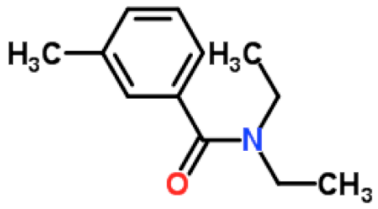
INTRODUCTION TO ORGANIC CHEMISTRY

UNIT 01: Nomenclature

Organic chemistry involves the study of compounds of carbon. Carbon atoms have the ability to form stable bonds with one another, and hence make chains of virtually any length as well as complex ring and cage structures. This means that there are a colossal number of compounds containing carbon and hydrogen *only*, and millions of other organic compounds containing carbon, hydrogen and other elements such as nitrogen, oxygen, sulfur and the halogens, e.g.,

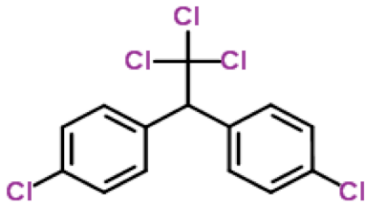
'DEET'



N,N-diethyl-3-methyl-Benzamide

$C_{12}H_{17}NO$

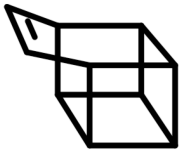
'DDT'



1,1'-(2,2,2-Trichloro-1,1-ethanediyl)bis(4-chlorobenzene)

$C_{14}H_9Cl_5$

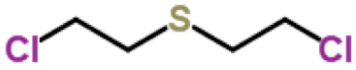
'Basketene'



Pentacyclo[4.4.0.0^{2,5}.0^{3,8}.0^{4,7}]dec-9-ene

$C_{10}H_{10}$

'Sulfur Mustard'



1-Chloro-2-[(2-chloroethyl)sulfanyl]ethane

$C_4H_8Cl_2S$

Image Credits: www.chemspider.com

The individual atoms that may up these compounds may be bonded to one another in many different ways, but there is a consistency regarding the *total* number of covalent bonds each type of atom may make. In all of the organic compounds *that we will consider*, all carbon atoms must make four covalent bonds, all oxygen atoms must make two covalent bonds, all nitrogen atoms must make three covalent bonds and all hydrogen atoms & all halogens must make one covalent bond.

With millions of compounds comes the need to *distinguish* between millions of compounds. One such way that you are familiar with is to use chemical formula. That's fine, but when you consider that www.chemspider.com lists twenty-four compounds with the relatively simple molecular formula C_8H_{18}, it quickly becomes apparent that we need a finer filter. Each of those twenty-four compounds of C_8H_{18}, and each of the millions and millions of others, requires a unique name. In the examples above you will see two names for each compound. DEET, DDT, Basketene and Sulfur Mustard are the trivial (non-systematic and often common) names of the compounds, and below each drawn structure you will see the systematic name for each (along with the molecular formula). A quick look at those systematic names shows why trivial names are often favored!

The systematic naming of organic compounds is governed by the International Union of Pure and Applied Chemistry (IUPAC) and the rules for naming compounds are summarized in their 'Blue Book'. Those rules can be found at www.acdlabs.com/iupac/nomenclature/. From simple origins, those rules become seemingly infinitely complex extremely quickly, and we will only scratch the surface here, but this TOPIC is devoted to the science/art of organic nomenclature.

- **Characteristics of organic compounds**

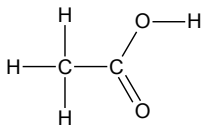
We will consider six identifying characteristics of organic compounds.

1. Empirical formula – showing the simplest whole number ratio of the atoms of each element in that compound e.g., CH_2O for ethanoic acid, meaning there is a ratio of 1:2:1 of C:H:O atoms in a single molecule of ethanoic acid.

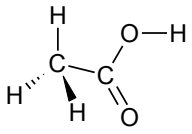
2. Molecular formula - showing the actual number of atoms of each element in one molecule of that compound e.g., $C_2H_4O_2$ for ethanoic acid, meaning there are two carbon atoms, four hydrogen atoms and two oxygen atoms present in a single molecule of ethanoic acid.

