agents.

The difference between ethanol and ethane is the functional group—the OH or hydroxyl group. We know that these chemical properties (being able to react with acids, bases, and oxidizing agents) are properties of the hydroxyl group and not just of ethanol because other compounds containing OH groups (in other words, other alcohols) have similar properties, whatever their hydrocarbon frameworks.

Your understanding of functional groups will be the key to your understanding of organic chemistry. We shall therefore now go on to meet some of the most important functional groups. We won't say much about the properties of each group; that will come in Chapter 5 and later. Your task at this stage is to learn to recognize them when they appear in structures, so make sure you learn their names. The classes of compound associated with some functional groups also have names: for example, compounds containing the hydroxyl group are known as alcohols. Learn these names too as they are more important than the systematic names of individual compounds. We've told you a few snippets of information about each group to help you to get to know something of the group's character.

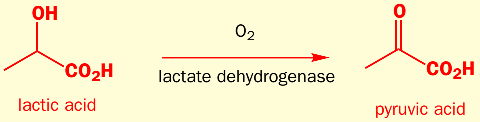
Ethanol

The reaction of ethanol with oxidizing agents makes vinegar from wine and sober people from drunk ones. In both cases, the oxidizing agent is oxygen from the air, catalysed by an enzyme in a living system. The oxidation of ethanol by microorganisms that grow in wine left open to the air leads to acetic acid (ethanoic acid) while the oxidation of ethanol by the liver gives acetaldehyde (ethanal).



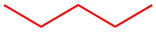
Human metabolism and oxidation

The human metabolism makes use of the oxidation of alcohols to render harmless other toxic compounds containing the OH group. For example, lactic acid, produced in muscles during intense activity, is oxidized by an enzyme called lactate dehydrogenase to the metabolically useful compound pyruvic acid.

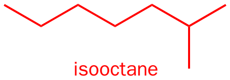


Alkanes contain no functional groups

The alkanes are the simplest class of organic molecules because they contain no functional groups. They are extremely unreactive, and therefore rather boring as far as the organic chemist is concerned. However, their unreactivity can be a bonus, and alkanes such as pentane and hexane are often used as solvents, especially for purification of organic compounds. Just about the only thing alkanes will do is burn—methane, propane, and butane are all used as domestic fuels, and petrol is a mixture of alkanes containing largely isooctane.

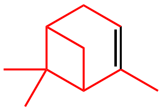

pentane

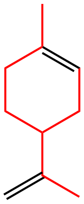

hexane


isooctane

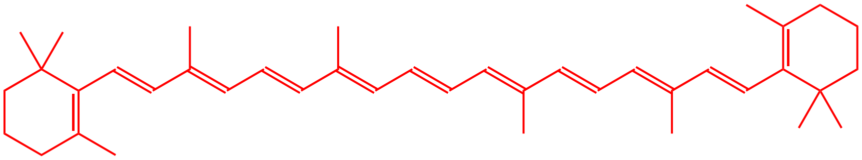
Alkenes (sometimes called olefins) contain C=C double bonds

It may seem strange to classify a type of bond as a functional group, but you will see later that C=C double bonds impart reactivity to an organic molecule just as functional groups consisting of, say, oxygen or nitrogen atoms do. Some of the compounds produced by plants and used by perfumers are alkenes (see Chapter 1). For example, pinene has a smell evocative of pine forests, while limonene smells of citrus fruits.


α-pinene



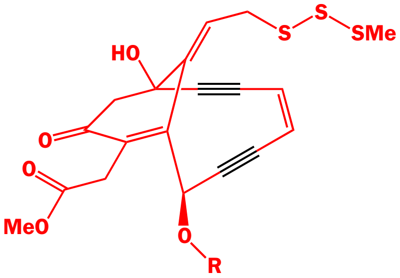
You've already met the orange pigment β-carotene. Eleven C=C double bonds make up most of its structure. Coloured organic compounds often contain chains of C=C double bonds like this. In Chapter 7 you will find out why this is so.



β-carotene

Alkynes contain C≡C triple bonds

Just like C=C double bonds, C≡C triple bonds have a special type of reactivity associated with them, so it's useful to call a C≡C triple bond a functional group. Alkynes are **linear** so we draw them with four carbon atoms in a straight line. Alkynes are not as widespread in nature as alkenes, but one fascinating class of compounds containing C≡C triple bonds is a group of antitumour agents discovered during the 1980s. Calicheamicin is a member of this group. The high reactivity of this combination of functional groups enables calicheamicin to attack DNA and prevent cancer cells from proliferating. For the first time we have drawn a molecule in three dimensions, with two bonds crossing one another—can you see the shape?



calicheamicin
(R = a string of sugar molecules)

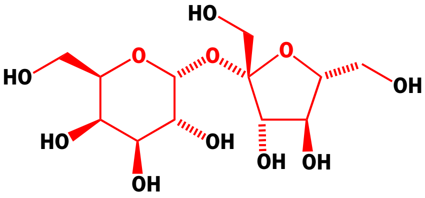
▶ Saturated and unsaturated carbon atoms

In an alkane, each carbon atom is joined to four other atoms (C or H). It has no potential for forming more bonds and is therefore **saturated**. In alkenes, the carbon atoms making up the C=C double bond are attached to only three atoms each. They still have the potential to bond with one more atom, and are therefore **unsaturated**. In general, carbon atoms attached to four other atoms are saturated; those attached to three, two, or one are unsaturated.

Alcohols (R–OH) contain a hydroxyl (OH) group

We've already talked about the hydroxyl group in ethanol and other alcohols. Carbohydrates are peppered with hydroxyl groups; sucrose has eight of them for example (a more three-dimensional picture of the sucrose molecule appears in Chapter 1).

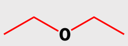
Molecules containing hydroxyl groups are often soluble in water, and living things often attach sugar groups, containing hydroxyl groups, to otherwise insoluble organic compounds to keep them in solution in the cell. Calicheamicin, a molecule we have just mentioned, contains a string of sugars for just this reason. The liver carries out its task of detoxifying unwanted organic compounds by repeatedly hydroxylating them until they are water-soluble, and they are then excreted in the bile or urine.



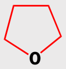
sucrose

■ Remember that R can mean any alkyl group.

■ If we want a structure to contain more than one 'R', we give the R's numbers and call them R^1, R^2... Thus R^1–O–R^2 means an ether with two different unspecified alkyl groups. (*Not* R$_1$, R$_2$..., which would mean 1 × R, 2 × R...)



diethyl ether
"ether"



THF

Ethers (R^1–O–R^2) contain an alkoxy group (–OR)

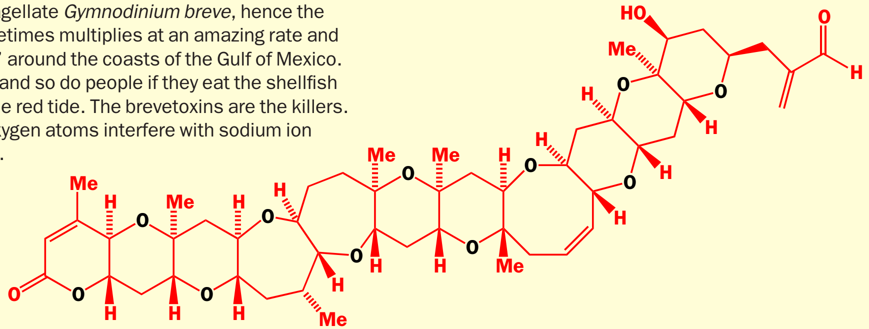
The name **ether** refers to any compound that has two alkyl groups linked through an oxygen atom. 'Ether' is also used as an everyday name for diethyl ether, Et$_2$O. You might compare this use of the word 'ether' with the common use of the word 'alcohol' to mean ethanol. Diethyl ether is a highly flammable solvent that boils at only 35 °C. It used to be used as an anaesthetic. Tetrahydrofuran (THF) is another commonly used solvent and is a cyclic ether.

Brevetoxin B is a fascinating naturally occurring compound that was synthesized in the laboratory in 1995. It is packed with ether functional groups in ring sizes from 6 to 8.

■ Another common laboratory solvent is called 'petroleum ether'. Don't confuse this with diethyl ether! Petroleum ether is in fact not an ether, but a mixture of alkanes. 'Ether', according to the *Oxford English Dictionary*, means 'clear sky, upper region beyond the clouds', and hence used to be used for anything light, airy, and volatile.

Brevetoxin B

Brevetoxin B is one of a family of polyethers found in a sea creature (a dinoflagellate *Gymnodinium breve*, hence the name) which sometimes multiplies at an amazing rate and creates 'red tides' around the coasts of the Gulf of Mexico. Fish die in shoals and so do people if they eat the shellfish that have eaten the red tide. The brevetoxins are the killers. The many ether oxygen atoms interfere with sodium ion (Na^+) metabolism.

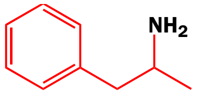


brevetoxin

Amines (R–NH_2) contain the amino (NH_2) group

We met the amino group when we were discussing the amino acids: we mentioned that it was this group that gave these compounds their basic properties. Amines often have powerful fishy smells: the smell of putrescine is particularly foul. It is formed as meat decays. Many neurologically active compounds are also amines: amphetamine is a notorious stimulant.

putrescine

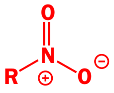


amphetamine

Nitro compounds (R–NO_2) contain the nitro group (NO_2)

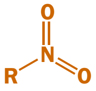
The nitro group (NO_2) is often incorrectly drawn with five bonds to nitrogen which you will see in Chapter 4, is impossible. Make sure you draw it correctly when you need to draw it out in detail. If you write just NO_2 you are all right!

Several nitro groups in one molecule can make it quite unstable and even explosive. Three nitro groups give the most famous explosive of all, TNT (trinitrotoluene), its kick.

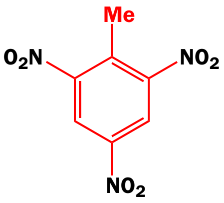


the nitro group

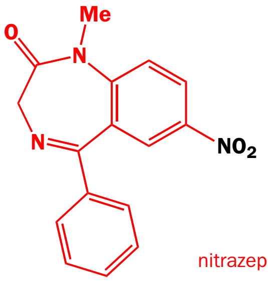
nitrogen cannot have five bonds!



incorrect structure for the nitro group



TNT

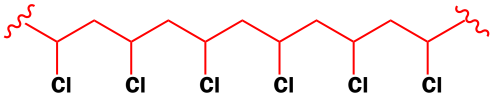


nitrazepam

However, functional groups refuse to be stereotyped. Nitrazepam also contains a nitro group, but this compound is marketed as Mogadon®, the sleeping pill.

Alkyl halides (fluorides R–F, chlorides R–Cl, bromides R–Br, or iodides R–I) contain the fluoro, chloro, bromo, or iodo groups

These three functional groups have similar properties—though alkyl iodides are the most reactive and alkyl fluorides the least. PVC (polyvinyl chloride) is one of the most widely used polymers—it has a chloro group on every other carbon atom along a linear hydrocarbon framework. Methyl iodide (MeI), on the other hand, is a dangerous carcinogen, since it reacts with DNA and can cause mutations in the genetic code.



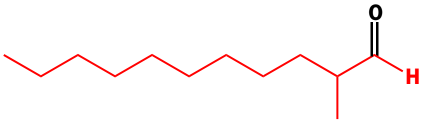
a section of the structure of PVC

■ These compounds are also known as haloalkanes (fluoroalkanes, chloroalkanes, bromoalkanes or iodoalkanes).

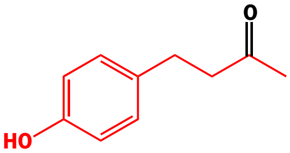
▶ Because alkyl halides have similar properties, chemists use yet another 'wild card' organic element, X, as a convenient substitute for Cl, Br, or I (sometimes F). So R–X is any alkyl halide.

Aldehydes (R–CHO) and ketones (R^1–CO–R^2) contain the carbonyl group C=O

Aldehydes can be formed by oxidizing alcohols—in fact the liver detoxifies ethanol in the bloodstream by oxidizing it first to acetaldehyde (ethanal, CH_3CHO). Acetaldehyde in the blood is the cause of hangovers. Aldehydes often have pleasant smells—2-methylundecanal is a key component of the fragrance of Chanel No 5™, and 'raspberry ketone' is the major component of the flavour and smell of raspberries.



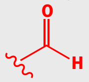
2-methylundecanal



"raspberry ketone"

■

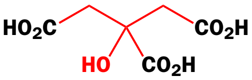
–CHO represents:



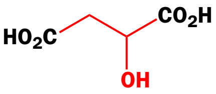
When we write aldehydes as R–CHO, we have no choice but to write in the C and H (because they're part of the functional group)—one important instance where you should ignore Guideline 3 for drawing structures. Another point: always write R–CHO and never R–COH, which looks too much like an alcohol.

Carboxylic acids (R–CO_2H) contain the carboxyl group CO_2H

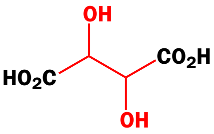
As their name implies, compounds containing the carboxylic acid (CO_2H) group can react with bases, losing a proton to form carboxylate salts. Edible carboxylic acids have sharp flavours and several are found in fruits—citric, malic, and tartaric acids are found in lemons, apples, and grapes, respectively.



citric acid



malic acid

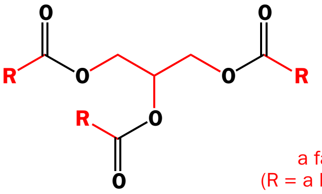


tartaric acid

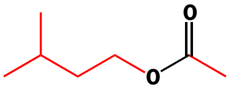
Esters (R^1–CO_2R^2) contain a carboxyl group with an extra alkyl group (CO_2R)

Fats are esters; in fact they contain three ester groups. They are formed in the body by condensing glycerol, a compound with three hydroxyl groups, with three fatty acid molecules.

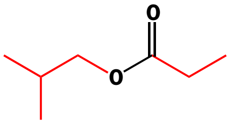
Other, more volatile esters, have pleasant, fruity smells and flavours. These three are components of the flavours of bananas, rum, and apples:



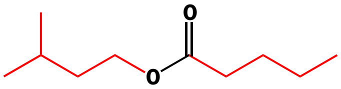
a fat molecule
(R = a long alkyl chain)



isopentyl acetate
(bananas)



isobutyl propionate
(rum)



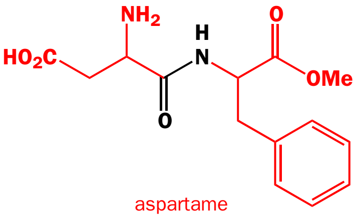
isopentyl valerate
(apples)

■

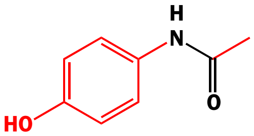
The terms 'saturated fats' and 'unsaturated fats' are familiar—they refer to whether the R groups are saturated (no C=C double bonds) or unsaturated (contains C=C double bonds)—see the box on p. 000. Fats containing R groups with several double bonds (for example, those that are esters formed from linoleic acid, which we met at the beginning of this chapter) are known as 'polyunsaturated'.

Amides (R–$CONH_2$, R^1–$CONHR^2$, or $R^1CONR^2R^3$)

Proteins are amides: they are formed when the carboxylic acid group of one amino acid condenses with the amino group of another to form an amide linkage (also known as a peptide bond). One protein molecule can contain hundreds of amide bonds. Aspartame, the artificial sweetener marketed as NutraSweet®, on the other hand contains just two amino acids, aspartic acid and phenylalanine, joined through one amide bond. Paracetamol is also an amide.



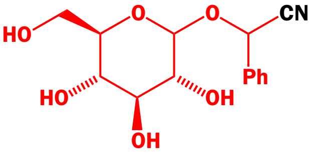
aspartame



paracetamol

Nitriles or cyanides (R–CN) contain the cyano group –C≡N

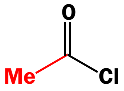
Nitrile groups can be introduced into molecules by reacting potassium cyanide with alkyl halides. The organic nitrile group has quite different properties associated with lethal inorganic cyanide: Laetrile, for example, is extracted from apricot kernels, and was once developed as an anticancer drug. It was later proposed that the name be spelt 'liar-trial' since the results of the clinical trials on laetrile turned out to have been falsified!



laetrile

Acyl chlorides (acid chlorides) (R–COCl)

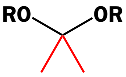
Acyl chlorides are reactive compounds used to make esters and amides. They are derivatives of carboxylic acids with the –OH replaced by –Cl, and are too reactive to be found in nature.



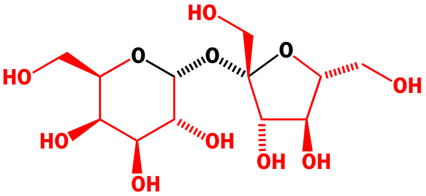
acetyl chloride

Acetals

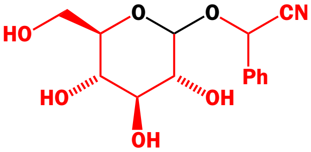
Acetals are compounds with two single bonded oxygen atoms attached to the same carbon atom. Many sugars are acetals, as is laetrile which you have just met.



an acetal



sucrose



laetrile

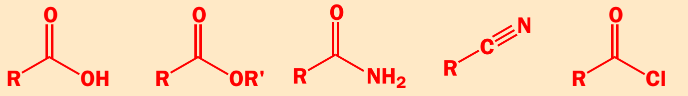
Carbon atoms carrying functional groups can be classified by oxidation level

All functional groups are different, but some are more different than others. For example, the structures of a carboxylic acid, an ester, and an amide are all very similar: in each case the carbon atom carrying the functional group is bonded to two **heteroatoms**, one of the bonds being a double bond. You will see in Chapter 12 that this similarity in structure is mirrored in the reactions of these three types of compounds, and in the ways in which they can be interconverted. Carboxylic acids, esters, and amides can be changed one into another by reaction with simple reagents such as water, alcohols, or amines plus appropriate catalysts. To change them into aldehydes or alcohols requires a different type or reagent, a reducing agent (a reagent which adds hydrogen atoms). We say that the carbon atoms carrying functional groups that can be interconverted without the need for reducing agents (or oxidizing agents) have the same oxidation level—in this case, we call it the 'carboxylic acid oxidation level'.

▶ **A heteroatom is an atom that is not C or H**

You've seen that a functional group is essentially any deviation from an alkane structure, either because the molecule has fewer hydrogen atoms than an alkane (alkenes, alkynes) or because it contains a collection of atoms that are not C and not H. There is a useful term for these 'different' atoms: heteroatoms. A **heteroatom** is any atom in an organic molecule other than C or H.

● **The carboxylic acid oxidation level**

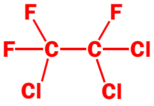


carboxylic acids esters amides nitriles acyl chlorides

In fact, amides can quite easily be converted into nitriles just by dehydration (removal of water), so we must give nitrile carbon atoms the same oxidation level as carboxylic acids, esters, and amides. Maybe you're beginning to see the structural similarity between these four functional groups that you could have used to assign their oxidation level? In all four cases, the carbon atom has *three* bonds to heteroatoms, and only one to C or H. It doesn't matter how many heteroatoms there are, just how many bonds to them. Having noticed this, we can also assign both carbon atoms in 'CFC-113', one of the environmentally unfriendly aerosol propellants/refrigerants that have caused damage to the earth's ozone layer, to the carboxylic acid oxidation level.

▶

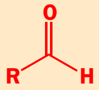
Don't confuse oxidation **level** with oxidation **state**. In all of these compounds, carbon is in oxidation state +4.

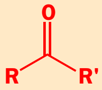


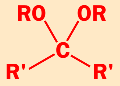
"CFC-113"

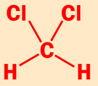
Aldehydes and ketones contain a carbon atom with *two* bonds to heteroatoms; they are at the 'aldehyde oxidation level'. The common laboratory solvent dichloromethane also has two bonds to heteroatoms, so it too contains a carbon atom at the aldehyde oxidation level, as do acetals.

The aldehyde oxidation level


aldehydes


ketones


acetals


dichloromethane

Alcohols, ethers, and alkyl halides have a carbon atom with only *one* single bond to a heteroatom. We assign these the 'alcohol oxidation level', and they are all easily made from alcohols without oxidation or reduction.

The alcohol oxidation level

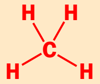

alcohols


ethers


alkyl halides

Lastly, we must include simple alkanes, which have no bonds to heteroatoms, as an 'alkane oxidation level'.

The alkane oxidation level

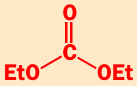

methane

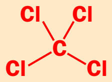
The small class of compounds that have a carbon atom with four bonds to heteroatoms is related to CO_2 and best described as at the carbon dioxide oxidation level.

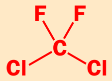
The carbon dioxide oxidation level

$O{=}C{=}O$

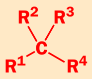
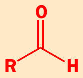
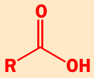
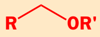
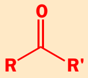
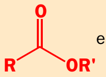
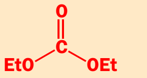
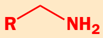
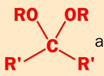
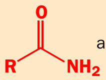
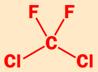
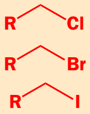
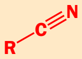
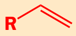
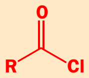
carbon dioxide


diethyl carbonate
useful reagent for
adding ester groups


carbon tetrachloride
formerly used as a
dry cleaning fluid


"CFC-12"
one of the refrigerants/
aerosol propellants which
has caused damage to the
earth's ozone layer

Summary: Important functional groups and oxidation levels

Zero bonds to heteroatoms Alkane oxidation level	**One bond to heteroatom Alcohol oxidation level**	**Two bonds to heteroatoms Aldehyde oxidation level**	**Three bonds to heteroatoms Carboxylic acid oxidation level**	**Four bonds to heteroatoms Carbon dioxide oxidation level**
 alkanes	 alcohols	 aldehydes	 carboxylic acids	$O{=}C{=}O$ carbon dioxide
	 ethers	 ketones	 esters	 diethyl carbonate
	 amines	 acetals	 amides	 CFC-12
	 alkyl halides		 nitriles	
	 alkenes	 alkynes	 acyl chlorides	

Alkenes and alkynes obviously don't fit easily into these categories as they have no bonds to heteroatoms. Alkenes can be made from alcohols by dehydration without any oxidation or reduction so it seems sensible to put them in the alcohol column. Similarly, alkynes and aldehydes are related by hydration/dehydration without oxidation or reduction.

Naming compounds

So far, we have talked a lot about compounds by name. Many of the names we've used (palytoxin, muscone, brevetoxin…) are simple names given to complicated molecules without regard for the actual structure or function of the molecule—these three names, for example, are all derived from the name of the organism from which the compound was first extracted. They are known as **trivial names**, not because they are unimportant, but because they are used in everyday scientific conversation.

Names like this are fine for familiar compounds that are widely used and referred to by chemists, biologists, doctors, nurses, perfumers alike. But there are over 16 million known organic compounds. They can't all have simple names, and no one would remember them if they did. For this reason, the IUPAC (International Union of Pure and Applied Chemistry) have developed **systematic nomenclature**, a set of rules that allows any compound to be given a unique name that can be deduced directly from its chemical structure. Conversely, a chemical structure can be deduced from its systematic name.

The problem with systematic names is that they tend to be grotesquely unpronounceable for anything but the most simple molecules. In everyday speech and writing, chemists therefore do tend to disregard them, and use a mixture of systematic and trivial names. Nonetheless, it's important to know how the rules work. We shall look next at systematic nomenclature, before going on to look at the real language of chemistry.

Systematic nomenclature

There isn't space here to explain all the rules for giving systematic names for compounds—they fill several desperately dull volumes, and there's no point knowing them anyway since computers will do the naming for you. What we will do is to explain the principles underlying systematic nomenclature. You should understand these principles, because they provide the basis for the names used by chemists for the vast majority of compounds that do not have their own trivial names.

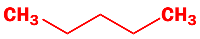
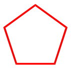
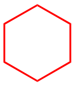
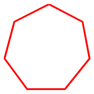
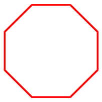
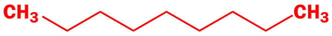
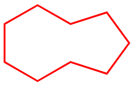
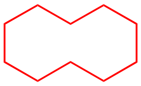
Systematic names can be divided into three parts: one describes the hydrocarbon framework; one describes the functional groups; and one indicates where the functional groups are attached to the skeleton.

You have already met the names for some simple fragments of hydrocarbon framework (methyl, ethyl, propyl…). Adding a hydrogen atom to these alkyl fragments and changing -yl to -ane makes the alkanes and their names. You should hardly need reminding of their structures:

Names for the hydrocarbon framework

one carbon	methane	CH_4		
two carbons	ethane	$CH_3—CH_3$		
three carbons	propane		cyclopropane	

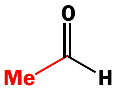
Names for the hydrocarbon framework *(continued)*

four carbons	butane		cyclobutane	
five carbons	pentane		cyclopentane	
six carbons	hexane		cyclohexane	
seven carbons	heptane		cycloheptane	
eight carbons	octane		cyclo-octane	
nine carbons	nonane		cyclononane	
ten carbons	decane		cyclodecane	

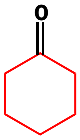
The name of a functional group can be added to the name of a hydrocarbon framework either as a suffix or as a prefix. Some examples follow. It is important to count all of the carbon atoms in the chain, even if one of them is part of a functional group: so pentanenitrile is actually BuCN.

CH_3OH

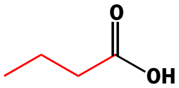
methanol



ethanal

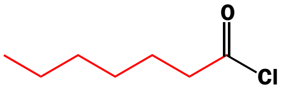


cyclohexanone



butanoic acid

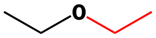
pentanenitrile



heptanoyl chloride

$HC \equiv CH$

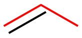
ethyne



ethoxyethane

CH_3NO_2

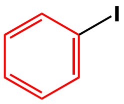
nitromethane



propene

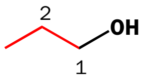
Compounds with functional groups attached to a benzene ring are named in a similar way.

iodobenzene

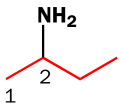


Numbers are used to locate functional groups

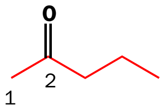
Sometimes a number can be included in the name to indicate which carbon atom the functional group is attached to. None of the above list needed a number—check that you can see why not for each one. When numbers are used, the carbon atoms are counted from one end. In most cases, either of two numbers could be used (depending on which end you count from); the one chosen is always the lower of the two. Again, some examples will illustrate this point. Notice again that some functional groups are named by prefixes, some by suffixes, and that the number always goes directly before the functional group name.



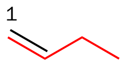
propan-1-ol



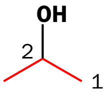
2-aminobutane



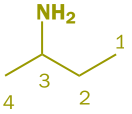
pentan-2-one



but-1-ene

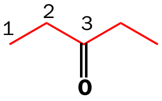


propan-2-ol

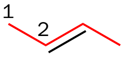


NOT CORRECT

(not 3-aminobutane)

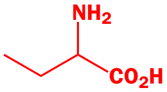


pentan-3-one



but-2-ene

Here are some examples of compounds with more than one functional group.



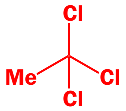
2-aminobutanoic acid

1,6-diaminohexane

hexanedioic acid

CBr_4

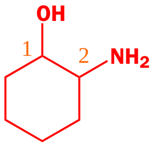
tetrabromomethane



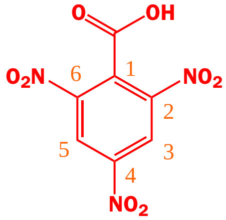
1,1,1-trichloroethane

Again, the numbers indicate how far the functional groups are from the end of the carbon chain. Counting must always be from the same end for each functional group. Notice how we use di-, tri-, tetra- if there are more than one of the same functional group.

With cyclic compounds, there isn't an end to the chain, but we can use numbers to show the distance between the two groups—start from the carbon atom carrying one of the functional groups, then count round.

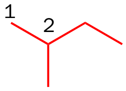


2-aminocyclohexanol

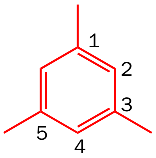


2,4,6-trinitrobenzoic acid

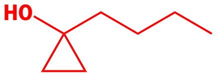
These rules work for hydrocarbon frameworks that are chains or rings, but many skeletons are branched. We can name these by treating the branch as though it were a functional group:



2-methylbutane



1,3,5-trimethyl benzene

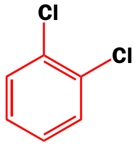


1-butylcyclopropanol

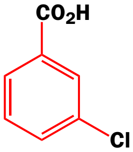
Ortho, *meta*, and *para*

With substituted benzene rings, an alternative way of identifying the positions of the substituents is to use the terms *ortho*, *meta*, and *para*. *Ortho* compounds are 1,2-disubstituted, *meta* compounds are 1,3-disubstituted, and *para* compounds are 1,4-disubstituted. Some examples should make this clear.

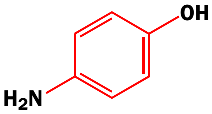
▶ *ortho*, *meta*, and *para* are often abbreviated to *o*, *m*, and *p*.



1,2-dichlorobenzene
or *ortho*-dichlorobenzene
or *o*-dichlorobenzene



3-chlorobenzoic acid
or *meta*-chlorobenzoic acid
or *m*-chlorobenzoic acid



4-aminophenol
or *para*-aminophenol
or *p*-aminophenol

▶ Beware! *Ortho*, *meta*, and *para* are used in chemistry to mean other things too: you may come across orthophosphoric acid, metastable states, and paraformaldehyde—these have nothing to do with the substitution patterns of benzene rings.

The terms *ortho*, *meta*, and *para* are used by chemists because they're easier to remember than numbers, and the words carry with them chemical meaning. '*Ortho*' shows that two groups are next to each other on the ring even though the atoms may not happen to be numbered 1 and 2. They are one example of the way in which chemists don't always use systematic nomenclature but revert to more convenient 'trivial' terms. We consider trivial names in the next section.

What do chemists really call compounds?

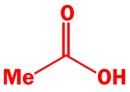
The point of naming a compound is to be able to communicate with other chemists. Most chemists are happiest communicating chemistry by means of structural diagrams, and structural drawings are far more important than any sort of chemical nomenclature. That's why we explained in detail how to draw structures, but only gave an outline of how to name compounds. Good diagrams are easy to understand, quick to draw, and difficult to misinterpret.

> ● **Always give a diagram alongside a name unless it really is something very simple, such as ethanol.**

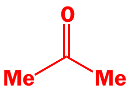
But we do need to be able to communicate by speech and by writing as well. In principle we could do this by using systematic names. In practice, though, the full systematic names of anything but the simplest molecules are far too clumsy for use in everyday chemical speech. There are several alternatives, mostly based on a mixture of trivial and systematic names.

Names for well known and widely used simple compounds

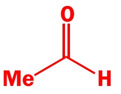
A few simple compounds are called by trivial names not because the systematic names are complicated, but just out of habit. We know them so well that we use their familiar names.



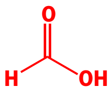
You may have met this compound before (left), and perhaps called it ethanoic acid, its systematic name. But in a chemical laboratory, everyone would refer to this acid as acetic acid, its trivial name. The same is true for all these common substances.



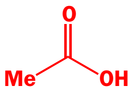
acetone



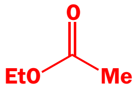
acetaldehyde



formic acid



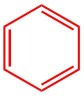
acetic acid



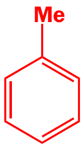
ethyl acetate

Et_2O

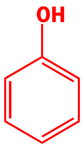
ether,
or diethyl ether



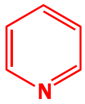
benzene



toluene



phenol



pyridine

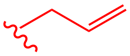
▶ We haven't asked you to remember any trivial names of molecules yet. But these 10 compounds are so important, you must be able to remember them. Learn them now.

Trivial names like this are often long-lasting, well understood historical names that are less easy to confuse than their systematic counterparts. 'Acetaldehyde' is easier to distinguish from 'ethanol' than is 'ethanal'.

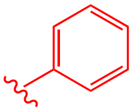
Trivial names also extend to fragments of structures containing functional groups. Acetone, acetaldehyde, and acetic acid all contain the acetyl group (MeCO-, ethanoyl) abbreviated Ac and chemists often use this 'organic element' in writing AcOH for acetic acid or EtOAc for ethyl acetate.

Chemists use special names for four fragments because they have mechanistic as well as structural significance. These are vinyl and allyl; phenyl and benzyl.

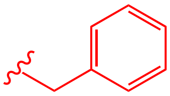
the vinyl group



the allyl group

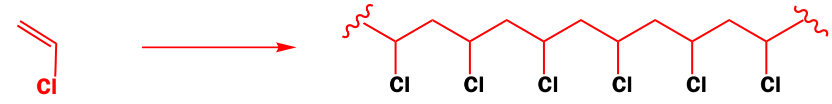


the phenyl group: Ph



the benzyl group: Bn

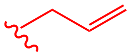
Giving the vinyl group a name allows chemists to use simple trivial names for compounds like vinyl chloride, the material that polymerizes to give PVC (poly vinyl chloride) but the importance of the name lies more in the difference in reactivity (Chapter 17) between vinyl and allyl groups.



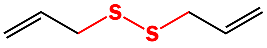
vinyl chloride

a section of the structure of PVC - Poly Vinyl Chloride

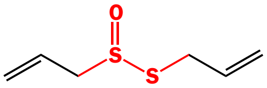
The allyl group gets its name from garlic (*Allium* sp.), because it makes up part of the structure of the compounds responsible for the taste and smell of garlic.



the allyl group



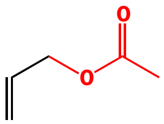
diallyl disulfide



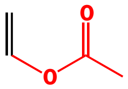
allicin

Allyl and vinyl are different in that the vinyl group is attached directly to a double bonded C=C carbon atom, while the allyl group is attached to a carbon atom *adjacent* to the C=C double bond. The difference is extremely important chemically: allyl compounds are typically quite reactive, while vinyl compounds are fairly unreactive.

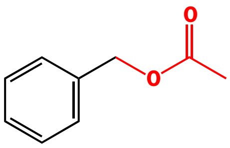
For some reason, the allyl and vinyl groups have never acquired organic element symbols, but the benzyl group has and is called Bn. It is again important not to confuse the benzyl group with the phenyl group: the phenyl group is joined through a carbon atom in the ring, while the benzyl group is joined through a carbon atom attached to the ring. Phenyl compounds are typically unreactive but benzyl compounds are often reactive. Phenyl is like vinyl and benzyl is like allyl.



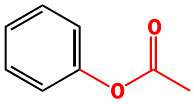
allyl acetate



vinyl acetate



benzyl acetate



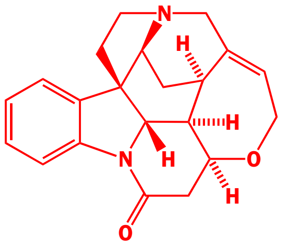
phenyl acetate

We shall review all the organic element element symbols you have met at the end of the chapter.

Names for more complicated but still well known molecules

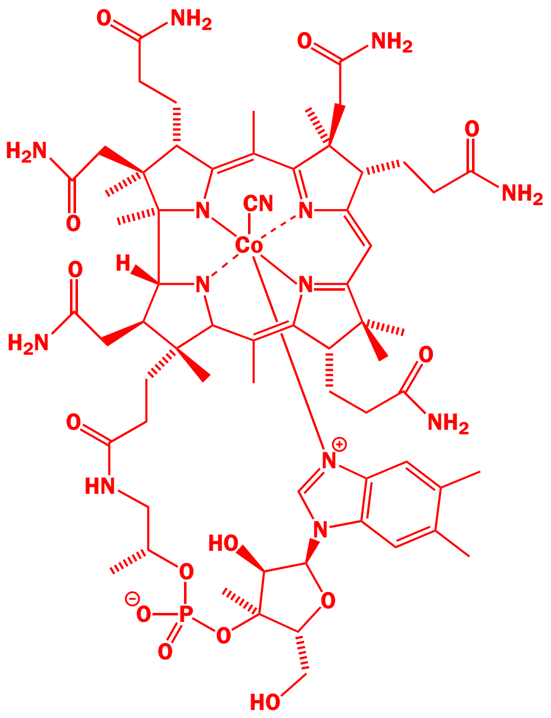
Complicated molecules that have been isolated from natural sources are always given trivial names, because in these cases, the systematic names really are impossible!

Strychnine is a famous poison featured in many detective stories and a molecule with a beautiful structure. All chemists refer to it as strychnine as the systematic name is virtually unpronounceable. Two groups of experts at IUPAC and *Chemical Abstracts* also have different ideas on the systematic name for strychnine. Others like this are penicillin, DNA, and folic acid.



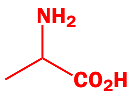
strychnine, or
(1R,11R,18S,20S,21S,22S)-12-oxa-8.17-diazaheptacyclo [15.5.$0^{1,8}$.$0^{2,7}$.$0^{15,20}$] tetracosa-2,4,6,14-tetraene-9-one (IUPAC)
or
4aR-[4aα,5aα,8aR*,15aα,15bα,15cβ]-2,4a,5,5a,7,8,15,15a,15b,15c-decahydro-4,6-methano-6H,14H-indolo[3,2,1-*ij*]oxepino [2,3,4-*de*]pyrrolo[2,3-*h*]quinolone (*Chemical Abstracts*)

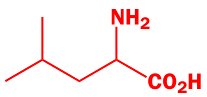
But the champion is vitamin B_{12}, a complicated cobalt complex with a three-dimensional structure of great intricacy. No chemist would learn this structure but would look it up in an advanced textbook of organic chemistry. You will find it in such books in the index under vitamin B_{12} and not under its systematic name. We do not even know what its systematic name might be and we are not very interested. This is vitamin B_{12}.

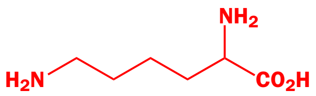


vitamin B_{12}, or....

Even fairly simple but important molecules, the amino acids for example, that have systematic names that are relatively easy to understand are normally referred to by their trivial names which are, with a bit of practice, easy to remember and hard to muddle up. They are given in full in Chapter 49.


alanine, or
2-aminopropanoic acid

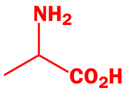

leucine, or
2-amino-4-methylpentanoic acid

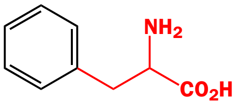

lysine, or
2,6-diaminohexanoic acid

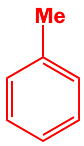
A very flexible way of getting new, simple names for compounds can be to combine a bit of systematic nomenclature with trivial nomenclature.

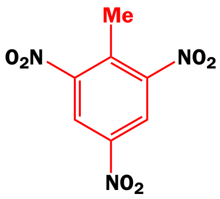
Alanine is a simple amino acid that occurs in proteins. Add a phenyl group and you have phenylalanine a more complex amino acid also in proteins.

Toluene, the common name for methylbenzene, can be combined (both chemically and in making names for compounds!) with three nitro groups to give the famous explosive trinitrotoluene or TNT.


alanine

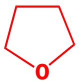

phenylalanine

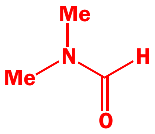

toluene

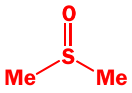

2,4,6-trinitrotoluene

Compounds named as acronyms

Some compounds are referred to by acronyms, shortened versions of either their systematic or their trivial name. We just saw TNT as an abbreviation for TriNitroToluene but the commoner use for acronyms is to define solvents and reagents in use all the time. Later in the book you will meet these solvents.

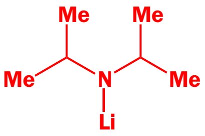

THF
(TetraHydroFuran)

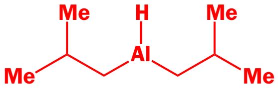

DMF
(DiMethylFormamide)

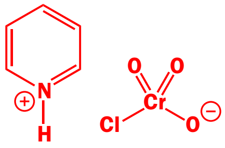

DMSO
(DiMethylSulfOxide)

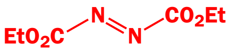
▶
The names and structures of these common solvents need learning too.

The following reagents are usually referred to by acronym and their functions will be introduced in other chapters so you do not need to learn them now. You may notice that some acronyms refer to trivial and some to systematic names. There is a glossary of acronyms for solvents, reagents, and other compounds on p. 000.


LDA
Lithium Di-isopropylAmide

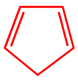

DIBAL
Di-IsoButylALuminiumhydride

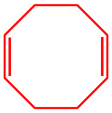

PCC
Pyridinium ChloroChromate

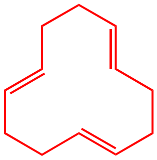

DEAD
DiEthyl Azo-Dicarboxylate

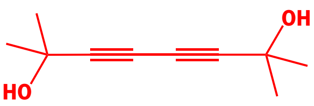
Compounds for which chemists use systematic names

You may be surprised to hear that practising organic chemists use systematic names at all in view of what we have just described, but they do! Systematic names really begin with derivatives of pentane (C_5H_{12}) since the prefix pent- means five, whereas but- does not mean four. Chemists refer to simple derivatives of open chain and cyclic compounds with 5 to about 20 carbon atoms by their systematic names, providing that there is no common name in use. Here are some examples.


cyclopentadiene


cyclo-octa-1,5-diene


cyclododeca-1,5,9-triene


2,7-dimethyl-3,5-octadiyne-2,7-diol

11-bromo-undecanoic acid

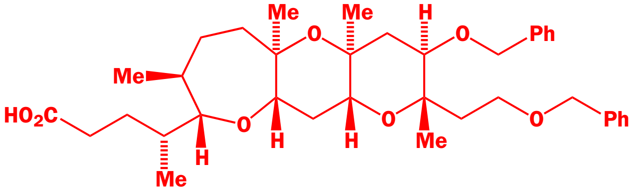
non-2-enal

These names contain a syllable that tells you the framework size: penta- for C_5, octa- for C_8, nona- for C_9, undeca- for C_{11}, and dodeca- for C_{12}. These names are easily worked out from the structures and, what is more important, you get a clear idea of the structure from the name. One of them might make you stop and think a bit (which one?), but the others are clear even when heard without a diagram to look at.

Complicated molecules with no trivial names

When chemists make complex new compounds in the laboratory, they publish them in a chemical journal giving their full systematic names in the experimental account, however long and clumsy those names may be. But in the text of the paper, and while talking in the lab about the compounds they have made, they will just call them 'the amine' or 'the alkene'. Everyone knows which amine or alkene is meant because at some point they remember seeing a chemical structure of the compound. This is the best strategy for talking about almost any molecule: draw a structure, then give the compound a 'tag' name like 'the amine' or 'the acid'. In written chemistry it's often easiest to give every chemical structure a 'tag' number as well.

To illustrate what we mean, let's talk about this compound.



19

This carboxylic acid was made and used as an intermediate when chemists in California made brevetoxin (see p. 000) in 1995. Notice how we can call a complicated molecule 'this acid'—a 'tag' name—because you've seen the structure. It also has a tag number (19), so we can also call it 'compound 19', or 'acid 19', or 'brevetoxin fragment 19'. How much more sensible than trying to work out its systematic name.

How should you name compounds?

So what should you call a compound? It really depends on circumstances, but you won't go far wrong if you follow the example of this book. We shall use the names for compounds that real

● Our advice on chemical names—six points in order of importance

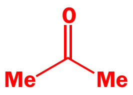
- Draw a structure first and worry about the name afterwards
- Learn the names of the *functional groups* (ester, nitrile, etc.)
- Learn and use the names of a few simple compounds used by all chemists
- In speech, refer to compounds as 'that acid' (or whatever) while pointing to a diagram
- Grasp the principles of systematic (IUPAC) nomenclature and use it for compounds of medium size
- Keep a notebook to record acronyms, trivial names, structures, etc. that you might need later

chemists use. There's no need to learn all the commonly used names for compounds now, but you should log them in your memory as you come across them. Never allow yourself to pass a compound name by unless you are sure you know what chemical structure it refers to. You will find many of the commonly used names for compounds on the endpapers of this book. Refer to these, or to the shorter glossary on p. 000 to refresh your memory should you ever need to.

