NAMING ORGANIC COMPOUNDS

1. ALKANES

Table 1 Straight-chain alkane (*n*-alkane) names

Number of carbons (n)	*Name*	*Formula* (C_nH_{2n+2})	*Number of carbons (n)*	*Name*	*Formula* (C_nH_{2n+2})
1	Methane	CH_4	6	Hexane	C_6H_{14}
2	Ethane	C_2H_6	7	Heptane	C_7H_{16}
3	Propane	C_3H_8	8	Octane	C_8H_{18}
4	Butane	C_4H_{10}	9	Nonane	C_9H_{20}
5	Pentane	C_5H_{12}	10	Decane	$C_{10}H_{22}$

The suffix *-ane* is added to the end of each name to show that the compound is an alkane. Thus, but*ane* is the four-carbon alkane, hept*ane* is the seven-carbon alkane, and so on. The names of the first ten alkanes, given in Table 1, should be memorized. Larger alkanes, such as icos*ane* ($C_{20}H_{42}$), have more complicated names and are outside the scope of this course.

If one hydrogen atom is removed from an alkane, the remaining part of the molecule is called an **alkyl group**. Alkyl groups are named by replacing the *-ane* ending of the parent alkane by an *-yl* ending. For example, removing a hydrogen atom from methane gives the **methyl group**. Similarly, removal of a hydrogen from an **end** carbon of any *n*-alkane produces the series of straight-chain alkyl (*n*-alkyl) groups show in Table 2.

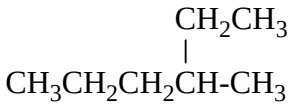
Table 2 Straight-chain alkyl (*n*-alkyl) groups (Abbreviations in parentheses)

Alkane	*Alkyl group*	*Alkane*	*Alkyl group*
CH_4	CH_3-	$CH_3CH_2CH_3$	$CH_3CH_2CH_2-$ or *n*-C_3H_7
Methane	Methyl (Me)	Propane	Propyl(Pr)
CH_3CH_3	CH_3CH_2-	$CH_3CH_2CH_2CH_3$	$CH_3CH_2CH_2CH_2-$ or *n*-C_4H_9
Ethane	Ethyl (Et)	Butane	Butyl (Bu)

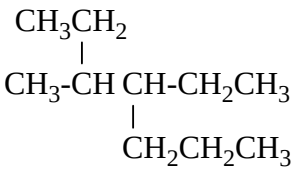
Using the IUPAC (International Union of Pure and Applied Chemistry) rules, most branched-chain alkanes can be named by the following four steps. For more complex alkanes, a fifth step is needed.

Step 1 Find the parent hydrocarbon

a) Find the *longest continuous carbon chain* present in the molecule and use the name of that chain as the parent name. The longest chain may not always be obvious from how the structure is written, as shown below.

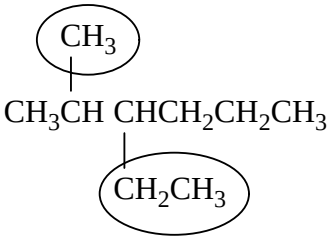


Named as a substituted hexane



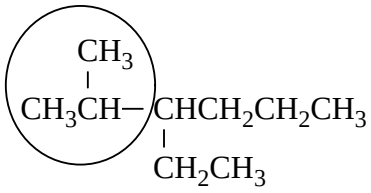
Named as a substituted heptane

b) If two different chains of equal length are present, choose the one with the most substituents as the parent:



Named as a hexane with **two** substituents

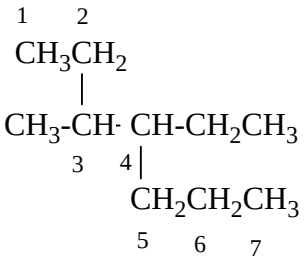
NOT



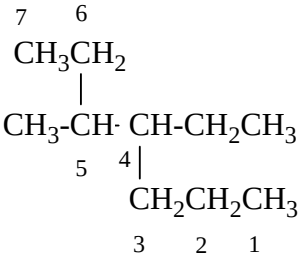
as a hexane with **one** substituent

Step 2 Number the atoms in the main chain

a) Beginning at the end *nearer the first branch*, number each carbon atom in the longest chain:

