Nomenclature of organic compounds

The term nomenclature means the system of naming of organic compounds. In the early stages of the development of organic chemistry, Chemists knew only a few organic compounds, thus they gave names to those compounds individually and unsystematically i.e. any new compound that was discovered was given a separate name .Such names often depend on one or the other property of the compound or on the basis of the source of the compound like

- Methane was derived from the word methu. In Greek the word methu means spirit. The molecular formula is CH_3-OH. If the –OH group of methu is replaced by hydrogen atom it becomes methane, the first member of saturated hydrocarbon.

  Methane is also called marsh gas or natural gas or gobar gas in nepali because it is found in marshy places or damp places.
- Formic acid (HCOOH) was named so because it was first time obtained by the distillation of red ants (Latin-formicus means red ant)
- Acetic acid (CH_3COOH) was derived from Acetum (Latin: acetum means vinegar)
- Propane was derived from mixture of some fruits.
- Butane was derived from butter.

  These names are called trivial names or common which have no systematic basis.

In the beginning, the number of organic compounds were only few and thus the common name system worked well but with the development of organic chemistry chemist knew infinite number of organic compounds and thus it became impossible to name them without any systematic basic as it became difficult to memorize their individual name. Moreover, different common name were given to same the compound in different country and thus when

many organic compounds were discovered the common name system could not work which also affected indirectly in the development of organic chemistry.

Therefore , chemists from different parts of the world met at a congress held in Geneva in the year 1892 to decide on an international system of naming of organic compounds .The system then finalized was called Geneva System of nomenclature. This system was revised from time to time and in 1930, the council of the international union of chemistry adopted a modified Geneva system which was called International Union of chemistry and this system was known as IUC system. The international body of pure and applied chemists again modified the IUC system and the revised international union of chemistry was published in 1967. This has been accepted as a modern systematic system of nomenclature which is known as IUPAC system which stands for International Union of Pure and Applied Chemistry. This IUPAC system gives the exact name to exact structure of the molecule. This revised system of nomenclature has been accepted all over the world and is abbreviated as IUPAC system of nomenclature of organic compounds.

IUPAC System of Nomenclature:-

According to the IUPAC system of nomenclature, the name of an organic compound consists of the following parts:

1. Word root
2. Primary suffix

3. Secondary suffix
4. Prefix

IUPAC name = Prefix(s) + word root + primary suffix + Secondary suffix

1. Word Root:-

Word Root is the basic unit of the name, it represents the number of carbon atoms in the parent chain. The largest continuous carbon chain including the functional group (if any) or multiple bonds (if any) is selected as the principle chain or parent chain. Depending upon the number of carbon atoms in this principle chain, the compound is assigned as a word root. For the carbon chains up to four carbon atoms, special word root based on their common name is used. For the chains consisting of five or more carbon atoms, the word root are derived from Greek numerical as shown below.

Chain length	**Word root**
C_1	Meth
C_2	Eth
C_3	Prop
C_4	But
C_5	Pent

C_6	Hex
C_7	Hept
C_8	Oct
C_9	Non
C_{10}	Dec

2. Primary Suffix:-

Primary suffix is used to represent saturation or unsaturation in the carbon chain and is added to the word. It indicates the nature of bonding between the carbon atoms. Eg

Nature of carbon chain(Parent chain)	**Primary suffix**
Saturated carbon chain(C-C)	-ane
Unsaturated chain with one carboncarbon double bond (C=C)	-ene
Unsaturated chain with two double bonds	-diene
Unsaturated chain with three double bonds	-triene
Unsaturated with one triple bond	-yne
Unsaturated with two triple bonds	-diyne

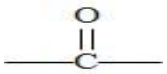
Example:-

Formula of compound	**No. of C atom in parent chain**	**Word root**	**Primary suffix**	**IUPAC name**
$CH_3CH_2CH_2CH_3$	4	But	ane	Butane
$CH_3CH = CH_2$	3	Prop	ene	Propene
$CH \equiv CH$	2	Eth	yne	Ethyne
$CH_2 = CH - CH = CH_2$	4	But a	diene	Butadine
$HC \equiv C - C \equiv CH$	4	But a	diyne	Butadiyne

Extra ‘a’ has been added to the word root in the 4th and 5th case since primary suffix i.e. diene or diyne begins with a numerical prefix .

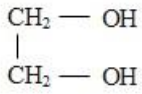
3. Seconddary suffix:-

Secondary suffix is used to indicate the functional group in organic compound. It is added to the primary suffix.