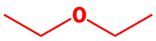
We've met a great many molecules in this chapter. Most of them were just there to illustrate points so don't learn their structures! Instead, learn to recognize the names of the functional groups they contain. However, there were 10 names for simple compounds and three for common solvents that we advised you to learn. Cover up the structures on the rest of this page and draw the structures for these 13 compounds.

Important structures to learn

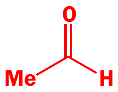
acetone



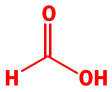
ether or
diethyl ether



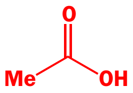
acetaldehyde



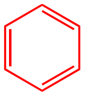
formic acid



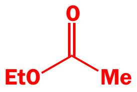
acetic acid
or AcOH



benzene

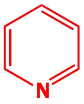


ethyl acetate
or EtOAc

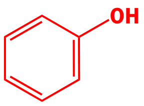


toluene

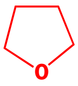
pyridine



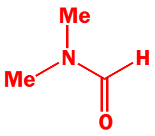
phenol



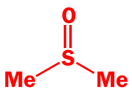
THF
(tetrahydrofuran)



DMF or
(dimethylformamide)
Me$_2$NCHO

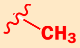
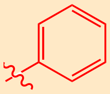
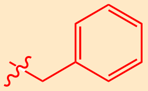
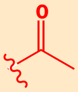
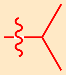
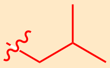
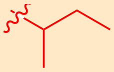


DMSO
(dimethylsulfoxide)



That's all we'll say on the subject of nomenclature—you'll find that as you practise using these names and start hearing other people referring to compounds by name you'll soon pick up the most important ones. But, to reiterate, make sure you never pass a compound name by without being absolutely sure what it refers to—draw a structure to check.

● Review box: Table of fragment names and organic elements

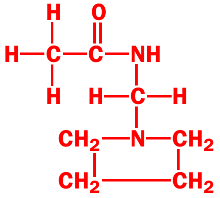
R	alkyl		***t*-Bu**	*tert*-butyl	
Me	methyl		**Ar**	aryl	any aromatic ring
Et	ethyl		**Ph**	phenyl	
Pr (or *n*-Pr)	propyl		**Bn**	benzyl	
Bu (or *n*-Bu)	butyl		**Ac**	acetyl	
***i*-Pr**	isopropyl			vinyl	
***i*-Bu**	isobutyl			allyl	
***s*-Bu**	*sec*-butyl		**X**	halide	**F, Cl, Br, or I**

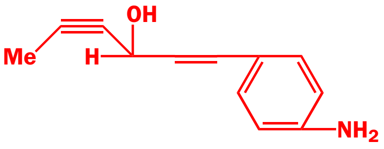
Problems

1. Draw good diagrams of saturated hydrocarbons with seven carbon atoms having (a) linear, (b) branched, and (c) cyclic frameworks. Draw molecules based on each framework having both ketone and carboxylic acid functional groups.

2. Study the structure of brevetoxin on p. 000. Make a list of the different types of functional group (you already know that there are many ethers) and of the numbers of rings of different sizes. Finally study the carbon framework—is it linear, cyclic, or branched?

3. What is wrong with these structures? Suggest better ways of representing these molecules.





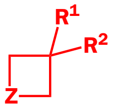
4. Draw structures corresponding to these names. In each case suggest alternative names that might convey the structure more clearly to someone who is listening to you speak.

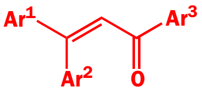
(**a**) 1,4-di-1(1-dimethylethyl)benzene

(**b**) 2-(prop-2-enyloxy)prop-1-ene

(**c**) cyclohexa-1,3,5-triene

5. Draw one possible structure for each of these molecules, selecting any group of your choice for the 'wild card' substituents.





6. Translate these very poor 'diagrams' of molecules into more realistic structures. Try to get the angles about right and, whatever you do, don't include any square coplanar carbon atoms or other bond angles of 90°!

$C_6H_5CH(OH).(CH_2)_4COC_2H_5$

$O(CH_2CH_2)_2O$

$(CH_3O)_2CHCH{=}CHCH(OMe)_2$

7. Suggest at least six different structures that would fit the formula C_4H_7NO. Make good realistic diagrams of each one and say which functional group(s) are present.

8. Draw and name a structure corresponding to each of these descriptions.

(**a**) An aromatic compound containing one benzene ring with the following substituents: two chlorine atoms having a *para* relationship, a nitro group having an *ortho* relationship to one of the chlorine atoms, and an acetyl group having a *meta* relationship to the nitro group.

(**b**) An alkyne having a trifluoromethyl substituent at one end and a chain of three carbon atoms at the other with a hydroxyl group on the first atom, an amino group on the second, and the third being a carboxyl group.

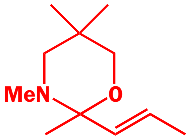
9. Draw full structures for these compounds, displaying the hydrocarbon framework clearly and showing all the bonds present in the functional groups. Name the functional groups.

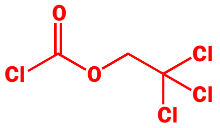
$AcO(CH_2)_3NO_2$

$MeO_2C.CH_2.OCOEt$

$CH_2{=}CH.CO.NH(CH_2)_2CN$

10. Identify the oxidation level of each of the carbon atoms in these structures with some sort of justification.





11. If you have not already done so, complete the exercises on pp. 000 (drawing amino acids) and 000 (giving structures for the 10 common compounds and three common solvents).

Determining organic structures

3

Connections

Building on:

- What sorts of structure organic molecules have **ch2**

Arriving at:

- Determining structure by X-ray crystallography
- Determining structure by mass spectrometry
- Determining structure by ^{13}C NMR spectroscopy
- Determining structure by infrared spectroscopy

Looking forward to:

- ^{1}H NMR spectroscopy **ch11**
- Solving unknown structures spectroscopically **ch15**

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:**
X-rays are scattered	measures these interactions
Radio waves make nuclei resonate	plots charts of absorption
Infrared waves are absorbed	relates interactions with structure

X-rays give bond lengths and angles. **Nuclear magnetic resonance** tells us about the carbon skeleton of the molecule. **Infrared spectroscopy** tells us about the types of bond in a 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 molecule and atomic composition)
- Nuclear magnetic resonance (NMR) spectroscopy (to determine carbon skeleton of molecule)
- Infrared spectroscopy (to determine functional groups in molecule)

Of these, NMR is more important than all the rest put together and so we shall return to it in Chapter 11. Then in Chapter 15, 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. If you would like more details of any of the spectroscopic methods we discuss, you should refer to a specialized book.

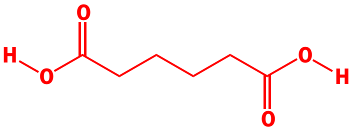
X-ray is the final appeal

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.

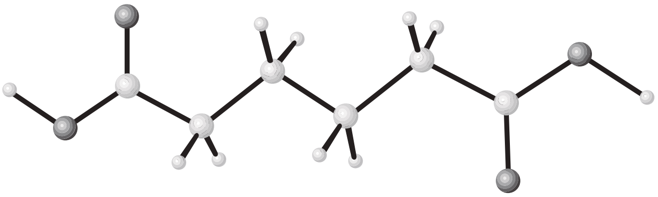
▶ X-ray crystal structures are determined by allowing a sample of a crystalline compound to diffract X-rays. From the resulting diffraction pattern, it is possible to deduce the precise spatial arrangement of the atoms in the molecule—except, usually, the hydrogen atoms, which are too light to diffract the X-rays and whose position must be inferred from the rest of the structure.

$HO_2C{-}(CH_2)_4{-}CO_2H$

hexanedioic acid



shape of hexanedioic acid

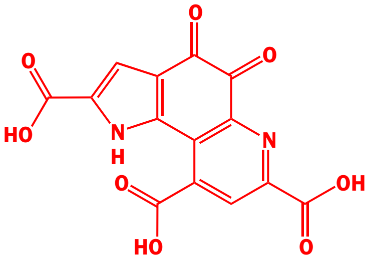


data for structure taken from Cambridge Crystallographic Data Centre

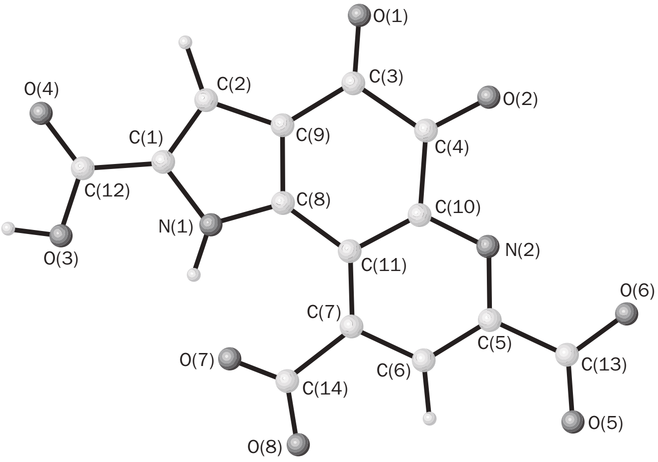
This is one question that X-ray answers better than any other method: what shape does a molecule have? Another important problem it can solve is the structure of a new unknown compound. There are bacteria in oil wells, for example, that use methane as an energy source. It is amazing that bacteria manage to convert methane into anything useful, and, of course, chemists really wanted to know how they did it. Then 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. This is a more complex molecule than hexanedioic acid but X-ray crystallographers routinely solve much more complex structures than this.

▶ **Coenzymes** are small molecules that work hand-in-hand with enzymes to catalyse a biochemical reaction.

■ If you like systematic names, you can call methoxatin 4,5-dihydro-4,5-dioxo-1*H*-pyrrolo[2,3-*f*]quinoline-2,7,9-tricarboxylic acid. But you may feel, like us, that 'methoxatin' and a diagram or the tag name 'the tricarboxylic acid' are better.



methoxatin



data for the X-ray structure taken from the Cambridge Crystallographic Data Centre

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.

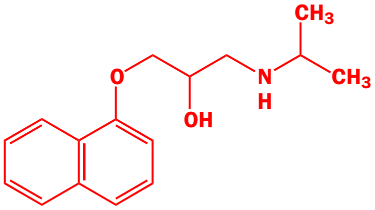
- X-ray crystallography is a science in its own right, a separate discipline from chemistry because it requires specific skills, and a structure determination can take a long time. Modern methods have reduced this time to a matter of hours or less, but nonetheless by contrast a modern NMR machine with a robot attachment can run more than 100 spectra in an overnight run. So we normally use NMR routinely and reserve X-rays for difficult unknown structures and for determining the detailed shape of important molecules.

Outline of structure determination by spectroscopy

Put yourself in these situations.

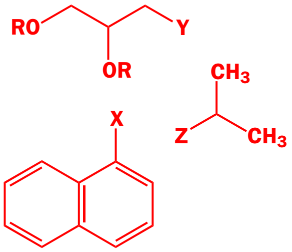
- Finding an unknown product from a chemical reaction
- Discovering an unknown compound from Nature
- Detecting a suspected food contaminant
- Routinely checking purity during the manufacture of a drug

In all cases except perhaps the second you need a quick and reliable answer. Suppose you are trying to identify the heart drug propranolol, one of the famous 'beta blockers' used to reduce high blood pressure and prevent heart attacks. You would first want to know the molecular weight and atomic composition and this 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.



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

fragments of propranolol from the NMR spectrum



There are many ways in which these fragments 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 functional groups such as CN or NO_2. This still leaves a variety of possible structures, and these could finally be distinguished by another technique, ^1H NMR. We are in fact going to avoid using ^1H NMR in this chapter, because it is more difficult, but you will learn just how much information can be gained from mass spectra, IR spectra, and ^{13}C NMR spectra.

^1H NMR makes an entrance in Chapter 11.

Now we must go through each of these methods and see how they give the information they do. For this exercise, we will use some compounds you encounter in everyday life, perhaps without realizing it.

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 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

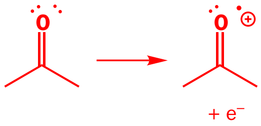
Mass spectrometry

Mass spectrometry weighs the molecule

A mass spectrometer 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 usually use positive ions. Two methods are used to convert neutral molecules into cations: electron impact and chemical ionization.

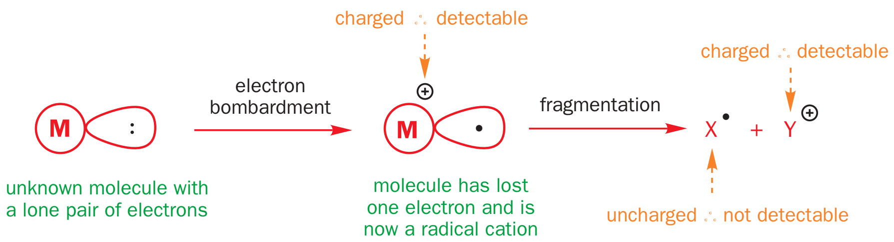
▶ Mass spectrometry uses a different principle from the other forms of spectroscopy we discuss: what is measured is not absorption of energy but the mass of the molecule or fragments of molecule.

Mass spectrometry by electron impact



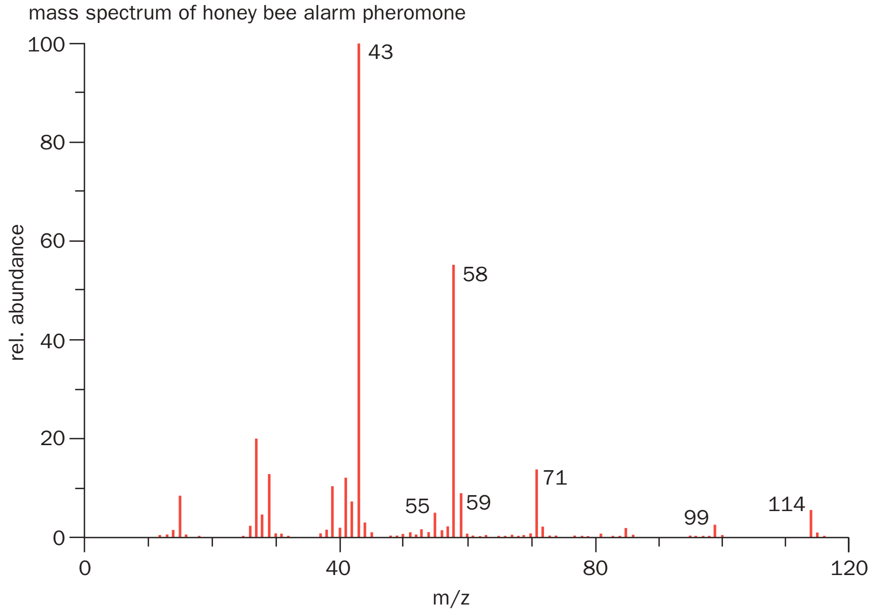
loss of one electron leaves a radical cation

In **electron impact (E.I.) 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 do not stick to the wall but knock loose bricks off the top of the wall. Losing a single electron leaves behind a radical cation: 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 this will typically be one not involved in bonding, for example, an electron from a lone pair. Thus ammonia gives $NH_3^{+\bullet}$ and a ketone gives $R_2C{=}O^{+\bullet}$. If the electron beam is not too high in energy, some of these rather unstable radical cations will survive the focusing operation and get to the detector. Normally two focusing operations are used: the beam is bent magnetically and electrostatically to accelerate the cations on their way to the detector and it takes about 20 μs for the cations to get there. But if, as is often the case, the electron beam supplies more than exactly the right amount of energy to knock out the electron, the excess energy is dissipated by fragmentation of the radical cation. Schematically, an unknown molecule first forms the radical cation $M^{+\bullet}$ which then breaks up (fragments) to give a radical $X^{\bullet}$ and a cation Y^{+}. Only charged particles (cations in most machines) can be accelerated and focused by the magnetic and electrostatic fields and so the detector records only the molecular ion $M^{+\bullet}$ and positively charged fragments Y^{+}. Uncharged radicals $X^{\bullet}$ are not recorded.



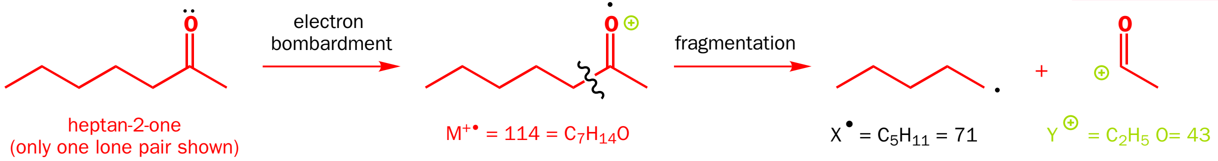
A typical result is the E.I. mass spectrum for the alarm pheromone of the honey bee. The bees check every insect coming into the hive for strangers. If a strange insect (even a bee from another hive) is detected, an alarm pheromone is released and the intruder is attacked. The pheromone is a simple volatile organic molecule having this mass spectrum.

■ Insects communicate by releasing compounds with strong smells (to the insect!). These have to be small volatile molecules, and those used to communicate between members of the same species are called **pheromones**.



The strongest peak, at 43 mass units in this case, is assigned an 'abundance' of 100% and called the **base peak**. The abundance of the other peaks is shown relative to the base peak. In this spectrum, there is only one other strong peak (58 at 50%) and the peak of highest mass at 114 (at 5%) is the molecular ion corresponding to a structure $C_7H_{14}O$. The main fragmentation is to a $C_5H_{11}^{\bullet}$ radical (not observed as it isn't charged) and a cation $C_2H_3O^+$, which forms the base peak. The pheromone is the simple ketone heptan-2-one.

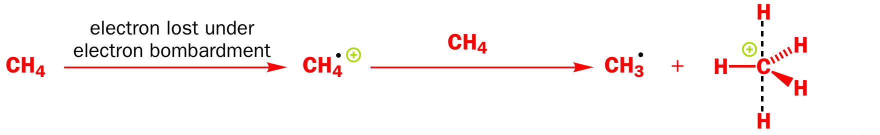
▶ Mass spectroscopy requires minute quantities of sample—much less than the amounts needed for the other techniques we will cover. Pheromones are obtainable from insects only on a microgram scale or less.



The problem with E.I. is that for many radical cations even 20 µs is too long, and all the molecular ions have decomposed by the time they reach the detector. The fragments produced may be useful in identifying the molecule, but even in the case of the bee alarm pheromone it would obviously be better to get a stronger and more convincing molecular ion as the weak (5%) peak at 114 might also be a fragment or even an impurity.

Mass spectrometry by chemical ionization

In **chemical ionization** (C.I.) **mass spectrometry** the electron beam is used to ionize a simple molecule such as methane which in turn ionizes our molecule by collision and transfer of a proton. Under electron bombardment, methane loses a bonding electron (it doesn't have any other kind) to give $CH_4^{+\bullet}$ which reacts with an unionized methane molecule to give $CH_3^{\bullet}$ and CH_5^{+}. Before you write in complaining about a mistake, just consider that last structure in a bit more detail. Yes, CH_5^{+} does have a carbon atom with *five* bonds. But it has only eight electrons! These are distributed between five bonds (hence the + charge) and the structure is thought to be trigonal bipyramidal. This structure has not been *determined* as it is too unstable. It is merely proposed from theoretical calculations.

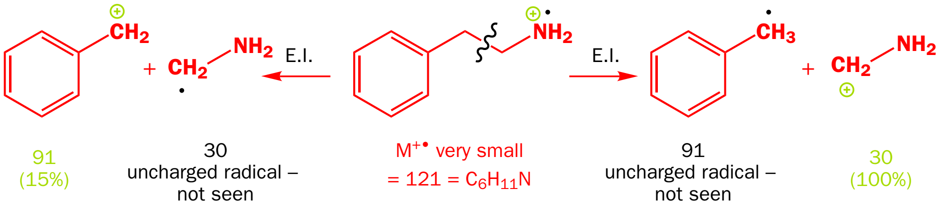


proposed structure of $CH_5^{\oplus}$
the two black bonds share two electrons

This unstable compound is a powerful acid, and can protonate just about any other molecule. When it protonates our sample, a proton has been added rather than an electron removed, so the resulting particles are simple cations, not radical cations, and are generally more stable than the radical cations produced by direct electron impact. So the molecular ion has a better chance of lasting the necessary 20 μs to reach the detector. Note that we now observe $[M + H]^+$ (i.e. one more than the molecular mass) rather than M^+ by this method.

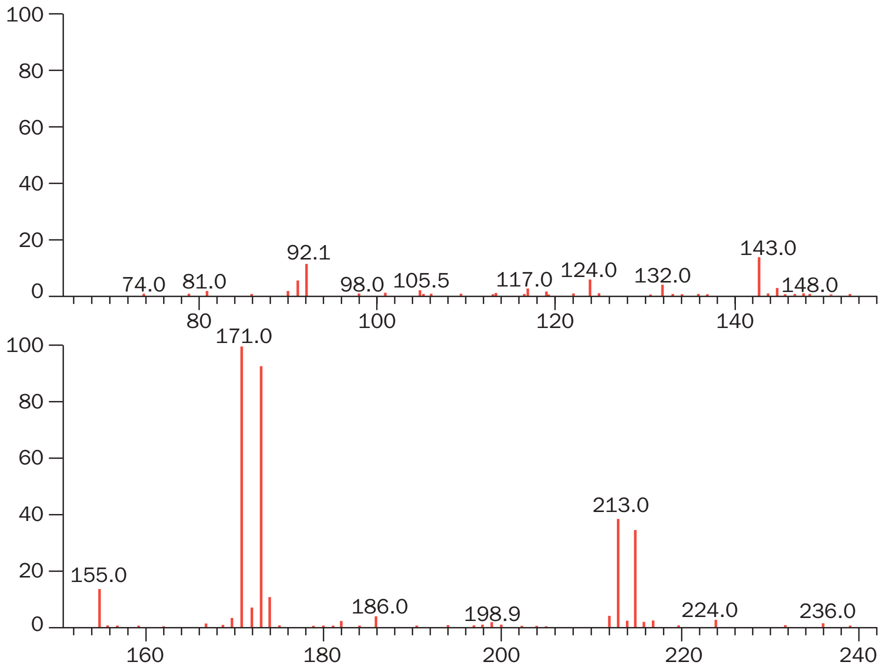
Having more functional groups helps molecular ions to decompose. The aromatic amine 2-phenylethylamine is a brain active amine found in some foods such as chocolate, red wine, and cheese and possibly implicated in migraine. It gives a poor molecular ion by E.I., a base peak with a mass as low as 30 and the only peak at higher mass is a 15% peak at 91. The C.I. mass spectrum on the other hand has a good molecular ion: it is $[M + H]^+$ of course. Normally a fragmentation gives one cation and another radical, only the cation being detected. It is relatively unusual for one bond to be able to fragment in either direction, but here it does, which means that both fragments are seen in the spectrum.

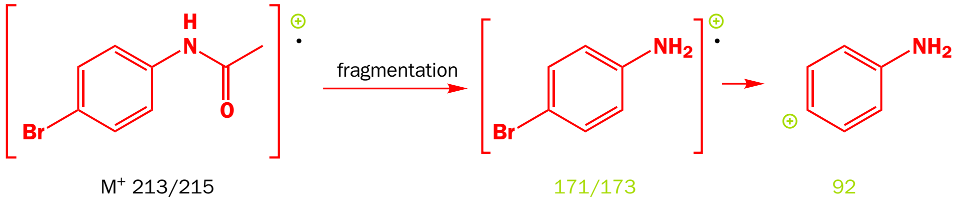
the radical cation can fragment in two ways



Mass spectrometry separates isotopes

You will know in theory that most elements naturally exist as mixtures of isotopes. If you didn't believe it, now you will. Chlorine is normally a 3:1 mixture of ^{35}Cl and ^{37}Cl (hence the obviously false 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!). Mass spectrometry separates these isotopes so that you get true not average molecular weights. The molecular ion in the E.I. mass spectrum of the bromo-amide below has two peaks at 213 and 215 of roughly equal intensity. This might just represent the loss of molecular hydrogen from a molecular ion 215, but, when we notice that the first fragment (and base peak) has the same pattern at 171/173, the presence of bromine is a more likely explanation. All the smaller fragments at 155, 92, etc. lack the 1:1 pattern and also therefore lack bromine.





The mass spectrum of chlorobenzene (PhCl, C_6H_5Cl) is very simple. There are two peaks at 112 (100%) and 114 (33%), a peak at 77 (40%), and very little else. The peaks at 112/114 with their 3:1 ratio are the molecular ions, while the fragment at 77 is the phenyl cation (Ph^+ or $C_6H_5^+$).

▶ It's worth remembering that the Ph^+ weighs 77: you'll see this mass frequently.

The mass spectrum of DDT is very revealing. This very effective insecticide became notorious as it accumulated in the fat of birds of prey (and humans) and was phased out of use. It can be detected easily by mass spectrometry because the five chlorine atoms produce a complex molecular ion at 252/254/256/258/260 with ratios of 243:405:270:90:15:1 (the last is too small to see). The peak at 252 contains nothing but ^{35}Cl, the peak at 254 has four atoms of ^{35}Cl and one atom of ^{37}Cl, while the invisible peak at 260 has five ^{37}Cl atoms. The ratios need some working out, but the first fragment at 235/237/239 in a ratio 9:6:1 is easier. It shows just two chlorine atoms as the CCl_3 group has been lost as a radical.

Table 3.1 Summary table of main isotopes for mass spectra

Element	Carbon	Chlorine	Bromine
isotopes	^{12}C, ^{13}C	^{35}Cl, ^{37}Cl	^{79}Br, ^{81}Br
rough ratio	1.1% ^{13}C (90:1)	3:1	1:1

▶ Remember: mass spectroscopy is very good at detecting minute quantities.

Isotopes in DDT

The ratio comes from the 3:1 isotopic ratio like this:

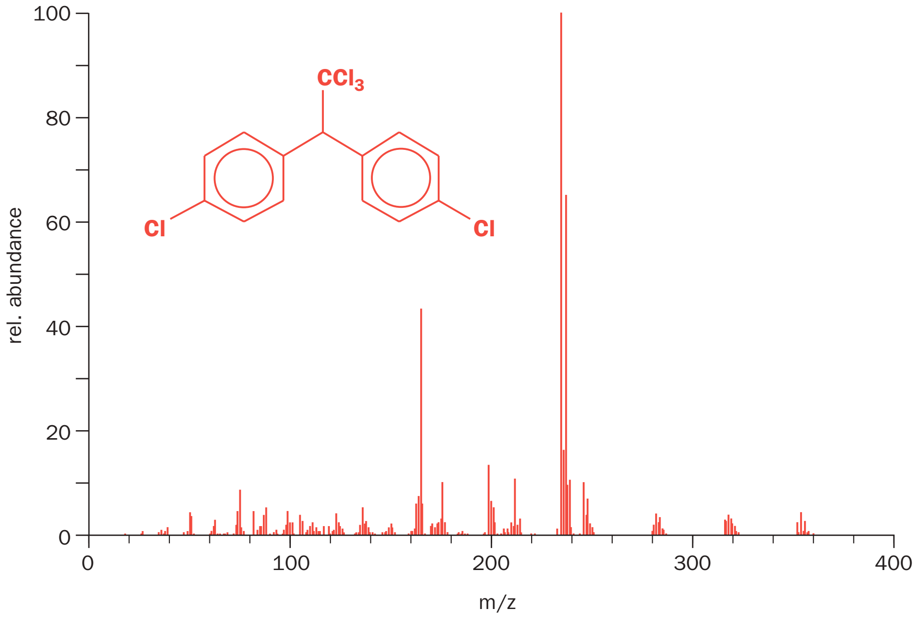
- chance of one ^{35}Cl in the molecule: $\frac{3}{4}$
- chance of one ^{37}Cl in the molecule: $\frac{1}{4}$

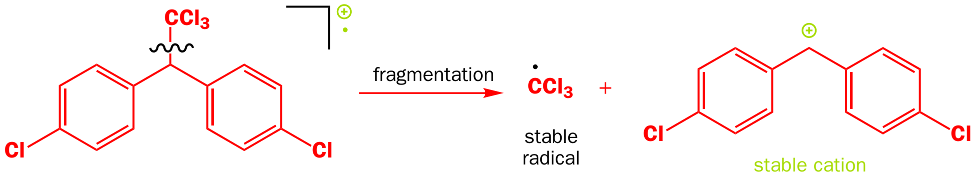
If the molecule or fragment contains two chlorine atoms, as does our $C_{13}H_9Cl_2$, then

- chance of two ^{35}Cls in the molecule: $\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$
- chance of one ^{35}Cl and one ^{37}Cl in the molecule: $\left[\frac{3}{4} \times \frac{1}{4}\right] + \left[\frac{1}{4} \times \frac{3}{4}\right] = \frac{6}{16}$
- chance of two ^{37}Cls in the molecule: $\frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$

The ratio of these three fractions is 9:6:1, the ratio of the peaks in the mass spectrum.

mass spectrum of DDT





■ $\rceil^{\oplus\bullet}$ denotes the cation radical produced by E.I.

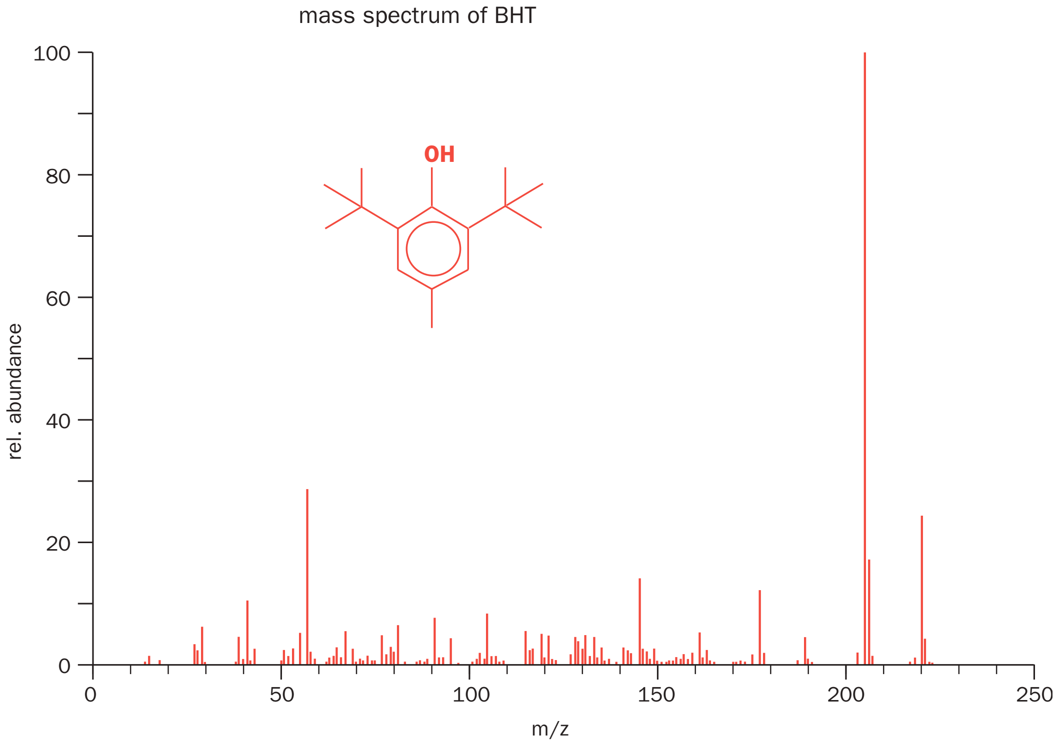
Carbon has a minor but important isotope ^{13}C

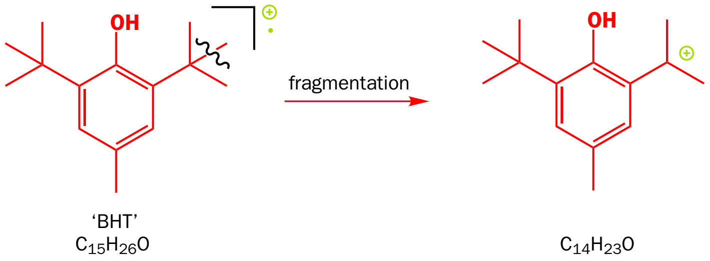
Many elements have minor isotopes at below the 1% level and we can ignore these. One important one we cannot ignore is the 1.1% of ^{13}C present in ordinary carbon. The main isotope is ^{12}C and you may recall that ^{14}C is radioactive and used in carbon dating, but its natural abundance is minute. The stable isotope ^{13}C is not radioactive, but it is NMR active as we shall soon see. If you look back at the mass spectra illustrated so far in this chapter, you will see a small peak one mass unit higher than each peak in most of the spectra. This is no instrumental aberration: these are genuine peaks 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. If there are *n* carbon atoms in a molecular ion, then the ratio of M^+ to $[M + 1]^+$ is 100: (1.1) × *n*.

The electron impact mass spectrum of BHT gives a good example. The molecular ion at 220 has an abundance of 34% and $[M + 1]^+$ at 221 has 5–6% abundance but is difficult to measure as it is so weak. BHT is $C_{15}H_{26}O$ so this should give an $[M + 1]^+$ peak due to ^{13}C of 15 × 1.1% of M^+, that is, 16.5% of M^+ or $34 \times 16.5 = 5.6\%$ actual abundance. An easier peak to interpret is the base peak at 205 formed by the loss of one of the six identical methyl groups from the *t*-butyl side chains (don't forget what we told you in Chapter 2—all the 'sticks' in these structures are methyl groups and not hydrogen atoms). The base peak (100%) 205 is $[M—Me]^+$ and the ^{13}C peak 206 is 15%, which fits well with $14 \times 1.1\% = 15.4\%$.

BHT

BHT is used to prevent the oxidation of vitamins A and E in foods. It carries the E-number E321. There has been some controversy over its use because it is a cancer suspect agent, but it is used in some 'foods' like chewing gum. BHT stands for 'Butylated HydroxyToluene', but you can call it 2,6-di-*t*-butyl-4-methylphenol if you want to, but you may prefer to look at the structure and just call it BHT. You met BHT briefly in Chapter 2 when you were introduced to the tertiary butyl group.





Other examples you have seen include the DDT spectrum, where the peaks between the main peaks are ^{13}C peaks: thus 236, 238, and 240 are each 14% of the peak one mass unit less, as this fragment has 13 carbon atoms. If the number of carbons gets very large, so does the ^{13}C peak; eventually it is *more* likely that the molecule contains one ^{13}C than that it doesn't. We can ignore the possibility of two ^{13}C atoms as 1.1% of 1.1% is very small (probability of 1.32×10^{-5}).

Table 3.2 summarizes the abundance of the isotopes in these three elements. Notice that the ratio for chlorine is not exactly 3:1 nor that for bromine exactly 1:1; nevertheless you should use the simpler ratios when examining a mass spectrum. Always look at the heaviest peak first: see whether there is chlorine or bromine in it,

Table 3.2 Abundance of isotopes for carbon, chlorine, and bromine

Element	Major isotope: abundance	Minor isotope: abundance
carbon	^{12}C: 98.9%	^{13}C: 1.1%
chlorine	^{35}Cl: 75.8%	^{37}Cl: 24.2%
bromine	^{79}Br: 50.5%	^{81}Br: 49.5%

and whether the ratio of M^+ to $[M + 1]^+$ is about right. If, for example, you have what seems to be M^+ at 120 and the peak at 121 is 20% of the supposed M^+ at 120, then this cannot be a ^{13}C peak as it would mean that the molecule would have to contain 18 carbon atoms and you cannot fit 18 carbon atoms into a molecular ion of 120. Maybe 121 is the molecular ion.

Atomic composition can be determined by high resolution mass spectrometry

Ordinary mass spectra tell us the molecular weight (MW) of the molecule: we could say that the bee alarm pheromone was MW 114. When we said it was $C_7H_{14}O$ we could not really speak with confidence 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). Table 3.3 gives these to five decimal places, which is the sort of accuracy you need for meaningful results. Such accurate mass measurements are 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).

For the bee alarm pheromone, the accurate mass turns out to be 114.1039. Table 3.4 compares possible atomic compositions, 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.

Table 3.3 Exact masses of common elements

Element	Isotope	Atomic weight	Exact mass
hydrogen	^1H	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

A more important case is that of the three ions at 28: nitrogen, carbon monoxide, and ethylene (ethene, $CH_2{=}CH_2$). Actually mass spectra rarely go down to this low value because some nitrogen is usually injected along with the sample, but the three ions are all significant and it is helpful to see how different they are. Carbon monoxide CO is 27.9949, nitrogen N_2 is 28.0061, and ethylene 28.0313.

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.

Table 3.4 Exact mass determination for the bee alarm pheromone

Composition	Calculated M^+	Observed M^+	Error in p.p.m.
$C_6H_{10}O_2$	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

One thing you may have noticed in Table 3.4 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 *one nitrogen in it has an odd molecular weight*. Look back at the mass spectrum of the compounds giving good molecular ions by C.I. for an example. The nitro compound had M = 127 and the amine M = 121. This is because C, O, and N all have even atomic weights—only H has an odd atomic weight. Nitrogen is the only element from C, O, 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. Molecules with only C, H, and O or with even numbers of nitrogen atoms have even molecular weights.

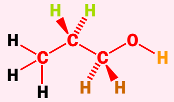
This rule holds as long as there are only C, H, N, O, S atoms in the molecule. It doesn't work for molecules with Cl or P atoms for example.

If we are talking about fragments, that is, cations or radicals, the opposite applies. A fragment has, by definition, an unused valency. Look back at the fragments in this section and you will see that this is so. Fragments with C, H, O alone have odd molecular weights, while fragments with one nitrogen atom have even molecular weights.

Nuclear magnetic resonance

What does it do?

Nuclear magnetic resonance (NMR) allows us to detect atomic nuclei and say what sort of environment they are in, within their molecule. Clearly, the hydrogen of, say, propanol's hydroxyl group is different from the hydrogens 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. Moreover, it can also distinguish between all the other different sorts of hydrogen atoms present. Likewise, carbon (or rather ^{13}C) NMR can easily distinguish between the three different carbon atoms. In this chapter we shall look at ^{13}C NMR spectra and then in Chapter 11 we shall look at proton (^{1}H) NMR spectra in detail.



Proton NMR can distinguish between the different coloured hydrogens. Carbon NMR can distinguish between all the carbons.

NMR is incredibly 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. We need first to spend some time explaining the principles of NMR.

When NMR is used medically it is usually called Magnetic Resonance Imaging (MRI) for fear of frightening 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. One effect would be to make navigation much harder since all compasses would be useless. They would be free to point in whatever direction they wanted to and, if we turned the needle round, it would simply stay where we left it. 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 and have different energy levels when placed in a magnetic field. The 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, our atomic nucleus is more restricted—its energy levels are quantized, just like the energy levels of an electron, which you will meet in the next chapter, and there are only certain specific energy levels it can adopt. This is like allowing our compass needle to point, say, only north or south. Some nuclei (including 'normal' carbon-12) do not interact with a magnetic field at all and cannot be observed in an NMR machine. The nuclei we shall be looking at, ^{1}H and ^{13}C, do interact and have just two different energy levels. When we apply a magnetic field to these nuclei, they can either align themselves with it, which would be the lowest energy state, or they can align themselves against the field, which is higher in energy.

Nuclei that interact with magnetic fields are said to possess **nuclear spin**. The exact number of different energy levels a nucleus can adopt is determined by this nuclear spin, I, of the particular isotope. The nuclear spin I can have various values such as 0, $\frac{1}{2}$, 1, $\frac{3}{2}$ and the number of energy levels is given by $2I + 1$. Some examples are: ^{1}H, $I = \frac{1}{2}$; ^{2}H (= D), $I = 1$; ^{11}B, $I = \frac{5}{2}$; ^{12}C, $I = 0$.

Let us return to the compass for a moment. We have already seen that if we could switch off the earth's magnetic field it would be easy to turn the compass needle round. When it is back on we need to push the needle (do work) to displace it from north. If we turned up the earth's magnetic field still more, it would be even harder to displace the compass needle. Exactly how hard it is to turn the compass needle depends on how strong the earth's magnetic field is and also on how well our 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, with our 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 also on the properties of the nucleus itself. The stronger the magnetic field we put our nucleus in, the greater the energy difference between the two alignments. Now here 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.

NMR machines contain very strong electromagnets

The earth's magnetic field has a field strength of 2×10^{-5} tesla. 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

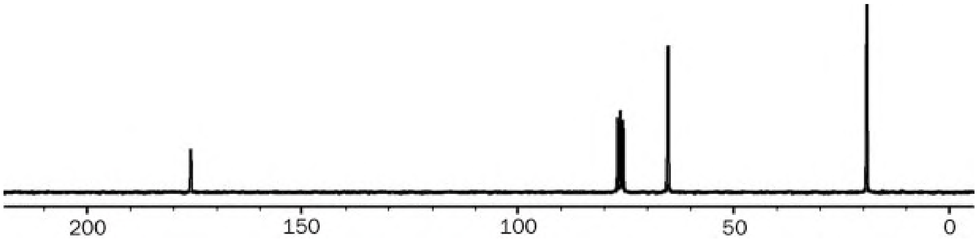
A ^{1}H or ^{13}C nucleus in a strong magnetic field can have two energy levels. We could do work to make our nucleus align against the field rather than with it (just like turning the compass needle round). But since the energy difference between the two states is so small, we don't need to do much work. In fact, the amount of energy needed to flip the nucleus can be provided by electromagnetic radiation of radio-wave frequency. Radio waves flip the nucleus from the lower energy state to the higher state. The nucleus now wants to return to the lower energy state and, when it does so, the energy comes out again and this (a tiny pulse of radiofrequency 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 and put in a very strong magnetic field. Any atomic nuclei with a nuclear spin now have 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. We then detect the energy given out when the nuclei fall back down to the lower energy level using what is basically a sophisticated radio receiver

<**Insert Graphic 03.022** (picture of NMR and diagram of components; not provided)>

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

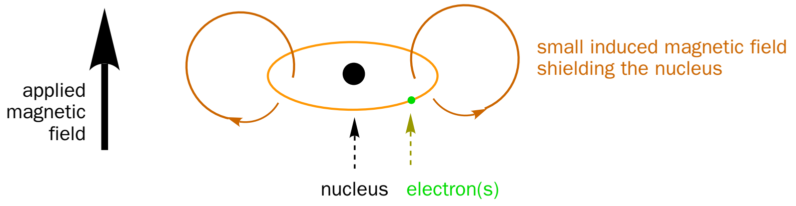


Why do chemically distinct nuclei absorb energy at different frequencies?

In the spectrum you see above, each line 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 carbon-13 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 there are different types of carbon atom is that their nuclei 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, and so does the resonating frequency of the ^{13}C nuclei. Now, you will see shortly (in Chapter 5) that a change in electron density at a carbon atom also alters the *chemistry* of that carbon atom. NMR tells us about the chemistry of a molecule as well as about its structure.

shielding of nuclei from an applied magnetic field by electrons:

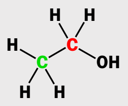


- Changes in the **distribution of electrons** around a nucleus affect:
  - the *local magnetic field* that the nucleus experiences
  - the *frequency* at which the nucleus resonates
  - the *chemistry* of the molecule at that atom

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

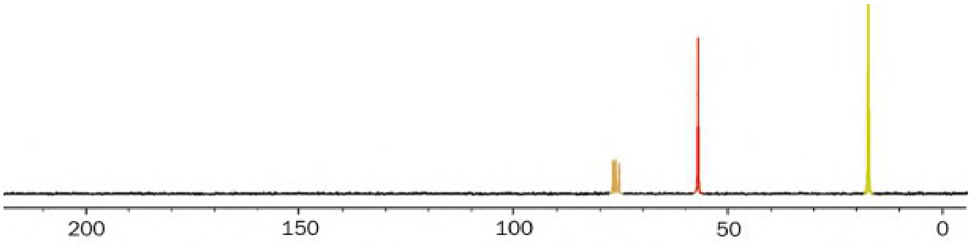
As an example, consider ethanol, CH_3CH_2OH. The carbon attached to the OH group will have relatively fewer electrons around it compared to the other carbon since the oxygen atom is more electronegative and draws electrons towards it, away from the carbon atom.

