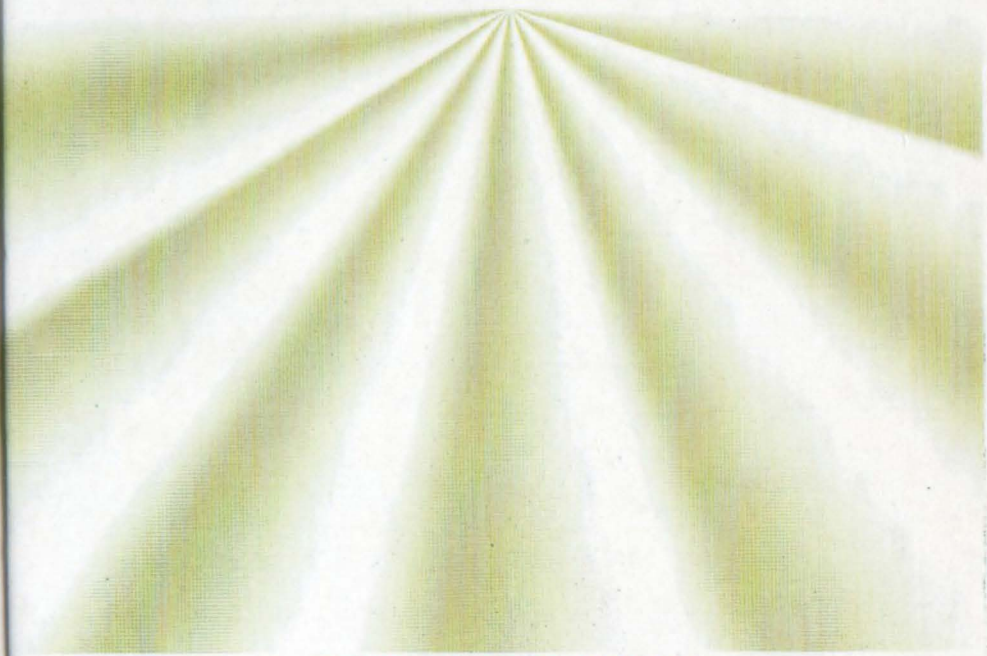
ADVANCED BIOPHYSICS



S.K. AGARWAL

BOOKS OF SIMILAR INTEREST

Air Pollution: *S.K. Agarwal*	Rs. 795/-
Eco-Informatics (in 5 Vols.): *S.K. Agarwal*	Rs. 3495/-
Vol. I: Environmental Management: New Concepts	
Vol. II: Green Management	
Vol. III: Wealth from Waste	
Vol. IV: Environmental Monitoring	
Vol. V: Non-Conventional Energy Systems	
Environmental Geography: *S.K. Agarwal*	Rs. 695/-
Environmental Monitoring: *S.K. Agarwal*	Rs. 795/-
Environmental Scenario for 21st Century: *S.K. Agarwal*	Rs. 995/-
Environmental Controversies: *S.K. Agarwal*	Rs. 800/-
Global Warming and Climate Change: Past, Present and Future: *S.K. Agarwal*	Rs. 895/-
Green Management: *S.K. Agarwal*	Rs. 795/-
Non-Conventional Energy Systems: *S.K. Agarwal*	Rs. 395/-
Noise Pollution: *S.K. Agarwal*	Rs. 395/-
Nuclear Energy: Principles, Practices & Prospects: *S.K. Agarwal*	Rs. 995/-
Pollution Management (in 5 Vols.): *S.K. Agarwal*	Rs. 3495/-
Vol. I: Air Polluiton	
Vol. II: Water Pollution	
Vol. III: Pesticide Pollution	
Vol. IV: Heavy Metal Pollution	
Vol. V: Noise Pollution	
Water Pollution: *S.K. Agarwal*	Rs. 895/-
Wealth from Waste: *S.K. Agarwal*	Rs. 895/-

ISBN 81-7648-987-5

APH PUBLISHING CORPORATION
5, Ansari Road, Darya Ganj, New Delhi-110 002
E-mail: aphbooks@vsnl.net

ADVANCED BIOPHYSICS

Dr. S.K. Agarwal
Principal
Modi Institute of Management & Technology
Modi Education Complex,
Dadabari Extn.,
Kota 324009

A P H PUBLISHING CORPORATION
5 ANSARI ROAD, DARYA GANJ
NEW DELHI-110 002

Published by
S.B. Nangia
A P H Publishing Corporation
5 Ansari Road, Darya Ganj
New Delhi-110002
☎ 23274050
Email : aphbooks@vsnl.net

ISBN 81-7648-987-5

2005

Typesetting at
Paragon Computers
B-36, Chanakya Place
New Delhi-110 059
☎ 25509417

Printed in India at
Efficient Offset Printers
New Delhi-110 035

PREFACE

Biophysics is a interdisciplinary science which involves the study of biological organisms at all levels from molecules and cells to the biosphere and ecosphere as a whole, from the perspective of the physical sciences. This is the field of science that explains how physical phenomena affect/account for the structure and function of biological systems. In the last two decades biophysics has made great advances in explaining and interpreting life processes.

The recognition of biophysics as a separate field is relatively recent, having been brought about by the inventions of physical tools, which greatly facilitate biological research. These tools are peculiarly adapted to the study of problems of great current importance to medicine, and problems related to disease.

The book discusses various aspects of biophysics, giving equal emphasis to each of the major aspect. The text has been written, as far as possible, in a easy, simple and understandable language, which is self contained, fully derived and critically discussed, it is supported by a large number of diagrams, which the author felt necessary.

All chapters have been carefully updated to account for the advances in biophysics. Chapter first describes Bioenergetics, followed by Photobiology, Molecular interactions. Sensory receptors, Photo-regulatory signal regulation. Biophysics of sonic vibrations, Membrane conductivity. X-ray imaging. X-ray diffraction imaging. Ultrasound imaging, Computerized axial tomography, Electrocardiography, Electroencephalography, Radiocarbon dating, DNA fingerprinting, and Chemical fingerprinting of plants, Amazing facts have been described in each chapter, which introduce the readers to the wonders of biophysics and biomolecular structures.

I am indebted to my students for their direct and indirect role in the preparation of this book, I am thankful to University Grants Commission, New Delhi for helping me to get acquainted with people working in the field of Biotechnology. Thanks are also due to Prof., R.R. Das, Jiwaji University, Gwalior, Prof. K. Dharmalingam, Madurai Kamraj University, Madurai; Prof. Gopinathan, Institute of Basic Medical Science, Madras; Prof. B.A. Ravishankar, CFIRI, Bangalore; Prof. P.S. Dubey, Vikram University, Ujjain; and Prof. L.N. Vyas, MLS University, Udaipur for their constant help in literature survey and critical discussions.

In the preparation of this book, I have relied heavily on many excellent treatises on different aspects, most of which are mentioned in the references. My attempt was to put them in unambiguous words and to find certain correlations which were not yet explored, I am greatly indebted to all the authors for their diagrams and tables that has been incorporated in this book.

It is hoped that this book would serve the needs of teachers as well as undergraduate and postgraduate students,

CONTENTS

Chapter 1

BIOENERGETICS

All life is based on energy. How we get it, and how we use it are very important. Bioenergetics is the study of energy transformations in biological systems. While this may appear to be a very dry (and boring) topic, a basic understanding of energy and its use in living organisms is invaluable. Living organisms must perform work to stay alive, grow, and reproduce. All living organisms must possess the ability to obtain energy and be able to transform that energy into a form that can be used by its cells. Practically speaking, knowing the fundamentals of bioenergetics aids in the understanding of cell functions and allows us to understand why and how the cells are able to harness energy.

The study of energy changes accompanying biochemical reactions is termed as bioenergetics or biochemical thermodynamics. The living organism is a highly organised arrangement of matter. The synthesis and maintenance of the system is possible only by the availability of certain reactions. These reactions include degradation of complex molecules of food material to simple molecules, and the synthesis of complex molecules from the simple compounds. In the living systems, heat is stored in the form of energy rich or high energy compounds which are mainly phosphates containing molecules. Whenever the energy is needed by metabolic processes, these high energy compounds undergo decomposition and the energy liberated is utilised. Further, the synthesis of these compounds take place as a result of various metabolic activities. Such spontaneous reactions produce useful energy for the performance of various vital activities of the

organisms. The raw materials are used by the organisms as forces of energy and as building blocks for its tissues.

زیندہ وزہ زانی کی بنچینڈی

BASIC BIOENERGETICS

Bioenergetics is the study of energy supply, utilization, and dissipation in living organisms. Energy is the capacity for performing work. Nutrients contain *chemical energy* which is yielded upon chemical breakdown and can be used in the body to perform chemical, mechanical, electrical, or osmotic work. The efficiency of conversion of chemical energy to work energy is less than 25 percent - the remaining 75 percent is converted to thermal energy.

The internationally agreed *unit of energy* is the Joule (J). J= $kg/(m^2S^2)$. The older unit, still used in the united States is the calorie (cal). 1 calorie = heat required to increase the temperature of 1 gram of water from 14.5 to 15.5 °C. 1 cal = 4.184 J.

زیندہ وزہ زانی کی خلیاتی

CELLULAR BIOENERGETICS;

Some properties of ATP are: ATP = ADP + Pi + 7400 cal. ATP is highly liable and is not stored in cells. It therefore needs continuous regeneration. ATP is synthesized mainly in mitochondria by the process of oxidative phosphorylation. For ATP synthesis the cells need (a) nutrients and their metabolites derived from carbohydrates (e.g., glucose), lipids (e.g., fatty acids), and protein (amino acids); (b) the initial steps are specific to the chemical nature of these components and can give a modest yield of ATP (e.g., glycolytic breakdown of glucose to pyruvate or its reversible intermediate (lactate) gives a net yield of 2 moles of ATP per mole of glucose; (c) much more ATP is generated via complete oxidation of carbohydrates, fatty acids or amino acids to Carbon dioxide and water. This involves two major associated processes within the mitochondria: (i) Tricarboxylic acid (TCA) cycle (also called Krebs cycle or Citric acid cycle), and (ii) Oxidative phosphorylation.

Oxidizable substances enter the TCA cycle after conversion to the two carbon compound, acetyl-coenzyme-A (acetyl.CoA). Their catabolism yields Carbon dioxide, a small amount of *ATP* and most important, reduced coenzyme (NADH, FADH2). NAD and PAD are derived from vitamin B, nicotinamide (niacin) and reboflavin respectively.

Most ATP is produced by the process of oxidative phosphorylation, in which hydrogen carried by reduced co-enzymes are finally oxidized to water after reaction with a series of enzymes and cofactors called the cytochrome chain (also called the respiratory or electron transport chain). The energy released by removal of hydrogen from reduced co-enzymes is used to drive the synthesis of ATP from ADP.

$$ADP + Pi + 7400 \text{ cal} = ATP.$$

ریښه وزه زانګی ی ته واوی-لاړښه

WHOLE-BODY BIOENERGETICS

Energy can be neither created nor destroyed. This applies to all living organisms of the universe. Energy can be interconverted between different forms, but thermal energy (heat) can not be converted to other forms in the body. All biochemical reactions involve net exchange of free energy (ΔG). The reactions which release free energy are termed *exergonic* where as the reactions which use free energy are termed *endergonic.* It is possible, with difficulty to measure ΔG for individual reactions within cells, but impossible to measure ΔG for multiple reactions, let alone all of the reactions which aggregate to represent whole-body energy exchange. Fortunately, for all common biochemical reactions, the heat of reaction (ΔH) = ΔG (e.g., for glucose oxidation ΔG= –688 kcal/mole; ΔH = –671 kcal/mole). This means that whole-body, organ or cellular free energy exchange can be estimated from heat production of the whole-body, organ or cell, which is much easier to measure than ΔG.

The sources of energy loss as arbitrary approximations for ruminant and non-ruminant animals in terms of gross energy (GE), *digestible energy* (DE), *metabolizable energy* (ME), and *net energy* (NE) (Table 1), show that the gross energy is the total energy content of a given weight of feed. The digestible energy is faecal energy. The metabolic energy is urine and/or gas energy. The net energy is heat increment of feeding. The net energy is used first to meet the animal's maintenance energy requirement. Any surplus which is deposited in the body tissues (growth), milk, eggs etc. is defined as *retained energy* (RE):

$$RE = ME - H$$

Table 1: Sources of energy loss in ruminants and non-ruminants.

Sources	*Ruminants*	*Non-ruminants*
Gross energy (GE)	100	100
Digestible energy (DE)	70	90
Metabolizable energy (ME)	<60	>85
Net energy (NE)	<40	>60

ENERGY AND ITS VARIOUS FORMS

Energy of a system may be defined as the capacity to do work. This capacity may be bound in the molecules. Energy exists in many forms, such as heat, light, chemical energy, and electrical energy. In living organisms the main source of energy is the *solar energy* which gets fixed up in the form of *chemical energy* of carbohydrate molecules during photosynthesis in green plants. Other organisms have to depend directly or indirectly upon this process (photosynthesis) for a constant supply of energy to maintain their structures and to perform their numerous functions. Energy is not only required for mechanical work, maintenance of body temperature and osmotic work but also to drive the numerous reactions. Various forms of energy are as follows:

(a) ***kinetic energy*** - the first type of energy is called kinetic energy (energy of motion). It includes light, heat, and the movement of molecules. It is the energy contained within a boundary by virtue of the motion of the parts contained therein.

(b) ***potential energy*** - the second type of energy is called potential energy (energy that is stored). An example is a rock rolling down a hill. The energy that was in the rock at the top of the hill is potential energy. This energy is inside the rock, due to gravity pulling down on the rock. As the rock rolls down the hill, this potential energy is the rock is converted to kinetic energy. Other examples of potential energy include water behind a dam, and energy stored in a molecule of sugar.

Energy can be converted from potential to kinetic and back to potential, and so on. If you push the rock back up to the hill,

the potential energy in your body is converted to kinetic energy as you push on the rock, and this kinetic energy is converted to potential energy within the rock which can later be converted to kinetic energy.

(c) ***Heat energy*** - in terms of kinetic theory, identically equal to the kinetic energy of motion (rotations, vibrations, translocations) of the component molecules.

(d) ***Specific heat*** - the heat energy required to raise gram of a substance one degree in temperature. A particularly important heat is that of water, by which the unit of heat energy is defined: one calorie is the amount of heat energy required to raise 1 gram of pure water 1 °C from 3.5 - 4.5 °C (where it is the most dense) at 1 atmospheric pressure.

(e) ***Heat capacity*** - the heat energy required to raise 1 molecular weight of a substance 1 °C. The unit of specific heat are cal per °C and of heat capacity are cal per °C mole.

POTENTIAL VS KINETIC ENERGY

Potential energy, as the name implies, is energy that is not yet been used, thus the term potential. Kinetic energy is the energy in use. A tank of gasoline has a certain potential energy that is converted into kinetic energy by the engine. Batteries, when new or recharged, have a certain potential. When placed into a tape recorder and played at a loud volume, the potential in the batteries is transformed into kinetic energy to drive the speakers. When the potential is all used up, the batteries are dead. In case of rechargeable batteries, their potential is re-elevated or restored. In the hydrologic cycle, the sun is the ultimate source of energy, evaporating water (in a fashion raising its potential above water in the ocean). When the water falls as rain (or snow) its potential energy is decreased. Without the sun, the water would eventually still reach sea-level, but never be evaporated to recharge the cycle.

Chemicals may also be considered from a potential energy standpoint. One kilogram of sugar has a certain potential energy. If that kilogram of sugar is burned the energy is released all at once. The energy released is kinetic energy (heat). So much is released that organisms would burn up if all the energy was

released at once, organisms must release the energy a little bit at a time.

Cells convert potential energy, usually in the form of C-C covalent bonds or ATP molecules, into kinetic energy to accomplish cell division, growth, biosynthesis, and active transport, among other things.

SCHEME OF BIOENERGETICS

The molecules in motion possess capacity to carry out work, the energy possessed by a molecule or a body by virtue of its motion is termed as "kinetic energy. This type of energy depends upon the mass of the body and its velocity of motion in accordance to the following reaction:

$$K.E. = 1/2mv^2$$

Where KE represents the kinetic energy, m the mass of the body and v the velocity of the body,

The degradation of the foodstuffs of high potential energy to products of low energy coupled to a mechanism for collection and storage energy (Figure l), This mechanism consists of the conversion of inorganic phosphate compound into high energy phosphate compound. This high energy phosphate compound can in turn be degraded back into inorganic phosphate after releasing energy, Another coupling mechanism ensures that this energy is made into useful energy,

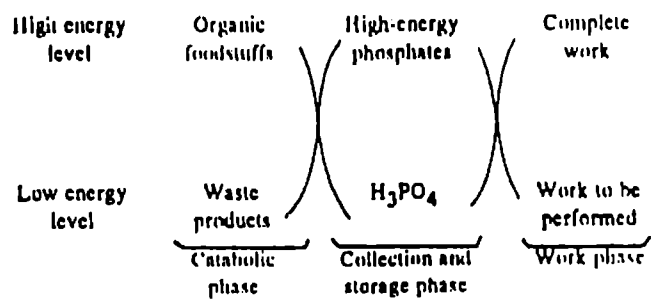


Figure 1: Scheme of bioenergetics.

Catabolism of food stuffs may be divided into three phases: *Firstly,* food material of high molecular weight (polysaccharides, lipids) are digested to products more easily absorbed. These reactions are hydrolytic and the small amount of energy liberated is not recoverable but is lost as heat. Products of this stage of catabolism include mainly common food sources like glycerol, fatty acids and many amino acids.

Second phase of catabolism converts this diverse collection of compounds into small number of intermediates for final oxidation in the next stage. Hexose and fatty acids are firstly activated by the formation of esters with phosphate and co-enzyme-A respectively. At this stage some free energy is liberated. In addition to the anaerobically producing lactic acid or alcohol depending upon the organism, lactic acid formation (glycolysis) makes available about 50,000 calorie and alcohol formation (fermentation) about 62,000 calorie and the rest of the energy is available during later oxidation. Higher animals in general can oxidise glucose completely to Carbon dioxide and water, and liberate much free energy. Under aerobic conditions pyruvic acid rather than lactic acid is the main product and this is further oxidised to acetyl coenzyme-A. Acetyl coenzyme-a, is also produced from fatty acids, the glycerol of fat joins the carbohydrate path at the triose state.

Certain amino acids are catabolized via special pathway, but initially deaminated to form the corresponding keto acids. The nitrogen of amino group follows its own pathway to eventually appear in the urine has urea and ammonia. Sulphur contained in the amino acid (for example methionine) is oxidized to inorganic sulphate, the carbon skeleton of a few amino acids are converted to ketone bodies or acetyl coenzyme-A, But many, amino acids either yield pyruvic acid, α-keto glutaric acid, or oxalacetic acid directly on deamination or form products readily convertible to those α-ketoacids. During this second phase a small amount of recoverable free energy is liberated.

During the *third* phase, all of the products of the second phase are channelled into a single mechanism known as tricarboxylic acid or Krebs cycle (Figure 2), is a sequence of reactions which releases the carbon dioxide, while it feeds pair of their hydrogen into oxidative chain to form water.

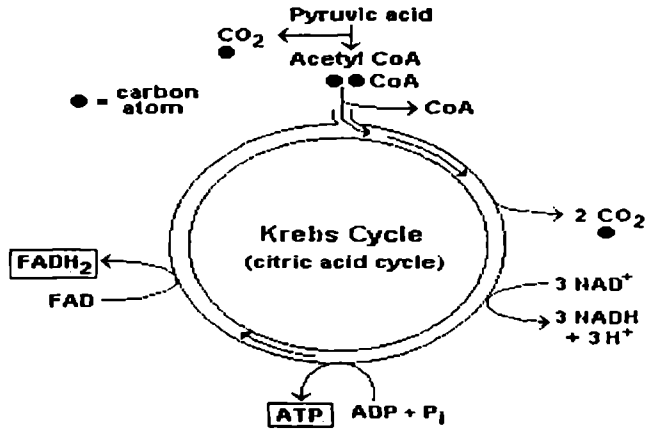


Figure 2: A summary diagram of the Krebs cycle.

Free energy liberated or consumed in any reaction is the difference between the free energies of formation and reactants in the reaction.

LAWS OF THERMODYNAMICS

A quick review of the laws of thermodynamics is in order. While these laws should be familiar to you, you should make sure that you have a good understanding of what these laws mean,

The first law of thermodynamics

Energy can be changed from one form to another/ but it cannot be created or destroyed. The total amount of energy and matter in the universe remains constant, merely changing from one form to another. The first law of thermodynamics (conservation) states that energy is always conserved, it cannot be created or destroyed.

The first law of thermodynamics has been defined in various ways:

1. The total energy in an isolated system always remains constant, although there may be a change from one form to another.
2. The energy of an isolated system remains constant and whenever a quantity of some form of energy disappears, an

exactly equivalent quantity of some other form of energy must be produced.

3. Any gain or loss of energy by the system must be exactly equivalent to the loss or gain respectively by the surroundings of the system.
4. Whenever a certain quantity of energy is produced, an equivalent amount of other form of energy must be used up.
5. Energy can neither be created nor destroyed, the only change which energy can undergo is a transformation from one form to another.
6. The total energy of a system, plus its surroundings remains constant,

These definitions differ slightly from one another but the basic idea is that energy can neither be created nor destroyed.

This law is telling you that energy cannot come from nowhere. Therefore, a reaction which requires the input of energy cannot occur unless there is a source for the needed energy. Imagine that there is pen on the table next to the computer. You would never expect the pen to suddenly rise off the table and move towards the ceiling. That is because such a movement requires energy to move the pen against the force of gravity. Unless someone gives the pen this energy, it will not move. You could pick up the pen and throw it so that it hits the ceiling. In this case, you are giving energy to the pen. This energy ultimately comes from the nutrients that you eat and convert into the muscle and fuel the muscle that you used to move the pen.

Einstein has shown that matter and energy are equivalent and interrelationship is given by the famous Einstein equation:

$$E = mc^2$$

where E is the energy, m is the mass and c^2 is the velocity of light. Hence, 1 gm of matter is equal to 2.1×10^{13} calories, a stupendously large amount of energy.

Suppose we put some amount of heat in a system. Since the heat energy cannot be lost, it must remain either wholly or partly as internal energy in the system, or can wholly or partly be used

up by the system in doing mechanical work. In the general case, when the heat absorbed goes both to increase the internal energy and to produce some mechanical external work, we must have Heat absorbed = Increase of internal energy + work done by the system. If the final and the initial internal energies of the above systems are E_2 and E_1 respectively, then the increase in internal energy is $E_2 - E_1 = \Delta E$ (the symbol Δ always signifies increase, that is Final - Initial, irrespective of whether it is positive or negative). If the heat absorbed is q and the work done by the system W, then by substituting these values in the foregoing equation we get:

$$E_2 - E_1 = \Delta E = q - W$$

The first law of thermodynamics includes more specific mass of constant heat sums which is of great importance in energy transformations. It states that the total amount of heat produced or absorbed if a chemical reaction is carried out in stages is equal to the total amount of heat evolved or consumed when the reaction occurs directly. A good example is the oxidation of glucose to Carbon dioxide and water.

1. $C_6 - H_{12}O_6 + 6O_2 + 6H_2O + 6CO_2 + 673$ (kcal)
2. $C_6H_{12}O_6 = 2C_2H_5OH + 18$ (kcal)
   $2C_2H_5OH + 6O_2 = 6H_2O + 4CO_2 + 655$ (kcal)

Thus, no matter what pathway a particular reaction follows, the total amount of heat evolved or absorbed is always the same.

LIMITATIONS OF FIRST LAW OF THERMODYNAMICS

There are three main limitations of the first law of thermodynamics. These are as follows:

(a) The first law of thermodynamics simply establishes equivalence between different forms of energy. However, this law does not tell us under what conditions and to what extent it is possible to bring about conversion of one form of energy into the other.

(b) The first law of thermodynamics is a quantitative statement which does not prelude the existence of either a heat engine or a refrigerator. The first law does not contradict the existence

of a 100 percent efficient heat engine or a self-acting refrigerator. In practice, these two are not attainable.

(c) The first law fails to explain why chemical reactions do not proceed to completion. If one mole of CO and one mole of H_2O were in the reaction vessel, then if the reaction went to completion, one mole of CO_2 and one mole of H_2 would be produced; However, if the reaction vessel is examined at equilibrium, the yield is less than 100 percent.

Hence, there must be some other law beside the first law, which governs the direction of flow of heat and the extent of its convertibility into work. The limitation forms the basis of the second law of thermodynamics.

The second law of thermodynamics

It states that "in all energy exchanges, if no energy enters or leaves the system, the potential energy of the state will always be less than that of the initial state". There are different ways of stating it:

- The entropy of the universe tends to increase.
- Any process whose net effect is the raising of a weight and - the cooling of an object is impossible.
- Systems tend to proceed from ordered (low entropy) to disordered (high entropy) states.
- The entropy of a system and its surroundings is unchanged by a reversible process and the entropy of a system and its surrounding increases for irreversible processes.
- All natural processes proceed towards equilibrium.
- If you think the world is a mess now, just wait.
- At room temperature.... conversion of a single calorie of thermal energy completely into potential energy is a less likely event than the production of Shakespeare's complete works fifteen quadrillion times in succession without error by a tribe of wild monkeys punching randomly on a set of typewriters.

From these different definitions of the second law of thermodynamics, it implies that whenever energy is changed, some of the energy will be wasted which will lead to an increase

in the disorder of the universe. When dealing with biological systems, this wasted energy is usually in the form of heat. Thus, whenever, any reactions occurs, some of the energy in the reaction will be released as heat. Thus, a watchspring-driven watch will run until the potential energy in the spring is converted, and not again until energy is reapplied to the spring to rewind it. A car that has run out of gas/petrol will not run again until you refuel the car. One the potential energy locked in carbohydrates is converted into kinetic energy, the organism will get no more until energy is input again. In the process of energy transfer, some energy will dissipate as heat. The flow of energy maintains order and life.

The difference between the first and the second law of thermodynamics (see Box 1) lies in the fact that the first law is concerned with the accounting of various kinds of energy involved in a given process, while the second law is concerned with the availability of energy of a given system for doing useful work,

Third law of thermodynamics

The third law of thermodynamics states that "the entropy of a perfect crystal is zero when the temperature of the crystal is equal to absolute zero (OK)". The crystal must be perfect, or else there will be some inherent disorder. It also must be at OK, otherwise there will be thermal motion within the crystal, which leads to disorder.

As the crystal warms to temperatures above OK, the particles in the crystal start to move, generating some disorder. The entropy of the crystal gradually increases with temperature as the average kinetic energy of the particles increases. At the melting point, the entropy of the system increases abruptly as the compound is transformed into a liquid, which is not as well ordered as the solid. The entropy of the liquid gradually increases as the liquid becomes warmer because of the increase in the vibrational, rotational, and translational motion of the particles. At the boiling point, there is another abrupt increase in the entropy of the substance as it is transformed into a random, chaotic gas,

To put these laws of thermodynamics into perspective, let's think of the universe as a whole. No matter where you look, there is energy. In stars, in planets, and in the space between stars and

COMPARISON OF 1ST AND 2ND LAW OF THERMODYNAMICS

The 'modus operandi' of the first law and that of the second law are basically similar.

1st law	*2nd law*
1. It introduces a new concept a quantity called *internal energy.*	1. It introduces a new function, called *entropy.*
2. We measure the change of internal energy by the energy we put into the system minus the energy we take out of it.	2. Entropy change is measured by the heat energy we put into the system divided by the corresponding temperature, summed up over the whole path of the change, provided the whole change is conducted reversibly.
3. The concept of internal energy is as much a logical invention of the 1st law.	3. The concept of entropy is as much a logical invention of the 2nd law.
4. The energy is indestructive and so easily appeals to our material comprehension.	4. Entropy always increases in any natural processes, which is unlike any material change we are familiar with.
5. It tells us without any bias towards any particular direction of change, which processes are energetically	5. It points out the direction in which these processes could go.
6. The first law enforces a necessary condition. for a change to take place.	6. The second law prescribes a sufficient condition

planets. The energy content of the universe is constant, and never changes. This is a very important aspect of the first law, in that there is only so much energy everywhere. The distribution of energy varies widely (like the difference between the energy in your body and the energy in the chair you sit on), but the amount or content in the universe will never change. Energy can be transformed, but it can't be created or destroyed. Each transformation of energy results in the loss of energy, and this loss of energy increases the entropy of the universe. Think of entropy like a jigsaw puzzle. When you throw the pieces all over the table, the puzzle has very high entropy. But as you add pieces into their right place, the entropy of the puzzle decreases, until the puzzle is complete, and its entropy is at its minimum. Heat is a good example of entropy, because right now, your body is losing heat from its chemical reactions that are taking place. This energy is the most random form of energy, therefore has the highest entropy. This heat energy, as a general rule, is lost to the universe, and cannot be used again. That is why the entropy of the universe is increasing as a whole. You can decrease your entropy while you grow and develop, but as you do so, you increase the entropy of the universe around you by eating food and losing heat.

ENTHALPY

The total energy present within a molecule is its internal energy or the intrinsic energy, it includes both the kinetic energy and the potential energy of the electrons, nuclear components and the molecules, de kinetic energy of the electrons about the nuclei. In molecules the kinetic energy may be translational (by which the molecules move in a liquid or gaseous state), rotational (the energy due to the rotation of the molecules about an axis passing through the centre of gravity), or vibrational (by which the atoms in a molecule vibrate with respect to one another). The kinetic energy does not contribute much to the total energy of the system. The potential energy of the electrons is due to electrostatic interactions between the electrons and the nuclei, electrons and electrons, or nuclei and nuclei of various atoms. The energy associated with such interactions is the bonding energy. The *nuclear energy* is the force by which the nuclear components (protons and neutrons) are held together.

The energy change taking place during a chemical reaction is the *chemical energy* which is mostly related to the bond energies of a molecule. Normally in a chemical reaction, energy is transacted in terms of heat.

The heat content of a system varies with the temperature and pressure. Enthalpy (H) is the heat content of a system which includes the internal energy (ΔU) and the product of pressure (P) × volume (V):

$$\Delta H = U \times PV$$
$$\Delta H = \Delta U + P \times \Delta V$$

where ΔH is the change in enthalpy, ΔU is the total energy change and ΔV is the change in volume. When the enthalpy change is more than zero, energy enters into the system and the reaction is endothermic. When the energy is lost from a system the enthalpy change is less than zero, and the reaction is exothermic.

The change in '*heat content*' or enthalpy is measured calorimetrically. Thus, ΔH for the oxidation of glucose is –673,000 cal/mole of glucose at 20 °C and at 1 atmospheric pressure. The change in enthalpy varies with temperature and therefore it is customary to denote the particular ΔH value with a subscript that indicates the appropriate absolute temperature in the Kelvin scale (0 °C = 273.1 °K; 20 °C = 273 + 23 = 293 °K)

$$C_2H_{12}O_6\ (s) + 6O_2\ (g) \longrightarrow 6H_2O\ (l) + 6CO_2\ (g)$$
$$\Delta H_{293} = 673 \text{ Kcals.}$$

This value is also known as '*heat of combustion*' because it is the heat released by the complete oxidation/burning in oxygen. The heat of combustion of some organic substances have been shown in Table 2.

Table 2: Heats of combustion of certain biochemical substances at 20°C and at 1 atm.

Substance	*H Kcal/mole*	*Substance*	*H Kcal/mole*
Glucose (s)	673	Stearic acid (s)	2680
Galactose (s)	670	Oleic acid (1)	2657
Maltose (s)	1350	Glycine (s)	234
Sucrose (s)	1349	Tyrosine (s)	1070
Palmitic acid (s)	2380	Urea (s)	152

The energy present in a material can be divided into two forms: *(1) isothermally unavailable energy,* and *(2) isothermally available energy.* An example of isothermally unavailable energy is the kinetic energy of the water molecule stored in a beaker. The molecules of water exhibit molecular motion due to their energy and such energy cannot be made to do work at room temperature. Thus this energy is the isothermally unavailable energy, isothermally available energy is the energy available for work. This is named *free energy* which is analogous to the potential energy in a mechanical system. Free energy is given the symbol G (also known as F) and the change in free energy is represented by ΔG or ΔF. The free energy change (ΔG) in a chemical reaction is the driving force of the reaction.

GIBBS FREE ENERGY

Now, these two laws are relatively easy to repeat and understand - but how can we use them to answer questions dealing with biological reactions. Through our experience, you have a feeling for how these laws apply to macro events that surround us (such as the pen moving toward the ceiling). You do not, however, have an understanding about how these laws apply to biochemical reactions that not only you can't see but you didn't even know were occurring. For example, the following is a reaction that every cell performs, the adding of a phosphate group (PO_4^{2-}) to the sixth carbon of the monosaccharide, glucose.

PO_4^{2-} + glucose $\longleftrightarrow$ glucose-6-phosphate

Is this reaction possible? Most likely, you have no idea as you have little practical experience dealing with this reaction. You might assume that since every cell can phosphorylate glucose, this reaction is thermodynamically permissible. But in actuality, this reaction, as written, is (usually) impossible. This is because this reaction requires energy to occur and there is no source for this energy (thus this reaction is analogous to the pen moving to the ceiling). How were you to know this? You couldn't. At some time, a scientist measured the change in the energy between the reactants (PO_4^{2-} + glucose) and the product (glucose-6-phosphate) and found that under certain conditions, the product contained more energy (I bet that got the scientist tenure!). Note: the scientist's

measurements took into account both actual energy changes as well as entropy (disorder) changes. These measurements are reported using a value known as Gibbs Free Energy, abbreviated ***delta* G.** For a reaction in which the products have more energy than the reactants (thus a gain in energy), the *delta G* for that reaction is positive. That means that a reaction with a positive *delta G* cannot occur. For a reaction in which the products have less energy than the reactants (thus a loss in energy), the *delta G* for that reaction is negative and the reaction is thermodynamically permissible. You should remember this!

This should leave you wondering - the phosphorylation of glucose (as shown above) has a positive *delta G* and thus cannot occur, but every cell phosphorylates glucose! How can this be? In order for the reaction to occur, a source of energy is needed (like the pen to the ceiling). A common source of energy that cells use is a compound called ATP (adenosine triphosphate). ATP can undergo the hydrolysis (breaking off) of one of the phosphate groups to produce adenosine diphosphate, ADP and a free phosphate. This is written as:

$$ATP \longleftrightarrow ADP + PO_4^{2-}$$

This reaction has a negative *delta G.* Thus if the cell breaks up ATP at the same time that it is adding the phosphate to glucose, the cell can use the energy released from ATP to fuel the addition of the phosphate to glucose. This requires that the amount of energy released from the hydrolysis of ATP is greater than the amount of energy needed to add the phosphate. If this occurs, then these reactions are said to be coupled. If so, then the reaction would be written as:

$$ATP + PO_4^{2-} + \text{glucose} \longleftrightarrow \text{glucose-6-phosphate} + ADP + PO_4^{2-}$$

This reaction can be simplified (just like in Algebra) by noticing that there is PO_4^{2-} on each side of the reaction and can be cancelled.

$$ATP + \text{glucose} \longleftrightarrow \text{glucose-6-phosphate} + ADP$$

Now, where do cells get the ATP? Obviously, if the breakdown of ATP has a negative *delta* G (gives off energy), the formation of ATP must have a positive *delta* G.

$$ADP + PO_4^{2-} \longleftrightarrow ATP$$

Where do cells get the energy to make ATP (remember, most of the cellular reactions that require energy use the hydrolysis of ATP as the source - ATP is like money to the cell). The ultimate source of this energy is from the nutrients eaten by organisms. The breakdown of nutrients such as sugars and fats into carbon dioxide and water has a negative *delta* G. The cells are capable of taking these nutrients, breaking them down and using the energy in them to make ATP.

ENTROPY

Entropy is defined as the degree of freedom that particles of matter have. Gases have a greater entropy than pure liquids which have a greater entropy than pure solids. Solutions have a greater entropy than pure liquids as the particles in a solution are more separated and solvent molecules separate the solute particles.

The word entropy is sometimes confused with energy. Although they are related quantities, they are distinct. Energy measures the capability of an object or system to do work. Enthalpy on the other hand, is a measure of the "disorder" of a system. What disorder refers to is really the number of different microscopic states a system can be in, given that the system has a particular fixed composition, volume, energy, pressure, and temperature. By "Microscopic states", we mean the exact states of all the molecules making up the system.

The idea here is that just knowing the composition volume, energy, pressure, and temperature doesn't tell you very much about the exact state of each molecule making up the system. For even a very small piece of matter, there can be trillions of different microscopic states, all of which correspond to the sample having the same composition, volume, energy, pressure, and temperature. But you are ignorant of exactly which one of the system is in at any given moment - and that turns out to be important.

For those who are technically inclined, the exact definition of entropy is:

Entropy = (Boltzmann's constant k) × logarithm number of possible states
= k log (N)

Since the logarithm of a number always increases as the number increases, we see that the more possible states that the system can be in (given that it has a particular volume, energy, pressure, and temperature), then the greater the entropy.

Again, because we cannot see which particular microscopic state of a system is in, people often like to say that entropy is quantitative measure of how uncertain or ignorant one is about the exact, detailed, microscopic state of a system, or another popular way of saying this is that entropy measures the microscopic disorder of a system.

As a simple example, suppose that you put a marble in a large box, and shook the box around, and you did not look inside afterwards, Then the marble could be anywhere in the box. Because the box is large, there are many possible places inside the box that the marble could be, so the marble in the box has a high entropy. Now suppose you put the marble in a tiny box and shook up the box. Now even though you shook the box, you pretty much know where the parable is, because the box is small, in this case we say that the marble in the box has low entropy.

The same idea applies to the arrangements of atoms in a jar at room temperature. The smaller the jar, the lower the entropy. But keep in mind that we also have to consider the *velocities* of the gas particles to have full knowledge of their states, the higher the temperature of the gas, the faster the gas particles move on an average, so the wider the range of possible velocities for the gas particles, and hence, the more uncertainty we have about the velocity of any particular particle. Thus, higher temperature, as well as greater volume, mean higher entropy.

Entropy like energy, volume, temperature, and pressure, is another *thermodynamic state variable* of a system. It turns out that, for a simple system, if you know any two of these state variable, then the others are all determined. Although the word entropy might seem like a mysterious concept, it is really not. Remember that it is really just a measure of the number of states a system can be in, given the constraints on the system.

What is entropy good for ? Knowing the entropy of a system can tell us many things about what can and cannot happen. In particular, it is the basis for the second law of thermodynamics; the universe evolves such that its total entropy always stays the same or increases.

Why is this so ? In fact, the basic idea of entropy is simple to understand. Suppose you are floating out in space and you have a jar containing a particular gas, say argon, then you open the jar for a moment, the argon will almost certainly escape out into space. After the argon has escaped, its entropy is greatly increased (and it continues to increase as the gas expands). How do I know that the entropy has increased ? This is because the number of states that the argon gas can be in when it occupies a much larger volume is much greater than when it was confined to the Jar. So, the entropy of the gas increases when the argon escapes. But why must the argon escape ? Well, in fact, prior to opening the jar, if you arranged the microscopic states of the argon molecules in just the right way, you could open the jar for a moment and *not* have the argon escape. The point is that it is highly *improbable* that the argon is in one of these special non-escaping states when you open the jar - most of the states lead to the gas escaping. This is really the content of the second law - that if you begin not knowing the microscopic state of a system, then the system is more than likely to evolve to state where you are even more ignorant of its exact microscopic state. Just knowing the thermodynamic state variables of a system, such as its temperature and pressure, means you are in fact ignorant about the initial exact microscopic state - all you know from the states it can be in, that is the entropy. Hence, for most situations we encounter, chances are that entropy will increase with time.

It is very interesting to compare the behaviour of entropy compared to energy, unlike energy, entropy can be created (but not generally destroyed). In fact, your body is creating right now as it generates heat. One of the reasons that your body temperature has to be higher than the surroundings air or that you have a sweat off water if it is not, is that you have to get rid of the extra entropy (otherwise, you would become disorganized and eventually die), the energy that your warm body radiates carries away the extra

entropy. It does this because losing this energy decreases the number of microscopic states that the atoms and molecules of your body can be in.

Another practical example of entropy is the use of a source of heat, say, from steam generated by heating water, to drive some kind of turbine. Then, it turns out, by considering extropy, that the maximum efficiency of our process will be less than 100 percent. The reason that this is so is because when heat is brought into the turbine, it carries with it some entropy. We cannot keep this entropy in the turbine, because the turbine would become microscopically disordered and eventually break. So, some heat energy has to be released to the outside world to get rid of this entropy to protect the turbine, The heat released for this purpose therefore cannot be converted into work. We get rid of the unwanted entropy by rejecting this heat to the outside world at a lower temperature than we brought the heat in at. The reason for the lower temperature is that the heat released into a low temperature environment carries out more entropy from the turbine than the entropy this same amount of heat carries into the turbine at a high temperature, this is because heat disrupts a cold system more than a hot one, because the hot one is already more disordered. Thus, we must only sacrifice some of the heat carried into the turbine to get rid of the entropy imported into the turbine by that heat in the first place, one can see from this discussion, why power plants need a cold temperature environment to dump their waste heat.

In other words, the kinetic energy of the steam molecules is large (because the steam is hot), but the directions of the molecules are disordered. Somehow, to convert all of the energy of the steam into useful work, we have to line them all up in the same direction. But we are ignorant of the exact configuration at any given instant, and even if we are not, how are we going to get in there and actually do it for each molecule? Clearly, the microscopic disorder is a barrier.

The above example demonstrates now heat energy, because it cannot be completely converted to mechanical energy in a turbine, is, in a sense, of lesser quality than mechanical energy.

Entropy can also be applied to many other situations. For example, it can be used to predict the direction that a chemical reaction will proceed.

WHY DO WE NEED THE CONCEPT OF ENTROPY?

The conservation of energy is the bedrock of our understanding of nature. Conservation of energy, however, only tells us that in any process the total energy remains constant, that is $\Delta E = 0$ Energy conservation, however, does not specify what can enter and leave the energy conservation equation, only that the sum of energy entering and leaving a system must always be the same. Furthermore, energy conservation does not tell us what are the various forms of energy, and how energy **transforms** from one form into another. From our study of waves, light, electric and magnetic fields and so on, we found that energy plays a central role in all these phenomena. However, only the specific study of the various phenomena could unravel the specific forms that energy takes. The concept of energy also does not tell us what fraction of the total energy is kinetic and what fraction is electrical, chemical and so on. A new concept in addition to energy, is needed for a more complete understanding of physical processes, in particular those involving the *transformation of energy*. In all the examples we studied, energy could be completely transformed from one form into another. For example, in our study of magnetic fields, we found that we could fully convert mechanical energy into electrical currents by Faraday's Law of Induction. We may well ask: can all forms of energy be fully transformed from one form to another? The answer is unexpectedly complicated. In some cases the answer is "yes", and in other cases the answer is "no". The concept **of entropy** addresses this question regarding the transformation of the forms of energy, and tells us how much energy can be converted from one form into another, and in particular, how much energy is available for doing useful work. Heat had defied all attempts to explain it using the concept of force, and, in general using Newtonian mechanics. The fact that heat is a form of energy presented a great breakthrough in the understanding of heat, and was fundamental to the emergence of

energy as one of the most fundamental ideas in physics. It was, however, soon realized that heat was not an ordinary form of energy such as gravitational or electrical energy. Rather, heat is a unique form of energy in that only a **certain fraction** of heat energy can be transformed into other forms of energy. It is for understanding the unique problems presented by heat that new concepts such as entropy, which go beyond energy, were developed. The idea of entropy was introduced into physics in the nineteenth century and is essential for understanding the phenomenon of heat. Entropy is derived from the Greek word meaning "transformation content". Famous names such as Sadi Carnot, R.J. Clausius, Count Rumford, James Joule, Ludwig Boltzmann and so on contributed greatly to the ideas of heat and entropy. Entropy is a concept which does not answer all the questions left unanswered by energy, but it does address a crucial aspect of energy, namely how much energy can be **transformed** from one form into another. In particular, if a certain amount of heat energy is given, then entropy tells us how much "useful" work can be extracted from heat energy, in other words, what is the maximum possible efficiency of a heat engine. In general, since energy is conserved, one would think there is no need to be concerned about recklessly wasting energy. But we know from daily life that "useful" energy, or equivalently, energy "available" for use, is a scarce resource. This intuitive understanding that useful energy is a precious resource is explained by the concept of entropy. Entropy is also related to the concept or order and disorder, and a phase transition from say water to gas is permissible only if entropy increases in such a transformation. For any system, entropy is a physically measurable quantity. Furthermore, it is an experimentally observed fact that for **all** processes in nature, **the entropy of the total system can never decrease.** Note this empirical law concerning entropy is weaker than that of energy, since it does not specify, for any physical process, by **how much** does entropy increase. The concept of entropy has many other applications and vast ramifications in other disciplines. Only those chemical processes and reactions for which the total entropy of the system does not decrease are allowed by nature. All living entities must maintain their low entropic state to stay alive. More abstract applications of entropy occur in the physics of black

holes, and it was only in 1997 that the entropy of black holes could be calculated from first principles using results from string theory. The concept of entropy also has a key application in information theory. It was shown by Claude Shannon in 1950 that the information content of a message is determined by a suitable application of the idea of entropy to the discipline of information science. The current understanding of entropy is that it is a large scale (macroscopic) manifestation of the **atomic** nature of matter. For example, one cubic centimetre of air contains about 10^{20} number of atoms. This is an unimaginably large number of particles. Even if under some circumstances we can view these atoms as classical particles, it is futile to apply Newton's laws to such a large collection of particles. Instead, the best we can do is to acknowledge our **total ignorance** and **assume** that all the atoms are moving **randomly,** that is, the position and velocity of every atom is a **random variable.** Furthermore, for a system of atoms that reaches a stable state, called equilibrium, **every possible** velocity and position of the atoms in one cubic centimetre of air is **equally likely.** The empirical law that entropy never decreases is simply a statement that for a system consisting of a collection of atoms which is **not** in equilibrium, the collection of atoms will always move towards a state that is the **more probable,** and will only reach equilibrium on arriving at the **most probable** state available to the system. It should be noted that the field of thermodynamics developed in the eighteenth and nineteenth centuries without the concept of the atomic composition of matter. It is remarkable that very general results on how bulk matter behaves were obtained without any idea of the microscopic composition of matter. From the point of view of scientific methodology, the progress of thermodynamics shows that there are concepts describing the large scale properties of matter that have a logic of their own, and which later on are seen to match on smoothly to a deeper and more complete understanding of the same phenomenon.

ENTROPY AND DISORDER

Entropy is frequently used as a synonym for disorder. It can be explained that entropy only strictly increases in an isolated system (closed system). An isolated system is a system that does not interact in any way with its surroundings, in practice there are really no completely isolated systems in nature except, perhaps, the universe as a whole, therefore, the total entropy of the universe is always increasing.

Consider two chambers, separated by a dividing wall (Figure 3). Assume, we shot 25 balls into the chamber 1, each with 5 J kinetic energy. The balls will bounce around in the chamber and hit the wall and each other. If the walls of the chamber are perfectly hard and the coefficient of restitution of the ball is 1, then the average kinetic energy of the balls in chamber 1 will stay 5 J, even so some balls will gain and some will loose energy in the collisions. Assume we shot 25 balls into chamber *2,* each with 15 J of kinetic energy. The average kinetic energy of these balls will stay 15 J. So long as chamber 1 and chamber 2 are separated by a dividing wall, the balls on one side will be "hot" and the balls on the other side will be "cool". If we cut a hole into the dividing wall, big enough for a ball to pass through (Figure 4) and wait long enough, the average kinetic energy of the balls on either side of the wall will be approximately 10 J. There *will* be "hot" balls, with energies above 10 J, and "cool" balls, with energies below 10 J, on either side of the wall.

While Newton's law do not forbid all the hot balls to gather on one side and all the cool ones on the other side, the probability that this will happen is practically zero. There are very large number of ways to distribute the energy among all the balls. *Any one specific way is equally likely or unlikely.* It is just as unlikely for each ball to have exactly 10 j of kinetic energy than for one ball to have 500 j and all others to have O J. There are many ways of having a disorderly arrangement than of having an orderly arrangement.

Disorder is more likely than order. The amount of disorder is the number of ways the inside of a system can be arranged so that from the outside things looks the same. It turns out that the logarithm of the number of ways is proportional to the entropy.

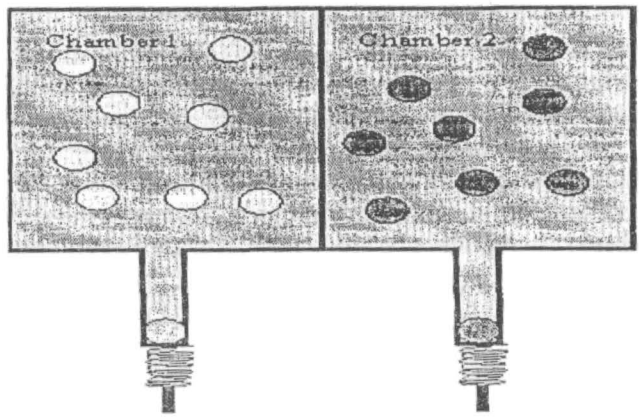


Figure 3: Shot ball game with separate chambers.

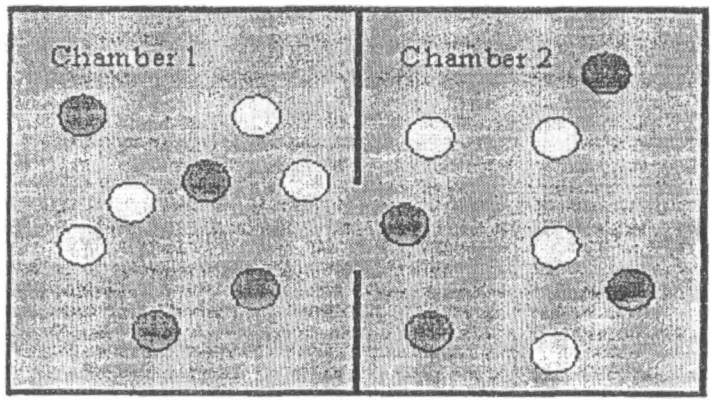


Figure 4: Shot ball game with partially separate chambers.

We can define the entropy as the logarithm of the disorder times some constant of proportionality, then we change the entropy of a substance by an amount $\Delta S = \Delta QT$, we change the disorder of the substance. Entropy always increases, because a high amount of disorder, is more likely than a low amount of disorder.

Here are some situations in which entropy always increases:

(a) The entropy increases whenever heat flows from a hot object to a cold object.

(b) The entropy increases when ice melts, water is heated, water boils, water evaporates.

(c) The entropy increases when a gas flows from a container under high pressure into a region of lower pressure.

(d) The entropy increases when we spray something out of an aerosol can or we let air out of a tire.

The entropy decreases when water freezes, this does not violate the second law of thermodynamics. The second law does not say that entropy can never decrease anywhere. It just says that the total entropy of the universe can never decrease. Entropy can decrease somewhere, provided it increases somewhere else atleast as much. The entropy of a system decreases only when it interacts with some other system whose entropy increases in the process.

Does the entropy of a closed system always increase, or could it possibly decrease the standard answer to this question is that entropy (disorder) will increase, but there are atleast two ways in which entropy can decrease in a closed system.

The laws of probability allow a closed system's entropy to decrease, but with such a low likelihood that the odds would make it very unlikely. Making the system small enough, however, by decreasing the number of possible states can help improve the odds.

Take for example, a movie of a billiards game 'break' shot. The ordered arrangement of balls become disordered, but running the film in reverse would show each individual collision obeying the usual physical laws. The time reversal would be apparent, when all the balls ended up in an ordered collection. Although that result could conceivably occur by chance, it is very unlikely. Reducing the example to just two balls would make the odds of an orderly arrangement occurring more likely.

For an example of decreasing entropy, start with a closed system large enough to allow significant gravitational forces among its components. Gravity provides a 'negative energy' that can take a completely disordered system and organise it into a radically symmetric arrangement around a common centre of gravity.

ENTROPY AND EVOLUTION

Entropy refers to increased disorder in any natural process. Evolution refers to diversification of form. Both involve a time - dependent relationship between what is actual and what is possible.

Many discussions about evolution are based on inaccurate considerations of entropy and complexity, one common fallacy asserts that entropy is a universal tendency towards disorder, that evolution is a process leading to greater order, and therefore the evolution of life is impossible because this would require an increase in order, whereas according to entropy, in any natural process the amount of disorder increases.

The complexity is sometimes defined as the number of bits needed to precisely describe a system. This definition implies that orderly systems can be described more compactly than systems that seem more disorderly, that would seem to contradict the usual notions that entropy is some kind of tendency toward disorder.

A more useful definition of entropy is that it is a measure of the tendency of energy to become less concentrated over time by occupying a greater number of microstates. The more concentrated a source of energy, the greater the likelihood that it will move into configuration in which it is more diffuse. Energy tends to dissipate or spread out as time passes, and so transformations leading to more dispersed are more probable than transformations leading to more concentrated energy.

Entropy brings about unceasing change. Arrangements of matter and energy continuously appear and disappear as energy flows and disperses. Any particular arrangement is usually only one of a vast number of possible arrangements, and is unlikely either to persist or arise again unless it leads to greater dispersal of energy. Therefore, those arrangements that do happen to be stable will accumulate over time as other arrangements disappear.

Persistence is an important factor in any evolutionary process. Evolutionary processes depend on some sort of selective barrier - reversibility. Replication is a possible mechanism for achieving this, but other frameworks for establishing memory are theoretically possible, in general, evolutionary processes also nearly always

involve selecting higher activation thresholds, either thresholds for chemical reactions or thresholds for analogous transformations in whatever space the evolutionary process is unfolding.

Evolution is the condensation of time itself. By definition of space, separate events cannot occur in precisely the same place at the same time. Therefore, a definite increment of time is ineradicably necessary for a finite number of possible configurations to be traversed and explored, the incompressibility of space is at the core of evolutionary processes in general.

A system evolves when it moves through a pattern space it has traversed, the more highly evolved the system. As a system becomes more highly evolved, it accumulates records that reduce the number of fruitless trials and enables it to consent directly to regions of space that are more richly connected to other patterns within that space. The evolving system builds upon and exploits its experience to discover wormholes through the pattern space.

In this sense, evolution is analogous to entropy, although evolution is morphological rather than energetic.

ROLE OF HIGH ENERGY PHOSPHATES

The high energy phosphates are important for the following reasons:

1. The phosphate group can be transferred to another organic molecule, in this transfer much of the high energy is not dissipated as heat. The product, which is a phosphorylated molecule, has a total energy content exceeding that of the non-phosphorylated molecule. For example, the interaction of ATP with glucose:

   glucose + ATP ⟷ glucose - P + ADP

2. High energy phosphate is used (a) to effect chemical synthesis; (b) to perform work (muscular, osmotic, secretory), (c) to liberate heat maintaining the body temperature.

HIGH ENERGY ESTERS, COENZIME-A

Thio-esters can be considered to be of the high energy type. This transfers the acyl residue of acetic acid and other carboxylic acids, Acyl group bound to Co-A (A for acetylation) have a high potential for group transfer. Hydrolysis of acetyl Co-A compounds is exergonic to the extent of about 8 - 10 kcal/g mol. The most important Co-A compound is acetyl Co-A activated acetate, Here the acetyl residue CH_3CO is bound to the free SH group. Then pyruvate is oxidized metabolically to acetate, the released energy is not wasted as heat. Instead, the actual product is the coenzyme A ester of acetic acid (acetyl Co-A, CH_3-Co-S-Co-A), a high energy compound, Acetyl Co-A can readily undergo reactions which would require an outside supply of energy, if free acetic acid was involved.

SOURCES OF HEAT

The fundamental requirement of a bioenergetic system is the source of energy, The terrestrial organisms are immersed in a radiation flux at all times, There may be direct sunlight incident on the organism, scattered sky light coming from thermosphere above, and sunlight plus skylight reflected off the ground and surrounding surface whose reflectivity is 'r'. The atmosphere emits radiation to the organism and the ground surface emits infrared radiation, A building or room is like a black-body cavity and the primary energy which a person inside receives is the infrared radiation emitted by the walls, ceiling, floor and other objects. The internal sources of energy of a biological system is a compound capable of being degraded to products of lower potential energy. As most of the energy yielding reactions are oxidative, the material are oxidizable foodstuffs. These reactions are accompanied by the liberation of energy as the reacting system moves from a higher to a lower energy level. Generally, the energy liberated is in the form of heat. The phosphorylated compounds store and distribute the liberated energy. Therefore, it is possible to carry out synthetic reactions, muscular contraction, nerve conduction, secretion, synthesis etc. by obtaining energy from chemical or coupling to oxidative reactions. In other words, the final mechanism of utilization of stored energy is the conversion of free energy into such type of work as muscle contraction, secretion, synthesis etc.

ENERGY COUPLING AND ENERGY TRANSFORMATION

The chemical energy of foodstuffs represents stored solar energy. This energy is locked in the different covalent bonds between the atoms of a molecule. Foods are metabolized in the body to other end products with an overall decrease in free energy which is utilized for synthetic processes and other work (Figure 5). These reactions are exergonic and take place with decrease in free energy. There are reactions called endergonic in which there is a free energy ΔF, increase of the system, in order to continue the endergonic reactions, they are coupled with exergonic reactions, so that in the overall process there is a net decrease in free energy. Thus, the free energy derived from one reaction may be utilized to drive another reaction. The most important mechanism of coupling is the formation of an energy rich intermediate compound, called adenosine triphosphate (ATP).

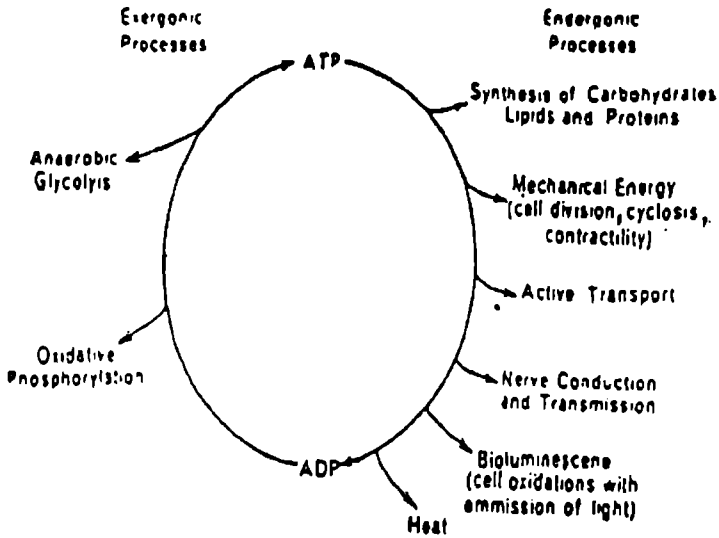


Fig. 5: Energy wheel showing the relationship between exergonic an endergonic process through ATP

Coupling of endergonic reactions may be illustrated by the reactions involved in glycogen synthesis:

ATP + Glucose = Glucose-6-P + ADP (DF D= –4.4 Kcal/mol)
Glucose - 6 - P = Glucose-1-P (DF = + 1.7 Kcal/mol)
(Hexose) n+ Glucose-1-P (Hexose) n+1+Pi (AF = -0.6 Kcal/mol)
(Hexose) n+ Glucose+ATP (Hexose) n+1+ADP+Pi (DF = -3.3 Kcal/mol)

In the foregoing reactions

ATP= the phosphorylating agent adenosine triphosphate,

Pi = inorganic phosphate, (Hexose) n= a branched glycogen molecule and (Hexose) n+1 = the glycogen molecule to which a glucose group has been added.

HIGH ENERGY BONDS

The energy rich bonds have a large potential energy to be used in all internal operations of living matter. They release a large amount of energy rich bonds which is symbolised by the sign '~' Free chemical energy is trapped in the form of these energy rich bonds and then utilized in various tasks.

The cell obtains usable energy primarily by utilizing the high energy phosphate compounds, the most important high energy compound is adenosine triphosphate (ATP). ATP contains purine base adenine, ribose and three molecules of phosphoric acid (Figure 6). the simplified formula of ATP and its transformation into ADP may be written as follows:

A=P~P~P ⟷ A-P~P+Pi + 7000 calories

Where A = adenosine; p a phosphate. The above reaction shows that the release of the terminal phosphate of ATP produces about 7000 calories instead of 3000 calories from common chemical bonds.

Other nucleotides having high energy bonds, such as cytosine triphosphate (CTP), uridine triphosphate (UTP) and guanosine triphosphate (GTP) are involved in nucleoside triphosphates. ATP is the energy source for these nucleoside triphosphates. The

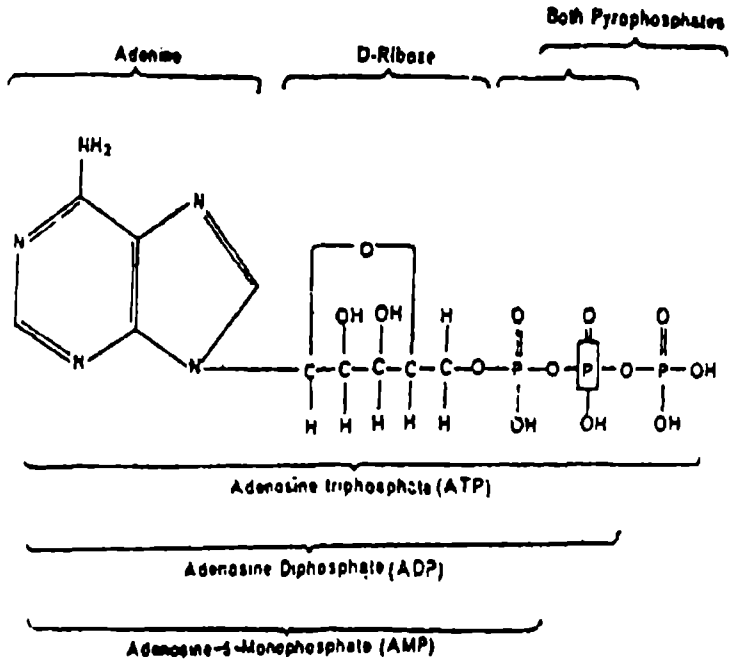


Fig. 6: Molecules of ATP, ADP and AMP

nucleoside triphosphates of the rlbose and deoxyribose type (i.e., dATP) are used as energy sources for -the synthesis of biological compounds (Fig. 7). High energy phosphate bonds are also found in diphosphoglyceric acid, acetyl phosphate, phosphoenol-pyruvic acid and creatine phosphate.

BIGENERGETIC PATHWAYS

There are two main kinds of reactions in our body or in our cells. The first are called endergonic reactions, and these require an input of energy. The second reactions are called exergonic reactions, and these reactions give off energy. An ice pack which we squeeze together to make very cold is an example of an endergonic reaction. It is cold because it is absorbing energy (and heat) from its surroundings. A heat peak, which gets hot when you squeeze it together, is an example of an exergonic reaction.

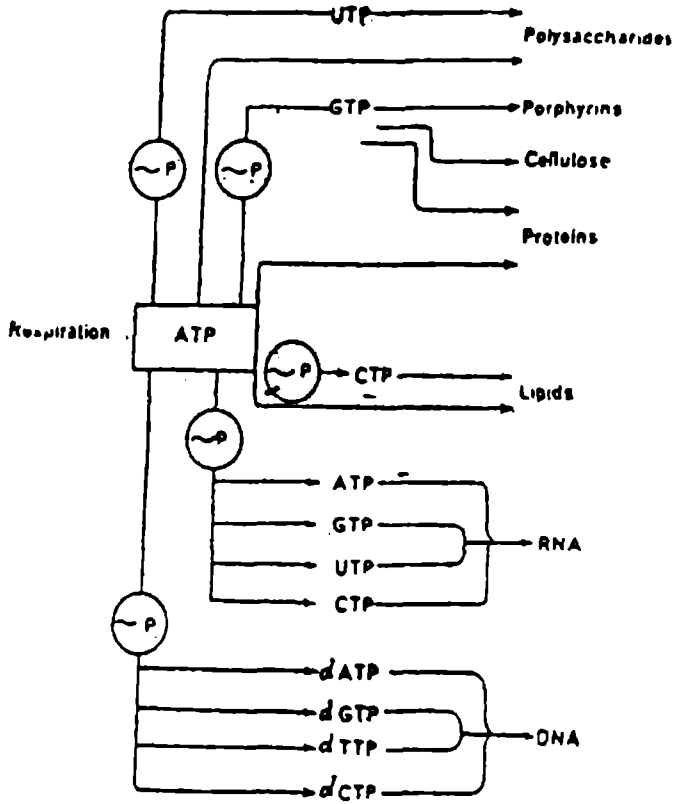


Fig. 7: Channeling of phosphate bond energy by ATP into specific biosynthetic

Many reactions are coupled which means that the energy given off from a exergonic reaction can be used by an endergonic reaction. Adenosine triphosphate (ATP) is a good example of a source of energy in our body, it is the main source of energy for our cells to perform work. Structurally, ATP consists of the adenine nucleotide (ribose sugar, adenine base and three phosphate groups (Figure 8).

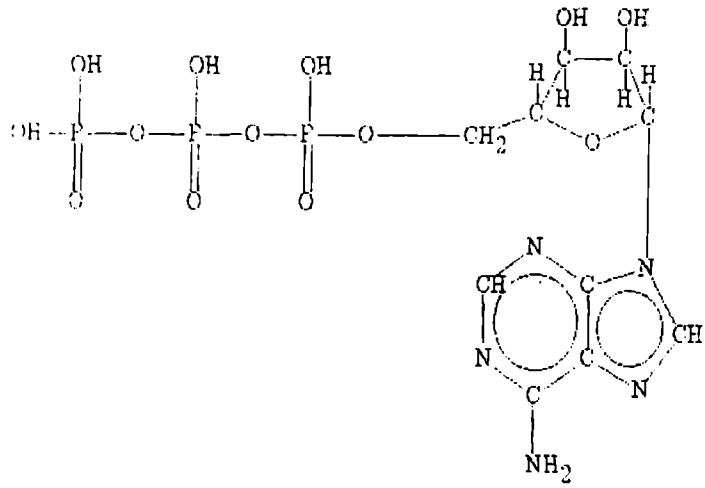


Figure 8: A 2-D stick view of the structure of ATP.

ENERGY MOBILIZATION IN LIVING SYSTEMS

Energy mobilization in living systems takes place by coupled flows of metabolites. The flow of metabolites is coupled to a flow of electrons and protons up and down the electronic/orotonic gradients via the inter-conversion of ATP and ADP. The energy of the proton absorbed by chlorophyll is coupled to electron transport. Electron transport if coupled to the translocation of protons across the energy transducing membrane, The proton gradient thereby created supplies the proton motive force for the synthesis of ATP from ADP and pi and finally the hydrolysis of ATP back to ADP and pi is coupled to practically all thermodynamically uphill or energy requiring reactions. All coupled flows are vectorial, the flows are in the direction of their of their respective forces or gradients. In addition, two features may be noted.

First, the coupling are *symmetrical* for the most energetically efficient processes. It means that the forces have reciprocal effects on the coupled flows, and also, if the forces are reversed, so are

the flows. This applies to ATP synthesis from ADP and pi, coupled to proton transport in oxidative and photosynthetic phosphorylation; as well as ATP splitting coupled to the molecular motor in muscle contraction. ATP is split into ADP and pi by the ATP synthase embedded in the membrane when the proton gradient is run in reverse, just as ATP is synthesised by the molecular motor when ADP and pi are supplied.

The second notable feature of the coupled flows of energy and material is that they are *cyclical*, as a casual glance at a metabolic chart will convince us. Cycles differ in lengths from the tricaroxylic acid cycle of core metabolism to the relatively short redox cycles in the elements of the electron transport chain and the two state interconversions of intermediates such as NADH/ NAD and ATP/ADP. Are the two features -symmetrical coupling and cyclical flows - predicted from thermodynamics ?

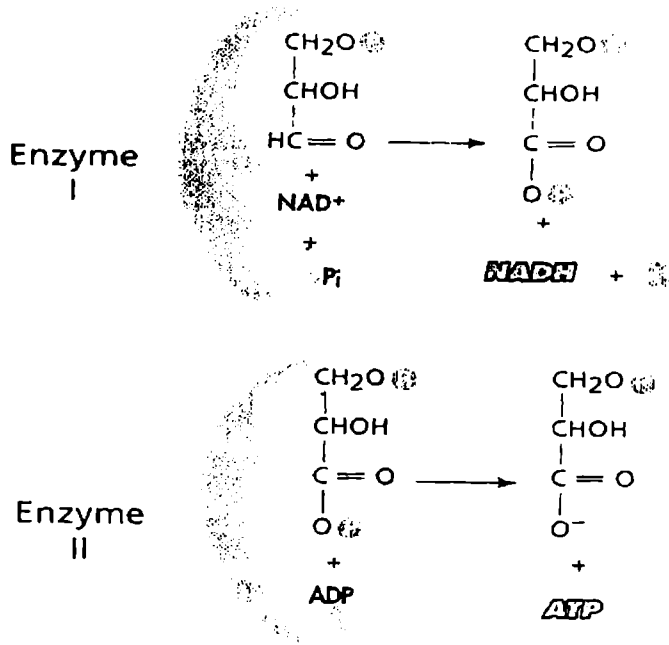


Figure 9: Enzymes and the formation of NADH and ATP.

Energy is stored in the *covalent bonds* between phosphates, with the greatest amount of energy (approximately 7 k cal/mole) in the bond between the second and third phosphate groups. This covalent bond is known as a phosphate bond. The chemical reactions for the formation of ATP and breakdown of ATP are as follows

$$ADP + Pi + energy \longrightarrow ATP$$
$$ATP \longrightarrow ADP + energy + Pi$$

The analogy between ATP and rechargeable batteries is appropriate. The batteries are used, giving *up* their potential energy until it has all been converted into kinetic energy and heat/unusable energy. Recharged batteries (into which energy has been put) can be used after the input of additional energy. Thus, ATP is the higher energy form (the recharged battery) while ADP is the lower energy form (the used battery). When the terminal (third) phosphate is cut loose, ATP becomes ADP, and the stored energy is released for some biological process to utilize. The input of additional energy (plus a phosphate group) recharges ADP to ATP.

ENEDERGONIC AND EXERGONIC REACTIONS

Energy releasing processes that 'generate' energy, are termed exergonic reactions. Reactions that require energy to initiate the reaction are known as endergonic reactions. All natural processes tend to proceed in such a direction that the disorder or randomness of the universe increases (the second law of thermodynamics).

OXIDATION/REDUCTION REACTIONS

Biochemical reactions in living organisms are essentially energy transfers, often they occur together, "linked", in what are referred to as oxidation/reduction reactions, *Reduction* is the gain of an electron. Sometimes we also have H ions along for the ride, so reduction also becomes the gain of H, *Oxidation* is the loss of an electron (or hydrogen). In oxidation reduction reactions, one chemical is oxidized, and its electrons are passed (like a hot potato) to another (reduced) chemical. Such coupled reactions are referred to as redox reactions, The metabolic processes *glycolysis*, *Krebs cycle*, and *Electron transport phosphorylation* involve the transfer of electrons (as varying energy states) by redox reactions.

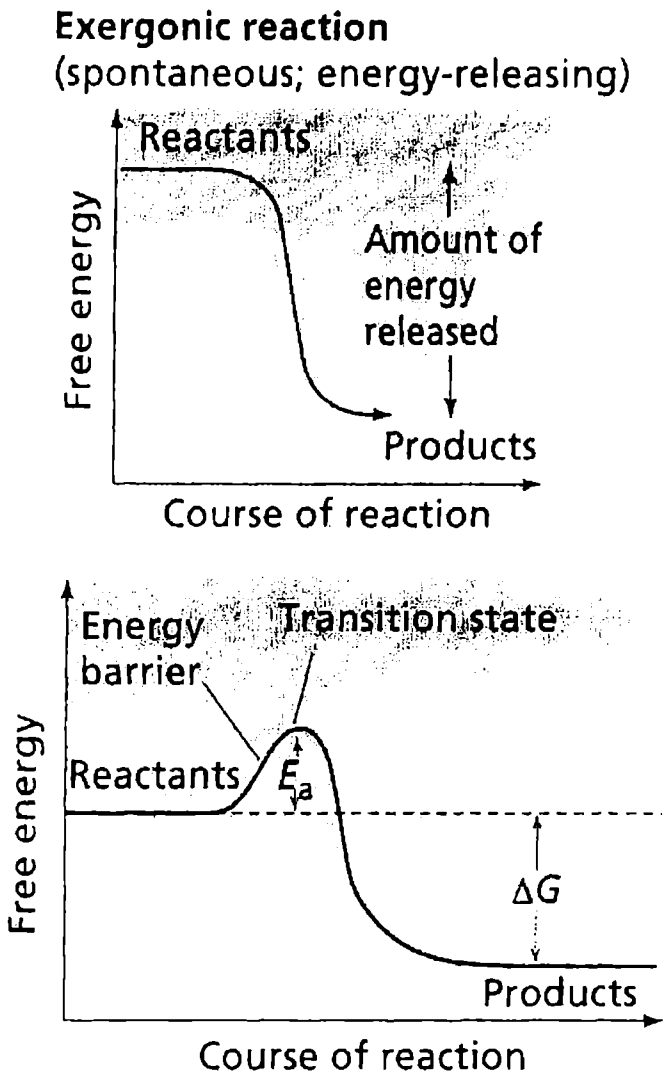


Figure 10: Time-energy graphs of an exergonic reaction (top) and endergonic reaction (bottom)

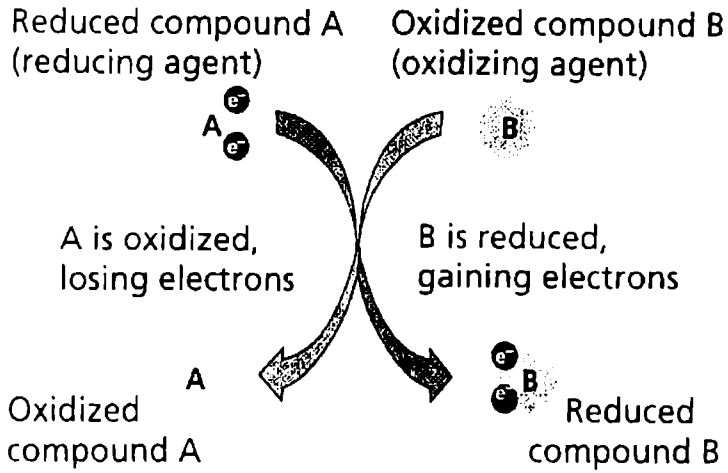


Figure 11: Passage of electrons from compound A to B, when A loses its electrons it is oxidized; when B gains the electrons it is reduced.

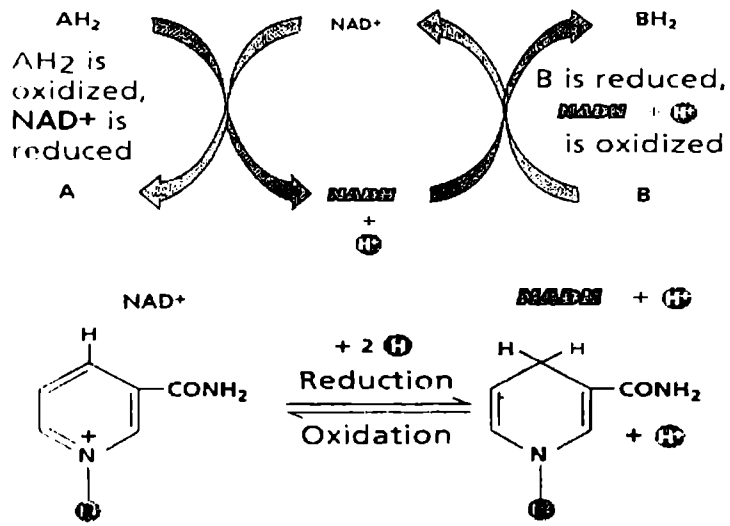


Figure 12: Oxidation/reduction reaction via an intermediary (energy carrier) compound.

CATABOLISM AND ANABOLISM

Anabolism is the total series *of* chemical reactions involved n synthesis of organic compounds. Autotrophs must be able to nanufacture (synthesize) all the organic compounds they need. Heterotrophs can obtain some of their compounds in their diet along with their energy). For example humans can synthesize 12 of the 20 amino acids, we must obtain the other 8 in our diet. *Catabolism* is the series of chemical reactions that breakdown arger molecules. Energy is released this way, some of it can be ıtilized for anabolism. Products of catabolism can be reassembled by anabolic processes into new anabolic molecules.

ENZYMES (ORGANIC CATALYSTS)

Enzymes allow many chemical reactions to occur within the ıomeostasis constraints of a living system. Enzymes function as ırganic catalysts. A catalyst is a chemical involved in, but not hanged by, a chemical reaction. Many enzymes function by owering the *activation energy* of the reaction. By bringing the eactants closer together, chemical bonds may be weakened and eactions will proceed faster than without the catalyst Figure 13).

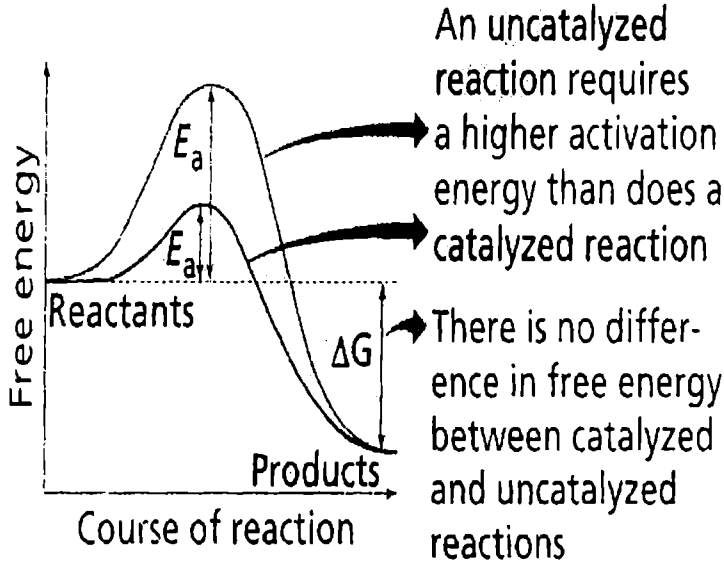


Figure 13: The use of enzymes can lower the activation energy of a reaction E_a.

Enzymes can act rapidly, as in the case of Carbonic anhydrase (enzymes typically end in the –ase suffix), which causes the chemicals to react 107 times faster than without the enzyme present. Carbonic anhydrase speeds up the transfer of Carbon dioxide from cells to the blood, There are over 2000 known enzymes, each of which is involved with one specific chemical reaction. Enzymes are substrate specific, the enzyme peptidase (which breaks peptide bonds in proteins) will not work on starch (which is broken down by human-produced amylase in the mouth).

Enzymes are proteins, the functioning of the enzymes is determined by the shape of the protein. The arrangement of molecules on the enzyme produces an area known as the active site within which the specific substrate (s) will "fit". It recognizes, confines and orients the substrate in a particular direction.

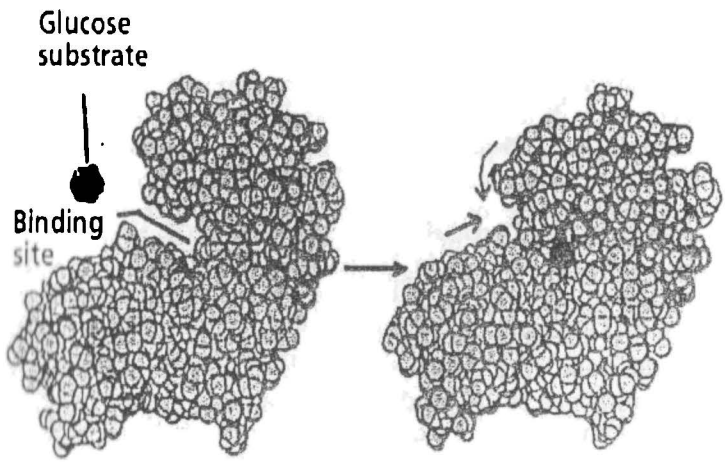


Figure 14: Space filling model of an enzyme working on glucose. Note the shape change in the enzyme (indicated by the arrows) after glucose has fit into the binding or active site.

The induced fit hypothesis suggests that the binding of the substrate to the enzyme alters the structure of the enzyme, placing some strain on the substrate and further facilitating the reaction. Cofactors are nonproteins essential for enzyme activity. Ions such

as K^+ and Ca^{2+} are cofactors. Coenzymes are nonprotein organic molecules bound to enzymes near the active site NAD (nicotinamide adenine dinucleotide).

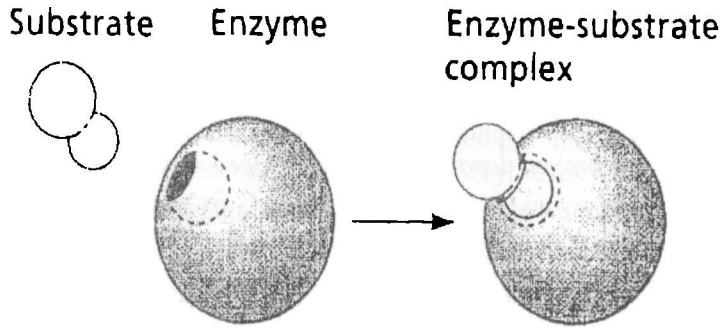


Figure 15: The formation of an enzyme-substrate complex.

Enzymatic pathways form as a result of the common occurrence of a series of dependent chemical reactions. In one example, the end product depends on the successful completion of five reactions, each mediated by a specific enzyme, the enzymes in a series can be located adjacent to each other (in an organelle or in the membrane of an organelle), thus speeding the reaction process. Also, intermediate products tend not to accumulate, making the process more efficient. By removing intermediates (and by inference end products) from the reactive pathways, equilibrium (the tendency of reactions to reverse when concentrations of the products build up to a certain level) effects are minimized, since equilibrium is not attained, and so the reactions will proceed in the "preferred" direction (Figure 16).

TEMPERATURE

Increases in temperature will speed up the rate of non-enzyme mediated reactions, and so temperature increase speeds up enzyme mediated reactions, but only to a point. When heated too much, enzymes (since they are proteins dependent on their shape) become denatured. When the temperature drops, the enzyme regains its shape. Thermolabile enzymes, such as those responsible for the colour distribution in Siamese cats and colour

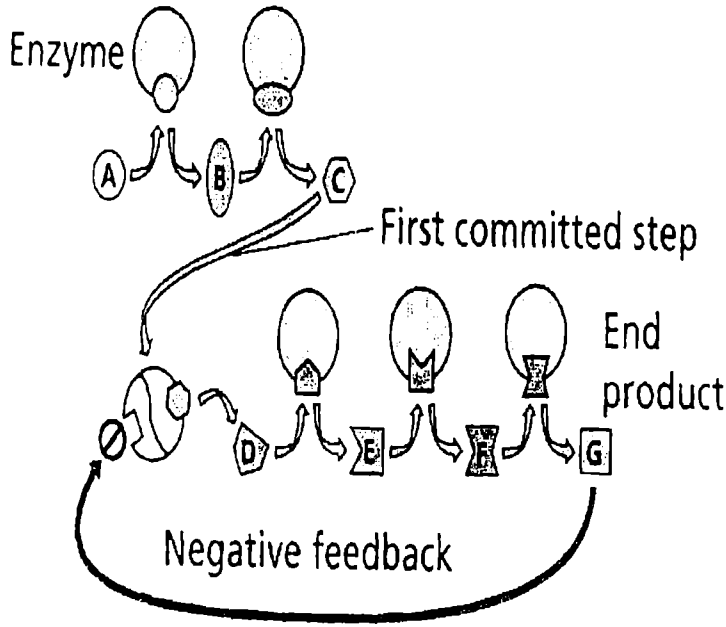


Figure 16: Negative feedback and a metabolic pathway. The production of the end product (G) in sufficient quantity to fill the square feedback slot in the enzyme will turn off this pathway between step C and D.

camouflage of the Arctic fox, work better (or work at all) at lower temperatures.

CONCENTRATION OF SUBSTRATE AND PRODUCT

The concentration of substrate and product also control the rate of reaction, providing a biofeedback mechanism

Activation

As in the case of chymotrypsin, activation protects a cell from the hazards or damage the enzyme might cause.

CHANGE IN PH

Change in pH will also denature the enzyme by changing the shape of the enzyme. Enzymes are also adapted to operate at a specific pH or pH range (Fig. 17).