3. Structural/Condensed formula - showing how those atoms are arranged e.g., CH_3CO_2H for ethanoic acid, meaning three hydrogen atoms are attached to one carbon atom, which in turn is attached to another carbon atom, which in turn has two oxygen atoms attached to it, one of which has a hydrogen atom attached to it.

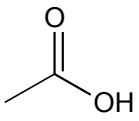
4. Graphical (Lewis) formula - showing how these atoms are arranged in space and the bonds between them. Lines represent covalent bonds (shared pairs of electrons) between atoms e.g., for ethanoic acid a double line (=) represents a double covalent bond and a single line (–) represents a single covalent bond. Lone pairs may (or may not) be shown, and in the case below, two lone pairs are omitted from each of the O atoms.



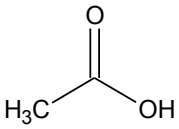
Additionally wedges and dotted lines can be used in order to represent covalent bonds and their 3D positions in space. A wedge represents a bond coming toward the viewer, and a dotted line represents one going away from the viewer e.g.,



5. Skeletal formula – showing an abbreviated form of the carbon chain, with each line segment understood to have a carbon atom at each end. These structures may (or may not) show terminal (end of chain) carbons. e.g., for ethanoic acid



or



6. Name - based upon an accepted system for naming compounds e.g., CH_3CO_2H, is named, "ethanoic acid" in IUPAC nomenclature. However, before diving into nomenclature, note the following very carefully!

As we have seen above, many compounds have (for good reason) trivial names that have arisen from (among other reasons) their history and origins. In addition, there have been other attempts to produce systematic nomenclature systems for example the Chemical Abstracts (CA) system. As a result, compounds often have several different, *entirely acceptable names,* and the modern IUPAC name may not be the one most commonly used. Furthermore, there are a number of inconsistencies that persist within the literature of even IUPAC naming that have the potential to confuse. One such example is the positioning of locants (numbers). For example, strictly speaking (according to post 1993 IUPAC nomenclature recommendations), the "correct" name for $CH_3CH_2CH_2CH_2OH$ is butan-1-ol – the '1' denotes the position of the –OH group on the carbon chain. Despite this, many texts persist with the (pre-'93) name, 1-butanol. Although the first is "better" than the second, there is no ambiguity so they both exist and are both acceptable - lack of ambiguity is the key.

In short, since nomenclature is simply a system for communication amongst chemists, if a compound can be reasonably named in more than one way *without ambiguity*, then different names are likely to be entirely acceptable (assuming that in any particular case there is no *specific* insistence on the use of the formal, IUPAC name).

- **Nomenclature**

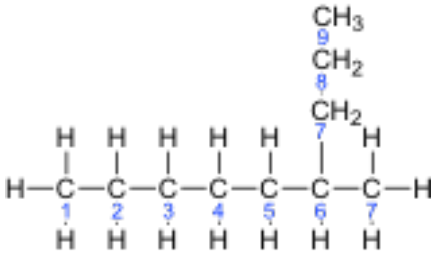
As mentioned above, the complete set of rules for naming *all* organic compounds is extremely long-winded and much too long to consider in completeness here (or indeed almost anywhere else). You will pick up the ideas with experience but *greatly* simplified, and in summary, the rules are as follows.

1. Chain Length

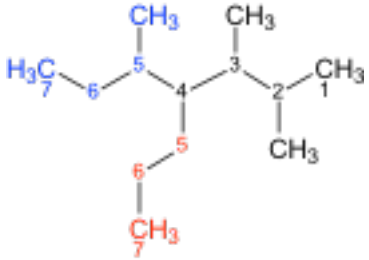
The root of the name of the compound is based upon the number of carbon atoms in the **longest, continuous chain**.