The magnetic field that this (red) carbon nucleus feels will therefore be slightly greater than that felt by the (green) carbon with more 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, there will be a greater energy difference between the two alignments of its nucleus. The greater the energy difference, the higher the resonant 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.



ethanol

We have shown all the Cs and Hs here because we want to talk about them.

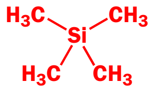


The peaks at 77 p.p.m., coloured brown, are those of the usual solvent ($CDCl_3$) and can be ignored for the moment. We shall explain them in Chapter 15.

The chemical shift scale

When you look at an NMR spectrum you will see that the scale does not appear to be in magnetic field units, nor in frequency units, but in ‘parts per million’ (p.p.m.). There is an excellent reason for this and we need to explain it. 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

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, 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!

Silicon and oxygen have opposite effects on an adjacent carbon atom: silicon shields; oxygen deshields.

Electronegativities: Si: 1.8; C: 2.5; O: 3.5.

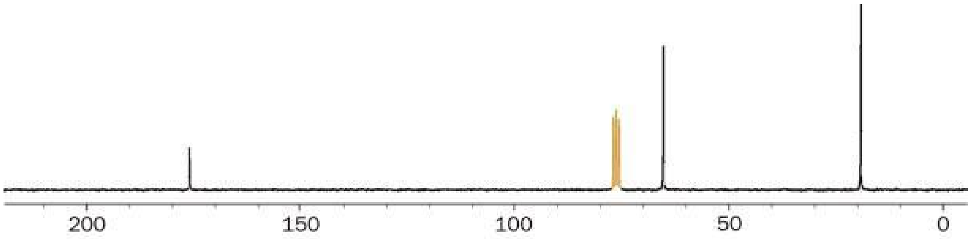
The chemical shift, δ, in parts per million (p.p.m.) of a given nucleus in our sample is defined in terms of the resonance frequency as:

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

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 p.p.m. whilst the (green) carbon of the methyl group resonates at 18.2 p.p.m. Notice that by definition TMS itself resonates at 0 p.p.m. The carbon nuclei in most organic compounds resonate at greater chemical shifts, normally between 0 and 200 p.p.m.

Now, let's return to the sample spectrum you saw on p. 000 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 p.p.m. of the field of the magnet, so each unit is one p.p.m. of 100 MHz, that is, 100 Hz. We can tell immediately from the three peaks at 176.8, 66.0, and 19.9 p.p.m. that there are three different types of carbon atom in the molecule.

Again, ignore the brown solvent peaks—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.

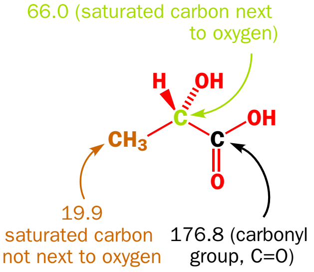


But we can do better than this: we can also work out what sort of chemical environment the carbon atoms are in. All ^{13}C spectra can be divided into four major regions: saturated carbon atoms (0–50 p.p.m.), saturated carbon atoms next to oxygen (50–100 p.p.m.), unsaturated carbon atoms (100–150 p.p.m.),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

Regions of the ^{13}C NMR spectrum (scale in p.p.m.)

Unsaturated carbon atoms next to oxygen (C=O)	Unsaturated carbon atoms (C=C and aromatic carbons)	Saturated carbon atoms next to oxygen (CH_3O, CH_2O, etc.)	Saturated carbon atoms (CH_3, CH_2, CH)
δ = 200–150	δ = 150–100	δ = 100–50	δ = 50–0

lactic acid (2-hydroxypropanoic acid)



The spectrum you just saw is in fact 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.

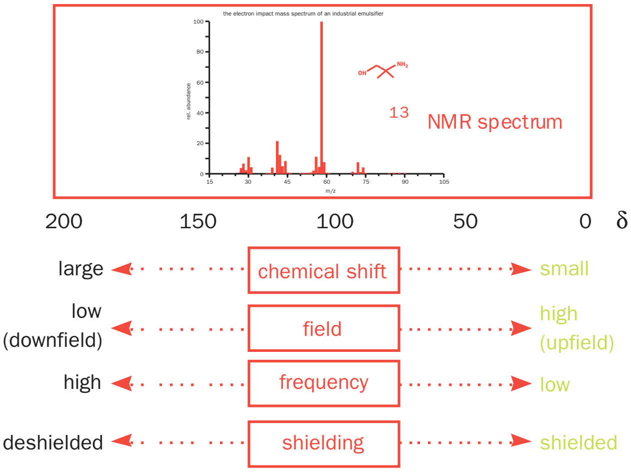
Different ways of describing chemical shift

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).

▶ 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 and low field' are a relic from the days of scanning by field variation.

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 p.p.m.) and resonate at high field. One electronegative oxygen atom moves the chemical shift downfield into the 50–100 p.p.m. region. The nucleus has become deshielded. Unsaturated carbon atoms experience even less shielding (100–150 p.p.m.) because of the way in which electrons are distributed around the nucleus. If the π bond is to oxygen, then the nucleus is even more deshielded and moves to the largest chemical shifts around 200 p.p.m. The next diagram summarizes these different ways of talking about NMR spectra.

■ If you are coming back to this chapter after reading Chapter 4 you might like to know that unsaturated C atoms are further deshielded because a π bond has a *nodal plane*. π Bonds have a plane with no electron density in at all, so electrons in π bonds are less efficient at shielding the nucleus than electrons in π bonds.

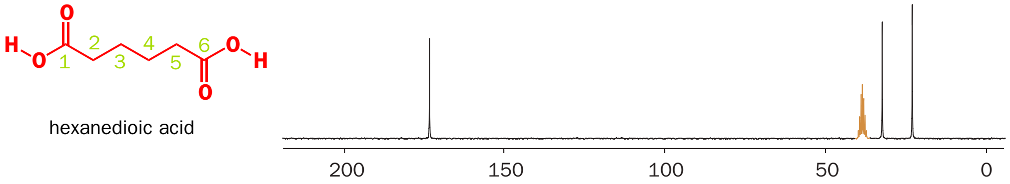


A guided tour of NMR spectra of simple molecules

We shall first look at NMR spectra of a few simple compounds before looking at unknown structures. 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 p.p.m. By the same token C2 and C5 are identical while C3 and C4 are identical. These are all in the saturated region 0–50 p.p.m. but it is likely that the carbons next to the electron-withdrawing CO_2H group are more deshielded than the others. So we assign C2/C5 to the peak at 33.2 p.p.m. and C3/C4 to 24.0 p.p.m.

■ 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.

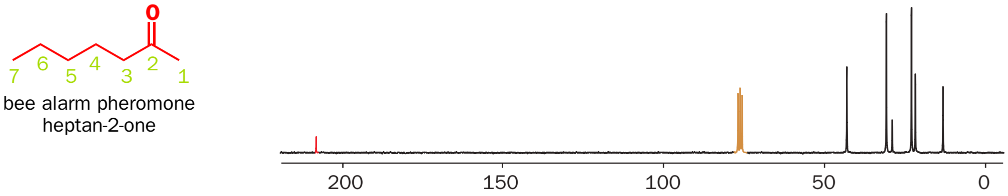
This spectrum was run in a different solvent, DMSO (DiMethyl SulfOxide); hence the brown solvent peaks are in a different region and have a different form. Whenever you first look at a spectrum, identify the peaks due to the solvent!



hexanedioic acid

The bee alarm pheromone (heptan-2-one) has no symmetry so all its seven carbon atoms are different. The carbonyl group is easy to identify (208.8 p.p.m., highlighted in red) but the rest are more difficult. Probably the two carbon atoms next to the carbonyl group come at lowest field, while C7 is certainly at highest field (13.9 p.p.m.). It is important that there are 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.

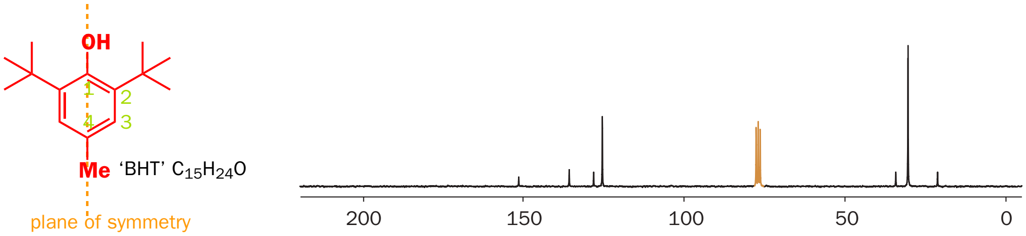
The peak due to the carbonyl carbon is particularly small in this spectrum. This is typical for quaternary carbons, i.e. carbons that have no protons attached to them.



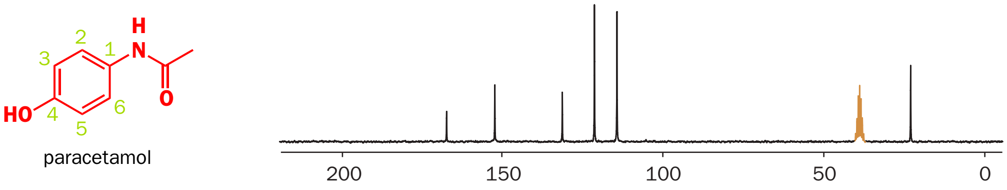
bee alarm pheromone
heptan-2-one

You met BHT on p. 000: 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. The very strong signal at $\delta = 30.4$ p.p.m. belongs to the six identical methyl groups on the *t*-butyl groups and the other two signals in the 0–50 p.p.m. range are the methyl group at C4 and the 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.

The NMR spectrum of BHT tells us that the *t*-butyl groups must also rotate rapidly as the three methyl groups give only one signal.



Paracetamol is a familiar painkiller with a simple structure—it too is a phenol but in addition it has an amide on the benzene ring. Its NMR spectrum contains one saturated carbon atom at 24 p.p.m. (the methyl group of the amide side chain), one carbonyl group at 168 p.p.m., and four other peaks at 115, 122, 132, and 153 p.p.m. These are the carbons of the benzene ring. Why four peaks? The two sides of the benzene ring are the same because the NHCO·CH_3 side chain can rotate rapidly so that C2 and C6 are the same and C3 and C5 are the same. Why is one of these aromatic peaks in the C=O region at 153 p.p.m.? This must be C4 as it is bonded to oxygen, and it just reminds us that carbonyl groups are not the only unsaturated carbon atoms bonded to oxygen (see the chart on p. 000), though it is not as deshielded as the true C=O group at 168 p.p.m.



paracetamol

The effects of deshielding within the saturated carbon region

We have mentioned deshielding several times. The reference compound TMS (Me_4Si) has very shielded carbon atoms because silicon is more electropositive than carbon. Oxygen moves a saturated carbon atom downfield to larger chemical shifts (50–100 p.p.m.) because it is much more electro*negative* than carbon and so pulls electrons away from a carbon atom by polarizing the C–O bond. In between these extremes was a CO_2H group that moved its adjacent carbon down to around 35 p.p.m. These variations in chemical shift within each of the 50 p.p.m. regions of the spectrum are a helpful guide to structure as the principle is simple.

- **Electro*negative* atoms move adjacent carbon atoms *down*field (to larger δ)by *de*shielding.**

For the carbon atom next to the carboxylic acid, the oxygen atoms are, of course, no longer adjacent but one atom further away, so their deshielding effect is not as great.

The reverse is true too: electro*positive* atoms move adjacent carbon atoms upfield by shielding. This is not so important as there are few atoms found in organic molecules that are more electropositive than silicon and so few carbons are more shielded than those in Me_4Si. About the only important elements like this are the metals. When a carbon atom is more shielded than those in TMS, it has a negative δ value. There is nothing odd about this—the zero on the NMR scale is an arbitrary point. Table 3.5 shows a selection of chemical shift changes caused to a methyl group by changes in electronegativity.

Table 3.5 Effect of electronegativity on chemical shift

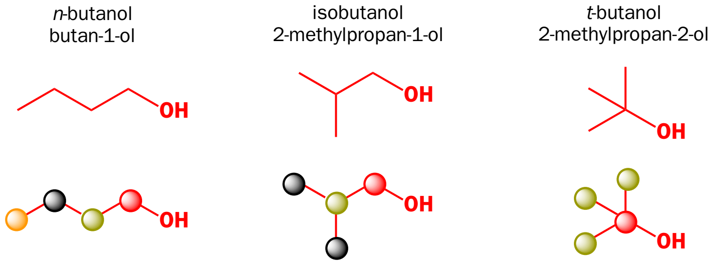
Electronic effect	Electronegativity of atom bonded to carbon	Compound	δ(CH_3)	δ(CH_3) – 8.4
donation	1.0	CH_3–Li	–14	–22.4
↑	2.2	CH_3–H	–2.3	–10.7
weak	1.8	CH_3–$SiMe_3$	0.0	–8.4
no effect	2.5	CH_3–CH_3	8.4	0
weak	3.1	CH_3–NH_2	26.9	18.5
↓	—	CH_3–COR	~30	~22
↓	3.5	CH_3–OH	50.2	41.8
withdrawal	4.1	CH_3–F	75.2	66.8

The last column in Table 3.5 shows the effect that each substituent has when compared to ethane. In ethane there is no electronic effect because the substituent is another methyl group so this column gives an idea of the true shift caused by a substituent. These shifts are roughly additive. Look back at the spectrum of lactic acid on p. 000: the saturated carbons occur at 19.9 and 66.0. The one at 66.0 is next both to an oxygen atom and a carbonyl group so that the combined effect would be about 42 + 22 = 64—not a bad estimate.

■ We shall look at similar but more detailed correlations in Chapters 11 and 15.

NMR is a powerful tool for solving unknown structures

Simple compounds can be quickly distinguished by NMR. These three alcohols of formula $C_4H_{10}O$ have quite different NMR spectra.

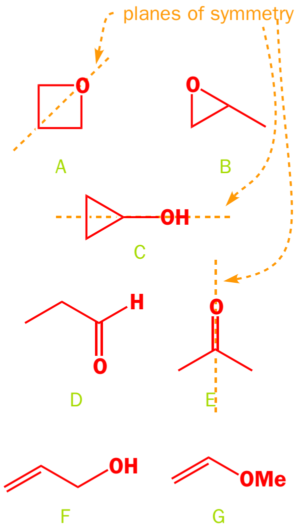


	***n*-butanol**	**isobutanol**	***t*-butanol**
🔴	62.9	70.2	69.3
🟢	36.0	32.0	32.7
⚫	20.3	20.4	—
🟠	15.2	—	—

▶ The C atoms have been arbitrarily colour-coded.

■ The meanings of *n*-, iso-, and *t*- were covered in Chapter 2 (p. 000).

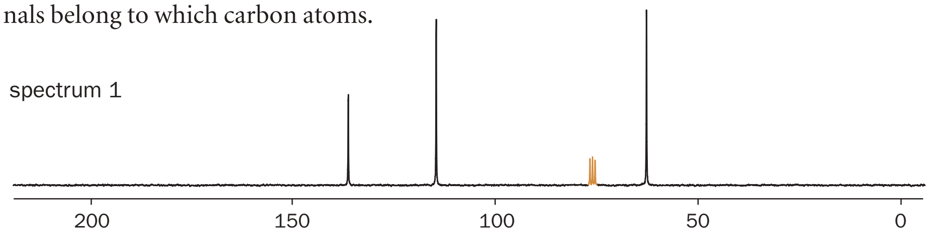
Each alcohol has a saturated carbon atom next to oxygen, all close together. Then there are carbons next door but one to oxygen: they are back in the 0–50 p.p.m. region but at its low field end—about 30–35 p.p.m.. Notice the similarity of these chemical shifts to those of carbons next to a carbonyl group (Table 3.5 on p. 000). In each case we have C–C–O and the effects are about the same. Two of the alcohols have carbon(s) one further away still at yet smaller chemical shift (further upfield, more shielded) at about 20 p.p.m., 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.

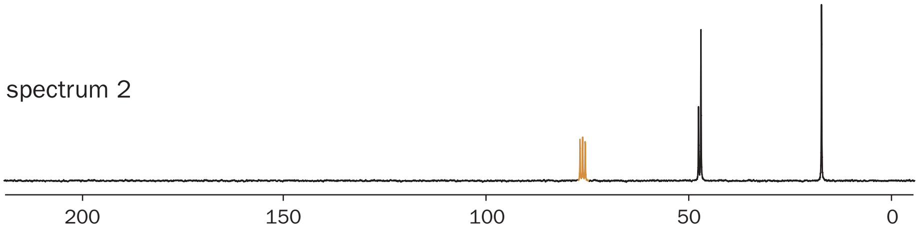


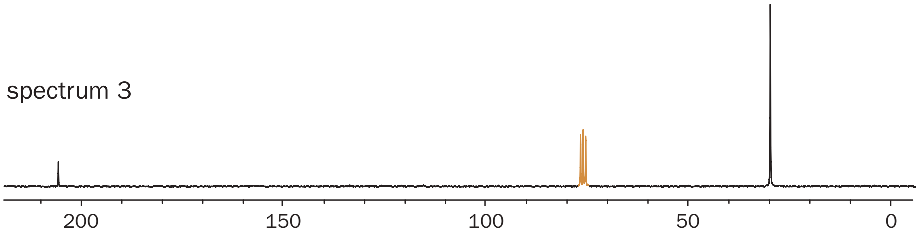
■ An epoxide is a three-membered cyclic ether.

A more realistic example would be an unknown molecule of formula C_3H_6O. There are seven reasonable structures, as shown. Simple symmetry can distinguish structures A, C, and E from the rest as these three have only two types of carbon atom. A more detailed inspection of the spectra makes identification easy. The two carbonyl compounds, D and E, each have one peak in the 150–220 p.p.m. region but D has two different saturated carbon atoms while E has only one. The two alkenes, F and G, both have one saturated carbon atom next to oxygen, but F has two normal unsaturated carbon atoms (100–150 p.p.m.) while the enol ether, G, has one normal alkene and one unsaturated carbon joined to oxygen. The three saturated compounds (A–C) present the greatest problem. The epoxide, B, has two different carbon atoms next to oxygen (50–100 p.p.m.) and one normal saturated carbon atom. The remaining two both have one signal in the 0–50 and one in the 50–100 p.p.m. regions. Only proton NMR (Chapter 11) and, to a certain extent, infrared spectroscopy (which we will move on to shortly) will distinguish them reliably.

Here are NMR spectra of three of these molecules. Before looking at the solutions, cover up the rest of the page and see if you can assign them to the structures above. Try also to suggest which signals belong to which carbon atoms.







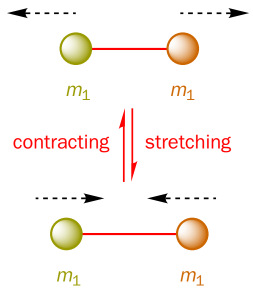
These shouldn't give you too much trouble. The only carbonyl compound with two identical carbons is acetone, Me_2CO (E) so spectrum 3 must be that one. Notice the very low field C=O signal (206.6 p.p.m.) typical of a simple ketone. Spectrum 1 has two unsaturated carbons and a saturated carbon next to oxygen so it must be F or G. In fact it has to be F as both unsaturated carbons are similar (137 and 116 p.p.m.) and neither is next to oxygen (>150 p.p.m., cf. 206.6 in spectrum 3). This leaves spectrum 2, which appears to have no carbon atoms next to oxygen as all chemical shifts are less than 50 p.p.m. No compound fits that description (impossible for C_3H_6O anyway!) and the two signals at 48.0 and 48.2 p.p.m. are suspiciously close to the borderline. They are, of course, next to oxygen and this is compound B.

Infrared spectra

Functional groups are identified by infrared spectra

Some functional groups, for example, C=O or C=C, can be seen in the NMR spectrum because they contain carbon atoms, while the presence of others like OH can be inferred from the chemical shifts of the carbon atoms they are joined to. Others cannot be seen at all. These might include NH_2 and NO_2, as well as variations around a carbonyl group such as COCl, CO_2H, and $CONH_2$. Infrared (IR) spectroscopy provides a way of finding 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.

bond vibration in the infrared



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, corresponds to rather shorter wavelengths. These wavelengths lie in the infrared, that is, heat radiation just to the long wavelength side of visible light. When the carbon skeleton of a molecule vibrates, all the bonds stretch and relax in combination and these absorptions are unhelpful. However some bonds stretch essentially independently of the rest of the molecule. 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** 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.

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

The equation shows that the frequency of the vibration ν is proportional to the (root of) a **force constant** f—more or less the bond strength—and inversely proportional to the (root of) a **reduced mass** μ, that is, the product of the masses of the two atoms forming the bond divided by their sum.

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

Stronger bonds vibrate faster and so do lighter atoms. You may at first think that stronger bonds ought to vibrate more slowly, but a moment's reflection will convince you of the truth: which stretches and contracts faster, a tight steel spring or a slack steel spring?

Infrared spectra are simple absorption spectra. The sample is exposed to infrared radiation and the wavelength scanned across the spectrum. Whenever energy corresponding to a specific wavelength is absorbed, the intensity of the radiation reaching a detector momentarily decreases, and this is recorded in the spectrum. Infrared spectra are usually recorded using a frequency measurement called **wavenumber** (cm^{-1}) which is the inverse of the true wavelength λ in centimetres to give convenient numbers (500–4000 cm^{-1}). Higher numbers are to the left of the spectrum because it is really wavelength that is being scanned.

We need to use another equation here:

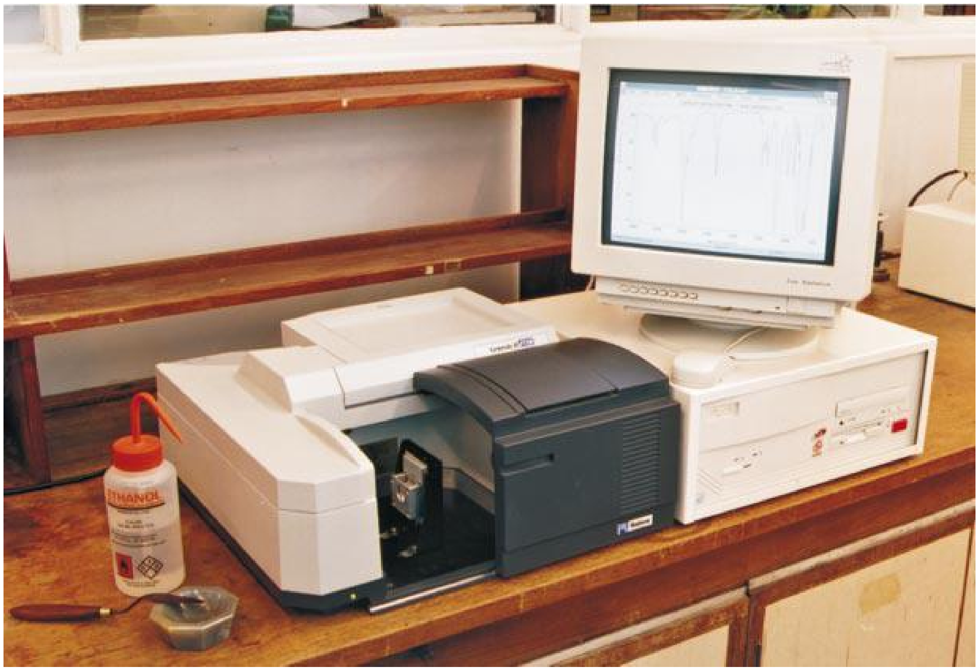
$$E = h\nu = h\frac{c}{\lambda} \text{ since } \lambda = \frac{c}{\nu}$$

The energy, E, required to excite a bond vibration can be expressed as the inverse of a wavelength λ or as a frequency ν. Wavelength and frequency are just two ways of measuring the same thing. More energy is needed to stretch a strong bond and you can see from this equation that larger E means higher wavenumbers (cm^{-1}) or smaller wavelength (cm).

To run the spectrum, the sample is either dissolved in a solvent such as $CHCl_3$ (chloroform) that has few IR absorptions, pressed into a transparent disc with powdered solid KBr, or ground into an oily slurry called a **mull** with a hydrocarbon oil called 'Nujol'. Solutions in $CHCl_3$ cannot be used for looking at the regions of C–Cl bond stretching nor can Nujol mulls be used for the region of C–H stretching. Neither of these is a great disadvantage, especially as nearly all organic compounds have some C–H bonds anyway.

h is **Planck's constant** and *c* the velocity of light.

You should always check the way the spectrum was run before making any deductions!



We shall now examine the relationship between bond stretching and frequency in more detail. Hooke's law told us to expect frequency to depend on both mass and bond strengths, and we can illustrate this double dependence with a series of bonds of various elements to carbon.

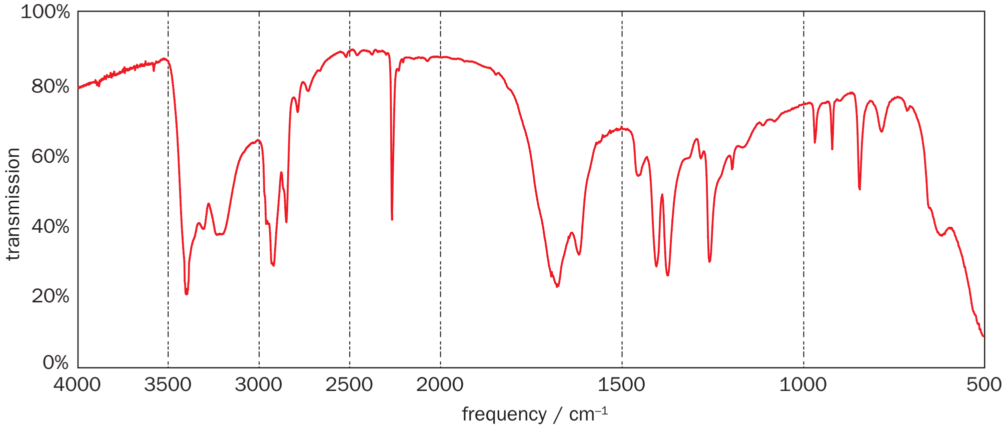
Values chiefly affected by mass of atoms: (lighter atom, higher frequency)

C–H	C–D	C–O	C–Cl
3000 cm^{-1}	2200 cm^{-1}	1100 cm^{-1}	700 cm^{-1}

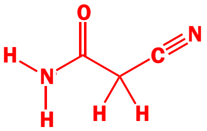
Values chiefly affected by bond strength (stronger bond, higher frequency)

C≡O	C=O	C–O
2143 cm^{-1}	1715 cm^{-1}	1100 cm^{-1}

Just because they were first recorded in this way, infrared spectra have the baseline at the top and peaks going downwards. You might say that they are plotted upside down and back to front. At least you are now accustomed to the horizontal scale running backwards as that happens in NMR spectra too. A new feature is the change in scale at 2000 cm^{-1} so that the right-hand half of the spectrum is more detailed than the left-hand half. A typical spectrum looks like this.



IR spectra are plotted 'upside down' because they record **transmission** (the amount of light reaching the detector) rather than absorbance.



cyanoacetamide
(spectrum taken as a Nujol mull)

There are four important regions of the infrared spectrum

You will see at once that the infrared spectrum contains many lines, particularly at the right-hand (lower frequency) end; hence the larger scale at this end. Many of these lines result from several bonds vibrating together and it is actually the left-hand half of the spectrum that is more useful.

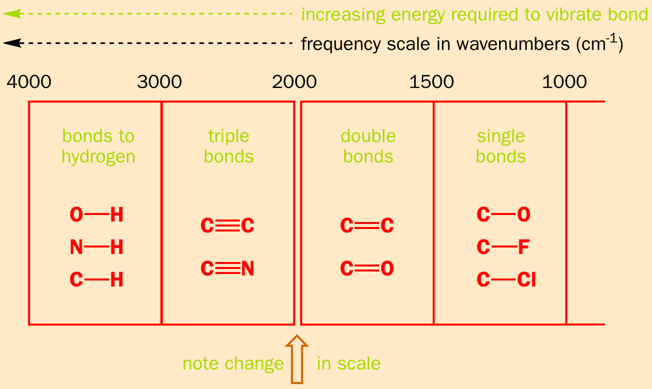
The first region, from about 4000 to about 2500 cm^{-1} is the region for C–H, N–H, and O–H bond stretching. Most of the atoms in an organic molecule (C, N, O, for example) are about the same weight. Hydrogen is an order of magnitude lighter than any of these and so it dominates the stretching frequency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

The concept of reduced mass was introduced on p. 000.

Even the strongest bonds—triple bonds such as C≡C or C≡N—absorb at slightly lower frequencies than bonds to hydrogen: these are in the next 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: double bonds such as C=C and C=O from about 1900–1500 cm^{-1} and single bonds at the right-hand end of the spectrum. These regions are summarized in this chart, which you should memorize.

Remember. Hooke's law says that frequency depends on both mass and a force constant (bond strength).

The regions of the infrared spectrum



Looking back at the typical spectrum, we see peaks in the X–H region at about 2950 cm^{-1} which are the C–H stretches of the CH_3 and CH_2 groups. The one rather weak peak in the triple bond region (2270 cm^{-1}) is of course 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—and, similarly, it cannot be 'interpreted'. It is indeed called the **fingerprint region**. The useful information from this spectrum is the presence of the CN and C=O groups and the exact position of the C=O absorption.

The X–H region 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} (though they are of little use as virtually all organic compounds have C–H bonds), N–H stretches occur at about 3300 cm^{-1}, and O–H stretches higher still. 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.

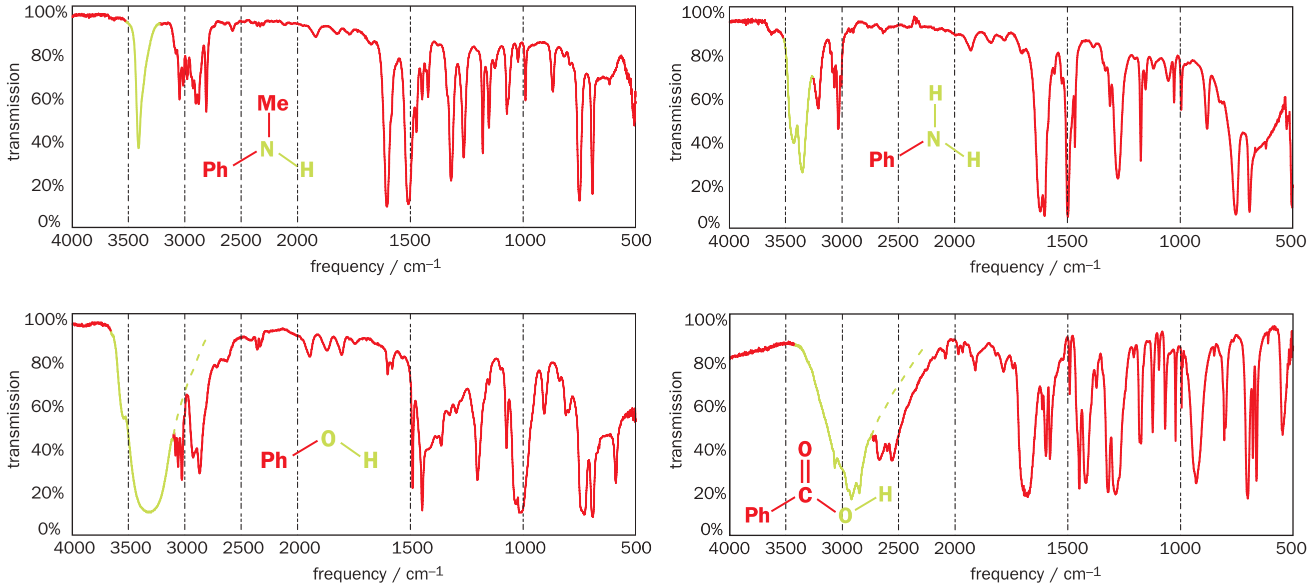
Table 3.6 IR bands for bonds to hydrogen

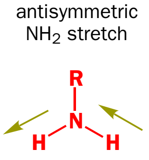
Bond	Reduced mass, μ	IR frequency, cm^{-1}	Bond strength, kJ mol^{-1}
C–H	12/13 = 0.92	2900–3200	CH_4: 440
N–H	14/15 = 0.93	3300–3400	NH_3: 450
O–H	16/17 = 0.94	3500–3600[a]	H_2O: 500

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

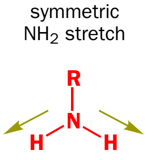
■ This may surprise you: you may be used to thinking of O–H as more reactive than CH. This is, of course, true but, as you will see in Chapter 5, factors other than bond strength control reactivity. Bond strengths will be much more important when we discuss radical reactions in Chapter 39.

The X–H IR stretches are very different in these four compounds.





about
3400 cm^{-1}

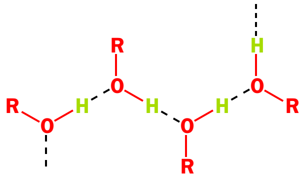


about
3300 cm^{-1}

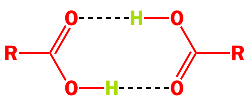
The IR peak of an NH group is different from that of an NH_2 group. A group gives an independent vibration only if both bond strength and reduced mass are different from those of neighbouring bonds. In the case of N–H, this is likely to be true and we usually get a sharp peak at about 3300 cm^{-1}, whether 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, 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. The sharp absorption at 3600 cm^{-1} is the non-hydrogen-bonded OH and the lower the absorption 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 (though 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 typically give a rounded absorption at about 3300 cm^{-1} (contrast the sharp N–H stretch in the same region). 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 very broad V-shaped absorbances.

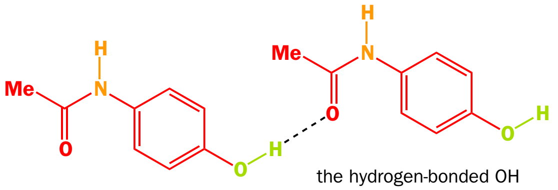


hydrogen bonding in an alcohol

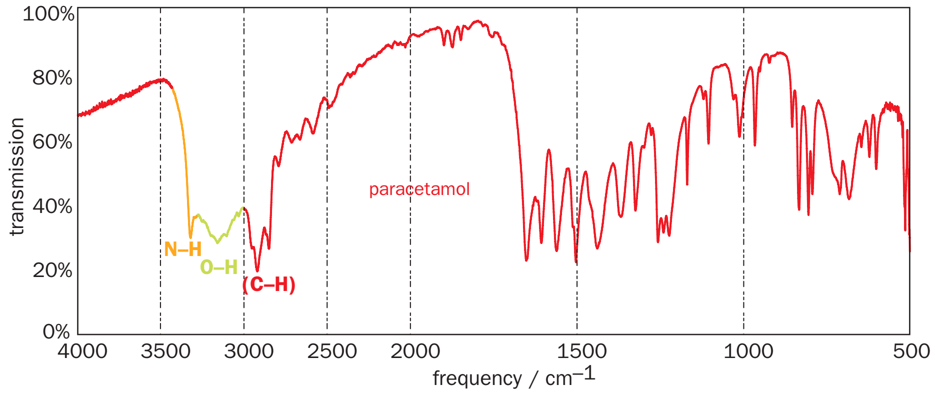


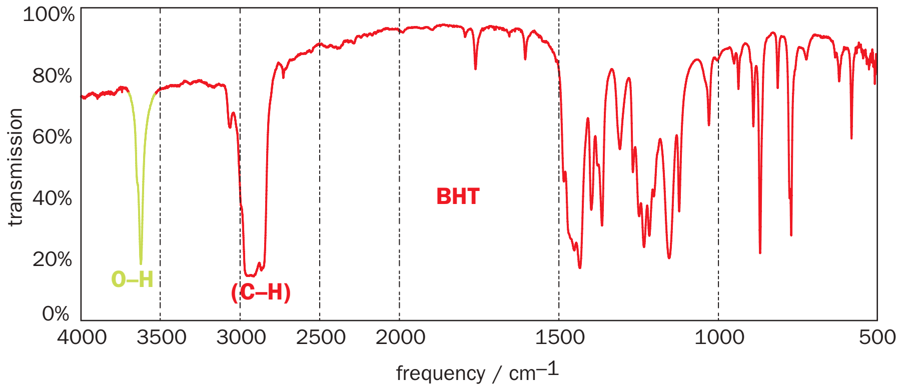
the hydrogen-bonded dimer of a carboxylic acid

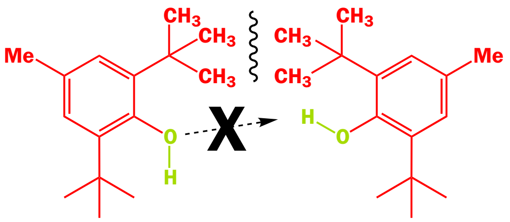
Good examples are paracetamol and BHT. Paracetamol has a typical sharp peak at 3330 cm^{-1} for the N–H stretch and then a rounded absorption for the hydrogen-bonded O–H stretch from 3300 down to 3000 cm^{-1} in the gap between the N–H and C–H stretches. By contrast, BHT has a sharp absorption at 3600 cm^{-1} as the two large and roughly spherical *t*-butyl groups prevent the normal H bond from forming.



the hydrogen-bonded OH group in paracetamol



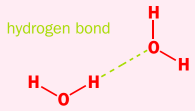




in BHT H-bonding is prevented by large *t*-butyl groups

▶

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.

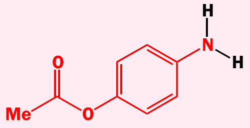


■

The ^{13}C NMR spectra of these two compounds are on pp. 000 and 000.

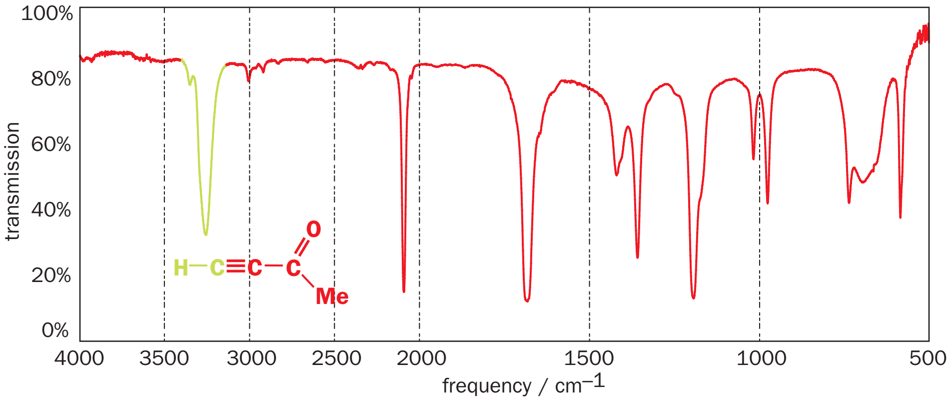
▶

We can use the N–H and O–H absorptions to rule out an alternative isomeric structure for paracetamol: an ester with an NH_2 group instead of an amide with NH and OH. This structure must be wrong as it would give two similar sharp peaks at about 3300 cm^{-1} instead of the one sharp and one broad peak actually observed.



alternative and wrong structure for paracetamol

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 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.



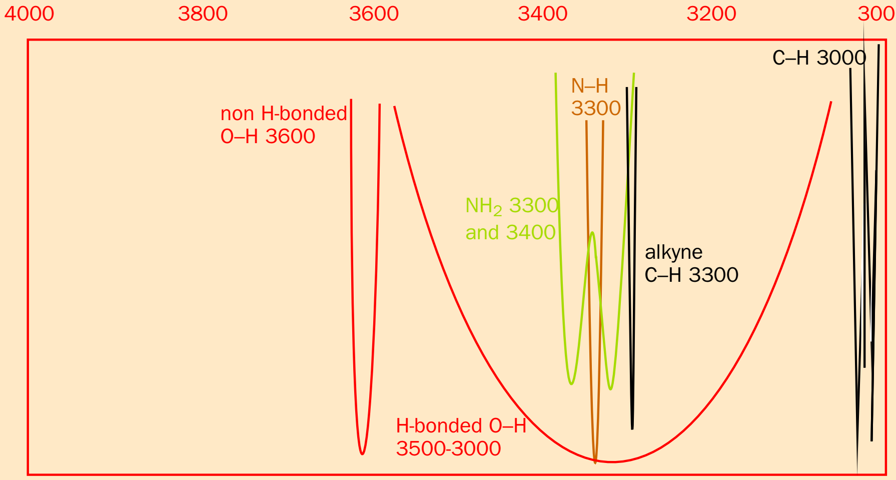
▶

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.

■

What are the other peaks in this spectrum?

● **The summary chart shows some typical peak shapes and frequencies for X–H bonds in the region 4000–3000 cm^{-1}.**



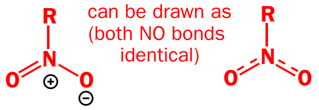
The double bond region is the most important in IR spectra

In the double bond region, there are three important absorptions, those of the carbonyl (C=O), alkene (C=C), and nitro (NO_2) groups. All give rise to sharp bands: C=O to one strong (i.e. intense) band anywhere between 1900 and 1500 cm^{-1}; C=C to one weak band at about 1640 cm^{-1}; and NO_2 to two strong (intense) bands in the mid-1500s and mid-1300s cm^{-1}. The number of bands is easily dealt with. 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 appears to contain two different groups, N^+–O^- and N=O, but delocalization means they are identical and we see absorption for symmetrical and antisymmetrical stretching vibrations. As with NH_2, more work is needed for the antisymmetrical vibration which occurs at higher frequency (>1500 plus cm^{-1}).

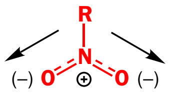
▶

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

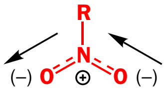
delocalization in the nitro group



symmetric
NO_2 stretch (~1350)



antisymmetric
NO_2 stretch (~1550)



The strength of an IR absorption depends on dipole moment

Now what about the variation in strength (i.e. intensity, the amount of energy absorbed)? The strength of an IR absorption varies with the change of dipole moment 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 is weaker in the IR. Indeed it may be absent altogether in a symmetrical alkene. By contrast the carbonyl group is very polar (Chapter 4) and stretching it causes a large change in dipole moment and C=O stretches are usually the strongest peaks in the IR spectrum. You may also have noticed that O–H and N–H stretches are stronger than C–H stretches (even though most organic molecules have many more C–H bonds than O–H or N–H bonds): the reason is the same.

▶ Contrast the term 'strength' applied to absorption and to bonds. A stronger absorption is a *more intense* absorption—i.e. one with a big peak. A strong *bond* on the other hand has a *higher frequency* absorption (other things being equal).

Dipole moments

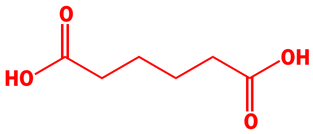
Dipole moment depends on the variation in distribution of electrons along the bond, and also its length, which is why stretching a bond can change its dipole moment. For bonds between unlike atoms, the larger the difference in electronegativity, the greater the dipole moment, and the more it changes 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 are measured using Raman spectra. This is an IR-based technique using scattered light that relies on 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.

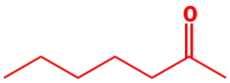
● Absorptions in IR spectra

Position of band depends on ⟶	reduced mass of atoms	**light** atoms give **high** frequency
	bond strength	**strong** bonds give **high** frequency
Strength 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 **wide** peak

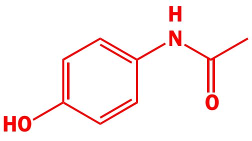
We have seen three carbonyl compounds so far in this chapter and they all show peaks in the right region (around 1700 cm^{-1}) even though one is a carboxylic acid, one a ketone, and one an amide. We shall consider the exact positions of the various carbonyl absorptions in Chapter 15 after we have discussed some carbonyl chemistry.



hexanedioic acid
1720 cm^{-1}



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



paracetamol
1667 cm^{-1}

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.

In addition, C–C bonds are likely to be 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 and different enough (Table 3.7) 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) 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 the 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.

Table 3.7 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}

Table 3.8 Useful deformations (bending vibrations)

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

Finally in this section, we summarize all the useful absorptions in the fingerprint region. Please be cautious in applying these as there are other reasons for bands in these positions.

■ You may not yet understand all the terms in Table 3.9, but you will find it useful to refer back to later.

