Course Code/name	Block	Unit Name	Page number
DCECHE -108 ORGANIC CHEMISTRY -III (SELECTED TOPICS IN ORGANIC CHEMISTRY)	Block- 1	**Unit 1: NMR (PMR) Spectroscopy**	**1-33**
		Unit 2: Organometallic Compounds	**1-14**
		Unit 3: Sulphur Containing Compounds	**1-43**
		Unit 4: Amino Acids, Peptides, Proteins and Nucleic Acids	**1-58**
	Block- 2	**Unit 5: Active Methylene Group**	**1-24**
		Unit 6: Carbohydrates	**1-69**
		Unit 7: Problem based on Spectroscopy (UV-Vis., IR and PMR)	**1-39**

Curriculum Design Committee

Dr. Ashutosh Gupta,
Coordinator
Director, School of Sciences, UPRTOU, Allahabad
Prof. Jagdamba Singh
Member
Dept. of Chemistry.,
University of Allahabad, Allahabad
Prof. S.S. Narvi,
Member
Dept. of Chemistry.,
MNNIT, Allahabad
Prof. Arun K. Srivastava
Member
Dept. of Chemistry.,
University of Allahabad, Allahabad
Dr. Dinesh Kumar Gupta
Member Secretary
Academic Consultant-Chemistry
School of Science, UPRTOU, Allahabad

Course Preparation Committee

Prof. Nameer Al-Hasan **Author (Unit-1 and 2)**
Dept. of Chemistry,
Sambhunath college of Engg. And Technology, Prayagraj
Dr. Brijesh Kumar Singh , **Author (Unit-3, 4,5 and 6)**
Dept. of Chemistry,
BBS College of Engg. And Tech, Prayagraj, U. P., India

Dr. Santosh Singh ,
Author (Unit-7)
Dept. of Chemistry, DDU Govt. PG College,
Saidabad,Prayagraj, U. P., India
Prof. N. D. Pandey
Editor(Unit-1 and 2)
Ex. Head, Dept. of Chemistry
M.N.N.I.T, Prayagraj
Prof. M. D. Pandey **Editor (Unit-3, 4,5,6 and 7)**
Dept. of Chemistry, Institute of Science
BHU, Varanasi, U. P., India
Dr. Dinesh Kumar Gupta, **SLM Coordinator**
Academic Consultant- Chemistry School of Science,
UPRTOU, Prayagraj, U. P., India

Course Code- **DCECHE -108**
Course Name- **ORGANIC CHEMISTRY III (SELECTED TOPICS IN ORGANIC**
©U.P. Rajarshi Tandon Open University
Year- 2022

UNIT. 1 N M R SPECTROSCOPY

Structure

1.1 Introduction

Objective

1.2 Theory of NMR

1.3 Number of signals

1.4 Position of signals (Chemical shift)

1.5 Shielding and Deshielding effects

1.6 Peak area and Proton counting

1.7 Spin – spin coupling

1.8 Splitting of the signals

1.9 Coupling constants (J)

1.10 Important tips for interpretation of NMR spectrum

1.11 Summary

1.12 Interpretation of spectra of some common compounds

1.13 Terminal Questions

1.14 Answers to Terminal Questions

1.1 INTRODUCTION

Spectroscopy is the cornerstone of Organic chemistry since the second half of twentieth century. One of the analytical method that has the greatest impact on science has been Nuclear Magnetic Resonance (NMR) Spectroscopy. The effect of NMR has come in three distinct ways. Although the first of application came in 1950s, the tool was not widely used until the advent of Varian Associates a-60 spectrometer in the early 60s. The new experiment provided the final piece of the structural puzzle in many cases. The effect on organic chemistry was immediate and electrifying, with virtually every journal and publication dealing with organic chemistry included the NMR data as the most important structural evidence. The pace of progress in structure determination increased perceptibly.

The second identifiable stage in the introduction of NMR into structural chemistry resulted from application of the Fourier Transform (FT) method to the NMR experiment. It allows use of NMR to C13 isotope also.

The third surge of activity has come in this decade and is a daughter of the computer revolution. "Two Dimensional" (2-D) NMR is a still burgeoning aspect of the method that may influence the way chemists think about structure determination more than any of the earlier spectroscopic techniques.

Electronic spin resonance, also using radio frequency signals, detects unpaired electrons and can be used to measure the distribution of electron density in radicals. Mass spectrometry measures the mass – to – charge ratio of organic ions. Structural information comes from the moderately predictable fragmentation organic molecules undergo; the masses of the fragment ions can often be related to likely structures.

NMR spectroscopy is an important tool in the hands of an organic chemist for getting structural information from the spectrum of an unknown compound. It also helps in studying the stereo chemical details within the molecule. Although important, it cannot replace other techniques such as ultraviolet, infra-red, Mass etc. Leaving aside the functional groups, a large part of an organic molecule consists of carbon-hydrogen skeleton and this tool is most useful in the investigation of this and is gaining importance in the quantitative analysis of the compounds. For example, the ratio of alcohol and water in the given sample can be determined from the values of delta. It becomes possible due to the property of chemical exchange

Objective

After studying this unit you should be able to:

- Draw structures of compounds as per the spectrum
- Predict the structure from an NMR spectrum
- Predict the spectrum from a given structure
- Help in qualitative analysis of a compound
- Help in quantitative analysis of a compound and check purity
- Identify the reaction progress

1.2 THEORY OF NMR

The nuclei with non-zero spins like a Hydrogen atom (Proton) behaves as a spinning bar magnet because it possesses both electric and magnetic spin. Like any other spinning charged body, the nucleus of Hydrogen atom also generates a magnetic field

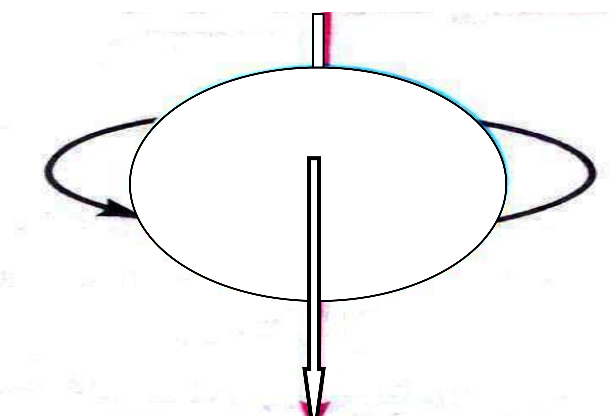


Fig 1.1 Spinning charge in nucleus generates magnetic field

Nuclear Magnetic Resonance involves the interaction between and oscillating magnetic field of electromagnetic radiation and the magnetic energy of the hydrogen nucleus or some other type of nuclei when these are placed in and external static magnetic field. The sample absorbs electromagnetic radiations in radio wave region in different frequencies since absorption depends upon the type of protons or certain nuclei contained in the sample. Consider a spinning top. It also performs a slower waltz like motion, in which the spinning axis of the top moves slowly around the vertical axis. This is **precessional motion** and the top is said to be precessing around the vertical axis of earth's gravitational field. The precession arises from the interaction of spin with earth's gravity acting vertically downwards. It is called **gyroscopic motion**. Considering a small spinning magnet in an external field it is found that the proton (tiny magnet) precesses about the axis of the external magnetic field in the same manner in which a spinning gyroscope precesses under the influence of gravity.

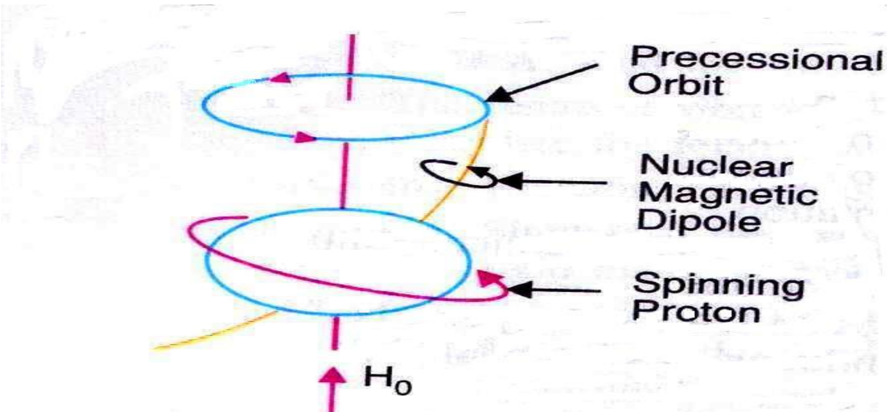


Fig (1.2) Proton precessing in magnetic field Ho

It has been found that

$$\omega = \gamma H_o \qquad \ldots \text{(i)}$$

Where ω = angular precessional velocity

H_o = applied field in Gauss

γ = Gyromagnetic ratio = $2\pi\mu / hI$

Here μ = magnetic moment of the spinning magnet

I = spin quantum number of the spinning magnet

h = Planck's constant

According to the fundamental NMR equation which correlates electromagnetic frequencies with the magnetic field, we say that

$$\gamma H_o = 2\pi\nu \qquad \ldots \text{(ii)}$$

here, ν is the frequency of electromagnetic radiation.

From (i) and (ii),

Angular precessional velocity $\omega = 2\pi\nu$

The value of this frequency (ν) inserted is called **Precessional Frequency.**

The precessional frequency may be defined as the number of revolutions per second made by the magnetic moment vector of the nucleus around the external field H_o. Alternatively, the precessional frequency of the spinning bar magnet (nucleus) may be defined as equal to the frequency of electromagnetic radiations in megacycles per second necessary to induce a transition from one spin state to another.

All nuclei carry a charge. So, they will possess spin angular momentum, which is quantized, i.e., only those nuclei which have a finite value of spin quantum number ($I > 0$) will precess along the axis of the rotation. It is known that the spin quantum number I is associated with mass number and atomic number of the nuclei.

Mass number	*Atomic number*	*Spin quantum number I*
Odd	odd or even	½, 3/2, 5/2, …
Even	even	0
Even	odd	1, 2, 3, …

The circulation of the nuclear charge generates a magnetic moment along the axis. The intrinsic magnitude of the generated dipole is expressed in terms of magnetic moment μ.

If a proton is placed in a magnetic field, then it starts precessing at a certain frequency in the radiowave region and thus, will be capable of taking up one of the two orientations with respect to the axis of the external field.

(i) Alignment with the field and

(ii) Alignment against the field.

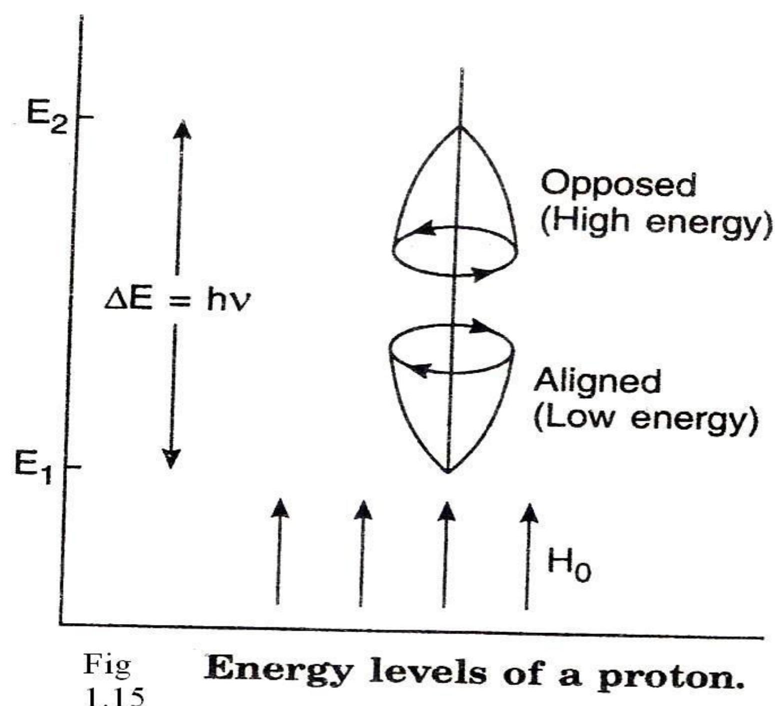
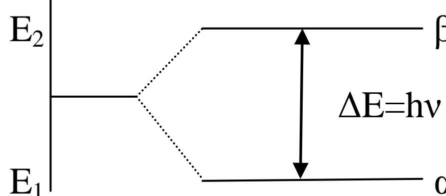


Fig 1.15 **Energy levels of a proton.**

If a proton is precessing in the aligned orientation, it can pass into the opposed orientation by absorbing energy. From the high energy opposed orientation, it comes back to the low energy aligned orientation (more stable) by losing energy. The transition from one energy state to the other is called **flipping** of the proton. The transition between the two energy states can be brought about by the absorption of a quantum of electromagnetic radiation in the radio wave region with energy $h\nu$.



The energy required to bring about the transition ($\Delta E = h\nu$) or to flip the proton depends upon the strength of the external field. Stronger the field, greater will be the tendency of the nuclear magnet to remain lined up with it and higher

will be the frequency of the radiation needed to flip the proton to the higher energy state.

We know that $\nu = \gamma H_o / 2\pi$

Where ν = frequency in cycles per second or Hz

H_o = strength of the magnetic field in Gauss

γ = nuclear constant or Gyromagnetic ration and is equal to 26750 for the proton.

In a field of 14092 gauss, the energy required to cause flipping corresponds to electromagnetic radiation of frequency 60 million cycles per second or 600 MHz (Radiowave region). The energy is much lower than that possessed by Infra Red radiation. If we irradiate the precessing nuclei with a beam of electromagnetic radiation of desired frequency, then the low energy nuclei will absorb it and move to higher energy state. The precessing proton will absorb energy from the radiofrequency source only if the precessing frequency is same as the frequency of the radio frequency beam, i.e., when the quantum energy ($h\nu$) of electromagnetic radiation matches up the energy difference between the two energy states at the magnetic field H_o. When this occurs, the nucleus and the radio frequency beam are said to be in resonance. Hence, the term **Nuclear Magnetic Resonance**.

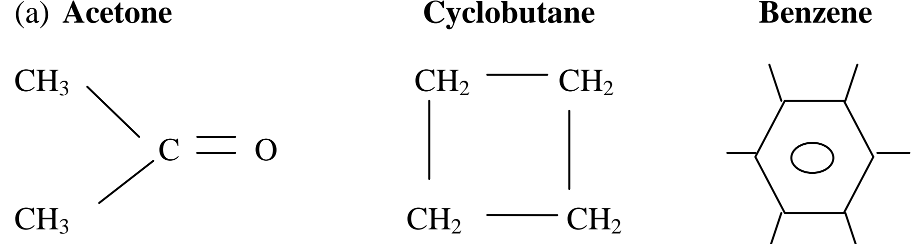
This technique consists of placing the proton (substance) in an organic molecule to a powerful field. The protons will precess at different frequencies. Now, we irradiate these precessing protons with steadily changing frequencies (for promoting or flipping protons from low energy state to high energy state) and observe the frequency / frequencies at which absorption occur. It is generally more convenient to keep the radio frequency constant and the strength of the magnetic field is constantly varied. At some value of the field strength, the energy required to flip the protons matches the energy of the radiation. Absorption occurs and a signal is observed. Such a spectrum is called **Nuclear Magnetic Spectrum.**

It may be noted that all protons do not absorb at the same applied field but absorption depends upon the magnetic field which a particular proton feels. Clearly, the effective field strength is different for different sets of protons as one set of protons will have slightly different **environment** from any other set of protons. Thus, at a given radio frequency, different protons (different sets of equivalent protons) will require slightly different applied field strengths to produce the same effective field strength which causes absorption. In the NMR spectrum, we measure the applied field strengths for each set of protons and the absorption peaks are plotted. The number of signals at different applied field strengths are equal to the different sets of equivalent protons.

1.3 NUMBER OF SIGNALS

The number of signals in an NMR spectrum tells the number of sets of **equivalent** protons in a molecule. Each signal corresponds to a set of equivalent protons. **Magnetically equivalent protons are chemically equivalent**, e.g.

(a) **Acetone** **Cyclobutane** **Benzene**



In acetone, all six protons have exactly similar environment, therefore, only one signal is observed. Similar is the case in cyclobutane and benzene where all the protons being identical give only one signal.

(b) Compounds showing more than one signal are :

Methanol, methyl acetate, p-xylene, etc

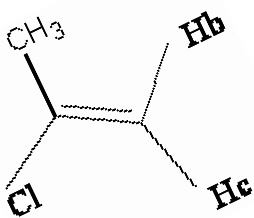
$CH_3 - CH_2 - OH$

Here, in ethanol, 3 types of protons are there so three signals and in methyl acetate and p-xylene only two signals will be obtained.

(c) Chemically equivalent protons must also be stereo-chemically equivalent, i.e., a particular set of protons are said to be Chemically equivalent only if they remain in exactly similar environment when the streo-chemical formula of the molecule under consideration is written.

Considering the case of 2-chloropropene, CH_3-C(Cl)=CH_2

One can expect two sets of equivalent protons (Two peaks), but stereo chemical formula reveals three sets of protons in it.



Here, H^a and H^c are not in exactly similar environment.

Considering methyl cyclopropane, one can expect three sets of protons. But actually four signals are observed as is clear from the stereochemical formula.

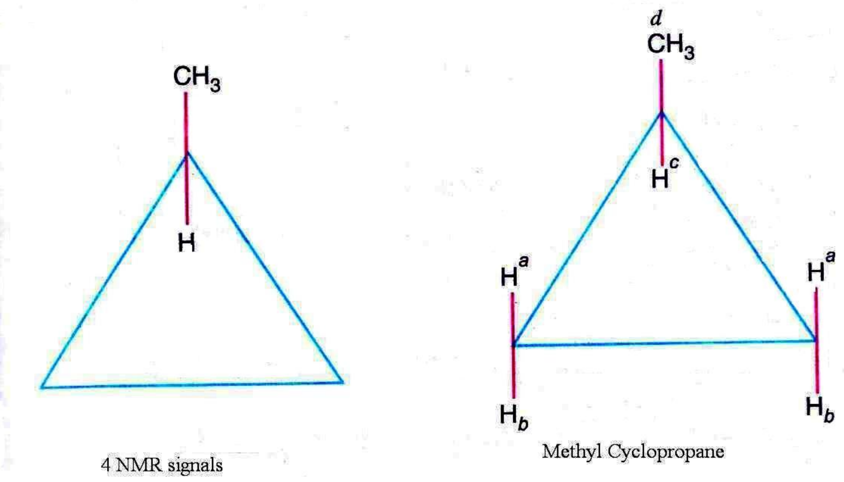


Fig (1.4) Number of signals in Methyl Cyclopropane

In 1,2-dichloropropane, CH_3 CH(Cl)-CH_2Cl one should expect three signals, but actually four signals are observed because two hydrogen atoms attached with C1 are not in exactly similar environment. Rotation around C-C single bond in this molecule cannot bring similar environment for the said hydrogen atoms.

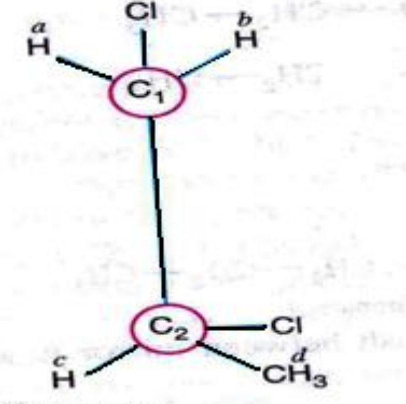


Fig (1.5) 1,2-dichloropropane
Four signals in 1,2-dichloropropane

1.4 POSITION OF SIGNALS (CHEMICAL SHIFT)

The number of signals in an NMR spectrum tell the number of the sets of equivalent protons in a molecule.

The position of the signals in the spectrum helps us to know the nature of the protons, viz., aromatic, aliphatic, acetylenic, vinylic, adjacent to some electron attracting or electron releasing group etc. Each of these type of protons will have different electronic environments, they absorb in different electronic environments and thus, they absorb at different applied field strengths. The electronic environment indicates where a proton shows absorption in the spectrum.

When a molecule is placed in a magnetic field, its electrons are caused to circulate and thus, they produce secondary magnetic fields i.e., induced magnetic fields. Rotation of electrons about the protons itself generates a field in such a way that at the proton, it opposes the applied field. Thus, the field felt by the proton, diminishes and the proton is said to be ***shielded***. Rotation of electrons (especially π) about the nearby nuclei generates a field that can either oppose or reinforce the applied field at the proton. If the induced field opposes the applied field, then proton is said to be shielded, but if the induced field reinforces the applied field, the proton feels higher field strength and such a proton is said to be ***deshielded***. Shielding shifts the absorption upfield and deshielding shifts the absorption downfield to get effective field strength necessary for absorption.

Such shifts (compared to a standard reference) in the positions of NMR absorptions which arise due to the shielding or deshielding of protons by the electrons are called ***Chemical shifts***. For measuring various shifts in a molecule, the signal for tetra methyl silane (TMS) is taken as a reference. Due to the low electro negativity of Silicon, the shielding of equivalent protons in TMS is greater than most of the organic molecules.

Therefore, NMR signal for tetramethyl silane is taken as a reference and chemical shift for different kinds of protons are measured relative to it. Clearly, the NMR signal for a particular proton in a molecule will appear with different field strengths compared to a signal from TMS.

This difference in the absorption position of the proton with respect to TMS signal is called chemical shift (denoted by δ - value). It is not measured in Gauss but is measured in equivalent frequency units which is then divided by the frequency of the spectrometer used. This gives the value of delta (δ). Protons with the same chemical shift are called equivalent protons. Non equivalent protons have different chemical shifts. δ (Delta) or **τ**(Tau) scales are commonly used. In majority of organic compounds, protons resonate at a lower field than the protons of tertra methyl silane (TMS). Thus assigning delta δ value for TMS equal to zero, a scale can be defined in which most proton resonances are of the same sign. Any proton or sets of protons which absorb a

lower field than TMS is given a positive value for δ, Tetramethyl silane is the most convenient reference and has the following characteristics.

(i) It is miscible with almost all organic substances,
(ii) It is highly volatile(low boiling point) and is readily removed from the system,
(iii) It does not take part in intermolecular associations with the sample

The values of δ for a substance with respect to TMS can be obtained by measuring $\nu_s - \nu_{TMS}$,

where, ν_s = Resonance frequency of the sample, and

ν_{TMS} = Resonance frequency of TMS

keeping the radio frequency constant, NMR signal for a particular set of proton in the sample will appear at different field strength than the signal obtained from TMS. Thus, the value for $\nu_s - \nu_{TMS}$ can be expressed in the corresponding field strength in gauss. Since, the operating frequency of the instrument is directly proportional to the strength of the magnetic field, we can define δ as

$$\delta = (\nu_{sample} - \nu_{reference}) \text{ / operating frequency in megacycles (per sec)}$$

$$= \Delta\nu \text{ / operating frequency in megacycles (per sec)}$$

where $\Delta\nu$ is the frequency shift. The value δ is expressed in parts per million (ppm). Most chemical shifts have δ values between 0 and 10. In the Tau scale, signal for the standard reference TMS is taken as 1- ppm. It has been found that signals for fluorine resonances (absorption due to fluorine nuclei) do not appear in this range but are observed downfield by about 50 ppm or even more.

The two scales are related by the expression:

$$T = 10 - \delta$$

NMR signal is usually plotted with magnetic field strength increasing to the right. Thus, signal for TMS (highly shielded) appears at the extreme right of spectrum with $\delta = 0$ ppm. **Greater the deshielding of the protons, larger will be the value of δ.** For most of the organic compounds, signals appear downfield to the left of TMS signal. The value of chemical shifts for protons in different environments are:

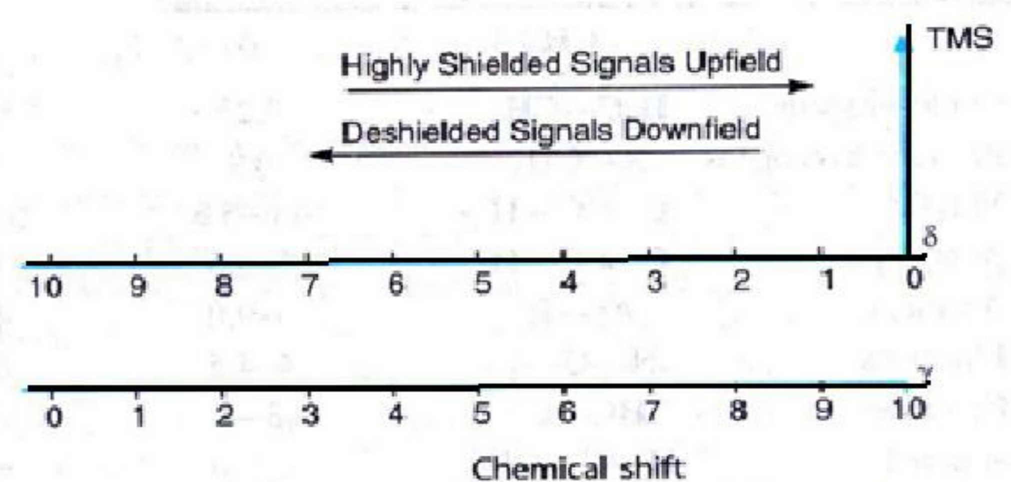


Fig (1.6) Chemical shifts in different environments

TABLE – 1.1

APPROX CHEMICAL SHIFTS OF PROTONS

Type of Protons	δ	τ
$CH_3 - C$	0.9	9.1
$CH_3 - C - C = C$	0.1	8.9
$CH_3 - C - O$	1.6	8.4
$CH_3 - C = C$	1.4	8.6
$CH_3 - Ar$	2.3	7.7
$CH_3 - Oar$	3.8	6.2
$CH_3 - N$	2.3	7.7
$CH_3 - Cl$	3.0	7.0
$CH_3 - Br$	2.7	7.3
$C - CH_2 - C$	1.3	8.7
$C - CH_2 - C - C = C$	1.7	8.3
$C - CH_2 - C - O$	1.9	8.1
$C - CH_2 - C = C$	2.3	7.7
$C - CH_2 - Ar$	2.7	7.3
$- CH_2 - Cl$	3.4	6.6
O.CH – Ar	3.0	7.0
$CHCl_{3SS}$	7.25	2.75
H_2O	5.0	5.0
Ar – H	7.2	2.8

If the observed shift from TMS is 250 Hz and the operating frequency of the instrument is 100 MHz, then the chemical shift in terms of δ is given by the expression:

$$\delta = 250 \text{ Hz} / (100 \times 10^{6} \text{Hz}) = 2.5 \times 10^{-6}$$

The frequency ration is then multiplied by 10^6 and the resulting value is chemical shift which is expressed as parts per million (10^6) of the operating frequency.

$$\delta \text{ (Chemical shift)} = [\nu_s - \nu_{TMS} \text{ (in Hz)} / \text{Operating frequency of instrument}] \times 10^6 \text{ ppm}$$

$$= 2.5 \times 10^{-6} \times 10^6 \quad = \quad 2.5 \text{ ppm}$$

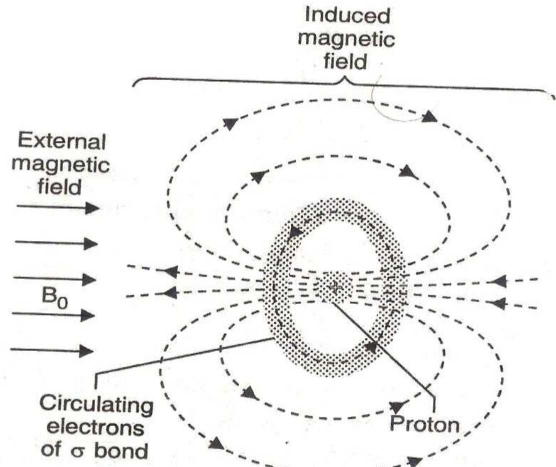
Thus, the chemical shift is independent of the operating frequency of the instrument but depends on the solvent. Thus, it becomes essential that the value of chemical shift is reported alongwith solvent used for running the spectrum.

1.5 SHIELDING AND DESHIELDING EFFECTS

Hydrogen nuclei in a molecule are surrounded by the electronic charge which shields the nucleus from the influence of the applied field. Thus, to overcome the shielding effect and to bring the protons to resonance, greater external field is required. Therefore, greater the electron density around the proton, greater will be the induced secondary magnetic field (Local diamagnetic effect) which opposes the applied field and thus, greater external field will cause proton absorption. The extent of shielding is represented in terms of shielding parameter α. When absorption occurs, the field H felt by the proton is represented by:

$$H = H_0(1 - \alpha) \qquad \ldots(1)$$

where H_0 is the applied field strength. Thus, the field felt by the proton is not the same as the applied field.



The circulations of the electrons of a C—H bond under the influence of an external magnetic field. The electron circulations generate a small magnetic field (an induced field) that shields the proton from the external field.

Greater the value of α, greater will be the value of the field strength which has to be applied to get the effective field required for absorption and vice versa.

Also $\nu = \gamma H / 2\pi$ …(2)

From (1) and (2), $\nu = [\gamma H_0(1 - \alpha)] / 2\pi$

From this relation it is clear that the protons with different electronic environments or with different shielding parameter can be brought into resonance in two ways:

(i) The strength of the external field is kept steady and the radio-frequency is constantly varied.

(ii) The radio-frequency is kept steady and the strength of applied field is constantly varied.

Thus, at constant radio-frequency, shielding shifts the upfield in the molecules where there is spherical distribution of electrons around the proton. It is called **positive shielding**.

Presence of the electromagnetic atoms or groups cause reduction in electron density around a proton (deshielding) and thus, absorption is shifted downfield (Tau decreases). Hence, we say that if local diamagnetic currents were responsible, then the value of proton absorption should depend upon the electro negativities of the attached atoms or groups. But there are some other factors, viz, the diamagnetic and the para magnetic effects from the neighboring atoms and also the effects from the intra-atomic currents which result in the proton absorption shift.

It has been found that the absorption depends upon the manner in which the π-electrons circulate under the influence of the applied field. In case of ethane oriented at right angles to the direction of the applied field the induced magnetic field (due to the circulation of π-electrons) is diamagnetic (opposes external field) around the carbon atoms and is paramagnetic (in the direction of the applied field) in the region of protons. Thus, protons feel more magnetic field (deshielding) and hence absorption occurs at low field. Similarly, aldehydic protons appear at low field.

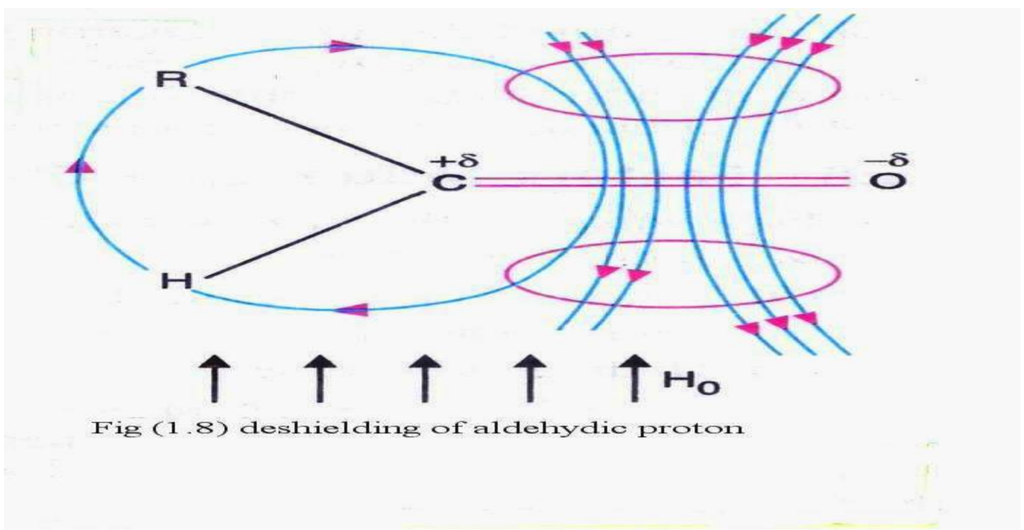


Fig (1.8) deshielding of aldehydic proton

In the case of acetylene, the protons experience a magnetic shielding effect. The induced magnetic field at the protons is diamagnetic but paramagnetic at the carbon atoms. Thus, protons are said to be shielded (feel smaller field) and hence absorption occurs upfield (Tau value increases).

In case of olefins, acetylenes, aldehydes, ketones, acids, esters, nitriles, etc., if the proton is present in the positive region, it will be shielded and absorption occurs upfield. On the other hand if the proton lies in the negative region, its absorption lies downfield. In case of benzene and other aromatic compounds, strong diamagnetic currents are induced (in the loops of the π-electrons) by the field applied. This causes paramagnetic shielding in the aromatic protons and thus, absorption occurs downfield. On the other hand, any group which is present above or below the plane of the aromatic nucleus will be shielded and absorption for such a proton occurs upfield. Ring current effects as is seen in benzene and other aromatic compounds is absent in saturated hydrocarbons.

In saturated cyclic hydrocarbons, diamagnetic currents are induced as a result of the circulating bond electrons. Thus, effect is weaker relative to that in aromatic nuclei (mobility of electrons). Although weak, yet it is responsible to distinguish between the axial and the equatorial protons in cyclohexane. Axial protons are comparatively shielded and such protons absorb 0.5 ppm upfield as compared to those equatorial protons.

The deshielding effect due to a particular group decreases as its distance from the absorbing proton increases.

Due to the decreasing deshielding influence, the Tau values of absorption decrease in the following order:

Compound A	Compound B
CH_2-CH_3 (benzene ring)	CH_3-CH_2-CH_2-Cl
a > b > c	c > b > a

For saturated molecules, the diamagnetic contribution is most important and the absorption positions of the protons depend upon the electronegativities of the attached atoms or groups. The deshielding effect due to electronegative atom falls with distance and there is a little effect on the position of absorption of protons which are three bonds away. In unsaturated compounds, paramagnetic effects become important. In addition to their distance, the effect depends upon the relative orientation of the protons with respect to the groups present in the molecule.

1.6 PEAK AREA AND PROTON COUNTING

In an NMR spectrum, various peaks represent equivalent sets of protons. The size or the area of each peak tells the number of protons in each set present in the compound under study. The area under an NMR signal is directly proportional to the number of protons giving rise to signal. For flipping over a proton, a quantum of energy is absorbed in the same effective magnetic field. Greater the number of protons that flip over at a particular frequency, greater will be the energy absorbed and greater is the area under the absorption peak. Squares under each peak are simply counted and from this, the ratio between various kinds of protons is found out. These ratios are then converted into whole numbers. These whole numbers (or multiple of them) tell the number of protons represented by the various NMR signal.

In the spectrum of Toluene, it shows two types of protons as is clear from the two signals. If the number of squares under each signal is counted, it will be found that the areas under the peaks have the ration 5:3. Thus, in toluene, 5 protons are of one kind and 3 protons are of another kind. Hence, we say that the NMR spectrum of toluene represents 2 kinds of protons which are in the ratio 5 : 3.

(i) 5 proton signals (downfield due to deshielding) and
(ii) 3 proton signals (up field)

Note: The students are advised to draw small squares in the NMR spectrum of a compound like the squares in a graph paper. The number of squares under each signal is carefully counted and then the said ratio is found out.

Integration. It is a process by which relative areas under spectral peaks are measured. These areas are proportional to the number of equivalent protons giving the signal. The NMR instrument draws an horizontal line in case there is no signal. On receiving a signal, the line ascends and levels off as the signal ends. The relative distance from plateau gives the relative area. Hence, the integration process tells the relative number of protons. Consider the PMR spectrum of Benzyl acetate :

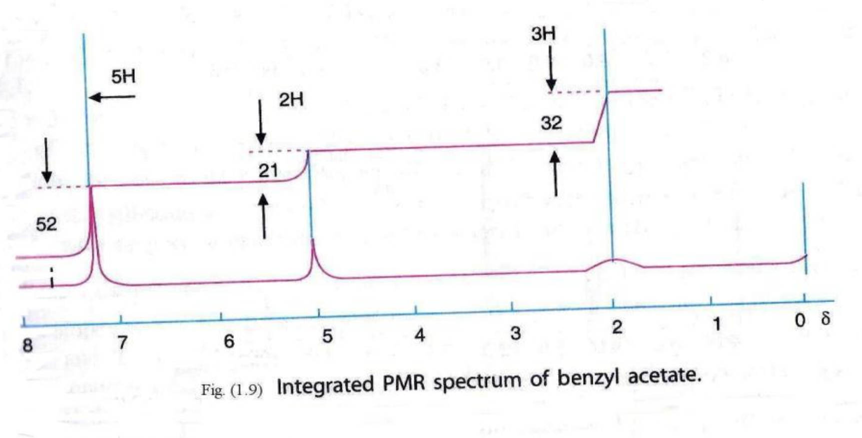


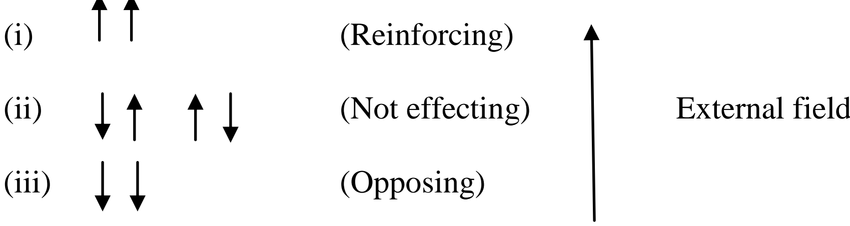
Fig. (1.9) Integrated PMR spectrum of benzyl acetate.

$$CH_3\text{-}\overset{\overset{\displaystyle O}{||}}{C}\text{-}O\text{-}CH_2C_6H_5$$

It is found that the three types of protons give three signals in the integral ratio 52:21:32 or 5:2:3.

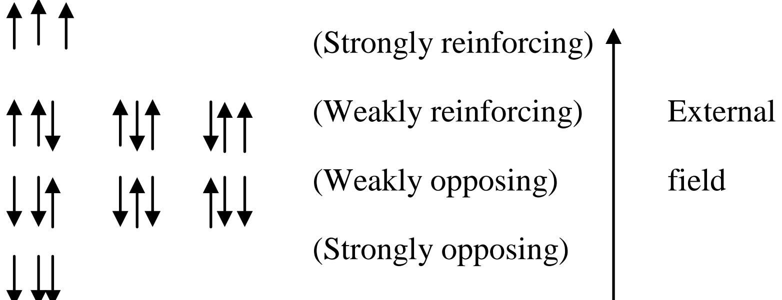
1.7 SPIN – SPIN COUPLING

To understand it properly, considering a molecule of ethyl bromide (CH_3CH_2Br), the spin of two protons (- CH_2 -) can couple with the adjacent Methyl group (CH_3 -) in three different ways relative to the external field. The three different ways of alignment are:



Thus, a triplet of peaks results with the intensity ratio of 1:2:1 which corresponds to the distribution ratio of alignment.

Similarly, the spin of three protons (CH_3 -) can couple with adjacent methylene group (- CH_2 -) in 4 different ways relative to the external field.



Thus, quartet of peaks result with an intensity ration of 1:3:3:1 which corresponds to the distribution ratio of the all the alignments. The relative intensities of the individual lines of a multiplet correspond to the numerical coefficient of the lines in the binomial expression:

$$(1 + X)^n = 1 + X \text{ if } n = 1$$

If n=2, then $(1+ X)^2 = 1 + 2X + X^2$. Thus, the lines of the triplet have relative intensities 1: 2: 1.

If n=3, then $(1 + X) = 1 + 3X^2 + 3X + X^3$. Thus the line of the quatet formed due to the influence of three equivalent protons will have relative intensities 1: 3 : 3: 1.

Similarly, the lines of the pentad (quintet) formed will have relative intensities 1 : 4 : 6 : 4 : 1.

The splitting of protons and the peak intensities can be readily obtained by the **Pascal's triangle** as:

No of protons causing splitting	Splitting pattern (relative intensities)	Name
0	1	Singlet
1	1 : 1	Doublet
2	1 : 2 : 1	Triplet
3	1 : 3 : 3 : 1	Quartet
4	1 : 4 : 6 : 4 : 1	Quintet
5	1 : 5 : 10 : 10 : 5 : 1	Hextet

Hence, the splitting of a signal is due to different environment of the absorbing proton not with respect to electrons but with respect to the nearby protons (protons attached to the adjacent carbon atom). Let us consider the case of 1 : 1: 2 trichloro-ethane

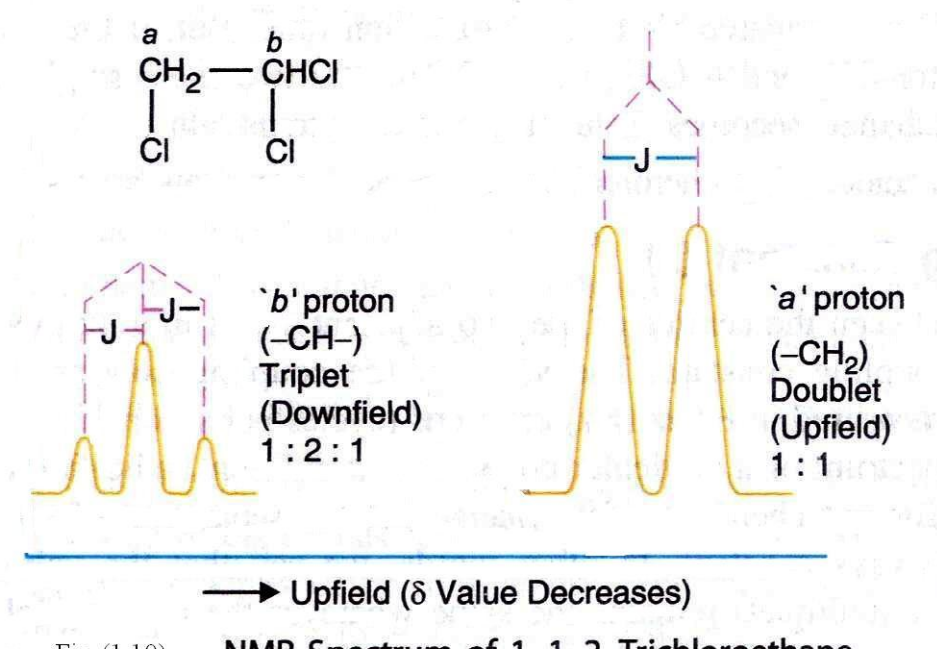


Fig. (1.10) NMR Spectrum of 1, 1, 2–Trichloroethane.

1.8 SPLITTING OF THE SIGNALS

It is already pointed out that each signal an NMR spectrum represents one kind or one set of identical protons in a molecule. It is found that in certain molecules a single peak (singlet) is not observed but instead a multiplet (group of peaks) is observed. Consider the molecule of CH_3CH_2Br, ethyl bromide. This molecule has two kinds of protons in it and thus, two signals are expected in its NMR spectrum. It has been observed that for each kind of protons, we do not get singlets but a group of peaks are observed. For 'CH_3' protons , a triplet i.e. a group of three peaks is observed and a quartet (group of four peaks) is noticed for protons -CH_2.

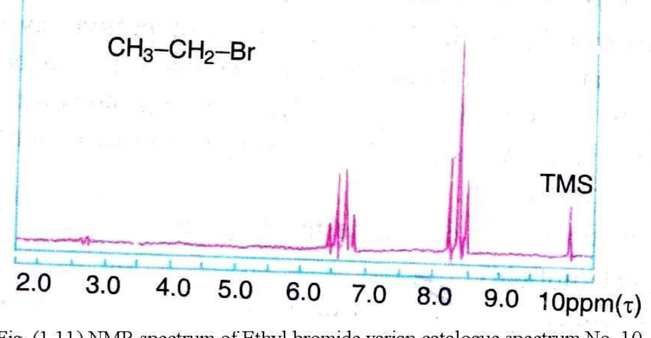


Fig. (1.11) NMR spectrum of Ethyl bromide.varian catalogue spectrum No. 10

Fig. 5.21. NMR spectrum of Ethyl bromide. Varian catalogue spectrum No. 10.
Signals and their absorption positions

(i) A three proton triplet 8.35 τ

(ii) A two proton quartet 6.6 τ

Similarly, in the NMR spectrum of 1 : 1 : 2 trichloroethane (Cl_2CH-CH_2Cl), two signals are observed. For –CH_2 - kind of protons, a doublet (a group of two lines) is observed while a triplet (a group of three lines) is notice for –CH- kind of protons.

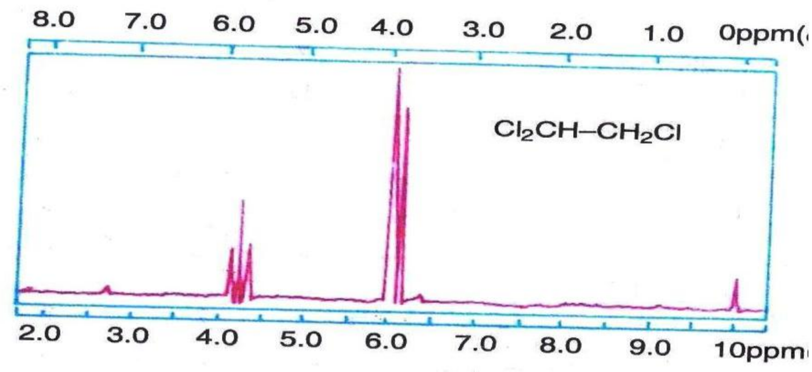


Fig. (1.12) NMR spectrum of 1:1:2 trichloro ethane. Varian catalogue spectrum No.2

NMR spectrum of 1 : 1: 2 trichloro ethane. Varian catalogue spectrum No. 2

(i) A three proton doublet 6.05 τ

(ii) A two proton triplet 4.2 τ

1.9 COUPLING CONSTANT (J)

The distance between the centres of the two adjacent peaks in a multiplet is usually constant and is called the coupling constant. The value of the coupling constant is independent of the external field. It is measured in Hertz(Hz) or in cps (cycles per second).it is denoted by the letter J. If we work the spectrum of a particular compound at different radio-frequencies, the separation of signals due to deferent chemical shifts change but the separation of two adjacent peaks in a multiplet remains always constant. In other words we say that the value of J(separation of adjacent peaks in the multiplet) remains the same whatever may be the applied field. From the value of coupling constant, one can distinguish between the two singlets and one doublet and also a quartet from two doublet. It can be done by simply recording the spectrum at two different radio-frequencies. If the separation (in Hz) between the lines (value of J) does not change, then the signal is a doublet. On the other hand , if the separation between the lines increases with increasing frequency, then the signal in fact, will be two singlets. The value of 'J' generally lies between 0 and 20Hz. Same explanation can be given to distinguish a quartet from the doublets.

Now let us consider a compound, >CH-CH_2-. In this compound two signals are expected in the NMR spectrum. Under the influence of two equivalent protons 'a', the signal for proton 'b' will appear as a triplet. The distance between any

two adjacent peaks in a multiplet will be exactly the same. The triplet formed due to spin-spin coupling is shown as:

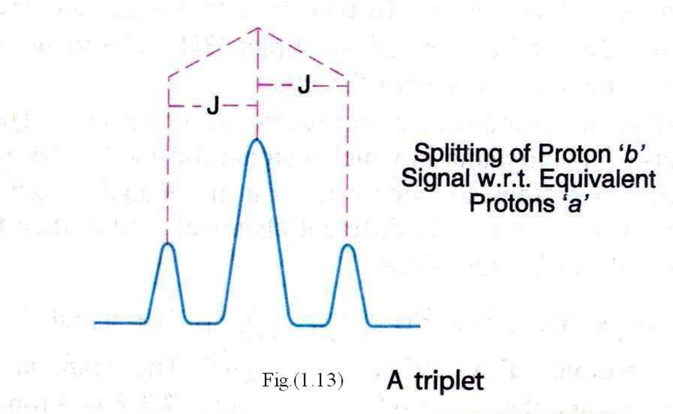


Fig.(1.13) A triplet

Similarly, consider a compound, CH_3-CH_3. In the NMR spectrum of this compound, proton 'b' is under the influence of three equivalent protons 'a'. thus, due to spin-spin coupling, the signal for proton 'b' will appear as a quartet with intensity ratio 1 : 3 : 3: 1. The quartet for 'b' can be represented as

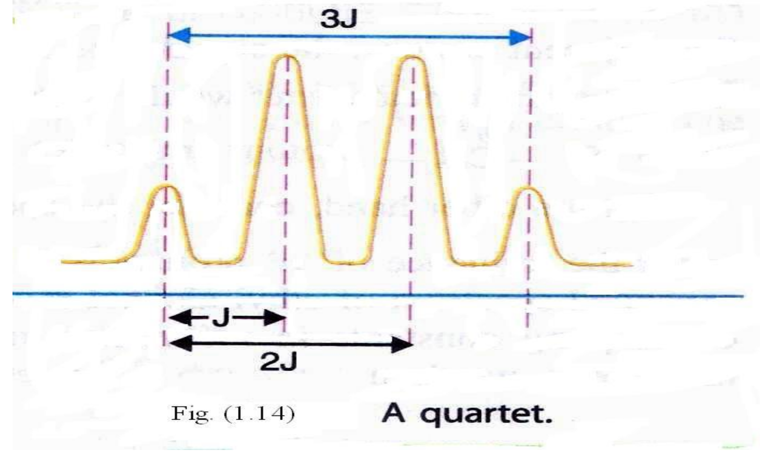


Fig. (1.14) A quartet.

.

From the constant value of 'J', we say that the multiplets formed are of symmetrical nature. It has been noted that the departure from the symmetrical nature of the multiplet result when the absorption position (chemical shifts) of the interacting groups are close enough.

In the NMR spectrum of 1, 2, 2- trichloroethane (see Fig. 5.30) two multiplets are observed. The value J in each of its multiplets is found to be constant.
Following are some of the compounds which have constant value of J in each of their multiplets.

(a) In the spectrum of $\begin{array}{l} CH_2\text{—}O \\ | \quad\quad | \\ CH_2\text{—}CH_2 \end{array}$: a low field triplet (4H) and an upfield quintet (2H) are observed.

(b) The value of H (peak separation) in the triplet as well as in the quintet is found to be the same, i.e. 7.1 cps.

(c) In the spectrum of Isopropyl cyanide; a doublet (6H) and a septet (1H) are formed. The value of 'J' in each of these multiplets formed is the same, i.e.6.7cps.

(d) Consider the spectrum Butanone-2. In this case, two multiplets are formed.

Similarly, a large number of examples can be quoted in which the value of 'J' remains the same for any two multiplets of a compound in which a signal for one kind of protons is split under the influence of another kind of protons and vice versa. But if a signal is split due to influence of two different sets of protons (protons with different chemical shifts), then the peak separations in the multiplet observed will not be equidistant.

Consider an example of propyl iodide CH_3- CH_2-CH_2I. The signal for 'b' kind of protons are under the influence of protons 'a' as well as protons 'c'. the examination of the expanded spectrum of propyl iodide reveals, that J_{ab} =6.8 cps and J_{bc}=7.3 cps. Form these values of the coupling constants, the multiplet pattern of 'b' protons can be found out. Under the influence of five adjacent protons, the signal for –CH_2- protons will be split into a sextet. If Jab were equal to J_{bc}, a symmetrical sextet would have been observed. Since the values of H are close enough, the signal for –CH_2- protons (multiple) will resemble a sextet.

On the other hand, consider the spectrum of pure ethyl alcohol CH_3CH_2OH. 'b' protons are under the influence of 'a' protons (OH protons) and the 'c' protons. The examination of the expanded spectrum of ethyl alcohol reveals that J_{ab} =5.0 cps and J_{bc} =7.2cps. since the values of coupling constants in methylene multiplet are much different (assuming the symmetry of multiplets), we do not observe a quintet normally expected for 'b' protons in pure ethyl alcohol.

Actually, a group of eight lines are observed in its multiplet. Quintet would have been formed if J_{ab} were equal to J_{bc}.

It may be clearly noted that the value of coupling constant depends partly on the number of covalent bonds through which protons may interact and also upon the structural relationships between the coupled protons.

1.10 IMPORTANT TIPS FOR INTERPRETING AN NMR SPECTRUM

a. Following points regarding the value of chemical shift may be useful:

(i) Tau value of methyl, methylene and methyne protons have the order:

Mehyl > Methylene > Methyne

(ii) Tue value depends upon the nature of the substituent on the carbon atom bearing the proton. Greater the electronegativity of the substiuent, lower is the value of tau.

(iii) The value of tau depends upon the type hybrid orbital holding the proton.

$$sp^3 > sp > sp^2$$

(iv) The tau value for aromatic protons is always less than 4 ppm. The value depends upon the degree and the nature of substitution.

(v) Tau value for the aldehydic protons are generally lower, i.e., 0.8 ppm or lower.

(vi) Tau value of protons in a cyclic compound is always higher than that of any other protons.

The set of protons in cyclopropane has the maximum tau value.

(vii) The chemical shift of the protons in O-H group and also in $–NH_2$ group depend upon temperature, solvent, concentration and the neighbouring group. For example, the alcoholic (-OH) proton gives a singlet at 4.5-9.0 τ, the phenolic OH absorbs at -2 to 6 τ. if C=O group is present in the ortho positions in phenol, then due to interamoleclular hydrogen bonding, absorption occurs even at negative value of tau. The OH proton in the enolic form absorbs at -4 to -5 τ.

(viii) The absorption due to –COOH group appears at -0.5 to -2.0 Tau (10.5 – 12 Delta)

b. The number of signals in an NMR spectrum tells the number of sets of the protons in different chemical environments.

c. It also tells the number of equivalent protons causing the splitting of a signal.

Lets us consider a few cases:

(i) In a spectrum, two signals, say, singlets are observed. This spectrum can be interpreted by saying that there are two kinds of protons and the two carbon atoms carrying the two sets of protons are not the adjacent carbon atoms otherwise splitting of signals would have taken place.

(ii) The spectrum consisting of one doublet and one triplet must be due to the molecule having $–CH_2$-CH- as its part.

(iii) Formation of two triplets must be due to two methylene groups whose protons are in different environments.

$$X\text{-}CH_2\text{-}CH_2\text{-}Y$$

(iv) The formation of a septet and a doublet must be due to $(CH_3)_2CH$- group which form a part of the molecule and so on. From the intensities of the signals (peak areas) the ratio between different kinds of protons can be found.

1.11 SUMMARY:

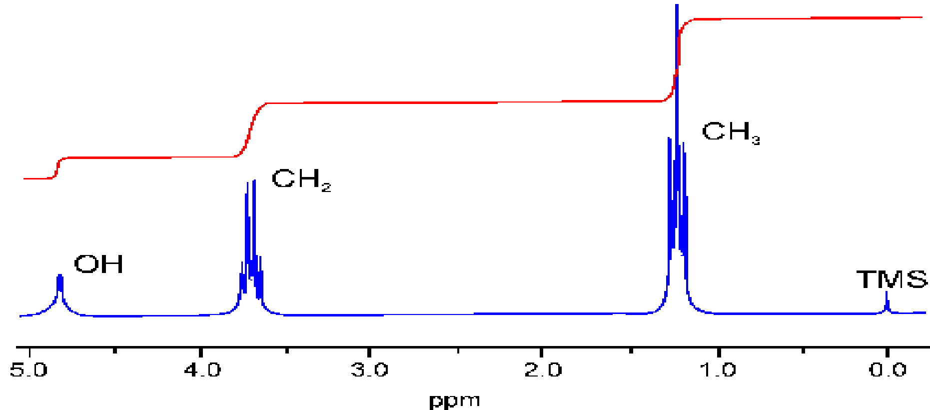
1. A nucleus with odd atomic number or an odd mass number has a nuclear spin which can be observed by the NMR spectroscopy. A wide range of nuclei including 1H, 13C, 15N, 19F and 31P have a suitable magnetic properties and can be studied by NMR spectroscopy.
2. In presence of an external magnetic field H_o, a proton has a spin quantum number I of +1/2 (alpha-spin state aligned with the external magnetic field) or-1/2 (Beta-spin state aligned against external field).
3. The Beta spin state is higher in energy than alpha spin state.
4. In the absence of magnetic field there are random orientations of the proton magnetic moments i, e., spin in all directions with no energy difference.
5. When a photon with just right amount of electromagnetic energy strikes the proton, the protons spin can flip from alpha to Beta and vice versa. The strength of earth's magnetic field is about 0.57 gauss.
6. The energy differenced between the two spin state (alpha and Beta) is not very large, and is characteristic of a particular type of a nucleus and the strength of magnetic field experienced by that nucleus. For a proton in a field if 14000.00 gauss (typical of normally available NMR spectrometers) the value is 5.7 x 10^{-6} Kcal/mole.
7. Relaxation is the process by which nuclei in the Beta spin state can return to alpha spin state. Absence of relaxation leads to saturation, where there is no net excess of nuclei in the ground state and one will have exactly equal populations in spin states and net absorption ceases.
8. The usual procedure for determining PMR spectra involves dissolving the sample in a suitable solvent containing no protons (e. g. CCl_4)etc.
9. A PMR spectrum is usually recorded by sweeping the magnetic field whilst irradiating the sample with radiation of constant frequency. (absorption-n) of signal occurs when the field reaches the value hν/2μ.
10. Solid samples give broad diffuse peak while solutions and liquids give sharp narrow peaks.
11. Chemical shift (δ) ppm is the position of an absorption peak relative to that of a reference compound (usually TMS).
12. Chemical shift is largely depended on the presence (shielding) or absence (deshielding) of electron density.
13. Multiple bonded systems also give rise to induced magnetic fields. These induced field lines curve around and consequently divide the immediate environment into two regions. In one region the proton is shielded and in the other its de-shielded from the applied field. In other words in one region the induced field acts to oppose the external field while in the other the induced field adds to external field.
14. Induced magnetic field of circulating aromatic electrons is larger than in alkenes due to a large effective ring of electrons. eg. in benzene. As a

consequence the benezenoid protons and aromatic protons in general have greater δ value (*6-8.5) ppm than alkenic proton. This forms the basic of the use of PMR spectroscopy to detect aromaticity.