Class of organic compound	**Functional group**	**Secondary suffix**
Alcohols	-OH	-ol
Aldehydes	-CHO	-al
Ketones		-one
Carboxylic acids	-COOH	-oic acid
Esters	-COOR	-oate
Acid oxides	$-CONH_2$	-amide
Amines	$-NH_2$	-amines
Acid chlorides	-COCl	-oyl chloride
Cyanide	-CN	-nitrile

NB

- If the secondary suffix begins with a vowel then terminal e of primary suffix is dropped while adding the secondary suffix.
- If secondary suffix begins with a consonant then terminal e is retained and secondary suffix is added after e
- If the secondary suffix has some numerical prefix like di, tri etc before it then the terminal e of primary suffix is retained

Formula	**No. of C**	**Word root**	**Primary suffix**	**Secondary suffix**	**IUPAC name**
CH_3CH_2OH	2	Eth	ane	ol	Ethanol
$CH_3CH_2CH_2NH_2$	3	Prop	ane	amine	Propanamine
CH_3CH_2CN	3	Prop	ane	nitrile	Propanenitrile
$CH \equiv CCOOH$	3	Prop	yne	oic acid	Propynoic acid
	2	Eth	ane	diol	Ethanediol

4. Prefix:-

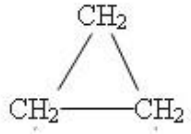
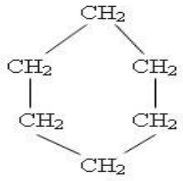
Prefix is a part of the name which appears before word root. It is of two types

- Primary prefix
- Secondary prefix

 Primary prefix:-

A primary prefix is used to distinguish alicyclic (aliphatic cyclic compounds) from acyclic compounds (open chain compounds). If the given compound is an alicyclic compound then a prefix cyclo is added just before the word root. Eg:

Compound	**No of C**	**Word root**	**Primary suffix**	**Primary prefix**	**Name**

	3	Prop	ane	cyclo	Cyclo propane
	6	Hex	ane	cyclo	Cyclo Hexane

The absence of primary prefix cyclo in the name of organic compounds indicates that it is an acyclic compound

- Secondary prefix

Secondary prefix is a part of the name which appears before the word root. Prefixes are used to indicate the names of alkyl groups or some functional group which are regarded as substituents. The secondary prefixes of some subsistent groups are as follows.

Subsituent group	**Secondary prefix**
-X(-Cl, -Br, -I)	halo
$-CH_3$	methyl
$-C_2H_5$ or CH_3-CH_2-	ethyl
$-C_3H_8$ or $CH_3-CH_2-CH_2-$	n-propyl
$(CH_3)_2CH-$	iso propyl
$(CH_2)_3C-$	tert. butyl
$-NO_2$	Nitro
$-OR$	alkoxy
$-OCH_3$	methoxy
$-OCH_2CH_3$	ethoxy
$-NH_2$	amino

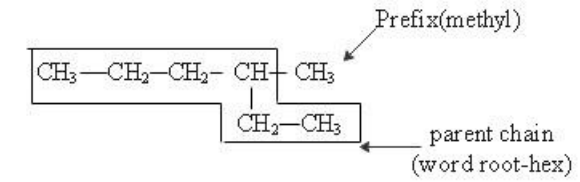
While adding a secondary prefix, following rules should be followed

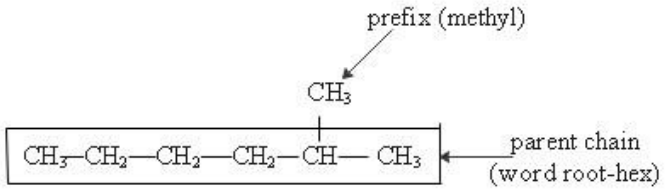
a. In case of alicyclic (aliphatic cyclic) compounds the secondary prefix is added before primary prefix (i.e. before the word cyclo) in alphabetical order.
b. In the case of open chain(acyclic) compounds the secondary prefix is added just before word root in alphabetical order

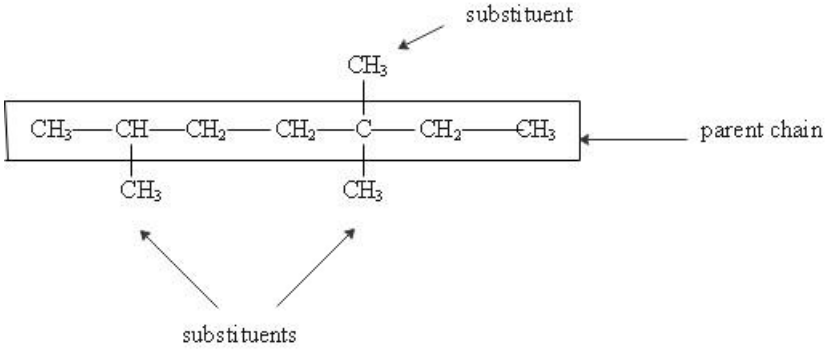
Rules for writing IUPAC names of saturated hydrocarbon or Alkanes:-