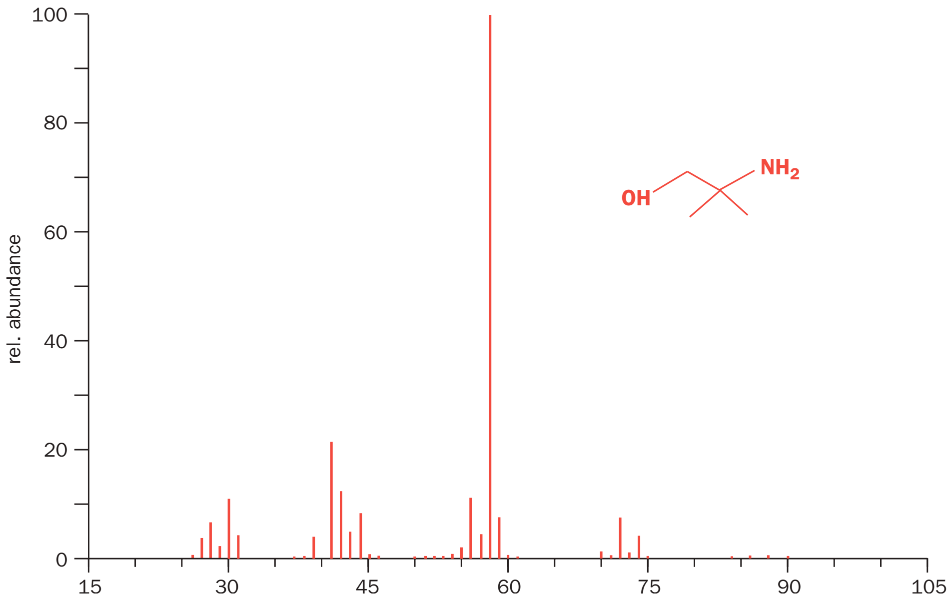
Table 3.9 Useful absorptions in the fingerprint region

Frequency, cm^{-1}	Strength	Group	Comments
1440–1470	medium	CH_2	deformation (present in nujol)
~1380	medium	CH_3	deformation (present in nujol)
~1350	strong	NO_2	symmetrical N=O stretch
1250–1300	strong	P=O	double bond stretch
1310–1350	strong	SO_2	antisymmetrical S=O stretch
1120–1160	strong	SO_2	symmetrical S=O stretch
~1100	strong	C–O	single bond stretch
950–1000	strong	C=CH	*trans* alkene (out-of-plane deformation)
~690 *and* ~750	strong	Ar–H	five adjacent Ar–H (out-of-plane)
~750	strong	Ar–H	four adjacent Ar–H (out-of-plane)
~700	strong	C–Cl	single bond stretch

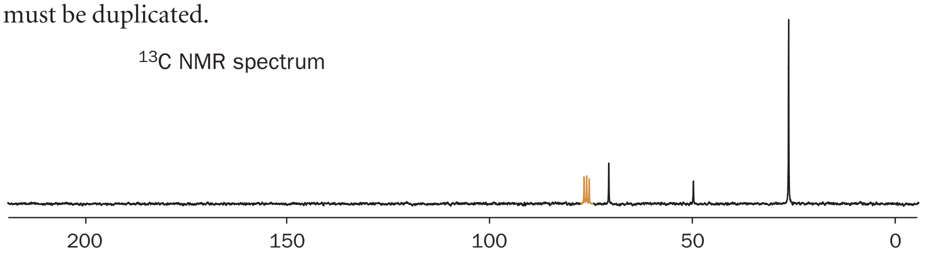
Mass spectra, NMR, and IR combined make quick identification possible

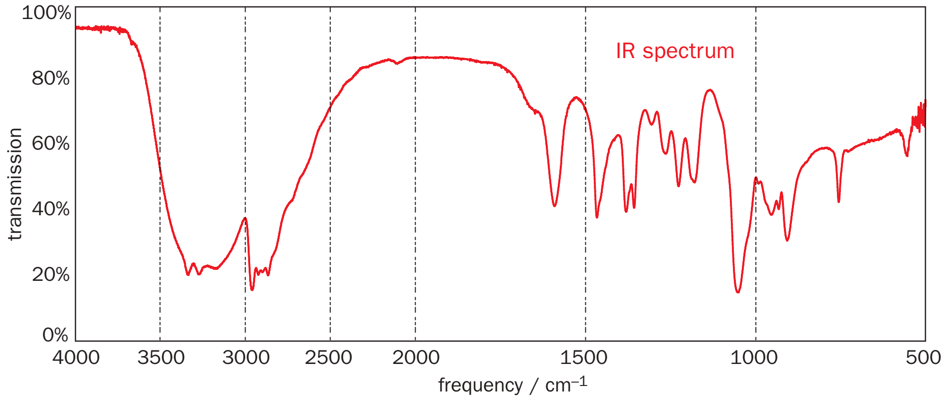
If these methods are each as powerful as we have seen on their own, how much more effective they must be together. We shall finish this chapter with the identification of some simple unknown compounds using all three methods. The first is an industrial emulsifier used to blend solids and liquids into smooth pastes. Its electron impact mass spectrum has peaks at 75 and 74 (each about 20%) and a base peak at 58. The two peaks at 75 and 74 cannot be isotopes of bromine as the separation is only one mass unit, nor can 75 be a ^{13}C peak as it is far too strong. It looks as though 75 might be the molecular ion and 74 an unusual loss of a hydrogen atom. However a chemical ionization mass spectrum reveals a molecular ion at 90 (MH^+) and hence the true molecular ion at 89. An odd molecular weight (89) suggests one nitrogen atom, and high resolution mass spectrometry reveals that the formula is $C_4H_{11}NO$.

the electron impact mass spectrum of an industrial emulsifier

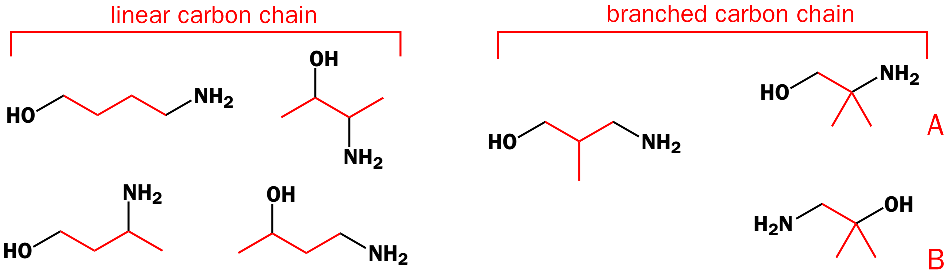


The ^{13}C NMR spectrum has only three peaks so two carbon atoms must be the same. There is one signal for saturated carbon next to oxygen, and two for other saturated carbons, one more downfield than the other. The IR spectrum reveals a broad peak for an OH group with two sharp NH_2 peaks just protruding. 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 one of the remaining carbons must be duplicated.


^{13}C NMR spectrum

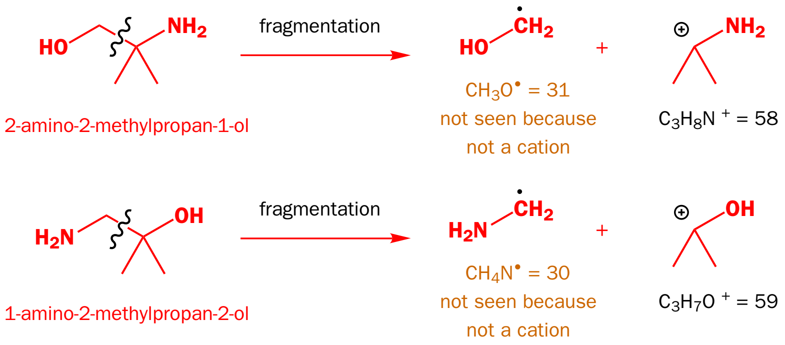


The next stage is one often overlooked. We don't seem to have much information, but try and put the two fragments together, knowing the molecular formula, and there's very little choice. The carbon chain (shown in red) could either be linear or branched and that's it!



▶ By **chain terminating** we mean only attachable to one other atom.

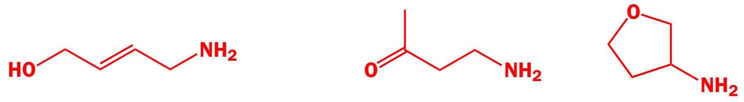
There is no room for double bonds or rings because we need to fit in the eleven hydrogen atoms. We cannot put N or O in the chain because we know from the IR that we have the *chain terminating* groups OH and NH_2. Of the seven possibilities only the last two, A and B, are serious since they alone have two identical carbon atoms (the two methyl groups in each case); all the other structures would have four separate signals in the NMR. How can we choose between these? The base peak in the mass spectrum was at 58 and this fits well with a fragmentation of one structure but not of the other: the wrong structure would give a fragment at 59 and not 58. The industrial emulsifier is 2-amino-2-methylpropan-1-ol.



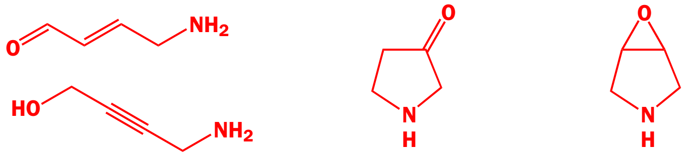
Double bond equivalents help in the search for a structure

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. This is usually expressed as 'double bond equivalents'. It may seem obvious to you that, if $C_4H_{11}NO$ 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



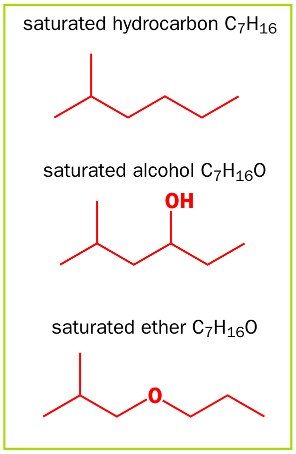
some structures for C_4H_7NO



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**).

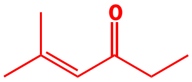
You can work out how many DBEs there are in a given atomic composition just by making a drawing of one possible structure (all possible structures have the same number of DBEs). Alternatively, you can calculate the DBEs if you wish. A saturated hydrocarbon with n carbon atoms has $(2n + 2)$ hydrogens. 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.

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:



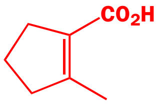
All have $(2n + 2)$ H atoms

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

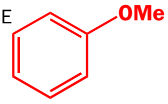


$C_7H_{12}O$ = two DBE

The unsaturated ketone does indeed have an alkene and a carbonyl group. The unsaturated cyclic acid has: $16 - 10 = 6$ divided by $2 = 3$ DBEs and it has one alkene, one C=O and one ring. Correct. 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.

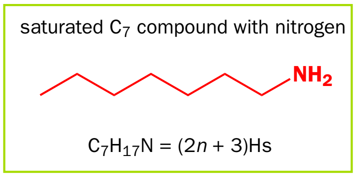


$C_7H_{10}O$ = two DBE



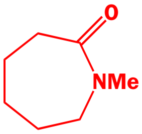
C_7H_8O = four DBE

Nitrogen makes a difference. Every nitrogen adds *one extra hydrogen* atom because nitrogen can make three bonds. This is one fewer hydrogen to subtract. 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.

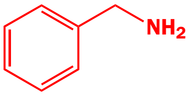


$C_7H_{17}N = (2n + 3)$Hs

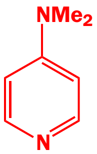
$C_7H_{15}NO_2$ = one DBE



$C_7H_{13}NO$ = two DBE



C_7H_9NO= 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. The last compound (we shall meet this later as ‘DMAP’) has:

1	Maximum number of H atoms for 7 Cs	$2n + 2 = 16$
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. You would be wise to check 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 finding one possible structure before you really start! One good tip is that if you have few hydrogens relative to the number of carbon atoms (and at least four DBEs) then there is probably an aromatic ring in the compound.

►

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: $2n + 2$ Hs if C, H, O only

(b) For C_nN_m there would be $2n + 2 + m$ Hs

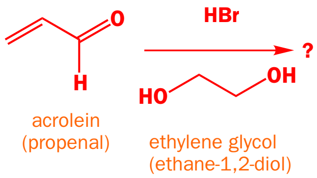
2 Subtract the actual number of Hs and divide by 2. This gives the DBEs

3 If there are other atoms (Cl, B, P, etc.) it is best to draw a trial structure

4 If there are few Hs, e.g. less than the number of Cs, suspect a benzene ring

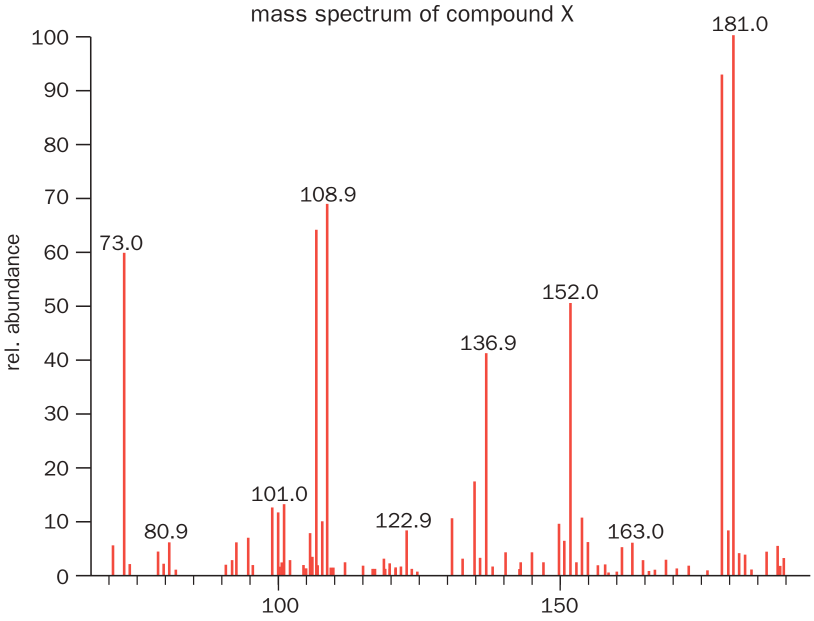
5 A benzene ring has *four* DBEs (three for the double bonds and one for the ring)

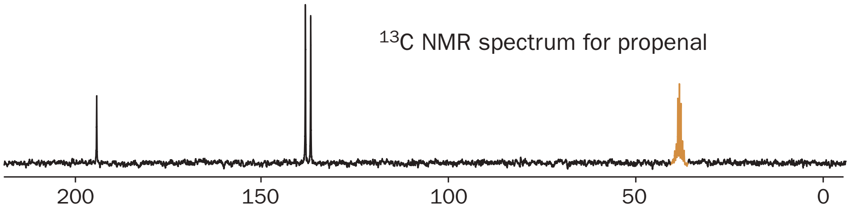
6 A nitro group has *one* DBE only

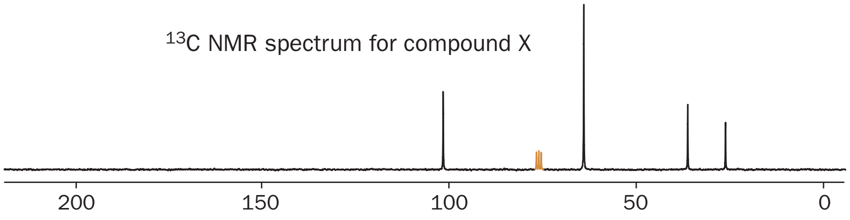


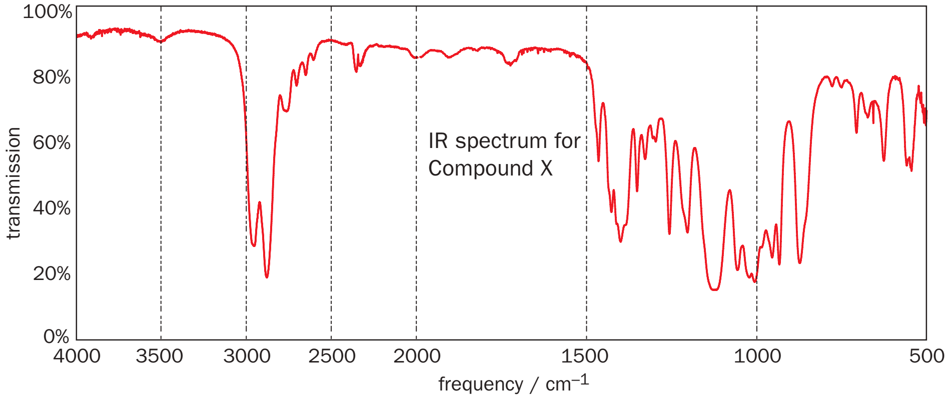
An unknown compound from a chemical reaction

Our second 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 one hour at room temperature. Distillation of the reaction mixture gives a colourless liquid, compound X. What is it?


mass spectrum of compound X


^{13}C NMR spectrum for propenal


^{13}C NMR spectrum for compound X



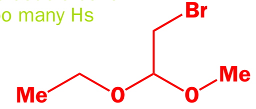
The mass spectrum shows a molecular ion (181) much heavier than that of the starting material, C_3H_4O = 56. Indeed it shows two molecular ions at 181/179 typical of a bromo-compound, so it looks as if HBr has added to the aldehyde somehow. High resolution reveals a formula of $C_5H_9BrO_2$ and the five 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. A trial structure reveals *one* DBE.

C_3H_4O + $C_2H_6O_2$ + HBr = $C_5H_{11}BrO_3$ = $C_5H_9BrO_2$ + H_2O

The next thing is to see what remains of the propenal. The 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 p.p.m. just creeping into the double bond region. It can't be an alkene as an alkene is impossible with only one carbon atom! 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. Two of these could be the same, but where are the rest?

We can solve this problem with a principle you may have guessed at before. If one oxygen atom takes a saturated carbon atom downfield to 50 p.p.m. or more, what could take a carbon downfield to 100 p.p.m. or more? We have established that chemical shifts are roughly additive so two oxygen atoms would just do. This would give us a fragment C–O–C–O–C accounting for three of the five carbon atoms. If you try and join the rest up with this fragment, you will find that you can't do it without a double bond, for example, the structure in the margin.

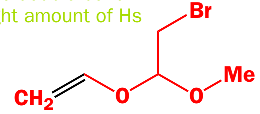
no double bond – too many Hs



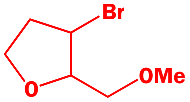
$C_5H_{11}BrO_2$

But 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 join the skeleton of the present molecule up in many rings of various sizes like this one in the margin.

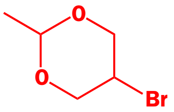
one double bond – right amount of Hs

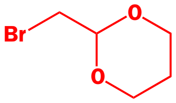


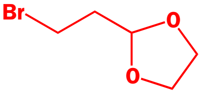
$C_5H_9BrO_2$

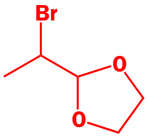


But this won't do as it would have five different carbon atoms. It is much more likely that the basic skeletons of the organic reagents are preserved, that is, that we have a two-carbon and a three-carbon fragment joined through oxygen atoms. This gives four possibilities.

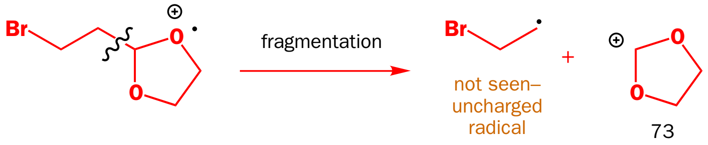






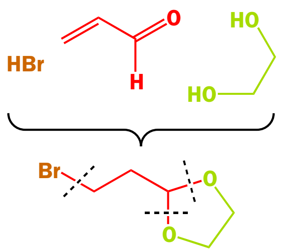


These are all quite reasonable, though we might prefer the third as it is easier to see how it derives from the reagents. A decision can easily be reached from the base peak in the mass spectrum at 73. This is a fragment corresponding to the five-membered ring and not to the six-membered ring. The product is in fact the third possibility.



Looking forward to Chapters 11 and 14

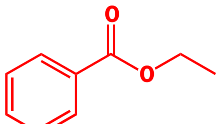
We have only begun to explore the intricate world of identification of structure by spectroscopy. It is important that you recognize that structures are assigned, not because of some theoretical reason or because a reaction ‘ought’ to give a certain product, but because of sound evidence from spectra. You have seen three powerful methods—mass spectra, ^{13}C NMR, and IR spectroscopy in this chapter. In Chapter 11 we introduce the most important of all—proton (^{1}H) NMR and, finally, in Chapter 14 we shall take each of these 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 proton 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.



Problems

1. How does the mass spectrum give evidence of isotopes in the compounds of bromine, chlorine, and carbon. Assuming the molecular ion of each of these compounds is of 100% abundance, what peaks (and in what intensity) would appear around that mass number? (a) C_2H_5BrO, (b) C_{60}, (c) C_6H_4BrCl? Give in cases (a) and (c) a possible structure for the compound. What compound is (b)?

2. The ^{13}C NMR spectrum for ethyl benzoate contains these peaks: 17.3, 61.1, 100–150 p.p.m. (four peaks), and 166.8 p.p.m. Which peak belongs to which carbon atom?


ethyl benzoate

3. The thinner used in typists’ correction fluids is a single compound, $C_2H_3Cl_3$, having ^{13}C NMR peaks at 45.1 and 95.0 p.p.m. What is its structure? A commercial paint thinner gives two spots on thin layer chromatography and has ^{13}C NMR peaks at 7.0, 27.5, 35.2, 45.3, 95.6, and 206.3 p.p.m. Suggest what compounds might be used to make up this thinner.

4. The ‘normal’ O–H stretch (i.e. without hydrogen bonding) comes at about 3600 cm^{-1}. What is the reduced mass (μ) for O–H? What happens to the reduced mass when you double the atomic weight of each atom in turn, that is, what is μ for O–D and what is μ for S–H? In fact, both O–D and S–H stretches come at about 2500 cm^{-1}.

5. Four compounds, each having the formula C_3H_5NO, have the IR spectra summarized here. What are their structures? Without ^{13}C NMR data, it may be easier to tackle this problem by first writing down all the possible structures for C_3H_5NO. In what specific ways would ^{13}C NMR data help?

(**a**) One sharp band above 3000 cm^{-1}; one strong band at about 1700 cm^{-1}

(**b**) Two sharp bands above 3000 cm^{-1}; two bands between 1600 and 1700 cm^{-1}

(**c**) One strong broad band above 3000 cm^{-1}; a band at about 2200 cm^{-1}

6. Four compounds having the molecular formula $C_4H_6O_2$ have the IR and ^{13}C NMR spectra given below. How many DBEs are there in $C_4H_6O_2$? What are the structures of the four compounds? You might again find it helpful to draw out some or all possibilities before you start.

(**a**) IR: 1745 cm^{-1}; ^{13}C NMR: 214, 82, 58, and 41 p.p.m.

(**b**) IR: 3300 (broad) cm^{-1}; ^{13}C NMR: 62 and 79 p.p.m.

(**c**) IR: 1770 cm^{-1}; ^{13}C NMR: 178, 86, 40, and 27 p.p.m.

(**d**) IR: 1720 and 1650 (strong) cm^{-1}; ^{13}C NMR: 165, 131, 133, and 54 p.p.m.

7. Three compounds of molecular formula C_4H_8O have the IR and ^{13}C NMR spectra given below. Suggest a structure for each compound, explaining how you make your deductions.

compound A IR: 1730 cm^{-1}; ^{13}C NMR: 13.3, 15.7, 45.7, and 201.6 p.p.m.

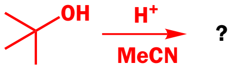
compound B IR: 3200 (broad) cm^{-1}; ^{13}C NMR: 36.9, 61.3, 117.2, and 134.7 p.p.m.

compound C IR: no peaks except CH and fingerprint; ^{13}C NMR: 25.8 and 67.9 p.p.m.

compound D IR: 3200 (broad) cm^{-1}; ^{13}C NMR: 15.2, 20.3, 36.0, and 62.9 p.p.m.

Compound A reacts with $NaBH_4$ to give compound D. Compound B reacts with hydrogen gas over a palladium catalyst to give the same compound D. Compound C reacts with neither reagent. Suggest a structure for compound D from the data given and explain the reactions. (*Note.* H_2 reduces alkenes to alkanes in the presence of a palladium catalyst.)

8. You have dissolved *t*-BuOH (Me_3COH) in MeCN with an acid catalyst, left the solution overnight, and found crystals with the following characteristics there in the morning. What are they?

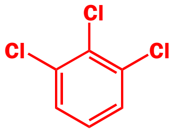


IR: 3435 and 1686 cm^{-1}

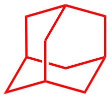
^{13}C NMR: 169, 50, 29, and 25 p.p.m.

mass spectrum (%): 115 (7), 100 (10), 64 (5), 60 (21), 59 (17), 58 (100), and 56 (7). (Don't try to assign all of these!)

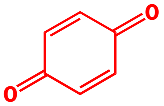
9. How many isomers of trichlorobenzene are there? The 1,2,3-trichloro isomer is illustrated. Could they be distinguished by ^{13}C NMR?



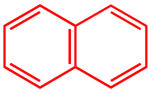
10. How many signals would you expect in the ^{13}C NMR of the following compounds?



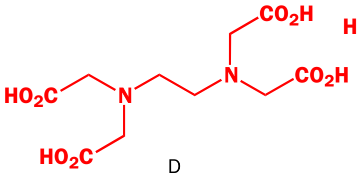
A



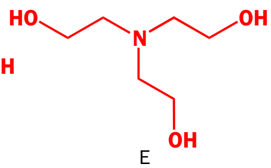
B



C

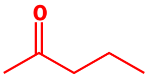


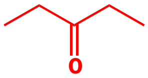
D

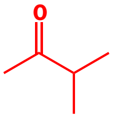


E

11. How would mass spectra help you distinguish these structures?







Structure of molecules

4

Connections

Building on:

- How organic structures are drawn ch2
- Evidence used to determine organic structure ch3

Arriving at:

- How we know that electrons have different energies
- How electrons fit into atomic orbitals
- How atomic orbitals combine to make molecular orbitals
- Why organic molecules have linear, planar, or tetrahedral structures
- Connection between shape and electronic structure
- A true system of molecular orbital energies for simple molecules
- Why such rigour is not possible for typical organic molecules
- Predicting the locations of lone pairs and empty orbitals
- Interaction between theory and experiment

Looking forward to:

- Mechanisms depend on molecular orbitals ch5
- Conjugation ch7
- ^{1}H NMR involves molecular orbitals ch11
- Reactivity derives from energies of molecular orbitals ch3

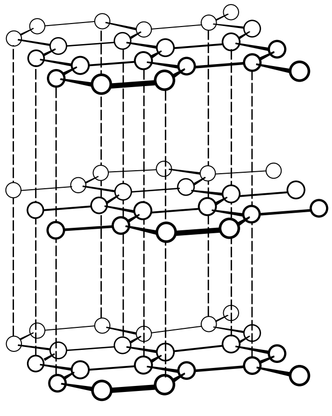
Note from the authors to all readers

This chapter contains mathematical material that some readers may find daunting. Organic chemistry students come from many different backgrounds since organic chemistry occupies a middle ground between the physical and the biological sciences. We hope that those from a more physical background will enjoy the material as it is. If you are one of those, you should work your way through the entire chapter. If you come from a more biological background, especially if you have done little maths at school, you may lose the essence of the chapter in a struggle to understand the equations. We have therefore picked out the more mathematical parts in boxes and you should abandon these parts (and any others!) if you find them too alien. The general principles behind the chapter—why molecules have the structures they do—are obviously so important that we cannot omit this essential material but you should try to grasp the principles without worrying too much about the equations. The ideas of atomic orbitals overlapping to form bonds, the molecular orbitals that result, and the shapes that these orbitals impose on organic molecules are at least as central for biochemistry as they are for organic chemistry. Please do not be discouraged but enjoy the challenge.

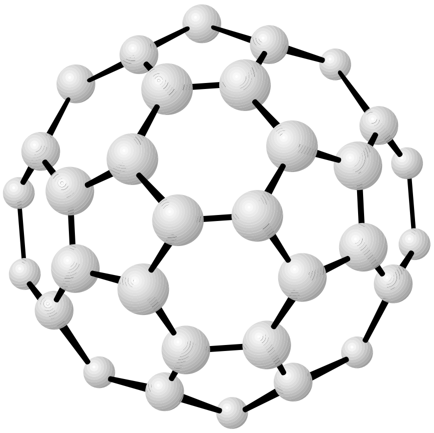
Introduction

You may recognize the model above as DNA, the molecule that carries the genetic information for all life on earth. It is the exact structure of this compound that determines precisely what a living thing is—be it man or woman, frog, or tree—and even more subtle characteristics such as what colour eyes or hair people have.

What about this model?

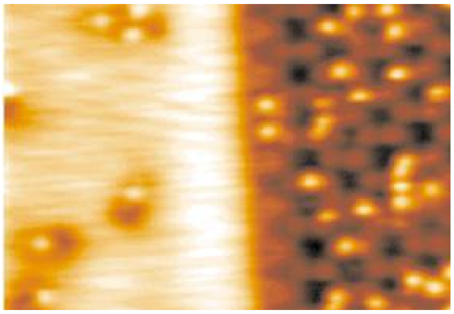


graphite



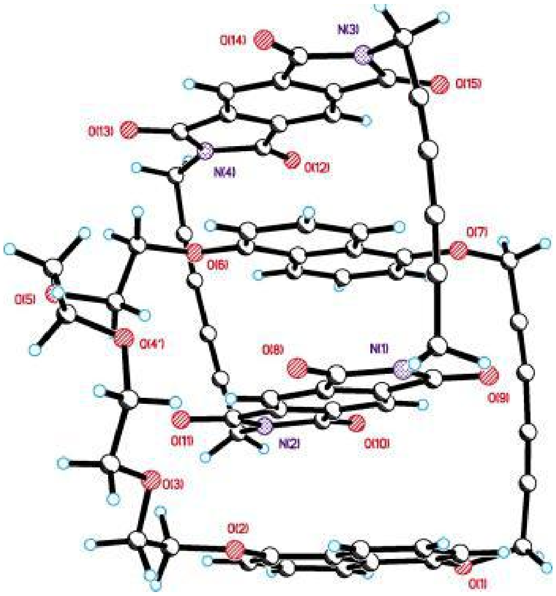
You may also have recognized this molecule as buckminsterfullerene, a form of carbon that received enormous interest in the 1980s and 1990s. The question is, how did you recognize these two compounds? You recognized their *shapes*. All molecules are simply groups of atoms held together by electrons to give a definite three-dimensional shape. What exactly a compound might be is determined not only by the atoms it contains, but also by the arrangement of these atoms in space—the shape of the molecule. Both graphite and buckminsterfullerene are composed of carbon atoms only and yet their properties, both chemical and physical, are completely different.

There are many methods available to chemists and physicists to find out the shapes of molecules. One of the most recent techniques is called **Scanning Tunnelling Microscopy** (STM), which is the closest we can get to actually 'seeing' the atoms themselves.



► The dark brown blobs in this STM picture recorded at a temperature of 4 K are individual oxygen atoms adsorbed on a silver surface. The light blobs are individual ethylene (ethene) molecules. Ethylene will only adsorb on silver if adjacent to an oxygen atom. This is an atomic scale view of a very important industrial process—the production of ethylene oxide from ethylene and oxygene using a silver catalyst.

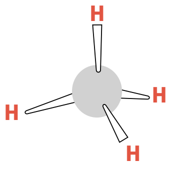
Most techniques, for example, X-ray or electron diffraction, reveal the shapes of molecules indirectly.



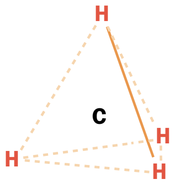
► The picture on the right is an X-ray structure of a catenane—a molecule consisting of two interlocking rings joined like two links in a chain. The key to the synthesis depends on the self-stacking of the planar structures prior to ring closure.

In Chapter 3 you met some of the spectroscopic methods frequently used by organic chemists to determine the shape of molecules. Spectroscopy would reveal the structure of methane, for example, as tetradral—the carbonatom in the centre of a regular tetra-hedron with the hydrogen atoms at the corners. In this chapter we are going to discuss *why* compounds adopt the shapes that they do.

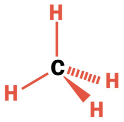
This tetrahedral structure seems to be very important—other molecules, both organic and inorganic, are made up of many tetrahedral units. What is the origin of this tetrahedral structure? It could simply arise from four pairs of electrons repelling each other to get as far as possible from each other. That would give a tetrahedron.



methane is tetrahedral



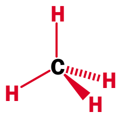
the H atoms form a tetrahedron



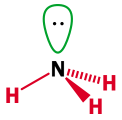
methane is tetrahedral

This simple method of deducing the structure of molecules is called **Valence Shell Electron Pair Repulsion Theory** (**VSEPRT**). It says that all electron pairs, both bonding and nonbonding, in the outer or valence shell of an atom repel each other. This simple approach predicts (more or less) the correct structures for methane, ammonia, and water with four electron pairs arranged tetrahedrally in each case.

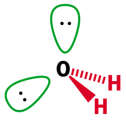
VSEPRT seems to work for simple structures but surely there must be more to it than this? Indeed there is. If we really want to understand *why* molecules adopt the shapes they do, we must look at the atoms that make up the molecules and how they combine. By the end of this chapter, you should be able to predict or at least understand the shapes of simple molecules. For example, why are the bond angles in ammonia 107°, while in hydrides of the other elements in the same group as nitrogen, PH_3, AsH_3, and SbH_3, they are all around 90°? Simple VSEPRT would suggest tetrahedral arrangements for each.



tetrahedral methane
four bonds and no lone pairs



tetrahedral ammonia
three bonds and one lone pair



tetrahedral water
two bonds and two lone pairs

Atomic structure

You know already what makes up an atom—protons, neutrons, and electrons. The protons and neutrons make up the central core of an atom—the nucleus—while the electrons form some sort of cloud around it. As chemists, we are concerned with the electrons in atoms and more importantly with the electrons in molecules: chemists need to know how many electrons there are in a system, where they are, and what energy they have. Before we can understand the behaviour of electrons in molecules, we need to look closely at the electronic structure of an atom. Evidence first, theory later.

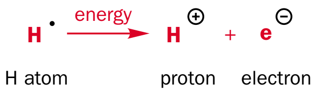
Atomic emission spectra

Many towns and streets are lit at night by sodium vapour lamps. You will be familiar with their warm yellow-orange glow but have you ever wondered what makes this light orange and not white? The normal light bulbs you use at home have a tungsten filament that is heated white hot. You know that this white light could be split by a prism to reveal the whole spectrum of visible light and that each of the different colours has a different frequency that corresponds to a distinct energy. But where does the orange street light come from? If we put a coloured filter in front of our white light, it would absorb some colours of the spectrum and let other colours through. We could make orange light this way but that is not how the street lights work—they actually generate orange light and orange light only. Inside these lights is sodium metal. When the light is switched on, the sodium metal is slowly vaporized and, as an electric current is passed through the sodium vapour, an orange light is emitted. This is the same colour as the light you get when you do a flame test using a sodium compound.

Quantum mechanics tells us that energy is quantized. Light does not come in a continuous range of energies but is divided up into minute discrete packets (quanta) of different noncontinuous (discrete) energies. The energy of each of these packets is related to the frequency of the light by a simple equation: $E = h\nu$ (E is the energy, ν the frequency of the light, and h is Planck's constant). The packet of light released from sodium atoms has the frequency of orange light and the corresponding energy.

The point is that only one colour light comes from a sodium lamp and this must have one specific frequency and therefore one energy. It doesn't matter what energy source is used to generate the light, whether it be electricity or a Bunsen burner flame; in each case light of one specific energy is given out. Looking at the orange sodium light through a prism, we see a series of very sharp lines with two particularly bright orange lines at around 600 nm. Other elements produce similar spectra—indeed two elements, rubidium and cesium, were discovered by Robert Bunsen after studying such spectra. They are actually named after the presence of a pair of bright coloured lines in their spectra—cesium from the Latin *caesius* meaning bluish grey and rubidium from the Latin *rubidus* meaning red. Even hydrogen can be made to produce an atomic spectrum and, since a hydrogen atom is the simplest atom of all, we shall look at the atomic spectrum of hydrogen first.

If enough energy is supplied to a hydrogen atom, or any other atom, an electron is eventually knocked completely out of the atom. In the case of hydrogen a single proton is left. This is, of course, the ionization of hydrogen.



What if we don't quite give the atom enough energy to remove an electron completely? It's not too hard to imagine that, if the energy is not enough to ionize the atom, the electron would be 'loosened' in some way—the atom absorbs this energy and the electron moves further away from the nucleus and now needs less energy to remove it completely. The atom is said to be in an **excited state**. This process is a bit like a weight lifter lifting a heavy weight—he can hold it above his head with straight arms (the excited state) but sooner or later he will drop it and the weight will fall to the ground. This is what happens in our excited atom—the electron will fall to its lowest energy, its **ground state**, and the energy put in will come out again. This is the origin of the lines in the atomic spectra not only for hydrogen but for all the elements. The flame or the electric discharge provides the energy to promote an electron to a higher energy level and, when this electron returns to its ground state, this energy is released in the form of light.

Line spectra are composed of many lines of different frequencies, which can only mean that there must be lots of different energy transitions possible, but not just *any* energy transitions. Quantum mechanics says that an electron, like light, cannot have a continuous range of energies, only certain definite energies, which in turn means that only certain energy transitions are possible. This is rather like trying to climb a flight of stairs—you can jump up one, two, five, or even all the steps if you have enough energy but you cannot climb up half or two-thirds of a step. Likewise coming down, you can jump from one step to any other—lots of different combinations are possible *but there is a finite number, depending on the number of steps*. This is why there are so many lines in the atomic spectra—the electron can receive energy to promote it to a higher energy level and it can then fall to any level below and a certain quantity of light will be released.

We want to predict, as far as we can, where all the electrons in different *molecules* are to be found including the ones not involved with bonding. We want to know where the molecule can accommodate *extra* electrons and from where electrons can be removed most easily. Since most molecules contain many electrons, the task is not an easy one. However, the electronic structure of atoms is somewhat easier to understand and we can approximate the electronic structure of *molecules* by considering how the component atoms combine.

The next section is therefore an introduction to the electronic structure of *atoms*—what energies the electrons have and where they may be found. Organic chemists are rarely concerned with atoms themselves but need to understand the electronic structure in atoms before they can understand the electronic structure in molecules. As always, evidence first!

The atomic emission spectrum of hydrogen

The atomic emission spectrum of hydrogen is composed of many lines but these fall into separate sets or series. The first series to be discovered, not surprisingly, were those lines in the visible part of the spectrum. In 1885, a Swiss schoolmaster, Johann Balmer, noticed that the wavelengths, λ, of the lines in this series could be predicted using a mathematical formula. He did not see why; he just saw the relationship. This was the first vital step.

$$\lambda = \text{constant} \times \frac{n^2}{n^2 - 2^2} \quad (n \text{ is an integer greater than } 2)$$

As a result of his work, the lines in the visible spectrum are known as the Balmer series. The other series of lines in the atomic emission spectrum of hydrogen were discovered later (the next wasn't discovered until 1908). These series are named after the scientists who discovered them; for example, the series in the ultraviolet region is known as the Lyman series after Theodore Lyman.

Balmer's equation was subsequently refined to give an equation that predicts the frequency, ν, of any of the lines in any part of the hydrogen spectrum rather than just for his series. It turns out that his was not the most fundamental series, just the first to be discovered.

$$\nu = \text{constant} \times \left(\frac{1}{n_1^2} - \frac{1}{n_2^2} \right)$$

Each series can be described by this equation if a particular value is given for n_1 but n_2 is allowed to vary. For the Lyman series, n_1 remains fixed at 1 while n_2 can be 2, 3, 4, and so on. For the Balmer series, n_1 is fixed at 2 while n_2 can be 3, 4, 5, and so on.

Atomic emission spectra are evidence for electronic energy levels

Atomic emission spectra give us our first clue to understanding the electronic energy levels in an atom. Since the lines in the emission spectrum of hydrogen correspond to the electron moving between energy levels and since frequency is proportional to energy, $E = h\nu$, the early equations must represent just the difference between two energy levels. This in turn tells us that the electron's energy levels in an atom must be inversely proportional to the square of an important integer 'n'. This can be expressed by the formula

$$E_n = -\frac{\text{constant}}{n^2}$$

where E_n is the energy of an electron in the nth energy level and n is an integer ≥ 1 known as the **principal quantum number**. Note that, when $n = \infty$, that is, when the electron is no longer associated with the nucleus, its energy is zero. All other energy levels are lower than zero because of the minus sign in the equation. This is consistent with what we know already—we must put energy in to ionize the atom and remove the electron from the nucleus.

Electronic energy levels

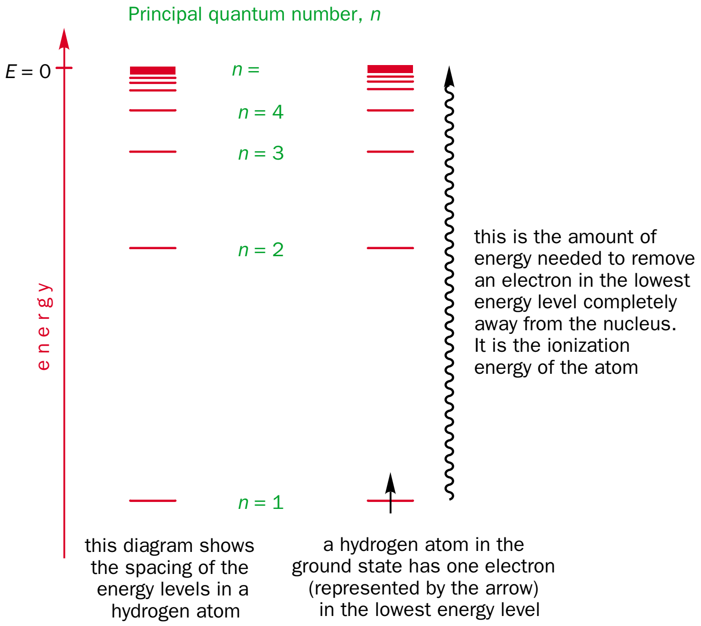
In more detail, the constant in this equation can be broken down into a universal constant, the Rydberg constant R_H, which applies to any electron on any atom, and a constant Z which has a particular value for each atom.

$$E_n = -\frac{R_H Z^2}{n^2}$$

The Rydberg constant R_H, is measured in units of energy. For a given atom (i.e. Z is constant) there are many different energy levels possible (each corresponding to a different value of n). Also, as n gets bigger, the energy gets smaller and smaller and approaches zero for large n. The energy gets smaller as the electron gets further away from the nucleus.For electrons in the same energy level but in different atoms, (i.e. keeping n constant but varying Z), the energy of an electron depends on the square of the atomic number. This makes sense too—the more protons in the nucleus, the more tightly the electron is held in the atom.

- **The electrons in any atom are grouped in energy levels whose energies are universally proportional to the inverse square of a very important number n. This number is called the principal quantum number and it can have only a few integral values ($n = 1, 2, 3\ldots$). The energy levels also depend on the type of atom.**

An energy level diagram gives some idea of the relative spacing between these energy levels.



▶ Notice how the spacing between the energy levels gets closer and closer. This is a consequence of the energy being inversely proportional to the square of the principal quantum number. It tells us that it becomes easier and easier to remove an electron completely from an atom as the electron is located in higher and higher energy levels. As we shall see later, the increasing value of the principal quantum number also correlates with the electron being found (on average) further and further from the nucleus and being easier and easier to remove. This is analogous to a rocket escaping from a planet—the further away it is, the less it experiences the effects of gravity and so the less energy it requires to move still further away. The main difference is that there seems to be no quantization of the different energy levels of the rocket—it appears (to us in our macroscopic world at least) that any energy is possible. In the case of the electron in the atom, only certain values are allowed.

Three quantum numbers come from the Schrödinger equation

There is no doubt about the importance of n, the principal quantum number, but where does it come from? This quantum number and two other quantum numbers come from solving the **Schrödinger equation**. We are not going to go into any details regarding Schrödinger's equation or how to solve it—there are plenty of more specialized texts available if you are interested in more detail.

Solutions to Schrödinger's equation come in the form of **wave functions** (symbol Ψ), which describe the energy and position of the electrons thought of as waves. You might be a little unsettled to find out that we are describing electrons using waves but the same wave–particle duality idea applies to electrons as to light. We regularly think of light in terms of waves with their associated wavelengths and frequencies but light can also be described using the idea of photons—individual little light 'particles'. The same is true of the electron; up to now, you will probably have thought of electrons only as particles but now we will be thinking of them as waves.

