

INTRODUCTION TO STRUCTURE DETERMINATION

There are so many organic compounds in existence that it can be difficult to establish exactly what compounds, or what combinations of compounds, are present in a given sample. Chemical tests can be used to distinguish between different functional groups, and melting point and boiling point data can provide some further information as to the identity and purity of a compound, but they can generally only confirm suspected structures and can not usually be used to identify new compounds. Chemical techniques also require fairly large quantities of sample and are not always effective at distinguishing between similar compounds.

Chemists have developed a number of other, more precise, analytical techniques to determine the exact structure of organic compounds. Knowledge of three of these techniques is required for AQA A-level Chemistry. These are **mass spectrometry**, **infra-red spectroscopy** and **nuclear magnetic resonance (nmr) spectroscopy**.

As infra-red spectroscopy is covered at AS-level, this topic is concerned with a more detailed consideration of mass spectrometry and nmr spectroscopy.

At A-level, these analytical techniques must be used to determine the structure of organic molecules containing carbon, hydrogen and oxygen only and generally containing no more than six carbon atoms. The molecule being analysed will be an alkane, alkene, alcohol, ether, carbonyl, carboxylic acid or ester.

Usually, information is given which enables the empirical formula to be deduced before the analysis begins. This is usually given in the form of composition by mass:

Eg A compound contains 40.0% carbon, 53.3% oxygen and 6.7% hydrogen.

Mole ratio:	C	$40/12 = 3.3$
	H	$6.7/1 = 6.7$
	O	$53.3/16 = 3.3$

Simplest whole number ratio:	C	$3.3/3.3 = 1$
	H	$6.7/3.3 = 2$
	O	$3.3/3.3 = 1$

So empirical formula = CH₂O

The molecular formula of the molecule, and its structure, can then be deduced from one or more of the analytical techniques.

USING THE ANALYTICAL TECHNIQUES TOGETHER

If all the different spectra are available, then the most effective analytical strategy is to use them together. There are a number of useful steps to deducing the structure:

If composition data is available:

1. Use the composition data to work out the empirical formula.
2. Use the mass spectrum to deduce the relative molecular mass, and hence work out the molecular formula.
3. Use the infra-red spectrum to check for the presence of C=O and O-H bonds, and hence identify the functional group.
4. Make a list of the possible isomers consistent with the molecular formula and functional group.
5. Use the number of peaks in the proton nmr spectrum to deduce the number of hydrogen environments. Eliminate the isomers inconsistent with this number.
6. Use the integration factors in the proton nmr spectrum to work out the number of hydrogen atoms in each environment. Eliminate the isomers inconsistent with this distribution.
7. Compare the splitting of the peaks in the nmr spectrum with the expected splitting patterns of the remaining possible isomers. Eliminate the isomers which do not give this splitting pattern.
8. If necessary, compare the chemical shifts of the peaks in the nmr spectrum with the expected values. Eliminate the isomers which are inconsistent with the chemical shifts.

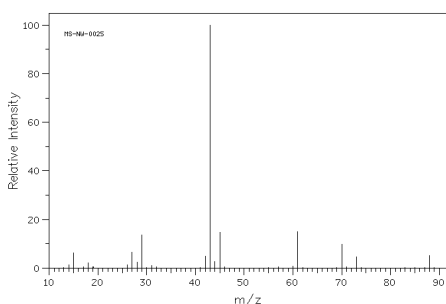
If composition data is not available, the molecular formula must be found by a trial and error method.

1. Deduce the relative molecular mass from the mass spectrum.
2. Deduce the functional groups present from the infra-red spectrum and hence establish the likely number of oxygen atoms present.
3. Add up the integration factors in the proton nmr spectra. The number of hydrogen atoms present is an integral multiple of this number.
4. Use the relative molecular mass and the information on hydrogen and oxygen atoms to deduce the molecular formula.

A worked example of combined analysis is given on the following page (no composition data available).