15. Spin splitting i.e. magnetic coupling is the interaction of the magnetic field of two or more nuclei, both through their connecting bonds and space. Spin-Spin splitting causes PMR signals to be split and to appear as two or more peaks as a multiplet.
16. The separation of components in a multiplet is called the coupling constant and given by symbol J> (values measured in Hz).
17. A signal that is being split by n equivalent protons appear as a multiplet with n+1 individual peaks known as n+1 rule. The chemical shift equivalent hydrogens show no spin-spin splitting.
18. Diastereotopic protons can be distinguished by PMR spectroscopy since these are not equivalent. Using the replacement test. The replacement of diastereotopic proton gives diasteromers. The diasterotopic protons can split each other unless there are accidentally equivalent.
19. Enantiotropic protons (i.e. protons which by using replacement test giver enantiomers) cannot be distinguished by PMR spectroscopy, this being an achiral technique. The enantiotropic protons can be detected by only using chiral probe.
20. The non equivalent protons are turned accidentally equivalent when these absorb at the same chemical shift. Protons with same chemical shift do not split each other.
21. Branching diagram is a useful method to determine the spin-spin coupling patterns.
22. Complexity in coupling patterns arise when ($\delta\Delta \leq J$) complicated spectra can be improved by using.
    a. Higher field strengths
    b. Spin decoupling by double resonance
    c. Using paramagnetic shift reagents
    d. Substitution by deuterium
23. Most of protons in organic compound experience a rapidly changing magnetic environment due to intra molecular motions. Compared toPMR absorption process intramolecular motions are rapid and therefore the pmr spectrum records the position of each proton.

1.12 INTERPRETAION OF SPECTRA OF SOME COMMON COMPOUNDS

SAQ -1: Spectrum of Ethanol



The three main features of NR spectroscopy ,i.e., chemical shift, signal intensities and spin- spin coupling can be studied in this spectrum.

Three – proton atmosphere, all three protons lying in different magnetic atmosphere, due to motion of valence electrons and neighbouring atoms in response to magnetic field.

(i) CH_3 – (three protons)
(ii) – CH_2 – (two protons)
(iii) - OH (one proton)

The effects arising due to the motions of the electrons which will be different for each kind of hydrogen will be at different field strengths.

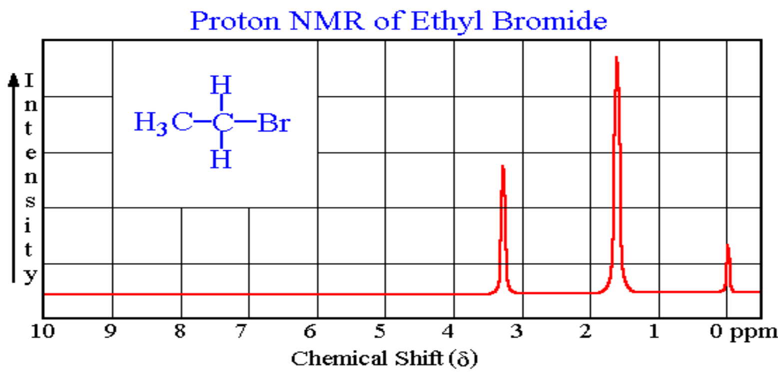
A plot of signals' against field strengths thus shows three principle groups of lines for ethyl alcohol.

The applied signal is a three proton triplet of methyl group. The chemical shift of this signal is (1.22) for all the three identical methyl group protons have two neighbouring protons i.e. (n = 2) and according to (n+1) rule one sees a three proton triplet, since three identical hydrogen of methyl group couple equally with the protons of the methylene group. Similarly, the two identical protons of methylene group are adjacent to the hydroxyl group and as a consequence come into resonance at lower field than the methyl group. The signal is a quartet since the methylene protons are equally coupled to three hydrogens of CH_3.

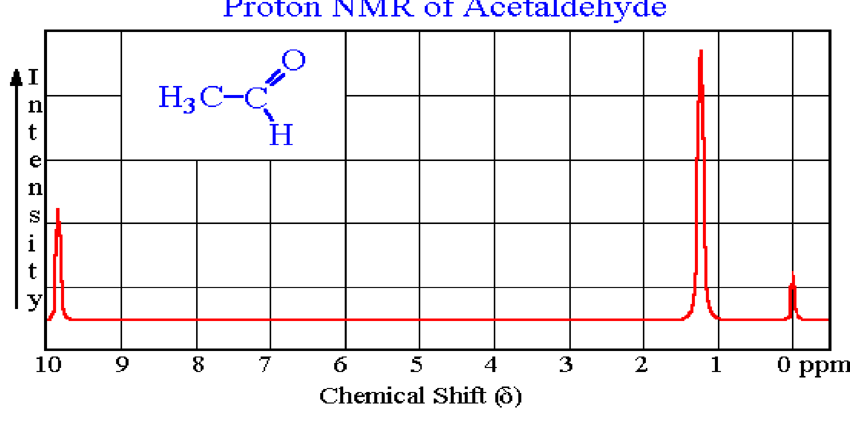
The hydroxyl proton appears as a singlet, i.e, no interaction is seen between the hydroxyl proton and neighbouring methylene group. It may be sufficient to remember that the spin-spin splitting is not observed normally for a proton of a

hydroxyl or amino group. Such protons are relatively acidic and a s a result these rapidly interchange between molecules.

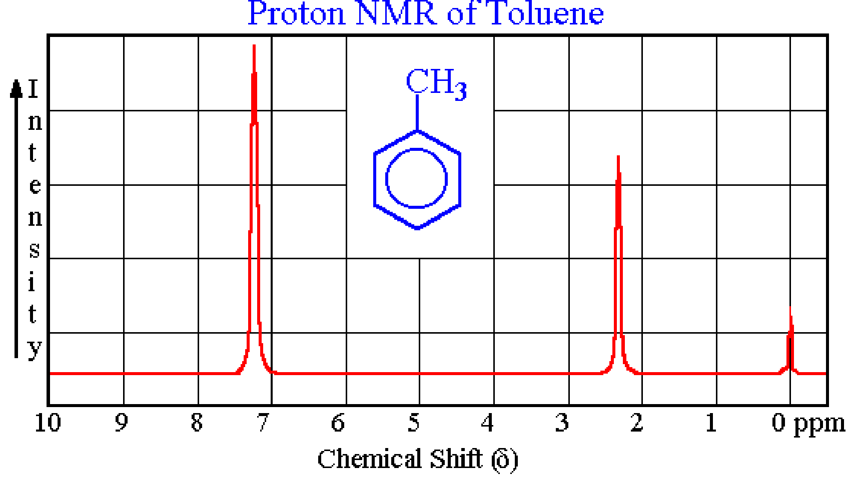
SAQ – 2: Spectrum of Ethyl Bromide



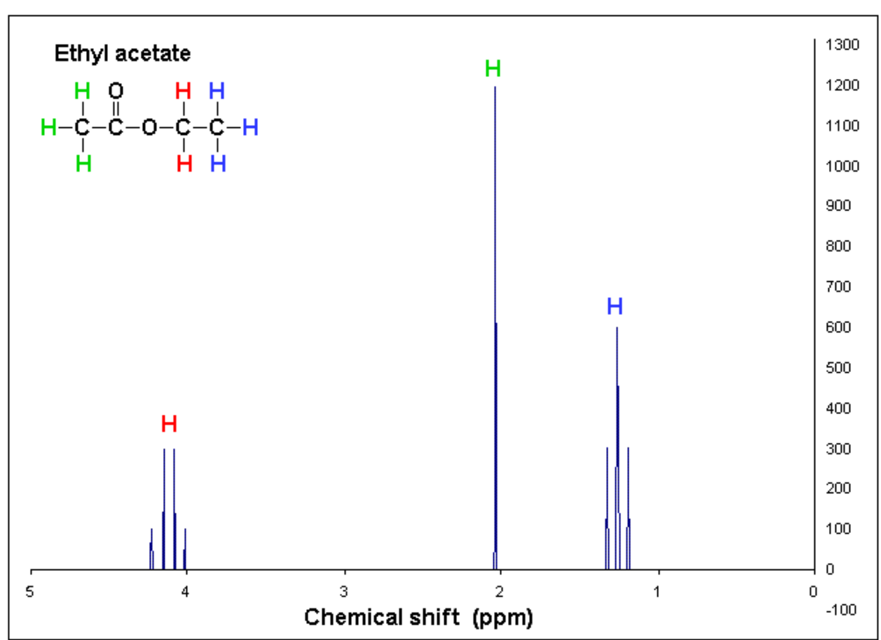
SAQ – 3: Spectrum of Acetaldehyde



SAQ – 4: Spectrum of Toluene



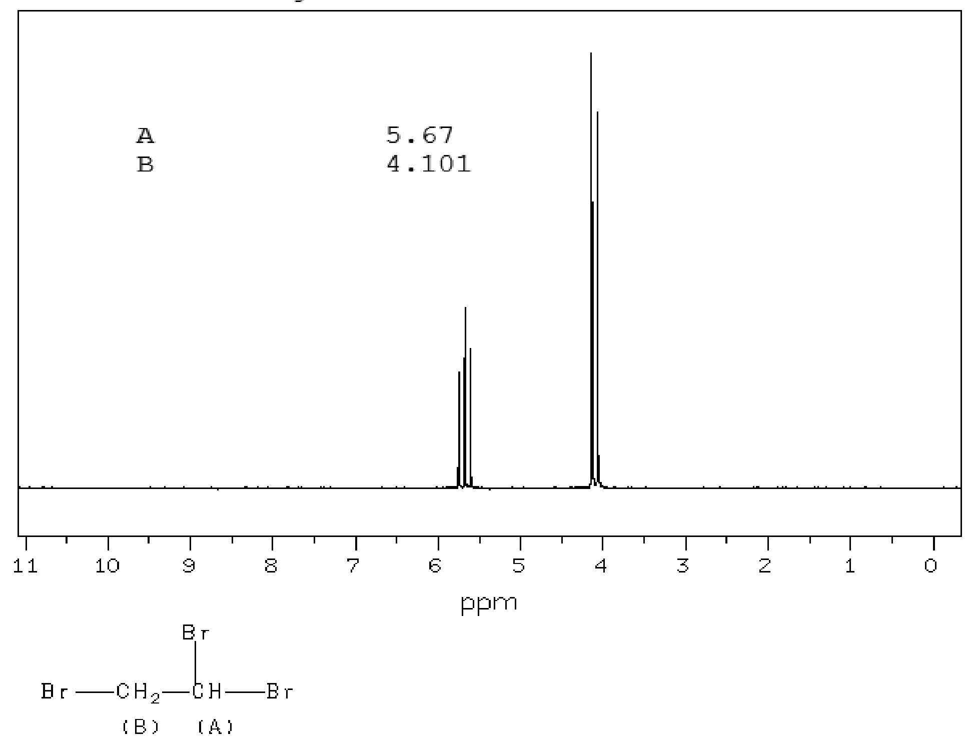
SAQ - 5: Spectrum of Ethyl Acetate



SAQ - 6: Spectrum of 1,1,2-tri bromo ethane

89.56 MHz

0.04 ml : 0.5 ml $CDCl_3$



1.13 TERMINAL QUESTIONS

PROBLEM-1. A compound with molecular mass 158 absorbs in the ultraviolet region at 225 nm ε_{max} (hexane). In infra-red spectrum, absorption bands are formed at 3077-2857 cm^{-1}(m), 1828 cm^{-1}(s), 1757 cm-1 (m) and 1457 cm^{-1} (m). In NMR, two signals are observed (i) 7.30τ septet (J=6.7 cps, 6.4squares) and (ii) 8.80 τ doublet (J = 6.7 cps, 37.2 squares).

PROBLEM 2. A compound molecular mass 164 absorbs at 220 nm ε_{max} 1800. In infra-redspectrum, absorption bands are formed at 3077 cm^{-1} (w), 2976 cm^{-1} (s) 1608 cm^{-1} (m), 1497 cm^{-1} (m) and 1456 cm^{-1}(m).
In NMR, the signals formed are (i) 2.7 τ singlet (16.5 squares), (ii) 5.70 τ triplet (J = 7.3 cps, 6.2 squares), (iii) 7.07 triplet (J = 7.3 cps, 6.7 squares) and (iv) 7.98 τ singlet (10.2 squares).

PROBLEM 3. An organic compound with molecular mass 174 shows absorption in ultraviolet region at 213nm ε_{max} 60. In infra-red, absorption bands are formed at 2941- 2857 cm^{-1} (m), 1745 cm^{-1} (s) and 1458 cm^{-1} (m).
In NMR, the signals observed are (i) 5.86 τ quartet (J = 7.2cps, 10.4 squares), (ii) 7.40 τ singlet (10.8 squares) and (iii) 8.73 τ triplet (J = 7.2 cps, 16.0 squares).

PROBLEM 4. A organic compound with molecular formula $C_{14}H_{19}N$ give the following spectral data:
UV : (i) 222 nm ε_{max} 20,400
(ii) 210 nm ε_{max} 20,000
IR : 3022 (m), 1600 (m), 1510 (m), 1680 (w), 750 and 695 cm^{-1}(m)
NMR : (i) singlet 7.15 τ (3H), (ii) singlet 8.5 τ (3H), (iii) complicated 8.0 τ (4H), (iv) complicated 8.35 τ (4H), multiplet; 3.45 τ (3H) and multiplet 2.95 τ (2H)
Determine the structural formula of the compound.

PROBLEM 5: An organic compound with molecular formula C_4H_9NO gives the following spectral data:

UV : λ_{max} 220 mμ ε_{max} 63
IR : 3500 (m), 3402 (m), 2960 (w), 1682 (s), 1610 (s)
NMR : 9.0 τ doublet (23.2 squares); 7.9 τ septet (3.8 squares); 1.92 τ singlet (7.5 squares).
Determine the structural formula of the compound.

PROBLEM 6. An organic compound contains 66.6% Carbon, 11.1% Hydrogen. In UV, it gave a characteristic band at 275 mμ ε_{max} 17. In infra-red, bands are formed at 2941-2857 (m), 1715(s) and 1460 cm^{-1} (m). In NMR, three

signals appear at (i) 7.52 τ quartet, (2H), 7.88 singlet, (3H) and 8.93 τ triplet, (3H). Determine the structure formula of the compound.

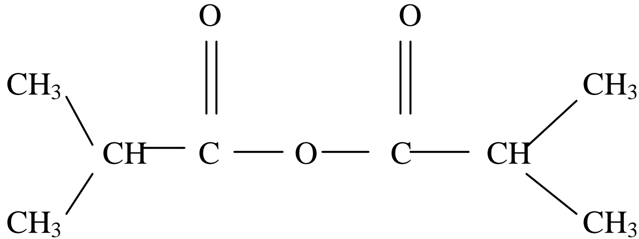
1.14 ANSWERS TO Terminal Questions

SOLUTION-1. In the NMR spectrum, a septet and a doublet with the same value of coupling constant shows $(CH_3)_2CH$- as a part of the structure.

In the infra-red spectrum, a strong band at 1828 cm^{-1} and a medium band at 1757 cm^{-1} shows

The presence of anhydride— $\overset{\overset{O}{||}}{C}$ — O — $\overset{\overset{O}{||}}{C}$.

Thus, $(CH_3)_2CH$- and –CO-O-CO- amounts to 115 mass units. The remaining 43 mass unit may also be the isopropyl group. Hence, the most probable structure of the compound is:



A band at 3077-2857 cm^{-1} is due to C-H str and that at 1456 cm^{-1} is due to C-C str. An absorption band at 225 nm ε_{max} 50 does not give any authentic information.

SOLUTION - 2. In NMR, the signals are formed in the proton ratio 2.5:1:1:1.5 or 5:2:2:3. The formation of two triplets with the same value of coupling constant shows that $-CH_2$-CH_2- is a part of the structure.

A five proton singlet at 2.71 τ shows the presence of an aromatic ring C_6H_5- .

A three proton upfield singlet shows $-CH_3$ group in the compound. Thus, the units C_6H_5 - , -CH_2-CH_2- and CH_3- amount to 120 mass unit. A strong band at 1745 cm^{-1} in the infra-red spectrum suggests that the compound is an ester. Thus, -COO (of ester) amounts to the remaining 44 mass unit of the compound of molecular mass 164. The bands at 1608 cm^{-1}(m) and 1497 cm^{-1} (m) support the presence of benzene ring. The band at 1456 cm^{-1} (m) is due to C-C str. Thus, the probable structure of the compound may be

O
||
CH_2—CH_2— C — OCH_3

Or

O
||
CH_2—CH2— C — O — CH_3

The structure II looks more probable because three proton singlets are relatively less deshielded.

SOLUTION - 3. In the NMR, the signals are formed in the proton ratio 2:2:3. The formation of a quartet and a triplet with the same value of coupling constant suggests that CH_3CH_2- is a part of the structure. Also a two proton singlet is formed. It may be due to –CH_2- group in the structure.

In infra-red spectrum, the formation of a strong band at 1745 cm^{-1} shows the presence of an ester. Thus, -COO is a part of the structure. Thus, CH_3CH_2-, -CH_2- and –COO units amount to 87 mass unit. The molecular mass 174 is just double of it. Thus, the probable structure of the compound is:

O
||
CH_2—C — O — CH_2—CH_3
|
CH_2—C — O — CH_2—CH_3
||
O

A low field triplet indicates that its carbon is directly linked with oxygen atom.

An absorption band at 2941-2857 cm^{-1} indicates C-H str in infra-red and that at 1458 cm^{-1} indicates C-C str.

SOLUTION - 4. From the absorption in the ultraviolet spectrum, nothing can be said with certainty.

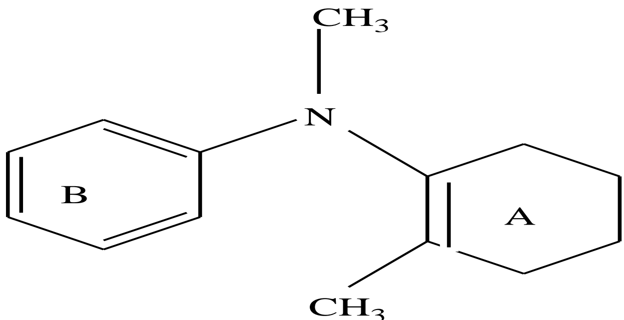
The infrared absorption at 3022 cm^{-1} includes =C-H stretching from an aromatic ring. The appearance of bands at 1600 (m), 1510 cm^{-1} (m) indicates the presence of an aromatic ring. Further the bands at 750 and 695 cm^{-1}(m) shows that the aromatic ring must be substituted. The weak band at 1680 cm^{-1} is characteristic of C=C stretching. The presence of the aromatic is further confirmed by multiplets in the region 2.95 – 3.35 τ. The two proton multiplet at

2.95 τ must be due to two ortho positions with respect to the N-substituted group on the ring. The remaining three protons show another slightly less deshielded multiplet at 3.45 τ.

The two signals each of three protons singlet at 7.15 τ and at 8.5 τ indicate two methyl groups. The singlet at 7.15 τ is slightly deshielded being attached with nitrogen which in turn is attached with the aromatic ring.

The remaining two signals (complicated) corresponding to 8 protons in almost similar environments indicate 4 methylene groups in ring A. the four proton signals at 8.0 τ corresponds to 2 methylene groups which are attached with the carbon atoms carrying a a double bond.

From this, the structural formula of the compound is probably.



SOLUTION - 5. From the absorption at 220mμ ε_{max} 60 in the ultraviolet region, nothing can be said with certainty. The presence of a primary amide is indicated by two bands at 3500 and 3402 cm^{-1}. These two bands are the result of N-H stretching. A strong band at 1682 cm^{-1} also indicates the presence of an amide group. The band at 2960 cm^{-1} is due to C-H stretching and that at 1610 cm^{-1} (s) is due to N-H bending. From this much data, we say that $–CONH_2$ is a part of the structure.

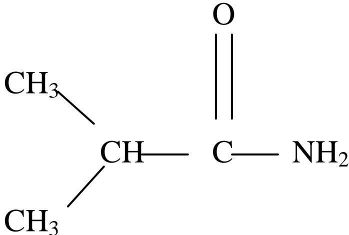
The NMR spectrum is most informative. The doublet at 9.0 τ and a septet at 7.9 τ in the ratio 6 : 1 indicates clearly an isopropyl group i.e., $(CH_3)_2$ CH- group.

Another two proton singlet appears at a very low field due to much deshielding. This signal is due to two protons of the primary amide group. The appearance of the signals at a very low field may be explained by writing the resonating structure of amide as:

$$\underset{\text{I}}{—\overset{\overset{\displaystyle O}{||}}{C}—NH_2} \longleftrightarrow \underset{\text{II}}{—\overset{\overset{\displaystyle O^-}{|}}{C}=NH_2}$$

The double bond character between C and N and also the positive charge developed on nitrogen atom in structure II case greater deshielding of protons attached with nitrogen atom and hence, signal appears at 1.92 τ.

Thus, the structure of the compound is



From the given molecular formula (C_4H_9NO), the number of ring and double bond equivalents are calculated as:

$$4+1-\frac{9-1}{2} = 5 - 4 = 1$$

This is in accordance with the structure which is determined.

SOLUTION - 6 The compound contains -

C = 66.6%
H = 11.1%
O = 100-(66.6+11.1) = 22.3%

From this data, the empirical formula of the compound is found to be C_4H_8O. This must be the molecular formula since eight hydrogen atoms are shown by NMR spectrum.

(i) The absorption at 275 mμ ε_{max} 17 is characteristic of a carbonyl group due to $n \longrightarrow \pi^*$ transition (forbidden band).

(ii) The absorption at 2941-2857cm^{-1} (m) in the IR spectrum is due to C-H stretching, at 1715cm^{-1} (s) is characteristic of saturated ketonic group and that at 1460 cm^{-1} (m) may be due to C—H bending.

(iii) The NMR spectrum reveals three kinds of protons.

The presence of a triplet at 8.93 τ and a quartet at 7.52 τ is characteristic of $CH_3 – CH_2 –$ group in the compound. The singlet at 7.88 τ is due to methyl group adjacent to a carbonyl group. Hence, the probable structure of the compound is $CH_3 – CH_2 – CO – CH_3$. The structure contains one double bond. Also, the number of DBE from the molecular formula:

$C_4H_8O = 4 + 1 – 8/2 = 1$

UGCHE08

BLOCK - I

UNIT – 1: NMR SPECTROSCOPY

UNIT. 2 ORGANOMETALLIC COMPOUNDS

Structure

2.1 Introduction

2.2 Organo Magnesium Halide [Grignard Reagent]

2.2.1 Structure of Grignard Reagent

2.2.2 Formation / Synthesis of Grignard Reagent

2.2.3 Application / Chemical Reactions of Grignard Reagent

2.3 Organo Zinc Compounds [Dialkyl Zincs]

2.3.1 Formation / Synthesis of Organo Zinc Compounds

2.3.2 Application / Chemical Reactions of Organo Zinc Compounds

2.4 Organo Lithium Compounds [Alkyl Lithiums]

2.4.1 Formation / Synthesis of Organo Lithium Compounds

2.4.2 Application / Chemical Reactions of Organo Lithium Compounds

2.5 Summary

2.6 Terminal Questions

2.7 Answers to Terminal Questions

Objective

After studying this unit you should be able to:

- Know what organometallic compounds are
- Have an idea about the various important Organometallic compounds
- Structure of the most important Organometallic compound (Grignard Reagent)
- The syntheses of the various Organometallic compounds
- The synthetic applications of the various Organometallic compounds

2.1 INTRODUCTION

ORGANO-METALLIC COMPOUNDS

The term organometallic or metallo-organic is used nowadays for a wide range of compounds containing metal to carbon bond (the Carbon atom is directly linked to the Metal atom). Organometallic compounds thus include various hydrocarbon derivatives, carbonyls, certain chelate complexes, carbides and related compounds, but do not include carbonates of the metals, complexes of the metals with organic amines and metallic salt of organic acids.

e.g.

$CH_3 - Li$; $C_2H_5 - Zn - C_2H_5$; Mg – Cl ; R_4Si; $C_6H_5 - Hg - Cl$; etc.

The Carbon – Metal bond have a high degree of ionic character. The greater the ionic character the more is the reactivity of metal alkyl.

2.2 ORGANO – MAGNESIUM HALIDE -[Grignard Reagents]

Of all the organometallic compounds known organo-magnesium halides (R-Mg-X) are the most important and were discovered by the French chemist Victor Grignard in 1900. These compounds are now called Grignard reagents and have vast and varied synthetic uses.

2.2.1 Structure of Grignard reagent:

The dimeric structure is the most important one and is in equilibrium with dialkyl / magnesium compound and magnesium dihalide

X
X
RMg MgR ⟷ 2RMgX ⟷ R_2Mg + MgX_2 ⟷ RMg
MgR
X
X

2.2.2 Preparation of Grignard's Reagents (Synthesis):

A Grignard reagent is generally prepared by direct reaction between Magnesium (metal) and alkyl/ aryl halide (RX or ArX) in anhydrous ether.

$$R - X + Mg \xrightarrow{(C_2H_5)_2O} R - Mg - X$$

$$Ar - X + Mg \xrightarrow{(C_2H_5)_2O} R - Mg - X$$

Examples:

$$(H_3C)_2CH - Br + Mg \xrightarrow{Ether} (H_3C)_2CHMgBr$$

Isopropyl magnesium bromide

$$CH_2 = CHCH_2I + Mg \xrightarrow{Ether} CH_2 = CHCH_2MgI$$

Allyl Magnesium Iodide

The ease with which an alkyl halide forms a Grignard reagent depends on a number of factors. For a given alkyl group the ease of formation is alkyl iodide > alkyl bromide > alkyl chloride. The formation of Grignard reagent becomes increasingly difficult as the number of carbon atoms in the alkyl group increases. The ease of formation is:

$$CH_3X > C_2H_5X > C_3H_7X$$

Note: (1) All mono halides except allyl and benzyl halides form Grignard reagents.

Grignard reagents of allyl halide and benzyl halide are not prepared due to greater reactivity of these halides as these readily react with their Grignard reagent and form dibenzyl compounds.

$$C_6H_5 - CH_2 - X \xrightarrow[ether]{Mg} C_6H_5 - CH_2 - MgX \xrightarrow{C_6H_5 - CH_2 - X} C_6H_5 - CH_2 - CH_2 - C_6H_5$$

Dibenzyl

$$CH_2 = CH - CH_2 - X \xrightarrow[ether]{Mg} CH_2 = CH - CH_2 - MgX \xrightarrow{CH_2 = CH - CH_2 - X} CH_2 = CH - CH_2 - CH_2 - CH = CH_2$$

Dibenzyl

(2) Halo compounds having acidic hydrogen do not form Grignard Reagent.

2.2.3 Applications (Chemical Reactions of Grignard reagent):

(i) Formation of hydrocarbon

(a) Alkanes

When a Grignard Reagent is treated with any compound containing active Hydrogen (water, alcohols, amines), a hydrocarbon is produced

$$RMgX + HOH \longrightarrow RH + Mg(OH)X$$

$$[CH_3\text{-}Mg\text{-}I + HOH \longrightarrow CH_4 + OH\text{-}Mg\text{-}X]$$

Methyl magnesium iodide methane

$$RMgX + C_2H_5OH \longrightarrow RH + C_2H_5OMgX$$

$$[C_2H_5 - Mg - I + C_2H_5OH \longrightarrow C_2H_6 + C_2H_5\,O\text{-}Mg\text{-}I]$$

Ethyl Magnesium iodide Ethanol Ethane

$$CH_3\text{-}Mg\text{-}I + C_2H_5NH_2 \longrightarrow CH_4 + C_2H_5NH - I$$

Methyl magnesium iodide Ethylamine methane

(b) Higher alkanes:

When Grignard Reagent is treated with alkyl halide

$$CH_3\text{-}Mg\text{-}I + IC_2H_5 \longrightarrow H_3C.C_2H_5 + MgI_2$$

Ethyl iodide Propane

(c) Synthesis of higher alkenes

By action of unsaturated halides on the Grignard reagent

$CH_3MgBr + BrH_2CHC = CH_2 \longrightarrow CH_3CH_2CH{=}CH_2 + MgBr_2$

Methyl magnesium bromide Alkyl bromide 1-Butene

(d) Synthesis of higher alkynes:

By treating alkyl halide with Grignard reagent. The product is a higher alkyne

$CH_3C \equiv CH + CH_3MgI \longrightarrow CH_3C \equiv C - MgI + CH_4$

Propyne

$CH_3C \equiv C - MgI + CH_3I \longrightarrow CH_3C \equiv C\,CH_3 + MgI_2$

2-butyne

(ii) Formation of alcohols

A Grignard Reagent can be used to synthesise an alcohol by treating it with dry oxygen and decomposing the product with acid.

$$RMgX \xrightarrow{O_2} RO_2MgX \xrightarrow{RMgX} 2ROMgX \xrightarrow{H_3O^+} 2ROH$$

(a) When a Grignard Reagent is treated with formaldehyde or ethylene oxide, followed by hydrolysis, a **primary alcohol** is formed

$$R - MgI + HCHO \longrightarrow \begin{matrix} H & & R \\ & C & \\ H & & OMgI \end{matrix} \xrightarrow{H_3O^+} RCH_2OH$$

H : OH dil HCl

$HCH = O + CH_3MgI \longrightarrow CH_3CH_2O : MgI \longrightarrow$

$CH_3CH_2OH + HOMgI$

Formaldehyde Addition product

Ethanol (Pr. alcohol)

CH_3 H: OH dil HCl

$O + CH_3MgI \longrightarrow CH_3CH_2CH_2O : Mg \longrightarrow$

$CH_3CH_2CH_2OH + HOMgI$

CH_3 n- propyl alcohol

Ethylene oxide

(b) When a Grignard Reagent is treated with any aldehyde other than formaldehyde, a **secondary alcohol** is formed.

OMgX H_3O^+

$R^1CHO + RMgX \longrightarrow R' - CH \longrightarrow R^1CHOHR$

R

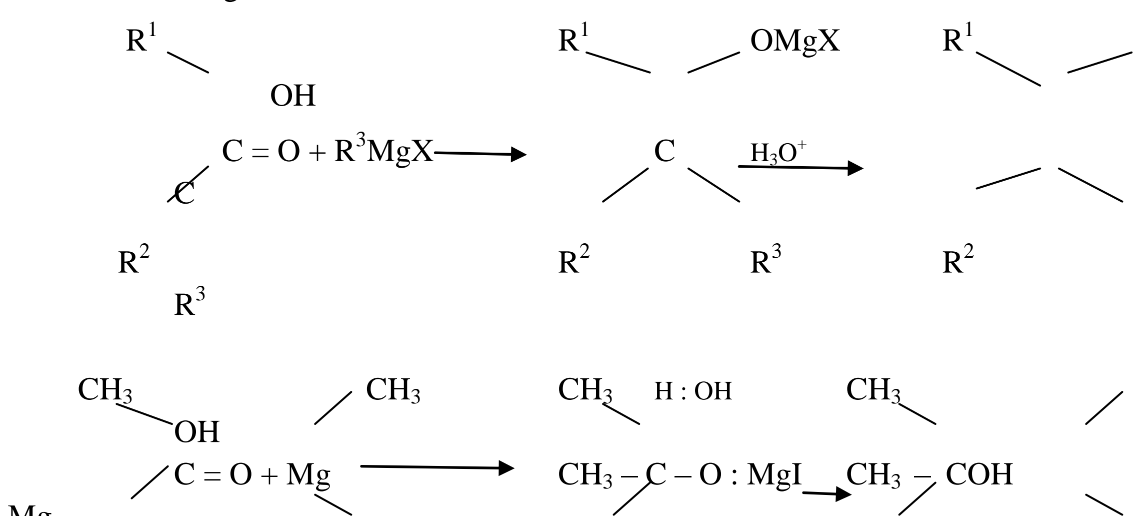
CH_3 H : OH CH_3

$CH_3CH = O + CH_3MgI \longrightarrow CH - O : MgI \longrightarrow CHOH + HOMgI$

CH_3 CH_3

Acetaldehyde Iso-propyl alcohol (Sec)

(c) A **tertiary alcohol** may be prepared by action of a Grignard Reagent on a ketone:



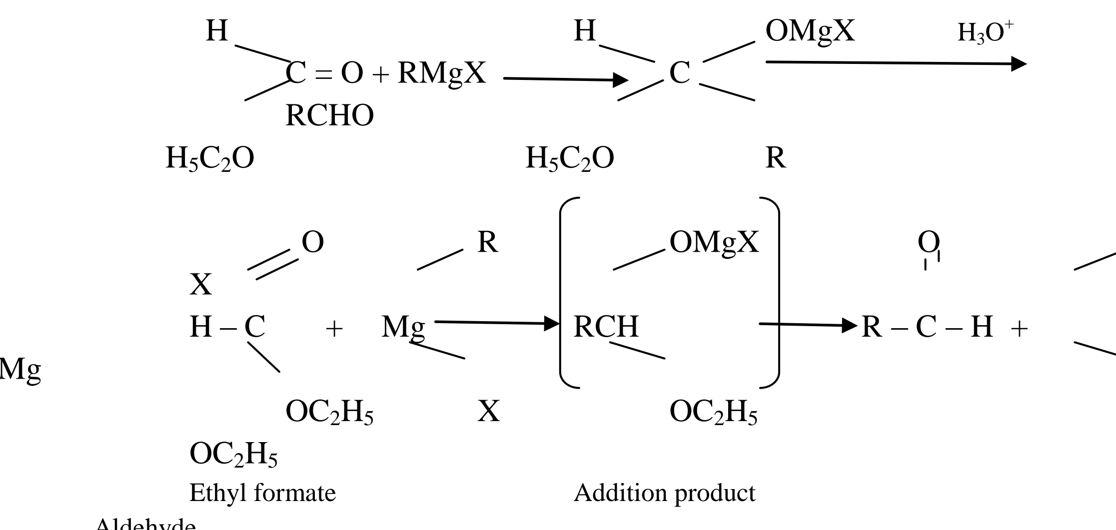
CH_3 I CH_3 CH_3

I

Acetone tert-butyl alcohol

(iii) Formation of aldehydes

An aldehyde may be prepared by the action of Grignard Reagent on ethyl formate or Hydrogen cyanide:



$$HC \equiv N + CH_3MgI \rightarrow CH_3CH = N : MgI \xrightarrow{2H_2O} CH_3CHO + NH_3 + HOMgI$$

O : H_2 H : OH

Hydrogen cyanide Addition product

⇨ If the Grignard Reagent is in excess a secondary alcohol is formed. Hence, to avoid this, the Grignard Reagent is added to the ester.

(iv) **Synthesis of higher ethers:** by action of Grignard Reagent on a mono chloro ether.

$$ROCH_2Cl + R'MgX \longrightarrow ROCH_2R' + ClMgX$$

Monochloroether Higher ether

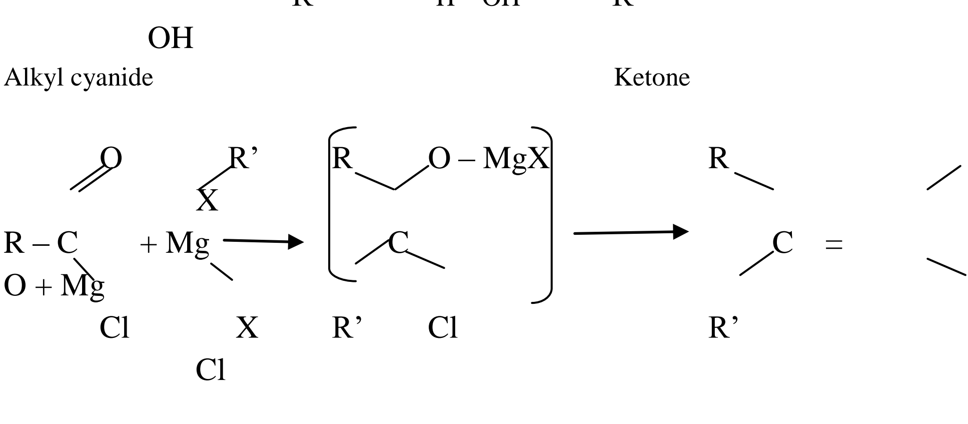
(v) Synthesis of Ketones:

By treating Grignard reagent with alkyl cyanide or acid chloride.

R H2 R

X

$$RC \equiv N + R'MgX \longrightarrow C = O\ N\ MgX \longrightarrow C = O + NH_3 + Mg$$



(vi) Synthesis of carboxylic acids:

By reacting Grignard reagent with solid CO2 (dry ice) followed by decomposing the addition product with dilute acid.

CH$_3$ — Mg — I + O=C=O ⟶ CH$_3$ – C(=O)–O – MgI ⟶ CH$_3$ – C(=O)–OH + Hg(I)(O$^-$)

Methyl acetate

ion

(vii) Synthesis of esters:

By reacting Grignard reagent with ethyl chloroformate.

$$C_2H_5MgI + C_2H_5COOCl \longrightarrow C_2H_5COOC_2H_5 + ClMgI$$

Ester

(viii) Synthesis of alkylcyanides:

By interacting cyanogens chloride or chloramines with Grignard Reagent.

$$C_2H_5MgI + N\,C\,Cl \longrightarrow C_2H_5CN + Mg\begin{matrix} I \\ Cl \end{matrix}$$

Cyanogen Chloride Ethyl cyanide

$$C_2H_5MgI + NH_2Cl \longrightarrow C_2H_5NH_2 + ClMgI$$

Chloramine Ethylamine

(ix) Synthesis of alkyl iodides:

By treating RMgCl or RMgBr with iodine.

$$Mg\begin{matrix} R \\ Br \end{matrix} + I - I \longrightarrow \underset{\text{alkyl iodide}}{RI} + Mg\begin{matrix} I \\ Br \end{matrix}$$

(x) Synthesis of other organometallic compounds:

By reacting metal chlorides with Grignard reagent.

$$HgCl_2 + 2C_2H_5MgI \longrightarrow \underset{\text{Diethyl mercury}}{(C_2H_5)_2Hg} + 2[ClMgI]$$

$$2PbCl_2 + 4C_2H_5MgI \longrightarrow \underset{\text{Tetraethyl Lead}}{(C_2H_5)_4Pb} + 4[ClMgI]$$

(xi) Formation of Silicon polymer

A Grignard Reagent reacts with silicon tetrachloride and the product on hydrolysis forms silicon polymer

$$2RMgX + SiCl_4 \longrightarrow R_2SiCl_2 + 2MgXCl$$

Dichloro dialkyl silane

$$R_2SiCl_2 + H2O \longrightarrow -\overset{R}{\underset{R}{Si}} - O - \overset{R}{\underset{R}{Si}} -$$

Silicon polymer

2.3 ORGANO – ZINC COMPOUNDS

Dialkylzincs (R – Zn – R) were the first organic compounds discovered by Frankland in 1849, when he was trying to prepare the ethyl radical ($CH_3CH_2^o$), by removing iodine from ethyliodide by means of Zn.

2.3.1 Formation / Synthesis of Dialkylzincs:

By heating alkyliodide (RI) with Zn in an atmosphere of CO_2, followed by distilling the alkylzinc iodide so formed in an atmosphere of CO_2

$$RI + Zn \longrightarrow R - Zn - I$$

$$2R - Zn - I \longrightarrow R_2Zn + ZnI_2$$

Dialkylzinc

e.g. $$2C_2H_5I \xrightarrow{2Zn} 2C_2H_5ZnI \xrightarrow{-ZnI_2} Zn(C_2H_5)_2$$

2.3.2 Applications / Chemical Reactions of dialkylzincs:

(i) Synthesis of hydrocarbons containing quarternary carbon atom.

Action of dimethyl zinc on *tert.*- Butyl chloride yields neo-pentane

$$H_3C - \underset{CH_3}{\overset{CH_3}{\underset{|}{\overset{|}{C}}}} - Cl + (CH_3)_2Zn \longrightarrow H_3C - \underset{CH_3}{\overset{CH_3}{\underset{|}{\overset{|}{C}}}} - CH_3 + CH_3ZnCl$$

Neo-heptane

(ii) Synthesis of ketones

By reaction of dialkylzinc with acid chloride.

$$R'COCl + R_2Zn \longrightarrow R'COR + RZnCl$$

Acyl chloride → Ketone

2.4 ORGANO – LITHIUM COMPOUNDS

Alkyllithiums are very reactive and have been found to be more useful than the Grignard reagent for some syntheses.

2.4.2 Formation / Synthesis of Alkyllithiums:

(i) Some organo-lithiums can be prepared by heating alkyl or aryl halide in ether or benzene with Zn metal.

$$RX + 2Li \longrightarrow RLi + LiX$$

Alkyllithium

$$CH_3(CH_2)_3Cl + 2Li \longrightarrow CH_3(CH_2)_3Li + LiCl$$

n-butyllithium

(ii) Many alkyl-lithiums can be prepared by halogen-metal exchange by heating a suitable alkyl-lithium (usually *n*-butyl-lithium) with an appropriate alkyl halide.

$$RLi + R'X \longrightarrow R'Li + RX$$

Where *R* is usually *n*-butyl radical

2.4.2 Applications / Chemical Reactions of of dialkyl-lithiums:

(i) Synthesis of alkanes

$$\underset{n\text{-butyllithium}}{C_4H_9Li} + H_2O \longrightarrow \underset{\text{n-butane}}{C_4H_{10}} + LiOH$$

(ii) Synthesis of Carboxylic acids

$$RLi \xrightarrow{CO_2} R - \overset{O}{\overset{||}{C}} - OLi \xrightarrow{H^+/H_2O} R - \overset{O}{\overset{||}{C}} - COH + LiOH$$

Addition Product hydrolysis Carboxylic acid

(iii) Synthesis of ketones

$$\underset{[\text{From(ii)}]}{R- \overset{O}{\overset{||}{C}}- OLi} \xrightarrow{RLi} R- \overset{O}{\overset{||}{C}}- OLi \xrightarrow{H^+/H_2O} \underset{\text{Ketone}}{R - \overset{R}{\overset{|}{C}} = O} + 2\,LiOH$$

(R, Li labels above the intermediate)

(iii) Synthesis of Other Organometallics

$$3CH_3Li + \underset{\text{Auric bromide}}{AuBr_3} \longrightarrow \underset{\text{Trimethyl gold}}{(CH_3)_3Au} + 3LiBr$$

(iv) Synthesis of ketones

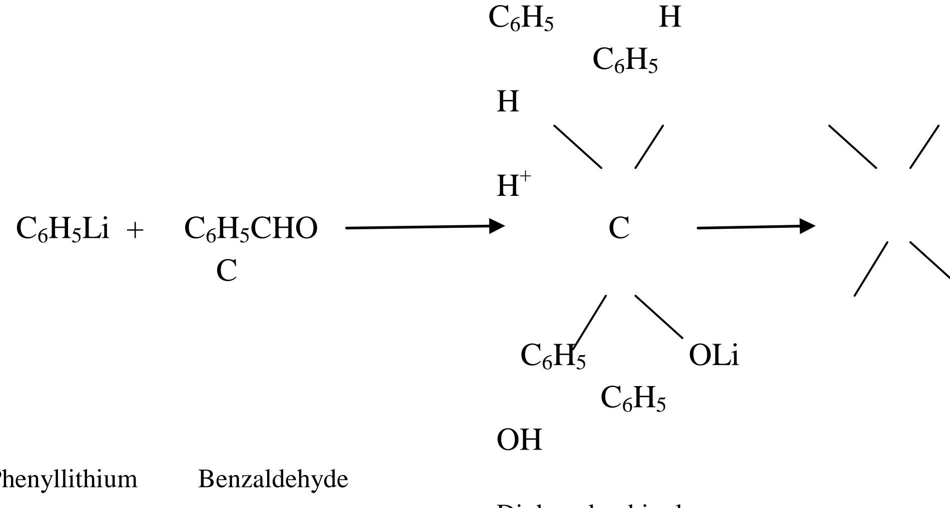
$$R - \overset{O}{\overset{||}{C}}- OH + LiR \longrightarrow R- \overset{O}{\overset{||}{C}}- OLi \xrightarrow{LiR} R- \overset{R}{\overset{|}{\underset{R}{\underset{|}{C}}}}- OLi \xrightarrow{H+,\ \text{Hydrolysis}} C=O$$

(R labels as printed; "OLi" struck through)

OLi

Ketone

(v) Synthesis of Sec Alcohols



(vi) Synthesis of higher alcohols

$CH_2C_4H_9$

$C_4H_9Li \longrightarrow CH_2 = CH_2 \xrightarrow[\text{pressure}]{\text{High}} \quad | \quad \xrightarrow{HCHO} CH_2Li$

$C_4H_9(CH_2)_2CH_2OH$

Butyl lithium Ethylene

Higher alcohol

2.5 SUMMARY:

Compounds in which the central metal atom is bonded directly to the Carbon atom of the Hydrocarbon radical or molecule are called Organometallic compounds. The metals bonded are generally less electronegative than Carbon, e.g., B, As, Si, Mg, Li, Zn, Cd, Pb.

Owing to their versatile reactivity, these compounds are exhaustively utilized in conversions and product formations in different environments.

2.6 TERMINAL QUESTIONS

QUESTION No. 1: What are organ- metallic compounds?

QUESTION No. 2: Give the preparation and properties of organo-lithium Compounds

QUESTION No. 3: Complete the following reactions:

(a) $CH_3 - CH_2 - CH_2 - Br \xrightarrow{\text{Li / Ether}}$ **?**

(b) $C_6H_5 - CH_2 - CH_2 - Br \xrightarrow{\text{Mg / Ether}}$ **?**

(c) $CH_3 - CH_2 - C \equiv C - H + CH_3MgBr \longrightarrow$ **?**

QUESTION No. 4: Identify the products [A], [B] and [C]

(a) $C_6H_5 - CH = CH - CHO + CH \equiv C - MgBr \xrightarrow{HOH / H^+}$ **[A]**

(b) $C_6H_5 - CHO + CH_3NH_2 \xrightarrow{H+ / \Delta}$ **[B]** $\xrightarrow[\text{(ii) HOH}]{\text{(i) } C_6H_5 - CH_2 MgBr}$ **[C]**

QUESTION No. 5: Synthesise the following with the help of Grignard's regents. Give equations only.

(a) Ethane from Ethyl Bromide
(b) Isopropyl Alcohol from acetaldehyde
(c) *tert* – butyl alcohol from acetone
(d) Acetone form acetyl chloride
(e) Ethyl alcohol from formaldehyde

2.7 ANSWERS TO TERMINAL QUESTIONS

SOLUTION No 1: *refer page no 3 {Section 2.1} of the SLM*

SOLUTION No 2: *refer page nos 10 & 11{Section 2.4} of the SLM*

SOLUTION No 3: Reaction completion:

(a) $CH_3 - CH_2 - CH_2 - Br \xrightarrow{\text{Li / Ether}} CH_3 - CH_2 - CH_2 - Li + LiBr$

(b) $C_6H_5 - CH_2 - CH_2 - Br \xrightarrow{\text{Mg / Ether}} C_6H_5 - CH_2 - CH_2 - MgBr$

(c) $CH_3 - CH_2 - C \equiv C - H + CH_3MgBr \longrightarrow CH_3 - CH_2 - C \equiv C - H + CH_4$

SOLUTION No 4: Product identification –

[A] is : $C_6H_5 - CH = CH - CHOH - C \equiv CH$

[B] is : $C_6H_5 - CH \equiv N - CH_3$

[C] is : $C_6H_5 - CH(C_6H_5) - N - CH_3$

SOLUTION No 5:

(a) $4C_2H_5Br + Mg \xrightarrow{\text{Dry ether}} C_2H_5 : MgBr \xrightarrow{H : OH} C_2H_6 + Mg(OH)Br$

Ethyl Bromide — Ethane

(b) $CH_3CH{=}O + CH_3MgI \longrightarrow (CH3)(CH3)CH\text{-}O{:}MgI \xrightarrow{H:OH} (CH3)(CH3)CHOH + HOMgI$

Acetaldehyde — Iso-propyl alcohol (Sec)

(c) $(CH_3)_2C = O + CH_3MgI \longrightarrow (CH_3)_3C - O : MgI \xrightarrow{H : OH} (CH_3)_3COH + Mg(OH)I$

Acetone — tert-butyl alcohol

(d) $CH3 - C(=O)Cl + CH3MgX \longrightarrow [(CH3)(CH3)C(Cl)(O - MgX)] \xrightarrow{H : OH,\ dil\ HCl} (CH3)(CH3)C = O + Mg(X)Cl$

Acetyl chloride — Addition product — Ketone

(e) $HCH = O + CH_3MgI \longrightarrow CH_3CH_2O : MgI \longrightarrow CH_3CH_2OH + HOMgI$

Formaldehyde Addition product Ethanol (Pr. alcohol)

Unit 3: Sulphur Containing Compounds

3.1 Introduction

Objective

3.2 Classification of Sulphur Containing Compounds

3.3 The special characteristics of Organosulfur Compounds

3.4 Methods of formation and chemical reactions of Thiols

3.5 Biological importance of Thiols

3.6 Methods of formation and chemical reactions of Thioethers

3.7 Methods of formation and chemical reactions of Sulphonic acids

3.8 Methods of formation and chemical reactions of Sulphonamides

3.9 Methods of formation and chemical reactions of Sulphaguamidine

3.10 Summary

3.11 Terminal Questions

3.1 Introduction

Sulfur is essential to the life and growth of all organisms from microbes to man. Plants and microorganisms have the capacity to utilise inorganic sulfur, since they can reduce oxidised forms of sulfur to the sulfur amino acids (1) and (2). Mammals, on the other hand, cannot utilise inorganic forms of sulfur and consequently have to obtain the essential sulfur amino acids from plant sources. The sulfur amino acids (1)–(3) are used by living organisms as the source of a vast array of biochemically vital organosulfur compounds, e.g. proteins, glutathione, coenzyme A, the vitamins biotin and thiamine (part of the vitamin B complex), lipoic acid, and plant and fungal metabolites like penicillin.

Objective

Learner is able to define

- To Classify Sulphur Containing Compounds
- Methods of formation and chemical reactions of Thiols
- Thioethers
- Sulphonic acids
- Sulphonamides
- Sulphaguamidine

Basics of Sulphur Containing Compounds

A molecule containing one or more carbon-sulfur bonds is known as organosulfur compounds. Sulfur itself has been known since ancient time and it was mostly found in volcanic eruptions, which was referred ‘brimstone’ (burning stone). In the third century, Chinese alchemists used a mixture of sulfur with saltpetre (KNO_3) as primitive gunpowder. Elemental sulfur is widely distributed in nature and also as compounds such as hydrogen sulfide, sulfur dioxide, sulfates, e.g. those of calcium and magnesium, and sulfide ores.

Sulfur was recognized as a common element in proteinaceous material in the early eighteenth century, and the disulfide cystine (1) was probably the first sulfur amino acid was discovered in year 1899 while its structure was elucidated by Friedmann (1903).

$SCH_2CH(NH_2)CO_2H$
|
$SCH_2CH(NH_2)CO_2H$

Disulphide cysteine
(1)

$MeSCH_2CH_2CH(NH_2)CO_2H$

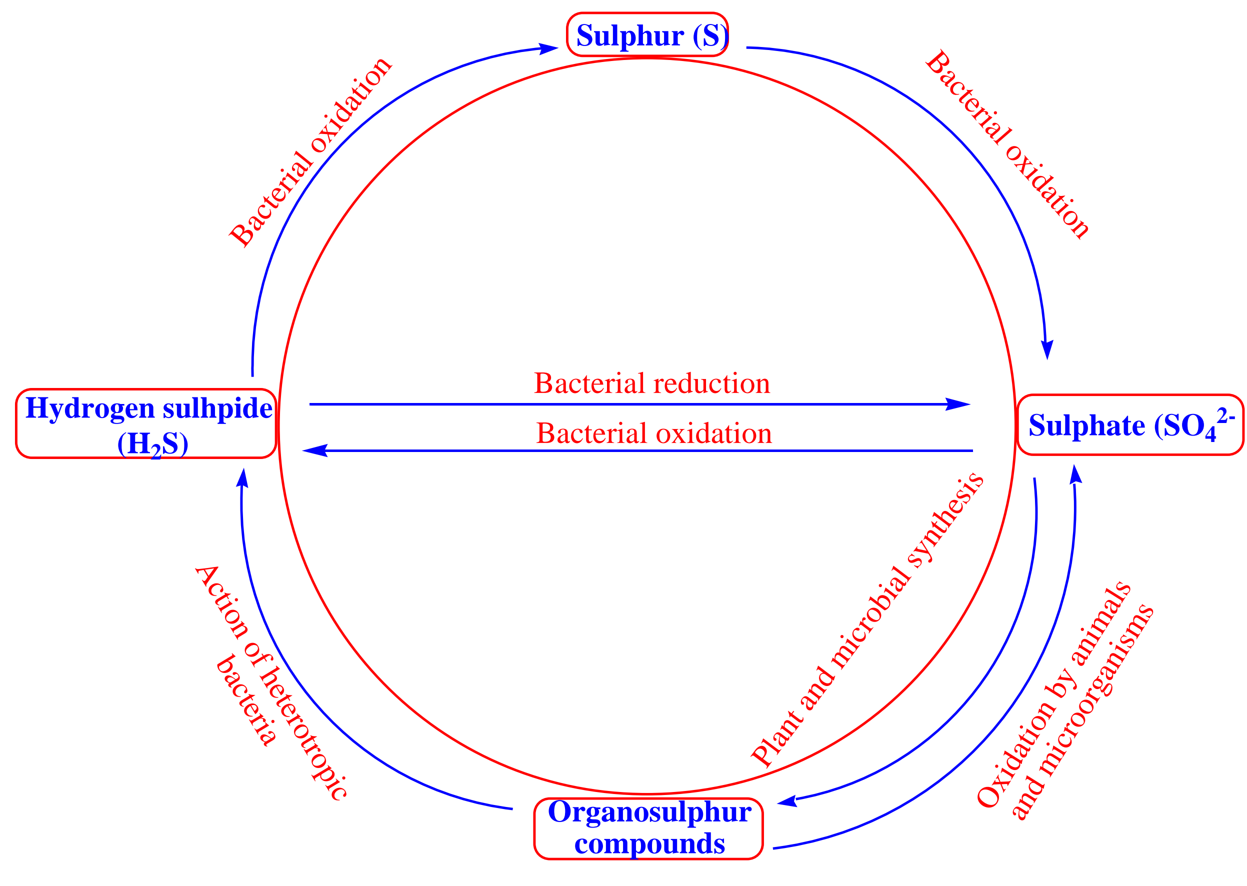
Methionine
(2)

SH
|
CH_2
|
$CH(NH_2)$
|
CO_2H

Pure cysteine
(3)

Methionine (2) was isolated by Mueller (1922) and it was then that the nutritional significance of the relationship of (1), (2) and (3) emerged; pure cysteine (3) is difficult to isolate owing to its facile oxidation to cystine (1).

The relationship between elemental sulfur and the oxidised and reduced forms in nature is depicted in the sulfur cycle. This demonstrates that sources of oxidized sulfur (e.g. inorganic sulfates) and reduced sulfur (e.g. thiols) in the biosphere can be interconverted in living organisms; the special chemistry of sulfur allows these oxido-reductive transformations to occur relatively easily.