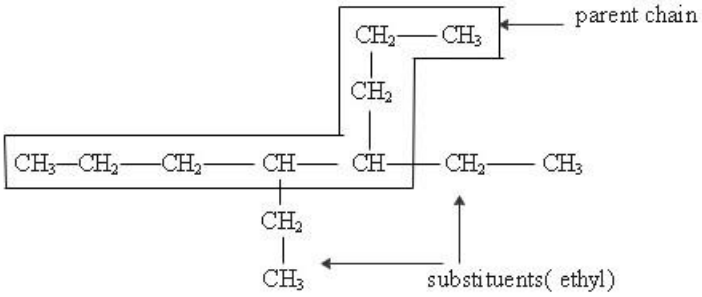
1. **Longest chain rule**:-

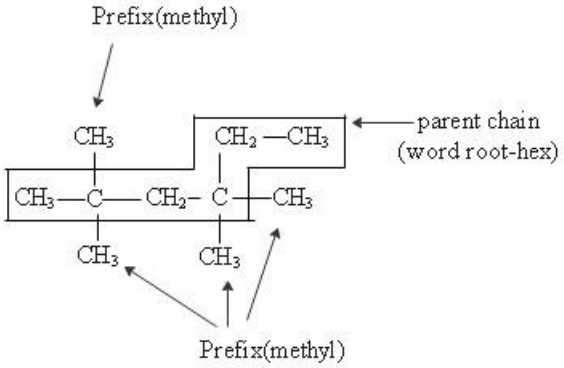
   The longest continuous carbon chain in the molecule of the given compound is selected as the principle chain or parent chain. This longest chain may or may not be straight but it must be continuous. The number of carbon atoms in this principle chain determines the word root. The carbon atoms which are not included in the parent chain are considered as alkyl substituent and determines prefixes.

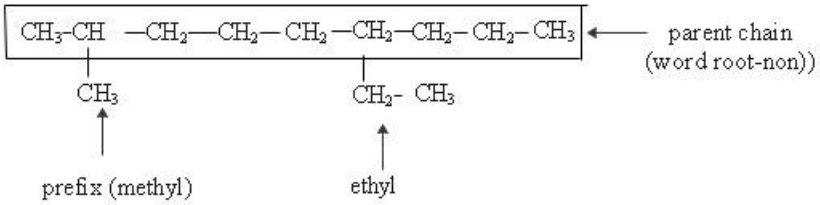




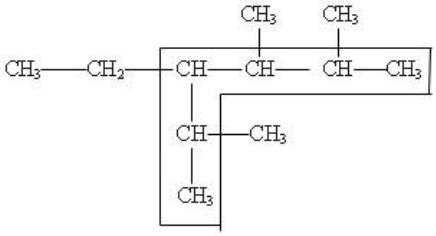




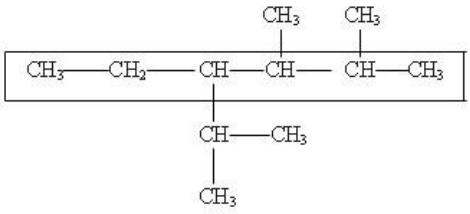




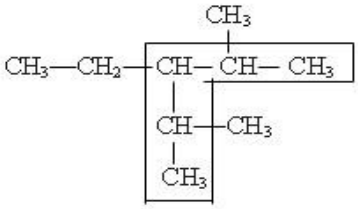
If two equally long chain are possible the chain with maximum number of side chain is selected as a parent chain.



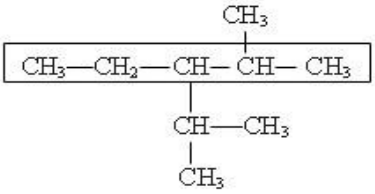
Correct chain
(Longest chain with 4 substituents)



Wrong chain
(Longest chain but only 3 substituents)



correct
(longest chain with three substituents)

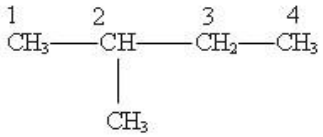


wrong
(longest chain but with only two substituents)

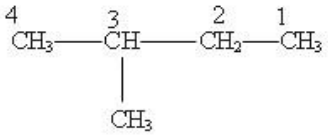
2. Substituent rule:-

If a substituent is present in the parent chain the carbon atoms in the selected parent chain are numbered as 1,2,3…. from one end to the other in such a way that the carbon atom caring the side chain or substituent gets the lowest possible number i.e. the number of carbon atoms are counted from the end which gives the lowest number to the substituent bearing carbon atom. The number indicating the position of substituent or side chain on the main chain is called locator number.