■ Don't worry about the rather fancy names of these quantum numbers; just accept that the three numbers define a given wave function.

It turns out that there is not one specific solution to the Schrödinger equation but many. This is good news because the electron in a hydrogen atom can indeed have a number of different energies. It turns out that each wave function can be defined by three quantum numbers (there is also a fourth quantum number but this is not needed to define the wave function). We have already met the principal quantum number, n. The other two are called the **orbital angular momentum quantum number** (sometimes called the azimuthal quantum number), ℓ, and the **magnetic quantum number**, $m\ell$.

ℓ

A specific wave function solution is called an **orbital**. The different orbitals define different energies and distributions for the different electrons. The name 'orbital' goes back to earlier theories where the electron was thought to orbit the nucleus in the way that planets orbit the sun. It seems to apply more to an electron seen as a particle, and orbitals of electrons thought of as particles and wave functions of electrons thought of as waves are really two different ways of looking at the same thing. Each different orbital has its own individual quantum numbers, n, ℓ, and m_ℓ.

Summary of the importance of the quantum numbers

What does each quantum number tell us and what values can it adopt? You have already met the principal quantum number, n, and seen that this is related to the energy of the orbital.

The principal quantum number, n

Different values for n divide orbitals into groups of similar energies called **shells**. Numerical values for n are used in ordinary speech. The first shell ($n = 1$) can contain only two electrons and the atoms H and He have one and two electrons in this first shell, respectively.

The orbital angular momentum quantum number, ℓ

The orbital angular momentum quantum number, ℓ, determines, as you might guess, the angular momentum of the electron as it moves in its orbital. This quantum number tells us the shape of the orbital, spherical or whatever. The values that ℓ can take depend on the value of n: ℓ can have any value from 0 up to $n - 1$: $\ell = 0, 1, 2, \ldots, n - 1$. The different possible values of ℓ are given letters rather than numbers and they are called s, p, d, and f.

value of n	1	2	3	4
possible values of ℓ	0	0, 1	0, 1, 2	0, 1, 2, 3
name	1s	2s, 2p	3s, 3p, 3d	4s, 4p, 4d, 4f

The magnetic quantum number, m_ℓ

The magnetic quantum number, $m\ell$, determines the spatial orientation of the angular momentum. In simple language it determines where the orbitals are in space. Its value depends on the value of ℓ, varying from $-\ell$ to $+l$: $m_\ell = \ell, \ell - 1, \ell - 2, \ldots, -\ell$. The different possible values of m_l are given suffixes on the letters

value of n	1	2	2
value of ℓ	0	0	1
name	1s	2s	2p
possible values of $m\ell$	0	0	+1, 0, –1
name	1s	2s	$2p_x$, $2p_y$, $2p_z$

defining the quantum number ℓ. These letters refer to the direction of the orbitals along the *x*-, *y*-, or *z*-axes. Organic chemists are concerned mostly with s and p orbitals ($\ell = 0$ or 1) so the subdivisions of the d orbitals can be omitted.

Each quantum number gives subdivisions for the one before. There are no subdivisions in the lowest value of each quantum number: and the subdivisions increase in number as each quantum number increases. Now we need to look in more detail at the meanings of the various values of the quantum numbers.

Atomic orbitals

Nomenclature of the orbitals

For a hydrogen atom the energy of the orbital is determined only by the principal quantum number, *n*, and *n* can take values 1, 2, 3, and so on. This is the most fundamental division and is stated first in the description of an electron. The electron in a hydrogen atom is called $1s^1$. The 1 gives the value of *n*: the most important thing in the foremost place. The designation s refers to the value of ℓ. These two together, 1s, define and name the orbital. The superscript 1 tells us that there is one electron in this orbital.

The orbital angular momentum quantum number, ℓ, determines the shape of the orbital. Instead of expressing this as a number, letters are used to label the different shapes of orbitals. s orbitals have $\ell = 0$, and p orbitals have $\ell = 1$.

Using both these quantum numbers we can label orbitals 1s, 2s, 2p, 3s, 3p, 3d, and so on. Notice that, since *l* can only have integer values up to $n - 1$, we cannot have a 1p or 2d orbital.

s, p, d, f

These letters hark back to the early days of spectroscopy and refer to the appearance of certain lines in atomic emission spectra: 's' for 'sharp', 'p' for 'principal', 'd' for 'diffuse', and 'f' for 'fundamental'. The letters s, p, d, and f matter and you must know them, but you do not need to know what they originally stood for.

Value of *l*	Name of orbital
0	s
1	p
2	d
3	f

● Names of atomic orbitals

- **The first shell ($n = 1$) has only an s orbital, 1s**
- **The second shell ($n = 2$) has s and p orbitals 2s and 2p**
- **The third shell ($n = 3$) has s, p, and d orbitals, 3s, 3p, and 3d**

One other point to notice is that, for the hydrogen atom (and, technically speaking, any one-electron ion such as He^+ or Li^{2+}), a 2s orbital has exactly the same energy as a 2p orbital and a 3s orbital has the same energy as the 3p and 3d orbitals. Orbitals that have the same energy are described as **degenerate**. In atoms with more than one electron, things get more complicated because of electron–electron repulsion and the energy levels are no longer determined by *n* alone. In such cases, the 2s and 2p or the 3s, 3p, and 3d orbitals or any other orbitals that share the same principal quantum number are no longer degenerate. In other words, in multielectron atoms, the energy of a given orbital depends not only on the principal quantum number, *n*, but also in some way on the orbital angular momentum quantum number, ℓ.

Values of the magnetic quantum number, m_l, depend on the value of ℓ. When $\ell = 0$, m_ℓ can only take one value (0); when $\ell = 1$, m_l has three possible values (+1, 0, or –1). There are five possible values of m_l when $\ell = 2$ and seven when $\ell = 3$. In more familiar terms, there is only one sort of s orbital; there are three sorts of p orbitals, five sorts of d orbitals, and seven sorts of f orbitals. All three p orbitals are degenerate as are all five d orbitals and all seven f orbitals (for both single-electron and multielectron atoms). We shall see how to represent these orbitals later.

There is a fourth quantum number

The spin of an electron is the angular momentum of an electron spinning *about its own axis*, although this is a simplified picture. This angular momentum is different from the angular momentum, ℓ, which represents the electron's angular momentum about the nucleus. The magnitude of the electron's spin is constant but it can take two orientations. These are represented using the fourth quantum number, the **spin angular momentum quantum number**, m_s, which can take the value of +1 or –1 in any orbital, regardless of the values of *n*, ℓ, or m_ℓ. Each

■ You have already come across another spin—the nuclear spin—which gives rise to NMR. There is an analogous technique, **electron spin resonance**, ESR, which detects unpaired electrons.

orbital can hold a maximum of two electrons and then only when the electrons have different 'spin', that is, they must have different values of m_s, +1 or –1. The rule that no more than two electrons may occupy any orbital (and then only if their spins are paired) is known as the **Pauli exclusion principle**.

● Every electron is unique!

If electrons are in the same atom, they must have a unique combination of the four quantum numbers. Each orbital, designated by three quantum numbers, n, ℓ, and m_ℓ, can contain only two electrons and then only if their spin angular quantum numbers are different.

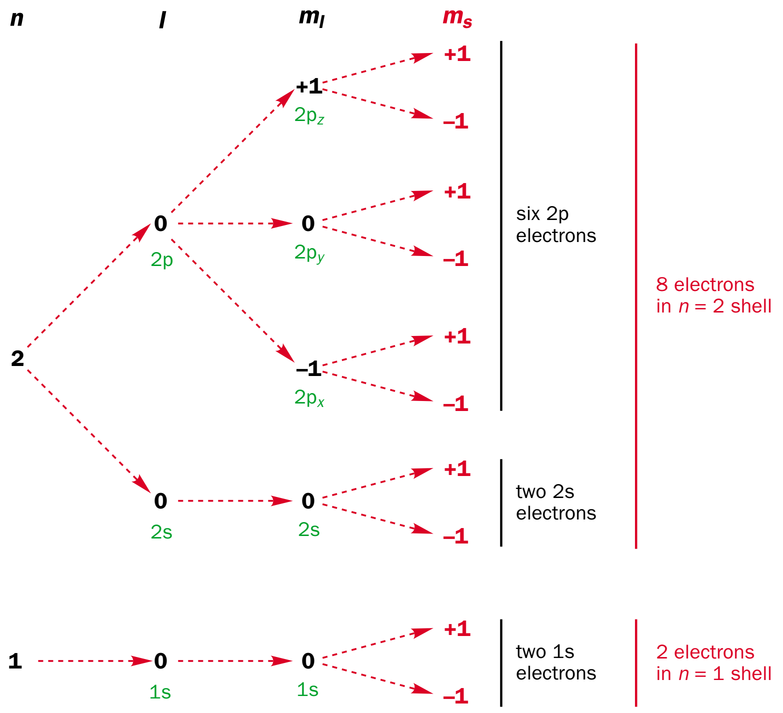
How the periodic table is constructed

All the quantum numbers for all the electrons with $n = 1$ and 2 can now be shown in a table like the ones earlier in this chapter. Though we have so far been discussing the hydrogen atom, in fact, the H atom never has more than two electrons. Fortunately, the energy levels deduced for H also apply to all the other elements with some minor adjustments. This table would actually give the electronic configuration of neon, Ne.

In this table, the energy goes up from left to right, though all the 2p orbitals are degenerate. To add $n = 3$, one column for the 3s, three columns for the 3p, and five columns for the 3d orbital would be needed. Then all five 3d orbitals would be degenerate.

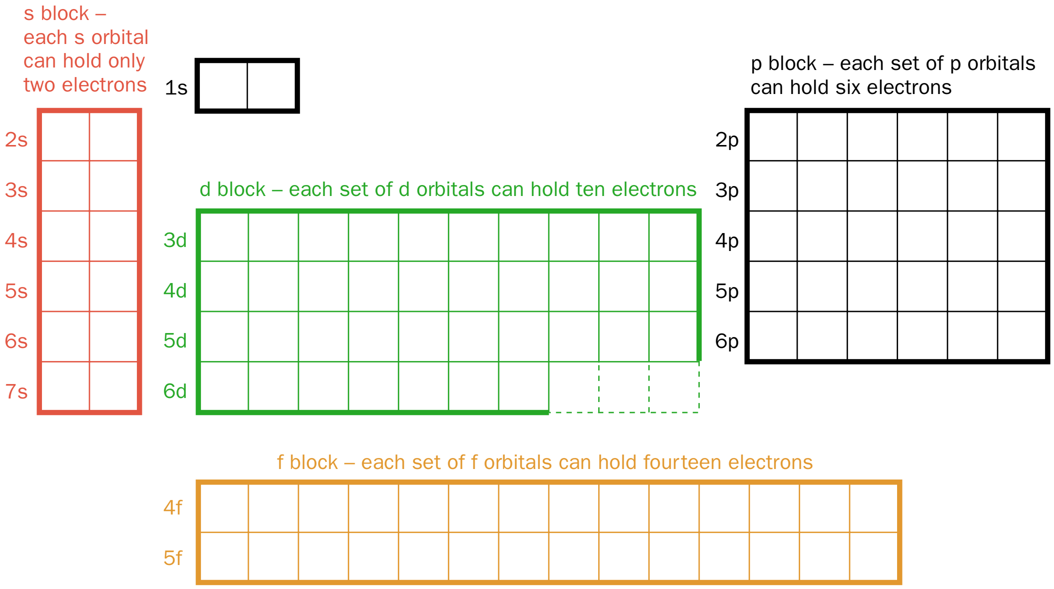
value of n	1	2	2	2	2
value of ℓ	0	0	1	1	1
name	1s	2s	2p	2p	2p
possible values of m_ℓ	0	0	+1	0	–1
name	1s	2s	$2p_x$	$2p_y$	$2p_z$
possible values of m_s	+1, –1	+1, –1	+1, –1	+1, –1	+1, –1
electrons	$1s^2$	$2s^2$	$2p_x^2$	$2p_y^2$	$2p_z^2$

Another way to show the same thing is by an energy diagram showing how the quantum numbers divide and subdivide.



These numbers explain the shape of the periodic table. Each element has one more electron (and one more proton and perhaps more neutrons) than the one before. At first the lowest energy shell ($n = 1$) is filled. There is only one orbital, 1s, and we can put one or two electrons in it. There are therefore two elements in this block, H and He. Next we must move to the second shell ($n = 2$), filling 2s first so we start the top of groups 1 and 2 with Li and Be. These occupy the top of the red stack marked ‘s block’ because all the elements in this block have one or two electrons in their outermost s orbital and no electrons in the outermost p orbital. Then we can start on the 2p orbitals. There are three of these so we can put in six electrons and get six elements B, C, N, O, F, and Ne. They occupy the top row of the black p block. Most of the elements we need in this book are in those blocks. Some, Na, K, and Mg for example, are in the s block and others, Si, P, and S for example, are in the second row of the p block.

The layout of the periodic table

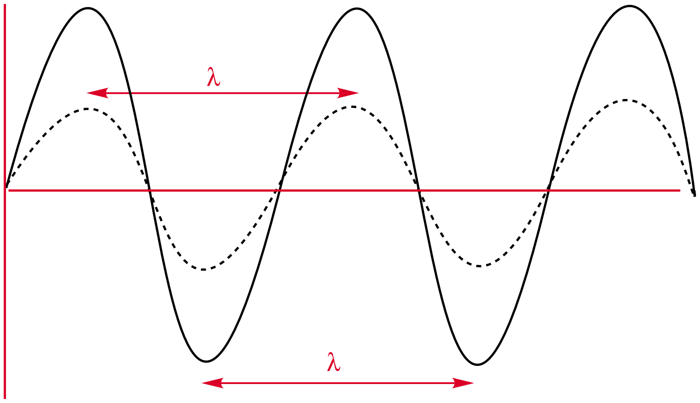


Other orbitals

Organic chemists are really concerned only with s and p orbitals since most of the elements we deal with are in the second row of the periodic table. Later in the book we shall meet elements in the second row of the p block (Si, P, S) and then we will have to consider their d orbitals, but for now we are not going to bother with these and certainly not with the f orbitals. But you may have noticed that the 4s orbital is filled before the 3d orbitals so you may guess that the 4s orbital must be slightly lower in energy than the 3d orbitals. Systems with many electrons are more complicated because of electron repulsion and hence the energies of their orbitals do not simply depend on *n* alone.

Graphical representations of orbitals

One problem with wave functions is trying to visualize them: what does a wave function look like? Various graphs of wave functions can be plotted but they are not much help as Ψ itself has no physical meaning. However the *square* of the wave function, Ψ^2, does have a practical interpretation; it is proportional to the probability of finding an electron at a given point. Unfortunately, we can’t do



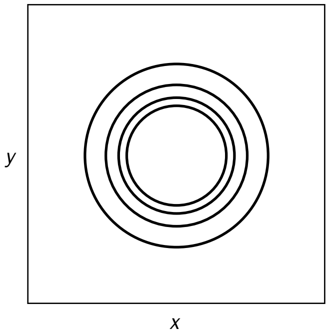
These two waves both have the same wavelength, λ, but the dashed wave is less intense than the other wave. The intensity is proportional to the amplitude squared.

▶

There is some justification for this interpretation that the wave function squared is proportional to the probability of finding an electron. With light waves, for example, while the wavelength provides the colour (more precisely the energy) of the wave, it is the amplitude squared that gives the brightness.

But this is looking at light in terms of waves. In terms of particles, photons, the intensity of light is proportional to the density of photons.

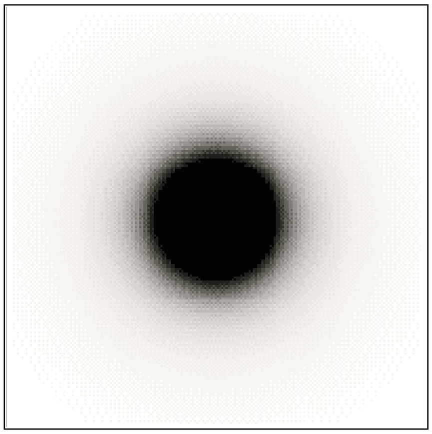
better than probability as we are unable to say exactly where the electron is at any time. This is a consequence of **Heisenberg's uncertainty principle**—we cannot know both the exact position and the exact momentum of an electron simultaneously. Here we know the momentum (energy) of the electron and so its exact position is uncertain.



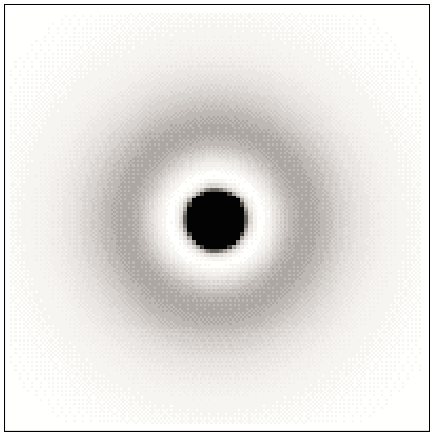
contour diagram of 1s orbital

How do we depict a probability function? One way would be to draw contours connecting regions where there is an equal probability of finding the electron. If Ψ^2 for a 1s orbital is plotted, a three-dimensional plot emerges. Of course, this is a two-dimensional representation of a three-dimensional plot—the contours are really spherical like the different layers of an onion. These circles are rather like the contour lines on a map except that they represent areas of equal probability of finding the electron instead of areas of equal altitude.

Another way to represent the probability is by a **density plot**. Suppose we could see exactly where the electron was at a given time and that we marked the spot. If we looked again a little later, the electron would be in a different place—let us mark this spot too. Eventually, if we marked enough spots, we would end up with a fuzzy picture like those shown for the 1s and 2s orbitals. Now the *density* of the dots is an indication of the probability of finding an electron in a given space—the more densely packed the dots (that is, the darker the area), the greater the probability of finding the electron in this area. This is rather like some maps where different altitudes are indicated by different colours.



density plot of 1s orbital



density plot of 2s orbital

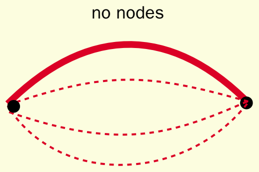
The 2s orbital, like the 1s orbital, is spherical. There are two differences between these orbitals. One is that the 2s orbital is bigger so that an electron in a 2s orbital is more likely to be found further away from the nucleus than an electron in a 1s orbital. The other difference between the orbitals is that, within the 2s orbital but not within the 1s orbital, there is a region where there is no electron density at all. Such a region is called a **nodal surface.** In this case there is no electron density at one set radius from the nucleus; hence this is known as a **radial node.** The 2s orbital has one radial node.

Nodes are important for musicians

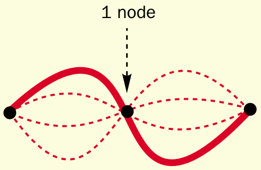
You can understand these nodal surfaces by thinking in terms of waves. If a violin or other string instrument is plucked, the string vibrates. The ends cannot move since they are fixed to the instrument. The note we hear is mainly due to the string vibrating as shown in the diagram for the first harmonic.

However, there are other vibrations of higher energy known as **harmonics**, which help to give the note its timbre (the different timbres allow us to tell the difference between, say, a flute and a violin playing the same note). The second and third harmonics are also shown.

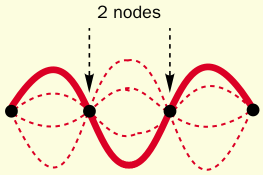
Each successive harmonic has one extra node—while the first harmonic has no nodes (if you don't count the end stops), the second harmonic has one and the third has two and so on. These are points where the string does not vibrate at all (you can even 'select out' the second harmonic on a stringed instrument if you gently press halfway along the vibrating string).



1st harmonic



2nd harmonic

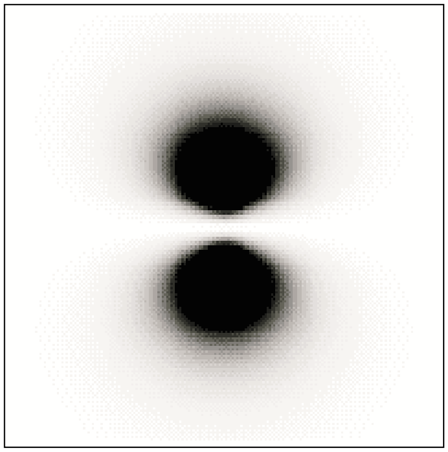


3rd harmonic

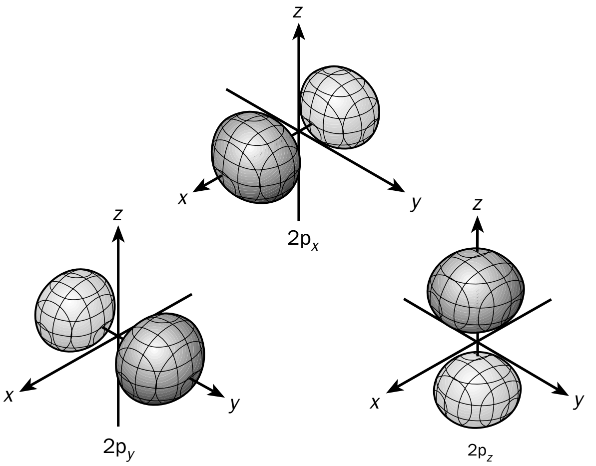
● Shapes of s orbitals

- The 1s orbital is spherically symmetrical and has no nodes
- The 2s orbital has one radial node and the 3s orbital two radial nodes. They are both spherically symmetrical

What does a Ψ^2 for a 2p orbital look like? The probability density plot is no longer spherically symmetrical. This time the shape is completely different—the orbital now has an orientation in space and it has two lobes. Notice also that there is a region where there is no electron density between the two lobes—another nodal surface. This time the node is a plane in between the two lobes and so it is known as a **nodal plane**. One representation of the 2p orbitals is a three-dimensional plot, which gives a clear idea of the true shape of the orbital.



density plot of 2p orbital

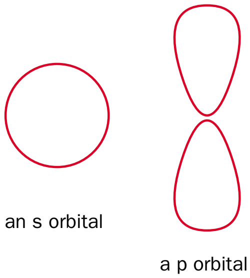


three-dimensional plot of the 2p orbitals

▶ You might have noticed that each orbital in the *n*th energy level has the same total number of nodes, $n - 1$. The total number of nodes is the sum of the numbers of radial nodes and nodal planes. Thus both the 2s and 2p have one node (a radial node in the case of the 2s and a nodal plane in the case of the 2p) while the 3s, 3p, and 3d orbitals each have two nodes (the 3s have two radial nodes, the 3p orbitals each have one radial node and one nodal plane, and the 3d orbitals each have two nodal planes).

Plots of 3p and 4p orbitals are similar—each has a nodal plane and the overall shape outlined in each is the same. However, the 3p orbital also has a radial node and the 4p has two radial nodes and once again the size of the orbital increases as the principal quantum number increases.

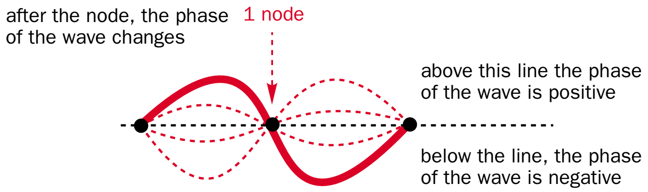
All this explains why the shape of an orbital depends on the orbital angular quantum number, ℓ. All s orbitals ($\ell = 0$) are spherical, all p orbitals ($\ell = 1$) are shaped like a figure eight, and d orbitals ($\ell = 2$) are yet another different shape. The problem is that these probability density plots take a long time to draw—organic chemists need a simple easy way to represent orbitals. The contour diagrams were easier to draw but even they were a little tedious. Even simpler still is to draw just one contour within which there is, say, a 90% chance of finding the electron. This means that all s orbitals can be represented by a circle, and all p orbitals by a pair of lobes.



■ Remember that the orbitals are three-dimensional and that these drawings represent a cross-section. A three-dimensional version would look more like a sphere for an s orbital and an old-fashioned hour-glass for a p orbital. Actually, each lobe of a p orbital is much more rounded than the usual representation, but that is not so important.

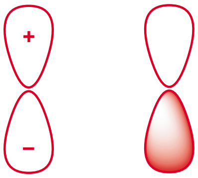
The phase of an orbital

The wave diagrams need further discussion to establish one fine point—the **phase** of an orbital.



Just as an electromagnetic wave, or the wave on a vibrating string, or even an ocean wave possesses different 'phases' (for example, the troughs and peaks of an ocean wave) so too do the atom's wave functions—the orbitals. After each node in an orbital, the phase of the wave function changes. In the

2p orbital, for example, one lobe is one phase; the other lobe is another phase with the nodal plane in between. In the standing wave above the different phases are labelled positive and negative. The phases of a p orbital could be labelled in the same way (and you may sometimes see this) but, since chemists use positive and negative signs to mean specific charges, this could get confusing. Instead, one half of the p orbital is usually shaded to show that it has a different phase from the other half.

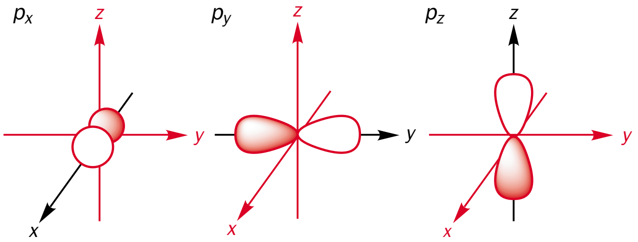


here the different phases of the p orbital are labelled positive and negative – this can be confusing and so is best avoided

here the different phases of the p orbital are shown by shading one half and not the other

The magnetic quantum number, m_ℓ

The magnetic quantum number, m_l, determines the spatial orientation of the orbital's angular momentum and takes the values $-\ell$ to $+\ell$. An s orbital ($\ell = 0$), being spherical, can only have one orientation in space—it does not point in any one direction and hence it only has one value for $m_\ell(0)$. However, a p orbital could point in any direction. For a p orbital ($\ell = 1$) there are three values of m_l: –1, 0, and +1. These correspond to the p orbitals aligned along the mutually perpendicular x-, y-, and z-axes. These orbitals, designated p_x, p_y, and p_z, are all degenerate. They differ only in their spatial orientations.



the three degenerate p orbitals are aligned along perpendicular axes

Summary so far

- Electrons in atoms are best described as waves
- All the information about the wave (and hence about the electron) is in the wave function, Ψ, the solution to the Schrödinger equation
- There are many possible solutions to the Schrödinger equation but each wave function (also called an orbital) can be described using three quantum numbers
- The principal quantum number, n, is largely responsible for the energy of the orbital (in one-electron systems, such as the hydrogen atom, it alone determines the energy). It takes integer values 1, 2, 3, 4, and so on, corresponding to the first, second, third, and so on shells of electrons
- The orbital angular momentum quantum number, ℓ, determines the angular momentum that arises from the motion of an electron moving in the orbital. Its value depends on the value of n and it takes integer values $0, \ldots, n-1$ but the orbitals are usually known by letters (s when $\ell = 0$, p when $\ell = 1$, d when $\ell = 2$, and f when $\ell = 3$). Orbitals with different values of ℓ have different shapes—s orbitals are spherical, p orbitals are shaped like a figure of eight
- The magnetic quantum number, m_ℓ, determines the spatial orientation of the orbital. Its value depends on the value of ℓ and it can take the integer values: $-\ell, \ldots .\ 0, \ldots .+\ell$. This means that there is only one type of s orbital, three different p orbitals (all mutually perpendicular), five different d orbitals, and seven different f orbitals. The three different p orbitals are all degenerate, that is, they have the same energy (as do the five d orbitals and the seven f orbitals)

- There is also a fourth quantum number, the spin angular momentum quantum number, m_s, which can take values of +1 or –1. The spin is not a property of orbitals but of the electrons that we put in the orbitals
- No two electrons in any one atom can have all four quantum numbers the same—this means that each orbital as described by the (first) three quantum numbers can hold a maximum of two electrons and then only if they have opposing spins
- We usually use a shorthand notation to describe an orbital such as 1s or $2p_y$

the number tells us the principal quantum number, *n*

$2s$ $3p_x$ $4p_y$

the letter tells us the orbital angular momentum quantum number, *l*

the subscript letter tells us the magnetic quantum numer, m_l

these three quantum numbers, *n*, *l*, and m_l, define an orbital

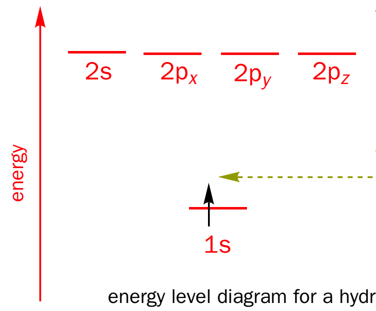
A few points are worth emphasis. Orbitals do not need to have electrons in them—they can be vacant (there doesn't have to be someone standing on a stair for it to exist!). So far we have mainly been talking about the hydrogen atom and this has only one electron. Most of the time this electron is in the 1s orbital (the orbital lowest in energy) but if we give it enough energy we can promote it to a vacant orbital higher in energy, say, for example, the $3p_x$ orbital.

Another point is that the electrons may be found anywhere in an orbital except in a node. In a p orbital containing one electron, this electron may be found on either side but never in the middle. When the orbital contains two electrons, one electron doesn't stay in one half and the other electron in the other half—both electrons could be anywhere (except in the node).

Finally, remember that all these orbitals are superimposed on each other. The 1s orbital is *not* the middle part of the 2s orbital. The 1s and 2s orbitals are separate orbitals in their own rights and each can hold a maximum of two electrons but the 2s orbital does occupy some of the same space as the 1s orbital (and also as the 2p orbitals, come to that). Neon, for example, has ten electrons in total: two will be in the 1s orbital, two in the 2s orbital, and two in each of the 2p orbitals. All these orbitals are superimposed on each other but the pairs of electrons are restricted to their individual orbitals. If we tried to draw all these orbitals, superimposed on each other as they are, in the same diagram the result would be a mess!

Putting electrons in orbitals

Working out where the electrons are in any atom, that is, which orbitals are populated, is easy. We simply put two electrons into the lowest energy orbital and work upwards. This 'building up' of the different atoms by putting electrons in the orbitals until they are full and then filling up the orbital next lowest in energy is known as the **Aufbau principle** (*Aufbau* is German for 'building up'). The first and only electron in the hydrogen atom must go into the 1s orbital. In this sort of diagram the energy levels are represented as horizontal lines stacked roughly in order with the lowest energy at the bottom. Electrons are represented as vertical arrows. Arrows pointing upwards show one spin (m_s = +1 or –1) and arrows pointing downwards the other (which is which doesn't matter).

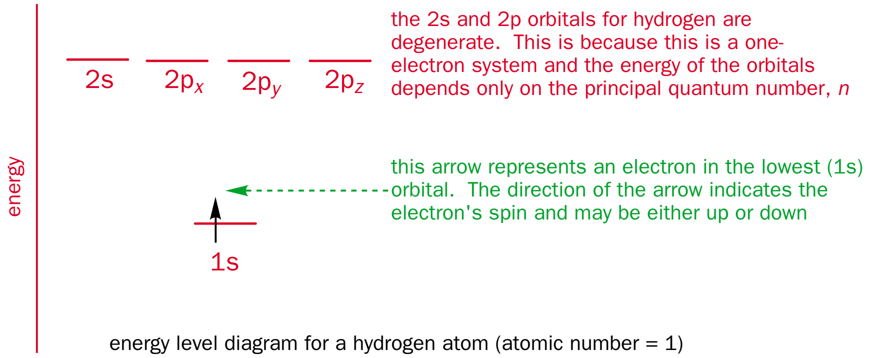


the 2s and 2p orbitals for hydrogen are degenerate. This is because this is a one-electron system and the energy of the orbitals depends only on the principal quantum number, *n*

this arrow represents an electron in the lowest (1s) orbital. The direction of the arrow indicates the electron's spin and may be either up or down

energy level diagram for a hydrogen atom (atomic number = 1)

The helium atom has two electrons and they can both fit into the 1s orbital providing they have opposite spins. The other change to the diagram is that, with two electrons and electron repulsion a factor, the 2s orbital is now lower in energy than the three 2p orbitals, though these three are still degenerate.

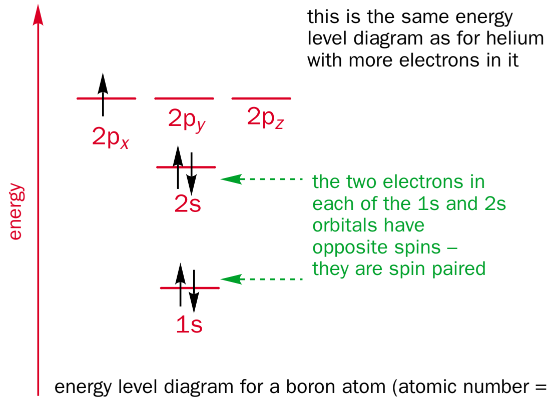


the 2s and 2p orbitals for hydrogen are degenerate. This is because this is a one-electron system and the energy of the orbitals depends only on the principal quantum number, *n*

this arrow represents an electron in the lowest (1s) orbital. The direction of the arrow indicates the electron's spin and may be either up or down

energy level diagram for a hydrogen atom (atomic number = 1)

Lithium has one more electron but the 1s orbital is already full. The third electron must go into the next lowest orbital and that is the 2s. In this three-electron system, like that of the two-electron He atom, the three 2p orbitals are higher in energy than the 2s orbital. By the time we come to boron, with five electrons, the 2s is full as well and we must put the last electron into a 2p orbital. It doesn't matter which one; they are degenerate.



this is the same energy level diagram as for helium with more electrons in it

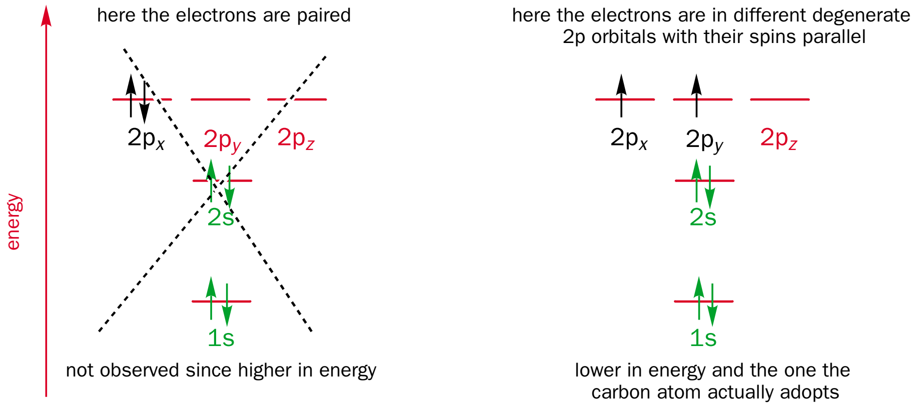
the two electrons in each of the 1s and 2s orbitals have opposite spins – they are spin paired

energy level diagram for a boron atom (atomic number = 5)

Carbon has one more electron than boron but now there is a bit of a problem—where does the last electron go? It could either be paired with the electron already in one of the p orbitals or it could go into one of the other degenerate p orbitals. It turns out that the system is lower in energy (electron–electron repulsion is minimized) if the electrons are placed in different degenerate orbitals with their spins parallel (that is, both spins +1 or both –1). Another way of looking at this is that putting two electrons into the same orbital with their spins paired (that is, one +1, one –1) requires some extra amount of energy, sometimes called **pairing energy**.

▶

This is known as **Hund's rule**. An atom adopts the electronic configuration that has the greatest number of unpaired electrons in degenerate orbitals. Whilst this is all a bit theoretical in that isolated atoms are not found very often, the same rule applies for electrons in degenerate orbitals in molecules.



here the electrons are paired

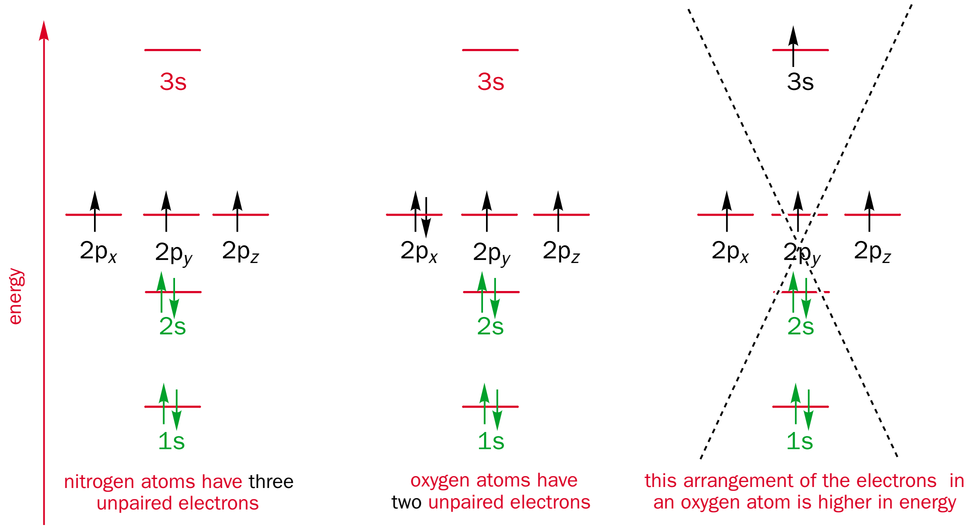
here the electrons are in different degenerate 2p orbitals with their spins parallel

not observed since higher in energy

lower in energy and the one the carbon atom actually adopts

the two possible arrangements for the electrons in a carbon atom

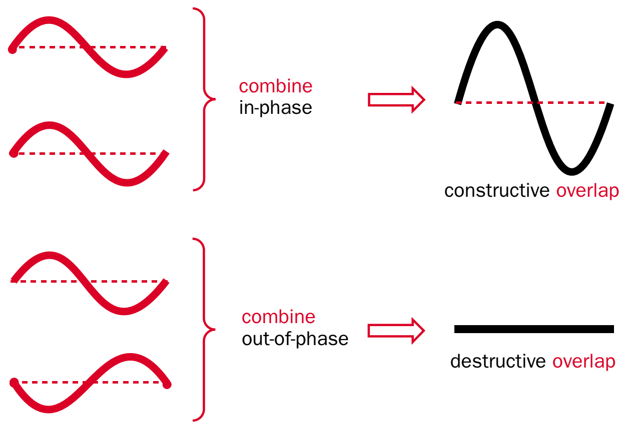
Nitrogen, with one more electron than carbon, has a single electron in each of the 2p orbitals. The new electron pairs up with another already in one of the 2p orbitals. It doesn't enter the 3s orbital (the orbital next lowest in energy) since this is so much higher in energy and to enter the 3s orbital would require more energy than that needed to pair up with a 2p electron.



Molecular orbitals—homonuclear diatomics

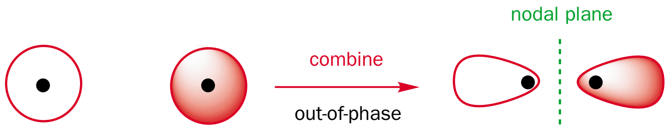
So far the discussion has concerned only the shapes and energies of **atomic orbitals** (**AOs**). Organic chemists really need to look at the orbitals for whole molecules. One way to construct such **molecular orbitals** (**MOs**) is to combine the atomic orbitals of the atoms that make up the molecule. This approach is known as the **Linear Combination of Atomic Orbitals** (**LCAO**).

Atomic orbitals are wave functions and the different wave functions can be combined together rather in the way waves combine. You may be already familiar with the ideas of combining waves—they can add together constructively (in-phase) or destructively (out-of-phase).

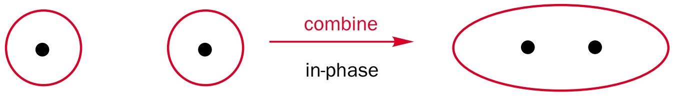


the two ways of combining a simple wave – in-phase and out-of-phase

Atomic orbitals can combine in the same way—in-phase or out-of-phase. Using two 1s orbitals drawn as circles (representing spheres) with dots to mark the nuclei and shading to represent phase, we can combine them in-phase, that is, add them together, or out-of-phase when they cancel each other out in a nodal plane down the centre between the two nuclei. The resulting orbitals belong to both atoms—they are molecular rather than atomic orbitals. As usual, the higher energy orbital is at the top.



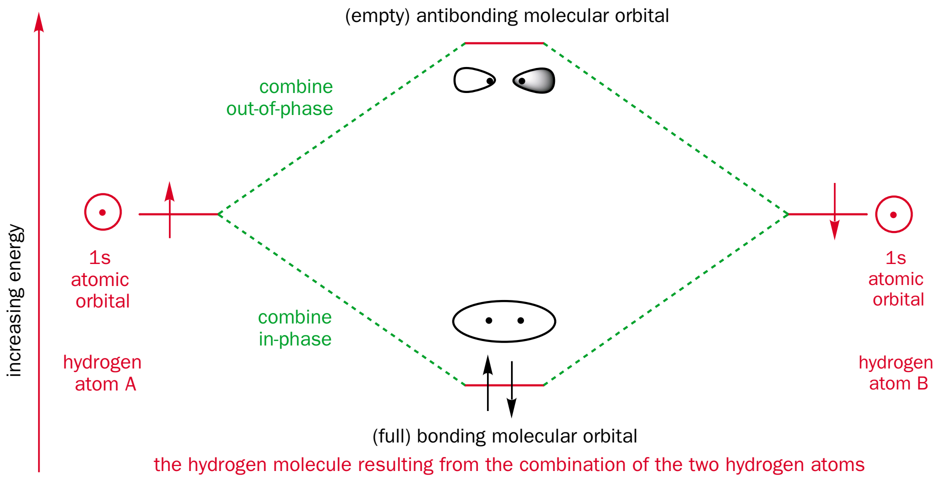
the two 1s orbitals combining out-of-phase to give an **antibonding orbital**



the two 1s orbitals combining in-phase to give a **bonding orbital**

When the two orbitals combine out-of-phase, the resulting molecular orbital has a nodal plane between the two nuclei. This means that if we were to put electrons into this orbital there would be no electron density in between the two nuclei. By contrast, if the molecular orbital from in-phase combination contained electrons, they would be found in between the two nuclei. Two exposed nuclei repel each other as both are positively charged. Any electron density between them helps to bond them together. So the in-phase combination is a **bonding molecular orbital**. As for the electrons themselves, they can now be shared between two nuclei and this lowers their energy relative to the 1s atomic orbital. Electrons in the orbital from the out-of-phase combination do not help bond the two nuclei together; in fact, they hinder the bonding. When this orbital is occupied, the electrons are mainly to be found anywhere *but* between the two nuclei. This means the two nuclei are more exposed to each other and so repel each other. This orbital is known as an **antibonding molecular orbital** and is higher in energy than the 1s orbitals.

The combination of the atomic 1s orbitals to give the two new molecular orbitals is simply shown on an energy level diagram. With one electron in each 1s orbital, two hydrogen atoms combine to give a hydrogen molecule.



the hydrogen molecule resulting from the combination of the two hydrogen atoms

There are several points to notice about this diagram.

