SECTION 6

NOMENCLATURE AND STRUCTURE OF ORGANIC COMPOUNDS

Many organic compounds have common names which have arisen historically, or have been given to them when the compound has been isolated from a natural product or first synthesised. As there are so many organic compounds chemists have developed rules for naming a compound systematically, so that it structure can be deduced from its name. This section introduces this systematic nomenclature, and the ways the structure of organic compounds can be depicted more simply than by full Lewis structures. The language is based on Latin, Greek and German in addition to English, so a classical education is beneficial for chemists!

Greek and Latin prefixes play an important role in nomenclature:

	Greek	*Latin*
½	hemi	semi
1	mono	uni
1½		sesqui
2	di	bi
3	tri	ter
4	tetra	quadri
5	penta	quinque
6	hexa	sexi
7	hepta	septi
8	octa	octo
9	ennea	nona
10	deca	deci

Organic compounds: Compounds containing the element carbon [e.g. methane, butanol]. (CO, CO_2 and carbonates are classified as inorganic.) See *page 1-4.*

Special characteristics of many organic compounds are chains or rings of carbon atoms bonded together, which provides the basis for naming, and the presence of many carbon-hydrogen bonds. The valency of carbon in organic compounds is 4.

Hydrocarbons: Compounds containing only the elements C and H.

Straight chain hydrocarbons are named according to the number of carbon atoms: CH_4, methane; C_2H_6 or $H_3C\text{-}CH_3$, ethane; C_3H_8 or $H_3C\text{-}CH_2\text{-}CH_3$, propane; C_4H_{10} or $H_3C\text{-}CH_2\text{-}CH_2\text{-}CH_3$, butane; C_5H_{12} or $CH_3CH_2CH_2CH_2CH_3$, pentane; C_6H_{14} or $CH_3(CH_2)_4CH_3$, hexane; C_7H_{16}, heptane; C_8H_{18}, octane; C_9H_{20}, nonane; $C_{10}H_{22}$, $CH_3(CH_2)_8CH_3$, decane.

Saturated compound: One having only single bonds [e.g. ethane, C_2H_6].

Alkane: A saturated hydrocarbon [e.g. all the above compounds].
Systematic **nomenclature** is largely based on the above series of saturated alkanes with the number of carbon atoms associated with the stems *meth*-, (1); *eth*-, (2); *prop*-, (3); *but*-, (4);

pent-, (5); *hex-*, (6); *hept-*, (7): *oct-*, (8); *non-*, (9); *dec-*, (10); *alk-*, general. The ending *-ane* means no **unsaturation** (no double or triple bonds). Alkanes may be non-cyclic (**acyclic**) or **cyclic** (contain rings). The general formula for an acyclic alkane is C_nH_{2n+2} and for one containing one ring C_nH_{2n}. In cyclic alkanes the stem gives the number of carbon atoms in the ring. [e.g. *c*-C_6H_{12} is cyclohexane, where *c*- means cyclic]

Unsaturated compound: A compound with one or more multiple (double or triple) bonds [e.g. ethene (ethylene), $CH_2=CH_2$].

Alkene: A hydrocarbon containing a double bond [e.g. C_3H_6, CH_3-$CH=CH_2$, propene].

Alkyne: A hydrocarbon containing a triple bond [e.g. C_4H_6 or $CH_3CH_2C\equiv CH$, but-1-yne]. The endings *-ene* and *-yne* are for the double or triple bond respectively. The general formula C_nH_{2n+2} loses two H's for each ring or each double bond and four H's for each triple bond. The position of the multiple bond is shown by a number in the name, numbering from the end of the chain to give the smallest number [e.g $CH_3CH_2CH_2CH=CH_2CH_2CH_3$ is hept-3-ene (formerly 3-heptene) not hept-4-ene].

Alkyl group: In general, an alkane minus one hydrogen atom and represented by R [e.g. CH_3- is methyl (sometimes shown as Me); CH_3CH_2- is ethyl (sometimes shown as Et); $CH_3CH_2CH_2$- is propyl (sometimes shown as Pr); $CH_3CH_2CH_2CH_2$- is butyl (sometimes shown as Bu)].

In straight chain alkanes the non-terminal carbon atoms are bonded to two other carbon atoms. In a <u>branched</u> alkane one or more carbons are bonded to three or four other carbon atoms.

Primary carbon atom: A carbon atom bonded to only one other C atom.

Secondary carbon atom: One bonded to two other C atoms.

Tertiary carbon atom: One bonded to three other C atoms.

Quaternary carbon atom: One bonded to four other C atoms.

Branched hydrocarbons are named after the longest chain (saturated) or the longest chain containing the double or triple bond (unsaturated) with the branched group given by its alkyl name. [e.g. $CH_3C(CH_3)_2CH_2CH_3$ is 2,2-dimethylbutane.]

Isomers: Compounds with the same molecular formula but with their atoms arranged differently [e.g. hexane and 2,2-dimethylbutane, both C_6H_{14}].