Number of carbon atoms in the longest continuous chain	Root of name
1	meth-
2	eth-
3	prop-
4	but-
5	pent-
6	hex-
7	hept-
8	oct-
9	non-
10	dec-
11	undec-
12	dodec-

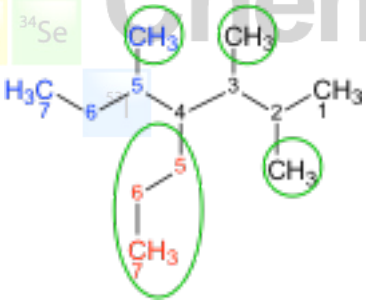
The longest, continuous chain can "go around corners", i.e., although the carbons in the longest chain must be connected, they do not have to be written in a "straight line". For example, in the molecule shown below the longest chain is formed by nine carbon atoms, not by just the seven that run from left right straight across the structure.



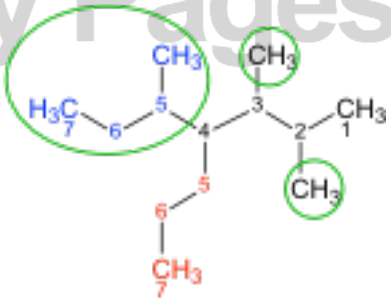
In a compound where there are two chains that are equally long, priority is given to chain with the greatest number of side chains (branching). For example, in the molecule shown below, the longest chain that runs continuously from right to left when considering the blue carbon atoms labeled 5–7 in the diagram, *or* from right to bottom when considering the red carbon atoms 5–7.



Since numbering from right to left through the blue carbons maximizes the side chains (4), rather than minimizing them (3) by numbering from right to bottom through the red carbons, the former is the preferred method.



4 side-chains, preferred



3 side-chains, not preferred

There are many other rules that follow on from these basic ideas for finding the longest carbon chain in an aliphatic compound, but they are beyond the scope of this course.

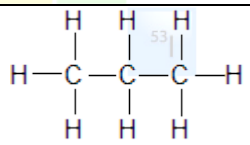
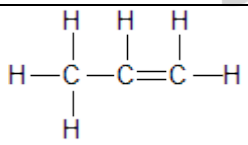
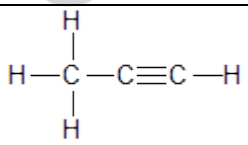
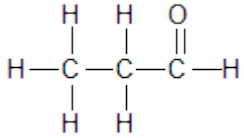
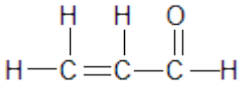
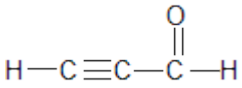
Cyclo compounds - those containing carbon atoms joined together in rings - are named using the cycloalkane as the root (if there are greater number of carbons in the ring than there are in the alkyl substituents), *or* using the cyclo part of the compound as a prefix to the main, (greater carbon) chain. If there are two or more substituents on a ring, number these so as to keep the lowest numbers with the alphabetical order.

2. Functional Groups

Organic compounds contain various functional groups. A functional group is an atom, or group of atoms, that occur together and as a whole possess their own characteristic properties. The existence of each functional group is shown by adding a suffix (an ending) to the root (determined by the longest chain length), and if necessary, an additional prefix (a beginning).

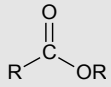
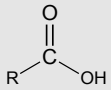
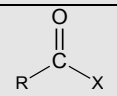
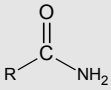
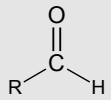
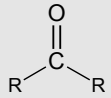
The root and suffix are linked together using either 'an', 'en' or 'yn', according to the construct of the carbon chain. Chains with carbon-carbon single bonds use 'an', those with double bonds use 'en', and those with triple bonds use 'yn'. Where no functional groups exist (other than carbon-carbon single, double or triple bonds) there is effectively an absence of a suffix, and the compound will end with '–ane' for single bonded alkanes, '-ene' for double bonded alkenes, and '-yne' for triple bonded alkynes.

For example, here are some compounds with three membered carbon chains that have no functional groups (other than C-C, C=C or C≡C bonds), AND compounds with three membered carbon chains with the aldehyde functional group (see below) that has the suffix '-al'.