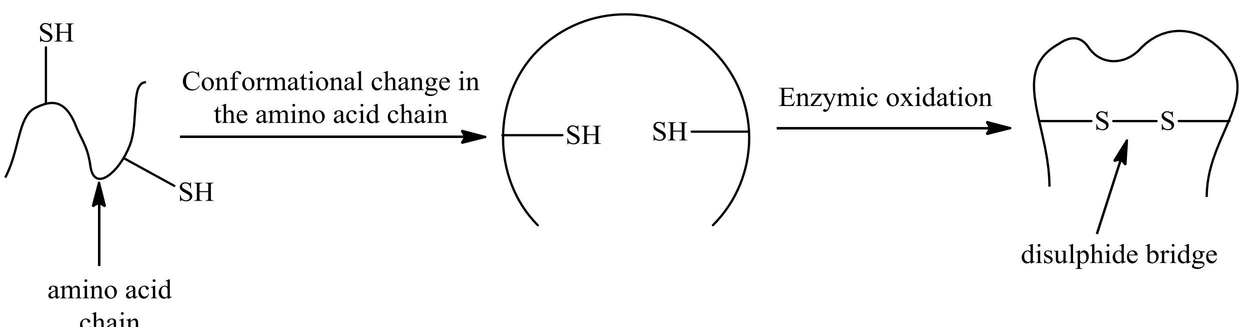
The Sulphur cycle

Thiols (RSH) are easily oxidized under mild conditions to yield disulfides (RSSR), and this process can be reversed by reduction with, for instance, zinc-dilute acid. The reversible disulfide formation from thiols is an important biological process, since many proteins and peptides possess free thiol (SH) groups that form bridging disulfide links. This mechanism is exploited in nature to achieve the intermolecular and intramolecular joining together of amino acid chains. Thiols are of considerable biological significance since many vital enzymes in plants, animals and fungi contain the SH group.

$$2\ RSH \underset{\text{Reduction } (Zn/H^{\oplus})}{\overset{\text{Oxidation } (I_2)}{\rightleftharpoons}} RSSR + 2HI$$

The formation of ether by the reaction between ethanol and sulfuric acid was investigated as early as 1540, but the intermediate, ethyl hydrogensulfate, was first observed by Dabit in 1820. Carbon disulfide was originally prepared by Lampadius (1796); this was a very significant discovery since carbon disulfide

can be used as the starting material for the synthesis of many organosulfur compounds. The xanthate reaction was discovered by Zeise (1815) and really marks the beginning of organosulfur chemistry.



The xanthate reaction provides an easy preparative route to dithio acids (4) and was of major industrial importance in the manufacture of rayon and cellophane, which are obtained by treatment of cellulose with carbon disulfide and alkali.

$$ROH + KOH + CS_2 \longrightarrow ROC(=S)S{-}K^{\oplus} \xrightarrow{H^{\oplus}} ROC(=S)SH \quad (4)$$

3.2 Classification of Sulphur Containing Compounds

Classification of organosulphur compounds: The organosulfur compounds may be classified into three types, *viz.*

(i) Bivalent organosulfur compounds e.g. thiols, sulfides, disulfides, sulfenic acids and their derivatives, thiocarbonyl compounds, thioamides, thioesters etc.

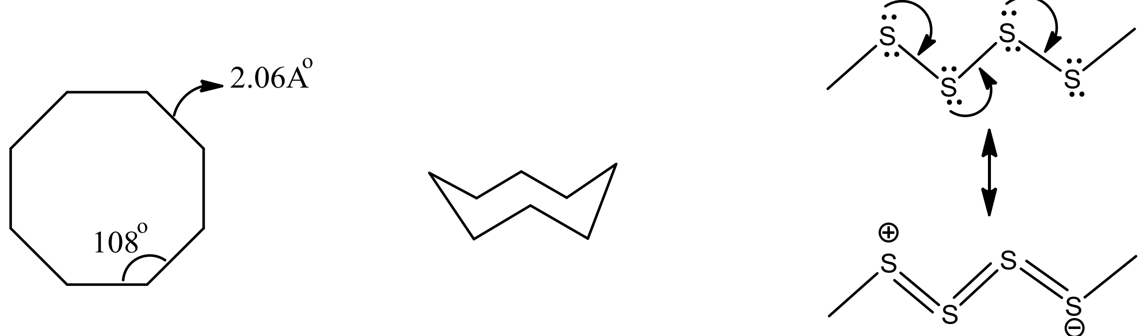
(ii) Tricoordinated organosulfur compounds e.g., sulfoxides, sulfinic acids and their derivatives, thiosulfinate esters etc.

(iii) Tetracoordinated organosulfur compounds e.g., sulfone, sulfurane, sulfonic acids and their derivatives, sulfonate esters etc.

Structure and bonding in sulfur and organosulfur compounds

Elemental sulfur is a bright yellow crystalline solid, it is the second element of group VI in the periodic table of the elements, just below oxygen, and has the electronic configuration: ${}_{16}S=1s^2 2s^2 2p^6 3s^2 3p^4$.

Sulfur shows a marked tendency to react with itself (catenation) and consequently can exist in a large number of acyclic and cyclic species. All the chain and ring forms of sulfur are thermodynamically less stable than cyclooctasulfur (S_8) at 25°C. Cyclooctasulfur is the most common form of sulfur and this occurs in the crown conformation. The sulfur-sulfur bond in elemental sulfur is probably a resonance hybrid, so the S_8 ring exists in the energetically favoured crown shape. Such electronic delocalisation would also account for the colour of sulfur and the special features of Sulfur-Sulfur bonds in the ultraviolet spectrum.



Ortharhombic sulphur in S_8 form

When diatomic sulfur (S_2) is dissolved in a polar solvent like methanol or acetonitrile, an equilibrium is established in which ~1% of the sulfur exists as the S_6 and S_7 ring forms. Since these are much more reactive than the normal S_8 rings, they may provide a pathway for the reactions of sulfur in polar solvents. Cyclic sulfur allotropes also dissolve in carbon disulfide, benzene and cyclohexane. Sulfur reacts with amines, e.g. piperidine, giving coloured solutions containing polythiobisamines (1).

$$2RR'NH + S_8 \longrightarrow \underset{(1)}{(RR'N)_2S_7} + H_2S$$

Sulfur reacts with many organic molecules, and in such sulfur-sulfur bond-breaking reactions free radicals may be involved; the reactions are often catalysed by amines and Lewis acids. Amines and other bases activate sulfur by formation of nucleophilic sulfur species, while Lewis acids cleave the sulfur-sulfur bond to give electrophilic sulfur moieties. Saturated hydrocarbons can be dehydrogenated by heating with sulfur; thus, cyclohexane (2) is converted to benzene (3).

$$\text{(2)} \xrightarrow{S_8 > 200^{o}C} \text{(3)}$$

In some cases, organic compounds may cyclise on heating with sulfur to give thiophenes (4).

$$\text{–CH–CH–, –CH}_2\text{H}_2\text{C–} \xrightarrow{S_8 \ \text{Heat}} \text{(4)}$$

In the presence of bases, mild thiolation of organic compounds by treatment with sulfur provides an efficient route to organosulfur compounds. The reaction with alkenes is of great industrial importance in the vulcanisation of rubber. Many common reactions of sulfur can be interpreted as a nucleophilic attack on the sulfur-sulfur bonds; for instance, the reaction with cyanide anion. The initial ring opening (i) is followed by a series of nucleophilic reactions on the Sulfur-Sulfur bonds (ii) with displacement of the thiocyanide anion leading to the overall reaction (iii).

(i) $S_8 + CN^- \longrightarrow S^{\ominus}\text{–}S_6\text{–}SCN$

(ii) $S_6^{\ominus}\text{—S—SCN} + CN \longrightarrow S_6^{\ominus}\text{—SCN} + SCN^{\ominus}$

(iii) $S_8 + 8CN^{\ominus} \longrightarrow 8SCN^{\ominus}$

3.3 The special characteristics of organosulfur compounds

Since sulfur lies just below oxygen in group VI of the periodic table, the chemistry of organosulfur compounds should parallel that of the oxygen analogues; indeed, there are many similarities, e.g. between alcohols and thiols and ethers and sulfides, because both elements possess the same outer electronic configuration (s^2p^4). However, the prediction is not fully realized since many factors serve to differentiate sulfur from oxygen.

(i) Sulfur (electronegativity 2.44) is appreciably less electronegative than oxygen (3.5), and this lessens the ionic character of organosulfur compounds in comparison with the analogous oxygen derivatives and decreases the importance of hydrogen bonding.

(ii) Sulfur, like most elements in the second and higher rows of the periodic table, is reluctant to form normal π-double bonds; thus, thiocarbonyl (C=S) compounds are comparatively rare and are usually unstable with a tendency to polymerise. This is a result of the relatively low effectiveness of *pπ-dπ* bonding involving lateral overlap of the *3p*-orbitals and arising from the larger size of the sulfur atom as compared with carbon. For sulfur, unlike oxygen, the diatomic S_2 molecule is relatively unstable; the most stable form of sulfur is the cyclic S_8 molecule known as cyclooctasulfur.

(iii) The sulfur atom (atomic radius 1.02Å) is larger than oxygen (atomic radius 0.73 Å), and its outermost electrons are therefore more shielded from the attractive force of the positive nucleus. Sulfur is consequently more polarisable than oxygen; the sulfur lone pairs of electrons are better nucleophiles but weaker bases in reactions with acids. The outer electronic shell in sulfur contains not only *s*-electrons and *p*-electrons but also vacant *3d*-orbitals which can be utilized in bonding. The valency of sulfur, therefore, unlike that of oxygen, is not limited to two. Sulfur can expand its octet and form many hypervalent compounds of higher oxidation states, e.g. SO_2, SO_3, SF_4, SF_6, organic sulfoxides, sulfones and sulfonic acids. The bonding in such four-valent and six-valent sulfur compounds may involve promotion of *3p*-electrons and *3s*-electrons into the *3d*-orbitals by formation of either four sp^3d or six sp^3d^2 hybrid bonds; however, a three-centre four-electron bond, termed a hypervalent bond, is now generally favored.

Although normal *pπ-pπ* double bonds are not generally found in sulfur compounds, another type of π-bonding arising from overlap involving the

vacant sulfur *d*-orbitals is often very important. The possibility of such π-bonding arises in compounds containing sulfur attached to an element like oxygen, in which, in addition to a π-bond between the atoms, a π-bond may result by utilizing an unshared electron pair on the oxygen atom and the vacant sulfur *d*-orbitals. The resultant 'double bond' is correctly described as involving the *p*π-*d*π overlap and differs from a π-double bond because it involves expansion of the valence shell of the sulfur atom to accommodate more than eight electrons. A typical example is dimethyl sulfoxide (DMSO) (5) which may be regarded as a resonance hybrid of (5a) and (5b).

Me_2SO (5) $Me_2S{=}O$ (5a) $Me_2S^{\oplus}{=}O^{\ominus}$ (5b)

S=O bonds involving *p*π-*d*π overlap do not require coplanarity of the groups attached to the sulfur atoms as distinct from normal double bonds, so DMSO (5) differs from acetone in possessing a non-planar pyramidal configuration. Divalent sulfur compounds, because of the available lone electron pairs and vacant *d*-orbitals which permit expansion of the sulfur valence shell, are readily oxidized to a wide variety of compounds differing in the oxidation state of the sulfur atom. In such hypervalent compounds, the oxidation site is sulfur and not carbon; for example, thiols (6) and sulfides (7) are oxidized to sulfonic acids (8), sulfoxides (9) or sulfones (10).

$$\underset{(6)}{RSH} \xrightarrow[\text{[} KMnO_4 \text{]}]{\text{Strong oxidants}} \underset{(8)}{R{-}S(=O)_2{-}OH}$$

$$\underset{(7)}{RSR} \xrightarrow[MeCO_3H]{\text{Peracid}} \underset{(9)}{R{-}S(=O){-}R} \xrightarrow{\text{further oxidation}} \underset{(10)}{R{-}S(=O)_2{-}R}$$

The sulfur-oxygen bonds are polar, so oxidation results in changes in the chemical behaviour of the parent compounds. On oxidation, the nucleophilicity is decreased but acidity, polar character, water solubility and leaving group

capacity increases. These changes are important in the metabolism of organosulfur compounds; the higher oxidation state compounds like sulfones can function as electrophiles and are good leaving groups in nucleophilic displacement reactions.

Bivalent organosulfur compounds:

Thiols: Several thiols occur naturally; for example, skunk secretion contains 3-methyl-butanethiol and cut onions evolve 1-propanethiol, and the thiol group of the natural amino acid cysteine plays a vital role in the biochemistry of proteins and enzymes**.**

A thiol (RSH) is structurally similar to an alcohol (ROH) but contains a sulfur atom in place of the oxygen atom normally found in an alcohol. The outstanding feature of thiol derivatives is the foul smell. The simplest thiol is hydrogen sulfide (H_2S) and its oxygen analog is water (H_2O). It can be detected by the human nose at a concentration of a few parts per billion and is readily identifiable as having the odor of rotten eggs. Ethane thiol (CH_3CH_2SH) is added in trace amounts to natural gas to give it a detectable odor.

Oxidation states of sulfur compounds: Oxygen and sulfur both belongs from the same group but oxygen shows two oxidation states i.e. -1 in peroxides and -2 in other compounds while sulfur have different oxidation states which is given in following table:

Oxidation states of sulphur in organic compounds					
-2	**-1**	**0**	**+2**	**+4**	**+6**
H_2S Hydrogen sulfide R–S–H Thiols R–S–R Sulfides R_3S^{+} Sulfonium ions	R–S–S–R Disulfides	S Elemental sulphur R–S(=O)–R Sulfoxides R–S–OH Sulfenic acids	R–S(=O)(=O)–R Sulfones R–S(=O)–OH Sulfinic acids	SO_2 Sulphur dioxide R–S(=O)(=O)–OH Sulfonic acids R–O–S(=O)–O–R Sulfite esters	SO_3 Sulphur trioxide R–O–S(=O)(=O)–O–R Sulfate esters

3.4 Methods of formation and chemical reactions of thiols

Nomenclature of thiols:

The sulfur analog of an alcohol is called a thiol (thi- from the Greek: theion, sulfur) or, its ancient name mercaptan, which means “mercury capturing.” Thiols react with Hg^{2+} in aqueous solution to give sulfide salts as insoluble precipitates. Thiophenol, C_6H_5SH, for example, gives $(C_6H_5S)_2Hg$.

Sulfur is one row down from oxygen in the periodic table. Compounds that contain sulfur instead of oxygen might be expected to bear similar characteristics of the oxygen compound. But, the chemistry of the SH function is much different from the OH function. The SH group is called a sulfhydryl function, but organic compounds containing the SH are called thiols in the IUPAC system, or called mercaptans in older nomenclature which means “mercury capturing.” Thiols react with Hg^{2+} in aqueous solution to give sulfide salts as insoluble precipitates. Thiophenol, C_6H_5SH, for example, gives $(C_6H_5S)_2Hg$. Thiol names in the IUPAC system are derived from the parent alkane by dropping the ‘e’ followed by adding -thiol. A number is used to locate the position of the thiol. Older mercaptan names are obtained by naming the alkyl group followed by a second word, mercaptan.

CH_3CH_2SH	$CH_3CH(CH_3)CH_2SH$	$HSCH_2CH_2OH$
Ethanthiol (Ethyl mercaptan)	2-Methyl-1-propanethiol (Isobutyl mercaptan)	2-Mercaptoethanol

Sulfur analogs of ethers (thioethers) are named by using the word sulfide to show the presence of the –S- group. Following are common names of two sulfides:

CH_3SCH_3	$CH_3CH_2SCH(CH_3)CH_3$
Dimethyl sulfide	Ethyl isopropyl sulfide

Formation of thiols: Thiols, or thioalcohols (RSH) are sulfur analogues of alcohols and phenols in which the oxygen atom has been replaced by sulfur; they are derivatives of hydrogen sulfide in the same way that alcohols may be regarded as being derived from water. The volatile thiols have very unpleasant smells rather like the smell of hydrogen sulfide. In comparison with oxygen,

sulfur has a larger atomic size, more diffuse electronic orbitals and appreciably lower electronegativity; consequently, sulfur-hydrogen bonds are weaker than oxygen-hydrogen bonds. Thiols are therefore more acidic than the corresponding alcohols and form much weaker hydrogen bonds; this is shown by their relatively lower boiling points and aqueous solubility. Thiols can be prepared by following methods:

1. Formation from alkanes

(a) Hydrogen sulphide additions: This is in principle the simplest process for preparing alkanethiols. Usually the use of a high hydrogen sulphide ratio favours the formation of the thiol. Two isomeric thiols are formed from an unsymmetric alkene.

$$RCH{=}CH_2 + H_2S \longrightarrow \underset{(1)}{RCH(SH)CH_3} + \underset{(2)}{RCH_2CH_2SH}$$

Where; product 1 is termed the Markownikoff product and 2, the anti-Markownikoff adduct.

(a) Additions of other suphur acids: Additions of other sulphur acids to alkenes are frequently used in thiol preparations; the required thiol being obtained from the intermediate compounds after hydrolysis or dealkylation.

$$R^1CH{=}CHR^2 + R^3SH \longrightarrow R^1CH(SR^3)CH_2R^2 \longrightarrow R^1CH(SH)CH_2R^2$$

Where, R^3= alkyl, aryl or acyl

2. Formation from alcohols

(a) Using hydrogen sulphide: Direct reaction between hydrogen sulphide and alcohols normally requires the presence of a catalyst. Several processes involve basic catalysts, high temperatures and high pressures.

$$ROH + H_2S \xrightarrow{\text{Catalyst}} RSH + H_2O$$

(b) Using phosphorus pentasulphide: C_4-C_{16} alcohols can be converted to dialkyldithiophosphates, which on acid hydrolysis give the corresponding thiols. Yields greater than 70% can be obtained if the sulphides formed in the reaction are dealkylated to give thiols as well.

P_2S_5 Pyridine, Heat

1*H*-imidazo[4,5-*d*]pyridazine-4,7-diol → 1*H*-imidazo[4,5-*d*]pyridazine-4,7-dithiol

P_2S_5 Pyridine, Heat

3,5-diiodopyridin-2(1*H*)-one → 3,5-diiodopyridine-2-thiol

3. Formation from halides: When alkyl halides are heated with an alcoholic solution of sodium or potassium hydrogen sulphide; thioalcohols are formed.

$$RX + KSH \xrightarrow{Heat} RSH + KX$$

$$C_2H_5Cl + NaSH \xrightarrow{Heat} C_2H_5SH + KX$$

4. Formation from thiourea: This method, generally involves the reaction of a halide with thiourea to give an iso-thiouronium salt is formed which on treatment with sodium hydroxide gives thiol.

$$R{-}X + S{=}C(NH_2)_2 \longrightarrow R{-}S{-}C({=}\overset{\oplus}{N}H_2)NH_2\ \overset{\ominus}{X} \longrightarrow RSH$$

5. Formation from quinones: Reaction of p-benzoquinones with thiosulphate led to 1,4-dihydroxyphenyl thiosulphates, which on reduction with zinc and hydrochloric acid gave the mercaptodihydroxybenzenes.

$$\text{R-C}_6\text{H}_3\text{O}_2 \ (\text{R-benzoquinone}) \xrightarrow{S_2O_3^{2-}} \text{R-C}_6\text{H}_2(\text{OH})_2\text{-O-S(=O)(=S)-O}^- \xrightarrow{Zn/H_3O^{\oplus}} \text{R-C}_6\text{H}_2(\text{OH})_2\text{-SH}$$

R= H, Me and Ph

6. Formation from Grignard reagent: When sulphur is treated with Grignard reagent, thiocomplexes are formed which on acid hydrolysis gives thiol.

$$S_8 + R\text{—}MgX \longrightarrow R\text{—}S\text{—}MgX \xrightarrow{HOH/H^{\oplus}} R\text{—}S\text{—}H$$

Chemical reactions of thiols: The thiol group (-SH) is the sulfur analog of the hydroxyl group (-OH) found in alcohols. Since sulfur and oxygen belong to the same periodic table group, they share some similar chemical bonding properties. Like alcohol, in general, the deprotonated form RS^- (called a thiolate) is chemically more reactive than the protonated thiol form RSH. The chemistry of thiols is thus related to the chemistry of alcohols: thiols form thioethers, thioacetals and thioesters, which are analogous to ethers, acetals and esters respectively.

1. Acidic character: The sulfur atom of a thiol is quite nucleophilic, rather more than that of oxygen atom of alcohols. The thiol group is fairly acidic with a usual pK_a around 10 to 11. In the presence of a base, a thiolate anion is formed which is a very powerful nucleophile. The group and its corresponding anion are readily oxidized by reagents such as bromine to give an organic disulfide (R-S-S-R). Oxidation by more powerful reagents, such as sodium hypochlorite or hydrogen peroxide yields sulfonic acids (RSO_3H).

$$2\,R\text{—}SH + Br_2 \longrightarrow RSSR + 2HBr$$

$$R\text{—}SH + 3\,H_2O_2 \longrightarrow RSO_3H + 3H_2O$$

Acidity is described as the ability to lose H^+ ion in solution. In case of thiols (R-SH) after losing H^+ ions; negative charge is generated on sulphur which is more stabilized. So, if the resulting state of compound is more stable, its activity to reach there is faster and hence thiols lose H^+ ions easily as compared to alcohols for which negative charge on oxygen is unstable and will not let H^+ go that easily. Hence, thiols are more acidic in nature than alcohols.

2. Reaction with alkali metals: Thiols, unlike alcohols, form insoluble salts (mercaptides) by reaction with heavy metals like mercury and lead. The former name 'mercaptan' for thiols comes from the Latin mercurium captans which means mercury seizing. In modern nomenclature the name 'thiol' is preferred over 'mercaptan', although the prefix 'mercapto' is still allowed for the unsubstituted -SH radical.

$$2\,RSH + HgO \longrightarrow Hg(SR)_2 + H_2O$$

$$2\,RSH + Pb(OCOMe)_2 \longrightarrow Pb(SR)_2 + 2\,MeCO_2H$$

The reaction accounts for the high toxicity of lead and mercury to living organisms because they react with vital cellular thiol enzymes, thereby poisoning them.

3. Reaction with carbonyl compound: Thiols will undergo nucleophilic addition to aldehydes and ketones, whereas alcohols only react with aldehydes. Thiols also react with acid chlorides to yield esters. The thioacetals (R = H) and thioketals can be used for the protection of aldehydes and ketones since on treatment with dilute acid they are converted back to the original carbonyl substrates. They may also be applied in the conversion of a carbonyl to a methylene group by reaction with Raney nickel which is a common procedure of desulfurization.

$$R_2C{=}O + R'SH \xrightarrow{H^{\oplus}} R_2C(SR')(OH) \xrightarrow[H^{\oplus}]{R'SH} R_2C(SR')_2 \xrightarrow{\text{Raney Ni}} R'_2CH_2$$

$$R_2C(SR')_2 \xrightarrow{\text{dilute acid}} R_2C{=}O$$

4. Oxidation of thiols: The oxidation of thiols follows a completely different course as compared with the oxidation of alcohols, because the capacity of the sulfur atom to form hypervalent compounds allows it to become the site of

oxidation. Thiols are readily oxidised to disulfides by mild oxidants such as atmospheric oxygen, halogens or iron (III) salts. This type of reaction is unique to thiols and is not undergone by alcohols, it is a consequence of the lower bond strength of the S-H as compared with the O-H bond, so that thiols are oxidised at the weaker S-H bonds, whereas alcohols are preferentially oxidised at the weaker C-H bonds. The mechanism of oxidation of thiols may be either radical or polar or both. The polar mechanism probably involves transient sulfenic acid intermediates.

$$2\,RSH \xrightarrow[\text{halogens or iron (III) salts}]{\text{mild oxidation by oxygen,}} RSSR$$

The mechanism of above reaction could be understood by either by free radical mechanism or ionic mechanism which involves following steps:

(a) Radical mechanism:

$$RSH + [O] \longrightarrow RS^{\bullet} + H[O]$$

$$2\,RS^{\bullet} \xrightarrow{\text{dimerisation}} RSSR$$

(b) Ionic mechanism:

$$RSH + [O] \longrightarrow RSOH \xrightarrow{RS^-} RSSR + OH^-$$

$$RSH + X_2 \longrightarrow RSX \xrightarrow{RSH} RSSR + HX$$

In contrast, thiols react with more powerful oxidants, like potassium permanganate, concentrated nitric acid or hydrogen peroxide, to yield the corresponding sulfonic acids.

$$R-CH_2-OH + [O] \xrightarrow[KMnO_4]{oxidant} \left[R-CH_2-O-H \right] \xrightarrow{-H_2O} R-CHO$$

Primary alcohol

$$R-CH(R')-OH + [O] \xrightarrow[KMnO_4]{oxidant} \left[R-C(R')(OH)-O-H \right] \xrightarrow{-H_2O} R-CO-R'$$

Secondary alcohol

5. Reaction with acid and their derivatives: Thiols react with acids, acid chlorides, acid anhydrides, etc. and forms thioesters.

$$RSH + R'-CO-OH \underset{}{\overset{H^{\oplus}}{\rightleftharpoons}} R'-CO-S-R + H_2O$$

$$RSH + R'-CO-Cl \longrightarrow R'-CO-S-R + HCl$$

$$RSH + (R'-CO)_2O \xrightarrow{\triangle} R'-CO-S-R + R'COOH$$

6. Addition to carbon-carbon double bonds: In the presence of an acid, thiols get added to a carbon-carbon double bond and gives thioethers.

$$R-CH=CH_2 + R'-SH \xrightarrow{H^{\oplus}} R-CH(SR')-CH_3$$

3.5 Biological importance of Thiols

As the functional group of the amino acid cysteine, the thiol group plays an important role in biological systems. When the thiol groups of two cysteine residues (as in monomers or constituent units) are brought near each other in the course of protein folding, an oxidation reaction can create a cystine unit with a disulfide bond (-S-S-). Disulfide bonds can contribute to a protein's tertiary structure if the cysteines are part of the same peptide chain, or contribute to the quaternary structure of multi-unit proteins by forming fairly strong covalent bonds between different peptide chains. The heavy and light chains of antibodies are held together by disulfide bridges. Also, the kinks in curly hair are a product of cystine formation. Permanents take advantage of the oxidizability of cysteine residues. The chemicals used in hair straightening are reductants that reduce cystine disulfide bridges to free cysteine sulfhydryl groups, while chemicals used in hair curling are oxidants that oxidize cysteine sulfhydryl groups to form cystine disulfide bridges. Sulfhydryl groups in the active site of an enzyme can form noncovalent bonds with the enzyme's substrate as well, contributing to catalytic activity. Active site cysteine residues are the functional unit in cysteine proteases.

3.6 Methods of formation and chemical reactions of Thioethers

The sulfur analogs of ethers are commonly called sulfides by chemists. A thioether is a functional group in organosulfur chemistry with the connectivity C-S-C. Like many other sulfur-containing compounds, volatile thioethers have foul odors. A thioether is similar to an ether except that it contains a sulfur atom in place of the oxygen. The grouping of oxygen and sulfur in the periodic table suggests that the chemical properties of ethers and thioethers are somewhat similar. The general structure of a thioether may be represented as follows:

$$R^1{-}S{-}R^2$$

Nomenclature of thioethers:

Thioethers or sulfides are analogues of ethers. They have the general formula R-S-R’ where R and R’ are alkyl groups which may be different or same. The functional group in the thioethers is –S. The common name of a sulphide is constructed by naming the two alkyl groups on the sulphide sulphur, followed

by the word sulphide or thioether. In the IUPAC system sulphides are named as alkylthioalkanes.

$H_3C—S—CH_3$
Dimethyl sulfide

SCH_3
methyl(phenyl)sulfide

SCH_2CH_3
ethyl(4-methylpent-3-en-2-yl)sulfide

Formation of thioethers:

1. From alkyl hailides: Sulfides may be prepared by a method analogous to the Williamson ether synthesis. The nucleophile is a thiolate anion rather than an alkoxide. Thiolate ions, RS– , are better nucleophiles than alkoxides because sulfur is more polarizable than oxygen. Thus, thiolate ions displace halide ions from alkyl halides by an SN2 reaction to give good yields of sulfides.

$$H_3C—\overset{\ominus}{S}—\overset{\oplus}{Na} + CH_3CH_2—Br \longrightarrow CH_3SCH2CH_3 + NaBr$$

There are two important differences between reactions to form ethers and those that form sulfides.

1. Because thiolates are better nucleophiles and weaker bases than alkoxides, elimination reactions do not compete much with substitution reactions. Even secondary alkyl halides can be used to form sulfides.

2. Second, because thiols are more acidic (pKa = 8) than alcohols (pKa = 16), they are quantitatively converted to thiolates by sodium hydroxide.

Therefore, it is not necessary to prepare the thiolate in a separate reaction with a strong base, as is required in the reaction of alcohols with sodium hydride. Sulfides are usually prepared by adding the alkyl halide to a basic alcoholic solution of the thiol.

2. From Grignard reagent: Grignard reagents react with disulphides to form sulphides.

$$R—MgX + R'—S—S—R' \xrightarrow{\text{ether}} R—S—R' + R'\overset{\ominus}{S}\overset{\oplus}{Mg}X$$

This reaction can be used to synthesize sulphides that are easily prepared by the SN^2 reaction.

(a) C_6H_5MgBr + (2-methylphenyl)–S–S–(2-methylphenyl) $\xrightarrow{THF}$ C_6H_5–S–(2-methylphenyl)

(b) H_3C–C(CH_3)$_2$–MgBr + C_6H_5–S–S–C_6H_5 $\xrightarrow{THF}$ H_3C–C(CH_3)$_2$–S–C_6H_5

3. From ethers: Ethers when heated with phosphorus pentasulphide form sulphides.

$$5\,C_2H_5—O—C_2H_5 + P_2S_5 \longrightarrow 5\,C_2H_5—S—C_2H_5 + P_2O_5$$

4. From aryl diazonium salts: Aryl diazonium salts on treatment with K_2S or RSNa form aryl and aryl alkyl thioethers.

$$2\,C_6H_5—\overset{\oplus}{N}\equiv N\overset{\ominus}{X} + K_2SH \longrightarrow C_6H_5—S—C_6H_6 + N_2 + 2KX$$

$$C_6H_5—\overset{\oplus}{N}\equiv N\overset{\ominus}{X} + R—\overset{\ominus}{S}\overset{\oplus}{Na} \longrightarrow C_6H_5—S—R + NaX + N_2$$

5. From thioalcohols: Thioethers are formed when vapours of thioalcohols are passed over heated catalyst at 300°C.

$$2\,C_2H_5SH \xrightarrow[300^{o}C]{Al_2O_3/ZnS} C_2H_5—S—C_2H_5 + H_2S$$

Chemical reactions: Thioethers are more reactive than their oxygen analogs, but have a chemistry that is less complex than that of the thiols. As a result, they are of lesser biochemical significance than their thiol cousins. Despite this, certain individual thioethers assume great importance. Such compounds include the vitamins, thiamine and biotin, and the penicillin and cephalosporin classes of antibiotics.

A major difference between ethers and thioethers is the ability of the sulfur in the latter to expand its octet and accept electrons into the d orbitals. As a result, whereas a thioether can donate an electron pair to form a sulfonium ion,

analogous to the less stable oxonium ions formed by ethers, the sulfur in thioethers can also stabilize a negative charge on an adjacent position by providing an electron sink. Such a process is not possible with ethers. As a consequence of this phenomenon, the α-hydrogens adjacent to the heteroatom are much more acidic in thioethers than in ethers. A stable carbanion may be formed, or there may be elimination of a group β to the sulfur to form an alkene. Because of the polarizibility of sulfur, in unsaturated thioethers there is an electron shift towards the β-position, resulting in electrophilic additions occurring at this position. Paradoxically, however, thioethers also add nucleophiles at the β-position, although nucleophilic additions in ethers occur at the α-position.

1. Oxidation of thioethers: Oxidation reaction of thioethers involve two steps:

(a) Oxidation of thioethers to sulfoxide: In the first step of the reaction, the sulfur atom attacks the terminal oxygen of the peroxide group. Breakage of the peroxide bond results in the formation of an alkoxy anion and a sulfoxide protonated at the oxygen atom. Proton exchange yields the sulfoxide and the alcohol corresponding to the peroxide.

(b) Oxidation of sulfoxide to sulfone: In acidic or neutral solution, the oxidation follows the same mechanism as that described for the oxidation of sulfide to sulfoxide. In basic solution, however, a nucleophilic attack of the peroxy anion takes place at the sulfur atom.

2. Reaction with alkyl halide: Sulfur is an excellent nucleophile because its electron cloud is polarizable. As a result, a thioether reacts readily with an alkyl halide to form a sulfonium ion, whereas an ether does not do the equivalent

reaction because oxygen is not as nucleophilic as sulfur and cannot accommodate a positive charge as well as sulfur can.

$$H_3C-\ddot{S}-CH_3 + H_3C-I \longrightarrow H_3C-\overset{\oplus}{S}(CH_3)-CH_3 \; I^{\ominus}$$

Dimethyl sulfide (Thioether) — Trimethylsulfonium iodide, Sulfonium ion

The positively charged group of a sulfonium ion is an excellent leaving group, so a sulfonium ion readily undergoes nucleophilic substitution reactions. Like other SN^2 reactions, the reaction works best if the group undergoing nucleophilic attack is a methyl group or a primary alkyl group.

$$:\ddot{O}H^{\ominus} + H_3C-\overset{\oplus}{S}(CH_3)-CH_3 \longrightarrow CH_3-\ddot{O}H \quad H_3C-\ddot{S}-CH_3$$

Dimethyl sulfide

3. Addition of halogens: Thioethers form addition products of halogens.

$$R-\ddot{S}-R + X_2 \longrightarrow R_2SX_2$$

$$R_2\ddot{S} + Br-Br \longrightarrow R_2\overset{\oplus}{S}-Br + Br^{\ominus} \longrightarrow R_2SBr_2$$

4. Addition of metal salts: Thioethers form an insoluble complex with mercuric chloride. Sulfides bind to metals to form coordination complexes.

$$R-\ddot{S}-R + HgCl_2 \longrightarrow R_2\overset{\oplus}{S}-\overset{\ominus}{Hg}Cl_2$$

5. Hydrolysis: When boiled with alkali, thioethers are hydrolyzed to alcohols.

$$C_2H_5-\ddot{\underset{..}{S}}-C_2H_5 \quad + \quad H_2O \xrightarrow[\Delta]{\overset{\ominus}{OH}} 2C_2H_5OH \quad + \quad H_2S$$

6. Hydrogenolysis: Hydrogenolysis leads to desulphurization and occur when thioethers are heated with hydrogen in presence of Raney nickel.

$$R-S-R' \; + \; 2H_2 \longrightarrow RH \; + \; R'H \; + \; H_2S$$

3.7 Methods of formation and chemical reactions of Sulphonic acids

Organic acids containing sulfur and having the general formula RSO_3H, in which R is an organic combining group, are known as sulfonic acids. The sulfonic acids are the most important organosulfur compounds which are widely used as catalysts (free acids) in organic synthesis, while the salts and other derivatives are used for manufacturing of detergents, water-soluble dyes and catalysts, sulfonamide pharmaceuticals, and ion-exchange resins. Sulfonic acid groups can greatly enhance the water solubility of compounds and their metal complexes are used as homogeneous catalysts for the syntheses of organic compounds in two-phase systems (e.g., in a mixture of water and an organic solvent) in industry and research laboratory.

A very few sulfonic acids occur naturally, for example; the essential nutrient taurine, the sulfobacins, sulfonolipids and echinosulfonic acid. The aliphatic sulfonic acids; methanesulfonic acid and trifluoromethanesulfonic acid (triflic acid) are also commercially important reagents and catalysts. Triflic acid is one of the strongest known organic acids and is used as a polymerization catalyst. However, sulfonic acid is also used in fuel cells, in gasoline production, and in the synthesis of organic and organometallic compounds. Aromatic sulfonic acids are obtained generally by treating aromatic compounds with concentrated sulfuric acid and added sulfur trioxide (“oleum”), the process being called sulfonation.

Nomenclature of sulfonic acids: Names of organosulfur acids are analogous to those of cyclic carboxylic acids. The suffix name sulfonic acid is used for the functional groups —SO_2OH. This suffix is used substitutively with names of hydrocarbon and heterocyclic systems.

benzenesulfonic acid   2-hydroxybenzenesulfonic acid   4-methylbenzenesulfonic acid

The corresponding acid groups in which one or more oxygen atoms bound to the central sulfur atom have been replaced by a sulfur atom are named by prefixing the term thio, dithio, or trithio, as appropriate, to the suffix name. When an unambiguous description of structure is desired, the term S-acid (for - SH) or O-acid (for - OH) may be used.

Benzenedithiosulfonic acid   Benzenedithiosulfonic S-acid   Benzenedithiosulfonic O-acid

Formation of sulfonic acid: Several methods are available for the synthesis of aliphatic and aromatic sulfonic acids. However, here are some selected methods are given for the synthesis of sulfonic acids.

1. By sulfonation: Sulfonic acids are strong acids comparable in strength with sulfuric acid; they are stable and do not decompose on heating. Sulfonic acids may be obtained by direct sulfonation of the appropriate aliphatic or aromatic substrate by treatment with sulfuric acid. Direct sulfonation is generally used in the preparation of aromatic sulfonic acids, however, sulfonation in aliphatic compounds is harder than aromatic substrates.

$$RH + H_2SO_4 \rightleftharpoons R-SO_2-OH + H_2O$$

Sulfonation is a bimolecular electrophilic substitution reaction and the electrophile is sulfur trioxide. Sulfur trioxide is a powerful electrophile because of the electron-withdrawing effect of the three double-bonded oxygen atoms. Consequently, oleum (fuming sulfuric acid), which contains approximately 10% of excess sulfur trioxide, is a much more powerful sulfonating agent than concentrated sulfuric acid. Sulfur trioxide is a sufficiently powerful electrophile to attack benzene directly. The mechanism of the sulfonation of benzene by hot concentrated sulfuric acid to give benzenesulfonic acid is shown below.

$$RH + H_2SO_4 \rightleftharpoons R-SO_2-OH + H_2O$$

$$2\,H_2SO_4 \xrightleftharpoons{\text{Step-1}} SO_3 + H_3O^{\oplus} + HSO_4^{\ominus}$$

$$C_6H_6 + SO_3 \underset{\text{(slow)}}{\xrightleftharpoons{\text{Step-2}}} \text{sigma complex } [C_6H_6SO_3^{\ominus}]^{+}$$

sigma complex

Step-3: $HSO_4^{\ominus}$ (fast), $(-H^{\oplus})$

$$C_6H_5SO_3H \underset{(-H_2O)}{\overset{H^{\oplus}\text{ (fast)}}{\rightleftharpoons}} C_6H_5SO_3^{\ominus} + H_3O^{\oplus}$$

2. By Strecker synthesis: Aliphatic or benzylic sulfonic acids can be obtained in good yields by reaction of an alkyl or benzylic halide or an epoxide with sodium sulfite or sodium hydrogen sulfite. The sulfite anion is nucleophilic at

sulfur, which accounts for these reactions which go best with primary alkyl halides.

$$RCH_2X + Na_2SO_3 \xrightarrow[\text{aqueous EtOH}]{\text{Reflux in}} RCH_2SO_3Na + NaX$$

$$PhCH_2X + NaHSO_3 \longrightarrow PhCH_2SO_3H + NaX$$

$$R{-}\underset{H}{C}H(O)CH_2 + HSO_3^{\ominus} \xrightarrow{NaHSO_3} RCH(OH)CH_2SO_3Na$$

The Strecker reaction is also successful with aromatic compounds containing reactive halogen atoms, e.g. 2,4-dinitrochlorobenzene. In this compound, the chlorine atom is activated with respect to nucleophilic attack by the strongly electron-withdrawing (-I, -M) nitro groups.

$$\text{2,4-}(NO_2)_2C_6H_3Cl \xrightarrow[\text{aqueous EtOH}]{Na_2SO_3 \text{ in hot}} \text{2,4-}(NO_2)_2C_6H_3SO_3Na + NaCl$$

3. By oxidation of thiols: Sulfonic acids may also be obtained by oxidation of the appropriate thiols. The oxidants may be halogens, hydrogen peroxide, nitric acid, potassium permanganate or chromic anhydride. An illustrative example is provided by the oxidation of 4-chloro-2-methylbutane-2-thiol to the corresponding sulfonic acid.

$$RSH \xrightarrow[3[O]]{Oxidation} RSO_3H$$

$$ClCH_2CH_2C(Me)_2SH \xrightarrow{H_2O_2,\ HOAc} ClCH_2CH_2C(Me)_2SO_3H$$

The method can also be applied to the conversion of 4-chloropyridine to pyridine-4-sulfonic acid.

$$\text{4-chloropyridine (Cl)} \xrightarrow{NaSH,\ aq.\ EtOH} \text{4-mercaptopyridine (SH)} \xrightarrow{hot\ conc.\ HNO_3} \text{pyridine-4-sulfonic acid } (SO_3H)$$

4. By addition of sodium bisulphate to alkenes: In the presence of oxygen or peroxides, anti-Markovnikov's addition of sodium bisulphate to alkenes and gives the sodium salt of the corresponding alkanesulphonic acid.

$$R{-}CH{=}CH_2 + NaSO_3H \xrightarrow{H_2O_2\ or\ O_2} R{-}SO_3Na + NaX$$

Reactions of sulfonic acid: Due to sulfonic acid derivatives importance, they are prepared by various synthetic routes.

1. Formation of sulfonic acid esters: Sulfonic acids may be used to produce sulfonic acid esters, which are derived from epoxides, olefins, alkynes, allenes, and ketenes.

2. Formation of sulfonamide: The combination of H_2O_2 and $SOCl_2$ is a highly reactive reagent for the direct oxidative conversion of thiol derivatives to the corresponding sulfonyl chlorides through oxidative chlorination. Upon reaction with amines, the corresponding sulfonamides are obtained in excellent yields.

$$R{-}SH \xrightarrow[\text{MeCN, } 25^{o}C]{H_2O_2,\ SOCl_2} [RSO_2Cl] \xrightarrow[\text{Pyridine}]{R'{-}NH_2} RSO_2NHR'$$

3. Hydrolysis of sulfonic acid: Through the hydrolysis of sulfonic acids, benzene can be produced. Benzene sulphonic acid is subjected to superheated steam in this process, resulting in the synthesis of benzene. Benzene sulphonic acid is treated with an aqueous solution of NaOH. This mixture is then mixed with solid NaOH and fused at high temperature. The resulting product is then acidified to yield phenol.

SO_3H $\xrightarrow[H_2O]{NaOH}$ $SO_3^-Na^+$ $\xrightarrow[350^oC]{NaOH}$ O^-Na^+ $\xrightarrow{H_3O^+}$ OH

3.8Methods of formation and chemical reactions of Sulphonamides

Sulfonamides or sulfa drugs are a class of antibiotics that target bacteria causing infections. These classes of drugs are generally broad-spectrum antibiotics that act on a wide range of bacterial types and are therefore employed in treating many kinds of bacterial infections. Sulphonamides do not kill bacteria, but it interferes with the ability of bacteria to grow and multiply (bacteriostatic).

Folic acid is a key component of bacteria that it uses for growth and multiplication. Sulfa drugs block the ability of the bacteria to use folic acid, thereby inhibiting the growth process. As a result, the bacteria fail to reproduce and spread. Because humans do not make folic acid and completely rely on diet for getting their folic acid, they remain safe against the adverse effects of inhibition of folate production.

Sulfonamides are oily liquids or crystalline solids that are almost always prepared by the reaction of a sulfonyl chloride with ammonia or an amine, most commonly in the presence of caustic alkali.

$C_6H_5SO_2Cl + CH_3NH_2 \xrightarrow[-H_2O\ -Cl^{\ominus}]{OH^{\ominus}} C_6H_5SO_2NHCH_3$

Benzenesulfonyl chloride Methanamine N-methylbenzenesulfonamide

The first sulfonamide drug, introduced in 1932, was a red azo dye called Prontosil (*q.v.*). As new sulfonamides were synthesized, more effective and less toxic agents were discovered. Some, which are not absorbed, can be administered orally to treat specific localized infections in the gastrointestinal tract. Others are absorbed slowly or excreted slowly and therefore are longer acting.

Synthesis of sulfonamide

Sulfonamide is synthesised by following routes:

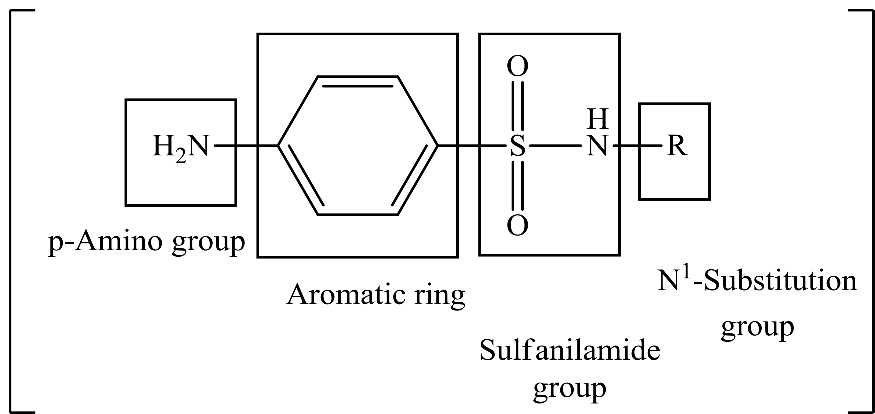
Benzene → (HNO_3/H_2SO_4, Nitration) → Nitrobenzene (NO_2) → (Sn/HCl, [H]) → Aniline (NH_2) → ($(CH3COO)_2O$, Amine protection) → Acetanilide (NHCOCH)

Acetanilide → (Cl—S($=O$)$_2$—OH, $-H_2O$) → $NHCOCH_3$ → ($R-NH_2$, -HCl) → $NHCOCH_3$ → (NaOH, $-CH_3COOH$) → NH_2

Structural Activity Relationship (SAR) of sulphonamide:

H_2N—(4) benzene ring (positions 1–6) (1)—SO_2NHR

The major features of SAR of sulphonamides include the following:



- Sulphanilamide skeleton is the minimum structural requirement for antibacterial activity.
- The amino- and sulphonyl-groups on the benzene ring are essential and should be in 1 and 4 position.
- The N-4 amino group could be modified to be prodrugs, which are converted to free amino function *in vivo.*
- Sulphur atom should be directly linked to the benzene ring.
- Replacement of benzene ring by other ring systems or the introduction of additional substituents on it decreases or abolishes its activity.
- Exchange of the $–SO_2NH$ group by –CONH reduces the activity.
- On N^1-substituted sulphonamides, activity varies with the nature of the substituent at the amino group. With substituents imparting electron-rich characters to SO_2 group, bacteriostatic activity increases.
- Heterocyclic substituents lead to highly potent derivatives, while sulphonamides, which contain a single benzene ring at N-1 position, are considerably more toxic than heterocyclic ring analogues.
- The free aromatic amino groups should reside para to the sulphonamide group. Its replacement at ortho or meta position results in compounds devoid of antibacterial activity.
- The active form of sulphonamide is the ionized, maximum activity that is observed between the pKa values 6.6-7.4.

- Substitutions in the benzene ring of sulphonamides produced inactive compounds.
- Substitution of free sulphonic acid (–SO_3H) group for sulphonamido function destroys the activity, but replacement by a sulphinic acid group (–SO_2H) and acetylation of N-4 position retains back the activity.
- Meta-Sulphonamides bind to the basic centres of arginine, histidine, and lysine sites of proteins. The binding groups are alkyl, alkoxy, and halides. The binding affects the activity of sulphonamides; protein binding appears to modulate the availability of the drug and its half-life.
- The lipid solubility influences the pharmacokinetic and antibacterial activity, and so increases the half-life and antibacterial activity in vitro.

Classification of sulphonamides

1. On the basis of the site of action

(a) Sulphonamides for general infection: Sulphanilamide, Sulphapyridine, Sulphadiazine, Sulphamethoxacine, Sulphamethoxazole.

(b) Sulphonamides for urinary tract infections: Sulphaisoxazole, Sulphathiazole.

(c) Sulphonamides for intestinal infections: Phthalylsulphathiazole, Succinyl sulphathiazole, Sulphasalazine.

(d) Sulphonamides for local infections: Sulpahacetamide, Mafenamide, Silver sulphadiazine.

(e) Sulphonamides for dermatitis: Dapsone, Solapsone.

(f) Sulphonamides in combination: Trimethoprim with Sulphamethoxazole.

2. On the basis of the pharmacokinetic properties

(a) Poorly absorbed sulphonamides (locally acting sulphonamides): Sulphasalazine, Phthalylsulphathiazole, Sulphaguanidine, Salicylazo sulphapyridine, Succinyl sulpha thiazole.

(b) Rapidly absorbed and rapidly excreted (systemic sulphanamides): Sulphamethoxazole, Sulphaisoxazole, Sulphadiazine, Sulphadimidine,

Sulphafurazole, Sulphasomidine, Sulphamethiazole, Sulphacetamide Sulphachlorpyridazine.

(c) Topically used sulphonamides: Sulphacetamide, Mafenide, Sulphathiazole, Silver sulphadiazine.

3. On the basis of the pharmacological activity

(a) Antibacterial agents: Sulphadiazine, Sulfi soxazole.

(b) Drugs used in dermatitis: Dapsone.

4. On the basis of the duration of action

(a) Extra-long-acting sulphonamides (half-life greater than 50 h): Sulphasalazine, Sulphaclomide, Sulphalene.

(b) Long-acting sulphonamides (half-life greater than 24 h): Sulphadoxine, Sulphadimethoxine, Sulphamethoxy pyridazine, Sulphamethoxydiazine, Sulphaphenazole, Sulphamethoxine.

(c) Intermediate-acting sulphonamides (half-life between 10-24 h): Sulphasomizole, Sulphamethoxazole.

(d) Short-acting sulphonamides (half-life less than 20 h): Sulphamethiazole, sulphaisoxazole.

(e) Injectable (soluble sulpha drugs): Sulphafurazole, Sulphadiazine, Sulphamethoxine.

5. On the basis of the chemical structure

(a) N-1 substituted sulphonamide: Sulphadiazine, Sulphacetamide, Sulphadimidine.

(b) N-4 substituted sulphonamides (prodrugs): Prontosil.

(c) Both N-1 and N-4 substituted sulphonamides: Succinyl sulphathiazole, Phthalylsulphathiazole.

(d) Miscellaneous: Mefenide sodium.

Sulfonamide derivatives:

N-1 Substituted sulphonamides

Name	R	R[1]
Sulphanilamide	–H	–H
Sulphacetamide	–H	$-COCH_3$
Sulphadiazine	–H	N, N
Sulphamethoxazole	–H	CH_3, O, N

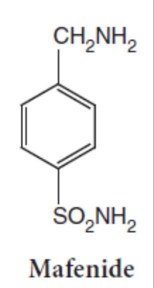
N-4 substituted suphonamides

H_2N–N=N–SO_2NH_2, NH_2

Prontosil

N-1 and N-4 substituted suphonamides

Name	R	R_1
Succinyl sulphathiazole	O, C, CH_2, HO–C–CH_2, O	S, CH, C, N, CH
Phthalylsulphathiazole	O, C, HOOC	S, N



Mafenide

H_2N–SO_2–N–, Ag, N, N

Silver sulphadiazine

H_3CO, CH_2, NH_2, N, N, NH_2, H_3CO, OCH_3

Trimethoprim

Adverse reactions:

(a) The most common manifestation of a hypersensitivity reaction to sulfa drugs are rash and hives. However, there are several life-threatening manifestations of hypersensitivity to sulfa drugs, including Stevens-Johnson syndrome, toxic epidermal necrolysis, agranulocytosis, hemolytic anemia, thrombocytopenia, and fulminant hepatic necrosis, among others.

(b) The sulfonamide antibiotic chemical structures are implicated in the hypersensitivity reactions associated with the class.

➢ The first is the N1 heterocyclic ring, which causes a type I hypersensitivity reaction.

➢ The second is the N4 amino nitrogen that, in a stereospecific process, forms reactive metabolites that cause either direct cytotoxicity or immunologic response.

3.8 Methods of formation and chemical reactions of Sulphaguamidine

Sulfaguanidine is a guanidine derivative of sulfanilamide used in veterinary medicine. Sulfaguanidine is poorly absorbed from the gut but is well suited for the treatment of bacillary dysentery and other enteric infections. Sulphaguanidine was independently prepared by Marshall, Bratton and White in 1940, and introduced for the treatment of bacillary dysentery on the basis of its poor absorption from the gut. IUPAC name of Sulfaguanidine is 2-(4-aminophenyl) sulfonylguanidine.

Synthesis of sulfaguanidine:

NO_2 NH_2 $NHCOCH_3$

HNO_3/H_2SO_4 Nitration; Sn/HCl [H]; $(CH3COO)_2O$ Amine protection

Benzene Nitrobenzene Aniline Acetanilide

O Cl—S—OH O $-H_2O$

NH_2 H_2N NH_2 NH NH_2 $NHCOCH_3$ $NHCOCH_3$

NaOH $-CH_3COOH$; R-NH_2 -HCl

O=S=O SO_2NH-R SO_2NH-R SO_2Cl

N Sulfanilamide

H_2N NH_2

Sulfaguadine

Sulfaguanidine action: Sulfaguanidine is a sulfonamide with properties similar to those of sulfamethoxazole. Sulfonamides have a similar structure to para-aminobenzoic acid (PABA) and interfere with the synthesis of nucleic acid in

sensitive microorganisms by blocking the coversion of PABA to the co-enzyme dihydrofolic acid, a reduced form of folic acid. Their actions are primarily bacteriostatic, although bactericidal effects are exerted where concentrations of thymine are low in the surrounding medium. Sulfonamides are active against Gram positive bacteria and Gram-negative bacteria. Mode of resistance is via the alteration of dihydropteroate synthase or alternative pathway for folic acid synthesis.

3.11 Summary

In this Unit of Sulphur Containing Compounds learners are able to define basics of Nomenclature, structural information and Methods of formation as well as chemical reactions of thiols, thioethers, sulphonic acids, sulphonamides and Sulphaguamidine. Also able to define application and importance of Sulphur Containing Compounds for our daily life.

3.12 Terminal Questions

Q1. Oxidation states of sulfur organic compounds?

Ans. A molecule containing one or more carbon-sulfur bonds is known as organosulfur compounds.

Oxidation states of sulphur in organic compounds					
-2	-1	0	+2	+4	+6
H_2S Hydrogen sulfide R–S–H Thiols R–S–R Sulfides R_3S^+ Sulfonium ions	R–S–S–R Disulfides	S Elemental sulphur R–S(=O)–R Sulfoxides R–S–OH Sulfenic acids	R–S(=O)₂–R Sulfones R–S(=O)–OH Sulfinic acids	SO_2 Sulphur dioxide R–S(=O)₂–OH Sulfonic acids R–O–S(=O)–O–R Sulfite esters	SO_3 Sulphur trioxide R–O–S(=O)₂–O–R Sulfate esters

Q2. Methods of formation and chemical reactions of thiols?

Ans.

1. Formation from alkanes

(a) Hydrogen sulphide additions: This is in principle the simplest process for preparing alkanethiols. Usually the use of a high hydrogen sulphide ratio favours the formation of the thiol. Two isomeric thiols are formed from an unsymmetric alkene.

$$RCH{=}CH_2 + H_2S \longrightarrow \underset{(1)}{RCH(SH)CH_3} + \underset{(2)}{RCH_2CH_2SH}$$

Where; product 1 is termed the Markownikoff product and 2, the anti-Markownikoff adduct.

(b) Additions of other suphur acids: Additions of other sulphur acids to alkenes are frequently used in thiol preparations; the required thiol being obtained from the intermediate compounds after hydrolysis or dealkylation.

$$R^1CH{=}CHR^2 + R^3SH \longrightarrow R^1CH(SR^3)CH_2R^2 \longrightarrow R^1CH(SH)CH_2R^2$$

Where, R^3= alkyl, aryl or acyl

2. Formation from alcohols

(a) Using hydrogen sulphide: Direct reaction between hydrogen sulphide and alcohols normally requires the presence of a catalyst. Several processes involve basic catalysts, high temperatures and high pressures.

$$ROH + H_2S \xrightarrow{\text{Catalyst}} RSH + H_2O$$

(b) Using phosphorus pentasulphide: C_4-C_{16} alcohols can be converted to dialkyldithiophosphates, which on acid hydrolysis give the corresponding thiols. Yields greater than 70% can be obtained if the sulphides formed in the reaction are dealkylated to give thiols as well.

1*H*-imidazo[4,5-*d*]pyridazine-4,7-diol —(P_2S_5 Pyridine, Heat)→ 1*H*-imidazo[4,5-*d*]pyridazine-4,7-dithiol

3,5-diiodopyridin-2(1*H*)-one —(P_2S_5 Pyridine, Heat)→ 3,5-diiodopyridine-2-thiol

3. Formation from halides: When alkyl halides are heated with an alcoholic solution of sodium or potassium hydrogen sulphide; thioalcohols are formed.

$$RX + KSH \xrightarrow{\text{Heat}} RSH + KX$$

$$C_2H_5Cl + NaSH \xrightarrow{\text{Heat}} C_2H_5SH + KX$$

4. Formation from thiourea: This method, generally involves the reaction of a halide with thiourea to give an iso-thiouronium salt is formed which on treatment with sodium hydroxide gives thiol.

$$R-X + S=C(NH_2)_2 \longrightarrow R-S-C(=\overset{\oplus}{N}H_2)NH_2\ \overset{\ominus}{X} \longrightarrow RSH$$

5. Formation from quinones: Reaction of p-benzoquinones with thiosulphate led to 1,4-dihydroxyphenyl thiosulphates, which on reduction with zinc and hydrochloric acid gave the mercaptodihydroxybenzenes.

$$\text{R-p-benzoquinone} \xrightarrow{S_2O_3^{2-}} \text{R-2,5-dihydroxyphenyl-O-S(=O)(=S)-O}^{\ominus} \xrightarrow{Zn/H_3O^{\oplus}} \text{R-2,5-dihydroxy-C}_6\text{H}_2\text{-SH}$$

R= H, Me and Ph

6. Formation from Grignard reagent: When sulphur is treated with Grignard reagent, thiocomplexes are formed which on acid hydrolysis gives thiol.

$$S_8 + R-MgX \longrightarrow R-S-MgX \xrightarrow{HOH/H^{\oplus}} R-S-H$$

Q3. Chemical reactions of thiols?