Eg.



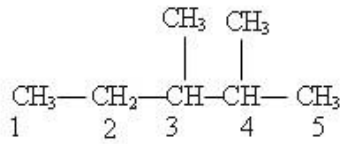
2-methyl butane (correct)



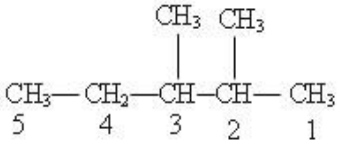
3-methyl butane (incorrect)

3. Lowest sum rule:-

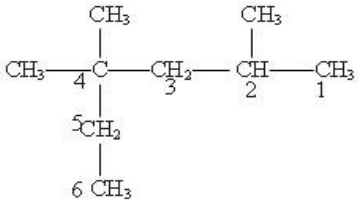
If two or more same substituents are present in a parent chain the numbering of carbon atoms are done in such a way that gives the lowest possible sum of the numbers assigned to the set of substituents.



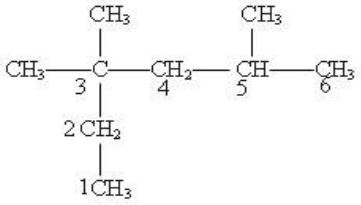
3,4-dimethyl pentane(wrong)
3 + 4 = 7



2,3-dimethyl pentane(correct)
2 + 3 = 5

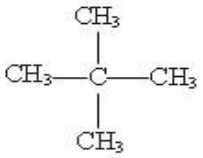


2 + 4 + 4 = 10
(correct)
2,4,4-tri methyl hexane



3 + 3 + 5 = 11
(wrong)
3,3,5-tri methyl hexane

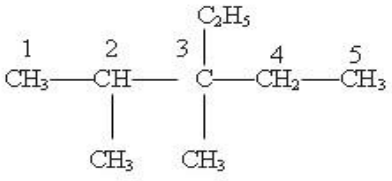
When same substituent is present more than once on the chain then the prefixes di, tri and tetra etc. are used together with a set of locants-one for each substituent and is separated by commas.



2,2-dimethyl propane

4. Alphabetical order:-

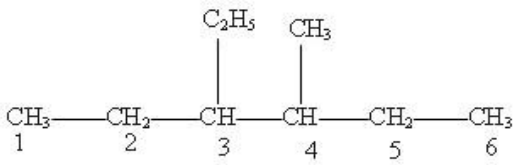
If two different substituents are present in parent chain their names are written in alphabetical order. However, the numerical prefixes such as di, tri,



3-ethyl-2,3-dimethyl pentane

tetra etc. are not considered for comparison.

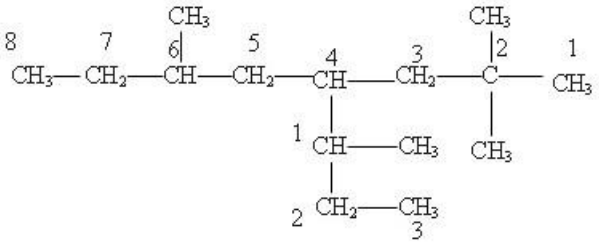
In case if two different substituents are present at an equivalent position then numbering of the chain is done in such a way that substituent which comes first in alphabetical order get the lower position. For e.g. If ethyl and methyl groups are present at an equivalent position, then carbon bearing ethyl group gets the lower number.



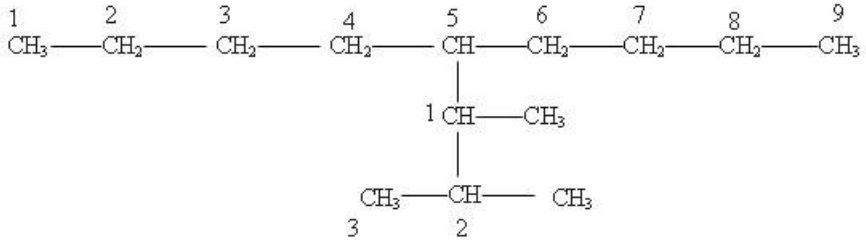
3- ethyl-4-methyl hexane

5. Complex alkyl substituents rule:-

If a substituent on the parent chain is complex i.e. it has a branch chain then it is named as a substituted alkyl group. The substituted alkyl group is then named as a separate carbon chan. The carbon atom of the substituted alkyl groups are separately numbered in such a way that carbon atom attached directly to the parent chain is given no. 1. The name of such a substituent is enclosed in bracket to avoid confusion with the numbering in the parent chain.