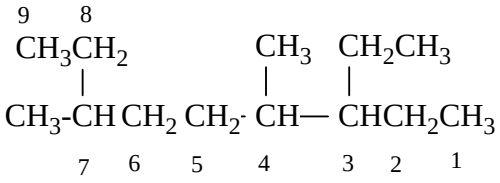


NOT

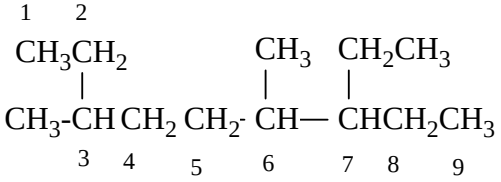


The first branch occurs at C3 in the proper numbering system but at C4 in the improper system.

b) If there is branching an equal distance from both ends of the longest chain, number from the end nearer the *second* branch:



NOT

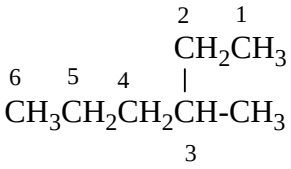


Step 3 Identify and number the substituents

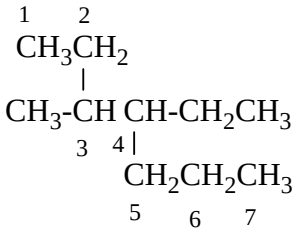
a) Using the numbering arrived at in step 2, assign a number to each substituent according to its point of attachment to the main chain:

b) If there are two substituents on the same carbon, assign them both the same number. There must always be as many numbers in the name as there are substituents:

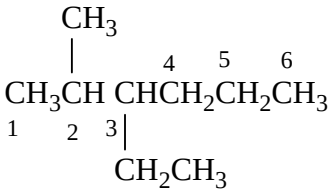
Step 4 Write out the name as a single word, using hyphens to separate the different prefixes and using commas to separate numbers. If two or more *different* substituents are present, cite them in alphabetical order. If two or more *identical* substituents are present, use one of the prefixes *di-*, *tri-*, *tetra-*, and so forth. Do **not** use these prefixes for deriving the alphabetical order, however. Some examples are shown below.



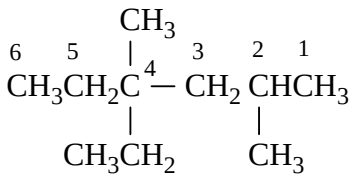
3-Methylhexane



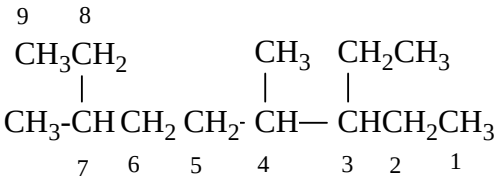
4-Ethyl-3-methylheptane



3-Ethyl-2-methylhexane



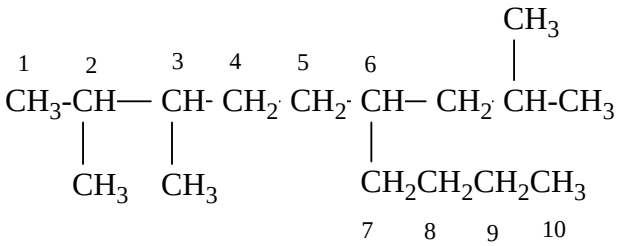
4-Ethyl-2,4-dimethylhexane



3-Ethyl-4,7-dimethylnonane

Complex substituents

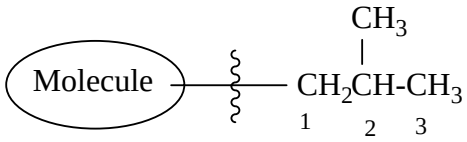
Application of the preceeding four steps allows us to name many alkanes. However, in some very complex cases a fifth step is needed. This occurs when a substituent is, itself, branched (i.e., has sub-branching). Such a substituent is called a *complex substituent*. An example is shown below.