Use the following spectra to deduce the structure of A:

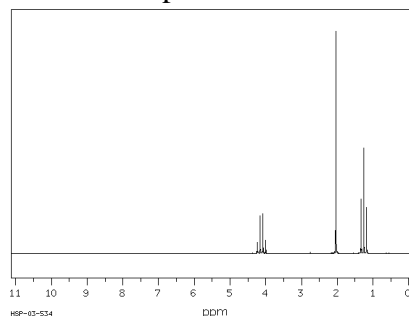
Mass spectrum:



Molecular ion peak = 88

Most intense fragments: 43, 29, 45, 61, 70

Proton nmr spectrum:



three peaks:

$\delta = 1.3$, triplet, integration factor 3

$\delta = 2.0$, singlet, integration factor 3

$\delta = 4.1$, quartet, integration factor 2

Infra-red spectrum: sharp absorption at 1715 cm^{-1} , no broad absorptions between 1500 and 3500 cm^{-1}

ANSWER:

From mass spectrum, $m_r = 88$

From infra-red spectrum, C=O present

From proton nmr spectrum, sum of integration factors = 8

So 8, 16 or 24 hydrogen atoms present and at least one oxygen

If one O atom present, remaining mass = 62 and cannot make this using 8/16/24 H atoms

If two O atoms present, remaining mass = 56; can make this from 4 carbons and H

hydrogens

So likely molecular formula = $\text{C}_4\text{H}_8\text{O}_2$.

From infra-red spectrum, no O-H is present, so it is not a carboxylic acid.

So is probably an ester.

Possible structures:

Methyl propanoate, ethyl ethanoate, propyl methanoate, methylethyl methanoate

From proton nmr spectrum:

There are 3 hydrogen environments so it cannot be propyl methanoate (which has 4).

The integration factors are 3:3:2 so cannot be methylethyl methanoate (which has the ratio 6:1:1).

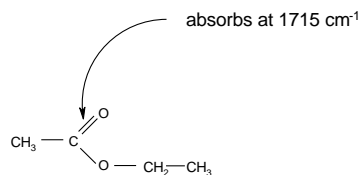
The coupling (3H triplet, 2H quartet, 3H singlet) is consistent with both remaining possible structures.

But the 3H singlet has a chemical shift of 2.0, which is consistent with CH_3CO and not CH_3O (for which the chemical shift would be 3.3 – 4.0)

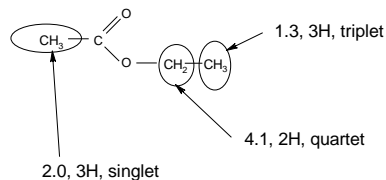
The 2H quartet has a chemical shift of 4.1, which is consistent with $\text{CH}_3\text{CH}_2\text{O}$ and not $\text{CH}_3\text{CH}_2\text{CO}$ (for which the chemical shift would be 2.0 – 2.5)

So it cannot be methyl propanoate and therefore the structure is ethyl ethanoate.
Full explanation of spectra:

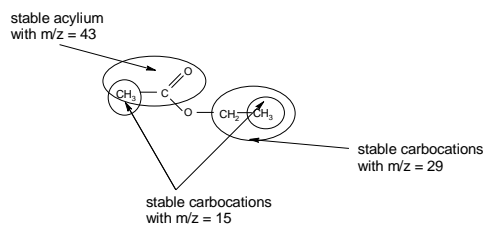
Infra-red spectrum:



Proton nmr spectrum:



Mass spectrum:



$[\text{CH}_3\text{COOCH}_2\text{CH}_3]^+$ is the molecular ion peak

$[\text{CH}_3\text{COOCH}_2\text{CH}_3]^+ \rightarrow [\text{CH}_3\text{CO}]^+ + [\text{OCH}_2\text{CH}_3]^\cdot$ accounts for the peak at $m/z = 43$

$[\text{CH}_3\text{COOCH}_2\text{CH}_3]^+ \rightarrow [\text{CH}_3\text{CH}_2]^\cdot + [\text{CH}_3\text{COO}]^+$ accounts for the peak at $m/z = 29$

$[\text{CH}_3\text{COOCH}_2\text{CH}_3]^+ \rightarrow [\text{CH}_3]^\cdot + [\text{CH}_3\text{COOCH}_2]^\cdot$ accounts for the peak at $m/z = 15$

$[\text{CH}_3\text{COOCH}_2\text{CH}_3]^+ \rightarrow [\text{CH}_3]^\cdot + [\text{COOCH}_2\text{CH}_3]^\cdot$ accounts for the peak at $m/z = 15$