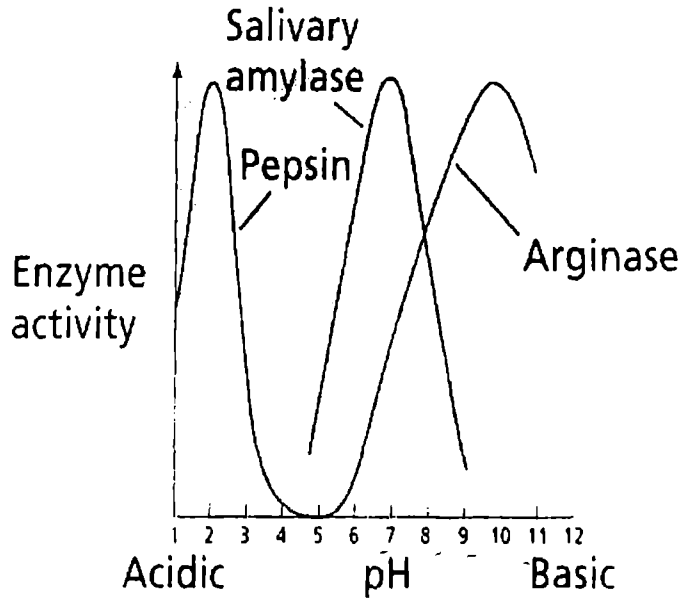


Figure 17: Enzyme activity as a function of pH for several enzymes. Each enzyme has a range of pH at which it is most active.

ALLOSTERIC INTERACTIONS

Allosteric interactions may allow an enzyme to be temporarily inactivated. Binding of an allosteric effector changes the shape of the enzyme, inactivating it while the effector is still bound. Such a mechanism is commonly employed in feedback inhibition, often one of the products, either an end or near-end product act as an allosteric effector, blocking or shunting the pathway (Figure 18).

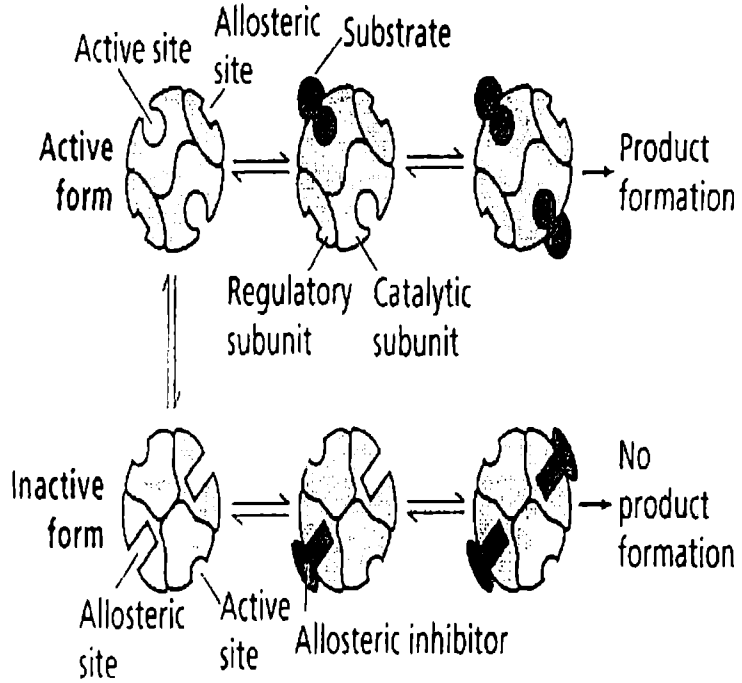


Fig. 18: Action of an allosteric inhibitor control on the action of an enzyme.

COMPETITIVE INHIBITION

Competitive inhibition works by the competition of the regulatory compound and substrate for the binding site, if enough regulatory compound molecules bind to enough enzymes, the pathway is shut or atleast slowed down, PABA, a chemical essential to a bacteria that infects animals, resembles a drug, sulfanilamide, that competes with PABA, shutting down an essential bacterial (but not animal) pathway (Figure 19).

NONCOMPETITIVE INHIBITION

Noncompetitive inhibition occurs when the inhibitory chemical, which does not have to reasonable the substrate, binds to the enzyme other than at the active site. Lead binds to SH

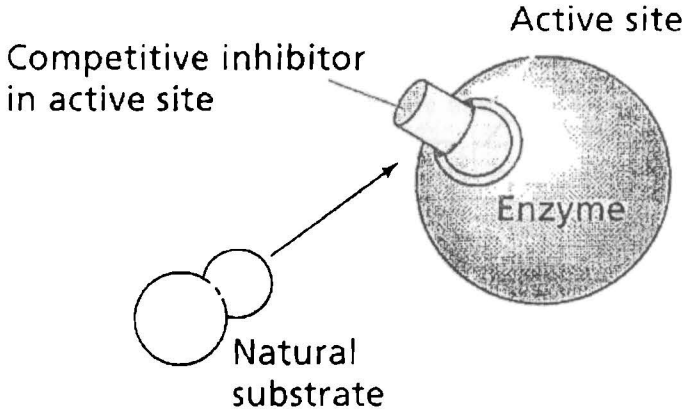


Figure 19: Competitive inhibition in the active site of an enzyme normally occupied by the natural substrate.

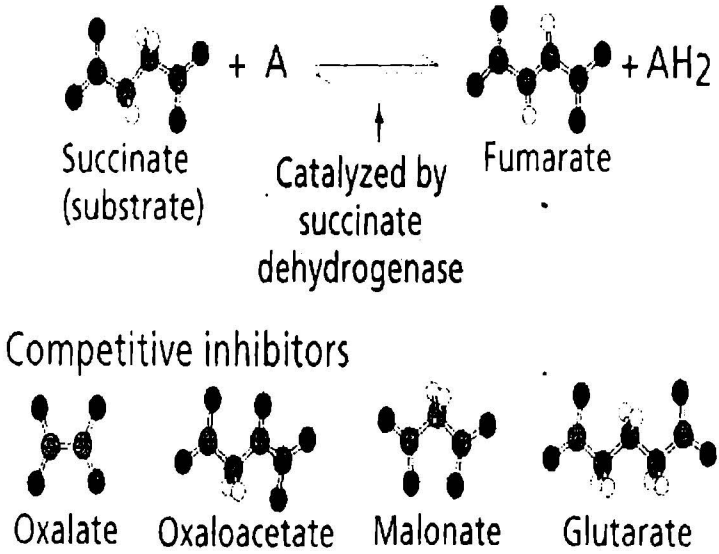


Figure 20: Specific case of succinate dehydrogenase and its natural substrate (succinate) and competitors (oxalate etc).

groups in this fashion. Irreversible inhibition occurs when the chemical either permanently binds to or massively denatures the enzyme so that the tertiary structure cannot be restored. Nerve gas permanently blocks pathways involved in nerve message transmission, resulting in death. Penicillin, the first of the "wonder drug" antibiotics, permanently blocks the pathways certain bacteria use to assemble their cell wall components.

BIOENERGETICS AND BIOCOMMUNICATION

Organisms are so enigmatic from the thermodynamic point of view that Lord Kelvin, co-inventor of the Second Law of thermodynamics, specifically excluded them from its dominion (Ehrenberg, 1967), while Schrodinger (1944) suggested they feed upon "negative entropy" to free themselves from all the entropy they cannot help producing.

Lord Kelvin was impressed with how organisms seem to have energy at will, whenever required, and in a perfectly coordinated way. That is at once the problem of bioenergetics—how organisms can have energy so readily—and of biocommunication—how the energy mobilizing activities are organised as a whole. Similarly, Schrodinger alluded to the ability of organisms to use the energy they feed on to build up and maintain their dynamic organization. The intuition of both physicists is that energy and organization are intimately linked.

Schrodinger was reprimanded by Linus Pauling and others, for using the term "negative entropy", which does not correspond to any rigorous thermodynamic entity (Gnaiger, 1994). However, the idea that open systems can "self-organise" under energy flow became more concrete in the discovery of *dissipative structures* (Prigogine, 1967), that depend on the flow and dissipation of energy, such as the Benard convection cells and the laser. In both cases, energy input results in a phase transition to global dynamic order in which all the molecules or atoms in the system move coherently. From these considerations negative entropy of Schrodinger have been identified as "stored mobilizable energy in a space time structured system" (Ho, 1994 b, 19952).

ENERGY STORAGE AND, MOBILIZATION IN LIVING SYSTEMS

The key to understanding the thermodynamics of the living systems is not energy flow or energy dissipation, but the energy storage under energy flow. Energy flow is of no consequence unless the energy is trapped and stored within the system where it circles before being dissipated. A reproducing life cycle, i.e., an organism/arises when the loop of circulating energy closes. At that point, we have a life cycle within which the stored energy is mobilized, remaining stored as it is mobilized, and coupled to the energy flow.

$$Ji + \sum kLik = Xk2 \quad (1)$$

Where Ji is the flow of the ith process (i = 1,2,3...n), xk is the kth thermodynamic force (k = 1,2,3.....n), and Lik are the proportionality coefficients (where i =k) and coupling coefficients (where i=k), the couplings for which the xks are invariant with time reversal (i.e., velocity reversal) will be symmetrical; in other words,

$$Lik = Lki \quad (2)$$

so long as the Js and the Xs satisfy $\tau g = SJiXi$ where g is the rate of entropy increase per unit volume.

Morowitz theorem states that the flow of energy through the system from a source to a sink will lead to atleast one cycle in the system at steady state, *provided that the energy is trapped and stored within the system,* This important theorem captures a key aspect of the steady state, and implies that title steady state – at which global balance in maintained– may harbour non-linear processes (Ho, 1993).

Onsagar's reciprocity relationship has been extended to the far from equilibrium regime: "for multienzyme systems and more recently, by Sewell (1991) for infinite quantum systems. However, the validity and the theoretical basis for the extension of Onsager's reciprocity relationship to biological systems are still under debate show that while linear non-equilibrium thermodynamics can describe an autocatalytic system, the matrix of phenomenological coefficients is nonsymmetrical. They conclude therefore, that it is the symmetry property (Onsager's reciprocity relationshfp) and

not the linearity of the flow-force relations in the near equilibrium domain the precludes oscillations; and conversely, a system with oscillations cannot at the same time satisfy the symmetry property.

Ho (1996) believe some form of Onsager's reciprocity relationship does hold in living systems if only to account for the ready mobilization of energy on the one hand– why we can have energy at will - and on the other hand, for the linear relationships between steady-state flows and conjugate thermodynamic forces outside the range of equilibrium actually observed in many biological systems.

According to Rothschild et al (1980) linearity in biological processes can arise in enzymes operating near a multidimensional inflection point far away from thermodynamic equilibrium, if some of the rate constants are linked. That is realistic for living systems which are now known to have highly organized flows in the cytoplasmic matrix (Welch, 1985). Sewell shows how Onsager's reciprocity relationship applies to locally linearized combinations of forces and flows, which nonetheless behave globally in non-linear fashion. That is particularly relevant to the living system, where nested compartments and microcompartments ensure that many processes may be operating locally at thermodynamic equilibria even though the system or subsystem as a whole is far away from equilibrium (Ho, 1995 a). Furthermore, as each process is ultimately connected to every other in the metabolic net through catenation of space and time, even if truly symmetrical couplings are localized to a limited number of metabolic/energy transducing junctions, the effects will eventually be shared or delocalized throughout the system, so that symmetry will apply to appropriate combinations of forces and flows over a sufficiently macroscopic space time scale (Sewell, 1991). That is perhaps the most important consideration. As real processes take time. Onsager's reciprocity relationship cannot be true for an arbitrarily short instant, but must apply at a sufficiently macroscopic time interval when overall balance holds.

To summarize, non-linearity does not preclude symmetry on the appropriate scale, and local linearity does not exclude the possibility for self-organization at a more global level. Hence the contension that oscillations typical of self-organized systems is incompatible with symmetry properties, may be irrelevant when

the entire system or subsystem of balanced flows and forces is taken into account. This will become clear as we consider the origins of the thermodynamics of the steady state.

THERMODYNAMICS OF THE STEADY STATE VS THERMODYNAMICS OF ORGANISED COMPLEXITY

Denbigh (1951) defines the steady state as one in which "the macroscopic parameters, such as temperature, pressure and composition, have time independent values at every point of the system, despite the occurrence of a dissipative process". That is too restrictive to apply to the living system, which has coupled processes spanning the whole gamut of relaxation times and volumes (Ho, 1993). A less restrictive formulation—one consistent with a "thermodynamics of organized complexity" (Ho, 1994 a) – might be to define the living system, to first approximation, as a *dynamic equilibrium in which the macroscopic parameters, such as temperature, pressure and composition have time–independent values despite the occurrence of dissipative processes.* The present formulation omits the phrase, "at every point of the system". Microscopic homogeneity is not crucial for the formulation of any thermodynamic state, as the thermodynamic parameters are macroscopic, entities quite independent of the microscopic interpretation (Ho, 1993). Like the principle of microscopic reversibility, it is extraneous to the phenomenological laws of thermodynamics (Denbigh, 1951).

The first incursion into the thermodynamics of the steady state was W. Thomson's (Lord Kelvin) treatment of the thermoelectric effect (Denbigh, 1951). This involves a circuit in which heat is absorbed and rejected at two junctions (the peltier heat), and in addition, heat is absorbed and given off (the Thomson heat) due to current flows between two parts of the same metal at different temperatures. Both of these heat effects are reversible, in that they change sign when the direction of the current is reversed, on the other hand, there are two other effects which are not reversible- heat conduction along the wires and dissipation due to the resistance, it is thus impossible to devise a reversible thermoelectric circuit even in principle. Nevertheless, Thomson took the step of assuming that, at steady state, those heat effects

that are reversible i.e. the peltier heat and Thomson heat balance each other so that no net entropy is generated,

$$\Delta Sp + \Delta ST = 0$$

On that basis, he derived the well-known relations between the Peltier and Thomson heats and the temperature coefficient of the electromotive force, it was a bold new departure in the application of the Second Law, but one which was subsequently justified by experimental evidence.

Very similar methods were used by Helmholz in his treatment of the electromotive force and transport in the concentration cell, where he states clearly that the two irreversible process in the cell, heating and diffusion, are to be disregarded and the Second law to be applied to those parts of the total process which are reversible, A virtual flow of current is supposed to take across the liquid junction, resulting in a displacement of the ions. The process is taken to be reversible and to generate no entropy. The justification, according to Guggenheim (cited in Denbigh, 1951), is that the two processes, diffusion and flow of current the junction, "take place at rates which vary according to different laws" when the composition gradient across the boundary is altered, and so it seems reasonable to suppose that the two processes are merely superposed, and that the one may be ignored when considering the other; Thus, the steady state is treated *as if there were no dissipative processes,* and it is this assumption which is later validated by Onsager's reciprocity relationship.

The living system is free from immediate thermodynamic constraints.

The living system have been treated as a superposition of dissipative irreversible processes and non-dissipative processes, so that Onsager's reciprocity relationship applies only to the latter, in other words, it applies to coupled processes for which the net entropy production is balanced or zero,

$$Sk\ DSk = 0 \quad (3)$$

This will include most living processes because of the coupled cycles, for which the net entropy production balances out to zero, the principle applies in the smallest unit cycle in the living system

- enzyme catalysis - on which all energy transduction in the living system is absolutely dependent. Over the past 30 years, Lumry and his coworkers have shown convincingly how the flexible enzyme molecule balances out entropy with enthalpy to conserve free energy during catalysis. The organism is, in effect, a closed, self-sufficient energetic domain of cyclic non-dissipative processes coupled to the dissipative processes (Ho, 1995 b). In the formalism of conventional thermodynamics, the life cycle, or more precisely, the living system in dynamic equilibrium, consists of all cyclic processes for which the net entropy change is zero, coupled to dissipative processes necessary to keep it going, for which the net entropy change is greater than zero.

Consequently, the organism is free from the immediate constraints of energy conservation–the First Law and the Second Law of thermodynamics. *Their is always energy; available within the system, which is mobilized at close to maximum efficiency and over all space-time modes.* This in turn creates the conditions for rapid, sensitive and specific intercommunication throughout the system.

ENERGY SELF-SUFFICIENCY AND SENSITIVITY

The distinguishing feature of the living system is its exquisite sensitivity to weak signals. For example, the eye can detect single photons falling on the retina, and the presence of several molecules of pheromones in the air is sufficient to attract male insects to their mates. That exquisite sensitivity applies to all levels of 'information processing' in the organism, and is the direct consequence of its energy self-sufficiency. No part of the system has to be pushed or pulled into action, nor be subjected to mechanical regulation and control. Instead, coordinated action of all the parts depends on rapid *intercommunication* throughout the system. The organism is a system of "excitable media" (Goodwin, 1994; 1995), or excitable cells and tissues poised to respond specifically and disproportionately to weak signals because the large amount of energy stored can amplify weak signals into macroscopic actions. It is by virtue of its energy self-sufficiency, therefore, that an organism is a *sentient* being – a system of sensitive parts all set to intercommunicate, to respond and to act appropriately as a whole to any contingency.

THE POLYCHROMATIC WHOLE

Evidence for constant intercommunication throughout the living system may already exist in the physiological literature. The *deterministic chaos* which has been used to describe many living functions from the complex, locally unpredictable patterns of brain activities (Freeman, 1995). A different understanding of the complex activity spectrum of the healthy state is that it is polychromatic, approaching "white" in the ideal, in which all the modes of stored energy are equally represented. It corresponds to the so-called f(1) = const rule that Fritz Popp (1986) has generalised from the spectrum of light or "biophotones" found to be emitted from all open systems capable of energy storage naturally evolve (Ho, 1994 b). It is a state of both maximum and minimum entropy—maximum because energy is equally distributed over all space-time modes; and minimum because the modes are coupled together to give, in effect, a single degree of freedom (Popp, 1986; Ho, 1993). In a system with no impedance to energy mobilization, all the modes are intercommunicating and hence all frequencies are represented,. But when coupling is imperfect, or when the sub-system, say, the heart, or the brain, is not communicating properly, it falls back on its own modes, leading to impoverishment of its activity spectrum. The living system is necessarily a polychromatic whole, it is full of variegated complexity that nevertheless cohere into a singular being, and that is the ultimate problem of biocommunication that needs to be addressed.

Advances in biochemistry, cell biology and genetics are giving us a concrete picture of the organism as an interconnected, intercommunicating whole. It is becoming increasingly clear that living organism cannot be understood in terms of mechanistic controls, nor of endless processings of genetic information.

A MOLECULAR DEMOCRACY OF DISTRIBUTED CONTROL

Henrik Kacser, (1987) was among the first to realize that in a network, especially one as complicated as the metabolic network, it is unrealistic to think that there could be special enzymes controlling the flow of metabolites under all circumstances. He

pioneered *metabolic control analysis* to discover how the network is actually regulated. After more than 20 years of investigation, it is now generally acknowledged that so called 'control' is invariably distributed over many enzymes (and metabolites) in the network, and moreover, the distribution of control differs under different conditions. The metabolic network turns out to be a "molecular democracy" of distributed control.

LONG RANGE ENERGY CONTINUA IN CELLS AND TISSUES

Studies over the past 25 years have also revealed that energy mobilization in living systems is achieved by protein or enzyme molecules acting as "flexible molecular energy machines" (Ho, 1995 a) transferring energy directly from the point of release to the point of utilization, without thermalization or dissipation.

These direct energy transfers are carried out in collective modes extending from the molecular to the macroscopic domain. The flow of metabolites is channelled coherently at the molecular level, directly from one enzyme to the next in sequence, in multienzyme complexes (Welch and Clegg, 1987). At the same time, high voltage electron microscopy and other physical measurement techniques reveal that the cell is more like a 'solid state' than the 'bag of dissolved enzymes, that generations of biochemists had previously supposed (Clegg, 1984). Not only are almost all enzymes bound to an intricate "microtrabecular lattice", but a large proportion of metabolites as well as water molecules are also structured on the enormous surfaces available. Aqueous channels may be involved in the active transport of solutes within the cell in the way that the blood stream transports metabolites and chemical messengers within the organism (Wheatley and Clegg, 1991).

As Welch and Berry (1985) proposed, the whole cell is linked by "long-range energy continua of mechanical interactions, electric and electrochemical fluxes and in particular, proton currents that form a "protoneural network", whereby metabolism is regulated instantly and down to minute detail. Cells are in turn interconnected by electrical and other cytoplasmic junctions. There is increasing

evidence that cells and tissues are also linked by electromagnetic photons and protons (popp, Li and Gu, 1992; Ho, 1993; Ho, Popp & Warnke, 1994). The cell as well as organism is not so much a "solid state" as liquid crystalline. Living systems, therefore, possess just the conditions favouring the rapid propagation of influences or 'information in all directions, which are naturally elated in cascades (Ho, 1993) by the relaxation space-times of the processes involved. These are precisely the conditions that can yield linear flow force relationship in a system globally far from thermodynamic equilibrium? Global phase transitions may often take place, which can be initiated at any point within the system or subsystem. Abrupt, phase transition like changes in the electrical activities of whole area of the brain are indeed frequently observed in simultaneous recordings with a large array of electrodes (Freeman, 1975), for which no definite centre (s) of origin can be identified.

ORGANISM AND ENVIRONMENT - A MUTUAL PARTNERSHIP

Biology today remains dominated by the genetic paradigm. The genome is seen as the repository of genetic information controlling the development of the organism, but otherwise isolated from the environment, and passed on unchanged to the next generation except for rare random mutations. The much publicised Human Genome Project is being promoted on that basis (Ho, 1995e). The genetic paradigm has already been fatally undermined at least ten years ago, when a plethora of 'fluid genome' processes were first discovered, and many more have come to light since. These processes destabilize and alter genes and genomes in the course of development, some of the genetic changes are so well correlated with the environment that they are referred to as "directed mutations". Many of the genetic changes are passed on to the next generation. Heredity can no longer be seen to reside solely in the DNA passed on from one generation to the next. Instead, the stability and repeatability of development - which we recognize as heredity - is distributed in the whole gamut of dynamic feedback interrelationships between organism and environment from the socio-ecological to the genetic. All of these may leave imprints that are passed on to the subsequent generations; as cultural traditions or artefacts, maternal or cytoplasmic effects,

gene expression states, as well as genetic (DNA sequence) changes (Ho, 1386; 1896).

THE DISTRIBUTED ORGANIC WHOLE

The essence of the organic whole Is that it is distributed throughout its constituent parts, with no centre of control, no governors, no hierarchical levels of line-managers or regulators processing information down the line of command. Instead, pervasive, moment to moment intercommunication throughout the system renders part and whole, local and global completely indistinguishable. The existing mechanistic framework is most inadequate in coming to grips with the organic whole.

THE COHERENCE OF ORGANISM

The living system is necessarily a polychromatic whole — a variegated complexity that nevertheless cohere into a singular being.. The wholeness of the organism is the ultimate problem of biocommunication how to account for the continuity that encompasses the activities of elementary particles and atoms, molecules and cells, tissues and organs all the way to the organism itself (Joseph Needham, 1936). The problem has never been adequately addressed until Herbert Frohlich (1968, 1980) presented the first detailed theory of *coherence.* He argued that as organisms are made up of strongly dipolar molecules packed rather densely together, electric and elastic forces will constantly interact. Metabolic pumping will excite macromolecules such as proteins and nucleic acids as well as cellular membranes (which typically have an enormous electric field of some 10^7 V/m across them). These will start to vibrate and eventually build up into collective modes, or *coherent excitations,* of both photones and photons extending over macroscopic distances within, and perhaps also outside the organism.

The emission of electromagnetic radiation from coherent lattice vibrations in a solid-state semiconductor has been experimentally demonstrated for the first time (Dekorsy et al, 1295). The possibility that organisms may use electromagnetic radiations to communicate between cells was already entertained by Soviet biologist Gurwitsch (1925). This hypothesis was revived by Popp and his coworkers in the late 1970s, and there is now

a large and rapidly growing literature on "biophotons" believed to be emitted from a coherent photon field within the living system (Popp, Li and Gu, 1992). It has been found that a single, brief exposure of synchronously developing early fruitfly embryos to white light results in the re-emission of relatively intense and prolonged flashes of light, some tens of minutes and even hours after the light exposure (Ho et al, 1992 b). The phenomenon is reminiscent of phase-correlated collective emission, or *super-radiance,* in atomic systems, although the time scale orders of magnitude longer, perhaps in keeping with the coherence times of organisms. For phase correlation to build up over the entire population, one must assume that each embryo has a *collective phase* of all its activities, in other at fords, each embryo must be considered a highly coherent domain, despite its multiplicity of activities (Ho, Zhou and Haffegee, 1995).

During the same period of early development, exposure of the embryos to weak static magnetic fields also cause characteristic global transformation of the normal segmental body pattern to helical configurations in the larvae emerging 24 hours later (Ho et al, 1992 a). As the energies involved are several order of magnitude below the thermal threshold, we conclude that there can be no effect unless the external field is acting on a coherent domain where charges are moving in phase, or where magnetically sensitive liquid crystals are undergoing phase alignment globally (Ho et al, 1994). Liquid crystals may indeed be the material basis of many, if not all aspects of biological organisation (Ho et al., 1595).

ORGANISMS AS POLYPHASIC LIQUID CRYSTALS

Liquid crystals are phases of matter between the solid and the liquid states, hence the term *masophases* (De Cannes, 1974). Liquid crystalline mesophases possess long range orientational order, and often also varying degrees of translational order, in contrast to solid crystals, liquid crystals are mobile and flexible, and above all, highly responsive, they undergo rapid changes in orientation or phase transitions when exposed to electric and magnetic fields (Binov, 1983) or to changes in temperature, pressure, pH, hydration, and concentrations of inorganic ions (Collings,

1990; Knight, 1993). These properties are ideal for organisms (Gray, 1993; Knight, 1993). Liquid crystals in organisms include all the major constituents of the organism: the amphiphilic lipids of cellular membranes, the DNA in chromosomes, all proteins, especially cytoskeletal proteins, muscle proteins, collagens and proteohlycans of connective tissues. These adopt a multiplicity of meso-phases that may be crucial for biological structure and function at all levels of organisation (Ho et al, 1995) from channelling metabolites in the cell to pattern determination and the coordinated locomotion of whole organisms.

The importance of lipid crystals for living organisation was recognized by Joseph Needham (1936) among others. He suggested that living systems actually are liquid crystals, and that many liquid crystalline mesophases may exist in. The cell although they cannot be detected. Indeed, there has been no direct evidence that extensive liquid crystalline mesophases exist in living organisms or in the cytoplasm until our recent discovery of a non-invasive optical technique (Ho and Lawrence, 1993; Ho and Saunders, 1994; Newton, Haffegee and Ho, 1995). This enables us to obtain high resolution and high contrast coloured images of live organisms based on visualising just the kind of coherent liquid crystalline mesophases which Needham and others had predicted.

The technique amplifies small birefringences typical of biological liquid crystals, enabling us to see the whole living organism down to the phase alignment of the molecules that make up its tissues. Brilliant interference colours are generated, specific for each tissue, dependent on the birefringence of the molecules and their degree of coherence phase alignment. The colours are generated even as the molecules in the tissues are moving about, busily transforming energy, That is possible because visible light vibrates much faster than the molecules can move, so the tissues will appear indistinguishable from static crystals to the light passing through so long as the movements of the constituent molecules are sufficiently coherent. With this imaging technique, one can see that the organism is *thick with activities at every level, coordinated in a continuum from the macroscopic to the molecular.* And that is what the coherence of the organism entails.

These images also bring out another aspect of the wholeness of the organism: all organisms, from protozoa to vertebrates without exception, are polarized along the anteroposterior axis, so that all the colours in different parts of the body are maximum when the anteroposterior axis is appropriately aligned, and they change in concert as the organism is rotated from that position. The anteroposterior axis is the optical axis of the whole organism, which is, in effect, a single (uniaxial) crystal. This leaves us in little doubt that the organism is a singular whole, despite the diverse multiplicity and polychromatic nature of its constituent parts.,

The tissues not only maintain their crystalline order when they are actively transforming energy, the degree of order seems to depend on energy transformation, in that the more active and energetic the organism, the more intensely colourful it is, implying that the molecular motions are all the more coherent (Ho and Saunders, 1994). The coherence of the organism is closely tied up with its energetic status, as argued in the beginning : energy and organisation are intimately linked. The coherent whole is full of energy – it is a vibrant coherent whole.

QUANTUM COHERENCE IN LIVING ORGANISMS

The above considerations and observations convince that the wholeness of organisms is only fully captured by quantum coherence (Ho, 1993). An intuitive way to understand quantum coherence is to think of the 'I' that each and every one of us experience of our own being, we know that our body is a multiplicity of organs and tissues, composed of many billions of cells and astronomical numbers of molecules of many different kinds, all capable of working automatically, and yet somehow cohering into the singular being of our private experience. That is just the stuff of quantum coherence. Quantum coherence does not mean that everybody or every element of the system must be doing the same thing all the time, it is more akin to a grand ballet, or better yet, a very large jass band where everyone is doing his or her own thing while being perfectly in step and in time with the whole,

A quantum coherent system maximizes both global cohesion and local freedom (Ho, 1993). This property, technically referred

to as factorisability, enables the body to be performing all sorts of different coordinated functions simultaneously (Ho, 1995 b). It also enables *instantaneous (nonlocal) and noiseless intercommunication to take place throughout the system.* For example, the digestive system works independently, the metabolism dully transforms chemical energy in cells, stores fats and glycogen, converts most of it into readily utilizable forms as ATP. Similarly, muscles keep in tone and allow to work on keyboard, neurons fire in coherent patterns in the brain.

The coherent organism is, in the ideal, a quantum superposition of activities – organised according to their characteristic space-times – each itself coherent, so that it can couple coherently to the rest (Ho, 1995 b), It is, in effect a vast array of Frohlich systems all coupled together. This picture is fully consistent with the earlier proposal that the organism stores energy over all space-time domains each intercommunicating with the rest. It is also consistent with Onsager's reciprocity relationship or symmetrical coupling between all energy modes. Furthermore, quantum superposition enables the system to maximise its potential degrees of freedom so that the single degree of freedom required for coherent action can be instantaneously accessed.

The main implication of quantum coherence for living organization is that, in maximizing both local freedom and global intercommunication, the organism is in a very real sense completely free. Nothing is in control, and yet every thing is in control. Thus, it is the failure to transcend the mechanistic framework that makes people persist in enquiring which parts are in control, or issuing instructions or information. These questions are meaningless when one understands what is to be a coherent, organic whole, global and local are so thoroughly implicated as to be indistinguishable, and where each part is as much in control as it is sensitive and responsive. The challenge is to rethink information processing in the context of the coherent organic whole.

BIOENERGETICS AND THE DIMENSIONALITY OF ECOSYSTEMS

A major topic in ecology is the understanding of the emergence of dynamics of diversity in ecosystems. Cascades of evolutionary

branching in a community are a possible explanation for an increase in diversity, but maximum diversity is bound by the dimentionality of the environment (Meszena & Metz, 1999).

Since the dimentionality of the abiotic environment is exogenous to the basic community, a persistent increase in diversity requires an increase in dimensionality in the biotic environment. A simultaneous increase in diversity and dimensionality can mean the branching (Geritz et al, 1997) of a primitive population into two populations, genetically and functionally isolated, such that a new resource has appeared.

BIOENERGETICS AND NATURAL SELECTION

An individual is fully characterised through initial conditions and a set of two dynamic equations:

$$\frac{dl}{dt} = f(l, e, X, \phi, \psi)$$

$$\frac{de}{dt} = g(l, e, X, \phi, \psi)$$

The state variables are l and e, respectively scaled volumetric length and scaled reserve density. X stands for the generalized resource, Φ stands for the set of traits of the individual and χ for the set of parameters of the environment (such as temperature or salinity, i.e., all properties of the environment that are not *X).*

ϕ is the set of traits or "meta-parameters" of which the true parameters are a function of. For example, extensive parameters are linearly dependent on maximal body size while intensive parameters are independent. The functional response terms (saturation coefficient and maximal ingestion rate) co-vary in a special relation, etc. The existence of these constraints (Kooijman, 2000,) reduces the set of traits to a number smaller than the set of parameters.

A structured population (Kooijman, 2000,) can be characterized by p, the number of individuals in an infinitesimal interval in the phase-space of l, e, and a, age, and by the total size of the population, N:

$$\frac{dl}{dt}\ p(l,e,a) = f(l,e,a,X,\phi\ \psi)$$

$$\frac{de}{dt}\ (\chi,\phi,\psi) = \int\rho\ dl\ de\ da$$

To obtain these equations explicitly it is necessary to solve the McKendrick equations, which is by no means trivial (Gurney and Nisbet, 1998). But for the purpose of this essay we assume that these equations exist.

Until now we considered that the population was physiologically structured but that all its individuals shared the same set of parameters, ϕ. Ignoring the specific mechanism of reproduction and assuming that mutations are rare and small, natural selection can be introduced as follows.

Let there be a resident population, *x*, exhibiting parameter or trait set Φ_x. Let there appear, by mutation, an invading population, *y*, exhibiting trait set Φ_y. In a first moment, the population size of the invader is small and so the environment is fully characterized by the resident population. If the growth rate of the invader in this environment is smaller than that of the resident, the invader cannot thrive and the invader trait goes extinct. If it is larger, than the invader grows and two situations are possible. When population *x* becomes rare, it cannot grow in the environment set by *y*, *x* becomes extinct and *y* establishes itself as the new resident. Otherwise, if when *x* is rare it can grow in the environment set by *y*, then both traits can be mutually invaded and both will survive, evolutionary branching has occurred (Geritz et al., 1997).

DISCRETE DIMENSION OF A SMALL COMMUNITY

Consider a population of Vl-morphs, with 1 life stage, such that the population dynamics are simply the dynamics of an individual scaled to the population size. Consider a spatially homogeneous environment closed to mass transfer and with input of one form of energy (e.g., light). Consider furthermore that the population is composed of mixotrophs, that can alternatively process the external inflow of energy, be heterotrophs (cannibals

in this case) or decomposers, converting dead organic matter into non-organic materials (Mulder et al., 2001).

Without natural selection, a functional community may arise (Kooijman, 2000, exhibiting cyclical fluctuations, since originally autotrophs thrive but as inorganic matter is exhausted growth must halt. Heterotrophism and decomposition become feasible, when organic mass is abundant. Hence, under some conditions, defined by the parameters, it is possible that a three-stage community arises. Each stage processes a different type of resource (environmental dimension). Autotrophy is supported by the abiotic environment (either the external energy source or inorganic matter) and decomposition is supported by the first biotic dimension (biomass). The number of heterotrophic levels depends on the number of different preys, which is one in the absence of natural selection, since all individuals are identical.

Assume that different individuals take different specializations and do not change it throughout their lives. Natural selection, through trait change, turns functional diversity into genetic divergence, since the traits that are more advantageous for each specialization are different. Loss of metabolic plasticity is advantageous for prokaryotes since it allows for a reduction in DNA size and hence an increase in population growth rate, which favours specialization, if the environment is stable (Kooijman, 2000). The emergence of genetically isolated populations allows in turn the coevolution of preys and predators, leading to stabilization of traits (Doebeli and Dieckmann, 2000). Thus, specialization of several consumer levels is expectable.

Specialization at the bottom and at the top of the ecosystem are not so plausible. The energy source is assumed to be unique and even though the decomposers can also specialise in structure or reserves, or the corpses of some particular "species", structural homeostasis implies that species that are genetically similar are also biochemically similar (Kooijman, 2000) - hence leaving little room to specialization.

Assume that each evolutionary branching occurs at a time, since mutations are rare. Then the adaptive dynamics of a new trait y in a population x are described in the (ϕ) phase-space, by

the sign of the invasion exponent, $S_x(y)$, which is the growth rate of a rare mutant y in an environment set by population x (Geritz et al., 1997). The invasion exponent is a function of X, ϕ, ψ, and ¥. Now X is the vector of different resources (the environment), which is 1 + n dimensional, where n is the number of isolated populations in the community. ψ is a set of 1 + n sets of parameters, containing the parameters of the abiotic environment plus the sets of traits of each isolated population.

If the dimensionality of the environment is measured discretely, the complexity of the problem grows in direct proportion to the number of different coexisting populations. In a model, where mass and energy conservation are explicitly considered, the maximal number of coexisting populations (hence the dimensionality of the environment) is constrained by the energy available to fuel the ecosystem, and only a finite (and expectably small) number of coexisting species is possible.

BIOENERGETICS PROBLEMS

R = 1.987 cal/mole °K

Remember to use T in °K (25° C = 298° K)

1. Given the following ΔG°'s, calculate the Keq's:
   a. ΔG°' = -4.8 kcal/mole
   b. ΔG°'=1.7kcal/mole
   c. ΔG°' = -95.6 cal/mole
2. Given the following Keq's, calculate the AG°'s:
   a. Keq = 20
   b. Keq = 0.003
   c. Keq= 1.4
3. Given the following chemical equations and AG°'s, calculate the AG and the Keq for the isomerization of glucose-1 - phosphate to fructose-6-phosphate

   Glucose-IP -> Glucose-6P ΔG0< = -1.74 kcal/mole

   Fructose-6P —> Glucose-6P ΔG°' = -0.4 kcal/mole

4. Calculate the *actual* DG of ATP hydrolysis in muscle if the concentrations of reactants and products are those given below: (Remember to use the correct temperature)

   ATP ⟶ ADP + Pi Keq = 2.22×10^5

   3 mM 6 mM 4 mM

5. How do muscle cells keep up their ATP supplies:

   In the short term? (seconds)

   In the long term? (minutes)

   Why do you think muscles fatigue?

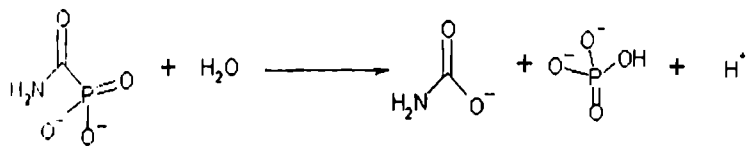
6. The $\Delta G^{\circ\prime}$ of ATP hydrolysis is -7.3 kcal/mole. The $\Delta G^{0\prime}$ of glucose-6-phosphate hydrolysis is –3.3 kcal/mole. What is the $\Delta G^{0\prime}$ of the following reaction:

ATP + glucose ⟶ glucose-6-phosphate + ADP?

What relative concentration of glucose-6-phosphate and ADP would be necessary to cause the reverse reaction to occur spontaneously?

7. Calculate the $\Delta G0'$ in kcal/mole for the following reaction:

Carbamoyl phosphate + H_2O ⟶ Carbamate + HPO_4^{2-} + H+



$K_{eq} = 4.5 \times 10^8$

What chemical characteristics of the reactants and products generate this magnitude of free energy?

8. The Keq for the forward movement of stubborn Jack Longears is 2.15×10^{-26}. A certain impetus can be provided by applying an energy boost from the following reaction, catalyzed by the infamous rubberbandase:

Limb band + 10-inch stretch + target ⟶ sharp thwack + loose band + no stretch

$\Delta G^{\circ\prime}$ = -20 kcal/mule

How many rubberbandase reactions are required to fuel any forward motion?

9. How does ATP –> AMP + PPi allow formation of a new thioester?

   a. The reaction is:

   $ATP + RCOOH + CoA\text{-}SH \longrightarrow AMP + PPi + RCO\text{-}S\text{-}CoA$

   fatty acid fatty acyl CoA

   $\Delta G^{\circ\prime} = 0.2$ kcal/mole

   b. Will this reaction tend to go forward?

   c. The hydrolysis of ATP isn't yet finished:

(1) The first step is

$ATP + H_2O \quad AMP + PPi \quad \Delta G^{\circ\prime} = -7.7$ kcal/mole

(2) In a second step, PPi is hydrolysed, as well:

$PPi + H_2O \longrightarrow 2Pi \quad \Delta G^{\circ\prime} = -6.9$ kcal/mole

(3) The total reaction delivers:

$ATP + H_2O \longrightarrow AMP + 2Pi \quad \Delta G^{\circ\prime} =$ _____kcal/mole

d. So the true $\Delta G^{\circ\prime}$ for fatty acid activation with pyrophosphate cleavage is

10. What happens to patients in whom the respiratory chain has been poisoned to become uncoupled from generation of ATP? What symptoms would they suffer? Could the problem be alleviated? How?

Hints for solving the Bioenergetics problems

Hints for solving the Bioenergetics problems

Questions 1, 2 and 8:

Use the equation: $\Delta G' = -2.303RT \log K_{eq}$

Questions 3 and 9:

ΔG's for summed reactions - write the reactions in their correct order and direction, then decide the correct ΔG's to use.

Question 4: Calculating the actual AG of a reaction. Use the equation: AG = AG0' + 2.303RT log [prod]/[react]

Question 5: Remember the only other compound in the middle of the energy "slope" with ATP.

Question 6: See #3 above; also, realize what actual conditions must occur for the reaction to reverse direction. (i.e., decide what conditions the actual AG must satisfy)

Question 7: This is right from your notes.

Question 10: A preview - read ahead to oxidative phosphorylation.

Chapter 2

PHOTOBIOLOGY

This branch of science deals with the study of light energy, and its interaction with living organisms, for example, origin of light, nature of light, energy content of light, primary photochemical reactions, photosynthesis - the process of conversion of light energy into chemical energy, vision - interaction of light with cells in retina of the eye, so that organisms can see. Some organisms give off light as a result of chemical reactions, which take place in specialized organs in their bodies. The intensity of light at different times of a day has different effects on the organisms - on the production of certain hormones, vitamins and even on the behaviour. Light energy can exert harmful effects on organisms, for example, too much exposure to sunlight can cause skin cancer, as well as premature aging of the skin and the eyes, ultraviolet light energy can cause harmful mutations in cells, as well as on gene expression.

Living organisms use light energy in two ways: (1) in the conversion of light energy into chemical energy by the process of photosynthesis - a characteristic feature of green plants; (2) in the visual sensation of the organisms. It is important to observe the molecular mechanisms of highly sensitive photosynthetic and visual responses in biological systems.

NATURE OF LIGHT

Light comprises a small region of continuous *electromagnetic spectrum* of radiation energy emitted by the sun. It has both electric and magnetic properties. This radiation contains gamma rays, x-

rays, ultraviolet rays, visible rays, infrared rays, and radio waves (Figure 21). The visible light region occupies a very small portion of the electromagnetic spectrum. Light moves through vacuum at 3,0 x 10^8 m/s (speed of light). It has a wave *like* characteristic. The electromagnetic spectrum extends from gamma rays with wavelengths of one hundredth of a nanometer to radiowaves with wavelengths of one meter or greater.

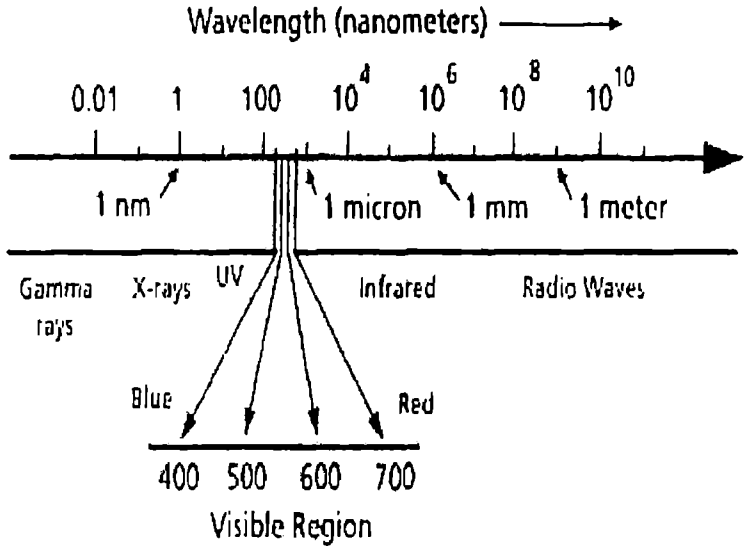


Figure 21: The electromagnetic spectrum of light energy.

The nanometer is a unit of distance in the metric scale and is abbreviated as 'nm'. One nanometer (nm) equals one thousand millionths of a meter (m) or 1 nm 10^{-9} m. One nanometer is a distance too small to be resolved in an optical microscope, but one micron (μm) or one thousand nanometer can be resolved (1 micron = 1000 nm). The wavelengths of visible light are smaller than common objects such as the thickness of a sheet of paper or the diameter of a human hair. Both of these are about one hundred microns thick which translates to distances greater than one hundred wavelengths of a visible light.

As we move through the visible light spectrum of violet, indigo, blue, green, yellow, orange, and red, the wavelengths become longer (Figure 22). The range of wavelengths (400 - 700 nm) of visible light is centrally located in the electromagnetic spectrum, infrared and radio waves are at the long wavelength side, while ultraviolet (UV), x-rays and gamma rays lie at the short wavelength side of the electromagnetic spectrum. Radiation with wavelength shorter than 400 nm cannot be sensed by the eye. Light with wavelength longer than 700 nanometers is also invisible.

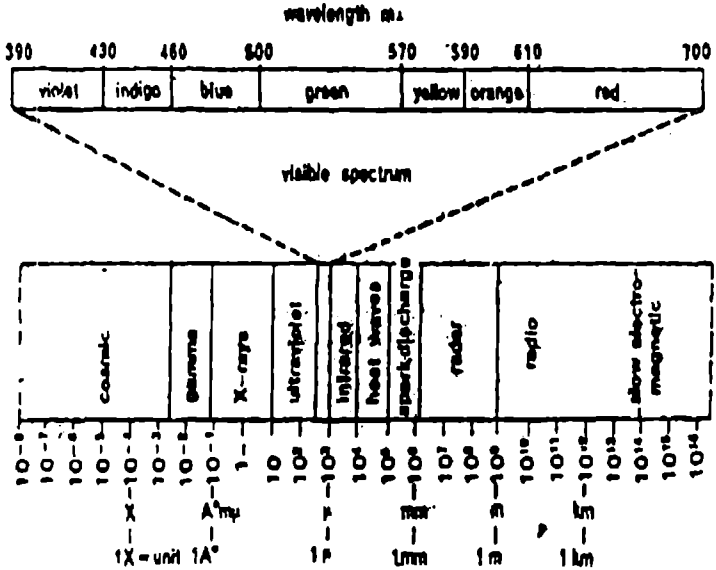


Figure 22: The electromagnetic spectrum of radiant energy (after Vernberg and Vernberg, 1970).

Light behaves both as a wave phenomenon and as discrete particles of energy called *photons*. If we look at light as a wave phenomena, it has peaks and troughs. We can assign it a *wavelength* (the distance from one peak of the wave to the next) and an *amplitude* (the distance the wave oscillates from its centre line) (Figure 23). Different wavelengths of light have different characteristic energies and properties. Light can also travel at various speeds in different media, producing a *frequency* at which

the wave travel. The energy contained in a wave of light is related to its frequency:

$$E = \frac{hc}{\lambda}$$

Where E is energy, h is planck's constant energy (6.626196* 10^-34 Joule-seconds), and c is the speed of light. Short wavelengths have higher energies and long wavelengths have lower energies.

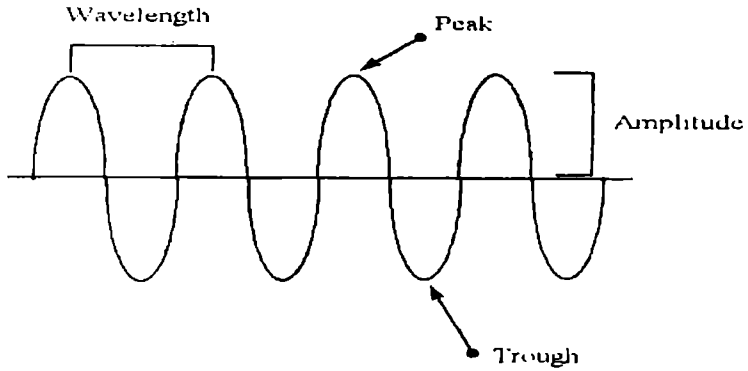


Figure 23: The wave nature of light.

If the wavelength is long, there will be fewer cycles passing a given point per second (Figure 24). Thus, the frequency will be low. If the wavelength is short, there will be more cycles passing a given point per second (Figure 25), and the frequency will be high.

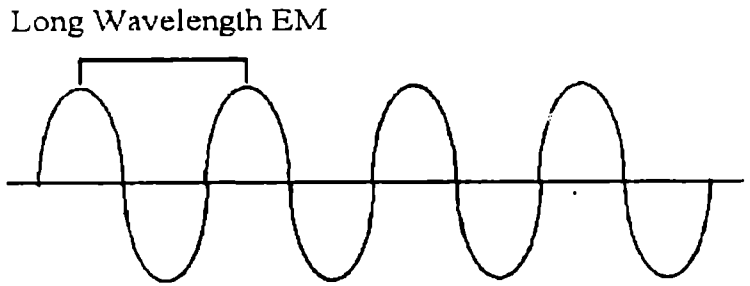


Figure 24: Long wavelength of light.

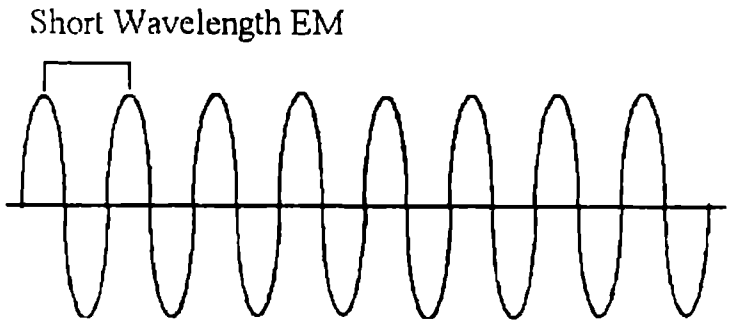


Figure 25: Short wavelength of light

The number of complete wavelengths or cycles, that pass a given point in one second is the frequency of the wave.

frequency = cycles per second

There is an inverse relationship between wavelength and frequency:

$$\text{Frequency} = \left(\frac{1}{\text{wavelength}}\right) \times \text{speed of light}$$

$$v = \left(\frac{1}{\lambda}\right) \times c$$

$$v\lambda = c$$

We can describe light as electromagnetic waves with colour identified by its wavelength, we can also consider light as a stream of minute packets of energy - photons - which create a pulsating electromagnetic disturbance, *A* single photon of one colour differs from a photon of another colour only by its energy.

The most convenient unit of energy is the electron volt (ev). The electron volt is the energy gained by an electron that moves across a positive voltage of one volt (V). Visible light is composed of photons in the energy range of around 2 to 3 ev (Figure 26). As the energy of light increases, the wavelength decreases. It is the energy range of 1.0 to 3.1 ev which triggers the photo receptors

in the eye. Lower energies (longer wavelengths) are not detected by the human eye but can be detected by special infrared sensors. Higher energies (shorter wavelengths) such as x-rays are detected by sensitive photographic film.

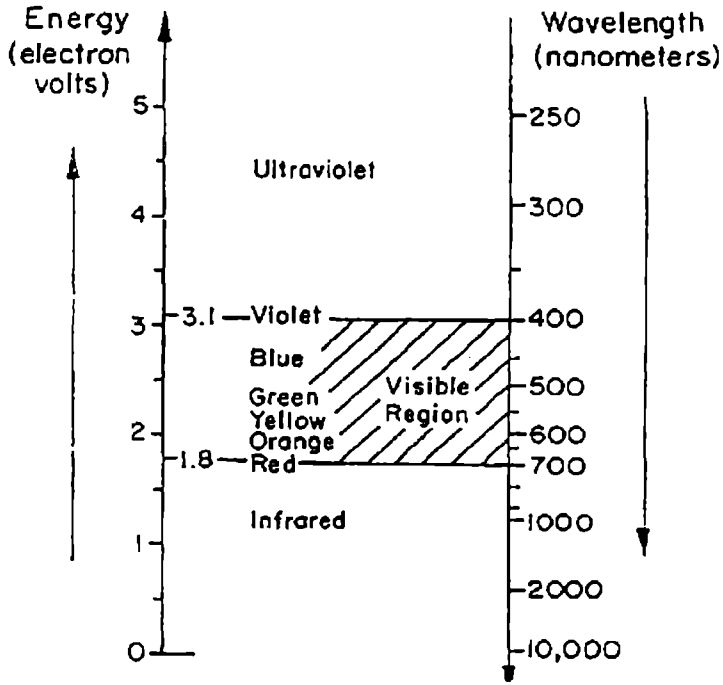
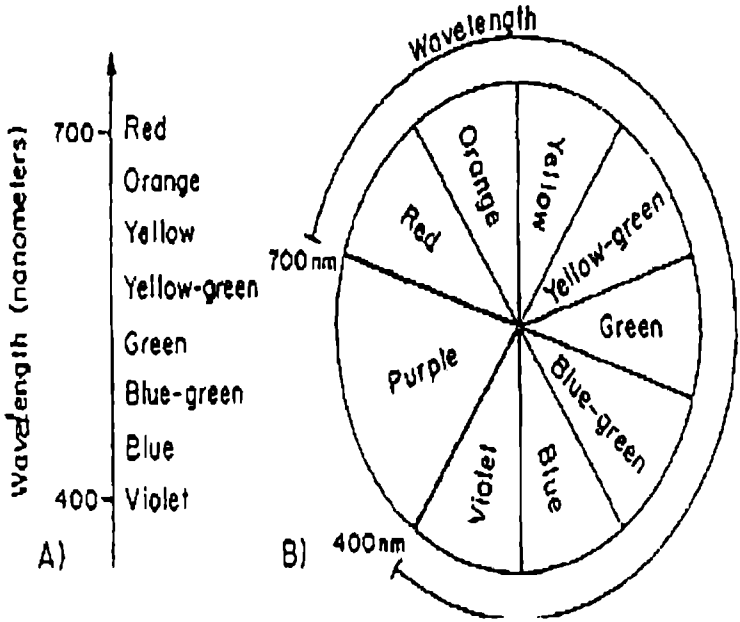


Figure 26: Diagram showing the visible spectrum in terms of wavelength and corresponding energies.

Light show the major spectral colours as a linear and circular sequence from violet at 400 nm to red (at 700 nm). The progression of colours from violet through red is identical to that of the linear scale. The circular wavelength scale outside the colour wheel shows the wavelength connection between the linear and the circular sequences. The purple region in the colour wheel is a notable difference between the two sequence: Colours in this purple portion of the colour wheel are composed of mixtures of wavelength and cannot be represented by a single wavelength.



. Figure 27: Linear (A), and circular (B) colour arrangement in visible light.

ENERGY CONTENT OF LIGHT

The total amount of solar radiation falling on the surface of the earth in the form of photon is enormous which is estimated to exceed 2 x 10^{25} cal per year. Only 12 percent of this energy is available to plants and animals, the rest is outside the visible range, or is absorbed by the atmosphere, or gets absorbed by the non-living substances of the earth's surface.

Photon of ultraviolet wavelength have the maximum energy content, and that of the infrared region of spectrum have the minimum energy content (Table 3). Photon of shorter wavelength, at the violet, visible limit of the spectrum have greatest energy content, while that of the red-visible limit have a smallest energy content.

Table 3: Photon energy as a function of wavelength.

Region of Spectrum	*Microns*	*E, kilocalories per g-mole*
Ultraviolet	0.3	95.3
Violet, visible limit	0.38	75.1
Visible	0.7	40.8
Red-visible limit	0.76	37.6
Infrared	1	28.6

In the visible range a photon carries from 40.9 to 71.5 k cal of energy, depending on the wavelength of light (Table 4).

Table 4: Energy content of visible range of solar radiation.

Wavelength(nm)	*Colour*	*Kilocalories*	*kilojoules*
700	Farred	40.9	171
	Red		
600	Orange	47.7	199
	Yellow		
500	Green	57.2	239
	Blue		
400	Violet	71.5	299

It has been found that the energy content of a mole of photons regardless of wavelength, is considerably greater than the amount of energy required to synthesize one mole of ATP from ADP and phosphate, which is 7.3 k cal under standard thermodynamic conditions. The ability of a compound to absorb photons depend on the atomic structure, particularly on the arrangement of electrons surrounding its atomic nuclei.

CHANGE IN MOLE AFTER LIGHT ABSORPTION

When a molecule in ordinary state absorbs a quantum of light, then this energy-rich molecule is known as *excited molecules*. The change of electron distribution in the excited molecule makes

it different from the ordinary state molecule in its physical and chemical properties, for example, in size and sometimes shape for which it is regarded as a new molecule. This excited molecule with electronic and vibrational changes may undergo the following changes:

(1) It may emit light immediately as fluorescence and drop to a lower energy level.

(2) It may pass to a different state which may emit light for a longer time as phosphorescence.

(3) It may get the whole or part of its excess energy converted to heat, that is, kinetic energy of the surrounding molecules.

(4) It may react chemically as it is in an energetically activated state.

(5) It is not only electronically excited but also have excess vibrational energy, that is, nuclear motions, bounding and non-bounding electrons are energised.

PRIMARY PHOTOCHEMICAL REACTIONS

The initial steps in photochemical reactions is the absorption of a photon by an atom, molecule, free radical or ion. The result of this absorption is strongly dependent on the energy, in other words the wavelength of the photon. Visible and ultraviolet wavelength of light is required to start the photochemical reactions. The absorption of light can generate dissociation, internal rearrangement, fluorescence, or excited species. In photochemical reactions, the excited molecule may show either of the two pathways. It may either directly undergo a chemical change:

$$A + h\nu \longrightarrow A'$$

or it may form active intermediates, usually free radicals:

$$A' \longrightarrow A_1 + A_2$$

These two reactions are caused by two direct result of light absorption (hv); hence are called primary photochemical reactions. All subsequent reactions of these free radicals are called secondary photochemical reactions, primary photochemical reactions occur because of photon absorption and dissociation. Secondary photochemical reactions occur due to availability of these primary products.

Species which absorb a photon and then dissociate are the fundamental in the occurrence of *photochemical smog*, in Los Angeles, in 1950s, a sharp increase in NO and non-methane hydrocarbon concentration was noticed due to starting of the automobile traffic. By late morning hydrocarbons and NO concentrations began to decrease whereas, NO_2 concentration began to increase. At mid-day an increase of the concentration of NO_2 occurs along with the rising in the concentrations of oxidants, especially ozone. As the afternoon proceeded and the sun-set start, the decrease in the ozone concentration occur. In the late afternoon, the decline of the NO and NO_2 levels was observed. These observations proved that the availability of a relationship between sunlight, hydrocarbons/volatile organic compounds, NO and NO_2.

It was also observed that CO show an increase in the early morning and late afternoon periods. CO takes an important role in the smog formation process by reacting with OH free radicals to produce free radicals to generate a free hydrogen atom. Consequently, free hydrogen atom reacts with oxygen rapidly to form the hydroperoxy free radical (HO_2). This radical gets involved in the formation of ozone.

DISSOCIATION OF NO_2

Photodissociation of NO_2 is a specific example of photochemical reaction, NO_2 absorbs over the whole of the visible and ultraviolet range of the solar spectrum with a decrease in absorption in the longer wavelength visible portion. The colour of the gas is reddish brown during the reaction. The energy requirement to break the bond between the NO and NO_2 is approximately 72 k cal/gmole at 25 °C. The dissociation is dependent on wavelength. Above 4200 A, due to insufficient energy to achieve dissociation, other photochemical effects like fluorescence occur. Below about 3700 A the rate of molecules undergoing process per photon absorbed is more than 90 percent. In typical sunlight the half life of NO_2 is approximately two minutes.

CONTRIBUTION OF HYDROCARBONS TO SMOG FORMATION

In terms of reactivity, some hydrocarbons emitted are worse than the others. It means that some hydrocarbons are more likely to get involved in these chemical reactions than other groups. However, the criteria should be defined for reactivity, in the event of using nitric oxide photooxidation rate as a basis, then the reactivity of several hydrocarbons can be calculated (Figure 27). Photooxidation implies the rate at which the hydrocarbons cause NO to be oxidised be NO_2, given in parts per billion per minute.

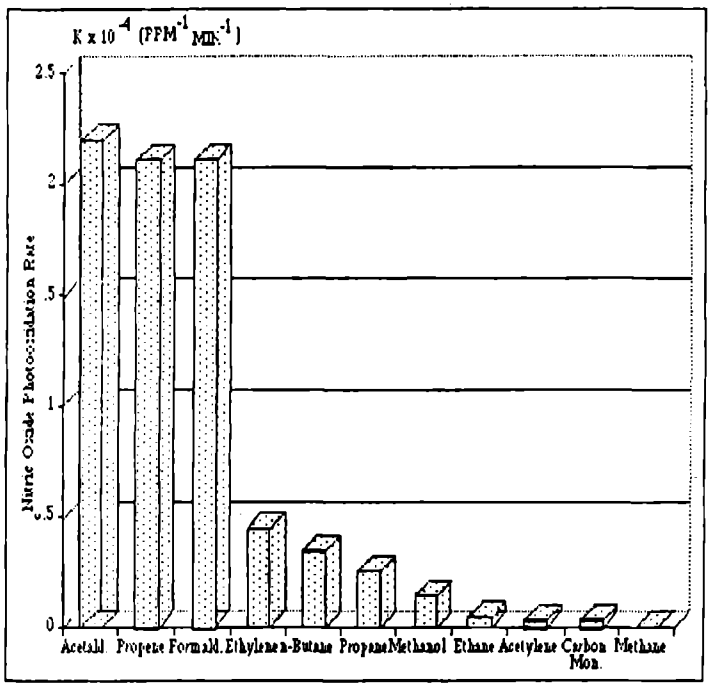


Figure 28: Reactivity index of several compounds.

To control the emission of more reactive hydrocarbon compounds means that the control of the photochemical smog,

This can be possible by reducing the emission of internally double-bonded olefins, the diolefins, and the cycloalkanes. It should be noted that the level of different hydrocarbons varies during the day. The concentrations of more reactive hydrocarbons like olefins and diolefins (alkenes and alkadiens) like ethylene, propylene, butadiene decrease dramatically, the drop in the concentrations of cycloalkenes such as cyclopentene is apparent. However, the alkanes such as the methane and ethane are collected in the atmosphere, since they do not react with other compounds.

PHOTOLYTIC CYCLE

In first step, NO_2 is dissociated into NO and an free radical oxygen atom by ultraviolet light. Then, appearing oxygen atom quickly makes a combination with molecular oxygen to form ozone (Figure 29).

$$NO_2 + hv = NO + O^*$$

$$O^* + O_2 + M = O_3 + H$$

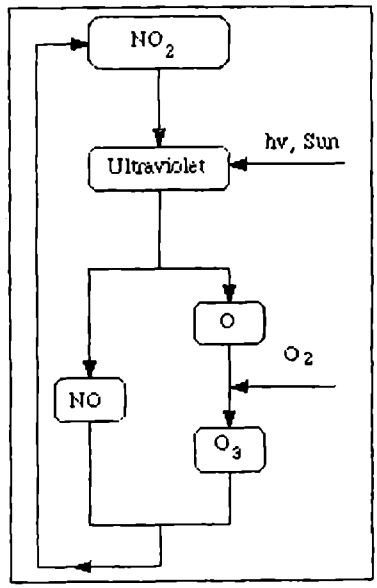


Fig. 29: Photolysis of NO_2 and generation of O_3.

M represents any other molecule especially N_2 or O_2 which absorb the energy of the reaction, without the M body, only oxygen exchange within an oxygen molecule would occur. Although a triple reaction is required, the reaction is kinetically fast. The third reaction completes the cycle.

$$O_3 + NO = NO_2 + O_2$$

This reaction also occurs fast. Constant level of each compounds, NO, NO_2 and O_3 could be formed when those three reaction are happened. Steady-state ozone formation can be predicted as a function of initial NO_2 concentration. O_3 steady-state concentration increases with decreasing concentration of nitric oxide and vice-versa.

$$(O_3) = \frac{K_2(NO_2)}{K_1(NO)}$$

K_2 and K_1 indicate the rate constant for the reactions, is approximately 1.2 ppm for the Los Angeles noonday condition.

Calculation show that 10 pphm NO_2 causes appropriately 2.7 pphm ozone. In fact, most of the NO_2 emitted from combustion process is NO and levels does not usually exceed 10 pphm. But ozone level usually reaches to 50 pphm for 1 hour peak average.

Further research has showed that there should be a mechanism which convert NO to NO_2 without consuming ozone.

The overall higher production of ozone can be explained by the impact of reactive hydrocarbons. Olefins is the most reactive group because of double bond, oxygen atom attack olefin and divide it into two parts (ozone can also do that but reaction with oxygen is faster). Highly reactive free radical which is an incomplete hydrocarbon appears and continues to get involved in other reactions (Figure 30).

$$O + \text{olefin} \Rightarrow R^{\cdot} + R^{\cdot}O$$

$$R^{*} + O_2 \Rightarrow RO_2^{\cdot}$$

$$RO_2^{\cdot} + NO \Rightarrow NO_2 + RO^{\cdot}$$

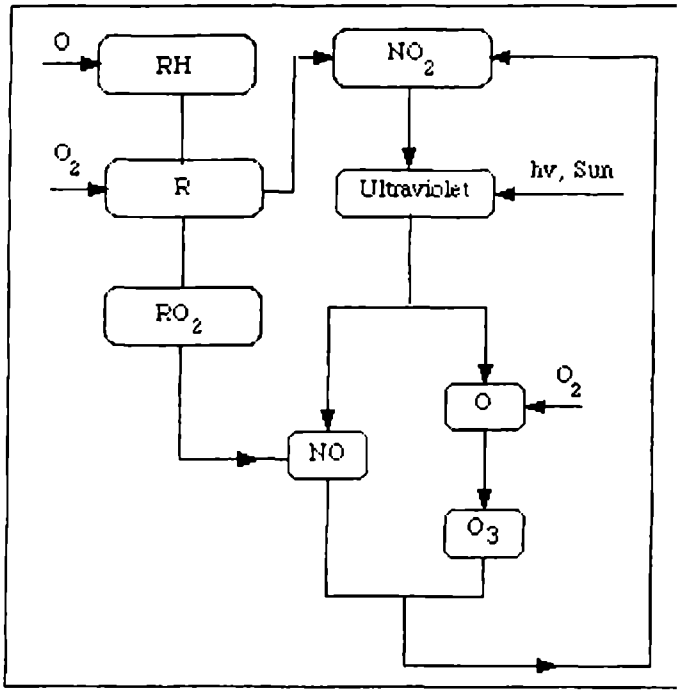


Figure 30: The effect of hydrocarbons in the ozone cycle.

Compound which is noted RO_2 stands for peroxy radica Because, each hydrocarbon molecule requires one oxygen ato to start its oxidation, one hydrocarbon molecule dissociated cau more than one NO molecule to convert NO_2.

then original reactions occur,

$NO + hv \Rightarrow NO + O^*$

$O^* + O_2 + M \Rightarrow O_3 + M$

where; h : Ultraviolet radiation

R : Hydrocarbon group like CH_3

* : Free radical

The other radical denoted as R'O might be an aldehyde, since aldehydes are among the products of the reaction between olefins and O.

Photolysis of the formaldehyde results in:

$$HCHO + hv \Rightarrow H + HCO$$

Consequently, the fast reaction of the products with O_2 :

$$H + O_2 \Rightarrow HO_2$$

$$HCH + O_2 \Rightarrow HO_2 + CO$$

The reaction of formaldehyde with OH radical yields:

$$HCHO + OH \Rightarrow H_2O\ HCO$$

$$HCO + O_2 \Rightarrow HO_2 + CO$$

For higher aldehydes, the process is as follow:

$$RCHO + OH \Rightarrow RCO' + H_2O$$

$$RCO' + O_2 \Rightarrow RCOOO'$$

RCOOO' can react with either NO or NO_2

$$RCOOO' + NO \Rightarrow RCOO + NO_2$$

$$RCOO' + NO_2 \Rightarrow R' + CO_2$$

In the event of reaction with NO_2

$$RCOOO' + NO_2 \Rightarrow RCOOONO_2$$

results in peroxyacetyl nitrate (PAN)

OTHER MECHANISMS

Although several mechanisms are available for photochemical smog formation, the NO_2-NO-O_3-NO_2 cycle is the main part in all the models. Other mechanisms can be listed as follow:

ACETALDEHYDES

The reaction of acetaldehyde with atomic oxygen generates hydroxyl radicals (Figure 31)

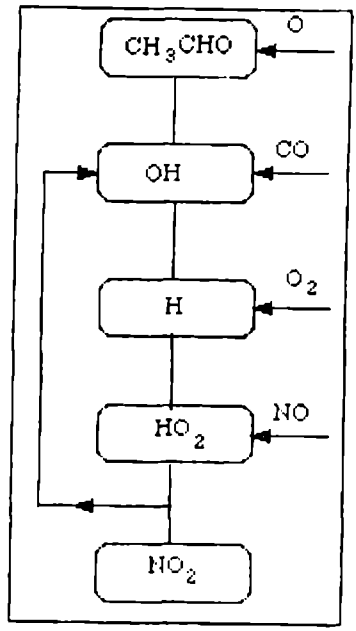


Figure 31: The impact of acetaldehyde on ozone formation.

$CH_3CHO + O \Rightarrow CH_3CO + OH$

The second step is the reaction of OH radical with CO

$OH + CO \Rightarrow CO_2 + H$

The hydrogen can react with molecular oxygen

$H + O_2 + M \Rightarrow HO_2 + M$

Consequently, hydroperoxyl radical oxidize nitric oxide

$HO_2 + NO \Rightarrow NO_2 + OH$

It can be seen that the original nitrogen-oxygen cycle is valid in this mechanism. It should be noted the OH is first used as a reactant and generated as a product.

While nitric dioxide participate in the buildup of ozone, regenerated OH and available CO keep the reaction go on and on. The role of CO in the ozone formation should be underlined.

ALKANES

Alkanes react with OH radicals (daytime) and NO_3 (night-time)

$$RH + OH \Rightarrow R' + H_2O$$

$$RH + NO_3 \Rightarrow R' + HNO_3$$

Reaction of R* with O_2 :

$$R' + O_2 \Rightarrow RO_2'$$

The reaction of alky peroxy radical with NO occurs in two ways.

i) For the compounds < C4

$$RO_2' + HO \Rightarrow RO' + NO_2$$

RO: Alkoxy radical

ii) For larger alky peroxy radicals:

$$RO' + NO \Rightarrow RONO_2$$

ALKENES

Gas-phase alkenes react with OH radicals, NO_3 radicals, O_3 radicals. Among these, reaction with OH radicals is fast. For example, for propylene:

$$CH_3CH = CH_2 + OH \Rightarrow CH_3CHCH_2OH + CH_3CHCH_2\text{-}OH$$

The first product is the dominant in this process. The reaction of this radical with O_2 results in

$$CH_3CHCH_2OH + O_2 \Rightarrow CH_3CCH_2OH\text{–}O'$$

beta-hydroxyalkoxy follows two pathways, this radical can react with O_2 or decompose. It should be noted that isomerization is not important for smaller alkanes.

$$CH_2CHCH_2\ OH - O' + O_2 \Rightarrow CH_3CCH_2OH = O + HO_2$$

or decomposition of the product

$$CH_3CHCH_2OH - O' \Rightarrow CH_3CHO + CH_2OH$$

The available data shows that the decomposition of -hydroxyalkoxy is dominant over the first pathway. Then, it can be concluded that OH radical reactions with alkenes lead to the formation of aldehyde and/or ketones.

BASIC LAWS OF PHOTOCHEMICAL REACTIONS

There are two basic laws of photochemical reactions:

(a) Grothus-Drapper photochemical law or first law of photochemical reaction

The law states that only light which is absorbed is photochemically active. Each light absorbing substance has one or more specific band or wavelength or absorption maxima (λ max) at which it absorbs light maximally and only then this light may produce maximal activation of the substance and produce maximum photochemical reaction. The absorbed light is not always utilised or fully utilised in photochemical reactions and may re-emit wholly or it may be transformed into heat.

(b) Einstein's law of photochemical equivalence or second law of photochemical reaction

The law states that the primary process in photochemical reaction is absorption of one quantum of light energy *hv* by one molecule or atom to form an excited molecule or atom. The proportional ratio between the number of quanta and the number of excited molecules is 1:1. This is deviation from the physics of wave nature of light according to which the light energy should have been spread out over a number of molecules and not one quantum per molecule.

From Einstein's law, each mole of light-absorbing substance must absorb the Avogadro number of photons or quanta of light, this amount of energy is known as *Einstein* (E):

$$E = Nhc/\lambda$$

where hc/λ = Energy of a photon or quantum

N = Avogadro number or the number of molecules per mole

The quantum is given by hv; v is the frequency of the light, is related to the wavelength, χ, by $v\chi = c$ where c is the velocity of light, so that the quantum is $hc = \chi$. If each molecule absorbs

one quantum, the activation energy for one mole is Nhc/χ, where N is Avogadro's number

$N = 6.023.10^{23}$ $h = 6.624.10^{-27}$ and $c = 2.9977.10^{10}$ cm/sec

The energy will be given in ergs

Thus $E = 1.96 \times 10^{8}/\lambda$ ergs/mole

when λ is given in cm.

Again on dividing by $4.184.10^{10}$, the result comes out in calories (Kcal). Thus

$E = 2.859.10^{10}/\lambda$ cal/mole.

In the visible range, 1 Einstein carries 40-72 Kcal of energy, depending on the wavelength of the light.

The observed photochemical change is the result of secondary reactions of the excited molecules or atoms or free radicals produced by the secondary reaction as governed by Einstein's law.

QUANTUM EFFICIENCY

The interpretation of the Einstein equivalence law indicates that the quantum efficiency, that is, the number of molecules reaching per quantum absorbed, would be equal to unity;

$$\text{Quantum efficiency} \quad \phi = \frac{\text{number of molecules decomposed}}{\text{Number of qunata absorbed}}$$

If the sole reaction is the decomposition of this excited molecule, we should get unity for the quantum efficiency for all photochemical reactions. Actually, very few reactions exhibit this efficiency for example, decomposition of HI by ultraviolet light show an efficiency of 2, whereas many other have values less than unity. The value ranges from very low to very high. Low values are due to some kind of non-utilisation of excited molecules for the reaction and high efficiency in the decomposition of HI is due to secondary reactions following the primary reaction, which consists in the splitting of the molecules into two atoms:

$HI + h\nu \rightarrow H + I$

Following this the atoms may react in a variety of ways, the principal ones being:

$$H + HI \rightarrow H_2 + I$$

$$I + I \rightarrow I_2$$

Thus, one quantum causes the splitting of two molecules. A quantum efficiency of less than unity is given *by Inelastic collisions* of the excited molecules with normal molecules, the absorbed light energy being dissipated as heat; alternatively, some of the light may be lost by fluorescence, the excited molecule retaining its energy for some 10^{-7} to 10^{-8} sec, and then re-emitting most of a longer wavelength,

PHOTO-SENSITIZATION AND PHOTOSENSITIZER

According to the first law of photochemistry only the absorbed light is photo-chemically reactive. But sometimes, an external substance makes a system sensitive to some region of spectrum to which the system is sensitive because the system does not have absorption in this region. The external substance, that is, sensitizer, absorbs the light and makes it available to the system. Such a phenomenon is called photosensitization, and the substance responsible for photosensitization is called photosensitizer or a sensitizer. Thus, ozone does not decompose in visible light because it has no absorption band in the visible light, but if a trace of Chlorine is present, it is rapidly decomposed by visible light which is absorbed by Chlorine. The great example in biology is photosynthesis in which Chlorophyll acts as a photosensitizer.

PHOTOSYNTHESIS

Photosynthesis is the physico-chemical process by which primary producers (mostly plants, algae and photosynthetic bacteria) use light energy to drive the synthesis of organic compounds, in plants, algae and certain types of bacteria, the photosynthetic process results in the release of molecular oxygen and the removal of Carbon dioxide from the atmosphere that is used to synthesize carbohydrates *(oxygenic photosynthesis).* Other types of bacteria use light energy to create organic compounds but do not produce

oxygen *(anoxygenic photosynthesis)*. Photosynthesis provides the energy and reduced carbon required for the survival of virtually all life on our planet, as well as the molecular oxygen necessary for the survival of oxygen consuming organisms. In addition, the fossil fuels currently being burned to provide energy for human activity were produced by ancient photosynthetic organisms. Although photosynthesis occurs in cells and organells that are typically only a few microns across, the process has a profound impact on the earth's atmosphere and climate. Each year more than 10 percent of the total atmospheric Carbon dioxide is reduced to carbohydrates by photosynthetic organisms. Most if not all, of the reduced carbon is returned to the a atmosphere as Carbon dioxide by microbial, plant, and animal metabolism, and by biomass combustion. In turn, the performance of photosynthetic organisms depends on the earth's atmosphere and climate. Over the next century, the large increase in the amount of atmospheric Carbon dioxide created by human activity is certain to have a profound impact on the performance and competition of photosynthetic organisms. Knowledge of the physico-chemical process of photosynthesis is essential for understanding the relationship between living organisms and the atmosphere and the balance of life on earth.

The overall equation for photosynthesis in green plants as proposed by Dutch microbiologist van Nile (1941):

$$CO_2 + 4H_2O \xrightarrow[\text{Chlorolphyll}]{\text{Light}} CH_2O + 3H_2O + O_2$$

This equation is deceptively simple. In fact, a complex set of physico-chemical reactions must occur in a coordinated manner for the synthesis of carbohydrates It involves the sum of three individual reactions;

(1) $4\ H_2O \xrightarrow[\text{Chlorolphyll}]{\text{Light}} 4(OH) + 4H$

(2) $4\ H + CO_2 \rightarrow (CH_2O) + H_2O$

(3) $4\ (OH) \rightarrow 2\ H_2O + O_2$

Where CH_2O represent a carbohydrate (for example, a six carbon sugar glucose), the synthesis of carbohydrates from Carbon and water requires a large input of light energy. The standard free

energy for reduction of one mole of CO_2 to the level of glucose is + 478 KJ/mol. Because glucose, a six-carbon sugar, is often an intermediate product of photosynthesis, the net equation of photosynthesis is frequently written as;

$$6CO_2\ 12H_2O + \text{light} \xrightarrow{\text{Chlorolphyll}} C_6H_{12}O_6 + 6O_2 + 6H_2O$$

The reaction steps discussed above clearly show that the process of photosynthesis can be split into two distinct phases: (a) photolysis (the photo-chemical stage), and (b) the *Calvin cycle* (the thermochemical stage).

(a) Photolysis

Not surprisingly, early scientists studying photosynthesis concluded that 'the oxygen released by plants came from Carbon dioxide, which was thought to be split by light energy. In the 1930s comparison of bacterial and plant photosynthesis lead Cornelis van Niel to propose the general equation of photosynthesis that applies to plants, algae and bacteria. Van Niel was aware that some photosynthetic bacteria could use hydrogen sulphide (H_2S) instead of water for photosynthesis and that these organisms released sulphur instead of oxygen. Van Niel concluded that photosynthesis depends on electron donation and acceptor reactions and that the oxygen released during photosynthesis comes from the oxidation, of water,

$$CO_2 + 2\ H_2A \xrightarrow[\text{Chlorolphyll}]{\text{Light}} CH_2O + 2A + H_2O$$

In oxygenic photosynthesis, 2A is oxygen, whereas in anoxygenic photosynthesis, which occurs in some photosynthetic bacteria, the electron donor can be an inorganic hydrogen donor, such as H_2S (in which case A, is elemental sulphur) or an organic hydrogen donor such as succinate (in which case A is fumarate).

The photolysis part of photosynthesis occurs in the *granum* of a *chloroplast* where light is absorbed by *chlorophyll;* a type of photosynthetic pigment that converts the light energy to chemical energy. This reacts with water (H_2O) and splits the oxygen and hydrogen molecules apart.

From this dissection of water, the oxygen is released as a by-product while the reduced hydrogen acceptor makes its way to the second stage of photosynthesis, the Calvin cycle.

Overall, since the water is oxidised (hydrogen is removed) and energy is gained in photolysis, which is required in the Calvin cycle.

(b) Calvin Cycle

This stage is also known as the Carbon fixation stage, this part of the photosynthetic process occurs in the *stroma* of chloroplasts. The carbon made available from breathing in Carbon dioxide enters this cycle (Figure 32).

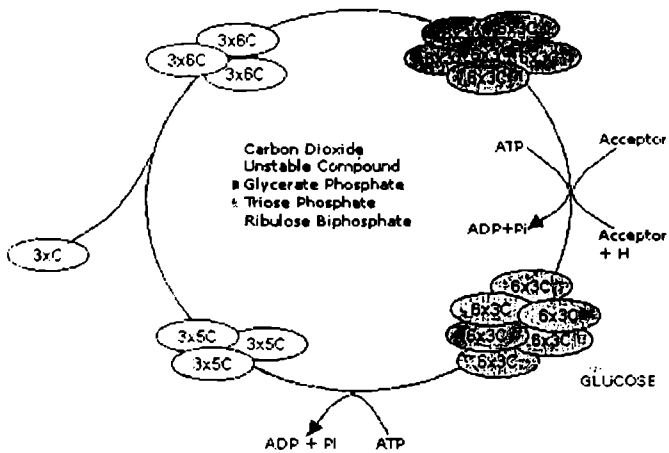


Figure 32: The Calvin cycle.

In Calvin cycle, a substrate is manipulated into various carbon compounds to produce energy. In case of photosynthesis, the following steps occur, which create glucose for respiration from the carbon dioxide introduced into the cycle:

- Carbon from CO_2 enters the cycle combining with Ribuloso Biphosphate (RuBp);
- A compound formed is unstable and breaks down from its 6 carbon nature to a 3 carbon compound called glycerate phosphate (GP);
- Energy is used to break down GP into triose phosphate, while a hydrogen acceptor reduces the compound therefore requiring energy;

- Triose phosphate is the end product of this, a 3 carbon compound which can double up to form glucose, which can be used in respiration;
- The cycle is completed, when the leftover GP molecules are met with a carbon acceptor and than turned into RuBp which is to be joined with the carbon dioxide molecules to re-begin the process.

The energy that is used up in the Calvin cycle is the energy that is made available during photolysis, the glucose that is made via *GP* can be used in respiration or a building block in forming starch and cellulose, materials that are commonly in demand in plants.

THE EVOLUTION OF PHOTOSYNTHESIS

Life theoretically originated on earth 3.5 to 4 billion years ago. The atmosphere was thin; composed of methane, carbon oxide, and water vapour. Any gaseous oxygen has been used up in the combustion (or oxidation) of materials when the earth was very hot.

The cooling water collected in pools, assimilating the nutrients from the rocks. As water evaporated, the nutrients concentrated, forming a rich soup. The first organisms would have made a good living off this food source, breaking down the complex molecules into water and carbon dioxide through respiration. Eventually, as life grew, the need arose to somehow synthesize complex compounds, both to eat and to use for structure and function. Some organisms learned how to use the sun's energy to synthesize large molecules from small molecules. Other organisms learned to use other sources of reductive power, these organisms who learned how to build the building blocks of life are called *autotrophs*, or self-feeders. Autotrophs are found in the bacterial and in the plant kingdom.

CLASSIFICATION OF PHOTOSYNTHETIC ORGANISMS

All life can be divided into three domains - Archaea, Bacteria and Eucarya; which originated from a common ancestor (Woese

et al, 1990). Historically, the term photosynthesis has been applied to organisms that depend on chlorophyll (Gest, 1993). These include the organisms from bacteria (photosynthetic bacteria) and Eucarya (algae and higher plants). The most primitive domain - Archaea includes organisms known as halo-bacteria, that convert light energy into chemical free energy.

(a) Oxygenic Photosynthetic organisms

The photosynthetic process in all plants and algae as well as in certain types of photosynthetic bacteria involves the reduction of CO_2 to carbohydrates and removal of electrons from H_2O, which results in the release of O_2. In this process, water is oxidized by the photosystem II reaction centre, a multi-subunit protein located in the photosynthetic membrane.

(b) Anoxygenic photosynthetic organisms

Some photosynthetic bacteria can use light energy to extract electrons from molecules other than water. These organisms are of ancient origin, presumed to have evolved before oxygenic photosynthetic organisms, and have representatives in four phyla purple bacteria, green sulphur bacteria, green gliding bacteria, and gram positive bacteria.

CHLOROPHYLL AND ACCESSORY PIGMENTS

A pigment is any substance that absorbs light. The colour of the pigment comes from the wavelengths of light reflected (in other words, those not absorbed), Chlorophyll, the green pigment common to all photosynthetic cells, absorbs all wavelengths of visible light except green, which it reflects to be detected by our eyes. Black pigment absorb all of the wavelengths that strike them. White pigment/lighter colour reflects all or almost all of the energy striking them.

Chlorophyll is a complex molecule (Figure 33). Several modifications of chlorophyll occur among plants and other photosynthetic organisms. All photosynthetic organisms (plants, certain protistans, prochlorobacteria, and cyanobecteria) have Chlorophyll. Accessory pigments absorb energy that chlorophyll-a does not absorb, Accessory pigments include chlorophyll-b (also

c, d, and e in algae and protistans), carotenoids (each as beta-carotene), and phycobilins (Table 5). Of these pigments chlorophylls are the most abundant pigments playing the most active role in photosynthesis, while carotenoids and phycobilins are called accessory pigments since the solar energy absorbed by them is transferred to chlorophyll.