1. Acidic character: The sulfur atom of a thiol is quite nucleophilic, rather more than that of oxygen atom of alcohols. The thiol group is fairly acidic with a usual pK_a around 10 to 11. In the presence of a base, a thiolate anion is formed which is a very powerful nucleophile. The group and its corresponding anion are readily oxidized by reagents such as bromine to give an organic disulfide (R-S-S-R). Oxidation by more powerful reagents, such as sodium hypochlorite or hydrogen peroxide yields sulfonic acids (RSO_3H).

$$2\,R-SH + Br_2 \longrightarrow RSSR + 2HBr$$

$$R-SH + 3\,H_2O_2 \longrightarrow RSO_3H + 3H_2O$$

Acidity is described as the ability to lose H^+ ion in solution. In case of thiols (R-SH) after losing H^+ ions; negative charge is generated on sulphur which is more stabilized. So, if the resulting state of compound is more stable, its activity to reach there is faster and hence thiols lose H^+ ions easily as compared to alcohols for which negative charge on oxygen is unstable and will not let H^+ go that easily. Hence, thiols are more acidic in nature than alcohols.

2. Reaction with alkali metals: Thiols, unlike alcohols, form insoluble salts (mercaptides) by reaction with heavy metals like mercury and lead. The former name 'mercaptan' for thiols comes from the Latin mercurium captans which means mercury seizing. In modern nomenclature the name 'thiol' is preferred over 'mercaptan', although the prefix 'mercapto' is still allowed for the unsubstituted -SH radical.

$$2\,RSH + HgO \longrightarrow Hg\,(SR)_2 + H_2O$$

$$2\,RSH + Pb(OCOMe)_2 \longrightarrow Pb\,(SR)_2 + 2\,MeCO_2H$$

The reaction accounts for the high toxicity of lead and mercury to living organisms because they react with vital cellular thiol enzymes, thereby poisoning them.

3. Reaction with carbonyl compound: Thiols will undergo nucleophilic addition to aldehydes and ketones, whereas alcohols only react with aldehydes. Thiols also react with acid chlorides to yield esters. The thioacetals (R = H) and thioketals can be used for the protection of aldehydes and ketones since on treatment with dilute acid they are converted back to the original carbonyl substrates. They may also be applied in the conversion of a carbonyl to a methylene group by reaction with Raney nickel which is a common procedure of desulfurization.

$$R_2C{=}O + R'SH \xrightarrow{H^{\oplus}} R_2C(SR')(OH) \xrightarrow[H^{\oplus}]{R'SH} R_2C(SR')_2 \xrightarrow{\text{Raney Ni}} R'_2CH_2$$

$$R_2C(SR')_2 \xrightarrow{\text{dilute acid}} R_2C{=}O$$

4. Oxidation of thiols: The oxidation of thiols follows a completely different course as compared with the oxidation of alcohols, because the capacity of the sulfur atom to form hypervalent compounds allows it to become the site of

oxidation. Thiols are readily oxidised to disulfides by mild oxidants such as atmospheric oxygen, halogens or iron (III) salts. This type of reaction is unique to thiols and is not undergone by alcohols, it is a consequence of the lower bond strength of the S-H as compared with the O-H bond, so that thiols are oxidised at the weaker S-H bonds, whereas alcohols are preferentially oxidised at the weaker C-H bonds. The mechanism of oxidation of thiols may be either radical or polar or both. The polar mechanism probably involves transient sulfenic acid intermediates.

$$2\,RSH \xrightarrow[\text{halogens or iron (III) salts}]{\text{mild oxidation by oxygen,}} RSSR$$

The mechanism of above reaction could be understood by either by free radical mechanism or ionic mechanism which involves following steps:

(a) Radical mechanism:

$$RSH + [O] \longrightarrow RS^{\bullet} + H[O]$$

$$2\,RS^{\bullet} \xrightarrow{\text{dimerisation}} RSSR$$

(b) Ionic mechanism:

$$RSH + [O] \longrightarrow RSOH \xrightarrow{RS^-} RSSR + OH^-$$

$$RSH + X_2 \longrightarrow RSX \xrightarrow{RSH} RSSR + HX$$

In contrast, thiols react with more powerful oxidants, like potassium permanganate, concentrated nitric acid or hydrogen peroxide, to yield the corresponding sulfonic acids.

$$R-CH_2-OH + [O] \xrightarrow[KMnO_4]{oxidant} \left[R-\underset{H}{\overset{H}{C}}-O-H \right] \xrightarrow{-H_2O} R-\overset{O}{\overset{\|}{C}}-H$$

Primary alcohol

$$R-\underset{H}{\overset{R'}{C}}-OH + [O] \xrightarrow[KMnO_4]{oxidant} \left[R-\underset{OH}{\overset{R'}{C}}-O-H \right] \xrightarrow{-H_2O} R-\overset{O}{\overset{\|}{C}}-R'$$

Secondary alcohol

5. Reaction with acid and their derivatives: Thiols react with acids, acid chlorides, acid anhydrides, etc. and forms thioesters.

$$RSH + R'-\overset{O}{\overset{\|}{C}}-OH \underset{}{\overset{H^{\oplus}}{\rightleftharpoons}} R'-\overset{O}{\overset{\|}{C}}-S-R + H_2O$$

$$RSH + R'-\overset{O}{\overset{\|}{C}}-Cl \longrightarrow R'-\overset{O}{\overset{\|}{C}}-S-R + HCl$$

$$RSH + (R'CO)_2O \xrightarrow{\triangle} R'-\overset{O}{\overset{\|}{C}}-S-R + R'COOH$$

6. Addition to carbon-carbon double bonds: In the presence of an acid, thiols get added to a carbon-carbon double bond and gives thioethers.

$$R-\underset{H}{C}=CH_2 + R'-SH \xrightarrow{H^{\oplus}} R-\underset{SR'}{\underset{|}{CH}}-CH_3$$

Q4. Biological importance of Thiols?

1. **On the basis of the site of action**

(a) Sulphonamides for general infection: Sulphanilamide, Sulphapyridine, Sulphadiazine, Sulphamethoxacine, Sulphamethoxazole.

(b) Sulphonamides for urinary tract infections: Sulphaisoxazole, Sulphathiazole.

(c) Sulphonamides for intestinal infections: Phthalylsulphathiazole, Succinyl sulphathiazole, Sulphasalazine.

(d) Sulphonamides for local infections: Sulpahacetamide, Mafenamide, Silver sulphadiazine.

(e) Sulphonamides for dermatitis: Dapsone, Solapsone.

(f) Sulphonamides in combination: Trimethoprim with Sulphamethoxazole.

2. On the basis of the pharmacokinetic properties

(a) Poorly absorbed sulphonamides (locally acting sulphonamides): Sulphasalazine, Phthalylsulphathiazole, Sulphaguanidine, Salicylazo sulphapyridine, Succinyl sulpha thiazole.

(b) Rapidly absorbed and rapidly excreted (systemic sulphanamides): Sulphamethoxazole, Sulphaisoxazole, Sulphadiazine, Sulphadimidine, Sulphafurazole, Sulphasomidine, Sulphamethiazole, Sulphacetamide Sulphachlorpyridazine.

(c) Topically used sulphonamides: Sulphacetamide, Mafenide, Sulphathiazole, Silver sulphadiazine.

3. On the basis of the pharmacological activity

(a) Antibacterial agents: Sulphadiazine, Sulfi soxazole.

(b) Drugs used in dermatitis: Dapsone.

4. On the basis of the duration of action

(a) Extra-long-acting sulphonamides (half-life greater than 50 h): Sulphasalazine, Sulphaclomide, Sulphalene.

(b) Long-acting sulphonamides (half-life greater than 24 h): Sulphadoxine, Sulphadimethoxine, Sulphamethoxy pyridazine, Sulphamethoxydiazine, Sulphaphenazole, Sulphamethoxine.

(c) Intermediate-acting sulphonamides (half-life between 10-24 h): Sulphasomizole, Sulphamethoxazole.

(d) Short-acting sulphonamides (half-life less than 20 h): Sulphamethiazole, sulphaisoxazole.

(e) Injectable (soluble sulpha drugs): Sulphafurazole, Sulphadiazine, Sulphamethoxine.

5. On the basis of the chemical structure

(a) N-1 substituted sulphonamide: Sulphadiazine, Sulphacetamide, Sulphadimidine.

(b) N-4 substituted sulphonamides (prodrugs): Prontosil.

(c) Both N-1 and N-4 substituted sulphonamides: Succinyl sulphathiazole, Phthalylsulphathiazole.

(d) Miscellaneous: Mefenide sodium.

Unit 4: Amino Acids, Peptides, Proteins and Nucleic Acids

4.1 Introduction

Objective

4.1. 2 Classification, structure and stereochemistry of amino acids

4.1.3 Chemical properties of amino acids

4.1.5 Functions of Amino acids

4.1.6 Stereochemistry of amino acids

4.1.7 Spectral analysis of amino acids

4.1.8 Acid-base behavior of Amino acids

4.1.9 Isoelectric Point (pI)

4.1.10 Electrophoresis

4.1.11 Preparation of α-amino acid

4.1.12 Reactions of α–amino acids

4.1.13 Reagents used for the identification of Amino acids

4.1.14 Determination of the structures of Polypeptide

4.1.15 Sequencing the Peptide-Terminal Residue Analysis

4.1.16 Peptide Synthesis

4.2 Introduction of Proteins

4.2.1 Classification of proteins

4.2.2 Difference between polypeptides and proteins

4.2.3 Structure of proteins

4.3 Introduction of Nucleic Acids

4.3.1 Types of Nucleic Acids

4.3.2 Difference between nucleoside and nucleotide

4.3.3 Components of nucleic acid

4.3.4 Primary Structure of Nucleic Acids

4.3.5 Secondary Structure of Nucleic Acids

4.4 Summary

4.5 Terminal Questions

4.1 Introduction

In this unit learner able to classify, structure and stereochemistry of amino acids. Its acid-base behaviour, Isoelectric point and electrophoresis, Preparation and reactions with α–amino acids. Further learners able to give Structure and nomenclature of peptides and proteins. Classification of proteins, Peptide structure determination, end group analysis, selective hydrolysis of peptides. Classical peptide synthesis, solid-phase peptide synthesis. Structures of peptides and proteins. Levels of protein structure, Protein denaturation/renaturation. With all this learner came to describe Nucleic acids, Introduction, Constituents of nucleic acids. Ribonucleosides and ribonucleotides. The double helical structure of DNA.

Objective

Learner able to define

- Classification, structure and stereochemistry of amino acids.
- Acid-base behaviour, Isoelectric point and electrophoresis.
- Preparation and reactions of α–amino acids.
- Structure and nomenclature of peptides and proteins.
- Classification of proteins, Peptide structure determination.
- End group analysis, selective hydrolysis of peptides.
- Classical peptide synthesis, solid-phase peptide synthesis.
- Structures of peptides and proteins.
- Levels of protein structure, Protein denaturation/renaturation.
- Nucleic acids introduction.
- Constituents of ncleic acids. Ribonucleosides and ribonucleotides.
- The double helical structure of DNA.

Amino acid

Amino acids are organic molecules which on linking together with other amino acids forms protein molecule. Amino acids are essential to life because the proteins they form are involved in virtually all cell functions. Some proteins function as enzymes, antibodies, while others provide structural support. Although there are hundreds of amino acids found in nature, only about 20 amino acids are needed to make all the proteins found in the human body.

Main highlights of amino acids

- Almost all cell functions involve proteins. These proteins are composed of organic molecules called amino acids.

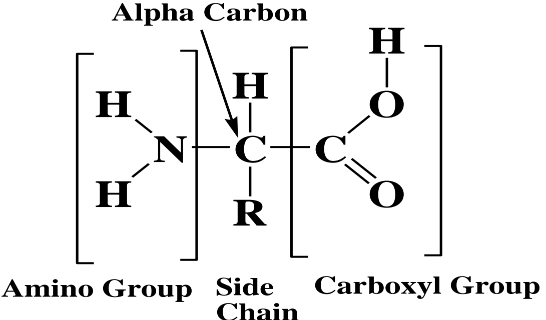
- While there are many different amino acids in nature, all proteins are formed from twenty amino acids.

- From a structural perspective, amino acids are typically composed of a carbon atom, a hydrogen atom, a carboxyl group along with an amino group and a variable group.

- Based on the variable group, amino acids can be classified into four major categories: nonpolar, polar, negatively charged, and positively charged amino acids.

Out of twenty amino acids, eleven can be made naturally by the body and are termed nonessential amino acids. Those amino acids which can not be naturally synthesized by our body are called essential amino acids.

An amino acid structure:



Structural explanation of amino acids

Generally, amino acids have the following structural properties:

- A carbon (C, essentially alpha carbon)
- A hydrogen atom (H)
- A Carboxyl group (-COOH)
- An Amino group (-NH_2)
- A "variable" group or "R" group

All amino acids have the alpha carbon bonded to a hydrogen atom, carboxyl group, and amino group. The "R" group varies among amino acids and determines the differences between these protein monomers. The amino acid

sequence of a protein is determined by the information found in the cellular genetic code. The genetic code is the sequence of nucleotide bases in nucleic acids (DNA and RNA) which codes for amino acids. These gene codes not only determine the order of amino acids in a protein, but they also determine a protein's structure and function. There are some exceptions also:

- Glycine does not have a side chain and its α-carbon contains two hydrogen atoms.
- Proline, in which the nitrogen is part of a ring.

Thus, each amino acid has an amine group at one end and an acid group at the other and a distinctive side chain. The backbone is the same for all amino acids while the side chain differs from one amino acid to the next.

Classification of amino acids on the basis of R-groups

1. **Nonpolar, aliphatic amino acids**: The R groups in this class of amino acids are nonpolar and hydrophobic. Examples are: Glycine, Alanine, Valine, Leucine, Isoleucine, Methionine and Proline.

Glycine Alanine Valine

Leucine Methionine Isoleucine

2. **Aromatic amino acids:** Phenylalanine, Tyrosine and Tryptophan, with their aromatic side chains, are relatively nonpolar (hydrophobic). All can participate in hydrophobic interactions.

Phenylalanine tyrosine tryptophan

3. **Polar, uncharged amino acids:** The R groups of these amino acids are more soluble in water, or more hydrophilic, than those of the nonpolar amino acids, because they contain functional groups that form hydrogen bonds with water. This class of amino acids includes Serine, Threonine, Cysteine, Proline, Asparagine and Glutamine.

serine threonine cysteine

proline asparagine glutamine

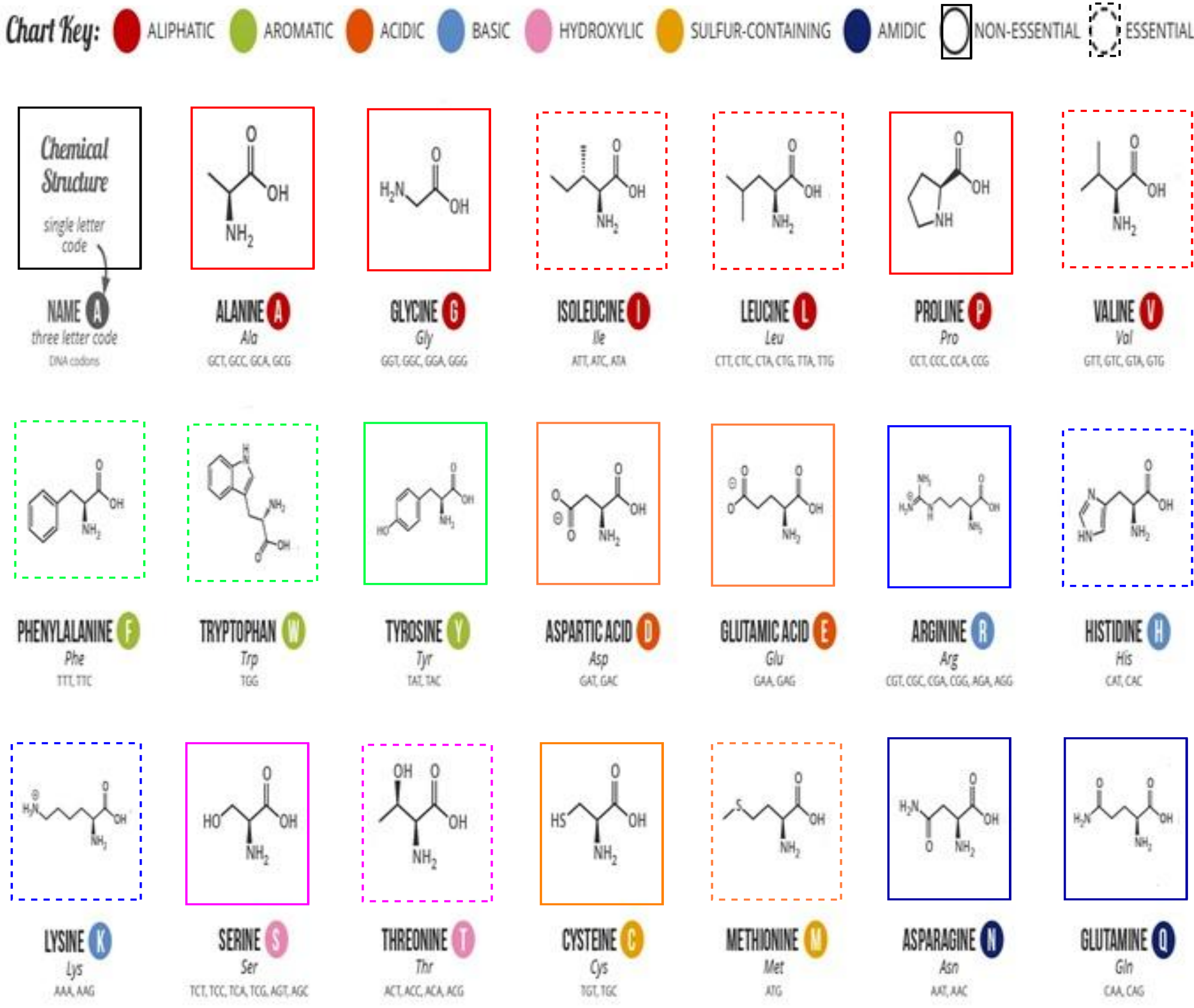
4. **Acidic amino acids:** Amino acids in which R-group is acidic or negatively charged. Examples are- Glutamic acid and Aspartic acid

Glutamic acid Aspartic acid

5. **Basic amino acids:** Amino acids in which R-group is basic or positively charged. Examples are- Lysine, Arginine, Histidine.

Lysine Arginine Histidine

A summary of all the amino acids with respect to their property and structure can be represented by following chart.



Properties of amino acids

Physical properties

- Amino acids are colorless, crystalline solid.
- All amino acids have a high melting point greater than 200ºC.

- Amino acids are soluble in water, slightly soluble in alcohol and sparingly soluble in organic solvents such as methanol, ethanol and propanol. R-group of amino acids and pH of the solvent plays important role in solubility.
- On heating at high temperatures, amino acids decompose slowly.
- All amino acids (except glycine) are optically active.
- Amino acids can connect with a peptide bond involving their amino and carboxylate groups. A covalent bond formed between the α-amino group of one amino acid and an α-carboxyl group of other forming -CO-NH- linkage. Peptide bonds are planar and partially ionic.

Chemical properties

1. **Zwitterionic property**: A zwitterion is a molecule with functional groups, of which at least one has a positive and one has a negative electrical charge. The net charge of the entire molecule is zero. Amino acids are the best known examples of zwitterions. They contain an amine group (basic) and a carboxylic group (acidic). The $-NH_2$ group is the stronger base, and so it picks up H^+ from the -COOH group to leave a zwitterion. The zwitterion (neutral) is the usual form amino acids exist in solution.
2. **Amphoteric property**: Amino acids are amphoteric in nature; they act as both acids and bases since due to the presence of amine and carboxylic group in same molecule.
3. **Ninhydrin test**: When 1 ml of Ninhydrin solution is added to a 1 ml protein solution and heated constantly, the solution becomes violet in color which indicates the presence of an α-amino acid.
4. **Xanthoproteic test:** The xanthoproteic test is performed for the detection of aromatic amino acids (Tyrosine, Tryptophan and Phenylalanine) in a protein solution. The nitration of benzoid radicals present in the amino acid chain occurs due to reaction with nitric acid, giving the solution yellow coloration.
5. **Reaction with Sanger's reagent**: Sanger's reagent (1-fluoro-2, 4-dinitrobenzene) reacts with a free amino group in the peptide chain in a mild alkaline medium under cold conditions.

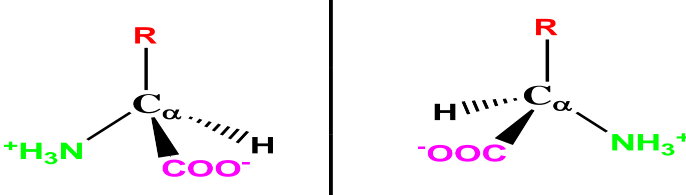
6. **Reaction with nitrous acid**: Nitrous acid reacts with the amino group to liberate nitrogen and form the corresponding hydroxyl compound.

Functions of Amino acids

- All the amino acids are crucial for life as they contain peptides and proteins which are known to be building blocks for all living things.
- The linear sequence of amino acid residues in a polypeptide chain determines the three-dimensional configuration of a protein, and the structure of a protein determines its function.
- Amino acids are imperative for sustaining the health of the human body. They largely promote; production of hormones, structure of muscles, human nervous system's healthy functioning, the health of vital organs and normal cellular structure.
- The amino acids are used by various tissues to synthesize proteins and to produce nitrogen-containing compounds (e.g., purines, heme, creatine, epinephrine), or they are oxidized to produce energy.
- The breakdown of both dietary and tissue proteins yields nitrogen-containing substrates and carbon skeletons.
- The nitrogen-containing substrates are used in the biosynthesis of purines, pyrimidines, neurotransmitters, hormones, porphyrins, and nonessential amino acids.
- The carbon skeletons are used as a fuel source in the citric acid cycle, used for gluconeogenesis, or used in fatty acid synthesis.

Stereochemistry of amino acids

With the exception of glycine, all the 19 other common amino acids have a uniquely different functional group on the central tetrahedral alpha carbon (i.e. Cα). The Cα is termed "chiral" to indicate there are four different constituents and that the Cα is asymmetric. Since the Cα is asymmetric there exist two possible, non-superimposable, mirror images of the amino acids:



The D, L system

The four bonds of the central (alpha) carbon (C) of an amino acid are directed towards the four corners of a tetrahedron. With respect to the carboxyl (COOH) and amino (NH_2) groups, there are two possible arrangements of the H atom and radical group. These arrangements are literally mirror images of each other, and are called stereoisomers (enantiomers). Stereoisomers are designated D (dextro-rotatory) or L (levo-rotatory) according to the direction in which the crystalline forms rotate polarized light, to the right and left, respectively. Naturally-occurring proteins comprise exclusively the L forms of amino acids.

Every amino acid (except glycine) can occur in two isomeric forms, because of the possibility of forming two different enantiomers (stereoisomers) around the central carbon atom. By convention, these are called L-and D-forms, analogous to left-handed and right-handed configurations. Only L-amino acids are manufactured in cells and incorporated into proteins. Some D-amino acids are found in the cell walls of bacteria, but not in bacterial proteins.

Glyceraldehyde contains a chiral carbon, and therefore, there are two enantiomers of this molecule. One is labeled the "L" form, and the other the "D" form. This is the frame of reference used to describe amino acid enantiomers as being either the "L" or "D" form.

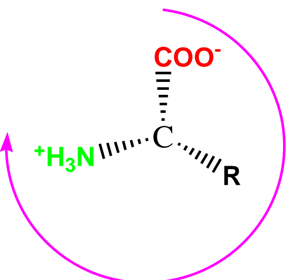
CHO | CHO
HO, C, H, CH_2OH | HOH_2C, C, H, OH

L-Glyceraldehyde | **D-Glyceraldehyde**

COO^- | COO^-
^+H_3N, C, H, R | R, C, H, NH_3^+

L-Amino Acid | **D-Amino Acid**

Even though the two enantiomers would seem to be essentially equivalent to each other, all common amino acids are found in the "L" enantiomeric form in living systems.



Starting with the carbonyl functional group, and going clockwise around the CαCα of the L-enantiomer, the three functional groups spell out the word CORN which is a meaningless word.

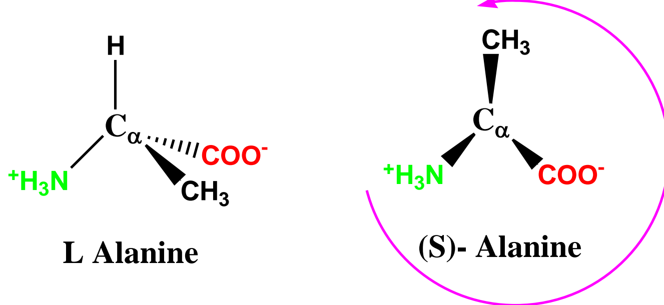
The R, S system of naming chiral centers

A relative ranking of the "priority" of various functional groups is given below:

$$\mathbf{SH > OH > NH_2 > COOH > CHO > CH_2OH > CH_3 > H}$$

A chiral center has four different functional groups. Firstly, lowest priority functional group is identified then find out the chiral center down the bond from the chiral center to the lowest priority atom. Assign priorities to other functional groups connected to the chiral center, using the above ranking. If the priority of

these other groups goes in a clockwise rotation, the chirality is "R". If the priorities of these other groups goes counterclockwise, the chirality is "S”.



Spectral analysis of amino acids

- Spectral analysis refers to the ability of amino acids to absorb or emit electromagnetic radiations at different wavelengths (i.e. energies)
- No any amino acid absorbs light in the visible spectrum which represents that they are "colorless"
- If proteins have color e.g. hemoglobin is red; it is because they contain iron in the structure
- All the amino acids are absorbed in the infrared region (longer wavelengths, weaker energy than visible light)
- Some amino acids are absorbed in the ultraviolet spectrum (shorter wavelengths, higher energy than visible light)
- Electrons in aromatic ring structures absorb in the UV spectrum. Such structures comprise the side chains of tryptophan, tyrosine and phenylalanine

Acid-base behavior of Amino acids

- Amino acids are amphoteric (or amphiprotic); they can react either as an acid or as a base.
- An amino acid can have several forms depending on the pH of the system. At low pH or acidic conditions, the amino group ($-NH_2$) is protonated by the addition of a proton (H^+) from the acid.

- At high pH or basic conditions, the carboxylic acid (-COOH) is deprotonated by the removal of a proton.

- There is an internal transfer of a hydrogen ion from the -COOH group to the $-NH_2$ group to leave an ion with both a negative charge and a positive charge.

- In an aqueous solution at a certain compound specific pH, this structure may change so that a proton from the -COOH (carboxylic acid group), transfers to the $-NH_2$ (amino group), leaving an ion with both a negative charge and a positive charge. The transfer of electrons or ions results in a net neutral charge because the number of protonated ammonium groups with a positive charge and deprotonated carboxylate groups with a negative charge are equal.

- All amino acid contains an acidic carboxylic group and a basic amino group. It can form a zwitter ion at pH=7. Zwitter ion forms when carboxylic group releases it proton and remains in an anionic form and NH_2 group takes a proton and remains in a cationic form. At this pH it is called isoelectronic point.

Zwitter ion

At this isoelectric point, the positive ion does not move towards anode and negative charge does not move towards cathode.

When a base is added to the amino acid increasing the pH of the solution, then positive charge on the NH_3^+ is removed and the molecule becomes an anion. This anion moves toward the positively charged anode.

OH^-

Anionic species

In the presence of an acid, the carboxylate anion takes one proton from the acid solution and becomes neutral. The net amino acid becomes cationic charged and moves towards the anionic charged cathode.

H^+

Cationic species

Isoelectric Point (pI)

The pH at which all amino acids are in the zwitterions form and contains very low and equal concentrations of the positive and negative ions is known as Isoelectric point.

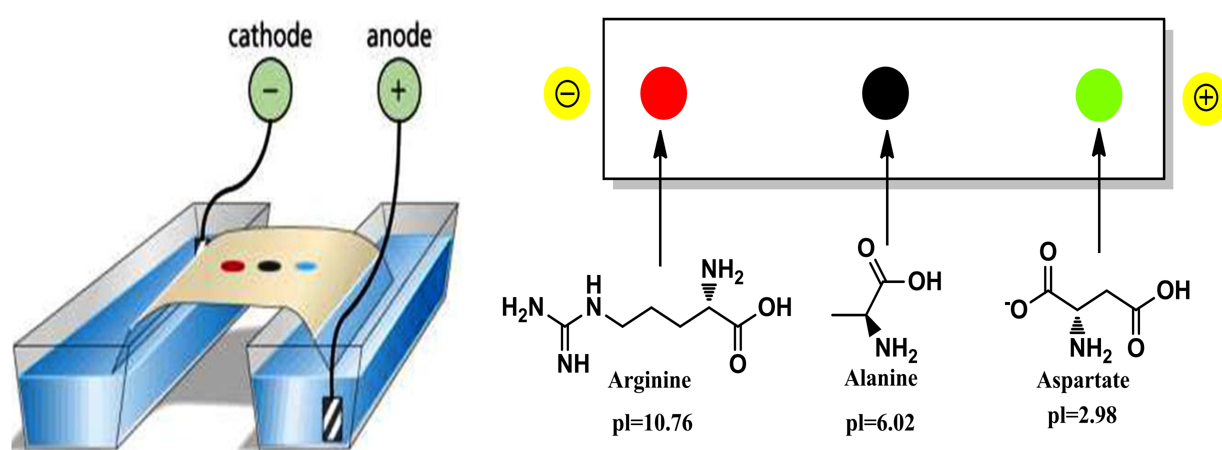
- In more acidic media (pH<pI), the concentration of ions increases while the concentration of the zwitterion decreases.
- In a more basic media (pH>pI), the concentration of negative ions increases while the concentration of the zwitterion decreases.
- More than 98% of the amino acid is in zwitterion from over a pH range of ±2 of pI.
- At pH=pI, the amino acid present as the zwitterion with one amine or carboxyl group in uncharged form.
- Isoelectric points found at the values ranging from 7.8 to 10.8 (basic)
- Isoelectric points found at the values ranging from 4.86 to 6. 3 (neutral)
- Isoelectric points found at the values ranging from 2.8 to 3.3 (acidic)

Isoelectric Points of some amino acids

Amino Acid	Isoelectric point (pI)
Arganine (Arg)	10.8
Lysine (Lys)	9.7
Alanine (Ala)	6.0
Glycine (Gly)	6.0
Serine (Ser)	5.7
Glutamic acid (Glu)	3.2
Aspartic acid (Asp)	2.9

Electrophoresis

Electrophoresis is an analytical method for identifying amino acids by observing their migration as a function of pH under an applied electric field gradient. At its pI, the amino acid is present in the zwitterion form with no net charge′ and will not migrate in electrophoresis. At pH < pI, the amino acid carries a positive charge and will migrate to the negative electrode. At pH, pI, the amino acid carries a negative charge and will migrate to the positive electrode.



Preparation of α-amino acid: Following methods are usually used for the preparation of α–amino acids

1. **Amination of α–halo acids:** α-chloro or α-bromo acids on treatment with aqueous ammonia give the α-amino acids.

$$R-\underset{H}{\overset{Br}{C}}-COOH \xrightarrow[(ii)\ Dil.\ HCl]{(i)\ NH_3(excess)} R-\underset{H}{\overset{NH_2}{C}}-COOH$$

It may also be prepared by Hell-Volhard-Zelinsky reaction.

$$R-\underset{H}{\overset{H}{C}}-COOH \xrightarrow{P(red) + Br_2} R-\underset{H}{\overset{Br}{C}}-COOH$$

2. **Amination of α–halo ester- Gabriel Phthalimide synthesis:** In this method, phthalimide is treated with α–halo ester to introduce amino group.

$$C_6H_4(CO)_2\overset{\ominus}{N}\overset{\oplus}{K} + Cl-CH_2COOEt \longrightarrow C_6H_4(CO)_2N-CH_2COOEt$$

$$C_6H_4(CO)_2N-CH_2COOEt \xrightarrow{HCl/H_2O} C_6H_4(COOH)_2 + \overset{\oplus}{N}H_3-\overset{H_2}{C}-COOH$$

3. **Strecker Synthesis:** An elegant procedure, known as the Strecker synthesis, assembles an alpha-amino acid from ammonia (the amine precursor), cyanide (the carboxyl precursor), and an aldehyde. This reaction is essentially

an imino analog of cyanohydrin formation. The alpha-amino nitrile formed in this way can then be hydrolyzed to an amino acid by either acid or base catalysis.

$$R\text{-}CHO + NH_3 \rightleftharpoons H_2O + R\text{-}CH\text{=}NH \rightleftharpoons R\text{-}C(NH_2)(H)\text{-}CN \xrightarrow[\text{Acid/Base}]{H_2O} R\text{-}C(NH_3^{\oplus})(H)\text{-}CO_2^{\ominus}$$

Reactions of α–amino acids

Amino acids have two reactive functional groups: Amino group and carboxylic acid group. It reacts with different kind of reagents because of these two groups either separately or in combined form.

A. Reactions due to amine group of Amino acid

1. **Reaction with strong inorganic acids:** The amine group of amino acid reacts with strong inorganic acids and produces respective salt.

$$R\text{-}C(NH_2)(H)\text{-}COOH + HCl \longrightarrow R\text{-}C(NH_3^{\oplus}\,Cl^{\ominus})(H)\text{-}COOH$$

2. **Amine acylation:** In order to convert the amine function of an amino acid into an amide, the pH of the solution must be raised to 10 or higher so that free amine nucleophiles should present in the reaction system. Carboxylic acids are converted to carboxylate anions at such a high pH, and do not interfere with amine acylation reactions. The following two reactions are illustrative. In the first, an acid chloride serves as the acylating reagent. This is a good example of the superior nucleophilicity of nitrogen in acylation reactions, since water and hydroxide anion are also present as competing nucleophiles. The second reaction employs an anhydride-like reagent for the acylation. This is a particularly useful procedure in peptide synthesis. Since amides are only weakly basic (pKa~ -1), the resulting

amino acid derivatives do not display zwitterionic character, and may be converted to a variety of carboxylic acid derivatives.

$$R{-}CH(NH_3^{\oplus}){-}CO_2^{\ominus} \xrightarrow[C_6H_5COCl]{NaOH} \left[R{-}CH(NH_2){-}CO_2^{\ominus} \right] \longrightarrow R{-}CH(HN{-}CO{-}C_6H_5){-}CO_2^{\ominus} \xrightarrow{H_3O^{\oplus}} R{-}CH(HN{-}CO{-}C_6H_5){-}CO_2H$$

$$HC(H_3C)_2{-}CH(NH_3^{\oplus}){-}CO_2^{\ominus} + C(H_3C)_3{-}O{-}CO{-}O{-}CO{-}O{-}C(CH_3)_3 \xrightarrow[THF\ and\ H_2O]{NaOH} HC(H_3C)_2{-}CH(HN{-}CO{-}O{-}(CH_3)_3C){-}CO_2^{\ominus}$$

3. **Reaction with Nitrous acid:** Amino acids react with nitrous acid and produces respective hydroxyl acids with their original configuration. Estimation of the nitrogen evolved in the reaction is the basis of Van Slyke estimation of amino acids.

$$R{-}CH(NH_2){-}COOH + HNO_2 \longrightarrow R{-}CH(OH){-}COOH + N_2 + H_2O$$

4. **Reaction with formaldehyde:** Formaldehyde reacts with amino group of amino acid and forms mixture of products.

$$R{-}CH(NH_2){-}COOH + H{-}C(=O){-}H \longrightarrow R{-}CH(N{=}CH_2){-}COOH$$

$$R{-}CH(CH_3){-}COOH + 2\ H{-}C(=O){-}H \longrightarrow R{-}CH(N(CH_2OH)(HOH_2C)){-}COOH$$

B. Reactions due to Carboxylic acid group of Amino acid

1. **Reaction with alkali:** Amino acid reacts with alkali and forms salt.

$$R-CH(NH_2)-COOH + NaOH \longrightarrow R-CH(NH_2)-COO^{\ominus}Na^{\oplus}$$

2. **Esterification of Amino acids:** Amino acids undergo most of the chemical reactions characteristic of each function, assuming the pH is adjusted to an appropriate value. Esterification of the carboxylic acid is usually conducted under acidic conditions, as shown in the two equations written below. Under such conditions, amine functions are converted to their ammonium salts and carboxyic acids are not dissociated. The first equation is a typical Fischer esterification involving methanol. The initial product is a stable ammonium salt. The amino ester formed by neutralization of this salt is unstable, due to acylation of the amine by the ester function. The second reaction illustrates benzylation of the two carboxylic acid functions of aspartic acid, using p-toluenesulfonic acid as an acid catalyst. Once the carboxyl function is esterified, zwitterionic species are no longer possible and the product behaves like any 1°-amine.

$$R-CH(NH_3^{\oplus})-CO_2^{\ominus} \xrightarrow[CH_3OH]{HCl} [R-CH(NH_3^{\oplus})-CO_2H] \xrightarrow[HCl]{CH_3OH} R-CH(NH_3^{\oplus})-CO_2CH_3 \xrightarrow[\text{or any Base}]{NaHCO_3} R-CH(NH_2)-CO_2CH_3$$

Unstable

$$^{\ominus}O_2C-CH_2-CH(NH_3^{\oplus})-CO_2^{\ominus} \xrightarrow[2\ C_6H_5CH_2OH]{C_7H_7SO_3H} C_6H_5CH_2O-CO-CH_2-CH(NH_3^{\oplus})-CO-OCH_2C_6H_5$$

Decarboxylation of Amino acid: When amino acids are heated with BaO_2, they yield corresponding amines.

$$R-CH(NH_3^{\oplus})-COO^{\ominus} \xrightarrow{BaO_2/\Delta} RCH_2NH_2 + CO_2$$

3. **Reduction of amino acid:** Amino acids can be reduced to corresponding amino alcohols by reduction with lithium aluminium hydride (LAH).

$$R-CH(NH_3^{\oplus})-COO^{\ominus} \xrightarrow{LAH} R-CH(CH_2OH)-NH_2$$

4. **Formation of acid chlorides:** In the formation of amino acid chlorides, the amino group is first protected by treating it with acetic anhydride. The acetylated amino acid formed is converted to the acetylated amino acid chloride with phosphorus penta chloride.

$$R-CH(NH_3^{\oplus})-COO^{\ominus} + Ac_2O \longrightarrow R-CH(NHAc)-COOH \longrightarrow R-CH(NHAc)-C(=O)-Cl$$

Cysteine oxidation: The sulfhydryl group of cysteine is highly reactive. The most common reaction of this group is a reversible oxidation that forms a disulfide. Oxidation of two molecules of cysteine forms cystine, a molecule that contains a disulfide bond. When two cysteine residues form such a bond, it is referred to as a disulfide bridge. This bond can occur in a single chain to form a ring or between two separate chains to form an intermolecular bridge. Disulfide bridges help in stabilizing many polypeptides and proteins.

$$2\ HS-CH_2-CH(NH_2)-COOH \underset{[H]}{\overset{[O]}{\rightleftharpoons}} HOOC-CH(NH_2)-CH_2-S-S-CH_2-CH(NH_2)-COOH$$

Cysteine　　Cystine

Reagents used for the identification of Amino acids: Following reagents can be used to identify amino acids and measure their amounts.

1. **Ninhydrin Reaction:** The most common reagent for quantitative and qualitative analysis of amino acids is ninhydrin. It is strong oxidizing agent and causes oxidative decarboxylation of the amino acids. Each molecule of amino acid reacts with 2 molecules of ninhydrin to produce a purple pigment that can be estimated colorimetrically at 750nm.

Ninhydrin

$- 2\ H_2O$

$-CO_2$

H_2O

- RCHO

Ninhydrin

$-H_2O$

Purple or Blue

Proline and hydroxyproline produce yellow colour instead of blue with ninhydrin.

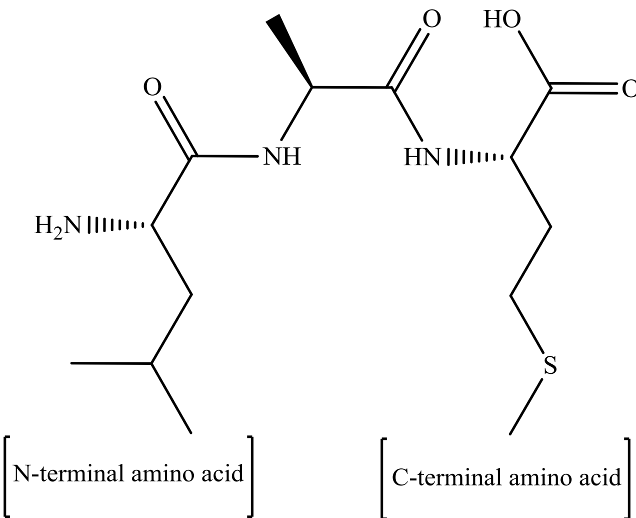
2. **Sanger reaction:** The amino acids react with 2,4-dinitrofluorobenzene (DNFB) also known as Sanger reagent to produce N-(dinitrophenyl) amino acid (DNPAA). The obtained product is yellow in colour and can be identified by chromatographic techniques.

Structure and nomenclature of peptides and proteins: A peptide is a short chain of amino acids. The amino acids in a peptide are connected to one another in a sequence by bonds called peptide bonds. Typically, peptides are distinguished from proteins by their shorter length (less than 50 amino acids). Meanwhile, proteins are long molecules made up of multiple peptide subunits (more than 50 amino acids) and are also known as polypeptides. Proteins can be digested by enzymes (other proteins) into short peptide fragments.

$$n\ H_2N-\underset{R}{\overset{H}{C}}-COOH \longrightarrow H_2N-\underset{H}{\overset{R}{C}}-\overset{O}{\overset{\|}{C}}-\left(\overset{H}{N}-\underset{H}{\overset{R}{C}}-\overset{O}{\overset{\|}{C}}\right)_{n-2}-\overset{H}{N}-\underset{H}{\overset{R}{C}}-COOH$$

Polypeptide

In naming peptides, the amino acid at the end of the polymer bearing a free $-NH_2$ group is known as “N-terminal amino acid”. The end amino acid bearing the free carboxylic group is called with the N-terminal amino acid, which is numbered 1. By convention, the N-terminal end is written to the left and the C-terminal end to the right. The amino acids are then named left to right



Leucylalanyl methionine

Instead of writing out full name of amino acid in a polypeptide sequence, scientists find it convenient to use three letter abbreviation in the interest of brevity.

For example: Leucylalanyl methionine or Leu. Ala. Met.

Determination of the structures of Polypeptide: The structure of a polypeptide can be determined by knowing (i) the name of amino acids constituting the peptide, (ii) the number of molecules of individual amino acids, and (iii) their sequence in the peptide chain.

1. Cleavage of disulfide linkages: The first step in structure determination is to break all the disulfide bonds, opening any disulfide-linked rings and separating the individual peptide chains. The individual peptide chains are then purified and analyzed separately. Cystine bridges are easily cleaved by reducing them to the thiol (cysteine) form. These reduced cysteine residues have a tendency to reoxidize and re-form disulfide bridges, however, a more permanent cleavage involves oxidizing the disulfide linkages with peroxyformic acid. This oxidation converts the disulfide bridges to sulfonic acid groups (-SO_3H). The oxidized cysteine units are called cysteic acid residues.

Oxidation of a protein by peroxyformic acid cleaves all the disulfide linkages by oxidizing cystine to cysteic acid.

2. Determination of the Amino acid composition: Once the disulfide bridges have been broken and the individual peptide chains have been separated and purified, the structure of each chain must be determined. The first step is to determine which amino acids are present and in what proportions. To analyze the amino acid composition, the peptide chain is completely hydrolyzed by

boiling it for 24 hours in 6 M HCl. The resulting mixture of amino acids (the hydrolysate) is placed on the column of an amino acid analyzer.

In the amino acid analyzer, the components of the hydrolysate are dissolved in an aqueous buffer solution and separated by passing them down an ion-exchange column. The solution emerging from the column is mixed with ninhydrin, which reacts with amino acids to give the purple ninhydrin color. The absorption of light is recorded and printed out as a function of time. The time required for each amino acid to pass through the column (its retention time) depends on how strongly that amino acid interacts with the ion-exchange resin. The retention time of each amino acid is known from standardization with pure amino acids. The amino acids present in the sample are identified by comparing their retention times with the known values. The area under each peak is nearly proportional to the amount of the amino acid producing that peak, so we can determine the relative amounts of amino acids present.

Sequencing the Peptide-Terminal Residue Analysis: The amino acid analyzer determines the amino acids present in a peptide, but it does not reveal their sequence: the order in which they are linked together. The peptide sequence is destroyed in the hydrolysis step. To determine the amino acid sequence, we must cleave just one amino acid from the chain and leave the rest of the chain intact. The cleaved amino acid can be separated and identified, and the process can be repeated on the rest of the chain. The amino acid may be cleaved from either end of the peptide (either the N terminus or the C terminus), and we will consider one method used for each end. This general method for peptide sequencing is called terminal residue analysis.

Sequencing from the N Terminus

1. The Edman Degradation

The most efficient method for sequencing peptides is the Edman degradation. A peptide is treated with phenyl isothiocyanate, followed by acid hydrolysis. The products are the shortened peptide chain and a heterocyclic derivative of the N-terminal amino acid called a phenylthiohydantoin.

This reaction takes place in three stages. First, the free amino group of the Nterminal amino acid reacts with phenylisothiocyanate to form a phenylthiourea. Second, the phenylthiourea cyclizes to a thiazolinone and expels the shortened peptide chain. Third, the thiazolinone isomerizes to the more stable phenylthiohydantoin.

Step-1: Nucleophilic attack by the free amino group on phenyl isothiocyanate, followed by a proton transfer, gives a phenylthiourea.

Phenylthiourea

Step 2: Treatment with HCl induces cyclization to a thiazolinone and expulsion of the shortened peptide chain.

Protonated thiourea

thiazolinone

Step 3: In acid, the thiazolinone isomerizes to the more stable phenylthiohydantoin.

HCl

thiazolinone phenylthiohydantoin

The phenylthiohydantoin derivative is identified by chromatography, by comparing it with phenylthiohydantoin derivatives of the standard amino acids.

This gives the identity of the original N-terminal amino acid. The rest of the peptide is cleaved intact, and further Edman degradations are used to identify additional amino acids in the chain. This process is well suited to automation, and several types of automatic sequencers have been developed.

2. The Sanger method

The Sanger method for N-terminus determination is a less common alternative to the Edman degradation. In the Sanger method, the peptide is treated with the Sanger reagent, 2,4-dinitrofluorobenzene, and then hydrolyzed by reaction with 6 M aqueous HCl. The N-terminal amino acid is recovered as its 2,4-dinitrophenyl derivative and identified.

1-fluoro-2,4-dinitrobenzene (Sanger reagent) + Peptide → (6M HCl, Heat) → 2,4-dinitrophenyl derivative + Amino acid

C-Terminal Residue Analysis

There is no efficient method for sequencing several amino acids of a peptide starting from the C terminus. In many cases, however, the C-terminal amino acid can be identified using the enzyme carboxypeptidase, which cleaves the C-terminal peptide bond. The products are the free C-terminal amino acid and a shortened peptide. Further reaction cleaves the second amino acid that has now become the new C terminus of the shortened peptide. Eventually, the entire peptide is hydrolyzed to its individual amino acids.

$$\boxed{\text{Peptide}}-\ddot{N}H-\underset{R^{n-1}}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-\overset{H}{N}-\underset{R^{n}}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-OH \xrightarrow[H_2O]{\text{Carboxypeptidase}} \boxed{\text{Peptide}}-\ddot{N}H-\underset{R^{n-1}}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-OH + H_2N-\underset{R^{n}}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-OH$$

(Further cleavage) Free amino acid

A peptide is incubated with the carboxypeptidase enzyme, and the appearance of free amino acids is monitored. In theory, the amino acid whose concentration increases first should be the C terminus, and the next amino acid to appear should be the second residue from the end. In practice, different amino acids are cleaved at different rates, making it difficult to determine amino acids past the C terminus and occasionally the second residue in the chain.

Breaking the Peptide into Shorter Chains:

Partial Hydrolysis Before a large protein can be sequenced, it must be broken into smaller chains, not longer than about 30 amino acids. Each of these shortened chains is sequenced, and then the entire structure of the protein is deduced by fitting the short chains together like pieces of a jigsaw puzzle.

Partial cleavage can be accomplished either by using dilute acid with a shortened reaction time or by using enzymes, such as trypsin and chymotrypsin, that break bonds between specific amino acids. The acid-catalyzed cleavage is not very selective, leading to a mixture of short fragments resulting from cleavage at various positions. Enzymes are more selective, giving cleavage at predictable points in the chain.

- **Trypsin:** Cleaves the chain at the carboxyl groups of the basic amino acids lysine and arginine.
- **Chymotrypsin:** Cleaves the chain at the carboxyl groups of the aromatic amino acids phenylalanine, tyrosine, and tryptophan.

To explain this phenomenon, partial hydrolysis of oxytocin can be taken as an example. Oxytocin could be sequenced directly by C-terminal analysis and a series of Edman degradations, but it provides a simple example of how a structure can be pieced together from fragments. Acid-catalyzed partial

hydrolysis of oxytocin (after cleavage of the disulfide bridge) gives a mixture that includes the following peptides:

Ile-Gln-Asn-Cys   Gln-Asn-Cys-Pro   Pro-Leu-Gly $.NH_2$ Cys-Tyr-Ile-Gln-Asn   Cys-Pro-Leu-Gly

When we match the overlapping regions of these fragments, the complete sequence of oxytocin appears:

Cys-Tyr-Ile-Gln-Asn

Ile-Gln-Asn-Cys

Gln-Asn-Cys-Pro

Cys-Pro-Leu-Gly

Pro-Leu-Gly.NH_2

Hence, complete structure would be

Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly.NH_2

The two Cys residues in oxytocin may be involved in disulfide bridges, either linking two of these peptide units or forming a ring. By measuring the molecular weight of oxytocin, we can show that it contains just one of these peptide units; therefore, the Cys residues must link the molecule in a ring.

Peptide Synthesis:

Peptide synthesis requires the formation of amide bonds between the proper amino acids in the proper sequence. With simple acids and amines, we would form an amide bond simply by converting the acid to an activated derivative (such as an acyl halide or anhydride) and adding the amine.

$$R-C(=O)-X + H_2\ddot{N}-R' \longrightarrow R-C(=O)-NH-R' + H-X$$

Where X is good leaving group, preferably electron-withdrawing

Each amino acid has both an amino group and a carboxyl group. If we activate the carboxyl group, it reacts with its own amino group. If we mix some

amino acids and add a reagent to make them couple, they form every conceivable sequence. Also, some amino acids have side chains that might interfere with peptide formation. For example, glutamic acid has an extra carboxyl group, and lysine has an extra amino group. As a result, peptide synthesis always involves both activating reagents to form the correct peptide bonds and protecting groups to block formation of incorrect bonds.

Several methods have been developed by chemists which falls under two major groups. The solution-phase method involves adding reagents to solutions of growing peptide chains and purifying the products as needed. The solid-phase method involves adding reagents to growing peptide chains bonded to solid polymer particles.

1. Solution-Phase Method

Consider the structure of alanylvalylphenylalanine, a simple tripeptide:

Alanyl Valyl Phenylalanine

Alanylvalylphenylalanine

Solution-phase peptide synthesis begins at the N terminus and ends at the C terminus, or left to right as we draw the peptide. The first major step is to couple the carboxyl group of alanine to the amino group of valine. This cannot be done simply by activating the carboxyl group of alanine and adding valine. If we activated the carboxyl group of alanine, it would react with another molecule of alanine. To prevent side reactions, the amino group of alanine must be protected to make it nonnucleophilic. It is observed that an amino acid reacts with benzyl chloroformate (also called benzyloxycarbonyl chloride) to form a urethane, or carbamate ester, that is easily removed at the end of the synthesis.

This protecting group has been used for many years, and it has acquired several names. It is called the benzyloxycarbonyl group, the carbobenzoxy group (Cbz), or simply the Z group (abbreviated Z).

Preliminary step: Protect the amino group with Z.

C_6H_5–CH_2–O–C(=O)–Cl + H_2N–HC(CH_3)–C(=O)–OH → [Z Group: C_6H_5–CH_2–O–C(=O)]–NH–CH(CH_3)–C(=O)–OH + HCl

Benzyl chloroformate Z-Cl; Alanine Ala; Benzyloxycarbonyl alanine

The amino group in Z-Ala is protected as the nonnucleophilic amide half of a carbamate ester. The carboxyl group can be activated without reacting with the protected amino group. Treatment with ethyl chloroformate converts the carboxyl group to a mixed anhydride of the amino acid and carbonic acid. It is strongly activated toward nucleophilic attack.

Step 1: Activate the carboxyl group with ethyl chloroformate.

Z–NHCH(CH_3)–C(=O)–OH + HCl–C(=O)–OCH_2CH_3 → Z–NHCH(CH_3)–[C(=O)–O–C(=O)–OCH_2CH_3] + HCl

Protected alanine; Ethyl chloroformate; anhydride of carbonic acid; Mixed anhydride

When the second amino acid (valine) is added to the protected, activated alanine, the nucleophilic amino group of valine attacks the activated carbonyl of alanine, displacing the anhydride and forming a peptide bond.

Step 2: Form an amide bond to couple the next amino acid.

Z–NHCH(CH_3)–C(=O)–[O–C(=O)–OCH_2CH_3] + H_2N–CH($CH(CH_3)_2$)–C(=O)–OH → Z–NHCH(CH_3)–C(=O)–NH–CH($CH(CH_3)_2$)–C(=O)–OH + CO_2↑ + CH_3CH_2OH

Protected, activated alanine; Valine; Z-Ala-Val

To make a larger peptide, repeat these two steps for the addition of each amino acid residue: 1. Activate the C terminus of the growing peptide by

reaction with ethyl chloroformate. 2. Couple the next amino acid. The final step in the solution-phase synthesis is to deprotect the N terminus of the completed peptide. The N-terminal amide bond must be cleaved without breaking any of the peptide bonds in the product. Fortunately, the benzyloxycarbonyl group is partly an amide and partly a benzyl ester, and hydrogenolysis of the benzyl ester takes place under mild conditions that do not cleave the peptide bonds. This mild cleavage is the reason for using the benzyloxycarbonyl group (as opposed to some other acyl group) to protect the N terminus.

Final step: Remove the protecting group

$$C_6H_5{-}CH_2{-}O{-}C(=O){-}NHCH(CH_3)C(=O){-}Val{-}Phe \xrightarrow{H_2,\ Pd} NH_2CH(CH_3)C(=O){-}Val{-}Phe + CO_2 + Ph{-}CH_3$$

Z-Ala-Val-Phe Ala-Val-Phe

The solution-phase method works well for small peptides, and many peptides have been synthesized by this process. A large number of chemical reactions and purifications are required even for a small peptide, however although the individual yields are excellent, with a large peptide, the overall yield becomes so small as to be unusable, and several months (or years) are required to complete so many steps. It needs too much time and several purification steps for getting the final result. For larger peptides and proteins, solid-phase peptide synthesis is usually preferred.

Solid-phase peptide synthesis: Three reactions are crucial for solid-phase peptide synthesis. These reactions attach the first amino acid to the solid support, protect each amino group until its time to react, and form the peptide bonds between the amino acids.

Attaching the Peptide to the Solid Support: The greatest difference between solution-phase and solid-phase peptide synthesis is that solid-phase synthesis is done in the opposite direction: starting with the C terminus and going toward the N terminus, right to left as we write the peptide. The first step is to attach the last amino acid (the C terminus) to the solid support. The solid support is a special polystyrene bead in which some of the aromatic rings have chloromethyl

groups. This polymer, often called the Merrifield resin, is made by copolymerizing styrene with a few percent of p-(chloromethyl) styrene.

Formation of the Merrifield resin:

Styrene p(chloromethyl)styrene Polymer abbreviation

Like other benzyl halides, the chloromethyl groups on the polymer are reactive toward attack. The carboxyl group of an N-protected amino acid displaces chloride, giving an amino acid ester of the polymer. In effect, the polymer serves as the alcohol part of an ester protecting group for the carboxyl end of the C-terminal amino acid. The amino group must be protected, or it would attack the chloromethyl groups.

Attachment of the C-terminal amino acid

Once the C-terminal amino acid is fixed to the polymer, the chain is built on the amino group of this amino acid.