2,2,6-tri methyl-4-(1-methyl propyl) octane



5-(1,2-dimethyl propyl) nonane

Examples:-

1. $CH_3-CH_2-CH_2-CH_3$

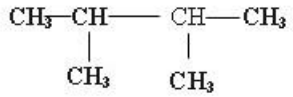
Butane

Word root-But
Pri.suffix-ane

2. $CH_3-CH(CH_3)-CH_2-CH_3$

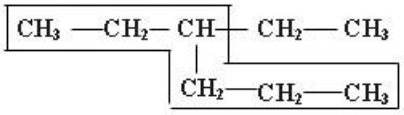
2-methyl butane

Word root-But
pri. suffix-ane
sec prefix-methyl



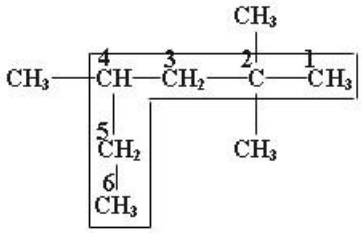
Word root-But
pri. suffix-ane
sec prefix-dimethyl

2,3-dimethyl butane



Word root-Hex
pri. suffix-ane
sec. prefix- ethyl

2,3-dimethyl butane



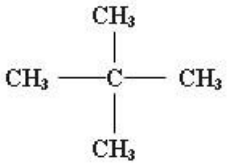
Word root-Hex
pri. suffix-ane
sec. prefix- trimethyl

2,2,4-trimethyl hexane

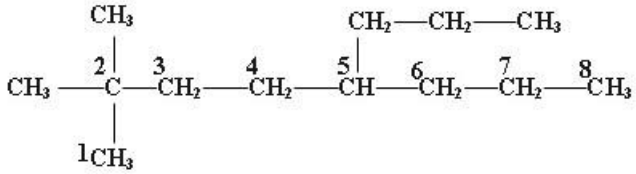
3.

4.

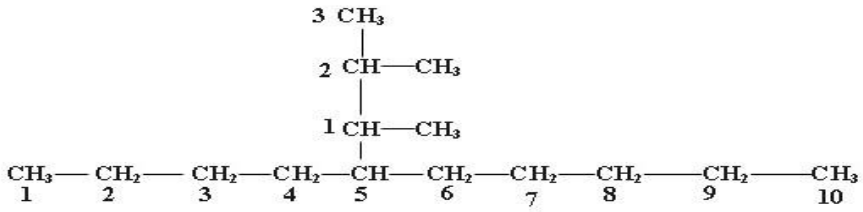
5.

6. 

2,2-dimethyl propane

7. 

5-propyl-2,2-dimethyl octane

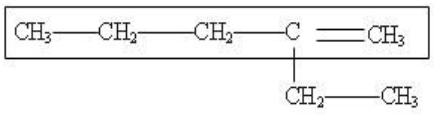
8. 

5-(1,2-dimethyl propyl) decane

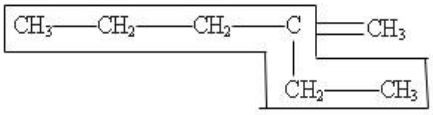
Nomenclature of unsaturated hydrocarbons (Alkenes and Alkynes):-

1. Longest Chain Rule:-

If the molecule contains multiple bonds i.e. double or triple bonds then the longest continuous carbon chain containing multiple bonds is selected as the parent chain. The parent chain must contain double or triple bond even if it is not the longest.



correct

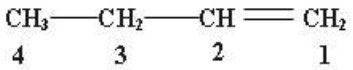


wrong

On the basis of the number of carbon atoms in the parent chain the parent alkane is determined and suffix –ane of the parent alkane is replaced by – ene for alkene and –yne for alkynes.

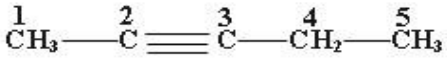
2. Lowest number rule:

If the parent chain contains multiple bonds then the carbon atoms of the selected principle chain are given the locator number so as to give the lowest number to the carbon atom joining to the double bond or triple bond.



1-butene
but-1-ene
Butene-1

The position of the multiple bonds can be indicated by adding the serial number of the first carbon of the multiple bond before suffix-ene.

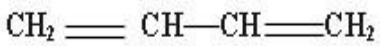


2-pentyne
pentyne-2
pent-2-yne

3. Primary suffix for the compounds containing more than one multiple bonds:-

If there are two or more multiple bonds in the parent chain, then the suffix – diene, triene respectively are used. The position of the multiple bonds are indicated by the numerical indicating the carbon atoms at which the multiple bonds starts. The numbering of the chain is done so that the sum of the numerical assigned to the carbon atoms linked by multiple bonds is the lowest.

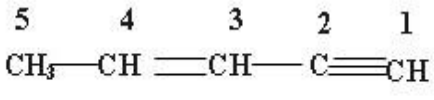
Eg.