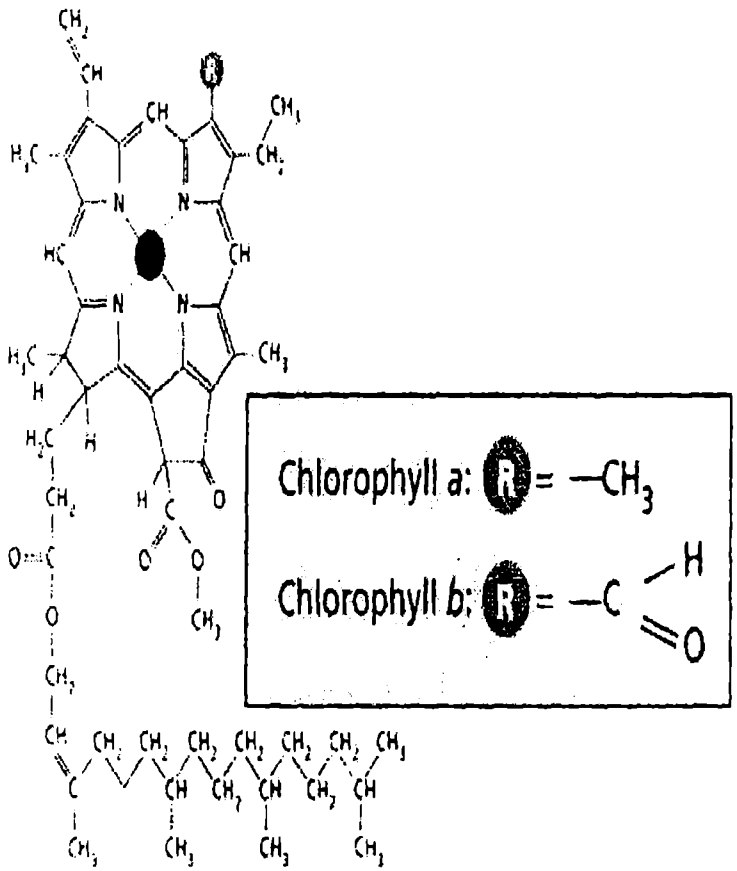


Fig. 33: The molecular structure of chlorophylls.

Table 5: Types and Occurrence of Photosynthetic Pigments

Type of pigment	Occurrence
Chlorophylls	
Chlorophyll *a* (all O_2-evolving organisms)	All higher plants and algae
Chlorophyll *b*	All higher plants and green algae
Chlorophyll *c*	Diatoms and brown algae
Chlorophyll *d*	Red algae
Carotenoids	
β-carotene	Higher plants and most algae
α-carotene	Most plants and some algae
Luteol	Green and red algae and higher plants
Violaxanthol	Higher plants
Fucoxanthol	Diatoms and brown algae
Phycobilins	
Phycocyanins	Blue-green algae and some red algae
Phycoerythrins	Red algae and some blue-green algae
Allophycocyanins	Blue-green and red algae
Bacteriochlorophyll *a*	*Chlorobium*
Bacteriochlorophyll *b*	*Chromatium, Rhodospirillum*

All photosynthesising cells except photosynthetic bacteria contain chlorophylls. In higher plants and algae chlorophyll-b occurs as another form of chlorophyll, which is generally found associated with chlorophyll-a. Besides chlorophyll-a and -b, which are universally present in land plants, several other closely related chlorophylls have been isolated. Chlorophyll-c is found in brown algae and diatoms. In red algae, chlorophyll-d has been found, Bacterio-chlorophyll is the main pigment of purple bacteria. Green bacteria contains chlorobium-chlorophyll or bacterioviridin and also traces of bacterio-chlorophyll. In one species of Xanthophyta chlorophyll-e has been observed.

Although, chlorophyll-a and -b closely resemble each other in molecular structure, they differ in their physical and chemical properties. For example, there is difference in the absorption spectra of the two pigments. Chlorophyll-a has the absorption maxima at 430 and 662 nm, while the corresponding peaks for chlorophyll-b occurs at 453 and 642 nm (Figure 34). Besides this, the solubility properties of the two pigments also differ; petroleum ether is a good solvent for chlorophyll-a, while methyl alcohol is a good solvent for chlorophyll-b.

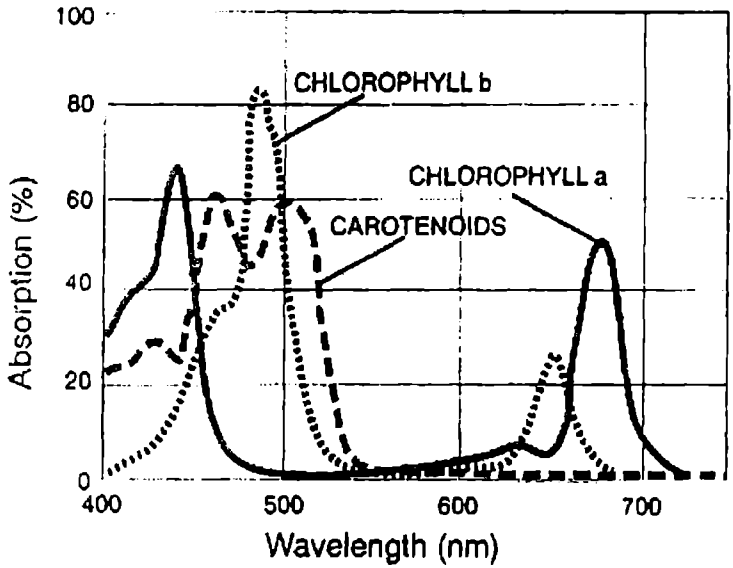


Figure 34: Estimated absorption spectra of chlorophyll- a, -b and carotenoid in chloroplasts.

Carotenoids are the second group of photosynthetic pigments, They are of two types - the *carotenes* and the *Xanthophylls*. Carotenes in higher plants are *beta-carotenes* with which small amount of *alpha-carotene* may occur. The major xanthophylls are *lutin, violaxanthin* and *neoxanthin*, of which lutin is most abundant. In photosynthetic green bacteria E-carotene is the major carotenoid. The major xanthophylls are found in higher plants and in algae, The brown algae and diatoms contain fucoxanthin. Generally, the

carotenoids are yellow or orange. Two or three absorption bands of carotenoids are located in the blue-violet region of the spectrum. Like chlorophyll, the carotenoids also play a basic role in the photosynthetic process. Two distinct roles have been envisaged for carotenoids: (1) they protect the chlorophyll against photodynamic action of light, and (2) the light energy absorbed by the carotenoid is not directly utilised.

The third group of photosynthetic pigments is phycobilins. They are not found widely distributed as chlorophyll and carotenoid. They occur in algae-Cyanophyta, Rhodophyta and Cryptophyta. Phycocyanin and phycoerythrin are major phycobilins. Phycocyanin has the main absorption band at 630 nm, while phycoerythrin has absorption bands at about 500, 545 and 570 nm (Figure 35). These pigments are also accessory pigments. Here also the light absorbed is transferred to chlorophyll-a before it takes part in photosynthesis.

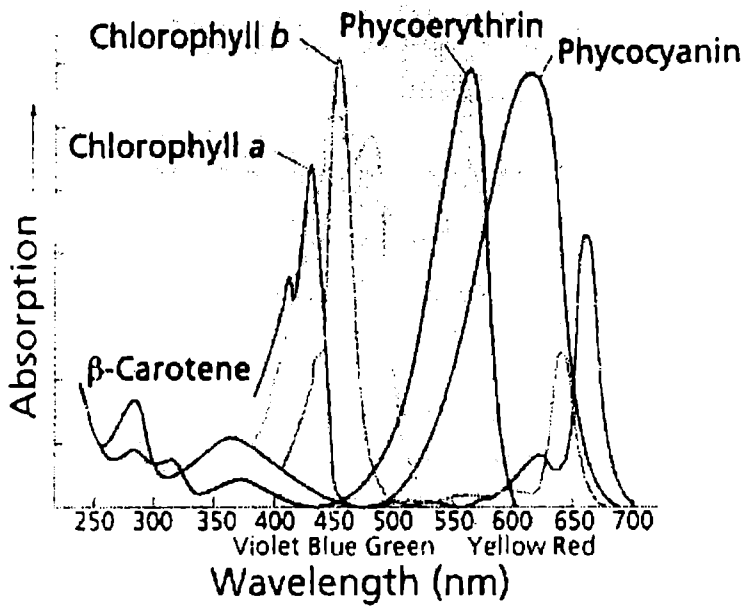


Figure 35: Absorption spectra of various photo synthetic pigments.

PHOTOSYNTHETIC ENERGY TRANSFORMATION

The energy that drives photosynthesis originates in the centre of the sun where mass is converted to heat by the fusion of hydrogen, over -time, the heat energy reaches the sun's surface, where some of it is converted to light by black body radiation that reaches the earth. A small portion of the visible light incident on the earth is absorbed by plants. Through a series of transducing reactions, photosynthetic organisms are able to -transform light energy into chemical free energy in a stable form that can last for hundreds of millions of years (as fossil fuels).

The photosynthetic process in plants and algae occurs in small orgenells known as chloroplasts that are located inside cells. The more primitive photosynthetic organisms, for example oxygenic Cyanobacteria, prochlorophytes and anoygenic photosynthetic bacteria lack chloroplasts. The photosynthetic reactions are traditionally divided into two stages - the *light reactions*, which consist of electron and proton transfer reactions and the *dark reaction* which consist of the biosynthesis of carbohydrates from Carbon dioxide. The light reactions occur in a complex membrane system (the photosynthetic membrane) that is made up of protein complexes, electron carriers, and lipid molecules. The photosynthetic membrane is surrounded by water and can be thought of as a two-dimensional surface that defines a closed space, with an inner and outer water phase. A molecule or ion must pass through the photosynthetic membrane to go from the inner space to the outer space. The protein complexes embedded in the photosynthetic membrane have a unique orientation with respect to the inner and outer phase, the asymmetrical arrangement of the protein complexes allows some of the energy released during electron transport to create an electrochemical gradient of protons across the photosynthetic membrane, photosynthetic electron transport consists of a series on individual electron transfer steps from one electron carrier to another, the electron carrier are metal ion complexes and aromatic groups. The metal ion complexes and most of the aromatic groups are bound with proteins. Most of the proteins involved in photosynthetic electron transport are composed of numerous polypeptide chains that lace through the

membrane, providing a scaffolding for metal ions and aromatic groups. An electron enters a protein complex at a specific site, is transferred within the protein from one carrier to another, and exits the protein at a different site. The protein controls the pathway of electrons between the carriers by determining the location and environment of the metal ion complexes and aromatic groups. By setting the distance between electron carriers and controlling the electronic environment surrounding a metal ion complex or aromatic group, the protein controls pairwise electron transfer reactions. Between proteins, electron transfer is controlled by distance and free energy, as for intraprotein transfer, and by the probability that the two proteins are in close contact. Protein association is controlled by a number of factors, including the structure of the two proteins, their surface electrical and chemical properties and the probability that they collide with one another. Not all electron carriers are bound to proteins. The reduced forms of plastoquinone or ubiquinone and nicotinamide adenine dinucleotide phosphate (NADPH) or NADH act as mobile electron carrier operating between protein complexes. For electron transfer to occur, those small molecules must bind to special pockets in the proteins known as binding sites. The binding sites are highly specific and are a critical factor in controlling the rate and pathway of electron transfer.

The light reaction convert energy into several forms (Figure 36). The first step is the conversion of a photon to an excited electronic state of an antenna pigment molecule located in the antenna system. The antenna system consists of hundreds of pigment molecules (mainly chlorophyll or bacterio-chlorophyll and carotenoids) that are anchored to proteins within the photosynthetic membrane and serve a specialised protein complex known as reaction centre. The electronic excited state is transferred over the antenna molecules as an exciton. Some excitons are converted back into photons and emitted as fluorescence, some are converted to heat, and some are trapped by a reaction centre protein. Excitons trapped by a reaction centre provide the energy for the primary photochemical reaction of photosynthesis - the transfer of an electron from a donor molecule to an acceptor molecule. Both the donor and acceptor molecules are attached to the reaction centre protein complex, once primary change

separation occurs, the subsequent electron transfer reactions ai energetically downhill.

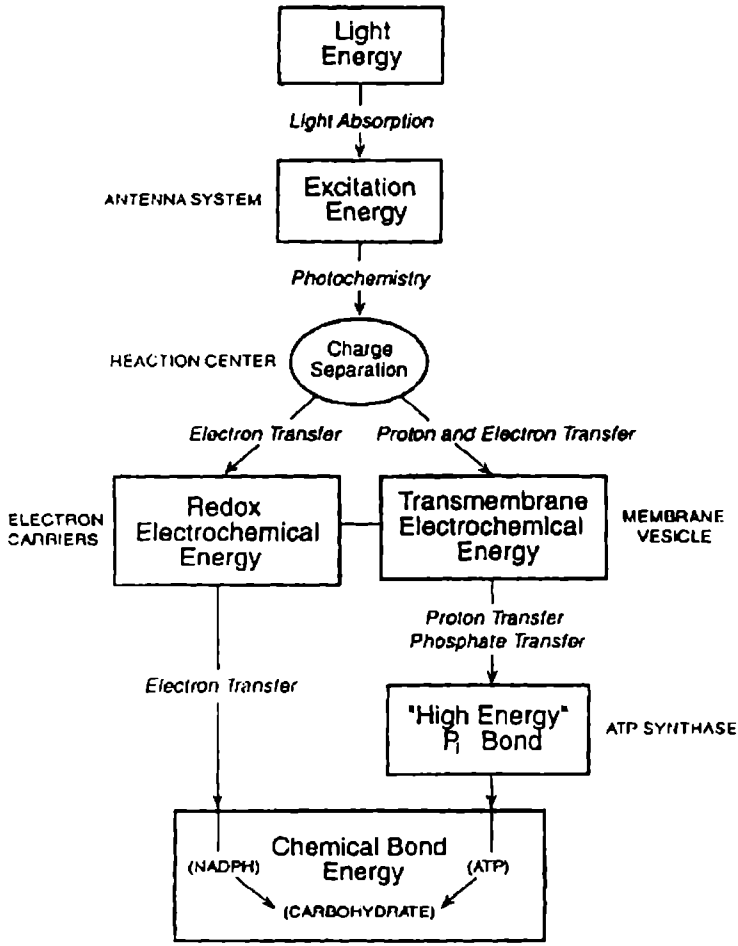


Figure 36: Energy transformation in photosynthesis.

In oxygenic photosynthetic organisms, two different reactio centres, known as photosystem II and photosystem I, wor concurrently but in series. In the light photosystem II feeds electron

to photosystem I. The electrons are transferred from photosystem II to photosystem I by intermediate carrier, the next reaction is the transfer of electrons from a water molecule to $NADP^+$, producing the reduced form NADPH. In the photosynthetic process, much of the energy initially provided by light energy is stored as redox free energy (a form of chemical free energy) in NADPH, to be used later in the reduction of carbon, in addition electron transfer reactions concentrate protons inside the membrane vesicle and create an electric field across the photosynthetic membrane. In this process the electron transfer reactions convert redox free energy into an electrochemical potential of protons. The energy stored in the proton electro-chemical potential is used by a membrane bound protein complex (ATP-synthase) to covalently attach a phosphate group to adenosine diphosphate (ADP), forming adenosine triphosphate (ATP). Protons pass through the ATP-synthase protein complex that transforms electrochemical free energy into a type of chemical free energy known as phosphate group-transfer potential (Klotz, 1967). The energy stored in ATP can be transferred to another molecule by transferring the phosphate group. The net effect of light reaction is to convert radiant energy into redox free energy in the form of NADPH and phosphate group-transfer energy in the form of ATP. In the light reactions, the transfer of a single electron from water to $NADP^+$ involves 30 metal ions and 7 aromatic groups. The metal ions include 19 Fe, 5 Mg. 4 Mn and 1 Co. The aromatics include quinones, pheophytin, NADPH, tyrosine and a flavoprotein. The NADPH and ATP formed by the light reactions provide the energy for the dark reactions of photosynthesis, known as Calvin cycle or the photosynthetic carbon reduction cycle.

CHLOROPLAST STRUCTURE AND ORGANISATION

In plants the photosynthetic process occurs inside chloroplasts, which are organelles found in certain cells. Chloroplasts provide the energy and reduced carbon needed for plant growth and development, while the plant provides the chloroplasts with carbon dioxide, water, nitrogen, organic molecules and minerals necessary for the chloroplast biogenesis. Most of the chloroplasts are located in specialised leaf cells, which often contain 50 or more chloroplasts

per cell. Such chloroplast is defined by an inner and an outer envelope membrane and is shaped like a meniscus convex lens that is 5-10 microns in diameter (Figure 37), although many different shapes and sizes can be found in plants. 'The inner envelope membrane acts as a barrier, controlling the flux of organic and charged molecules and out of the chloroplast. Water passes freely through the envelope membranes, as do other small neutral molecules like carbon dioxide and oxygen. The chloroplast have some DNA necessary for their assembly, but much of the DNA necessary for their biosynthesis is located in the cell nucleus.

The organisation of the chloroplasts can be described as a group of stacks of chapati bread (pancakes), known collectively as *grana,* interconnected by non-stacked membranes that protrude from the edges of the stacks, which are referred to as *stroma.* It is not known why the photosynthetic membranes form such a convoluted structure.

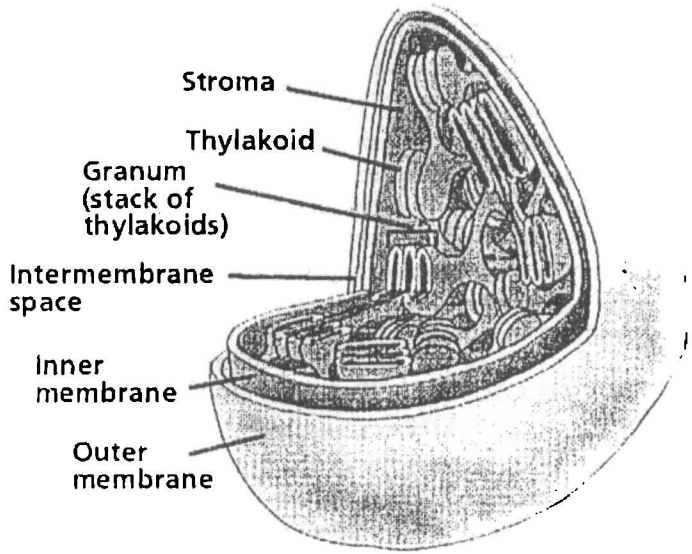


Figure 37: Diagrammatic representation of the electronmicrographic structure of a chloroplast.

Inside the chloroplasts is a complicated membrane system, known as photosynthetic membrane (or thylakoid membrane), that contains most of the proteins required for the light reactions. The proteins required for the fixation and reduction of carbon dioxide are located outside the photosynthetic membrane in the surrounding aqueous phase. The photosynthetic membrane is composed mainly of glycerol lipids and proteins. The glycerol lipids are a family of molecules characterized by a polar head group that is hydrophilic and two fatty acid side chains that are hydrophobia. In membranes, the lipid molecules arrange themselves in a bilayer, with the polar head toward the water phase and the fatty acid chains aligned inside the membrane forming a hydrophobia core (Figure 38).

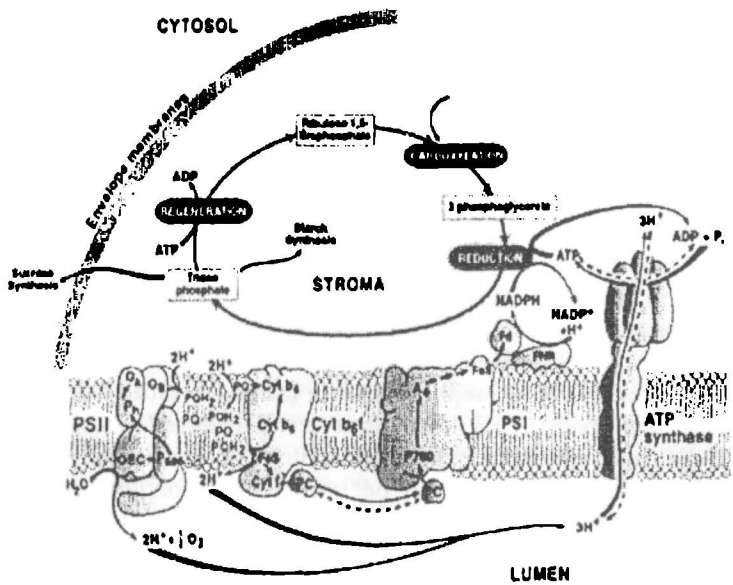


Figure 38: Model of photosynthetic membrane of plants showing four major protein complexes and electron transport component and the ATP synthase enzyme.

STAGES OF PHOTOSYNTHESIS

Photosynthesis is a two stage process. The first stage is the light dependent process (*light reaction*), requires the direct energy of light to make energy-carrier molecules that are used in the second stage. The light independent process (*or dark reaction*) occurs when the products of the light reaction are used to form C - C covalent bonds of carbohydrates. The dark reactions can usually occur in the dark, if the energy carrier from the light process are present. The light reactions occur in the *grana* and the dark reactions take place in the *stroma* of the chloroplasts.

The initial election transfer reaction in the photosynthetic reaction centre sets into motion a long series of reduction-oxidation reaction, passing the electron along a chain of cofactors and filling up the *"electron hole"* on the chlorophyll. All photo synthetic organisms that produce oxygen have two types of reaction centres, named photosystem II and photosystem I (PS II and PS I), both of which are pigment/protein complexes that are located in specialised membranes called thylakoids. PS II is the complex where water splitting and oxygen evolution occur. Upon oxidation of the reaction centre chlorophyll in PS II, an electron is pulled from a nearby aminoacid (tyrosine) which is part of the surrounding protein, which in turn gets an electron from the water-splitting complex. From the PS II reaction centre, electrons flow to free electron carrying molecules (plastoquinone) in the thylakoid membrane, and from there to another membrane-protein complex, the cytochrome complex. The other photosystem PS I, also catalyses light induced charge separation in a fashion basically similar to PS II. However, in PS 1 electrons are transferred eventually to NADP (nicotinamzzzxid adenosine dinucleotide phosphate) the reduced form of which can be used for carbon fixation. The oxidized reaction centre chlorophyll eventually receives another electron from the cytochrome complex. Therefore, electron transfer through PS II and PS I results in water oxidation (producing oxygen) and NADP reduction, with the energy for this process provided by light (2 quanta for each electron transported through the whole chain). A schematic overview of these processes have been shown in Figure 39.

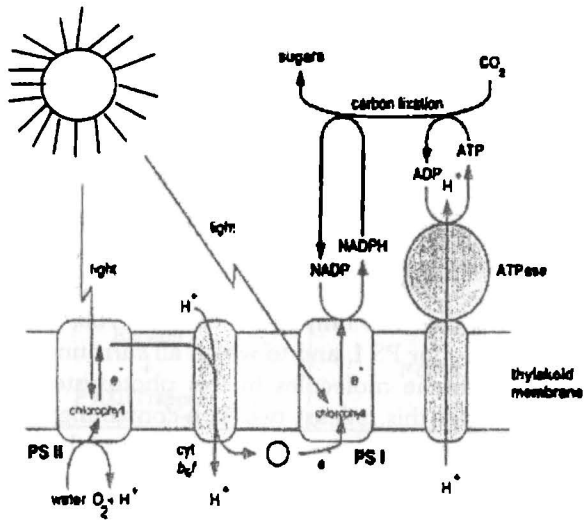
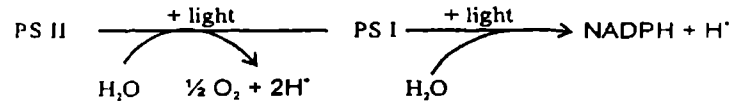


Figure 39: Overview of photosynthetic processes as they occur in plants, algae, and cyanobacteria.

LIGHT REACTION

Emerson and Rabinowitch (1960) explained the emerson enhancement effect and put forward the hypothesis that atleast two photochemical acts are involved in photosynthesis which are preferentially sensitized by different pigments and that each electron must be photoactivated twice on its path from the primary donor water, to the ultimate acceptor Carbon dioxide, Thus, they postulated the existence of two types of Chlorophyll-a, one is associated with the reductant and another associated with an oxidant, and one of these is closer to the accessory pigment then the other - these are termed photosystem I and II. Long red wavelengths are absorbed by only one photosystem, called photosystem I. The second photosystem, photosystem II absorb short red wavelengths (shorter than 690 nm) and for maximum

photosynthesis both systems must function together. The following equation gives the idea how photosystem I and II use energy to oxidise water and transfer two available electrons cooperatively to NADP, thus forming NADPH.



Photosystem I (PS I) contains chlorophyll-a, small amounts of chlorophyll-b and some beta carotene attached to several proteins. One of the chlorophyll-a molecule is special so that it absorbs light near 700 nm, and is called P 700, This P 700 is the reaction centre for PS I, and to which all surrounding chlorophyll a and b, carotene molecules in this photosystem transfer their energy. Besides this, atleast two iron-containing proteins similar to ferredoxin are also present in which each of the four iron atoms in each protein are bound to sulphur atoms - these are called Fe-S protein. The Fe-S proteins are the primary electron acceptors for PS I, that is, electrons are first transferred from the reaction centre of PS I (P 700) to one of these protein.

Photosystem II also contains chlorophyll-a and beta carotene and a little chlorophyll-b is present. The reaction centre is P 680 which is another special chlorophyll-a molecule different from P 700 and other chlorophyll-a molecules, PS II also contains its primary electron acceptor, which is a colourless chlorophyll-a that lacks Mg^{2+}. This molecule is called phaeophytin. There is a quinone called Q which is able to quench the fluorescence of P 680 by accepting its excited electron. Finally P II contains manganese bound proteins, called mangano-proteins. Four Mn^{2+} ions are bound to one or more proteins in PS II and a Cl^- ion bridge two Mn^{2+} together. The mangano-protein is part of the inner side of the thylakoid membrane and is involved directly in water oxidation.

DARK REACTION

Carbon fixing reactions are also known as the dark reactions (or light independent reactions). Carbon dioxide enters the single celled and aquatic autotrophs through general body surface, diffusing into the cells. Land plants have evolved stomata to allow

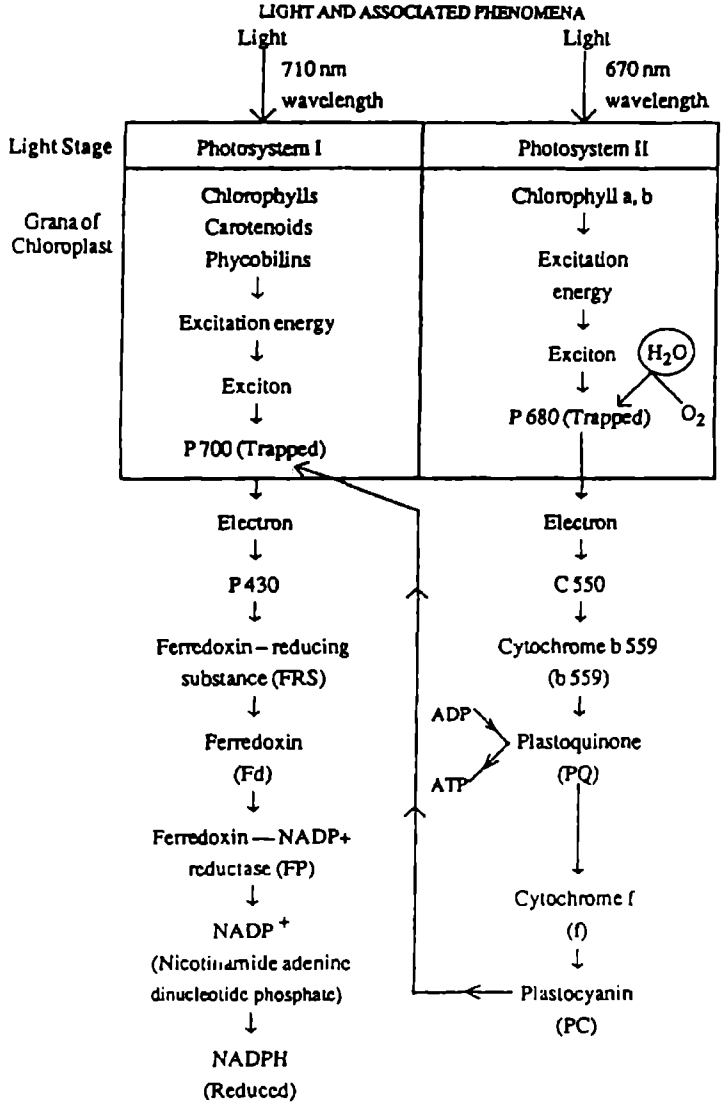


Fig. 40: Non-cyclic light-induced or photosynthetic electron transport.

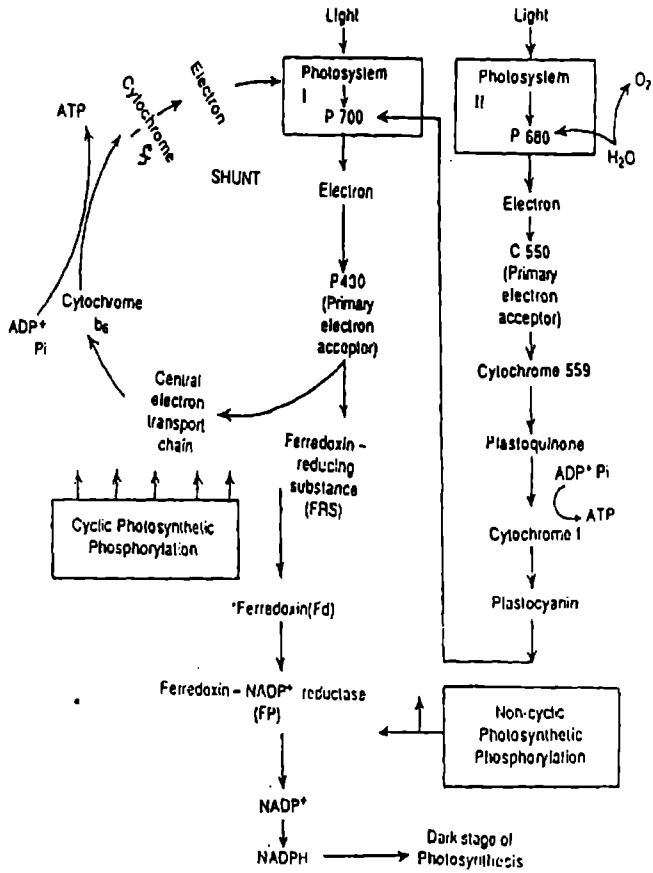


Fig. 41: Cyclic and noncyclic electron flow.

gas to enter the leaf. The *Calvin cycle* (Figure 42) occur in stroma of chloroplasts. Carbon dioxide is captured by the chemical Ribulose Biphosphate (RuBP). The cycle runs 6 times, each time incorporating a new carbon. Ribulose is a 5 carbon sugar and the glyceraldehyde are 3 carbon sugars. Six molecules of carbon dioxide enter the Calvin cycle, eventually producing one molecule of glucose.

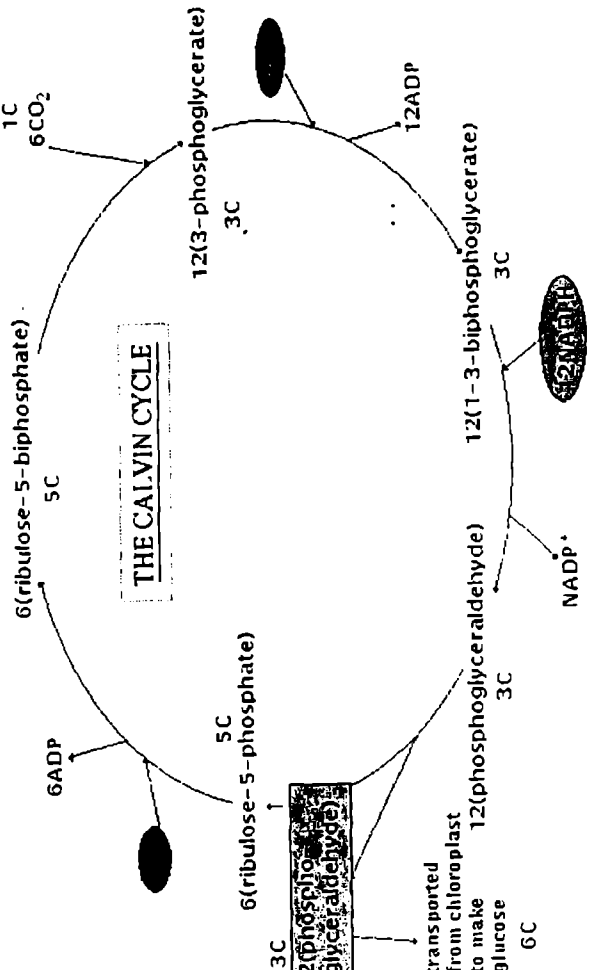


Figure 42: Dark stage showing the Calvin cycle.

The first stable product of the Calvin cycle is *phosphoglycerate* (PGA), a 3 carbon chemical. The energy from ATP and NADPH carriers generated by the photosystems is used to attach phosphates to *phosphorylate* the PGA. Eventually there are 12 molecules of glyceraldehyde phosphate (also known as phosphoglyceraldehyde or PGAL, a 3 carbon), two of which are removed from the cycle to make a *glucose*. The remaining PGAL molecules are converted by ATP energy to reform 6 RuBP molecules, thus start the cycle again. Each reaction in this process, is catalysed by a different reaction specific enzyme.

Those 6 carbon dioxide are reduced to glucose by the conversion of NADPH to $NADP^+$. Glucose can now serve as a building block to make polysaccharides, other monosaccharides, fats, amino acids, nucleotides, and all the molecules living organisms require. The dark reaction is an endergonic process and energy comes from ATP. The hydrogen for reducing carbon dioxide is supplied by the reduced NADP (Figure 43).

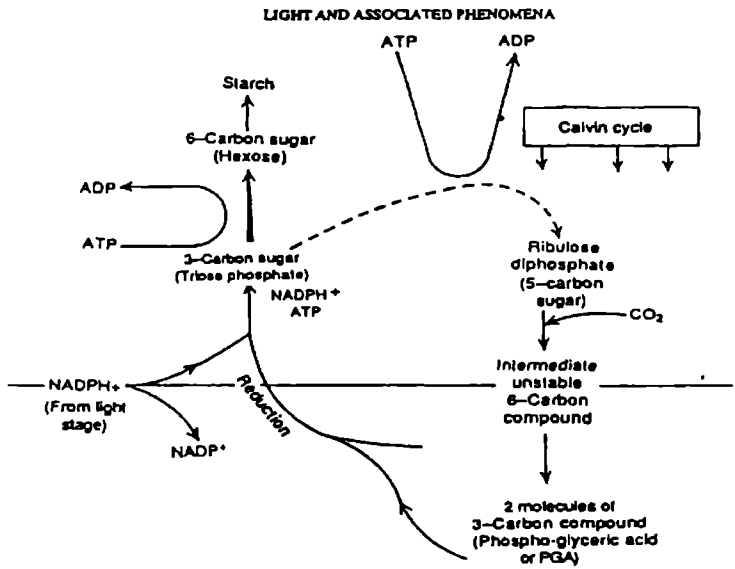


Figure 43: Dark stage showing sugar formation which is cyclical known as Calvin Cycle.

The reduction of Carbon dioxide and subsequent synthesis of carbohydrates takes place in a series of small steps, each controlled by a specific enzyme. In the first step the Carbon dioxide combines with 5 Carbon compound called ribulose diphosphate which serves as Carbon dioxide acceptor for fixing the Carbon dioxide. This combination gives an unstable 6 carbon compound which splits immediately into two molecules of 3 carbon compound called phosphoglyceric acid (PGA). PGA is reduced to 3 carbon sugar called triose phosphate. The hydrogen for reduction comes from NADPH for light reaction and energy from ATP. The triose phosphate converts into 6 sugar (Hexose) through a series of reactions and ultimately formation of starch takes place for storage. The starch is not the only end product of photosynthesis but lipids and aminoacids are also formed. The dark reaction occurs in the stroma of the chloroplasts.

FLOWERING

Flower initiation starts when the stimulation reaches a certain minimal or threshold value and then continues until one or a few flowers are formed. It is a common observation that each species blossoms over a rather restricted period, some regularly flowering very early in the spring, other late in the spring, some in midsummer or late summer and still others in early or late autumn. In majority of higher plants, flower initiation is under the control of the environment, whereby light and temperature affect the rate of flowering. It has been found that photoperiod is a major factor which determines the time of flowering in some plants. Photoperiodism means the response by a plant to the duration and the order of the sequence of light and dark periods, that is the influence of the relative lengths of the day and night on the activities of the living organisms. This was first reported in plants by Garner and Allard (1920).

The relationship between photoperiod and flowering in plants was demonstrated by exposing plants *(Nicotiana tobaccum)* to a brief flash of light in the middle of night. If the procedure was continued for a number of days, the flowering was inhibited. If the plants were to be made to flower early than extra exposure to darkness was required and in some cases one long night was

sufficient. However, reverse results may be observed in some plants, such as light treatment induces early flowering whereas dark treatment delays it. Garner and Allard described three basic response types: (I) short-day plants, (2) Long-day plants, and (3) day-neutral plants. Subsequently, two more types were recognised, these are long short-day plants, and short long-day plants.

(1) Short-day plants

Those require short days, that is they bloom with less than 12 hours of light per day and long nights. These plants flower only when the light period is shorter than a critical length in each 24 hour cycle, for example, tobacco *(Nicotiana tobaccum),* cocklebur *(Xanthium strumerium),* and biloxi *(Glycine max.).*

(2) Long-day plants

Those require long days and bloom with 12 hours of light per day and short nights. These plants flower only when the light period exceeds a certain critical length in each 24 hour cycle which varies on an average about 10 hours, for example black henbane *(Hyosocyamus niger),* spinach *(Spinacea oleracea),* and beet *(Beta vulgaris).*

(3) Day neutral plants

Those which are indifferent to the photoperiod with respect to flowering behaviour and will flower over almost any photoperiod for example, tomato *(Lycopersicum esculentum),* cucumber *(Cucumis sativus),* and balsam *(Impatiens balsamina).*

(4) Long short-day plants

Those plants which require to flower a long photoperiod followed by a short photoperiod for example, *Bryophyllum daigremontianum* and *Cestrum nocturnum.*

(5) Short long-day plants

Those plants which are induced to flower when short photoperiods are succeeded by long photoperiods, for example, winter rye *(Secale cereale),* and candytuft *(Iberis durandii).*

There is a definite demarcation between these classes of plants Table 6 shows the flowering behaviour of some selected plants in response to different photoperiods.

Table 6: Flowering behaviour of a few common plants (Salisbury, 1963).

Long-Day Plants	
Plants specifically requiring long days	
Hyoscyamus niger	Black henbane
Beta vulgaris	Sugar beet
Hordeum vulgare	Winter barley
Spinacea oleracea	Spinach
plants promoted by long days	
Ricinus communis	Castor bean
Lactuca sativa	Lettuce
Short-Day Plants	
Plants specifically requiring short days	
Nicotiana tobacum cv. Maryland Mammoth	Tobacco
Xanthium strumarium	Cocklebur
Pharbitis nil	Japanese morning glory
Chrysanthemum morifolium	Chrysanthemum
Plants promoted by short days	
Gossypium hirsutum	Cotton
Cosmos bipinnatus	Cosmos
Day-Neutral Plants	
Impatiens balsamina	Balsam
Zea mays	Maize
Long Short-day plants	
Bryophyllum daigremontianum	Bryophyllum
Cestrum nocturnum	Night-blooming jasmine
short long-day plants	
Secale cereale	winter rye
Iberis durandii	Candytuft

PERCEPTION OF LIGHT

The pigment phytochrome is responsible for the perception of light in photoperiodism. It is possibly the only photoreceptor in the flowering process. Parker et al (1946) showed that the red region of the visible spectrum extending from 600 to 680 nm is the most effective region in preventing flowering. A second less effective region is in the blue near 400 nm. It has been observed

that red light inhibits flowering of short-day plants and this inhibitory action can be removed by following the red treatment with far-red light (Borthwick et al, 1952). Far-red light reconverts P 725 back to p 665. It seems that short day plants will only flower if a sufficient proportion of its phytochrome is in the p 665 form. The stimulation to flowering could be either a high accumulation of P 665 or a low concentration of P 725. In long day plants, the reverse process is observed, that is, accumulation of P 725, due to long exposure to light, stimulates flowering (Figure 44)

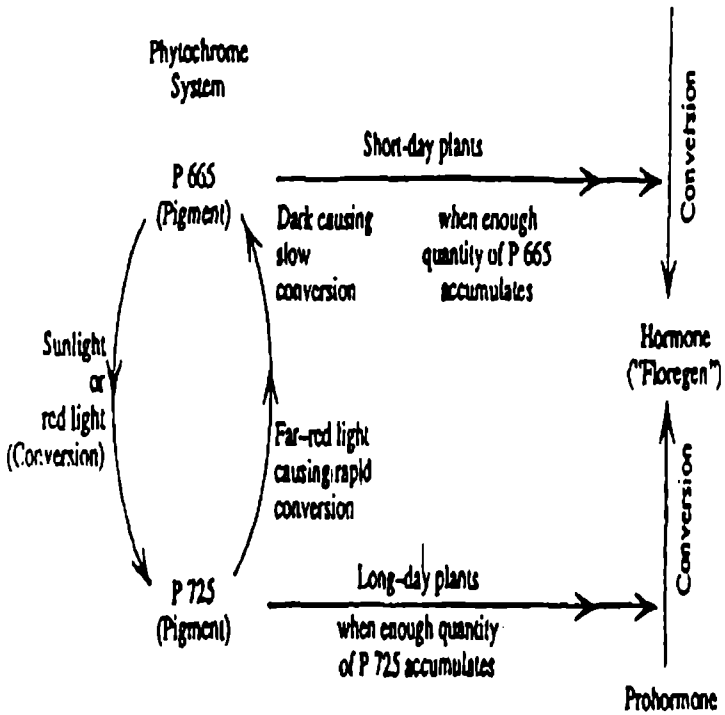


Figure 44: Schematic representation of photoperiodic control of flowering.

The absorption spectra of the two forms of phytochrome, that is P 660 and P 730 overlap considerably (Figure 45). This overlap is the reason why total photochemical conversion is not possible when irradiated with either red or far-red. If we irradiate the system with red light (660 nm), about 75 percent of the total phytochrome can be present as P 730 at photochemical equilibrium. Under irradiation with far-red (730 nm), the proportion of P 730 to P 660 is usually 3 percent or lesser at photostationary state.

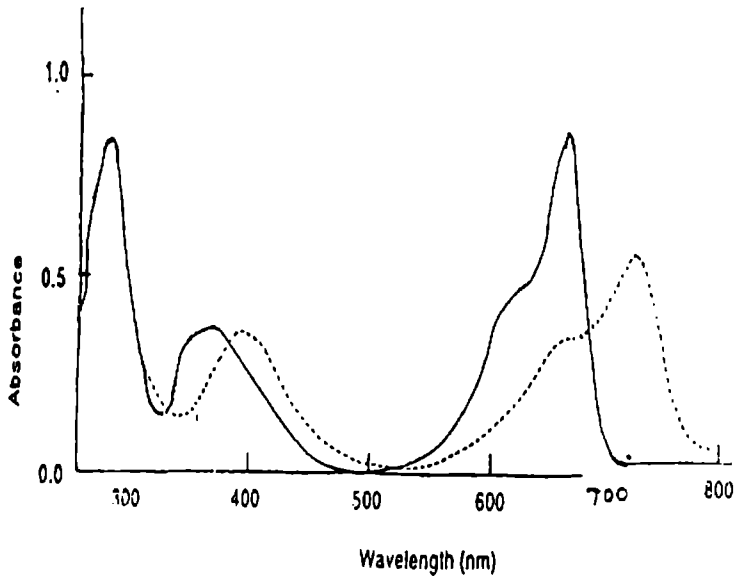


Fig. 45: Absorption spetra of P_{ro} (–) and P_{fr} (–) forms in isolated phytochrome

Besides photochemical conversions, non-photochemical reactions also occur in-vivo (Pratt, 1978). Thus P_{fr} may undergo dark reversion to P_r. Since natural white light acts like red, phytochrome will remain mainly in the pfr form at the end of the day. It has been observed that after several hours of darkness, plants become sensitive to red indicating that pfr is present in a large amount. Thus, it is inferred that pfr is converted spontaneously to Pr in darkness. Phytochrome decay or destruction is also a

dominant irreversible process in a seedling which is thermochemical transformation of Pfr to an inactive form (Figure 46).

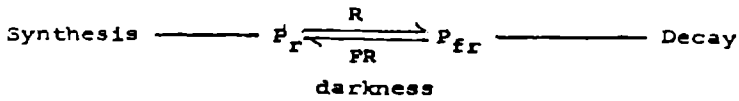


Fig. 46: Phytochrome synthesis and decay model.

SOLARIZATION

It is known for a long time that excessive light causes injury to the green plants, and this damaging effect is known as solarization or photo-oxidation. This damaging effect may be due to transition from one type of photochemical reaction to another. The damage is more pronounced in Carbon dioxide deficiency in the atmosphere. Excessive illumination causes stoppage of oxygen evolution and allows absorption of oxygen and the amount of oxygen absorption depends on the partial pressure of oxygen in the atmosphere. Similar photo-oxydation can be observed if plants are killed by boiling in water or plant juices. The important fact is that the concentration of Carbon dioxide is critical in determining capability of a given light intensity to cause photo-oxidation. This suggests that, in the absence of suitable concentration of substrate, the light energy absorbed by chlorophyll, is used in a less specific manner to oxidise the constituents of cells. If this oxidation proceeds further without any inhibition, the structural components of the leaf, that is, proteins and carbohydrates will be involved in the process and ultimately the damage will be permanent. Non-specific photo-oxidation normally runs parallel with, photosynthesis, a certain amount of the intermediate products of photosynthesis in the plant being thereby wasted, but not sufficient to prejudice the normal metabolism; during Carbon dioxide deficiency, however, the concentration of intermediates is reduced and the photooxidation process becomes dangerous, attacking the vital structures. The carotenoids by directing light energy from this destructive process, may protect photosynthetic plants and bacteria.

PHOTODYNAMISM

It is the sensitization of the organism to light involving photo-sensitized oxidation. Some dyes (for example eosin or methylene blue) when subjected to radiation undergo the sensitization:

$$D + hv = D^*$$

where D = Dye molecule before exposure
hv = Energy of quantum of light
D* = Molecule of dye after it has been activated

The activated dye (D*) molecule may now react with a molecule of its substrate X to form a reactive substrate molecule Xr thus:

$$D^* + X \rightarrow Xr + D$$

This reactive molecule (Xr) may now be oxidised to the molecule Xox

$$Xr + O2 \rightarrow Xox$$

In many cases this abnormal type of oxidation may be harmful to the organism, such as:

1. eosin $\xrightarrow{\text{Light}}$ killed
2. A frog muscle + dye (eosin) $\xrightarrow{\text{Light}}$ Light series of twitches $\rightarrow$ contracture.
3. Enzymes e.g. invertase, peroxidase and catalase + dye $\xrightarrow{\text{Light}}$ inactivation.
4. Plasma protein + dye $\xrightarrow{\text{Light}}$ Denatured, so that the fibrinogen $\rightarrow$ fails to clot after addition of thrombin.
5. RBC + dye $\xrightarrow{\text{Light}}$ Haemolysis.

In all these cases, the effects may be completely prevented by the removal of oxygen from the system either by anaerobiosis or by the addition of reducing agents such as sulphite. The dye stuffs which are commonly used belong to the fluorescent series, for example, eosin, rose-bengal, erythrosine, methylene blue, but chlorophyll and haematoporphyrin are also effective, Photodynamic dyes are more or less fluorescent in visible light for which these are suitable photosynthesitizers. This is because a fluorescent substance can hold its quantum of absorbed light energy for about 10^{-7} - 10^{-9} second and re-emit light of a longer wavelength. In these cases, not all of the light energy is re-emitted, and some of it remains to provide energy for the reactions. In contrast to non-fluorescent substances which lose their absorbed energy by collision

much more rapidly, the excited molecule has thus sufficient time to pass its quantum on to energy-absorbing reactants.

Visible light cannot penetrate to deeper tissues, therefore, photodynamic action can be produced in animals by injection of dye and subsequent irradiation. Rats and mice exhibited symptoms like generalized stimulation of the sensory nerve endings. But it is known that cattle and sheep develop skin diseases after eating certain weeds (for example *Hypericum)* due to the presence of photo dynamically active pigment named *Hypericum.* Other plants for example *Lippia* and *Tribulus* species having toxic substance are not photodynamically active but they damage liver and cause various breakdown products of chlorophyll which appear in the blood stream and phylloerythrin is one of them and photodynamically active. Another plant called buckwheat *(Fagopyrum esculentum)* causes sensitization of animals and the condition is known as *Fagopyrism.*

LIMITING FACTORS IN PHOTOSYNTHESIS

Some factors affect the rate of photosynthesis in plants, as follows

- *Temperature* plays a role in affecting the rate of photosynthesis. Enzymes involved in the photosynthetic process are directly affected by the temperature of the organism and its environment
- Light Intensity is also a *limiting factor,* if there is no sunlight, then the photolysis of water cannot occur without the light energy required.
- Carbon dioxide concentration also plays a factor, due to the supplies of carbon dioxide required in the Calvin cycle stage.

Overall, this is how a plant produces energy which supplies a rich source of glucose for respiration and the building blocks for more complex materials. While animals get their energy from food, plants get their energy from the sun.

Chapter 3

کارلیـــله کـــه ردیبوکان

MOLECULAR INTERACTION

Matter is made up of atoms. Atoms are composed of a nucleus (that is made up of protons and electrons) and one or more electrons. The nucleus is "positively" charged, while electrons are "negatively charged", A neutron is made up of an electron and a proton. A proton is made up of atleast three quarks. Electrons are very small particles, protons are larger than electrons, neutrons are about the same mass as protons, since they are made of a proton and electron.'

The elements that life is primarily made up of are: Carbon (C), Hydrogen (H), oxygen (0), Nitrogen (N), and phosphorous (p). Many other elements can be found in trace amounts in most organisms (for example metals).

The atoms form molecules by various interactions and bonding; covalent interactions, ionic interactions, hydrogen interactions, hydrophobia interactions, and van der Waals interactions.

The ability of cells to adhere to other cells or to extracellular matrix indeed plays a critical role in numerous processes, including embryogenesis, development, hemostasis, inflammation, and immune responsiveness, with the elucidation of specific molecular interactions responsible for adhesion has come the realization that these interactions are complex, well orchestrated and under sophisticated control. For example, the migration of leucocytes from blood vessels into a time (Figure 47) involves a cascade of steps. The loss of an adhesion interaction may result in disease.

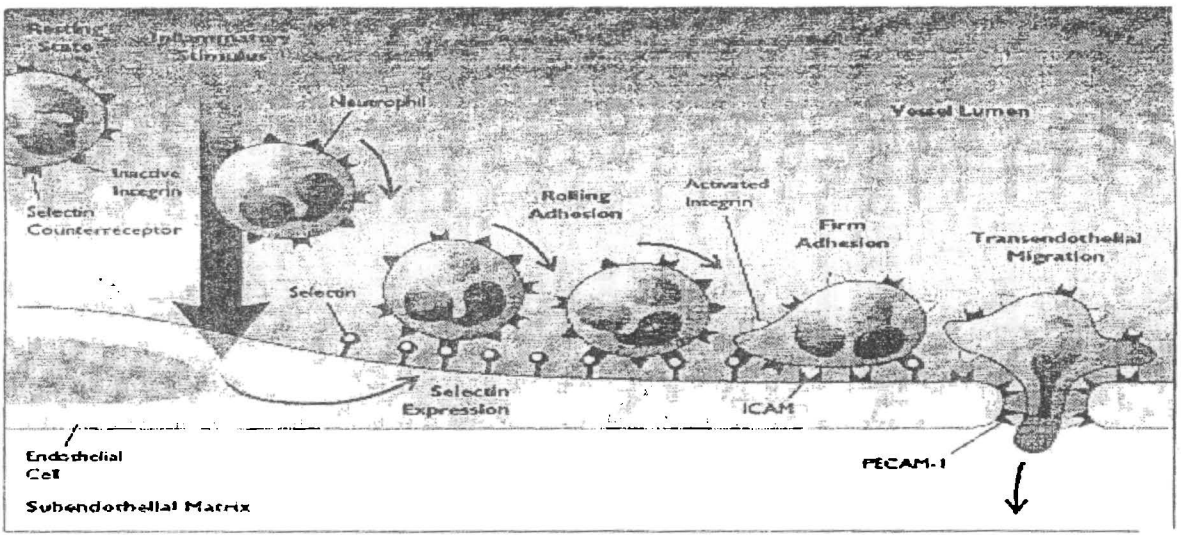
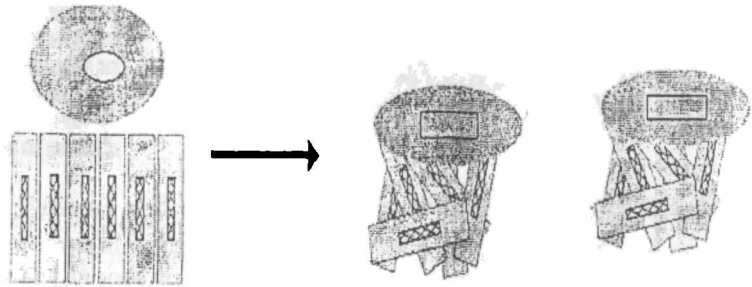


Figure 47: Cell-la-cell adhesions that enable a neutrcphul to leave the circulation begin with both the neiilropnit and Ihe vascular ereothelium in a resting, noninlsracfive slate. Activated by an inflarnmalcHy stimulus, the endolneliuni expresses seleciins, whose binding to Iheir receplors on initiates a tolling adhesion ol neutropdils to Ihe vessel's lurninal wall. The neulropttls adivaie their integrins, which b«nd to erxyothelial ICAMs. permilling a lirmer, stationary acfresion. Transep-ithefial migration may be guided by further adhesive interactions, partials involwng molecules such as PECAM-1, which endolhelial cells express at jitter-cellular junctions.

The biomolecular interactions in membrane systems can occur in a variety of ways:

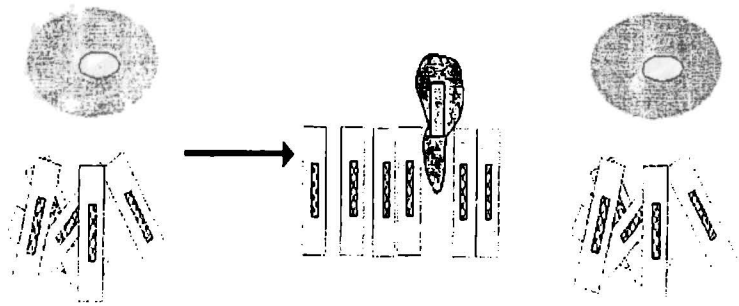
LARGE LIGAND BINDING: SCENARIO-II

Schematic representation of macromolecule binding to an ordered molecular film, resulting in both mass increase and conformational changes in both the macromolecule and the film that decrease orientational order.



LARGE LIGAND BINDING: SCENARIO-II

Schematic representation of a macromolecule binding to a disordered molecular film resulting in both mass increase and conformational changes that increase orientational order in the film.



● *Small Ligand Binding: Scenario-I*

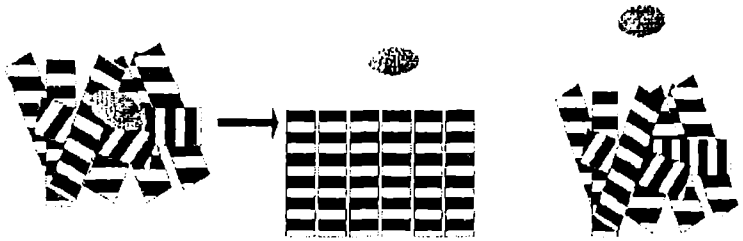
A schematic representation of small molecule binding to an ordered molecular film, resulting in conformational changes that **decrease** orientational order in the film, with **negligible** mass changes.

Example:

- *Small Ligand Binding: Scenario-II*

A schematic representation of a small molecule binding to a disordered molecular film,resulting in conformational change that **increase** orientational order in the film, with **negligible** mass changes.

Example:

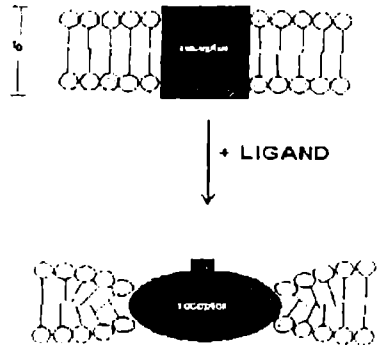


- *Structural Changes upon Receptor-Ligand Binding: Scenario-*

A Schematic representation of an integral membrane protein in a lipid bilayer which induces **negative** membrane curvature upon ligand binding due to protein conformational changes.

This results in a **decrease** in membrane thickness as shown below. Example:

Model for sturctural changes occuring upon lingand Binding to receptor



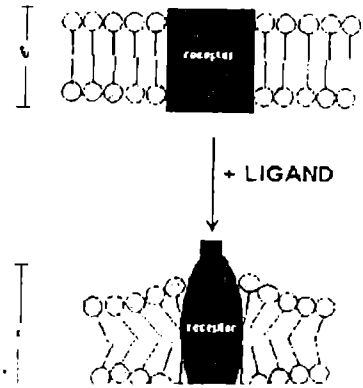
Structural Changes upon Receptor-Ligand Binding: Scenario-II

Schematic representation of an integral membrane protein in a lipid bilayer which induces **positive** membrane curvature upon ligand binding due to protein conformational changes.

This results in an **increase** in membrane thickness.

Model for structural changes occurring upon ligand binding to receptor

Today, the specific cell-surface proteins involved in mediating



adhesion have been classified on the basis of structural or functional similarities into many groups including integrins, cadherins, selectins, and immuno-globulin superfamily members. Some are involved in homotypic adhesion events, such as the recruitment of platelets at a cite of plug formation. Others take part in heterotypic adhesion, for example between leucocytes and vascular endothelial cells. Among the adhesion molecules, integrins are the most important group known to participate in all types of adhesion. While serving a fundamental role in cell-to-cell adhesion, they also appear to be the primary mediator of cellular adhesion to extracellular matrix.

INTRAMOLECULAR AND INTERMOLECULAR FORCES

Types of Molecular Forces: The nature and structure of a substance determine the type and strength of molecular forces that operate within it. We divide such forces into two broad categories. Intramolecular forces operate within the molecules or fundamental units of a substance. These are the covalent bonds within a molecule of a molecular substance, the forces between ions in an ionic compound, and the forces between atoms in a metal. Intermolecular forces operate between, rather than within, the molecules of a covalent substance; the atoms of a monatomic element; or the ions of one substance and the molecules of another. As a rule, intramolecular forces are much stronger than intermolecular forces. Both types of force, however, give rise to potential energy and a potential well. A listing of the important forces, arranged within these categories, follows:

Intramolecular Forces

- Attraction between positive and negative ions in a crystal of an ionic compound.
- Covalent bonds linking atoms in a network structure, such as that of silicon dioxide (sand).
- Covalent bonds in molecular substances.

  Metallic bonds

Intermolecular Forces

- Attractions exerted by one molecule of a molecular substance on another, such as the force of attraction between water molecules in ice.
- Attractions between atoms of the noble gas elements, helium through radon.
- Attractions between molecules of one substance and molecules of another, as when two liquids are mixed, or a molecular solid such as sugar is dissolved in a liquid.
- Attraction between molecules of one substance and ions of another, as when an ionic compound dissolves in a liquid.

THE ATTRACTIVE AND REPULSIVE FORCES

These forces are fundamentally electrostatic in nature that bind (1) electrons to nuclei in atoms, (2) atoms to atoms in molecules, and (3) molecules to molecules in liquids and solids.

Attractive forces are interactions between cations and anions, which are functional groups with formal charge. Charge-charge interactions are very strong, and cause the vapour pressure of Sodium chloride, for example, to be very low. '-the coordinates of Sodium chloride are located here. These interactions, the attractive force between the sodium cation and chloride anion in Sodium chloride, are frequently called "electrostatic interactions" However, all molecular interactions are inherently electrostatic in nature, so it is better to describe them as attractive force. For example, the amino acid aspartic acid (an anion) and lysine (a cation) engage in attractive interactions.

The repulsive forces are very short range in nature, and are only important when atoms are in very close proximity. 'The repulsive force rises sharply as distance decreases. It is reasonable to think of atoms as hard sphere with radii and surfaces. When two atoms approach each other their surfaces make contact when their distance reaches the sum of their radii (here we assume a lack of bonding interactions). The same principle applies to interactions between molecules although the shapes are more complex. The smallest distance between two non-bonding atoms is the sum of the radii of the two atoms.

The radius of carbon is evident from the spacing between the layers in graphite, The distance between atoms in different layers of graphite is never less than twice the radius of carbon (2 x 1.7 = 3.4 A). The atoms within a graphite layer are covalently linked and so are in violation of the van der Waals radius.

STRONG AND WEAK INTERACTIONS

It is convenient to divide the interactions into two classes: strong, and the weak. The yardstick by which we measure whether an interaction is strong or weak is thermal agitation. If the interaction is strong, the two groups held together by it will be stable against the random thermal bombardment of the surrounding medium. Two groups held together by a weak force will not have this stability, but it may be possible for many of these weak interactions to join in concert and so make the equivalent of a strong interaction. The backbone or primary structure of a large molecule, such as protein, is held together by strong interactions. The relative configuration or secondary structure of this backbone, whether it be an extended chain or a helix, is determined by weak interactions. When a protein is denatured by heat, the weak bond give way first, Thus, the secondary and tertiary structure of the protein is lost, and the characteristic position of the polypeptide chain reverts to a purely random one, although the chain itself stays intact.

It is of course, impossible to measure directly the force between two small objects such as atoms or even molecules, but in many cases we can measure the energy of interaction between the two. Two atoms which attract each other will loose potential energy, and by various physical arid chemical means it is possible to measure this energy loss. Hence, the primary observable quantity is the *interaction* energy.

The molecular interactions can be strong interaction (Such as covalent interaction), or weak interaction (such as ionic interactions, hydrogen interaction, hydrophobia interaction, or van der Waals interactions).

COVALENT INTERACTIONS (BONDS)

They provide the strong bonds that holds the atoms of biopolymers together, Covalent bond energies are on the order

of 50 to 110 k cal/mole between two atoms, in which one pair of electrons are shared. A double covalent bond is one where two pairs of electrons are shared. Covalent bonds are the bonds that keep atoms together in organic molecules. Carbon can form upto four covalent bonds; nitrogen three; oxygen two; and hydrogen only one. Some examples are:

C–C, C=C, C–H, C–O, C=O, O–H

The simplest example of a strong interatomic force is that shown by the interaction of two hydrogen atoms in forming a hydrogen molecule. As two hydrogen atoms approach each other, the spin of the two electrons may be either parallel or antiparallel. Moreover, because one electron cannot be distinguished from another, it is not correct to say that electron number 1 is always associated with nucleus A and electron 2 is associated with the other nucleus. Both electrons are associated with both nucleic Because the electrons are associated with two nuclei, they have a larger region of space which they may occupy and enjoy, The potential uncertainty is greater than if the two atoms were separated. The average momentum and kinetic energy of the combined molecule may be lower than for the separated atoms.: the two atoms approach each other from a great distance, the kinetic energy decreases because of the sharing and exchange of electrons. If the nuclei get too close together, the kinetic energy rises because the system is confined to small volume.

More complex atoms, such as Carbon, Oxygen and Nitrogen can also interact with one another. In these cases there may be more than one pair of interacting electrons. The dissociation energies for these covalent bonds may be quite large. They too can be expressed in terms of the energy which would have to be put into the molecule in order to put the nuclei apart.

IONIC INTERACTIONS

They are much weaker interactions. Ionic bond energies are 5 k cal/mole in comparison to that of covalent bonds, and are formed when one atom gives up an electron to the other. An ion is atom that has either lost or gained atleast one electron. Table salt (Sodium chloride) when dissolved in water, forms ions that is sodium ions (Na^+) and chloride ions (Cl^-).

The ionic bond represent an extreme case of the covalent bond. The electrons are shared in a very asymmetrical fashion so that the majority of the force may be regarded as purely electrostatic, as in the case of Sodium chloride. A decrease in this electrostatic force, as when the molecules are placed in a medium of high dielectric constant such as water, can cause dissociation.

HYDROGEN INTERACTIONS

They are also weak interactions (bonds) with energies from 4 to 5 k cal/mole. A hydrogen bond forms when a molecule that has a hydrogen atom is attracted to another atom (for example an oxygen atom) on an adjacent molecule. The two atoms have to be very close for weak interactions to take effect.

Hydrogen bond strengths form a continuum, Strong hydrogen bonds of 20 - 40 k cal/mol, generally formed between charged donors and acceptors, are nearly as strong as covalent bonds, weak hydrogen bonds of 1 - 5 k cal/mole, sometimes formed with carbon as proton donors are no stronger than van der Waals interaction, moderate hydrogen bonds, which are the most common are formed between neutral donors and acceptors are from 5-15 k cal/mole.

Hydrogen bonding in biological molecules is important because proteins, nucleic acids, and polysaccharides contain many groups which may participate in hydrogen bonding and which may be in a specific order in the primary structure of such molecules, Thus if the molecular energy in those molecules is to be minimized, the number of hydrogen bonds must be maximized without increasing the energy of, or distorting other types of bonds. Obviously, hydrogen bonding is one of the major determinants of secondary and tertiary structures. Hydrogen bonding determines the structures such as pleated sheet and α-helix of proteins and how it is crucial element in determining the pairing of the purine and pyrimidine bases in the double helix of DNA. Life would be very difficult without hydrogen bonds.

Hydrogen bonds are not necessarily two-centred. They can be three-centred and four-centred in geometry. In biological systems, hydrogen bonds are frequently cooperative. For example, in the hydrogen bonded system below, hydrogen bond 1 increases both

the acidity of the hydrogen, and the basicity of the oxygen, in hydrogen bond 2.

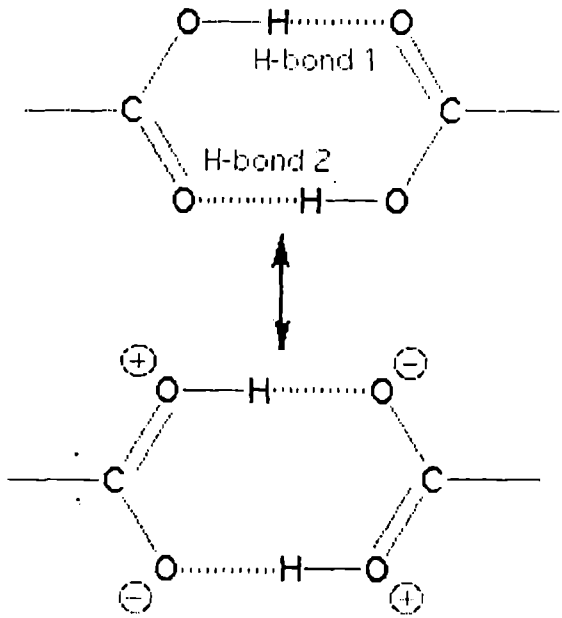


Figure 48: Resonance stabilisation of the hydrogen bonds of an Acetic acid Dimer.

HYDROPHOBIC INTERACTIONS

They are also weak interactions with an energy of 3 k cal/ mole; which occur when non-polar molecules (such an lipids or fats) are excluded by polar molecules such as water. "Polar" means that the molecule (as a whole) has some uneven distribution of electrons – the molecule then, is said to have a "dipole" (that is two.poles). At a result, the molecule is more negatively charged on one side and in effect, the molecule is "polar". An analogy could be the earth – - it has a North and South pole, Water molecules from Hydrogen bonds with other water molecules (since water molecules are polar and they are made up of Hydrogen and oxygen), which overcomes and therefore excludes non-polar

molecules from the polar molecules. Fats and oils are non-polar molecules that form little globs called "micelles" in water. You may have seen this in chicken soup, where you have little globs of fats on the liquid surface, This is a result of hydrophobic interaction.

Most living organisms are around 80 percent by weight water. Water is a reactive substance with unusual properties and is a crucial determinant in the structure and properties of other cellular components. Water is a uniquely powerful solvent power for ions and of equal importance in a uniquely weak solvent for non-polar substance.

The forces of attraction between water molecules in the liquid state are usually high. The melting point, boiling point, heat of vaporisation, heat of fusion, and surface tension of water are higher than those of similar substances. For example, the heat of vaporisation of water (540 cal/g) is over twice that of methanol (263) and nearly tentimes that of chloroform (59).

Electrostatic interactions are highly attenuated in water. The attractive force between two oppositely charged ions in solution is inversely proportional to the dielectric constant of the solvent. The dielectric constant of water (80.0) is over twice that of methanol (33.1) and over five-times that of ammonia (15.5).

Molecular structure of water in the crystalline state is tetrahedral in shape with either a hydrogen atom or a lone pair of electrons at each apex of the tetrahedron. Oxygen, which is highly electronegative, withdraws electrons density from the hydrogen atoms. The charge distribution of a water molecule have been shown in Figure 49.

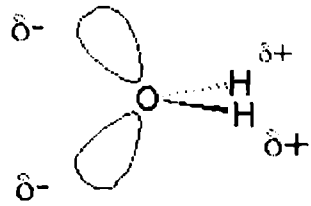


Figure 49: A Water Molecule

X-rays and neutron diffraction of a crystalline ice shows that each water molecule engaged in four hydrogen bonds with intramolecular oxygen-oxygen distances of 2.76 Å. Each oxygen atom is located at the centre of a tetrahedron formed by four other oxygen atoms. Such hydrogen atom lies on a line between two oxygen atoms and forms a covalent bond to one oxygen (bond length: 1.00 Å) and a hydrogen bond to the other (hydrogen bond length: 1.76 Å). The hydrogen atoms are not located midway between oxygen atoms (Figure 50).

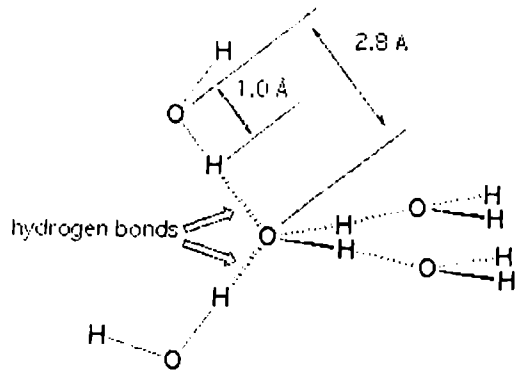


Figure 50: Hydrogen bonding degrees of freedom in between water molecules.

There are many degrees of freedom in hydrogen bond donor/accepter relationship which are inter-converted by cooperative rotations. Ice is rather distorted in this respect. The water molecules in the crystalline state are not closely packed and there is empty space within the crystal. The loose packing arrangement is the reason that water decreases in density upon freezing (i.e., ice floats).

The molecular structure of water in liquid state is not nearly as ordered as the diagram above might indicate.

At 0 °C a time averaged water molecule is involved in around 3.5 intermolecular hydrogen bonds. Some of them are three- and four- centred.

Hydrophobic structures are those which are highly soluble in non-polar solvents but only slightly soluble in water. This definition excludes substances which are generally insoluble because of strong intermolecular cohesion. Hydrophobic substances are non-polar and non-hydrogen bonding.

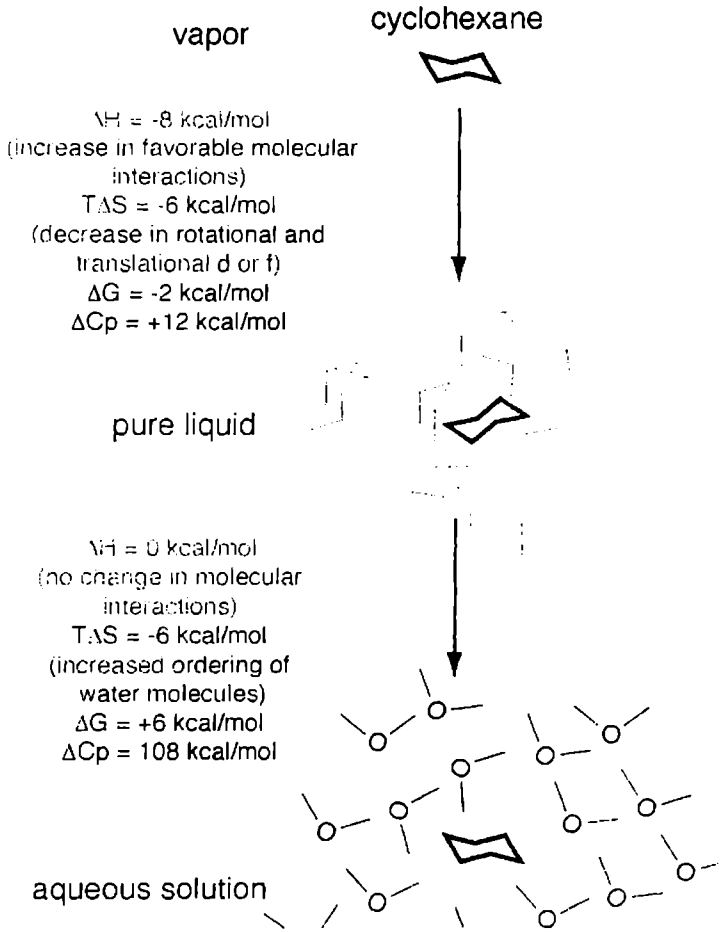


Figure 51: Consequences of conversion of cyclohexane vapour to neat liquid and aqueous phase.

The schematic diagram above illustrates what happens when a hydrophobia substance (cyclohexane) is converted from vapour to neat liquid to aqueous phase. In the first step from vapour to neat liquid there is an increase in intramolecular interactions and a decrease in rotational and translational degrees of freedom. Therefore one expects, and sees, a favourable enthalpy contribution (negative delta H) and an unfavourable entropy contribution (negative delta S) for the condensation. In the second step from neat liquid to dilute aqueous phase, the change in stability conferred from intramolecular interactions is a wash, no gain or loss. But the water loses entropy. Somehow water is more highly ordered in the vicinity of a cyclohexane molecule, therefore, for this step delta S is negative and so is delta G.

In the aqueous phase a region of relatively low entropy (high order) water forms at the interface between the aqueous solvents and a hydrophobic solute (Figure 52).

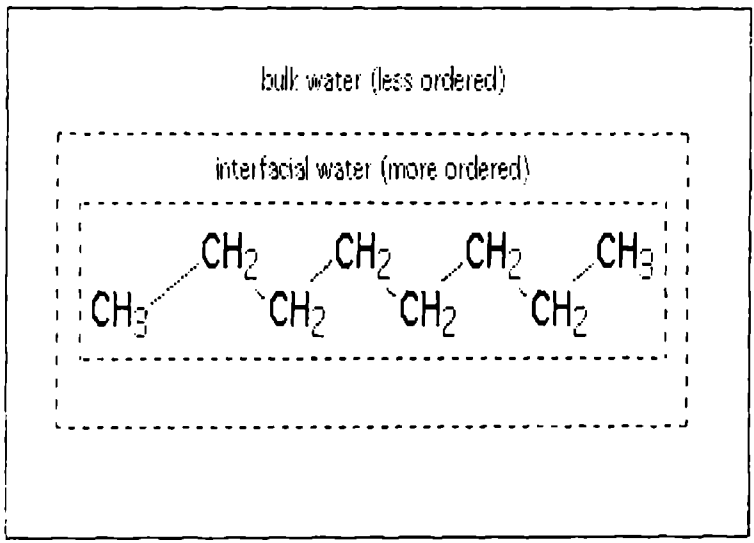


Figure 52: The change in entropy in the aqueous phase of cyclohexane.

The decrease in entropy at the interface arises from the strong intermolecular forces between water molecules, in bulk water, these forces are isotropic (extending in all directions). At the interface these forces are anisotropic because the cyclohexane molecule does not form hydrogen bonds, (thus water at the interface is rotationally and translationally constrained.

When isolated octane molecule aggregate in aqueous solution, the total volume of interfacial water decreases. Thus the driving force for aggregation of hydrophobic substances arises from an increase in entropy of the aqueous phase. The driving force for aggregation does not arise from intrinsic attraction between hydrophobic solute molecules.

If one considers the entropy of the cyclohexane molecules alone, a dispersed solution appears to be of greater entropy, and more stable, than an aggregated state. Similarly, a protein may appear to have greater entropy in a random coil than in a native state, only when the entropy of the aqueous phase is factored into the equation can one understand the separation of water and oil into two phases, and the folding of a protein into a native state.

VAN DER WAALS INTERACTIONS

They are very weak interactions with an energy of less than 1 k cal/mole of which there are two types: (1) dipole interactions, and (2) London dispersion or fluctuating dipole interaction.

DIPOLE INTERACTIONS

In a polar molecule with unlike atoms, the electrons are not distributed equally, therefore uncharged polar molecules can form dipole interactions. The tendency of any atom to pull electrons away from other atoms is characterised by a quantity called electronegativity (Table 7).

Table 7: Electronegativity of atoms.

Atom	*Electronegativity*
H	2.13
C	2.55
M	2.98
O	3.45
S	2.53

In a molecule composed of atoms of various electronegativities the atoms with lowest eloctronegativities hold partial positive charges and the atoms with the greater electronegativities hold partial negative charges (Figure 53).

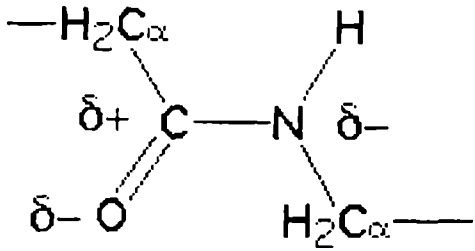


Figure 53: Partial charges within a peptide bond.

Table 8: Partial charge within a peptide bond.

Atom	*partial charge (e)*
N	– 0.36
HN	+ 0.18
Cα	+ 0.06
C	+ 0.45
O	– 0.38
Hα	+ 0.02

The extent of charge separation within a molecule is characterised by the dipole moment (μ). To express dipole moments, charges are expressed in esu's and distances in centimetres. The dipole moment of an electron and a proton separated by 1 Å is given by: $(4.8 \times 10^{\wedge} - 10^{-10}$ esu$)(10^{\wedge}-8$ cm) = $4.8 \times 10^{\prime}-18$ esu cm = 4.8 Debye. The orientation of the dipole moment of a peptide is approximately parallel to the N-H bond and in magnitude is around 3.7 Debye. That is relatively large dipole moment.

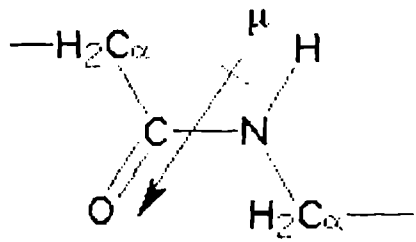


Figure 54: Dipole moment of a peptide.

The dipole moment of HCl is 1.0 Debye, that of CH_3Cl is 1.9 and that of HCN is 2.9. The large dipole moment of a peptide bond should lead one to expect that dipolar interactions are important in protein conformation and interactions. The large dipole of peptide bond can be attributed in part to resonance (Figure 55).

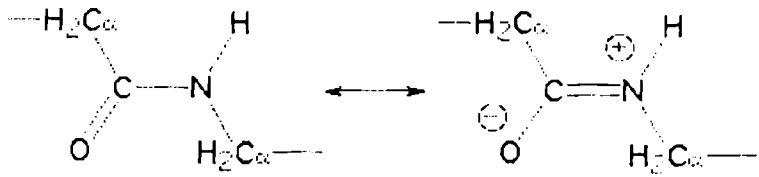


Figure 55: Resonant structure of a peptide

A dipole can interact with point charges (called Charge-Dipole interaction), other dipoles (called Dipole-Dipole interaction), ad can induce charge distribution in surrounding molecules (called Dipole-induced Dipole interaction).

Dipole-Dipole interactions

The interaction energy between two dipoles can be either positive or negative and can be calculated with Coulombs Law. Listed below are the energies of interaction for two dipoles with moments of 1 Debye at a distance of 5 Å in a medium of E. Dipole-dipole interactions fall off with 1/r^3.

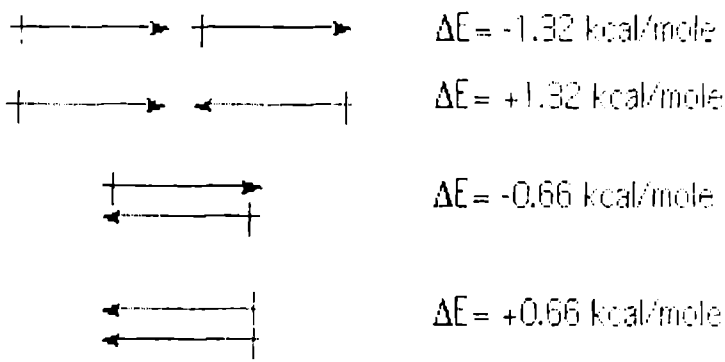


Figure 56: Dipole-dipole interaction.

DIPOLE INDUCED DIPOLE

A molecule with a permanent dipole can induce a dipole in a second molecule that is located nearby in space The strength of the interaction depends on the dipole moment of the first molecule and the polarizability of the second. Molecules with pi electrons, such as benzene and phenyl alanine are more polarisable than molecules without pi electrons (Figure 57). Dipole induced dipole interactions are always attractive and can contribute as much as 0.5 k cal/mole to stabilisation.

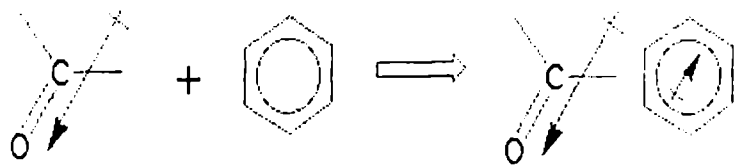


Figure 57: Dipole induced dipole interaction.

LONDON DISPERSION OR FLUCTUATING DEPOLE INTERACTION

Non-polar molecules have a relatively even electron distribution. They are symmetrical, and have a relatively even

electron distribution. They have same types of atoms on either side, and so there is no dipoles. However, nonpolar molecules can become temporarily polar due to electron moment within the covalent bond. A dipole is induced (this is called London dispersion or fluctuating dipole interaction).

Molecules behave like oscillating dipoles. In molecules that are located nearby to each other the oscillators are coupled. Coupled fluctuating dipoles experience favourable electrostatic interaction known as dispersive interactions. The strength of the interaction is related to the polarisabilities of the two molecules (Figure 58).

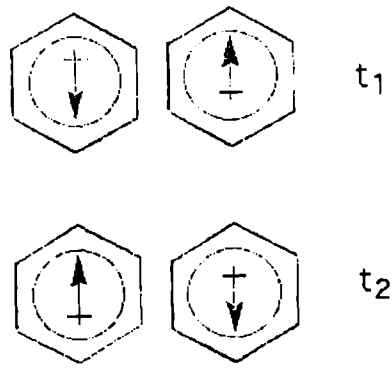


Figure 58: Fluctuating dipole of liquid Benzene at two time points.

Dispersive interactions are always attractive and occur between any pair of molecules, even non–polar molecules. Dispersive interactions provide the only attractive force between molecules in liquid N_2, which boils at 77 K. For a given atom-atom contact the energy of stabilisation provided by dispersive interactions is very small (0.05 k cal/mole). However, the total number of contacts within a protein is generally enormous, so that dispersive interactions can make a large contribution to stability. Fluctuating dipole interactions fall off with 1/r^6.

DNA PROTEIN INTERACTION

DNA-protein interaction occur specifically or non-specifically, in case of non-specific interactions, the sequence of nucleotides

does not matter, as far as the bonding interactions are concerned. Histone-DNA interactions are an example of such interactions, and they occur between functional groups on the protein and the sugar-phosphate backbone of DNA. Specific DNA– protein interactions, however, depend upon the sequence of bases in the DNA and on the orientation of the bases that can vary with twisting and writhing,

These DNA-protein interactions are strong and are mediated by: (1) Hydrogen bonding - can be direct H-bonds, or indirect, mediated by water molecules; (2) ionic interactions -salt bridges, protein side chains - DNA backbone interactions; and (3) other forces - such as van der Waals, and hydrophobic forces.

Nucleic acids in aqueous solutions can be considered rod like polyanions surrounded by inorganic cations (Na^+, K^+ and Mg^+) and polyamines. The high axial density of negatively charged phosphate groups on DNA causes strong radial electric fields. These electric fields lead to steep radial gradients of the surrounding cations concentrations. Theoretical considerations (counter ion condensation) predict that the local concentration of a monovalent cation such as K^+ near the surface of DNA is around 2 Molar. This local concentration (i.e., near the DNA) of K^+ is largely independent of K^+ concentration in bulk solution.

Release or uptake of local cations accompanies processes such as helix–coil and other conformational transitions that change the axial charge density. Release or uptake of local cations also accompanies binding of proteins, and intramolecular compaction or collapse of DNA. These effects explains the dramatic salt dependencies of DNA melting and DNA-protein complexation.