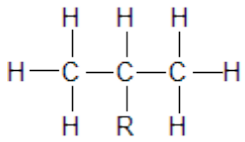
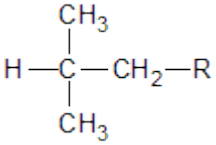
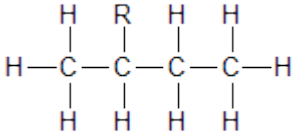
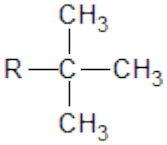
		
Propane	Propene	Propyne
		
Propanal	Propenal	Propynal

FUNCTIONAL GROUPS

(R represents an alkyl group, see below)

Priority	Structure	Name	Prefix	Suffix
1		**Ester**	-	-oate
2		**Carboxylic acid**	carboxy-	-oic acid
3	 X = F, Cl, Br, I	**Acid Halide**	-	-oyl halide
4		**Amide**	carbamide	-amide
5	R-C≡N	**Nitrile**	cyano-	-nitrile
6		**Aldehyde**	oxo-	-al
7		**Ketone**	oxo-	-one
8	R-OH	**Alcohol**	hydroxy-	-ol
9	$R\text{-}NH_2$	**Amine**	amino-	-amine
10	R-X X = F, Cl, Br, I	**Halogeno**	halo-	-
11	R-	**Alkyl** Alkyl groups are derived from the corresponding alkane with an H atom removed. Methane (CH_4) ➔ methyl (CH_3-), ethane (C_2H_6) ➔ ethyl (CH_3CH_2-), propane (C_3H_8) ➔ propyl ($CH_3CH_2CH_2$-)* & butane (C_4H_{10}) ➔ butyl ($CH_3CH_2CH_2CH_2$-)*	alkyl-	-

In addition to the propyl and butyl straight chain alkyl groups, there are other, related, branched groups. These names are not in line with the official, modern IUPAC nomenclature, but are useful to know since they persist in the literature. In each case, R represents the remainder of the molecule, whatever that may be.

Propyl isomer	**Butyl isomers**		
			
Isopropyl (*i*-Pr-)	Isobutyl (*i*-Bu)	sec-butyl (s-butyl)	tert-butyl (t-butyl)
R connected to CH in propyl group	R connected to the CH_2 in the butyl group	R connected to CH in the butyl group	R connected to C in butyl group

3. Choice of suffix and prefix.

A suffix is chosen first, according to its priority. As noted above, then –an, –en or –yn is used in conjunction with the chosen suffix in order to link the root to the chosen suffix. In the absence of a suffix, the name ends simply ends in –ane, ene or -yne.

Any functional groups that remain unaccounted for by the use of one suffix, are specified by prefixes, and are written in alphabetical order. Once a suffix has been chosen, as many prefixes as required may be added.

4. Final considerations.

Locants (numbers) precede the prefix names in order to indicate their position on the longest carbon chain, and 'N' can be used instead of a number to indicate an alkyl group connected directly to a nitrogen atom, rather than a carbon atom).

If one type of functional group appears more than once, the multiplying prefixes (di, tri, tetra etc.) must be Added but prefixes are listed alphabetically **before** any multiplying prefixes (di, tri, tetra etc.) have been added.

The chain is numbered from the end where the principle functional group is situated (the suffix), and in a manner that results in the lowest numbers being used in the name.

- **Homologous series**

Compounds with the same functional group, differing only in the number of CH_2 units in their skeleton, are said to be in the same *homologous series*. They are part of the same family of compounds with similar reactions and characteristics that differ only in the number of carbon atoms that make up the chain. Members of a series will have formulae that are linked by a general formula. For example, all alkanes have the general formula C_nH_{2n+2}, where n is an integer that represents the number of carbon atoms in the molecule.

- **Aromatics**

Aromatics are planar (flat), cyclic compounds where all carbons are sp^2 hybridized and the number of delocalized pi electrons[1] conform to the 4n + 2 (Hückel rule) where n = 0 or another integer. There are many different types of aromatics, but we will restrict our study to those containing a benzene ring (C_6H_6). The ring contains six sp^2 hybridized carbon atoms joined in a hexagonal ring, with one hydrogen atom attached to each carbon. It may be represented in one of two ways.