1,3-butadiene
or buta-1,3-diene

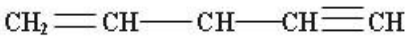
If both the double and triple bonds are present in the same molecule their locants are written before their respective suffixes. The name is ended in yne. The terminal e from suffix ene is dropped while writing complete name.

Eg.



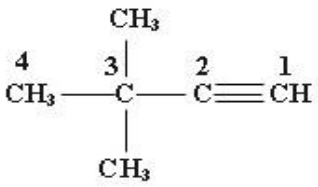
Pent-3-en-1-yne

If the lowest sum for multiple bonds i.e. double and triple bonds becomes same from both the ends then the priority is given to the double bond.

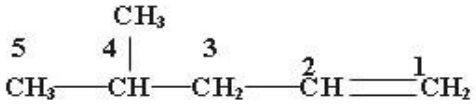


pent-1-en-4-yne

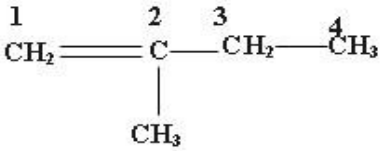
Examples:-

1. 

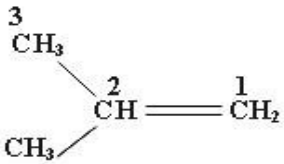
3,3-dimethyl-but-1-yne

2. 

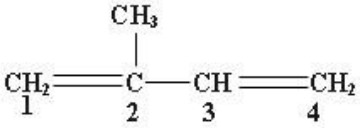
4-methyl-pent-1-ene

3. 

2-methyl but-1-ene

4. 

2-methyl propane

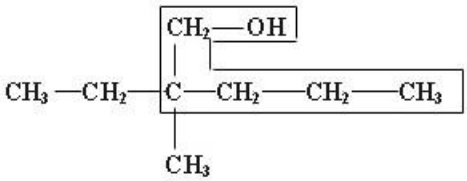
5. 

2 -methyl-1,3-but-a-diene

Nomenclature of compounds containing functional groups:-

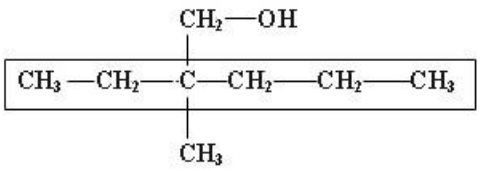
1. Longest chain rule:-

For a compound containing one functional group the longest continuous carbon chain including the functional group is taken as the parent chain. If the functional group contain carbon atom, that carbon atom is also counted as a part of the principle chain. Therefore the longest continuous chain of carbon atom including the carbon atoms of the functional group like –CHO, -CN, -CO etc. or those carbon atoms which are joined to the functional groups like –OH, -NH_2, -Cl etc. is selected as the parent chain. The total number of carbon atoms in this longest continuous chain including functional group determines the word root.



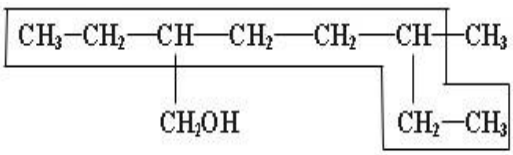
word root-pent

correct



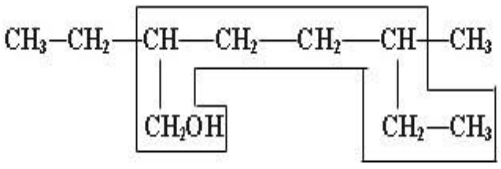
word root-hex

wrong



word root-oct

wrong



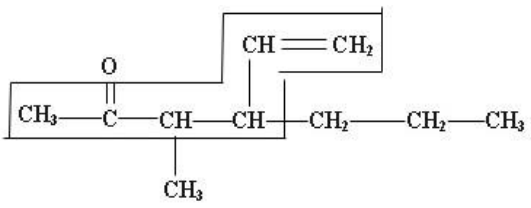
word root- hept

correct

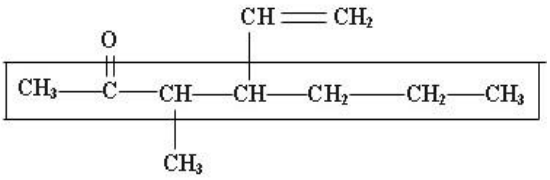
Here, the chain containing the functional group is taken as the parent chain even if it is not the longest.

If the compound contains both the functional group and double or triple bond, then the carbon chain containing both the functional group and

multiple bond is taken as the parent chain even if it is not the longest carbon chain.