Using the tert-Butyloxycarbonyl (Boc) Protecting Group: The benzyloxycarbonyl group (the Z group) cannot be used with the solid-phase process because the Z group is removed by hydrogenolysis in contact with a solid catalyst. A polymer-bound peptide cannot achieve the intimate contact with a solid catalyst required for hydrogenolysis. The N-protecting group used in the Merrifield procedure is the tert-butyloxycarbonyl group, abbreviated Boc or t-Boc. The Boc group is similar to the Z group, except that it has a tert-butyl group in place of the benzyl group. Like other tert-butyl esters, the Boc protecting group is easily removed under acidic conditions. The acid chloride of the Boc group is unstable, so we use the anhydride, di-tertbutyldicarbonate, to attach the group to the amino acid.

Protection of the amino group as its Boc derivative

di-tert-butyldicarbonate + Amino Acid → Boc-amino acid + CO_2 + $(CH_3)_3C{-}OH$

The Boc group is easily cleaved by brief treatment with trifluoroacetic acid (TFA), CF_3COOH Loss of a relatively stable tert-butyl cation from the protonated ester gives an unstable carbamic acid. Decarboxylation of the carbamic acid gives the deprotected amino group of the amino acid. Loss of a proton from the tert-butyl cation gives isobutylene.

CF_3COOH

Boc-amino acid → Protonated →

Carbamic acid → Free amino acid + Isobutylene + CO_2 ↑

Use of DCC as a Peptide Coupling Agent:

The final reaction needed for the Merrifield procedure is the peptide bond-forming condensation. When a mixture of an amine and an acid is treated with N,N'-dicyclohexylcarbodiimide (abbreviated DCC), the amine and the acid couple to form an amide. The molecule of water lost in this condensation converts DCC to N,N'-dicyclohexyl urea (DCU).

Acid + Amine + N,N'-dicyclohexylcarbodiimide (DCC) → Amide + N,N'-dicyclohexyl urea (DCU)

The mechanism for DCC coupling is not as complicated as it may seem. The carboxylate ion adds to the strongly electrophilic carbon of the diimide, giving an activated acyl derivative of the acid. This activated derivative reacts readily with the amine to give the amide. In the final step, DCU serves as an excellent leaving group. The cyclohexane rings are miniaturized for clarity.

Formation of an activated acyl derivative

Activated

Coupling with the amine and loss of DCU

Amide

N,N'-dicyclohexyl urea (DCU)

At the completion of the synthesis, the ester bond to the polymer is cleaved by anhydrous HF. Because this is an ester bond, it is more easily cleaved than the amide bonds of the peptide.

Cleavage of the finished peptide

Proteins

All proteins are mixed polymers of amino acids and as such all contain carbon, hydrogen, oxygen and nitrogen. Most of the proteins contain sulphur, a large number contain phosphorus and some contain various other mineral elements such as Cu^{++}, Zn^{++}, Fe^{++} etc. They have very high molecular weights, generally above 10,000 and may even up to millions. Proteins are large peptides which have several times higher molecular weights than peptides. The basic characteristics of proteins are as follows:

- Like amino acid, proteins also have amino and carboxylic terminus and are amphoteric in nature, i.e. they behave both as cations and anions under the influence of electric field. Therefore, they also have isoelectric point characteristic of amino acids.

- All proteins are optically active because of the presence of chiral centers at α-position of the amino acid residue.

- Proteins can be identified easily by following colour reactions:

(i) **Ninhydrin test:** When proteins are treated with pyridine solution of ninhydrin gives colour ranging from deep blue to violet pink or red in exceptional cases.

(ii) Biuret test: This test is characteristic of all compounds having a peptide linkage and indicates the presence of peptide bond in proteins. An alkaline solution of protein is treated with a drop of aqueous copper sulphate solution which gives bluish violet colour confirming the presence of peptide bond.

(iii) Millon's test: Millon's test is characteristics of those proteins having phenol group (e.g., those having tyrosine unit). When protein is treated with Millon's reagent ($HgNO_3$ in HNO_3 containing a little HNO_2), a white precipitate is obtained.

(iv) Xanthoproteic test: When proteins are treated with nitric acid gives yellow or orange colour. This is why our skin turns yellow when it comes in contact with nitric acid.

Classification of proteins: Classification of proteins is done on the basis of shape, constitution and nature of molecules.

1. On the basis of shape

(a) Fibrous protein (Scleroprotein): These proteins are foun in animals and are insoluble in water. Fibrous proteins are resistant to proteolytic enzymes and are coiled and exist in threadlike structures to form fibres. e.g. collagen, actin, and myosin, keratin in hair, claws, feathers, etc.

(b) Globular proteins: These proteins, unlike fibrous proteins are soluble in water. They are made up of polypeptides that are coiled about themselves to form oval or spherical molecules e.g. albumin, insulin, and hormones like oxytocin, etc.

2. On the basis of Constitution

(a) Simple proteins: These proteins are made up of amino acids only. e.g. albumins, globulins, prolamins, etc.

(b) Conjugated proteins: These are complex proteins that are combined with the characteristic of non–amino acid substance called as a prosthetic group. These are of following types:

(i) **Nucleoproteins:** Combination of protein and nucleic acid

(ii) **Mucoproteins:** Combination of proteins and carbohydrates (>4%)

(iii) **Glycoproteins:** Combination of proteins and carbohydrates (<4%)

(iv) **Chromoproteins:** Combination of proteins and coloured pigments.

(v) **Lipoproteins:** Combination of proteins and lipids.

(vi) **Metalloprotein:** Combination of proteins and metal ions.

(vii) **Phosphoprotein:** Combination of proteins and phosphate group.

(c) Derived proteins: When proteins are hydrolyzed by acids, alkalies or enzymes, the degradation products obtained from them are called derived proteins.

3. On the basis of nature of molecules

(a) Acidic proteins: They exist as anion and contain acidic amino acids. e.g. blood groups.

(b) Basic proteins: They exist as cations and are rich in basic amino acids e.g. lysine, arginine etc.

Difference between polypeptides and proteins: The difference between a polypeptide and a protein is not very distinct because both are polyamides constructed from amino acids. The basic difference between these two is summarized below:

1. **Difference due to definition:** A polypeptide is a polymer formed by a defined sequence of amino-acids linked together through covalent peptide bonds whereas a polypeptide is a polymer formed by a defined sequence of amino-acids linked together through covalent peptide bonds.

2. **Difference due to structure:** A polypeptide presents a simple structure and consists of the polypeptide backbone formed by the repeating sequence of atoms at the core of the linked amino-acids chain. Attached to the polypeptide backbone are the amino-acids specific side chains, the R

group. A protein, on the other side, is a complex molecule consisting of one or more polypeptide chains folding into secondary, tertiary or quaternary structure. The protein shape is held stable by three types of weak non covalent bonds: hydrogen bonds, ionic bonds, and van der Waal bonds.

3. **Difference due to function:** The main function of a polypeptide is that they are primary structure of more complex proteins. Polypeptides lack the three-dimensional structure which enables a protein to bind to a ligand and be functional. On the other side, the structural complexity of a protein, its stable shape with its ligand-binding sites enables it to bind specifically and with high affinity to particular ligands, to be regulated, and to participate in many vital cellular metabolic pathways.

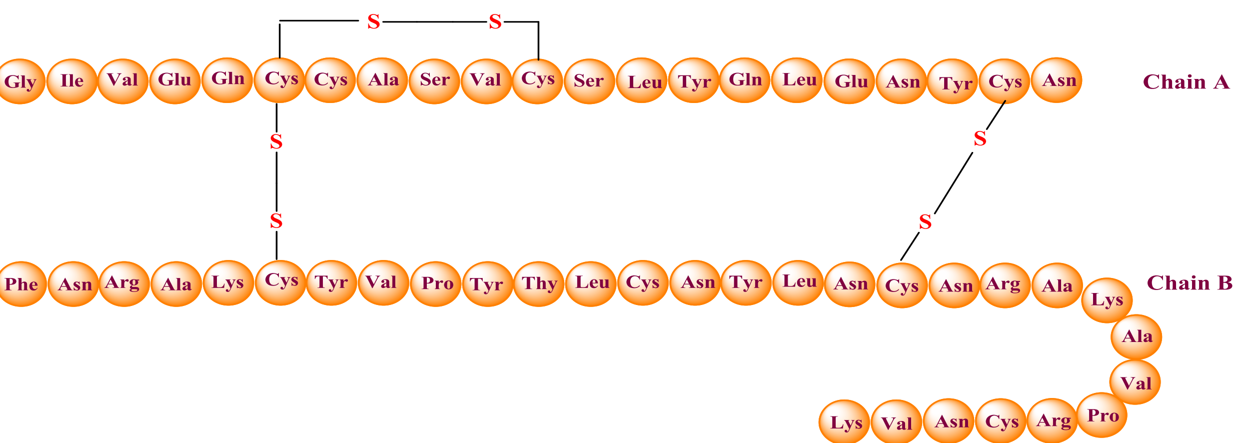
Structure of proteins: Proteins are biological polymers composed of amino acids. Amino acids, linked together by peptide bonds, form a polypeptide chain. One or more polypeptide chains twisted into a 3-D shape form a protein. Proteins have complex shapes that include various folds, loops, and curves. Folding in proteins happens spontaneously. Chemical bonding between portions of the polypeptide chain aid in holding the protein together and giving it its shape. There are two general classes of protein molecules: globular proteins and fibrous proteins. Globular proteins are generally compact, soluble, and spherical in shape. Fibrous proteins are typically elongated and insoluble. These proteins may exhibit one or more of four types of protein structure which may be distinguished from one another by the degree of complexity in the polypeptide chain. A single protein molecule may contain one or more of the protein structure types: primary, secondary, tertiary and quaternary structure.

1. **Primary structure**: It describes the unique order in which amino acids are linked together to form a protein. Proteins are constructed from a set of 20 amino acids. Generally, amino acids have the following structural properties:

- A carbon (the alpha carbon) bonded to the four groups below:
- A hydrogen atom (H)
- A Carboxyl group (-COOH)

- An Amino group (-NH_2)
- A "variable" group or "R" group

All amino acids have the alpha carbon bonded to a hydrogen atom, carboxyl group, and an amino group. The "R" group varies among amino acids and determines the differences between these protein monomers. The amino acid sequence of a protein is determined by the information found in the cellular genetic code. The order of amino acids in a polypeptide chain is unique and specific to a particular protein. Altering a single amino acid causes a gene mutation, which most often results in a non-functioning protein. For example, the hormone insulin has two polypeptide chains, A and B, shown in diagram below. Each chain has its own set of amino acids, assembled in a particular order. For instance, the sequence of the A chain starts with glycine at the N-terminus and ends with asparagine at the C-terminus, and is different from the sequence of the B chain.

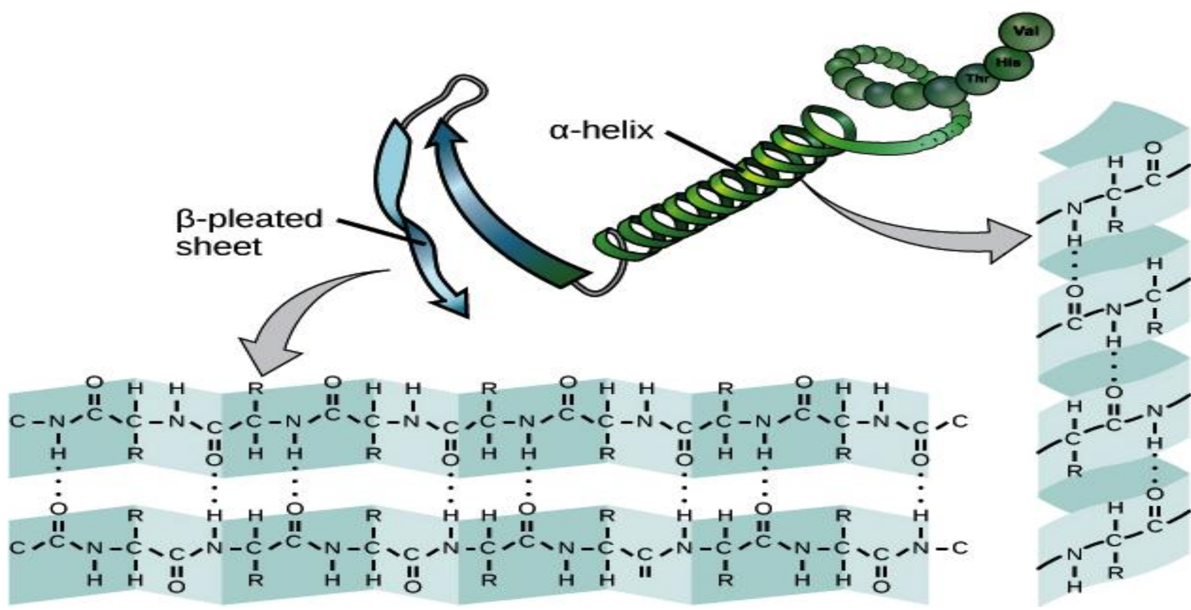


2. **Secondary structure:** In α helix, the carbonyl (C=O) of one amino acid is hydrogen bonded to the amino (H-N-H) of an amino acid that is four down the chain. (E.g., the carbonyl of amino acid 1 would form a hydrogen bond to the N-H of amino acid 5.) This pattern of bonding pulls the polypeptide chain into a helical structure that resembles a curled ribbon, with each turn of the helix containing 3.6 amino acids. The R

groups of the amino acids stick outward from α helix, where they are free to interact.

In β pleated sheet, two or more segments of a polypeptide chain line up next to each other, forming a sheet-like structure held together by hydrogen bonds. The hydrogen bonds form between carbonyl and amino groups of backbone, while the R groups extend above and below the plane of the sheet. The strands of a β pleated sheet may be parallel, pointing in the same direction (meaning that their N- and C-termini match up), or antiparallel, pointing in opposite directions (meaning that the N-terminus of one strand is positioned next to the C-terminus of the other).

Certain amino acids are more or less likely to be found in α-helices or β pleated sheets. For instance, the amino acid proline is sometimes called a “helix breaker” because its unusual R group (which bonds to the amino group to form a ring) creates a bend in the chain and is not compatible with helix formation. Proline is typically found in bends, unstructured regions between secondary structures. Similarly, amino acids such as tryptophan, tyrosine, and phenylalanine, which have large ring structures in their R groups, are often found in β pleated sheets, perhaps because the β pleated sheet structure provides plenty of space for the side chains.

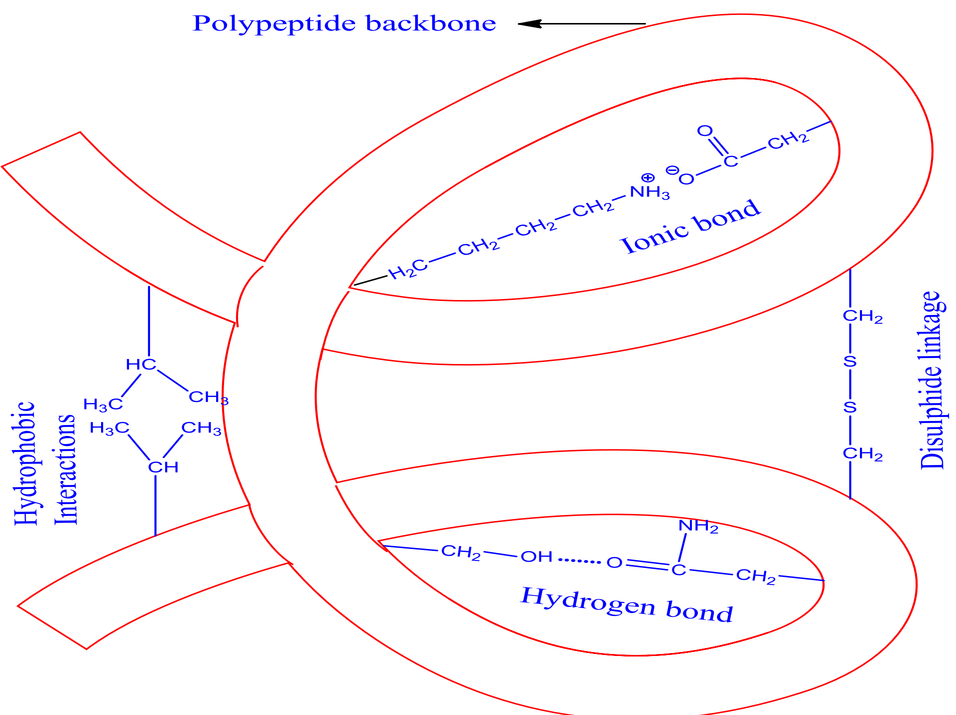


Secondary structure of Protein

3. Tertiary structure: The overall three-dimensional structure of a polypeptide is called its tertiary structure. The tertiary structure is primarily due to interactions between the R groups of the amino acids that make up the protein.

R group interactions that contribute to tertiary structure include hydrogen bonding, ionic bonding, dipole-dipole interactions, and London dispersion forces – basically, the whole gamut of non-covalent bonds. For example, R groups with same charges repel one another, while those with opposite charges can form an ionic bond. Similarly, polar R groups can form hydrogen bonds and other dipole-dipole interactions. One of the important characteristics of tertiary structure is hydrophobic interactions, in which amino acids with nonpolar, hydrophobic R groups cluster together on the inside of the protein, leaving hydrophilic amino acids on the outside to interact with surrounding water molecules.

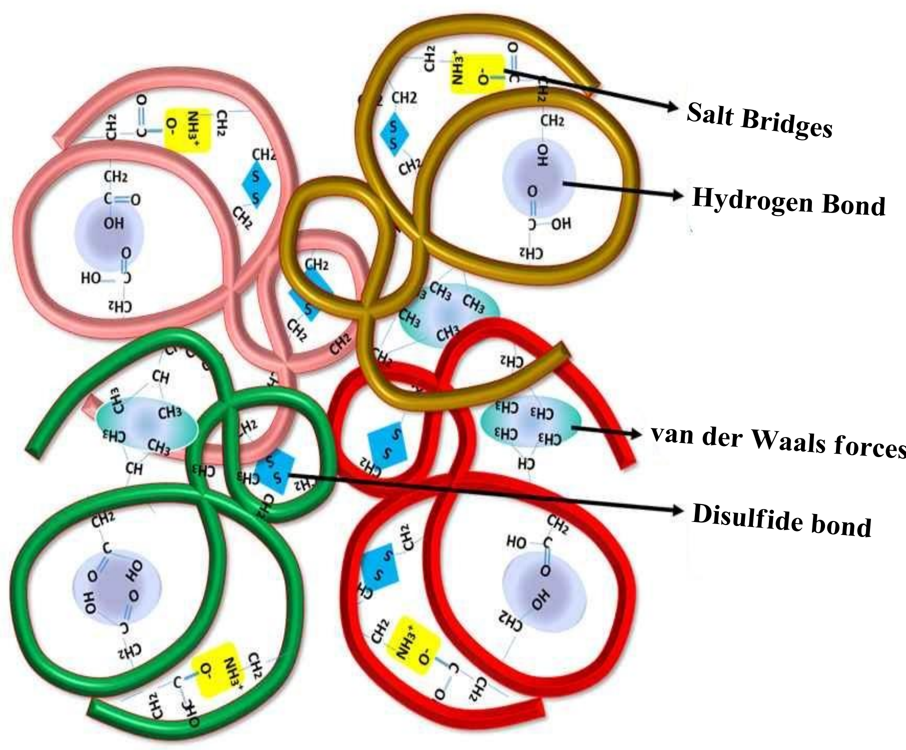
Finally, there's one special type of covalent bond that can contribute to tertiary structure: the disulfide bond. Disulfide bonds, covalent linkages between the sulfur-containing side chains of cysteines, are much stronger than the other types of bonds that contribute to tertiary structure. They act like molecular "safety pins," keeping parts of the polypeptide firmly attached to one another.



Tertiary structure of protein

Quaternary structure: Many proteins are made up of a single polypeptide chain and have only three levels of structure. However, some proteins are made up of multiple polypeptide chains, also known as subunits. When these subunits come together, they give the protein its quaternary structure.

Hemoglobin is one of the best example of quaternary protein structure. Hemoglobin carries oxygen in the blood and is made up of four subunits, two each of the α and β types. Another example is DNA polymerase, an enzyme that synthesizes new strands of DNA and is composed of ten subunits. In general, the same types of interactions that contribute to tertiary structure (mostly weak interactions, such as hydrogen bonding and London dispersion forces) also hold the subunits together to give quaternary structure.



Quaternary structure of protein

Nucleic Acids

Nucleic acid is a naturally occurring chemical compound that can be broken into smaller units such as phosphoric acid, sugars, and a mixture of organic bases (purines and pyrimidines). Nucleic acids are the main information-carrying molecules of the cell and by directing the process of protein synthesis; they determine the inherited characteristics of every living thing. The two main classes of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is the master blueprint for life and constitutes the genetic material in all free-living organisms and most viruses. RNA is the genetic material of certain viruses, but it is also found in all living cells, where it plays an important role in certain processes such as the making of proteins.

Types of Nucleic Acids

The two main types of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA is the genetic material found in all living organisms, ranging from single-celled bacteria to multicellular mammals. It is found in the nucleus of eukaryotes and in the chloroplasts and mitochondria. In prokaryotes, the DNA is not enclosed in a membranous envelope, but rather free-floating within the cytoplasm.

The entire genetic content of a cell is known as its genome and the study of genomes is genomics. In eukaryotic cells, but not in prokaryotes, DNA forms a complex with histone proteins to form chromatin, the substance of eukaryotic chromosomes. A chromosome may contain tens of thousands of genes. Many genes contain the information to make protein products; other genes code for RNA products. DNA controls all of the cellular activities by turning the genes "on" or "off."

The other type of nucleic acid, RNA, is mostly involved in protein synthesis. In eukaryotes, the DNA molecules never leave the nucleus but instead use an intermediary to communicate with the rest of the cell. This intermediary is the messenger RNA (mRNA). Other types of RNA—like rRNA, tRNA, and micro RNA-are involved in protein synthesis and its regulation.

Nucleotides

DNA and RNA are made up of monomers known as nucleotides. The nucleotides combine with each other to form a polynucleotide: DNA or RNA. Each nucleotide is made up of three components:

1. a nitrogenous base

2. a pentose (five-carbon) sugar

3. a phosphate group

Each nitrogenous base in a nucleotide is attached to a sugar molecule, which is attached to one or more phosphate groups.

Nitrogenous Base

The nitrogenous bases are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus, decreasing the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).

Adenine and guanine are classified as purines. The primary structure of a purine consists of two carbon-nitrogen rings. Cytosine, thymine, and uracil are classified as pyrimidines which have a single carbon-nitrogen ring as their primary structure. Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, the nitrogenous bases are simply known by their symbols A, T, G, C, and U. DNA contains A, T, G, and C whereas RNA contains A, U, G, and C.

Pyrimidines

Cytosine
C

Thymine
T

Uracil
U

Purines

Adenine
A

Guanine
G

Five-Carbon Sugar

The pentose sugar in DNA is deoxyribose and in RNA it is ribose. The difference between the sugars is the presence of the hydroxyl group on the second carbon of the ribose and hydrogen on the second carbon of the deoxyribose. The carbon atoms of the sugar molecule are numbered as 1′, 2′, 3′, 4′, and 5′ (1′ is read as "one prime").

Ribose

Deoxyribonucleoside

Deoxyribonucleotide

Ribonucleoside

Phosphate Group

The phosphate residue is attached to the hydroxyl group of the 5′ carbon of one sugar and the hydroxyl group of the 3′ carbon of the sugar of the next nucleotide, which forms a 5′3′ phosphodiester linkage. The phosphodiester linkage is not formed by simple dehydration reaction like the other linkages connecting monomers in macromolecules: its formation involves the removal of

two phosphate groups. A polynucleotide may have thousands of such phosphodiester linkages.

5'

Guanine

Cytosine

phosphodiester bond

Adenine

3'

Difference between nucleoside and nucleotide:

A nucleoside consists of a nitrogenous base covalently attached to a sugar (ribose or deoxyribose) but without the phosphate group. A nucleotide consists of a nitrogenous base, a sugar (ribose or deoxyribose) and one to three phosphate groups.

Nucleoside = Sugar + Base

Nucleotide = Sugar + Base + Phosphate

Comparatively it may be represented as follows:

Nucleotide	Nucleoside
The chemical composition of nucleotide consists of a phosphate group, a sugar and a nitrogenous base.	A nucleoside has a chemical composition that consists of a sugar and a base without the phosphate group.
They are one of the major causes for cancer causing agents to this very day.	They are used as agents in medicine that are primarily used against viruses and cancer causing agents.
Some of the major examples of nucleotides are adenosine, guanosine etc.	Some of the key examples of nucleosides are the same as nucleotides only with the addition of phosphate groups.

Components of nucleic acid: All the components of nucleic acid are summarized in following table.

	DNA only	DNA & RNA	RNA only
Nitrogen Bases	Thymine	Adenine, Guanine, Cytosine	Uracil
Sugars & Phosphate	2-Deoxyribose	Phosphate	Ribose

Nucleotides Structure

Primary Structure of Nucleic Acids

The sequence or order of the nucleotides defines the primary structure of DNA and RNA. The nucleotides of the polymer are linked by phosphodiester bonds connecting through the oxygen on the 5' carbon of one to the oxygen on the 3' carbon of another. The Oxygen and Nitrogen atoms in the backbone give DNA and RNA "polarity".

Secondary Structure of Nucleic Acids

In the early 1950's the primary structure of DNA was well established, but a firm understanding of its secondary structure was lacking. There were many discrepancies observed in establishing secondary structure of DNA and up to then it was conceded that the molar equivalences of base pairs (A & T and C &

G) discovered by Chargaff becomes an important factor. Researcher Rosalind Franklin, obtained X-ray diffraction evidence that suggested a long helical structure of uniform thickness. At the same time Francis Crick and James Watson, at Cambridge University, considered hydrogen bonded base pairing interactions, and arrived at a double stranded helical model that satisfied most of the known facts, and has been confirmed by subsequent findings.

Base pairing: Careful examination of the purine and pyrimidine base components of the nucleotides reveals that three of them could exist as hydroxy pyrimidine or purine tautomers, having an aromatic heterocyclic ring. Despite the added stabilization of an aromatic ring, these compounds prefer to adopt amide-like structures. The tautomerism can be represented by following scheme where right handed tautomer shows greater stability.

.

4-amino-2-hydroxypyrimidine (Cytosine)

2,4-dihydroxypyrimidines R=H, (Uracil) R=CH_3 (Thymine)

2-amino-6-hydroxypurine (Guanidine)

This tautomerism phenomenon can be understood taking 2-hydroxypyridine as an example where this structure might be expected to have phenol-like characteristics, such as an acidic hydroxyl group. However, the boiling point of the actual substance is 100º C greater than phenol and its acidity is 100 times less than expected (pKa = 11.7). These differences agree with the 2-pyridone tautomer, the stable form of the zwitterionic internal salt. Note that this tautomerism reverses the hydrogen bonding behavior of the nitrogen and oxygen functions (the N-H group of the pyridone becomes a hydrogen bond donor and the carbonyl oxygen an acceptor).

2-hydroxypyridine 2-pyridone

Once they had identified the favored base tautomers in the nucleosides, Watson and Crick were able to propose a complementary pairing, via hydrogen bonding, of guanosine (G) with cytidine (C) and adenosine (A) with thymidine (T). This pairing, which is shown in the following diagram, explained Chargaff's findings very clearly, and led them to suggest a double helix structure for DNA.

Hydrogen bonded base pairs: A purine base always pairs with a pyrimidine base or more specifically Guanosine (G) with Cytosine (C) and Adenine (A) with Thymine (T) or Uracil (U).

A==T G≡C

It may be observed from above structures that G-C pair has three hydrogen bonds while the A-T pair has two hydrogen bonds.

Double helical structure of DNA: After many trials and modifications, Watson and Crick conceived an ingenious double helix model for the secondary structure of DNA. Two strands of DNA were aligned anti-parallel to each other, i.e. with opposite 3' and 5' ends, as shown in part **a** of the following diagram.

Complementary primary nucleotide structures for each strand allowed intra-strand hydrogen bonding between each pair of bases. Coiling these coupled strands then leads to a double helix structure, shown as cross-linked ribbons in part b of the diagram. The double helix is further stabilized by hydrophobic attractions and pi-stacking of the bases.
The helix shown here has ten base pairs per turn, and rises 3.4 Å in each turn. This right-handed helix is the favored conformation in aqueous systems, and has been termed the B-helix. As the DNA strands wind around each other, they leave gaps between each set of phosphate backbones. Two alternating grooves result, a wide and deep major groove (*ca.* 22Å wide), and a shallow and narrow minor groove (*ca.* 12Å wide). Other molecules, including polypeptides, may insert into these grooves, and in so doing perturb the chemistry of DNA.

Double helix structure for DNA

4.4 Summary

In this unit learner able to classify amino acids, structure and stereochemistry of amino acids. Chemical properties of amino acids as well as functions of amino acids. Stereochemistry, spectral analysis of amino acids and acid-base behavior of amino acids. Some important applications of amino acids as Isoelectric Point (pI) and Electrophoresis.

Preparation and reactions of α-amino acid, reagents used for the identification of Amino acids and determination of the structures of Polypeptide. Analysis peptides via sequencing the peptide-terminal residue analysis and peptide synthesis. Some introduction and classification of proteins and difference between polypeptides and proteins. About structures of proteins. Introduction and types of Nucleic Acids and difference between nucleoside and nucleotide. Components of nucleic acid, primary structure of nucleic acids and secondary structure of nucleic acids

4.5 Terminal Questions

Q1. Chemical properties of amino acids?

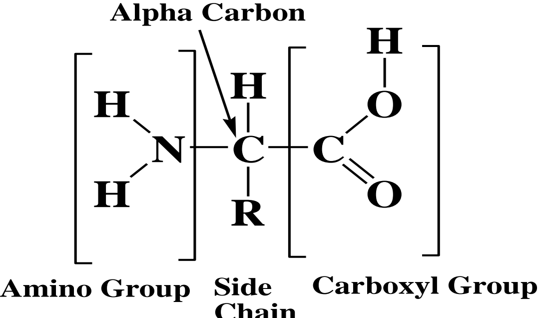
Ans. Amino acids are organic molecules which on linking together with other amino acids forms protein molecule. Amino acids are essential to life because the proteins they form are involved in virtually all cell functions. Some proteins function as enzymes, antibodies, while others provide structural support. Although there are hundreds of amino acids found in nature, only about 20 amino acids are needed to make all the proteins found in the human body.

Main highlights of amino acids

- Almost all cell functions involve proteins. These proteins are composed of organic molecules called amino acids.
- While there are many different amino acids in nature, all proteins are formed from twenty amino acids.
- From a structural perspective, amino acids are typically composed of a carbon atom, a hydrogen atom, a carboxyl group along with an amino group and a variable group.
- Based on the variable group, amino acids can be classified into four major categories: nonpolar, polar, negatively charged, and positively charged amino acids.

Out of twenty amino acids, eleven can be made naturally by the body and are termed nonessential amino acids. Those amino acids which can not be naturally synthesized by our body are called essential amino acids.

An amino acid structure:



Structural explanation of amino acids

Generally, amino acids have the following structural properties:

- A carbon (C, essentially alpha carbon)
- A hydrogen atom (H)
- A Carboxyl group (-COOH)
- An Amino group (-NH_2)
- A "variable" group or "R" group

Q2. Spectral analysis of amino acids

Ans.

- Spectral analysis refers to the ability of amino acids to absorb or emit electromagnetic radiations at different wavelengths (i.e. energies)
- No any amino acid absorbs light in the visible spectrum which represents that they are "colorless"
- If proteins have color e.g. hemoglobin is red; it is because they contain iron in the structure
- All the amino acids are absorbed in the infrared region (longer wavelengths, weaker energy than visible light)
- Some amino acids are absorbed in the ultraviolet spectrum (shorter wavelengths, higher energy than visible light)
- Electrons in aromatic ring structures absorb in the UV spectrum. Such structures comprise the side chains of tryptophan, tyrosine and phenylalanine

Q3. **Isoelectric Point (pI)**

The pH at which all amino acids are in the zwitterions form and contains very low and equal concentrations of the positive and negative ions is known as Isoelectric point.

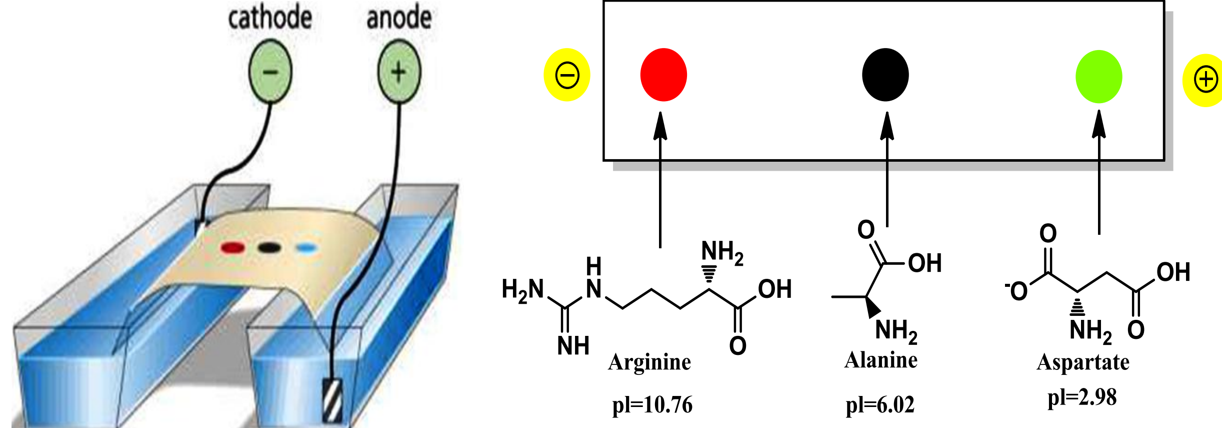
- In more acidic media (pH<pI), the concentration of ions increases while the concentration of the zwitterion decreases.
- In a more basic media (pH>pI), the concentration of negative ions increases while the concentration of the zwitterion decreases.
- More than 98% of the amino acid is in zwitterion from over a pH range of ±2 of pI.
- At pH=pI, the amino acid present as the zwitterion with one amine or carboxyl group in uncharged form.
- Isoelectric points found at the values ranging from 7.8 to 10.8 (basic)
- Isoelectric points found at the values ranging from 4.86 to 6. 3 (neutral)
- Isoelectric points found at the values ranging from 2.8 to 3.3 (acidic)

Isoelectric Points of some amino acids

Amino Acid	**Isoelectric point (pl)**
Arganine (Arg)	10.8
Lysine (Lys)	9.7
Alanine (Ala)	6.0
Glycine (Gly)	6.0
Serine (Ser)	5.7
Glutamic acid (Glu)	3.2
Aspartic acid (Asp)	2.9

Q5. Use of Amino acids in Electrophoresis

Electrophoresis is an analytical method for identifying amino acids by observing their migration as a function of pH under an applied electric field gradient. At its pI, the amino acid is present in the zwitterion form with no net charge′ and will not migrate in electrophoresis. At pH < pI, the amino acid carries a positive charge and will migrate to the negative electrode. At pH, pI, the amino acid carries a negative charge and will migrate to the positive electrode.



Q6. Define Nitrogenous Bases?

Ans.

The nitrogenous bases are organic molecules and are so named because they contain carbon and nitrogen. They are bases because they contain an amino group that has the potential of binding an extra hydrogen, and thus, decreasing the hydrogen ion concentration in its environment, making it more basic. Each nucleotide in DNA contains one of four possible nitrogenous bases: adenine (A), guanine (G) cytosine (C), and thymine (T).

Adenine and guanine are classified as purines. The primary structure of a purine consists of two carbon-nitrogen rings. Cytosine, thymine, and uracil are classified as pyrimidines which have a single carbon-nitrogen ring as their primary structure. Each of these basic carbon-nitrogen rings has different functional groups attached to it. In molecular biology shorthand, the nitrogenous bases are simply known by their symbols A, T, G, C, and U. DNA contains A, T, G, and C whereas RNA contains A, U, G, and C.

Pyrimidines

Cytosine
C

Thymine
T

Uracil
U

Purines

Adenine
A

Guanine
G

UNIT 5: ACTIVE METHYLENE COMPOUNDS

5.1 Introduction

Objective

5.2 Properties of Active Methylene Compounds

5.3 Tautomerism of Active Methylene Compounds

5.4 Preparation and synthetic applications of ethyl acetoacetate

5.5 Preparation and synthetic applications of diethyl malonate

5.6 Summary

5.7 Terminal Questions

5.1 Introduction

We are very well aware of methane, methyl group and methyne group but methylene group are not studied in previous classes. In this unit we would study

about active methylene compound, preparation and application of their derivatives.

Active methylene compounds are those compounds in which methylene group (-CH_2-) is attached with two electron withdrawing groups (EWG). Examples are; ethyl acetoacetate (EAA), malonic ester, etc.

$$H_2C(COCH_3)(COOC_2H_5)$$

Ethyl acetoacetate (EAA)

$$H_2C(COOC_2H_5)_2$$

Malonic ester

Objective

- Tautomerism of Active Methylene Compounds
- Preparation and synthetic applications of ethyl acetoacetate
- Preparation and synthetic applications of diethyl malonate

5.2 Properties of Active Methylene Compounds

Most common electron withdrawing groups are;

$-\overset{O}{\overset{\|\|}{C}}-$	$-\overset{O}{\overset{\|\|}{C}}-OEt$	$-\overset{O}{\overset{\|\|}{C}}-NH_2$	$-CN$	$-COOH$	$-NO_2$	All -M groups
Keto group	Ester group	Amide group	Cyano group	Carboxyl group	Nitro group	

Graphically active methylene compounds may be represented as follows;

$$H_2C\begin{matrix} \diagup \text{EWG} \\ \diagdown \text{EWG} \end{matrix}$$

Due to presence of two electron withdrawing groups on methylene carbon, both hydrogens of methylene carbon become acidic in nature.

$$H_3C-\overset{O}{\overset{\|}{C}}-CH_3 \xrightarrow[OH^{\ominus}]{Base} \left[H_3C-\overset{O}{\overset{\|}{C}}-\overset{\ominus}{C}H_2 \longleftrightarrow H_3C-\overset{O^{\ominus}}{\overset{|}{C}}=CH_2 \right]$$

Ketone — Carbanion-I — Carbanion-II (Enolate ion of ketone)

$$R-O-\overset{O}{\overset{\|}{C}}-CH_3 \xrightarrow[OH^{\ominus}]{Base} \left[R-O-\overset{O}{\overset{\|}{C}}-\overset{\ominus}{C}H_2 \longleftrightarrow R-O-\overset{O^{\ominus}}{\overset{|}{C}}=CH_2 \right]$$

Ester — (Enolate of ester)

5.3 Tautomerism of Active Methylene Compounds

1. Resonating structure (B) and (C) is more stable than (A) because negative charge is present on electronegative oxygen atom.
2. Resonating structure (B) is more stable than (C) because C=O group of carbonyl compound have more stabilised negative charge than the C=O group of ester. The reason behind this is; the C=O group of ester stabilizes both negative charge of carbon atom as well as loan pair of oxygen atom.

(B)

Base

$C_2H_5\overset{\ominus}{O}\,\overset{\oplus}{Na}$

Ethyl acetoacetate (EAA)

(A)

Stability order: (B) > (C) > (A)

Reactivity order: (A) > (C) > (B)

(C)

Acidic nature of methylene group: Methylene group in EAA is flanked by two electron withdrawing groups, viz, an acetyl group and an ester group. The hydrogens of this methylene group are ionisable due to the electron withdrawing effect of the surrounding groups. Also, the negative ion obtained after losing the proton gets stabilized due to resonance which is schematically represented below.

$H_3C-\overset{O}{\overset{\|}{C}}-\underset{H_2}{C}-\overset{O}{\overset{\|}{C}}-OCH_2CH_3 \underset{}{\overset{H^+}{\rightleftharpoons}}$

Active methalene group

$H_3C-C=C=C-OCH_2CH_3$ (with O, O and $\ominus$ above)

Active methalene group

Equivalent to $\equiv$

$[\, CH_3OC(=O)-\overset{\ominus}{C}H-C(=O)OC_2H_5 \leftrightarrow CH_3OC(\overset{\ominus}{O})=CH-C(=O)OC_2H_5 \leftrightarrow C_2H_5OC(=O)-CH=C(\overset{\ominus}{O})OC_2H_5 \,]$

Characteristics of Tautomerism

1. When two structural isomers are mutually interconvertible and exist in dynamic equilibrium, they are called **“TAUTOMERS”**.
2. Tautomers are discrete chemical entities, capable of isolation under suitable conditions.
3. Tautomers differ from each other in stability. The less stable form is called labile form. The relative proportion of two forms in a tautomeric mixture varies from compound to compound and also with temperature, solvent etc. Tautomeric transformations are also catalyzed by acids and bases.
4. Tautomers exist in dynamic equilibrium. Their separation can be achieved only by special methods.

Difference between Tautomerism and Resonance: As a matter of fact there is no point of similarity between tautomerism and resonance. The former is the phenomenon describing the dynamic equilibrium between isomers whereas the latter is a concept for representing structure of molecules. Some of the important differences have been pointed out below:

1. Tautomerism involves tautomers which have real existence whereas the resonance involves contributing structures which are hypothetical and do not exist.
2. Tautomerism involves the migration of atoms whereas for resonance to occur all the atoms must occupy the same position in all the contributing

structures. Contributing structures differ from each other in the placement of electrons only.

3. The tautomers are in dynamic equilibrium. Since the contributing structures in resonance are hypothetical hence no such equilibrium is possible.

4. The resonance involves shortening of single bond lengths and lengthening of double bond lengths but this is not so ¡n tautomerism.

5. The resonance lowers the energy of real molecule (i.e. stabilizes it) but this does not happen in tautomerism.

6. The molecule should be planar for resonance to occur while it is not necessary in tautomerism.

7. All the molecules of compound exhibiting resonance have the same structure, regardless of the fact whether that can be represented or not, whereas compounds exhibiting tautomerism have two or more types of molecules present as equilibrium mixture.

8. The tautomerism is shown by placing reversible arrows ($\rightleftharpoons$) sign of equilibrium between tautomers whereas resonance is depicted by placing double headed arrow (↔) between contributing structures.

5.3 Preparation and synthetic applications of Ethyl Acetoacetate (EAA) or Aceto acetic ester (AAE)

EAA offers a classical and most thoroughly investigated example of keto-enol tautomerism.

$$H_3C-\underset{\|}{\overset{}{C}}(=O)-CH_2-C(=O)-OEt \rightleftharpoons H_3C-C(OH)=CH-C(=O)-OCH_2CH_3$$

Keto form   Enol form

This compound was first prepared by Geuther who assigned it the enol structure, while Frankland and Duppa showed that EAA had keto structure. The presence of each of the keto and enol forms in EAA was supported by two sets of reactions.

Method of preparation: Acetoacetic ester is synthesized by self-condensation of ethyl acetate in presence of a base (Claisen ester condensation).

$$2\ H_3C-\overset{\overset{\displaystyle O}{\|}}{C}-OC_2H_5 \xrightarrow[\text{(ii) HCl}]{\text{(i) } C_2H_5\overset{\ominus}{O}\overset{\oplus}{Na}} H_2C\begin{cases} \overset{\overset{\displaystyle O}{\|}}{C}-CH_3 \\ \underset{\underset{\displaystyle O}{\|}}{C}-OC_2H_5 \end{cases}$$

(Ethyl acetate)

Ethyl acetoacetate (EAA)

Mechanism:

1. In the mechanism, first three steps are reversible and equilibrium favour backward reaction. This is because of less stability of enolate ester.

2. Step four is mainly responsible for the Claisen-ester condensation. In this step resonance stabilized carbanion intermediate form and the reaction becomes almost irreversible towards forward direction.

3. Only the esters which contains at least two α-hydrogen gives Claisen-ester condensation reaction.

$H_3C-C(=O)-OC_2H_5$ Ethyl acetate (First molecule) $\underset{\text{Step-1}}{\overset{C_2H_5O^{\ominus}}{\rightleftharpoons}}$ [$H_2\overset{\ominus}{C}-C(=O)-OC_2H_5 \longleftrightarrow H_2C=C(O^{\ominus})-OC_2H_5$] Resonance stabilized carbanion

$H_3C-C(=O)-OC_2H_5$ Ethyl acetate (Second molecule)

Step-2

$H_3C-C(O^{\ominus})(OC_2H_5)-CH_2-C(=O)-OC_2H_5$

Step-3 $-C_2H_5O^{\ominus}$

$H_3C-C(=O)-CH_2-C(=O)-OC_2H_5$

Step-4 $C_2H_5O^{\ominus}$, $-C_2H_5OH$

$H_3C-C(=O)-CH=C(O^{\ominus})-OC_2H_5$

HCl

[$H_3C-C(=O)-\overset{\ominus}{C}H-C(=O)-OC_2H_5 \longleftrightarrow H_3C-C(O^{\ominus})=CH-C(=O)-OC_2H_5$] Resonance stabilized carbanion intermediate $\longleftrightarrow$ $H_3C-C(=O)-CH_2-C(=O)-OC_2H_5$ Ethyl acetoacetate (EAA)

Physical Properties of EAA

1. Ethyl acetoacetate is colorless liquid and has fruity odour.
2. Boiling point is 181°C
3. Sparingly soluble in water but readily soluble in ethanol, ether and most organic solvents.
4. Neutral to litmus.

5. Soluble in dilute NaOH and it is enol form which dissolves to give sodium salt.
6. Refractive index is 1.4232.
7. Gives reddish violate colour with $FeCl_3$.

Keto-Enol tautomerism of ethyl acetoacetate: The structural isomers which are in equilibrium with each other and are capable of rapidly changing into each other when the equilibrium is disturbed, are known as tautomers and this phenomenon is known as tautomerism. If one of the tautomer is in keto form and other in enol form; then tautomerism is called keto-enol tautomerism.

Aceto acetic ester is a mixture of keto and enol form and hence shows keto-enol tautomerism. The reason behind keto-enol tautomerism of ethyl acetoacetate is due to the presence of intra-molecular H-bonding but equilibrium slightly shifted towards keto form. Carbonyl group is more electron withdrawing group than ester, hence reaction goes through carbonyl side.

Intramolecular H-bonding

Keto form

Enol form

Chemical properties of EAA:

Reactions supporting keto form of EAA:

(a) Reaction with $NaHSO_3$ and HCN: EAA forms addition products with HCN and $NaHSO_3$, indicating presence of carbonyl (>C=O) group.

$H_3C-\overset{O}{\overset{\|}{C}}-\overset{H_2}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5$ — Addition reaction

NaHSO₃ → $H_3C-\underset{SO_3Na}{\overset{OH}{C}}-\overset{H_2}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5$

Bisulphite

HCN → $H_3C-\underset{CN}{\overset{OH}{C}}-\overset{H_2}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5$

Cyanohydrin

Mechanism:

$H_3C-\overset{O}{\overset{\|}{C}}-\overset{H_2}{C}-\overset{O}{\overset{\|}{C}}-OEt + H-CN \longrightarrow H_3C-\underset{CN}{\overset{OH}{C}}-\overset{H_2}{C}-\underset{O}{\underset{\|}{C}}-OEt$

$EtO-\overset{O}{\overset{\|}{C}}-\overset{H_2}{C}-\overset{O}{\overset{\|}{C}}-CH_3 + HO-\overset{O}{\overset{\|}{S}}-O^- Na^+ \longrightarrow H_3C-\underset{SO_3Na}{\overset{OH}{C}}-\overset{H_2}{C}-\underset{O}{\underset{\|}{C}}-OEt$

(b) Reaction with hydroxylamine and phenyl hydrazine: It reacts with hydroxylamine (NH_2OH) and phenyl hydrazine ($PhNHNH_2$) to form oxime and phenyl hydrazone respectively. Formation of oxime and hydrazone is characteristic of compounds containing ketone group. This reaction indicates presence of keto group in EAA.

$CH_3CH_2CO-CH_2-C(=O)-CH_3$ (with C=O) + $HO-NH_2$ (Hydroxyl amine) $\xrightarrow{-H_2O}$ $CH_3CH_2CO-CH_2-C(=N-OH)-CH_3$

Hydroxyl amine

Oxime

$CH_3CH_2CO-CH_2-C(=O)-CH_3$ + $C_6H_5-NH-NH_2$ $\xrightarrow{-H_2O}$ $CH_3CH_2CO-CH_2-C(=N-NH-C_6H_5)-CH_3$

Phenyl hydrazine

Phenyl hydrazone

(c) Hydrolysis of ethyl acetoacetate: EAA is hydrolysed to acetone when treated with dilute acid or alkali, which indicates presence of –CH2- group.

$$H_3C-CO-CH_2-CO-OCH_2CH_3 \xrightarrow{\text{Acid/Alkali}} H_3C-CO-CH_3 + CO_2 + C_2H_5OH$$

Acetone

(d) Reaction with sodium ethoxide: It forms mono and dialkyl derivative indicating the presence of active methylene group.

$$H_3C-CO-CH_2-CO-OCH_2CH_3 \xrightarrow[\text{(ii) RI}]{\text{(i) } C_2H_5ONa} H_3C-CO-CH(R)-CO-OCH_2CH_3$$

Monoalkyl derivative

(e) In ethyl acetoacetate, carbonyl carbon is more electrophilic than carbonyl carbon of ester because alkoxy group (RO-) of ester decreases electrophilicity of carbonyl carbon of ester by increasing π electron density (+M effect).

(f) **Reduction of EAA:** $NaBH_4$ is a mild reducing agent, so it reduces only carbonyl group whereas $LiAlH_4$ is a strong reducing agent, so it reduces both carbonyl group as well as ester group.

$H_3C{-}CO{-}CH_2{-}CO{-}OC_2H_5 \xrightarrow[(H^+)]{NaBH_4} H_3C{-}CH(OH){-}CH_2{-}CO{-}OC_2H_5$

$H_3C{-}CO{-}CH_2{-}CO{-}OC_2H_5 \xrightarrow{LiAlH_4} H_3C{-}CH(OH){-}CH_2{-}CH_2OH$

Strong reducing agent

$H_3C{-}CO{-}CH_2{-}CO{-}OC_2H_5 \xrightarrow[HCl]{Na\text{-}Hg} H_3C{-}CH(OH){-}CH_2{-}CO{-}OC_2H_5$

Strong electrophilic centre

Weak electrophilic centre

+M effect decreases electrophilicity of C=O group of ester

oxygen of ethoxide gives his lone pair to more electrophilic carbon atom

(g) **Reaction with $PhNHNH_2$ / NH_2OH:**

$$CH_3COCH_2COOC_2H_5 \xrightarrow{PhNH.NH_2} H_3C-C(=N-NHPh)-CH_2-C(=O)-OC_2H_5 + H_2O$$

$$\xrightarrow{-C_2H_5OH} \text{3-methyl-1-phenyl-pyrazol-5-one} \xrightarrow{CH_3I \text{ Alkaline solution}} \text{Antipyrine}$$

$$CH_3COCH_2COOC_2H_5 \xrightarrow{H_2N-OH} H_3C-C(=N-OH)-CH_2-C(=O)-OC_2H_5 + H_2O$$

$$\xrightarrow{-C_2H_5OH} H_3C-C=N-N(Ph)-C(=O)-CH_2 \text{ (ring)}$$

All the above reactions indicate that EAA exists in keto form.

2. Reactions supporting keto form of EAA:

(a) Reaction with sodium ethoxide:

$$H_3C-C(OH)=CH-C(=O)-OC_2H_5 \xrightarrow{C_2H_5O^{\ominus} Na^{\oplus}} H_3C-C(O^{\ominus}Na^{\oplus})=CH-C(=O)-OC_2H_5 + C_2H_5OH$$

(b) Reaction with bromine: Bromine gives anti-addition on double bond of enolic form of ethyl acetoacetate.

$$H_3C-C(OH)=CH-C(=O)-OC_2H_5 \xrightarrow{Br_2} H_3C-C(OH)(Br)-CH(Br)-C(=O)-OC_2H_5$$

alpha, beta-dibromoacetoacetic ester

(c) Reaction with organomagnesium and organolithium compounds:

$$H_3C-C(OH)=CH-C(=O)-OC_2H_5 \xrightarrow[-R\text{-}H]{R\text{-}MgX} H_3C-C(OMgX)=CH-C(=O)-OC_2H_5$$

$$H_3C-C(OH)=CH-C(=O)-OC_2H_5 \xrightarrow[-R\text{-}H]{R\text{-}Li} H_3C-C(OLi)=CH-C(=O)-OC_2H_5$$

(d) Reaction with diazomethane: Diazomethane gives methylation with enol form of EAA.

$$H_3C-C(OH)=CH-C(=O)-OC_2H_5 \xrightarrow[-N_2]{CH_2N_2} H_3C-C(O^{\ominus})=CH-C(=O)-OC_2H_5 \longrightarrow H_3C-C(OCH_3)=CH-C(=O)-OC_2H_5$$

Ethyl crotonate

(e) Reaction with urea: EAA forms 4-methyl uracil with urea.

$\xrightarrow[-C_2H_5OH]{-H_2O}$

4-methyl uracil

(f) Alkylation of EAA:

$$H_3C-\overset{O}{\overset{\|}{C}}-\overset{H_2}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5 \xrightarrow{C_2H_5\overset{\ominus}{O}\,\overset{\oplus}{Na}} H_3C-\overset{O}{\overset{\|}{C}}-\overset{\ominus}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

$$\updownarrow$$

$$H_3C-\overset{O}{\overset{\|}{C}}-\underset{H}{C}=\overset{\overset{\ominus}{O}}{\overset{|}{C}}-OC_2H_5 \longleftrightarrow H_3C-\overset{\overset{\ominus}{O}}{\overset{|}{C}}=\underset{H}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

Elimination reaction

Tertiary alkyl hailide

$$H_3C-\underset{CH_3}{\overset{CH_3}{C}}-X$$

$$H_3C-\overset{CH_3}{\overset{|}{C}}=CH_2$$

Primary / secondary alkyl hailide

$$H_3C-\overset{O}{\overset{\|}{C}}-\underset{R}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

Monosubstituted

$$\Big\downarrow\, C_2H_5\overset{\ominus}{O}\,\overset{\oplus}{Na}$$

$$H_3C-\overset{O}{\overset{\|}{C}}-\underset{R}{\overset{R}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5 \xleftarrow{R-X} H_3C-\overset{O}{\overset{\|}{C}}-\underset{R}{\overset{\ominus}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

Disubstituted

Note: In the alkylation of ethyl acetoacetate either 1° or 2° alkyl halide is used while 3° alkyl halide is not used because it gives elimination reaction with conjugated base of ethyl acetoacetate.

All the above reactions indicate that EAA exists in enol form.