- *Two* atomic orbitals (AOs) combine to give *two* molecular orbitals (MOs)
- By LCAO we add the two AOs to make the bonding orbital and subtract them to make the antibonding orbital
- Since the two atoms are the same, each AO contributes the same amount to the MOs
- The bonding MO is *lower* in energy than the AOs
- The antibonding MO is *higher* in energy than the AOs
- Each hydrogen atom initially had one electron. The spin of these electrons is unimportant

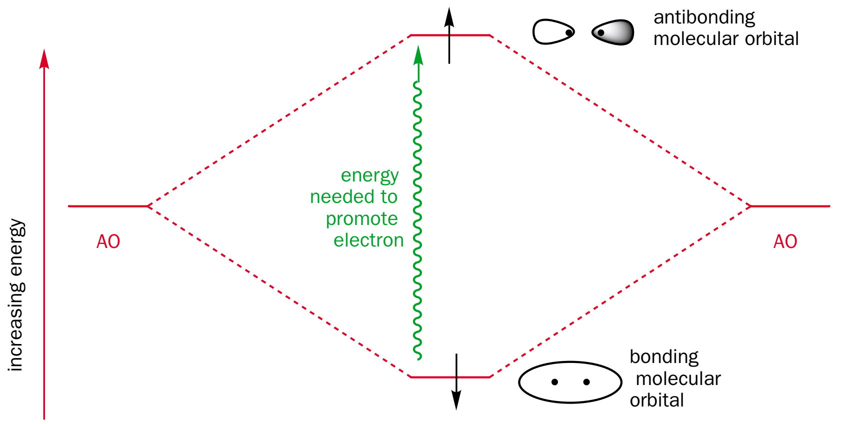
- The two electrons end up in the MO lowest in energy. This is the bonding MO
- Just as with AOs, each MO can hold two electrons *as long as the electrons are spin paired*
- The two electrons between the two nuclei in the bonding MO hold the molecule together—they are the chemical bond
- Since these two electrons are lower in energy in the MO than in the AOs, energy is given out when the atoms combine
- Or, if you prefer, we must put in energy to separate the two atoms again and to break the bond

From now on, we will always represent molecular orbitals in energy order—the highest-energy MO at the top (usually an antibonding MO) and the lowest in energy (usually a bonding MO and the one in which the electrons are most stable) at the bottom. We suggest you do the same.

When we were looking at the electronic configuration of atoms, we simply filled up the atomic orbitals starting from the lowest in energy and worked up. With molecules we do the same: we just fill up the molecular orbitals with however many electrons we have, starting from the lowest in energy and remembering that each orbital can hold two electrons and then only if they are spin paired.

Breaking bonds

If an atom is supplied with energy, an electron can be promoted to a higher energy level and it can then fall back down to its ground state, giving that energy out again. What would happen if an electron were promoted in a hydrogen molecule from the lowest energy level, the bonding MO, to the next lowest energy level, the antibonding MO? Again, an energy level diagram helps.



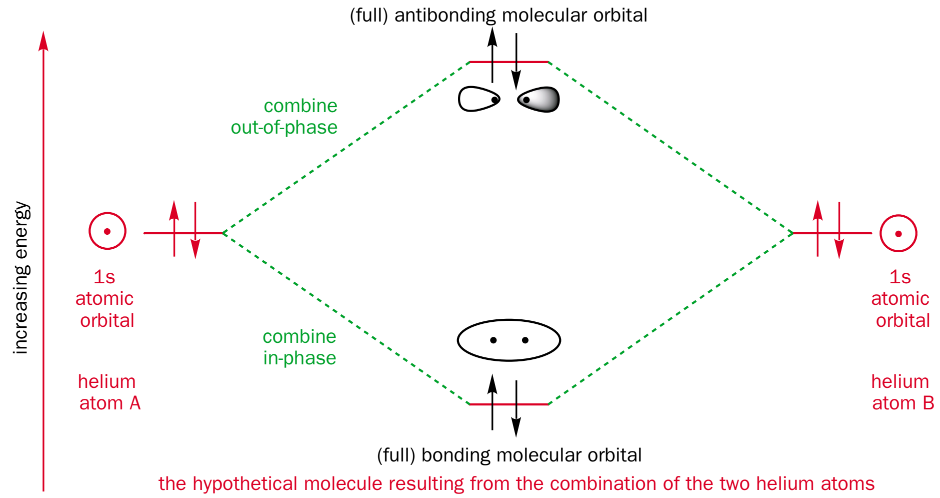
we can supply energy to promote an electron from the bonding MO to the antibonding MO

Now the electron in the antibonding orbital 'cancels out' the bonding of the electron in the bonding orbital. Since there is no overall bonding holding the two atoms together, they can drift apart as two separate atoms with their electrons in 1s atomic orbitals. In other words, promoting an electron from the bonding MO to the antibonding MO breaks the chemical bond. This is difficult to do with hydrogen molecules but easy with, say, bromine molecules. Shining light on Br_2 causes it to break up into bromine atoms.

■ This idea will be developed in Chapters 5 and 6 when we look at bond-breaking steps in organic reaction mechanisms.

Bonding in other elements: helium

A hydrogen molecule is held together by a single chemical bond since the pair of electrons in the bonding orbital constitutes this single bond. What would the MO energy level diagram for He_2 look like? Each helium atom has two electrons ($1s^2$) so now both the bonding MO and the antibonding MO are full. Any bonding due to the electrons in the bonding orbital is cancelled out by the electrons in the antibonding orbital.



the hypothetical molecule resulting from the combination of the two helium atoms

There is no overall bonding, the two helium atoms are not held together, and He_2 does not exist. Only if there are more electrons in bonding MOs than in antibonding MOs will there be any bonding between two atoms. In fact, we define the number of bonds between two atoms as the **bond order** (dividing by two since *two* electrons make up a chemical bond).

$$\text{bond order} = \frac{(\text{no. of electrons in bonding MOs}) - (\text{no. of electrons in antibonding MOs})}{2}$$

Hence the bond orders for H_2 and He_2 are

$$\text{bond order } (H_2) = \frac{2-0}{2} = 1 \quad \text{i.e. a single bond}$$

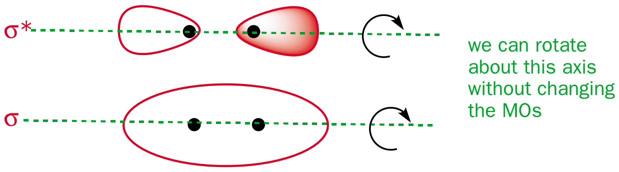
$$\text{bond order } (He_2) = \frac{2-2}{2} = 0 \quad \text{i.e. no bond}$$

Bond formation using 2s and 2p atomic orbitals

So far we have been looking at how we can combine the 1s atomic orbitals to give the molecular orbitals of simple molecules. However, just as there are lots of higher, vacant energy levels in atoms, so there are in molecules too. Other atomic orbitals combine to give new molecular orbitals and the 2s and 2p orbitals concern organic chemistry most of all. The 2s AOs combine in exactly the same way as the 1s orbitals do and also give rise to a bonding and an antibonding orbital. With p orbitals as well, there are more possibilities.

Since we are beginning to talk about lots of different MOs, we shall need to label them with a little more thought. When s orbitals combine, the resulting MOs, both bonding and antibonding, are totally symmetrical about the axis joining the two nuclei.

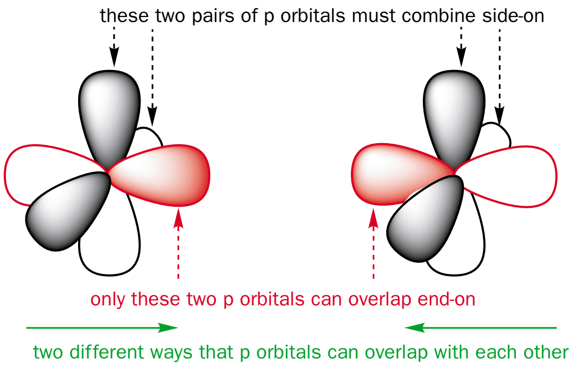
■ Antibonding orbitals are designated with a * e.g. σ^*, or π^*



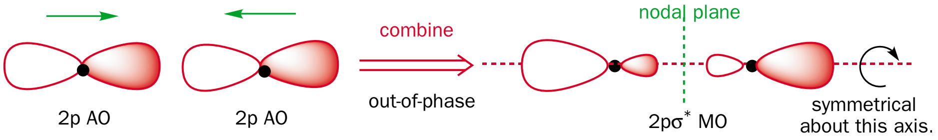
both MOs have rotational symmetry about the axis through the two nuc

When orbitals combine in this end-on overlap to give cylindrically symmetrical MOs, the resulting orbitals are said to possess **sigma (σ) symmetry**. Hence the bonding MO is a **sigma orbital** and electrons in such an orbital give rise to a **sigma bond**. In the hydrogen molecule the two hydrogen atoms are joined by a σ bond.

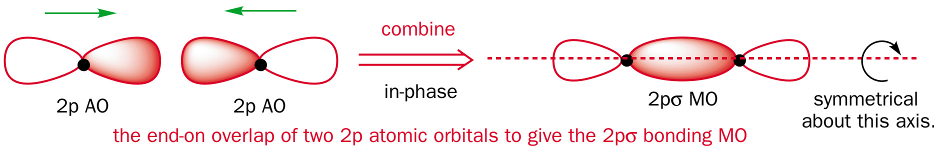
What MOs result from the combination of two p orbitals? There are three mutually perpendicular p orbitals on each atom. As the two atoms approach each other, these orbitals can combine in two different ways—one p orbital from each atom can overlap end-on, but the other two p orbitals on each atom must combine side-on.



The end-on overlap (in-phase and out-of-phase) results in a pair of MOs that are cylindrically symmetrical about the internuclear axis—in other words, these combinations have σ symmetry. The two molecular orbitals resulting from the end-on combination of two 2p orbitals are labelled the 2pσ and the 2pσ* MOs.

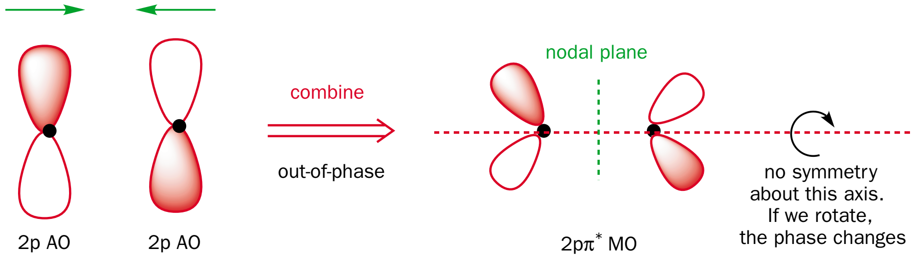


the end-on overlap of two 2p atomic orbitals to give the 2pσ* antibonding MO

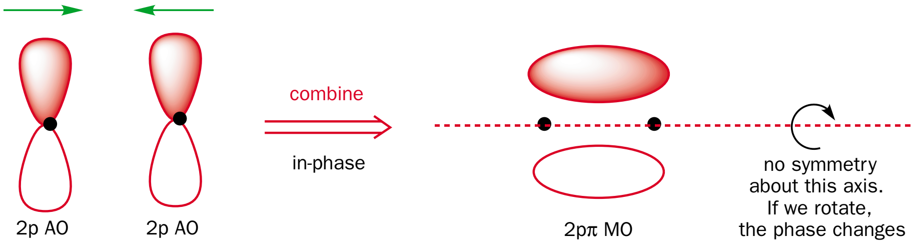


the end-on overlap of two 2p atomic orbitals to give the 2pσ bonding MO

The side-on overlap of two p orbitals forms an MO that is no longer symmetrical about the internuclear axis. If we rotate about this axis, the phase of the orbital changes. The orbital is described as having **π symmetry**—a **π orbital** is formed and the electrons in such an orbital make up a **π bond**. Since there are *two mutually perpendicular pairs* of p orbitals that can combine in this fashion, there are a *pair* of degenerate mutually perpendicular π bonding MOs and a *pair* of degenerate mutually perpendicular π* antibonding MOs.

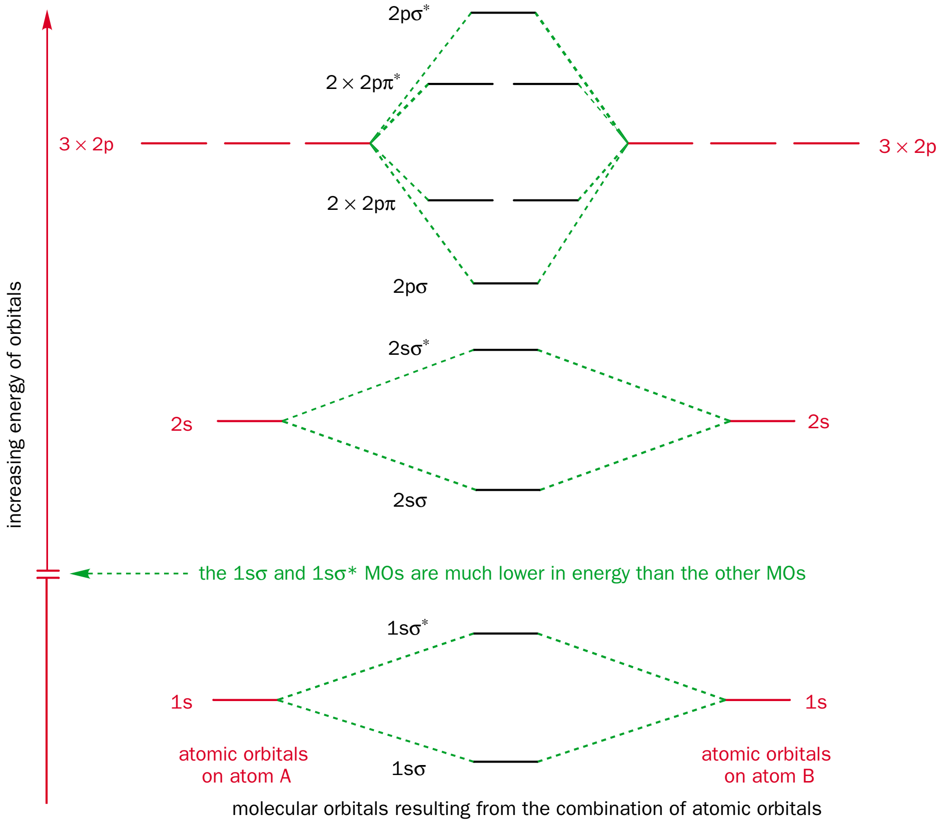


the side-on overlap of two 2p atomic orbitals to give the 2pπ* antibonding MO



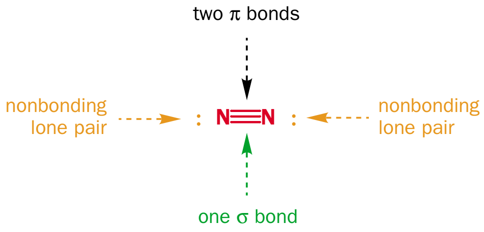
the side-on overlap of two 2p atomic orbitals to give the 2pπ bonding MO

The two sorts of molecular orbitals arising from the combinations of the p orbitals are not degenerate—more overlap is possible when the AOs overlap end-on than when they overlap side-on. As a result, the pσ orbital is lower in energy than the pπ orbital. We can now draw an energy level diagram to show the combination of the 1s, 2s, and 2p atomic orbitals to form molecular orbitals.



Let us now look at a simple diatomic molecule—nitrogen. A nitrogen molecule is composed of two nitrogen atoms, each containing seven electrons in total. We shall omit the 1s electrons because they are so much lower in energy than the electrons in the 2s and 2p AOs and because it makes no difference in terms of bonding since the electrons in the 1sσ* cancel out the bonding due to the electrons in the 1sσ MO. The electrons in the 1s AOs and the 1s MOs are described as core electrons and so, in discussing bonding, we shall consider only the electrons in the outermost shell, in this case the 2s and 2p electrons. This means each nitrogen contributes five bonding electrons and hence the molecular orbitals must contain a total of ten electrons.

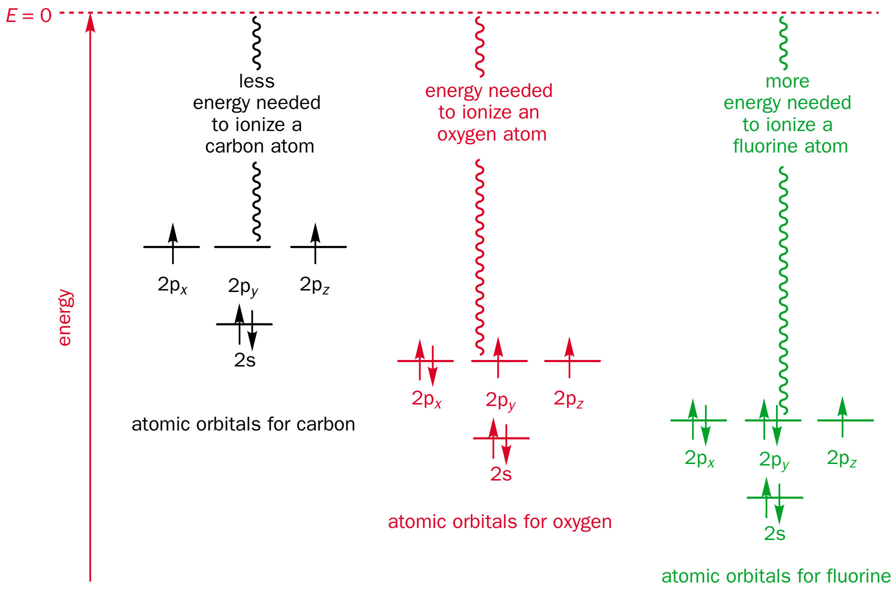
The electrons in the σ and σ* MOs formed from the 2s MOs also cancel out—these electrons effectively sit on the atoms, two on each, and form **lone pairs**—nonbonding pairs of electrons that do not contribute to bonding. All the bonding is done with the remaining six electrons. They fit neatly into a σ bond from two of the p orbitals and two π bonds from the other two pairs. Nitrogen has a triple bonded structure.



▶ Homonuclear and heteronuclear refer to the nature of the atoms in a diatomic molecule. In a **homonuclear molecule** the atoms are the same (such as H_2, N_2, O_2, F_2) while in a **heteronuclear molecule** they are different (as in HF, CO, NO, ICl).

Heteronuclear diatomics

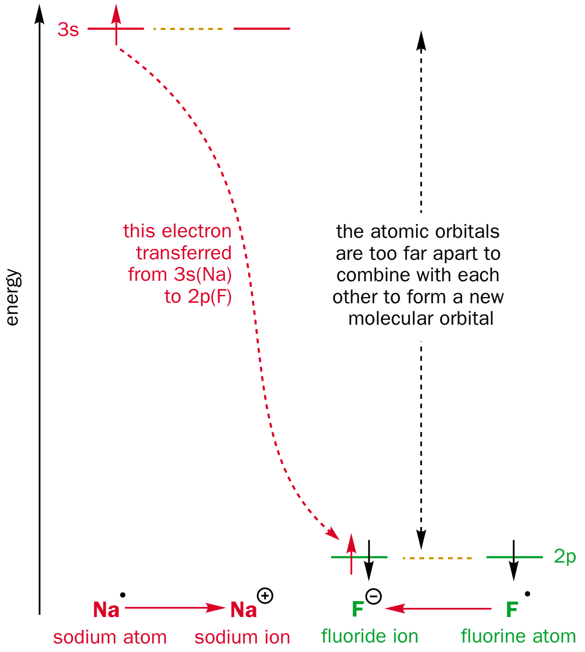
Up to now we have only considered combining two atoms of the same element to form homonuclear diatomic molecules. Now we shall consider what happens when the two atoms are different. First of all, how do the atomic orbitals of different elements differ? They have the same sorts of orbitals 1s, 2s, 2p, etc. and these orbitals will be the same shapes but the orbitals will have different energies. For example, removing an electron completely from atoms of carbon, oxygen, or fluorine (that is, ionizing the atoms) requires different amounts of energy. Fluorine requires most energy, carbon least, even though in each case we are removing an electron from the same orbital, the 2p AO. The energies of the 2p orbitals must be lowest in fluorine, low in oxygen, and highest in carbon.



We are talking now about electronegativity. The more electronegative an atom is, the more it attracts electrons. This can be understood in terms of energies of the AOs. The more electronegative an atom is, the lower in energy are its AOs and so any electrons in them are held more tightly. This is a consequence of the increasing nuclear charge going from left to right across the periodic table. As we go from Li across to C and on to N, O, and F, the elements steadily become more electronegative and the AOs lower in energy.

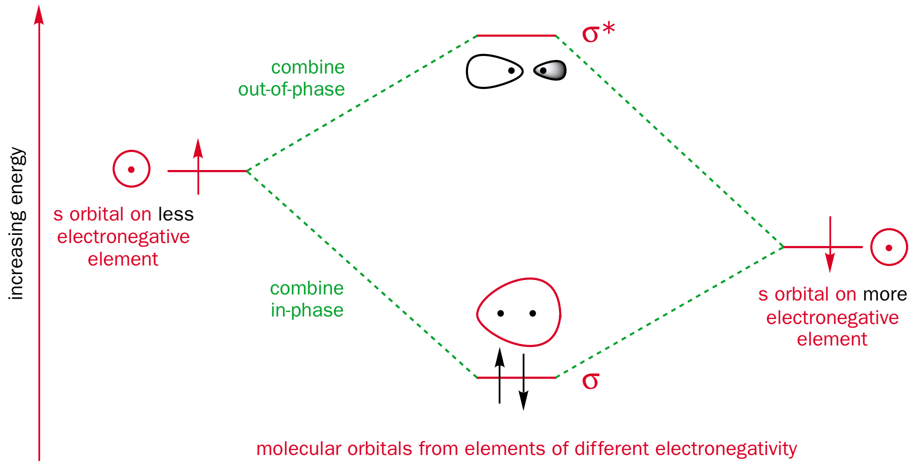
So what happens if two atoms whose atomic orbitals were vastly different in energy, such as Na and F, were to combine? An electron transfers from sodium to fluorine and the product is the ionic salt, sodium fluoride, Na^+F^-.

The important point is that the atomic orbitals are too far apart in energy to combine to form new molecular orbitals and no covalent bond is formed. The ionic bonding in NaF is due simply to the attraction between two oppositely charged ions. When the atomic orbitals have exactly the same energy, they combine to form new molecular orbitals, one with an energy lower than the AOs, the other with an energy higher than the AOs. When the AOs are very different in energy, electrons are transferred from one atom to another and ionic bonding results. When the AOs are *slightly* different in energy, they do combine and we need now to look at this situation in more detail.



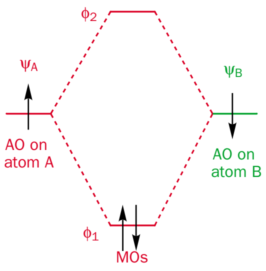
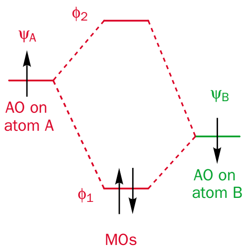
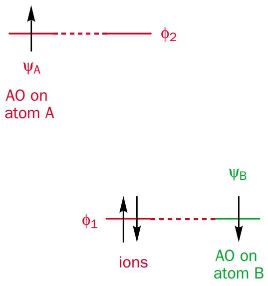
both electrons in sodium fluoride end up in the fluorine's 2p orbital

The AOs combine to form new MOs but they do so unsymmetrically. The more electronegative atom, perhaps O or F, contributes more to the bonding orbital and the less electronegative element (carbon is the one we shall usually be interested in) contributes more to the antibonding orbital. This applies both to σ bonds and to π bonds so here is an idealized case.



molecular orbitals from elements of different electronegativity

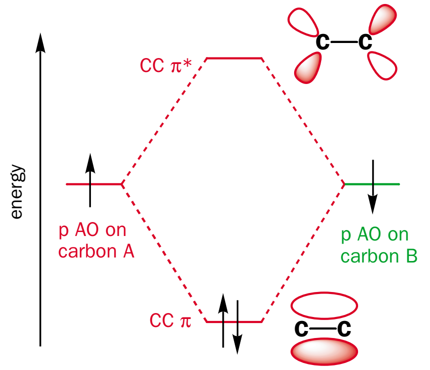
These three different cases where the two combining orbitals differ greatly in energy, only a little, or not at all are summarized below.

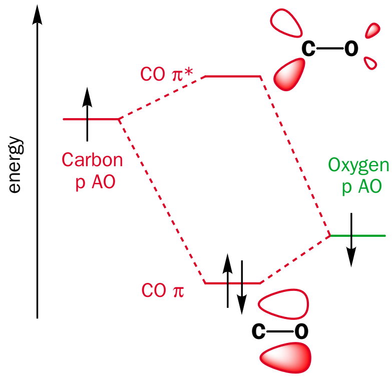
Energies of AOs both the same	**AO on atom B is a *little* lower in energy than AO on atom A**	**AO on atom B is a *lot* lower in energy than AO on atom A**
		
large interaction between AOs	less interaction between AOs	AOs are too far apart in energy to interact
bonding MO much lower in energy than AOs	bonding MO is lowered only by a small amount relative to AO on atom B	the filled orbital on the anion has the same energy as the AO on atom B
antibonding MO is much higher in the energy than the AOs	antibonding MO is raised in energy by only a small amount relative to AO on atom B	the empty orbital on the cation has same energy as the AO on atom A
both AOs contribute equally to the MOs	the AO on B contributes more to the bonding MO and the AO on A	only one AO contributes to each 'MO'
electrons in bonding MO are shared equally between the two atoms	electrons in bonding MO are shared between atoms but are associated more with atom B than A	electrons in the filled orbital are located only on atom B
bond between A and B would classically be described as purely covalent	bond between A and B is covalent but there is also some electrostatic (ionic) attraction between atoms	bond between A and B would classically be described as purely ionic
easiest to break bond into two radicals (homolytic fission). Heterolytic fission of bond is possible and could give either A^+ and B^- or A^- and B^+	easiest to break bond into two ions, A^+ and B^-, although it is also possible to give two radicals	compound already exists as ions A^+ and B^-

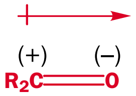
▶

Homolytic and heterolytic refer to the fate of the electrons when a bond is broken. In **heterolytic fission** one electron goes to each atom. In **homolytic fission** both electrons go to the same atom. Thus I_2 easily gives two iodine atoms by homolytic fission ($I_2 \rightarrow 2I^\bullet$) while HI prefers heterolytic fission ($HI \rightarrow H^+ + I^-$). The dot in $I^\bullet$ means a single unpaired electron.

As an example of atomic orbitals of equal and unequal energies combining, let us consider the π bonds resulting from two carbon atoms combining and from a carbon atom combining with an oxygen atom. With the C–C π bond, both p orbitals have the same energy and combine to form a symmetrical π bond. If the bonding MO (π) is occupied, the electrons are shared equally over both carbon atoms. Compare this with the π bond that results from combining an oxygen p AO with a carbon p AO.



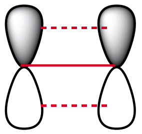




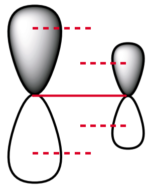
Now the bonding MO (π) is made up with a greater contribution from the oxygen p orbital than from the carbon p orbital. If this MO contained electrons, there would be more electrons around the oxygen atom than around the carbon. This C–O π bond is covalent but there is also some electrostatic contribution to its bond strength. This electrostatic interaction actually makes a C–O double bond much stronger than a C–C double bond (bond strength for C=O, about 725–60 kJ mol^{-1}; for C=C, 600–25 kJ mol^{-1}: compare also a C–O single bond, 350–80 kJ mol^{-1} with a C–C single bond, 340–50 kJ mol^{-1}). Because the electrons in the populated MO (π) are associated more with the oxygen atom than with the carbon, it is easier to break this bond heterolytically with both electrons moving completely on to the oxygen atom than it is to break it homolytically to get a diradical with one electron moving on to the carbon and one on to the oxygen atom. This will be the first chemical reaction we study in detail in Chapters 5 and 6.

Other factors affecting degree of orbital interaction

Having similar energies is not the only criterion for good interaction between two atomic orbitals. It also matters how the orbitals overlap. We have seen that p orbitals overlap better in an end-on fashion (forming a σ bond) than they do side-on (forming a π bond). Another factor is the size of the atomic orbitals. For best overlap, the orbitals should be the same size—a 2p orbital overlaps much better with another 2p orbital than it does with a 3p or 4p orbital.

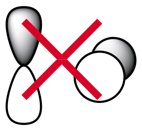


efficient overlap of p orbitals of the same size (same principal quantum number *n*)



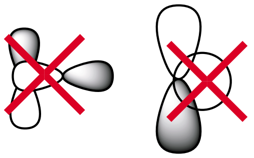
inefficient overlap of p orbitals of different size (different principal quantum numbers *n*)

A third factor is the symmetry of the orbitals—two atomic orbitals must have the appropriate symmetry to combine. Thus a $2p_x$ orbital cannot combine with a $2p_y$ or $2p_z$ orbital since they are all perpendicular to each other (they are **orthogonal**). In one case the two p orbitals have no overlap at all; in the other case any constructive overlap is cancelled out by equal amounts of destructive



p_z and p_x

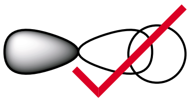
these two p orbitals cannot combine because they are perpendicular to each other



p_z and p_y

p and s (side-on)

here any constructive overlap is cancelled out by equal amounts of destructive overlap

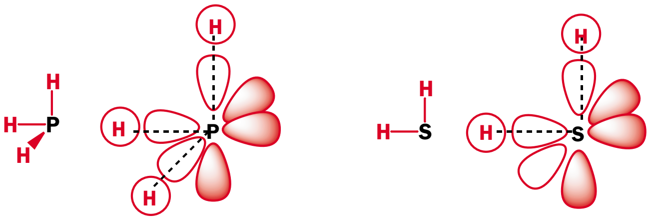


p and s (end-on)

however, s and p orbitals can overlap end-on overlap. Likewise, an s orbital can overlap with a p orbital only end-on. Sideways overlap leads to equal amounts of bonding and antibonding interactions and no overall gain in energy.

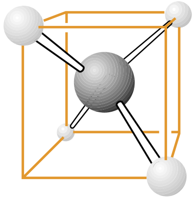
Molecular orbitals of molecules with more than two atoms

We now need to look at ways of combining more than two atoms at a time. For some molecules, such as H_2S and PH_3, that have all bond angles equal to 90°, the bonding should be straightforward—the p orbitals (which are at 90°) on the central atom simply overlap with the 1s orbitals of the hydrogen atoms.



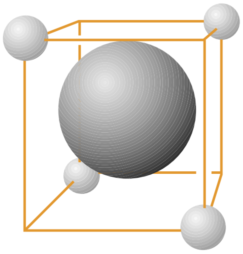
the 90° angles in PH_3 and H_2S come from the overlap of the hydrogen 1s AO with the p AO of the phosphorus or sulfur

But how do we account for the bond angles in water (104°) and ammonia (107°) when the only atomic orbitals are at 90° to each other? All the covalent compounds of elements in the row Li to Ne raise this difficulty. Water (H_2O) and ammonia (NH_3) have angles between their bonds that are roughly tetrahedral and methane (CH_4) is exactly tetrahedral but how can the atomic orbitals combine to rationalize this shape? The carbon atom has electrons only in the first and second shells, and the 1s orbital is too low in energy to contribute to any molecular orbitals, which leaves only the 2s and 2p orbitals. The problem is that the 2p orbitals are at right angles to each other and methane does not have any 90° bonds. (So don't draw any either! Remember Chapter 2.). Let us consider exactly where the atoms are in methane and see if we can combine the AOs in such a way as to make satisfactory molecular orbitals.



a molecule of methane enclosed in a cube

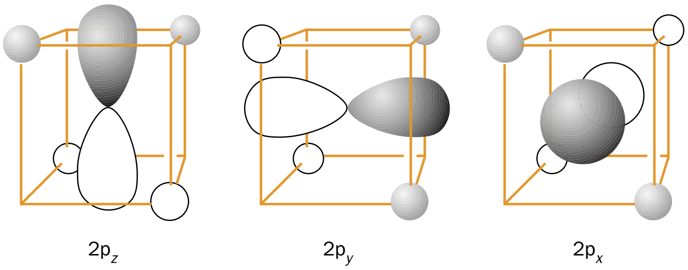
Methane has a tetrahedral structure with each C–H bond 109 pm and all the bond angles 109.5°. To simplify things, we shall draw a molecule of methane enclosed in a cube. It is possible to do this since the opposite corners of a cube describe a perfect tetrahedron. The carbon atom is at the centre of the cube and the four hydrogen atoms are at four of the corners.



the carbon 2s AO can overlap with all four hydrogen 1s AOs at once

Now, how can the carbon's 2s and 2p atomic orbitals combine with the four hydrogen 1s atomic orbitals? The carbon's 2s orbital can overlap with all four hydrogen 1s orbitals at once with all the orbitals in the same phase. In more complicated systems like this, it is clearer to use a diagram of the AOs to see what the MO will be like.

Each of the 2p orbitals points to opposite faces of the cube. Once more all four hydrogen 1s orbitals can combine with each p orbital but this time the hydrogen AOs on the opposite faces of the cube must be differently phased.



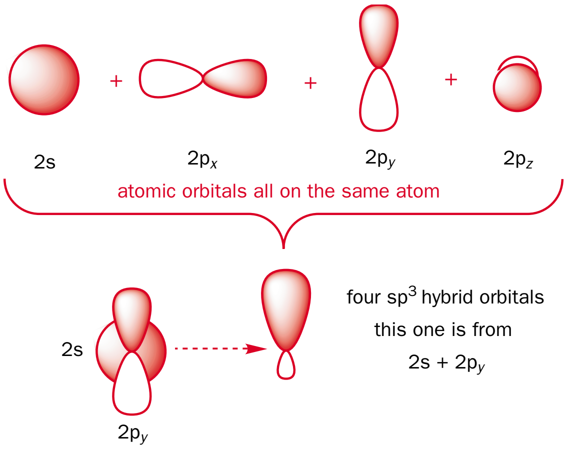
the hydrogen 1s orbitals can overlap with the three 2p orbitals

Again we are not going to draw these three molecular orbitals but you can see from the AO diagrams what they look like. They are degenerate (that is, they have the same energy) and each orbital has one nodal plane (it is easiest to see in the middle diagram passing vertically down the middle of the cube and dividing shaded orbitals on the right from unshaded orbitals on the left). Only the bonding overlap between the AOs is shown but of course there is an antibonding interaction for every bonding interaction, which means there are eight MOs altogether (which is correct since there were eight AOs to start with).

Organic chemists can just about understand this 'correct' MO picture of methane and theoretical chemists are able to construct correct MOs for very much more complex molecules than methane. There is experimental evidence too that these pictures are correct. Other experiments reveal that all four C–H bonds in methane are exactly the same and yet the MOs for methane are not all the same. *There is no contradiction here!* The molecular orbital approach tells us that there is one MO of one kind and three of another but the electrons in them are shared out over all five atoms. No one hydrogen atom has more or less electrons than any other—they are all equivalent. Techniques that tell us the structure of methane do not tell us where bonds are; they simply tell us where the atoms are located in space—*we* draw in bonds connecting atoms together. Certainly the *atoms* form a regular tetrahedron but exactly where the electrons are is a different matter entirely. The classical picture of two atoms held together by a pair of electrons is not necessarily correct—the five atoms in methane are held together by electrons but these are in molecular orbitals, which spread over all the atoms. We are going to need the classical picture when we draw mechanisms. Methane only has one carbon atom—imagine what it would be like with larger compounds that can contain hundreds of carbon atoms! Fortunately, there is another, simpler method we can use to describe bonding that preserves the important points from this theory.

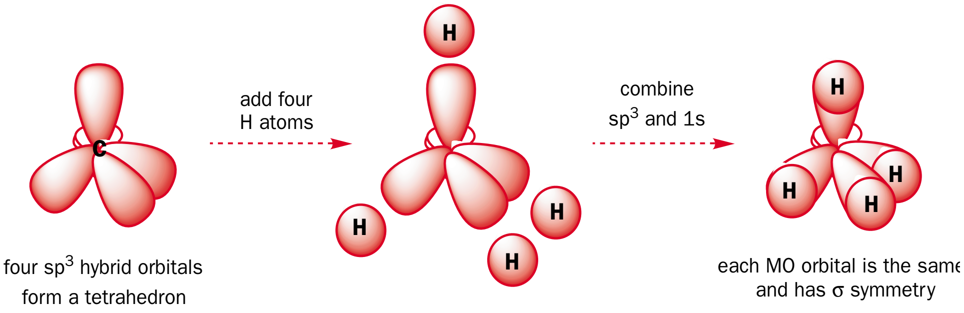
Hybridization of atomic orbitals

For most of organic chemistry, it is helpful to consider the molecule as being made up of atoms held together by bonds consisting of a pair of electrons. When working out the MOs for methane, we used the carbon 2s and all three of the 2p orbitals to combine with the hydrogen 1s orbitals. Each orbital combined with all the hydrogen orbitals equally. Another way to consider the bonding would be to combine the carbon 2s and 2p orbitals first to make four new orbitals. Each of these orbitals would be exactly the same and be composed of one-quarter of the 2s orbital and three-quarters of one of the p orbitals. The new orbitals are called sp^3 hybrid orbitals to show the proportions of the AOs in each. This process of mixing is called **hybridization.**

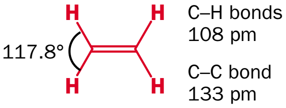


Combining four atomic orbitals on the same atom gives the same total number of hybrid orbitals. Each of these has one-quarter s character and three-quarters p character. The sp^3 orbital has a planar node through the nucleus like a p orbital but one lobe is larger than the other because of the extra contribution of the 2s orbital, which adds to one lobe but subtracts from the other.

The four sp^3 orbitals on one carbon atom point to the corners of a tetrahedron and methane can be formed by overlapping the large lobe of each sp^3 orbital with



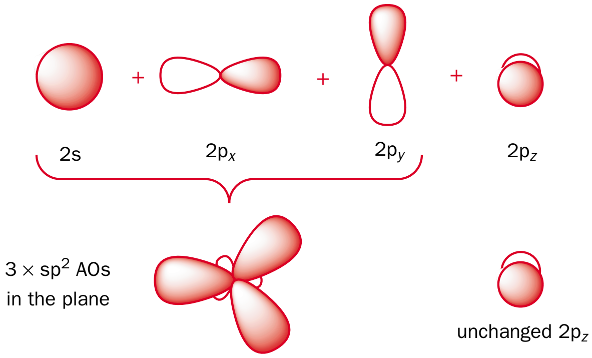
the 1s orbital of a hydrogen atom. Each overlap forms an MO ($2sp^3 + 1s$) and we can put two electrons in each to form a C–H σ bond. There will of course also be an antibonding MO, σ* ($2sp^3 – 1s$) in each case, but these orbitals are empty.



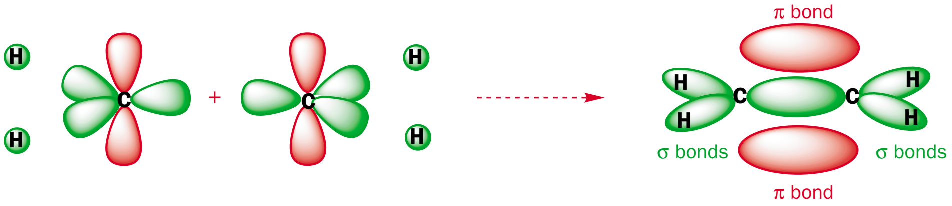
ethene
(ethylene)

The great advantage of this method is that it can be used to build up structures of much larger molecules quickly and without having to imagine that the molecule is made up from isolated atoms. So it is easy to work out the structure of ethene (ethylene) the simplest alkene. Ethene is a planar molecule with bond angles close to 120°. Our approach will be to hybridize all the orbitals needed for the C–H framework and see what is left over. In this case we need three bonds from each carbon atom (one to make a C–C bond and two to make C–H bonds).

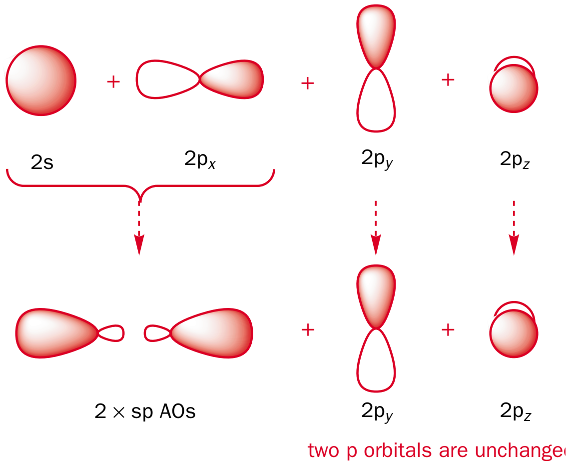
Therefore we need to combine the 2s orbital on each carbon atom with two p orbitals to make the three bonds. We could hybridize the 2s, $2p_x$, and $2p_y$ orbitals (that is, all the AOs in the plane) to form three equal sp^2 hybrid atomic orbitals, leaving the $2p_z$ orbital unchanged. These sp^2 hybrid orbitals will have one-third s character and only two-thirds p character.



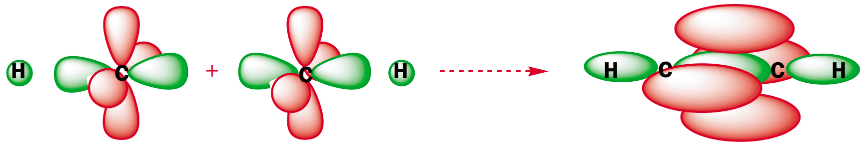
The three sp^2 hybrid atomic orbitals on each carbon atom can overlap with three other orbitals (two hydrogen 1s AOs and one sp^2 AO from the other carbon) to form three σ MOs. This leaves the two $2p_z$ orbitals, one on each carbon, which combine to form the π MO. The skeleton of the molecule has five σ bonds (one C–C and four C–H) in the plane and the central π bond is formed by two $2p_z$ orbitals above and below the plane.



Ethyne (acetylene) has a C–C triple bond. Each carbon bonds to only two other atoms to form a linear CH skeleton. Only the carbon 2s and $2p_x$ have the right symmetry to bind to only two atoms at once so we can hybridize these to form two sp hybrids on each carbon atom leaving the $2p_y$ and $2p_z$ to form π MOs with the 2p orbitals on the other carbon atom. These sp hybrids have 50% each s and p character and form a linear carbon skeleton.



We could then form the MOs as shown below. Each sp hybrid AO overlaps with either a hydrogen 1s AO or with the sp orbital from the other carbon. The two sets of p orbitals combine to give two mutually perpendicular π MOs.

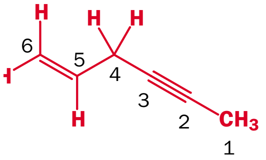


2sp hybridized carbon atoms
$2p_y$ and $2p_z$ orbitals remain

linear σ bonds form skeleton
two perpendicular π bonds
$H-C\equiv C-H$

Hydrocarbon skeletons are built up from tetrahedral (sp^3), trigonal planar (sp^2), or linear (sp) hybridized carbon atoms. It is not necessary for you to go through the hybridization process each time you want to work out the shape of a skeleton. In real life molecules are not made from their constituent atoms but from other molecules and it doesn't matter how complicated a molecule might be or where it comes from; it will have an easily predictable shape. All you have to do is count up the single bonds at each carbon atom. If there are two, that carbon atom is linear (sp hybridized), if there are three, that carbon atom is trigonal (sp^2 hybridized), and, if there are four, that carbon atom is tetrahedral (sp^3 hybridized).

This hydrocarbon (hex-5-en-2-yne) has two linear sp carbon atoms (C2 and C3), two trigonal sp^2 carbon atoms (C5 and C6), a tetrahedral sp^3 CH_2 group in the middle of the chain (C4), and a tetrahedral sp^3 methyl group (C1) at the end of the chain. We had no need to look at any AOs to deduce this—we needed only to count the bonds.



hex-5-en-2-yne

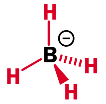
▶ Notice that atoms 1–4 are drawn in a straight line. Alkynes are linear—draw them like that!

If you had drawn the molecule more professionally as shown in the margin, you would have to check that you counted up to four bonds at each carbon. Of course, if you just look at the double and triple bonds, you will get the right answer without counting single bonds at all. Carbon atoms with no π bonds are tetrahedral (sp^3 hybridized), those with one π bond are trigonal (sp^2 hybridized), and those with two π bonds are linear (sp hybridized). This is essentially the VSEPRT approach with a bit more logic behind it.