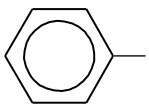
Kekulé structure

This structure shows the six carbon atoms joined together in a ring with alternate single and double bonds

Delocalized structure

This structure shows the more realistic structure where resonance leads to the delocalization of the alternate single and double bonds to create a π cloud of electron density above and below the ring

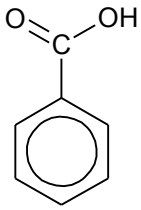
When a benzene ring is attached to some other group, it loses an H atom and becomes C_6H_5, the phenyl (Ph or Arene) group. This is precise analogous with the alkyl groups listed on page 9.



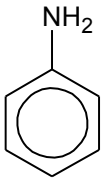
[1] See UNIT 03 for a full explanation

Trivial (traditional or common) names of benzene derivatives

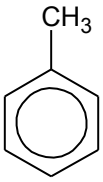
Many benzene compounds are named by their trivial names (not systematic IUPAC nomenclature). These include (but are not limited to),



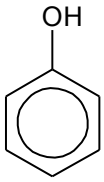
Benzoic Acid



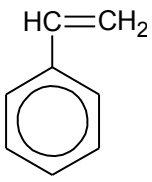
Aniline



Toluene



Phenol

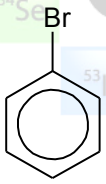


Styrene

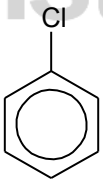
These compounds can also be named according to regular IUPAC nomenclature, for example, toluene may be called methylbenzene.

Monosubstituted benzenes

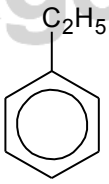
Monosubstituted benzene rings that do not have trivial names that are accepted by IUPAC, are named as derivatives of benzene, for example,



Bromobenzene



chlorobenzene

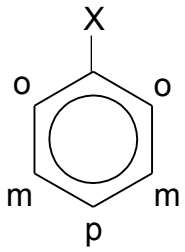


Ethylbenzene

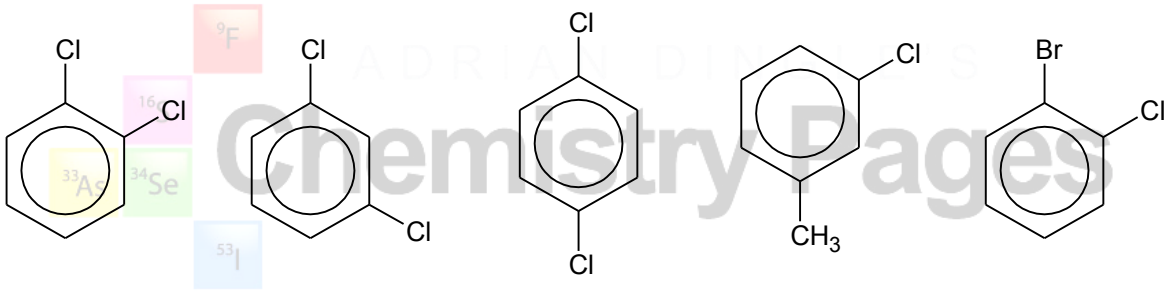
Disubstituted benzenes

When two substituents are present on the ring, the position of the second substituent relative to the first substituent, can be named in two separate ways. Either,

(i) Using ortho, meta and para prefixes where ortho represents the 2 and 6 positions on the ring, meta represents the 3 and 5 positions and p represents the 4 position on the ring.



For example,



ortho-dichlorobenzene | meta-dichlorobenzene | para-dichlorobenzene | meta-chlorotoluene | ortho-bromochlorobenzene

Notes:

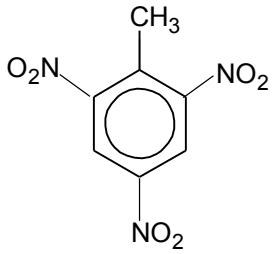
Ortho, meta and para positions are relative to the first substituent. So, in meta-chlorotoluene the methyl group is considered to be position 1 (regardless of have the molecule is drawn) and the chlorine atom is positioned at position 3.