Genomic DNAs are very long molecules. The 160,000 base pairs of T4 phage DNA extend to 54 microns. The 4.2 million base pairs of the *E.coli* chromosome extend to 1.4 millimetres. In biological systems, long DNA molecules must be competed to fit into very small spaces inside a cell, nucleus or virus particle. The energetic barriers to tight packaging of DNA arise from decreased configurational entropy, bending the stiff double helix, and

intermolecular (or intersegment) electrostatic repulsion of the negatively charged DNA phosphate groups. Yet extended DNA chains condense spontaneously by collapse into very compact, very orderly particles. In the condensed state, DNA helixes are separated by one or two layers of water. Condensed DNA particles are commonly compact toroid.s, DNA condensation in aqueous solution requires highly charged cation such as spermine (+4) or spermidine (+3). Divalent cations will condense DNA in water-alcohol mixtures. The role of the cations is to decrease electrostatic repulsion of adjacent negatively charged DNA segments. The source of the attraction between nearby DNA segment is not so easy to understand, one possible source of attraction are fluctuations of ion atmospheres in analogy with fluctuating dipolos between molecules (London forces).

When thinking about the free energy of protein folding (folding is the trasition from denatured state to native state). It is important to understand that the stability of a folded state in a biological system can only be understood as a difference between the free energies of the native and denatured states.

$$\Delta G\text{folding} = \Delta G\text{native state} - \Delta G\text{denatured state}$$

In the native state many of these interactions are intramolecular, for example one part of the protein will form a hydrogen bond with another part of the protein. In the denatured state, interactions are intermolecular (i.e., between the protein and water molecules, cations, and anions). A intramolecular hydrogen bond observed in the native state will be replaced in the denatured state by several hydrogen bonds between the protein and water molecules. The same is essentially true for charge-discharge, dipole-dipole, dipole-induced, and fluctuations dipole interactions.

One of the primary driving forces for protein folding is solvent release. In the transition from denatured to native state, a large number of solvent molecules are released. Release is defined as the conversion from bound state to free state.

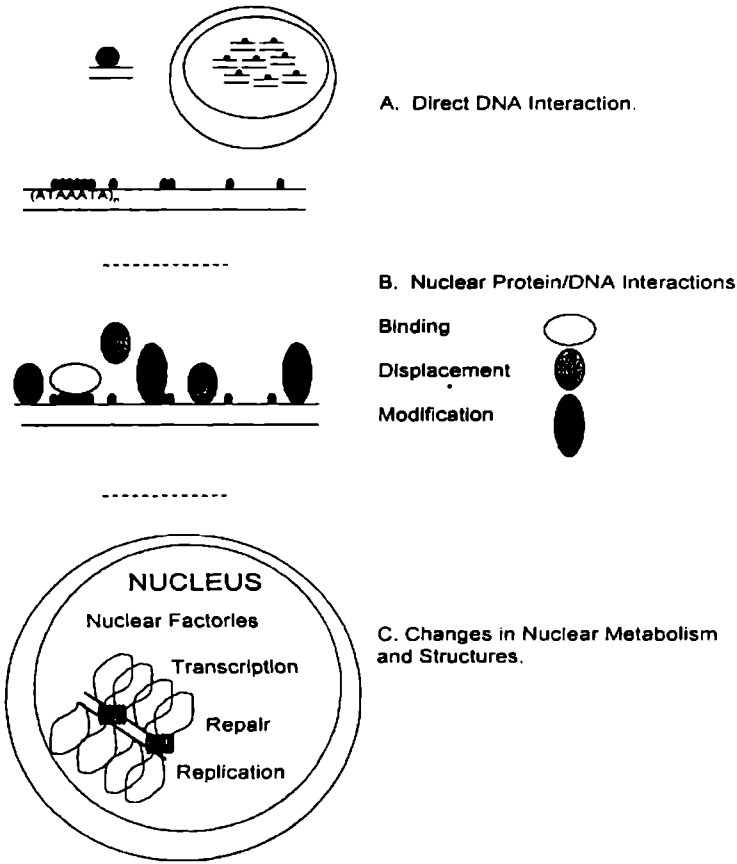


Figure 59: Levels of Drug/DNA Interactions: *A) The first level of interaction is the direct drug interaction with the DNA. Drug binding frequency and sequence preferences are two examples of evaluations of this level. B) The second level involves the disruption of DNA interactions with other molecules that result from the drug/DNA interaction such as DNA/protein interactions. Another possibility (not shown) is the alteration or stabilization of DNA secondary structures. C) The third level is the changes in nuclear activities such as transcription, replication, and repair or disruption in complex ternary structures. Green circles, drug/DNA adduct; black lines, DNA strands, chromosome (with loops); yellow oval, red oval, and pink circle, DNA associated proteins; blue square, protein modification such as phosphorylation or methylation.*

DRUG-DNA INTERACTION

DNA as carrier of genetic information is a major target for drug interaction because of the ability to interfere with transcription (gene expression and protein synthesis) and DNA replication, a major step in cell growth and division. The latter is central for tumorigenesis and pathogenesis.

There are three principally different ways of drug-binding. First, through control of transcription factors and polymerases. Here, the drugs interact with the proteins that bind to DNA. Second, through RNA binding to DNA double helices to form nucleic acid triple helical structures or RNA hybridization (sequence specific binding) to exposed DNA single strand regions forming DNA-RNA hybrids that may interfere with transcriptional activity. Third, small **aromatic ligand molecules** that bind to DNA double helical structures by (i) intercalating between stacked base pairs thereby distorting the DNA backbone conformation and interfering with DNA-protein interaction or (ii) the minor groove binders. The latter cause little distortion of the DNA backbone. Both work through non covalent interaction.

The small ligand drug approach offers a simple solution. The synthesis and screening of synthetic compounds that do not exist in nature, work much like pharmacological ligand for cell surface recepters in excitable tissue, and appear to be more readily delivered to cellular targets than large RNA or protein ligands. The lack of sequence specificity for intercalating molecules, however, does not allow to target specific genes, but rather certain cellular states or physiological and pathological conditions, like rapid cell growth and division that can be selectively suppressed as compared to non growing or slowly growing healthy tissue.

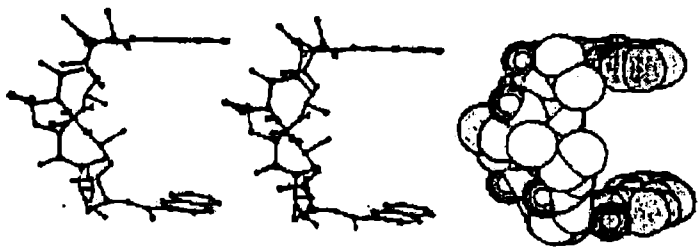
Modelling DNA-ligand interaction of intercalating ligands

The following properties have been identified as important for the successful modelling of ligand-DNA interaction:

- degrees of freedom
- role of base pair sequence
- counter ion effects

- role of solvent ligand-receptor binding
- degrees of freedom

This problem is analogous to that of protein ligand interaction. The major requirement for intercalating agents is the planar aromatic ring structure. This structure fits between to adjacent base pair planes and can have some, although much restricted, rotational freedom within the plane of the ring. The ligand itself may have flexibility of structural parts outside the DNA binding site and may contain more than one intercalating side chain:



The structure of the antibiotic triostin A shows the presence of two quinoxaline (groups to the right; double aromatic rings) units linked through a cyclic peptide structure (centre left) which is stabilized at its center by a cystein pair (disulfhydril covalent bond).

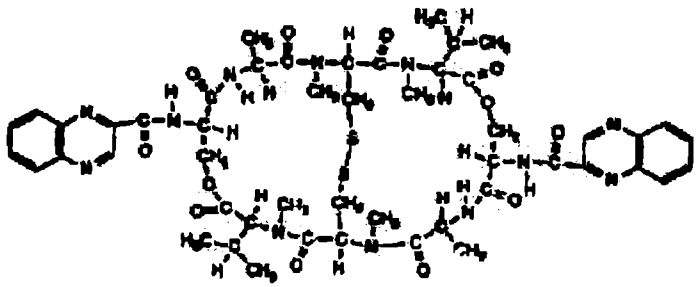


Figure 60: Chemical structure of triostatin A

The space filled side view indicates how the two quinoxaline rings are positioned by the linker peptide in co-planar fashion suitable for intercalating with DNA base pairs. As a rule, the more

intercalating sidechains are linked within a single ligand structure, the stronger the expected binding affinity.

Triostatin A belongs to a family of antibiotics which are characterized by cross-linked octapeptide rings bearing two quinoxaline chromophores. Since the spacing between the chromophores is 3.5A, the intercalation process sandwiches two base pairs between the two quinoxalines. This phenomenon is called *bis-intercalation* and has first been described for echinomycin by showing that *bis-intercalating* drugs cause twice the DNA helix extension and unwinding seen as compared to single intercalating molecule like ethidium. The latter is a chromosphere which is activated by UV light and is used by molecular biologists to label nucleic acids in gel electrophoresis or ion gradient centrifugation.

Role of base pair sequence

Experimental evidence suggests that base pair sequence does not play a large role on the specific nature of most intercalating complexes. As the structure of triostatin A suggests, however, the linker peptide structure may well promote specific interaction with the DNA surface. The major group specific readout sequence of H-bond donor and acceptor could be involved in triostatin A binding. The figure graphically shows the direct readout of the DNA base sequence on a double helical structure.

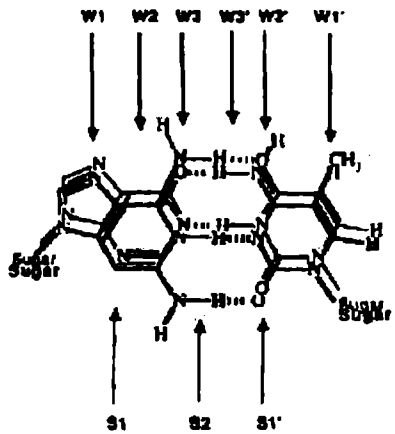


Figure 61: Overlap of AT and GC base pairs

Note: readout sequence of minor (S) and major groove (W) side as they are available for protein interaction.

The following characteristics of non covalent bond formation are associated with the binding sites indicated above:

binding site	*GC base pair*	*AT base pair*
Wl	H-bond acceptor	H-bond acceptor
W2 ;	blank	blank
W3	H-bond acceptor	H-bond donor
W3'	blank	blank
W2' :	H-bond donor	H-bond acceptor
w1' ;	C-H weak hydrophobic	CHS, strong hydrophobic

While the interaction on the major groove side is distinct for the direction of the base pair (e.g. AT vs TA), there is no directionality at the minor groove side.

The molecular basis of specific recognition between echinomycin and DNA is due to the hydrogen bonding between the ligand alanine carbonyl groups and the 2-amino group of guanine. This is consistent with the observation that the preferred binding site is the sequence CG.

Counter ion effect

DNA is a negatively charged polyanion attracting counter ions, positively charged Na+, or Ca++ and Mg++ ions as well as basic residues of proteins. The presence of small counter ion affect drug binding, since the counter ions can screen and shield the negative backbone surface allowing non electrolytes as well as positively charged ligand to interact more strongly with the DNA target. High ionic strength, however, reduces non covalent interaction mediated by hydrogen bonds and electrostatic interactions.

Role of solvent ligand-receptor binding

There are three general classes of interactions that must be considered in solvated ligand-receptor binding.

(a) ligand solvent interaction (e.g. hydration shell), (b) receptor solvent interaction, and (c) ligand-DNA complex with solvent interaction. The three classes basically describe the sequence of events of free ligand interacting with its receptor and the change in overall solvent interaction before and after binding. We have seen that the hydrophobic effect is completely described by this system and the contribution of the entropy of free bulk water is the major driving force of hydrophobic ligand receptor interaction. This type of interaction is found in intercalating substrates because the hydrophobic, aromatic sidechains interactive favourably with the aromatic environment of the base pair stacking. The total amount of surface bound water is reduced in the after complex formation.

Rational for drug design

When a compound intercalates into nucleic acids, there are changes which occur in both the DNA and the compound during complex formation that can be used to study the ligand DNA interaction. The binding is of course an equilibrium process because no covalent bond formation is involved. The binding constant can be determined by measuring the free and DNA bound form of the ligand. Since many of the intercalating substrates are aromatic chromophores, this can be done spectroscopically. Also, DNA double helix structures are found to be more stable with intercalating agents present and show a reduced heat denaturation. Correlating these biophysical parameters with cytotoxicity is used to support the antitumor activity of these drugs as based on their ability to intercalate in DNA double helical structures.

Improvement of anticancer drugs based on intercalating activity is not only focussed on DNA-ligand interaction, but also on tissue distribution and toxic side effects on the heart (cardiac toxicity) due to redox reduction of the aromatic rings and subsequent free radical formation. Free radical species are thought to induce destructive cellular events such as enzyme inactivation, DNA strand cleavage and membrane lipid peroxidation.

Modelling DNA-ligand interaction of minor groove binders

Hairpin minor grove binding molecules have been identified and synthesized that bind to GC rich nucleotide sequences. Hairpin polyamides are linked systems that exploit a set of simple recognition rules for DNA base pairs through specific orientation of imidazole (Im) and pyrrole (Py) rings. The hairpin polyamides originated from the discovery of the three-ring Im-Py-Py molecule that bound to minor groove DNA as an antiparallel side by side dimer.

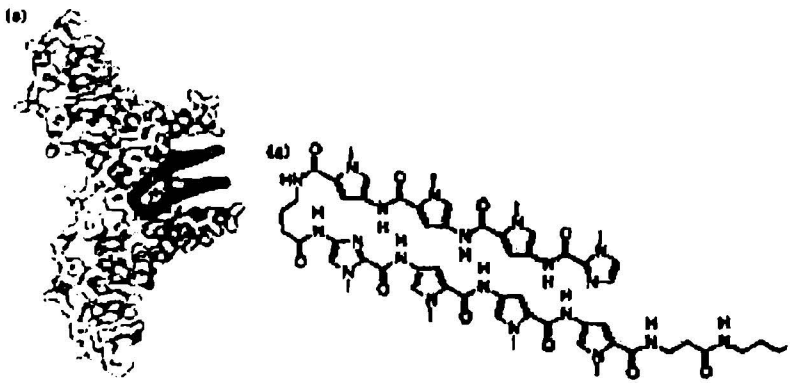


Figure 62: Structure of hairpin ligand (right) on DNA minor groove (left) from JB Chairs, Current Opinion in Structural Biology, 1998, 8:314-320

The compound was found to recognize GC base pairs. Solid phase synthesis of polyamides of variable length has produced efficient ligands, e.g. the eight ring hairpin polyamide ImPyPyPy-g-ImPyPyPy-b-Dp (Dp dimethylamino propylamide) shown in the figure above. This small synthetic molecule has an binding constant in the order of 0.03nM.

The optimal goal of polyamide ligand design has been reached with finding structures able to recognize DNA sequences of specific genes. The structure shown above inhibits the expression

of 5S RNA in fibroblast cells (skin cancer cells) by interfering with the transcription factor IIIA-binding site.

A new strategy of rational drug design exploits the combination of polyamides with bis-intercalating structures. WP631 is a dimeric analog of the clinically proven anthracycline antibiotic daunorobuicin.

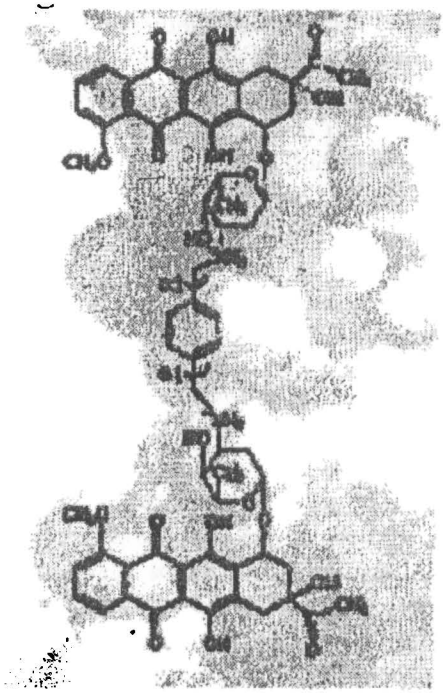


Figure 63: Structure of WP631 from JB Chairs, Current Opinion in Structural Biology, 1998, 8:314-320

This new synthetic compound shows an affinity of 10pM and also showed to be resistant against multidrug resistance mechanisms often encountered in antitumor therapy.

Multidrug resistance is a phenomenon where small aromatic compounds are efficiently expelled from the cell by cell membrane transport proteins commonly referred to as ABC transporters (or ATP Binding Cassette proteins).

Drugs interfere DNA

Drugs that interfere with DNA function by chemically modifying specific nucleotides are Mitomycin C, Cisplatin, and Anthramycin.

Mitomycin C is a well characterized antitumor antibiotic which forms a covalent interaction with DNA after reductive activation. The activated antibiotic forms a cross-linking structure between guanine bases on adjacent strands of DNA thereby inhibiting single strand formation (this is essential for mRNA transcription and DNA replication).

Anthramycin is an antitumor antibiotic which bind covalently to N-2 of guanine located in the minor groove of DNA. Anthramycin has a preference of purine-G-purine sequences (purines are adenine and guanine) with bonding to the middle G.

Cisplatin is a transition metal complex *cis-diamine-dichloro-platinum* and clinically used as anticancer drug.

The effect of the drug is due to the ability to platinate the N-7 of guanine on the major groove site of DNA double helix. This chemical modification of platinum atom crosslinks two adjacent guanines on the same DNA strand interfering with the mobility of DNA polymerases.

DRUG-DRUG INTERACTIONS

It is important to ask the patient about any medication they are currently taking, before prescribing any medication. This is necessary because it is not uncommon to have clinically significant drug-drug interactions which involve drugs used in the management of health. For example, verapamil which is indicated in the treatment of stable or unstable angina, hypertension, and certain cardiac arrhythmias can interact with topical or systemic beta-blockers. The combination of systemically absorbed topical timolol with a calcium antagonists (calcium channel blocker) such as verapamil would normally be additive and provide increased hypotensive effects. In some patients however, the result will be a decrease in the contractility of the heart muscle or exaggerated depression of nerve conduction within the heart. The patient may

experience bradycardia, syncope, orthostatic hypotension, or loss of left ventricle function, the mechanism for this interaction seems to be two-fold. Both an inhibition of the liver metabolism of the beta-blocker (with increased drug blood levels) and an inhibition of the sympathetic nervous system to moderate the effects of the calcium channel blocker occur. Other significant drug-drug interactions that involve drugs used in the treatment have been listed in Table 9.

Table 9: A partial list of clinically significant drug-drug interactions.

Precipitant Drug (1)	Object Drug (2)	Description	Risk
Tetracycline	Digoxin	Increased Digoxin blood levels	Cardiac arrhythmias
Erythromycin	Terfenadine / Astemizole	Decreased antihistamine metabolism	Cardiotoxicity
Propoxyphene	Warfarin	Decreased warfarin metabolism	Risk of bleeding
Aspirin	Sodium Valpronate	Displaced from plasma protein	Tremor, ataxia, nystagmus

(1) drug causing the reaction
(2) Drug affected

Also ask about any untoward reactions due to previously prescribed medications. Previous allergic reactions or adverse drug reactions should also be documented. If the patient does have a positive history for drug allergies, ask specifically "what happened" during any reported reactions.

What is the best way to avoid drug-drug interaction First, know the characteristic of any drag you are prescribing. Second, know the drugs that the patient is currently taking. Third, expect the unexpected. It is often very difficult to predict which patient will have a problem with a particular drug combination.

Chapter 4

SENSORY RECEPTERS

Organisms are subjected to many influences from their surroundings constituting the environment. All changes in the environment, both external and internal are known as *stimuli* Organs of the body that detect these changes or stimuli are called *receptors* or *sense* organs.

Structurally, sensory organs are somatic, when they are in skin, muscle, tendon or bones, and are visceral when they are attached to sensory nerve fibres in visceral organs. Most of the sensory neurons are dendrites, and most cell bodies also have an axon that links to another part of the nervous system. Ends of neurons or entire neurons are exposed, encapsulated with connective tissue, or part of an elaborate sense organ. Receptors associated with connective, epithelial or muscular tissue can be moderated or enhanced by that surrounding tissue.

Functionally, sensory receptors detect stimuli (temperature, pain, pressure, body position, equilibrium, hearing, vision, smell and taste), magnify or amplify a stimulus, in both the internal and the external environment and transduce the stimuli into a graded potential in the dendrites and cell bodies of neurons. Free sensory receptors repost about painful sensations, touch, and sometimes hot and cold. Encapsulated receptors inform about pressure, hot and cold, and also touch. Associated nerve endings like a hair follicle receive all mechanical sensations. As a stimulus overcomes a threshold level, the sensation travels towards the central nervous system. In turn, the central nervous system integrates the incoming

information and sends out messages via efferent motor nerve fibres to effector organs in the brain which responds in appropriate manner. If the signal was small, it may end at the spinal cord, producing a reflex action, like a jerk away from a hot object.

Animals respond to the messages they receive from the world around them. Their reactions to the outside world depend on how the data collected from their surroundings are correctly coded into signals that can be received and processed by neurons in the brain (Figure 64). The sensory organs provide the only means of communication from the environment to the nervous system. Sensations arise when signals are detected by sensory receptor cells and transmitted through the nervous system to the designated part of the brain. Various organs and cells are designated to receive specific stimuli. The major categories of sensory reception are mechanoreception, Chemoreception, Photoreception and phonoreception.

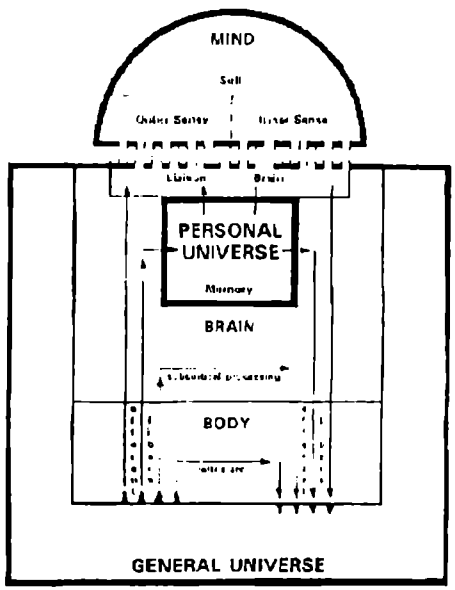


Figure 64: Diagram showing information flow at which afferent (sensory) data is processed and efferent (motor) responses are initiated. The mind can only become aware of information via the liaison brain.

COMMON SENSE

Atleast the following five senses are most common although more senses are recognised by the biologists:

1. *Touch* - includes contact, pressure, heat, cold etc.
2. *Taste* - for certain substances in solution.
3. *Smell* - for volatile chemicals and gases in air.
4. *Hearing* - for vibrations in air, water or solid.
5. *Sight* - for light waves.

The senses can be classified as either general or special:

1. General senses

Include pain, temperature, touch, pressure, vibration, position. These have relatively simple receptors and those receptors are located all over the body.

2. Special senses

Sight, smell, taste, hearing and equilibrium. These have relatively complex receptors and those receptors are located in the head, central nervous system - hypothalamus and medulla oblongata. The hypothalamus contains receptors for detecting temperature, plasma glucose concentrations, and the concentration of blood. The medulla oblongata detects pH of the cerebrospinal fluids

LEVELS OF SENSATION

Sensation is a conscious or unconscious awareness of external and internal stimuli. *Perception* is the interpretation of conscious sensations. The nature of sensation (and reactions to it) vary with the level of the central nervous system in which the sensation is translated. Sensory (afferent) pathways that end in the spinal cord can generate reflex action; if the pathway goes to the brainstem complex subconscious motor reactions can be the result (like medulla changing heart rate or breathing rate); if the path leads to the thalamus, crude identification of the type of sensation occur (pain, touch, hearing); but not until the path reaches the cerebral cortex do we get precise identification and localization (perception) of the sensory experience. Pair really occurs in the brain, as does

sight, smell etc., but we project the stimulus back to the receptor *(so* we think we feel in our hands, hear in our ears, but that perception occurs in the brain, not at the source).

تایبەتمەندیان هاریکاریی لەهەستکردنا

SPECIFICITY OR MODALITY OF SENSATION

This is the property by which one sensation is distinguished from another, in general, a given sensory neuron carries only one modality. For example, a neuron conveying touch to the somato-sensory cortex is not capable of also carrying information about temperature. This is because receptors are specialized having a low threshold to specific stimullus, and a high threshold to all other stimuli. Some free nerve endings in the skin are not specific, but may be stimulated by a variety of stimuli (such as chemical, mechanical, temperature, and pressure). The receptive field is the area (for example of the skin) that is monitored by a single receptor cell.

بەشەکانی هەستکردن

COMPONENTS OF SENSATION

Four events must occur for a sensation to occur:

1. ***Stimulation.*** A stimulus (change in the environment) capable of initiating a nerve impulse must be present.
2. ***Tranduction***: Sensory receptor (sense organ) must pick up the stimulus and transduce it to a nerve impulse (convert it from one energy form, like light, to another energy form, the electrical nerve impulse), by way of what's called a generator potential.
3. ***Conductions*** The impulse must be conducted along a neural pathway from the receptor to the brain.
4. ***Translation:*** A region of the brain or spinal cord most translate the impulse into a sensation.

پۆلێنکردنی ئەندامەکانی هەستن

CLASSIFICATION OF SENSE ORGANS

Sense organs can be classified in many ways, by complexity, by nature, by the type of stimuli, by the location (Table).

پۆلێنکردن بەگوێرەی ئالۆزیی

CLASSIFICATION BY COMPLEXITY

Simple receptors are associated with the general senses (like pain, temperature, touch). They may be free nerve endings or a

dendrite wrapped in some connective or epithelial tissue. Complex receptors are associated with the special senses (like vision, hearing, equilibrium) and have complex receptors in specific localized organs of the body (eye, ear).

CLASSIFICATION BY NATURE

General receptors are cutaneous sense organs that are distributed widely upon or within the body especially the skin. Their exact functions are not clear and any one of them can not be related to a single sensation alone. On the other hand special receptors include tongue, nose, eye and ears. They are concentrated in small areas particularly on the cephalia end of the body. They respond to particular type of stimuli or special senses and their functions are better understood.

CLASSIFICATION BY STIMULUS

The *mechanoreceptors* are sensitive to stimuli that distort cell membranes. Those membranes contain mechanically-gated ion channels that open or close in response to stretching, twisting, compression etc. *Thermoreceptors* are free nerve endings. Separate thermoreceptors respond to hot and cold stimuli, but there are no known structural differences between the two types. They are mainly in the skin, but also in skeletal muscle, the liver, and hypothalamus. Cold outnumber warm receptors by about three or four times. *Nociceptors* are pain receptors, especially common in superficial parts of skin, in joint capsules, around walls of blood vessels. They consist of free nerve endings located in nearby every body tissue. *Chemoreceptors* detect smell of a chemical substance, or odours in and (nose), and taste (tongue); *photoreceptors* detect photons of light waves or sight (eye).

CLASSIFICATION BY LOCATION

Exteroreceptors are at or near body's surface and detect change in the external environment (like touch, pressure, taste, heat etc.). These include eyes, ears, nose, taste buds and cutaneous sense organs. They inform the organism about food, mate or enemy. *Interoreceptors* lie in various internal organs such as vessels or viscera and provide the information about the internal environment,

Table 10: Types of Receptors or Sense Organs According to Stimuli and Location.

Sense Organs or receptors	*According to type of stimulus*	*According to location of stimulus*	*Stimuli*	*Functions*
1. Skin (cutaneous)	*Mechanoreceplors* *Thermoreceptors*	*Exteroceplors*	Contact Temperature	Detecting touch, hot and cold, etc.
2. Muscles (kin esthetic)	*Mechanoreceplors*	*Proprioceptor*	Mechanical stretch	Feeling and gauging pressures.
3. Tongue (gusutoiy)	*Chemoreceplor*	*Exteroceplors*	Dissolved chemicals -	Tasting
4. Nose (Olfactory)	*Chemoreceptor*	*Exteroceptor*	Volatile chemicals gases in air.	melling
5. Eyes (visual)	*Photoreceptors*	*Exteroceptors*	Light	Seeing
6. Ears (auditory)	*Slatoacoustic*	*Exteroceplors*	Sound and gravity	Hearing and balancing
7. –	–	*Interoceplors*	Pain, fullness, CO_2 level, blood composition ,etc.	Maintaining internal body . environment

such as carbon dioxide concentration, blood composition, pain etc. They are responsible for maintaining an appropriate internal body environment necessary for the continued survival of the organism. *Propioreceptors* are stretch receptors present in muscle spindles, tendon organs, connective and skeletal tissues. They supply information about the so-called "kinesthetic" (pressure and acceleration/deceleration of joints) sense of equilibrium and orientation. They act like pressure gauge and are responsible for maintenance of body posture.

WHAT ARE THE GENERAL PROPERTIES OF SENSORY RECEPTORS AND HOW ARE THESE MESSAGES TRANSMITTED TO THE CENTRAL NERVOUS SYSTEM ?

Animals require a constant detection of information from their surroundings. Such information, is the animals link to the outside world. Sensory input is initially detected by sensory receptors, Some receptors are very complex, with many individual receptors along with other structures being organized into sensory organs such as vertebrate eye.

Sensory receptors are transducers; they convert stimuli into electric signals. In most cases, they do not directly generate *action potential.* Instead, sensory receptors generate receptor potentials, which vary in intensity with the intensity of the stimulus. These changes in membrane potential are passed to adjacent sensory neurons, which may generate an action potential if the incoming stimuli are sufficient for the neurons to reach threshold. Increases in receptor potential intensity are translated into a higher frequency of action potentials in the sensory neurons.

Sensory receptors are specialized to respond to only certain stimuli, which will activate the receptor with weak or moderate levels of intensity. This signal is then chemically amplified within the receptor cells. In order for the signal to be effective the intracellular chemical signal must cause membrane channels to open, This produces an electric signal that will be transmitted to the central nervous system.

HOW CAN SENSORY SYSTEM DETECT SUCH A BROAD RANGE OF STIMULI INTENSITY ?

The sensitivity range of a sensory organ is much broader than the range of a single receptor cell. This is because individual afferent fibres of the sensory system cover different parts of the sensitivity spectrum. For example, only the most sensitive receptor cells will respond to a low level stimulus. As the stimulus intensity continues to increase, the receptors become fully activated (saturated), but a group of less sensitive receptor becomes stimulated. This recruitment of additional receptors continues as the stimulus intensity increases, until all receptors are fully saturated. This subdivision of the total range of response by receptor cells of different sensitivities is called *range fractionation* because individual receptors cover only a fraction of the total range of the sensory system. When all receptors are fully active, the system is not capable of detecting any further increase in stimulus intensity,

Some sensory systems (receptors and their neurons) generate a rather constant "background" rate of action potential. If stimulated further, the rate of action potential generation increases due to increased levels of depolarization of the neurons. Therefore, the system does not rely on a minimal level of sensory input in order to respond, which makes the system more sensitive.

WHAT IS MECHANORECEPTION?

Mechanoreception is sensing physical contact on the surface of the skin or movement of the surrounding environment (such as sound waves in air or water). The simplest mechanoreceptors are nerve endings of skin's connective tissue. The most complex example of mechanoreception occurs in the middle and inner ears of vertebrates. The hair cell is the basic unit of vertebrate mechanoreception.

WHAT ARE THE STRUCTURAL MECHANISMS OF THE VERTEBRATE EAR ?

Sound waves enter the external ear of a vertebrate aided by the pinna and the tragus, the entire external structure has a function similar to that of a funnel, amplifying and then concentrating sound waves. Vibrations from sound waves cause changes in air pressure, which travel from the external ear, down the auditory canal and then move the eardrum (tympanum). This

energy is then conducted through the malleus, incus, and stapes, the three small bones that constitute the rest of the middle ear. These three bones are key in the conversion from airborne vibrations to fluid movements. Beneath the stapes is membrane called the oval window, which opens into the cochlea of the spiral shaped, fluid filled inner ear. This entire process serves to amplify sound stimuli upto 22 times before it reaches the cochlea.

HOW DOES THE EAR CHANGE VIBRATION WAVES TO MECHANICAL SOUND?

The ear converts energy to sound into nerve impulses. This process begins at the tympanic membrane. The vibrations that move the eardrum, and then consequently the three additional bones of the middle ear, are transmitted to the oval windows. These vibrations in turn move the fluid of the cochlea. The cochlea is divided into three longitudinal chambers. The two outer chambers are called the scala tympani, and the scala vestibuli, and they are both filled with a liquid perilymph that contains high sodium concentrations. The scala media is filled with a fluid endolymph that had high concentration of potassium. It also contains the organ of corti.

The sound vibrations that pass by the oval window into the cochlear chambers and vibrate the bacterial and basilar membranes, eventually dissipate through the membrane of the sound window.

The floor of the cochlea contains the basilar membrane, and the scala media, organ of corti, where these vibrations undergo the conversion to neuronal impulses. The organ of corti contains sensory hair cells against an overhanging tectorial membrane, and then pull them away. These hair cells are just across synopses from sensory neurons, and this action provides a stimulus that opens sodium channels in the sensory cell membranes. This provides for an action potential in the environment of high potassium concentrations that the endolymph has. Auditory nerves located in the spiral ganglion carry the action potential to the brain.

The frequency of impulses from action potentials relays information on sound to the brain. The louder a sound is, the greater height or amplitude of the vibrations in the sound wave, the more movement of hair cells, and thus the more action

potentials. Pitch can be distinguished through differences in sound wave frequencies. Different areas of the basilar membrane are sensitive to different pitches due to different levels of flexibility along the membrane. Higher frequencies stimulate the basilar membrane closest to the oval window, lower frequencies stimulate areas further along. These regions then stimulate neurons to send the sound signals to specific areas of the brain, and that leads to the perception of a certain pitch.

WHAT ARE THE INSECT'S SYSTEM OF MECHANORECEPTION ?

Insect mechanoreceptors can be found almost anywhere on the surface of an insect's body. They may act as tactile receptors, detecting movement of objects in the environment, or they may provide sensory input about the position or orientation of the body and its appendages. These receptors are innervated by one or more sensory neurons that fire in response to stretching, bending, compression, vibration or other mechanical disturbance. Some mechanoreceptors produce a *phasic response* when stimulated – that is why, they fire once when activated and again when deactivated. Other receptors generate a *tonic response,* firing repeatedly as long as a stimulus persists.

Most insects have 'ears' in their legs. A common structure consists of tracts called tracheae that leads to a membrane stretched over an internal air chamber. Sound waves directly activates nerve impulses in attached receptor cells. These nerve impulses then travel to the central nervous system. Neural processing centres in the brain or segmental ganglia interpret the combinations of tonic and phasic signals sent from nearby receptors.

Trichoform sensilla are probably the simplest mechanoreceptors (Figure 65) these are tactile hairs (setae) that are innervated by a sensory neurons. Dendrites of the neurons attach near the base of the hair and generate a nerve impulse whenever they detect movement. *Hair beds* (clusters of tactile setae) are often found behind the head, on the legs, or near joints where they respond to movements of the body.

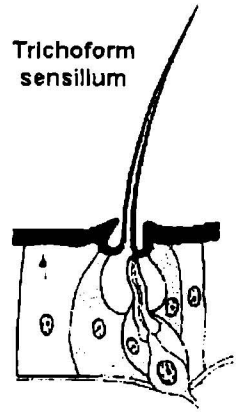


Figure 65: Trichoform sensilla mechanoreceptor of insects.

Campaniform sensilla (Figure 66) are flattened oval discs that usually serve as *flex receptors* in the exoskeleton. They respond whenever mechanical stress causes the exoskeleton to bend. They are found throughout the body especially on the legs, near the base of the wings, and along sutures where two sclerites of the exoskeleton meet.

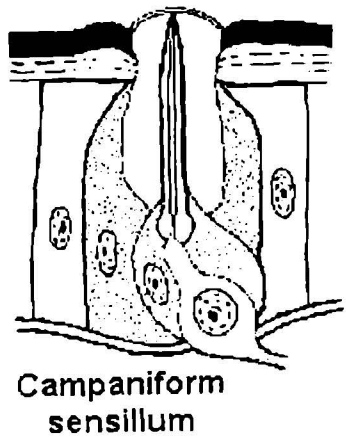


Figure 66. Campaniform sensilla mechanoreceptor of insects.

Some insects such as the noctuid moths have chordotoral ears specially adapted to avoid their predators such as bats. The ear structure consists of a tympanic cavity, a membrane, and three neurons in a *scolopida* formation. The system is stimulated by the ultrasonic vibrations of bat cries. One specific neuron is sensitive to the low intensity vibrations picked up from distant predators, and a different neuron is sensitive to the strong vibrations of a nearby predator. When the neurons are stimulated they send an action potential along the tympanic nerve, and the moth can move according to which neurons have been stimulated (Figure 67).

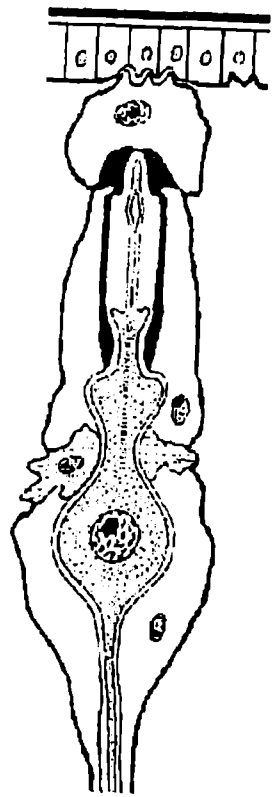


Figure 67: Chordotonal mechanoreceptos of moths.

The Chordotonal organs are of various types. Common types of chordotonal organs include:

(a) Subgenual organs – located in the legs of many insects, these receptors contain relatively few scolopidia yet they appear to be very sensitive to substrate vibrations. Insects may lack specialized sound receptors, yet they can still "hear" vibrations transmitted through the substrate.

(b) Tympanal organs – lie beneath a drum-like membrane (the tympanum) where they respond to sound vibrations. These "ears" may be located on the thorax (in some Hemiptera), on the abdomen (in grasshoppers, cicadas, and some moths), or on the front tibia (in crickets and katydids).

(c) Johnston's organs – found within the pedicel of each antenna. In some insects, they function as a proprioceptors, supplying information on position or orientation of the antennae. In mosquitoes and midges, they respond to certain frequencies of airborne sound by detecting resonant vibrations in antennal hairs. (Shorter hairs near the tip of the antennae resonate to higher frequencies than longer hairs near the base).

MACHANORECEPTION IN FISH

Many fishes and amphibians have a lateral line system enabling than to experience mechanoreception. Pores run up both sides of the fish, through Which moving water enters the lateral line system. This leads to stimulation of neuromasts, the receptor cells of fishes these neuromasts are located throughout the skin, in channels beneath the scales of the main body, and in the dermal bones of the head. These function like sensory hair cells, and vibrations in the water indicating nearby objects or organisms can be detected* Similar to vertebrate, their stimulation leads to an action potential. Eventually these nerve impulses travel to the brain through sensory neurons.

Fish also have ears to extend their hearing to higher frequencies. Sound waves in the water surrounding the fish are conducted as vibrations through the skull. Then, they travel to chambers similar to those located in the cochlea of the vertebrate, and move small granules called *otoliths.* These granules then stimulate sensory hair cells; Most fish tissue has the same approximate density as water, so vibrations in water travel right

through a fish's body. Any structures in the fish that have a significantly different density vibrate differently. The otolith provides this sensory detection in the inner ear of a fish. It's membranes pass the signal on the neighbouring sensory hair cells, and eventually trigger action potentials in the neurons of the auditory nerve.

The gas bladders of some fish also provide an area of variable density. Vibrations pass through the gas bladder,- and travel through a pathway of small bones called *Weberian ossicles.* These serve to connect the gas bladder directly with the inner ear of the fish.

CHEMORECEPTION

Chemoreception is the ability to perceive specific molecules in air or water. These molecules are important clues to the presence of specific objects in the environment. It is essential to many animals in finding food, locating a mate, and avoiding danger.

Chemoreception is divided into two main categories: *gustation* (taste), and *olfaction* (smell). Gustatory receptors respond to dissolved molecules that come in contact with the receptors. Olfactory receptors respond to airborne molecules from a source at a distance.

Gustatory receptors (Figure 68) are commonly described as thick walled hairs, pegs, or pits where the densrites of several (usually upto five) sensory neurons are exposed to the environment througn a single opening (pore) in the cuticle. Each neuron appears to respond to a different range of compounds (for example sugar, salt, water, protein, acid etc.). They are most abundant on the mouth parts, but may also be found on the antennae, tarsi, and genitalia (especially near the tip of the female's ovipositor).

The *peripheral gustatory* or *taste organs* consists of certain modified epithelial cells arranged in flask-shaped *gustatory calyculi* (taste-buds). Which are found on the tongue and adjacent parts. They are also found on the fungiform papillae over the back parts and sides of the tongue. Each taste bud is flask like in shape. It has a broad base and its neck opens by an orifice, the *gustatory* pore. The bud is formed by two kinds of cells: supporting cells, and gustatory cells.

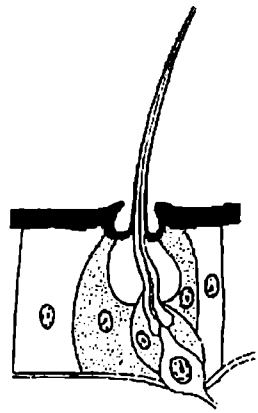


Figure 68: Gustatory receptor of insects.

Olfactory receptors (Figure 69) are usually thin walled pegs, cones or plates with numerous pores through which airborne molecules diffuse. Dendrites of sensory neurons branch profusely within these pores and may respond to very low concentration of detectable compound (for example sex pheromones). Some receptors respond to a wide range of substance while others are highly specific. They are most common on the antennas, but also be associated with the mouth parts or external genitals of insects.

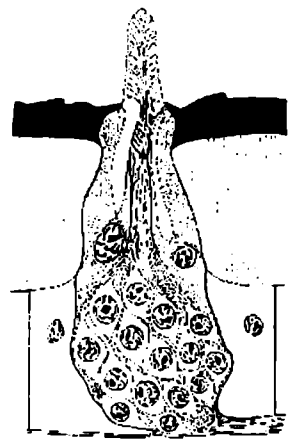


Figure 69: Olfactory receptor of insects.

In vertebrates the olfactory organ or organ of smell consist of two parts: an outer, the *external* nose, which projects from the center of the face, and an internal, the *nasal cavity,* which is divided by a septum into two parts, right and left *nasal chambers.* The external nose is pyramidal in shape, and its upper angle in connected directly with the forehead. Its free angle is termed the apex. Its base is perforated by two elliptical orifices –the *nares,* separated by an antero-posterior septum - the *columella.* The margins of nares have a number of stiff hairs which arrest the passage of foreign substances carried with the current of respiratory air.

DIFFERENCES IN CHEMORECEPTION IN INVERTEBRATES AND VERTEBRATES

Vertebrates detect chemicals using general receptors and two types of specialized receptors -gustatory and olfactory. Many aquatic vertebrates have generalized chemical receptors scattered over their body surface. Vertebrates usually accomplish chemoreception by moving chemically rich air or water into a canal or sac that contains the chemical receptors.

Chemoreception is much different in invertebrates than in vertebrates. For example, planarians find food by following chemical gradients in their surroundings. Their simple chemoreceptors are found in pits on their bodies, over which they move water with cilia. Insects have chemoreceptors in their body surface, mouth parts, antennae, forelegs, and in some cases the ovipositor. Moths smell with thousands of sensory hairs on the antennae. About 70 percent of the adult male receptors are made to respond to one molecule called *"bombkyol"* a sex attractant released by females of the species. The molecules enter the tiny pores of the hair, or sensillum where the olfactory receptors are found.

THE GUSTATORY MECHANISM

The receptors for the gustatory nerves are known as taste buds located on the tongue and the roof of the mouth (Figure 70) . Sweet, sour, bitter, and salty are the four basic taste sensations resulting from stimulation of the taste buds and the stimulation of olfactory receptor. That is why it is harder to taste when one has a cold, These four basic taste may be evolutionarily developed to show some basic food properties. Sweet taste signals foods high

in calories, salty food signals for food that helps maintain water balances, sour taste may help to signal foods that could be dangerous if eaten in excess, and bitter taste sensations signal toxic foods. Umami taste signals foods that are meaty (non-vegetarian).

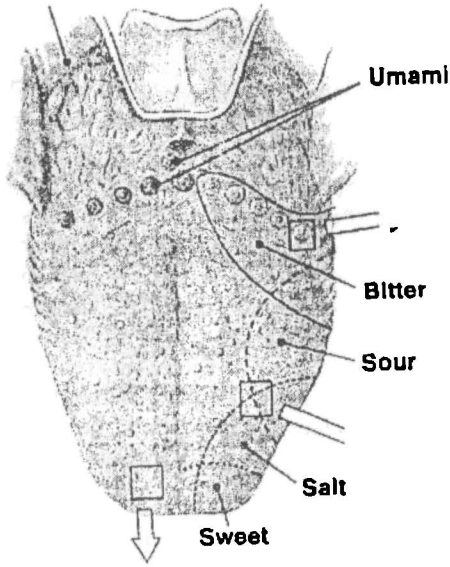


Figure 70: Location of taste buds on the tongue.

THE OLFACTORY MECHANISM

The receptors for the olfactory nerves are located in the upper part of the nasal cavity. The olfactory sense organ consists of hair-like cells at the end of a neuron and is simple compared to the complex visual and auditory organs. They are very sensitive to stimuli, however, they also become very fatigued. This explains why odours seem to go away after being easily noticeable. Canals lined with sheets of receptors with the nasal cavity are called turbinates. Protruding from the end of the nerve are thin cilia that are covered by mucus. Molecules are absorbed into the mucous layer and passed to the cilia where the chemical is detected.

PHOTORECEPTORS

The photoreceptors are either simple or compound:

SIMPLE PHOTORECEPTORS

Simple photoreceptors are of two types: *dorsal ocelli* and *lateral ocelli*. Although both types of ocelli are similar in structure, they are believed to have separate phylogentic and embryological origin.

The dorsal ocelli are commonly found in adults and in the immature stages (nymphs) of many hemimetabolous insects (Figure 71). They are not independent visual organs and never occur in species that tack compound eyes. Whenever present, dorsal ocelli appear as two or three small, convex swellings on the dorsal or facial regions of the head; they differ from compound eyes in having only a single corneal lens covering an array of several dozen rhabdom like sensory rods. These simple eyes do not form an image or perceive objects in the environment, but they are sensitive to a wide range of wavelengths, react to the polarization of light, and respond quickly to changes in light intensity. No exact function has been clearly established, but they act as an "iris mechanism" – adjusting the sensitivity of the compound eyes to different levels of light intensity.

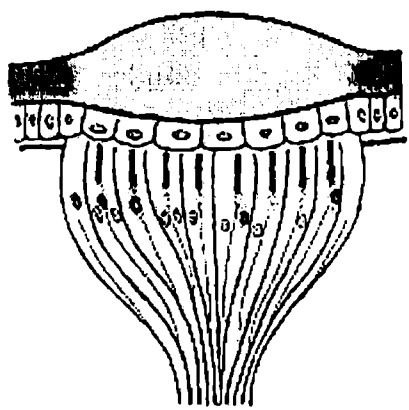


Figure 71: The dorsal ocellus.

The lateral ocelli are the sole visual organ of holometabolous insects. They always occur laterally on the head, and vary in number from one to six on each side. Structurally, they are similar to dorsal ocelli, but often have a crystalline cone under the cornea and fewer sensory rods. Larvae use these simple eyes to sense light intensity, detect outlines of nearby objects, and even track the movements of predators or prey.

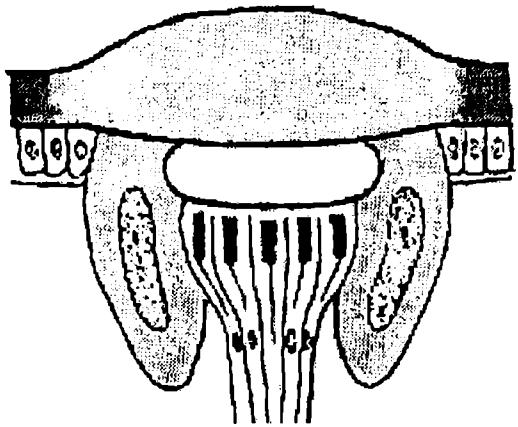


Figure 72: The lateral ocellus.

Ommatidia are the structural and functional units of vision. Externally, each ommatidium is masked by a convex thickening of transparent cuticle, the corneal lens. Beneath the lens, there is often a crystalline cone secreted by a pair of semper cells, Together the lens and the crystalline cone form a diopteric apparatus that refracts incoming light down into a receptor region containing visual pigments. The light sensitive part of an ommatidium is called the rhabdom. It is a rod like structure, The rhabdom contains an array of closely packed micro-tubules where light sensitive pigments (rhodopsin) are stored. These pigments absorb certain wavelengths of incident light and generate nerve impulses through a photochemical process.

COMPOUND PHOTORECEPTORS

A pair of compound eyes are the principle visual organs of most insects. As the name suggests, compound eyes are composed of many similar, closely-packed simple eyes. The number of

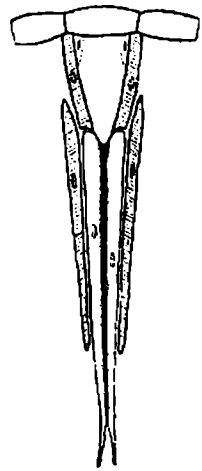


Figure 73: The ommatidium.

ommatidia in a compound eye varies considerably from species to species. Some ants have less than 6, while some dragonflies may have more than 25000 ommatidia. They produce a brighter and *mosaic image* of the environment. Since, insects cannot form a true (focused) image, their visual activity is relatively poor as compared to that of vertebrates. On the other hand, their ability to sense movement, by tracking objects from ommatidium TO ommatidium, is superior to most other animals. Temporal resolution of flicker

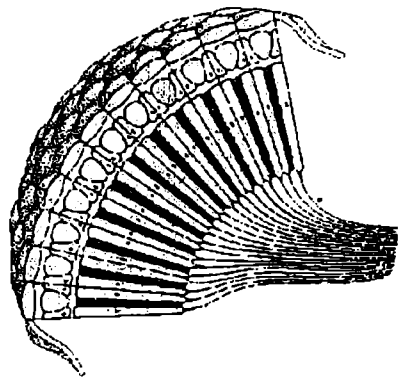


Figure 74: The compound photoreceptor.

is as high as 200 images/second in some bees and flies. They can detect polarization patterns in sunlight, and discriminate wavelengths in a range from ultraviolet to yellow.

STRUCTURAL MECHANISM OF THE VERTEBRATE EYE

In vertebrates such as humans, the surface of the eyeball is made up of the sclera, a white connective tissues and under that a thin pigmented layer called the choroid. The sclera contains the cornea which is transparent, and is where light initially enters the eye, and the choroid contains the iris which contracts and expands to regulate the amount of light entering the whole in its center, known as the pupil. The rear internal surface of the eye is the retina, which contains the actual photoreceptor cells. Between the cornea and the rest of the eyeball is a clear protein lens. The rest of the eyeball consists of a mass of 'jelly like' vitreous humour, which functions as an additional liquid lens through which to focus light images.

OPERATION OF VERTEBRATE EYE

Visual perception in all animals is based on a conserved mechanism. Specific protein molecules make up an optical pathway in which light is directed towards a certain photorecceptive surface in which photoreceptors capture photons. Light initially enters the eye through the cornea. During this process^ light rays are bent, and are then further refracted upon passage through the lens to form an inverted image on the retina. To focus an image, most vertebrates change the curve and thickness of the lens, this action is controlled through ciliary muscles surrounding the lens. They relax, and the lens flattens out when the organism is viewing a distant object and they contract to provide a rounded lens through which to view closer images. Strong ocular muscles direct both left and right eyes so that images received by each eye travel to the same spots on the two retinas, producing binocular convergence.

In the retina, there are two types of photoreceptor cells, *rods* and *cones* (Figure 75); Rods are more numerous than cones in the periphery of the retina. There are about 125 million rod cells and 6 million cone cells in human retina. Rods are for dim light, and cones are for bright light and are responsible for colour vision.

Rods and cones contain visual pigments called *opsins,* which control the pigments absorbed by each receptor cell. In the rods, protein is called *rhodopsin.* In bright light the opsin and retina separate thus making the rodsinactive. Rods are more sensitive to light than cones, that is why they work better in dim light but do not distinguish colour. They enable us to see only in black and white. This sensitivity is due to the connections with increased convergence leads to greater magnification of a weak stimuli. When only rods are stimulated, such as in dim light, we only see in black and white. Additionally, because rods are not part of the fovea, where images are best focussed, we can see images better at night when we do not look directly at them. In human beings there are three kinds of cones blue, green and orange. Each type of cone has specific photopigment molecules, and each molecule experiences maximum absorption at a different wavelength. All other colours are perceived by the stimulation of two or more cone types.

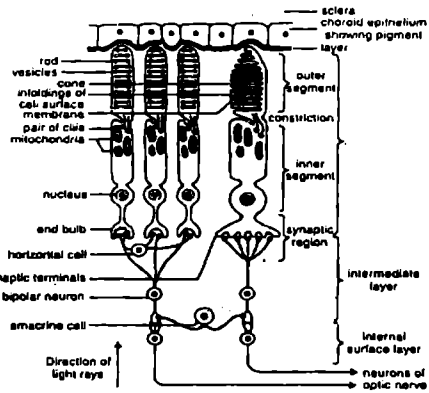


Figure 75: Rod and cone cells of retina.

In vertebrates (and human beings), there is a specialized portion of the retina called the *fovea* (Figure 76). This area provides most clear vision, in fovea; there are no rods. This region contains only cones packed together than in the rest of the retina. Blood vessels and nerve fibres go around the fovea so light has a direct path to the photoreceptors. This area of the eye is most efficiently taken advantage of during day when the photoreceptor cells of the cones that dominate the fovea, are best able to absorb light. Light hitting the receptor cells, rods or cones, produces a

charge gradient across the membrane, but this is not an action potential. These cells however synapse with neurons that then synapse with ganglion cells. These convey the image message as an action potential to the brain along optic nerves. The optic nerves from the right and left eyes meet at an optic chiasma in the brain.

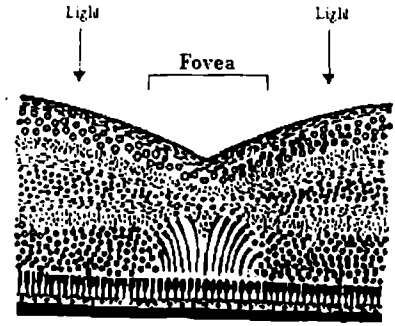


Figure 76: The fovaa region of the retina.

One part of the retina does not contain any photoreceptors. This is our *blind spot*. Any image that falls on this region will not be seen. It is in this region that the optic nerves come together and exit the eye on their way to the brain (Figure 77).

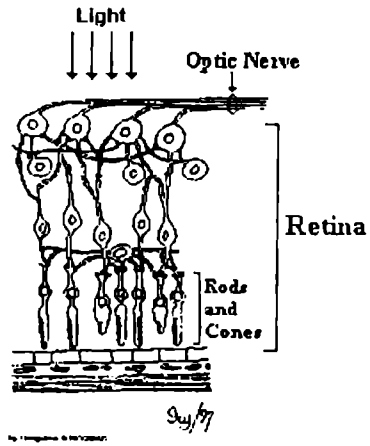


Figure 77: Blind spot of photoreceptor.

To demonstrate the blind spot, draw a small dot '•' about 6-8 microns from a small '+' on the right side, on a piece of paper. Hold the image about 20 inches away. Cover your left eye and look at the below directly with your right eye. You should also be able to see the even though you a looking directly at it. Now slowly move closer to the screen, keeping your right eye focussed on the '•'. The image of the '+' should disappear, then reappear as you continue to move closer. This is because the image of '+' moved across your blind spot as you moved closer....

Here is another image that will help you find your blind spot.

For this image, close your right eye. With your left eye, look at the red circle. Slowly move your head closer to the image. At a certain distance, the broken line will not look broken!!

WHY DO SOME ANIMALS SEE IN BLACK AND WHITE?

Many animals are nocturnal, and have increased amount of rods in their optical systems. The cones that control colour vision, are really unnecessary or are needed in extremely small quantity.

Photoreception systems of insects

Vertebrate and insect eyes have vastly different morphology and structure, although they operate under very similar photochemical systems. The compound eye of most insects has many facets. Behind the corneal lens of each facet, there are functional units called ommatidium.

The receptor cells within the ommatidium each detect a very small fraction of the spectrum of light that the eye as a whole is exposed to, like the rods and cones of the vertebrate eye. In compound eyes, the photoreception cells are called retinular cells, and they surround a single eccentric cell. The receptor cells have a specific position of membrane, designated as rhabdomere, which has a high density of microvilli. Rhodopsin, a photoreceptive

pigment molecule, is contained in this rhabdomere, and this protein absorbs the photons of light energy that enter the eye. This then provides amplification of light signals through a G-protein directed reactions ion channels in the cells open, allowing calcium ions to enter the cell. This is the basis for a current travelling down the receptor cell axon, which crosses gap junctions and reaches the dendrite of the adjacent eccentric cell. This eccentric cell then depolarizes and generates action potentials. These travel through the optic nerve to the central nervous system.

With each ommatidium, different retinular cells are sensitive to different colours due to protein variations within the rhodopsin. Most insects are equipped to see further along the short wavelength end of the colour spectrum, towards ultraviolet. However, they do not see into the rods, which make-up the longer wavelengths that vertebrates can see.

OPTICAL SYSTEM IN FISH

The optical system in fish is very similar to that of land vertebrates, however, there are some differences. The fish has a more spherical shaped lens than the land dwellers. Fish focus by changing the relative distance between the lens and the retina, whereas other vertebrates change the curvature of their more flexible lens. Fish have choroids which contain a special structure - the *tapetum lucidum,* and this contains very reflective guanine crystals to aid dim light vision. This is very important because of the lowered amount of light that penetrates the fish aquatic environment. Additionally, many deep-sea fish have only these and rods, for increased low light sensitivity. They even have epithelial layers for the specific purpose of protection from bright light.

Fish with cones, generally have four types of cones - red, green, blue, and ultraviolet. Some fish have two or three types of cones, fishes with all four cones usually live close to the water surface. Some fish have upwardly directed eyes, especially those that are preyed upon by birds. Some deep sea fish have tubular eyes, which helps to concentrate the limited light that penetrates to great depths. The south American "four-eyed fish" swims along the surface, with its eyes protruding partly out of the water. Each of its two eyes are split into an upper half for vision in air and a lower half for under water vision.

SPECIAL MECHANISMS OF THE CONVERSION OF LIGHT STIMULATION TO NEURONAL IMPULSES

The visual pigment molecules are the specific structures that absorb photons of light. These pigment molecules are made-up of an opsin such as rhodopsin, and an actual light absorbing component, which is usually retinal. When a photon of light hits this molecule, the normal cis-configuration of the retinae is isomerized into a transconfiguration. This in turn leads to a separation between the opsin and the retinal molecule, and eventually to changes in the opsin's conformation too. When the light is absorbed by the retina, proteins that are associated with the cell membrane are activated. This alters the cell's membrane potential, and can eventually lead to an action potential which is carried to the brain via a sensory neuron.

At the cells resting state, there is a certain concentration of each ion in and outside of the cell membrane. This provides for potential diffusions across the membrane. However, there are also charges on each of these ions, and there are polar gradients that do not necessarily correlate to the diffusion gradients. There is usually a stronger negative charge within the cell and a stronger positive charge outside the cell. Positive ions are attracted to the cell membrane because of the negative interior, and vice-versa. The only way that these ions can travel through the selectively permeable membrane, though, is through specific ion channels. Some of the channels are normally open, and stimulation of the cell closes them, and others are closed at resting state, and open in response to stimulation. Such a stimulus could be absorption of a photon of light by retina. When this occurs, there are two possible results - the cell can become hyperpolarized or depolarized, in retina, sodium channels close in the presence of light, and potassium continues to move out of the cell, This makes the cell environment even more positive, and the inside more negative. This is hyperpolarization. A lessening in the charge of difference between the inside and outside of the cell would conversely be depolarization. Each of these conditions leads to charges that travel across synapses to neuronal cells. This alters the firing rate of action potentials in the adjacent neurons.

THERMORECEPTION

The thermoreceptors are free nerve endings. Separate thermoreceptors respond to hot and cold stimuli, but there are no known structural differences between the two types. They are mainly in the dermis, but are also found in skeletal muscle, the liver and hypothalamus. Cold receptors outnumber the warm receptors by about three to four times.

Temperature receptors in some snakes can be extremely sensitive, the infrared detectors in the facial pits of rattle snakes are an excellent example. The sensory axons from the pit organs increase the rate of action potentials when the temperature inside the facial pit increases by 0.002 °C. A rattle snake can detect the body heat of a mouse standing 40 cm away if the mouse's body temperature is atleast 10 °C above the surrounding air temperature. Because the snakes have a pit on each each side of its head, they can tell the direction of the source of heat.

The beetle *Melanophila acuminata* seeks out forest that have just been ravaged by fires so that it can lay its eggs in the nutritious, freshly burnt wood. These insects are capable of detecting fires upto 32 kilometers distant. They do not see the fire with their eyes but instead detect the thermal (infrared) radiation with a special organ on their chest.

Chapter 5

PHOTOREGULATORY SIGNAL REGULATION

Plants use several different photoreceptors to enable them to sense the quality and quantity of light in the surrounding environment. In response to this information, they adjust their growth and development accordingly. In plants (such as *Arabidopsis thaliana)* cryptochrome photoreceptors selectively senses blue and ultraviolet light.

Among microbes such as *Clamydomonas, Cladophora Volvex, Euglena* and *Protozoa*, the photoreceptors are the most primitive eye like. They are called stigma or eye spot. It is simply a mass of granules; circular, oval or sublinear in outline; orange-red in colour. It is situated frequently in the anterior portion of the cell, near the insertion of the flagella. Thus, there are two sides - blind, and the seeing side of the organism, enabling it to orient to light. The eye spot or stigma is a photoreceptive organ concerned with the direction of the movement of flagella (Figure 78). These simple eye spots produce no image but detect directions.

The euglenoid eyespot is part of a fascinating system. The red eyespot is actually haematochrome (red pigment) granule filled concave-convex colourless shield that lies next to the base of the flagellum. Although the exact mechanism is not known, the eyespot appears to work in concert with the flagellum, allowing the euglenid to move in response to light, and to find optimal light conditions for photosynthesis. In other words, the eyespot (and

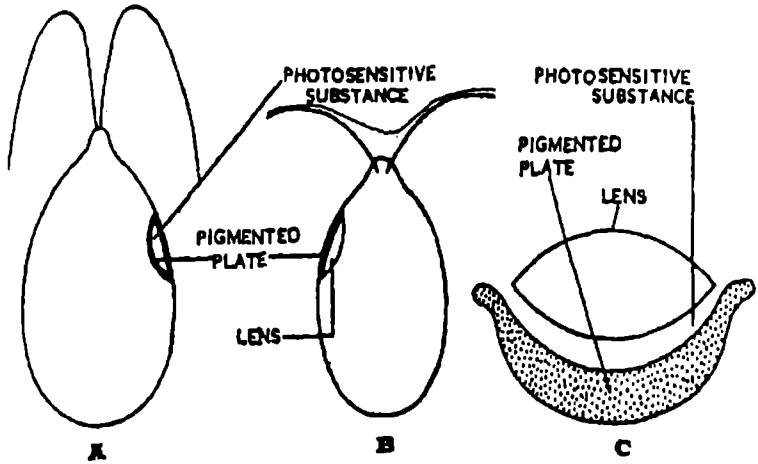


Figure 78: Types of eyespots, A: Chlamydomonas; B: Cladophora; C: Volvox

parts of the flagellum) are a very primitive eye that evolved in ancient single cell organisms - the ancestors of living euglenids. Do these primitive eyes have any relation to our eyes and the eyes of other animals ? It's a fascinating question; one that may be answered in part by studying the genes responsible for euglenid eyespots and for animal eyes also need to know more about the relation of our known single-celled ancestors, Choanoflagellates, with the eyespot-equipped euglenids.

There are great variety of photoreceptors in the animal kingdom from simple cluster of cells that detect only the direction and intensity of light to the complex organs that form images. The image forming eyes are capable of portraying to the brain a true picture of the object in the environment. Despite their many differences, all photoreceptors give a definite idea about the objects in the environment. The sensitivity to light is attained in two general ways. Firstly, the light is concentrated with the help of lenses. Secondly, the pigments present in the photoreceptors absorb light.

In arthropods there are two types of eyes: simple eye, and the compound eye. The simple eye have a single ommatidium, while the compound eye have several ommatidia.

The *Hirudeneria granulosa* (Annelida) there are five pairs of eyes having the appearance of blackspots. They are arranged in a semi-oval line on the dorsal surface of the five segments. The eyes vary in size. Those of the first and second pairs become progressively smaller. The eyes also differ in direction, so as to receive light from all sides. The first pair faces forwards, the second pair faces forwards and outwards, the third pair faces upwards, the fourth pair faces backwards and outwards and the fifth pair faces backwards. Each simple eye is cylindrical in form, embedded in the skin with its long axis at right angles to the body surface. It consists of a number of large, refractile, light sensitive cells arranged in vertical rows (Figure 79). Each cell encloses a large crescentric hyaline optic organelle or lens that reduces its cytoplasm to a thin peripheral layer with nucleus pushed to one side. The refractile cells are invested by a layer of pigment cells, forming a deep cup. The latter is covered by a convex transparent cap or cornea consisting of tall epidermal cells and cuticle. The optic nerva enters the cup at its base and extends through it as a central axis, sending a nerve fibril to each refractile cell. This simple eye, perhaps only enable the leech to distinguish light from darkness. Their ability to form images is doubtful.

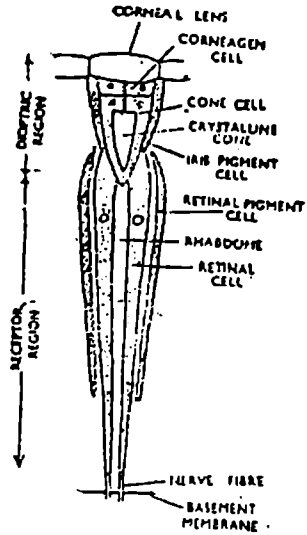


Figure 79: L.S. eye of Hirudineria

Compound eyes of various shapes are found in animal kingdom. They are flat in Aurelia (jelly fish), cup-shaped in Sarsia (gastropods), vesicular in polychaetes, molluscs and some vertebrates, convex eyes in certain molluscs (Figure 80), telescopic eyes in certain fishes. Two different types of image forming eyes are found in animals. They are known as mosaic eyes and camera eyes.

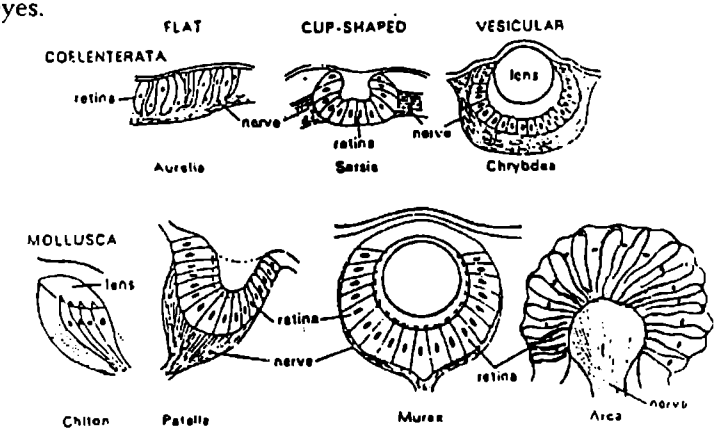


Figure 80: Patterns of eyes in Coelenterata and Mollusca representatives

The compound eye has many ommatidia, arranged parallely along their long axis. Thus, the image seen by the eye is formed on several points lying side by side, one in each ommatidium. Such an image is called the apposition or *mosaic image* of the objects, during bright light. However, daring dim light, the iris pigment cells and the retinal pigment cells retract apart and expose the dooptric parts of the ommatidia. With the result, the rays of light entering several adjacent corneal lenses can reach the same rhabdome. Thus, the image seen by the eye is formed by overlapping points of light. Such an image is called *superposition image.* This image is not as sharp as the mosaic image, but objects give a better impression with superposition image in dim light as no light rays are absorbed.

In *Nereis* (Annelida) the eye consist of two pairs of black spots on the dorsal surface of the prostomium. Each eye (Figure 81) has the form of a cup, whose light sensitive wall is called the retina, and a small circular opening - the pupil. The retina is composed

of a single layer of long narrow, pigmented cells, each having two parts: an inner clear hyaline part known as the rod and an outer pigmented part known as base. The base contains the nucleus, and is drawn out externally into a fine nerve-fibre. The nerve fibre of all cells converge to form a short optic nerve, that joins middle part of the brain. The cavity of the cup is occupied by a gelatinous transparent lens that converges the light rays on the retina. The epidermis and the cuticle of the body wall cover the pupil and function as the cornea.

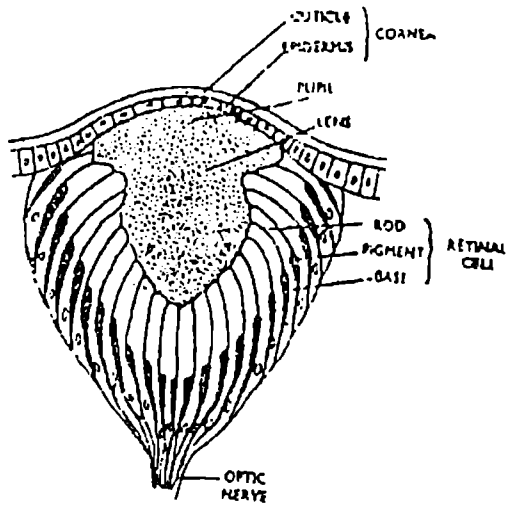


Figure 81: V.S. Eye of Nereis.

In *Pheretima posthuma* (Annelida) the photoreceptors are situated In the deeper part of the epidermis. They are more numerous on the prostomium but occur on other segments also, though in small numbers. A photoreceptor is short but broad cell with a nucleus and clear cytoplasm. It contains a small, curved transparent rod, the optic organelle or lens, which focusses light from almost any angle, on fine neurofibrils that lie in the lower part of the cell (Figure 82). The neurofibrils may be linked to the retina in the functional base. They converge into an afferent fibre which leaves the cell at its base and proceeds to the central nervous system. The photoreceptors enable the earthworm to detect change in the intensity of light. Being nocturnal in habit, the earthworm avoids strong light.

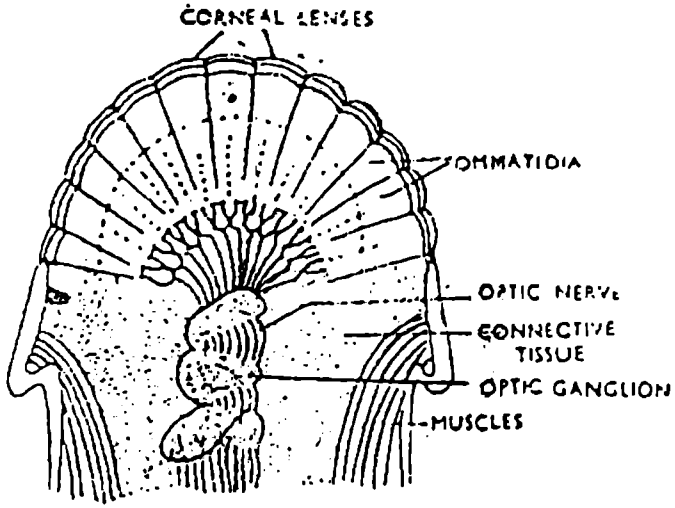


Figure 82: Photoreceptor of Pheretima posthuma.

In *pila globosa* (Mollusca) there are a pair of small black eyes, each situated on a short stalk, the ommatophore, arising outside the true tentacle of its side. Each eye consists of an egg shaped optic vesicle, embedded in the connective tissue a little below the surface of the skin (Figure 83). The wall of the vesicle is composed of a modified connective tissue lined by pigmented retinal cells. The latter are of two types: slender *supporting cells* and the broad *visual cells* bearing a tuft of hair like processes on their outer ends. On the anterior side of the optic vesicle, the retinal cells are unpigmented, thin and transparent, forming the inner cornea or *pellucida interna,* The cavity of the optic vesicle is filled up by a hyaline, gelatinous mass, the lens. The optic nerve enters the optic vesicle from behind to innervate the visual cells. The general epithelium of the ommatophore extends over the eye, where it becomes thin and transparent to form an *outer cornea* or *pellucida externa.* Inspite of their elaborate structure, the eye of pila only detect changes in the light intensity.

The human eye is nearly spnericai is shape and is divided into two chambers by a double convex eye lens. In front of the

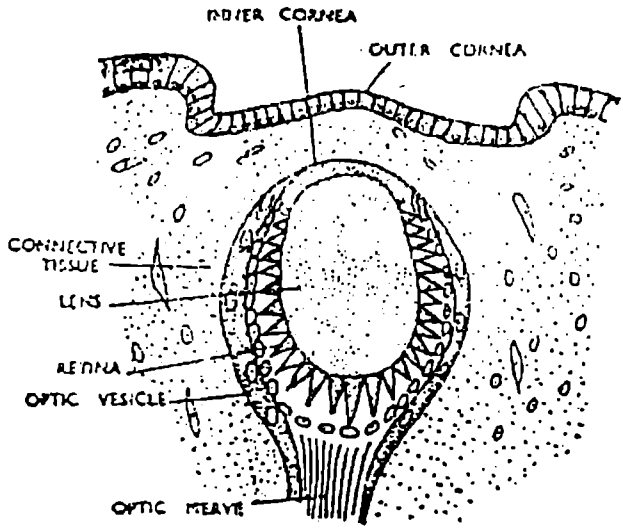


Figure 83: V.S. eye of Pila globosa

lens is a chamber filled with a liquid known as aqueous humour. Chamber behind the lens is filled with a transparent jelly like fluid known as vitreous humour. Iris placed before the lens acts like a stop, which limits the rays entering the eye to a narrow central point. Also it serve to regulate the quantity of light entering the eye, contracting when light is strong and expanding when light is weak. Front of the eye is bounded by a transparent substance known as cornea; Outer surface is opaque and white and serves as protecting cover, while iris is coloured, the colour differing in persons of different countries and nationality e.g. Mongolians are brown eyed, Americans are blue eyed, Indians are dark eyed and so on. Central part of iris has a circular aperture which is dark black in shade and is known as pupil of the eye. Behind it is the eye lens, a double convex lens, inner surface being more convex than the outer one. It is held by ciliary muscles above and below the lens. Inner surface of the eye is covered with transparent nerve fibres and is very sensitive to light. It is called retina and is connected to brain by optic nerves which carry sensation to the brain. Almost in the middle of retina there is a yellowish depressed spot which is exceptionally sensitive to light. It is called the yellow spot.

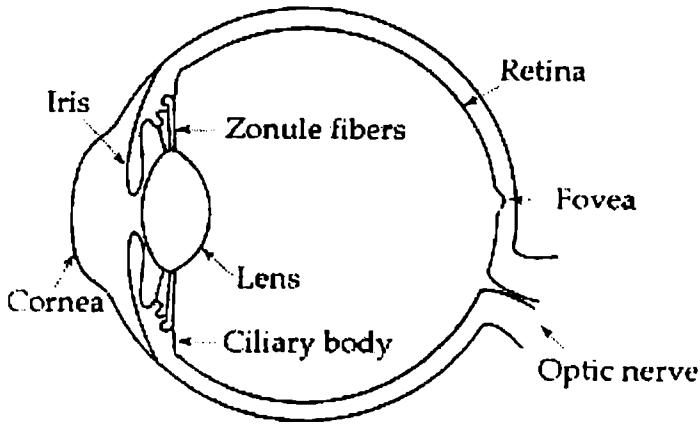


Figure 84: V.S. of the human eye showing its structure.

MOVING PARTS OF THE EYE

The iris is really a shutter that can be closed down to regulate the amount of light entering the eye. This process is controlled by two muscles with distinct innervation:

- the *pupillary sphincter* muscle constricts the pupil like a purse-string and is under the control of the parasympathetic system. Therefore it is innervated by fibres from the oculomotor nerve which originate in the Edinger-Westphal nucleus of the midbrain.
- the *pupillary dilator* muscle is *composed* of radial fibres which pull the pupil open, and is controlled by the sympathetic system. Therefore it is innervated by post-ganglionic sympathetics from the superior cervical ganglion.

The lens is a naturally elastic structure. If it had its way, it would round up into a more spherical shape. Under normal conditions, however, an array of radial fibres - the *zonule fibres* - hold the lens stretched out into a more disc-like shape. This shape allows for far-focusing. What happens when you need to near-focus? A this point the ciliary body, a hoop-like structure

that supports the zonule fibres, comes into plany. Imagine a spiderweb built into the opening of a drawstring purse, suspending a disk in the opening. When tfte purse is open, the spiderweb is taut. It you pull the drawstring, however, the web will go slack and collapse on itself. The ciliary body is the drawstring purse, in this analogy. The *cilliary muscle* within it is the drawstring. When the ciliary muscle contracts, the zonule fibers go slack, the suspended lens is released from their tension, and it is free to round up. This change is necessary for near-focusing. The entire process of adjusting the focus to different distances in called *accomodation.*

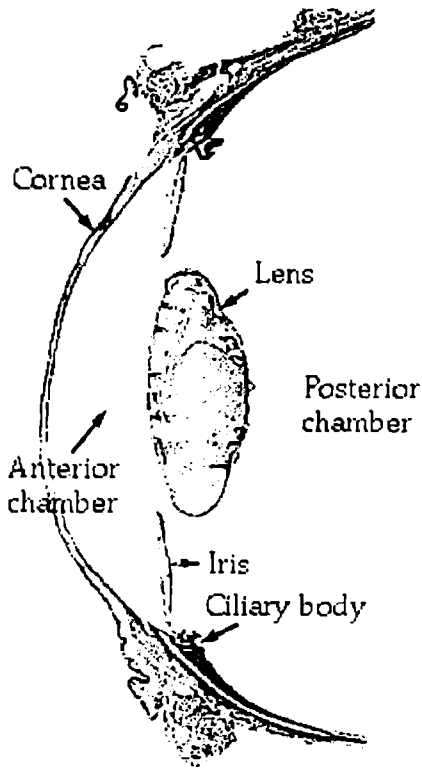


Figure 85: Moving patts of the eye.

THE RETINA

The retina is a seven layered structure (Figure 86) involved in signal transduction. In general, dark "nuclear" or "cell" layers contain cell bodies, while pale "plexiform" layers contain axons and dendrites.

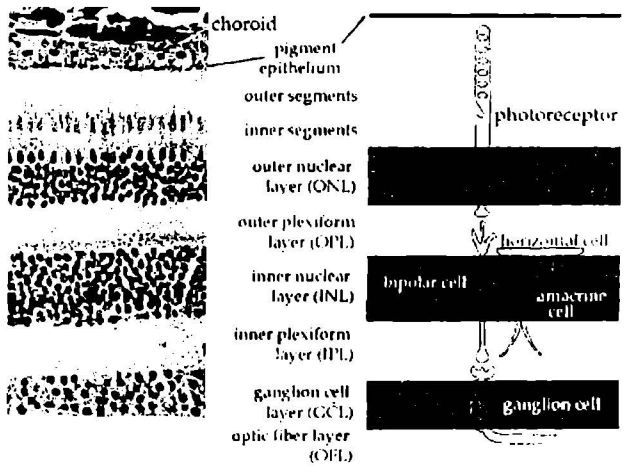


Figure 86: The structure of retina.

Trace the signal through the retina:

- Light enters from the GCL side first, and must penetrate all cell types before reaching the rods and cones.
- The outer segment of the rods and cones transduce the light and send the signal through the cell bodies of the ONL and out to their axons.
- In the OPL photoreceptor axons contact the dendrites of bipolar cells and horizontal cells. Horizontal cells are interneurons which aid in signal processing. The bipolar cells in the INL process input from photo-receptors and horizontal cells, and transmit the signal to their axons.
- In the IPL, bipolar axons contact ganglion cell dendrites and amacrine cells, another class of interneurons.
- The ganglion cells of the GCL send their axons through the OFL to the optic disk to make up the optic nerve. They travel all the way to the lateral geniculate nucleus.

SPECIALIZATIONS OF THE RETINA

The **fovea** defines the centre of the retina, and is the region of highest visual acuity. The fovea is directed towards whatever object you wish to study most closely - this sentence, at the moment. In the fovea there are almost exclusively cones, and they are at their highest density.

The ratio of ganglion cells : photoreceptors is about 2 :1 here, the highest in the eye. In addition, at the fovea all of the other cell types squeeze out of the way to allow the most light to hit the cones. This makes the fovea visible microscopically. The blood vessels also skirt a wide margin around the fovea. The area in and around the fovea has a pale yellow pigmentation that is visible through an ophthalmoscope, and is called the macula.

The ganglion cell axons all leave the eyeball at one location, the *optic disk*. At the optic disk all photoreceptors and accessory cells are pushed aside so the axons can penetrate the choroid and the sclera. This creates a hole in our vision, the *blind spot*. Normally each eye covers for the blind spot of the other, and the brain fills in missing information with whatever pattern surrounds the hole. Therefore we are not conscious of the blind spot.

Photoreceptors are not distributed evenly throughout the retina. Most cones lie in the fovea, whereas peripheral vision is dominated by rods. Overall, rods greatly outnumber cones. Review the characteristics of rods (black and white vision, very sensitive to low light) and cones (colour vision, not so sensitive) and explain these phenomena:

1. To see a faint star, you cannot look directly at it, but must look slightly to the side.
2. A person with macular degeneration can become functionally blind, yet their night vision is not really affected. How would their colour perception be?

INTERESTING ANATOMICAL FACTS

— **The cornea** is continuous with the **sclera,** which in turn is continuous with the **dura.**

— **The choroid,** a highly vascular, highly pigmented layer between the sclera and the retina, is continuous with the ciliary body and the iris. Do not confuse it with the pigment epithelium.

— **The pigment epithelium** is a single cell layer thick, and comes from the outer layer of the original **optic cup** (a classic embryological "pushed-in ball"). In the mature retina it is pushed directly up next to the neural retina, which came from the inner layer of the optic cup. They are not fused together, however, and can separate along the old plane - a "separated" or "detached" retina.

SIGNAL PROCESSING IN THE RETINA

If you were to record from a photoreceptor, you would find that it was "ON" (hyperpolarized, paradoxically) whenever light shone on it. If you recorded from a ganglion cell instead, you would find that diffuse light did little to the cell. However, the cell would respond well to a small spot of light, a small ring of light, or a light-dark edge. We say that this cell has a **centre-surround receptive field** - the centre must be mainly light and the surround mainly dark, or vice versa. What happens between the outer segment and the ganglion cell? This complex receptive field is created by the interneurons of the retina: the bipolar cells and the horizontal cells, primarily.

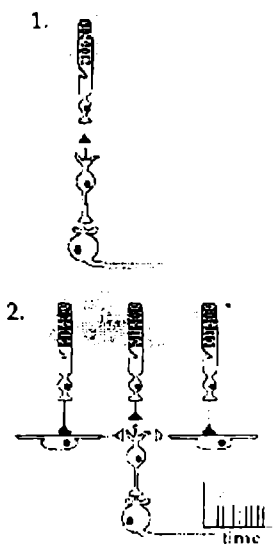
Let's trace a signal through

1. Light hyperpolarizes the cone (or rod). For simplicity's sake, we will just say that turns ON the cone, and thereby excites the bipolar cell directly underneath. That bipolar cell then excites its ganglion cell.

   The same thing is happening to neighbouring cells.

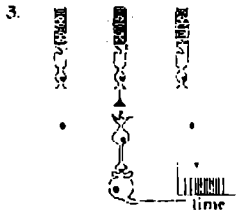
2. However, here's the trick. The neighbour cones also excite horizontal cells. The horizontal cells send processes laterally and inhibit the centre bipolar cell.

   So, what does diffuse light do? It excites the central bipolar cell,



but also inhibits it via the neighbours. Result - the ganglion cell does not get excited. It continues to tick along at its normal, tonic rate.

3. A small spot of light, however, excites the bipolar cell but not its neighbours. There is no inhibition, so it is free to get really excited and excite the ganglion cell, which fires like crazy.

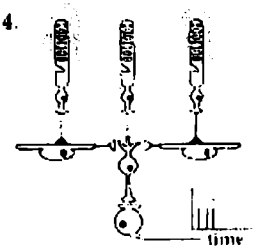


4. A ring of light excites only the neighbours. Now, the bipolar cell is strongly inhibited, with no excitation.

   In response to this strong silencing of the bipolar cell, the ganglion cell shuts down as well. It will not turn on again until the light is turned off, at which time you will see a rebound "off-response".

   This is an ON-centre cell.

   The reverse of this entire scenario can be created by reversing all the signals (which we can do with different receptors to the same neurotransmitter) - you then have an OFF-centre cell.



This unique centre-surround receptive field is also a property of lateral geniculate neurons. Things get even more complicated up in the cortex.