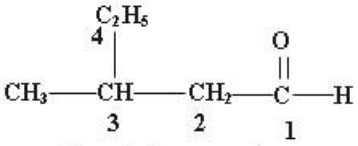
correct



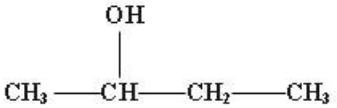
wrong

2. Lowest number rule:-

If the compound contains functional group like –COOH, -CHO in the parent chain the carbon atom of the parent chain are counted in such a way that carbon atom of the functional group gets no. 1. In case the functional group does not have carbon atom then the carbon atom of the parent chain attached to the functional group gets the lowest possible number.

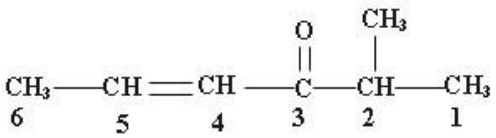


3-methyl pentanal



Butanol-2

If the parent chain contains functional group and multiple bond in same chain then the numbering is done in such a way so as to give lowest possible number to the carbon containing functional group or carbon atom of the functional group followed by the multiple bond.

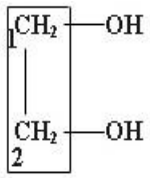


2-methyl hex-4-en-3-one

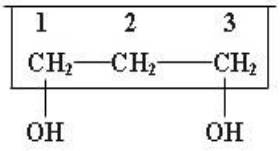
3. Lowest sum rule:-

If the parent chain contains more than one similar functional group the numbering of the carbon atoms are done from one end to the other to give the lowest possible sum of the number assigned to the set of

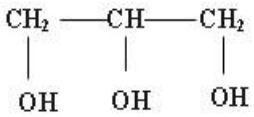
functional groups. The numerical prefixes di, tri etc. are added before the secondary suffix which indicates the functional group and while adding such words terminal e of the primary suffix is retained.



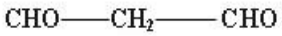
1,2-ethane diol



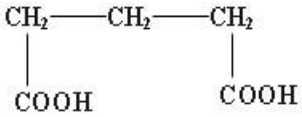
1,2,3-propane diol



1,2,3 propaane triol



1,3 propane dial



1,5 pentane dioic acid

4. Polyfunctional rule:-

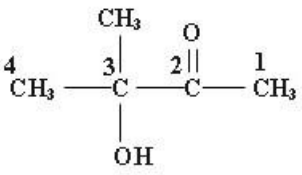
If the parent chain contains more than one dis-similar functional groups, the number of carbon atoms are counted so as to give the lowest number to the functional group of higher priority. The priority order of various functional groups are as follows.

-COOH > -COOR > $-SO_3H$ > -COX > $-CONH_2$ > -CHO > -CN > >C=O > -OH > $-NH_2$ > $-NO_2$ > -X > -R

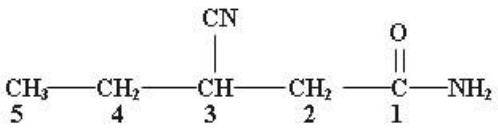
Priority or Seniority Table

Group	Prefix	Suffix
-COOH	carboxy	oic acid
-COOR	alkoxy carbonyl	oate
$-SO_3H$	sulpho	sulphonic acid
-COX	halo formyl	oyl halide
$-CONH_2$	carbamoyl	amide
-CHO	formyl or aldo	al
-CN	cyano	nitrile
-NC	isocyano	isonitrile
>C=O	keto or oxo	one
-OH	hydroxy	ol
$-NH_2$	amino	amine
-OR	alkoxy	-
>C=C<	-	ene
-C≡C-	-	yne
$-NO_2$	nitro	-
-X	halo(chloro,bromo,iodo)	-
-R	alkyl	-

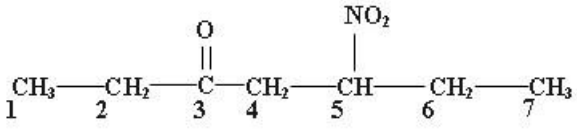
Examples:-



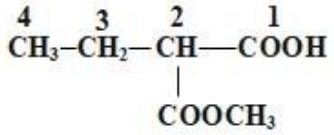
3 hydroxy-3-methyl-but-an-2-one



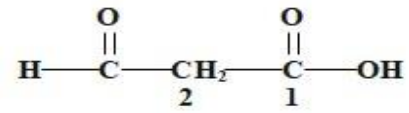
3 cyano-1-pentamide



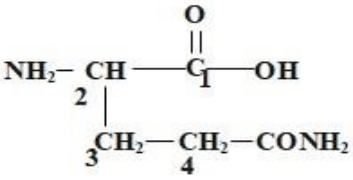
5 nitro-3-heptanone
or
5-nitro-heptan-3-one