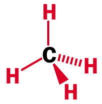
- **All normal compounds of carbon have eight electrons in the outer shell ($n = 2$) of the carbon atom, all shared in bonds. It doesn't matter where these electrons come from; just fit them into the right MOs on sp, sp^2, or sp^3 atoms.**

We can hybridize any atoms

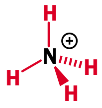
Hybridization is a property of AOs rather than specifically of carbon and, since all atoms have AOs, we can hybridize any atom. A tetrahedral arrangement of atoms about any central atom can be rationalized by describing the central atom as sp^3 hybridized. The three molecules shown here all have a tetrahedral structure and in each case the central atom can be considered to be sp^3 hybridized.



borohydride anion



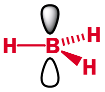
methane



ammonium cation

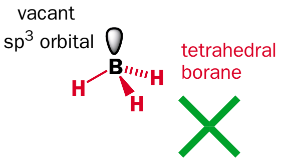
Each of these three molecules has four equivalent σ bonds from the central tetrahedral sp^3 atom, whether this is B, C, or N, and the same total number of bonding electrons—the molecules are said to be **isoelectronic**. These three elements come one after the other in the periodic table so each nucleus has one more proton than the last: B has 5, C has 6, and N has 7. This is why the charge on the central atom varies.

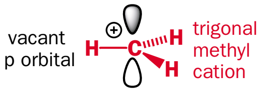
Compounds of the same three elements with only three bonds are more complicated. Borane, BH_3, has only three pairs of bonding electrons. The central boron atom bonds to only three other atoms. We can therefore describe it as being sp^2 hybridized with an empty p orbital.



vacant p orbital
trigonal Borane

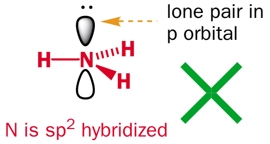
Each of the B–H bonds results from the overlap of an sp^2 orbital with the hydrogen 1s orbital. The p orbital is not needed and contains no electrons. Do not be tempted by the alternative structure with tetrahedral boron and an empty sp^3 orbital. You want to populate the lowest energy orbitals for greatest stability and sp^2 orbitals with their greater s character are lower in energy than sp^3 orbitals. Another way to put this is that, if you have to have an empty orbital, it is better to have it of the highest possible energy since it has no electrons in it and doesn't affect the stability of the molecule.



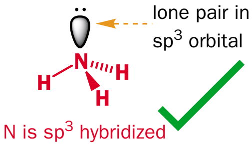


Borane is isoelectronic with the methyl cation, CH_3^+. All the arguments we have just applied to borane also apply to Me^+ so it too is sp^2 hybridized with a vacant p orbital. This will be very important when we discuss the reactions of carbocations in Chapter 17.

Now what about ammonia, NH_3? Ammonia is *not* isoelectronic with borane and Me^+! As well as three N–H bonds, each with two electrons, the central nitrogen atom also has a lone pair of electrons. We have two choices: either we could hybridize the nitrogen atom sp^2 and put the lone pair in the p orbital or we could hybridize the nitrogen sp^3 and have the lone pair in an sp^3 orbital.



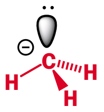
N is sp^2 hybridized



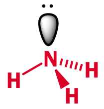
N is sp^3 hybridized

This is the opposite of the situation with borane and Me^+. The extra pair of electrons *does* contribute to the energy of ammonia so it should be in the lower-energy orbital, sp^3, rather than pure p. Experimentally the H–N–H bond angles are all 107.3°. Clearly, this is much closer to the 109.5° sp^3 angle than to the 120° sp^2 angle. But the bond angles are not exactly 109.5°, so ammonia cannot be described as pure sp^3 hybridized. VSEPRT says the lone pair repels the bonds more than they repel each other. Alternatively, you could say that the orbital containing the lone pair must have slightly more s character while the N–H bonding orbitals must have correspondingly more p character.

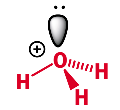
The methyl anion, CH_3^-, and hydronium ion, H_3O^+, are both isoelectronic with ammonia so that all share the same pyramidal structure. Each is approximately tetrahedral with a lone pair in an sp^3 orbital. These elements follow each other in the periodic table so the change in charge occurs because each nucleus has one more proton than the last. VSEPRT also gives this answer.



methyl anion



ammonia



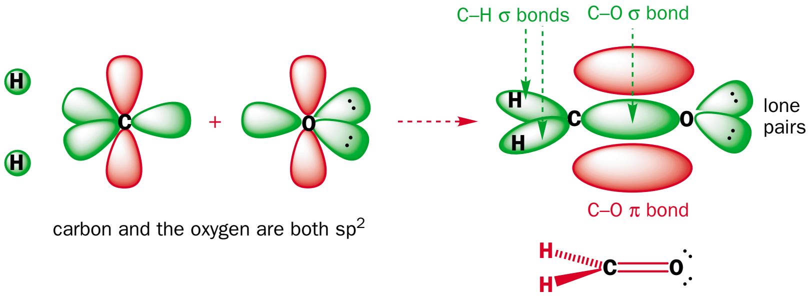
hydronium ion

Shape of phosphine

Phosphine, PH_3, has bond angles of about 90° and there is no need for hybridization. The three H 1s AOs can overlap with the three 3p orbitals of the phosphorus atom, which leaves the lone pair in the 3s orbital. This 'pure s' lone pair is less energetic and therefore less reactive than the sp^3 lone pair in ammonia which explains why ammonia is more basic than phosphine (see Chapter 8). In general atoms from Na to Ar are less likely to be hybridized than those from Li to Ne because the longer bonds mean the substituents are further from the central atom and steric interaction is less. VSEPRT does *not* give this answer.

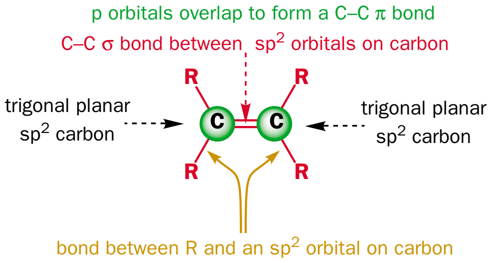
Double bonds to other elements

The C=O double bond is the most important functional group in organic chemistry. It is present in aldehydes, ketones, acids, esters, amides, and so on. We shall spend Chapters 5–10 discussing its chemistry so it is important that you understand its electronic structure. As in alkenes, the two atoms that make up this double bond are sp^2 hybridized. The carbon atom uses all three sp^2 orbitals for overlap with other orbitals to form σ bonds, but the oxygen uses only one for overlap with another orbital (the sp^2 orbitals on the carbon atom) to form a σ bond. However, the other two sp^2 orbitals are not vacant—they contain the oxygen's two lone pairs. A p orbital from the carbon and one from the oxygen make up the π bond which also contains two electrons.

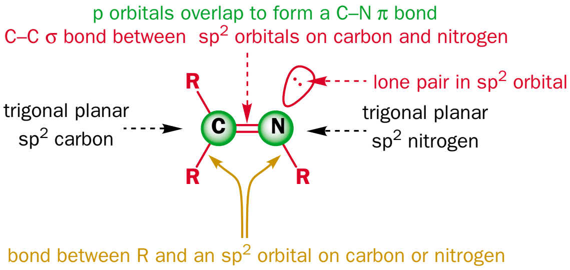


The less important double bonds to nitrogen (imines) are very similar but now there is only one lone pair on nitrogen and a second σ bond to whatever substituent is on the nitrogen atom. Looking down on the planar structures of alkenes, imines, and ketones we see only the ends of the p orbitals but the rest of the structures are clearly related.

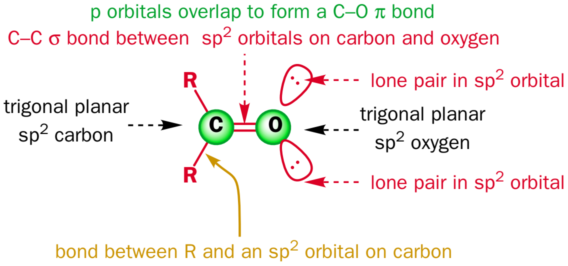
Alkenes have a planar trigonal framework of sp^2 carbon atoms. Each uses one sp^2 orbital to form a σ bond to the other carbon atom and two sp^2 orbitals to form σ bonds to the substituents (here the general 'R'). Two carbon p orbitals are used for a C–C π bond. There are no lone pairs of electrons on either carbon atom.



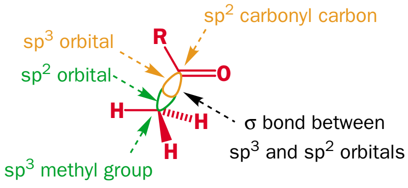
Imines have a planar trigonal framework of an sp^2 carbon atom and an sp^2 nitrogen atom. Each uses one sp^2 orbital to form a σ bond to the other atom and a p orbital to form a π bond to the other atom. The carbon uses two sp^2 orbitals and the nitrogen one to form σ bonds to the substituents (here the general 'R'). There is one lone pair of electrons on the nitrogen atom.

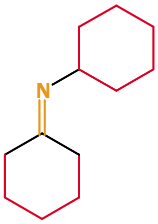


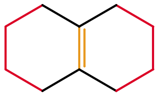
Carbonyl compounds have a planar trigonal framework of an sp^2 carbon atom and an sp^2 oxygen atom. Each uses one sp^2 orbital to form a σ bond to the other atom and a p orbital to form a π bond to the other atom. The carbon uses two sp^2 orbitals to form σ bonds to the substituents (here the general 'R'). There are two lone pairs of electrons on the oxygen atom.



Where 'R' is joined to the double bond through a carbon atom, the nature of R determines which orbital will be used to pair up with the sp^2 orbital. In all the compounds shown below a saturated carbon atom with four bonds is joined to the double bond. The C–C single bond is a σ bond between

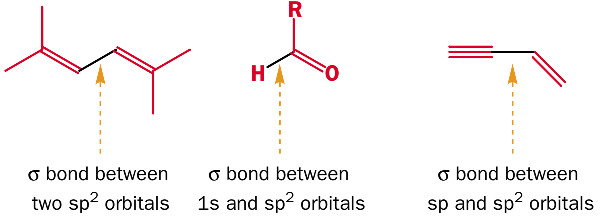




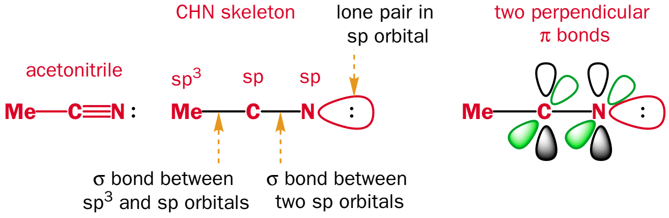


an sp^2 orbital on the ketone, imine, or alkene and an sp^3 orbital on the substituent. It doesn't make any difference that the second two compounds contain rings. In all cases the black bond joins a saturated, tetrahedral, sp^3 carbon atom to the double bond and all the black σ bonds are between sp^2 and sp^3 carbons or nitrogens.

All the other combinations are possible—here are just a few. It should be clear by now that σ bonds can form between any sort of orbitals that can point towards each other but that π bonds can form only between p orbitals.



Triple bonds can be formed between carbon and other elements too. The most important is the CN triple bond present in cyanides or nitriles. Both C and N are sp hybridized in these linear molecules, which leaves the lone pair on nitrogen in an sp orbital too. You will see (Chapter 8) how this affects the basicity of nitriles.



● **All normal compounds of nitrogen have eight electrons in the outer shell ($n = 2$) of the nitrogen atom, six shared in bonds and two in a lone pair. All normal compounds of oxygen have eight electrons in the outer shell ($n = 2$) of the oxygen atom, four shared in bonds and four in lone pairs. It doesn't matter where these electrons come from; just fit them into the right MOs on sp, sp^2, or sp^3 atoms.**

Conclusion

We have barely touched the enormous variety of molecules, but it is important that you realize at this point that these simple ideas of structural assembly can be applied to the most complicated molecules known. We shall use AOs and combine them into MOs to solve the structure of very small molecules and to deduce the structures of small parts of much larger molecules. With the additional ideas in Chapter 7 (conjugation) you will be able to grasp the structures of any organic compound. From now on we shall use terms like AO and MO, 2p orbital, sp^2 hybridization, σ bond, energy level, and populated orbital without further explanation. If you are unsure about any of them, refer back to this chapter for an explanation.

Problems

1. In the (notional and best avoided in practice) formation of NaCl from a sodium atom and a chlorine atom, descriptions like this abound in textbooks: 'an electron is transferred from the valency shell of the sodium atom to the valency shell of the chlorine atom'. What is meant, in quantum number terms, by 'valency shell'? Give a complete description in terms of all four quantum numbers of that transferred electron: (**a**) while it is in the sodium atom and (**b**) after it has been transferred to the chlorine atom. Why is the formation of NaCl by this process to be discouraged?

2. What is the electronic structure of these species? You should consult a periodic table before answering.

$H^{\ominus}$ $HS^{\ominus}$ $K^{\oplus}$ Xe

3. What sort of bonds can be formed between s orbitals and p orbitals? Which will provide better overlap, 1s + 2p or 1s + 3p? Which bonds will be stronger, those between hydrogen and C, N, O, and F on the one hand or those between hydrogen and Si, P, S, and Cl on the other? Within the first group, bond strength goes in this order: HF > OH > NH > CH. Why?

4. Though no helium 'molecule' He_2 exists, an ion He_2^+ does exist. Explain.

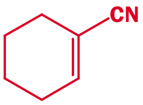
5. You may be surprised to know that the molecule CH_2, with divalent carbon, can exist. It is of course very unstable but it is known and it can have two different structures. One has an H–C–H bond angle of 180° and the other an angle of 120°. Suggest structures for these species and say which orbitals will be occupied by all bonding and nonbonding electrons. Which structure is likely to be more stable?

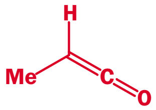
6. Construct an MO diagram for the molecule LiH and suggest what type of bond it might have.

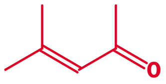
7. Deduce the MOs for the oxygen molecule. What is the bond order in oxygen and where are the 2p electrons?

8. Construct MOs for acetylene (ethyne) without hybridization.

9. What is the shape and hybridization of each carbon atom in these molecules?







10. Suggest detailed structures for these molecules and predict their shapes. We have deliberately made noncommittal drawings to avoid giving away the answer to the question. Don't use these sorts of drawing in your answer.

CO_2, $CH_2{=}NCH_3$, CHF_3, $CH_2{=}C{=}CH_2$, $(CH_2)_2O$

Organic reactions

5

Connections

Building on:

- **Drawing molecules realistically ch2**
- **Ascertaining molecular structure spectroscopically ch3**
- **What determines molecular shape and structure ch4**

Arriving at:

- **Why molecules generally *don't* react with each other!**
- **Why sometimes molecules *do* react with each other**
- **In chemical reactions electrons move from full to empty orbitals**
- **Molecular shape and structure determine reactivity**
- **Representing the movement of electrons in reactions by curly arrows**

Looking forward to:

- **The rest of the chapters in this book**

Chemical reactions

Most molecules are at peace with themselves. Bottles of water, or acetone (propanone, $Me_2C{=}O$), or methyl iodide (iodomethane CH_3I) can be stored for years without any change in the chemical composition of the molecules inside. Yet when we add chemical reagents, say, HCl to water, sodium cyanide (NaCN) to acetone, or sodium hydroxide to methyl iodide, chemical reactions occur. This chapter is an introduction to the reactivity of organic molecules: why they don't and why they do react; how we can understand reactivity in terms of charges and orbitals and the movement of electrons; how we can represent the detailed movement of electrons—the mechanism of the reaction—by a special device called the curly arrow.

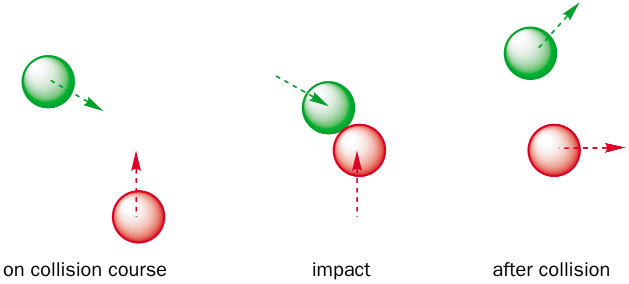
To understand organic chemistry you must be familiar with two languages. One, which we have concentrated on so far, is the structure and representation of molecules. The second is the description of the reaction mechanism in terms of curly arrows and that is what we are about to start. The first is static and the second dynamic. The creation of new molecules is the special concern of chemistry and an interest in the mechanism of chemical reactions is the special concern of organic chemistry.

Molecules react because they move. They move internally—we have seen (Chapter 3) how the stretching and bending of bonds can be detected by infrared spectroscopy. Whole molecules move continuously in space, bumping into each other, into the walls of the vessel they are in, and into the solvent if they are in solution. When one bond in a single molecule stretches too much it may break and a chemical reaction occurs. When two molecules bump into each other, they may combine with the formation of a new bond, and a chemical reaction occurs. We are first going to think about collisions between molecules.

▶ The **activation energy**, also called the **energy barrier** for a reaction, is the minimum energy molecules must have if they are to react. A population of a given molecule in solution at room temperature has a range of energies. If the reaction is to occur, some at least must have an energy greater than the activation energy. We shall discuss this concept in more detail in Chapter 13.

Not all collisions between molecules lead to chemical change

All organic molecules have an outer layer of many electrons, which occupy filled orbitals, bonding and nonbonding. Charge–charge repulsion between these electrons ensures that all molecules repel each other. Reaction will occur only if the molecules are given enough energy (the activation energy for the reaction) for the molecules to pass the repulsion and get close enough to each other. If two molecules lack the required activation energy, they will simply collide, each bouncing off the electrons on the surface of the other and exchanging energy as they do so, but remain chemically unchanged. This is rather like a collision in snooker or pool. Both balls are unchanged afterwards but are moving in different directions at new velocities.



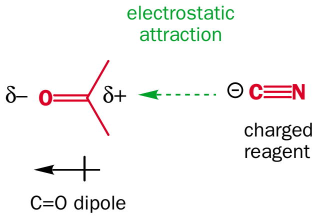
Charge attraction brings molecules together

In addition to this universal repulsive force, there are also important attractive forces between molecules if they are charged. Cations (+) and anions (–) attract each other electrostatically and this may be enough for reaction to occur. When an alkyl chloride, RCl, reacts with sodium iodide, NaI, in acetone (propanone, $Me_2C{=}O$) solution a precipitate of sodium chloride forms. Sodium ions, Na^+, and chloride ions, Cl^-, ions in solution are attracted by their charges and combine to form a crystalline lattice of alternating cations and anions—the precipitate of crystalline sodium chloride.

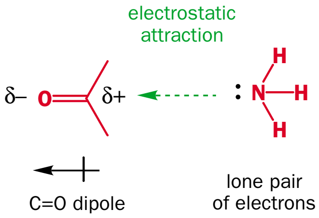
■ We saw why these atoms form an ionic compound in Chapter 4.

This inorganic style of attraction is rare in organic reactions. A more common cause of organic reactions is attraction between a charged reagent (cation or anion) and an organic compound that has a dipole. An example that we shall explore in this chapter is the reaction between sodium cyanide (a salt, NaCN) and a carbonyl compound such as acetone. Sodium cyanide is made up of sodium cations, Na^+, and cyanide anions, CN^-, in solution. Acetone has a carbonyl group, a C=O double bond, which is polarized because oxygen is more electronegative than carbon. The negative cyanide ion is attracted to the positive end of the carbonyl group dipole.

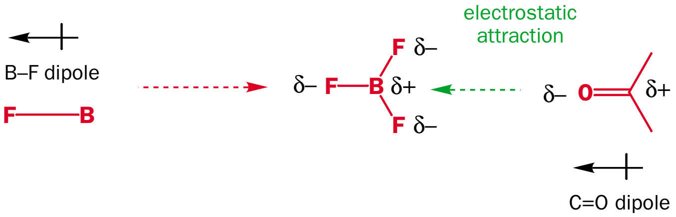
■ We analysed the orbitals of the carbonyl group in Chapter 4 and established that the reason for the polarity is the greater electronegativity of the oxygen atom.



It is not even necessary for the reagent to be charged. Ammonia also reacts with acetone and this time it is the **lone pair** of electrons —a pair of electrons not involved in bonding and concentrated on the nitrogen atom of the uncharged ammonia molecule—that is attracted to the positive end of the carbonyl dipole.



Polarity can arise from σ bonds too. The most electronegative element in the periodic table is fluorine and three fluorine atoms on electropositive boron produce a partially positively charged boron atom by σ bond polarization. The negative end of the acetone dipole (the oxygen atom) is attracted to the boron atom in BF_3.

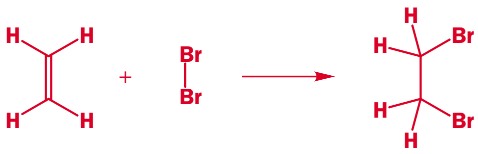


But we have not told you the whole story about BF_3. Boron is in group 3 and thus has only six electrons around it in its trivalent compounds. A molecule of BF_3 is planar with an empty p orbital. This is the reverse of a lone pair. An empty orbital on an atom does *not* repel electron-rich areas of other molecules and so the oxygen atom of acetone is attracted electrostatically to the partial positive charge and one of the lone pairs on oxygen can form a bonding interaction with the empty orbital. We shall develop these ideas in the next section.

So, to summarize, the presence of a dipole in a molecule represents an imbalance in the distribution of the bonding electrons due to polarization of a σ bond or a π bond or to a pair of electrons or an empty orbital localized on one atom. When two molecules with complementary dipoles collide and together have the required activation energy to ensure that the collision is sufficiently energetic to overcome the general electronic repulsion, chemical change or reaction can occur.

Orbital overlap brings molecules together

Other organic reactions take place between completely uncharged molecules with no dipole moments. One of the old 'tests' for unsaturation was to treat the compound with bromine water. If the brown colour disappeared, the molecule was unsaturated. We don't use 'tests' like these any more (spectroscopy means we don't need to) but the reaction is still an important one. A simple symmetrical alkene combines with symmetrical bromine in a simple addition reaction.



The only electrons that might be useful in the kind of attraction we have discussed so far are the lone pair electrons on bromine. But we know from many experiments that electrons flow out of the alkene towards the bromine atom in this reaction—the reverse of what we should expect from electron distribution. The attraction between these molecules is not electrostatic. In fact, we know that reaction occurs because the bromine molecule has an empty orbital available to accept electrons. This is not a localized atomic orbital like that in the BF_3 molecule. It is the antibonding orbital belonging to the Br–Br σ bond: the σ* orbital. There is therefore in this case an attractive interaction between a full orbital (the π bond) and an empty orbital (the σ* orbital of the Br–Br bond). The molecules are attracted to each other because this one interaction is between an empty and a full orbital and leads to bonding, unlike all the other repulsive interactions between filled orbitals. We shall develop this less obvious attraction as the chapter proceeds.

Most organic reactions involve interactions between full and empty orbitals. Many also involve charge interactions, and some inorganic reactions involve nothing but charge attraction. Whatever the attraction between organic molecules, reactions involve electrons moving from one place to another. We call the details of this process the **mechanism of the reaction** and we need to explain some technical terms before discussing this.

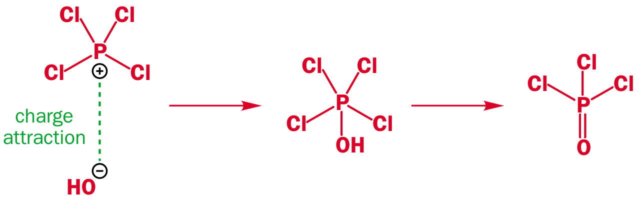
■ Terms such as σ bond, σ* orbital, π bond, π* orbital, lone pair, atomic and molecular orbital, and bonding and antibonding orbital, are all explained in Chapter 4.

Electron flow is the key to reactivity

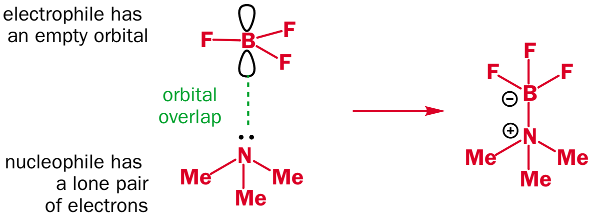
The vast majority of organic reactions are polar in nature. That is to say, electrons flow from one molecule to another as the reaction proceeds. The electron donor is called a **nucleophile** (nucleus-loving) while the electron acceptor is called the **electrophile** (electron-loving). These terms come from the idea of charge attraction as a dominating force in reactions. The nucleophile likes nuclei because they are positively charged and the electrophile likes electrons because they are negatively charged. Though we no longer regard reactions as controlled only by charge interactions, these names have stuck.

▶ Nucleophiles do not really react with the nucleus but with empty electronic orbitals. Even so, electrostatic attraction (and repulsion) may well play a crucial role in determining the course of the process. If a molecule has a positive charge, it is because there are more protons in its nuclei than there are electrons around them.

Examples of reactions where the nucleophile is an anion and the electrophile is a cation and a new bond is formed simply by charge attraction leading to the combination of opposite charges include the reaction of sodium hydroxide with positively charged phosphorus compounds. The new bond between oxygen and phosphorus is formed by the donation of electrons from the nucleophile (hydroxide ion HO^-) to the electrophile (the positively charged phosphorus atom).



More often, reaction occurs when electrons are transferred from a lone pair to an empty orbital as in the reaction between an amine and BF_3. The amine is the nucleophile because of the lone pair of electrons on nitrogen and BF_3 is the electrophile because of the empty p orbital on boron.



The kind of bond formed in these two reactions used to be called a 'dative covalent bond' because both electrons in the bond were donated by the same atom. We no longer classify bonds in this way, but call them σ bonds or π bonds as these are the fundamentally different types of bonds in organic compounds. Most new bonds are formed by donation of both electrons from one atom to another.

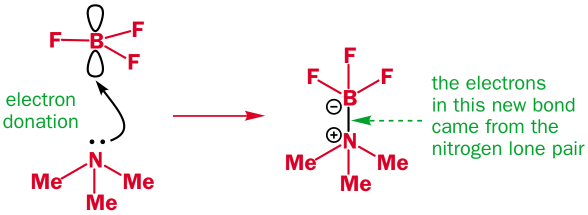
■ A 'dative covalent bond' is just an ordinary σ bond whose electrons happen to come from one atom. Most bonds are formed by electron donation from one atom to another and a classification that makes it necessary to know the history of the molecule is not useful. Forget 'dative bonds' and stick to σ bonds or π bonds.

These simple charge or orbital interactions may be enough to explain simple inorganic reactions but we shall also be concerned with nucleophiles that supply electrons out of bonds and electrophiles that accept electrons into antibonding orbitals. For the moment accept that polar reactions usually involve electrons flowing *from a nucleophile* and *towards an electrophile.*

● **In reaction mechanisms**

- **Nucleophiles donate electrons**
- **Electrophiles accept electrons**

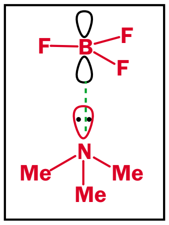
Since we are describing a dynamic process of electron movement from one molecule to another in this last reaction, it is natural to use some sort of arrow to represent the process. Organic chemists use a curved arrow (called a 'curly arrow') to show what is going on. It is a simple and eloquent symbol for chemical reactions.

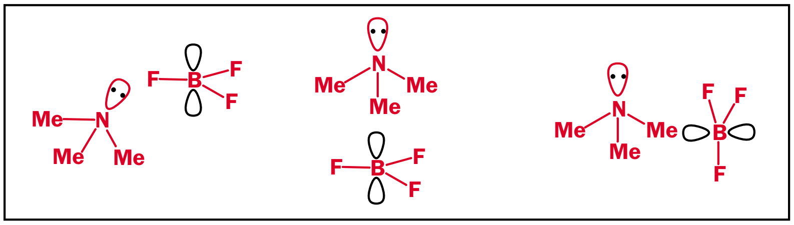


The curly arrow shows the movement of a pair of electrons from nitrogen into the gap between nitrogen and boron to form a new σ bond between those two atoms. This representation, what it means, and how it can be developed into a language of chemical reactions is our main concern in this chapter.

Orbital overlap controls angle of successful attack

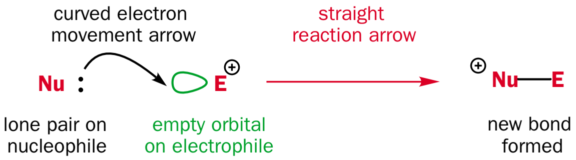
Electrostatic forces provide a generalized attraction between molecules in chemical reactions. In the reaction between chloride anions and sodium cations described above, the way in which these two spherical species approached one another was unimportant because the charges attracted one another from any angle. In most organic reactions the orbitals of the nucleophile and electrophile are directional and so the molecular orbitals of the reacting molecules exert important control. If a new bond is to be formed as the molecules collide, the orbitals of the two species must be correctly aligned in space. In our last example, only if the sp^3 orbital of the lone pair on nitrogen points directly at the empty orbital of the BF_3 can bond formation take place. Other collisions will not lead to reaction. In the first frame a successful collision takes place and a bond can be formed between the orbitals. In the second frame are three examples of unsuccessful collisions where no orbital overlap is possible. There are of course many more unproductive collisions but only one productive collision. Most collisions do not lead to reaction.



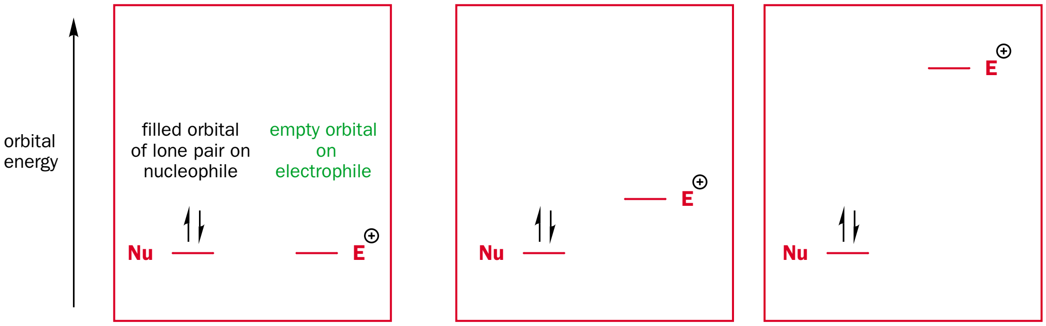


The orbitals must also have about the right amount of energy to interact profitably. Electrons are to be passed from a full to an empty orbital. Full orbitals naturally tend to be of lower energy than empty orbitals—that is after all why they are filled! So when the electrons move into an empty orbital they have to go up in energy and this is part of the activation energy for the reaction. If the energy gap is too big, few molecules will have enough energy to climb it and reaction will be bad. The ideal would be to have a pair of electrons in a filled orbital on the nucleophile and an empty orbital on the electrophile of the same energy. There would be no gap and reaction would be easy. In real life, a small gap is the best we can hope for.

Now we shall discuss a generalized example of a neutral nucleophile, Nu, with a lone pair donating its electrons to a cationic electrophile, E, with an empty orbital. Notice the difference between the curly arrow for electron movement and the straight reaction arrow. Notice also that the nucleophile has given away electrons so it has become positively charged and that the electrophile has accepted electrons so it has become neutral.



If we look at different possible relative energies for the lone pair orbital and the empty orbital, we might have equal energies, a small gap, or a large gap. Just as in Chapter 4, the horizontal lines represent energy levels, the arrows on them represent electrons, and the vertical scale is energy with high energy at the top and low energy at the bottom.



orbitals have the same energy

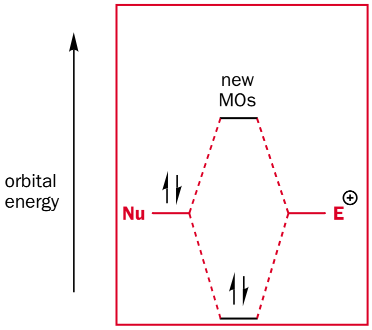
small difference in energy of filled and empty orbitals

large difference in energy of filled and empty orbitals

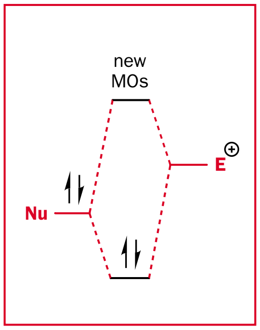
At first this picture suggests that the electrons will have to climb up to the empty orbital if it is higher in energy than the filled orbital. This is not quite true because, when atomic orbitals interact, their energies split to produce two new molecular orbitals, one above and one below the old orbitals. This is the basis for the static structure of molecules described in the last chapter and is also the key to reactivity. In these three cases this is what will happen when the orbitals interact (the new molecular orbitals are shown in black between the old atomic orbitals).

■

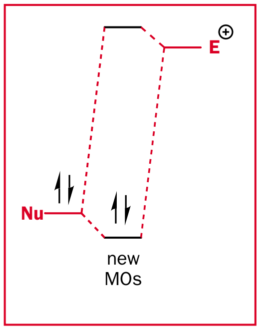
These diagrams of molecular energy levels combining to form new bonding and antibonding orbitals are almost identical to those we used in Chapter 4 to make molecular orbitals from atomic orbitals.



orbitals have the same energy



small difference in energy
of filled and empty orbitals



large difference in energy
of filled and empty orbitals

■ We saw exactly the same response when we combined AOs of different energies to make MOs in Chapter 4.

In each case there is actually a gain in energy when the electrons from the old lone pair drop down into the new stable bonding molecular orbital formed by the combination of the old atomic orbitals. The energy gain is greatest when the two orbitals are the same and least when they are very far apart in energy. The other new MO is higher in energy than either of the old AOs but it does not have to be occupied.

Only the highest-energy occupied orbitals of the nucleophile are likely to be similar in energy to only the lowest unoccupied orbitals of the electrophile. This means that the lower-lying completely filled bonding orbitals of the nucleophile can usually be neglected and only the **highest occupied molecular orbital** (**HOMO**) of the nucleophile and the **lowest unoccupied molecular orbital** (**LUMO**) of the electrophile are relevant. These may be of about the same energy and can then interact strongly. Orbital overlap—of both direction and energy—is therefore an important requirement for successful reaction between two organic molecules.

● Molecules repel each other because of their outer coatings of electrons.

Molecules attract each other because of:

- attraction of opposite charges
- overlap of high-energy filled orbitals with low-energy empty orbitals

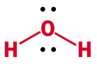
For reaction, molecules must approach each other so that they have:

- enough energy to overcome the repulsion
- the right orientation to use any attraction

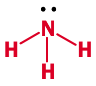
We need now to look at which types of molecules are nucleophiles and which types are electrophiles. When you consider the reactivity of any molecule, this is the first question you should ask: is it nucleophilic or electrophilic?

Nucleophiles donate high-energy electrons to electrophiles

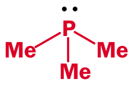
Nucleophiles are either negatively charged or neutral species with a pair of electrons in a high energy filled orbital that they can donate to electrophiles. The most common type of nucleophile has a nonbonding lone pair of electrons. Usually these are on a heteroatom such as O, N, S, or P.



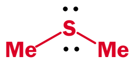
water



ammonia



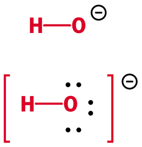
trimethylphosphine

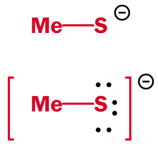


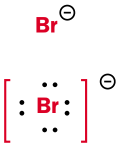
dimethylsulfide

These four neutral molecules, ammonia, water, trimethylphosphine, and dimethylsulfide, all have lone pairs of electrons in sp^3 orbitals and in each case this is the donor or nucleophilic orbital. The group VI atoms (O and S) have two lone pairs of equal energy. These are all nonbonding electrons and therefore higher in energy than any of the bonding electrons.

Anions are often nucleophiles too and these are also usually on heteroatoms such as O, S, or halogen which may have several lone pairs of equal energy. The first diagram for each of our examples shows the basic structure and the second diagram shows all the lone pairs. It is not possible to allocate the negative charge to a particular lone pair as they are the same.

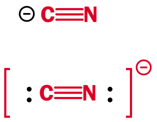

hydroxide

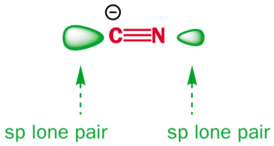

methane thiolate


bromide

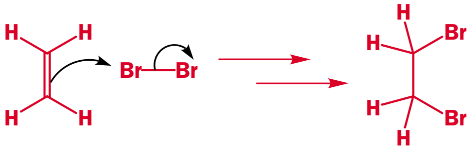
There are a few examples of carbon nucleophiles with lone pairs of electrons, the most famous being the cyanide ion. Though linear cyanide has a lone pair on nitrogen and one on carbon, the nucleophilic atom is usually anionic carbon rather than neutral nitrogen as the sp orbital on carbon has a higher energy than that on the more electronegative nitrogen. Most anionic nucleophiles containing carbon have a heteroatom as the nucleophilic atom such as the anion methane thiolate shown above.

■ This point will be important in Chapter 6 as well.

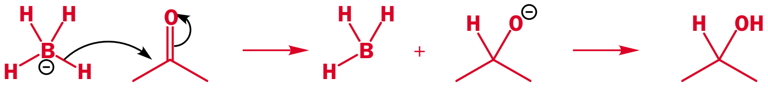

cyanide ion



Neutral carbon electrophiles usually have a π bond as the nucleophilic portion of the molecule. When there are no lone pair electrons to supply high-energy nonbonding orbitals, the next best is the lower-energy filled π orbitals rather than the even lower-energy σ bonds. Simple alkenes are weakly nucleophilic and react with strong electrophiles such as bromine. In Chapter 20 we shall see that the reaction starts by donation of the π electrons from the alkene into the σ* orbital of the bromine molecule (which breaks the Br–Br bond) shown here with a curly arrow. After more steps the dibromoalkane is formed but the molecules are attracted by overlap between the full π orbital and the empty σ* orbital.



It is possible for σ bonds to act as nucleophiles and we shall see later in this chapter that the borohydride anion, BH_4^-, has a nucleophilic B–H bond and can donate those electrons into the π* orbital of a carbonyl compound breaking that bond and eventually giving an alcohol as product. The first stage of the reaction has electrons from the B–H single bond of nucleophilic anion BH_4^-, which lacks lone pair electrons or π bonds, as the nucleophile.



In this section you have seen lone pairs on anions and neutral molecules acting as nucleophiles and, more rarely, π bonds and even σ bonds able to do the same job. In each case the nucleophilic electrons came from the HOMO—the highest occupied molecular orbital—of the molecule. Don't worry if you find the curly arrows strange at the moment. They will soon be familiar. Now we need to look at the other side of the coin—the variety of electrophiles.

Electrophiles have a low-energy vacant orbital

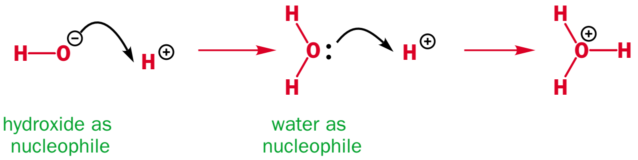
Electrophiles are neutral or positively charged species with an empty atomic orbital (the opposite of a lone pair) or a low-energy antibonding orbital. The simplest electrophile is the proton, H^+, a species without any electrons at all and a vacant 1s orbital. It is so reactive that it is hardly ever found and almost any nucleophile will react with it.


proton

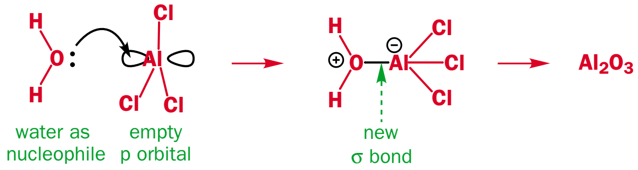

empty 1s orbital


reaction with anionic nucleophile

Each of the nucleophiles we saw in the previous section will react with the proton and we shall look at two of them together. Hydroxide ion combines with a proton to give water. This reaction is governed by charge control. Then water itself reacts with the proton to give H_3O^+, the true acidic species in all aqueous strong acids.

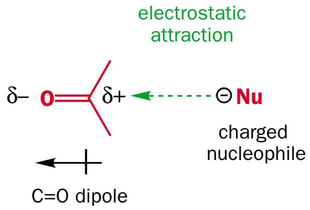


We normally think of protons as acidic rather than electrophilic but an acid is just a special kind of electrophile. In the same way, Lewis acids such as BF_3 or $AlCl_3$ are electrophiles too. They have empty orbitals that are usually metallic p orbitals. We saw above how BF_3 reacted with Me_3N. In that reaction BF_3 was the electrophile and Me_3N the nucleophile. Lewis acids such as $AlCl_3$ react violently with water and the first step in this process is nucleophilic attack by water on the empty p orbital of the aluminium atom. Eventually alumina (Al_2O_3) is formed.

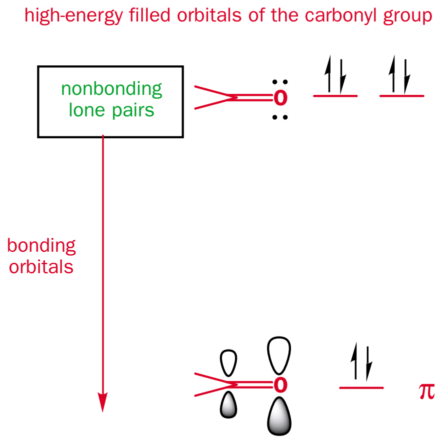


Protic and Lewis acids

Protic acids (also known as Brønsted acids) are electrophiles (like HCl) that can donate protons (H^+) to nucleophiles. They will be discussed in detail in Chapter 8. Lewis acids are also electrophiles but they donate more complicated cations to nucleophiles. They are usually metal halides such as LiCl, BF_3, $AlCl_3$, $SnCl_4$, and $TiCl_4$. We shall meet them in many later chapters, particularly in Chapters 22–8 when we discuss carbon–carbon bond formation.

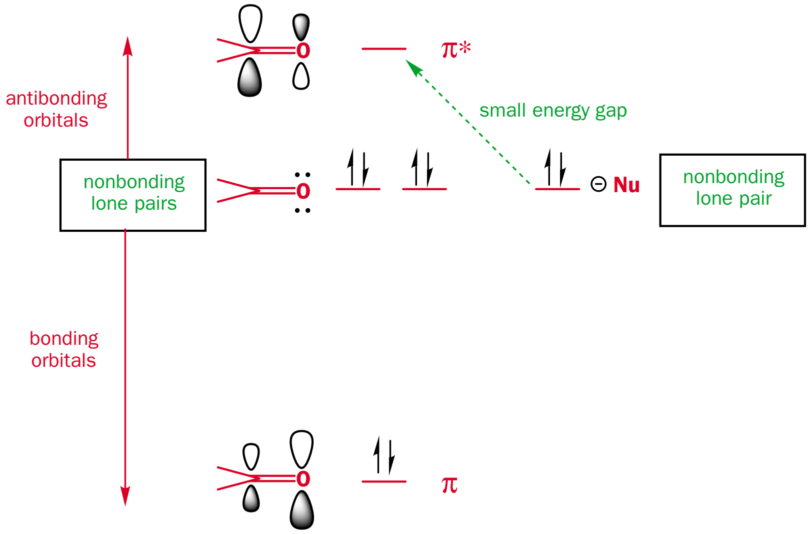


Few organic compounds have vacant atomic orbitals and most organic electrophiles have low-energy antibonding orbitals. The most important are π* orbitals as they are lower in energy than σ* orbitals and the carbonyl group (C=O) is the most important of these—indeed it is the most important functional group of all. It has a low-energy π* orbital ready to accept electrons and also a partial positive charge on the carbon atom. Previously we said that charge attraction helped nucleophiles to find the carbon atom of the carbonyl group.

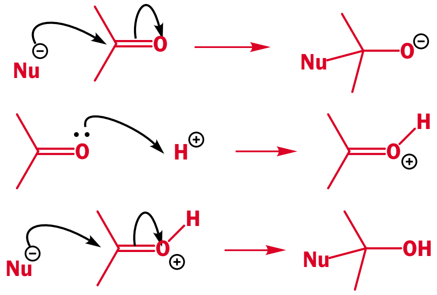


Charge attraction is important in carbonyl reactions but so are the orbitals involved. Carbonyl compounds have a low-energy bonding π orbital. Carbonyl compounds have a dipole because in this filled orbital the electrons are more on electronegative oxygen than on carbon. The same reason (electronegative oxygen) makes this an exceptionally low-energy orbital and the carbonyl group a very stable structural unit. This orbital is rarely involved in reactions. Going up the energy scale we next have two degenerate (equal in energy) lone pairs in nonbonding orbitals. These are the highest-energy electrons in the molecule (HOMO) and are the ones that react with electrophiles.