What is the point? Well, our entire visual system exists to see borders and contours. We see the world as a pattern of lines, even things as complex as a face. We judge colours and brightness by comparison, not by any absolute scale. (Don't believe it? Put that

teal scarf next to a blue shirt, you'll call it green. On a green coat, you'll call it blue.) This system of lateral inhibition in the retina is the first step towards sharpening contours and picking up on borders between light and dark. Diffuse light is ignored by the ganglion cell, but a sharp dot will really turn it on. Higher up in the cortex, all these dots will be combined into lines, which will be combined into curves, etc., etc.

THE EYE AS A LIVING CAMERA

When we think of eyes, a natural question comes before us - How do eyes make images ? Our eyes do not see, but we see with our eyes, The visual sense depends entirely on reflected light. Objects in our environment reflect light that enters our eyes, forms an image and transmits information to our brain for processing. This being said is important to recognize the difference between photoreception and vision and thus a difference between photoreceptors and eyes. Photoreception and phototaxis (reaction to light) is observed in single cell animals, microorganisms and lower invertebrates, who neither sense nor are conscious of the light. The term "eye" is generally reserved for organs with a photo-receptory epithelium (the retina) and a lens to focus the light into an image, vision occurs when this image is electrically transmitted to the brain for conscious analysis and response.

Thus, our own advanced human vision is due to a highly complex brain working with a simple lens eye. Other vertebrates who lack our cerebral development depend on a more elaborate eye. Nonetheless, most lens eyes function under the same optical principles.

When light rays hit a medium with a higher optical density, they are slowed and refracted (or bent with a slight change in direction). In the case of a convex surface such as a lens, the angle of refraction depends on the degree of curve. Thus, referring to the figure 87. Each parallel light ray hits the lens at a different angle depending on its distance from the centre, the rays entering the curved lens further from the centre are bent at a sharper angle and those nearer the centre are bent less. At a fixed distance beyond the lens, all the rays onverge on a single point called the focus. This optical principle is used by the eye to focus images on the retina.

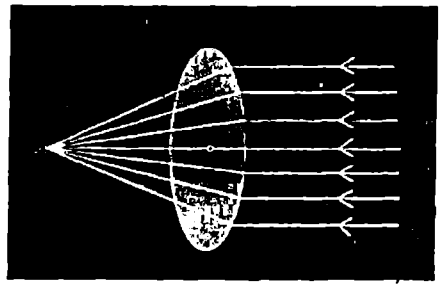


Figure 87: Refraction of light through a lens.

FORMATION OF AN IMAGE

It is useful to think of an image as a collection of points all corresponding to a point on the object. Light reflecting off the object moves in straight lines in all directions away from the object. The points of light from the object that hit the lens of an eye are bent or refracted) and brought into focus as a point of light on the retinal image. Thus, the image is a mosaic of light points that are brought to focus by the refraction or the lens. An image is projected at a fixed focal distance from the lens and appears small and upside-down (Figure 88) . The projection can be captured on a screen, the retina, if the object moves closer or farther from the lens, the image will blur unless the curvature of the lens is changed. Luckily for us, small muscles attached to the eye lens are constantly changing its curvature (a process called accomodation), and our world appears as a sharply focussed image, the points of light of the image then stimulate the retinal cells to induce information transferal to the brain.

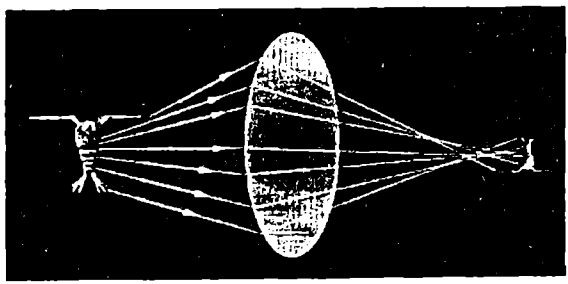


Figure 88: Image projection from a lens.

HUMAN VISION AND COLOUR PERCEPTION

Human stereo colour vision is a very complex process that is not completely understood, despite hundreds of years of intense study. Vision involves the nearly simultaneous interaction of the two eyes and the brain through a network of neurons, receptors, and other specialized cells. The first step in this sensory process are the stimulation of light receptors in the eyes, conversion of the light stimuli or images into signals, and transmission of electrical signals containing the vision information from each eye to the brain through the optic nerves. This information is processed in several stages, ultimately reaching the *visual cortices* of the cerebrum.

The human eye is equipped with a variety of optical components including the cornea, iris, pupil, aqueous and vitreous humours, a variable-focus lens, and the retina, together these elements work to form images of the objects that fall into the field of view for each eye, When an object is observed, it is first focused through the convex cornea and lens elements, forming an inverted image on the surface of the retina, a multi-layered membrane that contains millions of light-sensitive cells. In order to reach the retina, light rays focused by the cornea must successively traverse the aqueous humour (in the anterior chamber), the crystalline lens, the gelatinous vitreous body, and the vascular and neuronal layers of the retina before they reach the photosensitive outer segments of the cone and rod cells, These photosensory cells detect the image and translate it into a series of electrical signals for transmission to the brain.

Despite some misconceptions due to the wide spectrum of terminology employed for describing eye anatomy, it is the cornea, not the lens, which is responsible for the major part of the total refractive power of the eye. Being smooth and clear as glass, yet as flexible and durable as plastic, the anterior, strongly curved, transparent part of the exterior wall of the eyeball allows the image-forming light rays to pass through to the interior. The cornea also protects the eyes by providing a physical barrier that shields the inside of the eye from microorganisms, dust, fibres, chemical, and other harmful materials. Although much thinner in width than the crystalline lens, the cornea provides about 65 percent of the eye's refractive power. Most of the power to bend

light resides near the centre of the cornea which is rounder and thinner than the peripheral portions of the tissue.

As the window that controls the entry of light into the eye, the cornea (Figure 89) is essential to good vision and also acts as an ultraviolet filter. The cornea removes some of the most damaging ultraviolet wavelengths present in sunlight, thereby further protecting the highly susceptible retina and crystalline lens from damage, if the cornea is curved too much, as in the case of near sightedness, distant objects will appear as blurry images, because of imperfect light refraction to the retina. In a condition known as *astigmatism,* imperfections or irregularities in the cornea result in unequal refraction, which creates distortion of images projected onto the cornea.

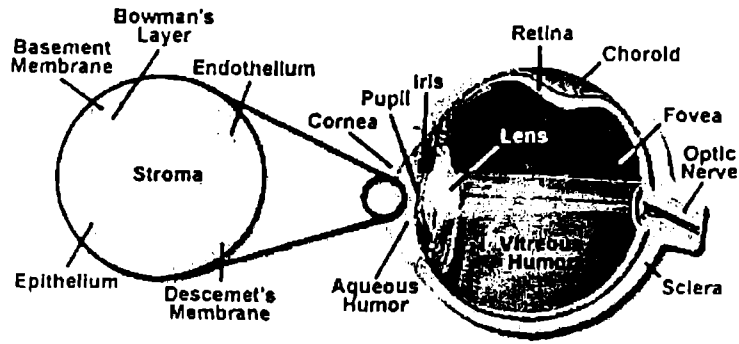


Figure 89: Tissue and structure of the human cornea.

Unlike most tissues of the body, the cornea does not contain blood vessels for nourishment or to protect it against infection. Even the smallest capillaries would interfere with the precise refraction process. The cornea receives its nourishment from tears and the aqueous humour, which fills the chambers behind the structure. The outer epithelial layers of the cornea is packed with thousands of small nerve endings, making the cornea extremely sensitive to pain when rubbed or scratched. Comprising about 10 percent of the tissue's thickness, the epithelial layer of the cornea blocks foreign matter from entering the eye while providing a smooth surface for oxygen and nutrient absorption. The central layer of the cornea, known as the *stroma,* comprises about 90 percent of the tissue, and consists of a water-saturated fibrous

protein network that provides strength, elasticity, and form to support the epithelium. Nourishing cells complete the remainder of the stroma layer. Because the stroma tends to absorb water, the endothelium tissue's primary task is to pump excess water from the stroma. Without this pumping action, the stroma would swell with water, become hazy, and ultimately turn the cornea opaque, rendering the eye blind.

The partial or complete loss of transparency by the crystalline lens, or its capsule, results in a common condition known as *cataracts*. Cataracts are the leading cause of blindness worldwide and represent an important cause of visual impairment. Development of cataracts in adults is related to normal aging, sunlight exposure, smoking, poor nutrition, eye trauma, systemic disease such as diabetes and glaucoma, and undesirable side effects from some pharmaceuticals, including steroids. In the early stages, an individual suffering from cataracts perceives the world as bluny or out of focus. Clear vision is prevented by a reduction in the amount of light that reaches the retina and by clouding of the image (through diffraction and light scattering) as though the individual were observing the environment through a fog or haze (Figure 90). Removal of the opaque lens during cataract surgery, with subsequent replacement by a plastic lens (intraocular lens implants), often resits in corrected vision for unrelated conditions such as nearsightedness or farsightedness.

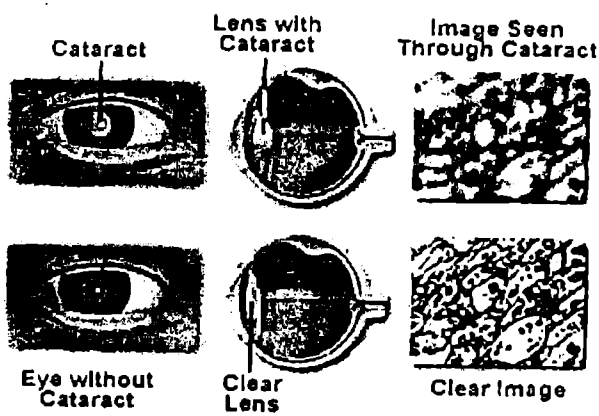


Figure 90: Cataracts in the human visual system.

The Function of the retina is similar to the combination of a digital image sensor [such as a charge-coupled device (CCD)] with an analog-to-digital converter, as featured in modern digital camera systems, the image capturing receptors of the eyes, known as rods and cones, are connected with the fibres of the optic nerve bundle through a series of specialized cells that coordinate the transmission of signals to the brain. The amount of light allowed to enter each eye is controlled by the iris, a circular diaphragm that opens wide at low light levels and closes to protect the pupil and retina at very high levels of illumination.

As illumination changes, the diameter of the pupil (positioned in front of the crystalline lens) reflexively varies between a size of about 2 to 8 millimetres, modulating the amount of light that reaches the retina. When illumination is very bright, the pupil narrows and peripheral portions of the refractile elements are excluded from the optical pathway. The result is that fewer aberrations are encountered by image-forming light rays, and the image on the retina becomes sharper. A very narrow pupil (approximately 2 millimetres) produces diffraction artifacts that spread the image of a point source on the retina.

In the brain, the neural fibres of the optic nerves from each eye cross at the *optic chiasma* where visual information from both retinas travelling in parallel pathways is correlated, somewhat like the function of a time base correction generator in a digital video tape recorder. From there, the visual information travels through the *optic tract* to the knee-shaped *lateral geniculate nuclei* in the *thalamus,* where the signals are distributed via the *optic radiations* to the two *visual cortices* located on the lower rear section of each half of the *cerebrum.* In the lower layers of the cortex, the information from each eye is maintained as columnar ocular *dominance stripes.* As the visual signals are transmitted to the upper layers of the cortex, information from the two eyes is merged and binocular vision is formed, in abnormal ophthalmic conditions such as *phorias* (misalignments) of the eyes, including *strabismus* (crossed eyes), stereovision is disrupted as are the individuals bearings and depth perception, in cases where ophthalmic surgery is not warranted, prismatic lenses mounted in spectacles can correct some of these anomalies. Causes of interruption to the binocular fusion nay be head or birth trauma, neuromuscular disease or congenital defects.

The central fovea is located in an area near the centre of the retina, and positioned directly along the *optical axis* of each eye. Known also as the "yellow spot", the fovea is small (less than 1 square millimetre), but very specialized. These areas contain exclusively high-density, tightly packed cone cells (greater than 200,000 cone per square millimetre in adult humans (Figure 91). The central fovea is the area of sharpest vision, and produces the maximum resolution of space (spatial resolution), contrast, and colour. Each eye is populated with approximately seven million cone cells; which are very thin 3 micrometers in diameter and elongated. The density of the cone cells decreases outside of the fovea as the ratio of rod cells to cone cells gradually increases. At the periphery of the retina, the total number of both types of light receptors decreases substantially, causing a dramatic loss of visual sensitivity at the retinal borders. This is offset by the fact that humans constantly scan objects in the field of view (due to involuntary rapid eye movements resulting in a perceived image that is uniformly sharp, in fact, when the image is prevented from moving relative to the retina (via an optical fixation device), the eye no longer senses an image after a few seconds.

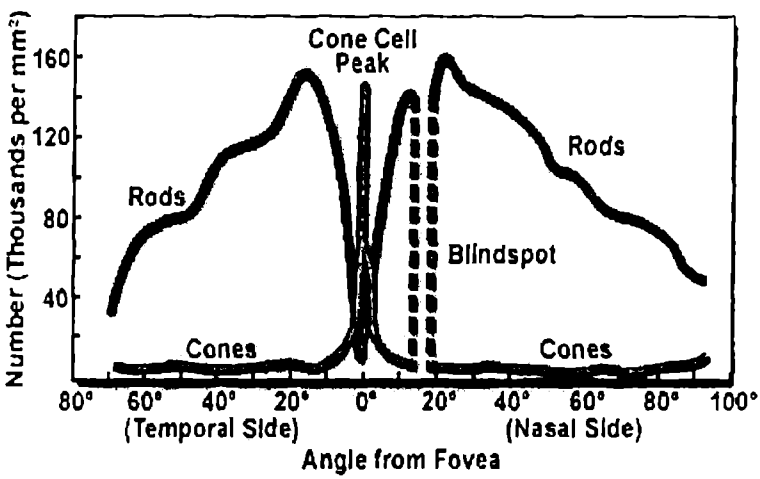


Figure 91: Cone and rod cell distribution in retina.

The arrangement of sensory receptors in the outer segments of the retina partially determine the limit of resolution in different regions of the eye. In order to resolve an image, a row of less-stimulated photoreceptors must be interposed between two rows of photoreceptors that are highly stimulated, otherwise, it is impossible to distinguish whether the stimulation originated from two closely spaced images or from a single image that spans the two receptors rows. With a centre-to-centre spacing ranging between 1.5 and 2 micrometers for the cones in the central fovea, optical stimuli having a separation of approximately 3 to 4 micrometers should produce a resolvable set of intensities on the retina. The radius of the first minimum for a diffraction pattern formed on the retina is about 4.6 micrometers with 550 nanometer light and a pupil diameter of 2 millimetres. Thus, the arrangement of sensory elements in the retina will determine the limiting resolution of the eye. Another factor, termed *visual acuity* (the ability of the eye to detect small objects and resolve their separation), varies with many parameters, including the definition of the term and the method by which acuity is measured. Over the retina, visual acuity is generally highest in the central fovea, which spans a visual field of about 1.4 degrees.

The spatial arrangement of rod and cone cells and their connection to neurons within the retina is presented in Figure 92 . Rod cells, containing only the photopigment *rhodopsin,* have a peak sensitivity to blue-green light (wavelength of about 500 nanometers), although they display a broad range of response throughout the visible spectrum. They are the most common visual receptor cells with each eye containing about 125-130 million rod cells. The light sensitivity of rod cells is about 1,000 times that of cone cells. However, the images generated by rod stimulation alone are relatively unsharp and confined to shades of grey, similar to those found in a black and white soft-focus photographic image. Rod vision in commonly referred to as *scotopic* or *twilight* vision because in low light conditions, shapes and the relative brightness of objects can be distinguished, but not their colours. This rechanism of *dark adaptation* enables the detection of potential prey and predators via shape and motion in a wide spectrum of vertebrates.

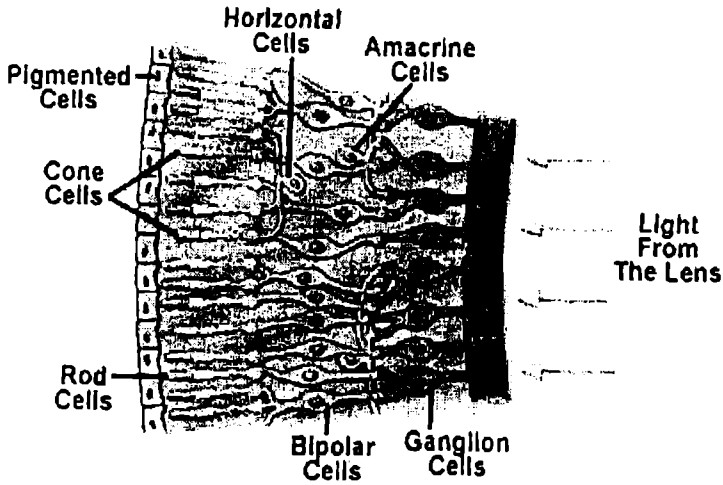


Figure 92: Microscopic anatomy of the retina.

The human visual system response is logarithmic, not linear, resulting in the ability to perceive an incredible brightness range of over 10 decades, in broad daylight, humans can visualize objects in the glaring light from the sun, while at night large objects can be detected by starlight when the moon is dark. At *threshold* sensitivity, the human eye can detect the presence of about 100-150 photons of blue-green light entering the pupil. For the upper seven decades of brightness, *photopic* vision predominates, and it is the retinal cones that are primarily responsible for photoreception. In contrast, the lower four decades of brightness, termed *scotopic* vision, are controlled by the rod cells.

Adaptation of the eye enables vision to function under such extremes of brightness. However, during the interval of time before adaptation occurs, individuals can sense a range of brightness covering only about three decades. Several mechanisms are responsible for the ability of the eye to adapt to a high range of brightness levels. Adaptation can occur in seconds or may take several minutes, depending upon the level of brightness change. Pull cone sensitivity is reached in about 5 minutes, whereas it

requires about 30 minutes to adapt from moderate photopic sensitivity to the full scoptic sensitivity by the rod cells.

When fully light adapted, the human eye features a wavelength response from around 400 to 700 nanometers, with a peak sensitivity at 555 nanometers. The dark adapted eye responds to a lower range of wavelength between 380 and 650 nanometers, with the peak occurring at 507 nanometers. For both photopic and scoptic vision, these wavelengths are not absolute, but vary with the intensity of light. The transmission of light through the eye becomes progressively lower at shorter wavelengths. In the blue-green region (500 nm), only about 50 percent of light entering the eye reaches the image point on the retina. At 400 nm, this value is reduced to a scant 10 percent, even in a young eye. Light scattering and absorption by elements in the crystalline lens contributes to a further loss of sensitivity in the far blue.

Cones consist of three cell types, each "tuned" to a distinct wavelength response maximum centered at either 430, 535 or 590 nm. The basis for the individual maxima is the utilization of three different photo-pigments, each with a characteristic visible light absorption spectrum, the photopigments alter their conformation when a photon is detected, enabling them to react with a *transducin* to initiate a cascade of visible events. Transducin is a protein that resides in the retina and is able to effectively convert light energy into an electrical signal. The population of cone cells is much smaller than rod cells, with each eye containing between 5 and 7 million of these colour receptors. True colour vision is induced by the stimulation of cone cells. The relative intensity and wavelength distribution of light impacting on each of the three cone receptor types determines the colour that is imaged (as a mosaic), in a manner comparable to an additive RGB video monitor or COD colour camera.

A beam of light that contains mostly short-wavelength blue radiation stimulates the cone cells that respond to 430 nm light to a far greater extent than the other two cone types. This beam will activate the blue colour pigment in specific cones, and that light is perceived as blue. Light with a majority of wavelengths centred around 550 nm is seen as green, and a beam containing mostly 600 nm wavelengths or longer is visualized as red. Pure

cone vision is referred to as photopic vision and is dominant at normal light levels, both indoors and out. Most mammals are *dichromats,* usually able to only distinguish between bluish and greenish colour components. In contrast, some primates (humans) exhibit *trichromatic* colour vision, with significant response to red, green and blue light stimuli.

The absorption spectra of the four human visual pigments, display maxima in the expected red, green, and blue regions of the visible light spectra (Figure 93). When all three types of cone cell are stimulated equally, the light is perceived as being *achromatic* or white. For example, noon sunlight appears as white light to humans, because it contains approximately equal amounts of red green and blue light. An excellent demonstration of the colour spectrum from sunlight is the interception of the light by a glass prism, which *refracts* (or bends) different wavelengths to varying degrees, spreading out the light into its component colours. Human colour perception is dependent upon the interaction of all receptor cells with light, and this combination results in nearly trichromic stimulation. There are shifts in colour sensitivity with variations in light levels, so that blue colours look relatively brighter in dim light and red colours look brighter in bright light. This effect can be observed by pointing a flashlight onto a colour print, which will result in the reds suddenly appearing much brighter and more saturated.

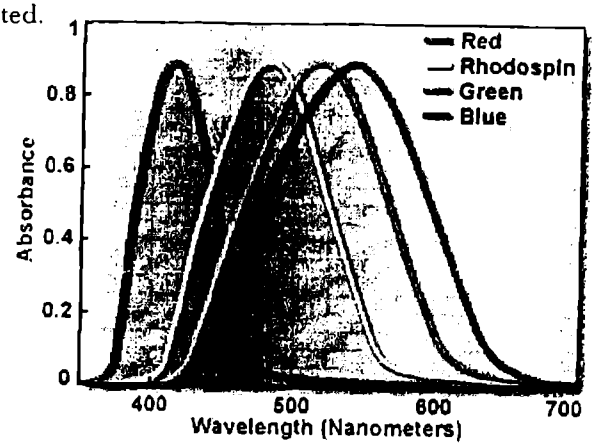


Figure 93: Absorption spectra of human visual pigments.

In recent years, consideration of human colour visual sensitivity has led to changes in the long-standing practice of painting emergency vehicles, such as fire trucks and ambulances, entirely red. Although the colour is intended for the vehicles to be easily seen and responded to, the wavelength distribution is not highly visible at low light levels and appears nearly black at night. The human eye is much more sensitive to yellow-green or similar hues, particularly at night, and now most new emergency vehicles are at least partially painted a vivid yellowish green or white, often retaining some red highlights in the interest of tradition.

When only one or two types of cone cells are stimulated, the range of perceived colours is limited. For example, if a narrow band of green light (540 to 550 nm) is used to stimulate all of the cone cells, only the ones containing green photoreceptors will respond to produce a sensation of seeing the colour green. Human visual perception of primary substractive colours, such as yellow, can arise in one or two ways. If the red and green cone cells are simultaneously stimulated with monochromatic yellow light having a wavelength of 580 nm the cone cell receptors each respond almost equally because their absorption spectral overlap is approximately the same in this region of the visible light spectrum The same colour sensation can be achieved by stimulating the red and green cone cells individually with a mixture of distinct red and green wavelengths selected from regions of the receptor absorption spectra that do not have significant overlap. The result, in both cases, is simultaneous stimulation of red and green cone cells to produce a sensation of yellow colour, even though the end result is achieved by two different mechanisms. The ability to perceive other colours requires the stimulation of one, two, or all three types of cone cells, to various degrees, with the appropriate wavelength palette.

Although the human visual system features three types of cones cells with their respective colour pigments plus light-receptive rod cells for scotopic vision, it is the human brain that compensates for variations of light wavelengths and light sources in the perception of colour. *Metamers* are pairs of different light spectra perceived as the same colour by the human brain, interestingly, colour that are interpreted as the same or similar by a human are sometimes readily distinguishable by other animals, most notably birds.

Intermediary neurons that ferry visual information between the retina and the brain are not simply connected one-to-one with the sensory cells. Each cone and rod cell in the fovea sends signals to atleast three bipolar cells, whereas in the more peripheral regions of the retina, signals from large numbers of rod cells converge to a single ganglion cell. Spatial resolution in the outer portions of the retina is compromised by having a large number of rod cells feeding a single channel, but having many sensory cells participate in capturing weak signals significantly improves the threshold sensitivity of the eye. This feature of the human eye is somewhat analogous to the consequence of *binning* in slow-scan CCD digital camera systems.

The sensory, bipolar cells, and ganglion cells of the retina are also interconnected to other neurons, providing a complex network of inhibitory and excitatory pathways. As a result, the signals from the 5 to 7 million cones and 125 million rods in the human retina are processed and transported to the visual cortex by only about 1 million myelinated optic nerve fibres. The eye muscles are stimulated and controlled by ganglion cells in the *lateral geniculate body,* which acts a feedback control between the retina and the visual cortex.

The complex network of excitatory and inhibitory pathways in the retina are arranged in three layers of neuronal cells that arise from a specific region of the brain during embryonic development. These circuits and feedback loops result in a combination of effects that produce edge sharpening, contrast enhancement, spatial summation, noise averaging, and other forms of signal processing, perhaps including some that have not yet been discovered, in human vision, a significant degree of image processing takes place in the brain, but the retina itself also is involved in a wide range of processing tasks,

In *colour invariance* of human vision, the colour or gray value of an object does not appear to change over a wide range of luminance, in 1672, Sir Isaac Newton demonstrated colour invariance in human visual sensation and provided clues for the classical theory of colour perception and the nervous system. Edwin H. Land, founder of the Polaroid Corporation, proposed the retinex theory of colour vision, based on his observations of colour invariance. As long as colour (or a gray value) is viewed

under adequate lighting, a colour patch does not change its colour even when the luminance of the scene is changed. In this case, a gradient of illumination across the scene does not alter the perceived colour or gray-level tone of a patch, if the illumination level reaches the threshold for scotopic or twilight vision, the sensation of colour vanishes, in Land's algorithm, the lightness values of coloured areas are computed, and the energy at a particular area in the scene is compared with all the other areas in the scene for that waveband. The calculations are performed three times, one for each waveband (long wave, short wave, and middle wave), and the resulting triplet of lightness values determines a position for the area in the three-dimensional colour space defined space defined by the Retinex theory.

The term colour blindness (Figure 94) is something of a misnomer, being widely used in colloquil conversation to ferer to any difficulty in distinguishing between colours. True colour blindness, or the inability to see any colour, is extremely rare, although as many as 8 percent of men and 0.5 percent of women are borne with some form of colour defect (Table 11). Inherited deficiencies in colour vision are usually the result of defects in the photoreceptor cells in the retina, a neuro-membrane that functions as the imaging surface at the rear of the eye. Colour vision defects can also be acquired, as a result of disease, side effects of certain medications, or through normal aging processes, and these deficiencies may affect parts of the eye other than the photoreceptors.

Table 11: Colour Blindness Incidence and Causes

Classification	*Cause of Defect*	*Incidence(%)*
Anomalous Tricromacy		6.0
Protanomaly	Abnormal Red-Sensing Pigment	1.0
Deuteranomaly	Abnormal Green-Sensing Pigment	5.0
Tritanomaly	Abnormal Blue-Sensing Pigment	0.0001
Dichromacy		2.1
Protanopia	Absent Red-Sensing Pigment	1.0
Deuteranopia	Absent Green-Sensing Pigment	1.1
Tritanopia	Absent Blue-Sensing Pigment	0.001
Red Monochromacy	No functioning cones	< 0.0001

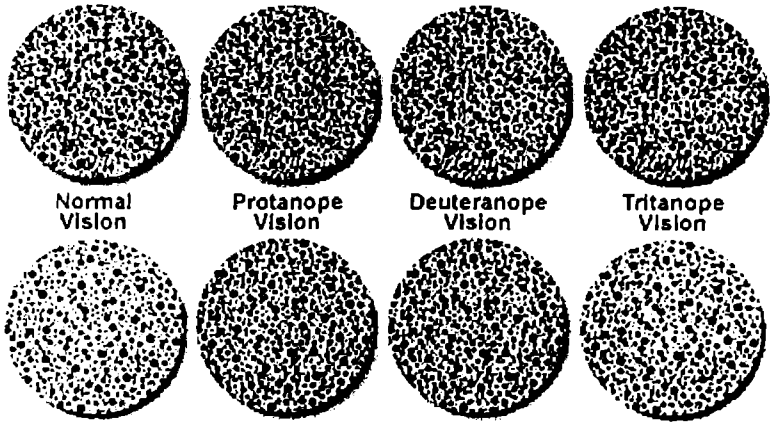


Figure 94. Ishihara colour blindness test.

Normal cones and pigments sensitivity enable an individual to distinguish all the different colours as well as subtle mixtures of hues. This type of normal colour vision is known as *trichromacy* and relies upon mutual interaction from the overlapping sensitivity ranges of ail three types of photoreceptor cone. A mild colour vision deficiency occurs when the pigment in one of the three cone types has a defect, and its peak sensitivity is shifted to another wavelength, producing a visual deficiency termed *anomalous trichromacy*, one of three broad categories of colour vision defects. *Dichromacy*, a more severe form of colour blindness, or colour deficiency, occurs when one of the pigments is seriously deviant in its absorption characteristics, or the particular pigment has not been produced at all. The complete absence of colour sensation, or *monochromacy* extremely rare, but individuals with total colour blindness see only varying degrees of brightness, and the world appears In black, whits, and shades of grey. This condition occurs only in individuals who inherit a gene for the disorder from both parents.

Dichromats can distinguish some colours, and are therefore less affected in their daily lives than monochromats, but they are

usually aware that they have a problem with their colour vision. Dichromacy is subdivided into three types, *protanopia, deuteranopia* and *tritanopia.* Approximately two percent of the male population inherits one of the first two types, with the third occurring much more rarely.

Protanopia is a red-green defect, resulting from loss of red sensitivity, which causes a lack of perceptible difference between red, orange, yellow, and green. In addition, the brightness of red, orange, and yellow colours is dramatically reduced in comparison to normal levels. The reduced intensity effect can result in red traffic lights appearing dark (unlit), and red hues, appearing as black or dark gray. Protanopia often learn to correctly distinguish between red and green, and red from yellow, primarily based on their apparent brightness, rather than on any perceptible hue difference. Green generally appears lighter than red to these individuals. Because red light occurs at one end of the visible spectrim, there is little overlap in sensitivity with the other two cone types, and people with protanopia have a pronounced loss of sensitivity to light at the long wavelength end of the spectrum, individuals with this colour vision defect can discriminate between blues and yellows, but lavender, violet, and purple cannot be distinguished from various shades of blue, due to the attenuation of the red component in these hues.

Individuals with deuteranopia, which is a loss of green sensitivity, have many of the same problems with hue discrimination as do protanopia, but have a fairly normal level of sensitivity across the visible spectrum. Because of the location of green light in the centre of the visible light spectrum, and the overlapping sensitivity curves of the cone receptors, there is some response of the red and blue photoreceptors to green wavelengths. Although deuteranopia is associated with atleast a brightness response to green light, the names red, orange, yellow, and green seem to the deuteranope to be too many terms for colours that appear the same, in a similar fashion, blues, violets, purples, and lavenders are not distinguishable to individuals with this colour vision defect,

Tritanopia is the absence of blue sensitivity, and functionally produces a blue-yellow defect in colour vision, individuals with this deficiency cannot distinguish blues and yellows, but do register

a difference between red and green, The condition is quite rare, and occurs about equally in both sexes. Tritanopes usually do not have as much difficulty in performing everyday tasks as do individuals with either of the red-green variants of dichromacy. Because blue wavelengths occur only at one end of the spectrum, and there is little overlap in sensitivity with the other two cone types, total loss of sensitivity across the spectrum can be quite severe with this condition.

When there is a loss of sensitivity by a cone receptor, but the cones are still functional, resulting colour vision deficiencies are considered anomalous trichromacy, and, they are categorized in a similar manner to the dichromacy types. Confusion often arises because these conditions are named similarly, but appended, with a suffix derived from the term *anomaly*. Thus, *protanomaly* and *deuteranomaly* produce hue recognition problems that are similar to the red-green dichromacy defects, though not as pronounced. Protanomaly is considered a "red weakness" of colour vision, with red (or any colour having a red component) being visualized as lighter than normal, and hues shifted towards green. A deuternomalous individual exhibits "green weakness", and has similar difficulties in discriminating between small variations in hues falling in the red, orange, yellow, and green region of the visible spectrum. This occurs because the hues appear to be shifted toward red. In contrast, deuteranomalous individuals do not have the brightness loss defect that accompanies protanomaly. Many people with these anomalous trichromacy variants have little difficulty performing tasks that require normal colour vision, and some may not even be aware that their colour vision is impaired. *Tritanomaly* of blue weakness, has not been reported as an inherited defect. In the few cases in which the deficiency has been identified, it is thought to have been acquired rather than inherited.

Several eye diseases (such as glaucoma which attacks the blue cones) can result in tritanomaly. Peripheral blue cone loss is most common in these diseases.

Inspite of the limitations, there are some visual acuity advantages to colour blindness, such as the increased ability to discriminate camouflaged objects. Outlines, rather than colours,

are responsible for pattern recognition, and improvements in night vision may occur due to certain colour vision deficiencies, in the military, colour-blind snipers and spotters are highly value for these reasons, During the early 1900s, in an effort to evaluate abnormal human colour vision, the Nagel anomalouscope was developed. Utilizing this instrument, the observer manipulates control knobs to match two coloured fields for colour and brightness. Another evaluation method, the Ishihara pseudoisochromatic plate test for colour blindness, named for Dr.. Shinobu ishihara, discriminates between normal colour vision and red-green colour blindness. A test: subject with normal colour vision can detect the hue difference between the figure and background. To an observer with red-green deficiency, the plates appear isochromatic with no discrimination between the figures and the design pattern.

As a natural part of the aging process, the human eye begins to perceive colours differently in later years, but does not become "colour-blind" in the true sense of the term. Aging results in the yellowing and darkening of the crystalline lens and cornea, degenerative effects that are also accompanied by a shrinking of the pupil size. With yellowing, shorter wavelengths of visible light are absorbed, so blue hues appear darker. As a consequence, elderly individuals often experience difficulty discriminating between colours that differ primarily in their blue content, such as blue and gray or red and purple. At age 60, when compared to the visual efficiency of a 20 year old, only 33 percent of the light incident on the cornea reaches the photoreceptors in the retina, This value drops to around 12.5 percent by the mid-70s.

Accommodation of the eye refers to the act of physiologically adjusting the crystalline lens element to alter the refractive power and bring objects that are closer to the eye into sharp focus. Light rays initially refracted at the surface of the cornea are further converged after passing through the lens. During accommodation, contraction of the ciliary muscles relaxes tension on the lens, resulting in changes of the shape of the transparent and elastic tissue, while also moving it slightly forward. The net effect of the lens alterations is to adjust the focal length of the eye to bring the image exactly into focus onto the photosensitive layer of cells residing in the retina, Accommodation also relaxes the tension

applied to the lens by the zonule fibres, and allows the anterior surface of the lens to increase its curvature. The increased degree of refraction, coupled with slight forward shift in the position of the lens, brings objects that are closer to the eye into focus.

Focus in the eye *is* controlled by a combination of elements including the iris, lens, cornea, md muscle tissue, which can alter the shape of the lens so the eye can focus on both nearby and distant objects. However, in some instances these muscles do not work properly or the eye is slightly altered in shape, and the focal point does not interact with the retina (*convergent vision*), As individuals age, the lens becomes harder and cannot be properly focused, leading to poor vision, If the point of focus falls short of the retina, the condition is referred to as nearsightedness or *myopia,* and individuals with this affliction cannot focus on distant objects. In cases where the focal point is behind the retina, the eye will have trouble focusing on nearby objects, creating a condition known as farsightedness or *hypermetropia*, These malfunctions of the eye can usually be corrected with eyeglasses (Figure 95) using a concave lens to treat myopia and a convex lens to treat hypermetropia.

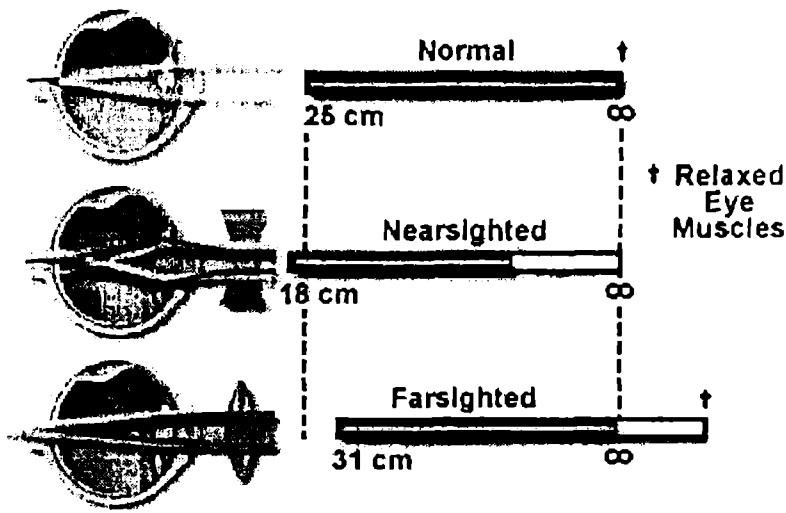


Figure 95: Human eye accomodation range

Convergent vision is not totally physiological and can be influenced by training, if the eyes are not defective. Repetitive procedures can be utilized to develop strong convergent vision. Athletes, such as baseball shorts tops, have well-developed convergent vision. In every movement, the two eyes have to translate in unison to preserve binocular vision, with an accurate and responsive neuromuscular apparatus that is not usually subject to fatigue, controlling their motility and coordination. Changes in ocular convergence or head motion are considered in the calculations made by the complex ocular system to produce the proper neural inputs to the eye muscles. An eye movement of 10 degrees may be completed in about 40 milliseconds, with the calculations occurring faster than the eye can reach its intended target. Small eye movements are known as *saccades* and the larger movements from one point to another are termed *versions*.

The human visual system must not only detect light and colour, but as an optical system, must be able to discern differences among objects, or an object and its background. Known as *physiological contrast or contrast discrimination,* the relationship between the apparent brightness of two objects that are seen either at the same time (*simultaneous contrast*) or sequentially (*successive contrast*) against a background, may or may not be the same. In the human visual system, contrast is reduced in environmental darkness and with individuals suffering from colour deficiencies such as red-green colour blindness. Contrast is dependent on binocular vision, visual acuity, and image processing by the visual cortex the brain. An object with low contrast, which cannot be distinguished from the background unless it is moving, is considered *camouflaged*. However, colour-blind individuals are often able to detect camouflaged objects because of increased rod vision and loss of misleading colour cones. Increasing value for contrast is usually expressed as a percentage or ratio, under optimal conditions, the human eye can barely detect the presence of two percent contrast.

With human vision, an apparent increase in contrast is perceived in a narrow zone on each side of the boundary between two areas of different brightness and/or chromaticity. At the end of the nineteenth century. French Physicist Michel Eugene Chevreul

discovered simultaneous contrast. As a special function of human visual perception, the edges or contour of an object are highlighted, setting the object away from its background and easing spatial orientation. When positioned over a bright background, the region at the edge of a dark object appears lighter than the rest of the background. With this perception phenomenon, the colour with the strongest contrast, the complementary colour, is created (by the brain) at the edge. Because the colour and its complement are perceived simultaneously, the effect is known as *simultaneous contrast.* Borders and other lines of demarcation that separate the contrasting areas tend to lessen the effect (*optical illusion*) by eliminating marginal contrast. Many forms of optical microscopy, most notably phase contrast illumination, take advantage of these features of the human visual system. By increasing the physical contrast of an image without having to change the object via staining or other technique, the phase contrast specimen is protected from damage or death.

HUMAN/MACHINE VISION ANALOGY

Comparing machine vision with human vision is useful for establishing a perspective to understand machine vision. It is also useful, with limits, to help assess what can or should be done with machine vision.

As a model (Figure 96) we can consider the human vision system composed of the eye(s), the optic nerve, ad the brain. The eye forms an optical image with its lens, and senses this optical image with the retina. The optic nerve transmits the image information to part of the brain which analyses and extracts the image information. Another part of the brain uses this information to control the body's muscles.

A model of machine vision is analogous to this simplistic model of the human vision system, A camera with a lens forms the optical image onto an image sensor, a video signal travels through a cable to a computer that analyses the image information to extract the necessary information. This information is then sent to a controller which operates some machinery.

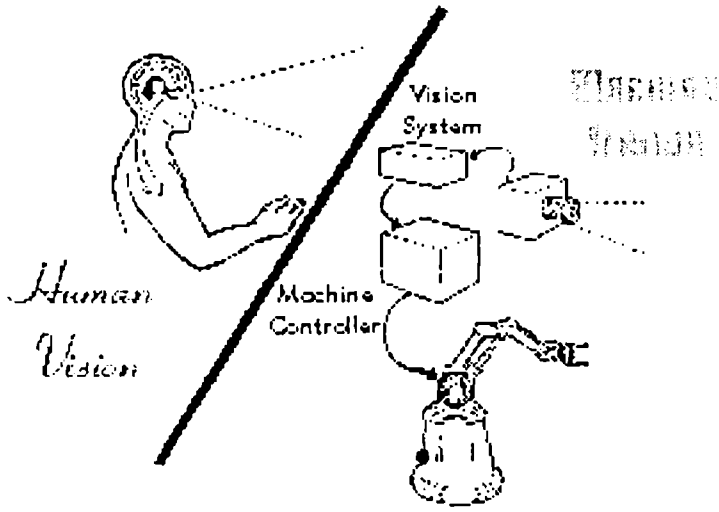


Figure 96. Human/Machine vision analogy.

These are very simplified models, both of machine vision and especially of human vision, but they suffice as a beginning point,

There is much that is not understood about the workings of human vision, and we can only marvel at how incredibly well the human vision system works. With the best technology and the most powerful computers, it is still not practical for any artificial vision system to guide a car through traffic, looking for street signs and avoiding sudden dangers. But even here, research is pushing forward.

Human vision is still an unsolved mystery. Analysing the human vision system we find that its apparent speed would not allow humans to see the ball in most sports, We find that human visual acuity is too coarse to explain how inspectors can grade surface finish on a finely machined part; Yet both activities are very common.

Human vision provides a natural starting point for development of machine vision, Most machine vision equipment has an architecture that is modelled, to some degree, after human vision,

In a person the eye senses an image, information from the image is extracted by one part of the brain, another part of the brain accepts the processed information and commands the muscles to make certain movements. In a machine vision system, a camera or other image replaces the eye, a processor, specially constructed and programmed to analyse image information, processes the camera's output, and a machine controller accepts the output of the image analyser and directs the associated mechanisms in performing the work.

As remarkable as human vision is, the capabilities of the human vision are not an acceptable criteria for the functioning of machine vision. The human eye has on the order of 100,000,000 discrete light sensing elements. Today, common machine vision systems use around 250,000 pixels or picture elements. Many thousands of machine vision systems in the field are operating proof that much can be done with what is available. Technology is advancing, and we now have cameras delivering 1,000,000 to 16,000,000 pixels.

No vision system is presently capable of reliably guiding a car through traffic. Yet the human system does this relatively well. Fortunately most visual tasks in manufacturing are simple when compared to the requirement for driving a car. It is performing these simple visual tasks that machine vision excepts and unshackles the worker. Studies show human inspectors are capable of achieving 80 percent reliability in visual inspection. Machine vision systems can perform with reliability of 99.7 percent and better.

If machine vision were limited to only emulating human capabilities, no one would have ever tried, and succeeded, to make measurements with machine vision; the unaided human eye is an unreliable ruler. Machine vision would be limited to working only with visible light, instead of also exploiting the potential of infrared and x-ray imaging.

BIOPHYSICS OF VISION

The most important photobiological process is photoreception. The eye is a self-regulating system and this regulatory mechanism optimize the operation of the eye. These are essential for (1)

focusing the image on the retina, and (2) regulation of the amount of light which falls on the retina.

The image is the replica of the outside world formed by the optical lens on the retina. The geometric relationship between the outer environment and the retinal image is based on the projection system. The centre of the pupil entrance acts as centre of projection. An object sends out light energy in many directions but only a small proportion which passes through the pupil, contributes to the formation of the retinal image. The optical lens focuses the image on the retina is inverted and reversed way with respect to the object (Fig. 97). The mind perceives the object in the upright condition, despite the upside-down orientation on the retina because the brain is trained to consider an inverted image as the normal.

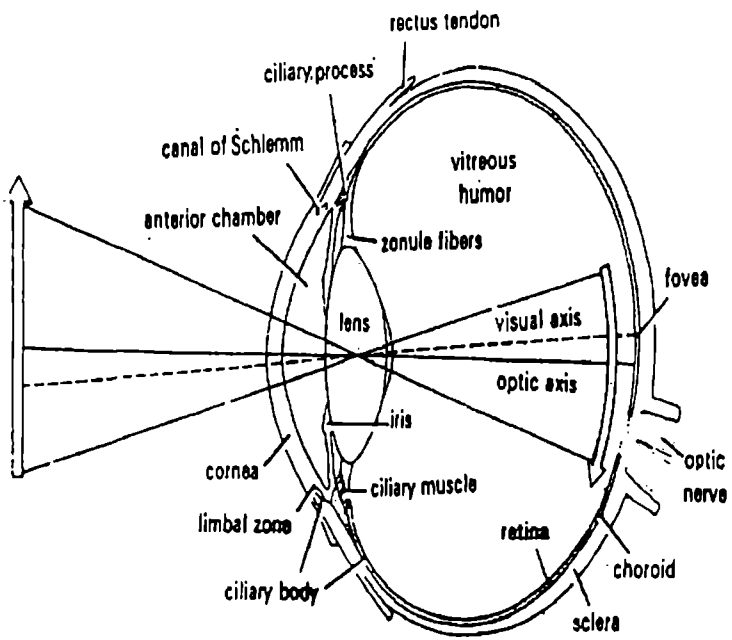


Figure 97: Image formation on the retina of mammalian eye.

Before the actual absorption of light by the photochemical pigments, there are several steps of attenuation of the light incident on the cornea. Therefore, not all of the 54-148 quanta actually are involved in vision.

When light from a distant object falls on the eye lens, it converges on the retina and forms an image there, because for normal eye under perfectly relaxed condition distance of retina from eye lens is equal to its focal length. So, theoretically, normal eye can see objects situated at infinite distance. When eye looks at nearer objects, image is formed behind retina, and image will be either invisible or almost indistinct. Then ciliary muscles compress the eye lens and its focal length gets reduced and image is formed on the retina. Thus, nearer the object to be seen, greater is the strain on the eye lens for adjustment of focal length to form the image on retina.

The radius of curvature of inner surface of eye lens in normal eye can alter upto 40 percent. This adoption of the eye lens in adjusting its focal length so as to always form the image on the retina for various positions of the object is called *Power of accomodation* of the eye lens. Accomodation and pupil size are functions of the smooth muscle cells of the ciliary muscle and the dialator and splincter muscle of the iris. They are termed *intrinsic eye muscle,* since they are inside the eye ball. Convergence of the eye is a function of the voluntary muscle attached to the outside of the eyeball. These muscle are called the *extrinsic eye muscles.*

The normal eye, known as *emmetropin eye,* can sufficiently refract light rays from an object 6 meter away to focus a clear object on the retina. Many individuals, however, do not have this ability because of abnormalities related to improper refraction.

SHORTEST DISTANCE OF DISTINCT VISION

It is that smallest distance upto which eye can see objects distinctly. It can also be defined as the point which is at a distance equal to the shortest distance of distinct vision from the eye is called *near point* of the eye. For normal eye, shortest distance of distinct vision is 25 cm.

The image formed on the retina is inverted but by some unknown mechanism brain interprets the objects to be erect.

TYPES OF VISION

Monocular vision is formed is fishes, amphibians, reptiles and in most of the birds. In *binocular vision* both the eyes can be focussed upon the same object. Birds of prey such as owls and to a lesser degree hawks and eagles have binocular vision. Many mammals possess binocular vision, but it is best developed in human and other primates. In human beings, the vision is binocular as well as *stereoscopic,* that gives three dimensional effect. Certain deep see fishes posses *telescopic vision,* that help them to form image of the objects situated at great distance.

CAUSES AND CONTROL OF VISION LOSS

Following are a few definitions in context of vision loss:

Visual impairement; Includes trouble in seeing with one or both eyes even when wearing glasses or contact lenses.

Severe visual impairement; Inability to read ordinary newsprint even with the best correction glasses or contact lenses;

Low vision: Vision that cannot be further improved by corrective lenses or medical, or surgical intervention, although low vision rehabilitation may help someone to use remaining sight more effectively.

Legal blindness; A central visual equity for distance of 20/200 or poorer in the better eye with correction, or a field of vision no greater than 20 degrees in widest diameter.

Functional blindness; No useful vision, clinically measured light perception less.

Following are a few vision faults with their correction mechanisms;

(a) Aging eye
(b) Myopia
(c) Hypermetropia
(d) Presbyopia
(e) Astigmatism
(f) Colour blindness

(a) Aging eye

Human vision declines with advancing age. Although these are neutral losses, the major decline is due to changes in the eye's optics. First, the lens becomes yellowish, making discrimination of blue colours more difficult. More importantly less light entering the eye reaches the photoreceptors. One problem is that the lens and other optical media become opaque. Further, the pupil shrinks, allowing less light to enter the eye (Table 12). These data show that pupil's response to dim light also decreases with age and becomes virtually nil by the age of 80 years. This means the elderly person have especially large vision problems in low light environments.

Table 12: Shrinking of pupil size with age.

Age (yr)	*Day (mm)*	*Night (mm)*	*Diff (mm)*
20	4.7	8.0	3.3
30	4.3	7.0	2.7
40	3.9	6.0	2.1
50	3.5	5.0	1.5
60	3.1	4.1	1.0
70	2.7	3.2	0.5
80	2.3	2.5	0.2

As a result of all these factors, at the age of 60 years, the amount of light reaching the photoreceptors is only 33 percent of the amount seen at the age of 20 years. By the late seventies, the amount falls to 12 percent. Further, aging reduces light transmission even more as the effect accelerates.

The *contrast sensitivity* of the eyes declines with age. Figure 98 shows how contrast must be increased with age. Using the sensitivity of a 20 year old as the base line, the graph shows the factor by which contrast must be increased in order to maintain visibility level. The required contrast increases gradually to a factor two in the 60's. The loss of contrast sensitivity then accelerates reaching a factor of 6 by age 80.

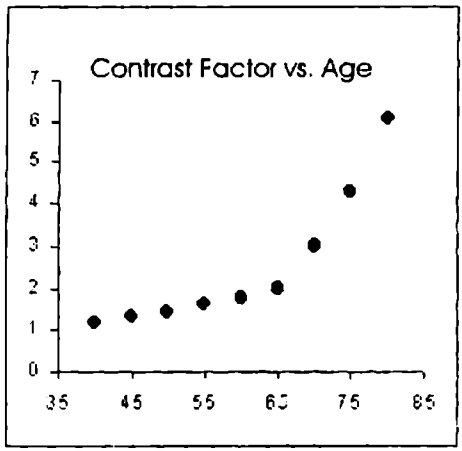


Figure 98: Loss of eye contrast sensitivity with age.

A nutrition supplement have been recommended to give relief, it includes consumption of spinach, lutin (6-12 mg), vitamin E (200-400 iu), selenium (200 meg); vitamin B12 (300 meg), magnesium (400 meg), vitamin C (500-2000 mg), sulphur rich food such as garlic, eggs, asparagus; onion.

Avoid high dose Calcium supplements without balancing magnesium, avoid hydrogenated fats, and avoid very low-fat diets.

(b) Myopia (Or short sightedness)

A person suffering from myopia cannot see distant object. Rays coming from distant object intersect in front of retina and eye is incapable of increasing its focal length. So without external aid, eye cannot see distant objects. Myopia is a result of either the eye lens becoming too much convergent or the eye ball becoming too much elongated.

Obviously, then the rays should be made more divergent before entering the eye so that they might meet on the retina. The patient is given spectacles with concave lenses of suitable power.

Far point of myopic eye is not infinity but at some finite distance.

There is however, no difficulty in seeing nearer objects. Near point of myopic eye also shifts towards the eye i.e., his least distance of distinct vision is less than that for normal eye.

CORRECTION OF SHORT SIGHTEDNESS

Suppose D is the distance of far point for the myopic eye. Let x be the distance of retina from the eye lens; and f_e the focal length of the eye lens. An object at a distance D from the eye produces an image on the retina when eye in unaided.

$$\therefore \frac{1}{(-x)} - \frac{1}{D} = \frac{1}{f_e}$$

Now suppose he puts on glasses, with which he can see distant objects and let f be focal length of this correcting lens

$$\therefore \frac{1}{(-x)} - \frac{1}{\alpha} = \frac{1}{f_e} \frac{1}{f}$$

Substracting (1) from (2) and putting $1/\infty = 0$

we have $1/D + 1/f$

$\therefore f = + D$

Hence, correcting lens for myopic eye should be concave (Fig. 99) and of focal length equal to the distance of its far point.

(c) Hypermetropia (or long sightedness)

In this case, a person can see distant objects without any difficulty but cannot see objects at nearer distances at which normal eye would see. It is the result of either the eye lens becoming less convergent or the eye ball becoming much contracted.

When object is at infinity, image is formed a little behind the retina, but on account of power of accomodation eye brings the image on the retina and so distant objects can be seen.

When an object is placed at a distance of 25 cm, image is formed too far behind the retina, so that even with almost

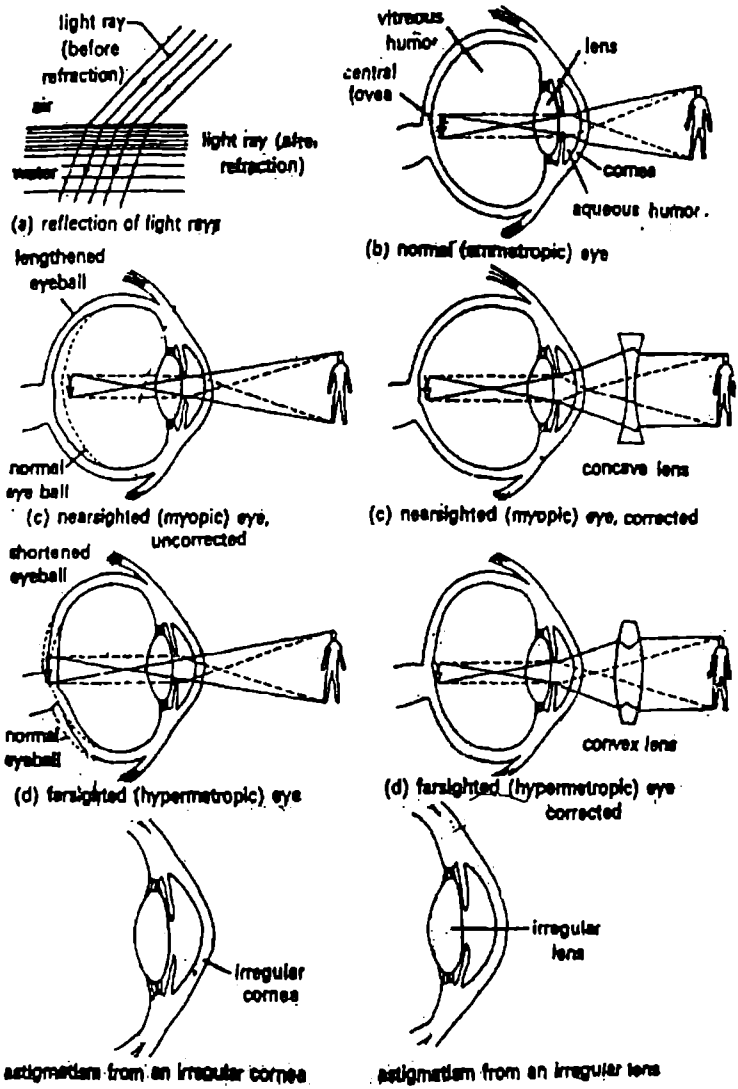


Figure 99: Normal and abnormal refraction in the eyeball.

accomodation image cannot be brought to the retina, and hence object is not visible clearly. Near point shifts farther from the eye.

Obviously, rays should be made more convergent before entering the eye. Hence, a convex lens of suitable focal length should be used. Suppose a person cannot see objects nearer than d cm, which is less than D, the shortest distance of distinct vision for normal eye. Let f be the focal length of the correcting lens which enables him to read like a normal person. Let x be the distance of retina from eye lens and fe the focal length of the eye lens

Then for unaided eye $\therefore \frac{1}{(-x)} - \frac{1}{d} = \frac{1}{f_e}$

and for aided eye $\therefore \frac{1}{(-x)} - \frac{1}{D} = \frac{1}{f_e} + \frac{1}{f}$

substracting $\frac{1}{D} - \frac{1}{d} = -\frac{1}{f}$

D < d, so focal length is negative, which means a convex lens should be used.

(d) Presbyopia

Usually in advanced age eye looses partially its power of accomodation. Neither he can see distant objects nor the objects at normal near distance. So he feels difficulty both in reading as well as in seeing ahead. This type of combined defect in vision is called presbyopia. The person is given two spectacles, one with convex lenses for reading purposes and the other with concave lenses to see ahead. When both the lenses are fitted in the same frame, the system is called *bifocal.*

(e) Astigmatism

In this case eye cannot see with equal distinctness the horizontal and the vertical lines at the same distance. So if he looks at a network of wires, either horizontal or vertical wires are seen more distinctly. The defect is often due to unequal curvatures of the

horizontal and vertical sections of cornea or retina. This is corrected by using *cylindrical glasses.* Sometimes astigmatism is combined with myopia or hypermetropia. In such cases *supra-cylindrical glasses* are recommended.

(f) Trachoma

It is the greatest single cause of serious and progressive loss of sight in the world, often leading to total blindness. It is caused by Virus Conjunctivitis. Trachoma is characterised by many granulations or fleshy projections on the eyelids. The disease produces an excessive growth of sub-conjunctival tissue and the invasion of blood vessels in front of cornea. The cornea may be ulcerated and vision is lost. The disease can also be spread by flies which settle on the eyes.

(g) Conjuctivitis

Inflammation of the conjunctiva may be acute or chronic and generally affects both eyes. Conjunctivitis can be caused by microorganisms or by a number, of irritants including dust, smoke, wind, air pollution and excessive glare.

(h) Ptosis

It is the falling and drooping of the eyelids. In this the upper lid is paralysed and drops down over the eye. The condition may be congenital due to weakness of the muscle of the eyelids.

(g) Strabismus

An eye muscle diorder commonly known as *"crossed eyes"* or *"squint"*. The eye bal do not move in unison and the image does not fall upon corresponding points of the two retinas. It may be caused by lack of coordination of the extrinsic eye muscle. As a result, two images are seen.

(j) Iritis

It is the inflammation of the iris and is usually associated with the inflammation of the ciliary body. It is caused by spread of inflammation from surrounding structures or by injury and sometimes it is syphilitic in origin. In this abnormality, the aqueous

fluid, unable to escape, may cause the iris to bulge forwards. Pains and photophobia may be present. There may be a little lacrymation and blurring of vision.

(k) Colour blindness

The yellowing of the eye optics changes colour vision. The yellow optics blocks blue light and makes blue objects appear black. Colours that contain blue will also look different. Purple and meganta (red-blue) for example, will appear red. Older people cannot distinguish shades of blue very well.

Eyes with this detect fail to have correct impression of colours, if similar paper strips of different colours are placed and the patient is directed to pick-up the strip of a blue, purple or magenta colour he cannot do so.

RETINAL CHIP VISION SENSORS;

In an effort to illuminate the perpetually dark world of the blind, researchers are turning to technology. Advances in semiconductor fabrication technology has led to a new generation of miniature photosensors, the so-called *eye-chips* by combining optics, human vision, and microprocessors are advancing ophthalmology through the new field of *optobionics*. Damaged retinas resulting from debilitating visual diseases, such as *retinitis pigmentosa* and *macular degeneration* as well as aging and injuries to the retina, which rob vision, are being corrected with the implanted eye chips.

The retinal chips are electronic devices that are used to bypass various eye defects, and send impulse directly to the visual part of the brain. The initial experiment were done on development of a retinal chip implant to provide vision to totally blind person. In these experiments, thin platinum wires were inserted into the eye and placed against the retinal surface. Controlled electrical pulses were then applied to the retina. When current was applied through the electrodes, the patients could see the stimulation. Later, electrical stimulation was done by using an array with 25 electrodes, which allowed a blind person to see a large letter. An array of 8 by 8 pixels, that is **64** pixels, was used for optical

character recognition, and an array of 25 by 25 pixels would give a crude image. Not the silicon eye chip contain approximately 3500 miniature light detectors attached to metal electrodes that mimic the function of the human rods and cones. The light detectors absorb incident light refracted by the comes and lens and produce a small quantity of electrical charge that stimulates the retinal neurons. Featuring a diameter of two millimetres (Figure 100) the replacement retina is half as thick as a typical piece of paper and is implanted into a pocket under the damaged retina.

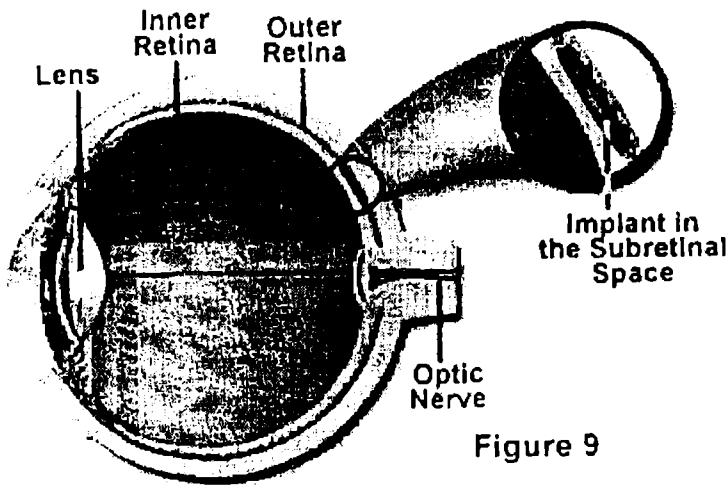


Figure 100: Eye chip synthetic visual sensor.

The artificial vision system that can electronically capture and transmit images to the brain to create sight, consist of a computerized miniature video camera mounted on a pair of glasses, and a sophisticated computer chip surgically implanted in the eye. The camera would transmit its image to the computer chip, which would be connected to tissue in the back of the eye called the retina, so it could transmit image to the brain (Figure 101).

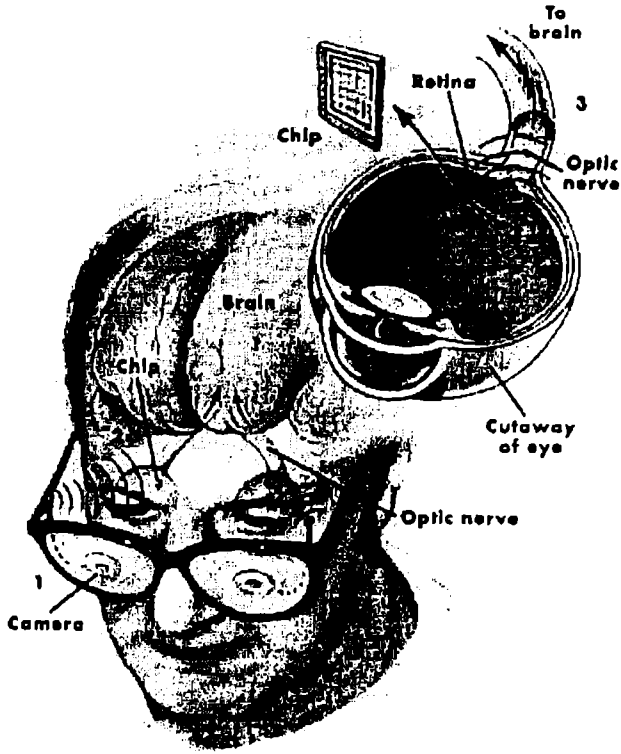


Figure 101: Some artificial vision strategies under investigation intend to replace damaged cells in the eye's retina with a computer chip. For example, one possible system would detect images in a video camera mounted on a pair of eyeglasses (1). Next, the images would be converted into specially patterned electrical signals that wirelessly transmit to a computer chip implanted in the retina area (2). The chip then would stimulate the remaining healthy portions of the retina and visual pathway and, if all goes well, produce the perception of an image in the brain (3).

People most likely to benefit from retinal chip biosensor are those suffering from retinitis pigmentosa, are inherited disorder that causes blindness by destroying the retina, and age related macular degeneration, another condition that causes a deterioration of the central part of the retina in the elderly.

DOBELLE ARTIFICIAL VISION SYSTEM

Dobelle artificial vision system consists of a sub-miniature television camera and ultrasonic distance sensor mounted on a pair of eyeglasses worn by the patient (Figure 102). The sensors are connected to a miniature computer, which is worn on a belt-pack. The computer processes the video and distance signals, reduces the noise, and triggers a microcomputer to transmit pulses to a 68 platinum electrode array implanted on the surface of the visual cortex. Each electrode produces 1-4 closely spaced phosphenes. The external computer package is about the size of a dictionary and weighs 10 pounds, with the computer battery lasting about 3 hours. It requires recalibration each day, and ongoing training of the user. The purpose of the Dobelle artificial vision system is to promote independent mobility, but not for reading. It is contraindicated in blind people with severe chronic infections and those blinded by stroke or cortical trauma.

Results of Dobelle artificial vision system implanted in one 62 year old person, who lost vision in one eye at the age of 22 and in the other eye because of trauna at the age of 36 years. The patient could count fingers and achieved a 20/1200 visual acuity with tunnel vision. With the addition of a more powerful computer and further magnification, the patient's visual acuity improved to 20/400.

Dobelle institute, Portugal plans further testing of this technology, using larger electrode array, more powerful computers and more sophisticated image processing, The Doubelle artificial vision system is not commercially available. Additional basic research to prove the concept and clinical trials to demonstrate safety and effectiveness will probably take several years and is critical for developing a base of scientific evidence in order to make decisions regarding clinical applications.

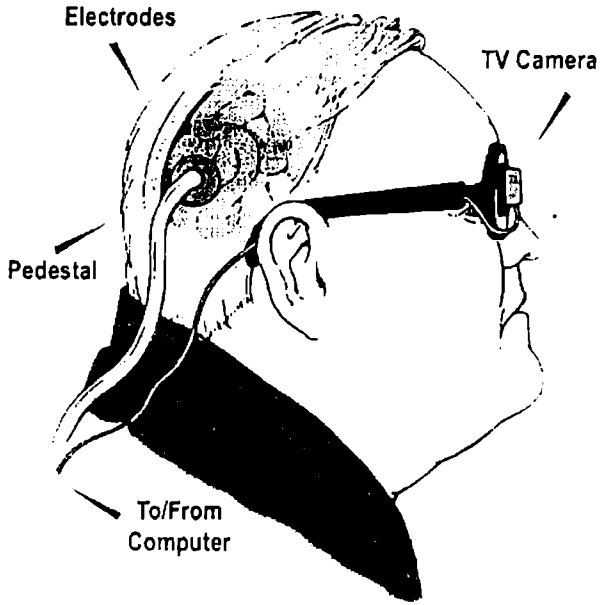


Figure 102: The Dobelle arteficial vision system,

BIOLUMINANCE

It is the emission of visible light at ordinary temperature due to exergonic chemical reaction and represents a reversal of photochemical reactions. Apart of the heat of reaction raises some molecules to a higher electronic state from which they return to the low energy state by emitting visible light of special wavelengths. Chemiluminescence is the cause of bioluminescence in many biological organisms.

Out of the 33 phyla and 80 classes, bioluminescence is found in 13 phyla and 28 classes. These include among animals -fireflies, protozoa, flagellates, crustaceans, deep-sea fishes like whale fish, lantern fish, hatch fish and viper fish, adult and larval glow worms, deep rat-tail fish, sponges, hydroids, marine worms, sea spiders. Plants show bioluminescence too, among bacteria, yeast, fungi, rusts and mushrooms.

SITE OF BIOLUMINESCENCE

In animals, the light is emitted from the following structures:

1. Specialized glandular organs from which secretory material conies out for the production of light and this secretion is squeezed out on stimulation of contracting muscles;
2. Glandular cells present on the surface of the body which secretes materials on stimulation;
3. Intracellular luminescence which is produced due to nervous or hormonal stimulation; or
4. Small particles scattered throughout the cell contents which may be broken by any kind of stimulation.

The light may be turned off and on in all of these animals The special organ which emit light continuously may have some mechanism for turning the organ so that its luminous surface alternatively open or close or may cover with flaps down over the organ. In some species, the luminescence may occur only when two or more substances are brought into contact producing a "flash". Bacteria are generally continuously luminescent, in some cases the luminescence of animals is the result of symbiotic relationship between the bacteria and animal.

STRUCTURE OF LIGHT ORGAN

The firefly *(photinus pyralis)* is an example of an organism containing a definite light organ, which may be activated by nervous influences to emit a bright flash of light, the luminescen being intracellular. In the male, the organ occupies the whole ventral surface of the 6 and 7 segments whilst in the female only about 2/3 of these segments are occupied. The structure of the organ of *Photinus* consists of two layers (a) reflecting layer, an (b) photogenic layer (Fig 103)

The cells of the reflecting layer contain crystals of some urate which give the layer of reflecting property, whilst the photogenic layer is made up of large cells filled with yellow granules. An adequate supply of oxygen is ensured by the numerous trachae which penetrate both layers and branch profusely in the photogenic layer; each branch ends in a trachael end cell which give rise to tracheal capillaries or tracheoles which are considered to enter the

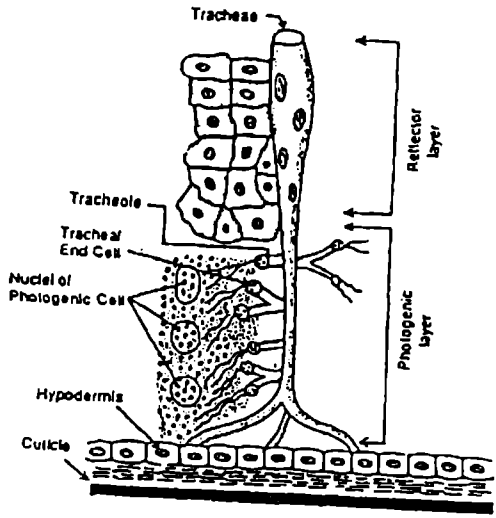


Figure 103: Showing structure of firefly light-organ

photogenic cells. The end cell is peculiar to the tracheal system of the light organ, and is remarkable for the large number of mitochondria it contains. The tracheolar twig is surrounded by the end cell and, more directly, by a rounded body that has been called the end bulb. Before emerging from this, it breaks into two tracheoles which pass between the photogenic cells, the membrane and cytoplasm of the end-bulb investing the outer layer of the photogenic cells are differentiated from the inner regions by being free of photogenic granules and containing numerous small mitochondria. The photogenic granules have a laminar basis.

MECHANISM OF BIOLUMINESCENCE

Bioluminescence is produced due to chemical reactions which follow the usual kinetic rate theory. Chemiluminescence drives energy for light production from the chemical reaction. Most luminescence reactions occur with oxidation. Dubois (1887) postulated that the essential feature of luminescence was the oxidation of a substance called *"luciferin"* - the constituent of the coloured granules of photogenic cells of certain, beetles, ostracods, molluscs and decapods by means of an enzyme *"luciferase"* found

in water extracts of the same organ. Luciferase catalyses the breakdown of luciferin. Luciferin from one species does not react with luciferase of another species.

In some glands, the enzyme and substrate are secreted by separate cells, but only the combined secretions can exhibit potentiality for luminescence. In others, the enzyme and substrate are secreted by the same cell. In intracellular luminescence, both enzyme and substrate must be produced in the same cell.

All the bioluminescence reaction appear to be enzyme substrate type wherein from luciferin, oxyluciferin is produced. The activated oxyluciferin molecule reverts back to the low ground state by emitting the excess energy as visible light. As much as 90 percent of the free energy change is transformed to light in fireflies. The aerobic oxidation of each luciferin molecule generates 1 photon of light. The emitted light flashes about 10-20 times per second, every time for a fraction of a second. It has intensity of less than 0.025 of a candle-power. Variation of wavelength being probably due to the species differences in the primary structure of luciferase. Some animals emit bluish light, others emit greenish or yellowish and in some cases reddish light may be seen.

SIGNIFICANCE OF BIOLUMINESCENCE

The bioluminescence of *Phontius pyralis* (firefly) has significance in the reproduction. The female is wingless and in the evening preches on a blade of grass whilst the male, who during the day remain quiescent, fly about emitting flashes. The female never flashes spontaneously, but only in response to the flash of a male within some 3-4 meters, if the female makes this response, the male immediately turns directly towards her and flashes again, the female flashes in response and after about five of these exchanges the insect mate. The ability of the male fly to distinguish a flesh emitted by a female, from one emitted by a male is not due to any difference in the spectral quality of the light emitted by the two sexes, but depends entirely on the interval elapsing between his own flash and the answering one. If this is precisely 2 seconds, the male responds to it, whereas, he ignores any flash occurring at other intervals after his own.

Chapter 6

BIOPHYSICS OF SONIC VIBRATIONS

The word "sound" denotes a periodic mechanical disturbance in gases, fluids or solids. Sound waves are longitudinal mechanical waves. The material atoms and molecules transmitting such a wave oscillate in the direction of propagation of the wave itself. There is a large range of frequencies within which longitudinal mechanical waves can be generated, sound waves being confined to the frequency range which can stimulate the human ear and brain to the sensation of hearing. This range is from about 20 Hz to about 20 kHz and is called the *audible range.* A longitudinal mechanical wave whose frequency is below the audible range is called an *infrasonic wave,* and one whose frequency is above the audible range is called an *ultrasonic wave.*

Infrasonic vibrations are usually generated by large sources like earthquake, thunder, high winds, ocean waves and volcanic eruptions. The ultrasonic vibrations may be produced by elastic vibrations of a quartz crystal induced by resonance with an applied alternating electric field (piezoelectric effect). It is possible to produce ultrasonic vibrations as high as 6x108 Hz in this way.

Audible vibrations originate in vibrating strings (violin, human vocal cords), vibrating air columns (organ, clarinet), and vibrating plates and membranes (xylophones, loudspeaker, drum). All these vibrating elements alternately compress the surrounding air on a forward movement and rarefy it on a backward movement.

The air transmits these disturbances outward from the source as a wave. Upon entering the ear, these waves produce the sensation of sound. Waveforms which are approximately periodic components give rise to a pleasant sensation. Sound whose waveform is non-periodic is heard as noise.

Audible sound is the effect on the ear of a wave like motion of an elastic medium caused by vibrations. Vibrations impinge on the ear drum of a human and/or animal and setup a nervous disturbance, which we call sound, to hear a sound we need a vibrating object for generating sound waves as well as receiver - the ear - for absorbing the energy which has been passing through the transmitting medium.

The word sound is commonly used in two different senses: (1) to denote the *sensation* perceived by means of the ear when the auditory nerves are excited; and (2) to denote the external physical disturbance which, under ordinary condition, suitably excites the auditory nerves.

Three things are imperative for the production of sound:

(a) some medium to receive and transmit the vibratory motion;
(b) the parts of the body in vibratory motion should have such shape, size and motion as to cause a disturbance to advance through the medium (air, water etc.), and
(c) our ears should "enable us to perceive the sensation of sound only when effected by to-and-fro movements, whose number per second lies between certain limits".

Vibrations are defined as the period which is no other than the time from the instant when the vibrating point passes through any position to the instant when it next passes through the same position, moving in the same direction. Sound vibrations may be continuous, intermittent, impulsive or explosive. Every unit of vibration is possessed by a *frequency,* and the frequency of a vibration is the number of vibrations per unit time. The *amplitude* of a vibration is the maximum displacement, assumed by a vibrating point in the course of its motion. All sound vibrations are characterised by three features -*pitch, intensity* and *quality.*

The pitch of a sound depends upon the period or frequency of the vibrations constituting the sound. The greater the frequency, higher the pitch. Pitch is specified in two distinct ways: (a) scientifically, by the statement of the period or frequency of the vibrations, or by the logarithm of its frequency, and (b) by the statement of the period or frequency of the vibration, or by assigning to the sound in question its position in a certain accepted series of sounds.

The intensity of the sound waves is purely physical quantity. It is proportional to the wave energy passing per unit time through unit area. Thus we find that a pitch depends upon the frequency, where as an intensity upon the amplitude of the vibration. The loudness of a sound depends upon the intensity of the waves producing it, and increases and decreases with the intensity for a sound of given frequency.

The quality of a sound serves to distinguish between sounds of the same pitch and intensity.

PHONORECEPTORS

The phonoreceptors are organs concerned with the sensation of hearing. Stimulation of these receptors takes place mainly by means of sound waves in water or air so as to excite the nerve impulse. Fishes and amphibians are capable of perceiving the sound wave in water. Other organisms perceive sound waves in air.

In invertebrate animals, the phonoreceptors are called *'statocysts'*; which are compact mechanosensitive organs consisting of a solid material 'statolith' (heavy calcarious structure), usually held together by a gelatinous substance. Delicate hairs are protruded into the statocyst lumen from the statocyst wall. These hair like filaments are activated by sound vibrations. The impulses are relayed to nervous system to provide information.

In vertebrates, the most primitive phonoreceptors is lateral line system found in fishes and few amphibians. This is *statoreceptor.* The individual cell of each group of neuromast (Fig. 105) have widely different threshold to mechanical disturbance in water. A neuromast of a fish comprises of a cluster of pear shaped secondary

sensory cells supported in a basket like arrangement of tail epithelial cells.

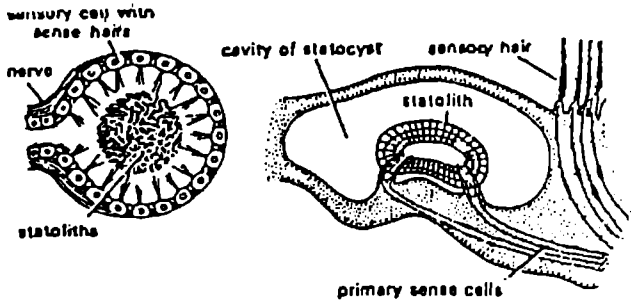


Figure104: Statocysts of invertibrate animals, Left, *Pecten*; Right, Leptomysis

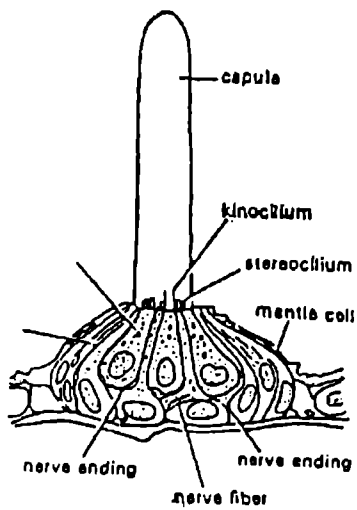


Figure 105: Structure of a neuromast of fish.

In Arachnids the phonoreceptors are called *Lyriform organs*. They are comprised of parallel slits in the cuticle. The slits are of varying length, like the strings of a lyre, covered by a thin cuticle, containing nerve endings from the sensory neurons (Fig. 106).

In Orthoptera, Hemiptera and Lepidoptera the phonoreceptors are called *hair sensilla,* which vibrate in their tiny sockets and pull on the nerves. Their sensory axons respond to sound frequency between 50 to 400 c/s.

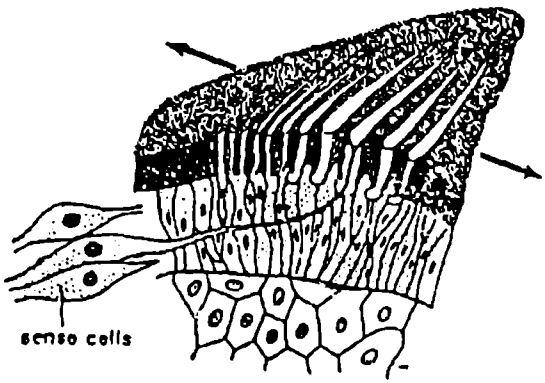


Figure 106: A lyriform organ of spider.

In Mammals the phonoreceptors are called *ear.* The ear is differentiated into three parts: external, middle, and inner ear (Fig. 107).

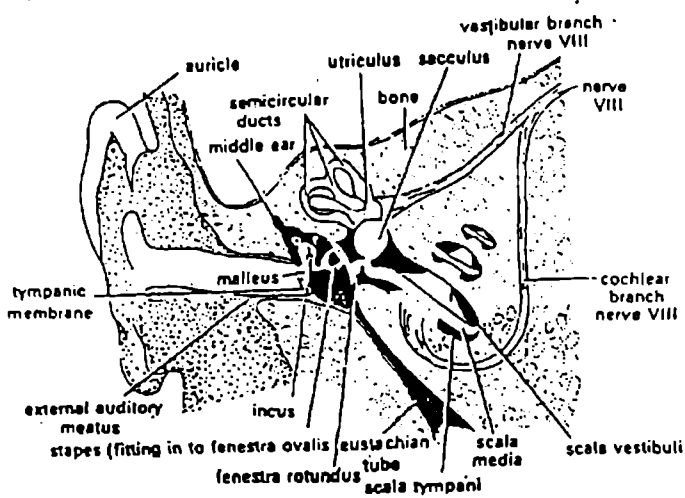


Figure 107: Mammalian ear.

The external ear has two parts: *external pinna* is found in most of the mammals (the exceptions include prototherians, marine mammals such as whale, seals and tortoises), and an *ear-canal* or *auditory meatus*. The pinna is used to select sound waves from certain frequencies and concentrate them in the ear canal.

The ear canal is an S-shaped canal in the temporal bone that extends from the pinna to the eardrum, a distance of about 4 cm in human beings. Its lining contains numerous *ceruminous glands* which secrete *cerumen* (ear wax). The wax lubricates the canal and prevents foreign objects from entering the ear. Overproduction of ear wax can block the canal and impair hearing.

The *Tympanic membrane* or the eardrum is a delicate, thin, semitransparent, elliptical disc situated between the external and middle ear. The greater part of the eardrum which remains fitted into the bony furrow of the tympanic ring is taut, and is called the *pars tensa,* the other smaller part of the eardrum facing froward and upwards and directly attached to the incisure in the squamosa known as the notch of Rivinus is lax and is called the *pars flaccida* or *Shrapnells membrane* (Figure 108). The eardrum consists of three layers - an outer or epidermal layer continuous with that of the auditory meatus, a middle layer of radiating and circular connective tissue fibres, and an inner layer of mucosa, continuous with the mucous membrane of the tympanic cavity. Shrapnells membrane consists only of two layers and lacks the middle stratum of fibrous tissue.

In the early childhood, the drum is comparatively thick owing to the pressure of a loose submucous layer. It grows compact with time and in old age becomes quite thin. The drum is placed obliquely to the long axis of the auditory meatus, so that it faces forward, downward and inwards. In the new born and breast-fed babies, it is placed in a almost horizontal position.

Middle ear is not present in caecilians, caudate amphibians, and a few salientians. Snakes also lack this part. It is well developed in birds and mammals. The cavity of the middle ear (tympanic cavity) is connected with pharynx by means of *eustachian duct.* In tongueless toads and in birds both the eustachian ducts open by a single, median outlet into the pharynx. The eustachian ducts

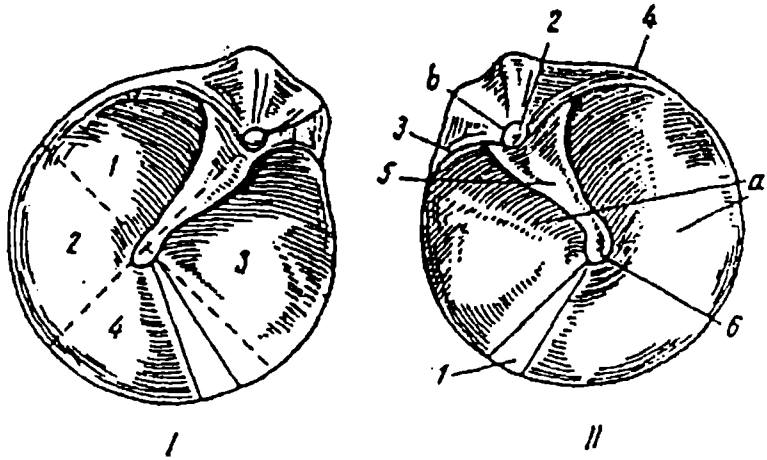


Figure 108: Tympanic membrane, right and left sides.

I. Right drum is divided into four quadrants
(i). posterosuperior. (ii). posterointerior
(iii). anterosuperior (iv). anteroinferior
II. Left drum (a) pars tensa, (b) pars flaccida,
(i) light reflex, (ii) short process of malleus, (3&4) anterior and posterior folds, (5) handle of malleus, (6) umbo.

equalize air pressure on both sides of the tympanic membrane. If the air pressure is not equalized, the tympanic membrane is unable to vibrate in response to sound waves and may even rupture.

Within the tympanic cavity are three small bones or *ossicles:* the *malleus* (hammer), *incus* (anvil), and *stapes* (stirrup). These bones are modified from articular, quadrate and hyomandibular respectively. Birds have only one ossicle (the *columelld).* In mammals, the three bones act as a liver system performing an amplification function. The *tensor tympani* inserts on the malleus and is innervated by the trigeminal nerve. The *stapedius muscle* inserts on the stapes and is innervated by the facial nerve. Protection against loud voices is provided by these muscle. The tensor

tympani tightens the eardrum at tympanic membrane to reduce its vibrations and the stapedius pulls the stapes away from the oval window to reduce the intensity of the fluid waves.