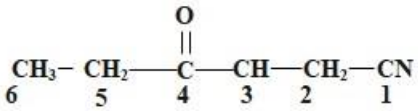
2-Carbomethoxy-butanoic acid



2-Aldo-ethanoic acid
or
2-Formyl ethanoic acid



2-Amino-4-carbamido-butanoic acid



4-keto- hexanenitrile

Nomenclature of compound having more than two same terminal functional groups (-COOH, -CHO, -CN, -COCl)

When more than two same **terminal functional groups** are present in the compound, then there are two conditions for naming them

Condition-1

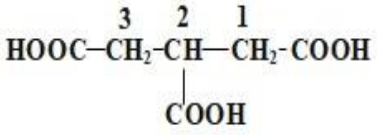
If all the functional groups are directly bonded to the parent chain, carbon atoms of the terminal functional groups are not counted as the part of principle chain and named by giving special suffix like

Functional group	Special suffix
-COOH	-carboxylic acid
-COOR	-carboxalate
-COCl	-carbonyl chloride
$-CONH_2$	-carboxamide
-CN	-carbonitrile
-CHO	-carbaldehyde

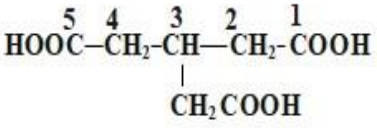
Condition-2

If one of the terminal functional group is not directly bonded to the parent chain, then the longest continuous carbon chain between two same terminal functional groups is selected as the parent chain. The remaining functional groups are treated as substituent.

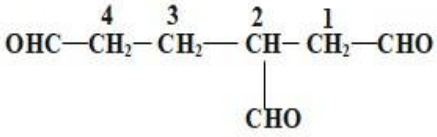
For example:-



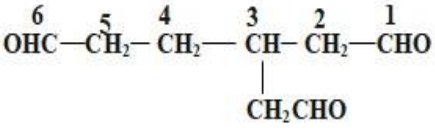
Propane-1,2,3-tricarboxylic acid



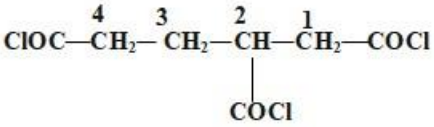
3-(carboxymethyl)pentanedioic acid



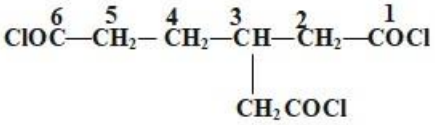
Butane-1,2.4-tricarbaldehyde



3-(Formylmethyl)hexanedial

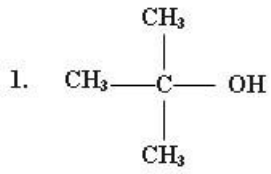


Butane-1,2,3-tricarbonyl chloride



3-(Chloroformylmethyl)hexanedioyl chloride

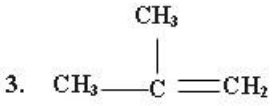
Some more examples:-

1. 

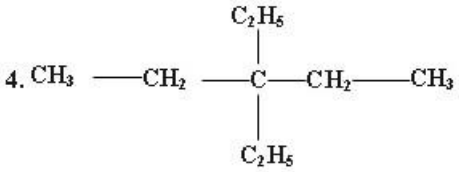
2 methyl-2-propanol

2. $CH_3\,CH_2\,CONH_2$

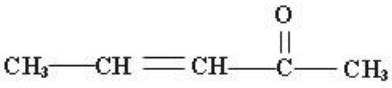
Propanamide

3. 

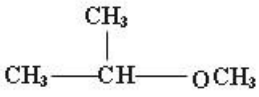
2 methyl-prop-1-ene

4. 

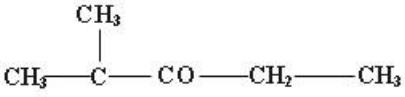
3,3-diethyl pentane

5. 

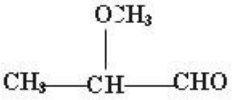
Pent-3-en-2-one

6. 

2-methoxy propane

7. 

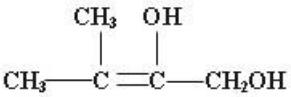
2 methyl pent-an-3-one

8. 

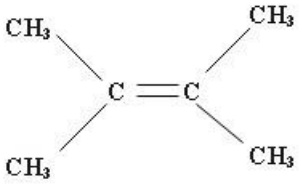
2 methoxy propanal

9. $HCOOCH_3$

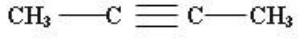
methyl methanoate

10. 

3 methyl but-2-en-1,2 diol

11. 

2,3-dimethyl but-2-ene

12. 

But -2 -yne
or
Butyne-2