Problem: How to convert…???

$$CH_3COCH_2COOC_2H_5 \xrightarrow{(CH_3)_3C-Br \,/\, AgNO3} CH_3CO-CH(C(CH_3)_3)-COOC_2H_5$$

ethyl 3-oxobutanoate

ethyl 2-acetyl-3,3-dimethylbutanoat

Mechanism:

$$H_3C-CO-CH_2-CO-OC_2H_5 \rightleftharpoons H_3C-C(OH)=CH-CO-OC_2H_5$$

Keto form

Enol form (behaves as base)

$$(CH_3)_3C-Br + AgNO3 \longrightarrow (CH_3)_3C^{\oplus} + AgBr$$

$$H_3C-C(OH)=CH-CO-OC_2H_5 + H_2C-C(CH_3)_2-Br \longrightarrow H_3C-CH(OH^{\oplus})-CH(C(CH_3)_3)-CO-OC_2H_5$$

$$\xrightarrow{-H^{\oplus}} H_3C-CO-CH(C(CH_3)_3)-CO-OC_2H_5$$

Problem: How to convert….???

$$H_3C-\overset{OH}{\overset{|}{C}}=\underset{H}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5 \xrightarrow{\text{Phenylation}} H_3C-\overset{O}{\overset{\|}{C}}-\underset{Ph}{\underset{|}{\overset{H}{C}}}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

ethyl 3-hydroxybut-2-enoate → ethyl 3-oxo-2-phenylbutanoate

Mechanism:

$$H_3C-\overset{OH}{C}=\underset{H}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5 \xrightarrow{C_2H_5\overset{\ominus}{O}\ \overset{\oplus}{Na}} H_3C-\overset{O}{\overset{\|}{C}}-\overset{\ominus}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5 + C_2H_5OH$$

$Br^{\oplus}$ ⟷ Br (bromobenzene)

No Reaction

$$C_6H_4 \text{ (benzyne)} + H_3C-\overset{OH}{\underset{H}{C}}-\overset{\ominus}{\overset{H}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5 \longrightarrow H_3C-\overset{O}{\overset{\|}{C}}-\overset{H}{\underset{C_6H_4^{\ominus}}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

$$\xrightarrow{H^{\oplus}} H_3C-\overset{O}{\overset{\|}{C}}-\overset{H}{\underset{C_6H_5}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

$$C_6H_4(F)(H) + H_3C-\overset{OH}{C}=\underset{H}{C}-\overset{O}{\overset{\|}{C}}-OC_2H_5 \xrightarrow{NaNH_2} H_3C-\overset{O}{\overset{\|}{C}}-\overset{H}{\underset{C_6H_5}{C}}-\overset{O}{\overset{\|}{C}}-OC_2H_5$$

Properties of Keto and Enol forms:

Keto form	Enol form
1. Long colorless needles. M.p. -39°C	1. Colorless oily liquid at -78°C
2. Refractive index is 1.4171	**2.** Refractive index is 1.4432
3. Density is 1.0368[10]	**3.** Density is 1.0119[10]
4. No coloration with $FeCl_3$	**4.** Intense coloration with $FeCl_3$ (red-violet)
5. It decolorizes bromine solution	**5.** It does not decolorizes bromine solution

5.4 Preparation and synthetic applications of Diethyl malonate $CH_2(COOC_2H_5)$

Diethyl malonate is also known as malonic ester. It is prepared by boiling sodium or potassium cyanoacetate with alcohol and concentrated hydrochloric acid. Diethyl malonate is a diester derivative of malonic acid, a dicarboxylic acid with two carboxyl groups (-COO-) separated by one methylene group (-CH_2-). Diethyl malonate is formed by the replacement of the hydroxyl groups (-OH) of malonic acid with ethoxy groups (-OCH_2CH_3). The hydrogen atoms on the methylene carbon between the two carboxyl groups make this compound acidic. Because of its unique structure, diethyl malonate is reactive and functions as a reagent for organic synthesis and to make products such as barbiturates, pigments, and agrochemicals. Volatile esters are known to have fruity scents and are often used as fragrances and flavorings. Diethyl malonate is a volatile diester that occurs naturally in fruits such as grapes, strawberries, guava, melon, pineapple, and blackberries.

$$H_2C(CN)(COOK) + 2H_2O + 2HCl \longrightarrow H_2C(COOH)(COOH) + KCl + NH_4Cl$$

Potassium cyanoacetate — Malonic acid

$$H_2C(COOH)(COOH) + C_2H_5OH \xrightarrow{HCl} H_2C(COOC_2H_5OH)(COOC_2H_5OH) + 2H_2O$$

Malonic acid — Diethyl malonate (Malonic ester)

Potassium cyanoacetate required for above reaction is obtained by following method.

$$CH_3COOH \xrightarrow[\text{(HVZ reaction)}]{Cl2/P} Cl{-}CH_2OOH \xrightarrow{K_2CO_3} 2HC(Cl)(COOK) \xrightarrow{KCN} 2HC(CN)(COOK)$$

Potassium cyanoacetate

Physical properties:

- Character colorless liquid, with a sweet ether odor
- Melting point -50℃
- Boiling point 199.3℃
- Relative density 1.0551
- Refractive index 1.4135
- Flash point 100℃
- Diethyl malonate is miscible with alcohols and ethers, soluble in chloroform, benzene and other organic solvents. It is slightly soluble in water (2.08g/100ml at 20°C).

Chemical Properties:

Like ethylacetoacetate, diethyl malonate contains a methylene group joined to two carbonyl groups. The H atom of the CH_2 group is acidic. This is attributed to two factors:

(a) The electron attracting power of the electronegative oxygen of the carbonyl group (inductive effect) and

(b) The resonance stabilization of the resultant cation.

$$C_2H_5OC(=O)-CH_2-C(=O)OC_2H_5 \rightleftharpoons H^+ + C_2H_5OC(=O)-\overset{-}{C}H-C(=O)OC_2H_5$$

Diethyl malonate   Diethyl malonate anion

The diethyl malonate anion is highly resonance stabilized so that so that its negative charge is delocalized into the two carbonyl groups.

$$H_3C-C(=O)-CH_2-C(=O)-OCH_2CH_3$$

Active methalene group

H^+

$$CH_3OC(=O)-\overset{\ominus}{C}H-COC_2H_5 \leftrightarrow CH_3OC(-O^{\ominus})=CH-C(=O)OC_2H_5 \leftrightarrow C_2H_5OC(=O)-CH=C(-O^{\ominus})OC_2H_5$$

Equivalent to

$$H_3C-C(O)=C=C(O)-OCH_2CH_3 \ (\ominus)$$

Active methalene group

1. **Salt formation:** Diethyl malonate mostly reacts with strong bases and forms salt because CH_2 group is acidic in nature.

$$H_2C(COOC_2H_5)_2 + C_2H_5ONa \longrightarrow \overset{+}{Na}\ \overset{-}{C}H(COOC_2H_5)_2 + C_2H_5OH$$

Diethyl malonate (Malonic ester) → Sodium diethyl malonate

2. Alkylation: Diethyl malonate anion works as good nucleophile and reacts with halogens and forms diethyl alkylmalonate.

$$\overset{+}{Na}\ \overset{-}{C}H_2(COOC_2H_5)_2 + R-X \longrightarrow R-\underset{}{\overset{H}{C}}(COOC_2H_5)_2 + NaBr$$

Sodium diethyl malonate → Diethyl alkylmalonate

3. **Hydrolysis and Decarboxylation:** Diethyl malonate undergoes hydrolysis with dilute hydrochloric acid and gives malonic acid. Those acids which have two -COOH groups separated by a carbon, on heating gives monocarboxylic acid.

$$R-HC(COOC_2H_5)_2 \xrightarrow{\overset{\oplus}{H} / H_2O} R-HC(COOH)_2 \xrightarrow[-CO_2]{\triangle} R-\overset{H_2}{C}-COOH$$

Diethyl alkylmalonate → Alkylacetic acid

5.6 Summary

After Studying this unit learner may able to explain

get knowledge of Active Methylene Compounds Structure, Properties including physical and chemical and Tautomerism in these compounds. Also come to explain Preparation and synthetic applications of ethyl acetoacetates and diethyl malonates and there derivatives.

5.7 Terminal Questions

Q1. Give Physical Properties of Ethyl acetoacetate (EAA)?

Ans.

1. Ethyl acetoacetate is colorless liquid and has fruity odour.
2. Boiling point is 181°C
3. Sparingly soluble in water but readily soluble in ethanol, ether and most organic solvents.
4. Neutral to litmus.
5. Soluble in dilute NaOH and it is enol form which dissolves to give sodium salt.
6. Refractive index is 1.4232.
7. Gives reddish violate colour with $FeCl_3$.

Q2. Give mechanism of Keto-Enol tautomerism of ethyl acetoacetate.

Ans. The structural isomers which are in equilibrium with each other and are capable of rapidly changing into each other when the equilibrium is disturbed, are known as tautomers and this phenomenon is known as tautomerism. If one of the tautomer is in keto form and other in enol form; then tautomerism is called keto-enol tautomerism.

Aceto acetic ester is a mixture of keto and enol form and hence shows keto-enol tautomerism. The reason behind keto-enol tautomerism of ethyl acetoacetate is due to the presence of intra-molecular H-bonding but equilibrium slightly shifted towards keto form. Carbonyl group is more electron withdrawing group than ester, hence reaction goes through carbonyl side.

Intramolecular H-bonding

Keto form

Enol form

Q3. Method of preparation of Acetoacetic ester?

Ans. Acetoacetic ester is synthesized by self-condensation of ethyl acetate in presence of a base on Claisen ester condensation.

$$2\ H_3C-\overset{\overset{\displaystyle O}{\|}}{C}-OC_2H_5 \xrightarrow[\text{(ii) HCl}]{\text{(i) } C_2H_5\overset{\ominus}{O}\overset{\oplus}{Na}} H_2C\begin{matrix} \diagup \overset{\overset{\displaystyle O}{\|}}{C}-CH_3 \\ \diagdown \underset{\underset{\displaystyle O}{\|}}{C}-OC_2H_5 \end{matrix}$$

(Ethyl acetate)

Ethyl acetoacetate (EAA)

Mechanism:

1. In the mechanism, first three steps are reversible and equilibrium favour backward reaction. This is because of less stability of enolate ester.
2. Step four is mainly responsible for the Claisen-ester condensation. In this step resonance stabilized carbanion intermediate form and the reaction becomes almost irreversible towards forward direction.
3. Only the esters which contains at least two α-hydrogen gives Claisen-ester condensation reaction.

$H_3C-C(=O)-OC_2H_5$ Ethyl acetate (First molecule) $\xrightleftharpoons[\text{Step-1}]{C_2H_5O^{\ominus}}$ [$^{\ominus}H_2C-C(=O)-OC_2H_5 \longleftrightarrow H_2C=C(O^{\ominus})-OC_2H_5$] Resonance stabilized carbanion

$H_3C-C(=O)-OC_2H_5$ Ethyl acetate (Second molecule)

Step-2

$H_3C-C(O^{\ominus})(OC_2H_5)-CH_2-C(=O)-OC_2H_5$

Step-3 $-C_2H_5O^{\ominus}$

$H_3C-C(=O)-CH_2-C(=O)-OC_2H_5$

Step-4 $C_2H_5O^{\ominus}$, $-C_2H_5OH$

$H_3C-C(=O)-CH=C(O^{\ominus})-OC_2H_5$

HCl

[$H_3C-C(=O)-CH^{\ominus}-C(=O)-OC_2H_5 \longleftrightarrow H_3C-C(O^{\ominus})=CH-C(=O)-OC_2H_5$] Resonance stabilized carbanion intermediate $\longleftrightarrow$ [$H_3C-C(=O)-CH_2-C(=O)-OC_2H_5$] Ethyl acetoacetate (EAA)

Q4. How EAA forms 4-methyl uracil.

Ans. Reaction with urea EAA forms 4-methyl uracil.

Q5. Give reaction of Diethyl malonate Hydrolysis and Decarboxylation.

Ans. Gives monocarboxylic acid.

Q6. Diethyl malonate form Salt why?

Ans. Because CH_2 group is acidic in nature.

UNIT 6: CARBOHYDRATES

6.1 Introduction
Objective
6.2 Classification of Carbohydrates
6.3 Classification of monosaccharides
6.3.1 Erythro and Threo Diastereomers
6.4 The cyclic structure of Glucose
6.5 Interconversion of straight-chain and ring forms of sugars
6.6 Configuration of monosaccharides
6.6.1 Conversion of glucose into mannose
6.6.2 Formation of glycosides
6.6.3 Ether formation (Williamson ether synthesis)
6.6.4 Ester formation
6.7 Determination of ring size of monosaccharides
6.7.1 Ribose
6.7.2 Deoxyribose
6.8 Disaccharides
6.8.1 Maltose
6.8.2 Lactose
6.8.3 Sucrose
6.9 Polysaccharides
6.9.1 Starch
6.9.1.1 Amylose
6.9.1.2 Amylopectin
6.9.2 Cellulose
6.10 Summary
6.11 Terminal Questions

6.1 Introduction

In this unit learner is able to classification and nomenclature of Monosaccharides, mechanism of osazone formation, interconversion of glucose and fructose, chain lengthening and chain shortening of aldoses as well as Configuration of monosaccharides, Erythro and threo diastereomers, Conversion of glucose into mannose. Formation of glycosides, ethers and esters. Determination of ring size of monosaccharides. Cyclic nature and structure of D(+)-glucose. Mechanism of mutarotation. Structures of ribose and deoxyribose and introduction to disaccarides with reference to maltose, sucrose and lactose and polysaccharides with reference to starch and cellulose without involving structure determination.

Objective

Learner is able to define

- Classification and nomenclature of Monosaccharides,
- Mechanism of osazone formation, interconversion of glucose and fructose.
- Chain lengthening and chain shortening of aldoses.
- Configuration of monosaccharides. Erythro and threo diastereomers,
- Conversion of glucose into mannose.
- Formation of glycosides, ethers and esters.
- Determination of ring size of monosaccharides.
- Cyclic structure of D(+)-glucose.
- Mechanism of mutarotation.
- Structures of ribose and deoxyribose.
- An introduction to disaccarides (maltose, sucrose and lactose)
- Polysaccharides (starch and cellulose) without involving structure determination.

Carbohydrate

Carbohydrate is a class of naturally occurring compounds which are mostly derived from them substances such as wood, starch, and linen however these substances are composed mainly molecules containing atoms of carbon (C), hydrogen (H), and oxygen (O) in fixed ratios. The general chemical formula of a carbohydrate is $C_x(H_2O)_y$, which denotes that some carbons (C) attached with some water molecules (H_2O) and hence the word carbohydrate means “hydrated carbon.” Simple carbohydrates are also

known as sugars or saccharides and the ending of the names of most sugars in –ose. Examples are: glucose, fructose, maltose, sucrose, arabinose, etc. carbohydrates are usually defined as optically active polyhydroxy aldehydes and ketones or in other words "compounds which on hydrolysis gives optically active polyhydroxy aldehydes and ketones."

Carbohydrates are products of photosynthesis, an endothermic reductive condensation of carbon dioxide requiring light energy and the pigment chlorophyll.

$$n\,CO_2 + n\,H_2O + \text{Energy} \xrightarrow[\text{Chlorophyll}]{\text{Sunlight}} C_nH_{2n}O_n + n\,O_2$$

If we do analysis at the molecular level, we found that most carbohydrates are polyhydroxy aldehydes, polyhydroxy ketones, or compounds that yield either of these after hydrolysis. Therefore, the chemistry of carbohydrates is essentially the chemistry of hydroxyl groups and carbonyl groups, and of the acetal bonds formed between these two functional groups.

Classification of Carbohydrates: Carbohydrates can be classified into three major groups which further divided into several groups. A schematic representation is given as follows:

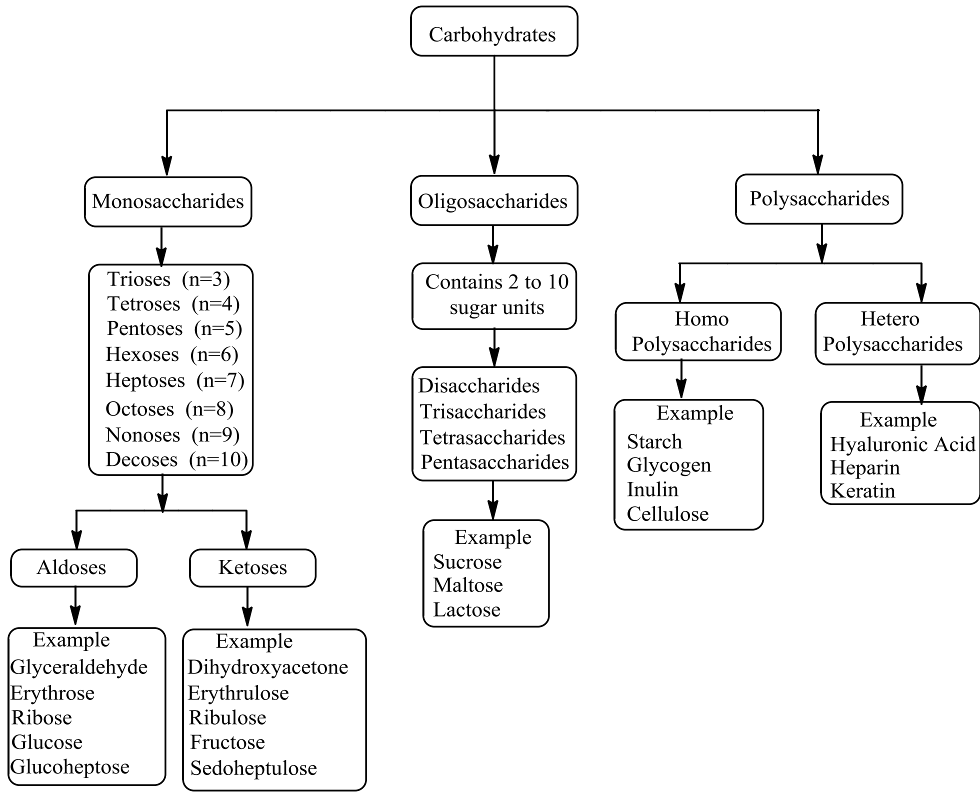


Figure 1: Classification of carbohydrates

1. **Monosaccharides:** The simplest carbohydrates that cannot be hydrolyzed into simple carbohydrates, are called monosaccharides.

2. **Oligosaccharides:** Carbohydrates that hydrolyze to yield 2 to 10 molecules of monosaccharides are called oligosaccharides. The most common oligosaccharides are disaccharides which include maltose, lactose, sucrose, etc.

3. **Polysaccharides:** Carbohydrates that yield a large number of monosaccharide molecules (more than 10 units), are known as polysaccharides. The common examples are starch, cellulose, glycogen, etc.

Classification of monosaccharides: Monosaccharides are basically classified on three criterions:

1. **Whether the monosaccharide contains a ketonic or an aldehydic group:** Monosaccharides are named usually having suffix –ose. Since monosaccharides are polyhydroxy aldehydes or ketones, they are divided into aldoses (Aldehyde + ose) and ketoses (ketone + ose).

2. **The number of carbon atoms in the carbon chain:** The number of carbon atoms in the monosaccharide molecule is indicated by a Greek prefix. The sub-classes of monosaccharides are listed below:

No. of carbon atoms in monosaccharide	**Empirical formula**	**Sub-class**	
		Aldose	**Ketose**
3	$C_3H_6O_3$	Aldotriose: glyceraldehyde	--
4	$C_4H_8O_2$	Aldotetrose: erythrose and threose	Ketotetrose: erythrulose
5	$C_5H_{10}O_5$	Aldopentose: arabinose, xylose, ribose, etc.	Ketopentose: xylulose, ribulose, etc.

6	$C_6H_{12}O_6$	Aldohexose: glucose, mannose, galactose, etc.	Ketohexose: fructose, sorbose, etc.

3. **The stereochemical configuration of the chiral carbon which have maximum distance from the carbonyl group (D and L configuration of monosaccharides):** The simplest member of monosaccharide family is glyceraldehyde which contains a stereocenter and hence exists in two enantiomeric forms.

CHO
|
H——C——OH
|
CH_2OH

R-(+)glyceraldehyde

CHO
|
OH——C——H
|
CH_2OH

S-(+)glyceraldehyde

When absolute configuration of organic compounds was not known, another system of stereochemical designations was introduced. According to this system (+)- glyceraldehyde is designated as D-(+) glyceraldehyde and (-)- glyceraldehyde is designated as L-(-) glyceraldehydes. These two compounds serve as configurational standards for all monosaccharides.

CHO
|
CHOH
|
CHOH
|
CHOH
|
H——C——OH
|
CH_2OH

D-aldohexose

CHO
|
CHOH
|
CHOH
|
CHOH
|
H——C——OH
|
CH_2OH

L-ketohexose

A monosaccharide whose chiral carbon atom is farthest from the carbonyl group has the same configuration as D-(+)-glyceraldehyde is designated as D-sugar. In the similar way a monosaccharide whose chiral carbon atom is farthest from the carbonyl group has the same configuration as L-(-)-glyceraldehyde is designated as L-sugar.

Erythro and Threo Diastereomers:

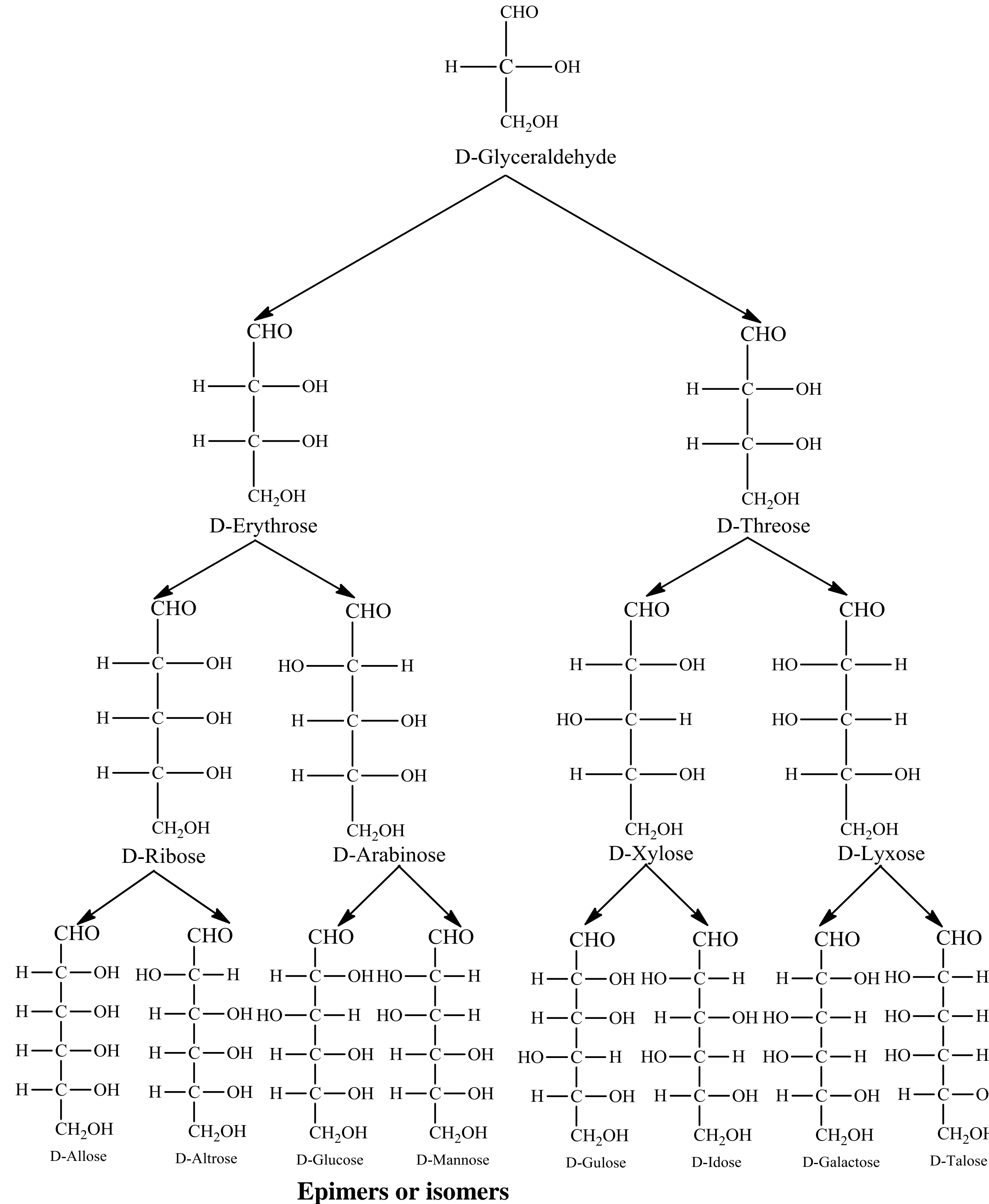
Erythro and *Threo* system of nomenclature is used for aldotetroses only. Aldotetrose have two chiral centres and hence shows four stereoisomers out of which two are D-sugars and remaining two stereoisomers are L-sugars. When Fisher projections are drawn for stereoisomers with two adjacent chiral centers, the pair of enantiomers with similar groups on the same side of the carbon chain is called the *erythro* enantiomers. The pair of enantiomers with similar groups on opposite sides is called the *threo* enantiomers. *Erythrose* and *threose* are diastereoisomers.

CHO	CHO	CHO	CHO
H—C—OH	HO—C—H	HO—C—H	H—C—OH
H—C—OH	HO—C—H	H—C—OH	HO—C—H
CH_2OH	CH_2OH	CH_2OH	CH_2OH
D-Erythrose	*L-Erythrose*	*D-Threose*	*L-Threose*

Epimers: The two L isomers; L-erythrose and L-threose is the enantiomer of D-threose. L-threose is a diastereomer of both D and L-erythrose, and L-erythrose is a diastereomer of both D and L-threose. Diastereomers that differ from each other in configuration at only one chiral carbon are called epimers; erythrose and D-threose are epimers. Glucose and galactose are also examples of epimers.

Aldopentoses have three chiral carbons and there are 2^3 or 8 possible stereoisomers; four D- form and four L- form, similarly aldohexoses will have 16 stereoisomers. Some of the possible stereoisomers are much more common in nature than others. For example, D sugars predominates in nature rather than

L sugars. Most sugars present in nature, especially in foods, contain either five or six carbon atoms. Here D-glucose (an aldohexose) is taken as an example because it is most abundant monosaccharide in nature.



Epimers or isomers

Structure of monosaccharides: It has been observed that aldehydes and ketones react with alcohols to form hemiacetals and these cyclic hemiacetals

form very readily when hydroxyl and carbonyl groups are part of the same molecule and their interaction can form a five- or six membered ring. For example, 4-hydroxypentanal forms a five-membered cyclic hemiacetal.

Rearranging -OH and COOH to come close with each other

New chiral centre

A cyclic hemiacetal

It is evident from the above structural rearrangement that 4-hydroxypentanal contains one chiral center whereas second chiral center is generated at carbon 1 as a result of hemiacetal formation. Monosaccharides have hydroxyl and carbonyl groups in the same molecule as a result, they too exist almost exclusively as five- and six-membered cyclic hemiacetals.

The cyclic structure of Glucose: Before formulation of hemiacetals, it was proposed that saccharides are straight-chain open compounds but after the synthesis of glucose it was observed that glucose having open chain structure proposed by E. Fischer dose not account for all the reactions. Thus:

1. Despite having the aldehyde group, glucose does not give Schiff's test and it does not form the hydrogen sulphite addition product with $NaHSO_3$.

$$R-CHO + NaHSO_3 \longrightarrow R-CH(OH)(SO_3Na)$$

2. The pentaacetate of glucose does not react with hydroxylamine indicating the absence of free —CHO group.

3. Glucose forms crystalline products when refluxed with methanolic hydrogen chloride: methyl α and methyl β-D-glucosides, two stereoisomers with α isomer $(\alpha)_D$ + 158°, m.p. 166° and β-isomer $(\alpha)_D$- 34°, m.p. 108°.

These glucosides have no reducing properties. In the D-series the sugar isomer with the most positive rotation is the α-isomer, the isomer with the lower rotation being called β-isomer.

In case of normal aldehyde, hemiacetal is formed whereas acetal with methanolic hydrogen chloride.

$$R-CHO \xrightarrow{CH_3OH} R-CH(OCH_3)(OH) \xrightarrow{CH_3OH} R-CH(OCH_3)_2$$

Whereas only one (OCH_3) group is introduced into glucose to form the glucoside (hemi-acetal).

4. Corresponding with the glucosides two α- and β-modifications of D-glucose itself were isolated: α-$(\alpha)_D$ + 112°, m.p. 146° and β-$(\alpha)_D$ + 18.7°, m.p. 156°. The two forms are interconvertible in solution when either form is dissolved in water the rotation gradually changes (mutarotation) until an equilibrium value (+52.7°) is reached. Mutarotation occurs throught the aldehyde form as intermediates this form being present normally to the extent of less than 1%. The low concentration of the aldehyde form dose not favours reaction with Schiff's reagent.

Alpha-D-Glucopyranose ⇌ D-Glucose ⇌ Beta-D-Glucopyranose

The α- and β-forms which differ in configuration at C1 only are known as anomers (anomer is actually an epimer that differs in configuration at the acetal/hemiacetal carbon). The hydroxyl group on C1 is cis to the hydroxyl on C2 in the α-form and trans in the β-form was deduced by Boeseken who found that the α form of D-glucose increases conductivity of boric acid solution considerably and is greater for the cis- than the trans- arrangement of hydroxyl group.

The ring structure of glucose (1,5) is a six-membered ring and is described as a pyranose ring by similarity to pyran while the five-membered ring is described as a furanose ring.

Furan Pyran

The glycosides of the pyranose sugars are described as pyranosides and those of the furanose sugars as furanosides.

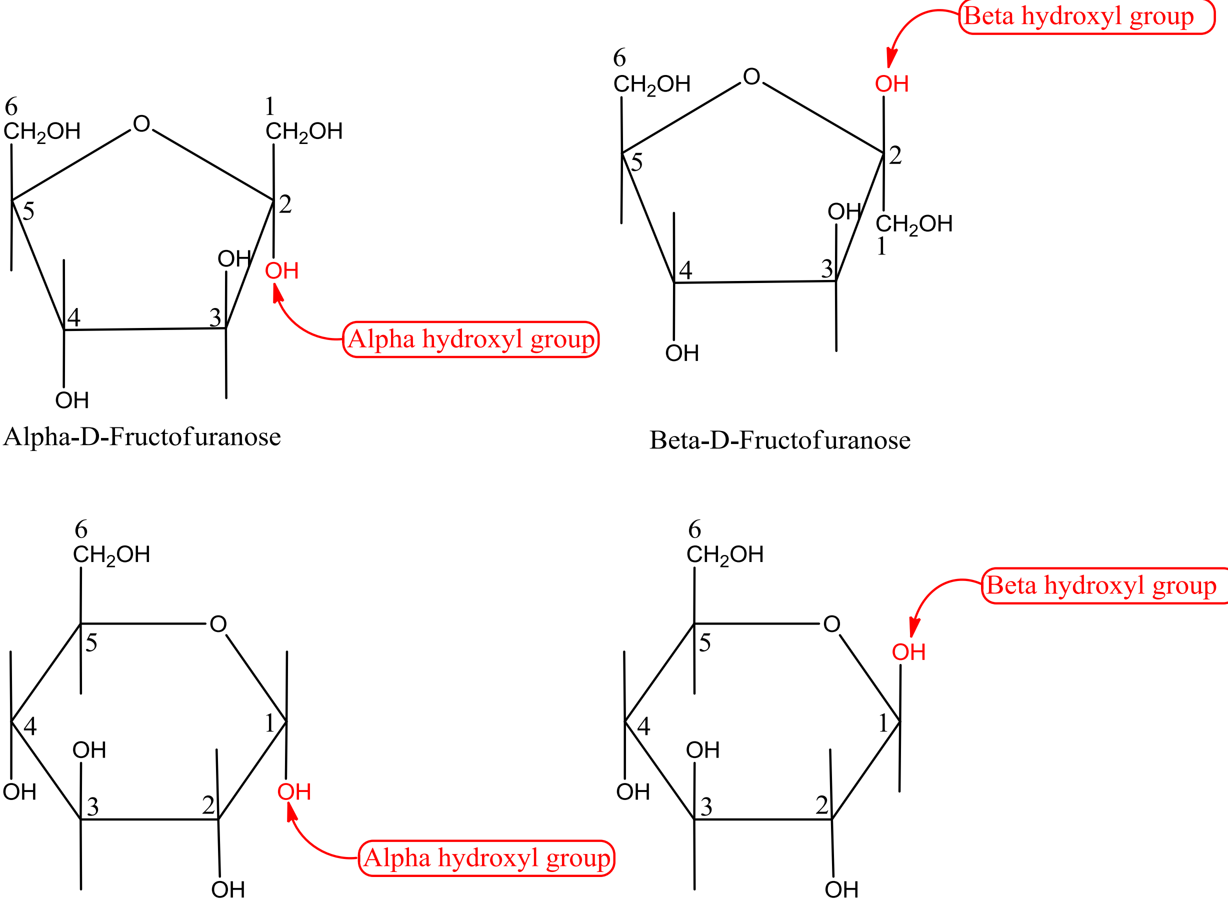
A monosaccharide existing as a five-membered ring is a furanose; one existing as a six-membered ring is a pyranose. A pyranose is most commonly drawn as either a Haworth projection or a chair conformation.

D-Glucose

Beta-D-Glucopyranose
(Beta-D-Glucose)

Howorth projections: The stereochemistry of the cyclic forms of sugars is often represented by their Haworth projections (A standardized way of

depicting the positions of hydroxyl groups in space. In viewing Haworth projections, envision the plane of the ring as tilted perpendicular to the plane of the paper. The attached groups are above and below the plane of the ring. Here we have taken D-glucopyranose and D-fructofuranose as an example for understanding Haworth projections. Two forms of each of these sugars are possible; these are designated as alpha and beta. Two cyclic forms are possible because, in going from a straight chain to a ring, a new asymmetric carbon is introduced at carbon 1 (the hemiacetal carbon) of aldoses and carbon 2 (the hemiketal carbon) of ketoses. In the Haworth projections of D-glucopyranose, for example, the hydroxyl group at carbon 1, which is formed from the aldehyde functional group of the straight-chain sugar, may end up below (alpha or Greek a) or above (beta or Greek F) the plane of the pyranose ring.



Anomers are sugars that differ in stereochemistry only at the hemiacetal or hemiketal carbon. The alpha and beta anomers of the cyclic forms of sugars

have different melting points and different abilities to rotate plane polarized light. Alpha-D-glucose melts at 146°C whereas beta-D-glucose melts at 150°C. These differences help to demonstrate again how small changes in molecular shape or structure dramatically affect the physical properties of molecules.

Interconversion of straight-chain and ring forms of sugars: The straight-chain sugar forms are in equilibrium with the ring forms. The ring forms are usually quite predominant. For example, if stereochemically pure α-D-glucopyranose is dissolved in an acidic solution, the ring will open and close repeatedly. In reclosing, some β-D-glucopyranose is formed. The final equilibrium mixture consists of about 63% β-D-glucopyranose, about 37% α-D-grucopyranose, and only a tiny amount of the straight-chain aldehyde. From the percentages of products formed, it can be concluded that β-D-glucopyranose is only slightly more preferred than α-D-grucopyranose and that both D-glucopyranose anomers are much more preferred than the straight-chain aldehyde form of D-glucose.

Alpha-D-Glucopyranose

(About 37%)

D-Glucose

(Less than 1%)

Beta-D-Glucopyranose

(About 63%)

Mutarotation: It is possible to obtain a sample of crystalline glucose in which all the molecules have the α structure or all have the β structure. The α form melts at 146°C and has a specific rotation of +112°, while the β form melts at 150°C and has a specific rotation of +18.7°. When the sample is dissolved in water, an interesting change in the specific rotation is observed. When the α-anomer dissolves, its specific rotation gradually decreases from an initial value of +112° to +52.6°. When the pure β-anomer dissolves, its specific rotation gradually increases from +18.7° to the same value of +52.6°. This change in the value of specific rotation to a mutual value is called mutarotation. Mutarotation occurs because in solution the two anomers can interconvert. When either of the

pure anomers dissolves in water, the rotation of the pure anomer gradually changes to an intermediate rotation that results from the equilibrium concentrations of the anomers. Thus mutarotation is due to the conversion of either anomer to an equilibrium mixture of both. At equilibrium, the mixture consists of about 36% α-D-glucose, 64% β-D-glucose, and less than 0.02% of the open-chain aldehyde form. The observed rotation of this solution is +52.7°.

Alpha anomer | **Equilibrium mixture of alpha & beta anomer** | **Beta anomer**

Crystalize below 98°C | Crystalize above 98°C

CHO
H—C—OH
HO—C—H
H—C—OH
H—C—OH
CH_2OH

Alpha-D-(+) Glucopyranose | **D-(+) Glucose (Open chain form)** | **Beta-D-(+) Glucopyranose**

Mechanism of mutarotation: Mutarotation occurs by opening of the ring to the free carbonyl form. The mechanism starts as the reverse of hemiacetal (or hemiketal) formation. A 180° rotation about the bond to the carbonyl group permits attack of the hydroxyl group at C-5 on the opposite face of the carbonyl carbon. Hemiacetal formation then gives the other anomer. Mutarotation is catalysed by both acid and base.

Alpha-D-Glucopyranose

Open chain form

Rotation, 180°

Beta-D-Glucopyranose

Open chain form

Mechanism of Mutarotation

Chain lengthening and chain shortening of aldoses

Chain lengthening and chain shortening process involves two independent and different procedures which are explained here under two sub-headings.

1. Chain Elongation: The Kiliani–Fischer Synthesis

The carbon chain of an aldose can be increased by one carbon in a Kiliani–Fischer synthesis. For example, tetroses can be converted into pentoses, and pentoses can be converted into hexoses and so on. In the first step of the synthesis (the Kiliani portion), the aldose is treated with sodium cyanide and HCl. Addition of cyanide ion to the carbonyl group creates a new asymmetric carbon. Consequently, two cyanohydrins that differ only in configuration at C-2 are formed. The configurations of the other asymmetric carbons do not change, because no bond to any of the asymmetric carbons is broken during the course of the reaction. Kiliani went on to hydrolyze the cyanohydrins to aldonic acids, and Fischer had previously developed a method to convert aldonic acids to aldoses. This reaction sequence was used for many years, but the method currently employed to convert the cyanohydrins to aldoses was developed by Serianni and Barker in 1979; it is referred to as the modified Kiliani–Fischer

synthesis. Serianni and Barker reduced the cyanohydrins to imines, using a partially deactivated palladium (on barium sulfate) catalyst so that the imines would not be further reduced to amines. The imines could then be hydrolyzed to aldoses.

The modified Kiliani–Fischer synthesis:

D-erythrose + $C{\equiv}N^{-}$ $\xrightarrow{HCl}$ cyanohydrins $\xrightarrow[Pd/BaSO_4]{H_2}$ imines $\xrightarrow[H_2O]{HCl}$ D-ribose + NH_4^{+} and D-arabinose + NH_4^{+}

It may be noticed that the synthesis leads to a pair of C-2 epimers because the first step of the reaction converts the carbonyl carbon in the starting material to an asymmetric carbon. Therefore, the OH bonded to C-2 in the product can be on the right or on the left in the Fischer projection. The two epimers are not obtained in equal amounts, however, because the first step of the reaction produces a pair of diastereomers and diastereomers are generally formed in unequal amounts (Section 5.19).

2. Chain Shortening: The Ruff Degradation

The Ruff degradation is the opposite of the Kiliani–Fischer synthesis. Thus, the Ruff degradation shortens an aldose chain by one carbon: Hexoses are converted into pentoses, and pentoses are converted into tetroses. In the Ruff degradation, the calcium salt of an aldonic acid is oxidized with hydrogen peroxide. Ferric ion catalyzes the oxidation reaction, which cleaves the bond

between C-1 and C-2, forming and an aldehyde. It is known that the reaction involves the formation of radicals, but the precise mechanism is not well understood.

The Ruff degradation:

$$COO^{\ominus}(Ca^{2+})_{1/2}\text{–}CH(OH)\text{–}CH(OH)\text{–}CH(OH)\text{–}CH(OH)\text{–}CH_2OH + H_2O_2 \xrightarrow{Fe^{3+}} HC{=}O\text{–}CH(OH)\text{–}CH(OH)\text{–}CH(OH)\text{–}CH_2OH + CO_2$$

Calcium D-gluconate → D-arabinose

The calcium salt of the aldonic acid necessary for the Ruff degradation is easily obtained by oxidizing an aldose with an aqueous solution of bromine and then adding calcium hydroxide to the reaction mixture.

$$HC{=}O\text{–}CH(OH)\text{–}CH(OH)\text{–}CH(OH)\text{–}CH(OH)\text{–}CH_2OH \xrightarrow[2.\ Ca(OH)_2]{1.\ Br_2,\ H_2O} HC{=}O\text{–}CH(OH)\text{–}CH(OH)\text{–}CH(OH)\text{–}CH(OH)\text{–}CH_2OH$$

D-glucose → Calcium D-gluconate

Configuration of monosaccharides:

1. **Configuration of trioses and tetroses**

Glyceraldehyde has one chiral centre and hence exists in two enantiomeric forms which is represented by Fischer projection formula.

$$\begin{array}{ccc} & CHO & \\ & | & \\ H— & C & —OH \\ & | & \\ & CH_2OH & \end{array} \qquad \begin{array}{ccc} & CHO & \\ & | & \\ HO— & C & —H \\ & | & \\ & CH_2OH & \end{array}$$

D(+) glyceraldehyde L(-) glyceraldehyde

The D and L indicates the configurational relationship whereas (+) and (-) signs shows whether the substance is dextro- or levorotatory. Any compound which could be either converted or obtained from D-(+)-glyceraldehyde belong to D-series and in similar way L series compounds could be obtained from or converted to L-(-)-glyceraldehyde. In deciding the configurational relationship of the molecule containing more than one asymmetric carbon atoms, the asymmetric carbon atom of the glyceraldehyde is drawn at the bottom and the rest of the molecule built at upper part.

Kilianisynthesis

D(+) glyceraldehyde → D(-) erythrose + D(-) threose

(D-Series)

Kilianisynthesis

L(-) glyceraldehyde → L(-) threose + L(-) erythrose

(L-Series)

Thus D (+) glyceraldehyde can be converted into two sugars; D-erythrose and D-threose while L (-) glyceraldehyde converts into L-threose and L-erythrose. From above structures it is evident that D and L-erythrose and D and L-threose are enantiomers. The enantiomeric D- and L-threose give D- and L-tartaric acid however, both D- and L-erythrose give *meso*-tartaric acid on oxidation with HNO_3.

After kiliani synthesis from glyceraldehyde, two kinds of erythroses are obtained; one has (R,R) configuration and another one has (S,S) configuration. But there was no any accurate method to find out which structure has what kind of configuration. In 1954 a chemist Bijvoet determined the absolute configuration of D-tartaric acid by crystallographic technique for the first time. The configuration of D-tartaric acid was found as follows:

COOH
HO—C—H
H—C—OH
COOH

D-tartaric acid

The configurations of D-threose and L-threose are as follows:

CHO
HO—C—H
H—C—OH
CH_2OH
D-threose

$\xrightarrow{HNO_3}$

COOH
HO—C—H
H—C—OH
COOH
D-tartaric acid

CHO
H—C—OH
HO—C—H
CH_2OH
L- threose

$\xrightarrow{HNO_3}$

COOH
HO—C—H
H—C—OH
COOH
D-tartaric acid

D-threose is obtained from D-glyceraldehyde by Kiliani synthesis. In this conversion there is no change in the configuration of C-2 carbon, therefore the absolute configuration of D-glyceraldehyde is as follows:

CHO
H—C—OH
CH_2OH

D-glyceraldehyde

D-erythrose is also obtained from D-glyceraldehyde by Kiliani synthesis, thus its configuration should be as follows:

CHO

H—C—OH

H—C—OH

CH_2OH

D-erythrose

D-erythrose and L-erythrose are enantiomers therefore the configuration of L-erythrose is as follows:

CHO

HO—C—H

HO—C—H

CH_2OH

L-erythrose

2. **Configuration of aldopentoses:** Aldopentoses have three dissimilar asymmetric carbon atoms and hence 2^3 =8 stereoisomers or four pairs of enntiomers. The structure of D-erythrose has already been established and when it is stepped up; it gives D-(-)-ribose and D-(-)-arabinose. Similarily D-threose gives D-(+)-xylose and D-(-)-lyxose.

D-erythrose

Kiliani syntheis

O=C—H

H—C—OH

H—C—OH

H—C—OH

CH_2OH

(A)

O=C—H

HO—C—H

H—C—OH

H—C—OH

CH_2OH

(B)

A and B must be ribose and arabinose but which is which? If we consider ribose, it gives optically active trihydroglutaric acid on oxidation and it is possible only when arabinose has the structure B because only compound B leads to the formation of an optically active dibasic acid whereas compound A leads to the formation of optically inactive dibasic acid. Hence, compound A is ribose and B is arabinose.

D-erythrose

Kiliani syntheis

(C)	(D)
CHO	CHO
H—C—OH	HO—C—H
HO—C—H	HO—C—H
H—C—OH	H—C—OH
CH_2OH	CH_2OH

(C) **(D)**

Similarly out of C and D the one which gives an optically active dicarboxylic acid is known to be D-xylose and the other which gives an optically inactive dibasic acid is D-lyxose. Accordingly C is D-xylose and D is D-lyxose. Other four sugars being enantiomers, their configuration can be established in the same way.

3. **Configuration of aldohexoses:** Structure of aldohexose contains four asymmetric chiral carbons and so there are 2^4= 16 isomers or 8 pairs of enantiomers.

By stepping D-ribose gives D-allose and D-altrose, D-arabinose gives D-glucose and D-mannose, D-xylose gives D-gulose and D-idose whereas D-lyxose gives D-galactose and D-talose. It could be represented in following chain like structure:

CHO
H—C—OH
H—C—OH
H—C—OH
CH_2OH
D-Ribose

Kiliani Synthesis

CHO
H—C—OH
H—C—OH
H—C—OH
H—C—OH
CH_2OH
D-Allose (1)

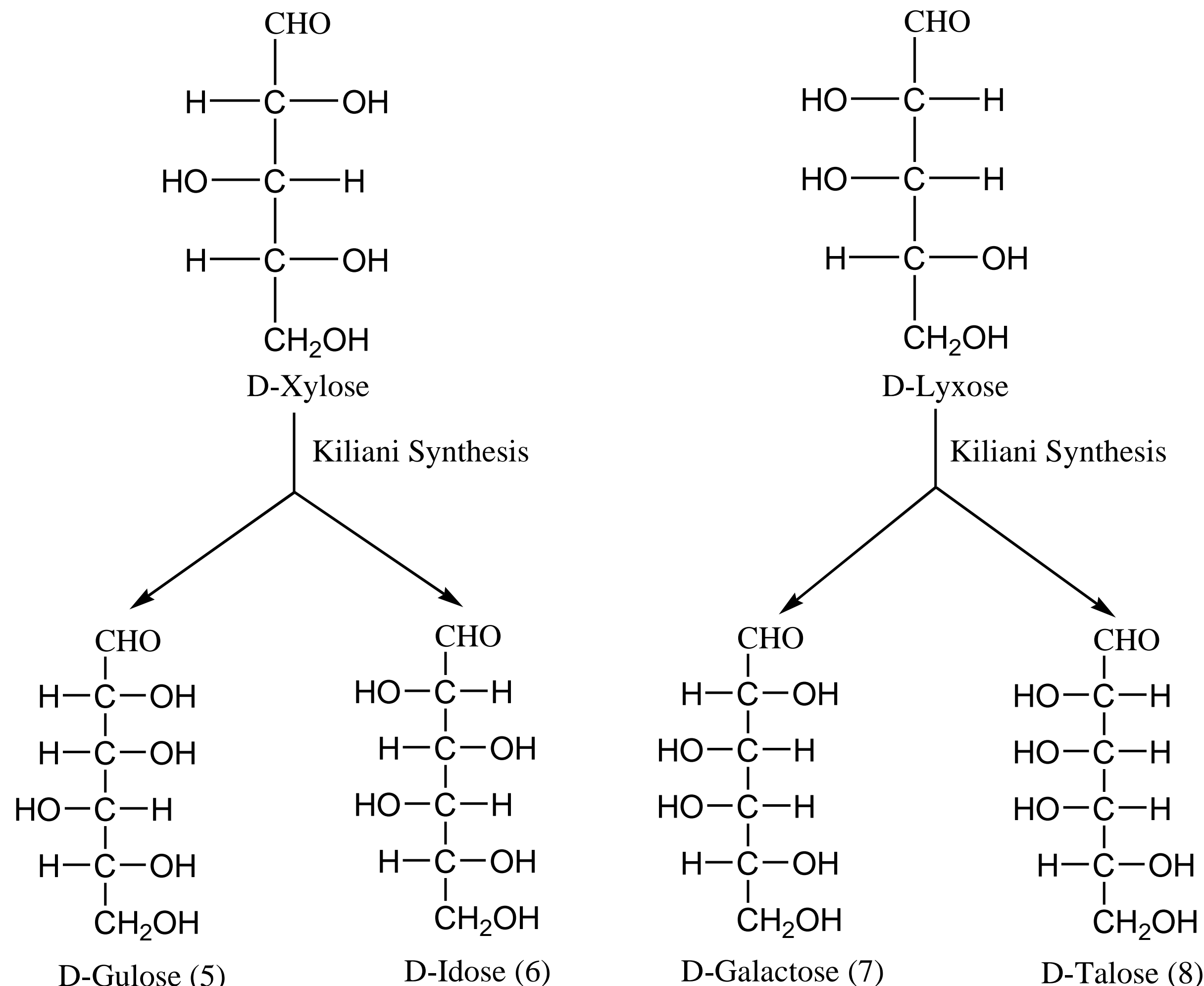
CHO
HO—C—H
H—C—OH
H—C—OH
H—C—OH
CH_2OH
D-Altrose (2)

CHO
HO—C—H
H—C—OH
H—C—OH
CH_2OH
D-Arabinose

Kiliani Synthesis

CHO
H—C—OH
HO—C—H
H—C—OH
H—C—OH
CH_2OH
D-Glucose (3)

CHO
HO—C—H
HO—C—H
H—C—OH
H—C—OH
CH_2OH
D-Mannose (4)



(1) and (2) must be allose and altrose, but which one is which? On oxidation allose is known to give an optically inactive allomucic acid whereas altrose on similar treatment gives an optically active talomucic acid. Hence, (1) must be D (+) allose and (2) must be D (+) altrose.

D-arabinose gives D (+) glucose and D (+) mannose. Thus (3) and (4) must be glucose and mannose, but which one is which? On oxidation both give optically active acids. Thus oxidation method cannot be applied for the determination of the configuration of glucose and mannose.

D-Glucose (3)

D-Mannose (4)

HNO_3

HNO_3

Optically active

Optically active

D-Glucose (3)

HNO_3

△

(A)

(B)

Mixture of lactones

By stepping up D-lyxose gives D-(+)-galactose and D-(-)-talose. Thus (7) and (8) must be D-(+)-galactose and D-(-)-talose, but which one is which? On oxidation D-galactose is known to give an optically inactive mucic acid and D-talose an optically active talomucic acid. Hence (7) must be galactose and (8) must be talose. Now to determine which one is D-glucose and which one is D-mannose of (3) and (4), Fischer tried to explain by following mechanism. Fischer interchanged the –CHO and –CH_2OH groups of each (3) and (4) and found that (3) on such an interchange as stated gave a new aldohexose of L-series while (4) on similar interchange gave the same aldohexose. He then

undertook an experiment to interchange –CHO and CH_2OH and found that D-glucose gave a new L-aldohexose [L-(+) gulucose] which is enantiomer of D-glucose while D-mannose gave the same aldohexose , D-mannose. Therefore, glucose is (3) and mannose is (4)

(A) $\xrightarrow[\text{HOH}]{\text{Na-Hg}}$ $\xrightarrow[\text{pH 3-5}]{\text{Na-Hg }(CO_2)}$ $\xrightarrow[\text{the plane of paper}]{180^0\text{ rotation on}}$ L-gluase

Interchange of groups by chemical method will be obtained by lactone which is formed by the carboxylic group produced by the oxidation of $–CH_2OH$.

(4) $\xrightarrow[\text{[O]}]{HNO_3}$ $\xrightarrow{\Delta}$ + (A)

(A) $\xrightarrow{\text{Na-Hg-HOH}}$ $\xrightarrow{\text{Na-Hg}/H_2O}$ $\xrightarrow{180^O\text{rotation}}$ (4)

Conversion of glucose into mannose: When glucose is treated with very dilute alkalies or organic bases such as pyridine or quinoline, epimers of glucose are obtained and thus glucose yields a mixture of mannose, fructose and unreacted

glucose. The mechanism of this reaction was suggested by Lobrg de Druyn and Van Ekenstein involving the formation of an intermediate enediol.

The hydrogen atom attached to the carbon to the carbonyl (C2 in glucose) enolysis to form an enediol, thus destroying the asymmetry of C2. On ketonisation the 2 epimeric aldoses are formed. If the second hydrogen on C2 migrates to C1 a ketose is formed (fructose). Thus we can convert glucose into mannose and fructose.

The reaction is best carried out by the epimerization of aldonic acids which are more stable towards alkaline medium. Thus the aldose is first oxidized to the aldonic acid/ gluconic acid, which is then heated with an organic base like pyridine or quinoline and thus it converted to mannonic acid which is then lactonised and reduced to give mannose.

Glucose $\xrightarrow[Br_2/H_2O]{[O]}$ Gluconic acid $\xrightarrow[\text{Heat}]{\text{Organic base}}$ Mannonic acid $\xrightarrow[-H_2O]{\text{Heat}}$ lactone $\xrightarrow{[H]}$ Mannose

Formation of glycosides: It is evident that when an aldehyde reacts with an equivalent of an alcohol to form a hemiacetal, the hemiacetal reacts with a second equivalent of alcohol to form
an acetal.

$$\text{Aldehyde or Ketone} + 2\,ROH \underset{}{\overset{H_3O^+}{\rightleftharpoons}} \text{Acetal}$$

The mechanism for understanding formation of hemiacetal or acetal is given in following steps.

i. Protonation of the carbonyl group

ii. **Nucleophilic attack by the alcohol**

iii. **Deprotonation to form a hemiacetal**

Hemiacetal

iv. **Protonation of the alcohol**

v. **Removal of water**

+ H_2O

vi. Nucleophilic attack by the alcohol

vii. Deprotonation by water

+ H_3O^+

Acetal

Similarly, the cyclic hemiacetal (or hemiketal) formed by a monosaccharide can react with an alcohol to form an acetal (or ketal). The acetal (or ketal) of a sugar is called a glycoside, and the bond between the anomeric carbon and the alkoxy oxygen is called a glycosidic bond. Glycosides are named by replacing the "e" ending of the sugar's name with "ide." Thus, a glycoside of glucose is a glucoside, a glycoside of galactose is a galactoside, etc. If the pyranose or furanose name is used, the acetal is called a pyranoside or a furanoside.

CH_3CH_2OH, HCl

A glycosidic bond

Beta-D-glucose
Beta-D-glucopyranose

Ethylbeta-D-glucoside
Ethylbeta-D-glucopyranoside

Ethylalpha-D-glucoside
Ethylalpha-D-glucopyranoside

It was noticed that the reaction of a single anomer with an alcohol leads to the formation of both the α- and β- glycosides. The mechanism of the reaction shows why both glycosides are formed. The OH group bonded to the anomeric carbon becomes protonated in the acidic solution, and a lone pair on the ring oxygen helps expel a molecule of water. The anomeric carbon in the resulting oxocarbenium ion is sp^2 hybridized, so that part of the molecule is planar. (An oxocarbenium ion has a positive charge that is shared by a carbon and an oxygen.) When the alcohol comes in from the top of the plane, the β- glycoside is formed; when the alcohol comes in from the bottom of the plane, the α-glycoside is formed.

Mechanism of glycoside formation:

CH3CH2OH comes in from the top

CH3CH2OH comes in from the bottom

A beta glycoside

An alpha glycoside
(Major Product)

Surprisingly, D-glucose forms more of the α- glycoside than the β-glycoside. Similar to the reaction of a monosaccharide with an alcohol is the reaction of a monosaccharide with an amine in the presence of a trace amount of acid. The product of the reaction is an *N*-glycoside. An *N*-glycoside has a

nitrogen in place of the oxygen at the glycosidic linkage. The subunits of DNA and RNA are β-*N*- glycosides.

CH$_2$OH O OH OH OH

Analine, Catalytic H$^+$

CH$_2$OH O NH OH OH

\+

CH$_2$OH NH O OH OH

N-phenyl-beta-D-ribosylamine
An beta-N-glycoside

N-phenyl-alpha-D-ribosylamine
An alpha-N-glycoside

Ether formation (Williamson ether synthesis): The anomeric hydroxyl group can be replaced by an alkoxy group under mild acidic conditions. This reaction produces ether and is known as glycosidation. The remaining hydroxyl groups of the sugar are unaffected by this process because such a process would involve a primary or secondary carbocation, rather than the far more stable oxonium ion that is involved in glycoside formation. The other hydroxyl groups can be converted into ethers by an application of Williamson ether synthesis. The most common ethers are the methyl ethers which are prepared by treating the sugar with 30% aqueous sodium hydroxide and $(CH_3)_2SO_4$ or with AgI and CH_3I. Since the free aldehyde form of sugar is not stable to strongly basic conditions, it is customary to protect the anomeric hydroxy group by converting the sugar into the methyl glycoside. The glycosidic linkage then be cleaved by mild acid hydrolysis because the normal ether linkages are stable under these conditions.

$(CH_3)_2SO_4/NaOH$ or CH_3I/AgI →

Methyl 2, 3, 4, 6-tetra-O-methyl beta-glucopyranoside

↓ Dil. Hcl

2, 3, 4, 6-tetra-O-methyl alpha-D-glucose + 2, 3, 4, 6-tetra-O-methyl beta-D-glucose

Methylation is very useful method for determining the size of the sugar ring. It can be summarized under following steps.

(a) Acid-catalysed methylation

(b) Williamson etherification or permethylation

(c) Hydrolysis of permethylated derivative

Applications of step (a), (b) and (c) to glucose produces 2, 3, 4, 6-tetra-o-methyl glucose. This compound contains free hydroxyl group at C-5. Thus the ring must be pyranose rather than furanose, i.e. C-5 hydroxyl group is involved in hemiacetal formation.

$$\text{D-glucose} \xrightarrow[\text{HCl}]{CH_3OH} \text{Methyl glucoside} \xrightarrow{(CH_3)_2SO_4/NaOH} \text{Methyltetra-O-methylglucoside}$$

$$\downarrow \; HOH / H^{\oplus}$$

CHO
H — C — OCH_3
H_3CO — C — H
H — C — OCH_3
H — C — OH
CH_2OCH_3

2, 3, 4, 6-tetra-O-methyl-D-glucose

Ester formation: The hydroxyl groups of sugars are easily esterified. Acetylation is the most common esterification reaction. Acetylation is achieved by using acetic anhydride with acidic (e.g., sulphuric acid or $ZnCl_2$) or basic (e.g., sodium acetate or pyridine) catalysts. This reaction acetylates all the hydroxyl groups, including that of the hemiacetal (or hemiketal) on the anomeric carbon. The anomeric C-O bond is not broken in the acetylation, and the stereochemistry of the anomeric carbon atom is usually preserved.