The inner ear consists of two main divisions: (a) *bony labyrinth,* and (b) *membranous labyrinth* that fits in the bony layrinth. The bony labyrinth consists of a system of canals and cavities in the pertrous portion of the temporal bone. It has the *cochlea* (snail shell), *vestibule* and *semicircular canals.* The bony labyrinth is lined with periosteum and contains a fluid called *perilymph.* The membraneous labyrinth is lined with epithelial cells and contains a fluid called *endolymph.* The four major divisions of membranous labyrinth are the *cochlear duct,* which is suspended in the *cochlea.* The *utricule* and *saccule,* which are suspended in semi-circular canals. Both utricule and saccule are connected to each other by a small duct, the *utriculo-saccular duct.* From this duct, a slender endolymphatic duct extends to the inner surface of the pertrous bone where it enlarges into the *endolymphatic sac* lying just beneath the dura. The saccule is also connected with the cochlear duct by a slender canal, the *ducts reuniens.*

From the vestibule, three bony semicircular canals project upwards. Each is arranged at approximately right angles to the other two. On the basis of their positions, they are called the superior, posterior, and lateral canals. One end of each canal enlarges into a swelling called the *ampulla.* Inside the bony semicircular canals lie membranous *semicircular canals* or *semicircular ducts.* In the fish *(Myxine)* only one semicircular duct is present. In lamprey *(Petromyzon)* two vertical ducts are present. Birds and animals have three semicircular canals (Figure 109).

The bony cochlear canal, in which lies the cochlear duct, spirals about a conical axis, the *modiolus.* The number of turns in the cochlea of various mammals varies. The cochlear duct becomes attached along two sides of the bony canal thus dividing the perilymphatic space into the *scala vestibuli* and *scala tympani.* Near the apex of cochlea, the vestibular and tympanic scalae becomes continuous and this common space is known as the *helicotrema* that responds to vibrations of low frequencies. The

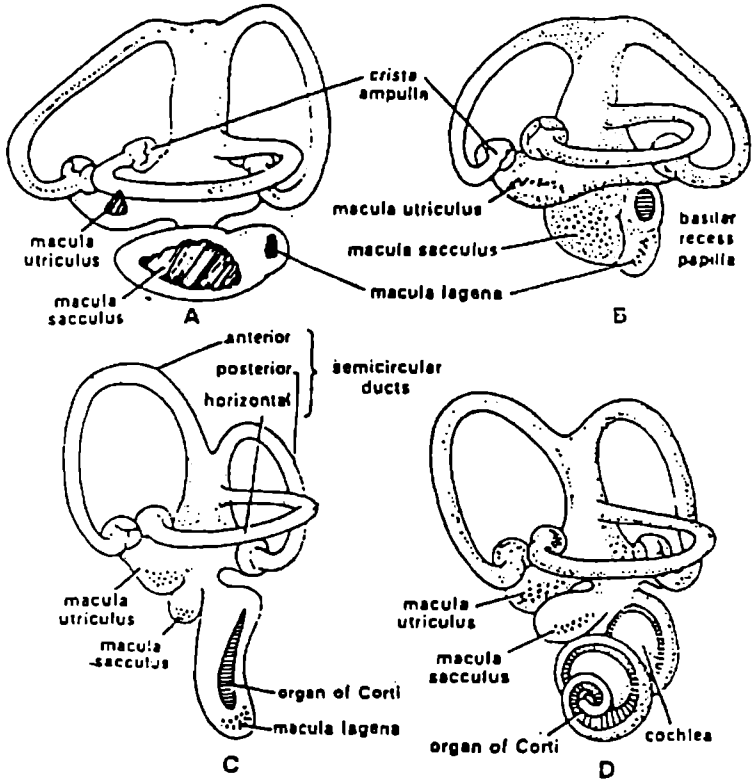


Figure 109: Membranous labyrinth (A) in Myxine fish, (B) Petromyzon turtle, (C) birds, and (D) Mammals.

scala tympani terminates at the *round window.* At the basal end of the scala vestibuli is the *oval window.* The cochlear duct forms the *scala media* which is separated from the scala vestibuli by the vestibular membrane or *Reissner's membrane* and from the scala tympani by the *basilar's membrane.* Resting on the basilar's membrane is the spiral organ or organ of Corti or organ of hearing (Figure 110).

The organ of Corti is a series of epithelial cells on the inner surface of the basilar membrane. The epithelial cells are modified into *supporting cells* and *hair cells,* which are receptors for auditory

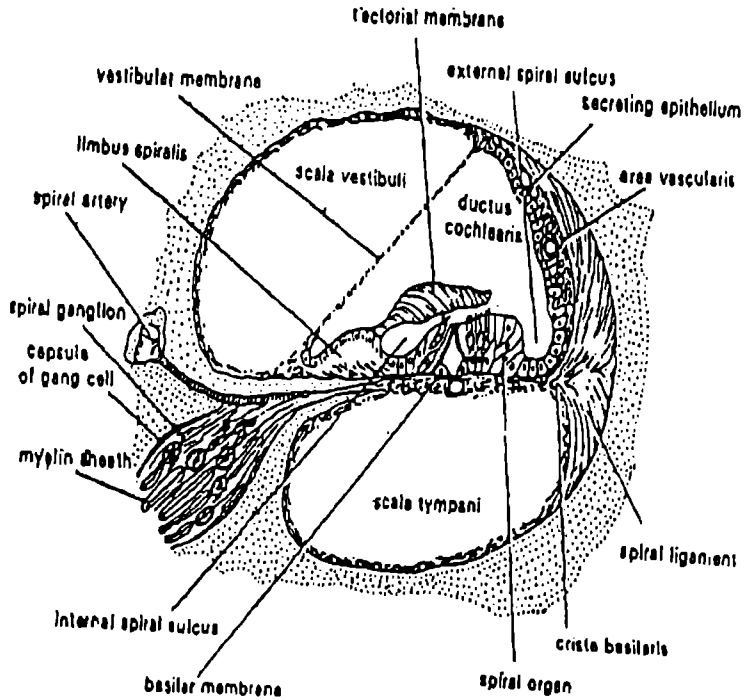


Figure 110: The structure of a mammalian cochlea.

sensations (Fig. 111). The hair cells in the human ear are estimated to be from 13000 to 54000, each one with perhaps 40 cilia or hair at the receptive end, projecting in the endolymph. The hair cells are arranged in two bands, the hair cells form a single median row that extends the entire length of the cochlea. In outer band the hair cells are arranged in several rows, there being three in the basal coil of the cochlea, four in the middle coil and five in the apical coil.

AUDITORY FUNCTION

The auditory function of the ear consists in the conduction of sounds through the external and middle ears or cranial bones and their reception by the organ of Corti, the receptor of the auditory analyzer. The external and middle ears makes up the

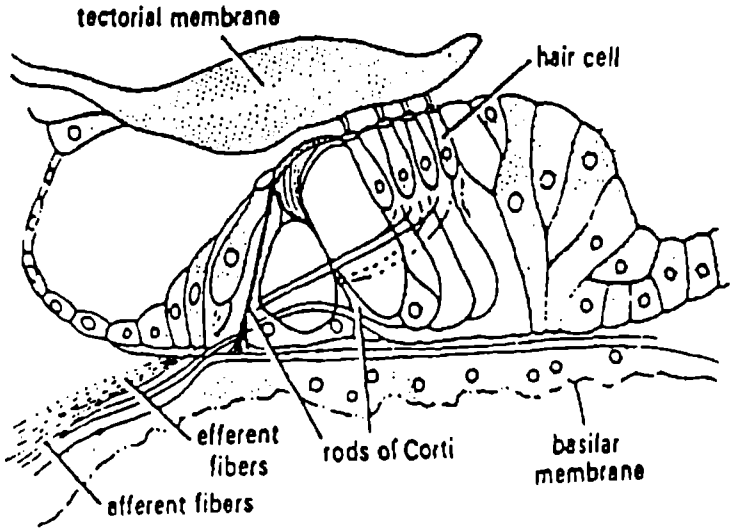


Figure 111: Arrangement of hair cells in the Organ of Corti.

sound-conducting apparatus, whereas the internal ear, specifically, the organ of Corti, belongs to the sound perceiving apparatus.

The auricle in man is of lesser importance than in animals, and yet there is no doubt that it plays a certain role in collecting sounds and determining their direction.

The external auditory meatus conducts sound waves from the outer medium to the tympanic membrane. The meatus diameter has nothing to do with hearing activity. Its atresia, however, as well as its complete obstruction, as occurs in ear wax impaction, hinders the passage of sound waves and considerably impairs the hearing.

Sound waves striking the tympanic membrane set it into vibration. The drum being connected to the handle of the malleus, these vibrations are transmitted to the ossicular window of the labyrinth, rocks in and out of the oval window according to the phase of sound vibrations. The vibration of the foot plate of the stapes in the oval window sets up vibrations in the perilymph.

These vibrations are transmitted to the basilar membrane and the Organ of Corti which it supports.

The vibration of the basilar membrane causes the hair cells of the spiral Organ of Corti to get in touch with the overhanging tectorial membrane. At the same time, the mechanical energy of vibrations change into the physiological process of nervous excitation which is conveyed to the most delicate receptors of the auditory nerve to be passed further to its nuclei in the medulla oblongata and through appropriate canals to the cortical auditory centres in the temporal brain lobes where nervous impulses are interpreted as sounds heard.

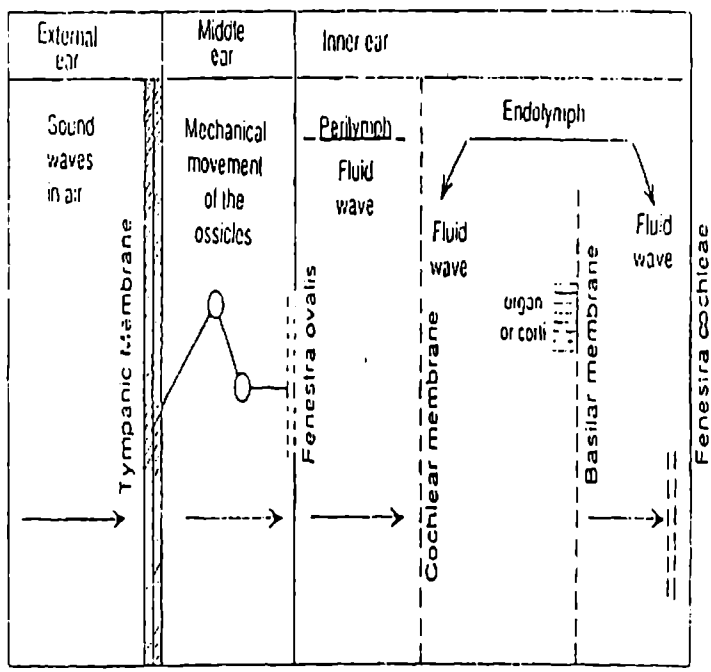


Figure 112: Showing transmission of sound through the ear.

The internal ear performs the most important functions of the ear, because it is here that sound perception takes place.

Normal hearing depends on the normal condition of the apparatus for sound perception and conduction.

The tension of the drum and the ossicular chain necessary for normal sound conduction is maintained by the combined action of the tympanic muscle. For normal vibration the tympanic membrane requires a constant equilibrium between air pressure in the middle ear cavity and in the outer air, that is on both sides of the drum. This is maintained by the passage of air through the Eustachian tube during swallowing. Disturbance of air supply to the middle ear through the Eustachian tube cause air in the middle ear to be sucked in and the drum to be indrawn, which is followed by deterioration of hearing. The normal condition of the sound-conducting apparatus is extremely important for the transmission of low tones to the labyrinth, that is, sounds with a low frequency of vibrations per second.

There are two ways of conducting sound waves to the labyrinth; air conduction (through the external auditory meatus, the tympanic membrane and the chain of ossicles); and bone conduction (directly through the cranial bones and the stapes).

High tones, that is, sound of a high vibration frequency are easily conducted to the labyrinth not only through the tympanic membrane and the ossicular chain, but through the cranial bones and the stapes as well.

Human being can hear external sounds with a frequency of 16 to 20000 cycles per second.

The human ear can differentiate between sounds of different pitch, intensity and timbre. There are a number of theories which seek to explain the essence of hearing and the ability of the ear to differentiate between sounds:

(a) RESONANCE THEORY

This is the oldest theory advanced by Helmholtz in 1863 and based on the physical phenomenon of sympathetic vibrations. According to this theory, the fibres of the basilar membrane

vibrate in unison with sounds, similar to the action of strings in certain musical instruments, such as the Piano or the Harp. The short, thin tighter fibres on the basilar membrane which lie in the basal turn of the cochlea resonate in response to low tones. If the tone is soft, the basilar membrane vibrates gently, and a stream of impulses of low frequency ascends the cochlear nerve. If the tone is loud, then vibrations of the basilar membrane are stronger and the frequency of the nerve impulse is greater.

There are a number of serious objections to the resonance theory as it oversimplifies the essence of hearing as a physiological process by describing it from the mechanical aspect alone, and fails to give a picture of the physiological properties of the auditory analyzer as a whole. It should be noted, however, that the localization of perception of high and low tones in the basal and apical cochlear turns respectively, on which the resonance theory is based, has been confirmed by experiments and clinical observations.

(b) TELEPHONE THEORY

This theory asserts that the basilar membrane vibrates all over like a diaphragm in a telephone, which transmits whatever sound enter the ear without any change by the cochlear nerve to the brain. It denies any analysis of sound being made in the peripheral receptor contained in the cochlea. This concept has been disapproved by clinical and by experimental research.

LOCATING THE ORIGIN OF SOUND

The faculty of locating the origin of sounds, the so called *Ototopia* depends upon the binaural hearing. It is largely lost in people with unilateral hearing, who have to turn their heads in various directions to locate the origin of sound. People with two healthy ears can easily determine the direction of sounds without turning their heads.

The ability to find the direction of sounds is a function of the Central Nervous System. If a sound comes from one side, it arrives at the ear on the other side with an insignificant delay of 0.0006 second. This delay makes it possible to determine the direction of sound.

ECOLOCATION

It is a mechanism of orientation in which an energy form emitted by an animal interacts with features of the environment and is returned to the animals where any change in the energy can be detected. In most cases, the energy is in the form of high pitched sounds.

Moths can hear when they approach within 120 feet as their tympanic organ responds to sounds with frequencies between 10 and 200 kHz. Bats can detect moths when about 15 to 30 feet distance. Bats while flying emit ultrasonic sounds.

The laryngeal apparatus is uniquely developed in bat. In this, the crico-thyroid muscles are greatly developed and the arytenoid cartilages are ossified and fused. Some bats fly with their mouth continuously open, while others fly with closed mouths and emit cries through the nose.

Some marine animals use echolocation system to detect the location of objects. Tortoises can emit high frequency sounds both at rest and while swimming under water. Tortoise is without vocal cords so it emits 'whistles" or clicks which are used for ecolocation. Whales also produce sounds in the form of 'songs', which are specific sound frequencies repeated constantly. In the depth at which these whales live,, such sound may carry for hundreds of miles. Whale sound are presumably a form of communication rather than an echolocation system. The oilbird *(Steatornis* of Venezuela) flies in dark caves echolocating on clicks it can, produce with a frequency range upto 10 kHz. Small shrews use echolocation system in moving about underground.

SENSITIVITY OF A DETECTOR AND THE WEBER-FECHNER LAW

It is a fact that whether or not a receiver will detect a signal depends upon how much the signal differs from the background noise. The dependence is not a simple proportionality, but rather a logarithmic one. Thus, the sensation or loudness L is given by

$$L = \log I/I^{\circ}$$

Where I° is background intensity, and I is the intensity, over background, of the signal to be detected. This is the basic form

of the Weber-Fechner Law. It has many manifestations. For instance, if there are two signals equally along, with different backgrounds, the resolution of (difference in loudness), L_2-L_1 is related to the ratio of the intensities of the two backgrounds I and I_2 as follows:

$$L_2\text{-}L_1 \propto \log I^\circ_1/I^\circ_2$$

Because of this logarithmic law, it is convenient to express power ratios by a logarithmic unit, so that sensation becomes approximately linearly proportional to this unit. This unit is called the "bel"(b) and is equal to the logarithm of the ratio of two sound intensities if they are in a ratio of 10:1. The number of bels then is given by

$$b = \log I/I^\circ$$

For sound, the value 1° is arbitrarily chosen to be the lowest one which a human ear can detect (10^{-16} w/cm^2; or in pressure units 0.0002 dynes/cm^2, since the same conversion factor applies to numerator and denominator). The bel unit is too large for convenience, and the decibel, one tenth of a bel, has received wider use. Therefore, the number of decibels is

$$dB = 10 \log 1/1^\circ$$

Another form of the Weber-Fechner law, then is

$$L = dB$$

It holds true for all sensory receptors.

Sensitivity S of a detector, or discernment per decibel of signal over background, is defined as

$$S = \log I^\circ/\Delta I$$

Where $\Delta I_1 = I_1 - I^\circ$. Sensitivity is higher, the smaller is the value of AI_1 Usually when S is determined at different values of an independent variable, the result is expressed as the sensitivity relative to the maximum value taken as unity (S/S_{max}). The sensitivity of the average ear have been shown in Fig 113. The sensitivity curve varies from individual to individual.

Deformation in the structure of the ear, or failure of the ear to respond to sonic vibrations, is corrected by amplification of the

signal reaching the tympanic membrane (using hearing aids), or by stimulation of the bone structure to free the "frozen" level system by surgery.

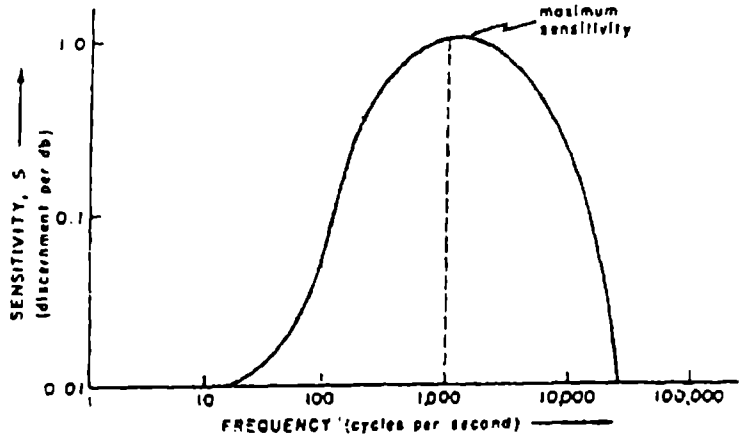


Figure 113: Sensitivity of the human ear at different frequencies of sound

HEARING AIDS

It is probable that several thousands of years have passed since man first cupped his hand behind his ear to enhance his failing hearing, but it remained for the unprecedented technologic advancements of the present century to spawn the modern hearing aid. Throughout the attempts to produce hearing aids that operated on electric power. The early instruments were bulky, and nonportable. By 1920-1930 miniaturised hearing aids were introduced. Fig. 114 shows a block diagram of the functional parts of a basic hearing aid which represent the majority of hearing aids that are manufactured today.

Basic Hearing aid Components

The basic hearing aid components include transducers, amplifiers, volume controls, switches, power sources and electromagnetic devices.

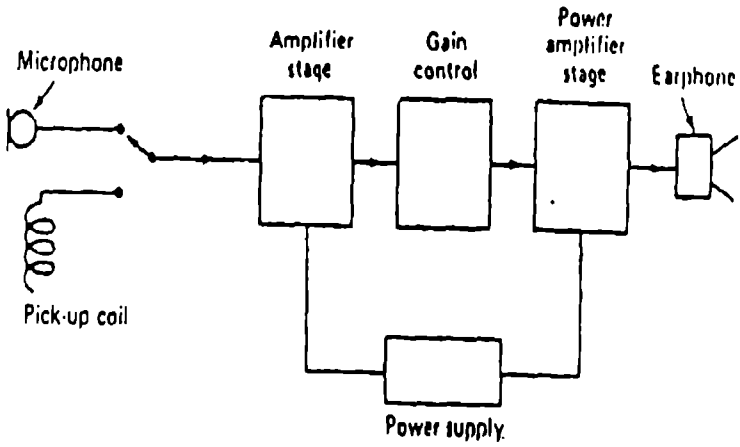


Figure 114: A functional block diagram of a hearing aid with microphone and pickup coil input.

Transducers

The hearing aid microphone and the receiver function to transform one type of physical energy into another, so they are referred to collectively as transducers. The microphone, or input transducer, provides the means for sound entry into the hearing aid. It responds to vibratory sound waves in the vicinity and translates these vibrations into electrical impulses, which are ultimately processed by the hearing aid circuitry. The receiver, or output transducer is a subminiaturised loudspeaker which reverses this process. It converts the electrical impulses back into vibratory sound waves capable of stimulating the ear.

The operational performance of any transducer is determined by its size and various aspects of its composition. Limitations imposed by such features preclude any reasonable expectation for any transducer to operate with equal efficiency over an infinite range of sound frequencies and intensities, so there is inherent distortion at any point in the sound processing system at which transduction occurs. However, the transducers used in modern hearing aids have been designed to accomplish the necessary energy conversion with remarkably minimal distortion.

Although sound amplifying systems are susceptible to several different types of distortion, the main types ordinarily considered are frequency distortion, and amplification distortion. Frequency distortion is the result of resonances and other factors which cause a transducer to respond better to some sound frequencies than to others, and to still others, not at all.

Thus, transducers may be classified with regard to their frequency range, the range of frequencies over which the response of the transducer is relatively equal. Amplitude distortion also occurs as a result of physical limitations that prevent the transducer from responding with equal frequency to infinite variations in the intensity of the stimulus. The range of amplitudes over which a transducer follows changes in sound intensity with relative fidelity is referred to as the dynamic range of that transducer, and the intensity beyond which it ceases to amplify further is referred to as its saturation level.

Amplifiers

Input sound pressures detected by the hearing aid microphone are translated into very weak electrical impulses, of magnitude expressed in thousandths of a volt. Because the amplifier must exaggerate these minute signals in order that they may stimulate the ear adequately, it is the heart of the hearing aid system. The physical limitations mentioned for transducers also apply, atleast in part, to amplifiers, for they are not capable of equally efficient operation over an infinite range of sound frequencies or intensities.

Volume Controls

The volume control provides the means by which the patient may modify the intensity of the amplified sound stimulating the ear. It is mandatory that the control be accessible, for it must be adjusted periodically to meet changing acoustic conditions. In outward appearance, the volume control is usually a small wheel which must be manually rotated in one direction to increase loudness or in the opposite direction to decrease it. In few hearing aids, the volume control is a slide mechanism. In either case, movement of the wheel or slide is mechanically transferred to a variable resistor, altering the electrical resistance and, thus, the voltage developed at that point in the circuit.

Switches

Every hearing aid is provided with some means of interrupting the flow of power to the circuit. Whether this is accomplished by a conventional on-off switch or by physically disconnecting the battery from the circuit, this function is necessary to save the loss of energy when the hearing aid is not in use.

Power Sources

Every hearing aid operates on electrical power provided by some type of battery. Every battery has a finite operating life and must be replaced periodically. Rechargeable batteries made from nickel-cadmium alloy are used these days. The inadequacy of sunlight at various geographical locations, has limited the use of solar battery in hearing aids.

Electromagnetic Devices

Most modern hearing aids incorporate an electromagnetic device known as inductance coil. The coil is a metal core wrapped with fine wire and has the capacity to detect magnetic energy in its vicinity. Such coils are installed within hearing aids circuits solely for the purpose of enhancing hearing via the telephone, and are therefore often referred to as telephone pickups.

The number of hearing aids on the commercial market is very large. The wide range of hearing aids can be grouped into four general categories.

(1) POCKET TYPE

The pocket instrument (Fig. 115) comprises a small case called the transmitter, an electrical connection cord and a receiver. The transmitter houses the microphone, amplifiers and all controls, and is connected to the output transducer by means of a cord, which may be 12 to 36 or more inches in length. The capacity for producing acoustic power is greater with the pocket aid than with any other type, so these are often utilized in cases of extremely serve or profound hearing loss.

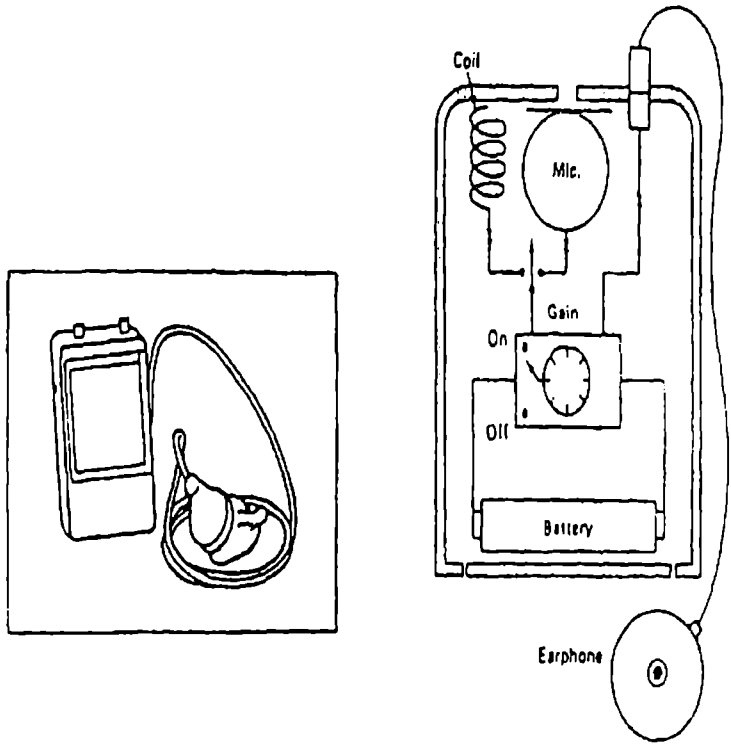


Figure 115: Pocket type hearing aid (a) External view, and (b) functional diagram.

(2) POSTAURICLE TYPE

The postauricle or behind the ear hearing aid (Figure 116) is by far the prevalent style in use today. All components are contained in a case designed to fit behind the ear, in the niche formed by the attachment of the pinna to the temporal area of the skull. A flexible, hollow tube conducts the amplified sound from the instrument to the patient's ear and is usually, but not always, retained in the ear by a molded earliece. Because the input transducer is located in the vicinity of the ear, and not at chest level, as it is with a pocket instrument, postauricle, as well as eyeglass and all-in-the-ear aids are often referred to collectively as ear-level instruments.

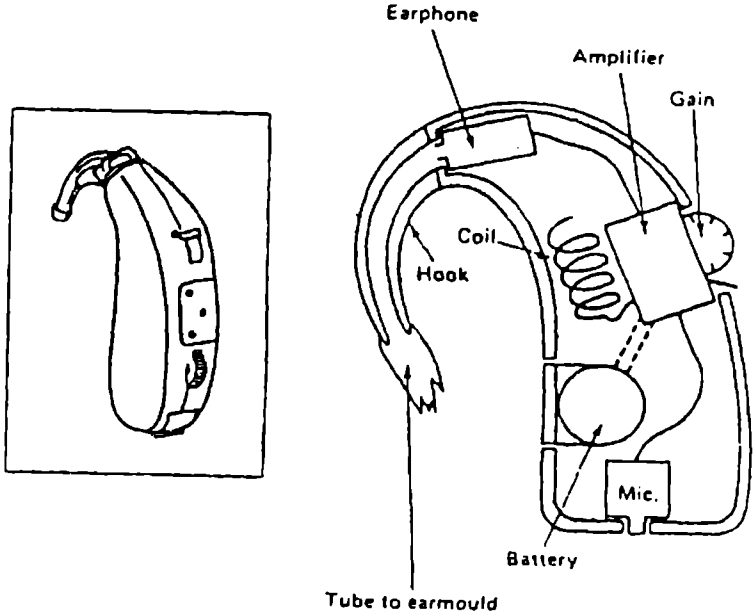


Figure 116: Postauricle type hearing aid (a) External view, and (b) functional diagram.

The postauricle hearing aid can provide more amplification than smaller devices due to the stronger amplifier and the larger battery. It is available in several colours for hair and skin tone matching.

(3) EYEGLASS TYPE

Eyeglass hearing instruments (Fig. 117) are electronically identical to their postauricle counterparts, but the components are enclosed within a case designed in the form of an eyeglass temple. Means are provided for attaching the temple to most optical front frames and for accommodating variations in the required ear.

The person who wears glasses only on a part time basis is not a candidate for eyeglass hearing aid. Even when glasses are worn full time, the eyeglass hearing aid may not be satisfactory because the physical connection of the eyeglass to the hearing aid can present problems.

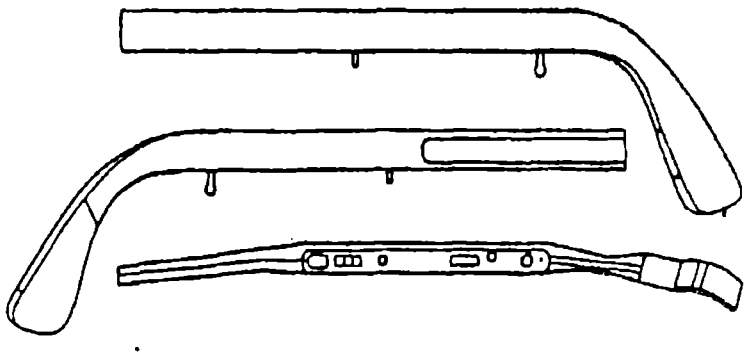


Figure 117: Eyeglass hearing aids.

(4) ALL-IN-THE EAR TYPE

Miniaturization in the construction of components has made it possible to enclose all necessary elements within a hollow earpiece molded to fit the patients ear (Fig. 118). This type of unit ordinarily fills the canal of the external ear and, aside from its cosmetic desirability, sometimes provides the only possible solution in patients with inadequate space behind the ear for a postauricle or eyeglass instrument.

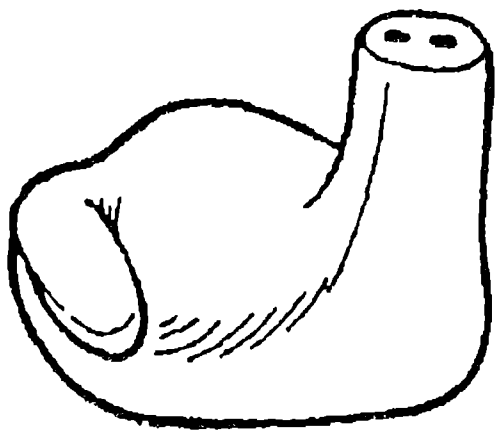


Figure 118: All-in-the-ear hearing aid.

These hearing aids can be used for a wide range of hearing losses. Due to their size, they can allow larger sound amplifiers and more features such as a telephone switch. They are also much easier to handle.

(5) IN-THE-CANAL HEARING AID

A bit larger hearing aid also fit far into the ear canal. Canal hearing aids use a slightly larger battery. This aid is used for mild to moderate hearing loss.

(6) COMPLETELY-IN-THE-CANAL HEARING AID

They are the smallest hearing aid types and are almost invisible in the ear. All components are housed in a small case that fits far into the ear canal. This takes advantage of the ear's own natural sound collecting design and offers convenient telephone usage. They are restricted to persons with ear canal large enough to accommodate the insertion depth of the instrument into the ear. This type of hearing aid is not suitable for persons with severe hearing loss.

(7) DIGITAL HEARING AIDS

They represent the most advanced technology available today. These instruments contain a computer chip that is programmed by a computer. This offers the best way currently available to match a particular hearing loss with the most prescriptive amplification needed. Complete flexibility and amazing fine tuning capability are among the many benefits. This unique technology separates the incoming sound into bands and channels; and processes each band and channel independently. This class of hearing instrument utilizes separate circuit paths to independent processes in different frequency regions of sound. This is the most advanced technology that allows the most precise prescriptive fitting available today.

Chapter 7

MEMBRANE CONDUCTIVITY

The significance of biomembranes is being increasingly felt these days. Not only do they delineate the protoplast but also compartmentalise areas of different metabolic activities and regulate the influx of nutrients and the efflux of waste products to and from the cells as well as the organelles. In addition to these, membranes form the sites of a number of reactions.

The plasma membrane have a number of important functions which are based on the principles of physics and physical chemistry, to the analysis of the life processes at a quantitative and mechanistic level. Some of these functions are: (1) Diffusion of molecules from a higher concentration to a lower concentration; (2) Transport of ions at the expanse of free energy liberated in chemical reactions; (3) Osmosis of solvent molecules where its activity is high to another region where its activity is low, through a semi-permeable membrane; (4) Absorption and uptake of substances; (5) Biosorption of substances, metals and radionuclide particulates; (6) Transpiration or loss of water and other substances; (7) Ultrafiltration: and (8) Electrical properties of plasma membranes.

Each of these function is being elaborated in the following description to explain their mechanism and significance.

DIFFUSION

When the particles, ions, or molecules of a solute are in the body fluids (solvents), both solute and solvent show rapid random movement and frequent collisions occur between the molecules,

ions, or particles. Each particle moves in its own separate way and the greater the motion, the higher is the temperature and it never ceases except at absolute zero temperature.

Diffusion refers to the process by which molecules show net movement from an area of its own high concentration to another area of its low concentration. Diffusion involves the process by which molecules intermingle as a result of their kinetic energy of random motion. Consider two containers of gas A and B separated by a partition.

The molecules of both gases are in constant collisions with the partition. If the partition is removed as in the lower illustration, the gases will mix because of the random velocities of their molecules, in time a uniform mixture of A and B molecules will be produced in the container.

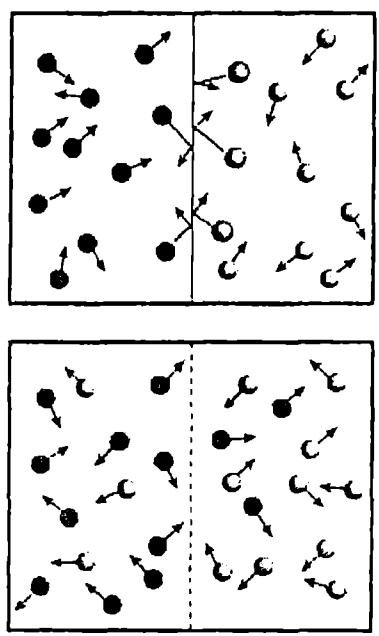


Figure 119: Dual gas container (A) with partition; and (B) without partition.

The flow of energy or matter from a higher concentration to a lower concentration, result in a homogeneous distribution. If one end of a rod in heated or electrically charged, the heat or electricity will diffuse from the hot or charged portion to the cool or uncharged portion. If the bar is made of metal, this diffusion will be rapid for heat and almost instantaneous for electricity; if the bar is made of asbestos, the diffusion will be slow for heat and extremely slow for electricity. Diffusion of matter occurs most rapidly in gases, more slowly in liquids, and most slowly in solids. The spreading of odoriferous molecules (a smell) throughout a room is a common example of gaseous diffusion. A solid may dissolve and diffuse through a liquid, as when a lump of sugar is placed in a cup of water. This process is much slower than the diffusion of a gas; if water is not stirred", it may take weeks for the solution to become homogeneous. An example of the slowest diffusion process, is a solid diffusing into a solid, for example when gold is plated on copper. The gold will diffuse slowly into the surface of the copper; however, diffusion of an appreciable amount of gold more than a microscopic distance normally requires thousands of years.

All these types of diffusion follow the same laws. In all cases, the rate of diffusion is proportional to the cross-sectional area and to the gradient of concentration, temperature, or charge. Thus, heat will travel four times as fast through a rod 2 cm in diameter as through a rod 1 cm in diameter, and when the temperature gradient is 10 ° per cm, heat will diffuse twice as fast as when the gradient is only 5° per cm. The rate of diffusion is also proportional to a specific property of the substance, which in the case of heat or electricity is called conductivity; in the case of matter, this property is called diffusivity or diffusion coefficient. The amount of material that diffuses in a certain time, or the distance it traverses, is proportional to the square root of the time; thus if it takes sugar one week to diffuse through water 1 cm from its starting point, it will take four weeks to diffuse through 2 cm.

As distinguished from stirring, which is a process of mixing masses of materials, diffusion is a molecular process, depending solely on the random motion of individual molecules. The rate

of diffusion of matter is therefore directly proportional to the average velocity of the molecules. In the case of gases, this average speed is greater for smaller molecules, in proportion to the square root of the molecular weight, and is greatly increased by rise in temperature. Metallic *thorium*, for example, diffuses rapidly through metallic *tungsten* at temperatures around 2000 °C, the operation of certain vacuum tubes is based on this diffusion.

If one molecule is four times as heavy as another, it will, in the case of gases, move half as fast and its rate of diffusion will be half as great. Advantage can be taken of this difference to separate substances of different molecular weights, and in particular to separate different isotopes of the same substance, if a gas containing two isotopes is forced through a fine porous barrier, the lighter isotopes, which have a higher average speed, will pass through the barrier faster than the heavier ones. The gas with the greater concentration of lighter isotopes is then diffused through a series of such barriers for large-scale separation. This technique, known as the gaseous-diffusion process, is widely used in the separation of the fissionable uranium (u-235) from the non-fissionable (u-238). In another isotope-separation technique, called the thermal-diffusion process, the separation depends upon thermal effects exhibited by some gases; if such gases are enclosed in a chamber subjected to temperature gradient, the heavier isotopes tend to concentrate in the cool region.

Diffusion results in a shift from a more ordered to a disordered state. Diffusion is therefore, spontaneous process driven by an increase in entropy and is thermodynamically favourable process. The free energy change during diffusion depends on the extent of the concentration gradient. As the concentration gradient decreases, the free energy difference also decreases until at equilibrium it becomes zero.

The net movement of molecules results due to the development of a pressure which may be defined as *diffusion pressure* (DP). The magnitude of the pressure is inversely proportional to the average distance between the molecules or directly proportional to the concentration of diffusible particles, that is, higher the concentration of the particles, the greater their diffusion pressure, it is also directly proportional to the absolute

temperature, that: is, the average energy of a particle in a homogeneous substance rises as the temperature increases but is constant for various substances at a given temperature. The particle velocities in gases can be easily calculated, but it is much more difficult to obtain values for liquids and solids.

Pure solvents show maximum diffusion pressure (that is zero). Whenever, solute is added, the chemical potential of the solvent decreases (becomes negative) and a *diffusion pressure deficit* (DPD) results. Hence, DPD of the solvent is proportional to the amount of the solute added to it.

Diffusion is of immense significance in plant-water relations, with evaporation, a diffusion process, being the overall driving force for most water movement through the plant. All living cells exist in an aqueous environment and are separated by a differentially permeable membrane which allows only certain substances in and out of the cell through the membrane, is a matter of diffusion. The phenomenon of diffusion is directly or indirectly involved in all physiological processes. All the aerial intake of different gases - Carbon dioxide and Oxygen from the atmosphere, as well as the movement of gases through intercellular spaces of the tissue takes place by this process. Absorption of water and minerals from the soil by the higher plants also takes place by diffusion phenomenon although a more complicated mechanism is involved in their absorption and accumulation by the root hairs. The loss of molecules from the plants and animals body is also a diffusional process. Thus, transpiration and loss of Carbon dioxide and Oxygen take place by diffusion.

PASSIVE DIFFUSION

Figure 120 shows passive diffusion. The dashed line is intended to indicate a membrane that is permeable to the molecules or ions as illustrated by dots, initially all of the dots are within the membrane. As time passes, there is net diffusion of the dots out of the membrane, following their concentration gradient. When the concentration of dots become same inside and outside of the membrane the net diffusion ceases. However, the dots still diffuse into and out of the membrane, but the rate of inward and outward diffusion are the same resulting in a net diffusion of zero.

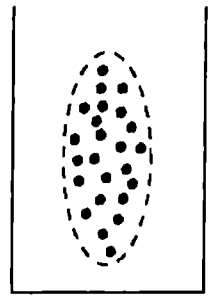


Figure 120: Illustration showing initial stage of passive diffusion.

FACILITATED DIFFUSION

Facilitated diffusion of ions takes place through proteins or assemblies of proteins, embedded in the plasma membrane. These trans-membrane proteins form a water-filled channel through which the ion can pass down its concentration gradient. The trans-membrane channels that permit facilitated diffusion can be opened or closed. They are said to be "gated". Some types of gated ion channels are:

- Ligand-gated
- Mechanically-gated
- Voltage-gated

Ligand-gated ion channels: Many ion channels open or close in response to binding a small signalling molecule or ligand (Figure 121). Some ion channels are gated by extracellular ligands; some by intracellular ligands. In both cases, the ligand is not the substance that is transported when the channel opens. External ligands bind to a site on the extracellular side of the channel. For example, Acetylcholine (ACh). The binding of neurotransmitter acetylcholine at certain synopses open channels that admit Na^+ and initiate a *nerve impulse* or *muscle contraction*. Similarly, binding of gamma amino butyric acid (GABA) at certain synopsis in *the* central nervous system admits Cl^- ions into the cell and inhibits the creation of a nerve impulse.

Internal ligands bind to a site on the channel protein exposed to *cytosol*. Second messengers like cyclic AMP (cAMP) and cyclic GMP (cGMP) regulate channels involved in the initiation of

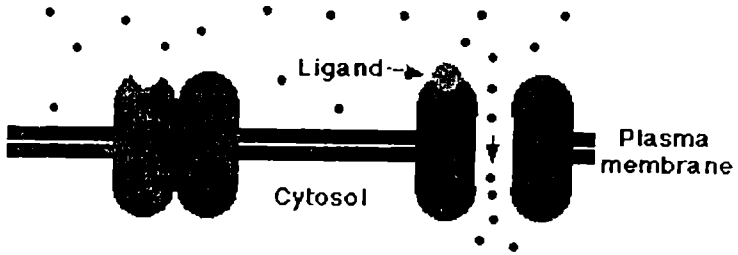


Figure 121: Facilitated diffusion through a ligand-gated channel

impulses in neurons involved in the initiation of impulses in neurons responding to odours and light respectively. Similarly, ATP is needed to open the channel that allows chloride (Cl^-) and bicarbonate (HCO^-_3) ions out of the cell. This channel is defective in patients with *cystic fibrosis*. Although the energy liberated by the hydrolysis of ATP is needed to open the channel, this is not an example of active transport; the *ions* diffuse through the open channel following their concentration gradient.

Mechanically-gated ion channels: Some channels open or close in response to mechanical stimuli. For example, the sound waves bend the cilia-like projections on the hair cells of the inner ear, and open up ion channels leading to the creation of nerve impulses that the brain interprets as sound. Similarly, mechanical deformation of the cells of stretch receptors open ion channels leading to the creation of nerve impulse.

Voltage-gated ion channels: In excitable cells like neurons and muscle cells, some channels open or close in response to changes in the charge across the plasma membrane. For example, as an impulse passes down a neuron, the reduction in the voltage opens sodium channels in the adjacent portion of the membrane. This allows the influx of Na^+ into the neuron and thus the continuation of the nerve impulse. Some 7000 sodium ions pass through each channel during the brief period (about 1 millisecond) that it remains open.

FACTORS AFFECTING DIFFUSION

(1) Molecular size

If the molecular size is greater, the rapidity with which a molecule diffuses from one side to another is less, because the large particles are not forced so intensely by collisions with other molecules. Diffusion is inversely related to the diameter of diffusing molecules or ions.

(2) Molecular weight

The rate of diffusion is approximately inversely proportional to the square root of the molecular weight. The less the square root of the molecular weight, the greater is the rate of diffusion.

(3) Shape of the molecule

Elongated molecules encounter higher frictional resistance in a solution and consequently has a low rate of diffusion than a spherical molecule having same molecular weight.

(4) Concentration difference

The greater the concentration difference, the greater is the rate of diffusion.

(5) Diffusion distance

The rate of diffusion is inversely proportional to the distance over which it has to take place. It indicates that the thicker the tissue the slower is the rate at which oxygen reaches the cell farthest from the surface. In small organisms, the diffusion distance is short. In large organisms, the problem has been overcome by development of a circulatory system, or by reducing the diffusion distance.

(6) Cross-section of the container

The greater the cross-section of the container in which diffusion takes place, the greater is the rate of diffusion.

(7) Temperature

The greater the temperature, the greater is the molecular motion and also, the greater is the rate of diffusion.

(8) Viscosity of solvents

The more viscous is the solvent, the greater is its frictional resistance against free movement of solute, hence the lower is the rate of diffusion and mean free path.

(9) Mean molecular velocity

In case of solutes and also for gases, diffusion varies directly with the mean molecular velocity. All the molecules do not move at the same velocity, they may possess an average velocity called the mean molecular velocity and mean free path.

DIFFUSION ACROSS CELL MEMBRANES

Diffusion occurs when a membrane separates two fluids differing in concentration at the same hydrostatic pressure. When the membrane is freely permeable to all the solutes, then all the solutes are freely diffusible through the membrane and an equilibrium is reached with equal concentrations on each side and concentration gradient disappears: If one solution contains substance which is not diffusible to equilibrium may result in an unequal distribution of freely-diffusible solutes producing Donnan membrane equilibrium of ions.

In living animals, beneath the membrane, there is a layer of relatively unstirred water of 100-400 μm thick. Solutes pass through this unstirred water layer by diffusion (Figure. 122)

BIOLOGICAL MEMBRANES

Biological membrane have been studied in different ways: for example, by looking at them with high powered microscopes or by observing their various transport or signalling properties with

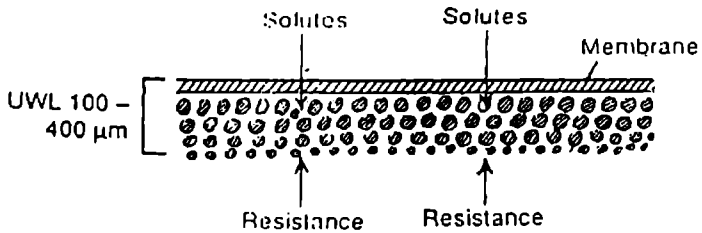


Figure 122: Unstirred water layer (UWL) beneath the membrane.

electrical instruments and enzymatic assays. More basically, we begin our study of membranes structure and function by examining how membranes are constructed? What is their composition?

The *fluid-mosaic model* of biological membranes was proposed by Singer and Nicolson (1972) - views the membrane as a fluid bilayer of amphipathic complex lipids with proteins embedded in and/or spanning the bilayer (Fig. 123). The model is based on the thermodynamic hypothesis that the membrane structure is dominated by the hydrophobic effect. The lipid bilayer is permeable to polar and large non-polar molecules. The lipids provide the fluid matrix in which the integral membrane proteins are embedded. The proteins can either be localised at an interface or traverse the membrane, being exposed at both interfaces. They provide the catalysts for the transport and transduction reactions. The alkyl chains of the polar lipids are sequestered into the hydrophobic region of the bilayer away from water.

The plasma-membrane is a lipid matrix in which amphipathic lipids (phospholipids, glycolipids and sterols are arranged in a bilayer to form a fluid, liquid-crystalline matrix or core). Globular proteins are embedded in the matrix and form a mosaic on the matrix. They may be partially embedded on either side or may completely penetrate the membrane. This mosaic is not fixed or static, because the proteins are free to diffuse laterally in two dimensions. Water and dissolved substances pass through the protein part of the membrane freely, but the lipid layer is the limiting boundary for the fluids between the two sides of the membrane.

The bilayer structure maximises both the hydrophobic interactions of the alkyl chains and the hydrophilic attraction of the polar groups with the aqueous medium, and thereby provides a thermodynamically stable structure. The absence of strong attractive forces between the adjacent alkyl chains keep the bilayer structure fluid and deformable and the hydrocarbon tails have considerable conformational mobility. These properties are influenced by the degree of unsaturation of the hydrocarbon tails. The range of fluidity in membranes is determined by the proportions of saturated and unsaturated fattyacids. About 50 percent of the fatty acids isolated from all sources are unsaturated. In bacteria,

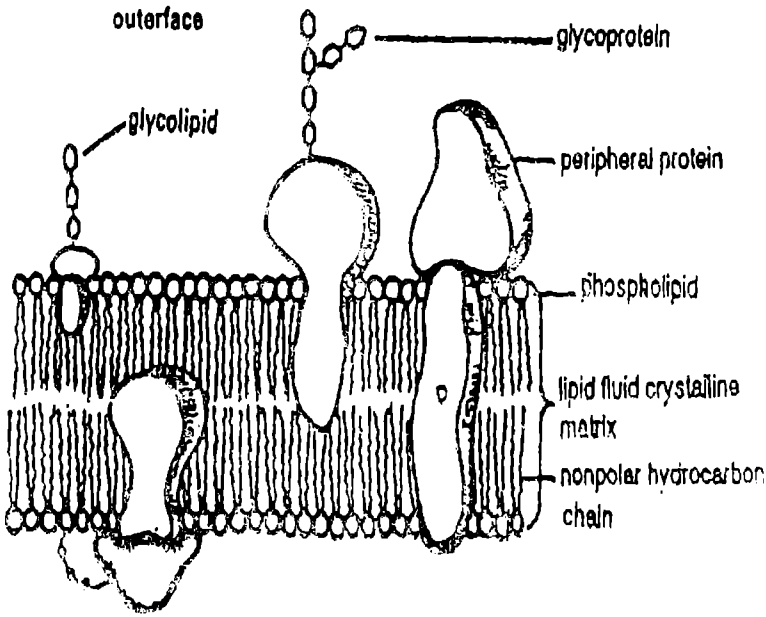


Figure 123: Fluid Mosaic model of membrane; with partly embedded protein. and folly embedded protein

the percentage of unsaturated fatty acids in membranes is inversely proportional to the temperature of the growth medium or environment. The methyl side chains, when present in fatty acids, lower the melting point of the lipid bilayer.

The polar head groups in the lipid bilayer may get attracted electrostatically towards each other especially if they are dipolar. Within the membrane, the phospholipids are distributed asymmetrically, between the inner and outer halves of the bilayer. This asymmetry is due to steric considerations as well as due to electrostatic effects. Since the membrane is often highly curved, there is lesser room in the inner half than in the outer half. The similar ionic heads, when brought very close, repel one another more on the inner side of the membrane than on the outer side.

The erythrocyte membrane is composed of two major structural units, the membrane bilayer and the underlying meshwork of peripheral membrane proteins, called the membrane associated cytoskeleton or membrane skeleton. The bilayer has little structural strength and fragments readily by vesiculation. The lipid bilayer contains phospholipids and cholesterol as the major lipid constituents in almost equimolar amounts. The amount of choline containing phospholipids is greater than the aminophospholipids.

An increase in the level of unsaturated fatty acids in membrane lipids is also seen in many plants. This is interpreted as an adaptation to resist frost hardening and freezing injury.

Sterols of all forms (sterol, sterol glycoside, acylated steryl glycoside) are abundant in membranes especially in the endoplasmic reticulum. In mitochondrial membranes the outer membrane contains about three times as much sterols than in the inner membrane. The steryl and the acylated moieties are embedded in the hydrophobic phase while the glucose and other polar head groups joins the hydrophobic phase. Steroid skeleton favours the extended chain conformations of adjacent hydrocarbon chains of the fatty acids which help in their close packing and produce a more crystalline array of molecules within the membrane.

The compartmentation of the metabolic activity, which is essential for the metabolic regulation, is possible because of the plasmamembranes present in the cells as well as in the organelles. These membrane partition the metabolites among the organells and prevent the loss of dissolved materials from these compartments. They also regulate the influx of nutrients and the efflux of waste products. Materials move across the membranes either by diffusion or by active transport.

The plashmamembranes from quite different cells have similar compositions Table 13. These outer envelops consist, mainly of protein and lipids, with very little carbohydrate is present as covalently linked to either protein or lipid, making these latter molecules, respectively, glycoprotein or glycolipid. Basically, then the plasmamembranes of most eucaryotic cells contain equal weights of lipid and protein (with and without attached carbohydrate), but there are many more lipid than protein molecules in these membranes.

Table 13: Amounts of protein, lipid and carbohydrate (Approx. % of dry weight) in different biological membranes (Guidotti, 1972)

Membranes	*Protein*	*Lipid*	*Carbohydrate*
Plasma membrane:			
Red blood cells	49	43	8
Liver cells	54	36	10
Amoeba	54	42	4
Myelin	18	79	3
Nuclear envelope	66	32	2
Endoplasmic reticulum	62	27	10
Golgi complex	64	26	10
Mitochondrion			
Outer membrane	55	45	trace
Inner membrane	78	22	-
Chloroplast inner membrane	70	30	-

A weight fraction of lipid contains many more molecules than an equal weight fraction of protein simply because lipids are much smaller than proteins. At the other extreme, inner mitochondrial membranes (and the plasma membranes of many bacteria) are relatively rich in protein, containing three times more protein than lipids on a weight basis.

The rate at which the substance diffuse inwards is proportional to the concentration of molecules present outside (Figure 124).

The concentration determines the number of molecules knocking the outer part of the pore each second. The rate at which the molecules diffuse outward is proportional to their concentration inside the cell membrane. Therefore, the fate of net diffusion into the cell is proportional to the concentration on the outside (Co) minus the concentration on the inside (Ci):

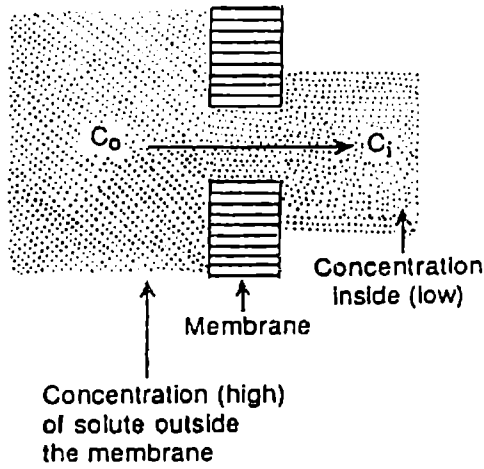


Figure 124: Effect of concentration difference on diffusion of molecules and ions across a membrane.

Net diffusion = P (Co-Ci)

where Co = Concentration on the outside

Ci = Concentration on the inside

P = Permeability of the membrane for the substance

When an electrical potential difference is applied across the membrane, the ions for their electrical charges move through the membrane though there is no concentration difference between two sides of the membrane (Figure 125).

In A, the negative ions concentration are same on both sides but when a negative charges is applied to the left and positive charge to the right, an electrical gradient is developed across the membrane. The positive charge attracts the negative ions while negative charge repels them. So the net diffusion occurs from left to right. In B due to concentration of large amount of negative ions on the right side, a considerable concentration difference of the same ions, they start moving in the direction opposite to the electrical potential difference.

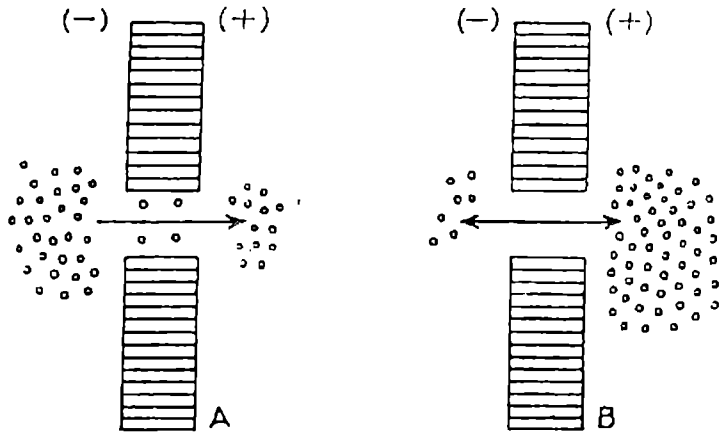


Figure 125: Effect of electrical potential difference on diffusion of molecules and ions across the cell membrane.

In addition to diffusion through cell membranes by simple passive transport, some substances which are very insoluble in lipids (sugar, amino acid) can pass through lipid matrix very rapidly by *Carrier-mediated diffusion* or facilitated diffusion. At first glucose (or amino acid) combines with specific carrier at the outer border of the cell membrane, to form sugar carrier-complex. The sugar carrier complex transfers glucose to the outer border of the membrane, where it is dissociated to deliver the sugar into the fluid on that side and the carrier molecule moves back to the original site again, to pick up still more sugar and transport it to the inside and act as ferry-boat (Figure 126)

Co-transport of glucose and sodium takes place in the membrane which is not active in the absence of sodium. When sodium is present, the carrier is activated by allosteric modification and its affinity for glucose also increase. It is the sodium gradient across the membrane that provides energy to promote the transport of sodium and glucose. Since energy is provided from sodium gradient glucose can be transported against a concentration gradient. Thus, if there is low concentration of glucose in the human intestine or renal tubules it will be transported to the interior of

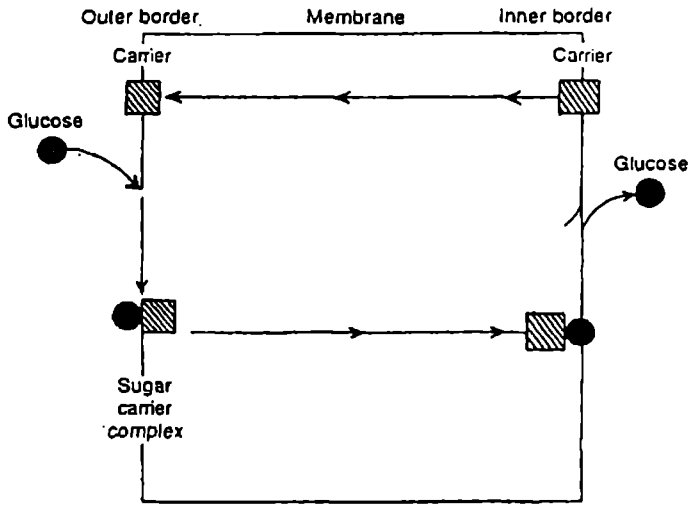


Figure 126: Carrier model for sugar transport across cell membrane (Facilitated diffusion).

the cell where glucose concentration is high. At the inner side of the cytoplasmic membrane the sodium ion concentration is kept low by Na^+/K^+ exchange pump (Fig. 127).

Sodium leaves the sodium-glucose carrier complex and residues the affinity of the carrier for glucose, glucose is released from the carrier to the cytoplasm. There are different types of sodium Co-transport carriers. The nature of carrier system and chemical reactions that takes place for sodium co-transport is not known.

ACTIVE TRANSPORT

The concentration of a substance is high in the extra-cellular fluid but it is present in low concentration in intra-cellular fluid, such as sodium. Where as some substance is present in low concentration in the extra-cellular fluid, but they are present in high concentration in the intra-cellular fluid such as potassium. Diffusion takes place through a cell membrane from a high concentration region to low concentration region, but reverse process also occurs in the system.

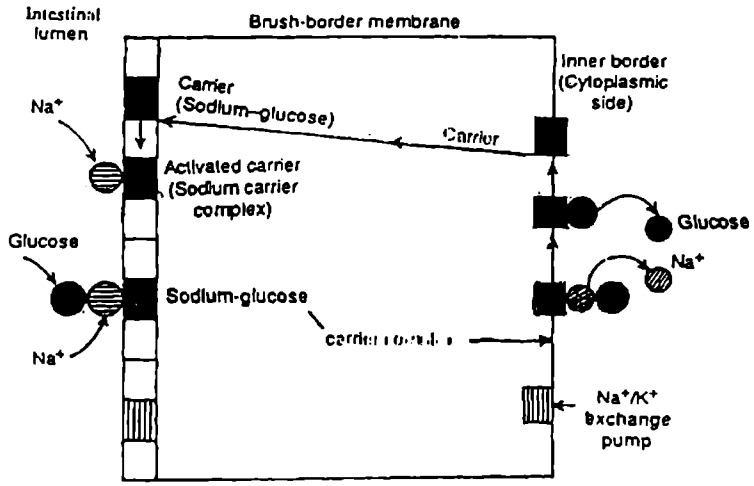


Figure 127: Model for co-transport of sodium and glucose.

When the movement of the substance takes place through cell membrane from low concentration region to higher concentration region, that is, against concentration gradient and energy is required for the molecule, the process is called *"Active transport"*.

The mechanism of active transport is similar *to facilitated diffusion* but having difference, that is: (a) active transport works against concentration gradient but not in the case of facilitated diffusion, and (b) active transport requires metabolic energy whereas facilitated diffusion does not require energy. Some sequence of events in active transport (Figure 128) are as follows-

The most important substance, Na+, is actively transported in this manner. When Na+ ions are in the lateral space, their positive electrostatic charges pull negatively charged Cl- ions to follow electroneutrality. When concentration of both ions increase in the lateral space, which causes osmosis of water out of cell and into the lateral space. Na+ and Cl" flow along with water to the connective tissue where and ions diffuse into the capillaries.

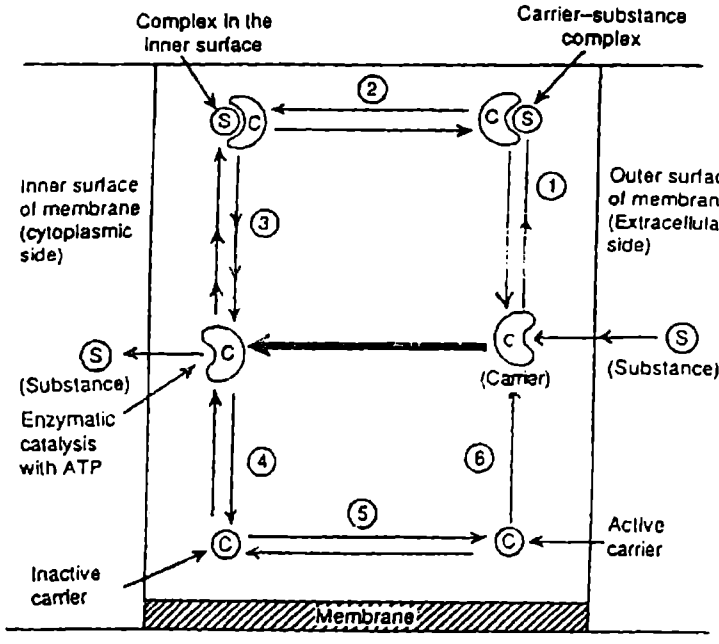


Figure 128: Sequence of events in active transport.

Glucose and aminoacids are transported through the epithelial cells of intestine or renal tubules by a mechanism that is neither pure diffusion nor active transport but a mixture of the so called Secondary active transport or *sodium co-transport.*

Sodium concentration is very high and potassium is low outside the cell and reverse concentration is present inside the cell. Some minute quantities of Na^+ and K^+ can diffuse through the pores of the cell membrane and if it is allowed to continue there will be equal distribution between the two sides of the membrane. This distribution is maintained by the mechanism of active transport - *sodium-potassium* pump (Fig. 129).

(1) First, the energy is supplied to the inner surface of the cell membrane from high energy substances for example ATP from cytoplasm.

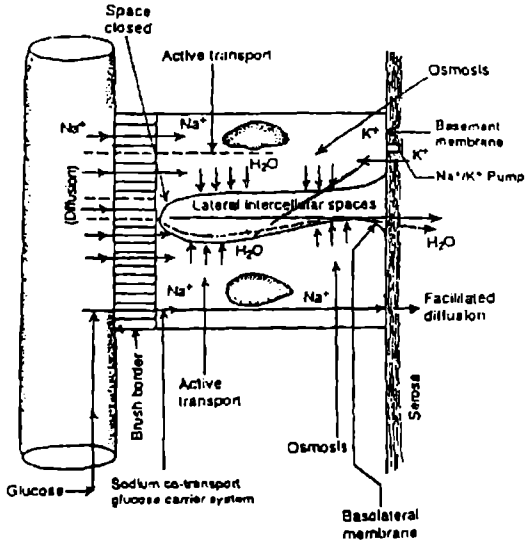


Figure 129: Active transport through a layer of cells or cellular sheets, e.g. intestinal epithelium.

(2) Second, active transport obeys the usual laws for chemical combination of one substance (which is to be transported) with another (carrier).

(3) Third, a specific carrier molecule is required to transport each type or each class of substance.

(4) Fourth, a specific enzyme (or enzymes) is required to promote active transport. The enzyme catalyses the reaction utilizing energy from ATP to split the substance away from the carrier. But the released substance being insoluble in the membrane, cannot diffuse backward through the lipid matrix of the membrane. Hence, the substance is discharged to the cytoplasm. The carrier is enzymatically degraded into an inactive form. The inactive carrier diffuses back to the outer side for transportation of another molecule. The carrier is activated by an energy-requiring process. Carrier for this mechanism is called *Sodium Potassium A·TPase* which can also split ATP to provide energy. This ATPase is composed of two protein molecules, globulin and glycoprotein

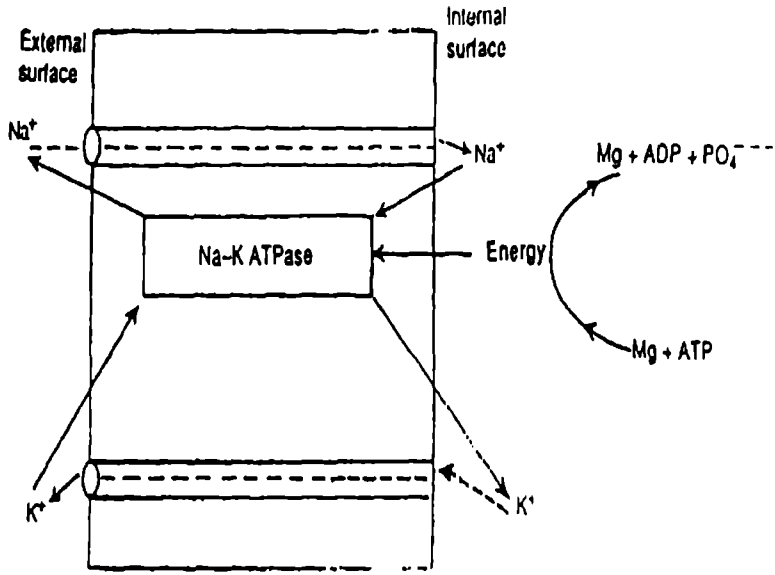


Fig. 130: Active transport of An and K+ (Sodium-Potassium pump).

The active transport of sodium-potassium ion pump operates asymmetrically or vectorially. Sodium ion are pumped out of the cell while the potassium ions are pumped into the cell against concentration gradient. The intra-cellular concentration of Na+ and K+ are about 10 mM and 100 mM respectively, in the ratio of 1:10. On the contrary the extracellular fluids have about 140 mM and 5 mM K+ in a ratio of 28:1. The energy required for the functioning of the pump comes from the hydrolysis of ATP and a phosphorylated protein is identified as an intermediate in the process. The hydrolysis of ATP is coupled to phosphorylation of the transport protein. The hydrolysis of phosphoprotein is believed to cause a conformational change that opens a pore, then drives the transport of Na+ and K+. The phosphorylation of protein might alter the conformation of the Na+ and K+ binding sites hi the protein. It has been calculated that the free energy change of 9.3 KJ mol/l (the energy released by the hydrolysis of the phosphoprotein) will drive a concentration gradient of 50:1

uphill. For each ATP hydrolysed, three sodium ions are pumped out to pump in two potassium ions. The pump is a reciprocal one, so that transport of one type of ion cannot occur without the simultaneous movement of the other species in the opposite direction. Unlike the protein mediated facilitated diffusion system (which will carry the substance equally well in both directions), the active transport system can drive the specific ion to one direction only.

OSMOSIS

Osmosis is a special type of diffusion in which solution and semi permeable membrane are involved. The flow of one constituent of a solution through a semi-permeable while the other constituent are bloated and unable to pass through the membrane constitutes osmosis. If two solutions of different concentrations are separated by a semi-permeable membrane which is permeable to the smaller solvent molecules but not to the larger solute molecules, then the solvent will tend to diffuse across the membrane from less concentration to the more concentrated solution. This process is known as osmosis (Figure 131).

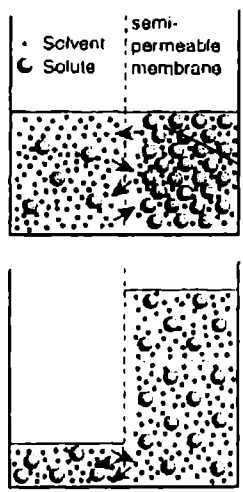


Figure 131: Osmosis through semipermeable membrane.

Experimentation is necessary to determine which membranes permit selective flow, or osmosis, because not all membranes act in this way. Many membranes allow all or non of the constituents of a solution to pass through; only a few allow a selective flow. In the classic demonstration of osmosis, a vertical tube containing a solution, with its lower end closed off by a semi-permeable membrane, is placed in a container of water. As the water passes through the membrane into the tube, the level of sugar solution in the tube rises visibly. A semipermeable membrane that may be used for such a demonstration is the membrane found just inside the shell of an egg, this is, the film that keeps the white of the egg from direct contact with the shell. In this demonstration, the water moves in both directions through the membrane, the flow is greater from the vessel of pure water, however, because the concentration of water is greater there, that is, fewer dissolved substances exist in this solution than in the sugar solution. The level of liquid in the tube of sugar solution will eventually rise.

Large quantities of water molecules constantly move across cell membranes by simple diffusion, but, in general, net movement of water into or out of cells is negligible. For example, it has been estimated that an amount of water equivalent to roughly 250 times the volume of the cell diffuses across the red blood cell membrane every second, the cell does not lose or gain water because equal amounts do in and out.

There are, however, many cases in which net flow of water occurs across cell membranes and sheet of cells. An example of great importance is the secretion and absorption of water in small intestine, in such situations, water still moves across membranes by simple diffusion, but the process is important enough to warrant a distinct name - osmosis.

Osmosis is the net movement of water across a selectively permeable membrane driven by a difference in solute concentrations on the two sides of the membrane. A selectively permeable membrane is one that allows unrestricted passage of water, but not solute molecules or ions.

Different concentrations of solute molecules leads to different concentrations of free water molecules on either side of the

membrane, on the side of the membrane with higher free water concentration (that is, a lower concentration of solute), more water molecules will strike the pores in the membrane in a given interval of time. More strike equates to more molecules passing through the pores, which in turn results in net diffusion of water from the compartment with high concentration of free water to that with low concentration of free water.

The key to remember about osmosis is that water flows from the solution with the lower solute concentration into the solution with higher solute concentration. This means that water flows in response to differences in molarity across a membrane. The size of the solute particles does not influence osmosis. Equilibrium is reached once sufficient water has moved to equalize the solute concentration on both sides of the membrane, and at that point, net flow of water ceases. For example, two containers of equal volume are separated by a membrane that allows free passage of water, but totally restricts passage of solute molecules. Solution A has 3 molecules of the protein albumin (molecular weight 66,000) and solution B contains 15 molecules of glucose (molecular weight (80). Water will flow from solution A to B.

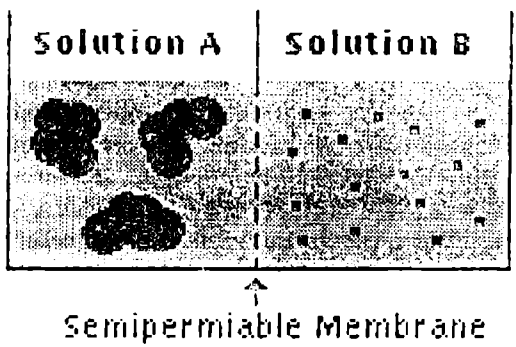


Figure 132: The phenomenon of osmosis.

When thinking about osmosis, we are always comparing solute concentrations between two solutions, and some standard terminology is commonly used to describe these differences (Figure 133): *isotonic,* the solutions being compared with equal concentration of solutes; *Hypertonic,* the solution with the higher concentration

of solutes; and *Hypo tonic,* the solution with the lower concentration of solutes, in the above example solution A and B are isotonic (with each other), solution A and B are both hypertonic as compared to solution C, and solution C is hypertonic relative to solution A and B.

Figure 133: Comparison of solute concentration between two solutions.

OSMOSIS IN ANIMAL CELLS

The classic demonstration of osmosis is to immerse red blood cells in solutions of varying osmolarity. Blood serum is isotonic with respect to the cytoplasm, and red cells in that solution assume the shape of a biconvex disk. When the red cells were diluted in the serum, the beautiful biconcave shape of the cells as they circulate in blood were noticed. Then the cells in serum were diluted with water, at 200 milliosmols (mos) the cells become visibly swollen and have lost their biconcave shape, and at 100 mos, most of the cells swell so much that they become ruptured, leaving what are called red blood cell ghosts. This is because in a hypotonic solution, water rushes into red blood cells; However, when a concentrated sodium chloride solution was mixed with the cells and serum to increase osmolarity. At 400 mos and especially at 500 mos, water flows out of the cells, causing them to collapse and assume the spiky appearance (Figure 134)

OSMOTIC PRESSURE

Osmosis is the selective diffusion process driven by the *internal energy* of the solvent molecules, it is convenient to express the available energy per unit volume in terms of "osmotic pressure", it is customary to express this tendency toward solvent transport in pressure units relative to the pure solvent,

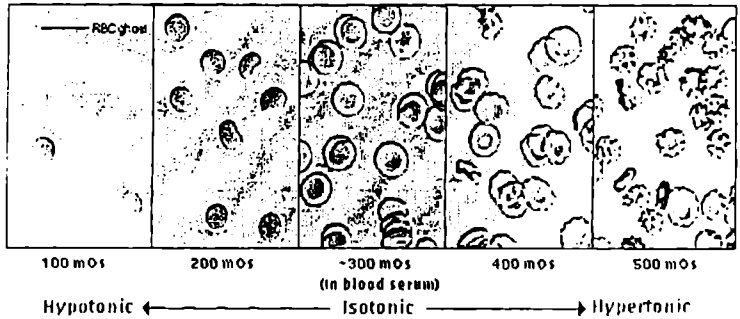


Figure 134: Sequence of events when red blood cells are placed in isotonic, hypotonic and hypertonic solutions.

If pure water on both sides of the membrane were there, the osmotic pressure difference would be zero. But if normal human blood were on the right side of the membrane, the osmotic pressure would be about seven atmospheres (Figure 135)

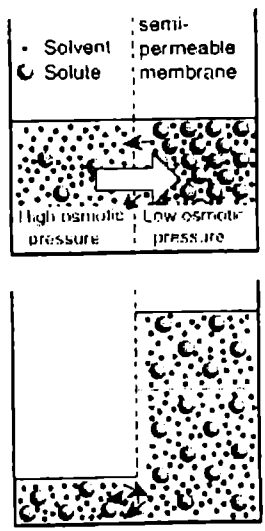


Figure 135: Diffusion from high osmotic pressure to that of low osmotic pressure.

The decision about which side of the membrane to call "high" osmotic pressure is a troublesome one. The choice made here is the opposite of that made in many biology texts, which attribute "high" osmotic pressure to the solution and zero osmotic pressure to pure water. The rational for the choice is that the energy which drives the fluid transfer is the *thermal energy* of the water molecules, and that energy density is higher in the pure solvent since there are more water molecules, the thermal energy of the solute molecules does not contribute to transport, presuming that the membrane is impermeable to them. The choice is also influenced by the observed direction of fluid movement, since under this choice the fluid transport is from high "pressure" to low, congruent with normal fluid flow through pipes from high pressure to low. The final rationale has to do with the measurement of osmotic pressure, if the pressure in the compartment into which water is flowing is raised to the equivalent of the osmotic pressure, movement of water will stop. This pressure is often called *hydrostatic* (water-stopping) *pressure* (Figure 136).

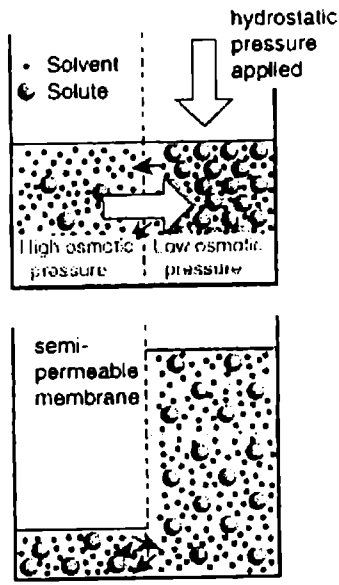


Figure 136: Demonstration of hydrostatic pressure.

The term *osmolarity* is used to describe the number of solute particles in a volume of fluid. Osmols are used to describe the concentration in terms of number of particles - a 1 osmolar solution contains 1 mole of osmotically active particles (molecules and ions) per litre.

The instrument used for the study of osmosis is called *osmometer.* The mouth of the thistle funnel is covered with a semipermeable membrane such as dead pig's bladder. The funnel is then filled with a strong, that is, concentrated sugar (sucrose) solution and immersed in a beaker of pure water. After some time, the level of the solution in the funnel starts to rise. Analysis of the solution of the funnel shows that it gradually becomes more dilute indicating that the water is passing into the funnel from the beaker (Fig. 137).

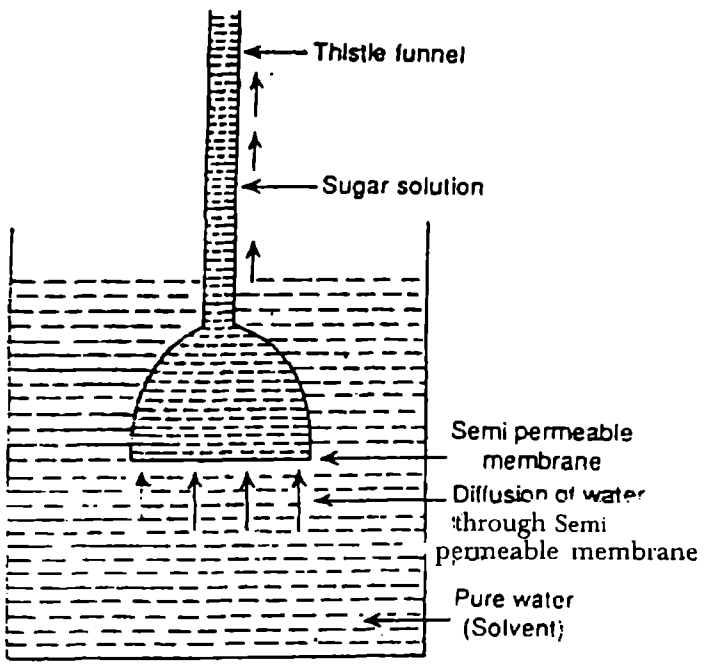


Figure 137: Simple osmometer (Abbe Nollet's Experiments)

This unequal diffusion through the membrane will continue, accompanied by a continuous rise of the volume of solution in the funnel until the pressure of the solution column in the funnel called *hydrostatic pressure* balances the excess pressure of the water entering the solution through the membrane over and above that of water leaving the solution. This hydrostatic pressure presses down the membrane and eventually stops further net movement of water molecules from the beaker as well as from the solution and an equilibrium is established. Since this hydrostatic pressure is produced by a process of osmosis, it is termed *osmotic pressure.*

This hydrostatic pressure varies with the concentration of the solution and becomes a measure of the osmotic pressure of said solution, since the greater the osmotic pressure of the solution, the higher must be the column of water necessary to overcome the excess pressure of the pure water against that of the less concentrated water of the solution. The concentration can be reversed if the solvent, that is, pure water, is kept inside the funnel and the solution in the beaker, the inequality of water pressure against the membrane would be reversed, hence more water would pass from funnel through the membrane to the beaker.

Osmotic pressure (π) of a solution is defined as the positive pressure exerted on the solution necessary to prevent a net flux of water between the solution and the pure solvent, when these are separated by a perfect semipermeable membrane. At the pore aperture of the membrane diffusion of water molecules into the solution will occur, due to the difference in chemical potential, which creates a localised decrease in pressure resulting in mass flow of water along the pore.

The term osmotic pressure is misleading. It is in fact an irrational statement. No isolated solution can possess an osmotic pressure since the phenomenon is only demonstrable in a system in which pure solvent and solution are separated by a semipermeable membrane. An unconfirmed solution has an osmotic pressure although no pressure in the literal sense is exerted. Since the osmotic pressure is the pressure that must be imposed upon the solution to maintain its solvent in equilibrium with pure solvent at the same temperature, it is confusing to refer to this pressure as if it were exhibited by the solution. There appears to be a need for anew and agreed term to denote *'solution potential'* or *osmotic potential* of a solution.

The pressure which a solution is potentially capable of developing if separated from pure water by a semipermeable membrane is referred to as its *osmotic potential.* The osmotic potential is directly related to the concentration of solute molecules in the solution.

If the applied pressure is higher than the osmotic pressure of the solution, solvent will flow from the solution through the membrane to the solvent side. This is called *reverse osmosis.* This process has been applied for obtaining pure water from sea water.
Mechanism of osmotic pressure

In the experiment described earlier, it is seen that water molecules can pass through the membrane in both directions, but sugar molecules cannot pass through it. The rate at which water molecule are allowed to pass through the membrane is determined by the concentration, or activity of water molecules which are in contact with the membrane. It is a fact that there are more water molecules per unit volume in pure water (solvent) than the water molecules present in the sugar solution. As a result, the diffusion pressure of water molecules from pure water through the membrane is higher than the diffusion pressure of water molecules from the sugar solution through the membrane. Thus, there is net movement of water from pure water to the sugar solution and this dilutes the solution. As a result, the level of sugar solution rises above the level of the pure water. The level of solution continues to rise until the hydrostatic pressure equalizes the diffusion pressure of the water molecules in the sugar solution and in the pure water (solvent), at which point water passes through the membrane in opposite directions at the same rate and the system is in equilibrium. In this process, the solute molecules decrease the diffusion pressure of the solvent in direct proportion to their concentration, or number and independently of the kind of molecules. The higher the osmotic pressure shown by a solution, the lower is the diffusion pressure of its solvent molecules. Thus, the osmotic pressure is the representation of the difference between the diffusion pressure of pure water (solvent) molecules and the diffusion pressure of solvent molecules present in the sugar solution. The more concentrated the sugar solution, the greater is the difference in diffusion pressure and the greater is the osmotic pressure.

The vapour pressure or diffusion pressure of the solvent molecules in the more dilute solution is greater than that of the solvent molecules in the more concentrated solution, so the net result is that the solvent molecules distil from the more concentrated solution until the concentration and vapour pressure of the two solutions are equal. The use of a semipermeable membrane in determining the osmotic pressure is the only means of estimating the diffusion of vapour of solvent molecules. When we insert a semipermeable membrane between the solutions, only water molecules can pass and there is the phenomenon of osmotic pressure. But when a membrane is permeable to sugar molecules and not the water molecules, we get an osmotic pressure due to diffusion of sugar molecules in it, a direction opposite to the osmotic pressure of water molecules. If there is a membrane which separates a solution from the solvent, and is readily permeable to the solvent and slowly permeable to the solute, a transitory rate of osmotic pressure, followed by a fall, is seen. When the membrane is equally readily permeable to both solute and solvent, no osmotic effect is seen, though osmotic pressure is there (Figure 138)

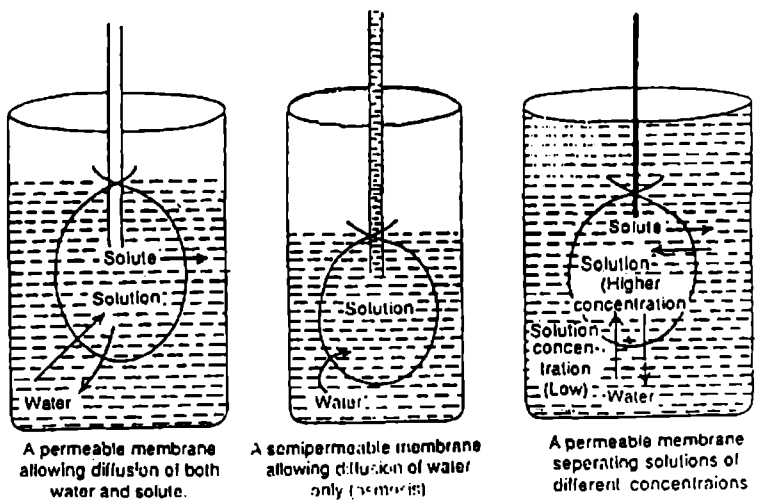


Figure 138: Diagrammatic representation of diffusion and osmosis across a membrane.

KINETICS OF OSMOTIC PRESSURE

The surface of a solution shows actual pressure at sea level which is equal to atmospheric pressure that is, 760 mm mercury. But the solutions between the two sides of the semipermeable membrane show wide difference of pressure of many thousand mm mercury, when the solution are different from each other. The capacity of the solution producing such pressure is called *potential pressure* of the molecules and ions of each solution. The potential pressure of each molecule and ion is proportional to its *chemical potential* in the solution. The amount of chemical potential of each type of molecule or ion is directly proportional to its fractional molar concentration in the solution.

The potential pressure of water molecules (solvent) only is greater than the potential pressure of water molecules in solution. This difference is equal to the hydrostatic pressure required to prevent osmosis through semipermeable membrane (Fig. 139).