Constitutional (structural) isomers: Isomers having their atoms joined together in a different sequence. (Some chemists restrict this term for isomers which have different **functional groups** [e.g. hexene and cyclohexane]. They would classify isomers containing the same functional groups as **positional isomers** [e.g. 2-methylpentane, $CH_3CH(CH_3)CH_2CH_2CH_3$ and 3-methylpentane, $CH_3CH_2CH(CH_3)CH_2CH_3$]).

Organic compounds are classified by the **functional groups** they contain.

Functional group: An atom or group of atoms which give the compound distinctive chemical properties [e.g. -Cl, -OH, >C=C<, -CO_2H]. Thus all organic compounds except saturated hydrocarbons have one or more functional groups. The functional group determines the class of compound. In nomenclature the functional group may be identified by a prefix, a suffix, or by the class of compound. (See below)

Common functional groups and classes of compounds are:

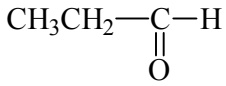
-F, *fluoro*-; -Cl, *chloro*-; -Br, *bromo*-; -I, *iodo*-; generally called **haloalkanes** (prefix) or **alkyl halides** (class of compound). [e.g. CH_3CH_2Cl is chloroethane or ethyl chloride. $CH_3CHFCH_2CH_3$ is 2-fluorobutane or secondary butyl fluoride.]

-OH, *hydroxy*-, giving rise to **alcohols**. The -OH group can be named as the prefix *hydroxy*-, as the suffix -*ol* replacing the -*e* of the alkane or as an alcohol. [e.g. $CH_3CH(OH)CH_2CH_3$ is butan-2-ol or secondary butyl alcohol; $(CH_3)_3COH$ is 2-methylpropan-2-ol or tertiary butyl alcohol].

-NH_2, *amino*-, giving rise to **amines**. The -NH_2 group can be named as the prefix *amino*-, with the suffix -*amine* replacing the -*e* of the alkane or as an amine. [e.g. $CH_3CH_2NH_2$ is aminoethane, ethanamine or ethylamine.] Amines can be considered as ammonia with hydrogens replaced by alkyl groups. If just one H is replaced the amine is a **primary amine**, as above. If two hydrogens are replaced the amine is a **secondary amine** [e.g. $CH_3CH_2NHCH_3$, N-methylethanamine (the N shows the methyl substituent is bonded to nitrogen) or ethyl methyl amine]. If all three hydrogens are replaced the amine is a **tertiary amine** [e.g. $(CH_3)_3N$ is trimethylamine].

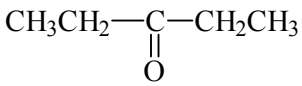
(Note that for alkyl halides and alcohols 1^o, 2^o or 3^o refers to the carbon atom to which the halo or hydroxy group is attached. For amines it refers to the number of alkyl groups on the nitrogen. Thus $(CH_3)_3COH$ is a tertiary alcohol, but $(CH_3)_3CNH_2$ is a primary amine.)

>C=O, carbonyl. If this is at the end of a chain the compound is an **aldehyde**, and named with the suffix -*al*. [e.g.

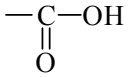


is propanal or propyl aldehyde. It is more simply written as CH_3CH_2CHO, but never as CH_3CH_2COH. You need to understand that the oxygen is bonded to the carbon. This -CHO representation is used because the other would look too much like an alcohol.]

If the CO group is not at the end of a chain the compound is a **ketone**, and named with the suffix *–one*. [e.g.



is pentan-3-one, (or diethyl ketone). Propanone (dimethyl ketone) is commonly called acetone. It is more simply written as CH_3COCH_3].



, carboxyl, giving rise to **carboxylic acids**. The suffix -*oic* and the word *acid*

are used in naming. [e.g. $CH_3CH_2CH_2CO_2H$ is called butanoic acid. Methanoic acid, HCO_2H, is commonly called formic acid, and ethanoic acid, CH_3CO_2H, is commonly called acetic acid.] The group is also often written as -COOH instead of $-CO_2H$. The CH_3CO- group is commonly called the acetyl group. RCO- is an acyl group.

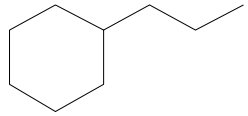
The product of the reaction of a carboxylic acid with a base is a **carboxylic acid salt**, an ionic compound. The name of the cation is given first followed by the acid with the suffix -*oate* replacing -*oic*. [e.g. $CH_3CH_2CH_2CO_2^-Na^+$ is sodium butanoate; $CH_3CO_2^-NH_4^+$ is ammonium ethanoate or ammonium acetate.] The general name for the anion is carboxylate.