When we consider the carbonyl group as an electrophile, we must look at antibonding orbitals too. The only one that concerns us is the relatively low-energy π* orbital of the C=O double bond (the LUMO). This orbital is biased towards the carbon to compensate for the opposite bias in the filled π orbital. How do we know this if there are no electrons in it? Simply because nucleophiles, whether charged or not, attack carbonyl groups at the carbon atom. They get the best overlap with the larger orbital component of the π* orbital.



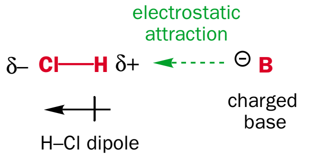
So now we can draw a mechanism for the attack of a nucleophile on the carbonyl group. The lone pair electrons on the nucleophile move into the π* orbital of the C=O double bond and so break the π bond, though not, of course, the σ bond. Here is that process in curly arrow terms.

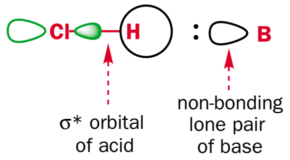


The lone pair electrons on oxygen interact better with empty orbitals such as the 1s of the proton and so carbonyl compounds are protonated on oxygen.

The resulting cation is even more electrophilic because of the positive charge but nucleophiles still attack the carbon atom of the carbonyl group because the π* orbital still has more contribution from carbon. The positive charge is neutralized even though the nucleophile does not attack the positively charged atom.

Even σ bonds can be electrophilic if the atom at one end of them is sufficiently electronegative to pull down the energy of the σ* orbital. Familiar examples are acids where the acidic hydrogen atom is joined to strongly electronegative oxygen or a halogen thus providing a dipole moment and a relatively low-energy σ* orbital.





These two diagrams suggest two different ways of looking at the reaction between a base and an acid, but usually both interactions are important. Notice that an acid is just an electrophile that has an electrophilic hydrogen atom and a base is just a nucleophile that acts on a hydrogen atom. This question is explored more in Chapter 8. Bonds between carbon and halogen are also polarized in some cases though the electronegativity difference is sometimes very small.

It is easy to exaggerate the importance of single-bond polarization. The electronegativity difference between H and Cl is 0.9 but that

Quick guide to important electronegativities

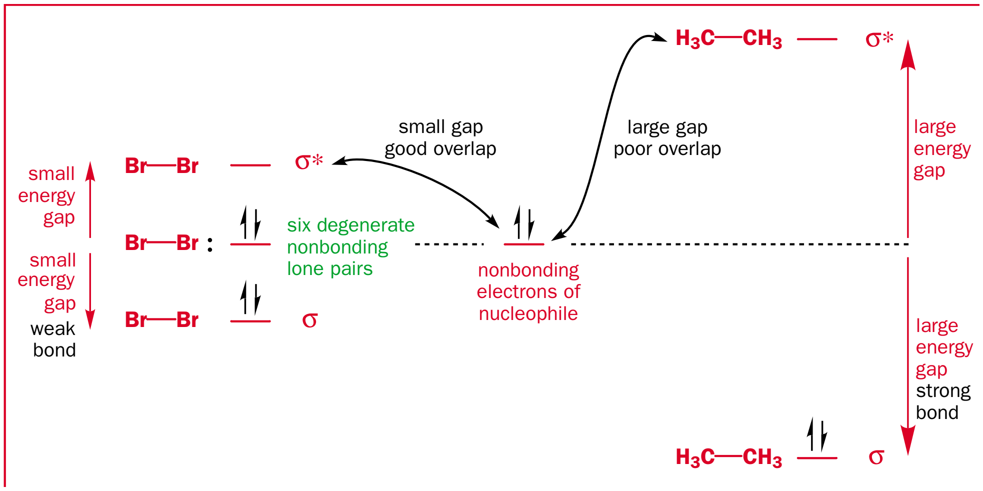
		H 2.1				
Li 1.0		**B** 2.0	**C** 2.5	**N** 3.0	**O** 3.5	**F** 4.0
	Mg 1.2	**Al** 1.5	**Si** 1.8	**P** 2.1	**S** 2.5	**Cl** 3.0
						Br 2.8
						I 2.5

▶

These are Pauling electronegativities, calculated by Linus Pauling (1901–94) who won the chemistry Nobel prize in 1954 and the Nobel peace prize in 1983 and from whose ideas most modern concepts of the chemical bond are derived. Born in Portland, Oregon, he worked at 'CalTech' (the California Institute of Technology at Pasadena) and had exceptionally wide-ranging interests in crystallography, inorganic chemistry, protein structure, quantum mechanics, nuclear disarmament, politics, and taking vitamin C to prevent the common cold.

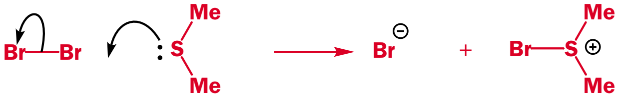
between C and Br only 0.3 while the C–I bond is not polarized at all. When carbon–halogen σ bonds act as electrophiles, polarity hardly matters but a relatively low-energy σ* orbital is vitally important. The bond strength is also important in these reactions too as we shall see.

Some σ bonds are electrophilic even though they have no dipole at all. The halogens such as bromine (Br_2) are examples. Bromine is strongly electrophilic because it has a very weak Br–Br σ bond. Symmetrical bonds have the energies of the σ orbital and the σ* orbital roughly evenly distributed about the nonbonding level. A weak symmetrical σ bond means a small energy gap while a strong symmetrical σ bond means a large energy gap. Bromine is electrophilic but carbon–carbon σ bonds are not. Reverting to the language of Chapter 4, we could say that the hydrocarbon framework is made up of strong C–C bonds with low-energy populated and high-energy unpopulated orbitals, while the functional groups react because they have low LUMOs or high HOMOs.



An example would be the rapid reaction between a sulfide and bromine. No reaction at all occurs between a sulfide and ethane or any other simple C–C σ bond. Lone pair electrons are donated from sulfur into the Br–Br σ* orbital, which makes a new bond between S and Br and breaks the old Br–Br bond.

■ Notice how putting charges in circles (Chapter 2) helps here. There is no problem in distinguishing the charge on sulfur (in a ring) with the plus sign (not in a ring) linking the two products of the reaction.



Summary: interaction between HOMO and LUMO leads to reaction

Organic reactions occur when the HOMO of a nucleophile overlaps with the LUMO of the electrophile to form a new bond. The two electrons in the HOMO slot into the empty LUMO. The reacting species may be initially drawn together by electrostatic interaction of charges or dipoles but this is not necessary. Thus at this simplest of levels *molecular recognition* is required for reaction. The two components of a reaction must be matched in terms of both charge–charge attraction and the energy and orientation of the orbitals involved.

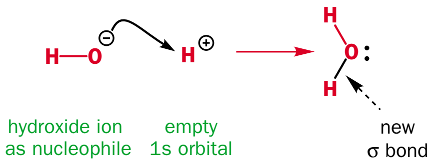
Nucleophiles may donate electrons (in order of preference) from a lone pair, a π bond, or even a σ bond and electrophiles may accept electrons (again in order of preference) into an empty orbital or into the antibonding orbital of a π bond (π* orbital) or even a σ bond (σ* orbital). These antibonding orbitals are of low enough energy to react if the bond is very polarized by a large electronegativity difference between the atoms at its ends or, even for unpolarized bonds, if the bond is weak.

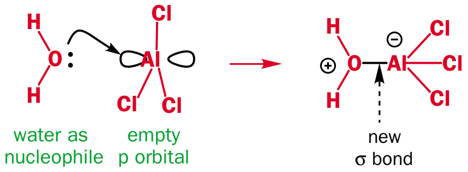
The hydrocarbon framework of organic molecules is unreactive. Functional groups such as NH_2 and OH are nucleophilic because they have nonbonding lone pairs. Carbonyl compounds and alkyl halides are electrophilic functional groups because they have low-energy LUMOs (π* for C=O and σ* for C–X, respectively).

Organic chemists use curly arrows to represent reaction mechanisms

You have seen several examples of curly arrows so far and you may already have a general idea of what they mean. The representation of organic reaction mechanisms by this means is so important that we must now make quite sure that you do indeed understand exactly what is meant by a curly arrow, how to use it, and how to interpret mechanistic diagrams as well as structural diagrams.

A **curly arrow** represents the *actual movement of a pair of electrons* from a filled orbital into an empty orbital. You can think of the curly arrow as representing a pair of electrons thrown, like a climber's grappling hook, across from where he is standing to where he wants to go. In the simplest cases, the result of this movement is to form a bond between a nucleophile and an electrophile. Here are two examples we have already seen in which lone pair electrons are transferred to empty atomic orbitals.

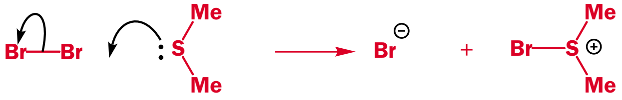




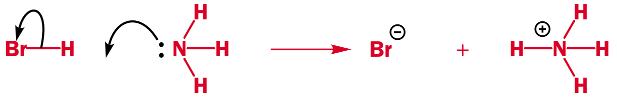
Note the exact position of the curly arrow as the value of this representation lies in the precision and uniformity of its use. The arrow always starts with its tail on the source of the moving electrons, representing the filled orbital involved in the reaction. The head of the arrow indicates the final destination of the pair of electrons—the new bond between oxygen and hydrogen or oxygen and aluminium in these examples. As we are forming a new bond, the head of the arrow should be drawn to a point on the line between the two atoms.

▶ Some chemists prefer to place this point halfway between the atoms but we consider that the representation is clearer and more informative if the arrowhead is closer to the atom to which the new bond is forming. For these examples the difference is minimal and either method is completely clear but in more complex situations, our method prevents ambiguity as we shall see later. We shall adopt this convention throughout this book.

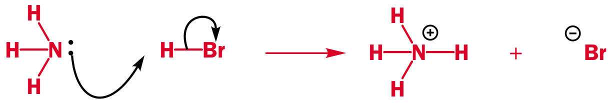
When the nucleophile attacks an antibonding orbital, such as the weak Br–Br bond we have just been discussing, we shall need two arrows, one to make the new bond and one to break the old.



The bond-making arrow is the same as before but the bond-breaking arrow is new. This arrow shows that the two electrons in the bond move to one end (a bromine atom) and turn it into an anion. This arrow should start in the centre of the bond and its head should rest on the atom (Br in this case) at the end of the bond. Another example would be the attack of a base on the strong acid HBr.

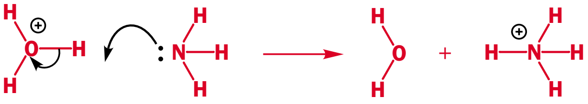


It is not important how much curvature you put into the arrows or whether they are above or below the gaps of the bonds, both on the same side, or on opposite sides so long as they begin and end in the right places. All that matters is that someone who reads your arrows should be able to deduce exactly what is happening in the reaction from your arrows. We could have drawn the ammonia/HBr reaction like this if we had wished.



Charge is conserved in each step of a reaction

In all these examples we have reacted neutral molecules together to form charged species. Because the starting materials had no overall charge, neither must the products. If we start with neutral molecules and make a cation, we must make an anion too. Charge cannot be created or destroyed. If our starting materials have an overall charge—plus or minus—then the same charge must appear in the products.



starting materials have overall positive charge

products must also have overall positive charge

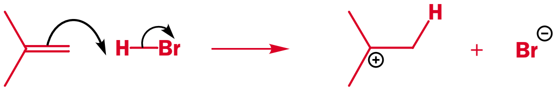
When it is a π bond that is being broken rather than a σ bond, only the π bond is broken and the σ bond should be left in place. This is what commonly happens when an electrophilic carbonyl group is attacked by a nucleophile. Just as in the breaking of a σ bond, start the arrow in the middle of the π bond and end by putting the arrowhead on the more electronegative atom, in this case oxygen rather than carbon.

C–O σ bond remains

π bond is broken

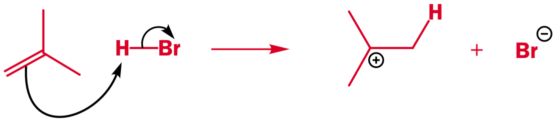
In this case the starting materials had an overall negative charge and this is preserved as the oxyanion in the product. The charge disappears from the hydroxide ion because it is now sharing a pair of electrons with what was the carbonyl carbon atom and a charge appears on what was the carbonyl oxygen atom because it now has both of the electrons in the old π bond.

Electrons can be donated from π bonds and from σ bonds too. The reaction of an alkene with HBr is a simple example of a C–C π bond as nucleophile. The first arrow (on the nucleophile) starts in the middle of the π bond and goes into the gap between one of the carbon atoms and the hydrogen atom of HBr. The second arrow (on the electrophile) takes the electrons out of the H–Br σ bond and puts them on to the bromine atom to make bromide ion. This sort of reaction make us place alkenes among the functional groups as well as as part of the framework of organic molecules.

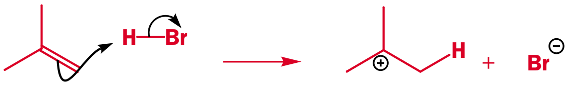


■ We have drawn in the hydrogen atom that was part of HBr. It is not necessary to do this but you may wish to show what has happened to one particular hydrogen atom among many in a reaction mechanism and this is another instance of ignoring, for a good reason, one of the guidelines from Chapter 2!

Notice that it was important to draw the two reagents in the right orientation since both are unsymmetrical and we want our arrow to show which end of the alkene reacts with which end of HBr. If we had drawn them differently we should have had trouble drawing the mechanism. Here is a less satisfactory representation.



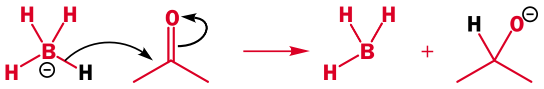
If you find yourself making a drawing like this, it is worth having another go to see if you can be clearer. Drawing mechanisms is often rather experimental—try something and see how it looks: if it is unclear, try again. One way to avoid this particular problem is to draw an *atom-specific* curly arrow passing through the atom that reacts. Something like this will do.



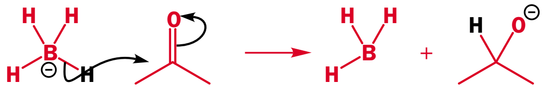
This reaction does not, in fact, stop here as the two ions produced (charge conservation requires one cation and one anion in this first reaction) now react with each other to form the product of the reaction. This reaction is pretty obvious as the anion is the nucleophile and the cation, with its empty p orbital, is the electrophile.

The reaction that occurs between the alkene and HBr occurs in two stages—the formation of the ions and their combination. Many reactions are like this and we call the two stages **steps** so that we talk about 'the first step' and 'the second step', and we call the ions **intermediates** because they are formed in one step and disappear in the next. We shall discuss these intermediates in several later chapters (for example, 17 and 19).

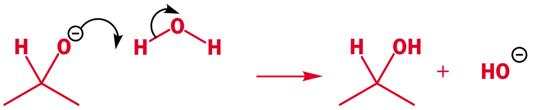
When σ bonds act as nucleophiles, the electrons also have to go to one end of the σ bond as they form a new bond to the electrophile. We can return to an earlier example, the reaction of sodium borohydride ($NaBH_4$) with a carbonyl compound, and complete the mechanism. In this example, one of the atoms (the hydrogen atom) moves away from the rest of the BH_4 anion and becomes bonded to the carbonyl compound. The LUMO of the electrophile is, of course, the π* orbital of the C=O double bond.



The arrow on the nucleophile should again start in the middle of the bond that breaks and show which atom (the black H in this case) is transferred to the electrophile. The second arrow we have seen before. Here again you could use an atom-specific arrow to make it clear that the electrons in the σ bond act as a nucleophile through the hydrogen and not through the boron atom.



This reaction also occurs in two steps and the oxyanion is an intermediate, not a product. The reaction is normally carried out in water and the oxyanion reacts with water by proton transfer.

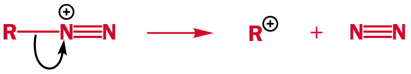


We shall discuss this reaction, the reduction of carbonyl compounds by $NaBH_4$, in detail in Chapter 6.

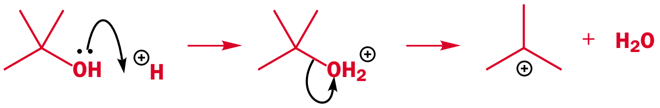
The decomposition of molecules

So far we have described reactions involving the combination of one molecule with another. Many reactions are not like this but involve the spontaneous decomposition of one molecule by itself without any assistance from any other molecule. In these reactions there is no electrophile or nucleophile. The usual style of reaction consists of a weak, often polarized σ bond breaking to give two new molecules or ions. The dissociation of a strong acid HX is a simple example.

In organic chemistry spontaneous dissociation of diazonium salts, compounds containing the N_2^+ group, occurs very easily because one of the products, nitrogen gas ('dinitrogen') is very stable. It does not much matter what R is (alkyl or aryl); this reaction happens spontaneously at room temperature.



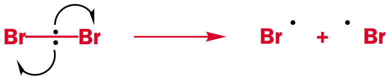
This is not, of course, the end of the reaction as R^+ is very reactive and we shall see the sort of things it can do in Chapters 17 and 19. More commonly, some sort of catalysis is involved in decomposition reactions. An important example is the decomposition of tertiary alcohols in acid solution. The carbon–oxygen bond of the alcohol does not break by itself but, after the oxygen atom has been protonated by the acid, decomposition occurs.



This two-step mechanism is not finished because the positive ion (one particular example of R^+) reacts further (Chapter 17). In the decomposition step the positive charge on the oxygen atom as well as the fact that the other product is water helps to break the strong C–O σ bond. In these three examples, the functional group that makes off with the electrons of the old σ bond (X, N_2^+, and OH_2^+) is called the **leaving group**, and we shall be using this term throughout the book. The spontaneous decomposition of molecules is one of the clearest demonstrations that curly arrows mean the movement of two electrons. Chemical reactions are dynamic processes, molecules really do move, and electrons really do leave one atomic or molecular orbital to form another.

▶ Each bromine radical has an unpaired electron in an atomic orbital.

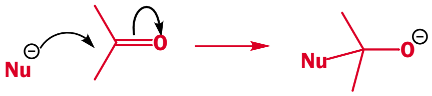
These three examples all have the leaving group taking both electrons from the old σ bond. This type of decomposition is sometimes called **heterolytic fission** or simply **heterolysis** and is the most common in organic chemistry. There is another way that a σ bond can break. Rather than a pair of electrons moving to one of the atoms, one electron can go in either direction. This is known as **homolytic fission** as two species of the same charge (neutral) will be formed. It normally occurs when similar or indeed identical atoms are at each end of the σ bond to be broken. Both fragments have an unpaired electron and are known as radicals. This type of reaction occurs when bromine gas is subjected to sunlight.



The weak Br–Br bond breaks to form two bromine **radicals**. This can be represented by two single-headed curly arrows, **fish hooks**, to indicate that only one electron is moving. This is virtually all you will see of this special type of curly arrow until we consider the reactions of radicals in more detail (Chapter 39). When you meet a new reaction you should assume that it is an ionic reaction and use two-electron arrows unless you have a good reason to suppose otherwise.

Curly arrows also show movement of electrons within molecules

So far all the mechanisms we have drawn have used only one or two arrows in each step. In fact, there is no limit to the number of arrows that might be involved and we need to look at some mechanisms with three arrows. The third arrow in such mechanisms usually represents movement of electrons inside of the reacting molecules. Some pages back we drew out the addition of a nucleophile to a carbonyl compound.

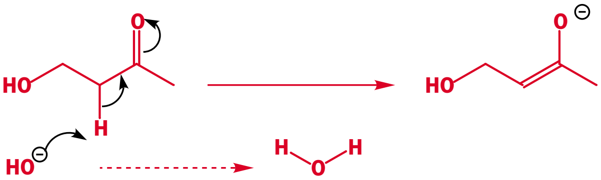


▪ Don't be alarmed—these mechanisms will all be discussed in full later in the book, this particular one in Chapter 10.

This is a two-arrow mechanism but, if we lengthen the structure of the carbonyl compound by adding a double bond in the right position, we can add the nucleophile to a different position in the molecule by moving electrons within the molecule using a third arrow.

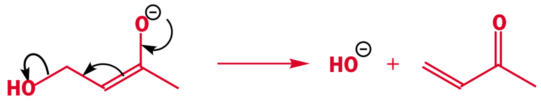
The first arrow from the nucleophile makes a new σ bond and the last breaks the carbonyl π bond. The middle arrow just moves the C–C π bond along the molecule. If you inspect the product you will see that its structure follows precisely from the arrows. The middle arrow starts in the middle of a π bond and ends in the middle of a σ bond. All it does is to move the π electrons along the molecule. It turns the old π bond into a σ bond and the old σ bond into a π bond. We shall discuss this sort of reaction in Chapter 10.

In some mechanisms there is a second step in the mechanism and both are three-arrow processes. Here is the first step in such a mechanism. See if you can understand each arrow before reading the explanation in the next paragraph.

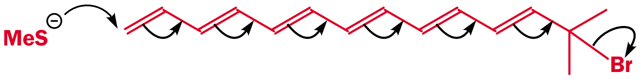


The arrow from the hydroxide ion removes a proton from the molecule making a new O–H bond in a molecule of water. The middle arrow moves the electrons of a C–H bond into a C–C bond making it into a π bond and the third arrow polarizes the carbonyl π bond leaving an oxyanion as the product. Charge is conserved—an anion gives an anion. In fact this 'product' is only an intermediate and the second step also involves three arrows.

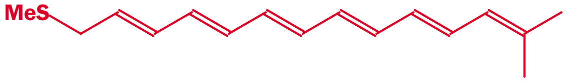
Starting from the oxyanion, the first arrow re-forms the carbonyl group, the middle arrow moves a π bond along the molecule, and the third arrow breaks a C–O σ bond releasing hydroxide ion as one of the products of the reaction. We shall meet this sort of reaction in detail later (Chapters 19 and 27).



Mostly for entertainment value we shall end this section with a mechanism involving no fewer than eight arrows. See if you can draw the product of this reaction without looking at the result.



The first arrow forms a new C–S σ bond and the last arrow breaks a C–Br σ bond but all the rest just move π bonds along the molecule. The product is therefore:



We shall not be discussing this reaction anywhere in the book! We have included it just to convince you that, once you understand the principle of curly arrows, you can understand even very complicated mechanisms quite easily. At this stage we can summarize the things you have learned about interpreting a mechanism drawn by someone else.

Summary: what do curly arrows mean?

- A curly arrow shows the movement of a pair of electrons
- The tail of the arrow shows the source of the electron pair, which will be a filled orbital (HOMO)
  - such as a lone pair or a π bond or a σ bond
- The head of the arrow indicates the ultimate destination of the electron pair which will either be:
  - an electronegative atom that can support a negative charge (a *leaving group*)
  - *or* an empty orbital (LUMO) when a new bond will be formed
  - *or* an antibonding orbital (π* or σ*) when that bond will break
- Overall charge is always conserved in a reaction. Check that your product obeys this rule

Now would be a good time to do Problems 1 and 2 at the end of the chapter, which will give you practice in the interpretation of mechanisms.

Drawing your own mechanisms with curly arrows

Curly arrows must be drawn carefully! The main thing you need to remember is that curly arrows must start where there is a pair of electrons and end somewhere where you can leave a pair of electrons without drawing an absurd structure. That sounds very simple—and it is—but you need some practice to see what it means in detail in different circumstances. Let us look at the implications with a reaction whose products are given: the reaction of triphenylphosphine with methyl iodide.

$$MeI + Ph_3P \longrightarrow Ph_3\overset{\oplus}{P}{-}Me + I^{\ominus}$$

First observe what has happened: a new bond has been formed between the phosphorus atom and the methyl group and the carbon–iodine bond has been broken. Arrows represent movement of electron pairs *not* atoms so the reactants must be drawn within bonding distance before the mechanism can be drawn. This is analogous to the requirement that molecules must collide before they can react. First draw the two molecules so that the atoms that form the new bond (P and C) are near each other and draw out the bonds that are involved (that is, replace 'MeI' with a proper chemical structure).

$$Ph_3P \quad CH_3{-}I$$

Now ask: which is the electrophile and which the nucleophile (and why)? The phosphorus atom has a lone pair and the carbon atom does not so Ph_3P must be the nucleophile and the C–I bond of MeI must be the electrophile. All that remains is to draw the arrows.

Admittedly, that was quite an easy mechanism to draw but you should still be pleased if you succeeded at your first try.

Warning! Eight electrons is the maximum for B, C, N, or O

We now ought to spell out one thing that we have never stated but rather assumed. Most organic atoms, if they are not positively charged, have their full complement of electrons (two in the case of hydrogen, eight in the cases of carbon, nitrogen, and oxygen) and so, if you make a new bond to one of those elements, *you must also break an existing bond*. Suppose you just 'added' Ph_3P to MeI in this last example without breaking the C–I bond: what would happen?

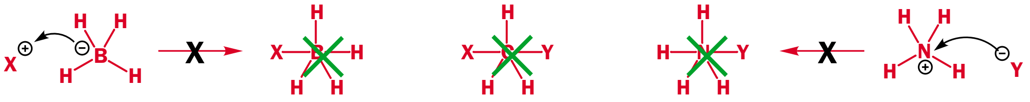
wrong mechanism

impossible structure
carbon has five bonds

This structure must be wrong because carbon cannot have five bonds—if it did it would have ten electrons in the 2s and the 2p orbitals. As there are only four of those (2s, $2p_x$, $2p_y$, and $2p_z$) and they can have only two electrons each, eight electrons is the maximum and that means that four bonds is the maximum.

● If you make a new bond to uncharged H, C, N, or O you must also break one of the existing bonds in the same step.

There is a nasty trap when a charged atom has its full complement of electrons. Since BH_4^- and NH_4^+ are isoelectronic with methane and have four σ bonds and hence eight electrons, no new bonds can be made to B or N. The following attractive mechanisms are impossible because boron has no lone pair in BH_4^- and nitrogen has no empty orbital in NH_4^+.



impossible reaction

impossible structure
boron has five bonds

impossible structure
carbon has five bonds

impossible structure
nitrogen has five bonds

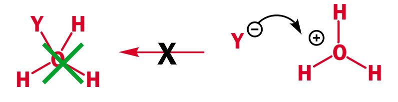
impossible reaction

Reactions with BH_4^- always involve the loss of H and a pair of electrons using the BH bond as nucleophile and reactions with NH_4^+ always involve the loss of H without a pair of electrons using the NH bond as electrophile.

correct mechanism

correct mechanism

Similarly, nucleophiles do not attack species like H_3O^+ at oxygen, even though it is the oxygen atom that carries the positive charge. Reaction occurs at one of the protons, which also neutralizes the positive charge. Or, to put it another way, H_3O^+ is an acid (electrophilic at hydrogen) and not electrophilic at oxygen.



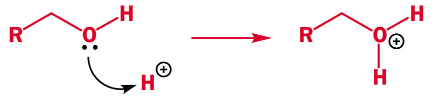
impossible structure
oxygen has four bonds

impossible reaction

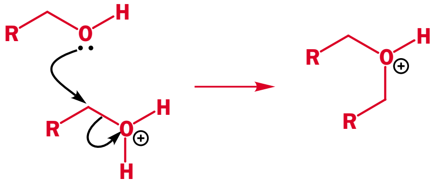
correct mechanism

Try a simple example: primary alcohols can be converted into symmetrical ethers in acid solution. Suggest a mechanism for this acid-catalysed conversion of one functional group into another.

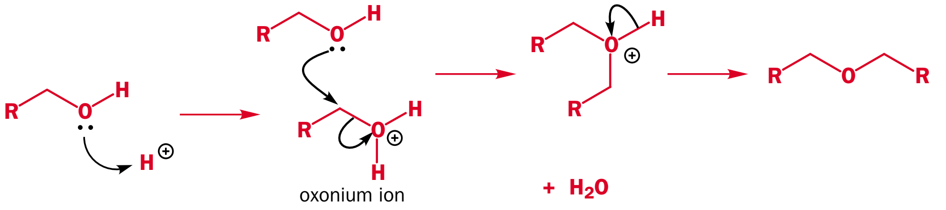
The reaction must start by the protonation of something and the only candidate is the oxygen atom as it alone has lone pair electrons. This gives us a typical oxonium ion with three bonds to oxygen and a full outer shell of eight electrons.



To make the ether a second molecule of alcohol must be added but we must not now be tempted to attack the positively charged oxygen atom with the nucleophilic OH group. The second molecule could attack a proton, but that would just make the same molecules. Instead it must attack at carbon expelling a molecule of water as a leaving group and creating a new oxonium ion.

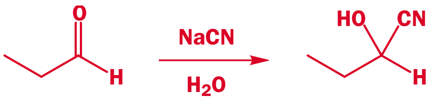


Finally, the loss of the proton from the new oxonium ion gives the ether. Though this is a three-step mechanism, two of the steps are just proton transfers in acidic solution and the only interesting step is the middle one. Here is the whole mechanism.



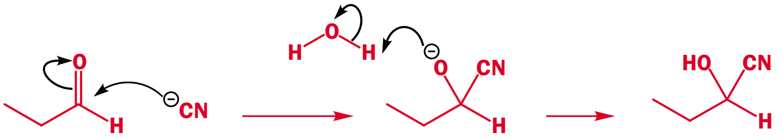
Drawing a two-step mechanism: cyanohydrin formation

Now what about this slightly more complicated example? Sodium cyanide is added to a simple aldehyde in aqueous solution. The product is a **cyanohydrin** and we shall discuss this chemistry in Chapter 6.



This reaction is presented in a style with which you will become familiar. The organic starting material is written first and then the reagent over the reaction arrow and the solvent under it. We must decide what happens. NaCN is an ionic solid so the true reagent must be cyanide ion. As it is an anion, it must be the nucleophile and the carbonyl group must be the electrophile. Let us try a mechanism.

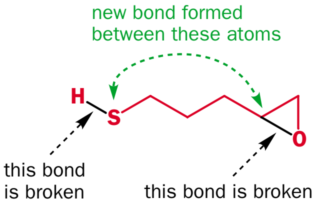
This is a good mechanism but it doesn't quite produce the product. There must be a second step in which the oxyanion picks up a proton from somewhere. The only source of protons is the solvent, water, so we can write:



This is the complete mechanism and we can even make a prediction about the reaction conditions from it. The second step needs a proton and water is not a very good proton donor. A weak acid as catalyst would help.

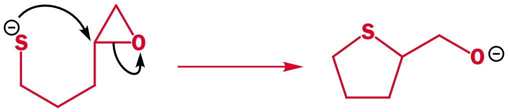
Now for a real test: can you draw a mechanism for this reaction?

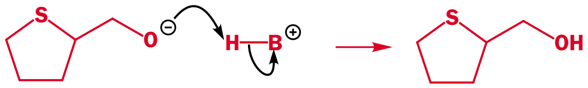
You might well protest that you don't know anything about the chemistry of three-membered rings or of either of the functional groups, SH and cyclic ether. Be that as it may, you can still draw a mechanism for the reaction. It is important that you are prepared to try your hand at mechanisms for new reactions as you can learn a lot this way. Ask first of all: which bonds have been formed and which broken? Clearly the S–H bond has been broken and a new S–C bond formed. The three-membered ring has gone by the cleavage of one of the C–O bonds. The main chain of carbon atoms is unchanged. We might show these ideas in some way such as this.



Now you could continue in many ways. You might say 'what breaks the SH bond?' This must be the role of the base as a base removes protons. You might realize that the reaction cannot happen while the sulfur atom is so far away from the three-membered ring (no chance of a collision) and redraw the molecule so that the reaction can happen.

Now draw the mechanism. It is easy once you have done the preparatory thinking. The sulfur anion must be the nucleophile so the C–O bond in the three-membered ring must be the electrophile. Here goes!



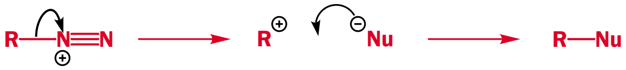


That is not quite the product so we must add a proton to the oxyanion. Where can the proton come from? It must be the proton originally removed by the base as there is no other. We can write B for the base and hence BH$^+$ for the base after it has captured a proton.

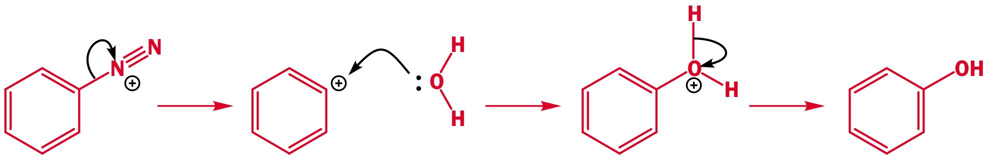
Your mechanism probably didn't look as neat as the printed version but, if you got it roughly right, you should be proud. This is a three-step mechanism involving chemistry unknown to you and yet you could draw a mechanism for it. Are you using coloured arrows, by the way? We are using black arrows on red diagrams but the only point of that is to make the arrows stand out. We suggest you use any colour for your arrows that contrasts with your normal ink.

Decide on a 'push' or a 'pull' mechanism

In one step of a reaction mechanism electrons flow from a site rich in electrons to an electron-deficient site. When you draw a mechanism you must make sure that the electrons flow in one direction only and neither meet at a point nor diverge from a point. One way to do this is to decide whether the mechanism is 'pushed' by, say, a lone pair or an anion or whether it is 'pulled' by, say, a cation, an empty orbital, or by the breaking of a reactive weak π bond or σ bond. This is not just a device either. Extremely reactive molecules, such as fluorine gas, F_2, react with almost anything—in this case because of the very electrophilic F–F σ bond (low energy F–F σ^* orbital). Reactions of F_2 are 'pulled' by the breaking of the F–F bond. The nearest thing in organic chemistry is probably the reactions of carbon cations such as those formed by the decomposition of diazonium salts.



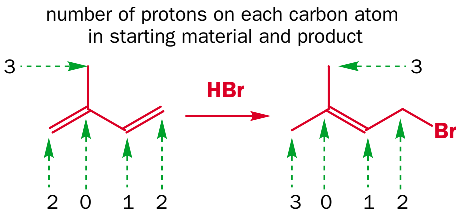
In the first step the electrons of the σ bond are pulled away by the positive charge and the very stable leaving group, N_2. In the second step lone pair electrons are pulled into the very reactive cation by the nonbonding empty orbital on carbon. Even very weak nucleophiles such as water will react with such cations as a real example shows.



In all our previous examples we have drawn the first arrow from the nucleophile, anion, lone pair, or whatever and pushed the electrons along the chain of arrows. This is a natural thing to do; indeed the skill of drawing mechanisms is sometimes derisively referred to as 'electron pushing', but some mechanisms are more easily understood as 'electron pulling'. In general, if a cation, an acid, or a Lewis acid is a reagent or a catalyst, the reaction is probably *pulled.* If an anion or a base is involved as a reagent, the reaction is probably *pushed.* In any case it isn't so important which approach you adopt as that you should do one or the other and not muddle them up.

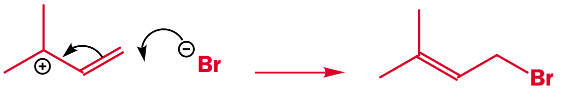
A more interesting example of a pull mechanism is the reaction of isoprene (2-methylbutadiene) with HBr. The product is an unsaturated alkyl bromide (a bromoalkene).

What has happened? HBr has clearly added to the diene while one of the double bonds has vanished. However, the remaining double bond, whichever it is, has moved to a new position in the middle of the molecule. So how do we start? HBr is a strong acid so the reaction must begin with the protonation of some atom in the diene by HBr. Which one? If you examine the product you will see that one atom has an extra hydrogen and this must be where protonation occurs.

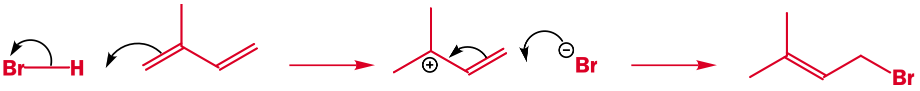


The only change is at the left-hand end of the molecule where there is an extra proton. We must add the proton of HBr to that atom. The highest-energy orbital at that atom is the rather unreactive alkene π bond so we must use that as the nucleophile, though the electrons are really being pulled out of the π bond by reactive HBr.

It is not necessary to draw in that hydrogen atom in the product of this step. It is, of course, necessary to put the positive charge on the carbon atom in the middle that has lost electrons. Now we can add bromide ion (the other product of the first step) to this cation but not where we have written the plus charge as that will not give us the right product. We must move the remaining double bond along the molecule as we add the bromide ion. This too is a 'pulled' reaction as the unstable plus charge on carbon pulls electrons towards itself.



So this is a two-step reaction and the driving force for the two steps is a strongly acidic electrophile in the first and a strongly electrophilic cation at carbon in the second. Here is the full mechanism.



Now we can summarize the extra points we have made in this section as a series of guidelines.

Extra guidelines for writing your own mechanisms

- Decide on the structure of any ambiguous reagents, for example, salt or a covalent compound?
- Decide which is the nucleophilic and which the electrophilic atom
- Decide whether to think in a *push* or a *pull* manner
- Mark lone pairs on the nucleophilic atom
- Draw the molecule(s) in a spatial arrangement that makes reaction possible
- Curly arrows always move in the same direction. They never meet head on!
- If you make a new bond to H, C, N, or O you must also break one of the existing bonds in the same step
- Draw your arrows in colour to make them stand out
- Mark charges clearly on reactants and intermediates
- Make sure that overall charge is conserved in your mechanism

We have only given you a preliminary trial run as a learner driver of curly arrows in this section. The way forward is practice, practice, practice.

Curly arrows are vital for learning organic chemistry

Curly arrows can be used to explain the interaction between the structure of reactants and products and their reactivity in the vast majority of organic reactions, regardless of their complexity. When used correctly they can even be used to predict possible outcomes of unknown processes and hence to design new synthetic reactions. They are thus a powerful tool for understanding and developing organic chemistry and it is vital that you become proficient in their use. They are the dynamic language of organic reaction mechanisms and they will appear in every chapter of the book from now on.

Another equally important reason for mastering curly arrows now, before you start the systematic study of different types of reactions, is that the vast number of 'different reactions' turn out not to be so different after all. Most organic reactions are ionic; they therefore all involve nucleophiles and electrophiles and two-electron arrows. There are relatively few types of organic electrophiles and nucleophiles and they are involved in all the 'different' reactions. If you understand and can draw mechanisms, the similarity between seemingly unrelated reactions will become immediately apparent and thus the number of distinct reaction types is dramatically reduced.

Drawing curly arrow mechanisms is a bit like riding a bike. Before you've mastered the skill, you keep falling off. Once you've mastered the skill, it seems so straightforward that you wonder how you ever did without it. You still come across busy streets and complex traffic junctions, but the basic skill remains the same.

If you still feel that drawing mechanisms for yourself is difficult, this stage-by-stage guide may help you. Once you've got the idea, you probably won't need to follow it through in detail.

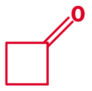
A guide to drawing mechanisms with curly arrows

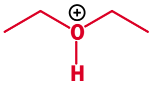
1. Draw out the reagents as clear structures following the guidelines in Chapter 2. Check that you understand what the reagents and the solvent are under the conditions of the reaction, for example, if the reaction is in a base, will one of the compounds exist as an anion?
2. Inspect the starting materials and the products and assess what has happened in the reaction. What new bonds have been formed? What bonds have been broken? Has anything been added or removed? Have any bonds moved around the molecule?
3. Identify the nucleophilic centres in all the reactant molecules and decide which is the most nucleophilic. Then identify the electrophiles present and again decide which is the most electrophilic
4. If the combination of these two centres appears to lead to the product, draw the reactants, complete with charges, so as to position the nucleophilic and electrophilic centres within bonding distance ensuring that the angle of attack of the nucleophile is more or less consistent with the orbitals involved
5. Draw a curly arrow from the nucleophile to the electrophile. It must start on the filled orbital or negative charge (show this clearly by just touching the bond or charge) and finish on the empty orbital (show this clearly by the position of the head). You may consider a 'push' or a 'pull' mechanism at this stage
6. Consider whether any atom that has been changed now has too many bonds; if so one of them must be broken to avoid a ridiculous structure. Select a bond to break. Draw a curly arrow from the centre of the chosen bond, the filled orbital, and terminate it in a suitable place
7. Write out the structures of the products specified by the curly arrows. Break the bonds that are the sources of the arrows and make those that are the targets. Consider the effect on the charges on individual atoms and check that the overall charge is not changed. Once you have drawn the curly arrows, the structure of the products is already decided and there is no room for any further decisions. Just write what the curly arrows tell you. If the structure is wrong, then the curly arrows were wrong so go back and change them
8. Repeat stages 5–7 as required to produce a stable product

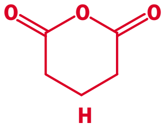
When you have read through all the different types of reaction mechanism, practise drawing them out with and without the help of the book. Complete the exercises at the end of the chapter and then try to devise mechanisms for other reactions that you may know. You now have the tools to draw out in the universal pictorial language of organic chemists virtually all the mechanisms for the reactions you will meet in this book and more besides!

Problems

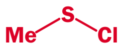
1. Each of these molecules is electrophilic. Identify the electrophilic atom and draw a mechanism for reaction with a generalized nucleophile Nu$^-$, giving the product in each case.

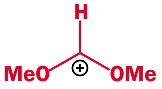






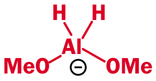
$Cl—Cl$



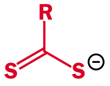


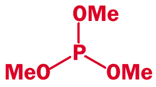
2. Each of these molecules is nucleophilic. Identify the electrophilic atom and draw a mechanism for reaction with a generalized electrophile E$^+$, giving the product in each case.

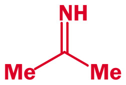
$R—C\equiv C^{\ominus}$



$H_2N—NH_2$

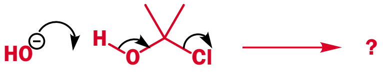




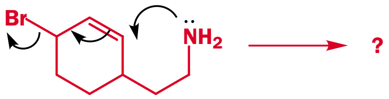


3. Complete these mechanisms by drawing the structure of the products in each case.

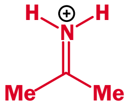
(a)

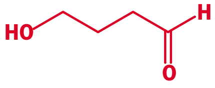


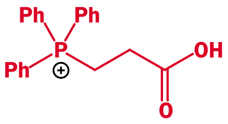
(b)



4. Each of these electrophiles could react with a nucleophile at (at least) two different atoms. Identify these atoms and draw a mechanism for each reaction together with the products from each.

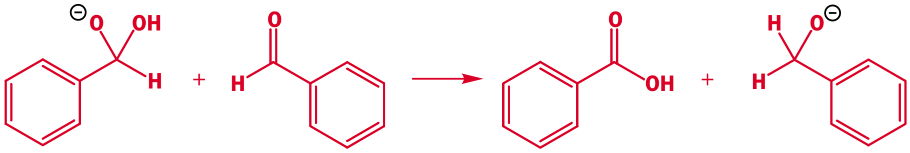






5. Put in the arrows on these structures (which have been drawn with all the atoms in the right places!) to give the products shown.

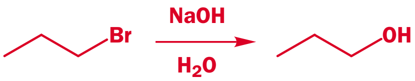
(a)



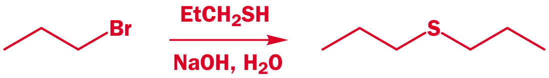
(b)

6. Draw mechanisms for these reactions. The starting materials have not necessarily been drawn in a helpful way.

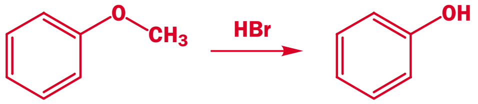
(a)



(b)



(c)



7. Draw a mechanism for this reaction.

$$PhCHBr.CHBr.CO_2H + NaHCO_3 \longrightarrow PhCH{=}CHBr$$

Hints. First draw good diagrams of the reagents. $NaHCO_3$ is a salt and a weak base—strong enough only to remove which proton? Then work out which bonds are formed and which broken, decide whether to push or pull, and draw the arrows. What are the other products?