The proportion of α- and β- glucose pentaacetates produced on acetylation can be controlled by varying experimental conditions. Above room temperature α- and β- acetates are interconverted by acids to produce a mixture of 90% α and 10% β.

Ester formation of alpha-anomer

At temperature below 0°C and in the presence of a basic catalyst, the rate of acetylation is much faster than interconversion of the anomers. Thus at low temperature the reaction occurs stereospecifically: the α- anomer gives the α- acetate and β- anomer gives β- acetate. Since the equatorial hydroxy group is more reactive than the axial hydroxy group, base catalyzed acetylation produces mainly the β- monomer.

$Ac_2O + AcO^{\ominus}$

Fast

Slow at 0°C

Ester formation of beta-anomer

The more stable pentaacetate is actually the α- form but equilibrium is established only under still more drastic condition, i.e., at high temperature and in acidic medium.

$$\text{Glucose pentaacetate (Beta-anomer)} \underset{\substack{ZnCl_2 \\ K = 6.7}}{\overset{K}{\rightleftharpoons}} \text{Glucose pentaacetate Alpha-anomer)}$$

Generally, β- anomer is more stable than the α- anomer. In the case of acetylation α- anomer is more stable than the β- anomer. The stability of the α-anomer is an example of anomeric effect. The preference for the axial geometry at C-1 is brought about by destabilizing dipolar repulsion between heterocyclic oxygen and the substituent at C-1 in the case of certain substituents in equatorial geometry (e.g., RO, AcO, halogen, RCOO, RS).

No dipolar repulsion (More stable)

Dipolar repulsion (Less stable)

The anomeric effect depends very much on the nature of the electronegative (i.e.,-I group) substituents. The decreasing order of anomeric effect among the following groups is:

$$X > ArCOO > AcO > RO > RS$$

The anomeric effect has been found to be less in the solvents of high polarity. The anomeric effect is more important in nonpolar solvents.

Esters are used as protecting groups in organic synthesis with sugars. Esters can be deprotected because they are easily hydrolysed by base. Ester groups are also easily removed by transesterification as in the equation below:

$$\text{(penta-OAc pyranose)} + C_2H_5OH\ (\text{Excess}) \xrightarrow{C_2H_5O^{\ominus}} \text{(pyranose, all OH)} + 5CH_3-\overset{\overset{\displaystyle O}{\|}}{C}-OC_2H_5$$

Determination of ring size of monosaccharides: Two different procedures can be used to determine what size ring a monosaccharide forms. In the first procedure, treatment of the monosaccharide with excess methyl iodide and silver oxide converts all the OH groups to OCH_3 groups. Acid-catalyzed hydrolysis of the acetal then forms a hemiacetal, which is in equilibrium with its open chain form. The size of the ring can be determined from the structure of the open-chain form because the sole OH group is the one that had formed the cyclic hemiacetal.

In the second procedure used to determine a monosaccharide's ring size, an acetal of the monosaccharide is oxidized with excess periodic acid.

The α-hydroxyaldehyde formed when periodic acid cleaves a 1,2,3-diol is further oxidized to formic acid and another aldehyde.

The products obtained from periodate cleavage of a six-membered-ring acetal are different from those obtained from cleavage of a five-membered-ring acetal.

Five-membered ring

Ribose: Ribose is a simple sugar and carbohydrate with molecular formula $C_5H_{10}O_5$ and the linear-form composition H−(C=O)−(CHOH)$_4$−H. Ribose naturally occurs in the form of D-ribose which is an essential component of the ribonucleotides from which RNA is built. It has a structural analog, deoxyribose, which is a similarly essential component of DNA. Emil Fischer and Oscar Piloty in year 1891, first time synthesized an unnatural sugar L-Ribose. Like other sugars, ribose also exists as mixture of cyclic and linear forms in equilibrium and are readily interconvertible in aqueous solution. Ribose in its linear form is known as pentose sugar with all of its hydroxyl functional groups (-OH) on the same side in its Fischer projection. Cyclisation of ribose occurs via hemiacetal formation due to attack on the aldehyde by the C-4 hydroxyl group to produce a furanose form or by the C-5 hydroxyl group to produce a pyranose form. At room temperature, about 76% of D-ribose is present in pyranose forms and 24% in the furanose forms while only about 0.1% of the linear form.

Beta-D-ribofuranose | Alpha-D-ribopyranose | D-ribose | L-ribose

Structure of ribose:

1. The molecular formula of ribose as determined by elemental analysis and molecular weight is $C_5H_{10}O_5$.
2. Ribose reacts with hydrogen cyanide to form cyanohydrin and forms an oxime with hydroxyl amine, thus indicating the presence of carbonyl group.
3. Ribose gives tetrahydroxypentanoic acid on oxidation with bromine water. Hence the carbonyl group present in the ribose is –CHO.
4. Ribose on successive reduction with $NaBH_4$ and HI/red phosphorus, gives n-pentane. This shows that ribose has a straight chain of five carbon atoms.
5. Ribose forms ribose tetracetate with acetic anhydride indicating the presence of four hydroxyl groups. Since ribose is a stable compound, no two hydroxyl groups are present on the same carbon atom.

On the basis of above explained reactions of ribose, its structure may be written as:

$$CH_2OH — CHOH — CHOH — CHOH — CHO$$

Configuration: Ribose has three asymmetric carbon atoms in the molecule, hence 2^3=8 isomers are possible. Its configuration can be established from the fact that it forms the same osazone as arabinose. As only C-1 and C-2 are involved in the osazone formation , the configuration of arabinose and ribose must be identical at C-3 and C-4. Thus the configuration of ribose is:

CHO
H—C—OH
H—C—OH
H—C—OH
CH_2OH

Ribose

CHO
HO—C—H
H—C—OH
H—C—OH
CH_2OH

Arabinose

Although open chain structure, as established above could account for most of the properties of the ribose but it fails to explain the following properties:

(a) Unlike other aldehydes, ribose does not give Schiff's test neither does it form addition product with $NaSO_3H$.

(b) Ribose exists in two isomeric forms: α-D-ribose and β-D-ribose. If D-ribose has a free aldehyde group, it should not have two isomers.

(c) It shows mutarotation.

(d) Ribose reacts with one molecule of methanol in the presence of hydrochloric acid gas and forms two isomeric methyl ribosides namely methyl- α-D-riboside and methyl- β-D-riboside. These two compounds display the properties of acetals. These two forms did not give properties corresponding to aldehyde group. This shows that glycosides are fully acetals.

Beta-D-ribopyranose

Open chain form

Alpha-D-ribopyranose

So, D-ribose probably does not contain a free aldehydic carbonyl group and on the other hand it may be a hemiacetal since it adds only one molecule of an alcohol to form a full acetal. This is not impossible because the carbonyl group may undergo intramolecular hemiacetal formation with any one of its four hydroxyl groups. Thus, ribose must have a cyclic structure having one oxygen atom as the ring constituent =. Oxidative degradation and periodate oxidation of methyl-D-ribosides have proved that there are six-membered pyran-like ring in D-glucose.

An important consequence of the formation of this cyclic hemiacetal is the generation of new chiral centre at C-1. Two structures are thus possible differing in configuration at C-1. In aqueous solution these two hemiacetals undergo facile hydrolysis to open chain form which then again forms hemiacetal ring yielding either of the anomers. Thus, equilibrium mixture contains both the anomers as well as the open-chain form. This interconversion explains the change in rotation.

Basically, ribose exists in both pyranose as well as furanose form. Pyranose form is more stable than the furanose form. Equilibrium mixture of ribose contains 56% β-D-ribopyranose, 20% α-D-ribopyranose, 18% β-D-ribofuranose and 6% α-D-ribofuranose.

Deoxyribose: Deoxyribose, also known as D-Deoxyribose and 2-deoxyribose, is a pentose sugar that is a key component of the nucleic acid deoxyribonucleic acid (DNA). It is derived from the pentose sugar ribose. Deoxyribose has the chemical formula $C_5H_{10}O_4$. Deoxyribose is the sugar component of DNA, just as ribose serves that role in RNA. Alternating with phosphate bases, deoxyribose forms the backbone of the DNA, binding to the nitrogenous bases adenine, thymine, guanine, and cytosine. In RNA, uracil is the base rather than thymine.

Structure of Deoxyribose:

1. The molecular formula of deoxy-D-ribose is $C_5H_{10}O_4$.
2. Deoxy-D-ribose reacts with hydrogen cyanide to form cyanohydrin and forms an oxime with hydroxylamine, thus indicating the presence of carbonyl group.
3. Deoxy-D-ribose gives trihydroxypentanoic acid on oxidation with bromine water. Hence the carbonyl group present in deoxy-D-ribose is –CHO.
4. Deoxy-D-ribose on successive reduction with $NaBH_4$ and HI/red phosphorus gives n-pentane. This shows that ribose has a straight chain of five carbon atoms.
5. Deoxy-D-ribose forms triacetate derivative with acetic anhydride indicating the presence of only three hydroxyl groups. This shows that

one of the carbon of deoxy-D-ribose has no hydroxyl group. Since deoxy-D-ribose is a stable compound, no two hydroxyl groups are present on the same carbon atom.

6. Deoxy-D-ribose forms phenylhydrazone. But it does not form osazone. This shows that deoxy-D-ribose has no hydroxyl group at C-2 carbon.

On the basis of above reaction findings, the structure for deoxy-D-ribose may be written as follows:

$$CH_2OH \text{---} CHOH \text{---} CHOH \text{---} \overset{H_2}{C} \text{---} CHO$$

Deoxy-D-ribose has two asymmetric carbon atoms in the molecule, hence 2^2=4 isomers are possible.

Configuration of deoxy-D-ribose: Its configuration has been established by converting glucose into deoxy-D-ribose. On this basis the configuration of deoxy-D-ribose is:

Deoxy ribopyranose (74.3%) ⇌ Deoxy ribose (0.7%) ⇌ Deoxy ribofuranose (25.0%)

Chemical equilibrium of deoxyribose in solution form

Although open chain structure, as established above, could account for most of the properties of deoxy-D-ribose but it fails to explain the following properties:

(a) Unlike the aldehydes, deoxy-D-ribose does not give Schiff's test neither does it form addition product with sodium bisulphate.

(b) Deoxy-D-ribose exists in two isomeric forms: α- deoxy-D-ribose and β-deoxy-D-ribose. If deoxy-D-ribose has a free aldehyde group, it should not have two isomers.

(c) It shows mutarotation.

(d) Deoxy-D-ribose reacts with one molecule of methanol in the presence of hydrochloric acid gas and forms two isomeric methyl deoxy-D-ribosides

namely methyl α- deoxy-D-riboside and methyl β- deoxy-D-riboside. These two compounds show properties of acetals. These two forms do not give properties corresponding to aldehyde group. This shows that glycosides are fully acetals.

So deoxy-D-ribose probably does not contain a free aldehyde carbonyl group and on the other hand, it may be a hemiacetal since it adds only one molecule of an alcohol to form a full acetal. This is not impossible because the carbonyl group may undergo intramolecular hemiacetal formation with any one of its three hydroxyl groups. Thus deoxy-D-ribose must have a cyclic structure having one oxygen atom as the ring constituent. Oxidative degradation and periodate oxidation of methyl deoxy-D-ribosides have proved that there are six membered pyran like ring in 2-deoxy-D-ribose.

An important consequence of the formation of this cycle hemiacetal is the generation of new chiral centre at C-1. Two structures are thus possible differing in configuration at C-1. In aqueous solution these two hemiacetals undergo facile hydrolysis to open chain form which then again forms hemiacetal ring yielding either of the anomers. Thus, equilibrium mixture contains both the anomers as well as the open chain form. This interconversion explains the change in rotation.

Beta-deoxy-D-ribopyranose   Open chain form   Alpha-2-deoxy-D-ribopyranose

In fact, 2-deoxy-D-ribose exists in both pyranose as well as furanose form but pyranose form is more stable than the furanose form. In DNA 2-deoxy-D-ribose is present in furanose form.

Disaccharides:

On hydrolysis, disaccharides give two monosaccharide units which may be similar or dissimilar in nature. For example, sucrose on hydrolysis gives one molecule each of glucose and fructose, whereas maltose gives two molecules of glucose.

$$C_{12}H_{22}O_{11} + H_2O \xrightarrow{\text{Acid/enzyme}} C_6H_{12}O_6 + C_6H_{12}O_6$$

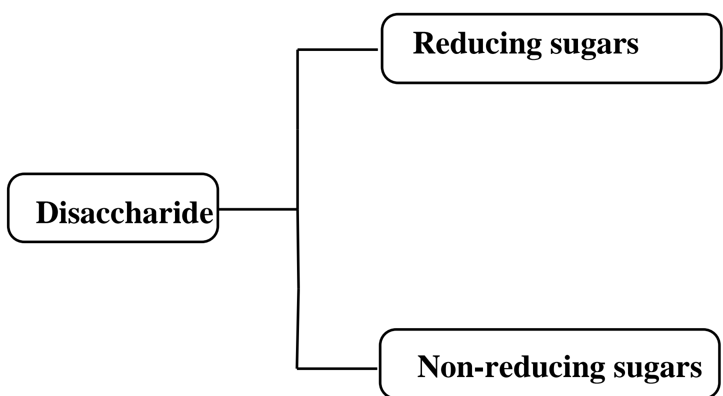
Sucrose Glucose Fructose

$$C_{12}H_{22}O_{11} + H_2O \xrightarrow{\text{Acid/enzyme}} 2\ C_6H_{12}O_6$$

Maltose Glucose

Because glycosides are acetals (or ketals), they are not in equilibrium with the open chain aldehyde (or ketone) in neutral or basic aqueous solutions. Because they are not in equilibrium with a compound with a carbonyl group, they cannot be oxidized by reagents such as Ag^+ or Br_2 Glycosides, therefore, are nonreducing sugars-they cannot reduce Ag^+ or Br_2.

Hemiacetals (or hemiketals) are in equilibrium with the open-chain sugars in aqueous solution. So as long as a sugar has an aldehyde, a ketone, a hemiacetal, or a hemiketal group, it is able to reduce an oxidizing agent and therefore is classified as a reducing sugar. Without one of these groups, it is a nonreducing sugar.



Reducing disaccharides: Those saccharides which have free aldehyde or keto group are known as reducing disaccharides. For example: maltose, lactose, etc. Reducing sugars give positive test for Benedict's, Fehlings, and Tollen's reagent because of free aldehyde present. Sugars with a hemiacetal functional group give positive tests since they are in equilibrium with an open-chain aldehyde. Monosaccharides with a hemacetal are also act as reducing sugars because their open-chain form contains aldehyde. Disaccharide sugars which do

not give positive tests for Benedict, Fehlings, and Tollens reagents are known as reducing sugars.

Alpha-D-(+) Glucopyranose

D-(+) Glucose

(Open chain form)

Beta-D-(+) Glucopyranose

Acetal

Hemiacetal

Lactose

Non-reducing disaccharides: Those saccharides which do not have free aldehyde or keto group are known as non-reducing disaccharides. For example: sucrose, trehalose, etc. There are three main criterion for being non-reducing saccharides:

(a) Mono and di-saccharides which lack a hemiacetal

(b) Polysaccharides where the ratio of hemiacetals to acetal linkages is very low (e.g. starch)

(c) Disaccharide sugars which do not give positive tests for Benedict, Fehlings, and Tollens reagents are known as non-reducing sugars.

If the hemiacetal group of a monosaccharide forms an acetal by reacting with an alcohol group of another monosaccharide, the glycoside that is formed is a disaccharide. Disaccharides are compounds consisting of two

monosaccharide subunits hooked together by an acetal linkage. For example, maltose is a disaccharide obtained from the hydrolysis of starch. It contains two D-glucose subunits hooked together by an acetal linkage. This particular acetal linkage is called an α-1,4'-glycosidic linkage. The linkage is between C-1 of one sugar subunit and C-4 of the other. The "prime" superscript indicates that C-4 is not in the same ring as C-1. The linkage is an α-1,4'-glycosidic linkage because the oxygen atom involved in the glycosidic linkage is in the α-position. Remember that the α-position is axial when a sugar is shown in a chair conformation and is down when the sugar is shown in a Haworth projection; the β-position is equatorial when a sugar is shown in a chair conformation and is up when the sugar is shown in a Haworth projection.

The configuration of this carbon is not specified

An alpha-1,4'-glycisidic linkage

Maltose

Maltose: The molecular formula of maltose is $C_{12}H_{22}O_{11}$ and its systemic name is 4-O-α-D-glucopyranosyl-D-glucose. Maltose, or malt sugar, is a disaccharide formed from two units of glucose joined with an α(1→4) linkage. It is the second member of an important biochemical series of glucose chains. The addition of another glucose unit yields maltotriose; further additions will produce dextrins (also called maltodextrins) and eventually starch.

Maltose can be broken down into two glucose molecules by hydrolysis. In living organisms, the enzyme maltase can achieve this very rapidly. In the laboratory, heating with a strong acid for several minutes will produce the same result. The production of maltose from germinating cereals, such as barley, is an important part of the brewing process. When barley is malted, it is brought into a condition in which the concentration of maltose-producing amylases has been maximized. Mashing is the process by which these amylases convert the cereal's

starches into maltose. Metabolism of maltose by yeast during fermentation then leads to the production of ethanol and carbon dioxide.

$$\text{Starch} \xrightarrow{\text{Beta-amylase}} \underset{\text{Maltose}}{C_{12}H_{22}O_{11}}$$

Structure of Maltose:

1. The molecular formula of maltose is $C_{12}H_{22}O_{11}$.

2. One mole of maltose gives two moles of glucose on acid hydrolysis. Thus maltose is disaccharide of glucose.

3. Maltose is a reducing sugar; it gives positive test with Fehling's and Tollens solutions. Maltose also reacts with phenylhydrazine to form a monophenylosazone.

4. Maltose exists in two anomeric forms: α-maltose; [α] = +168° and β-maltose; [α] = +112°.

5. Maltose on oxidation with bromine water gives monobasic acid, maltobionic acid. Points 3, 4 and 5 indicate one of the glucose reduces of maltose is present in a hemiacetal (reducing unit) form; the other, therefore must be present as a glucoside (aetal unit).

6. Maltose is hydrolysed by yeast enzyme maltase (or α-glucosidase) which is characteristic of α-glucosides. This result shows that the glycosidic linkage in maltose is α.

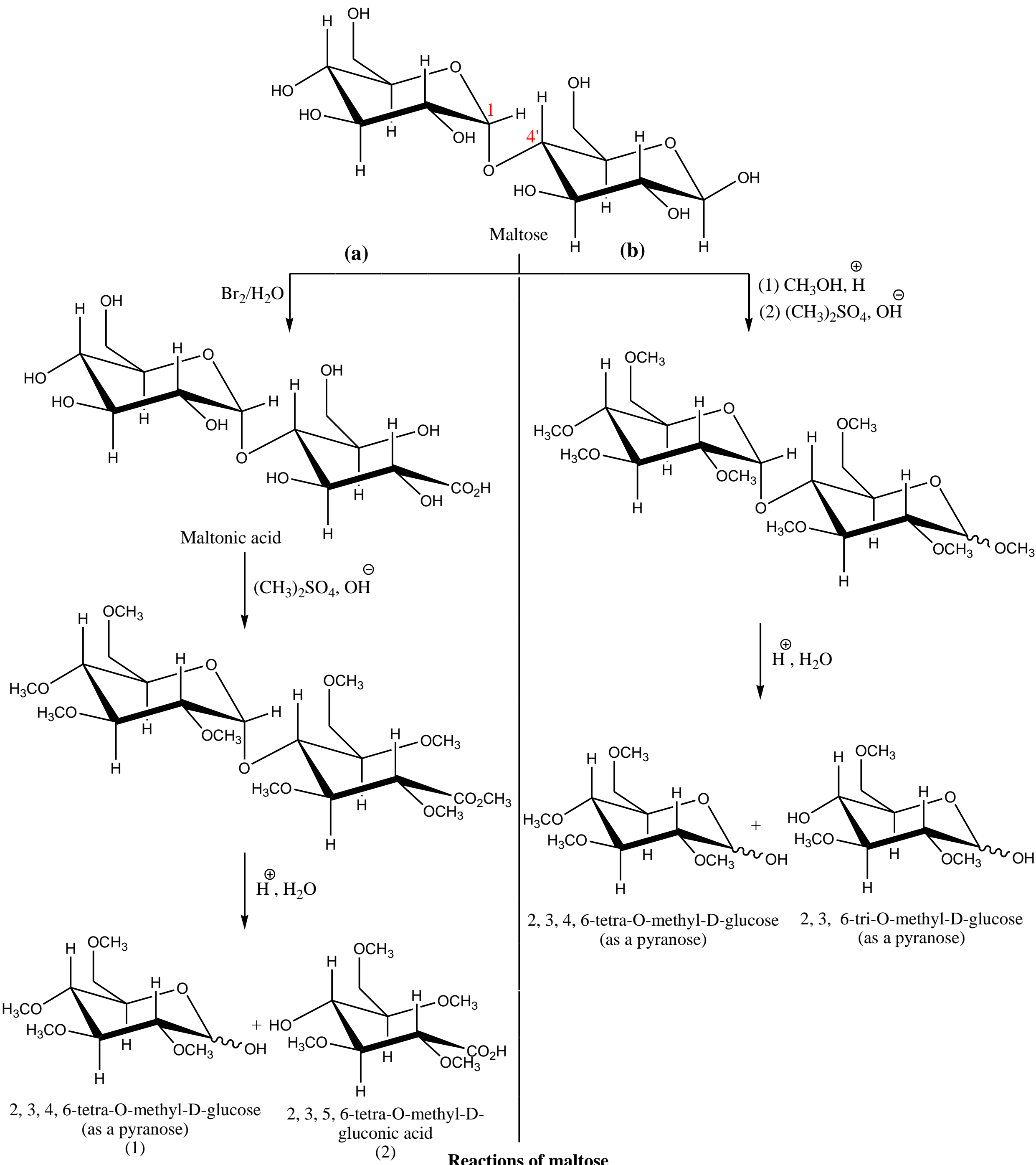
7. Maltose on oxidation with bromine water gives maltobionic acid. Maltobionic acid on methylation with $(CH_3)_2SO_4$ in the presence of NaOH gives octa-O-methylmaltobionic acid which on acid hydrolysis gives 2, 3, 4, 6-tetra-O-methylglucose and 2, 3, 5, 6-tetra-O-methylgluconic acid.

Formation of first product (1) indicates that it is obtained from the non reducing glucose portion because it has no carboxylic group. The non reducing glucose portion is present in pyranose form because compound (1) has hydroxyl group at C-5.

That the second product, 2, 3, 5, 6-tetra-O-methylgluconic acid has a free hydroxy at C-4 indicates that C-4 was involved in a glycosidic linkage with the non reducing glucose. Thus, this experiment confirms that intersugar linkage is

1→4 and non reducing glucose portion is present in pyranose form. This result does not indicate the ring size of the reducing glucose portion.

8. Maltose on complete methylation (i.e., premethylation of maltose) yields octa-O-methylmaltose which has no reducing properties. This on hydrolysis with very dilute acid gives hepta-O-methylmaltose. This hepta-O-methylmaltose has reducing properties. Thus octa-O-methylmaltose is methyl hepta-O-methylmaltoside. This also confirms that out of two glucose units present in the maltose, one is reducing and other is non-reducing.



Reactions of maltose

Step (a): Oxidation of maltose to maltonic acid followed by methylation and hydrolysis
Step (b): Methylation and subsequent hydrolysis of maltose itself

Hydrolysis of hepta-O-methylmaltose or Octa-O-methylmaltose gives 2, 3, 4, 6-tetra-O-methylglucose and 2, 3, 6-tri-O-methylglucose**.** The free -OH at C-5 in 2, 3, 6-tri-O-methylglucose indicates that it must involved in hemiacetal formation (i.e., oxide ring because C-4 –OH is involved in a glycosidic linkage) and that the reducing glucose is present as a pyranose.

Thus, maltose is 4-O-(α-D-glucopyranosyl)-D-glucopyranose and its structure can be represented as follows:

(+)-Maltose (beta-anomer)
4-O-(alpha-D-glucopyranosyl)-D-glucopyranose

This structure of maltose explains all the properties of maltose.

Lactose: Lactose is one of the main constituents of human and animal milk. As a disaccharide consisting of glucose and galactose, lactose acts as an energy-carrier in milk. Due to its physiological and functional characteristics, industrially manufactured lactose is used today in a large number of foodstuffs as well as in the pharmaceutical industry. Lactose is produced from whey, a byproduct of cheese making and casein production, by crystallizing an oversaturated solution of whey concentrate. Lactose is the primary disaccharide in virtually all mammalian milks. It is unique among the major dietary sugars because of the β-1→4 linkage between its component monosaccharides, galactose and glucose. Lactose production in nature is limited to the mammalian breast, which contains the enzyme system (lactose synthase) necessary to create this linkage. Human milk contains approximately 7% lactose by weight, which is among the highest lactose concentrations of all mammalian milks while cow's milk contains 4 to 5% lactose.

Structure of lactose:

1. The molecular formula of lactose is $C_{12}H_{22}O_{11}$.

2. One mole of lactose gives one mole of glucose and one mole of galactose on acid hydrolysis. Thus lactose is disaccharide of glucose and galactose.

3. Lactose is a reducing sugar; it gives positive test with Fehling's and Tollens solutions. Lactose also reacts with phenylhydrazine to form a monophenylosazone.

4. Lactose exists in two anomeric forms: α-lactose; and β-lactose.

5. Lactose on oxidation with bromine water gives lactobionic acid.

Points 3, 4 and 5 indicate one of the monosaccharide of lactose is present in a hemiacetal (reducing unit) form; the other, therefore must be present as a glycoside (aetal unit).

6. Lactose is hydrolysed by the enzyme emulsion (or β-galactosidase) which is specific for the hydrolysis of β-galactoside links. This enzymic hydrolysis shows that the glycosidic linkage in lactose is β.

7. Lactose on oxidation with bromine water gives lactobionic acid. Lactobionic acid on permethylation with $(CH_3)_2SO_4$ in the presence of NaOH followed by hydrolysis gives two compounds; 2, 3, 4, 6-tetra-O-methylgalactose (1) and 2, 3, 5, 6-tetra-O-methyl-D-gluconic acid (2).

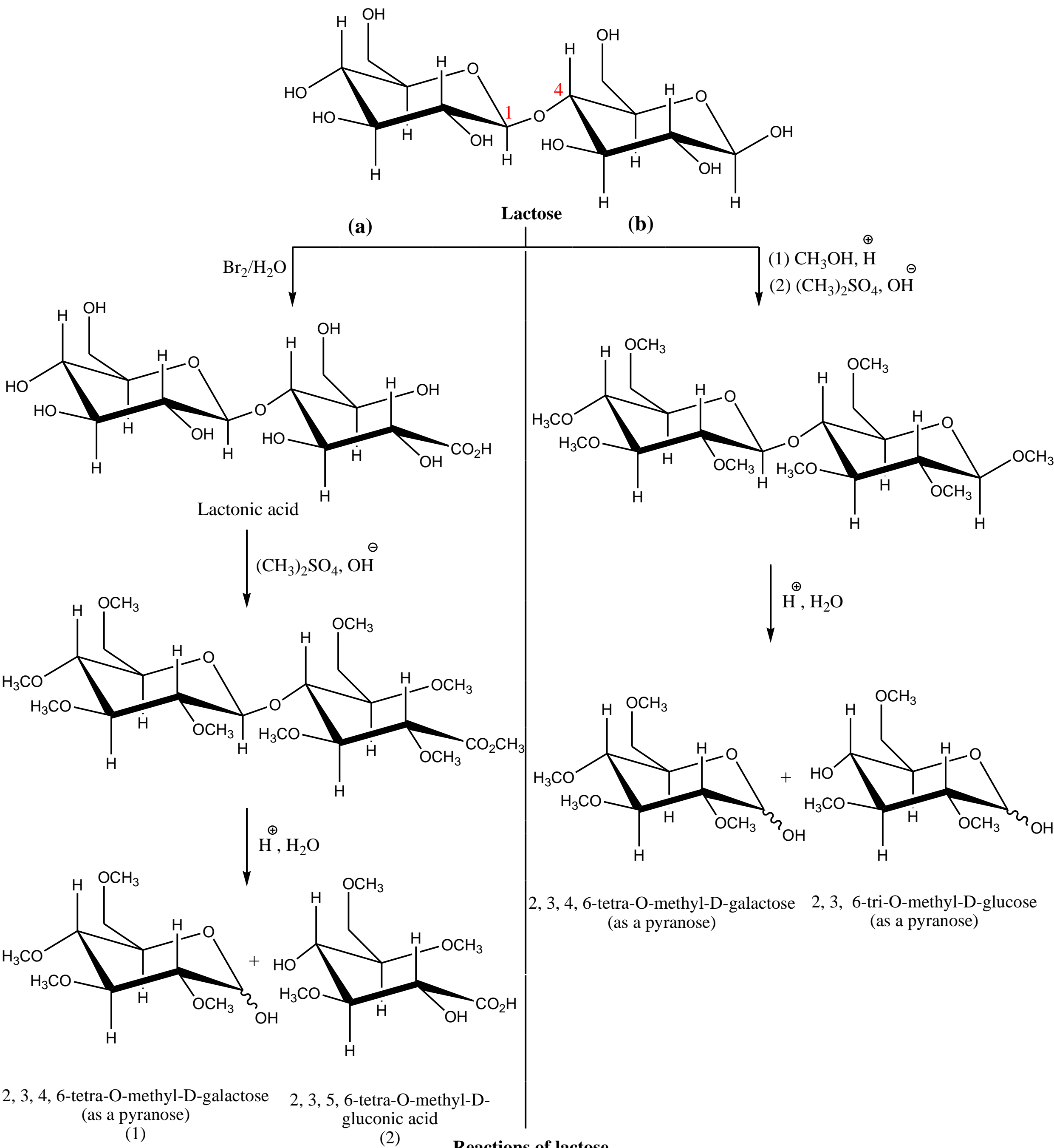
Formation of compound (2) indicates that the glucose unit is present in the hemiacetal form. Similarly formation of compound (1) indicates that the galactose unit is present in the galactoside (acetal) form. Identification of the compounds (1) and (2) indicates that:

(i) D-galactose has a pyranose ring because C-5 OH is free (1) and

(ii) Glucose unit is bonded to galactose unit by C-4 OH since this OH is free (2)

This experiment confirms that inter sugar linkage is 1→4 and non reducing galactose unit is present in pyranose form. This result does not indicate the ring size of the reducing glucose portion.

8. Lactose on complete methylation gives methyl hepta-O-methyllactoside which on stepwise acid hydrolysis gives 2, 3, 6-tri-O-methylglucose and 2, 3, 4, 6-tetramethyl-O-galactose indicating that both the rings are pyranose rings since in both the case C-5 OH groups are free in addition to C-4 OH of the glucose unit.



Reactions of lactose

Step (a): Oxidation of lactose to lactonic acid followed by methylation and hydrolysis

Step (b): Methylation and subsequent hydrolysis of lactose

On the basis of the above results structure of lactose can be represented as follows:

Lactose
4-O-(beta-D-galactopyranosyl)-beta-D-glucopyranose

The above structure of lactose explains all the properties of the lactose.

Sucrose: Sucrose is a nonreducing disaccharide composed of glucose and fructose linked via their anomeric carbons. It is obtained commercially from sugarcane, sugar beet (beta vulgaris), and other plants and used extensively as a food and a sweetener. Sucrose is derived by crushing and extraction of sugarcane (Saccharum officinarum) with water or extraction of the sugar beet (Beta vulgaris) with water, evaporating, and purifying with lime, carbon, and various liquids. Sucrose is also obtainable from sorghum. Sucrose occurs in low percentages in honey and maple syrup. Sucrose is used as a sweetener in foods and soft drinks, in the manufacture of syrups, in invert sugar, confectionery, preserves and jams, demulcent, pharmaceutical products, and caramel. Sucrose is also a chemical intermediate for detergents, emulsifying agents, and other sucrose derivatives. Sucrose is wiidespread in seeds, leaves, fruits, flowers and roots of plants, where it functions as an energy store for metabolism and as a carbon source for biosynthesis. The annual world production of sucrose is in excess of 90 million tons mainly from the juice of sugar cane (20%) and sugar beet (17%). In addition to its use as a sweetner, sucrose is used in food products as a preservative, antioxidant, moisture control agent, stabilizer and thickening agent.

Structure of sucrose:

1. The molecular formula of sucrose is $C_{12}H_{22}O_{11}$.

2. On hydrolysis, one mole of sucrose gives one mole of glucose and one mole of fructose. Sucrose has a specific rotation of + 65.5°. When it is hydrolyzed, the resulting equimolar mixture of glucose and fructose has specific rotation of -22.0°. Because of the change in the sign of the rotation when sucrose is hydrolyzed, a 1:1 mixture of glucose and fructose

is called invert sugar. The enzyme that catalyses the hydrolysis of sucrose is called invertase. The most common form of invert sugar is honey, a supersaturated mixture of glucose and fructose hydrolyzed from sucrose by the invertase enzyme of honeybees.

3. Sucrose gives negative test with Fehling's and Tollens solutions. It does not form an osazone and does not undergo mutarotation. Sucrose does not undergo oxidation by bromine water.

These results show that in sucrose the hemiacetal/hemiketal group is absent and the sucrose is a non reducing sugar. Thus, the two monosaccharides must have a glycosidic linkage that involves C-1 of glucose and C-2 of fructose.

4. Since sucrose is non reducing sugar, the stereochemistry of glycosidic linkage in sucrose may be any of the following:

i. α, α (α-anomer of glucose combines with an α-anomer of fructose)
ii. α, β (α-anomer of glucose combines with a β -anomer of fructose)
iii. β, α (β -anomer of glucose combines with an α -anomer of fructose)
iv. β, β (β -anomer of glucose combines with a β -anomer of fructose)

The stereochemistry of glycosidic linkage in sucrose can be known from the enzymic hydrolysis. Sucrose is hydrolyzed by an α-glucosidase obtained from yeast but not by β-glucosidase enzymes. This hydrolysis indicates an α configuration at the glucoside portion. Sucrose is also hydrolysed by sucrose, an enzyme known to hydrolyze β fructosides. This hydrolysis indicates a β-configuration at the fructoside portion.

5. Sucrose on premethylation gives methyl octa-O-methylsucrose which on stepwise acid hydrolysis gives 2, 3, 4, 6-tetra-O-methylglucose and 1, 3, 4, 6-tetra-O-methylfructose. The identification of these products indicates that the glucose is present in a pyranose form and fructose is present in a furanose form.

Thus, the structure of sucrose can be represented as follows:

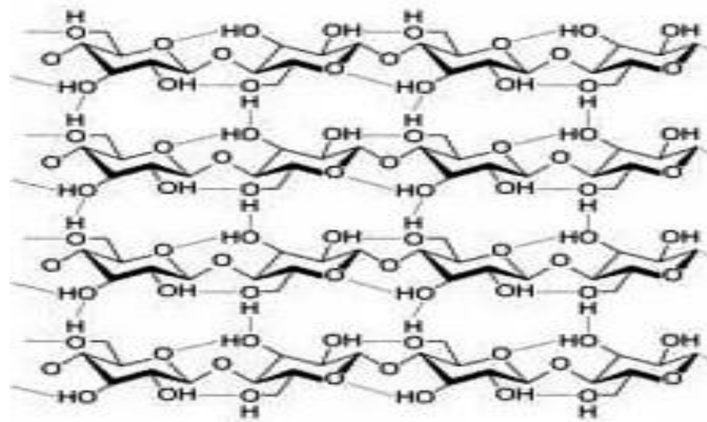
Glucose

Glycosidic bond

Fructose

Polysaccharides:

A polysaccharide is a type of carbohydrate. It is a polymer made of chains of monosaccharides that are joined by glycosidic linkages. Polysaccharides are also known as glycans. By convention, a polysaccharide consists of more than ten monosaccharide units, while an oligosaccharide consists of three to ten linked monosaccharides. Special enzymes bind these small monomers together creating large sugar polymers, or polysaccharides. A polysaccharide can be a homopolysaccharide, in which all the monosaccharides are the same, or a heteropolysaccharide in which the monosaccharides vary. Depending on which monosaccharides are connected, and which carbons in the monosaccharides connects, polysaccharides take on a variety of forms. A molecule with a straight chain of monosaccharides is called a linear polysaccharide, while a chain that has arms and turns is known as a branched polysaccharide. Natural saccharides are generally of simple carbohydrates called monosaccharides with general formula $(CH_2O)_n$ where n is three or more. Examples of monosaccharides are glucose, fructose, and glyceraldehyde. Polysaccharides, meanwhile, have a general formula of $C_x(H_2O)_y$ where x is usually a large number between 200 and 2500. When the repeating units in the polymer backbone are six-carbon monosaccharides, as is often the case, the general formula simplifies to $(C_6H_{10}O_5)_n$, where typically $40 \leq n \leq 3000$.



The glycosidic bonds between monosaccharides consist of an oxygen molecule bridging two carbon rings. The bond is formed when a Hydroxyl group is lost from the carbon of one molecule, while the hydrogen is lost by the hydroxyl group of another monosaccharide. The carbon on the first molecule will substitute the oxygen from the second molecule as its own, and glycosidic bond is formed. Because two molecules of hydrogen and one oxygen is expelled, the reaction produced a water molecule as well. This type of reaction is called a dehydration reaction as water is removed from the reactants.

Out of many naturally occurring polysaccharides, the two most common and important are starch and cellulose.

Starch:

The starches occur widespread as reserve carbohydrate in tubers such as potatoes, in many fruits, grains and seeds. In the grains, the starch is arranged in concentric layers. When starch grains are treated with boiling water, the substance in the center passes into solution, but the greater part of the grain is not soluble. This insoluble portion swells as it absorbs water and the whole mass becomes starch paste. Both the soluble portion and the insoluble portion are heterogeneous mixtures. The soluble fraction is referred to as amylose, and the insoluble fraction as amylopectin. Most starches contain 80-90 per cent amylopectin and 10-20 per cent amylose. Amylose and amylopectin can be separated by taking advantage of the difference in solubility in water. Both amylose and amylopectin are polymers of glucose and upon hydrolysis with acid they give D-glucose as the product.

1. Amylose: Amylose is an unbranched long chain polymer in which the glucose residues are linked through α-1, 4- glycosidic linkages. The amylose structure may be ragarded as a repeated maltose structure with a free sugar group (acetal group) at one end. This is also known as the reducing end (nth residue), whereas, the opposite end (first residue) is referred to as the non-reducing end. Any particular preparation of amylose usually consists of a mixture of populations of molecules, which differ widely in chain length (number of glucose residues per chain). The molecular weight of amylose is about 40,000, i.e., a polymer containing 200 glucose units. The α-1, 4-glycosidic linkage is easily hydrolysed by an α-glucosidase enzyme, present in all animals.

Alpha-1,4'-glycosidic linkage

2. Amylopectin:

Amylopectin is a branched chain polymer containing about 10^6 D-glucose units. Like amylase, it is composed of D-glucose units joined by $\alpha-1\rightarrow4$ glucose linkages. Unlike amylose, amylopectin also contains $\alpha-1\rightarrow6$ glycosidic linkages. These linkages create the branches in the amylopectin. About every 20 to 30 glucose units, there is a branch point.

Cellulose:

Cellulose is the structural material of higher plants. Cotton, for example, is composed of about 90% cellulose, and wood is about 50% cellulose. Like amylose, cellulose is composed of unbranched chains of D-glucose units. Unlike amylose, however, the glucose units in cellulose are joined by β-1,4'-glycosidic linkages rather than by α-1,4'-glycosidic linkages.

All the glucose molecules in cellulose have the beta-configuration at the C1 atom, so all the glycosidic bonds that join the glucose molecules together are also of the beta type. This means that the cellulose molecule is straight, and many such molecules can lay side by side in a parallel series of rows. Tiny forces called hydrogen bonds hold the glucose molecules together, and the chains in close proximity. Although each hydrogen bond is very, very weak, when thousands or millions of them form between two cellulose molecules the result is a very stable, very strong complex that has enormous strength.

6.10 Summary

In this unit learner is able to define Carbohydrates, Classification of Carbohydrates, Classification of monosaccharides, Erythro and Threo Diastereomers, the cyclic structure of Glucose, Interconversion of straight-chain and ring forms of sugars, Configuration of monosaccharides. Conversion of glucose into mannose, Formation of glycosides, Ether formation via Williamson ether synthesis, Ester formation and Determination of ring size of monosaccharides. About Ribose, Deoxyribose, Disaccharides such as Maltose, Lactose and Sucrose. Polysaccharides such as Starch, Amylose, Amylopectin and Cellulose.

6.11 Terminal Questions

Q1. Give brief on Classification of Carbohydrates?

Q2. Define Epimers?

Q3. Explain Mutarotation in Carbohydrates?

Q4. Explain Conversion of glucose into mannose?

Q5. Define Erythro and Threo Diastereomers?

Q6. Define Lactose.

Q7. Ribose sugar and Structure of ribose

Solution Q1: Classification of Carbohydrates: Carbohydrates can be classified into three major groups which further divided into several groups. A schematic representation is given as follows:

1. **Monosaccharides:** The simplest carbohydrates that cannot be hydrolyzed into simple carbohydrates, are called monosaccharides.

2. **Oligosaccharides:** Carbohydrates that hydrolyze to yield 2 to 10 molecules of monosaccharides are called oligosaccharides. The most common oligosaccharides are disaccharides which include maltose, lactose, sucrose, etc.

3. **Polysaccharides:** Carbohydrates that yield a large number of monosaccharide molecules (more than 10 units), are known as polysaccharides. The common examples are starch, cellulose, glycogen, etc.

Solution Q2: The two L isomers; L-erythrose and L-threose is the enantiomer of D-threose. L-threose is a diastereomer of both D and L-erythrose, and L-erythrose is a diastereomer of both D and L-threose. Diastereomers that differ from each other in configuration at only one chiral carbon are called epimers; erythrose and D-threose are epimers. Glucose and galactose are also examples of epimers.

Solution Q3: Mutarotation is possible to obtain a sample of crystalline glucose in which all the molecules have the α structure or all have the β structure. The α form melts at 146°C and has a specific rotation of +112°, while the β form melts at 150°C and has a specific rotation of +18.7°. When the sample is dissolved in water, an interesting change in the specific rotation is observed. When the α-anomer dissolves, its specific rotation gradually decreases from an initial value of +112° to +52.6°. When the pure β-anomer dissolves, its specific rotation gradually increases from +18.7° to the same value of +52.6°. This change in the value of specific rotation to a mutual value is called mutarotation. Mutarotation occurs because in solution the two anomers can interconvert. When either of the pure anomers dissolves in water, the rotation of the pure anomer gradually changes to an intermediate rotation that results from the equilibrium

concentrations of the anomers. Thus mutarotation is due to the conversion of either anomer to an equilibrium mixture of both. At equilibrium, the mixture consists of about 36% α-D-glucose, 64% β-D-glucose, and less than 0.02% of the open-chain aldehyde form. The observed rotation of this solution is +52.7°.

Solution Q4: When glucose is treated with very dilute alkalies or organic bases such as pyridine or quinoline, epimers of glucose are obtained and thus glucose yields a mixture of mannose, fructose and unreacted glucose. The mechanism of this reaction was suggested by Lobrg de Druyn and Van Ekenstein involving the formation of an intermediate enediol.

The hydrogen atom attached to the carbon to the carbonyl (C2 in glucose) enolysis to form an enediol, thus destroying the asymmetry of C2. On ketonisation the 2 epimeric aldoses are formed. If the second hydrogen on C2 migrates to C1 a ketose is formed (fructose). Thus we can convert glucose into mannose and fructose.

The reaction is best carried out by the epimerization of aldonic acids which are more stable towards alkaline medium. Thus the aldose is first oxidized to the aldonic acid/ gluconic acid, which is then heated with an organic base like pyridine or quinoline and thus it converted to mannonic acid which is then lactonised and reduced to give mannose.

Glucose → $[O]$, Br_2/H_2O → Gluconic acid → Organic base, Heat → Mannonic acid → Heat, $-H_2O$ → lactone → $[H]$ → Mannose

Solution Q5: *Erythro* and *Threo* system of nomenclature is used for aldotetroses only. Aldotetrose have two chiral centres and hence shows four stereoisomers out of which two are D-sugars and remaining two stereoisomers are L-sugars. When Fisher projections are drawn for stereoisomers with two adjacent chiral centers, the pair of enantiomers with similar groups on the same side of the carbon chain is called the *erythro* enantiomers. The pair of enantiomers with similar groups on opposite sides is called the *threo* enantiomers. *Erythrose* and *threose* are diastereoisomers.

D-Erythrose *L-Erythrose* *D-Threose* *L-Threose*

Solution Q6: Lactose is one of the main constituents of human and animal milk. As a disaccharide consisting of glucose and galactose, lactose acts as an

energy-carrier in milk. Due to its physiological and functional characteristics, industrially manufactured lactose is used today in a large number of foodstuffs as well as in the pharmaceutical industry. Lactose is produced from whey, a byproduct of cheese making and casein production, by crystallizing an oversaturated solution of whey concentrate. Lactose is the primary disaccharide in virtually all mammalian milks. It is unique among the major dietary sugars because of the β-1→4 linkage between its component monosaccharides, galactose and glucose. Lactose production in nature is limited to the mammalian breast, which contains the enzyme system (lactose synthase) necessary to create this linkage. Human milk contains approximately 7% lactose by weight, which is among the highest lactose concentrations of all mammalian milks while cow's milk contains 4 to 5% lactose.

Structure of lactose:

1. The molecular formula of lactose is $C_{12}H_{22}O_{11}$.
2. One mole of lactose gives one mole of glucose and one mole of galactose on acid hydrolysis. Thus lactose is disaccharide of glucose and galactose.
3. Lactose is a reducing sugar; it gives positive test with Fehling's and Tollens solutions. Lactose also reacts with phenylhydrazine to form a monophenylosazone.
4. Lactose exists in two anomeric forms: α-lactose; and β-lactose.
5. Lactose on oxidation with bromine water gives lactobionic acid.

Solution Q7: Ribose is a simple sugar and carbohydrate with molecular formula $C_5H_{10}O_5$ and the linear-form composition $H-(C=O)-(CHOH)_4-H$. Ribose naturally occurs in the form of D-ribose which is an essential component of the ribonucleotides from which RNA is built. It has a structural analog, deoxyribose, which is a similarly essential component of DNA. Emil Fischer and Oscar Piloty in year 1891, first time synthesized an unnatural sugar L-Ribose. Like other sugars, ribose also exists as mixture of cyclic and linear forms in equilibrium and are readily interconvertible in aqueous solution. Ribose in its linear form is known as pentose sugar with all of its hydroxyl functional groups (-OH) on the same side in its Fischer projection. Cyclisation of ribose occurs via hemiacetal formation due to attack on

the aldehyde by the C-4 hydroxyl group to produce a furanose form or by the C-5 hydroxyl group to produce a pyranose form. At room temperature, about 76% of D-ribose is present in pyranose forms and 24% in the furanose forms while only about 0.1% of the linear form.

Beta-D-ribofuranose | Alpha-D-ribopyranose | D-ribose | L-ribose

Structure of ribose:

6. The molecular formula of ribose as determined by elemental analysis and molecular weight is $C_5H_{10}O_5$.
7. Ribose reacts with hydrogen cyanide to form cyanohydrin and forms an oxime with hydroxyl amine, thus indicating the presence of carbonyl group.
8. Ribose gives tetrahydroxypentanoic acid on oxidation with bromine water. Hence the carbonyl group present in the ribose is –CHO.
9. Ribose on successive reduction with $NaBH_4$ and HI/red phosphorus, gives n-pentane. This shows that ribose has a straight chain of five carbon atoms.
10. Ribose forms ribose tetracetate with acetic anhydride indicating the presence of four hydroxyl groups. Since ribose is a stable compound, no two hydroxyl groups are present on the same carbon atom.

On the basis of above explained reactions of ribose, its structure may be written as:

$$CH_2OH — CHOH — CHOH — CHOH — CHO$$

Configuration: Ribose has three asymmetric carbon atoms in the molecule, hence 2^3=8 isomers are possible. Its configuration can be established from the fact that it forms the same osazone as arabinose. As only C-1 and C-2 are

involved in the osazone formation , the configuration of arabinose and ribose must be identical at C-3 and C-4. Thus the configuration of ribose is:

Ribose	Arabinose
CHO	CHO
H—C—OH	HO—C—H
H—C—OH	H—C—OH
H—C—OH	H—C—OH
CH_2OH	CH_2OH

Although open chain structure, as established above could account for most of the properties of the ribose but it fails to explain the following properties:

(e) Unlike other aldehydes, ribose does not give Schiff's test neither does it form addition product with $NaSO_3H$.

(f) Ribose exists in two isomeric forms: α-D-ribose and β-D-ribose. If D-ribose has a free aldehyde group, it should not have two isomers.

(g) It shows mutarotation.

(h) Ribose reacts with one molecule of methanol in the presence of hydrochloric acid gas and forms two isomeric methyl ribosides namely methyl- α-D-riboside and methyl- β-D-riboside. These two compounds display the properties of acetals. These two forms did not give properties corresponding to aldehyde group. This shows that glycosides are fully acetals.

Unit 7: Problem based on Spectroscopy (UV-Vis., IR and PMR)

7.1 Introduction

Objective

7.2 Problem based on Spectroscopy (UV-Vis.)

7.3 Problem based on Spectroscopy (IR)

7.4 Problem based on Spectroscopy (PMR)

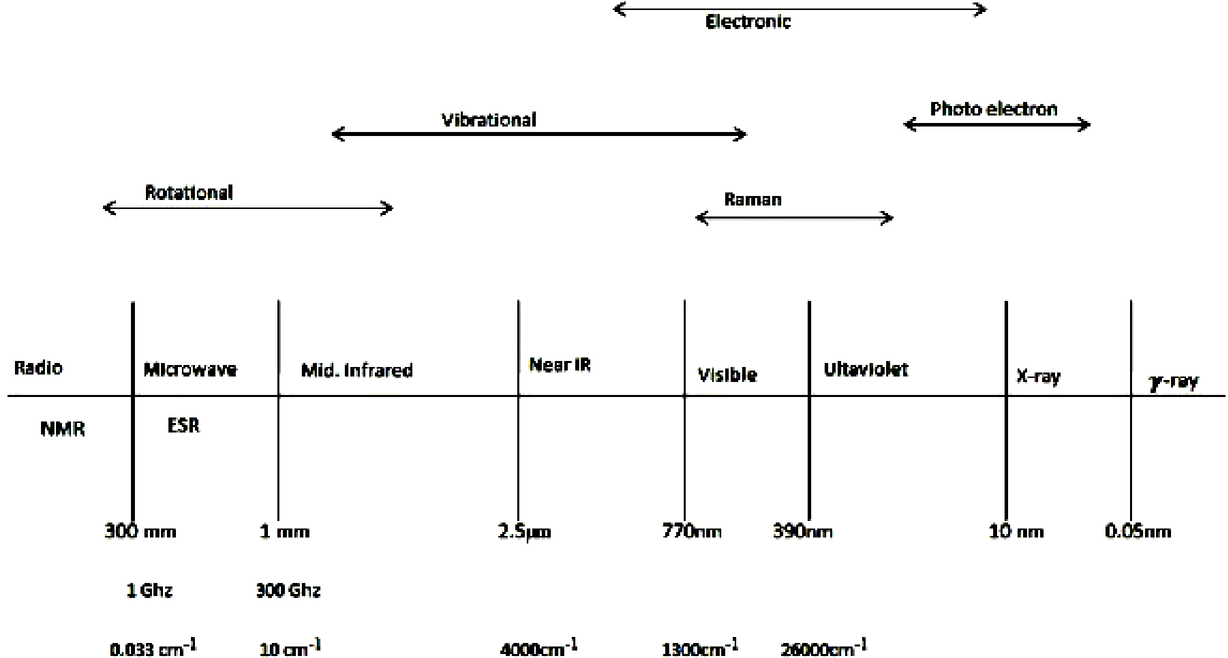
7.5 Problem based on Spectroscopy (UV-Vis., IR and PMR)

7.6 Summary
7.7 Terminal Questions

7.1 Introduction

In this unit we learn about problems based on UV-Vis Spectroscopy, IR Spectroscopy and NMR Spectroscopy as we studied earlier.

Electromagnetic Radiations: All types of radiations have the same velocity (2.998 x 10^8 m/s in vacuum) and require no medium for their propagation, i.e. they can travel even through vacuum. Electromagnetic radiations are characterized by frequencies, wavelengths and wave-numbers. Electromagnetic radiation includes or commonly refers to as 'light', radiation of longer and shorter wavelengths. Frequency v is defined as the number of waves which can pass through a point in one second, measured in cycles per second (cps) or hertz (Hz) (1Hz = 1 cps). Wavelength λ is defined as the distance between two consecutive crests [C] or troughs [T] measured in micrometer (μm) or micron (μ) (1 μm = 1, μ =10^{-6}m) or in nanometer (nm).



Regions of the electromagnetic spectrum

Objective

Learners have to practice to solve problems based on

- Problem based on Spectroscopy (UV-Vis.)
- Problem based on Spectroscopy (IR)
- Problem based on Spectroscopy (PMR)
- Problem based on Spectroscopy (UV-Vis., IR and PMR)

7.2 Problems based on UV spectrophotometry

Chromatophores are pigment-containing cells or groups of cells, found in a wide range of animals including amphibians, fish, reptiles, crustaceans and cephalopods. Mammals and birds, in contrast, have a class of cells called melanocytes for coloration. Chromatophores are largely responsible for generating skin and eye colour in ectothermic animals and are generated in the neural crest during embryonic development. Mature chromatophores are grouped into subclasses based on their colour (more properly "hue") under white light: xanthophores (yellow), erythrophores (red), iridophores (reflective/ iridescent), leucophores (white), melanophores (black/brown), and cyanophores (blue).

Chromophore	**Example**	**λ_{max} (nm)**	**Transition**
>C=C<	**Ethylene**	**171**	**π-π***
–C≡C–	**Acetylene**	**150**	**π-π***
>C=O	**Aectaldehyde**	**160**	**n-σ***
		180	**π-π***
		290	**n-π***
	Acetone	**166**	**n-σ***
		188	**π-π***
		279	**n-π***
–COOH	**Acetic acid**	**204**	**n-π***
$–CONH_2$	**Acetamide**	**178**	**π-π***
$–NO_2$	**Niromethane**	**201**	**π-π***
		274	**n-π***

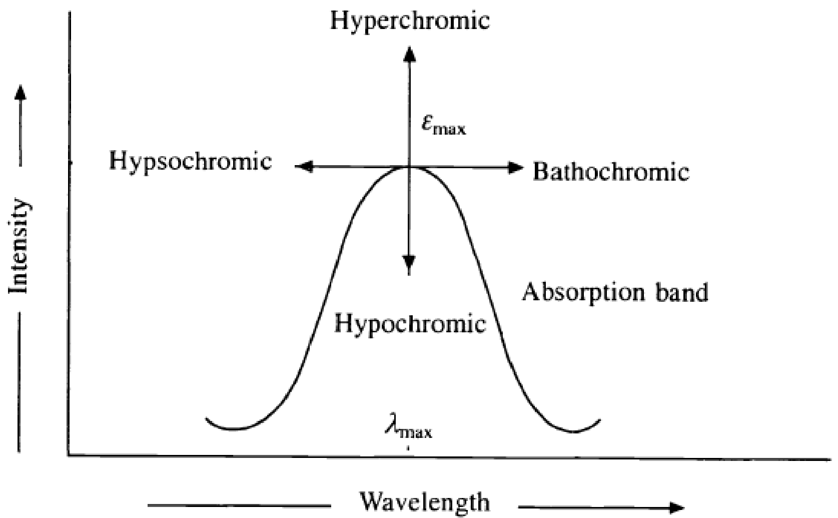
Auxochrome

A covalently saturated group when attached to a chromophore, changes both the wavelength and the intensity of the absorption maximum is known as

auxochrome, e.g. NH_2, OH, SH, halogens etc. Auxochromes generally increase the value of λ_{max} as well as ε_{max} by extending the conjugation through resonance. These are also called ***colour enhancing groups***. An auxochrome itself does not show absorption above 200 nm. Actually, the combination of chromophore and auxochrome behaves as a new chromophore having different values of λ_{max} and ε_{max}. For example, benzene shows λ_{max} 256 nm, ε_{max} 200, whereas aniline shows λ_{max} 280 nm, ε_{max} 1430 (both increased). Hence, NH_2 group is an auxochrome which extends the conjugation involving the lone pair of electrons on the nitrogen atom resulting in the increased values of λ_{max} and ε_{max}.