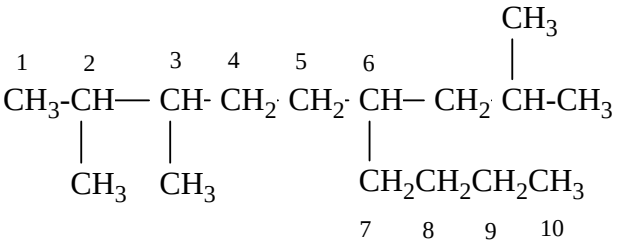
Named as a 2,3,6-trisustituted decane

In this case, the substituent at C6 is a branched four-carbon unit. To name the compound fully, the complex substituent must first be named.

Step 5 Name the complex substituent. A complex substituent is named by applying the four steps described above exactly as if it were a compound itself. In the present case, the complex substituent is a substituted propyl group.

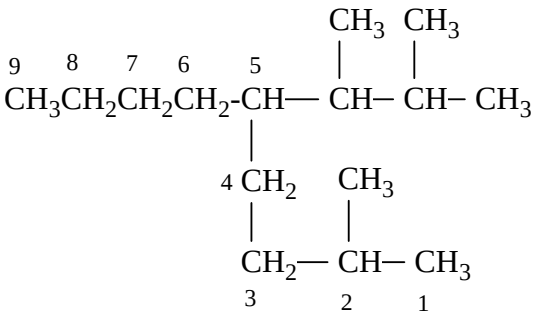


Begin numbering *at the point of attachment* to the main chain. At C2, there is a methyl group. The complex substituent is therefore a 2-methylpropyl group. To avoid confusion, the complex group name is put in parentheses when the name of the complete molecule is written.

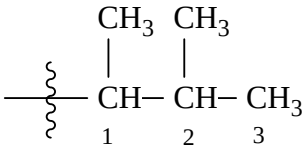


2,3-Dimethyl-6-(2-methylpropyl)decane

Another example is given below.