Where the compound is not based upon a trivial name (e.g., ortho/meta/para-chlorotoluene), the substituents are listed by priority then alphabetically.

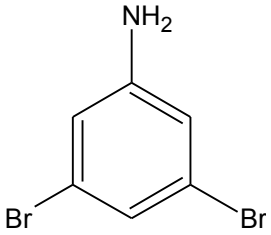
(ii) Simply using the numbering system without ortho, meta and para. For example meta-dibromobenzene could also be called 1,3-dibromobenzene and meta-chlorophenol could be called 3-chlorophenol.

Polysubstituted benzenes

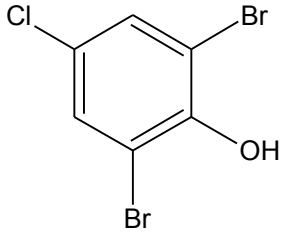
The numbering system is used to indicate position. e.g.,



2,4,6-trinitrotoluene (TNT)
(2,4,6-trinitromethylbenzene)



3,5-dibromoaniline



2,6-dibromo-4-chlorophenol

Task 01a

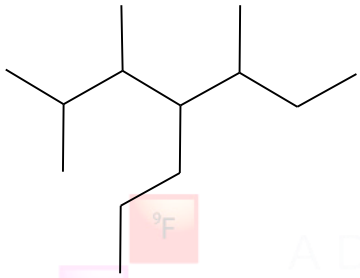
These questions come in pairs. The heading identifies the functional group(s). The first question in the pair asks for a name from a formula, then the second question asks for a formula from a name.

ALKANES

1. ***Name the following compounds.***

   (i) $CH_3C(CH_3)_2CH_2CH_3$

   (ii)



   (iii) $CH_3CH_2C(C_2H_5)_2CH_2CH_2CH_3$

   (iv) $CH_3CH(CH_3)CH_2CH(CH_3)CH_3$

   (v) $CH_3CH_2CH(CH_3)CH(C_2H_5)CH_2C(CH_3)_2CH_3$

2. ***Draw graphical formulae for the following compounds***

   (i) ***2-methylpropane***

   (ii) ***3-methylheptane***

   (iii) ***2,3-dimethylbutane***

   (iv) ***2,2,6-trimethyloctane***

   (v) ***1,3-diethylcyclobutane***

ALKENES & ALKYNES

3. Name the following compounds

(i) $CH_3CH_2CH=CH_2$

(ii) $CH_3C{\equiv}CCH_3$

(iii) $CH_3CH_2CH(CH_3)C{\equiv}CCH_3$

(iv) $CH_3C{\equiv}CCH(CH_3)_2$

(v) $CH_2=CHCH=CH_2$

4. Draw graphical formulae for the following compounds

(i) 2,3-dimethylbut-2-ene

(ii) 4-methylhex-1-yne

(iii) 3,4-diethylhex-2-ene

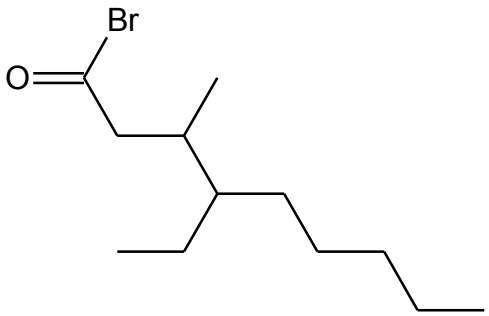
(iv) 4-ethylcyclohex-1-ene

(v) 3-methylcyclooct-1-ene

HALOGEN COMPOUNDS

5. Name the following compounds

(i) $CH_3CH_2CH_2CH_2CH_2I$



(ii)