When the OH of the carboxyl group is replaced by another group the compound is a **carboxylic acid derivative**. If the OH is replaced by OR of an alcohol the compound is called an **ester**. The R group is given first followed by the acid with the suffix -*oate* replacing -*oic* [e.g. $CH_3CH_2CO_2CH_2CH_3$ is ethyl propanoate]. When the OH group is replaced by NH_2 the compound is a **primary amide**. The suffix -*amide* replaces -oic. [e.g. CH_3CONH_2 is ethanamide, more commonly called acetamide]. If the OH has been replaced by an RNH the compound is a **secondary amide**, or by an RR'N group a **tertiary amide**, and the alkyl group of the amine named as such with the prefix N to show it is bonded to the nitrogen atom [e.g. $CH_3CH_2CONHCH_3$ is N-methylpropanamide]. If the OH has been replaced by a halo group the compound is an **acyl halide**, -*oic becoming -oyl* [e.g. CH_3CH_2COCl is propanoyl chloride]. If the OH has been replaced by a carboxylate group, OCOR, the compound is an **acid anhydride**. [e.g. $CH_3COOCOCH_3$ is ethanoyl anhydride or acetic anhydride. An anhydride in general is a substance formed by removing the elements of water from the compound.
[e.g. $2CH_3CO_2H \rightarrow CH_3COOCOCH_3 + H_2O$]

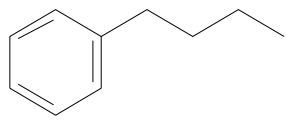
Multifunctional compound: A compound with more than one functional group.

Nomenclature of multifunctional compounds: The longest chain containing the suffix is chosen, the priority for choosing the suffix being carboxylic acid, $-CO_2H$, > carboxylic acid derivative, -COX > aldehyde, -CHO > ketone, -CO-, > alcohol, -OH > amine, $-NH_2$. The second and other groups are labelled as substituents. [e.g. $CH_3CH(OH)CH_2CO_2H$ is 3-hydroxybutanoic acid; $HOCH_2CH_2CH_2COCH_3$ is 5-hydroxypentan-2-one; $CH_3CH(OH)CH_2C(CH_3)(NH_2)CH_3$ is 4-amino-4-methylpentan-2-ol; CH_3COCO_2H is 2-oxopropanoic acid, (the =O of an aldehyde or ketone is called **oxo** when it has to be named as a substituent).] The carbon-carbon double and triple bonds are always incorporated in the chain, with lower priority than the other groups. [e.g. $CH_2=CHCH(OH)CH_3$ is but-3-en-2-ol; $CH_3C\equiv CCH_2CO_2H$ is pent-3-yn-oic acid.]

For compounds with larger carbon skeletons a further condensation of structural may be used.

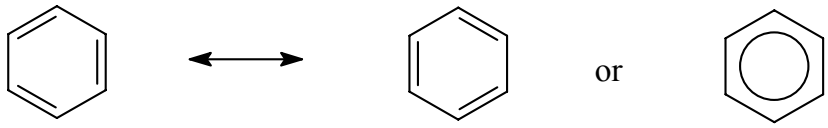


represents propylcyclohexane. Each line represents two carbon atoms joined by a single bond, and hydrogens which are present are not shown. The number of H's is such to satisfy the valency of carbon, 4.



is butylbenzene.

Benzene is C_6H_6 and is the parent of **aromatic** compounds. Each carbon in the benzene ring has one hydrogen attached. As a second **resonance structure** with the double bonds in the other three positions can be drawn, the **resonance hybrid** of benzene is often represented as a hexagon with a circle inside:
i.e.



Aromatic compound: An organic compound containing one or more benzene rings.

Phenyl group: Benzene minus one hydrogen, C_6H_5- (sometimes shown as Ph). [e.g. $C_6H_5CH_2CH_2CO_2H$ is 3-phenylpropanoic acid.]

Benzyl group: $C_6H_5CH_2$- [e.g. $C_6H_5CH_2Cl$ is benzyl chloride; $C_6H_5CO_2H$ is benzoic acid.]

Many substituted benzenes have common names. PhOH is phenol, $PhCH_3$ is toluene; $PhNH_2$ is aniline; $PhCH=CH_2$ is styrene.

Vinyl Group: Ethene minus a hydrogen, $CH_2=CH$-. [e.g. $CH_2=CHCl$ is vinyl chloride, $CH_3CO_2CH=CH_2$ is vinyl acetate.] Vinyl compounds are the monomers of vinyl polymers. [e.g. PVC, polyvinyl chloride; PVA, polyvinyl acetate.]

EXERCISES

Write condensed formulae, but showing all double and triple bonds, for the compounds:

1. *Example*: 2-chlorohex-3-ene

   Answer: $CH_3CHClCH=CHCH_2CH_3$

2. pent-2-ene 3. oct-4-yne 4. 5-bromohept-2-one

5. 4-aminobutanal 6. pentyl propanoate 7. 2-propyl-4-hydroxyhexanoic acid

8. nonanamide 9. N-ethylmethanamide

10-18. Give the systematic name of the compounds in questions 9-17 of *section 5*.