5-(1,2-Dimethylpropyl)-2-methylnonane

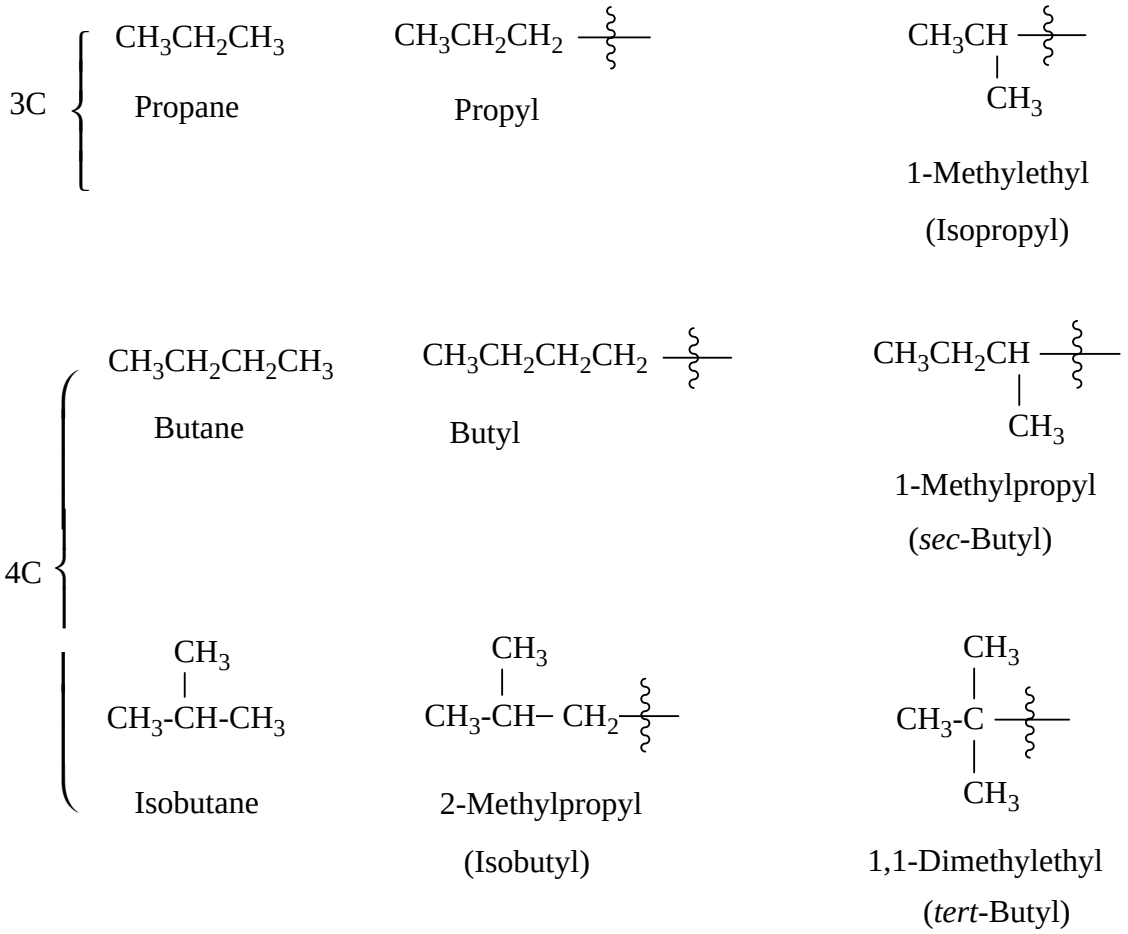


(1,2-Dimethylpropyl)-

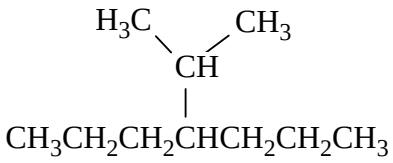
N.B.: A **complex substituent** is alphabetized under the first letter of its name.

Branched alkyl groups

On page 40, it was shown that straight-chain alkyl (*n*-alkyl) groups are formed by removal of a terminal (end) hydrogen atom from straight-chain alkanes. It is also possible to generate a large number of *branched* alkyl groups by removing *internal* hydrogen atoms from alkanes. For example, there are two possible three-carbon alkyl groups and four possible four-carbon alkyl groups, as shown below. The possibilities increase at an enormous rate as the number of carbon atoms increases. Branch-chain alkyl groups can be systematically named as discussed above (step 5). These groups are always numbered so that the point of attachment to the rest of the molecule is C1, with the longest continuous chain *beginning from the point of attachment* taken as the parent. For historical reasons, some of the simpler branched-chain alkyl groups also have nonsystematic (or *common*, or *trivial*) names. These names are shown in parentheses for the three-carbon and four-carbon alkyl groups.



The common names of the simple branched alkyl groups are so well entrenched in the chemical literature that the IUPAC rules make allowance for them. Thus, the following compound may be properly named *either* 4-(1-methylethyl)heptane or 4-isopropylheptane. These common names, and a few others that will be encountered, must be memorized.



4-(1-Methylethyl)heptane or 4-Isopropylheptane

N.B.: When writing an alkane name, the prefix iso is considered to be part of the alkyl group name for alphabetizing purposes, but the hyphenated prefixes *sec*- and *tert*- (or *t*-) are not. Thus isopropyl and isobutyl are listed alphabetically under **i** , but *sec*-butyl and *tert*-butyl are listed under **b**.

The iso prefix can be applied to any branched alkyl group with a $(CH_3)_2CH$- group connected to a straight chain. Thus, $(CH_3)_2CH_2CH_2CH_2CH_2$- is the isohexyl group.

2. **CYCLOALKANES**

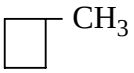
(a) named as cyclo...ane, based on the number of carbon atoms **in the ring**.