(iii) C_6H_5Br

(iv) $C_6H_5CH_2Cl$

(v) $(CH_3)_2CHBr$

6. Draw graphical formulae for the following compounds

(i) 2-iodo-2-methylpentane

(ii) 1-chloro-1-bromoethane

(iii) triiodomethane

(iv) 1-bromo-3-chloro-4-methylhexane

(v) 1-chloro-1-phenylethane

ALCOHOLS

7. Name the following compounds

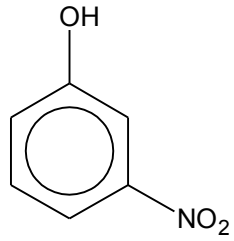
(i) $CH_3CH_2CH(OH)CH_2CH_3$

(ii) $CH_3C(CH_3)(OH)CH_2CH_3$

(iii) $CH_3CH(OH)CHBrCH_3$

(iv) $CH_3CH(C_6H_5)CH_2CH_2OH$

(v)



8. Draw graphical formulae for the following compounds

(i) pentan-2-ol

(ii) 3-methylhexan-1-ol

(iii) cyclohexane-1,3-diol

(iv) 2-phenylpropan-2-ol

(v) 4-chlorophenol

ALDEHYDES & KETONES

9. Name the following compounds

(i) $CH_3CH_2CH_2CHO$

(ii) $CH_3CH(CH_3)COCH_3$

(iii) $CH_3CH(CH_3)CHO$

(iv) $CH_3CH(CH_3)COCH_3$

(v) $C_6H_5CH_2CH_2CHO$

10. Draw graphical formulae for the following compounds

(i) cyclohexanone

(ii) 3-methylbutanal

(iii) 2,2-dimethylpropanal

(iv) 3-chlorobutan-2-one

(v) 4-phenylpentanal

CARBOXYLIC ACIDS & DERIVATIVES

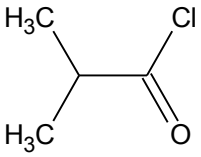
11. Name the following compounds

(i) $CH_3CH_2CO_2H$

(ii) $CH_3CH_2CH(OH)CH_2COOH$

(iii) $(CH_3)_3CCO_2H$

(iv)



(v) $C_6H_5CH(CH_3)CO_2H$

12. Draw graphical formulae for the following compounds

(i) ***pentanoic acid***

(ii) ***3-aminopropanoic acid***

(iii) ***butenedioic acid***

(iv) ***chloroethanoic acid***

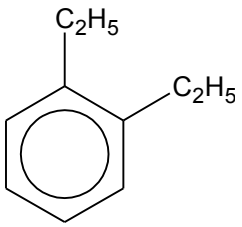
(v) ***methylpropanoate***

AROMATICS

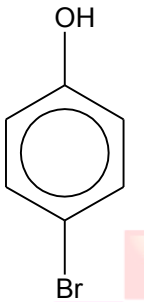
13. Name the following compounds

(i) **$PhCH_2Cl$**

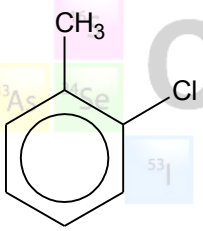
(ii)



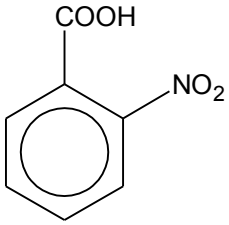
(iii)



(iv)



(v)



14. Draw graphical formulae for the following compounds

(i) 2-phenylpropane

(ii) 1,2-dihydroxybenzene

(iii) 4-nitrophenol

(iv) 3-chlorobenzoic acid

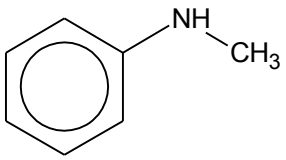
(v) 2,4,6-tribromophenol

NITROGEN COMPOUNDS

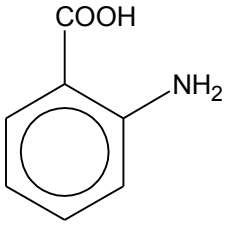
15. Name the following compounds

(i) CH_3CH_2CN

(ii)



(iii)



(iv) $(C_2H_5)_3N$

(v) $CH_3CH(C_6H_5)CH_2CN$

16. Draw graphical formulae for the following compounds

(i) 2-methylbutanonitrile

(ii) 3-aminopentane

(iii) phenylmethylamine

(iv) N,N-dimethylphenylamine

(v) N-phenylpropanamide