Absorption and Intensity Shifts

Bathochramic Shift: The shift of an absorption maximum to a longer wavelength due to the presence of an auxochrome or solvent effect is called a *bathochromic shift* or *red shift.* Example: Benzene shows λ_{max} 256 nm and Aniline shows λ_{max} 280 nm. Thus, there is a bathochromic shift of 24 nm in the λ_{max} of benzene due to the presence of the auxochrome NH_2. Similarly, a bathochromic shift of n -π^* band is observed in carbonyl compounds on decreasing solvent polarity, e.g. λ_{max} of acetone is at 264.5 nm in water as compared to 279 nm in hexane.



Shifts in absorption position and intensity

Hypsochromic Shift: The shift of an absorption maximum to a shorter wavelength is called *hypsochromic* or *blue shift*. This is caused by the removal of conjugation or change in the solvent polarity. Example: Aniline shows λ_{max} 280 nm, whereas anilinium ion (acidic solution of aniline) shows λ_{max} 254 nm. This hypsochromic shift is due to the removal of *n-π* conjugation of the lone pair of electrons of the nitrogen atom of aniline with the n-bonded benzene ring on protonation because the protonated aniline (anilinium ion) has no lone pair of electrons for conjugation. Similarly, there is a hypsochromic shift of 10-20 nm in the Amax of n-π* bands of carbonyl compounds on going from ethanol as solvent to hexane, i.e. on decreasing solvent polarity.

Hyperchromic Effect:An effect which leads to an increase in absorption intensity ε_{max} is called *hyperchromic effect*. The introduction of an auxochrome usually causes hyperchromic shift. Example: Benzene shows B-band at 256 nm, ε_{max} 200, whereas aniline shows B-band at 280 nm, ε_{max} 1430. The increase of 1230 in the value ε_{max} of aniline compared to that of benzene is due to the hyperchromic effect of the auxochrome NH_2.

Hypochromic Effec: An effect which Ieads to a decrease in absorption intensity ε_{max} is called *hypochromic effect*. This is caused by the introduction of a group which distorts the chromophore. For example, biphenyl shows λ_{max} 252 nm, ε_{max} 19,000, whereas 2,2'-dimethylbiphenyl shows λ_{max} 270 nm, ε_{max} 800. The decrease of 18,200 in the value of ε_{max} of 2,2'-dimethylbiphenyl is due to the hypochromic effect of the methyl groups which distort the chromophore by forcing the rings out of coplanarity resulting in the loss of conjugation.

Self Assessment Questions

Q1. What are the possible electronic transitions in following compounds?

(i) CH_4 (ii) CH_3Br (iii) CH_3COCH_3 (iv) CH_3CHO

Hint:

S. No.	Compound	Possible electronic transitions
(i)	CH_4	□□□□□□□
(ii)	CH_3Br	□□□□□□□□□□*n*□□□□□
(iii)	CH_3COCH_3	*n*□□□□□□□□□□□□□□□□□*n*□□□□□□

(iv)	CH_3CHO	n□□□□□□□□□□□□□□□□□□n□□□□□□

Q2. Calculate the frequency and energy of a photon having wavelength of 500 nm.

Hint:

$$\because c = \lambda\nu \therefore \nu = c/\lambda$$

$$\nu = \frac{2.998\times10^{8} m/s}{500\times10^{-9} m} = 5.976\times10^{14} Hz$$

$$E = hc/\lambda$$

$$E = \frac{(6.626\times10^{-34} J/s)(2.998\times10^{8} m/s)}{500\times10^{-9} m} = 3.973\times10^{-19} J$$

Q3. A UV-vis spectrophotometer is showing range of 195 nm to 900 nm. What will be this range in Hertz?

Hint: Before calculation change wavelength in meters.

When λ= 195 nm; then

$$\because c = \lambda\nu \therefore \nu = c/\lambda$$

$$\nu = \frac{3.0\times10^{8} m/s}{1.95\times10^{-7} m} = 1.53\times10^{15} s^{-1}$$

When λ= 195 nm; then

$$\nu = \frac{3.0\times10^{8} m/s}{9.00\times10^{-7} m} = 3.33\times10^{14} s^{-1}$$

Q4. Explain why it is considered a "best practice" to keep the total absorbance below 1 when conducting quantitative work.

Hint: In practice the slope of A vs c is usually linear if the total absorbance is kept below 1. When the absorbance exceeds 1, the slope of the line often curves downward and flattens out. Remember that absorbance is -log (T) so at an absorbance of 1, only 10% of the radiant power is reaching the detector. For this reason, it is considered a best practice to keep absorbance values below 1.

Q5. Explain why it is important that the ionic strength of your blank be similar to the ionic strength of your samples, even if the salt does not absorb light at the frequency of interest.

Hint: If your analyte has strong intermolecular forces of attraction for itself (for example it is capable of hydrogen bonding), or if there are other chemical

species in solution that can polarize the orbitals of your solute (such as a salt) then it is possible that ε will vary with concentration. Ionic strength can also affect the density of the solution and therefore influence the refractive index of the solution.

Q6. What are the two most common sources used in a UV-vis spectrophotometer?

Hint: A deuterium discharge lamp and a tungsten-halogen bulb.

Q7. Why is it a common design feature to see two different sources used in a UV-vis spectrophotometer?

Hint: In the most common design, the UV portion of the source radiation is provided by a deuterium discharge lamp and the visible and near infrared radiation is provided by the tungsten-halogen bulb.

Q8. Why is iodine sometimes added to tungsten lamps?

Hint: The addition of a small amount of iodine to the lamp extends the life span of the lamp and allows for the operation of the lamp at higher voltages. With iodine present, the sublimed tungsten atoms react to form a gaseous W-I compound; when the W-I compound strikes the hot filament, it decomposes and the tungsten atoms are deposited back onto the filament.

Q9. What is the wavelength range of a typical UV-vis spectrophotometer? Briefly discuss the physical constraints that limit a typical UV-vis spectrometer to this wavelength range.

Hint: For spectroscopic purposes, the ultraviolet spectrum begins at about 195 nm and ends at 400 nm. Although the region extends further, below 195 nm, the air (oxygen, nitrogen, and moisture) begins to absorb the radiant light and therefore background interference becomes a limiting factor for a typical UV-vis spectrometer.

Q10. To what does the term *n-electrons* refer? What is the spectroscopic significance of *n-electrons*?

Hint: A spectroscopist will refer to lone pair electrons as *n-electrons*. The "n" stands for **n**on-bonding. If a lone pair of electrons is conjugated with a π-system, the existence of the n-electrons produces a low energy n → π* transition.

Q11. Why are the peaks in molecular UV-vis spectroscopy broad relative to atomic UV-vis spectroscopy?

Hint: The electronic transitions in molecular UV-vis spectroscopy have superimposed vibrational transitions (vibronic) whereas the transitions in atomic UV-vis spectroscopy are purely electronic in nature.

Q12. What is the percent transmittance if the absorbance is:

(a) 0.5 (b) 1.0 (c) 2.5 (d) 0.8

Hint: (a) %T= 31.6%

(b) %T= 10%

(c) %T=0.316%

(d) %T=

$$A = -\log(\mathrm{T})$$
$$\therefore \mathrm{T} = 10^{-\mathrm{A}}$$

Q13. The 3s →3p transition in sodium is observe at about 5893 Å. What is this wavelength in nanometers? What is the frequency of the absorbed photon? What is the energy of the absorbed photon?

$$\because c = \lambda\nu \therefore \nu = c/\lambda$$
$$\lambda = 589.3nm$$
$$hence,$$
$$\nu = \frac{2.998\times10^{8} m/s}{589.3\times10^{-9} m} = 5.0874\times10^{14} Hz$$

$$E = hc/\lambda$$
$$E = \frac{(6.626\times10^{-34} J/s)(2.998\times10^{8} m/s)}{589.3\times10^{-9} m} = 3.3709\times10^{-19} J$$

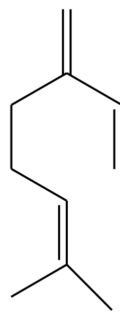
Woodward-Fieser Rules for Calculating λ_{max} in Conjugated Dienes and Trienes

Base value for acyclic or heteroannular diene	214nm
Base value for homoannular diene	253nm
Increment for each:	
Alkyl substituent or ring residue	5nm
Exocyclic conjugated double bond	5nm
Double bond extending conjugation	30nm
-OR (alkoxy)	6nm
-Cl,-Br	6nm

-OCOR (acyloxy)	0nm
-SR (alkylthio)	30nm
-NR2 (dialkylamino)	60nm
In the same double bond is exocyclic to two rings	
Simultaneously	10nm
Solvent correction	0nm
Calculated* λ_{max} of the compound	Total = nm

For π-π transition (K-band).

Q14: Calculate the wavelength of the maximum UV absorption for **Myrcene**

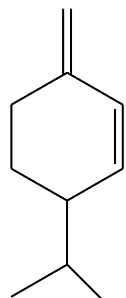


Myrcene

Since, it is an acyclic diene with one alkyl substituent, thus

Base value for heteroannular diene		214nm
One alkyl substituent		5nm
Calculated	**λmax**	**219nm**
Observed	**λmax**	**224nm**

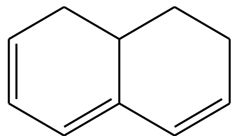
Q 15: Calculate the value of λ_{max} for **β-phellandrene**



This is a heteroannular diene with conjugated double bonds are not in the same ring with two ring residues and one exocyclic double bond, hence

Base value for heteroannular diene	214nm
Two alkyl substituents (2 x 5nm)	10nm
One exocyclic double bond	5nm
Predicted λ_{max}	**229nm**
Observed λ_{max}	**232nm**

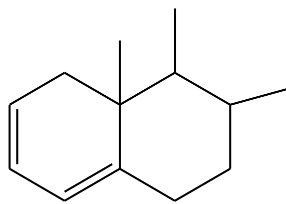
Q16: Expected λ_{max} values of compound



This is a homoannular diene with conjugated double bonds are not in the same ring with two ring residues and one exocyclic double bond, hence

Base value for homoannular diene	253nm
Three alkyl substituents (3 x 5nm)	15nm
One exocyclic double bond	5nm
Predicted λ_{max}	**273nm**

Q17: Absorption maximum for the ethanolic solution of given compound



This is a homoannular diene with three ring residues and one exocyclic double bond, thus

Base value for homoannular diene	253nm
Three ring residues (3 x 5)	15nm
One exocyclic double bond	5nm
Calculated λmax	**273nm**
Observed λmax	**275nm**

Q18: Calculate λmax for the ethanolic solution of given compound

This is a heteroannular diene with three ring residues and one exocyclic double bond, thus

Base value for v	214nm
Three ring residues (3 x 5)	15nm
One exocyclic double bond	5nm
Calculated λmax	**234nm**
Observed λmax	**235nm**

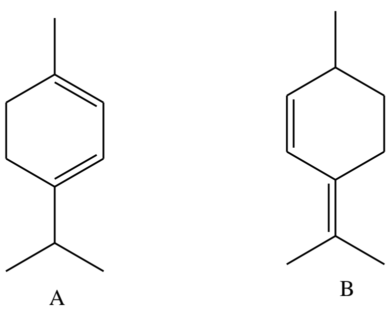
Q19: Calculate the value of absorption maximum for

It contains both homoannular and heteroannular diene systems but the calculation of its will be based on the homoannular diene system. There are six ring residues attached to the carbon atoms of the entire conjugated system, one double bond extending conjugation, two exocyclic double bonds and one double bond exocyclic to two rings simultaneously. Thus, Amax of compound is calculated as:

Base value for homoannular	253nm
Six ring residues (6 x 5)	30nm
One double bond extended conjugation	30nm

Two exocyclic double bond (2 x5)	10nm
One double bond expcyclic to two ring	
Simultaneously	10
Calculated λ_{max}	333nm

Question 7 Calculate λ_{max} for the compound A and B



Solution:

For compound A

Homoannular diene

Base value	253nm
Two ring residues (2 x 5)	10nm
Two alkyl substituents	10nm
Calculated λ_{max}	273nm

For compound B

Hetertroannular diene

Base value	214nm
Two ring residues (2 x 5)	10nm
Two alkyl substituents	10nm
One exo cyclic bond	5nm
Calculated λ_{max}	239nm

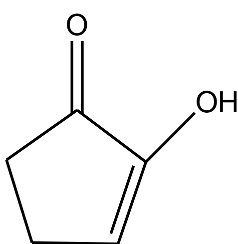
Question 8- Predict the value of λ_{max} (hexane) for

Hint:

This is a six-membered cyclic α,ß-unsaturated ketone with one *α*- and one *ß* alkyl substituents. Hence,

Base value	215 nm
One α-alkyl substituent	10nm
One β-alkyl substituent	12nm
Predicted λmax (EtOH)	237nm
Observed λmax (EtOH)	**249nm**
Calculated λmax (hexane)	**234nm**
Solvent correction	+11nm
	248 nm

Q20: Calculate λmax (EtOH) for



Hint:

1t is a five-membered cyclic α,β-unsaturated ketone with one α-hydroxy and one β-ring residue. Thus,

Base value	202nm
One α-hydroxy groups	35nm
One β-ring residue	12nm
Calculated Amax (EtOH)	**249nm**
Observed Amax (EtOH)	**247nm**

7.3 Problem based on Spectroscopy (IR)

Origin of Infrared Spectra

IR absorption spectra originate from transitions in vibrational and rotational energy levels within a molecule. On absorption of IR radiation, vibrational and rotational energies of the molecule are increased. When a molecule absorbs IR radiation below 100 cm^{-1}, the absorbed radiation causes transitions in its rotational energy levels. Since these energy levels are quantized, a molecular rotational spectrum consists of discrete lines.

When a molecule absorbs IR radiation in the range 100-10,000 cm^{-1}, the absorbed radiation causes transitions in its vibrational energy Ievels. These energy levels are also quantized, but vibratonal spectra appear as bands rather than discrete lines. Thus, a single transition in vibrational energy levels is accompanied by a large number of transitions in rotational energy Ievels and so the vibrational spectra appear as vibrational-rotational bands instead of discrete lines. Organic chemists are mainly concerned with these vibrational-rotational bands, especially with those occurring in the region 4000-667 cm^{-1}.

Atoms in a molecule are not still but they vibrate. The two types (modes) of fundamental molecular vibrations known are: (a) stretching and (b) bending vibrations.

Stretching Vibrations

Fingerprint Region

It is not possible for any two different compounds to have exactly the same IR spectrum(except enantiomers). Therefore, the IR spectrum of a compound is called its *fingerprint.* The region below 1500 cm^{-1} is called fingerprint region because every compound has unique absorption pattern in this region, just as every person has unique fingerprints. The fingerprint region contains many absorption bands caused by bending vibrations as well as absorption bands caused by C-C, C-O like in alcohols, ethers, esters, etc. and C-N (e.g. in amines, amino acids, amides, etc.) Stretching vibrations.

Since the number of bending vibrations in a molecule is much greater than its Stretching vibrations, the fingerprint region is rich in absorption bands and shoulders. Thus, the superimposability of IR bands of the spectra of any two

different compounds becomes impossible in this region. However, similar compounds may show very similar spectra above 1500 cm^{-1}.

Some characteristic group frequencies along with effects of structural environments on them are discussed as follows:

Type	Group	Absorption frequency (cm^{-1})	Intensity*	Assignment and remarks
Alkanes	$—CH_3$	2840-3000	m→ s	C-H stretch; two or three bands
	$>CH_2$ $—CH<$			
	$>C—C<$	800-1200	w	C-C stretch; little value
Cycloalkanes	$—CH_2—$	2840-3950	m	Asym. And Sym. C-H stretch, two bands
	$—CH_2—$ (cyclopropane)	3040-3060	m	Asym. C-H stretch
		2975-2985	m	Sym. C-H stretch
		1015-1045	m	Skeletal vibration
Alkenes	$>C=C<$	1620-1680	v	C=C stretch; diene,tiene.etc. 1650(s) and 1600(s)
	HC=CH cis	1655-1660	m	

Class	Structure	Frequency	Intensity	Assignment
	$HC{=}CH$ (trans)	1670-1675	w	
	$C{=}CH$	3000-3100	m	C-H stretch; almost the same position in the cis and trans isomer
Alkynes	$-C{\equiv}C-$	2100-2260	v	$C{\equiv}C$ stretch
	$-C{\equiv}C-H$	2100-2140	s	$C{\equiv}C$ stretch
Halogen Compounds	H_3C-X X=F,Cl,Br,I	near 3000	s	Asym. And sum. C-H
	$C-F$	1000-1400	s	C-F stretch
	$C-Cl$	600-800	s	C-Cl stretch
	$C-Br$	500-750	s	C-Br stretch
	$C-I$	-500	s	C-I stretch
	H_3C-X	1441-3100	v	C-H Asym. bending
	X=F,Cl,Br,I	1255-1475	v	C-H Sym. bending
Aromatic	C-H	3000-3100	m	C-H stretch
	C-C	1600±5	v	C=C skeletal stretch
		1580±5	m	Sleletak stretch; present when ring is further conjugated

Alcohols and Phenols	O-H	3590-3650	v	Free O-H stretch
		3200-3600	v	Intermediate hydrogen bonded O-H stretch
		2500-3200	s	Intermediate hydrogen bonded O-H stretch
	C-O	1000-1200	m→ s	C-O stretch
		-1050	s	C-O stretch pri. alcohol
		-1100	s	C-O stretch sec. alcohol
		- 1150	s	C-O stretch Ter. Alcohol
		-1200	s	C-O stretch phenols
	O-H	1339-1420	s	In-plane O-H stretch
		650-769	s	Out-of-plane O-H bending
Aldehydes, aliphatic		1720-1740	s	C=O stretch;saturated
	R—C═O H	1695-1715	s	C=O stretch; aromatic
		1680-1705	s	C=O stretch;α,β-unsaturated aliphatic
Ketones	R R C═O	1705-1725	s	C=O stretch; saturated, acyclic
		1680-1700	s	C=O stretch; aryl
		1660-1670	s	C=O stretch; diaryl
Carboxylic Acid	-COOH	1700-1725	s	C=O stretch;saturated aliphatic
		1690-1715	s	C=O stretch; α,β-unsaturated aliphatic

		1680-1700	s	C=O stretch; aryl
Carboxylate $–COO^-$		1550-1650	s	C=O asym. stretch
Ions		1300-1400	s	C=O sym. Stretch
Esters	$R-COOR'$	1735-1750	s	C=O stretch; saturated acyclic
		1650	s	C=O stretch;β-ketoester enolic
Acid halides	-COCl	1790-1815	s	C=O stretch;fluorides higher,bromides and iodine repectively lower
Acid Anhydrides	-CO-O-CO-	1800-1850	s	C=O stretch; two bands
		1740-1780	s	
Amides	$R-CONH_2$	-1690	s	Pri. amide, C=O stretch
	R-CONHR	1670-700	s	Sec. amide, C=O stretch
	$R-CONR_2$	1630-1670	s	Ter. amide, C=O stretch
Nitro Compounds And Nitriles	$R-NO_2$ (aliphatic)	1550-1570	s	Asym., N=O stretch
		1370-1380	s	Sym., N=O stretch
	$Ar-NO_2$ (aromatic)	1500-1570	s	Asym., N=O stretch
		1300-1370	s	ym., N=O stretch
Nitriles And Related Compounds	$R-C\equiv N$	200-2260	v	$C\equiv N$ stretch
	$R-N\equiv C$ (isonitriles)	2070-2220	m	$N\equiv C$ stretch
	$R-S-C\equiv N$ (thiocyanates)	2140-2175	s	$C\equiv N$ stretch

Sulphur	-SH	2550-2600	w	S-H stretch;less affected by H-bonding
Compounds	>C=S	1050-2100	s	C=S stretch
Phosphorus	P-H	2350-2440	s	P-H stretch
Compounds	P-O-R	1030-1240	s	P-O-C stretch

Question **1. Ethane (C_2H_6) molecule**

Non-linear molecule ethane (C_2H_6), the vibrational degrees of freedom can be calculated as:

Number of atoms (n) = 8

Total degrees of freedom (3n) = 3 x 8 = 24

Rotational degrees of freedom = 3

Translational degrees of freedom = 3

Hence, vibrational degrees of freedom = 24- 3- 3 = 18

Question **2. Benzene (C_6H_6) molecule**

Benzene (C_6H_6), the number of vibrational degrees of freedom can be calculated as follows:

Number of atoms (n) = 12

Total degrees of freedom (3n) = 3 x 12 = 36

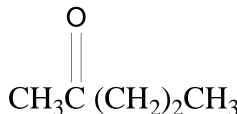
Rotational degrees of freedom = 3

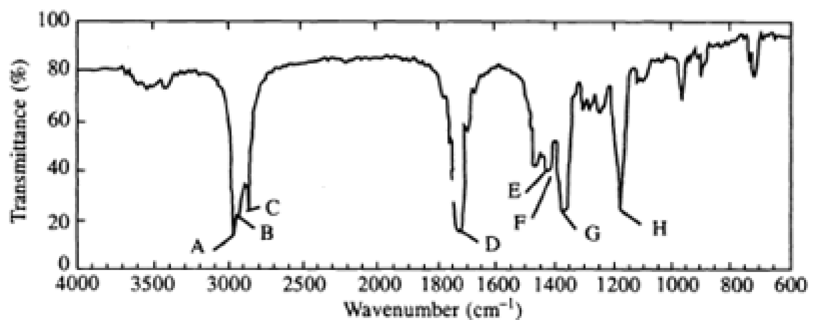
Translational degrees of freedom = 3

Therefore, vibrational degrees of freedom = 36 - 3 - 3 = 30

Thus, theoretically, there should be 30 fundamental vibrational bands in the IR spectrum of benzene.

Question **3. Interpretation of IR-spectra of 2-pantanone**





A: v_{as}, methyl, 2955 cm^{-1} E: δ_{as}, CH_3, -1430 cm^{-1}

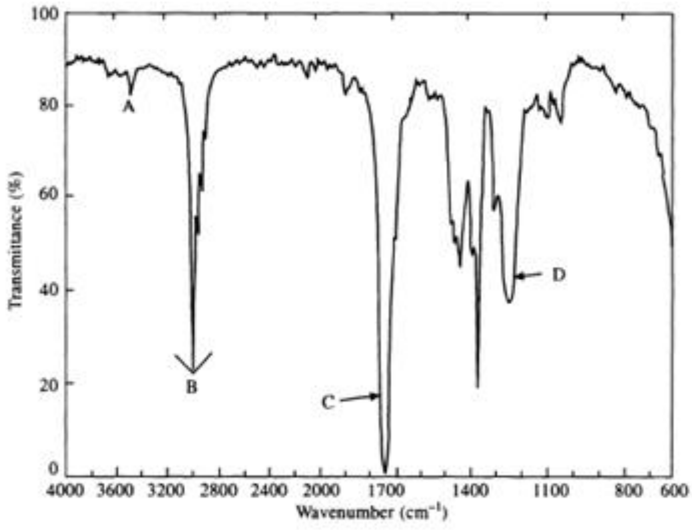
B: v_{as}, methylene, 2930 cm^{-1} F: δ_{s}, CH_2, -1430 cm^{-1}

C: v_{s}, methyl, 2866 cm^{-1} G: δ_{s}, CH_3 of CH_3CO unit, -1370 cm^{-1}

D: Normal C=O Stretching, 1725 cm^{-1} H: C-CO-C Stretching and bending, 1172 cm^{-1}

Question **4. Interpretation of spectrum of ethyl acetate**

$CH_3COO\ CH_2CH_3$



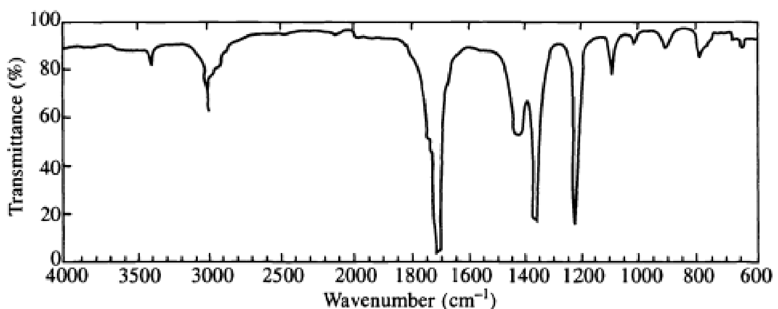
A: Overtone band of C=O Stretching, frequency twice that of C=O stretching, 3478 cm^{-1}

B: Methyl and methylene stretching bands, around 2900 cm^{-1}

C: Normal ester C=O Stretching, 1740 cm^{-1}

D: C-O-C Stretching, 1259 cm^{-1}

Question 5. IR spectrum of molecular formula C_2H_6O. Assign its structure.



7.4 Problem based on Spectroscopy (PMR)

NUMBER OF SIGNALS

The number of signals in an NMR spectrum tells the number of sets of **equivalent** protons in a molecule. Each signal corresponds to a set of equivalent protons. **Magnetically equivalent protons are chemically equivalent**, e.g.

Acetone **Cyclobutane** **Benzene**

CH_3 CH_2 — CH_2

C = O

CH_3 CH_2 — CH_2

In acetone, all six protons have exactly similar environment, therefore, only one signal is observed. Similar is the case in cyclobutane and benzene where all the protons being identical give only one signal.

(d) Compounds showing more than one signal are :
Methanol, methyl acetate, p-xylene, etc
CH_3

O

CH_3 - CH_2-OH O CH_3

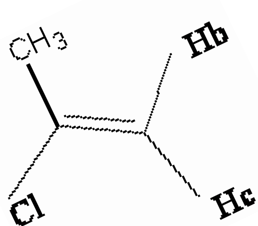
Here, in ethanol, 3 types of protons are there so three signals and in methyl acetate and p-xylene only two signals will be obtained.

(e) Chemically equivalent protons must also be stereo-chemically equivalent, i.e., a particular set of protons are said to be Chemically equivalent only if they remain in exactly similar environment when the streo-chemical formula of the molecule under consideration is written.

Considering the case of 2-chloropropene, CH_3-C(Cl)=CH_2

Cl

One can expect two sets of equivalent protons (Two peaks), but stereo chemical formula reveals three sets of protons in it.



Here, H^a and H^c are not in exactly similar environment.

Considering methyl cyclopropane, one can expect three sets of protons. But actually four signals are observed as is clear from the stereochemical formula.

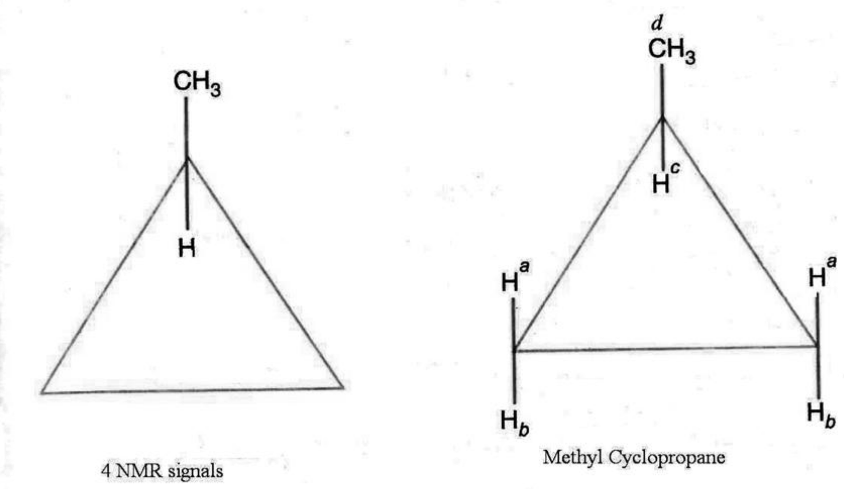
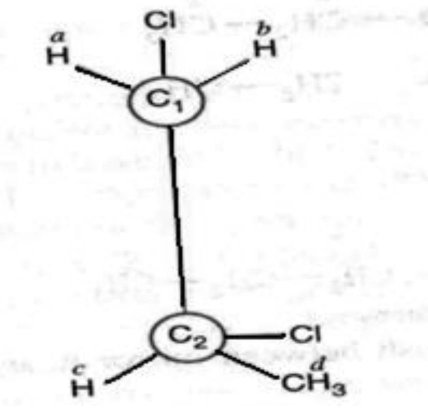


Fig (1.4) Number of signals in Methyl Cyclopropane

In 1,2-dichloropropane, CH_3 CH(Cl)-CH_2Cl one should expect three signals, but actually four signals are observed because two hydrogen atoms attached with C1 are not in exactly similar environment. Rotation around C-C single bond in this molecule cannot bring similar environment for the said hydrogen atoms.



2,3-dichloropropane

When a molecule is placed in a magnetic field, its electrons are caused to circulate and thus, they produce secondary magnetic fields i.e., induced magnetic fields. Rotation of electrons about the protons itself generates a field in such a way that at the proton, it opposes the applied field. Thus, the field felt by the proton, diminishes and the proton is said to be ***shielded***. Rotation of electrons (especially π) about the nearby nuclei generates a field that can either oppose or reinforce the applied field at the proton. If the induced field opposes the applied field, then proton is said to be shielded, but if the induced field reinforces the applied field, the proton feels higher field strength and such a proton is said to be ***deshielded***. Shielding shifts the absorption upfield and deshielding shifts the absorption downfield to get effective field strength necessary for absorption.

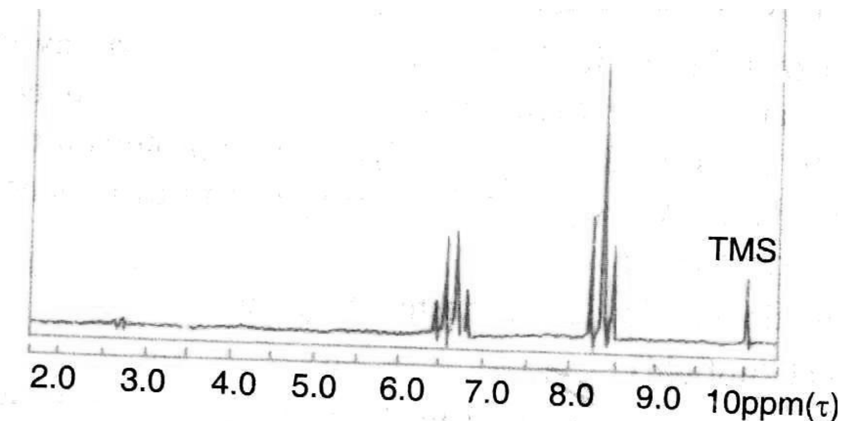
Such shifts (compared to a standard reference) in the positions of NMR absorptions which arise due to the shielding or deshielding of protons by the electrons are called ***Chemical shifts***. For measuring various shifts in a molecule, the signal for tetra methyl silane (TMS) is taken as a reference. Due to the low electro negativity of Silicon, the shielding of equivalent protons in TMS is greater than most of the organic molecules.

Therefore, NMR signal for tetramethyl silane is taken as a reference and chemical shift for different kinds of protons are measured relative to it. Clearly, the NMR signal for a particular proton in a molecule will appear with different field strengths compared to a signal from TMS.

Self Assessment Questions

Question **1. Molecule of CH_3CH_2Br, ethyl bromide:**

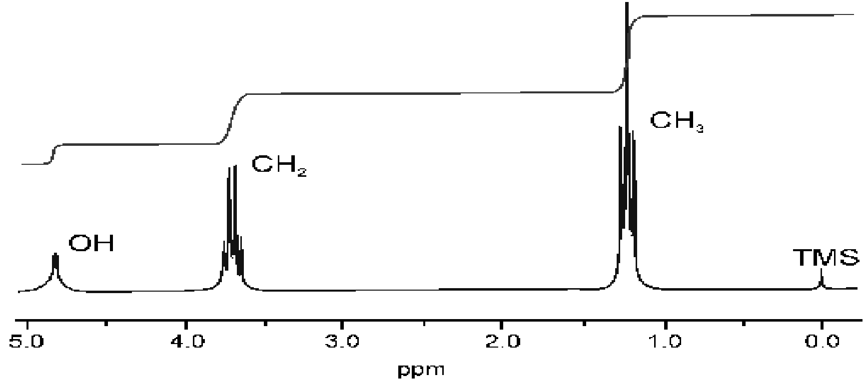
This molecule has two kinds of protons in it and thus, two signals are expected in its NMR spectrum. It has been observed that for each kind of protons, we do not get singlets but a group of peaks are observed. For 'CH_3' protons, a triplet i.e. a group of three peaks is observed and a quartet (group of four peaks) is noticed for protons -CH_2.



NMR spectrum of Ethyl bromide. Signals and their absorption positions

Question **-2:** Spectrum of Ethanol

CH_3CH_2OH



The three main features of NMR spectroscopy ,i.e., chemical shift, signal intensities and spin- spin coupling can be studied in this spectrum.

Three – proton atmosphere, all three protons lying in different magnetic atmosphere, due to motion of valence electrons and neighbouring atoms in response to magnetic field.

(iv) CH_3 – (three protons)

(v) – CH_2 – (two protons)

(vi) - OH (one proton)

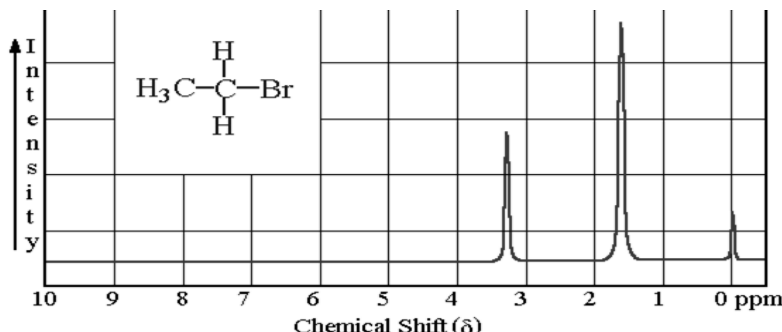
The effects arising due to the motions of the electrons which will be different for each kind of hydrogen will be at different field strengths.

A plot of signals' against field strengths thus shows three principle groups of lines for ethyl alcohol.

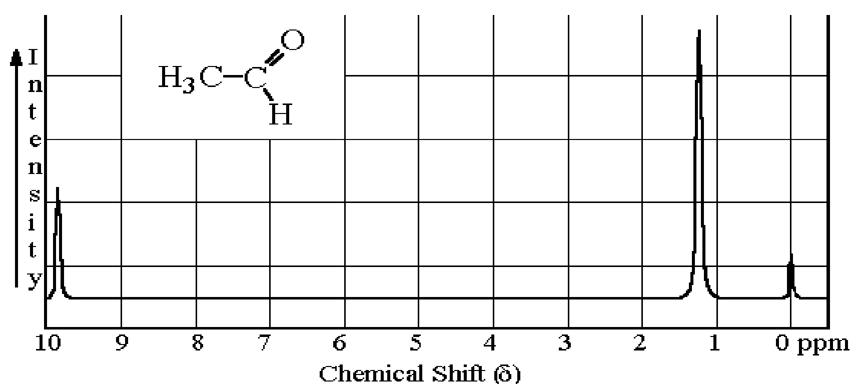
The applied signal is a three proton triplet of methyl group. The chemical shift of this signal is (1.22) for all the three identical methyl group protons have two neighbouring protons i.e. (n = 2) and according to (n+1) rule one sees a three proton triplet, since three identical hydrogen of methyl group couple equally with the protons of the methylene group. Similarly, the two identical protons of methylene group are adjacent to the hydroxyl group and as a consequence come into resonance at lower field than the methyl group. The signal is a quartet since the methylene protons are equally coupled to three hydrogens of CH_3.

The hydroxyl proton appears as a singlet, i.e, no interaction is seen between the hydroxyl proton and neighbouring methylene group. It may be sufficient to remember that the spin-spin splitting is not observed normally for a proton of a hydroxyl or amino group. Such protons are relatively acidic and a s a result these rapidly interchange between molecules.

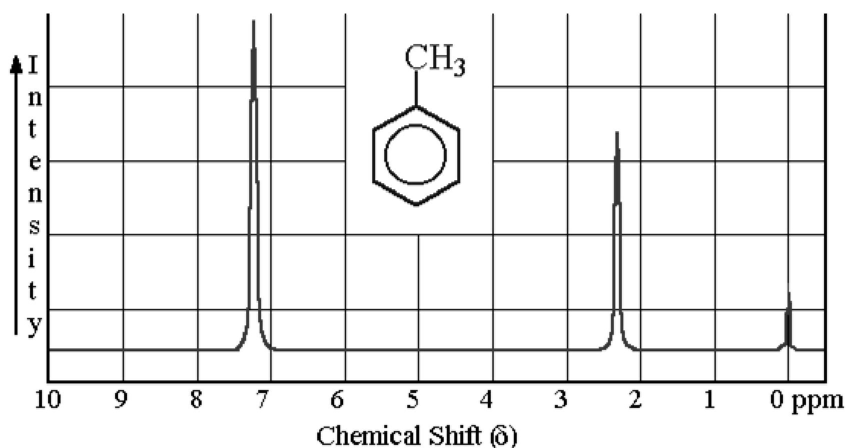
Question – **3: Spectrum of Ethyl Bromide**



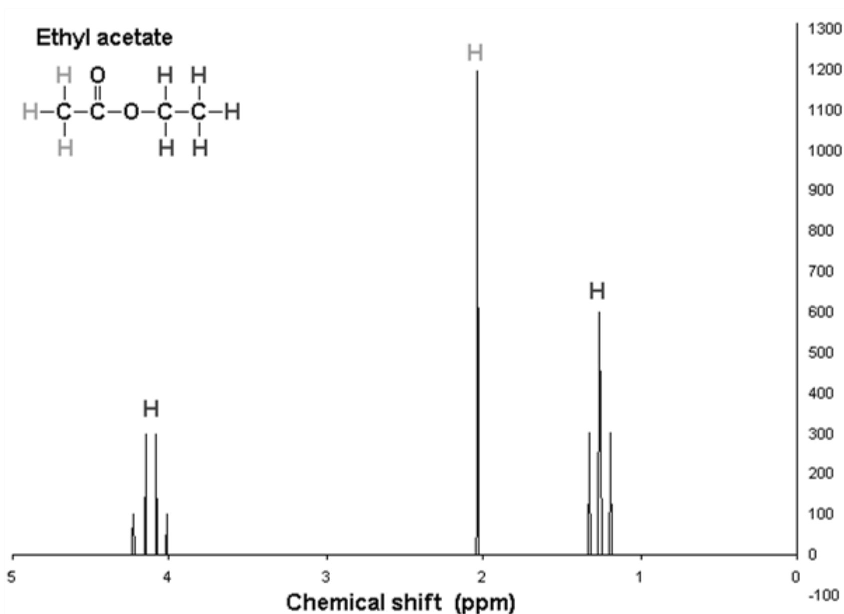
Question – **3: <u>Spectrum of Acetaldehyde</u>**



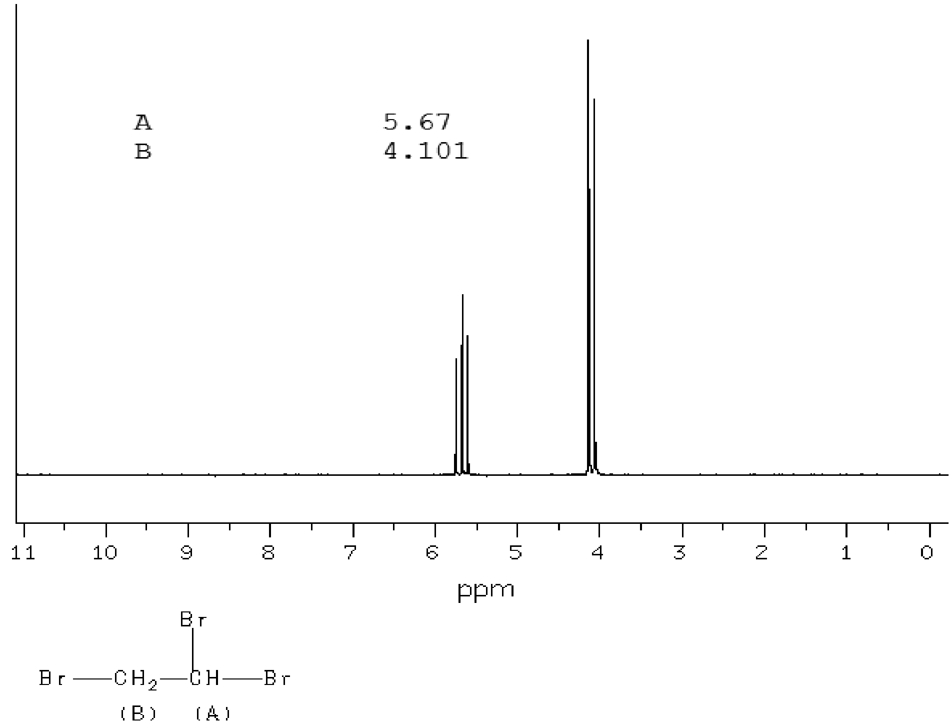
Question – **4: <u>Spectrum of Toluene</u>**



Question **- 5: <u>Spectrum of Ethyl Acetate</u>**



Question **- 6: <u>Spectrum of 1,1,2-tri bromo ethane</u>**



Question **-7**. A compound with molecular mass 158 absorbs in the ultraviolet region at 225 nm ε_{max} (hexane). In infra-red spectrum, absorption bands are formed at 3077-2857 cm^{-1}(m), 1828 cm^{-1}(s), 1757 cm-1 (m) and 1457 cm^{-1} (m).

In NMR, two signals are observed (i) 7.30τ septet (J=6.7 cps, 6.4squares) and (ii) 8.80 τ doublet (J = 6.7 cps, 37.2 squares).

Question **8.** A compound molecular mass 164 absorbs at 220 nm ε_{max} 1800. In infra-redspectrum, absorption bands are formed at 3077 cm^{-1} (w), 2976 cm^{-1} (s) 1608 cm^{-1} (m), 1497 cm^{-1} (m) and 1456 cm^{-1}(m).

In NMR, the signals formed are (i) 2.7 τ singlet (16.5 squares), (ii) 5.70 τ triplet (J = 7.3 cps, 6.2 squares), (iii) 7.07 triplet (J = 7.3 cps, 6.7 squares) and (iv) 7.98 τ singlet (10.2 squares).

Question **9.** An organic compound with molecular mass 174 shows absorption in ultraviolet region at 213nm ε_{max} 60. In infra-red, absorption bands are formed at 2941- 2857 cm^{-1} (m), 1745 cm^{-1} (s) and 1458 cm^{-1} (m).

In NMR, the signals observed are (i) 5.86 τ quartet (J = 7.2cps, 10.4 squares), (ii) 7.40 τ singlet (10.8 squares) and (iii) 8.73 τ triplet (J = 7.2 cps, 16.0 squares).

Question **10.** A organic compound with molecular formula $C_{14}H_{19}N$ give the following spectral data:

UV : (i) 222 nm ε_{max} 20,400

(ii) 210 nm ε_{max} 20,000

IR : 3022 (m), 1600 (m), 1510 (m), 1680 (w), 750 and 695 cm^{-1}(m)

NMR : (i) singlet 7.15 τ (3H), (ii) singlet 8.5 τ (3H), (iii) complicated 8.0 τ (4H), (iv) complicated 8.35 τ (4H), multiplet; 3.45 τ (3H) and multiplet 2.95 τ (2H)

Determine the structural formula of the compound.

Question **11: A**n organic compound with molecular formula C_4H_9NO gives the following spectral data:

UV : λ_{max} 220 mμ ε_{max} 63

IR : 3500 (m), 3402 (m), 2960 (w), 1682 (s), 1610 (s)

NMR : 9.0 τ doublet (23.2 squares); 7.9 τ septet (3.8 squares); 1.92 τ singlet (7.5 squares).

Determine the structural formula of the compound.

Question **12**. An organic compound contains 66.6% Carbon, 11.1% Hydrogen. In UV, it gave a characteristic band at 275 mμ ε_{max} 17. In infra-red, bands are formed at 2941-2857 (m), 1715(s) and 1460 cm^{-1} (m). In NMR, three signals appear at (i) 7.52 τ quartet, (2H), 7.88 singlet, (3H) and 8.93 τ triplet, (3H). Determine the structure formula of the compound.

SOLUTION-7. In the NMR spectrum, a septet and a doublet with the same value of coupling constant shows $(CH_3)_2CH$- as a part of the structure.

In the infra-red spectrum, a strong band at 1828 cm^{-1} and a medium band at 1757 cm^{-1} shows

The presence of anhydride —C(=O)—O—C(=O)—

Thus, $(CH_3)_2CH$- and –CO-O-CO- amounts to 115 mass units. The remaining 43 mass unit may also be the isopropyl group. Hence, the most probable structure of the compound is:

$(H_3C)_2CH—C(=O)—O—C(=O)—CH(CH_3)_2$

A band at 3077-2857 cm^{-1} is due to C-H str and that at 1456 cm^{-1} is due to C-C str. An absorption band at 225 nm ε_{max} 50 does not give any authentic information.

SOLUTION - 8. In NMR, the signals are formed in the proton ratio 2.5:1:1:1.5 or 5:2:2:3. The formation of two triplets with the same value of coupling constant shows that $–CH_2$-CH_2- is a part of the structure.

A five proton singlet at 2.71 τ shows the presence of an aromatic ring C_6H_5- .

A three proton upfield singlet shows $–CH_3$ group in the compound. Thus, the units C_6H_5 - , -CH_2-CH_2- and CH_3- amount to 120 mass unit. A strong band at 1745 cm^{-1} in the infra-red spectrum suggests that the compound is an ester. Thus, -COO (of ester) amounts to the remaining 44 mass unit of the compound

of molecular mass 164. The bands at 1608 cm^{-1}(m) and 1497 cm^{-1} (m) support the presence of benzene ring. The band at 1456 cm^{-1} (m) is due to C-C str. Thus, the probable structure of the compound may be

O

H_2C — C — C — OCH_3

H_2

The structure II looks more probable because three proton singlets are relatively less deshielded.

SOLUTION - 9. In the NMR, the signals are formed in the proton ratio 2:2:3. The formation of a quartet and a triplet with the same value of coupling constant suggests that CH_3CH_2- is a part of the structure. Also a two proton singlet is formed. It may be due to –CH_2- group in the structure.

In infra-red spectrum, the formation of a strong band at 1745 cm^{-1} shows the presence of an ester. Thus, -COO is a part of the structure. Thus, CH_3CH_2-, -CH_2- and –COO units amount to 87 mass unit. The molecular mass 174 is just double of it. Thus, the probable structure of the compound is:

O

H_2C — C — O — CH_2CH_3

H_2C — C — O — CH_2CH_3

O

A low field triplet indicates that its carbon is directly linked with oxygen atom.

An absorption band at 2941-2857 cm^{-1} indicates C-H str in infra-red and that at 1458 cm^{-1} indicates C-C str.

SOLUTION - 10. From the absorption in the ultraviolet spectrum, nothing can be said with certainty.

The infrared absorption at 3022 cm^{-1} includes =C-H stretching from an aromatic ring. The appearance of bands at 1600 (m), 1510 cm^{-1} (m) indicates the presence of an aromatic ring. Further the bands at 750 and 695 cm^{-1}(m) shows that the aromatic ring must be substituted. The weak band at 1680 cm^{-1} is characteristic of C=C stretching. The presence of the aromatic is further confirmed by multiplets in the region 2.95 – 3.35 τ. The two proton multiplet at 2.95 τ must be due to two ortho positions with respect to the N-substituted group on the ring. The remaining three protons show another slightly less deshielded multiplet at 3.45 τ.

The two signals each of three protons singlet at 7.15 τ and at 8.5 τ indicate two methyl groups. The singlet at 7.15 τ is slightly deshielded being attached with nitrogen which in turn is attached with the aromatic ring.

The remaining two signals (complicated) corresponding to 8 protons in almost similar environments indicate 4 methylene groups in ring A. the four proton signals at 8.0 τ corresponds to 2 methylene groups which are attached with the carbon atoms carrying a a double bond.

From this, the structural formula of the compound is probably.

SOLUTION - 11. From the absorption at 220mμ ε_{max} 60 in the ultraviolet region, nothing can be said with certainty. The presence of a primary amide is indicated by two bands at 3500 and 3402 cm^{-1}. These two bands are the result of N-H stretching. A strong band at 1682 cm^{-1} also indicates the presence of an amide group. The band at 2960 cm^{-1} is due to C-H stretching and that at 1610

cm^{-1} (s) is due to N-H bending. From this much data, we say that $-CONH_2$ is a part of the structure.

The NMR spectrum is most informative. The doublet at 9.0 τ and a septet at 7.9 τ in the ratio 6 : 1 indicates clearly an isopropyl group i.e., $(CH_3)_2$ CH- group.

Another two proton singlet appears at a very low field due to much deshielding. This signal is due to two protons of the primary amide group. The appearance of the signals at a very low field may be explained by writing the resonating structure of amide as:

$$-\overset{\overset{\displaystyle O}{\|}}{C}-NH_2 \longleftrightarrow -\overset{\overset{\displaystyle O^-}{|}}{C}=NH_2$$

I II

The double bond character between C and N and also the positive charge developed on nitrogen atom in structure II case greater deshielding of protons attached with nitrogen atom and hence, signal appears at 1.92 τ.

Thus, the structure of the compound is

$$(H_3C)_2CH-\overset{\overset{\displaystyle O}{\|}}{C}-NH_2$$

From the given molecular formula (C_4H_9NO), the number of ring and double bond equivalents are calculated as:

$$4+1-\frac{9-1}{2} = 5 - 4 = 1$$

This is in accordance with the structure which is determined.

SOLUTION - 12 The compound contains -

C = 66.6%

H = 11.1%

O = 100-(66.6+11.1) = 22.3%

From this data, the empirical formula of the compound is found to be C_4H_8O. This must be the molecular formula since eight hydrogen atoms are shown by NMR spectrum.

(iv) The absorption at 275 mμ ε_{max} 17 is characteristic of a carbonyl group due to n π* transition (forbidden band).

(v) The absorption at 2941-2857cm^{-1} (m) in the IR spectrum is due to C-H stretching, at 1715cm^{-1} (s) is characteristic of saturated ketonic group and that at 1460 cm^{-1} (m) may be due to C—H bending.

(vi) The NMR spectrum reveals three kinds of protons.

The presence of a triplet at 8.93 τ and a quartet at 7.52 τ is characteristic of $CH_3 - CH_2 -$ group in the compound. The singlet at 7.88 τ is due to methyl group adjacent to a carbonyl group. Hence, the probable structure of the compound is $CH_3 - CH_2 - CO - CH_3$. The structure contains one double bond. Also, the number of DBE from the molecular formula:

$C_4H_8O = 4 + 1 - 8/2 = 1$

7.5 Problem based on Spectroscopy (UV-Vis., IR and PMR)

1. The chemical shifts of a doublet signal for a proton in a spectrum are 4.08 and 4.06 using a 400 MHz NMR spectrometer. The coupling constant (in Hz) is:

(a) 0.02 (b) 8.0

(c) 8.14 (d) 10.0

2. The two characteristic stretching frequencies (cm^{-1}) observed in the IR spectrum of compounds containing NO_2 group is:

(a) 3400 and 3300 (b) 1860 and 1760

(c) 1550 and 1350 (d) 2250 and 1760

3. The UV-light source used in UV-visible spectrophotometer is:

(a) Mercury lamp (b) Tungsten lamp

(c) Deuterium lamp (d) Sodium lamp

4. The ^{1}H NMR spectrum of 1, 4-dimethoxybenzene will have

(a) ten singlets (b) two singlets

(c) two doublets and one singlet (d) two doublets and two singlets

(d)

5. The pattern of ^{1}H NMR spectrum of 1, 4-dichlorobenzene is:

(a) AX (b) AM

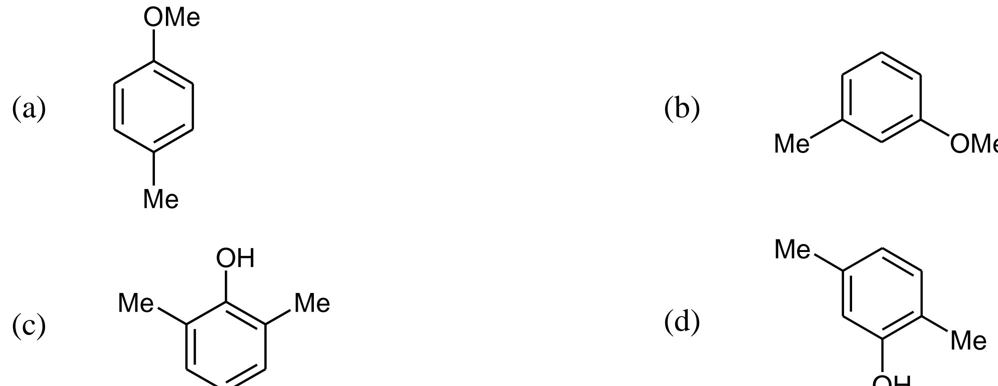
(c) AB (d) A_4

6. Which of the following molecules will not absorb infrared radiation?

(a) CO (b) Cl_2

(c) $COCl_2$ (d) CO_2

7. An organic compound with molecular formula $C_8H_{10}O$ exhibited 6 peaks in its broad-band decoupled ^{13}C NMR spectrum. The possible structure of the compound is:



8. The presence of hydrogen bonding in an organic compound can easily be established utilizing the following technique:

(a) IR spectroscopy (b) Mass spectra

(c) Cyclic voltametry (d) CD-ORD

9. Which of the following will most conveniently confirm if a known solid sample is impure?

(a) NMR (b) Mass spectrum

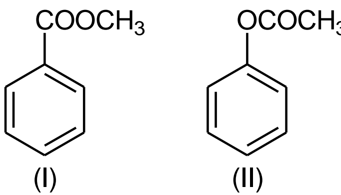
(c) IR spectrum (d) Melting point

10. The molecule azulene has an absorption maximum at 700 nm, the red end of the visible spectrum. The next shortest wavelength occurs at 357 nm. The predicted colour that azulene can exhibit is:

(a) blue (b) red

(c) indigo (d) green

11. In the ^{1}H NMR spectra, signals due to methyl group in isomeric compounds I and II appear respectively at δ.



(a) 1.25 and 3.9 (b) 3.9 and 2.1

(c) 3.9 and 7.25 (d) 7.25 and 2.1

12. How many signals will be observed in the proton decoupled ^{13}C NMR for hexamethylbenzene?

(a) 1 (b) 2

(c) 3 (d) 4

13. The ^{1}H NMR spectrum of an organic compound of molecular formula C_4H_8 exhibited only a singlet at $\delta = 1.9$ ppm. The compound is:

(a) 1-butane (b) cis-2-butene

(c) trans-2-butene (d) cyclobutane

14. The number of IR active vibrational modes in ammonia is

(a) 6 (b) 4

(c) 2 (d) 3

15. The ^{1}H NMR spectrum of $(\eta^5\text{-}C_5H_5)_2$ Fe recorded at room temperature has

(a) one singlet (b) one multiplet

(c) two singlets (d) two multiplets

16. The infrared spectrum of CO_2 exhibits the following number of absorptions :

(a) One (b) Two

(c) Three (d) Four

17. Which of the following electronic transitions is disallowed?

(a) $\pi \rightarrow \pi$ (b) $\sigma \rightarrow \sigma$

(c) $n \rightarrow \pi$ (d) $\delta \rightarrow \delta$

18. In the UV spectrum of cyclohexanone, the absorption at λ_{max} ~215 nm is due to the transition

(a) $\sigma \rightarrow \sigma$ (b) $\sigma \rightarrow n$

(c) $\pi \rightarrow n$ (d) $\pi \rightarrow \pi$

7.6 Summary

In this unit learner learn about problem based on Spectroscopy (UV-Vis.), Problem based on Spectroscopy (IR), Problem based on Spectroscopy (PMR) and Problem based on Spectroscopy (UV-Vis., IR and PMR).

7.7 Terminal Questions

1. Absorption occurs at...

(a) All wavelengths in the spectrum

(b) A characteristic wavelength dependent on the molecule

(c) The UV region

(d) The visible region

2. Transmittance is...

(a) The amount of radiation absorbed by the sample

(b) The amount of radiation initially divided by the amount of radiation passing through a sample

(c) The amount of radiation passing through the sample divided by the initial amount

(d) The wavelength used that promotes an electron

3. The Beer-Lambert Law...

(a) Relates absorbance, concentration, path length and molar absorption coefficient

(b) Tells us the volume of the sample

(c) Relates frequency and wavelength

(d) Allows us to calculate how conjugated the system is

4. Conjugated systems tend to absorb in the visible region because...

(a) electrons are coloured

(b) overlapping pi orbitals increase the energy gap between orbitals

(c) overlapping pi orbitals reduce the energy gap between orbitals

(d) 100% transmittance occurs

5. UV-Visible spectrometer uses a prism to...

(a) Focus all wavelengths on the sample simultaneously

(b) Separate radiation into its constituent wavelengths

(c) Reduce the amount of radiation passing through the sample

(d) Stop any radiation going through the sample

6. UV-Vis. Spectroscopy of organic compounds is usually concerned with which electronic transition(s)?

(a) s to s*

(b) s to s*

(c) n to p* and p to p*

(d) n to s*