(b) for polysubstituted cycloalkanes, use the lowest-possible numbering sequence; where two such sequences are possible, the **alphabetical order** of the substituents takes precedence.

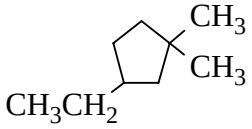
(c) cycloalkanes with two substituents on the same side are named *cis* and on opposite sides are named *trans*.

(d) rings as substituents can be named as **cycloalkyl** groups.

cyclopentane

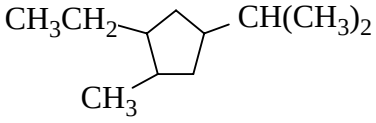


methylcyclobutane

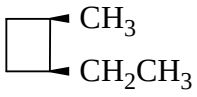


3-ethyl-1,1-dimethylcyclopentane

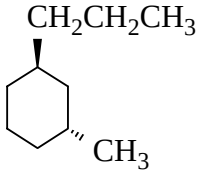
not 1-ethyl-3,3-dimethyl-cyclopentane



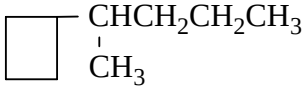
1-ethyl-4-isopropyl-2-methylcyclopentane



cis-1-ethyl-2-methylcyclobutane



trans-1-methyl-3-propyl-cyclohexane



2-cyclobutylpentane *or* (1-methylbutyl)cyclobutane

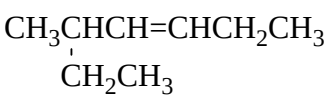
3. **ALKENES**

(a) use the longest chain **containing** the C=C and replace **-ane** with **-ene**.

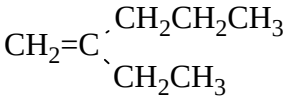
(b) number from the end that gives the first carbon of the C=C the **lowest** number; only the first carbon of the C=C is numbered.

(c) isomers with the same group on the same side of the C=C are named *cis* and those with the same group on opposite sides are named *trans*.

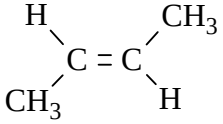
(d) alkenes with two or three C=C are named as **dienes** or **trienes**.



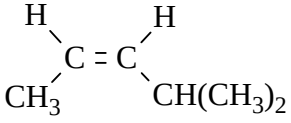
5-methyl-3-heptene



2-ethyl-1-pentene



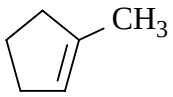
trans-2-butene



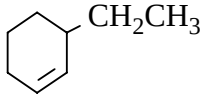
cis-4-methyl-2-pentene

CH_2=CH-CH_2-CH=CH-CH_3 1,4-hexadiene

4. **CYCLOALKENES** (C=C in the ring)
   (a) one of the carbon atoms of the C=C is always numbered C1.
   (b) number so as to give the first substituent the lowest number

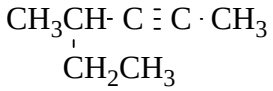


1-methylcyclopentene



3-ethylcyclohexene

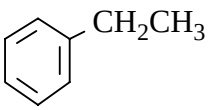
5. **ALKYNES**
   (a) use the longest chain **containing** the C≡C and replace **-ane** with **-yne**.
   (b) number from the end to give the first carbon of the C≡C the lowest number.



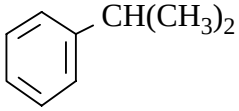
4-methyl-2-hexyne

CH_3-C≡C-C≡C-CH_2-C≡C-CH_3 2,4,7-nonatriyne

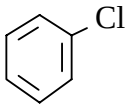
6. **AROMATIC COMPOUNDS (ARENES)**
Many are named as substituted benzenes. Examples of monosubstituted benzenes are shown below.