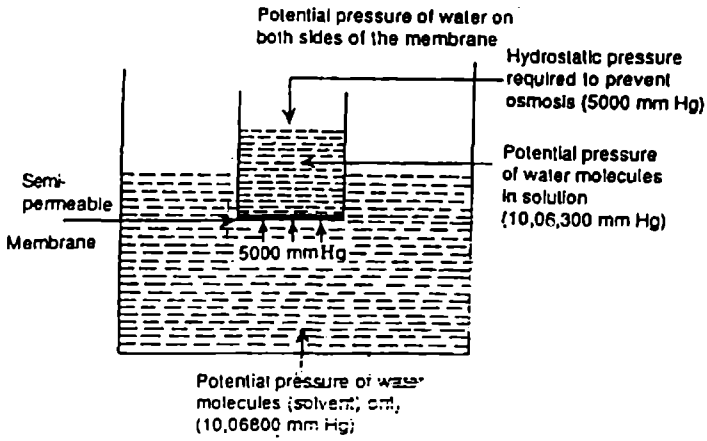


Figure 139: Potential pressure

OSMOSIS IN PLANT CELLS

Plant cells contains higher solute concentration than the surroundings and these solutes are present in the sap vacuole. The sap vacuole is surrounded by two semipermeable membranes, one on each side of the protoplast, but the cellulose wall is fully permeable.

When a partially turgid plant cells is placed in a solution having low osmotic pressure, that is, hypotonic solution, water enters into the sap vacuole by osmosis called *endosmosis* and the cell wall does not burst because the cellulose is fully stretched and resists the further stretching of the cell.

While the water enters the sap vacuole by osmosis, the protoplast presses against the cellulose wall and an internal pressure is developed called *turgor pressure* which is equal to the opposite pressure exerted known as *wall pressure*. When the turgor pressure reaches its optimum and the cell wall cannot stretch more, the cell becomes fully turgid (Figure 140)

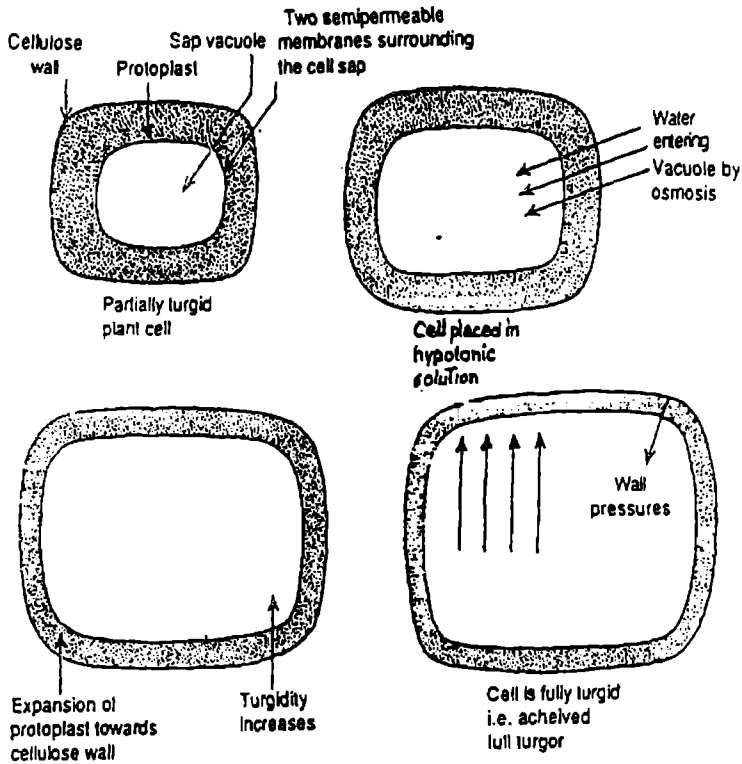


Figure 140: Sequence of events when partially turgid plant cell placed in hypotonic solution.

However, when a plant cell is placed in a hypertonic solution, that is, solution having high osmotic pressure, for example, strong sugar solution, water comes out with a gradual decrease in cell volume. Within a few minutes, the protoplast shrinks in such a way that it pulls away from the cellulose wall, leaving a gap between the wall and the outer membrane of the protoplast - *incipient plasmolysis.* Shrinkage of protoplast continues until the sap vacuole disappears and the protoplast looks like a round ball plasmolysis (figure 141)

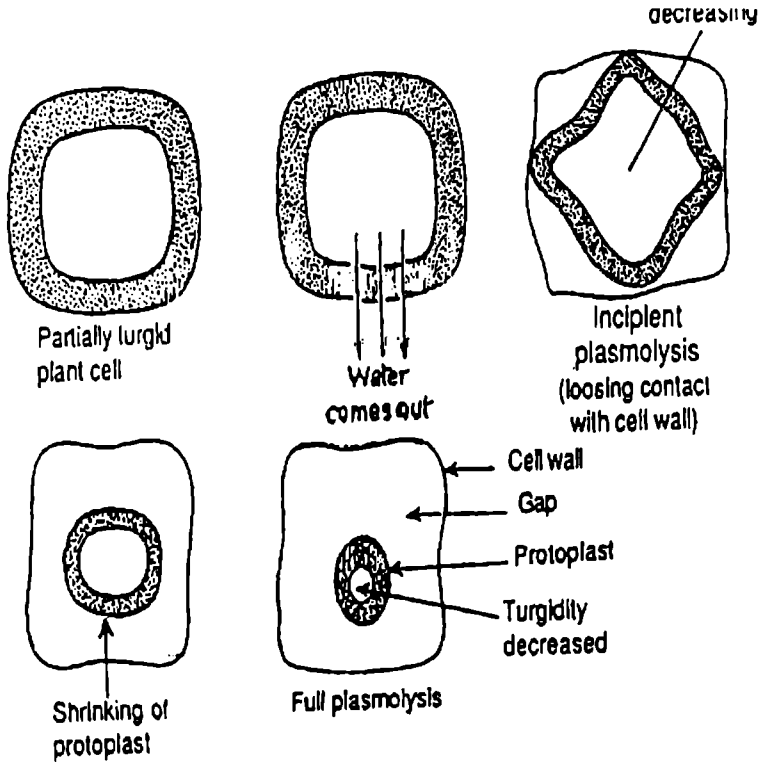


Figure 141: Sequence of events when partially turgid plant cell is placed in hypertonic solution.

DIFFUSION PRESSURE DEFICIT

When a plasmolysed cell is placed in pure water (hypotonic), water rushes into the sap vacuole and the protoplast starts expansion. The force at which water enters may be called *suction pressure* or *diffusion pressure deficit.*

Due to influx of water, the protoplast goes on expanding until it comes into contact with the cellulose wall and during this period the diffusion pressure deficit is opposed by the inward pressure of the cellulose wall, resisting the expansion of protoplast.

Diffusion pressure deficit (DPD)	=	Osmotic potential of the cell sap (OP)	–	Wall pressure or inward pressure exerted by cellulose wall (WP)

If the solute concentration of the cell and its surrounding medium, that is, the osmotic pressure in the cell and outside are the same, there will be no osmosis in either direction and the external solution is termed *isotonic solution.* A solution of 0.9 percent NaCl is isotonic to the human cell, 0.65 percent NaCl for amphibians, and is known as *normal saline,* which is used in medical and biological purposes.

ABSORPTION

It is the act of absorbing; the state or process of being absorbed; assimilation, incorporation of small forms into one big form; uptake of substances by a tissue; reception by molecular or chemical action; removal of energy or particles from a beam by the medium through which the beam propagates; and complete attention or preoccupations deep engrossment, in one's work.

Absorption has been variously defined by various authors. Some of these definitions are as follows:

- The penetration of a substance into or through another. For example in air pollution control, absorption is the dissolving of a soluble gas, present in an emission, in a liquid which can be extracted.

- The penetration of atoms, ions, or molecules into the bulk mass of a substance.
- The uptake of water or dissolved chemicals by a cell or an organism (as tree roots absorb dissolved nutrients in the soil).
- The uptake of water, other fluids, or dissolved chemicals by a cell or an organism (as tree roots absorb dissolved nutrients in soil).

ABSORPTION BY ATMOSPHERE

Absorption is the process by which "incident radiant energy is retained by a substance". In this case, the substance is the atmosphere. When the atmosphere absorbs energy, the result is an irreversible transformation of radiation into another form of energy. This energy is transformed according to the nature of the medium doing absorption,

The absorbing medium can also do much more. The medium will only absorb a portion of the total energy. The other energy will either be reflected, refracted, or scattered. The energy absorbed can also be transmitted back into other parts of the atmosphere,

The atmosphere, due to the many different gases and particles contained therein, absorbs and transmits many different wavelengths of electromagnetic radiation, The wavelengths that pass through the atmosphere unabsorbed constitute the "atmospheric windows" (Figure 142). The graph shows the lines of transmission through the atmosphere. The valleys, like at the left end of the scale for visible light, are the "windows" where there is very little attenuating of the radiation by the medium it passes through.

Absorption is mainly caused by three different atmospheric gases. Contrary to popular belief, water vapour causes the most absorption, followed by Carbon dioxide and then Ozone.

Atmospheric absorption of electromagnetic radiation (Figure 143) helps the earth in two main ways. First, absorption helps people by preventing high-energy radiation from reaching the surface which limits our exposure to harmful radiation. The atmosphere absorbs most of the radiation from the ultraviolet region through the x-ray region.

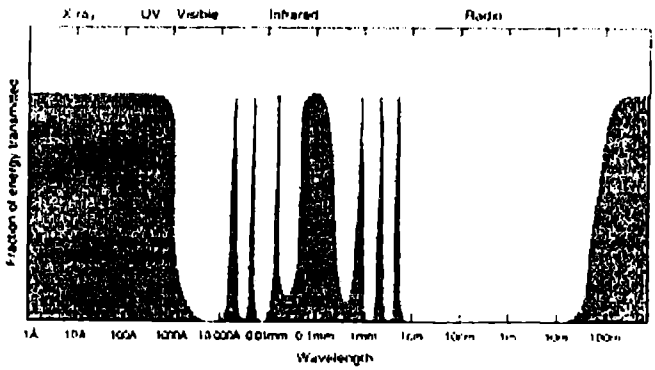


Figure 142: Windows through the Earth's atmosphere.

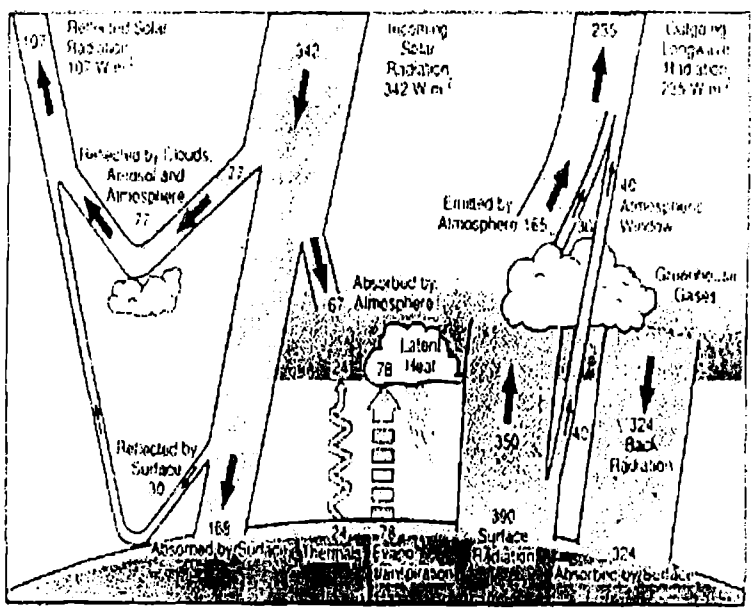


Figure 143: Atmospheric absorption of radiation.

The second way in which absorption helps the earth is as a heat source for it. If one were to take a vertical cross section of the entire atmosphere, one would note that the temperature generally increases with height. This increase in temperature is caused by an increase in absorption of electromagnetic radiation with height due to higher concentration of high-energy wavelength absorbing gases present at higher atmospheric levels.

ABSORPTION BY SOIL

Household not served by community public sewers often depend on septic tank-soil absorption systems to treat and dispose of wastewater. The septic tank removes most settleable and floatable solids from the wastewater; the soil absorption system filters and treats the clarified septic tank effluent. By removing most solids, the septic tank protects the soil absorption system from clogging and premature failure. To work properly, the septic tank needs periodic maintenance.

Soil absorption system can be used in areas where the percolation rate of the soil is between 3 and 60 minutes per inch (soil permeability between 1 and 20 inches per hour). At least 4 feet of suitable soil is required under the soil absorption system to provide adequate treatment of the septic tank effluent. To accommodate the construction of the system and provide adequate soil cover to grade, a minimum of 5.5 to 6.5 feet of suitable soil is needed above the limiting layer. A limiting layer may be bedrock, an impervious soil layer (hardpan, fragipan) or a seasonally high water table (gray soil or mottles). The soil absorption system must be atleast 8 feet from any drain line on the lot, 50 feet from a water supply, and 10 feet from the property line, right-of-ways and the house, septic systems cannot be placed on the flood plain and are limited to areas with less than 15 percent slope.

A septic tank is watertight container constructed of a sound, durable material resistant to corrosion or decay. Septic tanks may have one or two compartments. The compartment tanks or two single tanks in series do a better job of settling and are required for homes with four bedrooms or more.

Among the most important components of a septic tank are the baffles. The inlet baffle forces wastewater down into the tank,

preventing short-circuiting across the top. The outlet baffle keeps the scum layer from moving into the soil absorption system. Septic tanks have inspection ports for checking the condition of the baffles and a manhole for cleaning the tank (Figure 144).

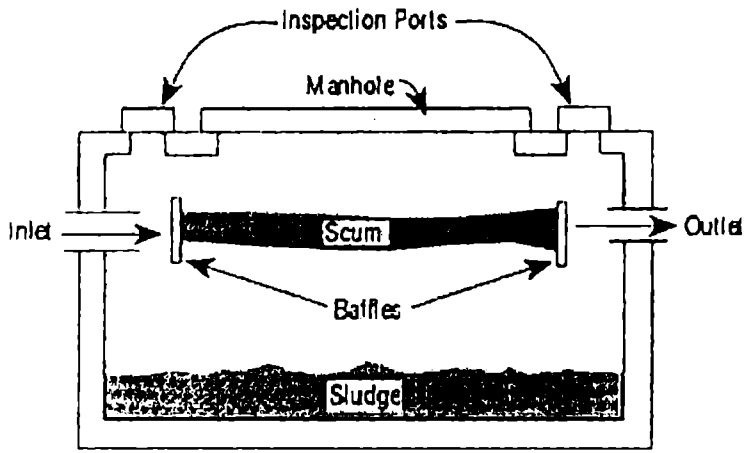


Figure 144: Cross section of a septic tank.

The capacity of the septic tank is based on the size of the house. a 1000 gallon tank is required for a house with one or two bedrooms. For a three bedroom house, a 1500 gallon tank is required. A 2000 gallon tank is required for 4 to 5 bedroom and a 2500 gallon tank is required for 6 or more bedrooms,

The soil absorption system receives effluents from the septic tank and filters and treats the effluent before it enters the ground water. Atleast 4 feet of unsaturated soil beneath the soil absorption system is needed to renovate wastewater before it reaches a limiting layer, A limiting layer may be bedrock, an impervious soil layer or the seasonal high water table,

The soil absorption system is a set of trenches 18 to 30 inches deep, atleast 8 inches wide, and placed atleast 6 feet apart. The maximum length of a trench is 150 feet. The bottom of these trenches must be level; construct them to follow the contours of the lot (Figure 145).

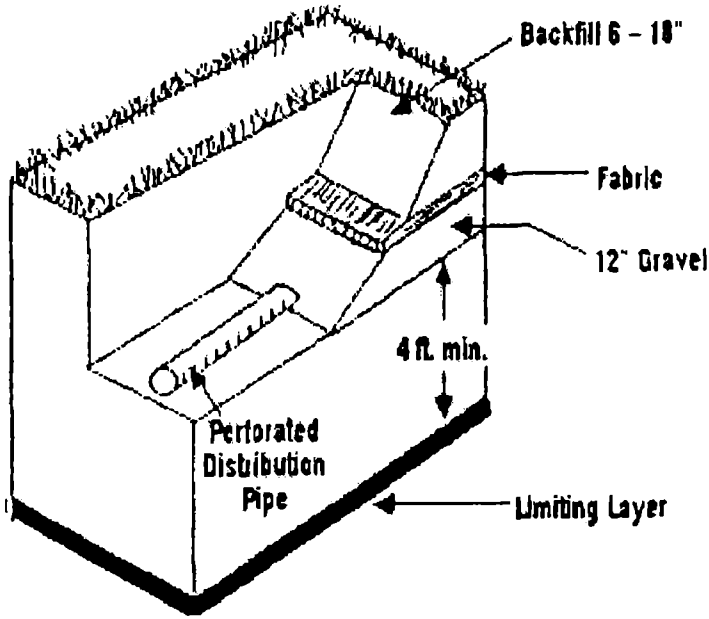


Figure 145: Trench soil absorption system.

The bottom of each trench is filled with 6 inches of clean gravel, A 4 inch perforated pipe is placed on top of the gravel and covered with 2 more inches of gravel. The top of the gravel is covered with synthetic building fabric before the trench is backfilled with native cover soil. This prevents the soil particles from moving down into the gravel. Two inches of straw or a layer of untreated building paper are still sometimes used in place of the fabric. The cover soil should be mounded to account for settling, graded to avoid ponding of rainwater, and seeded with grass to prevent erosion.

The size of a soil absorption system is based on the size of the house and the percolation rate of the soil. The soil absorption system can be divided into two equal sections, A diversion device is used to alternate the flow of wastewater from one side of the system to the other (Figure 146). The diversion device should be easily accessible for annual switching.

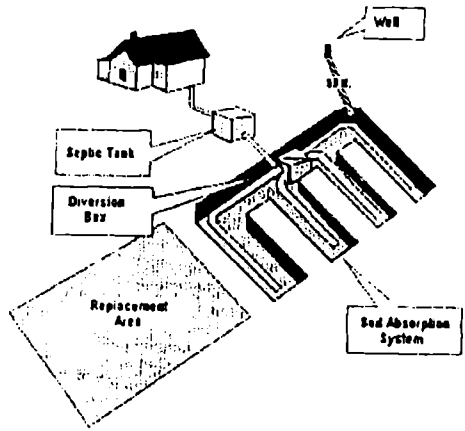


Figure 146: Soil absorption system with diversion box.

ABSORPTION BY ANIMALS

Absorption is the passage of various substances involving *ingestion, inhalation*, and *absorption,* Merely being exposed will not cause harm if the hazard does not actually enter the body. For example, a pack of cigarettes in a man's shirt pocket does not cause harm to him because nothing from the cigarettes has entered his body, if, however, he smokes one of the cigarette, the smoke has entered his body through his lungs and can cause harm.

Chemicals that are ingested enter the body by being eaten. From the digestive track, they can go to the liver or the lymphatic system and then to the bloodstream. Some chemicals are not absorbed by the digestive track, so they pass through the body and are excreted in the urine and faeces.

Chemicals can be breathed into the lungs - inhalation. The inside surface of the lungs is very large and is a poor chemical barrier. Many chemicals that are inhaled can easily and quickly enter the bloodstream from the lung stream.

Chemicals can enter the body by moving through the skin, The skin is a very good barrier and provides protection from

many hazards, but some substances can penetrate the skin, then enter the blood stream and can be carried to all parts of the body.

Absorption from the *mouth* and *eosophagus* is practically negligible. Because food is not broken by enzymes and it is not retained for sufficient time. However, atropine, nitroglycerine (drug for heart attack are given sublingually) and epinephrine are absorbed in the mouth.

Most of the substances are not absorbed by the *stomach* due to the absence of villi-like structures and the junction between the epithelium is very light. However, those substances which do not require digestion, for example, water, alcohol, saline, glucose, and certain easily diffusible drugs are absorbed by stomach.

Small-intestine is the chief site for absorption. After 5 to 8 hours of ingestion most of the semi-digested food passing from the small-intestine into the *large-intestine* hardly contain any absorbable substances except water, which is vigorously absorbed from large-intestine. Different portions of the small-intestine vary in their ability to absorb different substances. The upper portion absorbed sugar faster than water. In the lower portion water is absorbed faster than sugar. Sodium chloride is absorbed in the lower portion even faster than water. Fats and bile salts are absorbed in the ileum. The absorption in the small intestine is due to: (1) Greater length of the small intestine which provide large surface area; (2) Presence of villi, which provide increased absorptive surface area; (3) Presence of brush-border epithelium which further increase surface area for absorption; (4) Presence of arterial and emphatic vessels in the villi help absorption of non-lipids and lipid substances; and (5) Lymph vessels which drain the intestinal walls, digests and absorbs fats.

Large-intestine absorbs water, salts and glucose, human large-intestine absorbs 80 ml of water per hour.

ABSORPTION BY PLANTS

Plants generally absorb water by the root system from the soil. Depending on the plant species and ecological conditions some plants have deeply penetrating root system while others have shallow spreading root system. The root hair zone below the root tip is mainly responsible for water absorption.

The root hair is a unicellular structure and it is the elongated projection of certain cells of the epiblema. It is generally less than 1 mm to about 1 cm in length and is about 10 nm is diameter. The outer wall of the root hair is made up of pectic compounds and the inner layer is cellulosic. Inside the wall a thin cytoplasmic layer is present, which is continuous with the protoplasmic layer of the main cell of the epiblema. Extending from the main cell to the hair there is a central large vacuole, containing cell sap. The thin protoplasmic layer acts as a differentially permeable membrane.

Root hair absorb water from the soil. Water is drawn into the root hair, mainly by osmosis. Due to presence of sugars and other metabolites, the concentration of solutes in the sap vacuole is greater than that of the surrounding soil water. Water molecules pass through the permeable cellulose wall and semipermeable protoplast into the vacuole.

From root hairs, water passes to the vascular tissue in the centre of the root via the parenchyma cells. There are three possible mechanisms: (a) through the sap vacuole, water is drawn from one vacuole to the next by osmosis (osmotic gradient): (b) through cytoplasm from one to another; and the (c) along the cell walls through the cellulose of adjacent cells and through the small intercellular spaces between them, water diffuses (diffusion gradient) (Figure 147)

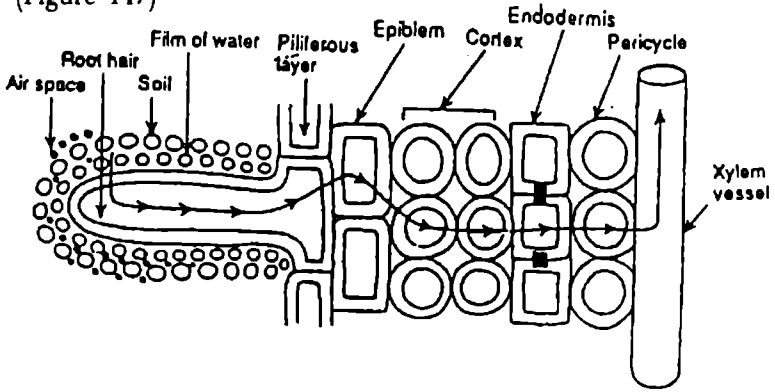


Figure 147: Absorption and transport of water by plants.

Transport of water from parenchyma cells into the vascular tissues of the root is due to high osmotic pressure developed in the leaf cells. As a result of transpiration the water is drawn into and along the vascular tissues, that is pull exerted by the leaves. Besides that there is a force, pushing water up the stem from the roots is called root pressure.

The mineral salts remain in the soil solution in dissociated condition. The essential ions are absorbed by plants in different amounts by the root hairs and are then translocated through the xylem stream to the different parts of the plant body. Previously, it was assumed that inorganic salts were passively absorbed en mass along with water. However, it has shown that salt absorption is largely dependent on metabolic energy.

There are two major mechanisms for ion absorption-*non-metabolic or passive* and *metabolic or active.* An ion will move in a given direction only if driven by some force. If ion transport in and out of cells occurs spontaneously down a gradient of potential energy, it is called non-metabolic or passive absorption. When ions are driven up such a gradient by some process directly coupled to metabolism, it is called metabolic or active absorption.

ADSORPTION

The process of accumulation of a substance on the surface of another substance in called adsorption. The substance that sticks or adheres on the surface of another substance is called *adsorbate* and the substance on the surface of which the adsorbate settles is called *adsorbent,* The common surfaces between the two phases where the adsorbed molecules concentrate is called the *interface,*

The use of solids for removing substances from either gaseous or liquid solutions has been widely used since biblical times. This process, known as adsorption, involves nothing more than the preferential partitioning of substances from the gaseous or liquid phase onto the surface of a solid substrate. The binding to the surface is usually weak and reversible, just about anything including the fluid that dissolves or suspends. The material is bound, but compounds with colour and those that have taste or odour bind

strongly. Compounds that contain chromogenic groups (atomic arrangements that vibrate at frequencies in the visible spectrum) very often are strongly absorbed on one char (*activated carbon*), Decolourization of sugar solution and other foods can be wonderfully efficient by adsorption and with negligible loss of other materials, The use of activated carbon for adsorbing nerve gas from the battlefield is a useful tool, among thousands of other applications,

The most common industrial adsorbents are activated carbon, silica gel, and alumina, because they present enormous surface area per unit weight. Activated carbon is produced by roasting organic material to decompose it to granules of carbon-coconut shell, wood, and bone are common sources, silica gel is a matrix of hydrated, silicon dioxide. Alumina is mined or precipitated aluminium oxide and hydroxide. Although activated carbon is a magnificent material for adsorption, its black colour persists and adds a gray tinge if even trace amounts are left after treatment; however, filter materials with fine pores remove carbon quite well.

A surface already heavily contaminated by adsorbates is not likely to have much capacity for additional binding. Freshly prepared activated carbon has a clean surface. Charcoal made from roasting wood differs from activated carbon in that its surface is contaminated by other products, but further heating will drive off these compounds to produce a surface with high adsorptive capacity. Although carbon atoms and linked carbon are most important for adsorption, the mineral structure contributes to shape and to mechanical strength.

Adsorption phenomenon are operative in most natural physical, biological, and chemical systems. Adsorption operations employing solids such as activated carbon and synthetic resins are used widely in industrial applications and for purification of water and waste-waters.

The process of adsorption involves separation of a substance from one phase accompanied by its accumulation or concentration at the surface of another, adsorbing phase is the adsorbent, and the material concentrated or adsorbed at the surface of that phase is the adsorbate.

Physical adsorption is caused by van der Waals forces and electrostatic forces between adsorbate molecules and the atoms which compose the adsorbent surface. Thus, adsorbents are characterized first by surface properties such as surface area and polarity.

A large specific area is preferable for providing large adsorption capacity, but creation of large internal surface area in a limited volume inevitably give rise to large number of small sized pores between adsorption surfaces. The size of the minipores determines the accessibility of *adsorbate* molecules to the internal adsorption surfaces, so the pore size distribution of *micropores* is another important property for characterising *adsorptivity* of adsorbents. Especially materials such as zeolite and carbon molecular sieves can be specifically engineered with precise pore size distributions and hence tuned for particular separation.

Surface polarity corresponds to affinity with polar substances such as water or alcohols. Polar adsorbents are thus called "hydrophillic" and alumino-silicates such as zeolite, porous alumina, silica gel, or silica-alumina are examples of adsorbents of this type, on the other hand, non-polar adsorbents are adsorbents which are generally "hydrophobic". Carbonaceous adsorbents, polymer adsorbents and silicates are non-polar adsorbents. These adsorbents have more affinity with oil or hydrocarbons than water.

TYPES OF ADSORPTIONS

Adsorption can be of three types: (a) physio-sorption (by van der Waals and electrostatic forces), (b) Chemisorption (by chemical binding), and (c) Biosorption (by living organisms).

(a) Physiosorption

This is the most common form of adsorption, The molecules are attracted by van der waals forces, and attach themselves to the surface of the solid. The molecules remain intact, and can be freed easily (the forces are small, and short range).

The electric field outside the surface, produced by the spill out of the electrons, polarizes the adsorbed atom, creating an induced dipole, The interaction of the induced dipole with the surface charge is attractive, and similar to the Lennard-Jones interaction:

$$V_{JL} = -c/r^3$$

first discussed by John Bardeen (Twice Noble prize winner for transistor and superconductivity). The so-called dispersion forces due to oscillating charge fall faster, according to the London potential:

$$V_L = \alpha_1\alpha_2/r^6$$

Where α_1 and α_2 are the polarizabilities.

Table 10. Polarizabilities (in A^3)

Ne	0.39	CO_2	2.65
Ar	1.63	H_2S	3.78
Kr	2.46	N_2O	2.26
Xe	4.00	CH_4 (methane)	3.00
H2	0.79	C_2H_6 (ethane)	2.60
N2	1.76	C_2H_4 (ethylene)	4.26
02	1.60	C_6H_6 (benzene)	10.32
CO	1.95	CH_3 $COCH_3$(acetone)	6.33

Molecules with permanent dipole moments (Table 14) μ can further polarize each other and give a potential:

$$V_{Debye} = \alpha_1\mu_2/r^6 - \alpha_2\mu_1^2/r^6$$

Direct dipole interaction gives:

Table 15: Dipole moments (in Debyes)

H_2O	1.84
H_2S	0.89
NO	0.16
CO	0.12
N_2O	0.17
HF	1.91
HCl	1.08

NH_3	1.45
CH_3OH	1.68
CH_3 CHO	2.72
CH_3 $COCH_3$	2.90

Adsorption is governed also by the interaction between the adsorbed molecules, which have the same origin.

- Van der Waals attraction, due to the correlated charge fluctuations, occur for all absorbed molecules. It is very weak and thus important only at low temperatures.
- Dipole forces are related to the permanent dipole moment of the adsorbate (for example H_2O, NH_3) or dipoles induced by charge transfer from the surface. The interactions are repulsive between parallel dipoles and attractive otherwise.
- Repulsion due to orbital overlap, important in densely packed layers.
- Substrate mediated forces, produced when the surface modification induced by one adsorbate acts on another adsorbate.

The main aspect of the adsorption process in the bonding of the incoming molecules to the surface, its motion across it or to the interior (diffusion) and the evaporation (desorption).

(b) Chemisorption

Metals such as copper and aluminium could be considered as chemisorptive absorbents, since they collect molecules like sulphur in air, and bind them to their surfaces by chemical reactions. This results in the greenish compound - copper sulphate, that collects on the surface of copper, and the whitish surface of aluminium, which is aluminium oxide.

Activated charcoal in commonly used to remove odour (Figure 148) molecules from air in gas masks.

Adsorptive materials can be used to control the production of reactants in a chemical reaction, in this case, the adsorbent is unaffected by the gases passing over it. This is what a catalytic converter on an automobile does - it binds some of the pollutants from the exhaust gases to its surface.

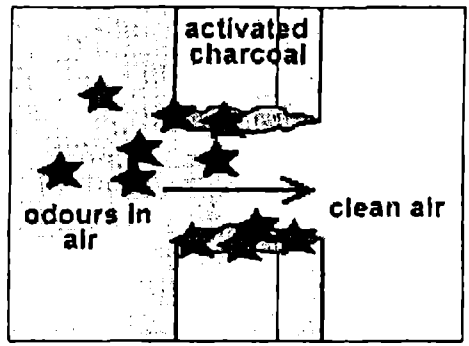


Figure148: Adsorption of odours by activated charcoal.

Chemisorption is characterised by high binding energies. The surface chemical bond results from charge transfer or charge redistribution involving the surface and the absorbate. High adsorption energies are found for transition metals, due to their incomplete d-bands. The range of energies for hydrogen adsorption is around 2.5 ev and oxygen adsorption is 4 to –10 ev.

Noble metals like gold are noble partially because they do not contain empty localised levels, which can participate in the chemical bonding. In addition, the large electron cloud of gold strongly repels the electron cloud of the incoming atom or molecule, and also leads to the population of anti-bonding (repulsive) states that provide a repulsive force to the gold.

In dissociative chemisorption the internal molecular bond is broken by interactions with the substrate. In the interaction of the generic molecule X_2 with the surface S, the electronic potential curve X_2 · –S, describing the energy of the system as a function of the distance X_2 to the surface crosses the corresponding potential curve X + X – S, which describes the interaction of the dissociated fragments with the surface. At the crossing, there can be a dissociative transition, unless the energy at the crossing is positive. In that case, χ_2S might stay at a local minimum until the barrier may be overcome by activation.

Description of chemisorption in terms of energy-distance curve is an over simplification since more coordinates are needed. Not only depends also on the coordinate along the surface but also on the interatomic distance in the molecule, its orientation with respect to the surface, and individual coordinates of the dissociation products. Thus, the problem of obtaining the relevant energies for the chemisorption process is extremely complicated.

The chemisorbed state can be studied in several ways. Changes in work function give information on charge transfer. An increase (decrease) in work function corresponds to a dipole with the positive (negative) side towards the surface. Changes in the electronic density of states are best studied by ultraviolet photoelectron spectroscopy.

ADSORPTION VS ABSORPTION

Adsorption is fairly rapid process and surface phenomenon, whereas, absorption is a slow process and involves diffusion into the interior of the material. Thus, a chalk stick, when dipped in ink, adsorbs the ink and when broken it is found to be white from within. On the other hand, water is adsorbed by Calcium chloride crystals. Both adsorption and absorption often go side by side and it is difficult to distinguish between the two (Figure 149)

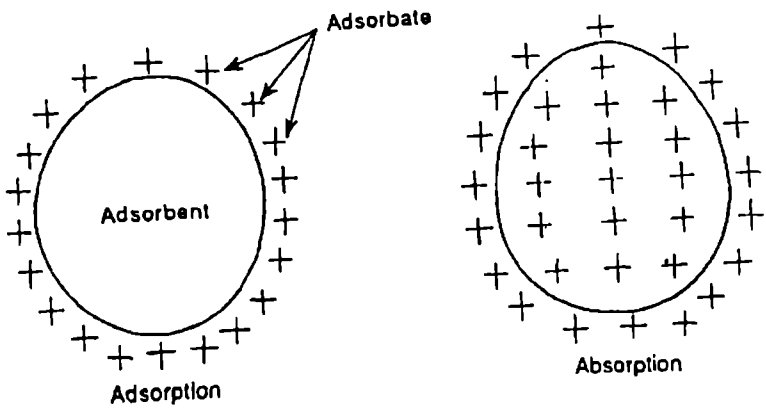


Figure 149: Process of adsorption showing adsorbate and adsorbent and absorption.

Activated carbon, silica gel, and diatomaceous earth are some of the important adsorbents. The attractive forces between atoms, molecules and ions which hold together a solid are unsatisfied at the surface of a solid and thus available for holding particles or other materials such as gases and liquids. Adsorption is reversible process, the same adsorbent can be used more than once. Rise in temperature decreases adsorption due to the increased kinetic energy of the adsorbed molecule, which causes them to escape more readily from the adsorbing surface. The amount of adsorption is proportional to the surface area and varies with the nature of the surface of the adsorbent and the substance adsorbed. Due to adsorption, the residual forces decrease, and therefore, the surface energy gets decreased considerably. This energy is lost as heat energy. Thus, adsorption is always accompanied by evolution of heat. When one gram mole of gas is adsorbed on the surface of the solid, it is called *heat of adsorption.*

(c) Biosorption

Metabolism independent binding or adsorption of heavy metals to living or dead cells, extracellular polysacchrides, capsules and slime layers are referred to as "biosorption". Walls and envelops of bacteria, yeasts, algae, fungi are very efficient in biosorption due to the charged groups present in them. Biosorption is influenced by environmental factor such as pH and ion competition and some metabolism aspects which can affect the microenvironment of cells. Variation in chemical behaviour of metal species, composition of microbial cell wall and extracellular material also effect the biosorptive capacity. Gold and silver crystalize due to the reduction while uranium and thorium crystalize due to the formation of hydrolytic products. Metals may be adsorbed around cells in the form of phosphates, sulphides or oxides.

The microbial removal of potentially toxic and/or valuable metal and metalloid species, radionuclides, organometalloids, metal particulates can be done from effluents, industrial waste waters. This can result in detoxification and safe environmental discharge. Subsequent treatment of loaded biomass can enable recovery and/or containment of toxic/radioactive metal elements.

To prove economically successful, biosorption should result in 99 percent removal with metal loading greater than 150 mg/gm of biomass. Biosorption outperforms chemical and physical techniques in many respects. Still there is reluctance in the application of biosorption which may be due to:

(1) Chemical engineers are more accustomed to non-biological processes:

(2) Units are operating on non-biological processes and turning to biological processes will be expensive;

(3) Process engineers find themselves comfortable in handling already handled physico-chemical system.

USE OF BACTERIA IN BIOSORPTION

(1) *Pseudomonas fluorescence* immobilized on polyvinyl chloride granules and packed into columns at the rate of 588 g/lit column volume was used primarily to remove nitrate but was found to be capable of removing metals such as lead and zinc. This system was used for the treatment of effluents containing zinc, chromium, barium, aluminium, nickel, lead and nitrate after reducing metal concentration to about 1 mcg/ml by precipitation with lime.

(2) *Pseudomonas aeruginosa* immobilized on an oxygen plasma treated polypropylene web with a 1.5 gm web containing 0.3 gm cells/gm of polypropylene were fed at the rate of 15 ml/h with plutonium at a concentration of 1.7 nCi/100 ml for 12 hours. This gave 75-80 percent removal efficiency. With 2 columns in series, efficiency was 98-99 percent with a useful column life of 2 weeks. This system was not used on an industrial scale and no attempt was made to recover plutonium. It was suggested that plutonium loaded biosorbent could be packed in drums and dumped.

(3) Accumulation of heavy metals by "resting cells" of *Citrobacter* species is mediated by surface located phosphatase enzyme that releases $(HPO_4)_2$ from supplied substrate (for example glycerol 2 phosphatase) and precipitates divalent cations (M^{+2}) as $MHPO_4$ at the cell surface. The process is non-specific and depends on the fact that certain metal phosphates are insoluble. Cadmium, Lead, Copper can be precipitated singly or in

combination. *Citrobacter* process is capable of long term use, functional with a high metal load, possibility of metal recovery and biomass regeneration are the advantages. Possibility of enzyme (phosphatase) inhibition by effluent components is disadvantage. *Citrobacter* species are immobilized in polyacrylamide gel columns (10 cm Ú 3 cm). When 3 columns are used in series and 10 lit of 200 mcg Cd/ml passed, 90 percent removal could be achieved. Mixture of cadmium copper and Lead in above system show 87, 83 and 85 percent removal when inflow was 112 meg Cd, 0.63 meg Cu and 2.1 mcg lead per ml. This system has commercial exploitation potential. Uranium from 200 meg/ml solution could be removed by this process. Uranium removal was 91 percent when 24 lit solution was passed. 4.5 gm uranium could accumulate in the column. The metal could be removed from the loaded columns by dilution with 1M citrate buffer pH 4.33 gm uranium was released in 5 litre effuent. For the recovery of cadmium, citrate buffer (0.5 M pH 5) can be used and 85 percent efficiency was retained on reuse of columns.

(4) Granulated *Bacillus* are used at the rate of 20 kg in fixed bed reactors for small flows of less than 15 lit per minute, where as larger system with 80-90 kg biomass is used for larger flows of more than 35 lit/minute. Metals are removed from the biomass using H_2SO_4, NaOH or complexing agents and is recovered using electrowinning. Regeneration of granules may be achieved by alkali treatment.

(5) Metals like gold and palladium present in industrial/mining waste water can be removed by adsorption even in ppm amounts. Precious metals thus trapped by adsorption can be easily recovered by using thiourea as an intermediary. The recovery rate is 98.6 percent. *Pseudomonas, Micrococcus leutus, Streptomyces phacermogenes* are capable of acting as adsorbents. Both live and dead microorganisms can be used. One gram of bacteria can recover as much as 180 mg of gold. Recovery occurs in 5-10 minutes.

(6) Heterotrophic bacteria - *Sphaerotilus natans* accumulates metals in mucilagenous layer outside sheath. It accumulates iron, copper, magnesium, cobalt, cadmium when present in sulfates but growth is inhibited if these metals are present as chlorides.

Although the ability of *Sphaerotilus-Leptothrix* group to adsorb metals is enormous, it is not commercially exploited as yet.

(7) *Streptomyces virido-chromogenes* immobilized in polyacrylamide gels is used for uranium adsorption. It removes 100 percent uranium from 200 ml of 20 mcg/ml at pH 6. Bound uranium could be released with 10 ml of 0.1 M sodium carbonate and adsorption capacity is retained through 5 cycles of loading and unloading of uranium. For sea water with 10 meg/ml of uranium, efficiency is 80 percent (4 cycles).

(8) Bacterium *Deinococcus radiodurans* adsorbs doses of nuclear radiations many thousands of times stronger than those that would kill a person. Scientists have established that the microbe doesn't simply shield its DNA from radiation. It instead has an unprecedented ability to repair genetic damage. A dose of 500 to 1000 rads is lethal to the average person. In contrast, *D. radiodurans* thrives after exposure of upto 1.5 million rads. Cool and freeze microbe may survive upto 3.0 million rads. Thus the bacterium is supreme in radiation resistance and can be used for the bioremediation of the sites contaminated by nuclear weapons production and development of nuclear reactors. Such sites are polluted with radioactive elements such as uranium and plutonium and an array of heavy metals. For clean up of such sites *D. radiodurans* my offer a solution.

USE OF FUNGI IN BIOSORPTION

The removal of heavy metals or radionuclides from effluents is done with live or dead fungal mycelium. This can help for cleaning polluted effluents as well as recovery of precious metal ions.

Toxic heavy metals like Lead, Cadmium, Zinc, Cobalt, Chromium, Nickel and various actinide elements like uranium, thorium, plutonium can be adsorbed and removed by fungal biomass. Fungal biomass can also be used in effluent treatment to adsorb other particulate matter like metal particulates etc. Colloidal gold can be removed and recovered. Particulate adsorption by fungal cell wall occurs independent of cellular metabolism. Young actively growing mycelial tips are more effective.

Mycelial biomass as bioadsorbent serves two purposes : (1) Removal of toxic metals from effluents; (2) Overcoming internal negative charge, passive influx of K+ from surrounding cells takes place (Levitt, 1974). The potassium malate in the vacuole gives osmotic potentiality to guard cells causing stomatal opening (Figure 150).

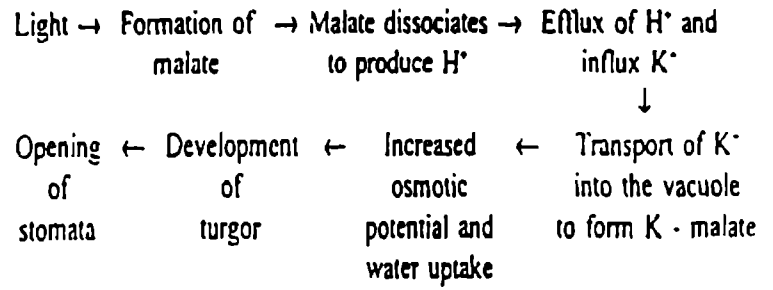


Figure 150: Series of events leading to stomatal opening by K+ transport hypothesis.

In dark, the sequence of events is reversed leading to stomatal closure. The K^+ and Cl^- ions are transported out of the guard cells in dark. The role of Cl^- is not clear, though its influx during opening of stomata during the day is obvious atleast in some cases.

ULTRA FILTRATION

Filtration is a process that is used to separate one or more components from a fluid stream. In membrane filtration that has selective permeability, physical difference among solution components influence the retention or transport through the membrane. Ultra-filtration (UF) is a pressure driven membrane separation process that uses molecular size difference to separate macromolecules and colloidal matter from solvents and small solutes.

The differential driving potential used to transport solvents across ultrafiltration and reverse osmosis (RO) membranes is the hydrostatic pressure. The difference between the two processes is the applied pressure range, ultrafiltration is a low pressure process usually less than 10 atm, while RO operates at pressures

above 40 atm. The particle size range for ultrafiltration technology applications extends from 10 A to 200 A and roughly corresponds to a molecular weight range from 500 to 500,000 amu. On the other hand RO is used to separate molecules as small as ionic species in size.

Ultrafiltration is form of filtration that uses membranes to preferentially separate different fluids or ions. Ultrafiltration is not as fine a filtration process as ***nanofiltration,*** but it also does not require the same energy to perform the separation. Ultrafiltration also uses a membrane that is partially permeable to perform the separation, but the membrane's pores are typically much larger than the membrane pores that are used in nanofiltration.

Ultrafiltration is more commonly used to separate a solution that has a mixture of some desirable components and some that are not desirable. One of the uses that demonstrates the usefulness of ultrafiltration is electro-deposition paint recovery. In this instance the paint, composed of a resin, a pigment and water are separated into two streams that can he reused. The first stream includes the water and a small amount of the paint resin, which can be used to rinse the parts later in the process, The paint pigment is separated from that stream and can be reused in the paint path, allowing the bath to be concentrated to a usable level.

Ultrafiltration is capable of concentrating bacteria, some proteins, some dyes, and constituents that have a larger molecular weight of greater than 10,000 daltons. Ultrafiltration is only somewhat dependent upon the charge of the particle and is much more concerned with the size of the particle. Ultrafiltration is typically not effective at separating organic streams.

Colloidal solution can pass easily through ordinary filter paper because its pores are not small enough to retain the colloidal particles. But a semipermeable membrane made of collodion, cellophane, inert cellulose esters or hardened gelatine supported properly on a grid of wire mesh or deposited on a piece of porcelain or filter paper does not allow the colloidal particles to pass through these membranes.

Ultrafiltration is the process of separation of particles of colloidal system from electrolytes-water, glucose etc. by filtering

the mixed solution through a membrane of different pore sizes under, pressure. Ultra-filtration is the direct opposite of osmosis: osmosis causes migration of water from lower to higher concentration and is *mixing* process whereas ultrafiltration causes migration of water from higher to the lower concentration and is a *separation* process.

In order for ultrafiltration to occur two criteria must by fulfilled: (1) the barrier membrane, and (2) pressure to force fluid the barrier. The two biological membranes - capillary membrane and glomerular membrane in the kidney fulfil both the criteria.

The capillary membrane is composed of a unicellular layer of endothelial cells which are surrounded by a basement membrane on the outside. The total thickness of the wall is 0.5 micron and the diameter of the capillary is 5-9 microns through which a blood cell can manage to squeeze through.

The intercellular cleft which is a thin slit or pore at the junction between the adjacent endothelial cells. The space between the endothelial cells, for example, width of the intercellular slit pore is 9-7 nanometers filled with loose reticular fibrillae composed of proteoglycans mainly hyaluronic acid.

There are picnocytic vesicles at one surface of the endothelial cell and more to the opposite surface carrying large molecules and solid particles through the basement membrane to discharge their contents. Sometimes these picnocytic vesicles join with each other to form a continuous channel through the endothelial membrane, called *picnocytic channel.* There vesicles are largely responsible for the exchange across the capillary wall

Diffusion results from the thermal motion of the water molecules and the dissolved substances in the fluid. Lipid soluble substances diffuse directly through the capillary membrane, their rate of diffusion is two times greater than water soluble substances.

Water molecules diffuse through endothelial cell and membrane pores.

ELECTRICAL CONDUCTIVITY

Electrical properties on the surface determines the permeability of the membrane to ions which also posses electrical charges. Cell

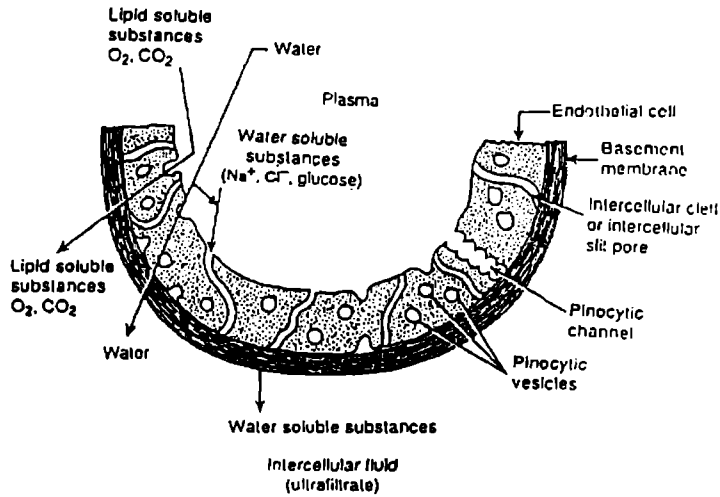


Figure 151: Capillary membrane pores and ultrafiltration.

membrane with a positively charged surface repels positively charged particles and thereby prevents their passage, while negatively charged particles or uncharged particles even larger than ions might readily enter or leave the cell. Migration of negatively charged ion into the cell could not occur unless there was simultaneous migration of equal numbers of negatively charged particles. Three types of electrical forces effect the ion permeability:

(a) The intensity of electrical charge lining the ionic channel

(b) The hydration energy of ions.

(c) The difference of potential across the membrane.

It has been known for 200 years that nerve impulses are electrical in their nature. However, unlike electricity as we usually think of it, the electric signals produced by cells are carried by ions rather than electrons. There are many implications that follow from this distinction. For example, electrical impulses in living cells do not travel at the speed of light as electrons do. Also, since ions can carry either positive or negative charges, the electrical signals sent within cells are more variable than those that are sent in man-made electrical circuits.

Electrical currents are not only the basic for nerve signals; they also trigger and initiate muscle contraction. So, before we get into discussing the physiological mechanism of muscle contractions and nerve signalling we must have a look at the physical and molecular basis for electrical phenomenon in cells (Fig. 152)

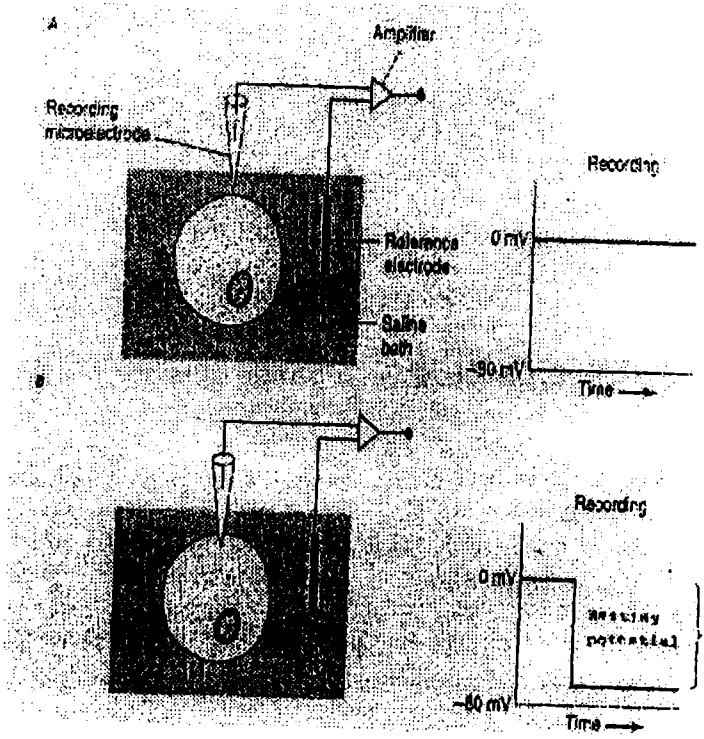


Figure 152: Electrical measurement in cells.

With modern technique it is fully possible to measure the electrical potential difference across plasma membranes. To measure the transmembrane potential of a cell we typically need an electrode with a very fine tip. Usually a glass capillary pipette microelectrode is used. The lumen of this hollow glass electrode is filled with a salt solution that can conduct electricity to the silver or platinum wire. We also need a sensitive voltmeter. For most applications, an oscilloscope will do the trick. By convention, the transmembrane potential is always given as the intracellular potential

relative to the electrical potential outside the cell. Here the reference potential is recorded by a silver electrode that is submerged in the bath.

Before we insert the tip into the cell, the recording microelectrode and the reference electrode are at the same electrical potential. As the microelectrode penetrates the cell membrane and the cytosol makes contact with the salt solution inside the electrode, the recording on the voltmeter suddenly shifts and indicates a negative potential inside the cell.

This negative potential recorded by the micrbelectrode is the resting potential of the cell. The vast majority of cell types in eucaryotes have a negative resting potential, like this one The negative potential varies greatly, but resting potentials are between -50 and -80 mV most commonly.

If we expand our experimental setup with a second electrode, we can study the cell membrane of a nerve cell responds to pertubations in the membrane potential. Through this second electrode we can inject a current into the cell and the changes in membrane potential can be recorded with our first microelectrode (Fig.153).

By adding positive charges through this external electrode, we make the outside of the cell more positive, while the inside becomes more negative than before because positive charges are removed from the interior of the cell by the current electrode. The result is that the inside of the cell becomes even more negative, compared with the outside. We say that the cell has become hyperpolarised.

If positive charges are added to the inside of the cell by the current electrode, the potential difference across the membrane will decrease and the cell is now depolarised.

The upper panel shows the registration on the voltmeter and the lower panel denotes the pulses given through the current electrode. The downward registration indicate that we take away positive charges from the interior of the cell and the registrations above the line illustrate stimulations where positive charges have been injected into the cell.

When we apply a hyperpolarising current, the registration on the voltmeter shows that the negative membrane potential becomes

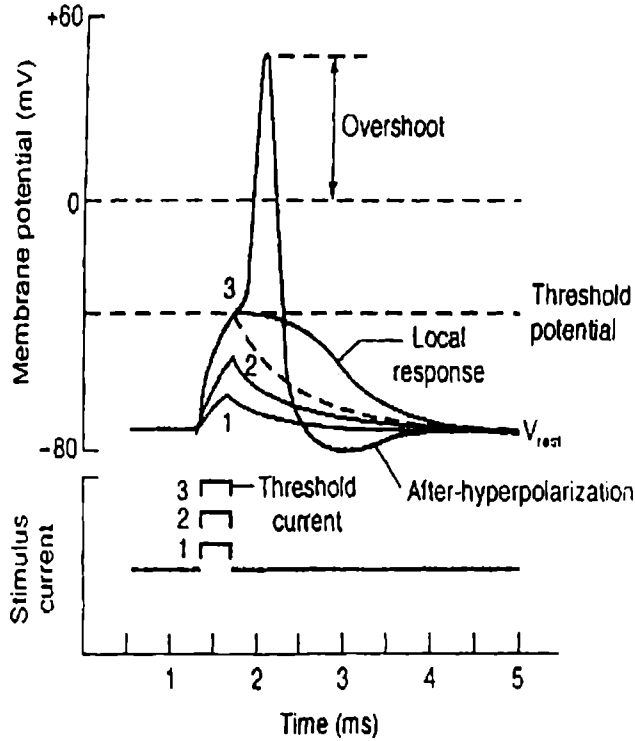


Figure 153: Passive and active membrane electrical responses in nerve cells.

larger. That is the inside of the membrane becomes even more negative than before. And the stronger the stimulus, the stronger the hyperpolarisation. However, the membrane potential is back at the resting value as soon as a few milliseconds after the current is terminated. We say that the membrane repolarises.

If we add a weak depolarising current to the interior of the cell - that is we add positive charges to the cell- a similar effect is obtained, but with opposite polarity than before. The voltmeter registers a slight depolarisation of the plasma membrane that dies out shortly after the termination of the stimulating current. So far, this nerve cell has not responded differently from any other cell.

However, if we now give a stronger dipolarising current to the cell, something new will happen. Instead of following the predicted pattern of depolarisation and subsequent repolarisation, the recording electrode registers a vigorous depolarisation of the membrane. For a moment, the inside of the nerve cell is actually positively charged as compared with the exterior.

What we have seen is a ***positive feed-back*** where the depolarisation of the nerve cell membrane triggers the opening of Na+ channels in the membrane. The massive depolarisation is a result of positively charged Na+ ions that flow into the cell down their electro-chemical gradients. For this positive feed back response to occur, the plasma membrane must first be depolarised above a given threshold and this is what we did when we applied this second stronger dipolarising current.

The whole phenomenon is called *action potential* (AP) and it constitutes the basis for long-range nerve potentials and here they serve as a trigger that initiates a sequence of events that finally leads to the contraction of the muscle.

An action potential is an active electric response that is self-propagating. That is, once an AP has been initiated it will continue to run until it reaches the end of the cell. It takes energy in the form of an ATP to run an AP because it is dependent upon the maintenance of a Na+ gradient across the cell membrane.

In contrast, the small waves of hyperpolarisation and depolarisation, that were observed at all stimulations, which resulted in membrane potentials below the threshold, were all passive electrical responses. These responses are not dependent on any physiological changes, such as opening or closing of ion channels. This is the main difference between active and passive electrical responses.

There are two structural features of cell membranes that give them two characteristic electrical properties (Fig. 154 and 155) (1) Especially, in nerve cells, extensive portions of the membrane contains very few ion channel proteins. Since, the lipid bilayer itself is highly impermeable to ions, this means that large areas of the nerve cell membrane are virtually impermeable to ions. Instead negative ions accumulate on the inside of the thin plasma-

membrane attracted by the positive charges on the outside. In essence the plasma membrane therefore has the properties of a capacitor in an electric circuit.

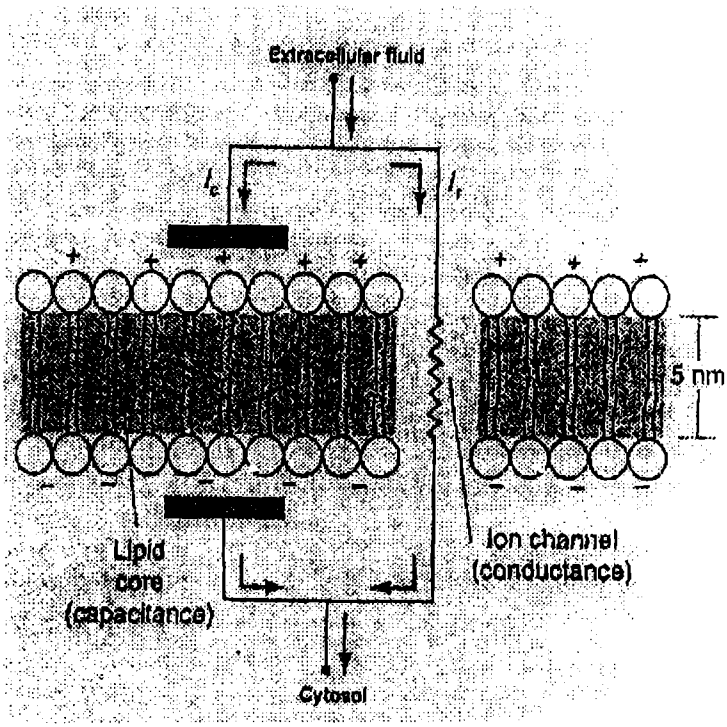


Fig. 154: Plasma membrane as a capacitor in an electrical circuit.

(2) Certain areas of the nerve cell membrane do not have dense populations of ions channels. Ions are electrically charged and so the channels carry electrical charges across the membrane and they give the membrane a certain conductance to ions. The conductance of a membrane is a measure of its permeability to ions. The greater the conductance of the membrane to an ion the easier can that ion pass the plasma membrane. The analogy for these ion channels in an electric circuit would be a resister.

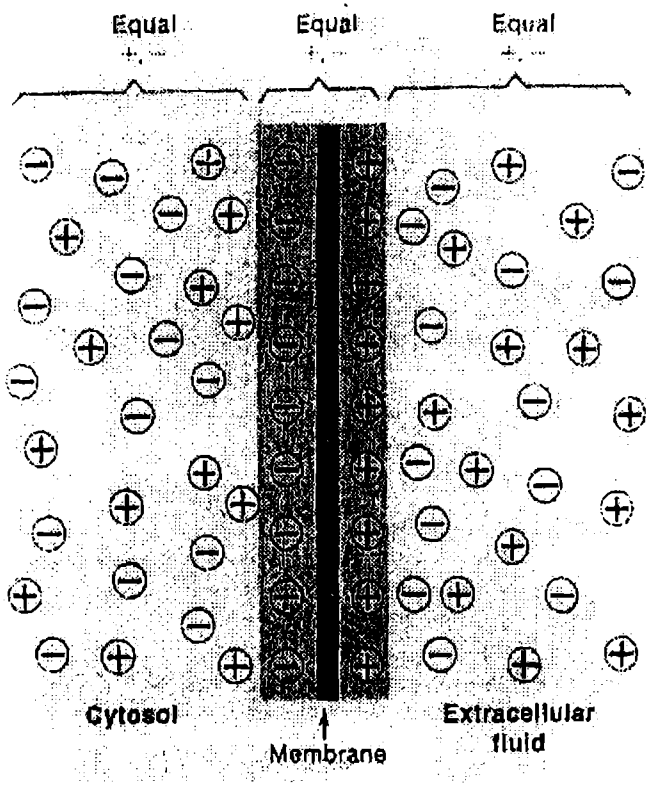


Fig. 155: Ion channel carrying electrical charges across the membrane.

It is these two features that are responsible for the passive electrical properties of the cell membranes.

The basis of all nervous signals in an animal body is the maintenance of electrochemical gradients across cell membranes (Fig. 156). Electrochemical gradients originate in: (1) the presence of negatively charged proteins in the cell, (2) the pumping of selected ions across the plasma membrane, (3) a relative low permeability of the plasma membrane lipid bilayer to electrolytes, and (4) selective permeability of ion channels in the plasma membrane.

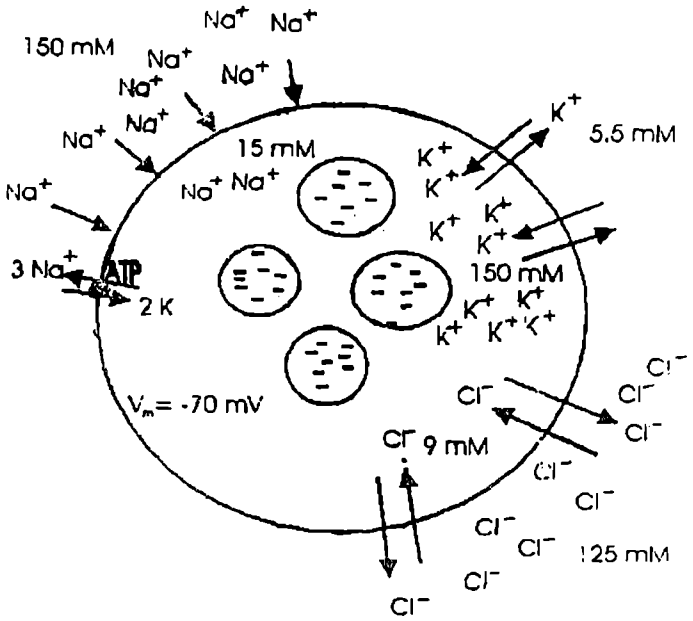


Figure 156: Electrochemical gradient across the plasma membrane of a mammalian nerve cell.

With this schematic drawing attempt is being made to explain how electrical and chemical gradients are generated across cell membranes. The figure is a representation of a typical animal cell. The major inorganic ions in our body are Na+, K+ And Cl−, and these are the most important ions in creating a membrane potential. In addition to these inorganic ions, the cystosol of cells is loaded with negative charges in the form of proteins (the vast majority of all intracellular proteins are anionic).

Cell membranes are far more permeable to K+ than to Na+. In addition, Na+ is actively pumped out from the cell by the Na^+, K^+-ATPase (the sodium pump), leaving very little Na^+ inside the cell. The permeability to Cl^- varies. In some cells it is similar to that of K^+ and in other cells, it is lower. The membrane is of course largely impermeable to the large negatively charged protein molecules inside the cell.

1. The negative charges of proteins in the cell attract positive charges, such as Na+ and K+ and repel negative charges such as Cl^-.
2. Since the mebrane permeability to Na+ is very low and that of Na^+/K^+-ATPase continuously pumps out Na^+ from the cell and K^+ into the cell, K^+ is the ion that will accumulate in the cell to balance the negatively charged proteins.

The result is that the most concentrated inorganic ion in the cytoplasm is K^+, which typically is 10-30 times as concentrated as in the ECF. Coversely, the intracellular concentration of Na^+ and Cl^- are typically less than 1/10th of those in the ECF.

If it wasn't for the active pumping of Na^+ and K^+ the system would be in equilibrium, the because of the activity of the Na^+/K^+-ATPase the ions are instead in a steady state. As we will see, this asymmetrical distribution generates the membrane potential, Vm, which in the cell is -70 mV.

All electrical phenomena in neurons and other cells depend on membrane potentials. These are voltage differences that result from different concentration of dissolved salts (ions) in solution separated by a semipermeable, and therefore selective membrane. The voltage difference, Vm, is proportional to the ratio of the concentration difference on either side of the membrane and is described by the Nernst equation for Na^+ as:

Vm = $(RT/zF) \times \ln (Na_{out})/(Na^+_{in})$, where

R = gas constant = 8.3144 JK/mole

T = temperature in degrees Kelvin,

°C = 273

z = valance (charge) on the ion

F = Faraday's constant = 96500 coulombs/gram - equivalent charge

Note: In = natural log; to convert to $\log_{10}$ multiply RT/zF by 2.303

The Nernst equation only applies when one ionic species is responsible for determining Vm. However, in reality all cell

membranes are permeable to varying degrees of Several types of ions, and all permeating ions can contribute to setting the voltage difference across the membrane. In this situation the Goldman-Hodgkin-Katz (GHK) equation is more appropriate and should be used to calculate an estimate of Vm.

Even though a mebrane can be permeable to more than one ionic species. This does not mean it is equally permeable to all ions, rather, real cell membrane and artificial membrane are selectively permeable. The GHK equation therefore includes a term NNa, Pk etc. which is the permeability constant for each ionic species.

We can conclude that "If the permeability of the piasmamembrane for any of the major inorganic ions suddenly is increased, the affected ions will pass through the plasma-membrane, driven by its electromotive force, towards the equilibrium potential for that ion. The result will be an altered membrane potential"- and this is what neurological signalling is all about.

The changes in permeability of the plasma-membrane are brought about by ion-specific ion channels (Fig. 157) of all suet channel, the voltage gated Na+ channel is the most well known.

Sodium channel has two gates - an activation gate and an mactivation particle. The Na^+ channel is sensitive to depolarisation.

(a) Before depolarisation, the activation gate is closed, which prevents Na^+ from passing through the channel.

(b) Depolarisation of the plasma membrane allows positive charges to accumulate at the inside of the membrane. This activates the channel by opening the activation gate.

(c) Sodium now flows into the cell, driven bv the electromotive force. The Na^+ channel is selctive for Na^+ and the for the Na^+ from passing through the channel.

(d) With 2 ms of the activation, while the membrane still is depolarised, the motivation particle closes and inactivates he channel. In this state, the channel will not open even if the depolansation continues or is enhanced

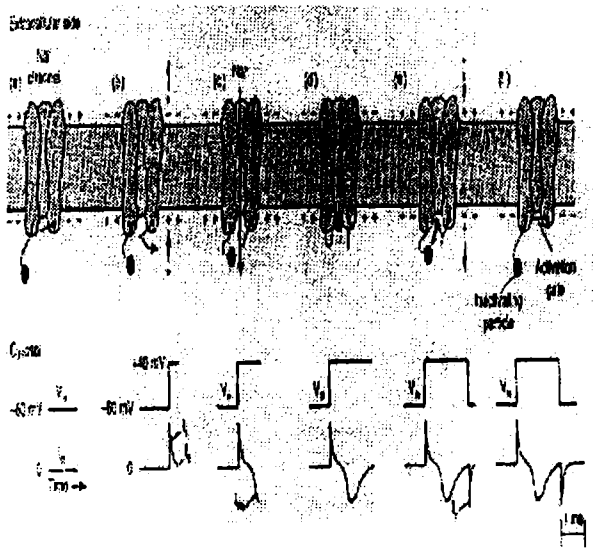


Fig 157: Sodium ion channels in the plasma-membrane.

(e) As the membrane repolarises, negative charges are a-ain accumulating at the inside of the membrane The mactmtion particle moves out of the pore returning the channel m its original closed state in which it is responsive to membrane depolarisation (Fig. 158).

The relationship between the degree of depolarisation and the number of Na^+ channels that activated constitutes the very basis for the generation of an action potential.

A depolarisation of the plasma membrane activates Na+ channels. Na+ enters the cell channels, driven by the electromotive force. Since, the Na^+ ion is positive, the influx of Na+ depolarises the membrane further. A larger depolarisation increases the Na^+ conductivity of the membrane by the opening of more Na^+ channels.

Thus, this system, which is called the Hodgkin cycle, feed back on itself in a positive manner. Na^+ will continue to enter the cell until the inactivation particle start to close and the membrane potential approaches the equilibrium potential for Na^+.

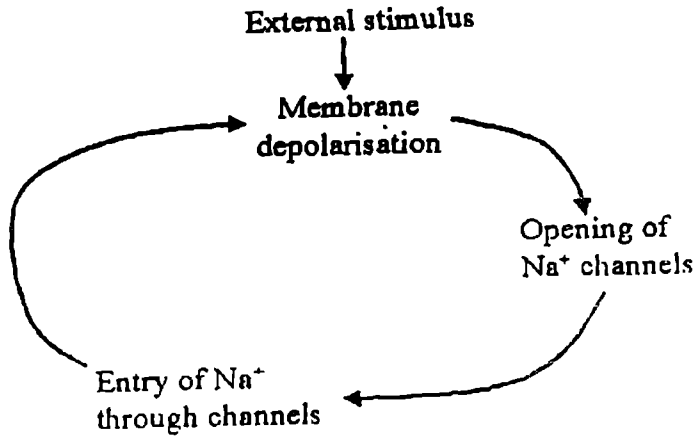


Figure 158: Hodgkin cycle

When a nerve cell membrane is depolarised, a rapid inward movement of ions occur (figure 159). This inward current is the flow of Na^+ into the cell through the Na^+ channels. A few milliseconds later, the current changes direction into an outward flux of ions. This delayed outward current has been identified as an efflux of K^+ and its function is to break the Hodgkin cycle and return the membrane potential to its initial polarised state.

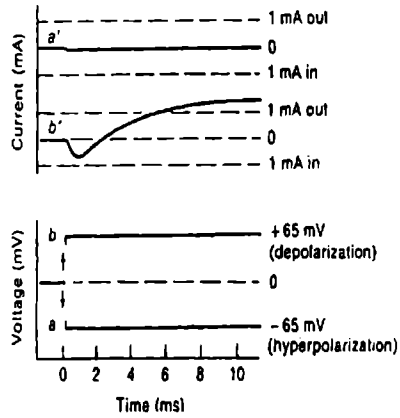
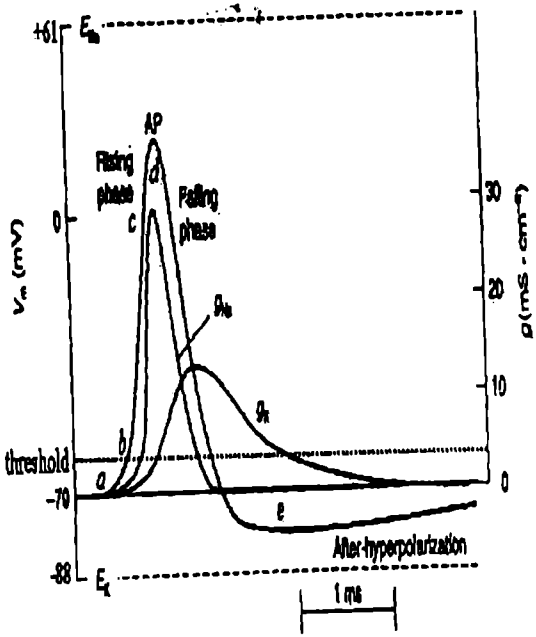


Figure 159: The delayed outward current of a cell membrane.

Action potentials are produced by the plasma membranes of neorons and muscle cells as wells some receptor cells, and sensory cells. They serve two purposes:

1. Rapid propagation of signals over long distances in nerves and muscles.
2. Control of effector responses, such as muscle concentration and activation of hormone secretion.

If a membrane is slightly depolarised by stimulus from an electrode or from another nerve, for example, Na^+ will run into the cell through opened Na^+ channels, but the inward current of Na^+ will be cancelled by a K^+ efflux potential, the conductivity for K^+ is 30-100 times higher than that for Na^+. The result is a local response that dies out (Fig. 160).



Figurs 160: Action potential conductivities of plasma membrane.

However, if the stimulus is strong enough to cause a depolarisation that reaches the action potential the Na^+ influx will exceed the K^+ efflux and the Hodgkin cycle will be initiated. The threshold potential is just the break point where Na^+ influx > K^+ efflux.

Now the Na+ ions that enter the cell with their positive charges recruit the opening of further Na^+ channels, and so it goes until the inside of the cell eventually is positive in comparison with the outside.

Here the delayed outward movement of K^+ exceeds the influx of Na^+ and the membrane potential is driven back towards and transiently even below the resting potential.

Chapter 8

وینہ گرتنی تیشکی-X (ایکس)

X-RAY IMAGING

W.C. Roentgen in 1895 discovered the X-rays when he was studying the phenomenon of discharge of electricity through rarefied gases. He found that when the pressure in the discharge tube was reduced to 0.001 mm of mercury and an electric discharge was passed between cathode and anode, the glass wall behind the cathode began to glow with a greenish yellow colour. During his experiments, he also observed that a fluorescent screen placed close to the discharge tube continued to fluorescent even if the discharge tube was completely covered with a black paper. When a plate of iron was placed, it cast a shadow on the screen, showing that certain radiations are coming out from the discharge tube. Roentgen concluded that when a beam of fast moving electrons strikes a solid target, an invisible high penetrating radiation was produced. Because of their unknown nature Roentgen called these radiations as X-rays. Actually, X-rays are electromagnetic waves of very short wavelengths.

PRODUCTION OF X-RAYS

The X-rays are produced when fast moving electrons strike a target of suitable material. The basic requirements for their production are: (i) a source of electrons, (ii) effective means of accelerating the electrons, and (iii) a target of suitable material of high atomic weight. There are two types of X-ray tubes:

(a) The gas filled or Roentgen X-ray tube, and

(b) The modern coolidge tube or hot filament tube.

The X-ray tube designed by Coolidge (Figure 161) consists of highly evacuated hard glass bulb containing a cathode and anode, the cathode consists of a tungsten filament and is heated by passing a current through it from a low tension battery of 5-10 volts or a separate tapping from a transformer. The electrons are emitted by the process of thermionic emission from the cathode. The filament is surrounded by a molybdenum cylinder kept at a negative potential to the filament. Hence, the electrons emitted from the filament are collimated into a fine pencil of electronic beam.

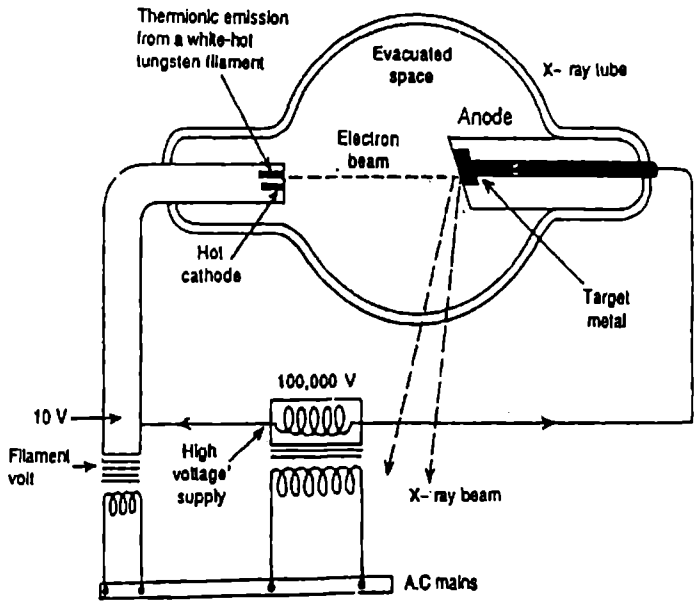


Figure 161: Schematic diagram of a common X-ray tube

The target metal consists of copper block in which a piece of tungsten or molybdenum is fitted. The anode should have the following characteristics:

(i) high atomic weight - to reduce hard X-rays,

(ii) high melting point - so that it is not melted due to the bombardment of fast moving electrons which cause lot of heat generation,

(iii) high thermal conductivity - to carry away the generated heat.

The target metal is placed at an angle of 45° with the path of electron beam. The target metal is cooled by flowing water into hollow tube attached with the target. A high alternating current potential of about 20,000 volts is applied between filament and the target. This is obtained with the help of a step up transformer. Due to this high potential difference, the electrons emitted from the filament are accelerated. When these accelerated electrons strike the target, they give up their kinetic energy and theory produce X-rays, only a small percentage of the electrons at the target are converted into x-rays, while the rest is dissipated as heat. Due to this heating the target metal is heated. In order to save the target it is constantly cooled by cooling arrangement.

The intensity of X-rays depends upon the number of electrons which strike the target, or the rate of emission of electrons from the filament, or thermionic emission. This can be controlled by varying the filament current with the help of a rheostat included in filament circuit.-

The quality of X-ray, which is measured by their penetrating power is a function of potential difference between cathode and target. Higher is the accelerating voltage, higher is the speed of striking electrons and consequently more penetrating X-rays are produced. High penetrating X-rays are termed as *hard X-rays,* while low penetrating X-rays are termed *soft X-rays.*

X-RAY IMAGING

X-ray imaging is one of the fastest and easiest way for a physician to view the internal organs and structure of the body. X-ray imaging has been available for 100 years and is an excellent tool for assessing broken bones, for diagnosing the gastro-intestinal system (digestive tract), for high resolution diagnostic imaging of the breasts (mammography), and for comprehensive imaging of the thoracic cavity including lungs and heart. A host of other applications for X-ray imaging are also available including imaging

the kidneys, teeth, and jaws; the fine structure of the ear, nose, and throat. X-ray imaging is used to detect and guide breast cancer. X- ray imaging is also an important part of bone density measurement.

An X-ray source is turned on, and X-rays are radiated through the body part of interest and onto a film cassette positioned under or behind the body part. A special phosphor coating inside the cassette glows and exposes the film. The resulting film is then developed much like a regular photographic film. It is the special energy and wavelength of the X-rays which allows them to pass through the body part and create the image of the internal structures like bones of the hand. As the x-rays pass through the hand, for instance, they are weakened by the different density tissues they encounter. Bone is very dense and absorbs or attenuates . or absorbs far less X-rays energy, it is these differences in absorption and the corresponding varying exposure level of the film that creates the image which can show broken bones, clogged blood vessels, cancerous tissues and other abnormalities.

Fluorescopic imaging yields a moving X-ray picture or movie. The physician can watch the fluorescent screen which registers the X-rays and emitted glowing light, and see a moving image of the patient's body (for example the beating heart).

The latest X-ray system have the ability to acquire the radiograph or fluorescent movie using digital acquisitions. This gives the radiologist additional control over image quality and interpretation can allow lower doses of radiation. Digital imaging allows the final image to be networked to various locations for additional consultation or interpretation.

Various types of radiography and fluoroscopy are available to image the anatomy and physiology of a wide variety of organs:

1. Angiography (imaging of the blood vessels)
2. Arthrography (imaging of the joints)
3. Bariun X-ray (imaging of gastro-intestinal tract)
4. Chest film (imaging of the thoracic cavity and heart)
5. Cholangiography (imaging of bile ducts)

6. Cholecystography (imaging of the gall bladder)
7. Dental X-rays (imaging of the teeth and jaw)
8. Lyraphangiography (imaging of the lungs)
9. Mammography (imaging of the breast)
10. Myelography (imaging of the spinal cord)
11. Pyelography (imaging of the urinary tract)
12. Skeletal X-rays (imaging of bones)
13. Urography (imaging of the kidney and bladder).

ADVERSE EFFECTS OF X-RAY DIAGNOSIS

The use of X– rays for diagnosis introduces a serious question of the extent of damage done by the rays absorbed. Within a few years after 1895, many effects of X-rays on adults human beings were observed, and other adverse effects were imagined and foreseen, the early workers and their patients, suffered from skin burns, radiation sickness, warts, deformed fingers, loss of hair, and finally the onset of various forms of cancer. Although immediately measurable damage appears only if the dose is hundreds of times higher, more subtle effects, such as malignant growths may show up years or even generations later if the greatest caution is not exercised. The effects of absorbed X– ray radiation dose can be cumulative.

X-RAY SAFETY AND RISKS

X-rays are a type of electromagnetic radiation, are invisible and create no sensation when they pass through the body* Special care is taken during X-ray examinations to ensure maximum safety for the patient. Women should always inform the radiologist if there is any possibility of being pregnant.

During an X-ray, patient may wear a lead apron which will shield the other parts of their body (not being imaged) from radiation. Head is a very dense material and can absorb 100 percent of the X-rays passing through it. X-ray technicians and radiologists are particularly sensitive to the hazards of constant exposure to X-rays and wear special sensitized film badges which monitor the amount of radiation they receive over time.

Modern, state-of-the-X-ray systems have very tightly controlled X-ray beams with significant filtration and X-ray dose control methods. Thus, scattered or stray radiation is minimized and those parts of a patient's body not being imaged receive minimal exposure.

Chapter 9

X-RAY DIFFRACTION IMAGING

X-ray diffraction imaging provides a method for lensless, diffraction-limited, aberration free X-ray imaging of nano-objects in three-dimensions at high resolution.

The rapid growth of nano-science has produced an urgent need for techniques capable of revealing the internal structure in three dimensions, of inorganic nano-structures and large: molecules which cannot be crystallized (such as membrane proteins of vital importance for drug delivery). The development of this method would have a decisive impact on several fields of science.

THE TECHNIQUE

This technique is based on the diffraction of radiation when they encounter small obstacles. If a ray of white light (wavelength averaging 0.5 μm impinge upon a diffraction grating that has 1000 times per millimetre 1 μm spacing). It will be diffracted and will show the various bands of the spectrum. If the wavelength of the light is known the spacing can be calculated from the diffracted angles, and vice- versa. This type of grating would be too wide for X-rays and no diffraction would be produced.

Grating of much smaller dimension, such as those found in natural crystals, would be necessary for the diffraction of X-rays. The atoms, ions or molecules in crystals constitute a true lattice of molecular dimensions capable of diffracting radiations of this wavelength. This technique has its widest application in the study

of inorganic and organic crystals, in which it is possible to determine the precise spatial relationship between the constituent atoms. An analysis of the structure of complex organic molecules, such as proteins and nucleic acids, is much more difficult because of the great number of atoms involved in a single molecule and the irregularities in three-dimensional architecture that most of these large and complex molecules have. However, this configuration of molecules is so vitally important to the understanding of biological function. The study of molecular structure, such as of haemoglobin, myoglobin, DNA and collagen, has been of fundamental importance in the development of molecular biology.

X-ray spectrophotometer are either *rotating crystal type,* or *powder diffraction type.* In the rotating crystal assembly, a single crystal is rotated about one of its axes when a monochromatic X-radiation is incident on it (Figure 162). By rotating the crystal about

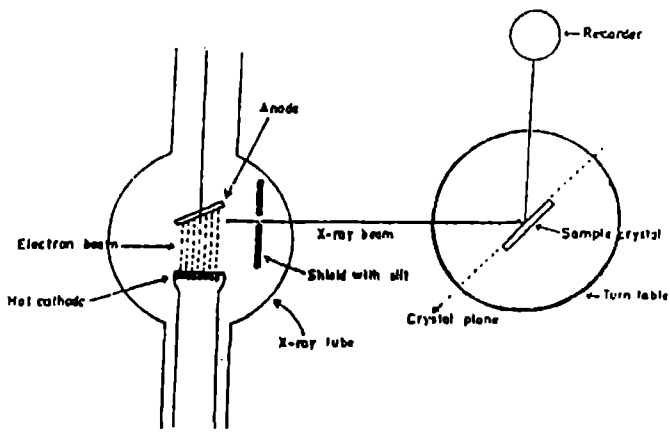


Figure 162: Components of a X-ray spectrometer

various axes the three unit cell dimensions are obtained. In the powder type assembly, the crystal is replaced by the powder of the sample which contains a large number of very small randomly oriented crystals. Since the powder contains crystals of all the planes, there is no need to rotate the powder assembly (Figure 163)

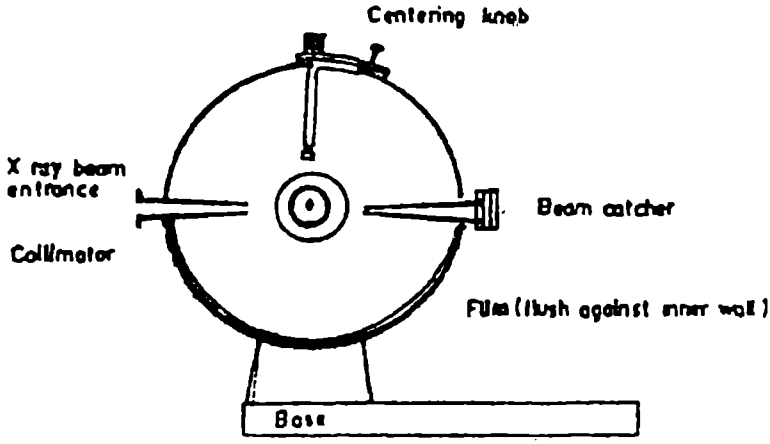


Figure 163: X-ray powder diffraction assembly

In the rotating crystal assembly, the crystal is generally affixed to a thin glass capillary which, in turn, is fastened to a brass pin. This assembly is mounted on a turn tube which is rotated gradually to change the glancing angle of the x-rays incident on a crystal. The X-rays reflected from the crystal fall on a detector. The detector may be a photographic plate or a diffractometer.