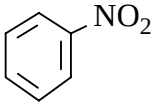
ethylbenzene



isopropylbenzene

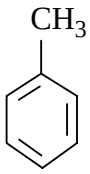


chlorobenzene

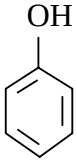


nitrobenzene

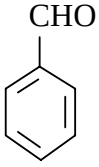
However, trivial names are used for the following:



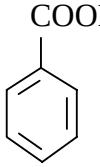
toluene



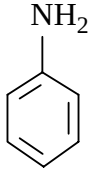
phenol



benzaldehyde

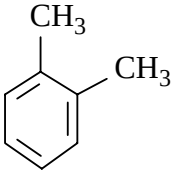


benzoic acid

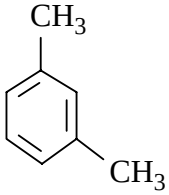


aniline

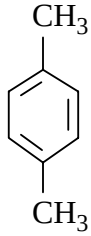
There are 3 possible isomers for a disubstituted benzene. Those for dimethylbenzene are shown and named below. (*o*, *m* and *p* stand for *ortho*, *meta* and *para*, respectively)



1,2-dimethylbenzene
o-dimethylbenzene
o-xylene

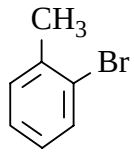


1,3-dimethylbenzene
m-dimethylbenzene
m-xylene

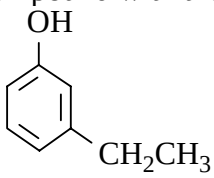


1,4-dimethylbenzene
p-dimethylbenzene
p-xylene

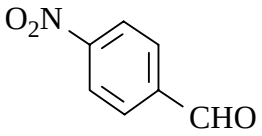
Monosubstituted toluenes, phenols, benzaldehydes, benzoic acids and anilines are named as derivatives of the parent compound with the base substituent at C1.



2-bromotoluene
or o-bromotoluene

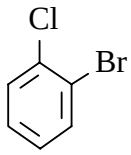


3-ethylphenol
or m-ethylphenol

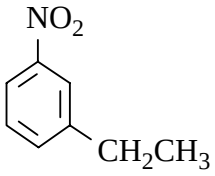


4-nitrobenzaldehyde
or p-nitrobenzaldehyde

Other disubstituted benzenes are named by putting the substituent groups in alphabetical order. If numbers are used, the first group is located at C1.

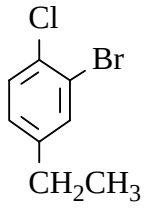


1-bromo-2-chlorobenzene
or o-bromochlorobenzene

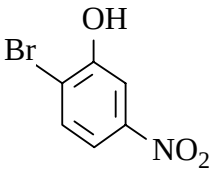


1-ethyl-3-nitrobenzene
or m-ethylnitrobenzene

When there are **three** or more substituents on the benzene ring, numbers **must** be used. Begin numbering at the substituent with the nearest next substituent. However, in substituted toluenes, phenols, etc, the base substituent is taken as C1.

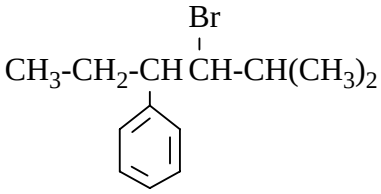


2-bromo-1-chloro-4-ethylbenzene
not 1-bromo-2-chloro-5-ethylbenzene
not 3-bromo-4-chloro-1-ethylbenzene



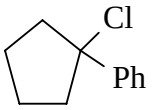
2-bromo-5-nitrophenol

A benzene ring can also be named as a **phenyl** group.



3-bromo-2-methyl-4-phenylhexane

Ph is often used as an abbreviation for the benzene ring in drawing structures of such compounds.



1-chloro-1-phenylcyclopentane

7. HALOGEN DERIVATIVES

These are named as halo derivatives as in the examples above. Simple compounds are often named as alkyl halides.

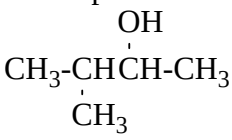
CH_3Cl	CH_3CH_2Br	$CH_3CHBrCH_3$
methyl chloride *or* chloromethane	ethyl bromide *or* bromoethane	isopropyl bromide *or* 2-bromopropane

8. ALCOHOLS

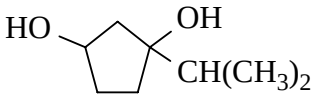
(a) use the longest chain containing the OH group and replace **-ane** by **-anol**.