In the powder assembly, the powder of the sample is taken in a test tube and a continuous cone of refracted rays are produced. The X-rays diffraction assembly is circular in shape. The film is in the form of a circular arc flush against the inner circumference of the assembly.

The process of determining a crystal structure can be divided into two parts: *experimental* and *computational.* The first involves choosing a crystal, determining its lattice geometry and symmetry, and measuring the relative intensities of a large number of diffracted rays, The second transforms the crude data to a representation of the electronic distribution in the crystal from which the desired information about the molecules can be deducted.

The scanning power of macromolecules such as haemoglobin, myoglobin and DNA can be increased by introducing heavy atoms (such as mercury) into known points of the organic molecules.

Thus, it is possible to construct a three-dimensional representation of the object.

APPLICATIONS

The applications of X-ray diffraction imaging include:

1. The visualization of the internal labyrinth structure of the new mesoporous framework structures,
2. Imaging the complex tangles of dislocation lines which are responsible for work hardening.
3. Imaging the castles within duplex steels, responsible for their very high uniform extension.
4. Three-dimension imaging defect structures in magnetic multilayers.
5. The tomographic imaging of misfit dislocations at interfaces, free of the thin-film elastic relaxation process which distort the images obtained by transmission electron microscopy.
6. Imaging of the three-dimensional arrangement of Orowan dislocation loops, which by entanglement with particles provide the dispersion-hardening of copper alloys.
7. The imaging of precipitates in metal-matrix composite materials.
8. The imaging of electronic device elements for future computing schemes, such as quantum computing.

X-ray diffraction is one of the most important tools in molecular biology because it permits the biologists to determine not only the orientation of the molecules, but also the exact distances that separate them and even to recognise their atomic orientation. The technique has the ability to determine the internal structure of assemblies of macro-molecules, protein complexes, and virus particles at a resolution sufficient to recognise known proteins and determine their ralation-ships to each other.

Chapter 10

ULTRASOUND IMAGING

The term *"ultrasound waves"* or *"supersonic waves"* denotes mechanical vibrations at frequencies above the limit of human audibility, that is, from about 16 KHz to about 10 MHz, "The wavelengths of ultrasonic waves are very small as compared to audible sound. They require an elastic medium for transmission. They are generated by stressing the medium. Their velocity depends on the medium through which they propagate (Table 16 and 17).

Table 16. Main difference between ultrasound and X-rays

	Ultrasound	*X-rays (radiology)*
wave type	longitudinal mechanical waves	electromagnetic waves
transmission requirements	elastic medium	No medium
generation	stressing the medium	accelerating electric charges
velocity	depends on the medium through which it propagates	It is relatively constant: (299,792.456.2 m/s)
similar waves	seismic, acoustic	radio, light

Table 17. Velocity of sound in some biological materials.

Velocity of sound in some Biological Materials

Material	*Velocity of Sound (m/s)*	*Impedance (Rayl x 10-6)*
Air	330	0.0004
Fat	1450	1.38
Water	1480	1.48
Average Human	1540	1.63
Soft Tissue		
Brain	1540	NA
Liver	1550	1.65
Kidney	1560	1.62
Blood	1570	1.61
Muscle	1580	1.7
Lens of eye	1620	N A
Skull Bone	4080	7.8

PRODUCTION OF ULTRASONIC WAVES

The ultrasonic waves cannot be produced by our usual method of a diaphragm loudspeaker fed with alternating current. This is due to the fact that at very high frequencies the inductive effect of the loudspeaker coil is so large that practically no current passes through it. Moreover, the diagram of a loudspeaker cannot vibrate at such high frequencies. Hence, other materials are used for the production of ultrasonic waves. There are two important methods namely *magneto-striction* (Figure 164) and *piezo electric method* (Figure 165), which are mostly used. Magnetostriction method is used when frequencies upto 100 KHz are needed while piezo electric generators are used mostly for frequencies above that.

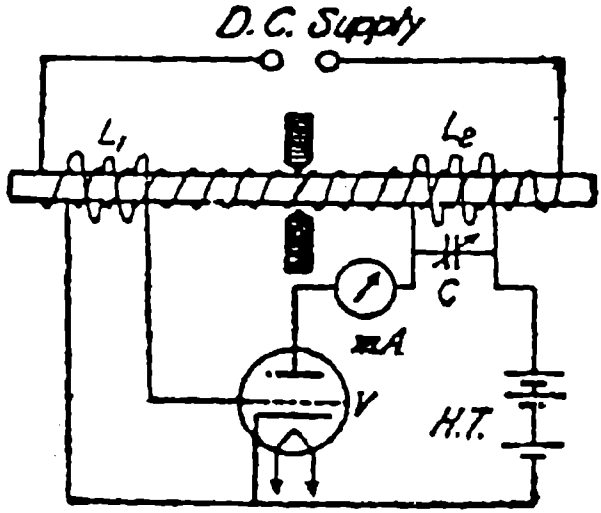


Figure 164: Manetostriction assemble for production of ultrasound waves

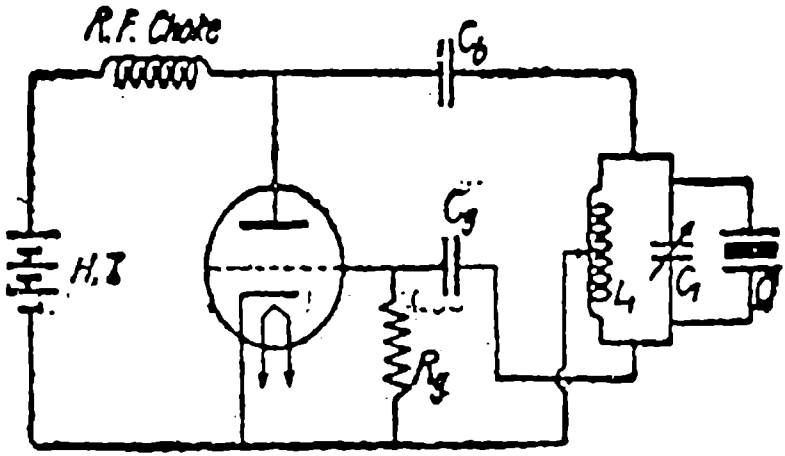


Figure 165: Piezo electric assembly for production of ultrasound waves.

A transducer converts electrical energy into ultrasonic energy. In the diagnostic ultrasound examination, very high frequency sound is directed into the body from a transponder placed in contact with the skin. As the sound travels through the body, it is reflected by tissue interfaces to produce echoes, which are picked up by the same transponder and converted into an electrical signal.

Air base and other heavily calcified materials absorb nearly all the ultrasonic beams, ultrasonic waves plays little part in the diagnosis of lung or bone disease. It is more often used to detect whether a structure is solid or cystic. Cyst or other fluid filled structures produce large echoes from their walls but no echoes from fluid contained between them. Also the tissue behind the cyst and the effect is known as *"acoustic enhancement"*. Conversely in a catalysed structure e.g.., a gall stone, there is a great reduction in the sound, that will pass through, so a band of reduced echoes referred to as an *"acoustic shadow"* is seen behind the stone.

Ultrasound is produced by causing a special crystal to oscillate at a pre-detected frequency. Very short pulses of sound lasting about a millionth of a second are transmitted about 500 times each second. The crystal not only transmits the sound but also listens to the returning echoes which are electronically amplified to be recorded as signals on a cathode ray oscilloscope. Photographic reproduction of image on oscilloscope can provide a permanent record.

The time taken for each echo to return to the transducer is α to distance travelled. Knowledge of depth of the interface responsible for echoes allows an image to be produced. Also by knowing the velocity of sound in tissues, it is possible to measure the distance between interfaces.

The equipment consists of a transducer and a monitoring system. The transducer is a small, hand held device that resembles a microphone. A lubricating gel is spread on the body where the organs to be sonographed are located, and then it is pressed firmly against the skin. The ultrasound image is immediately visible on a nearby screen that looks much like a computer or television monitor.

APPLICATIONS OF ULTRASOUND

Most of the applications of ultrasound waves have been possible on account of their small wavelengths. Although ultrasound is a relatively new energy source, its versatility has led to its widespread use in various industrial, medical, scientific, and consumer applications. Industrial ultrasound is used for cleaning and welding plastics, drying fine powders, emulsification, detecting flaws in materials, and non-destructive testing. The use of ultrasonic devices in medicine includes both, therapy (surgery) and diagnosis (image detection). Ultrasound sensor remotely detects and Images concealed weapons. The breadboard sensor can detect metallic and non-metallic weapons concealed on a human body under heavy clothing at ranges upto 8 m and can image concealed weapons at ranges upto 5 m. The commercially available knife made of hard lexan plastic passes through an airport metal detector with no chance of detection.

ULTRASOUND SCANNERS

Ultrasound scanners consists of a console containing a computer and electronics, a video display screen and a transducer that is used to scan the body. The transducer is a small hand-held device about the size of a bar of soap, attached to the scanner by a cord. The ultrasound image is visible on the screen that looks much like a computer or television monitor. The physician or technologist watches this screen during an examination and captures representative images for storage, often the patient is able to see it as well. An example of the ultrasound imaging equipment has been shown in Figure 166.

ULTRASOUND IMAGINGS

Ultrasound imaging is based on the same principle as the sonar used by bats, ships at sea, and anglers with fish detectors. As a controlled sound wave bounces against objects, its echoing waves can be used to identify how far away the object is, how large it is, and how dense is it?

Ultrasound imaging (also called ultrasound scanning or sonography) is a relatively inexpensive, fast and radiation-free imaging modality. It is a method of obtaining images from inside

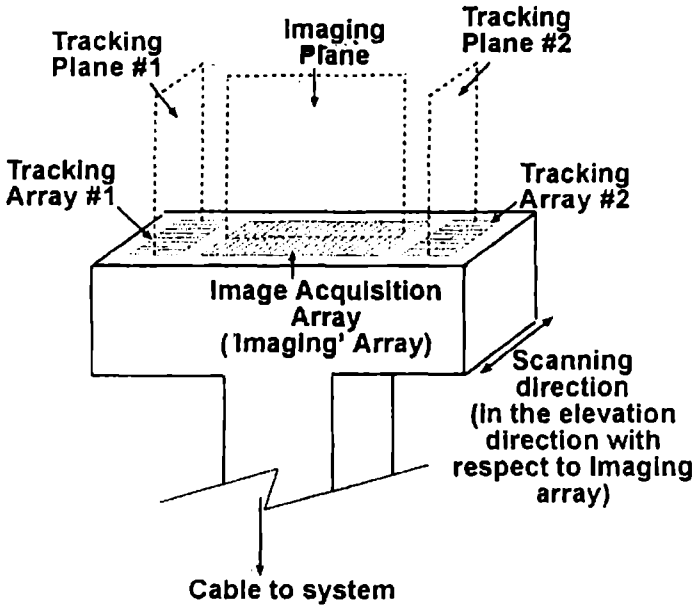


Figure 166: Ultrasound imaging equipment.

the human body through the use of high frequency sound waves. No ionising radiation (X-ray) is involved in ultrasound imaging. Because high frequency sound waves cannot penetrate bone or air, they are especially useful in imaging soft tissues and fluid filled spaces. Ultrasound is good at non-invasively imaging a number of soft tissue organs such as heart, pelvis and reproductive organs/ kidneys, liver, pancreas, gall bladder, eye, thyroid, blood vessels, and fetus.

Ultrasound imaging involve the transmission of high frequency sound waves Approximately over 20,000 Hz). Defined bursts of ultrasonic waves are emitted from the transducer. The ultrasound transducer combines functions a stereo loudspeaker and a microphone in one device; it can transmit sound and receive sound. Ultrasound waves are aimed through a body part and conversion of echoes into electrical impulses, which are displayed as a pattern on an oscilloscope screen. These waves are received

by the ultrasound machine and turned into live pictures with the use of computers and reconstruction software.

The purpose of ultrasound Imaging is to detect and evaluate lesions, venases insufficiency, thrombases, cholelithiasis, jaundice, tumours, abscesses, cysts, enlarged lymph nodes, cataracts, pancreatitis, carcinoma, foreign bodies, and to confirm single or multiple pregnancies and condition of the fetus. Ultrasound Imaging and ultrasound angiography are finding a greater role in the detection, diagnosis and treatment of heart disease, heart attack, acute stroke. Ultrasound can also be used to guide fine needle, tissue biopsy to facilitate sampling cells from an organ for laboratory testing.

Most ultrasound examinations are similar, involving the following steps:

1. Patient preparation involves removing any articles of clothing or jewellery surrounding the area to be imaged. In some cases, the patient may be asked to wear a patient gown.
2. The patient is positioned by the technologist on an examination table. A clear gel (which helps "connect" the ultrasound transducer to the skin) is applied to the area to be examined for example the abdomen.
3. The technologist then brings the transducer into contact with the skin and sweeps it back and forth to image the area of interest (e.g., the fetal baby). The patient is simply required to relax and stay calm during the examination.
4. The technologist will ask the patient to get dressed and wait while the ultrasound images are reviewed, either on film or a TV monitor. In many cases, the technologist or physician reviews the ultrasound images in real time as they are acquired.
5. After the ultrasound images are reviewed, the patient will be released from the imaging department or centre, in some cases, more images will need to be taken.
6. Other preparations depends on the type of examination you will have. For some scans, you may be instructed not to eat or drink for as many as 12 hours before your appointment. For still others, you may be asked to drink upto six glasses of water 2 hours prior to examination and avoid urinating, so that your bladder is full when the scan begins.

RADIOGRAPHIC FILM

Descriptive terms are used in ultrasound imaging. Tissues that strongly reflect ultrasound are *hyperechoic* or of increased echogenicity. Poorly reflecting tissues are *hypoechoic,* while fluid, which does not reflect sound, is *anechoic* or *sonolucent.* Tissue behind am area of sonolucency will appear hyperechoic because of acoustic enhancement. On the other hand, through-transmission of ultrasound beam will be blocked by a strongly hyperechoic object (such as a rib); thus, an acoustic shadow (where no images appear) is cast behind the object.

ADVANTAGES OF ULTRASOUND IMAGING

There are many advantages to imaging the body with ultrasound. Most importantly, there is no ionizing radiation as with X-rays, so that ultrasound is used extensively during pregnancy. Furthermore, soft tissues, such as liver, spleen, kidneys and pancreas can be imaged directly without the injection of any sort of radio-opaque substances or isotopes to males them visible, in addition, the entire abdomen and pelvis can be rapidly scanned while the patient is lying on the table.

DRAWBACKS OF ULTRASOUND IMAGING

There are many drawbacks of ultrasound imaging, probably the most serious is the fact that sound is not able to travel through certain organs; their surfaces reflect almost 100 percent of sound waves, so that the interiors of these organs and those lying directly beneath them cannot be imaged, organs filled with air such as the lungs, stomach and intestines are opaque to sound as are hard tissues such as bones.

BENEFITS VS RISKS OF ULTRASOUND SCANNING

Benefits

- Ultrasound scanning is non–invasive (no needles or injections, in most cases) and is usually painless.
- Ultrasound is widely available and easy to use.

- Ultrasound uses no ionizing radiation, and is the preferred image modality for diagnosis and monitoring of pregnant women and their unborn infants.
- Ultrasound provides real-time imaging, making it a good tool for guiding minimally invasive procedures such as needle biopsies.
- Ultrasound images can visualise structure, movement and live function in the body's organs and blood vessels,

Risks

- For standard diagnostic ultrasound there are no known harmful effects on humans.

LIMITATIONS OF GENERAL ULTRASOUND IMAGING

Ultrasound has difficulty in penetrating bone and therefore can only see the outer surface of bony structure and not what lies within. For visualizing bone or internal structure of certain joints, waves do not reflect clearly from bone or air. For visualization of bone, other imaging modalities, such as magnetic resonance imaging (MRI), may be selected.

Chapter 11

COMPUTERIZED AXIAL TOMOGRAPHY (CAT) SCAN

While X-rays provide a flat, two dimensional picture, CAT scanning gives a three-dimensional image. During CAT scanning X-rays are directed at the object from all sides to create images of slices through it. A computer uses these to generate a three dimensional reconstruction of the inside of the object. A CT scan can image many parts of the body but is commonly used to look at the brain, it shows epidural, subdural, and intracerebral haemorrhages and deformities of ventricular system from mass lesions, and demonstrate tumours as well as areas of brain edema and infraction, hydrocephalus and brain atrophy. The simplicity of this non-invasive procedure have virtually revolutionised diagnostic neurology and neurosurgery.

CT SCANNER

The CT scanner consists of an X—ray source (an X-ray tube) mounted opposite an array of detectors. Generally there are 300 to 600 detectors. Both the X-ray tube and the detector bank are mounted on a revolving gantry. The patient being scanned lies in the aperture of the gentry, in this system the X-ray tube and detectors are made to rotate around the part being scanned. The gentry rotates one revolution for each scan, while the X-ray tube active, in CT the X-ray beam is rotated through atleast 180°. Depending upon the scan time (from 2.8 to 9.0 seconds), between 300 to 1000 views are produced. As the X-ray pass through the

tissues, they are dissipated by varying amounts related to the density of the tissues. This raw data is used to produce an image by computer. The result is a cross-sectional image of the part giving excellent special and contrast resolution. Most of the imaging is done in the horizontal (axial) plane, hence the term Computerized Axial Tomography (CAT).

Image intensifiers are used -where there is need to see a time-continuous image, such as during the setting of a fracture. Electrons displaced from the photocathode by incoming X-rays photon (Figure 167) are amplified and focussed on an output phosphor screen, video and film may be recorded from this screen.

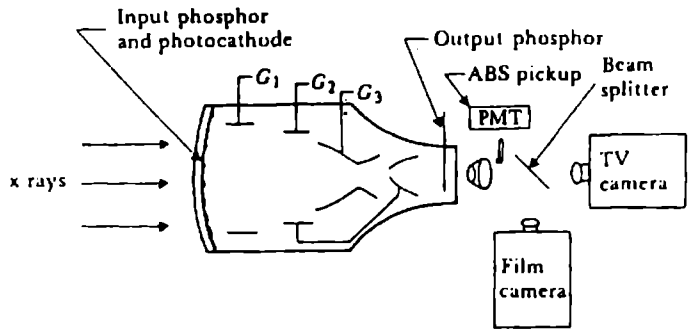


Figure 167: Image intensifier of CT scan.

Radio opaque dyes (iodine compounds) may be injected into the body to enhance different structures. Depending on the compounds with -which the iodine is combined, different structures or organs may be imaged. If the iodine is injected into an aertery and remains in the arterial system, arteries in the brain, heart, kidneys or legs are most commonly imaged. Compounds that, are excreted by the liver or kidneys may be chosen in order to enhance the gall bladder and bile ducts or the kidneys, ureters and bladder.

In 1973 EMI Ltd. Britain marketed the first machine for CAT. The machine was constructed to image cross sections of the body

from X-ray scans (Figure 168). An X-ray source is collimated into a narrow beam and scanned through the plane of interest. The transmitted photons are received by a detector which tracks the scanning motion of the source. After each scan, the source and detector are rotated by 1 degree around the patient. This is repeated until a 180° circuit has been covered. Compared with plain X-rays a high dose of radiation is given to provide data for reconstructing a two-dimensional slice through the patient.

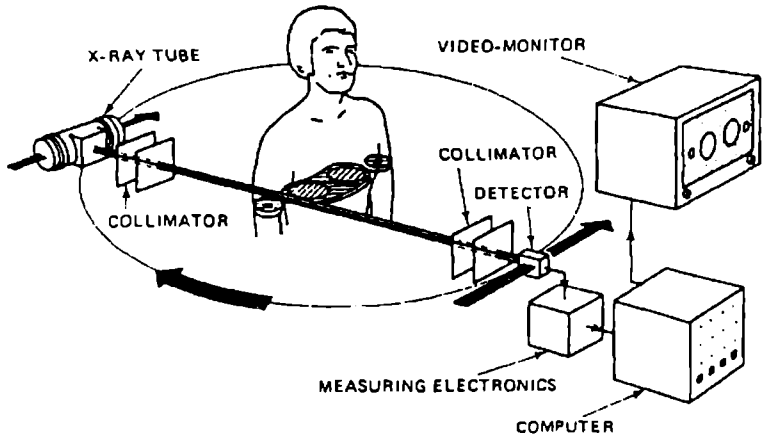


Figure 168: First generation computer tomography.

Since the first generation scanner, three further types have been developed, The first generation scanner takes a few minutes to collect the data. Hence, it can only be used for immobile body parts such as the head. However, it is still popular because of its low cost.

In the second generation scanners, a single source of X-rays and multiple detectors are used. However, rotation is in steps of 10 degrees instead of 1 degree, because the multiple detectors collect more of the source radiation each time.

The third generation scanner has a fan beam and a circular arc of detectors. The fan beam and the detectors rotate about a point in the patient.

The fourth generation scanner has a stationary 360 degree ring of detectors surrounding the patient. A source with a fan beam rotates around the patient, inside the detector ring. This allows a more rapid scan.

The fifth generation scanner allows scanning in milliseconds (Figure 169). It is able to produce images of the beating heart.

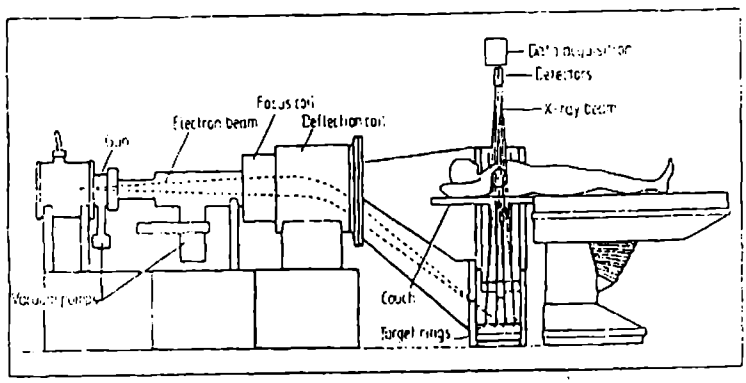


Figure 169: Fifth generation CT scanner.

The sixth generation scanners are now in common use. The patient is moved through the scanner while scanning is in progress. This has enabled a pseudo three dimensional images to be constructed.

CT SCAN PROCEDURE

When a patient arrive, he has to check in with the receptionist so that the radiographer knows that the patient is there, Then the patient usually takes a seat in the waiting room until he/she is called for the scan.

When the patient is called, he/she may first go to a cubicle to take off his/her outer clothing. He/she have to strip down his/her underwear and put on a hospital gown. If the patient is to have a CT of his/her head, he/she may not be asked to undress. The patient must take off any jewellery that is in the area to be scanned because metal interferes with the machine.

When the patient is ready, the radiographer or helper will take him into the scanning room. The patient have to lie down on the machine couch on his/her back. Sometimes the scan is done with the patient with him on his side or lying on his/her front. The patient need to lie as still as he/she can, but breathe normally.

Once the patient is in theright position on the couch, the radiographer will leave the room. This is because there will be X-rays in the room and it would be dangerous for the staff to be exposed to these. They see patients having X-rays and CT scans all day, every day and if they stayed in the room, they would be exposed to far more X-rays than any patient.

The radiographer will be able to see you on a TV screen and you can talk to each other through an intercom. The radiographer will control the position of the couch from outside. The couch can move automatically through the CT scanner so that the part of the body to be scanned is in the machine. The radiographer will tell the patient that he/she is about to start the scan and remind the patient to keep as still as he/she can. When the scan is over, the radiographer will come back into the room and help the patient down from the couch.

Most scans take about half an hour. A lot of that is for setting up the scan, rather than actually taking it. Lying still for that long can be uncomfortable, if the patient is getting stiff and need to move, tell the radiographer through the intercom. During the actual scan, the patient have to try to keep as still as possible, and not cough or swallow, particularly if his/her head is being scanned. Mostly, the patient can breathe quietly but normally throughout the scan. For some scans, the radiographer may ask the patient to hold his/her breath at various times during the scan. If this is going to happen, they will tell the patient beforehand. The patient should be able to go home as soon the scan is over.

SUGGESTIONS FOR CT SCAN

If necessary, the patient can have a tablet or injection to calm him down before the scan. Asking the patient to close his/her eyes sometimes helps. Some CT scans need spatial preparation beforehand. This is explained below for scans of different parts of the body.

ABDOMINAL CT SCANS

If you are having a CT scan of your abdomen, you may be asked

- Not to eat or drink after midnight the night before the scan
- To drink a liquid 'contrast medium' 24 hours before the scan
- To drink more of the liquid in the X-ray department

You may have the contrast medium by injection either instead of, or as well as, the drink. The contrast medium makes the digestive system (gut) show up more clearly in the scan. It does not have any side effects.

CT SCANS OF THE HEAD

For some brain scans, you may be given an injection of the 'contrast medium' dye beforehand to make the scan clearer.

CT SCANS OF THE CHEST

For some chest (thoracic) scans, you may be given an injection of the 'contrast medium' dye beforehand. This is to help show up the tissues in the area containing the cancer, for example blood vessels. It may help to show whether the cancer can be removed with surgery or not.

PELVIC CT SCANS

If you are having a CT scan of the pelvis, you may be asked

- Not to eat or drink after midnight the night before the scan
- Have an injection of 'contrast medium' just before the scan

Depending on the part of your pelvis being scanned, you may have an injection of a drug to slow down the normal movement of your bowel. This movement (called 'peristalsis') can distort the scan and make it more difficult to read.

Occasionally, for a rectal scan, you need to have an enema of the 'contrast medium' dye. This shows up on the X- ray and makes the outline of the bowel stand out more on the scan. This 'rectal contrast' isn't used very often. It may make you constipated. Your first couple of bowel motions will be white, but there are no other side effects.

There is a very detailed scan of the bowel called a 'virtual colonoscopy'. If you are having one of these, you will be asked

- Not to eat or drink for 36 hours before the scan
- Take 2 doses of a strong laxative the day before the scan

If you are diabetic, your doctor may want you to come into hospital for the day or two before this. It may not be sensible for you to avoid eating and drinking for 36 hours without medical supervision.

Just before the scan, you'll have two injections. One is the 'contrast medium' dye to show up the body tissues more clearly. The other is a drug to slow down the normal movement of your bowel which can make the scan less clear. You will also have a tube put into your back passage. The radiographers will put air through this tube to inflate your bowel and make the scan clearer. Apart from the obvious and slightly embarrassing after effects of having air pumped into your bowel, there are no other side effects of this. Try to take it all in your stride. The staff are professionals and are used to doing this type of test. They won't be embarrassed by it so there is no need for you to be.

For virtual colonoscopy, you will have two scans - one on your back and one on your front.

CAN A CT SCAN BE DANGEROUS?

Like any X-ray, you should not have a CT scan if you are pregnant as it could be dangerous for the baby. Other than this situation, a CT scan is not dangerous. It involves being exposed to a small amount of radiation, but this is less than you would be exposed to in an aeroplane flight across the atlantic.

Very, very rarely, someone has an allergic reaction to the contrast injection. The reaction most often starts with weakness, sweating and difficulty breathing. It is possible to react to any injection in this way, and the doctors and radiographers will know what to do if you do have this type of reaction.

Chapter 12

هێڵکاری کارەبایی دڵ

ELECTROCARDIOGRAPHY (ECG)

Electrocardiography is a commonly used, non-invasive procedure for recording electrical changes in the heart. An electrocardiogram (ECG or EKG), shows the series of waves that relate to the electrical impulses which occur during each beat of the heart. The results are printed on paper or displayed on a monitor. The waves in a normal record are P, Q, R, S, and T and follow in alphabetical order. The number of waves may vary, and other waves may be present.

The heart is a muscular pump made up of four chambers (Figure 170). The two upper chambers are called atria, and the two lower chambers are called ventricles. A natural electrical system causes the heart muscle to contract and pump blood through the heart to the lungs and rest of the body.

When the heart contracts, electrical current is produced and distributed throughout the body to the skin, just like the spreading waves of a pool of water into which a stone has been dropped. The electrocardiography is a starting point for detecting many cardiac problems. It is used routinely in physical examinations and for monitoring the patients condition during and after surgery, as well as during intensive care, it is used to evaluate causes of symptoms such as chest pain, shortness of breath, and palpitations.

The electrocardiogram is a graphic recording of the electrical activity of the heart as detected on the body surface by a group of electrodes. The electrodes are small metal discs, attached to the skin on the chest, arras, and legs. The electrodes are also

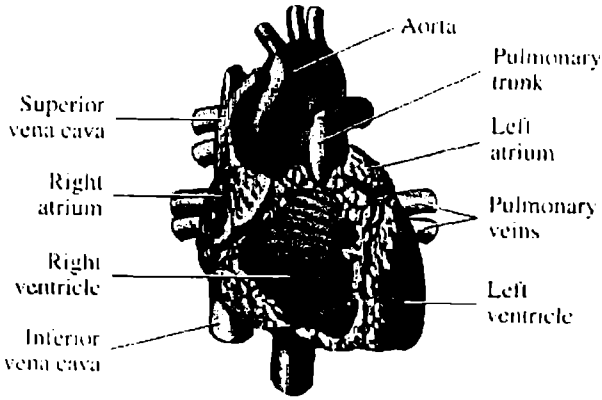


Figure 170: Showing the heart is a muscular pump made up of four chambers.

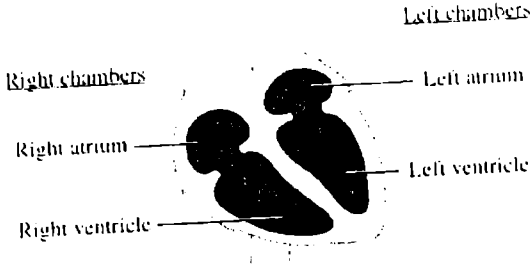


Figure 171: Diagrammatic sketch showing four chambers of the heart.

connected to a machine that translates the electrical activity into line tracings on paper. Most cardiac cells maintain a resting membrane polarization of 90 mV, the inside of the cell in negative with respect to the outside, but the cells are electrically active, and a sufficient stimuli is able to initiate a depolarization, cardiac tissue is unique among other electrically active tissues in its significant time delay excitability or repolarization. When an excitatory

process flows towards a unipolar electrode, a positive or upward detection is produced, and when it flows from the electrode, a negative or downward deflection is produced (Figure 172).

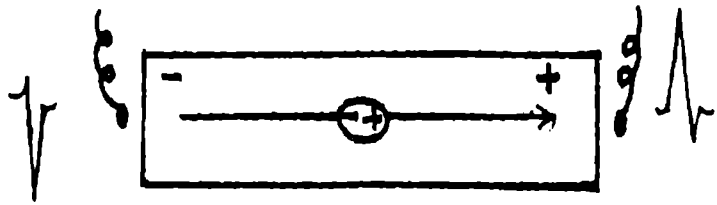


Figure 172: The electrical behaviour of a single cell or strip of muscle, indicated by a rectangle from the two ends of which the potentials are led off to a galvanometer.

The ECG lead system is composed of five electrodes, one on each of the four limbs and one placed at various sites on the pericardium. Each lead is a continuous recording of the change in electrical potential during the cardiac cycle between two of the electrodes, or between one electrode and a combination of the others. The right leg electrode is an inactive ground electrode in all leads.

The term lead is used to denote the connection of the galvanometer by wires to the electrodes and also for the actual tracing obtained. Two types of leads are usually recorded - *bipolar* and *unipolar.*

Originally bipolar standard leads were selected by Einthoven to record the electrical pattern in the frontal plane based on K law which states that, "the algebraic sum of all patterns difference in a closed circuit equals zero". The electrodes in their case were applied to right arm, left leg, and left arm (Figure 173).

Lead I is obtained by applying the electrodes to left and right arm, lead II connects the left leg to right arm, and lead III connects the left leg to left arm. The relationship between the three leads can be expressed by the equation Lead II = lead I + lead III. It

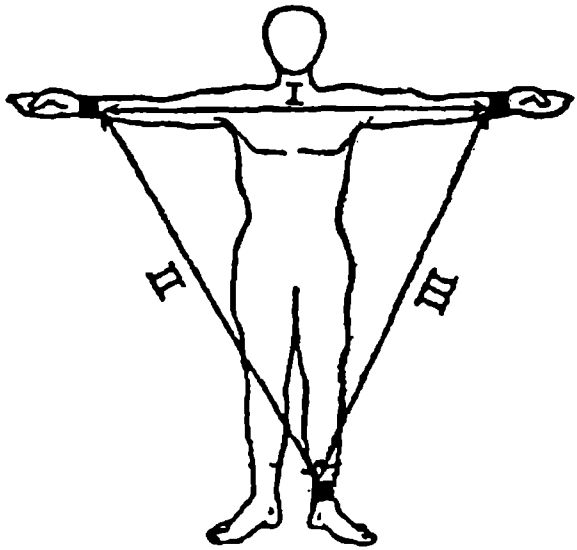


Figure 173: Connections of standard limb leads.

other words one can reconstruct lead II from the algebraic sum of the waves of I and III.

Unipolar leads are obtained by placing one electrode (exploring electrode) in close proximity to the heart, while the other indifferent electrode is far removed from it so that its potential is more or less reduced to zero. There are two types of unipolar leads - unipolar limb leads and unipolar chest leads.

In unipolar limb leads one electrode in placed in turn over one of the three extremities used in recording standard leads, while the other is connected to the central terminal. With this technique however, the amplitude of the deflection is so small that their interpretation is difficult. The amplitude can be increased by 50 percent if the exploring electrode is attached to either the right arm, left arm or left leg and connected to one pole of the galvanometer, whilst the other pole is connected to an indifferent lead points - a central terminal connecting the remaining three extremities not being explored (Fig. 174).

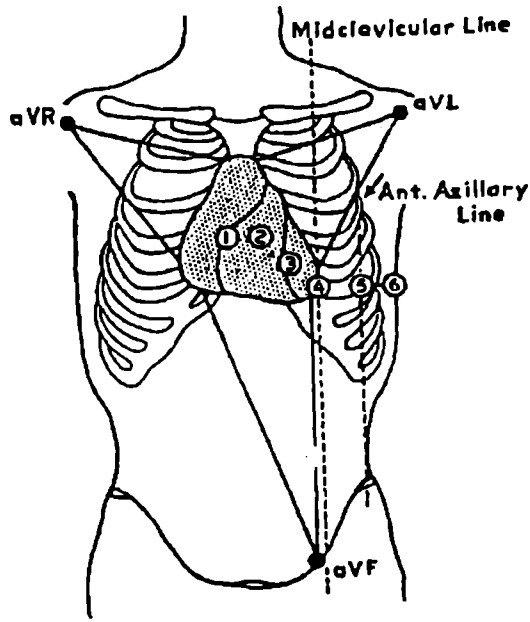


Figure 174: Diagram showing unipolar limb leads and position of pericardial leads.

In unipolar pericardial or chest leads, direct leads from the various points on the heart itself present the most detailed information regarding the spread of the excitation wave though the heart.

The normal ECG deflections are designated by the letters P, Q, R, S, T. The P wave signifies arterial depolarization and it is normally an upright small round deflection. It precedes QRS complex due to electrical stimulation of the ventricles. The Q wave is a negative deflection which follows the R wave. The R wave is an initial upward deflection. S wave downward deflection following R. The T wave signifies ventricular polarization and normally it is upright. U wave is a small upright rounded deflection following T wave, its pathogenesis is uncertain.

A positive or negative deflection is called *monophasic.* A positive and negative deflection is called *diphasic.*

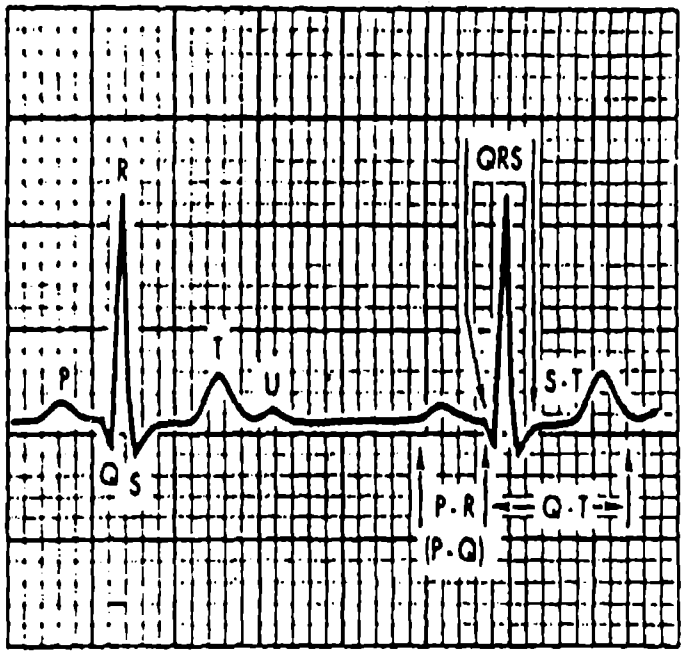


Figure 175: The waves of the electrocardiogram. P, Q, R, S, T and U are indicated. The measurements of the P-R interval, QRS complex, S-T segment and QT interval are identified on the right.

The ECG paper is a specially prepared paper. The horizontal and vertical lines are marked on all papers. The vertical lines represent time and are divided into large and small squares. Every fifth line is marked deeper than the others. Each large square indicates 0.2 seconds and each small square is one-fifth of this i.e., 0.04 seconds. The horizontal lines represent voltage, 1 mm being equal to 0.1 millivolt when the instrument is properly standardized. The ECG paper moves at a rate such that in 1 minute 300 large square or 1500 small square are covered. Hence, to calculate the heart rate, one must divide 1500 by the number of small squares between the two heart beats.

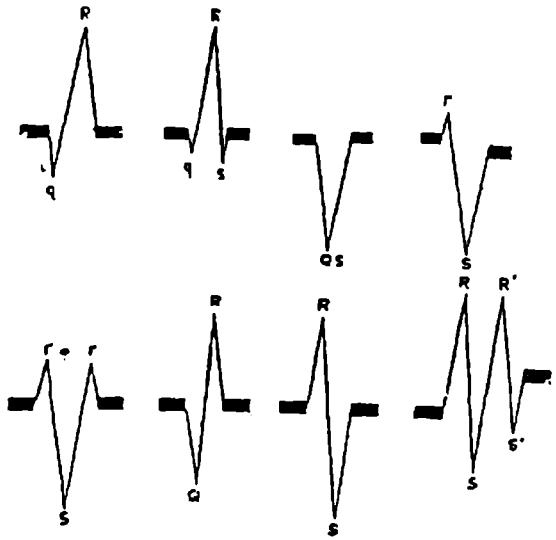


Figure 176: Diagram of various configurations of QRS complex. Capital letters are used to describe the relatively large waves and small letters for relatively small waves.

An electrocardiogram may show:

- Evidence of heart enlargement
- Signs of insufficient blood flow to the heart.
- Signs of a new or previous injury to the heart.
- Heart rhythm problems.
- Changes in the electrical activity of the hearty caused by an electrolyte imbalance in the body.
- Signs of inflammation of the sac surrounding the heart.

However, electrocardiography cannot predict whether a person will have a heart attack.

WHY ELECTROCARDIOGRAPHY IS DONE ?

Electrocardiography is done to:

1. Evaluate unexplained chest pain, especially when a heart attack is possibility, identified by a irregular heart beat (arrhythmias), a heart chamber with thickened walls (hypertrophy), inflammation of the surrounding the heart (pericarditis), and reduced blood flow to the heart muscle (ischemia).
2. Monitor the electrical activity of heart.
3. Determine the thickening of the walls of ventricle.
4. Monitor the effectiveness and possible side effects of certain modifications that may affect the heart.
5. Check the function of mechanical devices implanted into the heart (pacemakers or defibrillators).

HOW ELECTROCARDIOGRAPHY IS DONE?

An electrocardiogram is usually done by a nurse, ad the resulting ECG is interpreted by a doctor.

You may receive an ECG as part of a physical examination by your doctor during a series of tests. ECG equipment is usually portable, so the test can be done almost anywhere. It is also possible to continuously monitor the ECG, this process is called telemetry.

You should remove all jewellery from your neck, arms, and wrists. You will also need to remove your clothing above the waist and keep your forearms and lower legs exposed. You will be given a cloth or paper covering to use during the test.

During electrocardiography you will lie on a bed or table. You may rest for 5 to 15 minutes before the test to assure accurate test results. Areas on your arms, legs, and chest where electrodes will be placed are cleaned and possibly shaved to provide a clean, smooth surface to attach the electrode discs. A special ECG paste or small pads soaked in alcohol will be placed between the electrodes and your skin to improve conduction of the electrical impulses.

Several (usually 12) metal electrodes or "leads" are attached to the skin on each arm and leg and on your chest. After the ECG is complete, the electrode paste is wiped off.

It is important not to move or talk during the ECG recording because muscular activity can cause inaccurate results. For best results, lie very still and breathe normally. Sometimes you may be asked to hold your breath.

An electrocardiogram usually takes 5 to 10 minutes to complete, sometimes a longer period of recording is done to measure your heart's rhythm for a minute or longer.

RISKS ASSOCIATED WITH ECOCARDIOGRAPHY

There are no risk associated with ecocardiography. It is a completely safe test. The electrodes only detect electrical impulses by your heart. No electricity passes through your body from the machine, and there is no danger of getting an electric shock. Electrocardiography tracings show a characteristic pattern of electrical impulses that are generated by the heart. The different parses ox an ECG tracing of a heart beat are called the P wave, the QRS complex, the ST segment, and the T wave,

- The P wave is a record of the movement of electrical activity through the upper heart chambers and is recorded when they contract,
- The QRS complex is a record of the movement of electrical impulses through the lower heart chambers and is recorded when they contract,
- The ST segment usually appears as a straight, level line between the QRS complex and the T wave. Elevated or lowered SB segments may mean the heart muscle is damaged or not receiving enough blood,
- The T wave corresponds to the period when the lower heart chambers are relaxing and preparing for their next muscle contraction.

Chapter 13

د مغز د برقی کارۍ معاینه
ELECTROENCEPHALOGRAPHY (EEG)

The first recording of the electric field of the human brain was made by the German psychiatrist Hans Berger (Berger, 329). The electrical activity show that it may be either *spontaneous activity,* or *evoked activity.*

Spontaneous activity is measured on the scalp or on the brain, the amplitude of spontaneous activity is about 100 μV when measured on the scalp, and about 1-2 mV when measured on the surface of the brain. The bandwidth of this signal is from under 1 Hz to about 50 Hz in a normal EEG (Figure 177) . As the phrase "spontaneous activity" implies, this activity goes on continuously in the living organism.

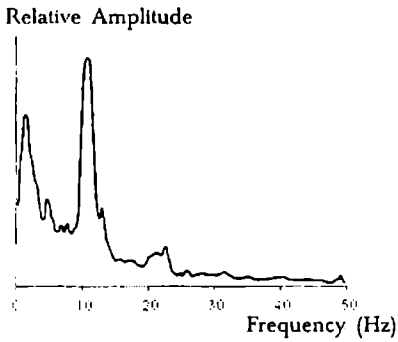


Fig. 177: Frequency spectrum of normal EEG.

The normal electrocardiographic record in adults is usually easy to identify. The frequency pattern shows somewhat asymmetric 8 to 12 Hz, 50 μV sinusoidal *alpha* waves in both occipital and parietal regions. These waves wax and vane spontaneously and disappear promptly when the parent opens his eyes and fixes his attention on something. Faster waves than 13 Hz or lower amplitude (10 to 12 μV), called *beta* waves are seen symmetrically in the frontal region. Very slow *delta* waves or other unusual pattern are absent in a normal record. When the normal subject fills asleep, the rhythm slows symmetrically, and characteristic wave form appear.

Evoked activity are those components of the EEC that arise in response to a stimulus which may be electric, auditory, visual etc.). Such signals are usually below the noise level and thus not readily distinguishable, and one must use a train of stimuli and signals averaging to improve the signals-to-noise ratio.

Noise may lead to *sleep disturbance* or *awakening sleep* disturbance even though it may not lead to full awakening, is frequently cited as the main causes of annoyance. Electroencephalography provides a means of studying the effect of noise on sleep. Depth of sleep is effected by noise, and periods of very deep sleep may be reduced in length by impulsive noise of very short duration, whose intensity is 20 dB greater than that of the background noise. Above 78 dB, sound of no more than 300 milliseconds duration can interrupt deep sleep, and acoustic stimuli (white noise) of short duration cause EEG changes. In the same way, when individuals sleeping naturally are subjected to continuous noise of 70 dB, periods of deep sleep may be appreciably reduced in length. The probability that an individual will be awakened by pulse levels of 40 dB is 5 percent, and this rises to 30 percent for noise levels of 70 dB. Apparently, noise not only affects depth of sleep in the early part of the night, appears to be compensated by a longer period of deep sleep in the second half of the night.

EEG LEAD SYSTEM

The internationally standardized 10–20 system is usually employed to record the spontaneous EEG. In this system 21 electrodes are located on the surface of the scalp. Reference points

are *nasion,* when is at the top of the nose, level with the eyes, and *incon,* which is the bony lump at the base of the skull on the midline at the back of the head. From these points, the skull parameters are measured in the transverse and median planes. In addition to the 21 electrodes, intermediate 10 percent electrode positions are also used.

The *queen Square System* of electrode placement has also been proposed as a standard in recording the pattern of evoked potentials (Blumhardt et al., 1977).

Bipolar and unipolar electrodes have also been used in EEG measurement, in the former method the potential difference between a pair of electrodes is measured, in the latter method the potential of each electrode is compared either to a neutral electrode or to the average of all electrodes.

EEG DIAGNOSIS

EEC diagnosis is a technique for the recording and analysis of gross electrical signals of the brain. In this technique, small pellets of solder, or other metal contact electrodes, preferably non-polarizable (in which the voltage with respect to some reference remains unchanged when current is passed through the electrode). A silver disk coated with a thin layer of silver chloride, which makes contact with the chloride containing body salts, is non-polarizable electrode in EEG work, where the currents are very small (< 10 amp). These electrodes are placed on the symmetrical points of the scalp and fastened there with a binder such as collodion. 12-26 leads cover the scalp in localization experiments, overlying each important lobe of the brain, and even different portions of each lobe. Voltage between these and some reference positions, such as a lead to the every lobe, are fed into standard high gain amplifiers, and traced by pen recorders. 5-6 seconds of recording gives patterns (Figure 178), which quite empirically have been catalogued as coming from normal or diseased tissue, patterns take on an individual vary with the emotional state of the patient. A creative person is Said to have patterns quite different from one who lacks new ideas. However, the fine structure of these waves is not well understood. Recorded spikes are only bout 150 μV high. Characteristic spike of different shapes and frequencies have been named beta, alpha, theta, delta.

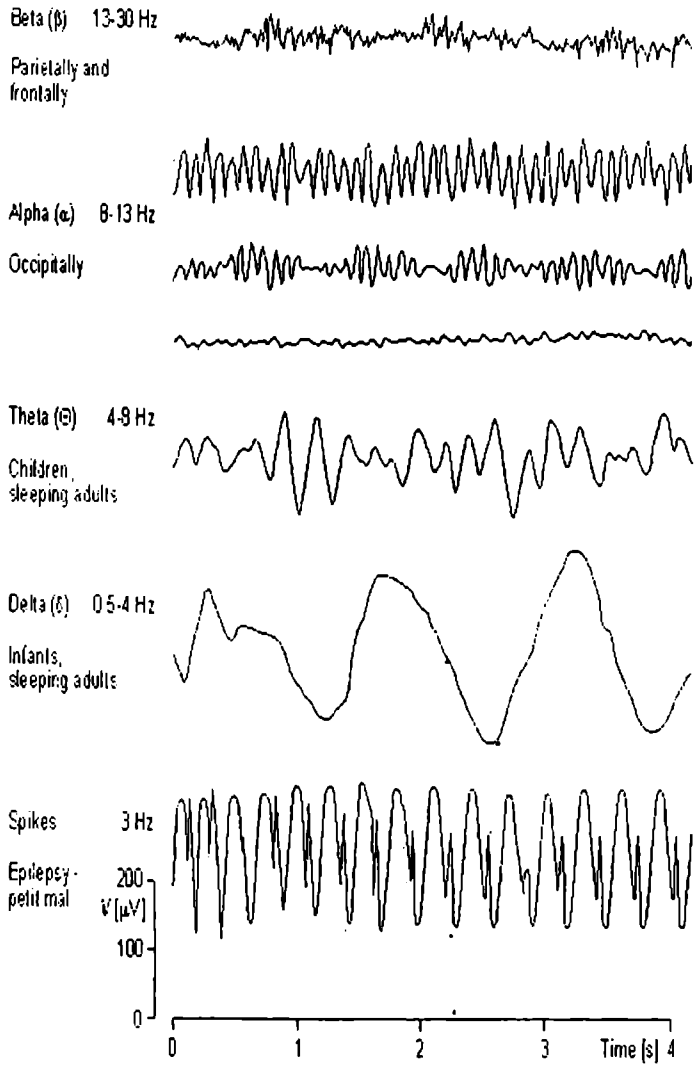


Figure 178 : Some examples of EEG waves.

The alpha waves have frequency spectrum of 8-13 Hz and can be measured from the occipital region in an awake person when the eyes are closed. The frequency band of the beta waves is 13-30 Hz; these are detectable over the parietal and frontal lobes. The delta waves have the frequency range of 0.5-4 Hz and are detectable in infants and sleeping adults. The theta waves have the frequency range of 4-8 Hz and are obtained from children and sleeping adults.

The EEG signal is closely related to the level of consciousness of the person. As the activity increases, the EEG shifts to higher dominating frequency and lower amplitude, when the eyes are closed, the alpha waves begin to dominate the EEG. When the person falls asleep, the dominant EEG frequency decreases, in a certain phase of sleep, rapid eye movement called (REM) sleep, the person dreams and has active movements of the eyes, which can be seen as a characteristic EEG signal. In deep sleep, the EEG has large and slow deflections called delta waves (Figure 179). No cerebral activity can be detected from a patient with complete cerebral death.

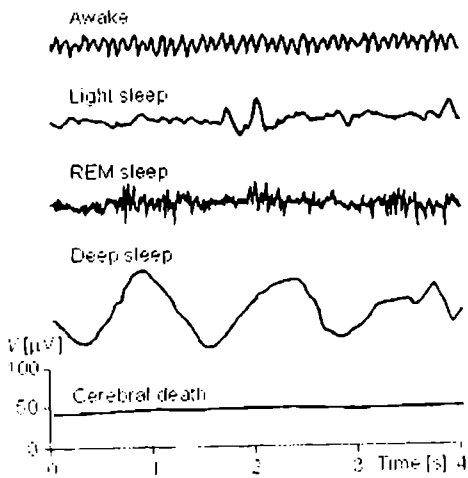


Figure 179 : EEG activity is dependent on the level of consciousness.

Acute intoxication or cerebral pathogenecity leads to disappearance of the EEG pattern and its replacement by "electrocerebral silence", which means that the electrical activity of the cortical mantle, measured at t scalp, is below 2 and may be absent.

Location of tumours, via predominance of the delta waves (Table 18) has been particularly successful, with 73-90 percent accuracy claimed. Other abnormalities, such as epilepsy, have been studied by this technique.

Table 18: Classification of electroencephalograph waves.

Names of waves	frequency (cps)	Association
delta	0.5 to 3.5	"disease, degeneration, death; defence" *
theta	6 to 7	
alpha	8 to 13	a scanning mechanism
beta	14 to 30	alertness; active response

While all the encompassing phenomenological technique of EEC have been making useful contribution to life, studies of individual neurons, via micro-electrodes in the cortex, and studies of the properties of synopses and ganglia in the spinal cord have demonstrated interesting phenomena such as: inhibition of transmission across nerve endings (strong signals passed through one nerve ending reduce the effectiveness on one close by); post tectonic potentiation (faster and more energetic transmission through a particular nerve path following a rapid succession of pulses through that path); and the promotion of epileptic - like seizures and peculiar mental images in human being by electrical stimulation of particular spots in the cortex via micro-electrodes.

Chapter 14

RADIOCARBON DATING

Radiocarbon dating, discovered in 1947 by chemist Willard Libby, can be used to work out the age of any material that was once alive, from coal - which comes from trees - to human bones.

All living organisms contain carbon which they either absorbed from the atmosphere (plants) or ingested in as their food (animals). Most of that carbon is a stable form called C^{12} but a tiny proportion is an unstable, radioactive form called C^{14} or radiocarbon. Radiocarbon decays very slowly at a constant rate.

The proportion of radiocarbon in any plant or animal remains constant throughout its life - some decays but this is replaced by breathing or eating. But once the organisms dies, the proportion of C^{14} starts to fall as the decaying molecules are no longer replaced.

As each atom decays, it changes from carbon to nitrogen and in the process emits a particle called an electron - this is radioactivity. These radioactive, emissions can be measured to find out how much radiocarbon has decayed and how much still remains. C^{14} has a half-life of 5,730 years, which means that the level of radioactivity will fall by half its original amount over that period. So from the proportion of C^{14} to C^{12} in the organism, scientists can work out reliably how long ago it died.

Chapter 15

پہ نجه فوړکړی د نا

DNA FINGER PRINTING

More accurately called DNA fingerprinting, this technique was first used to identify criminals. Now it is an important tool in forensic archaeology used to identify individuals, such as the Romanovs and also to gain information about people in the past.

DNA is the genetic material found in every cell of every living organism, The DNA molecule is a double helix - it looks like a ladder that has been twisted into a spiral. The uprights are the same for everyone, but the sequence of the rungs is unique for each person, apart from identical twins.

Forensic scientists use enzymes to cut the DNA into short pieces, and these fragments vary in length, depending on individual's genetic code. The fragments are separated according to their size, using an electric current, then they are labelled with a radioactive substance. An X-ray of the labelled fragments produces a DNA fingerprint that looks rather like a bar code,

The DNA from the cell nucleus is fragile and can be difficult to read if the samples come from very cold bodies. But another, more stable form of DNA is found in the *mitochondria* - the structure that produce the cell's energy, Mitochondrial DNA is harder to test and can only indicate relationships through the maternal line, but it provides important information about archaeological finds.

The technique of DNA fingerprinting was developed by Alec Jeffery and his colleagues at Leicester university. United Kingdom, and that of autoantibody fingerprinting described by A.K. Francoeur (1988). In this field, identity of a person with the help of blood stains, semen (sperm) stains, hair roots, tears, saliva and perspiration are possible with almost absolute certainty. The various techniques of DNA fingerprinting are as follows;

DNA FINGERPRINTING USING MINI-SATELLITE DNA

In this technique, DNA is isolated from blood stains, semen stains or root hairs, since hairs contain less amount of DNA, it can be produced on a large amount by polymerase chain reaction (PCR). Red blood corpuscle (RBC) do not have DNA, therefore, White blood corpuscle (WBC) are the source of DNA. The DNA isolated is cut with restriction enzymes and subjected to Southern Blotting. The DNA bands appearing on membrane are hybridized with 32 p DNA Probe, washed in water to remove unhybridized DNA, and passed through X-ray. The hybridized complementary DNA sequences develop images (prints). Identical images appearing on two X-ray films are identified, and thus identity is confirmed.

The DNA probes in this technique correspond to hyperveriable mini-satellites in DNA, each made up of tendom repeats of short sequences. A large number of these mini-satellites are scattered throughout the human genome which were first detected in an intron of human myoglobin gene. This reveals polymorphism in DNA, which has a very stable inheritance.

It has been speculated that the above technique will allow the identification of criminals in murder cases, rapist in rape cases, and of mother and/or father in case of doubtful parentage. This technique allows identification even when the stains on victims clothes are several years old and with more certainty than has hitherto been possible through techniques of blood groups etc. since the number of blood groups available becomes a limitation, The technique of DNA fingerprinting reveals such a great polymorphism that the possibility of two persons having the same pattern of DNA fingerprints is very remote.

AUTOANTIBODY FINGERPRINTING USING DIPSTICS

Autoantibodies that react with cellular components occur in high frequency in patients with systemic rheumatic diseases. However, a noval class of auto-antibodies that react with cellular components has now been identified in normal humans. Unlike the disease associated autoantibodies, that are restricted in number, these human autoantibodies increase in number from birth upto the age of 2 years and then remain constant for decades, if not lifelong. The complement of these autoantibodies present in an individual is unique, and for this reason they have been named *individual specific* (IS) autoantibodies. These is autoantibodies when physically separated comprise an *antibody fingerprint* that can serve to identify people just like DNA fingerprints. For these autoantibody fingerprints, body fluids such as blood, semen, tears, saliva, and perspiration can be used.

The advantage of autoantibody fingerprinting over DNA fingerprinting include the following:

1. *Sensitivity* (less than 10 μm blood needed);
2. *Rapidity* (only few hours needed; no DNA extraction required, which may take time);
3. Simplicity (no equipment required);
4. Cost effectiveness;
5. Portability;
6. Autoantibody fingerprints though are similar in the newborn child and the mother, can distinguish between genetically identical individuals like identical twins which cannot be distinguished through DNA fingerprinting.

The autoantibody fingerprinting protocol involves: (a) preparation of antigen, and (b) antibody fingerprint assay. Following steps are involved;

(i) A panel of antigen is prepared from an extract of human cells (e.g., Hela cells) cultured in-vitro; the antigens are first separated by molecular mass with denaturing polyacrylaraide gel electrophoresis (SDS-PAGE);

(ii) The antigen is then transferred elctrophoretically to a nylon membrane;

(iii) unbound sites on the membrane are blocked with the help of a blocking agent;

(iv) The membranes are cut into strips called *dipsticks,* which are

(v) Incubated with dilutions of sera or plasma by one hour;

(vi) washed with buffer;

(vii) incubated in detector molecule (1 mCl/assay of 125 I protein, which binds the PC*protein of human igG in the immune complex formed) for one hour.

(viii) Rewashed and dried; and

(ix) Finally used for autoradiography on X-ray film or scanned with a gamma scanner or an optical scanner.

The results obtained can be utilized for the identification of individuals (human, dogs, mice, cows, horses, rabbits etc.).

BANDED KRAIT MINOR SATELLITE FINGERPRINTING

Dr. Lalji Singh, has developed an indigenous technique where BKM-DNA probe is used for hybridization. While he was working on sex determination in snakes for his Ph.D. programme (in BHU, Varanasi), he found a contrast result, unlike human, the female snake contains XY and the male YY sex chromosome. A segment of DNA was isolated from sex determining Y chromosome of banded krait *(Bungarus fasciatus),* an Indian poisonous snake. This segment was named as 'banded krait minor satellite' (BKM). It is similar to the sex determining chromosome in humans. The probe which is used for this purpose is, therefore, called as BKM-DNA probe.

In India, facilities for DNA fingerprinting are available at Centre for Cellular and Molecular Biology (CCMB), Hyderabad. At this centre, a disputed matter of parentage at Madras was solved by the High Court based on the report of CCMB. A four year old girl (Laxmi) was stolen in June 1988, and named as Merry by the so-called parent. The identity of the real parent was confirmed through DNA fingerprinting.

In 1993, DNA fingerprinting centre was set up at Hyderabad with the help of CSIR, Department of Biotechnology and Ministry of Home Affairs.

Chapter 16

CHEMICAL FINGERPRINTING OF PLANTS

The term fingerprinting has been widely used in a number of sciences: medicine, biology, genetics and pharmaceutical studies, in geology, archeometry and environmental studies, chemical fingerprinting identifies the distribution of chemical elements within a matrix and thus defines its unique portrait in comparison to similar matrices. Fingerprinting implies the determination of as many elements as possible (Djingova et al., 2004).

A number of environmental materials (rocks, soils, and sediments) have already been fingerprinted by numerous investigations. The genetic relationship between them, which began with the formation of the earth and followed the process from magma to igneous rocks to sediments and soils, makes the task relatively easy in comparison to biota. Thus, today the fingerprints of the basic types of rocks and sediments are well known and the variations within one type are amazingly small. Although there are many different types of soil and it is quite difficult to classify them, once the parent material is known, certain observations can be done immediately (Pfeifer et al., 2000).

The chemical fingerprinting of a plant demands the determination of a large number of elements at the background level, but also in the polluted regions and in relation to other species. The possibility to make fingerprints of plants in respect to the distribution of chemical elements is very natural since their

composition depends to a great extent on the geochemical features of the environment where plants grew. Since it is possible to determine the inorganic components of this environment, it should also be possible to determine the composition of the plants growing there, of course the biochemical processes within a living organism are very complicated and might not allow the definition of narrow intervals of chemical distribution within the organism.

The possibility of fingerprinting the chemical composition of a plant is supported by several facts. First plants can exist under specific conditions determined by the nutrient supply and levels of toxic substances. The uptake of specific nutrients is dependant on their amount and availability but the plant can survive only within certain concentrations of nutrients otherwise there are deficiency symptoms and possibly death. As far as toxic elements are concerned no matter how toxicotolerant a plant is, there is a certain limit of endurance after which the plant dies (Markert, 1996). The fact that such limits are necessary for the normal existence of plants indicates the existence of reasonable concentration intervals in a plant.

Another very important fact is the Biological System of Elements (BSE). The BSE implies that elements are generally correlated in the organism according to their physiological role in the organism; and the uptake form of individual elements by living organisms.

FINGER PRINTING OF PLANTS

Chemical fingerprints of plants may be considered as normal concentrations in a plant at background levels described with the respective interval of biological variation. The fingerprints of plants might facilitate research in a number of areas of environmental, biomonitoring, and biological research. Some possible applications are: (i) better understanding of the accumulative, indicative, and regular behaviour of a plant and evolution of the distribution pattern of the elements in the different compartments of the ecosystem by comparison of plant fingerprints; (ii) quick assessment of local pollution levels and quick risk assessment for ecosystems by direct comparison of the experimental results to the plant's fingerprints; (iii) the possibility to compare, type and systematise groups of plants.

The use of sophisticated methods of analysis during the past 20 years has led to new opportunities in environmental research. Methods such as inductively coupled plasma mass spectrometry (ICPES), instrumental neutron activation analysis (INAA), inductively coupled plasma atomic emission spectrometry (ICP-AES), graphite furnace atomic absorption spectrometry (GFAA), and X-ray methods permit the routine determination of more than 15 elements in plants with reasonable accuracy and overall certainty. The possibility of determining a large number of elements within short time intervals facilitated the concept of using distribution patterns of chemical elements to produce fingerprints of environmental matrices. However, two extremely important conditions should always be kept in mind. The chemical characterization demands that the method used must have high accuracy and precision and this is only possible for specialized laboratories.

The production of correct analytical data is meaningless if the intrinsic dynamics of the plant systems relative to location and time are overlooked. Therefore, representative sampling programs adapted to the needs of biological and environmental research should be developed. Otherwise the correct analytical data will only be a snapshot of a given element in a given plant at a given point in time at a given location and will be useless for comparison and further conclusions (Markert, 1996). A problem that resulted from multi-element analysis is how to make a presentation of the data obtained from the investigations of plants. The approach has been the graphic presentation of chemical fingerprints either individually or in respect to some basic concentrations.

BASIC PROBLEMS IN CHEMICAL FINGERPRINTING OF PLANTS

Multi-element analysis of plants and attempts to generate the fingerprints of plants are continuously being done. Upto now, however, success has been achieved in a limited number of cases. The reasons are:

1. Unspecified or non-standardized sampling conditions concerning minimum plant quantity for representative sample, sampling season, sampling pre-treatment, which in most cases, makes the results not representative enough for a fingerprint.

2. Lack of background investigations. Most of the species are only investigated in polluted regions and these results cannot be used for fingerprints or for evaluation of the behaviour of the plant to pollution stress.
3. The limited number of elements that are determined mainly due to the wide use of atomic absorption analysis for these purposes, in majority of studies less than 10 elements are reported, usually there is no evidence for quality assurance and control or estimation of the uncertainty, which makes any comparisons meaningless.
4. The lack of inter species comparisons at background areas and in polluted areas, as well as the lack of ecosystem approach hinders not only the fingerprinting of a single plant but makes it impossible to specify its position among other species in the plant kingdom.

STATE OF THE ART IN THE FINGERPRINTING OF PLANTS

The state of the art in the fingerprinting of plants such as lichens, mosses, and higher plants is as follows:

(a) Fingerprinting of lichens

There are enormous variety of lichens - over 15000 species. Depending on the growing mode they are classified as crustose, foliose, or fruticose. Classification according to place of growth subdivides lichens into epigeic (growing on soil), epilithic (growing on rocks), epiphytic (growing on trees). Lichens derive nutrients and minerals from rainwater, air, through fall but also from the substrate, producing special acid substances (Tyler, 1989). In this way lichens are considered among the best indicators of atmospheric pollution. Therefore, the construction of a fingerprint of lichen species should not be difficult. However, there are a number of problems:

1. Over 25 species have been used to evaluate local pollution or regional distribution of heavy metals, and Cs-137 and there is not very high repeatability in the species used. Different lichen types have been used in most studies and fingerprint comparisons are impossible.

2. In comparison to the numerous data from polluted regions the data from background areas are rather limited.
3. Sampling parameters are also not comparable, in some of the studies lichens are washed before sampling, and/or after sampling (Garty et al., 1977), in other studies lichen samples are mechanically cleaned (Chiarenzelli et al, 1997) and in many cases the samples are not treated at all. Therefore, concentration differences in many lithophilic elements are difficult to be interpreted. Are they species specific or springing from dust?
4. Comparisons between species from the different groups - epigeic (growing on soil), epilithic (growing of rocks), or epiphytic (growing on trees) are also impossible due to the role of the substrate. There is also some evidence of within one group concentration differences depending on which tree the lichen grows.

The observations of Chiarenselli et al (1997) and Gough et al (1988 a,b) might be considered good examples of studies aiming to fingerprint lichens (Table 19).

The observations for Arsenic, Cadmium, Chromium, Copper, Nickel, Lead, Antimony, Vanadium and Zinc indicate that substrate rocks do not significantly influence the element concentration of the lichen. The concentrations are comparable to the observations for remote arctic area, Canada, Finland and united states. Crustose lichens are characterised by higher metal concentrations than foliose lichens and the lowest concentrations were observed in fruicose lichens.

(b) Fingerprinting of mosses

Mosses are among the most commonly used plants as bioindicators. They represent the simple art form of terrestrial plants. They are subdivided into endohydric, ectohydric and mixohydric. Ectohydric mosses have no internal conductive system or cuticle. Endohydric mosses possess an internal conductive system and water uptake from the soil is possible while the cuticle layer on the leaves of some permit absorption of water via the leaf surface. Ectohydric mosses are favoured in biomonitoring.

Table 19: Baseline concentrations (mg kg^{-1}) of several lichen species.

Element	*Hypogymnia enteromorpha*	*Usnea* spp	*Parmelia sulcata*	*Caloplaca cellaris*	*Caloplaca nivalis*	*Parmelia mougeotii*	*Usnea polyphylla*	*Caloplaca aurantia*
Al	1100	260	2000	–	–	–	–	–
As	–	–	0.97	0.19	0.19	0.48	0.54	
Ba	24	16	79	–	–	–	–	–
Cd	<0.08	<0.05	–	0.07	0.24	0.24	0.19	1.92
Ca	3800	3300	4400	–	–	–	–	–
Ce	0.55	0.22	–	–		–	–	–
Cr	4.9	1.0	7.3	1.24	1.18	12.4	3	45
Co	0.32	0.19	–	–	–	–	–	–
Cu	3.7	2.6	24	1.39	1.98	5.08	5.21	25.5
Ga	0.31	0.09	–	–	–	–	–	–
Fe	850	170	2700					11537
K	1800	1800	–	–	–	–	–	–
La	0.37	0.11	–	–	–	–	–	–
Li	0.33	0.097	–	–			–	–
Mg	1300	1600	730	–	–	–	–	–
Mn	89	97	72	–	–	–	–	162
Mo	<0.09	<0.05	–	–	–		–	–
Na	320	300	–	–	–	–	–	–
Nd	0.30	<0.09	–	–	–		–	–
Ni	11	6	6.6	1.83	1.92	6.4	2.42	36
P	670	420	790	–	–	–	–	–
Pb	12	7.5	26	0.75	3.23	12.5	10.4	46
S	470	18	1200	–	–	–	–	–
Sb	–	–	–	0.023	0.020	0.041	0.024	–
Sc	0.26	0.08	–	–	–	–	–	–
Sn	<0.9	<0.5	–	–		–	–	–
Sr	18	18	28	–	–	–	–	–
Ti	51	16	16	–	–	–	–	–
V	2.5	0.53	4.4	1.14	0.92	19.3	5.73	–
Y	0.22	0.10	2.5	–	–	–	–	–
Zn	25	21	95	15.6	25.3	24	22.2	304

Mosses have been used to estimate atmospheric heavy metal deposition in Europe (Table 19).

Sampling strategies of mosses are unified and as a rule of thumb, the samples are not being washed. Inter-comparisons between moss texa have been carried out and calibration factors for some elements have been derived. However, the conclusions about the possibility of using one species instead of another within one study are still contradictory.

(c) Fingerprinting of higher plants

Although lichens and mosses have been used in a large number of monitoring studies, large scale screening of higher plants are few. Generally higher plants are much more v/idely distributed, have a better morphology than lower plants, most of them are toxico-tolerant and are the basis of human alimentation (Wittig, 1993). This means that studying higher plants leads to much easier standardization of exposure methods, sampling methods, large scale screening, and comparative studies between places with very different degrees of anthropogenic influence, in an attempt to look for plants suitable for biomonitoring, trees such as *Populus nigra* italica, *Picea abies, P. silvestris,* and *Fagus sylvatica* have been screened for heavy metals. Among grasses *Lolium multiflorum* have gained popularity. Weed species are not well suited for fingerprinting of heavy metals.

Taraxcum officinale has been introduced as a biomonitor of heavy metal and toxic element pollution in a number of studies in Poland, Germany, Canada and Hungary. All these studies have contributed to the information obtained about the fingerprinting of chemical composition . Table 20 presents the concentrations for 39 elements background level. For some elements (for example Gold, Manganese, Caesium and Strontium) the intervals are larger and this is especially typical for elements whose concentration is highly dependent on the type and composition of soil. All studies were carried out in relatively unpolluted regions.

Populus nigra has been standardized as a biomonitor (Wagner, 1987). It has been used for screening of 40 elements (Table *21*). The element content of this plant depends on the type of soil. This

Table 20: Finger of several mass species(mg kg^{-1})

Element	*Pleurozium schreberi*	*Hylocomium splendens*	*Sphagnum* spp.	*Polytrichum formosum*
Al	–	322	488	400 (405)[3]
As	0.2	–	–	–
Ba	–	–	–	5 (7.8)
Br	1.6	7.4	12.6	–
Ca	–	–	2003	1000 (2710)
Cd	0.2	0.3	0.27	0.2
Ce	0.6	–	1.41	–
Co	0.2	–	–	–
Cr	0.9	1.07	1.24	3 (1)
Cs	0.12	–	–	–
Cu	4.5	4.9	2.5	12 (16.6)
Dy	–	–	0.1	–
Eu	–	–	0.023	–
Fr	–	–	0.066	–
Fe	150	210	352	400 (165)
Gd	–	–	0.1	–
Ho	–	–	0.025	–
K	–	–	5227	7500 (15490)
La	0.32	–	0.72	–
Mg	–	–	1088	500 (240)
Mn	110	235	101	–
Na	130	322	760	–
Ni	0.6	1.4	1.5	15
Pb	5.7	9.1	9.4	10 (2.7)
Pr	–	–	0.16	–
Rb	12	–	–	–
Sb	0.34	–	–	–
Sc	0.062	–	–	–
Se	0.34	–	–	–
Sm	0.04	–	0.11	–
Sr	–	–	–	4 (3.8)
Tb	–	–	0.026	–
Th	0.069	–	–	–
Ti	4	4.5	–	8
Tm	–	–	0.011	
V	1.4	1.75	2.36	
Yb	–	–	0.065	
Zn	25	26.5	26	30 (46)

Table 21: Fingerprint of *Taraxacum officinale* (mg kg^{-1})

Element	Concentration interval for background regions in Europe and USA	Element	Concentration interval for background regions in Europe and USA
Al	60–300	Mn	15–200
As	0.1–0.4	Mo	0.6–2.9
Au	0.004–0.03	N%	2.2–3.3
Ba	14–80	Na	50–400
Br	7–30	Ni	0.3–4
Ca%	1.1–2.0	P	2000–4000
Cd	0.2–0.8	Pb	0.3–6
Ce	0.3–0.6	Rb	24–160
Co	0.1–0.2	S	2200–5000
Cr	0.1–0.5	Sc	0.05–0.1
Cs	0.04–0.2	Se	0.05–0.2
Cu	5–20	Si	70–500
Eu	<0.005–0.02	Sm	0.05–0.2
Fe	60–500	Sr	10–45
Ga	0.16	Ti	5.6
Hg	<0.1–0.2	Th	<0.3–0.5
K%	2.1–4	Tl	0.025
La	0.2–0.8	V	0.18
Mg%	2.0–3.0	Zn	30–100

Table 22: Fingerprint of *populus nigra* Italica (mg kg^{-1})

Element	Background concentration interval	Element	Background concentration interval
As	0.2–0.4	La	0.1–0.54†
Be	0.0047–0.008†	Mn	26–126†
Bi	0.0016–0.019†	Na	64–300
Br	5.6–8.1	Nd	0.084–0.35†
Ca%	1.73–2.38	Ni	0.35–3.2
Cd	0.1–0.5	Pb	0.5–2
Ce	0.18–0.89†	Pr	0.02–0.26
Co	0.6–0.8	Rb	2.8–10.5
Cr	0.12–0.67	Sb	0.02–0.06
Cs	0.05–0.15	Sc	0.03–0.05
Cu	1.9–8.2	Sm	0.016–0.08†
Dy	0.011–0.056†	Sr	67–169
Er	0.0056–0.021†	Tb	0.002–0.027†
Eu	0.0066–0.028†	Te	0.0065–0.20†
Ga	0.61–1.19†	Tl	0.0083–0.061†
Gd	0.016–0.06†	U	0.015–0.083†
Ho	0.0019–0.0066†	V	0.15–0.79†
Fe	62–130	Y	0.067–0.36†
K%	0.66–1.35	Yb	0.0045–0.0145†

†Dependent on soil concentration.

information permits the fingerprint of this plant to include rarely observed elements such as Be, Te, Tl, U, Th.

The comparison between the fingerprints of *T. officinale* and *P. nigra* indicate that at background levels *T. officinale* has higher concentrations of Br, Cd, Cu, K, Kb, and Sm ad *P. nigra* has higher concentration of Ca, Co, and Sr. The concentration of the rest of the elements were similar in both the species.

Fingerprinting of the chemical composition of plants might be very helpful for structural, physiological, and environmental studies. However, many prerequisites are necessary to attain a successful fingerprint; standardized sampling procedures, accurate and precise analytical methods; determination of as many elements as possible, investigation of a large number of samples from background regions as well as from regions with different types of pollution and soils, and the determination of the relationship between the investigated plant within the same ecosystem. For lichens and mosses such studies have been more common but for higher plants the scope of these investigations should definitely be enlarged.

BIBLIOGRAPHY

Berger, H. 1929. Uber das Elektroenkepha logram des Menschen. *Arch.F.Psychiat.* 87: 527-570.

Blinov, L.M. 1983. *Electro-optical and magneto-optical principles of liquid crystals.* John Wiley & Sons, London.

Blumhardt, L.D., Barrett, G., Halliday, A.M., and Kriss, A. 1977. The asymmetrical visual evoked potential to pattern reversal in one half field and its significance for the analysis of visual field effects, *Br. J. Ophthalmol.* 61: 454-461.

Borthwick, H.A., Hindricks, S.B., and Parker, M.W. 1952. The reaction controlling floral initiation. *Proc. Natl. Acad. Sci.* USA. 38: 929-934.

Chiarenzelli, J., Aspler, L., Ozarko, D., Hall, G.E.M., Powis, K., and Danaldson, J. 1997. Heavy metals in lichens, southern district of Keewatin, Northwest territories, Canada. *Chemosphere,* 35: 1329-1341.

Clegg, J.S. 1984. *Amer. J. Physiol.* 246: 133-151.

Collings, P.J. 1990. *Liquid crystals, nature's delicate phase of matter.* Princeton University Press, Princeton.

Dekorsy, T., Auer, H., Waschke, C., Bakker, H.J., Roskos, H.G., and Kurz, H. 1995. *Physiol. Rev. Letters,* 74: 738-741.

Denbig, D. 1951. *The thermodynamics of the steady state.* Methuen & Co. Ltd., New York.

De Gennes, P.G. 1974. *The physics of liquid crystals.* Clarendon Press, Oxford.

Djingova, P., Kuleff, I., and Marlcert, B. 2004. Chemical fingerprinting of plants. *Ecological Research,* 19: 3-11.

Doebeli, M., Dieclamann, U. 2000. Evolutionary branching and sympatric speciation caused by different kinds of ecological interactions. *IIASA Interim Report.* IR-00-040.

Ehrenberg, W. 1967. *Scient. Amer.* 217: 103-110.

Freeman, W.J, 1995. *Societies of brains; A study in the neuroscience of love and hate.* Lawrence Erlbaum Associates, Hove.

Frohlich, H.1968. *Int. J. Quan. Chem.* 2: 641- 649.

Frohlich, H. 1980. *Adv. Electronics and Electron.* Phys. 53: 85-152.

Garner, W.W., and Allard, H.A. 1920. Effect of the relative length of day and night and other factors of the environment on growth and reproduction in plants. *Jour. Agric. Res.* 18: 553-606.

Garty, J., Galun, M., Fuchs, C., and Zisapel. N.1977. Heavy metals in the lichen *Caloplaca aurantia* from urban, suburban and rural regions in Israel (a comparative study). *Water, Air and Soil Pollution,* 8: 71-188.

Geritz, S.A.H., Metz, J.A.J., Kisdi, E., and Messena, a. 1997. Dynamics of adaptation and evolutionary branching, *Physical Review Letters,* 78: 2024-2027.

Gnaiger, E. 1994. *Modern trends in biothermokinetics.* 3: 62-70.

Goodwin, B.C. 1994. *How the leopard changed its spots: The evolution of complexity,* Weidenfeld and Nicolson, London.

Goodwin, B.C. 1995. In *Biocommunication* S 327 *Living processes.* Open University Press, Milton Keynes. 183-230.

Gouch, L., Jackson, L. , and Sacklon, J. 1988 a. Determination of baseline element composition of lichens. 11. *Hypogymmia enetromorpha.* and *Usnea* spp at Redwood National Park, California. *Water, Air and Soil Pollution,* 38: 169-180.

Gouch, L., Severson, R., and Jackson, L. 1988 b. Determination of baseline element composition of lichens. *I. Parmella sulcata* at theotore Roosevelt National Park. North Dacota, *Water, Air and Soil Pollution.* 38: 157-167.

Gray, G. 1993. British Association for the advancement of Science, Chemistry session: Molecular self-assembly in science and life. Sept. 1, Keele.

Gurney, W.S.C., and Nisbet, R.M. 1998. *Ecological dynamics.* Oxford University Press, New York.

Gurwitsche, A.G. 1925. *Bot. Gaz.* 80: 224-226.

Ho, M.W. 1986. *Riviata di Biologia.* 79: 407-447.

Ho, M.W. 1993. *The rainbow and the worm; The physics of organisms.* World Scientific, Singapore.

Ho, M.W. 1994 a. in *New metaphysical foundations of modern science.* Eds. W.Hasman and J.clark. Institute of Noetic sciences, Sausalito.

Ho, M.W. 1994 b. *Modern trends in biothermokinetics.* 3: 50-61.

Ho, M.W. ed. 1995 a. *Bioenergetics* S 327 *Living processes, University third level Science Course,* Open University Press, Milton Keynes.

Ho, M.W. 1995 b. *Neural Network World.* 5: 733-750.

Ho, M.W. 1995 C. In *Encyclopedia of Comparative Psychology. Eds.* G. Green berg and M. Haraway. Garland Publishing, New York.

Ho, M.W. 1995 e. *Soundings,* 1: 77-98.

Ho, M.W. 1996. *Annal Human Genetics.* 60: 81-84.

Ho, M.W., French, A., Haffegee, J., and Saunders, P.T. 1994. in *Bioetectrodynamics and biocommunication.* Eds. M.W. Ho, F.A. Popp and U. Warnke. World Scientific, Singapore. 195-207.

Ho, M.W., and Fox, eds. 1988. *Revolutionary processes and metaphors.* Wiley, London.

Ho, M.W., Popp, F.A., and Warnke, eds. 1994. *Bioelectro-magnetism and biocoimnunications.* World Scientific, Singapore.

Ho, M.W., and Saunders, PT. 1994. In *Bioelectromagnetism and biocommunication.* Eds. M.W. Ho, F.A. Popp and U. Warnke. World Scientific, Singapore. 213-225.

Ho, M.W., Stone, T.A., Jerraan, I., Bolton, J., Bolton, H., Goodwin, B.C., Saunders, P.T., and Robertsom, F. 1992. *Physics in medicine and biology.* 37: 1171-1179.

Ho, M.W., Xu, K., Ross, S., and Saunders, P. T. 1992 b. In *Advances in biophotons research. Eds.* F.A. Popp, K.H.Li and Q. Gu. World Scientific, Singapore. 287-306.

Ho, M.W., Zhom, Y.M., and Haffegee, J. 1995. In *Physics in biology.* Eds. L. Trainer and C. Lumsden. Academic Press.

Kaeser, H. 1987. In *The organisation* of *cell metabolism,* Eds. G.R. Welch and J.S. Clegg. Plenum Publishing Corporation, New York.

Knight, D. 1993. British association for the advancement of science, chemistry session: Molecular self-assembly in science and life. Sept. 1, Keele.

Kooijman, S.A.L.M. 2000. *Dynamics and mass budget in biological systems,* 2nd Edn. Cambridge University Press, Cambridge.

Markert, B. 1996. *Instrumental element and multi-element analysis of plants.* John Wiley and Sons, Chichester.

McClare, C.W.P. 1971. *J. Theor. Biol.* 30: 1-34.

Meszema. G., and Metz, J.A.J. 1999. Species diversity and population regulation: The importance of environmental feedback dimensionality. *IIASA Interim Report.* IR-99-045.

Mulder, C., Zonneveld, C., and Kooijman, S.A.L.M. 2001. Self organisation and emergent properties of the trophic structure and nutrient cycling of ecosystems. www.taio.vu.nl/thb/research/index.html.

Needham, J. 1936. *Orders and Life.* MIT Press, Cambridge, Mass.

Newton, R., Haffegee, J., and Ho, M.W. 1995. J. *Microscopy.* 180: 127-130.

Parker, M.W., Hendricks, S.B., Borthwick, H.A., and Scully, N.J. 1946. Action spectrum for the photoperiodic control of floral initiation of short day plants. *Bot. Gaz.* 108: 1-26.

Pfeifer, H.R., Durson, M.H., Rey, D., Schlegel, C., Atteia, O., Dallapiazza, R., Dubois, J.P., and Mandia, Y, 2000. Natural trace element input to soil-sediment-water-plant system: Examples of background and contaminated situations in Switzerland, Eastern France and North Italy. In *Trace elements, their distribution and effects in the environment. Eds.* B. Markert and K. Friase. Elsevier Science publishers, Amsterdam. 33-86.

Popp, F.A. 1986. In *Disequilibrium and self-organization.* Ed. C.W. Kilmester, Reidel, Dordrecht.p. 207.

Popp, F.A., Li, K.H., and Q. Gu eds. 1992. *Recent Advances in Biophoton Research and its applications.* World Scientific, Singapore.

Pratt, L.H. 1978. Molecular properties of phytochrome. *Photochem. Photobiol.* 27: 81-105.

Prigogine, I. 1967. *Introduction to thermodynamics of irreversible processes.* John Wiley & Sons, New York.

Schrodinger, E. 1944. What is life? Cambridge University Press, Cambridge.

Sewell, G.L. 1991, In *Large scale molecular systems; Quantum and stochastic aspects.* Eds. W. Gans, A. Blumen and A. Amann. Plenum Press, New York. 1-48.

Tyler, G. 1989. Uptake, retention and toxicity of heavy metals in lichens. *Water, Air and Soil Pollution,* 47: 321-333.

Wagner, G. 1987. Entwicklung einer methode zur grossracumigen uberwachung mittels standartiesierte pappelnblattproben von Beispiel von Blei, Cadmium und Zink. In *Beitrace gur* umwolf probebank. NPc - 5. Eds. M.Stoeppler and H.W.Duerback. 412, Juel,Sper.

Welch, G.R., and Berry, M.N. 1985. In *organized multi-enzyme systems,* Academic Press, New York.

Welch, G.R., add Glegg, J.S. Eds. 1987. *The organization* of cell metabolism. Plenum Press, New York.

Wheatley, D., and Clegg, J.S. 1991. *Am. Zool.* 31: 504-513.

Wittig, R. 1993. General aspects of biomonitoring heavy metals by plants. In *Plants as bioindicators.* Ed. B. Markert. VCH, Weinheim. 3-27.

SUBJECT INDEX

D

I

K

L

M

N

S

T

U

V

W

X

Z