(b) number from the end nearest the OH group.

(c) compounds with more than one OH are named as **diols** (2 OH), **triols** (3 OH), etc.



3-methyl-2-butanol



1-isopropyl-1,3-cyclopentanediol

(d) simple alcohols are also named as *alkyl alcohols*; for example, CH_3OH is methanol or methyl alcohol and CH_3CH_2OH is ethanol or ethyl alcohol.

9. ETHERS (simple ethers only)

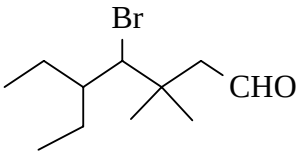
The simple ethers are named as **alkyl alkyl ethers** or **dialkyl ethers**.

CH_3-O-CH_2CH_3	$(CH_3)_2CH$-O-$CH(CH_3)_2$
ethyl methyl ether	diisopropyl ether

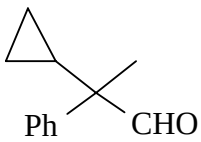
10. ALDEHYDES

(a) use the longest chain containing the CHO group and replace **-ane** by **-anal.**

(b) number the CHO as C1.



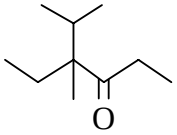
4-bromo-5-ethyl-3,3-dimethylheptanal



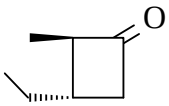
2-cyclopropyl-2-phenylpropanal

(c) the simple aldehydes are also known by their common names; e.g, CH_2O is formaldehyde and CH_3CHO is acetaldehyde.

11. KETONES
    (a) use the longest chain containg the C=O group and replace **-ane** by **-anone**.
    (b) number from the end nearest the C=O group.



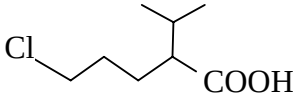
4-ethyl-4,5-dimethyl-3-hexanone



trans-3-ethyl-2-methylcyclobutanone

    (c) a few simple ketones have frequently used common names; e.g., CH_3COCH_3 is acetone and $CH_3COCH_2CH_3$ is methyl ethyl ketone (abbreviated MEK in industry).

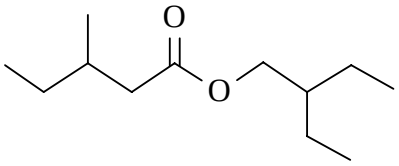
12. CARBOXYLIC ACIDS
    (a) use the longest chain containing the COOH group and replace **-ane** by **-anoic acid**.
    (b) number the COOH as C1.



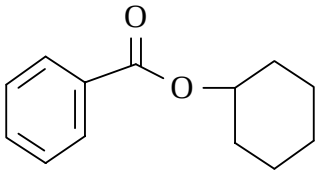
5-chloro-2-isopropylpentanoic acid

    (c) some common names are formic acid for HCOOH and acetic acid for CH_3COOH.

13. ESTERS
    (a) replace **-anoic acid** in the corresponding acid by **-anoate**.
    (b) write the name of the alkyl group in the ester **first** as a **separate word**.
    (c) number the COOR as C1.



2-ethylbutyl 3-methylpentanoate



cyclohexyl benzoate

    (d) some common names are based on the common names of acid [cf. 12(c)]; $HCOOCH_3$ is methyl formate and $CH_3COOCH_2CH_3$ is ethyl acetate.

14. AMINES (simple amines only)
    Simple amines are named as **alkylamines**, **dialkylamines**, **trialkylamines**, ***N*-alkylalkylamines** (the first alkyl group is the smaller), etc.

CH_3NH_2	$(CH_3)_2NH$	$(CH_3CH_2)_3N$
methylamine	dimethylamine	triethylamine

$CH_3CH_2NHCH_3$	$CH_3CH_2CH_2N(CH_3)_2$
N-methylethylamine	*N*,*N*-dimethylpropylamine